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. Output edited.

```
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 NUMBER DESCRIPTION VALUE
---------- ----------------------------- ------------
1 CONSTANT( 1) 17.066420
2 CONSTANT( 2) 25.039345
3 MOVING AVERAGE ( 1, 1, 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

<table>
<thead>
<tr>
<th>SERIES</th>
<th>NAME</th>
<th>MEAN</th>
<th>STD DEV</th>
<th>DIFFERENCE</th>
<th>ORDER(S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Z1</td>
<td>17.0664</td>
<td>2.3334</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Z2</td>
<td>25.0393</td>
<td>1.5875</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NUMBER OF OBSERVATIONS = 250  (EFFECTIVE NUMBER = NOBE = 250)

MODEL SPECIFICATION WITH PARAMETER VALUES

<table>
<thead>
<tr>
<th>PARAMETER</th>
<th>PARAMETER</th>
<th>PARAMETER</th>
</tr>
</thead>
<tbody>
<tr>
<td>NUMBER</td>
<td>DESCRIPTION</td>
<td>VALUE</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
<td>-----------</td>
</tr>
<tr>
<td>1</td>
<td>CONSTANT( 1)</td>
<td>17.152506</td>
</tr>
<tr>
<td>2</td>
<td>CONSTANT(2)</td>
<td>25.107561</td>
</tr>
<tr>
<td>3</td>
<td>MOVING AVERAGE ( 1, 1, 1)</td>
<td>0.192436</td>
</tr>
<tr>
<td>4</td>
<td>MOVING AVERAGE ( 1, 1, 2)</td>
<td>0.409907</td>
</tr>
<tr>
<td>5</td>
<td>MOVING AVERAGE ( 1, 2, 1)</td>
<td>-0.574374</td>
</tr>
<tr>
<td>6</td>
<td>MOVING AVERAGE ( 1, 2, 2)</td>
<td>1.124622</td>
</tr>
</tbody>
</table>

-----------------------

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

\(-2*(\text{LOG LIKELIHOOD AT FINAL ESTIMATES})\) IS  0.82972019E+03

miden r1,r2. maxl 12.

TIME PERIOD ANALYZED . . . . . . . . . . . . 1 TO 250
EFFECTIVE NUMBER OF OBSERVATIONS (NOBE) . . 250

<table>
<thead>
<tr>
<th>SERIES</th>
<th>NAME</th>
<th>MEAN</th>
<th>STD. ERROR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>R1</td>
<td>-0.0137</td>
<td>2.2055</td>
</tr>
<tr>
<td>2</td>
<td>R2</td>
<td>0.0002</td>
<td>1.0072</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>SERIES: Z1</th>
<th>Z2</th>
</tr>
</thead>
<tbody>
<tr>
<td>TIME</td>
<td>FORECAST</td>
</tr>
<tr>
<td>251</td>
<td>18.019</td>
</tr>
<tr>
<td>253</td>
<td>17.083</td>
</tr>
<tr>
<td>254</td>
<td>17.083</td>
</tr>
</tbody>
</table>

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
\[
\hat{\Sigma} = \begin{bmatrix}
  4.89 & 1.13  \\
  1.13 & 1.01
\end{bmatrix}.
\]
4 Vector ARMA Models

A $k$-dimensional time series $z_t$ follows a vector ARMA, VARMA($p,q$), model if

$$\phi(B)z_t = c + \theta(B)a_t$$  \hspace{1cm} (1)

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

1. $\phi(B)$ and $\theta(B)$ are left coprime, i.e. if $u(B)$ is a left common factor of $\phi(B)$ and $\theta(B)$, then $|u(B)|$ is a non-zero constant. Such a polynomial matrix is called a unimodular matrix. In theory, $u(B)$ is unimodular if and only if $u^{-1}(B)$ 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 $\phi_p$ and $\theta_q$ satisfy the condition that $\text{Rank}[[\phi_p, \theta_q]] = \text{dim}(z_t)$.

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 $\phi(B)$ and $\theta(B)$ 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,

$$z_t = \mu + \sum_{i=0}^{\infty} \psi_i a_{t-i},$$  \hspace{1cm} (2)

where $\psi_0 = I$, and $\{a_t\}$ is an iid sequence of random vectors with mean zero and positive-definite covariance matrix $\Sigma$. A VARMA model is said to be identifiable if the matrix polynomials $\phi(B)$ and $\theta(B)$ are uniquely determined by the $\psi_i$ in Eq. (2). There are cases for which multiple pairs of AR and MA matrix polynomials give rise to the same $\psi_i$s. We use simple bivariate models to discuss the issue.

**Example.** Consider the VMA(1) model

$$\begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} - \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a_{1,t-1} \\ a_{2,t-1} \end{bmatrix}.$$  \hspace{1cm} (3)

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

$$\begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} - \begin{bmatrix} 0 & -2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_{1,t-1} \\ z_{2,t-1} \end{bmatrix} = \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}.$$  \hspace{1cm} (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
\[
\begin{bmatrix}
  z_{1t} \\
  z_{2t}
\end{bmatrix}
- \begin{bmatrix}
  0.8 & 2 \\
  0 & 0
\end{bmatrix}
\begin{bmatrix}
  z_{1,t-1} \\
  z_{2,t-1}
\end{bmatrix}
= \begin{bmatrix}
  a_{1t} \\
  a_{2t}
\end{bmatrix}
- \begin{bmatrix}
  0.3 & 0 \\
  0 & 0
\end{bmatrix}
\begin{bmatrix}
  a_{1,t-1} \\
  a_{2,t-1}
\end{bmatrix},
\]
(5)
It is easy to see that the model is identical to
\[
\begin{bmatrix}
  z_{1t} \\
  z_{2t}
\end{bmatrix}
- \begin{bmatrix}
  0.8 & 2 + \omega \\
  0 & \beta
\end{bmatrix}
\begin{bmatrix}
  z_{1,t-1} \\
  z_{2,t-1}
\end{bmatrix}
= \begin{bmatrix}
  a_{1t} \\
  a_{2t}
\end{bmatrix}
- \begin{bmatrix}
  0.3 & \omega \\
  0 & \beta
\end{bmatrix}
\begin{bmatrix}
  a_{1,t-1} \\
  a_{2,t-1}
\end{bmatrix},
\]
(6)
for any \( \omega \neq 0 \) and \( \beta \neq 0 \). From Eq. (5), we have
\[
\begin{align*}
  z_{1t} &= 0.8z_{1,t-1} + 2z_{2,t-1} + a_{1t} - .3a_{1,t-1} \\
  z_{2t} &= a_{2t}.
\end{align*}
\]
(7)\( (8)\)
Thus, \( z_{2t} \) is a white noise series. Multiplying Eq. (8) by \( \omega \) 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 constraints, 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 \( \text{Rank}[\phi_1, \theta_1] = 1 \), which is smaller than the dimension of \( z_t \). 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, \( \mathbf{u}(B) = \text{diag}\{1, 1 - \beta B\} \) is a left-common factor. Finally, the identifiability problem can occur even if none of the components of \( z_t \) 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 $a_t$ has mean zero and covariance matrix $\Sigma$, which is positive definite.

1. Stationarity condition

Similar to a VAR($p$) model, the necessary and sufficient condition for weak stationarity of the $z_t$ process in Eq. (1) is that all zeros of the polynomial $|\phi(B)|$ 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

$$\phi(B)(z_t - \mu) = \theta(B)a_t,$$

where $\mu$ satisfies $\phi(1)\mu = \phi_0$ and is the mean vector of $z_t$. Let $\tilde{z}_t = z_t - \mu$ be the mean-corrected series. The VARMA($p,q$) model then becomes

$$\phi(B)\tilde{z}_t = \theta(B)a_t. \quad (10)$$

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

When $z_t$ is stationary, it has an MA representation as

$$\tilde{z}_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \ldots = \psi(B)a_t, \quad (11)$$

where $\psi(B) = \sum_{i=0}^{\infty} \psi_i B^i$ with $\psi_0 = I_k$, the $k \times k$ identity matrix. The coefficient matrices $\psi_i$ can obtain recursively by equating the coefficients of $B$ in

$$\phi(B)\psi(B) = \theta(B).$$

For simplicity, let $\theta_i = 0$ if $i > q$, $\phi_i = 0$ if $i > p$, and $m = \max\{p,q\}$. Then, the coefficient matrices $\psi_i$ can be obtained recursively as follows:

$$\begin{align*}
\psi_1 &= \phi_1 - \theta_1 \\
\psi_2 &= \phi_1 \psi_1 + \phi_2 - \theta_2 \\
\psi_3 &= \phi_1 \psi_2 + \phi_2 \psi_1 + \phi_3 - \theta_3 \\
& \vdots \\
\psi_m &= \phi_1 \psi_{m-1} + \phi_2 \psi_{m-2} + \ldots + \phi_{m-1} \psi_1 + \phi_m - \theta_m \\
\psi_\ell &= \phi_1 \psi_{\ell-1} + \phi_2 \psi_{\ell-2} + \ldots + \phi_m \psi_{\ell-m}, \quad \ell > m.
\end{align*}$$

In particular, we have

$$\psi_i = \sum_{j=1}^{\min\{p,i\}} \phi_j \psi_{j-i}, \quad i > q.$$
As discussed in the VAR case, these $\psi_i$ matrices are useful in calculating the impulse response function of the VARMA$(p,q)$ model.

The $\psi$ matrices are also useful in forecasting. Using the MA representation, it is easy to see that the $\ell$-step ahead forecast error at the forecast origin $t$ is

$$e_t(\ell) = a_{t+\ell} + \psi_1 a_{t+\ell-1} + \cdots + \psi_{\ell-1} a_{t+1}.$$  

Consequently, the covariance matrix of $\ell$-step ahead forecast error is

$$\text{Cov}[e_t(\ell)] = \Sigma + \psi_1 \Sigma \psi_1' + \cdots + \psi_{\ell-1} \Sigma \psi_{\ell-1}' .$$

2. Invertibility condition

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

$$\pi(B)(z_t - \mu) = a_t,$$  

(12)

where $\pi(B) = I - \pi_1 B - \pi_2 B^2 - \cdots$ and the coefficient matrices $\pi_i$ can be obtained by equating the coefficients of $B^i$ in

$$\theta(B)\pi(B) = \phi(B).$$

Again, let $\phi_i = 0$ if $i > p$, $\theta_j = 0$ if $j > q$, and $m = \max\{p,q\}$. It is easy to see that

$$\pi_1 = \phi_1 - \theta_1,$$

$$\pi_2 = \phi_2 + \theta_1 \pi_1 - \theta_2,$$

$$\pi_3 = \phi_3 + \theta_1 \pi_2 + \theta_2 \pi_1 - \theta_3,$$

$$\vdots = \vdots$$

$$\pi_m = \phi_m + \theta_1 \pi_{m-1} + \theta_2 \pi_{m-2} + \cdots + \theta_{m-1} \pi_1 - \theta_m,$$

$$\pi_\ell = \theta_1 \pi_{\ell-1} + \theta_2 \pi_{\ell-2} + \cdots + \theta_m \pi_{\ell-m}, \quad \ell > m.$$  

Note that $\pi(B)\psi(B) = I = \psi(B)\pi(B)$.

3. Moment equations

For a stationary VARMA$(p,q)$ series $z_t$, one can use the MA representation to obtain

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

(13)

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

$$\Gamma_\ell - \sum_{i=1}^p \phi_i \Gamma_{i-\ell} = \begin{cases} 
-\sum_{j=0}^q \theta_j \Sigma \psi'_{\ell-j} & \text{if } \ell = 0,1,\ldots,q, \\
0 & \text{if } \ell > q. 
\end{cases}$$  

(14)
where, for convenience, $\theta_0 = -I$. In particular,

$$\Gamma_\ell = \phi_1 \Gamma_{\ell-1} + \ldots + \phi_p \Gamma_{\ell-p}, \quad \ell = q + 1, \ldots, q + p,$$

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-multiply 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

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

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 $z_t$, the necessary and sufficient conditions that $z_{1t}$ is an input variable and $z_{2t}$ is an output 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. Univariate component models

For a stationary VARMA($p,q$) series $z_t$, 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. Nevertheless, 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 $E(z_t) = 0$ or equivalently, we use the model in Eq. (10).

Suppose that the data set is $\{z_1, \ldots, z_n\}$. Let $Z = (z_1, z_2, \ldots, z_n)'$ be the $kn \times 1$ vector of the data, and $A = (a_1', a_2', \ldots, a_n')'$. Also, let $Z_0 = (z_1-p, z_2-p, \ldots, z_0)'$ be the $kp \times 1$ vector of presample data and $A_0 = (a_1'-q, a_2'-q, \ldots, a_0)'$ be the $kq \times 1$ vector of presample innovations. Finally, let $U_0 = (Z_0, A_0)'$ be the vector of presample variables. Using these newly defined variables, the sample of the model can be written as

$$\Phi Z = \Theta A + PU_0,$$

(16)

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

$$P = \begin{bmatrix} G_p & H_q \\ 0 & 0 \end{bmatrix},$$

where

$$G_p = \begin{bmatrix} \phi_p & \phi_{p-1} & \cdots & \phi_1 \\ 0 & \phi_p & \cdots & \phi_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \phi_p \end{bmatrix}, \quad H_q = -\begin{bmatrix} \theta_q & \theta_{q-1} & \cdots & \theta_1 \\ 0 & \theta_q & \cdots & \theta_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \theta_q \end{bmatrix}.$$

Since $A$ is independent of $U_0$ and $\text{Cov}(A) = I_n \otimes \Sigma$, the covariance matrix of $Z$ in Eq. (16) is

$$\Gamma_Z = \Phi^{-1}[\Theta(I_n \otimes \Sigma)\Theta'] + P\Xi P' (\Phi')^{-1},$$

(17)

where $\Xi = \text{Cov}(U_0)$. For the stationary VARMA($p,q$) model $z_t$ in Eq. (1), the covariance matrix $\Xi$ can be written as

$$\Xi = \begin{bmatrix} \Gamma(p) & C' \\ C & I_q \otimes \Sigma \end{bmatrix}.$$
where $\Gamma(p)$ is the $kp \times kp$ covariance matrix of $Z_0$ such that the $(i,j)$th block of which is $\Gamma_{i-j}$, and $C = \text{Cov}(A_0,Z_0)$ that can be obtained by using Eq. (13) as $C = (I_q \otimes \Sigma)\Psi$ with

$$\Psi = \begin{bmatrix}
\psi'_{-p} & \psi'_{-1-p} & \ldots & \psi'_{-1} \\
\psi'_{-2-p} & \psi'_{-2} & \ldots & \psi'_{-2} \\
\vdots & \vdots & \ddots & \vdots \\
\psi'_{-p} & \psi'_{-p} & \ldots & I_k
\end{bmatrix},$$

being a $kq \times kp$ matrix of the $\psi$-weight matrices of $z_t$, where it is understood that $\psi_j = 0$ if $j < 0$.

Using the matrix inversion formula

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

we obtain the inverse of $\Gamma_Z$ as

$$\Gamma_Z^{-1} = \Phi'(\Omega^{-1} - \Omega^{-1}PQ^{-1}P'\Omega^{-1})\Phi,$$

(18)

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

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

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

$$\Xi^{-1} = \begin{bmatrix} \Delta^{-1} & -\Delta^{-1}\Psi' \\ -\Psi\Delta^{-1} & (I_q \otimes \Sigma^{-1}) + \Psi\Delta^{-1}\Psi' \end{bmatrix},$$

where $\Delta = \Gamma(p) - \Psi'(I_q \otimes \Sigma)\Psi$.

We are ready to write down the likelihood function of the sample $Z$. Using the property of determinant

$$\begin{vmatrix} W & J \\ B & D \end{vmatrix} = |W| |D - BW^{-1}J|,$$

and the definition of $\Xi$, we obtain

$$|\Xi| = |(I_q \otimes \Sigma)||\Gamma(p) - \Psi'(I_q \otimes \Sigma)\Psi| = |(I_q \otimes \Sigma)||\Delta| = |\Sigma|^q|\Delta|.$$

For a given order $(p,q)$, let $\vartheta$ be the set of all parameters in a VARMA$(p,q)$ model. Let $X = (x'_1, \ldots, x'_n)' = \Phi Z$. That is, $x_t = z_t - \sum_{i=1}^{t-1} \phi_i z_{t-i}$ for $t = 1, \ldots, p$ and $x_t = z_t - \sum_{i=1}^{p} \phi_i z_{t-i}$, where it is understood that the summation term disappears if the lower limit exceeds the upper limit. From Eq. (17) and using the identity

$$|A_{22}| |A_{11} - A_{12}A_{22}^{-1}A_{21}| = |A_{11}| |A_{22} - A_{21}A_{11}^{-1}A_{12}|,$$
where the inverses involved exist, we obtain the determinant of $\Gamma_Z$ as

$$|\Gamma_Z| = |\Phi|^{-2}\Omega + P\Xi P' = |\Omega| |\Xi| |\Xi^{-1} + P'\Omega^{-1}P| = |\Omega| |\Xi| |Q| = |\Sigma|^{n} |\Xi| |Q|,$$

where we have used $|\Phi| = 1$ and $|\Omega| = |\Sigma|^n$. Consequently, ignoring the normalizing factor involving $2\pi$, the exact likelihood function of $Z$ can be written as

$$L(\theta; Z) = |\Sigma|^{-(n+q)/2}|Q|^{-1/2}|\Delta|^{-1/2}\exp\left\{-\frac{1}{2}X'\left[\Omega^{-1} - \Omega^{-1}PQ^{-1}P'\Omega^{-1}\right]X\right\}. \quad (19)$$

The quadratic form in the exponent of Eq. (19) can also be written as

$$Z'\Gamma_Z^{-1}Z = X'\Omega^{-1}X - \hat{U}_0'Q\hat{U}_0,$$

where $\hat{U}_0 = Q^{-1}P'\Omega^{-1}X$.

### 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 + +
STANDARD ERRORS
.020  .007
.036  .012

RESIDUAL COVARIANCE MATRIX S(1)
0.102E+00
0.894E-01 0.343E+00

RESIDUAL CORRELATION MATRIX R(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
### Autoregressive Fitting on Lags 7 Through 12

**Autoregressive Fitting on Lags (1)**

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Standard Error</th>
<th>Sign</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.813</td>
<td>0.040</td>
<td>+</td>
</tr>
<tr>
<td>-0.076</td>
<td>0.022</td>
<td>-</td>
</tr>
<tr>
<td>0.235</td>
<td>0.055</td>
<td></td>
</tr>
<tr>
<td>1.446</td>
<td>0.030</td>
<td>+</td>
</tr>
</tbody>
</table>

**Standard Errors**

- 0.040
- 0.022
- 0.055
- 0.030

**Autoregressive Fitting on Lags (2)**

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Standard Error</th>
<th>Sign</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.961</td>
<td>0.048</td>
<td>-</td>
</tr>
<tr>
<td>0.046</td>
<td>0.019</td>
<td>+</td>
</tr>
<tr>
<td>-0.642</td>
<td>0.065</td>
<td></td>
</tr>
<tr>
<td>-0.580</td>
<td>0.027</td>
<td>-</td>
</tr>
</tbody>
</table>

**Standard Errors**

- 0.048
- 0.019
- 0.065
- 0.027

**Residual Covariance Matrix S(2)**

<table>
<thead>
<tr>
<th></th>
<th>0.367E-01</th>
<th>-0.381E-02</th>
<th>0.683E-01</th>
</tr>
</thead>
</table>

**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**

<table>
<thead>
<tr>
<th>Correlation</th>
<th>0.12</th>
<th>0.02</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.11</td>
<td>0.17</td>
<td></td>
</tr>
</tbody>
</table>

**Lag = 2**

<table>
<thead>
<tr>
<th>Correlation</th>
<th>-0.02</th>
<th>0.10</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.15</td>
<td>0.12</td>
<td></td>
</tr>
</tbody>
</table>

...  

**Lag = 5**

<table>
<thead>
<tr>
<th>Correlation</th>
<th>0.04</th>
<th>-0.01</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.13</td>
<td>0.01</td>
<td></td>
</tr>
</tbody>
</table>

**Lag = 6**

<table>
<thead>
<tr>
<th>Correlation</th>
<th>0.14</th>
<th>-0.11</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.10</td>
<td>0.18</td>
<td></td>
</tr>
</tbody>
</table>
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) ===
STANDARD ERRORS

0.137  0.077
0.176  0.099

=== PHI(4) ===

0.102  0.017
0.283  0.163

STANDARD ERRORS

0.082  0.033
0.105  0.042

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  -0.065
0.075  1.562

STANDARD ERRORS

0.059  0.046
0.075  0.059

=== PHI(2) ===

-1.193  0.117
-0.148  -0.615

STANDARD ERRORS

0.128  0.085
0.164  0.108

=== PHI(3) ===

0.147  -0.053
-0.414  -0.201

STANDARD ERRORS

0.147  0.088
0.187  0.113

=== PHI(4) ===

-0.024  -0.012
-0.008  0.185

STANDARD ERRORS

0.139  0.078
0.178  0.099
PHI(5)

0.094  0.011  
0.213 -0.017  

STANDARD ERRORS
0.084  0.033  
0.107  0.043  

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  -0.051  +
0.064   1.545  +

STANDARD ERRORS
0.058  0.046  
0.075  0.059  

PHI(2)

-1.204   0.100  -
-0.135  -0.593  -

STANDARD ERRORS
0.127  0.085  
0.162  0.108  

PHI(3)

-0.167   -0.079  
-0.438  -0.171  -

STANDARD ERRORS
0.145  0.088  
0.185  0.113  

PHI(4)

-0.158   0.027  
0.150   0.132  

STANDARD ERRORS
0.146  0.088  
0.187  0.112  

PHI(5)

-0.379   -0.042  +
-0.120   0.057  

STANDARD ERRORS
0.138  0.077  

17
<table>
<thead>
<tr>
<th></th>
<th>1.76</th>
<th>0.099</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHI(6)</td>
<td>-0.214</td>
<td>0.031</td>
</tr>
<tr>
<td></td>
<td>0.250</td>
<td>-0.042</td>
</tr>
</tbody>
</table>

**STANDARD ERRORS**

<table>
<thead>
<tr>
<th></th>
<th>0.084</th>
<th>0.033</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.108</td>
<td>0.042</td>
</tr>
</tbody>
</table>

**RESIDUAL COVARIANCE MATRIX S(6)**

<table>
<thead>
<tr>
<th></th>
<th>0.343E-01</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.229E-02</td>
<td>0.560E-01</td>
</tr>
</tbody>
</table>

**RESIDUAL CORRELATION MATRIX RS(6)**

<table>
<thead>
<tr>
<th></th>
<th>1.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.05</td>
<td>1.00</td>
</tr>
</tbody>
</table>

**EIGENVALUES AND EIGENVECTORS OF S(6)**

<table>
<thead>
<tr>
<th>EIGENVALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.034</td>
</tr>
<tr>
<td>0.056</td>
</tr>
</tbody>
</table>

**CROSS CORRELATION MATRICES IN TERMS OF +,-,.**

**LAGS 1 THROUGH 6**

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**LAGS 7 THROUGH 12**

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**AUTOREGRESSIVE FITTING ON LAG(S) 1 2 3 4 5 6 7**

<table>
<thead>
<tr>
<th>PHI(1)</th>
<th>1.924</th>
<th>-0.047</th>
<th>+</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.059</td>
<td>1.547</td>
<td>+</td>
</tr>
</tbody>
</table>

**STANDARD ERRORS**

<table>
<thead>
<tr>
<th></th>
<th>0.059</th>
<th>0.046</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.075</td>
<td>0.059</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PHI(2)</th>
<th>-1.188</th>
<th>0.095</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.123</td>
<td>-0.594</td>
<td>-</td>
</tr>
</tbody>
</table>

**STANDARD ERRORS**

<table>
<thead>
<tr>
<th></th>
<th>0.128</th>
<th>0.085</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.164</td>
<td>0.108</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PHI(3)</th>
<th>0.156</th>
<th>-0.085</th>
<th>.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.446</td>
<td>-0.158</td>
<td>-</td>
</tr>
</tbody>
</table>

**STANDARD ERRORS**

|         | 0.145 | 0.089 |
.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
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 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 $b = 3$.  

Note: Chi-Squared critical values with 4 degrees of freedom are
5 percent: 9.5  1 percent: 13.3
Table 1: VAR(6) Fit via Least Squares Method for the Gas-Furnace Data

<table>
<thead>
<tr>
<th></th>
<th>$\phi_1$</th>
<th>$\phi_2$</th>
<th>$\phi_3$</th>
<th>$\phi_4$</th>
<th>$\phi_5$</th>
<th>$\phi_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.93</td>
<td>-.05</td>
<td>-1.20</td>
<td>.10</td>
<td>-.08</td>
<td>.03</td>
</tr>
<tr>
<td></td>
<td>.06</td>
<td>1.55</td>
<td>-1.44</td>
<td>-.59</td>
<td>-1.17</td>
<td>.13</td>
</tr>
<tr>
<td>Standard error</td>
<td></td>
<td>.15</td>
<td>.44</td>
<td>.10</td>
<td>-.12</td>
<td>.06</td>
</tr>
<tr>
<td></td>
<td>.06</td>
<td>.05</td>
<td>.13</td>
<td>.09</td>
<td>.15</td>
<td>.09</td>
</tr>
<tr>
<td></td>
<td>.08</td>
<td>.06</td>
<td>.16</td>
<td>.11</td>
<td>.19</td>
<td>.11</td>
</tr>
<tr>
<td>Simplified notation</td>
<td>+</td>
<td>.</td>
<td>-</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>.</td>
<td>+</td>
<td>.</td>
<td>-</td>
<td>.</td>
<td>+</td>
</tr>
</tbody>
</table>

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:**

```plaintext
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**

<table>
<thead>
<tr>
<th>VARIABLE DIFFERENCING</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
</tr>
<tr>
<td>Y</td>
</tr>
<tr>
<td>PARAMETER</td>
</tr>
<tr>
<td>FACTOR</td>
</tr>
<tr>
<td>ORDER</td>
</tr>
<tr>
<td>CONSTRAINT</td>
</tr>
<tr>
<td>1  C1 CONSTANT</td>
</tr>
<tr>
<td>2  P1 REG AR 1 CP1</td>
</tr>
<tr>
<td>3  P2 REG AR 2 CP2</td>
</tr>
<tr>
<td>4  P3 REG AR 3 CP3</td>
</tr>
<tr>
<td>5  P4 REG AR 4 CP4</td>
</tr>
<tr>
<td>6  P5 REG AR 5 CP5</td>
</tr>
<tr>
<td>7  P6 REG AR 6 CP6</td>
</tr>
</tbody>
</table>

```plaintext
input cp1. nrow 2. ncol 2. <= One way to put constraint
```

```plaintext
input cp2=cp1
```

```plaintext
input p1. nrow 2. ncol 2. <= Put initial values
```

```plaintext
p2=p1
```
cp3=cp2

\[ cp3(2,1) = 0 \]  \text{Another way to put constraint}

p3=p1

\[ p3(2,1) = 0.1 \]  \text{Another way to put initial value}

p4=p3

p5=p4

p6=p5

mestim m1. hold resi(r1,r2)

### SUMMARY FOR THE MULTIVARIATE ARMA MODEL

<table>
<thead>
<tr>
<th>SERIES</th>
<th>NAME</th>
<th>MEAN</th>
<th>STD DEV</th>
<th>DIFFERENCE ORDER(S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X</td>
<td>-0.0568</td>
<td>1.0710</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Y</td>
<td>53.5091</td>
<td>3.1967</td>
<td></td>
</tr>
</tbody>
</table>

**NUMBER OF OBSERVATIONS = 296 (EFFECTIVE NUMBER = NOBE = 290)**

### MODEL SPECIFICATION WITH PARAMETER VALUES

<table>
<thead>
<tr>
<th>PARAMETER NUMBER</th>
<th>PARAMETER DESCRIPTION</th>
<th>VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CONSTANT( 1)</td>
<td>-0.022734</td>
</tr>
<tr>
<td>2</td>
<td>CONSTANT( 2)</td>
<td>21.426382</td>
</tr>
<tr>
<td>3</td>
<td>AUTOREGRESSIVE ( 1, 1, 1)</td>
<td>0.100000</td>
</tr>
<tr>
<td><em>FIXED</em></td>
<td>AUTOREGRESSIVE ( 1, 1, 2)</td>
<td>0.000000</td>
</tr>
<tr>
<td>4</td>
<td>AUTOREGRESSIVE ( 1, 2, 2)</td>
<td>0.100000</td>
</tr>
<tr>
<td>18</td>
<td>AUTOREGRESSIVE ( 6, 2, 2)</td>
<td>0.100000</td>
</tr>
</tbody>
</table>

|------------------|-------------------------------------|-------------|

### ERROR COVARIANCE MATRIX
\begin{verbatim}
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 + .
\end{verbatim}
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 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 -----  

<table>
<thead>
<tr>
<th>Estimates of PHI(1) Matrix and Significance</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.977 .000 +</td>
<td>.000 1.529 . +</td>
</tr>
</tbody>
</table>

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


\[ 
\begin{array}{ccc}
1 & 2 \\
1 & 0.035927 \\
2 & -0.002878 & 0.056344 \\
\end{array} \\
--
\]

\[ \text{miden r1,r2. maxl 12.} \]

\[ \text{TIME PERIOD ANALYZED . . . . . . . . . . . . 7 TO 296} \]

\[ \text{EFFECTIVE NUMBER OF OBSERVATIONS (NOBE)... 290} \]

<table>
<thead>
<tr>
<th>SERIES</th>
<th>NAME</th>
<th>MEAN</th>
<th>STD. ERROR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>R1</td>
<td>0.0000</td>
<td>0.1895</td>
</tr>
<tr>
<td>2</td>
<td>R2</td>
<td>0.0000</td>
<td>0.2374</td>
</tr>
</tbody>
</table>

\[ \text{NOTE: THE APPROX. STD. ERROR FOR THE ESTIMATED CORRELATIONS BELOW IS (1/NOBE**.5) = 0.05872} \]

\[ \text{SAMPLE CORRELATION MATRIX OF THE SERIES} \]

\[ \begin{array}{ccc}
1.00 \\
-0.06 & 1.00 \\
\end{array} \]

\[ \text{CROSS CORRELATION MATRICES IN TERMS OF +,-,.} \]

\[ \begin{array}{cccccccc}
& \text{Lags 1 Through 6} & & & & & & \\
& . . . . . . & - & . . . . . . & & & & \\
& . . . . . . . & . . . . . . . . & & & & \\
\end{array} \]

\[ \begin{array}{cccccccc}
& \text{Lags 7 Through 12} & & & & & & \\
& . . . . . . & . . . . . . & + & . . . . . . & & & & \\
\end{array} \]

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

\[ (1 - 1.98B + 1.38B^2 - .35B^2)x_t = a_{1t} \]  \hspace{1cm} (21)

\[ (1 - 1.53B + .58B^2 + 1.16B^3)y_t = 3.76 + (-.46B^3 + .23B^6)x_t + a_{2t} \]  \hspace{1cm} (22)

where \( a_{1t} \) and \( a_{2t} \) are essentially uncorrelated. Note that Eq. (21) is basically the AR(3) model for input gas rate. From Eq. (22), we have

\[ y_t = \frac{3.76}{1 - 1.53 + .58 + .16} + \frac{- .46B^3 + .23B^6}{1 - 1.53B + .58B^2 + 1.16B^3} x_t + \frac{1}{1 - 1.53B + .58B^2 + 1.16B^3} a_{2t}. \]

One can compare the above approximation model with transfer function model built in Lecture 1 using the estimated impulse 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 \( z_t = (z_{1t}, z_{2t}, z_{3t})' \). 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
```

<table>
<thead>
<tr>
<th>SERIES</th>
<th>NAME</th>
<th>MEAN</th>
<th>STD. ERROR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X</td>
<td>3.1063</td>
<td>1.0000</td>
</tr>
<tr>
<td>2</td>
<td>Y</td>
<td>2.8365</td>
<td>1.0000</td>
</tr>
<tr>
<td>3</td>
<td>Z</td>
<td>16.5473</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

NOTE: THE APPROX. STD. ERROR FOR THE ESTIMATED CORRELATIONS BELOW IS \( (1/\text{NOBE}^{**.5}) = 0.12700 \)

SAMPLE CORRELATION MATRIX OF THE SERIES

\[
\begin{pmatrix}
1.00 & 0.88 & -0.40 \\
0.88 & 1.00 & -0.24 \\
-0.40 & -0.24 & 1.00
\end{pmatrix}
\]

SUMMARIES OF CROSS CORRELATION MATRICES USING +,-,.

\[
\begin{array}{c|ccc}
1 & 2 & 3 \\
\hline
1 & ++++++++ & ++++++++ & ---------- \\
2 & ++++++++ & ++++++++ & ---------- \\
3 & ----------- & ----------- & ++++++++ \\
\end{array}
\]

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

\[
\begin{array}{cccccccc}
+ & + & - & + & + & - & + & + & - & + & + & - & + & + & -
\end{array}
\]
LAGS 7 THROUGH 12

+ + - + + - + + - + + - + + - + + - + + - + + -
- . + - . + - . + - . + - . + - . + - . + -

DETERMINANT OF S(0) = 0.146394E+00

AUTOREGRESSIVE FITTING ON LAG(S) 1

SUMMARIES OF CROSS CORRELATION MATRICES USING +,-,.

\[
\begin{array}{ccc}
1 & 2 & 3 \\
\cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots \\
+ & \cdots & \cdots \\
\cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots \\
\end{array}
\]

CROSS CORRELATION MATRICES IN TERMS OF +,-,.

LAGS 1 THROUGH 6

LAGS 7 THROUGH 12

====== STEPWISE AUTOREGRESSION SUMMARY ======

<table>
<thead>
<tr>
<th>LAG</th>
<th>RESIDUAL</th>
<th>EIGENVAL</th>
<th>CHI-SQ</th>
<th>SIGNIFICANCE</th>
<th>VARIANCES</th>
<th>OF SIGMA</th>
<th>TEST</th>
<th>AIC</th>
<th>OF PARTIAL AR COEFF.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.449E-01</td>
<td>.338E-01</td>
<td>292.93</td>
<td>-7.319</td>
<td>+</td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.900E-01</td>
<td>.788E-01</td>
<td></td>
<td></td>
<td>+</td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.164E+00</td>
<td>.186E+00</td>
<td></td>
<td></td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.849E-01</td>
<td>.775E-01</td>
<td></td>
<td></td>
<td>+</td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.123E+00</td>
<td>.141E+00</td>
<td></td>
<td></td>
<td>I</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>.819E-01</td>
<td>.759E-01</td>
<td></td>
<td></td>
<td>I</td>
<td>I</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

30
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 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**

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>DIFFERENCING</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Z</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PARAMETER</th>
<th>FACTOR</th>
<th>ORDER</th>
<th>CONSTRAINT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>C</td>
<td>0</td>
<td>CC</td>
</tr>
<tr>
<td>2</td>
<td>P1</td>
<td>1</td>
<td>CP1</td>
</tr>
<tr>
<td>3</td>
<td>T1</td>
<td>1</td>
<td>CT1</td>
</tr>
</tbody>
</table>

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

**SUMMARY FOR THE MULTIVARIATE ARMA MODEL**

<table>
<thead>
<tr>
<th>SERIES</th>
<th>NAME</th>
<th>MEAN</th>
<th>STD DEV</th>
<th>DIFFERENCE ORDER(S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X</td>
<td>3.1063</td>
<td>1.0000</td>
<td></td>
</tr>
</tbody>
</table>
NUMBER OF OBSERVATIONS = 62 (EFFECTIVE NUMBER = NOBE = 61)
MODEL SPECIFICATION WITH PARAMETER VALUES

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

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.957295</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.833973</td>
<td>0.952684</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-0.347628</td>
<td>-0.180279</td>
<td>0.945902</td>
</tr>
</tbody>
</table>

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  

32
.221  .168  .107
.277  .214  .129

-----------------------
ERROR COVARIANCE MATRIX
-----------------------

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>.036735</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.021639</td>
<td>.077807</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.013216</td>
<td>.021723</td>
<td>.130048</td>
</tr>
</tbody>
</table>

-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  -  -  

33
STANDARD ERRORS

-- .092 --
.168 -- --
.246 .150 .122

-----------------------
ERROR COVARIANCE MATRIX
-----------------------

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.043320</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.022633</td>
<td>0.079942</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.017799</td>
<td>0.021173</td>
<td>0.130828</td>
</tr>
</tbody>
</table>

-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

<table>
<thead>
<tr>
<th>SERIES</th>
<th>NAME</th>
<th>MEAN</th>
<th>STD DEV</th>
<th>DIFFERENCE</th>
<th>ORDER(S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X</td>
<td>3.1063</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Y</td>
<td>2.8365</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Z</td>
<td>16.5473</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MODEL SPECIFICATION WITH PARAMETER VALUES

<table>
<thead>
<tr>
<th>PARAMETER NUMBER</th>
<th>PARAMETER DESCRIPTION</th>
<th>PARAMETER VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CONSTANT( 1)</td>
<td>0.109976</td>
</tr>
<tr>
<td>2</td>
<td>CONSTANT( 2)</td>
<td>0.724382</td>
</tr>
<tr>
<td>3</td>
<td>CONSTANT( 3)</td>
<td>2.630637</td>
</tr>
<tr>
<td>4</td>
<td>AUTOREGRESSIVE ( 1, 1, 1)</td>
<td>0.922758</td>
</tr>
<tr>
<td><em>FIXED</em></td>
<td>AUTOREGRESSIVE ( 1, 1, 2)</td>
<td>0.000000</td>
</tr>
<tr>
<td>5</td>
<td>AUTOREGRESSIVE ( 1, 2, 2)</td>
<td>0.921368</td>
</tr>
<tr>
<td>6</td>
<td>AUTOREGRESSIVE ( 1, 3, 3)</td>
<td>0.838594</td>
</tr>
</tbody>
</table>
MOVING AVERAGE (1, 3, 1) -0.503301
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

<table>
<thead>
<tr>
<th></th>
<th>1.00</th>
<th>0.39</th>
<th>0.25</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.39</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td>0.22</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

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

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

Consequently, except for the contemporaneous dependence 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 three series are financial indices. This example demonstrates that VARMA modeling can show the relationship between variables.
Table 2: Estimation of VARMA(1,1) model CGK data set: exact likelihood method. The number in parentheses denotes standard error.

<table>
<thead>
<tr>
<th>c</th>
<th>(\phi_1)</th>
<th>(\theta_1)</th>
<th>(\Sigma)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(a) Initial estimation: Full model</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.11</td>
<td>.81</td>
<td>.16</td>
<td>-.06</td>
</tr>
<tr>
<td>(.64)</td>
<td>(.08)</td>
<td>(.08)</td>
<td>(.04)</td>
</tr>
<tr>
<td>1.86</td>
<td>-.08</td>
<td>.99</td>
<td>-.09</td>
</tr>
<tr>
<td>(.82)</td>
<td>(.10)</td>
<td>(.10)</td>
<td>(.05)</td>
</tr>
<tr>
<td>4.09</td>
<td>-.33</td>
<td>.30</td>
<td>.76</td>
</tr>
<tr>
<td>(1.5)</td>
<td>(.18)</td>
<td>(.17)</td>
<td>(.08)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(b) An intermediate model</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.11</td>
<td>.92</td>
<td>.07</td>
<td>0</td>
</tr>
<tr>
<td>(.08)</td>
<td>(.05)</td>
<td>(.05)</td>
<td>–</td>
</tr>
<tr>
<td>0.72</td>
<td>0</td>
<td>.92</td>
<td>-.03</td>
</tr>
<tr>
<td>(.64)</td>
<td>–</td>
<td>(.04)</td>
<td>(.04)</td>
</tr>
<tr>
<td>2.63</td>
<td>0</td>
<td>0</td>
<td>.84</td>
</tr>
<tr>
<td>(1.1)</td>
<td>–</td>
<td>–</td>
<td>(.07)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(c) A refined model</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.13</td>
<td>.98</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(.08)</td>
<td>(.03)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>0.24</td>
<td>0</td>
<td>.93</td>
<td>0</td>
</tr>
<tr>
<td>(.11)</td>
<td>–</td>
<td>(0.04)</td>
<td>–</td>
</tr>
<tr>
<td>2.54</td>
<td>0</td>
<td>0</td>
<td>.84</td>
</tr>
<tr>
<td>(1.1)</td>
<td>–</td>
<td>–</td>
<td>(.06)</td>
</tr>
</tbody>
</table>
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 $z_t$ be endogenous variables and $x_t$ be exogenous variables. A simultaneous equation model can be written as

$$\Omega z_t = \Phi_0 + \sum_{i=1}^{p} \Phi_i z_{t-i} + \sum_{i=0}^{r} V_i x_{t-i} + b_t,$$  
(23)

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

Given $z_t$ and $x_t$, the VARMA models can be generalized as

$$z_t = \phi_0 + \sum_{i=1}^{p} \phi_i z_{t-i} + \sum_{i=0}^{r} v_i x_{t-i} + a_t,$$  
(24)

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

Given Eq. (23), one can easily obtain Eq. (24) by pre-multiplying the model via $\Omega^{-1}$. The parameters are related as

$$\Omega^{-1} \Phi_i = \phi_i, \quad i = 0, \ldots, p; \quad \Omega^{-1} V_i = v_i, \quad i = 0, \ldots, r; \quad \Omega^{-1} \Sigma_b(\Omega^{-1})' = \Sigma.$$

On the other hand, obtaining Eq. (23) from Eq. (24) is more involved. We focus on the last element of $z_t$. From the simultaneous equation model, we have

$$z_{kt} + \sum_{i=1}^{k-1} \Omega_{ki} z_{it} = \Phi_{0,k} + \sum_{i=1}^{p} \sum_{j=1}^{k} \Phi_{i,kj} z_{j,t-i} + \sum_{i=0}^{r} \sum_{j=1}^{m} V_{i,kj} x_{j,t-i} + b_{kt},$$  
(25)

where $A_{t,ij}$ denotes the $(i,j)$th element of matrix $A_t$. Because $\Sigma$ of the VARMAX model is positive definite, by Cholesky Decomposition, we have $\Sigma = L G L'$, where $G$ is a diagonal matrix and $L$ is a lower triangular matrix with unit diagonal elements.

Define $b_t = L^{-1} a_t$. Then, $E(b_t) = 0$ and

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

Thus, the covariance matrix of $b_t$ is a diagonal matrix. Multiplying $L^{-1}$ from left to Eq. (24), we have

$$L^{-1} z_t = L^{-1} \phi_0 + \sum_{i=1}^{p} L^{-1} \phi_i z_{t-i} + \sum_{i=0}^{r} L^{-1} v_i x_{t-i} + L^{-1} a_t.$$
Since $L^{-1}$ is also lower triangular with unit diagonal elements, the $k$th row of $L^{-1}$ is in the form $(\Omega_{k1}, \cdots, \Omega_{k,k-1}, 1)$. Define

$$\Phi_i = L^{-1}\phi_i, \quad i = 0, \cdots, p; \quad V_i = L^{-1}v_i, \quad i = 0, \cdots, r.$$

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. (25). Thus, we can obtain a simultaneous equation model from a given vector ARMA model.

**Example.** Consider the bivariate model

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

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

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

It is easy to verify that $G = \text{diag}\{2,0.5\}$ 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

$$\begin{bmatrix} z_{2t} \\ z_{1t} \end{bmatrix} = \begin{bmatrix} 0.4 & 0.2 \\ 0.2 & 1.0 \end{bmatrix} + \begin{bmatrix} 1.1 & -0.6 \\ 0.3 & 0.2 \end{bmatrix} \begin{bmatrix} z_{2,t-1} \\ z_{1,t-1} \end{bmatrix} + \begin{bmatrix} a_{2t} \\ a_{1t} \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}.$$ 

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

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

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

39