LISTA DE EXERCÍCIOS - AJUSTE DE MÍNIMOS QUADRADOS
Cálculo Numérico para Geociências - 2009 - Prof. Eduardo Colli

Em todos os casos, convencionamos ter um conjunto de dados (x_i, y_i), com $i = 1, \ldots, N$. Faça o gráfico dos pontos dados em todos os exercícios, plote a função ajustada e calcule a diferença quadrática. Use sempre 5 algarismos significativos de precisão em todas as contas.

1. O ajuste por mínimos quadrados da família a um parâmetro $f(x) = c$ leva a que resultado para c? (Já falei na aula...)

2. A dependência $y = ax$ pode surgir em diversas situações. Dois exemplos: (i) x é o tempo e y é o deslocamento, para um movimento retilíneo uniforme, com $y = 0$ quando $x = 0$; esse tipo de movimento não é tão abstrato assim, basta um corpo em queda livre com uma força de resistência proporcional à velocidade, pois rapidamente o corpo atingirá a velocidade de equilíbrio que iguala as duas forças opostas (gravidade e resistência); (ii) solte uma bola de borracha de uma altura h_i, deixe-a bater no chão e subir, até a altura máxima h_f; se $x = h_i$ e $y = h_f$ então pode-se constatar experimentalmente que a dependência é aproximadamente $y = ax$.

Ajuste $f(x) = ax$ por mínimos quadrados aos dados fictícios

x_i	y_i
0.0	0.13
1.1	1.55
2.3	3.94
2.9	5.01
3.7	5.93
5.5	9.30

3. Agora suponha uma queda livre sem resistência do ar, partindo do repouso no instante $x = 0$ (x mede o tempo); a evolução da altura do corpo é medida por y, com y crescente de cima para baixo, e com $y = 0$ no ponto onde o corpo é solto. Neste caso, $y = ax^2$, onde $a = \frac{g}{2}$. Ajuste $f(x) = ax^2$

por mínimos quadrados e ache g.

x_i	y_i
0.0	0.05
1.1	5.73
2.3	26.3
2.9	41.2
3.7	66.9
5.5	149.1

4. Uma bola de naftalina exposta ao ar perde massa constantemente. A evolução do tamanho da bola pode ser modelada da seguinte forma. Primeiro supomos que o formato é sempre esférico e que o material tem densidade homogênea. Assim o tamanho da bola é acompanhado pelo seu raio em função do tempo: $r(t)$. Sua massa $m(t)$ é dada por $\frac{4\rho}{3}\pi r(t)^3$, onde ρ é a densidade. A segunda hipótese é que a perda de massa se dá proporcionalmente à área da superfície, que é $4\pi r(t)^2$. Então

$$\dot{m}(t) = -\alpha \cdot 4\pi r(t)^2 \,.$$

Por outro lado, derivando a expressão de $m(t)$, obtemos

$$\dot{m}(t) = 4\pi r(t)^2 \dot{r}(t) \,.$$

Igualando-se as duas, resulta $\dot{r}(t) = -\alpha$. Isto quer dizer que o raio da bolinha diminui a uma taxa constante. Em outras palavras,

$$r(t) = r_0 - \alpha t \,,$$

que é uma função *afim* de t. Suponha que a naftalina foi deixada ao ar livre às 12h do dia 01 de julho, e foram medidos os seguintes dados de r (em centímetros) e de t (em dias). Faça o ajuste de mínimos quadrados e responda: qual era o raio inicial da bola, aproximadamente? Quando ela desaparecerá?

t_i	r_i
7.	5.16
9.	5.05
10.	5.25
14.	5.20
17.	5.10
23.	4.58
30.	4.20
40.	3.97

5. Ajuste $f(x) = ax + bx^2$ aos seguintes dados.

x_i	y_i
1.	1.5
2.	4.6
3.	5.8
4.	7.8
5.	9.2
6.	10.2
7.	9.8

6. Ajuste um polinômio cúbico aos seguintes dados.

x_i	y_i
-3.1	-1.4
-2.4	1.6
-1.1	2.3
0.1	0.65
1.4	1.7
2.8	9.0
3.2	15.3

7. Se a evolução temporal de $y(t)$ parece ter um 'drift' do tipo $a_0 + a_1 t$, mais uma oscilação senoidal de período 1, é razoável fazer um ajuste pela família

$$f(t) = a_0 + a_1 t + a_2 \sin(2\pi t) + a_3 \cos(2\pi t)\,.$$

Faça esse ajuste para os dados seguintes e responda: qual é a amplitude da oscilação em torno da reta média?

t_i	y_i
0.0	-0.15
0.2	-0.44
0.4	0.84
0.6	1.51
0.8	1.26
1.0	0.14
1.2	0.35
1.4	2.13
1.6	2.40

Análise harmônica. Se os dados apresentam indício de periodicidade T, mas T não é um parâmetro ajustável, a análise harmônica consiste em minimizar a diferença quadrática com funções que são combinações lineares de funções

$$\cos\left(k\frac{2\pi}{T}x\right)\ ,\ \sin\left(k\frac{2\pi}{T}x\right)\ ,$$

com $k \geq 0$ (com $k = 0$, a função cosseno é simplesmente a função constante igual a 1 e a função seno é a função nula, que deve ser ignorada). Todas elas são T periódicas, mas para um k fixo as funções (cosseno e seno) são, de fato (também) $\frac{T}{k}$-periódicas.

A ortogonalidade tem um papel importante quando se trata de facilitar os cálculos, mas neste caso há uma *hipótese fundamental: o número de pontos N tem que ser par, digamos $N = 2m$, e os dados devem ter espaçamento regular igual a $\frac{T}{N}$*. Dizemos que f e g são ortogonais para esses pontos se

$$\langle f, g\rangle = \sum_{i=1}^{N} f(x_i)g(x_i) = 0\,.$$

É possível mostrar que, se vale a hipótese acima sobre os pontos, então são ortogonais entre si as N seguintes funções:

$$\begin{array}{cccccc} 1 & \cos(\frac{2\pi}{T}x) & \cos(2\cdot\frac{2\pi}{T}x) & \ldots & \cos((m-1)\frac{2\pi}{T}x) & \cos(m\frac{2\pi}{T}x) \\ & \sin(\frac{2\pi}{T}x) & \sin(2\cdot\frac{2\pi}{T}x) & \ldots & \sin((m-1)\frac{2\pi}{T}x) & \end{array}$$

Daí resulta que a matriz de coeficientes do sistema linear em que recai o ajuste de mínimos quadrados é *diagonal* e de fácil solução, para qualquer família de funções que seja uma combinação linear de um sobconjunto dessas funções. Para sabermos o que aparece na diagonal, temos que saber os produtos internos dessas funções por elas mesmas. Eles são

$$\begin{array}{cccccc} N & \frac{N}{2} & \frac{N}{2} & \ldots & \frac{N}{2} & N \\ & \frac{N}{2} & \frac{N}{2} & \ldots & \frac{N}{2} & \end{array}$$

Um *ajuste harmônico de ordem k* $(k < m)$ é definido como a combinação linear (que minimiza a diferença quadrática) das $2k+1$ funções

$$\begin{array}{ccccc} 1 & \cos(\frac{2\pi}{T}x) & \cos(2\cdot\frac{2\pi}{T}x) & \ldots & \cos(k\frac{2\pi}{T}x) \\ & \sin(\frac{2\pi}{T}x) & \sin(2\cdot\frac{2\pi}{T}x) & \ldots & \sin(k\frac{2\pi}{T}x) \end{array}$$

Os respectivos parâmetros podem ser denotados por

$$\begin{array}{ccccc} a_0 & a_1 & a_2 & \ldots & a_k \\ & b_1 & b_2 & \ldots & b_k \end{array}$$

O *harmônico de ordem* j é o pedaço

$$a_j \cos(j\frac{2\pi}{T}x) + b_j \sin(j\frac{2\pi}{T}x)$$

que pode ser escrito na forma

$$A_j \sin\left(j\frac{2\pi}{T}(x - \beta_j\right) ,$$

onde

$$A_j = \sqrt{a_j^2 + b_j^2}$$

é a *amplitude* desse harmônico e

$$\beta_j = \frac{T}{2\pi j} \arctan(-a_j/b_j)$$

é sua fase.

8. Faça o ajuste harmônico de segunda ordem para os seguintes dados, supondo período $T = 3$. Calcule amplitudes e fases. O que dá o ajuste somente até primeira ordem? Compare as diferenças quadráticas obtidas nos dois casos. O que é preciso calcular para fazer o ajuste até terceira ordem?

t_i	y_i
0.0	0.47
0.3	0.51
0.6	-0.061
0.9	-1.32
1.2	-0.18
1.5	2.19
1.8	4.7
2.1	4.0
2.4	2.6
2.7	1.2
3.0	0.94

Não lineares. Um truque. Suponha uma relação esperada entre x e y dada por $y = f_{a_0,\ldots,a_k}(x)$ que não seja linear nos parâmetros $a_0, \ldots, a_k$. Para exemplificar, usaremos

$$y = \frac{a_0}{1 + a_1 x} .$$

O ajuste por mínimos quadrados pode ser formulado, mas achar $(a_0, a_1, \ldots, a_k)$ que minimiza a diferença quadrática não recai mais num sistema linear. Simplesmente há que se achar o mínimo dessa função com outros métodos.

O truque aqui é mudar o objetivo. Como? Transforme a relação $y = f_{a_0,\ldots,a_k}(x)$ numa relação linear equivalente do tipo

$$h(x,y) = b_0 h_0(x,y) + \ldots + b_m h_m(x,y)\,,$$

onde os novos parâmetros $b_0, \ldots, b_m$ dependem dos anteriores, resolva o sistema linear

$$\begin{pmatrix} \langle h_0, h_0\rangle & \cdots & \langle h_0, h_m\rangle \\ \vdots & \ddots & \vdots \\ \langle h_m, h_0\rangle & \cdots & \langle h_m, h_m\rangle \end{pmatrix} \begin{pmatrix} b_0 \\ \vdots \\ b_m \end{pmatrix} = \begin{pmatrix} \langle h, h_0\rangle \\ \vdots \\ \langle h, h_m\rangle \end{pmatrix}$$

e depois recupere os $a_0, \ldots, a_k$ a partir dos $b_0, \ldots, b_m$. Por exemplo,

$$y = \frac{a_0}{1 + a_1 x}$$

é equivalente (exceto pelos problemas causados quando o denominador se anula) a

$$y = a_0 - a_1 x y\,.$$

Então

$$h(x,y) = y\,,\ h_0(x,y) = 1\,,\ h_1(x,y) = xy\,,\ b_0 = a_0\,,\ b_1 = -a_1\,.$$

Temos que resolver o sistema linear

$$\begin{pmatrix} N & \sum x_i y_i \\ \sum x_i y_i & \sum x_i^2 y_i^2 \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum x_i y_i^2 \end{pmatrix}$$

Depois de achados b_0 e b_1, lembrar que $a_0 = b_0$ e $a_1 = -b_1$.

Note que se houver parâmetros $(a_0, a_1, \ldots, a_k)$ tais que $y_i = f_{a_0,\ldots,a_k}(x_i)$ para todo $i = 1, \ldots, N$ (uma interpolação, de fato), então a relação linear que daí deriva também será válida para todos os pontos. Neste caso, a diferença quadrática do problema linear será minimizada (e igual a zero) pelo mesmo conjunto de parâmetros que a diferença quadrática do problema não linear original. Em geral, no entanto, as soluções diferem um pouco.

9. Ajuste para $y = \frac{a_0}{1+a_1 x}$.

x_i	y_i
-1	-8.8
0	-1.64
1	-0.829
2	-0.97
3	-0.488
4	-0.35
5	-0.555

10. Ajuste para $y = 1 + ax$.

x_i	y_i
0.0	1.13
1.1	2.55
2.3	4.94
2.9	6.01
3.7	6.93
5.5	10.30

11. Ajuste para $y = ae^{bx}$ (exponenciais). Dica: tomar logaritmos nos dois lados.

x_i	y_i
0.0	4.60
0.3	2.66
0.6	1.50
0.9	1.13
1.2	0.53
1.5	0.27

12. Ajuste para $y = ax^b$. Dica: logaritmos de novo.

x_i	y_i
0.0	0.00017
0.5	0.0076
1.0	0.020
1.5	0.029
2.0	0.041
2.5	0.074

13. Ajuste para $y = ax^b$. Dica: logaritmos de novo.

x_i	y_i
0.5	1.82
1.0	0.340
1.5	0.112
2.0	0.0793
2.5	0.0403
3.0	0.0328

14. Ajuste para $y = \frac{x}{a+bx+cx^2}$.

x_i	y_i
0.5	0.285
1.0	0.322
1.5	0.308
2.0	0.327
2.5	0.222
3.0	0.257

Ajuste linear com pesos. É frequente termos dados em que a ordenada y_i é acompanhada de uma estimativa de erro. Essa estimativa de erro não é um intervalo de certeza, mas sim um suposto desvio-padrão σ_i daquele dado. A rigor, esse σ_i deveria ser obtido da seguinte forma. O dado x_i tem que ser preciso (pelo menos muito mais preciso do que o dado y_i), e, para um mesmo x_i, deve-se tomar várias medidas de y. O valor de y_i será a média dessas medidas e σ_i será o desvio-padrão. Normalmente, no entanto, isso não é feito. Toma-se uma única medida e estima-se o erro por argumentos heurísticos.

Uma vez dados os erros é preciso isso levar em conta na diferença quadrática. Pensemos assim. Uma mesma diferença $y_i - f(x_i)$ é um "erro" mais grave se σ_i for pequena do que se σ_i for grande. Então, para levar isso em conta, tomamos σ_i como a "unidade" da diferença, como se introduzíssemos uma escala em y diferente para cada ponto, em que σ_i corresponda a uma unidade. Portanto, o erro passa a ser medido por

$$\frac{y_i - f(x_i)}{\sigma_i}$$

e a diferença quadrática passa a ser

$$Q = \sum_{i=1}^{N} \frac{(y_i - f(x_i))^2}{\sigma_i^2} \,.$$

Agora pode-se desenvolver a mesma teoria para ajuste de famílias lineares nos parâmetros $f(x) = a_0 g_0(x) + \ldots + a_k g_k(x)$, para essa diferença quadrática "adaptada". O resultado será o mesmo sistema linear que antes, exceto que os "produtos internos" terão que ser definidos por

$$\langle u, v \rangle = \sum_{i=1}^{N} \frac{u_i v_i}{\sigma_i^2} \,,$$

para $u = (u_1, \ldots, u_N)$ e $v = (v_1, \ldots, v_N)$.

15. Ajuste um polinômio quadráticos aos dados, com barras de erro distintas.

x_i	y_i	σ_i
$-2.$	-7.0	0.3
$-1.$	-5.7	0.2
$0.$	-1.3	0.1
$1.$	0.11	0.03
$2.$	-0.17	0.03

16. Convença-se de que, se todos os desvios-padrão forem iguais, então é indiferente considerá-los para minimizar a diferença quadrática.