MthStat 465, Spring 2005, Lecture Number 26
Least Squares Fit in Linear Models

In the linear model we are given a **design matrix** B and we are hoping to estimate the vector $\vec{A}$ so that the obeserved values $\vec{y}$ are well-approximated by $B\vec{A}$. By well-approximated we mean that the distance from $\vec{y}$ to $B\vec{A}$ is as small as possible. If we measure the length of a vector as the square root of the sum of the squares of its components we are just saying that we want to apply the **least squares criterion** for a good approximation.

Consider, for a moment, the familiar problem of fitting a line of the form $y = A_1 + A_2 x$ to the points $(1, 2)$, $(2, 3)$ and $(3, 7)$ using the least squares criterion. We want to choose A_1 and A_2 so that

$$(2 - (A_1 + A_2 \cdot 1))^2 + (3 - (A_1 + A_2 \cdot 2))^2 + (7 - (A_1 + A_2 \cdot 3))^2$$

is as small as possible. If we let $\| \cdot \|$ denote the length of a vector we can represent this expression using vectors as

$$\left\| \begin{bmatrix} 2 \\ 3 \\ 7 \end{bmatrix} - \left(A_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + A_2 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right) \right\|^2 .$$

We can think of the geometrically. The vectors

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

span a plane, and we are trying to find the closest vector, $\vec{P}$, in this plane, written as

$$\vec{P} := A_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + A_2 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},$$

to the vector

$$\begin{bmatrix} 2 \\ 3 \\ 7 \end{bmatrix}.$$

It is clear that to do so we should project the latter vector straight down into this plane, and, as a result, the vector

$$\begin{bmatrix} 2 \\ 3 \\ 7 \end{bmatrix} - \left(A_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + A_2 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right)$$

should be perpendicular to each of the two vectors that define the plane. Since perpendicularity can be checked by checking if the dot products of the vectors are equal to 0, we have the equations

$$\begin{aligned} \left(\begin{bmatrix} 2 \\ 3 \\ 7 \end{bmatrix} - \left(A_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + A_2 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right) \right) \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} &= 0; \\ \left(\begin{bmatrix} 2 \\ 3 \\ 7 \end{bmatrix} - \left(A_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + A_2 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right) \right) \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} &= 0. \end{aligned}$$

This can be summarized as the single matrix equation

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \left(\begin{bmatrix} 2 \\ 3 \\ 7 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Observe that the matrix form of our model was

$$\begin{bmatrix} 2 \\ 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{bmatrix}.$$

The design matrix B is the matrix

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$$

and our equation for the coefficient vector $\vec{A}$ reads

$$B^t \left(\vec{y} - B\vec{A} \right) = \vec{0},$$

where B^t denotes the transpose of the matrix B. If B^tB can be inverted, then we can solve for $\vec{A}$:

$$\vec{A} = (B^tB)^{-1}B^t\vec{y}. \tag{1}$$

We need to check that the result in (1) does indeed give the solution t the following least squares problem. Suppose that B is a matrix with k columns and N rows, with $N \geq k$. Suppose that B^tB is invertible, and $\vec{y}$ is a column vector with N rows. Then for any column vector $\vec{a}$ with k rows,

$$\|\vec{y} - B\vec{a}\|^2 \geq \|\vec{y} - B\vec{A}\|^2 \tag{2}$$

where $\vec{A}$ is given by (1), and we equality if and only if $\vec{a} = \vec{A}$. This is straightforward to demonstrate:

$$\begin{aligned} \|\vec{y} - B\vec{a}\|^2 &= \|(\vec{y} - B\vec{A}) + (B\vec{A} - B\vec{a})\|^2 \\ &= \|(\vec{y} - B\vec{A})\|^2 + \|B(\vec{A} - \vec{a})\|^2 + 2B(\vec{A} - \vec{a}) \cdot (\vec{y} - B\vec{A}). \end{aligned}$$

Now

$$\begin{aligned} B(\vec{A} - \vec{a}) \cdot (\vec{y} - B\vec{A}) &= (\vec{A} - \vec{a})^t B^t (\vec{y} - B\vec{A}) \\ &= (\vec{A} - \vec{a})^t B^t (\vec{y} - B(B^tB)^{-1}B^t\vec{y}) \\ &= (\vec{A} - \vec{a})^t (B^t\vec{y} - (B^tB)(B^tB)^{-1}B^t\vec{y}) \\ &= (\vec{A} - \vec{a})^t (B^t\vec{y} - B^t\vec{y}) \\ &= (\vec{A} - \vec{a})^t (\vec{0}) \\ &= 0 \end{aligned}$$

so

$$\|\vec{y} - B\vec{a}\|^2 = \|(\vec{y} - B\vec{A})\|^2 + \|B(\vec{A} - \vec{a})\|^2 \geq \|(\vec{y} - B\vec{A})\|^2. \tag{3}$$

Furthermore, if $\|B(\vec{A} - \vec{a})\|^2 = 0$, then $B(\vec{A} - \vec{a}) = \vec{0}$. Multiplying both sides by B^t gives $(B^tB)(\vec{A} - \vec{a}) = \vec{0}$. Now multiply both sides by the inverse of B^tB and we get $\vec{A} - \vec{a} = \vec{0}$, so $\vec{a} = \vec{A}$. So, given the design matrix B and the observed values $\vec{y}$ we can find the least squares coefficients using (1).