Cochran's Theorem and Analysis of Variance

Theorem 5.1:(Cochran's Theorem)

Suppose $\boldsymbol{Y} \sim N(\boldsymbol{\mu}, \sigma^2 \boldsymbol{I})$ and $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_k$ are symmetric and idempotent matrices with

$$rank(\boldsymbol{A}_i) = s_i \quad \forall\, i = 1, \ldots, k.$$

Then $\sum_{i=1}^{k} \boldsymbol{A}_i = \underset{n \times n}{\boldsymbol{I}} \Longrightarrow \frac{1}{\sigma^2} \boldsymbol{Y}' \boldsymbol{A}_i \boldsymbol{Y}$ $(i = 1, \ldots, k)$ are independently distributed as $\chi^2_{s_i}(\phi_i)$, with

$$\phi_i = \frac{1}{2\sigma^2} \boldsymbol{\mu}' \boldsymbol{A}_i \boldsymbol{\mu}, \; \sum_{i=1}^{k} s_i = n.$$

Proof of Theorem 5.1:

By Result 5.15,

$$\frac{1}{\sigma^2}\boldsymbol{Y}'\boldsymbol{A}_i\boldsymbol{Y} \sim \chi^2_{s_i}(\phi_i)$$

with

$$\phi_i = \frac{1}{2\sigma^2}\boldsymbol{\mu}'\boldsymbol{A}_i\boldsymbol{\mu}$$

$\because \frac{1}{\sigma^2}\boldsymbol{A}_i\sigma^2\boldsymbol{I} = \boldsymbol{A}_i$ is idempotent and has rank $s_i \quad \forall\, i = 1, \ldots, k$.

$$
\begin{aligned}
\sum_{i=1}^{k} s_i &= \sum_{i=1}^{k} rank(\boldsymbol{A}_i) \\
&= \sum_{i=1}^{k} trace(\boldsymbol{A}_i) \\
&= trace(\sum_{i=1}^{k} \boldsymbol{A}_i) \\
&= trace(\underset{n\times n}{\boldsymbol{I}}) \\
&= n.
\end{aligned}
$$

By Lemma 5.1, $\exists \underset{n\times s_i}{\boldsymbol{G}_i} \ni$

$$\boldsymbol{G}_i\boldsymbol{G}_i' = \boldsymbol{A}_i,\ \boldsymbol{G}_i'\boldsymbol{G}_i = \underset{s_i\times s_i}{\boldsymbol{I}} \quad \forall\, i = 1,\ldots,k.$$

Now let $\boldsymbol{G} = [\boldsymbol{G}_1,\ldots,\boldsymbol{G}_k]$.

$\because \boldsymbol{G}_i$ is $n \times s_i \quad \forall\, i = 1,\ldots,k$ and $\sum_{i=1}^{k} s_i = n$, it follows that $\boldsymbol{G}$ is $n \times n$.

Moreover,

$$\boldsymbol{G}\boldsymbol{G}' = \begin{bmatrix} \boldsymbol{G}_1 & \dots & \boldsymbol{G}_k \end{bmatrix} \begin{bmatrix} \boldsymbol{G}_1' \\ \vdots \\ \boldsymbol{G}_k' \end{bmatrix}$$

$$= \sum_{i=1}^{k} \boldsymbol{G}_i\boldsymbol{G}_i' = \sum_{i=1}^{k} \boldsymbol{A}_i = \boldsymbol{I}.$$

Thus, $\underset{n\times n}{\boldsymbol{G}}$ has $\boldsymbol{G}'$ as its inverse; i.e., $\boldsymbol{G}^{-1} = \boldsymbol{G}'$. Thus $\boldsymbol{G}'\boldsymbol{G} = \boldsymbol{I}$.

Now we have

$$\boldsymbol{I} = \boldsymbol{G}'\boldsymbol{G} = \begin{bmatrix} \boldsymbol{G}_1' \\ \vdots \\ \boldsymbol{G}_k' \end{bmatrix} \begin{bmatrix} \boldsymbol{G}_1 & \cdots & \boldsymbol{G}_k \end{bmatrix}$$

$$= \begin{bmatrix} \boldsymbol{G}_1'\boldsymbol{G}_1 & \boldsymbol{G}_1'\boldsymbol{G}_2 & \dots & \boldsymbol{G}_1'\boldsymbol{G}_k \\ \boldsymbol{G}_2'\boldsymbol{G}_1 & \boldsymbol{G}_2'\boldsymbol{G}_2 & \dots & \boldsymbol{G}_2'\boldsymbol{G}_k \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{G}_k'\boldsymbol{G}_1 & \boldsymbol{G}_k'\boldsymbol{G}_2 & \dots & \boldsymbol{G}_k'\boldsymbol{G}_k \end{bmatrix}.$$

$\therefore \boldsymbol{G}_i'\boldsymbol{G}_j = \boldsymbol{0} \quad \forall\ i \neq j.$

$\therefore \boldsymbol{G}_i\boldsymbol{G}_i'\boldsymbol{G}_j\boldsymbol{G}_j' = \boldsymbol{0} \quad \forall\ i \neq j$

$\therefore \boldsymbol{A}_i\boldsymbol{A}_j = \boldsymbol{0} \quad \forall\ i \neq j$

$\therefore \sigma^2\boldsymbol{A}_i\boldsymbol{A}_j = \boldsymbol{0} \quad \forall\ i \neq j$

$\therefore \boldsymbol{A}_i(\sigma^2\boldsymbol{I})\boldsymbol{A}_j = \boldsymbol{0} \quad \forall\ i \neq j$

$\therefore$ Independence, hold by Corollary 5.4, $\forall$ pair of quadratic forms $\boldsymbol{Y}'\boldsymbol{A}_i\boldsymbol{Y}/\sigma^2$ and $\boldsymbol{Y}'\boldsymbol{A}_j\boldsymbol{Y}/\sigma^2$.

However, we can prove more than pairwise independence.

$$\begin{bmatrix} \boldsymbol{G}_1'\boldsymbol{Y} \\ \vdots \\ \boldsymbol{G}_k'\boldsymbol{Y} \end{bmatrix} = \begin{bmatrix} \boldsymbol{G}_1' \\ \vdots \\ \boldsymbol{G}_k' \end{bmatrix} \boldsymbol{Y} = \begin{bmatrix} \boldsymbol{G}_1 & \dots & \boldsymbol{G}_k \end{bmatrix}' \boldsymbol{Y} = \boldsymbol{G}'\boldsymbol{Y}$$

$$\sim N(\boldsymbol{G}'\boldsymbol{\mu}, \boldsymbol{G}'(\sigma^2\boldsymbol{I})\boldsymbol{G} = \sigma^2\boldsymbol{G}'\boldsymbol{G} = \sigma^2\boldsymbol{I}).$$

By Result 5.4, $\boldsymbol{G}_1'\boldsymbol{Y}, \dots, \boldsymbol{G}_k'\boldsymbol{Y}$ are mutually independent.

$G_1'Y, \ldots, G_k'Y$ mutually independent,

$$\implies \|G_1'Y\|^2, \ldots, \|G_k'Y\|^2 \text{ mutually independent}$$

$$\implies Y'G_1G_1'Y, \ldots, Y'G_kG_k'Y \text{ mutually independent}$$

$$\implies Y'A_1Y/\sigma^2, \ldots, Y'A_kY/\sigma^2 \text{ mutually independent.}$$

□

Example:

Suppose $Y_1, \ldots, Y_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$.

Find the joint distribution of $n\bar{Y}_{\cdot}^2$ and $\sum_{i=1}^n (Y_i - \bar{Y}_{\cdot})^2$.

Let $\boldsymbol{A}_1 = \boldsymbol{P}_1 = \mathbf{1}(\mathbf{1}'\mathbf{1})^{-}\mathbf{1}' = \frac{1}{n}\mathbf{1}\mathbf{1}'$.

Let $\boldsymbol{A}_2 = \boldsymbol{I} - \boldsymbol{P}_1 = \boldsymbol{I} - \frac{1}{n}\mathbf{1}\mathbf{1}'$.

Then

$$rank(\boldsymbol{A}_1) = 1 \quad \text{and} \quad rank(\boldsymbol{A}_2) = n - 1.$$

Also, $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ are each symmetric and idempotent matrices $\ni$

$$\boldsymbol{A}_1 + \boldsymbol{A}_2 = \boldsymbol{P}_1 + \boldsymbol{I} - \boldsymbol{P}_1 = \boldsymbol{I}.$$

Let $\boldsymbol{Y} = (Y_1, \ldots, Y_n)' \ni E(\boldsymbol{Y}) = \mu\mathbf{1},\ \mathrm{Var}(\boldsymbol{Y}) = \sigma^2\boldsymbol{I}$.

Cochran's Thenorem $\Longrightarrow \frac{1}{\sigma^2}\boldsymbol{Y}'\boldsymbol{A}_i\boldsymbol{Y} \overset{\text{IND}}{\sim} \chi^2_{s_i}(\phi_i)$, where

$$s_i = rank(\boldsymbol{A}_i) \quad \text{and} \quad \phi_i = \frac{1}{2\sigma^2}\boldsymbol{\mu}'\boldsymbol{A}_i\boldsymbol{\mu} = \frac{\mu^2}{2\sigma^2}\mathbf{1}'\boldsymbol{A}_i\mathbf{1}.$$

For $i = 1$, we have

$$\frac{1}{\sigma^2}\boldsymbol{Y}'\mathbf{1}\mathbf{1}'\boldsymbol{Y}/n = \frac{n}{\sigma^2}\bar{Y}^2_{\cdot}, \quad s_1 = 1, \quad \text{and}$$
$$\phi_1 = \frac{\mu^2}{2\sigma^2}\mathbf{1}'\left(\frac{1}{n}\mathbf{1}\mathbf{1}'\right)\mathbf{1} = \frac{n\mu^2}{2\sigma^2}.$$

For $i = 2$, we have

$$
\begin{aligned}
\frac{1}{\sigma^2}\boldsymbol{Y}'\boldsymbol{A}_2\boldsymbol{Y} &= \frac{1}{\sigma^2}\boldsymbol{Y}'\boldsymbol{A}_2'\boldsymbol{A}_2\boldsymbol{Y} \\
&= \frac{1}{\sigma^2}\|\boldsymbol{A}_2\boldsymbol{Y}\|^2 \\
&= \frac{1}{\sigma^2}\left\|\left(\boldsymbol{I} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}'\right)\boldsymbol{Y}\right\|^2 \\
&= \frac{1}{\sigma^2}\|\boldsymbol{Y} - \boldsymbol{1}\bar{Y}_\cdot\|^2 \\
&= \frac{1}{\sigma^2}\sum_{i=1}^{n}(Y_i - \bar{Y}_\cdot)^2.
\end{aligned}
$$

Also, $s_2 = n - 1$ and

$$\begin{aligned}\phi_2 &= \frac{\mu^2}{2\sigma^2}\mathbf{1}'\left(\boldsymbol{I} - \frac{1}{n}\mathbf{1}\mathbf{1}'\right)\mathbf{1} \\ &= \frac{\mu^2}{2\sigma^2}\left(\mathbf{1}'\mathbf{1} - \frac{1}{n}\mathbf{1}'\mathbf{1}\mathbf{1}'\mathbf{1}\right) \\ &= \frac{\mu^2}{2\sigma^2}\left(n - \frac{n^2}{n}\right) = 0.\end{aligned}$$

Thus,

$$n\bar{Y}_{\cdot}^2 \sim \sigma^2 \chi_1^2 \left(\frac{n\mu^2}{2\sigma^2} \right)$$

independent of

$$\sum_{i=1}^{n} (Y_i - \bar{Y}_{\cdot})^2 \sim \sigma^2 \chi_{n-1}^2 .$$

It follows that

$$\frac{n\bar{Y}_{\cdot}^2/\sigma^2}{\frac{\sum_{i=1}^n (Y_i - \bar{Y}_{\cdot})^2}{n-1}/\sigma^2} = \frac{n\bar{Y}_{\cdot}^2}{\frac{\sum_{i=1}^n (Y_i - \bar{Y}_{\cdot})^2}{n-1}}$$

$$\sim F_{1,n-1}\left(\frac{n\mu^2}{2\sigma^2}\right).$$

If $\mu = 0$, $\frac{n\bar{Y}_{\cdot}^2}{\frac{\sum_{i=1}^n (Y_i - \bar{Y}_{\cdot})^2}{n-1}} \sim F_{1,n-1}$.

Thus, we can test $H_0 : \mu = 0$ by comparing $\frac{n\bar{Y}_{\cdot}^2}{\frac{\sum_{i=1}^n (Y_i - \bar{Y}_{\cdot})^2}{n-1}}$ to $F_{1,n-1}$ distribution and rejecting H_0 for large values.

Note

$$\frac{n\bar{Y}_{\cdot}^2}{\frac{\sum_{i=1}^{n}(Y_i-\bar{Y}_{\cdot})^2}{n-1}} = t^2,$$

where

$$t = \frac{\bar{Y}_{\cdot}}{\sqrt{s^2/n}},$$

the usual t-statistics for testing $H_0 : \mu = 0$.

Example:

ANalysis Of VAriance (ANOVA):

Consider $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, where

$$\boldsymbol{X} = [\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_m], \ \boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_m \end{bmatrix}, \ \boldsymbol{\varepsilon} \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{I}).$$

Let $\boldsymbol{P}_j = \boldsymbol{P}_{[\boldsymbol{X}_1,\ldots,\boldsymbol{X}_j]}$ for $j = 1, \ldots, m$.

Let

$$\begin{aligned} \boldsymbol{A}_1 &= \boldsymbol{P}_1 \\ \boldsymbol{A}_2 &= \boldsymbol{P}_2 - \boldsymbol{P}_1 \\ \boldsymbol{A}_3 &= \boldsymbol{P}_3 - \boldsymbol{P}_2 \\ &\vdots \\ \boldsymbol{A}_m &= \boldsymbol{P}_m - \boldsymbol{P}_{m-1} \\ \boldsymbol{A}_{m+1} &= \boldsymbol{I} - \boldsymbol{P}_m. \end{aligned}$$

Note that $\sum_{j=1}^{m+1} \boldsymbol{A}_j = \boldsymbol{I}$.

Then $\boldsymbol{A}_j$ is symmetric and idempotent $\forall\, j = 1, \ldots, m+1$.

$$
\begin{aligned}
s_1 &= rank(\boldsymbol{A}_1) = rank(\boldsymbol{P}_1) \\
s_2 &= rank(\boldsymbol{A}_2) = rank(\boldsymbol{P}_2) - rank(\boldsymbol{P}_1) \\
&\vdots \\
s_m &= rank(\boldsymbol{A}_m) = rank(\boldsymbol{P}_m) - rank(\boldsymbol{P}_{m-1}) \\
s_{m+1} &= rank(\boldsymbol{A}_{m+1}) = rank(\boldsymbol{I}) - rank(\boldsymbol{P}_m) \\
&= n - rank(\boldsymbol{X}).
\end{aligned}
$$

It follows that

$$\frac{1}{\sigma^2}\boldsymbol{Y}'\boldsymbol{A}_j\boldsymbol{Y} \overset{\text{IND}}{\sim} \chi^2_{s_j}(\phi_j), \quad \forall\, j = 1, \ldots, m+1,$$

where

$$\phi_j = \frac{1}{2\sigma^2}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{A}_j\boldsymbol{X}\boldsymbol{\beta}.$$

Note

$$\begin{aligned}\phi_{m+1} &= \frac{1}{2\sigma^2}\boldsymbol{\beta}'\boldsymbol{X}'(\boldsymbol{I} - \boldsymbol{P}_X)\boldsymbol{X}\boldsymbol{\beta} \\ &= \frac{1}{2\sigma^2}\boldsymbol{\beta}'\boldsymbol{X}'(\boldsymbol{X} - \boldsymbol{P}_X\boldsymbol{X})\boldsymbol{\beta} \\ &= \frac{1}{2\sigma^2}\boldsymbol{\beta}'\boldsymbol{X}'(\boldsymbol{X} - \boldsymbol{X})\boldsymbol{\beta} = 0.\end{aligned}$$

Thus,

$$\frac{1}{\sigma^2}\boldsymbol{Y}'\boldsymbol{A}_{m+1}\boldsymbol{Y} \sim \chi^2_{n-rank(\boldsymbol{X})}$$

and

$$\begin{aligned} F_j &= \frac{\boldsymbol{Y}'\boldsymbol{A}_j\boldsymbol{Y}/s_j}{\boldsymbol{Y}'\boldsymbol{A}_{m+1}\boldsymbol{Y}/(n-rank(\boldsymbol{X}))} \\ &\sim F_{s_j,n-rank(\boldsymbol{X})}\left(\frac{1}{2\sigma^2}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{A}_j\boldsymbol{X}\boldsymbol{\beta}\right) \quad \forall\, j = 1,\ldots,m. \end{aligned}$$

We can assemble the ANOVA table as below:

Source	Sum of Squares	DF	Mean Square	Expected Mean Square	F
$\boldsymbol{A}_1$	$\boldsymbol{Y}'\boldsymbol{A}_1\boldsymbol{Y}$	s_1	$\boldsymbol{Y}'\boldsymbol{A}_1\boldsymbol{Y}/s_1$	$\sigma^2 + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{A}_1\boldsymbol{X}\boldsymbol{\beta}/s_1$	F_1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\boldsymbol{A}_m$	$\boldsymbol{Y}'\boldsymbol{A}_m\boldsymbol{Y}$	s_m	$\boldsymbol{Y}'\boldsymbol{A}_m\boldsymbol{Y}/s_m$	$\sigma^2 + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{A}_m\boldsymbol{X}\boldsymbol{\beta}/s_m$	F_m
$\boldsymbol{A}_{m+1}$	$\boldsymbol{Y}'\boldsymbol{A}_{m+1}\boldsymbol{Y}$	s_{m+1}	$\boldsymbol{Y}'\boldsymbol{A}_{m+1}\boldsymbol{Y}/s_{m+1}$	σ^2	
$\boldsymbol{I}$	$\boldsymbol{Y}'\boldsymbol{I}\boldsymbol{Y}$	n			

This ANOVA table contains sequential (a.k.a type I) sum of squares.

$\boldsymbol{A}_{m+1}$ corresponds to “error.”

We can use F_j to test

$$
\begin{aligned}
H_{0j}: &\quad \frac{1}{2\sigma^2}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{A}_j\boldsymbol{X}\boldsymbol{\beta} = 0 \iff \\
H_{0j}: &\quad \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{A}_j'\boldsymbol{A}_j\boldsymbol{X}\boldsymbol{\beta} = 0 \iff \\
H_{0j}: &\quad \boldsymbol{A}_j\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{0}.
\end{aligned}
$$

Now

$$
\begin{aligned}
\boldsymbol{A}_j \boldsymbol{X} \boldsymbol{\beta} &= \boldsymbol{A}_j [\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m] \begin{bmatrix} \boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_m \end{bmatrix} \\
&= \boldsymbol{A}_j \sum_{i=1}^{m} \boldsymbol{X}_i \boldsymbol{\beta}_i \\
&= \sum_{i=1}^{m} \boldsymbol{A}_j \boldsymbol{X}_i \boldsymbol{\beta}_i \\
&= \sum_{i=1}^{m} (\boldsymbol{P}_j - \boldsymbol{P}_{j-1}) \boldsymbol{X}_i \boldsymbol{\beta}_i \quad \forall\, j = 1, \ldots, m \text{ (where } \boldsymbol{P}_0 = \boldsymbol{0} \text{).}
\end{aligned}
$$

Recall $\boldsymbol{P}_j = \boldsymbol{P}_{[X_1,\ldots,X_j]}$.

Thus, $\boldsymbol{P}_j\boldsymbol{X}_i = \boldsymbol{X}_i \quad \forall\ i \leq j$.

It follows that

$$(\boldsymbol{P}_j - \boldsymbol{P}_{j-1})\boldsymbol{X}_i = \boldsymbol{0} \text{ whenever } i \leq j-1.$$

Therefore

$$\begin{aligned}\boldsymbol{A}_j\boldsymbol{X}\boldsymbol{\beta} &= \sum_{i=1}^{m}(\boldsymbol{P}_j - \boldsymbol{P}_{j-1})\boldsymbol{X}_i\boldsymbol{\beta}_i \\ &= \sum_{i=j}^{m}(\boldsymbol{P}_j - \boldsymbol{P}_{j-1})\boldsymbol{X}_i\boldsymbol{\beta}_i.\end{aligned}$$

For $j = m$, this simplifies to

$$\begin{aligned} \boldsymbol{A}_m \boldsymbol{X} \boldsymbol{\beta} &= (\boldsymbol{P}_m - \boldsymbol{P}_{m-1}) \boldsymbol{X}_m \boldsymbol{\beta}_m \\ &= (\boldsymbol{I} - \boldsymbol{P}_{m-1}) \boldsymbol{P}_m \boldsymbol{X}_m \boldsymbol{\beta}_m \\ &= (\boldsymbol{I} - \boldsymbol{P}_{m-1}) \boldsymbol{X}_m \boldsymbol{\beta}_m. \end{aligned}$$

Now

$$\begin{aligned} & (\boldsymbol{I} - \boldsymbol{P}_{m-1}) \boldsymbol{X}_m \boldsymbol{\beta}_m = \boldsymbol{0} \\ \iff & \boldsymbol{X}_m \boldsymbol{\beta}_m \in \mathcal{N}(\boldsymbol{I} - \boldsymbol{P}_{m-1}) \\ \iff & \boldsymbol{X}_m \boldsymbol{\beta}_m \in \mathcal{C}(\boldsymbol{P}_{m-1}) \\ \iff & \boldsymbol{X}_m \boldsymbol{\beta}_m \in \mathcal{C}([\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{m-1}]). \end{aligned}$$

Now

$$
\begin{aligned}
\boldsymbol{X}_m\boldsymbol{\beta}_m &\in \mathcal{C}([\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{m-1}]) \\
\iff E(\boldsymbol{Y}) &= \boldsymbol{X}\boldsymbol{\beta} \\
&= \sum_{i=1}^{m} \boldsymbol{X}_i\boldsymbol{\beta}_i \\
&\in \mathcal{C}([\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{m-1}])
\end{aligned}
$$

$\iff$ Explanatory variables in $\boldsymbol{X}_m$ are irrelevant in the presence of $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{m-1}$.

For the special case where $\boldsymbol{X}$ is of full column rank,

$$
\begin{aligned}
& \boldsymbol{X}_m\boldsymbol{\beta}_m \in \mathcal{C}([\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{m-1}]) \\
\iff & \boldsymbol{X}_m\boldsymbol{\beta}_m = \boldsymbol{0} \\
\iff & \boldsymbol{\beta}_m = \boldsymbol{0}.
\end{aligned}
$$

Thus, in this full column rank case, F_m tests

$$H_{0m} : \boldsymbol{\beta}_m = \boldsymbol{0}.$$

Explain what F_j tests for the special case where

$$\boldsymbol{X}_k'\boldsymbol{X}_{k^*} = \boldsymbol{0} \quad \forall\, k \neq k^*.$$

Now suppose

$$\boldsymbol{X}_k'\boldsymbol{X}_{k^*} = \boldsymbol{0} \quad \forall\, k \neq k^*.$$

Then

$$\boldsymbol{P}_j\boldsymbol{X}_i = \begin{cases} \boldsymbol{X}_i & \text{for } \; i \leq j \\ \boldsymbol{0} & \text{for } \; i > j \end{cases}$$

$$\because [\boldsymbol{X}_1, \ldots, \boldsymbol{X}_j]'\boldsymbol{X}_i = \begin{bmatrix} \boldsymbol{X}_1'\boldsymbol{X}_i \\ \vdots \\ \boldsymbol{X}_j'\boldsymbol{X}_i \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \vdots \\ \boldsymbol{0} \end{bmatrix} \quad \text{for } i > j.$$

It follows that

$$\boldsymbol{A}_j\boldsymbol{X}\boldsymbol{\beta} = \sum_{i=1}^{m}(\boldsymbol{P}_j - \boldsymbol{P}_{j-1})\boldsymbol{X}_i\boldsymbol{\beta}_i = \boldsymbol{P}_j\boldsymbol{X}_j\boldsymbol{\beta}_j = \boldsymbol{X}_j\boldsymbol{\beta}_j.$$

Thus, for the orthogonal case, F_j can be used to test

$$H_{0j} : \boldsymbol{X}_j \boldsymbol{\beta}_j = \boldsymbol{0} \quad \forall\, j = 1, \ldots, m.$$

If we have $\boldsymbol{X}$ full column rank in addition to the orthogonality condition, F_j tests

$$H_{0j} : \boldsymbol{\beta}_j = \boldsymbol{0} \quad \forall\, j = 1, \ldots, m.$$