Chapter 2
SINGLE USE CONFIDENCE REGIONS*

2.1 Introduction

In this chapter we shall construct a single use confidence regions for $\theta$ in the model (1.1.3) and $\eta$ in the model (1.1.5) and also the variance covariance matrix $\Sigma$ is a completely unknown positive definite matrix. Construction of our confidence region is based on a pivotal statistic, and the choice this pivotal quantity is quite natural. However, based on our pivotal statistic, it is not possible to obtain an exact confidence region, because the distribution of the pivot depends on the nuisance parameter $A_1$ (where $A_1$ is an appropriate transformation of $A$). Consequently, the confidence region that we have constructed are only conservative, i.e. its coverage probability is more than assumed confidence level. The computation of the required confidence region is easily done by applying suitable transformations. This is explained in section (2.5). We have used several ideas from Mathew and Zha (1996). We take the same canonical form used by Mathew and Zha (1996) for

* Part of the results included in this chapter was presented in the form of a paper entitled ‘Single Use Conservative Confidence Regions in Multivariate Controlled Calibration’ in the Young Statistician Award Competition in the 24th ISPS Conference held at St. Thomas College, Pala in Nov 2004.
constructing the required confidence regions. This is given in section (2.2). In section (2.4) we have extended our procedure to the model (1.1.5). More details are available in Jose and Jikcey (2004 a), b.

### 2.2 Canonical Forms

We shall first consider the models (1.1.3) and (1.1.4). Let

\[
Y_1 = Y X^T (XX^T)^{-1/2}, \quad S = Y(I - X^T (XX^T)^{-1} X) Y^T, \\
A_1 = A (XX^T)^{1/2}, \quad \theta_1 = (XX^T)^{1/2} \theta
\] (2.2.1)

Then

\[
Y_1 \sim N(A_1, I_q \otimes \Sigma), \quad y \sim N(A_1 \theta_1, \Sigma), \quad S \sim W_p(\Sigma, N-q) \] (2.2.2)

where \( W_p(\Sigma, r) \) denotes p-dimensional Wishart distribution with r degrees of freedom and dispersion matrix \( \Sigma \).

Clearly, \( Y_1 \) in (2.2.1) is the least square estimator of \( A \), suitably scaled. Furthermore, \( S \) in (2.2.2) is the matrix of residuals. In other words, \( Y_1 \) and \( S \) form a set of sufficient statistics for the model (1.1.3). (2.2.2) is the canonical form that we shall work with. Once we obtain a confidence region for \( \theta_1 \), the transformation in (2.2.1) can be used to obtain a confidence region for \( \theta \).

For the models (1.1.1) the canonical form similar to (2.2.2) is
Single Use Confidence Regions

\[ Y_1 \sim \mathcal{N}(A_1, I_q \otimes \Sigma), \quad y \sim \mathcal{N}(A_1 t(\eta), \Sigma), \quad S \sim W_p(\Sigma, (N-q)) , \]  

\[ \text{(2.2.3)} \]

where, \( t_1(\eta) = (X X^\top)^{-\frac{1}{2}} t(\eta) \)  

\[ \text{(2.2.4)} \]

and the other terms in (2.2.3) are defined in (2.2.1).

Note that in (2.2.2) and (2.2.3) \( Y_1, y_1 \) and \( S \) are independently distributed. Also, for the positive definiteness of \( S \), we need the condition \( N - q \geq p \). This assumption will be made in our analysis.

We also note that the statistical inference concerning \( \theta_1 \) in (2.2.2) and \( \eta \) in (2.2.3) is invariant under the non-singular transformation.

\[ Y_1 \rightarrow CY_1, \quad y \rightarrow Cy \quad \text{and} \quad S \rightarrow CS^\top \]  

\[ \text{(2.2.5)} \]

where \( C \) is a \( p \times p \) non-singular matrix. Hence it is natural and desirable to require that the confidence regions that we develop be invariant under the above transformation.

### 2.3 Confidence Regions for the Model (2.2.2)

In the model (2.2.2) \( Y_1 \sim \mathcal{N}(A_1, I_q \otimes \Sigma), \quad y \sim \mathcal{N}(A_1 \theta_1 , \Sigma), \quad S \sim W_p(\Sigma, N-q) \) where we assume that \( N - q \geq p \), and \( Y_1, y \) and \( S \) are independently distributed.

The confidence region that we shall construct will be based on the statistic
Single Use Confidence Regions

\[(Y-Y_i\theta)^\top S^{-1}Y_i (Y_i^\top S^{-1}Y_i)^{-1} Y_i^\top S^{-1}(y-y_i\theta_i)\]

\[= (\hat{\theta}_i - \theta_i)^\top Y_i^\top S^{-1}Y_i(\hat{\theta}_i - \theta_i), \quad (2.3.1)\]

where \(\hat{\theta}_i = (Y_i^\top S^{-1}Y_i)^{-1} Y_i^\top S^{-1}y\) \( (2.3.2)\)

Clearly it is reasonable to use (2.3.2) to obtain a confidence region for \(\theta_i\). The possibility of using (2.3.1) is mentioned in Williams (1959) and Wood (1982). Fujikoshi and Nishii (1984) and Davis and Hayakawa (1987) discussed the asymptotic results based on (2.3.1). These authors have derived the conditional distribution of the statistic in (2.3.1), conditionally given \(Y_1\) and \(y^\top [S^{-1}SY_i(Y_i^\top S^{-1}Y_i)^{-1} Y_i^\top S^{-1}] y\) (see Fujikoshi and Nishii (1984, Theorem 3), and Davis and Hayakawa (1987, page 145)). They have shown that if

\[
K(\theta_i) = 
\left(\frac{N-p-q+1}{q}\right) \left(y - Y_i\theta_i\right)^\top S^{-1}Y_i \left(Y_i^\top S^{-1}Y_i\right)^{-1} Y_i^\top S^{-1}(y - Y_i\theta_i) 
\right)
\left[1 + y^\top [S^{-1} - S^{-1}Y_i(Y_i^\top S^{-1}Y_i)^{-1} Y_i^\top S^{-1}] y \right]^{-1}
\]

Then, conditionally given \(Y_1\) and \(y^\top [S^{-1} - S^{-1} Y_i(Y_i^\top S^{-1} Y_i)^{-1} Y_i S^{-1}] y\), we have

\[
K(\theta_i) \sim F_{q, N-p-q+1}(A_1, \theta_i), \quad (2.3.4)
\]
the non central F distribution with degrees of freedom \((q,N-p-q+1)\) and non centrality parameter \(\Lambda (A_1, \theta_1)\) given by

\[
\Lambda(A_1, \theta_1) = \frac{\theta_1^I (A_1 - Y_1)^I \Sigma^{-1} Y_1 (Y_1^I \Sigma^{-1} Y_1)^I Y_1^I \Sigma^{-1} (A_1 - Y_1) \theta_1}{1 + y^I \left( S^{-1} Y_1 (Y_1^I S^{-1} Y_1)^I Y_1^I S^{-1} \right) y}
\]

(2.3.5)

For completeness, this result is proved below.

**Lemma 2.1:** Let \(K(\theta_1)\) be as given in (2.3.3). Then, conditionally given \(Y_1\) and \(y^I \left[ S^{-1} - S^{-1} Y_1 (Y_1^I S^{-1} Y_1)^I Y_1^I S^{-1} \right] y\),

\[
K(\theta_1) \sim F_{q, N-p-q+1}(\Lambda (A_1, \theta_1))
\]

the non-central F-distribution with degrees of freedom \((q, N - p - q + 1)\) and non-centrality parameter \(\Lambda(A_1, \theta_1)\) given by

\[
\Lambda(A_1, \theta_1) = \frac{\theta_1^I (A_1 - Y_1)^I \Sigma^{-1} Y_1 (Y_1^I \Sigma^{-1} Y_1)^I Y_1^I \Sigma^{-1} (A_1 - Y_1) \theta_1}{1 + y^I \left( S^{-1} Y_1 (Y_1^I S^{-1} Y_1)^I Y_1^I S^{-1} \right) y}
\]

*Proof.* Define

\[
L_1 = (y - Y_1 \theta_1)^I S^{-1} Y_1 (Y_1^I S^{-1} Y_1)^I Y_1^I S^{-1} (y = Y_1 \theta_1), \text{ and}
\]

\[
L_2 = y^I \left[ S^{-1} - S^{-1} Y_1 (Y_1^I S^{-1} Y_1)^I Y_1^I S^{-1} \right] y
\]

(2.3.6)

Let \(T\) be an \(q \times q\) non-singular matrix and \(\Gamma\) be a \(p \times p\) orthogonal matrix such that...
\[ \Sigma^{-1/2} Y_1 = \Gamma \begin{pmatrix} I_q \\ 0 \end{pmatrix} \Gamma' \]  

(2.3.7)

Note that \( \Gamma \) depends on \( Y_1 \). Then

\[ W = \Gamma' \Sigma^{-1/2} S \Sigma^{-1/2} \Gamma = \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix} \sim W_p (I_p, N - q) \]

(2.3.8)

and is distributed independently of \( Y_1 \). In (2.3.8), the dimensions are \( W_{11} : q \times q \), \( W_{12} : q \times (p - q) \), \( W_{21} = W'_{12} : (p - q) \times q \) and \( W_{22} : (p - q) \times (p - q) \). Also given \( Y_1 \),

\[ z = \Gamma' \Sigma^{-1/2} y = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \sim N(\psi, I_p) \]

(2.3.9)

where

\[ \psi = \Gamma' A_1 \theta_1 = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \]

(2.3.10)

\( \psi_1 \) and \( \psi_2 \) being \( q \times 1 \) and \((p - q) \times 1\) vectors. Using the expression for the inverse of a partitioned matrix, we can write

\[ W^{-1} = \begin{pmatrix} W^{-1}_{11,2} & W^{-1}_{12} - W^{-1}_{11,2} W_{12}^{-1} W_{22} \\ -W^{-1}_{22} W_{21} W^{-1}_{11,2} W_{12} & -W^{-1}_{22} + W_{21} W^{-1}_{11,2} W_{12} \end{pmatrix} \]

(2.3.11)

where

\[ W_{11,2} = W_{11} - W_{12} W_{22}^{-1} W_{21} \]. We can thus represent \( L_1 \) in (2.3.6) as
Single Use Confidence Regions

\[ L_1 = (y - Y_1\theta_1)' \Sigma^{-1/2} R W^{-1} \begin{pmatrix} I_m \\ 0 \end{pmatrix} W^{-1} \begin{pmatrix} I_m \\ 0 \end{pmatrix}' \Sigma^{-1/2} (y - Y_1\theta_1) \]

\[ = a' W_{11.2}^{-1} a \quad (2.3.12) \]

where

\[ a = z_1 - T'\theta_1 - Rv \quad (2.3.13) \]

\[ R = W_{12} W_{22}^{-1/2} \text{ and } v = W_{22}^{-1/2} z_2. \] In order to arrive at the second expression in (2.3.12), we have used (2.3.11) along with (2.3.9). Recall that \( R = W_{12} W_{22}^{-1/2} \sim N(0, I_p \otimes I_q) \) and is distributed independently of \( W_{22} \) and \( W_{11.2} \) (see Muirhead (1982, p. 93)). Thus \( R \) is distributed independently of \( v = W_{22}^{-1/2} z_2. \) Hence, using (2.3.9) and (2.3.10), we see that conditionally given \( Y_1 \) and \( v, \)

\[ a \sim N (\psi_1 - T'\theta_1, (1 + v'v) I_q) \quad (2.3.14) \]

and is distributed independently of \( W_{11.2}. \)

Also, \( W_{11.2} \sim W_q (I_q, N - q - p + 1). \) Thus, conditionally given \( Y_1 \) and \( v, \)

\[ \frac{N - q - p + 1}{q} \frac{a' W_{11.2}^{-1} a}{1 + v'v} \quad (2.3.15) \]

has the distribution of a non-central Hotelling’s \( T^2 \) statistic (i.e., non-central F) with degrees of freedom \( (q, N - q - p + 1) \) and non-centrality parameter.
Using (2.3.8) to (2.3.11), we can write $L_2$ in (2.3.6) as

$$L_2 = z' W^{-1} z - (z_1 - W_{12} W_{22}^{-1/2} z_2)' W_{11,2}^{-1} (z_1 - W_{12} W_{22}^{-1/2} z_2)$$

$$= z' W_{22}^{-1/2} z_2 = v'v$$

(2.3.17)

and

$$\Gamma' (A_1 - Y_1) \Sigma^{-1/2} \theta_1 = \psi_1 - T'\theta_1$$

(2.3.18)

Using (2.3.17) and (2.3.18), the non-centrality parameter in (2.3.16) reduces to $\Lambda(A_1, \theta_1)$ in (2.3.5). In view of (2.3.6), (2.3.12) and (2.3.17), the statistic (2.3.15) coincides with $K(\theta_1)$ in (2.3.3). This completes the proof of the lemma (2.1).

Now for constructing our confidence region we shall use the pivot $K(\theta_1)$ in (2.3.3). The distribution of $K(\theta_1)$ depends on both $A_1$ and $\theta_1$ and hence exact confidence region for $\theta_1$ cannot be obtained using $K(\theta_1)$. We shall proceed as follows.

Let $\beta_\alpha (A_1, \theta_1)$ be such that

$$P[K(\theta_1) \leq \beta_\alpha (A_1, \theta_1)] = 1 - \alpha$$

(2.3.19)
ie., \( \{ \theta_1: K(\theta_1) \leq \beta_\alpha (A_1, \theta_1) \} \) is a 100 \((1 - \alpha)\) % confidence region for \( \theta_1 \) of course, this confidence region cannot be computed since \( \beta_\alpha (A_1, \theta_1) \) depends on the nuisance parameter \( A_1 \), we shall obtain two quantities \( U_\alpha (\theta_1) \) and \( V_\alpha (\theta_1) \) which depend on \( \theta_1 \) (but not \( A_1 \)) satisfying

\[
U_\alpha (\theta_1) \leq \beta_\alpha (A_1, \theta_1) \leq V_\alpha (\theta_1) \tag{2.3.20}
\]

for all \( A_1 \), then we have

\[
P[K(\theta_1) \leq U_\alpha (\theta_1)] \leq P[K(\theta_1) \leq \beta_\alpha (A_1, \theta_1)] \leq P[K(\theta_1) \leq V_\alpha (\theta_1)] \tag{2.3.21}
\]

from (2.3.19), (2.3.20) and (2.3.21) we get

\[
P[K(\theta_1) \leq V_\alpha (\theta_1)] \geq 1 - \alpha \tag{2.3.22}
\]

In other words the region

\[
\{ \theta_1: K(\theta_1) \leq V_\alpha (\theta_1) \} \tag{2.3.23}
\]

is a conservative confidence region for \( \theta_1 \), with coverage probability at least \( 1 - \alpha \). It should be noted that if \( V_\alpha (\theta_1) \) depends on \( \theta_1 \), the region in (2.3.23) can be obtained once the functional form of \( V_\alpha (\theta_1) \) is known. The same idea is used in Oman (1988) to obtain a conservative confidence region. Note that the value

\[
\hat{\theta}_1 = (Y_1^\top S^{-1} Y_1)^{-1} Y_1^\top S^{-1} y
\]
always belongs to the confidence region (2.3.23)

Hence the region is non empty. Also \( Y \rightarrow CY \) and \( y \rightarrow Cy \) (or equivalently \( Y \rightarrow CY_1, S \rightarrow CSC^1 \) and \( y \rightarrow Cy \)) where \( C \) is a \( p \times p \) non-singular matrix shows that the region is invariant under the transformation. Since \( K(\theta_1) \) is an invariant statistic, the construction of \( V_\alpha(\theta_1) \) in (2.3.23) is based on the following theorem (see Mathew and Zha (1996)).

This theorem leads us to take a suitable approximation for constructing our confidence region. In order to establish the theorem, we shall consider the following result, which is used in the proof of the theorem.
Lemma 2.2: Consider the model (2.2.2) and let $F_{r,s}(\gamma)$ denote a non-central F random variable with degrees of freedom of $(r, s)$ and non-centrality parameter $\gamma$. Then, conditionally given $Y_1$,

$$\frac{N-p+1}{p-q}y'[S^{-1} - S^{-1}Y_1(Y_1'S^{-1}Y_1)^{-1}Y_1'S^{-1}]y \sim F_{p-q,N-p+1}(\lambda)$$ (2.3.24)

where

$$\lambda = \theta_1'A_1[\Sigma^{-1} - \Sigma^{-1}Y_1(Y_1'S^{-1}Y_1)^{-1}Y_1'S^{-1}]A_1\theta_1$$ (2.3.25)

Proof: The result follows from the representation (2.3.17) for the quantity $L_2$ in (2.3.6). Now we shall establish the main result.

Theorem 2.1

Consider the model (2.2) and let $K(\theta_1)$ be as defined in (2.4). Let $f_{n_1}^{(\cdot)}$ and $g_{n_2,n_3}^{(\cdot)}$ respectively, denote the probability density functions of a central chi-square random variable with $n_1$ degrees of freedom and a central F random variable with $(n_2, n_3)$ degrees of freedom. Also let $F_{n_2,n_3}^{(\cdot)}$ and $F_{n_2,n_3}^{(\gamma)}$ respectively denote a central F random variable and a non central F random variable with non centrality parameter $\gamma$, both having degrees of freedom $(n_2, n_3)$.
Then for any \( a > 0 \),

\[
P[F_{q,N-p-q+1} \leq a] \geq P[K(\theta_1) \leq a]
\]

\[
\geq \int_0^\infty \int_0^\infty P[F_{q,N-p-q+1}(\begin{bmatrix} \theta_1^1 & \theta_1^1 \end{bmatrix} w / \lambda + (p-q)Z/(N-p+1)) \leq a) \]
\]

\[
\times f_p(w)g_{p-q,N-p+1\ (Z)} \ dwdz
\]

\[
\geq \int_0^\infty P[F_{q,N-p-q+1}(\theta_1^1, \theta_1^1, w) \leq a] f_p(w) \ dw
\]

**Proof**

From Fujikoshi and Nishii (1984) and Davis and Hayakawa (1987), it follows that conditionally given \( Y_1 \),

\[
\frac{N-p+1}{p-q} y^1 [S^{-1} - S^{-1}Y_1(Y_1^1 S^{-1}Y_1 Y_1) Y_1^1 S^{-1}]y \sim F_{p-q,N-p+1}(\lambda) \quad (2.3.27)
\]

where \( \lambda = \theta_1^1 A_1^1 \Sigma^{-1}Y_1(Y_1^1 \Sigma^{-1}Y_1) Y_1^1 \Sigma^{-1}A_1 \theta_1 \).

We note that the quantity \( y^1 [S^{-1} - S^{-1}Y_1(Y_1^1 S^{-1}Y_1) Y_1^1 S^{-1}]y \)

which appears on the left hand side of (2.13) is the quantity considered in Fujikoshi and Nishii (1984) and Davis and Hayakawa (1987).

Since \( \Sigma^{-1} \geq \Sigma^{-1} Y_1(Y_1^1 \Sigma^{-1}Y_1) Y_1^1 \Sigma^{-1} \) we have the following inequality concerning the numerator of \( \Lambda(A_1, \theta_1) \) in (2.3.5).
\[
\begin{align*}
\theta_1^1 (A_1 - Y_1) \Sigma^{-1} Y_1 (Y_1^\dagger \Sigma^{-1} Y_1)^{-1} Y_1^\dagger \Sigma^{-1} (A_1 - Y_1) \theta_1 \\
\leq \theta_1^1 (A_1 - Y_1) \Sigma^{-1} (A_1 - Y_1) \theta_1
\end{align*}
\] (2.3.28)

Also, given \( Y_1 \), the non central F random variable

\[
\frac{N - p + 1}{p - q} y^\dagger [S^{-1} - S^{-1} Y_1 (Y_1^\dagger S^{-1} Y_1)^{-1} Y_1^\dagger S^{-1}] y
\]

is stochastically larger than the central F random variable \( F_{p-q, N-p+1} \). This fact, along with (2.3.28) shows that \( \Lambda(A_1, \theta_1) \) in (2.3.5) satisfies (conditionally given \( Y_1 \))

\[
\Lambda(A_1, \theta_1) \leq \text{st} \frac{\theta_1^1 (A_1 - Y_1) \Sigma^{-1} (A_1 - Y_1) \theta_1}{1 + \frac{p - q}{N - p + 1} z}
\] (2.3.29)

where \( Z \) denotes a central F random variable with \( (p-q, N-p+1) \) degrees of freedom and \( \leq \text{st} \) denotes stochastically smaller. The first inequality in (2.3.26) follows from (2.3.4) using the fact that \( P[F_{n_2, n_3} \leq a] \) is a decreasing function of \( \gamma \). Furthermore, their property, along with (2.3.4) and (2.3.26) gives the following inequality, conditionally given \( Y_1 \)

\[
P[K(\theta_1) \leq a] \geq \int_0^P \left[ F_{q, N-p+1-q} \left( \frac{\theta_1^1 (A_1 - Y_1) \Sigma^{-1} (A_1 - Y_1) \theta_1}{1 + \frac{p - q}{N - p + 1} z} \right) \right] \leq a \]

\[\times g_{p-q, N-p+1} \, dz \] (2.3.30)
Using the distribution of $Y_1$ given in (2.3.1), it follows that 

$$(A_1 - Y_1) \theta_1 \sim N(0, (\theta_1^T \theta_1) \Sigma).$$

Hence

$$W = \frac{\theta_1^T (A_1 - Y_1)^T \Sigma^{-1} (A_1 - Y_1) \theta_1}{\theta_1^T \theta_1} \sim \chi^2_p$$

(2.3.31)

Thus (2.3.30) and (2.3.31) together give the second inequality in (2.3.26). The last inequality in (2.3.26) follows by noting that

$$\frac{\theta_1^T \theta_1 w}{1 + \frac{p - q}{N - p + 1} z} \leq \theta_1^T \theta_1 w$$

and $P[F_{n_2, n_3}^{(\gamma)} \leq a]$ is a decreasing function of $\gamma$. This completes the proof of the theorem.

The approach adopted by Mathew and Zha (1996) is to compute a function of $\theta_1^T \theta_1$, say $V_\alpha (\theta_1^T \theta_1)$ satisfying

$$\int_0^\infty P[F_{p, N-p-q+1}(\theta_1^T \theta_1 w) \leq V_\alpha (\theta_1^T \theta_1)] f_p(w) dw = 1 - \alpha$$

(2.3.32)

From theorem 1 it follows that $P(K(\theta_1) \leq V_\alpha (\theta_1^T \theta_1)) \geq 1 - \alpha$. In other words, $\{\theta_1: K(\theta_1) \leq V_\alpha (\theta_1^T \theta_1)\}$ is a conservative confidence region for $\theta_1$. However, the function $V_\alpha (\theta_1^T \theta_1)$ has to be numerically determined and the computation is quite involved. Here we shall develop a simple approximation.
Note that the expressions in Theorem 2.1 involve non-central F distributions with random non-centrality parameters. Our approximation consists of replacing these random non-centrality parameters by their expected values. Thus let
\[ \delta = E \begin{bmatrix}
    W \\
    1 + \frac{p-q}{N-p+1} Z
\end{bmatrix}. \]

Also, since \( W \sim \chi^2_p \), \( E(W) = p \).

Now let \( V_{\alpha}^{(1)}(\theta' \theta_1) \) and \( V_{\alpha}^{(2)}(\theta' \theta_1) \) satisfy
\[
P \left[ F_{q,N-p-q+1} (\theta' \theta_1 \ p) \leq V_{\alpha}^{(1)}(\theta' \theta_1) \right] = 1 - \alpha
\]
\[
P \left[ F_{q,N-p-q+1} (\theta' \theta_1 \ \delta) \leq V_{\alpha}^{(2)}(\theta' \theta_1) \right] = 1 - \alpha \tag{2.3.33}
\]

Hence our approximate confidence regions for \( \theta_1 \) are given by
\[ \{ \theta_1: K(\theta_1) \leq V_{\alpha}^{(i)}(\theta_1') \} \ i = 1, 2. \]
We shall also numerically investigate the accuracy of the above approximation.

2.4 Confidence Region for the Model (2.2.3)

Now we shall describe the construction of a confidence region for the parameter \( \eta \) in the model (2.2.3). Remember that \( t(\eta) \) is a \( q \times 1 \) vector and \( \eta \) is a \( s \times 1 \) vector, \( s \leq q \). The model is such that
\[
Y_i \sim N(A_i, I_q \otimes \Sigma), \ y \sim N(A_i t(\eta), \Sigma), \ S \sim W_p (\Sigma_i (N-q)) \tag{2.4.1}
\]
where, \( t_1(\eta) = (X'X)^{-1/2} t(\eta) \).

Note that \( t_1(\eta) \) is a \( q \times 1 \) vector and \( \eta \) is a \( s \times 1 \), \( s \leq q \) and the other terms in (2.4.1) are defined in (2.2.2). We shall assume that the component of \( t(\eta) \) [and hence those of \( t_1(\eta) \)] are differentiable functions of \( \eta \). Let \( T(\eta) \) be the \( p \times s \) matrix defined by

\[
T(\eta) = \frac{Y_i \partial t_1(\eta)}{\partial \eta}
\]  

(2.4.2)

We require the condition that the rank of \( T(\eta) \) is \( r \);

Similar to \( K(\theta_1) \) in (2.4) define

\[
K(\eta) = \frac{N - q - p + 1}{s} \left[ y - Y_i t_1(\eta) \right]^1 S^{-1} T(\eta)\left[ T(\eta) S^{-1} T(\eta) \right]^{-1} \times T(\eta) S^{-1} \left[ y - Y_i t_1(\eta) \right] \\
\times \left[ 1 + y^1 \{ S^{-1} - S^{-1} T(\eta) \} \left[ T(\eta) S^{-1} T(\eta) \right]^{-1} T(\eta) S^{-1} \right] y
\]

(2.4.3)

Let \( V^{(1)}_\alpha (t_1(\eta)' t_1(\eta)) \) and \( V^{(2)}_\alpha (t_1(\eta)' t_1(\eta)) \) satisfy

\[
P \left[ F_{s,N-q-p+1} t_1(\eta)' t_1(\eta) p \leq V^{(1)}_\alpha (t_1(\eta)' t_1(\eta)) \right] = 1 - \alpha, \quad \text{and}
\]

\[
P \left[ F_{s,N-q-p+1} t_1(\eta)' t_1(\eta) \delta \leq V^{(2)}_\alpha (t_1(\eta)' t_1(\eta)) \right] = 1 - \alpha,
\]

(2.4.4)

where \( \delta = E \left[ \frac{W}{1 + \frac{p-q}{N-p+1} Z} \right] \).
We then have the following conservative confidence regions for $\eta$ analogous to (2.3.29) given by

$$\{\eta : K(\eta) \leq V\alpha^{(i)} t_1(\eta)^1 t_1(\eta)\}, \ i = 1, 2. \quad (2.4.5)$$

Also notice that the regions in (2.4.5) are non empty, since they contain values of $\eta$ minimizing, $(y - Y_1 t_1(\eta))^1 S^{-1} (y - Y_1 t_1(\eta))$ so that the values of $\eta$ satisfy $T(\eta)^1 S^{-1} (y - Y_1 t_1(\eta)) = 0$.

### 2.5 Computation of the Confidence Region

We have to obtain $V_{\alpha}^{(1)} (\theta'_1 \theta_1)$ and $V_{\alpha}^{(2)} (\theta'_1 \theta_1)$ satisfying (2.3.27) and (2.3.28). Also we have to compute $V_{\alpha}^{(i)} [t_1(\eta) t_1(\eta)]$ in (2.4.5).

Recall that $\delta$ is the expected value of

$$\begin{bmatrix} W \\ 1 + \frac{p-q}{N-p+1} Z \end{bmatrix} \quad \text{where} \ W \sim \chi_p^2$$

and $Z \sim F_{p-q, N-p+1}$

We shall first derive $\delta$. For this note that

$$\begin{bmatrix} W \\ 1 + \frac{p-q}{N-p+1} Z \end{bmatrix} = \begin{bmatrix} W \\ 1 + \frac{p-q}{N-p+1} U_1 \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$$
Single Use Confidence Regions

\[
\frac{W}{1 + \frac{U_1}{U_2}} = \frac{W \cdot U_2}{U_1 + U_2}, \text{ where } U_1 \sim \chi^2_{p-q}, \ U_2 \sim \chi^2_{N-p+1} \quad (2.5.1)
\]

Hence

\[
\delta = E\left[ \frac{W}{1 + \frac{p-q}{N-p+1} Z} \right]
\]

\[
= E\left[ \frac{W \cdot U_2}{U_1 + U_2} \right] \quad (2.5.2)
\]

\[
= E(W) \ E\left( \frac{U_2}{U_1 + U_2} \right)
\]

\[
= p \times \frac{N-p+1}{N-q+1}
\]

This is obtained using the facts that \( W \) and \( Z \) are independent and

\[
E(W) = p, \ E\left( \frac{U_2}{U_1 + U_2} \right) = \frac{N-p+1}{N-q+1}, \text{ since } W \sim \chi^2_p \text{ and }
\]

\[
\left( \frac{U_2}{U_1 + U_2} \right) \sim \text{Beta} \left( \frac{N-p+1}{2}, \frac{p-q}{2} \right).
\]
Thus \( \delta = p \times \frac{N - p + 1}{N - q + 1} \). We note that \( \delta \) will be very close to \( p \) unless \( N \) is small and \( p \) and \( q \) are very much different. In other words, we expect very little difference between \( V^{(1)}_\alpha(\theta' \theta_1) \) and \( V^{(2)}_\alpha(\theta' \theta_1) \).

For the computation of the confidence region, we require the functional form of \( V^{(i)}_\alpha(\theta' \theta_1) \). This can be obtained as follows. Arguing as in Mathew and Zha (1996), we can assume an upper bound for \( d = \theta' \theta_1 \), say \( d_0 \). For a specific value of \( d(0 \leq d \leq d_0) \), \( V^{(1)}_\alpha(d) \) can be computed as the \((1 - \alpha)^{th}\) percentile of a non-central F distribution with non-centrality parameter \( \theta' \theta_1 p \) and \( df = (q, N - p - q + 1) \). We can compute \( V^{(1)}_\alpha(d) \) for a few values of \( d \) and these values can be plotted against \( d \) and a suitable function can be fitted. A polynomial of appropriate degree should provide an adequate fit. This provides the function \( V^{(1)}_\alpha(\theta' \theta_1) \) to be used for computing the confidence region. \( V^{(2)}_\alpha(\theta' \theta_1) \) and \( V^{(i)}_\alpha(t_1(\eta)' t_1(\eta)) \) can be similarly computed.

The following lemma show how the distribution of the pivotal statistic in section (2.3) and (2.4) depend on \( A_1 \) and \( \Sigma \).

**Lemma 2.3:** The distribution of \( K(\theta_1) \) in (2.3.3) and \( K(\eta) \) in (2.4.3) depend on \( A_1 \) and \( \Sigma \) only through \( A_1' \Sigma^{-1} A_1 \).
Proof: Since the distribution of $K(\theta_1)$ is invariant under the transformation

$Y_1 \rightarrow CY_1, \ y \rightarrow Cy$ and $S \rightarrow CSC'$, and since $(A_1' \Sigma^{-1} A_1, \ \theta_1)$ is a maximal invariant parameter, the lemma follows for the distribution of $K(\theta_1)$. The proof of $K(\eta)$ is similar.

The observation in lemma 2.3 can be useful for simulating the coverage probabilities of our confidence regions. The computations were carried out using the results in Abdel – Aty (1954), Tikku (1985) and Lechner et al. (1982). Similar works are reported in Liberman et al. (1967), Liman and Thomas (1988) Lundberg and Mare’ (1980), Mee et al. (1991).