Universidade Federal do Paraná

Gabriel Alves de Lima

Decomposição em valores singulares de operadores compactos em espaços de Hilbert

Curitiba
2021

Gabriel Alves de Lima

Decomposição em valores singulares de operadores compactos em espaços de Hilbert

Trabalho de conclusão apresentado ao curso de Licenciatura em Matemática da Universidade Federal do Paraná como requisito parcial à obtenção do grau de licenciado em Matemática.

Orientador: Prof. Dr. Cleber de Medeira

Curitiba

2021

Agradecimentos

Primeiramente gostaria de dedicar este trabalho à Deus, meu criador e guia.

À minha família, que me apoiou, apoia e me apoiará em todas as minhas escolhas. Em especial a minha mãe, que está sempre disposta a me ajudar e sem ela minha caminhada teria sido muito diferente.

Aos meus "antigos" amigos, que apesar de não termos mais tanto contato, ainda se fazem presente na minha vida e estão sempre dispostos a compartilhar bons momentos.

Aos meus "novos" amigos, que vivenciaram boa parte dessa jornada comigo, tanto no curso, quanto em momentos descontraídos, livres de preocupações acadêmicas.

Ao PET Matemática e todos seus participantes, que me proporcionaram um aprendizado único na graduação.

Aos professores do Departamento de Matemática da UFPR, que foram essenciais para meu crescimento, em especial ao meu orientador, que me guiou neste trabalho, que a princípio parecia impossível.

Por fim, gostaria de agradecer a banca examinadora e todos os demais que estiveram comigo de alguma forma.

"Hard work beats talent when talent fails to work hard."

– Kevin Durant

Resumo

Em espaços vetoriais de dimensão finita, é bem conhecida a decomposição em valores singulares (SVD) de operadores lineares. De modo geral, essa “boa propriedade” pode ser generalizada para operadores lineares compactos em espaços de Hilbert. Provamos esse resultado com o aporte da teoria básica de espaços funcionais e de operadores lineares. Por fim, estudamos algumas aplicações onde esse tipo de decomposição é bastante útil.

Palavras-chave: Espaço de Hilbert, operador linear compacto, decomposição em valores singulares.

Abstract

In finite-dimensional vector spaces, it is well known the singular value decomposition (SVD) for linear operators. In general, this “good property” can be generalized to compact linear operators in Hilbert spaces. We prove this result by using the basic theory of functional spaces and linear operators. Lastly, we study applications for which this decomposition is quite useful.

Keywords: Hilbert space, compact linear operator, singular value decomposition.

Sumário

Introdução

A Análise Funcional é uma área da Matemática relativamente nova, com avanços significativos a partir do fim do século XIX que contou com nomes importantes como David Hilbert, Stefan Banach, Henry Poincaré, Frigyes Riesz, Maurice Frechét, dentre outros. Uma característica desta área é o estudo de operadores em *espaços funcionais*. Em particular destacamos os operadores compactos e veremos que essa classe de operadores tem propriedades interessantes, notavelmente a decomposição em valores singulares.

No Capítulo 1 estabelecemos a teoria necessária para desenvolver os resultados principais do trabalho. Mais precisamente, estudamos noções de espaços métricos, espaços normados, espaços de Banach e operadores lineares limitados. Veremos que, diferente do que é observado na Álgebra Linear, muitas propriedades que são garantidas em espaços vetoriais de dimensão finita nem sempre são verdadeiras num contexto mais geral. Também apresentamos uma introdução à teoria espectral de operadores lineares.

O segundo capítulo trata de espaços com produto interno. Uma das principais características desses espaços é a ideia de ortogonalidade que recupera várias noções geométricas conhecidas dos espaços euclidianos. Apresentamos uma série de resultados sobre esses espaços que serão de extrema importância para o estudo sobre operadores lineares compactos que veremos no último capítulo. No caso em que o espaço com produto interno é completo, em relação a métrica induzida, o chamamos de *espaço de Hilbert*. Por fim, fazemos uma breve discussão sobre *convergência fraca* em espaços normados.

Os dois primeiros capítulos constituem uma boa parte da teoria estudada inicialmente em um curso de Análise Funcional, sendo nossas principais referências os livros [1, 8] e [4]. No Capítulo 3 teremos como objeto principal os operadores lineares compactos em espaços de Hilbert. A principal referência para essa parte do trabalho é o livro [11]. Iniciamos considerando operadores compactos de forma geral e posteriormente, em espaços de Hilbert, apresentamos um estudo sistemático sobre a decomposição desses operadores em termos de seus valores singulares.

Finalizamos o trabalho aplicando a decomposição em valores singulares na obtenção de alguns resultados, em especial uma caracterização para os operadores de Hilbert-Schmidt e também para a classe Traço, onde usamos como principal referência a dissertação [9].

Capítulo 1

Espaços métricos e normados

Neste capítulo faremos um apanhado geral sobre espaços métricos, apresentando definições e alguns resultados importantes para o desenvolvimento desse trabalho. Posteriormente discutiremos espaços normados e espaços de Banach.

1.1 Espaços métricos

Um *espaço métrico* é um par (X, d) onde X é um conjunto e d uma função

$$\begin{aligned} d : X \times X &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto d(x, y), \end{aligned}$$

denominada *métrica* (ou *distância* em X), que satisfaz para quaisquer x, y e z em X as seguintes propriedades:

(M1) d é não-negativa;

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

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

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

De forma geral, segue de (M4) que para quaisquer $x_1, \ldots, x_n$ em X vale

$$d(x_1, x_n) \leq d(x_1, x_2) + d(x_2, x_3) + \cdots + d(x_{n-1}, x_n).$$

Se $Y \subset X$, então podemos induzir em Y uma métrica considerando a restrição $\bar{d} \doteq d|_{Y \times Y}$. Dessa forma, $(Y, \bar{d})$ é um subespaço métrico, sendo $\bar{d}$ a métrica induzida em Y por d.

Muitas vezes, quando ficar subentendido qual a métrica envolvida e para não sobrecarregar a notação, vamos nos referir ao espaço métrico (X, d) simplesmente por X.

1.1.1 Exemplos de espaços métricos

A seguir apresentamos alguns exemplos clássicos de espaços métricos. Nesse exemplos as propriedades (M1), (M2) e (M3) são simples de serem verificadas, sendo mais complicada, em geral, a desigualdade triangular (M4). A demonstração dessa propriedade para os exemplos a seguir encontra-se no Capítulo 1 da referência [4].

Exemplos 1.1.1. 1. *Reta real:*

O conjunto dos números reais $\mathbb{R}$ juntamente com a métrica:

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

forma um espaço métrico.

2. *Espaço* $\mathbb{R}^2$*:*

Tome $x = (\xi_1, \xi_2)$, $y = (\eta_1, \eta_2) \in \mathbb{R}^2$ e considere a métrica euclidiana dada por

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

Dessa forma, temos que $(\mathbb{R}^2, d)$ é um espaço métrico. Lembremos que $\mathbb{R}^2$ possui um produto escalar natural dado por

$$\begin{aligned} \langle\ ,\ \rangle\ &: \mathbb{R}^2 \longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \langle x, y \rangle = \xi_1 \cdot \eta_1 + \xi_2 \cdot \eta_2. \end{aligned}$$

Neste caso, podemos ver que $d(x, y) = \sqrt{\langle x - y, x - y \rangle}$, para todo $(x, y) \in \mathbb{R}^2$.

Além da euclidiana, existem outras métricas clássicas em $\mathbb{R}^2$, por exemplo,

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

Nesse caso, temos que $(\mathbb{R}^2, d_1)$ é também um espaço métrico.

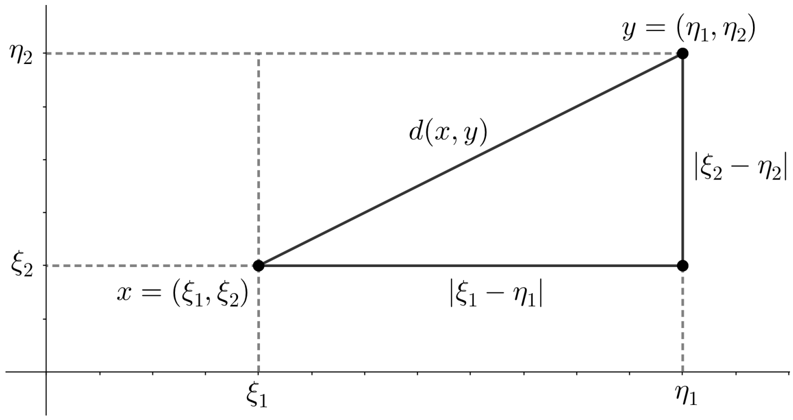


Figura 1.1: Distância entre x e y a partir das métricas d e d_1.

3. *Espaços* $\mathbb{R}^n$ *e* $\mathbb{C}^n$*:*

   Podemos generalizar os dois últimos exemplos, tomando n-uplas $x = (\xi_1, \ldots, \xi_n)$ e $y = (\eta_1, \ldots, \eta_n)$ em $\mathbb{R}^n$ (ou em $\mathbb{C}^n$) com a métrica euclidiana:

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

   Desta forma $\mathbb{R}^n$ (e também $\mathbb{C}^n$) são espaços métricos munidos da métrica acima.

   Vejamos exemplos mais complexos de espaços métricos. Novamente, a demonstração das propriedades (M1) - (M4) pode ser encontrada no Capítulo 1 de [4].

4. *Espaço métrico discreto:*

   Seja X um conjunto qualquer não vazio. Definimos uma métrica d em X por:

$$d(x, x) = 0, \ d(x, y) = 1 \text{ se } x \neq y.$$

   Para esse caso, as propriedades (M1) - (M4) são de fácil verificação e serão omitidas. Essa métrica é conhecida como métrica discreta ou métrica zero-um.

5. *Espaço das funções contínuas:*

   Considere o espaço das funções reais e contínuas definidas no intervalo fechado $J = [a, b]$.

Uma métrica natural nesse espaço é definida por:

$$d(x, y) = \max_{t \in J} |x(t) - y(t)|. \tag{1.1}$$

Denotamos esse espaço métrico por $C[a, b]$.

6. *Espaço de sequências l^∞:*

   Denotemos por l^∞ o espaço de todas as sequências complexas limitadas, ou seja, $x = (\xi_j) \in l^\infty$, se existe $c_x > 0$ tal que $|\xi_j| \leq c_x$, para todo $j \in \mathbb{N}$. Consideramos em l^∞ a seguinte métrica

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

   onde $x = (\xi_j), y = (\eta_j) \in \ l^\infty$.

1.1.2 Espaços l^p, desigualdades de Hölder e Minkowski

Definição 1.1.2. *Seja $p \geq 1$ um número real fixo. Considere o conjunto*

$$\ell^p \doteq \left\{ x = (\xi_j) \text{ sequência real ou complexa; } \sum_{j=1}^{\infty} |\xi_j|^p < \infty \right\}.$$

Definimos em l^p a seguinte métrica

$$d(x, y) = \left(\sum_{j=1}^{\infty} |\xi_j - \eta_j|^p \right)^{1/p}, \tag{1.2}$$

onde $x = (\xi_j)$ e $y = (\eta_j)$ são elementos de ℓ^p. Podemos tomar somente sequências reais ou somente sequências complexas, assim obtendo o espaço real ℓ^p ou o espaço complexo ℓ^p. O caso $p = 2$ é denotado por espaço ℓ^2 das sequências de Hilbert.

Iremos demonstrar que ℓ^p é um espaço métrico, para isto observe que, pela definição da distância, é fácil ver que as condições (M1) até (M3) são satisfeitas. A propriedade (M4) será demonstrada a partir dos resultados a seguir.

Lema 1.1.3. *Sejam $p > 1$ e q definido por*

$$\frac{1}{p} + \frac{1}{q} = 1. \tag{1.3}$$

Os números p e q são ditos conjugados. Então, dados α e β números reais não negativos temos que:

$$\alpha\beta \leq \frac{\alpha^p}{p} + \frac{\beta^q}{q}.$$

Demonstração. Se $\alpha = 0$ ou $\beta = 0$ o resultado é imediato, suponha então que $\alpha \neq 0$ e $\beta \neq 0$. De 1.3 temos que,

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

Assim, $u = t^{p-1}$ implica em $t = u^{q-1}$. Notando que $\alpha\beta$ é o retângulo nas figuras 1.2 e 1.3 segue que

$$\alpha\beta \leq \int_0^\alpha t^{p-1}\,dt + \int_0^\beta u^{q-1}\,du = \frac{\alpha^p}{p} + \frac{\beta^q}{q}. \tag{1.4}$$

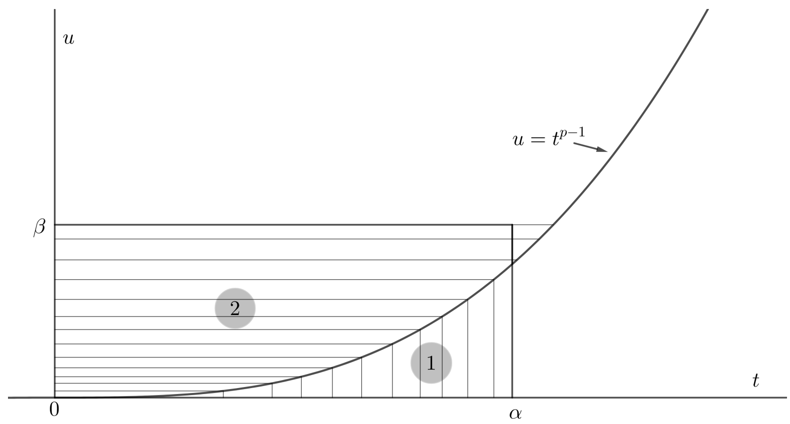


Figura 1.2: (1) corresponde à primeira integral de (1.4) e (2) à segunda.

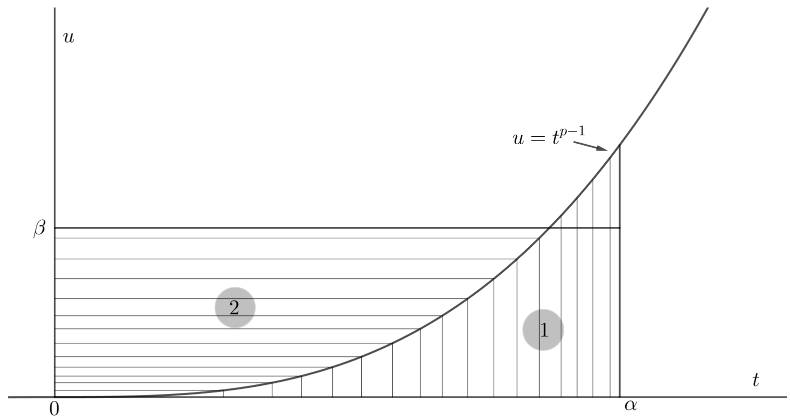


Figura 1.3: (1) corresponde à primeira integral de (1.4) e (2) à segunda.

$\square$

Lema 1.1.4. *(Desigualdade de Hölder) Sejam $p > 1$ e q seu conjugado, tal como no Lema 1.1.3. Dados $x = (\xi_j) \in \ell^p$ e $y = (\eta_j) \in \ell^q$ temos que $z \doteq (\xi_j \eta_j) \in l^1$ e além disso*

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

Demonstração. Sejam $(\tilde{\xi_j}) \in \ell^p$ e $(\tilde{\eta_j}) \in \ell^q$ definidas por

$$\tilde{\xi_j} = \frac{\xi_j}{\left(\sum |\xi_j|^p\right)^{1/p}} \quad \text{e} \quad \tilde{\eta_j} = \frac{\eta_j}{\left(\sum |\eta_j|^q\right)^{1/q}} .$$

Assim, temos que $\sum |\tilde{\xi_j}|^p = 1$ e $\sum |\tilde{\eta_j}|^q = 1$. Denotando $\alpha = |\tilde{\xi_j}|$ e $\beta = |\tilde{\eta_j}|$, temos do Lema 1.1.3, que

$$|\tilde{\xi_j} \tilde{\eta_j}| \leq \frac{1}{p} |\tilde{\xi_j}|^p + \frac{1}{q} |\tilde{\eta_j}|^q.$$

Somando sobre j segue que

$$\frac{\sum |\xi_j \eta_j|}{\left(\sum |\xi_j|^p\right)^{1/p} \left(\sum |\eta_j|^q\right)^{1/q}} = \sum |\tilde{\xi_j} \tilde{\eta_j}| \leq \frac{1}{p} \sum |\tilde{\xi_j}|^p + \frac{1}{q} \sum |\tilde{\eta_j}|^q = \frac{1}{p} + \frac{1}{q} = 1.$$

Por fim, multiplicando ambos os lados da desigualdade acima por $\left(\sum |\xi_j|^p\right)^{1/p} \left(\sum |\eta_j|^q\right)^{1/q}$, segue o resultado. $\square$

No Lema anterior, no caso particular em que $p = q = 2$, temos a chamada *Desigualdade de*

Cauchy-Schwarz

$$\sum_{j=1}^{\infty} |\xi_j \eta_j| \leq \sqrt{\sum_{j=1}^{\infty} |\xi_j|^2} \sqrt{\sum_{j=1}^{\infty} |\eta_j|^2}.$$

Lema 1.1.5. *(Desigualdade de Minkowski) Sejam $x = (\xi_j), y = (\eta_j) \in \ell^p$, onde $p \geq 1$. Então,*

$$\left(\sum_{j=1}^{\infty} |\xi_j + \eta_j|^p \right)^{1/p} \leq \left(\sum_{j=1}^{\infty} |\xi_j|^p \right)^{1/p} + \left(\sum_{j=1}^{\infty} |\eta_j|^p \right)^{1/p}. \tag{1.5}$$

Demonstração. Se $p = 1$ a desigualdade é verdadeira pois se reduz à desigualdade triangular para números. Se $p > 1$ escrevemos $\omega_j = \xi_j + \eta_j$. Segue da desigualdade triangular para números:

$$|\omega_j|^p = |\xi_j + \eta_j||\omega_j|^{p-1} \leq (|\xi_j| + |\eta_j|)|\omega_j|^{p-1}.$$

Somando de j até um n fixo qualquer temos:

$$\sum_{j=1}^{n} |\omega_j|^p \leq \sum_{j=1}^{n} |\xi_j||\omega_j|^{p-1} + \sum_{j=1}^{n} |\eta_j||\omega_j|^{p-1}. \tag{1.6}$$

Para a primeira soma da direita usamos a desigualdade de Hölder:

$$\sum_{j=1}^{n} |\xi_j||\omega_j|^{p-1} \leq \left(\sum_{j=1}^{n} |\xi_j|^p \right)^{1/p} \left(\sum_{j=1}^{n} (|\omega_j|^{p-1})^q \right)^{1/q} = \left(\sum_{j=1}^{n} |\xi_j|^p \right)^{1/p} \left(\sum_{j=1}^{n} |\omega_j|^p \right)^{1/q}.$$

A última parcela da desigualdade é verdadeira pois $(p-1)q = p$, pela demonstração do Lema 1.1.3.

Fazendo o mesmo com a última soma de (1.6) temos

$$\sum_{j=1}^{n} |\eta_j||\omega_j|^{p-1} \leq \left(\sum_{j=1}^{n} |\eta_j|^p \right)^{1/p} \left(\sum_{j=1}^{n} |\omega_j|^p \right)^{1/q}.$$

Juntando as desigualdades anteriores, temos

$$\sum_{j=1}^{n} |\omega_j|^p \leq \left[\left(\sum_{j=1}^{n} |\xi_j|^p \right)^{1/p} + \left(\sum_{j=1}^{n} |\eta_j|^p \right)^{1/p} \right] \left(\sum_{j=1}^{n} |\omega_j|^p \right)^{1/q}.$$

Agora dividimos ambos os lados por $\left(\sum_{j=1}^{n} |\omega_j|^p \right)^{1/q}$ e notando que $1 - 1/q = 1/p$, obtemos a Desigualdade (1.5), mas com n no lugar de ∞. Basta agora fazer $n \longrightarrow \infty$ e notar que as séries no lado direito da desigualdade convergem pois $x, y \in \ell^p$, logo a série que aparece no lado esquerdo da desigualdade também converge. $\square$

Finalmente provemos a desigualdade triangular em ℓ^p. Dados $x = (\xi_j)$, $y = (\eta_j)$ e $z = (\zeta_j)$ em ℓ^p, temos pela desigualdade de Minkowski que a série em (1.2) é convergente. Usando novamente a desigualdade de Minkowski, juntamente com a desigualdade triangular para números, obtemos:

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

o que prova que ℓ^p é um espaço métrico munido da métrica definida em (1.2).

1.1.3 Noções topológicas em espaços métricos

O intuito dessa seção é apresentar algumas noções topológicas em espaços métricos. Vamos estabelecer algumas notações e nomenclaturas que serão usadas no decorrer do trabalho.

Definição 1.1.6. *Sejam X um espaço métrico, x_0 um elemento em X e $r > 0$ uma constante. Definimos:*

1. $B(x_0; r) = \{x \in X;\ d(x, x_0) < r\}$ *(bola aberta);*
2. $\overline{B}(x_0; r) = \{x \in X;\ d(x, x_0) \leq r\}$ *(bola fechada);*
3. $S(x_0; r) = \{x \in X;\ d(x, x_0) = r\}$ *(esfera).*

Nos conjuntos acima, o elemento x_0 é dito centro e o número r raio. Note também que a seguinte relação é válida $S(x_0; r) = \overline{B}(x_0; r) \setminus B(x_0; r)$.

Definição 1.1.7. *Sejam X um espaço métrico e M um subconjunto de X. Dizemos então que*

- *$x_0 \in M$ é um ponto interior de M se existir $r > 0$, tal que $B(x_0; r) \subset M$;*
- *M é um conjunto aberto se todo ponto de M é um ponto interior de M;*
- *M é fechado se $X \setminus M$ é um conjunto aberto.*

O interior de M é o conjunto de todos os pontos interiores de M e será denotado por $Int(M)$ ou por M°. Temos em particular que $Int(M)$ é um conjunto aberto, sendo o maior subconjunto de M com essa propriedade.

Observação 1.1.8. Iremos nos referir a bola aberta $B(x_0;\epsilon)$ por $\epsilon-$vizinhança de x_0, sendo $\epsilon > 0$. Além disso, chamaremos de vizinhança de x_0 qualquer subconjunto de X que contenha uma $\epsilon-$vizinhança de x_0. Obviamente, se $N \subset M$ e N é uma vizinhança de x_0, então M também é vizinhança de x_0.

A coleção $\mathfrak{T}$ formada por todos os subconjuntos abertos de X possui as seguintes propriedades:

1. $\emptyset \in \mathfrak{T}$ e $X \in \mathfrak{T}$;

2. União de elementos de $\mathfrak{T}$ é um elemento de $\mathfrak{T}$;

3. Interseção finita de elementos de $\mathfrak{T}$ é um elemento de $\mathfrak{T}$.

De fato, para verificar (1) note que obviamente X é aberto. Além disso, $\emptyset$ é também aberto, pois caso não fosse existiria um ponto em $\emptyset$ que não é interior, o que é falso uma vez que $\emptyset$ não possui elementos.

Para provar (2) considere $x \in \bigcup A_{\lambda_i}$, onde $A_{\lambda_i} \in \mathfrak{T}$, então $x \in A_{\lambda_j}$ para algum j, mas A_{λ_j} é aberto, portanto contém uma bola aberta, digamos $B \subset A_{\lambda_j} \subset \bigcup A_{\lambda_i}$, logo $\bigcup A_{\lambda_i}$ é um aberto.

Tome $y \in \bigcap_{i=1}^{n} A_i$, onde cada A_i é aberto. Assim, para cada $i = 1, \ldots, n$, existe uma bola centrada em y que está contida em A_i. Tome dentre essas n bolas a que possui o menor raio. Portanto, essa bola estará contida em A_i, para todo $i = 1, \ldots, n$, logo estará contida na interseção $\bigcap_{i=1}^{n} A_i$, o que prova (3).

Definição 1.1.9. *Sejam $X = (X, d)$ e $Y = (Y, \tilde{d})$ espaços métricos. Uma aplicação $T : X \to Y$ é contínua em $x_0 \in X$ quando dado $\epsilon > 0$, existe $\delta > 0$ tal que $d(x, x_0) < \delta$ implica $\tilde{d}(Tx, Tx_0) < \epsilon$.*

Diremos que a aplicação T é contínua quando for contínua em todo ponto de X.

Teorema 1.1.10. *Uma aplicação $T : X \to Y$ entre os espaços métricos X e Y é contínua se, e somente se a imagem inversa de qualquer conjunto aberto de Y é um conjunto aberto de X.*

Demonstração. Por um lado, temos que T é contínuo. Seja $S \subset Y$ um subconjunto aberto e $T^{-1}(S) = S_0$. Se $S_0 = \emptyset$, então S_0 é aberto.

Consideremos que $S_0 \neq \emptyset$ e tome $x_0 \in S_0$, digamos que $y_0 = Tx_0$.

Como S é aberto, contém uma ϵ-vizinhança N de y_0, assim como T é contínuo, segue que existe uma δ-vizinhança N_0 tal que $T : N_0 \to N$. Como $N \subset S$, temos que $N_0 \subset S_0$, logo S_0 é aberto.

Por outro lado, assumimos que a imagem inversa de qualquer conjunto aberto de Y é um conjunto aberto em X.

Desse modo, dado $x_0 \in X$ e uma ϵ-vizinhança N de Tx_0, temos que $T^{-1}(N) = N_0$ é aberto pois N é aberto e N_0 contém x_0. Portanto N_0 contém uma δ-vizinhança de x_0 que é mapeada em N (pois N_0 é mapeado em N), pela definição de continuidade, segue que T é continuo em x_0 (onde x_0 é arbitrário). $\square$

Definição 1.1.11. *Seja M um subconjunto do espaço métrico X. Um ponto $x_0 \in X$ (não necessariamente em M) é dito ponto de acumulação de M, se toda vizinhança de x_0 contém pelo menos um $y \in M$, $y \neq x_0$.*

Definição 1.1.12. *O conjunto formado por todos os pontos de M juntamente com os pontos de acumulação de M é chamado de fecho do conjunto M e denotado $\overline{M}$. Em particular temos que $\overline{M}$ é fechado, sendo o menor conjunto fechado que contém M.*

Definição 1.1.13. *Seja M subconjunto de X, onde X é espaço métrico. M é dito denso em X se $\overline{M} = X$.*

X é separável se tem um subconjunto contável que é denso em X.

Observação 1.1.14. Se M é denso em X, então toda bola em X contém pontos de M.

1.1.4 Sequências e convergência em espaços métricos

Quando nos referimos a sequência em espaços métricos, usamos uma definição totalmente análoga à vista em um curso de análise real; aqui ao invés da reta, os elementos da sequência podem estar em um espaço métrico qualquer. Veremos que nem todo resultado válido na reta, usando o módulo de um número real, continua valido para espaços métricos quaisquer. Por exemplo, nem toda sequência de Cauchy é convergente. Começamos com a definição de convergência em um espaço métrico.

Definição 1.1.15. *Uma sequência (x_n) em um espaço métrico $X = (X, d)$ converge, ou é convergente, se existe $x \in X$ tal que*

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

nesse caso x é dito limite de (x_n) e escrevemos

$$\lim_{n\to\infty} x_n = x,$$

ou simplesmente $x_n \longrightarrow x$.

Se (x_n) não converge para nenhum $x \in X$, dizemos que (x_n) é divergente.

Observação 1.1.16. Note que escrever $x_n \longrightarrow x$ em X significa que a sequência de números reais $a_n = d(x_n, x)$ converge para zero, quando $n \to \infty$. Além disso, isso equivale afirmar que para cada $\epsilon > 0$, existe $N = N(\epsilon)$ em $\mathbb{N}$ tal que, todo x_n com $n > N$, pertence à $\epsilon-$vizinhança $B(x;\epsilon)$ de x.

Por exemplo, se tomarmos o subconjunto $X = (0,1)$ de $\mathbb{R}$ com a métrica $d(x,y) = |x-y|$, teremos que a sequência

$$\left(\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\right)$$

não converge, pois o seu "limite" seria 0, mas $0 \notin X$.

Definição 1.1.17. *Dizemos que um subconjunto $M \subset X$ é limitado quando seu diâmetro*

$$\delta(M) = \sup_{x,y\in M} d(x,y)$$

é finito. Se M é limitado, então existe um número $r > 0$ suficientemente grande tal que para qualquer $x_0 \in X$, $M \subset B(x_0; r)$.

Nesse sentido, uma sequência (x_n) em X é dita limitada se o seu conjunto de pontos correspondente for limitado.

Lema 1.1.18. *Seja $X = (X, d)$ um espaço métrico. Então:*

1. *Uma sequência convergente em X é limitada e seu limite é único.*

2. *Se $x_n \longrightarrow x$ e $y_n \longrightarrow y$ em X, então*

$$d(x_n, y_n) \longrightarrow d(x,y).$$

Demonstração. 1. Suponha que $x_n \longrightarrow x$.

Tomando $\epsilon = 1$ sabemos que existe N tal que $d(x_n, x) < 1$, para todo $n > N$.

Segue pela desigualdade triangular que

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) \leq 2a \quad (\text{para todo } m, n \in \mathbb{N}),$$

onde

$$a = \max\{1, d(x_1, x), \ldots, d(x_N, x)\}.$$

Ou seja, (x_n) é limitada.

Assumindo que $x_n \longrightarrow x$ e $x_n \longrightarrow z$, obtemos novamente pela desigualdade triangular

$$0 \leq d(x, z) \leq d(x, x_n) + d(x_n, z) \longrightarrow 0 + 0 = 0.$$

Ou seja, $d(x, z) = 0$ e segue da propriedade de métrica (M2) que $x = z$.

2. Temos, pela desigualdade triangular generalizada,

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

Portanto,

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

Logo

$$|d(x_n, y_n) - d(x, y)| \leq d(x_n, x) + d(y_n, y) \xrightarrow{n \longrightarrow \infty} 0.$$

□

Lembre-se que uma sequência (x_n) de números reais, ou complexos, converge em $\mathbb{R}$, ou $\mathbb{C}$, se e somente se, satisfaz o *Critério de Convergência de Cauchy*, ou seja, dado $\epsilon > 0$ existe $N = N(\epsilon)$ tal que

$$|x_m - x_n| < \epsilon, \quad \text{para todo } \ m, n > N.$$

Chamamos (x_n) de *sequência de Cauchy*. Veremos que, para um espaço métrico qualquer, esse critério não vale de modo geral, porém os espaços com essa propriedade têm grande importância na Análise Funcional.

Se X tem essa propriedade, dizemos que X é um *espaço métrico completo*, ou seja, toda sequência de Cauchy em X é uma sequência convergente. Para prosseguir usaremos em diversos momentos o resultado a seguir (a demonstração pode ser encontrada em [5])).

Teorema 1.1.19. *A reta real $\mathbb{R}$ e o plano complexo $\mathbb{C}$ são espaços métricos completos.*

Observação 1.1.20. Se considerarmos a métrica em $\mathbb{R}$ dada por $d(x, y) = |x - y|$, temos que os seguintes espaços métricos $(M, \overline{d})$, sendo $\overline{d}$ a métrica induzida por d, não são completos.

- $M = \mathbb{R} \setminus \{a\}$, com $a \in \mathbb{R}$;
- $M = \mathbb{Q}$;
- $M = (a, b) \subset \mathbb{R}$, com $a < b$.

Teorema 1.1.21. *Toda sequência convergente em um espaço métrico é uma sequência de Cauchy.*

Demonstração. Se $x_n \longrightarrow x$, então para todo $\epsilon > 0$ existe $N = N(\epsilon)$ tal que

$$d(x_n, x) < \frac{\epsilon}{2}, \quad \text{para todo } n > N.$$

Pela desigualdade triangular, obtemos para $m, n > N$:

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

Logo (x_m) é de Cauchy. □

Teorema 1.1.22. *Seja $M \neq \emptyset$ subconjunto de um espaço métrico (X, d) e $\overline{M}$ seu fecho, então:*

1. *$x \in \overline{M}$ se, e somente se, existe uma sequência (x_n) em M tal que $x_n \longrightarrow x$.*
2. *M é fechado se, e somente se, dada (x_n) em M com $x_n \longrightarrow x$, então $x \in M$.*

Demonstração. 1. Seja $x \in \overline{M}$. Se $x \in M$, tome a sequência $(x_n) = (x, x, \ldots, x, \ldots)$ assim teremos $x_n \longrightarrow x$. Se $x \notin M$, x é ponto de acumulação de M.

Portanto para cada $n = 1, 2, \ldots$, a bola $B(x; \frac{1}{n})$ contém um $x_n \in M$ e $x_n \longrightarrow x$ pois $\frac{1}{n} \longrightarrow 0$, quando $n \longrightarrow \infty$.

Por outro lado, se (x_n) está em M e $x_n \longrightarrow x$, então $x \in M$ ou cada vizinhança de x contém pontos $x_n \neq x$, ou seja, x é ponto de acumulação de M. Logo $x \in \overline{M}$.

2. Segue diretamente de (1).

$\square$

Teorema 1.1.23. *Um subespaço M de um espaço métrico completo X também é completo se, e somente, se o conjunto M é fechado em X.*

Demonstração. Suponha que $M \subset X$ é completo. Dada uma sequência (x_n) em M tal que $x_n \to x$, concluímos que (x_n) é de Cauchy. Como M é completo, segue que $x \in M$. Segue do Teorema anterior que M é fechado.

Por outro lado, considere M fechado e (x_n) uma sequência de Cauchy em M. Como X é completo, temos que existe $x \in X$ com $x_n \longrightarrow x$, assim $x \in \overline{M}$, porém $M = \overline{M}$, por hipótese. Portanto, $x \in M$, o que prova a completude de M. $\square$

Teorema 1.1.24. *Uma função $T : X \longrightarrow Y$ de um espaço métrico (X, d) para um espaço métrico $(Y, \tilde{d})$ é contínua em um ponto $x_0 \in X$ se, e somente se, $x_n \longrightarrow x_0$ implica $Tx_n \longrightarrow Tx_0$.*

Demonstração. Assuma que T é contínua em x_0, ou seja, dado $\epsilon > 0$ existe $\delta > 0$ tal que $d(x, x_0) < \delta \Rightarrow \tilde{d}(Tx, Tx_0) < \epsilon$.

Se $x_n \longrightarrow x_0$, então existe N tal que para todo $n > N$ temos

$$d(x_n, x_0) < \delta.$$

Portanto, para todo $n > N$, $\tilde{d}(Tx_n, Tx_0) < \epsilon$. Logo, segue pela definição que $Tx_n \longrightarrow Tx_0$.

Por outro lado, assumimos que $x_n \longrightarrow x_0$ implica $Tx_n \longrightarrow Tx_0$ e provaremos que T é contínua em x_0. Suponha que é falso, ou seja, existe $\epsilon > 0$ tal que para todo $\delta > 0$ existe $x \neq x_0$ satisfazendo $d(x, x_0) < \delta$ mas $\tilde{d}(Tx, Tx_0) \geq \epsilon$.

Em particular, se $\delta = \frac{1}{n}$ existe x_n que satisfaz $d(x_n, x_0) < \frac{1}{n}$ mas $\tilde{d}(Tx_n, Tx_0) \geq \epsilon$.

Claramente $x_n \longrightarrow x_0$ mas (Tx_n) não converge para Tx_0, o que contradiz a hipótese. $\square$

Vamos ver alguns exemplos de espaços métricos com a propriedade de que toda sequência de Cauchy é convergente, ou seja, espaços métricos completos.

Exemplos 1.1.25. **(1)** Os espaços $\mathbb{R}^n$ e $\mathbb{C}^n$ são completos se considerarmos a métrica

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

onde $x = (\xi_j)$ e $y = (\eta_j)$.

Considere uma sequência de Cauchy (x_m) em $\mathbb{R}^n$, escrevemos $x_m = (\xi_1^{(m)}, \dots, \xi_n^{(m)})$. Como (x_m) é de Cauchy, para todo $\epsilon > 0$ existe N tal que

$$d(x_m, x_r) = \left(\sum_{j=1}^{r}(\xi_j^m - \xi_j^r)^2\right)^{\frac{1}{2}} < \epsilon \quad (m, r > N). \tag{1.7}$$

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

$$(\xi_j^{(m)} - \xi_j^{(r)})^2 < \epsilon^2 \quad \Rightarrow \quad |\xi_j^{(m)} - \xi_j^{(r)}| < \epsilon.$$

Ou seja, para cada $0 \leq j \leq n$ a sequência $(\xi_j^{(1)}, \xi_j^{(2)}, \dots)$ é sequência de Cauchy de números reais. Como já vimos, esta sequência converge, digamos

$$\xi_j^{(m)} \xrightarrow{m \longrightarrow \infty} \xi_j.$$

Definimos $x = (\xi_1, \xi_2, \dots, \xi_n)$. Note que $x \in \mathbb{R}^n$. De (1.7), fazendo $r \longrightarrow \infty$:

$$d(x_m, x) \leq \epsilon \quad (m > N).$$

Ou seja, x é o limite de (x_m), logo $\mathbb{R}^n$ é completo, pois (x_m) era arbitrária. Para $\mathbb{C}^n$ usamos a mesma métrica e a prova é análoga.

(2) ℓ^∞ é completo.

De fato, seja (x_m) sequência de Cauchy em ℓ^∞, onde $x_m = (\xi_1^{(m)}, \xi_2^{(m)}, \dots)$ e usamos a métrica:

$$d(x,y) = \sup_j |\xi_j - \eta_j|,$$

onde $x = (\xi_j)$ e $y = (\eta_j)$.

Deste modo, como (x_m) é de Cauchy, para todo $\epsilon > 0$ existe N tal que para todo $m, n > N$,

$$d(x_m, x_n) = \sup_j |\xi_j^{(m)} - \xi_j^{(n)}| < \epsilon.$$

Para cada j fixado:

$$|\xi_j^{(m)} - \xi_j^{(n)}| < \sup_j |\xi_j^{(m)} - \xi_j^{(n)}| < \epsilon \quad (m, n > N). \tag{1.8}$$

Logo, para cada j, a sequência $(\xi_j^{(1)}, \xi_j^{(2)}, \ldots)$ é sequência de Cauchy de números, portanto converge, digamos

$$\xi_j^{(m)} \xrightarrow{m \longrightarrow \infty} \xi_j.$$

Usando os limites $\xi_1, \xi_2, \ldots$ definimos $x = (\xi_1, \xi_2, \ldots)$ e mostraremos que $x \in \ell^\infty$ e $x_m \longrightarrow x$.

De (1.8) e fazendo $n \longrightarrow \infty$, temos:

$$|\xi_j^{(m)} - \xi_j| \leq \epsilon \quad (m > N), \tag{1.9}$$

como $x_m = (\xi_j^{(m)}) \in \ell^\infty$, existe $k_m \in \mathbb{R}$ tal que $|\xi_j^{(m)}| \leq k_m$ para todo j.

Portanto, pela desigualdade triangular:

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

A desigualdade é verdadeira para todo j e o lado direito não envolve j, portanto (ξ_j) é uma sequência limitada de números. Logo $x = (\xi_j) \in \ell^\infty$. Também, de (1.9), obtemos

$$d(x_m, x) = \sup_j |\xi_j^{(m)} - \xi_j| \leq \epsilon \quad (m > N).$$

Logo $x_m \longrightarrow x$. Como (x_m) era sequência de Cauchy arbitrária, ℓ^∞ é completo.

(3) O espaço c é completo, onde c denota o espaço das sequências convergentes de números complexos com a métrica induzida de ℓ^∞.

De fato, c é subespaço de ℓ^∞, assim mostraremos que c é fechado em ℓ^∞.

Considere $x = (\xi_j) \in \bar{c}$. Portanto existe $x_n = (\xi_j^{(n)}) \in c$ tal que $x_n \longrightarrow x$ (pelo teorema 1.1.22).

Dado $\epsilon > 0$, existe N tal que para $n \geq N$ e para todo j temos:

$$|\xi_j^{(n)} - \xi_j| \leq d(x_n, x) < \frac{\epsilon}{3}.$$

Em particular, para $n = N$. Como $x_N \in c$, seus termos $\xi_j^{(N)}$ formam uma sequência convergente. Tal sequência é de Cauchy, portanto existe N_1 tal que

$$|\xi_j^{(N)} - \xi_k^{(N)}| < \frac{\epsilon}{3} \quad (j, k \geq N_1).$$

Pela desigualdade triangular para $j, k \geq N_1$:

$$|\xi_j - \xi_k| \leq |\xi_j - \xi_j^{(N)}| + |\xi_j^{(N)} - \xi_k^{(N)}| + |\xi_k^{(N)} - \xi_k| < \epsilon.$$

Desta forma $x = (\xi_j)$ é convergente, portanto $x \in c$. Logo c é fechado em ℓ^∞, ou seja, c é completo.

(4) ℓ^p é completo, $1 \leq p \leq \infty$.

De fato, seja (x_m) sequência de Cauchy qualquer em ℓ^p onde $x_m = (\xi_1^{(m)}, \xi_2^{(m)}, \ldots)$, deste modo para todo $\epsilon > 0$ existe N tal que para todo $m, n > N$

$$d(x_m, x_n) = \left(\sum_{j=1}^{\infty} |\xi_j^{(m)} - \xi_j^{(n)}| \right)^{\frac{1}{p}} < \epsilon. \tag{1.10}$$

Segue para todo $j = 1, 2, \ldots$ que

$$|\xi_j^{(m)} - \xi_j^{(n)}| < \epsilon \quad (m, n > N).$$

Escolhemos um j fixo. Da expressão, vemos que $(\xi_j^{(1)}, \xi_j^{(2)}, \ldots)$ é uma sequência de Cauchy de números, portanto converge pois $\mathbb{R}$ e $\mathbb{C}$ são completos.

Digamos,

$$\xi_j^{(m)} \xrightarrow{m \longrightarrow \infty} \xi_j.$$

Vamos mostrar que $x \in \ell^p$ e $x_m \longrightarrow x$. Por (1.10) temos para todo $m, n > N$

$$\sum_{j=1}^{k} |\xi_j^{(m)} - \xi_j^{(n)}|^p < \epsilon^p \quad (k = 1, 2, \ldots).$$

Fazendo $n \longrightarrow \infty$, obtemos para $m > N$:

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

Fazendo $k \longrightarrow \infty$, obtemos para todo $m > N$

$$\sum_{j=1}^{\infty} |\xi_j^{(m)} - \xi_j|^p \leq \epsilon^p. \tag{1.11}$$

Portanto $x_m - x = (\xi_j^{(m)} - \xi_j) \in \ell^p$. Como $x_m \in \ell^p$, segue que:

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

A série em (1.11) representa $\left[d(x_m, x)\right]^p$, ou seja, $x_m \longrightarrow x$. Logo ℓ^p é completo.

(5) $C[a, b]$ (espaço das funções contínuas no intervalo fechado $[a, b]$ em $\mathbb{R}$) é completo.

De fato, seja (x_m) de Cauchy em $C[a, b]$.

Dado $\epsilon > 0$, existe N tal que para todo $m, n > N$ temos:

$$d(x_m, x_n) = \max_{t \in [a,b]} (x_m(t) - x_n(t)) < \epsilon. \tag{1.12}$$

Portanto, para qualquer $t = t_0$ em $[a, b]$,

$$|x_m(t_0) - x_n(t_0)| < \epsilon \quad (m, n > N).$$

Portanto $(x_1(t_0), x_2(t_0), \ldots)$ é sequência real de Cauchy e converge.

Digamos,

$$x_m(t_0) \xrightarrow{m \longrightarrow \infty} x(t_0).$$

Assim, associamos cada $t \in [a, b]$ à um único número real $x(t)$. Isso define uma função x no intervalo $[a, b]$. Mostraremos que $x \in C[a, b]$ e $x_m \longrightarrow x$.

De (1.12) com $n \longrightarrow \infty$:

$$\max_{t \in [a,b]} |x_m(t) - x(t)| \leq \epsilon.$$

Portanto para todo $t \in [a,b]$,

$$|x_m(t) - x(t)| \leq \epsilon \quad (m > N).$$

Deste modo $(x_m(t))$ converge para $x(t)$ uniformemente em $[a,b]$. Como os x_m são contínuos em $[a,b]$ e a convergência é uniforme, a função limite x é contínua em $[a,b]$, portanto $x \in C[a,b]$, também segue $x_m \longrightarrow x$. Logo $c[a,b]$ é completo.

Observação: O mesmo resultado vale se considerarmos funções complexas e a demonstração é praticamente análoga.

Podemos destacar o último resultado do exemplo anterior no seguinte teorema.

Teorema 1.1.26. *A convergência $x_n \longrightarrow x$ em $C[a,b]$ é uma convergência uniforme, isto é, (x_m) converge uniformemente em $[a,b]$ para x.*

1.2 Espaços normados e espaços de Banach

Essa seção tem como principal objetivo discutir espaços cuja métrica está definida em termos de uma *norma*. Aqui usaremos livremente resultados de álgebra linear.

Definição 1.2.1. *Dado um espaço vetorial X (real ou complexo), uma norma é uma função real*

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

a qual possui as propriedades:

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

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

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

(N4) $\|x + y\| \leq \|x\| + \|y\|$ *(desigualdade triangular).*

Onde x e y são vetores quaisquer em X e α é um escalar arbitrário.

Uma norma em X define uma métrica d, dada por:

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

Dizemos que d é a métrica induzida pela norma.

Definição 1.2.2. *Um espaço vetorial X é dito espaço normado se existe uma norma em X. Um espaço de Banach é um espaço normado e completo na métrica induzida pela norma.*

Denotamos o espaço normado por $(X,\|\cdot\|)$ ou simplesmente X.

Espaços de Banach terão um papel importante nos próximos capítulos.

Segue de (N4) que

$$\begin{aligned} &\|y-x\| + \|x\| \geq \|y\| \\ &\Rightarrow \|y\| - \|x\| \leq \|y-x\|. \end{aligned}$$

De maneira análoga conseguimos a expressão:

$$-\left(\|y\| - \|x\| \right) \leq \|y-x\|.$$

Logo,

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

Deste modo, segue que a norma é contínua, ou seja, a aplicação $x \longrightarrow \|x\|$ de X em $\mathbb{R}$ é contínua.

Vejamos alguns exemplos de espaços normados e de Banach.

Exemplos 1.2.3. Veremos exemplos de espaços de Banach, os quais são completos (demonstrado anteriormente). Não provaremos que as expressões definem de fato normas, a verificação das propriedades (N1)-(N4) pode ser encontrada em [4].

(1) Os espaços $\mathbb{R}^n$ e $\mathbb{C}^n$ são de Banach usando a norma:

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

O espaço é completo com a métrica:

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

(2) ℓ^p é um espaço de Banach com a norma:

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

Essa norma induz a métrica:

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

(3) ℓ^∞ é espaço de Banach com a norma:

$$\|x\| = \sup_i |\xi_j|.$$

(4) $C[a,b]$ é espaço de Banach com a norma:

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

Lema 1.2.4. *Uma métrica d, induzida por uma norma em um espaço normado X, satisfaz:*

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

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

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

Demonstração. Temos que

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

Além disso

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

□

Vejamos mais algumas propriedades interessantes, mas antes precisamos definir alguns conceitos básicos.

Definição 1.2.5. *Um subespaço Y de um espaço normado X é um subespaço vetorial de X, com a norma obtida restringindo a norma em X ao subconjunto Y. Essa norma é dita induzida pela norma de X. Se Y é fechado em X, então Y é chamado de subespaço fechado de X.*

Um subespaço Y de um espaço de Banach X é, portanto, um subespaço vetorial com a norma induzida de X, mas Y não é necessariamente completo.

Teorema 1.2.6. *Um subespaço Y de um espaço de Banach X é completo se, e somente se, o conjunto Y é fechado em X.*

Demonstração. Segue diretamente do Teorema 1.1.23. □

Agora podemos falar de *convergência* de sequências em termos da norma. Iremos usar os resultados já provados e a métrica $d(x, y) = \|x - y\|$. Logo,

- uma sequência (x_n) em um espaço normado X é *convergente* se existe $x \in X$ tal que

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

  Escrevemos $x_n \longrightarrow x$ e x é dito *limite* de (x_n).

- Uma sequência (x_n) em um espaço normado X é de *Cauchy* se para todo $\epsilon > 0$ existe N tal que

$$\|x_m - x_n\| < \epsilon, \quad \text{para todo } m, n > N.$$

Em espaços normados podemos considerar *séries*. Seja (x_k) uma sequência em um espaço normado X, associamos (x_n) à sequência (s_n) das *somas parciais*, onde

$$s_n = x_1 + x_2 + \cdots + x_n \quad (n = 1, 2, \ldots).$$

Se (s_n) é convergente, digamos $s_n \longrightarrow s$, ou seja, $\|s_n - s\| \longrightarrow 0$, então a *série*

$$\sum_{k=1}^{\infty} x_k = x_1 + x_2 + \cdots \tag{1.13}$$

é dita *convergente* e s é dita *soma* da série e escrevemos:

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

Se $\|x_1\| + \|x_2\| + \cdots$ converge dizemos que a série (1.13) é *absolutamente convergente*. Vejamos agora um importante resultado sobre convergência absoluta.

Lema 1.2.7. *Dado um espaço de Banach X temos que toda série absolutamente convergente é convergente.*

Demonstração. Seja $\sum x_j$ uma série absolutamente convergente, portanto a sequência das somas parciais de $\sum\|x_j\|$ é de Cauchy, isto é, para todo $\epsilon > 0$ e para todo $p \in \mathbb{N}$, temos para n suficientemente grande o seguinte

$$\left|\sum_{j=1}^{n}\|x_j\| - \sum_{j=1}^{n+p}\|x_j\|\right| = \sum_{j=n+1}^{n+p}\|x_j\| \leq \epsilon.$$

Pela desigualdade triangular obtemos:

$$\left\|\sum_{j=1}^{n} x_j - \sum_{j=1}^{n+p} x_j\right\| = \left\|\sum_{j=n+1}^{n+p} x_j\right\| \leq \sum_{j=n+1}^{n+p}\|x_j\| \leq \epsilon.$$

Ou seja, a sequência das somas parciais de $\sum x_j$ é de Cauchy, como X é completo, segue que a série é convergente. □

Usamos o conceito de convergência para definir uma "base" para um espaço normado. Se existe uma série (e_n) no espaço normado X, de modo que para todo $x \in X$ existe uma única sequência de escalares (α_n) tal que

$$\|x - (\alpha_1 e_1 + \cdots + \alpha_n e_n)\| \xrightarrow{n \longrightarrow \infty} 0. \tag{1.14}$$

Então (e_n) é dita *base de Schauder* de X. A série

$$\sum_{k=1}^{\infty} \alpha_k e_k,$$

cuja soma é x, é dita *expansão* de x com respeito à (e_n) e escrevemos

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

Exemplo 1.2.8. Em ℓ^p se consideramos (e_n) definida por $e_n = \delta_{nj}$, ou seja,

$$\begin{aligned} e_1 &= (1, 0, 0, 0, 0, \dots) \\ e_2 &= (0, 1, 0, 0, 0, \dots) \\ e_3 &= (0, 0, 1, 0, 0, \dots) \\ \vdots &\qquad \vdots \end{aligned}$$

temos que (e_n) é uma base de Schauder de ℓ^p.

Se um espaço normado X possui uma base de Schauder, então X é separável. De fato, basta tomar $M \subset X$ como sendo o conjunto das combinações lineares dos elementos de (e_n). Temos que M é contável, pois (e_n) é contável e segue diretamente que para todo $x \in X$, dado $\epsilon > 0$ existe N tal que

$$\left\|x - (\alpha_1 e_1 + \cdots + \alpha_n e_n)\right\| < \epsilon \quad (n > N).$$

Basta tomar $y = (\alpha_1 e_1 + \cdots + \alpha_n e_n) \in M$ e teremos

$$\|x - y\| < \epsilon.$$

Logo, M é denso em X.

A recíproca não é verdadeira. O próprio matemático Banach levantou a questão ao desenvolver a teoria. Apesar da maioria dos espaços separáveis de Banach estudados até agora terem uma base de Schauder, o matemático sueco P. Enflo (1973) demonstrou com a construção de um exemplo a invalidade da recíproca.

1.2.1 Espaços normados de dimensão finita

Discutiremos importantes propriedades sobre espaços normados com dimensão finita, estas que serão de grande importância em resultados futuros.

Lema 1.2.9. *Seja $\{x_1, \ldots, x_n\}$ um conjunto linearmente independente de vetores em um espaço normado X (de qualquer dimensão). Então existe um número $c > 0$ tal que para toda escolha de escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$ temos*

$$\|\alpha_1 x_1 + \cdots + \alpha_n x_n\| \geq c(|\alpha_1| + \cdots + |\alpha_n|).$$

Demonstração. Escrevemos $s = |\alpha_1| + \cdots + |\alpha_n|$. Se $s = 0$, $\alpha_j = 0$ para todo $j = 1, \ldots, n$, então segue a igualdade desejada para qualquer valor de c. Suponha então que $s > 0$, queremos provar que

$$\frac{\|\alpha_1 x_1 + \cdots + \alpha_n x_n\|}{s} \geq c$$

e escrevendo $\beta_j = \frac{\alpha_j}{s}$, a desigualdade é equivalente à

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

Portanto é suficiente mostrar a desigualdade acima, ou seja, a existência de um $c > 0$ tal que a desigualdade é verdadeira para todas as $n-$uplas de escalares $\beta_1, \ldots, \beta_n$ com $\sum |\beta_j| = 1$.

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

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

tal que

$$\|y_m\| \xrightarrow{m \longrightarrow \infty} 0.$$

Como $\sum |\beta_j^{(m)}| = 1$, temos $|\beta_j^{(m)}| \leq 1$. Portanto para todo j fixo, a sequência

$$(\beta_j^{(m)}) = (\beta_j^{(1)}, \beta_j^{(2)}, \ldots)$$

é limitada. Pelo Teorema de Bolzano Weierstrass, $(\beta_1^{(m)})$ tem subsequência convergente. Denote β_1 o limite dessa subsequência e seja $(y_{1,m})$ a subsequência correspondente de (y_m). Pelo mesmo argumento, $(y_{1,m})$ tem subsequência $(y_{2,m})$ para a qual a subsequência correspondente de escalares $(\beta_2^{(m)})$ converge; seja β_2 o limite dessa sequência. Continuando com esse processo, após n passos obtemos um sequência $(y_{n,m}) = (y_{n,1}, y_{n,2}, \ldots)$ de (y_m) cujos termos são da forma

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

com escalares $\gamma_j^{(m)}$ satisfazendo $\gamma_j^{(m)} \xrightarrow{m \longrightarrow \infty} \beta_j$. Assim,

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

onde $\sum |\beta_j| = 1$, então nem todo β_j pode ser zero. Como $\{x_1, \ldots, x_n\}$ é um conjunto linearmente independente, temos $y \neq 0$. Por outro lado, $y_{n,m} \longrightarrow y$ implica $\left\|y_{n,m}\right\| \longrightarrow \|y\|$, pela continuidade da norma. Como $\|y_m\| \longrightarrow 0$ por hipótese e $(y_{n,m})$ é subsequência de (y_m), devemos ter

$\|y_{n,m}\| \longrightarrow 0$. Portanto $\|y\| = 0$, então $y = 0$ por (N2); o que contradiz $y \neq 0$ e prova o lema. $\square$

Vamos aplicar esse lema no próximo resultado.

Teorema 1.2.10. *Todo subespaço vetorial de dimensão finita $Y \subset X$ (X é espaço normado) é completo. Em particular, todo espaço normado de dimensão finita é completo.*

Demonstração. Considere uma sequência de Cauchy arbitrária (y_m) em Y e vamos mostrar que converge em Y e denotaremos o limite por y. Seja $\dim Y = n$ e $\{e_1, \ldots, e_n\}$ uma base qualquer de Y. Então cada y_m possui uma representação única da forma

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

Como (y_m) é sequência de Cauchy, para todo $\epsilon > 0$ existe um N tal que $\|y_m - y_r\| < \epsilon$ onde $m, r > N$. Disso e do Lema 1.2.9 segue que temos para algum $c > 0$:

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

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

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

Isso mostra que cada uma das n sequências

$$(\alpha_j^{(m)}) = (\alpha_j^{(1)}, \alpha_j^{(2)}, \ldots)$$

é de Cauchy em $\mathbb{R}$ ou $\mathbb{C}$. Portanto converge, digamos que α_j denota o limite. Usando os n limites $\alpha_1, \ldots, \alpha_n$, definimos

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

Note que $y \in Y$. Além disso,

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

Temos que $\alpha_j^{(m)} \xrightarrow{m \longrightarrow \infty} \alpha_j$. Portanto $\|y_m - y\| \longrightarrow 0$, ou seja, $y_m \longrightarrow y$. O que mostra que

(y_m) converge em Y. $\square$

Agora, usando esse resultado e o Teorema 1.1.23 segue diretamente o teorema:

Teorema 1.2.11. *Todo subespaço Y de dimensão finita de um espaço normado X é fechado em X.*

Note que nem todo subespaço com dimensão infinita é fechado. Por exemplo seja $X = C[0,1]$ e $Y = \text{span}\{x_0, x_1, \ldots\}$ onde $x_j(t) = t^j$, ou seja, Y é o conjunto dos polinômios. Y não é fechado em X. De fato, pelo teorema da aproximação de Stone-Weierstrass, qualquer função real contínua em $[0,1]$ pode ser aproximada por polinômios, ou seja, é limite de uma sequência de polinômios. Então se Y fosse fechado teríamos que $Y = C[0,1]$, o que é uma contradição.

Agora veremos que em um espaço vetorial X com dimensão finita, os conjuntos abertos de X são os mesmos, independente da escolha de norma.

Definição 1.2.12. *Uma norma $\|\cdot\|$ em um espaço vetorial X é dita equivalente à norma $\|\cdot\|_0$ em X se existem números positivos a e b tais que para todo $x \in X$ temos*

$$a\|x\|_0 \leq \|x\| \leq b\|x\|_0 .$$

Segue do Lema 1.2.9 o teorema a seguir.

Teorema 1.2.13. *Em um espaço vetorial de dimensão finita X, qualquer norma $\|\cdot\|$ é equivalente à qualquer outra norma $\|\cdot\|_0$.*

Demonstração. Seja $\dim X = n$ e $\{e_1, \ldots, e_n\}$ uma base qualquer de X. Então todo $x \in X$ tem uma representação única da forma

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

Pelo Lema 1.2.9 existe uma constante positiva c tal que

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

Por outro lado, segue da desigualdade triangular

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

Deste modo, $a\|x\|_0 \leq \|x\|$, onde $a = \frac{c}{k} > 0$. Para a outra desigualdade basta trocar os papéis de $\|\cdot\|_0$ e $\|\cdot\|$. $\square$

Esse teorema, em especial, garante que a convergência de sequências em espaços normados de dimensão finita não depende da norma escolhida.

1.2.2 Conjuntos compactos

Iremos definir conjuntos compactos e demonstrar alguns resultados envolvendo espaços normados de dimensão finita. Começamos com a seguinte definição.

Definição 1.2.14. *Um espaço métrico X é dito compacto (ou sequencialmente compacto) se toda sequência em X possui uma subsequência convergente. Um subconjunto M de X é dito compacto se M visto como subespaço de X é compacto, ou seja, toda sequência em M possui subsequência convergente, a qual tem limite em M.*

Provemos uma propriedade geral dos conjuntos compactos.

Lema 1.2.15. *Um subconjunto compacto M de um espaço métrico é fechado e limitado.*

Demonstração. Para todo $x \in \overline{M}$ existe uma sequência $(x_n) \in M$ tal que $x_n \longrightarrow x$. Como M é compacto, então $x \in M$. Portanto M é fechado. Vamos provar que M é limitado. Se M fosse não limitado, teria uma sequência não limitada (y_n) tal que $d(y_n, b) > n$, onde b é um elemento fixo. Essa sequência não poderia ter uma subsequência convergente, pois uma subsequência convergente deve ser limitada pelo Lema 1.1.18. Logo segue o resultado. $\square$

A recíproca do teorema anterior, em geral, é falsa. Para provar isso considere uma sequência (e_n) em ℓ^2, onde $e_n = \delta_{nj}$ (função delta de Kronecker). Essa sequência é limitada, pois $\|e_n\| = 1$. Além disso, seus termos constituem um conjunto de pontos que é fechado, pois não possui pontos de acumulação. Pela mesma razão, esse conjunto de pontos não é compacto.

Em contrapartida, podemos garantir o seguinte para espaços normados de dimensão finita.

Teorema 1.2.16. *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. Do Lema 1.2.15, se um conjunto é compacto então é fechado e limitado, agora vamos provar a recíproca. Seja $M \subset X$ fechado e limitado. Digamos que $\dim X = n$ e $\{e_1, \ldots, e_n\}$

base de X. Consideramos uma sequência $(x_m) \in M$. 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 $\|x_m\| \leq k$ para todo m. Pelo Lema 1.2.9,

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

onde $c > 0$. Portanto a sequência de números $(\xi_j^{(m)})$ (para j fixo) é limitada e, pelo teorema de Bolzano-Weierstrass, tem um ponto de acumulação ξ_j, onde $1 \leq j \leq n$. Concluímos 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 subsequência convergente, a qual converge em M. Logo M é compacto. $\square$

Vejamos um lema útil em aplicações.

Lema 1.2.17 (Riesz)**.** *Sejam Y e Z subespaços de um espaço normado X (de qualquer dimensão), e suponha que Y é fechado e é conjunto próprio de Z. Então para todo número real $\theta \in (0,1)$ existe um $z \in Z$ tal que*

$$\|z\| = 1, \quad \|z - y\| \geq \theta \text{ para todo } y \in Y.$$

Demonstração. Considere qualquer $v \in (Z - Y)$ e denote a distância de v a Y por a, ou seja

$$a = \inf_{y \in Y} \|v - y\|.$$

Observe que $a > 0$, pois Y é fechado. Considere um $\theta \in (0,1)$. Pela definição de ínfimo existe $y_0 \in Y$ tal que

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

Temos que $\frac{a}{\theta} > a$, pois $0 < \theta < 1$. Seja

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

Então $\|z\| = 1$. Agora mostraremos que $\|z - y\| \leq \theta$, para todo $y \in Y$. Temos

$$\|z - y\| = \big\|c(v - y_0) - y\big\| = c\Big\|v - y_0 - c^{-1}y\Big\| = c\|v - y_1\|\,.$$

Aqui $y_1 = y_0 + c^{-1}y$.

Portanto $y_1 \in Y$ (pois $y, y_0 \in Y$). Deste modo, $\|v - y_1\| \geq a$ (segue da definição de a). Finalmente podemos concluir que

$$\|z - y\| = c\|v - y_1\| \geq ca = \frac{a}{\|v - y_0\|} \geq \frac{a}{a/\theta} = \theta.$$

□

Espaços normados de dimensão finita tem a particularidade que a bola unitária fechada é compacta (pelo Teorema 1.2.16). Reciprocamente, segue do Lema 1.2.17 o próximo resultado.

Teorema 1.2.18. *Seja X um espaço normado com a propriedade que a bola unitária $B_X = \{x \in X;\ \|x\| \leq 1\}$ é compacta, então X tem dimensão finita.*

Demonstração. Assumimos que B_X é compacto, mas $\dim X = \infty$ e chegaremos em uma contradição. Escolha um x_1 tal que $\|x_1\| = 1$, considere $X_1 = \operatorname{span}\{x_1\} \subsetneq X$, temos que X_1 é fechado (pois tem dimensão finita, então é compacto). Pelo Lema 1.2.17 existe um $x_2 \in X$, com $\|x_2\| = 1$ tal que

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

Usando x_1, x_2 consideramos o espaço $X_2 = \operatorname{span}\{x_1, x_2\} \subsetneq X$ ($\dim X_2 = 2$), o qual é fechado. Novamente, pelo Lema 1.2.17, existe um $x_3 \in X$, com $\|x_3\| = 1$ tal que para todo $x \in X$

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

Em particular,

$$\|x_3 - x_1\| \geq \frac{1}{2},$$

$$\|x_3 - x_2\| \geq \frac{1}{2}.$$

Continuando por indução obtemos uma sequência (x_n) de elementos $x_n \in B_X$ tal que

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

Deste modo, (x_n) não pode ter uma subsequência convergente, o que contradiz o fato de B_X ser compacto. Logo $\dim X = \infty$ é falso e segue o resultado. $\square$

Um resultado importante sobre compacidade está relacionado com operadores contínuos, veremos que conjuntos compactos têm imagens compactas.

Teorema 1.2.19. *Sejam X e Y espaços métricos e $T : X \longrightarrow Y$ um operador contínuo. Então a imagem de um conjunto compacto por T também é um conjunto compacto.*

Demonstração. Considere $M \subset X$ um conjunto compacto qualquer. É suficiente mostrar que toda sequência (y_n) em $T(M) \subset Y$ contém subsequência convergente em $T(M)$. Como $y_n \in T(M)$, temos que $y_n = Tx_n$ para algum $x_n \in M$. M é compacto, então (x_n) possui subsequência (x_{n_k}) convergente em M. A imagem de (x_{n_k}) é uma subsequência de (y_n) que converge em $T(M)$, pois T é contínua. Logo $T(M)$ é compacto. $\square$

Do teorema acima segue o próximo resultado, que é um clássico de cálculo, e se mantém para espaços métricos quaisquer.

Corolário 1.2.20. *Um operador contínuo $T : M \subset X \longrightarrow \mathbb{R}$, onde M é subconjunto compacto de X, assume um máximo e um mínimo para alguns pontos em M.*

Demonstração. $T(M) \subset \mathbb{R}$ é compacto pelo Teorema 1.2.19 e fechado pelo Lema 1.2.15 (aplicado à $T(M)$), então $\inf T(M) \in T(M)$, $\sup T(M) \in T(M)$, e as imagens inversas desses pontos consistem em pontos de M, os quais Tx é mínimo ou máximo, respectivamente. $\square$

1.3 Introdução aos operadores lineares limitados

1.3.1 Operadores lineares: definições e exemplos

Nessa seção iremos discutir algumas propriedades dos operadores lineares definidos sobre espaços vetoriais. Uma vez que os conceitos vistos aqui são muito análogos aos vistos em Álgebra Linear, a abordagem será mais direta e algumas demonstrações não serão apresentadas.

Definição 1.3.1. *Sejam X e Y espaços vetoriais definidos sobre o mesmo corpo de escalares. Um operador linear, definido em um subespaço vetorial $\mathscr{D}(T)$ de X, é uma aplicação $T : \mathscr{D}(T) \longrightarrow Y$*

que satisfaz para todo $x, y \in \mathscr{D}(T)$ e escalar α as seguintes igualdades:

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

ou equivalentemente, para todo $x, y \in \mathscr{D}(T)$ e escalares α, β

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

Dois importantes conjuntos são considerados com relação a um operador linear T. O primeiro deles é chamado de *núcleo* de T e é definido por

$$\ker(T) \doteq \{x \in \mathscr{D}(T)\ ;\ Tx = 0\},$$

enquanto que o segundo é chamado de *imagem* de T a qual é definida por

$$\mathscr{R}(T) \doteq \{y \in Y\ ;\ y = Tx \ \text{ para algum } x \in \mathscr{D}(T)\}.$$

Veremos no Teorema 1.3.3 que $\ker(T)$ e $\mathscr{R}(T)$ são subespaços vetoriais de X e Y, respectivamente.

Sempre é valida a propriedade $0 \in \ker(T)$. De fato, $T0 = T(0+0) = T0 + T0$, subtraindo $T0$ de ambos os lados segue que $0 = T0$. Usando esse fato, veremos no Teorema 1.3.5 que T é *injetivo* se, e somente se, $\ker(T) = \{0\}$. O operador T é *sobrejetivo* quando $\mathscr{R}(T) = Y$.

A seguir apresentamos alguns exemplos de operadores lineares.

Exemplos 1.3.2. (1) (Operador Identidade) O operador $I_X : X \longrightarrow X$ dado por $I_X(x) = x$ para todo $x \in X$ é um operador linear. A prova é imediata.

Também denotamos I_X por I quando não há risco de ambiguidades.

(2) (Operador Nulo) O operador $0 : X \longrightarrow Y$ dado por $0x = 0$ para todo $x \in X$ é linear e a prova é imediata.

(3) (Diferenciação) Seja X o espaço vetorial de polinômios definidos em $[a, b]$ e considere o operador $T : X \longrightarrow X$ dado por $T(x(t)) = x'(t)$. Usando diretamente as propriedades de derivada de polinômios concluímos que T é linear.

(4) (Produto Cruzado) Em $\mathbb{R}^3$ o "produto escalar por um fator fixo" também define um operador linear, ou seja, $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$, dado por $Tx = x \cdot a = \xi_1\alpha_1 + \xi_2\alpha_2 + \xi_3\alpha_3$, onde

$x = (\xi_1, \xi_2, \xi_3) \in \mathbb{R}^3$ é um vetor qualquer e $a = (\alpha_1, \alpha_2, \alpha_3)$ está fixado. De fato, dados γ escalar, $x = (\xi_j)$ e $y = (\eta_j)$, com $j \in \{1, 2, 3\}$, temos que

$$\begin{aligned} T(\gamma x + y) &= \alpha_1(\gamma\xi_1 + \eta_1) + \alpha_2(\gamma\xi_2 + \eta_2) + \alpha_3(\gamma\xi_3 + \eta_3) \\ &= (\gamma\alpha_1\xi_1 + \gamma\alpha_2\xi_2 + \gamma\alpha_3\xi_3) + (\alpha_1\eta_1 + \alpha_2\eta_2 + \alpha_3\eta_3) \\ &= \gamma(\alpha_1\xi_1 + \alpha_2\xi_2 + \alpha_3\xi_3) + (\alpha_1\eta_1 + \alpha_2\eta_2 + \alpha_3\eta_3) = \gamma Tx + Ty. \end{aligned}$$

(5) (Matrizes) Uma matriz real $A = (\alpha_{jk})$ com r linhas e n colunas define um operador $T : \mathbb{R}^n \longrightarrow \mathbb{R}^r$, dado por $Tx = Ax = y$ (estamos identificando o vetor x com a matriz coluna das coordenadas). Ou seja, se $x = (\xi_j) \in \mathbb{R}^n$ e $y = (\eta_j) \in \mathbb{R}^r$ temos

$$\begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_r \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{r1} & \cdots & \alpha_{rn} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix}.$$

Segue das propriedades de matrizes que T é linear.

Definido operador linear o próximo passo é relembrar resultados relevantes. Começamos com um resultado cuja consequência é conhecida como *Teorema do Núcleo e da Imagem.*

Teorema 1.3.3. *Seja T um operador linear. Então:*

1. *A imagem $\mathscr{R}(T)$ é espaço vetorial;*

2. *Se $\dim \mathscr{D}(T) = n < \infty$, então $\dim \mathscr{R}(T) \leq n$;*

3. *O núcleo $\ker(T)$ é um espaço vetorial.*

Demonstração.

1. Tome $y_1, y_2 \in \mathscr{R}(T)$. Temos que $y_1 = Tx_1$ e $y_2 = Tx_2$ para certos $x_1, x_2 \in \mathscr{D}(T)$, além disso $\alpha x_1 + \beta x_2 \in \mathscr{D}(T)$ para quaisquer α, β escalares. Portanto $T(\alpha x_1 + \beta x_2) = \alpha Tx_1 + \beta Tx_2 = \alpha y_1 + \beta y_2 \in \mathscr{R}(t)$, logo $\mathscr{R}(T)$ é um espaço vetorial.

2. Escolha $n + 1$ elementos $y_1, \ldots, y_{n+1} \in \mathscr{R}(T)$, deste modo existem $x_1, \ldots, x_{n+1} \in \mathscr{D}(T)$ tais que $y_1 = Tx_1,\ y_2 = Tx_2,\ \ldots\ ,\ y_{n+1} = Tx_{n+1}$. Como $\dim \mathscr{D}(T) = n$, $\{x_1, \ldots, x_{n+1}\}$ é linearmente dependente. Logo, na equação

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

existe pelo menos um $\alpha_j \neq 0$ ($j \in \{1, \dots, n+1\}$). Como T é linear e $T0 = 0$ temos

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

ou seja, $\{y_1, \dots, y_{n+1}\}$ é linearmente dependente (pois nem todo α_j é zero). Logo, não existe nenhum conjunto linearmente independente com $n+1$ elementos em $\mathscr{R}(T)$, ou seja, $\dim \mathscr{R}(T) \leq n$.

3. Tome $x_1, x_2 \in \ker(T)$, então $Tx_1 = Tx_2 = 0$. Logo

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

Portanto $\alpha x_1 + \beta x_2 \in \ker(T)$.

□

Corolário 1.3.4. *Dado X espaço vetorial de dimensão finita e $T : \mathscr{D}(T) \longrightarrow \mathscr{R}(T)$ um operador linear, temos que* $\dim \mathscr{D}(T) = \dim \ker(T) + \dim \mathscr{R}(T)$.

Demonstração. (Ideia da demonstração) Denote $\dim \mathscr{D}(T) = n$ e $\dim \mathscr{R}(T) = m$, sabemos que $m \leq n$. Tome um conjunto linearmente independente $\{x_1, \dots, x_m\}$ de modo que $Tx_j \neq 0$ para todo $j = 1, \dots, m$ (se isso não fosse possível, chegaríamos à uma contradição com a dimensão da imagem). Para concluir o teorema observamos que se $m = n$ nada precisa ser feito, e caso $m < n$ completamos o conjunto à uma base de $\mathscr{D}(T)$ e observa-se que é necessário exatamente de $\dim \ker(T)$ elementos para tal. □

Se T é um operador linear injetivo, então existe a função $T^{-1} : \mathscr{R}(T) \longrightarrow \mathscr{D}(T)$, onde $T^{-1}y = x$ (tal que $Tx = y$). Além disso segue que

$$T^{-1}Tx = x \text{ e } TT^{-1}y = y, \quad \text{para todos} \quad x \in \mathscr{D}(T), y \in \mathscr{R}(T).$$

Vamos discutir alguns resultados envolvendo operadores inversos.

Teorema 1.3.5. *Sejam X e Y espaços vetoriais reais ou complexos. Seja $T : \mathscr{D}(T) \longrightarrow Y$ operador linear com domínio $\mathscr{D}(T) \subset X$ e imagem $\mathscr{R}(T) \subset Y$. Então:*

1. *O operador inverso $T^{-1} : \mathscr{R}(T) \longrightarrow \mathscr{D}(T)$ existe se, e somente se,* $\ker(T) = \{0\}$.

2. *Se T^{-1} existe, então T^{-1} é operador linear;*

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

Demonstração. 1. Suponha que $\ker(T) = \{0\}$. Seja $Tx_1 = Tx_2 \Rightarrow T(x_1 - x_2) = 0$, então, por hipótese, temos que $x_1 - x_2 = 0 \Rightarrow x_1 = x_2$. Portanto T é injetiva, logo existe a inversa T^{-1} definida na imagem de T.

Por outro lado, suponha que T^{-1} existe. Assim, se $x \in \ker(T)$, então $Tx = 0$ donde segue que $x = T^{-1}(Tx) = 0$.

2. Supondo que T^{-1} existe, já vimos que $\mathscr{R}(T)$ é espaço vetorial (é o domínio de T^{-1}). Considere $x_1, x_2 \in \mathscr{D}(T)$ e $y_1 = Tx_1$, $y_2 = Tx_2$. Portanto, $x_1 = T^{-1}y_1$ e $x_2 = T^{-1}y_2$. Como T é linear, $\alpha y_1 + \beta y_2 = \alpha Tx_1 + \beta Tx_2 = T(\alpha x_1 + \beta x_2)$. Portanto,

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

3. Como $\dim \mathscr{D}(T) = n < \infty$, temos pelo Teorema 1.3.3 que $\dim \mathscr{R}(T) \leq \dim \mathscr{D}(T)$, da mesma forma $\dim \mathscr{D}(T) = \dim \mathscr{R}(T^{-1}) \leq \dim \mathscr{D}(T^{-1}) = \dim \mathscr{R}(T)$, logo $\dim \mathscr{R}(T) = \dim \mathscr{D}(T)$.

□

Lema 1.3.6. *Seja* $T : X \longrightarrow Y$ *e* $S : Y \longrightarrow Z$ *operadores lineares bijetivos, onde* X, Y *e* Z *são espaços vetoriais. O operador inverso* $(ST)^{-1} : Z \longrightarrow X$ *do produto* ST *existe e vale a seguinte igualdade* $(ST)^{-1} = T^{-1}S^{-1}$.

Demonstração. O operador $ST : X \longrightarrow Z$ é bijetivo. De fato, é injetivo pois se $(ST)x = 0$, então $Tx = 0$ (pois S é injetivo), ademais $x = 0$ (pois T é injetivo). É sobrejetivo, pois dado $z \in Z$ temos que existe $y \in Y$ de modo que $Sy = z$, além disso existe um $x \in X$ tal que $Tx = y$, logo $(ST)x = z$, portanto (ST) é sobrejetivo. Logo, $(ST)^{-1}$ existe.

Temos, $ST(ST)^{-1} = I_Z$ (operador identidade). Aplicando S^{-1} e usando $S^{-1}S = I_Y$, obtemos:

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

Agora, aplicando T^{-1} e usando que $T^{-1}T = I_X$, segue que

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

□

1.3.2 Operadores lineares limitados

Dando sequência ao estudo de operadores lineares, veremos algumas propriedades dos operadores lineares limitados definidos entre espaços normados. Diferente de funções limitadas, as quais simplesmente tem a sua imagem contida em um conjunto limitado, operadores lineares limitados são obtidos no seguinte sentido: existe uma constante real de forma que para todo x a norma da imagem Tx não excede a norma de x multiplicada por essa constante, mais precisamente temos a seguinte definição.

Definição 1.3.7. *Sejam X e Y espaços normados e $T : \mathscr{D}(T) \longrightarrow Y$ um operador linear, onde $\mathscr{D}(T) \subset X$. O operador T é dito limitado se existe uma constante $c > 0$ tal que, para todo $x \in \mathscr{D}(T)$, vale*

$$\|Tx\| \leq c\|x\|\,. \tag{1.15}$$

Segue imediatamente da definição anterior que um operador linear limitado T leva conjunto limitado em conjunto limitado. Além disso, veja que se T é um operador linear limitado, então existe $c > 0$ tal que $\frac{\|Tx\|}{\|x\|} \leq c$, para todo $x \in \mathscr{D}(T)/\{0\}$. Portanto o conjunto

$$\left\{\frac{\|Tx\|}{\|x\|} \,:\, x \in \mathscr{D}(T)/\{0\}\right\}$$

é limitado e dessa forma o menor $c > 0$ que satisfaz (1.15) será dado por

$$\|T\| \doteq \sup_{x \in \mathscr{D}(T)/\{0\}} \frac{\|Tx\|}{\|x\|}. \tag{1.16}$$

Se $\mathscr{D}(T) = \{0\}$, definimos $\|T\| = 0$ (neste caso $T = 0$, pois $T0 = 0$). Por (1.15) com $c = \|T\|$ temos que $\|Tx\| \leq \|T\|\|x\|$, para todo $x \in \mathscr{D}(T)$.

A notação usada em (1.16) é motivada pelo fato de que são válidas todas as propriedades de norma conforme mostra o seguinte lema.

Lema 1.3.8. *Seja T operador linear limitado, como definido anteriormente, então:*

1. *Uma fórmula alternativa para (1.16) é*

$$\|T\| = \sup_{\|x\|=1} \|Tx\|\,,$$

   sendo $x \in \mathscr{D}(T)$.

2. A fórmula definida em (1.16) satisfaz as propriedades de norma.

Demonstração. 1. Como T é linear, temos:

$$\sup_{x\in\mathscr{D}(T)/\{0\}} \frac{\|Tx\|}{\|x\|} = \sup_{x\in\mathscr{D}(T)/\{0\}} \left\|\frac{1}{\|x\|}Tx\right\| = \sup_{x\in\mathscr{D}(T)/\{0\}} \left\|T\left(\frac{x}{\|x\|}\right)\right\| = \sup_{y\in\mathscr{D}(T),\ \|y\|=1} \|Ty\|.$$

2. - Se $\|T\| = 0$, então, pelo item anteior

$$\sup_{\|x\|=1} \|Tx\| = 0 \ \Rightarrow \ T = 0.$$

- Dado α escalar, temos

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

- Dados T_1 e T_2 operadores lineares limitados, definidos sobre o mesmo domínio e contradomínio, temos que para todo x com $\|x\| = 1$ vale

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

Dessa forma

$$\|T_1 + T_2\| = \sup_{\|x\|=1} \left\|(T_1 + T_2)x\right\| \leq \|T_1\| + \|T_2\|.$$

□

Vejamos alguns exemplos de operadores lineares limitados.

Exemplo 1.3.9 (Operador identidade)**.** O operador linear $I : X \longrightarrow X$ dado por $Ix = x$ é limitado e

$$\|I\| = \sup_{\|x\|=1} \|Ix\| = \sup_{\|x\|=1} \|x\| = 1.$$

Exemplo 1.3.10 (operador nulo)**.** O operador linear $0 : X \longrightarrow Y$ é obviamente limitado e $\|0\| = 0$ por definição.

Exemplo 1.3.11 (operador diferenciação)**.** Seja X o espaço normado de polinômios definido no intervalo $[0, 1]$ com norma $\|x\| = \max |x(t)|$, para $t \in [0, 1]$. Definimos T em X, por $Tx(t) = x'(t)$. Temos que T é linear, mas não é limitado. De fato, tome (x_n) uma sequência em X dada por

$x_n(t) = t^n$, assim

$$Tx_n(t) = x_n'(t) = nt^{n-1} \;\Rightarrow\; \|Tx_n\| = n \;\Rightarrow\; \frac{\|Tx_n\|}{\|x_n\|} = \frac{n}{1} = n.$$

Quando $n \longrightarrow \infty$ vemos que não é possível encontrar uma constante $c > 0$ que satisfaça (1.15), logo o operador linear T não é limitado.

Exemplo 1.3.12 (operador integral)**.** Considere o espaço das funções contínuas $C\,[0,1]$ e o operador

$$\begin{aligned} T : C\,[0,1] &\longrightarrow C\,[0,1] \\ x \mapsto Tx = y \quad &\text{onde} \quad y(t) = \int_0^1 k(t,\tau)x(\tau)d\tau, \end{aligned}$$

onde k é uma função dada, contínua no quadrado fechado $[0,1] \times [0,1]$, a qual é chamada de *núcleo integral* do operador T. Vejamos que T é um operador linear limitado.

Primeiramente temos que T é linear pois

$$\begin{aligned} \alpha T(x_1) + \beta T(x_2) = \alpha y_1 + \beta y_2 &= \alpha \int_0^1 k(t,\tau)x_1(\tau)d\tau + \beta \int_0^1 k(t,\tau)x_2(\tau)d\tau \\ &= \int_0^1 k(t,\tau)(\alpha x_1(\tau) + \beta x_2(\tau))d\tau = T(\alpha x_1 + \beta x_2). \end{aligned}$$

Agora vejamos que T é limitado. De fato, primeiro note que como k é uma função contínua no compacto $[0,1] \times [0,1]$, então k é limitada, ou seja, existe $C > 0$ tal que $|k(t,\tau)| \leq C$, para todo $(t,\tau) \in [0,1] \times [0,1]$. Como $|x(t)| \leq \max\{|x(t)| : t \in [0,1]\} = \|x\|$ segue que,

$$\begin{aligned} \|y\| = \|Tx\| = \max_{t\in[0,1]} \left| \int_0^1 k(t,\tau)x(\tau)d\tau \right| &\leq \max_{t\in[0,1]} \int_0^1 |k(t,\tau)||x(\tau)|d\tau \leq C\|x\| \\ \Rightarrow \|Tx\| \leq C\|x\| \;\Rightarrow\; \frac{\|Tx\|}{\|x\|} \leq C, &\quad \text{para todo} \quad x \in C\,[0,1]\,. \end{aligned}$$

Logo, na desigualdade acima tomando o supremo sobre todos o valores de $x \neq 0$ segue que $\|T\| \leq C$.

Exemplo 1.3.13 (Matriz)**.** Dada uma matriz $A = (\alpha_{jk})$, de entradas reais, com r linhas e n

colunas, definimos

$$T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$
$$Ax = y,$$

onde $x = (\xi_j),\ y = (\eta_j)$ são vetores coluna. Em decorrência, as componentes η_j são dadas por

$$\eta_j = \sum_{k=1}^{n} \alpha_{jk}\xi_k \quad (j = 1, \ldots, r).$$

Já sabemos que T é um operador linear, pois a multiplicação de matrizes é uma operação linear, então provemos que é limitado. Lembre-se que $\|x\| = \left(\sum_{m=1}^{n} \xi_m^2\right)^{\frac{1}{2}}$, para $x \in \mathbb{R}^n$ e $\|y\|$ é similar. Logo, pela desigualdade de Cauchy-Schwarz segue que

$$\|Tx\|^2 = \sum_{j=1}^{r} \eta_j^2 = \sum_{j=1}^{r} \left[\sum_{k=1}^{n} \alpha_{jk}\xi_k\right]^2 \leq \sum_{j=1}^{r} \left[\left(\sum_{k=1}^{n} \alpha_{jk}^2\right)^{\frac{1}{2}} \left(\sum_{m=1}^{n} \xi_m^2\right)^{\frac{1}{2}}\right]^2 = \|x\|^2 \sum_{j=1}^{r}\sum_{k=1}^{n} \alpha_{jk}^2.$$

Como a última soma não depende de x, temos:

$$\|Tx\|^2 \leq c^2\|x\|^2 \quad \text{onde} \quad c^2 = \sum_{j=1}^{r}\sum_{k=1}^{n} \alpha_{jk}^2,$$

logo T é limitado.

O próximo resultado nos dirá que todo operador linear em um espaço normado de dimensão finita é limitado.

Teorema 1.3.14. *Seja X espaço normado com dimensão finita, então todo operador linear definido em X é limitado.*

Demonstração. Sejam $n = \dim X$ e $\{e_1, \ldots, e_n\}$ uma base de X. Tome $x = \sum \xi_j e_j$ e considere T operador linear em X, como T é linear:

$$\|Tx\| = \left\|\sum \xi_j Te_j\right\| \leq \sum |\xi_j|\ \left\|Te_j\right\| \leq \max_k \|Te_k\| \sum |\xi_j|,$$

as somas são de 1 à n. Vamos aplicar o Lema 1.2.9 na última soma, assim obtendo:

$$\sum |\xi_j| \leq \frac{1}{c}\left\|\sum \xi_j e_j\right\| = \frac{1}{c}\|x\| \ \Rightarrow\ \|Tx\| \leq \gamma\|x\|\,,$$

onde $\gamma = \frac{1}{c}\max_k \|Te_k\|$. Logo, T é limitado. $\square$

Sejam X, Y espaços normados e $\mathscr{D}(T) \subset X$ um subespaço vetorial. Lembremos que um operador $T : \mathscr{D}(T) \longrightarrow Y$ é contínuo em um ponto $x_0 \in \mathscr{D}(T)$ se para todo $\epsilon > 0$, existe um $\delta > 0$ tal que $\|x - x_0\| < \delta \Rightarrow \|Tx - Tx_0\| < \epsilon$. Se for contínuo em todo ponto, o operador T é dito contínuo. O próximo teorema mostra que a limitação de um operador liner está intimamente ligado a continuidade desse operador.

Teorema 1.3.15. *Seja $T : \mathscr{D}(T) \longrightarrow Y$ operador linear, $\mathscr{D}(T) \subset X$ e X, Y são espaços normados. Então:*

1. *T é contínuo se, e somente se, é limitado;*

2. *Se T é contínuo em um ponto, então T é contínuo.*

Demonstração. 1. Se $T = 0$ o resultado é trivial. Vamos considerar então apenas $T \neq 0$.

Assumimos inicialmente que T é limitado e consideremos $x_0 \in \mathscr{D}(T)$. Dado $\epsilon > 0$, como T é linear, para cada $x \in \mathscr{D}(T)$ tal que $\|x - x_0\| < \delta$, onde $\delta = \epsilon / \|T\|$, obtemos

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

Como x_0 é arbitrário, segue que T é contínuo.

Por outro lado, assuma que T é contínuo em $x_0 \in \mathscr{D}(T)$ arbitrário. Dado $\epsilon > 0$, existe um $\delta > 0$ tal que $\|Tx - Tx_0\| \leq \epsilon$, para todo $x \in \mathscr{D}(T)$ que satisfaz $\|x - x_0\| \leq \delta$.

Tome $y \neq 0$ em $\mathscr{D}(T)$ e seja $x = x_0 + (\delta/\|y\|)y \;\Rightarrow\; x - x_0 = (\delta/\|y\|)y \;\Rightarrow\; \|x - x_0\| = \delta$.

Como T é linear, temos $\|Tx - Tx_0\| = \big\|T(x - x_0)\big\| = \Big\|T\big((\delta/\|y\|)y\big)\Big\| = (\delta/\|y\|)\|Ty\|$, portanto

$$\frac{\delta}{\|y\|}\|Ty\| = \|Tx - Tx_0\| \leq \epsilon \;\Rightarrow\; \|Ty\| \leq \frac{\epsilon}{\delta}\|y\| \,.$$

Logo, T é limitado, pois $y \neq 0$ é arbitrário. Note que se $y = 0$ a desigualdade acima é imediata.

2. Se T é contínuo em um ponto, pelo item anterior temos que T é limitado (veja que na demonstração usamos somente um ponto fixado), portanto novamente pelo item anterior T é contínuo em todo ponto.

$\square$

Corolário 1.3.16. *Seja T um operador linear limitado. Então:*

1. *Se $x_n \longrightarrow x$ ($x_n, x \in \mathscr{D}(T)$), então $Tx_n \longrightarrow Tx$;*

2. *O núcleo* $\ker(T)$ *é fechado.*

Demonstração. 1. Temos que $\|Tx_n - Tx\| = \big\|T(x_n - x)\big\| \leq \|T\| \cdot \|x_n - x\| \longrightarrow 0$, logo $Tx_n \longrightarrow Tx$.

2. Para cada $x \in \overline{\ker(T)}$ existe uma sequência $(x_n) \subset \ker(T)$ tal que $x_n \longrightarrow x$, portanto, pelo item anterior segue que $Tx_n \longrightarrow Tx$. Veja que,

$$\|Tx\| \leq \|Tx - Tx_n\| + \|Tx_n\| = \|Tx - Tx_n\| \longrightarrow 0.$$

Logo, $\|Tx\| = 0 \Rightarrow Tx = 0$, ou seja, $x \in \ker(T)$ e segue que $\ker(T)$ é fechado.

□

Dados operadores lineares limitados $T : X \longrightarrow X$, $T_2 : X \longrightarrow Y$ e $T_1 : Y \longrightarrow Z$, sendo X, Y e Z espaços normados, temos as seguintes desigualdades:

$$\|T_1 T_2\| \leq \|T_1\| \|T_2\| \quad \text{e} \quad \|T^n\| \leq \|T\|^n . \tag{1.17}$$

De fato, para a primeira basta observar que dado $x \in X$ temos:

$$\big\|T_1 T_2(x)\big\| \leq \|T_1\| \|T_2 x\| \leq \|T_1\| \|T_2\| \|x\| .$$

Para a segunda desigualdade basta considerar $X = Y = Z$ e $T_1 = T_2$ e usar Indução Matemática.

Definição 1.3.17. *Dois operadores T_1 e T_2 são ditos iguais se $\mathscr{D}(T_1) = \mathscr{D}(T_2)$ e $T_1 x = T_2 x$ para todo $x \in \mathscr{D}(T_1) = \mathscr{D}(T_2)$. Denotamos $T_1 = T_2$.*

Definição 1.3.18. *A restrição de um operador $T : \mathscr{D}(T) \longrightarrow Y$ à um subespaço $B \subset \mathscr{D}(T)$, denotada por $T|_B$, é o operador:*

$$T|_B : B \longrightarrow Y, \quad T|_B(x) = Tx, \ \forall x \in B.$$

Definição 1.3.19. *Uma extensão de T à um espaço $M \supset \mathscr{D}(T)$ é um operador $\widetilde{T} : M \longrightarrow Y$ tal que $\widetilde{T}|_{\mathscr{D}(T)} = T$.*

O próximo teorema nos garante, sob algumas hipóteses, que podemos "aumentar" o domínio de um operador limitado para o seu fecho.

Teorema 1.3.20. *Seja $T : \mathscr{D}(T) \longrightarrow Y$ um operador linear limitado, onde $X \supset \mathscr{D}(T)$ é um espaço normado e Y é um espaço de Banach. Então, T possui uma extensão $\widetilde{T} : \overline{\mathscr{D}(T)} \longrightarrow Y$ limitada e $\|\widetilde{T}\| = \|T\|$.*

Demonstração. Considere $x \in \overline{\mathscr{D}(T)}$. Existe uma sequência (x_n) em $\mathscr{D}(T)$ tal que $x_n \longrightarrow x$. Como T é linear e limitado, dados $n, p \in \mathbb{N}$:

$$\left\|Tx_n - Tx_{n+p}\right\| = \left\|T(x_n - x_{n+p})\right\| \leq \|T\|\left\|x_n - x_{n+p}\right\| \xrightarrow{n \longrightarrow \infty} 0$$

Logo, (Tx_n) é de Cauchy, pois (x_n) é de Cauchy uma vez que converge. Por hipótese, Y é completo, portanto (Tx_n) converge, digamos, $Tx_n \longrightarrow y \in Y$. Vamos definir $\widetilde{T}$, por $\widetilde{T}x = y$.

Vejamos que a definição anterior independe da sequência em $\mathscr{D}(T)$ que convergindo para x. Suponha que $x_n \longrightarrow x$ e $z_n \longrightarrow x$. Então $v_m \longrightarrow x$, onde $v_n = (x_1, z_1, x_2, z_2, \ldots)$. Deste modo, a sequência (Tv_n) converge, como (Tx_n) e (Tz_n) são subsequências, elas devem convergir para o mesmo valor. Assim $\widetilde{T}$ está unicamente definido para todo $x \in \overline{\mathscr{D}(T)}$.

Vejamos que $\widetilde{T}$ é linear. Tome x_1, x_2, veja que existem $(x_{1,n}), (x_{2,n})$ tais que $x_{1,n} \longrightarrow x_1$ e $x_{2,n} \longrightarrow x_2$, assim $Tx_{1,n} \longrightarrow Tx_1$ e $Tx_{2,n} \longrightarrow Tx_2$. Note que $\alpha\widetilde{T}x_1 + \beta\widetilde{T}x_2 = \alpha y_1 + \beta y_2$. Portanto a sequência $(T(\alpha x_{1,n} + \beta x_{2,n})) = (\alpha T(x_{1,n}) + \beta T(x_{2,n}))$ converge para $\alpha y_1 + \beta y_2$. Logo

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

Se $x \in \mathscr{D}(T)$ veja que $(x_n) = (x, x, x, \ldots)$ converge para x e $Tx_n \longrightarrow Tx$, assim $\widetilde{T}x = Tx$, portanto $\widetilde{T}$ é uma extensão de T.

Resta mostrar que $\|\widetilde{T}\| = \|T\|$. De fato, temos que $\|Tx_n\| \leq \|T\|\|x_n\|$, fazendo $n \longrightarrow \infty$, $Tx_n \longrightarrow y = \widetilde{T}x$. Como $x \mapsto \|x\|$ define uma função contínua, obtemos $\|\widetilde{T}x\| \leq \|T\|\|x\|$, portanto $\widetilde{T}$ é limitado e $\|\widetilde{T}\| \leq \|T\|$. Obviamente, $\|\widetilde{T}\| \geq \|T\|$ (por ser uma extensão, a menor cota possível é $\|T\|$). Logo, $\|\widetilde{T}\| = \|T\|$.

□

Finalizamos esta seção definindo o *espaço dos operadores lineares limitados*. Sejam X e Y espaços normados. Denotamos por $B(X, Y)$ o espaço dos operadores limitados de X em Y

munido com a norma

$$\|T\| = \sup_{x\in X\setminus\{0\}} \frac{\|Tx\|}{\|x\|}.$$

As operações de soma e multiplicação por escalar em $B(X,Y)$ são as usuais para funções, ou seja,

$$\begin{aligned}
T_1+T_2 \doteq (T_1+T_2) : X \longrightarrow Y, &\quad \text{onde } (T_1+T_2)(x) = T_1x + T_2x \quad \text{e}\\
\alpha T \doteq (\alpha T) : X \longrightarrow Y, &\quad \text{onde } (\alpha T)(x) = \alpha Tx.
\end{aligned}$$

Pelo Lema 1.3.8 temos que $\|T\|$ dado acima é de fato uma norma em $B(X,Y)$, sendo essa a norma natural considerada nesse espaço. No caso em que $X = Y$ usaremos a notação $B(X)$.

O próximo teorema, cuja demonstração não será discutida aqui, mostra uma propriedade interessante do espaço definido anteriormente e terá aplicação na seção em que discutiremos *funcionais lineares.*

Teorema 1.3.21. *Dados X, Y espaços normados, se Y é um espaço de Banach, então $B(X,Y)$ é um espaço de Banach.*

A demonstração do Teorema 1.3.21 pode ser encontrada no Capítulo 2 de [4].

1.4 Funcionais lineares

Uma classe especial de operadores lineares são aqueles cujo contradomínio é o corpo de escalares sobre o qual o domínio, como espaço vetorial, está definido.

Dados um espaço vetorial normado X, definido sobre um corpo de escalares $\mathbb{K}$, e $\mathscr{D}(f)$ um subespaço vetorial de X, chamaremos de *funcional linear* um operador linear $f : \mathscr{D}(f) \longrightarrow \mathbb{K}$. Em nosso contexto vamos considerar $\mathbb{K}$ como sendo $\mathbb{R}$ ou $\mathbb{C}$.

Definição 1.4.1. *Um funcional linear limitado é um operador linear limitado com imagem em um corpo de escalares. Se $f : \mathscr{D}(f) \subset X \longrightarrow \mathbb{K}$, então*

$$\|f\| = \sup_{x\neq 0} \frac{|f(x)|}{\|x\|} = \sup_{\|x\|=1} |f(x)|.$$

Da definição de operador limitado, temos $|f(x)| \leq \|f\|\|x\|$.

Observação 1.4.2. Se um operador tem imagem no corpo de escalares e não é linear, ainda assim podemos usar o termo funcional. Por exemplo, a norma $\|\cdot\| : X \longrightarrow \mathbb{R}$ em um espaço normado $(X,\|\cdot\|)$ é um funcional não linear em X.

Como consequência do Teorema 1.3.15, podemos destacar o próximo teorema.

Teorema 1.4.3. *Um funcional linear f, com domínio em um espaço normado, é contínuo se, e somente se, f é limitado.*

Vejamos alguns exemplos de funcionais lineares limitados.

Exemplo 1.4.4. Fixado $a = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$ considere o funcional $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, definido por $f(x) = x \cdot a = \xi_1\alpha_1 + \cdots + \xi_n\alpha_n$, onde $x = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$. Vejamos que f é uma funcional linear limitado e calculemos a sua norma.

Obviamente f é linear. Vamos mostrar que é limitado. Usando a Desigualdade de Cauchy-Schwarz segue que:

$$|f(x)| = |x \cdot a| \leq \|x\|\|a\|.$$

Portanto, $\|f\| \leq \|a\|$. Por outro lado, tomando $x = a$, temos que

$$\|f\| = \sup_{x\neq 0} \frac{|f(x)|}{\|x\|} \geq \frac{|f(a)|}{\|a\|} = \frac{\|a\|^2}{\|a\|} = \|a\| \Rightarrow \|f\| = \|a\|.$$

Exemplo 1.4.5. Considere o espaço normado das funções contínuas $C[a,b]$ e o funcional f dado por

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

Segue das propriedades de integral que f é linear. Vamos provar que f é limitado e que $\|f\| = b - a$. De fato,

$$|f(x)| = \left|\int_a^b x(t)dt\right| \leq (b-a) \max_{t\in[a,b]} |x(t)| = (b-a)\|x\|.$$

Portanto, $\|f\| \leq (b-a)$. Tome a função $x_1 = 1$, note que $\|x_1\| = 1$, assim,

$$\|f\| \geq \frac{|f(x_1)|}{\|x_1\|} = |f(x_1)| = \left|\int_a^b 1dt\right| = (b-a).$$

Logo, $\|f\| = b - a$.

Exemplo 1.4.6. Sejam $p, q \in \mathbb{N}$ tais que $1/p + 1/q = 1$, ou seja, p e q são conjugados. Fixado

$a = (\alpha_j) \in \ell^q$, definimos $f : \ell^p \to \mathbb{R}$ por

$$f(x) = \sum_{j=1}^{\infty} \xi_j \alpha_j \quad \text{onde } x = (\xi_j) \in \ell^p.$$

Se $a = 0$ temos o funcional linear nulo cuja norma $\|f\| = 0$. Caso contrário, pela Desigualdade de Hölder segue que

$$|f(x)| = \left| \sum_{j=1}^{\infty} \xi_j \alpha_j \right| \leq \sum_{j=1}^{\infty} |\xi_j \alpha_j| \leq \left(\sum_{j=1}^{\infty} |\xi_j|^p \right)^{1/p} \left(\sum_{j=1}^{\infty} |\alpha_j|^q \right)^{1/q} = \|x\|_p \|a\|_q .$$

Logo, f está bem definido e $\|f\| \leq \|a\|_q$.

Definição 1.4.7. *(Espaço Dual X') Seja X um $\mathbb{K}$-espaço normado. Então o espaço $B(X, \mathbb{K})$ com a norma*

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

é dito dual de X e denotado por X'.

Do teorema 1.3.21 segue diretamente o próximo resultado.

Teorema 1.4.8. *Dado X um espaço normado qualquer, temos que X' é um espaço de Banach.*

1.5 Introdução a teoria espectral de operadores lineares

No caso de espaços normados com dimensão finita a teoria espectral de operadores lineares se resume basicamente ao estudo de autovalores e autovetores de matrizes. Iremos discutir nesta seção uma breve introdução da teoria espectral de operadores lineares em espaços de dimensão qualquer. Além disso, vamos usar diversos resultados básicos sobre autovalores e autovetores de matrizes.

Consideramos $X \neq \{0\}$ um espaço vetorial complexo normado e $T : \mathscr{D}(T) \longrightarrow X$ um operador linear, onde $\mathscr{D}(T) \subset X$ é um subespaço vetorial. Dado um número complexo λ, associamos ao operador T o seguinte operador

$$T_\lambda = T - \lambda I,$$

sendo I o operador identidade em $\mathscr{D}(T)$. Se T_λ possui inversa, a denotamos por $R_\lambda(T)$, ou seja,

$$R_\lambda(T) = T_\lambda^{-1} = (T - \lambda I)^{-1}.$$

O operador $R_\lambda(T)$ é chamado de *operador resolvente de* T ou apenas *resolvente* de T. Quando for claro quem é o operador T simplesmente escrevemos R_λ.

O operador resolvente R_λ é linear pois T e I são lineares e o operador inverso de um operador linear também é linear pelo Teorema 1.3.5. Ainda pelo Teorema 1.3.5 temos que o resolvente $R_\lambda : \mathrm{Im}(T_\lambda) \longrightarrow \mathscr{D}(T_\lambda)$ existe se, e somente se, $\ker T_\lambda = \{0\}$.

Definição 1.5.1. *Sejam $X \neq \{0\}$ um espaço normado complexo e $T : \mathscr{D}(T) \subset X \longrightarrow X$ um operador linear. Um valor regular λ de T é um número complexo que satisfaz:*

(R_1) $R_\lambda(T)$ existe;

(R_2) $R_\lambda(T)$ é limitado;

(R_3) R_λ (T) está definido em um conjunto denso em X.

O conjunto de todos os valores regulares λ de T é chamado de *conjunto resolvente* e é denotado por $\rho(T)$. O complementar de $\rho(T)$ em $\mathbb{C}$ é chamado de *espectro* de T e é denotado por $\sigma(T) = \mathbb{C} \setminus \rho(T)$. Além disso $\lambda \in \sigma(T)$ é chamado de *valor espectral* de T. Vejamos que $\sigma(T)$ pode ser subdividido em três subconjuntos disjuntos:

- O *espectro discreto* (ou espectro pontual) $\sigma_p(T)$ é o conjunto dos números λ tais que $R_\lambda(T)$ não existe. Um elemento $\lambda \in \sigma_p(T)$ é dito *autovalor* de T.

- O *espectro contínuo* $\sigma_c(T)$ é o conjunto dos números λ tais que $R_\lambda(T)$ existe e satisfaz (R3) mas não satisfaz (R2), ou seja, $R_\lambda(T)$ não é limitado.

- O *espectro residual* $\sigma_r(T)$ é o conjunto dos números λ tais que $R_\lambda(T)$ existe (limitado ou não) mas não satisfaz (R3), ou seja, o domínio de $R_\lambda(T)$ não é denso em X.

No caso de dimensão finita, as propriedades (R2) e (R3) são satisfeitas para qualquer operador, portanto temos que $\sigma_c(T) = \sigma_r(T) = \emptyset$. A tabela abaixo irá auxiliar a compreender as propriedades de cada conjunto discutido na definição anterior.

Satisfaz	Não satisfaz	λ pertence
(R1),(R2) e (R3)		$\rho(T)$
	(R1)	$\sigma_p(T)$
(R1) e (R3)	(R2)	$\sigma_c(T)$
(R1)	(R3)	$\sigma_r(T)$

Note que os conjuntos acima são disjuntos e vale

$$\mathbb{C} = \rho(T) \cup \sigma(T) = \rho(T) \cup \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T).$$

Veja que se $T_\lambda x = (T - \lambda I)x = 0$ para algum $x \neq 0$, então $\lambda \in \sigma_p(T)$, ou seja, λ é autovalor de T. O elemento x é chamado de *autovetor* de T (ou *autofunção* de T se X é um espaço de funções) correspondente ao autovalor λ. O subespaço de $\mathscr{D}(T)$ formado por $x = 0$ e todos os autovalores de T correspondentes a λ é dito *autoespaço* de T relacionado ao autovalor λ.

O exemplo a seguir irá nos auxiliar a compreender algumas particularidades dos valores espectrais em espaços de dimensão infinita. Veremos que T pode possuir valores espectrais que não são autovalores.

Exemplo 1.5.2. No espaço de Hilbert $X = l^2$ definimos o operador linear $T : l^2 \longrightarrow l^2$ por

$$(\xi_1, \xi_2, \ldots) \longmapsto (0, \xi_1, \xi_2, \ldots),$$

onde $x = (\xi_j) \in l^2$. O operador T é chamado de *operador de deslocamento a direita*. O operador T é limitado e $\|T\| = 1$ pois:

$$\|Tx\|^2 = \sum_{j=1}^{\infty} |\xi_j|^2 = \|x\|^2 .$$

A imagem de T é o conjunto $Im(T) = \{(\xi_j) \in l^2 : \xi_1 = 0\}$. Portanto, o operador inverso $R_0(T) = T^{-1} : Im(T) \longrightarrow X$ existe e é dado pelo *operador de deslocamento a esquerda* definido por

$$(\xi_1, \xi_2, \ldots) \longmapsto (\xi_2, \xi_3, \ldots).$$

Vejamos que R_0 não satisfaz (R3), pois $Im(T)$ não é denso em X. De fato, se $\tilde{x} = (\tilde{x_1}, \tilde{x_2}, \ldots) \in l^2$ é tal que $\tilde{x_1} \neq 0$, então para $0 < \epsilon < |\tilde{x_1}|^2$ temos que $Im(T) \cap B(\tilde{x}; \epsilon) = \emptyset$, pois

$$\|\tilde{x} - y\| = |\tilde{x_1}|^2 + \sum_{k=2}^{\infty} |\tilde{x}_k - y_k|^2 > \epsilon, \text{ para todo } y = (0, y_2, y_3, \ldots) \in Im(T).$$

Logo $\lambda = 0$ é um valor espectral de T. Porém, $\lambda = 0$ não é autovalor, pois $Tx = 0$ implica $x = 0$. Mais precisamente, nesse caso temos que $\lambda = 0$ pertence ao espectro residual $\sigma_r(T)$.

Agora discutiremos algumas propriedades espectrais de operadores lineares limitados.

Teorema 1.5.3. *Seja $T \in B(X)$, onde X é um espaço de Banach. Se $\|T\| < 1$ então $(I - T)^{-1}$*

existe, é um operador linear limitado no espaço X e

$$(I-T)^{-1} = \sum_{j=0}^{\infty} T^j = I + T + T^2 + \cdots . \tag{1.18}$$

A série (1.18) é convergente na norma do espaço $B(X)$.

Demonstração. Temos por (1.17) que $\|T^j\| \leq \|T\|^j$. Note que que a série geométrica $\sum_{j=0}^{\infty} \|T\|^j$ converge absolutamente pois $\|T\| < 1$. Assim, a série (1.18) converge absolutamente quando $\|T\| < 1$. Além disso, como X é completo, então $\mathrm{B}(X)$ também o é. Logo, a convergência absoluta implica na convergência da série (1.18).

Vamos denotar a série de (1.18) por S. Logo, temos

$$\begin{aligned}
(I-T)(I+T+\cdots+T^n)& \\
&= (I+T+\cdots+T^n)(I-T) \\
&= I - T^{n+1}.
\end{aligned}$$

Temos que $T^{n+1} \xrightarrow{n\to\infty} 0$ em $B(X)$ pois é o termo geral de uma série convergente. Portanto,

$$(I-T)S = S(I-T) = I,$$

donte obtemos $S = (I-T)^{-1}$. $\square$

O teorema a seguir é uma importante aplicação do resultado que acabamos de provar.

Teorema 1.5.4. *O conjunto resolvente $\rho(T)$ de um operador linear limitado T definido em um espaço de Banach complexo X é aberto. Logo, o espectro $\sigma(T)$ é fechado.*

Demonstração. Se $\rho(T) = \emptyset$, então é um conjunto aberto. Suponha que $\rho(T) \neq \emptyset$. Para um $\lambda_0 \in \rho(T)$ fixo e um $\lambda \in \mathbb{C}$ qualquer segue

$$\begin{aligned}
T - \lambda I &= T - \lambda_0 I - (\lambda - \lambda_0) I \\
&= (T-\lambda_0 I)\left[I - (\lambda-\lambda_0)(T-\lambda_0 I)^{-1}\right].
\end{aligned}$$

Denotando $V = I - (\lambda-\lambda_0)(T-\lambda_0 I)^{-1} = I - (\lambda-\lambda_0)R_{\lambda_0}$ segue que

$$T_\lambda = T_{\lambda_0} V.$$

Como $\lambda_0 \in \rho(T)$ e T é limitado, segue do Lema 7.2-3 em [4] que $R_{\lambda_0} = T_{\lambda_0}^{-1} \in \mathrm{B}(X)$. Pelo Teorema 1.5.3, existe V^{-1}, para todo λ tal que $\left\|(\lambda - \lambda_0)R_{\lambda_0}\right\| < 1$, ou seja,

$$|\lambda - \lambda_0| < \frac{1}{\left\|R_{\lambda_0}\right\|}. \tag{1.19}$$

Pelo mesmo teorema, segue que

$$V^{-1} = \sum_{j=0}^{\infty} \left[(\lambda - \lambda_0)R_{\lambda_0}\right]^j = \sum_{j=0}^{\infty} (\lambda - \lambda_0)^j R_{\lambda_0}^j,$$

em $\mathrm{B}(X)$. Como $T_{\lambda_0}^{-1} = R_{\lambda_0} \in \mathrm{B}(X)$, temos para todo λ que satisfaz (1.19) que o operador T_λ tem inversa dada por

$$R_\lambda = T_\lambda^{-1} = (T_{\lambda_0} V)^{-1} = V^{-1} R_{\lambda_0}.$$

Logo, (1.19) nos dá uma vizinhança de λ_0 consistindo em valores regulares λ de T. Como $\lambda_0 \in \rho(T)$ era arbitrário, temos que $\rho(T)$ é aberto, logo seu complementar $\sigma(T) = \mathbb{C} \setminus \rho(T)$ é fechado.

$\square$

O teorema a seguir segue diretamente do Teorema 1.5.4.

Teorema 1.5.5. *Para um espaço complexo de Banach X, um operador linear limitado T e todo $\lambda_0 \in \rho(T)$ o resolvente $R_\lambda(T)$ tem representação:*

$$R_\lambda = \sum_{j=0}^{\infty} (\lambda - \lambda_0)^j R_{j_0}^{j+1},$$

onde essa série é absolutamente convergente para todo λ no disco aberto (em $\mathbb{C}$) dado por:

$$|\lambda - \lambda_0| < \frac{1}{\left\|R_{\lambda_0}\right\|}.$$

Esse disco é um subconjunto de $\rho(T)$.

O resultado a seguir também é consequência do Teorema 1.5.3.

Teorema 1.5.6. *O espectro $\sigma(T)$ de um operador linear limitado $T : X \longrightarrow X$ em um espaço complexo de Banach X é compacto e pertence ao disco dado por:*

$$|\lambda| \leq \|T\|. \tag{1.20}$$

Logo, o conjunto resolvente $\rho(T)$ não é vazio.

Demonstração. Sejam $\lambda \neq 0$ e $\tau = \frac{1}{\lambda}$. Pelo Teorema 1.5.3 obtemos uma representação

$$R_\lambda = (T - \lambda I)^{-1} = -\frac{1}{\lambda}(I - \tau T)^{-1} = -\frac{1}{\lambda}\sum_{j=0}^{\infty}(\tau T)^j = -\frac{1}{\lambda}\sum_{j=0}^{\infty}\left(\frac{1}{\lambda}T\right)^j,$$

sendo que pelo Teorema 1.5.3 essa série é convergente para todo λ tal que

$$\frac{\|T\|}{|\lambda|} = \left\|\frac{1}{\lambda}T\right\| < 1 \quad \Longleftrightarrow \quad |\lambda| > \|T\|.$$

Também segue que qualquer λ que satisfaz essa propriedade pertence à $\rho(T)$. Logo, o espectro $\sigma(T) = \mathbb{C} \setminus \rho(T)$ pertence ao disco (1.20), assim $\sigma(T)$ é limitado. Além disso, $\sigma(T)$ é fechado pelo Teorema 1.5.4, portanto $\sigma(T)$ é compacto. $\square$

Capítulo 2

Espaços de Hilbert

No capítulo anterior estudamos espaços vetoriais normados, conseguindo generalizar as noções de distâncias e comprimento de vetor. Neste capítulo estamos interessados em estudar *espaços com produto interno*, em outras palavras, espaços vetoriais com uma operação análoga ao *produto escalar* visto em espaços de dimensão finita.

Esses espaços, quando completos, serão chamados de *espaços de Hilbert*. Além disso, veremos noções de *perpendicularidade* e generalizamos alguns resultados da Álgebra Linear, além de novos resultados, que terão importantes aplicações no último capítulo.

2.1 Espaços com produto interno e espaços de Hilbert

Um *produto interno* em um espaço vetorial X, definido sobre um corpo $\mathbb{K}$, é uma função

$$\begin{aligned} \langle\ ,\ \rangle\ : X \times X &\longrightarrow\ \mathbb{K} \\ (x,y) &\longmapsto\ \langle x,y\rangle, \end{aligned}$$

que satisfaz as seguintes propriedades:

$$\langle x+y,z\rangle = \langle x,z\rangle + \langle y,z\rangle; \tag{2.1}$$

$$\langle \alpha x,y\rangle = \alpha\langle x,y\rangle; \tag{2.2}$$

$$\langle x,y\rangle = \overline{\langle y,x\rangle},\ \ \langle x,x\rangle \geq 0; \tag{2.3}$$

$$\langle x,x\rangle = 0 \Leftrightarrow x = 0. \tag{2.4}$$

Se X for um espaço vetorial com produto interno, diremos simplesmente que X é um *espaço com produto interno*. Veremos mais a frente que uma norma em X é obtida da seguinte forma:

$$\|x\| \doteq \sqrt{\langle x, x\rangle}, \quad x \in X. \tag{2.5}$$

Consequentemente, definimos a distância de dois pontos em X por:

$$d(x,y) = \|x-y\| = \sqrt{\langle x-y, x-y\rangle}, \quad x, y \in X.$$

Portanto, X é um espaço métrico com a *métrica induzida* pelo produto interno. Quando X for um espaço com produto interno que é completo em relação a essa métrica, dizemos que X é um *espaço de Hilbert*.

Verifiquemos agora que, de fato, a expressão (2.5) define uma norma em X. Dados $x \in X$ e $\alpha \in \mathbb{K}$ temos que

- $\|x\| \geq 0$ segue diretamente da definição.
- $\|\alpha x\|^2 = \langle \alpha x, \alpha x\rangle = \alpha\overline{\alpha}\langle x, x\rangle = |\alpha|^2\|x\|^2 \Longrightarrow \|\alpha x\| = |\alpha|\|x\|$.
- A desigualdade triangular seguirá do lema a seguir.

Lema 2.1.1. *Um produto interno e sua norma correspondente satisfazem:*

1. $|\langle x, y\rangle| \leq \|x\|\ \|y\|$ *(Desigualdade de Schwarz).*

   A igualdade é valida se, e somente se, $\{x, y\}$ *é linearmente dependente.*

2. $\|x+y\| \leq \|x\| + \|y\|$ *(Desigualdade triangular).*

Demonstração. 1. Se $y = 0$ o resultado é imediato. Suponhamos então $y \neq 0$. Segue para todo α escalar que

$$0 \leq \|x-\alpha y\|^2 = \langle x-\alpha y, x-\alpha y\rangle = \langle x, x\rangle - \overline{\alpha}\langle x, y\rangle - \alpha\left[\langle y, x\rangle - \overline{\alpha}\langle y, y\rangle\right].$$

A última parte, em colchetes, irá ser zero se escolhemos $\overline{\alpha} = \dfrac{\langle y, x\rangle}{\langle y, y\rangle}$. Logo,

$$0 \leq \|x\|^2\|y\|^2 - |\langle x, y\rangle|^2 \Longleftrightarrow |\langle x, y\rangle|^2 \leq \|x\|^2\|y\|^2 \Longleftrightarrow |\langle x, y\rangle| \leq \|x\|\|y\|.$$

Suponha $|\langle x, y\rangle| = \|x\| \cdot \|y\|$, então teremos que $y = 0$ (não é o caso) ou $\|x - \alpha y\| = 0$, ou seja, $x = \alpha y$, logo $\{x, y\}$ é linearmente dependente.

2. Temos que

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

Pelo item anterior $|\langle x, y\rangle| = |\langle y, x\rangle| \leq \|x\| \cdot \|y\|$, usando essa desigualdade segue que

$$\|x+y\|^2 \leq \|x\|^2 + 2|\langle x, y\rangle| + \|y\|^2 \leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 = \big(\|x\| + \|y\| \big)^2 .$$

Portanto $\|x+y\| \leq \|x\| + \|y\|$.

□

A seguir estabelecemos algumas fórmulas úteis, estas não serão demonstradas no presente trabalho, mas a demonstração para a maioria é simples sendo a mais complicada a identidade do paralelogramo que pode ser encontrada em [8].

- $\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle$;
- $\langle x, \alpha y\rangle = \overline{\alpha}\langle x, y\rangle$;
- $\langle x, \alpha y + \beta z\rangle = \overline{\alpha}\langle x, y\rangle + \overline{\beta}\langle x, z\rangle$;
- $\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$ (lei do paralelogramo).

Esta última igualdade é interessante para testar se uma norma é induzida de um produto interno, ou seja, se uma norma não satisfaz essa igualdade, então ela não é induzida de um produto interno.

Definição 2.1.2. *Sejam X espaço com produto interno e $x \in X$, dizemos que x é ortogonal à $y \in X$ se $\langle x, y\rangle = 0$ e denotamos por $x \perp y$.*

Se $A, B \subset X$, escrevemos $x \perp A$ se $x \perp a$, para todo $a \in A$ e $A \perp B$ se $a \perp b$, para todo $a \in A$ e $b \in B$.

Vejamos agora alguns exemplos de produto interno.

Exemplos 2.1.3. 1. $\mathbb{R}^n$:

Sendo $x = (\xi_j) = (\xi_1, \ldots, \xi_n)$ e $y = (\eta_j) = (\eta_1, \ldots, \eta_n)$ elementos de $\mathbb{R}^n$, definimos o produto interno de x por y por

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

Assim,

$$\begin{aligned} \|x\| &= \langle x, x\rangle^{\frac{1}{2}} = \left(\xi_1^2 + \cdots + \xi_n^2\right)^{\frac{1}{2}} \quad \text{e} \\ d(x,y) &= \|x - y\| = \langle x - y, x - y\rangle^{\frac{1}{2}} = \left[(\xi_1 - \eta_1)^2 + \cdots + (\xi_n - \eta_n)^2\right]. \end{aligned}$$

Já vimos que $\mathbb{R}^n$ é completo com essa norma.

2. $\mathbb{C}^n$:

Tome $x = (\xi_j)$ e $y = (\eta_j)$ elementos de $\mathbb{C}^n$, definimos o produto interno entre x e y por

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

Temos,

$$\|x\| = \left(\xi_1\overline{\xi_1} + \cdots + \xi_n\overline{\xi_n}\right)^{\frac{1}{2}} = \left(|\xi_1|^2 + \cdots + |\xi_n|^2\right)^{\frac{1}{2}}.$$

Também vimos que $\mathbb{C}^n$ é completo.

3. l^2 :

Sejam $x = (\xi_j) = (\xi_1, \ldots, \xi_n, \ldots)$ e $y = (\eta_j) = (\eta_1, \ldots, \eta_n, \ldots)$, definimos um produto interno em l^2 por

$$\langle x, y\rangle = \sum_{j=1}^{\infty} \xi_j\overline{\eta_j}.$$

A série é convergente (segue por Cauchy-Schwarz), e temos que

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

Além disso, l^2 é completo.

4. l^p, $p \neq 2$:

Temos que l^p, $p \neq 2$ não é de Hilbert, com a norma

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

Iremos, na verdade, provar que ℓ^p não é um espaço com produto interno na norma acima, concluindo que não satisfaz a lei do paralelogramo.

Tome $x = (1, 1, 0, 0, \ldots) \in l^p$ e $y = (1, -1, 0, 0, \ldots) \in l^p$.

Segue que $\|x\| = \|y\| = 2^{\frac{1}{p}}$ e $\|x + y\| = \|x - y\| = 2$.

Então,

- $\|x + y\|^2 + \|x - y\|^2 = 2^2 + 2^2 = 8 \quad$ e
- $2(\|x\|^2 + \|y\|^2) = 2(4^{\frac{1}{p}} + 4^{\frac{1}{p}})$, essa última é diferente de 8 para $p \neq 2$.

Como as expressões são diferentes, segue o resultado.

Listaremos mais algumas propriedades do produto interno, as quais serão importantes em futuros resultados.

- $\langle x, y \rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)$ (produto interno real);
- $Re\langle x, y \rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)$;
- $Im\langle x, y \rangle = \frac{1}{4}(\|x + iy\|^2 - \|x - iy\|^2)$.

As duas últimas são conhecidas como *identidades de polarização.*

Lema 2.1.4. *Em um espaço com produto interno, se $x_n \longrightarrow x$ e $y_n \longrightarrow y$, então temos que $\langle x_n, y_n \rangle \longrightarrow \langle x, y \rangle$.*

Demonstração. Usando propriedades do produto interno, a desigualdade triangular de números e a desigualdade de Schwarz temos

$$\begin{aligned} |\langle x_n, y_n \rangle - \langle x, y \rangle| &= |\langle x_n, y_n \rangle - \langle x_n, y \rangle + \langle x_n, y \rangle - \langle x, y \rangle| \\ &\leq |\langle x_n, y_n - y \rangle| + |\langle x_n - x, y \rangle| \\ &\leq \|x_n\|\|y_n - y\| + \|x_n - x\|\|y\| \xrightarrow{n \to \infty} 0, \end{aligned}$$

uma vez que $\|y_n - y\| \xrightarrow{n\to\infty} 0$, $\|x_n - x\| \xrightarrow{n\to\infty} 0$ e (x_n) é limitada, pois é uma sequência convergente.

$\square$

Definição 2.1.5. *Sejam X e Y espaços com produto interno sob o mesmo corpo. Um operador $T : X \longrightarrow Y$ é isomorfismo se é bijetivo e além disso o operador T preserva produto interno, isto é,*

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

para todo $x, y \in X$. Nesse caso, dizemos que X é isomorfo à Y.

Observação 2.1.6. Seja $Y \subset X$ um subespaço vetorial, então Y é um espaço com produto interno (o mesmo de X, mas restrito à $Y\times Y$). Se X for um espaço de Hilbert não necessariamente Y será espaço de Hilbert.

Teorema 2.1.7. *Seja $Y \subset H$ um subespaço do espaço de Hilbert H, então:*

1. *Y é completo se, e somente se, Y é fechado em H;*
2. *se $\dim Y < \infty$, então Y é completo.*

A demonstração do item (a), no teorema acima, segue diretamente de 1.2.6, enquanto que o item (b) segue de 1.2.10.

Agora continuaremos discutindo alguns conceitos básicos que darão mais ferramentas para a discussão sobre espaços de Hilbert.

Definição 2.1.8. *Seja X espaço métrico, a distância δ entre $x \in X$ e $M \subset X$, um subconjunto, é dada por:*

$$\delta = \inf_{\tilde{y}\in M} d(x, \tilde{y}) \quad (M \neq \emptyset).$$

Em um espaço normado, essa distância será dada por:

$$\delta = \inf_{\tilde{y}\in M} \|x - \tilde{y}\|.$$

Vale ressaltar que nem sempre existe $y \in M$ tal que $\delta = \|x - y\|$. Basta por exemplo, considerar na reta real o subconjunto M como um intervalo aberto e x um ponto fora desse intervalo.

Definição 2.1.9. *Em um espaço vetorial X definimos o segmento que une x e y por*

$$\{z \in X \mid z = \alpha x + (1-\alpha)y\} \quad (0 \leq \alpha \leq 1).$$

Definição 2.1.10. *Um subconjunto $M \subset X$ é convexo se para todo $x, y \in M$ o segmento que une x e y pertence à M.*

Obviamente todo subespaço vetorial é um conjunto convexo. Além disso, se $M \cap M'$ são convexos, então $M \cap M'$ também é convexo.

Teorema 2.1.11. *Seja X um espaço com produto interno e $M \neq \emptyset$ um subconjunto convexo completo, na métrica induzida pelo produto interno. Então, para cada $x \in X$, existe um único $y \in M$ tal que*

$$\delta = \inf_{\tilde{y} \in M} \|x - \tilde{y}\| = \|x - y\|.$$

Demonstração. (Existência)

Pela definição de inf, existe uma sequência (y_n) em M tal que $\delta_n \to \delta$, sendo $\delta_n = \|x - y_n\|$. Vamos mostrar que y_n é de Cauchy.

Defina v_n por $v_n = y_n - x$, assim $\|v_n\| = \delta_n$, então

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

A última desigualdade segue pois $\frac{1}{2}(y_n + y_m) \in M$ devido à M ser convexo.

Além disso $y_n - y_m = v_n - v_m$, agora usando a lei do paralelogramo obtemos:

$$\begin{aligned}\|y_n - y_m\|^2 = \|v_n - v_m\|^2 &= -\|v_n + v_m\|^2 + 2\left(\|v_n\|^2 + \|v_m\|^2\right)\\ &\leq -(2\delta)^2 + 2({\delta_n}^2 + \delta_m^2) \longrightarrow -4\delta^2 + 2(\delta^2 + \delta^2) = 0.\end{aligned}$$

Portanto y_n é de Cauchy, logo é convergente uma vez que M é completo, digamos $y_n \to y \in M$. Como $y \in M$, temos que $\|x - y\| \geq \delta$.

Também temos $\|x - y\| \leq \|x - y_n\| + \|y_n - y\| = \delta_n + \|y_n - y\| \longrightarrow \delta$. Ou seja, $\|x - y\| \leq \delta$, logo $\|x - y\| = \delta$.

(Unicidade)

Suponha que existem $y, y_0 \in M$ tais que $\|x-y\| = \delta$ e $\|x-y_0\| = \delta$, pela lei do paralelogramo:

$$\begin{aligned}\|y-y_0\|^2 &= \left\|(y-x)-(y_0-x)\right\|^2 \\ &= 2\|y-x\|^2 + 2\|y_0-x\|^2 - \left\|(y-x)+(y_0-x)\right\|^2 \\ &= 2\delta^2 + 2\delta^2 - 2^2\left\|\frac{1}{2}(y+y_0)-x\right\|^2.\end{aligned}$$

Como M é convexo, segue que $\frac{1}{2}(y+y_0) \in M$, então $\left\|\frac{1}{2}(y+y_0)-x\right\| \geq \delta$, portanto $\|y-y_0\|^2 \leq 2\delta^2 + 2\delta^2 - 2^2 \cdot \delta^2 = 0$. Logo $\|y-y_0\| = 0$, donde segue que $y = y_0$.

□

Lema 2.1.12. *Sejam $Y = M$ como no teorema anterior, sendo Y subespaço completo e $x \in X$ fixo. Então $z = x - y$ é ortogonal à Y, para todo $y \in Y$.*

Demonstração. Suponha, por contradição, que a propriedade $z \perp y$ é falsa, então existe $y_1 \in Y$ tal que $\langle z, y_1\rangle = \beta \neq 0$.

Claramente, $y_1 \neq 0$. Tomando um α escalar temos:

$$\begin{aligned}\|z-\alpha y_1\|^2 = \langle z-\alpha y_1, z-\alpha y_1\rangle &= \langle z,z\rangle - \overline{\alpha}\langle z, y_1\rangle - \alpha\left[\langle y_1, z\rangle - \overline{\alpha}\langle y_1, y_1\rangle\right] \\ &= \langle z,z\rangle - \overline{\alpha}\beta - \alpha\left[\overline{\beta} - \overline{\alpha}\langle y_1,y_1\rangle\right].\end{aligned}$$

Escolhendo $\overline{\alpha} = \dfrac{\overline{\beta}}{\langle y_1, y_1\rangle}$, segue que $\|z-\alpha y_1\|^2 = \langle z,z\rangle - \overline{\alpha}\beta = \langle z,z\rangle - \dfrac{|\beta|^2}{\langle y_1,y_1\rangle}$.

Do Teorema 2.1.11, segue que $\|z\| = \|x-y\| = \delta$. Portanto,

$$\|z-\alpha y_1\|^2 = \|z\|^2 - \frac{|\beta|^2}{\langle y_1,y_1\rangle} = \delta^2 - \frac{|\beta|^2}{\langle y_1,y_1\rangle} < \delta^2.$$

Observação: A última desigualdade é verdadeira pois $z - \alpha y_1 = x - (\alpha y_1 + y)$ e $(\alpha y_1 + y) \in Y$ então $\|z-\alpha y_1\|^2 \geq \delta^2$. Logo, $\|z-\alpha y_1\|^2 \geq \delta$ só acontecerá quando $\dfrac{|\beta|^2}{\langle y_1,y_1\rangle} = 0$, ou seja, $|\beta| = 0$ e, então, $\beta = 0$, uma contradição. Finalmente, concluímos que $\langle z, y\rangle = 0$ para todo $y \in Y$.

□

Agora iremos discutir alguns resultados que serão importantes para entender a estrutura dos espaços de Hilbert. Faremos uma abordagem análoga a que é comumente vista em Álgebra

Linear, utilizando a ortogonalidade como meio de representar, ou aproximar, elementos em um espaço de Hilbert.

Definição 2.1.13. *Um espaço vetorial X é dito soma direta dos espaços vetoriais Y e Z se $Y \cap Z = \{0\}$ e para todo $x \in X$, existem únicos $y \in Y$ e $z \in Z$ tais que $x = y + z$.*

Nesse caso usamos a notação $X = Y \oplus Z$. Dizemos que Z é o complemento algébrico de Y em X e Y é o complemento algébrico de Z em Y. Além disso, (Y, Z) é chamado de par complementar de subespaços em X.

Exemplo 2.1.14. Um exemplo clássico é tomar o espaço $\mathbb{R}^2$ (usando o produto interno usual) como sendo soma direta dos espaços $\operatorname{span}\{(1,0)\}$ e $\operatorname{span}\{(0,1)\}$, ou seja,

$$\mathbb{R}^2 = \operatorname{span}\{(1,0)\} \oplus \operatorname{span}\{(0,1)\}.$$

Definição 2.1.15. *Dado $Y \subset H$ subespaço vetorial fechado, onde H é espaço de Hilbert, definimos o complemento ortogonal de Y, denotado $Y^\perp$, do seguinte modo:*

$$Y^\perp \doteq \{\, z \in H \mid z \perp Y\}.$$

Teorema 2.1.16. *Seja $Y \subset H$ subespaço fechado do espaço de Hilbert H. Então, $H = Y \oplus Y^\perp$.*

Demonstração. Como H é completo e Y é fechado, segue que Y é completo pelo Teorema 1.2.6. Como Y é convexo, pois é subespaço vetorial, segue de 2.1.11 e 2.1.12 que para todo $x \in H$, existe $y \in Y$ tal que

$$x = y + z, \text{ onde } z \in Y^\perp. \tag{2.6}$$

Suponha que $x = y + z = y_1 + z_1$, então $y + z = y_1 + z_1$, assim $y - y_1 = z_1 - z = w$. Como Y e $Y^\perp$ são espaços vetoriais, segue que $w \in Y \cap Y^\perp$, logo $w = 0$, ou seja, $y = y_1$ e $z = z_1$.

□

Chamamos y em (2.6) de *projeção ortogonal* de x em y. Definiremos projeção agora com mais detalhes.

Definição 2.1.17. *Considere*

$$P : H \longrightarrow Y$$
$$x = y + z \longmapsto y,$$

onde $H = Y \oplus Y^{\perp}$. P é chamado de projeção ou operador projeção de H em Y.

Observação 2.1.18. Segue diretamente da definição de P que

1. $P(H) = Y$;
2. $P(Y) = Y$;
3. $P(Y^{\perp}) = 0$;
4. $P^2 = P$.

Lema 2.1.19. *O complemento ortogonal $Y^{\perp}$ de Y, onde Y é subespaço fechado de H (H é de Hilbert) é o núcleo de P, ou seja, $Y^{\perp} = \ker(P)$, onde $P : H \longrightarrow Y$ é projeção ortogonal.*

Definição 2.1.20. *Definimos o aniquilador de um conjunto $M \neq \emptyset$, denotado por $M^{\perp}$ o conjunto $M^{\perp} = \{x \in X \mid x \perp M\}$, portanto se $x \in M^{\perp}$ $\langle x, v\rangle = 0$, para todo $v \in M$.*

Observação 2.1.21. $M^{\perp}$ é espaço vetorial, fechado e $(M^{\perp})^{\perp} \doteq M^{\perp\perp}$. Temos que $M \subset M^{\perp\perp}$, pois $x \in M$ implica em $x \perp M^{\perp}$, logo $x \in (M^{\perp})^{\perp}$.

Lema 2.1.22. *Sejam H um espaço de Hilbert e $Y \subset H$ subespaço fechado, então $Y = Y^{\perp\perp}$.*

Demonstração. Temos que $Y \subset Y^{\perp\perp}$, resta mostrar que $Y^{\perp\perp} \subset Y$.

Podemos escrever $H = Y \oplus Y^{\perp\perp}$, pois Y é fechado. Agora considere $x \in Y^{\perp\perp}$, então $x = y+z$, onde $y \in Y \subset Y^{\perp\perp}$. Também temos que $z = x - y \in Y^{\perp\perp}$ (pois $x, y \in Y^{\perp\perp}$ e $Y^{\perp\perp}$ é espaço vetorial).

Assim, $z \perp Y^{\perp}$, mas $z \in Y^{\perp}$, portanto $z \perp z$, então $z = 0$ e segue que $x = y \in Y$. □

Observação 2.1.23. Denotando $Z = Y^{\perp}$, podemos escrever $H = Z \oplus Z^{\perp}$ e de modo análogo existe uma projeção $P_z : H \longrightarrow Z$.

Lema 2.1.24. *Para qualquer subconjunto $M \neq \emptyset$ de um espaço de Hilbert H, temos que spanM é denso em H se, e somente se, $M^{\perp} = \{0\}$.*

Demonstração. Seja $x \in M^{\perp}$ e suponha que $V = \operatorname{span} M$ é denso em H. Portanto $\overline{V} = H$ e $x \in \overline{V} = H$.

Existe uma sequência (x_n) em V tal que $x_n \longrightarrow x$, portanto $\langle x_n, x\rangle = 0$ para todo $n \in \mathbb{N}$.

Pela continuidade do produto interno, $\langle x_n, x\rangle \longrightarrow \langle x, x\rangle$, logo $\langle x, x\rangle = \|x\|^2 = 0$, assim $x = 0$, ou seja, $M^{\perp} = \{0\}$.

Por outro lado, suponha que $M^{\perp} = \{0\}$.

Se $x \perp V$, temos que $\langle x, y\rangle = 0$ para todo $y \in \operatorname{span} M = V$, em especial para todo $y \in M$, assim $x \perp M$, portanto $x \in M^{\perp}$ e segue que $x = 0$.

Logo, como $V \subseteq H$ é subespaço, segue que $H = \overline{V} \oplus \overline{V}^{\perp}$ e $\overline{V}^{\perp} = 0$, deste modo $H = \overline{V}$. Observação: Para concluir que $\overline{V}^{\perp} = \{0\}$ usamos que se $V^{\perp} = \{0\}$ então $\overline{V}^{\perp} = 0$. $\square$

2.2 Conjunto e sequências ortonormais

Definição 2.2.1. *Dizemos que $M \subset X$ é um conjunto ortogonal em um espaço com produto interno X, se os elementos de M são dois-a-dois ortogonais. Dizemos que M é ortonormal se seus elementos tem norma igual a 1, ou seja,*

$$\langle x, y\rangle = \begin{cases} 0, & x \neq y \\ 1, & x = y. \end{cases}$$

Dessa forma, um conjunto indexado ou uma família (x_α), $\alpha \in I$, é *ortogonal* se $x_\alpha \perp x_\beta$ para todo $\alpha, \beta \in I, \alpha \neq \beta$. É *ortonormal* se é ortogonal e

$$\langle x_\alpha, x_\beta\rangle = \delta_{\alpha\beta} = \begin{cases} 0, & \alpha \neq \beta \\ 1, & \alpha = \beta. \end{cases}$$

Em particular, se M é contável podemos arranjá-lo em uma sequência (x_n), essa sequência é dita *sequência ortogonal* ou *sequência ortonormal* no caso em que seus elementos tem norma igual a 1.

Uma propriedade interessante que os conjuntos ortogonais respeitam é a *relação de Pitágoras* dada pela expressão abaixo,

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2,$$

onde x e y são ortogonais.

De modo geral, se $\{x_1, x_2, \ldots, x_n\}$ é um conjunto ortogonal, então:

$$\|x_1 + \cdots + x_n\|^2 = \|x_1\|^2 + \cdots + \|x_n\|^2.$$

De fato, se $\langle x_j, x_k\rangle = 0$, $k \neq j$, teremos

$$\begin{aligned}\|x_1 + \cdots x_n\|^2 &= \langle x_1 + \cdots + x_n,\ x_1 + \cdots + x_n\rangle \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\langle x_j, x_j\rangle = \sum_{i=1}^{n}\langle x_i, x_i\rangle = \sum_{i=1}^{n}\|x_i\|^2 .\end{aligned}$$

Lema 2.2.2. *Todo conjunto ortonormal é linearmente independente.*

Demonstração. Sejam $x_1, x_2, \ldots, x_n$ elementos quaisquer do conjunto (onde $x_i \neq x_j, sei \neq j$) e considere uma combinação linear nula qualquer: $\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = 0$.

Basta tomar o produto interno da expressão acima com x_i para cada $i = 1, \ldots, n$, ou seja:

$$\begin{aligned}0 &= \langle x_i, \alpha_1 x_1 + \cdots + \alpha_i x_i + \cdots + \alpha_n x_n\rangle \\ &= \alpha_1\langle x_i, x_1\rangle + \cdots + \alpha_i\langle x_i, x_i\rangle + \cdots + \alpha_n\langle x_i, x_n\rangle \\ &= \alpha_i\langle x_i, x_i\rangle = \alpha_i.\end{aligned}$$

Segue que todos α_i são nulos, logo $x_1, x_2, \ldots, x_n$ são linearmente independentes. □

Exemplos 2.2.3.

1. Considere $\mathbb{R}^3$ com o produto interno usual, o conjunto

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

é ortonormal e, pelo lema acima, segue que é linearmente independente. Além disso, este conjunto é conhecido como *base canônica de* $\mathbb{R}^3$, pois é o conjunto ortonormal mais simples de ser obtido.

2. Considere no espaço l^2 a sequência (e_n), onde $e_n = (\delta_{nj})_{j\in\mathbb{N}}$. A sequência (e_n) é ortonormal, utilizando o produto interno definido na seção anterior, isto é,

$$\langle e_n, e_m\rangle = \sum_{j=1}^{\infty} e_m\overline{e_n} \quad (m \neq n).$$

Como o produto interno em l^2 é feito de entrada em entrada, se na entrada n de e_n estiver 1, em e_m terá 0, e na entrada m de e_m terá 1 enquanto em e_n terá 0, e as outras entradas são 0 pela definição da função Delta de Kronecker, logo todas as entradas serão 0 e segue a afirmação.

3. Seja X o espaço das funções reais contínuas em $[0, 2\pi]$ com produto interno

$$\langle x, y\rangle = \int_0^{2\pi} x(t)y(t)\ dt.$$

Uma sequência ortogonal em X é (u_n), onde

$$u_n(t) = \cos(nt), \quad n = 0, 1, 2, \ldots$$

Outra sequência ortogonal em X é v_n, onde

$$v_n(t) = \text{sen}(nt), \quad n = 0, 1, 2, \ldots$$

De fato, vamos provar que (u_n) é ortogonal. Para isso, vamos integrar

$$\langle u_m, u_n\rangle = \int_0^{2\pi} \cos(mt)\cos(nt)\ dt.$$

Usaremos a expressão $\cos(A)\cos(B) = \dfrac{\cos(A+B) + \cos(A-B)}{2}$ para $A \neq B$.

Vamos dividir em casos:

(a) Se $m = 0 = n$: temos que $\langle u_m, u_n\rangle = \int_0^{2\pi} \cos(0)\cos(0)\ dt = \int_0^{2\pi} 1\ dt = 2\pi$.

(b) Se $m = n \neq 0$: $\langle u_m, u_n\rangle = \int_0^{2\pi} cos^2(mt)\ dt = \frac{1}{2}\left[t + \text{sen}(2t)\right]_0^{2\pi} = \pi$.

(c) Se $m \neq n$: Para resolver a integral, vamos considerar a identidade trigonométrica relembrada acima.

Assim,

$$\begin{aligned}
\langle u_m, u_n\rangle &= \int_0^{2\pi} \cos(mt)\cos(nt)\ dt \\
&= \frac{1}{2}\left[\int_0^{2\pi} \cos\left[(m+n)t\right]\ dt + \int_0^{2\pi} \cos\left[(m-n)t\right]\ dt\right] \\
&= \frac{1}{2}\left[\frac{\text{sen}\left[(m+n)t\right]}{m+n} + \frac{\text{sen}\left[(m-n)t\right]}{m-n}\right]_0^{2\pi} \\
&= \frac{1}{2}\left[\frac{\text{sen}\left[(m+n)2\pi\right]}{m+n} + \frac{\text{sen}\left[(m-n)2\pi\right]}{m-n}\right].
\end{aligned}$$

Independente dos valores de m e n a expressão anterior será igual a 0, pois $\operatorname{sen}(2\pi j) = 0$, para todo $j \in \mathbb{N}$.

Resumindo,

$$\langle u_m, u_n \rangle = \begin{cases} 0, & m \neq n \\ \pi, & m = n = 1, 2, \dots \\ 2\pi, & m = n = 0. \end{cases}$$

Obtemos um resultado parecido para v_n.

Portanto, uma sequência ortonormal (e_n) em X é dada por:

$$e_0(t) = \frac{1}{\sqrt{2\pi}},\ e_n(t) = \frac{u_n(t)}{\|u_n\|} = \frac{\cos(nt)}{\sqrt{\pi}} \quad \text{, para } n = 1, 2, \dots$$

De v_n obtemos uma sequência ortonormal $(\tilde{e}_n)$, onde

$$\tilde{e}_n(t) = \frac{v_n(t)}{\|v_n\|} = \frac{\operatorname{sen}(nt)}{\sqrt{\pi}}, \text{ para } n = 1, 2, \dots$$

Também temos que $u_m \perp v_n$ para todo $m, n \in \mathbb{N}$. Para provar, basta observar que

$$\int_0^{2\pi} \cos(mt) \operatorname{sen}(nt)\, dt = 0.$$

Abaixo vamos listar algumas vantagens da ortonormalidade.

1. *Coeficientes simples de representar:*

   Se $\{e_1, e_2, \dots, e_n\}$ é um conjunto ortonormal e $x \in \operatorname{span}\{e_1, e_2, \dots, e_n\}$, então existem α_i escalares tais que $x = \sum_{i=1}^n \alpha_i e_i$.

   Deste modo,

   $$\langle x, e_i \rangle = \sum_j \alpha_j \langle e_j, e_i \rangle = \alpha_i \langle e_i, e_i \rangle = \alpha_i.$$

   Logo,

   $$x = \sum_{j=1}^n \langle x, e_i \rangle e_i. \tag{2.7}$$

2. *Podemos representar "parte" de um elemento:*

   Dado $x \in X$ qualquer, não necessariamente em $Y_n = \operatorname{span}\{e_1, \dots, e_n\}$, definimos $y \in Y_n$ por $y = \sum_{k=1}^n \langle x, e_k \rangle e_k$, onde n está fixo.

   Agora, defina z de modo que $x = y + z$, assim $z = x - y$.

   Afirmação: $z \perp y$.

De fato, primeiramente note que $y = \sum_{k=1}^{n} \alpha_k e_k$, $\alpha_k = \langle y, e_k \rangle$. Também temos que $\alpha_k = \langle x, e_k \rangle$.

Logo, $\|y\|^2 = \langle \sum \langle x, e_k \rangle e_k, \sum \langle x, e_m \rangle e_m \rangle = \sum |\langle x, e_k \rangle|^2$.

Agora, podemos mostrar que $z \perp y$:

$$\begin{aligned}
\langle z, y \rangle = \langle x - y, y \rangle &= \langle x, y \rangle - \langle y, y \rangle \\
&= \langle x, \sum \langle x, e_k \rangle e_k \rangle - \sum |\langle x, e_k \rangle|^2 \\
&= \sum \overline{\langle x, e_k \rangle} \langle x, e_k \rangle - \sum |\langle x, e_k \rangle|^2 \\
&= \sum |\langle x, e_k \rangle|^2 - \sum |\langle x, e_k \rangle|^2 = 0.
\end{aligned}$$

Logo, dado um conjunto ortonormal de dimensão finita, sempre é possível escrever $x \in X$ como uma soma de um elemento como em 2.7 com um "resto" ortogonal a essa soma.

3. *Desigualdade de Bessel:*

   Seja (e_k) uma sequência ortonormal em um espaço com produto interno X, então para todo $x \in X$:
$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2 \leq \|x\|^2 . \tag{2.8}$$

   De fato, pela relação de Pitágoras e pelo item anterior, segue que $\|z\|^2 = \|x\|^2 - \|y\|^2 = \|x\|^2 - \sum |\langle x, e_k \rangle|^2$. Como $\|z\| \geq 0$, temos para todo $n = 1, 2, \ldots$

$$\sum_{k=1}^{n} |\langle x, e_k \rangle|^2 \leq \|x\|^2 .$$

   A soma da esquerda possui termos não negativos, logo a sequência formada pelas somas parciais será monótona não decrescente e limitada por $\|x\|^2$, o que prova a desigualdade.

Observação 2.2.4. Os produtos internos $\langle x, e_k \rangle$ são chamados de *Coeficientes de Fourier* de x com respeito a sequência ortonormal (e_k).

Observação 2.2.5. Note que para conjuntos linearmente independentes enumeráveis é sempre possível encontrar um conjunto ortonormal de mesma cardinalidade, para isso usamos o *Processo de Ortonormalização de Gram-Schmidt*, o qual não descreveremos nesse texto, mas pode ser encontrado em livros de Álgebra Linear.

2.3 Séries relacionadas a sequências e conjuntos ortonormais

Nesta seção iremos seguir o roteiro abaixo:

- Motivação para o termo "Coeficientes de Fourier";
- Estudar séries infinitas relacionadas a sequências ortonormais;
- Estudar um pouco sobre conjuntos ortogonais incontáveis.

Vamos começar discutindo as *séries de fourier*.

Uma *série trigonométrica* é uma serie da forma abaixo:

$$a_0 + \sum_{k=1}^{\infty}(a_k \cos(kt) + b_k \operatorname{sen}(kt)), \tag{2.9}$$

onde a_0, a_k, b_k são escalares.

Definição 2.3.1. *Uma função real x é periódica se existe $p > 0$, tal que $x(t+p) = x(t)$ para todo $t \in \mathbb{R}$. O número positivo p é chamado de período da função x.*

Seja x uma função contínua de período $p = 2\pi$. Por definição, a *Série de Fourier de x* é a série trigonométrica com coeficientes a_k e b_k dados pelas *fórmulas de Euler*:

$$a_0 = \frac{1}{2\pi}\int_0^{2\pi} x(t)\ dt \tag{2.10}$$

$$a_k = \frac{1}{\pi}\int_0^{2\pi} x(t)\cos(kt)\ dt \quad k = 1, 2, \ldots \tag{2.11}$$

$$b_k = \frac{1}{\pi}\int_0^{2\pi} x(t)\operatorname{sen}(kt)\ dt \quad k = 1, 2, \ldots \tag{2.12}$$

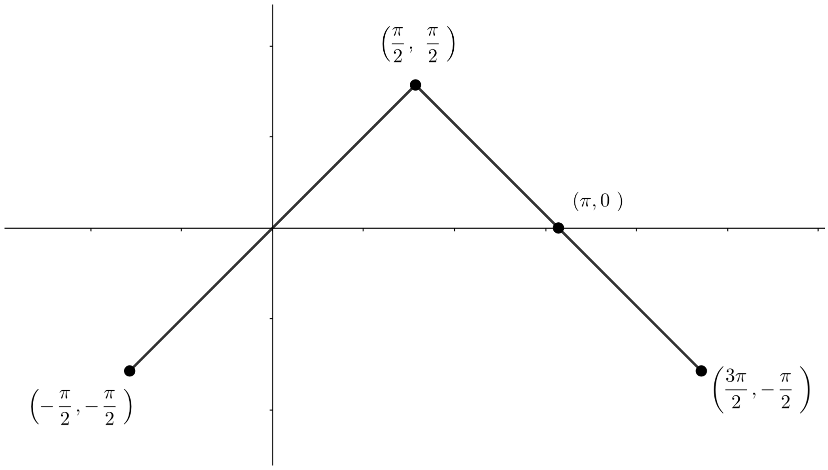
Os números a_0, a_k e b_k são chamados de *coeficientes de Fourier de x*.

Se a série de Fourier de x converge para todo t e tem soma $x(t)$, escrevemos

$$x(t) = a_0 + \sum_{k=1}^{\infty}(a_k \cos(kt) + b_k \operatorname{sen}(kt)).$$

Como x é periódica de período 2π, podemos substituir o intervalo de integração $[0, 2\pi]$ por $[-\pi, \pi]$.

Exemplo 2.3.2. Considere a função x dada por $x(t) = \begin{cases} t, & -\frac{\pi}{2} \leq t \leq \frac{\pi}{2} \\ \pi - t, & \frac{\pi}{2} \leq t \leq \frac{3\pi}{2}. \end{cases}$



Temos que $x(t+2\pi) = x(t)$. De (2.10) e (2.11) notando que $x(t)$ é ímpar, veremos que as integrais usadas para determinar a_0 e a_k irão se anular, logo $a_0 = a_k = 0$ (fizemos uma alteração no intervalo de integração, neste caso usamos $\left[-\frac{\pi}{2}, \frac{3\pi}{2}\right]$).

Para determinar b_k, vamos usar (2.12) e integração por partes:

$$b_k = \frac{1}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} t\,\mathrm{sen}(kt)\,dt + \frac{1}{\pi}\int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} (\pi - t)\,\mathrm{sen}(kt)\,dt$$

$$= -\frac{1}{\pi k}\Big[t\cos(kt)\Big]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} + \frac{1}{\pi k}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos(kt)\,dt - \frac{1}{\pi k}\Big[(\pi - t)\cos(kt)\Big]_{\frac{\pi}{2}}^{\frac{3\pi}{2}} - \frac{1}{\pi k}\int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \cos(kt)\,dt$$

$$= \frac{4}{\pi k^2}\,\mathrm{sen}\left(\frac{k\pi}{2}\right), \quad k = 1, 2, \ldots$$

Logo

$$x(t) = \frac{4}{\pi}\left(\mathrm{sen}(t) - \frac{1}{3^2}\,\mathrm{sen}(3t) + \frac{1}{5^2}\,\mathrm{sen}(5t) - \cdots + \cdots\right).$$

No exemplo sobre funções contínuas, estudado anteriormente, relembre as sequências (u_n) e (v_n) onde $u_k(t) = \cos(kt)$; $v_k(t) = \mathrm{sen}(kt)$. Deste modo, é possível mostrar a igualdade:

$$x(t) = a_0 u_0(t) + \sum_{k=1}^{\infty}\left[a_k u_k(t) + b_k v_k(t)\right].$$

Agora iremos utilizar a ortogonalidade de u_k e v_k, além disso vamos assumir que integração termo a termo é possível (segue da convergência uniforme). Com essas informações podemos calcular o produto interno de u_j qualquer com x:

$$\begin{aligned}\langle x, u_j\rangle &= a_0\langle u_0, u_j\rangle + \sum \left[a_k\langle u_k, u_j\rangle + b_k\langle v_k, u_j\rangle\right] = a_j\langle u_j, u_j\rangle \\ &= a_j\left\|u_j\right\|^2 = \begin{cases} 2\pi a_0, & j = 0 \\ \pi a_j, & j = 1, 2, \ldots \end{cases}\end{aligned}$$

De modo análogo, aplicando o produto interno por um v_j segue:

$$\langle x, v_j\rangle = b_j\left\|v_j\right\|^2 = \pi b_j, \quad j = 1, 2, \ldots$$

Portanto $a_j = \dfrac{1}{\left\|u_j\right\|^2}\langle x, u_j\rangle = \dfrac{1}{\left\|u_j\right\|}\langle x, e_j\rangle, \quad b_j = \dfrac{1}{\left\|v_j\right\|^2}\langle x, v_j\rangle = \dfrac{1}{\left\|v_j\right\|}\langle x, \tilde{e}_j\rangle.$

O que nos dá, $a_k u_k(t) = \dfrac{1}{\|u_k\|}\langle x, e_k\rangle u_k(t) = \langle x, e_k\rangle e_k(t)$ e uma expressão análoga para $b_k v_k(t)$.

Por fim, segue que

$$x = \langle x, e_0\rangle e_0 + \sum_{k=1}^{\infty}\left[\langle x, e_k\rangle e_k + \langle x, \tilde{e_k}\rangle \tilde{e}_k\right].$$

Observando a última expressão podemos notar que aparecem os coeficientes de Fourier, definidos anteriormente, o que nos dá uma justificativa para tal nome.

Como é possível estender essa ideia para outras sequências ortonormais? E o que é possível dizer sobre a convergências dessas séries?

Veremos a seguir um teorema que garante uma representação parecida com o que viemos estudando utilizando coeficientes de Fourier.

Lembremos que dada uma sequência ortonormal (e_k) em um espaço de Hilbert H, podemos considerar a série $\sum_{k=1}^{\infty} \alpha_k e_k$, onde $\alpha_1, \ldots, \alpha_k, \ldots$ são escalares. Essa série é *convergente* se existe $S \in H$, tal que $\|S_n - S\| \xrightarrow{n\to\infty} 0$, onde $S_n = \alpha_1 e_1 + \cdots + \alpha_n e_n$.

Teorema 2.3.3. *(convergência) Seja (e_k) uma sequência ortonormal em um espaço de Hilbert H, então:*

1. *A série $\sum_{k=1}^{\infty} \alpha_k e_k$ converge (na norma de H) se, e somente se, $\sum_{k=1}^{\infty} |\alpha_k|^2$ converge.*

2. *Se $\sum_{k=1}^{\infty} \alpha_k e_k$ converge, então α_k são os coeficientes de Fourier $\langle x, e_k\rangle$, onde $x = \sum_{k=1}^{\infty} \alpha_k e_k$.*

3. *Para todo $x \in H$, a série $\sum_{k=1}^{\infty} \alpha_k e_k$, $\alpha_k = \langle x, e_k\rangle$, converge na norma de H.*

Demonstração. 1. Seja $S_n = \alpha_1 e_1 + \cdots + \alpha_n e_n$ e $\sigma_n = |\alpha_1|^2 + \cdots + |\alpha_n|^2$.

Usando a ortonormalidade de (e_k), para todo $m < n$:

$$\begin{aligned}\|S_n - S_m\|^2 &= \|\alpha_{m+1}e_{m+1} + \cdots + \alpha_n e_n\|^2 \\ &= |\alpha_{m+1}|^2 + \cdots + |\alpha_n|^2 = \sigma_n - \sigma_m.\end{aligned}$$

A igualdade acima garante que (S_n) é de Cauchy em H se, e somente se, se (σ_n) é de Cauchy em $\mathbb{R}$, como ambos H e $\mathbb{R}$ são completos, é o mesmo que (S_n) converge se, e somente se, (σ_n) converge.

2. Considere S_n como no item anterior, deste modo

$$\langle S_n, e_j\rangle = \langle \alpha_1 e_1 + \cdots \alpha_n e_n, e_j\rangle = \alpha_j, \quad j = 1, \ldots, k.$$

Por hipótese, sabemos que $S_n \xrightarrow{n\to\infty} x$. Segue da continuidade do produto interno que $\alpha_j = \langle S_n, e_j\rangle \to \langle x, e_j\rangle$, ou seja, $\alpha_j = \langle x, e_j\rangle$.

3. Pela desigualdade de Bessel, vemos que $\sum_{k=1}^{\infty} |\langle x, e_k\rangle|^2$ converge, por (1) segue o resultado.

□

Observação 2.3.4. Se uma família ortonormal (e_λ), $\lambda \in I$, em um espaço com produto interno X é não contável (ou seja, I é não contável), ainda podemos formar os coeficientes de Fourier $\langle x, e_\lambda\rangle$, $x \in X$, $\lambda \in I$. Relembremos a desigualdade

$$\sum_{k=1}^{n} |\langle x, e_k\rangle|^2 \leq \|x\|^2.$$

Da desigualdade segue que, para todo $m = 1, 2, \ldots$ fixo, o número de coeficientes de Fourier tal que $|\langle x, e_\lambda\rangle| > \frac{1}{m}$ deve ser finito (podemos pegar um coleção enumerável de coeficientes de Fourier e utilizar a desigualdade de Bessel para chegar a uma contradição), logo podemos escolher os coeficientes de Fourier.

Portanto, em um espaço com produto interno X, dado $x \in X$, segue que *x possui no máximo uma quantidade contável de coeficientes de Fourier* $\langle x, e_\lambda\rangle \neq 0$, com respeito à uma família ortonormal $(e_\lambda)_{\lambda\in I}$, em X.

Logo, para todo $x \in H$ fixo, associamos

$$x = \sum_{\lambda \in I} \langle x, e_\lambda \rangle e_\lambda$$

e arranjamos os e_λ em uma sequência $(e_1, e_2, \ldots)$, a convergência segue do teorema anterior e para garantir que a forma arranjada não faz diferença provamos o lema a seguir.

Lema 2.3.5. *Nos termos da observação (2.3.4), a ordem que os e_λ são arranjados não afeta o resultado da soma.*

Demonstração. Seja (w_m) um rearranjo de (e_n), ou seja, por definição existe uma bijeção

$$\begin{aligned} \mathbb{N} &\longrightarrow \mathbb{N} \\ n &\mapsto m(n) \end{aligned}$$

onde $w_{m(n)} = e_n$.

Defina $\alpha_n = \langle x, e_n \rangle$, $\beta_m = \langle x, w_m \rangle$, $x_1 = \sum_{n=1}^{\infty} \alpha_n e_n$ e $x_2 = \sum_{m=1}^{\infty} \beta_m w_m$.

Pelo teorema anterior (2),

$$\begin{aligned} \alpha_n &= \langle x, e_n \rangle = \langle x_1, e_n \rangle \\ \beta_m &= \langle x, w_m \rangle = \langle x_2, w_m \rangle \end{aligned}$$

Como $e_n = w_{m(n)}$, obtemos

$$\langle x_1 - x_2, e_n \rangle = \langle x_1, e_n \rangle - \langle x_2, w_{m(n)} \rangle = \langle x, e_n \rangle - \langle x, w_{m(n)} \rangle = 0$$

Analogamente, $\langle x_1 - x_2, w_m \rangle = 0$, portanto

$$\begin{aligned} \|x_1 - x_2\|^2 &= \langle x_1 - x_2, \sum \alpha_n e_n - \sum \beta_m w_m \rangle \\ &= \sum \overline{\alpha_n} \langle x_1 - x_2, e_n \rangle - \sum \overline{\beta_m} \langle x_1 - x_2, w_m \rangle = 0. \end{aligned}$$

Logo $x_1 - x_2 = 0$, ou seja, $x_1 = x_2$. □

2.4 Conjuntos e sequências totalmente ortonormais

Em espaços de Hilbert, são interessantes conjuntos cujos elementos podem ser usados para representar qualquer elemento do espaço, ou pelo menos para "aproximar" esse elemento em um certo sentido. Essa propriedade é mais delicada em espaços de dimensão infinita, veremos alguns casos.

Definição 2.4.1. *Um conjunto total, ou conjunto fundamental, em um espaço normado X é um subconjunto $M \subset X$ tal que $\overline{span\, M} = X$.*

Se M for ortonormal (conjunto, sequência ou família) em um espaço com produto interno X e total em X, M é dito conjunto totalmente ortonormal.

Uma família totalmente ortonormal em X é dita, por vezes, base ortonormal de X o que não é equivale à base de Schauder, a menos que $\dim X < \infty$.

Afirmação 2.4.2. *Todo espaço de Hilbert $H \neq \{0\}$ possui subconjunto totalmente ortonormal.*

De fato, se $\dim H < \infty$, basta usar o processo de Gram-Schimidt em uma base qualquer. De modo análogo, se $\dim H = \infty$ e H for separável, por definição, H possui um subconjunto denso e enumerável, portanto basta usar o processo de Gram-Schimidt nos elementos deste subconjunto. Agora se $\dim H = \infty$ e H é não separável, a demonstração irá seguir do *Lema de Zorn*, o qual não tratamos neste trabalho (pode ser encontrado em [4]).

Observação 2.4.3. Todos os conjuntos totalmente ortonormais em $H \neq \{0\}$ têm mesma *cardinalidade*, iremos chamá-la de *dimensão de Hilbert* ou *dimensão ortogornal* e se $H = \{0\}$, essa dimensão vale 0 por definição.

O Teorema a seguir prova o que foi discutido na observação.

Teorema 2.4.4. *Seja X espaço com produto interno, $M \subset X$ subconjunto. Então:*

1. *Se M é total em X, não existe $0 \neq x \in X$ que é ortogonal a todo elemento de M, ou seja, se $x \perp M$, então $x = 0$.*

2. *Se X é completo, então vale a recíproca do item anterior.*

Demonstração. 1. Segue do Teorema 3.2-3 em [4] que existe um espaço de Hilbert H, tal que X é isomorfo à algum subespaço de H, além disso X é denso em H (visto como subespaço), como M é denso em X segue que é denso em H. Pelo Lema 2.1.24 segue que $M^{\perp}$ (complementar em H) é igual a $\{0\}$, ou seja, se $x \in X$ e $x \perp M$ então $x = 0$.

2. Se X é um espaço de Hilbert e M satisfaz a condição $M^{\perp} = \{0\}$, segue pelo Lema 2.1.24 que M é total em X.

$\square$

Vamos discutir um outro critério para totalidade. Vamos considerar um conjunto ortonormal M em um espaço de Hilbert H. Pela Observação 2.3.4 temos que todo $x \in H$ tem no máximo uma quantia contável de coeficientes de Fourier não nulos. Deste modo, vamos arranjá-los em uma sequência, digamos $\langle x, e_1 \rangle, \langle x, e_2 \rangle, \ldots, \langle x, e_n \rangle, \ldots$

Relembrando a desigualdade de Bessel:

$$\sum_k |\langle x, e_k \rangle|^2 \leq \|x\|^2 .$$

Onde a soma é finita ou uma série convergente. Se mudarmos o sinal para um sinal de igualdade, denotaremos a igualdade por *identidade de de Parseval.*

$$\sum_k |\langle x, e_k \rangle|^2 = \|x\|^2 . \tag{2.13}$$

E obtemos um novo critério, descrito a seguir.

Teorema 2.4.5. *Um conjunto ortonormal M em um espaço de Hilbert H é total se, e somente, se para todo $x \in H$ a identidade de Parseval é verdadeira.*

Demonstração. Por um lado, temos que a identidade de Parseval é verdadeira. Vamos supor que M não é total, ou seja, existe um $x \neq 0$ em H de modo que $x \perp M$. Deste modo, na igualdade (2.13) os termos $\langle x, e_k \rangle$ serão iguais a 0, o que torna a igualdade nula, ou seja, $\|x\|^2 = 0$, o que não pode ocorrer, logo M tem que ser total em H.

Por outro lado, suponhamos que M é total em H. Considere $x \in H$ qualquer e seus coeficientes de Fourier arranjados em uma sequência $\big(\langle x, e_n \rangle\big)_{n \in \mathbb{N}}$ ou em uma ordem qualquer caso sejam finitos. Definimos y por

$$y = \sum_k \langle x, e_k \rangle e_k . \tag{2.14}$$

Caso seja uma série, a convergência segue do Teorema 2.3.3. Vamos mostrar que $x - y \perp M$, fazendo isso, como M é total temos que $M^{\perp} = \{0\}$, o que irá quase concluir a prova.

Para todo e_j na igualdade (2.14), usando a ortogonalidade, temos:

$$\langle x-y, e_j\rangle = \langle x, e_j\rangle - \sum_k \langle x, e_k\rangle\langle e_k, e_j\rangle = \langle x, e_j\rangle - \langle x, e_j\rangle = 0.$$

Agora, para todo $v \in M$ que não está em (2.14), temos que $\langle x, v\rangle = 0$, então

$$\langle x-y, v\rangle = \langle x, v\rangle - \sum_k \langle x, e_k\rangle\langle e_k, v\rangle = 0 - 0 = 0.$$

Logo, $x = y$ como queríamos mostrar. Para concluir a prova, utilizando ortonormalidade segue que,

$$\|x\|^2 = \left\langle \sum_k \langle x, e_k\rangle e_k, \sum_m \langle x, e_m\rangle e_m \right\rangle = \sum_k \langle x, e_k\rangle \overline{\langle x, e_k\rangle} = \sum_k |\langle x, e_k\rangle|^2.$$

□

Vejamos agora algumas propriedades de separabilidade em espaços de Hilbert.

Teorema 2.4.6. *Seja H um espaço de Hilbert. Logo:*

1. *Se H é separável, então todo conjunto ortonormal em H é contável.*

2. *Se H possui uma sequência ortonormal que é total em H, então H é separável.*

Demonstração. 1. Seja H separável, $B \subset H$ qualquer e denso em H e M um conjunto ortonormal. Então para quaisquer $x, y \in M$ temos:

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

Ou seja, a distância entre $x, y \in M$ é sempre $\sqrt{2}$. Portanto dadas vizinhanças esféricas N_x e N_y de x e y respectivamente de raios $\frac{\sqrt{2}}{3}$, temos que $N_x \cap N_y = \emptyset$, assim existe $b \in B$ que está em N_x e $\tilde{b} \in B$ que está em N_y e $b \neq \tilde{b}$. Logo, se M fosse incontável, teríamos incontáveis pares de vizinhanças esféricas disjuntas, desse modo B seria incontável. Como B é qualquer, teríamos que H não possui nenhum subconjunto denso e contável, contradição com a separabilidade. Logo, M deve ser contável.

2. Seja (e_k) uma sequência ortonormal e total em H e A o conjunto das combinações lineares da forma:

$$\gamma_1^{(n)} e_1 + \cdots + \gamma_n^{(n)} e_n, \qquad n = 1, 2, \ldots$$

Onde $\gamma_k^{(n)} = a_k^{(n)} + ib_k^{(n)}$ e $a_k^{(n)}$ e $b_k^{(n)}$ são racionais ($b_k^{(n)} = 0$ se H é real). Temos que A é contável. Vamos mostrar que A é denso em H mostrando que para todo $x \in H$ e $\epsilon > 0$ existe um $v \in A$ tal que $\|x - v\| < \epsilon$.

Como (e_k) é total em H, existe n tal que $Y_n = \text{span}\,\{e_1, e_2, \ldots, e_n\}$ contém um ponto cuja distância à x é menor que $\frac{\epsilon}{2}$. Em particular, $\|x - y\| < \frac{\epsilon}{2}$ para

$$y = \sum_{k=1}^{n} \langle x, e_k \rangle e_k,$$

ou seja, para a projeção ortogonal de x em Y_n. Portanto temos

$$\left\| x - \sum_{k=1}^{n} \langle x, e_k \rangle e_k \right\| < \frac{\epsilon}{2}.$$

Como $\mathbb{Q}$ é denso em $\mathbb{R}$, para cada $\langle x, e_k \rangle$ existe um γ_k^n (com partes reais e imaginarias sendo racionais) tal que

$$\begin{aligned}\left\| \sum_{k=1}^{n} \left[\langle x, e_k \rangle - \gamma_k^{(n)} \right] e_k \right\| &\leq \sum_{k=1}^{n} \left\| (\langle x, e_k \rangle - \gamma_k^{(n)}) e_k \right\| \\ &\leq \sum_{k=1}^{n} |\langle x, e_k \rangle - \gamma_k^{(n)}| \leq \frac{\epsilon}{2}.\end{aligned}$$

Assim, defina $v \in A$ por:

$$v = \sum_{k=1}^{n} \gamma_k^{(n)} e_k.$$

Dessa forma, v satisfaz:

$$\begin{aligned}\|x - v\| &= \left\| x - \sum \gamma_k^{(n)} e_k \right\| \\ &\leq \left\| x - \sum \langle x, e_k \rangle e_k \right\| + \left\| \sum \langle x, e_k \rangle e_k - \sum \gamma_k^{(n)} e_k \right\| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.\end{aligned}$$

O que prova que A é denso em H e, como A é contável, temos que H é separável.

□

Definição 2.4.7. *Um isomorfismo de um espaço de Hilbert H em um espaço de Hilbert $\widetilde{H}$,* *ambos sobre o mesmo corpo, é um operador linear bijetivo $T : H \longrightarrow \widetilde{H}$ tal que*

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

para todo $x, y \in H$. Os espaços H e $\widetilde{H}$ são ditos isomorfos.

Teorema 2.4.8. *Dois espaços de Hilbert H e $\widetilde{H}$ (ambos ou reais ou complexos) são isomorfos se, e somente se eles têm a mesma dimensão de Hilbert.*

Demonstração. Se H é isomorfo à $\widetilde{H}$ sendo $T : H \longrightarrow \widetilde{H}$ um isomorfismo, então se $x \perp y$ segue que $Tx \perp Ty$ diretamente da definição de isomorfismo. Como T é bijetivo, concluímos que T mapeia todos os conjuntos totalmente ortonormais de H para conjuntos totalmente ortonormais de $\widetilde{H}$. Logo H e $\widetilde{H}$ tem a mesma dimensão de Hilbert.

Por outro lado, suponha que H e $\widetilde{H}$ tem a mesma dimensão de Hilbert. O caso $H = \{0\} = \widetilde{H}$ é trivial. Suponha então que $H \neq \{0\}$, portanto $\widetilde{H} \neq \{0\}$, além disso quaisquer conjuntos totalmente ortonormais M em H e $\tilde{M}$ em $\widetilde{H}$ tem a mesma cardinalidade, então podemos indexa-los usando o mesmo índice. Vamos escrever $M = (e_k)$ e $\widetilde{M} = (\tilde{e}_k)$, onde $\{k\}$ é o conjunto de índices.

Vamos construir um isomorfismo entre H e $\widetilde{H}$. Para cada $x \in H$, temos

$$x = \sum_k \langle x, e_k\rangle e_k.$$

A parte direita é finita ou uma série, além disso $\sum_k |\langle x, e_k\rangle|^2 < \infty$ pela desigualdade de Bessel. Definimos

$$\tilde{x} = Tx = \sum_k \langle x, e_k\rangle \tilde{e_k}.$$

A convergência segue do Teorema 2.3.3, assim $\tilde{x} \in \widetilde{H}$. O operador T é linear (segue da linearidade do primeiro fator do produto interno), além disso T é isométrico. De fato,

$$\|\tilde{x}\|^2 = \|Tx\|^2 = \sum_k |\langle x, e_k\rangle|^2 = \|x\|^2 .$$

Disso, segue que T preserva produto interno. Além disso, T é injetivo, de fato, se $Tx = Ty$ então

$$\|x - y\| = \big\|T(x - y)\big\| = \|Tx - Ty\| = 0.$$

assim $x = y$, ou seja, T é injetivo.

Por fim, mostraremos que T é sobrejetivo. Dado qualquer $\tilde{x} = \sum_k \alpha_k \tilde{e}_k$ em $\widetilde{H}$, nós temos $\sum |\alpha_k|^2 < \infty$ pela desigualdade de Bessel. Logo

$$\sum_k \alpha_k e_k.$$

é uma soma finita ou uma série convergente para um $x \in H$ por 2.3.3, além disso $\alpha_k = \langle x, e_k \rangle$ pelo mesmo teorema. Então, segue que $\tilde{x} = Tx$. Como $\tilde{x} \in H$ era arbitrário, isso mostra que T é sobrejetivo. □

2.5 Representação de funcionais lineares em espaços de Hilbert

Dados um espaço de Hilbert e um ponto $z \in H$, podemos definir o funcional definido por $f(x) = \langle x, z \rangle$. Temos que f é obviamente linear pela definição de produto interno, além disso, pela desigualdade de Schwarz, f é limitado. Veremos a seguir que f nos dá uma caracterização geral, ou seja, qualquer funcional linear limitado pode ser representado dessa forma em termos do produto interno em H.

Teorema 2.5.1 (1º Teorema de representação de Riesz). *Qualquer funcional linear limitado f em um espaço de Hilbert H pode ser representado com relação ao produto interno em H, do seguinte modo:*

$$f(x) = \langle x, z \rangle, \tag{2.15}$$

onde z é unicamente determinado por f e

$$\|z\| = \|f\|. \tag{2.16}$$

Demonstração. Seguimos o seguinte roteiro:

1. Provamos que f pode ser representado por (2.15),

2. z é único,

3. A equação (2.16) é verdadeira.

1. Se $f = 0$ então (1), (2) e (3) são imediatas tomando $z = 0$. Suponha que $f \neq 0$. Antes de prosseguir com a prova, vamos discutir algumas propriedades básicas que z precisa

satisfazer. Precisamos que $z \neq 0$ (caso contrário $f = 0$, o que não queremos), além disso $\langle x, z\rangle = 0$ para todo x tal que $f(x) = 0$, ou seja, para todo $x \in \ker(f)$ devemos ter $z \perp \ker(f)$. Assim, tomemos os espaços $\ker(f)$ e $\ker(f)^{\perp}$.

Temos que $\ker(f)$ é um espaço vetorial e fechado (segue de 1.3.16 e 1.3.3). Como temos $f \neq 0$, $\ker(f) \neq \{0\}$, portanto $\ker(f)^{\perp} \neq \{0\}$. Logo existe um $z_0 \in \ker(f)^{\perp}$ tal que $z_0 \neq 0$. Fixando $x \in H$ qualquer, definimos

$$v = f(x)z_0 - f(z_0)x,$$

aqui v depende de x. Aplicando f:

$$f(v) = f(x)f(z_0) - f(z_0)f(x) = 0.$$

Segue que $v \in \ker(f)$. Como $z_0 \perp \ker(f)$, temos que

$$\begin{aligned} 0 = \langle v, z_0\rangle &= \langle f(x)z_0 - f(z_0)x, z_0\rangle \\ &= f(x)\langle z_0, z_0\rangle - f(z_0)\langle x, z_0\rangle. \end{aligned}$$

Sabendo que $\langle z_0, z_0\rangle = \|z_0\|^2 \neq 0$, segue que

$$f(x) = \frac{f(z_0)}{\langle z_0, z_0\rangle}\langle x, z_0\rangle.$$

Ou equivalentemente, definimos

$$z = \frac{\overline{f(z_0)}}{\langle z_0, z_0\rangle} z_0.$$

Como $x \in H$ era qualquer, segue o item (1).

2. Suponha que para todo $x \in H$,

$$f(x) = \langle x, z_1\rangle = \langle x, z_2\rangle$$

Deste modo, $\langle x, z_1 - z_2\rangle = 0$ para todo x. Em particular se $x = z_1 - z_2$, temos que $\langle z_1 - z_2, z_1 - z_2\rangle = 0$, ou seja, $z_1 - z_2 = 0$ então $z_1 = z_2$.

3. Temos que $f \neq 0$ (o caso $f = 0$ já foi discutido), deste modo $z \neq 0$ e escolhendo $x = z$

temos

$$\|z\|^2 = \langle z, z\rangle = f(z) \leq \|f\|\|z\|$$

Dividindo por $\|z\| \neq 0$ segue que $\|z\| \leq \|f\|$.

Resta mostrar que $\|z\| \geq \|f\|$. Pela desigualdade de Schwarz segue que

$$|f(x)| = |\langle x, z\rangle| \leq \|x\|\|z\|\,.$$

Logo,

$$\|f\| = \sup_{\|x\|=1} |\langle x, z\rangle| \leq \|z\|\,.$$

□

Lema 2.5.2. *Se $\langle v_1, w\rangle = \langle v_2, w\rangle$ para todo $w \in X$, onde X é espaço com produto interno, então $v_1 = v_2$. Em particular, $\langle v_1, w\rangle = 0$ para todo $w \in X$ implica em $v_1 = 0$.*

Demonstração. Por hipótese, para todo w:

$$\langle v_1 - v_2, w\rangle = \langle v_1, w\rangle - \langle v_2, w\rangle = 0.$$

Para $w = v_1 - v_2$, temos $\|v_1 - v_2\|^2 = 0$. Portanto $v_1 - v_2 = 0$ implica $v_1 = v_2$. Em particular se $\langle v_1, w\rangle = 0$ para todo w, usando $w = v_1$ segue que $\|v_1\|^2 = 0$, então $v_1 = 0$. □

Definição 2.5.3. *Sejam X e Y espaços vetoriais sobre o mesmo corpo $\mathbb{K}$ (reais ou complexos). Uma forma sesquilinear (ou funcional sesquilinear) h em $X \times Y$ é uma transformação*

$$h : X \times Y \longrightarrow K$$

tal que para todo $x, x_1, x_2 \in X$ e $y, y_1, y_2 \in Y$ e escalares α, β:

- $h(x_1 + x_2, y) = h(x_1, y) + h(x_2, y)$;
- $h(x, y_1 + y_2) = h(x, y_1) + h(x, y_2)$;
- $h(\alpha x, y) = \alpha h(x, y)$;
- $h(x, \beta y) = \overline{\beta} h(x, y)$.

Note que h é linear na primeira entrada e é dita *linear conjugada* na segunda. Se $\mathbb{K} = \mathbb{R}$ temos que h é linear nas duas entradas, assim h é dita *bilinear*.

Se X e Y são espaços normados e se existe um número real c tal que para todo x,y

$$|h(x,y)| \leq c\|x\|\|y\|\,,$$

dizemos que h é *limitado*, e o número

$$\|h\| = \sup_{x\in X/\{0\},\, y\in Y/\{0\}} \frac{|h(x,y)|}{\|x\|\|y\|} = \sup_{\|x\|=1,\, \|y\|=1} |h(x,y)|$$

é chamado de *norma* de h.

Note que o produto interno é um exemplo de forma sesquilinear. Veremos que, em certo sentido, produto interno e forma sesquilinear são a "mesma coisa".

Um resultado análogo ao visto anteriormente para representação de funcionais lineares segue para formas sesquilineares, iremos discuti-lo a seguir.

Teorema 2.5.4 (2º Teorema de representação de Riesz)**.** *Sejam H_1 e H_2 espaços de Hilbert sobre o corpo $\mathbb{K}$ (reais ou complexos) e*

$$h : H_1 \times H_2 \longrightarrow \mathbb{K}$$

uma forma sesquilinear limitada. Então h pode ser representada da seguinte maneira:

$$h(x,y) = \langle Sx, y\rangle, \tag{2.17}$$

onde $S : H_1 \longrightarrow H_2$ é um operador linear limitado. S é unicamente determinado por h e

$$\|S\| = \|h\|\,. \tag{2.18}$$

Demonstração. Consideramos $\overline{h(x,y)}$. Ao fixar x, h é linear em y devido ao conjugado. Deste modo usaremos o Teorema 2.5.1 em h, ou seja, com x fixo segue que:

$$\overline{h(x,y)} = \langle y, z\rangle.$$

Portanto

$$h(x, y) = \langle z, y \rangle.$$

Também segue pelo teorema que $z \in H_2$ é único (mas depende do $x \in H_1$ fixo). Usando essa última expressão com x variável, temos um operador linear:

$$\begin{aligned} S : H_1 &\longrightarrow H_2 \\ x &\mapsto Sx = z. \end{aligned}$$

Deste modo $h(x, y) = \langle z, y \rangle = \langle Sx, y \rangle$ como queríamos. Resta mostrar que S é operador linear limitado e $\|S\| = \|h\|$. S é linear, de fato:

$$\begin{aligned} \langle S(\alpha x_1 + \beta x_2), y \rangle &= h(\alpha x_1 + \beta x_2, y) \\ &= \alpha h(x_1, y) + \beta h(x_2, y) \\ &= \alpha \langle Sx_1, y \rangle + \beta \langle Sx_2, y \rangle \\ &= \langle \alpha Sx_1 + \beta Sx_2, y \rangle. \end{aligned}$$

Isso vale para todo $y \in H_2$, logo pelo Lema 2.5.2

$$S(\alpha x_1 + \beta x_2) = \alpha Sx_1 + \beta Sx_2.$$

Vejamos que S é limitado. De fato (ignorando o caso trivial $S = 0$) temos que:

$$\|h\| = \sup_{x \neq 0,\ y \neq 0} \frac{|\langle Sx, y \rangle|}{\|x\| \|y\|} \leq \sup_{x \neq 0} \frac{\|Sx\| \|y\|}{\|x\| \|y\|} = \|S\|.$$

Também temos $\|h\| \geq \|S\|$ pela desigualdade de Schwarz.

Agora vejamos que S é único. De fato, assuma que existe $T : H_1 \longrightarrow H_2$ operador linear tal que, para todo $x \in H_1$ e $y \in H_2$ temos

$$h(x, y) = \langle Sx, y \rangle = \langle Tx, y \rangle,$$

segue pelo Lema 2.5.2 que $Sx = Tx$, para todo $x \in H_1$. Logo $S = T$. $\square$

2.6 Operador Hilbert-adjunto

Com os resultados discutidos na seção anterior iremos estudar algumas classes especiais de operadores lineares, começaremos com operadores Hilbert-adjuntos.

Definição 2.6.1. *Seja $T : H_1 \longrightarrow H_2$ um operador linear limitado, onde H_1 e H_2 são espaços de Hilbert. O operador Hilbert-adjunto T^* de T é o operador definido por*

$$T^* : H_2 \longrightarrow H_1,$$

tal que para todo $x \in H_1$ e $y \in H_2$, vale

$$\langle Tx, y \rangle = \langle x, T^* y \rangle.$$

Teorema 2.6.2. *Sejam H_1, H_2 espaços de Hilbert e $T : H_1 \longrightarrow H_2$ um operador linear limitado. O operador Hilbert-adjunto T^*, como na definição anterior, existe, é único e limitado com norma*

$$\|T^*\| = \|T\|.$$

Demonstração. Considere a fórmula

$$h(y, x) = \langle y, Tx \rangle. \tag{2.19}$$

Essa fórmula define uma forma sesquilinear em $H_2 \times H_1$ pois o produto interno é sesquilinear e T é linear.

Temos que h é limitada. De fato, pela desigualdade de Schwarz,

$$|h(y, x)| = |\langle y, Tx \rangle| \leq \|y\| \|Tx\| \leq \|T\| \|x\| \|y\|,$$

ou seja, dividindo ambos os lados por $\|x\| \|y\|$ e tomando sup, segue que $\|h\| \leq \|T\|$. Além disso, temos $\|h\| \geq \|T\|$, de fato:

$$\|h\| = \sup_{x \neq 0,\ y \neq 0} \frac{|\langle y, Tx \rangle|}{\|y\| \|x\|} \geq \sup_{x \neq 0,\ Tx \neq 0} \frac{|\langle Tx, Tx \rangle|}{\|Tx\| \|x\|} = \|T\|.$$

Logo $\|h\| = \|T\|$. Pelo Teorema 2.5.4 temos uma representação de Riesz para h; denotando o

S do teorema por T^* temos

$$h(y, x) = \langle T^* y, x \rangle, \tag{2.20}$$

e sabemos que $T^* : H_2 \longrightarrow H_1$ é operador linear limitado unicamente determinado e tem norma

$$\|T^*\| = \|h\| = \|T\|.$$

Portanto a fórmula está provada. Também temos, por (2.19) e (2.20), $\langle y, Tx \rangle = \langle T^* y, x \rangle$ e tomando o conjugado dos dois lados segue a propriedade descrita na definição anterior. □

O próximo lema será útil no estudo dos operadores adjuntos.

Lema 2.6.3. *Sejam X e Y espaços com produto interno e $Q : X \longrightarrow Y$ um operador linear limitado. Então:*

- *$Q = 0$ se, e somente se, $\langle Qx, y \rangle = 0$, para todo $x \in X$ e $y \in Y$.*
- *Seja $Q : X \longrightarrow X$, onde X é um espaço complexo. Se $\langle Qx, x \rangle = 0$ para todo $x \in X$, então $Q = 0$.*

Demonstração. • Para o primeiro item, $Q = 0$ então $Qx = 0$ para todo x, assim

$$\langle Qx, y \rangle = \langle 0, y \rangle = 0.$$

Analogamente se $\langle Qx, y \rangle = 0$ para todo x e y então $Qx = 0$ para todo x, pelo Lema 2.5.2, logo $Q = 0$.

• Para o segundo item, temos, por hipótese, que $\langle Qv, v \rangle = 0$ para todo $v = \alpha x + y \in X$, ou seja

$$\begin{aligned} 0 &= \langle Q(\alpha x + y), \alpha x + y \rangle \\ &= |\alpha|^2 \langle Qx, x \rangle + \langle Qy, y \rangle + \alpha \langle Qx, y \rangle + \overline{\alpha} \langle Qy, x \rangle \\ &= |\alpha|^2 \cdot 0 + 0 + \alpha \langle Qx, y \rangle + \overline{\alpha} \langle Qy, x \rangle \\ &= \alpha \langle Qx, y \rangle + \overline{\alpha} \langle Qy, x \rangle. \end{aligned}$$

Se $\alpha = 1$ temos

$$\langle Qx, y \rangle + \langle Qy, x \rangle = 0.$$

Tomando $\alpha = i$ segue que $\overline{\alpha} = -i$ e

$$\langle Qx, y\rangle - \langle Qy, x\rangle = 0.$$

Somando as duas últimas expressões segue que $\langle Qx, y\rangle = 0$. Finalmente, pelo item anterior, $Q = 0$.

$\square$

Veja que para provar o segundo item do lema anterior é essencial que X seja um espaço complexo. De fato, um contraexemplo para o caso em que X é um espaço real pode ser formulado usando $X = \mathbb{R}^2$ e Q uma rotação em um ângulo reto. Temos que Q é linear e $Qx \perp x$ para todo $x \in \mathbb{R}^2$, portanto $\langle Qx, x\rangle = 0$ para todo $x \in \mathbb{R}^2$, mas $Q \neq 0$.

A seguir são listadas algumas propriedades básicas e importantes de operadores Hilbert-adjuntos.

Teorema 2.6.4. *Seja H_1, H_2 espaços de Hilbert, $S : H_1 \longrightarrow H_2$ e $T : H_1 \longrightarrow H_2$ operadores lineares limitados e α escalar. Segue que:*

1. $\langle T^*y, x\rangle = \langle y, Tx\rangle$ *para todo* $x \in H_1$ *e* $y \in H_2$*;*
2. $(S + T)^* = S^* + T^*$*;*
3. $(\alpha T)^* = \overline{\alpha} T^*$*;*
4. $(T^*)^* = T$*;*
5. $\|T^*T\| = \|TT^*\| = \|T\|^2$*;*
6. $T^*T = 0 \Leftrightarrow T = 0$*;*
7. $(ST)^* = T^*S^*$ *(assumindo* $H_1 = H_2$*) .*

Demonstração. Os itens (1)-(4) seguem diretamente das definições e dos Lemas 2.5.2 e 2.6.3 usando operações algébricas simples. Para o item (5) vejamos que $T^*T : H_1 \longrightarrow H_1$, mas $TT^* : H_2 \longrightarrow H_2$. Pela desigualdade de Schwarz,

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle T^*Tx, x\rangle \leq \left\|T^*Tx\right\| \|x\| \leq \left\|T^*T\right\| \|x\|^2$$

Tomando supremo sobre x com norma 1, obtemos $\|T\|^2 \leq \|T^*T\|$. Logo

$$\|T\|^2 \leq \left\|T^*T\right\| \leq \left\|T^*\right\|\|T\| = \|T\|^2 .$$

Portanto $\|T\|^2 = \|T^*T\|$. Substituindo T por T^*:

$$\left\|T^{**}T^*\right\| = \left\|T^*\right\|^2 = \|T\|^2 .$$

logo, o item (5) está provado.

O item (6) segue diretamente de (5), e para o item (7) basta repetir a definição da adjunta repetidas vezes, ou seja,

$$\langle x, (ST)^*y\rangle = \langle (ST)x, y\rangle = \langle Tx, S^*y\rangle = \langle x, T^*S^*y\rangle.$$

Logo $(ST)^*y = T^*S^*y$ por 2.5.2. □

2.7 Operadores autoadjuntos, unitários e normais

Iremos discutir agora operadores Hilbert-adjuntos que serão úteis no estudo de operadores lineares limitados.

Definição 2.7.1. *Um operador linear limitado $T : H \longrightarrow H$ em um espaço de Hilbert H é dito:*

- *Auto-adjunto ou Hermitiano se $T^* = T$;*
- *Unitário se T é bijetivo e $T^* = T^{-1}$;*
- *Normal se $TT^* = T^*T$.*

Observação 2.7.2. 1. Se T é auto-adjunto, temos que $\langle Tx, y\rangle = \langle x, Ty\rangle$.

2. Se T é auto-adjunto ou unitário, segue que T é normal.

Exemplo 2.7.3. Considere $\mathbb{C}^n$ com o produto interno

$$\langle x, y\rangle = x^t\overline{y},$$

onde x e y são vetores coluna, t é a transposição, ou seja, $x^t = (\xi_1, \ldots, \xi_n)$ e a multiplicação de matrizes é a usual.

Seja $T : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ operador linear (que é limitado pelo Teorema 1.3.14). Dada uma base de $\mathbb{C}^n$ podemos representar T e T^* por duas matrizes quadradas de ordem n, digamos A e B, respectivamente.

Usando a definição de produto interno acima e a regra $(Bx)^t = x^t B^t$, obtemos:

$$\langle Tx, y\rangle = (Ax)^t\overline{y} = x^t A^t \overline{y} \quad \text{e}$$

$$\langle x, T^* y\rangle = x^t \overline{B}\overline{y}.$$

Segue, da definição de operador Hilbert-adjunto, que $x^t A^t \overline{y} = x^t \overline{B}\overline{y}$ para todo $x, y \in \mathbb{C}^n$, portanto $A^t = \overline{B}$, conjugando ambos os lados segue que:

$$B = \overline{A^t}$$

Acabamos de provar o seguinte resultado:

Dada uma base de $\mathbb{C}^n$ e um operador linear em $\mathbb{C}^n$ representado por uma matriz, então o operador Hilbert-adjunto é representado pela conjugada transposta dessa matriz.

Deste modo, as matrizes que representam operadores podem ser classificadas do seguinte modo:

- *Hermitiana* se T é auto-adjunto (Hermitiano);
- *Unitária* se T é unitário;
- *Normal* se T é normal.

De modo similar, para operadores $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, as matrizes que representam T podem ser classificadas:

- *Simétrica real* se T é auto-adjunto;
- *Ortogonal* se T é unitário.

Agora, relembremos alguns conceitos. Uma matriz $A = (\alpha_{jk})$ é dita:

- Simétrica se $A^t = A$ (A é real), portanto $\alpha_{kj} = \alpha_{jk}$
- Anti-simétrica se $A^t = -A$ (A real), portanto $\alpha_{kj} = -\alpha_{jk}$.
- Ortogonal se $A^t = A^{-1}$.

Portanto uma matriz Hermitiana real é uma matriz simétrica. Uma matriz anti-Hermitiana real é uma matriz anti-simétrica. Uma matriz real unitária é uma matriz ortogonal.

Teorema 2.7.4. *Seja $T : H \longrightarrow H$ operador linear limitado em um espaço de Hilbert H. Então:*

1. *Se T é auto-adjunto, então $\langle Tx, x\rangle$ é real para todo $x \in H$.*
2. *Se H é complexo e $\langle Tx, x\rangle$ é real para todo $x \in H$, então o operador T é auto-adjunto.*

Demonstração.

1. Temos, para todo $x \in H$,

$$\overline{\langle Tx, x\rangle} = \langle x, Tx\rangle = \langle Tx, x\rangle.$$

Logo, $\langle Tx, x\rangle$ é igual seu conjugado, ou seja, é real.

2. Se $\langle Tx, x\rangle$ é real para todo $x \in H$, então

$$\langle Tx, x\rangle = \overline{\langle Tx, x\rangle} = \overline{\langle x, T^*x\rangle} = \langle T^*x, x\rangle$$

Assim, $\langle Tx - T^*x, x\rangle = \langle (T - T^*)x, x\rangle = 0$ para todo x, segue pelo Lema 2.6.3 (b) que $T - T^* = 0$, logo $T = T^*$.

□

Teorema 2.7.5. *Dados S, T operadores lineares auto-adjuntos limitados em um espaço de Hilbert H, temos que o produto de S e T é auto-adjunto se, e somente se, os operadores comutam, ou seja,*

$$ST = TS.$$

Demonstração. Usando as propriedades já provadas de operadores adjuntos, temos que

$$(ST)^* = T^*S^* = TS.$$

Portanto, $ST = (ST)^*$ se, e somente se, $ST = TS$.

□

Agora mostraremos que o limite de uma sequência de operadores auto-adjuntoss ainda é um operador auto-adjunto.

Teorema 2.7.6. *Seja (T_n) uma sequência de operadores lineares auto-adjuntos limitados, onde $T_n : H \longrightarrow H$, H espaço de Hilbert. Suponha que (T_n) converge, digamos:*

$$T_n \longrightarrow T, \quad \textit{ou seja}, \quad \|T_n - T\| \longrightarrow 0,$$

onde $\|.\|$ é a norma no espaço $B(H,H)$ (espaço dos operadores lineares limitados em H). Então o operador limite T é um operados linear auto-adjunto limitado em H.

Demonstração. Para provar o resultado, vamos mostrar que $\|T - T^*\| = 0$.

Temos que

$$\left\|T_n^* - T^*\right\| = \left\|(T_n - T)^*\right\| = \|T_n - T\|.$$

Pela desigualdade triangular em $\mathrm{B}(H,H)$ segue:

$$\begin{aligned}\left\|T - T^*\right\| &\leq \|T - T_n\| + \left\|T_n - T_n^*\right\| + \left\|T_n^* - T^*\right\| \\ &= \|T - T_n\| + 0 + \|T_n - T\| \\ &= 2\|T_n - T\| \xrightarrow{n\to\infty} 0.\end{aligned}$$

Portanto $\|T - T^*\| = 0$, logo $T = T^*$. □

Teorema 2.7.7. *Sejam os operadores unitários $U : H \longrightarrow H$ e $V : H \longrightarrow H$, H é espaço de Hilbert. Então:*

1. *U é isométrico, ou seja, $\|Ux\| = \|x\|$ para todo $x \in H$;*

2. *$\|U\| = 1$ quando $H \neq \{0\}$;*

3. *$U^{-1} = U^*$ é unitário;*

4. *UV é unitário;*

5. *U é normal;*

6. *Um operador linear limitado T em um espaço de Hilbert complexo H é unitário se, e somente se T é isométrico e sobrejetivo.*

Demonstração. 1. $\|Ux\|^2 = \langle Ux, Ux\rangle = \langle x, U^*Ux\rangle = \langle x, Ix\rangle = \|x\|^2$.

2. $\|U\| = \sup_{x\neq 0} \frac{\|Ux\|}{\|x\|} = \sup_{x\neq 0} \frac{\|x\|}{\|x\|} = 1.$

3. Temos que U é bijetivo, portanto U^{-1} também o é. Logo

$$(U^{-1})^* = U^{**} = U = (U^{-1})^{-1}.$$

4. UV é bijetivo, deste modo:

$$(UV)^* = V^*U^* = V^{-1}U^{-1} = (UV)^{-1}.$$

5. $UU^* = UU^{-1} = I = U^{-1}U = U^*U.$

6. Suponha que T é isométrico e sobrejetivo. A injetividade segue do fato de T ser isométrico, assim T é bijetivo. Note que:

$$\langle T^*Tx, x\rangle = \langle Tx, Tx\rangle = \langle x, x\rangle = \langle Ix, x\rangle.$$

Portanto

$$\langle (T^*T - I)x, x\rangle = 0.$$

Segue pelo Lema 2.6.3 que $T^*T = I$. Logo

$$TT^* = TT^*(TT^{-1}) = T(T^*T)T^{-1} = TIT^{-1} = I.$$

Portanto $T^*T = TT^* = I$, com isso e pela bijetividade segue que T é unitária. O outro lado é evidente. De fato, pelo item 1. T é isométrico e é bijetivo por definição.

□

2.8 Convergência fraca

Vamos finalizar esse capítulo apresentando uma nova noção de convergência num espaço normado. Assim como em cálculo existem diversas noções de convergência (pontual, uniforme, absoluta, condicional, etc) estamos interessados aqui em uma nova versão de convergência que seja mais "flexível" que a noção de *convergência forte* dada pela norma. Iremos chamar de *convergência fraca* essa nova noção, a qual é precisamente descrita na definição a seguir.

Definição 2.8.1. *Uma sequência (x_n) em um espaço normado X é dita fracamente convergente* *se existe $x \in X$ tal que para todo $f \in X'$ vale*

$$\lim_{n \longrightarrow \infty} f(x_n) = f(x).$$

O elemento $x \in X$ é dito limite fraco de (x_n) e dizemos que (x_n) converge fracamente para x.

Em geral, convergência fraca não implica convergência forte. De fato, tome (e_n) uma sequência ortonormal em um espaço de Hilbert H. Todo funcional linear $f \in H'$ tem uma representação de Riesz $f(x) = \langle x, z \rangle$. Portanto $f(e_n) = \langle e_n, z \rangle$. Usando a Desigualdade de Bessel segue que

$$\sum_{n=1}^{\infty} |\langle e_n, z \rangle|^2 \leq \|z\|^2 .$$

Deste modo, a série é convergente e portanto o termo geral tende a zero, o que implica

$$f(e_n) = \langle e_n, z \rangle \longrightarrow 0.$$

Como $f \in H'$ é arbitrário, concluímos que $e_n \longrightarrow 0$ fracamente. Porém, (e_n) não converge fortemente, pois

$$\|e_m - e_n\|^2 = \langle e_m - e_n, e_m - e_n \rangle = \langle e_m, e_m \rangle + \langle e_n, e_n \rangle = 2 \quad (m \neq n).$$

No que segue desta seção vamos utilizar alguns resultados que não estão descritos nesse trabalho, dessa forma apenas indicaremos onde podem ser encontrados.

Lema 2.8.2. *Seja (x_n) uma sequência fracamente convergente em um espaço normado X, digamos $x_n \longrightarrow x$ fracamente. Então:*

1. *O limite fraco x de (x_n) é único;*
2. *Toda subsequência de (x_n) converge fracamente para x;*
3. *A sequência $(\|x_n\|)$ é limitada.*

Demonstração. 1. Suponha que $x_n \longrightarrow x$ fracamente e também $x_n \longrightarrow y$ fracamente. Então, para todo $f \in X'$ temos $f(x_n) \longrightarrow f(x)$ e $f(x_n) \longrightarrow f(y)$. Como $(f(x_n))$ é uma sequência de números, seu limite é único. Portanto $f(x) = f(y)$, ou seja,

$$f(x) - f(y) = f(x - y) = 0,$$

para todo $f \in X'$, logo como consequência do Corolário 4.3-4 em [4], temos $x - y = 0$, o que mostra a unicidade do limite fraco.

2. Segue do fato que $(f(x_n))$ é uma sequência convergente de números, então toda subsequência de $(f(x_n))$ converge e tem o mesmo limite da sequência.

3. Como para cada $f \in X'$ temos que $(f(x_n))$ é uma sequência convergente de números, então é uma sequência limitada, digamos $|f(x_n)| \leq c_f$, para todo n, onde c_f é uma constante que depende de f mas não de n. Usando a transformação canônica $J : X \longrightarrow X''$ dada por $x \mapsto Jx = g_x$, onde $g_x(f) = f(x)$, $f \in X'$, definimos para cada $n \in \mathbb{N}$ o funcional $g_n \in X''$ dado por $g_n(f) = f(x_n)$, para cada $f \in X'$.

   Logo, para cada n temos $|g_n(f)| = |f(x_n)| \leq c_f$, ou seja, a sequência $(|g_n(f)|)$ é limitada para todo $f \in X'$. Como X' é completo, segue pelo Princípio da Limitação Uniforme (veja Teorema 4.7-4 em [4]) que a sequência $(\|g_n\|)$ é limitada. Finalmente, temos $\|g_n\| = \|x_n\|$ pelo Lema 4.6-1 em [4] donde segue o resultado desejado.

□

Para finalizar esta seção apresentamos um resultado que nos dá uma relação entre convergência fraca e forte.

Teorema 2.8.3. *Seja (x_n) uma sequência em um espaço normado X. Então:*

1. *Convergência forte implica convergência fraca (com o mesmo limite);*

2. *Se $\dim X < \infty$, então convergência fraca implica convergência forte.*

Demonstração. 1. Suponha que $x_n \longrightarrow x$, logo por definição isso significa que $\|x_n - x\| \longrightarrow 0$, ou seja, dado $f \in X'$ temos

$$|f(x_n) - f(x)| = |f(x_n - x)| \leq \|f\| \|x_n - x\| \longrightarrow 0.$$

   Portanto, $x_n \longrightarrow x$ fracamente. Já havíamos discutido anteriormente que a recíproca desse resultado, em geral, não é verdadeira.

2. Suponha que $x_n \longrightarrow x$ fracamente e $\dim X = k$. Seja $\{e_1, \ldots, e_k\}$ uma base qualquer de X e digamos que $x_n = \alpha_1^{(n)} e_1 + \cdots + \alpha_k^{(n)} e_k$ e $x = \alpha_1 e_1 + \cdots + \alpha_k e_k$. Por hipótese,

$f(x_n) \longrightarrow f(x)$, para todo $f \in X'$. Em particular, tome $f_1, \ldots, f_k$ definidos por

$$f_j(e_n) = \begin{cases} 1, & j = n \\ 0, & j \neq n. \end{cases}$$

Então, $f_j(x_n) = \alpha_j^{(n)}$ e $f_j(x) = \alpha_j$. Como $f_j(x_n) \longrightarrow f_j(x)$, segue que $\alpha_j^{(n)} \longrightarrow \alpha_j$. Logo

$$\|x_n - x\| = \left\| \sum_{j=1}^{k} (\alpha_j^{(n)} - \alpha_j) e_j \right\| \leq \sum_{j=1}^{k} |\alpha_j^{(n)} - \alpha_j| \|e_j\| \xrightarrow{n \longrightarrow \infty} 0.$$

□

Capítulo 3

Operadores lineares compactos

3.1 Operadores lineares compactos em espaços normados

Operadores compactos formam uma classe importante dentre os operadores lineares limitados, com diversas aplicações em Física e em Matemática, em especial na Análise Funcional. Nesta seção iremos definir operadores compactos em espaços normados e discutir suas principais propriedades, as quais serão necessárias para o desenvolvimento do restante do trabalho.

Definição 3.1.1. *Sejam X e Y espaços normados e $B_X = \{x \in X : \|x\| \leq 1\}$ a bola unitária centrada na origem em X. Dizemos que um operador linear $T : X \longrightarrow Y$ é compacto quando $T(B_X)$ é pré-compacto, ou seja, $\overline{T(B_X)}$ é compacto em Y.*

Ao longo do texto usaremos certas caracterizações que descrevem um operador compacto. O próximo teorema apresenta algumas dessas equivalências.

Teorema 3.1.2. *Sejam X e Y espaços normados e $T : X \longrightarrow Y$ um operador linear. São equivalentes:*

1. *T é compacto;*
2. *$T(M)$ é pré-compacto em Y para todo $M \subset X$ limitado;*
3. *Para toda sequência limitada $(x_n) \subset X$, a sequência $(Tx_n) \subset Y$ possui subsequência convergente.*

Demonstração. (1) $\Rightarrow$ (2): Suponha que $T : X \longrightarrow Y$ é compacto e $M \subset X$ é limitado. Então, existe $c > 0$ tal que $\|x\| \leq c$, para todo $x \in M$. Logo, $\frac{1}{c}M \subset B_X$ e assim $\frac{1}{c}\overline{T(M)} \subset \overline{T(B_X)}$, donde segue que $\frac{1}{c}\overline{T(M)}$ é compacto, consequentemente $\overline{T(M)}$ é compacto.

(2) $\Rightarrow$ (3): Seja $(x_n) \subset X$ uma sequência limitada. Considere então o conjunto limitado $M = \{x_n;\ n \in \mathbb{N}\} \subset X$. Então, por hipótese, $\overline{T(M)}$ é compacto em Y e consequentemente $T(x_n)$ possui uma subsequência convergente.

(3) $\Rightarrow$ (1): Vamos mostrar que $\overline{T(B_X)}$ é compacto. Seja $(y_n) \subset \overline{T(B_X)}$. Para cada $n \in \mathbb{N}$, existe $x_n \in B_X$ tal que $\left\|y_n - T(x_n)\right\| < \frac{1}{n}$. Como (x_n) é uma sequência limitada, então, por hipótese, existe uma subsequência $(T(x_{n_k}))$ convergente. Suponha que $T(x_{n_k}) \longrightarrow y$ em Y. Logo,

$$\begin{aligned} \left\|y_{n_k} - y\right\| &\leq \left\|y_{n_k} - T(x_{n_k})\right\| + \left\|T(x_{n_k}) - y\right\| \\ &< \frac{1}{n_k} + \left\|T(x_{n_k}) - y\right\| \xrightarrow{k \longrightarrow \infty} 0. \end{aligned}$$

Logo (y_{n_k}) é uma subsequência convergente de (y_n), consequentemente $\overline{T(B_X)}$ é compacto.

□

Denotaremos o conjunto de operadores compactos de X em Y por $K(X,Y)$. No caso em que $X = Y$ usaremos a notação $K(X)$.

O próximo lema abrange mais propriedades sobre operadores compactos.

Lema 3.1.3. *Sejam X e Y espaços normados. Então:*

1. *Todo operador linear compacto $T : X \longrightarrow Y$ é limitado, portanto contínuo.*

2. *Se $\dim X = \infty$, então o operador identidade $I : X \longrightarrow X$ (que é um operador contínuo) não é compacto.*

Demonstração. A esfera unitária $S = \{x \in X;\ \|x\| = 1\}$ é limitada. Como T é compacto, por definição temos que $\overline{T(S)}$ é compacto, logo é um conjunto limitado. Segue que

$$\sup_{\|x\|=1} \|Tx\| < \infty.$$

Portanto T é um operador limitado, logo também é contínuo.

Suponha agora que $\dim X = \infty$. Temos que a bola unitária $B_X = \{x \in X;\ \|x\| \leq 1\}$ é limitada. Pelo Teorema 1.2.18 temos que B_X não pode ser compacto, logo $I(B_X) = B_X = \overline{B_X}$ não é um conjunto compacto em Y. Segue que o operador identidade I não é compacto.

□

Outra caracterização importante para operadores compactos é a seguinte:

Teorema 3.1.4. *Sejam X e Y espaços normados reflexivos e $T : X \longrightarrow Y$ um operador linear. Então T é compacto se, e somente se, $\|Tx_n\| \longrightarrow 0$ sempre que $x_n \longrightarrow 0$ fracamente em X.*

Demonstração. Suponha inicialmente que T seja compacto e $x_n \longrightarrow 0$ fracamente em X. Afirmamos que $\|Tx_n\| \longrightarrow 0$. De fato, caso contrário existe $\epsilon > 0$ tal que $\|Tx_{n_k}\| > \epsilon$, para todo $k \in \mathbb{N}$, para alguma subsequência (x_{n_k}). Então pelo Lema 2.8.2 temos que $(\|x_{n_k}\|)$ é limitada. Assim, pelo Teorema 3.1.2 existe uma subsequência $(T(x_{n_{k_j}}))$ convergente, digamos para $y \in Y$. Em particular, $Tx_{n_{k_j}} \longrightarrow y$ fracamente em Y.

Como $x_n \longrightarrow 0$ fracamente, então também a subsequência $x_{n_{k_j}} \longrightarrow 0$ fracamente. Logo, dado $g \in Y'$ temos que $g \circ T \in X'$, pois T é limitado. Assim $(g \circ T)(x_{n_{k_j}}) \longrightarrow 0$ quando $j \longrightarrow \infty$ e segue que $Tx_{n_{k_j}} \longrightarrow 0$ fracamente em Y. Pela unicidade do limite temos que $y = 0$.

Como $Tx_{n_{k_j}} \longrightarrow y = 0$ em Y seque que $\left\|Tx_{n_{k_j}}\right\| \longrightarrow 0$, o que contraria o fato de que $\|Tx_{n_k}\| > \epsilon$, para todo $k \in \mathbb{N}$.

Reciprocamente, suponha que $\|Tx_n\| \longrightarrow 0$ sempre que $x_n \longrightarrow 0$ fracamente e que T não é compacto. Logo, existe $(z_n) \subset B_X = \{x \in X : \|x\| \leq 1\}$ tal que (Tz_n) não possui subsequência convergente. Por outro lado, pelo Teorema de Banach-Alaoglu (veja o Teorema 3.16 em [1]) existe uma subsequência (z_{n_k}) tal que $z_{n_k} \longrightarrow z_0$ fracamente para algum $z_0 \in X$. Logo, a sequência $x_k = z_{n_k} - z_0$ converge fracamente para zero, o que implica por hipótese que

$$\|T(z_{n_k}) - Tz_0\| = \|T(z_{n_k} - z_0)\| = \|Tx_k\| \xrightarrow{k \longrightarrow \infty} 0,$$

ou seja, $Tz_{n_k} \longrightarrow Tz_0$, o que é uma contradição.

□

Observação 3.1.5. Uma consequência imediata do Teorema 3.1.4 é que se T_1 e T_2 são operadores lineares compactos de X em Y e α é um escalar, então $T_1 + T_2$ e αT_1 também são operadores compactos. Segue então que $K(X, Y)$ é um subespaço vetorial de $B(X, Y)$.

Vejamos agora que no caso particular em que Y é um espaço de Banach, então $K(X, Y)$ é um subespaço vetorial fechado em $B(X, Y)$.

Teorema 3.1.6. *Sejam X um espaço normado e Y um espaço de Banach. Se $T_n : X \longrightarrow Y$ é uma sequência de operadores compactos, tal que $T_n \longrightarrow T$ em $B(X, Y)$ então $T \in K(X, Y)$.*

Demonstração. Vamos usar um "método diagonal" para mostrar que para qualquer sequência limitada (x_m) em X a sua imagem (Tx_m) possui uma subsequência convergente em Y.

Como T_1 é compacto, (x_m) possui uma subsequência $(x_{(1,m)})$ tal que $(T_1 x_{(1,m)})$ é de Cauchy. De modo semelhante $(x_{(1,m)})$ possui uma subsequência convergente $(x_{(2,m)})$ tal que $(T_2 x_{(2,m)})$ é de Cauchy. Continuando com esse processo, vemos que a sequência diagonal $(y_m) = (x_{(m,m)})$ é subsequência de (x_m) tal que para todo $n \in \mathbb{N}$ fixo a sequência $(T_n y_m)_{m\in\mathbb{N}}$ é de Cauchy. Seja $C > 0$ tal que $\|x_m\| \leq C$, para todo m. Portanto $\|y_m\| \leq C$ para todo m.

Dado $\epsilon > 0$, como $T_m \longrightarrow T$, existe $n = p$ tal que $\|T - T_p\| < \frac{\epsilon}{3C}$. Como $(T_p y_m)_{m\in\mathbb{N}}$ é de Cauchy, existe n_0 tal que

$$\|T_p y_j - T_p y_k\| < \frac{\epsilon}{3}, \quad j,k > n_0.$$

Logo, para todo $j, k > n_0$:

$$\begin{aligned}
\|Ty_j - Ty_k\| &\leq \|Ty_j - T_p y_j\| + \|T_p y_j - T_p y_k\| + \|T_p y_k - Ty_k\| \\
&\leq \|T - T_p\|\|y_j\| + \frac{\epsilon}{3} + \|T_p - T\|\|y_k\| \\
&< \frac{\epsilon}{3C}C + \frac{\epsilon}{3} + \frac{\epsilon}{3C}C = \epsilon.
\end{aligned}$$

Isso mostra que (Ty_m) é uma sequência de Cauchy, logo convergente pois Y é completo. Como (y_m) é uma subsequência da sequência limitada (x_n), concluímos que T é compacto pelo Teorema 3.1.2. □

Definição 3.1.7. *Um operador $T \in B(X,Y)$ é chamado de operador de posto finito quando* $\dim Im(T) < \infty$.

Teorema 3.1.8. *Sejam X e Y espaços normados e $T : X \longrightarrow Y$ um operador linear.*

1. *Se $T \in B(X,Y)$ tem posto finito, então $T \in K(X,Y)$;*

2. *Se $\dim X < \infty$, então $T \in K(X,Y)$.*

Demonstração. Iniciemos provando o item 1. Seja (x_n) um sequência limitada qualquer em X. Então, como T é limitado, segue que $\|Tx_n\| \leq \|T\|\|x_n\|$, isso mostra que $A_n = \{Tx_n;\ n \in \mathbb{N}\}$ é um conjunto limitado. Portanto $\overline{A_n}$ é fechado e limitado e como a dimensão de $T(X)$ é finita segue que $\overline{A_n}$ é compacto, ou seja, A_n é relativamente compacto. Deste modo (Tx_n) possui subsequência convergente e portanto T é compacto pelo Teorema 3.1.2.

Para provar o segundo item, uma vez que $\dim X < \infty$ temos que T é limitado e como $\dim T(X) \leq \dim X < \infty$, segue pelo resultado anterior que T é compacto.

□

Exemplo 3.1.9. Vejamos que o operador $T : l^2 \longrightarrow l^2$ definido por $x = (\xi_j) \longmapsto y = (\eta_j) = Tx$, onde $\eta_j = \frac{\xi_j}{j}$ para $j = 1, 2, \ldots$, é compacto.

De fato, obviamente temos que T é linear e se $x = (\xi_j) \in l^2$, então $y = (\eta_j) \in l^2$, logo T está bem definido. Considere para cada $n \in \mathbb{N}$ o operador linear $T_n : l^2 \longrightarrow l^2$ definido por

$$T_n x = \left(\xi_1, \frac{\xi_2}{2}, \frac{\xi_3}{3}, \ldots, \frac{\xi_n}{n}, 0, 0, \ldots\right).$$

Como T_n tem posto finito, segue pelo Teorema 3.1.8 que para cada n o operador T_n é compacto. Além disso,

$$\begin{aligned}\big\|(T - T_n)x\big\|^2 &= \sum_{j=n+1}^{\infty} |\eta_j|^2 = \sum_{j=n+1}^{\infty} \frac{1}{j^2}|\xi_j|^2 \\ &\leq \frac{1}{(n+1)^2} \sum_{j=n+1}^{\infty} |\xi_j|^2 \leq \frac{\|x\|^2}{(n+1)^2}.\end{aligned}$$

Tomando supremo sobre todo x tal que $\|x\| = 1$ segue que

$$\|T - T_n\| \leq \frac{1}{n+1}.$$

Portanto $T_n \longrightarrow T$ e T é compacto por 3.1.6.

3.1.1 Introdução a teoria espectral para operadores compactos

Agora apresentaremos algumas propriedades espectrais dos operadores compactos. O primeiro teorema irá tratar de uma importante característica sobre os autovetores de operadores compactos.

Teorema 3.1.10. *O conjunto de autovalores de um operador linear compacto $T : X \longrightarrow X$ em um espaço normado X é enumerável (pode ser finito ou vazio), e o único ponto de acumulação possível é $\lambda = 0$.*

Demonstração. Vamos mostrar que para todo $k > 0$ real, temos que o conjunto dos $\lambda \in \sigma_p(T)$ tal que $|\lambda| \geq k$ é finito.

Suponha que o contrário acontece para algum $k_0 > 0$. Então existe uma sequência infinita (λ_n) de autovalores distintos tal que $|\lambda_n| \geq k_0$. Também temos que existem $x_n \neq 0$ tais que $Tx_n = \lambda_n x_n$. O conjunto formado pelos autovetores x_n é linearmente independente, pois são autovetores que correspondem a autovalores distintos. Tome $M_n = \text{span}\{x_1, \ldots, x_n\}$. Deste modo, cada $x \in M_n$ tem uma única representação

$$x = \alpha_1 x_1 + \cdots + \alpha_n x_n.$$

Aplicando $T - \lambda_n I$ na igualdade acima e usando que $Tx_j = \lambda_j x_j$ obtemos:

$$(T - \lambda_n I)x = \alpha_1(\lambda_1 - \lambda_n)x_1 + \cdots + \alpha_{n-1}(\lambda_{n-1} - \lambda_n)x_{n-1}.$$

Note que x_n não aparece mais do lado direito, portanto

$$(T - \lambda_n I)x \in M_{n-1}, \text{ para todo } x \in M_n. \tag{3.1}$$

Temos que os subespaços M_n são fechados, pois têm dimensão finita. Pelo Lema de Riesz (1.2.17) existe uma sequência (y_n) tal que

$$y_n \in M_n, \ \|y_n\| = 1, \quad \|y_n - x\| \geq \frac{1}{2}, \text{ para todo } x \in M_{n-1}.$$

Mostraremos que

$$\|Ty_n - Ty_m\| \geq \frac{1}{2}k_0, \qquad (n > m),$$

desta forma (Ty_n) não possui subsequência convergente pois $k_0 > 0$.

Isso contradiz a compacidade de T, pois a sequência (y_n) é limitada.

Observe que:

$$Ty_n - Ty_m = (Ty_n - Ty_m) + \lambda_n y_n - \lambda_n y_n = \lambda_n y_n - \tilde{x},$$

onde $\tilde{x} = \lambda_n y_n - Ty_n + Ty_m$.

Seja $m < n$. Vamos mostrar que $\tilde{x} \in M_{n-1}$. Como $m \leq n-1$, vemos que $y_m \in M_m \subset M_{n-1} = \text{span}\{x_1, \ldots, x_{n-1}\}$. Portanto $Ty_m \in M_{n-1}$, pois $Tx_j = \lambda_j x_j$. Segue, por (3.1), que

$$\lambda_n y_n - Ty_n = -(T - \lambda_n I)y_n \in M_{n-1}.$$

Deste modo, $\tilde{x} = (\lambda_n y_n - Ty_n) + Ty_m \in M_{n-1}$. Uma vez que $|\lambda_n| \geq k_0$ também segue que

$x = \lambda_n^{-1}\tilde{x} \in M_{n-1}$. Assim,

$$\|\lambda_n y_n - \tilde{x}\| = |\lambda_n|\|y_n - x\| \geq \frac{1}{2}k_0.$$

Como $Ty_n - Ty_m = \lambda_n y_n - \tilde{x}$ segue que $\|Ty_n - Ty_m\| \geq \frac{1}{2}k_0$, como queríamos. Logo, é falso que existem infinitos autovalores que satisfazem $|\lambda_n| \geq k_0$ para algum $k_0 > 0$ e o teorema está demonstrado.

□

Observação 3.1.11. O Teorema 3.1.10 nos permite concluir que se um operador linear compacto possui infinitos autovalores, podemos arranjá-los em uma sequência que converge para 0.

O próximo resultado irá garantir que a composição de um operador linear compacto com um operador linear limitado é um operador linear compacto.

Lema 3.1.12. *Seja $T : X \longrightarrow X$ operador linear compacto e $S : X \longrightarrow X$ operador linear limitado em um espaço normado X. Então TS e ST são operadores compactos.*

Demonstração. Considere um conjunto limitado $B \subset X$ qualquer. Como S é operador linear limitado, temos que $S(B)$ é um conjunto limitado, deste modo $T(S(B))$ é relativamente compacto, pois T é operador compacto. Logo TS é compacto pelo Teorema 3.1.2.

Vejamos agora que ST também é compacto. Considere (x_n) uma sequência limitada qualquer em X. Então (Tx_n) possui subsequência convergente (Tx_{n_k}), além disso (STx_{n_k}) converge, pois S é contínuo. Assim ST é compacto pelo Teorema 3.1.2.

□

O próximo resultado evidencia a semelhança da teoria espectral para operadores lineares compactos com operadores lineares em espaços de dimensão finita, o que basicamente se resume ao estudo dos autovalores de matrizes.

Teorema 3.1.13. *Seja $T : X \longrightarrow X$ um operador linear compacto em um espaço normado X. Então para todo $\lambda \neq 0$ o núcleo do operador $T_\lambda = T - \lambda I$ tem dimensão finita.*

Demonstração. Note que basta mostrar que a bola fechada unitária $B_{\ker(T_\lambda)}$ em $\ker(T_\lambda)$ é compacta e aplicar o resultado do Teorema 1.2.18.

Seja (x_n) uma sequência em $B_{\ker(T_\lambda)}$. Temos que (x_n) é limitada, pois $\|x_n\| \leq 1$, e (Tx_n) possui uma subsequência convergente (Tx_{n_k}), pois T é compacto. Por outro lado, como $\lambda \neq 0$

segue que

$$T_\lambda x_n = Tx_n - \lambda x_n = 0 \;\;\Rightarrow\;\; x_n = \lambda^{-1} T x_n \;\; .$$

Consequentemente $(x_{n_k}) = (\lambda^{-1} T x_{n_k})$ é uma subsequência convergente e o seu limite está em $B_{\ker(T_\lambda)}$, pois $B_{\ker(T_\lambda)}$ é fechado. Logo, por definição, temos que o conjunto $B_{\ker(T_\lambda)}$ é compacto. Isso prova que $\dim\ker(T_\lambda) < \infty$ pelo Teorema 1.2.18. □

Corolário 3.1.14. *Como consequência do Teorema 3.1.13 temos que*

$$\dim\ker(T_\lambda^n) < \infty, \qquad n = 1, 2, \ldots \tag{3.2}$$

e

$$\{0\} = \ker(T_\lambda^0) \subset \ker(T_\lambda) \subset \ker(T_\lambda^2) \subset \cdots . \tag{3.3}$$

Demonstração. Para provar (3.3) note que como T_λ é linear então $T_\lambda 0 = 0$. Assim para cada $n \in \mathbb{N}$ temos que $T_\lambda^n x = 0$ implica $T_\lambda^{n+1} x = 0$.

Para provar (3.2) fazemos uso do Teorema Binomial obtendo

$$\begin{aligned} T_\lambda^n = (T - \lambda I)^n &= \sum_{k=0}^{n} \binom{n}{k} T^k (-\lambda)^{n-k} \\ &= (-\lambda)^n I + T \sum_{k=1}^{n} \binom{n}{k} T^{k-1} (-\lambda)^{n-k} \\ &= W - \mu I, \end{aligned}$$

onde $\mu = (-1)^{n+1}\lambda^n$ e $W = TS = ST$, com $S = \sum_{k=1}^{n} \binom{n}{k} T^{k-1}(-\lambda)^{n-k}$.

Como T é compacto e S é limitado, então usando o Lema 3.1.12 segue que W é compacto. Finalmente, aplicando o Teorema 3.1.13 em $W - \mu I$ segue (3.2). □

Lembremos que o núcleo de um operador linear limitado T sempre é fechado, porém a sua imagem não é necessariamente fechada. Porém, se T é compacto, então $\mathrm{Im}(T_\lambda)$ é fechada para todo $\lambda \neq 0$ e o mesmo vale para $T_\lambda^2, T_\lambda^3, \ldots$, ou seja, $T_\lambda^n(X)$ é fechada, para todo $n \in \mathbb{N}$. Vamos inicialmente provar o caso $n = 1$.

Teorema 3.1.15. *Seja $T : X \longrightarrow X$ um operador linear compacto em um espaço normado X. Então para todo $\lambda \neq 0$ a imagem do operador $T_\lambda = T - \lambda I$ é fechada.*

Demonstração. De maneira indireta vamos provar o teorema por contradição, assumindo que

$T_\lambda(X)$ não é fechada.

Seguem em detalhes a demonstração em etapa.

(1ª etapa) Se $T_\lambda(X)$ não é fechado, então existe $y \in \overline{T_\lambda(X)}$ tal que $y \notin T_\lambda(X)$ e uma sequência $(x_n) \in X$ tal que

$$y_n = T_\lambda x_n \longrightarrow y.$$

Uma vez que $0 \in T_\lambda(X)$ e $y \notin T_\lambda(X)$, então $y \neq 0$. Portanto $y_n \neq 0$ e $x_n \notin \ker(T_\lambda)$ para todo n suficientemente grande. Considerando uma subsequência, podemos assumir sem perda de generalidade que isso ocorre para todo n. Como $\ker(T_\lambda)$ é fechado, então para cada n a distância de x_n até $\ker(T_\lambda)$ é positiva, ou seja,

$$\delta_n \doteq \inf_{z \in \ker(T_\lambda)} \|x_n - z\| > 0.$$

Pela definição de ínfimo existe uma sequência $(z_n) \subset \ker(T_\lambda)$ tal que

$$a_n = \|x_n - z_n\| < 2\delta_n. \tag{3.4}$$

(2ª etapa) Vejamos agora que

$$a_n = \|x_n - z_n\| \longrightarrow \infty \qquad (n \longrightarrow \infty).$$

Suponha que a afirmação acima é falsa. Então $(x_n - z_n)$ possui uma subsequência limitada. Como T é compacto, segue que $(T(x_n - z_n))$ possui uma subsequência convergente. Como $\lambda \neq 0$ temos que $T_\lambda = T - \lambda I$ implica $I = \lambda^{-1}(T - T_\lambda)$. Lembrando que $z_n \in \ker(T_\lambda)$ temos $T_\lambda z_n = 0$, desta forma

$$x_n - z_n = \frac{1}{\lambda}(T - T_\lambda)(x_n - z_n) = \frac{1}{\lambda}\left[T(x_n - z_n) - T_\lambda x_n\right].$$

Uma vez que $(T(x_n - z_n))$ possui subsequência convergente e $(T_\lambda x_n)$ converge (conforme visto na 1ª etapa), portanto $(x_n - z_n)$ possui subsequência convergente, digamos $x_{n_k} - z_{n_k} \longrightarrow v$.

Como T é contínuo, pois T é compacto, então T_λ é contínuo. Assim,

$$T_\lambda(x_{n_k} - z_{n_k}) \longrightarrow T_\lambda v.$$

Por outro lado, temos que $T_\lambda z_{n_k} = 0$, pois $z_n \in \ker(T_\lambda)$, assim

$$T_\lambda(x_{n_k} - z_{n_k}) = T_\lambda x_{n_k} \longrightarrow y.$$

Desse modo, pela unicidade do limite temos que $T_\lambda v = y$. Logo $y \in T_\lambda(X)$, o que contraria o fato que $y \notin T_\lambda(X)$.

(3ª etapa) Uma vez que para cada n temos que a_n definido anteriormente não é nulo, definimos

$$w_n = \frac{1}{a_n}(x_n - z_n). \tag{3.5}$$

Logo, temos $\|w_n\| = 1$. Como $a_n \longrightarrow \infty$, $T_\lambda z_n = 0$ e $(T_\lambda x_n)$ converge, segue que

$$T_\lambda w_n = \frac{1}{a_n} T_\lambda x_n \longrightarrow 0.$$

Usando novamente que $I = \lambda^{-1}(T - T_\lambda)$, obtemos

$$w_n = \frac{1}{\lambda}(Tw_n - T_\lambda w_n).$$

Como T é compacto e (w_n) é limitado, então (Tw_n) possui uma subsequência convergente. Além disso, $(T_\lambda w_n)$ é convergente, conforme visto anteriormente. Assim, (w_n) possui uma subsequência convergente, digamos

$$w_{n_j} \longrightarrow w.$$

Como $T_\lambda w_{n_j} \longrightarrow 0$ segue que $T_\lambda w = 0$. Portanto, $w \in \ker(T_\lambda)$ e como $z_n \in \ker(T_\lambda)$, temos

$$u_n = z_n + a_n w \in \ker(T_\lambda).$$

Desta forma, para a distância entre x_n e u_n teremos $\|x_n - u_n\| \geq \delta_n$. Segue por (3.4) e (3.5) que

$$\delta_n \leq \big\|x_n - (z_n + a_n w)\big\| = \big\|(a_n w_n + z_n) - z_n - a_n w\big\| = a_n\|w_n - w\| < 2\delta_n\|w_n - w\|,$$

para todo $n \in \mathbb{N}$.

Dividindo a desigualdade acima por $2\delta_n > 0$, obtemos $\frac{1}{2} < \|w_n - w\|$ para todo $n \in \mathbb{N}$, o que contradiz o fato que existe uma subsequência $w_{n_j} \longrightarrow w$, finalizando a demonstração.

□

Corolário 3.1.16. *Sob as mesmas hipóteses do Teorema 3.1.15, a imagem de T_λ^n é fechada para todo $n = 0, 1, 2, \ldots$ Além disso,*

$$X = T_\lambda^0(X) \supset T_\lambda(X) \supset T_\lambda^2(X) \supset \cdots .$$

Demonstração. Para a primeira afirmação basta observar do Corolário 3.1.14 que $T_\lambda^n = W - \mu I$, onde W é compacto; deste modo basta utilizar o Teorema 3.1.15 em $W - \mu I$ e concluir o resultado. Para a segunda parte, usamos Indução Matemática. Temos que $T_\lambda^0(X) = I(X) \supset T_\lambda(X)$ e supondo que $T_\lambda^{n-1}(X) \supset T_\lambda^n(X)$, basta aplicar T_λ e concluiremos que $T_\lambda^n(X) \supset T_\lambda^{n+1}(X)$.

□

3.2 Decomposição SVD de operadores compactos em espaços de Hilbert

Nessa parte do trabalho estudaremos algumas propriedades dos operadores compactos em espaços de Hilbert. O principal resultado que apresentaremos será a *decomposição em valores singulares* desses operadores. No que segue H denotará um espaço de Hilbert separável e consideraremos operadores compactos $T \in K(H)$.

Teorema 3.2.1. *Seja $T \in B(H)$. Então T é compacto se, e somente se, T^* é compacto.*

Demonstração. Suponha que T é compacto e vejamos que T^* também é compacto. Dada uma sequência $(x_n) \subset H$, tal que $x_n \longrightarrow 0$ fracamente, então pelo Lema 2.8.2. existe $c > 0$ tal que $\|x_n\| \leq c$, para todo $n \in \mathbb{N}$. Como T^* é limitado, então $T^* x_n \longrightarrow 0$ fracamente em H. Logo, usando a desigualdade de Cauchy-Schwarz temos:

$$\begin{aligned}
\left\|T^* x_n\right\|^2 &= \langle T^* x_n, T^* x_n \rangle \\
&= \langle T T^* x_n, x_n \rangle \\
&\leq \left\|T T^* x_n\right\| \|x_n\| \\
&\leq c \left\|T T^* x_n\right\|, \quad \forall n \in \mathbb{N}.
\end{aligned}$$

Como T é um operador compacto e $T^* x_n \longrightarrow 0$ fracamente, então segue do Teorema 3.1.4 que $\left\|T(T^* x_n)\right\| \longrightarrow 0$. Portanto, $\|T^* x_n\| \longrightarrow 0$ e segue novamente do Teorema 3.1.4 que T^* é compacto.

Reciprocamente, se T^* é compacto então pelo que vimos anteriormente $T = (T^*)^*$ também é compacto. $\square$

Agora veremos como obter operadores compactos a partir de sequências ortonormais.

Proposição 3.2.2. *Sejam (v_n) e (u_n) sequências ortonormais no espaço de Hilbert H e (λ_n) uma sequência numérica (real ou complexa) tal que $\lambda_n \longrightarrow 0$. Então, o operador definido por*

$$Tx = \sum_{n=1}^{\infty} \lambda_n \langle x, v_n \rangle u_n, \quad x \in H,$$

é um operador compacto.

Demonstração. Seja $(x_n) \subset H$ tal que $x_n \longrightarrow 0$ fracamente. Logo, considerando os funcionais lineares $f_n(z) = \langle z, v_n \rangle$, $n \in \mathbb{N}$, temos que $f_n(x_k) \longrightarrow f_n(0) = 0$ quando $k \longrightarrow \infty$, ou seja, para cada n temos que $\langle x_k, v_n \rangle \longrightarrow 0$ se $k \longrightarrow \infty$. Além disso, existe $c > 0$ tal que $\|x_k\| \leq c$, para todo $k \in \mathbb{N}$. Dado $\epsilon > 0$, tome $N \in \mathbb{N}$ tal que $|\lambda_n| < \epsilon$ se $n > N$. Assim, pelo Teorema de Pitágoras temos

$$\begin{aligned}
\|Tx_k\|^2 &= \sum_{n=1}^{\infty} |\lambda_n|^2 |\langle x_k, v_n \rangle|^2 \\
&= \sum_{n=1}^{N} |\lambda_n|^2 |\langle x_k, v_n \rangle|^2 + \sum_{n=N+1}^{\infty} |\lambda_n|^2 |\langle x_k, v_n \rangle|^2 \\
&\leq \sum_{n=1}^{N} |\lambda_n|^2 |\langle x_k, v_n \rangle|^2 + \epsilon^2 \sum_{n=N+1}^{\infty} |\langle x_k, v_n \rangle|^2 \\
(*) &\leq \sum_{n=1}^{N} |\lambda_n|^2 |\langle x_k, v_n \rangle|^2 + \epsilon^2 \|x_k\|^2 \\
&\leq \sum_{n=1}^{N} |\lambda_n|^2 |\langle x_k, v_n \rangle|^2 + \epsilon^2 c^2,
\end{aligned}$$

onde $(*)$ segue pela desigualdade de Bessel.

Como a soma no lado esquerdo da desigualdade anterior é finita e $|\langle x_k, v_n \rangle|^2 |\lambda_n|^2 \longrightarrow 0$ quando $k \longrightarrow \infty$, segue que $\|Tx_k\| \longrightarrow 0$ quando $k \longrightarrow \infty$. Portanto, pelo Teorema 3.1.4, T é um operador compacto. $\square$

Uma questão natural que surge da proposição anterior é sobre a validade da sua recíproca, ou seja, dado um operador compacto T definido num espaço de Hilbert H, é possível encontrar sequências ortonormais (v_n), (u_n) e uma sequência numérica (λ_n), com $\lambda_n \longrightarrow 0$, de forma que

$$T(x) = \sum_{n=1}^{\infty} \lambda_n \langle x, v_n \rangle u_n,$$

para todo $x \in H$?

A resposta para esta questão é afirmativa e os próximos resultados são dedicados à prová-la. Começamos com o caso em que H tem dimensão finita. Neste caso todo operador linear é compacto. De fato, basta usar a caracterização por sequências dada no Teorema 3.1.2 e o fato que em um espaço vetorial de dimensão finita toda sequência limitada possui uma subsequência convergente.

Teorema 3.2.3. *Seja H um espaço de Hilbert com $\dim H = n < \infty$ e $T : H \longrightarrow H$ operador linear. Fixada uma base ortonormal $\{e_i\}_{i=1}^n$ em H, o operador T pode ser escrito como*

$$T = U\Sigma V^*,$$

sendo $U, V : H \longrightarrow H$ operadores lineares unitários e $\Sigma : H \longrightarrow H$ dado por $\Sigma(x) = \sum_{i=1}^n \alpha_i \sigma_i e_i$, com $x = \sum_{i=1}^n \alpha_i e_i$.

Demonstração. Seja $V = \{v_1, v_2, \ldots, v_n\} \in H$ uma base ortonormal em H formada por autovetores do operador autoadjunto T^*T e $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ os autovalores correspondentes. Denotemos por r o posto de T, logo $r \leq n$ é o número de autovalores não nulos de T. Então $\{Tv_1, Tv_2, \ldots, Tv_r\}$ forma uma base ortogonal para $T(H)$. De fato,

$$\langle Tv_j, Tv_k \rangle = \langle v_j, T^*Tv_k \rangle = \overline{\lambda_k} \langle v_j, v_k \rangle = 0, \qquad \text{se } j \neq k.$$

Logo, $\{Tv_j\}_{j=1}^n$ é ortogonal. Por outro lado, note que

$$\left\|Tv_j\right\|^2 = \langle Tv_j, Tv_j \rangle = \langle v_j, T^*Tv_j \rangle = \lambda_i \doteq \sigma_i^2, \tag{3.6}$$

ou seja, $\lambda_j \geq 0$, para cada $j = 1, \ldots, n$. Como estamos supondo que existem $r \leq n$ autovetores não nulos, então $\lambda_{r+1} = \cdots = \lambda_n = 0$.

Assim, dado $y = Tx \in H$, $x = \sum_{j=1}^n \beta_j v_j$, teremos por (3.6) o seguinte:

$$y = T\left(\sum_{j=1}^n \beta_j v_j\right) = \sum_{j=1}^n \beta_j T v_j = \sum_{j=1}^r \beta_j T v_j,$$

pois $Tv_j = 0$, $j = r+1, \dots, n$. Segue que $\{Tv_1, Tv_2, \dots, Tv_r\}$ gera $T(H)$.

Normalizando cada vetor Tv_i, $i = 1, \dots, r$, obtemos uma base ortonormal $\{u_1, \dots, u_r\}$ de $T(H)$, onde

$$u_i = \frac{Tv_i}{\|Tv_i\|} = \frac{Tv_i}{\sigma_i} \iff Tv_i = \sigma_i u_i \quad (1 \leq i \leq r).$$

Estendemos $\{u_1, \dots, u_r\}$ para uma base ortonormal $\{u_1, \dots, u_n\}$ de H. Como $\{e_i\}_{i=1}^n$ é base, todo $x \in H$ pode ser escrito na forma $x = \sum_{i=1}^n \alpha_i e_i$. Definimos os operadores U, V da seguinte forma

$$\begin{aligned} U : H \longrightarrow H, &\quad U(x) = \sum_{i=1}^n \alpha_i u_i, \\ V : H \longrightarrow H, &\quad V(x) = \sum_{i=1}^n \alpha_i v_i. \end{aligned}$$

Relembrando que $Tv_j = 0$ para $j = r+1, \dots, n$, temos que

$$\begin{aligned} TV(x) &= \sum_{i=1}^n T(\alpha_i v_i) = \sum_{i=1}^r \alpha_i T v_i = \sum_{i=1}^r \alpha_i \sigma_i u_i = \sum_{i=1}^r \alpha_i \sigma_i U(e_i) \\ &= \sum_{i=1}^r U(\alpha_i \sigma_i e_i) = U\left(\sum_{i=1}^r \alpha_i \sigma_i e_i\right) = U\Sigma(x), \quad x \in H, \end{aligned} \tag{3.7}$$

sendo $\Sigma : H \longrightarrow H$ definido por $\Sigma(x) = \sum_{i=1}^r \alpha_i \sigma_i e_i$.

Vamos agora mostrar que U e V são operadores unitários. Para isso, provaremos que $V^*V = VV^* = I$. Para o operador U os argumentos são análogos.

Primeiramente, vamos determinar V^*. Dados $x, y \in H$ temos

$$\begin{aligned} \langle Vx, y\rangle &= \left\langle \sum_{j=1}^n \alpha_j v_j, y \right\rangle = \sum_{j=1}^n \alpha_j \langle v_j, y\rangle = \sum_{j=1}^n \langle x, e_j\rangle \overline{\langle y, v_j\rangle} \\ &= \sum_{j=1}^n \langle x, \langle y, v_j\rangle e_j\rangle = \left\langle x, \sum_{j=1}^n \langle y, v_j\rangle e_j \right\rangle = \langle x, V^*(y)\rangle. \end{aligned}$$

Logo, $V^*(y) = \sum_{j=1}^n \langle y, v_j\rangle e_j$.

Vejamos que $V^*V(x) = VV^*(x) = x$. De fato, usando o fato que $\langle v_i, v_j\rangle = 0$ para $i \neq j$ e

$\langle v_j, v_j \rangle = 1$ obtemos

$$
\begin{aligned}
V^*V(x) = V^* \left(\sum_{i=1}^{n} \alpha_i v_i \right) &= \sum_{j=1}^{n} \left\langle \sum_{i=1}^{n} \alpha_i v_i, v_j \right\rangle e_j \\
&= \sum_{j=1}^{n} \sum_{i=1}^{n} \alpha_i \langle v_i, v_j \rangle e_j = \sum_{j=1}^{n} \alpha_j \langle v_j, v_j \rangle e_j = x.
\end{aligned}
$$

Para o outro lado, note que cada e_j tem uma representação $e_j = \sum_{k=1}^{n} \beta_k v_k$. Portanto,

$$
\langle e_i, v_j \rangle = \left\langle \sum_{k=1}^{n} \beta_k v_k, v_j \right\rangle = \sum_{k=1}^{n} \beta_k \langle v_k, v_j \rangle = \beta_j.
$$

Deste modo, temos que

$$
\begin{aligned}
VV^*(x) = V \left(\sum_{j=1}^{n} \langle x, v_j \rangle e_j \right) &= \sum_{j=1}^{n} \langle x, v_j \rangle v_j = \sum_{j=1}^{n} \left\langle \sum_{i=1}^{n} \alpha_i e_i, v_j \right\rangle v_j \\
&= \sum_{j=1}^{n} \sum_{i=1}^{n} \alpha_i \langle e_i, v_j \rangle v_j = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \langle e_i, v_j \rangle v_j = \sum_{j=1}^{n} \alpha_j \beta_j v_j = \sum_{j=1}^{n} \alpha_j e_j = x.
\end{aligned}
$$

Portanto, V é um operador unitário. Logo $V^* = V^{-1}$ e segue por (3.7) que

$$
T(x) = U\Sigma V^*(x), \quad x \in H.
$$

Dessa forma, podemos escrever

$$
T(x) = U\Sigma \left(\sum_{j=1}^{n} \langle x, v_i \rangle e_i \right) = U \left(\sum_{j=1}^{r} \sigma_i \langle x, v_i \rangle e_i \right) = \sum_{j=1}^{r} \sigma_i \langle x, v_i \rangle u_i.
$$

□

Uma consequência importante do Teorema 3.2.3 é a conhecida decomposição em valores singulares (*Singular Value Decomposition - SVD*) de matrizes. Mais precisamente temos o seguinte corolário.

Corolário 3.2.4 (SVD). *Se $A_{n\times n}$ é uma matriz real (ou complexa) e $\{e_k\}_{k=0}^{n}$ é a base canônica em $\mathbb{R}^n$ (ou $\mathbb{C}^n$), então existem matrizes ortogonais (ou unitárias) U, V e uma matriz diagonal Σ tal que*

$$
A = U\Sigma V^*.
$$

Nosso objetivo agora é generalizar o resultado obtido no Teorema 3.2.3 para operadores compactos definidos em espaços de Hilbert com dimensão qualquer. Comecemos provando esse resultado para operadores compactos auto-adjuntos.

Teorema 3.2.5. *Seja T um operador compacto e auto-adjunto definido no espaço de Hilbert H. Então, existem uma sequência de números reais (λ_n), com $\lambda_n \longrightarrow 0$, e uma sequência ortonormal $(e_n)_{n\in\mathbb{N}}$ em H tais que*

$$Tx = \sum_{n=1}^{\infty} \lambda_n \langle x, e_n\rangle e_n, \quad \forall x \in H.$$

Demonstração. Se $T = 0$ o resultado é imediato tomando $\lambda_n = 0$ e (e_n) uma sequência ortonormal qualquer. Suponha então $T \neq 0$. Vamos provar inicialmente que T possui autovalores não nulos. Como T é compacto sabemos pelo Teorema 3.1.10 que o conjunto dos autovalores de T é enumerável, podendo ser finito, e $\lambda = 0$ é o único ponto de acumulação possível. Além disso, como T é auto-adjunto então todos os autovalores são reais.

Primeira parte: Uma vez que T é auto-adjunto, então pelo Teorema 9.2-2 em [4], temos que

$$\|T\| = \max\{|m|, |M|\},$$

sendo

$$m = \inf_{\|x\|=1} \langle Tx, x\rangle \quad \text{e} \quad M = \sup_{\|x\|=1} \langle Tx, x\rangle.$$

Supondo inicialmente que $\|T\| = |M|$, temos que $M > 0$, caso contrário, $m \leq M \leq 0$ o que implicaria $\|T\| = |m|$, o que é uma contradição. Assim, $\|T\| = \sup\{\langle Tx, x\rangle : \|x\| = 1\}$.

Logo, para cada $n \in \mathbb{N}$, existe $x_n \in H$ com $\|x_n\| = 1$ tal que

$$\|T\| - \frac{1}{n} < \langle Tx_n, x_n\rangle \leq \|T\|.$$

Fazendo $n \longrightarrow \infty$ temos

$$\langle Tx_n, x_n\rangle \longrightarrow \|T\|. \tag{3.8}$$

No caso em que $\|T\| = |m|$, então $m < 0$. Nesse caso, temos

$$\|-T\| = \|T\| = -m = -\inf_{\|x\|=1} \langle Tx, x\rangle = \sup_{\|x\|=1} \langle -Tx, x\rangle.$$

Assim, sem perda de generalidade podemos supor que existe $\{x_n\} \subset H$ com $\|x_n\| = 1$ tal que $\langle Tx_n, x_n\rangle \longrightarrow \|T\|$, pois caso contrário basta considerar o operador $\widetilde{T} = -T$ para concluir o restante da demonstração.

Por outro lado, pela desigualdade de Cauchy-Schwarz temos

$$\langle Tx_n, x_n\rangle \leq \|Tx_n\|\|x_n\| = \|Tx_n\| \leq \|T\|\|x_n\| = \|T\|\,. \tag{3.9}$$

Logo de (3.8) e (3.9) temos

$$\lim_{n\to\infty} \|Tx_n\| = \|T\|\,. \tag{3.10}$$

Dessa forma, segue de (3.8) e de (3.10) que

$$\big\|Tx_n - \|T\|\,x_n\big\|^2 = \|Tx_n\|^2 - 2\langle Tx_n, x_n\rangle\|T\| + \|T\|^2\|x_n\|^2$$

e portanto

$$\lim_{n\to\infty} \big\|Tx_n - \|T\|\,x_n\big\|^2 = \|T\|^2 - 2\|T\|^2 + \|T\|^2 = 0.$$

Agora, usando o fato que T é compacto e $\{x_n\}$ é limitada, então existem uma subsequência $\{x_{n_k}\}$ e $y \in H$ tais que $Tx_{n_k} \longrightarrow y$. Assim, pelo que foi discutido acima temos

$$\big\|\, \|T\|\,x_{n_k} - y\big\| \leq \big\|\, \|T\|\,x_{n_k} - Tx_{n_k}\big\| + \big\|Tx_{n_k} - y\big\| \longrightarrow 0, \quad n \longrightarrow \infty.$$

Logo, $\|T\|\,x_{n_k} \longrightarrow y$. Assim, $x_{nk} \longrightarrow y/\|T\|$ e segue que $y \neq 0$ pois $\|x_{n_k}\| = 1$, para todo k.

Como T é um operador contínuo e $\|T\|\,x_{n_k} \longrightarrow y$, concluímos que

$$\|T\|\,Tx_{n_k} \longrightarrow Ty. \tag{3.11}$$

Por outro lado, como $Tx_{n_k} \longrightarrow y$, segue que

$$\|T\|\,Tx_{n_k} \longrightarrow \|T\|\,y. \tag{3.12}$$

De (3.11) e (3.12) temos que $Ty = \|T\|\,y$, ou seja, y é um autovetor e $\|T\|$ é um autovalor não nulo do operador T.

Segunda parte: Seja $(\mu_1, \mu_2, \ldots)$ a sequência (finita ou infinita) de todos os autovalores *distintos e não nulos* de T, de forma que $|\mu_1| \geq |\mu_2| \geq \ldots$. Assim, se $T_{\mu_j} = T - \mu_j I$ temos que $m_j = \dim\ker(T_{\mu_j}) < \infty$ pelo Teorema 3.1.13.

Considere agora a sequência dada da seguinte forma:

$$(\mu_1, \mu_1, \ldots, \mu_1, \mu_2, \mu_2, \ldots, \mu_2, \mu_3, \mu_3, \ldots, \mu_3, \ldots),$$

onde cada μ_j está repetido m_j vezes. Vamos denotar essa sequência por $(\lambda_1, \lambda_2, \lambda_3, \ldots)$.

Para cada $j \in \mathbb{N}$ existe uma base ortonormal para o autoespaço $\ker(T_{\mu_j})$. Como autovetores associados a autovalores distintos são ortogonais, uma vez que T é auto-adjunto, então obtemos uma sequência ortonormal

$$\{e_1, e_2, \ldots\} \text{ em } H,$$

a partir das bases ortonormais encontradas para os autoespaços $\ker(T_{\mu_j})$, de forma que $Te_j = \lambda_j e_j$, para cada $j \in \mathbb{N}$.

Como $\lambda_j \longrightarrow 0$, então pelo Teorema 3.2.2, consideramos o operador compacto $T_0 : H \longrightarrow H$ por

$$T_0 x = \sum_{j=1}^{\infty} \lambda_j \langle x, e_j \rangle e_j, \;\; x \in H.$$

Basta agora provar que $T = T_0$.

Seja $M \doteq \overline{\text{span}\{e_1, e_2, \ldots\}}$. Logo M é um subespaço fechado em H. Dado $x \in H$, temos que $S_N(x) = \sum_{j=1}^{N} \lambda_j \langle x, e_j \rangle e_j \in M$ e $S_N(x) \longrightarrow T_0 x$. Como M é fechado, então $T_0 x \in M$. Portanto, $T_0 H \subset M$.

Note também que, se $x \in \text{span}\{e_1, e_2, \ldots\}$ então

$$x = \sum_k a_k e_k \text{ (soma finita) } \Longrightarrow Tx = \sum_k a_k T e_k = \sum_k a_k \lambda_k e_k \in M.$$

Em geral, se $x \in M$, então existe $(x_n) \subset M$ tal que $x_n \longrightarrow x$. Segue que $Tx_n \longrightarrow Tx$ e como $Tx_n \in M$ para todo n, temos que $Tx \in M$, uma vez que M é fechado. Portanto, $TM \subset M$.

Se $x \in M^{\perp} = \{x \in H \mid \langle x, y \rangle = 0, \; \forall y \in M\}$, então $\langle x, e_n \rangle = 0$, para todo $n \in \mathbb{N}$. Logo

$$\langle Tx, e_n \rangle = \langle x, T^* e_n \rangle = \langle x, Te_n \rangle = \lambda_n \langle x, e_n \rangle = 0, \quad \forall n \in \mathbb{N}.$$

Segue que $Tx \perp e_n$, para todo $n \in \mathbb{N}$. Como $\text{span}\{e_1, e_2, \ldots\}$ é denso em M, segue que $Tx \in M^{\perp}$. Portanto, $T(M^{\perp}) \subset M^{\perp}$.

Afirmação: O operador $S \doteq T - T_0$ não possui autovalores não nulos.

Suponha, por contradição, que existe $\lambda \neq 0$ e $x \neq 0$ tais que $Sx = \lambda x$. Note primeiramente que $z = x - \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n \in M^{\perp}$, pois

$$\langle z, e_k \rangle = \langle x - \sum_n \langle x, e_n \rangle e_n, e_k \rangle = \langle x, e_k \rangle - \sum_n \langle x, e_n \rangle \langle e_n, e_k \rangle = \langle x, e_k \rangle - \langle x, e_k \rangle = 0$$

e como span$\{e_1, e_2, \ldots\}$ é denso em M, segue que $z \in M^{\perp}$.

Dessa forma,

$$Tz = T\left(x - \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n \right) = Tx - \sum_{n=1}^{\infty} \langle x, e_n \rangle Te_n = Tx - T_0 x = Sx = \lambda x.$$

Como $z \in M^{\perp}$ e $T(M^{\perp}) \subset M^{\perp}$ segue que $\lambda x = Tz \in M^{\perp}$. Assim, $x \in M^{\perp}$ e concluímos que $\langle x, e_n \rangle = 0$, para todo $n \in \mathbb{N}$, o que implica $T_0 x = 0$. Segue que $Tx = \lambda x$. Portanto, existe $k \in \mathbb{N}$ tal que $\lambda = \lambda_k$ e $x \in \ker(T_{\lambda_k}) \subset M$. Logo, $x \in M \cap M^{\perp}$, ou seja, $x = 0$, o que é uma contradição.

Afirmamos que T_0 é auto-adjunto. De fato, como λ_n é real, dados $x, y \in H$ temos

$$\begin{aligned} \langle T_0 x, y \rangle &= \langle \sum_n \lambda_n \langle x, e_n \rangle e_n, y \rangle = \sum_n \lambda_n \langle x, e_n \rangle \langle e_n, y \rangle = \sum_n \lambda_n \overline{\langle y, e_n \rangle} \langle x, e_n \rangle \\ &= \sum_n \lambda_n \langle x, \langle y, e_n \rangle e_n \rangle = \langle x, \sum_n \lambda_n \langle y, e_n \rangle e_n \rangle = \langle x, T_0 y \rangle. \end{aligned}$$

Segue que $T_0^* = T_0$.

Por fim, como $S = T - T_0$ é auto-adjunto e compacto, além disso, S não possui autovalores não nulos, então da primeira parte da demostração segue que $S = 0$, isto é, $T = T_0$. $\square$

Nosso objetivo é generalizar esse resultado para operadores compactos não necessariamente auto-adjuntos. Para isso, precisamos de algumas definições adicionais.

Definição 3.2.6. *Um operador auto-adjunto $T : H \longrightarrow H$ é dito positivo se $\langle Tx, x \rangle \geq 0$ para todo $x \in H$. Neste caso, escrevemos $T \geq 0$.*

Definição 3.2.7. *Seja $T : H \longrightarrow H$ um operador auto-adjunto, limitado e positivo. Dizemos que um operador limitado e auto-adjunto A é a raiz quadrada de T se,*

$$A^2 = T.$$

Se, além disso, o operador A é positivo, dizemos que ele é a raiz quadrada positiva de T e o *denotamos por*

$$A = T^{1/2}.$$

Pode ser mostrado que, para todo operador auto-adjunto positivo, existe um único operador A que é a raiz quadrada positiva de T como o definido acima (veja o Teorema 9.4-2 em [4]).

Dado um operador linear e limitado T definido em H, temos que T^*T é um operador limitado, auto-adjunto e positivo pois $\langle T^*Tx, x\rangle = \langle Tx, Tx\rangle \geq 0$. Assim, denotaremos por $|T| = (T^*T)^{1/2}$ a raiz quadrada positiva de T^*T.

Lembre-se que um operador linear T definido em um espaço de Hilbert H é chamado de *isometria* se $\langle x, y\rangle = \langle Tx, Ty\rangle$, para todo $x, y \in H$. Se T é isometria quando restrito ao subespaço $(\ker T)^\perp$, diremos que T é uma *isometria parcial.*

Teorema 3.2.8 (Decomposição Polar)**.** *Seja T um operador linear limitado definido no espaço de Hilbert H. Então existe uma isometria parcial U tal que $T = U|T|$. Além disso, U é unicamente determinado pela condição:* $\ker U = \ker T$ *e* $\mathscr{R}(U) = \overline{\mathscr{R}(T)}$.

Demonstração. Defina $U_0 : \mathscr{R}(|T|) \longrightarrow \mathscr{R}(T)$ por $U_0(|T|x) = Tx$, para todo $x \in H$. Vejamos que U_0 está bem definido e é uma isometria em $\mathscr{R}(|T|)$. De fato, para $x, y \in H$ temos

$$\langle U_0(|T|x), U_0(|T|y)\rangle = \langle Tx, Ty\rangle = \langle x, T^*Ty\rangle = \langle x, |T|^2y\rangle = \langle x, |T|^*|T|y\rangle = \langle |T|x, |T|y\rangle.$$

Logo, U_0 é uma isometria de $\mathscr{R}(|T|)$ em $\mathscr{R}(T)$.

Note que, em particular, se $x = y$ na igualdade acima temos que $\|Tx\| = \big\||T|x\big\|$. Assim, se $|T|x_1 = |T|x_2$, isto é, $|T|(x_1 - x_2) = 0$ concluímos que $\big\|T(x_1 - x_2)\big\| = \big\||T|(x_1 - x_2)\big\| = 0$. Portanto, $Tx_1 = Tx_2$ e segue que U_0 está bem definido.

Considerando $U_0 : \mathscr{R}(|T|) \longrightarrow \overline{\mathscr{R}(T)}$ temos, pelo Teorema 1.3.20, que existe uma extensão $\widetilde{U}_0 : \overline{\mathscr{R}(|T|)} \longrightarrow \overline{\mathscr{R}(T)}$, pois $\overline{\mathscr{R}(T)}$ é completo uma vez que é um subespaço fechado de H. Como U_0 é limitado, pois é uma isometria, então $\widetilde{U}_0$ também é limitado. Segue por continuidade de $\widetilde{U}_0$ e do produto interno que $\widetilde{U}_0$ também é uma isometria.

Considere agora a decomposição $H = \overline{\mathscr{R}(|T|)} \oplus \overline{\mathscr{R}(|T|)}^\perp$. Defina $U : H \longrightarrow H$ por $U(x) = U(x_1 + x_2) \doteq \widetilde{U}_0(x_1)$, onde $x = x_1 + x_2 \in H$, com $x_1 \in \overline{\mathscr{R}(|T|)}$ e $x_2 \in \overline{\mathscr{R}(|T|)}^\perp$. Claramente U é linear e limitado. Note que,

$$U_{\big|\mathscr{R}(|T|)} = U_0 \quad \text{e} \quad U_{\big|\overline{\mathscr{R}(|T|)}^\perp} = 0.$$

Portanto, para todo $x \in H$, temos $U(|T|x) = U_0(|T|x) = Tx$, ou seja, $U|T| = T$. Além disso, $\overline{\mathscr{R}(|T|)}^{\perp} \subset \ker U$.

Vejamos agora que U é uma isometria parcial e valem as condições do enunciado.

Afirmação: $\mathscr{R}(|T|)^{\perp} = \ker |T|^*$. De fato, como

$$\langle y, |T|x\rangle = \langle |T|^* y, x\rangle, \tag{3.13}$$

se $y \in \mathscr{R}(|T|)^{\perp}$ então $\langle y, |T|x\rangle = 0$, para todo $x \in H$. Logo, $|T|^* y = 0$ por (3.13).

Se $y \in \ker(|T|^*)$, então podemos concluir por (3.13) que $\langle y, |T|x\rangle = 0$, para todo $x \in H$, ou seja, $y \in \mathscr{R}(|T|)^{\perp}$. Portanto, segue a validade da afirmação anterior e sendo $|T|$ auto-adjunto, temos $\mathscr{R}(|T|)^{\perp} = \ker |T|^* = \ker |T|$.

Por outro lado, como $\|Tx\| = \big\||T|x\big\|$, para todo $x \in H$. Logo $Tx = 0$ se, e somente se, $|T|x = 0$. Assim, pela definição de U_0 temos $\ker U_0 = \ker T = \ker |T|$. Logo,

$$\mathscr{R}(|T|)^{\perp} = \ker T. \tag{3.14}$$

Além disso, como $\mathscr{R}(|T|)^{\perp} = \overline{\mathscr{R}(|T|)}^{\perp} \subset \ker U$, segue que $\ker T = \mathscr{R}(|T|)^{\perp} \subset \ker U$. Pela construção temos $\ker U \subset \ker T$. Logo,

$$\ker T = \ker U. \tag{3.15}$$

Deste modo, por (3.14) e (3.15) temos que $\ker U = \mathscr{R}(|T|)^{\perp}$ o que implica, tomando o ortogonal novamente, que $(\ker U)^{\perp} = \overline{\mathscr{R}(|T|)}$. Assim,

$$U\Big|_{(\ker U)^{\perp}} = U\Big|_{\overline{\mathscr{R}(|T|)}} = \widetilde{U}_0,$$

que é uma isometria. Segue que U é uma isometria parcial.

Por fim, provemos que $\mathscr{R}(U) = \overline{\mathscr{R}(T)}$. Pela definição do operador U temos diretamente que $\mathscr{R}(U) \subset \overline{\mathscr{R}(T)}$, logo basta provar a outra inclusão.

Se $y = \lim Tx_n \in \overline{\mathscr{R}(T)}$, então Tx_n é uma sequência de Cauchy, logo $|T|x_n$ também é uma sequência de Cauchy, uma vez que $\|Tx_n - Tx_m\| = \big\||T|x_n - |T|x_m\big\|$. Como $\overline{\mathscr{R}(|T|)}$ é completo, segue que $|T|x_n \longrightarrow z$, para algum $z \in \overline{\mathscr{R}(|T|)}$. Logo,

$$U(z) = \widetilde{U}_0(z) = \widetilde{U}_0(\lim |T|x_n) = \lim \widetilde{U}_0(|T|x_n) = \lim U_0(|T|x_n) = \lim Tx_n = y,$$

donde segue que $y \in \mathscr{R}(U)$.

Vejamos agora a unicidade. Suponha que U e V são isometrias parciais em H tais que $U|T| = T$ e $V|T| = T$ e além disso

$$\ker U = \ker T = \ker V \text{ e } \mathscr{R}(U) = \overline{\mathscr{R}(T)} = \mathscr{R}(V).$$

Assim, por (3.14) temos que $\ker U = (\mathscr{R}(|T|))^{\perp} = \ker V$. Logo, se $x \in \overline{\mathscr{R}(|T|)}^{\perp} = (\mathscr{R}(|T|))^{\perp}$, então $Ux = 0 = Vx$. Por outro lado, se $x \in \overline{\mathscr{R}(|T|)}$ então $x = \lim |T|x_n$, para $(|T|x_n) \subset \mathscr{R}(|T|)$. Nesse caso, segue que $Ux = U(\lim |T|x_n) = \lim U(|T|x_n) = \lim V(|T|x_n) = V(\lim |T|x_n) = Vx$. Portanto, $Ux = Vx$ para todo $x \in H$ e o teorema está demonstrado.

□

Proposição 3.2.9. *Seja T um operador linear limitado no espaço de Hilbert H. As seguintes afirmações são equivalentes:*

1. *T é compacto;*
2. *T^*T é compacto;*
3. *$|T| = (T^*T)^{1/2}$ é compacto.*

Demonstração. (1) $\Rightarrow$ (2): Se T é compacto, então pelo Teorema 3.2.1 o operador adjunto T^* também é compacto. Assim, pelo Lema 3.1.12 segue que T^*T é compacto.

(2) $\Rightarrow$ (3): Seja $S = |T|$, então $S^2 = T^*T$. Dada uma sequência (x_n) em H que converge fracamente para 0, segue da compacidade de S^2 e do Teorema 3.1.4 que $\left\|S^2x_n\right\| \longrightarrow 0$. Adicionalmente, pelo Lema 2.8.2 temos que a sequência $(\|x_n\|)$ é limitada. Portanto,

$$\|Sx_n\|^2 = \langle S^2x_n, x_n\rangle \leq \|x_n\| \, \|S^2x_n\| \longrightarrow 0, \quad (n \longrightarrow \infty).$$

Portanto $\|Sx_n\| \longrightarrow 0$ e novamente pelo Teorema 3.1.4 temos que S é compacto.

(3) $\Rightarrow$ (1): Suponha que $|T| = (T^*T)^{1/2}$ é compacto. Pelo Teorema 3.2.8 consideramos a decomposição polar de T dada por $T = U|T|$, onde U é uma isometria parcial, logo limitado. Portanto, pelo Lema 3.1.12 segue que T é compacto.

□

Agora já temos as ferramentas suficientes para generalizar o Teorema 3.2.5. Considere T um operador linear compacto e a sua decomposição polar $T = U|T|$, onde $|T| = (T^*T)^{1/2} \geq 0$.

Pelo Teorema 3.2.9, como T é compacto então o operador auto-adjunto $|T|$ também é compacto. Dessa forma, usando o Teorema 3.2.5, sabemos que existe uma sequência ortonormal (v_n) em H de modo que

$$|T|x = \sum_{n=1}^{\infty} \lambda_n \langle x, v_n \rangle v_n, \quad x \in H, \tag{3.16}$$

onde (λ_n) é uma sequência numérica não crescente que converge para zero.

Considere agora a sequência (u_n), onde $u_n = Uv_n$, para cada $n \in \mathbb{N}$. Vejamos que (u_n) é uma sequência ortonormal em H. De fato, lembramos primeiramente do Teorema 3.2.5 que $|T|v_n = \lambda_n v_n$, sendo λ_n um autovalor não nulo de $|T|$. Logo, $v_n \notin \ker U$ pois pelo Teorema 3.2.8 temos que $\ker U = \ker |T|$. Dessa forma, $v_n \in (\ker U)^{\perp}$, para todo $n \in \mathbb{N}$, e como U é uma isometria parcial, segue que a sequência (u_n) é ortonormal pois (v_n) é ortonormal e

$$\langle u_i, u_j \rangle = \langle Uv_i, Uv_j \rangle = \langle v_i, v_j \rangle.$$

Como U é uma isometria parcial, em particular é um operador contínuo, então aplicando U na igualdade (3.16) obtemos

$$Tx = U|T|x = \sum_{n=1}^{\infty} \lambda_n \langle x, v_n \rangle u_n, \quad x \in H. \tag{3.17}$$

Chamaremos a decomposição (3.17) de *Representação de Schmidt* do operador linear compacto T e os elementos da sequência (λ_n) são chamados de *valores singulares* de T e serão denotados por $\sigma_n(T)$.

Para T^*, temos a Representação de Schmidt é dada por:

$$T^*x = \sum_{n} \sigma_n(T) \langle x, u_n \rangle v_n, \quad x \in H.$$

A decomposição de operadores compactos em espaços de Hilbert, estudada ao longo dessa seção, é um importante resultado na teoria de operadores lineares limitados. Como primeira consequência da decomposição estudada anteriormente temos uma nova caracterização para operadores compactos em espaços de Hilbert.

Teorema 3.2.10. *Um operador linear T definido no espaço de Hilbert H é compacto se, e somente se, existe uma sequência de operadores de posto finito T_n tal que $\|T_n - T\| \longrightarrow 0$.*

Demonstração. Primeiramente, assuma que existe uma sequência de operadores de posto finito

T_n tal que $\|T_n - T\| \longrightarrow 0$. Nesse caso, para cada $n \in \mathbb{N}$, o operador T_n é compacto pelo Teorema 3.1.8 e então segue pelo Teorema 3.1.6 que T é compacto.

Reciprocamente, suponha que T é compacto. Portanto para cada x unitário em H temos que

$$Tx = \sum_{n=1}^{\infty} \sigma_n(T)\langle x, v_n\rangle u_n,$$

onde $\sigma_n(T) \longrightarrow 0$ e $(v_n), (u_n)$ são sequências ortonormais. Para cada $k \geq 1$, seja T_k o operador definido em H por

$$T_k x = \sum_{n=1}^{k} \sigma_n(T)\langle x, v_n\rangle u_n, \quad x \in H.$$

Cada T_k é de posto finito e temos que $\|T_k - T\| \longrightarrow 0$ quando $k \longrightarrow \infty$. De fato, para cada vetor unitário $x \in H$ temos,

$$\left\|(T_k - T)x\right\|^2 = \sum_{n=k+1}^{\infty} |\sigma_n(T)|^2 |\langle x, v_n\rangle|^2.$$

Como a sequência λ_n é convergente, em particular é limitada, ou seja, para $n > k$ temos $|\sigma_n(T)|^2 \leq \sup\{|\sigma_n(T)|^2 : n > k\}$. Usando isso e a Desigualdade de Bessel segue que

$$\left\|(T_k - T)x\right\|^2 \leq \sup\{|\sigma_n(T)|^2 : n > k\}.$$

O lado direito da desigualdade acima converge para zero quando $k \longrightarrow \infty$, logo $\|T_k - T\| \longrightarrow 0$.

□

Observação 3.2.11. Uma das implicações do Teorema 3.2.10 não é válida quando consideramos espaços de Banach. Mais precisamente, no artigo [2], o matemático sueco Per H. Enflo provou que, em geral, num espaço de Banach nem sempre é possível aproximar um operador compacto por operadores de posto finito.

3.3 Operadores de Hilbert-Schmidt e a classe traço

Vamos usar a representação dada na seção anterior para caracterizar duas importantes classes de operadores limitados. A primeira delas é formada pelos operadores de Hilbert-Schmidt, enquanto a segunda é chamada de classe traço. Ao longo desta seção iremos denotar por H um espaço de Hilbert separável arbitrário.

Definição 3.3.1. *A classe de Hilbert-Schmidt consiste de todos os operadores lineares limitados $T : H \longrightarrow H$ tais que para algum conjunto ortonormal total (y_n) de H, vale*

$$\sum_{n=1}^{\infty} \|Ty_n\|^2 < \infty.$$

Denotaremos por $\mathscr{HS}(H)$ o espaço dos operadores Hilbert-Schmidt definidos em H.

Exemplo 3.3.2. *Seja (e_n) o conjunto ortonormal completo canônico de ℓ^2. Considere o operador $T : \ell^2 \longrightarrow \ell^2$, dado por*

$$Tx = \sum_{n=1}^{\infty} \frac{1}{n} \langle x, e_n \rangle e_n.$$

Temos que $\|Te_n\|^2 = \frac{1}{n^2}$. Logo,

$$\sum_{n=1}^{\infty} \|Te_n\|^2 = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty,$$

ou seja, T é um operador de Hilbert-Schmidt em ℓ^2.

Teorema 3.3.3. *Seja $T \in B(H)$. Então T é um operador de Hilbert-Schmidt se, e somente se, T^* é um operador de Hilbert-Schmidt. Além disso, se (x_n) é um sequência ortonormal total arbitrária em H, então $\sum_{n=1}^{\infty} \|Tx_n\|^2 < \infty$, sendo que o valor dessa soma independe da escolha da sequência ortonormal (x_n) tomada.*

Demonstração. Sejam (y_n) um conjunto ortonormal completo em H tal que $\sum_{n=1}^{\infty} \|Ty_n\|^2 < \infty$ e (x_n) um conjunto ortonormal completo arbitrário. Pela Identidade de Parseval temos $\|Ty_n\|^2 = \sum_{m=1}^{\infty} |\langle Ty_n, x_m \rangle|^2$. Então,

$$\sum_{n=1}^{\infty} \|Ty_n\|^2 = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} |\langle Ty_n, x_m \rangle|^2 = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |\langle y_n, T^* x_m \rangle|^2 = \sum_{m=1}^{\infty} \left\| T^* x_m \right\|^2,$$

onde a segunda igualdade decorre da definição de T^* e da convergência absoluta da série

$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} |\langle Ty_n, x_m \rangle|^2.$$

Assim, T^* é um operador de Hilbert-Schmidt (basta fazer o mesmo cálculo com T^* no lugar de T e notar que $T^{**} = T$), além disso as somas possuem o mesmo valor independente do conjunto ortonormal total considerado. $\square$

Usando a desigualdade triangular e a propriedade da norma $\|\lambda T x\| = |\lambda| \|Tx\|$, temos que o conjunto dos operadores de Hilbert-Schmidt é um espaço vetorial e ainda podemos definir uma norma, a partir do Teorema 3.3.3, da seguinte forma:

$$\|T\|_{\mathscr{H}\mathscr{S}} = \left(\sum_{n=1}^{\infty} \|Tx_n\|^2 \right)^{1/2},$$

onde (x_n) é um conjunto ortonormal total qualquer. Ainda pelo Teorema 3.3.3, segue que $\|T\|_{\mathscr{H}\mathscr{S}} = \|T^*\|_{\mathscr{H}\mathscr{S}}$.

Teorema 3.3.4. *Seja $T \in B(H)$. São equivalentes:*

1. *$T \in \mathscr{H}\mathscr{S}(H)$.*
2. *T é compacto e $\sum_{n=1}^{\infty} |\sigma_n(T)|^2 < \infty$.*

Demonstração. Sejam $T \in \mathscr{H}\mathscr{S}(H)$ e (x_n) um conjunto ortonormal total em H. Como a série $\sum_{n=1}^{\infty} \|Tx_n\|^2$ é absolutamente convergente, dado $\epsilon > 0$ existe $N \in \mathbb{N}$ tal que

$$\sum_{n=N+1}^{\infty} \|Tx_n\|^2 < \epsilon^2.$$

Sejam $P_N x = \sum_{n=1}^{N} \langle x, x_n \rangle x_n$, a projeção ortogonal de x em $\operatorname{span}\{x_1, x_2, ..., x_N\}$ e $T_N \doteq T P_N$, logo T_N tem posto finito. Então, para $x \in H$ com $\|x\| = 1$, temos

$$\begin{aligned}
\|(T - T_N)x\| &= \left\| T \left(\sum_{n=1}^{\infty} \langle x, x_n \rangle x_n \right) - T \left(\sum_{n=1}^{N} \langle x, x_n \rangle x_n \right) \right\| \\
&= \left\| \sum_{n=1}^{\infty} \langle x, x_n \rangle T x_n - \sum_{n=1}^{N} \langle x, x_n \rangle T x_n \right\| = \left\| \sum_{n=N+1}^{\infty} \langle x, x_n \rangle T x_n \right\| \\
&\leq \sum_{n=N+1}^{\infty} |\langle x, x_n \rangle| \|Tx_n\| \leq \left(\sum_{n=N+1}^{\infty} |\langle x, x_n \rangle|^2 \right)^{1/2} \left(\sum_{n=N+1}^{\infty} \|Tx_n\|^2 \right)^{1/2} < \epsilon,
\end{aligned}$$

onde usamos as Desigualdades de Bessel e Hölder na penúltima desigualdade.

Portanto, $\|T - T_N\| < \epsilon$ donde segue que T é aproximado por operadores de posto finito. Então, pelo Teorema 3.2.10, o operador T é compacto, logo possui uma representação de Schmidt

$Tx = \sum_{j=1}^{\infty} \sigma_j(T)\langle x, u_j\rangle v_j$, onde (u_j) e (v_j) são sequências ortonormais. Em particular,

$$\|Tx_n\|^2 = \langle \sum_{j=1}^{\infty} \sigma_j(T)\langle x_n, u_j\rangle v_j, \sum_{l=1}^{\infty} \sigma_l(T)\langle x_n, u_l\rangle v_l\rangle = \sum_{j=1}^{\infty} \sigma_j(T)^2 |\langle x_n, u_j\rangle|^2.$$

Assim

$$\begin{aligned}\sum_{n=1}^{\infty} \|Tx_n\|^2 &= \sum_{n=1}^{\infty}\sum_{j=1}^{\infty} \sigma_j(T)^2 |\langle x_n, u_j\rangle|^2 = \sum_{j=1}^{\infty}\sum_{n=1}^{\infty} \sigma_j(T)^2 |\langle x_n, u_j\rangle|^2 \\ &= \sum_{j=1}^{\infty} \sigma_j(T)^2 \sum_{n=1}^{\infty} |\langle x_n, u_j\rangle|^2 = \sum_{j=1}^{\infty} \sigma_j(T)^2.\end{aligned}$$

A penúltima igualdade é válida pela convergência absoluta da série e a última igualdade pela Identidade de Parseval.

Por outro lado, seja T compacto e assuma que a soma $\sum_{n=1}^{\infty} |\sigma_n(T)|^2$ é convergente, então T possui uma representação de Schmidt e fazendo cálculos análogos aos anteriores concluímos que

$$\sum_{n=1}^{\infty} \|Tx_n\|^2 = \sum_{n=1}^{\infty} |\sigma_n(T)|^2 < \infty.$$

□

Definição 3.3.5. *A classe de traço consiste dos operadores lineares limitados $T : H \longrightarrow H$ que possuem uma representação, ou seja, podem ser escritos da forma*

$$Tx = \sum_{n=1}^{\infty} \langle x, y_n\rangle x_n,$$

onde $x_n, y_n \in H$ para todo $n \in \mathbb{N}$ e $\sum_{n=1}^{\infty} \|x_n\|\|y_n\| < \infty$. Escrevemos $T \in \mathscr{T}\mathscr{C}(H)$.

Considere

$$v(T) = \inf \left\{ \sum_{n=1}^{\infty} \|x_n\|\|y_n\|, \text{ onde } \sum_{n=1}^{\infty} \langle x, y_n\rangle x_n \text{ é uma representação para } T \right\}.$$

A aplicação $v : \mathscr{T}\mathscr{C}(H) \longrightarrow \mathbb{R}$ define uma norma em $\mathscr{T}\mathscr{C}(H)$.

Teorema 3.3.6. *Seja $T \in B(H)$. São equivalentes:*

1. *T é um operador traço.*

2. *T é compacto e $\sum_{n=1}^{\infty} \sigma_n(T) < \infty$.*

Demonstração. Se $T \in \mathscr{TC}(H)$, então T admite uma representação, portanto é limite de operadores de posto finito. Assim, T é compacto pelo Teorema 3.2.10 e dessa forma possui uma representação de Schmidt,

$$Tx = \sum_{n=1}^{\infty} \sigma_n(T)\langle x, x_n\rangle y_n, \text{ onde } (x_n), (y_n) \text{ são sequências ortonormais.}$$

Pela definição de $v(T)$, dado $\epsilon > 0$ existe uma representação de T, digamos $Tx = \sum_{n=1}^{\infty}\langle x, w_n\rangle z_n$, tal que $\sum_{n=1}^{\infty}\|w_n\|\|z_n\| < v(T)(1+\epsilon)$. Temos dessa forma

$$\sigma_n(T) = \langle Tx_n, y_n\rangle = \left\langle \sum_{m=1}^{\infty}\langle x_n, w_m\rangle z_m, y_n \right\rangle = \sum_{m=1}^{\infty}\langle x_n, w_m\rangle\langle z_m, y_n\rangle \in \mathbb{R}.$$

Portanto,

$$\begin{aligned}\sum_{n=1}^{\infty}\sigma_n(T) &= \sum_{m=1}^{\infty}\sum_{m=1}^{\infty}\langle x_n, w_m\rangle\langle z_m, y_n\rangle \leq \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}|\langle x_n, w_m\rangle||\langle z_m, y_n\rangle| \\ &\leq \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}|\langle x_n, w_m\rangle||\langle z_m, y_n\rangle| \leq \sum_{m=1}^{\infty}\left(\sum_{n=1}^{\infty}|\langle x_n, w_m\rangle|^2\right)^{1/2}\left(\sum_{n=1}^{\infty}|\langle z_m, y_n\rangle|^2\right)^{1/2} \\ &\leq \sum_{m=1}^{\infty}\|w_m\|\|z_m\| < v(T)(1+\epsilon).\end{aligned}$$

No penúltimo passo usamos a Desigualdade de Bessel. Com isso, mostramos a desigualdade desejada.

Reciprocamente, se T é compacto e $\sum_{n=1}^{\infty}\sigma_n(T) < \infty$, temos que T tem uma representação de Schmidt, digamos

$$Tx = \sum_{n=1}^{\infty}\sigma_n(T)\langle x, z_n\rangle x_n, \text{ onde } \|x_n\| = \|z_n\| = 1, \ n \in \mathbb{N}.$$

Assim,

$$Tx = \sum_{n=1}^{\infty}\langle x, y_n\rangle x_n, \text{ onde } y_n = \sigma_n(T) z_n. \tag{3.18}$$

Além disso, pelos cálculos anteriores segue que $\sum_{n=1}^{\infty}\|x_n\|\|y_n\| = \sum_{n=1}^{\infty}\sigma_n(T) < \infty$. Logo $T \in \mathscr{TC}(H)$.

$\square$

Por fim, gostaríamos de destacar que as caracterizações dadas nos Teoremas 3.3.4 e 3.3.6 para os operadores de Hilbert-Schmidt e na Classe Traço mostram que essas duas classes de operadores lineares limitados são exemplos particulares de operadores nas chamadas *classes de Schatten*. Mais precisamente, dado $p \geq 1$ temos a *$p-$classe de Schatten* definida por

$$S_p = \left\{ T \in K(H);\ \left(\sum_{n=1}^{\infty} |\sigma_n(T)|^p \right)^{1/p} < \infty \right\}.$$

O espaço S_p é completo com relação a norma

$$\|T\|_p = \left(\sum_{n=1}^{\infty} |\sigma_n(T)|^p \right)^{1/p}.$$

Como consequência dos Teoremas 3.3.4 e 3.3.6, no caso $p = 1$ temos a classe traço e no caso $p = 2$ temos os operadores de Hilbert-Schmidt.

Considerações Finais

O desenvolvimento deste trabalho proporcionou criar um material que pode ser usado como base para estudos iniciais em Análise Funcional e também a uma introdução à teoria de operadores. Devido a grande abrangência de conteúdos, alguns tópicos não foram abordados com maiores detalhes, como por exemplo o Teorema de Hahn-Banach e alguns dos clássicos teoremas da Análise Funcional, que podem ser facilmente encontrado nos livros [1] e [4].

Este trabalho teve como um "princípio" a seleção e o desenvolvimento dos conteúdos sem exigir muitos pré-requisitos, então com conhecimentos básicos em Álgebra Linear e Análise deve ser possível acompanhar os resultados propostos neste texto sem grandes complicações. Uma continuação natural desse trabalho é o estudo mais detalhado das classes de Schatten, as quais estendem as classes traço e os operadores de Hilbert-Schmidt.

Referências Bibliográficas

[1] BREZIS, H. **Functional Analysis, Sobolev Spaces and Partial Differential Equations**. Springer, 2011.

[2] ENFLO, P. A. A counterexample to the approximation problem in Banach spaces. **Acta Math**, v. 130, p. 309-317, 1973. Disponível em: https://projecteuclid.org/download/pdf_1/euclid.acta/1485889774.

[3] HÖNIG, C.S. **Aplicações da Topologia à Análise**. Rio de Janeiro: IMPA, 1976.

[4] KREYSZIG, E. **Introductory Functional Analysis with Applications.** Nova Iorque: John Wiley and Sons, 1989.

[5] LIMA, E.L. **Análise real. Vol 1.** Rio de Janeiro: IMPA, 14° ed., 2017.

[6] LIMA, E.L. **Análise real. Vol 2.** Rio de Janeiro: IMPA, 6° ed., 2016.

[7] LIMA, E.L. **Espaços Métricos.** Rio de Janeiro: IMPA, 1993.

[8] OLIVEIRA, C. R. **Introdução à Análise Funcional**. Rio de Janeiro, IMPA: SBM, 2015.

[9] TORRES, E. R. **Teoremas de Traço e os Operadores Irregulares**. 2011. 98f. Dissertação (Mestrado em Matemática) - Instituto de Matemática, Universidade Federal do Rio de Janeiro, Rio de Janeiro (RJ), 2011. Disponível em: http://www.pg.im.ufrj.br/teses/Matematica/Mestrado/309.pdf .

[10] ZELASKO, W. A short history of polish mathematics. **Mathematical Institute, Polish Academy of Sciences**, Warsaw, 2004. Disponível em: https://www.impan.pl/Sixty/polmat.pdf.

[11] ZHU, K. **Operator Theory in Function Spaces.** Nova Iorque: Marcel Dekker, INC, 1990.