

ANEXO A – Princípios básicos de regressão não-linear

Qual a curva que melhor se aproxima dos pontos (x_i, y_i) ? Faz sentido fazer uma regressão linear neste exemplo? Como determinar a equação de uma função **não-linear** que se aproxime destes pontos?

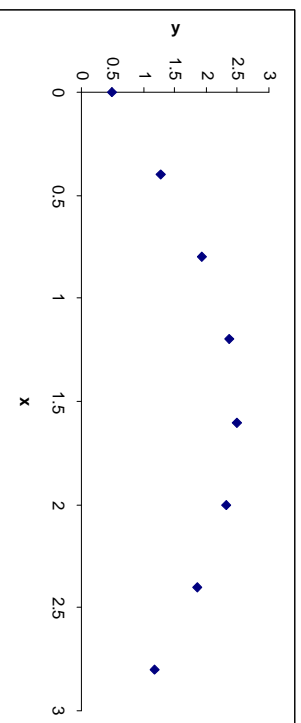
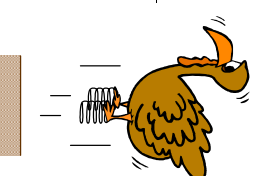


Figura A.1 – Pontos (x, y) descrevendo uma evolução não linear

Vamos admitir que a função de aproximação é um polinômio do 2º grau,

$$f(x) = a x^2 + b x + c$$

(A.1)

Quais os parâmetros **a**, **b** e **c**, que minimizam o erro de aproximação?

Critério dos Mínimos Quadrados

O erro cometido na aproximação em cada ponto **i** é dado pela diferença

$$e_i = f(x_i) - y_i$$

(A.2)

O erro quadrático total E_{qt} é dado por

$$E_{qt} = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (f(x_i) - y_i)^2$$

(A.3)

Critério: os parâmetros **a**, **b** e **c** são determinados de modo a minimizar E_{qt} .
...também pode ser aplicado com outras funções $f(x)$ (como sinusoidais, logarítmicas, exponenciais, etc) desde que sejam deriváveis.

E como determinar **a**, **b** e **c** de modo a minimizar E_{qt} ? Usando a eq. A.1 em A.3 resulta:

$$E_{qt} = \sum_{i=1}^n (a x_i^2 + b x_i + c - y_i)^2 = \min \quad (A.4)$$

Para minimizar A.4 vamos calcular a 1ª derivada de E_{qt} relativamente aos parâmetros **a**, **b** e **c** e igualar a zero. Desta operação resulta:

$$\begin{cases} \frac{\partial E_{qt}}{\partial a} = 2 \sum_{i=1}^n x_i^2 (a x_i^2 + b x_i + c - y_i) = 0 \\ \frac{\partial E_{qt}}{\partial b} = 2 \sum_{i=1}^n x_i (a x_i^2 + b x_i + c - y_i) = 0 \\ \frac{\partial E_{qt}}{\partial c} = 2 \sum_{i=1}^n (a x_i^2 + b x_i + c - y_i) = 0 \end{cases} \quad (A.5)$$



Expandindo as eq A.5 resulta

$$\begin{cases} a \sum_{i=1}^n x_i^4 + b \sum_{i=1}^n x_i^3 + c \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i^2 y_i \\ a \sum_{i=1}^n x_i^3 + b \sum_{i=1}^n x_i^2 + c \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i \\ a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i + c \sum_{i=1}^n 1 = \sum_{i=1}^n y_i \end{cases} \quad (A.6)$$

... bastando agora resolver o sistema (A.6) e determinando os parâmetros **a**, **b** e **c**.

A dimensão do sistema a resolver depende do n° de incógnitas. Se o polinómio for, por exemplo, de 3º grau teremos um sistema de 4 equações e 4 incógnitas.

Sugestão 1: A equação (2) pode ser rescrita da seguinte forma:

$$T = \frac{2\pi}{\sqrt{k}} \sqrt{m} = a \sqrt{m} \quad (A.9)$$

em que **a** é uma constante ($a = 2\pi / \sqrt{k}$). Sendo a equação (A.9) do tipo

$$f(x) = a \sqrt{x} \quad (A.10)$$

faz sentido tentar aproximar os pontos (m,T), obtidos experimentalmente, por uma curva deste género. Seguindo um procedimento semelhante ao anterior (polinómio de 2º grau), mostre que o parâmetro **a** é dado por

$$a = \frac{\sum_{i=1}^n \sqrt{x_i} y_i}{\sum_{i=1}^n x_i} \quad (A.11)$$

Sugestão 2: Imagine agora que pretende determinar a curva

$$f(x) = a\sqrt{x} + b \quad (A.12)$$

que melhor se aproxima dos pontos (x,y) segundo o critério dos mínimos quadrados.

Segundo um procedimento semelhante ao adoptado para o polinómio de 2º grau, mostre que **a** e **b** são determinados pelo seguinte sistema de equações:

$$\begin{cases} a \sum_{i=1}^n x_i + b \sum_{i=1}^n \sqrt{x_i} = \sum_{i=1}^n \sqrt{x_i} y_i \\ a \sum_{i=1}^n \sqrt{x_i} + b n = \sum_{i=1}^n y_i \end{cases} \quad (A.13)$$

ANEXO B – Usando o EXCEL para regressão não-linear

Exemplo: A Tabela B.1 representa medidas experimentais da distância (S) percorrida por um móvel em função do tempo (t). Pretende-se determinar a equação $S=f(t)$ que melhor¹ aproxima o conjunto de pontos. Neste exemplo, vamos considerar que a função f é do tipo polinomial de 2º grau, isto é, do tipo

$$S = at^2 + b t + c \quad (B.1)$$

onde **a**, **b** e **c** representam os parâmetros (coeficientes) que se pretende determinar.

Tabela B.1 – Espaço versus tempo

t (s)	1.0	3.0	5.0	7.0	9.0	11.0	13.0	15.0	17.0	19.0	21.0
S (m)	17.6	25.8	36.7	50.6	66.5	92.0	112.5	141.4	176.5	220.8	254.6

A Fig. B.1 (página seguinte) mostra que a relação entre S e t é claramente não-linear. Isto significa que não é possível aproximar estes pontos por uma recta sem cometer erros grosseiros.

¹ "Melhor" depende obviamente do critério adoptado. Neste caso, vai ser usado o Critério dos Mínimos Quadrados que se baseia na minimização do quadrado dos desvios (diferença entre os valores a aproximar e os valores correspondentes na função de aproximação).

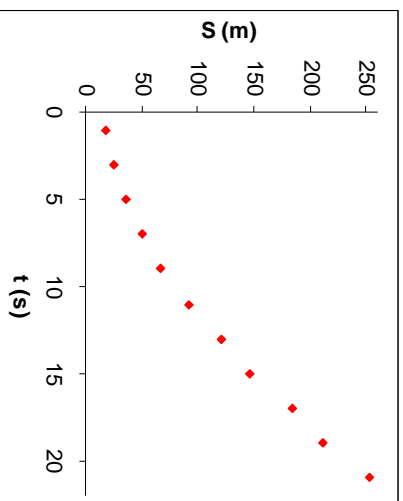


Figura B.1 – Espaço versus tempo

Para determinar os parâmetros **a**, **b** e **c** da equação (B.1) proceda do seguinte modo:

a. Coloque o rato sobre um dos pontos do gráfico e pressione o botão direito. Deve aparecer uma caixa de diálogo onde deve seleccionar “**Add trendline...**”. Aparece um novo quadro com a mesma designação. Este quadro tem 2 sub-menus: **Type** e

Options;

b. No sub-menu **Type**, seleccione a opção **Polynomial** com ordem (**Order**) igual a 2. Significa que vamos aproximar os pontos dados por um polinómio de 2º grau;

c. No sub-menu **Options** torne activa a opção “**Display equation on chart**”. Esta opção faz com que a equação da curva de aproximação seja também visualizada no próprio gráfico;

d. Prima **OK**.

Em resultado desta operação, é adicionada ao gráfico a curva do tipo (B.1) que melhor aproxima os pontos em questão, bem como a respectiva equação. A Fig. B.2 ilustra o resultado obtido.

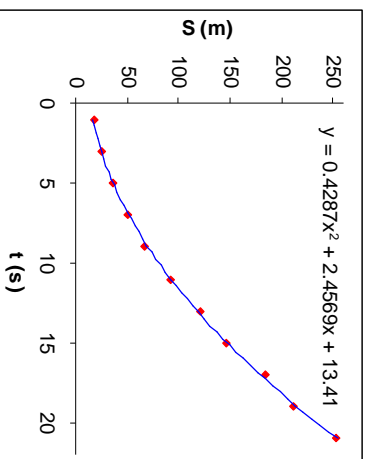


Figura B.2 – Curva (polinómio de 2º grau) que melhor aproxima os pontos dados