Shrinkage estimation of the regression parameters with multivariate normal errors

M. Arashi* and S.M.M. Tabatabaey
Department of Statistics,
Ferdowsi University of Mashhad, Mashhad

Abstract

In the linear model $y = X\beta + e$ with the errors distributed as normal, we obtain generalized least square (GLS), restricted GLS (RGLS), preliminary test (PT), Stein-type shrinkage (S) and positive-rule shrinkage (PRS) estimators for regression vector parameter $\beta$ when the covariance structure is known. We compare the quadratic risks of the underlying estimators and propose the dominance orders of the five estimators.

Keywords and phrases: GLS estimator, preliminary test estimator, stein-type shrinkage estimator, positive-rule shrinkage estimator.

AMS Subject Classification 2000: Primary 62J05; Secondary 62H12.

1 Introduction

The most important model belonging to the class of general linear hypotheses is the multiple regression model. The general purpose of multiple regression is

*E-mail: marashi.stat@yahoo.com
to learn more about the relationship between several independent or predictor variables and a dependent or criterion variable.

To deal with a common multiple regression equation, consider the linear model

$$ y = X\beta + e, $$

where $y$ is an $n$-vector of response, $X$ is an $n \times p$ design matrix with full rank $p$, $\beta = (\beta_1, \cdots, \beta_p)'$ is a $p$-vector of regression coefficients and $e = (e_1, \cdots, e_n)'$ is the $n$-vector of errors distributed as multivariate normal with location parameter zero and positive definite (p.d.) covariance matrix $\Sigma$, denoted by $e \sim N_n(0, \Sigma)$. Then directly

$$ y \sim N_n(X\beta, \Sigma). $$

Let us assume that in addition to the sample information $y$ in the model (1.1), that information also exists in the form of $q$ independent linear hypothesis about the unknown vector parameter $\beta$ where $q \leq p$. These general restrictions can be shown as

$$ H\beta = h, $$

where $H$ is a $q \times p$ known hypothesis design matrix of rank $q$ and $h$ is a $q \times 1$ vector of prespecified, hypothetical values.

The estimation of parameters of the multiple regression model is a common interest to many users. Often the properties of the estimators are of prime concern. Selection of any specific statistical property of any estimator often depends on the objective of the study. The choice of any particular estimator may very well be determined by the aim of the end users. It is well known that the ordinary least squares estimators are best linear unbiased. However, if the objective of any study is to minimize some specific risk function then other types of estimators perform better than the ordinary least squares estimator. Our primary object of this paper is to estimate $\beta$ when the p.d. covariance matrix $\Sigma$ is known under the subspace restriction (1.3); and then obtain shrinkage estimators of $\beta$ using the
likelihood ratio test (LRT) statistic of (1.3). For complete review of underlying
study in the special case \( \Sigma = \sigma^2 I_n \) for both known and unknown \( \sigma^2 \) and may \( \sigma^2 \)
have inverse gamma distribution, see Saleh and Han [9], Tabatabaey [13], Khan
[5, 6], Srivastava and Saleh [12] and Saleh [10].

2 Estimation

Given classical conditions (see Kuan [7]), it is well known that for known p.d.
covariance matrix \( \Sigma \), the generalized least square (GLS) estimator of \( \beta \) is

\[
\hat{\beta} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}y. \tag{2.1}
\]

Obtaining GLS estimator of \( \beta \) under the constraint \( H_0 : H\beta = h \), using method
of Lagrangian multipliers, the restricted GLS estimator of \( \beta \) subject to the linear
restriction \( H_0 : H\beta = h \) as \( \tilde{\beta} \) is given by

\[
\tilde{\beta} = \hat{\beta} - (X'\Sigma^{-1}X)^{-1}H'[H(X'\Sigma^{-1}X)^{-1}H']^{-1}(H\hat{\beta} - h). \tag{2.2}
\]

See Ravishanker and Dey [8].

Let \( G_1 = (X'\Sigma^{-1}X)^{-1} \) and \( G_2 = [HG_1H']^{-1} \), then simplifying (2.2) we obtain

\[
\tilde{\beta} = \hat{\beta} - G_1H'G_2(H\hat{\beta} - h). \tag{2.3}
\]

Now we consider the linear hypothesis \( H\beta = h \) in (1.3) and obtain the test
statistic for the null hypothesis \( H_0 : H\beta = h \).

Now let \( \omega = \{ \beta : \beta \in \mathbb{R}^p, \ H\beta = h, \ \Sigma > 0 \} \) and \( \Omega = \{ \beta : \beta \in \mathbb{R}^p, \ \Sigma > 0 \} \), then
the likelihood test statistic for underlying hypothesis is

\[
\lambda = \frac{\max_{\beta \in \omega} L(\beta, \Sigma)}{\max_{\beta \in \Omega} L(\beta, \Sigma)}
= \frac{\exp\left\{ \frac{1}{2} [(y - X\tilde{\beta})'\Sigma^{-1}(y - X\tilde{\beta})] \right\}}{\exp\left\{ \frac{1}{2} [(y - X\hat{\beta})'\Sigma^{-1}(y - X\hat{\beta})] \right\}}
= \exp\left\{ \frac{1}{2} [ (H\tilde{\beta} - h)'G_2(H\tilde{\beta} - h) ] \right\}.
\]
which is a decreasing function with respect to $\chi = (H\hat{\beta} - h)^{\prime}G_{2}(H\hat{\beta} - h)$.
Let $u = G_{2}^{1/2}(H\hat{\beta} - h)$; then using (1.2), $\chi = u^{\prime}u$ has non-central chi-square distribution with $q$ degrees of freedom and noncentrality parameter $\mu^{\prime}\mu/2$, where $\mu = G_{2}^{1/2}(H\beta - h)$.
Bancroft [2] defined the preliminary test estimator (PTE) of $\beta$ as a convex combination of $\hat{\beta}$ and $\tilde{\beta}$ by

$$
\hat{\beta}^{PT} = \tilde{\beta} + [1 - I(\chi \leq \chi^2(\alpha))] (\hat{\beta} - \tilde{\beta}),
$$

where $I(A)$ is the indicator of the set $A$ and $\chi^2(\alpha)$ is the upper 100$\alpha$ percentile of the central $\chi^2$ distribution with $q$ degrees of freedom.
The PTE has the disadvantage that it depends on $\alpha$ ($0 < \alpha < 1$), the level of significance and also it yields the extreme results, namely $\hat{\beta}$ and $\tilde{\beta}$ depending on the outcome of the test. Therefore we define an intermediate value as Stein-type shrinkage estimator (SE) of $\beta$, by

$$
\hat{\beta}^{S} = \tilde{\beta} + (1 - \rho \chi^{-1})(\hat{\beta} - \tilde{\beta}),
$$

where

$$
\rho = \frac{(q - 2)(n - p)}{q(n - p + 2)} \quad \text{and} \quad q \geq 3.
$$

The SE has the disadvantage that it has strange behavior for small values of $\chi$.
Also, the shrinkage factor $(1 - \rho \chi^{-1})$ becomes negative for $\chi < \rho$. Hence we define a better estimator by positive-rule shrinkage estimator (PRSE) of $\beta$ as

$$
\hat{\beta}^{S^+} = \tilde{\beta} + (1 - \rho \chi^{-1})I[\chi > \rho](\hat{\beta} - \tilde{\beta})
= \hat{\beta}^{S} - (1 - \rho \chi^{-1})I[\chi \leq \rho](\hat{\beta} - \tilde{\beta}).
$$

Note that this estimator is a convex combination of $\hat{\beta}$ and $\tilde{\beta}$.
The quadratic risk functions of the estimators are given in the following section and the dominance properties are studied in section 4.
3 Risk Evaluations

For a given non-singular matrix $W$, consider the *weighted quadratic error loss function* of the form

$$L(\beta^*; \beta) = (\beta^* - \beta)' W (\beta^* - \beta), \quad (3.1)$$

where $\beta^*$ is any estimator of $\beta$. Then the weighted quadratic risk function associated with (3.1) is defined as

$$R(\beta^*; \beta) = E[(\beta^* - \beta)' W (\beta^* - \beta)]. \quad (3.2)$$

In this section, using the risk function (3.2), we evaluate the quadratic risks of the five different estimators under study.

Direct computations using (1.2), (2.1) and (3.2) lead to

$$R(\hat{\beta}; \beta) = tr(G_1W). \quad (3.3)$$

Let $\delta = G_1H'G_2(H\beta - h)$, then using (2.4) we have

$$R(\hat{\beta}; \beta) = tr\{W[G_1(I_p - H'G_2H_1) + \delta\delta']\} = tr(G_1W) - tr\{W[G_1(H'G_2H_1)]\} + \delta'W\delta. \quad (3.4)$$

Note that $R = G_1^{1/2}H'G_2H_1^{1/2}$ is a symmetric idempotent matrix of rank $q \leq p$. Thus, there exists an orthogonal matrix $Q$ ($Q'Q = I_p$) (see Judge and Bock [4]) such that

$$QRQ' = \begin{bmatrix} I_q & 0 \\ 0 & 0 \end{bmatrix}, \quad (3.5)$$

$$QG_1^{1/2}WG_1^{1/2}Q' = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = A. \quad (3.6)$$

The matrices $A_{11}$ and $A_{22}$ are of orders $q$ and $p - q$, respectively.

Define the random variable

$$w = QG_1^{-1/2}\hat{\beta} - QG_1^{1/2}H'G_2h, \quad (3.7)$$
then

\[ w \sim N_q(\eta, I_p). \] (3.8)

Also

\[ \eta = QG_{1}^{-1/2}\beta - QG_{1}^{1/2}H'G_2h. \] (3.9)

Partitioning the vectors \( w = (w_1', w_2')' \) and \( \eta = (\eta_1', \eta_2')' \), where \( w_1 \) and \( w_2 \) are independent sub-vectors of orders \( q \) and \( p - q \) respectively, we obtain

\[ \hat{\beta} - \beta = G_{1}^{1/2}Q(w - \eta). \] (3.10)

Using (3.7) we can obtain

\[ \chi = w_1'w_1, \theta = \eta_1'\eta_1 = (H\beta - h)'G_2(H\beta - h). \] (3.11)

Now, we may write

\[
\begin{align*}
tr\{W[G_1H'G_2HG_1]\} & = tr\{QG_{1}^{1/2}WG_{1}^{1/2}Q'QR'\} \\
& = tr\left\{ \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I_q & 0 \\ 0 & 0 \end{bmatrix} \right\} \\
& = tr(A_{11}).
\end{align*}
\] (3.12)

Using (3.11) we have

\[ \delta'W\delta = (H\beta - h)'G_2HG_1W[G_1H'G_2(H\beta - h)] = \eta_1'A_{11}\eta_1. \] (3.13)

Therefore, we obtain

\[ R(\tilde{\beta}; \beta) = tr(G_1W) - tr(A_{11}) + \eta_1'A_{11}\eta_1. \] (3.14)

Using (2.5)

\[
\begin{align*}
R(\hat{\beta}^{PT}; \beta) & = E[(\hat{\beta}^{PT} - \beta)'W(\hat{\beta}^{PT} - \beta)] \\
& = E\{(\hat{\beta} - \beta)' - I(\chi \leq \chi_0^2)(\hat{\beta} - \tilde{\beta})W(\hat{\beta} - \beta) \\
& \quad - I(\chi \leq \chi_0^2)(\hat{\beta} - \tilde{\beta})\} \\
& = E[(\hat{\beta} - \beta)'(\hat{\beta} - \beta)] - 2E[I(\chi \leq \chi_0^2)(\hat{\beta} - \beta)'W(\hat{\beta} - \tilde{\beta})] \\
& \quad + E[I(\chi \leq \chi_0^2)(\hat{\beta} - \tilde{\beta})W(\hat{\beta} - \tilde{\beta})].
\end{align*}
\] (3.15)
Using (3.7)-(3.11) and (3.15)
\[ R(\hat{\beta}^{PT}; \beta) = tr(G_1W) - E[w'_1A_{11}w_1I(\chi \leq \chi^2_\alpha)] \\
-2E[w'_2A_{21}w_1I(\chi \leq \chi^2_\alpha)] + 2\eta'_1A_{11}E[w_1I(\chi \leq \chi^2_\alpha)] \\
+ 2\eta'_2A_{21}E[w_1I(\chi \leq \chi^2_\alpha)], \quad (3.16) \]

because \( w_1 \) and \( w_2 \) are independent

\[ E[w'_2A_{21}w_1I(\chi \leq \chi^2_\alpha)] = \eta'_2A_{21}E[w_1I(\chi \leq \chi^2_\alpha)], \quad (3.17) \]

using Lemma 1 in Appendix with \( \phi(\chi) \) as indicator function of \( \chi \), we get

\[ R(\hat{\beta}^{PT}; \beta) = tr(G_1W) - \chi^2_{q+2, \rho(\alpha)}(\hat{\beta}, \beta) \\
+ [2\chi^2_{q+2, \rho}(\alpha) - 2\chi^2_{q+4, \rho}(\alpha)]\eta'_1A_{11}\eta_1. \quad (3.18) \]

Using (2.6) and (2.7)
\[ R(\hat{\beta}^S; \beta) = E[(\hat{\beta}^S - \beta)'W(\hat{\beta}^S - \beta)] \]
\[ = E[(\hat{\beta} - \beta)'W(\hat{\beta} - \beta)] + \rho E[\chi^{-1}(\hat{\beta} - \beta)'W(\hat{\beta} - \beta)] \\
+ \rho^2 E[\chi^{-2}(\hat{\beta} - \beta)'W(\hat{\beta} - \beta)], \quad (3.19) \]

using (3.7)-(3.11) and (3.15)
\[ R(\hat{\beta}^S; \beta) = tr(G_1W) - 2\rho E[\chi^{-1}(w'_1A_{11}w_1 - \eta'_1A_{11}w_1 + w'_2A_{21}w_1) \\
- \eta'_2A_{21}w_1)] + \rho^2 E[\chi^{-2}(w'_1A_{11}w_1)]. \quad (3.20) \]

Using Lemma 1 in Appendix for \( \phi(\chi) = \chi^{-1} \), we have
\[ E[\chi^{-1}\eta'_1A_{11}w_1] = \eta'_1A_{11}\eta_1 E\left[ \frac{1}{\chi^2_{q+2, \rho}} \right], \quad (3.21) \]
\[ E[\chi^{-1}w'_1A_{11}w_1] = E\left[ \frac{1}{\chi^2_{q+2, \rho}} \right] tr(A_{11}) + E\left[ \frac{1}{\chi^2_{q+4, \rho}} \right] \eta'_1A_{11}\eta_1. \quad (3.22) \]

Using Lemma 1 in Appendix for \( \phi(\chi) = \chi^{-2} \), we have
\[ E[\chi^{-2}w'_1A_{11}w_1] = E\left[ \frac{1}{\chi^2_{q+2, \rho}} \right]^2 tr(A_{11}) + E\left[ \frac{1}{\chi^2_{q+4, \rho}} \right]^2 \eta'_1A_{11}\eta_1. \quad (3.23) \]
Using (3.21)-(3.23) one can obtain

\[
R(\hat{\beta}^S; \beta) = \text{tr}(G_1W) - \rho \left\{ 2E \left[ \frac{1}{\lambda_{q+2,\theta}} \right] - \rho E \left[ \frac{1}{\lambda_{q+4,\theta}} \right] \right\} \text{tr}(A_{11})
+ \rho \left\{ 2E \left[ \frac{1}{\lambda_{q+2,\theta}} \right] - 2E \left[ \frac{1}{\lambda_{q+4,\theta}} \right] \right\} \eta_1^t A_{11} \eta_1. \tag{3.24}
\]

Finally the risk of PRSE is given by

\[
R(\hat{\beta}^{S+}; \beta) = E[(\hat{\beta}^{S+} - \beta)^t W(\hat{\beta}^{S+} - \beta)]
\]
\[
= E[(\hat{\beta}^S - \beta - (1 - \rho \chi^{-1})I(\chi \leq \rho)(\hat{\beta} - \tilde{\beta}))^t W(\hat{\beta}^S - \beta - (1 - \rho \chi^{-1})I(\chi \leq \rho)(\hat{\beta} - \tilde{\beta})] - 2E[(\hat{\beta}^S - \beta)^t W(1 - \rho \chi^{-1})I(\chi \leq \rho)(\hat{\beta} - \tilde{\beta})]. \tag{3.25}
\]

But using (2.6)

\[
E[(\hat{\beta}^S - \beta)^t W(1 - \rho \chi^{-1})I(\chi \leq \rho)(\hat{\beta} - \tilde{\beta})]
\]
\[
= E[(\hat{\beta} - \beta)^t W(1 - \rho \chi^{-1})I(\chi \leq \rho)(\hat{\beta} - \tilde{\beta})]
\]
\[
= E\{ (\hat{\beta} - \beta)^t W[(1 - \rho \chi^{-1})I(\chi \leq \rho)(\hat{\beta} - \tilde{\beta})] \}
+ E[(1 - \rho \chi^{-1})^2 I(\chi \leq \rho)(\hat{\beta} - \tilde{\beta})^t W(\beta - \tilde{\beta})]. \tag{3.26}
\]

Thus, we obtain

\[
R(\hat{\beta}^{S+}; \beta) = R(\hat{\beta}^{S+}; \beta) - E[(1 - \rho \chi^{-1})^2 I(\chi \leq \rho)(\hat{\beta} - \tilde{\beta})^t W(\beta - \tilde{\beta})]
- 2E\{ (\hat{\beta} - \beta)^t W[(1 - \rho \chi^{-1})I(\chi \leq \rho)(\hat{\beta} - \tilde{\beta})] \}. \tag{3.27}
\]
Using (3.21)-(3.23), and Lemma 1 in Appendix for \( \phi(\chi) = (1 - \rho \chi^{-1})I(\chi \leq \rho), \)
(i = 1, 2) we get
\[
\begin{align*}
R(\hat{\beta}^S; \beta) &= R(\hat{\beta}^T; \beta) - E \left[ \left( 1 - \frac{\rho}{\chi_{q+2, \theta}} \right)^2 I(\chi_{q+2, \theta}^2 \leq \rho) \right] \text{tr}(A_{11}) \\
&+ E \left[ \left( 1 - \frac{\rho}{\chi_{q+4, \theta}} \right)^2 I(\chi_{q+4, \theta}^2 \leq \rho) \right] \eta_1'^i A_{11} \eta_1 \\
&- 2E \left[ \left( \frac{\rho}{\chi_{q+2, \theta}} - 1 \right) \chi_{q+2, \theta}^2 \leq \rho \right] \eta_1'^i A_{11} \eta_1
\end{align*}
\]
(3.28)

4 Comparison

Providing risk analysis of the underlying estimators with the weight matrix \( W \),
we have (see e.g. Searle [11])
\[
\theta ch_1(A_{11}) \leq \eta_1'^i A_{11} \eta_1 \leq \theta ch_q(A_{11}),
\]
(4.1)
where \( ch_1(A_{11}) \) and \( ch_q(A_{11}) \) are the minimum and maximum eigenvalues of \( A_{11} \)
respectively. Then by (3.3) and (3.14) one may easily see that
\[
R(\hat{\beta}; \beta) - \text{tr}(A_{11}) + \theta ch_1(A_{11}) \leq R(\tilde{\beta}; \beta) \leq R(\hat{\beta}; \beta) - \text{tr}(A_{11}) + \theta ch_q(A_{11}).
\]

By (3.11) and (3.30), under the null hypothesis \( H_0 : H \beta = h \), we conclude
\[
R(\tilde{\beta}; \beta) \leq R(\hat{\beta}; \beta).
\]

In general by (3.30), \( \tilde{\beta} \) performs better than \( \hat{\beta} \) whenever
\[
\theta \leq \frac{\text{tr}(A_{11})}{ch_q(A_{11})} = \frac{\sum_{i=1}^q ch_i(A_{11})}{ch_q(A_{11})} \leq q.
\]

Using (3.3) and (3.18) we have
\[
R(\hat{\beta}^{PT}; \beta) - R(\hat{\beta}; \beta) = [2\chi_{q+2, \theta}^2(\alpha) - \chi_{q+4, \theta}^2(\alpha)]\eta_1'^i A_{11} \eta_1 \\
- \chi_{q+2, \theta}^2(\alpha)\text{tr}(A_{11}).
\]
(4.2)
Therefore \( \tilde{\beta}^{PT} \) performs better than \( \hat{\beta} \) whenever

\[
\theta \leq \frac{tr(A_{11}) \times \chi^2_{q+2,\theta}(\alpha)}{ch_q(A_{11}) \times [2\chi^2_{q+2,\theta}(\alpha) - \chi^2_{q+4,\theta}(\alpha)]}.
\]

(4.3)

Because \( tr(A_{11}) = q \), (3.33) satisfies for \( W = X'\Sigma^{-1}X \).

Also under the null hypothesis \( H_0 \), since (3.32) is negative for all \( \alpha \), \( \tilde{\beta}^{PT} \) performs better than \( \hat{\beta} \).

Using (3.14), (3.18) and the risks difference we can conclude that \( \tilde{\beta}^{PT} \) performs better than \( \tilde{\beta} \) whenever

\[
\theta \geq \frac{[1 - \chi^2_{q+2,\theta}(\alpha)]tr(A_{11})}{[1 - 2\chi^2_{q+2,\theta}(\alpha) + \chi^2_{q+4,\theta}(\alpha)]ch_q(A_{11})}.
\]

(4.4)

Thus, the dominance order of the three estimator \( \hat{\beta}, \tilde{\beta} \) and \( \tilde{\beta}^{PT} \), under the null hypothesis \( H_0 \) is given by

\( \tilde{\beta} \succ \tilde{\beta}^{PT} \succ \hat{\beta} \),

where the notation \( \succ \) means dominate.

Under the null hypothesis,

\[
R(\tilde{\beta}^S; \beta) - R(\hat{\beta}; \beta) = -\rho tr(A_{11}) \frac{2(q - 2)}{q(q - 2)}.
\]

By the direct computations using the fact \( n \geq p \), we get \( \rho \leq 2(q - 2) \). Therefore, the risk difference \( R(\tilde{\beta}^S; \beta) - R(\hat{\beta}; \beta) \) is negative and \( \tilde{\beta}^S \succ \hat{\beta} \) uniformly.

Under the null hypothesis \( H_0 \), we have

\[
R(\hat{\beta}^S; \beta) = R(\tilde{\beta}; \beta) + tr(A_{11})f(n, q, p),
\]

where \( f(n, q, p) = \frac{q^2 - 2q(q - 2) + q(q - 2)}{q(q - 2)} \).

The function \( f(n, q, p) \) is positive for \( q \geq 3 \). Thus \( R(\hat{\beta}^S; \beta) > R(\tilde{\beta}; \beta) \). However, as \( \eta_1 \) moves away from 0, \( \eta_1' A_{11} \eta_1 \) increases and the risk of \( \tilde{\beta} \) becomes unbounded while the risk of \( \tilde{\beta}^S \) remains below the risk of \( \hat{\beta} \); thus \( \tilde{\beta}^S \) dominates \( \tilde{\beta} \) outside an interval around the origin.
Comparing $\hat{\beta}^S$ and $\hat{\beta}^{PT}$, under $H_0$, we get
\[
R(\hat{\beta}^S; \beta) = R(\hat{\beta}^{PT}) + \left[ \frac{1}{\chi^2_{q+2,0}^2(\alpha)} - \frac{2\rho E[\chi^2_{q+2,0}]}{\chi^2_{q+2,0}^2} + \frac{\rho^2 E[\chi^2_{q+2,0}]}{\chi^2_{q+2,0}^2} \right] \text{tr}(A_{11}) \\
= R(\hat{\beta}^{PT}) + \left[ \frac{1}{\chi^2_{q+2,0}^2(\alpha)} - \frac{2\rho}{q} + \frac{\rho^2}{q(q-2)} \right] \text{tr}(A_{11}) \\
\geq R(\hat{\beta}^{PT}),
\]
for all $\alpha$ such that \( l = \frac{1}{\chi^2_{q+2,0}^2(\alpha)} - \frac{2\rho}{q} + \frac{\rho^2}{q(q-2)} \geq 0 \) and $R(\hat{\beta}^S; \beta) \leq R(\hat{\beta}^{PT})$ for all $\alpha$ such that $l \leq 0$.

Because $w_1$ is independent of $w_2$, we get
\[
R(\hat{\beta}^{S+}; \beta) - R(\hat{\beta}^S; \beta) = -E[(1 - \rho \chi^{-1})^2 I(\chi \leq \rho) w_1 A_{11} w_1] \\
-2E[(1 - \rho \chi^{-1}) I(\chi \leq \rho) (w_1 A_{11} w_1 - \eta_1^T A_{11} \eta_1)].
\]

(4.5)

Note that for such $\theta$ under which $\chi^2_{q+2,0, \theta} \leq \rho$ we have
\[
E\left[ (1 - \frac{\rho}{\chi^2_{q+2,0, \theta}}) I(\chi^2_{q+2,0, \theta} \leq \rho) \right] \leq 0.
\]

Moreover, the expectation of a positive random variable, is positive, then one can obtain the risk difference in (4.5) is negative. Therefore, for all $\beta$, $\tilde{\beta}^S \succ \hat{\beta}^S$ and under $H_0$, $\tilde{\beta} \succ \hat{\beta}^{S+}$.

However, as $\eta_1$ moves away from 0, $\eta_1^T A_{11} \eta_1$ increases and the risk of $\tilde{\beta}$ becomes unbounded while the risk of $\hat{\beta}^{S+}$ remains below the risk of $\hat{\beta}$; thus $\tilde{\beta}^{S+}$ dominates $\tilde{\beta}$ outside an interval around the origin.

Under the conditions are given above, it can be found that the dominance order of five estimators of $\beta$ can be categorized in the following two orders:

1. $\tilde{\beta} \succ \hat{\beta}^{PT} \succ \hat{\beta}^{S+} \succ \hat{\beta}^S \succ \hat{\beta}$

\hspace{1cm} (4.6)

and

2. $\tilde{\beta} \succ \hat{\beta}^{S+} \succ \hat{\beta}^S \succ \hat{\beta}^{PT} \succ \hat{\beta}$.

(4.7)
5 Illustrative Example

For an illustrative example of domination orders of five estimators under study, we proceed with numerical and graphical examples.

**Numerical Example** Now for an illustrative example of domination orders given in the previous section, we accomplish with a numerical example from Seale [11]. Suppose we have the following five sets of observations (including \( x_{i0} = 1 \) for \( i = 1, \ldots, 5 \)).

<table>
<thead>
<tr>
<th>( i )</th>
<th>( y_i )</th>
<th>( x_{i0} )</th>
<th>( x_{i1} )</th>
<th>( x_{i2} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>62</td>
<td>1</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>60</td>
<td>1</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>57</td>
<td>1</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>48</td>
<td>1</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>5</td>
<td>23</td>
<td>1</td>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>

Then the model can be represented as

\[
\begin{bmatrix}
62 \\
60 \\ 57 \\ 48 \\ 23
\end{bmatrix} = \begin{bmatrix}
1 & 2 & 6 \\
1 & 9 & 10 \\
1 & 6 & 4 \\
1 & 3 & 13 \\
1 & 5 & 2
\end{bmatrix} \begin{bmatrix}
\beta_1 \\
\beta_2 \\
\beta_3
\end{bmatrix} + \begin{bmatrix}
e_1 \\
e_2 \\
e_3 \\
e_4 \\
e_5
\end{bmatrix},
\]

where the covariance structure of the error term has the form \( \Sigma = \sigma^2 R \) for \( R = (1 - \rho)I_5 + \rho J_5 \), when \( \rho = \frac{1}{2} \) and \( \sigma^2 = 2 \), which satisfies the condition under which \( \Sigma^{-1} \) exists. Then \( \Sigma^{-1} = I_5 - \frac{1}{\rho} J_5 \).

Moreover, assume that we want to test the null hypothesis

\[
H_0: \begin{cases}
\beta_2 = 0.5 \\
2\beta_1 - \beta_2 + 3\beta_3 = 2 \\
\beta_1 = -1
\end{cases}
\]
In this approach we have

\[
H = \begin{bmatrix}
0 & 1 & 0 \\
2 & -1 & 3 \\
1 & 0 & 0
\end{bmatrix}, \quad h = \begin{bmatrix}
0.5 \\
2 \\
-1
\end{bmatrix}.
\]

Direct algebraic computations lead to

\[
G_1 = \begin{bmatrix}
2.64583 & -0.16667 & -0.0875 \\
-0.16667 & 0.03333 & 0.0000 \\
-0.08750 & 0.00000 & 0.0125
\end{bmatrix}, \quad G_2 = \begin{bmatrix}
83.7037 & 23.1481 & -40.1852 \\
23.1481 & 13.4259 & -24.9074 \\
-40.1852 & -24.9074 & 46.7593
\end{bmatrix}.
\]

Using (2.1) and (2.2) we obtain

\[
\hat{\beta} = \begin{bmatrix}
37.0 \\
0.5 \\
1.5
\end{bmatrix}, \quad \tilde{\beta} = \begin{bmatrix}
0.5 \\
2 \\
-1
\end{bmatrix} \quad \text{and} \quad \chi = 1203.3.
\]

Consider in this example \( n = 5, p = 3 \) and \( q = 3 \). Therefore using (2.6) we get \( \rho = \frac{1}{5} \). Then using (2.4), (2.5) and (2.7) we have

\[
\hat{\beta}^{PT} = \begin{bmatrix}
0.5 \\
2 \\
-1
\end{bmatrix} + \left[ 1 - I(1203 \leq \chi^2_3(\alpha)) \right] \begin{bmatrix}
38 \\
0 \\
0
\end{bmatrix},
\]

\[
\hat{\beta}^s + = \hat{\beta}^s = \begin{bmatrix}
0.5 \\
2 \\
-1
\end{bmatrix} + 0.9998 \begin{bmatrix}
38 \\
0 \\
0
\end{bmatrix}.
\]

In order to compare the risks of the above five estimators, suppose the weight matrix is given by

\[
W = \begin{bmatrix}
0 & -1 & -1 \\
1 & 0 & 1 \\
1 & 1 & 1
\end{bmatrix}.
\]

Then from (3.12) we get \( tr(A_{11}) = 0.0125 \).

Using (3.3) and (3.14) we have \( R(\hat{\beta}; \beta) = 0.0125 \) and \( R(\tilde{\beta}; \beta) = \theta \). Clearly \( \tilde{\beta} \).
performs better than \( \hat{\beta} \) whenever \( \theta < 0.0125 \). Using Lemma 2 from Appendix we can determine the risk functions for different values \( \alpha \) and \( \theta \). We will continue with large values of \( \theta \), to do better comparisons, which result in large unreasonable risks’ values. The results are given in Table 1.

**Table 1: Risks’ comparison**

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>( \theta )</th>
<th>( R(\hat{\beta}; \beta) )</th>
<th>( R(\tilde{\beta}; \beta) )</th>
<th>( R(\hat{\beta}^{PT}; \beta) )</th>
<th>( R(\hat{\beta}^S; \beta) )</th>
<th>( R(\hat{\beta}^{S+}; \beta) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>0</td>
<td>0.0125</td>
<td>0</td>
<td>0.0044</td>
<td>0.0113</td>
<td>0.0113</td>
</tr>
<tr>
<td>0.001</td>
<td>0.0125</td>
<td>0.001</td>
<td>0.0045</td>
<td>0.0110</td>
<td>0.0110</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>0.0125</td>
<td>0.1</td>
<td>0.0187</td>
<td>0.0221</td>
<td>0.0221</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.0125</td>
<td>1</td>
<td>0.3041</td>
<td>0.0694</td>
<td>0.0694</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.0125</td>
<td>10</td>
<td>37.1286</td>
<td>0.0995</td>
<td>0.0995</td>
<td></td>
</tr>
</tbody>
</table>

From the Table 1, it can be easily seen that

1. Under \( H_0 (\theta = 0) \), the domination order given in (4.6) satisfies.
2. For \( \theta \leq 0.001 \), the risks of PRSE and SE have decreasing trends and for \( \theta \geq 0.1 \) those change to increasing.
3. For \( \theta \geq 0.1 \), GLSE performs better than both RGLSE and PTE, and PTE performs better than RGLSE.

**Graphical Example** Some graphical perspectives of the risks of estimators \( \hat{\beta}, \tilde{\beta}, \hat{\beta}^{PT}, \hat{\beta}^S \) and \( \hat{\beta}^{S+} \) can be shown using approximations of (3.24) and (3.28).

In this approach, we use lemma 2 in Appendix to compute (3.34) and (3.35)

Then substituting suitable expression in (3.24) and (3.28), we compute underlying risks approximately using packages MATLAB release 7.2 and MAPLE release 9.5.

For special case \( n = 20, p = 5 \) and \( q = 3 \), when \( W = X'\Sigma^{-1}X \), the graphical displays are as follow (Because changing values \( \alpha \) in (3.18), does not clear graphically we use just \( \alpha = 0.3 \). Note that when \( \alpha \) increases \( R(\hat{\beta}^{PT}; \beta) \) decreases).

In Figure 1, the horizontal axis are the values of \( \theta \) and

\[ R_1 = R(\hat{\beta}; \beta), \quad R_2 = R(\tilde{\beta}; \beta), \quad R_3 = R(\hat{\beta}^{PT}; \beta), \quad R_4 = R(\hat{\beta}^S; \beta), \quad R_5 = R(\hat{\beta}^{S+}; \beta). \]
Figure 1: Risks Comparison
6 Appendix

Lemma 6.1 Assume the random variable $w$ is normally distributed with mean vector $\tau$ and covariance matrix $I_j$ and $A$ is any p.d. symmetric matrix. Also assume $\phi(.)$ is a Borel measurable function, then

$$E[\phi(w'w)w] = E[\phi(\chi_{j+2,r,r/2}^2)]\tau,$$
$$E[\phi(w'w)Aw] = E[\phi(\chi_{j+2,r,r/2}^2)]tr(A) + E[\phi(\chi_{j+4,r,r/2}^2)]\tau'Ar.$$

Proof. For the proof see Appendix B.2. in Judge and Bock [4].

Lemma 6.2 Let $p$ be an integer greater than $2m$ ($p > 2m$) then

$$E[(1 - \frac{\rho}{\chi_{q,t/2}^2})^2I(\chi_{q,t/2}^2 \leq \rho)] = \chi_{q,t/2}^2(\rho) + \Upsilon,$$

where

$$\Upsilon = \sum_{r=0}^{\infty} \frac{\rho^r(\rho - 2q - 4r + 8)}{r!(q + 2r - 2)(q + 2r - 4)}e^{-\theta/4}(theta/4)^r \chi_{q+2r,0}^2(\rho).$$

Proof. Using the series expansion for inverse non-central chi-square distribution (see Johnson and Kotz [3]), we have

$$E[\left( \frac{1}{\chi_{q,t}^2} \right)^m] = \sum_{r=0}^{\infty} \frac{e^{-\theta/2}(\theta/2)^r}{r!} E[\left( \frac{1}{\chi_{q+2r,0}^2} \right)^m]$$
$$= \sum_{r=0}^{\infty} \frac{e^{-\theta/2}(\theta/2)^r}{2^r r!} \times \frac{\Gamma(q/2 + r - m)}{\Gamma(q/2 + r)}.$$

Thus we can obtain

$$E[(\frac{1}{\chi_{q,t}^2})^mI(\chi_{q,t}^2 \leq \rho)] = \sum_{r=0}^{\infty} \frac{e^{-\theta/2}(\theta/2)^r}{r!} E[\left( \frac{1}{\chi_{q+2r,0}^2} \right)^mI(\chi_{q+2r,0}^2 \leq \rho)]$$
$$= \sum_{r=0}^{\infty} \frac{e^{-\theta/2}(\theta/2)^r}{2^r r!} \times \frac{\Gamma(q/2 + r - m)}{\Gamma(q/2 + r)} \times \chi_{q+2r,0}^2(\rho).$$
Therefore

\[
E[(1 - \frac{\rho}{\chi^2_{q,\theta}})^2 I(\chi^2_{q,\theta} \leq \rho)] = \chi^2_{q,\theta} + 2 \rho E[(\frac{1}{\chi^2_{q,\theta}})I(\chi^2_{q,\theta} \leq \rho)] + \rho^2 E[(\frac{1}{\chi^2_{q,\theta}})^2 I(\chi^2_{q,\theta} \leq \rho)]
\]

\[
= \chi^2_{q,\theta} + \sum_{r=0}^{\infty} \frac{e^{-\theta^2/2(\theta/2)^r}}{r!} \chi^2_{q+2r,\theta}(\rho)
\]

\[
= \chi^2_{q,\theta} + \sum_{r=0}^{\infty} \frac{\rho(\rho - 2q - 4r + 8)e^{-\theta^2/2(\theta/2)^r} \chi^2_{q+2r,\theta}(\rho)}{r!(q+2r-2)(q+2r-4)}.
\]

Acknowledgments

We would like to thank the referees for their comments and suggestions which substantially improved the presentation and the quality of the paper.

Partial support from the “Ordered and Spatial Data Center of Excellence” of Ferdowsi University of Mashhad, Iran is acknowledged.

References


