The Least Square Method

Problem: Suppose we are given n data points, (x_1, y_1), $(x_2, y_2), \ldots, (x_n, y_n)$ (where the x_i's are distinct), and an integer $m \le n$. Find the polynomial $f(x)$ of degree $\le m-1$ which "best approximates" these points in the sense that the mean square deviation

$$\Delta(f) = \frac{1}{n}\left[(y_1 - f(x_1))^2 + \ldots + (y_n - f(x_n))^2\right]$$

is minimal (among all such polynomials).

Method: Put

$$A = \begin{pmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{m-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{m-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^{m-1} \end{pmatrix}, \vec{w} = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{m-1} \end{pmatrix}, \vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Theorem 4: The best approximate solution is $f(x) = a_0 + a_1 x + \ldots + a_{m-1}x^{m-1}$, where

$$\vec{w} = (A^t A)^{-1} A^t \vec{y}.$$

Note: If the data points all lie on the curve $y = f(x)$, then we would have $A\vec{w} = \vec{y}$, but in general this system is inconsistent.