Multiple Linear Regression

The Gauss-Markov Theorem

Math 392

Gauss-Markov Theorem

Claim: $\hat{\beta}_{LS}$ are the best linear unbiased estimates (BLUE) of β.

What do we mean by *linear*?

Linear refers to estimates that are a linear function of the random variable, which in the regression setting is Y.

$$\hat{\beta}_j = c_{1,j}Y_1 + c_{2,j}Y_2 + \ldots c_{n,j}Y_n$$

Or in matrix form:

$$\hat{\beta} = CY$$

In the least squares case, $C = (X'X)^{-1}X'$.

What do we mean by *unbiased*?

The same thing that usually mean:

$$E(\hat{\beta}) = \beta.$$

What do we mean by *best*?

Best = lowest MSE. Among unbiased estimators, this will be the one with the lowest variance. How do we compare variances two *vectors*?

$\hat{\beta}_{LS}$ is best if $Var(\hat{\beta}_{LS}) - Var(\tilde{\beta})$ is positive semi-definite for all other estimates $\tilde{\beta}$.

A $p \times p$ matrix M is positive semi-definite if

$$a' M a \geq 0$$

for all non-zero vectors $a \in \mathbb{R}^p$.

Assumptions

The proof assumes that

1. $Y = X\beta + \epsilon$
2. $E(\epsilon) = 0$
3. $Var(\epsilon) = \sigma^2 I$

Proof

Let $\tilde{\beta} = CY, C = (X'X)^{-1}X' + D$ for any $p \times n$ non-zero matrix D.

$$\begin{aligned} E(\tilde{\beta}) = E(CY) &= CE(Y) \\ &= CE(X\beta + \epsilon) \\ &= CE(X\beta) + CE(\epsilon) \\ &= CX\beta \\ &= ((X'X)^{-1}X' + D)X\beta \\ &= (X'X)^{-1}X'X\beta + DX\beta \\ &= \beta + DX\beta \end{aligned}$$

Since we're interested in the class of unbiased estimators, DX must be 0.

Proof, cont.

$$\begin{aligned}
Var(\tilde{\beta}) &= Var(CY) \\
&= CVar(Y)C' \\
&= C\sigma^2 IC' \\
&= \sigma^2((X'X)^{-1}X' + D)[(X'X)^{-1}X' + D]' \\
&= \sigma^2((X'X)^{-1}X' + D)(X(X'X)^{-1} + D') \\
&= \sigma^2[(X'X)^{-1}X'X(X'X)^{-1} + (X'X)^{-1}X'D' + DX(X'X)^{-1} + DD'] \\
&= \sigma^2(X'X)^{-1} + \sigma^2 DD' \\
&= Var(\hat{\beta}_{LS}) + \sigma^2 DD'
\end{aligned}$$

Since $\sigma^2 DD'$ is a positive semi-definite matrix, this shows that *any* other linear estimator of β will have a variance at least as large as the least squares estimates.

Postscript: Why we love OLS

- As a loss function, RSS has some intuitive appeal.
- $\hat{\beta}_{LS}$ has a closed form.
- $\hat{\beta}_{LS}$ are BLUE (Gauss-Markov).