THE UNIVERSITY OF CHICAGO
Graduate School of Business
Business 41914, Spring Quarter 2007, Mr. Ruey S. Tsay

Lecture 3: Vector ARMA Models (continued.)

Reference: Chapter 14 of Pena, Tiao and Tsay.

Example. Analysis of the vector MA(1) data set via SCA.

```
input z1,z2. file 'clsma1.dat'
 --
mtsm vma. series z1,z2. model series=c+(i-t1*b)noise.
 --
mestim vma. hold resi(r1,r2)
 SUMMARY FOR THE MULTIVARIATE ARMA MODEL
 SERIES    NAME            MEAN         STD DEV    DIFFERENCE ORDER(S)
   1     Z1              17.0664         2.3334
   2     Z2              25.0393         1.5875
 NUMBER OF OBSERVATIONS =  250  (EFFECTIVE NUMBER = NOBE =  250)
 MODEL SPECIFICATION WITH PARAMETER VALUES

 PARAMETER              PARAMETER                   PARAMETER
  NUMBER               DESCRIPTION                   VALUE
 ----------    -----------------------------     ------------
    1                CONSTANT( 1)                   17.066420
    2                CONSTANT( 2)                   25.039345
    3           MOVING AVERAGE ( 1, 1, 1)           0.100000
    4           MOVING AVERAGE ( 1, 1, 2)           0.100000
    5           MOVING AVERAGE ( 1, 2, 1)           0.100000
    6           MOVING AVERAGE ( 1, 2, 2)           0.100000
 -----------------------
 ERROR COVARIANCE MATRIX
 -----------------------
               1           2
   1     5.095209
   2     1.069979    2.807721
 ITERATIONS TERMINATED DUE TO:
   RELATIVE CHANGE IN DETERMINANT OF COVARIANCE MATRIX .LE. 0.100E-03
   TOTAL NUMBER OF ITERATIONS IS    8
```

```
 FINAL MODEL SUMMARY WITH CONDITIONAL LIKELIHOOD PARAMETER ESTIMATES
 ----- CONSTANT VECTOR (STD ERROR) -----
    17.153  (     0.102  )
    25.108  (     0.076  )
 ----- THETA MATRICES -----
 ESTIMATES OF  THETA(  1 ) MATRIX AND SIGNIFICANCE
      .192     .410      + +
     -.574    1.125      - +
 STANDARD ERRORS
      .039     .093
      .029     .066
 -----------------------
 ERROR COVARIANCE MATRIX
 -----------------------
               1           2
   1     4.877274
   2     1.108076    1.083260
 --
miden r1,r2. maxl 12.

 TIME PERIOD ANALYZED . . . . . . . . . . . . .  1  TO   250
 EFFECTIVE NUMBER OF OBSERVATIONS (NOBE). . .          250
 SERIES   NAME           MEAN       STD. ERROR
    1    R1            -0.0993      2.2062
    2    R2             0.0063      1.0408
 SAMPLE CORRELATION MATRIX OF THE SERIES
   1.00
   0.48  1.00
 SUMMARIES OF CROSS CORRELATION MATRICES USING +,-,.
          1              2
  1   ...-........   ............

  2   .......-....   ............

 CROSS CORRELATION MATRICES IN TERMS OF +,-,.
 LAGS  1 THROUGH  6
      . .      . .      . .      - .      . .      . .
      . .      . .      . .      . .      . .      . .
 LAGS  7 THROUGH 12
      . .      . .      . .      . .      . .      . .
      . .      - .      . .      . .      . .      . .
```

```
 --
mestim vma. hold resi(r1,r2). method exact.

 SUMMARY FOR THE MULTIVARIATE ARMA MODEL
 SERIES    NAME           MEAN        STD DEV    DIFFERENCE ORDER(S)
   1     Z1             17.0664        2.3334
   2     Z2             25.0393        1.5875

 NUMBER OF OBSERVATIONS =  250  (EFFECTIVE NUMBER = NOBE =  250)
 MODEL SPECIFICATION WITH PARAMETER VALUES

 PARAMETER            PARAMETER                  PARAMETER
  NUMBER             DESCRIPTION                  VALUE
 ----------    -----------------------------     ------------
    1               CONSTANT( 1)                 17.152506
    2               CONSTANT( 2)                 25.107561
    3          MOVING AVERAGE ( 1, 1, 1)          0.192436
    4          MOVING AVERAGE ( 1, 1, 2)          0.409907
    5          MOVING AVERAGE ( 1, 2, 1)         -0.574374
    6          MOVING AVERAGE ( 1, 2, 2)          1.124622
 -----------------------
 ERROR COVARIANCE MATRIX
 -----------------------
               1           2
   1     4.877274
   2     1.108076    1.083260
 ITERATIONS TERMINATED DUE TO:
   CHANGE IN  (-2*LOG LIKELIHOOD)/NOBE  .LE. 0.100E-03
   TOTAL NUMBER OF ITERATIONS IS    4

 FINAL MODEL SUMMARY WITH MAXIMUM LIKELIHOOD PARAMETER ESTIMATES
 ----- CONSTANT VECTOR (STD ERROR) -----
    17.083  (    0.104  )
    25.052  (    0.078  )
 ----- THETA MATRICES -----
 ESTIMATES OF  THETA(  1 ) MATRIX AND SIGNIFICANCE
      .173     .447      + +
     -.604    1.208      - +
 STANDARD ERRORS
      .038     .095
      .028     .067
 -----------------------
```

```
 ERROR COVARIANCE MATRIX
 -----------------------
               1           2
   1     4.887855
   2     1.125056    1.014653

 -2*(LOG LIKELIHOOD AT FINAL ESTIMATES)  IS   0.82972019E+03
 --
miden r1,r2. maxl 12.

 TIME PERIOD ANALYZED . . . . . . . . . . . . .  1  TO   250
 EFFECTIVE NUMBER OF OBSERVATIONS (NOBE). . .           250
 SERIES   NAME           MEAN       STD. ERROR
    1    R1             -0.0137      2.2055
    2    R2              0.0002      1.0072
 SAMPLE CORRELATION MATRIX OF THE SERIES
   1.00
   0.51  1.00
 CROSS CORRELATION MATRICES IN TERMS OF +,-,.
 LAGS  1 THROUGH  6
      . .      . .      . .      . .      . .      . .
      . .      . .      . .      . .      . .      . .
 LAGS  7 THROUGH 12
      . .      . .      . .      . .      . .      . .
      . .      - .      . .      . .      . .      . .
 --
mfore vma. nofs 4.
 ----------------------------------------
   4 FORECASTS, BEGINNING AT ORIGIN = 250
 ----------------------------------------
 SERIES:       Z1                  Z2
 TIME    FORECAST  STD ERR   FORECAST  STD ERR
  251      18.019    2.211     27.965    1.007
  252      17.083    2.326     25.052    1.624  <== Means after 1-step.
  253      17.083    2.326     25.052    1.624
  254      17.083    2.326     25.052    1.624
```

For the data set, the fitted model of exact MLE is

$$\begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \begin{bmatrix} 17.08 \\ 25.05 \end{bmatrix} + \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} - \begin{bmatrix} 0.17 & 0.45 \\ -0.60 & 1.21 \end{bmatrix} \begin{bmatrix} a_{1,t-1} \\ a_{2,t-1} \end{bmatrix},$$

where the covariance matrix of the residuals is

$$\hat{\boldsymbol{\Sigma}} = \begin{bmatrix} 4.89 & 1.13 \\ 1.13 & 1.01 \end{bmatrix}.$$

4 Vector ARMA Models

A k-dimensional time series $\boldsymbol{z}_t$ follows a vector ARMA, VARMA(p, q), model if

$$\boldsymbol{\phi}(B)\boldsymbol{z}_t = \boldsymbol{c} + \boldsymbol{\theta}(B)\boldsymbol{a}_t \tag{1}$$

where $\boldsymbol{c}$ is a constant vector, $\boldsymbol{\phi}(B) = \boldsymbol{I} - \sum_{i=1}^{p} \boldsymbol{\phi}_i B^i$ and $\boldsymbol{\theta}(B) = \boldsymbol{I} - \sum_{i=1}^{q} \boldsymbol{\theta}_i B^i$ are two matrix polynomials, and $\boldsymbol{a}_t$ is a sequence of independent and identically distributed random vectors with mean zero and positive definite covariance matrix $\boldsymbol{\Sigma}$. In Eq. (1), we require two additional conditions:

1. $\boldsymbol{\phi}(B)$ and $\boldsymbol{\theta}(B)$ are left coprime, i.e. if $\boldsymbol{u}(B)$ is a left common factor of $\boldsymbol{\phi}(B)$ and $\boldsymbol{\theta}(B)$, then $|\boldsymbol{u}(B)|$ is a non-zero constant. Such a polynomial matrix is called a *unimodular* matrix. In theory, $\boldsymbol{u}(B)$ is unimodular if and only if $\boldsymbol{u}^{-1}(B)$ exists and is a matrix polynomial.

2. The MA order q is as small as possible and the AR order p is as small as possible for that q, and the matrices $\boldsymbol{\phi}_p$ and $\boldsymbol{\theta}_q$ satisfy the condition that Rank$[\boldsymbol{\phi}_p, \boldsymbol{\theta}_q] = \dim(\boldsymbol{z}_t)$.

These two conditions are sufficient conditions for VARMA models to be identifiable. In the literature, these conditions are referred to as *block identifiability*. Condition (2) is discussed in Dunsmuir and Hannan (1976, Advances in App. Prob.). It can be refined by considering column degrees of $\boldsymbol{\phi}(B)$ and $\boldsymbol{\theta}(B)$ instead of the overall degrees p and q; see Hannan and Deistler (1988, Sec. 2.7).

4.1 Identifiability

Unlike the VAR or VMA models, VARMA models encounter the problem of identifiability. For a given linear vector process,

$$\boldsymbol{z}_t = \boldsymbol{\mu} + \sum_{i=0}^{\infty} \boldsymbol{\psi}_i \boldsymbol{a}_{t-i}, \tag{2}$$

where $\boldsymbol{\psi}_0 = \boldsymbol{I}$, and $\{\boldsymbol{a}_t\}$ is an iid sequence of random vectors with mean zero and positive-definite covariance matrix $\boldsymbol{\Sigma}$. A VARMA model is said to be identifiable if the matrix polynomials $\boldsymbol{\phi}(B)$ and $\boldsymbol{\theta}(B)$ are uniquely determined by the $\boldsymbol{\psi}_i$ in Eq. (2). There are cases for which multiple pairs of AR and MA matrix polynomials give rise to the same $\boldsymbol{\psi}_i$s. We use simple bivariate models to discuss the issue.

Example. Consider the VMA(1) model

$$\begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} - \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a_{1,t-1} \\ a_{2,t-1} \end{bmatrix}. \tag{3}$$

This is a well-defined VMA(1) model. However, it can also be written as the VAR(1) model

$$\begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} - \begin{bmatrix} 0 & -2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_{1,t-1} \\ z_{2,t-1} \end{bmatrix} = \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}. \tag{4}$$

To see this, the VMA(1) model in Eq. (3) implies that

$$z_{1t} = a_{1t} - 2a_{2,t-1} \quad \text{and} \quad z_{2t} = a_{2t}.$$

In other words, z_{2t} is a white noise series. As such, we have

$$z_{1t} = a_{1t} - 2a_{2,t-1} = a_{1t} - 2z_{2,t-1}.$$

Consequently, we have

$$z_{1t} + 2z_{2,t-1} = a_{1t} \quad \text{and} \quad z_{2t} = a_{2t},$$

which is precisely the VAR(1) model in Eq. (4). This type of non-uniqueness in model specification is harmless because either model can be used in a real application.

Example. Consider the VARMA(1,1) model

$$\begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} - \begin{bmatrix} 0.8 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_{1,t-1} \\ z_{2,t-1} \end{bmatrix} = \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} - \begin{bmatrix} 0.3 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a_{1,t-1} \\ a_{2,t-1} \end{bmatrix}. \tag{5}$$

It is easy to see that the model is identical to

$$\begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} - \begin{bmatrix} 0.8 & 2+\omega \\ 0 & \beta \end{bmatrix} \begin{bmatrix} z_{1,t-1} \\ z_{2,t-1} \end{bmatrix} = \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} - \begin{bmatrix} 0.3 & \omega \\ 0 & \beta \end{bmatrix} \begin{bmatrix} a_{1,t-1} \\ a_{2,t-1} \end{bmatrix}, \tag{6}$$

for any $\omega \neq 0$ and $\beta \neq 0$. From Eq. (5), we have

$$z_{1t} = 0.8z_{1,t-1} + 2z_{2,t-1} + a_{1t} - .3a_{1,t-1} \tag{7}$$

$$z_{2t} = a_{2t}. \tag{8}$$

Thus, z_{2t} is a white noise series. Multiplying Eq. (8) by ω for $t-1$ and adding Eq. (7), we obtain the first equation of Eq. (6). The second equation in Eq. (6) holds because $z_{2,t-1} = a_{2,t-1}$. This type of identifiability is serious because, without proper constriants, the likelihood function of the VARMA(1,1) model is not uniquely defined. In other words, without proper constraints, one cannot estimate the VARMA(1,1) model.
For the VARMA(1,1) model in Eq. (5), we have $\text{Rank}[\boldsymbol{\phi}_1, \boldsymbol{\theta}_1] = 1$, which is smaller than the dimension of $\boldsymbol{z}_t$. This is a clear violation of the condition of Eq. (1). Also, the two

polynomial matrices of Eq. (6) are not left-coprime. For instance, $\boldsymbol{u}(B) = \text{diag}\{1, 1-\beta B\}$ is a left-common factor. Finally, the identifiability problem can occur even if none of the components of $\boldsymbol{z}_t$ is white noise.
Identifiability is an important issue of VARMA models. It implies that model specification of VARMA models involves more than identifying the order (p, q). Indeed, model specification in VARMA models must include structural specification to overcome the problem of identifiability. In the literature, two approaches are available to perform structural specification of VARMA models. The first approach uses state-space formulation with Kronecker indices, and the second approach uses scalar component models. See Tsay (1991, Statistica Sinica). We shall discuss these two approaches later.

4.2 Some basic properties of VARMA models

In this section, we discuss some properties of a VARMA(p, q) model. The model is assumed to be identifiable and the innovation $\boldsymbol{a}_t$ has mean zero and covariance matrix $\boldsymbol{\Sigma}$, which is positive definite.

1. Stationarity condition

   Similar to a VAR(p) model, the necessary and sufficient condition for weak stationarity the $\boldsymbol{z}_t$ model in Eq. (1) is that all zeros of the polynomial $|\boldsymbol{\phi}(B)|$ are outside the unit circle, i.e., they are greater than 1 in modulus.

   For a stationary VARMA(p, q) model, we can rewrite the model as

$$\boldsymbol{\phi}(B)(\boldsymbol{z}_t - \boldsymbol{\mu}) = \boldsymbol{\theta}(B)\boldsymbol{a}_t, \tag{9}$$

   where $\boldsymbol{\mu}$ satisfies $\boldsymbol{\phi}(1)\boldsymbol{\mu} = \boldsymbol{\phi}_0$ and is the mean vector of $\boldsymbol{z}_t$. Let $\breve{\boldsymbol{z}}_t = \boldsymbol{z}_t - \boldsymbol{\mu}$ be the mean-corrected series. The VARMA(p, q) model then becomes

$$\boldsymbol{\phi}(B)\breve{\boldsymbol{z}}_t = \boldsymbol{\theta}(B)\boldsymbol{a}_t. \tag{10}$$

   For ease in presetation, we often use Eq. (10) to derive properties of the VARMA(p, q) series.

   When $\boldsymbol{z}_t$ is stationary, it has an MA representation as

$$\breve{\boldsymbol{z}}_t = \boldsymbol{a}_t + \boldsymbol{\psi}_1\boldsymbol{a}_{t-1} + \boldsymbol{\psi}_2\boldsymbol{a}_{t-2} + \ldots = \boldsymbol{\psi}(B)\boldsymbol{a}_t, \tag{11}$$

   where $\boldsymbol{\psi}(B) = \sum_{i=0}^{\infty}\boldsymbol{\psi}_i B^i$ with $\boldsymbol{\psi}_0 = \boldsymbol{I}_k$, the $k \times k$ identity matrix. The coefficient matrices $\boldsymbol{\psi}_i$ can obtain recursively by equating the coefficients of B in

$$\boldsymbol{\phi}(B)\boldsymbol{\psi}(B) = \boldsymbol{\theta}(B).$$

   In particular, we have

$$\boldsymbol{\psi}_i = \sum_{j=1}^{p}\boldsymbol{\phi}_j\boldsymbol{\psi}_{j-i}, \quad i > q.$$

2. Invertibility condition

   A sufficient and necessary condition for $\boldsymbol{z}_t$ in Eq. (1) to be invertible is that all zeros of $|\boldsymbol{\theta}(B)|$ are outside the unit circle. If $\boldsymbol{z}_t$ is invertible, then it has an AR representation, i.e.,

$$\boldsymbol{\pi}(B)(\boldsymbol{z}_t - \boldsymbol{\mu}) = \boldsymbol{a}_t, \tag{12}$$

   where $\boldsymbol{\pi}(B) = \boldsymbol{I} - \boldsymbol{\pi}_1 B - \boldsymbol{\pi}_2 B^2 - \ldots$ and the coefficient matrices $\boldsymbol{\pi}_i$ can be obtained by equating the coefficients of B^i in

$$\boldsymbol{\theta}(B)\boldsymbol{\pi}(B) = \boldsymbol{\phi}(B).$$

3. Moment equations

   For a stationary VARMA(p, q) series $\boldsymbol{z}_t$, one can use the MA representation to obtain

$$E(\boldsymbol{a}_t \boldsymbol{z}'_{t-\ell}) = \begin{cases} \boldsymbol{0} & \text{if } \ell > 0, \\ \boldsymbol{\Sigma} & \text{if } \ell = 0, \\ \boldsymbol{\Sigma}\boldsymbol{\psi}'_\ell & \text{if } \ell < 0. \end{cases} \tag{13}$$

   Posting multiplying the model in Eq. (10) by $\breve{\boldsymbol{z}}'_{t-\ell}$, taking expectation, and using the result in Eq. (13), we obtain the moment equations for $\boldsymbol{z}_t$ as

$$\boldsymbol{\Gamma}_\ell - \sum_{i=1}^{p} \boldsymbol{\phi}_i \boldsymbol{\Gamma}_{\ell-i} = \begin{cases} -\sum_{j=\ell}^{q} \boldsymbol{\theta}_j \boldsymbol{\Sigma} \boldsymbol{\psi}'_{j-\ell} & \text{if } \ell = 0, 1, \ldots, q, \\ \boldsymbol{0} & \text{if } \ell > q, \end{cases} \tag{14}$$

   where, for convenience, $\boldsymbol{\theta}_0 = -\boldsymbol{I}$. In particular,

$$\boldsymbol{\Gamma}_\ell = \boldsymbol{\phi}_1 \boldsymbol{\Gamma}_{\ell-1} + \cdots + \boldsymbol{\phi}_p \boldsymbol{\Gamma}_{\ell-p}, \quad \ell = q+1, \ldots, q+p,$$

   which is referred to as the multivariate generalized Yule-Walker equations.

4. Relationship to transfer function models

   For a VARMA model, the block structure of zeros is only a sufficient condition for the model to become a transfer function model. Details are given below.

   Unidirectional relationship:
   To understand the relationship between transfer function models and VARMA models, we consider the 2-dimensional case. Here a VARMA model can be written as

$$\begin{bmatrix} \phi_{11}(B) & \phi_{12}(B) \\ \phi_{21}(B) & \phi_{22}(B) \end{bmatrix} \begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \begin{bmatrix} \theta_{11}(B) & \theta_{12}(B) \\ \theta_{21}(B) & \theta_{22}(B) \end{bmatrix} \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix},$$

   where $\phi_{ij}(B)$ are polynomials. Pre-multyplying the prior equation by

$$\begin{bmatrix} \phi_{22}(B) & -\phi_{12}(B) \\ -\phi_{21}(B) & \phi_{11}(B) \end{bmatrix},$$

and letting $d(B) = \phi_{11}(B)\phi_{22}(B) - \phi_{12}(B)\phi_{21}(B)$, we have

$$d(B)\begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \begin{bmatrix} \phi_{22}(B)\theta_{11}(B) - \phi_{12}(B)\theta_{21}(B) & \phi_{22}(B)\theta_{12}(B) - \phi_{12}(B)\theta_{22}(B) \\ \phi_{11}(B)\theta_{21}(B) - \phi_{21}(B)\theta_{11}(B) & \phi_{11}(B)\theta_{22}(B) - \phi_{21}(B)\theta_{12}(B) \end{bmatrix}\begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}. \tag{15}$$

Consequently, it is easily seen that $\phi_{22}(B)\theta_{12}(B) - \phi_{12}(B)\theta_{22}(B) = 0$ if and only if z_{1t} does not depend on the past of z_{2t}. In addition, if $\phi_{11}(B)\theta_{21}(B) - \phi_{21}(B)\theta_{11}(B) \neq 0$, then z_{1t} is an input variable and z_{2t} is an output variable. Note that $\phi_{12}(B) = \theta_{12}(B) = 0$, which results in having simultaneous lower triangular matrices in AR and MA components, is only a sufficient condition that z_{1t} does not depend on the past of z_{2t}. In summary, for the bivariate VARMA model $\boldsymbol{z}_t$, the necessary and sufficient conditions that z_{1t} is an input variable and z_{2t} is an ouput variable are

$$\begin{vmatrix} \phi_{12}(B) & \theta_{12}(B) \\ \phi_{22}(B) & \theta_{22}(B) \end{vmatrix} = 0, \quad \begin{vmatrix} \phi_{11}(B) & \theta_{11}(B) \\ \phi_{21}(B) & \theta_{21}(B) \end{vmatrix} \neq 0.$$

5. Unvariate component models

   For a stationary VARMA(p, q) series $\boldsymbol{z}_t$, the maximum order of each component z_{it} is an ARMA$(pk, p(k-1)+q)$. This result can easily be derived using the co-factors of a matrix.

4.3 Model specification

Before introducing structural specification, we mention one procedure that can help identify a VARMA model in practice. It is to fit successive VAR models and examine the residual cross-correlation matrices after each VAR fit. In theory, such a procedure can lead to overidentification. Nevertheness, the procedure is relatively simple.
Estimation of VARMA models can be performed using either the conditional or exact likelihood method.

4.4 Exact likelihood function of a VARMA model

The exact likelihood function of a stationary VARMA(p, q) model in Eq. (1) has been derived by several authors, e.g., Hillmer and Tiao (1979) and Nicholls and Hall (1979). In this section, we follow the approach of Reinsel (1993, Section 5.3) to derive the likelihood function. For simplicity, we assume $E(\boldsymbol{z}_t) = \boldsymbol{0}$ or equivalently, we use the model in Eq. (10). Suppose that the data set is $\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n\}$.
Let $\boldsymbol{Z} = (\boldsymbol{z}_1', \boldsymbol{z}_2', \ldots, \boldsymbol{z}_n')'$ be the $kn \times 1$ vector of the data, and $\boldsymbol{A} = (\boldsymbol{a}_1', \boldsymbol{a}_2', \ldots, \boldsymbol{a}_n')'$. Also, let $\boldsymbol{Z}_0 = (\boldsymbol{z}_{1-p}', \boldsymbol{z}_{2-p}', \ldots, \boldsymbol{z}_0')'$ be the $kp \times 1$ vector of presample data and $\boldsymbol{A}_0 = (\boldsymbol{a}_{1-q}', \boldsymbol{a}_{2-q}', \ldots, \boldsymbol{a}_0')'$ be the $kq \times 1$ vector of presample innovations. Finally, let $\boldsymbol{U}_0 = (\boldsymbol{Z}_0', \boldsymbol{A}_0')'$ be the vector of presample variables. Using these newly defined variables, the sample of the model can be written as

$$\boldsymbol{\Phi Z} = \boldsymbol{\Theta A} + \boldsymbol{P U}_0, \tag{16}$$

where $\boldsymbol{\Phi} = (\boldsymbol{I}_n \otimes \boldsymbol{I}_k) - \sum_{i=1}^{p} (\boldsymbol{L}^i \otimes \boldsymbol{\phi}_i)$, $\boldsymbol{\Theta} = (\boldsymbol{I}_n \otimes \boldsymbol{I}_k) - \sum_{j=1}^{q} (\boldsymbol{L}^j \otimes \boldsymbol{\theta}_j)$, and $\boldsymbol{L}$ is the $n \times n$ lag matrix that has 1 on its main subdiagonal elements and zeros elsewhere. More specifically, $\boldsymbol{L} = [L_{ij}]$ such that $L_{i,i-1} = 1$ for $i = 2, \ldots, n$ and $= 0$ otherwise. The matrix $\boldsymbol{P}$ in Eq. (16) is given by

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{G}_p & \boldsymbol{H}_q \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix},$$

where

$$\boldsymbol{G}_p = \begin{bmatrix} \boldsymbol{\phi}_p & \boldsymbol{\phi}_{p-1} & \cdots & \boldsymbol{\phi}_1 \\ \boldsymbol{0} & \boldsymbol{\phi}_p & \cdots & \boldsymbol{\phi}_2 \\ \vdots & \vdots & & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{\phi}_p \end{bmatrix}, \quad \boldsymbol{H}_q = - \begin{bmatrix} \boldsymbol{\theta}_q & \boldsymbol{\theta}_{q-1} & \cdots & \boldsymbol{\theta}_1 \\ \boldsymbol{0} & \boldsymbol{\theta}_q & \cdots & \boldsymbol{\theta}_2 \\ \vdots & \vdots & & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{\theta}_q \end{bmatrix}.$$

Since $\boldsymbol{A}$ is independent of $\boldsymbol{U}_0$ and $\text{Cov}(\boldsymbol{A}) = \boldsymbol{I}_n \otimes \boldsymbol{\Sigma}$, the covariance matrix of $\boldsymbol{Z}$ in Eq. (16) is

$$\boldsymbol{\Gamma}_Z = \boldsymbol{\Phi}^{-1}[\boldsymbol{\Theta}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma})\boldsymbol{\Theta}' + \boldsymbol{P}\boldsymbol{\Xi}\boldsymbol{P}'](\boldsymbol{\Phi}')^{-1}, \tag{17}$$

where $\boldsymbol{\Xi} = \text{Cov}(\boldsymbol{U}_0)$. For the stationary VARMA(p, q) model $\boldsymbol{z}_t$ in Eq. (1), the covariance matrix $\boldsymbol{\Xi}$ can be written as

$$\boldsymbol{\Xi} = \begin{bmatrix} \boldsymbol{\Gamma}(p) & \boldsymbol{C}' \\ \boldsymbol{C} & \boldsymbol{I}_q \otimes \boldsymbol{\Sigma} \end{bmatrix},$$

where $\boldsymbol{\Gamma}(p)$ is the $kp \times kp$ covariance matrix of $\boldsymbol{Z}_0$ such that the (i, j)th bolck of which is $\boldsymbol{\Gamma}_{i-j}$, and $\boldsymbol{C} = \text{Cov}(\boldsymbol{Z}_0, \boldsymbol{A}_0)$ that can be obtained by using Eq. (13) as $\boldsymbol{C} = (\boldsymbol{I}_q \otimes \boldsymbol{\Sigma})\boldsymbol{\Psi}$ with

$$\boldsymbol{\Psi} = \begin{bmatrix} \boldsymbol{\psi}'_{q-p} & \boldsymbol{\psi}'_{q+1-p} & \cdots & \boldsymbol{\psi}'_{q-1} \\ \boldsymbol{\psi}'_{q-1-p} & \boldsymbol{\psi}'_{q-p} & \cdots & \boldsymbol{\psi}'_{q-2} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{\psi}'_{1-p} & \boldsymbol{\psi}'_{2-p} & \cdots & \boldsymbol{I}_k \end{bmatrix},$$

being a $kq \times kp$ matrix of the ψ-weight matrices of $\boldsymbol{z}_t$, where it is understood that $\boldsymbol{\psi}_j = \boldsymbol{0}$ if $j < 0$.
Using the matrix inversion formula

$$[\boldsymbol{W} + \boldsymbol{B}\boldsymbol{D}\boldsymbol{B}']^{-1} = \boldsymbol{W}^{-1} - \boldsymbol{W}^{-1}\boldsymbol{B}(\boldsymbol{B}'\boldsymbol{W}^{-1}\boldsymbol{B} + \boldsymbol{D}^{-1})^{-1}\boldsymbol{B}'\boldsymbol{W}^{-1},$$

we obtain the inverse of $\boldsymbol{\Gamma}_Z$ as

$$\boldsymbol{\Gamma}_Z^{-1} = \boldsymbol{\Phi}'(\boldsymbol{\Omega}^{-1} - \boldsymbol{\Omega}^{-1}\boldsymbol{P}\boldsymbol{Q}^{-1}\boldsymbol{P}'\boldsymbol{\Omega}^{-1})\boldsymbol{\Phi}, \tag{18}$$

where $\boldsymbol{\Omega} = \boldsymbol{\Theta}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma})\boldsymbol{\Theta}'$ and $\boldsymbol{Q} = \boldsymbol{P}'\boldsymbol{\Omega}^{-1}\boldsymbol{P} + \boldsymbol{\Xi}^{-1}$. Note that we also have $\boldsymbol{\Omega}^{-1} = (\boldsymbol{\Theta}')^{-1}(\boldsymbol{I}_n \otimes \boldsymbol{\Sigma}^{-1})\boldsymbol{\Theta}^{-1}$. Next, using the following matrix inversion formula

$$\begin{bmatrix} \boldsymbol{W} & \boldsymbol{B} \\ \boldsymbol{B}' & \boldsymbol{D} \end{bmatrix}^{-1} = \begin{bmatrix} \boldsymbol{W}^{-1} + \boldsymbol{F}\boldsymbol{E}^{-1}\boldsymbol{F}' & -\boldsymbol{F}\boldsymbol{E}^{-1} \\ -\boldsymbol{E}^{-1}\boldsymbol{F} & \boldsymbol{E}^{-1} \end{bmatrix},$$

where $\boldsymbol{E} = \boldsymbol{D} - \boldsymbol{B}'\boldsymbol{W}^{-1}\boldsymbol{B}$ and $\boldsymbol{F} = \boldsymbol{W}^{-1}\boldsymbol{B}$, we have

$$\boldsymbol{\Xi}^{-1} = \left[\begin{array}{cc} \boldsymbol{\Delta}^{-1} & -\boldsymbol{\Delta}^{-1}\boldsymbol{\Psi}' \\ -\boldsymbol{\Psi}\boldsymbol{\Delta}^{-1} & (\boldsymbol{I}_q \otimes \boldsymbol{\Sigma}^{-1}) + \boldsymbol{\Psi}\boldsymbol{\Delta}^{-1}\boldsymbol{\Psi}' \end{array} \right],$$

where $\boldsymbol{\Delta} = \boldsymbol{\Gamma}(p) - \boldsymbol{\Psi}'(\boldsymbol{I}_q \otimes \boldsymbol{\Sigma})\boldsymbol{\Psi}$.
We are ready to write down the likelihood function of the sample $\boldsymbol{Z}$. Using the property of determinant

$$\left| \begin{array}{cc} \boldsymbol{W} & \boldsymbol{J} \\ \boldsymbol{B} & \boldsymbol{D} \end{array} \right| = |\boldsymbol{W}||\boldsymbol{D} - \boldsymbol{B}\boldsymbol{W}^{-1}\boldsymbol{J}|,$$

and the definition of $\boldsymbol{\Xi}$, we obain

$$|\boldsymbol{\Xi}| = |(\boldsymbol{I}_q \otimes \boldsymbol{\Sigma})||\boldsymbol{\Gamma}(p) - \boldsymbol{\Psi}'(\boldsymbol{I}_q \otimes \boldsymbol{\Sigma})\boldsymbol{\Psi}| = |(\boldsymbol{I}_q \otimes \boldsymbol{\Sigma})||\boldsymbol{\Delta}| = |\boldsymbol{\Sigma}|^q|\boldsymbol{\Delta}|.$$

For a given order (p,q), let $\boldsymbol{\vartheta}$ be the set of all parameters in a VARMA(p,q) model. Let $\boldsymbol{X} = (\boldsymbol{x}_1', \ldots, \boldsymbol{x}_n')' = \boldsymbol{\Phi}\boldsymbol{Z}$. That is, $\boldsymbol{x}_t = \boldsymbol{z}_t - \sum_{i=1}^{t-1} \boldsymbol{\phi}_i \boldsymbol{z}_{t-i}$ for $t = 1, \ldots, p$ and $\boldsymbol{x}_t = \boldsymbol{z}_t - \sum_{i=1}^{p} \boldsymbol{\phi}_i \boldsymbol{z}_{t-i}$, where it is understood that the summation term disapears if the lower limit exceeds the upper limit. From Eq. (17), the determinant of $\boldsymbol{\Gamma}_Z$ is

$$|\boldsymbol{\Gamma}_Z| = |\boldsymbol{\Phi}|^{-2}|\boldsymbol{\Omega} + \boldsymbol{P}\boldsymbol{\Xi}\boldsymbol{P}'|$$

(To continue!)

4.5 Examples

We discuss two examples to better understand VARMA models. The first example is the gas-furnace data analyzed before and the second example is an old data set in Coen, Gomme and Kendall (1969).

Gas-Furnace Example revisited:
Ignoring that the data set belongs to transfer function models, we treat it as a bivariate time series that allows for possible feedback relationship between component series.
SCA demonstration

```
input x,y. file 'gasfur.dat'
 --
miden x,y. no ccm. maxl 12. arfits 1 to 8. rccm 1 to 8. @
output level(deta)

 TIME PERIOD ANALYZED . . . . . . . . . . . . .  1  TO   296
 EFFECTIVE NUMBER OF OBSERVATIONS (NOBE). . .            296
 SERIES   NAME           MEAN       STD. ERROR
    1    X              -0.0568      1.0710
    2    Y              53.5091      3.1967
```

```
DETERMINANT OF S(0) = 0.943952E+01

NOTE: S(0) IS THE SAMPLE COVARIANCE MATRIX OF W(MAXLAG+1),...,W(NOBE)

AUTOREGRESSIVE FITTING ON LAG(S)  1
=== PHI( 1) ===
     .995     .030      + +
    -.495     .894      - +
STANDARD ERRORS
     .020     .007
     .036     .012
RESIDUAL COVARIANCE MATRIX S( 1)
 0.102E+00
 0.894E-01 0.343E+00
RESIDUAL CORRELATION MATRIX RS( 1)
  1.00
  0.48  1.00
EIGENVALUES AND EIGENVECTORS OF S( 1)
EIGENVALUES
     0.072     0.373
EIGENVECTORS
     0.950     0.313
    -0.313     0.950

DETERMINANT OF S(J) = 0.269959E-01
LEADING TO A VALUE OF THE TEST STATISTIC M = -W*LN(U) = 1666.31
APPROXIMATELY DISTRIBUTED AS A CHI SQUARE WITH  4 DF,
WHERE U = DET(S(J))/DET(S(J-1))
   S(J) = RESIDUAL COVARIANCE MATRIX AFTER JTH FIT
      W = (NOBE-MAXARF-1)-J*K-.5, AND DF = K*K.

SAMPLE CROSS CORRELATION MATRICES FOR THE RESIDUAL SERIES.
LAG = 1
  0.72  0.20
  0.66  0.74
LAG = 2
  0.30 -0.04
  0.53  0.30
LAG = 3
 -0.09 -0.19
  0.14 -0.14
LAG = 4
```

```
 -0.34 -0.24
 -0.31 -0.40
LAG = 5
 -0.38 -0.24
 -0.62 -0.44
CROSS CORRELATION MATRICES IN TERMS OF +,-,.
LAGS  1 THROUGH  6
     + +      + .      . -      - -      - -      - -
     + +      + +      + -      - -      - -      - -
LAGS  7 THROUGH 12
     - -      . .      . .      + +      + +      + +
     - -      - .      . .      + .      + +      + +

AUTOREGRESSIVE FITTING ON LAG(S)  1  2
=== PHI( 1) ===
    1.813    -.076      + -
     .235    1.446      + +
STANDARD ERRORS
     .040     .022
     .055     .030
=== PHI( 2) ===
    -.961     .046      - +
    -.642    -.580      - -
STANDARD ERRORS
     .048     .019
     .065     .027
RESIDUAL COVARIANCE MATRIX S( 2)
 0.367E-01
-0.381E-02 0.683E-01

DETERMINANT OF S(J) = 0.249342E-02
LEADING TO A VALUE OF THE TEST STATISTIC M = -W*LN(U) =  672.92

SAMPLE CROSS CORRELATION MATRICES FOR THE RESIDUAL SERIES.
LAG = 1
  0.12  0.02
 -0.11  0.17
LAG = 2
 -0.02  0.10
  0.15  0.12
...
LAG = 5
```

```
  0.04 -0.01
 -0.13  0.01
LAG = 6
  0.14 -0.11
 -0.10  0.18
CROSS CORRELATION MATRICES IN TERMS OF +,-,.
LAGS  1 THROUGH  6
     + .      . .      . .      - .      . .      + .
     . +      + +      . .      . .      - .      . +

AUTOREGRESSIVE FITTING ON LAG(S)  1  2  3
=== PHI( 1) ===
    1.938    -.055      + .
     .106    1.612      . +
STANDARD ERRORS
     .058     .043
     .077     .057
=== PHI( 2) ===
   -1.266     .062      - .
    -.320    -.934      - -
STANDARD ERRORS
     .114     .066
     .152     .087
=== PHI( 3) ===
     .228    -.023      + .
    -.196     .194      . +
STANDARD ERRORS
     .080     .031
     .106     .042
RESIDUAL COVARIANCE MATRIX S( 3)
 0.356E-01
-0.224E-02 0.626E-01
DETERMINANT OF S(J) = 0.222347E-02
LEADING TO A VALUE OF THE TEST STATISTIC M = -W*LN(U) =   32.14

AUTOREGRESSIVE FITTING ON LAG(S)  1  2  3  4
=== PHI( 1) ===
    1.927    -.055      + .
     .085    1.572      . +
STANDARD ERRORS
     .059     .045
     .075     .057
```

```
=== PHI( 2) ===
   -1.195     .098      - .
    -.144    -.652      . -
STANDARD ERRORS
     .128     .083
     .165     .107
=== PHI( 3) ===
     .093    -.069      . .
    -.552    -.182      - .
STANDARD ERRORS
     .137     .077
     .176     .099
=== PHI( 4) ===
     .102     .017      . .
     .283     .163      + +
STANDARD ERRORS
     .082     .033
     .105     .042
RESIDUAL COVARIANCE MATRIX S( 4)
 0.354E-01
-0.306E-02 0.582E-01
DETERMINANT OF S(J) = 0.204899E-02
LEADING TO A VALUE OF THE TEST STATISTIC M = -W*LN(U) =   22.76

AUTOREGRESSIVE FITTING ON LAG(S)  1  2  3  4  5
=== PHI( 1) ===
    1.922    -.065      + .
     .075    1.562      . +
STANDARD ERRORS
     .059     .046
     .075     .059
=== PHI( 2) ===
   -1.193     .117      - .
    -.148    -.615      . -
STANDARD ERRORS
     .128     .085
     .164     .108
=== PHI( 3) ===
     .147    -.053      . .
    -.414    -.201      - .
STANDARD ERRORS
     .147     .088
```

```
     .187     .113
=== PHI( 4) ===
    -.024    -.012      . .
    -.008     .185      . .
STANDARD ERRORS
     .139     .078
     .178     .099
=== PHI( 5) ===
     .094     .011      . .
     .213    -.017      . .
STANDARD ERRORS
     .084     .033
     .107     .043
RESIDUAL COVARIANCE MATRIX S( 5)
 0.352E-01
-0.339E-02 0.573E-01
DETERMINANT OF S(J) = 0.200726E-02
LEADING TO A VALUE OF THE TEST STATISTIC M = -W*LN(U) =    5.69

AUTOREGRESSIVE FITTING ON LAG(S)  1  2  3  4  5  6
=== PHI( 1) ===
    1.931    -.051      + .
     .064    1.545      . +
STANDARD ERRORS
     .058     .046
     .075     .059
=== PHI( 2) ===
   -1.204     .100      - .
    -.135    -.593      . -
STANDARD ERRORS
     .127     .085
     .162     .108
=== PHI( 3) ===
     .167    -.079      . .
    -.438    -.171      - .
STANDARD ERRORS
     .145     .088
     .185     .113
=== PHI( 4) ===
    -.158     .027      . .
     .150     .132      . .
STANDARD ERRORS
```

```
     .146     .088
     .187     .112
=== PHI( 5) ===
     .379    -.042      + .
    -.120     .057      . .
STANDARD ERRORS
     .138     .077
     .176     .099
=== PHI( 6) ===
    -.214     .031      - .
     .250    -.042      + .
STANDARD ERRORS
     .084     .033
     .108     .042
RESIDUAL COVARIANCE MATRIX S( 6)
 0.343E-01
-0.229E-02 0.560E-01
RESIDUAL CORRELATION MATRIX RS( 6)
  1.00
 -0.05  1.00
EIGENVALUES AND EIGENVECTORS OF S( 6)
EIGENVALUES
     0.034     0.056
CROSS CORRELATION MATRICES IN TERMS OF +,-,.
LAGS  1 THROUGH  6
     . .      . .      . .      . .      . .      . .
     . .      . .      . .      . .      . .      . .
LAGS  7 THROUGH 12
     . .      . .      . .      . .      . .      - .
     . .      . .      . .      . .      . .      . .

AUTOREGRESSIVE FITTING ON LAG(S)  1  2  3  4  5  6  7
=== PHI( 1) ===
    1.924    -.047      + .
     .059    1.547      . +
STANDARD ERRORS
     .059     .046
     .075     .059
=== PHI( 2) ===
   -1.188     .095      - .
    -.123    -.594      . -
STANDARD ERRORS
```

```
     .128     .085
     .164     .108
=== PHI( 3) ===
     .156    -.085      . .
    -.446    -.158      - .
STANDARD ERRORS
     .145     .089
     .186     .114
=== PHI( 4) ===
    -.147     .024      . .
     .149     .115      . .
STANDARD ERRORS
     .146     .089
     .187     .114
=== PHI( 5) ===
     .346    -.015      + .
    -.155     .013      . .
STANDARD ERRORS
     .146     .088
     .187     .113
=== PHI( 6) ===
    -.152    -.012      . .
     .349     .034      . .
STANDARD ERRORS
     .139     .078
     .178     .099
=== PHI( 7) ===
    -.047     .022      . .
    -.080    -.034      . .
STANDARD ERRORS
     .087     .033
     .111     .042
RESIDUAL COVARIANCE MATRIX S( 7)
 0.342E-01
-0.227E-02 0.558E-01

DETERMINANT OF S(J) = 0.190324E-02
LEADING TO A VALUE OF THE TEST STATISTIC M = -W*LN(U) =    1.76

AUTOREGRESSIVE FITTING ON LAG(S)  1  2  3  4  5  6  7  8
=== PHI( 1) ===
    1.930    -.047      + .
```

```
      .053    1.554       . +
STANDARD ERRORS
      .059     .046
      .075     .059
=== PHI( 2) ===
   -1.186     .088       - .
    -.119    -.598       . -
STANDARD ERRORS
      .128     .085
      .162     .108
=== PHI( 3) ===
      .136    -.077       . .
    -.440    -.161       - .
STANDARD ERRORS
      .146     .089
      .185     .113
=== PHI( 4) ===
    -.128     .037       . .
      .141     .082       . .
STANDARD ERRORS
      .146     .089
      .186     .113
=== PHI( 5) ===
      .326    -.012       + .
    -.132     .033       . .
STANDARD ERRORS
      .146     .089
      .185     .113
=== PHI( 6) ===
    -.095    -.070       . .
      .337     .168       . .
STANDARD ERRORS
      .147     .088
      .187     .112
=== PHI( 7) ===
    -.150     .114       . .
    -.113    -.256       . -
STANDARD ERRORS
      .140     .077
      .178     .098
=== PHI( 8) ===
      .080    -.046       . .
```

```
     .030     .104      . +
STANDARD ERRORS
     .087     .033
     .111     .042
RESIDUAL COVARIANCE MATRIX S( 8)
 0.339E-01
-0.178E-02 0.546E-01

==========  STEPWISE AUTOREGRESSION SUMMARY  ==========
----------------------------------------------------------------------
    I RESIDUAL I EIGENVAL.I CHI-SQ  I          I SIGNIFICANCE
LAG I VARIANCESI OF SIGMA I  TEST   I   AIC    I OF PARTIAL AR COEFF.
----+----------+----------+---------+----------+----------------------
  1 I .102E+00 I .724E-01 I 1666.31 I   -3.585 I + +
    I .343E+00 I .373E+00 I         I          I - +
----+----------+----------+---------+----------+----------------------
  2 I .367E-01 I .362E-01 I  672.92 I   -5.940 I - +
    I .683E-01 I .688E-01 I         I          I - -
----+----------+----------+---------+----------+----------------------
  3 I .356E-01 I .354E-01 I   32.14 I   -6.028 I + .
    I .626E-01 I .628E-01 I         I          I . +
----+----------+----------+---------+----------+----------------------
  4 I .354E-01 I .350E-01 I   22.76 I   -6.082 I . .
    I .582E-01 I .586E-01 I         I          I + +
----+----------+----------+---------+----------+----------------------
  5 I .352E-01 I .347E-01 I    5.69 I   -6.076 I . .
    I .573E-01 I .578E-01 I         I          I . .
----+----------+----------+---------+----------+----------------------
  6 I .343E-01 I .341E-01 I   12.83 I   -6.096 I - .
    I .560E-01 I .563E-01 I         I          I + .
----+----------+----------+---------+----------+----------------------
  7 I .342E-01 I .340E-01 I    1.76 I   -6.075 I . .
    I .558E-01 I .560E-01 I         I          I . .
----+----------+----------+---------+----------+----------------------
  8 I .339E-01 I .337E-01 I    8.24 I   -6.078 I . .
    I .546E-01 I .548E-01 I         I          I . +
----------------------------------------------------------------------
NOTE:  CHI-SQUARED CRITICAL VALUES WITH   4  DEGREES OF FREEDOM ARE
             5 PERCENT:   9.5     1 PERCENT:  13.3
```

Summary of the analysis:
From the summary table, a VAR(6) model is identified by AIC and the Chi-squared test at the 5% significance level. The estimated AR coefficient matrices are given in Table 1. Form

Table 1: VAR(6) Fit via Least Squares Method for the Gas-Furnace Data

ϕ_1		ϕ_2		ϕ_3		ϕ_4		ϕ_5		ϕ_6	
1.93	-.05	-1.20	.10	.17	-.08	-.16	.03	.38	-.04	-.21	.03
.06	1.55	-.14	-.59	-.44	-.17	.15	.13	-.12	.06	.23	-.04
Standard error											
.06	.05	.13	.09	.15	.09	.15	.09	.14	.08	.08	.03
.08	.06	.16	.11	.19	.11	.19	.11	.18	.10	.11	.04
Simplified notation											
+	.	−	.	.	.	.	.	+	.	−	.
.	+	.	−	−	.	.	.	.	.	+	.

the table, we make the following observations using the 5% significance level:

1. $\phi_{i,12}$ are insignificant for all $i \in \{1, \cdots, 6\}$, indicating that x_t (gas rate) does not depend on the past values of y_t (CO_2). That is, the transfer function structure is revealed.

2. $\phi_{1,21}$ and $\phi_{2,21}$ are insignificant, but $\phi_{3,21} \neq 0$. In addition, from the output, the correlation of the residuals a_{1t} and a_{2t} is very small, i.e. -0.05, suggesting that the residuals are essentially uncorrelated. These two facts indicate that the delay is $b = 3$.

Finally, we can further simplify the model via sequentially remove insignificant parameter estimates. This process is tedious, but there seems to be no simple alternative. For the gas-furnace data, I demonstrate how to remove parameters in SCA. See the output below. Note that the $\phi_{3,22}$ estimate can be removed too.

SCA demonstration:

```
mtsm m1. series x,y. model (1-p1*b-p2*b**2-p3*b**3-p4*b**4-@
p5*b**5-p6*b**6)series=c1+noise.
```

```
 SUMMARY FOR MULTIVARIATE ARMA MODEL --    M1
 VARIABLE   DIFFERENCING
  X
  Y
  PARAMETER     FACTOR     ORDER    CONSTRAINT
  1        C1   CONSTANT      0         CC1
  2        P1    REG AR       1         CP1
  3        P2    REG AR       2         CP2
  4        P3    REG AR       3         CP3
  5        P4    REG AR       4         CP4
  6        P5    REG AR       5         CP5
  7        P6    REG AR       6         CP6
```

```
 --
input cp1. nrow 2. ncol 2.     <== One way to put constraint
 --
cp2=cp1
 --
input p1. nrow 2. ncol 2.  <== Put initial values
 --
p2=p1
 --
cp3=cp2
 --
cp3(2,1)=0   <= Another way to put constraint
 --
p3=p1
 --
p3(2,1)=.1  <= Another way to put initial value
 --
p4=p3
 --
cp4=cp3
 --
p5=p4
 --
cp5=cp4
 --
p6=p5
 --
cp6=cp5
 --
mestim m1. hold resi(r1,r2)
 SUMMARY FOR THE MULTIVARIATE ARMA MODEL
 SERIES    NAME            MEAN         STD DEV    DIFFERENCE ORDER(S)
   1     X               -0.0568         1.0710
   2     Y               53.5091         3.1967
 NUMBER OF OBSERVATIONS =  296  (EFFECTIVE NUMBER = NOBE =  290)
 MODEL SPECIFICATION WITH PARAMETER VALUES

 PARAMETER              PARAMETER                 PARAMETER
  NUMBER               DESCRIPTION                 VALUE
 ----------    -----------------------------     ------------
    1                CONSTANT( 1)                -0.022734
    2                CONSTANT( 2)                21.426382
```

```
   3          AUTOREGRESSIVE ( 1, 1, 1)          0.100000
 *FIXED*      AUTOREGRESSIVE ( 1, 1, 2)          0.000000
   4          AUTOREGRESSIVE ( 1, 2, 2)          0.100000
  .....
  18          AUTOREGRESSIVE ( 6, 2, 2)          0.100000
-----------------------
ERROR COVARIANCE MATRIX
-----------------------
               1           2
  1      .638908
  2     -.012481    6.161229

FINAL MODEL SUMMARY WITH CONDITIONAL LIKELIHOOD PARAMETER ESTIMATES
----- CONSTANT VECTOR (STD ERROR) -----
   -0.004  (    0.011  )
    3.868  (    0.834  )
----- PHI MATRICES -----
ESTIMATES OF    PHI(  1 ) MATRIX AND SIGNIFICANCE
    1.934     .000      + .
     .000    1.538      . +
STANDARD ERRORS
     .058      --
      --      .058
ESTIMATES OF    PHI(  2 ) MATRIX AND SIGNIFICANCE
   -1.211     .000      - .
     .000    -.585      . -
STANDARD ERRORS
     .126      --
      --      .107
ESTIMATES OF    PHI(  3 ) MATRIX AND SIGNIFICANCE
     .188     .000      . .
    -.540    -.170      - .
STANDARD ERRORS
     .145      --
     .074     .112
ESTIMATES OF    PHI(  4 ) MATRIX AND SIGNIFICANCE
    -.127     .000      . .
     .172     .128      . .
STANDARD ERRORS
     .145      --
     .164     .112
ESTIMATES OF    PHI(  5 ) MATRIX AND SIGNIFICANCE
```

```
      .280     .000      + .
     -.120     .058      . .
 STANDARD ERRORS
      .127      --
      .175     .098
 ESTIMATES OF    PHI(  6 ) MATRIX AND SIGNIFICANCE
     -.116     .000      . .
      .252    -.042      + .
 STANDARD ERRORS
      .058      --
      .107     .042
 ----------------------
 ERROR COVARIANCE MATRIX
 ----------------------
               1           2
   1      .034792
   2     -.002320     .055795
 --
miden r1,r2. maxl 12
 TIME PERIOD ANALYZED . . . . . . . . . . . .  7  TO   296
 EFFECTIVE NUMBER OF OBSERVATIONS (NOBE). . .          290

 SERIES   NAME           MEAN       STD. ERROR
    1    R1             0.0000      0.1865
    2    R2             0.0000      0.2362
 SAMPLE CORRELATION MATRIX OF THE SERIES
   1.00
  -0.05  1.00
 CROSS CORRELATION MATRICES IN TERMS OF +,-,.
 LAGS  1 THROUGH  6
      . .      . .      . .      . .      . .      . .
      . .      . .      . .      . .      . .      . .
 LAGS  7 THROUGH 12
      . .      . .      . .      . .      + .      . .
      . .      . .      . .      . .      . .      . .
 --
p6(1,1)=0   <== Put model zero constraints
 --
cp6(1,1)=1
 --
p6(2,2)=0
 --
```

```
cp6(2,2)=1
 --
p5(2,2)=0
 --
cp5(2,2)=1
 --
p5(2,1)=0
 --
cp5(2,1)=1
 --
mestim m1. hold resi(r1,r2)
(** Result edited **)
 --
p5(1,1)=0
 --
cp5(1,1)=1
 --
p4(1,1)=0
 --
p4(2,1)=0
 --
cp4(1,1)=1
 --
cp4(2,1)=1
 --
mestim m1. hold resi(r1,r2)
 SUMMARY FOR THE MULTIVARIATE ARMA MODEL
 SERIES    NAME            MEAN         STD DEV    DIFFERENCE ORDER(S)
   1     X               -0.0568         1.0710
   2     Y               53.5091         3.1967
 NUMBER OF OBSERVATIONS =  296  (EFFECTIVE NUMBER = NOBE =  290)

 MODEL SPECIFICATION WITH PARAMETER VALUES
 PARAMETER              PARAMETER                  PARAMETER
  NUMBER               DESCRIPTION                  VALUE
 ----------    -----------------------------     ------------
    1                CONSTANT( 1)                -0.004119
    2                CONSTANT( 2)                 3.754644
    3          AUTOREGRESSIVE ( 1, 1, 1)          1.973142
  *FIXED*      AUTOREGRESSIVE ( 1, 1, 2)          0.000000
  ... (Use previous estimation results as initial estimates)
   11          AUTOREGRESSIVE ( 6, 2, 1)          0.226662
```

```
----------------------
ERROR COVARIANCE MATRIX
----------------------
               1           2
  1      .035817
  2     -.002865     .056348

FINAL MODEL SUMMARY WITH CONDITIONAL LIKELIHOOD PARAMETER ESTIMATES
----- CONSTANT VECTOR (STD ERROR) -----
   -0.004  (    0.011  )
    3.759  (    0.833  )
----- PHI MATRICES -----
ESTIMATES OF    PHI(  1 ) MATRIX AND SIGNIFICANCE
    1.977     .000      + .
     .000    1.529      . +
STANDARD ERRORS
     .055      --
      --      .054
ESTIMATES OF    PHI(  2 ) MATRIX AND SIGNIFICANCE
   -1.379     .000      - .
     .000    -.579      . -
STANDARD ERRORS
     .100      --
      --      .106
ESTIMATES OF    PHI(  3 ) MATRIX AND SIGNIFICANCE
     .346     .000      + .
    -.460    -.155      - .
STANDARD ERRORS
     .055      --
     .028     .097
ESTIMATES OF    PHI(  4 ) MATRIX AND SIGNIFICANCE
     .000     .000      . .
     .000     .135      . +
STANDARD ERRORS
      --       --
      --      .042
ESTIMATES OF    PHI(  5 ) MATRIX AND SIGNIFICANCE
     .000     .000      . .
     .000     .000      . .
STANDARD ERRORS
      --       --
      --       --
```

```
 ESTIMATES OF    PHI(  6 ) MATRIX AND SIGNIFICANCE
      .000     .000      . .
      .227     .000      + .
 STANDARD ERRORS
       --       --
      .047      --
 -----------------------
 ERROR COVARIANCE MATRIX
 -----------------------
               1           2
   1      .035927
   2     -.002878     .056344
 --
miden r1,r2. maxl 12.
 TIME PERIOD ANALYZED . . . . . . . . . . . . .  7  TO   296
 EFFECTIVE NUMBER OF OBSERVATIONS (NOBE). . .          290

 SERIES   NAME           MEAN       STD. ERROR
    1    R1              0.0000      0.1895
    2    R2              0.0000      0.2374
 NOTE: THE APPROX. STD. ERROR FOR THE ESTIMATED CORRELATIONS BELOW
       IS (1/NOBE**.5) =    0.05872

 SAMPLE CORRELATION MATRIX OF THE SERIES
   1.00
  -0.06  1.00
 CROSS CORRELATION MATRICES IN TERMS OF +,-,.
 LAGS  1 THROUGH  6
      . .      . .      . .      - .      . .      . .
      . .      . .      . .      . .      . .      . .
 LAGS  7 THROUGH 12
      . .      . .      . .      . .      + .      . .
      . .      . .      . .      . .      . .      . .
```

Let x_t be the Gas rate and y_t be the output CO_2 concentration. From the output, the fitted model is approximately

$$(1 - 1.98B + 1.38B^2 - .35B^2)x_t = a_{1t} \quad (19)$$

$$(1 - 1.53B + .58B^2 + .16B^3)y_t = 3.76 + (-.46B^3 + .23B^6)x_t + a_{2t} \quad (20)$$

where a_{1t} and a_{2t} are essentially uncorrelated. Note that Eq. (19) is basically the AR(3) model for input gas rate. From Eq. (20), we have

$$y_t = \frac{3.76}{1 - 1.53 + .58 + .16} + \frac{-.46B^3 + .23B^6}{1 - 1.53B + .58B^2 + .16B^3}x_t + \frac{1}{1 - 1.53B + .58B^2 + .16B^3}a_{2t}.$$

One can compare the above approximation model with transfer function model built in Lecture 1 using the estimated impluse response functions. They should be close.
Some discussions are in order:

1. The VARMA modeling does not require prior knowledge of input variables and can handle multiple inputs and multiple outputs.

2. The VARMA modeling, however, puts certain constraints on the formation of transfer function model and noise model. For the Gas-Furnace example, the above model using the same AR structure for the denominators of transfer function and noise model.

3. Finally, it is interesting to re-examine the results of consecutive VAR fits to the Gas-Furnace data. In particular, the unidirectional relationship does not appear in VAR(1) or VAR(2) model. Indeed, the fitted AR coefficient matrices are not lower triangular in VAR(1) or VAR(2) model. This indicates that under-specification of a model can result in misleading relationship between variables.

Example. Consider the quarterly data of the *Financial Times* ordinary share index, U.K. car production index, and the *Financial Times* commodity price index from 1952.III to 1967.IV. There are 62 observations and we denote the three series as $\boldsymbol{z}_t = (z_{1t}, z_{2t}, z_{3t})'$. Regression models were used in Coen et al. (1969) as prediction models for the series. We use VARMA models to analyze the data.
SCA demonstration: In SCA, the names of variables are x, y, and z.

```
miden x,y,z. arfits 1 to 6. rccm 1. maxl 12.
 TIME PERIOD ANALYZED . . . . . . . . . . . .  1  TO    62
 EFFECTIVE NUMBER OF OBSERVATIONS (NOBE). . .           62

 SERIES   NAME           MEAN       STD. ERROR
    1    X              3.1063      1.0000
    2    Y              2.8365      1.0000
    3    Z             16.5473      1.0000
 NOTE: THE APPROX. STD. ERROR FOR THE ESTIMATED CORRELATIONS BELOW
       IS (1/NOBE**.5) =    0.12700

 SAMPLE CORRELATION MATRIX OF THE SERIES
   1.00
   0.88  1.00
  -0.40 -0.24  1.00
 SUMMARIES OF CROSS CORRELATION MATRICES USING +,-,.
           1              2              3
  1   ++++++++++++   ++++++++++++   ------------

  2   ++++++++++++   ++++++++++++   ------------
```

```
 3   -------.....   ...........   ++++++++....

CROSS CORRELATION MATRICES IN TERMS OF +,-,.
LAGS  1 THROUGH  6
     + + -      + + -      + + -      + + -      + + -      + + -
     + + -      + + -      + + -      + + -      + + -      + + -
     - . +      - . +      - . +      - . +      - . +      - . +
LAGS  7 THROUGH 12
     + + -      + + -      + + -      + + -      + + -      + + -
     + + -      + + -      + + -      + + -      + + -      + + -
     - . +      . . +      . . .      . . .      . . .      . . .
DETERMINANT OF S(0) = 0.146394E+00

AUTOREGRESSIVE FITTING ON LAG(S)  1

SUMMARIES OF CROSS CORRELATION MATRICES USING +,-,.
          1              2              3
 1   ...........   ...........   ...........

 2   ...........   ...........-   ...........

 3   +..........   ...........   ...........

CROSS CORRELATION MATRICES IN TERMS OF +,-,.
LAGS  1 THROUGH  6
     . . .      . . .      . . .      . . .      . . .      . . .
     . . .      . . .      . . .      . . .      . . .      . . .
     + . .      . . .      . . .      . . .      . . .      . . .
LAGS  7 THROUGH 12
     . . .      . . .      . . .      . . .      . . .      . . .
     . . .      . . .      . . .      . . .      . . .      . - .
     . . .      . . .      . . .      . . .      . . .      . . .
==========  STEPWISE AUTOREGRESSION SUMMARY  ==========
----------------------------------------------------------------------
    I RESIDUAL I EIGENVAL.I CHI-SQ  I          I SIGNIFICANCE
LAG I VARIANCESI OF SIGMA I  TEST   I   AIC    I OF PARTIAL AR COEFF.
----+----------+----------+---------+----------+---------------------
  1 I .449E-01 I .338E-01 I  292.93 I   -7.319 I + . .
    I .900E-01 I .788E-01 I         I          I . + .
    I .164E+00 I .186E+00 I         I          I . . +
----+----------+----------+---------+----------+---------------------
```

```
  2 I .409E-01 I .310E-01 I   18.55 I   -7.411 I - . .
    I .849E-01 I .775E-01 I         I          I . . .
    I .123E+00 I .141E+00 I         I          I - + -
----+----------+----------+---------+----------+---------------------
  3 I .372E-01 I .261E-01 I    9.21 I   -7.323 I . . .
    I .819E-01 I .759E-01 I         I          I . . .
    I .122E+00 I .140E+00 I         I          I . . .
----+----------+----------+---------+----------+---------------------
  4 I .370E-01 I .251E-01 I    3.40 I   -7.113 I . . .
    I .801E-01 I .733E-01 I         I          I . . .
    I .120E+00 I .139E+00 I         I          I . . .
----+----------+----------+---------+----------+---------------------
  5 I .323E-01 I .235E-01 I   11.36 I   -7.110 I . + .
    I .713E-01 I .622E-01 I         I          I . + .
    I .113E+00 I .131E+00 I         I          I . . .
----+----------+----------+---------+----------+---------------------
  6 I .295E-01 I .226E-01 I   10.03 I   -7.095 I . . .
    I .569E-01 I .539E-01 I         I          I . . -
    I .109E+00 I .119E+00 I         I          I . . .
---------------------------------------------------------------------
NOTE:  CHI-SQUARED CRITICAL VALUES WITH   9  DEGREES OF FREEDOM ARE
             5 PERCENT:  16.9     1 PERCENT:  21.7
```

From the output, a VAR(2) model or a VARMA(1,1) model should be sufficient. The latter is identified because the residual cross-correlation matrices of VAR(1) fit show possible significant parameters in lag 1 only. We shall use VARMA(1,1) model in our analysis. VAR(2) analysis provides similar results.

The estimation of VARMA(1,1) model involves several steps in an effort to remove insignificant parameters. Details are given in below.

SCA demonstration: the variables are called x,y, and z.

```
input x,y,z. file 'cgk.dat'
 --
mtsm m1. series x,y,z. model (i-p1*b)series=c+(i-t1*b)noise.
 SUMMARY FOR MULTIVARIATE ARMA MODEL --    M1
 VARIABLE   DIFFERENCING
  X
  Y
  Z
  PARAMETER     FACTOR     ORDER    CONSTRAINT
  1        C   CONSTANT      0        CC
  2       P1    REG AR       1        CP1
  3       T1    REG MA       1        CT1
```

```
 --
mestim m1. method exact. hold resi(r1,r2,r3)

 SUMMARY FOR THE MULTIVARIATE ARMA MODEL
 SERIES    NAME           MEAN        STD DEV    DIFFERENCE ORDER(S)
   1     X              3.1063         1.0000
   2     Y              2.8365         1.0000
   3     Z             16.5473         1.0000
 NUMBER OF OBSERVATIONS =   62  (EFFECTIVE NUMBER = NOBE =   61)
 MODEL SPECIFICATION WITH PARAMETER VALUES

 PARAMETER              PARAMETER                  PARAMETER
  NUMBER               DESCRIPTION                  VALUE
 ----------    -----------------------------     ------------
     1               CONSTANT( 1)                   0.857298
     2               CONSTANT( 2)                   0.587504
     3               CONSTANT( 3)                  14.298246
     4         AUTOREGRESSIVE ( 1, 1, 1)            0.100000
 ....
    21         MOVING AVERAGE ( 1, 3, 3)            0.100000
 -----------------------
 ERROR COVARIANCE MATRIX
 -----------------------
               1           2           3
   1      .957295
   2      .833973     .952684
   3     -.347628    -.180279     .945902
 FINAL MODEL SUMMARY WITH MAXIMUM LIKELIHOOD PARAMETER ESTIMATES
 ----- CONSTANT VECTOR (STD ERROR) -----
     1.105  (    0.638  )
     1.864  (    0.821  )
     4.091  (    1.467  )
 ----- PHI MATRICES -----
 ESTIMATES OF    PHI(  1 ) MATRIX AND SIGNIFICANCE
      .811     .157    -.055      + + .
     -.082     .989    -.093      . + -
     -.331     .304     .760      . . +
 STANDARD ERRORS
      .079     .075     .035
      .102     .096     .046
      .183     .172     .081
 ----- THETA MATRICES -----
```

```
 ESTIMATES OF  THETA(  1 ) MATRIX AND SIGNIFICANCE
     -.297     .244     .067      . + .
     -.467     .221    -.157      - . .
     -.805     .597    -.430      - + -
 STANDARD ERRORS
      .153     .116     .073
      .221     .168     .107
      .277     .214     .129
 -----------------------
 ERROR COVARIANCE MATRIX
 -----------------------
              1           2           3
   1      .036735
   2      .021639     .077807
   3      .013216     .021723     .130048

 -2*(LOG LIKELIHOOD AT FINAL ESTIMATES)  IS  -0.31277415E+03
 --
input cp1. nrow 3. ncol 3.  <= Put constraints
 --
input ct1. nrow 3. ncol 3.
 --
p1(1,3)=0
 --
 ....
t1(2,3)=0
 --
mestim m1. method exact. hold resi(r1,r2,r3)

 FINAL MODEL SUMMARY WITH MAXIMUM LIKELIHOOD PARAMETER ESTIMATES
 ----- CONSTANT VECTOR (STD ERROR) -----
     0.110  (    0.079  )
     0.724  (    0.641  )
     2.631  (    1.097  )
 ----- PHI MATRICES -----
 ESTIMATES OF    PHI(  1 ) MATRIX AND SIGNIFICANCE
      .923     .065     .000      + . .
      .000     .921    -.028      . + .
      .000     .000     .839      . . +
 STANDARD ERRORS
      .053     .054      --
       --      .039     .037
```

```
      --       --      .066
 ----- THETA MATRICES -----
 ESTIMATES OF  THETA(  1 ) MATRIX AND SIGNIFICANCE
     .000     .112     .000      . . .
    -.202     .000     .000      . . .
    -.503     .220    -.444      - . -
 STANDARD ERRORS
      --      .092      --
     .168      --       --
     .246     .150     .122
 -----------------------
 ERROR COVARIANCE MATRIX
 -----------------------
              1           2           3
   1      .043320
   2      .022633     .079942
   3      .017799     .021173     .130828

 -2*(LOG LIKELIHOOD AT FINAL ESTIMATES)  IS  -0.30058643E+03
 --
cp1(1,2)=1   <== Put more constraints
 --
cp1(2,3)=1
 --
p1(1,2)=0
 --
...
t1(3,2)=0
 --
mestim m1. method exact. hold resi(r1,r2,r3)

 SUMMARY FOR THE MULTIVARIATE ARMA MODEL
 SERIES    NAME            MEAN         STD DEV    DIFFERENCE ORDER(S)
   1     X               3.1063         1.0000
   2     Y               2.8365         1.0000
   3     Z              16.5473         1.0000
 MODEL SPECIFICATION WITH PARAMETER VALUES

 PARAMETER              PARAMETER                  PARAMETER
  NUMBER               DESCRIPTION                  VALUE
 ----------    ----------------------------     ------------
    1              CONSTANT( 1)                  0.109976
```

```
   2              CONSTANT( 2)                 0.724382
   3              CONSTANT( 3)                 2.630637
   4        AUTOREGRESSIVE ( 1, 1, 1)          0.922758
 *FIXED*    AUTOREGRESSIVE ( 1, 1, 2)          0.000000
   5        AUTOREGRESSIVE ( 1, 2, 2)          0.921368
   6        AUTOREGRESSIVE ( 1, 3, 3)          0.838594
   7        MOVING AVERAGE ( 1, 3, 1)         -0.503301
   8        MOVING AVERAGE ( 1, 3, 3)         -0.443819
-----------------------
ERROR COVARIANCE MATRIX
-----------------------
               1           2           3
  1      .043320
  2      .022633     .079942
  3      .017799     .021173     .130828

FINAL MODEL SUMMARY WITH MAXIMUM LIKELIHOOD PARAMETER ESTIMATES
----- CONSTANT VECTOR (STD ERROR) -----
    0.128  (    0.082  )
    0.244  (    0.107  )
    2.537  (    1.067  )
----- PHI MATRICES -----
ESTIMATES OF    PHI(  1 ) MATRIX AND SIGNIFICANCE
     .977     .000     .000      + . .
     .000     .929     .000      . + .
     .000     .000     .844      . . +
STANDARD ERRORS
     .025      --       --
      --      .036      --
      --       --      .064
----- THETA MATRICES -----
ESTIMATES OF  THETA(  1 ) MATRIX AND SIGNIFICANCE
     .000     .000     .000      . . .
     .000     .000     .000      . . .
    -.402     .000    -.399      . . -
STANDARD ERRORS
      --       --       --
      --       --       --
     .226      --      .123
-----------------------
ERROR COVARIANCE MATRIX
-----------------------
```

```
              1           2           3
  1      .044685
  2      .024123     .085302
  3      .019532     .023950     .134203

 -2*(LOG LIKELIHOOD AT FINAL ESTIMATES)  IS  -0.29430693E+03
 --
miden r1,r2,r3. maxl 12.

 TIME PERIOD ANALYZED . . . . . . . . . . . . .  2  TO    62
 EFFECTIVE NUMBER OF OBSERVATIONS (NOBE). . .           61
 SERIES   NAME           MEAN       STD. ERROR
    1    R1             0.0004      0.2114
    2    R2             0.0003      0.2921
    3    R3             0.0009      0.3663
 NOTE: THE APPROX. STD. ERROR FOR THE ESTIMATED CORRELATIONS BELOW
       IS (1/NOBE**.5) =    0.12804
 SAMPLE CORRELATION MATRIX OF THE SERIES
   1.00
   0.39  1.00
   0.25  0.22  1.00
 CROSS CORRELATION MATRICES IN TERMS OF +,-,.
 LAGS  1 THROUGH  6
      . . .      . . .      . . .      . . .      . . .      . . .
      . . .      . . .      . . .      . . .      . . .      . . .
      . . .      . . .      . . .      . . .      . . .      . . .
 LAGS  7 THROUGH 12
      . . .      . . .      - . .      . . .      . . .      . . .
      . . .      . . .      . . -      . . .      . . .      . - .
      . . +      . . .      . . .      . . .      . . .      . . .
```

From the output, we summarize the estimation in Table 2. The residual cross-correlations indicate that the fitted model is adequate in modeling the dynamic dependence of the data. From the table, the fitted model shows that

$$
\begin{aligned}
z_{1t} &= 0.13 + 0.98 z_{1,t-1} + a_{1t} \\
z_{2t} &= 0.24 + 0.93 z_{2,t-1} + a_{2t} \\
z_{3t} &= 2.54 + 0.84 z_{3,t-1} + a_{3t} + 0.40 a_{1,t-1} + 0.40 a_{3,t-1}.
\end{aligned}
$$

Consequently, except for the contemporaneous depedence shown in residual covariance matrix, z_{1t} and z_{2t} are essentially random-walk with some weak time trend whereas z_{3t} follows ARMA(1,1) model with some minor dependence on the innovation $a_{1,t-1}$. In other words, the three index series are not dynamically correlated. This is not surprising given that the

Table 2: Estimation of VARMA(1,1) model CGK data set: exact likelihood method. The number in parentheses denotes standard error.

$\boldsymbol{c}$	$\boldsymbol{\phi}_1$			$\boldsymbol{\theta}_1$			$\boldsymbol{\Sigma}$		
(a) Initial estimation: Full model									
1.11	.81	.16	-.06	-.30	.24	.07	.03	.02	.01
(.64)	(.08)	(.08)	(.04)	(.15)	(.12)	(.07)			
1.86	-.08	.99	-.09	-.47	.22	-.16	.02	.08	.02
(.82)	(.10)	(.10)	(.05)	(.22)	(.17)	(.11)			
4.09	-.33	.30	.76	-.81	.60	-.43	.01	.02	.13
(1.5)	(.18)	(.17)	(.08)	(.28)	(.21)	(.13)			
(b) An intermediate model									
0.11	.92	.07	0	0	.11	0	.04	.02	.02
(.08)	(.05)	(.05)	–	–	(.09)	–			
0.72	0	.92	-.03	-.20	0	0	.02	.08	.02
(.64)	–	(.04)	(.04)	(.17)	–	–			
2.63	0	0	.84	-.50	.22	-.44	.02	.02	.13
(1.1)	–	–	(.07)	(.25)	(.15)	(.12)			
(c) A refined model									
0.13	.98	0	0	0	0	0	.04	.02	.02
(.08)	(.03)	–	–	–	–	–			
0.24	0	.93	0	0	0	0	.02	.09	.02
(.11)	–	(.04)	–	–	–	–			
2.54	0	0	.84	-.40	0	-.40	.02	.02	.13
(1.1)	–	–	(.06)	(.23)	–	(.12)			

three series are financial indices. This example demonstrates that VARMA modeling can show the relationship between variables.

5 Relation with Econometric Models

In this section, we discuss the relationship between VARMA and simultaneous equation models. The VARMA models are referred to as the *reduced form* models because the contemporaneous relationships between variables are embedded in the residual covariance matrix. On the other hand, simultaneous equation models use the concurrent values as one of the explanatory variables.

Let $\boldsymbol{z}_t$ be endogenous variables and $\boldsymbol{x}_t$ be exogenous variables. A *simultaneous equation model* can be written as

$$\boldsymbol{\Omega z}_t = \boldsymbol{\Phi}_0 + \sum_{i=1}^{p} \boldsymbol{\Phi}_i \boldsymbol{z}_{t-i} + \sum_{i=0}^{r} \boldsymbol{V}_i \boldsymbol{x}_{t-i} + \boldsymbol{b}_t, \tag{21}$$

where $\{\boldsymbol{b}_t\}$ is an iid sequence of random vectors with mean zero and $\text{Cov}(\boldsymbol{b}_t) = \boldsymbol{\Sigma}_b = \text{diag}\{\sigma_{b1}^2, \cdots, \sigma_{bk}^2\}$, and $\boldsymbol{\Omega}$ is a non-singular $k \times k$ matrix with diagonal elements being 1.

Given $\boldsymbol{z}_t$ and $\boldsymbol{x}_t$, the VARMA models can be generalized as

$$\boldsymbol{z}_t = \boldsymbol{\phi}_0 + \sum_{i=1}^{p} \boldsymbol{\phi}_i \boldsymbol{z}_{t-i} + \sum_{i=0}^{r} \boldsymbol{v}_i \boldsymbol{x}_{t-i} + \boldsymbol{a}_t, \tag{22}$$

where, for simplicity, we assume $\boldsymbol{\theta}(B) = \boldsymbol{I}$. This model is referred to as the vector ARMAX model in the literature, where "X" stands for exogenous variables.

Given Eq. (21), one can easily obtain Eq. (22) by pre-multiplying the model via $\boldsymbol{\Omega}^{-1}$. The parameters are related as

$$\boldsymbol{\Omega}^{-1}\boldsymbol{\Phi}_i = \boldsymbol{\phi}_i, \quad i = 0, \cdots, p; \quad \boldsymbol{\Omega}^{-1}\boldsymbol{V}_i = \boldsymbol{v}_i, \quad i = 0, \cdots, r; \quad \boldsymbol{\Omega}^{-1}\boldsymbol{\Sigma}_b(\boldsymbol{\Omega}^{-1})' = \boldsymbol{\Sigma}.$$

On the other hand, obtaining Eq. (21) from Eq. (22) is more involved. We focus on the last element of $\boldsymbol{z}_t$. From the simultaneous equation model, we have

$$z_{kt} + \sum_{i=1}^{k-1} \Omega_{ki} z_{it} = \Phi_{0,k} + \sum_{i=1}^{p}\sum_{j=1}^{k} \Phi_{i,kj} z_{j,t-i} + \sum_{i=0}^{r}\sum_{j=1}^{m} V_{i,kj} x_{j,t-i} + b_{kt}, \tag{23}$$

where $A_{\ell,ij}$ denotes the (i,j)th element of matrix $\boldsymbol{A}_\ell$. Because $\boldsymbol{\Sigma}$ of the VARMAX model is positive definite, by Cholesky Decomposition, we have $\boldsymbol{\Sigma} = \boldsymbol{LGL}'$, where $\boldsymbol{G}$ is a diagonal matrix and $\boldsymbol{L}$ is a lower triangular matrix with unit diagonal elements.

Define $\boldsymbol{b}_t = \boldsymbol{L}^{-1}\boldsymbol{a}_t$. Then, $E(\boldsymbol{b}_t) = \boldsymbol{0}$ and

$$\text{Cov}(\boldsymbol{b}_t = \boldsymbol{L}^{-1}\boldsymbol{\Sigma}(\boldsymbol{L}^{-1})' = \boldsymbol{G}.$$

Thus, the covariance matrix of $\boldsymbol{b}_t$ is a diagonal matrix. Multiplying $\boldsymbol{L}^{-1}$ from left to Eq. (22), we have

$$\boldsymbol{L}^{-1}\boldsymbol{z}_t = \boldsymbol{L}^{-1}\boldsymbol{\phi}_0 + \sum_{i=1}^{p} \boldsymbol{L}^{-1}\boldsymbol{\phi}_i \boldsymbol{z}_{t-i} + \sum_{i=0}^{r} \boldsymbol{L}^{-1}\boldsymbol{v}_i \boldsymbol{x}_{t-i} + \boldsymbol{L}^{-1}\boldsymbol{a}_t.$$

Since $\boldsymbol{L}^{-1}$ is also lower triangular with unit diagonal elements, the kth row of $\boldsymbol{L}^{-1}$ is in the form $(\Omega_{k1}, \cdots, \Omega_{k,k-1}, 1)$. Define

$$\boldsymbol{\Phi}_i = \boldsymbol{L}^{-1}\boldsymbol{\phi}_i, \quad i = 0, \cdots, p; \quad \boldsymbol{V}_i = \boldsymbol{L}^{-1}\boldsymbol{v}_i, \quad i = 0, \cdots, r.$$

We obtain

$$z_{kt} + \sum_{i=1}^{k-1} \Omega_{ki} z_{it} = \Phi_{0,k} + \sum_{i=1}^{p}\sum_{j=1}^{k} \Phi_{i,kj} z_{j,t-i} + \sum_{i=0}^{r}\sum_{j=1}^{m} V_{i,kj} x_{j,t-i} + b_{kt},$$

which is Eq. (23). Thus, we can obtain a simultaneous equation model from a given vector ARMA model.

Example. Consider the bivariate model

$$\begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \begin{bmatrix} 0.2 \\ 0.4 \end{bmatrix} + \begin{bmatrix} 0.2 & 0.3 \\ -0.6 & 1.1 \end{bmatrix} \begin{bmatrix} z_{1,t-1} \\ z_{2,t-1} \end{bmatrix} + \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

For this particular covariance matrix $\boldsymbol{\Sigma}$, the $\boldsymbol{L}^{-1}$ matrix is

$$\boldsymbol{L}^{-1} = \begin{bmatrix} 1.0 & 0.0 \\ -0.5 & 1.0 \end{bmatrix}.$$

It is easy to verify that $\boldsymbol{G}$ = diag{2,0.5} and the simultaneous equation for z_{2t} is

$$z_{2t} = 0.3 + 0.5z_{1t} - 0.7z_{1,t-1} + 0.95z_{2,t-1} + b_{2t}.$$

To obtain the simultaneous equation for z_{1t}, we can rewrite the VAR(1) model as

$$\begin{bmatrix} z_{2t} \\ z_{1t} \end{bmatrix} = \begin{bmatrix} 0.4 \\ 0.2 \end{bmatrix} + \begin{bmatrix} 1.1 & -0.6 \\ 0.3 & 0.2 \end{bmatrix} \begin{bmatrix} z_{2,t-1} \\ z_{1,t-1} \end{bmatrix} + \begin{bmatrix} a_{2t} \\ a_{1t} \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}.$$

In this instance, the $\boldsymbol{L}^{-1}$ matrix for $\boldsymbol{\Sigma}$ is

$$\boldsymbol{L}^{-1} = \begin{bmatrix} 1.0 & 0.0 \\ -1.0 & 0.0 \end{bmatrix}.$$

It is easy to verify that

$$r_{1t} = -0.2 + r_{2t} + 0.8r_{1,t-1} - 0.8r_{2,t-1} + b_{1t},$$

where $\text{Var}(b_{1t}) = 1.0$.