ISSN 1746-7659, England, UK
Journal of Information and Computing Science
Vol. 13, No. 4, 2018, *pp*.269-282

Confidence ellipsoids for the primary regression coefficients in m-equation seemingly unrelated regression models

Qingli Pan [1] and Chuanlin Zhang [2+]

[1] Department of Mathematics, The college of Information Science and Technology, Jinan University, Guangzhou, China

[2] Department of Mathematics, The college of Information Science and Technology, Jinan University, Guangzhou, China

(Received September 04 2018, accepted November 23 2018)

Abstract. For a m-equation seemingly unrelated regression(SUR) model, this paper derives two basic confidence ellipsoids(CEs) respectively based on the two-stage estimation and maximum likelihood estimation(MLE), and corrects the two CEs using the Bartlett correction method, resulting in four new CEs. In the meantime via using the partition matrix, we derive a new matrix-derivative-based formulation of Fisher's information matrix for calculating the maximum likelihood estimator of the m-equation SUR model. By Monte Carlo simulation, the coverage probabilities and average volumetric characteristics of CEs are compared under different sample values and different correlation coefficients. Moreover, it is proved that the CE based on the second bartlett correction method performs better even in the case of small samples. Finally, we apply these CEs to the actual data for analysis. The CEs of the SUR model with multiple equations are found to be more accurate than the case with only two equations.

Keywords: Seemingly unrelated regression, Confidence ellipsoids, Bartlett correction, Maximum likelihood estimation.

1. Introduction

The seemingly unrelated regressions (SUR) model, proposed by Arnold Zellner[1][2] in 1962, is a statistical model that is adopted broadly in various fields, such as economics, finance, medicine, biology and so on. In nature, the SUR model is a generalization of a linear regression model and consists of several regression equations with correlated error terms. The research on such models mainly uses the correlation of errors as additional information to improve the estimates efficiency of regression coefficients.

Zellner[2] used a generalized least squares approach to solve the regression coefficients of the SUR model. Moreover, a Bayesian estimation approach for this model was also introduced firstly by Zellner in[3]. Revankar[4][5] has proved properties of estimators via using residuals as estimates of the covariance matrix. V.K.Srivastava and A.K.Srivastava[6] have raised an improved estimation which is a convex combination of the ordinary-least-squares(OLS) estimator and a SUR estimator. In addition, V.Srivastava and D.Giles[7] have summarized kinds of estimation approaches and theories of SUR model at that time. Liu[8] has discussed two-stage estimators of the SUR model when the error terms defer to a general elliptical distribution. Zhao and Xu[9] have studied the high dimensional seemingly unrelated regression models recently.

In addition to OLS estimation and Bayesian estimation, scholars also studied another point estimation method - MLE approach. Phillips[10] has derived the exact distribution of the SUR estimators, which distribution is too complicated to use in practice. Thus, many scholars tried to use the approximate method to replace it. Park[11] has determined the suitable SUR-based MLE via using an iterative procedure. What's more, Fraser et al.[12] have analyzed the highly accurate likelihood method for the SUR model. However, this procedure is used only for scalar parameter and not for vector parameters.

Point estimate provides specific estimates of unknown parameters, which is easy to calculate and apply. However, its accuracy needs to be reflected by its distribution. In fact, interval estimation is the most intuitive way to measure the accuracy of a point estimate. In statistics, the confidence region is a multi-dimensional summary of confidence intervals. It is a set of points in n-dimensional space whose shape is usually an ellipsoid around a point. L. Le. Cam[13] has proposed the standard asymptotic confidence ellipsoids of Wald based on Hellinger distances. Lee. C. Adkins and R. Carter. Hill[14] have raised an improved confidence ellipsoid of the Stein-rule estimator for the linear regression model by bootstrap

+ Corresponding author. Tel.: +86-020-38214329; fax: +86-020-38214329.
E-mail address: tclzhang@jnu.edu.cn (P.D.Zhang).

method. Laurent El Ghaoui and Giuseppe Calafiore[15] have given a method of computing confidence ellipsoids for uncertain linear equations with structure. Erik Weyer and M.C.Campi[16] have considered the finite sample properties of least-squares system identification and have derived non-asymptotic confidence ellipsoids for the OLS estimator. However, this paper considers the CEs of the SUR model.

In recent years, the research on two equations of SUR model has made great progress in parameter estimation and statistical inference. However, there are relatively few studies and applications of the SUR model of $m(m>2)$ equations, especially for the study of confidence region of the estimator. Kent R.Riggs et al.[17] have learned confidence ellipsoids for the primary regression coefficients in two seemingly unrelated regression models and have advanced two new confidence ellipsoids. This paper mainly studies the parameter estimation of multiple equations of SUR model and the confidence region of the estimator. Here we considered using two-stage estimation method to estimate the unknown parameters. Meanwhile, according to the MLE covariant parameter method of the SUR model derived from the two equations of Kent R.Riggs et al, we generalize it to m equations of the SUR model, and obtain the MLE estimated covariant by using Fisher's information matrix. By using these two kinds of estimators, Wald statistics are constructed to obtain the corresponding CEs, and then the accuracy of the CEs are modified according to two bartlett correction methods.

Using Monte Carlo simulation, we find that the cover probability of the second modified CE is close to the theoretical value and the accuracy of the second modified CE is similar to that of the other two corresponding ellipsoid in the small sample size. Specifically, the CEs from MLE perform better basically than those which CEs from two-stage estimator in small sample size. At the same time, as the sample size increases, the coverage probabilities and accuracy of CEs increase accordingly. Furthermore, as ρ increases, the coverage probabilities of the CEs increase while the average volumes decrease.

The context of this paper is as follows. In section 2 we give the expression of m-equation SUR model. Then we list some of the commonly used estimators for this model, such as two-stage estimator and MLE in section 3. In section 4 we derive six different CEs via using Wald statistic and Slutsky's Theorem. In section 5 we examine properties of these CEs by using Monte Carlo simulation. In section 6 we use CEs estimation approaches to model a real data set and compare the accuracy under different number of equations. Also, we sum up some simple conclusions in section 7. Finally, we show generalize a matrix-derivative-based formulation of Fisher's information matrix to m equations of the SUR model in the Appendix.

2. The m-equation SUR model

Consider a m-equation SUR model

$$Y_i = X_i\beta_i + \varepsilon_i, i = 1,2,\cdots,m, \tag{1}$$

where $Y_i \in R_{n\times 1}$, $X_i \in R_{n\times p_i}$ with $rank(X_i) = p_i$, $\beta_i \in R_{p_i\times 1}$, $\varepsilon_i \in R_{n\times 1}$ such that $E(\varepsilon_i) = 0$, $COV(\varepsilon_i, \varepsilon_j) = \sigma_{i,j} I_n$, $i, j = 1,2,\cdots,m$, where $n_i > p$ and I_n is the identity matrix with the order n. Let

$$Y \equiv \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_m \end{bmatrix},\ X \equiv \begin{bmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & X_m \end{bmatrix},\ \beta \equiv \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix},\ \varepsilon \equiv \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_m \end{bmatrix}.$$

Then, the model (1) can be rewritten as

$$Y = X\beta + \varepsilon \tag{2}$$

Let

$$\Sigma \equiv \begin{pmatrix} \sigma_{1,1} & \sigma_{1,2} & \cdots & \sigma_{1,m} \\ \sigma_{2,1} & \sigma_{2,2} & \cdots & \sigma_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m,1} & \sigma_{m,2} & \cdots & \sigma_{m,m} \end{pmatrix} = \begin{pmatrix} \sigma_{1,1} & \Sigma_{1,2} \\ \Sigma_{2,1} & \Sigma_{2,2} \end{pmatrix}, \tag{3}$$

$$\rho = \left(\frac{\Sigma_{1,2}\Sigma_{2,2}^{-1}\Sigma_{2,1}}{\sigma_{1,1}} \right)^{\frac{1}{2}}, \tag{4}$$

where $\Sigma_{2,2}$ is $(m-1)\times(m-1)$ matrix, $\sigma_{1,1}$ is first main diagonal element of Σ. $\Sigma_{1,2}$ is $1\times(m-1)$ matrix, $\Sigma_{2,1}$ is $(m-1)\times 1$ matrix. And assumed $\varepsilon \sim N_{m\times n}(0, \Sigma \otimes I_n)$, where $\otimes$ represents the Kronecker product operator.

3. Parameter estimation of the m-equation SUR model

The coefficients of the regression equations in the SUR model (1) can be estimated by using some approaches, and the following is a list of some commonly used estimation methods. In general, due to the limitations of the paper's length, we only consider the estimation of β_1, the case of $\beta_i (i = 1,2,\cdots,m)$ is also similar.

3.1. Two-stage estimator

As we all known, the Gauss-Markov estimator of β proposed by Zellner[1] is

$$\beta^* = \left(X'\left(\Sigma^{-1}\otimes I\right)X\right)^{-1} X'\left(\Sigma^{-1}\otimes I\right)Y, \tag{5}$$

when Σ is known. In practice, Σ is often unknown. A relatively straightforward approach at this point is to use least-squares(LS) estimation

$$\hat{\beta}_{1LS} = \left(X_1' X_1\right)^{-1} X_1' Y_1, \tag{6}$$

which neglects the correlation between regression error vectors. Obviously, the LS estimator does not take full advantage of the sample information, so it is a vital work to seek more effective estimator. Thus, replacing Σ by its estimator $S = \left(s_{i,j}\right)_{m\times m}$. One of the popular estimates is given by

$$s_{i,j} = \frac{Y_i' N Y_j}{n-r}, i, j = 1,2,\cdots,m, \tag{7}$$

where $N = I - X_0\left(X_0' X_0\right)^+ X_0'$, $X_0 = \left(X_1, X_2, \cdots, X_m\right)$ and $r = rank(X_0)$. Assume that $n - r > m$, S is the so called unrestricted estimate. Partition S as

$$S \equiv \begin{pmatrix} s_{1,1} & s_{1,2} & \cdots & s_{1,m} \\ s_{2,1} & s_{2,2} & \cdots & s_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ s_{m,1} & s_{m,2} & \cdots & s_{m,m} \end{pmatrix} = \begin{pmatrix} s_{1,1} & S_{1,2} \\ S_{2,1} & S_{2,2} \end{pmatrix}, \tag{8}$$

where $S_{2,2}$ is $(m-1)\times(m-1)$ matrix. Substituting $\Sigma_{i,j}$ by $S_{i,j}$ yields the two-stage Zellner estimator or Aitken estimator

$$\beta_{TSZ}^* = \left(X'\left(S^{-1}\otimes I\right)X\right)^{-1} X'\left(S^{-1}\otimes I\right)Y. \tag{9}$$

However, it is hard to study the finite sample properties in case of the nonlinear of the two-stage Zellner estimator. Liu and Wang[18] have proposed a two-stage estimator for improved covariance

$$\tilde{\beta}_1 \equiv \hat{\beta}_{1LS} - \left(X_1' X_1\right)^{-1} X_1'\left(S_{1,2} S_{2,2}^{-1} \otimes N_*\right)Y_*, \tag{10}$$

where $N_* \equiv I - X_*\left(X_*' X_*\right)^+ X_*'$, $X_* = \left(X_2, X_3, \cdots, X_m\right)$, $Y_* = \left(Y_2', Y_3', \cdots, Y_m'\right)$ and $\hat{\beta}_{1LS}$, the LS estimator under single equation estimation, is defined by (6). This estimator has very well finite sample properties. According to Liu and Wang's explanation for SUR model (1), $\tilde{\beta}_1$ is unbiased and has covariance matrix

$$COV\left(\tilde{\beta}_1\right) = \sigma_{1,1}\left[\left(X_1' X_1\right)^{-1} - \left(\frac{n-r-1}{n-r-m}\hat{\rho}^2 - \frac{m-1}{n-r-m}\right)M\right], \tag{11}$$

where the matrix $M = (X_1'X_1)^{-1}X_1'N_*X_1(X_1'X_1)^{-1}$ and $\hat{\rho} = \left(\frac{S_{1,2}S_{2,2}^{-1}S_{2,1}}{s_{1,1}}\right)^{\frac{1}{2}}$.

3.2. Maximum likelihood estimator

Due to error items in the SUR model (1) obey the normal distribution, the log-likelihood function under the condition of giving the data Y and the model matrix X is

$$l(\beta,\Sigma \mid Y,X) = -n\ln(2\pi) - \frac{n}{2}\ln|\Sigma| - \frac{1}{2}(Y - X\beta)'(\Sigma^{-1}\otimes I_n)(Y - X\beta) \tag{12}$$

$$= -n\ln(2\pi) - \frac{n}{2}\ln|\Sigma| - \frac{1}{2}tr\left[A\Sigma^{-1}\right] \ , \tag{13}$$

where $A = \{A_{i,j}\} \equiv \{(Y_i - X_i\beta_i)'(Y_j - X_j\beta_j)\}$. If we let $X = \begin{bmatrix} X_1^* & X_2^* & \cdots & X_m^* \end{bmatrix}$, where

$$X_1^* \equiv \begin{bmatrix} X_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \cdots, \ X_m^* \equiv \begin{bmatrix} 0 \\ 0 \\ \vdots \\ X_m \end{bmatrix}, \tag{14}$$

then $Y - X\beta = Y - X_1^*\beta_1 - \cdots - X_m^*\beta_m$, and we can represent (12) as

$$\begin{aligned} &l(\beta,\Sigma \mid Y,X) = -n\ln(2\pi) - \frac{n}{2}\ln|\Sigma| - \\ &\frac{1}{2}(Y - X_1^*\beta_1 - \cdots - X_{\mathrm{m}}^*\beta_{\mathrm{m}})'(\Sigma^{-1}\otimes I_n)(Y - X_1^*\beta_1 - \cdots - X_{\mathrm{m}}^*\beta_{\mathrm{m}}). \end{aligned} \tag{15}$$

Suppose that $X_1, X_2, \cdots, X_m$ are full column rank. The MLE of β_1 and Σ are solve via deriving the likelihood function and make its derivative equal to zero. Let $\hat{\beta}_1, \ \hat{\beta}_2, \cdots, \ \hat{\beta}_{\mathrm{m}}$ and $vec(\hat{\Sigma})$ denote the MLEs of $\beta_1, \beta_2, \cdots, \beta_m$ and Σ, respectively. Meanwhile, $\hat{\beta}_1, \ \hat{\beta}_2, \cdots, \ \hat{\beta}_{\mathrm{m}}$ and $vec(\hat{\Sigma})$ can be obtained by using a numerical approach, which explained by N.Timm[19].

4. The confidence ellipsoids for β_1 in m-equation SUR model

In this section, the CEs for the primary regression coefficients β_1 in the SUR model (1) can be received by constructing Wald statistics of β_1. Moreover, we attach more importance to the inference on β_1 using a CE.

4.1. The confidence ellipsoids for β_1 from two-stage estimator

According to the section 3, $S_{1,2}$ and $S_{2,2}$ are consistent and the improved two-stage estimator $\tilde{\beta}_1$ in (10), proposed by Liu and Wang[18], is linear function of two multivariate normal vectors. We can see from Slutsky's Theorem that

$$\tilde{\beta}_1 \sim N\left(\beta_1, COV\left(\tilde{\beta}_1\right)\right) \tag{16}$$

for sufficiently large n. Thus, the Wald statistic can be built under this conclusion and yield its approximate distribution. Namely,

$$\left(\tilde{\beta}_1 - \beta_1\right)'\left[COV\left(\tilde{\beta}_1\right)\right]^{-1}\left(\tilde{\beta}_1 - \beta_1\right) \sim \chi^2_{p_1}. \tag{17}$$

From above approximate distribution, we can obtain an approximate $100(1\text{-}\alpha)\%$ CE

$$\left(\tilde{\beta}_1 - \beta_1\right)'\left[COV\left(\tilde{\beta}_1\right)\right]^{-1}\left(\tilde{\beta}_1 - \beta_1\right) \le \chi^2_{p_1}(\alpha), \tag{18}$$

where $\chi^2_{p_1}(\alpha)$ is the $100(1\text{-}\alpha)\%$ of a chi-squared distribution with p_1 degrees of freedom. However, the CE for β_1 in (17) may differ relatively from the nominal level for small sample sizes. Therefore, one can use a Bartlett correction for the $(1-\alpha)$ percentile to improve its accuracy. Cribari-Neto and Zarkos[20] have obtained size-corrected critical values to the order n^{-1} for adjusting the test statistics themselves. The chi-squared percentile can be adjusted by Bartlett-type to

$$\chi_{p_1}^{2(adj1)}(\alpha) \equiv \chi_{p_1}^2(\alpha)\left(1+\frac{m+p_1/m+1}{2n_2}+\frac{m-p_1-1}{4n_2(p_1+2)}\chi_{p_1}^2(\alpha)\right), \tag{19}$$

where $n_2 = n - p_1$. Then, a second CE can be built if one decides the values of β_1 that satisfy

$$\left(\tilde{\beta}_1-\beta_1\right)'\left[COV\left(\tilde{\beta}_1\right)\right]^{-1}\left(\tilde{\beta}_1-\beta_1\right) \le \chi_{p_1}^{2(adj1)}(\alpha). \tag{20}$$

Although the above correction method improves the accuracy to some extent, the correction effect is still relatively poor and the coverage probabilities of CEs is still quite different from nominal level in a small sample. Cardeiro and Ferrari[21] have indicated that another size-corrected critical values converge to chi-square at second-order speed. Gauss M.Cordeiro and Sóstenes L.lins[22] have derived some new results of corrected score statistics. Thus, we can adjust a new chi-squared percentile

$$\chi_{p_1}^{2(adj1)}(\alpha) \equiv \chi_{p_1}^2(\alpha)\left(1+\frac{m+p_1/m+1}{2n_2}+\frac{m-p_1-1}{4n_2(p_1+2)}\chi_{p_1}^2(\alpha)\right)+\frac{m-p_1/m-2}{2n_2(p_1+2)(p_1+4)}\left(\chi_{p_1}^2(\alpha)\right)^3. \tag{21}$$

A third approximate $100(1-\alpha)\%$ CE can be constructed when one finds the values of β_1 that satisfy

$$\left(\tilde{\beta}_1-\beta_1\right)'\left[COV\left(\tilde{\beta}_1\right)\right]^{-1}\left(\tilde{\beta}_1-\beta_1\right) \le \chi_{p_1}^{2(adj2)}(\alpha), \tag{22}$$

where $\chi_{p_1}^{2(adj2)}(\alpha)$ is the estimated $(1-\alpha)$ percentile of the statistic (16) determined via a new Bartlett-type correction method.

4.2. The confidence ellipsoids for β_1 from ML estimator

According to the asymptotic theory for MLEs, we can learn that for sufficiently large n,

$$\begin{bmatrix}\hat{\beta}_1\\ \hat{\beta}_2\\ \vdots\\ \hat{\beta}_m\\ \mathrm{vec}(\hat{\Sigma})\end{bmatrix} \sim N_{p+n}\left(\begin{bmatrix}\beta_1\\ \beta_2\\ \vdots\\ \beta_m\\ vec(\Sigma)\end{bmatrix},\left[I(\beta_1,\beta_2,\cdots,\beta_m,\Sigma)\right]^{-1}\right), \tag{23}$$

where $p=\sum_{i=1}^{n} p_i$, $\sim$ denotes `is approximately distributed as', and the covariance matrix, ie, its Fisher information matrix is

$$I(\beta_1,\beta_2,\cdots,\beta_m,\Sigma)=\begin{bmatrix} I_{\beta_1} & I_{\beta_1,\beta_2} & \cdots & I_{\beta_1,\beta_m} & I_{\beta_1,\Sigma}\\ I_{\beta_2,\beta_1} & I_{\beta_2} & \cdots & I_{\beta_2,\beta_m} & I_{\beta_2,\Sigma}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ I_{\beta_m,\beta_1} & I_{\beta_m,\beta_2} & \cdots & I_{\beta_m} & I_{\beta_m,\Sigma}\\ I_{\Sigma,\beta_1} & I_{\Sigma,\beta_2} & \cdots & I_{\Sigma,\beta_m} & I_{\Sigma}\end{bmatrix}=\begin{bmatrix} I_{\beta_1} & I_{1,2} & I_{1,3}\\ I_{2,1} & I_2 & I_{2,3}\\ I_{3,1} & I_{3,2} & I_{\Sigma}\end{bmatrix}, \tag{24}$$

where

$$I_{1,2}=\begin{bmatrix} I_{\beta_1,\beta_2} & \cdots & I_{\beta_1,\beta_m}\end{bmatrix},\ I_{3,2}=\begin{bmatrix} I_{\Sigma,\beta_2} & \cdots & I_{\Sigma,\beta_m}\end{bmatrix},$$

$$I_{2,1}=\begin{bmatrix} I_{\beta_2,\beta_1}\\ \vdots\\ I_{\beta_m,\beta_1}\end{bmatrix},\ I_{2,3}=\begin{bmatrix} I_{\beta_2,\Sigma}\\ \vdots\\ I_{\beta_m,\Sigma}\end{bmatrix},\ I_2=\begin{bmatrix} I_{\beta_2} & \cdots & I_{\beta_2,\beta_m}\\ \vdots & \ddots & \vdots\\ I_{\beta_m,\beta_2} & \cdots & I_{\beta_m,\beta_m}\end{bmatrix},$$

$I_{1,3} = I_{\beta_1,\Sigma}$ and $I_{3,1} = I_{\Sigma,\beta_1}$. Following the (23), we also have that $\hat{\beta}_1 \sim N_{p_1}\left(\beta_1, I^{\beta_1}\right)$, where I^{β_1} is the first main diagonal element of the covariance matrix. Via using the partition matrix, we derive a new matrix-derivative-based formulation of Fisher's information matrix for calculating the covariance matrix, and the derivation process is given in the Appendix. Hence,

$$I^{\beta_1} = \left[I_{\beta_1} - I_{1,2}I_2^{-1}I_{2,1}\right]^{-1}. \tag{25}$$

From Slutsky's Theorem, we know that

$$\left(\hat{\beta}_1 - \beta_1\right)'\left[\hat{I}^{\beta_1}\right]^{-1}\left(\hat{\beta}_1 - \beta_1\right) \dot{\sim} \chi^2_{p_1} \tag{26}$$

because $\hat{I}^{\beta_1}$ is a weakly consistent estimator of I^{β_1}. Therefore, an approximate $100(1-\alpha)\%$ CE can be constructed if the values of β_1 satisfies

$$\left(\hat{\beta}_1 - \beta_1\right)'\left[\hat{I}^{\beta_1}\right]^{-1}\left(\hat{\beta}_1 - \beta_1\right) \dot{\sim} \chi^2_{p_1}(\alpha), \tag{27}$$

where $\chi^2_{p_1}(\alpha)$ is the $100(1-\alpha)\%$ of a chi-squared distribution with p_1 degrees of freedom.

Also, an approximate $100(1-\alpha)\%$ CE can be constructed by using (19) if one determines the values of β_1 that satisfy

$$\left(\hat{\beta}_1 - \beta_1\right)'\left[\hat{I}^{\beta_1}\right]^{-1}\left(\hat{\beta}_1 - \beta_1\right) \dot{\sim} \chi^{2(adj1)}_{p_1}(\alpha), \tag{28}$$

where $\chi^{2(adj1)}_{p_1}(\alpha)$ is defined in (19).

A final approach to build CE by applying (21). This approximate $100(1-\alpha)\%$ CE can be showed when the values of β_1 satisfies

$$\left(\hat{\beta}_1 - \beta_1\right)'\left[\hat{I}^{\beta_1}\right]^{-1}\left(\hat{\beta}_1 - \beta_1\right) \dot{\sim} \chi^{2(adj2)}_{p_1}(\alpha). \tag{29}$$

5. Monte Carlo simulation

Next we use Monte Carlo simulation to simulate the coverage probabilities and average volumes of the above CEs. Noting that the notations such as $CE_{ts}(\beta_1)$, $CE_{ts}^{adj1}(\beta_1)$, $CE_{ts}^{adj2}(\beta_1)$, $CE_{ml}(\beta_1)$, $CE_{ml}^{adj1}(\beta_1)$ and $CE_{ml}^{adj2}(\beta_1)$ represent the six new CEs and are defined in (18), (20), (22), (27), (28) and (29), respectively. Anoop Chaturvedi et al.[24] have given the concrete manifestation of coverage probability and average volume.

In this section, we learned the properties of the m-equation SUR model via simulation and let $m = 5$. Assuming that the coefficient vectors of this model were $\beta_1 = (1,2,3)'$, $\beta_2 = (3,4,5)'$, $\beta_3 = (2,4,5)'$, $\beta_4 = (3,2,5)'$ and $\beta_5 = (1,2,5)'$. Moreover, we also supposed $Y \sim N(X\beta, \Sigma \otimes I_n)$ where $\beta = [\beta_1;\beta_2;\beta_3;\beta_4;\beta_5]$. Let $\sigma_{1,1} = 9$ and $\Sigma_{2,2} = pascal(4)$, while $\Sigma_{1,2}$ was varied to the correlation values of $\rho = 0, 0.1, 0.3, 0.5, 0.7, 0.9$. Meanwhile, we considered different sample sizes of $n = 20, 50, 100$. The last two columns for the design matrices X_1, X_2, X_3, X_4 and X_5 were generated by using $N\left((1,2)', 3I_2\right)$, $N\left((2,1)', 3I_2\right)$, $N\left((2,3)', 3I_2\right)$, $N\left((3,2)', 3I_2\right)$ and $N\left((2,4)', 3I_2\right)$, respectively, and $I_2 = [1,0;0,1]$. The first column of these design matrices were composed of $e = (1,1,\cdots,1)'$. In addition, the simulation repetition size was 10,000 and we performed all simulations by our own editing programs in Matlab. Table 1-3 show the results of our simulation and the nominal confidence level was 95%.

Table 1. Coverage probabilities(CP) and mean volumes(MV) for $n = 20$ and various values of ρ.

ρ		Two-stage estimator			Maximum Likelihood estimator		
		$CE_{ts}(\beta_1)$	$CE_{ts}^{adj1}(\beta_1)$	$CE_{ts}^{adj2}(\beta_1)$	$CE_{ml}(\beta_1)$	$CE_{ml}^{adj1}(\beta_1)$	$CE_{ml}^{adj2}(\beta_1)$
0	CP	0.8078	0.8596	0.8779	0.8528	0.8971	0.9237
	MV	7.7370	10.691	12.3664	5.7548	7.5207	8.9989
0.1	CP	0.8157	0.8549	0.8773	0.8519	0.9001	0.9287
	MV	7.8575	10.4907	12.2944	5.7960	7.4581	9.0198
0.3	CP	0.8073	0.8635	0.8792	0.8574	0.9003	0.9199
	MV	7.2968	9.9262	11.5608	5.6640	7.3658	8.8831
0.5	CP	0.8153	0.8661	0.8971	0.8691	0.9066	0.9276
	MV	6.5021	8.7278	10.1828	5.5187	7.0097	8.6030
0.7	CP	0.8275	0.8865	0.8979	0.8653	0.9054	0.9279
	MV	5.3617	7.2760	8.3863	5.2959	6.8383	8.1922
0.9	CP	0.8584	0.9088	0.9231	0.8617	0.9070	0.9366
	MV	3.7904	5.1669	5.8734	4.8882	6.3155	7.6836

From table 1, the coverage probabilities and average volumes of six CEs under different values of ρ and $n = 20$ can be showed. The coverage probability of each CE increased along with the increase of ρ, while its average volume decreased with the increase of ρ. Namely, the accuracy of the CE improved with increased ρ. If the value of ρ was fixed, $CE_{ml}^{adj2}(\beta_1)$ had more higher coverage probability and larger average volume than $CE_{ml}^{adj1}(\beta_1)$, while $CE_{ml}(\beta_1)$ had the lowest coverage probability and smallest average volume. Also, the case of $CE_{ts}(\beta_1)$, $CE_{ts}^{adj1}(\beta_1)$ and $CE_{ts}^{adj2}(\beta_1)$ were similar. These characteristics were rational because as ρ increased, more information can be obtained from the remaining regression error structures, which were essentially showed as the sample size increased. When the value of ρ was low, the coverage probabilities of CEs for MLE were generally higher than those CEs for two-stage estimator and their average volumes were smaller. When the value of ρ was high quite, the change of the average volumes were opposite to the above case.

For $n = 50$, the coverage probabilities of CEs were generally closer to the nominal level and their average volumes were closer to zero than those CEs under $n = 20$. The coverage probabilities of CEs from two-stage estimator were higher than those from MLEs under ρ took the lower value than 0.5, while their average volumes were also bigger. The above was the opposite under the case of that ρ took a larger value.

For $n = 100$ from table 3, compared to the coverage probabilities and average volumes of all six CEs for $n = 20,50$, the coverage probabilities when $n = 100$ were the most similar to the theoretical value and the average volumes were closest to zero, which accuracy reached $O(10^{-3})$. If ρ took the quite small value, the coverage probabilities of CEs from two-stage estimator were higher than those from MLEs, while their average volumes were also bigger. When ρ took a slightly larger value, the above was the opposite. Other changes of CEs under $n = 100$ were similar to those under $n = 50$. Kent R.Riggs et al.[17] have learned the coverage probabilities of CEs in two SUR model. We also examined the coverage probabilities and average volumes in two SUR model. For the SUR model with strong correlations under different sample sizes, the coverage probabilities and average volumes of multivariate model with $m = 5$ were greater than $m = 2$, while the conclusion was opposite when the correlation coefficient was small.

Table 2. Coverage probabilities(CP) and mean volumes(MV) for $n=50$ and various values of ρ.

ρ		Two-stage estimator			Maximum Likelihood estimator		
		$CE_{ts}(\beta_1)$	$CE_{ts}^{adj1}(\beta_1)$	$CE_{ts}^{adj2}(\beta_1)$	$CE_{ml}(\beta_1)$	$CE_{ml}^{adj1}(\beta_1)$	$CE_{ml}^{adj2}(\beta_1)$
0	CP	0.9191	0.9359	0.9454	0.9176	0.9330	0.9415
	MV	1.5890	1.7892	1.9111	1.4579	1.6303	1.7453
0.1	CP	0.9210	0.9331	0.9403	0.9173	0.9340	0.9418
	MV	1.5914	1.7684	1.8825	1.4646	1.6278	1.7364
0.3	CP	0.9190	0.9360	0.9493	0.9182	0.9279	0.9366
	MV	1.4767	1.6432	1.7662	1.4252	1.5740	1.7032
0.5	CP	0.9233	0.9374	0.9471	0.9188	0.9382	0.9481
	MV	1.2611	1.4110	1.4989	1.3550	1.5037	1.6090
0.7	CP	0.9222	0.9374	0.9475	0.9287	0.9470	0.9491
	MV	0.9351	1.0406	1.1103	1.2409	1.3841	1.4817
0.9	CP	0.9340	0.9442	0.9503	0.9358	0.9528	0.9537
	MV	0.4998	0.5631	0.5914	1.0927	1.2164	1.3036

Table 3. Coverage probabilities(CP) and mean volumes(MV) for $n=100$ and various values of ρ.

ρ		Two-stage estimator			Maximum Likelihood estimator		
		$CE_{ts}(\beta_1)$	$CE_{ts}^{adj1}(\beta_1)$	$CE_{ts}^{adj2}(\beta_1)$	$CE_{ml}(\beta_1)$	$CE_{ml}^{adj1}(\beta_1)$	$CE_{ml}^{adj2}(\beta_1)$
0	CP	0.9349	0.9410	0.9457	0.9327	0.9400	0.9447
	MV	0.5425	0.5717	0.5910	0.5212	0.5502	0.5678
0.1	CP	0.9364	0.9441	0.9462	0.9372	0.9409	0.9429
	MV	0.5363	0.5655	0.5843	0.5163	0.5458	0.5621
0.3	CP	0.9328	0.9439	0.9460	0.9361	0.9430	0.9467
	MV	0.4957	0.5268	0.5413	0.4993	0.5253	0.5465
0.5	CP	0.9386	0.9439	0.9483	0.9382	0.9452	0.9501
	MV	0.4152	0.4387	0.4532	0.4625	0.4853	0.5043
0.7	CP	0.9413	0.9462	0.9495	0.9457	0.9549	0.9566
	MV	0.2963	0.3126	0.3229	0.4060	0.4305	0.4429
0.9	CP	0.9429	0.9509	0.9508	0.9611	0.9622	0.9684
	MV	0.1365	0.1450	0.1489	0.3319	0.3521	0.3637

6. An application

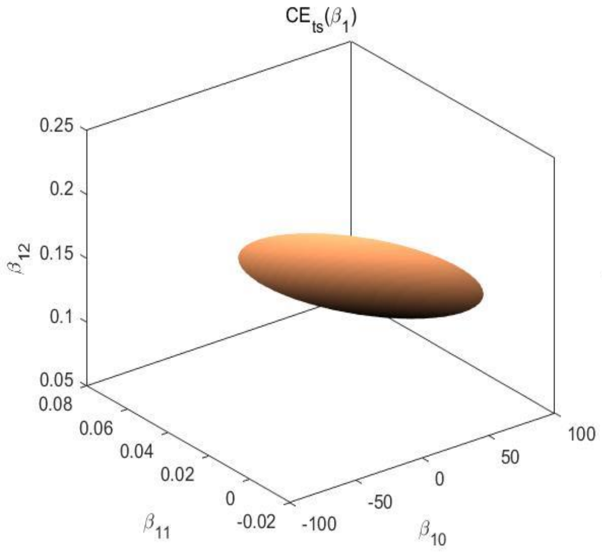
In this part, we consider five regression equations which model a company's total investment as a linear combination of its stock market value and outstanding shares value at the beginning of the year. The data set comes from Boot and de Wit[23]. These five companies of interest are General Electric(GE), Westinghouse(WH), General Motors Corporation(GMC), United States Steel Corporation(SC) and Chrysler Corporation(CC). Let $X_{i,1,j}(i=1,2,\cdots,5)$ denotes the stock market value at the beginning of the j-th year of GE, WH, GMC, SC and CC, respectively. Simultaneously, let $X_{i,2,j}(i=1,2,\cdots,5)$ denotes the outstanding shares value at the beginning of the j-th year of GE, WH, GMC, SC and CC, respectively. Also, the total investment of GE, WH, GMC, SC and CC are regarded as $Y_{i,j}(i=1,2,\cdots,5)$ of the jth year. Thus, the five-equation SUR model is

$$Y_{i,j} = \beta_{i,0} + X_{i,1,j}\beta_{i,1} + X_{i,2,j}\beta_{i,2} + \varepsilon_i \left(i = 1,2,\cdots,5,\ j = 1,2,\cdots,n\right), \tag{30}$$

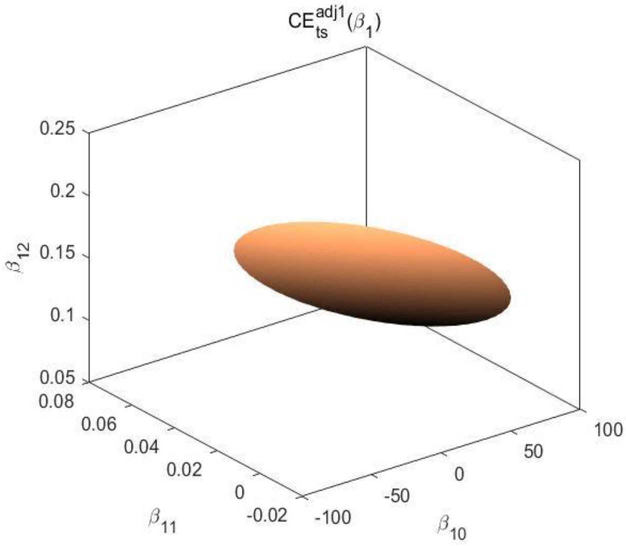
where $\beta_i = \left(\beta_{i,0}, \beta_{i,1}, \beta_{i,2}\right)' \in R_{3\times1}$, ε_i is a scalar and

$$X_i = \begin{pmatrix} 1 & X_{i,1,1} & X_{i,2,1} \\ 1 & X_{i,1,2} & X_{i,2,2} \\ \vdots & \vdots & \vdots \\ 1 & X_{i,1,n} & X_{i,2,n} \end{pmatrix}$$

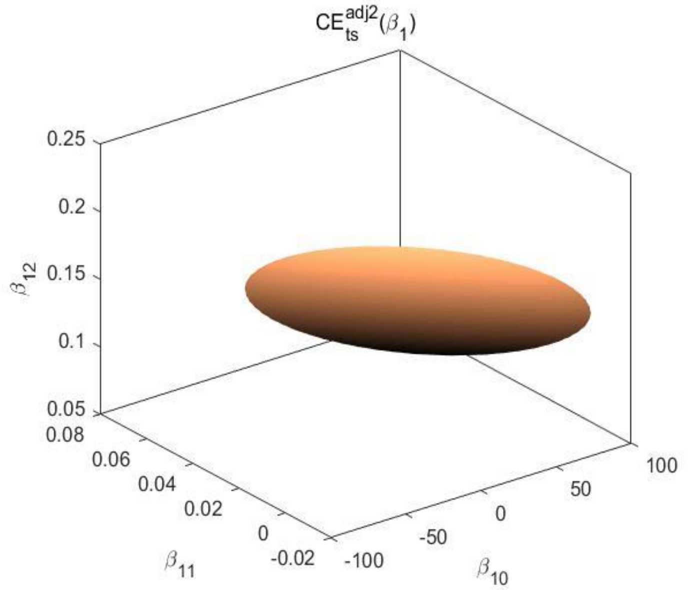
is $n\times3$ matrix, $i = 1,2,\cdots,5$. Here we mainly consider the main unknown parameter β_1 and the CEs for β_1. For this data set, $n = 20$.



(a)



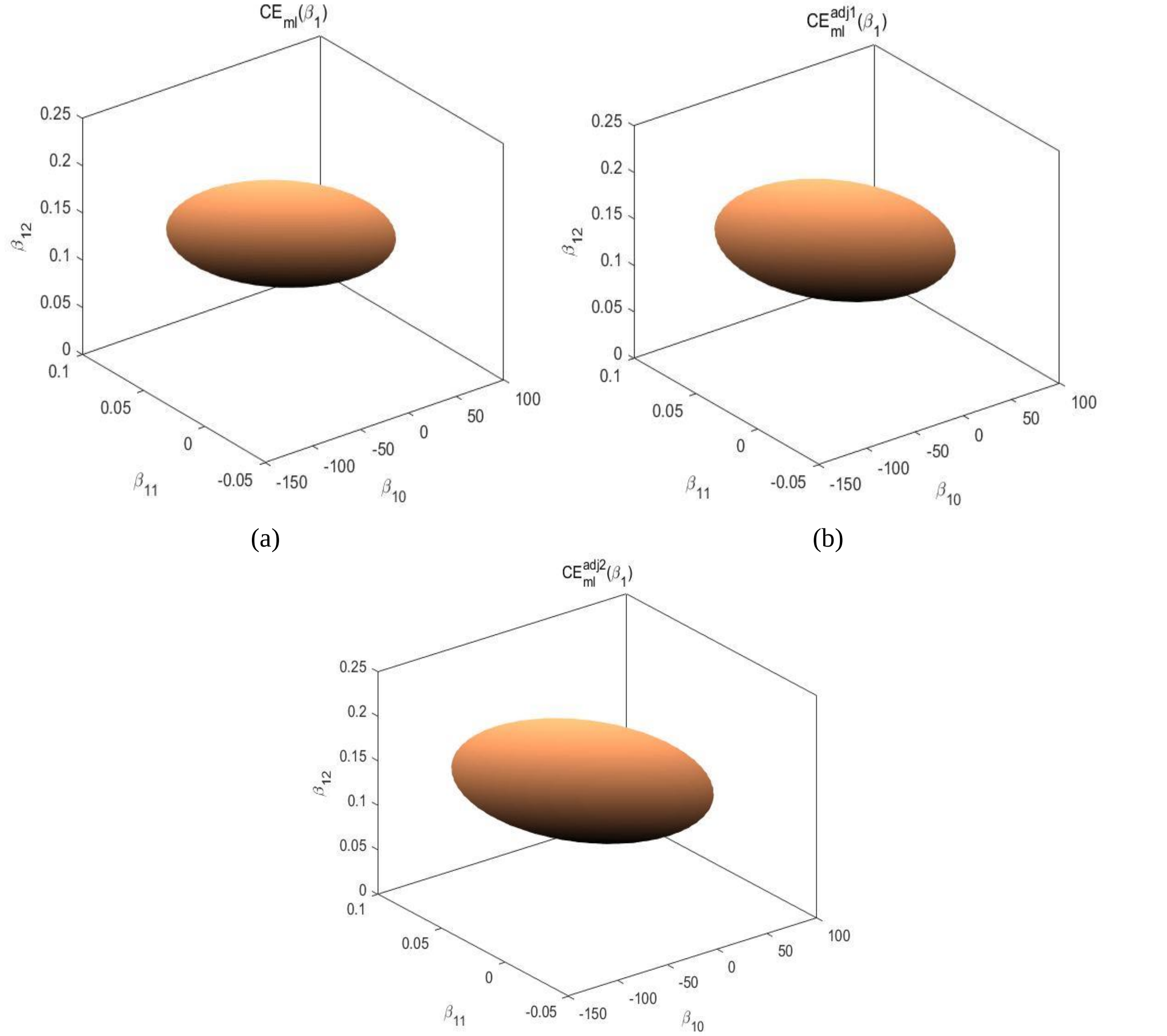
(b)



(c)

Fig. 1: Graphs of three approximate 95% CEs for β_1 from two-stage estimator in (30).

As we can see from the Fig 1 and 2 that $CE_{ts}^{adj2}\left(\beta_1\right)$ and $CE_{ml}^{adj2}\left(\beta_1\right)$ perform well because their coverage probabilities are very close to nominal values and yield quite small volumes of CEs. The estimators of the principal regression coefficients of the SUR model in (29) corresponding to the two types of CEs are $\tilde{\beta}_1 = (9.1816, 0.1561, 0.0158)'$ and $\hat{\beta}_1 = (-21.1807,\ 0.1287,\ 0.0371)'$, respectively.



(a) (b)

(c)

Fig. 2: Graphs of three approximate 95% CEs for β_1 from ML estimator in (30).

Table 4. Mean volumes(MV) of approximate 95% CEs for β_1.

	$CE_{ts}(\beta_1)$	$CE_{ts}^{adj1}(\beta_1)$	$CE_{ts}^{adj2}(\beta_1)$	$CE_{ml}(\beta_1)$	$CE_{ml}^{adj1}(\beta_1)$	$CE_{ml}^{adj2}(\beta_1)$
MV$(m=2)$	0.1462	0.1655	0.1482	0.1290	0.1460	0.1308
MV$(m=5)$	0.0279	0.0316	0.0433	0.1113	0.1260	0.1728

In the first row of the table 4, we consider only the two-equation SUR model constructed by GE and WH. We derived the corresponding average volumes of the CEs using the method proposed above. The second row shows the average volumes of the corresponding CEs resulting from considering the five companies in the above example. It can be seen from the table 4 that except for $CE_{ml}^{adj2}(\beta_1)$, the accuracy of CEs obtained in the case of $m=5$ is higher than that of $m=2$. This result is reasonable, because the greater the number of equations, the more information we can obtain and the higher the confidence of the CEs. At the same time, the CEs from two-stage estimator are quite small because the error term of the first equation has a very strong correlation with the error terms of other equations. The CEs from MLE are stable under strongly correlated conditions. $CE_{ts}^{adj2}(\beta_1)$ and $CE_{ml}^{adj2}(\beta_1)$ perform very well under \$m=2\$ in case of the high coverage and small average volumes. Although the average volumes of $CE_{ts}^{adj2}(\beta_1)$ and $CE_{ml}^{adj2}(\beta_1)$ are slightly larger in the case of $m=5$, the coverage probabilities are closest to the theoretical value, therefore they also behave well. Although the accuracy of $CE_{ts}^{adj2}(\beta_1)$ is better than $CE_{ml}^{adj2}(\beta_1)$, its coverage is low as the table 1 shows.

7. Discussion

In this paper, we have used two kinds of estimation methods to estimate the main regression coefficients of m-equation SUR model, and have obtained two new CEs by constructing wald statistics. We have also derived a new matrix form of maximum likelihood estimators by using Fisher's information matrix in the case of m regression equations. Through the Monte Carlo simulation, we have proved that the confidence ellipticity after correction was a little bit worse, but its confidence probability was improved under different sample values and different correlation coefficients. Applying these CEs to the actual data, we also got similar conclusions. This paper assumed that the rank of covariance in the m equations were equal and we can learn rank inequality in future studies. Moreover, we can also compare the features of the confidence ellipsoid constructed by other different estimation methods in future research.

8. Acknowledgement(s)

We would like to express our appreciation for the valuable comments made by editors and unknown referees so that we can make further improvements to this paper.

Appendix. A matrix-calculus derivation of Fisher's information matrix

Fisher's information matrix for the SUR model is

$$I(\beta_1,\beta_2,\cdots,\beta_m,\Sigma)=\begin{bmatrix} I_{\beta_1} & I_{\beta_1,\beta_2} & \cdots & I_{\beta_1,\beta_m} & I_{\beta_1,\Sigma} \\ I_{\beta_2,\beta_1} & I_{\beta_2} & \cdots & I_{\beta_2,\beta_m} & I_{\beta_2,\Sigma} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ I_{\beta_m,\beta_1} & I_{\beta_m,\beta_2} & \cdots & I_{\beta_m} & I_{\beta_m,\Sigma} \\ I_{\Sigma,\beta_1} & I_{\Sigma,\beta_2} & \cdots & I_{\Sigma,\beta_m} & I_{\Sigma} \end{bmatrix}=\begin{bmatrix} I_{\beta_1} & I_{1,2} & I_{1,3} \\ I_{2,1} & I_2 & I_{2,3} \\ I_{3,1} & I_{3,2} & I_{\Sigma} \end{bmatrix}, \tag{A.1}$$

where

$$I_{\beta_i}\equiv -E\left[\frac{\partial^2 l(\beta_1,\cdots,\beta_m,\Sigma)}{\partial\beta_i\partial\beta_i{'}}\right](i=1,2,\cdots,m),$$

$$I_{\Sigma}\equiv -E\left[\frac{\partial^2 l(\beta_1,\cdots,\beta_m,\Sigma)}{\partial vec(\Sigma)\partial vec(\Sigma)'}\right](i=1,2,\cdots,m),$$

$$I_{\beta_i,\beta_j}\equiv -E\left[\frac{\partial^2 l(\beta_1,\cdots,\beta_m,\Sigma)}{\partial\beta_i\partial\beta_j{'}}\right](i,j=1,2,\cdots,m,i\neq j),$$

$$I_{\beta_i,\Sigma}\equiv -E\left[\frac{\partial^2 l(\beta_1,\cdots,\beta_m,\Sigma)}{\partial\beta_i\partial vec(\Sigma)'}\right](i=1,2,\cdots,m),$$

$I'_{\beta_i,\beta_j}=I_{\beta_j,\beta_i}(i,j=1,2,\cdots,m,i\neq j)$ and $I'_{\beta_i,\Sigma}=I_{\Sigma,\beta_i}(i=1,2,\cdots,m)$. Srivastava and Giles[7]once gave a concrete representation of each element in (A.1), but such forms are not conducive to constructing the CEs of the principal regression coefficients of the m-equation SUR model. Thus, a new form of deriving (A.1) from Kent R.Riggs et al.[17] is as follows. Let

$$\Sigma^{-1} = \begin{bmatrix} \sigma^{(1,1)} & \sigma^{(1,2)} & \cdots & \sigma^{(1,m)} \\ \sigma^{2(1,1)} & \sigma^{(2,2)} & \cdots & \sigma^{(2,m)} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^{(m,1)} & \sigma^{(m,2)} & \cdots & \sigma^{(m,m)} \end{bmatrix}.$$

Then, according to $X_i^{*'}\left(\Sigma^{-1} \otimes I_n\right)X_j^* = \sigma^{(i,j)} X_i' X_j$, we have that

$$\frac{\partial l(\beta_1, \beta_2, \cdots, \beta_m, \Sigma)}{\partial \beta_i} = X_i^{*'}\left[\Sigma^{-1} \otimes I_n\right]\left(Y - X_1^*\beta_1 - X_2^*\beta_2 - \cdots - X_m^*\beta_m\right) = \sum_{j=1}^m \sigma^{(i,j)} X_i' Y_j - \sum_{j=1}^m \sigma^{(i,j)} X_i' X_j,$$

$i = 1,2,\cdots,m$, where $\sigma^{(i,j)}$ is the (i,j) element of $\Sigma^{-1}(i, j = 1,2,\cdots,m)$. Through simplification, we have that

$$\frac{\partial l(\beta_1, \beta_2, \cdots, \beta_m, \Sigma)}{\partial \Sigma} = -\frac{n}{2}\frac{\partial \ln|\Sigma|}{\partial \Sigma} - \frac{1}{2}\frac{\partial tr\left(A\Sigma^{-1}\right)}{\partial \Sigma} = -\frac{n}{2}\left[2\Sigma^{-1} - Diag\left(\Sigma^{-1}\right)\right] + \frac{1}{2}\Sigma^{-1}A\Sigma^{-1}.$$

Also,

$$\frac{\partial l(\beta_1, \beta_2, \cdots, \beta_m, \Sigma)}{\partial \beta_i \beta_j'} = -\sigma^{(i,j)} X_i' X_j,$$

and

$$\frac{\partial l(\beta_1, \beta_2, \cdots, \beta_m, \Sigma)}{\partial \beta_i \partial vec(\Sigma)'} = \frac{\partial}{\partial vec(\Sigma)'} vec\left[X_i^{*'}\left[\Sigma^{-1} \otimes I_n\right]\left(Y - X_1^*\beta_1 - X_2^*\beta_2 - \cdots - X_m^*\beta_m\right)\right]$$

$$= \left[\left(Y - \sum_{j=1}^m X_j^*\beta_j\right)' \otimes X_i^{*'}\right]\frac{\partial}{\partial vec(\Sigma)'} vec\left[\Sigma^{-1} \otimes I\right]$$

$$= -\left[\left(Y - \sum_{j=1}^m X_j^*\beta_j\right)' \otimes X_i^{*'}\right]\left(I_m \otimes K_{n,m} \otimes I_n\right)\left(\Sigma^{-1} \otimes \Sigma^{-1} \otimes I_n\right),$$

$i = 1,2,\cdots,m$, where $K_{n,m}$ denotes the commutation matrix defined such that $vec(Q') = K_{n,m} vec(Q)$ for any $Q \in R_{n\times m}$ and $vec(\Sigma) = \left(\sigma_{1,1}, \sigma_{1,2}, \cdots, \sigma_{1,m}, \cdots, \sigma_{m,1}, \sigma_{m,2}, \cdots, \sigma_{m,m}\right)'$. Empathy, we obtain that

$$\frac{\partial l(\beta_1, \beta_2, \cdots, \beta_m, \Sigma)}{\partial vec(\Sigma) \partial vec(\Sigma)'} = -\frac{n}{2}\frac{\partial^2 \ln|\Sigma|}{\partial vec(\Sigma) \partial vec(\Sigma)'} - \frac{1}{2}\frac{\partial^2 tr\left(A\Sigma^{-1}\right)}{\partial vec(\Sigma) \partial vec(\Sigma)'}$$
$$= -\frac{1}{2}\left[K_{m,m}\left(\Sigma^{-1}A\Sigma^{-1} \otimes \Sigma^{-1} + \Sigma^{-1} \otimes \Sigma^{-1}A\Sigma^{-1}\right)\right] + \frac{n}{2}\left[K_{m,m}\left(\Sigma^{-1} \otimes \Sigma^{-1}\right)\right],$$

where $A = \{A_{i,j}\} \equiv \left\{\left(Y_i - X_i\beta_i\right)'\left(Y_j - X_j\beta_j\right)\right\}$ .Thus, $I_{\beta_i,\Sigma} = 0$, $I_{\beta_i} = \sigma^{(i,i)} X_i' X_i$ and $I_{\beta_i,\beta_j} = I_{\beta_j,\beta_i} = \sigma^{(i,j)} X_i' X_j$. By calculation we can see $E[A] = n\Sigma$. Therefore,

$$I_\Sigma = \frac{1}{2}\left[K_{m,m}\left(n\Sigma^{-1} \otimes \Sigma^{-1} + \Sigma^{-1} \otimes n\Sigma^{-1}\right)\right] - \frac{n}{2}\left[K_{m,m}\left(\Sigma^{-1} \otimes \Sigma^{-1}\right)\right]$$
$$= \frac{n}{2}\left[K_{m,m}\left(\Sigma^{-1} \otimes \Sigma^{-1}\right)\right].$$

9. References

[1] A. Zellner, *An efficient method of estimating seemingly unrelated regression equations and tests for aggregation bias. Journal of American Statistical Association*, Journal of American Statistical Association, 57. 348-368, 1962.

[2] A. Zellner, *Estimators for seemingly unrelated regression equations: some exact finite sample results,* Journal of American Statistical Association, 58. 977--992, 1963.

[3] A. Zellner, *An introduction to Bayesian inference in econometrics,* New York:Weley, 1971.

[4] N. Revankar, *Some finite sample results in the context of two seemingly unrelated regression equations,* Journal of American Statistical Association, 69. 187-190, 1974.

[5] N. Revankar, *Use of restricted residuals in SUR systems:some finite sample results,* Econom.Theory, 12. 569-580, 1976.

[6] V. K. Srivastava, and A. K. Srivastava, *Improved estimationin a two equationseemingly unrelated regression model,* Statistica, 3. 417–422, 1984.

[7] V. Srivastava, and D.Giles, *Seemingly Unrelated Regression Equation Models. Marcel Dekker,* NewYork, 1987.

[8] A. Liu, *Two-stage estimators of seemingly unrelated regressions with elliptical distribution,* Comm. Statist. Theory Methods, 32. 1251–1263, 1993.

[9] Li Zhao, and Xingzhong Xu, *Generalized canonical correlation variables improved estimation in high dimensional seemingly unrelated regression models,* Statistics and Probability Letters, 126. 119-126, 2017.

[10] P. Phillips, *The exact distribution of SUR estimator,* Econometrica, 53.745–756, 1985.

[11] T. Park, *Equivalence of maximum likelihood estimation and iterative two-stage estimation for seemingly unrelated regression models,* Comm.Statist. Theory Methods, 22. 2285–2296, 1993.

[12] D. Fraser, M. Kekkas, and A. Wong, *Highly accurate likelihood analysis for the seemingly unrelated regression problem,* J.Econometrics1, 27. 17–33, 2005.

[13] L. Le. Cam, *On the standard asymptotic confidence ellipsoids of Wald,* International Statistical Review, 58. 129-152, 1990.

[14] Lee. C. Adkins, and R. Carter, *Hill. An improved confidence ellipsoid for the linear regression model,* Journal of Statistical Computation and Simulation, 36. 9-18, 1990.

[15] Laurent El Ghaoui, and Giuseppe Calafiore, *Confidence ellipsoids for uncertain linear equations with structure,* IEEE Conference on Decision and Control, 2. 1922-1927, 1999.

[16] Erik Weyer, and M.C.Campi, *Non-asymptotic confidence ellipsoids for the least-squares estimate,* Automatica, 38. 1539-1547, 2002.

[17] Kent R.Riggs, Phil D.Young, and Dean M.Young, *Confidence ellipsoids for the primary regression coefficients in two seemingly unrelated regression models,* Statistical Methodology, 32. 1-13, 2016.

[18] J. Liu, S. Wang, *Two-stage estimate of the parameters in seemingly unrelated regression models,* Prog. Nat. Sci, 9. 489–496, 1999.

[19] N.Timm, *Applied Multivariate Analysis,* Springer, New York, 2002.

[20] F.Cribari-Neto, and S.Zarkos, *Improved test statistics for multivariate regression,* Econom.Lett, 49. 113–120, 1995.

[21] G.M.Cordeiro, and S.L.P.Ferrari, *A modified score test statistic having chi-squared distribution to order* n^{-1}, Biometrika, 78. 573-582, 1991.

[22] Gauss M.Cordeiro, and Sóstenes L.lins, *New results on the likelihood ratio and score test for the von Miss distribution,* Journal of Statistical Computation and Simulation, 80. 915-926, 2010.

[23] J.C.G.Boot, and G.M.de Wit, *Investment demand: an empirical contribution to the aggregation problem,* International Economic Review, 1. 3-30, 1960.

[24] Anoop Chaturvedi, Suchita Gupta, and M.Ishaq Bhatti, *Confidence ellipsoids based on a general family of shrinkage estimators for a linear model with non-spherical disturbances,* Journal of Multivariate Analysis, 104.

140-158, 2012.

[25] Li Zhao, and Xingzhong Xu, *Generalized canonical correlation variables improved estimation in high dimensional seemingly unrelated regression models*, Statistics and Probability Letters, 126. 119-126, 2017.