Quadrados Mínimos Contínuos

Amadeus Haag Mazzini da Silva
UFPR - Cálculo Numérico - 2021
amadeus.mazzini@ufpr.br

RESUMO

Este artigo apresenta o problema de ajuste de curvas pelo método dos quadrados mínimos, para o caso contínuo, tal como a aplicação de Polinômios Ortogonais como método alternativo para facilitar o processo de ajuste. Será apresentado um processo de Ortogonalização e Polinômios ortogonais específicos. Todos os gráficos foram gerados em Linguagem Julia e o Algoritmo para plotagem de gráficos pode ser encontrada no seguinte repositório: `https://github.com/amadeusmazzini/Quad_Min_Continuos`.

PALAVRAS CHAVE. Quadrados Mínimos, Polinômios Ortogonais, Polinômios de Chebyshev.

1. Introdução

O estudo da teoria da aproximação por quadrados mínimos contínuos surge quando uma função é dada de forma explícita, mas queremos encontrar um tipo mais “simples” de função, tal como um polinômio, que possa ser utilizado para determinar valores aproximados da função dada. O método é uma técnica de otimização matemática que procura encontrar o melhor ajuste para uma função, tentando minimizar a área da diferenças dos quadrados do valor estimado e o valor observado.

2. Fundamentação teórica

2.1. Aproximação por Mínimos Quadrados

[1]Suponha que $f \in C[a, b]$ e que queiramos determinar um polinômio P_n de grau no máximo n que minimize o erro

$$\int_a^b [f(x) - P_n(x)]^2 dx$$

Para determinar um polinômio de aproximação por mínimos quadrados, ou seja, um polinômio para minimizar essa expressão, seja

$$P_n(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = \sum_{k=0}^{n} a_k x^k$$

O objetivo é que a área sob a curva do quadrado dos desvios seja mínima, ou seja

$$E \equiv E_2(a_0, a_1, \ldots, a_n) = \int_a^b [f(x) - \sum_{k=0}^{n} a_k x^k]^2 dx$$

O problema é encontrar os coeficientes reais $a_0, a_1, \ldots, a_n$ que minimizem E e uma condição necessária para que isso ocorra é que

$$\frac{\partial E}{\partial a_j} = 0, \quad \forall j = 0, 1, \ldots, n$$

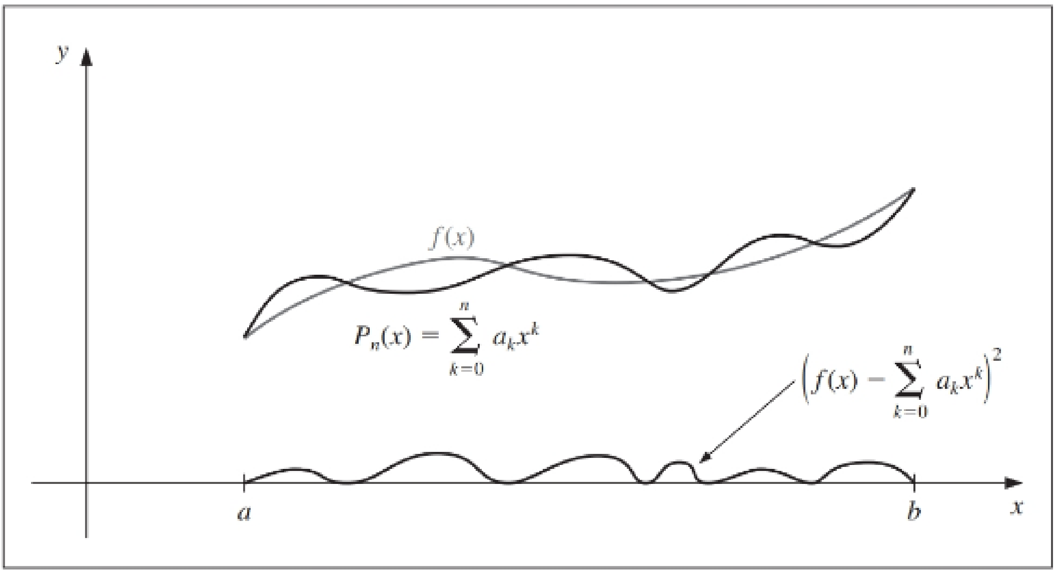


Figura 1

Aplicando o quadrado da diferença temos que

$$E = \int_a^b f(x)^2 dx \quad -2\sum_{k=0}^{n} a_k \int_a^b x^k f(x) dx + \quad \int_a^b (\sum_{k=0}^{n} a_k x^k)^2 dx$$

Logo temos que

$$\frac{\partial E}{\partial a_j} = -2\int_a^b x^j f(x) dx \quad +2\sum_{k=0}^{n} a_k \int_a^b x^{j+k} dx$$

Portanto para encontrarmos $P_n(x)$, as $(n+1)$ **Equações Normais** lineares

$$\sum_{k=0}^{n} a_k \int_a^b x^{j+k} dx = \int_a^b x^j f(x) dx \qquad \forall j = 0, 1, \ldots, n$$

devem ser resolvidas para se determinar as $(n+1)$ incógnitas a.

Alternativamente, podemos escrever [2]

$$\begin{cases} \int_a^b a_0 x^0 x^0 dx + \cdots + \int_a^b a_n x^0 x^n dx = \int_a^b f(x) x^0 dx \\ \int_a^b a_0 x^1 x^0 dx + \cdots + \int_a^b a_n x^1 x^n dx = \int_a^b f(x) x^1 dx \\ \quad \vdots \\ \int_a^b a_0 x^n x^0 dx + \cdots + \int_a^b a_n x^n x^n dx = \int_a^b f(x) x^n dx \end{cases}$$

Que é um sistema linear com $(n+1)$ equações e incógnitas $(a_0, a_1, \ldots, a_n)$. O sistema linear acima é chamado sistema das **Equações Normais** e elas sempre têm uma única solução desde que $f \in C[a, b]$.

Em termos matriciais, o sistema das equações normais pode ser escrito como

$$\mathbf{M}\ \boldsymbol{\alpha} = \mathbf{b}$$

Em que

$$\mathbf{M} = (m_{ij}) \in \mathbb{R}^{n\times n}, \quad \boldsymbol{\alpha} = (a_j) \in \mathbb{R}^n \quad \text{e} \quad \mathbf{b} = (b_i) \in \mathbb{R}^n$$

com

$$m_{ij} = \langle x^i, x^j \rangle = \int_a^b x^i x^j dx$$
$$b_i = \langle f(x), x^i \rangle = \int_a^b f(x) x^i$$
$$\forall i, j = 0, 1, \dots, n$$

2.2. Funções Ortogonais

O próximo passo será introduzir as noções de funções peso e ortogonalidade. Uma função integrável w é chamada função peso em um intervalo I se $w(x) \geqslant 0$, para todo x em I, porém $w(x) \not\equiv 0$ em qualquer subintervalo de I. A finalidade de uma função peso é atribuir graus variados de importância a aproximações em certas partes do intervalo.

Diz-se que $(\phi_0, \phi_1, \dots, \phi_n)$ é um conjunto de funções ortogonais[3] no intervalo $[a, b]$, com relação à função peso w, se

$$\int_a^b w(x)\phi_k(x)\phi_j(x)dx \begin{cases} 0, & \text{quando} \quad j \neq k \\ \alpha_j > 0, & \text{quando} \quad j = k \end{cases}$$
$$\forall j = 0, 1, \dots, n$$

2.3. Ortogonalização de Polinômios

Será apresentado um processo [1] que tem base no processo de ortogonalização Gram-Schmidt, que descreve como construir polinômios ortogonais em $[a, b]$ com relação a uma função peso w.

O conjunto de funções polinomiais $(\phi_0(x), \phi_1(x), \dots, \phi_n(x))$ definido da forma seguinte é ortogonal em $[a, b]$ em relação à função peso w:

$$\phi_0(x) \equiv 1 \qquad \phi_1(x) = x - B_1 \qquad \text{para cada } x \text{ em } [a, b]$$

Em que

$$B_1 = \frac{\int_a^b x w(x)[\phi_0(x)]^2 dx}{\int_a^b w(x)[\phi_0(x)]^2 dx}$$

e, quando $k \geqslant 2$

$$\phi_k(x) = (x - B_k)\phi_{k-1}(x) - C_k\phi_{k-2}(x) \qquad \text{para cada } x \text{ em } [a, b]$$

Em que

$$B_k = \frac{\int_a^b x w(x)[\phi_{k-1}(x)]^2 dx}{\int_a^b w(x)[\phi_{k-1}(x)]^2 dx}$$

e

$$C_k = \frac{\int_a^b x w(x)\phi_{k-1}(x)\phi_{k-2}(x) dx}{\int_a^b w(x)[\phi_{k-2}(x)]^2 dx}$$

2.4. Polinômios de Chebyshev

Os Polinômios de Chebychev[4] foram definidos como uma sequência de polinômios, relacionados com a fórmula de Moivre e facilmente obtíveis de forma recursiva pela relação:

$$T_0(x) = 1$$

$$T_1(x) = x$$

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$

E também:

$$T_n(x) = \cos(n \arccos(x))$$

Onde

$$T_n(\cos(\theta)) = \cos(n\theta)$$

A relação forma uma sequência de polinômios ortogonais com relação ao peso

$$w(x) = \frac{1}{\sqrt{1-x^2}}$$

no intervalo $[-1, 1]$, ou seja:

$$\int_{-1}^{1} \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}} dx \begin{cases} 0 & \text{quando} \quad n \neq m \\ \pi & \text{quando} \quad n = m = 0 \\ \frac{\pi}{2} & \text{quando} \quad n = m \neq 0 \end{cases}$$

Tal propriedade pode ser provada definindo $x = \cos(\theta)$ e usando a identidade

$$T_n(\cos(\theta)) = \cos(n\theta)$$

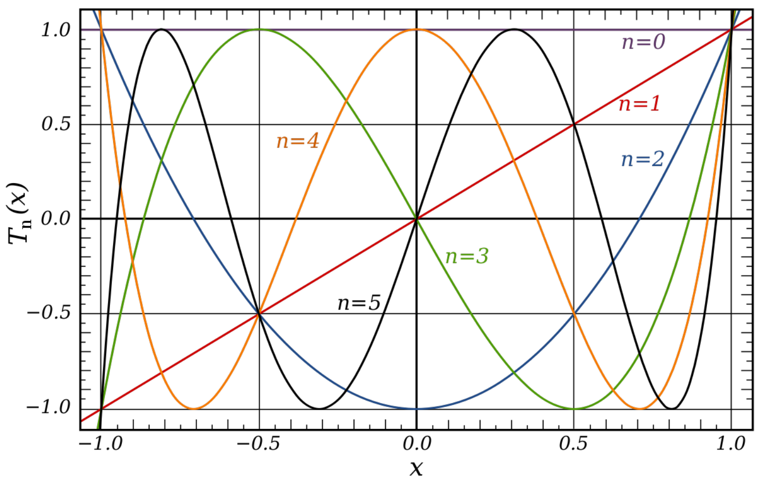


Figura 2: Gráfico de alguns polinômios de Chebyshev

3. Metodologia e Análise

3.1. Quadrados Mínimos utilizando Polinômios Ortogonais

Suponha que $(\phi_0(x), \phi_1(x), \ldots, \phi_n(x))$ seja um conjunto de funções linearmente independentes em $[a, b]$, que w seja uma função peso em $[a, b]$ e, para $f \in C[a, b]$, uma combinação linear

$$P(x) = \sum_{k=0}^{n} a_k \phi_k(x)$$

deve ser determinada para encontrar o erro

$$E = E(a_0, a_1, \ldots, a_n) = \int_a^b w(x) \left[f(x) - \sum_{k=0}^{n} a_k \phi_k(x) \right]^2 dx$$

Esse problema se reduz à situação considerada no início deste artigo, no caso especial no qual $w(x) \equiv 1$ e $\phi_k(x) = x^k$ para cada $k = 0, 1, \ldots, n$. Se for possível escolher as funções ϕ de forma que $(\phi_0, \phi_1, \ldots, \phi_n)$ sejam ortogonais, as equações normais são facilmente resolvidas por

$$a_j = \frac{\int_a^b w(x) f(x) \phi_j(x) dx}{\int_a^b w(x) \phi_j(x)^2 dx}, \quad \forall j = 0, 1, \ldots, n.$$

Sendo assim, o problema da aproximação por mínimos quadrados é muito simplificado quando as funções ϕ_j são escolhidas de modo a satisfazerem as condições de ortogonalidade. Embora as definições e teoremas apresentados permitam extensas categorias de funções ortogonais, somente conjuntos ortogonais de polinômios serão considerados neste artigo.

3.1.1. Aproximação utilizando Polinômios de Chebyshev

Os coeficientes a_n que multiplicam os n polinônimos de Chebyshev que aproximam uma função $f(x)$ no intervalo $[-1, 1]$ são:

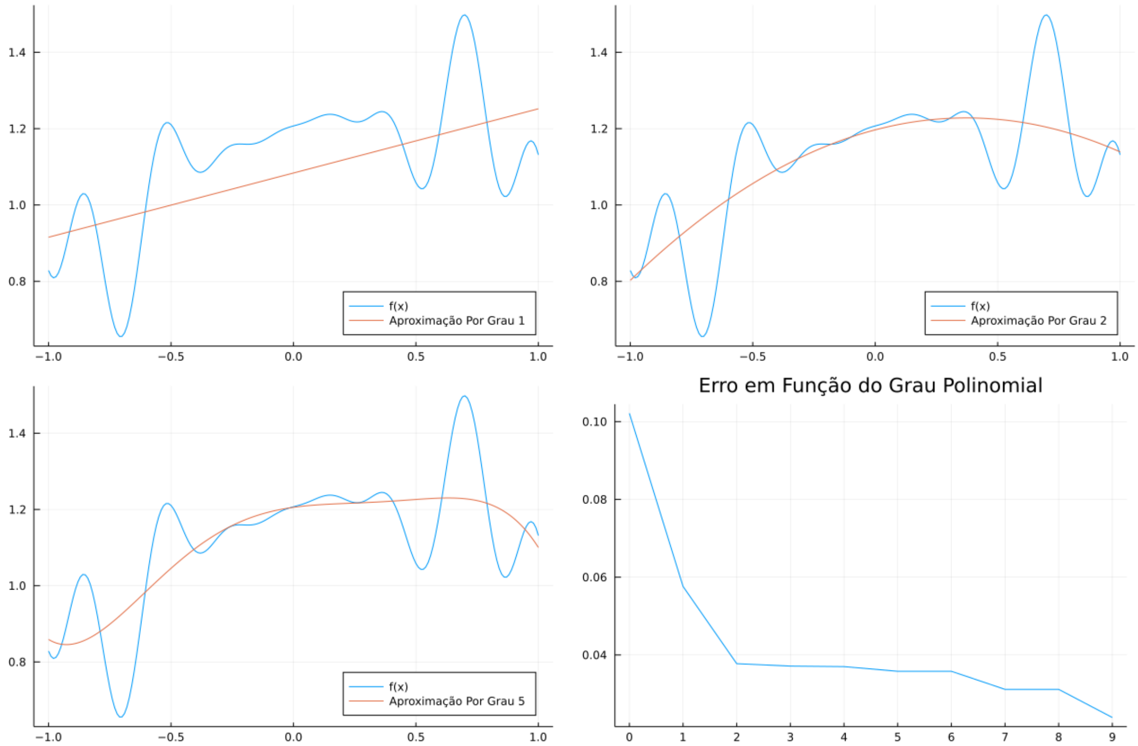
$$a_0 = \frac{\int_{-1}^{1} \frac{1}{\sqrt{1-x^2}} f(x) dx}{\pi}$$

$$a_j = \frac{2 \int_{-1}^{1} \frac{1}{\sqrt{1-x^2}} f(x) T_n dx}{\pi}, \quad \forall j = 1, \ldots, n$$

Exemplo: Vamos aproximar a função

$$f(x) = \frac{1 + \sqrt{2 + x}}{1 + \sqrt{1 + x^2}} + \frac{\sin(5\pi x) \cos(5x) x^3}{x^2 + 1}$$

No intervalo $[-1, 1]$ por polinômios de Chebyshev.



Nem sempre iremos querer uma aproximação no intervalo [-1,1], portanto, para fazer em um intervalo genérico [a,b] iremos precisar fazer uma transformação de variável da forma que

$$\int_{-1}^{1} f(t)dt = \int_{a}^{b} g(x)dx$$

Para isso ocorrer, é necessário que:

$$x = \frac{a+b}{2} + t\frac{(b-a)}{2}$$

Criando uma função z dessa forma

$$z(x) = \frac{a+b}{2} + x\frac{(b-a)}{2}$$

que tem como inversa:

$$z^{-1}(x) = \frac{(2x-a-b)}{(b-a)}$$

Chamaremos uma nova função

$$g(x) = f\Big(z(x)\Big)$$

Após aplicar as equações

$$\theta_0 = \frac{\int_{-1}^{1} \frac{1}{\sqrt{1-x^2}} g(x) dx}{\pi}$$

$$\theta_j = \frac{2 \int_{-1}^{1} \frac{1}{\sqrt{1-x^2}} g(x) T_j dx}{\pi}, \quad \forall j = 1, \ldots, n$$

É necessário re-transformar as variáveis, para isso, devemos aplicaremos a inversa de z no polinômio:

$$P(x) = \sum_{k=1}^{n-1} \theta_k T_{k-1}$$

Ou seja,

$$P\left(z^{-1}(x)\right) = \sum_{k=1}^{n-1} \theta_k T_{k-1}$$

E assim teremos concluido a aproximação de de uma função f que não esta no intervalo $[-1, 1]$ com Polinômios de Chebyshev.

4. Considerações finais

O método dos quadrados mínimos, caso contínuo, é usado para encontrar uma função $\phi(x) = \alpha_1\phi_1(x) + \alpha_2\phi_2(x) + \cdots + \alpha_n\phi_n(x)$, que melhor se aproxima de uma certa função f em um intervalo $[a, b]$. Tal como no caso discreto, os coeficientes $\alpha_1, \ldots, \alpha_n$ são obtidos resolvendo um sistema linear $\mathbf{M}\ \boldsymbol{\alpha} = \mathbf{b}$, conhecido como sistema das equações normais, que são consideravelmente simplificadas se as funções ϕ_n escolhidas forem ortogonais entre si. Além de Chebyshev, existem diversos Polinômios com a propriedade de ortogonalidade, abrindo possibilidades de aproximações variadas.

Referências

[1] Burden, R. L. (2016). *Análise Numérica - Tradução da 10ª edição norte-americana*. Cengage Learning Brasil. Acessado: 05/08/2021.

[2] Valle, M. E. (2015). Método dos quadrados mínimos – caso contínuo. `https://www.ime.unicamp.br/~valle/Teaching/2015/MS211/Aula18.pdf`. Acessado: 21/07/2021.

[3] WIKIPEDIA (2017). Funções ortogonais. `https://pt.wikipedia.org/wiki/Fun%C3%A7%C3%B5es_ortogonais`. Acessado: 21/07/2021.

[4] WIKIPEDIA (2017). Polinômios de tchebychev. `https://pt.wikipedia.org/wiki/Polin%C3%B4mios_de_Tchebychev`. Acessado: 26/07/2021.