Discussion of the Gauss-Markov Theorem

Introduction to Econometrics (C. Flinn)

October 1, 2004

We start with estimation of the linear (in the parameters) model

$$y = X\beta + \varepsilon,$$

where we assume that:

1. $E(\varepsilon|X) = 0$ for all X (mean independence)
2. $VAR(\varepsilon|X) = E(\varepsilon\varepsilon'|X) = \sigma_\varepsilon^2 I_N$ (homoskedasticity)

The Gauss-Markov Theorem states that

$$\hat{\beta} = (X'X)^{-1}X'y$$

is the Best Linear Unbiased Estimator (BLUE) if ε satisfies (1) and (2).

Proof: An estimator is "best" in a class if it has smaller variance than others estimators in the same class. We are restricting our search for estimators to the class of linear, unbiased ones. Since the data are the y (not the X), we are looking at estimators that are linear functions of y, or

$$\tilde{\beta} = m + My,$$

where β is a $k \times 1$ parameter vector, m is a $k \times 1$ vector of constants, M is a $k \times n$ matrix of constants, and the data vector y is $n \times 1$.

Second, we are restricting attention to the class of unbiased estimators, that is we require that

$$E(\tilde{\beta}) = \beta,$$

for any "valid" possible value β could take, i.e., for all β in the parameter space Ω_β.

First note that if $\tilde{\beta}$ is to be unbiased, then

$$\begin{aligned} E(\tilde{\beta}|X) &= m + ME(y|X) \\ &= m + ME(X\beta + \varepsilon|X) \\ &= m + MX\beta + ME(\varepsilon|X) \\ &= m + MX\beta, \end{aligned}$$

where the last line follows from the mean independence assumption. To be unbiased for any possible value of β then requires

$$m = 0$$

and

$$MX = I_k. \tag{1}$$

We note that the least squares estimator satisfies this requirement for unbiasedness, since for $\hat{\beta}$,

$$M = (X'X)^{-1}X',$$

and

$$MX = (X'X)^{-1}X'X = I_k.$$

Thus looking for linear unbiased estimators requires us to look for estimators of the form

$$\tilde{\beta} = My,$$

for an M that satisfies [1].

Without any loss of generality, we can redefine the M matrix to be of the form

$$M = (X'X)^{-1}X' + C,$$

where C is some $k \times n$ matrix. Using the condition [1] again, we know that for our estimator to be unbiased requires that

$$\begin{aligned} MX &= I_k \\ &\Rightarrow [(X'X)^{-1}X' + C]X = I_k \\ &\Rightarrow I_k + CX = I_k \\ &\Rightarrow CX = 0, \end{aligned}$$

that is, CX is a $k \times k$ matrix of $0's$.

Now we can compute the covariance matrix of all alternative estimator $\tilde{\beta}$. We can write

$$\begin{aligned} \tilde{\beta} &= My \\ &= M(X\beta + \varepsilon) \\ &= \beta + M\varepsilon. \end{aligned}$$

Thus

$$\tilde{\beta} - \beta = M\varepsilon.$$

Since it is unbiased by construction, $E(\tilde{\beta} - \beta|X) = 0$, so the covariance matrix of the estimator is

$$\begin{aligned} E(\tilde{\beta} - \beta)(\tilde{\beta} - \beta)'|X) &= E(M\varepsilon(M\varepsilon)'|X) \\ &= E(M\varepsilon\varepsilon'M'|X) \\ &= M(E\varepsilon\varepsilon'|X)M' \\ &= M\sigma_\varepsilon^2 I_n M' \\ &= \sigma_\varepsilon^2 MM', \end{aligned}$$

which is a $k \times k$ matrix. Now

$$\begin{aligned} MM' &= [(X'X)^{-1}X' + C][(X'X)^{-1}X' + C]' \\ &= (X'X)^{-1}X'X(X'X)^{-1} + (X'X)^{-1}X'C' \\ &\quad +CX(X'X)^{-1} + CC'. \end{aligned}$$

Since $CX = 0$ (and of course $X'C' = 0$ as well),

$$MM' = (X'X)^{-1} + CC'.$$

Now the matrix CC' is a $k \times k$ “cross products” matrix, which by construction cannot be negative definite. The best estimator in a class of estimators is the one with the “smallest” covariance matrix, where by small we mean that the covariance matrix associated with any other estimator in the class (that is, linear and unbiased in the current context) minus the covariance matrix of the best estimator is a positive definite matrix. Formally, the matrix difference

$$MM' - COV(\text{best estimator})$$

is positive definite. Since MM' is minimized when we set the matrix C equal to 0 (that is, it contains $k \times n$ $0's$), the best estimator in the class $\hat{\beta}$. Any other estimator M in this class (in which the C matrix does not contain $0's$ in every row and column) has a strictly “larger” covariance matrix. We conclude that the OLS estimator $\hat{\beta}$ is BLUE under the two conditions set forth (mean independence and homoskedastic).