# The Orthogonal Projection Matrix onto $\mathcal{C}(\boldsymbol{X})$ and the Normal Equations

# Theorem 2.1:

$\boldsymbol{P_X}$, the orthogonal projection matrix onto $\mathcal{C}(\boldsymbol{X})$, is equal to $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'$.

The matrix $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'$ satisfies

(a) $[\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'][\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'] = [\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}']$

(b) $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y} \in \mathcal{C}(\boldsymbol{X}) \; \forall \; \boldsymbol{y} \in \mathbb{R}^n$

(c) $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{x} = \boldsymbol{x} \; \forall \; \boldsymbol{x} \in \mathcal{C}(\boldsymbol{X})$

(d) $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}' = [\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}']'$

(e) $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'$ is the same $\forall$ GI of $\boldsymbol{X}'\boldsymbol{X}$.

It is useful to establish a few results before proving Theorem 2.1.

Show that if $\underset{n\times n}{\boldsymbol{A}}$ is a symmetric matrix and $\underset{n\times n}{\boldsymbol{G}}$ is a GI of $\underset{n\times n}{\boldsymbol{A}}$, then $\underset{n\times n}{\boldsymbol{G}'}$ is also GI of $\underset{n\times n}{\boldsymbol{A}}$.

# Proof:

$$
\begin{aligned}
AGA = A &\Rightarrow (AGA)' = A' \\
&\Rightarrow A'G'A' = A' \\
&\Rightarrow AG'A = A \quad (\because A = A') \\
&\therefore G' \text{ is a GI of } A.
\end{aligned}
$$

$\square$.

# Result 2.4:

$$X'XA = X'XB \iff XA = XB.$$

One direction is obvious ($\Longleftarrow$).

Can you prove the other ($\Longrightarrow$)?

# Proof ($\Longrightarrow$):

$$
\begin{aligned}
(\boldsymbol{XA} - \boldsymbol{XB})'(\boldsymbol{XA} - \boldsymbol{XB}) &= (\boldsymbol{A}'\boldsymbol{X}' - \boldsymbol{B}'\boldsymbol{X}')(\boldsymbol{XA} - \boldsymbol{XB}) \\
&= \boldsymbol{A}'\boldsymbol{X}'\boldsymbol{XA} - \boldsymbol{A}'\boldsymbol{X}'\boldsymbol{XB} - \boldsymbol{B}'\boldsymbol{X}'\boldsymbol{XA} + \boldsymbol{B}'\boldsymbol{X}'\boldsymbol{XB} \\
&= \boldsymbol{A}'\boldsymbol{X}'\boldsymbol{XA} - \boldsymbol{A}'\boldsymbol{X}'\boldsymbol{XA} - \boldsymbol{B}'\boldsymbol{X}'\boldsymbol{XA} + \boldsymbol{B}'\boldsymbol{X}'\boldsymbol{XA} \\
&= \boldsymbol{0}.
\end{aligned}
$$

Copyright ©2012 Dan Nettleton (Iowa State University)

$\therefore$ by Lemma A1,

$$\boldsymbol{XA} - \boldsymbol{XB} = \boldsymbol{0} \Rightarrow \boldsymbol{XA} = \boldsymbol{XB}.$$

$\square$

# Result 2.5:

Suppose $(\boldsymbol{X}'\boldsymbol{X})^-$ is any GI of $\boldsymbol{X}'\boldsymbol{X}$. Then $(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'$ is a GI of $\boldsymbol{X}$, i.e.,

$$\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{X}.$$

Copyright ©2012 Dan Nettleton (Iowa State University)

# Proof of Result 2.5

Since $(\boldsymbol{X}'\boldsymbol{X})^-$ is a GI of $\boldsymbol{X}'\boldsymbol{X}$,

$$\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{X}'\boldsymbol{X}.$$

By Result 2.4, the result follows.

(Take $\boldsymbol{A} = (\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{X}$ and $\boldsymbol{B} = \boldsymbol{I}$. Then
$\boldsymbol{X}\boldsymbol{A} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^-\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{X} = \boldsymbol{X}\boldsymbol{B}$.) $\square$

# Corollary to Result 2.5:

For $(X'X)^-$ any GI of $X'X$,

$$X'X(X'X)^-X' = X'.$$

# Proof of Corollary:

Suppose $(\boldsymbol{X}'\boldsymbol{X})^-$ is any GI of $\boldsymbol{X}'\boldsymbol{X}$.

$\because \boldsymbol{X}'\boldsymbol{X}$ is symmetric, $[(\boldsymbol{X}'\boldsymbol{X})^-]'$ is also a GI of $\boldsymbol{X}'\boldsymbol{X}$.

Thus Result 2.5 implies that

$$\begin{aligned}
&\boldsymbol{X}[(\boldsymbol{X}'\boldsymbol{X})^{-}]'\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{X} \\
&\Rightarrow \left[\boldsymbol{X}[(\boldsymbol{X}'\boldsymbol{X})^{-}]'\boldsymbol{X}'\boldsymbol{X}\right]' = [\boldsymbol{X}]' \\
&\Rightarrow \boldsymbol{X}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}' = \boldsymbol{X}'.
\end{aligned}$$

$\square$

# Now we can prove Theorem 2.1:

First show $X(X'X)^{-}X'$ is idempotent.

$$\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}' = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'$$ by Result 2.5.

Now show that

$$\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y} \in \mathcal{C}(\boldsymbol{X}) \;\forall\; \boldsymbol{y} \in \mathbb{R}^n.$$

$\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y} = \boldsymbol{X}\boldsymbol{z}$, where $\boldsymbol{z} = (\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y}$.

Thus, $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y} \in \mathcal{C}(\boldsymbol{X})$.

Now show that

$$\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{z} = \boldsymbol{z}\ \forall\ \boldsymbol{z} \in \mathcal{C}(\boldsymbol{X}).$$

If $\boldsymbol{z} \in \mathcal{C}(\boldsymbol{X}),\ \exists\ \boldsymbol{c} \ni \boldsymbol{z} = \boldsymbol{X}\boldsymbol{c}$. $\therefore$

$$\begin{aligned} \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{z} &= \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{c} \\ &= \boldsymbol{X}\boldsymbol{c} \text{ (By Result 2.5)} \\ &= \boldsymbol{z}. \end{aligned}$$

Thus, we have $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{z} = \boldsymbol{z}\ \forall\ \boldsymbol{z} \in \mathcal{C}(\boldsymbol{X})$.

Now we know $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'$ is a projection matrix onto $\mathcal{C}(\boldsymbol{X})$.

To be the unique projection matrix onto $\mathcal{C}(\boldsymbol{X})$, $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'$ must be symmetric.

To show symmetry of $X(X'X)^{-}X'$, it will help to first show that

$$X(X'X)_1^{-}X' = X(X'X)_2^{-}X'$$

for any two GIs of $X'X$ denoted $(X'X)_1^{-}$ and $(X'X)_2^{-}$.

By the Corollary to Result 2.5

$$X' = X'X(X'X)_2^- X'.$$

Therefore,

$$\begin{aligned} X(X'X)_1^- X' &= X(X'X)_1^- X'X(X'X)_2^- X' \\ &= X(X'X)_2^- X' \text{ by Result 2.5.} \end{aligned}$$

$\therefore X(X'X)^- X'$ is the same regardless of which GI for $X'X$ is used.

Now show that $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'$ is symmetric.

$$
\begin{aligned}
[X(X'X)^{-}X']' &= X[(X'X)^{-}]'X' \\
&= X(X'X)^{-}X'
\end{aligned}
$$

$\because$ symmetry of $X'X$ implies that $[(X'X)^{-}]'$ is a GI of $X'X$, and $X(X'X)^{-}X'$ is the same regardless of choice of GI. $\square$

We have shown that $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'$ is the orthogonal projection matrix onto $\mathcal{C}(\boldsymbol{X})$.

We know that it is the only symmetric projection matrix onto $\mathcal{C}(\boldsymbol{X})$ by Result A.16.

# Example:

Find the orthogonal projection matrix onto $\mathcal{C}(\underset{n\times 1}{\mathbf{1}})$.

Also, simplify $\boldsymbol{P_X y}$ in this case.

$\boldsymbol{X}'\boldsymbol{X} = \mathbf{1}'\mathbf{1} = n$. Thus $(\boldsymbol{X}'\boldsymbol{X})^{-} = \frac{1}{n}$.

$$\begin{aligned} \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}' &= \mathbf{1}\left[\frac{1}{n}\right]\mathbf{1}' \\ &= \frac{1}{n}\mathbf{1}\mathbf{1}'. \end{aligned}$$

Thus, when $\boldsymbol{X} = \mathbf{1}$, $\boldsymbol{P_X}$ is an $n \times n$ matrix whose entries are each $\frac{1}{n}$.

We saw previously the special case $\begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$.

$$\begin{aligned} \boldsymbol{P_X y} &= \frac{1}{n}\mathbf{1}\mathbf{1}'\boldsymbol{y} \\ &= \frac{1}{n}\mathbf{1}\sum_{i=1}^{n} y_i = \mathbf{1}\frac{1}{n}\sum_{i=1}^{n} y_i \\ &= \mathbf{1}\bar{y} = \begin{bmatrix} \bar{y} \\ \bar{y} \\ \vdots \\ \bar{y} \end{bmatrix}. \end{aligned}$$

By Corollary A.4 to Result A.16, we know that

$$\boldsymbol{I} - \boldsymbol{P_X} = \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'$$

is the orthogonal projection matrix onto $\mathcal{C}(\boldsymbol{X})^{\perp} = \mathcal{N}(\boldsymbol{X}')$.

This is Result 2.6.

# Fitted Values and Residuals

By Result A.4, any $\boldsymbol{y} \in \mathbb{R}^n$ may be written as

$$\boldsymbol{y} = \boldsymbol{s} + \boldsymbol{t},$$

where $\boldsymbol{s} \in \mathcal{C}(\boldsymbol{X})$ and $\boldsymbol{t} \in \mathcal{C}(\boldsymbol{X})^\perp = \mathcal{N}(\boldsymbol{X}')$.

Furthermore, the vectors $\boldsymbol{s} \in \mathcal{C}(\boldsymbol{X})$ and $\boldsymbol{t} \in \mathcal{C}(\boldsymbol{X})^\perp$ are unique.

# Fitted Values and Residuals

Because

$$\boldsymbol{y} = \boldsymbol{P_X y} + (\boldsymbol{I} - \boldsymbol{P_X})\boldsymbol{y}$$

$$\boldsymbol{P_X y} \in \mathcal{C}(\boldsymbol{X}) \text{ and } (\boldsymbol{I} - \boldsymbol{P_X})\boldsymbol{y} \in \mathcal{C}(\boldsymbol{X})^{\perp},$$

we get $\boldsymbol{s} = \boldsymbol{P_X y}$ and $\boldsymbol{t} = (\boldsymbol{I} - \boldsymbol{P_X})\boldsymbol{y}$.

Moreover,

$$\hat{\boldsymbol{y}} \equiv \boldsymbol{P_X y} = \text{ the vector of fitted values}$$

$$\hat{\boldsymbol{\varepsilon}} \equiv (\boldsymbol{I} - \boldsymbol{P_X})\boldsymbol{y} = \boldsymbol{y} - \hat{\boldsymbol{y}} = \text{ the vector of residuals.}$$

Let $\boldsymbol{P_W}$ and $\boldsymbol{P_X}$ denote the orthogonal projection matrices onto $\mathcal{C}(\boldsymbol{W})$ and $\mathcal{C}(\boldsymbol{X})$, respectively.

Suppose $\mathcal{C}(\boldsymbol{W}) \subseteq \mathcal{C}(\boldsymbol{X})$. Show that

$$\boldsymbol{P_W P_X} = \boldsymbol{P_X P_W} = \boldsymbol{P_W}.$$

# Proof:

$$\mathcal{C}(\boldsymbol{W}) \subseteq \mathcal{C}(\boldsymbol{X}) \Rightarrow \exists\, \boldsymbol{B} \ni \boldsymbol{XB} = \boldsymbol{W}.$$

$$\begin{aligned}
\therefore \boldsymbol{P_X P_W} &= \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{W}(\boldsymbol{W}'\boldsymbol{W})^{-}\boldsymbol{W}' \\
&= \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{XB}(\boldsymbol{W}'\boldsymbol{W})^{-}\boldsymbol{W}' \\
&= \boldsymbol{XB}(\boldsymbol{W}'\boldsymbol{W})^{-}\boldsymbol{W}' \\
&= \boldsymbol{W}(\boldsymbol{W}'\boldsymbol{W})^{-}\boldsymbol{W}' \\
&= \boldsymbol{P_W}.
\end{aligned}$$

Now

$$
\begin{aligned}
P_X P_W = P_W \Rightarrow (P_X P_W)' &= P_W' \\
\Rightarrow P_W' P_X' &= P_W' \\
\Rightarrow P_W P_X &= P_W.
\end{aligned}
$$

□

# Theorem 2.2:

If $\mathcal{C}(\boldsymbol{W}) \subseteq \mathcal{C}(\boldsymbol{X})$, then $\boldsymbol{P_X} - \boldsymbol{P_W}$ is the orthogonal projection matrix onto $\mathcal{C}((\boldsymbol{I} - \boldsymbol{P_W})\boldsymbol{X})$.

# Proof of Theorem 2.2:

$$(\boldsymbol{P}_X - \boldsymbol{P}_W)' = \boldsymbol{P}_X' - \boldsymbol{P}_W' = \boldsymbol{P}_X - \boldsymbol{P}_W.$$

$\therefore \boldsymbol{P}_X - \boldsymbol{P}_W$ is symmetric.

$$\begin{aligned}(\boldsymbol{P}_X - \boldsymbol{P}_W)(\boldsymbol{P}_X - \boldsymbol{P}_W) &= \boldsymbol{P}_X\boldsymbol{P}_X - \boldsymbol{P}_W\boldsymbol{P}_X - \boldsymbol{P}_X\boldsymbol{P}_W + \boldsymbol{P}_W\boldsymbol{P}_W \\ &= \boldsymbol{P}_X - \boldsymbol{P}_W - \boldsymbol{P}_W + \boldsymbol{P}_W \\ &= \boldsymbol{P}_X - \boldsymbol{P}_W.\end{aligned}$$

$\therefore \boldsymbol{P}_X - \boldsymbol{P}_W$ is idempotent.

Is $(\boldsymbol{P_X} - \boldsymbol{P_W})\boldsymbol{y} \in \mathcal{C}((\boldsymbol{I} - \boldsymbol{P_W})\boldsymbol{X}) \;\forall\; \boldsymbol{y}$?

$$\begin{aligned}(\boldsymbol{P_X} - \boldsymbol{P_W})\boldsymbol{y} &= (\boldsymbol{P_X} - \boldsymbol{P_W}\boldsymbol{P_X})\boldsymbol{y} \\ &= (\boldsymbol{I} - \boldsymbol{P_W})\boldsymbol{P_X}\boldsymbol{y} \\ &= (\boldsymbol{I} - \boldsymbol{P_W})\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y} \\ &\in \mathcal{C}((\boldsymbol{I} - \boldsymbol{P_W})\boldsymbol{X}).\end{aligned}$$

Is $(\boldsymbol{P_X} - \boldsymbol{P_W})\boldsymbol{z} = \boldsymbol{z} \quad \boldsymbol{z} \in \mathcal{C}((\boldsymbol{I} - \boldsymbol{P_W})\boldsymbol{X})$?

$$\begin{aligned}
(\boldsymbol{P_X} - \boldsymbol{P_W})(\boldsymbol{I} - \boldsymbol{P_W})\boldsymbol{X} &= \boldsymbol{P_X}(\boldsymbol{I} - \boldsymbol{P_W})\boldsymbol{X} - \boldsymbol{P_W}(\boldsymbol{I} - \boldsymbol{P_W})\boldsymbol{X} \\
&= (\boldsymbol{P_X} - \boldsymbol{P_X}\boldsymbol{P_W})\boldsymbol{X} - (\boldsymbol{P_W} - \boldsymbol{P_W}\boldsymbol{P_W})\boldsymbol{X} \\
&= (\boldsymbol{P_X} - \boldsymbol{P_W})\boldsymbol{X} - (\boldsymbol{P_W} - \boldsymbol{P_W})\boldsymbol{X} \\
&= (\boldsymbol{P_X} - \boldsymbol{P_W}\boldsymbol{P_X})\boldsymbol{X} \\
&= (\boldsymbol{I} - \boldsymbol{P_W})\boldsymbol{P_X}\boldsymbol{X} \\
&= (\boldsymbol{I} - \boldsymbol{P_W})\boldsymbol{X}.
\end{aligned}$$

Now $\boldsymbol{z} \in \mathcal{C}((\boldsymbol{I} - \boldsymbol{P_W})\boldsymbol{X})$.

$\Rightarrow \boldsymbol{z} = (\boldsymbol{I} - \boldsymbol{P_W})\boldsymbol{Xc}$ for some $\boldsymbol{c}$. Therefore,

$$\begin{aligned}(\boldsymbol{P_X} - \boldsymbol{P_W})\boldsymbol{z} &= (\boldsymbol{P_X} - \boldsymbol{P_W})(\boldsymbol{I} - \boldsymbol{P_W})\boldsymbol{Xc} \\ &= (\boldsymbol{I} - \boldsymbol{P_W})\boldsymbol{Xc} \\ &= \boldsymbol{z}.\end{aligned}$$

$\square$

We have previously seen that

$$Q(\boldsymbol{b}) = (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b})'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b})$$
$$= \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b}\|^2 \geq \|\boldsymbol{y} - \boldsymbol{P}_X\boldsymbol{y}\|^2$$

$\forall\ \boldsymbol{b} \in \mathbb{R}^p$ with equality iff $\boldsymbol{X}\boldsymbol{b} = \boldsymbol{P}_X\boldsymbol{y}$.

Now we know that $\boldsymbol{P_X} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'$.

Thus, $\hat{\boldsymbol{\beta}}$ minimizes $Q(\boldsymbol{b})$ iff $\boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y}$.

By Result 2.4, this equation is equivalent to

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

Because $\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}' = \boldsymbol{X}', \boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{X}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y}$ is equivalent to

$$\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{X}'\boldsymbol{y}.$$

This system of linear equations is known as the Normal Equations (NE).

We have established Result 2.3:

$\hat{\beta}$ is a solution to the NE $(\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b} = \boldsymbol{X}'\boldsymbol{y})$ iff $\hat{\beta}$ minimizes $Q(\boldsymbol{b})$.

# Corollary 2.1:

The NE are consistent.

# Proof of Corollary 2.1:

NE are $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b} = \boldsymbol{X}'\boldsymbol{y}$.

If we take $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y}$, then

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

□

By Result A.13, $\hat{\boldsymbol{\beta}}$ is a solution to $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b} = \boldsymbol{X}'\boldsymbol{y}$ iff

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y} + [\boldsymbol{I} - (\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{X}]\boldsymbol{z}$$

for some $\boldsymbol{z} \in \mathbb{R}^p$.

# Corollary 2.3:

$\boldsymbol{X}\hat{\boldsymbol{\beta}}$ is invariant to the choice of a solution $\hat{\boldsymbol{\beta}}$ to the NE, i.e., if $\hat{\boldsymbol{\beta}}_1$ and $\hat{\boldsymbol{\beta}}_2$ are any two solutions to the NE, then $\boldsymbol{X}\hat{\boldsymbol{\beta}}_1 = \boldsymbol{X}\hat{\boldsymbol{\beta}}_2$.

# Proof of Corollary 2.3:

$$\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}}_1 = \boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}}_2 (= \boldsymbol{X}'\boldsymbol{y})$$

$$\Rightarrow \boldsymbol{X}\hat{\boldsymbol{\beta}}_1 = \boldsymbol{X}\hat{\boldsymbol{\beta}}_2 \text{ by Result 2.4.}$$

□

We will finish this set of notes with some other results from Section 2.2 of the text.

Lemma 2.1: $\mathcal{N}(\boldsymbol{X}'\boldsymbol{X}) = \mathcal{N}(\boldsymbol{X})$.

Result 2.2: $\mathcal{C}(\boldsymbol{X}'\boldsymbol{X}) = \mathcal{C}(\boldsymbol{X}')$.

Corollary 2.2: $rank(\boldsymbol{X}'\boldsymbol{X}) = rank(\boldsymbol{X})$.

# Proof of Lemma 2.1:

$$\boldsymbol{Xc} = \boldsymbol{0} \Rightarrow \boldsymbol{X'Xc} = \boldsymbol{0}$$

$$\therefore \mathcal{N}(\boldsymbol{X}) \subseteq \mathcal{N}(\boldsymbol{X'X}).$$

$$\boldsymbol{X'Xc} = \boldsymbol{0} \Rightarrow \boldsymbol{c'X'Xc} = 0$$

$$\Rightarrow \boldsymbol{Xc} = \boldsymbol{0}$$

$$\therefore \mathcal{N}(\boldsymbol{X'X}) \subseteq \mathcal{N}(\boldsymbol{X}).$$

Thus, $\mathcal{N}(\boldsymbol{X'X}) = \mathcal{N}(\boldsymbol{X})$. $\square$

# Proof of Result 2.2:

$$X'Xc = X'(Xc) \Rightarrow \mathcal{C}(X'X) \subseteq \mathcal{C}(X').$$

$$\begin{aligned} X'c = X'P_X c &= X'X(X'X)^- X'c \\ &= X'X[(X'X)^- X'c] \\ &\Rightarrow \mathcal{C}(X') \subseteq \mathcal{C}(X'X). \end{aligned}$$

Thus, $\mathcal{C}(X'X) = \mathcal{C}(X')$. □

# Proof of Corollary 2.2:

$$\mathcal{C}(\boldsymbol{X}'\boldsymbol{X}) = \mathcal{C}(\boldsymbol{X}')$$
$$\Rightarrow dim(\mathcal{C}(\boldsymbol{X}'\boldsymbol{X})) = dim(\mathcal{C}(\boldsymbol{X}'))$$
$$\Rightarrow rank(\boldsymbol{X}'\boldsymbol{X}) = rank(\boldsymbol{X}')$$
$$\Rightarrow rank(\boldsymbol{X}'\boldsymbol{X}) = rank(\boldsymbol{X}).$$

□

(Can facilitate finding $(\boldsymbol{X}'\boldsymbol{X})^-$ needed for $\hat{\beta}$ and $\boldsymbol{P}_{\boldsymbol{X}}$.)