LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

INSTITUT FÜR STATISTIK

Shalabh & Christian Heumann

Simultaneous Prediction of Actual and Average Values of Study Variable Using Stein-rule Estimators

Technical Report Number 104, 2011
Department of Statistics
University of Munich

http://www.stat.uni-muenchen.de

Simultaneous Prediction of Actual and Average Values of Study variable Using Stein-rule Estimators

Shalabh

Department of Mathematics & Statistics

Indian Institute of Technology, Kanpur - 208016 (INDIA)

E-mail: shalab@iitk.ac.in, shalabh1@yahoo.com

Christian Heumann

Institute of Statistics

University of Munich, 80799 Munich (GERMANY)

E-mail: chris@stat.uni-muenchen.de

Abstract

The simultaneous prediction of average and actual values of study variable in a linear regression model is considered in this paper. Generally, either of the ordinary least squares estimator or Stein-rule estimators are employed for the construction of predictors for the simultaneous prediction. A linear combination of ordinary least squares and Stein-rule predictors are utilized in this paper to construct an improved predictors. Their efficiency properties are derived using the small disturbance asymptotic theory and dominance conditions for the superiority of predictors over each other are analyzed.

1 Introduction:

Traditionally the predictions from a linear regression model are made either for the actual values of study variable or for the average values. However, this may not be the case in many practical situations and one may be required to predict both the actual and average values of the study variable simultaneously; see, e.g. Rao et al. (2008), Shalabh (1995), and Zellner (1994). As an illustrative example, consider a new drug that promotes the duration of sleep in human beings. The manufacturer of such a drug will be more interested in knowing the average increase in the sleep duration by a specific dose, for example, in designing an advertisement or sale campaign and somewhat less interested in the actual increase of sleep duration. On the other strand, a user may be more interested in knowing the actual increase in sleep duration rather than the average duration. Suppose the statistician utilizes the theory of regression analysis for prediction. It is expected from the statistician to safeguard the interest of both the manufacturer and user who are interested in the prediction of the average and actual increase, respectively, although they may assign varying weight to prediction of actual and average increases of sleep attributable to the specific dose of new drug. The classical theory of prediction can predict either the actual value or the average value of study variable but not simultaneously.

In view of the importance of simultaneous prediction of actual and average values of study variable in a linear regression model, Shalabh (1995), see also Rao et al. (2008), has presented a framework for the simultaneous prediction of actual and average values of study variable. Shalabh (1995) has examined the efficiency properties of predictions arising from least squares and Stein-rule estimation procedures. The work on the issue of simultaneous prediction has been extended in various directions from various perspectives in different models in the literature. Toutenburg and Shalabh (2000), Shalabh and Chandra (2002), and Dube and Manocha (2002) analyzed the simultaneous prediction in restricted regression model, Chaturvedi and Singh (2000) and Chaturvedi et al. (2008) employed Stein-rule estimators for simultaneous prediction; Chaturvedi et al. (2002) discussed the issue of simultaneous prediction in a multivariate set up with an unknown covariance matrix of disturbance vector; Shalabh et al. (2008) considered the simultaneous prediction in measurement error models etc. In all such works, either the ordinary least squares (OLS) predictor or the Stein-rule (SR) predictor are utilized for predicting the actual and average values of study variable. They provide more efficient predictions under certain conditions depending on whether they are used for actual or average value predictions. So a natural question arises that can we utilize the good properties of the two predictors and obtain an improved estimator? Based on this, we have utilized the OLS and SR predictors together and have proposed two predictors in this paper. Their efficiency properties are derived and analyzed. The small disturbance asymptotic theory is utilized to derive the efficiency properties and dominance conditions for the superiority of predictors over each other are derived and analyzed.

The plan of this paper is as follows. Section 2 provides these predictions and presents their motivation. Their properties are analyzed in Sections 3 and 4 employing the small disturbance asymptotic theory. Some concluding remarks are placed in Section 5. Finally, derivation of main results is outlined in Appendix.

2 Model Specification And Predictions:

Let us postulate the following linear regression model:

$$y = X\beta + u \tag{2.1}$$

where y is a $n \times 1$ vector of n observations on the study variable, X is a $n \times p$ matrix of n observations on $p(> 2)$ explanatory variables, β is a $p \times 1$ vector of p regression coefficients and u is a $n \times 1$ vector of disturbances following a multivariate normal distribution with mean vector 0 and variance covariance matrix $\sigma^2 I_n$.

It is assumed that the scalar σ^2 is unknown and the matrix X has full column rank.

When the simultaneous prediction of average values $A_v = E(y)$ and actual values $A_c = y$ within the simple is to be considered, we may define our target function as

$$T = \lambda A_v + (1 - \lambda) A_c \tag{2.2}$$

where λ is a nonstochastic scalar between 0 and 1; see Shalabh (1995), Rao et al. (2008). The value of λ may reflect the weight being given to the prediction of average values in relation to the prediction of actual values.

The least squares estimator of β is given by

$$b_L = (X'X)^{-1} X'y \tag{2.3}$$

which is the best linear unbiased estimator of β. Sometimes the properties like linearity and unbiasedness may not be desirable. Under such situation, it may be possible to obtain an estimator with reduced variability by relaxing the properties of linearity and unbiasedness. The family of Stein-rule estimators gives rise to such estimators. The Stein-rule estimator of β is defined by

$$b_S = \left[1 - \frac{2(p-2)k}{(n-p+2)} . \frac{y'\bar{H}y}{y'Hy}\right] b_L \tag{2.4}$$

where $H = X(X'X)^{-1}X, \bar{H} = (I - H)$ and k is any positive nonstochastic scalar; see, e.g., Judge and Bock (1978), Saleh(2006).

Based on (2.3) and (2.4), predictions for the values of the study variable are obtained by Xb_L and Xb_S which can be used for both the average values $A_v = E(y)$ as well as actual values $A_c = y$.

For both the components A_v and A_c of T defined by (2.2), we may use Xb_L so that the vector of predictions for T is given by

$$T_{LL} = Xb_L. \tag{2.5}$$

Similarly, if we employ Xb_S for both A_v and A_c, we find the vector of predictions as

$$\begin{aligned} T_{SS} &= Xb_s \\ &= \left[1 - \frac{2(p-2)k}{(n-p+2)}.\frac{y'\bar{H}y}{y'Hy}\right] Xb_L. \end{aligned} \tag{2.6}$$

On the other hand, if we use Xb_L for A_c and Xb_S for A_v in T, we get the vector of predictions as

$$\begin{aligned} T_{SL} &= \lambda Xb_S + (1-\lambda)Xb_L \\ &= \left[1 - \frac{2(p-2)\lambda k}{(n-p+2)}.\frac{y'\bar{H}y}{y'Hy}\right] b_L. \end{aligned} \tag{2.7}$$

Similarly, utilizing Xb_L for A_v and Xb_S for A_c in T, we find yet another vector of predictions

$$\begin{aligned} T_{LS} &= \lambda Xb_L + (1-\lambda)Xb_S \\ &= \left[1 - \frac{2(p-2)(1-\lambda)k}{(n-p+2)}.\frac{y'\bar{H}y}{y'Hy}\right] b_L. \end{aligned} \tag{2.8}$$

Our motivation underlying the formulation of (2.7) and (2.8) is as follows. If we compare Xb_L and Xb_S with respect to the criterion of total mean squared error, it is well known that Xb_L is superior to Xb_S for all positive values of k when the aim is to predict A_c (the actual values of study variable). When the aim is to predict A_v (the average values of study variable), Xb_S is superior to Xb_L for positive values of k below one. Thus if we use superior predictions, i.e., Xb_L for A_c and Xb_S for A_v in T defined by (2.3), we get T_{SL}. Conversely, if we consider predictions, i.e. Xb_L for A_V and Xb_S for A_c, it leads to T_{LS}.

Looking at the expressions in (2.5), (2.6), (2.7) and (2.8). We may define a family of predictions for T as follows:

$$P_g = \left[1 - \frac{2(p-2)gk}{(n-p+2)} . \frac{y'\bar{H}y}{y'Hy}\right] Xb_L \tag{2.9}$$

where $0 \leq g \leq 1$ is any nonstochastic scalar characterizing the predictions.

Notice that assigning values $0, 1, \lambda$ and $(1-\lambda)$ to g in (2.9) yield T_{LL}, T_{SS}, T_{SL} and T_{LS} respectively.

Next, let us consider the problem of prediction outside the sample. For this purpose, let us assume to be given a matrix X_f of m (e.g., future) values of explanatory values corresponding to which the values of study variable are to be predicted. Further, we assume that the regression relationship continues to remain valid so that we can write

$$y_f = X_f\beta + u_f \tag{2.10}$$

where y_f denotes a $m \times 1$ vector of values of study variable to be predicted and u_f is a $m \times 1$ vector of disturbances having same distributional properties as u in (2.1).

Further, we assume that u and u_f are stochastically independent.

Defining the target function as

$$T_f = \lambda E(y_f) + (1-\lambda)y_f, \tag{2.11}$$

we can formulate the following vectors of predictions of T_f in the spirit of (2.5)-(2.8):

$$T_{fLL} = X_f b_L \tag{2.12}$$

$$T_{fSS} = X_f b_S \tag{2.13}$$

$$T_{fSL} = \left[1 - \frac{2(p-2)\lambda k}{(n-p+2)} . \frac{y'\bar{H}y}{y'Hy}\right] X_f b_L \tag{2.14}$$

$$T_{fLS} = \left[1 - \frac{2(p-2)(1-\lambda)k}{(n-p+2)} . \frac{y'\bar{H}y}{y'Hy}\right] X_f b_L \tag{2.15}$$

whence the following family of predictions can be defined:

$$P_{fg} = \left[1 - \frac{2(p-2)g_f k}{(n-p+2)} \cdot \frac{y'\bar{H}y}{y'Hy}\right] X_f b_L \tag{2.16}$$

where $0 \leq g_f \leq 1$ is a nonstochastic scalar characterizing the predictions.

If we set the value of g as 0,1,λ and $(1-\lambda)$, we obtain (2.12), (2.13), (2.14) and (2.15) respectively as special cases.

3 Asymptotic Efficiency Properties of Predictions Within The Sample:

It is easy to see that the predictions based on least squares are weakly unbiased in the sense that

$$E(T_{LL} - T) = 0. \tag{3.1}$$

Further, the second order moment matrix is

$$E(T_{LL} - T)(T_{LL} - T)' = \sigma^2 \left[\lambda^2 I_n + (1-2\lambda)\bar{H}\right]. \tag{3.2}$$

Similar exact expressions for the bias vector and second order moment matrix of P_g for any nonzero value of g can be derived following, for instance, Judge and Bock (1978) but they would be sufficiently intricate and would not permit to deduce any clear inference regarding the efficiency properties. We therefore propose to consider their asymptotic approximations employing the small disturbance asymptotic theory.

Theorem I: The asymptotic approximation for the bias vector of P_g for nonzero values of g to order $O(\sigma^2)$ is

$$\begin{aligned} B(P_g) &= E(P_g - T) \\ &= -\sigma^2 \left[\frac{2(n-p)(p-2)gk}{(n-p+2)\beta' X' X \beta}\right] X\beta \end{aligned} \tag{3.3}$$

while the difference between the second order moment matrices of P_g and $P_o \equiv T_{LL}$ to order $O(\sigma^4)$ is given by

$$\begin{aligned} D(P_g; P_o) &= E(P_o - T)(P_o - T)' - E(P_g - T)(P_g - T)' \\ &= \sigma^4 \left[\frac{4(n-p)(p-2)gk}{(n-p+2)\beta' X' X \beta} \right] XCX' \end{aligned} \tag{3.4}$$

where

$$C = \lambda (X'X)^{-1} - \frac{2\lambda + (p-2)gk}{\beta' X' X \beta} \beta \beta'. \tag{3.5}$$

These results are derived in Appendix.

From (3.3), we observe that P_g is not weakly unbiased. However, if we define the norm of bias vector to the order of our approximation as

$$B(P_g)' B(P_g) = \sigma^4 \left[\frac{4(n-p)^2 (p-2)^2 g^2 k^2}{(n-p+2)^2 \beta' X' X \beta} \right] \tag{3.6}$$

then with respect to the criterion of such a norm, P_g is superior than P_{g*} for g less than $g*$. In particular, both T_{SL} and T_{LS} are better than T_{SS} for positive λ. Further, T_{SL} is superior or inferior than T_{LS} when λ is less or greater than 0.5. When $\lambda = 0.5$, i.e., equal weight is assigned to the prediction of actual and average values of study variable, both T_{LS} and T_{SL} are equally good.

Next, let us compare the predictions with respect to the criterion of second order moment matrix to order $O(\sigma^4)$. For this purpose, we utilize the following two results for any $p \times 1$ vector a and $p \times p$ positive definite matrix A.

Result I: The matrix $(A^{-1} - aa')$ is positive definite if only only if $a'Aa$ is less than 1; see, e.g., Yancey, Judge and Bock (1974) for proof.

Result II: The matrix $(aa' - A^{-1})$ cannot be non-negative definite for $p > 1$; see, e.g., Gulkey and Price (1981).

Applying Result I to matrix C given by (3.5), we observe that it cannot be positive definite whence it follows from (3.4) that P_g cannot be superior to P_o

with respect to the criterion of second order moment matrix to the order of our approximation.

Similarly, using Result II, we find that the matrix C cannot be non-negative definite by virtue of our specification that p exceeds 2. In other words, P_o cannot be superior to P_g.

It is thus seen that P_g neither dominates P_o nor is dominated by P_o according to second order moment matrix criterion.

For the comparison of P_g and P_{g*}, we observe from (3.4) that

$$\begin{aligned} D(P_g; P_{g*}) &= E(P_{g*} - T)(P_{g*}T)' - E(P_g - T)(P_g - T)' \\ &= \sigma^4 \frac{4(n-p)(p-2)gk}{(n-p+2)\beta' XX\beta}(g - g^*) \\ &\quad \times \left[\lambda(X'X)^{-1} - \frac{2\lambda + (g+g^*)k}{\beta' X' X\beta}\beta\beta'\right]. \end{aligned} \tag{3.7}$$

Applying Result I and Result II, once again we find no clear dominance of P_g over P_{g*}.

Let us now compare the predictions with respect to the criterion of trace of second order moment matrix to order $O(\sigma^4)$.

From (3.4), we see that

$$trD(P_g; P_o) = \sigma^4 \frac{4(n-p)(p-2)^2 gk}{(n-p+2)\beta' XX\beta}(\lambda - gk) \tag{3.8}$$

which is positive when

$$k < \frac{\lambda}{g}. \tag{3.9}$$

Thus $P_1 \equiv T_{SS}$, $P_\lambda \equiv T_{SL}$ and $P_{1-\lambda} \equiv T_{LS}$ are better than $P_o \equiv T_{LL}$ when k is less than λ, 1 and $(1 - \lambda)$ respectively.

Just the reverse is true, i.e., T_{LL} beats $T_{SS}, T_{SL}$and T_{LS} for k exceeding λ, 1 and $(1 - \lambda)$ respectively which holds true at least so long as k exceeds 1.

Table 1: Choice of k for the superiority of predictions

Predictions	T_{LL}	T_{SS}	T_{SL}	T_{LS}
T_{LL}	$*$	$k > \lambda$	$k > 1$	$k > (1-\lambda)$
T_{SS}	$k < \lambda$	$*$	$k < \frac{\lambda}{1+\lambda}$	$k < \frac{\lambda}{2-\lambda}$
T_{SL}	$k < 1$	$k > \frac{\lambda}{1+\lambda}$	$*$	$k < \lambda \ for \ \lambda > \frac{1}{2}$ $k > \lambda \ for \ \lambda < \frac{1}{2}$
T_{LS}	$k < (1-\lambda)$	$k > \frac{\lambda}{2-\lambda}$	$k < \lambda < \frac{1}{2}$ $k > \lambda < \frac{1}{2}$	$*$
The entry in the $(i,j)^{th}$ cell gives the condition for the superiority of predictions in the i^{th} row over the predictions in the j^{th} column. For example, the entry in $(1,2)^{th}$ cell is which is the condition for the superiority of T_{LL} over T_{SS}.				

Similarly, we observe from (3.7) that

$$trD(P_g; P_{g*}) = \sigma^4 \left[\frac{4(n-p)(p-2)^2}{(n-p+2)\beta' X' X \beta} \right] (g - g*)[\lambda - (g+g*)k] \tag{3.10}$$

which is positive when

$$k < \left(\frac{\lambda}{g+g*} \right) \ for \ \ g > g* \tag{3.11}$$

$$k > \left(\frac{\lambda}{g+g*} \right) \ for \ \ g < g* \, . \tag{3.12}$$

This result can be used to study the relative performance of T_{SL} and T_{LS}The findings are assembled in Table 1.

It is interesting to observe that the superiority conditions presented in the tabular form are simple and easy to use in application.

4 Asymptotic Efficiency Properties Of Predictions Outside The Sample

Let us now consider the predictions for T_f specified by (2.11).

It is easy to see that

$$\begin{aligned} E(T_{fLL} - T_f) &= 0 && (4.1) \\ E(T_{fLL} - T_f)(T_{fLL} - T_f)' &= \sigma^2 \left[(1-\lambda)^2 I_n + X_f (X'X)^{-1} X_f' \right]. && (4.2) \end{aligned}$$

Assuming g to be different from zero, we observe from (2.16) that P_{fg} is not weakly unbiased for T_f unlike $P_{fo} \equiv T_{fLL}$ as is evident from (4.1).

Theorem II: The asymptotic approximation for the bias vector of P_{fg} to order $O(\sigma^2)$ is

$$\begin{aligned} B(P_{fg}) &= E(P_{fg} - T_f) \\ &= -\sigma^2 \left[\frac{2(n-p)(p-2)g_f k}{(n-p+2)\beta' X' X \beta} \right] X_f \beta && (4.3) \end{aligned}$$

and the difference between the second order moment matrices of P_{fg} and $P_{fo} \equiv T_{fLL}$ is

$$\begin{aligned} D(P_{fg}; P_{fo}) &= E(P_{fo} - T_f)(P_{fo} - T_f)' - E(P_{fg} - T_f)(P_{fg} - T_f)' \\ &= \sigma^4 \left[\frac{4(n-p)(p-2)g_f k}{(n-p+2)\beta' X' X \beta} \right] X_f C_f X_f' && (4.4) \end{aligned}$$

where

$$C_f = (X'X)^{-1} - \frac{2 + (p-2)g_f k}{\beta' X' X \beta} \beta \beta'. \qquad (4.5)$$

It is interesting to note from (4.3) and (4.4) that bias vector as well as the matrix difference is free from λ, at least to the order of our approximation. This means that our findings arising from these expressions remain valid whether the aim is to predict actual values or average values as well as both.

Taking the criterion to be the norm of bias vector to the order of our approximation, as in (3.6), it is observed that both T_{fSL} and T_{fLS} are better than T_{fSS}. Further, T_{fSL} is better than T_{fLS} when λ is less than 0.5. The reverse is true, i.e., T_{fLS} is better than T_{fSL} when λ exceeds 0.5.

If we choose the criterion to be second order moment matrix to order $O(\sigma^4)$ and use the Results I and II stated in preceding Section, it is found that neither P_{fg} is better than P_{fo} nor vice-versa.

Finally, let us take the criterion as trace of second order moment matrix to order $O(\sigma^4)$. Proceeding in the same manner as indicated in the preceding Section, we can easily find the conditions for the superiority one over the other. These are assembled in Table 2.

First we observe from (4.4) that

$$\begin{aligned} trD(P_{fg};P_{fo}) &= \sigma^4\left[\frac{4(n-p)(p-2)g_f k\beta' X_f' X_f\beta}{(n-p+2)(\beta' X' X\beta)^2}\right] \\ &\quad\times\left[\frac{\beta' X' X\beta}{\beta' X_f' X_f\beta}tr(X'X)^{-1}X_f' X_f - 2 - (p-2)g_f k\right]. \end{aligned} \tag{4.6}$$

The expression on the right hand side is positive when

$$k < \frac{1}{(p-2)g_f}\left[\frac{\beta' X' X\beta}{\beta' X_f' X_f\beta}tr(X'X)^{-1}X_f' X_f - 2\right] \tag{4.7}$$

provided that the quantity in the square brackets is positive.

If we define

$$q_1 = \left(\frac{1}{p-2}\right)\left[\frac{1}{\alpha_1}\sum_{i=2}^{p}\alpha_i - 1\right] \tag{4.8}$$

$$q_p = \left(\frac{1}{p-2}\right)\left[\frac{1}{\alpha_p}\sum_{i=1}^{p-1}\alpha_i - 1\right] \tag{4.9}$$

with $\alpha_1 \leq \alpha_2 \leq \ldots \leq \alpha_p$ denoting the eigenvalues of $X_f' X_f$ in the metric of $X'X$, we observe that the condition (4.7) is satisfied at least as long as

$$k < \frac{q_p}{g_f};\ q_p > 0. \tag{4.10}$$

This condition is easy to verify in any given application. Further, special cases of it provide the conditions stated in Table 2.

On the other hand, the expression on the right hand side of (4.6) is negative so long as

$$k > \frac{q_1}{g_f} \tag{4.11}$$

which lead to the results stated in the first row of Table 2.

Similarly, for the comparison of T_{fSS}, T_{fSL} and T_{fLS}, we observe that

$$\begin{aligned} trD(P_{fg}; P_{fg*}) &= \sigma^4 \left[\frac{4(n-p)(p-2)k\beta' X_f' X_f \beta}{(n-p+2)(\beta' X' X \beta)^2} \right] (g_f - g_f^*) \\ &\times \left[\frac{\beta' X' X \beta}{\beta' X_f' X_f \beta} tr(X'X)^{-1} X_f' X_f - 2 - k(p-2)(g_f + g_f^*) \right]. \end{aligned} \tag{4.12}$$

The expression on the right hand side is positive implying the superiority of P_{fg} over P_{fg*} when

$$k < \frac{1}{(p-2)(g_f + g_f^*)} \left[\frac{\beta' X' X \beta}{\beta' X_f' X_f \beta} tr(X'X)^{-1} X_f' X_f - 2 \right]; g_f > g_f^* \tag{4.13}$$

$$k > \frac{1}{(p-2)(g_f + g_f^*)} \left[\frac{\beta' X' X \beta}{\beta' X_f' X_f \beta} tr(X'X)^{-1} X_f' X_f - 2 \right]; g_f < g_f^* \tag{4.14}$$

provided that the expression on the right hand side of (4.13) is positive.

The conditions (4.13) and (4.14) will surely be satisfied so long as

$$k < \left(\frac{q_p}{g_f + g_f^*} \right);\ q_p > 0;\ g_f > g_f^* \tag{4.15}$$

$$k < \left(\frac{q_1}{g_f + g_f^*} \right);\quad g_f > g_f^*. \tag{4.16}$$

These conditions provide the remaining entries in Table 2.

Table 2: Choice of k for the superiority of predictions

Predictions	T_{fLL}	T_{fSS}	T_{fSL}	T_{fLS}
T_{fLL}	*	$k > q_1$	$k > \frac{q_1}{\lambda}$	$k > \frac{q_1}{1-\lambda}$
T_{fSS}	$k < q_p; q_p > 0$	*	$k < \left(\frac{q_p}{1+\lambda}\right); q_p > 0$	$k < \left(\frac{q_p}{2-\lambda}\right); q_p > 0$
T_{fSL}	$k < \frac{q_p}{\lambda}; q_p > 0$	$k > \left(\frac{q_1}{1+\lambda}\right)$	*	$k < \left(\frac{q_p}{2-\lambda}\right); q_p > 0;\ \lambda > \frac{1}{2}$ $k > \left(\frac{q_1}{2-\lambda}\right);\ \lambda < \frac{1}{2}$
T_{fLS}	$k < \frac{q_p}{(1-\lambda)}; q_p > 0$	$k > \left(\frac{q_1}{2-\lambda}\right)$	$k > \left(\frac{q_1}{2-\lambda}\right); \lambda < \frac{1}{2}$ $k < \left(\frac{q_p}{2-\lambda}\right); q_p > 0; \lambda > \frac{1}{2}$	*

The entry in the $(i,j)^{th}$cell gives the condition for the superiority of predictions in the i^{th} row over the predictions in the j^{th} column as in Table 1.

5 Some Concluding Remarks:

If we take the performance criterion to be total mean squared error, it is well-known that least squares predictions are better than Stein-rule predictions for the actual values of study variables while the opposite is true, i.e., Stein-rule predictions under some mild constraints are better than the least squares predictions for average values of study variable. This observation has prompted us to present two predictions when the objective is to predict both the actual and average values simultaneously.

The proposed predictions are based like Stein-rule predictions. However, if we look at the norms of bias vectors to the order of our approximation, both are found to be superior to Stein-rule predictions. Next, we have compared the predictions according to the criterion of second order moment matrix to the order of our approximation and have found that none of the four predictions is uniformly superior to the other . Finally, taking the criterion as trace of second order moment matrix, we have deduced conditions for the superiority of one over the other and have presented them in a tabular form. These conditions are elegant and easy to apply in developing efficient predictions.

It may be remarked that our investigations can be easy extended on the lines of Ullah, Srivastava and Chandra (1983) to the case when the disturbances are not necessarily normally distributed.

Appendix

In order to find small disturbance asymptotic approximations for the bias vectors and mean squared error matrices, we replace u in (2.1) by σv so that v has a multivariate normal distribution with mean vector 0 and variance covariance matrix I_n. Thus we can express

$$\begin{aligned}\frac{y'\bar{H}y}{y'Hy} &= \sigma^2\frac{v'\bar{H}v}{\beta'X'X\beta}\left[1+2\sigma\frac{\beta'X'v}{\beta'X'X\beta}+\sigma^2\frac{u'Hu}{\beta'X'X\beta}\right]^{-1}\\ &= \sigma^2\frac{v'\bar{H}v}{\beta'X'X\beta}-2\sigma^3\frac{v'\bar{H}v.\beta'X'v}{(\beta'X'X\beta)^2}+O_p(\sigma^4).\end{aligned}$$

Using it in (2.9) and observing that

$$b_L = \beta + \sigma(X'X)^{-1}X'v \tag{5.1}$$

we find

$$\begin{aligned}(P_g - T) &= \sigma(\lambda I_n - \bar{H})v - 2\sigma^2\frac{(p-2)gk\; v'\bar{H}v}{(n-p+2)\beta'X'X\beta}X\beta\\ &\quad -\sigma^3\frac{2(p-2)gk\; v'\bar{H}v}{(n-p+2)\beta'X'X\beta}\left(H-\frac{2}{\beta'X'X\beta}X\beta\beta'X'\right)v+O_p(\sigma^4).\end{aligned}$$

Thus the bias vector of P_g is given by

$$E(P_g - T) = -2\sigma^2\frac{(n-p)(p-2)gk}{(n-p+2)\beta'X'X\beta}+O(\sigma^3) \tag{5.2}$$

which provides the result (3.3) of Theorem I.

Similarly, dropping the terms with expectation as null matrix, the mean squared error matrix of P_g is

$$\begin{aligned}E(P_g - T)(P_g - T)' &= \sigma^2(\lambda I_n - \bar{H})E(vv')(\lambda I_n - \bar{H})\\ &\quad -\sigma^4\frac{2(p-2)gk}{(n-p+2)\beta'X'X\beta}\left[\left(H-\frac{2}{\beta'X'X\beta}X\beta\beta'X'\right)\right.\\ &\quad \times E(v'\bar{H}v.vv')(\lambda I_n - \bar{H}')\\ &\quad +(\lambda I_n - \bar{H})E(v'\bar{H}v.vv')\left(H-\frac{2}{\beta'X'X\beta}X\beta\beta'X'\right)\\ &\quad \left. -\frac{2(p-2)gk}{(n-p+2)\beta'X'X\beta}E(v'\bar{H}v)^2X\beta\beta'X'\right]+O(\sigma^5)\\ &= \sigma^2\left[\lambda^2 I_n + (1-2\lambda)\bar{H}\right]\\ &\quad -\sigma^4\frac{4(n-p)(p-2)gk}{(n-p+2)\beta'X'X\beta}\left[\lambda H - \frac{2\lambda+(p-2)gk}{\beta'X'X\beta}X\beta\beta'X'\right]\\ &\quad +O(\sigma^5)\end{aligned}$$

which leads to the result (3.4) of Theorem I.

The results of Theorem II can be derived in a similar manner.

References

Chaturvedi, Anoop, Alan T. K. Wan, Shri P. Singh (2002): 'Improved multivariate prediction in a general linear model with an unknown error covariance matrix', Journal of Multivariate Analysis, 83, 1, 166182.

Chaturvedi, Anoop and Shri Prakash Singh (2000): 'Stein rule prediction of the composite target function in a general linear regression model', Statistical Papers, 41, 3, 359367.

Chaturvedi, Anoop, Suchita Kesarwani, Ram Chandra (2008): 'Simultaneous prediction based on shrinkage estimator' in Recent Advances in Linear Models and Related Areas, (Eds. Shalabh and C.Heumann), 181204, Springer.

Dube, M. and Veena Manocha (2002): 'Simultaneous prediction in restricted regression models', Journal of Applied Statistical Science, 11, 4, 277288.

Guilkey, D.K. and J.M. Price (1981): 'On comparing restricted least squares estimators' Journal of Econometrics, 15, 397-404.

Rao, C. R., H. Toutenburg, Shalabh and C. Heumann (2008): : Linear Models and Generalizations - Least Squares and Alternatives, Springer.

Judge, G.G. and M.E. Bock (1978): The Statistical Implications of Pre-Test and Stein- Rule Estimators in Econometrics, North-Holland.

Saleh, A.K.Md.E. (2006): Theory of Preliminary Test and Stein-Type Estimation with Applications, Wiley

Shalabh (1995): 'Performance of Stein-rule procedure for simultaneous prediction of actual and average values of study variable in linear regression model' Jan Tinbergen award article in Proceedings of the 50th session of the International Statistical Institute, 1375-1390.

Shalabh, Chandra Mani Paudel and Narinder Kumar (2008): 'Simultaneous prediction of actual and average values of response variable in replicated measurement error models' in Recent Advances in Linear Models and Related Areas, (Eds. Shalabh and C.Heumann), 105133, Springer.

Shalabh and R. Chandra (2002): 'Prediction in Restricted Regression Models', Journal of Combinatorics, Information System and Sciences, Vol. 29, Nos. 1-4, pp. 229-238.

Toutenburg, H. and Shalabh (2000): 'Improved Prediction in Linear Regression Model with Stochastic Linear Constraints', Biometrical Journal, 42, 1, 71-86.

Ullah, A., V.K. Srivastava and R. Chandra (1983): 'Properties of shrinkage estimators in Linear regression when disturbances are not normal' Journal of Econometrics, 21, 389-402.

Yancey, T.A., G.G. Judge and M.E. Bock (1974): 'A mean square test when stochastic restrictions are used in regression' Communications in Statistics, 3, 755-768.