ON THE RESTRICTIONS OF THE FORM $Kq = m$

IN THE GENERAL LINEAR MODEL

GLM: $Y = X\beta + \varepsilon$  

$E(\varepsilon) = 0$; $\text{Cov}(\varepsilon) = \sigma^2 V$; $V$ known, positive definite matrix.

WITHOUT LOSS OF GENERALITY, ASSUME $V = I$.

We wish to examine the effect of the restrictions

$$K \begin{pmatrix} \beta \\ q \end{pmatrix} = m$$  

where $m \in \mathbb{R} (K)$.

However, (2) is equivalent to $\beta \in B = \{ \beta : \beta = K^+ m + (I - K^+ K)z \}$, where $z \in V_p$ is an arbitrary vector. Furthermore

$K^+ K$ is an $\perp$ projection matrix onto $\mathbb{R} (K')$

and

$(I - K^+ K)$ is an $\perp$ projection matrix onto $[\mathbb{R} (K')]^\perp$.

Now $\min_{\beta \in V_p} ||Y - X\beta||^2$ subject to (2)

$$\iff \min_{z \in V_p} ||Y - X(K^+ m + (I - K^+ K)z)||^2.$$  

$$\iff \min_{z \in V_p} ||Y^* - Xz||^2,$$  

where $Y^* = Y - XK^+ m$ and $X^* = X(I - K^+ K)$.

$$\iff \min_{K\beta = 0} ||Y^* - X\beta||^2.$$  

CASE 1:  

$\mathbb{R} (K') \subset [\mathbb{R} (K')]^\perp$ OR $[\mathbb{R} (K')]^\perp \supset \mathbb{R} (K')$.  

In this case $X^* = X$ and $XK^+ m = 0$. $[m = K\beta]$ for some $\beta = XK^+ K\beta = 0$. Hence,

$$S_w = \min_{\beta \in V_p} ||Y - X\beta||^2 \text{ Sub.to. (2)} = S_\Omega = \min_{\beta \in V_p} ||Y - X\beta||^2.$$  

$$S_{\Omega} = \min_{\beta \in V_p} ||Y - X\beta||^2.$$  

$$S_w = \min_{\beta \in V_p} ||Y - X\beta||^2 \text{ Sub.to. (2)} = S_\Omega = \min_{\beta \in V_p} ||Y - X\beta||^2.$$  

$$S_{\Omega} = \min_{\beta \in V_p} ||Y - X\beta||^2.$$
Therefore Model (1) subject to (2) is same as Model (1) without any restriction.

Furthermore, the NORMAL EQUATIONS

\begin{align}
\begin{pmatrix}
X'X & K' \\
K & 0
\end{pmatrix}
\begin{pmatrix}
\beta \\
\lambda
\end{pmatrix}
&= 
\begin{pmatrix}
X'Y \\
Y
\end{pmatrix}
\end{align}

(6)
(7)

for the "Constrained Minimization" are consistent only if \( \lambda = 0 \).

[The r.h.s. of (6) \( \in \mathbb{R}(X') \). The l.h.s. \( X'X_{\beta} + K'\lambda \leq \in \mathbb{R}(X') \) iff \( \lambda = 0 \).]

Then (6) and (7) reduce to

\begin{align}
X'X_{\beta} &= X'Y \\
K_{\beta} &= m
\end{align}

(8)

For a unique solution to (8), we need rank \( \begin{pmatrix} X'Y \\ K \end{pmatrix} \) = \( p \). Thus if \( q = p - \text{rank}(X) \), = rank(\( K \)), then (8) amounts to finding a particular g-inverse such that

\[ \hat{\beta} = (X'X)^{-1}X'Y \text{ with } K_{\beta} = m. \]

However, in this case the BLUE of \( \hat{\beta} \) does not depend on these conditions, since

\[ \hat{\beta} \] does not depend on the choice of g-inverse.

Conditions (2) which guarantee a unique solution of the unconstrained normal equations are called SIDE-CONDITIONS.

CASE 2. \( R(K') \subseteq \mathbb{R}(X') \), or, \( [R(K')]^\perp \supseteq [R(X')]^\perp \)

(9)

In this case the restrictions are on estimable functions of \( \beta \), i.e. \( K = A \otimes X \).

The solution of (6) and (7), or equivalently the solution of (3), is given by,

\begin{align}
(I-K^+K)^X(I-K^+K)^Z = (I-K^+K)^X Y^* \\
\end{align}

(10)

with \( \hat{\beta} = \hat{\beta}^+m + \hat{\beta}^+(I-K^+K)^Z \), and

\begin{align}
S_{\hat{\beta}} - S_{\Omega} = (K_{\hat{\beta}} - m)\left[K((X'X)^{-1})^{-1}(K_{\hat{\beta}} - m) \right] \\
\end{align}

(11)

where \( \hat{\beta} \) is an LSE of \( \beta \) under \( \Omega \).

CASE 3. \( \mathbb{R}(K') \notin \mathbb{R}(X') \) and \( \mathbb{R}(K') \notin [\mathbb{R}(X')]^\perp \).

In this case the rows of design matrix \( X^* \) in the model

\[ \ldots \]
\[ Y^* = X^*z + \varepsilon \quad (12) \]

are the projection of rows of \( X \) onto \([\mathcal{R}(K')]\) with \([\mathcal{R}(X^*)] = \mathcal{R}[X(K')^0]\), where \( A^0 \) denotes a matrix of maximum rank such that \( \mathcal{A}'A^0 = 0 \). Furthermore,

\[ \text{dim.}[\mathcal{R}(X^*)] = \text{rank}(X' : K') - \text{rank}(K') \quad (13) \]

Infact, \((K')^+ = (I-K'^+K)\) is one choice.

Let \( G \) be a matrix such that \( \mathcal{R}(G') = \mathcal{R}(X') \cap \mathcal{R}(K') \), then it can be shown that \( \text{dim.}[\mathcal{R}(X^*)] = \text{rank}(X') - \text{rank}(G') \).

**Lemma 1.** Let \( \mathcal{R}(G') = \mathcal{R}(X') \cap \mathcal{R}(K') \) so that \( G = AK \) for some \( A \). Then

\[ \min_{G \beta = A \mu} ||Y-X\hat{\beta}||^2 = \min_{K \beta = \mu} ||Y-X\beta||^2 \quad (15) \]

**Proof.** A solution \( K^+ \) of (2) is also a solution of

\[ G \hat{\beta} = A \mu \quad (16) \]

So, the problem on r.h.s. of (15) is equivalent to

\[ \min_{G \beta = 0} ||Y^* - X\beta||^2 \quad (17) \]

For (15) to hold true, we only have to show that the row space of \( K \) and \( G \) are same. However \( \mathcal{R}(G') \subset \mathcal{R}(K') \), implies that the result is valid if

\[ \text{dim.}[\mathcal{R}(X^*)] = \text{dim.}[\mathcal{R}(X(G')^0)] \]

which follows from (13) and (14).

Lemma 1 states that testing the hypothesis (2) is equivalent to testing (16). Hence if (16) is rejected, then (2) is rejected, since (2) \( \Rightarrow \) (16).

**NOTE:** \( \text{rank}(G) \) doesn't have to be \( q \). Since \( \mathcal{R}(G') \subset \mathcal{R}(X') \), it is obvious that

\[ S_w - S_\Omega = (G\hat{\beta} - A\mu)' [G(X'X)^{-1}G']^{-1} (G\hat{\beta} - A\mu) \]

For an arbitrary matrix \( K \), if we estimate \( \hat{\beta} \) by \( K\hat{\beta} \) where \( \hat{\beta} = (X'X)^{-1}X'Y \), then \( K\hat{\beta} \) depends on the choice of g-inverse. Furthermore
\[ E(\hat{\beta}) = \tilde{K}\hat{E}(\hat{\beta}) = \hat{\beta}(X'X)^{-1}X'\hat{\beta} \]

\[ = KH\tilde{\beta} \] where \( H = (X'X)^{-1}X'X \)

is a projection matrix onto \( \tilde{\beta} \) (X'), which is not necessarily orthogonal. In Case (2), \( KH = K \), and we have unbiased estimators of \( \hat{\beta} \). Otherwise \( \hat{\beta} \) is an unbiased estimator of \( (KH)\beta \).

Since \( \hat{\beta} \) depends on the choice of g-inverse, it is clear that

\[ S_W - \sum_{ii}^{X} \neq (\hat{\beta} - \hat{\beta})(X'X)^{-1}K'(K\hat{\beta} - \hat{\beta}) \] \hspace{1cm} (18)

Hence the F-statistic for testing (2) cannot be computed in a usual manner.

However, if we use a g-inverse such that \( H \) is symmetric, then \( KH = A \) where \( \tilde{A} = A + B \); rows of \( \tilde{A} \in \tilde{\beta}(X') \) and rows of \( \tilde{B} \in \tilde{\beta}(X') \). \( A \).

In fact the distribution of the r.h.s. of (18) may not even be chi-square (central or non-central) at all since

\[ \hat{\beta}_0 \sim N[0, \sigma^2(K(X'X)^{-1}X'(X'X)^{-1}K)] \] \hspace{1cm} (19)

For any g-inverse of \( (X'X) \) satisfying \( \tilde{A} = A \), the r.h.s. of (19) is a multiple of (central or non-central) \( \chi^2 \), but for an arbitrary choice of g-inverse, it is not necessarily so.