Tópico 6 - Ajuste de Curvas por Mínimos Quadrados

João B. Florindo

Instituto de Matemática, Estatística e Computação Científica
Universidade Estadual de Campinas - Brasil
jbflorindo@ime.unicamp.br

Outline

Introdução

Dados tabelados podem ser aproximados por funções contínuas em dois sentidos: ajuste e interpolação.

Na interpolação escolhe-se uma função que coincide exatamente nos pontos tabelados. Ideal para estimar valores intermediários não tabelados

Já no ajuste uma curva é escolhida com o objetivo de se aproximar do aspecto global dos dados tabelados e portanto esta curva não precisa coincidir nos pontos tabelados. Ideal para estimar valores fora do intervalo tabelado (extrapolação/regressão)

O problema do ajuste de curvas pode ser categorizado em dois casos: *discreto* e *contínuo*

Introdução-Discreto

No caso discreto, dados os valores fornecidos por uma tabela de pontos

$$(x_1, f(x_1)), (x_2, f(x_2)), \dots, (x_m, f(x_m)), \quad x_1, x_2, \dots, x_m \in [a, b],$$

escolher funções $g_1(x), g_2(x), \dots, g_n(x)$, contínuas em $[a, b]$, e obter constantes $\alpha_1, \alpha_2, \dots, \alpha_n$ tais que

$$\varphi(x) = \alpha_1 g_1(x) + \alpha_1 g_1(x) + \cdots + \alpha_n g_n(x) \approx f(x)$$

Este modelo é linear porque os coeficientes são escalares, mas as $g_k(x)$ não precisam ser lineares. Estas funções podem ser escolhidas olhando-se para o gráfico dos pontos tabelados (*diagrama de dispersão*).

Introdução-Discreto

x	−1.0	−0.75	−0.6	−0.5	−0.3	0	0.2	0.4	0.5	0.7	1.0
f(x)	2.05	1.153	0.45	0.4	0.5	0	0.2	0.6	0.512	1.2	2.05

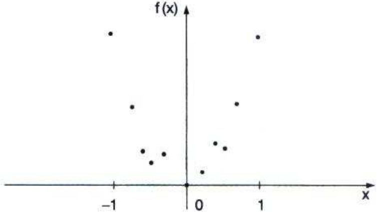

Figura: Fonte: Ruggiero&Lopes

Neste caso uma parábola de expressão geral αx^2 parece ser a função ideal. Resta encontrar o melhor α.

Mínimos quadrados minimiza o desvio $(f(x_i) - \varphi(x_i)), i = 1, 2, \ldots, m$.

Introdução-Contínuo

No caso contínuo, em vez de dados tabelados temos uma função contínua $f(x) : [a, b] \rightarrow \mathbb{R}$ e novamente queremos obter a aproximação em $[a, b]$

$$\varphi(x) = \alpha_1 g_1(x) + \alpha_1 g_1(x) + \cdots + \alpha_n g_n(x) \approx f(x)$$

O critério dos mínimos quadrados no caso consiste em minimizar o módulo da área sob a curva de $\varphi(x) - f(x)$

Outline

Mínimos Quadrados-Discreto

Vamos considerar que número m de pontos tabelados é maior ou igual ao número n de funções $g_k(x)$ usadas.

Podemos minimizar o desvio em termos de valores absolutos:

$$\max_{1 \leq i \leq m} |(f(x_i) - \varphi(x_i))|$$

mas esta abordagem, chamada de *minmax* não é tão trivial.

Já os mínimos quadrados, como o nome sugere, visam minimizar a soma dos quadrados dos desvios, expressa pela função $F(\alpha_1, \alpha_2, \ldots, \alpha_n)$:

$$F(\alpha_1, \alpha_2, \ldots, \alpha_n) = \sum_{k=1}^{m} [f(x_k) - \varphi(x_k)]^2 =$$

$$= \sum_{k=1}^{m} [f(x_k) - \alpha_1 g_1(x_k) - \alpha_2 g_2(x_k) - \cdots - \alpha_n g_n(x_k)]^2$$

Mínimos Quadrados-Discreto

Como cada termo dentro do somatório é não-negativo, cada parcela individualmente também deve ser minimizada.

Note que caso da interpolação, o modelo linear se encaixa exatamente aos dados tabelados faz o mínimo ser zero e portanto a interpolação é um caso especial de mínimos quadrados.

Sabemos do Cálculo que o mínimo de F se obtém igualando-se as derivadas parciais a zero:

$$\left.\frac{\partial F}{\partial \alpha_j}\right|_{\alpha_1,\alpha_2,\ldots,\alpha_n} = 0, j = 1, 2, \ldots, n$$

e estas derivadas são dadas por

$$\left.\frac{\partial F}{\partial \alpha_j}\right|_{\alpha_1,\alpha_2,\ldots,\alpha_n} = 2\sum_{k=1}^{m}[f(x_k)-\alpha_1 g_1(x_k)-\alpha_2 g_2(x_k)-\cdots-\alpha_n g_n(x_k)][-g_j(x_k)].$$

Mínimos Quadrados-Discreto

Substituindo na restrição acima:

$$\sum_{k=1}^{m}[f(x_k)-\alpha_1 g_1(x_k)-\alpha_2 g_2(x_k)-\cdots-\alpha_n g_n(x_k)][g_j(x_k)] = 0, j = 1, 2, \ldots, n$$

Isso leva ao seguinte sistema de equações:

$$\begin{cases} \sum_{k=1}^{m}[f(x_k) - \alpha_1 g_1(x_k) - \alpha_2 g_2(x_k) - \cdots - \alpha_n g_n(x_k)]g_1(x_k) = 0 \\ \sum_{k=1}^{m}[f(x_k) - \alpha_1 g_1(x_k) - \alpha_2 g_2(x_k) - \cdots - \alpha_n g_n(x_k)]g_2(x_k) = 0 \\ \vdots \\ \sum_{k=1}^{m}[f(x_k) - \alpha_1 g_1(x_k) - \alpha_2 g_2(x_k) - \cdots - \alpha_n g_n(x_k)]g_n(x_k) = 0 \end{cases}$$

ou ainda

$$\begin{cases} \left[\sum_{k=1}^{m} g_1(x_k)g_1(x_k)\right]\alpha_1 + \cdots + \left[\sum_{k=1}^{m} g_n(x_k)g_1(x_k)\right]\alpha_n = \sum_{k=1}^{m} f(x_k)g_1(x_k) \\ \left[\sum_{k=1}^{m} g_1(x_k)g_2(x_k)\right]\alpha_1 + \cdots + \left[\sum_{k=1}^{m} g_n(x_k)g_2(x_k)\right]\alpha_n = \sum_{k=1}^{m} f(x_k)g_2(x_k) \\ \vdots \\ \left[\sum_{k=1}^{m} g_1(x_k)g_n(x_k)\right]\alpha_1 + \cdots + \left[\sum_{k=1}^{m} g_n(x_k)g_n(x_k)\right]\alpha_n = \sum_{k=1}^{m} f(x_k)g_n(x_k) \end{cases}$$

O sistema acima é chamado de *sistema de equações normais*

Mínimos Quadrados-Discreto

Escrevendo na forma matricial $A\alpha = b$:

$$\left\{ \begin{array}{l} a_{11}\alpha_1 + a_{12}\alpha_2 + \cdots + a_{1n}\alpha_n = b_1 \\ a_{21}\alpha_1 + a_{22}\alpha_2 + \cdots + a_{2n}\alpha_n = b_2 \\ \vdots \\ a_{n1}\alpha_1 + a_{n2}\alpha_2 + \cdots + a_{nn}\alpha_n = b_n \end{array} \right.$$

tal que

$$a_{ij} = a_{ji} = \sum_{k=1}^{m} g_j(x_k) g_i(x_k) \text{ e } b_i = \sum_{k=1}^{m} f(x_k) g_i(x_k)$$

Lembrando o conceito de *produto escalar* entre 2 vetores x e $y \in \mathbb{R}^m$:

$$< x, y >= \sum_{i=1}^{m} x_i y_i$$

podemos re-escrever em notação compacta:

$$a_{ij} =< g_i, g_j > \text{ e } b_i =< f, g_i >$$

Mínimos Quadrados-Discreto

Na prática, é comum que os vetores $g_1, g_2, \ldots, g_n$ sejam ortogonais entre si, ou seja,

$$< g_i, g_j > = \begin{cases} 0 \text{ se } i \neq j \\ \neq 0 \text{ se } i = j \end{cases}$$

fazendo com que a matriz A seja diagonal e simplificando sobremaneira o cálculo da solução α.

Mínimos Quadrados-Polinômios de Gram

Os polinômios de Gram $\{P_{i,m}\}_{i=0}^{m}$ são ortogonais em conjuntos de pontos equidistantes $x_i = -1 + \frac{2i}{m}$. Por exemplo,

$$P_{0,0} = 1, \quad P_{0,1} = x, \quad P_{0,2} = x^2 - \frac{1}{2}$$

sobre o conjunto $X_5 = [-1, -1/2, 0, 1/2, 1]$.

$$\overline{g}_0 = P_{0,0}(x_i) = [1 \quad 1 \quad 1 \quad 1 \quad 1]^T$$

$$\overline{g}_1 = P_{0,1}(x_i) = [-1 \quad -\frac{1}{2} \quad 0 \quad \frac{1}{2} \quad 1]^T$$

$$\overline{g}_2 = P_{0,2}(x_i) = [\frac{1}{2} \quad -\frac{1}{4} \quad -\frac{1}{2} \quad -\frac{1}{4} \quad \frac{1}{2}]^T$$

Mínimos Quadrados-Polinômios de Gram

Produtos escalares:

$$< \overline{g}_0, \overline{g}_0 >= 5 \neq 0$$

$$< \overline{g}_0, \overline{g}_1 >= 1(-1) + 1(-\frac{1}{2}) + 1(0) + 1(\frac{1}{2}) + 1(1) = 0$$

$$< \overline{g}_0, \overline{g}_2 >= 1(\frac{1}{2}) + 1(-\frac{1}{4}) + 1(\frac{-1}{2}) + 1(\frac{1}{2}) = 0$$

$$\vdots$$

Mínimos Quadrados-Exemplo

Vamos resolver o exemplo da introdução usando uma parábola "completa": $\varphi(x) = \alpha_1 x^2 + \alpha_2 x + \alpha_3$.

Relembrando a tabela:

x	-1.00	-0.75	-0.60	-0.50	-0.30	0.00	0.20	0.40	0.50	0.70	1.00
y	2.05	1.15	0.45	0.40	0.50	0.00	0.20	0.60	0.51	1.20	2.05

Funções:

$$g_1(x) = x^2, \quad g_2(x) = x, \quad g_3(x) = 1$$

e os respectivos vetores:

g_1	1.00	0.56	0.36	0.25	0.09	0.00	0.04	0.16	0.25	0.49	1.0
g_2	-1.00	-0.75	-0.60	-0.50	-0.30	0.00	0.20	0.40	0.50	0.70	1.0
g_3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.0

A partir dos produtos escalares, obtemos:

$$\begin{array}{ccc} a_{11} =< g_1, g_1 >= 2.85, & a_{12} =< g_1, g_2 >= -0.25, & a_{13} =< g_1, g_3 >= 4.20, \\ a_{21} =< g_2, g_1 >= -0.25, & a_{22} =< g_2, g_2 >= 4.20, & a_{23} =< g_2, g_3 >= -0.35, \\ a_{31} =< g_3, g_1 >= 4.20, & a_{32} =< g_3, g_2 >= -0.35, & a_{33} =< g_3, g_3 >= 11, \end{array}$$

$$b_1 =< y, g_1 >= 5.87, \quad b_2 =< y, g_2 >= -0.11, \quad b_3 =< y, g_3 >= 9.11.$$

Mínimos Quadrados-Exemplo

Equações normais:

$$\begin{bmatrix} 2.85 & -0.25 & 4.20 \\ -0.25 & 4.20 & -0.35 \\ 4.20 & -0.35 & 11.00 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 5.87 \\ -0.11 \\ 9.11 \end{bmatrix}$$

cuja solução é

$$\alpha* = [1.94 \quad 0.10 \quad 0.09]^T$$

fazendo então com que a parábola que melhor ajusta os dados seja

$$\varphi(x) = 1.94x^2 + 0.10x + 0.09.$$

Erro da aproximação (soma dos desvios quadrados):

$$\sum_{k=1}^{m} ((1.94x_k^2 + 0.10x_k + 0.09) - y_k)^2 = 0.24$$

Mínimos Quadrados-Ajuste por reta

O ajuste por uma reta é um caso tão comum que merece atenção especial e a partir dos dados tabelados em pares $(x_i, y_i), i = 1, 2, \dots, n$ podemos deduzir uma fórmula genérica para os parâmetros. Temos então

$$\varphi(x) = \alpha_1 x + \alpha_2$$

$$g_1(x) = x, \qquad g_2(x) = 1$$

Produtos escalares:

$$a_{11} =< g_1, g_1 >= \sum_{i=1}^{n} x_i^2$$

$$a_{12} = a_{21} =< g_1, g_2 >= \sum_{i=1}^{n} x_i$$

$$a_{22} =< g_2, g_2 >= n$$

$$b_1 =< y, g_1 >= \sum_{i=1}^{n} x_i y_i, \quad b_2 =< y, g_2 >= \sum_{i=1}^{n} y_i.$$

Mínimos Quadrados-Ajuste por reta

Equações normais:

$$\begin{bmatrix} \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & n \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} x_i y_i \\ \sum_{i=1}^{n} y_i \end{bmatrix}$$

Por triangularização (ou substituição) pode-se verificar que um sistema 2×2 genérico tem solução

$$\alpha_2 = \frac{b_2 - b_1\left(\frac{a_{21}}{a_{11}}\right)}{a_{22} - a_{12}\left(\frac{a_{21}}{a_{11}}\right)} = \frac{b_2 a_{11} - b_1 a_{21}}{a_{22} a_{11} - a_{12} a_{21}}$$

$$\alpha_1 = \frac{b_1 - \alpha_2 a_{12}}{a_{11}} = \frac{b_1 - a_{12}\left(\frac{b_2 a_{11} - b_1 a_{21}}{a_{22} a_{11} - a_{12} a_{21}}\right)}{a_{11}} =$$

$$= \frac{b_1 a_{22} a_{11} - b_1 a_{12} a_{21} - b_2 a_{12} a_{11} + b_1 a_{12} a_{21}}{a_{11}^2 a_{22} - a_{11} a_{12} a_{21}}$$

Mínimos Quadrados-Ajuste por reta

No caso particular em que $a_{12} = a_{21}$:

$$\alpha_1 = \frac{b_1 a_{22} a_{11} - b_1 a_{12}^2 - b_2 a_{12} a_{11} + b_1 a_{12}^2}{a_{11}^2 a_{22} - a_{11} a_{12}^2} = \frac{b_1 a_{22} a_{11} - b_2 a_{12} a_{11}}{a_{11}^2 a_{22} - a_{11} a_{12}^2} =$$

$$= \frac{b_1 a_{22} - b_2 a_{12}}{a_{11} a_{22} - a_{12}^2}$$

Fazendo as substituições devidas:

$$\alpha_1 = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \left(\sum_{i=1}^n x_i^2\right) - \left(\sum_{i=1}^n x_i\right)^2}$$

$$\alpha_2 = \frac{\sum_{i=1}^n y_i \sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i y_i \sum_{i=1}^n x_i}{n \sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i \sum_{i=1}^n x_i} = \frac{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i - \sum_{i=1}^n x_i y_i \sum_{i=1}^n x_i}{n \left(\sum_{i=1}^n x_i^2\right) - \left(\sum_{i=1}^n x_i\right)^2}$$

Mínimos Quadrados-Ajuste por reta

EXEMPLO:

x_i	y_i	x_i^2	$x_i y_i$
1	1.3	1	1.3
2	3.5	4	7.0
3	4.2	9	12.6
4	5.0	16	20.0
5	7.0	25	35.0
15	21.0	55	75.9

$$\alpha_1 = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\left(\sum_{i=1}^{n} x_i^2\right) - \left(\sum_{i=1}^{n} x_i\right)^2} = \frac{5 \cdot 75.9 - 15 \cdot 21.0}{5 \cdot 55 - 15^2} = 1.29$$

$$\alpha_2 = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i y_i \sum_{i=1}^{n} x_i}{n\left(\sum_{i=1}^{n} x_i^2\right) - \left(\sum_{i=1}^{n} x_i\right)^2} = \frac{55 \cdot 21.0 - 75.9 \cdot 15}{5 \cdot 55 - 15^2} = 0.33$$

Mínimos Quadrados-Ajuste por polinômio

Aproximar dados $(x_i, y_i), i = 1, 2, \ldots, m$ com polinômio

$$P_n(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

tal que $n < m - 1$.

Vamos minimizar

$$F(a_0, a_1, \ldots, a_n) = \sum_{i=1}^{m} (y_i - P_n(x_i))^2$$

$$= \sum_{i=1}^{m} y_i^2 - 2 \sum_{i=1}^{m} P_n(x_i) y_i + \sum_{i=1}^{m} (P_n(x_i))^2$$

$$= \sum_{i=1}^{m} y_i^2 - 2 \sum_{i=1}^{m} \left(\sum_{j=0}^{n} a_j x_i^j \right) y_i + \sum_{i=1}^{m} \left(\sum_{j=0}^{n} a_j x_i^j \right)^2$$

Mínimos Quadrados-Ajuste por polinômio

$$= \sum_{i=1}^{m} y_i^2 - 2\sum_{i=1}^{m}\left(\sum_{j=0}^{n} a_j x_i^j\right) y_i + \sum_{i=1}^{m}\left(\sum_{j=0}^{n} a_j x_i^j \cdot \sum_{j=0}^{n} a_j x_i^j\right)$$

$$= \sum_{i=1}^{m} y_i^2 - 2\sum_{i=1}^{m}\left(\sum_{j=0}^{n} a_j x_i^j\right) y_i + \sum_{i=1}^{m}\left(\sum_{j=0}^{n}\sum_{k=0}^{n} a_j a_k x_i^{j+k}\right)$$

$$= \sum_{i=1}^{m} y_i^2 - 2\sum_{j=0}^{n} a_j \left(\sum_{i=1}^{m} y_i x_i^j\right) + \sum_{j=0}^{n}\sum_{k=0}^{n} a_j a_k \left(\sum_{i=1}^{m} x_i^{j+k}\right)$$

Novamente zerando a derivada:

$$0 = \frac{\partial F}{\partial a_j} = -2\sum_{i=1}^{m} y_i x_i^j + 2\sum_{k=0}^{n} a_k \sum_{i=1}^{m} x_i^{j+k}$$

Mínimos Quadrados-Ajuste por polinômio

levando às equações normais seguintes:

$$\sum_{k=0}^{n} a_k \sum_{i=1}^{m} x_i^{j+k} = \sum_{i=1}^{m} y_i x_i^j$$

Estas podem ser convenientemente escritas para cada j como:

$$\begin{cases} a_0 \sum_{i=1}^m x_i^0 + a_1 \sum_{i=1}^m x_i^1 + a_2 \sum_{i=1}^m x_i^2 + \cdots + a_n \sum_{i=1}^m x_i^n = \sum_{i=1}^m y_i x_i^0 \\ a_0 \sum_{i=1}^m x_i^1 + a_1 \sum_{i=1}^m x_i^2 + a_2 \sum_{i=1}^m x_i^3 + \cdots + a_n \sum_{i=1}^m x_i^{n+1} = \sum_{i=1}^m y_i x_i^1 \\ \vdots \\ a_0 \sum_{i=1}^m x_i^n + a_1 \sum_{i=1}^m x_i^{n+1} + a_2 \sum_{i=1}^m x_i^{n+2} + \cdots + a_n \sum_{i=1}^m x_i^{2n} = \sum_{i=1}^m y_i x_i^n \end{cases}$$

Mínimos Quadrados-Ajuste por polinômio

EXEMPLO:

x_i	y_i	x_i^2	x_i^3	x_i^4	$y_i x_i$	$y_i x_i^2$
0	1.0000	0	0	0	0	0
0.25	1.2840	0.0625	0.0156	0.0390	0.3210	0.0803
0.50	1.6487	0.2500	0.1250	0.0625	0.8244	0.4122
0.75	2.1170	0.5625	0.4219	0.3164	1.5878	1.1908
1.00	2.7183	1.0000	1.0000	1.0000	2.7183	2.7183
2.50	8.7680	1.8750	1.5625	1.3828	5.4515	4.4016

Mínimos Quadrados-Ajuste por polinômio

O sistema linear normal fica então:

$$\begin{bmatrix} 5 & 2.5 & 1.875 \\ 2.5 & 1.875 & 1.5625 \\ 1.875 & 1.5625 & 1.3828 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 8.6780 \\ 5.4515 \\ 4.4016 \end{bmatrix}$$

Resolvendo, teremos:

$$[a_0 \quad a_1 \quad a_2]^T = [1.0051 \quad 0.8647 \quad 0.8432]^T$$

Outline

Caso Não Linear

Quando as funções escolhidas são não-lineares em seus parâmetros, os mínimos quadrados podem ser aplicados fazendo-se transformações convenientes.

EX.:

$$y \approx \alpha_1 e^{-\alpha_2 x} \Rightarrow z = \ln(y) \approx \ln(\alpha_1) - \alpha_2 x$$

$$y \approx \frac{1}{\alpha_1 + \alpha_2 x} \Rightarrow z = \frac{1}{y} \approx \alpha_1 + \alpha_2 x$$

$$y \approx \alpha_1 \alpha_2^x \Rightarrow z = \ln(y) \approx \ln(\alpha_1) + x\ln(\alpha_2) = a_1 + a_2 x$$

$$y \approx \alpha_1 x^{\alpha_2} \Rightarrow z = \ln(y) \approx \ln(\alpha_1) + \alpha_2 \ln(x) = a_1 + a_2 \ln(x) \Rightarrow t = \ln(x) \approx a_1 + a_2 t$$

$$y \approx \alpha_1 + \alpha_2 cos(wx) \Rightarrow t = \cos(wx), \alpha_1 + \alpha_2 t$$

Caso Não Linear - Teste de Alinhamento

Para verificar se a transformação escolhida para a função não-linear foi aceitável seguimos os seguintes passos:

1. Linearização da função não-linear escolhida
2. Diagrama de dispersão dos dados transformados
3. Se os pontos no diagrama de dispersão estiverem alinhados, então a transformação escolhida está razoável

Caso Não Linear - Exemplo

Sejam os seguintes dados tabelados e seu respectivo diagrama de dispersão:

x	-1.00	-0.70	-0.40	-0.10	0.20	0.50	0.80	1.00
y	36.54	17.26	8.15	3.85	1.82	0.86	0.40	0.24

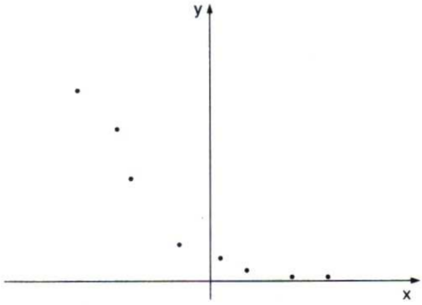

Figura: Fonte: Ruggiero&Lopes

A curva sugere um ajuste exponencial:

$$\varphi(x) = \beta_1 e^{\beta_2 x}$$

Caso Não Linear - Exemplo

Vamos linearizar aplicando a seguinte transformação

$$y \approx \beta_1 e^{\beta_2 x} \Rightarrow z = \ln(y) \approx \ln(\beta_1) + \beta_2 x,$$

donde recaímos no problema linear

$$z \approx \alpha_1 + \alpha_2 x$$

em que $\alpha_1 = \ln(\beta_1)$ e $\alpha_2 = \beta_2$.

A tabela dos dados linearizados fica:

x	-1.00	-0.70	-0.40	-0.10	0.20	0.50	0.80	1.00
$z = \ln(y)$	3.60	2.85	2.10	1.35	0.60	-0.15	-0.90	-1.40

e o diagrama de dispersão dos dados transformados como a seguir

Caso Não Linear - Exemplo

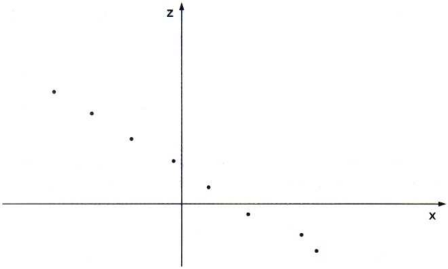

Figura: Fonte: Ruggiero&Lopes

As equações normais do problema linearizado ficam:

$$\begin{bmatrix} 8.00 & 0.30 \\ 0.30 & 3.59 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 8.00 \\ -8.68 \end{bmatrix}$$

cuja solução é

$$\alpha^* = [1.09 \quad -2.51]^T$$

Caso Não Linear - Exemplo

levando à aproximação

$$z \approx 1.09 - 2.51x$$

cujo desvio total é

$$\sum_{k=1}^{m}((1.09 - 2.51x_k) - z_k)^2 = 3.2 \times 10^{-4}$$

Voltando ao problema original:

$$\beta_1^* = e^{\alpha_1} = 2.99 \quad \beta_2^* = \alpha_2 = -2.51$$

e o ajuste final fica então

$$y \approx 2.99e^{-2.51x}$$

com desvio total

$$\sum_{k=1}^{m}((2.99e^{-2.51x_k}) - y_k)^2 = 0.038.$$

Caso Não Linear - Exemplo

NOTA

É importante destacar que os parâmetros β_1^* e β_2^* não são necessariamente ótimos para o problema original pois eles minimizam o problema linearizado!

Outline

Mínimos Quadrados-Contínuo

O processo no caso contínuo se assemelha ao discreto, exceto que agora o critério a ser minimizado é

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

Vamos desenvolver a expressão em um caso com apenas g_1 e g_2 para simplificar:

$$\int_a^b [f(x) - \varphi(x)]^2 dx = \int_a^b [f(x)^2 - 2f(x)\varphi(x) + \varphi(x)^2] dx$$

$$= \int_a^b \{f(x)^2 - 2f(x)[\alpha_1 g_1(x) + \alpha_2 g_2(x)] + \alpha^2 g_1^2(x) + 2\alpha_1\alpha_2 g_1(x)g_2(x) + \alpha_2^2 g_2^2(x)\} dx$$

$$= \int_a^b f(x)^2 dx - [2\int_a^b f(x)g_1(x)dx]\alpha_1 - [2\int_a^b f(x)g_2(x)dx]\alpha_2 + [\int_a^b g_1^2(x)dx]\alpha_1^2 +$$

$$+[2\int_a^b g_1(x)g_2(x)dx]\alpha_1\alpha_2 + [\int_a^b g_2^2(x)dx]\alpha_2^2 = F(\alpha_1, \alpha_2)$$

Mínimos Quadrados-Contínuo

Como fizemos no caso discreto, devemos novamente encontrar os pontos críticos:

$$\left.\frac{\partial F}{\partial \alpha_i}\right| = 0, i = 1, 2.$$

Mais precisamente:

$$\frac{\partial F}{\partial \alpha_1} = -2\int_a^b f(x)g_1(x)dx + [2\int_a^b g_1^2(x)dx]\alpha_1 + [2\int_a^b g_1(x)g_2(x)dx]\alpha_2$$

$$\frac{\partial F}{\partial \alpha_2} = -2\int_a^b f(x)g_2(x)dx + [2\int_a^b g_2^2(x)dx]\alpha_2 + [2\int_a^b g_1(x)g_2(x)dx]\alpha_1$$

Igualando as derivadas parciais a zero:

$$\begin{cases} \left[\int_a^b g_1^2(x)dx\right]\alpha_1 + \left[\int_a^b g_1(x)g_2(x)dx\right]\alpha_2 = \int_a^b f(x)g_1(x)dx \\ \left[\int_a^b g_2^2(x)dx\right]\alpha_2 + \left[\int_a^b g_1(x)g_2(x)dx\right]\alpha_1 = \int_a^b f(x)g_2(x)dx \end{cases}$$

Mínimos Quadrados-Contínuo

o qual pode ser representado na forma matricial $A\alpha = b$:

$$\begin{cases} a_{11}\alpha_1 + a_{12}\alpha_2 = b_1 \\ a_{21}\alpha_1 + a_{22}\alpha_2 = b_2 \end{cases}$$

dado que

$$a_{11} = \int_a^b g_1^2(x)dx \quad a_{12} = \int_a^b g_1(x)g_2(x)dx = \int_a^b g_2(x)g_1(x)dx = a_{21}$$

$$a_{22} = \int_a^b g_2^2(x)dx \quad b_1 = \int_a^b f(x)g_1(x)dx \quad b_2 = \int_a^b f(x)g_2(x)dx$$

Usando o conceito de produto escalar entre funções:

$$< p, q >= \int_a^b p(x)q(x)dx$$

temos

$$a_{ij} =< g_i, g_j > \qquad b_i =< f, g_i >$$

e novamente podemos ter funções ortogonais entre si envolvidas.

Polinômios de Legendre

Exemplo típico de funções ortogonais são os polinômios de Legendre, definidos por

$$P_0(x) \equiv 1, \quad P_k(x) = \frac{1}{2^k k!} \frac{d^{(k)}}{dx^{(k)}} [(x^2 - 1)]^k, k = 1, 2, \ldots$$

Pode-se mostrar que eles são ortogonais entre si em relação ao produto escalar no intervalo $[-1, 1]$. Vamos exemplificar com os três primeiros:

$$P_0(x) = 1 \qquad P_1(x) = x \qquad P_2(x) = \frac{1}{2}(3x^2 - 1)$$

Polinômios de Legendre

$$< P_0, P_0 >= \int_{-1}^{1} 1dx = x|_{-1}^{1} = 2$$

$$< P_0, P_1 >= \int_{-1}^{1} xdx = \frac{1}{2}x^2\Big|_{-1}^{1} = 0$$

$$< P_0, P_2 >= \int_{-1}^{1} \frac{1}{2}(3x^2 - 1)dx = \frac{1}{2}x^3 - \frac{1}{2}x\Big|_{-1}^{1} = 0$$

$$< P_1, P_1 >= \int_{-1}^{1} x^2dx = \frac{1}{3}x^3\Big|_{-1}^{1} = \frac{2}{3}$$

$$\vdots$$

Polinômios ortogonais costumam satisfazer uma fórmula de recorrência a três termos. No caso de Legendre:

$$P_{j+1}(x) = \left(\frac{2j+1}{j+1}\right) xP_j(x) - \left(\frac{j}{j+1}\right) P_{j-1}(x), j = 1, 2, \ldots$$

Mínimos Quadrados-Contínuo-Exemplo

Aproximar a função $f(x) = 4x^3$ por uma reta em $[0, 1]$

$$\varphi(x) = \alpha_1 x + \alpha_2 = \alpha_1 g_1(x) + \alpha_2 g_2(x)$$

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

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

$$a_{22} =< g_2, g_2 >= \int_a^b 1 dx = x|_0^1 = 1$$

$$b_1 =< f, g_1 >= \int_a^b 4x^4 dx = \frac{4}{5}x^5\Big|_0^1 = \frac{4}{5}$$

$$b_2 =< f, g_2 >= \int_a^b 4x^3 dx = x^4|_0^1 = 1$$

Mínimos Quadrados-Contínuo-Exemplo

O sistema linear é portanto

$$\begin{bmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} \frac{4}{5} \\ 1 \end{bmatrix}$$

cuja solução é

$$\alpha^* = [\frac{18}{5} \quad -\frac{4}{5}]^T$$

e a reta de melhor ajuste finalmente é

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