UFAL

Um passeio pela Análise Funcional

Autor: Victor Ferreira de Araújo Santos

Orientador: Márcio Cavalcante de Melo

Julho 2022

Victor Ferreira de Araújo Santos

Um passeio pela Análise Funcional

Trabalho de conclusão de curso de graduação apresentado ao Instituto de Matemática, da Universidade Federal de Alagoas, como parte dos requisitos necessários à obtenção do título de licenciado em Matemática.

Maceió - AL, 2022

Catalogação na Fonte
Universidade Federal de Alagoas
Biblioteca Central
Divisão de Tratamento Técnico
Bibliotecário: Marcelino de Carvalho Freitas Neto – CRB-4 – 1767

S237p Santos, Victor Ferreira de Araújo.
Um passeio pela análise funcional / Victor Ferreira de Araújo Santos. - 2022.
101 f. : il.

Orientador: Márcio Cavalcante Melo.
Monografia (Trabalho de Conclusão de Curso em Matemática : Licenciatura) – Universidade Federal de Alagoas. Instituto de Matemática. Maceió, 2022.

Bibliografia: f. 101.

1. Análise. 2. Análise matemática. 3. Análise funcional. 4. Dimensão infinita. I. Título.

CDU: 517.98

Victor Ferreira de Araújo Santos

Um passeio pela Análise Funcional

Trabalho de conclusão de curso apresentado ao corpo docente do Curso de Matemática Licenciatura da Universidade Federal de Alagoas, Campus A.C. Simões, como requisito parcial para obtenção do grau de Licenciado em Matemática, avaliado pela banca examinadora constituída pelos professores:

Prof. Dr. Márcio Cavalcante de Melo
Instituto de Matemática - UFAL
Orientador

Prof. Dr. Abraão Mendes do Rego Gouveia
Instituto de Matemática - UFAL
Examinador

Prof. Dr. Rafael Nóbrega de Olivera Lucena
Instituto de Matemática - UFAL
Examinador

Aprovado em: ______ de ____________ de 2022.

Resumo

A Análise funcional desempenha um papel cada vez mais importante nas ciências aplicadas, bem como na própria matemática pura. Consequentemente, é cada vez mais desejável introduzir ao aluno esse ramo da matemática numa fase inicial de estudo. Pensado como uma introdução a Análise funcional para o estudante de final da graduação em matemática, este trabalho tem como objetivo introduzir os Espaços de Banach e Hilbert, suas consequências e resultados, tanto para dimensão finita como em dimensão infinita, que é o grande objetivo da Análise Funcional. Como resultado final, o último capítulo deste trabalho apresenta quatro dos principais teoremas de Análise Funcional, sendo eles o Teorema da Representação de Riesz, O Teorema da Aplicação Aberta, O Teorema do Gráfico fechado e por fim, o Teorema de Hahn-Banach.

Palavras-chave: Análise, Análise Matemática, Análise Funcional, Espaços de Dimensão Infinita.

Abstract

Functional analysis plays an increasingly important role in applied sciences as well as in pure mathematics itself. Consequently, it is increasingly desirable to introduce the student to this branch of mathematics at an early stage of study. Thought as an introduction to Functional Analysis for the undergraduate student in mathematics, this work aims to introduce the Banach and Hilbert spaces, their consequences and results, both for finite and infinite dimensions, which is the main objective of Functional analysis. As a final result, the last chapter of this work presents four of the main theorems of Functional Analysis, namely the Riesz Representation Theorem, the Open Application Theorem, the Closed Graph Theorem and, finally, the Hahn-Banach Theorem.

Key-words: Analysis, Mathematical Analysis, Functional Analysis, Infinite Dimension Spaces.

Sumário

Introdução

A Análise Matemática foi desenvolvida no século XVII, durante a Revolução Científica, embora muitas das suas ideias remontam aos matemáticos de tempos mais antigos, como por exemplo, na Matemática grega antiga. Na segunda metade do século XII Leibniz e Newton introduziram o Cálculo diferencial e integral. Já no século XVIII, Euler introduziu a noção precisa de função. A Análise Real começou a surgir como disciplina independente quando o matemático Bernard Bolzano introduziu a definição moderna de continuidade em 1816. Ainda no século XVIII um ramo da Análise foi descoberto, centrado no chamado número imaginário i. Importantes matemáticos estão associados com números complexos, incluindo Euller, Gauss, Riemann Cauchy, Weiertrass, entre muitos outros. Já no século XIX, Cauchy ajudou a introduzir o Cálculo Infinitesimal em fundamentos lógicos firmes com a introdução do conceito de sequência de Cauchy. Outros matemáticos como, por exemplo, Poisson, Liouville e Fourier estudaram as equações em derivadas parciais e a Análise Harmônica. Em seguida, no fim do século XIX e início do século XX, Borel e Lebesgue desenvolveram a teoria da Medida e Integração, introduzindo ferramentas que são utilizadas atualmente em diversos ramos da Análise, como por exemplo, nas Equações Diferenciais Parciais.

Em um contexto mais recente, entre o fim do século XIX e início do século XX, surgiu A Análise Funcional, um ramo Análise Matemática que se originou da análise clássica. Seu desenvolvimento começou por volta do fim do século XIX e início do século XX, e hoje em dia os métodos analíticos funcionais e os resultados são importantes em vários campos da matemática e suas aplicações. O ímpeto veio da álgebra linear, equações diferenciais, cálculo de variações, teoria da aproximação e, em particular, equações lineares integrais, cuja teoria teve o maior efeito sobre o desenvolvimento e promoção das ideias modernas. Os matemáticos observaram que os problemas de diferentes campos, muitas vezes aproveitam os recursos e propriedades relacionados. Diante disto é claro a importância do estudo da Análise para a formação de um matemático.

O objetivo do presente trabalho é fornecer um material acessível para estudantes no fim de um curso de graduação em Matemática.

A seguir fazemos uma descrição rápida dos resultados abordados no presente trabalho.

Nas preliminares enunciaremos e provaremos alguns resultados fundamentais sobre Espaços Métricos que serão essências para a leitura do texto.

Um *espaço métrico* (cf. 0.0.1) é um conjunto X com uma *métrica* nele. A métrica associa a qualquer par de elementos (pontos) de X uma distância. A métrica é definida axiomaticamente, sendo os axiomas sugeridos por certos propriedades simples da distância familiar entre pontos na reta real $\mathbb{R}$ e o plano complexo $\mathbb{C}$. Exemplos básicos (0.0.1 a 0.0.8) mostram que o conceito de espaço métrico é notavelmente geral. Uma propriedade adicional muito importante é a de que um espaço métrico pode ter *completude*.

No Capítulo 1, trataremos dos Espaços de Banach, onde podemos destacar os seguintes conceitos:

Um *espaço normado* (cf. 1.2.1) é um espaço vetorial (cf. 1.1.1) com um métrica definida por uma norma (cf. 1.2.1). A norma generaliza o comprimento de um vetor no plano ou no espaço tridimensional. Um *espaço de Banach* (cf. 1.2.1) é um espaço normado que é um espaço métrico completo.

Uma aplicação de um espaço normado X para um espaço normado Y é chamado de *operador*. Uma aplicação de X para o corpo escalar $\mathbb{R}$ ou $\mathbb{C}$ é chamado de *funcional*. De particular importância são os chamados *operadores lineares limitados* (cf. 1.6.1) e *funcionais lineares limitados* (cf. 1.7.2) uma vez que são contínuos e aproveitam a estrutura do espaço vetorial. De fato, o Teorema 1.6.4 afirma que um operador linear é contínuo se e somente se for limitado. Este é um resultado fundamental. Espaços vetoriais são importantes aqui principalmente por causa da linearidade que operadores e funcionais carregam.

É básico que o conjunto de todos os operadores lineares limitados de um dado espaço normado X em um espaço normado Y pode ser transformado em um espaço normado (cf. 1.9.1), que é denotado por $B(X, Y)$. Da mesma forma, o conjunto de todos os funcionais lineares limitados em X torna-se um espaço normado, que é chamado de *espaço dual* X' de X (cf. 1.9.3).

No Capítulo 2, abordaremos os Espaços de Hilbert, onde um espaço de produto interno X (cf. Def. 2.1.1) é um espaço vetorial com um produto interno $\langle x, y\rangle$ definido nele. Este último generaliza o produto escalar de vetores no espaço tridimensional e é usado para definir

(I) uma norma $||\cdot||$ por $||x|| = \langle x, x\rangle^{1/2}$,

(II) Ortogonalidade por $\langle x, y\rangle = 0$.

Um espaço de Hilbert H é um espaço com produto interno completo. A teoria de produto interno e Espaços de Hilbert é mais rica do que em espaços normados e espaços de Banach gerais. As características distintivas são

(i) representações de H como uma soma direta de um subespaço fechado e seu *complemento ortogonal* (cf. Seç. 2.3),

(ii) *Conjuntos e sequências ortonormais* e representações correspondentes de elementos de H (cf. Seções 2.4, 2.5),

(iii) A *Representação de Riesz* 3.1.1 de funcionais lineares limitados por produtos internos.

Por fim o Capítulo 3 é a parte central do TCC, o qual iremos provar os Teoremas fundamentais, visto em cursos de Análise Funcional, sendo eles:

1. *Teorema da Representação de Riesz.* Como citado anteriormente, começamos provando que todo funcional linear limitado em um espaço de Hilbert pode ser representado por um produto interno.

2. *Teorema da Aplicação Aberta.* Este teorema afirma que um operador linear limitado T de um espaço de Banach para um espaço de Banach é uma aplicação aberta, ou seja, transforma conjuntos abertos em conjuntos abertos. Portanto, se T é bijetivo, T^{-1} é contínua.

3. *Teorema do Gráfico fechado.* Este teorema dá condições sob o qual um operador linear fechado é limitado.

4. *Teorema de Hahn-Banach (e suas variantes).* Este é um teorema de extensão para funcionais lineares em espaços vetoriais. Garante que um espaço normado seja ricamente suprido de funcionais lineares, que se obtém uma teoria adequada dos espaços duais.

Por fim, enunciamos uma versão alternativa do Teorema de Riesz para funcionais lineares limitados no espaço $C[a, b]$ onde afirma que todo funcional do espaço pode ser representado como uma integral de Riemann-Stieltjes.

Preliminares (Noções de Espaços Métricos)

Neste trabalho presume-se como pré-requisito o conhecimento de Análise na Reta e Álgebra Linear. Com relação as noções de espaços métricos, irei detalhar alguns conceitos a seguir que serão utilizados ao longo do texto, a exemplo do Teorema do Completamento para Espaços Métricos.

Definição 0.0.1 Um espaço métrico é um par (X, d) onde X é um conjunto e d é uma métrica em X, ou seja, uma função definida em $X \times X$ tal que para todo x, y e $z \in X$ nós temos:

(M1) d é um número real, finito e não-negativo ;

(M2) $d(x, y) = 0 \Leftrightarrow x = y$;

(M3) $d(x, y) = d(y, x)$;

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

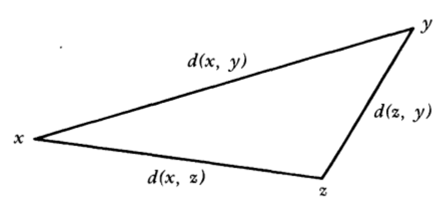

Figura 1: Desigualdade-triangular no plano.

A seguir daremos alguns exemplos de Espaços Métricos

Exemplo 0.0.1 **Reta real** $\mathbb{R}$. Este é o conjunto de todos os números reais, tomados com a métrica usual definida por

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

Exemplo 0.0.2 Plano euclidiano $\mathbb{R}^2$. O espaço métrico $\mathbb{R}^2$, chamado de espaço euclidiano plano, é obtido se tomarmos o conjunto dos pares ordenados de números reais, escrito $x = (\xi_1, \xi_2), y = (\eta_1, \eta_2)$ e a métrica euclidiana definida por

$$d(x, y) = \sqrt{(\xi_1 - \eta_1)^2 + (\xi_2 - \eta_2)^2}.$$

Outro espaço métrico é obtido se escolhermos o mesmo conjunto que antes, mas outra métrica d_1 definida por

$$d_1(x, y) = |\xi_1 - \eta_1| + |\xi_2 - \eta_2|.$$

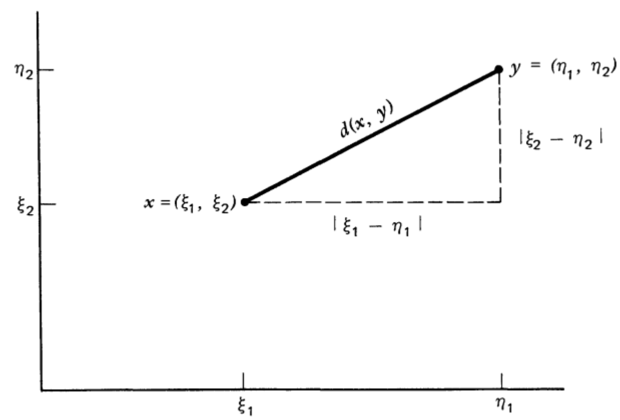

Figura 2: Métrica euclidiana no plano.

Exemplo 0.0.3 Espaço euclidiano tri-dimensional $\mathbb{R}^3$. Este espaço métrico consiste no conjunto das triplas ordenadas de números reais $x = (\xi_1, \xi_2, \xi_3), y = (\eta_1, \eta_2, \eta_3)$ e a métrica euclidiana definida por

$$d(x, y) = \sqrt{(\xi_1 - \eta_1)^2 + (\xi_2 - \eta_2)^2 + (\xi_3 - \eta_3)^3}.$$

Exemplo 0.0.4 Espaço euclidiano $\mathbb{R}^n$, Espaço unitário $\mathbb{C}^n$, Plano complexo $\mathbb{C}$. Os exemplos anteriores são casos especiais do espaço euclidiano n-dimensional $\mathbb{R}^n$. Este espaço é obtido se tomarmos o conjunto de todas as n-uplas ordenadas de números reais, escritas

$$x = (\xi_1, \cdots, \xi_n), \qquad y = (\eta_1, \cdots, \eta_n).$$

E a norma euclidiana definida por

$$d(x, y) = \sqrt{(\xi_1 - \eta_1)^2 + \cdots + (\xi_n - \eta_n)^2}.$$

O espaço unitário n-dimensional $\mathbb{C}^n$ é o espaço de todas as n-uplas ordenadas de números complexos com a métrica definida por

$$d(x, y) = \sqrt{|\xi_1 - \eta_1|^2 + \cdots + |\xi_n - \eta_n|^2}.$$

Quando $n = 1$ este é o plano complexo $\mathbb{C}$ com a métrica usual definida por

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

Exemplo 0.0.5 Espaço das sequências l^∞. Como um conjunto X, tomamos o conjunto de todas as sequências limitadas de números complexos; isto é, todo elemento de X é uma sequência complexa

$$x = (\xi_1, \xi_2, \cdots) \quad \text{ou} \quad x = (\xi_j)$$

tal que para todo $j = 1, 2, ...$ temos

$$|\xi_j| \leqslant c_x.$$

onde c_x é um número real que pode depender de x, mas não depende de j. Escolhemos a métrica definida por

$$d(x, y) = \sup_{j \in \mathbf{N}} |\xi_j - \eta_j|$$

onde $y = (\eta_j) \in X$ e $\mathbb{N} = \{1, 2, \cdots\}$. l^∞ é chamado espaço das sequências porque cada elemento de X (cada ponto de X) é uma sequência.

Exemplo 0.0.6 Espaço das funções C[a,b]. Como um conjunto X tomamos o conjunto de todas as funções de valor real $x, y, ...$ que são funções de uma variável real t e são definidas e contínuas em um dado intervalo fechado $J = [a, b]$. Escolhendo a métrica definida por

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

$C[a, b]$ é chamado espaço das funções contínuas em $[a, b]$ porque todo ponto de $C[a, b]$ é uma função contínua.

Exemplo 0.0.7 Espaço B(A) das funções limitadas. Por definição, cada elemento $x \in B(A)$ é uma função definida e limitada em um dado conjunto A, e a métrica é definida por

$$d(x, y) = \sup_{t \in A} |x(t) - y(t)|$$

Nós escrevemos $B[a, b]$ para $B(A)$ no caso de um intervalo $A = [a, b] \subset \mathbb{R}$.

Vamos mostrar que $B(A)$ é um espaço métrico. Claramente, (M1) e (M3) são satisfeitas. Além disso, $d(x, x) = 0$ é óbvio. Por outro lado, $d(x, y) = 0$ implica $x(t) - y(t) = 0$ para todo $t \in A$, de modo que $x = y$. Isso dá (M2). Além disso, para todo $t \in A$ temos

$$\begin{aligned} |x(t) - y(t)| &\leqq |x(t) - z(t)| + |z(t) - y(t)| \\ &\leqq \sup_{t \in A} |x(t) - z(t)| + \sup_{t \in A} |z(t) - y(t)|. \end{aligned}$$

Isso mostra que $x - y$ é limitado em A. Como o limite dado pela expressão na segunda linha não depende de t, podemos tomar o supremo à esquerda e obter (M4).

Exemplo 0.0.8 Espaço l^p, Espaço l^2, Desigualdades de Holder e Minkowski. Seja $p \geqslant 1$ um número real fixo. Por definição, cada elemento no espaço l^p é uma sequência $x = (\xi_j) = (\xi_1, \xi_2, \cdots)$ de números tais que $|\xi_1|^p + |\xi_2|^p + \cdots$ converge; portanto

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

e a métrica pe definida por

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

onde $y = (\eta_j)$ e $\sum |\eta_j| < \infty$. Se tomarmos apenas sequências reais satisfazendo (1), obtemos o espaço real l^p, e se tomarmos sequências complexas satisfazendo (1), obtemos o espaço complexo l^p. (Sempre que a distinção é essencial, podemos indicá-lo por um subscrito $\mathbb{R}$ ou $\mathbb{C}$, respectivamente.)
No caso $p = 2$ temos o famoso espaço sequencial de Hilbert l^2 com métrica definida por

$$d(x,y) = \sqrt{\sum_{j=1}^{\infty} |\xi_j - \eta_j|^2} \qquad (3)$$

Provaremos que l^p é um espaço métrico. Claramente, (2) satisfaz (M1) a (M3) desde que a série da direita seja convergente. Vamos provar que converge e que (M4) é satisfeito. Procedendo passo a passo, faremos:

(a) Desigualde de Young,

(b) Desigualdade de Holder,

(c) Desigualdade de Minkowski,

(d) Desigualdade triangular do espaço l^p.

Os detalhes da prova seguem a seguir:

(a) Seja $p > 1$ e defina q por

$$\frac{1}{p} + \frac{1}{q} = 1. \qquad (4)$$

Daí,

$$1 = \frac{p+q}{pq} \Rightarrow pq = p + q \Rightarrow (p-1)(q-1) = 1. \qquad (5)$$

Donde, $\dfrac{1}{p-1} = q-1$, de modo que

$$u = t^{p-1} \quad \text{implica} \quad t = u^{q-1}.$$

Sejam α e β quaisquer números positivos. Como $\alpha\beta$ é a área do retângulo na Fig. 3, obtemos assim por integração a desigualdade

$$\alpha\beta \leqq \int_0^\alpha t^{p-1}dt + \int_0^\beta u^{q-1}du = \frac{\alpha^p}{p} + \frac{\beta^q}{q}. \qquad \textbf{(Desigualdade de Young)} \qquad (6)$$

Note que a desigualdade é trivial no caso $\alpha = 0$ ou $\beta = 0$.

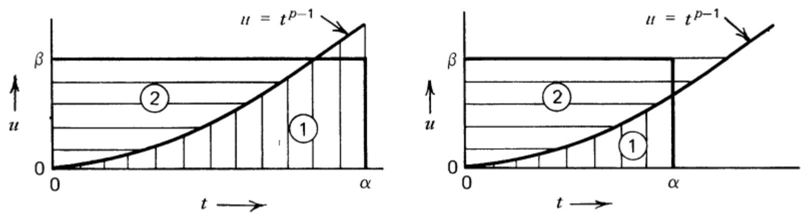

Figura 3: Desigualdade (6), onde a região 1 corresponde à primeira integral em (6) e a região 2 à segunda.

(b) Sejam $\tilde{\xi}_j$ e $\tilde{\eta}_j$ tais que

$$\sum |\tilde{\xi}_1|^p = 1, \qquad \sum |\tilde{\eta}_1|^p = 1. \qquad (7)$$

Tomando $\alpha = |\tilde{\xi}_1|$ e $\beta = |\tilde{\eta}_1|$ pela desigualdade de young,

$$\left|\tilde{\xi}_j\tilde{\eta}_j\right| \leqq \frac{1}{p}\left|\tilde{\xi}_j\right|^p + \frac{1}{q}\left|\tilde{\eta}_j\right|^q.$$

Se somarmos sobre j e usarmos (7) e (4), obtemos

$$\sum \left|\tilde{\xi}_j\tilde{\eta}_i\right| \leqq \frac{1}{p} + \frac{1}{q} = 1. \qquad (8)$$

Agora tomamos qualquer $x = (\xi_j) \in l^p$ diferente de zero e $y = (\eta_j) \in l^q$ e fazemos

$$\tilde{\xi}_j = \frac{\xi_j}{\left(\sum |\xi_k|^p\right)^{1/p}}, \qquad \tilde{\eta}_j = \frac{\eta_j}{\left(\sum |\eta_m|^q\right)^{1/q}}. \qquad (9)$$

Então (7) é satisfeito, para que possamos aplicar (8). Substituindo (9) em (8) e multiplicando a desigualdade resultante pelo produto dos denominadores em (9), chegamos à desigualdade de Holder para somas

$$\sum_{j=1}^{\infty} |\xi_j\eta_j| \leqq \left(\sum_{k=1}^{\infty} |\xi_k|^p\right)^{1/p} \left(\sum_{m=1}^{\infty} |\eta_m|^q\right)^{1/q} \qquad \textbf{(Desigualdade de Holder)} \qquad (10)$$

onde $p > 1$ e $1/p + 1/q = 1$.

Quando $p = 2 = q$ temos um caso particular conhecido como:

$$\sum_{j=1}^{\infty} |\xi_j \eta_j| \leqq \sqrt{\sum_{k=1}^{\infty} |\xi_k|^2} \sqrt{\sum_{m=1}^{\infty} |\eta_m|^2}. \quad \textbf{(Desigualdade de Cauchy-schwarz)} \qquad (11)$$

(c) Agora provaremos a desigualdade de Minkowski para somas

$$\left(\sum_{j=1}^{\infty} |\xi_j + \eta_j|^p\right)^{1/p} \leqq \left(\sum_{k=1}^{\infty} |\xi_k|^p\right)^{1/p} + \left(\sum_{m=1}^{\infty} |\eta_m|^p\right)^{1/p}. \quad \textbf{(Desigualdade de Minkowski)} \qquad (12)$$

Para p = 1, a desigualdade decorre prontamente da desigualdade triangular para números. Seja $p > 1$. Para simplificar as fórmulas, devemos escrever $\xi_j + \eta_j = \omega_j$. A desigualdade triangular para números dá

$$\begin{aligned} |\omega_j|^p &= |\xi_j + \eta_j|\, |\omega_j|^{p-1} \\ &\leqq (|\xi_j| + |\eta_j|)\, |\omega_j|^{p-1}. \end{aligned}$$

Somando j de 1 a qualquer n fixo, obtemos

$$\sum |\omega_j|^p \leqq \sum |\xi_j|\, |\omega_j|^{p-1} + \sum |\eta_j|\, |\omega_j|^{p-1}. \qquad (13)$$

À primeira soma à direita aplicamos a desigualdade de Holder, encontrando

$$\sum |\xi_j|\, |\omega_j|^{p-1} \leqq \left[\sum |\xi_k|^p\right]^{1/p} \left[\sum \left(|\omega_m|^{p-1}\right)^q\right]^{1/q}.$$

À direita temos $(p-1)q = p$ pois $pq = p + q$ por (5). Tratando a última soma em (13) de forma semelhante, obtemos

$$\sum |\eta_j|\, |\omega_j|^{p-1} \leqq \left[\sum |\eta_k|^p\right]^{1/p} \left[\sum |\omega_m|^p\right]^{1/q}.$$

Juntas,

$$\sum |\omega_j|^p \leqq \left\{\left[\sum |\xi_k|^p\right]^{1/p} + \left[\sum |\eta_k|^p\right]^{1/p}\right\} \left(\sum |\omega_m|^p\right)^{1/q}.$$

Dividindo pelo último fator à direita e notando que $1 - 1/q = 1/p$, obtemos (12) com n em vez de ∞. Agora fazemos $n \longrightarrow \infty$. À direita isso produz duas séries que convergem porque $x, y \in l^p$. Daí a série à esquerda também converge, e (12) está provado.

(d) De (12) segue que para $x, y \in l^p$ a série em (2) converge. (12) também produz a desigualdade triangular. Na verdade, tomando qualquer $x, y, z \in l^p$, escrevendo $z = (\xi_j)$ e usando a desigualdade triangular para números e então (12), obtemos

$$\begin{aligned} d(x,y) &= \left(\sum |\xi_j - \eta_j|^p\right)^{1/p} \\ &\leqq \left(\sum [|\xi_j - \zeta_j| + |\zeta_j - \eta_j|]^p\right)^{1/p} \\ &\leqq \left(\sum |\xi_j - \zeta_j|^p\right)^{1/p} + \left(\sum |\zeta_j - \eta_{ij}|^p\right)^{1/p} \\ &= d(x,z) + d(z,y). \end{aligned}$$

Isso completa a prova de que l^p é um espaço métrico.

Prosseguindo, podemos definir uma noção de continuidade entre dois espaços métricos, de maneira análoga a como ocorre na análise real:

Definição 0.0.2 Seja $X = (X, d)$ e $Y = (Y, \tilde{d})$ espaços métricos. A aplicação $T : X \longrightarrow Y$ é dita contínua em um ponto $x_0 \in X$ se para todo $\epsilon > 0$ existe um $\delta > 0$ tal que

$$d(Tx, Tx_0) < \varepsilon \qquad \text{para todo } x \text{ satisfazendo} \qquad d(x, x_0) < \delta. \tag{1}$$

T é dita contínua se for contínua em todos os pontos de X.

Chamamos x_0 de **ponto interior** de um conjunto $M \subset X$ se M é uma vizinhança de x_0. O interior de M é o conjunto de todos os pontos interiores de M e pode ser denotado por $int(M)$. Se $int(M) = M$ dizemos que M é **aberto**. $int(M)$ é aberto e é o maior conjunto aberto contido em M.

Uma Aplicação contínua pode ser caracterizado em termos de conjuntos abertos, seguindo o próximo resultado

Teorema 0.0.3 *Uma aplicação T de um Espaço Métrico X em um espaço métrico Y é dita contínua se, somente se, a imagem inversa de qualquer subconjunto aberto em Y for um subconjunto aberto em X.*

Espelhando-se no conceito de convergência de sequências na reta real, definiremos a convergência de sequências para espaços métricos

Definição 0.0.4 Uma sequência (x_n) em um espaço métrico $X = (X, d)$ é dita convergente se existe um $x \in X$ tal que

$$\lim_{n \to \infty} d(x_n, x) = 0. \tag{2}$$

Onde a convergência acima é na topologia dos reais.

Dizemos que x é o limite de (x_n) e podemos escrever

$$\lim_{n\to\infty} x_n = x \tag{3}$$

ou simplesmente $x_n \longrightarrow x$. Se (x_n) não for convergente, dizemos que ela é divergente.

Assim como na reta real, toda sequência convergente é limitada e o limite é único.

Teorema 0.0.5 *Seja $X = (X,d)$ um espaço métrico. Temos:*

(a) Uma sequência convergente em X é limitada e seu limite é único.

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

Vamos agora definir o conceito de completude de um Espaço Métrico, que será fundamental em nosso trabalho futuro.

Definição 0.0.6 Uma sequência (x_n) em espaço métrico $X = (X,d)$ é dita de Cauchy se para todo $\epsilon > 0$, existe um $N = N(\epsilon)$ tal que

$$d(x_m, x_n) < \epsilon, \qquad \forall\, m, n > N. \tag{4}$$

O espaço X é dito completo se toda sequência de Cauchy em X converge (isto é, tem um limite que é um elemento de X).

Com esta definição, conhecida como critério de Cacuhy, podemos enunciar o seguinte resultado

Teorema 0.0.7 *A reta real e o plano complexo são espaços métricos completos.*

De maneira mais geral, agora vemos diretamente da definição que Espaços Métricos completos são precisamente aqueles em que o critério de Cauchy continua sendo necessário e suficiente para a convergência. Dessa forma, torna-se necessário estabelecer uma relação intima entre convergência e sequências de Cauchy, como podemos ver no teorema a seguir

Teorema 0.0.8 *Toda sequência convergente em um espaço métrico é uma sequência de Cauchy.*

Vamos agora introduzir mais dois conceitos, que estão relacionados. Seja M um subconjunto de um espaço métrico X. Então um ponto x_0 de X (podendo não ser um ponto de M) é chamado de **ponto de acumulação** de M se cada vizinhança de x_0 contém pelo menos um ponto $y \in M$ distinto de x_0. O conjunto formado pelos pontos de M e os pontos de acumulação de M é chamado de **fecho** de M e é denotado por $\overline{M}$. Quando $M = \overline{M}$ diremos que M é fechado. Um subconjunto M de um espaço métrico X é dito **denso** em X se $\overline{M} = X$.

Veremos que um grande número de resultados básicos, por exemplo, em operadores lineares, vai depender da completude dos espaços correspondentes.

Iremos enunciar três teoremas relacionados a completude e a convergência que serão importantes para provar resultados futuros deste trabalho.

Teorema 0.0.9 *seja M um subconjunto não-vazio de um espaço métrico (X,d) e* $\overline{M}$ *o seu fecho. Então:*

(a) $x \in \overline{M}$ *se, somente se, existe uma sequência* (x_n) *em M tal que* $x_n \longrightarrow x$*.*

(b) M é fechado se, somente se, a situação $x_n \in M$*,* $x_n \longrightarrow x$ *implicar* $x \in M$*.*

Teorema 0.0.10 *um subespaço M de um espaço métrico completo X é ele mesmo completo se, somente se, o conjunto M é fechado em X.*

O último dos nossos três teoremas atuais mostra a importância da convergência de sequências em conexão com a continuidade de uma aplicação.

Teorema 0.0.11 *Uma aplicação* $T : X \longrightarrow Y$ *de um espaço métrico (X,d) em um espaço métrico (Y,*$\tilde{d}$*) é contínua em um ponto* $x_0 \in X$ *se, somente se,* $x_n \longrightarrow x_0$ *implicar* $Tx_n \longrightarrow Tx_0$*.*

Sabemos dos conceitos de Análise Real que o espaço $\mathbb{Q}$ dos números racionais apesar de ser incompleto, pode ser "completado"e assim obtemos o espaço $\mathbb{R}$ dos números reais, que é completo. Baseando-se nesse exemplo, levantamos a seguinte questão, espaços métricos incompletos podem ser completados até se formar um espaço completo? A resposta para essa questão é sim, como veremos a seguir com o Teorema do completamento para Espaços Métricos.

Para isso, primeiro iremos definir a noção de isometria entre Espaços Métricos.

Definição 0.0.12 (Espaços isométricos). Seja $X = (X, d)$ e $\tilde{X} = (\tilde{X}, \tilde{d})$ espaços métricos. Então:

(a) Uma aplicação T de X em $\tilde{X}$ é dito isométrico ou isometria se T preserva distância, isto é, se para todo $x, y \in X$, tem-se

$$\tilde{d}(Tx, Ty) = d(x, y). \tag{5}$$

(b) O espaço X é dito isométrico a $\tilde{X}$ se existir uma isometria bijetiva de X em $\tilde{X}$. Os espaços X e $\tilde{X}$ são então chamados espaços isométricos.

Portanto, os espaços isométricos podem diferir no máximo pela natureza de seus pontos, mas são indistinguíveis do ponto de vista da métrica. E em qualquer estudo em que a natureza dos pontos não importa, podemos considerar os dois espaços como idênticos, duas cópias do mesmo espaço "abstrato". Podemos agora enunciar e provar o resultado de que todo espaço métrico pode ser completado. O espaço $\tilde{X}$ que ocorre neste teorema é chamado de completamento do espaço X dado.

Teorema 0.0.13 ***(Teorema do completamento).*** *Para um espaço métrico $X = (X, d)$ existe um espaço métrico completo $\hat{X} = (\hat{X}, \hat{d})$ que tem um subespaço W que é isométrico com X e é denso em $\hat{X}$. Este espaço $\hat{X}$ é único exceto para isometrias, isto é, se $\tilde{X}$ é qualquer espaço métrico completo com um subespaço denso $\tilde{W}$ isométrico com X, então $\tilde{X}$ e $\hat{X}$ são isométricos.*

Demonstração:

Como a prova é extensa, vamos dividir esta em 4 etapas:

(a) Mostrar a existência de $\hat{X} = (\hat{X}, \hat{d})$.

(b) Construção da isometria T de X em W, onde $\overline{W} = \hat{X}$.

(c) Mostrar a completude de $\hat{X}$.

(d) Unicidade de $\hat{X}$, exceto por isometrias.

A grosso modo, nossa tarefa será a atribuição de limites adequados para sequências de Cauchy em X que não convergem. No entanto, não devemos introduzir muitos limites, mas ter em mente que certas sequências podem convergir para o mesmo limite, uma vez que os termos dessas sequências em última análise, aproximam-se arbitrariamente umas das outras. Segue os detalhes da prova:

(a) Construção do espaço $\hat{X} = (\hat{X}, \hat{d})$. Seja (x_n) e (x'_n) duas sequências de Cauchy em X. Defina (x_n) ser equivalente a (x'_n) escrevendo $(x_n) \sim (x'_n)$ se

$$\lim_{n\to\infty} d\left(x_n, x'_n\right) = 0. \tag{6}$$

Seja $\hat{X}$ o conjunto de todas as classes de equivalência $\hat{x}, \hat{y}, \cdots$ de sequências de Cauchy assim obtidas. Escrevemos que $(x_n) \in \hat{x}$ para indicar que (x_n) é um membro da classe $\hat{x}$ (um representante da classe $\hat{x}$). Nós agora definimos

$$\hat{d}(\hat{x}, \hat{y}) = \lim_{n\to\infty} d\left(x_n, y_n\right), \tag{7}$$

onde $(x_n) \in \hat{x}$ e $(y_n) \in \hat{y}$. Vamos mostrar que esse limite existe. Temos que

$$d\left(x_n, y_n\right) \leqq d\left(x_n, x_m\right) + d\left(x_m, y_m\right) + d\left(y_m, y_n\right), \tag{8}$$

daí obtemos

$$d(x_n, y_n) - d(x_m, y_m) \leqq d(x_n, x_m) + d(y_m, y_n) \tag{9}$$

e uma desigualdade semelhante com m e n trocados. Juntos,

$$|d(x_n, y_n) - d(x_m, y_m)| \leqq d(x_n, x_m) + d(y_m, y_n)\,. \tag{10}$$

Como (x_n) e (y_n) são de Cauchy, podemos tomar

$$d(x_n, x_m) \leqslant \frac{\varepsilon}{2}$$

e

$$d(y_m, y_n) \leqslant \frac{\varepsilon}{2}$$

de modo que

$$|d(x_n, y_n) - d(x_m, y_m)| \leqq d(x_n, x_m) + d(y_m, y_n) \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Isso implica que o limite em (7) existe porque $\mathbb{R}$ é completo.

Devemos também mostrar que o limite em (7) é independente da escolha particular dos representantes. De fato, se $(x_n) \sim (x'_n)$ e $(y_n) \sim (y'_n)$, então por (6),

$$|d(x_n, y_n) - d(x'_n, y'_n)| \leqq d(x_n, x'_n) + d(y_n, y'_n) \longrightarrow 0 \tag{11}$$

quando $n \longrightarrow 0$. isso implica,

$$\lim_{n\to\infty} d(x_n, y_n) = \lim_{n\to\infty} d(x'_n, y'_n)\,. \tag{12}$$

Provemos que d em (7) é uma métrica em X. Claramente d satisfaz a propriedade (M1), (M3) e (M4) da definição de métrica, e a propriedade (M2) é de fácil verificação, de fato

$$\hat{d}(\hat{x}, \hat{y}) = 0 \Rightarrow (x_n) \sim (y_n) \Rightarrow \hat{x} = \hat{y}. \tag{13}$$

(b) Construção da isometria $T : X \longrightarrow W \subset \hat{X}$. Para cada $b \in X$ associamos a classe $\hat{b} \in \hat{X}$ que contém a sequência de Cauchy constante $(b, b, \cdots)$. Isso define uma aplicação $T : X \longrightarrow W$ onde o subespaço $W = T(X) \subset \hat{X}$. A aplicação T é dado por $b \longmapsto \hat{b} = Tb$, onde $(b, b, \cdots) \in \hat{b}$. Temos que T é uma isometria, pois pela equação (7) temos

$$\hat{d}(\hat{b}, \hat{c}) = d(b, c); \tag{14}$$

aqui $\hat{c}$ é a classe de (y_n) onde $(y_n) = c$ para todo n. Qualquer isometria é injetiva, e $T : X \longrightarrow W$ é sobrejetiva, pois $T(X) = W$. Portanto, W e X são isométricos;

Vamos mostrar que W é denso em $\hat{X}$. Considere qualquer $\hat{x} \in \hat{X}$. Seja $(x_n) \in \hat{x}$. Para cada $\varepsilon > 0$ existe um $N > 0$ tal que se $n > N$ vale

$$d(x_n, x_N) < \frac{\varepsilon}{2} \tag{15}$$

Seja $(x_N, x_N, \cdots) \in \hat{x}_N$. Temos que $\hat{x}_N \in W$. Por (7),

$$\hat{d}\left(\hat{x}, \hat{x}_N\right) = \lim_{n\to\infty} d\left(x_n, x_N\right) \leqq \frac{\varepsilon}{2} < \varepsilon \tag{16}$$

Isso mostra que cada ε-vizinhança do arbitrário $\hat{x} \in \hat{X}$ contém um elemento de W. Portanto, W é denso em $\hat{X}$.

(c) Completude de $\hat{X}$. Seja $(\hat{x}_n)$ uma sequência de Cauchy arbitrária em $\hat{X}$. Como W é denso em $\hat{X}$ para cada $\hat{x}_n$ existe um $\hat{z}_n \in W$ tal que

$$\hat{d}(\hat{x}_n, \hat{z}_n) < \frac{1}{n} \tag{17}$$

Donde, pela desigualdade triangular,

$$\begin{aligned} \hat{d}\left(\hat{z}_m, \hat{z}_n\right) &\leqq \hat{d}\left(\hat{z}_m, \hat{x}_m\right) + \hat{d}\left(\hat{x}_m, \hat{x}_n\right) + \hat{d}\left(\hat{x}_n, \hat{z}_n\right) \\ &< \frac{1}{m} + \hat{d}\left(\hat{x}_m, \hat{x}_n\right) + \frac{1}{n} \end{aligned} \tag{18}$$

e isso é menor do que qualquer $\varepsilon > 0$ dado para m e n suficientemente grandes porque $(\hat{x}_m)$ é de Cauchy. Portanto $(\hat{z}_m)$ é de Cauchy. Uma vez que $T : X \longrightarrow W$ é isométrica e $\hat{z}_m \in W$, a sequência (z_m), onde $z_m = T^{-1}\hat{z}_m$, é de Cauchy em X. Seja $\hat{x} \in \hat{X}$ a classe à qual (z_m) pertence. Vamos mostrar que $\hat{x}$ é o limite de $\hat{x}_n$. Pela equação (17),

$$\begin{aligned} \hat{d}\left(\hat{x}_n, \hat{x}\right) &\leqq \hat{d}\left(\hat{x}_n, \hat{z}_n\right) + \hat{d}\left(\hat{z}_n, \hat{x}\right) \\ &< \frac{1}{n} + \hat{d}\left(\hat{z}_n, \hat{x}\right) \end{aligned} \tag{19}$$

Como $(z_m) \in \hat{x}$ e $\hat{z}_n \in W$, de modo que $(z_n, z_n, \cdots) \in \hat{z}_n$, a inequação acima nos dá que

$$\hat{d}\left(\hat{x}_n, \hat{x}\right) < \frac{1}{n} + \lim_{m\to\infty} d\left(z_n, z_m\right). \tag{20}$$

Como o lado direito é menor do que qualquer $\varepsilon > 0$ dado para n suficientemente grande. Portanto, a sequência arbitrária de Cauchy $\hat{x}_n \in \hat{X}$ tem o limite $\hat{x} \in \hat{X}$, e $\hat{X}$ é completo.

(d) Unicidade de $\hat{X}$ exceto por isometrias. Se $(\tilde{X}, \tilde{d})$ é outro espaço métrico completo com um subespaço $\tilde{W}$ denso em $\tilde{X}$ e isométrico com X, então para quaisquer $\tilde{x}, \tilde{y} \in \tilde{X}$, temos sequências $(\tilde{x}_n), (\tilde{y}_n) \in \tilde{W}$ tais que $\tilde{x}_n \longrightarrow \tilde{x}$ e $\tilde{y}_n \longrightarrow \tilde{y}$, donde

$$\tilde{d}(\tilde{x}, \tilde{y}) = \lim_{n\to\infty} \tilde{d}\left(\tilde{x}_n, \tilde{y}_n\right) \tag{21}$$

segue que,

$$\left|\tilde{d}(\tilde{x}, \tilde{y}) - \tilde{d}\left(\tilde{x}_n, \tilde{y}_n\right)\right| \leqq \tilde{d}\left(\tilde{x}, \tilde{x}_n\right) + \tilde{d}\left(\tilde{y}, \tilde{y}_n\right) \quad \longrightarrow \quad 0 \tag{22}$$

Como $\tilde{W}$ é isométrico com $W \subset \hat{X}$ e $\overline{W} = \hat{X}$, as distâncias em $\tilde{X}$ e $\hat{X}$ devem ser iguais. Portanto, $\tilde{X}$ e $\hat{X}$ são isométricos. ∎

Capítulo 1

Espaços de Banach

Espaços vetoriais desempenham um papel em muitos ramos da matemática e suas aplicações. De fato, em vários problemas práticos (e teóricos) nós temos um conjunto X cujos elementos podem ser vetores no espaço tridimensional, ou sequências de números, ou funções, e esses elementos podem ser somados e multiplicados por constantes (números) de forma natural, o resultado sendo novamente um elemento de X. Tais situações concretas sugerem o conceito de um espaço vetorial como definido abaixo. A definição vai envolver um corpo geral $\mathbb{K}$, mas na análise funcional, $\mathbb{K}$ será $\mathbb{R}$ ou $\mathbb{C}$. Os elementos de $\mathbb{K}$ são chamados escalares; portanto, no nosso caso, eles serão números reais ou complexos.

1.1 Espaços Vetoriais

Espaços vetoriais desempenham um papel em muitos ramos da matemática e suas aplicações. De fato, em vários problemas práticos (e teóricos) nós temos um conjunto X cujos elementos podem ser vetores no espaço tridimensional, ou sequências de números, ou funções, e esses elementos podem ser somados e multiplicados por constantes (números) de forma natural, o resultado sendo novamente um elemento de X. Tais situações concretas sugerem o conceito de um espaço vetorial como definido abaixo. A definição vai envolver um corpo geral $\mathbb{K}$, mas na análise funcional, $\mathbb{K}$ será $\mathbb{R}$ ou $\mathbb{C}$. Os elementos de $\mathbb{K}$ são chamados escalares; portanto, no nosso caso, eles serão números reais ou complexos.

Definição 1.1.1 Um espaço vetorial (ou espaço linear) sobre um corpo $\mathbb{K}$ é um conjunto não vazio X onde seus elementos são chamados de vetores e estão definidas duas operações: a adição que faz corresponder a cada par de vetores $x, y \in X$ o vetor $x + y \in X$ e a multiplicação por escalar que faz corresponder a cada $x \in X$ e $\alpha \in \mathbb{K}$ o vetor $\alpha x \in X$. Essas operações devem satisfazer, para quaisquer $x, y, z \in X$ e $\alpha, \beta \in \mathbb{K}$ as seguintes propriedades abaixo, que são chamadas de axiomas do espaço vetorial:

1. $x + y = y + x$

2. $x + (y + z) = (x + y) + z$

3. Existe um vetor 0, chamado vetor nulo, tal que $x + 0 = x$ para todo $x \in X$

4. Para cada $x \in X$, existe um $-x \in X$, chamado de inverso aditivo tal que $x+(-x) = (-x) + x = 0$

5. $\alpha(\beta x) = (\alpha\beta)x$

6. $1x = x$

7. $\alpha(x + y) = \alpha x + \alpha y$

8. $(\alpha + \beta)x = \alpha x + \beta x$

O vetor zero (ou nulo), é por definição, denotado por 0, entretanto ele não se refere ao escalar $0 \in \mathbb{K}$, no geral usaremos a mesma notação para os dois, sem risco de confusão, em casos específicos usaremos a notação 0_v para se referir ao vetor zero. Definimos o inverso aditivo de $x = (\xi_1, \xi_2, ..., \xi_n)$ como $-x = (-\xi_1, -\xi_2, ..., -\xi_n)$.

Como consequência dos axiomas acima, valem as seguintes propriedades:

1. Se $x+y = x+z$ então $y = z$. Em particular, $x+y = 0 \Rightarrow y = 0$ e $x+y = 0 \Rightarrow x = -y$.

2. $0x = 0_v$ e $\alpha 0_v = 0_v$

3. $(-1)x = -x$

Exemplo 1.1.1 Espaço $\mathbb{R}^n$. Este é o espaço euclidiano introduzido no exemplo 0.0.4, sendo o conjunto de todas as n-uplas de números reais, escritos $x = (\xi_1, \cdots, \xi_n), y = (\eta_1, \cdots, \eta_n)$, e agora vemos que este é um espaço vetorial com as duas operações algébricas definidas por:

$$x + y = (\xi_1 + \eta_1, \cdots, \xi_n + \eta_n)$$
$$\alpha x = (\alpha\xi_1, \cdots, \alpha\xi_n)$$

$(\alpha \in \mathbb{R})$.

Exemplo 1.1.2 Espaço $\mathbb{C}^n$. Este espaço foi definido no exemplo 0.0.4. Ele consiste de todas as n-uplas ordenadas de números complexos $x = (\xi_1, \cdots, \xi_n), y = (\eta_1, \cdots, \eta_n)$, e é um espaço vetorial complexo com as operações algébricas definidas como no exemplo anterior, onde agora $\alpha \in \mathbb{C}$.

Exemplo 1.1.3 Espaço C[a,b]. Este espaço foi definido no exemplo 0.0.6. Cada ponto deste espaço é uma função contínua com valor real em $[a, b]$. O conjunto de todas essas funções formam um espaço vetorial real com as operações algébricas definido da maneira usual:

$$(x + y)(t) = x(t) + y(t)$$
$$(\alpha x)(t) = \alpha x(t)$$

$(\alpha \in \mathbb{R})$. Na verdade, $x + y$ e αx são funções contínuas com valor real definidas em $[a, b]$ se x e y são funções contínuas e α é real.
Outros espaços vetoriais importantes de funções são:
(a) o espaço vetorial $B(A)$ definido no exemplo 0.0.7;
(b) o espaço vetorial de todas as funções diferenciáveis em R;
(c) o espaço vetorial de todas as funções reais em $[a, b]$ que são integráveis em algum sentido.

Exemplo 1.1.4 Espaço l^2. Este espaço foi introduzido no exemplo 0.0.8. É um espaço vetorial com as operações algébricas definidas como de costume em conexão com sequências, ou seja,

$$(\xi_1, \xi_2, \cdots) + (\eta_1, \eta_2, \cdots) = (\xi_1 + \eta_1, \xi_2 + \eta_2, \cdots)$$
$$\alpha\,(\xi_1, \xi_2, \cdots) = (\alpha\xi_1, \alpha\xi_2, \cdots)\,.$$

De fato, $x = (\xi_j) \in l^2$ e $y = (\eta_j) \in l^2$ implica $x + y \in l^2$, como segue prontamente da desigualdade de Minkowski; também temos $\alpha x \in l^2$.
Outros espaços vetoriais cujos pontos são sequências são l^∞ definido no exemplo 0.0.5 e l^p em 0.0.8, onde $1 \leqslant p < +\infty$.

Um **subespaço vetorial** de um espaço vetorial X é um subconjunto não-vazio Y de X tal que, para todo $y_1, y_2 \in Y$ e todos os escalares α, β temos que $\alpha y_1 + \beta y_2 \in Y$. Além disso, o vetor nulo também tem que pertencer a Y. Concluímos que Y é ele próprio um espaço vetorial, onde as duas operações algébricas são aquelas induzidas por X.

Uma **combinação linear** de vetores $x_1, ..., x_n$ de um espaço vetorial X é uma expressão da forma

$$\alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n$$

onde α_i são escalares quaisquer.

Para qualquer subconjunto não vazio $M \subset X$, o conjunto de todas as combinações lineares de vetores de M é chamada de span de M, escreveremos **SpanM**. Pode se mostrar rapidamente que $SpanM$ é um subespaço vetorial de X.

Definição 1.1.2 (Independência Linear, Dependência Linear) Independência e Dependência Linear de um conjunto M de vetores $x_1, ..., x_n$ (n⩾1) em um espaço vetorial X são definidos através da seguinte equação:

$$\alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n = 0. \qquad (1)$$

onde α_i são escalares. Claramente, a equação (1) vale para $\alpha_1 = \alpha_2 = ... = \alpha_n = 0$. Se esta é a única n-upla de escalares para qual a equação (1) vale, dizemos que M é linearmente independente (LI). Diz-se que M é linearmente dependente (LD) se M não for LI, isto é, se (1) vale para alguma n-upla de escalares, onde nem todos eles são iguais a 0.

Definição 1.1.3 (Espaços vetoriais de dimensão finita e infinita) Um espaço vetorial X é dito ser de dimensão infinita se existir um inteiro n tal que X contém um conjunto linearmente independente de n vetores, enquanto que qualquer conjunto de $n+1$ ou mais vetores de X é linearmente dependente. n é chamado de dimensão de X e escreveremos $dimX = n$. Por definição, $X = \{0\}$ tem dimensão finita e $dimX = 0$. Se X não é de dimensão finita, é dito ser de dimensão infinita.

Se $dimX = n$, uma n-upla linearmente independente de vetores de X é chamada de **base** de X (ou base em X). Se $(e_1, e_2, ..., e_n)$ é uma base para X, cada $x \in X$ tem uma representação única como uma combinação linear dos vetores da base:

$$x = e_1x_1 + e_2x_2 + ... + e_nx_n.$$

Por exemplo, uma base para o $\mathbb{R}^n$ é:

$$\begin{aligned}
&e_1 = (1, 0, 0, \cdots, 0)\\
&e_2 = (0, 1, 0, \cdots, 0)\\
&\vdots\\
&e_n = (0, 0, 0, \cdots, 1)
\end{aligned}$$

Chamamos a base acima de base canônica do $\mathbb{R}^n$.

Se X é qualquer espaço vetorial, não necessariamente de dimensão finita, e B é um subconjunto linearmente independente de X, tal que $spanB = X$, então B é chamado de **base** (ou **base de hamel**) para X. Portanto, se B é uma base para X, então todo $x \in X$ diferente de zero tem uma representação única como uma combinação linear de elementos de B com escalares diferentes de zero. Mostraremos futuramente que todo espaço vetorial X não identicamente nulo, possui uma base.

Teorema 1.1.4 ***(Dimensão de um subespaço)*** *Seja X um espaço vetorial n-dimensional. Então qualquer subespaço próprio Y de X tem dimensão menor que n.*

Demonstração:
Se $n = 0$, então $X = \{0\}$ e não tem subespaço próprio de X. Se $dimY = 0$, então $Y = \{0\}$ e $Y \neq X$ implica $dimX \geqslant 1$, donde $dimY \leqslant dimX = n$. Se $dimY$ fosse igual a n, então Y teria uma base de n elementos, que também seria uma base para X uma vez que $dimX = n$, de modo que $Y = X$. Isso mostra que qualquer conjunto linearmente independente de vetores de Y deve ter menos de n elementos, e $dimY < n$. ■

1.2 Espaços Normados (Espaços de Banach)

Definição 1.2.1 Um espaço normado X é um espaço vetorial com uma norma definida nele, um espaço de Banach é um espaço normado completo (completo na métrica definida pela norma). Aqui uma **norma** em um espaço vetorial (real ou complexo) X é uma função real em X cujo valor em cada $x \in X$ é denotado por

$$||x||.$$

Lê-se "norma de x".
E que tem as seguintes propriedades:

(N1) $\|x\| \geqq 0$;

(N2) $\|x\| = 0 \iff x = 0$;

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

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

Onde x e y são vetores arbitrários de X e α um escalar qualquer. A Propriedade (N4) é chamada de desigualdade-triangular.

Uma norma em X define uma métrica d em X definida por

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

e é chamada de métrica induzida pela norma.

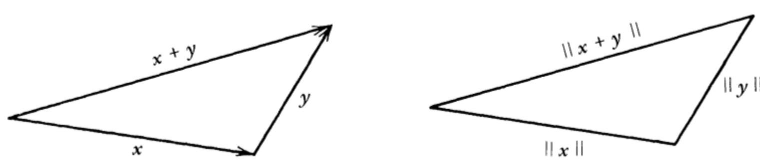

Figura 1.1: Ilustração da desigualdade-triangular (N4).

Proposição 1.2.2 *Seja X um espaço normado. Para todo $x, y \in X$ vale que $|\,||y|| - ||x||\,| \leqslant ||y - x||$*

Demonstração:
Note que,

$$||x|| = ||x - y + y|| \leqslant ||x - y|| + ||y|| \Rightarrow ||x|| - ||y|| \leqslant ||x - y||. \qquad (1)$$

Por outro lado,

$$||y|| = ||y-x+x|| \leqslant ||y-x||+||x|| \Rightarrow ||y||-||x|| \leqslant ||y-x|| \Rightarrow -(||x||-||y||) \leqslant ||x-y||. \qquad (2).$$

Daí, por (1) e (2) obtemos

$$| \, ||y|| - ||x|| \, | \leqslant ||y - x||$$

■

Exemplo 1.2.1 Espaço euclidiano $\mathbb{R}^n$ e espaço unitário $\mathbb{C}^n$. São espaços de Banach com norma definida por

$$\|x\| = \left(\sum_{j=1}^{n} |\xi_j|^2 \right)^{1/2} = \sqrt{|\xi_1|^2 + \cdots + |\xi_n|^2}.$$

A norma acima, produz a métrica definida no exemplo 0.0.4,

$$d(x, y) = ||x - y|| = \sqrt{|\xi_1 - \eta_1|^2 + \cdots + |\xi_n - \eta_n|^2} \qquad (\geqslant 0).$$

No caso particular do $\mathbb{R}^{\nvDash}$ nós temos

$$||x|| = |x| = \sqrt{(\xi_1)^2 + (\xi_2)^2 + (\xi_3)^2}.$$

Isso confirma nossa observação anterior de que a norma generaliza a noção elementar do comprimento $|x|$ de um vetor.

Exemplo 1.2.2 Espaço l^2. Este é o espaço definido no exemplo 0.0.8. É um espaço de Banach com norma definida por

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

Na verdade, esta norma induz a seguinte métrica

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

Exemplo 1.2.3 Espaço l^2. Este espaço foi definido no exemplo 0.0.5 e é um espaço de Banach, uma vez que sua métrica é obtida a partir da norma definida por

$$||x|| = \sup_j |\xi_j|.$$

Exemplo 1.2.4 Espaço C[a,b]. O espaço definido no exemplo 0.0.6 é um espaço de Banach definido com norma definida por

$$||x|| = \max_j |x(t)|$$

onde $j \in J = [a, b]$.

Lema 1.2.3 ***(Translação invariante)*** *Uma métrica d induzida por uma norma em um espaço normado X satisfaz*

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

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

para todo $x, y, a \in X$ e todo escalar α.

Demonstração:
Temos que,

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

e

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

■

Por definição, um subespaço Y de um espaço normado X é um subespaço de X considerado como um espaço vetorial, com a norma obtida restringindo a norma em X ao subconjunto Y. Esta norma em Y é induzida pela norma em X. Se Y é fechado em X, então Y é chamado de subespaço fechado de X.

Por definição, um subespaço Y de um espaço Banach X é um subespaço de X considerado como um espaço normado. Portanto, não exigimos que Y seja completo.

Nesse sentido, o Teorema 0.0.10 é útil, pois produz imediatamente o seguinte resultado

Teorema 1.2.4 *Um subespaço Y de um espaço de banach X é completo se, somente se, o conjunto Y é fechado em X.*

Definição 1.2.5 (i) Uma sequência (x_n) em um espaço normado X é dita convergente, se X contém um x tal que

$$\lim_{n \to \infty} \|x_n - x\| = 0$$

também escrevemos $x_n \longrightarrow x$ e dizemos que x é o limite de (x_n)

(ii) Uma sequência (x_n) em um espaço normado X é dita de *cauchy* se para cada $\epsilon > 0$, existe um N tal que para todo $m, n > N$ vale

$$\|x_m - x_n\| < \varepsilon$$

Definição 1.2.6 Se (x_k) é uma sequência em um espaço normado X, podemos associar com (x_k) a sequência s_n de somas parciais

$$s_n = x_1 + x_2 + \cdots + x_n$$

Se s_n é convergente, digamos $s_n \longrightarrow s$ então a série infinita, ou simplesmente, série

$$\sum_{k=1}^{\infty} x_k = x_1 + x_2 + \cdots$$

é dita convergente, s é chamado a soma da série e nos escrevemos

$$s = \sum_{k=1}^{\infty} x_k = x_1 + x_2 + \cdots \qquad (1)$$

Se $||x_1|| + ||x_2|| + \cdots$ a série (1) é dita absolutamente convergente. No entanto, alertamos o leitor que em um espaço normado X, convergência absoluta implica convergência se e somente se X é completo.

O conceito de convergência de uma série pode ser usado para definir uma base da seguinte forma. Se um espaço normalizado X contém uma sequência (e_n) com a propriedade de que para cada $x \in X$ existe uma única sequência de escalares (α_n) tal que

$$\|x - (\alpha_1 e_1 + \cdots + \alpha_n e_n)\| \longrightarrow 0 \quad (\text{ com } n \longrightarrow \infty)$$

então (e_n) é chamado de base de Schauder (ou base) para X. A série

$$\sum_{k=1}^{\infty} \alpha_k x_k$$

que tem a soma x é então chamada de expansão de x em relação a (e_n), e escrevemos

$$x = \sum_{k=1}^{\infty} \alpha_k x_k$$

Teorema 1.2.7 *Seja $X = (X, ||\cdot||)$ um espaço normado. Então existe um espaço de Banach $\hat{X}$ e uma isometria A de X em um subespaço W de $\hat{X}$ que é denso em $\hat{X}$. O espaço $\hat{X}$ é único exceto por isometrias.*

Demonstração:

Pelo teorema do completamento (teorema 0.0.13), existe um espaço métrico completo $\hat{X} = (\hat{X}, \hat{d})$ e uma isometria $A : X \longrightarrow W = A(X)$ onde W é denso em $\hat{X}$ e $\hat{X}$ é único exceto por isometrias. Para provar o presente teorema devemos fazer de $\hat{X}$ um espaço vetorial e introduzir uma norma adequada no mesmo.

Para definir em $\hat{X}$ as duas operações algébricas de um espaço vetorial, considere quaisquer $\hat{x}, \hat{y} \in \hat{X}$ e representantes arbitrários $(x_n) \in \hat{x}$ e $(y_n) \in \hat{y}$. Lembre-se que $\hat{x}$ e $\hat{y}$ são classes de equivalência de sequências de Cauchy em X. Definimos $z_n = x_n + y_n$. Então (z_n) é de Cauchy em X desde

$$\|z_n - z_m\| = \|x_n + y_n - (x_m + y_m)\| \leqq \|x_n - x_m\| + \|y_n - y_m\|.$$

Definimos a soma $\hat{z} = \hat{x} + \hat{y}$ de $\hat{x}$ e $\hat{y}$ para ser a classe de equivalência para qual (z_n) é um representante; assim $(z_n) \in \hat{z}$. Esta definição é independente da escolha particular de sequências de Cauchy pertencentes a $\hat{x}$ e $\hat{y}$. De fato, se $(x_n) \sim (x_n')$ e $(y_n) \sim (y_n')$ então $(x_n + y_n) \sim (x_n' + y_n')$ pois

$$\|x_n + y_n - (x_n' + y_n')\| \leqq \|x_n - x_n'\| + \|y_n - y_n'\|.$$

Da mesma forma, definimos o produto $\alpha\hat{x} \in \hat{X}$ de um escalar α e $\hat{x}$ para ser a classe de equivalência para a qual (αx_n) é um representante. Novamente, essa definição é independente da escolha particular de um representante de $\hat{x}$. O elemento zero de $\hat{X}$ é a classe de equivalência que contém todos as sequências de Cauchy que convergem para zero. Não é difícil ver que aquelas duas operações algébricas têm todas as propriedades exigidas pela definição, de modo que $\hat{X}$ é um espaço vetorial. Da definição segue que em W as operações do espaço vetorial induzidas de $\hat{X}$ concordam com aquelas induzidas de X por meio de A.

Além disso, A induz em W uma norma $||\cdot||_1$, cujo valor para cada $\hat{y} = Ax \in W$ é $||\hat{y}||_1 = ||x||$. A métrica correspondente em W é a restrição de $\hat{d}$ a W já que A é isometria. Podemos estender a norma $||\cdot||_1$ para $\hat{X}$ definindo $||\hat{x}||_2 = \hat{d}(\hat{0}, \hat{x})$ para todo $\hat{x} \in \hat{X}$. De fato, é de fácil verificação que $||\cdot||_2$ satisfaz as propriedades (N1) e (N2) da definição de norma, e os outros dois axiomas (N3) e (N4) seguem daqueles para $||\cdot||_1$ por um processo de limites.

■

1.3 Espaços Normados de Dimensão Finita e Subespaços

Lema 1.3.1 ***(Combinações Lineares).*** *Seja $\{x_1, x_2, ..., x_n\}$ um conjunto linearmente independente de vetores em um espaço normado X (de qualquer dimensão). Então existe um número $c > 0$ tal que para cada escolha de escalares $\alpha_1, \cdots, \alpha_n$, temos*

$$\|\alpha_1 x_1 + \cdots + \alpha_n x_n\| \geqq c\,(|\alpha_1| + \cdots + |\alpha_n|) \quad (c > 0) \qquad (1)$$

Demonstração:

Escrevemos $s = |\alpha_1| + \cdots + |\alpha_n|$. Se $s = 0$ todos os α_j são iguais a zero, e a equação (1) vale para qualquer c. Seja $s > 0$, então dividindo a equação (1) por s e fazendo $\beta_j = \frac{\alpha_j}{s}$, obtemos

$$\|\beta_1 x_1 + \cdots + \beta_n x_n\| \geqq c \quad \left(\sum_{j=1}^{n} |\beta_j| = 1\right) \qquad (2)$$

portanto basta provar que existe $c > 0$ tal que (2) vale para cada n-upla de escalares $\beta_1, \cdots, \beta_2$, com $\sum_{j=1}^{n} |\beta_j| = 1$.

Suponha que isso seja falso. Então existe uma sequência (y_m) de vetores

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

de tal modo que $||y_m|| \longrightarrow 0$ se $m \longrightarrow \infty$.

Note que, como $\sum_{j=1}^{n} \left|\beta_j^{(m)}\right| = 1$, temos $|\beta_j^{(m)}| \leqslant 1$. Por isso, para cada j fixo, a sequência $(\beta_j^{(m)}) = (\beta_j^{(1)}, \beta_j^{(2)}, \cdots)$ é limitada.

Consequentemente, Pelo Teorema de Bolzano-Weierstrass, $(\beta_1^{(m)})$ tem uma subsequência convergente. Seja β_1 o limite dessa subsequência e denote por $(y_{1,m})$ a subsequência correspondente de (y_m). Usando o mesmo argumento $(y_{1,m})$ possui uma subsequência $(y_{2,m})$ para qual a subsequência correspondente de escalares $(\beta_2^{(m)})$ converge. Seja β_2 o limite dessa subsequência. Prosseguindo dessa maneira, após n passos obtemos uma subsequência $(y_{n,m}) = (y_{1,m}, y_{2,m}, \cdots)$ de (y_n) cujos termos são da forma

$$y_{n,m} = \sum_{j=1}^{n} \gamma_j^{(m)} x_j \quad \left(\sum_{j=1}^{n} \left|\gamma_j^{(m)}\right| = 1\right)$$

com escalares $\gamma_j^{(m)}$ satisfazendo $\gamma_j^{(m)} \longrightarrow \beta_j$ quando $m \longrightarrow \infty$. Daí, como $m \longrightarrow \infty$, temos

$$y_{n,m} \longrightarrow y = \sum_{j=1}^{n} \beta_j x_j$$

onde $\sum_{j=1}^{n} |\beta_j| = 1$, de modo que nem todo β_j pode ser zero. Como $\{x_1, x_2, ..., x_n\}$ é um conjunto linearmente independente, temos assim que $y \neq 0$. Por outro lado, $y_{n,m} \longrightarrow y$ implica $||y_{n,m}|| \longrightarrow ||y||$ pela continuidade da norma. Como $||y_m|| \longrightarrow 0$ por suposição e $(y_{n,m})$ é subsequência de (y_m), devemos ter que $||y_{n,m}|| \longrightarrow 0$. Portanto $||y|| = 0$, de modo que $y = 0$ pela propriedade (N2) da definição de norma. Isso contradiz $y \neq 0$, e o lema está provado. ∎

A seguir mostramos a completude dos espaços vetoriais de dimensão finita.

Teorema 1.3.2 ***(Completude)*** *Todo subespaço Y de dimensão finita de um espaço normado X é completo. Em particular, todo espaço normado de dimensão finita é completo.*

Demonstração:

Considere uma sequência de Cauchy arbitrária (y_m) em Y e mostraremos que é convergente em Y; o limite será denotado por y. Seja $dimY = n$ e $\{e_1, e_2, ..., e_n\}$ qualquer base para Y. Então, cada y_m tem uma representação única da forma

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

Como (y_m) é uma sequência de Cauchy, para todo $\epsilon > 0$ existe um N tal que $||y_m - y_r|| < \epsilon$ quando $m, r > N$. Usando esse fato e o lema (1.3.1), temos que existe um $c > 0$ tal que

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

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

$$\left| \alpha_j^{(m)} - \alpha_j^{(r)} \right| \leqq \sum_{j=1}^{n} \left| \alpha_j^{(m)} - \alpha_j^{(r)} \right| < \frac{\varepsilon}{c} \quad (m, r > N)$$

Isso mostra que cada uma das n sequências $(\alpha_j^{(m)}) = (\alpha_j^{(1)}, \alpha_j^{(2)}, \cdots), 1 \leqslant j \leqslant n$, é de Cauchy em $\mathbb{R}$ ou $\mathbb{C}$, logo convergente; seja (α_j) esse limite. Usando esses n limites $\alpha_1, \cdots, \alpha_n$ definimos

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

claramente $y \in Y$. Além disso,

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

onde $\alpha_j^{(m)} \longrightarrow \alpha_j$. Donde, $||y_m - y|| \longrightarrow 0$, isto é $y_m \longrightarrow y$. Isso mostra que (y_m) é convergente em Y. Uma vez que (y_m) é uma sequência arbitrária e de Cauchy em Y, isto prova que Y é completo. ∎

Como consequência desse teorema e do teorema 0.0.10 podemos enunciar o seguinte

Corolário 1.3.3 ***(Fechamento)*** *Todo subespaço de dimensão finita Y de um espaço normado X, é fechado em X.*

Definição 1.3.4 (Normas equivalentes) Uma norma $|| \cdot ||$ em um espaço vetorial X é dita equivalente a uma norma $|| \cdot ||_0$ em X se existirem a e b positivos tais que para todo $x \in x$ vale

$$a\|x\|_0 \leqq \|x\| \leqq b\|x\|_0$$

Teorema 1.3.5 ***(Normas equivalentes)*** *Em um espaço vetorial de dimensão finita X, qualquer norma* $||\cdot||$ *é equivalente a qualquer outra norma* $||\cdot||_0$

Demonstração:
Seja $dimX = n$ e $\{e_1, e_2, ..., e_n\}$ uma base qualquer para X. Então, todo $x \in X$ tem um representação única da seguinte forma

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

pelo lema (1.3.1) existe uma constante $c > 0$ tal que

$$\|x\| \geqq c\,(|\alpha_1| + \cdots + |\alpha_n|)$$

Por outro lado, a desigualdade triangular nos fornece

$$\|x\|_0 \leqq \sum_{j=1}^{n} |\alpha_j|\, \|e_j\|_0 \leqq k \sum_{j=1}^{n} |\alpha_j|\,, \quad k = \max_j \|e_j\|_0$$

juntos, $a\|x\|_0 \leqq \|x\|$ onde $a = c/k > 0$. A outra desigualdade segue de maneira análoga, repetindo o mesmo argumento acima e trocando os papéis de $||\cdot||$ e $||\cdot||_0$. ■

1.4 Compacidade e dimensão finita

Algumas outras propriedades básicas de espaços normados de dimensão finita e subespaços estão relacionados com o conceito de compacidade. Este último é definido a seguir.

Definição 1.4.1 Diz-se que um espaço métrico X é dito compacto se toda sequência em X tem uma subsequência convergente. Um subconjunto M de X é dito compacto se M for compacto considerado como um subespaço de X, isto é, se toda sequência em M tem uma subsequência cujo limite é um elemento de M.

Uma propriedade geral de conjuntos compactos é expressa no seguinte lema

Lema 1.4.2 *Um subconjunto compacto M de um espaço métrico é fechado e limitado.*

Demonstração:
Para todo $x \in \overline{M}$ existe uma sequência (x_n) em M tal que $x_n \longrightarrow x$ (pelo teorema 0.0.9). Como M é compacto, segue que M é fechado, então $x \in M$. logo M é fechado pois $x \in \overline{M}$ é arbitrário. Vamos provar que M é limitado. Se M fosse ilimitado, ele conteria uma sequência ilimitada (y_n) tal que $d(y_n, b) > n$ onde b é qualquer elemento fixo. Esta sequência não pode tem uma subsequência convergente, uma vez que uma subsequência convergente deve ser limitada, pelo teorema 0.0.5. Portanto, M é limitado. ■

A recíproca deste lema é em geral falsa.

prova: Para provar este fato importante, consideramos a sequência $(e_n) \in l^2$ onde $e_n = (\delta_{nj})$ tem o enésimo termo igual a 1 e o restante igual a 0. Esta sequência é limitada pois $||e_n|| = 1$. Seus termos constituem um conjunto de pontos que é fechado porque não tem ponto de acumulação. Pela mesma razão, esse conjunto de pontos não é compacto.

No entanto, para um espaço normado de dimensão finita temos

Teorema 1.4.3 *Em um espaço normado de dimensão finita X, qualquer subconjunto $M \subset X$ é compacto se, e somente se, M é fechado e limitado.*

Demonstração:

Pelo lema anterior compacidade implica fechamento e limitação, logo a ida está feita. Vamos provar a recíproca. Suponha M fechado e limitado. Seja $dim X = n$ e $\{e_1, e_2, \cdots, e_n\}$ uma base qualquer de X. Consideramos qualquer sequência (x_m) em M. Cada x_m tem uma representação

$$x_m = \xi_1^{(m)} e_1 + \cdots + \xi_n^{(m)} e_n$$

Como M é limitado, x_m também é, digamos $||x_m|| \leqslant n$. Pelo lema 1.3.1,

$$k \geqq \|x_m\| = \left\| \sum_{j=1}^{n} \xi_j^{(m)} e_j \right\| \geqq c \sum_{j=1}^{n} \left| \xi_j^{(m)} \right|$$

onde $c > 0$. Portanto a sequência de números (ξ_j^m) (j fixo) é limitada, e pelo Teorema de Bolzano-Weierstrass, tem um ponto de acumulação ξ_j, onde $1 \leqslant j \leqslant n$. Como na prova do Lema 1.3.1 temos que concluir que (x_m) tem uma subsequência (z_m) que converge para $z = \sum \xi_j e_j$. Como M é fechado, $z \in M$. Isso mostra que a sequência arbitrária (x_m) em M tem uma subsequência que converge em M. Portanto, M é compacto. ∎

Nossa discussão mostra o seguinte. Em $\mathbb{R}^n$ (ou em qualquer outro espaço normalizado de dimensão finita) os subconjuntos compactos são precisamente os subconjuntos fechados e limitados, de modo que esta propriedade (fechamento e limitação) pode ser usado para definir compacidade. No entanto, isso pode não mais ser feito no caso de um espaço normado de dimensão infinita.

Uma fonte de outros resultados interessantes é o seguinte Lema de F. Riesz

Lema 1.4.4 ***(Lema de Riesz)*** *Sejam Y e Z subespaços de um espaço normado X(de qualquer dimensão). Considere Y um subespaço fechado próprio de Z. Então para qualquer número real $\theta \in (0, 1)$, existe $z \in Z$ tal que*

$(i) ||z|| = 1$

$(ii) ||z - y|| \geqslant \theta, \quad \forall \quad y \in Y$

Demonstração:

Tome $v \in (Z - Y)$ e seja $a = \inf_{y \in Y} ||v - y||$. $a > 0$, pois Y é fechado (caso contrário $v \in Y$). Tome $\theta \in (0, 1)$ daí por definição existe $y_0 \in Y$ tal que

$$a \leqslant ||v - y_0|| \leqslant \frac{a}{\theta}.$$

defina $z = c(v - y_0)$ onde $c = \frac{1}{||v - y_0||}$.

$\Rightarrow ||z|| = ||c(v - y_0)|| = \frac{1}{||v - y_0||} \cdot ||v - y_0|| = 1.$

Logo $||z|| = 1$. Note também que $||z - y|| \geqslant a$. Resta mostrar que $||z - y|| \geqslant \theta$.

Com efeito, $||z - y|| = ||c(v - y_0) - y|| = c||v - y_o - y \cdot c^{-1}|| = c||v - y_1|| \geqslant c \cdot a = \frac{a}{||v - y_0||} \geqslant \frac{a}{a/\theta} = \theta.$
onde $y_1 = y_0 - y \cdot c^{-1}$.

isso implica que $||z - y|| \geqslant \theta$. Como y era arbitrário a prova está completa.

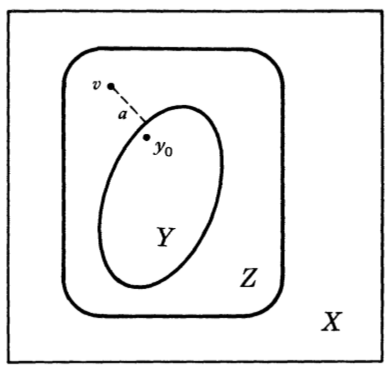

Figura 1.2: Ilustração representando a demonstração do Lema de Riesz.

■

Teorema 1.4.5 ***(dimensão finita).*** *Se um espaço normado X tem a propriedade de que a bola unitária fechada $M = \{x; ||x|| \leqslant 1\}$ é compacta, então X tem dimensão finita.*

Demonstração:

Assumimos que M é compacto, mas $dimX = \infty$, e mostramos que isso leva a uma contradição. Escolhemos qualquer x_1 de norma 1. Este x_1 gera um subespaço unidimensional X_1 de X, que é fechado e é um subespaço próprio de X já que $dimX = \infty$. Pelo Lema de Riesz, existe $x_2 \in X$ de norma 1 tal que

$$||x_2 - x_1|| \geqslant \theta = \frac{1}{2}.$$

Os elementos x_1, x_2 geram um subespaço fechado bidimensional próprio X_2 de X. Pelo lema de Riesz existe um x_3 de norma 1 tal que para todos os $x \in X_2$ temos

$$||x_3 - x|| \geqslant \frac{1}{2}.$$

em particular

$$||x_3 - x_1|| \geqslant \frac{1}{2},$$
$$||x_3 - x_2|| \geqslant \frac{1}{2}.$$

Procedendo por indução, obtemos uma sequência (x_n) de elementos $x_n \in M$ tal que

$$||x_m - x_n|| \geqslant \frac{1}{2}.$$

Obviamente, (x_n) não pode ter uma subsequência convergente. Isso contradiz a compacidade de M. Portanto, nossa suposição de que $dimX = \infty$ é falsa e $dimX < \infty$. ■

Conjuntos compactos são importantes porque são "bem comportados"; eles têm várias propriedades básicas semelhantes às dos conjuntos finitos e não compartilhados por conjuntos não compactos. No caso de aplicações contínuas, uma propriedade fundamental é que conjuntos compactos têm imagens compactas, como veremos a seguir.

Teorema 1.4.6 *Sejam X e Y espaços métricos e $T : X \longrightarrow Y$ uma aplicação contínua. Se M é um subconjunto de X for compacto, então sua imagem $T(M)$ é compacta.*

Demonstração:
Pela definição de compacidade, basta mostrar que cada sequência (y_n) na imagem $T(M) \subset Y$ contém uma subsequência que converge em $T(M)$. Como $y_n \in T(M)$, temos que $y_n = T(x_n)$ para algum $x_n \in M$. Como M é compacto, (x_n) tem uma subsequência (x_{n_k}) que converge em M. A imagem de (x_{n_k}) é uma subsequência de y_n que converge em $T(M)$ pelo teorema 0.0.1, pois T é contínua. Portanto, $T(M)$ é compacto. ■

A partir deste teorema concluímos que a seguinte propriedade, bem conhecida do cálculo para funções contínuas, se espelha para espaços métricos.

Corolário 1.4.7 ***(Máximo e Mínimo)*** *Uma aplicação contínua T de um subconjunto compacto M de um espaço métrico X em $\mathbb{R}$ assume um máximo e um mínimo para alguns pontos em M.*

Demonstração:
$T(M) \subset \mathbb{R}$ é compacto pelo teorema anterior e fechado e limitado pelo lema 1.4.2, de modo que $infT(M) \in T(M)$ e $supT(M) \in T(M)$. E as imagens inversas desses dois pontos consistem em pontos de M em que $T(x)$ é mínimo ou máximo, respectivamente. ■

1.5 Operadores Lineares

Definição 1.5.1 (Operador Linear). Um operador linear T é um operador tal que

(i) O domínio $D(T)$ de T é um espaço vetorial e a imagem $Im(T)$ está em um espaço vetorial sobre o mesmo corpo;

(ii) Para todo $x, y \in D(T)$ e $\alpha \in \mathbb{K}$ vale

$$T(x+y) = T(x) + T(y)$$

$$T(\alpha x) = \alpha T(x)$$

Por simplicidade, as vezes escreveremos o item (ii) como

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y) \tag{1.1}$$

Notação:
$D(T)$ denota o domínio de T.
$Im(T)$ denota a imagem de T.
$N(T)$ denota o núcleo de T.

Por definição, o núcleo de T é o conjunto de todos os $x \in D(T)$ tal que $T(x) = O$. Algumas propriedades dos operadores lineares podem ser vistas a seguir

(**P1**) $T(0) = 0$.

(**P2**) $T(-x) = -T(x)$.

(**P3**) $T(x-y) = T(x) - T(y)$.

Exemplo 1.5.1 O operador identidade, $I : X \longrightarrow X$ definido por $I(x) = x$ para todo $x \in X$ é linear.

Exemplo 1.5.2 O operador nulo $0 : X \longrightarrow Y$ definido por $0(x) = 0$ para todo $x \in X$ é linear.

Exemplo 1.5.3 Seja X o espaço vetorial de todos os polinômios em $[a, b]$. Podemos definir um operador linear T em X definindo

$$T(x(t)) = x'(t)$$

para todo $x \in X$, onde a linha denota diferenciação em relação a t. Este operador T mapeia X sobre si mesmo.

Exemplo 1.5.4 Um operador linear T de $C[a, b]$ em si mesmo pode ser definido por

$$T(x(t)) = \int_a^t x(\tau)d\tau \quad t \in [a, b]$$

Iremos Provar a seguir um importante teorema que relaciona a dimensão da imagem de um operador com a dimensão de seu núcleo.

Teorema 1.5.2 ***(Imagem e Núcleo).*** *Seja T um operador linear. Temos:*

(a) *A imagem $Im(T)$ é um espaço vetorial.*

(b) *Se $dimD(T) = n < \infty$, então $dimIm(T) \leqslant n$.*

(c) *O núcleo $N(T)$ é um espaço vetorial.*

Demonstração:
(a) Tome qualquer $y_1, y_2 \in Im(T)$ e vamos mostrar que $\alpha y_1 + \beta y_2 \in Im(T)$ para quaisquer escalares α e β. Seja $y_1, y_2 \in Im(T)$, temos que $y_1 = T(x_1)$ e $y_2 = T(x_2)$ para algum $x_1, x_2 \in D(T)$ e $\alpha x_1 + \beta x_2 \in D(T)$ pois $D(T)$ é um espaço vetorial. Como T é linear, obtemos

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

donde $\alpha y_1 + \beta y_2 \in Im(T)$. Portanto, $Im(T)$ é um espaço vetorial.

(b) Escolhemos $n + 1$ elementos $y_1, \cdots, y_{n+1} \in Im(T)$ arbitrários. Temos que $y_1 = T(x_1), \cdots, y_{n+1} = T(y_{n+1})$ para $n+1$ vetores $x_1, \cdots, x_{n+1} \in D(T)$. Seja $dimD(T) = n$, o conjunto $\{x_1, \cdots, x_{n+1}\}$ deve ser linearmente dependente. Por isso

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

onde $\alpha_1, \cdots, \alpha_{n+1}$ são nem todos iguais a zero. Como T é linear e $T(0) = 0$, aplicando T em ambos os lados, obtemos

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

isso mostra que $\{y_1, \cdots, y_{n+1}\}$ é linearmente dependente porque nem todos os α_i são iguais a zero. lembrando que este conjunto de $Im(T)$ foi escolhido arbitrariamente, donde concluímos que $Im(T)$ não possui subconjuntos linearmente independentes de $n+1$ elementos ou mais. Por definição isso significa que $dimIm(T) \leqslant n$.

(c) Tome qualquer $x_1, x_2 \in D(T)$. Então $T(x_1) = T(x_2) = 0$. Como T é linear, para quaisquer escalares α e β, temos que

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

concluímos que $\alpha x_1 + \beta x_2 \in N(T)$. Logo, $N(T)$ é um espaço vetorial. ■

Uma consequência imediata da parte (b) da prova é a de que operadores lineares preservam dependência linear.

Vamos agora nos voltar para a inversa de um operador linear. Lembra que um operador $T : D(T) \longrightarrow Im(T)$ é injetivo se, somente se, pontos diferentes do domínio tem diferentes imagens. isto é, $x_1, x_2 \in D(T)$,

$$x_1 \neq x_2 \Longrightarrow T(x_1) \neq T(x_2).$$

equivalentemente,

$$T(x_1) = T(x_2) \Longrightarrow x_1 = x_2.$$

Nesse caso, existe o operador $T^{-1} : Im(T) \longrightarrow D(T)$ onde para todo $x_0 \in D(T)$ existe um $y_0 \in Im(T)$ tal que $T^{-1}(y_0) = x_0$. Chamamos o operador T^{-1} de inversa de T.

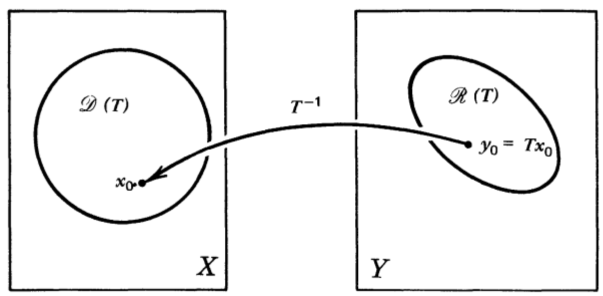

Figura 1.3: Ilustração representando a inversa de um operador linear.

É de verificação imediata os seguintes fatos,

$$T^{-1}(T(x)) = x \qquad \forall x \in D(T).$$

$$T(T^{-1}(y)) = y \qquad \forall y \in Im(T).$$

Em conexão com operadores lineares em espaços vetoriais, a situação é a seguinte, o inverso de um operador linear existe se, e somente se, o o núcleo do operador consiste apenas do vetor zero. Mais precisamente, temos o seguinte critério, que aplicaremos com bastante frequência.

Teorema 1.5.3 *Sejam X, Y espaços vetoriais, ambos reais ou ambos complexos. Seja $T : D(T) \longrightarrow Y$ um operador linear onde $D(T) \subset X$ e $Im(T) \subset Y$. Temos:*

(a) A inversa $T^{-1} : Im(T) \longrightarrow D(T)$ existe se, somente se, $T(x) = 0 \Rightarrow x = 0$.

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

(c) Se $dimD(T) = n < \infty$ e T^{-1} existe, então $dimIm(T) = dimD(T)$.

Demonstração:

(a) Suponha que $T(x) = 0 \Rightarrow x = 0$. Seja $T(x_1) = T(x_2)$. Como T é linear,

$$T(x_1 - x_2) = T(x_1) - T(x_2) = 0,$$

De modo que $x_1 - x_2 = 0$. Como $T(x_1) = T(x_2)$ implica que $x_1 = x_2$ temos que T é injetiva e existe T^{-1}. Por outro lado, suponha agora que T^{-1} exista, logo T é injetiva. Como T é injetiva vale que

$$T(x_1) = T(x_2) \Longrightarrow x_1 = x_2.$$

fazendo $x_2 = 0$, obtemos

$$T(x_1) = T(0) = 0 \Longrightarrow x_1 = 0$$

e a parte (a) está provada.

(b) Suponha que T^{-1} existe. Vamos mostrar que T^{-1} é linear. O domínio de T^{-1} é a $Im(T)$ que é um espaço vetorial pelo teorema 1.5.2(a). Consideremos qualquer $x_1, x_2 \in D(T)$ e suas imagens $y_1 = T(x_1)$ e $y_2 = T(x_2)$. Temos $x_1 = T^{-1}(y_1)$ e $x_2 = T^{-1}(y_2)$. T é linear, de modo que para quaisquer escalares α e β temos

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

Como $x_i = T^{-1}(y_i)$, isso implica

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

Portanto, T^{-1} é linear.

(c) Nós temos $dimIm(T) \leqslant dimD(T)$ pelo teorema 1.5.2(b). e $dimD(T) \leqslant dimIm(T)$ pelo mesmo teorema, aplicado a T^{-1}. Portanto $dimIm(T) = dimD(T)$. ■

Finalmente mencionamos uma fórmula útil para a inversa da composição de operadores lineares (O leitor talvez conheça esta fórmula para o caso de matrizes quadradas).

Lema 1.5.4 *Seja $T : X \longrightarrow Y$ e $S : Y \longrightarrow Z$ dois operadores lineares bijetivos e X, Y e Z espaços vetoriais. Então a inversa $(ST)^{-1} : Z \longrightarrow Y$ do produto (da composta) ST existe, e*

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

Demonstração:

O operador $(ST): X \longrightarrow Z$ é bijetivo, logo $(ST)^{-1}$ existe. Temos assim

$$ST(ST)^{-1} = I_Z$$

onde I_Z é o operador identidade em Z. Aplicando S^{-1} e usando que $S^{-1}S = I_Y$ (operador identidade em Y), nós obtemos

$$S^{-1}ST(ST)^{-1} = T(ST)^{-1} = S^{-1}I_Z = S^{-1} \tag{1.2}$$

Aplicando T^{-1} e usando que $T^{-1}T = I_X$ (operador identidade em X), nós obtemos o resultado desejado

$$T^{-1}T(ST)^{-1} = (ST)^{-1} = T^{-1}S^{-1} \tag{1.3}$$

e a prova está completa. ■

1.6 Operadores Lineares Contínuos e Limitados

Definição 1.6.1 (Operadores Lineares Limitados). Sejam X e Y espaços normados e $T : D(T) \longrightarrow Y$ um operador linear, onde $D(T) \subset X$. O operador T é dito limitado se existe um número real c tal que para todo $x \in D(T)$,

$$||T(x)|| \leqslant c||x||. \qquad (1)$$

Em (1) a norma à esquerda é aquela em Y, e a norma à direita é a norma em X. Por simplicidade, denotamos ambas as normas pelo mesmo símbolo $|| \cdot ||$, sem perigo de confusão. A fórmula (1) mostra que um operador linear limitado mapeia conjuntos limitados em $D(T)$ em conjuntos limitados em Y. Isso motiva o termo "operador limitado".

Ainda na equação (1) podemos nos perguntar qual o menor c tal que vale a mesma (desconsiderando $c = 0$), por divisão

$$\frac{||T(x)||}{||x||} \leqslant c.$$

e isso mostra que c é pelo menos tão grande quanto o supremo da expressão à esquerda tomada sobre $D(T) - 0$. Daí a resposta a nossa questão é que o menor c possível em (1) é esse supremo. Esse número é indicado por $||T||$; portanto

$$\|T\| = \sup_{\substack{x \in D(T) \\ x \neq 0}} \frac{\|Tx\|}{\|x\|}. \qquad (2)$$

$||T||$ é chamada de norma do operador T. Se $D(T) = 0$, definimos $||T|| = 0$.

Observe que (1) com c $= ||T||$ se transforma em

$$||T(x)|| \leqslant ||T||||x||. \qquad (3)$$

essa fórmula será aplicada com bastante frequência.

É claro que devemos justificar o uso do termo "norma"no contexto atual. Isso será feito no lema a seguir.

Lema 1.6.2 *Seja T um operador linear como definido em* 1.6.1, *temos:*

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

$$\|T\| = \sup_{\substack{x\in D(T)\\ ||x||=1}} \|Tx\| \qquad (4)$$

(b) *A norma definida em (2) satisfaz (N1) a (N4) da definição de norma.*

Demonstração:

(a) Escrevemos $||x|| = a$ e $y = (1/a)x$, onde $x \neq 0$. Então $||y|| = ||x||/a = 1$, e como T é linear, obtemos

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

trocando x por y, obtemos (4).

(b) (Nl) é de verificação imediata. De $||T|| = 0$ nós temos $T(x) = 0$ para todo $x \in D(T)$ de modo que $T = 0$. Portanto vale (N2). Além disso, (N3) é obtido de

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

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

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

$x \in D(T)$. ■

Exemplo 1.6.1 Operador diferencial. Seja X o espaço normado de todos polinômios em $J = [0, 1]$ com norma dada $||x|| = max|x(t)|, t \in J$. O operador de diferenciação T é definido em X por

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

onde a linha denota a diferenciação em com respeito a t. Este operador é linear, mas não limitado. De fato, seja $x_n = t^n$, onde $n \in \mathbb{N}$. Então $||x|| = 1$ e

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

de modo que $||T(x_n)|| = n$ e $||T(x_n)||/||x_n|| = n$, Como $n \in \mathbb{N}$ é arbitrário, isso mostra que não existe um número fixo c tal que $||T(x_n)||/||x_n|| \leqslant n$. A partir disso e (1) concluímos que T não é limitado.

Exemplo 1.6.2 Operador Integral. Podemos definir o operador integral $T : C[0,1] \longrightarrow C[0,1]$ por

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

Aqui k é uma função dada, que é chamada de kernel(não é o núcleo) de T e é assumida como contínua no quadrado fechado $G = J \times J$ no $t\tau$-plano, onde $J = [0,1]$. T é um operador linear.

Afirmação: T é limitado.

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

$$|x(t)| \leqq \max_{t \in J} |x(t)| = \|x\| \tag{1.4}$$

consequentemente

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

isso resulta que $||T(x)|| \leqslant k_0||x||$. Tomando $c = k_0$, obtemos que T é limitado.

Um resultado importante para operadores lineares em dimensão finita será provado a seguir.

Teorema 1.6.3 *Se um espaço normalizado X tem dimensão finita, então todo operador linear em X é limitado.*

Demonstração:

Seja $dimX = n$ e $\{e_1, \cdots, e_n\}$ uma base qualquer de X. Nós pegamos qualquer $x = \sum \xi_j e_j$ e consideremos qualquer operador linear T em X. Como T é linear,

$$\|Tx\| = \left\| \sum \xi_j T e_j \right\| \leqq \sum |\xi_j| \, \|Te_j\| \leqq \max_k \|Te_k\| \sum |\xi_j|$$

(somatório de 1 a n). para a última soma, aplicamos o lema das combinações lineares para $\alpha_j = \xi_j$ e $x_j = e_j$. Então obtemos,

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

juntado as duas equações

$$||T(x)|| \leqslant A||x|| \quad \text{onde} \quad A = \frac{1}{c} \max_k \|Te_k\|.$$

Concluímos que T é limitado. ∎

Vamos agora considerar importantes propriedades gerais de operadores lineares.

Os operadores são aplicações, de modo que a definição de continuidade se aplica a eles. É um fato fundamental que para um operador linear, continuidade e limitação tornam-se conceitos equivalentes. Dessa forma podemos enunciar o seguinte teorema,

Teorema 1.6.4 ***(Continuidade e Limitação)*** *Seja $T : D(T) \longrightarrow Y$ um operador linear, onde $D(T) \subset X$ e X e Y são espaços normados. temos:*

(a) *T é contínua se, somente se, T é limitada.*

(b) *Se T é contínua em um único ponto, ela é contínua.*

Demonstração:

(a) Se $T = 0$ a afirmação é trivial. Se $T \neq 0$ então $||T|| \neq 0$. Assumimos que T é limitado e consideramos qualquer $x_0 \in D(T)$. Dado $\varepsilon > 0$ qualquer, como T é linear, para todo $x \in D(T)$ temos que

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

nós obtemos

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

como $x_0 \in D(T)$ foi arbitrário, isso mostra que T é contínuo.

Por outro lado, suponha que T seja contínua em um $x_o \in D(T)$ arbitrário. Então, dado qualquer $\varepsilon > 0$, existe um $\delta > 0$ tal que se $||x - x_0|| \leqslant \delta$ para todo $x \in D(T)$ então $||T(x) - T(x_0)|| \leqslant \varepsilon$.
Agora tomamos qualquer $y \neq 0 \in D(T)$ e definimos

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

Donde $||x - x_0| = \delta$. Como T é linear, temos

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

que implica

$$\frac{\delta}{||y||}||T(y)|| \leqslant \varepsilon. \quad \text{portanto} \quad ||T(y)|| \leqslant \frac{\varepsilon}{\delta}||y||.$$

tomando $c = \frac{\varepsilon}{\delta}$, obtemos $||T(y)|| \leqslant c||y||$. Dessa forma, provamos que T é limitado.

(b) Continuidade de T em um ponto implica limitação de T pela segunda parte da prova de (a), que por sua vez implica continuidade de T por (a). ∎

Corolário 1.6.5 *Seja T um operador linear limitado. Então:*

(a) $x_n \longrightarrow x$ *(onde $x_n, x \in D(T)$) implica* $T(x_n) \longrightarrow T(x)$

(b) *O Núcleo $N(T)$ é fechado.*

Demonstração:

(a) Como T é linear, vale

$$||T(x_n) - T(x)|| = ||T(x_n - x)|| \leqslant ||T||||x_n - x|| \longrightarrow 0.$$

pois $x_n \longrightarrow x$ implica $||x_n - x|| \longrightarrow 0$.

(b) Para todo $x \in \overline{N(T)}$ existe $(x_n) \subset N(T)$ tal que $x_n \longrightarrow x$. Portanto, $T(x_n) \longrightarrow T(x)$ pela parte (a) desse corolário. temos que $T(x) = 0$, pois $T(x_n) = 0$ de modo que $x \in N(T)$. Como $x \in \overline{N(T)}$ era arbitrário, concluímos que $N(T)$ é fechado. ■

Operadores são aplicações, e alguns conceitos relacionados a aplicações foram discutidos, notadamente o domínio, a imagem e o núcleo de um operador. Dois outros conceitos (restrição e extensão) serão adicionados agora.

Vamos começar definindo igualdade de operadores da seguinte forma,

Dois operadores T_1 e T_2 são definidos como iguais, e escrevemos,

$$T_1 = T_2$$

se eles tem o mesmo domínio $D(T_1) = D(T_2)$ e $T_1(x) = T_2(x)$ para todo $x \in D(T_1) = D(T_2)$.

A restrição de um operador $T : D(T) \longrightarrow Y$ a um subconjunto $B \subset D(T)$ é denotado por

$$T|_B$$

e é definido por

$$T|_B : B \longrightarrow Y, \qquad T|_B(x) = T(x) \quad \text{para todo} \quad x \in B.$$

Uma extensão de T para um conjunto $M \supset D(T)$ é um operador

$$\tilde{T} : M \longrightarrow Y, \quad \text{tal que} \quad \tilde{T}|_{D(T)} = T.$$

isto é, $\tilde{T}(x) = T(x)$ para todo $x \in D(T)$.

Se $D(T)$ é um subconjunto próprio de M, então um operador linear T dado, tem muitas extensões. De interesse prático são geralmente aquelas extensões que preservam alguma propriedade básica, por exemplo linearidade (se T for linear) ou limitação (se $D(T)$ estiver em um espaço normado e T for limitado).

O seguinte teorema é muito importante com relação a esse respeito. Trata-se de uma extensão de um operador linear limitado T para o fecho $\overline{D(T)}$ do domínio tal que o operador estendido é novamente limitado e linear, e ainda tem a mesma norma. Isso inclui o caso de uma extensão de um conjunto denso em um espaço normado X para todo X. Também inclui o caso de uma extensão de um espaço normado X ao seu complemento (Teorema 1.2.7).

Teorema 1.6.6 *Seja*

$$T : D(T) \longrightarrow Y$$

um operador linear limitado. onde $D(T)$ está em um espaço normado X e Y é um espaço de Banach. Então T tem uma extensão

$$\tilde{T} : \overline{D(T)} \longrightarrow Y$$

onde T é um operador linear limitado de norma

$$||\tilde{T}|| = ||T||.$$

Demonstração:

Consideramos qualquer $x \in \overline{D(T)}$. Pelo Teorema 0.0.9(a), existe uma sequência (x_n) em $D(T)$ tal que $x_n \longrightarrow x$. Como T é linear e limitado, temos

$$||T(x_n) - T(x_m)|| \leqslant ||T(x_n - x_m)|| \leqslant ||T||||x_n - x_m||.$$

isso mostra que $(T(x_n))$ é de Cauchy, pois (x_n) é convergente. Por hipótese, Y é completo, de modo que $T(x_n)$ converge, digamos

$$T(x_n) \longrightarrow y \in Y.$$

definimos $\tilde{T}$ por

$$\tilde{T}(x) = y$$

Mostramos que esta definição é independente da escolha particular de uma sequência em $D(T)$ convergindo para x. Suponha que $x_n \longrightarrow x$ e $z_n \longrightarrow x$. Então $v_m \longrightarrow x$, onde (v_m) é a sequência

$$(x_1, z_1, x_2, z_2, \cdots).$$

Portanto, $T(v_m)$ converge pelo Corolário 1.6.5(a), e as duas subsequências $T(x_n)$ e $T(z_n)$ de $T(v_m)$ convergem para o mesmo limite. Isso prova que $\tilde{T}$ é único, definido para todo $x \in \overline{D(T)}$.

Claramente, $\tilde{T}$ é linear e $\tilde{T}(x) = T(x)$ para todo $x \in D(T)$, de modo que $\tilde{T}$ é uma extensão de T, agora usamos que

$$||T(x_n)|| \leqslant ||T||||x_n||$$

e seja $n \longrightarrow \infty$. Então $T(x_n) \longrightarrow y = \tilde{T}(x)$. Como $x \longmapsto ||x||$ define uma contínua, temos que

$$||\tilde{T}(x)|| \leqslant ||T||||x||.$$

Portanto $\tilde{T}$ é limitado, e $||\tilde{T}|| \leqslant ||T||$.
É claro que, $||\tilde{T}|| \geqslant ||T||$ porque a norma, sendo definida por um supremo, não pode diminuir em uma extensão. Juntos temos $||\tilde{T}|| = ||T||$. ∎

1.7 Funcionais Lineares

Um funcional é um operador cujo a imagem está na reta real $\mathbb{R}$ ou no plano complexo $\mathbb{C}$. E a Análise Funcional foi inicialmente a Análise de Funcionais, assim como o nome sugere. Funcionais são operadores, de modo que as definições anteriores se aplicam.

Nós precisaremos, em particular, das duas definições a seguir, porque a maioria dos funcionais a serem considerados serão lineares e limitados.

Definição 1.7.1 Um funcional linear f é um operador linear com domínio em um espaço vetorial X e imagem no corpo escalar $\mathbb{K}$ de X; portanto,

$$f : D(f) \longrightarrow \mathbb{K}$$

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

Definição 1.7.2 Um funcional linear limitado f é um operador linear limitado com a imagem em um corpo escalar do espaço normado X no qual o domínio $D(f)$ se encontra. Assim, existe um número real c tal que para todo $x \in D(f)$,

$$|f(x)| \leqslant c||x||. \qquad \textbf{(1)}$$

além disso, a norma de f é

$$\|f\| = \sup_{\substack{x\in D(f)\\ x\neq 0}} \frac{|f(x)|}{\|x\|}. \qquad \textbf{(2a)}$$

ou

$$\|f\| = \sup_{\substack{x\in D(f)\\ ||x||=1}} |f(x)|. \qquad \textbf{(2b)}$$

Analogamente ao que foi feito para operadores limitados, obtemos que

$$|f(x)| \leqslant ||f||||x||. \qquad \textbf{(3)}$$

e um caso especial do Teorema 1.6.4, pode ser enunciado para funcionais lineares.

Teorema 1.7.3 *Um funcional linear f com domínio $D(f)$ em um espaço normado é contínuo se, e somente se, f é limitado.*

Exemplo 1.7.1 A norma $||\cdot||: X \longrightarrow \mathbb{R}$ no espaço normado $(X, ||\cdot||)$ é um funcional em X mas não é linear.

Exemplo 1.7.2 O produto escalar define um funcional $f: \mathbb{R}^3 \longrightarrow \mathbb{R}$ da seguinte forma

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

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

f é linear e limitado. De fato,

$$|f(x)| = |x \cdot a| \leqq \|x\|\|a\|$$

de modo que $||f|| \leqslant ||a||$ segue pela equação (2b) tomando o supremo sobre todos os x de norma um. Por outro lado, tomando $x = a$ e usando (3) obtemos

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

concluímos que a norma de f é $||f|| = ||a||$.

Exemplo 1.7.3 A integral definida é um número se a considerarmos para uma única função, como fazemos no cálculo na maioria das vezes. No entanto, a situação muda completamente se considerarmos a integral para todas as funções em um determinado espaço de funções. Então a integral se torna um funcional nesse espaço. Seja $C[a, b]$ o espaço das funções contínuas. Defina f por

$$f(x) = \int_a^b x(t)dt$$

$x \in C[a, b]$.

f é linear. Vamos provar que f é limitado e tem norma $||f|| = b - a$.
De fato, tome $J = [a, b]$ e lembrando a norma em $C[a, b]$, nós obtemos

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

Tomando o supremo sobre todos os x de norma 1, obtemos $||f|| \leqslant b - a$. Para obter $||f|| \geqslant b - a$, tomamos em particular $x = x_0 = 1$, note que $||x_0|| = 1$ e usamos a equação (3):

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

Exemplo 1.7.4 Outro funcional importante no espaço $C[a,b]$ é obtido se escolhermos $t_0 \in J = [a,b]$ fixo e definirmos

$$f_1(x) = x(t_0)$$

$x \in C[a,b]$.

f_1 definida dessa forma, é linear. f_1 é limitado e sua norma é $||f|| = 1$. De fato,

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

isso implica que $||f|| \leqslant 1$ por (2). Por outro lado, para $x_0 = 1$ temos $||x_0|| = 1$ e obtemos de (3):

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

Exemplo 1.7.5 Podemos obter um funcional linear f no espaço de Banach l^2 escolhendo um fixo $a = (a_j) \in l^2$ e definindo

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

onde $x = (\xi_j) \in l^2$. Esta série converge absolutamente e f é limitada, pela desigualde de Cauchy-Schwarz para somas,

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

É de fundamental importância que o conjunto de todos os funcionais lineares definido em um espaço vetorial X pode ser transformado em um espaço vetorial. Este espaço é denotado por X^* e é chamado de **espaço dual algébrico** de X. Suas operações algébricas do espaço vetorial são definidas de forma natural do seguinte modo. A soma $f_1 + f_2$ de dois funcionais f_1 e f_2 é o funcional cujo valor em cada $x \in X$ é

$$s(x) = (f_1 + f_2)(x) = f_1(x) + f_2(x)$$

o produto αf de um escalar α e um funcional f é o funcional p cujo valor em $x \in X$ é

$$p(x) = (\alpha f)(x) = \alpha f(x)$$

Podemos dar um passo adiante e considerar o dual algébrico $(X)^{**}$ de X^*, cujos elementos são os funcionais lineares definidos em X^*. Nós chamaremos X^{**} de segundo espaço dual algébrico de X.

Podemos obter $g \in X^{**}$, que é um funcional linear definido em X^*, escolhendo um $x \in X$ fixo e definindo

$$g(f) = g_x(f) = f(x) \qquad \textbf{(4)}$$

onde $f \in X^*$ é variável. O subscrito x é um pequeno lembrete de que obtivemos g pelo uso de um certo $x \in X$. O leitor deve observar cuidadosamente que aqui f é a variável enquanto x é fixo. Tendo isso em mente, ele não deve ter dificuldades na compreensão de nossa presente consideração.

g_x conforme definido por (4) é linear. Isso pode ser visto a partir

$$g_x(\alpha f_1 + \beta f_2) = (\alpha f_1 + \beta f_2)(x) = \alpha f_1(x) + \beta f_2(x) = \alpha g_x(f_1) + \beta g_x(f_2)$$

Portanto, g_x é um elemento de X^{**}, pela definição de X^{**}.

A cada $x \in X$ corresponde um $g_x \in X^{**}$. Isso define uma aplicação

$$\begin{aligned} C : & X \longrightarrow X^{**} \\ & x \longmapsto g_x \end{aligned}$$

C é chamada de **aplicação canônica** de X em X^{**}.

C é linear, pois seu domínio é um espaço vetorial e temos

$$\begin{aligned} (C(\alpha x + \beta y))(f) &= g_{\alpha x + \beta y}(f) \\ &= f(\alpha x + \beta y) \\ &= \alpha f(x) + \beta f(y) \\ &= \alpha g_x(f) + \beta g_y(f) \\ &= \alpha (Cx)(f) + \beta (Cy)(f) \end{aligned}$$

C também é chamado de incorporação canônica de X em X^{**}. Para entender e motivar este termo, primeiro explicamos o conceito de "isomorfismo", que é de interesse geral.

Um isomorfismo T de um espaço vetorial X em um espaço vetorial $\tilde{X}$ sobre o mesmo corpo é uma aplicação bijetiva que preserva as duas operações de espaço vetorial; assim, para todos os $x, y \in X$ e escalares $\alpha \in \mathbb{K}$,

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

isto é, $T : X \longrightarrow \tilde{X}$ é um operador linear bijetivo. $\tilde{X}$ é então chamado isomorfo a X, e X e $\tilde{X}$ são chamados espaços vetoriais isomórficos.

Isomorfismos para espaços normados são isomorfismos de espaço vetoriais que também preservam as normas.

Pode-se mostrar que a aplicação canônica C é injetiva. Como C é linear, é um isomorfismo de X no intervalo $Im(C) \subset X^{**}$.

Se C é sobrejetivo (portanto bijetivo), de modo que $Im(C) = X^{**}$, então X é dito ser **algebricamente reflexivo**. Vamos provar na próxima seção que se X é de dimensão finita, então X é algebricamente reflexivo.

1.8 Operadores lineares e funcionais em espaços de dimensão finita.

Espaços vetoriais de dimensão finita são mais simples do que espaços de dimensão infinita e é natural perguntar que simplificação isso implica com respeito aos operadores lineares e funcionais definidos em tal espaço; esta é a questão a ser considerada, e a resposta esclarecerá o papel de matrizes (finitas) em conexão com operadores lineares, bem como a estrutura do dual algébrico X^* de um espaço vetorial dimensão finita X.

Operadores lineares em espaços vetoriais de dimensão finita podem ser representados em termos de matrizes, conforme explicado abaixo. Dessa forma, as matrizes tornam-se as ferramentas mais importantes para o estudo de operadores lineares no caso de dimensão finita.

Sejam X e Y espaços vetoriais de dimensão finita sobre o mesmo corpo e $T : X \longrightarrow Y$ um operador linear. Escolhemos uma base $E = \{e_1, \cdots, e_n\}$ para X e $B = \{b_1, \cdots, b_r\}$ uma base para Y, com os vetores dispostos em uma ordem definida que mantemos fixa. Então todo $xinX$ tem uma representação única

$$x = \xi_1 e_1 + \cdots + \xi_n e_n. \qquad (1)$$

Como T é linear, temos

$$y = T(x) = T\left(\sum_{k=1}^{n} \xi_k e_k\right) = \sum_{k=1}^{n} \xi_k T e_k. \qquad (2)$$

Como a representação (1) é única, temos nosso primeiro resultado:

T é determinado unicamente se as imagens $y_k = T(e_k)$ dos n vetores de base $e_1, \cdots, e_n$ são prescritos.

Como y e $y_k = T(e_k)$ estão em Y, eles têm representações únicas da forma

$$y = \sum_{j=1}^{r} \eta_j b_j \qquad (3a)$$

$$Te_k = \sum_{j=1}^{r} \tau_{jk} b_j \qquad (3b).$$

substituindo em (2), obtemos

$$y = \sum_{j=1}^{r} \eta_j b_j = \sum_{k=1}^{n} \xi_k T e_k = \sum_{k=1}^{n} \xi_k \sum_{j=1}^{r} \tau_{jk} b_j = \sum_{j=1}^{r} \left(\sum_{k=1}^{n} \tau_{jk} \xi_k\right) b_j.$$

Como os $b_j's$ formam um conjunto linearmente independente, os coeficientes de cada b_j à esquerda e à direita devem ser iguais, ou seja,

$$\eta_j = \sum_{k=1}^{n} \tau_{jk} \xi_k, \qquad j = 1, \cdots, r. \qquad (4)$$

Isso produz nosso próximo resultado:

A imagem $y = T(x) = \sum \eta_j b_j$ *de* $x = \sum \xi_k e_k$ *pode ser obtida por* (4).

Os coeficientes em (4) formam uma matriz

$$T_{EB} = (\tau_{jk})$$

com r linhas e n colunas. Se uma base E para X e uma base B para Y são dadas, com os elementos de E e B dispostos em alguma ordem definida (que é arbitrário, mas fixo), então a matriz T_{EB} é determinada exclusivamente pelo operador linear T. Dizemos que a matriz T_{EB} representa o operador T em relação a essas bases.

Ao introduzir os vetores coluna $\tilde{x} = (\xi_k)$ e $\tilde{y} = (\eta_j)$ podemos escrever (4) em notação matricial:

$$\tilde{y} = T_{EB}\tilde{x}. \qquad (4')$$

Da mesma forma, (3b) também pode ser escrito em notação matricial

$$T(e) = T_{EB}^t b$$

onde $T(e)$ é o vetor coluna com componentes $T(e_1), \cdots, T(e_n)$ (que são eles mesmos vetores) e b é o vetor coluna com componentes $b_1, \cdots, b_r$ e temos que usar a transposição T_{EB}^t de T_{EB} porque em (3b) somamos j, que é o primeiro subscrito, enquanto em (4) somamos sobre k, que é o segundo subscrito.

Nossa consideração mostra que um operador linear T determina uma única matriz representando T em relação a uma determinada base para X e uma determinada base para Y, onde os vetores de cada uma das bases são assumidos em uma ordem fixa. Por outro lado, qualquer matriz com r linhas e n colunas determina um operador linear que o representa com relação às bases dadas para X e Y.

Vamos agora aos funcionais lineares em X, onde $dimX = n$ e $\{e_1, \cdots, e_n\}$ é uma base para X, como antes. Esses funcionais constituem o espaço dual algébrico X^* de X, como sabemos da seção anterior. Para cada funcional f e cada $x = \sum \xi_j e_j \in X$ temos

$$f(x) = f\left(\sum_{i=1}^{n} \xi_j e_j\right) = \sum_{j=1}^{n} \xi_j f(e_j) = \sum_{j=1}^{n} \xi_j \alpha_j. \qquad (5a)$$

onde

$$\alpha_j = f(e_j), \quad j = 1, ..., n. \qquad (5b)$$

e f é determinado unicamente por seus valores α_j nos n vetores da base de X.

Por outro lado, toda n-upla de escalares $\alpha_1, \cdots, \alpha_n$ determina um funcional linear em X por (5). Em particular, tomemos as n-uplas

$$\begin{matrix} (1, & 0, & 0, & \cdots & 0, & 0) \\ (0, & 1, & 0, & \cdots & 0, & 0) \\ . & . & . & \cdots & . & . \\ (0, & 0, & 0, & \cdots & 0, & 1) \end{matrix}$$

Por (5) isso dá n funcionais, que denotamos por $f_1, \cdots, f_n$ com valores

$$f_k(e_j) = \delta_{jk} = \begin{cases} 0 & \text{se } j \neq k, \\ 1 & \text{se } j = k \end{cases} \qquad (6)$$

ou seja, f_k tem o valor 1 no k-ésimo vetor de base e 0 nos $n-1$ outros vetores de base. δ_{jk} é chamado de *delta de Kronecker.* $\{f_1, \cdots, f_n\}$ é chamada de **base dual** da base $\{e_1, \cdots, e_n\}$ para X. Isso é justificado pelo seguinte teorema.

Teorema 1.8.1 ***(Dimensão do X*).*** *Seja X um espaço vetorial n-dimensional e $E = \{e_1, \cdots, e_n\}$ uma base para X. Então $F = \{f_1, \cdots, f_n\}$ dado por (6) é uma base para o dual algébrico X^* de X, e $dimX^* = dimX = n$.*

Demonstração:

F é um conjunto linearmente independente, pois

$$\sum_{k=1}^{n} \beta_k f_k(x) = 0 \qquad (x \in X) \qquad (7)$$

para $x = e_j$ temos

$$\sum_{k=1}^{n} \beta_k f_k(e_j) = \sum_{k=1}^{n} \beta_k \delta_{jk} = \beta_j = 0,$$

para que todos os $\beta_k^{'}s$ em (7) sejam zero. Mostramos que todo $f \in X^*$ pode ser representado como uma combinação linear dos elementos de F de uma maneira única. Escrevemos $f(e_j) = \alpha_j$ como em $(5b)$. Por $(5a)$, temos

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

para todo $x \in X$. Por outro lado, por (6) obtemos

$$f_j(x) = f_j(\xi_1 e_1 + \cdots + \xi_n e_n) = \xi_j.$$

juntas,

$$f(x) = \sum_{j=1}^{n} \alpha_j f_j(x).$$

Daí a representação única do funcional linear arbitrário f em X em termos dos funcionais $f_1, \cdots, f_n$ é

$$f = \alpha_1 f_1 + \cdots + \alpha_n f_n.$$

■

Para nos prepararmos para uma aplicação interessante deste teorema, primeiro provaremos o seguinte lema.

Lema 1.8.2 *Seja X um espaço vetorial de dimensão finita. Se $x_0 \in X$ tem a propriedade de que $f(x_0) = 0$ para todo $f \in X^*$, então $x_0 = 0$.*

Demonstração:

Seja $\{e_1, \cdots, e_n\}$ uma base para X e $x_0 = \sum \xi_{0j} e_j$. Então por (5) temos

$$f(x_0) = \sum_{i=1}^{n} \xi_{0j} \alpha_j$$

Por suposição, isso é zero para todo $f \in X^*$, ou seja, para cada escolha de $\alpha_1, \cdots, \alpha_n$. Portanto, todos os ξ_{0j} devem ser zero. ■

Usando este lema, podemos agora obter o seguinte teorema,

Teorema 1.8.3 ***(Algebricamente reflexivo).*** *Todo espaço vetorial de dimensão finita é algebricamente reflexivo.*

Demonstração:

A aplicação canônica $C : X \longrightarrow X^{**}$ é linear, como provado na seção anterior. $Cx_0 = 0$ significa que para todo $f \in X^*$ nós temos

$$(Cx_0)(f) = g_{x_0}(f) = f(x_0) = 0,$$

pela definição de C. Isso implica $x_0 = 0$ pelo Lema 1.8.2. Consequentemente do Teorema 1.5.3 segue que a aplicação C tem uma inversa $C^{-1} : Im(C) \longrightarrow X$. Também temos $dimIm(C) = dimX$ pelo mesmo teorema. Agora pelo Teorema 1.8.1,

$$dimX^{**} = dimX^* = dimX.$$

Juntos, $dimIm(C) = dimX^{**}$. Portanto, $Im(C) = X^{**}$ porque $Im(C)$ é um espaço vetorial e um subespaço próprio de X^{**}. Como $Im(C)$ tem dimensão menor que $dimX^{**}$, pelo Teorema 1.1.4. Por definição isso prova reflexividade algébrica. ■

1.9 Espaço Dual

Sejam X e Y dois espaços vetoriais normados (ambos reais ou complexos). Considere o conjunto $B(X, Y)$ de todos os operadores lineares limitados de X em Y, ou seja, cada operador é definido em X e sua imagem está em Y. Queremos mostrar que $B(X,Y)$ pode ser transformado em um espaço normalizado.

A questão toda é bastante simples. Em primeiro lugar, $B(X,Y)$ torna-se um espaço vetorial se definirmos a soma T_1+T_2 de dois operadores $T_1, T_2 \in B(X,Y)$ de forma natural por

$$(T_1 + T_2)(x) = T_1(x) + T_2(x).$$

e o produto αT de um operador $T \in B(X,Y)$ por um escalar α da seguinte forma

$$(\alpha T)(x) = \alpha T(x).$$

Lembrando diretamente do lema 1.6.2, temos imediatamente o seguinte resultado:

Teorema 1.9.1 *O espaço vetorial $B(X,Y)$ de todos operadores lineares limitados de um espaço normado X para um espaço normado Y é ele próprio um espaço normado com norma definida por*

$$\|T\| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{\|Tx\|}{||x||} = \sup_{\substack{x \in X \\ ||x||=1}} \|Tx\| \qquad (1)$$

Em que caso $B(X,Y)$ será um espaço de Banach? Esta é um questão central, que é respondida no seguinte teorema. É notável que a condição do teorema não envolve X; ou seja, X pode ser ou não completo.

Teorema 1.9.2 *Se Y é um espaço de Banach, então $B(X,Y)$ é um espaço de Banach.*

Demonstração:

Consideramos uma sequência de Cauchy arbitrária (T_n) em $B(X,Y)$ e mostraremos que (T_n) converge para um operador $T \in B(X,Y)$.

Como (T_n) é de Cauchy, para cada $\varepsilon > 0$ existe um N tal que, $m, n > N$ então

$$||T_n - T_m|| < \varepsilon$$

donde para todo $x \in X$ e $n, m > N$ temos

$$\|T_n(x) - T_m(x)\| = \|(T_n - T_m)\, x\| \leqq \|T_n - T_m\|\, \|x\| < \varepsilon \|x\|. \qquad (2)$$

Agora, para qualquer x fixo e dado $\tilde{\varepsilon}$, podemos escolher $\varepsilon = \varepsilon_x$ de modo que $\varepsilon_x ||x|| < \varepsilon$. Então de (2) temos $||T_n(x) - T_m(x)|| < \varepsilon$ e vemos que $(T_n(x))$ é de Cauchy em Y. Como Y é completo, $(T_n(x))$ converge, digamos, $T_n(x) \longrightarrow y$. Claramente, o limite $y \in Y$ depende da escolha de $x \in X$. Isso define um operador $T : X \longrightarrow Y$, onde $y = T(x)$. O operador T é linear, pois

$$\lim T_n(\alpha x + \beta z) = \lim\, (\alpha T_n x + \beta T_n z) = \alpha \lim T_n x + \beta \lim T_n z$$

Provamos que T é limitado e $T_n \longrightarrow T$, ou seja, $||T_n - T|| \longrightarrow 0$.

Como (2) vale para todo $m > N$ e $T_m(x) \longrightarrow T(x)$, podemos fazer $m \longrightarrow \infty$. Usando a continuidade da norma, obtemos de (2) para todo $n > N$ e todo $x \in X$

$$\|T_n x - Tx\| = \left\| T_n x - \lim_{m\to\infty} T_m x \right\| = \lim_{m\to\infty} \|T_n x - T_m x\| \leqq \varepsilon \|x\|. \qquad (3)$$

Isso mostra que $(T_n - T)$ com $n > N$ é um operador linear limitado. Como T_n é limitado, $T = T_n - (T_n - T)$ é limitado, ou seja, $T \in B(X, Y)$. Além disso, se em (3) tomarmos o supremo sobre todos os x de norma 1, obtemos

$$||T_n - T|| < \varepsilon$$

$n > N$. Daí, $||T_n - T|| \longrightarrow 0$. ∎

Este teorema tem uma consequência importante no que diz respeito ao espaço dual X' de X, que é definido a seguir.

Definição 1.9.3 Seja X um espaço normado. Então o conjunto de todos os funcionais lineares limitados em X constitui um espaço com norma definido por

$$\|f\| = \sup_{\substack{x\in X \\ x\neq 0}} \frac{|f(x)|}{||x||} = \sup_{\substack{x\in X \\ ||x||=1}} |f(x)| \qquad (4)$$

que é chamado de espaço dual de X e é denotado por X'.

Como um funcional linear em X leva elementos de X em $\mathbb{R}$ ou $\mathbb{C}$ e como $\mathbb{R}$ ou $\mathbb{C}$, tomado com a métrica usual, é completo, vemos que X' e $B(X, Y)$ é completo ($Y = \mathbb{R}$ ou $\mathbb{C}$). Portanto O Teorema 1.9.2 é aplicável e implica diretamente o seguinte corolário.

Corolário 1.9.4 *O espaço dual X' de um espaço normado X é um espaço de Banach.*

fecharemos a seção refazendo a definição de isomorfismo, agora para espaços normados.

Definição 1.9.5 Um isomorfismo de um espaço normado X em um espaço normado $\tilde{X}$ é um operador linear bijetivo $T : X \longrightarrow \tilde{X}$ que preserva norma, isto é, para todo $x \in X$,

$$||T(x)|| = ||x||.$$

X é então chamado isomorfo à $\tilde{X}$, e X e $\tilde{X}$ são chamados espaços normados isomórficos. De um ponto abstrato de vista, X e $\tilde{X}$ são idênticos.

Capítulo 2

Espaços de Hilbert

Em um espaço normado podemos adicionar vetores e multiplicar vetores por escalares, como na álgebra vetorial elementar. Além disso, a norma sobre tal espaço generaliza o conceito elementar de comprimento de um vetor. No entanto, o que ainda falta em um espaço normado em geral, é que gostaríamos de ter, se possível, um análogo do produto escalar, usado regularmente na geometria analítica por exemplo, onde

$$a \cdot b = \alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3$$

e fórmulas resultantes,

$$|a| = \sqrt{a \cdot a}$$

e a condição de ortogonalidade (perpendicularidade)

$$a \cdot b = 0$$

que são ferramentas importantes em muitas aplicações. Daí a pergunta surge se o produto escalar e a ortogonalidade podem ser generalizados para espaços vetoriais arbitrários. Na verdade, isso pode ser feito e leva a espaços de produtos internos e espaços de produtos internos completos, chamados de espaços de Hilbert.

2.1 Espaços Com Produto Interno (Espaços de Hilbert)

Os espaços a serem considerados neste capítulo são definidos a seguir.

Definição 2.1.1 Um espaço com produto interno (ou espaço pré-Hilbert) é um espaço vetorial X com um produto interno definido em X. Um espaço de Hilbert é um espaço com produto interno completo (completo na métrica definida pelo produto interno). Aqui, um produto interno em X é uma aplicação de $X \times X$ no corpo escalar $\mathbb{K}$ de X; isto é,para cada par de vetores x e y existe associado um escalar que é escrito $\langle x, y\rangle$ e é chamado de produto interno de x e y, tal que para todos os vetores x, y, z e escalares α temos:

(IP1) $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$

(IP2) $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$

(IP3) $\langle x, y\rangle = \overline{\langle y, x\rangle}$

(IP4) $\langle x, x\rangle \geqq 0, \quad \langle x, x\rangle = 0 \quad \Longleftrightarrow \quad x = 0.$

Um produto interno em X define uma norma em X dada por

$$||x|| = \sqrt{\langle x, x\rangle} \qquad (1)$$

e uma métrica em X dada por

$$d(x, y) = ||x - y|| = \sqrt{\langle x - y, x - y\rangle}. \qquad (2)$$

Portanto, espaços de produto interno são espaços normados e espaços de Hilbert são espaços de Banach.

Em (IP3), a barra denota conjugação complexa. Assim, se X é um espaço vetorial real, temos simplesmente

$$\langle x, y\rangle = \langle y, x\rangle.$$

De (IP1) a (IP3) obtemos as seguintes fórmulas

$$\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle,$$

$$\langle x, \alpha y\rangle = \bar{\alpha}\langle x, y\rangle,$$

$$\langle x, \alpha y + \beta z\rangle = \bar{\alpha}\langle x, y\rangle + \bar{\beta}\langle x, z\rangle.$$

Proposição 2.1.2 ***(Identidade do Paralelogramo)*** *Seja X um espaço vetorial com produto interno. Então*

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right).$$

Demonstração:

$$\begin{aligned}\|x+y\|^2 + \|x-y\|^2 &= (\langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle) + (\langle x, x\rangle - \langle x, y\rangle - \langle y, x\rangle + \langle y, y\rangle) \\ &= 2\left(\|x\|^2 + \|y\|^2\right).\end{aligned}$$

■

Segue diretamente o seguinte resultado, derivado da Identidade do Paralelogramo

Corolário 2.1.3 *Seja X um espaço vetorial normado. Então a norma de X deriva de um produto interno se, e somente se, ela satisfaz a identidade do paralelogramo.*

Observação: Concluímos que se uma norma não satisfaz a identidade do paralelogramo, ela não pode ser obtida a partir de um produto interno. Veremos um exemplo futuramente.

A seguir, vamos introduzir a definição de dois vetores serem ortogonais em um espaço com produto interno

Definição 2.1.4 Um elemento x de um espaço com produto interno X é dito ortogonal a um elemento $y \in X$ se

$$\langle x, y\rangle = 0$$

Também dizemos que x e y são ortogonais e escrevemos $x \perp y$.

Exemplo 2.1.1 O espaço $\mathbb{R}^n$ é um espaço de Hilbert com produto interno definido por

$$\langle x, y\rangle = \xi_1\eta_1 + \cdots + \xi_n\eta_n$$

onde $x = (\xi_j)$ e $y = (\eta_j)$, $j = 1, ..., n$.
A norma desse espaço é proveniente do produto interno,

$$\|x\| = \langle x, x\rangle^{1/2} = \left(\xi_1^2 + \cdots + \xi_n\right)^{1/2}.$$

Exemplo 2.1.2 O espaço $\mathbb{C}^n$ é um espaço de Hilbert com produto interno definido por

$$\langle x, y\rangle = \xi_1\overline{\eta_1} + \cdots + \xi_n\overline{\eta_n}$$

onde a barra indica a conjugação complexa. Da equação acima, obtemos a norma definida da seguinte forma

$$\|x\| = \left(\xi_1\bar{\xi}_1 + \cdots + \xi_n\bar{\xi}_n\right)^{1/2} = \left(|\xi_1|^2 + \cdots + |\xi_n|^2\right)^{1/2}.$$

Exemplo 2.1.3 O espaço l^2 é um espaço de Hilbert com produto interno definido por

$$\langle x, y\rangle = \sum_{j=1}^{\infty} \xi_j\bar{\eta}_j.$$

A norma é definida por

$$\|x\| = \langle x, x\rangle^{1/2} = \left(\sum_{j=1}^{\infty} |\xi_j|^2\right)^{1/2}$$

Exemplo 2.1.4 *O espaço l^p com $p \neq 2$ não é um espaço com produto interno, portanto, não é um espaço de Hilbert.*

Nossa afirmação significa que a norma de l^p com $p \neq 2$ não pode ser obtida a partir de um produto interno. Provamos isso mostrando que a norma não satisfaz a identidade do paralelogramo. Na verdade, vamos tomar $x = (1, 1, 0, 0, \cdots) \in l^p$ e $y = (1, -1, 0, 0, \cdots) \in l^p$ e calcular

$$||x|| = ||y|| = 2^{1/p}, \qquad ||x + y|| = ||x - y|| = 2.$$

Daí,

$$||x + y||^2 + ||x - y||^2 = 4 + 4 = 8$$

e

$$2(||x||^2 + ||y||^2) = 2(2^{1/p} + 2^{1/p}) = 2(2 \cdot 2^{1/p}).$$

Onde

$$\|x + y\|^2 + \|x - y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$$

somente no caso em que $p = 2$. Logo a identidade do paralelogramo não é satisfeita para $p \neq 2$.

Exemplo 2.1.5 *O espaço $C[a, b]$ não é um espaço com produto interno,portanto, não é um espaço Hilbert.*

Mostraremos que a norma

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

não pode ser obtida a partir de um produto interno, uma vez que esta norma não satisfaz a identidade do paralelogramo. De fato, se tomarmos $x(t) = 1$ e $y(t) = (t - a)/(b - a)$, temos $||x|| = 1, ||y|| = 1$ e

$$x(t) + y(t) = 1 + \frac{t - a}{b - a}$$
$$x(t) - y(t) = 1 - \frac{t - a}{b - a}.$$

Como $||x + y|| = 2, ||x - y|| = 1$ e

$$||x + y||^2 + ||x - y||^2 = 5 \quad \text{mas} \quad 2(||x||^2 + ||y||^2) = 4.$$

Isso completa a prova.

Por fim, iremos enunciar um resultado importante para espaço com produto interno, chamado de identidade de polarização

Proposição 2.1.5 *Seja X um espaço vetorial com produto interno. Então para todo $x, y \in Xs$ vale*

(a) $\langle x, y\rangle = \frac{1}{4}\left(\|x + y\|^2 - \|x - y\|^2\right)$. *(caso real)*

(b) $\langle x, y\rangle = \frac{1}{4}\left[\|x + y\|^2 - \|x - y\|^2 + i\left(\|x + iy\|^2 - \|x - iy\|^2\right)\right]$. *(caso complexo)*

Demonstração:
(a) Considere as seguintes igualdades

$$\begin{cases} \|x+y\|^2 = \|x\|^2 + \langle x,y\rangle + \langle y,x\rangle + \|y\|^2 \\ \|x-y\|^2 = \|x\|^2 - \langle x,y\rangle - \langle y,x\rangle + \|y\|^2 \end{cases}$$

Subtraindo a segunda da primeira, temos

$$\|x+y\|^2 - \|x-y\|^2 = \|x\|^2 + \langle x,y\rangle + \langle y,x\rangle + \|y\|^2 - (\|x\|^2 - \langle x,y\rangle - \langle y,x\rangle + \|y\|^2)$$

$$\|x+y\|^2 - \|x-y\|^2 = \|x\|^2 + \langle x,y\rangle + \langle x,y\rangle + \|y\|^2 - \|x\|^2 + \langle x,y\rangle + \langle x,y\rangle - \|y\|^2$$

$$\|x+y\|^2 - \|x-y\|^2 = \langle x,y\rangle + \langle x,y\rangle + \langle x,y\rangle + \langle x,y\rangle$$

$$\|x+y\|^2 - \|x-y\|^2 = 4\langle x,y\rangle$$

$$\frac{1}{4}(\|x+y\|^2 - \|x-y\|^2) = \langle x,y\rangle$$

(b) Considere as igualdades

$$\begin{cases} \|x+iy\|^2 = \|x\|^2 + \langle x,iy\rangle + \langle iy,x\rangle + \|y\|^2 = \|x\|^2 - i\langle x,y\rangle + i\langle y,x\rangle + \|y\|^2 \\ \|x-iy\|^2 = \|x\|^2 + \langle x,-iy\rangle + \langle -iy,x\rangle + \|y\|^2 = \|x\|^2 + i\langle x,y\rangle - i\langle y,x\rangle + \|y\|^2 \end{cases},$$

Subtraindo a segunda da primeira obtemos

$$\|x+iy\|^2 - \|x-iy\|^2 = -2i\langle x,y\rangle + 2i\langle y,x\rangle. \qquad (1)$$

Da parte (a) da demonstração temos que:

$$\|x+y\|^2 - \|x-y\|^2 = 2\langle x,y\rangle + 2\langle y,x\rangle. \qquad (2)$$

Multiplicando (1) por i e somando com (2):

$$\begin{aligned} \|x+y\|^2 - \|x-y\|^2 + i\left(\|x+iy\|^2 - \|x-iy\|^2\right) &= 2\langle x,y\rangle + 2\langle y,x\rangle + 2\langle x,y\rangle - 2\langle y,x\rangle \\ &= 4\langle x,y\rangle, \end{aligned}$$

completando a demonstração. ■

2.2 Propriedades Dos Espaços com Produto Interno

Em primeiro lugar, devemos verificar que (1) na seção anterior define uma norma: (Nl) e (N2) da definição de norma (seção 1.2) seguem de (IP4). Além disso, (N3) é obtido pelo uso de (IP2) e (IP3); na verdade,

$$\|\alpha x\|^2 = \langle \alpha x, \alpha x\rangle = \alpha\bar{\alpha}\langle x,x\rangle = |\alpha|^2\|x\|^2$$

Finalmente (N4) é concluído de:

Proposição 2.2.1 *Seja X um espaço com Produto Interno. Para todo $x, y \in X$ vale:*

(a) $|\langle x, y\rangle| \leqslant ||x||||y||$ ***(Desigualdade de Cauchy-Schwarz)***

onde a igualdade vale se, e somente se, $\{x, y\}$ é um conjunto linearmente dependente.

(b) $||x+y|| \leqslant ||x|| + ||y||$. ***(Desigualdade Triangular)***

onde a igualdade vale se, e somente se, $y = 0$ ou $x = cy$ $\quad (c \geqslant 0)$.

Demonstração:

(a) Para todo $x, y \in X$ e $t \in \mathbb{R}$ vale

$$\langle x+ty, x+ty\rangle \geqslant 0$$

mas

$$\langle x+ty, x+ty\rangle = \|x\|^2 + 2\langle x, y\rangle t + \|y\|^2 t^2$$

logo o discriminante deste polinômio do segundo grau não pode ser positivo:

$$4\langle x, y\rangle^2 - 4\|x\|^2\|y\|^2 \leqslant 0$$

que implica

$$|\langle x, y\rangle| \leqslant ||x||||y||.$$

se $y = 0$ a igualdade se torna trivial. Se $0 = \langle x+ty, x+ty\rangle = ||x+ty||$ então $0 = x+ty$ que implica $x = -ty$ provando a dependência linear.

(b) Temos que

$$\|x+y\|^2 = \langle x+y, x+y\rangle = \|x\|^2 + \langle x, y\rangle + \langle y, x\rangle + \|y\|^2.$$

Pela Desigualdade de Cauchy-schwarz,

$$|\langle x, y\rangle| = |\langle y, x\rangle| \leqq \|x\|\|y\|$$

Dessa forma,

$$\begin{aligned}\|x+y\|^2 &\leqq \|x\|^2 + 2|\langle x, y\rangle| + \|y\|^2 \\ &\leqq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2\end{aligned}$$

Provando o resultado que queríamos. A igualdade resultado do mesmo argumento da letra (a). ∎

Outra propriedade frequentemente utilizada é a continuidade do produto interno:

Lema 2.2.2 *Seja X um espaço com produto interno. Se $x_n \longrightarrow x$ e $y_n \longrightarrow y$, então $\langle x_n, y_n\rangle \longrightarrow \langle x, y\rangle$*

Demonstração:

Subtraindo e somando um termo, usando a desigualdade triangular para números e, finalmente, a desigualdade de Cauchy-Schwarz, obtemos

$$\begin{aligned} |\langle x_n, y_n\rangle - \langle x, y\rangle| &= |\langle x_n, y_n\rangle - \langle x_n, y\rangle + \langle x_n, y\rangle - \langle x, y\rangle| \\ &\leqq |\langle x_n, y_n - y\rangle| + |\langle x_n - x, y\rangle| \\ &\leqq \|x_n\|\,\|y_n - y\| + \|x_n - x\|\,\|y\| \quad \longrightarrow \quad 0 \end{aligned} \tag{2.1}$$

pois $(x_n - x) \longrightarrow 0$ e $(y_n - y) \longrightarrow 0$ quando $n \longrightarrow \infty$. ■

Como primeira aplicação deste lema, vamos provar que todo espaço com produto interno pode ser completado. O completamento é um espaço de Hilbert e é único, exceto para isomorfismos. Aqui a definição de um isomorfismo é a seguinte:

Definição 2.2.3 Um isomorfismo T de um espaço com produto interno X em um espaço com produto interno $\tilde{X}$ sobre o mesmo corpo é um operador linear bijetivo $T : X \longrightarrow \tilde{X}$ que preserva o produto interno, ou seja, para todo $x, y \in X$,

$$\langle Tx, Ty\rangle = \langle x, y\rangle$$

$\tilde{X}$ é então chamado isomorfo à X, e X e $\tilde{X}$ são chamados espaços com produto interno isomórficos. Observe que a bijetividade e a linearidade garante que T é um isomorfismo do espaço vetorial X em $\tilde{X}$, de modo que T preserva toda a estrutura do espaço com produto interno. T também é isometria de X em $\tilde{X}$ porque as distâncias em X e $\tilde{X}$ são determinadas por as normas definidas pelos produtos internos em X e $\tilde{X}$.

O teorema sobre o completamento de um espaço com produto interno pode agora ser declarado, como segue.

Teorema 2.2.4 ***(Completamento).*** *Para qualquer espaço com produto interno X existe um espaço de Hilbert H e um isomorfismo A de X em um subespaço denso $W \subset H$. O espaço H é único exceto para isomorfismos.*

Demonstração:

Pelo teorema 1.2.7 existe um espaço de Banach H e uma isometria A de X em um subespaço W de H que é denso em H. Para razões de continuidade, sob tal isometria, somas e múltiplos escalares de elementos em X e W correspondem entre si, de modo que A é mesmo um isomorfismo de X em W, ambos considerados espaços normados. O Lema 2.2.2 mostra que podemos definir um produto interno em H por

$$\langle \hat{x}, \hat{y}\rangle = \lim_{n\to\infty} \langle x_n, y_n\rangle$$

as notações sendo como no Teorema 1.2.7 , isto é, (x_n) e (y_n) são representantes de $\hat{x} \in H$ e $\hat{y} \in H$, respectivamente. Tomando a identidade de polarização em conta, vemos que A é um isomorfismo de X em W, ambos considerados como espaços de produto interno.

O Teorema 1.2.7 também garante que H é único, exceto para isometrias, ou seja, dois completamentos H e $\tilde{H}$ de X estão relacionadas por uma isometria $T : H \longrightarrow \tilde{H}$. Raciocinando como no caso de A, concluímos que T é um isomorfismo do espaço de Hilbert H sobre o Espaço de Hilbert $\tilde{\tilde{H}}$. ∎

Um **subespaço** Y de um espaço com produto interno X é definido como um subespaço vetorial de X tomado com o produto interno em X restrito a $Y \times Y$.
Da mesma forma, um **subespaço** Y de um espaço de Hilbert H é definido como um subespaço de H, considerado como um espaço com produto interno. Observe que Y não precisa ser um espaço de Hilbert porque Y pode não ser completo. Na verdade, dos Teoremas 1.2.4 e 1.3.2 temos imediatamente as afirmações (a) e (b) no seguinte teorema.

Teorema 2.2.5 *Seja Y um subespaço de um espaço de Hilbert H. Então:*

(a) Y é completo se, e somente se, Y é fechado em H.

(b) Se Y é de dimensão finita, então Y é completo.

2.3 Complementos ortogonais e somas diretas

Definição 2.3.1 O segmento que une dois elementos dados x e y de um espaço vetorial X é definido como o conjunto de todos os $z \in X$ da forma

$$z = \alpha x + (1 - \alpha)y \quad (\alpha \in \mathbb{R}, 0 \leqq \alpha \leqq 1).$$

Um subconjunto M de X é convexo se para todo $x, y \in M$ o segmento unindo x e y está contido em M.

Teorema 2.3.2 *Seja X um espaço com produto interno e $M \neq \emptyset$ um subconjunto convexo que é completo (na métrica induzida pelo produto interno). Então para cada dado $x \in X$ existe um único $y \in M$ de tal modo que*

$$\delta = \inf_{\tilde{y} \in M} \|x - \tilde{y}\| = \|x - y\|. \qquad (1)$$

Demonstração:
(a) Existência. Pela definição de ínfimo existe uma sequência (y_n) em M tal que

$$\delta_n \longrightarrow \delta \qquad \text{onde} \qquad \delta_n = ||x - y_n||. \qquad (2)$$

vamos mostrar que (y_n) é de Cauchy. Escrevendo $y_n - x = v_n$, temos $||v_n|| = \delta_n$ e

$$\|v_n + v_m\| = \|y_n + y_m - 2x\| = 2\left\|\frac{1}{2}(y_n + y_m) - x\right\| \geqq 2\delta$$

porque M é convexo, de modo que $\frac{1}{2}(y_n + y_m) \in M$. Além disso, temos $y_n - y_m = v_n - v_m$. Portanto, pela igualdade do paralelogramo,

$$\begin{aligned}\|y_n - y_m\|^2 = \|v_n - v_m\|^2 &= -\|v_n + v_m\|^2 + 2\left(\|v_n\|^2 + \|v_m\|^2\right) \\ &\leqq -(2\delta)^2 + 2\left(\delta_n^2 + {\delta_m}^2\right),\end{aligned}$$

e (2) implica que (y_n) é de Cauchy. Como M é completo, (y_n) converge, digamos, $y_n \longrightarrow y \in M$. Como $y \in M$, temos $||x - y|| \geqslant \delta$. Também, por (2) temos

$$\|x - y\| \leqq \|x - y_n\| + \|y_n - y\| = \delta_n + \|y_n - y\| \quad \longrightarrow \quad \delta$$

Isto mostra que $||x - y|| = \delta$

(b) Unicidade. Assumimos que $y \in M$ e $y_0 \in M$ ambos satisfazem

$$||x - y|| = \delta \quad \text{e} \quad ||x - y_0|| = \delta$$

e mostraremos que $y_0 = y$. Pela igualdade do paralelogramo,

$$\begin{aligned}\|y - y_0\|^2 &= \|(y - x) - (y_0 - x)\|^2 \\ &= 2\|y - x\|^2 + 2\,\|y_0 - x\|^2 - \|(y - x) + (y_0 - x)\|^2 \\ &= 2\delta^2 + 2\delta^2 - 2^2\left\|\frac{1}{2}(y + y_0) - x\right\|^2\end{aligned}$$

onde a direita, $\frac{1}{2}(y + y_0) \in M$, de modo que

$$||\frac{1}{2}(y + y_0) - x|| \geqslant \delta.$$

Isso implica que o lado direito é menor ou igual a $2\delta^2 + 2\delta^2 - 4\delta^2 = 0$. Portanto, temos a desigualdade $||y - y_0|| \leqslant 0$. Claramente, $||y - y_0|| \geqslant 0$, de modo que devemos ter igualdade, e $y_o = y$. ∎

Segue diretamente desse teorema o seguinte lema sobre vetores perpendiculares.

Lema 2.3.3 *No Teorema 2.3.2, seja M um subespaço completo Y e $x \in X$ fixo. Então $z = x - y$ é ortogonal a Y.*

Demonstração:

Se $z \perp Y$ fosse falso, haveria um $y_1 \in Y$ tal que

$$\langle z, y_1 \rangle = \beta \neq 0. \qquad (3)$$

Claramente, $y_1 \neq 0$, caso contrário $\langle z, y_1 \rangle = 0$. Além disso, para qualquer escalar α,

$$\begin{aligned}\|z - \alpha y_1\|^2 &= \langle z - \alpha y_1, z - \alpha y_1 \rangle \\ &= \langle z, z \rangle - \bar{\alpha}\,\langle z, y_1 \rangle - \alpha\left[\langle y_1, z \rangle - \bar{\alpha}\,\langle y_1, y_1 \rangle\right] \\ &= \langle z, z \rangle - \bar{\alpha}\beta - \alpha\left[\bar{\beta} - \bar{\alpha}\,\langle \cdot y_1, y_1 \rangle\right]\end{aligned}$$

A expressão entre colchetes $[\cdots]$ é zero se escolhermos

$$\bar{\alpha} = \frac{\overline{\boldsymbol{\beta}}}{\langle y_1, y_1 \rangle}$$

De (1) temos $||z|| = ||x - y|| = \delta$, de modo que nossa equação agora produz

$$\|z - \alpha y_1\|^2 = \|z\|^2 - \frac{|\beta|^2}{\langle y_1, y_1 \rangle} < \delta^2.$$

Mas isso é impossível porque temos

$$z - \alpha y_1 = x - y_2 \quad \text{onde} \quad y_2 = y + \alpha y_1 \in Y,$$

de modo que $||z - \alpha y_1|| \geqslant \delta$ pela definição de δ. Portanto (3) não pode ser válido, e o lema está provado. ■

Nosso objetivo é uma representação de um espaço de Hilbert como uma soma direta que é particularmente simples e adequada porque utiliza ortogonalidade. Para entender a situação e o problema, vamos primeiro introduzir o conceito de soma direta. Esse conceito faz sentido para qualquer espaço vetorial e é definido como a seguir.

Definição 2.3.4 (Soma direta). Diz-se que um espaço vetorial X é a soma direta de dois subespaços Y e Z de X, escritos

$$X = Y \oplus Z$$

se cada $x \in X$ tem uma representação única

$$x = y + z$$

onde $y \in Y$ e $z \in Z$. Então Z é chamado de complemento algébrico de Y em X e vice-versa, e Y, Z é chamado de par complementar de subespaços em X.

Da mesma forma, no caso geral de um espaço de Hilbert H, o principal interesse diz respeito a representações de H como uma soma direta de um subespaço Y fechado e seu **complemento ortogonal**

$$Y^{\perp} = \{z \in H | z \perp Y\},$$

que é o conjunto de todos os vetores ortogonais a Y. Isso nos dá nosso principal resultado nesta seção, que às vezes é chamado de teorema da projeção, por razões a serem explicadas após a prova.

Teorema 2.3.5 ***(Soma direta).*** *Seja Y qualquer subespaço fechado de um Espaço de Hilbert H. Então*

$$H = Y \oplus Z \quad \textit{onde} \quad Z = Y^{\perp}. \qquad (4)$$

Demonstração:

Como H é completo e Y é fechado, Y é completo pelo Teorema 0.0.10. Como Y é convexo, O Teorema 2.3.2 e o Lema 2.3.3 implica que para todo $x \in H$ existe um $y \in Y$ tal que

$$x = y + z \quad \text{onde} \quad z \in Z = Y^{\perp}. \tag{5}$$

Para provar a unicidade, assumimos que

$$x = y + z = y_1 + z_1$$

onde $y, y_1 \in Y$ e $z, z_1 \in Z$. Então $y - y_1 = z_1 - z$. Desde $y - y_1 \in Y$ enquanto $z_1 - z \in Z = Y^{\perp}$, vemos que $y - y_1 \in (Y \cap Y^{\perp}) = 0$. Isso implica $y = y_1$. Daí também concluímos que $z = z_1$. ■

O vetor y em (5) é chamado a projeção ortogonal de x em Y (ou, brevemente, a projeção de x em Y). Este termo é motivado por conceitos elementares da geometria. [Por exemplo, podemos tomar $H = \mathbb{R}^2$ e projetar qualquer ponto $x = (\xi_1, \xi_2)$ no eixo ξ_1 que então desempenha o papel de Y; a projeção é $y = (\xi_1, 0)$.]

A equação (5) define a aplicação

$$\begin{aligned} P : & H \longrightarrow Y \\ & x \longmapsto y = P(x) \end{aligned}$$

P é chamada de **Projeção ortogonal** (ou operador projeção) de H em Y. P é um operador linear limitado.

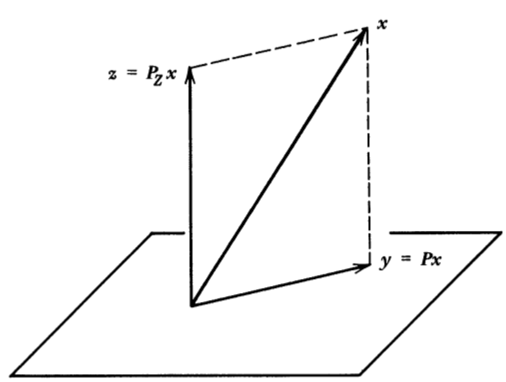

Figura 2.1: Notação em conexão com o Teorema 2.3.5.

P é uma aplicação que mapeia

$$H \text{ em } Y,$$

$$Y \text{ sobre si mesmo,}$$

$$Z = Y^{\perp} \text{ em } \{0\},$$

e é **idempotente**, ou seja,

$$P = P^2$$

assim, para todo $x \in H$,

$$P^2(x) = P(P(x)) = P(x)$$

Portanto, $P|_Y$ é o operador identidade em Y. E para $Z = Y^{\perp}$ temos imediatamente o seguinte lema

Lema 2.3.6 *O complemento ortogonal $Y^{\perp}$ de um subespaço fechado Y de um espaço de Hilbert H é o núcleo $N(P)$ da projeção ortogonal P de H em Y.*

Um complemento ortogonal é um espaço nulo especial, onde, por definição, o espaço nulo $M^{\perp}$ de um conjunto $M \neq \emptyset$ em um espaço com produto interno X é o conjunto

$$M^{\perp} = \{x \in X | x \perp M\}.$$

Assim, $x \in M^{\perp}$ se, e somente se, $\langle x, v \rangle = 0$ para todo $v \in M$. Isso explica o nome.

Observe que $M^{\perp}$ é um espaço vetorial, pois $x, y \in M^{\perp}$ implica para todo $v \in M$ e todos os escalares α, β

$$\langle \alpha x + \beta y, v \rangle = \alpha \langle x, v \rangle + \beta \langle y, v \rangle = 0$$

portanto, $\alpha x + \beta y \in M$.

$$M \subset M^{\perp\perp}. \qquad (6^*)$$

de fato, $x \in M \Rightarrow x \perp M^{\perp} \Rightarrow x \in M^{\perp\perp}$.

Proposição 2.3.7 *Se Y é um subespaço fechado de um espaço de Hilbert H, então*

$$Y = Y^{\perp\perp}. \qquad (6)$$

Demonstração:

$Y \subset Y^{\perp\perp}$ por (8*). Mostraremos que $Y^{\perp\perp} \subset Y$. Seja $x \in Y^{\perp\perp}$. Então $x = y + z$ pelo Teorema 2.3.5, onde $y \in Y \subset Y^{\perp\perp}$ por (8*). Como $Y^{\perp\perp}$ é um espaço vetorial e $x \in Y^{\perp\perp}$ por suposição, também temos $z = x - y \in Y^{\perp\perp}$, portanto $z \perp Y^{\perp}$. Mas $z \in Y^{\perp}$ por 2.3.5. Juntos $z \perp z$, portanto $z = 0$, de modo que $x = y$, isto é, $x \in Y$. Como $x \in Y^{\perp\perp}$ foi arbitrário, isso prova que $Y^{\perp\perp} \subset Y$. ∎

A propriedade (6) é a principal razão para o uso de subespaços fechados no contexto atual. Uma vez que $Z^\perp = Y^{\perp\perp} = Y$, a fórmula (4) também pode ser escrita

$$H = Z \oplus Z^\perp$$

Segue-se que $x \longmapsto z$ define uma projeção (Fig. 2.1)

$$P_Z : H \longrightarrow Z \qquad (9)$$

de H em Z, cujas propriedades são bastante semelhantes às da projeção P considerada antes.

Fecharemos a sessão com um resultado sobre conjuntos densos em espaços de Hilbert.

Proposição 2.3.8 *Para qualquer subconjunto $M \neq \emptyset$ de um espaço de Hilbert H, o $spanM$ é denso em H se, e somente se, $M^\perp = 0$.*

Demonstração:
(a) Seja $x \in M^\perp$ e suponha que $V = spanM$ seja denso em H. Então $x \in \overline{V} = H$. Pelo Teorema 0.0.9(a) existe uma sequência (x_n) em V tal que $x_n \longrightarrow x$. Como $x \in M^\perp$ e $M^\perp \perp V$, temos $\langle x_n, x\rangle = 0$. A continuidade do produto interno (Lema 2.2.2) implica que $\langle x_n, x\rangle \longrightarrow \langle x, x\rangle$. Juntos, $\langle x, x\rangle = ||x||^2 = 0$, de modo que $x = 0$. Como $x \in M^\perp$ foi arbitrário, isso mostra que $M^\perp = \{0\}$.

(b) Reciprocamente, suponha $M^\perp = \{0\}$. Se $x \perp V$, então $x \perp M$, de modo que $x \in M^\perp$ e $x = 0$. Portanto, $V^\perp = O$. Observando que V é subespaço de H, obtemos assim $\overline{V} = H$ do Teorema 2.3.5 com $Y = \overline{V}$. ■

2.4 Conjuntos e sequências ortonormais

Ortogonalidade dos elementos conforme definido na Sec. 3.1 desempenha um papel fundamental nos espaços com produto interno e espaços de Hilbert. Uma primeira impressão desse fato foi dada na seção anterior. De particular interesse são conjuntos cujos elementos são ortogonais em pares. Para entender isso, lembremo-nos uma situação familiar no espaço euclidiano $\mathbb{R}^3$. No espaço $\mathbb{R}^3$, um conjunto desse tipo é o conjunto formado por três vetores unitários nas direções positivas do eixos de um sistema de coordenadas retangulares; chame esses vetores de e_1, e_2, e_3. Esses vetores formam uma base para $\mathbb{R}^3$, de modo que todo $x \in \mathbb{R}^3$ tem uma representação única (Fig. 2.2)

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3.$$

Agora vemos uma grande vantagem da ortogonalidade. Dado x, podemos determinar prontamente os coeficientes desconhecidos $\alpha_1, \alpha_2, \alpha_3$ através do produto interno. De fato, para obter α_1 basta multiplicar a representação de x por e_1, ou seja,

$$\langle x, e_1\rangle = \alpha_1 \langle e_1, e_1\rangle + \alpha_2 \langle e_2, e_1\rangle + \alpha_3 \langle e_3, e_1\rangle = \alpha_1,$$

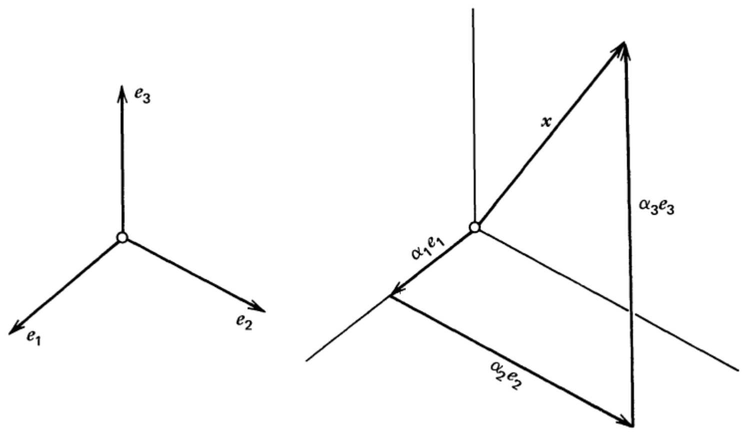

Figura 2.2: Conjunto Ortonormal $\{e_1, e_2, e_3\}$ em $\mathbb{R}^3$ e a representação $x = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3$.

e assim por diante. Em espaços de produto interno mais gerais existem outras possibilidades para o uso de conjuntos ortogonais, ortonormais e sequências, como explicaremos. De fato, a aplicação de tais conjuntos e sequências compõe uma parte substancial de toda a teoria de produtos internos e espaços de Hilbert. Comecemos nosso estudo sobre esta situação, introduzindo os conceitos necessários.

Definição 2.4.1 (Conjuntos Ortonormais e sequências). Um conjunto ortogonal M em um espaço com produto interno X é um subconjunto $M \subset X$ cujos elementos são ortogonais aos pares. Um conjunto ortonormal $M \subset X$ é um conjunto ortogonal definido em X cujos elementos têm norma 1, ou seja, para todo $x, y \in M$,

$$\langle x, y \rangle = \begin{cases} 0 & \text{se } x \neq y \\ 1 & \text{se } x = y \end{cases}. \qquad (1)$$

Se um conjunto ortogonal ou ortonormal M é enumerável, podemos organizá-lo em uma sequência (x_n) e denominar de sequência ortogonal ou ortonormal, respectivamente.

Mais geralmente, um conjunto indexado, ou família, (x_α), $\alpha \in I$, é chamado ortogonal se $x_\alpha \perp x_\beta$ para todo $\alpha, \beta \in I, \alpha \neq \beta$. A família chama-se ortonormal se for ortogonal e todos os (x_α) tiverem norma 1, de modo que para todos $\alpha, \beta \in I$ temos

$$\langle x_\alpha, x_\beta \rangle = \delta_{\alpha\beta} = \begin{cases} 0 & \text{se } \alpha \neq \beta \\ 1 & \text{se } \alpha = \beta. \end{cases}$$

Aqui, $\delta_{\alpha\beta}$ é o delta de Kronecker, como na Sec. 1.8.

Para os elementos ortogonais x, y temos $\langle x, y \rangle = 0$, de modo que vale o Teorema de Pitágoras

$$||x + y||^2 = ||x||^2 + ||y||^2. \qquad (3)$$

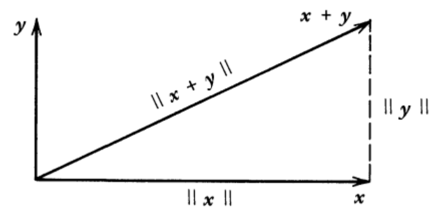

Figura 2.3: Teorema de Pitagoras (3) em $\mathbb{R}^2$

Mais geralmente, se $\{x_1, \cdots, x_n\}$ é um conjunto ortogonal, então

$$||x_1 + \cdots + x_n||^2 = ||x_1||^2 + \cdots + ||x_n||^2. \qquad (4)$$

De fato, $\langle x_j, x_k \rangle = 0$ se $j \neq k$; consequentemente,

$$\left\| \sum_j x_j \right\|^2 = \left\langle \sum_j x_j, \sum_k x_k \right\rangle = \sum_j \sum_k \langle x_j, x_k \rangle = \sum_j \langle x_j, x_j \rangle = \sum_j \|x_j\|^2$$

(soma de 1 até n). Também notamos

Lema 2.4.2 *Um conjunto ortonormal é linearmente independente.*

Demonstração:
Seja $\{e_1, \cdots, e_n\}$ um conjunto ortonormal e considere a equação

$$\alpha_1 e_1 + \cdots + \alpha_n e_n = 0$$

A multiplicação por um e_j fixo dá

$$\left\langle \sum_k \alpha_k e_k, e_j \right\rangle = \sum_k \alpha_k \langle e_k, e_j \rangle = \alpha_j \langle e_j, e_j \rangle = \alpha_j = 0$$

e prova a independência linear para qualquer conjunto ortonormal finito. este fato também implica independência linear se o conjunto ortonormal dado for infinito, pela definição de independência linear na Seção 1.1. ∎

Exemplo 2.4.1 O espaço $\mathbb{R}^3$ os três vetores unitários $(1, 0, 0), (0, 1, 0), (0, 0, 1)$ na direção dos três eixos de um sistema de coordenadas retangulares formam um conjunto ortonormal.

Exemplo 2.4.2 No espaço l^2, uma sequência ortonormal é (e_n), onde $e_n = (\delta_{nj})$ onde tem o n-ésimo elemento igual a 1 e todos os outros nulos.

Exemplo 2.4.3 Seja X o espaço com produto interno de todas funções contínuas de valores reais em $[0, 2\pi]$ com produto interno definido por

$$\langle x, y \rangle = \int_0^{2\pi} x(t)y(t)dt$$

Uma sequência ortogonal em X é (u_n), onde

$$u_n(t) = cos(nt) \qquad n = 0, 1, \cdots .$$

Outra sequência ortogonal em X é (v_n), onde

$$v_n(t) = sen(nt) \qquad n = 0, 1, \cdots .$$

De fato, por integração obtemos

$$\langle u_m, u_n \rangle = \int_0^{2\pi} \cos(mt)\cos(nt)dt = \begin{cases} 0 & \text{se } m \neq n \\ \pi & \text{se } m = n = 1, 2, \cdots \\ 2\pi & \text{se } m = n = 0 \end{cases} \qquad (5)$$

e similarmente para (v_n). Portanto, uma sequência ortonormal é (e_n), onde

$$e_0(t) = \frac{1}{\sqrt{2\pi}}, \quad e_n(t) = \frac{u_n(t)}{\|u_n\|} = \frac{\cos(nt)}{\sqrt{\pi}} \qquad (n = 1, 2, \cdots).$$

De (v_n) obtemos a sequência ortonormal $(\tilde{e}_n)$, onde

$$\tilde{e}_n(t) = \frac{v_n(t)}{\|v_n\|} = \frac{sen(nt)}{\sqrt{\pi}} \qquad (n = 1, 2, \cdots).$$

Uma grande vantagem de sequências ortonormais sobre sequências arbitrárias linearmente independentes é a seguinte. Se soubermos que um dado x pode ser representado como uma combinação linear de alguns elementos de uma sequência ortonormal, então a ortonormalidade torna a determinação dos coeficientes muito fácil. De fato, se $(e_1, e_2, \cdots)$ é uma sequência ortonormal em um espaço com produto interno X e temos $x \in span\{e_1, \cdots, e_n\}$, onde n é fixo, então pela definição de *span*,

$$x = \sum_{k=1}^{n} \alpha_k e_k, \qquad (6)$$

e se tomarmos o produto interno por um fixo e_j obtemos

$$\langle x, e_j \rangle = \left\langle \sum \alpha_k e_k, e_j \right\rangle = \sum \alpha_k \langle e_k, e_j \rangle = \alpha_j.$$

Com esses coeficientes, (6) se torna

$$x = \sum_{k=1}^{n} \langle x, e_k \rangle e_k. \qquad (7)$$

Isso mostra que a determinação dos coeficientes desconhecidos em (6) é simples. Outra vantagem da ortonormalidade se torna aparente se em (6) e (7) queremos adicionar outro termo $\alpha_{n+1}e_{n+1}$,

$$\tilde{x} = x + \alpha_{n+1}e_{n+1} \in span\{e_1, \cdots, e_{n+1}\};$$

então precisamos calcular apenas mais um coeficiente já que os outros coeficientes permanecem inalterados.

Mais geralmente, se considerarmos qualquer $x \in X$, não necessariamente em $Y_n = span\{e_1, \cdots, e_n\}$, podemos definir $y \in Y_n$ configurando

$$y = \sum_{k=1}^{n} \langle x, e_k \rangle e_k, \qquad (8a)$$

onde n é fixo como antes, e então defina z da seguinte forma

$$x = y + z, \qquad (8b)$$

isto é, $z = x - y$. Queremos mostrar que $z \perp y$. Para realmente entender o que está acontecendo, observe o seguinte, cada $y \in Y_n$ é uma combinação linear

$$y = \sum_{k=1}^{n} \alpha_k e_k.$$

Onde $\alpha_k = \langle y, e_k \rangle$, como segue do que discutimos antes. Nossa afirmação é que para a escolha particular $\alpha_k = \langle x, e_k \rangle$, $k = 1, \cdots, n$, temos que obter um y tal que $z = x - y \perp y$.

Para provar isso, notamos primeiro que, pela ortonormalidade,

$$\|y\|^2 = \left\langle \sum \langle x, e_k \rangle e_k, \sum \langle x, e_m \rangle e_m \right\rangle = \sum |\langle x, e_k \rangle|^2. \qquad (9)$$

Usando isso, podemos agora mostrar que $z \perp y$:

$$\begin{aligned}
\langle z, y \rangle = \langle x - y, y \rangle &= \langle x, y \rangle - \langle y, y \rangle \\
&= \left\langle x, \sum \langle x, e_k \rangle e_k \right\rangle - \|y\|^2 \\
&= \sum \langle x, e_k \rangle \overline{\langle x, e_k \rangle} - \sum |\langle x, e_k \rangle|^2 \\
&= 0.
\end{aligned}$$

Portanto, a relação pitagórica (3) dá

$$||z||^2 = ||y||^2 + ||z||^2. \qquad (10)$$

Por (9) segue que

$$||z||^2 = ||x||^2 + ||y||^2 = ||x||^2 - \sum |\langle x, e_k \rangle|^2 \qquad (11)$$

Como $||z|| \geqslant 0$, temos para todo $n = 1, 2, \cdots$

$$\sum_{k=1}^{n} |\langle x, e_k \rangle|^2 \leqslant ||x||^2. \qquad (12^*)$$

Essas somas têm termos não negativos, de modo que formam uma sequência monótona crescente. Esta sequência converge porque é limitada por $||x||^2$. Esta é a sequência das somas parciais de uma série infinita, que assim converge. Portanto (12^*) implica

Teorema 2.4.3 *Seja (e_k) uma sequência ortonormal em um espaço com produto interno X. Então para todo $xinX$*

$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2 \leqslant ||x||^2 \qquad \textbf{\textit{(Desigualdade de Bessel).}} \qquad (12)$$

Os produtos internos $\langle x, e_k \rangle$ em (12) são chamados de coeficientes de Fourier de x em relação à sequência ortonormal (e_k).

Observe que se X é de dimensão finita, então todo conjunto ortonormal em X deve ser finito porque é linearmente independente por 2.4.2. Daí em (12) temos então uma soma finita.

Processo de Ortonormalização de Gram-Schmidt

Vimos que sequências ortonormais são muito convenientes para trabalhar com elas. O problema prático restante é como obter uma sequência ortonormal de uma sequência arbitrária linearmente independente dada. Isso é realizado por um procedimento construtivo, o **processo de Gram-Schmidt** para ortonormalização de uma sequência linearmente independente (x_j) em um espaço com produto interno. A sequência ortonormal resultante (e_j) tem a propriedade de que para todo n,

$$span\{e_1, \cdots, e_n\} = span\{x_1, \cdots, x_n\}.$$

O processo é como se segue.

1º Passo. O primeiro elemento de (e_k) é

$$e_1 = \frac{1}{||x_1||} x_1.$$

2º Passo. x_2 pode ser escrito da seguinte forma:

$$x_2 = \langle x_2, e_1 \rangle e_1 + v_2$$

Então

$$v_2 = x_2 - \langle x_2, e_1 \rangle e_1$$

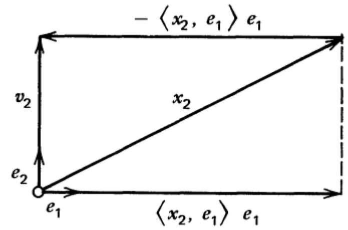

Figura 2.4: Processo de Gram-Schmidt, 2º Passo.

não é o vetor zero já que (x_j) é linearmente independente; temos $v_2 \perp e_1$ uma vez que $\langle v_2, e_1 \rangle = 0$, de modo que podemos tomar

$$e_2 = \frac{1}{||v_2||} v_2.$$

3º Passo. O vetor

$$v_3 = x_3 - \langle x_3, e_1 \rangle e_1 - \langle x_3, e_2 \rangle e_2$$

não é o vetor zero, e $v_3 \perp e_1$ bem como $v_3 \perp e_2$. Daí, tomamos

$$e_3 = \frac{1}{||v_3||} v_3.$$

nº Passo. O vetor

$$v_n = x_n - \sum_{k=1}^{n-1} \langle x_n, e_k \rangle e_k \qquad (13)$$

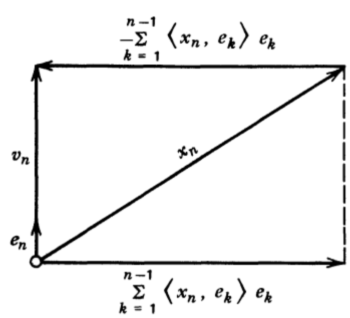

Figura 2.5: Processo de Gram-Schmidt, nº Passo.

não é o vetor zero e é ortogonal a $e_1, \cdots, e_{n-1}$. A partir dele nós obtemos

$$e_n = \frac{1}{||v_n||} v_n. \qquad (14)$$

Estas são as fórmulas gerais para o processo de Gram-Schmidt. Observe que a soma subtraída do lado direito de (13) é a projeção de x_n no $span\{e_1, \cdots, e_{n-1}\}$. Em outras palavras, em cada passo, subtraímos de x_n suas "componentes" nas direções dos vetores previamente ortogonalizados. Isso dá v_n, que é então multiplicado por $\frac{1}{||v_n||}$, de modo que obtemos um vetor de norma 1. v_n não pode ser o vetor zero para qualquer n. De fato, se n fosse o menor índice para o qual $v_n = 0$, então (13) mostra que x_n seria uma combinação linear de $e_1, \cdots, e_{n-1}$, portanto, uma combinação linear de $x_1, \cdots, x_{n-1}$, contradizendo a suposição de que $\{x_1, \cdots, x_n\}$ é linearmente independente.

2.5 Conjuntos e sequências ortonormais totais

Definição 2.5.1 (Conjunto ortonormal total). Um conjunto total (ou conjunto fundamental) em um espaço normado X é um subconjunto $M \subset X$ cuja extensão é densa em X. Assim, um conjunto ortonormal (ou sequência ou família) em um espaço com produto interno X que é total em X é chamado de conjunto ortonormal total (ou sequência ou família, respectivamente) em X.

M é total em X se, e somente se,

$$\overline{spanM} = X.$$

Isso é óbvio a partir da definição.

Uma família ortonormal total em X às vezes é chamada de base ortonormal para X. No entanto, é importante notar que esta não é uma base, no sentido da álgebra, para X como um espaço vetorial, a menos que X tenha dimensão finita.

Teorema 2.5.2 *Seja M um subconjunto de um espaço com produto interno X. Então:*

(a) *Se M é total em X, então não existe um $x \in X$ diferente de zero que é ortogonal a cada elemento de M; brevemente,*

$$x \perp M \Rightarrow x = 0. \qquad (1)$$

(b) *Se X é completo, essa condição também é suficiente para a totalidade de M em X.*

Demonstração:

(a) Seja H o completamento de X. Então X, considerado como um subespaço de H, é denso em H. Por suposição, M é total em X, de modo que o $spanM$ é denso em X, portanto, denso em H. A Proposição 2.3.8 implica que o complemento ortogonal de M em H é $\{0\}$. A princípio, se $x \in X$ e $x \perp M$, então $x = 0$.

(b) Se X é um espaço de Hilbert e M satisfaz essa condição, de modo que $M^{\perp} = 0$, então a Proposição 2.3.8 implica que M é total em X. ■

A completude de X em (b) é essencial. Se X não for completo, pode não existir um conjunto ortonormal $M \subset X$ tal que M seja total em X.

Do Teorema 2.4.3

$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2 \leqslant ||x||^2 \qquad \textbf{(Desigualdade de Bessel).} \qquad (2)$$

daremos um nome especial para o caso em que vale a igualdade na desigualdade de Bessel, chamaremos esse caso de relação de Parseval

$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2 = ||x||^2 \qquad \textbf{(Relação de Parseval).} \qquad (3)$$

e produz outro critério para a totalidade:

Teorema 2.5.3 *Um conjunto ortonormal M em um espaço de Hilbert H é total em H se, e somente se, para todo $x \in H$ a relação de Parseval (3) é válida (soma sobre todos os coeficientes de Fourier diferentes de zero de x em relação a M).*

Demonstração:

(a) Se M não é total, pelo Teorema 2.5.2 existe um vetor diferente de zero $x \perp M$ em H. Como $x \perp M$, em (3) temos $\langle x, e_k \rangle = 0$ para todo k, de modo que o lado esquerdo em (3) é zero, enquanto $||x||^2 \neq 0$. Isso mostra que (3) não vale. Portanto, se (3) vale para todo $x \in H$, então M deve ser total em H.

(b) Por outro lado, suponha que M seja total em H. Considere qualquer $x \in H$ e seus coeficientes de Fourier diferentes de zero dispostos em uma sequência $\langle x, e_1 \rangle, \langle x, e_2 \rangle, \cdots$, ou escrita em alguma ordem definida se houver são apenas um número finito deles diferente de zero. Agora definimos y por

$$y = \sum_k \langle x, e_k \rangle e_k, \qquad (4)$$

observando que no caso de uma série infinita, a convergência segue de Teorema 2.4.3. Vamos mostrar que $x - y \perp M$. Para cada e_j ocorrendo em (4) temos, usando a ortonormalidade,

$$\langle x - y, e_j \rangle = \langle x, e_j \rangle - \sum_k \langle x, e_k \rangle \langle e_k, e_j \rangle = \langle x, e_j \rangle - \langle x, e_j \rangle = 0$$

E para cada $v \in M$ não contido em (4) temos $\langle x, v \rangle = 0$, de modo que

$$\langle x - y, v \rangle = \langle x, v \rangle - \sum_k \langle x, e_k \rangle \langle e_k, v \rangle = 0 - 0 = 0$$

Daí $x - y \perp M$, isto é, $x - y \in M^\perp$. Como M é total em H, temos $M^\perp = 0$ de 2.3.8. Juntos, $x - y = 0$, ou seja, $x = y$. Usando (4) e novamente a ortonormalidade, obtemos (3) de

$$\|x\|^2 = \left\langle \sum_k \langle x, e_k \rangle e_k, \sum_m \langle x, e_m \rangle e_m \right\rangle = \sum_k \langle x, e_k \rangle \overline{\langle x, e_k \rangle}.$$

Isso completa a prova. ∎

Capítulo 3

Os 4 Teoremas do Apocalipse

Neste Capítulo mostraremos 4 teoremas fundamentais em Análise Funcional que são ferramentas eficazes na resolução de problemas de outras áreas da Matemática. Serão abordados o Teorema de Representação de Riesz, o Teorema da Aplicação aberta, o Teorema do gráfico fechado e o Teorema Hahn-Banach.

3.1 Teorema da Representação de Riesz

É de importância prática conhecer a forma geral de funcionais lineares em vários espaços. Isso foi apontado e explicado no Capítulo 1. Para espaços gerais de Banach, tais fórmulas e suas consequências às vezes pode ser complicado. No entanto, para um espaço de Hilbert, a situação é surpreendentemente simples.

Teorema 3.1.1 ***(Teorema da Representação de Riesz).*** *Todo funcional linear limitado f em um espaço de Hilbert H pode ser representado em termos do produto interno, ou seja,*

$$f(x) = \langle x, z \rangle \qquad (1)$$

onde z depende de f, é determinado exclusivamente por f e tem norma

$$||z|| = ||f||. \qquad (2)$$

Demonstração:

A demonstração seguirá os seguintes passos:

(a) f tem uma representação (1),

(b) z em (1) é único,

(c) $||z|| = ||f||$.

Seguem os detalhes a seguir.

(a) Se $f = 0$, então (1) e (2) valem se tomarmos $z = 0$. Seja $f \neq 0$. Para motivar a ideia da prova, vamos perguntar quais propriedades z deve ter se existir uma representação (1). Em primeiro lugar, $z \neq 0$, caso contrário $f = 0$. Segundo, $\langle x, z \rangle = 0$ para todo x para o qual $f(x) = 0$, ou seja, para todo x no núcleo $N(f)$ de f. Daí $z \perp N(f)$. Isso sugere que consideremos $N(f)$ e seu complemento ortogonal $N(f)^{\perp}$.

$N(f)$ é um espaço vetorial pelo Teorema 1.5.2 e é fechado pelo Teorema 1.6.5. Além disso, $f \neq 0$ implica $N(f) \neq H$, de modo que $N(f)^{\perp} \neq 0$ pelo Teorema 2.3.5. Portanto, $N(f)^{\perp}$ contém um $z_0 \neq 0$. Fazendo

$$v = f(x)z_0 - f(z_0)x$$

onde $x \in H$ é arbitrário. Aplicando f, obtemos

$$f(v) = f(x)f(z_0) - f(z_0)f(x) = 0.$$

Isso mostra que $v \in N(f)$. Como $z_0 \perp N(f)$, temos

$$\begin{aligned} 0 = \langle v, z_0 \rangle &= \langle f(x)z_0 - f(z_0)\, x, z_0 \rangle \\ &= f(x) \langle z_0, z_0 \rangle - f(z_0) \langle x, z_0 \rangle . \end{aligned}$$

Observando que $\langle z_0, z_0 \rangle = \|z_0\|^2 \neq 0$, podemos resolver para $f(x)$. O resultado é

$$f(x) = \frac{f(z_0)}{\langle z_0, z_0 \rangle} \langle x, z_0 \rangle$$

Isso pode ser escrito na forma (1), onde

$$z = \frac{\overline{f(z_0)}}{\langle z_0, z_0 \rangle} z_0.$$

Como $x \in H$ é arbitrário, (1) está provado.

(b) Vamos provar que z em (1) é único. Suponha que para todo $x \in H$,

$$f(x) = \langle x, z_1 \rangle = \langle x, z_2 \rangle .$$

Então $\langle x, z_1 - z_2 \rangle = 0$ para todo x. Escolhendo em particular $x = z_1 - z_2$, temos

$$\langle x, z_1 - z_2 \rangle = \langle z_1 - z_2, z_1 - z_2 \rangle = \|z_1 - z_2\|^2 = 0.$$

Daí $z_1 - z_2 = 0$, de modo que $z_1 = z_2$, provando a unicidade.

(c) Finalmente provaremos (2). Se $f = 0$, então $z = 0$ e (2) vale. Seja $f \neq 0$. Então $z \neq 0$. De (1) com $x = z$ e (3) na Seção 1.7 obtemos

$$\|z\|^2 = \langle z, z \rangle = f(z) \leqq \|f\| \|z\|.$$

Dividindo por $||z||neq0$ produz $||z|| \leqslant ||f||$. Resta mostrar que $||f|| \leqslant ||z||$. De (1) e da desigualdade de Schwarz (Seção 2.2) vemos que

$$|f(x)| = |\langle x, z \rangle| \leqq \|x\| \|z\|.$$

Isso implica

$$\|f\| = \sup_{\|x\|=1} |\langle x, z \rangle| \leqq \|z\|.$$

Completando dessa forma a demonstração. ∎

A ideia da prova de unicidade na parte (b) é digna de nota para uso posterior:

Corolário 3.1.2 *Se $\langle v_1, w \rangle = \langle v_2, w \rangle$ para todo w em um espaço com produto interno X, então $v_1 = v_2$. Em particular, $\langle v_1, w \rangle = 0$ para todo $w \in X$ implica $v_1 = 0$.*

Demonstração:
Por suposição, para todo w,

$$\langle v_1 - v_2, w \rangle = \langle v_1, w \rangle - \langle v_2, w \rangle = 0$$

Para $w = v_1 - v_2$ isso dá $||v_1 - v_2|| = 0$. Portanto $v_1 - v_2 = 0$, de modo que $v_1 = v_2$. Em particular, $\langle v_1, w \rangle = 0$ com $w = v_1$ dá $||v_1||^2 = 0$, de modo que $v_1 = 0$. ∎

3.2 Teorema da Aplicação Aberta

Nesta seção será discutido um resultado técnico importante em Análise Funcional, o Teorema da Aplicação Aberta, devido à Banach. Numa de suas consequências, ele dá condições suficientes para que uma aplicação linear entre espaços de Banach contínua e invertível tenha inversa contínua, em outras palavras, para que seja um homeomorfismo linear. Esse teorema será usado para demonstrar o Teorema do Gráfico Fechado na próxima seção.

Vamos começar introduzindo o conceito de aplicação aberta.

Definição 3.2.1 Sejam X e Y espaços métricos. Então $T : D(T) \longrightarrow Y$ com domínio $D(T) \subset X$ é chamado de aplicação aberta se para cada conjunto aberto em $D(T)$ a imagem for um conjunto aberto em Y.

Observe que se uma aplicação não é sobrejetiva, deve-se tomar cuidado para distinguir entre as afirmações de que a aplicação é aberta como uma aplicação de seu domínio:
(a) em Y,
(b) em sua imagem.

(b) é mais fraco que (a). Por exemplo, se $X \subset Y$, a aplicação $x \longrightarrow x$ de X em Y é aberto se, e somente se, X é um subconjunto aberto de Y, enquanto a aplicação $x \longrightarrow x$ de X em $Im(X)$ (que é X) é aberto em qualquer caso.

Definiremos agora o conceito de categoria em um espaço métrico.

Definição 3.2.2 Categoria Um subconjunto M de um espaço métrico X é dito ser

(a) raro (ou em nenhum lugar denso) em X se seu fecho $\overline{M}$ não tiver nenhum ponto interior.

(b) magro (ou de primeira categoria) em X se M é a união de conjuntos enumeráveis cada um dos quais é raro em X,

(c) não magro (ou de segunda categoria) em X se M não é magro em X.

A seguir provaremos o seguinte resultado de Espaços métricos, conhecido como Teorema da Categoria de Baire.

Lema 3.2.3 ***(Teorema da Categoria de Baire).*** *Se um espaço métrico $X \neq \emptyset$ é completo, ele não é magro em si mesmo. Portanto, se $X \neq \emptyset$ é completo e*

$$X = \bigcup_{k=1}^{\infty} A_k \qquad (A_k \textbf{\textit{ fechado}})$$

então pelo menos um A_k contém um subconjunto aberto não vazio.

Demonstração:
A ideia da prova é simples. Suponha que o espaço métrico completo $X \neq \emptyset$ é magro em si mesmo. Então

$$X = \bigcup_{k=1}^{\infty} M_k \qquad (1^*)$$

com cada M_k raro em X. Construiremos uma sequência de Cauchy (p_k) cujo limite p (que existe por completude) não está em M_k, portanto contrariando a representação (1^*).

Por suposição, M_1 é raro em X, de modo que, por definição, M_1 não contém um conjunto aberto não vazio. Mas X sim (por exemplo, ele mesmo). Este fato implica

$\overline{M_1} \neq X$. Portanto, o complemento $\overline{M_1}^c = X - \overline{M_1}$ de M_1 não é vazio e aberto. Podemos assim escolher um ponto p_1 em M_1 e uma bola aberta, digamos,

$$B_1 = B(p_1; \varepsilon_1) \subset \bar{M}_1^c, \qquad \varepsilon_1 < \frac{1}{2}.$$

Por suposição, M_2 é raro em X, de modo que M_2 não contém um conjunto aberto não vazio. Portanto, não contém a bola aberta $B\left(p_1; \frac{1}{2}\varepsilon_1\right)$. Isso implica que $\overline{M_2}^c \cap B\left(p_1; \frac{1}{2}\varepsilon_1\right)$ não é vazio e aberto, de modo que podemos escolher uma bola aberta neste conjunto, digamos,

$$B_2 = B(p_2; \varepsilon_2) \subset \bar{M_2}^C \cap B\left(p_1; \frac{1}{2}\varepsilon_1\right), \qquad \varepsilon_2 < \frac{1}{2}\varepsilon_1.$$

Por indução obtemos assim uma sequência de bolas

$$B_k = B(p_k; \varepsilon_k), \qquad \varepsilon_k < 2^{-k}.$$

tal que $B_k \cap M_k = 0$ e

$$B_{k+1} \subset B\left(p_k; \frac{1}{2}\varepsilon_k\right) \subset B_k, \qquad k = 1, 2, \cdots.$$

Como $\varepsilon_k < 2^{-k}$, a sequência (p_k) é de Cauchy e converge, digamos, $p_k \longrightarrow p \in X$ porque X é completo por suposição. Também, para cada m e $n > m$ temos $B_n \subset B\left(p_m; \frac{1}{2}\varepsilon_m\right)$, de modo que

$$\begin{aligned} d(p_m, p) &\leqq d(p_m, p_n) + d(p_n, p) \\ &< \frac{1}{2}\varepsilon_m + d(p_n, p) \quad \longrightarrow \quad \frac{1}{2}\varepsilon_m \end{aligned}$$

como $n \longrightarrow \infty$. Portanto, $p \in B_m$ para todo m. Como $B_m \subset \overline{M_m}^c$, vemos agora que $p \notin M_m$ para todo m, de modo que $p \notin \bigcup M_m = X$. Isso contradiz $p \in X$. Logo o teorema de Baire está provado.

■

A demonstração do Teorema da aplicação aberta, sairá do seguinte resultado,

Lema 3.2.4 *Um operador linear limitado T sobrejetivo de um espaço de Banach X em um espaço de Banach Y tem a propriedade de que a imagem $T(B_0)$ da bola unitária aberta $B_0 = B(0; 1) \subset X$ contém uma bola aberta de centro $0 \in Y$.*

Demonstração:

Procedendo passo a passo, provaremos:

(a) O fecho da imagem da bola aberta $B_1 = B(0; \frac{1}{2})$ contém uma bola aberta B^*.

(b) $\overline{T(B_n)}$ contém uma bola aberta V_n de centro $0 \in Y$, onde $B_n = B(0; 2^{-n}) \subset X$.

(c) $T(B_0)$ contém uma bola aberta em torno de $0 \in Y$.

Segue os detalhes da prova.

(a) Em conexão com os subconjuntos $A \subset X$, escreveremos αA (α um escalar) e $A + w$ ($w \in X$) para significar

$$\alpha A = \{x \in X | x = \alpha a, a \in A\} \qquad (1)$$

$$A + w = \{x \in X | x = a + w, a \in A\} \qquad (2)$$

e da mesma forma para subconjuntos de Y.

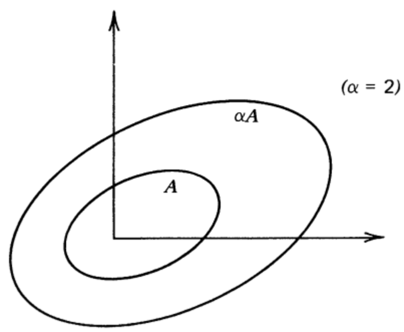

Figura 3.1: Ilustração da fórmula (1).

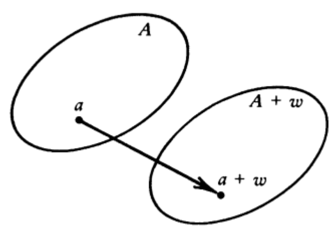

Figura 3.2: Ilustração da fórmula (2).

Consideramos a bola aberta $B_1 = B(0; \frac{1}{2}) \subset X$. Qualquer $x \in X$ fixo está em kB_1 com k real suficientemente grande ($k > 2||x||$). Por isso

$$X = \bigcup_{k=1}^{\infty} kB_k$$

Como T é sobrejetivo e linear,

$$Y = T(X) = T\left(\bigcup_{k=1}^{\infty} kB_1\right) = \bigcup_{k=1}^{\infty} kT(B_1) = \bigcup_{k=1}^{\infty} \overline{kT(B_1)}. \qquad (3)$$

Observe que, ao fazer os fechamentos, não adicionamos mais pontos à união já que essa união já era todo o espaço Y. Como Y é completo, é não magro em si, pelo teorema da categoria de Baire 3.2.3. Por isso, observando que (3) é semelhante a (1) em 3.2.3, concluímos que um $\overline{kT(B_1)}$ deve conter alguma bola aberta. Isso implica que $\overline{T(B_1)}$ também contém um bola aberta, digamos, $B^* = B(y_0; \varepsilon) \subset \overline{T(B_1)}$. Segue que

$$B^* - y_0 = B(0; \varepsilon) \subset \overline{T(B_1)} - y_0. \qquad (4)$$

(b) Provamos que $B^* - y_0 \subset \overline{T(B_0)}$, onde B_0 é dado no teorema. Isso fazemos mostrando que

$$\overline{T(B_1)} - y_0 \subset \overline{T(B_0)}. \qquad (5)$$

Seja $y \in \overline{T(B_1)} - y_0$. Então $y + y_0 \in \overline{T(B_1)}$, e lembramos que $y_0 \in \overline{T(B_1)}$, também. Por 0.0.9(a) existem

$$u_n = T(w_n) \in T(B_1) \qquad \text{de tal modo que} \qquad u_n \longrightarrow y + y_0,$$

$$v_n = T(z_n) \in T(B_1) \qquad \text{de tal modo que} \qquad v_n \longrightarrow y_0,$$

Como $w_n, z_n \in B_1$ e B_1 tem raio $\frac{1}{2}$, segue que

$$||w_n - z_n|| \leqslant ||w_n|| + ||z_n|| < \frac{1}{2} + \frac{1}{2} = 1,$$

de modo que, $w_n - z_n \in B_0$. A partir de,

$$T(w_n - z_n) = T(w_n) - T(z_n) = u_n - v_n \longrightarrow y.$$

nós vemos que $y \in \overline{T(B_0)}$. Como foi arbitrário, isso prova (5). De (4) temos assim

$$B^* - y_0 = B(0; \varepsilon) \subset \overline{T(B_0)}. \qquad (6)$$

Seja $B_n = B(0; 2^{-n}) \subset X$. Como T é linear, $\overline{T(B_n)} = 2^{-n}\overline{T(B_0)}$. De (6) obtemos assim

$$V_n = B(0; \varepsilon/2^n) \subset \overline{T(B_n)}. \qquad (7)$$

(c) Finalmente provaremos que

$$V_1 = B(0; \frac{1}{2}\varepsilon) \subset T(B_0)$$

mostrando que para todo $y \in V_1$ está em $T(B_0)$. Então seja $y \in V_1$. De (7) com $n = 1$ temos $V_1 \subset \overline{T(B_1)}$. Portanto, $y \in \overline{T(B_1)}$. Por 0.0.9(a) deve haver um $v \in T(B_1)$ próximo de y, digamos, $||y - v|| < \varepsilon/4$. Agora $v \in T(B_1)$ implica $v = T(x_1)$ para alguns $x_1 \in B_1$. Por isso

$$||y - T(x_1)|| < \frac{\varepsilon}{4}.$$

Disto e (7) com $n = 2$ vemos que $y - T(x_1) \in V_2 \subset \overline{T(B_2)}$. Como antes nós concluímos que existe um $x_2 \in B_2$ tal que

$$||(y - T(x_1)) - T(x_2)|| < \frac{\varepsilon}{8}.$$

Como $y - T(x_1) - T(x_2) \in V_3 \subset \overline{T(B_3)}$, e assim por diante. Na enésima etapa podemos escolher um $x_n \in B_n$ tal que

$$\left\| y - \sum_{k=1}^{n} Tx_k \right\| < \frac{\varepsilon}{2^{n+1}} \qquad (n = 1, 2, \cdots) \qquad (8)$$

Seja $z_n = x_1 + x_2 + \cdots + x_n$. Como $x_k \in B_k$, nós temos $||x_k|| < 1/2^k$. Isso resulta $n > m$

$$\|z_n - z_m\| \leqq \sum_{k=m+1}^{n} \|x_k\| < \sum_{k=m+1}^{\infty} \frac{1}{2^k} \quad \longrightarrow \quad 0$$

quando $m \longrightarrow \infty$. Portanto (z_n) é de Cauchy. (z_n) converge, digamos, $z_n \longrightarrow x$ porque X é completo. Também $x \in B_0$ já que B_0 tem raio 1 e

$$\sum_{k=1}^{\infty} \|x_k\| < \sum_{k=1}^{\infty} \frac{1}{2^k} = 1 \qquad (9)$$

Como T é contínuo, $T(z_n) \longrightarrow T(x)$, e (8) mostra que $T(x) = y$. Por isso $y \in T(B_0)$. ■

Teorema 3.2.5 ***(Teorema da Aplicação Aberta, Teorema da Inversa Limitada).*** *Um operador linear limitado T sobrejetivo de um espaço de Banach X sobre um espaço de Banach Y é uma aplicação aberta. Portanto, se T é bijetivo, T^{-1} é contínuo e limitado.*

Demonstração:

Provaremos que para cada conjunto aberto $A \subset X$ a imagem $T(A)$ é aberta em Y. Isso fazemos mostrando que para cada $y = T(x) \in T(A)$ o conjunto $T(A)$ contém uma bola aberta em torno de $y = T(x)$.

Seja $y = T(x) \in T(A)$. Como A é aberto, ele contém uma bola aberta com centro x. Portanto, $A - x$ contém uma bola aberta com centro 0. Tome o raio da bola sendo r e defina $k = 1/r$, de modo que $r = 1/k$. Então $k(A - x)$ contém a bola unitária aberta $B(0; 1)$. O Lema 3.2.4 agora implica que $T(k(A-x)) = k[T(A)-T(x)]$ contém uma bola aberta em torno de 0, e também $T(A) - T(x)$. Portanto, $T(A)$ contém uma bola aberta em torno de $T(x) = y$. Como $y \in T(A)$ foi arbitrário, $T(A)$ é aberto.
Finalmente, se $T^{-1} : Y \longrightarrow X$ existe, é contínuo pelo Teorema 0.0.3 porque T é aberto. Como T^{-1} é linear pelo Teorema 1.5.3,e é limitado pelo Teorema 1.6.4. ■

3.3 Teorema do Gráfico Fechado

Nesta seção definimos operadores lineares fechados em espaços normados e consideraremos algumas de suas propriedades, em particular em conexão com o importante Teorema do Gráfico Fechado que afirma sobre suficientes condições para que um operador linear fechado em um espaço de Banach seja limitado.

Comecemos pela definição.

Definição 3.3.1 (Operador linear fechado). Sejam X e Y espaços normados e $T : D(T) \longrightarrow Y$ um operador linear com domínio $D(T) \in X$. Então T é chamado de operador linear fechado se seu gráfico

$$G(T) = \{(x,y)|x \in D(T), y = T(x)\}$$

é fechado no espaço normado $X \times Y$, onde as duas operações algébricas de um espaço vetorial em $X \times Y$ são definidas como de costume, isto é

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$$

$$\alpha(x,y) = (\alpha x, \alpha y)$$

(α um escalar) e a norma em $X \times Y$ é definida por

$$||(x,y)|| = ||x|| + ||y||. \qquad (1)$$

Sob quais condições um operador linear fechado será limitado? A resposta é dada pelo seguinte teorema:

Teorema 3.3.2 ***(Teorema do Gráfico Fechado).*** *Sejam X e Y espaços de Banach e $T : D(T) \longrightarrow Y$ um operador linear fechado, onde $D(T) \subset X$. Se $D(T)$ é fechado em X, então o operador T é limitado.*

Demonstração:

Primeiro mostraremos que $X \times Y$ com norma definida por (1) é completo. Seja (z_n) uma sequência de Cauchy em $X \times Y$, onde $z_n = (x_n, y_n)$. Então para cada $\varepsilon > 0$ existe um N tal que

$$||z_n - z_m|| = ||x_n - x_m|| + ||y_n - y_m|| < \varepsilon \qquad (2)$$

$(m, n > N)$. Como (x_n) e (y_n) são de Cauchy em X e Y, respectivamente, e convergem, digamos, $x_n \longrightarrow x$ e $y_n \longrightarrow y$, porque X e Y são completos. Isso implica que $z_n \longrightarrow z = (x,y)$ já que de (2) com $m \longrightarrow \infty$ temos $||z_n - z|| \leqslant \varepsilon$ para $n > N$. Como a sequência de Cauchy (z_n) foi arbitrária, $X \times Y$ é completo.

Por suposição, $G(T)$ é fechado em $X \times Y$ e $D(T)$ é fechado em X. Portanto, $G(T)$ e $D(T)$ são completos pelo teorema 0.0.10. Consideramos agora a aplicação

$$\begin{aligned} P :& G(T) \longrightarrow D(T) \\ & (x, T(x)) \longmapsto x. \end{aligned}$$

P é linear. P é limitado por causa de

$$||P(x, T(x))|| = ||x|| \leqslant ||x|| + ||T(x)|| = ||(x, T(x))||.$$

P é bijetiva. Sua aplicação inversa é

$$\begin{aligned} P^{-1} :& D(T) \longrightarrow G(T) \\ & x \longmapsto (x, T(x)). \end{aligned}$$

Como $G(T)$ e $D(T)$ são completos, podemos aplicar o teorema da inversa limitada 3.2.5 e veja que P^{-1} é limitado, digamos, $||(x, T(x))|| \leqslant b||x||$ para algum b e todo $x \in D(T)$. Portanto, T é limitado porque

$$||T(x)|| \leqslant ||T(x)|| + ||x|| = ||(x, T(x))|| \leqslant b||x||$$

para todo $x \in D(T)$. ■

Por definição, $G(T)$ é fechado se, e somente se, $z = (x, y) \in G(T)$ implica $z \in G(T)$. Do Teorema 0.0.9(a) vemos que $z \in G(T)$ se, e somente se, existem $z_n = (x_n, T(x_n)) \in G(T)$ tais que $z_n \longrightarrow z$, portanto

$$x_n \longrightarrow x, \qquad T(x_n) \longrightarrow y; \qquad (3)$$

e $z = (x, y) \in D(T)$ se, e somente se, $x \in D(T)$ e $y = T(x)$. Isso prova o seguinte critério, que expressa uma propriedade que muitas vezes é tomada como uma definição de fecho de um operador linear.

Teorema 3.3.3 ***(Operador linear fechado).*** *Seja $T : D(T) \longrightarrow Y$ um operador linear, onde $D(T) \subset X$ e X e Y são espaços normados. Então T é fechado se, e somente se, tem a seguinte propriedade. Se $x_n \longrightarrow x$, onde $x_n \in D(T)$ e $T(x_n) \longrightarrow y$, então $x \in D(T)$ e $T(x) = y$.*

Exemplo 3.3.1 (Operador diferencial). Seja $X = C[0, 1]$ e

$$\begin{aligned} T :& D(T) \longrightarrow X \\ & x \longmapsto x'. \end{aligned}$$

onde a linha denota diferenciação e $D(T)$ é o subespaço das funções $x \in X$ que têm uma derivada contínua. Então T não é limitado, mas é fechado.

Demonstração:

Vemos do exemplo 1.6.1 que T não é limitado. Nós provaremos que T é fechado pela aplicação do Teorema 3.3.3. Seja (x_n) em $D(T)$ tal que ambos (x_n) e $(T(x_n))$ convergem, digamos,

$$x_n \longrightarrow x \quad \text{e} \quad T(x_n) = x'_n \longrightarrow y$$

Como a convergência na norma de $C[0,1]$ é convergência uniforme em $[0,1]$, de $x'_n \longrightarrow y$ temos

$$\int_0^t y(\tau)d\tau = \int_0^t \lim_{n\to\infty} x'_n(\tau)d\tau = \lim_{n\to\infty}\int_0^t x'_n(\tau)d\tau = x(t) - x(0),$$

isto é,

$$x(t) = x(0) + \int_0^t y(\tau)d\tau.$$

Isso mostra que $x \in D(T)$ e $x' = y$. O Teorema 3.3.3 agora implica que T é fechado. ■

Vale a pena notar que neste exemplo, $D(T)$ não é fechado em X já que T seria então limitado pelo teorema do gráfico fechado. Isso motiva o seguinte resultado:

Proposição 3.3.4 *Fechamento não implica limitação de um operador linear. Por outro lado, limitação não implica fechamento.*

Demonstração:
A primeira afirmação é ilustrada pelo exemplo 3.3.1 e a segunda pelo exemplo a seguir. Seja $T : D(T) \longrightarrow D(T) \subset X$ o operador identidade em $D(T)$, onde $D(T)$ é um subespaço denso adequado de um espaço normado X. Então é trivial que T seja linear e limitado. No entanto, T não é fechado. Isso segue imediatamente do Teorema 3.3.3 se tomarmos um $x \in X - D(T)$ e uma sequência (x_n) em $D(T)$ que converge para x. ■

encerrando a seção, temos um resultado decorrente do Teorema 3.3.3.

Lema 3.3.5 ***(Operador fechado).*** *Seja $T : D(T) \longrightarrow Y$ um operador linear limitado com domínio $D(T) \subset X$, onde X e Y são espaços normados. Então:*

(a) Se $D(T)$ é um subconjunto fechado de X, então T é fechado.

(b) Se T é fechado e Y é completo, então $D(T)$ é um subconjunto fechado de X.

Demonstração:
(a) Se (x_n) está em $D(T)$ e converge, digamos, $x_n \longrightarrow x$, e é tal que $(T(x_n))$ também converge, então $x \in \overline{D(T)} = D(T)$ já que $D(T)$ é fechado, e $T(x_n) \longrightarrow T(x)$ já que T é contínuo. Portanto, T é fechado pelo Teorema 3.3.3.

(b) Para $x \in \overline{N(T)}$ existe uma sequência (x_n) em $D(T)$ tal que $x_n \longrightarrow x$. Como T é limitado,

$$\|T(x_n) - T(x_m)\| = \|T(x_n - x_m)\| \leqq \|T\|\,\|x_n - x_m\|.$$

Isso mostra que $(T(x_n))$ é de Cauchy. $(T(x_n))$ converge, digamos, $T(x_n) \longrightarrow y \in Y$ porque Y é completo. Como T é fechado, $x \in D(T)$ pelo teorema 3.3.3 (e $T(x) = y$). Portanto, $D(T)$ é fechado porque $x \in \overline{D(T)}$ era arbitrário. ■

3.4 Lema de Zorn

Precisaremos do lema de Zorn na prova do Teorema de Hahn Banach, que é um teorema de extensão para funcionais lineares e é importante por razões que iremos expor quando formularmos o teorema. O lema de Zorn tem várias aplicações. Dois deles serão mostrados mais adiante nesta seção. Inicialmente precisaremos da seguinte definição.

Definição 3.4.1 (Conjunto parcialmente ordenado). Um conjunto parcialmente ordenado é um conjunto M no qual se define uma ordenação parcial, ou seja, uma relação binária que é denotada por $\leqslant$ e satisfaz as condições

(PO1) $a \leqslant a$ para todo $a \in M$. (Reflexividade)

(PO2) se $a \leqslant b$ e $b \leqslant a$ então $a = b$. (antissimétrica)

(PO2) se $a \leqslant b$ e $b \leqslant c$ então $a \leqslant c$. (transitividade)

O nome parcialmente enfatiza que M pode conter elementos a e b para os quais nem $a \leqslant b$ nem $b \leqslant a$ são válidos. Então a e b são chamados elementos incomparáveis. Em contraste, dois elementos a e b são chamados comparáveis se satisfazem $a \leqslant b$ ou $b \leqslant a$ (ou ambos).

Um conjunto ou cadeia totalmente ordenado 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 possui elementos incomparáveis.

Um limite superior de um subconjunto W de um conjunto parcialmente ordenado M é um elemento $u \in M$ tal que

$$x \leqslant u \qquad \text{para todo} \quad x \in M.$$

(Dependendo de M e W, tal u prefeito pode não existir.) Um elemento máximo de M é um $m \in M$ tal que

$$m \leqslant x \qquad \text{implica} \quad m = x.$$

(Mais uma vez, M pode ou pode não ter elementos máximos. Observe ainda que um elemento máximo não precisa ser um limite superior.)

Exemplo 3.4.1 Números Reais. Seja M o conjunto de todos os números reais e seja $x \leqslant y$ com seu significado usual. M é totalmente ordenado. M não tem elemento máximo.

Exemplo 3.4.2 Conjunto das partes. Seja $P(X)$ o conjunto das partes (conjunto de todos os subconjuntos) de um dado o conjunto X e seja $A \leqslant B$ usado para indicar que $A \subset B$, ou seja, A é um subconjunto de B. Então $P(X)$ é parcialmente ordenado. O único elemento máximo de $P(X)$ é X.

Usando os conceitos definidos em 3.4.1, podemos agora formular o lema, que consideramos como um axioma.

Lema 3.4.2 ***(Lema de Zorn).*** *Seja $M \neq 0$ um conjunto parcialmente ordenado. Suponha que toda cadeia $C \subset M$ tem um limite superior. Então M tem pelo menos um elemento máximo.*

A seguir vamos usar o Lema de Zorn para obter interessantes resultados de Análise Funcional.

Teorema 3.4.3 ***(Base de Hamel).*** *Todo espaço vetorial $X \neq \{0\}$ tem uma base de Hamel.*

Demonstração:

Seja M o conjunto de todos os subconjuntos linearmente independentes de X. Como $X \neq \{0\}$, ele tem um elemento $x \neq 0$ e $x \in M$, de modo que $M \neq \emptyset$. Defina inclusão de conjuntos como uma ordenação parcial em M(ex. 3.4.2). Cada cadeia $C \subset M$ tem um limite superior, ou seja, a união de todos os subconjuntos de X que são elementos de C. Pelo lema de Zorn, M tem um elemento B maximal. mostraremos que B é uma base de Hamel para X. Seja $Y = spanB$. Então Y é um subespaço de X, e $Y = X$, caso contrário $B \cup z, z \in X, z \notin Y$, seria um conjunto linearmente independente contendo B como um subconjunto próprio, contrariando a maximalidade de B ■

Teorema 3.4.4 ***(Conjunto Ortonormal Total).*** *Em todo espaço de Hilbert $H \notin \{0\}$ existe um conjunto ortonormal total.*

Demonstração:

Seja M o conjunto de todos os subconjuntos ortonormais de H. Como $H \notin \{0\}$, tem um elemento $x \neq 0$, e um subconjunto ortonormal de H é $\{y\}$, onde $y = ||x||^{-1}x$. Portanto, $M \neq \emptyset$. A inclusão de conjuntos define uma ordenação parcial em M. Toda cadeia $C \subset M$ tem um limite superior, ou seja, a união de todos os subconjuntos de X que são elementos de C. Pelo lema de Zorn, M tem um elemento máximo F. Provamos que F é total em H. Suponha que isso seja falso. Então pelo Teorema 2.5.2 existe um $z \in H$ diferente de zero tal que $z \perp F$. Daí $F_1 = F \cup \{e\}$, onde $e = ||z||^{-1}z$, é ortonormal e F é um subconjunto próprio de F_1 Isso contradiz a maximalidade de F. ■

3.5 Teorema de Hahn-Banach

O teorema de Hahn-Banach é um teorema que trata de extensão para funcionais lineares. Veremos na próxima seção que o teorema garante que um espaço normado é ricamente fornecido com funcionais lineares limitados e torna possível uma teoria adequada dos espaços duais, que é uma parte essencial da teoria geral dos espaços normados. Desta forma o teorema de Hahn-Banach torna-se um dos mais importantes teoremas em conexão com operadores lineares limitados. Além disso, nossa discussão mostrará que o teorema também caracteriza até que ponto os valores de um funcional linear podem ser pré-atribuídos.

De um modo geral, em um problema de extensão considera-se um objeto matemático (por exemplo, uma aplicação) definida em um subconjunto Z de um conjunto X e se quer estender o objeto de Z para todo X de tal forma que certas propriedades básicas do objeto continuem para o objeto estendido.

No teorema de Hahn-Banach, o objeto a ser estendido é um funcional f que é definido em um subespaço Z de um espaço vetorial X e tem uma certa propriedade de limitação que será formulada em termos de um **funcional sublinear.** Por definição, este é um funcional de valor real p em um espaço vetorial X que é **subaditivo**, isto é,

$$p(x+y) \leqslant p(x)+p(y) \qquad \text{para todo} \quad x, y \in X. \tag{1}$$

e homogêneo positivo, ou seja,

$$p(\alpha x) = \alpha p(x) \qquad \text{para todo} \quad \alpha \geqslant 0 \in \mathbb{R}, x \in X. \tag{2}$$

Vamos supor que o funcional f a ser estendido é o maior em Z por tal funcional p definido em X, e estenderemos f de Z para X sem perder a linearidade e a majoração, de modo que o funcional estendido $\tilde{f}$ em X ainda é linear e ainda majorado por p. A seguir enunciamos o Teorema de Hahn-Banach no caso em que X é um espaço vetorial real.

Teorema 3.5.1 ***(Teorema de Hahn-Banach).*** *Seja X um espaço vetorial real e p um funcional sublinear em X. Além disso, seja f um funcional linear que é definido em um subespaço Z de X e satisfaz*

$$f(x) \leqslant p(x) \qquad \textit{para todo} \quad x \in Z. \tag{3}$$

Então f tem uma extensão linear $\tilde{f}$ de Z a X satisfazendo

$$\tilde{f}(x) \leqslant p(x) \qquad \textit{para todo} \quad x \in Z. \tag{3*}$$

isto é, $\tilde{f}$ é um funcional linear em X, satisfaz (3^*) em X e $\tilde{f}(x) = f(x)$ para todo $x \in Z$.

Demonstração:

Procedendo passo a passo, provaremos:

(a) O conjunto E de todas as extensões lineares g de f satisfazendo $g(x) \leqslant p(x)$ em seu domínio $D(g)$ pode ser parcialmente ordenado e o lema de Zorn produz um elemento máximo $\tilde{f}$ de E.

(b) $\tilde{f}$ é definida no espaço X inteiro.

(c) Uma relação auxiliar que foi usada em (b).

Seguem os detalhes da prova:

(a) Seja E o conjunto de todas as extensões lineares g de f que satisfazem a condição

$$g(x) \leqslant p(x) \qquad \text{para todo} \quad x \in D(g),$$

Claramente, $E \neq \emptyset$ já que $f \in E$. Em E podemos definir uma ordenação parcial por $g \leqslant h$ significando que h é uma extensão de g, isto é, por definição, $D(h) \supset D(g)$ e $h(x) = g(x)$ para todo $x \in D(g)$.

Para qualquer cadeia $C \subset E$ definimos g por

$$\hat{g}(x) = g(x), \quad x \in D(g) \qquad (g \in C).$$

$\hat{g}$ é um funcional linear, sendo o domínio

$$D(\hat{g}) = \bigcup_{\mathbf{g} \in C} D(g),$$

que é um espaço vetorial, pois C é uma cadeia. A definição de g é não-ambígua. De fato, para um $x \in D(g_1) \cap D(g_2)$ com $g_1, g_2 \in C$ temos $g_1(x) = g_2(x)$ já que C é uma cadeia, então $g_1 \leqslant g_2$ ou $g_2 \leqslant g_1$. Claramente, $g \leqslant \hat{g}$ para todo $g \in C$. Portanto, $\hat{g}$ é um limite superior de C. Como $C \subset E$ foi arbitrário, o Lema de Zorn implica que E tem um elemento maximal $\tilde{f}$. Pela definição de E, esta é uma extensão linear de f que satisfaz

$$\tilde{f}(x) \leqslant p(x) \qquad x \in D(\tilde{f}). \qquad (4)$$

(b) Mostraremos agora que $D(\tilde{f})$ é todo X. Suponha que isso seja falso. Então podemos escolher um $y_1 \in X - D(\tilde{f})$ e considerar o subespaço Y_1 de X gerado por $D(\tilde{f})$ e y_1. Observe que $y_1 \neq 0$ desde $0 \in D(\tilde{f})$. Algum $x \in Y_1$ pode ser escrito

$$x = y + \alpha y_1 \qquad y \in D(\tilde{f}).$$

Esta representação é única. De fato, $y + \alpha y_1 = \tilde{y} + \beta y_1$ com $\tilde{y} \in D(\tilde{f})$ implica $y - \tilde{y} = (\beta - \alpha) y_1$ onde $y - \tilde{y} \in D(\tilde{f})$ enquanto $y_1 \notin D(\tilde{f})$, de modo que a única solução é $y - y_1 = 0$ e $\beta - \alpha = 0$. Isso prova a unicidade.
Um funcional g_1 em Y_1 é definido por

$$g_1(y + \alpha y_1) = \tilde{f}(y) + \alpha c \qquad (5)$$

onde c é qualquer constante real. Não é difícil ver que g_1 é linear. Além disso, para $\alpha = 0$ temos $g_1(y) = \tilde{f}(y)$. Portanto, g_1 é uma extensão de $\tilde{f}$, isto é, uma extensão tal que $D(\tilde{f})$ é um subconjunto próprio de $D(g_1)$. Consequentemente, se pudermos provar que $g_1 \in E$ mostrando que

$$g_1(x) \leqslant p(x) \qquad \text{para todo} \quad x \in D(g_1), \qquad (6)$$

isso irá contradizer a maximalidade de $\tilde{f}$, de modo que $D(\tilde{f}) \neq X$ é falso e $D(\tilde{f}) = X$ é verdadeiro.

(c) Assim, devemos finalmente mostrar que g_1 com um c adequado em (5) satisfaz (6). Consideramos qualquer y e z em $D(\tilde{f})$. De (4) e (1) obtemos

$$\begin{aligned}\tilde{f}(y) - \tilde{f}(z) = \tilde{f}(y-z) &\leqq p(y-z)\\ &= p\left(y + y_1 - y_1 - z\right)\\ &\leqq p\left(y + y_1\right) + p\left(-y_1 - z\right)\end{aligned}$$

Jogando o primeiro termo à direta para à esquerda e o último termo da esquerda para à direita, temos

$$-p\left(-y_1 - z\right) - \tilde{f}(z) \leqq p\left(y + y_1\right) - \tilde{f}(y) \qquad (7)$$

onde y_1 é fixo. Como y não aparece à esquerda e z não aparece à direita, a desigualdade continua valendo se tomarmos o supremo sobre $z \in D(\tilde{f})$ à esquerda (chame-o de m_0) e o ínfimo sobre $y \in D(\tilde{f})$ à direita (chame de m_1). Temos $m_0 \leqslant m_1$ e para um c com $m_0 \leqslant c \leqslant m_1$ temos de (7)

$$-p\left(-y_1 - z\right) - \tilde{f}(z) \leqq c \quad \text{para todo} \quad z \in D(\tilde{f}) \qquad (8a)$$

$$c \leqq p\left(y + y_1\right) - \tilde{f}(y) \quad \text{para todo} \quad y \in D(\tilde{f}). \qquad (8b)$$

Provaremos (6) primeiro para α negativo em (5) e depois para α positivo. Para $\alpha < 0$ usamos (8a) com z substituído por $\alpha^{-1}y$, ou seja,

$$-p\left(-y_1 - \frac{1}{\alpha}y\right) - \tilde{f}\left(\frac{1}{\alpha}y\right) \leqq c.$$

Multiplicando por $-\alpha > 0$ temos

$$\alpha p\left(-y_1 - \frac{1}{\alpha}y\right) + \tilde{f}(y) \leqq -\alpha c.$$

A partir disso e (5), usando $y + \alpha y_1 = x$ (veja acima), obtemos a desigualdade

$$g_1(x) = \tilde{f}(y) + \alpha c \leqq -\alpha p\left(-y_1 - \frac{1}{\alpha}y\right) = p\left(\alpha y_1 + y\right) = p(x)$$

Para $\alpha = 0$ temos $x \in D(\tilde{f})$ e não temos nada a provar. Para $\alpha > 0$ usamos (8b) com y substituído por $\alpha^{-1}y$ para obter

$$c \leqq p\left(\frac{1}{\alpha}y + y_1\right) - \tilde{f}\left(\frac{1}{\alpha}y\right)$$

Multiplicando por $\alpha > 0$ temos

$$\alpha c \leqq \alpha p\left(\frac{1}{\alpha}y + y_1\right) - \tilde{f}(y) = p(x) - \tilde{f}(y)$$

A partir disso e (5),

$$g_1(x) = \tilde{f}(y) + \alpha c \leqq p(x)$$

Dessa forma provamos a equação (6) e o teorema fica demonstrado. ■

O teorema de Hahn-Banach 3.5.1 diz respeito a espaços vetoriais reais. Uma generalização que inclui espaços vetoriais complexos será feita a seguir.

Teorema 3.5.2 ***(Teorema de Hahn-Banach Generalizado).*** *Seja X um espaço vetorial real ou complexo e p um funcional de valor real em X que é subaditivo, isto é, para todo $x, y \in X$,*

$$p(x+y) \leqslant p(x) + p(y) \qquad (1)$$

e para todo escalar α satisfaz

$$p(\alpha x) = |\alpha| p(x). \qquad (2)$$

Além disso, seja f um funcional linear definido em um subespaço Z de X que satisfaz

$$|f(x)| \leqslant p(x) \qquad \textit{para todo} \quad x \in Z. \qquad (3)$$

Então f tem uma extensão linear $\tilde{f}$ de Z a X satisfazendo

$$|\tilde{f}(x)| \leqslant p(x) \qquad \textit{para todo} \quad x \in X. \qquad (3^*)$$

Demonstração:

(a) Espaço vetorial real. Se X é real, a situação é simples. Temos que (3) implica $f(x) \leqslant p(x)$ para todo $x \in Z$. Assim, pelo Teorema de Hahn-Banach 3.5.1 existe uma extensão linear $\tilde{f}$ de Z a X tal que

$$\tilde{f}(x) \leqslant p(x) \qquad \text{para todo} \quad x \in X. \qquad (4)$$

A partir disso e (2) obtemos

$$-\tilde{f}(x) = \tilde{f}(-x) \leqq p(-x) = |-1| p(x) = p(x)$$

isto é, $\tilde{f}(x) \geqslant -p(x)$. Juntamente com (4) isso prova (3^*).

(b) Espaço vetorial complexo. Seja X complexo. Então Z é um espaço vetorial complexo também. Portanto, f é de valor complexo, e podemos escrever

$$f(x) = f_1(x) + i f_2(x) \qquad x \in Z$$

onde f_1 e f_2 são reais. Por um momento, consideramos X e Z como espaços vetoriais reais e denotaremos por X_r e Z_r, respectivamente; isto simplesmente significa que restringimos a multiplicação por escalares a números reais (em vez de números complexos). Como f é linear em Z e f_1 e f_2 são reais, f_1 e f_2 são funcionais lineares em Z_r. Também

$f_1(x) \leqslant |f(x)|$ porque a parte real de um número complexo não pode exceder o valor absoluto do mesmo. Portanto, por (3),

$$f_1(x) \leqslant p(x) \qquad \text{para todo} \quad x \in Z_r.$$

Pelo teorema de Hahn-Banach 3.5.1 existe uma extensão linear $\tilde{f}_1$ de f_1, de Z_r a X_r tal que

$$\tilde{f}_1(x) \leqslant p(x) \qquad \text{para todo} \quad x \in X_r. \qquad (5)$$

Isso resolve o caso de f_1 e agora nos voltamos para f_2. Retornando a Z e usando $f = f_1 + if_2$, temos para todo $x \in Z$

$$i\left[f_1(x) + if_2(x)\right] = if(x) = f(ix) = f_1(ix) + if_2(ix)$$

As partes reais de ambos os lados devem ser iguais:

$$f_2(x) = -f_1(ix) \qquad x \in Z. \qquad (6)$$

Portanto, se para todo $x \in X$ definirmos

$$\tilde{f}(x) = \tilde{f}_1(x) - i\tilde{f}_1(ix) \qquad x \in X, \qquad (7)$$

vemos em (6) que $\tilde{f}(x) = f(x)$ em Z. Isso mostra que $\tilde{f}$ é uma extensão de f de Z para X. Nossa tarefa restante é provar que

(i)) $\tilde{f}$ é um funcional linear no espaço vetorial complexo X,

(ii) $\tilde{f}$ satisfaz (3^*) em X

Que (i) vale pode ser visto a partir do seguinte cálculo que usa (7) e a linearidade de $\tilde{f}_1$ no espaço vetorial real X_r; aqui $a + ib$ é complexo e a e b é qualquer escalar real,

$$\begin{aligned}
\tilde{f}((a+ib)x) &= \tilde{f}_1(ax + ibx) - i\tilde{f}_1(iax - bx) \\
&= a\tilde{f}_1(x) + b\tilde{f}_1(ix) - i\left[a\tilde{f}_1(ix) - b\tilde{f}_1(x)\right] \\
&= (a+ib)\left[\tilde{f}_1(x) - i\tilde{f}_1(ix)\right] \\
&= (a+ib)\tilde{f}(x).
\end{aligned}$$

Resta provar (ii). Para qualquer x tal que $\tilde{f}(x) = 0$ isto é válido desde que $p(x) \geqslant 0$ por (1) e (2). Seja x tal que $\tilde{f}(x) \neq 0$. Então podemos escrever, usando a forma polar de números complexos,

$$\tilde{f}(x) = |\tilde{f}(x)|e^{i\theta}, \quad \text{portanto} \quad |\tilde{f}(x)| = \tilde{f}(x)e^{-i\theta} = \tilde{f}\left(e^{-i\theta}x\right)$$

Como $|\tilde{f}(x)|$ é real, a última expressão é real e, portanto, igual a sua parte real. Portanto, por (2),

$$|\tilde{f}(x)| = \tilde{f}\left(e^{-i\theta}x\right) = \tilde{f}_1\left(e^{-i\theta}x\right) \leqq p\left(e^{-i\theta}x\right) = \left|e^{-i\theta}\right| p(x) = p(x).$$

Isso completa a prova do teorema. ■

Embora o teorema de Hahn-Banach não diga nada diretamente sobre continuidade, uma aplicação principal do teorema lida com funcionais lineares. Isso nos traz de volta aos espaços normalizados, que é a nossa principal preocupação. De fato, o Teorema 3.5.2 implica no seguinte resultado.

Teorema 3.5.3 ***(Teorema de Hahn-Banach para espaços normados).*** *Seja f um funcional linear limitado em um subespaço Z de um espaço normalizado X. Então existe um funcional linear limitado $\tilde{f}$ em X que é uma extensão de f para X e tem a mesma norma,*

$$||\tilde{f}||_X = ||f||_Z \qquad (8)$$

Onde

$$\|\tilde{f}\|_X = \sup_{\substack{x\in X\\ \|x\|=1}} |\tilde{f}(x)|, \quad \|f\|_Z = \sup_{\substack{x\in Z\\ \|x\|=1}} |f(x)|$$

e $||f||_Z = 0$ no caso trivial onde $Z = \{0\}$.

Demonstração:

Se $Z = \{0\}$, então $f = 0$ e sua extensão é $\tilde{f} = 0$. Seja $Z \neq \{0\}$. Queremos usar o Teorema 3.5.2. Portanto, devemos primeiro descobrir um p adequado. Para todo $x \in Z$ temos

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

Que se encaixa na forma (3), onde

$$p(x) = ||f||_Z||x||. \qquad (9)$$

Vemos que p é definido em todos os pontos de X. Além disso, p satisfaz (1) em X pois pela desigualdade triangular,

$$p(x+y) = \|f\|_z\|x+y\| \leqq \|f\|_z(\|x\| + \|y\|) = p(x) + p(y).$$

p também satisfaz (2) em X pois

$$p(\alpha x) = \|f\|_z\|\alpha x\| = |\alpha|\|f\|_z\|x\| = |\alpha|p(x).$$

Portanto, podemos agora aplicar o Teorema 3.5.2 e concluir que existe um funcional linear $\tilde{f}$ em X que é uma extensão de f e satisfaz

$$|\tilde{f}(x)| \leqq p(x) = \|f\|_z\|x\| \qquad x \in X$$

Tomando o supremo sobre todos $x \in X$ de norma 1, obtemos a desigualdade

$$\|\tilde{f}\|_X = \sup_{\substack{x\in X\\ \|x\|=1}} |\tilde{f}(x)| \leqq \|f\|_Z$$

Como sob uma extensão a norma não pode diminuir, também temos $||\tilde{f}||_X \geqslant ||f||_Z$. Juntos obtemos (8) e o teorema é provado. ∎

Em casos especiais a situação pode se tornar muito simples. Espaços de Hilbert são um exemplo desse tipo. De fato, se Z é um subespaço fechado de um espaço de Hilbert $X = H$, então f possui uma representação de Riesz pelo teorema 3.1.1, digamos,

$$f(x) = \langle x, z \rangle \qquad z \in Z.$$

onde $||z|| = ||f||$. É claro que, como o produto interno é definido em todos os pontos de H, isso dá imediatamente uma extensão linear $\tilde{f}$ de f de Z a H, e $\tilde{f}$ tem a mesma norma que f porque $||\tilde{f}|| = ||z|| = ||f||$ pelo Teorema 3.1.1. Daí neste caso a extensão é imediata.

Do Teorema 3.5.3, derivaremos agora outro resultado útil, envolvendo funcionais lineares limitados.

Teorema 3.5.4 ***(Funcionais lineares limitados).*** *Seja X um espaço normalizado e seja $x_0 \neq 0$ qualquer elemento de X. Então existe um funcional linear limitado $\tilde{f}$ em X tal que*

$$||\tilde{f}|| = 1, \qquad \tilde{f}(x_0) = ||x_0||.$$

Demonstração:

Consideramos o subespaço Z de X consistindo de todos os elementos $x = \alpha x_0$ onde α é um escalar. Em Z definimos um funcional linear f por

$$f(x) = f(\alpha x_0) = \alpha ||x_0||.$$

f é limitado e tem norma $||f|| = 1$ porque

$$|f(x)| = |f(\alpha x_0)| = |\alpha| ||x_0|| = ||\alpha x_0|| = ||x||.$$

O Teorema 3.5.3 implica que f tem uma extensão linear $\tilde{f}$ de Z a X, de norma $||\tilde{f}|| = ||f|| = 1$. De (10) vemos que $\tilde{f}(x_0) = f(x_0) = ||x_0||$. ∎

Segue do teorema 3.5.4 o seguinte corolário:

Corolário 3.5.5 *Para cada x em um espaço normado X temos*

$$\|x\| = \sup_{\substack{f \in X' \\ f \neq 0}} \frac{|f(x)|}{\|f\|}$$

Portanto, se x_0 é tal que $f(x_0) = 0$ para todo $f \in X'$, então $x_0 = 0$.

Demonstração:

Do teorema 3.5.4 temos, escrevendo x para xo,

$$\sup_{\substack{f \in X' \\ f \neq 0}} \frac{|f(x)|}{\|f\|} \geqq \frac{|\tilde{f}(x)|}{\|\tilde{f}\|} = \frac{\|x\|}{1} = \|x\|,$$

e de $|f(x)| \leqslant ||f||||x||$ obtemos

$$\sup_{\substack{f \in X' \\ f \neq 0}} \frac{|f(x)|}{\|f\|} \leqq \|x\|.$$

■

3.6 Teorema de Representação de Riesz para funcionais limitados em C[a,b]

O Teorema de Hahn-Banach 3.5.3 tem muitas aplicações importantes. Uma delas foi considerado na seção anterior. Outro vai ser apresentado nesta seção. De fato, usaremos o Teorema 3.5.3 para obtenção de uma fórmula geral de representação para funcionais lineares limitados em $C[a,b]$, onde $[a,b]$ é um intervalo compacto fixo. No presente caso a representação será em termos de uma integral de Riemann-Stieltjes. Vamos recordar a definição e algumas propriedades desta integral, que é uma generalização da integral de Riemann familiar. Começamos com o seguinte conceito.

Definição 3.6.1 A função $w : [a,b] \longrightarrow \mathbb{R}$ é dita ser de **variação limitada** em $[a,b]$ se a sua *variação total* $Var(w)$ em $[a,b]$ é finita, onde

$$Var(w) = \sup \sum_{j=1}^{n} |w(t_j) - w(t_{j-1})|, \qquad (1)$$

o supremo é tomado sobre todas as partições

$$a = t_0 < t_1 < ... < t_n = b \qquad (2)$$

do intervalo $[a,b]$; aqui $n \in \mathbb{N}$ é arbitrário e também é a escolha dos valores $t_1, ..., t_{n-1}$. em $[a,b]$ e no entanto, deve satisfazer a condição (2).

Definição 3.6.2 Uma norma no espaço vetorial de todas as funções de variação limitada em $[a,b]$ é definida por

$$\|w\| = |w(a)| + Var(w). \qquad (3)$$

O espaço assim definido é denotado por $BV[a,b]$.

Seja $x \in C[a,b]$ e $w \in BV[a,b]$. Seja P_n uma partição de $[a,b]$ dada pela relação (2) e seja

$$\eta(P_n) = max(t_1 - t_0, ..., t_n - t_{n-1})$$

para cada partição P_n de $[a,b]$ nos consideramos a soma

$$s(P_n) = \sum_{j=1}^{n} x(t_j)[w(t_j) - w(t_{j-1})]. \qquad (4)$$

Existe um numero ρ com a propriedade de que para cada $\epsilon > 0, \exists \delta > 0$ tal que

$$\eta(P_n) < \delta \qquad (5)$$

implica

$$|\rho - s(P_n)| < \epsilon. \qquad (6)$$

ρ é chamada de integral de Riemann-Stieltjes de x sobre o intervalo $[a,b]$ com respeito a w e é denotada por

$$\int_a^b x(t)dw(t). \qquad (7)$$

Podemos obter (7) como o limite das somas (4) para uma sequência (P_n) de partições de $[a,b]$ satisfazendo $\eta(P_n) \longrightarrow 0$ quando $n \longrightarrow \infty$.

Sejam $x_1, x_2 \in C[a,b]$ e escalares α e β e $w, w_1, w_2 \in BV[a,b]$. Valem:

$$\int_a^b [\alpha x_1(t) + \beta x_2(t)]dw(t) = \alpha \int_a^b x_1(t)dw(t) + \beta \int_a^b x_2(t)dw(t)$$

e

$$\int_a^b x(t)d(\alpha w_1 + \beta w_2)(t) = \alpha \int_a^b x(t)dw_1(t) + \beta \int_a^b x(t)dw_2(t).$$

e

$$|\int_a^b x(t)dw(t)| \leqslant \max_{t \in [a,b]} |x(t)| Var(w). \qquad (8)$$

Teorema 3.6.3 ***(Teorema de Riesz para Funcionais em C[a,b])***

Todo funcional linear limitado f em $C[a,b]$ pode ser representado por uma integral de Riemann-Stieltjes

$$f(x) = \int_a^b x(t)dw(t) \qquad (9)$$

onde w é de variação limitada em $[a,b]$ e tem variação total

$$Var(w) = \|f\| \qquad (10)$$

Demonstração:

Do Teorema de Hahn-Banach para espaços normados, vemos que f *tem uma extensão* $\tilde{f}$ *de* $C[a,b]$ *para o espaço normado* $B[a,b]$ *consistindo de todas as funções limitadas em* $C[a,b]$ *com norma definida por*

$$\|x\| = \sup_{t \in [a,b]} |x(t)|.$$

Além disso, o funcional linear $\tilde{f}$ *é limitado e tem a mesma norma que* f*, isto é,*

$$\|\tilde{f}\| = \|f\|.$$

Seja x_t *a função característica do intervalo* $[a,t]$*. Definimos* w *por*

$$w(a) = 0 \qquad w(t) = \tilde{f}(x_t), \qquad t \in (a,b].$$

Seja $\zeta \in \mathbb{C}$*, então* $\zeta = |\zeta| e(\zeta)$ *onde* $e(\zeta) = \begin{cases} 1, \text{ se } \zeta = 0 \\ e^{i\theta}, \text{ se } \zeta \neq 0. \end{cases}$

Vemos que se $\zeta \neq 0$ *então* $|\zeta| = \zeta e^{-i\theta}$*. Portanto, qualquer* ζ*, zero ou não, temos*

$$|\zeta| = \zeta \overline{e(\zeta)}, \qquad (11)$$

onde a barra indica conjugação complexa.

Seja $\epsilon_j = \overline{e(w(t_j) - w(t_{j-1}))}$ *e seja* $x_{t_j} = x_j$*, dessa forma por (11) obtemos*

$$\begin{aligned} \sum_{j=1}^{n} |w(t_j) - w(t_{j-1})| &= |\tilde{f}(x_1)| + \sum_{j=2}^{n} |\tilde{f}(x_j) - \tilde{f}(x_{j-1})| \\ &= \epsilon_1 \tilde{f}(x_1) + \sum_{j=2}^{n} \epsilon_j [\tilde{f}(x_j) - \tilde{f}(x_{j-1})] \\ &= \tilde{f}\left(\epsilon_1 x_1 + \sum_{j=2}^{n} \epsilon_j [(x_j) - (x_{j-1})]\right) \\ &\leqslant \|\tilde{f}\| \left\| \epsilon_1 x_1 + \sum_{j=2}^{n} \epsilon_j [(x_j) - (x_{j-1})] \right\|. \end{aligned}$$

No lado direito, $\|\tilde{f}\| = \|f\|$ *e o outro fator* $\left\|\epsilon_1 x_1 + \sum_{j=2}^{n} \epsilon_j [(x_j) - (x_{j-1})]\right\| = 1$*, pois* $|\epsilon_j| = 1$ *e da definição dos* x_j*'s vemos que para cada* $t \in [a,b]$ *somente um dos termos* $x_1, x_2 - x_1, ...,$ *não é zero (e sua norma é 1). Do lado esquerdo, tomamos o supremo sobre todas as partições de* $[a,b]$*. Então temos*

$$Var(w) \leqslant \|f\|. \qquad (12)$$

Portanto w *é de variação limitada em* $[a,b]$*.*

Para cada partição P_n da forma (2) definimos uma função, que denotamos simplesmente por z_n ao invés de $z(P_n)$, por

$$z_n = x(t_0)x_1 + \sum_{j=2}^{n} x(t_{j-1})[x_j - x_{j-1}]. \qquad (13)$$

Então $z_n \in B[a,b]$. Pela definição de w,

$$\begin{aligned} \tilde{f}(z_n) &= x(t_0)\tilde{f}(x_1) + \sum_{j=2}^{n} x(t_{j-1})[\tilde{f}(x_j) - \tilde{f}(x_{j-1})] \\ &= x(t_0)w(t_1) + \sum_{j=2}^{n} x(t_{j-1})[w(t_j) - w(t_{j-1})] \qquad (14) \\ &= \sum_{j=1}^{n} x(t_{j-1})[w(t_j) - w(t_{j-1})], \end{aligned}$$

onde a ultima igualdade segue de $w(t_0) = w(a) = 0$. Agora escolhemos qualquer sequência (P_n) de partições do intervalo $[a,b]$ tal que $\eta(P_n) \longrightarrow 0$; quando $n \longrightarrow \infty$, a soma no lado direito de (14) se aproxima da integral da equação (9), e (9) segue, desde que $\tilde{f}(z_n) \longrightarrow \tilde{f}(x)$, que é igual à $f(x)$ já que $x \in C[a,b]$.

Vamos mostrar que $\tilde{f}(z_n) \longrightarrow \tilde{f}(x)$. Lembrando a definição de x_t, vemos que (13) nos fornece $z_n(a) = x(a)1$ já que a soma em (13) é nula em $t = a$. Portanto $z_n(a) - x(a) = 0$. Além disso, por (13), se $t_{j-1} < t \leqslant t_j$, então obtemos $z_n(t) = x(t_{j-1})1$. Segue daí que para este "t",

$$|z_n(t) - x(t)| = |x(t_{j-1}) - x(t)|.$$

Consequentemente, se $\eta(P_n) \longrightarrow 0$, então $\|z_n - x\| \longrightarrow 0$ porque x é contínua em $[a,b]$, portanto uniformemente contínua no intervalo $[a,b]$, já que $[a,b]$ é compacto. A continuidade de $\tilde{f}$ implica que $\tilde{f}(z_n) \longrightarrow \tilde{f}(x)$, e $\tilde{f} = f(x)$, então (9) está provado.

De (9) e (8) temos

$$|f(x)| \leqslant \max_{t \in [a,b]} |x(t)| Var(w) = \|x\| Var(w). \qquad (15)$$

Tomamos o supremo sobre todos os $x \in C[a,b]$ de norma 1, obtemos $\|f\| \leqslant Var(w)$. Junto com (12) isso fornece (10). ■

Referências Bibliográficas

[1] BOTELHO, G; PELLEGRINO, D; TEIXEIRA, E. **Fundamentos de Análise Funcional.** 2º edição. Rio de Janeiro: SBM, 2015.

[2] CALLIOLI, A; DOMINGUES, H; COSTA, R. **Álgebra Linear e Aplicações.** 6º edição. Editora Atual, 2009.

[3] KREYSZIG, E. **Introductory Functional Analysis with Applications.** 2º edição. Toronto: John Wiley Sons, 1978.

[4] LIMA, E. **Espaços Métricos** 5º edição. Rio de Janeiro: IMPA, 2017.

[5] LIMA, E. **Álgebra Linear.** 9º edição. Rio de Janeiro: IMPA, 2018.

[6] LIMA, E. **Análise Real Volume 1: Funções de Uma Variável** 12º edição. Rio de Janeiro: IMPA, 2018.

[7] LIMA, E. **Curso de Análise Vol. 2.** 11º edição. Rio de Janeiro: IMPA, 2018.

[8] LIMA, E. **Curso de Análise Vol. 1.** 15º edição. Rio de Janeiro: IMPA, 2019.

[9] OLIVEIRA, C. **Introdução à Análise Funcional.** 1º edição. Rio de Janeiro: IMPA, 2014.