2.1 INTRODUCTION

The general Gauss-Markoff model is written in the form

\[ Y = X\beta + \epsilon \]

where

- \( Y \) is an \( n \)-vector of observations,
- \( X \) is a given \( n \times p \) matrix,
- \( \beta \) is a \( p \)-vector of unknown fixed parameters,
- \( \epsilon \) is an \( n \)-vector of unknown error variables. Under the assumptions of

\[ E(\epsilon) = 0, \quad \text{Var}(\epsilon) = \Sigma \]

many authors have estimated \( \beta \) when \( \Sigma \) is known and non-singular matrix. For a singular matrix \( \Sigma \), amongst many Goldman and Zelen (1964), Mitra and Rao (1968), Zyskind and Martin (1969) have estimated \( \beta \). Later on Rao (1971c) has given the unified theory approach for the problem. Recently Rao (1985) has reviewed the entire unified theory from all aspects.

For testing prediction purposes the usual theory assumes the distribution of \( \epsilon \) as multivariate normal \( \mathcal{N}_n(0, \sigma^2 I) \). However a logarithmic transformation of
variables leads to linear expressions and it gives rise to lognormal linear model such as

\[ Z_i = \log Y_i = \mathbf{x}_i' \mathbf{\beta} + e_i, \quad i=1,2,\ldots,n, \]

where \( \mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{ip})' \) is a \( p \)-vector \((p < n)\) of non-stochastic regression values, \( \mathbf{\beta} \) is a conformable parametric \( p \)-vector and the independent random disturbances \( e_i \) has \( N(0, \sigma^2) \) distribution. Finney (1941), and Bradu and Mundlack (1970) derived minimum variance unbiased estimators of some parametric functions of this lognormal linear model.

Instead of \( e_i \) following normal or lognormal distribution one may think of \( e_i \) to follow exponential distribution, e.g. in life testing problem \( e_i \) may be failure time of the \( i^{th} \) component of an item in which each \( e_i \) is independent and has exponential probability density function such as

\[ f(e_i) = (1/\theta) \exp(-e_i/\theta), \quad 0 < e_i < \infty, \quad \theta > 0 \quad (2.1.1) \]

Further instead of failure time of the \( i^{th} \) component starting from zero if one assumes that it starts from \( y_i \) and \( y_i \) may be assumed as a linear function of parameters \( \beta_1, \beta_2, \ldots, \beta_p \) such as

\[ y_i = x_{i1} \beta_1 + x_{i2} \beta_2 + \ldots + x_{ip} \beta_p \]
then the failure time $Y_i$ of the $i^{th}$ component will be

$$Y_i = X_i\beta + e_i$$

where

$$X_i = (x_{i1}, x_{i2}, ..., x_{ip})$$ and $Y_i$ in this case are independent for $i = 1, 2, ..., n$ and $Y_i$ follows negative exponential distribution with probability density function as

$$f(Y_i) = \frac{1}{\theta} \exp\left(-\frac{Y_i - \gamma_i}{\theta}\right),$$

$$Y_i > \gamma_i, \theta > 0 \text{ for } i = 1, 2, ..., n.$$  

One may be interested not only in estimation of mean life $\theta$ but in estimation of $\beta$ and thereby $\gamma_i$ also. Further, while considering the Brownian motion of a gas, the model utilizes that the error term $e_i$ follows normal distribution with mean $\theta$ and variance $\theta^2$ and thus

$$E(e_i) = \theta, \text{ Var}(e_i) = \theta^2$$

gives rise to the above type of linear model.

In such cases we have following general linear model
\[ Y = X \beta + \varepsilon \quad (2.1.4) \]

where

\[ Y \] is an \( nx1 \) vector of random observations, \( X \) is a \( nxp \) matrix of known constants, \( \beta \) is a \( px1 \) vector of unknown fixed parameters and \( \varepsilon \) is an \( nx1 \) vector of error variables satisfying the unusual assumptions such as

\[ \mathbb{E}(\varepsilon) = \theta_1 I \quad \text{and} \quad \mathbb{D}(\varepsilon) = \theta^2 I \quad (2.1.5) \]

In this Chapter we shall consider a more general linear model defined by (2.1.4) with assumptions given in (2.1.5) and estimate linear parametric functions of \( \beta \) and \( \theta \). Since the same parameter \( \theta \) is involved both in the mean and in the variance of the error term, it creates difficulty in estimating \( \beta \) and \( \theta \) by method of maximum likelihood and by linear unbiased estimation method. The difficulty has been resolved in Section 2.2,2.3 and 2.4. In Section 2.2 maximum likelihood estimator of the parametric function \( \Delta = \mathbf{c}' \beta + \delta \theta \), where \( \mathbf{c} = (c_1, c_2, \ldots, c_p)' \) and \( \delta \) are known, is obtained together with its mean square error and comparison of it has been made with that of unbiased estimator. Section 2.3 deals with the linear unbiased estimators of \( \beta \), \( \theta \) and \( \Delta \) when the matrix involved is singular or non-singular. Further they have been compared with maximum likelihood estimators. Following Rao (1971c,1985) a unified theory approach has been developed for the
estimation of $\beta$, $\theta$ and $\Delta$ in Section 2.4.

2.2 MAXIMUM LIKELIHOOD ESTIMATION

With $X = (X_1, X_2, \ldots, X_n)'$, $Y = (Y_1, Y_2, \ldots, Y_n)'$, and $\epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n)'$,

(2.1.2) will be rewritten as

$$Y = X \beta + \epsilon.$$  

Then using the mean and variance of the distributions (2.1.1) and (2.1.3) in (2.1.2) we shall have

$$E(\epsilon) = 1 \theta,$$

$$E(Y) = X \beta + 1 \theta,$$ and

$$D(\epsilon) = D(Y) = \theta^2 I.$$ 

Let $Y_{(1)}, Y_{(2)}, \ldots, Y_{(n)}$ be the ordered sample of $Y_1, Y_2, \ldots, Y_n$. The likelihood $L(\beta, \theta)$ of $Y_1, Y_2, \ldots, Y_n$ at $Y_{(1)}, Y_{(2)}, \ldots, Y_{(n)}$ being

$$L(\beta, \theta) = \theta^{-n} \exp\left(-\frac{1}{\theta} \sum_{i=1}^{n} (Y_i - X_i \beta) / \theta\right)$$

33
the maximum likelihood estimator \( \hat{\beta} \) of \( \beta \) is given by

\[
X \hat{\beta} = \frac{1}{n} Y_{(1)} ,
\]

where

\[
i = (1, 1, \ldots, 1)'
\]
is an \( n \)-vector. Further

\[
\frac{\partial \log L(\hat{\beta}, \theta)}{\partial \theta} = 0
\]

lead to the maximum likelihood estimator \( \hat{\theta} \) of \( \theta \) as

\[
\hat{\theta} = \bar{y} - Y_{(1)} ,
\]

where

\[
\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i / n .
\]

For given \( p \)-vector \( \xi \) and a given constant \( \delta \), the maximum likelihood estimator \( \hat{\Delta} \) of \( \Delta = \xi' \beta + \delta \theta \) is

\[
\hat{\Delta} = \hat{\xi}' \frac{1}{n} Y_{(1)} + \delta \bar{y} ,
\]

where

\[
\hat{\xi}' = \xi' X - \delta \frac{1}{n} , \quad X^- \text{ is generalised inverse of matrix } X \text{ in the sense of } XX^-X = X .
\]
If \( \mathbf{c} \) belongs to the row space of \( X \) and \( \mathbf{1} \) belongs to the column space of \( X \) then \( \hat{\Delta} \) is obviously unique and unbiased.

By using

\[
E(\mathbf{Y}) = \mathbf{e} + \mathbf{1}' \times \mathbf{0}/n
\]

and

\[
E(\mathbf{Y}_\mathbf{w}) = \mathbf{1}/n + X \times \mathbf{0} \, ,
\]

\[
E(\hat{\Delta}) = \mathbf{c}' X X \times \mathbf{\theta} + \left[ \mathbf{c}' X \mathbf{1}/n + \delta - \delta/n \right] \times \mathbf{\theta}.
\]

\( \hat{\Delta} \) is unbiased for \( \Delta \) if \( \mathbf{c} \) belongs to the row space of \( X \) and \( \mathbf{1} \) belongs to the column space of \( X \).

Further, also \( \hat{\Delta} \) is unbiased for \( \Delta \) if

\[
\begin{align*}
\mathbf{c}' X - \mathbf{c}' \\
\mathbf{c}' X \mathbf{1} = \delta
\end{align*}
\]

(2.2.3)

Otherwise it is biased and the biased is

\[
(\mathbf{c}' X - \mathbf{c}') \times \mathbf{\theta} + \left[ (\mathbf{c}' X \mathbf{1}/n) - (\delta/n) \right] \times \mathbf{\theta}.
\]

Since the variance of the maximum likelihood estimator \( \hat{\Delta} \) of \( \Delta \) is
\[ \text{Var}(\hat{\Delta}) = (\xi' X^{-1} - \delta)^2 \sigma^2 / n^2 + \delta^2 \sigma^2 / n + 2\delta \sigma^2 (\xi' X^{-1} - \delta). \]

(2.2.4)

the variance of the unbiased estimator of \( \Delta \) under the conditions of estimability either given in (2.2.3) or \( \xi \) belongs to the row space of \( X \) and \( 1 \) belongs to the column space of \( X \), is \( \delta^2 \sigma^2 / n \), which does not involve \( \theta \).

We note that for fixed \( \xi, \delta, n, p \) the minimization of \( \text{Var}(\hat{\Delta}) \) is equivalent to the minimization of \( \delta^2 \sigma^2 / n \) when \( \xi' X^{-1} 1 = \delta \) holds.

Using the results for the negative exponential distribution defined in (2.1.3)

\[ E \left( Y_1 Y_2 \cdots Y_n \right) = \frac{1}{2} \theta^2 / n^2 + \left[ \theta (\theta / n) 1 + X\beta \right] \left[ \theta (\theta / n) 1 + X\beta \right]' \]

\[ E \left( Y_1 Y_2 \cdots Y_n \right) = (n+1) \theta^2 1 / n + \theta \left[ \theta (\theta / n) 1 + X\beta \right] \left[ \theta (\theta / n) 1 + X\beta \right] \frac{1}{2} X\beta X\beta / n, \]

and

\[ E(\hat{\gamma}) = \theta^2 / n + \left[ \theta + (1' X\beta / n) \right]^2, \]

the mean square error (MSE) of \( \hat{\Delta} \), the maximum likelihood estimator of \( \Delta \) and when \( \xi \) does not belong to the row space of \( X \) and \( 1 \) does not belong to the column space of \( X \), will be.
\[
\text{MSE}(\hat{\Delta}) = E(\hat{\Delta} - \Delta)^2
\]

\[
= \beta^2 E\left\{ Y, Y_{(\omega)} \right\} + 2\hat{\beta} E\left\{ Y_{(\omega)} \right\}
\]

\[
- 2\Delta \beta E\left\{ Y_{(\omega)} \right\} - 2\Delta E(\bar{Y}) + \sigma^2 E(\bar{Y}^2) + \Delta^2
\]

\[
= \sigma^2 \beta^2/n + (\langle \hat{\beta} \rangle \beta)^2 - 2(\langle \hat{\beta} \rangle \langle \hat{\beta} X^\top \beta \rangle + (\langle \hat{\beta} X^\top \beta \rangle)^2
\]

\[
+ 2(\beta_{\hat{\theta}} \beta_{1/n})^2 + 2\beta(\langle \hat{\beta} X^\top \beta \rangle (\beta_{1/n}) - 2\theta(\langle \hat{\beta} \rangle \beta_{1/n})/n
\]

\[
= \sigma^2 \beta^2/n + (\langle \hat{\beta} \rangle \beta - \langle \hat{\beta} X^\top \beta \rangle)^2 + \left(\langle \hat{\beta} X^\top \beta \rangle - \beta \right)^2/n
\]

\[
+ (\langle \hat{\beta} X^\top \beta \rangle)^2 + \left(\langle \hat{\beta} X^\top \beta \rangle - \beta \right)^2/n - \langle \hat{\beta} \beta \rangle^2
\]

\[
- (\langle \hat{\beta} \rangle \beta)^2 - (\langle \hat{\beta} X^\top \beta \rangle)^2 + 2\beta^2(\langle \hat{\beta} X^\top \beta \rangle - \beta)
\]

\[(2.2.5)\]

Under the condition of estimability either given in (2.2.3) or \( \beta \) belongs to the row space of \( X \) and \( 1 \) belongs to the column space of \( X \), the unbiased estimator of \( \Delta \) is better than the maximum likelihood estimator \( \hat{\Delta} \) given by (2.2.2) if

\[
\left[ \langle \hat{\beta} \rangle \beta - \langle \hat{\beta} X^\top \beta \rangle \right]^2 + \left[ \langle \hat{\beta} X^\top \beta \rangle - \beta \right]^2/n + \langle \hat{\beta} X^\top \beta \rangle \beta
\]
and the biased maximum likelihood estimator is better than unbiased one if the inequality in (2.2.6) is reversed.

2.3 LINEAR UNBIASED ESTIMATION

The linear model given in (2.1.2) may be written as

\[ Y = X \beta + \varepsilon \]  \hspace{1cm} (2.3.1)

where

\[ Y = (Y_1, Y_2, ..., Y_n) \] is an n-vector of random observations, \( X \) is an nxp matrix of known constants, \( \beta \) is a p-vector of unknown parameters and \( \varepsilon \) is an n-vector of random disturbances following multivariate exponential distribution with mean vector \( E(\varepsilon) = \theta I \) and dispersion matrix \( D(\varepsilon) = \theta^2 I \).

Restructuring the model as

\[ Y = X \beta + \varepsilon^* + \varepsilon, \]  \hspace{1cm} (2.3.2)
where
\[ e^* = e - \frac{1}{2} \theta \] such that
\[ \mathbb{E}(e^*) = 0 \]

and
\[ \text{Var}(e^*) = \sigma^2 I, \]

the model looks to be simple Gauss-Markoff fixed model but it is not so because it differs from Gauss-Markoff model with respect to the parameter \( \theta \) appearing not only in the linear part but it also appears in the variance of the disturbance term.

For the resticted model given in (2.3.2), one has to minimize \( \phi (\beta, \theta) \),

where
\[ \phi (\beta, \theta) = \left[ (Y - X \beta - \frac{1}{2} \theta) (Y - X \beta - \frac{1}{2} \theta) \right] / \sigma^2 \]

for all \( \theta \in \mathbb{R}^t, \beta \in \mathbb{R}^p \).

Minimization of \( \phi (\beta, \theta) \) leads to the equations

\[ X'X \beta + X' \frac{1}{2} \theta = X'Y \] \hspace{1cm} (2.3.3)

and

\[ \left( Y - X \beta \right)' \left( Y - X \beta \right) - \frac{1}{2} \left( Y - X \beta \right) \theta = 0 \] \hspace{1cm} (2.3.4)

which gives the least square estimators \( \tilde{\beta}, \tilde{\theta} \) of \( \beta, \theta \) respectively as
\[ \hat{\theta} = (X'X - (X'YX)/Y)' \begin{pmatrix} Y'Y - \{ X'X - (X'YX)/Y \} \end{pmatrix}^{-1} \begin{pmatrix} X'X - (X'YX)/Y \end{pmatrix}' \]

(2.3.5)

\[ \hat{\theta} = \begin{pmatrix} X'Y - (X'Y)/Y \end{pmatrix}^{-1} \begin{pmatrix} X'X - (X'YX)/Y \end{pmatrix}' \]

(2.3.6)

if the rank of \( X'X - X'YX \) is \( p \)

and

\[ \hat{\theta} = \begin{pmatrix} X'X - (X'YX)/Y \end{pmatrix}^{-1} \begin{pmatrix} X'Y - (X'Y)/Y \end{pmatrix}' \]

(2.3.7)

\[ \hat{\theta} = \begin{pmatrix} X'Y - (X'YX)/Y \end{pmatrix}^{-1} \begin{pmatrix} X'X - (X'YX)/Y \end{pmatrix}' \]

(2.3.8)

if the rank of \( X'X - X'YX \) is \( r(\leq p) \)

Thus \( \hat{\Delta} = \xi'\hat{\theta} + \delta \hat{\theta} \) is unique least square estimator of \( \Delta = \xi'\theta + \delta \theta \) if \( \xi \) belongs to the row space of \( X \) and \( 1 \) belongs to the column space of \( X \).

2.4 UNIFIED THEORY OF LINEAR UNBIASED ESTIMATION

Consider the model given in (2.3.1) with all the assumptions made therein. The problem is that of estimating
a linear parametric function \( a + \mathbf{L}^{'} \mathbf{Y} \) of \( \mathbf{Y} \) such that

\[
E \left[ a + \mathbf{L}^{'} \mathbf{Y} - \mathbf{c}^{'} \mathbf{\beta} - \delta \mathbf{\theta} \right] = 0 \quad (2.4.1)
\]

and

\[
(1/\theta^2)E \left[ a + \mathbf{L}^{'} \mathbf{Y} - \mathbf{c}^{'} \mathbf{\beta} - \delta \mathbf{\theta} \right] ^2 \quad (2.4.2)
\]

is minimum.

The divisor in (2.4.2) is due to making the model as usual Gauss–Markoff model, since \( \mathbf{\theta} \) is involved in mean vector and covariance matrix of error vector \( \mathbf{e} \).

The condition (2.4.1) leads to the equation

\[
a + \mathbf{L}^{'} \left( \mathbf{X} \mathbf{\beta} + \mathbf{1} \mathbf{\theta} \right) = \mathbf{c}^{'} \mathbf{\beta} + \delta \mathbf{\theta}
\]

which is equivalent to

\[
a = 0, \quad \mathbf{L}^{'} \mathbf{X} = \mathbf{c}^{'}, \quad \mathbf{L}^{'} \mathbf{1} = \delta \quad (2.4.3)
\]

or

\[
\mathbf{L}^{'} \mathbf{Z} = \mathbf{\eta}
\]

where

\( \mathbf{Z} = (\mathbf{X} ; \mathbf{1}) \) is a matrix of order \( n \times (p+1), \) \( \mathbf{\eta}^{'} = (\mathbf{c}^{'} = \delta) \)

is a \( (p+1) \)-row vector.

When (2.4.3) holds,

\[
\left(1/\theta^2\right)E \left[ a + \mathbf{L}^{'} \mathbf{Y} - \mathbf{c}^{'} \mathbf{\beta} - \delta \mathbf{\theta} \right] ^2 = \left(1/\theta^2\right) \left[ \theta^2 \mathbf{L}^{'} \mathbf{L} \right] \quad (2.4.4)
\]
The minimum of (2.4.4) subject to (2.4.3) is attained when \( L \) satisfies the equation

\[
\begin{pmatrix}
I_n & Z \\
Z' & 0
\end{pmatrix}
\begin{pmatrix}
L \\
\lambda
\end{pmatrix} =
\begin{pmatrix}
0 \\
\eta
\end{pmatrix}
\]  

(2.4.5)

where

\( \lambda \) is Lagrangian \((p+1)\)-vector.

Equation (2.4.5) is of the same type of equation which occurs in unified theory of linear estimation for fixed linear model given in Rao (1971c, 1973b pp. 298-300).

Let

\[
\begin{pmatrix}
I_n & Z \\
Z' & 0
\end{pmatrix}^{-1} =
\begin{pmatrix}
C_1 & C_2 \\
C_3 & -C_4
\end{pmatrix}
\]  

(2.4.6)

for one choice of the g-inverse. Then a solution \( L^* \) for \( L \) is

\[
L^* = C_2 \eta
\]

leading to the optimum estimator \( \Delta^* \) of \( \Delta \) as

\[
\Delta^* = \eta' C_2' Y
\]  

(2.4.7)

For given \( \eta' = (\xi; \delta) \), using the result given in Rao (1973b, pp. 296) and using \( Z = (X ; 1) \), (2.4.7) gives
\[ \Delta_\delta = \begin{pmatrix} X'X & X' \delta \\ \delta X & \delta \delta \end{pmatrix} \begin{pmatrix} X' \\ \delta \end{pmatrix} \] 

(2.4.8)

Let us consider some special cases of (2.4.8).

Case I:

Let

\[ \begin{vmatrix} X'X & X' \\ \delta X & \delta \delta \end{vmatrix} \neq 0 \text{ and } |X'X| \neq 0 \]

Then (2.4.8) reduces to

\[ \Delta_\delta = c' (X'X)^{-1} X'Y + \left[ \delta - c' (X'X)^{-1} X' \right] \begin{pmatrix} I_A X / I_A X \end{pmatrix} \]

(2.4.9)

and

\[ \text{Var}(\Delta_\delta) = \delta \left[ n c' (X'X)^{-1} c (\begin{pmatrix} I_A \\ \delta A \end{pmatrix}) + (c' BB c) n + \delta^2 I X B - 2n \delta c' B \right. \]

\[ + \left. \delta (\begin{pmatrix} I_A \\ \delta A \end{pmatrix}) / n (\begin{pmatrix} I_A \\ \delta A \end{pmatrix}) \right], \]

(2.4.10)

where

\[ A = I - X(X'X)^{-1} X', \quad B = (X'X)^{-1} X' \]

Comparing the variance of \( \Delta_\delta \) given by (2.4.10) with that of unbiased estimator based on maximum likelihood...
estimator, $\Delta_k$ is better than the unbiased estimator based on maximum likelihood estimator provided

$$(1^t A 1)\leq (X^t X)^{-1} + B B^t + \delta^2 (1^t X B)/n \leq 2\delta B$$ (2.4.11)

and $\Delta_k$ is not better if the inequality in (2.4.11) is reversed.

Case II (A):

Let$$\begin{bmatrix} X^t X & X^t 1 \\ 1^t X & 1^t 1 \end{bmatrix} = 0 \text{ and rank of } (X^t X) \text{ is } p.$$ Then by using the result of Rao and Mitra (1971a, pp. 40), one choice of the generalized inverse of

$$\begin{bmatrix} X^t X & X^t 1 \\ 1^t X & 1^t 1 \end{bmatrix}$$

will be

$$\begin{bmatrix} X^t X & X^t 1 \\ 1^t X & 1^t 1 \end{bmatrix}^{-1} = \frac{1}{k} \begin{bmatrix} (X^t X)^{-1} & -X^t (X^t X)^{-1} \\ X (X^t X)^{-1} & 1 \end{bmatrix}$$
\[-\frac{n}{k(k+n)} \left\{ \begin{array}{cc}
(X'X)^{-1} X' & \frac{1}{2} X'(X'X)^{-1} \\
\frac{1}{2} X(X'X)^{-1} & 1
\end{array} \right\} \]

where \(k = \frac{1}{2} X(X'X)^{-1} X' \frac{1}{2}\)

\[= \begin{pmatrix}
C_1 & C_2 \\
C_3 & -C_4
\end{pmatrix} \quad (2.4.12)\]

where

\[C_1 = (X'X)^{-1} - \frac{(k + 2n)}{k(k + n)} (X'X)^{-1} X' \frac{1}{2} X(X'X)^{-1}\]

\[C_2 = \frac{(X'X)^{-1} X' \frac{1}{2}}{k + n}, \quad C_3 = \frac{1}{2} X(X'X)^{-1}, \quad C_4 = \frac{1}{k + n} \cdot\]

Case II (B):

Next let

\[
\begin{bmatrix}
X'X & X' \frac{1}{2} \\
\frac{1}{2} X & \frac{1}{2} \frac{1}{2}
\end{bmatrix} = 0 \text{ and rank of } (X'X)
\]

be \(r < \rho\). Then one choice of g-inverse in this case will be

\[
\begin{pmatrix}
X'X & X' \frac{1}{2} \\
\frac{1}{2} X & \frac{1}{2} \frac{1}{2}
\end{pmatrix}^{-1} = \begin{pmatrix}
D_1 & D_2 \\
D_3 & -D_4
\end{pmatrix} \quad (2.4.13)
\]

where
\[ D_1 = (X'X)^{-1} - \frac{m + 2n}{m(m + n)} (X'X)^{-1} X' \frac{1}{m + n} X (X'X)^{-1}, \]
\[ D_2 = \frac{(X'X)^{-1} X'}{m + n}, \quad D_3 = \frac{1}{m + n} X (X'X)^{-1}, \quad D_4 = \frac{1}{m + n}, \]

where
\[ m = \frac{1}{n} X (X'X) X' \frac{1}{n} \]

Then with the use of (2.4.12) and (2.4.13) the best linear unbiased estimators \( \Delta^* \) and \( \Delta^{**} \) of \( \Delta = \xi \beta + \delta \theta \) in Case II(A) and II(B) are respectively,
\[ \Delta^* = \xi \beta^* + \delta \theta^*, \quad \Delta^{**} = \xi \beta^{**} + \delta \theta^{**}, \]
and their variances are
\[ \text{Var}(\Delta^*) = \theta^2 (\xi' M + \delta N') (\xi' M + \delta N')' \]
\[ \text{Var}(\Delta^{**}) = \theta^2 (\xi' P + \delta O') (\xi' P + \delta O')', \]
where
\[ \beta^* = MY, \quad M = C_X' + C_1' \]
\[ \theta^* = N'Y, \quad N' = C_X' + \frac{1}{2} \]
and
\[ \beta^{**} = PY, \quad P = D_X' + D_1' \]
Again writing $C_2 = (C_\beta : C_\theta)$, (2.4.7) will be

\[ C_2 = (C_\beta : C_\theta) \]

Let

\[ C_4 = \begin{pmatrix} C_{\beta\beta} & C_{\beta\theta} \\ C_{\theta\beta} & C_{\theta\theta} \end{pmatrix} \quad (2.4.16) \]

Then we shall have following remarks analogous to the remarks mentioned in Rao (1973b, pp. 298-300).

**Remark 4.1**: The best estimator of $C_\beta$ alone when estimable, i.e. when (2.4.3) holds with $\delta = 0$, is $C_\beta^*$, where $C_\beta^*$ is same as given in (2.4.15) and

\[ E(\xi C_\beta^* - \xi')^2 = \xi' C_{\beta\beta} \xi \quad (2.4.17) \]

**Remark 4.2**: The best estimator of $C_\theta$ when estimable, i.e. when (2.4.3) holds with $\xi = 0$, is $C_\theta^*$, where $C_\theta^*$ is same as given in (2.4.15) and
\[
E(\delta \theta^*_x - \delta \theta)^2 = \delta^2 C_{\theta \theta}
\] (2.4.18)

**Remark 4.3**: The optimum estimator \( \Delta^*_x \) of \( \Delta \) is given in (2.4.14) and

\[
E(\Delta^*_x - \Delta)^2 = \xi' C \beta \xi + \delta^2 C_{\theta \theta} - 2 \delta \xi' C_{\theta \theta} (2.4.19)
\]

The results (2.4.17), (2.4.18), and (2.4.19) depend only on the elements of the inverse partitioned matrix given in (2.4.16) and these elements can be obtained in the same way as is given in the unified theory for the Gauss-Markoff model [see Rao (1973b), pp. 298-300]. Thus we have a complete theory for the most general case without making any assumptions on the ranks of the matrices involved, which were done in Section 2.3.