Lecture 6: Structural Specification of VARMA Models

Structural specification seeks to find the underlying structure of a multivariate linear system so that a well-defined VARMA model can be identified. The specified model overcomes the difficulty of identifiability mentioned earlier and can reveal the hidden structure of the system. To a certain degree, structural specification is closely related to dimension reduction and can be regarded as a data mining method.

For simplicity, we focus on a \( k \)-dimensional zero-mean stationary multivariate linear system

\[
z_t = \sum_{i=0}^{\infty} \psi_i a_{t-i}
\]  

(4.1)

where \( \psi_0 = I \) and \( \{a_t\} \) is an iid sequence of random vectors with mean zero and positive-definite covariance matrix \( \Sigma \). In theory, the methods discussed continue to apply for unit-root nonstationary series, but the limiting distributions of the statistics involved will be different. The stationarity assumption implies that \( \sum_{i=0}^{\infty} \| \psi_i \|^2 < \infty \), where \( \|A\| \) is a matrix norm such as the largest singular value of the matrix \( A \). There are two methods available to perform structural specifications. The first method uses the idea of Kronecker index and the second method the concept of scalar component model. For further details, see Tsay (1991, Statistica Sinica, p.247-269) and the references therein. In real applications, both methods can be implemented using the canonical correlation analysis, which is a statistical method useful in multivariate analysis.

### 4.1 The Kronecker Index Approach

At the time index \( t \), define the past vector, \( P_{t-1} \), and future vector, \( F_t \), as

\[
P_{t-1} = (z'_{t-1}, z'_{t-2}, \cdots)' \quad F_t = (z'_t, z'_{t+1}, \cdots)'.
\]

In addition, define the Hankel matrix, \( H_\infty \), as

\[
H_\infty = E(F_t P'_{t-1}) = \begin{bmatrix}
\Gamma_1 & \Gamma_2 & \Gamma_3 & \cdots \\
\Gamma_2 & \Gamma_3 & \Gamma_4 & \cdots \\
\Gamma_3 & \Gamma_4 & \Gamma_5 & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{bmatrix},
\]  

(4.2)

where the subscript of \( H \) indicates that it is an \( \infty \)-dimensional matrix and \( \Gamma_j \) is the lag-\( j \) autocovariance matrix of \( z_t \). Note that a special feature of Hankel matrix is that, looking at blocks of non-overlapping \( k \) rows or \( k \) columns, the second block row is a subset of the first block row, the
second block column is a subset of the first block column, and so on. This is referred to as the Toeplitz form of the matrix. Also, for the linear system in Eq. (4.1),

$$\Gamma_t = \sum_{j=0}^\infty \psi_{t+j} \Sigma \psi'_j,$$

so that the Hankel matrix is uniquely defined.

**Lemma 1.** For the linear vector process $z_t$ in Eq. (4.1), $\text{Rank}(H_\infty) = m$ is finite if and only if $z_t$ follows a VARMA model.

**Proof:** If $z_t$ follows the VARMA($p,q$) model

$$\phi(B)z_t = \theta(B)\alpha_t,$$

then we have $\phi(B)\Gamma_j = 0$ for $j > q$, where the back-shift operator operates on the subscript $j$ of $\Gamma_j$, i.e. $B\Gamma_j = \Gamma_{j-1}$. These are the moment equations of $z_t$. Since $q$ is finite, the above moment equations implies that the rank of $H_\infty$ is finite. On the other hand, if $\text{Rank}(H_\infty) = m < \infty$, then there exists non-negative integer $q$ such that the $(q+1)$th block row $[\Gamma_{q+1}, \Gamma_{q+2}, \cdots]$ is a linear combination of $\{[\Gamma_i, \Gamma_{i+1}, \cdots]\}_{i=1}^q$. In other words, there exists matrices $\{\phi_i\}_{i=1}^q$ such that

$$\Gamma_j = \sum_{i=1}^q \phi_i \Gamma_{j-i}, \quad j > q,$$

where the special feature of Hankel matrix is used. This implies that $z_{t+\ell} - \sum_{i=1}^q \phi_i z_{t+\ell-i}$ is uncorrelated with the past vector $P_{t-1}$ for $\ell \geq q$. Consequently, $z_t$ should follow a VARMA($q,q$) model. The order $(q,q)$ is the maximum order. As it will be seen later, we can construct a well-defined lower-order VARMA model under the assumption that $\text{Rank}(H_\infty)$ is finite. $\qedsymbol$

From now on, we shall assume that $\text{Rank}(H_\infty)$ is finite so that $z_t$ is a VARMA process. Using the Toeplitz form of the Hankel matrix in Eq. (4.2), we can examine its row dependence by starting with the first row. That is, we can search a basis that expands the space generated by the rows of $H_\infty$ starting with the first row. Denote the $[(i-1)k + j]$th row of $H_\infty$ by $h(i,j)$, where $j = 1, \cdots, k$ and $i = 1, 2, \cdots$. From the definition, $h(i,j) = E(z_{j,t+i-1}P'_{t-1})$, measuring the linear dependence of $z_{j,t+i-1}$ on the past vector $P_{t-1}$. We say that $h(i,j)$ is a predecessor of $h(u,v)$ if $(i-1)k + j < (u-1)k + v$. Using the Toeplitz form, it is easy to obtain the following result.

**Lemma 2.** For the Hankel matrix in Eq. (4.2), if $h(i,j)$ is a linear combination of its predecessors $\{h(i_1,j_1), \cdots, h(i_s,j_s)\}$, then $h(i+1,j)$ is a linear combination of $\{h(i_1+1,j_1), \cdots, h(i_s+1,j_s)\}$.

**Definition:** For the $j$th component $z_{jt}$, the *Kronecker index* $k_j$ is defined as the smallest non-negative integer $i$ such that $h(i+1,j)$ is linearly dependent of its predecessors.

For instance, if $z_{1t} = a_{1t}$ is a white noise component, then the first row $h(1,1) = 0$, a zero row vector, so that $k_1 = 0$. As another example, consider the bivariate VAR(1) model

$$z_t - \phi z_{t-1} = \alpha_t,$$

with

$$\phi = \begin{bmatrix} 0.2 & 0.3 \\ -0.6 & 1.1 \end{bmatrix}.$$
Since each component \(z_{it}\) (for \(i = 1\) and \(2\)) follows a univariate ARMA(2,1) model with AR polynomial \((1 - 3B + 4B^2)\), it has non-zero autocorrelations. The first two rows of \(H_\infty\) are non-zero. On the other hand, from the moment equations \(\Gamma_j - \phi \Gamma_{j-1} = 0\) for \(j > 0\), we have \(\Gamma_j = \phi^{j-1} \Gamma_1\) for \(j > 1\). Consequently, \(\text{Rank}(H_\infty) = 2\), and the first two rows \(h(1,1)\) and \(h(1,2)\) span the row-space of \(H_\infty\). Therefore, the 3rd and 4th rows \(h(2,1)\) and \(h(2,2)\) are linearly dependent of their predecessors \(\{h(1,1), h(1,2)\}\), and we have \(k_1 = 1\) and \(k_2 = 1\).

To aid further discussion, it is helpful to consider the Hankel matrix \(H_\infty\) as follows:

<table>
<thead>
<tr>
<th>Block</th>
<th>Component</th>
<th>(P'<em>{t-1} = [z'</em>{t-1}, z'_{t-2}, \ldots])</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(z_{1t})</td>
<td>(h(1,1))</td>
</tr>
<tr>
<td></td>
<td>(z_{2t})</td>
<td>(h(1,2))</td>
</tr>
<tr>
<td></td>
<td>(\vdots)</td>
<td>(\vdots)</td>
</tr>
<tr>
<td></td>
<td>(z_{kt})</td>
<td>(h(1,k))</td>
</tr>
<tr>
<td>2</td>
<td>(z_{1,t+1})</td>
<td>(h(2,1))</td>
</tr>
<tr>
<td></td>
<td>(z_{2,t+1})</td>
<td>(h(2,2))</td>
</tr>
<tr>
<td></td>
<td>(\vdots)</td>
<td>(\vdots)</td>
</tr>
<tr>
<td></td>
<td>(z_{k,t+1})</td>
<td>(h(2,k))</td>
</tr>
<tr>
<td>(\vdots)</td>
<td>(\vdots)</td>
<td>(\vdots)</td>
</tr>
<tr>
<td>(k_j)</td>
<td>(z_{1,t+k_j-1})</td>
<td>(h(k_j,1))</td>
</tr>
<tr>
<td></td>
<td>(z_{2,t+k_j-1})</td>
<td>(h(k_j,2))</td>
</tr>
<tr>
<td></td>
<td>(\vdots)</td>
<td>(\vdots)</td>
</tr>
<tr>
<td></td>
<td>(z_{k,t+k_j-1})</td>
<td>(h(k_j,k))</td>
</tr>
<tr>
<td>(k_j+1)</td>
<td>(z_{1,t+k_j})</td>
<td>(h(k_j+1,1))</td>
</tr>
<tr>
<td></td>
<td>(\vdots)</td>
<td>(\vdots)</td>
</tr>
<tr>
<td></td>
<td>(z_{j-1,t+k_j})</td>
<td>(h(k_j+1,j-1))</td>
</tr>
<tr>
<td></td>
<td>(z_{j,t+k_j})</td>
<td>(h(k_j+1,j))</td>
</tr>
</tbody>
</table>

From the definition, a given Kronecker index \(k_j\) says that \(h(k_j + 1, j)\) is linearly dependent of its predecessors. From the prior structure of the Hankel matrix \(H_\infty\), there exist constants \(\{\alpha_{u,i,j}^*|u = 1, \ldots, k_j; i = 1, \ldots, k\}\) if \(k_j > 0\) and \(\{\alpha_{k_j+1,i,j}^*|i = 1, \ldots, j-1\}\) if \(j > 1\) such that

\[
h(k_j + 1, j) = \sum_{i=1}^{j-1} \alpha_{k_j+1,i,j}^* h(k_j + 1, i) + \sum_{u=1}^{k_j} \sum_{i=1}^{k} \alpha_{u,i,j}^* h(u, i), \tag{4.3}
\]

where it is understood that a summation is zero if its upper limit is smaller than its lower limit. In Eq. (4.3), the first term on the right hand side is summing over component \(i\) of \(z_{t+k_j}\) in \(F_t\) and the second term summing over \(z_{t+\ell}\) in \(F_t\) with \(0 \leq \ell < k_j\). In other words, the first term is summing over the \((k_j + 1)\)th block row whereas the second term over the block rows \(1, \ldots, k_j\).

**Definition:** The collection \(\{k_j\}_{j=1}^k\) is called the set of Kronecker indices of \(z_t\). The sum \(m = \sum_{j=1}^{k} k_j\) is called the McMillan degree of \(z_t\).
Next, consider the main question of how to obtain a well-defined VARMA model from the Kronecker indices of a given Hankel matrix $H_{\infty}$. To this end, we can simplify Eq. (4.3) by considering $\{k_i\}_{i=1}^k$ simultaneously. Rearranging terms in Eq. (4.3) according to the second argument of $h(i,j)$, that is, according to the components of $z_t$, we have

$$h(k_j + 1, j) = \sum_{i=1}^{j-1} \sum_{u=1}^{k+1} \alpha_{u,i,j}^* h(u, i) + \sum_{i=j}^{k} \sum_{u=1}^{k} \alpha_{u,i,j}^* h(u, i).$$

(4.4)

However, for each component $i$, $h(u, i)$ is a linear combination of its own predecessors if $u > k_i$. Therefore, Eq. (4.4) can be simplified as

$$h(k_j + 1, j) = \sum_{i=1}^{j-1} \sum_{u=1}^{\min(k_j+1, k_i)} \alpha_{u,i,j} h(u, i) + \sum_{i=j}^{k} \min(k_j, k_i) \sum_{u=1}^{\min(k_j, k_i)} \alpha_{u,i,j} h(u, i), \quad j = 1, \ldots, k,$$

(4.5)

where $\alpha_{u,i,j}$ are real numbers and are linear functions of $\{\alpha_{\ell,r,s}^*\}$. These $k$ linear combinations provide a way to specify a VARMA model.

Notice that the number of $\alpha_{u,i,j}$ coefficients in Eq. (4.5) is

$$\delta_j = \sum_{i=1}^{j-1} \min(k_j+1, k_i) + k_j + \sum_{i=j+1}^{k} \min(k_j, k_i).$$

(4.6)

Using Lemma 2 and considering Eq. (4.5) jointly for $j = 1, \ldots, k$, we obtain the following results:

1. The collection of rows

$${\mathcal{B}} = \{h(1,1), \ldots, h(k_1, 1); h(1,2), \ldots, h(k_2, 2); \ldots; h(1,k), \ldots, h(k,k)\}$$

forms a basis for the row-space of $H_{\infty}$, because all other rows are linear combinations of elements in $\mathcal{B}$. Note that if $k_j = 0$ then no row in the form of $h(\ell, j)$ belongs to $\mathcal{B}$.

2. $\text{Rank}(H_{\infty}) = \sum_{j=1}^{k} k_j$. That is, the sum of Kronecker indices is the rank of $H_{\infty}$. Thus, for the linear system $z_t$, the McMillan degree is the rank of its Hankel matrix.

3. From the definition of Kronecker index, the basis $\mathcal{B}$ consists of the first $m$ linearly independent rows $H_{\infty}$, starting from the top of the matrix.

Remark: From the definition, the Kronecker index $k_j$ depend on the ordering of the components of $z_t$. However, the rank of the Hankel matrix $H_{\infty}$ is invariant to the ordering of components of $z_t$. Thus, the set of Kronecker indices $\{k_j\}_{j=1}^k$ is invariant to the ordering of components of $z_t$.

4.1.1 A predictive interpretation

The linear combination in Eq. (4.5) has a nice predictive interpretation. Let $\alpha_j$ be an infinitely dimensional, real-valued vector with $\alpha_{u,i}^{(j)}$ as its $[(u-1)k + i]$th element. For each component $z_{jt}$, we can use Eq. (4.5), which is derived from Kronecker indices, to construct a vector $\alpha_j$ as follows:

1. Let $\alpha_{k_j+1,j}^{(j)} = 1$. That is, the $(k_j \times k + j)$th element of $\alpha_j$ is one.
2. For each $h(u, i)$ appearing in the right hand side of Eq. (4.5), let $\alpha_{u,i}^{(j)} = -\alpha_{u,i,j}$. By Eq. (4.5), we have
\[
\alpha_j^t H_\infty = 0. \tag{4.7}
\]

To see the implication of this result, we recall that $H_\infty = E(F_t P_{t-1}^\prime)$ and that the $(k_j \times k + j)$th element of $F_t$ is $z_{j,t+k_j}$. Define $w_{j,t+k_j} = \alpha_j^t F_t$. Then, from Eqs. (4.7) and (4.2), $w_{j,t+k_j}$ is uncorrelated with the past vector $P_{t-1}$. Thus, corresponding to each Kronecker index $k_j$, there is a linear combination of the future vector $F_t$ that is uncorrelated with the past $P_{t-1}$. Furthermore, using the MA representation in Eq. (4.1), $w_{j,t+k_j}$ is a linear function of $\{a_{\ell} | \ell \leq t + k_j\}$ whereas $P_{t-1}$ is a linear function of $\{a_{\ell} | t \leq \ell \leq t + k_j\}$. It is then easily seen that, from the zero correlation between $w_{j,t+k_j}$ and $P_{t-1}$, $w_{j,t+k_j}$ must be a linear function of $\{a_{\ell} | t \leq \ell \leq t + k_j\}$. Thus, we have
\[
w_{j,t+k_j} = \sum_{i=0}^{k_j} u_i^{(j)} a_{t+k_j-i}, \tag{4.8}
\]
where $u_i^{(j)}$'s are $k$-dimensional row vectors such that
\[
u_0^{(j)} = [\alpha_{k_j+1,1}^{(j)}, \ldots, \alpha_{k_j+1,j-1}^{(j)}, 1, 0, \ldots, 0]
\]
with 1 being in the $j$th position and it is understood that $\alpha_{k_j+1,i}^{(j)} = 0$ if $k_j < k_j + 1$ and $i < j$. In general, $u_i^{(j)}$ are functions of elements of $\psi_i$'s and non-zero elements of $\alpha_j$. Eq. (4.8) says that the scalar process $w_{j,t+k_j}$ is at most an MA($k_j$) process.

Next, from the definitions of $\alpha_j$ and $w_{j,t+k_j}$, we have
\[
w_{j,t+k_j} = z_{j,t+k_j} + \sum_{i=1}^{j-1} \sum_{u=1}^{\min(k_j+1,k_i)} \alpha_{u,i}^{(j)} z_{i,t+u-1} + \sum_{i=j}^{k} \sum_{u=1}^{\min(k_j,k_i)} \alpha_{u,i}^{(j)} z_{i,t+u-1}. \tag{4.9}
\]
Combining with Eq. (4.8) and noting that $\alpha_{k_j+1,i}^{(j)} = 0$ if $k_j < k_j + 1$ and $i < j$, we obtain
\[
z_{j,t+k_j} = a_{j,t+k_j} + \sum_{i=j}^{k} \alpha_{i,j}^{(j)} a_{i,j+k_j} + \sum_{i=1}^{k} u_i^{(j)} a_{i+k_j-i}. \tag{4.9}
\]
Taking the conditional expectation of Eq. (4.9) based on $P_{t-1}$, we have
\[
z_{j,t+k_j | t-1} = E(a_{j,t+k_j} | P_{t-1}) + \sum_{i=1}^{j-1} \sum_{u=1}^{\min(k_j+1,k_i)} \alpha_{u,i}^{(j)} E(z_{i,t+u-1} | P_{t-1}) \\
+ \sum_{i=j}^{k} \sum_{u=1}^{\min(k_j,k_i)} \alpha_{u,i}^{(j)} E(z_{i,t+u-1} | P_{t-1}) = 0. \tag{4.10}
\]
where \( z_{i,t+\ell} = E(z_{i,t+\ell} | P_{t-1}) \) is the conditional expectation of \( z_{i,t+\ell} \) given \( P_{t-1} \). Let \( F_{t|t-1} = E(F_{t} | P_{t-1}) \). The prior equation shows that, for each Kronecker index \( k_j \), there exists a linear relationship among the forecasts in \( F_{t|t-1} \). Since \( k_j \) is the smallest non-negative integer for Eq. (4.10) to hold, one can interpret \( k_j \) as the number of forecasts \( z_{j,t-1}, \ldots, z_{j,t+k_j-1} | t-1 \) needed to compute the forecasts \( z_{j,t+\ell} \) for any forecast horizon \( \ell \). Obviously, to compute \( z_{j,t+\ell} | t-1 \), one also needs forecasts \( z_{i,t+u} | t-1 \) with \( i \neq j \). However, these quantities are taken care of by the other Kronecker indices \( k_i \) with \( i \neq j \). In view of this, the McMillan degree \( m \) is the minimum number of quantities needed to compute all elements in \( F_{t|t-1} \). The Kronecker index \( k_j \) is the minimum number of those quantities that the component \( z_{j,t} \) contributes.

4.1.2 A VARMA specification

By the stationarity of \( z_t \), Eq. (4.9) can be rewritten as

\[
z_{j,t} = \sum_{i=1}^{j-1} \sum_{u=1}^{\min(k_j+1,k_i)} \alpha_{u,i}^{(j)} z_{i,t+u-1-k_j} + \sum_{i=j}^{k \min(k_j,k_i)} \alpha_{u,i}^{(j)} z_{i,t+u-1-k_j} + \sum_{i=j}^{k} \sum_{u=1}^{\min(k_j,k_i)} \alpha_{u,i}^{(j)} z_{i,t+u-1-k_j}
\]

\[
= a_{j,t} + \sum_{i<j,k_i<k_j+1} \alpha_{k_j+1,i}^{(j)} a_{i,t} + \sum_{i=1}^{k_j} \sum_{\ell=1}^{k_j} u_i^{(j)} a_{t-\ell}.
\]

(4.11)

Note that in Eq. (4.11), the number of coefficients \( \alpha_{u,i}^{(j)} \) on the left hand side is \( \delta_j \) given in Eq. (4.6) and the number of elements of \( u_i^{(j)} \)'s in the right hand side is \( k_j \times k \).

By considering Eq. (4.11) jointly for \( j = 1, \cdots, k \), we obtain a VARMA model for the \( z_t \) process as

\[
\Xi_0 z_t + \sum_{i=1}^{p} \Xi_i z_{t-i} = \Xi_0 a_t + \sum_{i=1}^{p} \Omega_i a_{t-i},
\]

where \( p = \max_j \{k_j\} \), \( \Xi_0 \) is a lower triangular matrix with unit diagonal elements, and its \( (j,i) \)th element is \( \alpha_{k_j+1,i}^{(j)} \) for \( i < j \) so that the \( (j,i) \)th element is unknown only if \( k_j + 1 \leq k_i \), and the coefficient matrices \( \Xi_i \) and \( \Omega_i \) are given in Eq. (4.11) for \( i = 1, \cdots, p \). More specifically, we have the following:

1. For \( \Omega_i \) matrix with \( i > 0 \): (a) the \( j \)th row is zero if \( k_j < i \leq p \); (b) all the other rows are unknown and must be estimated.

2. For \( \Xi_i \) matrix with \( i > 0 \): (a) the \( j \)th row is zero if \( k_j < i \leq p \); (b) the \( (j,j) \)th element is unknown if \( i \leq k_j \); (c) the \( (j,\ell) \)th element with \( j \neq \ell \) is unknown only if \( k_i + i > k_j \).

Eq. (4.12) gives rise to a VARMA representation for \( z_t \), the model for \( z_{jt} \) contains \( \delta_j \) unknown parameters in the AR polynomials and \( k \times k_j \) unknown parameters in the MA polynomials, where \( \delta_j \) is defined in Eq. (4.6). In summary, for the linear stationary process \( z_t \) in Eq. (4.1) with Kronecker indices \( \{k_j | j = 1, \cdots, k\} \) such that \( m = \sum_{j=1}^{k} k_j < \infty \), one can specify a VARMA representation to describe the process. Such a representation is given by Eq. (4.12) and contains

\[
N = m(1 + k) + \sum_{j=1}^{k} \sum_{i=1}^{j-1} \min(k_j+1,k_i) + \sum_{i=j+1}^{k} \min(k_j,k_i)
\]

(4.13)

unknown parameters in the AR and MA polynomials, where it is understood that the summation is zero if its lower limit is greater than its upper limit.
4.1.3 An illustrate example

To better understand the results of the preceding sections, we consider a simple example. Suppose that \( z_t \) is 3-dimensional and the Kronecker indices are \( \{k_1 = 3, k_2 = 1, k_3 = 2\} \). Here the basis \( \mathcal{B} \) for the row-space of the Hankel matrix \( H_\infty \) of \( z_t \) is

\[
\mathcal{B} = \{h(1,1), h(2,1), h(3,1); h(1,2); h(1,3), h(2,3)\}
\]

\[
= \{h(1,1), h(1,2), h(1,3); h(2,1), h(2,2), h(3,1)\},
\]

where the second representation is based on the row number of the Hankel matrix. Therefore, we have the following results:

1. The first linearly dependent row of \( H_\infty \) is \( h(2,2) \) that provides a model for \( z_{2t} \) as

\[
z_{2t} + a_{2,1} z_{1t} + \alpha_{1,3} z_{3,t-1} + \alpha_{1,2} z_{2,t-1} + \alpha_{1,1} z_{1,t-1} = a_{2,t} + a_{1,2} a_{t-1} + a_{1,1} a_{t-2},
\]

where, as defined before, \( u_1^{(2)} \) is a \( k \)-dimensional row vector. Based on Lemma 2, the rows \( h(\ell, 2) \) with \( \ell \geq 2 \) are all linearly dependent of their predecessors so that they can be removed from \( H_\infty \) for any further consideration.

2. The next linearly dependent row is \( h(3,3) \) that gives a model for \( z_{3t} \) as

\[
z_{3t} + a_{3,1} z_{1t-1} + a_{2,3} z_{3,t-2} + a_{1,2} z_{2,t-2} + a_{1,1} z_{1,t-2} = a_{3,t} + a_{3,1} a_{t-1} + u_{1}^{(2)} a_{t-1} + u_{2}^{(2)} a_{t-2}.
\]

Again, by Lemma 2, we can remove all rows \( h(\ell, 3) \) with \( \ell \geq 3 \) from \( H_\infty \) for any further consideration.

3. The next linearly dependent row is \( h(4,1) \) that shows a model for \( z_{1t} \) as

\[
z_{1t} + a_{3,1} z_{1,t-1} + a_{2,3} z_{3,t-2} + a_{1,2} z_{2,t-2} + a_{1,1} z_{1,t-3} = a_{1,t} + u_{1}^{(1)} a_{t-1} + u_{2}^{(1)} a_{t-2} + u_{3}^{(1)} a_{t-3}.
\]

The preceding three equations can be seen from the sketch of the Hankel matrix shown in Table 4.1. Putting the three equations together, we have a VARMA representation

\[
\begin{bmatrix}
1 & 0 & 0 \\
\alpha_{2,1} & 1 & 0 \\
\alpha_{3,1} & 0 & 1
\end{bmatrix}
\begin{bmatrix}
z_t \\
z_{t-1} \\
z_{t-2}
\end{bmatrix} =
\begin{bmatrix}
\alpha_{3,1} & 0 & 0 \\
\alpha_{1,2} & \alpha_{1,3} & 0 \\
\alpha_{2,1} & 0 & \alpha_{2,3}
\end{bmatrix}
\begin{bmatrix}
z_{t+1} \\
z_{t} \\
z_{t-1}
\end{bmatrix} +
\begin{bmatrix}
\alpha_{1,1} & 0 & 0 \\
0 & \alpha_{2,3} & 0 \\
0 & \alpha_{3,1} & \alpha_{3,2}
\end{bmatrix}
\begin{bmatrix}
z_{t-2} \\
z_{t-3} \\
z_{t-4}
\end{bmatrix}
\]

where \( u_{i,j}^{(v)} \) is the \( j \)th element of \( u_i^{(v)} \). In practice, the notation is not important and we can summarize the specified model as

\[
\begin{bmatrix}
1 & 0 & 0 \\
1 & 0 & 0 \\
0 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
z_t \\
z_{t+1} \\
z_{t+2}
\end{bmatrix} =
\begin{bmatrix}
X & 0 & 0 \\
X & X & X \\
X & 0 & X
\end{bmatrix}
\begin{bmatrix}
z_{t-1} \\
z_{t} \\
z_{t+1}
\end{bmatrix} +
\begin{bmatrix}
X & 0 & 0 \\
0 & 0 & 0 \\
X & X & X
\end{bmatrix}
\begin{bmatrix}
z_{t-2} \\
z_{t-3} \\
z_{t-4}
\end{bmatrix}
\]

7
Table 4.1: A Sketch of Model Specification Using Kronecker Indices, where the indices are \((k_1 = 3, k_2 = 1, k_3 = 2)\).

(a) Find the first linearly dependent row

<table>
<thead>
<tr>
<th>Block 1</th>
<th>(h(1,1))</th>
<th>(h(1,2))</th>
<th>(h(1,3))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block 2</td>
<td>(h(2,1))</td>
<td>(h(2,2))</td>
<td>linearly dependent ((\Rightarrow) model for (z_{2t}))</td>
</tr>
</tbody>
</table>

(b) Find the next linearly dependent row

<table>
<thead>
<tr>
<th>Block 1</th>
<th>(h(1,1))</th>
<th>(h(1,2))</th>
<th>(h(1,3))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block 2</td>
<td>(h(2,1))</td>
<td>(h(2,3))</td>
<td></td>
</tr>
<tr>
<td>Block 3</td>
<td>(h(3,1))</td>
<td>(h(3,3))</td>
<td>linearly dependent ((\Rightarrow) model for (z_{3t}))</td>
</tr>
</tbody>
</table>

(c) Find the next linearly dependent row

<table>
<thead>
<tr>
<th>Block 1</th>
<th>(h(1,1))</th>
<th>(h(1,2))</th>
<th>(h(1,3))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block 2</td>
<td>(h(2,1))</td>
<td>(h(2,3))</td>
<td></td>
</tr>
<tr>
<td>Block 3</td>
<td>(h(3,1))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Block 4</td>
<td>(h(4,1))</td>
<td>linearly dependent ((\Rightarrow) model for (z_{4t}))</td>
<td></td>
</tr>
</tbody>
</table>
\[
\begin{bmatrix}
1 & 0 & 0 \\
X & 1 & 0 \\
X & X & 0 \\
X & X & X \\
\end{bmatrix} a_t + \begin{bmatrix}
X & X & X \\
X & X & 0 \\
X & X & 0 \\
\end{bmatrix} a_{t-1} + \begin{bmatrix}
X & X & X \\
0 & 0 & 0 \\
X & X & 0 \\
\end{bmatrix} a_{t-2} + \begin{bmatrix}
X & X & X \\
0 & 0 & 0 \\
0 & 0 & 0 \\
\end{bmatrix} a_{t-3},
\]
where “X” denotes an unknown parameter that requires estimation. The total number of unknown parameters is \( N = 6(1 + 3) + 10 = 34 \).

### 4.1.4 The Echelon form

The model representation of Eq. (4.12) is a canonical form for the \( z_t \) process. It is referred to as a reversed Echelon form and has some nice properties that we discuss in the subsection.

**Degree of individual polynomial.** Let \( A_{jv}(B) \) be the \((j, v)\)th element of the matrix polynomial \( A(B) \). Let \( \text{deg}[A_{jv}(B)] \) be the degree of polynomial \( A_{jv}(B) \). Then, the degree of each polynomial of \( \Xi(B) = \Xi_0 + \sum_{i=1}^{k} \Xi_i B^i \) of Eq. (4.12) is \( \text{deg}[\Xi_{jv}(B)] = k_j \) for all \( v = 1, \ldots, k \). In other words, the Kronecker index \( k_j \) is the degree of all the polynomials in the \( j \)th row of \( \Xi(B) \). The same result holds for the individual polynomials in \( \Omega(B) \) of Eq. (4.12). In fact, \( k_j \) is the maximum order of \( \Xi_{jv}(B) \) and \( \Omega_{jv}(B) \). The actual order might be smaller after estimation or further analysis.

**Number of unknown coefficients of the individual polynomial.** Let \( n_{jv} \) be the number of unknown coefficients of \( \Xi_{jv}(B) \) in \( \Xi(B) \) of Eq. (4.12). Then, from the structure of \( \Xi(B) \), we have

\[
n_{jv} = \begin{cases} 
\min(k_j, k_v) & \text{if } j \leq v \\
\min(k_j + 1, k_v) & \text{if } j > v.
\end{cases} \tag{4.14}
\]

Similarly, let \( m_{jv} \) be the number of unknown coefficients of \( \Omega_{jv}(B) \) in \( \Omega(B) \). Then, we have

\[
m_{jv} = \begin{cases} 
k_j & \text{if } j \leq v \text{ or } (j > v \text{ and } k_j \geq k_v) \\
k_j + 1 & \text{if } j > v \text{ and } k_j < k_v.
\end{cases} \tag{4.15}
\]

Both counts include the lower triangular elements in \( \Xi_0 \), if they exist.

**Form of the individual polynomial.** Denote by \( A_{jv}^{(i)} \) the \((j, v)\)th element of the matrix \( A \). Using the degree and the equation form discussed earlier, one can easily specify the form of each individual polynomial in \( \Xi(B) \) of Eq. (4.12). Specifically, we have

\[
\Xi_{jj}(B) = 1 + \sum_{i=1}^{k_j} \Xi_{jj}^{(i)} B^i, \quad j = 1, \ldots, k \tag{4.16}
\]

\[
\Xi_{jv}(B) = \sum_{i=k_j+1}^{k_j+n_{jv}} \Xi_{jv}^{(i)} B^i, \quad j \neq v, \tag{4.17}
\]

where \( n_{jv} \) is defined in Eq. (4.14). For the polynomial in \( \Omega(B) \), the result is

\[
\Omega_{jj}(B) = 1 + \sum_{i=1}^{k_j} \Omega_{jj}^{(i)} B^i, \quad j = 1, \ldots, k \tag{4.18}
\]

\[
\Omega_{jv}(B) = \sum_{i=k_j+1}^{k_j+m_{jv}} \Omega_{jv}^{(i)} B^i, \quad j \neq v, \tag{4.19}
\]

9
where \( m_{jv} \) is defined in Eq. (4.15).

The preceding results show that, for a \( k \)-dimensional linear process \( z_t \) of Eq. (4.1), the Kronecker indices \( \{ k_j \mid j = 1, \ldots, k \} \) specify a VARMA representation (4.12) for \( z_t \). This VARMA specification is well-defined in the sense that (a) all the unknown parameters in the AR and MA matrix polynomials are identified and (b) each individual polynomial is specifically given. In the literature, such a VARMA representation is called a reversed Echelon form (Hannan and Deistler, 1988) and has the following nice properties:

**Theorem 1.** Suppose that \( z_t \) is a \( k \)-dimensional stationary time series of Eq. (4.1) with Kronecker indices \( \{ k_j \mid j = 1, \ldots, k \} \) such that \( m = \sum_{j=1}^{k} k_j < \infty \). Then, \( z_t \) follows the VARMA model in Eq. (4.12) with \( \Xi(B) \) and \( \Omega(B) \) specified by Eqs. (4.14)-(4.19). Furthermore, \( \Xi(B) \) and \( \Omega(B) \) are left coprime, and \( \text{deg}[\|\Xi(B)\|] + \text{deg}[\|\Omega(B)\|] \leq 2m \).

### 4.1.5 The example continued

For the 3-dimensional example of Subsection 4.1.3, the number of unknown coefficients in the individual polynomials is given by

\[
[n_{jv}] = \begin{bmatrix} 3 & 1 & 2 \\ 2 & 1 & 1 \\ 3 & 1 & 2 \end{bmatrix} \quad \text{and} \quad [m_{jv}] = \begin{bmatrix} 3 & 3 & 3 \\ 2 & 1 & 1 \\ 3 & 2 & 2 \end{bmatrix}.
\]

Since \( \Xi_0 = \Omega_0 \), the total number of unknown coefficients in the AR and MA polynomials is different from the sum of all \( n_{jv} \) and \( m_{jv} \), as the latter counts the unknown coefficients in \( \Xi_0 \) twice.

### 4.2 The Scalar-Component Approach

#### 4.2.1 Scalar component models

Scalar component models (SCM), developed by Tiao and Tsay (1989, JRSS Series B), generalize the concept of model-structure of a VARMA model to search for simplifying structure of the data. Consider the VARMA\((p, q)\) model

\[
\phi(B)z_t = \theta(B)a_t \tag{4.20}
\]

where \( z_t \) and \( a_t \) are defined in Eq. (4.1), and \( \phi(B) = I - \sum_{i=1}^{p} \phi_i B^i \) and \( \theta(B) = I - \sum_{i=1}^{q} \theta_i B^i \) are matrix polynomials of finite order \( p \) and \( q \), respectively. Let \( A^{(i)} \) be the \( i \)th row of the matrix \( A \). Then, the equation for the \( i \)th component \( z_{it} \) is

\[
I^{(i)} z_t - \sum_{j=1}^{p} \phi_j^{(i)} z_{t-j} = I^{(i)} - \sum_{j=1}^{q} \theta_j^{(i)} a_{t-j}.
\]

One way to interpret the prior equation is as follows: Given the vector \( v_0^{(i)} = I^{(i)} \), there exist \( p \) \( k \)-dimensional row vectors \( \{v_j^{(i)}\}_{j=1}^{p} \) with \( v_j^{(i)} = -\phi_j^{(i)} \) such that the linear combination

\[
w_{it} = v_0^{(i)} z_t + \sum_{j=1}^{p} v_j^{(i)} z_{t-j},
\]

10
is uncorrelated with \( \mathbf{a}_{t-\ell} \) for \( \ell > q \), because we also have
\[
    w_{it} = v_0^{(i)} a_t - \sum_{j=1}^{q} \theta_j^{(i)} a_{t-j},
\]
which is a linear combination of \( \{a_{t-j}\}_{j=0}^{q} \). In other words, the process \( w_{it} \), which is a linear combination of \( \{z_{t-j}\}_{j=0}^{p} \), has a finite memory in the sense that it is uncorrelated with \( \mathbf{a}_{t-\ell} \) for \( \ell > q \). For a \( k \)-dimensional process \( \mathbf{z}_t \), there are \( k \) such linearly independent processes of order \((p, q)\). However, we would like to keep the order \((p, q)\) of each linear combination as low as possible.

**Definition.** Suppose that \( \mathbf{z}_t \) is a stationary linear vector process of Eq. (4.1). A non-zero linear combination of \( \mathbf{z}_t \), denoted by \( y_t = v_0 \mathbf{z}_t \), is a scalar component of order \((r, s)\) if there exist \( r \) \( k \)-dimensional row vectors \( \{v_i\}_{i=1}^{r} \) such that (a) \( v_r \neq \mathbf{0} \) if \( r > 0 \), and (b) the scalar process
\[
    w_t = y_t + \sum_{i=1}^{r} v_i \mathbf{z}_{t-i} \quad \text{satisfies} \quad E(\mathbf{a}_{t-\ell} w_t) \quad \begin{cases} = \mathbf{0} & \text{if } \ell > s \\ \neq \mathbf{0} & \text{if } \ell = s. \end{cases}
\]

From the definition, \( y_t = v_0 \mathbf{z}_t \) is a SCM of \( \mathbf{z}_t \) if the scalar process \( w_t \) is uncorrelated with the past vector \( P_{t-\ell} \) for \( \ell > s \), but correlated with \( P_{t-s} \). The requirements of \( v_r \neq \mathbf{0} \), if \( r > 0 \), and \( E(\mathbf{a}_{t-s} w_t) \neq \mathbf{0} \) are used to reduce the order \((r, s)\).

Using Eq. (4.1) for \( \mathbf{z}_{t-i} \) and collecting the coefficient vectors of \( \mathbf{a}_{t-j} \), we can write
\[
    w_t = v_0 \mathbf{a}_t + \sum_{j=1}^{s} \mathbf{u}_j \mathbf{a}_{t-j}, \quad \text{(4.21)}
\]
where \( \mathbf{u}_j \)'s are \( k \)-dimensional row vectors, \( \mathbf{u}_s \neq \mathbf{0} \) if \( s > 0 \), and it is understood that the summation is zero if its upper limit is smaller than its lower limit. Thus, an SCM of order \((r, s)\) implies that there exists a non-zero linear combination of \( \mathbf{z}_t, \ldots, \mathbf{z}_{t-r} \) which is also a linear function of \( \mathbf{a}_t, \ldots, \mathbf{a}_{t-s} \).

With this interpretation, it is seen that a Kronecker index \( k_j \) of \( \mathbf{z}_t \) implies the existence of an SCM\((k_j, k_j)\) of \( \mathbf{z}_t \).

Note that \( y_t \) being an SCM of order \((r, s)\) does not necessarily imply that \( y_t \) follows a univariate ARMA\((r, s)\) model. The SCM is a concept within the vector framework and it uses all the components of \( \mathbf{z}_t \) in describing a model. On the other hand, a univariate model of \( y_t \) only depends on its own history \( y_{t-j} \) for \( j > 0 \). Also, from the definition, the order \((r, s)\) of an SCM \( y_t \) is not unique. For example, multiplying \( w_{t-m} \) with \( m > 0 \) by a non-zero constant \( c \), then adding it to \( w_t \), we obtain, from Eq. (4.21), a new scalar process
\[
    w_t^s = w_t + c w_{t-m} = v_0 \mathbf{a}_t + \sum_{j=1}^{s} \mathbf{u}_j \mathbf{a}_{t-j} + c \left( v_0 \mathbf{a}_{t-m} + \sum_{j=1}^{s} \mathbf{u}_j \mathbf{a}_{t-m-j} \right),
\]
which is uncorrelated with \( \mathbf{a}_{t-\ell} \) for \( \ell > s + m \). This type of redundancies should be eliminated, so that we require that the order \((r, s)\) of a SCM satisfies the condition that \( r + s \) is as small as possible. Note that, even with the requirement on the sum \( r + s \), the order of a given SCM is still not unique. For instance, consider the model
\[
    \mathbf{z}_t - \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix} \mathbf{z}_{t-1} = \mathbf{a}_t,
\]

11
which can be written equivalently as
\[ z_t = a_t - \begin{bmatrix} 0 & 0 \\ -2 & 0 \end{bmatrix} a_{t-1}. \]

It is easily seen that \( z_{2t} = 0,1] \) is an SCM of order (1,0) or (0,1). Here both orders satisfy \( r+s = 1 \), the lowest possible value. This type of non-uniqueness does not cause problems in model specification because the sum \( r+s \), which is fixed, plays an important role in the scalar-component approach. We shall discuss this issue in the next subsection.

4.2.2 Model specification via scalar component models

Suppose that \( y_{it} = (v_i^{(1)}, \cdots, v_i^{(k)}) \) is an SCM \((p_i,q_i)\) of \( z_t \), where \( i = 1, \cdots, k \). These \( k \) SCM’s are said to be linearly independent if the \( k \times k \) matrix \( T' = [v_0^{(1)}, \cdots, v_0^{(k)}] \) is non-singular, i.e. \( |T| \neq 0 \). For a \( k \)-dimensional process \( z_t \) in Eq. (4.1), a set of \( k \) linearly independent SCMs determines a VARMA model for \( z_t \). From the definition, for each SCM \( y_{it} \) there exist \( p_i \)-dimensional row vector \( \{v_j^{(i)}\}_{j=1}^{p_i} \) such that the scalar process \( w_{it} = \sum_{\ell=0}^{p_i} v_{\ell}^{(i)} z_{t-\ell} \) is uncorrelated with \( a_{t-j} \) for \( j > q_i \). Let \( w_t = (w_{1t}, \cdots, w_{kt})' \), \( r = \max\{p_i\} \), and \( s = \max\{q_i\} \). We have

\[ w_t = T z_t + \sum_{\ell=1}^{r} G_{\ell} z_{t-\ell}, \quad (4.22) \]

where \( G_{\ell} = [v_{\ell}^{(1)}, \cdots, v_{\ell}^{(k)}] \) with \( v_{\ell}^{(i)} = 0 \) for \( p_i < \ell \leq r \). Furthermore, from Eq. (4.21), \( w_t \) can also be written as

\[ w_t = T a_t + \sum_{\ell=1}^{s} U_{\ell} a_{t-\ell}, \quad (4.23) \]

where \( U_{\ell} = [u_{\ell}^{(1)}, \cdots, u_{\ell}^{(k)}] \) is a \( k \times k \) matrix whose \( i \)th row is zero if \( q_i < \ell \leq s \). Combining Eqs. (4.22) and (4.23), we have a VARMA \((r,s)\) model for \( z_t \). Furthermore, the row structure of the coefficient matrices for the specified model is available. More specifically, we have

\[ T z_t + \sum_{\ell=1}^{r} G_{\ell} z_{t-\ell} = T a_t + \sum_{\ell=1}^{s} U_{\ell} a_{t-\ell} \quad (4.24) \]

such that

1. the \( i \)th row of \( G_{\ell} \) is zero if \( p_i < \ell \leq r \),
2. the \( i \)th row of \( U_{\ell} \) is zero if \( q_i < \ell \leq s \), and
3. some further reduction in parameterization is possible under certain circumstances.

The last result is due to certain identifiable redundant parameters between AR and MA components in Eq. (4.24), which we shall discuss in the next subsection. From Eq. (4.24), a VARMA \((r,s)\) model for \( z_t \) is obtained. Note that by inserting \( T^{-1}T \) in the front of \( z_{t-\ell} \) and \( a_{t-\ell} \) in Eq. (4.24), one obtains a VARMA \((r,s)\) model for the transformed process \( y_t \)

\[ (I - \varphi_1 B - \cdots - \varphi_r B^r)y_t = (I - \Theta_1 B - \cdots - \Theta_s B^s)b_t, \quad (4.25) \]
where \( \mathbf{b}_t = T \mathbf{a}_t \), \( \varphi_j = -G_j T^{-1} \) and \( \Theta_j = -U_j T^{-1} \). Because multiplication from right does not change the structure of a zero-row, we see that \( \varphi_i \) and \( \Theta_j \) have the same row-structure as those of \( G_i \) and \( U_j \), respectively, for \( i = 1, \ldots, r \) and \( j = 1, \ldots, s \). From the model in Eq. (4.25), it is clear that the order \( (p_i, q_i) \) of an SCM signifies that one needs \( p_i + q_i \) unknown rows to describe the structure of \( y_t \) in the VARMA model of \( y_t \). Here, by unknown row we mean that its parameters require estimation. This is used in contrast with the other rows that are known to be zero.

### 4.2.3 Redundant parameters

In this subsection, we consider the possible redundant parameters in the VARMA representation of Eq. (4.25) and discuss a method that can easily identify such parameters when they exist. It is worth mentioning that redundant parameters can occur even without over-specifying the overall order \((r, s)\) of Eq. (4.25).

Suppose that the orders \((p_1, q_1)\) and \((p_2, q_2)\) of the first two SCM’s \( y_{1t} \) and \( y_{2t} \) satisfy \( p_2 > p_1 \) and \( q_2 > q_1 \). In this case, we can write the model structure for \( y_{1t} \) and \( y_{2t} \) as

\[
y_{it} - [\varphi_p^{(i)} B + \cdots + \varphi_p^{(i)} B^p] y_{it} = b_{it} - [\Theta_1^{(i)} B + \cdots + \Theta_q^{(i)} B^q] \mathbf{b}_t,
\]

where \( i = 1, 2 \) and \( A^{(i)} \) denotes the \( i \)th row of the matrix \( A \). Now for \( i = 2 \) we see from Eq. (4.26) that \( y_{2t} \) is related to \( y_{1,t-1}, \ldots, y_{1,t-p_2} \) and \( b_{1,t-1}, \ldots, b_{1,t-q_2} \) via

\[
(\varphi_{1,2} B + \cdots + \varphi_{p_2,2} B^{p_2}) y_{1t} - (\Theta_{1,2} B + \cdots + \Theta_{q_2,2} B^{q_2}) b_{1t},
\]

where \( A_{v,ij} \) denotes the \((i, j)\)th element of the matrix \( A_v \). On the other hand, from Eq. (4.26) with \( i = 1 \), we have

\[
B^\ell (y_{1t} - b_{1t}) = [\varphi_1^{(1)} B + \cdots + \varphi_{p_2}^{(1)} B^{p_1}] y_{t-\ell} - [\Theta_1^{(1)} B + \cdots + \Theta_q^{(1)} B^q] \mathbf{b}_{t-\ell}.
\]

Therefore, if all the \( y \)'s and \( b \)'s on the right hand side of Eq. (4.28) are in the component model for \( y_{2t} \), then either the coefficient of \( y_{1,t-\ell} \) or that of \( b_{1,t-\ell} \) is redundant given that the other is in the model. Consequently, if \( p_2 > p_1 \) and \( q_2 > q_1 \), then for each pair of parameters \((\varphi_{\ell,21}, \Theta_{\ell,21})\), \( \ell = 1, \ldots, \min\{p_2 - p_1, q_2 - q_1\} \), only one of them is needed.

The preceding method of spotting redundant parameters in a VARMA model of Eq. (4.25) is referred to as the rule of elimination in Tiao and Tsay (1989). In general, by considering an ARMA model constructed from SCM’s and applying the rule of elimination in a pairwise fashion, all redundant parameters of the model structure for \( y_{1t} \) in Eq. (4.25) can be eliminated. By applying the rule of elimination to each pair of SCM’s, we obtain

\[
\eta_i = \sum_{v=1}^k \max[0, \min\{p_i - p_v, q_i - q_v\}].
\]

Putting all results together, we see that the total number of unknown parameters in the coefficient matrices of Eq. (4.25) is

\[
P = k \times \sum_{i=1}^k (p_i + q_i) - \sum_{i=1}^k \eta_i,
\]

which can be much smaller than \( k^2(r + s) \).
Example. Suppose that $z_t$ is a bivariate linear process with two linearly independent scalar components $y_{1t} \sim \text{SCM}(1,0)$ and $y_{2t} \sim \text{SCM}(2,1)$. In this case, we have $r = \max(1, 2) = 2$ and $s = \max(0, 1) = 1$. The model for $y_t = (y_{1t}, y_{2t})'$ is a VARMA(2,1) model. Since $y_{1t}$ is SCM(1,0), we have

$$y_{1t} = \varphi_{1,11} y_{1,t-1} + \varphi_{1,12} y_{2,t-1} + b_{1t}, \tag{4.31}$$

where, again, $\varphi_{\ell,ij}$ denotes the $(i,j)$th element of the matrix $\varphi_{\ell}$. Since $y_{2t}$ is SCM(2,1), we have

$$y_{2t} = \varphi_{1,21} y_{1,t-1} + \varphi_{1,22} y_{2,t-1} + \varphi_{2,21} y_{1,t-2} + \varphi_{2,22} y_{2,t-2} + b_{2t} - \Theta_{1,21} b_{1,t-1} - \Theta_{1,22} b_{2,t-1}. \tag{4.32}$$

By time-invariance of the system, Eq. (4.31) gives

$$y_{1,t-1} = \varphi_{1,11} y_{1,t-2} + \varphi_{1,12} y_{2,t-2} + b_{1,t-1}. \tag{4.33}$$

Note that all terms of Eq. (4.33) appear in the right hand side of Eq. (4.32). This situation occurs because the orders of the two SCMs satisfy the condition that $p_2 > p_1$ and $q_2 > q_1$. Consequently, one can substitute either $y_{1,t-1}$ or $b_{1,t-1}$ of Eq. (4.32) by Eq. (4.33) to simplify the model. In fact, the parameters $\varphi_{1,21}$ and $\Theta_{1,21}$ are not identifiable in the sense that one of them is redundant given the other. Therefore, one can fix one of these two parameters to zero to simplify the model structure.

4.2.4 A VARMA model specification

Results of the preceding subsections enable us to specify an estimable VARMA model for the linear vector process $z_t$ of Eq. (4.1) if $k$ linearly independent SCMs are given. Obviously, the order $(p_i, q_i)$ of a given SCM must satisfy the condition that $p_i + q_i$ is as small as possible. In this subsection, we provide an demonstration. Suppose that $k = 4$ and $z_t$ has 4 linearly independent SCM’s of orders $(0,0)$, $(0,1)$, $(1,0)$ and $(2,1)$. Since $\max\{p_i\} = 2$ and $\max\{q_i\} = 1$, $z_t$ is a VARMA(2,1) process. Furthermore, one can easily write down the specified model for the transformed series $y_t = Tz_t$, where $T$ is the matrix of SCM’s. The model is given by

$$(I - \varphi_1 B - \varphi_2 B^2)y_t = (I - \Theta_1 B)b_t,$$

where the coefficient matrices are

$$\varphi_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ X & X & X & X \\ X & X & X & X \end{bmatrix}, \quad \varphi_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ X & X & X & X \\ X & X & X & X \end{bmatrix},$$

$$\Theta_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ X & X & X & X \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

where 0 denotes a zero parameter and $X$ denotes an unknown parameter. Note that the (4,1) and (4,3) elements of $\Theta_1$ are set to zero because, by applying the rule of elimination, they are redundant once the (4,1) and the (4,3) elements of $\varphi_1$ are in the model. Consequently, in this particular instance modeling the transformed series $y_t$ would involve 18 parameters in the coefficient matrices instead of 48 parameters for an unrestricted VARMA(2,1) model.
4.2.5 Transformation matrix

In practice, the transformation matrix $T$ must be estimated. A simple, but less efficient, method is to use a two-stage procedure. First, one estimates $T$ in the process of identifying SCM models. Second, using the estimated $T$, one can transform the data to perform further estimation.

A more suitable method is to estimate $T$ jointly with other parameters. This approach would require further study of properties of relationships between SCMs. For a $k$-dimensional process $z_t$, $T$ may contain as many as $k(k-1)$ parameters upon normalization. However, in some cases, $T$ can be reduced to an upper triangular matrix without changing the row-structure of the specified VARMA model. For instance, suppose that $k=2$, $(p_1, q_1) = (1,0)$ and $(p_2, q_2) = (1,1)$. Then, the model specified by the SCM approach for the original series $z_t$ is

$$Tz_t - \begin{bmatrix} X & X \\ X & X \end{bmatrix} Tz_{t-1} = Ta_t - \begin{bmatrix} 0 & 0 \\ X & X \end{bmatrix} Ta_{t-1}, \quad (4.34)$$

where $T = [v_0^{(1)'}, v_0^{(2)'}]$. By rearranging the order of the components of $z_t$ if necessary, we may assume that the first element of $v_0^{(1)}$ is nonzero. Define the lower triangular $2 \times 2$ matrix

$$G = \begin{bmatrix} 1 & 0 \\ -v_{0,1}^{(2)}/v_{0,1}^{(1)} & 1 \end{bmatrix},$$

where $v_{0,1}^{(i)}$ is the first element of $v_0^{(i)}$ for $i = 1, 2$. Pre-multiplying Eq. (4.34) by $G$ and inserting $G^{-1}G$ in the front of $Tz_{t-1}$ and $Ta_{t-1}$, we have

$$T^*z_t - \begin{bmatrix} X & X \\ X & X \end{bmatrix} T^*z_{t-1} = T^*a_t - \begin{bmatrix} 0 & 0 \\ X & X \end{bmatrix} T^*a_{t-1}, \quad (4.35)$$

where $T^* = GT$ is an upper triangular matrix. Thus, in this particular instance, one can make the transformation matrix upper triangular without changing the row-structure of the VARMA model of the transformed process. Furthermore, from Eq. (4.35), the orders of the two SCM’s are not altered. In general, whenever the orders of any two SCM’s are nested, namely $p_i \leq p_j$ and $q_i \leq q_j$, one can simplify the transformation matrix $T$ by eliminating a non-zero parameter without altering the row-structure of the SCM specification. More specifically, suppose that the orders of SCM’s $y_i$ of $z_t$ are $(p_i, q_i)$ for $i = 1, \cdots, k$. Then, to obtain further simplification in the transformation matrix $T$, one can simply examine the $k(k+1)/2$ pairs of SCM’s. For any nested pair, by using the technique illustrated in Eqs. (4.34) and (4.35), one can identify a zero parameter in $T$. Mathematically, the total number of zero parameters identified by such a procedure is

$$\tau = \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \text{Ind}[\min(p_j - p_i, q_j - q_i) \geq 0],$$

where $\text{Ind}(\cdot)$ is an indicator operator which assumes the value 1 if its argument is true and the value 0, otherwise.