Toutenburg, Shalabh:

A note on the comparison of minimax linear and mixed regression estimation of regression coefficients when prior estimates are available

Sonderforschungsbereich 386, Paper 238 (2001)

A note on the comparison of minimax linear and mixed regression estimation of regression coefficients when prior estimates are available

Shalabh
Department of Statistics
Panjab University

H. Toutenburg
Institute of Statistics
University of Munich

23rd May 2001

Abstract

When prior estimates of regression coefficients along with their standard errors or their variance covariance matrix are available, they can be incorporated into the estimation procedure through minimax linear and mixed regression approaches. It is demonstrated that the mixed regression approach provides more efficient estimators, at least asymptotically, in comparison to the minimax linear approach with respect to the criterion of variance covariance matrix.

keywords: linear regression; mixed model

1 Introduction

Recent studies, past experience and pilot investigations often provide some useful information in the form of estimates of coefficients in a linear regression model along with their standard errors or their variance covariance matrix. Such a prior information can be incorporated into the estimation procedure in two simple ways based on point and interval estimation of parameters. One is to express it as a set of stochastic linear restrictions and to apply the method of mixed regression estimation. And the other is to form a confidence ellipsoid for the regression coefficients and to apply the method of minimax linear estimation. The latter approach generally does not lead to a simple form of the estimators and iterative procedures are to be followed; see, e.g., Schipp (1990) for an interesting exposition. However, if we take the matrix involved in the quadratic loss function to be of rank one, the estimators have a closed form; see, e.g., Rao and Toutenburg (1999, Chap.3) . Here we restrict our attention to such estimators.

Superiority of estimators arising from both the mixed regression and minimax linear estimation procedures over the conventional estimators ignoring the prior information is well discussed in the literature but explicit attention does not seem to have been paid to the relative efficiency of one approach over the other; see, e.g., Rao and Toutenburg (1999) for the details. This has formed the subject matter of this note. Taking the performance criterion to be the variance covariance matrix it is found that the mixed regression approach provides more efficient estimators than the minimax linear approach at least asymptotically.

2 Main Result

Consider a linear regression model

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

wehere y is a $n \times 1$ vector of n observations on the study variable, X is a $n \times K$ matrix of n observations on K explanatory variables, β is a $K \times 1$ vector of regression coefficients and u is a $n \times 1$ vector of disturbances following a multivariate normal distribution with null mean vector as 0 and variance covariance matrix as σ^2 times an identity matrix.

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

Further, let us assume to be given the prior information specifying an unbiased estimate b_0 along with the variance covariance matrix $s_0^2(X_0'X_0)^{-1}$ based on n_0 observations from some extraneous source. This prior information can be expressed in two forms for the purpose of utilizing it in the estimation of β. From the viewpoint of point estimation, we may write it as

$$b_0 = \beta + v \tag{2.2}$$

where v is a $K \times 1$ random vector with mean vector as 0 and variance covariance matrix as $s_0^2(X_0'X_0)^{-1}$.

Alternatively, from the viewpoint of interval estimation, we can formulate it as a confidence ellipsoid given by

$$(b_0 - \beta)'X_0'X_0(b_0 - \beta) \leq s_0^2 n_0 C \tag{2.3}$$

where C denotes the $(1-\alpha)$ quantile of the χ^2 distribution with n_0 degrees of freedom with $(1-\alpha)$ indicating the level of confidence like 95% or 99%.

The least squares estimator of β in (2.1) is given by

$$b = (X'X)^{-1}X'y \tag{2.4}$$

which ignores the prior information all together.

If we use the formulation (2.2) of the prior information and employ the method of mixed regression introduced by Theil and Goldberger (1961), the estimator of β is given by

$$\hat{\beta} = \left(X'X + \frac{s^2}{s_0^2}X_0'X_0\right)^{-1}\left(X'y + \frac{s^2}{s_0^2}X_0'X_0 b_0\right) \tag{2.5}$$

where

$$s^2 = \frac{1}{n-K}(y - Xb)'(y - Xb) \tag{2.6}$$

Similarly, if we consider a quadratic loss function with loss matrix of rank one and apply the method of minimax linear estimation using the formulation (2.3) of the prior information on the lines of Kuks and Olman (1972), the following estimator of β is obtained

$$\tilde{\beta} = \left(X'X + \frac{s^2}{s_0^2 n_0 C} X_0' X_0 \right)^{-1} \left(X'y + \frac{s^2}{s_0^2 n_0 C} X_0' X_0 b_0 \right) \tag{2.7}$$

which can be termed as pseudo minimax estimator; see Rao and Toutenburg (1999, Chap. 3) for the details.

Efficiency properties of $\hat{\beta}$ and $\tilde{\beta}$ in relation to b are well discussed in the literature but a comparison of $\hat{\beta}$ and $\tilde{\beta}$ does not seem to have been made; see, e.g., Rao and Toutenburg (1999) for an interesting account.

It can be easily seen following Kakwani (1968) that all the three estimators b, $\hat{\beta}$ and $\tilde{\beta}$ are unbiased. Further, the variance covariance matrix of b is given by

$$\begin{aligned} \mathrm{V}(b) &= \mathrm{E}(b-\beta)(b-\beta)' \\ &= \sigma^2 (X'X)^{-1}. \end{aligned} \tag{2.8}$$

Exact expressions for the variance covariance matrices of $\hat{\beta}$ and $\tilde{\beta}$ can be obtained following Swamy and Mehta (1969) but their comparison fails to provide any clear inference regarding their efficiency. We therefore consider their large sample approximations. For this purpose, we assume that the explanatory variables are asymptotically cooperative so that the limiting form of the matrix $n^{-1}X'X$ is finite and nonsingular.

Using (2.1) and (2.2), we observe that

$$\begin{aligned} \mathrm{V}(\hat{\beta}) &= \mathrm{E}\left(\hat{\beta}-\beta\right)\left(\hat{\beta}-\beta\right)' \\ &= \mathrm{E}\left(X'X + \frac{s^2}{s_0^2} X_0'X_0 \right)^{-1} \left[X'uu'X + \frac{s^4}{s_0^4} X_0'X_0 vv' X_0' X_0 \right] \\ &\quad \cdot \left(X'X + \frac{s^2}{s_0^2} X_0'X_0 \right)^{-1} \qquad (2.9) \\ &= \mathrm{E}\left(X'X + \frac{s^2}{s_0^2} X_0'X_0 \right)^{-1} \left[\sigma^2 X'X + \frac{s^4}{s_0^2} X_0'X_0 \right] \left(X'X + \frac{s^2}{s_0^2} X_0'X_0 \right)^{-1} \end{aligned}$$

by virtue of mutual independence of $X'u, s^2$ and v.

If we assume that the limiting form of the matrix $n^{-1}X'X$ is finite and nonsingular, it is shown in the Appendix that

$$\mathrm{V}(\hat{\beta}) = \sigma^2 (X'X)^{-1} - \frac{\sigma^4}{s_0^2} (X'X)^{-1} X_0'X_0 (X'X)^{-1} + O\left(n^{-3}\right). \tag{2.10}$$

Similarly, the variance covariance matrix of the estimator $\tilde{\beta}$ is given by

$$\begin{aligned} \mathrm{V}(\tilde{\beta}) &= \sigma^2 (X'X)^{-1} - \frac{\sigma^4}{s_0^2 n_0 C} \left(2 - \frac{1}{n_0 C} \right) (X'X)^{-1} X_0'X_0 (X'X)^{-1} \\ &\quad + O(n^{-3}). \end{aligned} \tag{2.11}$$

From (2.8), (2.10) and (2.11), we observe that both $\hat{\beta}$ and $\tilde{\beta}$ are more efficient than b with respect to the criterion of variance covariance matrix to order $O(n^{-2})$.

Further, we have

$$\mathrm{V}(\tilde{\beta}) - \mathrm{V}(\hat{\beta}) = \frac{\sigma^4}{s_0^2}\left(1 - \frac{1}{n_0 C}\right)^2 (X'X)^{-1}X_0'X_0(X'X)^{-1} \tag{2.12}$$

whence it follows that the variance covariance matrix of $\tilde{\beta}$ exceeds the variance covariance matrix of $\hat{\beta}$ by a positive definite matrix implying the superiority of the mixed regression approach over the minimax linear approach for the estimation of regression coefficients.

It may be remarked that a similar comparison is made by Toutenburg and Srivastava (1996) but they deal with the case of interval constraints and consequently the estimators arising from the frameworks of mixed regression and minimax linear estimators do not have the same efficiency properties as $\hat{\beta}$ and $\tilde{\beta}$ in the present context.

Let us write

$$\epsilon = (s^2 - \sigma^2)$$

so that ϵ is of order $O_p(n^{-\frac{1}{2}})$.

If g is any fixed scalar, we can express

$$\begin{aligned}
&\left(X'X + \frac{s^2}{s_0^2 g}X_0'X_0\right)^{-1} \\
&= (X'X)^{-1}\left[I_K + \frac{\sigma^2}{s_0 g}X_0'X_0(X'X)^{-1} + \frac{\epsilon}{s_0 g}X_0'X_0(X'X)^{-1}\right]^{-1} \\
&= (X'X)^{-1}\left[I_K - \frac{\sigma^2}{s_0 g}X_0'X_0(X'X)^{-1} - \frac{\epsilon}{s_0 g}X_0'X_0(X'X)^{-1} + \dots\right] \\
&= (X'X)^{-1} - \frac{\sigma^2}{s_0 g}(X'X)^{-1}X_0'X_0(X'X)^{-1} - \frac{\epsilon}{s_0 g}(X'X)^{-1}X_0'X_0(X'X)^{-1} + O_p(n^{-3})
\end{aligned}$$

Using it along with

$$\begin{aligned}
&\left(\sigma^2 X'X + \frac{s^4}{s_0^2 g^2}X_0'X_0\right) \\
&= \sigma^2 X'X + \frac{\sigma^4}{s_0^2 g^2}X_0'X_0 + \frac{2\sigma^2\epsilon}{s_0^2 g^2} + \frac{\epsilon^2}{S_0^2 g^2}X_0'X_0
\end{aligned}$$

we find that

$$\begin{aligned}
&\mathrm{E}\left(X'X + \frac{s^2}{s_0^2 g}X_0'X_0\right)^{-1}\left(\sigma^2 X'X + \frac{s^4}{s_0^2 g^2}X_0'X_0\right)\left(X'X + \frac{s^2}{s_0^2 g}X_0'X_0\right)^{-1} \\
&= \sigma^2(X'X)^{-1} - \frac{\sigma^4}{s_0^2 g}(2 - \frac{1}{g})(X'X)^{-1}X_0'X_0(X'X)^{-1} + O(n^{-3}).
\end{aligned}$$

Substituting $g = 1$ and $g = n_0 C$, we obtain the expression (2.10) and (2.11).

References

Kakwani, N. C. (1968). Note on the unbiasedness of mixed regression estimation, *Econometrica* **36**: 610–611.

Kuks, J. and Olman, W. (1972). Minimax linear estimation of regression coefficients (II) (in Russian), *Iswestija Akademija Nauk Estonskoj SSR* **21**: 66–72.

Rao, C. R. and Toutenburg, H. (1999). *Linear Models: Least Squares and Alternatives*, 2 edn, Springer, New York.

Schipp, B. (1990). *Minimax Schätzer im simultanen Gleichungsmodell bei vollständiger und partieller Vorinformation*, Hain, Frankfurt/M.

Swamy, P. A. V. B. and Mehta, J. S. (1969). On Theil's mixed regression estimator, *Journal of the American Statistical Association* **64**: 273–276.

Theil, H. and Goldberger, A. S. (1961). On pure and mixed estimation in econometrics, *International Economic Review* **2**: 65–78.

Toutenburg, H. and Srivastava, V. K. (1996). Estimation of regression coefficients subject to interval constraints, *Sankhya, Series A* **58**: 273–282.