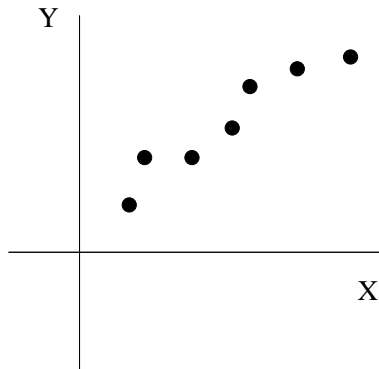
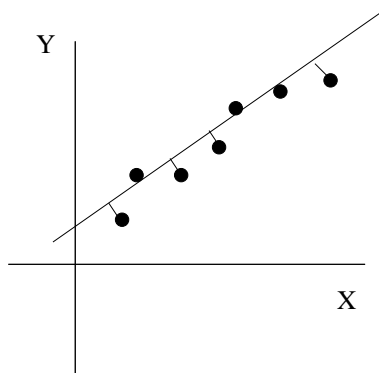


Least Squares Method

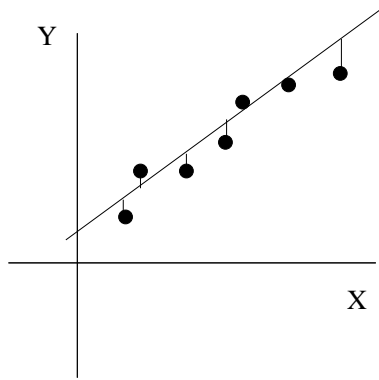
Suppose we have a set of data, that is a set of points (x_i, y_i) , which come from an experiment. See below



If we wish to compute the “error” in fitting a straight line, we could compute the shortest distance to a line (see below)



It is, however, more convenient to let the error be measured by the vertical deviation from the straight line, as shown below.



The vertical distance from a point (x_i, y_i) to the line $y = mx + b$ is given by

$$\sqrt{(x_i - x_i)^2 + (m * x_i + b - y_i)^2} = \sqrt{(m * x_i + b - y_i)^2}$$

And the total distance is

$$\sum_{i=1}^N \sqrt{(m * x_i + b - y_i)^2}$$

In order to obtain the “best” line, we want to make this error as small as possible. Minimizing the distance is the same as minimizing the distance squared, we define the error to be minimized as

$$E = E(m, b) = \sum_{i=1}^N (m * x_i + b - y_i)^2$$

Clearly, the error is a function of two variables, the slope (m) and the intercept (b). Since the error is a **quadratic** function of two variables, and is always positive, it must have a unique minimum! The minimum is achieved at the point \hat{m} and \hat{b} which satisfy

$$\begin{aligned} N\hat{b} + (\sum x_i)\hat{m} &= (\sum y_i) \\ (\sum x_i)\hat{b} + (\sum x_i^2)\hat{m} &= (\sum x_i y_i) \end{aligned}$$

The quantities $A = \sum x_i$, $B = \sum y_i$, $C = \sum x_i^2$ and $D = \sum x_i y_i$ are crucial to calculate. In terms of A, B, C, D, we have

$$\begin{aligned} N\hat{b} + A\hat{m} &= B \\ A\hat{b} + C\hat{m} &= D \end{aligned}$$

We can solve these equations by substitution or by cross multiplication and subtraction:

$$NA\hat{b} + A^2\hat{m} = AB$$

$$NA\hat{b} + NC\hat{m} = ND$$

therefore,

$$(A^2 - NC)\hat{m} = AB - ND$$

and

$$\hat{m} = \frac{AB - ND}{A^2 - NC} = \frac{(\sum x_i)(\sum y_i) - N(\sum x_i y_i)}{(\sum x_i)^2 - N(\sum x_i^2)}$$

Dividing numerator and denominator by N^2 we have

$$\hat{m} = \frac{\langle x \rangle \langle y \rangle - \langle xy \rangle}{\langle x \rangle^2 - \langle x^2 \rangle}$$

Solving for \hat{b} from the first equation, we get

$$\hat{b} = \frac{(B - A\hat{m})}{N} = \frac{\sum y_i - \hat{m} \sum x_i}{N} = \langle y \rangle - \hat{m} \langle x \rangle$$

These general formulas are valid for an arbitrary amount of data, and can easily be programmed in maple or any other computer language.