Aproximação de uma Função a uma Variável Real pelo Método dos Mínimos Quadrados

Márcio Antônio de Andrade Bortoloti

Cálculo Numérico

Departamento de Ciências Exatas e Tecnológicas - DCET
Universidade Estadual do Sudoeste da Bahia - UESB

Sumário

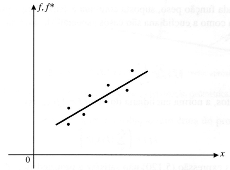
Introdução

Introdução

- Vamos tratar do problema da aproximação linear de uma função f por uma função f_h da forma
$$f_h(x) = \sum_{i=0}^{n} c_i \phi_i(x),$$
onde ϕ_i são funções escolhidas apropriadamente.
- Dado um conjunto de pontos $(x_i, f(x_i))$, $0 \le i \le m$, os coeficientes c_i são determinados tais que
$$f_h(x_i) = f(x_i), \quad 0 \le i \le m.$$
- Obtendo assim o seguinte sistema
$$\begin{cases} c_0\phi_0(x_0) + c_1\phi_1(x_0) + \cdots + c_n\phi_n(x_0) = f(x_0) \\ c_0\phi_0(x_1) + c_1\phi_1(x_1) + \cdots + c_n\phi_n(x_1) = f(x_1) \\ \vdots \\ c_0\phi_0(x_m) + c_1\phi_1(x_m) + \cdots + c_n\phi_n(x_m) = f(x_m) \end{cases}$$

Introdução

- Se $m = n$ e as funções ϕ_i forem linearmente independentes então o sistema anterior terá uma única solução. Nesse caso a função é chamada interpolante.
- Se $m > n$ o sistema tem mais equações do que incógnitas, sendo chamado de sobredeterminado. Nesse caso as equações serão satisfeitas aproximadamente. Essa situação é comum na prática e é utilizada para dois diferentes tipos de regularização:
  - Reduzir o efeito de erros nos valores das funções;
  - Dar a curva uma forma regular, entre pontos da malha de aproximação, como se pode observar na figura abaixo:



- O tratamento de sistemas sobredeterminados neste caso é feito pelo Método dos Mínimos Quadrados.
- Para isso, vamos utilizar algumas definições:

Definição
Seja $f : [a, b] \to \mathbb{R}$ contínua. Vamos definir as sguintes normas:

- *Norma do Máximo*

$$\|f\|_\infty = \max_{a \leq x \leq b} \{f(x)\}$$

- *Norma Euclidiana*

$$\|f\|_2 = \left[\int_a^b [f(x)]^2 dx \right]^{1/2}$$

- *Norma Euclidiana Ponderada*

$$\|f\|_{2,w} = \left[\int_a^b [f(x)]^2 w(x)\, dx \right]^{1/2}$$

Introdução

- O Método dos Mínimos Quadrados toma como base a minimização da função

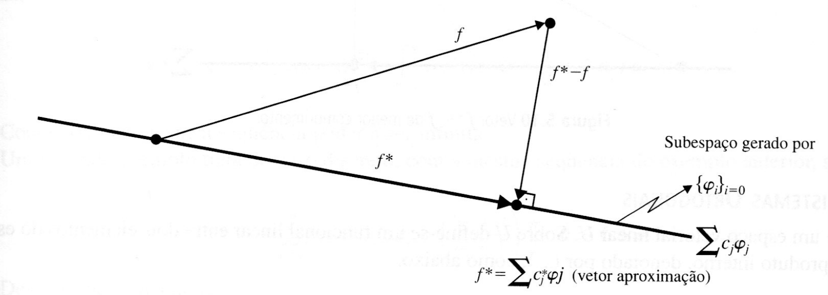
$$\phi(x) = \|f(x) - f_h(x)\|_p.$$

- Dessa forma considere um conjunto de $n+1$ funções linearmente independentes ϕ_i, $0 \leq i \leq n$ gerando um subespaço de dimensão $n+1$.
- O problema de aproximação fica sendo encontrar, no subespaço gerado por

$$\{\phi_i(x)\}_{i=0}^{n},$$

um vetor cuja distância em relação à f seja a menor possível.

- Quando a norma euclidiana for a escolhida, a solução do problema de aproximação será a generalização do que ocorre geometricamente no $\mathbb{R}^2$ ou no $\mathbb{R}^3$, isto é, a menor distância de um ponto a um subespaço linear é o comprimento do vetor que é perpendicular ao subespaço..



O Método dos Mínimos Quadrados

O Método dos Mínimos Quadrados

- Considere uma função $f : [a, b] \to \mathbb{R}$ contínua para ser aproximada por uma função

$$f_h(x) = \sum_{i=0}^{n} c_i \phi_i(x).$$

- Os coeficientes c_i serão determinados de modo que a norma euclidiana ponderada do vetor $f - f_h$

$$\|f - f_h\|_{2,w}^2 = \int_a^b |f_h(x) - f(x)|^2 w(x)\, dx.$$

seja a menor possível. No caso discreto,

$$\|f - f_h\|_{2,w}^2 = \sum_{i=0}^{n} |f_h(x_i) - f(x_i)|^2 w(x_i).$$

O Método dos Mínimos Quadrados

- Geometricamente, os coeficientes c_i são determinados para que $f - f_h$ seja ortogonal ao subespaço gerado por $\{\phi\}_{i=0}^{n}$.

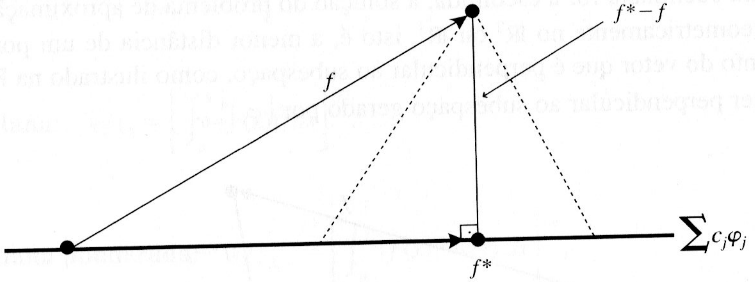


Figura 5.10 Vetor $f^* - f$ de menor comprimento.

Sistemas Ortogonais

Definição (Produto Interno)
Seja U um espaço vetorial e $f, g \in U$. O produto interno de f e g, (f, g), é um funcional que satisfaz às seguintes propriedades:

- $(f, g) = (g, f)$ *para toda* $f, g \in U$;
- $(c_1 f + c_2 g, \phi) = c_1(f, \phi) + c_2(g, \phi)$ *para todo* $c_1, c_2 \in \mathbb{R}$ *e* $\phi \in U$.
- $(f, f) \geq 0$ *para toda* $f \in U$;
- $(f, f) = 0$ *se e somente se* $f = 0$.

Com o produto interno definido, pode-se estabelecer uma métrica induzida por tal produto interno pondo:

$$\|f\|^2 = (f, f).$$

Definição (Ortogonalidade)
Duas funções f e g são ortogonais se $(f, g) = 0$. Uma sequência finita ou infinita de funções ϕ_i forma um sistema ortogonal se

$$(\phi_i, \phi_j) = 0 \text{ para todo } i \neq j \text{ e } \|\phi_i\| \neq 0, \text{ para todo } i.$$

Se além disso, $\|\phi_i\| = 1$ para todo i, então a sequência é chamada de sistema ortonormal.

Exemplo

Considere a sequência de funções definidas em $[0, \pi]$ por $\phi_j(x) = \cos jx$ para $j = 0, 1, 2, \cdots, m$, com o produto interno

$$(f, g) = \int_0^{\pi} f(x)g(x)\, dx.$$

Nesse caso tem-se

$$\begin{aligned}(\phi_j, \phi_k) &= \int_0^\pi \cos jx \cos kx \; dx = \int_0^\pi \frac{1}{2}[\cos(j-k)x + \cos(j+k)x]dx \\ &= \frac{1}{2}\Big[\frac{\operatorname{sen}(j-k)x}{j-k} + \frac{\operatorname{sen}(j+k)x}{j+k}\Big]_0^\pi, \quad j \neq k \\ &= 0, \quad j \neq k, \quad 1 \leq j, k \leq m.\end{aligned}$$

Tem-se ainda

$$\|\phi_j\|^2 = \|\cos jx\|^2 = \int_0^\pi \cos^2 jx \, dx = \frac{\pi}{2}, \quad 1 \leq j \leq m.$$

Solução do Problema de Aproximação

- Seja f_h a aproximação de f por mínimos quadrados escrita como

$$f_h(x) = \sum_{j=0}^{m} c_j^* \phi_j(x).$$

- Desejamos determinar os coeficientes c_j^* tais que $f - f_h$ seja o menor possível em uma dada norma.
- De modo equivalente $f_h - f$ deve ser ortogonal ao subespaço gerado por $\{\phi_j\}_{j=0}^{n}$.
- Vamos tomar o vetor $\sum_{j=0}^{n} c_j \phi_j - f(x)$ tal que $c_j \neq c_j^*$ para pelo menos um j.
- Temos então

$$\sum_{j=0}^{n} c_j \phi_j - f(x) = \sum_{j=0}^{n} (c_j - c_j^*)\phi_j + (f_h(x) - f(x)).$$

Solução do Problema de Aproximação

- Como $f_h - f$ é ortogonal a todo ϕ_i, tem-se $f_h - f$ ortogonal a

$$\sum_{j=0}^{n}(c_j - c_j^*)\phi_j(x).$$

- Assim

$$\Big\|\sum_{j=0}^{n} c_j\phi_j(x) - f(x)\Big\|^2 = \Big\|\sum_{j=0}^{n}(c_j - c_j^*)\phi_j(x) + [f_h(x) - f(x)]\Big\|^2$$

$$= \Big(\sum_{j=0}^{n}(c_j - c_j^*)\phi_j(x) + [f_h(x) - f(x)], \sum_{j=0}^{n}(c_j - c_j^*)\phi_j(x) + [f_h(x) - f(x)]\Big)$$

$$= \Big(\sum_{j=0}^{n}(c_j - c_j^*)\phi_j(x), \sum_{j=0}^{n}(c_j - c_j^*)\phi_j(x)\Big) + 2\Big(\sum_{j=0}^{n}(c_j - c_j^*)\phi_j(x), f_h(x) - f(x)\Big)$$
$$+\Big(f_h(x) - f(x), f_h(x) - f(x)\Big)$$

Solução do Problema de Aproximação

Segue então que

$$\begin{aligned}\Big\|\sum_{j=0}^{n} c_j\phi_j(x) - f(x)\Big\|^2 &= \Big\|\sum_{j=0}^{n}(c_j - c_j^*)\phi_j(x)\Big\|^2 + \|f_h(x) - f(x)\|^2 \\ &> \|f_h(x) - f(x)\|^2\end{aligned}$$

Assim o vetor erro com a menor norma é $f_h - f$ que é ortogonal às funções $\{\phi_i\}_{i=0}^n$. Isso nos dá uma condição de ortogonalidade

$$\Big(\sum_{j=0}^{n} c_j^*\phi_j(x) - f(x), \phi_k(x)\Big) = 0, \quad 0 \le k \le n.$$

Essa condição forma o seguinte sistema

$$\begin{cases}(\phi_0,\phi_0)c_0^* + (\phi_0,\phi_1)c_1^* + \cdots + (\phi_0,\phi_n)c_n^* = (\phi_0, f)\\ (\phi_1,\phi_0)c_0^* + (\phi_1,\phi_1)c_1^* + \cdots + (\phi_1,\phi_n)c_n^* = (\phi_1, f)\\ \vdots \\ (\phi_n,\phi_0)c_0^* + (\phi_n,\phi_1)c_1^* + \cdots + (\phi_n,\phi_n)c_n^* = (\phi_n, f)\end{cases}$$

Solução do Problema de Aproximação

- Vamos analisar o sistema anterior na sua forma homogênea, isto é,

$$\sum_{j=0}^{n}(\phi_k, \phi_j)c_j = 0, \quad 0 \le k \le n.$$

- Vamos admitir que o sistema admite solução não-trivial para ao menos um $c_j \neq 0$.
- Mas

$$\begin{aligned}
\Big\| \sum_{j=0}^{n} c_j \phi_j(x) \Big\|^2 &= \Big(\sum_{j=0}^{n} c_j \phi_j(x), \sum_{k=0}^{n} c_k \phi_k(x) \Big) \\
&= \sum_{k=0}^{n} \sum_{j=0}^{n} (\phi_k(x), \phi_j(x)) c_j c_k \\
&= \sum_{k=0}^{n} \Big(\sum_{j=0}^{n} (\phi_k(x), \phi_j(x)) c_j \Big) c_k = \sum_{k=0}^{n} 0 \, c_k = 0
\end{aligned}$$

- O que contradiz a hipótese de que $\{\phi_i\}_{i=0}^{n}$ é linearmente independente.
- Logo o sistema tem solução única trivial e daí as equações têm solução única.

Solução do Problema de Aproximação

Suponha que $f \in \mathcal{C}[a,b]$ e queiramos determinar um polinômio $P_n(x)$ de grau no máximo n que minimize o erro

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

Seja $P_n(x) = \sum_{k=0}^n a_k x^k$ e defina

$$E = E(a_0, a_1, \cdots, a_n) = \int_a^b \left(f(x) - \sum_{k=0}^n a_k x^k \right)^2 dx.$$

Uma condição necessária para determinar os a_is que minimizem E é

$$\frac{\partial E}{\partial a_j} = 0 \quad \text{para cada } j = 0, 1, \cdots, n.$$

Solução do Problema de Aproximação

Tem-se então

$$\frac{\partial E}{\partial a_j} = -2\int_a^b x^j f(x)\,dx + 2\sum_{k=0}^{n} a_k \int_a^b x^{j+k}\,dx \text{ para cada } j = 0, 1, \cdots, n$$

Portanto, para encontrar $P_n(x)$ as $n+1$ equações lineares

$$\sum_{k=0}^{n} a_k \int_a^b x^{j+k}\,dx = \int_a^b x^j f(x)\,dx \text{ para cada } j = 0, 1, \cdots, n$$

devem ser resolvidas.

Essas equações são denominadas **equações normais**.

Exemplo

Encontre o polinômio de aproximação por mínimos quadrados de segundo grau para a função $f(x) = \text{sen}\ (\pi x)$ no intervalo $[0, 1]$.

Seja $P_2(x) = a_2x^2 + a_1x + a_0$. Das equações normais temos

$$\begin{aligned}
a_0 \int_0^1 1\, dx + a_1 \int_0^1 x\, dx + a_2 \int_0^1 x^2\, dx &= \int_0^1 \text{sen}\ (\pi x)\, dx \\
a_0 \int_0^1 x\, dx + a_1 \int_0^1 x^2\, dx + a_2 \int_0^1 x^3\, dx &= \int_0^1 x\text{sen}\ (\pi x)\, dx \\
a_0 \int_0^1 x^2\, dx + a_1 \int_0^1 x^3\, dx + a_2 \int_0^1 x^4\, dx &= \int_0^1 x^2\text{sen}\ (\pi x)\, dx
\end{aligned} \tag{1}$$

Exemplo

Calculando as integrais obtemos

$$\begin{aligned} a_0 + \frac{a_1}{2} + \frac{a_2}{3} &= \frac{2}{\pi} \\ \frac{a_0}{2} + \frac{a_1}{3} + \frac{a_2}{4} &= \frac{1}{\pi} \\ \frac{a_0}{3} + \frac{a_1}{3} + \frac{a_2}{5} &= \frac{\pi^2 - 4}{\pi^3} \end{aligned}$$

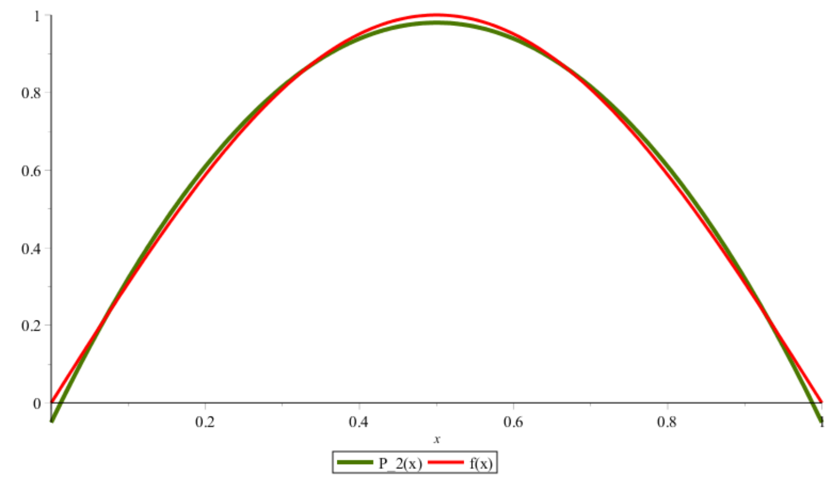
cuja solução é dada por

$$a_0 = \frac{12\pi^2 - 120}{\pi^3} \approx -0.050465 \text{ e } a_1 = -a_2 = \frac{720 - 60\pi^2}{\pi^3} \approx 4.12251.$$

Consequentemente,

$$P_2(x) = -4.12251x^2 + 4.12251x - 0.050465.$$

Exemplo



Polinômios Ortogonais

Polinômios Ortogonais

- Os elementos de um conjunto de polinômios, ortogonais entre si em relação a um produto interno e que podem ser gerados por meio de fórmulas recursivas, são denominados polinômios ortogonais.
- A utilização desses polinômios é frequente nas ciências e tecnologia em vista de possuirem boas qualidades de convergência; serem de fácil manipulação algébrica e computacional e fornecem, em geral, representações bem condicionadas de funções.

Definição (Polinômios de Legendre)
Os polinômios de Legendre são definidos em $[-1,1]$ *pela fórmula*

$$
\begin{aligned}
P_0(x) &= 1 \\
P_n(x) &= \frac{1}{2^n n!}\frac{d^n}{dx^n}[(x^2-1)^n], \quad n \geq 1
\end{aligned}
$$

Propriedades Exercício

- Os polinômios de Legendre são ortogonais em relação ao produto interno usual para funções contínuas no intervalo $[-1,1]$

$$(f,g) = \int_{-1}^{1} f(x)g(x)\,dx,$$

de onde se tem

$$(P_n, P_j) = \begin{cases} 0, & \text{se } n \neq j \\ \frac{2}{2^{n+1}}, & \text{se } n = j \end{cases}$$

Polinômios Ortogonais - Polinômios de Legendre

- Os polinômios de Legendre apresentam a seguinte simetria (exercício)

$$P_n(-x) = (-1)^n P_n(x).$$

- Os polinômios de Legendre podem ser dados pela seguinte recorrência: (exercício)

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= x \\ P_{n+1}(x) &= \left(\frac{2n+1}{n+1}\right) x P_n(x) - \frac{n}{n+1} P_{n-1}(x), \quad n \geq 1. \end{aligned}$$

- Para $x \in [-1, 1]$ tem-se (exercício)

$$|P_n(x)| \leq 1.$$

Definição (Polinômios de Tchebycheff)
A família de polinômios de Tchebycheff é definida em $[-1,1]$ *pela fórmula*

$$T_n(x) = \cos[n\cos^{-1}(x)], \quad n = 0, 1, 2, \cdots .$$

Propriedades Exercícios

- A fórmula recursiva para gerar esses polinômios é dada por

$$\begin{aligned} T_0(x) &= 1 \\ T_1(x) &= x \\ T_{n+1}(x) &= 2xT_n(x) - T_{n-1}(x), \quad n = 1, 2, 3, \cdots . \end{aligned}$$

- O coeficiente do termo de maior potência é dada por $A_0 = 1$ e $A_n = 2^{n-1}$, $n \geq 1$.

Polinômios Ortogonais - Polinômios de Tchebycheff

- Os polinômios de Tchebycheff apresentam a seguinte simetria

$$T_n(-x) = (-1)^n T_n(x).$$

- O polinômio de Tchebycheff de ordem n possui n zeros x_k dados por

$$x_k = \cos\left[\left(\frac{2k+1}{n}\right)\frac{\pi}{2}\right], \quad 0 \leq k \leq n-1.$$

- Os polinômios de Tchebycheff são ortogonais em relação ao produto interno

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

resultando em

$$(T_i, T_k) = \begin{cases} 0, \text{ se } j \neq k \\ \pi/2, \text{ se } j = k \neq 0 \\ \pi, \text{ se } j = k = 0 \end{cases}$$