Ajuste de Funções.

Em geral se aplica a um conjunto de dados experimentais: $\{(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)\}$, e deseja-se obter a lei $y = f(x)$ que relaciona x com y e a partir dela, calcular y para um certo x que não dos dados experimentais. O ajustamento traduz um comportamento médio (extrapolação com certa margem de segurança).

Para um conjunto de vários elementos a seleção da aproximação dos dados por interpolação pode ocupar muita memória e tempo de processamento.

Embora não conheçamos a função que originou os dados, pela origem do problema podemos saber o tipo da função que estabelece a relação entre x e y.

Critérios para a medida da qualidade de Ajuste. Método dos mínimos quadrados- Outros Nomes: Análise de Regressão, Suavização Dados, Otimização Linear.

No caso do ajuste de dados, o técnica dos mínimos quadrados é utilizada no cálculo de um conjunto de constantes c_j de uma função f^* que aplicada a uma função f, desconhecida, da forma:

$$f^*(x) = c_1 f_1(x) + \ldots + c_n f_n(x)$$

observe que as constantes desconhecidas c_j aparecem linearmente, embora as funções básicas $f_1, \ldots, f_n$ possam ser funções não lineares em x.

A escolha de f_1, ... , f_n é de acordo com a natureza do problema (dados experimentais e pode não ser fácil escolhe-las de modo a ter uma solução razoável). Então

$$R = \sum_{j=1}^{m} [\, y_i - f^*(x_i)\,]^2$$

onde R seja mínima.

R = R (c_1, ... , c_m) e portanto passarão por um mínimo quando as m derivadas parciais se anularem simultaneamente, ou seja, teremos:

$$\frac{\partial R}{\partial c_j} = 2\sum_{i=1}^{m}[\, f^*(x_i) - y_i\,]\frac{\partial f^*(x_i)}{\partial c_j} \equiv 0$$

para j =1, ... ,n (são mínimos pois a função $[f^*(x_i) - y_i]^2$ é uma grandeza sempre positiva), com

$$\frac{\partial f^*(x_i)}{\partial c_j} = f_j(x_i)$$

j = 1, ... ,n ; i = 0,1, ... , m ; m ≤ n.

Note que temos um sistema de n equações algébricas lineares e n incógnitas c_1, ... , c_m ; são denominadas de equações normais, portanto

$$\begin{cases} A_{11}C_1 + A_{12}C_2 + \ldots + A_{1n}C_n = B_1 \\ A_{21}C_1 + A_{22}C_2 + \ldots + A_{2n}C_n = B_2 \\ \vdots \\ A_{n1}C_1 + A_{n2}C_2 + \ldots + A_{mn}C_n = B_n \end{cases}$$

$$\therefore \; A_{kj} = A_{jk}$$

$$A_{kj} = \sum_{k=1}^{m} f_j(x_k).f_i(x_k)$$

C = ($C_1, C_2, \ldots, C_n)^T$;

$$\mathrm{B} = (\mathrm{B}_1, \mathrm{B}_2, \ldots, \mathrm{B}_n)^T$$

$$\mathrm{B}_i = \sum_{k=1}^{m} f(x_k).f_i(x_k)$$

Demonstra-se que se as funções escolhidas, $f_1(x), f_2(x), \ldots, f_n(x)$, forem linearmente independentes, o determinante da matriz A é diferente de zero e, portanto, o sistema linear admite solução única: $\bar{c}_1, \bar{c}_2, \ldots, \bar{c}_n$. Ainda mais, demonstra-se também que esta solução $\bar{c}_1, \bar{c}_2, \ldots, \bar{c}_n$ é o ponto em que a função *f*($c_1, c_2, \ldots, c_n$) atinge seu valor mínimo.

Exemplo 1: Considere a tabela

x	-1.0	-1.75	-0.6	-0.5	-0.3	0.0	0.2	0.4	0.5	0.7	1
$f(x)$	2.05	-1.153	0.45	0.4	0.5	0.0	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, $f(x) \approx \varphi(x) = \alpha\ x^2$ (neste caso temos apenas uma função $\phi(x) = x^2$).

Temos pois de resolver apenas a equação

$$| \sum_{k=1}^{11} f\ (x_k).f\ (x_k) \,|\, \alpha = \sum_{k=1}^{11} f(x_k).f\ (x_k)$$

$$| \sum_{k=1}^{11} f\ (x_k)^2 \,|\, \alpha = \sum_{k=1}^{11} f(x_k).f\ (x_k)$$

$$| \sum_{k=1}^{11} (x_k^2)^2 \,|\, \alpha = \sum_{k=1}^{11} x_k^2\ .f\ (x_k)$$

Usando a tabela com $g(x_k)g(x_k)$ $g(x_k)f(x_k)$: temos

$$a = \frac{5.8756}{2.8464} = 2.06442$$

Logo $f(x) \approx 2.0642\ x^2$ sendo esta párabola a que melhor se aproxima no sentido dos minimos quadrados a função tabelada.

Se tratar-se de caso não-linear $y = a_1 e^{-a_2 x}$ tem-ser aplicando log dos dois lados que

$$Y = a + BX$$

onde $Y = \ln y$ $a = \ln a_1$, $a_2 = a_2$

daí o metodo pode sr aplicado na resoluçãso do problema linearizado.