Constraints on the Parameter Vector

Suppose

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

where $E(\boldsymbol{\varepsilon}) = \mathbf{0}$ and $\boldsymbol{\beta} \in \mathbb{R}^p$ satisfies $\boldsymbol{H}'\boldsymbol{\beta} = \boldsymbol{h}$ for some known $\underset{p\times q}{\boldsymbol{H}}$ of rank q and some known $\underset{q\times 1}{\boldsymbol{h}}$.

Given that we know $\boldsymbol{\beta}$ satisfies $\boldsymbol{H}'\boldsymbol{\beta} = \boldsymbol{h}$, what functions $\boldsymbol{c}'\boldsymbol{\beta}$ are estimable, and how do we estimate them?

A linear estimator $d + \boldsymbol{a}'\boldsymbol{y}$ is unbiased for $\boldsymbol{c}'\boldsymbol{\beta}$ in the restricted model iff

$$E(d + \boldsymbol{a}'\boldsymbol{y}) = \boldsymbol{c}'\boldsymbol{\beta} \quad \forall\ \boldsymbol{\beta} \ni \boldsymbol{H}'\boldsymbol{\beta} = \boldsymbol{h}.$$

$\boldsymbol{c}'\boldsymbol{\beta}$ is estimable in the restricted model iff $\exists$ a linear estimator $d + \boldsymbol{a}'\boldsymbol{y} \ni$

$$E(d + \boldsymbol{a}'\boldsymbol{y}) = \boldsymbol{c}'\boldsymbol{\beta} \quad \forall\ \boldsymbol{\beta} \text{ satisfying } \boldsymbol{H}'\boldsymbol{\beta} = \boldsymbol{h},$$

i.e., iff $\exists$ a linear estimator that is unbiased for $\boldsymbol{c}'\boldsymbol{\beta}$ in the restricted model.

Result 3.7:

In the restricted model, $d + \boldsymbol{a}'\boldsymbol{y}$ is unbiased for $\boldsymbol{c}'\boldsymbol{\beta}$ iff

$$\exists\, \boldsymbol{l} \ni \boldsymbol{c} = \boldsymbol{X}'\boldsymbol{a} + \boldsymbol{H}\boldsymbol{l} \quad \text{and} \quad d = \boldsymbol{l}'\boldsymbol{h}.$$

Proof of Result 3.7:

$(\Longleftarrow)$ Suppose

$$\exists\, \boldsymbol{l} \ni \boldsymbol{c} = \boldsymbol{X}'\boldsymbol{a} + \boldsymbol{H}\boldsymbol{l} \quad \text{and} \quad d = \boldsymbol{l}'\boldsymbol{h}.$$

Then

$$\begin{aligned}
E(d + \boldsymbol{a}'\boldsymbol{y}) &= \boldsymbol{l}'\boldsymbol{h} + \boldsymbol{a}'\boldsymbol{X}\boldsymbol{\beta} \\
&= \boldsymbol{l}'\boldsymbol{H}'\boldsymbol{\beta} + \boldsymbol{a}'\boldsymbol{X}\boldsymbol{\beta} \quad \forall\, \boldsymbol{\beta} \ni \boldsymbol{H}'\boldsymbol{\beta} = \boldsymbol{h} \\
&= (\boldsymbol{l}'\boldsymbol{H}' + \boldsymbol{a}'\boldsymbol{X})\boldsymbol{\beta} \quad \forall\, \boldsymbol{\beta} \ni \boldsymbol{H}'\boldsymbol{\beta} = \boldsymbol{h} \\
&= (\boldsymbol{X}'\boldsymbol{a} + \boldsymbol{H}\boldsymbol{l})'\boldsymbol{\beta} \quad \forall\, \boldsymbol{\beta} \ni \boldsymbol{H}'\boldsymbol{\beta} = \boldsymbol{h} \\
&= \boldsymbol{c}'\boldsymbol{\beta} \quad \forall\, \boldsymbol{\beta} \ni \boldsymbol{H}'\boldsymbol{\beta} = \boldsymbol{h}.
\end{aligned}$$

($\Longrightarrow$) First note that

$$\{\boldsymbol{\beta} : \boldsymbol{H}'\boldsymbol{\beta} = \boldsymbol{h}\} = \{(\boldsymbol{H}')^{-}\boldsymbol{h} + (\boldsymbol{I} - (\boldsymbol{H}')^{-}\boldsymbol{H}')\boldsymbol{z} : \boldsymbol{z} \in \mathbb{R}^p\}$$
$$= \{\boldsymbol{b}^* + \boldsymbol{W}\boldsymbol{z} : \boldsymbol{z} \in \mathbb{R}^p\},$$

where $\boldsymbol{b}^* = (\boldsymbol{H}')^{-}\boldsymbol{h}$ is one particular solution to $\boldsymbol{H}'\boldsymbol{\beta} = \boldsymbol{h}$ and $\mathcal{C}(\boldsymbol{W}) = \mathcal{N}(\boldsymbol{H}')$ by Results A.12 and A.15, respectively.

Now suppose

$$E(d + \boldsymbol{a}'\boldsymbol{y}) = d + \boldsymbol{a}'\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{c}'\boldsymbol{\beta} \quad \forall\ \boldsymbol{\beta} \ni \boldsymbol{H}'\boldsymbol{\beta} = \boldsymbol{h}.$$

This is equivalent to

$$d + \boldsymbol{a}'\boldsymbol{X}(\boldsymbol{b}^* + \boldsymbol{W}\boldsymbol{z}) = \boldsymbol{c}'(\boldsymbol{b}^* + \boldsymbol{W}\boldsymbol{z}) \quad \forall\ \boldsymbol{z} \in \mathbb{R}^p$$

$$\iff d + \boldsymbol{a}'\boldsymbol{X}\boldsymbol{b}^* - \boldsymbol{c}'\boldsymbol{b}^* + (\boldsymbol{a}'\boldsymbol{X} - \boldsymbol{c}')\boldsymbol{W}\boldsymbol{z} = 0 \quad \forall\ \boldsymbol{z} \in \mathbb{R}^p$$

$$\iff d + \boldsymbol{a}'\boldsymbol{X}\boldsymbol{b}^* - \boldsymbol{c}'\boldsymbol{b}^* = 0 \quad \text{and} \quad \boldsymbol{W}'(\boldsymbol{X}'\boldsymbol{a} - \boldsymbol{c}) = \boldsymbol{0} \quad \text{by Result A.8.}$$

Now $\boldsymbol{W}'(\boldsymbol{X}'\boldsymbol{a} - \boldsymbol{c}) = \boldsymbol{0}$ implies that

$$\begin{aligned} \boldsymbol{X}'\boldsymbol{a} - \boldsymbol{c} \in \mathcal{N}(\boldsymbol{W}') &= \mathcal{C}(\boldsymbol{W})^{\perp} \\ &= \mathcal{N}(\boldsymbol{H}')^{\perp} \\ &= \mathcal{C}(\boldsymbol{H}). \end{aligned}$$

$\therefore \exists\, \boldsymbol{m} \ni \boldsymbol{H}\boldsymbol{m} = \boldsymbol{X}'\boldsymbol{a} - \boldsymbol{c}$

$\Rightarrow \exists\, \boldsymbol{m} \ni \boldsymbol{c} = \boldsymbol{X}'\boldsymbol{a} - \boldsymbol{H}\boldsymbol{m}$

$\Rightarrow \exists\, \boldsymbol{l} \ni \boldsymbol{c} = \boldsymbol{X}'\boldsymbol{a} + \boldsymbol{H}\boldsymbol{l}. \quad (\boldsymbol{l} = -\boldsymbol{m}.)$

Now

$$\begin{aligned} d + \boldsymbol{a}'\boldsymbol{X}\boldsymbol{b}^* - \boldsymbol{c}'\boldsymbol{b}^* = 0 \Rightarrow d &= \boldsymbol{c}'\boldsymbol{b}^* - \boldsymbol{a}'\boldsymbol{X}\boldsymbol{b}^* \\ &= (\boldsymbol{X}'\boldsymbol{a} + \boldsymbol{H}\boldsymbol{l})'\boldsymbol{b}^* - \boldsymbol{a}'\boldsymbol{X}\boldsymbol{b}^* \\ &= \boldsymbol{l}'\boldsymbol{H}'\boldsymbol{b}^* + \boldsymbol{a}'\boldsymbol{X}\boldsymbol{b}^* - \boldsymbol{a}'\boldsymbol{X}\boldsymbol{b}^* \\ &= \boldsymbol{l}'\boldsymbol{H}'\boldsymbol{b}^* \\ &= \boldsymbol{l}'\boldsymbol{h}. \end{aligned}$$

□

Recall that in the unrestricted case, $\boldsymbol{c}'\boldsymbol{\beta}$ is estimable iff $\boldsymbol{c} \in \mathcal{C}(\boldsymbol{X}')$.

Result 3.7 says that $\boldsymbol{c}'\boldsymbol{\beta}$ is estimable in the restricted case iff $\boldsymbol{c} \in \mathcal{C}([\boldsymbol{X}', \boldsymbol{H}])$.

Thus $\boldsymbol{c}'\boldsymbol{\beta}$ is estimable under unrestricted model $\Rightarrow \boldsymbol{c}'\boldsymbol{\beta}$ estimable under restricted model.

However, the converse doesn't hold.

If $\mathcal{C}(\boldsymbol{X}') \subset \mathcal{C}([\boldsymbol{X}', \boldsymbol{H}]), \exists$ functions $\boldsymbol{c}'\boldsymbol{\beta}$ estimable in restricted case but nonestimable in unrestricted case.

Example:

Consider the one-way ANOVA model

$$E(\boldsymbol{y}_{ij}) = \mu + \tau_i \quad i = 1, \ldots, t \quad \text{and} \quad j = 1, \ldots, n_i.$$

Show that $\boldsymbol{c}'\boldsymbol{\beta}$ is estimable $\forall\ \boldsymbol{c} \in \mathbb{R}^p$ under restriction

$$\tau_1 + \cdots + \tau_t = 0.$$

$$\boldsymbol{H} = \underset{t\times 1}{\begin{bmatrix} 0 \\ \mathbf{1} \end{bmatrix}}, \quad \boldsymbol{h} = [0], \quad \boldsymbol{\beta} = \begin{bmatrix} \mu \\ \tau_1 \\ \vdots \\ \tau_t \end{bmatrix}.$$

Then $\boldsymbol{H}'\boldsymbol{\beta} = \boldsymbol{h}$ is equivalent to $\sum_{i=1}^{t} \tau_i = 0$. We have

$$\mathcal{C}([\boldsymbol{X}', \boldsymbol{H}]) = \mathcal{C}\left(\begin{bmatrix} \mathbf{1}' & 0 \\ \underset{t\times t}{\boldsymbol{I}} & \mathbf{1} \end{bmatrix}\right) = \mathbb{R}^p,$$

where $p = t + 1$.

Thus, $\boldsymbol{c} \in \mathcal{C}([\boldsymbol{X}', \boldsymbol{H}]) \quad \forall\, \boldsymbol{c} \in \mathbb{R}^p$.

How do we know $\begin{bmatrix} \mathbf{1}' & 0 \\ \underset{t\times t}{\boldsymbol{I}} & 1 \end{bmatrix}$ has $\mathbb{R}^p$ as its column space?

$$rank\left(\begin{bmatrix} \mathbf{1}' \\ \underset{t\times t}{\boldsymbol{I}} \end{bmatrix}\right) = t = p - 1.$$

$$rank\left(\begin{bmatrix} 0 \\ \mathbf{1} \end{bmatrix}\right) = 1.$$

$$\mathcal{C}\left(\begin{bmatrix} \mathbf{1}' \\ \underset{t\times t}{\boldsymbol{I}} \end{bmatrix}\right) \cap \mathcal{C}\left(\begin{bmatrix} 0 \\ \mathbf{1} \end{bmatrix}\right) = \{\mathbf{0}\}.$$

Thus, $rank\left(\begin{bmatrix} \mathbf{1}' & 0 \\ \underset{t\times t}{\boldsymbol{I}} & 1 \end{bmatrix}\right) = rank\left(\begin{bmatrix} \mathbf{1}' \\ \underset{t\times t}{\boldsymbol{I}} \end{bmatrix}\right) + rank\left(\begin{bmatrix} 0 \\ \mathbf{1} \end{bmatrix}\right) = p.$

Alternatively, suppose

$$\begin{bmatrix} \mathbf{1}' & 0 \\ \underset{t\times t}{\boldsymbol{I}} & \mathbf{1} \end{bmatrix} \boldsymbol{z} = \mathbf{0}.$$

Then

$$\begin{cases} z_1 + \cdots + z_{p-1} & = 0 \\ z_1 + z_p & = 0 \\ \vdots & \vdots \\ z_{p-1} + z_p & = 0 \end{cases} \Rightarrow \begin{cases} z_1 + \cdots + z_{p-1} & = 0 \\ z_1 = \cdots = z_{p-1} & = -z_p \end{cases}$$

$$\Rightarrow z_1 = \cdots = z_p = 0$$

$$\therefore \begin{bmatrix} \mathbf{1}' & 0 \\ \boldsymbol{I} & \mathbf{1} \end{bmatrix} \quad \text{is of full-column rank.}$$

Now suppose we consider the constraints

$$\tau_1 = \tau_2 = \cdots = \tau_t.$$

What functions $\boldsymbol{c}'\boldsymbol{\beta}$ are estimable in this case?

$\tau_1 = \tau_2 = \cdots = \tau_t$ is equivalent to $\boldsymbol{H}'\boldsymbol{\beta} = \boldsymbol{h}$, where

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{0}' \\ \boldsymbol{1}' \\ \underset{(t-1)\times(t-1)}{-\boldsymbol{I}} \end{bmatrix}, \quad \boldsymbol{h} = \underset{(t-1)\times 1}{\boldsymbol{0}}, \quad \boldsymbol{\beta} = \begin{bmatrix} \mu \\ \tau_1 \\ \vdots \\ \tau_t \end{bmatrix}.$$

$\because \mathcal{C}(\boldsymbol{H}) \subseteq \mathcal{C}(\boldsymbol{X}'), \quad \mathcal{C}(\boldsymbol{X}') = \mathcal{C}([\boldsymbol{X}', \boldsymbol{H}])$.

$\therefore$ same functions estimable with or without restrictions.

The Restricted Normal Equations (RNE) are

$$\begin{bmatrix} \boldsymbol{X}'\boldsymbol{X} & \boldsymbol{H} \\ \boldsymbol{H}' & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{b} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}'\boldsymbol{y} \\ \boldsymbol{h} \end{bmatrix}.$$

Result 3.8:

The RNE are consistent.

Proof: First show

$$\begin{bmatrix} \boldsymbol{X}'\boldsymbol{y} \\ \boldsymbol{h} \end{bmatrix} \in \mathcal{C}\left(\begin{bmatrix} \boldsymbol{X}' & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{H}' \end{bmatrix}\right).$$

The constraint equations are consistent and thus have a solution, say $\boldsymbol{b}^*$, such that $\boldsymbol{H}'\boldsymbol{b}^* = \boldsymbol{h}$. Thus,

$$\begin{bmatrix} \boldsymbol{X}' & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{H}' \end{bmatrix} \begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{b}^* \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}'\boldsymbol{y} \\ \boldsymbol{H}'\boldsymbol{b}^* \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}'\boldsymbol{y} \\ \boldsymbol{h} \end{bmatrix}$$

$$\therefore \begin{bmatrix} \boldsymbol{X}'\boldsymbol{y} \\ \boldsymbol{h} \end{bmatrix} \in \mathcal{C}\left(\begin{bmatrix} \boldsymbol{X}' & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{H}' \end{bmatrix} \right).$$

Now suppose that we can show

$$\mathcal{N}\left(\begin{bmatrix} \boldsymbol{X}'\boldsymbol{X} & \boldsymbol{H} \\ \boldsymbol{H}' & \boldsymbol{0} \end{bmatrix}\right) \subseteq \mathcal{N}\left(\begin{bmatrix} \boldsymbol{X} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{H} \end{bmatrix}\right).$$

Explain why this implies the RNE are consistent.

$\mathcal{N}(\boldsymbol{A}) \subseteq \mathcal{N}(\boldsymbol{B}) \Rightarrow \mathcal{N}(\boldsymbol{B})^{\perp} \subseteq \mathcal{N}(\boldsymbol{A})^{\perp}$ by Result A.6.

Now $\mathcal{N}(\boldsymbol{B})^{\perp} \subseteq \mathcal{N}(\boldsymbol{A})^{\perp} \Rightarrow \mathcal{C}(\boldsymbol{B}') \subseteq \mathcal{C}(\boldsymbol{A}')$ by Result A.5. Thus,

$$\begin{aligned}
\mathcal{N}\left(\begin{bmatrix} \boldsymbol{X}'\boldsymbol{X} & \boldsymbol{H} \\ \boldsymbol{H}' & \boldsymbol{0} \end{bmatrix}\right) \subseteq \mathcal{N}\left(\begin{bmatrix} \boldsymbol{X} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{H} \end{bmatrix}\right) &\Rightarrow \mathcal{C}\left(\begin{bmatrix} \boldsymbol{X}' & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{H}' \end{bmatrix}\right) \subseteq \mathcal{C}\left(\begin{bmatrix} \boldsymbol{X}'\boldsymbol{X} & \boldsymbol{H} \\ \boldsymbol{H}' & \boldsymbol{0} \end{bmatrix}\right) \\
&\Rightarrow \begin{bmatrix} \boldsymbol{X}'\boldsymbol{y} \\ \boldsymbol{h} \end{bmatrix} \in \mathcal{C}\left(\begin{bmatrix} \boldsymbol{X}'\boldsymbol{X} & \boldsymbol{H} \\ \boldsymbol{H}' & \boldsymbol{0} \end{bmatrix}\right) \\
&\Rightarrow \text{RNE consistent.}
\end{aligned}$$

Now show that

$$\mathcal{N}\left(\begin{bmatrix} \boldsymbol{X}'\boldsymbol{X} & \boldsymbol{H} \\ \boldsymbol{H}' & \boldsymbol{0} \end{bmatrix}\right) \subseteq \mathcal{N}\left(\begin{bmatrix} \boldsymbol{X} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{H} \end{bmatrix}\right).$$

Suppose $\begin{bmatrix} \boldsymbol{X}'\boldsymbol{X} & \boldsymbol{H} \\ \boldsymbol{H}' & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix}$. Then

$$\boldsymbol{X}'\boldsymbol{X}\boldsymbol{v}_1 + \boldsymbol{H}\boldsymbol{v}_2 = \boldsymbol{0} \quad (1)$$

$$\text{and} \qquad \boldsymbol{H}'\boldsymbol{v}_1 = \boldsymbol{0} \quad (2)$$

Multiplying (1) on the left by $\boldsymbol{v}_1'$ gives

$$\boldsymbol{v}_1'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{v}_1 + \boldsymbol{v}_1'\boldsymbol{H}\boldsymbol{v}_2 = 0.$$

By (2), $\boldsymbol{v}_1'\boldsymbol{H} = \boldsymbol{0}'$. Thus $\boldsymbol{v}_1'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{v}_1 = 0$.

Now $\boldsymbol{v}_1'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{v}_1 = 0 \Rightarrow \boldsymbol{X}\boldsymbol{v}_1 = \mathbf{0}$.

Thus, (1) becomes

$$\boldsymbol{X}'\mathbf{0} + \boldsymbol{H}\boldsymbol{v}_2 = \mathbf{0} \Rightarrow \boldsymbol{H}\boldsymbol{v}_2 = \mathbf{0}.$$

$\therefore \begin{bmatrix} \boldsymbol{X} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{H} \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}$, and it follows that

$$\mathcal{N}\left(\begin{bmatrix} \boldsymbol{X}'\boldsymbol{X} & \boldsymbol{H} \\ \boldsymbol{H}' & \mathbf{0} \end{bmatrix}\right) \subseteq \mathcal{N}\left(\begin{bmatrix} \boldsymbol{X} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{H} \end{bmatrix}\right).$$

$\square$

Result 3.9:

If $\tilde{\beta}$ is the first p components of a solution to the RNE, then $\tilde{\beta}$ minimizes

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

over $\boldsymbol{b}$ satisfying $\boldsymbol{H}'\boldsymbol{b} = \boldsymbol{h}$.

Proof of Result 3.9:

Suppose $\boldsymbol{b}$ is any vector satisfying $\boldsymbol{H}'\boldsymbol{b} = \boldsymbol{h}$. Then

$$\begin{aligned} Q(\boldsymbol{b}) &= \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b}\|^2 \\ &= \|\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}} + \boldsymbol{X}(\tilde{\boldsymbol{\beta}} - \boldsymbol{b})\|^2 \\ &= \|\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}}\|^2 + \|\boldsymbol{X}(\tilde{\boldsymbol{\beta}} - \boldsymbol{b})\|^2 \\ &\quad + 2(\tilde{\boldsymbol{\beta}} - \boldsymbol{b})'\boldsymbol{X}'(\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}}). \end{aligned}$$

$1/2$ the cross product is

$$(\tilde{\boldsymbol{\beta}} - \boldsymbol{b})'\boldsymbol{X}'(\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}}).$$

Because $\tilde{\boldsymbol{\beta}}$ satisfies RNE, we have

$$\boldsymbol{X}'\boldsymbol{X}\tilde{\boldsymbol{\beta}} + \boldsymbol{H}\boldsymbol{\lambda} = \boldsymbol{X}'\boldsymbol{y}.$$

$$\therefore \boldsymbol{X}'\boldsymbol{y} - \boldsymbol{X}'\boldsymbol{X}\tilde{\boldsymbol{\beta}} = \boldsymbol{H}\boldsymbol{\lambda}.$$

Thus $1/2$ cross product is

$$\begin{aligned}(\tilde{\boldsymbol{\beta}} - \boldsymbol{b})'\boldsymbol{H}\boldsymbol{\lambda} &= \boldsymbol{\lambda}'\boldsymbol{H}'(\tilde{\boldsymbol{\beta}} - \boldsymbol{b}) \\ &= \boldsymbol{\lambda}'(\boldsymbol{H}'\tilde{\boldsymbol{\beta}} - \boldsymbol{H}\boldsymbol{b}) \\ &= \boldsymbol{\lambda}'(\boldsymbol{h} - \boldsymbol{h}) \\ &= \boldsymbol{0}.\end{aligned}$$

Thus, we have

$$
\begin{aligned}
Q(\boldsymbol{b}) &= \|\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}} + \boldsymbol{X}(\tilde{\boldsymbol{\beta}} - \boldsymbol{b})\|^2 \\
&= \|\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}}\|^2 + \|\boldsymbol{X}(\tilde{\boldsymbol{\beta}} - \boldsymbol{b})\|^2.
\end{aligned}
$$

$\therefore Q(\tilde{\boldsymbol{\beta}}) \leq Q(\boldsymbol{b})$ with equality iff $\boldsymbol{X}\tilde{\boldsymbol{\beta}} = \boldsymbol{X}\boldsymbol{b}$. $\square$

Result 3.10:

If $\tilde{\boldsymbol{\beta}}$ satisfies

$$\boldsymbol{H}'\tilde{\boldsymbol{\beta}} = \boldsymbol{h} \quad \text{and} \quad Q(\tilde{\boldsymbol{\beta}}) \leq Q(\boldsymbol{b}) \quad \forall\, \boldsymbol{b} \ni \boldsymbol{H}'\boldsymbol{b} = \boldsymbol{h},$$

then $\tilde{\boldsymbol{\beta}}$ is the first p components of a solution to the RNE.

Proof of Result 3.10:

Let $\boldsymbol{\beta}^*$ denote the first p components of any solution to the RNE.

By the proof of Result 3.9, it follows that $\boldsymbol{X}\boldsymbol{\beta}^* = \boldsymbol{X}\tilde{\boldsymbol{\beta}}$. Thus

$$\boldsymbol{X}'\boldsymbol{y} = \boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}^* + \boldsymbol{H}\boldsymbol{\lambda} = \boldsymbol{X}'\boldsymbol{X}\tilde{\boldsymbol{\beta}} + \boldsymbol{H}\boldsymbol{\lambda}$$

so that $\begin{bmatrix} \tilde{\boldsymbol{\beta}} \\ \boldsymbol{\lambda} \end{bmatrix}$ solves the RNE. $\square$