Multivariate processes

Andreas Jakobsson

VARMAX processes

In many applications, one may estimate many forms of data simultaneously, such that ones measurements are of the form

$$\mathbf{z}_t = \begin{bmatrix} z_{t,1} & \dots & z_{t,m} \end{bmatrix}^T$$

For such measurements, one generally wish to take the data from all m sources into account simultaneously. To do this, we introduce the notion of an m-dimensional stochastic process, $\mathbf{z}_t$.

VARMAX processes

In many applications, one may estimate many forms of data simultaneously, such that ones measurements are of the form

$$\mathbf{z}_t = \begin{bmatrix} z_{t,1} & \dots & z_{t,m} \end{bmatrix}^T$$

For such measurements, one generally wish to take the data from all m sources into account simultaneously. To do this, we introduce the notion of an m-dimensional stochastic process, $\mathbf{z}_t$.

The mean of the process is given as

$$\mathbf{m}_\mathbf{z} = E\{\mathbf{z}_t\} = \begin{bmatrix} E\{z_{t,1}\} & \dots & E\{z_{t,m}\} \end{bmatrix}^T$$

The covariance (matrix) function for the vector process is defined as

$$\begin{aligned} \mathbf{R}_\mathbf{z}(k) &= C\left\{\mathbf{z}_t, \mathbf{z}_{t-k}\right\} \\ &= E\left\{[\mathbf{z}_t - \mathbf{m}_\mathbf{z}]\,[\mathbf{z}_{t-k} - \mathbf{m}_\mathbf{z}]^*\right\} \\ &= \mathbf{R}_\mathbf{z}^*(-k) \end{aligned}$$

The notion of a white noise is formed as

$$\mathbf{R}_\mathbf{z}(k) = \begin{cases} \mathbf{\Sigma}_\mathbf{z} & k = 0 \\ \mathbf{0} & k \neq 0 \end{cases}$$

If the process is also uncorrelated among data souces, $\mathbf{\Sigma}_\mathbf{z} = \mathbf{I}$. This is called a *doubly white* process.

VARMAX processes

The multivariate, or vector, ARMAX (VARMAX) process is defined as

$$\mathbf{A}(z)\mathbf{y}_t = \mathbf{B}(z)\mathbf{x}_{t-d} + \mathbf{C}(z)\mathbf{e}_t$$

where $\mathbf{A}(z)$ and $\mathbf{C}(z)$ are monic $(m \times m)$ matrix polynomials of order p and q, respectively, $\mathbf{B}(z)$ is a $(m \times s)$ matrix polynomial of order r, and $\mathbf{x}_{t-d}$ is a d-step delayed s-dimensional exogenous input signal, i.e.,

$$\begin{aligned}
\mathbf{A}(z) &= \mathbf{I} + \mathbf{A}_1 z^{-1} + \ldots + \mathbf{A}_p z^{-p} \\
\mathbf{B}(z) &= \mathbf{B}_0 + \mathbf{B}_1 z^{-1} + \ldots + \mathbf{B}_r z^{-r} \\
\mathbf{C}(z) &= \mathbf{I} + \mathbf{C}_1 z^{-1} + \ldots + \mathbf{C}_q z^{-q}
\end{aligned}$$

and $\mathbf{e}_t$ is a multivariate zero-mean white noise process with variance $\mathbf{\Sigma}_e$. To ensure stability, we require that all roots of $\det\{\mathbf{A}(z)\} = 0$, with respect to z, lie within the unit circle.

Excluding some parts yields the VMA, VAR, VARMA, etc, models.

VARMAX processes

Example:
Consider a first-order vector MA (VMA) process $\mathbf{y}_t = \mathbf{C}(z)\mathbf{e}_t$, where $\mathbf{C}(z) = \mathbf{I} + \mathbf{C}_1 z^{-1}$, with $\mathbf{e}_t$ being a multivariate zero-mean white noise process with covariance matrix $\mathbf{\Sigma}_e$. Then,

$$
\begin{aligned}
\mathbf{R}_{\mathbf{y}}(k) &= C\left\{\mathbf{y}_t, \mathbf{y}_{t-k}\right\} \\
&= E\left\{\left(\mathbf{e}_t + \mathbf{C}_1\mathbf{e}_{t-1}\right)\left(\mathbf{e}_{t-k} + \mathbf{C}_1\mathbf{e}_{t-k-1}\right)^*\right\} \\
&= \begin{cases} \mathbf{\Sigma}_e + \mathbf{C}_1\mathbf{\Sigma}_e\mathbf{C}_1^* & k = 0 \\ \mathbf{C}_1\mathbf{\Sigma}_e & k = 1 \\ \mathbf{\Sigma}_e\mathbf{C}_1^* & k = -1 \\ \mathbf{0} & |k| > 1 \end{cases}
\end{aligned}
$$

Note that $\mathbf{R}_{\mathbf{y}}(1) = \mathbf{R}_{\mathbf{y}}^* - 1)$ and that $\mathbf{R}_{\mathbf{y}}(k)$ is zero after lag one.

VARMAX processes

Example:
Consider a first-order vector MA (VMA) process $\mathbf{y}_t = \mathbf{C}(z)\mathbf{e}_t$, where $\mathbf{C}(z) = \mathbf{I} + \mathbf{C}_1 z^{-1}$, with $\mathbf{e}_t$ being a multivariate zero-mean white noise process with covariance matrix $\mathbf{\Sigma}_e$. Then,

$$
\begin{aligned}
\mathbf{R}_{\mathbf{y}}(k) &= C\left\{\mathbf{y}_t, \mathbf{y}_{t-k}\right\} \\
&= E\left\{\left(\mathbf{e}_t + \mathbf{C}_1\mathbf{e}_{t-1}\right)\left(\mathbf{e}_{t-k} + \mathbf{C}_1\mathbf{e}_{t-k-1}\right)^*\right\} \\
&= \begin{cases} \mathbf{\Sigma}_e + \mathbf{C}_1\mathbf{\Sigma}_e\mathbf{C}_1^* & k = 0 \\ \mathbf{C}_1\mathbf{\Sigma}_e & k = 1 \\ \mathbf{\Sigma}_e\mathbf{C}_1^* & k = -1 \\ \mathbf{0} & |k| > 1 \end{cases}
\end{aligned}
$$

Note that $\mathbf{R}_{\mathbf{y}}(1) = \mathbf{R}_{\mathbf{y}}^* - 1)$ and that $\mathbf{R}_{\mathbf{y}}(k)$ is zero after lag one.

Example:
Consider a first-order VAR process $\mathbf{A}(z)\mathbf{y}_t = \mathbf{e}_t$, where $\mathbf{A}(z) = \mathbf{I} + \mathbf{A}_1 z^{-1}$, with $\mathbf{e}_t$ being a zero-mean white noise process with covariance matrix $\mathbf{\Sigma}_e$. Then,

$$
\begin{aligned}
\mathbf{R}_{\mathbf{y}}(k) &= C\left\{\mathbf{y}_t, \mathbf{y}_{t-k}\right\} \\
&= E\left\{\left(\mathbf{e}_t - \mathbf{A}_1\mathbf{y}_{t-1}\right)\mathbf{y}_{t-k}^*\right\} \\
&= E\left\{\mathbf{e}_t\mathbf{y}_{t-k}^*\right\} - \mathbf{A}_1 E\left\{\mathbf{y}_{t-1}\mathbf{y}_{t-k}^*\right\} \\
&= \begin{cases} \mathbf{\Sigma}_e - \mathbf{A}_1\mathbf{R}_{\mathbf{y}}(-1) & k = 0 \\ -\mathbf{A}_1\mathbf{R}_{\mathbf{y}}(k-1) & k > 0 \end{cases} \\
&= \begin{cases} \mathbf{\Sigma}_e - \mathbf{A}_1\mathbf{R}_{\mathbf{y}}(-1) & k = 0 \\ \left(-\mathbf{A}_1\right)^k\mathbf{R}_{\mathbf{y}}(0) & k > 0 \end{cases}
\end{aligned}
$$

Thus,

$$
\mathbf{\Sigma}_e = \mathbf{R}_{\mathbf{y}}(0) + \mathbf{A}_1\mathbf{R}_{\mathbf{y}}(-1) = \mathbf{R}_{\mathbf{y}}(0) - \mathbf{A}_1\mathbf{R}_{\mathbf{y}}(0)\mathbf{A}_1^*
$$

VARMAX processes

The autocorrelation matrix is formed as

$$\boldsymbol{\rho}_{\mathbf{z}}(k) = \mathbf{P}_{\mathbf{z}}^{-1/2} \mathbf{R}_{\mathbf{z}}(k) \mathbf{P}_{\mathbf{z}}^{-1/2}$$

where

$$\mathbf{P}_{\mathbf{z}} = \operatorname{diag}\left\{ \begin{bmatrix} [\mathbf{R}_{\mathbf{z}}(0)]_{1,1} & \ldots & [\mathbf{R}_{\mathbf{z}}(0)]_{m,m} \end{bmatrix} \right\}$$

with diag $\{\mathbf{x}\}$ denoting the diagonal matrix formed with the vector $\mathbf{x}$ along the diagonal.

	$k = 1$	$k = 2$	$k = 3$	$k = 4$
$\boldsymbol{\rho}_{\mathbf{y}}(k)$	$\begin{bmatrix} -0.67 & -0.36 \\ 0.31 & -0.57 \end{bmatrix}$	$\begin{bmatrix} 0.41 & 0.50 \\ -0.43 & -0.08 \end{bmatrix}$	$\begin{bmatrix} -0.14 & -0.37 \\ 0.33 & 0.43 \end{bmatrix}$	$\begin{bmatrix} 0.00 & 0.13 \\ -0.13 & -0.38 \end{bmatrix}$
	$\begin{bmatrix} - & - \\ + & - \end{bmatrix}$	$\begin{bmatrix} + & + \\ - & - \end{bmatrix}$	$\begin{bmatrix} - & - \\ - & + \end{bmatrix}$	$\begin{bmatrix} . & + \\ - & - \end{bmatrix}$

Example with $\pm 2/\sqrt{N} \approx \pm 0.063$.

We can use the ACF and the PACF for identification! The latter is formed using the multivariate Yule-Walker equations.

A useful trick

An often useful result when working with multivariate processes is that

$$\mathrm{vec}\left\{\mathbf{ABC}\right\} = (\mathbf{C}^T \otimes \mathbf{A})\mathrm{vec}\left\{\mathbf{B}\right\}$$

Example:

$$\mathrm{vec}\left\{\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}\right\} = \begin{bmatrix} x_{11} \\ x_{21} \\ x_{12} \\ x_{22} \end{bmatrix}$$

Example:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \otimes \begin{bmatrix} 0 & 5 & 1 \\ 6 & 7 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 5 & 1 & 0 & 10 & 2 \\ 6 & 7 & 0 & 12 & 14 & 0 \\ 0 & 15 & 3 & 0 & 20 & 4 \\ 18 & 21 & 0 & 24 & 28 & 0 \end{bmatrix}$$

A useful trick

An often useful result when working with multivariate processes is that

$$\text{vec}\,\{\mathbf{ABC}\} = (\mathbf{C}^T \otimes \mathbf{A})\text{vec}\,\{\mathbf{B}\}$$

Example:

$$\text{vec}\left\{\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}\right\} = \begin{bmatrix} x_{11} \\ x_{21} \\ x_{12} \\ x_{22} \end{bmatrix}$$

Example:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \otimes \begin{bmatrix} 0 & 5 & 1 \\ 6 & 7 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 5 & 1 & 0 & 10 & 2 \\ 6 & 7 & 0 & 12 & 14 & 0 \\ 0 & 15 & 3 & 0 & 20 & 4 \\ 18 & 21 & 0 & 24 & 28 & 0 \end{bmatrix}$$

Thus, we may express

$$\mathbf{\Sigma}_e = \mathbf{R}_{\mathbf{y}}(0) - \mathbf{A}_1 \mathbf{R}_{\mathbf{y}}(0) \mathbf{A}_1^*$$

as

$$\begin{aligned} \text{vec}\,\{\mathbf{\Sigma}_e\} &= \text{vec}\,\{\mathbf{R}_{\mathbf{y}}(0)\} - \text{vec}\,\{\mathbf{A}_1 \mathbf{R}_{\mathbf{y}}(0) \mathbf{A}_1^*\} \\ &= \text{vec}\,\{\mathbf{R}_{\mathbf{y}}(0)\} - \left(\bar{\mathbf{A}}_1 \otimes \mathbf{A}_1\right) \text{vec}\,\{\mathbf{R}_{\mathbf{y}}(0)\} \\ &= \left[\mathbf{I} - \bar{\mathbf{A}}_1 \otimes \mathbf{A}_1\right] \text{vec}\,\{\mathbf{R}_{\mathbf{y}}(0)\} \end{aligned}$$

where $\bar{\mathbf{A}}$ denotes the conjugate of $\mathbf{A}$.

A useful trick

Example:
Consider a first-order VAR process with $\mathbf{\Sigma}_e = \mathbf{I}$ and

$$\mathbf{A}_1 = \begin{bmatrix} 0.5 & 0.4 \\ 0.1 & 0.8 \end{bmatrix}$$

As

$$\det\left(\mathbf{I} + \mathbf{A}_1 z^{-1}\right) = 1 + 1.3z^{-1} + 0.36z^{-2} = 0$$

has the roots $z_1 = -0.9$ and $z_2 = -0.4$, which both lie within the unit circle, the process is stable.

A useful trick

Example:
Consider a first-order VAR process with $\mathbf{\Sigma}_e = \mathbf{I}$ and

$$\mathbf{A}_1 = \begin{bmatrix} 0.5 & 0.4 \\ 0.1 & 0.8 \end{bmatrix}$$

As

$$\det\left(\mathbf{I} + \mathbf{A}_1 z^{-1}\right) = 1 + 1.3z^{-1} + 0.36z^{-2} = 0$$

has the roots $z_1 = -0.9$ and $z_2 = -0.4$, which both lie within the unit circle, the process is stable. Setting

$$\mathbf{R}_{\mathbf{y}}(0) = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}$$

it holds that

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} - \begin{bmatrix} 0.5 & 0.4 \\ 0.1 & 0.8 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} \begin{bmatrix} 0.5 & 0.1 \\ 0.4 & 0.8 \end{bmatrix}$$

A useful trick

Example:
Consider a first-order VAR process with $\mathbf{\Sigma}_e = \mathbf{I}$ and

$$\mathbf{A}_1 = \begin{bmatrix} 0.5 & 0.4 \\ 0.1 & 0.8 \end{bmatrix}$$

As

$$\det\left(\mathbf{I} + \mathbf{A}_1 z^{-1}\right) = 1 + 1.3z^{-1} + 0.36z^{-2} = 0$$

has the roots $z_1 = -0.9$ and $z_2 = -0.4$, which both lie within the unit circle, the process is stable. Setting

$$\mathbf{R}_{\mathbf{y}}(0) = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}$$

it holds that

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} - \begin{bmatrix} 0.5 & 0.4 \\ 0.1 & 0.8 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} \begin{bmatrix} 0.5 & 0.1 \\ 0.4 & 0.8 \end{bmatrix}$$

or, using $\mathrm{vec}\left\{\mathbf{\Sigma}_e\right\} = \left[\mathbf{I} - \bar{\mathbf{A}}_1 \otimes \mathbf{A}_1\right] \mathrm{vec}\left\{\mathbf{R}_{\mathbf{y}}(0)\right\}$

$$\begin{bmatrix} 0.75 & -0.20 & -0.20 & -0.16 \\ -0.05 & 0.60 & -0.04 & -0.32 \\ -0.05 & -0.04 & 0.60 & -0.32 \\ -0.01 & -0.08 & -0.08 & 0.36 \end{bmatrix} \begin{bmatrix} r_{11} \\ r_{21} \\ r_{12} \\ r_{22} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

yielding

$$\mathbf{R}_{\mathbf{y}}(0) = \begin{bmatrix} 3.6028 & 2.6355 \\ 2.6355 & 4.0492 \end{bmatrix}$$

$$\mathbf{R}_{\mathbf{y}}(k) = (-1)^k \begin{bmatrix} 0.5 & 0.4 \\ 0.1 & 0.8 \end{bmatrix}^k \begin{bmatrix} 3.6028 & 2.6355 \\ 2.6355 & 4.0492 \end{bmatrix}$$

It is worth noting that, as expected, $r_{12} = r_{21}$.