A regression fit of a chosen function form to a set of data is obtained by picking coefficients that minimize the total squared difference (or error) between the function and the data. (If the function form is a polynomial, for example, the parameters are the coefficients of the polynomial.) The data points, \((\bar{x}_i, \bar{y}_i)\), are located at points \(\bar{x}_i\) in an \(N\)-dimensional space. We denote the function fit as \((\bar{x}_i, f(\bar{a}, \bar{x}_i))\), where \(\bar{a}\) contains the coefficients of \(f\).

\[
E = \text{SSE} = \text{Sum of Squared Errors of all observations}
\]

\[
E = \sum_{i=1}^{N} (\bar{y}_i - f(\bar{a}, \bar{x}_i))^2
\]

From calculus, the least squares solution is to set the derivatives of the total squared error with respect to \(a_1, \ldots, a_M\) equal to zero.

\[
\frac{dE}{da_j} = 2 \left\{ \sum_{i=1}^{N} \left[ \bar{y}_i - f(\bar{a}, \bar{x}_i) \right] \frac{df(\bar{a}, \bar{x})}{da_j} \bigg|_{\bar{x}=\bar{x}_i} \right\} = 0
\]

**NOTE:** The sum is over \(i\), but the derivative is with respect to \(j\).

**NOT'N:**

\[
f_i = f(\bar{a}, \bar{x}_i)
\]

\[
f_{ji} = \frac{df(\bar{a}, \bar{x})}{da_j} \bigg|_{\bar{x}=\bar{x}_i}
\]

Using this notation, we find \(a_1, \ldots, a_M\) by solving the following equation:

\[
\sum_{i=1}^{N} \left\{ [y_i - f_i] f_{ji} \right\} = 0, \quad j = 1, \ldots, M
\]

or

\[
\sum_{i=1}^{N} y_i f_{ji} = \sum_{i=1}^{N} f_i f_{ji}, \quad j = 1, \ldots, M
\]

**NOTE:** We get \(M\) equations in \(M\) unknowns—one for each \(a_j\).