Chapter 3

Multiple Regression Analysis - Gauss-Markov Theorem

Jim Bang

OLS in Matrix Form

The proofs in this tutorial are a bit easier if we express the estimator in matrix form. In matrix form, OLS solves

$$min_{\hat{\beta}}(X\hat{\beta} - y)'(X\hat{\beta} - y)$$

$$min_{\hat{\beta}}(\hat{\beta}'X'X\hat{\beta} - 2\hat{\beta}'X'y - y'y)$$

The first order condition solves

$$2X'X\hat{\beta} - 2X'y = 0$$

The solution to which is

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

The first part (the part with the inverse) is just the sums of squares of the X variables with themselves (the diagonal) and one another (off-diagonal). The rest is the transpose of the X matrix cross-multiplied with the vector of y values.

Exercise

Using the `wage1` dataset calculate the regression coefficients and their standard errors for the regression of `wage` on education (`educ`) and experience (`exper`).

Note you will need to define `n` (the number of rows in the data), `y`, `X`, `k`, and (after calculating $\hat{\beta}$) `uhat`.

```r
n <- nrow(wage1)
y <- wage1$wage
X <- cbind(1, wage1$educ, wage1$exper)
k <- ncol(X) - 1
bhat <- solve( t(X)%*%X ) %*% t(X)%*%y
uhat <- y - X %*% bhat
sigsqhat <- as.numeric( t(uhat) %*% uhat / (n-k-1) )
SER <- sqrt(sigsqhat)
Vbetahat <- sigsqhat * solve( t(X)%*%X )
se <- sqrt( diag(Vbetahat) )
```

Gauss-Markov Theorem

Under the assumptions of the classical regression model, the OLS estimates of $\hat{\beta}$ are BLUE.

- Best: Efficient, or minimum-variance (compared to estimates that fit the other criteria).
- Linear: Linear in the parameters.
- Unbiased: $E(\hat{\beta}_{OLS}) = \beta$.
- Estimator.

Since linearity is trivially true by OLS assumption #1, we will prove bestness and unbiasedness, starting with unbiasedness (OLS needs to be unbiased before it can be the *best* unbiased!)

But wait! Actually, Hansen (2022) shows that under the classical assumptions (without assuming normality in the errors as is assumed in your book) OLS is even best among *nonlinear* estimators (i.e. BUE or MVUE)!

Unbiased-ness

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

Substituting $X\hat{\beta} - u$ for y,

$$E(\hat{\beta}) = E[(X'X)^{-1}X'(X\beta + u)]$$

Distributing $(X'X)^{-1}X'$ through the parentheses,

$$E(\hat{\beta}) = E[(X'X)^{-1}X'X\beta + (X'X)^{-1}X'u)]$$

In the first term, note that the anything times its multiplicative inverse (one over something, but in this case the matrix inverse) cancels out (becomes an identity matrix, or just 1 for a scalar).

$$E(\hat{\beta}) = \beta + E[(X'X)^{-1}X'u]$$

In the second term, note that zero conditional expectation of the errors implies that X and u are uncorrelated, or $E(X'u) - E(X)E(u) = 0$. Since $E(u) = 0$, this implies $E(X'u) = 0$.

Efficiency Relative to Other Linear Estimators

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

Let $Z = (X'X)^{-1}X'$ and substitute $y = X\beta + u$, so we have

$$Var(\hat{\beta}) = Var[Z(X\beta + u)]$$

$$Var(\hat{\beta}) = Var(ZX\beta + Zu)$$

Since β is a constant (albeit unknown),

$$Var(\hat{\beta}) = Var(Zu)$$

Factoring out Z (which we have to "square" (in matrix terms) by the rules of the variance of a linear combination) we have

$$Var(\hat{\beta}) = Z'ZVar(u) = Z'Z\sigma^2$$

$$Var(\hat{\beta}) = X(X'X)^{-1}(X'X)^{-1}X'\sigma^2 = (X'X)^{-1}\sigma^2$$

Comparison to other linear, unbiased estimators

$$Var(\hat{\beta}) = Var(Zy) = Z'Z\sigma^2$$

Let $C = Z + D$ so that it gives a different linear estimate, $\tilde{\beta} = Cy$. For $\tilde{\beta}$ to be unbiased, DX must equal 0. Then through a similar set of steps as the previous slide,

$$Var(\tilde{\beta}) = Var(Cy) = C'C\sigma^2$$

Reverting to $Z + D$ for C,

$$Var(\tilde{\beta}) = (Z + D)'(Z + D)\sigma^2$$

$$Var(\tilde{\beta}) = (Z'Z + 2Z'D + D'D)\sigma^2$$

Since $\tilde{\beta}$ is still unbiased $Z'D = (X'X)^-1XD = 0$, so

$$Var(\tilde{\beta}) = (Z'Z + D'D)\sigma^2 = Var(\hat{\beta}) + D'D\sigma^2$$

Since $D'D$ is like squaring the matrix, D, it is *positive definite*, which means that the any linear adjustment to the OLS estimator that is also unbiased will have higher variance.