Journal of Modern Applied Statistical Methods

Volume 15 | Issue 1 Article 34

5-1-2016

Principal Component Preliminary Test Estimator in the Linear Regression Model

Sivarajah Arumairajan
Department of Mathematics and Statistics, University of Jaffna, Sri Lanka, arumais@gmail.com

Pushpakanthie Wijekoon
Department of Statistics & Computer Science, Faculty of Science, University of Peradeniya, Sri Lanka

Principal Component Preliminary Test Estimator in the Linear Regression Model

Cover Page Footnote
We thank the Postgraduate Institute of Science, University of Peradeniya, Sri Lanka for providing all facilities to do this research.

Journal of Modern Applied Statistical Methods
May 2016, Vol. 15, No. 1, 690-710.

ISSN 1538 − 9472

Principal Component Preliminary Test Estimator in the Linear Regression Model

Sivarajah Arumairajan
University of Jaffna
Sri Lanka

Pushpakanthie Wijekoon
University of Peradeniya
Sri Lanka

A Preliminary Test Estimator is introduced based on Principal Component Regression Estimator defined in the linear regression model when the stochastic restrictions are available in addition to the sample information, and when the explanatory variables are multicollinear. It is further developed as a large sample preliminary test estimator by using Wald (WA), Likelihood Ratio (LR), and Lagrangian Multiplier (LM) tests. Stochastic properties of this estimator based on F test as well as WA, LR, and LM tests are derived, and the performance of the estimator is compared using WA, LR, and LM tests with respect to Mean Square Error Matrix (MSEM). A Monte Carlo simulation is carried out to illustrate the theoretical findings.

Keywords: Principal Component Regression, Preliminary Test Estimator, Wald Test, Likelihood Ratio Test, Lagrangian Multiplier Test, Mean Square Error Matrix

Introduction

Instead of using the Ordinary Least Square Estimator (OLSE), some biased estimation procedures were developed in the literature to combat the multicollinearity problem in the linear regression model. Some of these are namely the Principal Component Regression Estimator (PCRE) (Massy, 1965), Ridge Estimator (RE) (Hoerl & Kennard, 1970) and Liu Estimator (LE) (Liu, 1993). Another way of solving the multicollinearity problem is to consider parameter estimation with some additional information on the unknown parameters such as the exact or stochastic restrictions. By adding exact restrictions to a sample model, the resulting Restricted Least Squares Estimator (RLSE) might again be better in the mean square error sense than the OLSE. By grafting the ridge regression philosophy into the RLSE, the Restricted Ridge

Sivarajah Arumairajan is a lecturer in the Department of Mathematics and Statistics. Email at: arumais@gmail.com. Pushpakanthie Wijekoon is a Professor in the Department of Statistics & Computer Science.

Estimator (RRE) has been proposed by Sarkar (1992). As done by Liu (1993), Kaçiranlar, Sakallioğlu, Akdeniz, Styan, & Werner (1999) proposed a biased estimator called Restricted Liu Estimator (RLE) by combining exact prior information with the sample information, and studied its properties. In the presence of stochastic restrictions, Theil and Goldberger (1961) proposed the Mixed Estimator (ME). By replacing OLSE by ME in the RE and LE respectively, the Stochastic Mixed Ridge Estimator (SMRE) (Li & Yang, 2010), and Stochastic Restricted Liu Estimator (SRLE) (Hubert & Wijekoon, 2006) were introduced.

When different estimators are available, preliminary test estimation procedure is adopted to select a suitable estimator. The preliminary test approach was first proposed by Bancroft (1944), and then has been studied by many researchers, such as Judge and Bock (1978), Wijekoon (1990), and Saleh and Kibria (1993). Later Golam Kibria and Saleh (2003) have discussed the performance of preliminary test ridge estimators based on large sample tests; WA (Wald, 1943), LR (Atchison & Silvey, 1958), and LM (Rao, 1947). Recently Arumairajan and Wijekoon (2013) proposed the Preliminary Test Stochastic Restricted Liu Estimator (PTSRLE) by combining Liu Estimator and Stochastic Restricted Liu Estimator. A Preliminary Test Principal Component Regression Estimator (PTPCRE) is proposed by combining the idea of Preliminary Test Estimator and Principal Component Regression Estimator.

Model Specification and Estimation

Consider the multiple linear model

$$y = X\beta + \varepsilon, \varepsilon \sim N\left(0, \sigma^2 I\right) \tag{1}$$

where y is an $n \times 1$ observable random vector, X is an $n \times p$ known design matrix of rank p, β is a $p \times 1$ vector of unknown parameters and ε is an $n \times 1$ vector of disturbances.

The Ordinary Least Squares Estimator (OLSE) for the model (1) is given as

$$\hat{\beta} = S^{-1} X' y \tag{2}$$

where $S = X'X$.

Consider the transformation for model (1):

$$y = XTT'\beta + \varepsilon = Z\alpha + \varepsilon \tag{3}$$

where $Z = XT$, $\alpha = T'\beta$ and $T = (t_1, t_2, \ldots, t_p) = (T_k, T_{p-k})$ is a $p \times p$ orthogonal matrix such that

$$(T_k, T_{p-k})' X'X (T_k, T_{p-k}) = \Lambda = \begin{pmatrix} \Lambda_k & 0 \\ 0 & \Lambda_{p-k} \end{pmatrix}$$

where $0 < k \leq p,$ $T_k = (t_1, t_2, ..., t_k)$, $T_{p-k} = (t_{k+1}, t_{k+2}, \ldots, t_p)$, $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_p)$, $\Lambda_k = diag(\lambda_1, \lambda_2, \ldots, \lambda_k)$, $\Lambda_{p-k} = diag(\lambda_{k+1}, \lambda_{k+2}, \ldots, \lambda_p)$, and $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p > 0$ are the eigenvalues of $X'X$. Note that $Z = XT = (z_1, z_2, \ldots, z_p) = (Z_k, Z_{k-p})$ is the $n \times p$ matrix of the principal components, where $z_i = Xt_i$ is the i^{th} principal component. When Z_{p-k} contains principal components corresponding to near zero eigenvalues, Z can be separated as Z_k and Z_{p-k}, where Z_{p-k} is to be deleted. Rewrite the model (3) as

$$y = XTT'\beta = XT_kT_k'\beta + XT_{p\text{-}k}T_{p\text{-}k}'\beta + \varepsilon = Z_k\alpha_k + Z_{p\text{-}k}\alpha_{p\text{-}k} + \varepsilon. \tag{4}$$

By omitting Z_{p-k}, the OLSE of α_k is obtained, and $\hat{\alpha}_k = (Z_k'Z_k)^{-1} Z_k'y$. Then PCRE of β is

$$\hat{\beta}_{PCRE} = T_k (T_k'ST_k)^{-1} T_k'X'y \tag{5}$$

Xu and Yang (2011) showed that the PCRE estimator could be rewritten as follows.

$$\hat{\beta}_{PCRE} = T_kT_k'\hat{\beta} = L_k\hat{\beta} \tag{6}$$

where $L_k = T_kT_k'$.

The RE was proposed by Hoerl and Kennard (1970) as

$$\hat{\beta}(k) = W\hat{\beta} \tag{7}$$

where $W = \left(I + kS^{-1}\right)^{-1}$ for $k \geq 0$.

The LE was introduced by Liu (1993) as

$$\hat{\beta}(d) = F_d \hat{\beta} \tag{8}$$

where $F_d = \left(S + I\right)^{-1}\left(S + dI\right)$ for $0 < d < 1$.

In addition to sample model (1), suppose some prior information was given about β in the form of a set of m independent stochastic linear restrictions as follows;

$$r = R\beta + \delta + \upsilon, \quad \upsilon \sim N\left(0, \sigma^2\Omega\right) \tag{9}$$

where r is an $m \times 1$ stochastic known vector, R is a $m \times p$ of full row rank $m \leq p$ with known elements, δ is non zero $m \times 1$ unknown vector, υ is an $m \times 1$ random vector of disturbances, and Ω is assumed to be known and positive definite. Further it is assumed that υ is stochastically independent of ε i.e. $E\left(\varepsilon\upsilon'\right) = 0$.

The Ordinary Least Squares Estimator (OLSE) for the model (1) and the Mixed Estimator (ME) (Theil & Goldberger, 1961) due to a stochastic prior restriction (9) are given by

$$\hat{\beta} = S^{-1}X'\,Y \text{ and } \hat{\beta}_m = \hat{\beta} + S^{-1}R'\left(\Omega + RS^{1}R'\right)^{-1}\left(r - R\hat{\beta}\right) \tag{10}$$

respectively. The expectation vector, and the mean square error matrix of $\hat{\beta}$ are given as $E\left(\hat{\beta}\right) = \beta$ and $MSE\left(\hat{\beta}\right) = \sigma^2 S^{-1}$ respectively.

The expectation vector, dispersion matrix, and the mean square error matrix of $\hat{\beta}_m$ are given as $E\left(\hat{\beta}_m\right) = \beta + H\delta$, $D\left(\hat{\beta}_m\right) = \sigma^2 S^{-1} - \sigma^2 G$ and $MSE\left(\hat{\beta}_m\right) = \sigma^2\left(S^{-1} - G\right) + H\delta\delta'H'$ respectively, where, $G = S^{-1}R'\left(\Omega + RS^{-1}R'\right)^{-1}RS^{-1}$, $H = S^{-1}R'\left(\Omega + RS^{-1}R'\right)^{-1}$ and $\delta = E\left(r\right) - R\beta$.

Li and Yang (2010) proposed the Stochastic Mixed Ridge Estimator (SMRE), and is given as

$$\hat{\beta}_{SMRE}\left(k\right) = W\hat{\beta}_m \;. \tag{11}$$

The Stochastic Restricted Liu Estimator (SRLE) was proposed by Hubert and Wijekoon (2006), and is given by

$$\hat{\beta}_{srd}\left(d\right)=F_d\hat{\beta}_m \ . \tag{12}$$

By using the similar idea used by Hubert and Wijekoon (2006) and Li and Yang (2010), write the Stochastic Restricted Principal Component Regression Estimator (SRPCRE) as

$$\hat{\beta}_{SRPCRE}=L_k\hat{\beta}_m \ . \tag{13}$$

Turn to the question of the statistical evaluation of the compatibility of sample and stochastic information. The classical procedures is to test the hypothesis

$$H_0:\delta=0 \text{ against } H_1:\delta\neq 0 \tag{14}$$

under linear model (1) and stochastic prior information (9).

The Ordinary Stochastic Pre Test Estimator (OSPE) of β (Wijekoon, 1990) is defined as

$$\hat{\beta}_{OSPE}=\begin{cases}\hat{\beta}_m & \text{if } H_0:\delta=0\\ \hat{\beta} & \text{if } H_1:\delta\neq 0\end{cases} \tag{15}$$

Further, we can write equation (15) as

$$\hat{\beta}_{OSPE}=\hat{\beta}_m I_{\left[0,F_{m,n-p}(\alpha)\right)}\left(F\right)+\hat{\beta}\, I_{\left[F_{m,n-p}(\alpha),\infty\right)}\left(F\right) \tag{16}$$

$$\text{where} \quad F=\frac{\left(r-R\hat{\beta}\right)'\left(\Omega+RS^{-1}R'\right)^{-1}\left(r-R\hat{\beta}\right)}{m\hat{\sigma}^2} \tag{17}$$

has a non-central $F_{m,n-p,\lambda}$ distribution under $H_1:\delta\neq 0$, with non-centrality parameter

$$\lambda = \frac{\delta'\left(\Omega + RS^{-1}R'\right)^{-1}\delta}{2\sigma^2} \quad \text{with} \quad \hat{\sigma}^2 = \frac{\left(Y - X\hat{\beta}\right)'\left(Y - X\hat{\beta}\right)}{n-p} \tag{18}$$

and $I_{\left[0,F_{m,n-p}(\alpha)\right)}\left(F\right)$ and $I_{\left[0,F_{m,n-p}(\alpha),\infty\right)}\left(F\right)$ are indicator functions which take the value one if F falls in the subscripted interval, and zero otherwise. $F_{m,n-p}\left(\alpha\right)$ is the upper α-level critical value from the central F distribution $F_{m,n-p,0}$.

The expectation vector, dispersion matrix, and the mean square error matrix of $\hat{\beta}_{OSPE}$ are derived by Wijekoon (1990) are given below:

$$E\left(\hat{\beta}_{OSPE}\right) = \beta + h_\lambda\left(2\right)H\delta \tag{19}$$

$$D\left(\hat{\beta}_{OSPE}\right) = \sigma^2 S^{-1} - \sigma^2 h_\lambda\left(2\right)G + \left[2h_\lambda\left(2\right) - h_\lambda\left(4\right) - h_\lambda^2\left(2\right)\right]H\delta\delta'H' \tag{20}$$

and

$$MSE\left(\hat{\beta}_{OSPE}\right) = \sigma^2 S^{-1} - \sigma^2 h_\lambda\left(2\right)G + \left[2h_\lambda\left(2\right) - h_\lambda\left(4\right)\right]H\delta\delta'H' \tag{21}$$

respectively.

where $h_\lambda\left(\ell\right) = \Pr\left(\frac{\chi^2_{m+\ell,\lambda}}{\chi^2_{n-p}} \le \frac{mF_{m,n-p}(\alpha)}{n-p}\right)$ for $\ell \in \mathrm{N}$.

When different estimators are available for the same parameter vector β in the linear regression model, one must solve the problem of their comparison. Usually as a simultaneous measure of covariance and bias, the mean square error matrix is used, and is defined by

$$MSE\left(\hat{\beta},\beta\right) = E\left[\left(\hat{\beta}-\beta\right)\left(\hat{\beta}-\beta\right)'\right] = D\left(\hat{\beta}\right) + B\left(\hat{\beta}\right)B'\left(\hat{\beta}\right), \tag{22}$$

where $D\left(\hat{\beta}\right)$ is the dispersion matrix and $B\left(\hat{\beta}\right) = E\left(\hat{\beta}\right) - \beta$ denotes the bias vector. Recall the Scalar Mean Square Error $SMSE\left(\beta,\hat{\beta}\right) = trace\left(MSE\left(\beta,\hat{\beta}\right)\right)$.

For any two given estimators $\hat{\beta}_1$ and $\hat{\beta}_2$, the estimator $\hat{\beta}_2$ is said to be superior to $\hat{\beta}_1$ under the MSEM criterion if and only if

$$M\left(\hat{\beta}_1, \hat{\beta}_2\right) = MSE\left(\hat{\beta}_1, \beta\right) - MSE\left(\hat{\beta}_2, \beta\right) \geq 0. \tag{23}$$

The Proposed Estimator

Now it is possible to propose the Preliminary Test Principal Component Regression Estimator (PTPCRE) as

$$\tilde{\beta}_{PTPCRE} = \begin{cases} L_k \hat{\beta}_m & \text{if} \quad H_0 : \delta = 0 \\ L_k \hat{\beta} & \text{if} \quad H_1 : \delta \neq 0 \end{cases} \tag{24}$$

Then the PTPCRE can be rewritten as follows.

$$\tilde{\beta}_{PTPCRE} = L_k \hat{\beta}_m I_{\left[0, F_{m,n-p}(\alpha)\right)}\left(F\right) + L_k \hat{\beta} I_{\left(F_{m,n-p}(\alpha), \infty\right]}\left(F\right) = L_k \hat{\beta}_{OSPE}. \tag{25}$$

When $k = p$, L_k becomes I_p and consequently $\tilde{\beta}_{PTPCRE}$ becomes $\tilde{\beta}_{OSPE}$.

Using equations given in (19) and (20), we can now obtain the expectation vector, bias vector, dispersion matrix and mean square error matrix as

$$E\left(\tilde{\beta}_{PTPCRE}\right) = L_k \beta + h_\lambda\left(2\right) L_k H \delta \tag{26}$$

$$B\left(\tilde{\beta}_{PTPCRE}\right) = \left(L_k - I\right) \beta + h_\lambda\left(2\right) L_k H \delta \tag{27}$$

$$\begin{aligned} D\left(\tilde{\beta}_{PTPCRE}\right) &= \sigma^2 L_k S^{-1} L_k' - \sigma^2 h_\lambda\left(2\right) L_k G L_k' \\ &\quad + \left[2h_\lambda\left(2\right) - h_\lambda\left(4\right) - \left(h_\lambda\left(2\right)\right)^2\right] L_k H \delta \delta' H' L_k' \end{aligned} \tag{28}$$

and

$$MSE\left(\tilde{\beta}_{PTPCRE}\right)=\sigma^2 L_k S^{-1} L_k' - \sigma^2 h_\lambda(2) L_k G L_k' \\ +\left[2h_\lambda(2)-h_\lambda(4)-\left(h_\lambda(2)\right)^2\right] L_k H\delta\delta' H' L_k' \\ +\left[\left(L_k - I\right)\beta + h_\lambda(2) L_k H\delta\right]\left[\left(L_k - I\right)\beta + h_\lambda(2) L_k H\delta\right]' \tag{29}$$

respectively.

The Proposed Estimator Based on WA, LR and LM Tests

In general, the finite sample tests such as t or F were used to define the preliminary test estimator. In the field of Econometrics, these finite sample tests are not used due to large samples. In this situation it is very useful to define preliminary test estimators based on large sample tests. The three large sample tests considered in the literature are WA, LR, and LM. The WA test offers the advantage of only requiring estimates of the unrestricted model, whereas LR test requires estimates of both unrestricted and the restricted model. The LM test only requires estimates of the restricted model. In different situations, we may use one of these tests which are easier to compute.

Judge and Bock (1978) have rewritten the model given in (1) and (9) to obtain the F statistics for testing the hypothesis in (14). Using the rewritten model we can derive the test statistics for the WA, the LR and the LM tests which are well employed for testing the hypothesis (14), and are given by

$$\varepsilon_{WA}=\frac{(n+m)mF}{n-p},\quad \varepsilon_{LR}=(n+m)\ln\left[1+\frac{mF}{n-p}\right],\ \text{and}\ \ \varepsilon_{LM}=\frac{(n+m)mF}{n-p+mF} \tag{30}$$

respectively (Evans & Savin, 1982).

It's known that under the null hypothesis H_0, the three test statistics have the same asymptotic chi-square distribution with m degrees of freedom (Evans & Savin, 1982). When the exact distribution is approximated by the asymptotic chi-square distribution, the critical value for an α level test of H_0 is approximated by the central chi-square critical value $\chi_m^2(\alpha)$ for large sample tests. Further Berndt and Savin (1977) showed that a symmetric numerical inequality $\varepsilon_{WA} \geq \varepsilon_{LR} \geq \varepsilon_{LM}$ exists between these three tests. This asymptotic chi-square distribution has wide applications in the field of Econometrics.

Based on the above tests, the PTPCRE takes the form

$$\tilde{\beta}_{PTPCRE}\left(\varepsilon_*\right) = L_k \hat{\beta}_m I_{[0,\chi_m^2(\alpha))}(\varepsilon_*) + L_k \hat{\beta}\ I_{[\chi_m^2(\alpha),\infty)}(\varepsilon_*) \tag{31}$$

where (*) stands for either WA, LR or LM tests values, and $\chi_m^2(\alpha)$ is the upper percentiles of the central χ^2 distribution with m degrees of freedom.
Using equation (26), (27), (28) and (29) we can obtain the stochastic properties of PTPCRE based on WA, LR and LM tests as follows.

$$E\left[\tilde{\beta}_{PTPCRE}\left(\varepsilon_*\right)\right] = L_k\beta + h_\lambda^*(2) L_k H\delta, \tag{32}$$

$$B\left[\tilde{\beta}_{PTPCRE}\left(\varepsilon_*\right)\right] = \left(L_k - I\right)\beta + h_\lambda^*(2) L_k H\delta, \tag{33}$$

$$\begin{aligned} D\left[\tilde{\beta}_{PTPCRE}\left(\varepsilon_*\right)\right] &= \sigma^2 L_k S^{-1} L_k' - \sigma^2 h_\lambda^*(2) L_k G L_k' \\ &\quad + \left[2h_\lambda^*(2) - h_\lambda^*(4) - \left(h_\lambda^*(2)\right)^2\right] L_k H\delta\delta' H' L_k', \end{aligned} \tag{34}$$

$$\begin{aligned} MSE\left[\tilde{\beta}_{PTPCRE}\left(\varepsilon_*\right)\right] &= \sigma^2 L_k S^{-1} L_k' - \sigma^2 h_\lambda^*(2) L_k G L_k' \\ &\quad + \left[2h_\lambda^*(2) - h_\lambda^*(4) - \left(h_\lambda^*(2)\right)^2\right] L_k H\delta\delta' H' L_k' \\ &\quad + \left[\left(L_k - I\right)\beta + h_\lambda^*(2) L_k H\delta\right]\left[\left(L_k - I\right)\beta + h_\lambda^*(2) L_k H\delta\right]', \end{aligned} \tag{35}$$

where $h_\lambda^*(\ell) = \Pr\left(\dfrac{\chi^2_{m+\ell,\lambda}}{\chi^2_{n-p}} \le \dfrac{mc^*}{n-p}\right)$ for $\ell \in N$ and c^* takes the value for WA, LR and LM tests as

$$c^{WA} = \frac{(n-p)\chi_m^2(\alpha)}{(n+m)m}, \quad c^{LR} = \frac{(n-p)\left(e^{\chi_m^2(\alpha)/(n+m)} - 1\right)}{m} \quad \text{and}$$

$$c^{LM} = \frac{(n-p)\chi_m^2(\alpha)}{m\left(n+m-\chi_m^2(\alpha)\right)}.$$

Mean Square Error Matrix Comparisons

The performance of PTPCRE will be compared using WA, LR and LM tests with respect to Mean Square Error Matrix (MSEM) sense for the two cases in which the stochastic restrictions are correct, and not correct.

Now we consider the following dispersion matrix differences.

$$D\left[\tilde{\beta}_{PTPCRE}\left(\varepsilon_{WA}\right)\right]-D\left[\tilde{\beta}_{PTPCRE}\left(\varepsilon_{LR}\right)\right]=\sigma^2\psi_1 L_k G L_k'-\xi_1 L_k H\delta\delta' H' L_k',$$

$$D\left[\tilde{\beta}_{PTPCRE}\left(\varepsilon_{LR}\right)\right]-D\left[\tilde{\beta}_{PTPCRE}\left(\varepsilon_{LM}\right)\right]=\sigma^2\psi_2 L_k G L_k'-\xi_2 L_k H\delta\delta' H' L_k',$$

$$D\left[\tilde{\beta}_{PTPCRE}\left(\varepsilon_{WA}\right)\right]-D\left[\tilde{\beta}_{PTPCRE}\left(\varepsilon_{LM}\right)\right]=\sigma^2\psi_3 L_k G L_k'-\xi_3 L_k H\delta\delta' H' L_k',$$

where

$$\psi_1=h_\lambda^{LR}(2)-h_\lambda^{WA}(2)\ge 0,\quad \psi_1^*=h_\lambda^{LR}(4)-h_\lambda^{WA}(4)\ge 0,\quad \tilde{\psi}_1=h_\lambda^{LR}(2)+h_\lambda^{WA}(2)\ge 0,$$

$$\psi_2=h_\lambda^{LM}(2)-h_\lambda^{LR}(2)\ge 0,\quad \psi_2^*=h_\lambda^{LM}(4)-h_\lambda^{LR}(4)\ge 0,$$

$$\tilde{\psi}_2=h_\lambda^{LM}(2)+h_\lambda^{LR}(2)\ge 0,\quad \psi_3=h_\lambda^{LM}(2)-h_\lambda^{WA}(2)\ge 0,$$

$$\psi_3^*=h_\lambda^{LM}(4)-h_\lambda^{WA}(4)\ge 0,\quad \tilde{\psi}_3=h_\lambda^{LM}(2)+h_\lambda^{WA}(2)\ge 0,\quad \xi_1=\left(2-\psi_1\right)\tilde{\psi}_1-\psi_1^*,$$

$$\xi_2=\left(2-\psi_2\right)\tilde{\psi}_2-\psi_2^* \text{ and } \xi_3=\left(2-\psi_3\right)\tilde{\psi}_3-\psi_3^*.$$

It is clear that $2-\psi_1\ge 1$ since $\psi_1\le 1$. This implies that $\left(2-\psi_1\right)\tilde{\psi}_1\ge\tilde{\psi}_1$ as $\tilde{\psi}_1\ge 0$. But we can show that $\tilde{\psi}_1-\psi_1^*\ge 0$. This implies that $\xi_1=\left(2-\psi_1\right)\tilde{\psi}_1-\psi_1^*\ge 0$ as $\left(2-\psi_1\right)\tilde{\psi}_1\ge\tilde{\psi}_1$. Similarly we can show that $\xi_2\ge 0$ and $\xi_3\ge 0$.

Write the following mean square error matrices differences.

$$\begin{aligned} &MSE\left[\tilde{\beta}_{PTPCRE}\left(\varepsilon_{WA}\right)\right]-MSE\left[\tilde{\beta}_{PTPCRE}\left(\varepsilon_{LR}\right)\right]. \\ &\qquad = L_k\left(D_1+b_{WA}b_{WA}'-b_{LR}b_{LR}'\right)L_k', \end{aligned} \tag{36}$$

$$MSE\left[\tilde{\beta}_{PTPCRE}\left(\varepsilon_{LR}\right)\right]-MSE\left[\tilde{\beta}_{PTPCRE}\left(\varepsilon_{LM}\right)\right] = L_k\left(D_2+b_{LR}b'_{LR}-b_{LM}b'_{LM}\right)L'_k, \tag{37}$$

$$MSE\left[\tilde{\beta}_{PTPCRE}\left(\varepsilon_{WA}\right)\right]-MSE\left[\tilde{\beta}_{PTPCRE}\left(\varepsilon_{LM}\right)\right] = L_k\left(D_3+b_{WA}b'_{WA}-b_{LM}b'_{LM}\right)L'_k, \tag{38}$$

where $D_i=\sigma^2\psi_i G-\xi_i H\delta\delta' H'$ for $i = 1, 2, 3$, $b_{WA}=\left(I-L_k^{-1}\right)\beta+h_\lambda^{WA}\left(2\right)H\delta$, $b_{LR}=\left(I-L_k^{-1}\right)\beta+h_\lambda^{LR}\left(2\right)H\delta$ and $b_{LM}=\left(I-L_k^{-1}\right)\beta+h_\lambda^{LM}\left(2\right)H\delta$.
Note that $b_{WA}=b_{LR}=b_{LM}=\left(I-L_k^{-1}\right)\beta$ when $\delta = 0$.

Based on the mean square error matrix differences the following theorems can be stated.

Theorem 1:

i) When the stochastic restrictions are true (i.e. $\delta = 0$), $\tilde{\beta}_{PTPCRE}\left(\varepsilon_{LR}\right)$ is always superior to $\tilde{\beta}_{PTPCRE}\left(\varepsilon_{WA}\right)$ in the mean square error matrix sense.

ii) When the stochastic restriction are not true (i.e. $\delta \neq 0$), and if $\lambda \leq \frac{\psi_1}{2\left[\left(2-\psi_1\right)\tilde{\psi}_1-\psi_1^*\right]}$ then the $\tilde{\beta}_{PTPCRE}\left(\varepsilon_{LR}\right)$ is superior to $\tilde{\beta}_{PTPCRE}\left(\varepsilon_{WA}\right)$ with respect to MSE matrix sense if and only if

$$\left\{\left[\left(I-L_k^{-1}\right)\beta+h_\lambda^{WA}\left(2\right)H\delta\right]' D_1^+\left[\left(I-L_k^{-1}\right)\beta+h_\lambda^{WA}\left(2\right)H\delta\right]+1\right\}$$
$$*\left\{\left[\left(I-L_k^{-1}\right)\beta+h_\lambda^{LR}\left(2\right)H\delta\right]' D_1^+\left[\left(I-L_k^{-1}\right)\beta+h_\lambda^{LR}\left(2\right)H\delta\right]-1\right\}$$
$$\leq\left\{\left[\left(I-L_k^{-1}\right)\beta+h_\lambda^{WA}\left(2\right)H\delta\right]' D_1^+\left[\left(I-L_k^{-1}\right)\beta+h_\lambda^{LR}\left(2\right)H\delta\right]\right\}^2.$$

Proof:

Consider the mean square error matrix difference (36) between WA and LR.

$$MSE\left[\tilde{\beta}_{PTPCRE}\left(\varepsilon_{WA}\right)\right]-MSE\left[\tilde{\beta}_{PTPCRE}\left(\varepsilon_{LR}\right)\right]=L_k\left(D_1+b_{WA}b'_{WA}-b_{LR}b'_{LR}\right)L'_k$$

where $D_1=\sigma^2\psi_1 G-\xi_1 H\delta\delta' H'$.

When the stochastic restrictions are true (i.e. $\delta = 0$), then the MSE matrix difference in (36) reduces to $\sigma^2\psi_1 L_k G L'_k$ which is clearly a nonnegative definite matrix since $\psi_1 \geq 0$, $G \geq 0$ and $L_k > 0$.

When the stochastic restrictions are not correct (i.e. $\delta \neq 0$), then $MSE\left[\tilde{\beta}_{PTPCRE}\left(\varepsilon_{WA}\right)\right]-MSE\left[\tilde{\beta}_{PTPCRE}\left(\varepsilon_{LR}\right)\right]$ is a nonnegative definite matrix if and only if $D_1+b_{WA}b'_{WA}-b_{LR}b'_{LR}$ is a nonnegative matrix. To apply lemma 3 (Appendix) we have to show that D_1 is a nonnegative definite matrix. We rewrite D_1 as $D_1=\xi_1 D_1^*$, where $D_1^*=\dfrac{\sigma^2\psi_1}{\xi_1}G-H\delta\delta' H'$. Then D_1 is a nonnegative definite matrix, if and only if D_1^* is nonnegative definite matrix. To show that D_1^* is nonnegative definite matrix, lemma 1 (Appendix) is used by setting

$$\gamma=\frac{\sigma^2\psi_1}{\xi_1},\ \ B=G\ \text{ and }\ a=H\delta.$$

Note that $G=S^{-1}R'\left(\Omega+RS^{-1}R'\right)^{-1}RS^{-1}\geq 0$, and the generalized inverse of G is $G^{-}=SR^{+}\left(\Omega+RS^{-1}R'\right)\left(R'\right)^{+}S$. Hence $GG^{-}H\delta=H\delta$. This implies that $H\delta\in\Re\left(G\right)$, where $\Re\left(.\right)$ denote the column space of the corresponding matrix and R^{+} is a Moore-Penrose matrix of R.

Then according to lemma 1, $D_1^*\geq 0$ if and only if

$$\delta' H' G^{-} H\delta\leq\frac{\sigma^2\psi_1}{\xi_1}. \tag{39}$$

After some straightforward calculations it can be shown

$$\delta' H' G^{-} H \delta = \delta' \left(\Omega + R S^{-1} R' \right)^{-1} \delta. \tag{40}$$

By substituting this result to (39) obtain

$$\frac{\delta' \left(\Omega + R S^{-1} R' \right)^{-1} \delta}{2\sigma^2} \le \frac{\psi_1}{2\xi_1}. \tag{41}$$

Using (18), this inequality can be rewritten as

$$\lambda \le \frac{\psi_1}{2\xi_1} = \frac{\psi_1}{2\left[\left(2 - \psi_1 \right) \tilde{\psi}_1 - \psi_1^* \right]}. \tag{42}$$

This implies that D_1^* is a nonnegative definite matrix if and only if $\lambda \le \frac{\psi_1}{2\left[\left(2 - \psi_1 \right) \tilde{\psi}_1 - \psi_1^* \right]}$. Therefore $D_1 = \xi_1 L_k D_1^* L_k'$ is nonnegative definite matrix if and only if $\lambda \le \frac{\psi_1}{2\left[\left(2 - \psi_1 \right) \tilde{\psi}_1 - \psi_1^* \right]}$, $\xi_1 \ge 0$.

To apply lemma 3 (Appendix), the Moore Penrose inverse of D_1 is obtained by using lemma 2 (Appendix), and is given by

$$D_1^{+} = \frac{1}{\sigma^2 \psi_1} \times \left[G^{+} + \frac{\xi_1}{\sigma^2 \psi_1 - \xi_1 \delta' H' G^{+} H \delta} G^{+} H \delta \delta' H' G^{+} \right] \tag{43}$$

After some straightforward calculations we can show that

$$\delta' H' G^{+} H \delta = 2\sigma^2 \lambda \tag{44}$$

Using (43) and (44) we can easily prove that $D_1 D_1^{+} = I_p$, where I_p is an identity matrix with order $(p \times p)$. This implies that $D_1 D_1^{+} b_{WA} = b_{WA}$ and $D_1 D_1^{+} b_{LR} = b_{LR}$. Then we have $b_{WA} \in \Re\left(D_1 \right)$ and $b_{LR} \in \Re\left(D_1 \right)$. To establish condition (a) in lemma 3, we find $f_{ij} = b_i' D_1^{-} b_j$ for *i*, *j*, = *WA*, *LR* such that

$$f_{WA,WA} = \left[\left(I - L_k^{-1}\right)\beta + h_\lambda^{WA}(2)H\delta\right]' D_1^{+}\left[\left(I - L_k^{-1}\right)\beta + h_\lambda^{WA}(2)H\delta\right]$$

$$f_{LR,LR} = \left[\left(I - L_k^{-1}\right)\beta + h_\lambda^{LR}(2)H\delta\right]' D_1^{+}\left[\left(I - L_k^{-1}\right)\beta + h_\lambda^{LR}(2)H\delta\right] \text{ and}$$

$$f_{WA,LR} = \left[\left(I - L_k^{-1}\right)\beta + h_\lambda^{WA}(2)H\delta\right]' D_1^{+}\left[\left(I - L_k^{-1}\right)\beta + h_\lambda^{LR}(2)H\delta\right]$$

Note that, instead of D_1^-, the Moore Penrose inverse D_1^+ of D_1 is used, since f_{ij} is invariant to the choice of D_1^-.

Now according to lemma 3 (Appendix) $MSE\left[\tilde{\beta}_{PTPCRE}(\varepsilon_{WA})\right] - MSE\left[\tilde{\beta}_{PTPCRE}(\varepsilon_{LR})\right] \geq 0$ if and only if

$$\begin{aligned}
&\left\{\left[\left(I - L_k^{-1}\right)\beta + h_\lambda^{WA}(2)H\delta\right]' D_1^{+}\left[\left(I - L_k^{-1}\right)\beta + h_\lambda^{WA}(2)H\delta\right] + 1\right\} \\
&* \left\{\left[\left(I - L_k^{-1}\right)\beta + h_\lambda^{LR}(2)H\delta\right]' D_1^{+}\left[\left(I - L_k^{-1}\right)\beta + h_\lambda^{LR}(2)H\delta\right] - 1\right\} \\
&\leq \left\{\left[\left(I - L_k^{-1}\right)\beta + h_\lambda^{WA}(2)H\delta\right]' D_1^{+}\left[\left(I - L_k^{-1}\right)\beta + h_\lambda^{LR}(2)H\delta\right]\right\}^2.
\end{aligned}$$

This completes the proof.

By considering the mean square error matrix differences given in equation (37) and (38), we can state Theorem 2 and Theorem 3 respectively. The proofs of these theorems are similar to the proof of Theorem 1.

Theorem 2:

i) When the stochastic restrictions are true (i.e. $\delta = 0$), $\tilde{\beta}_{PTPCRE}(\varepsilon_{LM})$ is always superior to $\tilde{\beta}_{PTPCRE}(\varepsilon_{LR})$ in the mean square error matrix sense.

ii) When the stochastic restriction are not true (i.e. $\delta \neq 0$), and if $\lambda \leq \dfrac{\psi_2}{2\left[\left(2-\psi_2\right)\tilde{\psi}_2-\psi_2^*\right]}$ then the $\tilde{\beta}_{PTPCRE}\left(\varepsilon_{LM}\right)$ is superior to $\tilde{\beta}_{PTPCRE}\left(\varepsilon_{LR}\right)$ with respect to MSE matrix sense if and only if

$$\left\{\left[\left(I-L_k^{-1}\right)\beta+h_\lambda^{LR}\left(2\right)H\delta\right]' D_2^+\left[\left(I-L_k^{-1}\right)\beta+h_\lambda^{LR}\left(2\right)H\delta\right]+1\right\}$$

$$*\left\{\left[\left(I-L_k^{-1}\right)\beta+h_\lambda^{LM}\left(2\right)H\delta\right]' D_2^+\left[\left(I-L_k^{-1}\right)\beta+h_\lambda^{LM}\left(2\right)H\delta\right]-1\right\}$$

$$\leq\left\{\left[\left(I-L_k^{-1}\right)\beta+h_\lambda^{LR}\left(2\right)H\delta\right]' D_2^+\left[\left(I-L_k^{-1}\right)\beta+h_\lambda^{LM}\left(2\right)H\delta\right]\right\}^2.$$

Theorem 3:

i) When the stochastic restrictions are true (i.e. $\delta = 0$), $\tilde{\beta}_{PTPCRE}\left(\varepsilon_{LM}\right)$ is always superior to $\tilde{\beta}_{PTPCRE}\left(\varepsilon_{WA}\right)$ in the mean square error matrix sense.

ii) When the stochastic restrictions are not true (i.e. $\delta \neq 0$), and if $\lambda \leq \dfrac{\psi_3}{2\left[\left(2-\psi_3\right)\tilde{\psi}_3-\psi_3^*\right]}$ then the $\tilde{\beta}_{PTPCRE}\left(\varepsilon_{LM}\right)$ is superior to $\tilde{\beta}_{PTPCRE}\left(\varepsilon_{WA}\right)$ with respect to MSE matrix sense if and only if

$$\left\{\left[\left(I-L_k^{-1}\right)\beta+h_\lambda^{WA}\left(2\right)H\delta\right]' D_3^+\left[\left(I-L_k^{-1}\right)\beta+h_\lambda^{WA}\left(2\right)H\delta\right]+1\right\}$$

$$*\left\{\left[\left(I-L_k^{-1}\right)\beta+h_\lambda^{LM}\left(2\right)H\delta\right]' D_3^+\left[\left(I-L_k^{-1}\right)\beta+h_\lambda^{LM}\left(2\right)H\delta\right]-1\right\}$$

$$\leq\left\{\left[\left(I-L_k^{-1}\right)\beta+h_\lambda^{WA}\left(2\right)H\delta\right]' D_3^+\left[\left(I-L_k^{-1}\right)\beta+h_\lambda^{LM}\left(2\right)H\delta\right]\right\}^2.$$

Monte Carlo Simulation

To illustrate the behavior of the proposed estimators, a Monte Carlo Simulation study was designed by considering different levels of multicollinearity. Following McDonald and Galarneau (1975) generate explanatory variables as follows:

$$x_{ij} = \left(1-\rho^2\right)^{1/2} z_{ij} + \rho z_{i,p+1}, \; i = 1, 2, \ldots, n, \; j = 1, 2, \ldots, p,$$

where z_{ij} is an independent standard normal pseudo random number, and ρ is specified so that the theoretical correlation between any two explanatory variables is given by ρ^2. A dependent variable is generated by using the equation.

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \varepsilon_i, \; i = 1, 2, \ldots, n,$$

where ε_i is a normal pseudo random number with mean zero and variance σ_i^2. Newhouse and Oman (1971) have noted that if the MSE is a function of σ^2 and β, and if the explanatory variables are fixed, then subject to the constraint $\beta'\beta = 1$, the MSE is minimized when β is the normalized eigenvector corresponding to the largest eigenvalue of the $X'X$ matrix. In this study we choose the normalized eigenvector corresponding to the largest eigenvalue of $X'X$ as the coefficient vector β, $n = 50$, $p = 4$ and $\sigma_i^2 = 1$. Three different sets of correlations are considered by selecting the values as $\rho = 0.7$, 0.8 and 0.9, and two various significance levels are taken as $\alpha = 0.01$ and 0.05. Further R, r and υ in equation (9) are taken as $R = (0, 1, 3, 1)$, $r = 0$ and $\upsilon \sim N\left(0, \Omega = \hat{\sigma}_\rho^2\right)$, where $\hat{\sigma}_\rho^2$ is estimated by using equation (18). The eigenvalues of the matrix S for $\rho = 0.7$, 0.8 and 0.9 are given in Table 1.

The first three principal components account for 91.29% and 93.65% of the total variance when $\rho = 0.7$ and 0.8 respectively, and also the first two principal components account for 91.8% of the total variance when $\rho = 0.9$. Therefore we choose the number of the principal components $k = 3$ when $\rho = 0.7$ and 0.8, and $k = 2$ when $\rho = 0.9$. Table 2, Table 3 and Table 4 show the scalar mean square errors (SMSE) obtained by using equation (35).

Table 1. Eigenvalues of the matrix *S* for ρ = 0.7, 0.8 and 0.9.

ρ	Eigenvalues	Proportion of Total Variance (%)	Cumulative Percentage of Total Variance (%)
0.7	0.12261	63.02	63.02
	0.03209	16.49	79.51
	0.02292	11.78	91.29
	0.01693	8.80	100
0.8	0.13843	73.15	73.15
	0.02255	11.92	85.07
	0.01624	8.58	93.65
	0.01201	6.35	100
0.9	0.15549	85.30	85.3
	0.01184	6.50	91.8
	0.00861	4.72	96.52
	0.00636	3.48	100

Table 2. Estimated SMSE of PTPCRE for WA, LR and LM tests for ρ = 0.7 and *k* = 3.

Estimators	SMSE at $\alpha = 0.01$	SMSE at $\alpha = 0.05$
$\tilde{\beta}_{PTPCRE}(\varepsilon_{WA})$	0.0584	0.0543
$\tilde{\beta}_{PTPCRE}(\varepsilon_{LR})$	0.0590	0.0544
$\tilde{\beta}_{PTPCRE}(\varepsilon_{LM})$	0.0597	0.0544

Table 3. Estimated SMSE of PTPCRE for WA, LR and LM tests for ρ = 0.8 and *k* = 3.

Estimators	SMSE at $\alpha = 0.01$	SMSE at $\alpha = 0.05$
$\tilde{\beta}_{PTPCRE}(\varepsilon_{WA})$	0.2444	0.1556
$\tilde{\beta}_{PTPCRE}(\varepsilon_{LR})$	0.2621	0.1592
$\tilde{\beta}_{PTPCRE}(\varepsilon_{LM})$	0.2841	0.1633

Table 4. Estimated SMSE of PTPCRE for WA, LR and LM tests for ρ = 0.9 and *k* = 2.

Estimators	SMSE at $\alpha = 0.01$	SMSE at $\alpha = 0.05$
$\tilde{\beta}_{PTPCRE}(\varepsilon_{WA})$	0.3227	1.7228
$\tilde{\beta}_{PTPCRE}(\varepsilon_{LR})$	0.3313	1.7098
$\tilde{\beta}_{PTPCRE}(\varepsilon_{LM})$	0.3409	1.6961

Based on Table 2, there are no big differences in the SMSE among the estimators when $\rho = 0.7$. Based on the Table 3, the PTPCRE based on WA test has the smallest SMSE. From Table 4, notice that when $\rho = 0.9$ and $\alpha = 0.01$, the PTPCRE based on WA test has the smallest SMSE. Then the LM test has the smallest SMSE when $\rho = 0.9$ and $\alpha = 0.05$.

Conclusion

A new Preliminary Test Estimator based on Principal Component Regression Estimator defined in the linear regression model when the stochastic restrictions are available in addition to the sample information, and when the explanatory variables are multicollinear. Based on the simulation study, we can conclude that the PTPCRE based on WA test has the smallest SMSE when $\rho = 0.9$ and $\alpha = 0.01$. The PTPCRE based on LM test has smallest SMSE when $\rho = 0.9$ and $\alpha = 0.05$.

References

Arumairajan, S & Wijekoon, P. (2013). Improvement of the preliminary test estimator when stochastic restrictions are available in linear regression model. *Open Journal of Statistics*, *3*(4), 283-292. doi:10.4236/ojs.2013.34033

Atchison, J & Silvey, S. D. (1958). Maximum likelihood estimation of parameters subject to restraints. *Annals of Mathematical Statistics*, *29*(3), 813-828. doi:10.1214/aoms/1177706538

Baksalary, J. K. & Kala, R. (1983). Partial orderings between matrices one of which is of rank one. *Bulletin of the Polish Academy of Science*, *Mathematics*, *31*, 5-7.

Baksalary, J. K. & Trenkler, G. (1991). Nonnegative and positive definiteness of matrices modified by two matrices of rank one. *Linear Algebra and its Applications*, *151*, 169-184. doi:10.1016/0024-3795(91)90362-Z

Bancroft, T. A. (1944). On biases in estimation due to use of preliminary tests of significance. *Annals of Mathematical Statistics*, *15*(2), 190-204. doi:10.1214/aoms/1177731284

Berndt, E. R. & Savin, N. E. (1977). Conflict among criteria for testing hypothesis in the multivariate linear regression model. *Econometrica*, *45*(5), 1263-1278. doi:10.2307/1914072

Evans, G. B. A. & Savin, N. E. (1982). Conflict among the criteria revisited; The W, LR and LM tests. *Econometrica*, *50*(3), 737 -748. doi:10.2307/1912611

Golam Kibria, B. M. & Saleh, A. K. Md. E. (2003) Effect of W, LR and LM tests on the performance of preliminary test ridge regression estimators. *Journal of the Japan Statistical Society, 33* (1), 119-136. doi:10.14490/jjss.33.119

Hoerl, A. E. & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55-67. doi:10.2307/1267351

Hubert, M. H. & Wijekoon, P. (2006). Improvement of the Liu estimator in linear regression model. *Statistical Papers, 47*(3), 471-479. doi:10.1007/s00362-006-0300-4

Judge, G. G. & Bock, M. E. (1978). *The statistical implications of pre-test and Stein-rule estimators in econometrics*. New York: North Holland.

Kaçiranlar, S., Sakallioğlu, S., Akdeniz, F., Styan, G. P. H. & Werner, H. J. (1999). A new biased estimator in linear regression and a detailed analysis of the widely analyzed dataset on Portland cement. *The Indian Journal of Statistics, 61*(3), 443-459.

Kibria, B. M. G. (2003). Performance of some new ridge regression estimators. *Communication in Statistics*, *32*(2), 419-435. doi:10.1081/SAC-120017499

Li, Y. & Yang, H. (2010). A new stochastic mixed ridge estimator in linear regression. *Statistical Papers*, *51*(2), 315-323. doi:10.1007/s00362-008-0169-5

Liu, K. (1993). A new class of biased estimate in linear regression. *Communications in Statistics-Theory and Methods, 22*(2), 393-402. doi:10.1080/03610929308831027

Massy, W. F. (1965). Principal components regression in exploratory statistical research. *Journal of the American Statistical Association*, *60*(309), 234-266. doi:10.1080/01621459.1965.10480787

McDonald, G. C. & Galarneau, D. I. (1975). A Monte Carlo evaluation of some ridge-type estimators. *Journal of American Statistical Association*, *70*(350), 407-416. doi:10.1080/01621459.1975.10479882

Newhouse, J. P. & Oman, S. D. (1971). An evaluation of ridge estimators. *Rand Report*, *R-716-Pr*, 1-28.

Rao, R. (1948). Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Proceedings of the Cambridge Philosophical Society*, *44*(1), 50-57.

Saleh, A. K. M. E. & Kibria, B. M. G. (1993). Performance of some new preliminary test ridge regression estimators and their properties. *Communications in Statistics—Theory and Methods, 22*(10), 2747-2764. doi:10.1080/03610929308831183

Sarkar, N. (1992). A new estimator combining the ridge regression and the restricted least squares methods of estimation. *Communication in Statistics-Theory Methods*, *21*(7), 1987-2000.doi:10.1080/03610929208830893

Theil, H. & Goldberger, A. S. (1961). On pure and mixed estimation in economics. *International Economic review, 2*(1), 65-78. doi:10.2307/2525589

Trenkler, G. (1985). Mean square error matrix comparisons of estimators in linear regression. *Communication in Statistics A*, *14*(10), 2495-2509. doi:10.1080/03610928508829058

Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transaction of the American Mathematical Society*, *54*(3), 426-482.doi:10.2307/1990256

Wijekoon, P. (1990). *Mixed estimation and preliminary test estimation in the linear regression model*. (Unpublished doctoral thesis). University of Dortmund, Germany.

Xu, J. & Yang, H. (2011). On the restricted *r–k* class estimator and the restricted *r–d* class estimator in linear regression. *Journal of Statistical Computation and Simulation*, *81*(6), 679-691. doi:10.1080/00949650903471023

Appendix

Lemma 1: (Baksalary & Kala, 1983)
Suppose B is a symmetric real $(n \times n)$ matrix, a is an $(n \times 1)$ real vector and γ is a positive real number. Then the following two properties are equivalent

i) $\gamma B - ad$ is nonnegative definite (n.n.d)
ii) B is n.n.d, $a \in \Re(B)$ and $a'B^{-}a \leq \gamma$.

Lemma 2: (Trenkler, 1985)
Let A be a symmetric $(n \times n)$ matrix, and let a, a_1, and a_2 be $(n \times 1)$ vectors. Suppose that

a) $a \in \Re(A)$, and the real numbers ϕ and ψ satisfy $\phi \neq 0$ and $\phi + \psi a'A^{+}a \neq 0$. Then we have the identity $\left[\phi A + \psi aa'\right]^{+} = \frac{1}{\phi}\left[A^{+} - \frac{\psi}{\phi + \psi a'A^{+}a} A^{+}aa'A^{+}\right]$
b) $a_j \in \Re(A)$, $j = 1, 2$, and the real number ρ satisfies $1 + \rho a_1'A^{+}a_1 \neq 0$. Then we have $a_2 \in \Re(A + \rho a_1 a_1')$.

Lemma 3: (Baksalary & Trenkler, 1991)
Let C be a nonnegative definite matrix and c_1, c_2 be linearly independent vectors. Furthermore for some generalized inverse C^{-} of C, let $f_{ij} = c_i'C^{-}c_j$; $i = 1, 2$, $j = 1, 2$ and let

$$s = \frac{c_2'\left(I - CC^{-}\right)'\left(I - CC^{-}\right)c_2}{c_1'\left(I - CC^{-}\right)\left(I - CC^{-}\right)c_1}$$

where $c_1 \in \Re(C)$ and $\Re(.)$ denote the column space of the corresponding matrix. Then we have $C + c_1c_1' - c_2c_2' \geq 0$ if and only if

a) $c_1 \in \Re(C), c_2 \in \Re(C)$ and $\left(f_{11} + 1\right)\left(f_{22} - 1\right) \leq f_{12}^2$ or
b) $c_1 \notin \Re(C), c_2 \in \Re(C, c_1)$ and $(c_2 - sc_1))C^{-}(c_2 - sc_1) \leq 1 - s^2$

and all expressions in (a) and (b) are independent of the choice of C^{-}.