



Non-linear Regression - Logarithmic Curve

The process of producing a mathematical function that provides the best fit to a series of experimental data points.

The equation of the logarithmic curve: $y = a + b \cdot \ln(x)$

Experimental data

Raw data variables (Independent variables) (dependent variables)

$data :=$	$\begin{bmatrix} 1 & 47.3 \\ 2 & 50.0 \\ 3 & 54.1 \\ 4 & 59.7 \\ 5 & 62.9 \\ 6 & 68.2 \\ 7 & 69.7 \\ 8 & 70.8 \\ 9 & 73.7 \\ 10 & 75.4 \\ 11 & 76.8 \end{bmatrix}$	$x := \text{col}(data, 1) =$	$\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \\ 11 \end{bmatrix}$	$y_d := \text{col}(data, 2) =$	$\begin{bmatrix} 47.3 \\ 50.0 \\ 54.1 \\ 59.7 \\ 62.9 \\ 68.2 \\ 69.7 \\ 70.8 \\ 73.7 \\ 75.4 \\ 76.8 \end{bmatrix}$
-----------	--	------------------------------	---	--------------------------------	--

Total number of data points $n := \text{length}(x) = 11$

finding the logs of the data will allow a linear relationship between x and y to be used

Mean values for x, y, log(x) and log(y)

$$x_{mean} := \sum_{i=1}^n x_i \cdot n^{-1} = 6$$

$$y_{mean} := \sum_{i=1}^n y_{d_i} \cdot n^{-1} = 64.4182$$

$$\ln x_{mean} := \sum_{i=1}^n \ln(x_i) \cdot n^{-1} = 1.5911$$

$$\ln y_{mean} := \sum_{i=1}^n \ln(y_{d_i}) \cdot n^{-1} = 4.1528$$

Least squares fitting - Sum of squares

$$SS_{\ln x} := \sum_{i=1}^n \ln(x_i) = 17.5023$$

$$SS_{\ln y} := \sum_{i=1}^n \ln(y_{d_i}) = 45.6808$$

$$SS_{y \ln x} := \sum_{i=1}^n y_{d_i} \cdot \ln(x_i) = 1202.8488$$

$$SS_{x \ln sq} := \sum_{i=1}^n \ln(x_i)^2 = 33.4001$$

$$SS_x := \sum_{i=1}^n x_i = 66$$

$$SS_y := \sum_{i=1}^n y_{d_i} = 708.6$$

$$SS_{x^2 y} := \sum_{i=1}^n x_i^2 \cdot y_{d_i} = 36466.1$$

$$SS_{xy \ln y} := \sum_{i=1}^n x_i \cdot y_{d_i} \cdot \ln(y_{d_i}) = 19490.7684$$



$$SS_{ylny} := \sum_{i=1}^n y_{d_i} \cdot \ln(y_{d_i}) = 2960.1985$$

$$SS_{xy} := \sum_{i=1}^n x_i \cdot y_{d_i} = 4588.5$$

Non-linear Regression (linearisation of a non-linear function)

$$b := \frac{n \cdot SS_{ylnx} - SS_y \cdot SS_{lnx}}{n \cdot SS_{xlnsq} - SS_{lnx}^2} = 13.5777$$

$$a_0 := lny_{mean} - b \cdot lnx_{mean} = -17.4509$$

Linearisation of
Non-linear function

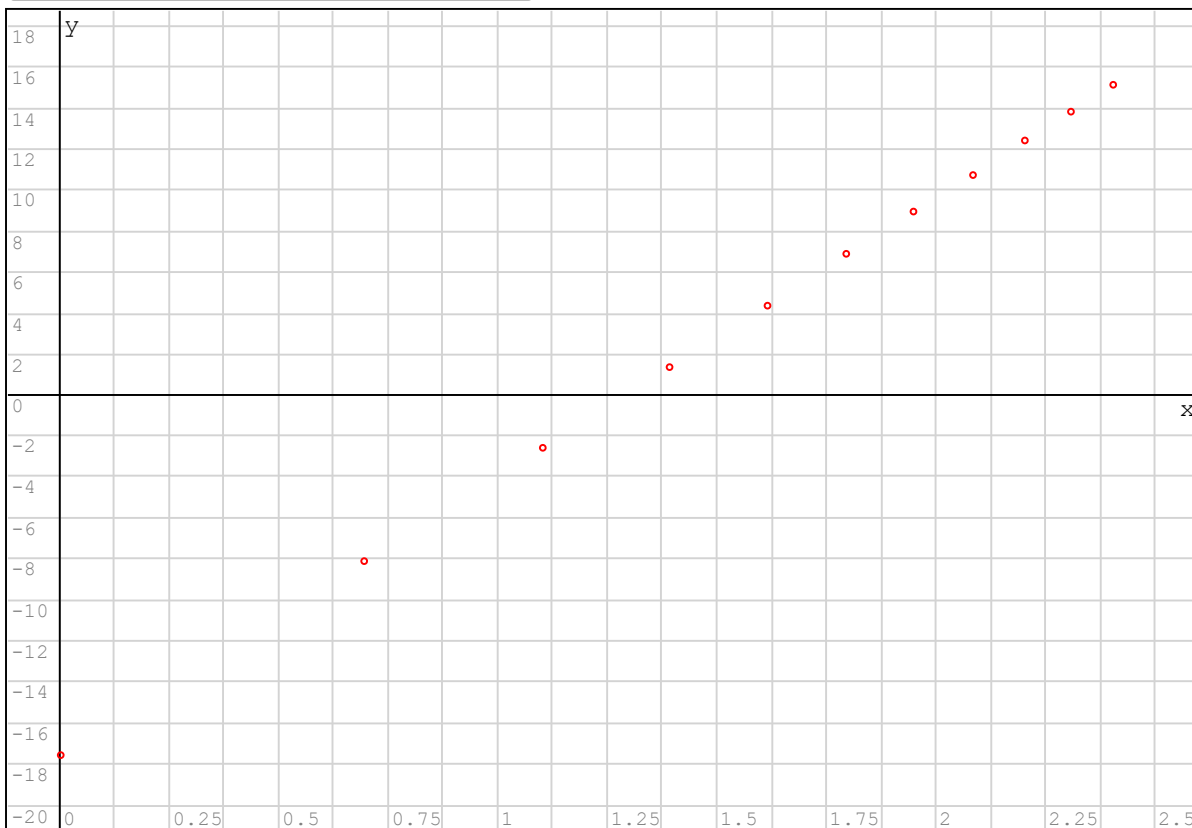
for $i \in [1..n]$
 $x_{ln_i} := \ln(x_i)$

for $i \in [1..n]$
 $y_{ln_i} := a_0 + b \cdot \ln(x_i)$

$$x_{ln} = \begin{bmatrix} 0 \\ 0.6931 \\ 1.0986 \\ 1.3863 \\ 1.6094 \\ 1.7918 \\ 1.9459 \\ 2.0794 \\ 2.1972 \\ 2.3026 \\ 2.3979 \end{bmatrix}$$

$$y_{ln} = \begin{bmatrix} -17.4509 \\ -8.0396 \\ -2.5343 \\ 1.3717 \\ 4.4015 \\ 6.877 \\ 8.97 \\ 10.7831 \\ 12.3823 \\ 13.8129 \\ 15.107 \end{bmatrix}$$

Linearisation of non-linear function
red points show data points





Non-linear Regression (Logarithmic curve)

$$y = a + b \cdot \ln(x)$$

$$a := \frac{SS_y - b \cdot SS_{\ln x}}{n} = 42.8144$$

$$b = 13.5777$$

see linearisation above

Check of results against the experimental data

$$\text{for } i \in [1..n] \\ y_i := a + b \cdot \ln(x_i)$$

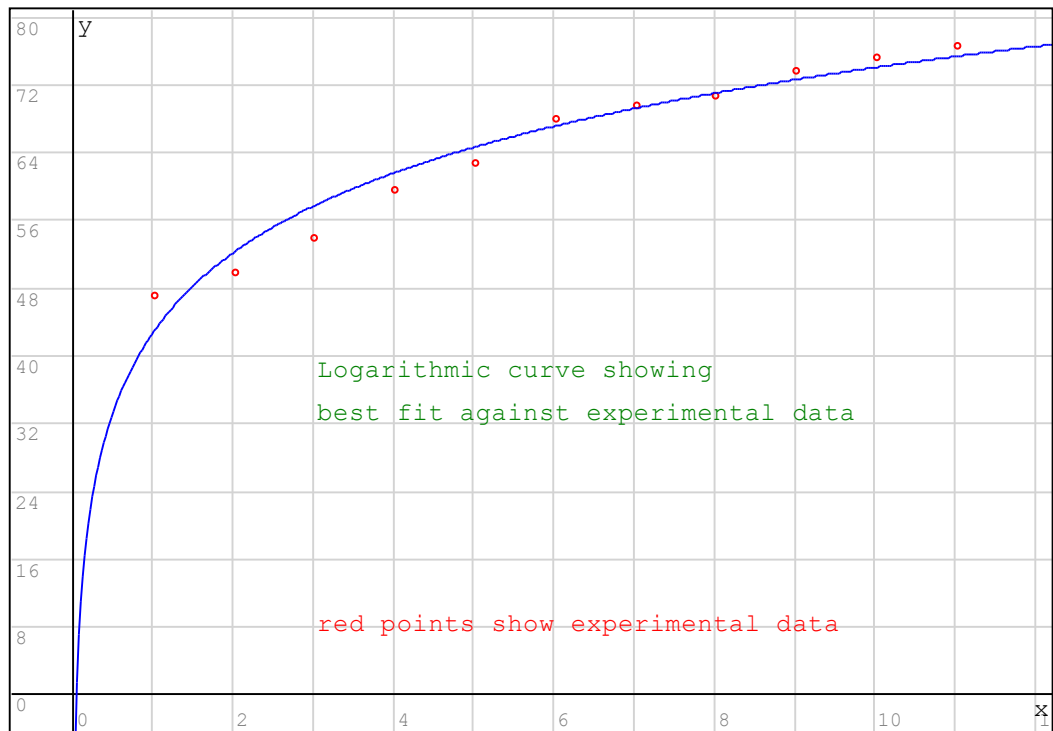
$$x = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ \vdots \end{bmatrix}$$

$$y_d = \begin{bmatrix} 47.3 \\ 50 \\ 54.1 \\ 59.7 \\ 62.9 \\ 68.2 \\ 69.7 \\ 70.8 \\ \vdots \end{bmatrix}$$

Calculated values

$$y = \begin{bmatrix} 42.8144 \\ 52.2258 \\ 57.7311 \\ 61.6371 \\ 64.6669 \\ 67.1424 \\ 69.2354 \\ 71.0485 \\ \vdots \end{bmatrix}$$

$$f(x) := a + b \cdot \ln(x)$$



Sum of squares of the residual

$$SS_{res} := \sum_{i=1}^n \left(y_{d_i} - (a + b \cdot \ln(x_i)) \right)^2$$

Total sum of the squares

$$SS_{tot} := \sum_{i=1}^n \left(y_i - y_{mean} \right)^2$$

R squared variance

$$R_{sq} := \frac{SS_{tot} - SS_{res}}{SS_{tot}} = 0.9498$$

Result for exponential curve

$$y := a + b \cdot \ln(x)$$

$$a = 42.8144$$

$$b = 13.5777$$