Lecture Notes: Normal Equations

The technique for fitting a line to set of points can be generalized to the fitting of any function consisting of a linear combination of terms. Such a function can be written as

$$y = a_1 f_1(x) + a_2 f_2(x) + ... + a_M f_M(x) \tag{1}$$

where $a_1 ... a_M$ are the unknowns (M of them). The terms $f_1(x), f_2(x), ..., f_M(x)$ are called basis functions. The basis functions do not need to be linear; they can be anything. However, the unknowns must all be linear constants.

The function for a line is a special example of equation 1. For example, let $a_1 = a$, $a_2 = b$, $f_1(x) = x$, and $f_2(x) = 1$. Then equation 1 simplifies to

$$y = a \cdot x + b \cdot 1 = ax + b \tag{2}$$

Given a set of points, we desire to find the general solution to equation 1 that best fits the data. Let the data be denoted as

$$(x_i, y_i) \quad i = 1...N \tag{3}$$

where N indicates the total number of data points.

We define the residual e_i for each point as:

$$e_i = \left(y_i - \sum_{j=1}^{M} a_j f_j(x_i) \right) \tag{4}$$

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

$$\chi^2(a_1, a_2, ..., a_M) = \sum_{i=1}^{N} \left(y_i - \sum_{j=1}^{M} a_j f_j(x_i) \right)^2 \tag{5}$$

In order to find the best possible values for the unknowns $a_1 ... a_M$ we use differential equations to solve for the minimum chi-squared error. We take the partial derivatives of χ^2 with respect to $a_1 ... a_M$, set them equal to zero, and solve for $a_1 ... a_M$. There are M partial derivative equations. Here are the first two:

$$\frac{\partial \chi^2}{\partial a_1} = \sum_{i=1}^{N} 2 \left(y_i - \sum_{j=1}^{M} a_j f_j(x_i) \right) (-f_1(x_i)) \tag{6}$$

$$\frac{\partial \chi^2}{\partial a_2} = \sum_{i=1}^{N} 2 \left(y_i - \sum_{j=1}^{M} a_j f_j(x_i) \right) (-f_2(x_i)) \tag{7}$$

In general form, the set of M equations can be written as:

$$\forall k = 1...M \quad \frac{\partial \chi^2}{\partial a_k} = \sum_{i=1}^{N} 2 \left(y_i - \sum_{j=1}^{M} a_j f_j(x_i) \right) (-f_k(x_i)) \tag{8}$$

To solve for the unknowns $a_1...a_M$ we set all these equations equal to zero:

$$\forall k = 1...M \quad \sum_{i=1}^{N} f_k(x_i) \left(y_i - \sum_{j=1}^{M} a_j f_j(x_i) \right) = 0 \tag{9}$$

Rearranging the terms and expanding the sums, we obtain

$$\forall k = 1...M \quad \sum_{i=1}^{N} f_k(x_i) y_i = \sum_{i=1}^{N} \sum_{j=1}^{M} f_k(x_i) f_j(x_i) a_j \tag{10}$$

In order to proceed we use matrix notation to simplify the equations. We define the following matrices:

$$A = \begin{bmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_M(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_M(x_2) \\ \vdots & & & \vdots \\ f_1(x_N) & f_2(x_N) & \cdots & f_M(x_N) \end{bmatrix} \tag{11}$$

$$x = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix} \tag{12}$$

$$b = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \tag{13}$$

Note that matrix A is $N \times M$ in size, x is $M \times 1$ and b is $N \times 1$. Using these matrices, equation 10 can be rewritten in matrix form as

$$A^T b = A^T A x \tag{14}$$

We desire to solve for the unknowns in matrix x. Reversing the equation puts x on the left side:

$$A^T A x = A^T b \tag{15}$$

Note that A^TA is by definition a square matrix and is therefore invertible. Multiplying both sides of equation 15 by this inverse gives

$$(A^TA)^{-1}A^TAx = (A^TA)^{-1}A^Tb \tag{16}$$

Any matrix multiplied by its inverse yields the identity matrix, so that the left side of this equation simplifies:

$$x = (A^TA)^{-1}A^Tb \tag{17}$$

Equation 17 is called the solution to the normal equations. Given properly constructed matrices A, x and b, the solution to any problem in the form of equation 1 can be found using equation 17.

For example, let us revisit the line fitting problem. Given (x_i, y_i) for $i = 1...N$ data, we can fit the model $y = ax + b$ by constructing the following three matrices:

$$A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_N & 1 \end{bmatrix} \tag{18}$$

$$x = \begin{bmatrix} a \\ b \end{bmatrix} \tag{19}$$

$$b = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \tag{20}$$

After constructing these matrices, the solution is found by solving equation 17 for the matrix x.

As a second example, consider the problem of fitting a circle to a set of points. Assume we are given (x_i, y_i) for $i = 1...N$ data. The desired model is of the form

$$(x - a)^2 + (y - b)^2 = r^2 \tag{21}$$

Unfortunately, this model is not linear in the unknowns a, b, r. It cannot directly be written in the form of equation 1. We can however use a trick to make it linear. Equation 21 can be rearranged as follows:

$$r^2 - (x - a)^2 - (y - b)^2 = 0 \tag{22}$$

Expanding this gives

$$r^2 - x^2 + 2xa - a^2 - y^2 + 2yb - b^2 = 0 \tag{23}$$

Now we use a trick to substitute a linear term for the set of non-linear terms. Let

$$\alpha = r^2 - a^2 - b^2 \tag{24}$$

Then equation 23 can be written as

$$\alpha - x^2 + 2xa - y^2 + 2yb = 0 \tag{25}$$

Rearranging this gives

$$2xa + 2yb + \alpha = x^2 + y^2 \tag{26}$$

which is linear in the unknowns a, b, α. We can therefore construct the following three matrices:

$$A = \begin{bmatrix} 2x_1 & 2y_1 & 1 \\ 2x_2 & 2y_2 & 1 \\ \vdots & \vdots & \\ 2x_N & 2y_N & 1 \end{bmatrix} \tag{27}$$

$$x = \begin{bmatrix} a \\ b \\ \alpha \end{bmatrix} \tag{28}$$

$$b = \begin{bmatrix} x_1^2 + y_1^2 \\ x_2^2 + y_2^2 \\ \vdots \\ x_N^2 + y_N^2 \end{bmatrix} \tag{29}$$

After constructing these matrices, the solution is found by solving equation 17 for the matrix x. Finally, the value of r is found by back-substitution using equation 24.