Linear Regression Example

This example assumes that you know how to do Matrix Multiplication for any size matrices. Matrix inversion for a 2x2 Matrix, Transpose of a Matrix, and Scalar Multiplication of a Matrix will all be covered below.

Transpose of a Matrix:

To calculate the Transpose of some Matrix A, we take each of its elements $a_{i,\,j}$ and create a new Matrix A^T where element $a_{i,\,j}$ from A is in position (j, i) in A^T.

Scalar Multiplication of a Matrix:

Some Matrix X multiplied by some scalar α is given by:

$$\alpha\boldsymbol{X} = \begin{bmatrix} \alpha x_{0,0} & \cdots & \alpha x_{0,j} \\ \vdots & \ddots & \vdots \\ \alpha x_{i,0} & \cdots & \alpha x_{i,j} \end{bmatrix}$$

Inverse of a 2x2 Matrix:

The inverse of some 2x2 Matrix X where:

$$\boldsymbol{X} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

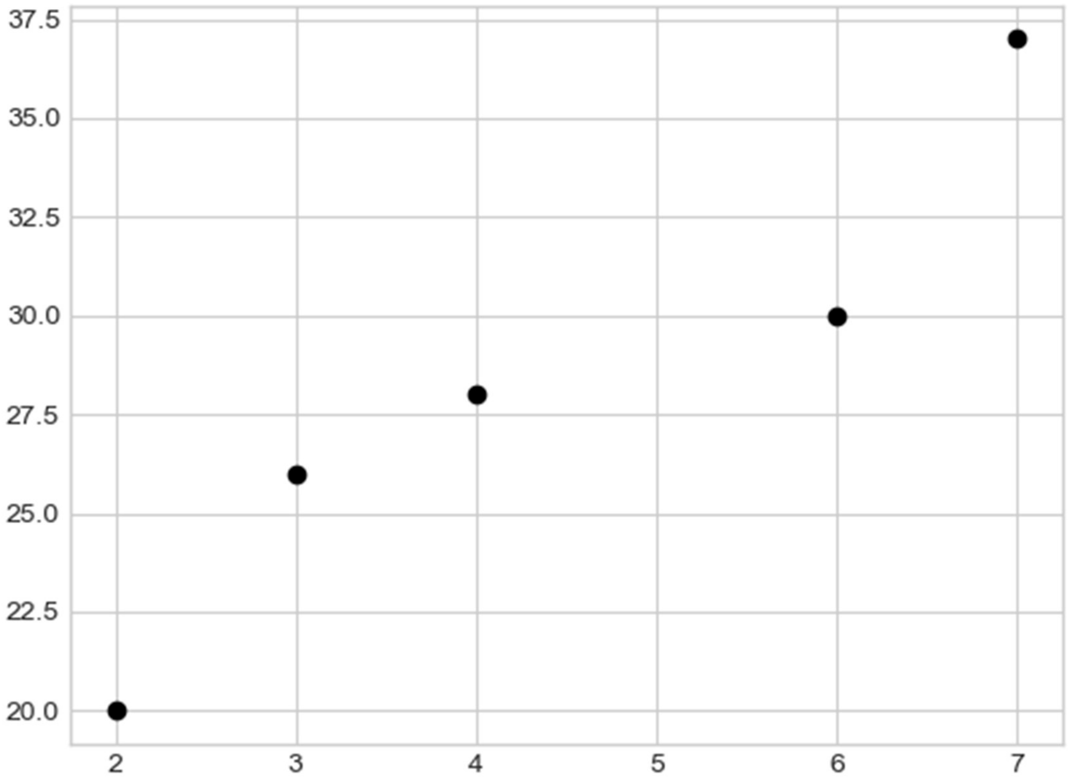
is given by[1]:

$$\boldsymbol{X}^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -c \\ -b & a \end{bmatrix}$$

[1] $ab - bc$ is the determinant of $\boldsymbol{X}$, and $\begin{bmatrix} d & -c \\ -b & a \end{bmatrix}$ is the Cofactor Matrix of $\boldsymbol{X}$. The inverse of $\boldsymbol{X}$ can only be calculated if the determinant is not 0, otherwise the equation for the inverse would involve dividing by 0.

Linear Regression Example:

The data covered by this example is shown in the graph below:



In this example the datapoints are (2, 20), (3, 26), (4, 28), (6, 30), and (7, 37).

In this case we are dealing with 2D data which means in order to create a Linear Regression Model we need two coefficients: a y-intersect, and a gradient.[2]

[2] For an N-Dimensional set of data, you will need N coefficients to create a Linear Regression Model, 1 for the y-intersect and N-1 for the gradients. This model will be of the form:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_{N-1} x_{N-1} + \beta_N x_N$$

Our model will be of the form:

$$y = \beta_0 + \beta_1 x$$

where β_0 and β_1 are obtained from the vector $\hat{\beta}$ where:

$$\hat{\beta} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y}$$

where:

$$\boldsymbol{X} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 6 \\ 1 & 7 \end{bmatrix}$$

$$\boldsymbol{Y} = \begin{bmatrix} 20 \\ 26 \\ 28 \\ 30 \\ 37 \end{bmatrix}$$

which have been obtained from our data points.

First, we need to calculate the Transpose of $\boldsymbol{X}$ or $\boldsymbol{X}^T$.

$$\boldsymbol{X}^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 3 & 4 & 6 & 7 \end{bmatrix}$$

Next, we need to calculate $\boldsymbol{X}^T\boldsymbol{X}$:

$$\boldsymbol{X}^T\boldsymbol{X} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 6 \\ 1 & 7 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 3 & 4 & 6 & 7 \end{bmatrix} = \begin{bmatrix} 5 & 22 \\ 22 & 114 \end{bmatrix}$$

The next step is to take the inverse of $\boldsymbol{X}^T\boldsymbol{X}$:

$$(\boldsymbol{X}^T\boldsymbol{X})^{-1} = \frac{1}{(5 \times 114) - (22 \times 22)} \begin{bmatrix} 114 & -22 \\ -22 & 5 \end{bmatrix} = \frac{1}{86} \begin{bmatrix} 114 & -22 \\ -22 & 5 \end{bmatrix} = \begin{bmatrix} \frac{57}{43} & \frac{-11}{43} \\ \frac{-11}{43} & \frac{5}{86} \end{bmatrix}$$

Next is to calculate $(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T$:

$$(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T = \begin{bmatrix} \frac{57}{43} & \frac{-11}{43} \\ \frac{-11}{43} & \frac{5}{86} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 3 & 4 & 6 & 7 \end{bmatrix} = \begin{bmatrix} \frac{35}{43} & \frac{24}{43} & \frac{13}{43} & \frac{-9}{43} & \frac{-20}{43} \\ \frac{-6}{43} & \frac{-7}{86} & \frac{-1}{43} & \frac{4}{43} & \frac{13}{86} \end{bmatrix}$$

The last step is to calculate $(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y}$ or $\hat{\beta}$:

$$\hat{\beta} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y} = \begin{bmatrix} \frac{35}{43} & \frac{24}{43} & \frac{13}{43} & \frac{-9}{43} & \frac{-20}{43} \\ \frac{-6}{43} & \frac{-7}{86} & \frac{-1}{43} & \frac{4}{43} & \frac{13}{86} \end{bmatrix} \begin{bmatrix} 20 \\ 26 \\ 28 \\ 30 \\ 37 \end{bmatrix} = \begin{bmatrix} \frac{678}{43} \\ \frac{243}{86} \end{bmatrix}$$

This gives us a Linear Regression Model of:

$$y = \frac{678}{43} + \frac{243}{86}x$$

Plotting this on the graph gives us the following: