Revista Eletrônica Paulista de Matemática

ISSN 2316-9664
Volume 18, jul. 2020
Iniciação Científica

Teoria da aproximação em espaços de Banach e de Hilbert

Theory of approximation in Banach and Hilbert spaces

Carlos Alberto do Nascimento
CEFAPRO-Centro de Formação e Atualização dos Profisionais da Educação
prof.math.edu@gmail

Marta Cilene Gadotti
UNESP-Universidade Estadual Paulista Júlio de Mesquita Filho
mc.gadotti@unesp.br

Resumo

O objetivo deste trabalho é revisitar resultados clássicos da Análise Funcional sobre Teoria da Aproximação em espaços de Banach e de Hilbert, além de apresentar alguns exemplos importantes.
Palavras-chave: Espaço Convexo. Condição de Haar. Polinômio de Chebyshev.

Abstract

The aim of this work is to revisit classic results of Functional Analysis on Approximation Theory in Banach and Hilbert spaces, besides presenting some important examples.
Keywords: Convex space. Haar condition. Chebyshev polinomial.

Artigo recebido em abr. 2020 e aceito em jul. 2020

1 Introdução

A teoria da aproximação trata de fornecer métodos para a resolução do problema de aproximar uma determinada função por funções em uma classe especial. Seu objetivo central é estabelecer um controle sobre o erro de aproximação. Embora existam cálculos aproximados muito antigos, por exemplo a aproximação do valor de π por Arquimedes, a teoria da aproximação é um ramo relativamente novo da Matemática, pois exige uma noção precisa de função, que só apareceu no final do século XVIII.

A referência [1] mostra como se deu a evolução histórica dos métodos e resultados da teoria da aproximação, a partir do trabalho de Euler em 1777 sobre a mi nimização de erros de distância nos mapas da Rússia e de Laplace, em 1843, sobre a descoberta do melhor elipsóide para a Terra e terminando com o trabalho de Bernstein. Vale ressaltar que o personagem central do livro é Chebyshev (1821-1894).

Muitos trabalhos que envolvem esse tema tratam de aproximação de funções por funções trigonométricas, pelo uso das Séries de Fourier e há poucas referências em português sobre a aproximação por funções polinomiais. Assim, neste trabalho, iremos revisitar resultados sobre existência e unicidade de uma melhor aproximação em espaços de Banach e de Hilbert e apresentar exemplos interessantes de melhor aproximação, incluindo os polinômios de Chebyshev. Na seção 2, apresentamos as provas de resultados conhecidos envolvendo espaços de Banach convexos e estritamente convexos. Na seção 3 discutimos os resultados em espaços de Hilbert e, finalmente, na seção 4 apresentamos um exemplo importante utilizado até os dias atuais, veja por exemplo [2].

2 Aproximação em espaço de Banach

Em muitos exemplos, podemos verificar que um espaço vetorial X pode também ser um espaço métrico, desde que consigamos definir uma norma ou um produto interno sobre ele. Contudo, se não há relação entre a estrutura algébrica de espaço vetorial e a topológica, não podemos avançar muito em termos de aplicações desta teoria, a qual combina conceitos algébricos e métricos. Para garantir tal relação entre essas propriedades sobre X, definimos norma, indicado por $\|x\|$, para todo $x \in X$. Esta seção é baseada nas referências [3] e [4].

Definição 1 *Seja X um espaço vetorial sobre o corpo $\mathbb{K}$, real ou complexo. Uma norma em X é uma aplicação $\|\cdot\| : X \longrightarrow [0, \infty)$ satisfazendo as seguintes propriedades:*

N1) $\|x\| > 0$ *se* $x \neq 0$. *E vale* $\|x\| = 0 \Leftrightarrow x = 0$;

N2) $\|\alpha x\| = |\alpha|\|x\|$, $\alpha \in \mathbb{K}$ *e* $x \in X$;

N3) $\|x + y\| \leq \|x\| + \|y\|$, $x, y \in X$.

Um espaço X munido de uma norma $\|\cdot\|$ é denominado espaço normado e denotado pelo par $(X, \|\cdot\|)$.

Definição 2 *Seja $X = (X, \|\cdot\|)$ um espaço normado. A distância δ de um elemento $x \in X$ a um subconjunto não-vazio $Y \subset X$ é definida por*

$$\delta = \inf_{y_0 \in Y} \|x - y_0\|.$$

Seja $X = (X, \|\cdot\|)$ um espaço normado e suponha que qualquer $x \in X$ pode ser aproximado por um $y \in Y$, onde Y é um subespaço de X. Seja δ a distância de x a Y. Assim, se existir um $y_0 \in Y$ tal que

$$\|x - y_0\| = \delta$$

então y_0 é chamado de uma *melhor aproximação* de x em Y. Uma melhor aproximação é um elemento de Y com distância mínima de x.

Teorema 3 (Existência da melhor aproximação) *Se Y é um subespaço de dimensão finita de um espaço normado $X = (X, \|\cdot\|)$, então para cada $x \in X$, existe uma melhor aproximação de x em Y.*

Demonstração. Seja $x \in X$ dado, $x \neq 0$. Considere a bola fechada

$$\widetilde{B} = \{y \in Y : \|y\| \leq 2\|x\|\}.$$

Como $\widetilde{B} \subset Y$ então $\delta(x, \widetilde{B}) \geq \delta(x, Y)$.

Note que $0 \in \widetilde{B}$, logo a distância de x a $\widetilde{B}$ satisfaz

$$\delta(x, \widetilde{B}) = \inf_{\widetilde{y} \in \widetilde{B}} \|x - \widetilde{y}\| \leq \|x - 0\| = \|x\|. \tag{1}$$

Agora, $x, y \in Y$ e $y \notin \widetilde{B}$, então $\|y\| > 2\|x\|$ e

$$\|x - y\| \geq \|y\| - \|x\| > 2\|x\| - \|x\| = \|x\| \geq \delta(x, \widetilde{B}). \tag{2}$$

Portanto, $\delta(x, \widetilde{B}) = \delta(x, Y)$. E este valor não pode ser assumido por um $y \in Y - \widetilde{B}$. Assim, se houver uma melhor aproximação, ela deve estar em $\widetilde{B}$. Por isso o motivo do uso de $\widetilde{B}$ em vez de todo subespaço Y.

Como $\widetilde{B}$ é compacto (fechado e limitado), Y é de dimensão finita e a função norma é contínua, segue que existe um $y_0 \in \widetilde{B}$ tal que $\|x - y\|$ assume um mínimo $y = y_0$, portanto, y_0 é uma melhor aproximação para x de Y. $\square$

O fato da dimensão de Y ser finita no Teorema 3 é essencial, pois na prova é utilizada a importante caracterização de espaços normados de dimensão finita através da compacidade dada por: um espaço normado possui dimensão finita se, e somente se, a bola fechada é compacta. Vejamos um exemplo em que o espaço é de dimensão infinita.

Seja Y o conjunto de todos os polinômios definidos no intervalo $[0, 1/2]$, de qualquer grau, considerado como um subespaço de $C([0, 1/2], \mathbb{R})$, onde $X = C([0, 1/2], \mathbb{R}) = \{f : [0, 1/2] \to \mathbb{R}$: f é contínua $\}$, ou seja, Y é um subespaço de X de dimensão infinita.

Dado

$$x(t) = (1 - t)^{-1} \in X,$$

mostremos que não existe em Y uma melhor aproximação de x. Note que, dado $\varepsilon > 0$, existe um $N \in \mathbb{N}$ tal que, definindo

$$y_n(t) = 1 + t + \cdots + t^n,$$

temos

$$\begin{aligned} \|x - y_n\| &= \left\|(1 - t)^{-1} - (1 + \cdots + t^n)\right\| \\ &= \left\|\frac{1 - (1 - t)(1 + t \cdots + t^n)}{1 - t}\right\| \end{aligned}$$

$$
\begin{aligned}
&= \left\| \frac{1 - [(1 + t + \cdots + t^n) - (t + t^2 + \cdots + t^{n+1})]}{1 - t} \right\| \\
&= \left\| \frac{1 - (1 - t^{n+1})}{1 - t} \right\| \\
&= \left\| \frac{t^{n+1}}{1 - t} \right\| \\
&< \epsilon, \forall n > N. \qquad (3)
\end{aligned}
$$

O que implica $\delta(x, Y) = 0$. Assim, vemos que não existe um $y_0 \in Y$ que seja polinômio de grau menor ou igual n, satisfazendo

$$\delta = \delta(x, Y) = \|x - y_0\| = 0.$$

Definição 4 *Um subconjunto M de X é convexo se para todos $x, y \in M$ o segmento definido por*

$$\gamma(\lambda) = (1 - \lambda)x + \lambda y,$$

$\lambda \in [0, 1]$*, está contido em M.*

O conjunto das melhores aproximações é convexo. Esse resultado será provado a seguir.

Lema 5 *Em um espaço normado $(X, \|\cdot\|)$ o conjunto M das melhores aproximações de um ponto x para um subespaço Y de X é convexo.*

Demonstração. Tomemos δ como a distância entre x e Y. A conclusão é válida se M for vazio ou se tiver apenas um ponto.

Supomos que M tenha mais de um ponto. Então para $y, z \in M$ temos, por definição

$$\|x - y\| = \|x - z\| = \delta.$$

Mostremos que isto implica em $w = \alpha y + (1 - \alpha)z \in M$, se $0 \leq \alpha \leq 1$. Com efeito, $\|x - w\| \geq \delta$, já que $w \in Y$. E também vale $\|x - w\| \leq \delta$, pois

$$
\begin{aligned}
\|x - w\| &= \|x - w + \alpha x - \alpha x\| \\
&= \|x - \alpha y - (1 - \alpha)z + \alpha x - \alpha x\| \\
&= \|x - \alpha y - z + \alpha z + \alpha x - \alpha x\| \\
&= \|\alpha(x - y) + (1 - \alpha)(x - z)\| \\
&\leq \alpha\|x - y\| + (1 - \alpha)\|x - z\| \\
&= \alpha\delta + (1 - \alpha)\delta \\
&= \alpha\delta + \delta - \alpha\delta \\
&= \delta, \qquad (4)
\end{aligned}
$$

usamos $\alpha \geq 0$ bem como $(1-\alpha) \geq 0$. Logo, $\|x-w\| = \delta$. Portanto, $w \in M$, o que mostra que M é convexo. □

Definição 6 *Uma norma estritamente convexa satisfaz para todos x, y de norma 1, $x \neq y$,*

$$\|x + y\| < 2.$$

Um espaço normado com tal norma é chamado de espaço estritamente convexo.

Exemplo 7 *O corpo de escalares $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ visto como um espaço normado sobre $\mathbb{K}$ é estritamente convexo. Mais geralmente, todo espaço normado cuja a dimensão é 0 ou 1 é estritamente convexo.*

Exemplo 8 *Sejam $x_1 = (1, 1, 0, 0, \ldots)$ e $x_2 = (1, -1, 0, 0, \ldots)$ elementos de c_0, espaço das sequências de números reais que convergem para 0. Temos que*

$$\|x_1\|_\infty = \sup_{j\in\mathbb{N}} |x_j| = 1 = \|x_2\|_\infty = \left\|\frac{x_1 + x_2}{2}\right\|_\infty .$$

Logo, o espaço c_0 não é estritamente convexo.

Observe que para $\|x\| = \|y\| = 1$ a desigualdade triangular fornece $\|x + y\| \leq 2$ e a convexidade estrita exclui o sinal de igualdade, exceto quando $x = y$.

Proposição 9 *Seja X um espaço normado. X é estritamente convexo se, somente se, ocorrer $\|x + y\| = \|x\| + \|y\|$, quando um dos vetores é múltiplo real não-negativo do outro (neste caso l.d.).*

Demonstração. Suponha que $x, y \in X$ sejam tais que $\|x + y\| = \|x\| + \|y\|$ e que X seja estritamente convexo. Sem perda de generalidade, podemos supor $1 = \|x\| \leq \|y\|$. Defina então o vetor $z = \dfrac{y}{\|y\|}$. Assim,

$$\begin{aligned}
\|x + z\| &= \left\|x + \frac{y}{\|y\|}\right\| \\
&= \left\|x - y + y + \frac{y}{\|y\|}\right\| \\
&= \left\|x + y - y\left(1 - \frac{1}{\|y\|}\right)\right\| \\
&\geq \|x + y\| - \|y\|\left(1 - \frac{1}{\|y\|}\right) \\
&= \|x\| + \|y\| - \|y\| - 1 \\
&= 2.
\end{aligned} \tag{5}$$

Por outro lado,

$$\begin{aligned}
\|x + z\| &\leq \|x\| + \|z\| \\
&= 1 + \|z\| \\
&\leq 1 + 1 \\
&= 2.
\end{aligned} \tag{6}$$

Segue destas duas desigualdades que $\left\|\frac{x+z}{2}\right\| = 1$. Pelo fato de X ser estritamente convexo segue que $x = z$, pois caso contrário deveríamos ter $\|x + z\| < 2$. Portanto, $x = \dfrac{y}{\|y\|}$.

A recíproca é imediata, já que se $x, y \in X$ com $\|x\| = 1 = \|y\|$ e nenhum é múltiplo real do outro, tem-se

$$\|x + y\| < \|x\| + \|y\| = 2.$$

□

Teorema 10 (Unicidade de melhor aproximação) *Em um espaço normado estritamente convexo X, existe no máximo uma melhor aproximação de um $x \in X$ em um subespaço, $x \notin Y$.*

Demonstração. Sejam $y \neq \widetilde{y} \in Y$, dois elementos de melhores aproximação de $x \in X$. Então,

$$\begin{aligned}\left\|x - \frac{y+\widetilde{y}}{2}\right\| &= \left\|\frac{1}{2}(x-y) + \frac{1}{2}(x-\widetilde{y})\right\| \\ &\leq \frac{1}{2}\|x-y\| + \frac{1}{2}\|x-\widetilde{y}\| \\ &= \frac{1}{2}\delta(x,Y) + \frac{1}{2}\delta(x,Y) \\ &= \delta(x,Y), \end{aligned} \tag{7}$$

implica que o ponto médio é elemento de melhor aproximação. Consequentemente, pela proposição anterior, existe $\alpha > 0$ tal que $x - y = \alpha(x - \widetilde{y})$.

Se $\alpha = 1$, então $y = \widetilde{y}$, que é uma contradição.

Se $\alpha \neq 1$, então

$$x - y = \alpha(x - \widetilde{y})$$

$$x - y = \alpha x - \alpha\widetilde{y}$$

$$x - \alpha x = y - \alpha\widetilde{y}$$

$$x(1-\alpha) = y - \alpha\widetilde{y}$$

$$x = \frac{y - \alpha\widetilde{y}}{1-\alpha} \quad \in Y,$$

o que mostra que $x \in Y$, o que é uma contradição, o que completa a demonstração. □

Porém, existem espaços que não são estritamente convexos e assim não é possível utilizar o Teorema 10.

Exemplo 11 *O espaço $X = C([a,b],\mathbb{R})$, não é estritamente convexo com a norma do máximo.*

Consideremos x_1 e x_2 definidos por

$$x_1(t) = 1, \qquad x_2(t) = \frac{t-a}{b-a}$$

onde $t \in [a,b]$. Claramente, $x_1, x_2 \in X$ e $x_1 \neq x_2$.

Temos também que $\|x_1\| = \|x_2\| = 1$ e $\|x_1 + x_2\| = \max_{t\in J}\left|1 + \dfrac{t-a}{b-a}\right| = 2$, onde $I = [a,b]$. O que mostra que X, não é estritamente convexo.

Uma vez que o espaço X não é estritamente convexo o problema da unicidade requer mais atenção. Por isso, o seguinte conceito será de grande utilidade.

Definição 12 *Um ponto extremo de um x em $X = C([a,b],\mathbb{R})$ é um $t_0 \in [a,b]$ tal que $|x(t_0)| = \|x\|$.*

Exemplo 13 *Seja $Y = P_2$ o conjunto dos polinômios de grau menor ou igual a 2 tal que, $Y \subset X = C([-2,2],\mathbb{R})$. A função $y(t) = t^2$ varia de $[0,4]$ para $t \in [-2,2]$, deste modo x possui dois pontos extremos $t = \pm 2$, pois*

$$\|x\| = \max_{t\in[-2,2]} |x(t)| = 4 = |x(2)|$$

e

$$\|x\| = \max_{t\in[-2,2]} |x(t)| = 4 = |x(-2)|.$$

Para garantir a unicidade da melhor aproximação no espaço $C([a,b],\mathbb{R})$, vamos precisar da condição de Haar como segue.

Definição 14 (Condição de Haar) *Um subespaço Y de dimensão finita do espaço $X = C([a,b],\mathbb{R})$ satisfaz a condição de Haar se cada $y \in Y, y \neq 0$, tiver no máximo $n-1$ zeros em $[a,b]$, onde $n = dim\ Y$.*

A condição de Haar é equivalente a condição de que para qualquer base $\{y_1, \ldots\ y_n\}$ de Y e qualquer n-upla de pontos distintos, $t_1, \cdots, t_n$ no intervalo $[a,b]$, o determinante da matriz

$$M = \begin{bmatrix} y_1(t_1) & y_1(t_2) & \cdots & y_1(t_n) \\ y_2(t_1) & y_2(t_2) & \cdots & y_2(t_n) \\ \vdots & \vdots & \ddots & \vdots \\ y_n(t_1) & y_n(t_2) & \cdots & y_n(t_n) \end{bmatrix},$$

é diferente de zero. De fato, todo $y \in Y$, possui uma representação $y = \sum_{k=1}^{n} \alpha_k y_k$. O subespaço Y satisfaz a condiçao de Haar se, e somente se, todo $y = \sum_{k=1}^{n} \alpha_k y_k \in Y$, com n ou mais zeros nos pontos $t_1, t_2, \cdots, t_n, \cdots$ em $[a,b]$ for identicamente nulo. Isso significa que as n condições

$$y(t_j) = \sum_{k=1}^{n} \alpha_k y_k(t_j) = 0, \qquad j = 1, \cdots, n, \tag{8}$$

devem implicar $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. Isto acontece se, e somente se, o determinante de M não é zero.

Exemplo 15 *O conjunto de todos os polinômios de grau menor ou igual a n definidos no intervalo $[a,b] \subset \mathbb{R}$, denotado por $Y = P_n([a,b],\mathbb{R})$, satisfaz a condição de Haar. De fato, analisemos o caso particular $Y_1 = P_2([-3,3],\mathbb{R})$. Seja a base $\{y_1, y_2, y_3\}$ tal que*

$$\{y_1, y_2, y_3\} = \{t^2 - t - 2, t^2 - 1, t^2 + 2t + 1\}.$$

Construindo a matriz M no intervalo $[-3,3]$, tomando $t_1 = -3, t_2 = 2, t_3 = 3$ temos

$$M = \begin{bmatrix} 10 & 0 & 4 \\ 8 & 3 & 8 \\ 4 & 8 & 6 \end{bmatrix}.$$

Como $\det M \neq 0$, *segue que*

$$y(t_j) = \sum_{k=1}^{3} \alpha_k y_k(t_j) = 0, \quad j = 1, 2, 3$$

implica $\alpha_1 = \alpha_2 = \alpha_3 = 0$. De fato, isso pode ser constatado pela definição ou seja,

$$\alpha_1(t^2 - t - 2) + \alpha_2(t^2 - 1) + \alpha_3(t^2 + 2t + 1) = 0,$$

temos

$$\begin{cases} \alpha_1 & + & \alpha_2 & + & \alpha_3 & = & 0 \\ -\alpha_1 & + & 0 & + & 2\alpha_3 & = & 0 \\ -2\alpha_1 & - & \alpha_2 & + & \alpha_3 & = & 0 \end{cases}$$

que implica $\alpha_1 = \alpha_2 = \alpha_3 = 0$.

De forma similar, podemos concluir que $Y = P_n([a,b],\mathbb{R})$ *satisfaz a condição de Haar.*

Lema 16 *Suponha que um subespaço* Y *do espaço* $X = C([a,b],\mathbb{R})$ *satisfaz a condição de Haar. Se para dados* $x \in X$ *e* $y \in Y$ *a função* $x - y$ *tem menos que* $n+1$ *pontos extremos, então* y *não é uma melhor aproximação de* x *para* Y*, onde* $\dim Y = n$.

Demonstração. Por hipótese, a função $v = x - y$ tem $m \leq n$ pontos extremos $t_1, t_2, t_3, \cdots, t_m$. Se $m < n$, escolhemos pontos adicionais t_j em $[a,b]$, até que tenhamos n pontos distintos $t_1, \cdots, t_n$. Usando esses pontos e uma base $\{y_1, \cdots y_n\}$ para Y, consideremos o sistema de equações lineares não homogêneo.

$$\sum_{k=1}^{n} \beta_k y_k(t_j) = v(t_j), \tag{9}$$

$j = 1, \cdots, n$ nas incógnitas $\beta_1, \cdots, \beta_n$. Como Y satisfaz a condição de Haar, segue que (9) possui uma única solução. Usando essa solução, definimos

$$y_0 = \beta_1 y_1 + \cdots + \beta_n y_n$$

e

$$\widetilde{y} = y + \varepsilon y_0, \qquad \varepsilon > 0.$$

Mostremos que quando $\varepsilon \to 0$, a norma da função $\widetilde{v} = x - \widetilde{y}$, satisfaz

$$\|\widetilde{v}\| = \|x - \widetilde{y}\| < \|x - y\| = \|v\| \tag{10}$$

e assim y não pode ser uma melhor aproximação de x para Y. De fato, vamos dividir $[a,b]$ em dois conjuntos N e $K = [a,b] - N$, onde N contém os pontos extremos $t_1, \cdots, t_m$ de v.

Nos pontos extremos, $|v(t_i)| = \|v\|$ e $\|v\| > 0$, desde que $v = x - y \neq 0$ e $y_0(t_i) = v(t_i)$ por (9) e pela definição de y_0.

Pela continuidade de v, para cada t_i, existe uma vizinhança N_i tal que em $N = N_1 \cup \cdots \cup N_m$, temos

$$\mu = \inf_{t \in N} |v(t)| > 0. \tag{11}$$

Pela definição de ínfimo e levando em conta que $y_0(t_i) = v(t_i)$ temos

$$\inf_{t \in N} |y_0(t)| \geq \frac{1}{2}\|v\|. \tag{12}$$

Como $y_0(t) = v(t) \neq 0$ para $t \in N$, temos $\dfrac{y_0(t)}{v(t)} > 0$, por (11) e, utilizando (12), obtemos

$$\frac{y_0(t)}{v(t)} = \frac{|y_0(t)|}{|v(t)|} \geq \inf \frac{|y_0(t)|}{\|v\|} \geq \frac{1}{2}.$$

Seja $M_0 = \sup_{t\in N} |y_0(t)|$. Então para $0 < \varepsilon < \dfrac{\mu}{M_0}$ e todo $t \in N$, obtemos

$$\frac{\varepsilon y_0(t)}{v(t)} = \varepsilon \frac{|y_0(t)|}{|v(t)|} \leq \frac{\varepsilon M_0}{\mu} < 1.$$

Como

$$\widetilde{v} = x - \widetilde{y} = x - y - \varepsilon y_0 = v - \varepsilon y_0,$$

usando essa desigualdades, para todo $t \in N$ e $0 < \varepsilon < \frac{\mu}{M_0}$, temos

$$\begin{aligned} |\widetilde{v}(t)| &= |v(t) - \varepsilon y_0(t)| \\ &= |v(t)| \cdot \left[1 - \frac{\varepsilon y_0(t)}{v(t)}\right] \\ &\leq \|v\| \cdot \left(1 - \frac{\varepsilon}{2}\right) \\ &< \|v\|. \end{aligned} \tag{13}$$

Em K, que é um conjunto fechado, definimos

$$M_1 = \max_{t\in K} |y_0(t)| \quad e \quad M_2 = \max_{t\in K} |v(t)|.$$

Como N contém todos os pontos extremos de v, segue que $M_2 < \|v\|$ e daí $\|v\| = M_2 + \gamma$, onde $\gamma > 0$.

Escolhendo um número positivo $\varepsilon < \frac{\gamma}{M_1}$, temos $\varepsilon M_1 < \gamma$ e obtemos para todo $t \in K$,

$$\begin{aligned} |\widetilde{v}(t)| &\leq |v(t)| + \varepsilon |y_0(t)| \\ &\leq M_2 + \varepsilon M_1 \\ &< \|v\|. \end{aligned} \tag{14}$$

Escolhendo $\varepsilon < \min\{\mu/M_0, \gamma/M_1\}$ e tomando o supremo de $\widetilde{v}$ em K, temos

$$\|\widetilde{v}\| < \|v\|,$$

o que conclui a demonstração. □

Exemplo 17 *Seja $Y = P_2$ subespaço do espaço $X = C([0,2], \mathbb{R})$ e note que Y satisfaz a condição de Haar. Tomemos $x(t) = t^3 \in X$ e $y(t) = t^2 \in P_2$. Logo,*

$$y(t) = x(t) - y(t) = t^3 - t^2.$$

Temos que $\dim P_2 = 3$ e que y varia de $[\frac{-4}{27}, 4]$ em $t \in [0,2]$, assim y possui um ponto extremo $t = 2$ pois,

$$\|y\| = \max_{t\in[0,2]} |y(t_0)| = 4 = |y(2)|.$$

Logo y não é uma melhor aproximação de x para Y. Podemos observar esse fato graficamente como ilustrado nas figuras que seguem. No entanto, seja $Y = P_2 \in X = C([0,2], \mathbb{R})$ que satisfaça a condição de Haar tal que, $y(t) = t^2 \in Y$ e $x(t) = t^2 + \dfrac{1}{4} \in X$. Temos que

$$z(t) = x(t) - y(t) = t^2 + \frac{1}{4} - t^2 = \frac{1}{4}.$$

Logo, a $\dim Y = 3$ e z possui infinitos pontos extemos, como ilustra a Figura 2. Portanto, y é uma melhor aproximação de x para Y.

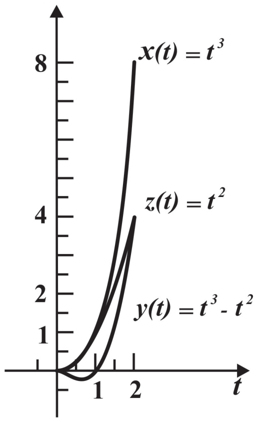

Figura 1: Gráfico das funções $x(t)$, $y(t)$ e $z(t)$.

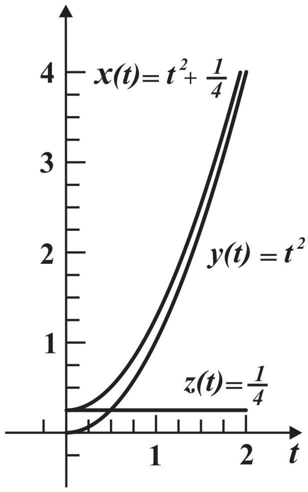

Figura 2: Gráfico das funções $x(t)$, $y(t)$ e $z(t)$.

Utilizando o lema anterior, obtemos o importante teorema a seguir.

Teorema 18 (Unicidade de Haar para melhor aproximação) *Seja Y um subespaço de dimensão finita do espaço X. Então uma melhor aproximação para Y é unica para todo $x \in C[a,b]$ se, e somente se, Y satisfaz a condição de Haar.*

Demonstração. Suponhamos que Y não satisfaz a condição de Haar, ou seja, $\det M = 0$, logo o sistema homogêneo

$$\gamma_1 y_k(t_1) + \cdots + \gamma_n y_k(t_n) = 0, \qquad k = 1, \cdots, n$$

possui solução não trivial $\gamma_1, \cdots, \gamma_n$. Usando essa solução e qualquer $y = \sum_{k=1}^{n} \alpha_k y_k \in Y$, temos

$$\sum_{j=1}^{n} \gamma_j y(t_j) = \sum_{k=1}^{n} \alpha_k \left[\sum_{j=1}^{n} \gamma_j y_k(t_j) \right] = 0. \tag{15}$$

Além disso, o sistema transposto

$$\beta_1 y_1(t_j) + \cdots + \beta_n y_n(t_j) = 0, \qquad j = 1, \cdots, n,$$

também possui uma solução não trivial $\beta_1, \cdots, \beta_n$. Usando esta solução definimos $y_0 = \sum_{k=1}^{n} \beta_k y_k$, o que implica $y_0 \neq 0$ e $y_0 = 0$ em $t_1, \cdots, t_n$.

Seja λ tal que $\|\lambda y_0\| \leq 1$ e seja $z \in X$, tal que $\|z\| = 1$ e

$$z(t_j) = sgn(\gamma_j) = \begin{cases} -1, & \text{se} \quad \gamma_j < 0 \\ 1, & \text{se} \quad \gamma_j \geq 0, \end{cases}$$

onde $sgn(\cdot)$ denota função sinal. Definimos $x \in X$ por

$$x(t) = z(t)(1 - |\lambda y_0(t)|).$$

Então

$$x(t_j) = z(t_j) = sgn(\gamma_j),$$

pois $y_0(t_j) = 0$ e $\|x\| = 1$. Mostremos que esta função x possui infinitas melhores aproximações para Y. Usando $|z(t)| \leq \|z\| = 1$ e $|\lambda y_0(t)| \leq \|\lambda y_0\| \leq 1$, para todo $\varepsilon \in [-1, 1]$, obtemos

$$\begin{aligned} |x(t) - \varepsilon\lambda y_0(t)| &\leq |x(t)| + |\varepsilon\lambda y_0(t)| \\ &= |z(t)(1 - |\lambda y_0(t)|)| + |\varepsilon\lambda y_0(t)| \\ &= |z(t)||1 - |\lambda y_0(t)|| + |\varepsilon\lambda y_0(t)| \\ &\leq 1 - |\lambda y_0(t)| + |\varepsilon\lambda y_0(t)| \\ &= 1 - |\lambda y_0(t)| + |\varepsilon||\lambda y_0(t)| \\ &= 1 - (1 - |\varepsilon|)|\lambda y_0(t)| \\ &\leq 1. \end{aligned}$$

Portanto, todo $\varepsilon\lambda y_0$, $-1 \leq \varepsilon \leq 1$, é uma melhor aproximação para x, desde que ocorra

$$\|x - y\| \geq 1, \qquad \forall y \in Y. \tag{16}$$

Vamos mostrar (16) para um arbitrário $y = \sum_{k=1}^{n} \alpha_k y_k \in Y$. Suponha que $\|x - \widetilde{y}\| < 1$, para um $\widetilde{y} \in Y$. Então as condições

$$\begin{cases} x(t_j) = sgn(\gamma_j) = \pm 1 \\ |x(t_j) - \widetilde{y}(t_j)| \leq \|x - \widetilde{y}\| < 1, \end{cases}$$

implicam que para todo $\gamma_j \neq 0$,

$$sgn(\widetilde{y}(t_j)) = sgn(x(t_j)) = sgn(\gamma_j).$$

De fato, se $x(t_j) = 1$, para todo j, assuma que

$$sgn(\widetilde{y}(t_j)) \neq sgn(x(t_j)).$$

Assim $sgn(\widetilde{y}(t_j)) = -1$. Segue que

$$\begin{aligned} |x(t_j) - \widetilde{y}(t_j)| &= |1 - \widetilde{y}(t_j)| \\ &= |1 - (-1)| \\ &> 1. \end{aligned}$$

Logo,
$$sgn(\widetilde{y}(t_j)) = 1 = sgn(x(t_j)).$$
De maneira análoga, se $x(t_j) = -1$, para todo j, assuma que
$$sgn(\widetilde{y}(t_j)) \neq sgn(x(t_j)).$$
Assim, $sgn(\widetilde{y}(t_j)) = 1$. Segue que
$$\begin{array}{rcl} |x(t_j) - \widetilde{y}(t_j)| & = & |-1-\widetilde{y}(t_j)| \\ & = & |-1-1| \\ & > & 1. \end{array}$$
Logo,
$$sgn(\widetilde{y}(t_j)) = -1 = sgn(x(t_j)).$$
Mas isso contradiz (15) com $y = \widetilde{y}$ pois $\gamma_j \neq 0$, para algum j, assim
$$\sum_{j=1}^{n} \gamma_j \widetilde{y}(t_j) = \sum_{j=1}^{n} \gamma_j sgn(\gamma_j) = \sum_{j=1}^{n} |\gamma_j| \neq 0, \tag{17}$$
Portanto, (16) acontece.

Reciprocamente, suponha que Y satisfaz a condição de Haar, mas ambos $y_1 \in Y$ e $y_2 \in Y$ são melhores aproximações para o mesmo fixo $x \in X$. Então, definindo
$$v_1 = x - y_1 \qquad e \qquad v_2 = x - y_2,$$
temos $\|v_1\| = \|v_2\| = \delta$, onde δ é a distância de x a Y, como anteriormente. O Lema 5, implica que $y = \dfrac{1}{2}(y_1 + y_2)$ é também uma melhor aproximação para x. Pelo Lema 16 a função
$$v = x - y = x - \frac{1}{2}(y_1 + y_2) = \frac{1}{2}(v_1 + v_2), \tag{18}$$
tem pelo menos $n+1$ pontos extremos $t_1, \cdots, t_{n+1}$. Em tais pontos temos
$$|v(t_j)| = \|v\| = \delta.$$
De (18) obtemos
$$2v(t_j) = v_1(t_j) + v_2(t_j) = 2\delta \qquad ou \qquad -2\delta.$$
Agora, $|v_1(t)| \leq \|v_1\| = \delta$ e também para v_2. Logo, ambos os termos devem ter o mesmo sinal e o valor absoluto máximo possível, isto é
$$v_1(t_j) = v_2(t_j) = \delta \qquad ou \qquad -\delta,$$
onde $j = 1, \cdots, n+1$. Isso implica que $y_1 - y_2 = v_2 - v_1$ tem $n+1$ zeros em $[a, b]$. Logo, $y_1 - y_2 = 0$, pela condição de Haar. Concluímos que $y_1 = y_2$, o que completa a demonstração. □

3 Aproximação em espaço de Hilbert

Um produto interno sobre um espaço vetorial de dimensão finita é uma generalização do produto escalar e, em termos de tal produto interno, pode-se também definir "comprimento" e "ângulo". Espaços reais com produto interno de dimensão finita são considerados como espaços normados e frequentemente denominados espaços euclidianos, pois defini-se a norma a partir do produto interno.

Definição 19 *Um espaço pré-hilbertiano é um espaço vetorial com produto interno. Um espaço de Hilbert é um espaço com produto interno completo com respeito à norma* $\|x\| = \sqrt{\langle x, x\rangle}$*, onde* $< \cdot >$ *indica o produto interno.*

Lema 20 *Todo espaço de Hilbert é estritamente convexo.*

Demonstração. Para todos x e $y \neq x$, distintos e de norma 1, temos $\|x - y\| = \alpha$, onde $\alpha > 0$, e a igualdade do paralelogramo nos mostra que

$$\begin{aligned} \|x+y\|^2 &= -\|x-y\|^2 + 2(\|x\|^2 + \|y\|^2) \\ &= -\alpha^2 + 2(1+1) < 4. \end{aligned} \tag{19}$$

Logo, $\|x + y\| < 2$. □

Teorema 21 *Para cada x dado em um espaço de Hilbert H e cada subespaço fechado Y de H, existe uma única melhor aproximação para x de Y.*

Segue do Teorema 10 e do Lema 20. □

Para aprofundar o conhecimento sobre melhor aproximação em espaço de Hilbert, o leitor pode consultar [3].

4 Aplicação sobre aproximação em espaços normados

Nesta seção vamos discutir uma aplicação sobre aproximação de funções em espaço normado, onde deduziremos os conhecidos *Polinômios de Chebyschev*. Apresentaremos inicialmente alguns resultados necessários para a compreensão da aplicação. Vamos considerar o espaço $X = C([a, b], \mathbb{R})$ com os escalares em $\mathbb{R}$. Iniciamos com a seguinte definição:

Definição 22 *Sejam* $X = C([a, b], \mathbb{R})$*,* $x \in X$ *e* $y \in Y$*, onde* Y *é um subespaço de* X*. Um conjunto de pontos* $t_0, \cdots, t_k$ *em* $[a, b]$*, onde* $t_0 < t_1 < \cdots < t_k$*, é chamado de conjunto alternante para* $x - y$ *se* $x(t_j) - y(t_j)$ *tem alternadamente os valores* $\|x - y\|$ *e* $-\|x - y\|$ *em consecutivos pontos* t_j*.*

Lema 23 *Seja* Y *um subespaço do espaço* $X = C([a, b], \mathbb{R})$*, que satisfaça a condição de Haar. Dado* $x \in X$ *seja* $y \in Y$*, tal que para* $x - y$*, existe um conjunto alternante de* $n + 1$ *pontos, onde dim* $Y = n$*. Então* y *é a melhor aproximação uniforme de* x *em* Y*.*

Demonstração. Pelo Teorema 3 e pelo Lema 18, existe uma única melhor aproximação de x em Y. Se esta solução não é y, ela é algum outro $y_0 \in Y$, então

$$\|x - y\| > \|x - y_0\|.$$

Esta desigualdade implica que, nesses $n+1$ pontos extremos, a função

$$y_0 - y = (x - y) - (x - y_0),$$

tem o mesmo sinal que $x - y$, pois $(x - y)$ é igual à $\pm\|x - y\|$ em tal ponto, enquanto que, $x - y_0$, nunca pode exceder $\|x - y_0\|$ em valor absoluto, o que é estritamente menor que $\|x - y\|$. Isto mostra que $y_0 - y$ está alternando entre positivo e negativo naqueles $n+1$ pontos, o que implica em n raízes em $[a, b]$, o que é impossível, a menos que $y_0 - y = 0$ pois $y_0 - y \in Y$ e Y satisfaz a condição de Haar. Assim, y deve ser a melhor melhor aproximação de x para Y. □

Um problema clássico muito interessante, o qual é uma aplicação deste último resultado, é a aproximação de $x \in C([-1, 1], \mathbb{R})$, definida por

$$x(t) = t^n, \tag{20}$$

$n \in \mathbb{N}$ fixo, para um subespaço gerado $Y = [y_0, \cdots, y_{n-1}]$, pelas funções

$$y_j(t) = t^j, j = 0, \cdots, n-1. \tag{21}$$

Queremos aproximar x por um polinômio y de grau menor que n, onde y tem a forma

$$y(t) = \alpha_{n-1}t^{n-1} + \cdots + \alpha_0.$$

Assim, para $z = x - y$, $y \in Y$, temos

$$z(t) = t^n - [(\alpha_{n-1})t^{n-1} + (\alpha_{n-2})t^{n-2} + \cdots + \alpha_0].$$

Desejamos determinar y tal que $\|z\|$ seja a menor possível. Ou seja, queremos encontrar o polinômio z entre todos os polinômios de grau n, com coeficiente igual a 1, que tenha o menor desvio máximo em $[-1, 1]$.

A resolução deste problema utiliza as funções chamadas de *Polinômios de Chebyshev*, do tipo 1 de ordem n.

Seja a função

$$t = \cos n\theta.$$

No intervalo $[0, \pi]$ a função definida por $x(n) = \cos n\theta$ possui $n+1$ pontos extremos, sendo os

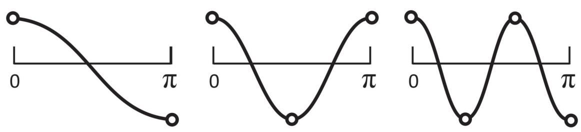

Figura 3: Gráfico de $t = \cos n\theta$, para $n = 1, 2, 3$. Fonte: Adaptado de Kreyszig (1978, p. 347).

valores alternados. Pelo Lema 23, podemos observar que $x(t) = \cos n\theta$, poderá ajudar a resolver o problema, se pudermos escrever $\cos n\theta$ como polinômio em $t = \cos\theta$, com $\theta \in [0, \pi]$.

Pelas identidades trigonométricas, sabemos que:

$$\cos 0\theta = \cos 0 = 1$$

$$\cos 1\theta = \cos\theta$$

$$\cos 2\theta = 2\cos^2\theta - 1$$
$$\cos 3\theta = 4\cos^3\theta - 3\cos\theta$$
$$\cos 4\theta = 8\cos^4\theta - 8\cos^2\theta + 1$$
$$\vdots$$

De forma geral, é possível escrever $\cos(n\theta)$ como um polinômio em $x = \cos\theta$:

$$\cos(n\theta) = \frac{1}{2^{n-1}}\cos^n\theta + \sum_{j=0}^{n-1}\beta_{n,j}\cos^j\theta, \qquad n = 1, 2, \ldots$$

onde $\beta_{n,j}$ são constantes. De fato, procedendo por indução finita, sabemos que a igualdade anterior é válida para $n = 1$. Suponhamos agora que seja válida para todo n e provemos que vale para $n+1$. Usando as identidades e os fatos de que a função seno é ímpar e a cosseno par, temos

$$\cos(n+1)\theta = \cos(n\theta + \theta) = \cos(n\theta)\cos\theta - sen(n\theta)sen\theta,$$
$$\cos(n-1)\theta = \cos(n\theta - \theta) = \cos(n\theta)\cos\theta + sen(n\theta)sen\theta$$

e somando ambos os lados, obtemos

$$\cos(n+1)\theta + \cos(n-1)\theta = 2\cos(n\theta)\cos\theta.$$

Por hipótese de indução, segue que

$$\begin{aligned}\cos(n+1)\theta &= 2\cos(n\theta)\cos\theta - \cos(n-1)\theta\\ &= 2\cos\theta(2^{n-1}\cos^n\theta + \sum_{j=0}^{n-1}\beta_{n,j}\cos^j\theta) - 2^{n-2}\cos^{n-1}\theta - \sum_{j=0}^{n-2}\beta_{n-1,j}\cos^j\theta.\end{aligned}$$

Portanto,

$$\cos(n+1)\theta = \frac{1}{2^n}\cos^{n+1}\theta + \sum_{j=0}^{n}\beta_{n+1,j}\cos^j\theta.$$

Agora, vamos substituir $\cos\theta = x$, podemos escrever

$$T_0(x) = \cos(0\theta) = \cos(0) = 1$$
$$T_1(x) = \cos(1\theta) = x$$
$$T_2(x) = \cos 2\theta = 2x(x) - 1 = 2x^2 - 1$$
$$T_3(x) = \cos 3\theta = 2x(2x^2 - 1) - x = 4x^3 - 3x$$
$$T_4(x) = \cos 4\theta = 2x(4x^3 - 3x) - (2x^2 - 1) = 8x^4 - 8x^2 + 1$$
$$T_5(x) = \cos 5\theta = 2x(8x^4 - 8x^2 + 1) - (4x^3 - 3x) = 16x^5 - 20x^3 + 5x$$
$$\vdots$$

.

A expressão geral de cada polinômio T_n é dada por

$$T_n(x) = \frac{n}{2}\sum_{j=0}^{[n:2]}(-1)^j\frac{(n-j-1)!}{j!(n-2j)!}(2x)^{n-2j}, \qquad n = 1, 2, \ldots$$

onde $[n : 2] = n/2$, se n é par e $[n : 2] = (n-1)/2$ se n é ímpar.

Além disso, podemos estabelecer a relação:

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x). \tag{22}$$

De fato, vamos usar novamente indução finita. Tomando $T_n(x) = \cos(n\theta)$, para $n = 1$,

$$T_1(x) = \cos(1\theta) = \cos(\theta).$$

Supondo que seja verdade para algum k, mostremos que é válido para $k+1$. Assim,

$$T_{k+1}(x) = \cos(k+1)\theta = \cos(k\theta+\theta) = \cos(k\theta)\cos\theta - sen(k\theta)sen\theta.$$

e

$$T_{k-1}(x) = \cos(k-1)\theta = \cos(k\theta-\theta) = \cos(k\theta)\cos\theta + sen(k\theta)sen\theta.$$

Adicionando ambos os lados, obtemos

$$T_{k+1}(x) + T_{k-1}(x) = 2\cos(k\theta)\cos\theta$$

$$T_{k+1}(x) = 2T_k(x)x - T_{k-1}(x)$$
$$T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x)$$

Esta fórmula de recursão nos oferece sucessivamente os gráficos das funções $T_n(x)$, alguns representados na Figura 4.

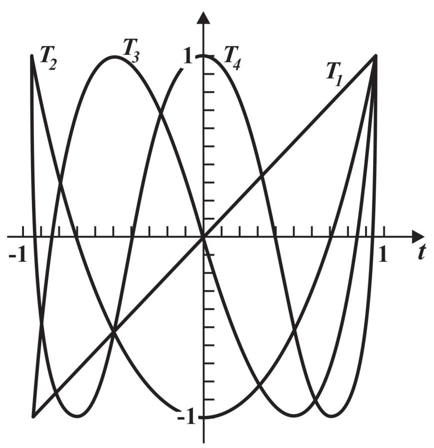

Figura 4: Gráfico de T_1, T_2, T_3 e T_4. Fonte: Adaptado de Kreyszig (1978, p. 350).

Portanto, obtemos a seguinte formulação do resultado que queríamos, que expressa a propriedade mais famosa dos *Polinômios de Chebyshev*.

NASCIMENTO, C. A.; GADOTTI, M. C. Teoria da aproximação em espaços de Banach e de Hilbert. **C.Q.D. – Revista Eletrônica Paulista de Matemática**, Bauru, v. 18, p. 112–128, jul. 2020. Edição Iniciação Científica.
DOI: 10.21167/cqdvol18ic202023169664canmcg112128 Disponível em: www.fc.unesp.br/departamentos/matematica/revista-cqd/

Teorema 24 (Polinômio de Chebyshev) *Seja o polinômio definido por*

$$\widetilde{T}_n(t) = \frac{1}{2^{n-1}} T_n(t) = \frac{1}{2^{n-1}} \cos(n \arccos t) \quad n \geq 1. \tag{23}$$

tem o menor desvio máximo de 0 sobre o intervalo $[-1, 1]$ *entre todos os polinômios reais considerados em* $C([-1, 1], \mathbb{R})$ *que possuem grau n e coeficiente 1.*

A melhor aproximação uniforme para a função $x \in C([-1, 1], \mathbb{R})$, definido por $x(t) = t^n$, para o gerado $Y = [\{y_o, \cdots, y_{n-1}\}]$ com y_j dado por (21), (isto é, a aproximação por um polinômio real de grau menor que n) é y definido por

$$y(t) = x(t) - \frac{1}{2^{n-1}} T_n(t).$$

Note que

$$|y(t) - x(t)| = |(1/2^{n-1}) T_n(t)| \to 0,$$

quando $n \to \infty$.

Neste trabalho, discutimos o que é uma melhor aproximação e analisamos as condições necessárias e suficientes para a existência e unicidade de melhores aproximações em espaços normados e espaços de Hilbert.

Como aplicação de aproximação em espaço normado, apresentamos uma das propriedades mais importantes dos polinômios de Chebyshev que é o fato de possuir o menor desvio máximo de 0 no intervalo $[-1, 1]$ entre todos os polinômios reais considerados em $C([-1, 1], \mathbb{R})$ que apresentam grau n e coeficiente 1. Para aprofundar o conhecimento sobre esta propriedade o leitor pode consultar [5].

5 Referências

[1] STEFFENS, K.-G. **The history of approximation theory**: from Euler to Bernstein. Boston: Birkhäuser, 2006.

[2] KAFASH, B.; DELAVARKHALAFI, A.; KARBASSI, S.M. Application of Chebyshev polynomials to derive efficient algorithms for the solution of optimal control problems. **Scientia Iranica**, v. 19, n. 3, p. 795-805, 2012.

[3] KREYSZIG, E. **Introductory functional analysis with applications**. New York: John Wiley & Sons, 1978.

[4] HÖNIG, C. S. **Análise funcional e aplicações**. São Paulo: IME - USP, 1970, 2 v.

[5] SZEGO, G. **Orthogonal polynomials**. 3rd ed. Providence: American Mathematical Society, 1967.