

## UMA DISCUSSÃO SOBRE O AJUSTAMENTO DE FUNÇÕES NÃO-LINEARES NOS PARÂMETROS

Romeu MAGNANI\*  
João SAHÃO JUNIOR\*

**RESUMO:** O problema de estimação de parâmetros no caso não-linear é abordado pela algoritmização de um método clássico (Newton), juntamente com modificações que visam superar algumas dificuldades nele encontradas. Sugestões no sentido de melhorar o desempenho deste método também são aqui enfatizadas.

**UNITERMOS:** Estimação de parâmetros; Modelos não-lineares; métodos iterativos do gradiente.

### INTRODUÇÃO

A análise de dados experimentais, frequentemente, envolve o ajustamento de uma relação funcional entre as variáveis em estudo. No caso mais simples, os dados consistem de valores  $y_1, y_2, \dots, y_n$  de uma variável dependente y obtidos para diversos valores  $x_1, x_2, \dots, x_n$  de uma variável independente x. Como exemplo, y poderia ser a resposta de um método aplicado para uma quantidade x de certa substância. O que se procura é relacionar as variáveis através de uma função matemática  $y = f(x)$ .

O tipo da função pode ser conhecido por considerações teóricas ou ser investigado a partir de indicações fornecidas pela representação gráfica dos pontos experimentais. Assim, os pontos representados nas Fig. 1 e Fig. 2 sugerem funções, respectivamente, linear e quadrática. Estas são casos particulares, para  $r = 1$  e  $r = 2$ , da função polinomial

$$y = b_0 + b_1x + b_2x^2 + \dots + b_r x^r$$

a qual é linear nos parâmetros  $b_0, b_1, b_2, \dots, b_r$ .

Se os parâmetros fossem conhecidos, seria sempre observada uma diferença entre o valor experimental de y e o correspondente valor teórico. Esse erro ocorre mesmo no laboratório, onde é possível, em geral, ser exercido um controle rigoroso sobre a experiência. Entretanto, os parâmetros são incógnitas a serem determinadas e, devido à impossibilidade de se conhecer exatamente os erros, são estabelecidas, apenas, estimativas dos parâmetros. A técnica comumente empregada é a dos "mínimos quadrados", que produz estimativas de modo a minimizar a soma de quadrados

$$S = \sum_{i=1}^n [y_i - f(x_i)]^2$$

a qual é função dos r parâmetros, pois estes são, agora, as variáveis. A diferença entre o valor experimental e o valor "ajustado",  $y_i - f(x_i)$ , representa uma aproximação do erro real e recebe o nome de desvio ou resíduo. A derivação de S, em

\* Departamento de Físico-Química — Instituto de Química — UNESP — 14.800 — Araraquara — SP.

relação a cada parâmetro, leva a um sistema de  $r$  equações, chamadas normais, cuja resolução fornece as estimativas (2,5).

Existem muitas situações na prática em que a relação entre as variáveis deve ser estudada por uma função não-linear nos parâmetros (1,2,5), tais como

$$y = b_1 e^{b_2 x}$$

$$y = b_1 + b_2 e^{b_3 x}$$

A primeira pode ser transformada em linear com a aplicação de logaritmos nos dois membros da igualdade. A segunda, entretanto, não permite uma linearização e o sistema de equações normais correspondente, que também será não-linear, é de difícil resolução.

Não raro, um conjunto de dados experimentais é analisado através de uma função linear nos parâmetros quando uma não-linear seria, provavelmente, a indicada. Ainda que a primeira, empregada devido as facilidades na estimação, tenha conduzido a resultados satisfatórios, o advento dos microcomputadores possibilita, e talvez imponha, que seja tentado um ajustamento da segunda. Neste artigo é discutido, para o caso não-linear, um método clássico de estimação de parâmetros. O objetivo é introduzir o assunto a pesquisadores que, pela formação não exclusivamente matemática, relutam em enfrentar o problema de estimação não-linear.

### DISCUSSÃO E CONCLUSÕES

Como acentuado anteriormente, o ajustamento da função  $f(x)$  aos dados experimentais  $(x_i, y_i)$ ,  $(i = 1, 2, \dots, n)$ , envolve a minimização de

$$S(B) = \sum_{i=1}^n [y_i - f(x_i)]^2$$

onde  $B = (b_1, b_2, \dots, b_r)$  é o vetor de parâmetros, cuja estimativa de mínimos qua-

drados  $\hat{B} = (\hat{b}_1, \hat{b}_2, \dots, \hat{b}_r)$  pretende-se encontrar. Para  $f(x)$  não-linear, um método iterativo clássico é o de Newton (1,5,8) o qual se resume nos quatro passos discutidos a seguir:

#### 1. Iniciar a iteração com um vetor $B_0 = (b_1^0, b_2^0, \dots, b_r^0)$

O vetor  $B_0$  é modificado a cada iteração pelo próprio processo, mas, no início da primeira iteração ele deve estar à disposição. A escolha de um vetor inicial de estimativas pode ser baseada na intuição ou no conhecimento prévio que se tenha da função em estudo. (6) De qualquer modo, o sucesso do método depende em grande parte dessa escolha inicial.

#### 2. Calcular $B_1 = B_0 - H^{-1} G_0$

A matriz  $H$  é a Hessiana de  $S(B)$ , isto é,

$$H = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1r} \\ h_{21} & h_{22} & \dots & h_{2r} \\ \dots & \dots & \dots & \dots \\ h_{r1} & h_{r2} & \dots & h_{rr} \end{bmatrix}$$

$$h_{ri} = h_{ir} \dots h_{rr}$$

com

$$h_{km} = \frac{\partial^2 S}{\partial b_k \partial b_m} \quad (k, m = 1, 2, \dots, r)$$

$$= 2 \sum_{i=1}^n \frac{\partial f}{\partial b_k} \frac{\partial f}{\partial b_m} - 2 \sum_{i=1}^n [y_i - f(x_i)] \frac{\partial^2 f}{\partial b_k \partial b_m}$$

e  $G$  é o vetor gradiente de  $S(B)$ , isto é,

com

$$G = (g_1, g_2, \dots, g_r)$$

$$g_k = \frac{\partial S}{\partial b_k} \quad (k = 1, 2, \dots, r)$$

$$= -2 \sum_{i=1}^n [y_i - f(x_i)] \frac{\partial f}{\partial b_k}$$

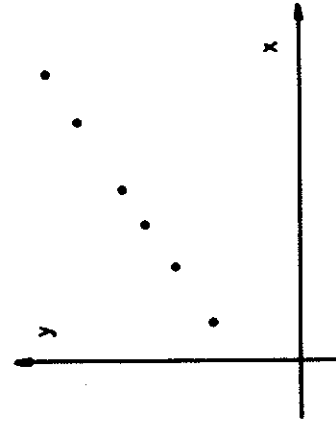


FIG. 1 — Dependência linear

O índice "0" em  $H_0$  e  $G_0$  indica que os elementos da matriz Hessiana e do vetor gradiente são calculados com  $b_1^0, b_2^0, \dots, b_r^0$ , respectivamente, em lugar de  $b_1, b_2, \dots, b_r$ . Obtém-se, assim, um novo vetor de estimativas  $B_1 = (b_1^1, b_2^1, \dots, b_r^1)$ .

Por exemplo, no caso de dois parâmetros a fórmula deste passo será

$$\begin{bmatrix} b_1^1 \\ b_2^1 \end{bmatrix} = \begin{bmatrix} b_1^0 \\ b_2^0 \end{bmatrix} - \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}^{-1} \begin{bmatrix} g_1 \\ g_2 \end{bmatrix}$$

com  $h_{km}$  e  $g_k$  calculados de acordo com as formulas acima. Invertendo  $H$ , chega-se no vetor  $B_1$  de componentes

$$b_1^1 = b_1^0 - \frac{h_{22} g_1 - h_{12} g_2}{h_{11} h_{22} - h_{12} h_{21}}$$

$$b_2^1 = b_2^0 - \frac{-h_{21} g_1 - h_{12} g_2}{h_{11} h_{22} - h_{12} h_{21}}$$

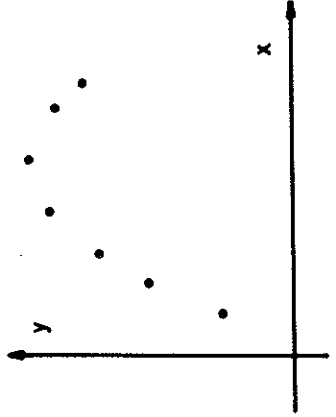


FIG. 2 — Dependência quadrática.

O procedimento adotado equivale a determinar o vetor  $B_1$  que minimiza uma aproximação  $s$  de  $S$ , definida por

$$s(B) = s(B_0) + G_0^T (B - B_0) + \frac{1}{2} (B - B_0)^T H_0 (B - B_0)$$

a qual consiste nos termos até segunda ordem do desenvolvimento em série de Taylor de  $S(B)$ , em torno de  $B_0$ . O expoente "1" indica a transposta.

É importante notar que  $B_1$  é resultante de uma correção aplicada em  $B_0$  (Fig. 3). A derivada direcional,  $-G_0^T H_0^{-1} G_0$ , de  $S$  em  $B_0$ , na direção do vetor correção, será negativa se  $H_0^{-1}$  for uma matriz definida positiva e  $G_0 \neq 0$ . Nestas condições, e levando-se em conta que a correção possui o gradiente,  $B_1$  será calculado na direção, a partir de  $B_0$ , em que  $S$  decresce mais rapidamente.

#### 3. Verificar se $S(B_1) < S(B_0)$

Esta desigualdade, se verdadeira, garante que a estimativa dos parâmetros obtida no passo 2 é melhor do que  $B_0$ . Isso representa uma condição para que o processo possa ser continuado. Quando esta condição falha, uma hipótese é que  $B_1$  te-

nha sido obtido por uma correção muito grande em  $B_0$ . Neste caso, basta definir:

$$B_1 = B_0 - \epsilon H_0^{-1} G_0$$

e procurar um escalar  $\epsilon$ , com  $0 < \epsilon < 1$ , de modo que  $S(B_1)$  seja menor do que  $S(B_0)$ . Entretanto, o método de Newton, apesar de teoricamente excelente, não se revela da mesma forma na prática, pois  $H_0^{-1}$  não é necessariamente definida positiva, exceto nas proximidades do mínimo de  $S$ . Mesmo que a inversa da Hessiana seja definida positiva, a acumulação de erros de arredondamento pode gerar  $B_1$ , numa direção inconveniente. Além disso, outra restrição ao método de Newton é a exigência, um tanto laboriosa, do cálculo de derivadas segundas. Essa pode ser, inclusive, a causa dos erros de arredondamento, que levam à divergência do processo. São discutidas, adiante, duas modificações para superar essas dificuldades, procurando manter as qualidades do método de Newton.

4. *Testar a finalização do processo iterativo, isto é, verificar se foi atingida a estimativa de mínimos quadrados B de B. Caso contrário, retornar ao passo 1 com B<sub>0</sub> = B<sub>1</sub> para nova iteração.*

A rigor, no final do procedimento deve-se ter a anulação do gradiente de  $S$ , o que dificilmente ocorre na prática. Portanto, é necessário um critério de interrupção das iterações, como por exemplo, parar quando o valor absoluto de cada componente do vetor correção  $B_1 - B_0$  for menor do que uma quantidade pré-fixada. Isto é, parar quando

$$|b_k - b_k^*| < \epsilon_k \text{ para } k = 1, 2, \dots, r, \text{ e aceitar } \hat{B} = B_1.$$

A seguir são discutidos dois métodos (4, 7, 9, 10) que, na realidade, são duas

modificações no passo 2 do método de Newton. Um é o método clássico de Gauss-Newton, que adota uma matriz  $H$  de elementos.

$$h_{km} = 2 \sum_{i=1}^n \frac{\partial f}{\partial b_k} \frac{\partial f}{\partial b_m} \quad (k, m = 1, 2, \dots, r)$$

Portanto, são simplesmente eliminados, nos elementos da matriz Hessiana, os termos que contêm as derivadas parciais de segunda ordem. Esses termos possuem os resíduos  $y_i - f(x_i)$  os quais, nas proximidades do mínimo de  $S$ , devem ser pequenos e, aí, o método de Gauss se aproxima do método de Newton.

Outro é o método de MARQUARDT (9) que modifica os elementos da matriz  $H$ , do método de Gauss, da seguinte forma:

$$h_{kk} = \begin{cases} (1 + \lambda)h_{kk} & (\text{se } h_{kk} \neq 0) \\ \lambda & (\text{se } h_{kk} = 0) \end{cases}$$

Agora, pode-se obter sempre uma matriz  $H^{-1}$  definida positiva, requerida no passo 2, escolhendo-se  $\lambda$  suficientemente grande. MARQUARDT (9) sugere iniciar a iteração com  $\lambda$  pequeno, por exemplo  $\lambda = 0,01$ , ou seja, iniciar com uma matriz  $H$

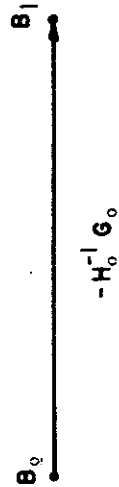


FIG. 3 — O vetor correção de Newton:  $-H_0^{-1} G_0$

próxima da que seria usada no método de Gauss e se, em uma iteração, a condição do passo 3 falhar, aumentar  $\lambda$  sucessivamente, até que a condição se verifique.

Deve-se ressaltar que, no método de Newton, a complexidade da função de ajustamento gera uma acumulação maior

de erros de arredondamento, tornando-se, às vezes, impossível alcançar o mínimo da soma de quadrados dos desvios, mesmo com as modificações introduzidas.

$$H_0(B_1 - B_0) = -G_0$$

por algum processo, sem recorrer à inversa de  $H_0$ . Portanto, o ajustamento de funções não-lineares nos parâmetros, constitui um problema sobre o qual existe, ainda, muito a estudar.

Entretanto, outras alterações podem ser pesquisadas (1, 3) tal como um processo mais adequado de inversão da matriz  $H_0$ ,

MAGNANI, R. & SAHÃO JUNIOR, J. — A discussion about nonlinear parameter estimation. *Ecl. Quím.*, São Paulo, 10:17-21, 1985.

ABSTRACT: An approach of the problem of nonlinear parameter estimation is presented by the algorithmization of the Newton Classic Method, with the discussion of modifications and suggestions for the improvement of the method.

KEY-WORDS: Parameter estimation; nonlinear model; gradient method.

REFERÊNCIAS BIBLIOGRÁFICAS

1. BARD, Y — *Nonlinear parameter estimation*. Nova York, Academic Press, 1974.
2. DRAPER, N.R. & SMITH, H. — *Applied regression analysis*. Nova York, John Wiley, 1981.
3. FLETCHER, R. & POWELL, M.J.D. — *The computer Journal*, 1963, 6, 163.
4. HARTLEY, H.O. — *Technometrics*, 1961, 3, 269.
5. HOFFMANN, R. & VIEIRA, S. — *Análise de regressão*. São Paulo, HUCITEC/USP, 1977.
6. KITRELL, J.R.; MEZAKI, R. & WATSON, C.C. — *Industrial and Engineering Chemistry*, 1965, 57, 18.
7. LEVENBERG, K. — *Quart. Appl. Math.*, 1944, 2, 164.
8. MAGNANI, R. — *Alguns aspectos da aplicação da Segunda Lei de Mitscherlich e da equação de regressão quadrática à maturação da cana-de-açúcar*. Piracicaba, ESALQ/USP, 1985 (Dissertação de Mestrado).
9. MARQUARDT, D.W. — *SIAM J. NUM. ANALYSIS*, 1963, 1, 431.
10. SMITH, F.B. & SHANNO, D.F. — *Technometrics*, 1971, 13, 63.

Recebido em 11.07.85