BLUE property of Least Squared Estimator

November 1, 2021

Preliminary: Let $\mathbf{A}$ be $k \times k$ symmetric matrix. We say $\mathbf{A}$ is positive semidefinite(definite) if for any $\mathbf{v} \in \mathbb{R}^k \setminus \{\mathbf{0}\}$, $\mathbf{v}^t\mathbf{A}\mathbf{v} \geq 0 (> 0)$. Note that for any $m \times n$ matrix $\mathbf{C}$, $\mathbf{C}^t\mathbf{C}$ is symmetric and positive semidefinite.

Assumption. $\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}$, where $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)^t$, $\mathbf{1} = (1, \ldots, 1)^t, \mathbf{x}_j = (x_{1j}, \ldots, x_{nj})^t$, $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)^t$ are n-dimensional vectors, $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_n)^t$ is a $(p+1)$-dimensional vector, $\mathbf{X} = (\mathbf{1}, \mathbf{x_1}, \ldots, \mathbf{x_p})$ is a $n \times (p+1)$ known matrix of full column rank and $\boldsymbol{\epsilon} \sim \boldsymbol{N}_n(0, \sigma^2 I_n)$ with $\sigma^2 > 0$.

In the lecture, we've mentioned the BLUE property of Least Squared Estimator. BLUE stands for "Best Linear Unbiased Estimator". The "Best" is in sense of variance. Usually, we prefer the estimator with smaller variance. So the estimator may be "Best", if it has the smallest variance among all possible estimators. However, generally, there is no nontrivial estimator with the smallest variance. Hence, we my put restriction on the class of estimators so that we can discuss estimator with the smallest variance among estimators in the given class.

Thus, to say Least Squared Estimator is BLUE, it means among the linear and unibased estimators for $\boldsymbol{\beta}$, Least Squared Estimator $\hat{\boldsymbol{\beta}}$ has the smallest variance. But as $\hat{\boldsymbol{\beta}}$ is $(p+1)$-dimensional vector, it may be quite awkward to say $\hat{\boldsymbol{\beta}}$ has the "smallest" variance, because variance of $\hat{\boldsymbol{\beta}}$ is matrix. So, to compare the variance, we have to clarify meaning of $\geq$. Let $\mathbf{A}, \mathbf{B}$ be $k \times k$ symmetric matrix. Then we may say $\mathbf{A} \geq \mathbf{B}$ if $\mathbf{A} - \mathbf{B}$ is positive semidefinite. We may use $\succeq$ instead of $\geq$ not to be confused with the usual meaning of $\geq$.

Now let us prove $\hat{\boldsymbol{\beta}}$ is indeed BLUE. Suppose $\mathbf{CY}$ is linear unbiased estimator of $\boldsymbol{\beta}$, where $\mathbf{C}$ is $(p+1) \times n$ matrix. Then, since $\mathbf{CY}$ is unbiased estimator,

$$\mathrm{E}(\mathbf{CY}) = \mathbf{C}\mathrm{E}\mathbf{Y} = \mathbf{CX}\boldsymbol{\beta} = \boldsymbol{\beta}$$

Because this equality holds for any possible $\boldsymbol{\beta}$, $\mathbf{CX} = \mathbf{I}_{p+1}$. Variance of $\mathbf{CY}$ is $\mathrm{Var}(\mathbf{CY}) = \mathbf{C}\mathrm{Var}(\mathbf{Y})\mathbf{C}^t = \mathbf{C}(\sigma^2\mathbf{I}_n)\mathbf{C}^t = \sigma^2\mathbf{CC}^t$. As $\mathrm{Var}(\hat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}^t\mathbf{X})^{-1}$, to show $\mathrm{Var}(\mathbf{CY}) \succeq \mathrm{Var}(\hat{\boldsymbol{\beta}})$, it suffices to show $\mathbf{CC}^t \succeq (\mathbf{X}^t\mathbf{X})^{-1}$.

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

Here the fourth equality holds for $\mathbf{CX} = \mathbf{I}_{p+1}$. Since $[\mathbf{C} - (\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t][\mathbf{C} - (\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t]^t$ is symmetric and positive semidefinite, as we've mentioned in **Preliminary**, and

$$\mathbf{CC}^t - (\mathbf{X}^t\mathbf{X})^{-1} = [\mathbf{C} - (\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t][\mathbf{C} - (\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t]^t,$$

we see that $\mathbf{CC}^t \succeq (\mathbf{X}^t\mathbf{X})^{-1}$ and this finishes the proof.