Multivariate time series modelling *

Jean-Marie Dufour †
Université de Montréal

First version: May 1999
Revised: March 2006
This version: March 28, 2006
Compiled: March 28, 2006, 9:55pm

* This work was supported by the Canada Research Chair Program (Chair in Econometrics, Université de Montréal), the Alexander-von-Humboldt Foundation (Germany), the Canadian Network of Centres of Excellence [program on *Mathematics of Information Technology and Complex Systems* (MITACS)], the Canada Council for the Arts (Killam Fellowship), the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, the Fonds de recherche sur la société et la culture (Québec), and the Fonds de recherche sur la nature et les technologies (Québec).

† Canada Research Chair Holder (Econometrics). Centre interuniversitaire de recherche en analyse des organisations (CIRANO), Centre interuniversitaire de recherche en économie quantitative (CIREQ), and Département de sciences économiques, Université de Montréal. Mailing address: Département de sciences économiques, Université de Montréal, C.P. 6128 succursale Centre-ville, Montréal, Québec, Canada H3C 3J7. TEL: 1 (514) 343 2400; FAX: 1 (514) 343 5831; e-mail: jean.marie.dufour@umontreal.ca . Web page: http://www.fas.umontreal.ca/SCECO/Dufour

Contents

List of Definitions, Propositions and Theorems

1. Multivariate time series models

Let

$$X(t) = \big(X_1(t), X_2(t), \ldots, X_m(t)\big)', \; t \in \mathbb{Z} \tag{1.1}$$

a vector of real-valued time series. By convention,

$$X_t \equiv X(t), \quad X_{it} \equiv X_i(t).$$

1.1 Definition MULTIVARIATE STRICT STATIONARITY. *$X(t)$ is* strictly stationary *iff*

$$\begin{aligned}&\big(X(t_1)', X(t_2)', \ldots, X(t_k)'\big)'\\ &\sim \big(X(t_1+\ell)', X(t_2+\ell)', \ldots, X(t_k+\ell)'\big)'\end{aligned}$$

for all $t_1, t_2, \ldots, t_k \in \mathbb{Z}, k \geq 1, \; \ell \geq 0.$

1.2 Definition L^2 VECTOR. *A random vector $X = (X, \ldots, X_m)'$ is in L^2 iff each one of its components is in L^2, i.e.,*

$$\mathsf{E}\big(X_i^2\big) < \infty, \; i = 1, \ldots, m.$$

1.3 Definition L^2 VECTOR PROCESS. *The stochastic process $\{X(t) : t \in \mathbb{Z}\}$ is in L^2 iff each one of the vectors $X(t), t \in \mathbb{Z}$, is in L^2.*

1.4 Definition MULTIVARIATE WEAK STATIONARITY. *$X(t)$ is* second-order stationary *(or* weakly stationary*) iff*

(a) $\mathsf{E}\left[X_i(t)^2\right] < \infty, \; i = 1, \ldots, m, \; \forall t;$

(b) $\mathsf{E}[X(t)] = \mu, \forall t;$

(c) $\mathsf{E}\left[(X(t) - \mu)(X(t+k) - \mu)'\right] = \Gamma_k, \; \forall t,$ *for all* $k \in \mathbb{Z}$.

1.5 Notation

$$\Gamma_k \equiv \Gamma(k) = \left[\gamma_{ij}(k)\right]_{i,j=1,\ldots,m} \tag{1.2}$$

is an $m \times m$ matrix whose elements are

$$\gamma_{ij}(k) = \mathsf{Cov}\left(X_i(t), X_j(t+k)\right). \tag{1.3}$$

In general,

$$\begin{aligned}\gamma_{ij}(k) &= \mathsf{Cov}\left(X_i(t), X_j(t+k)\right)\\ &\neq \mathsf{Cov}\left(X_j(t), X_i(t+k)\right)\end{aligned}$$

$$= \gamma_{ji}(k) \text{ for } i \neq j \tag{1.4}$$

so that

$$\Gamma_k \neq \Gamma_k'. \tag{1.5}$$

But

$$\begin{aligned} \gamma_{ij}(k) &= \mathsf{Cov}(X_i(t), X_j(t+k)) \\ &= \mathsf{Cov}(X_j(t+k), X_i(t)) \\ &= \mathsf{Cov}(X_j(t), X_i(t-k)) \\ &= \gamma_{ji}(-k) \end{aligned} \tag{1.6}$$

so that

$$\Gamma_k = \Gamma_{-k}' \tag{1.7}$$

and (usually)

$$\Gamma_{-k} = \Gamma_k' \neq \Gamma_k \, . \tag{1.8}$$

1.6 Definition MULTIVARIATE WHITE NOISE. *An m-dimensional process $\{a(t) : t \in \mathbb{Z}\}$ is a white noise if it satisfies the following properties:*

(a) $\mathsf{E}\left[a_i(t)^2\right] < \infty, \; i = 1, \ldots, m, \forall t;$

(b) $\mathsf{E}[a(t)] = 0, \; \forall t;$

(c) $\mathsf{E}\left[a(s)\,a(t)'\right] \begin{array}{l} = \Sigma, \text{ if } s = t \\ = 0, \text{ if } s \neq t. \end{array}$

1.7 Definition MEAN SQUARE CONVERGENCE FOR VECTORS. *Let $\{X_n : n \geq 1\}$ be a sequence of $m \times 1$ random vectors in L^2, and let X be another $m \times 1$ random vector in L^2. Then we say X_n converges to X in mean square as $n \to \infty$ $\left(X_n \underset{n\to\infty}{\overset{2}{\longrightarrow}} X\right)$ iff each component of X_n converges to the corresponding component of X in mean square, i.e.,*

$$\mathsf{E}\left[(X_{in} - X_i)^2\right] \underset{n\to\infty}{\longrightarrow} 0, \, i = 1, \ldots, m \, , \tag{1.9}$$

where

$$X_n = \left(X_{1n}, \ldots, X_{mn}\right)', \quad X = \left(X_1, \ldots, X_m\right)'.$$

Consider a process of the form

$$X(t) = \mu + \sum_{k=0}^{\infty} \psi_k a(t-k) \, , \; t \in \mathbb{Z} \tag{1.10}$$

where $\{a(t) : t \in \mathbb{Z}\}$ is an m-dimensional white noise and $\{\psi_k : k \geq 0\}$ is a sequence of $m \times m$ fixed matrices such that

$$\sum_{k=0}^{\infty} tr\left(\psi_k \psi_k'\right) < \infty \,. \tag{1.11}$$

Then the series $\sum_{k=0}^{\infty} \psi_k a(t-k)$ converges in mean square and the process $X(t)$ is second-order stationary. $X(t)$ is a vector $MA(\infty)$ process. When a process has a representation of the form (1.10) where the vectors $a(t)$, $t \in \mathbb{Z}$, are i.i.d., we say $X(t)$ is a *linear process*.

The $MA(\infty)$ model can also be written

$$X(t) = \mu + \psi(B)\, a(t) \tag{1.12}$$

where

$$\psi(B) = \sum_{k=0}^{\infty} \psi_k B^k. \tag{1.13}$$

1.8 Theorem MULTIVARIATE WOLD THEOREM. *Let $\{X(t) : t \in \mathbb{Z}\}$ be an m-dimensional second-order stationary process. Then $X(t)$ can be written in the form*

$$X(t) = \mu + \sum_{k=0}^{\infty} \psi_k a(t-k) + v(t)\,, \;\; t \in \mathbb{Z}\,, \tag{1.14}$$

where $a(t) \equiv (a_1(t), \ldots, a_m(t))'$ is a white noise process, $\{\psi_k : k \geq 0\}$ is a sequence of fixed $m \times m$ matrices such that the series $\sum_{k=0}^{\infty} \psi_k a(t-k)$ converges in mean square, and $v(t)$ is a deterministic process which is uncorrelated with $a(t-j)$, $j \geq 0$. Further, we can choose $a(t)$ and the ψ_k's such that

$$\psi_0 = I_m$$

and

$$a(t) = X(t) - P_L[X(t) \mid X(t-j), j \geq 1]\,.$$

1.9 Remark If $v(t) = 0$, $\forall t$, we say the process $X(t)$ is *strictly indeterministic*.

1.10 Definition MULTIVARIATE AR(p) PROCESS. *An m-dimensional vector process $\{X(t) : t \in \mathbb{Z}\}$ follows an* $\mathrm{AR}(p)$ *model [or a* $\mathrm{VAR}(p)$ *model] if it satisfies an equation of the form :*

$$X(t) = \mu + \sum_{k=1}^{p} \Phi_k X(t-k) + a(t)\;, \;\; \forall t \tag{1.15}$$

where $\Phi_1, \ldots, \Phi_p$ are $m \times m$ fixed matrices and $\{a(t) : t \in \mathbb{Z}\}$ is a white noise process.

1.11 Definition MULTIVARIATE ARMA(p, q). *An m-dimensional vector process $\{X(t) : t \in \mathbb{Z}\}$ follows an ARMA(p, q) model [or a VARMA(p, q) model] if it satisfies an equation of the form*

$$\Phi_p(B) X(t) = \bar{\mu} + \Theta_q(B) a(t) \tag{1.16}$$

where $\bar{\mu}$ is a fixed vector,

$$\begin{aligned} \Phi_p(B) &= I_m - \Phi_1 B - \cdots - \Phi_p B^p, \\ \Theta_q(B) &= I_m - \Theta_1 B - \cdots - \Theta_q B^q, \end{aligned}$$

$\Phi, \ldots, \Phi_p, \Theta, \ldots, \Theta_q$ are fixed $m \times m$ matrices, and $\{a(t) : t \in \mathbb{Z}\}$ is a white noise process.

The VARMA(p, q) model has a stationary solution which is causal in $a(t)$ if all the roots of the determinantal equation

$$\det[\Phi_p(z)] = 0 \tag{1.17}$$

are outside the unit circle. In other words, we must find the roots of the polynomial

$$\det[I_m - \Phi_1 z - \cdots - \Phi_p z^p]$$

and check whether their moduli are greater than one.

A stationary VARMA(p, q) model has a MA(∞) representation of the form

$$X(t) = \mu + \psi(B) a(t) \tag{1.18}$$

where

$$\begin{aligned} \psi(B) &= \Phi_p(B)^{-1} \Theta_q(B), \\ \mu &= \Phi_p(B)^{-1} \bar{\mu} = [I - \Phi_1 - \cdots - \Phi_p]^{-1} \bar{\mu}. \end{aligned} \tag{1.19}$$

The VARMA(p, q) model is invertible, i.e., $X(t)$ can be written in an autoregressive form

$$\Pi(B) X(t) = a(t)$$

where

$$\Pi(B) = I_m - \Pi_1 B - \Pi_2 B^2 - \cdots$$

$$= \quad I_m - \sum_{k=1}^{\infty} \varPi_k B^k$$

and $\varPi_k,\ k = 1,\ 2,\ \ldots$ are $m \times m$ fixed matrices, when the roots of the equation

$$\det \left[\varTheta_q \left(z \right) \right] = 0$$

are all outside the unit circle. In this case, the operator $\varPi \left(B \right)$ is given by

$$\varPi \left(B \right) = \varTheta_q \left(B \right)^{-1} \varPhi_p \left(B \right) .$$

2. Alternative representations

2.1. Processes of individual components

Given a stationary model

$$\varPhi_p \left(B \right) X \left(t \right) = \bar{\mu} + \varTheta_q \left(B \right) a \left(t \right) , \tag{2.1}$$

we can write

$$X \left(t \right) = \mu + \varPhi_p \left(B \right)^{-1} \varTheta_q \left(B \right) a \left(t \right) \tag{2.2}$$

where

$$\varPhi_p \left(B \right)^{-1} = \frac{1}{\det \left[\varPhi_p \left(B \right) \right]} \varPhi_p^* \left(B \right) \ ,$$

$\varPhi_p^* \left(B \right)$ is the adjoint matrix of $\varPhi_p \left(B \right)$ and $\det \left[\varPhi_p \left(B \right) \right]$ is a polynomial in B. If all the elements of $\varPhi_p \left(B \right)$ are polynomials in B, all the elements of $\varPhi_p^* \left(B \right)$ are also polynomials. On multiplying both sides of (2.1) by $\varPhi_p^* \left(B \right)$, we get :

$$\det \left[\varPhi_p \left(B \right) \right] X \left(t \right) = \tilde{\mu} + \varPhi_p^* \left(B \right) \varTheta_q \left(B \right) a \left(t \right)$$

and

$$\det \left[\varPhi_p \left(B \right) \right] X \left(t \right) = \tilde{\mu} + \bar{\theta} \left(B \right) a \left(t \right) \tag{2.3}$$

where $\tilde{\mu} = \varPhi_p^* \left(B \right) \bar{\mu}$.

Now consider separately each row of the vector $\det \left[\varPhi_p \left(B \right) \right] X \left(t \right)$:

$$\det \left[\varPhi_p \left(B \right) \right] X_i \left(t \right) = \tilde{\mu}_i + \bar{\theta}_{i\cdot} \left(B \right) a \left(t \right) , \ i = 1,\ \ldots\ , m \, ,$$

where $\bar{\theta}_{i\cdot} \left(B \right)$ is the i-th row of $\bar{\theta} \left(B \right)$. It is easy to see that :

$$\det \left[\varPhi_p \left(B \right) \right] \text{ is a polynomial in } B \text{ of degree } mp \text{ (or less)} \tag{2.4}$$

and

$$\begin{array}{c}\text{the elements of the matrix } \bar{\theta}(B) \\ \text{are all polynomials of degree } q + (m-1)\,p.\end{array} \tag{2.5}$$

Since $a(t)$ is a white noise process, $\bar{\theta}_i(B)\,a(t)$ is a moving average of order $q+(m-1)\,p$, so each component of $X(t)$ follows an ARMA$(mp,\ q+(m-1)\,p)$ model.

2.1 Remark If $X(t)$ satisfies a $VAR(p)$ model, its components do not usually follow AR processes.

2.2. Transfer functions

The matrix $\Phi_p(B)$ has the form

$$\Phi_p(B) = [\Phi_{pij}(B)] \tag{2.6}$$

where

$$\begin{aligned} \Phi_{pij}(B) &\equiv \delta_{ij} - \sum_{k=1}^{p} \varphi_{ijk} B^k \\ \delta_{ij} &= 1 \quad , \quad \text{if} \quad i = j \\ &= 0 \quad , \quad \text{if} \quad i \neq j. \end{aligned}$$

Now, consider the $m \times m$ diagonal matrix

$$\Delta(B) = \text{diag}\,[\Phi_{pii}(B)]_{i=1,\,\dots\,,\,m}$$

and multiply both sides of (2.1) by Δ^{-1} :

$$\Delta(B)^{-1}\,\Phi_p(B)\,X(t) = \bar{\bar{\mu}} + \Delta(B)^{-1}\,\Theta_q(B)\,a(t) \tag{2.7}$$

where

$$\Delta(B)^{-1}\,\Phi_p(B) = \left[\frac{\Phi_{pij}(B)}{\Phi_{pii}(B)}\right]_{i,\,j\equiv 1,\,\dots\,,\,m}$$

has all its diagonal elements equal to 1. From (2.7), we then see that

$$X_i(t) + \sum_{\substack{k=1 \\ k\neq i}}^{m} \frac{\Phi_{pik}(B)}{\Phi_{pii}(B)} X_k(t) = \bar{\bar{\mu}}_i + \varepsilon_i(t)\,, \quad i = 1,\ \dots\ ,\ m\,, \tag{2.8}$$

where

$$\varepsilon_i(t) = \sum_{k=1}^{m} \frac{\Theta_{qik}(B)}{\Phi_{pii}(B)} a_k(t), \ \ i = 1, \ldots, m.$$

Further, $\varepsilon_i(t)$ can be shown to have an ARMA representation, i.e.,

$$\varepsilon_i(t) = \frac{\theta_i(B)}{\varphi_i(B)} \eta_i(t) \quad , \ i = 1, \ldots, m, \tag{2.9}$$

where $\eta_i(t)$ is a white noise process. On simplifying the polynomial ratios $\Phi_{pi}(B)/\Phi_{pii}(B)$, we obtain a representation of the form :

$$X_i(t) = \bar{\bar{\mu}}_i + \sum_{\substack{k=1 \\ k\neq i}}^{m} \frac{\omega_{ik}(B)}{\delta_{ik}(B)} X_k(t) + \frac{\theta_i(B)}{\varphi_i(B)} \eta_i(t), \ \ i = 1, \ldots, m. \tag{2.10}$$

This is called a transfer function. It relates each variable of the system to current and past values of the other variables and to an autocorrelated noise.

3. VAR models

A VAR(p) model is a model of the form

$$X(t) = \bar{\mu} + \sum_{k=1}^{p} \Phi_k X(t-k) + a(t) \tag{3.11}$$

where $a(t)$ is a white noise process such that

$$a(t) \perp \{X(t-k), \, k \geq 1\}$$

and

$$\mathsf{V}[a(t)] = \Sigma, \ \det(\Sigma) \neq 0.$$

We can also write the model

$$\Phi(B) X(t) = \bar{\mu} + a(t)$$

where

$$\Phi(B) = I - \Phi_1 B - \cdots - \Phi_k B^k.$$

Let also

$$\psi(B) = \Phi(B)^{-1} = \sum_{k=0}^{\infty} \psi_k B^k,$$

the matrix defined by the equation

$$\Phi(B)\,\psi(B) = I_m$$

where

$$\psi_0 = I_m.$$

Suppose the polynomial $\Phi(z)$ satisfies the stationarity condition so that model (3.11) has a stationary solution of the form :

$$\begin{aligned} X(t) &= \mu + \psi(B)\,a(t) \\ &= \mu + \sum_{k=0}^{\infty} \psi_k a(t-k)\,. \end{aligned} \tag{3.12}$$

From the latter, we see that

$$\mathsf{V}[X(t)] = \sum_{k=0}^{\infty} \psi_k \Sigma \psi_k'. \tag{3.13}$$

The coefficients of the ψ_k matrices are called the *impulse response coefficients* associated with the innovations $a(t)$. If $a(t)$ is interpreted as a vector of "shocks", the elements of ψ_k can be interpreted as the "effects" of these shocks on $X(t)$.

Further, (3.13) provides a decomposition of the covariance matrix of $X(t)$ in terms of shocks at different lags. Let

$$\begin{aligned} \psi_k &= \left[\psi_{kij}\right]_{i,\,j=1,\,\ldots\,,\,m} \\ &= \begin{bmatrix} \psi_{k1\cdot} \\ \psi_{k2\cdot} \\ \vdots \\ \psi_{km\cdot} \end{bmatrix} \end{aligned}$$

Then we can write

$$\begin{aligned} X_i(t) &= \mu_i + \sum_{k=0}^{\infty} \psi_{ki\cdot} a(t-k) \\ &= \mu_i + \sum_{k=0}^{\infty} \left(\sum_{\ell=1}^{m} \psi_{ki\ell} a_\ell (t-k) \right) \end{aligned} \tag{3.14}$$

and

$$\mathsf{V}\left[X_i\left(t\right)\right] = \sum_{k=0}^{\infty} \psi_{ki\cdot} \Sigma \psi_{ki\cdot}^{'} \ .$$

$\psi_{ki\cdot} \Sigma \psi_{ki\cdot}^{'}$ is the contribution of $a\left(t-k\right)$ to the variance of $X_i\left(t\right)$: a proportion

$$v_{ik} = \psi_{ki\cdot} \Sigma \psi_{ki\cdot}^{'} / \mathsf{V}\left[X_i\left(t\right)\right]$$

of the variance $\mathsf{V}\left[X_i\left(t\right)\right]$ is accounted for by $a\left(t-k\right)$.

However, since the elements of $a\left(t\right)$ may be correlated (Σ is not generally a diagonal matrix), the MA representation does not allow one to separate the effects of different innovations. Since Σ is a positive definite matrix, we can find a non-singular matrix P such that

$$P \Sigma P' = I_m$$

hence

$$\begin{aligned} X\left(t\right) &= \mu + \sum_{k=0}^{\infty} \psi_k P^{-1} P a\left(t-k\right) \\ &= \mu + \sum_{k=0}^{\infty} \bar{\psi}_k \, \varepsilon\left(t-k\right) \end{aligned} \tag{3.15}$$

where

$$\begin{aligned} \bar{\psi}_k &= \psi_k P^{-1} \, , \\ \varepsilon\left(t\right) &= P a\left(t\right) \ , \\ \mathsf{V}\left[\varepsilon\left(t\right)\right] &= P \Sigma P' = I_m \, . \end{aligned}$$

Since the components of $\varepsilon\left(t\right)$ are uncorrelated, we have

$$\mathsf{V}\left[X\left(t\right)\right] = \sum_{k=0}^{\infty} \bar{\psi}_k \bar{\psi}_k^{'} \, . \tag{3.16}$$

Then we can write

$$\begin{aligned} X_i\left(t\right) &= \mu_i + \sum_{k=0}^{\infty} \bar{\psi}_{ki\cdot} \varepsilon\left(t-k\right) \\ &= \mu_i + \sum_{k=0}^{\infty} \left(\sum_{\ell=1}^{m} \bar{\psi}_{ki\ell} \varepsilon_\ell \left(t-k\right) \right) \end{aligned}$$

$$= \mu_i + \sum_{\ell=1}^{m} \left(\sum_{k=0}^{\infty} \bar{\psi}_{ki\ell} \varepsilon_\ell (t-k) \right) \tag{3.17}$$

where

$$\bar{\psi}_{ki\cdot} = \left(\bar{\psi}_{ki1}, \bar{\psi}_{ki2}, \ldots, \bar{\psi}_{kim} \right)'$$

hence

$$\begin{aligned} \mathsf{V}\left[X_i(t)\right] &= \sum_{\ell=1}^{m} \mathsf{V}\left[\sum_{k=0}^{\infty} \bar{\psi}_{ki\ell} \varepsilon_\ell (t-k) \right] \\ &= \sum_{\ell=1}^{m} \left(\sum_{k=0}^{\infty} \bar{\psi}_{ki\ell}^2 \right). \end{aligned} \tag{3.18}$$

A proportion

$$p_{i\ell} = \left(\sum_{k=0}^{\infty} \bar{\psi}_{ki\ell}^2 \right) / \mathsf{V}\left[X_i(t)\right] \tag{3.19}$$

of the variance of $X_i(t)$ can be attributed to the "shocks"

$$\varepsilon_\ell (t-k), \; k \geq 0 .$$

There is an infinity of ways of orthogonalizing the innovations of a VAR model. The most common one consists in using the Choleski decomposition. In other words, we choose

$$P = T$$

where T is a lower-triangular matrix :

$$T = \begin{bmatrix} T_{11} & 0 & 0 & \cdots & 0 \\ T_{12} & T_{22} & 0 & \cdots & 0 \\ T_{31} & T_{32} & T_{33} & \cdots & 0 \\ \vdots & & & & \vdots \\ T_{m1} & T_{m2} & T_{m3} & \cdots & T_{mm} \end{bmatrix} .$$

In other words, the matrix Σ is orthogonalized by using the Gram-Schmidt method.

Instead of decomposing the total variance $\mathsf{V}[X(t)]$, it is also possible to look at similar decompositions for forecast errors. Let

$$I(t) = \{X(s) : s \leq t\} .$$

Then it is easy to see that

$$\mathsf{P}_L\left[X\left(t+h\right) \mid I\left(t\right)\right] = \mu + \sum_{k=h}^{\infty} \psi_k\, a\left(t+h-k\right)$$

where $h \geq 0$, and

$$\begin{aligned} a^{(h)}\left(t\right) &\equiv X\left(t+h\right) - P_L\left[X\left(t+h\right) \mid I\left(t\right)\right] \\ &= \mu + \sum_{k=0}^{\infty} \psi_k\, a\left(t+h-k\right) - \mu - \sum_{k=h}^{\infty} \psi_k\, a\left(t+h-k\right) \\ &= \sum_{k=0}^{h-1} \psi_k\, a\left(t+h-k\right)\ , \end{aligned}$$

so that

$$\mathsf{V}\left[a^{(h)}\left(t\right)\right] = \sum_{k=0}^{h-1} \psi_k \Sigma \psi_k^{'} \underset{h\to\infty}{\longrightarrow} \mathsf{V}\left[X\left(t\right)\right]$$

and

$$\mathsf{V}\left[a_i^{(h)}\left(t\right)\right] = \sum_{k=0}^{h-1} \psi_{ki\cdot} \Sigma \psi_{ki\cdot}^{'} \underset{h\to\infty}{\longrightarrow} \mathsf{V}\left[X_i\left(t\right)\right]\ .$$

The proportion of the h-step ahead prediction error due to $a\left(t+h-k\right)$, where $0 \leq k \leq h-1$, is

$$v_{ik}^{(h)} = \psi_{ki\cdot} \Sigma \psi_{ki\cdot}^{'} / \mathsf{V}\left[a_i^{(h)}\left(h\right)\right] \underset{h\to\infty}{\longrightarrow} v_{ik\cdot}$$

Similarly, we can rewrite $a^{(h)}\left(t\right)$ in terms of orthogonalized innovations :

$$\begin{aligned} a^{(h)}\left(t\right) &= \sum_{k=0}^{h-1} \bar{\psi}_k\, \varepsilon\left(t+h-k\right)\ , \\ a_i^{(h)}\left(t\right) &= \sum_{k=0}^{h-1} \bar{\psi}_{ki\cdot}\, \varepsilon\left(t+h-k\right) = \sum_{\ell=1}^{m} \left(\sum_{k=0}^{h-1} \bar{\psi}_{ki\ell}\, \varepsilon_\ell\left(t-k\right)\right)\ , \end{aligned}$$

hence

$$\mathsf{V}\left[a^{(h)}\left(t\right)\right] = \sum_{k=0}^{h-1} \bar{\psi}_k \bar{\psi}_k^{'} \underset{h\to\infty}{\longrightarrow} \mathsf{V}\left[X\left(t\right)\right]\ , \tag{3.20}$$

$$\begin{aligned}\mathsf{V}\left[a_i^{(h)}(t)\right] &= \sum_{\ell=1}^{m} \mathsf{V}\left[\sum_{k=0}^{h-1} \bar{\psi}_{ki\ell}\, \varepsilon_\ell(t-k)\right] \\ &= \sum_{\ell=1}^{m} \left(\sum_{k=0}^{h-1} \bar{\psi}_{ki\ell}^2\right) \underset{h\to\infty}{\longrightarrow} \mathsf{V}\left[a_i^{(h)}(t)\right].\end{aligned}$$

A proportion

$$p_{i\ell}^{(h)} = \left(\sum_{k=0}^{h-1} \bar{\psi}_{ki\ell}^2\right) / \mathsf{V}\left[a_i^{(h)}(t)\right]$$

of the variance of $a_i^{(h)}(t)$ can be attributed to the shocks

$$\varepsilon_\ell(t-k),\ k \geq 0.$$

4. Bibliographic notes

To get more details on VAR and VARMA models, the reader may consult: Mills (1990, Chap. 13-14), Brockwell and Davis (1991, Sections 11.1-11.5, 13.1), Lütkepohl (1991), Hamilton (1994, Chap. 9, 18, 19, 20), Gouriéroux and Monfort (1997, Chap. VII, VIII, IX, X, XI, XIII) and Reinsel (1997). On VARMA models, see also Tiao and Box (1981) and Tiao and Box (1981).

References

Brockwell, P. J. and Davis, R. A. (1991), *Time Series: Theory and Methods*, second edn, Springer-Verlag, New York.

Gouriéroux, C. and Monfort, A. (1997), *Time Series and Dynamic Models*, Cambridge University Press, Cambridge, U.K.

Hamilton, J. D. (1994), *Time Series Analysis*, Princeton University Press, Princeton, New Jersey.

Lütkepohl, H. (1991), *Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin.

Mills, T. C. (1990), *Time Series Techniques for Economists*, Cambridge University Press, Cambridge, U.K.

Reinsel, G. C. (1997), *Elements of Multivariate Time Series Analysis*, second edn, Springer-Verlag, New York.

Tiao, G. C. and Box, G. E. P. (1981), 'Modeling multiple time series with applications', *Journal of the American Statistical Association* **76**(376), 802–816.