# Lecture Notes: Line Fitting

Given a set of points, we desire to find the line that *best fits* the data. In other words, the line should be as close as possible to the collective set of points. Figure 1 shows an example.

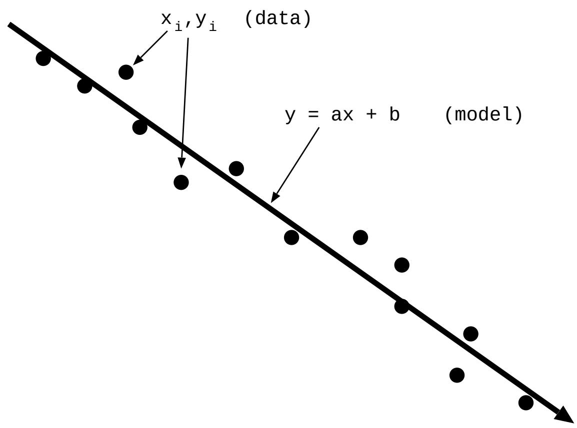


Figure 1: Fitting a line to a set of points.

Let the data be denoted as

$$(x_i, y_i) \quad i = 1...N \tag{1}$$

where $N$ indicates the total number of data points.

The model to be fit to the data is a line, denoted as

$$y = ax + b \tag{2}$$

where $a$ is the slope of the line and $b$ is the y-intercept. The unknowns in the model are $a$ and $b$.

We define the quality of the fit according to the residual $e_i$, which is the distance from each point to the line:

$$e_i = y_i - ax_i - b \tag{3}$$

This can be useful as an interpretation of how well the line fits the data. Lower values of $e_i$ indicate the given point is closer to the line, and a value of zero indicates the point is precisely on the line.

We define the chi-squared error metric as the difference between the best fitting line and the collective set of data:

$$\chi^2(a, b) = \sum_{i=1}^{N} (y_i - ax_i - b)^2 \tag{4}$$

In order to find the best possible values for $a$ and $b$, we use differential equations to solve for the minimum chi-squared error. We take the partial derivatives of $\chi^2(a,b)$ with respect to $a$ and $b$, set them equal to zero, and solve for $a$ and $b$:

$$\frac{\partial \chi^2}{\partial a} = \sum_{i=1}^{N} -2x_i(y_i - ax_i - b) = 0 \tag{5}$$

$$\frac{\partial \chi^2}{\partial b} = \sum_{i=1}^{N} -2(y_i - ax_i - b) = 0 \tag{6}$$

Distributing the summations yields the following two equations:

$$\sum_{i=1}^{N} x_i y_i - a\sum_{i=1}^{N} {x_i}^2 - b\sum_{i=1}^{N} x_i = 0 \tag{7}$$

$$\sum_{i=1}^{N} y_i - a\sum_{i=1}^{N} x_i - b\sum_{i=1}^{N} 1 = 0 \tag{8}$$

All the values inside the sums are known from the data. The unknowns are $a$ and $b$. We therefore have two linear equations with 2 unknowns, which can be solved using simple algebra.