EP/T 2 - Ajuste de uma exponencial por mínimos quadrados

Este trabalho envolve ajuste linear por mínimos quadrados, Método de Newton e integração de equações diferenciais por Runge-Kutta

É muito comum encontrar, na natureza, uma variável, digamos y, que depende de forma exponencial de outra, digamos x:

$$y = ae^{bx} \ .$$

Se no modelo teórico presumir-se que a relação entre as variáveis é essa, procura-se encontrar, com base em dados experimentais $(x_i, y_i)_{i=1,\ldots,N}$, os valores "ideais" de a e b, segundo algum critério razoável.

Neste curso, o critério é minimizar a diferença quadrática

$$Q(a,b) = \sum_{i=1}^{N} (y_i - ae^{bx_i})^2 \ ,$$

lembrando que, no entanto, a dependência de ae^{bx} não é linear.

Abordaremos quatro métodos para encontrar a e b, que devem ser implementados por um programa de computador.

1. Este método não encontra o par (a,b) que minimiza $Q(a,b)$, mas é amplamente utilizado pela facilidade de aplicá-lo. Veja que a relação $y = ae^{bx}$ pode ser reescrita como

$$\log y = \log a + bx \ ,$$

onde log é o logaritmo natural. Isto significa que a variável $z = \log y$ tem uma dependência afim com relação a x, e se chamarmos $A = \log a$, a dependência passa a ser linear com os parâmetros A e b. Portanto pode-se ajustar uma reta aos dados $(x_i, \log y_i)$, obtendo-se o par (a_0, b_0) que minimiza a diferença quadrática

$$\widetilde{Q}(a,b) = \sum_{i=1}^{N} (\log y_i - \log a - bx_i)^2 \ .$$

Como $Q(a,b)$ e $\widetilde{Q}(a,b)$ são funções diferentes, não se espera que tenham mínimo no mesmo ponto. Portanto $Q(a_0, b_0)$ não é o mínimo valor possível para Q, como deve ser conferido quando o programa for implementado com este e com os demais métodos.

Usaremos os valores de (a_0, b_0) encontrados neste primeiro método para dar as condições iniciais nos métodos seguintes.

2. No segundo método, a idéia é explorar a linearidade de $g(x) = ae^{bx}$ em a. O algoritmo é: escolhe-se um valor b_0 para b (pode ser aquele do método anterior) e fixa-se a função $f(x) = e^{b_0 x}$. Faz-se então um ajuste linear (a um parâmetro) para $g(x) = af(x)$, obtendo-se um valor a_1. Agora procura-se atualizar o valor de b_0 para b_1 fixando-se o valor de a_1 e procurando o mínimo, em b, para a diferença quadrática

$$Q_{a_1}(b) = \sum_{i=1}^{N} (y_i - a_1 e^{bx_i})^2 \ .$$

Acontece que esse mínimo pode ser encontrado resolvendo-se a equação $Q'_{a_1}(b) = 0$, problema para o qual pode ser usado o Método de Newton. Encontrado b_1 que satisfaz essa equação, fixa-se $f(x) = e^{b_1 x}$ e ajusta-se $g(x) = af(x)$ por mínimos quadrados para se obter a_2. E assim por diante, até parar por algum critério de convergência.

3. O terceiro método é parecido com o segundo, mas é mais direto. No mínimo de Q vale que as derivadas parciais com relação a a e b são nulas (embora isto não seja suficiente para garantir que é mínimo). A equação $\frac{\partial Q}{\partial a} = 0$ leva à relação

$$a = \frac{\sum y_i e^{bx_i}}{\sum e^{2bx_i}} . \tag{1}$$

Já $\frac{\partial Q}{\partial b} = 0$ leva a

$$a \left(\sum x_i e^{2bx_i} \right) = \sum x_i y_i e^{bx_i} \tag{2}$$

Substituindo (1) em (2), sai

$$F(b) = (\sum y_i e^{bx_i}) \cdot (\sum x_i e^{2bx_i}) - (\sum x_i y_i e^{bx_i}) \cdot (\sum e^{2bx_i}) . \tag{3}$$

Basta então aplicar o Método de Newton para achar b, e depois usar (1) ou (2) para achar a.

4. Se (a_*, b_*) é mínimo de $Q(a,b)$, então $\frac{\partial Q}{\partial a} = 0 = \frac{\partial Q}{\partial b}$. Isto é o mesmo que dizer que (a_*, b_*) é singularidade do (negativo do) campo gradiente $X(a,b) = (-\frac{\partial Q}{\partial a}(a,b), -\frac{\partial Q}{\partial b}(a,b))$. Além disso, (a_*, b_*) é singularidade atratora do campo X. Então o mínimo pode ser achado integrando-se X por algum método de integração de equações diferenciais.

Exercício-programa.

- Ler os dados (x_i, y_i) de um arquivo-texto, no formato "x_i espaço y_i quebra de linha".
- Calcular (a_0, b_0) pelo método 1, e a diferença quadrática $Q(a_0, b_0)$.
- Obter o ponto de mínimo de $Q(a,b)$ pelos métodos 2, 3 e 4, verificando se são iguais e comparando a diferença quadrática obtida com aquela obtida pelo primeiro método. Nos três últimos métodos, usar os valores obtidos no primeiro método como condição inicial.