Chapter 3 Testing

Hakbae Lee

The Department of Statistics and Data Science, Yonsei University

- Parameter-free approach

$$\mathbf{Y} = \mathrm{E}(\mathbf{Y}) + \mathbf{Y} - \mathrm{E}(\mathbf{Y}) = \mu + \boldsymbol{\epsilon}$$

where $\mathrm{E}(\mathbf{Y}) = \mu$ and $\boldsymbol{\epsilon} = \mathbf{Y} - \mathrm{E}(\mathbf{Y})$.

- Parameter approach

$$\mathbf{Y} = \mathrm{E}(\mathbf{Y}) + \mathbf{Y} - \mathrm{E}(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where $\mathrm{E}(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$ and $\boldsymbol{\epsilon} = \mathbf{Y} - \mathbf{X}\boldsymbol{\beta}$.

-

$$\mathrm{E}(\boldsymbol{\epsilon}) \quad = \quad 0,\ \mathrm{Cov}(\boldsymbol{\epsilon}) = \sigma^2 I \quad \text{Ordinary Least Square(OLS)}$$

or

$$\mathrm{E}(\boldsymbol{\epsilon}) \quad = \quad 0,\ \mathrm{Cov}(\boldsymbol{\epsilon}) = \sigma^2 \Sigma \quad \text{Generalized Least Square(GLS)}$$

More About Models

- Consider

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon, \quad \mathrm{E}(\epsilon) = 0, \quad \mathrm{Cov}(\epsilon) = \sigma^2 I$$

- $\mathcal{C}(\mathbf{X}) =$ Estimation space
- $M = \mathbf{X}(\mathbf{X}^T\mathbf{X})^-\mathbf{X}^T =$ orthogonal projection onto $\mathcal{C}(\mathbf{X})$
- $\mathrm{E}(\mathbf{Y}) = \mathbf{X}\beta \in \mathcal{C}(\mathbf{X})$, $\mathrm{Cov}(\mathbf{Y}) = \sigma^2 I$
- $\mathcal{C}(\mathbf{X})^\perp =$ Error space
- $\mathbf{I} - M = \mathbf{I} - \mathbf{X}(\mathbf{X}^T\mathbf{X})^-\mathbf{X}^T =$ orthogonal projection onto $\mathcal{C}(\mathbf{X})^\perp$
- $\epsilon \in \mathcal{C}(\mathbf{X})^T$, $\mathrm{Cov}(\epsilon) = \sigma^2 I$
- Any two linear models with the same estimation space are the same model.

More About Models

- *One-Way ANOVA*

$$
\begin{aligned}
y_{ij} &= \mu_i + \epsilon_{ij} = \mu + \alpha_i + \epsilon_{ij}, \text{ with } \mu_i = \mathrm{E}(y_{ij}) = \mu + \alpha_i \\
\bar{\mu} &= \mu + \bar{\alpha}_+ \\
\mu_1 - \mu_2 &= \alpha_1 - \alpha_2
\end{aligned}
$$

  The parameters in the two models are different, but they are related.

- *Simple Linear Regression*

$$
\begin{aligned}
y_i &= \beta_0 + \beta_1 x_i + \epsilon_i \\
&= \gamma_0 + \gamma_1 (x_i - \bar{x}) + \epsilon_i
\end{aligned}
$$

where

$$
\mathrm{E}(y_i) = \beta_0 + \beta_1 x_i = \gamma_0 + \gamma_1 (x_i - \bar{x})
$$

More About Models

- $$\mathcal{C}(\mathbf{X}_1) = \mathcal{C}(\mathbf{X}_1) \Rightarrow \mathbf{X}_1 = \mathbf{X}_2 T$$

  so that

  $$\mathbf{X}_1 \boldsymbol{\beta}_1 = \mathbf{X}_2 T \boldsymbol{\beta}_1 = \mathbf{X}_2 \boldsymbol{\beta}_2$$

  which implies that

  $$\boldsymbol{\beta}_2 = T\boldsymbol{\beta}_1 + \nu \quad \text{for } \nu \in \mathcal{C}(\mathbf{X}_2^T)^\perp$$

- NOTE: A unique parameterization for $\mathbf{X}_j, j = 1, 2$ occurs if and only if $\mathbf{X}_j^T \mathbf{X}_j$ is nonsingular.

- Exercise: Show that a unique parameterization for $\mathbf{X}_j, j = 1, 2$ means $\mathcal{C}(\mathbf{X}_2^T)^\perp = \{0\}$.

Testing Models

- Consider

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon, \quad \epsilon \sim N(0, \mathbf{I}_n)$$

Let's partition $\mathbf{X}$ into $\mathbf{X} = (\mathbf{X}_0, \mathbf{X}_1)$: $\mathcal{C}(\mathbf{X}_0) \subset \mathcal{C}(\mathbf{X})$

$$\begin{aligned} \mathbf{Y} &= \mathbf{X}_0\beta_0 + \mathbf{X}_1\beta_1 + \epsilon && : \text{Full Model(FM)} \\ \mathbf{Y} &= \mathbf{X}_0\gamma + \epsilon && : \text{Reduced Model(RM)} \end{aligned}$$

- Hypothesis testing procedure can be described as

$$H_0 : \text{Reduced Model(RM)} \quad H_1 : \text{Full Model(FM)}$$

- Example 3.2.0: pp. 52–54

Testing Models

- Let M and M_0 be the orthogonal projection onto $\mathcal{C}(\mathbf{X})$ and $\mathcal{C}(\mathbf{X}_0)$ respectively.
- Note that with $\mathcal{C}(\mathbf{X}_0) \subset \mathcal{C}(\mathbf{X})$, $M - M_0$ is the orthogonal projection onto the orthogonal complement of $\mathcal{C}(\mathbf{X}_0)$ with respect to $\mathcal{C}(\mathbf{X})$, that is,

$$\mathcal{C}(X_0)^{\perp}_{\mathcal{C}(\mathbf{X})} = \mathcal{C}(M - M_0) = \mathcal{C}(M \cap M_0^{\perp}) \qquad \text{check!}$$

and

$$\begin{aligned} \hat{\mu} &= \hat{\mathrm{E}}(\mathbf{Y}) = M\mathbf{Y} \qquad \text{under FM} \\ \hat{\mu}_0 &= \hat{\mathrm{E}}(\mathbf{Y}) = M_0\mathbf{Y} \qquad \text{under RM} \end{aligned}$$

- If RM is true, then $M\mathbf{Y} - M_0\mathbf{Y} = (M - M_0)\mathbf{Y}$ should be reasonably small.
- Note that $\mathrm{E}(M - M_0)\mathbf{Y} = 0$

Testing Models

The decision about whether RM is appropriate hinges on deciding whether the vector $(M - M_0)\mathbf{Y}$ is large.

An obvious measure of the size of $(M - M_0)\mathbf{Y}$ is

$$[(M - M_0)\mathbf{Y}]^T[(M - M_0)\mathbf{Y}] = \mathbf{Y}^T(M - M_0)\mathbf{Y}.$$

A reasonable measure of the size of $(M - M_0)\mathbf{Y}$ is given by $\mathbf{Y}^T(M - M_0)\mathbf{Y}/r(M - M_0)$

Note that

$$\mathrm{E}\left[\frac{\mathbf{Y}^T(M - M_0)\mathbf{Y}}{r(M - M_0)}\right] = \sigma^2 + \frac{\beta^T\mathbf{X}^T(M - M_0)\mathbf{X}\beta}{r(M - M_0)}$$

Testing Models

Theorem 3.2.1. Consider $\mathbf{Y} = \mathbf{X}\beta + \epsilon, \quad \epsilon \sim N(0, \mathbf{I}_n)$ with $\mathcal{C}(\mathbf{X}_0) \subset \mathcal{C}(\mathbf{X})$

$$\begin{aligned} \mathbf{Y} &= \mathbf{X}_0\beta_0 + \mathbf{X}_1\beta_1 + \epsilon && : \text{Full Model(FM)} \\ \mathbf{Y} &= \mathbf{X}_0\gamma + \epsilon && : \text{Reduced Model(RM)} \end{aligned}$$

(i) Under the Full Model(FM),

$$\frac{\mathbf{Y}^T(M - M_0)\mathbf{Y}/r(M - M_0)}{\mathbf{Y}^T(\mathbf{I} - M)\mathbf{Y}/r(\mathbf{I} - M)} \sim F\left(df_1, df_2, \frac{\beta^T\mathbf{X}^T(M - M_0)\mathbf{X}\beta}{2\sigma^2}\right)$$

where $df_1 = r(M - M_0)$, $df_2 = r(\mathbf{I} - M)$

(ii) Under the Reduced Model(RM),

$$\frac{\mathbf{Y}^T(M - M_0)\mathbf{Y}/r(M - M_0)}{\mathbf{Y}^T(\mathbf{I} - M)\mathbf{Y}/r(\mathbf{I} - M)} \sim F(df_1, df_2, 0)$$

where $df_1 = r(M - M_0)$, $df_2 = r(\mathbf{I} - M)$

Note:

$$\begin{aligned} M - M_0 &= (I - M_0) - (I - M) \\ \mathbf{Y}^T(M - M_0)\mathbf{Y} &= \mathbf{Y}^T(I - M_0)\mathbf{Y} - \mathbf{Y}^T(I - M)\mathbf{Y} \\ &= \text{SSE}_{\text{RM}} - \text{SSE}_{\text{FM}} \end{aligned}$$

Example 3.2.2.; pp. 58–59

A Generalized Test Procedure

Assume that

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon \text{ is correct.}$$

Want to test the adequacy of a model

$$\mathbf{Y} = \mathbf{X}_0\gamma + \mathbf{X}b + \epsilon$$

where $\mathcal{C}(\mathbf{X}_0) \subset \mathcal{C}(\mathbf{X})$ and some known vector $\mathbf{X}b =$ offset

Example 3.2.3.; Multiple Regression

$$\mathbf{Y} = \beta_0 J + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \epsilon$$

To test $H_0 : \beta_2 = \beta_3 + 5, \beta_1 = 0$,....

A Generalized Test Procedure

$$
\begin{aligned}
\mathbf{Y} &= \mathbf{X}\beta + \epsilon \quad \text{(FM)} \\
\mathbf{Y}^* &\equiv \mathbf{Y} - \mathbf{X}b = \mathbf{X}\beta - \mathbf{X}b + \epsilon = \mathbf{X}\beta^* + \epsilon \quad \text{(FM)}
\end{aligned}
$$

where $\beta^* = \beta - b$ and let's consider

$$
\begin{aligned}
\mathbf{Y} &= \mathbf{X}_0\gamma + \mathbf{X}b + \epsilon \quad \text{(RM)} \\
\mathbf{Y}^* &= \mathbf{Y} - \mathbf{X}b = \mathbf{X}\gamma + \epsilon \quad \text{(RM)}
\end{aligned}
$$

In addition,

$$
\frac{\mathbf{Y}^{*T}(M - M_0)\mathbf{Y}^*/r(M - M_0)}{\mathbf{Y}^{*T}(I - M)\mathbf{Y}^*/r(I - M)} \sim F(r(M - M_0), r(I - M), \delta^2)
$$

where the noncentrality parameter δ^2 is

$$
\delta^2 = \frac{1}{2\sigma^2}\beta^{*T}\mathbf{X}^T(M - M_0)\mathbf{X}\beta^*
$$

$$0 = \beta^{*T}\mathbf{X}^T(M - M_0)\mathbf{X}\beta^*$$

if and only if

$$0 = (M - M_0)\mathbf{X}\beta^* \qquad \text{why?}$$

if and only if

$$\mathbf{X}\beta = M_0(\mathbf{X}\beta - \mathbf{X}b) + \mathbf{X}b \qquad \text{why?}$$

which holds if

$$\gamma = (\mathbf{X}_0^T\mathbf{X}_0)^-\mathbf{X}_0(\mathbf{X}\beta - \mathbf{X}b) = (\mathbf{X}_0^T\mathbf{X}_0)^-\mathbf{X}_0\mathbf{X}\beta^*$$

Furthermore,

$$\mathbf{Y}^{*T}(M - M_0)\mathbf{Y}^* = \mathbf{Y}^{*T}(I - M_0)\mathbf{Y}^* - \mathbf{Y}^{*T}(I - M)\mathbf{Y}^*$$

and

$$\mathbf{Y}^{*T}(I - M)\mathbf{Y}^* = \mathbf{Y}^T(I - M)\mathbf{Y}$$

Example 3.2.3.

Testing Linear Parametric Functions

$$H_0 : \mathbf{Y} = \mathbf{X}\beta + \epsilon \quad \text{and} \quad \Lambda^T \beta = 0 \qquad (1)$$

$$\begin{aligned}
\Lambda^T \beta = 0 &\iff \beta \in \mathcal{N}(\Lambda^T) = \mathcal{C}(X)^\perp \\
&\iff \beta \perp \mathcal{C}(\Lambda) \\
&\iff \beta \perp \mathcal{C}(\Gamma) \quad \text{if } \exists\, \Gamma \text{ so that } \mathcal{C}(\Gamma) = \mathcal{C}(\Lambda) \\
&\iff \beta \perp \mathcal{C}(U) \quad \text{if } \exists\, U \text{ so that } \mathcal{C}(U) = \mathcal{C}(\Lambda)^\perp \\
&\iff \beta = U\gamma \quad \text{for some } \gamma \qquad (2)
\end{aligned}$$

Thus, letting $\mathbf{X}_0 = \mathbf{X}U$, (in general, $\mathcal{C}(\mathbf{X}_0) \subset \mathcal{C}(\mathbf{X})$)

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon = \mathbf{X}U\gamma + \epsilon = \mathbf{X}_0\gamma + \epsilon \qquad (3)$$

Suppose $\mathcal{C}(\mathbf{X}_0) = \mathcal{C}(\mathbf{X})$. Then there is nothing to test and $\Lambda^T\beta = 0$ involves only arbitrary side conditions that do not affect the model.

EXAMPLE 3.3.1. pp. 62–64

$$\text{Estimable } \Lambda^T\beta \quad \Longleftrightarrow \quad \Lambda = \mathbf{X}^T P \quad \text{for some } P$$

Remark:

$$\mathcal{C}(MP) \equiv \mathcal{C}(M - M_0) = \mathcal{C}(\mathbf{X} - \mathbf{X}_0) = \mathcal{C}(\mathbf{X}) \cap \mathcal{C}(\mathbf{X}_0)^{\perp} = \mathcal{C}(\mathbf{X}_0)^{\perp}_{\mathcal{C}(\mathbf{X})}$$

Thus, its distribution for testing $H_0 : \Lambda^T\beta = 0$ is given by

$$\frac{\mathbf{Y}^T M_{MP}\mathbf{Y}/r(M_{MP})}{\mathbf{Y}^T(I - M)\mathbf{Y}/r(I - M)} \sim F(r(M_{MP}), r(I - M), \delta^2) \qquad (5)$$

where $\delta^2 = \beta^T\mathbf{X}^T M_{MP}\mathbf{X}\beta$

Proposition 3.3.2

$$\mathcal{C}(M - M_0) = \mathcal{C}(\mathbf{X}_0)^{\perp}_{\mathcal{C}(\mathbf{X})} \equiv \mathcal{C}(\mathbf{X}U)^{\perp}_{\mathcal{C}(\mathbf{X})} = \mathcal{C}(MP)$$

$$
\begin{aligned}
& H_0 : \mathbf{Y} = \mathbf{X}\beta + \epsilon \quad \text{and} \quad \Lambda^T\beta = 0 \\
\iff \quad & H_0 : \mathbf{Y} = \mathbf{X}\beta + \epsilon \quad \text{and} \quad P^T\mathbf{X}\beta = 0 \\
\iff \quad & H_0 : \mathbf{Y} = \mathbf{X}\beta + \epsilon \quad \text{and} \quad P^T M\mathbf{X}\beta = 0 \quad (M\mathbf{X} = \mathbf{X}) \\
\iff \quad & \mathrm{E}(\mathbf{Y}) \in \mathcal{C}(\mathbf{X}) \quad \text{and} \quad \mathrm{E}(\mathbf{Y}) \perp \mathcal{C}(MP) \\
\iff \quad & \mathrm{E}(\mathbf{Y}) \in \mathcal{C}(\mathbf{X}) \cap \mathcal{C}(MP)^\perp
\end{aligned}
$$

and

$$
\begin{aligned}
& \mathcal{C}(\mathbf{X}_0) = \mathcal{C}(\mathbf{X}) \cap \mathcal{C}(MP)^\perp = \mathcal{C}(MP)^\perp_{\mathcal{C}(\mathbf{X})} \\
\Rightarrow \quad & \mathcal{C}(\mathbf{X}_0)^\perp_{\mathcal{C}(\mathbf{X})} = \mathcal{C}(MP) \\
\iff \quad & \mathbf{X}_0 = (I - M_{MP})\mathbf{X}
\end{aligned}
$$

Theorem 3.3.3

$$
\mathcal{C}[(I - M_{MP})\mathbf{X}] = \mathcal{C}(\mathbf{X}) \cap \mathcal{C}(MP)^\perp = \mathcal{C}(\mathbf{X}) \cap \mathcal{C}(P)^\perp
$$

EXAMPLE 3.3.4.: pp.66–67

$$
\begin{aligned}
\Lambda^T\beta \quad & \text{is} \quad \text{estimable, i.e., } \Lambda = \mathbf{X}^T P \\
\mathcal{C}(\Lambda) &= \mathcal{C}(\mathbf{X}^T P) = \mathcal{C}(MP)
\end{aligned}
$$

and $\mathbf{X}\hat{\beta} = M\mathbf{Y}$, and $\Lambda^T\hat{\beta} = P^T\mathbf{X}\hat{\beta} = P^T M\mathbf{Y}$

$$
\begin{aligned}
\mathbf{Y}^T M_{MP}\mathbf{Y} &= \mathbf{Y}^T M(P^T MP)^- MP\mathbf{Y} \\
&= \hat{\beta}^T\Lambda[P^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^-\mathbf{X}^T P]^-\Lambda^T\hat{\beta} \\
&= \hat{\beta}^T\Lambda[\Lambda^T(\mathbf{X}^T\mathbf{X})^-\Lambda]^-\Lambda^T\hat{\beta}
\end{aligned}
$$

Thus,

$$
(5) \quad = \quad \frac{\hat{\beta}^T\Lambda[\Lambda^T(\mathbf{X}^T\mathbf{X})^-\Lambda]^-\Lambda^T\hat{\beta}/r(\Lambda)}{MSE} \sim F(r(MP), r(I-M), \delta^2)
$$

where

$$
\begin{aligned}
\delta^2 &= \frac{\beta^T\Lambda[\Lambda^T(\mathbf{X}^T\mathbf{X})^-\Lambda]^-\Lambda^T\beta}{2\sigma^2} \\
\mathrm{Cov}(\Lambda^T\hat{\beta}) &= \sigma^2\Lambda^T(\mathbf{X}^T\mathbf{X})^-\Lambda
\end{aligned}
$$

For $H_0 : \lambda^T \boldsymbol{\beta} = 0,\ \lambda \in \mathbf{R}^p$,

$$\begin{aligned}\mathbf{Y}^T M_{MP} \mathbf{Y} &= \hat{\boldsymbol{\beta}}^T \lambda[\lambda^T(\mathbf{X}^T\mathbf{X})^- \lambda]^- \lambda^T \hat{\boldsymbol{\beta}} \\ &= \frac{(\lambda^T \hat{\boldsymbol{\beta}})^2}{\lambda^T(\mathbf{X}^T\mathbf{X})^-\lambda}\end{aligned}$$

and, under $H_0 : \lambda^T \boldsymbol{\beta} = 0$

$$F = (5) \quad = \quad \frac{(\lambda^T \hat{\boldsymbol{\beta}})^2}{MSE[\lambda^T(\mathbf{X}^T\mathbf{X})^-\lambda]} \sim F(1, r(I - M))$$

Definition 3.3.5. The condition $\mathrm{E}(\mathbf{Y}) \perp \mathcal{C}(MP)$ is called the constraint by $\Lambda^T \boldsymbol{\beta} = 0$ where $\Lambda = \mathbf{X}^T P$: $\mathcal{C}(MP) =$ the constraint by $\Lambda^T \boldsymbol{\beta} = 0$.

Do Exercise 3.5: Show that a necessary and sufficient condition for $\rho_1^T \mathbf{X}\boldsymbol{\beta} = 0$ and $\rho_2^T \mathbf{X}\boldsymbol{\beta} = 0$ to determine orthogonal constraints on the model is that $\rho_1^T \mathbf{X} \rho_2 = 0$.

Theoretical Complements

- Consider testing $\Lambda^T\beta = 0$ when $\Lambda^T\beta$ is NOT estimable

- Let $\Lambda_0^T\beta$ be estimable part of $\Lambda^T\beta$.

- Λ_0 is chosen so that

$$\mathcal{C}(\Lambda_0) = \mathcal{C}(\Lambda) \cap \mathcal{C}(\mathbf{X}^T)$$

which means that $\Lambda^T\beta = 0$ implies that $\Lambda_0^T\beta = 0$ but $\Lambda_0^T\beta$ is estimable because

$$\mathcal{C}(\Lambda_0) \subset \mathcal{C}(\mathbf{X}^T).$$

Theorem 3.3.6. Let $\mathcal{C}(\Lambda_0) = \mathcal{C}(\Lambda) \cap \mathcal{C}(\mathbf{X}^T)$ and $\mathcal{C}(U_0) = \mathcal{C}(\Lambda_0)^\perp$. Then $\mathcal{C}(\mathbf{X}U) = \mathcal{C}(\mathbf{X}U_0)$. Thus $\Lambda^T\beta = 0$ and $\Lambda_0^T\beta = 0$ induce the same reduced model.

Proposition 3.3.7. Let $\Lambda_0^T\beta$ be estimable and $\Lambda \neq 0$. Then

$$\Lambda^T\beta = 0 \Longrightarrow \mathcal{C}(\mathbf{X}U) \neq \mathcal{C}(\mathbf{X}).$$

Corollary 3.3.8.

$$\mathcal{C}(\Lambda_0) = \mathcal{C}(\Lambda) \cap \mathcal{C}(\mathbf{X}^T) = \{0\} \Longleftrightarrow \mathcal{C}(\mathbf{X}U) = \mathcal{C}(\mathbf{X}).$$

A Generalized Test Procedure

Consider $H_0 : \Lambda^T\beta = d$, where $d \in \mathcal{C}(\mathbf{X}^T)$, which is solvable. Let b so that $\Lambda^T b = d$. Then

$$\Lambda^T\beta = \Lambda^T b = d \Longleftrightarrow \Lambda^T(\beta - b) = 0$$

$$\begin{aligned}
&\Longleftrightarrow && (\beta - b) \perp \mathcal{C}(\Lambda) \\
&\Longleftrightarrow && (\beta - b) \in \mathcal{C}(U) \quad \text{where} \quad \mathcal{C}(U) = \mathcal{C}(\Lambda)^{\perp} \\
&\Longleftrightarrow && \beta - b = U\gamma \quad \text{for some } \gamma \\
&\Longleftrightarrow && \mathbf{X}\beta - \mathbf{X}b = \mathbf{X}U\gamma \quad \text{i.e. } \mathbf{X}\beta = \mathbf{X}U\gamma + \mathbf{X}b
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{Y} &= \mathbf{X}\beta + \epsilon = \mathbf{X}U\gamma + \mathbf{X}b + \epsilon \\
&= \mathbf{X}_0\gamma + \mathbf{X}b + \epsilon \quad \text{where } \mathbf{X}_0 = \mathbf{X}U \qquad (6)
\end{aligned}$$

A Generalized Test Procedure

If $\Lambda = \mathbf{X}^T P$, then $\mathcal{C}(\mathbf{X}_0)^{\perp}_{\mathcal{C}(\mathbf{X})} = \mathcal{C}(MP)$ and its test statistic is

$$\begin{aligned} F &= \frac{(\mathbf{Y} - \mathbf{X}b)^T M_{MP} (\mathbf{Y} - \mathbf{X}b)/r(M_{MP})}{(\mathbf{Y} - \mathbf{X}b)^T (I - M)(\mathbf{Y} - \mathbf{X}b)/r(I - M)} \\ &= \frac{(\Lambda^T \hat{\beta} - d)^T [\Lambda^T (\mathbf{X}^T \mathbf{X})^- \Lambda]^- (\Lambda^T \hat{\beta} - d)/r(\Lambda)}{MSE} \sim F(?, ?, ?) \end{aligned}$$

Remark: If $\Lambda^T \beta = d$, the same reduced model results if we take $\Lambda^T \beta = d_0$ where $d_0 = d + \Lambda^T \nu$ and $\nu \perp \mathcal{C}(\mathbf{X}^T)$. Note that, in this construction, if $\Lambda^T \beta = d$ is estimable, $d_0 = d$ for any ν.

EXAMPLE 3.3.9.: pp.71–72