Método dos Mínimos Quadrados - Modelos não Lineares

O Método dos Mínimos Quadrados foi introduzido a partir de dois ingredientes:

- **Modelo**
  Consiste na descrição de uma variável dependente y em função de uma variável independente x utilizando uma combinação linear de funções $\{g_0(x), g_1(x), ..., g_m(x)\}$

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

- **Dados Empíricos**
  Consiste um um conjunto de observações de valores da variável dependente y em função de valores selecionados para a variável independente x.

k	x_k	y_k
1	x_1	y_1
2	x_2	y_2
$\vdots$	$\vdots$	$\vdots$
n	x_n	y_n

O objetivo do Método dos Mínimos Quadrados é selecionar valores para os parâmetros $\{a_0, a_1, ..., a_m\}$ de forma a obter o *melhor ajuste* do Modelo aos Dados Empíricos. Para isto consideramos

$$EQ(a_0, a_1, ..., a_m) =$$

$$\langle y - (a_0 g_0 + a_1 g_1 + ... + a_m g_m) \mid y - (a_0 g_0 + a_1 g_1 + ... + a_m g_m) \rangle =$$

onde $\langle \cdot \mid \cdot \rangle$ é um produto interno em $\mathbb{R}^n$.

A escolha de $\{a_0, a_1, ..., a_m\}$ correspondente ao *melhor ajuste* são os valores para os quais $EQ(a_0, a_1, ..., a_m)$ é mínimo, valores estes obtidos resolvendo o *sistema normal*:

$$\begin{pmatrix} \langle g_0 \mid g_0 \rangle & \cdots & \langle g_0 \mid g_m \rangle \\ \vdots & & \vdots \\ \langle g_m \mid g_0 \rangle & \cdots & \langle g_m \mid g_m \rangle \end{pmatrix} \begin{pmatrix} a_0 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} \langle g_0 \mid y \rangle \\ \vdots \\ \langle g_m \mid y \rangle \end{pmatrix}$$

Muitas vezes os parametros a serem escolhidos não aparecem de forma linear no modelo, como nos exemplos abaixo:

$$y = f(x) = a_0 e^{a_1 x}; \quad y = f(x) = \frac{a_0 + a_1 x}{1 + a_2 x}$$

Neste casos se considerarmos, no primeiro exemplo:

$$EQ(a_0, a_1) = \sum_{k=1}^{n} [y_k - a_0 e^{a_1 x_k}]^2$$

ou no segundo exemplo:

$$EQ(a_0, a_1, a_2) = \sum_{k=1}^{n} [y_k - \frac{a_0 + a_1 x_k}{1 + a_2 x_k}]^2$$

as equações:

$$\frac{\partial EQ}{\partial a_0}(a_0, a_1) = 0; \quad \frac{\partial EQ}{\partial a_1}(a_0, a_1) = 0$$

no primeiro exemplo e, no segundo exemplo:

$$\frac{\partial EQ}{\partial a_0}(a_0, a_1, a_2) = 0; \quad \frac{\partial EQ}{\partial a_1}(a_0, a_1, a_2) = 0; \frac{\partial EQ}{\partial a_2}(a_0, a_1, a_2) = 0$$

são **não** lineares tornando a escolha dos valores dos parâmetros correspondentes ao melhor ajuste um problema muito mais complexo.

Esta dificuldade em alguns casos pode ser contornada. Definindo no primeiro exemplo:

$$EQ(\tilde{a}_0, a_1) = \sum_{k=1}^{n} [\ln(y_k) - (\tilde{a}_0 + a_1 x_k)]^2; \quad \tilde{a}_0 = \ln(a_0)$$

isto porque $y_k \approx a_0 e^{a_1 x_k} \iff \ln(y_k) \approx \ln(a_0) + a_1 x_k$

no segundo exemplo podemos considerar:

$$EQ(a_0, a_1, a_2) = \sum_{k=1}^{n} [y_k\,(1 + a_2 x_k) - (a_0 + a_1 x_k)]^2$$

isto porque $y_k \approx \frac{a_0 + a_1 x_k}{1 + a_2 x_k} \iff y_k\,(1 + a_2 x_k) \approx a_0 + a_1 x_k$

Como em ambos os exemplos, $EQ(a_0, a_1)$ e $EQ(a_0, a_1, a_2)$ são funções quadráticas, suas derivadas são funções lineares resultando em um *sistema normal* linear.

No contexto de um problema de projeção, no primeiro exemplo, consideramos em $\mathbb{R}^n$ os vetores:

$$v^0 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}; \quad v^1 = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{e} \quad u = \begin{pmatrix} \ln(y_1) \\ \ln(y_2) \\ \vdots \\ \ln(y_n) \end{pmatrix}$$

e os valores de a_0 e a_1 ficam determinados pela projeção ortogonal (produto interno euclideano) de u no subespaço gerado pelas combinações lineares de v_0 e v_1

No segundo exemplo, consideramos em $\mathbb{R}^n$ os vetores:

$$v^0 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}; \quad v^1 = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}; v^2 = \begin{pmatrix} -y_1 x_1 \\ -y_2 x_2 \\ \vdots \\ -y_n x_n \end{pmatrix} \quad \text{e} \quad u = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

e os valores de a_0, a_1 e a_2 ficam determinados pela projeção ortogonal (produto interno euclideano) de u no subespaço gerado pelas combinações lineares de v_0, v_1 e v_2.

Para entender o sinal -"nas componentes de v_2, observe que:

$$[y_k\,(1 + a_2 x_k) - (a_0 + a_1 x_k)]^2 = [y_k - (a_0 \cdot 1 + a_1 x_k + a_2\,(-x_k y_k))]^2$$