Mathematics 3: Least-Squares fitting

Least-squares fitting is a widely used method of trying to obtain smooth approximations to data points. Any approximation of the form:

$$y(x) = a_1 f_1(x) + a_2 f_2(x) + \ldots + a_n f_n(x)$$

may be considered, where the functions $f_i(x)$ are given and the variables a_i are to be 'fit' by the procedure. Simple examples are:

$$\begin{aligned} y(x) &= a + bx \\ y(x) &= a + be^{-x} \\ y(x) &= a + bx + cx^2 \log x \end{aligned}$$

The idea is quite elementary, one minimises the sum of squares of the errors, hence 'least-squares fitting'. Given n points, (x_i, y_i), the sum of the squares of the errors is:

$$E = \frac{1}{n} \sum_{i=1}^{n} \left[y_i - \sum_{j=1}^{N} a_j f_j(x_i) \right]^2$$

since this is an elementary quadratic in the variables a_i, it is a simple matter to minimise over them:

$$\frac{\partial E}{\partial a_k} = \frac{1}{n} \sum_i (-2) f_k(x_i) \left[y_i - \sum_{j=1}^{N} a_j f_j(x_i) \right]$$

and setting all these equal to zero yields simultaneous linear equations for the a_i. The best way to formulate the problem is in matrix language:

$$E = \frac{1}{n} \sum_{i=1}^{n} y_i^2 - 2\mathbf{a}^T.\mathbf{B} + \mathbf{a}^T A \mathbf{a}$$

where the coefficients of the vector $\mathbf{a}$ are the a_i and for which the vector $\mathbf{B}$ and matrix A have components:

$$B_j = \frac{1}{n} \sum_{i=1}^{n} y_i f_j(x_i)$$

$$A_{jk} = \frac{1}{n} \sum_{i=1}^{n} f_j(x_i) f_k(x_i)$$

in terms of which the optimum solution satisfies:

$$A\mathbf{a}^* = \mathbf{B}$$

in terms of the optimal variables a_i^*, and the minimum error becomes:

$$E^* = \frac{1}{n} \sum_{i=1}^{n} y_i^2 - \mathbf{a}^{*T}.\mathbf{B}$$

The entire analysis reduces to a set of simultaneous linear equations.

For the common case of:

$$y(x) = a + bf(x)$$

we need the sums:

$$s_x = \frac{1}{n}\sum_{i=1}^{n} f(x_i)$$

$$s_y = \frac{1}{n}\sum_{i=1}^{n} y_i$$

$$s_{xx} = \frac{1}{n}\sum_{i=1}^{n} f(x_i)^2$$

$$s_{xy} = \frac{1}{n}\sum_{i=1}^{n} y_i f(x_i)$$

$$s_{yy} = \frac{1}{n}\sum_{i=1}^{n} y_i^2$$

in terms of which:

$$a = \frac{s_{xx}s_y - s_x s_{xy}}{s_{xx} - s_x^2}$$

$$b = \frac{s_{xy} - s_x s_y}{s_{xx} - s_x^2}$$

$$E = s_{yy} - s_y^2 - \frac{(s_{xy} - s_x s_y)^2}{s_{xx} - s_x^2}$$

which is the simplest non-trivial example of the analysis.