MS211 - Cálculo Numérico

Aula 20 – Método dos Quadrados Mínimos: Caso Discreto em $\mathbb{R}^N$ e Caso Contínuo.

Marcos Eduardo Valle
Matemática Aplicada
IMECC - Unicamp

Introdução

Na aula anterior, vimos o problema do ajuste de curvas pelo método dos quadrados mínimos para o caso discreto.

Introdução

Na aula anterior, vimos o problema do ajuste de curvas pelo método dos quadrados mínimos para o caso discreto.

Especificamente, dada uma tabela

x	x_1	x_2	$\dots$	x_K
y	y_1	y_2	$\dots$	y_K

e funções bases $g_1, \dots, g_M : \mathbb{R} \to \mathbb{R}$ escolhidas *a priori*, a função

$$\varphi(x) = \alpha_1 g_1(x) + \alpha_2 g_2(x) + \ldots + \alpha_M g_M(x),$$

que minimiza a soma dos quadrados dos desvios é obtida resolvendo o sistema linear $\mathbf{A}\alpha = \mathbf{b}$, em que

$$a_{ij} = \sum_{k=1}^{K} g_i(x_k) g_j(x_k) \quad \text{e} \quad b_i = \sum_{k=1}^{K} y_k g_i(x_k), \quad \forall i, j = 1, \dots, M.$$

Caso Discreto em $\mathbb{R}^N$

De forma mais geral, dada uma tabela

$$\begin{array}{c|cccc}
x_1 & x_{11} & x_{12} & \ldots & x_{1K} \\
x_2 & x_{21} & x_{22} & \ldots & x_{2K} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
x_N & x_{N1} & x_{N2} & \ldots & x_{NK} \\
\hline
\mathrm{y} & y_1 & y_2 & \ldots & y_K
\end{array}$$

considere funções bases $g_1, \ldots, g_M : \mathbb{R}^N \to \mathbb{R}$ escolhidas *a priori*.

Caso Discreto em $\mathbb{R}^N$

De forma mais geral, dada uma tabela

$$\begin{array}{c|ccccc} x_1 & x_{11} & x_{12} & \ldots & x_{1K} \\ x_2 & x_{21} & x_{22} & \ldots & x_{2K} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_N & x_{N1} & x_{N2} & \ldots & x_{NK} \\ \hline y & y_1 & y_2 & \ldots & y_K \end{array}$$

considere funções bases $g_1, \ldots, g_M : \mathbb{R}^N \to \mathbb{R}$ escolhidas *a priori*.

Defina a função $\varphi : \mathbb{R}^N \to \mathbb{R}$ pela equação

$$\varphi(\mathbf{x}) = \alpha_1 g_1(\mathbf{x}) + \alpha_2 g_2(\mathbf{x}) + \ldots + \alpha_M g_M(\mathbf{x}),$$

para $\mathbf{x} = [x_1, x_2, \ldots, x_N]^T \in \mathbb{R}^N$.

Tal como no caso anterior, os parâmetros $\alpha_1, \ldots, \alpha_M$ que minimizam a soma dos quadrados dos desvios

$$J(\alpha_1, \alpha_2, \ldots, \alpha_M) = \sum_{k=1}^{K} \left(y_k - \varphi(\mathbf{x}_k)\right)^2,$$

em que $\mathbf{x}_k = [x_{1k}, x_{2k}, \ldots, x_{Nk}]^T$, para todo $k = 1, \ldots, K$, é obtida resolvendo o sistema linear $\mathbf{A}\alpha = \mathbf{b}$, em que

$$a_{ij} = \sum_{k=1}^{K} g_i(\mathbf{x}_k) g_j(\mathbf{x}_k) \quad \text{e} \quad b_i = \sum_{k=1}^{K} y_k g_i(\mathbf{x}_k), \quad \forall i, j = 1, \ldots, M.$$

Tal como no caso anterior, os parâmetros $\alpha_1, \ldots, \alpha_M$ que minimizam a soma dos quadrados dos desvios

$$J(\alpha_1, \alpha_2, \ldots, \alpha_M) = \sum_{k=1}^{K} \left(y_k - \varphi(\mathbf{x}_k)\right)^2,$$

em que $\mathbf{x}_k = [x_{1k}, x_{2k}, \ldots, x_{Nk}]^T$, para todo $k = 1, \ldots, K$, é obtida resolvendo o sistema linear $\mathbf{A}\alpha = \mathbf{b}$, em que

$$a_{ij} = \sum_{k=1}^{K} g_i(\mathbf{x}_k) g_j(\mathbf{x}_k) \quad \text{e} \quad b_i = \sum_{k=1}^{K} y_k g_i(\mathbf{x}_k), \quad \forall i, j = 1, \ldots, M.$$

Em termos gerais, note que simplesmente substituímos o escalar x_k pelo vetor $\mathbf{x}_k = [x_{1k}, \ldots, x_{Nk}]^T$.

Exemplo

Considere a tabela de pontos

x	-1	-1	0	+1	+1
y	-1	+1	0	-1	+1
z	4.09	0.70	1.67	0.15	-2.38

e a função

$$\varphi(x,y) = \alpha_1 + \alpha_2 x + \alpha_3 y + \alpha_4 xy.$$

Em outras palavras, temos que

$$\varphi(x,y) = \alpha_1 g_1(x,y) + \alpha_2 g_2(x,y) + \alpha_3 g_3(x,y) + \alpha_4 g_4(x,y).$$

em que

$$g_1(x,y) = 1, \quad g_2(x,y) = x, \quad g_3(x,y) = y \quad \text{e} \quad g_4(x,y) = xy,$$

são as funções base.

Usando as fórmulas anteriores, temos que:

$$\begin{aligned}
a_{11} &= \sum_{k=1}^{5} g_1(x_k, y_k) g_1(x_k, y_k) \\
&= (1)(1) + (1)(1) + (1)(1) + (1)(1) + (1)(1) = 5, \\
a_{12} &= \sum_{k=1}^{5} g_1(x_k, y_k) g_2(x_k, y_k) \\
&= (1)(-1) + (1)(-1) + (1)(0) + (1)(1) + (1)(1) = 0, \\
a_{13} &= \sum_{k=1}^{5} g_1(x_k, y_k) g_3(x_k, y_k) \\
&= (1)(-1) + (1)(+1) + (1)(0) + (1)(-1) + (1)(1) = 0, \\
a_{14} &= \sum_{k=1}^{5} g_1(x_k, y_k) g_3(x_k, y_k) \\
&= (1)(1) + (1)(-1) + (1)(0) + (1)(-1) + (1)(1) = 0.
\end{aligned}$$

Prosseguindo, obtemos o sistema linear

$$\mathbf{A}\alpha = \mathbf{b},$$

em que

$$\mathbf{A} = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} \quad \text{e} \quad \mathbf{b} = \begin{bmatrix} 4.23 \\ -7.02 \\ -5.92 \\ 0.86 \end{bmatrix}.$$

cuja solução é

$$\alpha_1 = 0.85, \quad \alpha_2 = -1.75, \quad \alpha_3 = -1.48 \quad \text{e} \quad \alpha_4 = 0.21.$$

Logo, temos a função

$$\varphi(x, y) = 0.85 - 1.75x - 1.48y + 0.21xy.$$

Ajuste de Curvas – Caso Contínuo

O problema de ajuste de curvas pelo método dos quadrados mínimos também pode ser aplicado para o caso contínuo.

Problema de Ajuste de Curvas – Caso Contínuo

Considere uma função f contínua em um intervalo $[a, b]$. Escolhidas funções contínuas $g_1, g_2, \ldots, g_M$, chamadas **funções base**, desejamos encontrar coeficientes $\alpha_1, \ldots, \alpha_M$ de modo que

$$\varphi(x) = \alpha_1 g_1(x) + \alpha_2 g_2(x) + \ldots + \alpha_M g_M(x),$$

forneça a melhor aproximação de f em $[a, b]$.

Ajuste de Curvas – Caso Contínuo

O problema de ajuste de curvas pelo método dos quadrados mínimos também pode ser aplicado para o caso contínuo.

Problema de Ajuste de Curvas – Caso Contínuo

Considere uma função f contínua em um intervalo $[a, b]$. Escolhidas funções contínuas $g_1, g_2, \ldots, g_M$, chamadas **funções base**, desejamos encontrar coeficientes $\alpha_1, \ldots, \alpha_M$ de modo que

$$\varphi(x) = \alpha_1 g_1(x) + \alpha_2 g_2(x) + \ldots + \alpha_M g_M(x),$$

forneça a melhor aproximação de f em $[a, b]$.

Observe que no caso contínuo φ deve aproximar f em $[a, b]$ e não num conjunto discreto de pontos $x_1, x_2, \ldots, x_K$.

Formulação Matemática

No problema de quadrados mínimos – caso contínuo, a notação $\varphi \approx f$ em $[a, b]$ significa que a área sob a curva do quadrado dos desvios é mínima, ou seja,

$$J(\alpha_1, \ldots, \alpha_M) = \int_a^b (\varphi(x) - f(x))^2 dx,$$

é mínimo.

Formulação Matemática

No problema de quadrados mínimos – caso contínuo, a notação $\varphi \approx f$ em $[a, b]$ significa que a área sob a curva do quadrado dos desvios é mínima, ou seja,

$$J(\alpha_1, \ldots, \alpha_M) = \int_a^b (\varphi(x) - f(x))^2 dx,$$

é mínimo.

Tal como no caso discreto, devemos encontrar os pontos críticos de J, ou seja, escolher $\alpha_1, \ldots, \alpha_M$ de modo que

$$\frac{\partial J}{\partial \alpha_j} = 0, \quad \forall j = 1, \ldots, M.$$

Pela regra da cadeia, a derivada parcial é

$$\frac{\partial J}{\partial \alpha_j} = 2 \int_a^b \Big(\alpha_1 g_1(x) + \ldots + \alpha_M g_M(x) - f(x) \Big) g_j(x) dx.$$

Pela regra da cadeia, a derivada parcial é

$$\frac{\partial J}{\partial \alpha_j} = 2\int_a^b \Big(\alpha_1 g_1(x) + \ldots + \alpha_M g_M(x) - f(x)\Big) g_j(x)dx.$$

Dessa forma, devemos ter

$$\int_a^b \Big(\alpha_1 g_1(x) + \ldots + \alpha_M g_M(x) - f(x)\Big) g_j(x)dx = 0,$$

ou ainda,

$$\int_a^b \alpha_1 g_1(x) g_j(x)dx + \ldots + \int_a^b \alpha_M g_M(x) g_j(x)dx = \int_a^b f(x) g_j(x)dx,$$

para todo $j = 1, \ldots, M$.

Equações Normais

Alternativamente, podemos escrever

$$\begin{cases} \left(\int_a^b g_1(x)g_1(x)dx\right)\alpha_1 + \ldots + \left(\int_a^b g_M(x)g_1(x)dx\right)\alpha_M & = \int_a^b f(x)g_1(x)dx, \\ \left(\int_a^b g_1(x)g_2(x)dx\right)\alpha_1 + \ldots + \left(\int_a^b g_M(x)g_2(x)dx\right)\alpha_M & = \int_a^b f(x)g_2(x)dx, \\ & \vdots \\ \left(\int_a^b g_1(x)g_M(x)dx\right)\alpha_1 + \ldots + \left(\int_a^b g_M(x)g_M(x)dx\right)\alpha_M & = \int_a^b f(x)g_M(x)dx, \end{cases}$$

que é um sistema linear com M equações e incógnitas $\alpha_1, \ldots, \alpha_M$.

Equações Normais

Alternativamente, podemos escrever

$$\begin{cases} \left(\int_a^b g_1(x)g_1(x)dx\right)\alpha_1 + \ldots + \left(\int_a^b g_M(x)g_1(x)dx\right)\alpha_M & = \int_a^b f(x)g_1(x)dx, \\ \left(\int_a^b g_1(x)g_2(x)dx\right)\alpha_1 + \ldots + \left(\int_a^b g_M(x)g_2(x)dx\right)\alpha_M & = \int_a^b f(x)g_2(x)dx, \\ & \vdots \\ \left(\int_a^b g_1(x)g_M(x)dx\right)\alpha_1 + \ldots + \left(\int_a^b g_M(x)g_M(x)dx\right)\alpha_M & = \int_a^b f(x)g_M(x)dx, \end{cases}$$

que é um sistema linear com M equações e incógnitas $\alpha_1, \ldots, \alpha_M$.

O sistema linear acima é chamado sistema das **equações normais.**

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

$$\mathbf{A}\boldsymbol{\alpha} = \mathbf{b},$$

em que $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{M \times M}$, $\boldsymbol{\alpha} = (\alpha_j) \in \mathbb{R}^M$ e $\mathbf{b} = (b_i) \in \mathbb{R}^M$, com

$$a_{ij} = \langle g_i, g_j \rangle = \int_a^b g_i(x) g_j(x) dx,$$

e

$$b_i = \langle f, g_i \rangle = \int_a^b f(x) g_i(x) dx,$$

para todo $i, j = 1, \ldots, M$.

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

$$\mathbf{A}\boldsymbol{\alpha} = \mathbf{b},$$

em que $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{M \times M}$, $\boldsymbol{\alpha} = (\alpha_j) \in \mathbb{R}^M$ e $\mathbf{b} = (b_i) \in \mathbb{R}^M$, com

$$a_{ij} = \langle g_i, g_j \rangle = \int_a^b g_i(x) g_j(x) dx,$$

e

$$b_i = \langle f, g_i \rangle = \int_a^b f(x) g_i(x) dx,$$

para todo $i, j = 1, \ldots, M$.

Pode-se mostrar que a solução das equações normais, quando **A** é não-singular, é o mínimo global de $J(\alpha_1, \ldots, \alpha_M)$.

Exemplo 1

Encontre a reta que melhor aproxima $f(x) = 4x^3$ em $[0, 1]$.

Exemplo 1

Encontre a reta que melhor aproxima $f(x) = 4x^3$ em $[0, 1]$.

Resposta: A reta que melhor se aproxima é formulada como o problema de quadrados mínimos $\varphi \approx f$ em que

$$\varphi(x) = \alpha_1 x + \alpha_2 = \alpha_1 g_1(x) + \alpha_2 g_(x),$$

em que $g_1(x) = x$ e $g_2(x) = 1$ para todo $x \in [0, 1]$.

Exemplo 1

Encontre a reta que melhor aproxima $f(x) = 4x^3$ em $[0, 1]$.

Resposta: A reta que melhor se aproxima é formulada como o problema de quadrados mínimos $\varphi \approx f$ em que

$$\varphi(x) = \alpha_1 x + \alpha_2 = \alpha_1 g_1(x) + \alpha_2 g_(x),$$

em que $g_1(x) = x$ e $g_2(x) = 1$ para todo $x \in [0, 1]$.
Portanto,

$$a_{11} = \langle g_1, g_1 \rangle = \int_a^b x^2 dx = \frac{1}{3}x^3\Big|_0^1 = \frac{1}{3},$$

$$a_{12} = \langle g_1, g_2 \rangle = \int_a^b x dx = \frac{1}{2}x^2\Big|_0^1 = \frac{1}{2} = a_{21},$$

$$a_{22} = \langle g_2, g_2 \rangle = \int_a^b 1 dx = x|_0^1 = 1.$$

Além disso,

$$b_1 = \langle f, g_1 \rangle = \int_a^b 4x^4 dx = \frac{4}{5}x^5 \Big|_0^1 = \frac{4}{5},$$
$$b_2 = \langle f, g_2 \rangle = \int_a^b 4x^3 dx = \frac{4}{4}x^4 \Big|_0^1 = 1.$$

Dessa forma, temos o sistema linear

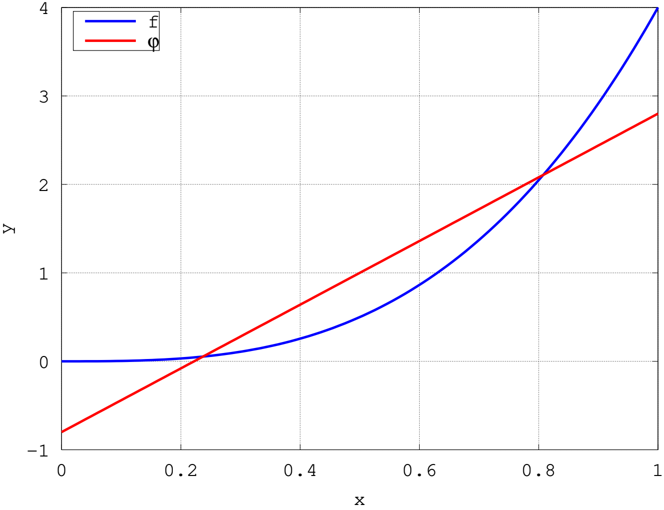
$$\underbrace{\begin{bmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}}_{\boldsymbol{\alpha}} = \underbrace{\begin{bmatrix} \frac{4}{5} \\ 1 \end{bmatrix}}_{\mathbf{b}},$$

cuja solução é

$$\boldsymbol{\alpha}^* = \begin{bmatrix} \frac{18}{5} & -\frac{4}{5} \end{bmatrix}^T.$$

Logo, a reta que melhor se aproxima de $f(x) = 4x^3$ em $[0, 1]$ é

$$\varphi(x) = \frac{18}{5}x - \frac{4}{5}.$$



Considerações Finais

Na aula de hoje, apresentamos duas variações do método dos quadrados mínimos:

Considerações Finais

Na aula de hoje, apresentamos duas variações do método dos quadrados mínimos:

- Caso discreto em $\mathbb{R}^N$ – usado para encontrar $\varphi : \mathbb{R}^N \to \mathbb{R}$, com

$$\varphi(\mathbf{x}) = \alpha_1 g_1(\mathbf{x}) + \alpha_2 g_2(\mathbf{x}) + \ldots + \alpha_M g_M(\mathbf{x}),$$

  que melhor se ajusta a uma tabela, i.e, $\varphi(\mathbf{x}_k) \approx y_k$, $k = 1, \ldots, K$.

Considerações Finais

Na aula de hoje, apresentamos duas variações do método dos quadrados mínimos:

- Caso discreto em $\mathbb{R}^N$ – usado para encontrar $\varphi : \mathbb{R}^N \to \mathbb{R}$, com
$$\varphi(\mathbf{x}) = \alpha_1 g_1(\mathbf{x}) + \alpha_2 g_2(\mathbf{x}) + \ldots + \alpha_M g_M(\mathbf{x}),$$
que melhor se ajusta a uma tabela, i.e, $\varphi(\mathbf{x}_k) \approx y_k$, $k = 1, \ldots, K$.
- Caso contínuo – usado para encontrar uma função
$$\varphi(x) = \alpha_1 g_1(x) + \alpha_2 g_2(x) + \ldots + \alpha_M g_M(x),$$
que melhor se aproxima de uma certa função f em $[a, b]$.

Considerações Finais

Na aula de hoje, apresentamos duas variações do método dos quadrados mínimos:

- Caso discreto em $\mathbb{R}^N$ – usado para encontrar $\varphi : \mathbb{R}^N \to \mathbb{R}$, com
$$\varphi(\mathbf{x}) = \alpha_1 g_1(\mathbf{x}) + \alpha_2 g_2(\mathbf{x}) + \ldots + \alpha_M g_M(\mathbf{x}),$$
que melhor se ajusta a uma tabela, i.e, $\varphi(\mathbf{x}_k) \approx y_k$, $k = 1, \ldots, K$.
- Caso contínuo – usado para encontrar uma função
$$\varphi(x) = \alpha_1 g_1(x) + \alpha_2 g_2(x) + \ldots + \alpha_M g_M(x),$$
que melhor se aproxima de uma certa função f em $[a, b]$.

Em ambos os casos, os coeficientes $\alpha_1, \ldots, \alpha_M$ são obtidos resolvendo um sistema linear $\mathbf{A}\boldsymbol{\alpha} = \mathbf{b}$.

Muito grato pela atenção!