Example 1:

Let

$$A = \begin{pmatrix} 4 & 0 & 1 \\ 1 & -5 & 1 \\ 6 & 1 & 0 \\ 1 & -1 & -5 \end{pmatrix}, \ \mathbf{b} = \begin{pmatrix} 9 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

The system $A\mathbf{x} = \mathbf{b}$ is inconsistent:

$$\begin{pmatrix} 4 & 0 & 1 & 9 \\ 1 & -5 & 1 & 0 \\ 6 & 1 & 0 & 0 \\ 1 & -1 & -5 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -1 & -5 & 0 \\ 0 & 1 & -3/2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 9 \end{pmatrix}.$$

So we find a least-squares solution. We compute:

$$\begin{aligned} A^T A &= \begin{pmatrix} 4 & 1 & 6 & 1 \\ 0 & -5 & 1 & -1 \\ 1 & 1 & 0 & -5 \end{pmatrix} \begin{pmatrix} 4 & 0 & 1 \\ 1 & -5 & 1 \\ 6 & 1 & 0 \\ 1 & -1 & -5 \end{pmatrix} = \begin{pmatrix} 54 & 0 & 0 \\ 0 & 27 & 0 \\ 0 & 0 & 27 \end{pmatrix} \\ A^T\mathbf{b} &= \begin{pmatrix} 4 & 1 & 6 & 1 \\ 0 & -5 & 1 & -1 \\ 1 & 1 & 0 & -5 \end{pmatrix} \begin{pmatrix} 9 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 36 \\ 0 \\ 9 \end{pmatrix}. \end{aligned}$$

The augmented matrix for the normal equations $A^T A\mathbf{x} = A^T\mathbf{b}$ is

$$\begin{pmatrix} 54 & 0 & 0 & 36 \\ 0 & 27 & 0 & 0 \\ 0 & 0 & 27 & 9 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 & 2/3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1/3 \end{pmatrix}.$$

so that the only least-squares solution is

$$\widehat{\mathbf{x}} = \begin{pmatrix} 2/ \\ 0 \\ 1/3 \end{pmatrix}.$$

Example 2:

Let
$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 7 \\ 2 \\ 3 \\ 6 \\ 5 \\ 4 \end{pmatrix}.$$

The system $A\mathbf{x} = \mathbf{b}$ is inconsistent:

$$\begin{pmatrix} 1 & 1 & 0 & 7 \\ 1 & 1 & 0 & 2 \\ 1 & 1 & 0 & 3 \\ 1 & 0 & 1 & 6 \\ 1 & 0 & 1 & 5 \\ 1 & 0 & 1 & 4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 0 & 7 \\ 0 & 0 & 0 & -5 \\ 1 & 1 & 0 & 3 \\ 1 & 0 & 1 & 6 \\ 1 & 0 & 1 & 5 \\ 1 & 0 & 1 & 4 \end{pmatrix}$$

So we find a least-squares solution. We compute:

$$A^T A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 3 & 3 \\ 3 & 3 & 0 \\ 3 & 0 & 3 \end{pmatrix}$$

$$A^T \mathbf{b} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 7 \\ 2 \\ 3 \\ 6 \\ 5 \\ 4 \end{pmatrix} = \begin{pmatrix} 27 \\ 12 \\ 15 \end{pmatrix}.$$

The augmented matrix for the normal equations $A^T A\mathbf{x} = A^T\mathbf{b}$ is

$$\begin{pmatrix} 6 & 3 & 3 & 27 \\ 3 & 3 & 0 & 12 \\ 3 & 0 & 3 & 15 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 1 & 5 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

so that the general least-squares solution is

$$\widehat{\mathbf{x}} = \begin{pmatrix} 5 \\ -1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

where x_3 is free.

Example 3:

Find the least-squares line that best fits the data points

$$(-1, 1) \ , \ (2, 3) \ , \ (3, 2) \ , \ (5, 3) \ , \ (6, 4).$$

The least-squares line has the form $y = \beta_0 + \beta_1 x$. The x-coordinates of our data points give

$$X = \begin{pmatrix} 1 & -1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 5 \\ 1 & 6 \end{pmatrix}$$

and the y-coordinates give

$$\mathbf{y} = \begin{pmatrix} 1 \\ 3 \\ 2 \\ 3 \\ 4 \end{pmatrix}.$$

We need the least squares solution to $X\beta = \mathbf{y}$. We compute:

$$X^T X = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & 2 & 3 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 5 \\ 1 & 6 \end{pmatrix} = \begin{pmatrix} 5 & 15 \\ 15 & 75 \end{pmatrix}$$

$$X^T \mathbf{y} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & 2 & 3 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 13 \\ 50 \end{pmatrix}.$$

Since $X^T X$ is 2×2 we can solve the normal equations $X^T X \beta = X^T \mathbf{y}$ by inversion:

$$
\begin{aligned}
\widehat{\beta} &= (X^T X)^{-1} X^T \mathbf{y} \\
&= \begin{pmatrix} 5 & 15 \\ 15 & 75 \end{pmatrix}^{-1} \begin{pmatrix} 13 \\ 50 \end{pmatrix} \\
&= \frac{1}{150} \begin{pmatrix} 75 & -15 \\ -15 & 5 \end{pmatrix} \begin{pmatrix} 13 \\ 50 \end{pmatrix} \\
&= \begin{pmatrix} 3/2 \\ 11/30 \end{pmatrix}.
\end{aligned}
$$

The least-squares line is therefore

$$y = \frac{3}{2} + \frac{11}{30} x.$$

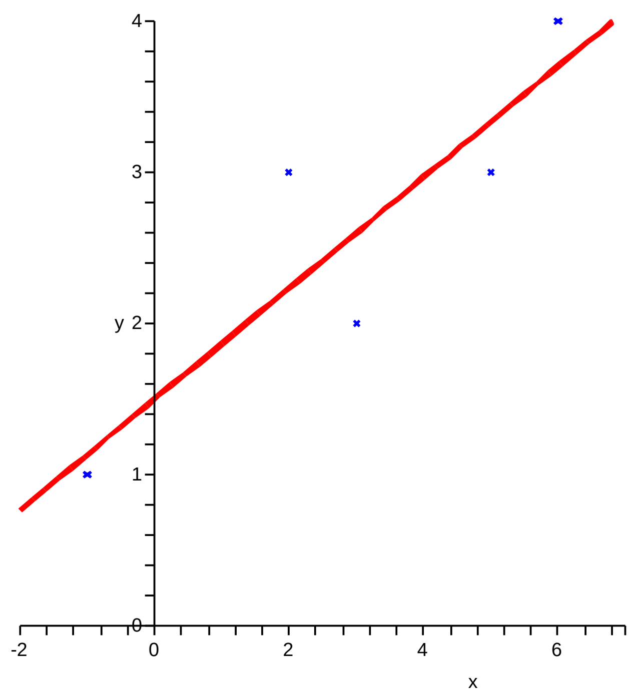

Figure 1: Data points and linear approximation

Example 4:

An experiment has produced the data points

$(4, 8.1)$, $(6, 10.4)$, $(8, 11.4)$, $(10, 10.8)$, $(12, 9.5)$, $(14, 10.5)$, $(16, 12.2)$, $(18, 15.9)$.

We want to use the least-squares method to find the curve of the form $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$ that best approximates this data.

If the data points (x_i, y_i) all sat on this curve then we would have

$$
\begin{aligned}
y_1 &= \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \beta_3 x_1^3 \\
y_2 &= \beta_0 + \beta_1 x_2 + \beta_2 x_2^2 + \beta_3 x_2^3 \\
y_3 &= \beta_0 + \beta_1 x_3 + \beta_2 x_3^2 + \beta_3 x_3^3 \\
&\vdots \\
y_8 &= \beta_0 + \beta_1 x_8 + \beta_2 x_8^2 + \beta_3 x_8^3
\end{aligned}
$$

which can be written as the matrix equation

$$
\begin{array}{ccc}
\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_8 \end{pmatrix} & = & \begin{pmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ 1 & x_3 & x_3^2 & x_3^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_8 & x_8^2 & x_8^3 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} \\
\mathbf{y} & = & X \qquad\qquad \beta
\end{array}
$$

The system $X\beta = \mathbf{y}$ is inconsistent in this case. The points *do not* all lie on a cubic curve. One way to approximate them by such a curve is to minimize the sum of the squares of the differences

$$|(\beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3) - y_i|$$

for $i = 1, 2, \ldots, 8$. This amounts to minimizing $||X\beta - \mathbf{y}||$, which is a least-squares problem!

So we need to find the least-squares solution to $X\beta = \mathbf{y}$ where

$$X = \begin{pmatrix} 1 & 4 & 16 & 64 \\ 1 & 6 & 36 & 216 \\ 1 & 8 & 64 & 512 \\ 1 & 10 & 100 & 1000 \\ 1 & 12 & 144 & 1728 \\ 1 & 14 & 196 & 2744 \\ 1 & 16 & 256 & 4096 \\ 1 & 18 & 324 & 5832 \end{pmatrix}, \ \mathbf{y} = \begin{pmatrix} 8.1 \\ 10.4 \\ 11.4 \\ 10.8 \\ 9.5 \\ 10.5 \\ 12.2 \\ 15.9 \end{pmatrix}.$$

We compute:

$$X^T X = \begin{pmatrix} 8 & 88 & 1136 & 16192 \\ 88 & 1136 & 16192 & 245312 \\ 1136 & 16192 & 245312 & 3866368 \\ 16192 & 245312 & 3866368 & 62617856 \end{pmatrix}$$

$$X^T \mathbf{y} = \begin{pmatrix} 88.8 \\ 1036.4 \\ 14014.4 \\ 207329.6 \end{pmatrix}.$$

The augmented matrix for the normal equations $X^T X\beta = X^T\mathbf{y}$ is therefore

$$\begin{pmatrix} 8 & 88 & 1136 & 16192 & 88.8 \\ 88 & 1136 & 16192 & 245312 & 1036.4 \\ 1136 & 16192 & 245312 & 3866368 & 14014.4 \\ 16192 & 245312 & 3866368 & 62617856 & 207329.6 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 & -5.595 \\ 0 & 1 & 0 & 0 & 5.331 \\ 0 & 0 & 1 & 0 & -0.546 \\ 0 & 0 & 0 & 1 & 0.018 \end{pmatrix}$$

so that the least-squares solution is

$$\widehat{\beta} = \begin{pmatrix} -5.595 \\ 5.331 \\ -0.546 \\ 0.018 \end{pmatrix}.$$

In other words, the cubic curve that best fits the given data is

$$y = -5.595 + 5.331x - 0.546x^2 + 0.018x^3.$$

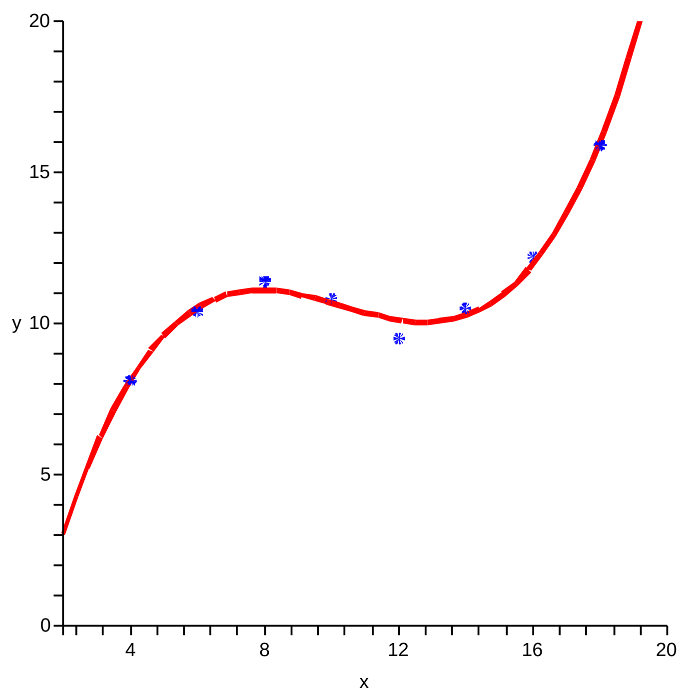

Figure 2: Data points and cubic approximation