Capítulo 5 : Ajuste de Funções.

5.1. Preliminares.

No capítulo 4 vimos como aproximamos uma função por um polinômio e por splines. O ajuste é outra técnica de aproximação de funções que tem características diferentes da interpolação.

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.

Para ajustar um conjunto de dados a uma função é preciso estabelecer o tipo de curva e calcular os parâmetros dessa curva. Se conhecermos os parâmetros dessa curva poderemos calcular $f(x_i)$ e comparar com o valor dado y_i.

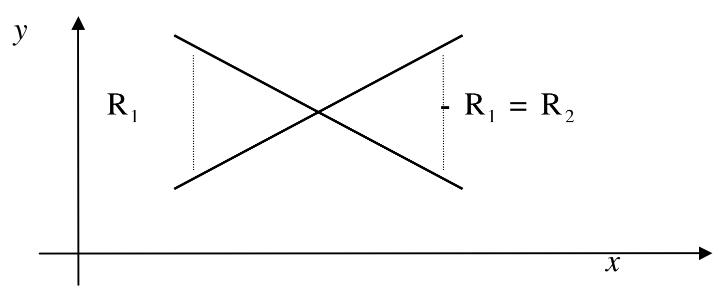
$$\therefore \quad \tilde{\mathrm{R}}_i = f^*(x_i) - y_i \qquad (5.1.1)$$

Resíduo (erro).

A primeira idéia para avaliarmos a qualidade do ajuste seria:

Critério 1: Todos os erros devem tender a zero. Tal critério não é utilizável, ou então cairíamos no problema de interpolação, onde exigimos que:
$f(x_i) = y_i = f^*(x_i)$, isto é, $\mathrm{R}_i = 0$ para $i = 0, 1, \ldots, \mathrm{m}$.

Critério 2: A soma do erro deve ser tão pequena quanto possível, isto é, $\sum_i \mathrm{R}_i$ deve ser minimizada. Tal critério não é utilizável, pois podemos ter uma soma de erros nula sem que o ajuste seja bom. A figura 5.1.1. ilustra tal fato:



Critério 3 : A soma do erro em valor absoluto deve ser a menor possível, isto é,

$\sum_{j=1}^{m}$ | $f^*(x_i) - y_i$ | deve ser mínima.

OBS : | . | não é diferenciável, problema achar o mínimo.

Critério 4 (difícil) : $\max_{1 \le i \le n}$ | $f^*(x_i) - y_i$ | deve ser o mínimo.

Critério 5: Método dos mínimos quadrados. $\sum_{i=1}^{n} \mathrm{R}_\mathrm{i}^2$ é minimizado.

Neste caso, nunca teremos uma soma nula (o que ocorreria só na interpolação) e a função "quadrado" é facilmente diferenciável. É o critério mais usado.

5.2. 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:

$$f^*(x) = \mathrm{c}_1\phi_1(x)+\ldots+\mathrm{c}_\mathrm{n}\phi_n(x) \qquad (5.1\text{-}1)$$

observe que as constantes desconhecidas c_j aparecem linearmente, embora as funções básicas $\phi_1, \ldots, \phi_n$ possam ser funções não lineares em *x*.

A escolha de $\phi_1, \ldots, \phi_n$ 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

$$\mathrm{R} = \sum_{j=1}^{m} [\, y_i - f^*(x_i)\,]^2 \qquad (5.1\text{-}2)$$

onde R seja mínima. De (4.1-1) e (5.1-2) vem: R = R ($c_1, \ldots, c_m$) e portanto passarão por um mínimo quando as m derivadas parciais se anularem simultaneamente, ou seja, teremos:

$$\frac{\partial R}{\partial c_j} = 2\sum_{i=1}^{m}[f^*(x_i) - y_i]\frac{\partial f^*(x_i)}{\partial c_j} \equiv 0 \qquad (5.1\text{-}3)$$

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

$$\frac{\partial f^*(x_i)}{\partial c_j} = \phi_j(x_i) \qquad (5.1\text{-}4)$$

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

Note (5.1-3) é um sistema de n equações algébricas lineares e n incógnitas $c_1, \ldots, c_m$; são denominadas de equações normais, portanto

$$\begin{cases} A_{11}C_1 + A_{12}C_2 + \ldots + A_{1n}C_n = B_1 \\ A_{21}C_1 + A_{22}C_2 + \ldots + A_{2n}C_n = B_2 \\ \vdots \\ A_{n1}C_1 + A_{n2}C_2 + \ldots + A_{mn}C_n = B_n \end{cases} \qquad (5.1\text{-}5)$$

$$\therefore\ A_{kj} = A_{jk} \qquad (5.1\text{-}6)$$

$$A_{kj} = \sum_{k=1}^{m} \phi_j(x_k).\phi_i(x_k) \qquad (5.1\text{-}7)$$

$$\mathrm{C} = (C_1, C_2, \ldots, C_n)^T \quad ; \quad \mathrm{B} = (\mathrm{B}_1, \mathrm{B}_2, \ldots, \mathrm{B}_n)^T$$

$$\mathrm{B_i} = \sum_{k=1}^{m} f(x_k).\phi_i(x_k) \qquad (5.1\text{-}8)$$

Demonstra-se que se as funções escolhidas, $\phi_1(x), \phi_2(x), \ldots, \phi_n(x)$, 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

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

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

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

Continuando a tabela com $\phi(x_k).\phi(x_k)$ e $\phi(x_k).f(x_k)$, temos: