MAP 3121 - Métodos Numéricos e Aplicações

Método dos Mínimos Quadrados
Caso geral e discreto - Aula I

Depto. Matemática Aplicada
Instituto de Matemática e Estatística
Universidade de São Paulo
São Paulo – SP

Abordaremos o problema de **aproximar** uma função f por outra g pertencente a uma **família** de funções G **previamente** escolhida.

- Ao **aproximar** f por g em G introduzimos um erro chamado **resíduo**:

$$r(x) = f(x) - g(x)$$

- É necessário **minimizar** r para que uma **boa** aproximação g seja **obtida**.

Mínimos Quadrados

Exigimos que a seguinte quantidade seja **mínima**:

$$\sum_x r(x)^2. \tag{1}$$

- Note que minimizar $\sum_x r(x)$ não é adequado pelo **cancelamento** de erros positivos e negativos.
- Poderíamos usar $\sum_x |r(x)|$. Por questões **geométricas/diferenciabilidade** escolhemos (1).

Abaixo vemos que o **método de mínimos quadrados** está associado a espaços vetoriais com **produto interno**. Na sequência estudamos métodos que são na verdade **casos particulares** deste resultado.

Formulação geométrica

Seja F um espaço vetorial real com **produto interno** $\langle \cdot, \cdot \rangle$

$$(f, g) \in F \times F \mapsto \langle f, g \rangle \in \mathbb{R},$$

$f \in F$ e $G \subset F$ um **subespaço** gerado pelos vetores g_0, g_1, ..., g_m.

- Ache $g \in G$ que **melhor aproxime** f, ou seja, $g \in G$ tal que o erro quadrático

$$E(f, g) := \langle f - g, f - g \rangle = \|f - g\|_F^2$$

seja o **menor possível**.

Exemplos de **espaços** com produto interno:

a) $\mathbb{R}^N = \{(x_1, x_2, ..., x_N) \ : \ x_i \in \mathbb{R}, \quad i = 1, ..., N\}$ com $\langle x, y\rangle = \sum_{i=1}^{N} x_i y_i$.
Assim:

$$E(x,y) = \sum_{i=1}^{N} (x_i - y_i)^2, \quad x, y \in \mathbb{R}^N.$$

Recordamos que este produto interno é muitas vezes chamado **produto escalar** e define a norma **Euclidiana** em $\mathbb{R}^N$ por $\|x\| = \sqrt{\langle x, x\rangle} = \sqrt{\sum_{i=1}^{N} x_i^2}$.

b) Seja $F = \{\alpha_0\, 1 + \alpha_1\, x + ... + \alpha_n\, x^n \ : \ \alpha_i \in \mathbb{R}, \quad i = 0, 1, ..., n\}$ o espaço dos **polinômios** de grau menor ou igual a n com

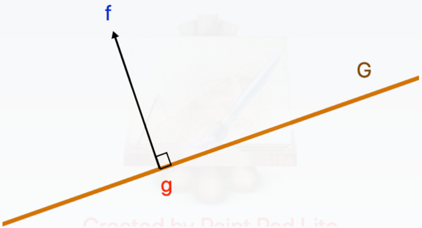
$$\langle f, g\rangle = \int_{-1}^{1} f(x)\, g(x)\, dx.$$

Assim:

$$E(f,g) = \int_{-1}^{1} \left\{ f(x) - g(x) \right\}^2 dx.$$

Ortogonalidade

Em **espaços vetorias** com produto interno podemos trabalhar com **ortogonalidade** e também minimizar distância.



O vetor de $g \in G$ que melhor **aproxima** f é o vetor tal que $f - g$ é **ortogonal** a G. Isto ocorre se, e somente se, $\langle f - g, g_j \rangle = 0$ para todo $j = 0, \ldots, m$ onde[1]

$$G = [g_0, g_1, \ldots, g_m] .$$

[1]Lembre-se que o subespaço G é gerado pelos vetores $g_0, g_1, \ldots, g_m$.

Como $g \in G$, existem números reais a_0, ..., a_m tais que

$$g = a_0 g_0 + ... + a_m g_m.$$

Assim, esses **números** devem resolver o seguinte **sistema**[2]

$$\begin{pmatrix} \langle g_0, g_0 \rangle & \cdots & \langle g_0, g_m \rangle \\ \vdots & \ddots & \vdots \\ \langle g_m, g_0 \rangle & \cdots & \langle g_m, g_m \rangle \end{pmatrix} \begin{pmatrix} a_0 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} \langle g_0, f \rangle \\ \vdots \\ \langle g_m, f \rangle \end{pmatrix} \qquad (2)$$

já que **para todo** j

$$\begin{aligned} 0 &= \langle f - g, g_j \rangle = \langle f - (a_0 g_0 + ... + a_m g_m), g_j \rangle \\ &= \langle f, g_j \rangle - a_0 \langle g_0, g_j \rangle - ... - a_m \langle g_m, g_j \rangle \\ &= \langle f, g_j \rangle - (\langle g_0, g_j \rangle, ..., \langle g_m, g_j \rangle) \cdot (a_0, ..., a_m). \end{aligned}$$

i) Particularmente **interessante** é o caso em que $\langle g_i, g_j \rangle = 0$ para $i \neq j$. Assim

$$a_j = \frac{\langle g_j, f \rangle}{\langle g_j, g_j \rangle} \quad j = 0, 1, ..., m.$$

ii) **Note** que a matriz que define o sistema é **simétrica**.

[2] Este sistema é muitas vezes chamado **sistema normal**.

Mínimos quadrados discreto

Chamamos o método de método dos mínimos quadrados **discreto** quando o espaço vetorial F **assume** a forma particular $F = \mathbb{R}^N$ para algum $N \in \mathbb{N}$.

O problema

Dados N pontos do gráfico de uma função

$$\{(x_1, y_1), ..., (x_N, y_N)\} \subset \mathbb{R}^2$$

x_j distintos, $y_j = f(x_j)$ e $m+1$ **funções**

$$g_0, ..., g_m : I \to \mathbb{R}$$

em que $I \subset \mathbb{R}$ é um intervalo que contém x_1, ..., x_N. **Encontre** a_0, a_1, ..., a_m tais que

$$y(x) = a_0 g_0(x) + a_1 g_1(x) + ... + a_m g_m(x)$$

seja a melhor função que **aproxime** esses pontos.

Para **resolvermos** este problema, definimos o seguinte **erro** quadrático:

$$E(a_0, a_1, ..., a_m) = \sum_{i=1}^{N} \left| y_i - (a_0 g_0(x_i) + a_1 g_1(x_i) + ... + a_m g_m(x_i)) \right|^2 .$$

Definimos também os **vetores**

$$f = (y_1, ..., y_N) \quad \text{e} \quad g_j = (g_j(x_1), ..., g_j(x_N)), \quad j = 0, 1, ..., m$$

e usamos o seguinte **produto interno** em $\mathbb{R}^N$

$$\langle z, w \rangle = \sum_{i=1}^{N} z_i w_i, \quad z, w \in \mathbb{R}^N.$$

Se $g = a_0 g_0 + ... + a_m g_m$ **temos** que

$$E\left(a_0, a_1, ..., a_m\right) = \langle f - g, f - g\rangle \,.$$

Além disso, os **valores** a_0,a_1, ..., a_m que **minimizam** E são soluções do sistema

$$\left(\begin{array}{ccc} \sum_{j=1}^{N} g_0\left(x_j\right) g_0\left(x_j\right) & \dots & \sum_{j=1}^{N} g_0\left(x_j\right) g_m\left(x_j\right) \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^{N} g_m\left(x_j\right) g_0\left(x_j\right) & \dots & \sum_{j=1}^{N} g_m\left(x_j\right) g_m\left(x_j\right) \end{array}\right)\left(\begin{array}{c} a_0 \\ \vdots \\ a_m \end{array}\right)$$

$$= \left(\begin{array}{c} \sum_{j=1}^{N} g_0\left(x_j\right) y_j \\ \vdots \\ \sum_{j=1}^{N} g_m\left(x_j\right) y_j \end{array}\right)$$

já que

$$\langle g_j, g_k\rangle = \sum_{i=1}^{N} g_j(x_i)\, g_k(x_i), \quad 0 \le j, k \le m.$$

1) O **sinal** de um osciloscópio corresponde a superposição de efeitos oscilatório e crescente. Nestas condições o **aproximamos** por uma função do tipo

$$g(x) = ax + b\cos x.$$

Medindo alguns valores deste sinal obtemos

x	0	1.5	3.0	4.5	6.0
$f(x)$	1.0	1.57	2.0	4.3	7.0

Usando notações **anteriores** temos $g_0(x) = x$ e $g_1(x) = \cos x$.

$$f = (1.0, 1.57, 2.0, 4.3, 7.0), \quad g_0 = (0, 1.5, 3.0, 4.5, 6.0)$$
$$\text{e} \;\; g_1 = (\cos(0), \cos(1.5), \cos(3.0), \cos(4.5), \cos(6.0))$$

que nos leva ao **sistema**

$$\begin{pmatrix} \langle g_0, g_0\rangle & \langle g_0, g_1\rangle \\ \langle g_1, g_0\rangle & \langle g_1, g_1\rangle \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \langle g_0, f\rangle \\ \langle g_1, f\rangle \end{pmatrix}$$

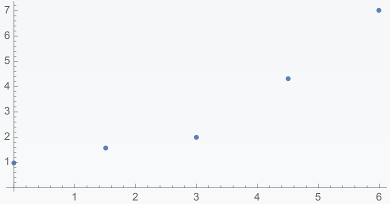
que é igual a

$$\begin{pmatrix} 67.5 & 1.949 \\ 1.949 & 2.952 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 69.71 \\ 4.946 \end{pmatrix}$$

cuja **solução** é: $(a, b) = (1.003, 1.013)$.

Observações dadas:

x	0	1.5	3.0	4.5	6.0
$f(x)$	1.0	1.57	2.0	4.3	7.0



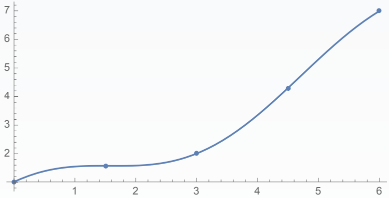


Figura: Gráfico da função $g(x) = 1.003x + 1.013\cos x$.

2) **Regressão linear.** Determine a **reta** que melhor ajusta a função abaixo segundo o método de **mínimos quadrados**

x	0	1	2	3	4
$f(x)$	0	1	1	4	4

Temos $g_0(x) = 1$ e $g_1(x) = x$.

$$f = (0,1,1,4,4), \quad g_0 = (1,1,1,1,1)$$
$$\text{e } \ g_1 = (0,1,2,3,4)$$

que nos leva ao **sistema**

$$\begin{pmatrix} \langle g_0, g_0\rangle & \langle g_0, g_1\rangle \\ \langle g_1, g_0\rangle & \langle g_1, g_1\rangle \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \langle g_0, f\rangle \\ \langle g_1, f\rangle \end{pmatrix}$$

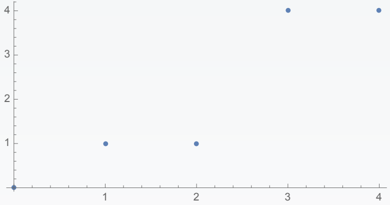
que é igual a

$$\begin{pmatrix} 5 & 10 \\ 10 & 30 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 10 \\ 31 \end{pmatrix}$$

cuja **solução** é: $(a,b) = (-1/5, 11/10)$.

Observações dadas:

x	0	1	2	3	4
$f(x)$	0	1	1	4	4



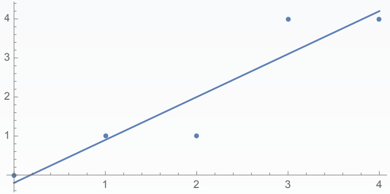


Figura: Gráfico da função $g(x) = -1/5 + (11/10)x$.

Note que a **Regressão linear** do conjunto

$$\{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\} \subset \mathbb{R}^2$$

é o caso geral previamente discutido **assumindo**

$$m = 1, \quad g_0(x) = 1 \quad \text{e} \quad g_1(x) = x.$$

Logo, é **equivalente** a resolver o seguinte **sistema** linear

$$\begin{pmatrix} N & \sum_{i=1}^N x_i \\ \sum_{i=1}^N x_i & \sum_{i=1}^N x_i^2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^N y_i \\ \sum_{i=1}^N x_i y_i \end{pmatrix}$$

já que se

$$\langle g_j, g_k \rangle = \sum_{i=1}^N g_j(x_i)\, g_k(x_i), \quad 0 \leq j, k \leq 2$$

$$g_0 = (1, 1, \ldots, 1), \quad g_1 = (x_1, x_2, \ldots, x_N) \quad \text{e} \quad f = (y_1, y_2, \ldots, y_N).$$

então

$$\langle g_0, g_0 \rangle = \sum_{i=1}^N 1^2, \quad \langle g_0, g_1 \rangle = \sum_{i=1}^N x_i \quad \text{e} \quad \langle g_1, g_1 \rangle = \sum_{i=1}^N x_i^2.$$