Available online at www.sciencedirect.com
ScienceDirect
Statistics & Probability Letters 78 (2008) 109–117
STATISTICS & PROBABILITY LETTERS
www.elsevier.com/locate/stapro

MSE superiority of Bayes and empirical Bayes estimators in two generalized seemingly unrelated regressions ☆

Lichun Wang[a,*], Noël Veraverbeke[b]

[a]Department of Mathematics, Beijing Jiaotong University, Beijing 100044, China
[b]Center for Statistics, Hasselt University, 3590 Diepenbeek, Belgium

Received 11 October 2006; received in revised form 21 March 2007; accepted 23 May 2007
Available online 8 June 2007

Abstract

This paper deals with the estimation problem in a system of two seemingly unrelated regression equations where the regression parameter is distributed according to the normal prior distribution $N(\beta_0, \sigma_\beta^2 \Sigma_\beta)$. Resorting to the covariance adjustment technique, we obtain the best Bayes estimator of the regression parameter and prove its superiority over the best linear unbiased estimator (BLUE) in terms of the mean square error (MSE) criterion. Also, under the MSE criterion, we show that the empirical Bayes estimator of the regression parameter is better than the Zellner type estimator when the covariance matrix of error variables is unknown.

MSC: primary: 62J05, 62F11; secondary: 62C10, 62C12

Keywords: Bayes method; Seemingly unrelated regressions; Covariance adjusted approach; Mean square error criterion

1. Introduction

The seemingly unrelated regression system was first introduced by Zellner (1962, 1963) and later developed by Kementa and Gilbert (1968), Mehta and Swamy (1976) and Wang (1988), etc. Recently, Gao and Huang (2000) established some finite sample properties of the Zellner estimator in the context of m seemingly unrelated regression equations, whereas Liu (2000) proposed a two stage estimator and proved its superiorities over the ordinary least square estimator and Zellner type estimator under mean square error matrix criterion.

Differing from the past works, in this paper we employ the Bayes and empirical Bayes approach to construct the estimators of the regression parameter and exhibit their MSE properties. Also, differing from the above regressions, here we do not make the same dimension assumption of observation vectors.

A system of two generalized seemingly unrelated regression equations is given by

$$y_1 = X_1\beta + u_1, \quad y_2 = X_2\gamma + u_2, \tag{1.1}$$

☆Partly supported by NNSF of China (10571001) and IUAP(P6/03) of Belgium.
*Corresponding author.
E-mail address: wlc@amss.ac.cn (L. Wang).

0167-7152/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.spl.2007.05.008

where y_1 and y_2 are $m \times 1$ and $n \times 1$ vectors of observations ($m \neq n$, without loss of generality, let $m > n$), X_1 and X_2 are $m \times p_1$ and $n \times p_2$ matrices with full column rank, β and γ are vectors of unknown parameters, u_1 and u_2 are $m \times 1$ and $n \times 1$ vectors of error variables, and

$$E(u_1) = 0, \quad E(u_2) = 0,$$
$$Cov(u_1, u_1) = \sigma_{11} I_m, \quad Cov(u_2, u_2) = \sigma_{22} I_n,$$
$$Cov(u_1, u_2) = \sigma_{12} \begin{pmatrix} I_n \\ 0 \end{pmatrix}, \quad Cov(u_2, u_1) = \sigma_{21}(I_n \vdots 0),$$

where $\Sigma^* = (\sigma_{ij})$ is a 2×2 non-diagonal positive definite matrix. Such a system (usually $m = n$) appears in many research fields and has received considerable attention including the above authors and Chen (1986), Lin (1991) and so on.

Denote $y = (y_1', y_2')'$, $X = \mathrm{diag}(X_1, X_2)$, $\alpha = (\beta', \gamma')'$, $u = (u_1', u_2')'$, $\Sigma_{ij} = Cov(u_i, u_j)$. Then (1.1) can be represented as

$$y = X\alpha + u, \quad E(u) = 0, \quad Cov(u) = \Sigma, \tag{1.2}$$

where $\Sigma = (\Sigma_{ij})_{2 \times 2}$ is a partitioned matrix.

In what follows, our main concern is how to estimate β better. To adopt the Bayes and empirical Bayes approach, we assume that the prior distribution of the parameter β is

$$\beta \sim \mathrm{N}(\beta_0, \sigma_\beta^2 \Sigma_\beta), \tag{1.3}$$

where Σ_β is a positive definite matrix (namely $\Sigma_\beta > 0$), β_0 and σ_β^2 are hyper-parameters.

Furthermore, assume

$$u \sim \mathrm{N}(0, \Sigma). \tag{1.4}$$

It follows from (1.3) and (1.4), that the posterior density of β given y_1 is (see Wang and Chow, 1994)

$$f(\beta | y_1) \propto \exp\left\{ -\frac{1}{2\sigma_{11}} (\beta - \bar{\beta})' \bar{\Sigma}^{-1} (\beta - \bar{\beta}) \right\}, \tag{1.5}$$

where

$$\bar{\beta} = \bar{\Sigma}(X_1' X_1 \hat{\beta} + \lambda \Sigma_\beta^{-1} \beta_0), \tag{1.6}$$

$\bar{\Sigma} = (X_1' X_1 + \lambda \Sigma_\beta^{-1})^{-1}$, $\lambda = \sigma_{11}/\sigma_\beta^2$ and $\hat{\beta} = (X_1' X_1)^{-1} X_1' y_1$. Thus, under any quadratic loss, the Bayes estimator (BE) of the parameter β would be the posterior expectation of β with given y_1, i.e.,

$$\hat{\beta}_{\mathrm{BE}} = E(\beta | y_1) = \bar{\beta}. \tag{1.7}$$

It is clear that $\hat{\beta}_{\mathrm{BE}}$ only contains the information of the first equation in the regressions (1.1) but that it does not make most use of all information of regressions since $\sigma_{12} \neq 0$.

As we know the estimation problems arise in many situations in statistics. An important concept is the minimum variance unbiased estimation (MVUE) and an interesting result is how to judge whether an estimator is MVUE or not: Let $\hat{g}(x)$ be an unbiased estimator (UE) of $g(\theta)$, and $Var_\theta(\hat{g}(x)) < \infty$, then $\hat{g}(x)$ is MVUE if and only if $Cov_\theta(\hat{g}(x), l(x)) = 0$ for any $\theta \in \Theta$ (parameter space), where $l(x)$ denotes any UE of zero. Obviously, if there exists an UE $l_0(x)$ of zero such that $Cov_\theta(\hat{g}(x), l_0(x)) \neq 0$, then $\hat{g}(x)$ must not be the MVUE of its mean. However, a problem is how we utilize the relationship between $l_0(x)$ and $\hat{g}(x)$ to obtain the MVUE of $g(\theta)$. Rao (1967) introduces the covariance adjusted approach to propose a UE of $g(\theta)$ whose variance is less than $\hat{g}(x)$, which is a linear combination of $\hat{g}(x)$ and $l_0(x)$.

In the followings, by virtute of the covariance adjustment technique, firstly, we use an UE of zero to improve $\hat{\beta}_{\mathrm{BE}}$ and get $\hat{\beta}_{\mathrm{BE}}^{(1)}$, secondly, we adjust $\hat{\beta}_{\mathrm{BE}}^{(1)}$ by another UE of zero. Repeating this process, we obtain the best BE of the parameter β, which contains all information of β in the regressions (1.1), and prove its MSE superiority over the BLUE of β. When σ_{ij} $(i,j = 1,2)$ and the hyper-parameters are unknown, we replace them by their consistent estimators in the best BE of β and present the corresponding empirical Bayes (EB) estimator and exhibit its MSE superiority, too.