Using matrix algebra in linear regression

Jackie Nicholas
Mathematics Learning Centre
University of Sydney

The regression model

Consider the linear regression model:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i \qquad\qquad i = 1, \cdots, n$$

We can write model in matrix form as,

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} \qquad \text{or } Y = X\beta + \epsilon,$$

where $Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$, $X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$, $\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$ and $\epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$.

The ordinary least squares estimate (OLS) of β

The sample regression equation is written as:

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\epsilon}_i \qquad\qquad i = 1, \cdots, n$$

which can be written in matrix form as:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} + \begin{bmatrix} \hat{\epsilon}_1 \\ \hat{\epsilon}_2 \\ \vdots \\ \hat{\epsilon}_n \end{bmatrix}$$

or in matrix notation as:

$$Y = X\hat{\beta} + \hat{\epsilon}.$$

The OLS estimate of β

The OLS estimate of β is obtained by minimising

$$\Sigma\hat{\epsilon}_i^2 = \Sigma(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

to get the normal equations (don't do this)

$$\begin{aligned} n\hat{\beta}_0 + \Sigma x_i\hat{\beta}_1 &= \Sigma y_i \\ \Sigma x_i\hat{\beta}_0 + \Sigma x_i^2\hat{\beta}_1 &= \Sigma x_i y_i. \end{aligned}$$

The normal equations can be written in matrix form as:

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Check by multiplying the matrices out.

Solving the matrix equation

This is written in matrix notation as

$$X'X\hat{\beta} = X'Y.$$

As the matrix X' is $2 \times n$ and X is $n \times 2$, $X'X$ is a 2×2 matrix.

If $(X'X)^{-1}$ exists, we can solve the matrix equation as follows:

$$\begin{aligned}
X'X\hat{\beta} &= X'Y \\
(X'X)^{-1}(X'X)\hat{\beta} &= (X'X)^{-1}X'Y \\
I\hat{\beta} &= (X'X)^{-1}X'Y \\
\hat{\beta} &= (X'X)^{-1}X'Y.
\end{aligned}$$

This is a fundamental result of the OLS theory using matrix notation. The result holds for a multiple linear regression model with $k-1$ explanatory variables in which case $X'X$ is a $k \times k$ matrix.

Example from Associate Professor Tim Fisher

Let $X = \begin{bmatrix} 1 & 2 \\ 1 & 4 \\ 1 & 6 \end{bmatrix}$, $Y = \begin{bmatrix} 3 \\ 6 \\ 7 \end{bmatrix}$ with $\hat{\beta} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix}$.

Show that $\hat{\beta} = \begin{bmatrix} 1.3333 \\ 1.0000 \end{bmatrix}$.

$$\begin{aligned} X'X &= \begin{bmatrix} 1 & 1 & 1 \\ 2 & 4 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 4 \\ 1 & 6 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 12 \\ 12 & 56 \end{bmatrix} \end{aligned}$$

$|X'X| = 3(56) - 12(12) = 24$, so $(X'X)^{-1}$ exists.

Example

So, $(X'X)^{-1} = \frac{1}{24}\begin{bmatrix} 56 & -12 \\ -12 & 3 \end{bmatrix}$.

$$\begin{aligned}
\hat{\beta} &= (X'X)^{-1}X'Y \\
&= \frac{1}{24}\begin{bmatrix} 56 & -12 \\ -12 & 3 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ 2 & 4 & 6 \end{bmatrix}\begin{bmatrix} 3 \\ 6 \\ 7 \end{bmatrix} \\
&= \frac{1}{24}\begin{bmatrix} 56 & -12 \\ -12 & 3 \end{bmatrix}\begin{bmatrix} 16 \\ 72 \end{bmatrix} \\
&= \frac{1}{24}\begin{bmatrix} 32 \\ 24 \end{bmatrix} \\
&= \begin{bmatrix} 1.3333 \\ 1.0000 \end{bmatrix} \qquad \text{as required.}
\end{aligned}$$