Least Squares Approximation

Least Squares Linear Fitting: If we are given a set of data points,

$$(x_i, y_i) \quad i = 1, 2, \ldots, n$$

can we use a line to fit these data points? The answer is positive.

If the line is expressed as

$$y = a_0 + a_1 x$$

where a_0 and a_1 are the two best values to be determined. Obviously, the error e_i of each point (x_i, y_i) with respect to $y = a_0 + a_1 x$ will be

$$e_i = y_i - y\,|_{x=x_i} = y_i - (a_0 + a_1 x_i)$$

The least squares criterion requires that

$$S = e_1^2 + e_2^2 + \ldots + e_n^2 = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - a_1 x_i - a_0)^2$$

be a minimum.

At a minimum for S, the two derivatives $\partial S/\partial a_0$ and $\partial S/\partial a_1$ will both be zero:

$$\left.\begin{aligned} \frac{\partial S}{\partial a_1} &= \sum_{i=1}^{n} 2(y_i - a_1 x_i - a_0)(-x_i) = 0, \\ \frac{\partial S}{\partial a_0} &= \sum_{i=1}^{n} 2(y_i - a_1 x_i - a_0)(-1) = 0, \end{aligned}\right\} \quad \begin{aligned} & a_1 \sum_{i=1}^{n} x_i^2 + a_0 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i, \\ & a_1 \sum_{i=1}^{n} x_i + a_0 n = \sum_{i=1}^{n} y_i \end{aligned}$$

Thus, a_0 and a_1 can be obtained so that the data points are linearly fitted.

In fact, we can write the above equations into a linear system:

$$\begin{bmatrix} \sum_{i=1}^{n} (x_i)^0 & \sum_{i=1}^{n} (x_i)^1 \\ \sum_{i=1}^{n} (x_i)^1 & \sum_{i=1}^{n} (x_i)^2 \end{bmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \end{pmatrix}$$

for solve it for a_0 and a_1.

Least Squares Polynomials:

Instead of matching the data in every node, the least square method is trying to fit n pairs of data by a polynomial of a pre-determined degree, say m,

$$y = a_0 + a_1 x + a_2 x^2 + \dots + a_m x = \sum_{j=0}^{m} a_j x^j$$

We define the fitting errors

$$e_i = y_i - \sum_{j=0}^{m} a_j x_i^j \qquad S = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left(y_i - \sum_{j=0}^{m} a_j x_i^j \right)^2$$

In order to achieve minimal error S (least square error), all the partial derivatives

$$\partial S / \partial a_0 ,\ \partial S / \partial a_1 , \dots , \partial S / \partial a_m$$

must equal to 0. Writing the equations for these given $m + 1$ equations:

$$\frac{\partial S}{\partial a_0} = \sum_{i=1}^{n} 2\left(y_i - \sum_{j=0}^{m} a_j x_i^j\right)(-1) = 0$$

$$\frac{\partial S}{\partial a_1} = \sum_{i=1}^{n} 2\left(y_i - \sum_{j=0}^{m} a_j x_i^j\right)(-x_i) = 0$$

$$\vdots$$

$$\frac{\partial S}{\partial a_m} = \sum_{i=1}^{n} 2\left(y_i - \sum_{j=0}^{m} a_j x_i^j\right)(-x_i^m) = 0$$

$$\left.\begin{array}{r} a_0 n + a_1 \sum x_i + \cdots + a_n \sum x_i^m = \sum y_i \\ a_0 \sum x_i + a_1 \sum x_i^2 + \cdots + a_n \sum x_i^{m+1} = \sum x_i y_i \\ a_0 \sum x_i^2 + a_1 \sum x_i^3 + \cdots + a_n \sum x_i^{m+2} = \sum x_i^2 y_i \\ \vdots \\ a_0 \sum x_i^n + a_1 \sum x_i^{n+1} + \cdots + a_n \sum x_i^{2m} = \sum x_i^m y_i \end{array}\right.$$

Or solving the following system

$$\begin{pmatrix} n & \Sigma x_i & \Sigma x_i^2 & \cdots & \Sigma x_i^m \\ \Sigma x_i & \Sigma x_i^2 & \Sigma x_i^3 & \cdots & \Sigma x_i^{m+1} \\ \Sigma x_i^2 & \Sigma x_i^3 & \Sigma x_i^4 & \cdots & \Sigma x_i^{m+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Sigma x_i^m & \Sigma x_i^{m+1} & \Sigma x_i^{m+2} & \cdots & \Sigma x_i^{2m} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} = \begin{bmatrix} \Sigma y_i \\ \Sigma x_i y_i \\ \Sigma x_i^2 y_i \\ \vdots \\ \Sigma x_i^m y_i \end{bmatrix}$$

for $a_0, a_1, \ldots, a_m$.

Example:

To demonstrate how the method is used, we would fit a quadratic to the following data:

x_i:	0.05	0.11	0.15	0.31	0.46	0.52	0.70	0.74	0.82	0.98	1.17
y_i:	0.956	0.890	0.832	0.717	0.571	0.539	0.378	0.370	0.306	0.242	0.104

These data are actually a perturbation of the relation

$$y = 1 - x + 0.2x^2$$

Obviously we have

$$\Sigma x_i = 6.01 \quad \Sigma x_i^2 = 4.6545 \quad \Sigma x_i^3 = 4.1150 \quad \Sigma x_i^4 = 3.9161$$

$$n = 11 \qquad \Sigma y_i = 5.9050 \quad \Sigma x_i y_i = 2.1839 \quad \Sigma x_i^2 y_i = 1.3357$$

Thus the equation system to be solved is:

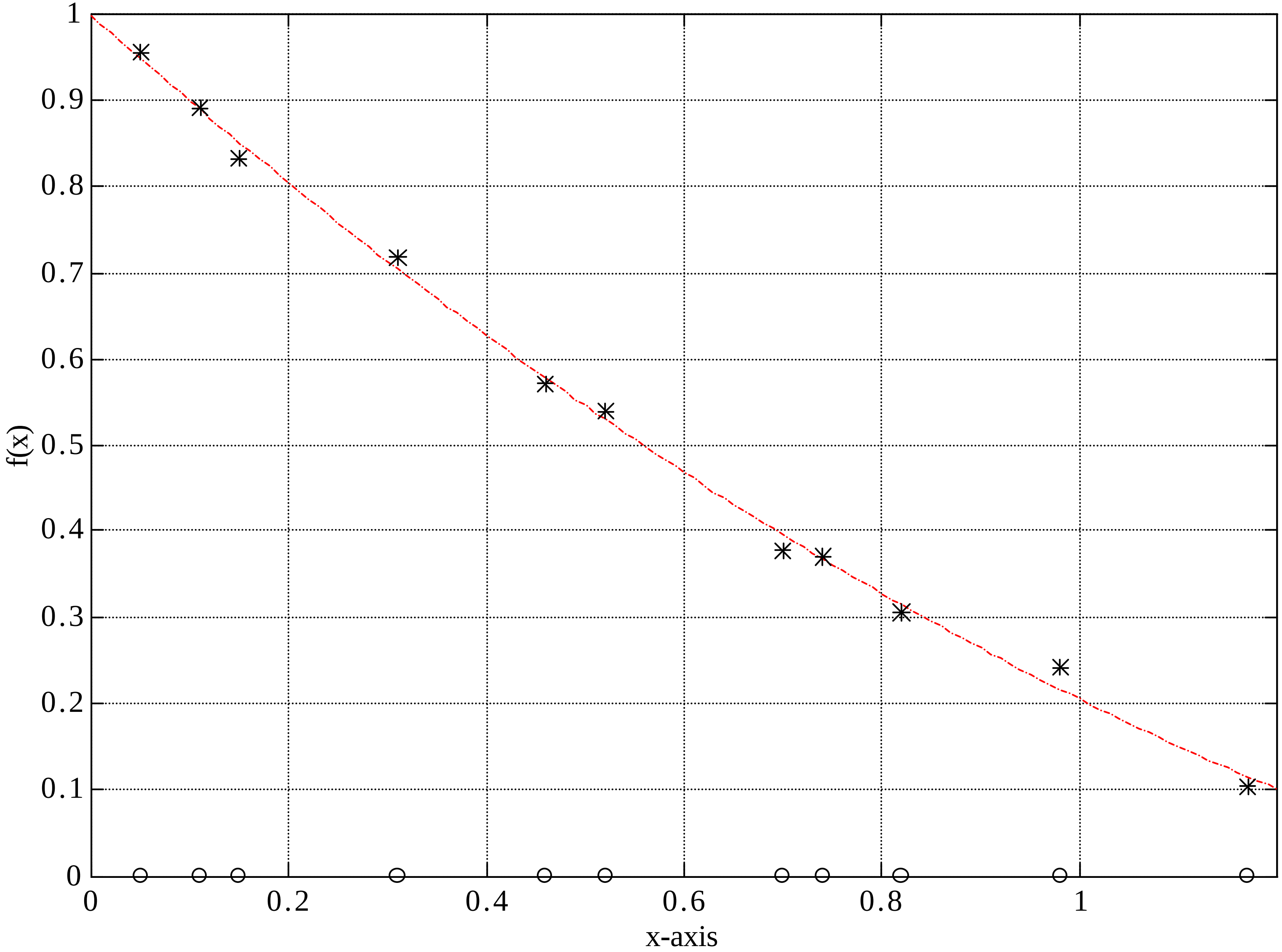
$$
\left.\begin{aligned}
11.0000a_0 + 6.0100a_1 + 4.6545a_2 &= 5.9050 \\
6.0100a_0 + 4.6545a_1 + 4.1150a_2 &= 2.1839 \\
4.6545a_0 + 4.1150a_1 + 3.9161a_2 &= 1.3357
\end{aligned}\right\} \quad
\begin{aligned}
a_0 &= 0.998 \\
a_1 &= -1.018 \\
a_2 &= 0.225
\end{aligned}
$$

The above linear system can be solved using methods given in Part I of of this course or using MATLAB software package.

Thus, the least square quadratic fit is given by

$$y = 0.998 - 1.018x + 0.225x^2$$

Compare this to $y = 1 - x + 0.2x^2$. We do not expect to reproduce the coefficients exactly because of the error in the data. Figure of next page shows a plot of the data and its fitting-curve.



n