

## Time Series Analysis: Assignment 5

Bowen Ma (12960780), Mianyun He (13605275), Shiyi Yang (13627295)

### Part 1

#### Exercise 1

a.

From the tutorial, we know the vectorization operation  $\text{vec}(\cdot)$  is defined by stacking the columns of a matrix on top of each other.

$$\begin{aligned}\text{vec}(\mathbf{Y}) &= \text{vec}(\mathbf{BZ} + \mathbf{E}) = \text{vec}(\mathbf{BZ}) + \text{vec}(\mathbf{E}) = \text{vec}(\mathbf{I}_m \mathbf{BZ}) + \text{vec}(\mathbf{E}) \\ &= (\mathbf{Z}' \otimes \mathbf{I}_m) \text{vec}(\mathbf{B}) + \text{vec}(\mathbf{E}) = (\mathbf{Z}' \otimes \mathbf{I}_m) \boldsymbol{\beta} + \boldsymbol{\varepsilon}\end{aligned}$$

b.

The covariance matrix of  $\boldsymbol{\varepsilon}$  is given by  $\boldsymbol{\Omega} = \mathbf{I}_n \otimes \boldsymbol{\Omega}_\varepsilon$ , GLS minimizes  $S(\boldsymbol{\beta}) = \boldsymbol{\varepsilon}' (\mathbf{I}_n \otimes \boldsymbol{\Omega}_\varepsilon)^{-1} \boldsymbol{\varepsilon}$

$$\begin{aligned}S(\boldsymbol{\beta}) &= (\mathbf{y} - (\mathbf{Z}' \otimes \mathbf{I}_m) \boldsymbol{\beta})' (\mathbf{I}_n \otimes \boldsymbol{\Omega}_\varepsilon)^{-1} (\mathbf{y} - (\mathbf{Z}' \otimes \mathbf{I}_m) \boldsymbol{\beta}) = (\mathbf{y}' - \boldsymbol{\beta}' (\mathbf{Z}' \otimes \mathbf{I}_m)') (\mathbf{I}_n \otimes \boldsymbol{\Omega}_\varepsilon)^{-1} (\mathbf{y} - (\mathbf{Z}' \otimes \mathbf{I}_m) \boldsymbol{\beta}) \\ &= \mathbf{y}' (\mathbf{I}_n \otimes \boldsymbol{\Omega}_\varepsilon)^{-1} \mathbf{y} + \boldsymbol{\beta}' (\mathbf{Z} \otimes \mathbf{I}_m) (\mathbf{I}_n \otimes \boldsymbol{\Omega}_\varepsilon)^{-1} (\mathbf{Z}' \otimes \mathbf{I}_m) \boldsymbol{\beta} \\ &\quad - \mathbf{y}' (\mathbf{I}_n \otimes \boldsymbol{\Omega}_\varepsilon)^{-1} (\mathbf{Z}' \otimes \mathbf{I}_m) \boldsymbol{\beta} - \boldsymbol{\beta}' (\mathbf{Z} \otimes \mathbf{I}_m) (\mathbf{I}_n \otimes \boldsymbol{\Omega}_\varepsilon)^{-1} \mathbf{y} \\ &= \mathbf{y}' (\mathbf{I}_n \otimes \boldsymbol{\Omega}_\varepsilon^{-1}) \mathbf{y} + \boldsymbol{\beta}' (\mathbf{Z} \otimes \boldsymbol{\Omega}_\varepsilon^{-1}) (\mathbf{Z}' \otimes \mathbf{I}_m) \boldsymbol{\beta} - \mathbf{y}' (\mathbf{Z}' \otimes \boldsymbol{\Omega}_\varepsilon^{-1}) \boldsymbol{\beta} - \boldsymbol{\beta}' (\mathbf{Z} \otimes \boldsymbol{\Omega}_\varepsilon^{-1}) \mathbf{y} \\ &= \mathbf{y}' (\mathbf{I}_n \otimes \boldsymbol{\Omega}_\varepsilon^{-1}) \mathbf{y} + \boldsymbol{\beta}' (\mathbf{Z} \mathbf{Z}' \otimes \boldsymbol{\Omega}_\varepsilon^{-1}) \boldsymbol{\beta} - 2 \boldsymbol{\beta}' (\mathbf{Z} \otimes \boldsymbol{\Omega}_\varepsilon^{-1}) \mathbf{y}\end{aligned}$$

c.

$$S(\boldsymbol{\beta}) = \mathbf{y}' (\mathbf{I}_n \otimes \boldsymbol{\Omega}_\varepsilon^{-1}) \mathbf{y} + \boldsymbol{\beta}' (\mathbf{Z} \mathbf{Z}' \otimes \boldsymbol{\Omega}_\varepsilon^{-1}) \boldsymbol{\beta} - 2 \boldsymbol{\beta}' (\mathbf{Z} \otimes \boldsymbol{\Omega}_\varepsilon^{-1}) \mathbf{y}$$

$$\begin{aligned}\text{F.O.C: } \frac{\partial S(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} &= 0 \\ \Rightarrow 2 (\mathbf{Z} \mathbf{Z}' \otimes \boldsymbol{\Omega}_\varepsilon^{-1}) \boldsymbol{\beta} - 2 (\mathbf{Z} \otimes \boldsymbol{\Omega}_\varepsilon^{-1}) \mathbf{y} &= 0 \\ \Rightarrow \hat{\boldsymbol{\beta}} &= (\mathbf{Z} \mathbf{Z}' \otimes \boldsymbol{\Omega}_\varepsilon^{-1})^{-1} (\mathbf{Z} \otimes \boldsymbol{\Omega}_\varepsilon^{-1}) \mathbf{y} \\ \Rightarrow \hat{\boldsymbol{\beta}} &= ((\mathbf{Z} \mathbf{Z}')^{-1} \otimes \boldsymbol{\Omega}_\varepsilon) (\mathbf{Z} \otimes \boldsymbol{\Omega}_\varepsilon^{-1}) \mathbf{y} \\ \Rightarrow \hat{\boldsymbol{\beta}} &= ((\mathbf{Z} \mathbf{Z}')^{-1} \mathbf{Z} \otimes \mathbf{I}_m) \mathbf{y} \\ \text{S.O.C: } \frac{\partial^2 S(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} &= 2 (\mathbf{Z} \mathbf{Z}' \otimes \boldsymbol{\Omega}_\varepsilon^{-1})\end{aligned}$$

We know that  $\boldsymbol{\Omega}_\varepsilon^{-1}$  is positive definite, and  $\mathbf{Z} \mathbf{Z}'$  is positive definite, since for  $\mathbf{a} \neq 0$ ,  $\mathbf{Z}' \mathbf{a} \neq 0$  and  $\mathbf{a}' \mathbf{Z} \mathbf{Z}' \mathbf{a} = ((\mathbf{Z}' \mathbf{a})' (\mathbf{Z}' \mathbf{a})) > 0$ . The Kronecker product of two p.d. matrix is also p.d.

d.

(i)

$$\begin{aligned}
\hat{\beta} &= \left( (ZZ')^{-1} Z \otimes I_m \right) \mathbf{y} \\
&= \left( (ZZ')^{-1} Z \otimes I_m \right) \left( (Z' \otimes I_m) \beta + \varepsilon \right) \\
&= \left[ (ZZ')^{-1} ZZ' \otimes I_m \right] \beta + \left[ (ZZ')^{-1} Z \otimes I_m \right] \varepsilon \\
&= \left[ (ZZ')^{-1} ZZ' \otimes I_m \right] \text{vec}(\mathbf{B}) + \left( (ZZ')^{-1} Z \otimes I_m \right) \varepsilon \\
&= \text{vec}(\mathbf{I}_m \mathbf{B} \mathbf{I}_{mp+1}) + \left( (ZZ')^{-1} Z \otimes I_m \right) \varepsilon \\
&= \text{vec}(\mathbf{B}) + \left( (ZZ')^{-1} Z \otimes I_m \right) \varepsilon = \beta + \left( (ZZ')^{-1} Z \otimes I_m \right) \varepsilon
\end{aligned}$$

(ii)

$$\begin{aligned}
\text{vec}(\hat{\mathbf{B}}) &= \hat{\mathbf{B}} = ((ZZ')^{-1} Z \otimes I_m) \mathbf{y} \\
&= \underbrace{((ZZ')^{-1} Z \otimes I_m)}_{C'} \underbrace{\text{vec}(\mathbf{Y})}_{A} \underbrace{\text{vec}(\mathbf{B})}_{B} \\
&= \text{vec}(\mathbf{I}_m \mathbf{Y} \mathbf{Z}' (ZZ')^{-1}) = \text{vec}(\mathbf{Y} \mathbf{Z}' (ZZ')^{-1})
\end{aligned}$$

(iii)

Since  $\text{vec}(\hat{\mathbf{B}}) = \text{vec}(\mathbf{Y} \mathbf{Z}' (ZZ')^{-1})$ ,

$$\begin{aligned}
\hat{\mathbf{B}} &= \mathbf{Y} \mathbf{Z}' (ZZ')^{-1} \\
&= (\mathbf{B} \mathbf{A} + \mathbf{E}) \mathbf{Z}' (ZZ')^{-1} \\
&= \mathbf{B} (ZZ') (ZZ')^{-1} + \mathbf{E} \mathbf{Z}' (ZZ')^{-1} \\
&= \mathbf{B} + \mathbf{E} \mathbf{Z}' (ZZ')^{-1}
\end{aligned}$$

e.

$$\Omega_{\varepsilon} = E[\varepsilon_t \varepsilon_t'], \sum_{t=1}^n \hat{\varepsilon}_t \hat{\varepsilon}_t' = \hat{\mathbf{E}} \hat{\mathbf{E}}', \text{ where } \hat{\mathbf{E}} = \mathbf{Y} - \hat{\mathbf{B}} \mathbf{Z} = \mathbf{Y} - \underbrace{\mathbf{Y} \mathbf{Z}' (ZZ')^{-1} \mathbf{Z}}_{\hat{\mathbf{B}}}$$

$$\begin{aligned}
\tilde{\Omega}_{\varepsilon} &= \frac{1}{n} \hat{\mathbf{E}} \hat{\mathbf{E}}' = \frac{1}{n} \left( \mathbf{Y} - \mathbf{Y} \mathbf{Z}' (ZZ')^{-1} \mathbf{Z} \right) \left( \mathbf{Y} - \mathbf{Y} \mathbf{Z}' (ZZ')^{-1} \mathbf{Z} \right)' \\
&= \frac{1}{n} \left( \mathbf{Y} - \mathbf{Y} \mathbf{Z}' (ZZ')^{-1} \mathbf{Z} \right) \left( \mathbf{Y}' - \mathbf{Z}' (ZZ')^{-1} \mathbf{Z} \mathbf{Y}' \right) \\
&= \frac{1}{n} \left[ \mathbf{Y} \mathbf{Y}' - \mathbf{Y} \mathbf{Z}' (ZZ')^{-1} \mathbf{Z} \mathbf{Y}' \right] = \frac{1}{n} \mathbf{Y} \left( \mathbf{I}_n - \mathbf{Z}' (ZZ')^{-1} \mathbf{Z} \right) \mathbf{Y}'
\end{aligned}$$

Correct for d.o.f. :  $\hat{\Omega}_{\varepsilon} = \frac{n}{n-mp-1} \tilde{\Omega}_{\varepsilon}$ .

f.

We know that  $\varepsilon_t \stackrel{iid}{\sim} \mathcal{N}(0, \Omega_{\varepsilon})$ , hence  $\mathbf{y}_t \stackrel{iid}{\sim} \mathcal{N}(\mathbf{B} \mathbf{Z}_t, \Omega_{\varepsilon})$

$$\begin{aligned}
f(\mathbf{y}_t) &= \frac{1}{(2\pi)^{\frac{m}{2}} |\Omega_{\varepsilon}|^{\frac{1}{2}}} \exp \left( -\frac{1}{2} \varepsilon_t' \Omega_{\varepsilon}^{-1} \varepsilon_t \right) = (\det(2\pi \Omega_{\varepsilon}))^{-\frac{1}{2}} \exp \left( -\frac{1}{2} \varepsilon_t' \Omega_{\varepsilon}^{-1} \varepsilon_t \right) \\
\log f(\mathbf{y}_t) &= -\frac{1}{2} \log(\det(2\pi \Omega_{\varepsilon})) - \frac{1}{2} \varepsilon_t' \Omega_{\varepsilon}^{-1} \varepsilon_t \\
\text{Log-likelihood: } l &= -\frac{n}{2} \log(\det(2\pi \Omega_{\varepsilon})) - \frac{1}{2} \sum_{t=1}^n \varepsilon_t' \Omega_{\varepsilon}^{-1} \varepsilon_t
\end{aligned}$$

MLE is obtained by maximize log-likelihood, which is equivalent to minimize  $\sum_{t=1}^n \boldsymbol{\varepsilon}_t' \boldsymbol{\Omega}_{\varepsilon}^{-1} \boldsymbol{\varepsilon}_t$ :

$$\min \sum_{t=1}^n \boldsymbol{\varepsilon}_t' \boldsymbol{\Omega}_{\varepsilon}^{-1} \boldsymbol{\varepsilon}_t = \min \boldsymbol{\varepsilon}' (\boldsymbol{I}_n \otimes \boldsymbol{\Omega}_{\varepsilon})^{-1} \boldsymbol{\varepsilon}$$

Hence, the GLS and MLE estimators of the coefficients coincide.

## Part 2

a.

The R script imports the daily data of the 13-week Treasury Bill and the 5-year Treasury Yield from January 2, 2002, up to and including December 29, 2023. Any null values in the dataset are replaced with NA and then removed. The script further processes the data to extract the adjusted closing prices, ensuring they are converted to the appropriate data types, and using them as our time series  $\{x_t\}$  and  $\{y_t\}$ .

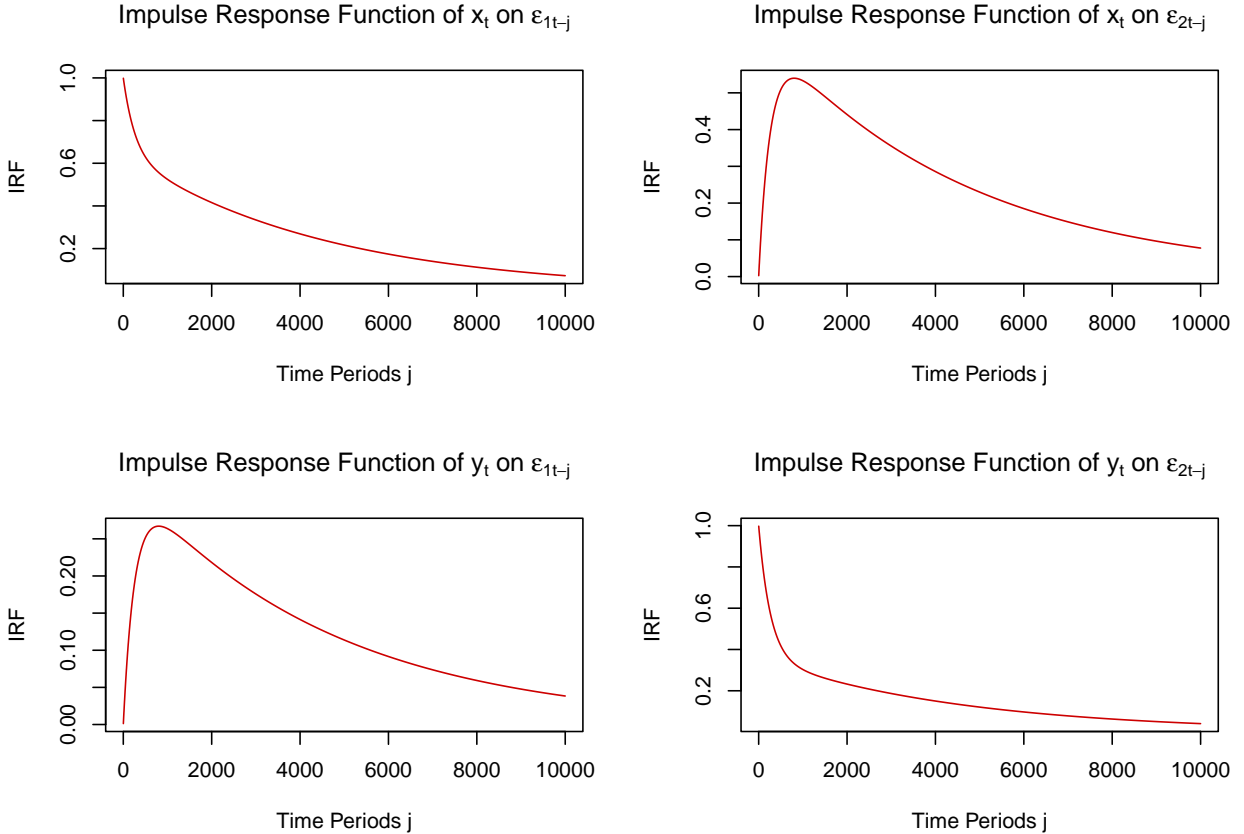
In section a, we estimate the parameters for the two-dimensional VAR(1) model by applying the formula in section d of part 1, which is  $\hat{\mathbf{B}} = \mathbf{Y}\mathbf{Z}'(\mathbf{Z}\mathbf{Z}')^{-1}$ , and we have:

$$\hat{\boldsymbol{\alpha}} = \begin{pmatrix} -0.003158298 \\ 0.004261147 \end{pmatrix} \quad \hat{\boldsymbol{\Phi}} = \begin{pmatrix} 0.99850877 & 0.002423684 \\ 0.00119926 & 0.997501019 \end{pmatrix}$$

b.

With the resulting coefficients in section a, it can be observed that  $\hat{\phi}_{12}$  and  $\hat{\phi}_{21}$  are very close to zero. This implies that it cannot be assured that  $\{y_t\}$  Granger causes  $\{x_t\}$  or the other way around. Hence, we need to test Granger causality with the appropriate F-test.

The impulse response function for the two-dimensional VAR(1) model is calculated as  $\frac{\partial \mathbf{Y}_t}{\partial \boldsymbol{\varepsilon}_{t-j}} = \boldsymbol{\Phi}^j$ . Based on this result, we have 4 plots of the impulse response functions:



From the first and fourth plots of the impulse response function, it is clear that the impact of the shock  $\varepsilon_{1t-j}$  on  $x_t$  dies out geometrically but slowly as  $j \rightarrow \infty$ , similarly to the effect of the shock  $\varepsilon_{2t-j}$  on  $y_t$ . However, the shock  $\varepsilon_{2t-j}$  on  $x_t$  initially increases and then decreases geometrically toward zero, again similarly to the behavior of the shock  $\varepsilon_{1t-j}$  on  $y_t$ . The eigenvalues of  $\hat{\boldsymbol{\Phi}}$  obtained in section a are close to 1, suggesting slow shock decay, as verified by our plots. The observed positive impact of shock in the second and third plots may results from delayed and indirect effects of the exotic shock from the other process on its own, for instance, the shock  $\varepsilon_{2t-j}$  from the process  $\{y_t\}$  on  $\{x_t\}$ .

c.

The covariance matrix  $\mathbf{\Omega}$  is estimated using  $\hat{\mathbf{\Omega}} = \frac{1}{n-mp-1} \mathbf{Y}(\mathbf{I}_n - \mathbf{Z}'(\mathbf{Z}\mathbf{Z}')^{-1}\mathbf{Z})\mathbf{Y}'$ , and we have:

$$\hat{\mathbf{\Omega}} = \begin{pmatrix} 0.0019047273 & 0.0007340099 \\ 0.0007340099 & 0.0036351834 \end{pmatrix}$$

The corresponding log-likelihood is computed using  $l = -\frac{n}{2} \log(\det(2\pi\mathbf{\Omega}_\epsilon)) - \frac{1}{2} \sum_{t=1}^n \epsilon_t' \mathbf{\Omega}_\epsilon^{-1} \epsilon_t$ , and we have 17379.99 for the estimated two-dimensional VAR(1) model

d.

In section d, we perform the estimations for VAR models with two and three lags, VAR(2) and VAR(3), and form a table with the AIC, the log-likelihood and the total number of parameters for the VAR(1), VAR(2) and VAR(3) models. When counting the number of parameters, we also include the unknown parameters in the covariance matrix  $\mathbf{\Omega}$ .

Table 1: Model selection

	AIC	log-likelihood	# parameters
VAR(1)	-11.96007	17379.99	9
VAR(2)	-11.96621	17399.82	13
VAR(3)	-11.97644	17430.94	17

From Table 1, we found that the two-dimensional VAR(3) model gives the lowest AIC and highest log-likelihood values, hence we choose VAR(3).

e.

Based on our results from section d, the estimated VAR(3) model is:

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} -0.0031 \\ 0.0038 \end{pmatrix} + \begin{pmatrix} 1.1000 & -0.0232 \\ 0.0276 & 0.9728 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} -0.1889 & 0.0604 \\ -0.0954 & -0.0184 \end{pmatrix} \begin{pmatrix} x_{t-2} \\ y_{t-2} \end{pmatrix} + \begin{pmatrix} 0.0875 & -0.0348 \\ 0.0688 & 0.0434 \end{pmatrix} \begin{pmatrix} x_{t-3} \\ y_{t-3} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}$$

We define  $\Phi(z) = I_2 - \Phi_1 z - \Phi_2 z^2 - \Phi_3 z^3$ , and  $C(z)$  as the cofactor matrix of  $\Phi(z)$ :

$$\Phi(z) = \begin{pmatrix} 1 - 1.1000z + 0.1889z^2 - 0.0875z^3 & 0.0232z - 0.0604z^2 + 0.0348z^3 \\ -0.0276z + 0.0954z^2 - 0.0688z^3 & 1 - 0.9728z + 0.0184z^2 - 0.0434z^3 \end{pmatrix}$$

$$C(z) = \begin{pmatrix} 1 - 0.9728z + 0.0184z^2 - 0.0434z^3 & -0.0232z + 0.0604z^2 - 0.0348z^3 \\ 0.0276z - 0.0954z^2 + 0.0688z^3 & 1 - 1.1000z + 0.1889z^2 - 0.0875z^3 \end{pmatrix}$$

Multiply  $C(L)$  on both sides, we have  $d(L)I_2 \mathbf{Y}_t = C(L)(\boldsymbol{\alpha} + \boldsymbol{\epsilon}_t)$ :

$$\begin{pmatrix} 0.006L^6 - 0.017L^5 + 0.145L^4 - 0.339L^3 + 1.278L^2 - 2.073L + 1 & 0 \\ 0 & 0.006L^6 - 0.017L^5 + 0.145L^4 - 0.339L^3 + 1.278L^2 - 2.073L + 1 \end{pmatrix} \begin{pmatrix} x_t \\ y_t \end{pmatrix} \\ = \begin{pmatrix} 0.0000023 \\ 0.0000022 \end{pmatrix} + \begin{pmatrix} 1 - 0.9728L + 0.0184L^2 - 0.0434L^3 & -0.0232L + 0.0604L^2 - 0.0348L^3 \\ 0.0276L - 0.0954L^2 + 0.0688L^3 & 1 - 1.1000L + 0.1889L^2 - 0.0875L^3 \end{pmatrix} \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}$$

Hence, we verify that both  $\{x_t\}$  and  $\{y_t\}$  have a ARMA(6,3) process.

f.

We perform the Johansen's trace test for cointegration on the two-dimensional VAR(3) model by computing the  $LR(r)$  statistics for  $r = 0$  and  $r = 1$  via the following formula. The interpretation of the canonical correlation coefficients  $\hat{\lambda}_j$ 's is that they measure the correlations between  $\Delta Y_t$  and  $Y_{t-1}$ :

$$LR(r) = 2\{\log(L_{max}(m)) - \log(L_{max}(r))\} = -(n - p) \sum_{j=r+1}^m \log(1 - \hat{\lambda}_j)$$

The results of the test statistics and the critical values, given in Exhibit 7.31 of Heij et al., are presented in the following table:

Table 2: Johansen's trace test for VAR(3)

	LR(r)	Critical Value
r=0	18.55	15.41
r=1	0.22	3.76

The interpretation of the test results is that we first reject the null hypothesis  $H_0 : r = 0$  at 5% significance level, which indicates there is no cointegration relation in the model. Then we fail to reject the null hypothesis  $H_0 : r = 1$  at 5 % significance level, which suggests the model has one cointegration relation. Hence, there exists one cointegration relation between  $\{x_t\}$  and  $\{y_t\}$  in the two-dimensional VAR(3) model.