

## **4.7 - AJUSTE DE CURVAS PELO MÉTODO DOS QUADRADOS MÍNIMOS**

### **Introdução**

Vimos, no capítulo anterior, que uma forma de se trabalhar com uma função definida por uma tabela de valores é a interpolação polinomial.

Contudo, a interpolação não é aconselhável quando:

- a) é preciso obter um valor aproximado da função em algum ponto fora do intervalo de tabelamento, ou seja, quando se quer extrapolar;
- b) os valores tabelados são resultados de algum experimento físico ou de alguma pesquisa, porque, nestes casos, estes valores poderão conter erros inerentes que, em geral, não são previsíveis.

Surge então a necessidade de se ajustar a estas funções tabeladas uma função que seja uma “boa aproximação” para os valores tabelados e que nos permita “extrapolar” com certa margem de segurança.

### **4.7.1 - Método dos quadrados mínimos**

#### **4.7.1.1- O Caso discreto**

Sejam dados os pontos  $(x_1, f(x_1)), (x_2, f(x_2)), \dots, (x_m, f(x_m))$  e as  $n$  funções  $g(x), g_2(x), \dots, g_n(x)$  escolhidas de alguma forma.

Consideraremos que o número de pontos  $m$ , tabelados, é sempre maior ou igual a  $n$  o número de funções escolhidas ou o número de coeficientes  $\alpha_i$  a se determinar.

Nosso objetivo é encontrar os coeficientes  $\alpha_1, \alpha_2, \dots, \alpha_n$  tais que a função  $\varphi(x) = \alpha_1 g_1(x) + \alpha_2 g_2(x) + \dots + \alpha_n g_n(x)$  se aproxime ao máximo de  $f(x)$ .

Seja  $d_k = f(x_k) - \varphi(x_k)$  o desvio em  $x_k$ . Vamos observar que, um conceito de proximidade é que  $d_k$  seja mínimo para todo  $k = 1, 2, \dots, m$ .

O método dos quadrados mínimos consiste em escolher os  $\alpha_j$ 's de tal forma que a soma dos quadrados dos desvios seja mínima. É claro que se a soma

$\sum_{k=1}^m d_k^2 = \sum_{k=1}^m (f(x_k) - \varphi(x_k))^2$  é mínima, teremos que cada parcela  $[f(x_k) - \varphi(x_k)]^2$  é

pequena, donde cada desvio  $[f(x_k) - \varphi(x_k)]$  é pequeno.

Portanto, dentro do critério dos quadrados mínimos, os coeficientes  $\alpha_k$ , que fazem com que  $\varphi(x)$  se aproxime ao máximo de  $f(x)$ , são os que minimizam a função

$$\begin{aligned} F(\alpha_1, \alpha_2, \dots, \alpha_n) &= \sum_{k=1}^m [f(x_k) - \varphi(x_k)]^2 = \\ &= \sum_{k=1}^m [f(x_k) - \alpha_1 g_1(x_k) - \alpha_2 g_2(x_k) - \dots - \alpha_n g_n(x_k)]^2 . \end{aligned}$$

Observamos que, se o modelo ajustar exatamente os dados, o mínimo da função acima será zero e, portanto, a interpolação é um caso especial dentro do método dos quadrados mínimos.

Usando o Cálculo Diferencial, sabemos que, para obter um ponto de mínimo de  $F(\alpha_1, \alpha_2, \dots, \alpha_n)$ , temos de, inicialmente, encontrar seus pontos críticos, ou seja, os  $(\alpha_1, \alpha_2, \dots, \alpha_n)$  tais que

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

Calculando estas derivadas parciais para cada  $j = 1, 2, \dots, n$ , temos

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

Impondo a condição

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

temos

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

Assim,

$$\left. \begin{aligned} \sum_{k=1}^m [f(x_k) - \alpha_1 g_1(x_k) - \dots - \alpha_n g_n(x_k)] g_1(x_k) &= 0 \\ \sum_{k=1}^m [f(x_k) - \alpha_1 g_1(x_k) - \dots - \alpha_n g_n(x_k)] g_2(x_k) &= 0 \\ &\vdots \\ \sum_{k=1}^m [f(x_k) - \alpha_1 g_1(x_k) - \dots - \alpha_n g_n(x_k)] g_n(x_k) &= 0 \end{aligned} \right\} \Rightarrow$$

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

que é um sistema linear com  $n$  equações e  $n$  incógnitas:  $\alpha_1, \alpha_2, \dots, \alpha_n$ .

As equações deste sistema linear são as chamadas *equações normais*.  
O sistema linear acima pode ser escrito na forma matricial  $A\alpha = b$ :

$$\begin{cases} a_{11}\alpha_1 + a_{12}\alpha_2 + \dots + a_{1n}\alpha_n = b_1 \\ a_{21}\alpha_1 + a_{22}\alpha_2 + \dots + a_{2n}\alpha_n = b_2 \\ \vdots \\ a_{n1}\alpha_1 + a_{n2}\alpha_2 + \dots + a_{nn}\alpha_n = b_n \end{cases}$$

onde  $A = (a_{ij})$  é tal que  $a_{ij} = \sum_{k=1}^m g_j(x_k)g_i(x_k) = a_{ji}$  (ou seja,  $A$  é simétrica)

$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)^t$  e  $b = (b_1, b_2, \dots, b_n)^t$  é tal que

$$b_i = \sum_{k=1}^m f(x_k)g_i(x_k).$$

Lembramos que, dados os vetores  $x$  e  $y \in \mathfrak{R}^m$ , o número real  $\langle x, y \rangle = \sum_{i=1}^m x_i y_i$  é

chamado de *produto escalar* de  $x$  por  $y$ .

Usando esta notação, o sistema normal  $A\alpha = b$  ficará expresso por

$$A = (a_{ij}) = \langle \bar{g}_i, \bar{g}_j \rangle \text{ e } b = (b_i) = \langle \bar{f}, \bar{g}_i \rangle \text{ onde}$$

$\bar{g}_\ell$  é o vetor  $(g_\ell(x_1)g_\ell(x_2) \dots g_\ell(x_m))^T$  e  $\bar{f}$ , o vetor  $(f(x_1)f(x_2) \dots f(x_m))^T$ .

Demonstra-se que, se as funções  $g_1(x), \dots, g_n(x)$  forem tais que os vetores  $\bar{g}_1, \bar{g}_2, \dots, \bar{g}_n$  sejam linearmente independentes, então o determinante da matriz  $A$  é diferente de zero e, portanto, o sistema linear

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

admite solução única:  $\bar{\alpha}_1, \dots, \bar{\alpha}_n$ . Ainda mais, demonstra-se também que esta solução  $\bar{\alpha}_1, \dots, \bar{\alpha}_n$  é o ponto em que a função  $F(\alpha_1, \dots, \alpha_n)$  atinge seu valor mínimo.

Observamos que, se os vetores  $\bar{g}_1, \bar{g}_2, \dots, \bar{g}_n$  tiverem uma propriedade suplementar de serem tais que  $\langle \bar{g}_i, \bar{g}_j \rangle: \begin{cases} = 0, i \neq j \\ \neq 0, i = j \end{cases}$ , o que, em linguagem de álgebra linear

se diz “se os vetores  $\bar{g}_1, \bar{g}_2, \dots, \bar{g}_n$  forem ortogonais entre si”, então a matriz  $A$  do sistema normal será matriz diagonal, com  $a_i \neq 0$  e, portanto, o sistema terá solução única, a qual será facilmente determinada.

Felizmente, dado um conjunto de pontos  $\{x_1, x_2, \dots, x_m\}$  é fácil construir polinômios de grau 0, 1, ...,  $n$  que são ortogonais, no sentido acima, em relação ao produto escalar

$$\langle \bar{g}_i, \bar{g}_j \rangle = \sum_{k=1}^m g_i(x_k)g_j(x_k).$$

Polinômios ortogonais constituem uma classe particular de funções ortogonais. Tais funções possuem várias propriedades muito interessantes e úteis. O leitor interessado em aprender sobre o assunto pode pesquisar, por exemplo, nos livros [5] e [27]. O estudo de funções ortogonais, em particular de polinômios ortogonais, merece um capítulo especial, o que será feito aqui.

#### Exemplo 4.7.1:

Seja o conjunto de pontos  $X_5 = \{-1, -1/2, 0, 1/2, 1\}$  e os polinômios

$$g_0(x) = 1; g_1(x) = x, g_2(x) = x^2 - 1/2$$

Então, os polinômios  $g_0(x)$ ,  $g_1(x)$  e  $g_2(x)$  são funções ortogonais em  $X_5$  com relação ao produto escalar  $\langle \bar{g}_i, \bar{g}_j \rangle = \sum_{k=1}^m g_i(x_k)g_j(x_k)$  pois os vetores

$$\bar{g}_0 = (g_0(x_i)) = (1 \ 1 \ 1 \ 1 \ 1)^T$$

$$\bar{g}_1 = (g_1(x_i)) = (-1 \ -1/2 \ 0 \ 1/2 \ 1)^T \text{ e}$$

$\bar{g}_2 = (g_2(x_i)) = (1/2 \ -1/4 \ -1/2 \ -1/4 \ 1/2)^T$  são ortogonais entre si, o que se verifica facilmente:

$$\langle \bar{g}_0, \bar{g}_0 \rangle = 5 \neq 0$$

$$\langle \bar{g}_0, \bar{g}_1 \rangle = 1(-1) + 1(-1/2) + 1(0) + 1(1/2) + 1(1) = 0$$

$$\langle \bar{g}_0, \bar{g}_2 \rangle = 1(1/2) + 1(-1/4) + 1(-1/2) + 1(-1/4) + 1(1/2) = 0$$

Fica a cargo do leitor fazer as demais verificações.

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

$$\text{Assim, } \left\langle P_{i,m}, P_{j,m} \right\rangle \begin{cases} = 0 & \text{se } i \neq j \\ \neq 0 & \text{se } i = j \end{cases}$$

#### Exemplo 4.7.2:

Seja a função tabelada

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

Feito o diagrama de dispersão, deve ser ajustada por uma parábola passando pela origem, ou seja,  $f(x) = \varphi(x) = \alpha x^2$  (neste caso temos apenas uma função  $g(x) = x^2$ ).

Temos, pois, de resolver apenas a equação

$$\left[ \sum_{k=1}^{11} g(x_k)g(x_k) \right] \alpha = \sum_{k=1}^{11} f(x_k)g(x_k)$$

$$\left[ \sum_{k=1}^{11} g(x_k)^2 \right] \alpha = \sum_{k=1}^{11} f(x_k)g(x_k)$$

$$\left[ \sum_{k=1}^{11} (x_k^2)^2 \right] \alpha = \sum_{k=1}^{11} (x_k^2)f(x_k)$$

Continuando a tabela com  $g(x_k)g(x_k)$  e  $g(x_k)f(x_k)$ , temos

x	-1.0	-0.75	-0.6	-0.5	-0.3	0	0.2	0.4	0.5	0.7	1.0	Somas
$(x^2)(x^2)$	1	0.3164	0.1296	0.0625	0.0081	0	0.0016	0.0256	0.0625	0.2401	1.0	2.8464
$f(x)x^2$	2.05	0.6486	0.162	0.1	0.045	0	0.008	0.096	0.128	0.0588	2.05	5.8756

$$\text{Assim, nossa equação é } 2.0642\alpha = 5.8756 \Rightarrow \alpha = \frac{5.8756}{2.8464} \approx 2.0642$$

Então  $\varphi(x) = 2.0642x^2$  é a parábola que melhor se aproxima, no sentido dos quadrados mínimos, da função tabelada.

#### 4.7.1.2- O Caso Contínuo

Para simplificar a notação, desenvolveremos aqui o caso em que “escolhemos” apenas duas funções.

Sejam então  $f(x)$  contínua em um intervalo  $[a, b]$  e  $g_1(x)$  e  $g_2(x)$  duas funções contínuas em  $[a, b]$  que foram escolhidas de alguma forma. É preciso encontrar duas constantes reais  $\alpha_1$  e  $\alpha_2$  tais que  $\varphi(x) = \alpha_1 g_1(x) + \alpha_2 g_2(x)$  esteja o “mais próximo possível” de  $f(x)$ .

Seguindo o critério dos quadrados mínimos para o conceito de proximidade entre  $\varphi(x)$  e  $f(x)$ , os coeficientes  $\alpha_1$  e  $\alpha_2$  a serem obtidos deverão ser tais que o valor de  $\int_a^b [f(x) - \varphi(x)]^2 dx$  seja o menor possível.

Geometricamente, isto significa que a área entre as curvas  $f(x)$  e  $\varphi(x)$  seja mínima. Portanto, o problema consiste em obter o mínimo para

$$\begin{aligned} \int_a^b [f(x) - \varphi(x)]^2 dx &= \int_a^b [f(x)^2 - 2f(x)\varphi(x) + \varphi(x)^2] dx = \\ &= \int_a^b \{f(x)^2 - 2f(x)[\alpha_1 g_1(x) + \alpha_2 g_2(x)] + \alpha_1^2 g_1^2(x) + \\ &\quad + 2\alpha_1 \alpha_2 g_1(x)g_2(x) + \alpha_2^2 g_2^2(x)\} dx \\ &= \int_a^b f(x)^2 dx - [2 \int_a^b f(x)g_1(x) dx] \alpha_1 - [2 \int_a^b f(x)g_2(x) dx] \alpha_2 + \\ &\quad + [\int_a^b g_1^2(x) dx] \alpha_1^2 + [2 \int_a^b g_1(x)g_2(x) dx] \alpha_1 \alpha_2 + [\int_a^b g_2^2(x) dx] \alpha_2^2 = F(\alpha_1, \alpha_2) \\ &\Rightarrow \int_a^b [f(x) - \varphi(x)]^2 dx = F(\alpha_1, \alpha_2) \end{aligned}$$

Com o mesmo argumento do caso discreto, temos de achar os pontos críticos de  $F$ , ou seja, achar  $(\alpha_1, \alpha_2)$  tal que

$$\left. \frac{\partial F}{\partial \alpha_i} \right|_{(\alpha_1, \alpha_2)} = 0, \quad i = 1, 2.$$

$$\begin{aligned} i = 1 \Rightarrow \left. \frac{\partial F}{\partial \alpha_1} \right|_{(\alpha_1, \alpha_2)} &= -2 \int_a^b f(x)g_1(x) dx + [2 \int_a^b g_1^2(x) dx] \alpha_1 + \\ &\quad + [2 \int_a^b g_1(x)g_2(x) dx] \alpha_2 \end{aligned}$$

$$\text{Assim, } \left. \frac{\partial F}{\partial \alpha_1} \right|_{(\alpha_1, \alpha_2)} = \left. \frac{\partial F}{\partial \alpha_2} \right|_{(\alpha_1, \alpha_2)} = 0 \Rightarrow$$

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

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

$$a_{22} = \int_a^b g_2^2(x) dx$$

$$b_1 = \int_a^b f(x) g_1(x) dx \quad \text{e} \quad b_2 = \int_a^b f(x) g_2(x) dx,$$

podemos escrever o sistema linear acima como

$$\begin{cases} a_{11} \alpha_1 + a_{12} \alpha_2 = b_1 \\ a_{21} \alpha_1 + a_{22} \alpha_2 = b_2 \end{cases} \quad \text{ou} \quad A \alpha = b, \quad \text{onde} \quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$\alpha = (\alpha_1 \alpha_2)^T, \quad b = (b_1 b_2)^T.$$

Demonstra-se que, se as funções escolhidas  $g_1(x)$  e  $g_2(x)$  forem linearmente independentes, o determinante da matriz  $A$  é diferente de zero, o que implica que o sistema linear admite única solução  $(\bar{\alpha}_1, \bar{\alpha}_2)$ . Ainda mais, demonstra-se também que esta solução é o ponto em que a função  $F(\alpha_1, \alpha_2)$  atinge seu valor mínimo.

Usando aqui a definição de produto escalar de duas funções  $p(x)$  e  $q(x)$  no intervalo  $[a, b]$  por

$$\langle p, q \rangle = \int_a^b p(x) q(x) dx,$$

teremos que, no caso em que queremos aproximar

$f(x) \approx \alpha_1 g_1(x) + \dots + \alpha_n g_n(x)$  o sistema normal  $A \alpha = b$  fica

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

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

Da mesma forma que no caso discreto, temos funções ortogonais com relação ao produto escalar, como mostrará o exemplo abaixo.

**Exemplo 4.7.3:**

Os polinômios de Legendre, definidos por

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

são ortogonais em  $[-1, 1]$ , com relação ao produto escalar  $\langle p, q \rangle = \int_a^b p(x)q(x)dx$ .

Fica como exercício a verificação de que os três primeiros polinômios de Legendre  $P_0(x) \equiv 1, P_1(x) = x$  e  $P_2(x) = \frac{1}{2}(3x^2 - 1)$  são ortogonais entre si.

Uma observação interessante é que, em geral, polinômios ortogonais satisfazem uma fórmula de recorrência de 3 termos, ou seja, dados  $P_0(x)$  e  $P_1(x)$ , conseguimos construir  $P_k(x), k = 2, 3, \dots$

No caso dos polinômios de Legendre, a fórmula de recorrência é

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

**Exemplo 4.7.4:**

Vamos aproximar  $f(x) = 4x^3$  por um polinômio do primeiro grau, uma reta, no intervalo  $[a, b] = [0, 1]$ .

$$\varphi(x) = \alpha_1 g_1(x) + \alpha_2 g_2(x) = \alpha_1 + \alpha_2 x, \quad \alpha_1, \alpha_2 \in \mathfrak{R}$$

$$(g_1(x) \equiv 1 \quad g_2(x) = x).$$

Pelo que vimos,  $(\alpha_1, \alpha_2)$  é a única solução de  $A\alpha = b$  onde

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \text{ sendo}$$

$$a_{11} = \int_a^b g_1^2(x) dx = \int_0^1 1 dx = 1$$

$$a_{12} = \int_a^b g_1(x)g_2(x) dx = \int_0^1 x dx = \frac{x^2}{2} \Big|_0^1 = \frac{1}{2} = a_{21}$$

$$a_{22} = \int_a^b g_2^2(x) dx = \int_0^1 x^2 dx = \frac{x^3}{3} \Big|_0^1 = \frac{1}{3}$$

$$b_1 = \int_a^b f(x)g_1(x) dx = \int_0^1 4x^3 dx = \frac{4x^4}{4} \Big|_0^1 = 1$$

$$b_2 = \int_a^b f(x)g_2(x) dx = \int_0^1 4x^3 x dx = \frac{4x^5}{5} \Big|_0^1 = \frac{4}{5}$$

Temos então o sistema

$$\begin{cases} 1\alpha_1 + \frac{1}{2}\alpha_2 = 1 \\ \frac{1}{2}\alpha_1 + \frac{1}{3}\alpha_2 = \frac{4}{5} \end{cases} \Rightarrow \alpha_1 = -\frac{4}{5}, \alpha_2 = \frac{18}{5}.$$

Logo, a aproximação por quadrados mínimos de  $f(x) = 4x^3$  no intervalo  $[0, 1]$ , por um polinômio de grau 1, é a reta  $\varphi(x) = \frac{18}{5}x - \frac{4}{5}$ .

#### **4.7.3- O Caso Não Linear**

Em alguns casos, a família de funções escolhidas pode ser não linear nos parâmetros, como, por exemplo, se ao diagrama de dispersão de uma determinada função se ajustar uma exponencial do tipo  $f(x) \approx \varphi(x) = \alpha_1 e^{-\alpha_2 x}$ ,  $\alpha_1$  e  $\alpha_2$  positivos.

Para se aplicar o método dos quadrados mínimos, é necessário que se efetue uma linearização do problema através de alguma transformação conveniente.

Por exemplo:

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

Se  $a_1 = \ln(\alpha_1)$  e  $a_2 = -\alpha_2 \Rightarrow \ln(y) \approx a_1 - a_2 x = \varphi(x)$  que é um problema linear nos parâmetros  $a_1$  e  $a_2$ .

O método dos quadrados mínimos pode então ser aplicado na resolução do problema linearizado. Obtidos os parâmetros deste problema, usaremos estes valores para calcular os parâmetros originais.

É importante observar que os parâmetros assim obtidos não são ótimos dentro do critério dos quadrados mínimos, isto porque estamos ajustando o problema linearizado por quadrados mínimos e não o problema original.

Portanto, no exemplo, os parâmetros  $a_1$  e  $a_2$  são os que ajustam a função  $\varphi(x)$  à função  $z(x)$  no sentido dos quadrados mínimos; não se pode afirmar que os parâmetros  $\alpha_1$  e  $\alpha_2$  (obtidos através de  $a_1$  e  $a_2$ ) são os que ajustam  $\varphi(x)$  à  $f(x)$  dentro do critério dos quadrados mínimos.

**Exemplo 4.7.5:**

Suponhamos que num laboratório obtivemos experimentalmente os seguintes valores para  $f(x)$  sobre os pontos  $x_i$ ,  $i = 1, 2, \dots, 8$ :

x	-1.0	-0.7	-0.4	-0.1	0.2	0.5	0.8	1.0
f(x)	36.547	17.264	8.155	3.852	1.820	0.860	0.406	0.246

Fazendo o diagrama de dispersão dos dados acima, obtemos

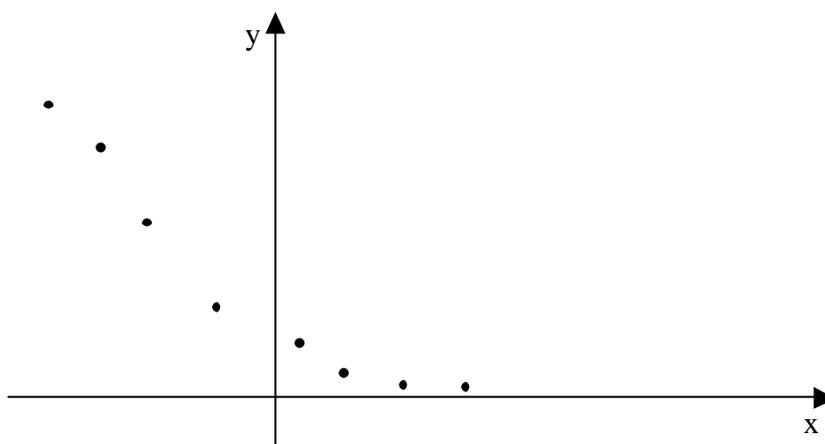


Figura 4.7.1 – diagrama de dispersão dos dados da tabela dada.

Os dados nos sugere um ajuste  $y \approx \varphi(x) = \alpha_1 e^{-\alpha_2 x}$ .

Conforme vimos anteriormente, a “linearização” a ser feita é

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

Assim, em vez de ajustarmos  $y$  por quadrados mínimos, ajustaremos  $z = \ln(y)$  por quadrados mínimos, encontrando  $\phi(x) = a_1 + a_2 x$ , onde  $a_1 = \ln(\alpha_1)$  e  $a_2 = -\alpha_2$ . (Aqui  $g_1(x) = 1$  e  $g_2(x) = x$ ).

Temos pois:

x	-1.0	-0.7	-0.4	-0.1	0.2	0.5	0.8	1.0
z = ln(y)	3.599	2.849	2.099	1.349	0.599	-0.151	-0.901	-1.402

e  $a_1$  e  $a_2$  serão a solução do sistema:

$$\begin{cases} [\sum_{k=1}^8 g_1(x_k)g_1(x_k)]a_1 + [\sum_{k=1}^8 g_2(x_k)g_1(x_k)]a_2 = \sum_{k=1}^8 z(x_k)g_1(x_k) \\ [\sum_{k=1}^8 g_1(x_k)g_2(x_k)]a_1 + [\sum_{k=1}^8 g_2(x_k)g_2(x_k)]a_2 = \sum_{k=1}^8 z(x_k)g_2(x_k) \end{cases}$$

$$g_1(x) = 1 \Rightarrow \sum_{k=1}^8 g_1(x_k)g_1(x_k) = \sum_{k=1}^8 1 = a_{11} = 8$$

$$g_2(x) = x \Rightarrow \sum_{k=1}^8 g_2(x_k)g_2(x_k) = \sum_{k=1}^8 x_k^2 = a_{22} = 3.59$$

$$\sum_{k=1}^8 g_1(x_k)g_2(x_k) = \sum_{k=1}^8 1x_k = a_{12} = a_{21} = 0.3$$

$$b_1 = \sum_{k=1}^8 z(x_k)g_1(x_k) = \sum_{k=1}^8 z(x_k) = 8.041$$

$$b_2 = \sum_{k=1}^8 z(x_k)g_2(x_k) = \sum_{k=1}^8 z(x_k)x_k = -8.646$$

donde

$$A = \begin{bmatrix} 8 & 0.3 \\ 0.3 & 3.59 \end{bmatrix} \quad b = \begin{bmatrix} 8.041 \\ -8.646 \end{bmatrix}$$

e o sistema fica

$$\begin{cases} 8.0a_1 + 0.3a_2 = 8.041 \\ 0.3a_1 + 3.59a_2 = -8.646 \end{cases} \Rightarrow a_1 = 1.099 \text{ e } a_2 = -2.5$$

$$\text{Agora, } \alpha_1 = e^{a_1} \Rightarrow \alpha_1 = e^{1.099} = 3.001$$

$$\alpha_2 = -a_2 \Rightarrow \alpha_2 = 2.5.$$

$$\text{Assim, a função } \varphi(x) = \alpha_1 e^{-\alpha_2 x} = 3.001 e^{-2.5x}$$

Assim, como no exemplo anterior, onde ajustamos aos dados a curva  $y \approx \alpha_1 e^{-\alpha_2 x}$ , é comum encontrarmos casos em que os dados tabelados, feito o diagrama de dispersão, devem ser ajustados por

$$1) \text{ Uma hipérbole: } y \approx \frac{1}{\alpha_1 + \alpha_2 x} = \varphi(x)$$

$$(z = \frac{1}{x} \approx \alpha_1 + \alpha_2 x)$$

2) Uma curva exponencial:  $y \approx \alpha_1 \alpha_2^x = \phi(x)$

$$(se\ y > 0, z = \ln(y) \approx \underbrace{\ln(\alpha_1)}_{a_1} + x \underbrace{\ln(\alpha_2)}_{a_2} = a_1 + a_2 x = \phi(x)).$$

3) Uma curva geométrica:  $y \approx \alpha_1 x^{\alpha_2} = \phi(x)$

$$(se\ x > 0\ e\ y > 0, z = \ln(y) \approx \underbrace{\ln(\alpha_1)}_{a_1} + \underbrace{\alpha_2}_{a_2} \ln(x) = a_1 + a_2 \underbrace{\ln(x)}_t)$$

$\Rightarrow z = \ln(y) \approx a_1 + a_2 t = \phi(t)$ . (Aqui minimizamos a soma dos quadrados dos desvios nos logaritmos de y, para os logaritmos de x.)

4) Uma curva trigonométrica:  $y \approx \alpha_1 + \alpha_2 \cos(wx) = \phi(x)$ . ( $t = \cos(wx) \Rightarrow \phi(t) = \alpha_1 + \alpha_2 t$  e, neste caso, estamos minimizando a soma dos quadrados dos desvios em y.)

#### 4.7.4- Teste de Alinhamento

Uma vez escolhida uma função linear em  $\alpha_1, \alpha_2, \dots, \alpha_n$  para ajustar uma função dada, uma forma de verificarmos se a escolha feita foi razoável é aplicarmos o *teste de alinhamento*, que consiste em:

- i) fazer a “linearização” da função não linear escolhida;
- ii) fazer o diagrama de dispersão dos novos dados;
- iii) se os pontos do diagrama (ii) estiverem alinhados, isto significará que a função não linear escolhida foi uma “boa escolha”.

Observamos que, devido aos erros de observação, e cálculos aproximados, consideramos satisfatório o diagrama de dispersão onde os pontos se distribuem aleatoriamente em torno de uma reta média.

No exemplo 4.7.5, temos

x	-1.0	-0.7	-0.4	-0.1	0.2	0.5	0.8	1.0
y	36.547	17.264	8.155	3.852	1.820	0.860	0.406	0.246
z = ln(y)	3.599	2.849	2.099	1.349	0.599	-0.151	-0.901	-1.402

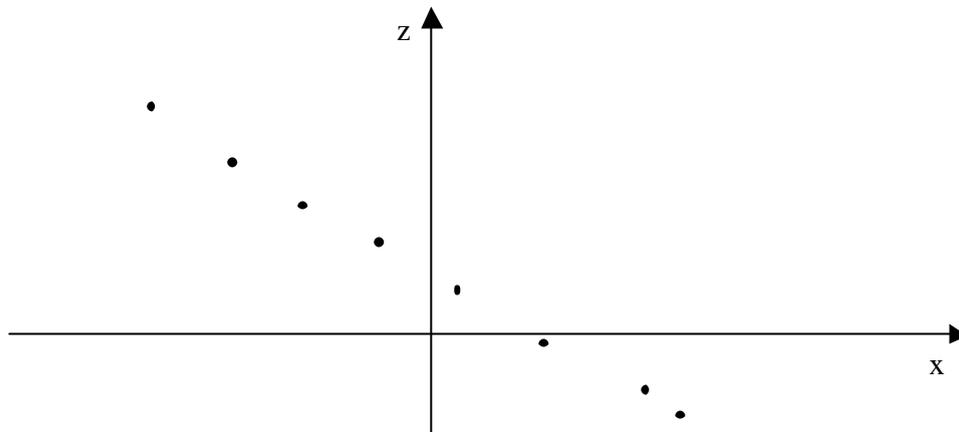


Figura 4.7.2- diagrama de dispersão dos dados da tabela dada.

#### 4.7.5- Exercícios

Ver Ruggiero (página 287 a 291 – exercícios 01 ao 13)