Shalabh, Toutenburg, Heumann:

Performance of Double k-class Estimators for Coefficients in Linear Regression Models with Non Spherical Disturbances under Asymmetric Losses


Online unter: http://epub.ub.uni-muenchen.de/
Performance of Double $k$-class Estimators for Coefficients in Linear Regression Models with Non Spherical Disturbances under Asymmetric Losses

Shalabh $^1$ H. Toutenburg $^2$ C. Heumann $^3$

Discussion Paper 509

Abstract

The risk of the family of feasible generalized double $k$-class estimators under LINEX loss function is derived in a linear regression model. The disturbances are assumed to be non-spherical and their variance covariance matrix is unknown.

1 Introduction

Specification of a suitable loss function is a matter crucial important in analyzing the data and therefrom deducing inferences. Loss functions that are generally employed in statistical practice are taken to be symmetric around zero such as squared error and absolute error loss functions which assign equal weight to positive and negative estimation errors of the same magnitude. They are favoured

$^1$Department of Mathematics and Statistics, Indian Institute of Technology, Kanpur - 208016 (INDIA), E-Mail : shalab@iitk.ac.in, shalabh1@yahoo.com

Acknowledgement: The first author is grateful to the Alexander von Humboldt Foundation for the support. This work is a part of the project under Humboldt Fellowship.

$^2$Department of Statistics, University of Munich, 80799 Munich (Germany), E-mail: Helge.Toutenburg@stat.uni-muenchen.de

$^3$Department of Statistics, University of Munich, 80799 Munich (Germany), E-mail: chris@stat.uni-muenchen.de
because of their comprehensibility and capacity of mathematical manipulations, and not because they are relevant all the time and reflect reality. They represent analyst’s convenience rather than practitioner’s preference. They may fail to capture the salient features of loss structures that are actually faced in practice where, for instance, a positive estimation error of a certain magnitude may have for reaching consequences in comparison to the negative estimation error of the same magnitude; see, e.g., Ashley (1990), Granger (1969), Granger and Newbold (1986), McCloskey (1985), Stockman (1987), Varian (1975) and Zellner and Geisel (1968) for some examples.

The requirement of asymmetry in the losses has led to the development of various asymmetric loss functions but majority of them are not analytically tractable and closed-form expressions for the risk are difficult to obtain. Relatively free from such limitations is the LINEX loss function, formulated by Varian (1975), which incorporates the asymmetric nature of losses in a simple manner and retains the property of analytical tractability at the same time besides having a close link between the traditional squared error loss functions.

In the context of estimation problems in a linear regression model, the LINEX loss function is utilized in a Bayesian framework by Bolfarine (1989), Varian (1975) and Zellner (1986); see also Zellner (1992). Employing the classical (non-Bayesian) framework, performance properties of some estimators under the LINEX loss function are studied by Giles and Giles (1993, 96), Ohtani (1995) and Srivastava and Rao (1992).

It is well known that the non-linear and biased estimators of regression coefficients in a linear regression model can have smaller risk than linear and unbiased estimators under mild constraints, see Judge and Bock (1978, 1983), Judge et al. (1985) and Saleh (2006). One such family is described by the double $k$-class estimators proposed by Ullah and Ullah (1978, 1981). The family of double $k$-class estimators is characterized by two characterizing scalars and encompasses many
estimators as its particular cases, including the Stein rule family of estimators proposed by James and Stein (1961). In the context of linear regression models, the family of double $k$-class estimators was proposed under the assumption of spherical or homoskedastic disturbances. Later, Wan and Chaturvedi (2001) extended it to the case when disturbances are non-spherical or heteroskedastic and their variance covariance matrix is also unknown. They proposed the family of feasible generalized double $k$-class estimators and analyzed the quadratic risk performance of several estimators arising from it as a particular cases under the large sample asymptotic approximation theory. The performance of the feasible generalized double $k$-class estimators under balanced loss function and general Pitman closeness criterion was studied by Chaturvedi and Shalabh (2004). What is the performance of the feasible generalized double $k$-class estimators to estimate the regression coefficients in a linear model under LINEX loss function under non-spherical disturbances with unknown variance covariance matrix constitutes the subject matter of this paper.

The plan of the paper is as follows. The model and the estimators are described in Section 2. The properties of the feasible generalized double $k$-class estimators for the coefficients in a linear regression model with non-spherical disturbances are derived and analyze under the LINEX loss function in Section 3.

2 Model Specification And The Estimators

Consider the following linear regression model with nonspherical disturbances:

$$y = X\beta + \epsilon$$  \hspace{1cm} (2.1)

where $y$ is a $T\times1$ vector of $T$ observations on the study variable, $X$ is a $T\times p$ matrix of $T$ observations on $p$-explanatory variables, $\beta$ is a $p \times 1$ vector of coefficients associated with them and $\epsilon$ is a $T \times 1$ vector of disturbances.
It is assumed that \( \epsilon \) has a multivariate normal distribution with mean vector \( 0 \) and variance covariance matrix \( \sigma^2 \Omega^{-1} \) where \( \sigma^2 \) is an unknown scalar. Further, the elements of \( \Omega \) are functions of an unknown parameter \( \theta \) belonging to an open subset of the \( q \)-dimensional Euclidian space. It is also assumed that a consistent estimator \( \hat{\theta} \) of \( \theta \) is available which permits to obtain a consistent estimator \( \hat{\Omega} \) of \( \Omega \).

The family of double \( k \)-class estimators proposed by Ullah and Ullah (1978) is

\[
\tilde{\beta}_{kk} = \left[ 1 - k_1 \frac{(y - X\tilde{\beta})(y - X\tilde{\beta})}{y'y - k_2(y - X\tilde{\beta})(y - X\tilde{\beta})} \right] \tilde{\beta}
\]

(2.2)

where \( \tilde{\beta} = (X'X)^{-1}X'y \) is the ordinary least squares estimator of \( \beta \), \( k_1 \) and \( k_2 \) are the non-stochastic characterizing scalers.

If we apply the method of generalized least squares for the estimation of \( \beta \) in (2.1) and replace \( \Omega \) by \( \hat{\Omega} \), a feasible generalized least squares (FGLS) estimator of \( \beta \) is given by

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

(2.3)

Similarly, the feasible generalized double \( k \)-class (FGKK) estimators presented by Wan and Chaturvedi (2001) are specified by

\[
\hat{\beta}_{kk} = \left[ 1 - \left( \frac{k_1^*}{T - p + 2} \right) \frac{(y - X\hat{\beta})(y - X\hat{\beta})}{y'y - k_2(y - X\hat{\beta})(y - X\hat{\beta})} \right] \hat{\beta}
\]

(2.4)

where \( k_1^* (> 0) \) and \( k_2 \) are the scalars characterizing the estimator.

It may be observed that the scalar \( k_1 \) in (2.2) and \( k_1^* \) in (2.4) are related by \( k_1 = k_1^*/(T - p + 2) \) which is due to Vinod and Srivastava (1995) who have established that the estimators arising from the family of double \( k \)-class estimators are equivalent if

\[
k_1 = k_1^*T^{-(j + \frac{1}{2})}
\]

where \( j \) is any positive number and \( k_1^* \) is a fixed scalar independent of \( T \).
The family (2.4) is quite flexible and encompasses several interesting estimations as special cases. For example, if we set $k_1^* = 0$, we get the FGLS estimator. Similarly, if we put $k_2 = 1$, we obtain the feasible generalized Stein-rule (FGSR) estimator of Chaturvedi and Shukla (1990). If we take $k_1^* = [1 + 2(T - p)^{-1}]$ and $k_2 = [1 - (T - p)^{-1}]$, we get the feasible generalized minimum mean squared error (FGMMS) estimator in the spirit of Farebrother (1975) while if we take $k_1^* = [1 + 2(T - p)^{-1}]$ and $k_2 = [1 - p(T - p)^{-1}]$, we find the adjusted feasible generalized minimum mean squared error (AFGMMSE) estimator in the light of Ohtani (1996). Stemming from the work reported in Carter, Srivastava and Chaturvedi (1993), another interesting estimator is specified by $k_1^* = (p - 2)$ and $k_2 = [1 - (p - 2)(T - p + 2)^{-1}]$ which can be abbreviated as FGKKCSC estimator; see Wan and Chaturvedi (2001, Sec. 4) for some other choices of $k_1^*$ and $k_2$.

3 Properties of Estimators

Assuming that the explanatory variables are asymptotically cooperative in the sense that the matrix $T^{-1}X'\Omega X$ tends to a finite nonsingular matrix as $T$ tends to infinity, it is found that the limiting distributions of the FGLS and FGKK estimators are identical, and thus the superiority of any estimators over the other cannot be examined on the basis of limiting distribution. On the other hand, if we consider their exact distributions, it can be well appreciated that they are difficult to derive. Even if we succeed in the derivation of exact distributions, they will be sufficiently complex and will not permit us to draw any clear inference.

The large sample properties of the FGKK estimators have been extensively studied by Wan and Chaturvedi (2001). In particular, they have derived the asymptotic distribution of the estimators under fairly general conditions. Also presented are the expressions for bias vector to order $O(T^{-1})$ and mean squared error matrix to order $O(T^{-2})$. Further, taking the performance criterion as risk
under a symmetric general quadratic loss function to order $O(T^{-2})$, they have compared various estimators and have found the conditions for the superiority of one estimator over the other. Various choices of characterizing scalars are also discussed. Finally, assuming the disturbances to follow a first order autoregressive scheme, the results of Monte Carlo experiment are reported.

Let us now analyze the performance of FGKK estimators with respect to the criterion of risk under LINEX has function which is asymmetric. For this purpose, let us consider the estimation of a linear parametric function $g'\beta$ where $g$ is any arbitrary column vector with known and nonstochastic elements. If we take all the elements of $g$ as unity, $g'\beta$ is equal to the sum of regression coefficients. Similarly, if we assume all the elements of $g$ to be zero except the $i^{th}$ element as unity, $g'\beta$ reduces to the regression coefficient associated with the $i^{th}$ explanatory variable in the model.

The LINEX loss function, introduced by Varian (1975), for the estimator of any scalar parameter $\delta$ by an estimator $\hat{\delta}$ is defined as

$$L(\hat{\delta}; \delta) = c \left[ \exp\left\{\alpha(\hat{\delta} - \delta)\right\} - \alpha(\hat{\delta} - \delta) - 1 \right]$$ (3.1)

where $\alpha$ and $c$ are the characterizing scalars with non-zero $\alpha$ and positive $c$.

The values of $c$ specifies the factor of proportionality while the value of $\alpha$ determines the relative losses associated with the positive and negative values of the estimation error $(\hat{\delta} - \delta)$. The LINEX loss function attains its minimum value as zero when $(\hat{\delta} - \delta) = 0$. Further, it rises approximately linearly on one side of zero and exponentially on the other side of zero. Zellner (1986) has prepared graphs of the LINEX loss function for some selected values of $\alpha$ and has observed that over-estimation of a certain magnitude leads to larger loss in comparison to the under-estimation of the same magnitude for positive values of $\alpha$ while the reverse is true for negative values of $\alpha$, i.e., over-estimation leads to comparatively smaller loss than under-estimation. Thus possibly unequal weight to under-estimation and over estimation can be assigned through an appropriate sign for the scalar $\alpha$. 
Regrading the magnitude of $\alpha$, the LINEX loss function is fairly symmetric like the squared error loss function for small values of $\alpha$. If the value of $\alpha$ is taken away from zero, asymmetry of loss function increases. In this manner, the behavior of loss function can be suitably tailored depending upon the requirements of problem in hand.

For analyzing the performance properties of FGKK estimators, let us specify the LINEX loss function as follows

$$L(\hat{\beta}; \beta) = \left[ \exp \left\{ \frac{\alpha \sqrt{T}}{\sigma} g'(\hat{\beta} - \beta) \right\} - \frac{\alpha \sqrt{T}}{\sigma} g'(\hat{\beta} - \beta) - 1 \right]$$

(3.2)

where $\hat{\beta}$ denotes an estimator of $\beta$ and the scaling factor $c$ in the LINEX loss function (3.1) is taken as unity without any loss of generality.

Following Wan and Chaturvedi (1998), let us make the following assumptions.

For all $j, k = 1, 2, \ldots, q$, let us define

$$\Omega_j = \frac{\partial \Omega}{\partial \theta_j}, \quad \Omega_{jk} = \frac{\partial^2 \Omega}{\partial \theta_j \partial \theta_k},$$

$$A = \frac{X'\Omega X}{T}, \quad A_j = \frac{X'\Omega_j X}{T}, \quad A_{jk} = \frac{X'\Omega_{jk} X}{T},$$

$$\alpha = \frac{X'\Omega \epsilon}{\sqrt{T}}, \quad \alpha_j = \frac{X'\Omega_j \epsilon}{\sqrt{T}}, \quad \text{and} \quad \alpha_{jk} = \frac{X'\Omega_{jk} \epsilon}{\sqrt{T}}.$$  

Furthermore, the set of matrices (or vectors) having the same number of indices is denoted by boldface letters subscripted in brackets by that number. For instance, $A_{(3)}$ denotes the set of matrices $\{A_{jkl} : j, k, l = 1, 2, \ldots, q\}$. The following regularity conditions are required to have a valid Edgeworth expansion of the distribution:

(i) Each matrix in the sets $A_{(1)}, A_{(2)}, \ldots, A_{(5)}$ and covariance matrix of each vector in $\alpha_{(1)}, \alpha_{(2)}, \ldots, \alpha_{(5)}$ converges to a finite matrix as $T \to \infty$;

(ii) $X'C^2X/T$ is bounded and tends to infinity for all $C$ in $\Omega_{(6)}$;

(iii) The estimator $\hat{\theta}$ has an expansion of the form $\sqrt{T}(\hat{\theta} - \theta) = e + O_p(T^{-1})$ such that the asymptotic distribution of $e$ is multivariate normal with mean vector of order $O(T^{-1})$ and variance covariance matrix as of $\Lambda + O(T^{-1})$.  

7
The third and higher order cumulants of $T^{-\frac{1}{2}} X^t \left( \frac{\partial^2 \Omega}{\partial \theta^2} \right) \epsilon$ are of order $O(T^{-1})$.

Under the above assumptions, Wan and Chaturvedi (2001) have demonstrated that the asymptotic distribution of the vector $Z = \sqrt{T} \sigma (\hat{\beta}_{kk} - \beta)$ is multivariate normal. Further, if we write

$$\phi = \frac{1}{T \sigma^2} \beta^t X^t \Omega X \beta,$$

we have the following results from Wan Chaturvedi (2001):

$$\mu = \frac{\sqrt{T}}{\sigma} E(\hat{\beta}_{kk} - \beta)$$  
$$= -k^*_1 \frac{\beta}{\sqrt{T} \sigma (\phi + 1 - k^*_2)} + O(T^{-1})$$  
$$\Sigma = \frac{T}{\sigma^2} E(\hat{\beta}_{kk} - \beta)(\hat{\beta}_{kk} - \beta)^t$$  
$$= A^{-1} + O(T^{-1}).$$

By virtue of asymptotic normality of $\sqrt{T} \sigma (\hat{\beta}_{kk} - \beta)$, we observe that

$$E \left[ \exp \left\{ \alpha \sqrt{T} \sigma (\hat{\beta}_{kk} - \beta) \right\} \right]$$  
$$= \int_{-\infty}^{\infty} \exp(\alpha g^t Z) \frac{\exp\left\{ -\frac{(g^t Z - g^t \mu)^2}{2g^t \Sigma g} \right\}}{(2\pi g^t \Sigma g)^{\frac{1}{2}}} \, dg^t Z$$  
$$= \exp \left( \alpha g^t \mu + \frac{\alpha^2}{2} g^t \Sigma g \right) \int_{-\infty}^{\infty} \frac{\exp\left\{ -\frac{(g^t Z - g^t \mu - \alpha g^t \Sigma g)^2}{2g^t \Sigma g} \right\}}{(2\pi g^t \Sigma g)^{\frac{1}{2}}} \, dg^t Z$$  
$$= \exp \left( \alpha g^t \mu + \frac{\alpha^2}{2} g^t \Sigma g \right)$$  
$$= \exp \left\{ -\frac{\alpha k^*_1 g^t \beta}{\sqrt{T} \sigma (\phi + 1 - k^*_2)} + O(T^{-1}) \right\} \exp \left( \frac{\alpha^2}{2} g^t A^{-1} g \right)$$  
$$= \left[ 1 - \frac{\alpha k^*_1 g^t \beta}{\sqrt{T} \sigma (\phi + 1 - k^*_2)} + O(T^{-1}) \right] \exp \left( \frac{\alpha^2}{2} g^t A^{-1} g \right).$$

Also,

$$\frac{\alpha \sqrt{T}}{\sigma} E[g^t (\hat{\beta}_{kk} - \beta)] = -\frac{\alpha k^*_1 g^t \beta}{\sqrt{T} \sigma (\phi + 1 - k^*_2)} + O(T^{-1}).$$  
(3.7)
Using these results, the risk associated with FGKK estimator under the LINEX loss function (3.2) to order $O(T^{-\frac{1}{2}})$ is given by

$$R(\text{FGKK}) = E \left[ \exp \left\{ \frac{\alpha' \sqrt{T}}{\sigma} g'(\hat{\beta}_{kk} - \beta) \right\} - \frac{\alpha' \sqrt{T}}{\sigma} g'(\hat{\beta}_{kk} - \beta) - 1 \right]$$

(3.8)

Setting $k_1^* = 0$, we find the risk associated with the FGLS estimator to order $O(T^{-\frac{1}{2}})$ as follows

$$R(\text{FGLS}) = \left[ \exp \left\{ \frac{\alpha^2}{2} g' A^{-1} g \right\} - 1 \right]$$

(3.9)

Comparing (3.8) and (3.9), we find that the FGKK estimator dominate the FGLS estimator when

$$\left( \frac{\alpha g' \beta}{\phi + 1 - k_2} \right) > 0$$

(3.10)

which holds true, for instance, as long as $\alpha$ and $g' \beta$ have same signs and $k_2$ does not exceed 1. An interesting implication this finding is that all FGSR, FGMMSE, AFGMMSE and FGKKCSC estimators will be better than FGLS estimator provided $\alpha$ and $g' \beta$ have same sign.

The opposite is true, i.e., the FGLS estimator dominates the FGKK estimator when the inequality (3.10) holds with a reversed sign. In particular, the FGLS estimator remains unbeaten by the FGSR, FGMMSE, AFGMMSE and FGKKCSC estimators.

Next, let us compare the FGSR estimator characterized by the scalars $k_1^*$ and the FGKK estimator characterized by the scalars $k_1^*$ and $k_2$. It can be easily seen from (3.6) that FGKK estimator dominates the FGSR estimator when

$$\frac{\alpha g' \beta (1 - k_2)}{\phi + 1 - k_2} < 0.$$

(3.11)

When the quantity on the left hand side of inequality (3.11) is positive, the FGKK estimator fails to dominate the FGSR estimator.
If thus follows from (3.10) and (3.11) that the FGKK estimator has better performance than both the FGLS and FGSR estimators when \( k_2 \) lies between 1 and \((1 + \phi)\) provided that \( \alpha \) and \( g'\beta \) both are either negative or positive. When \( \alpha \) and \( g'\beta \) have opposite signs, there does not exist a value of \( k_2 \), given \( k_1^* \), such that FGKK estimator dominates FGLS and FGSR estimators simultaneously.

Now a question arises whether given a FGSR estimator with characterizing scalar as \( k_1^* = k_1 \), can we find a FGKK estimator having superior performance than the given FGSR estimator? The answer is affirmative, and the condition turns out to be as follows:

\[
\left( \frac{\alpha g'\beta}{\phi + 1 - k_2} \right) [k_1^* - \left( 1 + \frac{1 - k_2}{\phi} \right) k_1] > 0. \quad (3.12)
\]

Thus, if the inequality (3.10) holds true, we may choose the characterizing scalar \( k_1^* \) in FGKK such that

\[
k_1^* > \left( 1 + \frac{1 - k_2}{\phi} \right) k_1 \quad (3.13)
\]

and then the FGKK estimator dominates not only the FGSR estimator but FGLS estimator too.

Not so interesting is the case when the condition (3.10) does not hold good so that both FGSR and FGKK are no better than the FGLS estimator. However, in this case, the FGKK succeeds in dominating the FGSR estimator when \( k_1^* \) is chosen to satisfy the condition (3.13) with a reversed inequality sign.

Next, let us restrict our attention to all those members of the feasible generalized double \( k \)-class such that they are specified by \( k_2 < 1 \) and they have better performance than FGLS estimator meaning thereby that the condition (3.10) is satisfied. Now consider two such FGKK estimators. One is FGKK\((k_1^*, k_2)\) specified by \( k_1^* \) and \( k_2 \) while the other is FGKK\((k_1^* + f_1, k_2)\) characterized by \((k_1^* + f_1) \) and \( k_2 \). From (3.8), it is interesting to see that all the FGKK\((k_1^* + f_1, k_2)\) estimators with \( f_1 > 0 \) are more efficient than the FGKK\((k_1^*, k_2)\) estimator. Similarly,
the FGKK($k_1^*, k_2$) estimator is dominated by all the FGKK($k_1^*, k_2 + f_2$) estimators with $0 < f_2 < (\phi + 1 - k_2)$.

Looking at the expression (3.8), we observe that a substantial reduction in the risk under the LINEX loss function to the order of our approximation may be achieved when $\alpha$ and $g'/\beta$ have same signs but large in magnitude, $k_1^*$ is large, $\sigma$ is small and $k_2$ is such that $(\phi + 1 - k_2)$ is positive and small.

Comparing the risks of FGMMSE with FGKK, we find that FGKK is better than FGMMSE when

$$\alpha g'/\beta \left[ \frac{k_1^* \{\phi(T - p) + 1\}}{(T - p + 2)(\phi + 1 - k_2) - 1} \right] > 0. \tag{3.14}$$

So when $\alpha$ and $g'/\beta$ have same signs, the condition (3.14) holds true for all choices of $k_1^*$ and $k_2$ such that

$$\frac{k_1^*}{\phi + 1 - k_2} > \frac{T - p + 2}{\phi(T - p) + 1}. \tag{3.15}$$

In case, $\alpha$ and $g'/\beta$ have opposite signs then FGKK is still better than FGMMSE as long as (3.15) is satisfied with a reverse inequality sign.

Similarly, FGKK has smaller risk than AFGMMSE when

$$\frac{k_1^*}{\phi + 1 - k_2} > \frac{(T - p + 2)p}{\phi(T - p) - p} \tag{3.16}$$

provided $\alpha$ and $g'/\beta$ have same signs. The reverse holds true if (3.16) holds true with a reverse inequality sign.

Lastly, we compare the risks of FGKK and FGKKCSC. The FGKK estimators have smaller risk than FGKKCSC when $\alpha$ and $g'/\beta$ have same signs and the characterizing scalars $k_1^*$ and $k_2$ are chosen to satisfy

$$\frac{\phi + 1 - k_2}{k_1^*} > \left( \frac{\phi}{T - p + 2} - \frac{1}{T - p + 2} \right). \tag{3.17}$$

The reverse holds true if (3.17) holds true with a reverse inequality sign.

In a similar manner, we can compare the FGKK estimator with other specific members arising from the family of the double $k$-class estimators and conditions for the superiority of one estimator over the other can be deduced.
References


