Mashhad J. Math. Sci., Vol.1(1)(2007) 83-96.

Shrinkage estimation of the regression parameters with multivariate normal errors

Mohammad Arashi*and Sayyed Mohammad Mehdi 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 β when the covariance structure in 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 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

$$y = X\beta + e \tag{1.1}$$

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 p-vector of regression coefficients and $e = (e_1, \cdots, e_n)'$ is the n-vector of

*Corresponding author

errors distributed as multivariate normal with location parameter zero and positive definite (p.d.) covariance matrix Σ, denoted by $e \sim N_n(0, \Sigma)$.
Then directly

$$y \sim N_n(X\beta, \Sigma). \tag{1.2}$$

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 β where $q \leq p$. These general restriction can be shown as

$$H\beta = h, \tag{1.3}$$

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 minimise 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 β when the p.d. covariance matrix Σ is known under the subspace restriction (1.3); and then obtain shrinkage estimators of β 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 σ^2 and may σ^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 Σ, the generalized least square (GLS) estimator of β is

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

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

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

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

$$\widetilde{\beta} = \hat{\beta} - G_1H'G_2(H\hat{\beta} - h). \tag{2.3}$$

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 $\omega = \{\beta : \beta \in \Re^p,\ H\beta = h,\ \Sigma > 0\}$ and $\Omega = \{\beta : \beta \in \Re^p,\ \Sigma > 0\}$, then the likelihood test statistic for underlying hypothesis is

$$\begin{aligned}\lambda &= \frac{\max_{\beta\in\omega} L(\beta,\Sigma)}{\max_{\beta\in\Omega} L(\beta,\Sigma)}\\ &= \frac{\exp\{\frac{-1}{2}\,[(y - X\widetilde{\beta})'\Sigma^{-1}(y - X\widetilde{\beta})]\}}{\exp\{\frac{-1}{2}\,[(y - X\hat{\beta})'\Sigma^{-1}(y - X\hat{\beta})]\}}\\ &= \exp\{\frac{-1}{2}\,[(H\hat{\beta} - h)'G_2(H\hat{\beta} - h)]\},\end{aligned}$$

which is a decreasing function with respect to (w.r.t.) $\chi = (H\hat{\beta} - h)'G_2(H\hat{\beta} - h)$.
Let $u = G_2^{1/2}(H\hat{\beta} - h)$; then using (1.2), $\chi = u'u$ has non-central chi-square distribution with q degrees of freedom and noncentrality parameter $\mu'\mu/2$, where $\mu = G_2^{1/2}(H\beta - h)$.
Bancroft [2] defined the preliminary test estimator (PTE) of β as a convex combination of $\hat{\beta}$ and $\widetilde{\beta}$ by

$$\hat{\beta}^{PT} = \widetilde{\beta} + [1 - I(\chi \leq \chi^2(\alpha))](\hat{\beta} - \widetilde{\beta}), \tag{2.4}$$

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

$$\hat{\beta}^{S} = \widetilde{\beta} + (1 - \rho\chi^{-1})(\hat{\beta} - \widetilde{\beta}), \tag{2.5}$$

where

$$\rho = \frac{(q-2)(n-p)}{q(n-p+2)} \quad \text{and} \quad q \geq 3. \tag{2.6}$$

The SE has the disadvantage that it has strange behavior for small values of χ. 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 β as

$$\begin{aligned}\hat{\beta}^{S+} &= \widetilde{\beta} + (1 - \rho\chi^{-1})I[\chi > \rho](\hat{\beta} - \widetilde{\beta})\\ &= \hat{\beta}^{S} - (1 - \rho\chi^{-1})I[\chi \leq \rho](\hat{\beta} - \widetilde{\beta}).\end{aligned} \tag{2.7}$$

Note that this estimator is a convex combination of $\hat{\beta}$ and $\widetilde{\beta}$.

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

Consider for a given non-singular matrix W, the *weighted quadratic error loss function* of the form

$$L(\beta^*;\beta) = (\beta^* - \beta)'W(\beta^* - \beta), \tag{3.1}$$

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

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

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

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

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

$$\begin{aligned} R(\widetilde{\beta};\beta) &= tr\{W[G_1(I_p - H'G_2HG_1) + \delta\delta']\} \\ &= tr(G_1W) - tr\{W[G_1(H'G_2HG_1)]\} + \delta'W\delta. \end{aligned} \tag{3.4}$$

Note that $R = G_1^{1/2}H'G_2HG_1^{1/2}$ is a symmetric idempotent matrix of rank $q \leq p$. Thus, there exist an orthogonal matrix Q $(Q'Q = I_p)$ (see Judge and Bock [4]) such that

$$QRQ' = \begin{bmatrix} I_q & 0 \\ 0 & 0 \end{bmatrix}, \tag{3.5}$$

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

The matrices A_{11} and A_{22} are of order q and $p - q$, respectively.
Define random variable

$$w = QG_1^{-1/2}\hat{\beta} - QG_1^{1/2}H'G_2h, \tag{3.7}$$

then

$$w \sim N_q(\eta, I_p). \tag{3.8}$$

Also

$$\eta \;=\; QG_1^{-1/2}\beta - QG_1^{1/2}H'G_2h. \tag{3.9}$$

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

$$\hat{\beta} - \beta \;=\; G_1^{1/2}Q'(w-\eta). \tag{3.10}$$

Using (3.7) we can obtain

$$\chi = w_1'w_1,\ \theta = \eta_1'\eta_1 = (H\beta - h)'G_2(H\beta - h). \tag{3.11}$$

Now, we can write

$$\begin{aligned} tr\{W[G_1H'G_2HG_1]\} &= tr\{QG_1^{1/2}WG_1^{1/2}Q'QRQ'\} \\ &= tr\left\{\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}\end{bmatrix}\begin{bmatrix} I_q & 0 \\ 0 & 0\end{bmatrix}\right\} \\ &= tr(A_{11}). \end{aligned} \tag{3.12}$$

Using (3.11) we have

$$\begin{aligned} \delta'W\delta &= (H\beta - h)'G_2HG_1WG_1H'G_2(H\beta - h) \\ &= \eta_1'A_{11}\eta_1. \end{aligned} \tag{3.13}$$

Therefore, we obtain

$$R(\widetilde{\beta};\beta) \;=\; tr(G_1W) - tr(A_{11}) + \eta_1'A_{11}\eta_1. \tag{3.14}$$

Using (2.5)

$$\begin{aligned} R(\hat{\beta}^{PT};\beta) &= E[(\hat{\beta}^{PT} - \beta)'W(\hat{\beta}^{PT} - \beta)] \\ &= E\{[(\hat{\beta} - \beta) - I(\chi \le \chi^2_\alpha)(\hat{\beta} - \widetilde{\beta})]'W[(\hat{\beta} - \beta) \\ &\quad - I(\chi \le \chi^2_\alpha)(\hat{\beta} - \widetilde{\beta})]\} \\ &= E[(\hat{\beta} - \beta)'(\hat{\beta} - \beta)] - 2E[I(\chi \le \chi^2_\alpha)(\hat{\beta} - \beta)'W(\hat{\beta} - \widetilde{\beta})] \\ &\quad + E[I(\chi \le \chi^2_\alpha)(\hat{\beta} - \widetilde{\beta})'W(\hat{\beta} - \widetilde{\beta})]. \end{aligned} \tag{3.15}$$

Using (3.7)-(3.11) and (3.15)

$$\begin{aligned} R(\hat{\beta}^{PT};\beta) &= tr(G_1W) - E[w_1'A_{11}w_1I(\chi \le \chi^2_\alpha)] \\ &\quad - 2E[w_2'A_{21}w_1I(\chi \le \chi^2_\alpha)] + 2\eta_1'A_{11}E[w_1I(\chi \le \chi^2_\alpha)] \\ &\quad + 2\eta_2'A_{21}E[w_1I(\chi \le \chi^2_\alpha)], \end{aligned} \tag{3.16}$$

because w_1 and w_2 are independent

$$E[w_2'A_{21}w_1I(\chi \le \chi^2_\alpha)] \;=\; \eta_2'A_{21}E[w_1I(\chi \le \chi^2_\alpha)], \tag{3.17}$$

using Lemma1 in Appendix with $\phi(\chi)$ as indicator function of χ, we get

$$\begin{aligned} R(\hat{\beta}^{PT};\beta) &= tr(G_1W) - \chi^2_{q+2,\theta}(\alpha)tr(A_{11}) \\ &\quad +[2\chi^2_{q+2,\theta}(\alpha) - \chi^2_{q+4,\theta}(\alpha)]\eta_1' A_{11}\eta_1. \end{aligned} \tag{3.18}$$

Using (2.6) and (2.7)

$$\begin{aligned} R(\hat{\beta}^{S};\beta) &= E[(\hat{\beta}^S - \beta)'W(\hat{\beta}^S - \beta)] \\ &= E\{[(\hat{\beta} - \beta) - \rho\chi^{-1}(\hat{\beta} - \widetilde{\beta})]'W[(\hat{\beta} - \beta) - \rho\chi^{-1}(\hat{\beta} - \widetilde{\beta})] \\ &= E[(\hat{\beta} - \beta)'W(\hat{\beta} - \beta)] - 2\rho E[\chi^{-1}(\hat{\beta} - \beta)'W(\hat{\beta} - \widetilde{\beta})] \\ &\quad +\rho^2 E[\chi^{-2}(\hat{\beta} - \widetilde{\beta})'W(\hat{\beta} - \widetilde{\beta})], \end{aligned} \tag{3.19}$$

using (3.7)-(3.11) and (3.15)

$$\begin{aligned} R(\hat{\beta}^{S};\beta) &= tr(G_1W) - 2\rho E[\chi^{-1}(w_1' A_{11}w_1 - \eta_1' A_{11}w_1 + w_2' A_{21}w_1 \\ &\quad -\eta_2' A_{21}w_1)] + \rho^2 E[\chi^{-2}(w_1' A_{11}w_1)]. \end{aligned} \tag{3.20}$$

Using Lemma1 in Appendix for $\phi(\chi) = \chi^{-1}$, we have

$$E[\chi^{-1}\eta_1' A_{11}w_1] = \eta_1' A_{11}\eta_1 E\left[\frac{1}{\chi^2_{q+2,\theta}}\right], \tag{3.21}$$

$$E[\chi^{-1}w_1' A_{11}w_1] = E\left[\frac{1}{\chi^2_{q+2,\theta}}\right] tr(A_{11}) + E\left[\frac{1}{\chi^2_{q+4,\theta}}\right]\eta_1' A_{11}\eta_1. \tag{3.22}$$

Using Lemma1 in Appendix for $\phi(\chi) = \chi^{-2}$, we have

$$E[\chi^{-2}w_1' A_{11}w_1] = E\left[\frac{1}{\chi^2_{q+2,\theta}}\right]^2 tr(A_{11}) + E\left[\frac{1}{\chi^2_{q+4,\theta}}\right]^2 \eta_1' A_{11}\eta_1. \tag{3.23}$$

Using (3.21)-(3.23) one can obtain

$$\begin{aligned} R(\hat{\beta}^{S};\beta) &= tr(G_1W) - \rho\left\{2E\left[\frac{1}{\chi^2_{q+2,\theta}}\right] - \rho E\left[\frac{1}{\chi^2_{q+2,\theta}}\right]^2\right\}tr(A_{11}) \\ &\quad +\rho\left\{2E\left[\frac{1}{\chi^2_{q+2,\theta}}\right] - 2E\left[\frac{1}{\chi^2_{q+4,\theta}}\right]\right. \\ &\quad \left. +\rho E\left[\frac{1}{\chi^2_{q+4,\theta}}\right]^2\right\}\eta_1' A_{11}\eta_1. \end{aligned} \tag{3.24}$$

Finally the risk of PRSE is given by

$$\begin{aligned} R(\hat{\beta}^{S+};\beta) &= E[(\hat{\beta}^{S+} - \beta)'W(\hat{\beta}^{S+} - \beta)] \\ &= E\{[(\hat{\beta}^S - \beta) - (1 - \rho\chi^{-1})I(\chi \le \rho)(\hat{\beta} - \widetilde{\beta})]'W[(\hat{\beta}^S - \beta) \\ &\quad -(1 - \rho\chi^{-1})I(\chi \le \rho)(\hat{\beta} - \widetilde{\beta})]\} \\ &= R(\hat{\beta}^S;\beta) + E[(1 - \rho\chi^{-1})^2 I(\chi \le \rho)(\hat{\beta} - \widetilde{\beta})'W(\hat{\beta} - \widetilde{\beta})] \\ &\quad -2E[(\hat{\beta}^S - \beta)'W(1 - \rho\chi^{-1})I(\chi \le \rho)(\hat{\beta} - \widetilde{\beta})]. \end{aligned} \tag{3.25}$$

But using (2.6)

$$\begin{aligned}
& E[(\hat{\beta}^S-\beta)'W(1-\rho\chi^{-1})I(\chi\leq\rho)(\hat{\beta}-\widetilde{\beta})] \\
= \; & E[(\widetilde{\beta}-\beta)+(1-\rho\chi^{-1})(\hat{\beta}-\widetilde{\beta})]'W(1-\rho\chi^{-1})I(\chi\leq\rho)(\hat{\beta}-\widetilde{\beta})] \\
= \; & E\{(\widetilde{\beta}-\beta)'W[(1-\rho\chi^{-1})I(\chi\leq\rho)(\hat{\beta}-\widetilde{\beta})]\} \\
& +E[(1-\rho\chi^{-1})^2 I(\chi\leq\rho)(\hat{\beta}-\widetilde{\beta})'W(\hat{\beta}-\widetilde{\beta})]. \qquad (3.26)
\end{aligned}$$

Thus we can obtain

$$\begin{aligned}
R(\hat{\beta}^{S+};\beta) \;=\; & R(\hat{\beta}^S;\beta)-E[(1-\rho\chi^{-1})^2 I(\chi\leq\rho)(\hat{\beta}-\widetilde{\beta})'W(\hat{\beta}-\widetilde{\beta})] \\
& -2E\{(\widetilde{\beta}-\beta)'W[(1-\rho\chi^{-1})I(\chi\leq\rho)(\hat{\beta}-\widetilde{\beta})]\}. \qquad (3.27)
\end{aligned}$$

Using (3.21)-(3.23), and Lemma1 in Appendix for $\phi(\chi)=(1-\rho\chi^{-1})^i I(\chi\leq\rho)$, $(i=1,2)$ we get

$$\begin{aligned}
R(\hat{\beta}^{S+};\beta) \;=\; & R(\hat{\beta}^S;\beta)-E\left[\left(1-\frac{\rho}{\chi^2_{q+2,\theta}}\right)^2 I(\chi^2_{q+2,\theta}\leq\rho)\right]tr(A_{11}) \\
& +E\left[\left(1-\frac{\rho}{\chi^2_{q+4,\theta}}\right)^2 I(\chi^2_{q+4,\theta}\leq\rho)\right]\eta_1'A_{11}\eta_1 \\
& -2E\left[\left(\frac{\rho}{\chi^2_{q+2,\theta}}-1\right) I(\chi^2_{q+2,\theta}\leq\rho)\right]\eta_1'A_{11}\eta_1 \qquad (3.28)
\end{aligned}$$

4 Comparison

Providing risk analysis of the underlying estimators with the weight matrix W, we have (see e.g. Searle [11])

$$\theta ch_1(A_{11})\leq\eta_1'A_{11}\eta_1\leq\theta ch_q(A_{11}), \qquad (4.1)$$

where $ch_1(A_{11})$ and $ch_q(A_{11})$ are the minimum and maximum eigenvalue of A_{11} respectively. Then by (3.3) and (3.14) one may easily seen

$$R(\hat{\beta};\beta)-tr(A_{11})+\theta ch_1(A_{11})\leq R(\widetilde{\beta};\beta)\leq R(\hat{\beta};\beta)-tr(A_{11})+\theta ch_q(A_{11}).$$

By (3.11) and (3.30), under the null hypothesis $H_0: H\beta=h$, we conclude

$$R(\widetilde{\beta};\beta) \;\leq\; R(\hat{\beta};\beta).$$

Generally by (3.30), $\widetilde{\beta}$ performs better then $\hat{\beta}$ whenever

$$\begin{aligned}
\theta \;\leq\; & \frac{tr(A_{11})}{ch_q(A_{11})} \\
= \; & \frac{\sum_{i=1}^q ch_i(A_{11})}{ch_q(A_{11})} \\
\leq \; & q.
\end{aligned}$$

Using (3.3) and (3.18) we have

$$R(\hat{\beta}^{PT};\beta) - R(\hat{\beta};\beta) = [2\chi^2_{q+2,\theta}(\alpha) - \chi^2_{q+4,\theta}(\alpha)]\eta_1' A_{11}\eta_1 - \chi^2_{q+2,\theta}(\alpha) tr(A_{11}). \quad (4.2)$$

Therefore $\hat{\beta}^{PT}$ performs better than $\hat{\beta}$ whenever

$$\theta \leq \frac{tr(A_{11})}{ch_q(A_{11})} \times \frac{\chi^2_{q+2,\theta}(\alpha)}{[2\chi^2_{q+2,\theta}(\alpha) - \chi^2_{q+4,\theta}(\alpha)]}. \quad (4.3)$$

For $W = X'\Sigma^{-1}X$, because $tr(A_{11}) = q$, (3.33) satisfies.
Also under the null hypothesis H_0, since (3.32) is negative for all α, $\hat{\beta}^{PT}$ performs better than $\hat{\beta}$.
Using (3.14) and (3.18), using the risks difference we can conclude that $\hat{\beta}^{PT}$ performs better than $\widetilde{\beta}$ whenever

$$\theta \geq \frac{[1 - \chi^2_{q+2,\theta}(\alpha)]tr(A_{11})}{[1 - 2\chi^2_{q+2,\theta}(\alpha) + \chi^2_{q+4,\theta/2}(\alpha)]ch_q(A_{11})}. \quad (4.4)$$

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

$$\widetilde{\beta} \succ \hat{\beta}^{PT} \succ \hat{\beta},$$

where the notation $\succ$ means dominate.

Under the null hypothesis,

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

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

Under the null hypothesis H_0, we have

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

where $f(n,q,p) = \frac{\rho^2 - 2\rho(q-2) + q(q-2)}{q(q-2)}$.
The function $f(n,q,p)$ is positive for $q \geq 3$. Thus $R(\hat{\beta}^S;\beta) > R(\widetilde{\beta};\beta)$. However, as η_1 moves away from 0, $\eta_1' A_{11}\eta_1$ increases and the risk of $\widetilde{\beta}$ becomes unbounded while the risk of $\hat{\beta}^S$ remains below the risk of $\hat{\beta}$; thus $\hat{\beta}^S$ dominates $\widetilde{\beta}$ outside an interval around the origin.
Comparing $\hat{\beta}^S$ and $\hat{\beta}^{PT}$, under H_0, we get

$$\begin{aligned} R(\hat{\beta}^S;\beta) &= R(\hat{\beta}^{PT}) + \left[\chi^2_{q+2,0}(\alpha) - 2\rho E[\frac{1}{\chi^2_{q+2,0}}] + \rho^2 E[\frac{1}{\chi^2_{q+2,0}}]^2\right] tr(A_{11}) \\ &= R(\hat{\beta}^{PT}) + \left[\chi^2_{q+2,0}(\alpha) - \frac{2\rho}{q} + \frac{\rho^2}{q(q-2)}\right] tr(A_{11}) \\ &\geq R(\hat{\beta}^{PT}), \end{aligned}$$

for all α such that $l = \chi^2_{q+2,0}(\alpha) - \frac{2\rho}{q} + \frac{\rho^2}{q(q-2)} \geq 0$ and $R(\hat{\beta}^S;\beta) \leq R(\hat{\beta}^{PT})$ for all α such that $l \leq 0$.

Because w_1 is independent of w_2, we get

$$\begin{aligned} R(\hat{\beta}^{S+};\beta) - R(\hat{\beta}^S;\beta) &= -E[(1-\rho\chi^{-1})^2 I(\chi \leq \rho) w_1' A_{11} w_1] \\ &\quad -2E[(1-\rho\chi^{-1}) I(\chi \leq \rho)(w_1' A_{11} w_1 - \eta_1' A_{11} w_1)]. \end{aligned} \tag{4.5}$$

Note that for such θ under which $\chi^2_{q+2,\theta} \leq \rho$ we have

$$E[(1 - \frac{\rho}{\chi^2_{q+2,\theta}}) I(\chi^2_{q+2,\theta}) \leq \rho)] \leq 0.$$

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 β, $\hat{\beta}^{S+} \succ \hat{\beta}^S$ and under H_0, $\widetilde{\beta} \succ \hat{\beta}^{S+}$.
However, as η_1 moves away from 0, $\eta_1' A_{11} \eta_1$ increases and the risk of $\widetilde{\beta}$ becomes unbounded while the risk of $\hat{\beta}^{S+}$ remains below the risk of $\hat{\beta}$; thus $\hat{\beta}^{S+}$ dominates $\widetilde{\beta}$ outside an interval around the origin.

Under the conditions are given above, it can be found that the dominance order of five estimators of β can be categorized in the two following order.

1. $$\widetilde{\beta} \succ \hat{\beta}^{PT} \succ \hat{\beta}^{S+} \succ \hat{\beta}^S \succ \hat{\beta} \tag{4.6}$$

and

2. $$\widetilde{\beta} \succ \hat{\beta}^{S+} \succ \hat{\beta}^S \succ \hat{\beta}^{PT} \succ \hat{\beta} \tag{4.7}$$

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 previous section, we accomplish with a numerical example from Searle [11]. Suppose we have the following five sets of observations (including $x_{i0} = 1$ for $i = 1, \cdots, 5$).

i	y_i	x_{i0}	x_{i1}	x_{i2}
1	62	1	2	6
2	60	1	9	10
3	57	1	6	4
4	48	1	3	13
5	23	1	5	2

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 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 Σ^{-1} exists. Then $\Sigma^{-1} = I_5 - \frac{1}{6}J_5$.

Moreover assume 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 can obtain

$$\hat{\beta} = \begin{bmatrix} 37.0 \\ 0.5 \\ 1.5 \end{bmatrix}, \ \widetilde{\beta} = \begin{bmatrix} 0.5 \\ 2 \\ -1 \end{bmatrix} \text{ and } \chi = 1203.\bar{3}$$

Consider in this example $n = 5$, $p = 3$ and $q = 3$. Therefore using (2.6) we get $\rho = \frac{1}{6}$. Then using (2.4), (2.5) and (2.7) we have

$$\hat{\beta}^{PT} = \begin{bmatrix} 0.5 \\ 2 \\ -1 \end{bmatrix} + [1 - I(1203 \leq \chi_3^2(\alpha))] \begin{bmatrix} 38 \\ 0 \\ 0 \end{bmatrix},$$

$$\hat{\beta}^{S+} = \hat{\beta}^{S} = \begin{bmatrix} 0.5 \\ 2 \\ -1 \end{bmatrix} + 0.9998 \begin{bmatrix} 38 \\ 0 \\ 0 \end{bmatrix},$$

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 $R(\widetilde{\beta};\beta) = \theta$. Clearly $\widetilde{\beta}$ performs better than $\hat{\beta}$ whenever $\theta < 0.0125$. Using Lemma2 from Appendix we can determine the risk functions for different values α and θ. We will continue with large values of θ, to do better comparisons, which result in large unreasonable risks' values. The results are given in Table1.

Table1: Risks' comparison

α	θ	$R(\hat{\beta};\beta)$	$R(\widetilde{\beta};\beta)$	$R(\hat{\beta}^{PT};\beta)$	$R(\hat{\beta}^{S};\beta)$	$R(\hat{\beta}^{S+};\beta)$
0.05	0	0.0125	0	0.0044	0.0113	0.0113
	0.001	0.0125	0.001	0.0045	0.0110	0.0110
	0.1	0.0125	0.1	0.0187	0.0221	0.0221
	1	0.0125	1	0.3041	0.0694	0.0694
	10	0.0125	10	37.1286	0.0995	0.0995

From the Table1, 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 $\hat{\beta}$, $\widetilde{\beta}$, $\hat{\beta}^{PT}$, $\hat{\beta}^{S}$ 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 (3.285), 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 α in (3.18), does not clear graphically we use just $\alpha = 0.3$. Note that when α increases $R(\hat{\beta}^{PT};\beta)$ decreases).
In Figure5.1, horizontal axis is the values of θ and
$R1 = R(\hat{\beta};\beta)$, $R2 = R(\widetilde{\beta};\beta)$, $R3 = R(\hat{\beta}^{PT};\beta)$, $R4 = R(\hat{\beta}^{S};\beta)$, $R5 = R(\hat{\beta}^{S+};\beta)$.

6 Appendix

Lemma1. Assume the r.v. w is normally distributed with mean vector τ and covariance matrix I_j and A is any p.d. symmetric matrix. Also assume $\phi(.)$ is a Borel measurable function, then

$$\begin{aligned} E[\phi(w'w)w] &= E[\phi(\chi^2_{j+2,\tau'\tau/2})]\tau, \\ E[\phi(w'w)w'Aw] &= E[\phi(\chi^2_{j+2,\tau'\tau/2})]tr(A) + E[\phi(\chi^2_{j+4,\tau'\tau/2})]\tau' A\tau. \end{aligned}$$

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

Lemma2. Let p is an integer greater that $2m$ ($p > 2m$) then

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

where

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

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

$$\begin{aligned} E[(\frac{1}{\chi^2_{q,\theta}})^m] &= \sum_{r=0}^{\infty} \frac{e^{-\theta/2}(\theta/2)^r}{r!} E[(\frac{1}{\chi^2_{q+2r,0}})^m] \\ &= \sum_{r=0}^{\infty} \frac{e^{-\theta/2}(\theta/2)^r}{2^m r!} \times \frac{\Gamma(q/2+r-m)}{\Gamma(q/2+r)}. \end{aligned}$$

Thus we can obtain

$$\begin{aligned} E[(\frac{1}{\chi^2_{q,\theta}})^m I(\chi^2_{q,\theta} \le \rho)] &= \sum_{r=0}^{\infty} \frac{e^{-\theta/2}(\theta/2)^r}{r!} E[(\frac{1}{\chi^2_{q+2r,0}})^m I(\chi^2_{q+2r,0} \le \rho)] \\ &= \sum_{r=0}^{\infty} \frac{e^{-\theta/2}(\theta/2)^r}{2^m r!} \times \frac{\Gamma(q/2+r-m)}{\Gamma(q/2+r)} \times \chi^2_{q+2r,0}(\rho) \end{aligned}$$

Therefore

$$\begin{aligned} E[(1-\frac{\rho}{\chi^2_{q,\theta/2}})^2 I(\chi^2_{q,\theta/2} \le \rho)] &= \chi^2_{q,\theta/2}(\rho) - 2\rho E[(\frac{1}{\chi^2_{q,\theta}}) I(\chi^2_{q,\theta} \le \rho)] \\ &\quad + \rho^2 E[(\frac{1}{\chi^2_{q,\theta}})^2 I(\chi^2_{q,\theta} \le \rho)] \\ &= \chi^2_{q,\theta/2}(\rho) + \sum_{r=0}^{\infty} \frac{e^{-\theta/2}(\theta/2)^r}{r!\Gamma(q/2+r)} \times \chi^2_{q+2r,0}(\rho) \\ &\quad \times \left[\frac{\rho^2\Gamma(q/2+r-2)}{4} - \rho\Gamma(q/2+r-1)\right] \\ &= \chi^2_{q,\theta/2}(\rho) + \sum_{r=0}^{\infty} \frac{\rho[\rho - 2q - 4r + 8)]e^{-\theta/2}(\theta/2)^r \chi^2_{q+2r,0}(\rho)}{r!(q+2r-2)(q+2r-4)}. \end{aligned}$$

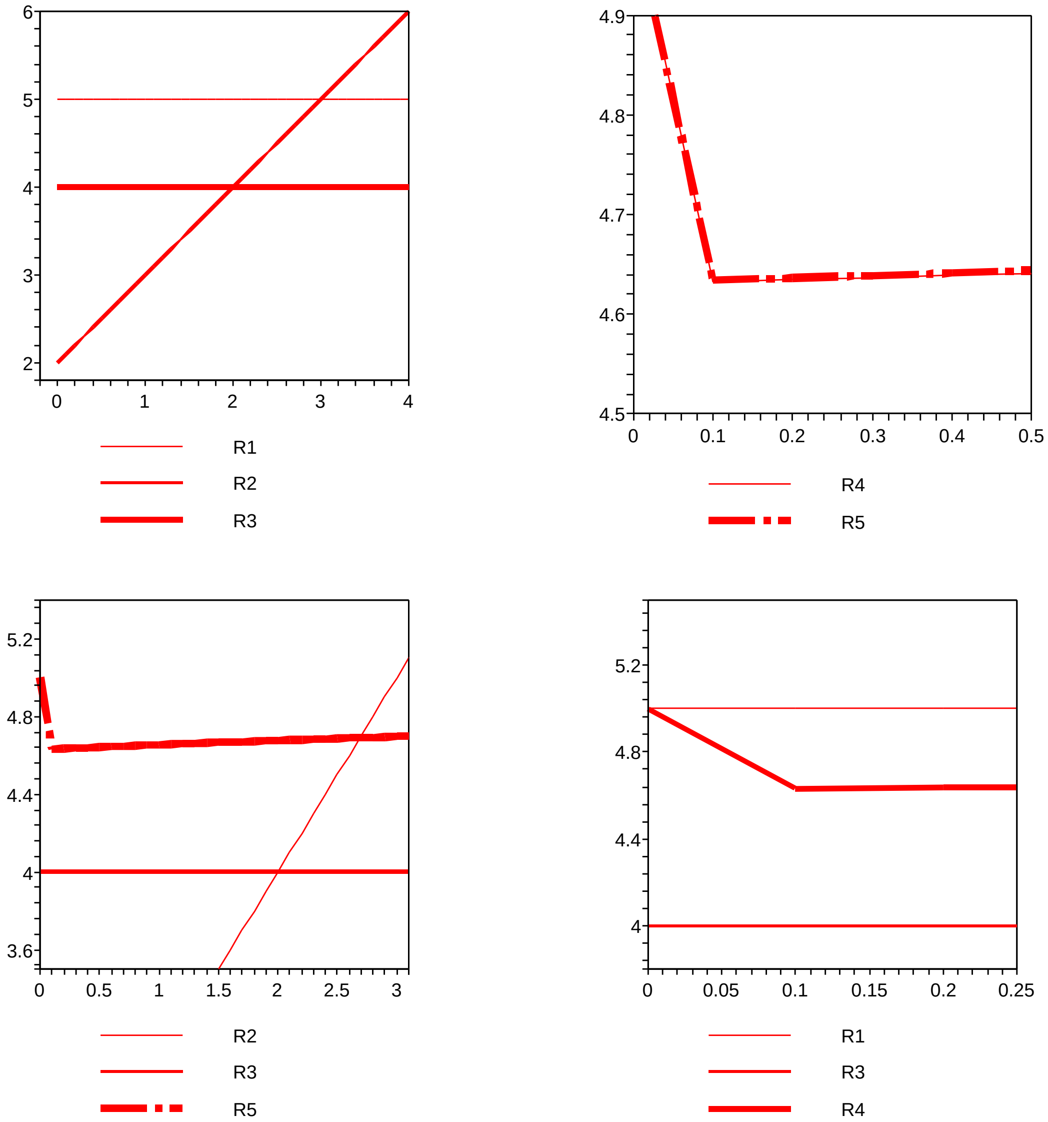

Figure 1: **Risks Comparison**

Acknowledgments

We would like to thank the two referees, editor in chief, Professor M. R. R. Moghaddam and Dr. Jabbari 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

[1] Anderson, T. W. (2003). *An introduction to multivariate statistical analysis*, 3rd ed., John Wiley and Sons, New York.

[2] Bancroft, T. A. (1944). On biases in estimation due to the use of preliminary test of significance, *Annals of Math. Stat.*, 15, 195–204.

[3] Johnson, N. L. and Kotz, S. (1970). *Continuous univariate distributions-2*, John Wiley, New York.

[4] Judge, W. and Bock, M. E. (1978). *The statistical implication of pre-test and Stein-rule estimators in Econometrics*, North-Holland, New York.

[5] Khan, Sh. (2000). Improved estimation of the mean vector for student-t model, *Commun. Statist.-Theory Meth.*, 29(3), 507–527.

[6] Khan, S. (2005). Estimation of parameters of the simple multivariate linear model with Student-t errors. *J. Statist. Res.*, 39 (2), 79-94.

[7] Kuan, Ch. M. (2004). *Statistics: Concepts and Methods*, 2nd edition, Hua-Tai, Taipei.

[8] Ravishanker, N. and Dey, D. K. (2002). *A first course in linear model theory*, Chapman and Hall/CRC.

[9] Saleh, A. K. Md. E. and Han, C. P. (1990). Shrinkage estimation in regression analysis, *Estadistica*, Vol. 42, 40–63.

[10] Saleh, A. K. Md. E. (2006), *Theory of Preliminary Test and Stein-type Estimation with Applications*, John Wiley, New York.

[11] Searle, S. R. (1982), *Matrix Algebra Useful for Statistics*, John Wiley, New York.

[12] Srivastava, M. S. and Ehsanes Saleh, A. K. Md. (2005). Estimation of the mean vector of a multivariate normal distribution: subspace hypothesis, *J. Multivariate Analysis*, 96, 55–72.

[13] Tabatabaey, S. M. M. (1995). *Preliminary test approach estimation: regression model with spherically symmetric errors*, Ph.D. Thesis, Carleton University, Canada.

e-mail: m-arashi-stat@yahoo.com
tabatab@math.um.ac.ir