Alastair Hall | **ECON61001: Econometric Methods** | Semester 1

Study Group Questions # 6

We argued earlier in the course that our preferred choice of estimator is the minimum variance unbiased estimator or equivalently the most efficient unbiased estimator. While this is a common approach to selection of estimators in classical statistics, it is not the only one. Here we explore an alternative criterion for estimator selection based on minimizing the *the mean-squared error* (MSE) of the estimator. This alternative approach can be relevant in so-called "big data" environments as we discuss in the Tutorial session. In this assignment, you compare the mean-squared error of the OLS and Ridge regression estimator.

As in the Tutorial, consider the Classical Linear regression that is,

$$y = X\beta_0 + u$$

where Assumptions CA1-CA5 hold. Let $\hat{\beta}$ and $\hat{\beta}_{RR}(\eta)$ denote respectively the OLS estimator and the Ridge Regression estimator of β_0.

Please enter the details of the group:

Group name	Group Matrix
Student ID 1	11335127
Student ID 2	10710007
Student ID 3	10704589
Student ID 4	11465531
Student ID 5	

1. Show that $E[\hat{\beta}_{RR}(\eta)] = A_\eta \beta_0$ where A_η is defined in the Tutorial session.

The Ridge Regression estimator is:

$$\hat{\beta}_{RR}(\eta) = (X'X + \eta I_K)^{-1} X'y$$
$$= A_\eta \hat{\beta}$$

with $A_\eta = \{I_K + \eta (X'X)^{-1}\}^{-1}$ and $\hat{\beta}$ is the OLS estimator of β_0.

$$\Rightarrow \quad E[\hat{\beta}_{RR}(\eta)] = E[A_\eta \cdot \hat{\beta}]$$
$$= E[A_\eta] \cdot E[\hat{\beta}] + \mathrm{cov}[A_\eta, \hat{\beta}]$$

As $CA_1 \rightarrow CA_5$ hold $\Rightarrow \mathrm{cov}[A_\eta, \hat{\beta}] = 0$

$$\Rightarrow E[\hat{\beta}_{RR}(\eta)] = E[A_\eta] \cdot E[\hat{\beta}]$$

$E[A_\eta] = A_\eta$ because X is fixed (CA_2)

$$E[\hat{\beta}] = E[(X'X)^{-1} X'(X\beta_0 + u)] = \beta_0 + (X'X)^{-1} X' E[u]$$
$$= \beta_0 \quad (CA_2, CA_4).$$

$$\Rightarrow E[\hat{\beta}_{RR}(\eta)] = A_\eta \beta_0.$$

2. Show that

$$Var[\hat{\beta}_{RR}(\eta)] \;=\; \sigma_0^2 \{ X'X \,+\, \eta I_k \}^{-1} X'X \{ X'X \,+\, \eta I_k \}^{-1}.$$

We have $\hat{\beta}_{RR}(\eta) = (X'X + \eta I_k)^{-1} X'y$

$= (X'X + \eta I_k)^{-1} X'(X\beta_0 + u)$

$\Rightarrow E[\hat{\beta}_{RR}(\eta)] = (X'X + \eta I_k)^{-1} X'X\beta_0$

$\Rightarrow \hat{\beta}_{RR}(\eta) - E[\hat{\beta}_{RR}(\eta)] = (X'X + \eta I_k)^{-1} X'u$

Examine $Var[\hat{\beta}_{RR}(\eta)] = E[(\hat{\beta}_{RR}(\eta) - E[\hat{\beta}_{RR}(\eta)])(\hat{\beta}_{RR}(\eta) - E[\hat{\beta}_{RR}(\eta)])']$

$= E[(X'X + \eta I_k)^{-1} X'uu'X(X'X + \eta I_k)^{-1}]$

$= (X'X + \eta I_k)^{-1} X' E[uu'] X (X'X + \eta I_k)^{-1}$

$= \sigma_0^2 (X'X + \eta I_k)^{-1} X'X (X'X + \eta I_k)^{-1}.$

Because, with $CA_1 \to CA_5$ assumptions, X is fixed and $Var[u] = \sigma_0^2$

3. Show that $Var[\hat{\beta}] - Var[\hat{\beta}_{RR}(\eta)]$ equals a positive definite matrix. *Hint: For two conformable matrices A and B, if $B^{-1} - A^{-1}$ is positive definite then $A - B$ is positive definite.*

$\hat{\beta}$ is the OLS estimator of β_0

$$\Rightarrow \hat{\beta} \sim N(0, \sigma_0^2 (X'X)^{-1})$$

$$Var[\hat{\beta}] = \sigma_0^2 (X'X)^{-1}$$

$$Var[\hat{\beta}_{RR}(\eta)] = \sigma_0^2 (X'X + \eta I_K)^{-1} X'X (X'X + \eta I_K)^{-1}$$

$$\Rightarrow Var[\hat{\beta}] - Var[\hat{\beta}_{RR}(\eta)] = \sigma_0^2 \{ \underbrace{(X'X)^{-1}}_{A} - \underbrace{(X'X + \eta I_K)^{-1} X'X (X'X + \eta I_K)^{-1}}_{B.} \}$$

Let's examine $B^{-1} - A^{-1}$ with $\begin{cases} B^{-1} = (X'X + \eta I_K)(X'X)^{-1}(X'X + \eta I_K) \\ A^{-1} = (X'X) \end{cases}$

$$\begin{aligned} B^{-1} - A^{-1} &= (X'X + \eta I_K)(X'X)^{-1}(X'X + \eta I_K) - (X'X) \\ &= [I_K + \eta (X'X)^{-1}](X'X + \eta I_K) - (X'X) \\ &= (X'X) + \eta I_K + \eta I_K + \eta^2 (X'X)^{-1} - (X'X) \\ &= 2\eta I_K + \eta^2 (X'X)^{-1}. \end{aligned}$$

As η is a positive constant $\rightarrow \eta I_K$ is positive definite.

As our assumptions hold, $X'X$ is positive definite

$\rightarrow (X'X)^{-1}$ is also positive definite.

$\Rightarrow B^{-1} - A^{-1}$ is positive definite

$\rightarrow A - B$ is positive definite

Hence $Var[\hat{\beta}] - Var[\hat{\beta}_{RR}(\eta)]$ is positive definite.

4. What can be say about the ranking of $\hat{\beta}$ and $\hat{\beta}_{RR}(\eta)$ in terms of MSE?

From Tutorial, it can be taken that for any estimator $\hat{\theta}$ of a true parameter θ, the MSE is:

$$MSE(\hat{\theta}) = Var[\hat{\theta}] + \{bias(\hat{\theta})\}^2.$$

First, compare the variance:

From Question 3, it can be seen that $\hat{\beta}_{RR}(\eta)$ has smaller variance than $\hat{\beta}$.

Yet, in terms of biasedness, as Ridge estimator introduces additional terms (ηI_k) to the equation, it's biased from β_0.

$$E[\hat{\beta}_{RR}(\eta)] = (X'X + \eta I_k)^{-1} X'X \beta_0 \neq \beta_0$$

$\rightarrow$ While $\{bias[\hat{\beta}]\}^2 = 0$; $\{bias[\hat{\beta}_{RR}(\eta)]\}^2 > 0$.

In conclusion, although $\hat{\beta}_{RR}(\eta)$ has smaller variance than $\hat{\beta}_{OLS}$, due to it's biasedness, we can't say much about their MSE.

$$\{Bias[\hat{\beta}_{RR}(\eta)]\}^2 = \beta_0 [(X'X + \eta I_k)^{-1} X'X - I_k] = A$$

$$Var[\hat{\beta}] - Var[\hat{\beta}_{RR}(\eta)] = 2\eta I_k + \eta^2 (X'X)^{-1}. \quad = B$$

If $A > B$, $MSE[\hat{\beta}] < MSE[\hat{\beta}_{RR}(\eta)]$

If $A < B$, $MSE[\hat{\beta}] > MSE[\hat{\beta}_{RR}(\eta)]$.