



Non-linear Regression - 3rd order Polynomial Curve

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

A 3rd order polynomial cubic regression: $y = a_0 + a_1 \cdot x + a_2 \cdot x^2 + a_3 \cdot x^3$

Experimental data

Raw data variables (Independent variables) (dependent variables)

$data :=$	$\begin{bmatrix} -5 & 58.5 \\ -4 & 37.6 \\ -3 & 19.7 \\ -2 & 11.8 \\ -1 & 6.3 \\ 0 & 3.6 \\ 1 & 7.7 \\ 2 & 13.6 \\ 3 & 27.2 \\ 4 & 40.9 \\ 5 & 61.1 \end{bmatrix}$	$x := col(data, 1) =$	$\begin{bmatrix} -5 \\ -4 \\ -3 \\ -2 \\ -1 \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}$	$y_d := col(data, 2) =$	$\begin{bmatrix} 58.5 \\ 37.6 \\ 19.7 \\ 11.8 \\ 6.3 \\ 3.6 \\ 7.7 \\ 13.6 \\ 27.2 \\ 40.9 \\ 61.1 \end{bmatrix}$
-----------	--	-----------------------	--	-------------------------	---

Summation of experimental data Total number of data points $n := length(x) = 11$

Sum of x values raised to power

Sum of x values multiplied by y values

Sum of y values

for $i \in [1..n]$

for $i \in [1..n]$

$$SS_i := \sum_{j=1}^n x_j^i$$

$$SS_{xy_i} := \sum_{j=1}^n x_j^i \cdot y_{d_j}$$

$$sum_y := \sum_{i=1}^n y_{d_i} = 288$$

$$SS = \begin{bmatrix} 0 \\ 110 \\ 0 \\ 1958 \\ 0 \\ 41030 \\ 0 \\ 9.2596 \cdot 10^5 \\ 0 \\ 2.1749 \cdot 10^7 \\ 0 \end{bmatrix}$$

$$SS_{xy} = \begin{bmatrix} 53.7 \\ 4783.7 \\ 754.5 \\ 99065.3 \\ 13385.7 \\ 2.2261 \cdot 10^6 \\ 2.7383 \cdot 10^5 \\ 5.2178 \cdot 10^7 \\ 6.0917 \cdot 10^6 \\ 1.2531 \cdot 10^9 \\ 1.4213 \cdot 10^8 \end{bmatrix}$$

$$\varphi(x) := \begin{bmatrix} 1 \\ x \\ x^2 \\ x^3 \end{bmatrix}$$

$$k := length(\varphi(x)) = 4$$

Mean value of x variables

Mean value of y variables

$$x_{mean} := \sum_{i=1}^n x_i \cdot n^{-1} = -5.5511 \cdot 10^{-17}$$

$$y_{mean} := \sum_{i=1}^n y_{d_i} \cdot n^{-1} = 26.1818$$



Matrices of least squares fitting - Sum of squares

```
for j ∈ [1..k]
  for i ∈ [1..n]
    Xij := φ(xi)j
```

$$M_1 := X^T \cdot X$$

$$M_2 := X^T \cdot y_d$$

$$M_1 = \begin{bmatrix} 11 & 0 & 110 & 0 \\ 0 & 110 & 0 & 1958 \\ 110 & 0 & 1958 & 0 \\ 0 & 1958 & 0 & 41030 \end{bmatrix}$$

$$M_2 = \begin{bmatrix} 288 \\ 53.7 \\ 4783.7 \\ 754.5 \end{bmatrix}$$

$$coeff := \text{eval}(M_1^{-1}) \cdot M_2$$

Coefficients

$$\begin{cases} a_0 := coeff_1 \\ a_1 := coeff_2 \\ a_2 := coeff_3 \\ a_3 := coeff_4 \end{cases} = \begin{cases} 3.99417 \\ 1.06838 \\ 2.21876 \\ -0.0326 \end{cases}$$

The residuals between the model and the experimental data

$$res := y_d - a_0 - a_1 \cdot x - a_2 \cdot x^2 = \begin{bmatrix} -184.2164 \\ -206.1848 \\ -225.1531 \\ -234.1215 \\ -240.6899 \\ -244.4583 \\ -241.4267 \\ -236.595 \\ -224.0634 \\ -211.4318 \\ -192.3002 \end{bmatrix}$$

Check of results against the experimental data

Calculated values

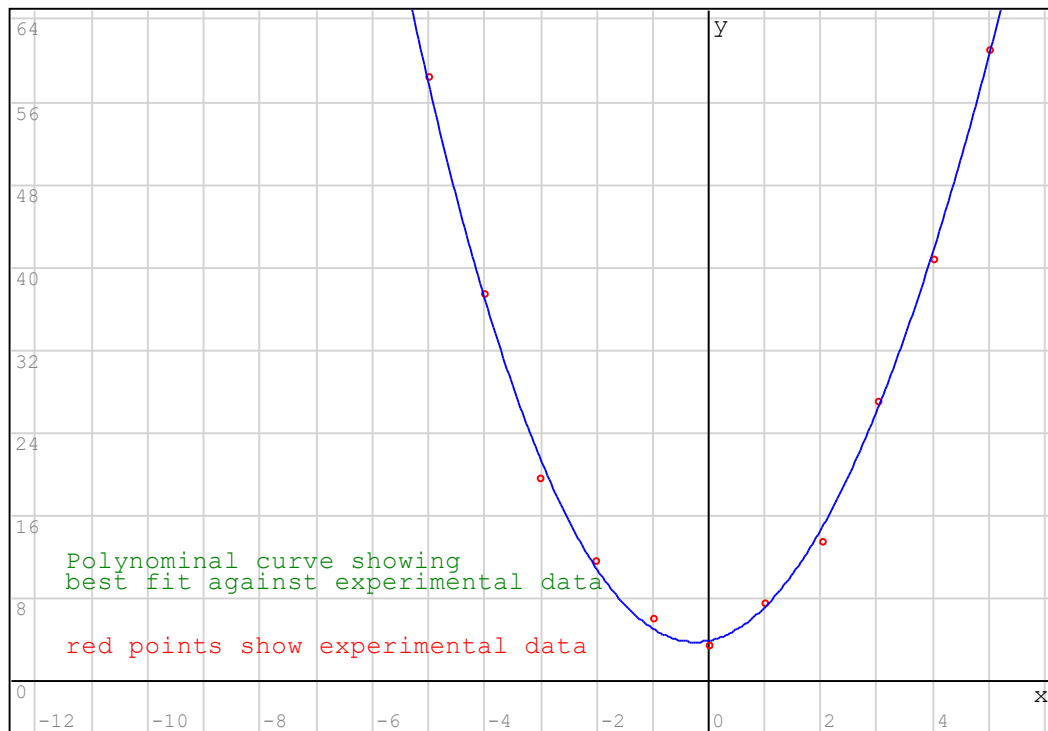
$$x = \begin{bmatrix} -5 \\ -4 \\ -3 \\ -2 \\ -1 \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}$$

$$y_d = \begin{bmatrix} 58.5 \\ 37.6 \\ 19.7 \\ 11.8 \\ 6.3 \\ 3.6 \\ 7.7 \\ 13.6 \\ 27.2 \\ 40.9 \\ 61.1 \end{bmatrix}$$

$$y := a_0 + a_1 \cdot x + a_2 \cdot x^2 + res = \begin{bmatrix} 58.5 \\ 37.6 \\ 19.7 \\ 11.8 \\ 6.3 \\ 3.6 \\ 7.7 \\ 13.6 \\ 27.2 \\ 40.9 \\ 61.1 \end{bmatrix}$$



$$f(x) := a_0 + a_1 \cdot x + a_2 \cdot x^2 + a_3 \cdot x^3$$



Sum of squares of the residual

$$SS_{res} := \sum_{i=1}^n \left(y_{d_i} - a_0 - a_1 \cdot x_i - a_2 \cdot x_i^2 - a_3 \cdot x_i^3 \right)^2$$

Total sum of the squares

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

R squared variance

$$R_{sq} := \frac{SS_{tot} - SS_{res}}{SS_{tot}} = 0.9979$$

Result for polynomial curve

$$y := a_0 + a_1 \cdot x + a_2 \cdot x^2 + a_3 \cdot x^3$$

$$a_0 = 3.9942 \quad a_1 = 1.0684 \quad a_2 = 2.2188 \quad a_3 = -0.0326$$