Pak. J. Statist.
2021 Vol. 37(4), 393-415

SEEMINGLY UNRELATED REGRESSIONS WITH LINEAR CONSTRAINT

Yupeng Sun, Jie Zhang and **Lichun Wang**[§]
School of Science, Beijing Jiaotong University, Beijing 100044, China
Email: [1] ypsun@bjtu.edu.cn
[2] jiezhangl@bjtu.edu.cn
[3] wlc@amss.ac.cn
[§] Corresponding author

ABSTRACT

In this paper, we study the parameter estimation problem in the system of two seemingly unrelated regressions with linear constraint. The main content contains: We prove that the covariance-adjusted constraint estimator and the constrained covariance adjustment estimator are better than the constrained least squares estimator and the covariance adjustment estimator, respectively. We define the two-stage covariance-adjusted constraint estimator and the two-stage constrained covariance adjustment estimator, and discuss their superiorities by theoretical proofs and numerical simulations. Employing the matrix power series, we obtain the estimator of regression parameter and further compare it with the limit of the sequence of the covariance-adjusted constraint estimator and the limit of the sequence of the constrained covariance adjustment estimator. For two important cases, we exhibit some interesting relationships between the limit of the covariance-adjusted constraint estimator sequence and the limit of the constrained covariance adjustment estimator sequence.

KEYWORDS

Seemingly unrelated regressions, constrained least squares estimator, covariance adjustment estimator, two-stage estimator.

MR Subject Classification: 62J05, 62H12.

1. INTRODUCTION

Consider the system of two seemingly unrelated regressions with linear constraint

$$\begin{cases} y_1 = X_1\beta + e_1, & (1) \\ y_2 = X_2\gamma + e_2, & (2) \\ H_1\beta = d, & (3) \end{cases} \tag{1.1}$$

where $y_i\,(i=1,2)$ are $n\times 1$ observation vectors, $X_i\,(i=1,2)$ are $n\times p_i$ design matrix with rank $(X_i)=p_i$, β and γ are $p_1\times 1$ and $p_2\times 1$ vectors of unknown regression parameters, respectively, e_1 and e_2 are $n\times 1$ correlated random error vectors with

$E(e_1)=0$, $Cov(e_1)=\sigma_{11}I_n$, $E(e_2)=0$, $Cov(e_2)=\sigma_{22}I_n$ and $Cov(e_1,e_2)=\sigma_{12}I_n$ and each row of (e_1,e_2) has the covariance matrix $\sum$, where $\sum=(\sigma_{ij})$ is a 2×2 non-diagonal positive definite matrix, and H_1 is $k\times p_1$ matrix with rank $(H_1)=k$.

If we ignore seemingly unrelated information (2) in the system (1.1), the system (1.1) will reduce to

$$\begin{cases} y_1 = X_1\beta+e_1, \\ H_1\beta = d, \end{cases} \tag{1.2}$$

which is a constrained regression model.

In the past, many papers discussed the parameter estimation problem under constraint condition. Mantel (1969) established some results for constrained least squares regression, Armstrong (1976) introduced a branch and bound solution of a constrained least squares problem. Other related results can be found in Baksalary (1979), Wang (1998) and Wang (2003), etc. Using the Lagrange multiplier method, the constrained least squares estimator of β in system (1.2) is

$$\hat{\beta}_H = \hat{\beta}_{LS} - (X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}(H_1\hat{\beta}_{LS}-d), \tag{1.3}$$

where $\hat{\beta}_{LS} = (X_1'X_1)^{-1}X_1'y_1$ denotes the LSE of β in the model $y_1 = X_1\beta+e_1$.

However, the estimator (1.3) does not make use of the seemingly unrelated information (2) of the system (1.1). Applying the covariance adjustment technique once to improve $\hat{\beta}_H$, we can obtain the one-step covariance-adjusted constraint estimator $\hat{\beta}_{H,CA}$, which is expected to be better than $\hat{\beta}_H$.

The covariance adjustment technique is introduced by Rao (1967), which is used to obtain an optimal unbiased estimation of a parameter via linearly combining an unbiased estimator and an unbiased estimator of a zero vector, with the use of a known covariance matrix. Here, an unbiased estimator $\hat{\theta}_1$ is more better than an unbiased estimator $\hat{\theta}_2$ means that $Cov(\hat{\theta}_1)$ is less than $Cov(\hat{\theta}_2)$.

On the other hand, if we ignore the linear constraint (3) in the system (1.1), the system of (1.1) will become

$$\begin{cases} y_1 = X_1\beta+e_1, \\ y_2 = X_2\gamma+e_2, \end{cases} \tag{1.4}$$

which is a system of two seemingly unrelated regression (SUR) models. SUR models have been widely applied to many fields including biological sciences, economics, geography and soon. Many papers have discussed the parameter estimation of SUR model and obtain some meaningful statistical properties. Zellner (1962) proposed an efficient method of estimating seemingly unrelated regressions. Zellner (1963) and

Revankar (1974) respectively proved that the two-stage estimator is superior to the ordinary LSE. Chen and Liao (1995) discussed the properties of an improved estimator when the design matrix is ill-conditioned. Similar works can be found in Wang (2001), Liu (2002) and Wang (2011), etc. In the system (1.4), applying the covariance adjustment technique once, the one-step covariance adjustment estimator for β would be

$$\hat{\beta}_{CA} = \hat{\beta}_{LS} - \frac{\sigma_{12}}{\sigma_{22}} (X_1' X_1)^{-1} X_1' N_2 y_2, \tag{1.5}$$

where $N_2 = I_n - X_2 (X_2' X_2)^{-1} X_2'$.

However, the estimator (1.5) does not give consideration to the constraint condition (3) of the system (1.1). Together with the constraint condition and $\hat{\beta}_{CA}$, we can obtain the constrained one-step covariance adjustment estimator $\hat{\beta}_{CA,H}$, which is also expected to be overwhelming than $\hat{\beta}_{CA}$.

In the case that the covariance matrix Σ is unknown, replacing it by its unrestricted estimate (see Zellner (1962, 1963), Revankar (1974)), we can define the two-stage covariance-adjusted constraint estimator $\hat{\beta}_{H,CA}(S)$ and the two-stage constrained covariance adjustment estimator $\hat{\beta}_{CA,H}(S)$. A natural problem is that they are still better than the estimators $\hat{\beta}_H$ and $\hat{\beta}_{CA}$ respectively?

Furthermore, if we use the covariance adjustment technique to improve $\hat{\beta}_H$ and $\hat{\beta}_{CA}$ repeatedly, we can obtain the sequence of the covariance-adjusted constraint estimator $\{\hat{\beta}_{H,CA}^{(n)}, n \geq 1\}$ and the sequence of the constrained covariance adjustment estimator $\{\hat{\beta}_{CA,H}^{(n)}, n \geq 1\}$, then what further conclusions will we obtain?

This paper is organized as follows. In Section 2, we introduce the one-step covariance-adjusted constraint estimator $\hat{\beta}_{H,CA}$ and the one-step constrained covariance adjustment estimator $\hat{\beta}_{CA,H}$, and prove that they are unbiased and much better than the estimators $\hat{\beta}_H$ and $\hat{\beta}_{CA}$ respectively. If Σ is unknown, we define the two-stage covariance-adjusted constraint estimators $\hat{\beta}_{H,CA}(S)$ and the two-stage constrained covariance adjustment estimator $\hat{\beta}_{CA,H}(S)$, and their superiorities are discussed by theoretical demonstration and numerical simulations. In Section 3, for the system (1.1) we firstly obtain the matrix series expression of the regression parameter β, say $\hat{\beta}_{CCA}$, and we also construct the sequence of the covariance-adjusted constraint estimator $\{\hat{\beta}_{H,CA}^{(n)}, n \geq 1\}$ and the sequence of the constrained covariance adjustment estimator

$\{\hat{\beta}^{(n)}_{CA,H}, n \geq 1\}$, subsequently we prove that $\hat{\beta}_{CCA}$ is equal to $\lim_{n\to\infty} \hat{\beta}^{(n)}_{CA,H}$ but not equal to $\lim_{n\to\infty} \hat{\beta}^{(n)}_{H,CA}$. In Section 4, for two special cases, we describe some interesting relationships between $\lim_{n\to\infty} \hat{\beta}^{(n)}_{H,CA}$, $\lim_{n\to\infty} \hat{\beta}^{(n)}_{CA,H}$ and $\hat{\beta}_{CCA}$.

2. THE PROPERTIES OF $\hat{\beta}_{H,CA}$ AND $\hat{\beta}_{CA,H}$

2.1 Σ is Known

In this Section we investigate the one-step covariance-adjusted constraint estimator $\hat{\beta}_{H,CA}$ and the one-step constrained covariance adjustment estimator $\hat{\beta}_{CA,H}$, and show that they are unbiased and more better than the estimators $\hat{\beta}_H$ and $\hat{\beta}_{CA}$, respectively.

Firstly, we introduce the following covariance adjustment theorem (see Wang (2003)).

Lemma 2.1

Let T_1, T_2 be $p\times 1$ and $q\times 1$ statistics, respectively, with $ET_1 = \theta$ and $ET_2 = 0$. Denote

$$Cov\begin{pmatrix} T_1 \\ T_2 \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \triangleq M . \tag{2.6}$$

If $M \geq 0$ and $M_{12} \neq 0$, then $\theta^* = T_1 - M_{12}M_{22}^- T_2$ is the best linear unbiased estimator among the class

$$\{A_1T_1 + A_2T_2 : A_1ET_1 + A_2ET_2 = \theta\}, \tag{2.7}$$

where A_1 and A_2 are undetermined matrices, and

$$Cov\left(\theta^*\right) = M_{11} - M_{12}M_{22}^- M_{21} \leq M_{11} = Cov\left(T_1\right), \tag{2.8}$$

where $A \geq B$ means the matrix $A - B$ is non-negative define and M_{22}^- donates any a generalized inverse of M_{22}. Hence, θ^* is better than T_1 in the sense of having less covariance.

Theorem 2.1

The one-step covariance-adjusted constraint estimator of β is

$$\hat{\beta}_{H,CA} = \hat{\beta}_{CA} - (X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}(H_1\hat{\beta}_{CA} - d). \tag{2.9}$$

Proof:

Let $T_1 = \hat{\beta}_H$ and $T_2 = N_2 y_2$. Thus, we have

$$
\begin{aligned}
Cov\left(T_1, T_2\right) &= Cov\left(\hat{\beta}_H, N_2 y_2\right) \\
&= \left[I_n - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1'\right)^{-1} H_1\right] (X_1'X_1)^{-1} X_1' Cov(y_1, y_2) N_2 \\
&= \sigma_{12} \left[I_n - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1'\right)^{-1} H_1\right] (X_1'X_1)^{-1} X_1' N_2, \qquad (2.10)
\end{aligned}
$$

and $Cov(T_2, T_2) = \sigma_{22} N_2$. Using the covariance adjustment technique once to improve T_1, we obtain the covariance-adjusted constraint estimator of β is

$$
\begin{aligned}
\hat{\beta}_{H,CA} &= T_1 - Cov\left(T_1, T_2\right)\left[Cov\left(T_2, T_2\right)\right]^{-} T_2 \\
&= \hat{\beta}_{LS} - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1'\right)^{-1} \left(H_1 \hat{\beta}_{LS} - d\right) \\
&\quad - \sigma_{12}\sigma_{22}^{-1} \left[I_n - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1'\right)^{-1} H_1\right] (X_1'X_1)^{-1} X_1' N_2 N_2^{-} N_2 y_2 \\
&= \left[I_n - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H'\right)^{-1} H_1\right] \left[\hat{\beta}_{LS} - \frac{\sigma_{12}}{\sigma_{22}} (X_1'X_1)^{-1} X_1' N_2 y_2\right] \\
&\quad + (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1'\right)^{-1} d \\
&= \hat{\beta}_{CA} - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1'\right)^{-1} (H_1 \hat{\beta}_{CA} - d). \qquad (2.11)
\end{aligned}
$$

□

Theorem 2.2

$\hat{\beta}_{H,CA}$ is an unbiased estimator of β, and the matrix $Cov(\hat{\beta}_H) - Cov(\hat{\beta}_{H,CA})$ is non-negative definite.

Proof:

From (1.5), we have

$$
\begin{aligned}
E[\hat{\beta}_{CA}] &= (X_1'X_1)^{-1} X_1' E[y_1] - \frac{\sigma_{12}}{\sigma_{22}} (X_1'X_1)^{-1} X_1' N_2 E[y_2] \\
&= \beta - \frac{\sigma_{12}}{\sigma_{22}} (X_1'X_1)^{-1} X_1' N_2 X_2 \gamma = \beta. \qquad (2.12)
\end{aligned}
$$

Thus,

$$
\begin{aligned}
E[\hat{\beta}_{H,CA}] &= E[\hat{\beta}_{CA}] - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1'\right)^{-1} \left(H_1 E[\hat{\beta}_{CA}] - d\right) \\
&= \beta - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1'\right)^{-1} (H_1 \beta - d) \\
&= \beta. \qquad (2.13)
\end{aligned}
$$

Hence, $\hat{\beta}_{H,CA}$ is an unbiased estimator of β.

By the covariance properties, we easily have $Cov(\hat{\beta}_{LS}) = \sigma_{11}(X_1'X_1)^{-1}$ and

$$Cov(\hat{\beta}_{CA}) = \sigma_{11}(X_1'X_1)^{-1} - \frac{\sigma_{12}^2}{\sigma_{22}}(X_1'X_1)^{-1}X_1'N_2X_1(X_1'X_1)^{-1}, \tag{2.14}$$

and

$$\begin{aligned} Cov(\hat{\beta}_H) &= Cov\left[\hat{\beta}_{LS} - (X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}(H_1\hat{\beta}_{LS} - d)\right] \\ &= \sigma_{11}\left[I_n - (X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}H_1\right](X_1'X_1)^{-1} \\ &\quad \times\left[I_n - (X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}H_1\right]' \\ &= \sigma_{11}(X_1'X_1)^{-1} - 2\sigma_{11}(X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}H_1(X_1'X_1)^{-1} \\ &\quad + \sigma_{11}(X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}H_1(X_1'X_1)^{-1}H_1' \\ &\quad \left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}H_1(X_1'X_1)^{-1} \\ &= \sigma_{11}(X_1'X_1)^{-1} - \sigma_{11}(X_1'X_1)^{-1}H_1'\left[H_1(X_1'X_1)^{-1}H_1'\right]^{-1}H_1(X_1'X_1)^{-1}. \end{aligned} \tag{2.15}$$

Then, the covariance matrix of $\hat{\beta}_{H,CA}$ is

$$\begin{aligned} Cov(\hat{\beta}_{H,CA}) &= Cov\left[\hat{\beta}_{CA} - (X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}(H_1\hat{\beta}_{CA} - d)\right] \\ &= \left[I_n - (X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}H_1\right]Cov(\hat{\beta}_{CA}) \\ &\quad \times\left[I_n - (X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}H_1\right]'. \end{aligned} \tag{2.16}$$

Comparing the covariance matrix of $\hat{\beta}_{H,CA}$ with that of $\hat{\beta}_H$, we have

$$\begin{aligned} &Cov(\hat{\beta}_H) - Cov(\hat{\beta}_{H,CA}) \\ &= \left[I_n - (X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}H_1\right]Cov(\hat{\beta}_{LS}) \\ &\quad \times\left[I_n - (X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}H_1\right]' \end{aligned}$$

$$-\left[I_n-(X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}H_1\right]Cov(\hat{\beta}_{CA})$$

$$\times\left[I_n-(X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}H_1\right]'$$

$$=\left[I_n-(X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}H_1\right]\left[Cov(\hat{\beta}_{LS})-Cov(\hat{\beta}_{CA})\right]$$

$$\times\left[I_n-(X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}H_1\right]'. \tag{2.17}$$

According to (2.14), it is easy to see that $Cov(\hat{\beta}_{LS})-Cov(\hat{\beta}_{CA})\geq 0$, which means the matrix (2.17) is nonnegative definite. This shows that the estimator $\hat{\beta}_{H,CA}$ is better than the estimator $\hat{\beta}_H$.

□

Theorem 2.3

The one-step constrained covariance adjustment estimator $\hat{\beta}_{CA,H}$ is

$$\hat{\beta}_{CA,H}=\hat{\beta}_{CA}-Cov(\hat{\beta}_{CA})H_1'\left(H_1Cov(\hat{\beta}_{CA})H_1'\right)^{-1}(H_1\hat{\beta}_{CA}-d). \tag{2.18}$$

Proof:

Put $T_1=\hat{\beta}_{CA}$ and $T_2=H_1\hat{\beta}_{CA}-d$, we have

$$Cov(T_1,T_2)=Cov(\hat{\beta}_{CA},H_1\hat{\beta}_{CA}-d)=Cov(\hat{\beta}_{CA})H_1'$$

$$=\left[\sigma_{11}(X_1'X_1)^{-1}-\frac{\sigma_{12}^2}{\sigma_{22}}(X_1'X_1)^{-1}X_1'N_2X_1(X_1'X_1)^{-1}\right]H_1', \tag{2.19}$$

$$Cov(T_2)=H_1Cov(\hat{\beta}_{CA})H_1'$$

$$=H_1\left[\sigma_{11}(X_1'X_1)^{-1}-\frac{\sigma_{12}^2}{\sigma_{22}}(X_1'X_1)^{-1}X_1'N_2X_1(X_1'X_1)^{-1}\right]H_1'. \tag{2.20}$$

From lemma 2.1, we obtain the one-step constrained covariance adjustment estimator as follows.

$$\hat{\beta}_{CA,H}=\hat{\beta}_{CA}-Cov(\hat{\beta}_{CA},H_1\hat{\beta}_{CA}-d)\left[Cov(H_1\hat{\beta}_{CA}-d)\right]^{-1}(H_1\hat{\beta}_{CA}-d)$$

$$=\hat{\beta}_{CA}-Cov(\hat{\beta}_{CA})H_1'\left(H_1Cov(\hat{\beta}_{CA})H_1'\right)^{-1}(H_1\hat{\beta}_{CA}-d). \tag{2.21}$$

□

Theorem 2.4

$\hat{\beta}_{CA,H}$ is an unbiased estimator of β, and the matrix $Cov(\hat{\beta}_{CA}) - Cov(\hat{\beta}_{CA,H})$ is non-negative definite.

Proof:

From (2.21), we have

$$\begin{aligned} E[\hat{\beta}_{CA,H}] &= E\left[\hat{\beta}_{CA} - Cov(\hat{\beta}_{CA})H_1'\left(H_1 Cov(\hat{\beta}_{CA})H_1'\right)^{-1}(H_1\hat{\beta}_{CA} - d)\right] \\ &= \beta - Cov(\hat{\beta}_{CA})H_1'\left(H_1 Cov(\hat{\beta}_{CA})H_1'\right)^{-1} E\left[H_1\hat{\beta}_{CA} - d\right] \\ &= \beta - Cov(\hat{\beta}_{CA})H_1'\left(H_1 Cov(\hat{\beta}_{CA})H_1'\right)^{-1}(H_1\beta - d) = \beta. \end{aligned} \tag{2.22}$$

Note that $Cov(\hat{\beta}_{CA})$ can be expressed as

$$\begin{aligned} Cov(\hat{\beta}_{CA}) &= (X_1'X_1)^{-1}\left(\sigma_{11}X_1'X_1 - \frac{\sigma_{12}^2}{\sigma_{22}}X_1'N_2X_1\right)(X_1'X_1)^{-1} \\ &= \sigma_{11}(X_1'X_1)^{-1}\left(X_1'X_1 - \frac{\sigma_{12}^2}{\sigma_{22}\sigma_{11}}X_1'N_2X_1\right)(X_1'X_1)^{-1} \\ &= \sigma_{11}(X_1'X_1)^{-1}\left[X_1'X_1 - \rho^2 X_1'\left(I_n - X_2(X_2'X_2)^{-1}X_2'\right)X_1\right](X_1'X_1)^{-1} \\ &= \sigma_{11}(X_1'X_1)^{-1}\left[(1-\rho^2)X_1'X_1 + \rho^2 X_1'X_2(X_2'X_2)^{-1}X_2'X_1\right](X_1'X_1)^{-1}, \end{aligned} \tag{2.23}$$

where $\rho^2 = \dfrac{\sigma_{12}^2}{\sigma_{11}\sigma_{22}} < 1$. Hence, $Cov(\hat{\beta}_{CA})$ is a positive definite matrix. Further, we have

$$\begin{aligned} Cov(\hat{\beta}_{CA,H}) &= Cov\left[\hat{\beta}_{CA} - Cov(\hat{\beta}_{CA})H_1'\left(H_1 Cov(\hat{\beta}_{CA})H_1'\right)^{-1}(H_1\hat{\beta}_{CA} - d)\right] \\ &= \left[I_n - Cov(\hat{\beta}_{CA})H_1'\left(H_1 Cov(\hat{\beta}_{CA})H_1'\right)^{-1}H_1\right]Cov(\hat{\beta}_{CA}) \\ &\qquad \times\left[I_n - Cov(\hat{\beta}_{CA})H_1'\left(H_1 Cov(\hat{\beta}_{CA})H_1'\right)^{-1}H_1\right]' \\ &= Cov(\hat{\beta}_{CA}) - 2Cov(\hat{\beta}_{CA})H_1'\left(H_1 Cov(\hat{\beta}_{CA})H_1'\right)^{-1}H_1 Cov(\hat{\beta}_{CA}) \\ &\qquad + Cov(\hat{\beta}_{CA})H_1'\left(H_1 Cov(\hat{\beta}_{CA})H_1'\right)^{-1}H_1 Cov(\hat{\beta}_{CA})H_1' \\ &\qquad \times\left(H_1 Cov(\hat{\beta}_{CA})H_1'\right)^{-1}H_1 Cov(\hat{\beta}_{CA}) \\ &= Cov(\hat{\beta}_{CA}) - Cov(\hat{\beta}_{CA})H_1'\left(H_1 Cov(\hat{\beta}_{CA})H_1'\right)^{-1}H_1 Cov(\hat{\beta}_{CA}). \end{aligned} \tag{2.24}$$

Comparing the covariance matrix of $\hat{\beta}_{CA,H}$ with that of $\hat{\beta}_{CA}$, we have

$$\begin{aligned} &Cov(\hat{\beta}_{CA}) - Cov(\hat{\beta}_{CA,H}) \\ &\quad = Cov(\hat{\beta}_{CA})H_1'\left(H_1 Cov(\hat{\beta}_{CA})H_1'\right)^{-1} H_1 Cov(\hat{\beta}_{CA}), \end{aligned} \tag{2.25}$$

which is positive definite since rank $(H_1) = k$. This means that the estimator $\hat{\beta}_{CA,H}$ is better than the estimator $\hat{\beta}_{CA}$. □

2.2 Σ is Unknown

In Section 2.1, we show that the facts that $\hat{\beta}_{H,CA}$ is better than $\hat{\beta}_H$ and $\hat{\beta}_{CA,H}$ is better than $\hat{\beta}_{CA}$, respectively. However the covariance matrix Σ may be unknown in many situations which cause that $\hat{\beta}_{H,CA}$ and $\hat{\beta}_{CA,H}$ are not available to use. We propose the two-stage estimator in this Section. Assuming the rows of (e_1, e_2) are independently distributed, and each has a bivariate normal distribution $N_2(0,\Sigma)$, we first define the estimator of Σ as follows

$$\begin{aligned} S &= \begin{pmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{pmatrix} = \frac{Y'N^*Y}{n-R} \\ &= \begin{pmatrix} \frac{y_1'N^*y_1}{n-R} & \frac{y_1'N^*y_2}{n-R} \\ \frac{y_2'N^*y_1}{n-R} & \frac{y_2'N^*y_2}{n-R} \end{pmatrix}, \end{aligned} \tag{2.26}$$

where $Y = (y_1, y_2)$, $N^* = I - X^*(X^{*\prime}X^*)^- X^{*\prime}$, $X^* = (X_1, X_2)$ and $R = \text{rank}(X^*)$.

Hence, the two-stage covariance-adjusted constraint estimator is given by

$$\hat{\beta}_{H,CA}(S) = \hat{\beta}_{CA}(S) - (X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}\left(H_1\hat{\beta}_{CA}(S) - d\right), \tag{2.27}$$

where

$$\hat{\beta}_{CA}(S) = \hat{\beta}_{LS} - \frac{s_{12}}{s_{22}}(X_1'X_1)^{-1}X_1'N_2 y_2. \tag{2.28}$$

Theorem 2.5

If the sample size $n > \frac{1-\rho^2}{\rho^2} + R + 2$, then the two-stage covariance-adjusted constraint estimator $\hat{\beta}_{H,CA}(S)$ is better than $\hat{\beta}_H$ in the sense of having less covariance.

Proof:

Note that

$$\frac{s_{12}}{s_{22}}=\sigma_{11}^{\frac{1}{2}}\sigma_{22}^{-\frac{1}{2}}\left(\rho+\frac{\sqrt{1-\rho^2}}{n\text{-}R}\times\frac{t_{12}}{t_{22}/(n-R)}\right), \tag{2.29}$$

where $\frac{t_{12}}{t_{22}/(n-R)}$ follows the student's t distribution with degree of freedom $n-R$ (also see Liu(2002)).From the properties of t-distribution, we have

$$E\left(\frac{s_{12}}{s_{22}}\right)=\frac{\sigma_{12}}{\sigma_{22}},\ E\left[\left(\frac{s_{12}}{s_{22}}\right)^2\right]=\sigma_{11}\sigma_{22}^{-1}\left(\rho^2+\frac{1-\rho^2}{n-R-2}\right), \tag{2.30}$$

and $E\left[\hat{\beta}_{CA}(S)\,|\,\frac{s_{12}}{s_{22}}\right]=\beta$ which makes $Cov\left(E\left[\hat{\beta}_{CA}(S)\,|\,\frac{s_{12}}{s_{22}}\right]\right)=0.$ Hence, we have

$$\begin{aligned}
Cov\left[\hat{\beta}_{CA}(S)\right]&=E\left(Cov\left[\hat{\beta}_{CA}(S)\,|\,\frac{s_{12}}{s_{22}}\right]\right)\\
&=\sigma_{11}(X_1'X_1)^{-1}+\sigma_{22}(X_1'X_1)^{-1}X_1'N_2X_1(X_1'X_1)^{-1}E\left[\left(\frac{s_{12}}{s_{22}}\right)^2\right]\\
&\qquad-2\sigma_{12}(X_1'X_1)^{-1}X_1'N_2X_1(X_1'X_1)^{-1}E\left[\frac{s_{12}}{s_{22}}\right]\\
&=\sigma_{11}(X_1'X_1)^{-1}+\sigma_{11}\left(\rho^2+\frac{1-\rho^2}{n-R-2}\right)(X_1'X_1)^{-1}X_1'N_2X_1(X_1'X_1)^{-1}\\
&\quad-2\sigma_{11}\rho^2(X_1'X_1)^{-1}X_1'N_2X_1(X_1'X_1)^{-1}\\
&=\sigma_{11}(X_1'X_1)^{-1}-\sigma_{11}\left(\rho^2-\frac{1-\rho^2}{n-R-2}\right)\\
&\qquad\times(X_1'X_1)^{-1}X_1'N_2X_1(X_1'X_1)^{-1}.
\end{aligned} \tag{2.31}$$

Thus, the covariance matrix of $\hat{\beta}_{H,CA}(S)$ is

$$\begin{aligned}
Cov\left[\hat{\beta}_{H,CA}(S)\right]&=\left[I_n-(X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}H_1\right]Cov\left(\hat{\beta}_{CA}(S)\right)\\
&\qquad\times\left[I_n-(X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}H_1\right]'
\end{aligned}$$

$$= \left[I_n - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1' \right)^{-1} H_1 \right]$$

$$\times \left[\sigma_{11} (X_1'X_1)^{-1} - \sigma_{11} \left(\rho^2 - \frac{1-\rho^2}{n-R-2} \right) (X_1'X_1)^{-1} X_1' N_2 X_1 (X_1'X_1)^{-1} \right]$$

$$\times \left[I_n - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1' \right)^{-1} H_1 \right]'$$

$$= \sigma_{11} (X_1'X_1)^{-1} - \sigma_{11} (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1' \right)^{-1} H_1 (X_1'X_1)^{-1}$$

$$+ \sigma_{11} \left(\rho^2 - \frac{1-\rho^2}{n-R-2} \right) \left[I_n - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1' \right)^{-1} H_1 \right]$$

$$\times (X_1'X_1)^{-1} X_1' N_2 X_1 (X_1'X_1)^{-1} \left[I_n - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1' \right)^{-1} H_1 \right]'$$

$$= Cov(\hat{\beta}_H) + \sigma_{11} \left(\rho^2 - \frac{1-\rho^2}{n-R-2} \right)$$

$$\times \left[I_n - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1' \right)^{-1} H_1 \right] (X_1'X_1)^{-1} X_1'$$

$$\times N_2 X_1 (X_1'X_1)^{-1} \left[I_n - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1' \right)^{-1} H_1 \right]' . \tag{2.32}$$

Comparing it with the covariance of $\hat{\beta}_H$, we have

$$Cov(\hat{\beta}_H) - Cov\left(\hat{\beta}_{H,CA}(S)\right)$$

$$= \sigma_{11} \left(\rho^2 - \frac{1-\rho^2}{n-R-2} \right) \left[I_n - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1' \right)^{-1} H_1 \right]$$

$$\times (X_1'X_1)^{-1} X_1' N_2 X_1 (X_1'X_1)^{-1} \left[I_n - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1' \right)^{-1} H_1 \right]' . \tag{2.33}$$

Obviously, if $\rho^2 - \frac{1-\rho^2}{n-R-2} > 0$ (or $n > \frac{1-\rho^2}{\rho^2} + R + 2$), then the matrix $Cov\left(\hat{\beta}_H\right) - Cov\left(\hat{\beta}_{H,CA}(S)\right)$ is nonnegative definite. This shows that the two-stage estimator $\hat{\beta}_{H,CA}(S)$ is better than $\hat{\beta}_H$. □

Similarly, we define the two-stage constrained covariance adjustment estimator for β as

$$\hat{\beta}_{CA,H}(S) = \hat{\beta}_{CA}(S) - \hat{C}ov(\hat{\beta}_{CA})H_1'\left(H_1\hat{C}ov(\hat{\beta}_{CA})H_1'\right)^{-1}\left(H_1\hat{\beta}_{CA}(S) - d\right), \tag{2.34}$$

where $\hat{\beta}_{CA}(S)=\hat{\beta}_{LS} - \frac{s_{12}}{s_{22}}(X_1'X_1)^{-1}X_1'N_2y_2$ and $\hat{C}ov(\hat{\beta}_{CA}) = s_{11}(X_1'X_1)^{-1} - \frac{s_{11}^2}{s_{22}}(X_1'X_1)^{-1}$ $X_1'N_2X_1(X_1'X_1)^{-1}$ is the estimator of $Cov(\hat{\beta}_{CA})$.

In the following, we numerically exhibit the superiorities of the two-stage constrained covariance adjustment estimator and the two-stage covariance-adjusted constraint estimator. Define the MSE's reduction percentage of $\hat{\beta}_{CA,H}(S)$ and $\hat{\beta}_{H,CA}(S)$, respectively, as

$$r_1 = \frac{\mathrm{MSE}[\hat{\beta}_{CA}] - \mathrm{MSE}\left[\hat{\beta}_{CA,H}(S)\right]}{\mathrm{MSE}[\hat{\beta}_{CA}]}, \tag{2.35}$$

$$r_2 = \frac{\mathrm{MSE}[\hat{\beta}_{H}] - \mathrm{MSE}\left[\hat{\beta}_{H,CA}(S)\right]}{\mathrm{MSE}[\hat{\beta}_{H}]}, \tag{2.36}$$

where MSE means the mean squares error.

Let linear constraint be $(1,1)\beta = 3$, that is $H = (1,1)$, $d = 3$ and $\beta = (1,2)^T$, $\gamma = (2,3,4)^T$, and set $n = 20, 50, 100$, and $\rho = 0.2, 0.5, 0.95, -0.95, -0.5, -0.2$. We generate $2n$ elements of $n\times 2$ matrix X_1 iid from the uniform distribution $\mathrm{U}(1,100)$ such that rank $\left(X_1\right) = 2$, generate $3n$ elements of $n\times 3$ matrix X_2 iid from the uniform distribution $\mathrm{U}(1,100)$ such that rank $\left(X_2\right) = 3$, and generate each row of $\left(e_1, e_2\right)$ from a bivariate normal distribution $N_2\left(0, \Sigma^*\right)$, where Σ^* is the covariance matrix with diagonal elements being 1 and non-diagonal elements being ρ.

From Table 2.1, Table 2.2 and Table 2.3, we find that the MSEs of the two-stage estimators are both influenced by the size of samples and the correlation coefficient. For the given correlation coefficient ρ, the MSEs are reducing as the sample size gets large. For the given sample size, the MSEs are reducing with the increasing of the absolute value of ρ. Comparing the MSE of $\hat{\beta}_{H,CA}(S)$ with that of $\hat{\beta}_H$, we see r_2 is always positive, which means $\hat{\beta}_{H,CA}(S)$ is always better than $\hat{\beta}_H$. Comparing the MSE of $\hat{\beta}_{CA,H}(S)$ with that of $\hat{\beta}_{CA}$, we also see r_1 is always positive, which means $\hat{\beta}_{CA,H}(S)$ is always better than $\hat{\beta}_{CA}$.

Table 2.1
The MSEs of Two-Stage Estimators with $n = 20$

n	ρ	MSE[$\hat{\beta}_{CA}$]	MSE[$\hat{\beta}_H$]	MSE[$\hat{\beta}_{CA,H}(S)$]	MSE[$\hat{\beta}_{H,CA}(S)$]	r_1	r_2
20	0.2	8.66×10^{-5}	8.07×10^{-5}	6.36×10^{-5}	6.47×10^{-5}	0.10	0.20
	-0.2	4.80×10^{-5}	5.42×10^{-5}	4.50×10^{-5}	4.50×10^{-5}	0.06	0.17
	0.5	9.85×10^{-5}	8.23×10^{-5}	7.34×10^{-5}	7.25×10^{-5}	0.25	0.12
	-0.5	8.92×10^{-5}	6.49×10^{-5}	5.73×10^{-5}	5.73×10^{-5}	0.35	0.11
	0.95	1.29×10^{-4}	3.30×10^{-5}	1.68×10^{-5}	1.68×10^{-5}	0.86	0.49
	-0.95	6.92×10^{-5}	2.67×10^{-5}	2.39×10^{-5}	2.15×10^{-5}	0.65	0.20

Table 2.2
The MSEs of Two-Stage Estimators with $n = 50$

n	ρ	MSE[$\hat{\beta}_{CA}$]	MSE[$\hat{\beta}_H$]	MSE[$\hat{\beta}_{CA,H}(S)$]	MSE[$\hat{\beta}_{H,CA}(S)$]	r_1	r_2
50	0.2	2.48×10^{-5}	2.72×10^{-5}	2.30×10^{-5}	2.31×10^{-5}	0.07	0.15
	-0.2	1.86×10^{-5}	1.88×10^{-5}	1.48×10^{-5}	4.50×10^{-5}	0.21	0.21
	0.5	2.46×10^{-5}	2.19×10^{-5}	1.90×10^{-5}	1.88×10^{-5}	0.22	0.14
	-0.5	2.34×10^{-5}	1.99×10^{-5}	1.62×10^{-5}	1.62×10^{-5}	0.31	0.18
	0.95	3.72×10^{-4}	7.09×10^{-6}	4.12×10^{-6}	4.15×10^{-6}	0.89	0.41
	-0.95	2.84×10^{-5}	7.44×10^{-6}	4.79×10^{-6}	4.82×10^{-6}	0.83	0.35

Table 2.3
The MSEs of Two-Stage Estimators with $n = 100$

n	ρ	MSE[$\hat{\beta}_{CA}$]	MSE[$\hat{\beta}_H$]	MSE[$\hat{\beta}_{CA,H}(S)$]	MSE[$\hat{\beta}_{H,CA}(S)$]	r_1	r_2
100	0.2	1.09×10^{-5}	1.13×10^{-5}	1.00×10^{-5}	1.00×10^{-5}	0.08	0.11
	-0.2	1.16×10^{-5}	1.24×10^{-5}	1.06×10^{-5}	1.06×10^{-5}	0.08	0.14
	0.5	1.45×10^{-5}	1.22×10^{-5}	1.05×10^{-5}	1.05×10^{-5}	0.22	0.14
	-0.5	1.03×10^{-5}	1.05×10^{-5}	8.30×10^{-5}	8.32×10^{-6}	0.19	0.20
	0.95	1.21×10^{-5}	3.09×10^{-6}	1.86×10^{-6}	1.87×10^{-6}	0.85	0.39
	-0.95	1.10×10^{-5}	3.17×10^{-6}	1.69×10^{-6}	1.65×10^{-6}	0.85	0.48

3. THE MATRIX SERIES AND COVARIANCE ADJUSTMENT SEQUENCES OF β

3.1 The Matrix Series of β

In this Section we employ the matrix power series to obtain the estimator of β in the system (1.1). Combining (1) and (2) in (1.1) yields a constrained regression model

$$\begin{cases} y = X\alpha + e, \\ H\alpha = d, \end{cases} \tag{3.37}$$

where $y = (y_1', y_2')'$, $H = (H_1, 0)$, $X = \text{diag}(X_1, X_2)$, $\alpha = (\beta', \gamma')'$, $e = (e_1', e_2')' \sim (0, \Sigma \otimes I_n)$.

Denote $\tilde{y}=(\Sigma\otimes I_n)^{-\frac{1}{2}}y$, $\tilde{X}=(\Sigma\otimes I_n)^{-\frac{1}{2}}X$, $\tilde{e}=(\Sigma\otimes I_n)^{-\frac{1}{2}}e$, thus we transform (3.37) to

$$\begin{cases}\tilde{y}=\tilde{X}\alpha+\tilde{e},\\ H\alpha=d,\end{cases} \tag{3.38}$$

where $\tilde{e}\sim(0,I_{2n})$. We build an auxiliary function by using the language multiplier method as follows

$$\begin{aligned}F(\alpha,\lambda)&=\left\|\tilde{y}-\tilde{X}\alpha\right\|^2+2\lambda'(H\alpha-d)\\ &=(y-X\alpha)'(\Sigma\otimes I_n)^{-1}(y-X\alpha)+2\lambda'(H\alpha-d),\end{aligned} \tag{3.39}$$

where $\lambda=(\lambda_1,\ldots,\lambda_k)'$ are lagrange multiplier.

Minimizing the function $F(\alpha,\lambda)$ with respect to α yields

$$\hat{\alpha}=\left(X'(\Sigma\otimes I_n)^{-1}X\right)^{-1}X'(\Sigma\otimes I_n)^{-1}y-\left(X'(\Sigma\otimes I_n)^{-1}X\right)^{-1}H'\hat{\lambda}, \tag{3.40}$$

where $\hat{\lambda}=\left[H\left(X'(\Sigma\otimes I_n)^{-1}X\right)^{-1}H'\right]^{-1}\left[H\left(X'(\Sigma\otimes I_n)^{-1}X\right)^{-1}X'(\Sigma\otimes I_n)^{-1}y-d\right].$

By the inverse of a partitioned matrix, we denote

$$\begin{aligned}(\Sigma\otimes I_n)^{-1}&=\begin{pmatrix}\sigma_{11}I_n & \sigma_{12}I_n\\ \sigma_{12}I_n & \sigma_{22}I_n\end{pmatrix}^{-1}\\ &=\begin{pmatrix}\sigma_{11}^{-1}(1-\rho^2)^{-1}I_n & -\sigma_{12}\sigma_{11}^{-1}\sigma_{22}^{-1}(1-\rho^2)^{-1}I_n\\ -\sigma_{12}\sigma_{11}^{-1}\sigma_{22}^{-1}(1-\rho^2)^{-1}I_n & \sigma_{22}^{-1}(1-\rho^2)^{-1}I_n\end{pmatrix}\\ &\triangleq\begin{pmatrix}\Sigma^{11} & \Sigma^{12}\\ \Sigma^{12} & \Sigma^{22}\end{pmatrix}\end{aligned} \tag{3.41}$$

Further we have

$$\begin{aligned}&\left[X'(\Sigma\otimes I_n)^{-1}X\right]^{-1}\\ &=\left[\begin{pmatrix}X_1' & 0\\ 0 & X_2'\end{pmatrix}\begin{pmatrix}\sigma_{11}^{-1}(1-\rho^2)^{-1}I_n & -\sigma_{12}\sigma_{11}^{-1}\sigma_{22}^{-1}(1-\rho^2)^{-1}I_n\\ -\sigma_{12}\sigma_{11}^{-1}\sigma_{22}^{-1}(1-\rho^2)^{-1}I_n & \sigma_{22}^{-1}(1-\rho^2)^{-1}I_n\end{pmatrix}\begin{pmatrix}X_1 & 0\\ 0 & X_2\end{pmatrix}\right]^{-1}\\ &\triangleq\begin{pmatrix}W^{11} & W^{12}\\ W^{12} & W^{22}\end{pmatrix}\triangleq W,\end{aligned} \tag{3.42}$$

where, employing the matrix power series, W^{11} and W^{12} can be expressed as (also see Wang (2011)).

$$W^{11} = \sigma_{11}(1-\rho^2)(X_1'X_1)^{-1}\sum_{i=0}^{\infty}\left[\rho^2 X_1'P_2X_1(X_1'X_1)^{-1}\right]^i, \tag{3.43}$$

$$W^{12} = \sigma_{12}(1-\rho^2)(X_1'X_1)^{-1}\sum_{i=0}^{\infty}\left[\rho^2 X_1'P_2X_1(X_1'X_1)^{-1}\right]^i X_1'X_2(X_2'X_2)^{-1}, \tag{3.44}$$

where $P_i = X_i(X_i'X_i)^{-1}X_i'$ $(i=1,2)$.

Note that (3.40) can be represented as

$$\begin{aligned}\hat{\alpha} &= WX'(\Sigma\otimes I_n)^{-1}y - WH'(HWH')^{-1}\left[HWX'(\Sigma\otimes I_n)^{-1}y - d\right]\\ &= \hat{\alpha}_c - WH'(HWH')^{-1}(H\hat{\alpha}_c - d),\end{aligned} \tag{3.45}$$

where

$$\begin{aligned}\hat{\alpha}_c &= WX'(\Sigma\otimes I_n)^{-1}y = \left(X'(\Sigma\otimes I_n)^{-1}X'\right)^{-1}X'(\Sigma\otimes I_n)^{-1}y\\ &= \begin{pmatrix} W^{11} & W^{12}\\ W^{12} & W^{22}\end{pmatrix}\begin{pmatrix} X_1' & 0\\ 0 & X_2'\end{pmatrix}\begin{pmatrix}\Sigma^{11} & \Sigma^{12}\\ \Sigma^{12} & \Sigma^{22}\end{pmatrix}\begin{pmatrix} y_1\\ y_2\end{pmatrix}\\ &= \begin{pmatrix}\left(W^{11}X_1'\Sigma^{11} + W^{12}X_2'\Sigma^{12}\right)y_1 + \left(W^{11}X_1'\Sigma^{12} + W^{12}X_2'\Sigma^{22}\right)y_2\\ \left(W^{12}X_1'\Sigma^{11} + W^{22}X_2'\Sigma^{12}\right)y_1 + \left(W^{12}X_1'\Sigma^{12} + W^{22}X_2'\Sigma^{22}\right)y_2\end{pmatrix}\\ &\triangleq \begin{pmatrix}\tilde{\beta}_c\\ \tilde{\gamma}_c\end{pmatrix}.\end{aligned} \tag{3.46}$$

According to the properties of partitioned matrix, from $\hat{\alpha}$ we can easily obtain the estimator of β as

$$\begin{aligned}\hat{\beta}_{CCA} = &\left(W^{11}X_1'\Sigma^{11} + W^{12}X_2'\Sigma^{12}\right)y_1 + \left(W^{11}X_1'\Sigma^{12} + W^{12}X_2'\Sigma^{22}\right)y_2\\ &-(W^{11}H_1'(H_1W^{11}H_1')^{-1}\Big\{H_1\Big[\left(W^{11}X_1'\Sigma^{11} + W^{12}X_2'\Sigma^{12}\right)y_1\\ &+\left(W^{11}X_1'\Sigma^{12} + W^{12}X_2'\Sigma^{22}\right)y_2\Big] - d\Big\}.\end{aligned} \tag{3.47}$$

Following (3.43) and (3.44) and employing the matrix power series, we have

$$\begin{aligned}\hat{\beta}_c &= \left(W^{11}X_1'\Sigma^{11} + W^{12}X_2'\Sigma^{12}\right)y_1 + \left(W^{11}X_1'\Sigma^{12} + W^{12}X_2'\Sigma^{22}\right)y_2\\ &= (X_1^1X_1)^{-1}X_1'\left[I_n - \rho^2\sum_{i=0}^{\infty}(\rho^2P_2P_1)^iP_2N_1\right]\left(y_1 - \frac{\sigma_{12}}{\sigma_{22}}N_2y_2\right).\end{aligned} \tag{3.48}$$

Thus, (3.47) is expressed as

$$\begin{aligned}
\hat{\beta}_{CCA} &= \hat{\beta}_c - W^{11}H_1'(H_1W^{11}H_1')^{-1}(H_1\hat{\beta}_c - d) \\
&= \left\{ I - W^{11}H_1'(H_1W^{11}H_1')^{-1}H_1 \right\}\hat{\beta}_c + W^{11}H_1'(H_1W^{11}H_1')^{-1}d \\
&= \Big\{ I - (X_1'X_1)^{-1}\left[I_{p_1} - \rho^2 X_1'P_2X_1(X_1'X_1)^{-1} \right]^{-1} H_1' \Big(H_1(X_1'X_1)^{-1} \\
&\quad \times \left[I_{p_1} - \rho^2 X_1'P_2X_1(X_1'X_1)^{-1} \right]^{-1} H_1' \Big)^{-1} H_1 \Big\} \\
&\quad \times (X_1'X_1)^{-1}X_1'\left[I_n - \rho^2 \textstyle\sum_{i=0}^{\infty}(\rho^2 P_2P_1)^i P_2N_1 \right]\left(y_1 - \frac{\sigma_{12}}{\sigma_{22}}N_2y_2 \right) \\
&\quad + (X_1'X_1)^{-1}\left[I_{p_1} - \rho^2 X_1'P_2X_1(X_1'X_1)^{-1} \right]^{-1} H_1' \Big(H_1(X_1'X_1)^{-1} \\
&\quad \times \left[I_{p_1} - \rho^2 X_1'P_2X_1(X_1'X_1)^{-1} \right]^{-1} H_1' \Big)^{-1} d, \qquad (3.49)
\end{aligned}$$

since $W^{11} = \sigma_{11}(1-\rho^2)(X_1'X_1)^{-1}\left[I_{p_1} - \rho^2 X_1'P_2X_1(X_1'X_1)^{-1} \right]^{-1}$.

3.2 The Sequences of Covariance Adjustment

In this subsection, we present the sequences of $\{\hat{\beta}_{H,CA}^{(n)}, n \ge 1\}$ and $\{\hat{\beta}_{CA,H}^{(n)}, n \ge 1\}$, which both are obtained by using the covariance adjustment technique repeatedly. Firstly, we prove the following useful results.

Theorem 3.1

The limit of the sequence of the covariance adjustment estimator of β is equal to $\hat{\beta}_c$.

Proof:

Let $T_1 = \hat{\beta}_{LS}$ and $T_2 = N_2y_2$. Firstly, by lemma 2.1 we can obtain the one-step covariance adjustment estimator $\hat{\beta}_{CA}^{(1)}$. Using the covariance adjustment technique repeatedly, we come to the following sequence $\{\hat{\beta}_{CA}^{(n)}, n \ge 1\}$. And it is easily shown that

$$\begin{cases}
\hat{\beta}_{CA}^{(2k-1)} = (X_1'X_1)^{-1}X_1'[P_1 + \rho^2 N_2N_1]^{k-1}y_1 - \dfrac{\sigma_{12}}{\sigma_{22}}(X_1'X_1)^{-1}X_1'[P_1 + \rho^2 N_2N_1]^{k-1}N_2y_2, \\
\hat{\beta}_{CA}^{(2k)} = (X_1'X_1)^{-1}X_1'[P_1 + \rho^2 N_2N_1]^{k}y_1 - \dfrac{\sigma_{12}}{\sigma_{22}}(X_1'X_1)^{-1}X_1'[P_1 + \rho^2 N_2N_1]^{k-1}N_2y_2,
\end{cases} \qquad (3.50)$$

where $k \ge 1$, and $N_1 = I_n - P_1$.

In fact, from the equation (3.50), let $T_1 = \hat{\beta}_{CA}^{(2k-1)}$ and $T_2 = N_2 y_2$, we obtain $\hat{\beta}_{CA}^{(2k)}$. Similarity, let $T_1 = \hat{\beta}_{CA}^{(2k)}$ and $T_2 = N_1 y_1$, we obtain $\hat{\beta}_{CA}^{(2k+1)}$. When $k \to \infty$, we have

$$\begin{aligned}\hat{\beta}_{CA}^{(\infty)} &= \lim_{k\to\infty} \hat{\beta}_{CA}^{(2k)} = \lim_{k\to\infty} \hat{\beta}_{CA}^{(2k-1)} \\ &= \lim_{m\to\infty} (X_1'X_1)^{-1} X_1' \left[P_1 + \rho^2 N_2 N_1\right]^m \left(y_1 - \frac{\sigma_{12}}{\sigma_{22}} N_2 y_2\right). \end{aligned} \quad (3.51)$$

Using the fact that

$$X_1'\left[P_1 + \rho^2 N_2 N_1\right]^m = X_1' \textstyle\sum_{i=0}^m \left[\rho^2 N_2 N_1\right]^i = X_1' - X_1' \sum_{i=0}^{m-1} (\rho^2 P_2 P_1)^i \rho^2 P_2 N_1,$$

we have

$$\lim_{m\to\infty} X_1'\left[P_1 + \rho^2 N_2 N_1\right]^m = X_1'\left[I_n - \rho^2 \textstyle\sum_{i=0}^{\infty} (\rho^2 P_2 P_1)^i P_2 N_1\right]. \quad (3.52)$$

We further have the following result

$$\begin{aligned}\hat{\beta}_{CA}^{(\infty)} &= \lim_{m\to\infty} (X_1'X_1)^{-1} X_1' \left[P_1 + \rho^2 N_2 N_1\right]^m \left(y_1 - \frac{\sigma_{12}}{\sigma_{22}} N_2 y_2\right) \\ &= (X_1'X_1)^{-1} X_1' \left[I_n - \rho^2 \textstyle\sum_{i=0}^{\infty} (\rho^2 P_2 P_1)^i P_2 N_1\right]\left(y_1 - \frac{\sigma_{12}}{\sigma_{22}} N_2 y_2\right) = \hat{\beta}_c. \end{aligned} \quad (3.53)$$

Theorem 3.1 has been proved. □

Theorem 3.2

The limit of the sequence of constrained covariance adjustment estimator of β is

$$\hat{\beta}_{CA,H}^{(\infty)} = \hat{\beta}_c - Cov(\hat{\beta}_c) H_1' \left(H_1 Cov(\hat{\beta}_c) H_1'\right)^{-1} (H_1 \hat{\beta}_c - d). \quad (3.54)$$

Proof:

Note that $\hat{\beta}_{CA}^{(\infty)}$ is defined in (3.53) and the facts that

$$E[\hat{\beta}_{CA}^{(\infty)}] = (X_1'X_1)^{-1} X_1' \left[I_n - \rho^2 \textstyle\sum_{i=0}^{\infty} (\rho^2 P_2 P_1)^i P_2 N_1\right] E\left(y_1 - \frac{\sigma_{12}}{\sigma_{22}} N_2 y_2\right) = \beta, \quad (3.55)$$

$$E[T_2] = E[H_1 \hat{\beta}_{CA}^{(\infty)} - d] = 0. \quad (3.56)$$

Combining the constraint condition and letting $T_1 = \hat{\beta}_{CA}^{(\infty)}$ and $T_2 = H_1 \hat{\beta}_{CA}^{(\infty)} - d$, we have

$$\hat{\beta}^{(\infty)}_{CA,H} = \hat{\beta}^{(\infty)}_{CA} - Cov(\hat{\beta}^{(\infty)}_{CA}, H_1\hat{\beta}^{(\infty)}_{CA})\left[Cov(H_1\hat{\beta}^{(\infty)}_{CA}, H_1\hat{\beta}^{(\infty)}_{CA})\right]^{-1}(H_1\hat{\beta}^{(\infty)}_{CA} - d)$$
$$= \hat{\beta}^{(\infty)}_{CA} - Cov(\hat{\beta}^{(\infty)}_{CA})H_1'\left(H_1 Cov(\hat{\beta}^{(\infty)}_{CA})H_1'\right)^{-1}(H_1\hat{\beta}^{(\infty)}_{CA} - d). \tag{3.57}$$

Thus, from Theorem 3.1 we have

$$\hat{\beta}^{(\infty)}_{CA,H} = \hat{\beta}_c - Cov(\hat{\beta}_c)H_1'\left(H_1 Cov(\hat{\beta}_c)H_1'\right)^{-1}(H_1\hat{\beta}_c - d). \tag{3.58}$$

Theorem 3.2 has been proved. □

Theorem 3.3

The limit of the sequence of the covariance-adjusted constraint estimator of β is

$$\hat{\beta}^{(\infty)}_{H,CA} = \hat{\beta}_c - (X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}(H_1\hat{\beta}_c - d). \tag{3.59}$$

Proof:

Let $T_1 = \hat{\beta}_H$ and $T_2 = N_2 y_2$. From lemma 2.1 we can obtain the one-step covariance-adjusted constrained estimator $\hat{\beta}^{(1)}_{H,CA}$. By using the covariance adjustment technique repeatedly, we get the sequence $\{\hat{\beta}^{(n)}_{H,CA}, n \geq 1\}$.

$$\begin{cases} \hat{\beta}^{(2k-1)}_{H,CA} = \left[I_{p_1} - (X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}H_1\right](X_1'X_1)^{-1}X_1'\left[P_1 + \rho^2 N_2 N_1\right]^{k-1} y_1 \\ \quad -\dfrac{\sigma_{12}}{\sigma_{22}}\left[I_{p_1} - (X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}H_1\right](X_1'X_1)^{-1} \\ \quad \times X_1'\left[P_1 + \rho^2 N_2 N_1\right]^{k-1} N_2 y_2 - (X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1} d, \\ \hat{\beta}^{(2k)}_{H,CA} = \left[I_{p_1} - (X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}H_1\right](X_1'X_1)^{-1}X_1'\left[P_1 + \rho^2 N_2 N_1\right]^{k} y_1 \\ \quad -\dfrac{\sigma_{12}}{\sigma_{22}}\left[I_{p_1} - (X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}H_1\right](X_1'X_1)^{-1}X_1' \\ \quad \times\left[P_1 + \rho^2 N_2 N_1\right]^{k-1} N_2 y_2 - (X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1} d, \end{cases} \tag{3.60}$$

where $k \geq 1$.

It is easily shown that if letting $T_1 = \hat{\beta}^{(2k-1)}_{H,CA}$ and $T_2 = N_2 y_2$, we get $\hat{\beta}^{(2k)}_{H,CA}$, and letting $T_1 = \hat{\beta}^{(2k)}_{H,CA}$ and $T_2 = N_1 y_1$, we can obtain $\hat{\beta}^{(2k+1)}_{H,CA}$.

Set $k \to \infty$ in (3.60), we have

$$\hat{\beta}_{H,CA}^{(\infty)} = \lim_{k\to\infty} \hat{\beta}_{H,CA}^{(2k)} = \lim_{k\to\infty} \hat{\beta}_{H,CA}^{(2k-1)}$$

$$= \lim_{n\to\infty} \left[I_{p_1} - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1' \right)^{-1} H_1 \right] (X_1'X_1)^{-1} X_1' \left[P_1 + \rho^2 N_2 N_1 \right]^n$$

$$\times \left(y_1 - \frac{\sigma_{12}}{\sigma_{22}} N_2 y_2 \right) - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1' \right)^{-1} d$$

$$= \hat{\beta}_{CA}^{(\infty)} - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1' \right)^{-1} (H_1 \hat{\beta}_{CA}^{(\infty)} - d)$$

$$= \hat{\beta}_c - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1' \right)^{-1} (H_1 \hat{\beta}_c - d). \tag{3.61}$$

Theorem 3.3 has been proved. □

Now we state the following important conclusion.

Theorem 3.4

$$\hat{\beta}_{CA,H}^{(\infty)} = \hat{\beta}_{CCA}.$$

Proof:

Using the partitioned matrix in (3.46), we have

$$Cov \begin{pmatrix} \hat{\beta}_c \\ \hat{\gamma}_c \end{pmatrix} = Cov(\hat{\alpha}_c) = WX'(\Sigma \otimes I_n)^{-1} (\Sigma \otimes I_n)(\Sigma \otimes I_n)^{-1} XW$$

$$= W \left(X'(\Sigma \otimes I_n)^{-1} X \right) W = WW^{-1}W = W$$

$$= \begin{pmatrix} W^{11} & W^{12} \\ W^{21} & W^{22} \end{pmatrix}. \tag{3.62}$$

Noting $Cov(\hat{\beta}_c) = W^{11}$ and the expressions of $\hat{\beta}_{CA,H}^{(\infty)}$ and $\hat{\beta}_{CCA}$, we have

$$\hat{\beta}_{CA,H}^{(\infty)} = \hat{\beta}_c - Cov(\hat{\beta}_c) H_1' \left(H_1 Cov(\hat{\beta}_c) H_1' \right)^{-1} (H_1 \hat{\beta}_c - d)$$

$$= \hat{\beta}_c - W^{11} H_1' (H_1 W^{11} H_1')^{-1} (H_1 \hat{\beta}_c - d)$$

$$\triangleq \hat{\beta}_{CCA}. \tag{3.63}$$

Thus, we conclude that $\hat{\beta}_{CA,H}^{(\infty)} = \hat{\beta}_{CCA}$.

Theorem 3.4 has been proved. □

Remark 3.1

Comparing the limit of the sequence of the covariance-adjusted constraint estimator $\hat{\beta}_{H,CA}^{(\infty)}$ with that of the sequence of the constrained covariance adjustment estimator $\hat{\beta}_{CA,H}^{(\infty)}$ and $\hat{\beta}_{CCA}$, we find that $\hat{\beta}_{CCA}$ is not equal to $\hat{\beta}_{H,CA}^{(\infty)}$ but equal to $\hat{\beta}_{CA,H}^{(\infty)}$.

4. SPECIAL CASES

In this Section we consider two special cases: (1) the column space of X_1 is orthogonal to that of X_2, i.e. $X_1'X_2 = 0$ (see Zellner (1962, 1963)). (2) X_2 is a proper subset of X_1, i.e. $X_1 = (X_2, L)$ (see Revankar (1974)).

Theorem 4.1

If $X_1'X_2 = 0$, then $\hat{\beta}_{H,CA}^{(\infty)} = \hat{\beta}_{CA,H}^{(\infty)} = \hat{\beta}_{CCA}$.

Proof:

If $X_1'X_2 = 0$, (3.42) can be simplified as

$$
\begin{aligned}
(W^{ij})_{2\times 2} &= \begin{pmatrix} \sigma_{11}^{-1}(1-\rho^2)^{-1}X_1'X_1 & -\sigma_{12}\sigma_{11}^{-1}\sigma_{22}^{-1}(1-\rho^2)^{-1}X_1'X_2 \\ -\sigma_{12}\sigma_{11}^{-1}\sigma_{22}^{-1}(1-\rho^2)^{-1}X_2'X_1 & \sigma_{22}^{-1}(1-\rho^2)^{-1}X_2'X_2 \end{pmatrix}^{-1} \\
&= \begin{pmatrix} \sigma_{11}^{-1}(1-\rho^2)^{-1}X_1'X_1 & 0 \\ 0 & \sigma_{22}^{-1}(1-\rho^2)^{-1}X_2'X_2 \end{pmatrix}^{-1} \\
&= \begin{pmatrix} \sigma_{11}(1-\rho^2)(X_1'X_1)^{-1} & 0 \\ 0 & \sigma_{22}(1-\rho^2)(X_2'X_2)^{-1} \end{pmatrix},
\end{aligned} \tag{4.64}
$$

and accordingly

$$
\begin{aligned}
\hat{\alpha}_c &= \begin{pmatrix} \hat{\beta}_c \\ \hat{\gamma}_c \end{pmatrix} = WX'(\Sigma\otimes I_n)^{-1}y \\
&= \begin{pmatrix} (X_1'X_1)^{-1}X_1' & -\sigma_{12}\sigma_{22}^{-1}(X_1'X_1)^{-1}X_1' \\ -\sigma_{11}^{-1}\sigma_{12}(X_2'X_2)^{-1}X_2' & (X_2'X_2)^{-1}X_2' \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\
&= \begin{pmatrix} (X_1'X_1)^{-1}X_1'y_1 - \sigma_{12}\sigma_{22}^{-1}(X_1'X_1)^{-1}X_1'y_2 \\ (X_2'X_2)^{-1}X_2'y_2 - \sigma_{12}\sigma_{11}^{-1}(X_2'X_2)^{-1}X_2'y_1 \end{pmatrix} \\
&\triangleq \begin{pmatrix} \hat{\beta}_{CA}^{*} \\ \hat{\gamma}_{CA}^{*} \end{pmatrix},
\end{aligned} \tag{4.65}
$$

and

$$
[HWH']^{-1} = \sigma_{11}^{-1}(1-\rho^2)^{-1}\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}. \tag{4.66}
$$

Denote

$$\hat{\beta}^*_{CA} = \hat{\beta}_{LS} - \frac{\sigma_{12}}{\sigma_{22}} (X_1'X_1)^{-1} X_1'y_2. \tag{4.67}$$

Then we obtain

$$\hat{\beta}_{CCA} = \hat{\beta}^*_{CA} - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1' \right)^{-1} (H_1 \hat{\beta}^*_{CA} - d). \tag{4.68}$$

Hence, under the assumption $X_1'X_2 = 0$, we have $\hat{\beta}_{CA} = \hat{\beta}^*_{CA}$ and $Cov(\hat{\beta}_{CA}) = Cov(\hat{\beta}^*_{CA}) = \sigma_{11}(1-\rho^2)(X_1'X_1)^{-1}$. Furthermore, $\hat{\beta}^{(\infty)}_{H,CA}$ and $\hat{\beta}^{(\infty)}_{CA,H}$ are simplified as

$$\hat{\beta}^{(\infty)}_{H,CA} = \hat{\beta}^*_{CA} - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1' \right)^{-1} (H_1 \hat{\beta}^*_{CA} - d), \tag{4.69}$$

$$\begin{aligned} \hat{\beta}^{(\infty)}_{CA,H} &= \hat{\beta}^*_{CA} - Cov(\hat{\beta}^*_{CA}) H_1' \left(H_1 Cov(\hat{\beta}^*_{CA}) H_1' \right)^{-1} (H_1 \hat{\beta}^*_{CA} - d) \\ &= \hat{\beta}^*_{CA} - (X_1'X_1)^{-1} H_1' \left(H_1 (X_1'X_1)^{-1} H_1' \right)^{-1} (H_1 \hat{\beta}^*_{CA} - d). \end{aligned} \tag{4.70}$$

From (4.68), (4.69) and (4.70), it is easy to see that $\hat{\beta}^{(\infty)}_{H,CA} = \hat{\beta}^{(\infty)}_{CA,H} = \hat{\beta}_{CCA}$.

Theorem 4.1 has been proved. □

Theorem 4.2

Under the condition that $X_1 = (X_2, L)$, $\hat{\beta}_{CCA} = \hat{\beta}^{(\infty)}_{CA,H} \neq \hat{\beta}^{(\infty)}_{H,CA}$.

Proof:

If $X_1 = (X_2, L)$, then $P_2 P_1 = P_2$ and $P_2 N_1 = 0$. Hence, we have

$$\begin{aligned} W^{11} &= \sigma_{11}(1-\rho^2)(X_1'X_1)^{-1} \textstyle\sum_{i=0}^{\infty} \left[\rho^2 X_1' P_2 X_1 (X_1'X_1)^{-1} \right]^i \\ &= \sigma_{11}(1-\rho^2)(X_1'X_1)^{-1} \left[I_{p_1} + \rho^2 X_1 P_2 X_1 (X_1'X_1)^{-1} + \rho^4 X_1 P_2 X_1 (X_1'X_1)^{-1} + \dots \right] \\ &= \sigma_{11} \Big[(X_1'X_1)^{-1} + \rho^2 (X_1'X_1)^{-1} X_1 P_2 X_1 (X_1'X_1)^{-1} + \rho^4 (X_1'X_1)^{-1} X_1 P_2 X_1 (X_1'X_1)^{-1} \\ &\quad + \rho^6 (X_1'X_1)^{-1} X_1 P_2 X_1 (X_1'X_1)^{-1} \dots - \rho^2 (X_1'X_1)^{-1} \\ &\quad - \rho^4 (X_1'X_1)^{-1} X_1 P_2 X_1 (X_1'X_1)^{-1} - \rho^6 (X_1'X_1)^{-1} X_1 P_2 X_1 (X_1'X_1)^{-1} \dots \Big] \\ &= \sigma_{11} \left[(X_1'X_1)^{-1} + \rho^2 (X_1'X_1)^{-1} X_1 P_2 X_1 (X_1'X_1)^{-1} - \rho^2 (X_1'X_1)^{-1} (X_1'X_1)(X_1'X_1)^{-1} \right] \\ &= \sigma_{11} \left[(X_1'X_1)^{-1} - \rho^2 (X_1'X_1)^{-1} X_1 (I_n - P_2) X_1 (X_1'X_1)^{-1} \right] \\ &= \sigma_{11} (X_1'X_1)^{-1} - \frac{\sigma_{12}^2}{\sigma_{22}} (X_1'X_1)^{-1} X_1' N_2 X_1 (X_1'X_1)^{-1} \\ &= Cov(\hat{\beta}_{CA}), \end{aligned} \tag{4.71}$$

and thus $\hat{\beta}_c$ becomes

$$\hat{\beta}_c = (X_1'X_1)^{-1}X_1'\left[I_n - \rho^2\sum_{i=0}^{\infty}(\rho^2 P_2 P_1)^i P_2 N_1\right]\left(y_1 - \frac{\sigma_{12}}{\sigma_{22}}N_2 y_2\right)$$
$$= \hat{\beta}_{LS} - \frac{\sigma_{12}}{\sigma_{22}}(X_1'X_1)^{-1}X_1'N_2 y_2$$
$$= \hat{\beta}_{CA}. \qquad (4.72)$$

Using (4.71) and (4.72), $\hat{\beta}_{CCA}$, $\hat{\beta}_{CA,H}^{(\infty)}$ and $\hat{\beta}_{H,CA}^{(\infty)}$ are respectively simplified as

$$\hat{\beta}_{CCA} = \hat{\beta}_{CA} - Cov(\hat{\beta}_{CA})H_1'\left(H_1 Cov(\hat{\beta}_{CA})H_1'\right)^{-1}(H_1\hat{\beta}_{CA} - d). \qquad (4.73)$$

$$\hat{\beta}_{CA,H}^{(\infty)} = \hat{\beta}_{CA} - Cov(\hat{\beta}_{CA})H_1'\left(H_1 Cov(\hat{\beta}_{CA})H_1'\right)^{-1}(H_1\hat{\beta}_{CA} - d). \qquad (4.74)$$

$$\hat{\beta}_{CA,H}^{(\infty)} = \hat{\beta}_{CA} - (X_1'X_1)^{-1}H_1'\left(H_1(X_1'X_1)^{-1}H_1'\right)^{-1}(H_1\hat{\beta}_{CA} - d). \qquad (4.75)$$

By (4.73), (4.74) and (4.75), it is easy to see the conclusion of Theorem 4.2 is true.

Theorem 4.2 has been proved. □

Remark 4.1

In fact, under the condition that $X_1 = (X_2, L)$, $\hat{\beta}_{CA,H}^{(\infty)}$ degenerates to $\hat{\beta}_{CA,H}$, which equals to the one-step constrained covariance adjustment estimator.

5. CONCLUSIONS

This paper investigates the parameter estimation problem in the system of two seemingly unrelated regressions (SUR) with linear constraint and obtains the following some interesting results. If the covariance matrix of the SUR system is known, then the one-step covariance-adjusted constraint estimator is better than the constrained least squares estimator in the sense of having less covariance matrix, and the one-step constrained covariance adjustment estimator is better than the covariance adjustment estimator, too. If the covariance matrix of SUR system is unknown, then the two-stage covariance-adjusted constraint estimator is better than the constrained least squares estimator and the two-stage constrained covariance adjustment estimator is better than the covariance adjustment estimator. Moreover, the covariance adjustment estimator with linear constraint can be expressed a matrix series by employing matrix series expansion, which is exactly equal to the limit of the sequence of the constrained covariance adjustment estimator but not equal to the limit of the sequence of the covariance-adjusted constraint estimator. Finally, under the Zellner (1962, 1963)'s assumption, the limit of the sequence of the covariance-adjusted constraint estimator, the limit of the sequence of the constrained covariance adjustment estimator and the covariance adjustment estimator with linear constraint are the same. Under the Revankar (1974)'s assumption, the covariance adjustment estimator with linear constraint is equal to the limit of the constrained covariance adjustment estimator sequence but not equal to the limit of the covariance-adjusted constraint estimator sequence.

ACKNOWLEDGEMENT

We would like to thank the editor and the three referees for their constructive suggestions and comments which improve the early version a lot. The research is partly supported by NNSF of China (11371051).

REFERENCES

1. Armstrong, R.D. and Frome, E.L. (1976). Branch and bound solution of a restricted least squares problem. *Technometrics*, 18(4), 447-450.
2. Baksalary, J.K. and Kala, R. (1979). Covariance adjustment when a vector of parameters is restricted to a given subspace. *Siam Journal on Applied Mathematics*, 37(1), 20-21.
3. Chen, S.J. and Liao, D.F. (1995). An improved estimator for the system of seemingly unrelated regression equations. *Applied Mathematics–A Journal of Chinese Universities*, 10(3), 303-312.
4. Liu, A.Y. (2002). Efficient estimation of two seemingly unrelated regression equations. *Journal of Multivariate Analysis*, 82(2), 445-456.
5. Mantel, N. (1969). Restricted least squares regression and convex quadratic programming. *Technometrics*, 11(4), 763-773.
6. Rao, C.R. (1967). Least square theory using an estimated dispersion matrix and its application to measurement of signals. *Berkeley Symposium on Mathematical Statistics and Probability*, 1, 355-372.
7. Revankar, N. (1974). Some finite sample results in the context of two seemingly unrelated regression equations. *J. Amer. Statist. Assoc.*, 69(345), 187-190.
8. Wang, L.C., Wang, H.M. and Chen, G.J. (2001). The covariance improvement estimation of general seemingly unrelated regression system. *Chinese Applied Probability and Statistics*, 17(2), 156-162.
9. Wang, L.C., Lian, H. and Singh, R.S. (2011). On efficient estimators of two seemingly unrelated regressions. *Statistics and Probability Letters*, 81(5), 563-570.
10. Wang, S.G. and Wang, A.J. (1998). Covariance adjustment technique and its application. *Applied Probability and Statistics*, 14(1), 99-107.
11. Wang, S.G., Shi, J.H. and Yin, S.J. (2003). *Introduction to linear models*, Science Press.
12. Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *J. Amer. Statist. Assoc.*, 57(298), 348-368.
13. Zellner, A. (1963). Estimators for seemingly unrelated regression equations: some exact finite sample results. *J. Amer. Statist. Assoc.*, 58(304), 977-992.