The Gauss-Markov Model

Recall that

$$\begin{aligned}
\text{Cov}(u, v) &= E((u - E(u))(v - E(v))) \\
&= E(uv) - E(u)E(v) \\
\text{Var}(u) &= \text{Cov}(u, u) \\
&= E(u - E(u))^2 \\
&= E(u^2) - (E(u))^2.
\end{aligned}$$

If $\underset{m\times 1}{\boldsymbol{u}}$ and $\underset{n\times 1}{\boldsymbol{v}}$ are random vectors, then

$$\begin{aligned} \text{Cov}(\boldsymbol{u}, \boldsymbol{v}) &= [\text{Cov}(u_i, v_j)]_{m\times n} \quad \text{and} \\ \text{Var}(\boldsymbol{u}) &= [\text{Cov}(u_i, u_j)]_{m\times m}. \end{aligned}$$

It follows that

$$\begin{aligned}
\text{Cov}(\boldsymbol{u}, \boldsymbol{v}) &= E((\boldsymbol{u} - E(\boldsymbol{u}))(\boldsymbol{v} - E(\boldsymbol{v}))') \\
&= E(\boldsymbol{u}\boldsymbol{v}') - E(\boldsymbol{u})E(\boldsymbol{v}') \\
\text{Var}(\boldsymbol{u}) &= E((\boldsymbol{u} - E(\boldsymbol{u}))(\boldsymbol{u} - E(\boldsymbol{u}))') \\
&= E(\boldsymbol{u}\boldsymbol{u}') - E(\boldsymbol{u})E(\boldsymbol{u}').
\end{aligned}$$

(Note that if $\boldsymbol{Z} = [z_{ij}]$, then $E(\boldsymbol{Z}) \equiv [E(z_{ij})]$; i.e., the expected value of a matrix is defined to be the matrix of the expected values of the elements in the original matrix.)

From these basis definitions, it is straightforward to show the following:

If $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{A}, \boldsymbol{B}$ are fixed and $\boldsymbol{u}, \boldsymbol{v}$ are random, then

$$\text{Cov}(\boldsymbol{a} + \boldsymbol{A}\boldsymbol{u}, \boldsymbol{b} + \boldsymbol{B}\boldsymbol{v}) = \boldsymbol{A}\text{Cov}(\boldsymbol{u}, \boldsymbol{v})\boldsymbol{B}'.$$

Some common special cases are

$$\begin{aligned}
\mathrm{Cov}(\boldsymbol{Au}, \boldsymbol{Bv}) &= \boldsymbol{A}\mathrm{Cov}(\boldsymbol{u}, \boldsymbol{v})\boldsymbol{B}' \\
\mathrm{Cov}(\boldsymbol{Au}, \boldsymbol{Bu}) &= \boldsymbol{A}\mathrm{Cov}(\boldsymbol{u}, \boldsymbol{u})\boldsymbol{B}' \\
&= \boldsymbol{A}\mathrm{Var}(\boldsymbol{u})\boldsymbol{B}' \\
\mathrm{Var}(\boldsymbol{Au}) &= \mathrm{Cov}(\boldsymbol{Au}, \boldsymbol{Au}) \\
&= \boldsymbol{A}\mathrm{Cov}(\boldsymbol{u}, \boldsymbol{u})\boldsymbol{A}' \\
&= \boldsymbol{A}\mathrm{Var}(\boldsymbol{u})\boldsymbol{A}'.
\end{aligned}$$

The Gauss-Markov Model

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

where

$$E(\boldsymbol{\varepsilon}) = \mathbf{0} \quad \text{and} \quad \text{Var}(\boldsymbol{\varepsilon}) = \sigma^2\boldsymbol{I} \text{ for some unknown } \sigma^2.$$

“Mean zero, constant variance, uncorrelated errors.”

More succinct statement of Gauss-Markov Model:

$$E(\boldsymbol{y}) = \boldsymbol{X}\boldsymbol{\beta}, \quad \mathrm{Var}(\boldsymbol{y}) = \sigma^2\boldsymbol{I}$$

or

$$E(\boldsymbol{y}) \in \mathcal{C}(\boldsymbol{X}), \quad \mathrm{Var}(\boldsymbol{y}) = \sigma^2\boldsymbol{I}.$$

Note that the Gauss-Markov Model (GMM) is a special case of the GLM that we have been studying.

Hence, all previous results for the GLM apply to the GMM.

Suppose $\boldsymbol{c}'\boldsymbol{\beta}$ is an estimable function.

Find $\mathrm{Var}(\boldsymbol{c}'\hat{\boldsymbol{\beta}})$, the variance of the LSE of $\boldsymbol{c}'\boldsymbol{\beta}$.

$$
\begin{aligned}
\mathrm{Var}(\boldsymbol{c}'\hat{\boldsymbol{\beta}}) &= \mathrm{Var}(\boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y}) \\
&= \mathrm{Var}(\boldsymbol{a}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y}) \\
&= \mathrm{Var}(\boldsymbol{a}'\boldsymbol{P_X}\boldsymbol{y}) \\
&= \boldsymbol{a}'\boldsymbol{P_X}\mathrm{Var}(\boldsymbol{y})\boldsymbol{P_X}'\boldsymbol{a} \\
&= \boldsymbol{a}'\boldsymbol{P_X}(\sigma^2\boldsymbol{I})\boldsymbol{P_X}\boldsymbol{a} \\
&= \sigma^2\boldsymbol{a}'\boldsymbol{P_X}\boldsymbol{P_X}\boldsymbol{a} \\
&= \sigma^2\boldsymbol{a}'\boldsymbol{P_X}\boldsymbol{a} \\
&= \sigma^2\boldsymbol{a}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{a} \\
&= \sigma^2\boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{c}.
\end{aligned}
$$

Example: Two-treatment ANCOVA

$$y_{ij} = \mu + \tau_i + \gamma x_{ij} + \varepsilon_{ij} \quad i = 1, 2;\ j = 1, \ldots, m$$

$$E(\varepsilon_{ij}) = 0 \quad \forall\ i = 1, 2;\ j = 1, \ldots, m$$

$$\text{Cov}(\varepsilon_{ij}, \varepsilon_{st}) = \begin{cases} 0 & \text{if } (i,j) \neq (s,t) \\ \sigma^2 & \text{if } (i,j) = (s,t). \end{cases}$$

x_{ij} is the known value of a covariate for treatment i and observation j $(i = 1, 2;\ j = 1, \ldots, m)$.

Under what conditions is γ estimable?

γ is estimable $\iff \cdots ?$

γ is estimable iff $x_{11}, \ldots, x_{1m}$ are not all equal, or $x_{21}, \ldots, x_{2m}$ are not all equal, i.e.,

$$\boldsymbol{x} = \begin{bmatrix} x_{11} \\ \vdots \\ x_{1m} \\ x_{21} \\ \vdots \\ x_{2m} \end{bmatrix} \notin \mathcal{C}\left(\begin{bmatrix} \underset{m\times 1}{\mathbf{1}} & \underset{m\times 1}{\mathbf{0}} \\ \underset{m\times 1}{\mathbf{0}} & \underset{m\times 1}{\mathbf{1}} \end{bmatrix} \right).$$

This makes sense intuitively, but to prove the claim formally...

$$\boldsymbol{X} = \begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{0} & \boldsymbol{x}_1 \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & \boldsymbol{x}_2 \end{bmatrix}, \quad \text{where} \quad \boldsymbol{x}_i = \begin{bmatrix} x_{i1} \\ \vdots \\ x_{im} \end{bmatrix} \quad i = 1, 2.$$

$$\boldsymbol{\beta} = \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \gamma \end{bmatrix}.$$

If $\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix} \in \mathcal{C}\left(\begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \right)$, then

$$\boldsymbol{X} = \begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{0} & a\mathbf{1} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & b\mathbf{1} \end{bmatrix} \quad \text{for some} \quad a, b \in \mathbb{R}.$$

$$\boldsymbol{z} \in \mathcal{N}(\boldsymbol{X}) \iff \begin{cases} z_1 + z_2 + az_4 & = 0 \\ z_1 + z_3 + bz_4 & = 0 \end{cases}$$

$$\Rightarrow \begin{bmatrix} 0 \\ -a \\ -b \\ 1 \end{bmatrix} \in \mathcal{N}(\boldsymbol{X}).$$

Now note that

$$\begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ -a \\ -b \\ 1 \end{bmatrix} = 1 \neq 0.$$

It follows that

$$\begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \gamma \end{bmatrix} = \gamma$$

is not estimable.

This proves that if γ is estimable, then $\boldsymbol{x} \notin \mathcal{C}\left(\begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}\right)$.

Conversely, if $\boldsymbol{x} \notin \mathcal{C}\left(\begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}\right)$, then $\exists\ x_{ij} \neq x_{ij'}$ for some i and $j \neq j'$.

Taking the difference of the corresponding rows of $\boldsymbol{X}$ yields

$$\begin{bmatrix} 0 & 0 & 0 & x_{ij} - x_{ij'} \end{bmatrix}.$$

Dividing by $x_{ij} - x_{ij'}$, yields

$$\begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}.$$

Thus,

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \in \mathcal{C}(\boldsymbol{X}'), \quad \text{and} \quad \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \gamma \end{bmatrix} = \gamma \text{ is estimable.}$$

To find the LSE of γ, recall $\boldsymbol{X} = \begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{0} & \boldsymbol{x}_1 \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & \boldsymbol{x}_2 \end{bmatrix}$. Thus

$$\boldsymbol{X}'\boldsymbol{X} = \begin{bmatrix} 2m & m & m & x_{\cdot\cdot} \\ m & m & 0 & x_{1\cdot} \\ m & 0 & m & x_{2\cdot} \\ x_{\cdot\cdot} & x_{1\cdot} & x_{2\cdot} & \boldsymbol{x}'\boldsymbol{x} \end{bmatrix}.$$

When $\boldsymbol{x}$ is not in $\mathcal{C}\left(\begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}\right)$, $rank(\boldsymbol{X}) = rank(\boldsymbol{X}'\boldsymbol{X}) = 3$. Thus, a GI of $\boldsymbol{X}'\boldsymbol{X}$ is

$$\begin{bmatrix} 0 & \underset{1\times 3}{\mathbf{0}} \\ \underset{3\times 1}{\mathbf{0}} & \boldsymbol{A} \end{bmatrix}, \quad \text{where} \quad \boldsymbol{A} = \begin{bmatrix} m & 0 & x_{1\cdot} \\ 0 & m & x_{2\cdot} \\ x_{1\cdot} & x_{2\cdot} & \boldsymbol{x}'\boldsymbol{x} \end{bmatrix}^{-1}.$$

Because $\begin{bmatrix} m & 0 & x_{1\cdot} \\ 0 & m & x_{2\cdot} \\ x_{1\cdot} & x_{2\cdot} & \boldsymbol{x}'\boldsymbol{x} \end{bmatrix}$ is not so easy to invert, let's consider a different strategy.

To find LSE of γ and its variance in an alternative way, consider

$$\boldsymbol{W} = \begin{bmatrix} \mathbf{1} & \mathbf{0} & \boldsymbol{x}_1 - \bar{x}_{1\cdot}\mathbf{1} \\ \mathbf{0} & \mathbf{1} & \boldsymbol{x}_2 - \bar{x}_{2\cdot}\mathbf{1} \end{bmatrix}.$$

This matrix arises by dropping the first column of $\boldsymbol{X}$ and applying GS orthogonalization to remaining columns.

$$
\begin{array}{ccc}
\boldsymbol{X} & \boldsymbol{T} & = \quad \boldsymbol{W} \\
\begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{0} & \boldsymbol{x}_1 \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & \boldsymbol{x}_2 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -\bar{x}_{1\cdot} \\ 0 & 1 & -\bar{x}_{2\cdot} \\ 0 & 0 & 1 \end{bmatrix} & = \begin{bmatrix} \mathbf{1} & \mathbf{0} & \boldsymbol{x}_1 - \bar{x}_{1\cdot}\mathbf{1} \\ \mathbf{0} & \mathbf{1} & \boldsymbol{x}_2 - \bar{x}_{2\cdot}\mathbf{1} \end{bmatrix}.
\end{array}
$$

$$\begin{bmatrix} \mathbf{1} & \mathbf{0} & \boldsymbol{x}_1 - \bar{x}_{1\cdot}\mathbf{1} \\ \mathbf{0} & \mathbf{1} & \boldsymbol{x}_2 - \bar{x}_{2\cdot}\mathbf{1} \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & \bar{x}_{1\cdot} \\ 1 & 0 & 1 & \bar{x}_{2\cdot} \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{0} & \boldsymbol{x}_1 \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & \boldsymbol{x}_2 \end{bmatrix}$$

$$\boldsymbol{W} \qquad\qquad \boldsymbol{S} \qquad = \qquad \boldsymbol{X}$$

$$\boldsymbol{W}'\boldsymbol{W} = \begin{bmatrix} \mathbf{1} & \mathbf{0} & \boldsymbol{x}_1 - \bar{x}_{1\cdot}\mathbf{1} \\ \mathbf{0} & \mathbf{1} & \boldsymbol{x}_2 - \bar{x}_{2\cdot}\mathbf{1} \end{bmatrix}' \begin{bmatrix} \mathbf{1} & \mathbf{0} & \boldsymbol{x}_1 - \bar{x}_{1\cdot}\mathbf{1} \\ \mathbf{0} & \mathbf{1} & \boldsymbol{x}_2 - \bar{x}_{2\cdot}\mathbf{1} \end{bmatrix}$$

$$= \begin{bmatrix} m & 0 & \mathbf{1}'\boldsymbol{x}_1 - \bar{x}_{1\cdot}\mathbf{1}'\mathbf{1} \\ 0 & m & \mathbf{1}'\boldsymbol{x}_2 - \bar{x}_{2\cdot}\mathbf{1}'\mathbf{1} \\ \mathbf{1}'\boldsymbol{x}_1 - \bar{x}_{1\cdot}\mathbf{1}'\mathbf{1} & \mathbf{1}'\boldsymbol{x}_2 - \bar{x}_{2\cdot}\mathbf{1}'\mathbf{1} & \sum_{i=1}^{2}\sum_{j=1}^{m}(x_{ij} - \bar{x}_{i\cdot})^2 \end{bmatrix}.$$

$$= \begin{bmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & \sum_{i=1}^{2} \sum_{j=1}^{m} (x_{ij} - \bar{x}_{i\cdot})^2 \end{bmatrix}$$

$$(\boldsymbol{W}'\boldsymbol{W})^{-1} = \begin{bmatrix} \frac{1}{m} & 0 & 0 \\ 0 & \frac{1}{m} & 0 \\ 0 & 0 & \frac{1}{\sum_{i=1}^{2} \sum_{j=1}^{m} (x_{ij} - \bar{x}_{i\cdot})^2} \end{bmatrix}.$$

$$
\boldsymbol{W}'\boldsymbol{y} = \begin{bmatrix} \mathbf{1} & \mathbf{0} & \boldsymbol{x}_1 - \bar{x}_{1\cdot}\mathbf{1} \\ \mathbf{0} & \mathbf{1} & \boldsymbol{x}_2 - \bar{x}_{2\cdot}\mathbf{1} \end{bmatrix}' \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{bmatrix}
$$

$$
= \begin{bmatrix} y_{1\cdot} \\ y_{2\cdot} \\ \sum_{i=1}^{2} \sum_{j=1}^{m} y_{ij}(x_{ij} - \bar{x}_{i\cdot}) \end{bmatrix}.
$$

$$
\begin{aligned}
\hat{\boldsymbol{\alpha}} &= (\boldsymbol{W}'\boldsymbol{W})^{-}\boldsymbol{W}'\boldsymbol{y} \\
&= \begin{bmatrix} \bar{y}_{1\cdot} \\ \bar{y}_{2\cdot} \\ \frac{\sum_{i=1}^{2}\sum_{j=1}^{m} y_{ij}(x_{ij}-\bar{x}_{i\cdot})}{\sum_{i=1}^{2}\sum_{j=1}^{m}(x_{ij}-\bar{x}_{i\cdot})^2} \end{bmatrix}.
\end{aligned}
$$

Recall that

$$\begin{aligned} E(\boldsymbol{y}) &= \boldsymbol{X\beta} = \boldsymbol{W\alpha} \\ &= \boldsymbol{WS\beta} = \boldsymbol{XT\alpha}. \end{aligned}$$

$\boldsymbol{c}'\boldsymbol{\beta}$ estimable $\Rightarrow \boldsymbol{c}'\boldsymbol{T\alpha}$ is estimable with LSE $\boldsymbol{c}'\boldsymbol{T}\hat{\boldsymbol{\alpha}}$.

Thus, LSE of $\boldsymbol{\beta}$ is

$$\overset{\boldsymbol{c}'}{\begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}} \overset{\boldsymbol{T}}{\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -\bar{x}_{1\cdot} \\ 0 & 1 & -\bar{x}_{2\cdot} \\ 0 & 0 & 1 \end{bmatrix}} \overset{\hat{\boldsymbol{\alpha}}}{\begin{bmatrix} \bar{y}_{1\cdot} \\ \bar{y}_{2\cdot} \\ \frac{\sum_{i=1}^{2}\sum_{j=1}^{m} y_{ij}(x_{ij}-\bar{x}_{i\cdot})}{\sum_{i=1}^{2}\sum_{j=1}^{m}(x_{ij}-\bar{x}_{i\cdot})^2} \end{bmatrix}}$$

$$= \frac{\sum_{i=1}^{2}\sum_{j=1}^{m} y_{ij}(x_{ij}-\bar{x}_{i\cdot})}{\sum_{i=1}^{2}\sum_{j=1}^{m}(x_{ij}-\bar{x}_{i\cdot})^2}.$$

Note that in this case

$$\boldsymbol{c}'\hat{\boldsymbol{\beta}} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \hat{\boldsymbol{\alpha}} = \hat{\alpha}_3.$$

$$\begin{aligned} \text{Var}(\boldsymbol{c}'\hat{\boldsymbol{\beta}}) &= \text{Var}\left(\begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \hat{\boldsymbol{\alpha}} \right) \\ &= \sigma^2 \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} (\boldsymbol{W}'\boldsymbol{W})^{-1} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ &= \sigma^2 \frac{1}{\sum_{i=1}^{2} \sum_{j=1}^{m} (x_{ij} - \bar{x}_{i\cdot})^2}. \end{aligned}$$

Theorem 4.1 (Gauss-Markov Theorem):

Suppose the GMM holds. If $\boldsymbol{c}'\boldsymbol{\beta}$ is estimable, then the LSE $\boldsymbol{c}'\hat{\boldsymbol{\beta}}$ is the best (minimum variance) linear unbiased estimator (BLUE) of $\boldsymbol{c}'\boldsymbol{\beta}$.

Proof of Theorem 4.1:

$\boldsymbol{c}'\boldsymbol{\beta}$ is estimable $\Rightarrow \boldsymbol{c}' = \boldsymbol{a}'\boldsymbol{X}$ for some vector $\boldsymbol{a}$.

The LSE of $\boldsymbol{c}'\boldsymbol{\beta}$ is

$$\begin{aligned}\boldsymbol{c}'\hat{\boldsymbol{\beta}} &= \boldsymbol{a}'\boldsymbol{X}\hat{\boldsymbol{\beta}} \\ &= \boldsymbol{a}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y} \\ &= \boldsymbol{a}'\boldsymbol{P}_{\boldsymbol{X}}\boldsymbol{y}.\end{aligned}$$

Now suppose $u + \boldsymbol{v}'\boldsymbol{y}$ is any other linear unbiased estimator of $\boldsymbol{c}'\boldsymbol{\beta}$.

Then

$$
\begin{aligned}
& E(u + \boldsymbol{v}'\boldsymbol{y}) = \boldsymbol{c}'\boldsymbol{\beta} \quad \forall\ \boldsymbol{\beta} \in \mathbb{R}^p \\
& \iff u + \boldsymbol{v}'\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{c}'\boldsymbol{\beta} \quad \forall\ \boldsymbol{\beta} \in \mathbb{R}^p \\
& \iff u = 0 \quad \text{and} \quad \boldsymbol{v}'\boldsymbol{X} = \boldsymbol{c}'.
\end{aligned}
$$

$$
\begin{aligned}
\text{Var}(u + \boldsymbol{v}'\boldsymbol{y}) &= \text{Var}(\boldsymbol{v}'\boldsymbol{y}) \\
&= \text{Var}(\boldsymbol{v}'\boldsymbol{y} - \boldsymbol{c}'\hat{\boldsymbol{\beta}} + \boldsymbol{c}'\hat{\boldsymbol{\beta}}) \\
&= \text{Var}(\boldsymbol{v}'\boldsymbol{y} - \boldsymbol{a}'\boldsymbol{P}_{\boldsymbol{X}}\boldsymbol{y} + \boldsymbol{c}'\hat{\boldsymbol{\beta}}) \\
&= \text{Var}((\boldsymbol{v}' - \boldsymbol{a}'\boldsymbol{P}_{\boldsymbol{X}})\boldsymbol{y} + \boldsymbol{c}'\hat{\boldsymbol{\beta}}) \\
&= \text{Var}(\boldsymbol{c}'\hat{\boldsymbol{\beta}}) + \text{Var}((\boldsymbol{v}' - \boldsymbol{a}'\boldsymbol{P}_{\boldsymbol{X}})\boldsymbol{y}) \\
&\quad + 2\text{Cov}((\boldsymbol{v}' - \boldsymbol{a}'\boldsymbol{P}_{\boldsymbol{X}})\boldsymbol{y}, \boldsymbol{c}'\hat{\boldsymbol{\beta}}).
\end{aligned}
$$

Now

$$\begin{aligned}
\text{Cov}((\boldsymbol{v}' - \boldsymbol{a}'\boldsymbol{P_X})\boldsymbol{y}, \boldsymbol{c}'\hat{\boldsymbol{\beta}}) &= \text{Cov}((\boldsymbol{v}' - \boldsymbol{a}'\boldsymbol{P_X})\boldsymbol{y}, \boldsymbol{a}'\boldsymbol{P_X}\boldsymbol{y}) \\
&= (\boldsymbol{v}' - \boldsymbol{a}'\boldsymbol{P_X})\text{Var}(\boldsymbol{y})\boldsymbol{P_X}\boldsymbol{a} \\
&= \sigma^2(\boldsymbol{v}' - \boldsymbol{a}'\boldsymbol{P_X})\boldsymbol{P_X}\boldsymbol{a} \\
&= \sigma^2(\boldsymbol{v}' - \boldsymbol{a}'\boldsymbol{P_X})\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{a} \\
&= \sigma^2(\boldsymbol{v}'\boldsymbol{X} - \boldsymbol{a}'\boldsymbol{P_X}\boldsymbol{X})(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{a} \\
&= \sigma^2(\boldsymbol{v}'\boldsymbol{X} - \boldsymbol{a}'\boldsymbol{X})(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{a} \\
&= 0 \because \boldsymbol{v}'\boldsymbol{X} = \boldsymbol{c}' = \boldsymbol{a}'\boldsymbol{X}.
\end{aligned}$$

$\therefore$ we have

$$\text{Var}(u + \boldsymbol{v}'\boldsymbol{y}) = \text{Var}(\boldsymbol{c}'\hat{\boldsymbol{\beta}}) + \text{Var}((\boldsymbol{v}' - \boldsymbol{a}'\boldsymbol{P_X})\boldsymbol{y}).$$

It follows that

$$\text{Var}(\boldsymbol{c}'\hat{\boldsymbol{\beta}}) \leq \text{Var}(u + \boldsymbol{v}'\boldsymbol{y})$$

with equality iff

$$\text{Var}((\boldsymbol{v}' - \boldsymbol{a}'\boldsymbol{P_X})\boldsymbol{y}) = 0;$$

i.e., iff

$$\begin{aligned}\sigma^2(\boldsymbol{v}' - \boldsymbol{a}'\boldsymbol{P}_X)(\boldsymbol{v} - \boldsymbol{P}_X\boldsymbol{a}) &= \sigma^2(\boldsymbol{v} - \boldsymbol{P}_X\boldsymbol{a})'(\boldsymbol{v} - \boldsymbol{P}_X\boldsymbol{a}) \\ &= \sigma^2\|\boldsymbol{v} - \boldsymbol{P}_X\boldsymbol{a}\|^2 \\ &= 0;\end{aligned}$$

i.e., iff

$$\boldsymbol{v} = \boldsymbol{P}_X\boldsymbol{a};$$

i.e., iff

$$u + \boldsymbol{v}'\boldsymbol{y} \text{ is } \boldsymbol{a}'\boldsymbol{P}_X\boldsymbol{y} = \boldsymbol{c}'\hat{\boldsymbol{\beta}}, \text{ the LSE of } \boldsymbol{c}'\boldsymbol{\beta}.$$

□

Result 4.1:

The BLUE of an estimable $\mathbf{c}'\boldsymbol{\beta}$ is uncorrelated with all linear unbiased estimators of zero.

Proof of Result 4.1:

Suppose $u + \boldsymbol{v}'\boldsymbol{y}$ is a linear unbiased estimator of zero. Then

$$
\begin{aligned}
& E(u + \boldsymbol{v}'\boldsymbol{y}) = 0 \quad \forall\, \boldsymbol{\beta} \in \mathbb{R}^p \\
\Longleftrightarrow\ & u + \boldsymbol{v}'\boldsymbol{X}\boldsymbol{\beta} = 0 \quad \forall\, \boldsymbol{\beta} \in \mathbb{R}^p \\
\Longleftrightarrow\ & u = 0 \quad \text{and} \quad \boldsymbol{v}'\boldsymbol{X} = \boldsymbol{0}' \\
\Longleftrightarrow\ & u = 0 \quad \text{and} \quad \boldsymbol{X}'\boldsymbol{v} = \boldsymbol{0}.
\end{aligned}
$$

Thus,

$$\begin{aligned}
\text{Cov}(\boldsymbol{c}'\hat{\boldsymbol{\beta}}, u + \boldsymbol{v}'\boldsymbol{y}) &= \text{Cov}(\boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y}, \boldsymbol{v}'\boldsymbol{y}) \\
&= \boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\text{Var}(\boldsymbol{y})\boldsymbol{v} \\
&= \boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'(\sigma^2\boldsymbol{I})\boldsymbol{v} \\
&= \sigma^2\boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{v} \\
&= \sigma^2\boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{0} \\
&= 0.
\end{aligned}$$

$\square$

Suppose $\boldsymbol{c}_1'\boldsymbol{\beta}, \ldots, \boldsymbol{c}_q'\boldsymbol{\beta}$ are q estimable functions $\ni \boldsymbol{c}_1, \ldots, \boldsymbol{c}_q$ are LI.

Let $\boldsymbol{C} = \begin{bmatrix} \boldsymbol{c}_1' \\ \vdots \\ \boldsymbol{c}_q' \end{bmatrix}$. Note that $rank(\underset{q \times p}{\boldsymbol{C}}) = q$.

$$\boldsymbol{C\beta} = \begin{bmatrix} \boldsymbol{c}_1'\boldsymbol{\beta} \\ \vdots \\ \boldsymbol{c}_q'\boldsymbol{\beta} \end{bmatrix}$$ is a vector of estimable functions.

The vector of BLUEs is the vector of LSEs

$$\begin{bmatrix} \boldsymbol{c}_1'\hat{\boldsymbol{\beta}} \\ \vdots \\ \boldsymbol{c}_q'\hat{\boldsymbol{\beta}} \end{bmatrix} = \boldsymbol{C}\hat{\boldsymbol{\beta}}.$$

Because $\boldsymbol{c}_1'\boldsymbol{\beta}, \ldots, \boldsymbol{c}_q'\boldsymbol{\beta}$ are estimable, $\exists$

$$\boldsymbol{a}_i \ni \boldsymbol{X}'\boldsymbol{a}_i = \boldsymbol{c}_i \quad \forall\, i = 1, \ldots, q.$$

If we let $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1' \\ \vdots \\ \boldsymbol{a}_q' \end{bmatrix}$, then $\boldsymbol{AX} = \boldsymbol{C}$.

Find $\mathrm{Var}(\boldsymbol{C}\hat{\boldsymbol{\beta}})$.

$$
\begin{aligned}
\text{Var}(\boldsymbol{C}\hat{\boldsymbol{\beta}}) &= \text{Var}(\boldsymbol{A}\boldsymbol{X}\hat{\boldsymbol{\beta}}) \\
&= \text{Var}(\boldsymbol{A}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y}) \\
&= \text{Var}(\boldsymbol{A}\boldsymbol{P}_{\boldsymbol{X}}\boldsymbol{y}) \\
&= \sigma^2\boldsymbol{A}\boldsymbol{P}_{\boldsymbol{X}}(\boldsymbol{A}\boldsymbol{P}_{\boldsymbol{X}})' \\
&= \sigma^2\boldsymbol{A}\boldsymbol{P}_{\boldsymbol{X}}\boldsymbol{A}' \\
&= \sigma^2\boldsymbol{A}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{A}' \\
&= \sigma^2\boldsymbol{C}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{C}'.
\end{aligned}
$$

Find $rank(\text{Var}(\boldsymbol{C}\hat{\boldsymbol{\beta}}))$.

$$
\begin{aligned}
q = rank(\boldsymbol{C}) &= rank(\boldsymbol{AX}) \\
&= rank(\boldsymbol{AP_X X}) \leq rank(\boldsymbol{AP_X}) \\
&= rank(\boldsymbol{P_X' A'}) = rank(\boldsymbol{AP_X P_X' A'}) \\
&= rank(\boldsymbol{AP_X P_X A'}) = rank(\boldsymbol{AP_X A'}) \\
&= rank(\boldsymbol{AX}(\boldsymbol{X'X})^{-}\boldsymbol{X'A'}) \\
&= rank(\boldsymbol{C}(\boldsymbol{X'X})^{-}\boldsymbol{C'}) \\
&\leq rank(\boldsymbol{C}) = q. \therefore rank(\sigma^2\boldsymbol{C}(\boldsymbol{X'X})^{-}\boldsymbol{C'}) = q.
\end{aligned}
$$

Note

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

is a $q \times q$ matrix of rank q and is thus nonsingular.

We know that each component of $\boldsymbol{C}\hat{\boldsymbol{\beta}}$ is the BLUE of each corresponding component of $\boldsymbol{C}\boldsymbol{\beta}$; i.e., $\boldsymbol{c}_i'\hat{\boldsymbol{\beta}}$ is the BLUE of $\boldsymbol{c}_i'\boldsymbol{\beta} \quad \forall\, i = 1, \ldots, q$.

Likewise, we can show that $\boldsymbol{C}\hat{\boldsymbol{\beta}}$ is the BLUE of $\boldsymbol{C}\boldsymbol{\beta}$ in the sense that

$$\text{Var}(\boldsymbol{s} + \boldsymbol{T}\boldsymbol{y}) - \text{Var}(\boldsymbol{C}\hat{\boldsymbol{\beta}})$$

is nonnegative definite for all unbiased linear estimators $\boldsymbol{s} + \boldsymbol{T}\boldsymbol{y}$ of $\boldsymbol{C}\boldsymbol{\beta}$.

Nonnegative Definite and Positive Definite Matrices

A symmetric matrix $\underset{n\times n}{A}$ is nonnegative definite (NND) if and only if

$$\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} \geq 0 \;\forall\; \boldsymbol{x} \in \mathbb{R}^n.$$

A symmetric matrix $\underset{n\times n}{A}$ is positive definite (PD) if and only if

$$\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} > 0 \;\forall\; \boldsymbol{x} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}.$$

Proof that $\text{Var}(\boldsymbol{s} + \boldsymbol{Ty}) - \text{Var}(\boldsymbol{C}\hat{\boldsymbol{\beta}})$ is NND:

$\boldsymbol{s} + \boldsymbol{Ty}$ is a linear unbiased estimator of $\boldsymbol{C\beta}$

$$\iff E(\boldsymbol{s} + \boldsymbol{Ty}) = \boldsymbol{C\beta} \quad \forall\ \boldsymbol{\beta} \in \mathbb{R}^p$$

$$\iff \boldsymbol{s} + \boldsymbol{TX\beta} = \boldsymbol{C\beta} \quad \forall\ \boldsymbol{\beta} \in \mathbb{R}^p$$

$$\iff \boldsymbol{s} = 0 \quad \text{and} \quad \boldsymbol{TX} = \boldsymbol{C}.$$

Thus, we may write any linear unbiased estimators of $\boldsymbol{C\beta}$ as $\boldsymbol{Ty}$ where $\boldsymbol{TX} = \boldsymbol{C}$.

Now $\forall\, \boldsymbol{w} \in \mathbb{R}^q$ consider

$$\begin{aligned} \boldsymbol{w}'[\text{Var}(\boldsymbol{Ty}) - \text{Var}(\boldsymbol{C}\hat{\boldsymbol{\beta}})]\boldsymbol{w} &= \boldsymbol{w}'\text{Var}(\boldsymbol{Ty})\boldsymbol{w} - \boldsymbol{w}'\text{Var}(\boldsymbol{C}\hat{\boldsymbol{\beta}})\boldsymbol{w} \\ &= \text{Var}(\boldsymbol{w}'\boldsymbol{Ty}) - \text{Var}(\boldsymbol{w}'\boldsymbol{C}\hat{\boldsymbol{\beta}}) \\ &\geq 0 \quad \text{by Gauss-Markov Theorem because} \ldots \end{aligned}$$

(i) $\boldsymbol{w}'\boldsymbol{C}\boldsymbol{\beta}$ is an estimable function:

$$\begin{aligned} \boldsymbol{w}'\boldsymbol{C} = \boldsymbol{w}'\boldsymbol{A}\boldsymbol{X} &\Rightarrow \boldsymbol{X}'\boldsymbol{A}'\boldsymbol{w} = \boldsymbol{C}'\boldsymbol{w} \\ &\Rightarrow \boldsymbol{C}'\boldsymbol{w} \in \mathcal{C}(\boldsymbol{X}'). \end{aligned}$$

(ii) $\boldsymbol{w}'\boldsymbol{T}\boldsymbol{y}$ is a linear unbiased estimator of $\boldsymbol{w}'\boldsymbol{C}\boldsymbol{\beta}$:

$$E(\boldsymbol{w}'\boldsymbol{T}\boldsymbol{y}) = \boldsymbol{w}'\boldsymbol{T}\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{w}'\boldsymbol{C}\boldsymbol{\beta} \quad \forall\ \boldsymbol{\beta} \in \mathbb{R}^p.$$

(iii) $\boldsymbol{w}'\boldsymbol{C}\hat{\boldsymbol{\beta}}$ is the LSE of $\boldsymbol{w}'\boldsymbol{C}\boldsymbol{\beta}$ and is thus the BLUE of $\boldsymbol{w}'\boldsymbol{C}\boldsymbol{\beta}$.

We have shown

$$\boldsymbol{w}'[\text{Var}(\boldsymbol{T}\boldsymbol{y}) - \text{Var}(\boldsymbol{C}\hat{\boldsymbol{\beta}})]\boldsymbol{w} \geq 0, \quad \forall\, \boldsymbol{w} \in \mathbb{R}^p.$$

Thus,

$$\text{Var}(\boldsymbol{T}\boldsymbol{y}) - \text{Var}(\boldsymbol{C}\hat{\boldsymbol{\beta}})$$

is nonnegative definite (NND). □

Is it true that

$$\text{Var}(\boldsymbol{T}\boldsymbol{y}) - \text{Var}(\boldsymbol{C}\hat{\boldsymbol{\beta}})$$

is positive definite if $\boldsymbol{T}\boldsymbol{y}$ is a linear unbiased estimator of $\boldsymbol{C}\boldsymbol{\beta}$ that is not the BLUE of $\boldsymbol{C}\hat{\boldsymbol{\beta}}$?

No. $\boldsymbol{Ty}$ could have some components equal to $\boldsymbol{C}\hat{\boldsymbol{\beta}}$ and others not.

Then

$$\exists\, \boldsymbol{w} \neq \boldsymbol{0} \ni \boldsymbol{w}'[\mathrm{Var}(\boldsymbol{Ty}) - \mathrm{Var}(\boldsymbol{C}\hat{\boldsymbol{\beta}})]\boldsymbol{w} = 0.$$

For example, suppose first row of $\boldsymbol{T}$ is $\boldsymbol{c}_1'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'$ but the second row is not $\boldsymbol{c}_2'(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'$.

Then $\boldsymbol{Ty} \neq \boldsymbol{C}\hat{\boldsymbol{\beta}}$ but

$$
\begin{aligned}
&\boldsymbol{w}'[\mathrm{Var}(\boldsymbol{T}\boldsymbol{y}) - \mathrm{Var}(\boldsymbol{C}\hat{\boldsymbol{\beta}})]\boldsymbol{w} \\
&= \mathrm{Var}(\boldsymbol{w}'\boldsymbol{T}\boldsymbol{y}) - \mathrm{Var}(\boldsymbol{w}'\boldsymbol{C}\hat{\boldsymbol{\beta}}) \\
&= 0 \quad \text{for} \quad \boldsymbol{w}' = [1, 0, \ldots, 0].
\end{aligned}
$$