Chapter 2

The Statistical Setting

This chapter introduces the basic notions regarding the multivariate stochastic processes. In particular, the reader will find the definitions of stationarity and of integration which are of special interest for the subsequent developments. The second part deals with principle stationary processes. The third section shows the way to integrated processes and takes a glance at cointegration. The last sections deal with integrated and cointegrated processes and related topics of major interest. An algebraic appendix and an appendix on the role of cointegration complete this chapter.

2.1 Stochastic Processes: Preliminaries

The notion of stochastic process is a dynamic extension of the notion of random variable. Broadly speaking a random process is a process running along in time and controlled by probabilistic laws. It can be properly defined as a family, an ordered sequence, of random variables $\boldsymbol{y}_t$, where the order is given by the (discrete) time variable t.

As a mirror image of the foregoing reading key, we can look at a stochastic process as a complex of like mechanisms, whose outcomes – to be identified with the notion of time series – exhibit distinguishing features and discrepancies which can be explained on a statistical basis.

By a multivariate stochastic process we mean a random vector, say

$$\underset{(1,\ n)}{\boldsymbol{y}_t'} = [y_{t1}, y_{t2}, \ldots, y_{tn}] \tag{2.1}$$

whose elements are scalar random processes.

In order to properly specify a stochastic process, the distribution functions of its elements, pairs of elements, …, k-ples of elements, for any k, should be given and satisfy the so-called symmetry and compatibility conditions (see, e.g., Yaglom 1962).

In practise, a short cut simplification is usually adopted and reference is made to the lower-order moments of the process, basically the mean and autocovariance functions that we are going to introduce next.

Denoting by E the averaging operator, otherwise known as expectation operator, the (unconditional) mean vector of the process is defined as

$$E\,(\boldsymbol{y}_t) \tag{2.2}$$

while the autocovariance matrices are defined as

$$E\,\{(\boldsymbol{y}_t - E\,(\boldsymbol{y}_t))\,(\boldsymbol{y}_\tau - E\,(\boldsymbol{y}_\tau))'\} \tag{2.3}$$

It is evident that formula (2.3) describes a family of functions when the pair of indices t and τ varies.

Restricting the attention to the principal moments, namely the mean vector and the autocovariance matrices, paves the way to the various notions of stationarity which enjoy prominent interest in econometrics.

In this connection, let us give the following definitions

Definition 1 – Stationary Processes

A stochastic process is called stationary insofar as – at least to some extent – it exhibits characteristics of permanence and satisfies statistical properties which are not affected by a shift in the time origin, which in turn grants some sort of temporal homogeneity (see, e.g., Blanc-Lapierre and Fortet 1953; Papoulis 1965).

The notion of stationary can actually assumes a plurality of facets: the ones reported below are of particular interest for the subsequent analysis.

Definition 2 – Stationarity in Mean

A process $\boldsymbol{y}_t$ is said to be stationary in mean if

$$E\,(\boldsymbol{y}_t) = \boldsymbol{\mu} \tag{2.4}$$

where $\boldsymbol{\mu}$ is a time-invariant vector.

Remark

If a process $\boldsymbol{y}_t$ is stationary in mean, the difference process $\nabla \boldsymbol{y}_t$ is itself a stationary process, whose mean is a null vector and vice versa.

Definition 3 – Covariance Stationarity

A process $\boldsymbol{y}_t$ is said to be covariance stationary if (2.3) depends only on the temporal lag $\tau - t$ of the argument processes.

Definition 4 – Stationarity in the Wide Sense

A process $\boldsymbol{y}_t$ is said to be stationary in the wide sense, or *weakly stationary*, when both stationary in mean and in covariance.

For a covariance stationary n-dimensional process the matrix

$$\Gamma(h) = E\{(\boldsymbol{y}_t - \mu)(\boldsymbol{y}_{t+h} - \mu)'\} \tag{2.5}$$

represents the autocovariance matrix of order h. It easy to see that for real processes the following holds

$$\Gamma(-h) = \Gamma'(h) \tag{2.6}$$

The autocorrelation matrix $\boldsymbol{P}(h)$ of order h is the matrix defined as follows

$$\boldsymbol{P}(h) = \boldsymbol{D}^{-1}\,\Gamma(h)\,\boldsymbol{D}^{-1} \tag{2.7}$$

where $\boldsymbol{D}$ is the diagonal matrix

$$\boldsymbol{D} = \begin{bmatrix} \sqrt{\gamma_{11}(0)} & 0 & 0 & 0 \\ 0 & \sqrt{\gamma_{22}(0)} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \sqrt{\gamma_{nn}(0)} \end{bmatrix} \tag{2.8}$$

whose diagonal entries are the standard error of the elements of the vector $\boldsymbol{y}_t$.

The foregoing covers what does really matter about stationarity for our purposes. Moving to non stationary processes, we are mainly interested in the class of so-called integrated processes, which we are now going to define.

Definition 5 – Integrated Processes

An integrated process of order d – written as $I(d)$ – where d is a positive integer, is a process $\boldsymbol{\zeta}_t$ such that it must be differenced d times in order to recover stationarity.

As a by-product of the operator identity

$$\nabla^0 = \boldsymbol{I} \tag{2.9}$$

a process $I(0)$ is trivially stationary.

2.2 Principal Multivariate Stationary Processes

This section displays the outline of principle stochastic processes and derives the closed-forms of their first and second moments.

We begin by introducing some prelimary definitions.

Definition 1 – White Noise

A white noise of dimension n, written as $\mathrm{WN}_{(n)}$, is a process $\boldsymbol{\varepsilon}_t$ with

$$E\,(\boldsymbol{\varepsilon}_t) = \boldsymbol{0} \tag{2.10}$$

$$E\,(\boldsymbol{\varepsilon}_t\ \boldsymbol{\varepsilon}'_\tau) = \delta_{\tau-t}\,\boldsymbol{\Sigma} \tag{2.11}$$

where $\boldsymbol{\Sigma}$ denotes a positive definite time-invariant dispersion matrix, and δ_v is the (discrete) unitary function, that is to say

$$\begin{cases} \delta_v = 1 & if \quad v = 0 \\ \delta_v = 0 & otherwise \end{cases} \tag{2.12}$$

The autocovariance matrices of the process turn out to be given by

$$\boldsymbol{\Gamma}_\varepsilon(h) = \delta_h \boldsymbol{\Sigma} \tag{2.13}$$

with the corollary that the following noteworthy relation holds for the autocovariance matrix of composite vectors (Faliva and Zoia 1999, p 23)

$$E = \left\{ \begin{bmatrix} \boldsymbol{\varepsilon}_t \\ \boldsymbol{\varepsilon}_{t-1} \\ \vdots \\ \boldsymbol{\varepsilon}_{t-q} \end{bmatrix} [\boldsymbol{\varepsilon}'_{t+h}, \boldsymbol{\varepsilon}'_{t-1+h}, \ldots, \boldsymbol{\varepsilon}'_{t-q+h}] \right\} = \boldsymbol{D}_h \otimes \boldsymbol{\Sigma} \tag{2.14}$$

where $\boldsymbol{D}_h$ is a matrix given by

$$\boldsymbol{D}_h = \begin{cases} \boldsymbol{I}_{q+1} & if \quad h = 0 \\ \boldsymbol{J}^h & if \quad 1 \le h \le q \\ (\boldsymbol{J}')^{|h|} & if \quad -q \le h \le -1 \\ \boldsymbol{0}_{q+1} & if \quad |h| > \mathrm{q} \end{cases} \tag{2.15}$$

Here $\boldsymbol{J}$ denotes the first unitary super diagonal matrix (of order $q + 1$), defined as

$$\underset{(q+1,\, q+1)}{\boldsymbol{J}} = [j_{nm}], \quad \text{with} \quad j_{nm} = \begin{cases} 1 & \mathit{if} \quad m = n+1 \\ 0 & \mathit{if} \quad m \neq n+1 \end{cases} \tag{2.16}$$

while $\boldsymbol{J}^h$ and $(\boldsymbol{J}')^h$, stand for, respectively, the unitary super and sub diagonal matrices of order $h = 1, 2, \ldots$.

Definition 2 – Vector Moving-Average Processes

A vector moving-average process of order q, denoted by VMA (q), is a multivariate process specified as follows

$$\underset{(n,1)}{\boldsymbol{y}_t} = \boldsymbol{\mu} + \sum_{j=0}^{q} \boldsymbol{M}_j\, \boldsymbol{\varepsilon}_{t-j}, \;\; \boldsymbol{\varepsilon}_t \sim WN_{(n)} \tag{2.17}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{M}_0, \boldsymbol{M}_1, \ldots, \boldsymbol{M}_q$ are, respectively, a constant vector and constant matrices.

In operator form this process can be expressed as

$$\boldsymbol{y}_t = \boldsymbol{\mu} + \boldsymbol{M}(L)\, \boldsymbol{\varepsilon}_t, \;\; \boldsymbol{M}(L) = \sum_{j=0}^{q} \boldsymbol{M}_j\, L^j \tag{2.18}$$

where L is the lag operator.

A VMA(q) process is weakly stationary, as the following formulas show

$$E\,(\boldsymbol{y}_t) = \boldsymbol{\mu} \tag{2.19}$$

$$\boldsymbol{\Gamma}(h) = \begin{cases} \sum_{j=0}^{q} \boldsymbol{M}_j \boldsymbol{\Sigma}\, \boldsymbol{M}'_j & \mathit{if} \quad h = 0 \\ \sum_{j=0}^{q-h} \boldsymbol{M}_j \boldsymbol{\Sigma}\, \boldsymbol{M}'_{j+h} & \mathit{if} \quad 1 \le h \le q \\ \sum_{j=0}^{q-|h|} \boldsymbol{M}_{j+|h|} \boldsymbol{\Sigma}\, \boldsymbol{M}'_j & \mathit{if} \quad -q \le h \le -1 \\ \boldsymbol{0} & \mathit{if} \quad |h| > q \end{cases} \tag{2.20}$$

Result (2.19) is easily obtained from the properties of the expectation operator and (2.10) above. Results (2.20) can be obtained upon noting that

$$y_t = \mu + [M_0, M_1, ..., M_q] \begin{bmatrix} \varepsilon_t \\ \varepsilon_{t-1} \\ \vdots \\ \varepsilon_{t-q} \end{bmatrix} \tag{2.21}$$

which in view of (2.14) and (2.15) leads to

$$\begin{aligned} \Gamma(h) &= E\left\{ [M_0, M_1, ..., M_q] \begin{bmatrix} \varepsilon_t \\ \varepsilon_{t-1} \\ \vdots \\ \varepsilon_{t-q} \end{bmatrix} [\varepsilon'_{t+h}, \varepsilon'_{t-1+h}, ..., \varepsilon'_{t-q+h}] \begin{bmatrix} M'_0 \\ M'_1 \\ \vdots \\ M'_q \end{bmatrix} \right\} \\ &= [M_0, M_1, ..., M_q]\, E\left\{ \begin{bmatrix} \varepsilon_t \\ \varepsilon_{t-1} \\ \vdots \\ \varepsilon_{t-q} \end{bmatrix} [\varepsilon'_{t+h}, \varepsilon'_{t-1+h}, ..., \varepsilon'_{t-q+h}] \right\} \begin{bmatrix} M'_0 \\ M'_1 \\ \vdots \\ M'_q \end{bmatrix} \\ &= [M_0, M_1, ..., M_q]\, (D_h \otimes \Sigma) \begin{bmatrix} M'_0 \\ M'_1 \\ \vdots \\ M'_q \end{bmatrix} \end{aligned} \tag{2.22}$$

whence (2.20).

It is also of interest to point out the staked version of the autocovariance matrix of order zero, namely

$$vec\, \Gamma(0) = \sum_{j=0}^{q} (M_j \otimes M_j)\ vec\, \Sigma \tag{2.23}$$

The first and second differences of a white noise process happen to play some role in time series econometrics and for that reason we have included definitions and properties in the next few pages.

Actually, such processes can be viewed as special cases of VMA processes, and enjoy the weak stationarity property accordingly, as the following definitions show.

Definition 3 – First Difference of a White Noise

Let the process $\boldsymbol{y}_t$ be specified as a VMA(1) in this fashion

$$\boldsymbol{y}_t = \boldsymbol{M\varepsilon}_t - \boldsymbol{M\varepsilon}_{t-1} \tag{2.24}$$

which is tantamount to saying as a first difference of a $WN_{(n)}$ process

$$\boldsymbol{y}_t = \boldsymbol{M}\nabla\boldsymbol{\varepsilon}_t \tag{2.25}$$

The following hold for the first and second moments of $\boldsymbol{y}_t$

$$E\,(\boldsymbol{y}_t) = \boldsymbol{0} \tag{2.26}$$

$$\Gamma\,(h) = \begin{cases} 2\boldsymbol{M}\Sigma\,\boldsymbol{M}' & if \quad h = 0 \\ -\,\boldsymbol{M}\Sigma\,\boldsymbol{M}' & if \quad |h| = 1 \\ \boldsymbol{0} & otherwise \end{cases} \tag{2.27}$$

as a by-product of (2.19) and (2.20) above.

Such a process can be referred to as an $I\,(-\,1)$ process upon the operator identity

$$\nabla = \nabla^{-(-1)} \tag{2.28}$$

Definition 4 – Second Difference of a White Noise

Let the process $\boldsymbol{y}_t$ be specified as a VMA(2) by

$$\boldsymbol{y}_t = \boldsymbol{M\varepsilon}_t - 2\boldsymbol{M\varepsilon}_{t-1} + \boldsymbol{M\varepsilon}_{t-2} \tag{2.29}$$

which is tantamount to saying as a second difference of a $WN_{(n)}$ process

$$\boldsymbol{y}_t = \boldsymbol{M}\nabla^2\boldsymbol{\varepsilon}_t \tag{2.30}$$

The following hold for the first and second moments of $\boldsymbol{y}_t$

$$E\,(\boldsymbol{y}_t) = \boldsymbol{0} \tag{2.31}$$

$$\boldsymbol{\Gamma}(h) = \begin{cases} 6\boldsymbol{M}\Sigma\boldsymbol{M}' & if \quad h = 0 \\ -4\boldsymbol{M}\Sigma\boldsymbol{M}' & if \quad |h| = 1 \\ \boldsymbol{M}\Sigma\boldsymbol{M}' & if \quad |h| = 2 \\ \boldsymbol{0} & otherwise \end{cases} \tag{2.32}$$

again as a by-product of (2.19) and (2.20) above.

Such a process can be read as an $I(-2)$ process upon the operator identity

$$\nabla^2 = \nabla^{-(-2)} \tag{2.33}$$

Remark

Should q tends to ∞, the VMA(q) process as specified in (2.17) is referred to as an infinite causal – i.e. unidirectional from the present backward to the past – moving average*,* (2.19) and (2.23) are still meaningful expressions, and stationarity is maintained accordingly provided both $\lim_{q\to\infty}\sum_{i=0}^{q}\boldsymbol{M}_i$ and $\lim_{q\to\infty}\sum_{i=0}^{q}\boldsymbol{M}_i \otimes \boldsymbol{M}_i$ exist as matrices with finite entries.

Definition 5 – Vector Autoregressive Processes

A vector autoregressive process of order p*,* written as VAR (p), is a multivariate process $\boldsymbol{y}_t$ specified as follows

$$\underset{(n,1)}{\boldsymbol{y}_t} = \boldsymbol{\eta} + \sum_{j=1}^{p}\boldsymbol{A}_j\boldsymbol{y}_{t-j} + \boldsymbol{\varepsilon}_t, \ \ \boldsymbol{\varepsilon}_t \sim WN_{(n)} \tag{2.34}$$

where $\boldsymbol{\eta}$ and $\boldsymbol{A}_1, \boldsymbol{A}_2, \dots, \boldsymbol{A}_p$, are a constant vector and constant matrices, respectively.

Such a process can be rewritten in operator form as

$$\boldsymbol{A}(L)\,\boldsymbol{y}_t = \boldsymbol{\eta} + \boldsymbol{\varepsilon}_t, \ \ \boldsymbol{A}(L) = \boldsymbol{I}_n - \sum_{j=1}^{p}\boldsymbol{A}_j L^j \tag{2.35}$$

and it turns out to be stationary provided all roots of the characteristic equation

$$det\, \boldsymbol{A}(z) = 0 \tag{2.36}$$

lie outside the unit circle (see, e.g., Lütkepohl 1991). In this circumstance, the polynomial matrix $\boldsymbol{A}^{-1}(z)$ is an analytical (matrix) function about $z = 1$ according to Theorem 4 of Sect. 1.7, and the process admits a causal VMA (∞) representation, namely

$$\boldsymbol{y}_t = \boldsymbol{\omega} + \sum_{\tau=0}^{\infty} \boldsymbol{C}_\tau \ \boldsymbol{\varepsilon}_{t-\tau} \tag{2.37}$$

where the matrices $\boldsymbol{C}_\tau$ are polynomials in the matrices $\boldsymbol{A}_j$ and the vector $\boldsymbol{\omega}$ depends on both the vector $\boldsymbol{\eta}$ and the matrices $\boldsymbol{C}_\tau$. Indeed the following hold

$$\boldsymbol{A}^{-1}(L) = \boldsymbol{C}\,(L) = \sum_{\tau=0}^{\infty} \boldsymbol{C}_\tau L^\tau \tag{2.38}$$

$$\boldsymbol{\omega} = \boldsymbol{A}^{-1}(L)\ \boldsymbol{\eta} = (\sum_{\tau=0}^{\infty} \boldsymbol{C}_\tau\ )\ \boldsymbol{\eta} \tag{2.39}$$

and the expressions for the matrices $\boldsymbol{C}_\tau$ can be obtained, by virtue of the isomorphism between polynomials in the lag operator L and in a complex variable z, from the identity

$$\begin{aligned} \boldsymbol{I} &= (\boldsymbol{C}_0 + \boldsymbol{C}_1 z + \boldsymbol{C}_2 z^2 + \ldots)\,(\boldsymbol{I} - \boldsymbol{A}_1 z + \ldots - \boldsymbol{A}_p z^p) \\ &= \boldsymbol{C}_0 + (\boldsymbol{C}_1 - \boldsymbol{C}_0\boldsymbol{A}_1)\, z + (\boldsymbol{C}_2 - \boldsymbol{C}_1\boldsymbol{A}_1 - \boldsymbol{C}_0\boldsymbol{A}_2)\, z^2\ \ldots,+ \end{aligned} \tag{2.40}$$

which implies the relationships

$$\begin{cases} \boldsymbol{I} = \boldsymbol{C}_0 \\ \boldsymbol{0} = \boldsymbol{C}_1 - \boldsymbol{C}_0\boldsymbol{A}_1 \\ \boldsymbol{0} = \boldsymbol{C}_2 - \boldsymbol{C}_1\boldsymbol{A}_1 - \boldsymbol{C}_0\boldsymbol{A}_2 \\ \qquad \ldots \end{cases} \tag{2.41}$$

The following recursive equations ensue as a by-product

$$\boldsymbol{C}_\tau = \sum_{j=1}^{\tau} \boldsymbol{C}_{\tau-j}\boldsymbol{A}_j \tag{2.42}$$

The case $p = 1$, which we are going to examine in some details, is of special interest not so much in itself but because of the isomorphic relationship between polynomial matrices and companion matrices (see, e.g., Banjeree et al., Lancaster and Tismenesky) which allows to bring a VAR model of arbitrary order back to an equivalent first order VAR model, after a proper reparametrization.

With this premise, consider a first order VAR model specified as follows

$$\boldsymbol{y}_t = \boldsymbol{\eta} + \boldsymbol{A}\boldsymbol{y}_{t-1} + \boldsymbol{\varepsilon}_t \quad \boldsymbol{\varepsilon}_t \sim WN_{(n)} \tag{2.43}$$

where $\boldsymbol{A}$ stands for $\boldsymbol{A}_1$.

The stationarity condition in this case entails that the matrix $\boldsymbol{A}$ is stable, i.e. all its eigenvalues lie inside the unit circle (see in this connection the considerations dealt with in Appendix A).

The useful expansion (see, e.g., Faliva 1987, p 77)

$$(\boldsymbol{I} - \boldsymbol{A})^{-1} = \boldsymbol{I} + \sum_{h=1}^{\infty} \boldsymbol{A}^h \tag{2.44}$$

holds accordingly, and the related expansions

$$(\boldsymbol{I} - \boldsymbol{A}z)^{-1} = \boldsymbol{I} + \sum_{h=1}^{\infty} \boldsymbol{A}^h z^h, \ |z| \leq 1 \Leftrightarrow (\boldsymbol{I} - \boldsymbol{A}L)^{-1} = \boldsymbol{I} + \sum_{h=1}^{\infty} \boldsymbol{A}^h L^h \tag{2.45}$$

$$[\boldsymbol{I}_{n^2} - \boldsymbol{A} \otimes \boldsymbol{A}]^{-1} = \boldsymbol{I}_{n^2} + \sum_{h=1}^{\infty} \boldsymbol{A}^h \otimes \boldsymbol{A}^h \tag{2.46}$$

ensue as by-products.

By virtue of (2.45) the VMA (∞) representation of the process (2.43) takes the form

$$\boldsymbol{y}_t = \boldsymbol{\omega} + \boldsymbol{\varepsilon}_t + \sum_{\tau=1}^{\infty} \boldsymbol{A}^{\tau} \boldsymbol{\varepsilon}_{t-\tau} \tag{2.47}$$

where

$$\boldsymbol{\omega} = (\boldsymbol{I} - \boldsymbol{A})^{-1} \boldsymbol{\eta} \tag{2.48}$$

and the principle moments of the process may be derived accordingly. For what concerns the mean vector, taking expectations of both sides of (2.47) yields

$$E(\boldsymbol{y}_t) = \boldsymbol{\omega} \tag{2.49}$$

As far as the autocovariances are concerned, observe first that the following remarkable staked form for the autocovariance of order zero

$$vec\, \boldsymbol{\Gamma}(0) = (\boldsymbol{I}_{n^2} - \boldsymbol{A} \otimes \boldsymbol{A})^{-1}\, vec\, \boldsymbol{\Sigma} \tag{2.50}$$

holds true because of (2.46) as a special case of (2.23) once $\boldsymbol{M}_0$ is replaced by $\boldsymbol{I}$, and $\boldsymbol{M}_j$ is replaced by $\boldsymbol{A}^j$ and we let q tend to ∞.

Bearing in mind (2.20) and letting q tend to ∞, simple computations lead to find the following expressions for the higher order autocovariance matrices

$$\boldsymbol{\Gamma}(h) = \boldsymbol{\Gamma}(0)(\boldsymbol{A}')^h \quad for \;\; h > 0 \tag{2.51}$$

$$\boldsymbol{\Gamma}(h) = \boldsymbol{A}^{|h|}\, \boldsymbol{\Gamma}(0) \quad for \;\; h < 0 \tag{2.52}$$

so that the recursive formulas

$$\boldsymbol{\Gamma}(h) = \boldsymbol{\Gamma}(h-1)\, \boldsymbol{A}' \quad for \;\; h > 0 \tag{2.53}$$

$$\boldsymbol{\Gamma}(h) = \boldsymbol{A}\boldsymbol{\Gamma}'(h-1) \quad for \;\; h < 0 \tag{2.54}$$

follow as a by-product.

The extensions of the conclusions just reached about higher order VAR processes, rely on the aforementioned companion-form analogue.

The stationary condition on the roots of the characteristic polynomial quoted for a VAR model has a mirror image in the so-called invertibility condition of a VMA model. In this connection we give the following definition.

Definition 6 – Invertible Processes

A VMA process is invertible if all roots of the characteristic equation

$$det\, \boldsymbol{M}(z) = 0 \tag{2.55}$$

lie outside the unit circle. In this case the matrix $\boldsymbol{M}^{-1}(z)$ is an analytical matrix function about $z = 1$ by Theorem 4 of Sect. 1.7, and therefore the process admits a (unique) representation as a function of its past, in the form of a VAR model.

Emblematic examples of non invertible VMA processes were given in Definitions 3 and 4 above.

One should be aware of the fact that it is immaterial to draw a distinction between invertible and non invertible processes for what concerns stationarity.

The property of invertibility is clearly related to the possibility of making predictions since it allows the process $\boldsymbol{y}_t$ to be specified as a convergent function of past random variables.

Should a VMA process be invertible according to Definition 6 above, the following VMA vs. VAR representation holds

$$\boldsymbol{y}_t = \boldsymbol{\mu} + \sum_{j=0}^{q} \boldsymbol{M}_j \, \boldsymbol{\varepsilon}_{t-j} \Rightarrow \boldsymbol{G}(L)\, \boldsymbol{y}_t = \boldsymbol{\nu} + \boldsymbol{\varepsilon}_t \tag{2.56}$$

where

$$\boldsymbol{\nu} = \boldsymbol{M}^{-1}(L)\, \boldsymbol{\mu} \tag{2.57}$$

$$\boldsymbol{G}(L) = \sum_{\tau=0}^{\infty} \boldsymbol{G}_\tau L^\tau = \boldsymbol{M}^{-1}(L) \tag{2.58}$$

The matrices $\boldsymbol{G}_\tau$ may be obtained through the recursive equations

$$\boldsymbol{G}_\tau = \boldsymbol{M}_\tau - \sum_{j=1}^{\tau-1} \boldsymbol{G}_{\tau-j}\, \boldsymbol{M}_j, \;\; \boldsymbol{G}_0 = \boldsymbol{M}_0 = \boldsymbol{I} \tag{2.59}$$

which are the mirror image of the recursive equations (2.42) and can be obtained in a similar manner.

Taking $q = 1$ in formula (2.17) yields a VMA (1) model specified as

$$\underset{(n,\,1)}{\boldsymbol{y}_t} = \boldsymbol{\mu} + \boldsymbol{M}\boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\varepsilon}_t, \;\; \boldsymbol{\varepsilon}_t \sim WN_{(n)} \tag{2.60}$$

where $\boldsymbol{M}$ stands for $\boldsymbol{M}_1$.

The following hold for the first and second moments in light of (2.19) and (2.20)

$$E(\boldsymbol{y}_t) = \boldsymbol{\mu} \tag{2.61}$$

$$\boldsymbol{\Gamma}(h) = \begin{cases} \Sigma + \boldsymbol{M}\,\Sigma\boldsymbol{M}' & if \quad h = 0 \\ \Sigma\boldsymbol{M}' & if \quad h = 1 \\ \boldsymbol{M}\Sigma & if \quad h = -1 \\ \boldsymbol{0} & if \quad h > 1 \end{cases} \tag{2.62}$$

The invertibilty condition in this case entails that the matrix $\boldsymbol{M}$ is stable, that is to say all its eigenvalues lie inside the unit circle.

The following noteworthy expansions

$$(\boldsymbol{I}+\boldsymbol{M})^{-1}=\boldsymbol{I}+\sum_{\tau=1}^{\infty}(-1)^{\tau}\boldsymbol{M}^{\tau} \tag{2.63}$$

$$(\boldsymbol{I}+\boldsymbol{M}z)^{-1}=\boldsymbol{I}+\sum_{\tau=1}^{\infty}(-1)^{\tau}\boldsymbol{M}^{\tau}z^{\tau} \Leftrightarrow (\boldsymbol{I}+\boldsymbol{M}L)^{-1}=\boldsymbol{I}+\sum_{\tau=1}^{\infty}(-1)^{\tau}\boldsymbol{M}^{\tau}L^{\tau} \tag{2.64}$$

where $|z| \leq 1$, hold for the same arguments as (2.44) and (2.45) above.

As a consequence of (2.64), the VAR representation of the process (2.60) takes the form

$$\boldsymbol{y}_t+\sum_{\tau=1}^{\infty}(-1)^{\tau}\boldsymbol{M}^{\tau}\boldsymbol{y}_{t-\tau}=\boldsymbol{\nu}+\boldsymbol{\varepsilon}_t \tag{2.65}$$

where

$$\boldsymbol{\nu}=(\boldsymbol{I}+\boldsymbol{M})^{-1}\boldsymbol{\mu} \tag{2.66}$$

Let us now introduce VARMA models which engender processes combining the characteristics of both VMA and VAR specifications.

Definition 7 – Vector Autoregressive Moving-Average Processes

A vector autoregressive moving-average process of orders p and q (where p is the order of the autoregressive component and q is the order of the moving-average component) – written as VARMA(p, q) – is a multivariate process $\boldsymbol{y}_t$ specified as follows

$$\underset{(n,1)}{\boldsymbol{y}_t}=\boldsymbol{\eta}+\sum_{j=1}^{p}\boldsymbol{A}_j\boldsymbol{y}_{t-j}+\sum_{j=0}^{q}\boldsymbol{M}_j\boldsymbol{\varepsilon}_{t-j},\ \ \boldsymbol{\varepsilon}_t \sim WN_{(n)} \tag{2.67}$$

where $\boldsymbol{\eta}$, $\boldsymbol{A}_j$ and $\boldsymbol{M}_j$ are a constant vector and constant matrices, respectively.

In operator form the process can be written as follows

$$\boldsymbol{A}(L)\,\boldsymbol{y}_t=\boldsymbol{\eta}+\boldsymbol{M}(L)\,\boldsymbol{\varepsilon}_t, \tag{2.68}$$

$$\boldsymbol{A}(L) = \boldsymbol{I}_n - \sum_{j=1}^{p} \boldsymbol{A}_j L^j, \ \boldsymbol{M}(L) = \sum_{j=0}^{q} \boldsymbol{M}_j L^j$$

The process is stationary if all roots of the characteristic equation of its autoregressive part, i.e.

$$det\, \boldsymbol{A}(z) = 0 \tag{2.69}$$

lie outside the unit circle. When this is the case, the matrix $\boldsymbol{A}^{-1}(z)$ is an analytical function in a neighbourhood of $z = 1$ by Theorem 4 in Sect. 1.7 and the process admits a causal VMA (∞) representation, namely

$$\boldsymbol{y}_t = \boldsymbol{\omega} + \sum_{\tau=0}^{\infty} \boldsymbol{C}_\tau \boldsymbol{\varepsilon}_{t-\tau} \tag{2.70}$$

where the matrices $\boldsymbol{C}_\tau$ are polynomials in the matrices $\boldsymbol{A}_j$ and $\boldsymbol{M}_j$ while the vector $\boldsymbol{\omega}$ depends on both the vector $\boldsymbol{\eta}$ and the matrices $\boldsymbol{A}_j$. Indeed, the following hold

$$\boldsymbol{\omega} = \boldsymbol{A}^{-1}(L)\, \boldsymbol{\eta} \tag{2.71}$$

$$\boldsymbol{C}(L) = \sum_{\tau=0}^{\infty} \boldsymbol{C}_\tau L^\tau = \boldsymbol{A}^{-1}(L)\, \boldsymbol{M}(L) \tag{2.72}$$

which, in turn, leads to the recursive formulas

$$\boldsymbol{C}_\tau = \boldsymbol{M}_\tau + \sum_{j=1}^{\tau} \boldsymbol{A}_j \boldsymbol{C}_{\tau-j}\ , \ \boldsymbol{C}_0 = \boldsymbol{M}_0 = \boldsymbol{I} \tag{2.73}$$

As far as the invertibility property is concerned, reference must be made to the VMA component of the process. The process is invertible if all roots of the characteristic equation

$$det\, \boldsymbol{M}(z) = 0 \tag{2.74}$$

lie outside the unit circle. Then again the matrix $\boldsymbol{M}^{-1}(L)$ is an analytical function in a neighbourhood of $z = 1$ by Theorem 4 in Sect. 1.7, and the VARMA process admits a VAR (∞) representation such as

$$\boldsymbol{G}(L)\, \boldsymbol{y}_t = \boldsymbol{\nu} + \boldsymbol{\varepsilon}_t \tag{2.75}$$

where

$$\boldsymbol{\nu} = \boldsymbol{M}^{-1}(L)\, \boldsymbol{\eta} \tag{2.76}$$

$$\boldsymbol{G}(L) = \sum_{\tau=0}^{\infty} \boldsymbol{G}_\tau L^\tau = \boldsymbol{M}^{-1}(L)\, \boldsymbol{A}(L) \tag{2.77}$$

and the matrices $\boldsymbol{G}_\tau$ may be computed through the recursive equations

$$\boldsymbol{G}_\tau = \boldsymbol{M}_\tau + \boldsymbol{A}_\tau - \sum_{j=1}^{\tau-1} \boldsymbol{M}_{\tau-j}\, \boldsymbol{G}_j,\ \boldsymbol{G}_0 = \boldsymbol{M}_0 = \boldsymbol{I} \tag{2.78}$$

Letting $p = q = 1$ in formula (2.67) yields a VARMA (1,1) specified in this way

$$\boldsymbol{y}_t = \boldsymbol{\eta} + \boldsymbol{A}\boldsymbol{y}_{t-1} + \boldsymbol{\varepsilon}_t + \boldsymbol{M}\boldsymbol{\varepsilon}_{t-1},\ \boldsymbol{A} \neq -\boldsymbol{M} \tag{2.79}$$

where $\boldsymbol{A}$ and $\boldsymbol{M}$ stand for $\boldsymbol{A}_1$ and $\boldsymbol{M}_1$ respectively, and the parameter requirement $\boldsymbol{A} \neq -\boldsymbol{M}$ is introduced in order to rule out the degenerate case of a first order dynamic model collapsing into that of order zero.

In this case the stationary condition is equivalent to assuming that the matrix $\boldsymbol{A}$ is stable whereas the invertibilty condition requires the stability of matrix $\boldsymbol{M}$.

Under stationarity, the following holds

$$\boldsymbol{y}_t = (\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{\eta} + (\boldsymbol{I} + \sum_{\tau=1}^{\infty} \boldsymbol{A}^\tau L^\tau)(\boldsymbol{I} + \boldsymbol{M}L)\boldsymbol{\varepsilon}_t \tag{2.80}$$

which tallies with the VMA (∞) representation (2.70) once we put

$$\boldsymbol{\omega} = (\boldsymbol{I} + \sum_{\tau=1}^{\infty} \boldsymbol{A}^\tau)\boldsymbol{\eta} \tag{2.81}$$

$$\boldsymbol{C}_\tau = \begin{cases} \boldsymbol{I} & if \quad \tau = 0 \\ \boldsymbol{A} + \boldsymbol{M} & if \quad \tau = 1 \\ \boldsymbol{A}^{\tau-1}(\boldsymbol{A} + \boldsymbol{M}) & if \quad \tau > 1 \end{cases} \tag{2.82}$$

Under invertibility, the following holds

$$(\boldsymbol{I} + \boldsymbol{M}L)^{-1}(\boldsymbol{y}_t - \boldsymbol{A}\boldsymbol{y}_{t-1}) = (\boldsymbol{I} + \boldsymbol{M})^{-1}\boldsymbol{\eta} + \boldsymbol{\varepsilon}_t \tag{2.83}$$

which tallies with the VAR(∞) representation (2.75) once we put

$$\boldsymbol{v} = \left(\boldsymbol{I} + \sum_{\tau=1}^{\infty} (-1)^\tau \boldsymbol{M}^\tau \right)\boldsymbol{\eta} \tag{2.84}$$

$$G_\tau = \begin{cases} \boldsymbol{I} & \text{if} \quad \tau = 0 \\ -\boldsymbol{M} - \boldsymbol{A} & \text{if} \quad \tau = 1 \\ -(-1)^{\tau-1}\boldsymbol{M}^{\tau-1}(\boldsymbol{M}+\boldsymbol{A}) & \text{if} \quad \tau > 1 \end{cases} \tag{2.85}$$

In order to derive the autocovariance matrices of a general VARMA (p, q) model of dimension n one may transform it into a VAR (1) model by virtue of the already mentioned companion form analogue.

So far we have considered only VAR and VARMA models, whose characteristic polynomial roots lie outside the unit circle.

Nevertheless, the case of a possibly repeated unit-root is worth considering also. As a matter of fact, this proves to stand as a gateway bridging the gap between stationarity and integrated processes as the Sect. 2.3 will clarify.

2.3 The Source of Integration and the Seeds of Cointegration

In this section we set out two theorems which bring to the fore the link between the unit-roots of a VAR model and the integration order of the engendered process and disclose the two-face nature of the model solution with cointegration finally appearing on stage.

Theorem 1

The order of integration of the process $\boldsymbol{y}_t$ *generated by a VAR model*

$$\boldsymbol{A}(L)\,\boldsymbol{y}_t = \boldsymbol{\eta} + \boldsymbol{\varepsilon}_t \tag{2.86}$$

whose characteristic polynomial has a possibly repeated unit-root, is the same as the degree of the principal part, i.e. the order of the pole, in the Laurent expansion for $\boldsymbol{A}^{-1}(z)$ *in a deleted neighbourhood of* $z = 1$.

Proof

A particular solution of the operational equation (2.86) is given by

$$\boldsymbol{y}_t = \boldsymbol{A}^{-1}(L)\,(\boldsymbol{\varepsilon}_t + \boldsymbol{\eta}) \tag{2.87}$$

By virtue of the isomorphism existing between the polynomials in the lag operator L and in a complex variable z (see, e.g., Dhrymes, p 23), the following holds

$$A^{-1}(z) \Leftrightarrow A^{-1}(L) \tag{2.88}$$

and the paired expansions

$$\sum_{j=1}^{\upsilon} \frac{1}{(1-z)^j} N_j + \sum_{i=0}^{\infty} z^i M_i \Leftrightarrow \sum_{j=1}^{\upsilon} \frac{1}{(I-L)^j} N_j + \sum_{i=0}^{\infty} L^i M_i \tag{2.89}$$

where υ stands for the order of the pole of $A^{-1}(z)$ at $z = 1$, are also true.

Because of (2.89) and by making use of sum-calculus identities such as

$$(I-L)^{-j} = \nabla^{-j} \quad j = 0, 1, 2, \ldots \tag{2.90}$$

where in particular (see (1.382) and (1.383) of Sect. 1.8)

$$\nabla^{-1} = \sum_{\tau \le t}, \; \nabla^{-2} = \sum_{\vartheta \le t} \sum_{\tau \le \vartheta} \tag{2.91}$$

the right-hand side of (2.87) can be given the informative expression

$$\begin{aligned} A^{-1}(L)(\varepsilon_t + \eta) &= (N_1\nabla^{-1} + N_2\nabla^{-2} + \ldots + N_K\nabla^{-\upsilon} + \sum_{j=0}^{\infty} M_j L^j)(\varepsilon_t + \eta) \\ &= N_1 \sum_{\tau \le t} \varepsilon_\tau + N_2 \sum_{\vartheta \le t} \sum_{\tau \le \vartheta} \varepsilon_\tau + \ldots + \sum_{j=0}^{\infty} M_j \varepsilon_{t-j} + N_1 \sum_{\tau \le t} \eta \\ &+ N_2 \sum_{\tau \le t} (t + 1 - \tau)\eta + \ldots + \sum_{j=0}^{\infty} M_j \eta \end{aligned} \tag{2.92}$$

By inspection of (2.92) the conclusion is easily drawn that the process engendered by the VAR model (1) is composed – stationary components apart – of integrated processes of progressive order.

Hence, the overall effect is that the solution y_t turns out to be an integrated process itself, whose order is the same as the order of the pole of $A^{-1}(z)$, that is to say

$$y_t \sim I(\upsilon) \tag{2.93}$$

□

Theorem 2

Let $z = 1$ be a possibly repeated root of the characteristic polynomial $\det A(z)$ of the VAR model

$$A(L)\, y_t = \eta + \varepsilon_t \tag{2.94}$$

and its solution $\boldsymbol{y}_t$ *be, correspondingly, an integrated process, say* $\boldsymbol{y}_t \sim I\,(d)$ *for some* $d > 0$.

Furthermore, let

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}' \tag{2.95}$$

be a rank factorization of the singular matrix $\boldsymbol{A}\,(1) = \boldsymbol{A}$.

Then the following decomposition holds

$$\boldsymbol{y}_t = \underbrace{(\boldsymbol{C}'_\perp)^g\,\boldsymbol{C}'_\perp \boldsymbol{y}_t}_{\text{maintained integrated component}} + \underbrace{(\boldsymbol{C}')^g\,\boldsymbol{C}'\boldsymbol{y}_t}_{\text{degenerate integrated component}} \tag{2.96}$$

where the maintained and degenerate components enjoy the integration properties

$$(\boldsymbol{C}'_\perp)^g\,\boldsymbol{C}'_\perp \boldsymbol{y}_t \sim I\,(d) \tag{2.97}$$

$$(\boldsymbol{C}')^g\,\boldsymbol{C}'\boldsymbol{y}_t \sim I\,(\delta),\ \ \delta \le d-1 \tag{2.98}$$

respectively.

The notion of cointegration fits with the process $\boldsymbol{y}_t$ *accordingly.*

Proof

In light of (1.247) of Sect. 1.6 and of isomorphism between polynomials in a complex variable z and in the lag operator L, the VAR model (2.94) can be rewritten in the more convenient form

$$\boldsymbol{Q}\,(L)\,\nabla \boldsymbol{y}_t + \boldsymbol{B}\boldsymbol{C}'\boldsymbol{y}_t = \boldsymbol{\eta} + \boldsymbol{\varepsilon}_t \tag{2.99}$$

where $\boldsymbol{Q}\,(z)$ is as defined by (1.248) of Sect. 1.6, and $\boldsymbol{B}$ and $\boldsymbol{C}$ are defined in (2.95).

Upon noting that

$$\boldsymbol{y}_t \sim I\,(d) \Rightarrow \nabla \boldsymbol{y}_t \sim I\,(d-1) \Rightarrow \boldsymbol{Q}\,(L)\,\nabla \boldsymbol{y}_t \sim I\,(\delta),\, \delta \le d-1 \tag{2.100}$$

the conclusion

$$\boldsymbol{C}'\boldsymbol{y}_t \sim I\,(\delta) \Leftrightarrow (\boldsymbol{C}')^g\,\boldsymbol{C}'\boldsymbol{y}_t \sim I\,(\delta) \tag{2.101}$$

is easily drawn, given that

$$\begin{aligned} \boldsymbol{B}\boldsymbol{C}'\boldsymbol{y}_t = -\,\boldsymbol{Q}\,(L)\nabla \boldsymbol{y}_t + \boldsymbol{\eta} + \boldsymbol{\varepsilon}_t \Leftrightarrow \boldsymbol{C}'\boldsymbol{y}_t \\ = -\boldsymbol{B}^g\,\boldsymbol{Q}\,(L)\nabla \boldsymbol{y}_t + \boldsymbol{B}^g\boldsymbol{\eta} + \boldsymbol{B}^g\boldsymbol{\varepsilon}_t \end{aligned} \tag{2.102}$$

by (2.99) and the integration order of $-\boldsymbol{B}^g \boldsymbol{Q}(L)\nabla \boldsymbol{y}_t + \boldsymbol{B}^g \boldsymbol{\eta} + \boldsymbol{B}^g \boldsymbol{\varepsilon}_t$ is at most that of $\boldsymbol{Q}(L)\nabla \boldsymbol{y}_t$, namely $\delta \leq d - 1$.

Insofar as a drop of integration order occurs when moving from the parent process $\boldsymbol{y}_t$ to its component $(\boldsymbol{C}')^g \boldsymbol{C}'\boldsymbol{y}_t$, the latter is a degenerate process with respect to the former.

The analysis of the degenerate component $(\boldsymbol{C}')^g \boldsymbol{C}'\boldsymbol{y}_t$ being accomplished, let us examine the complementary component $(\boldsymbol{C}'_\perp)^g \boldsymbol{C}'_\perp \boldsymbol{y}_t$.

To this end, observe that by virtue of (1.52) of Sect. 1.2, the following identity

$$\boldsymbol{I} = (\boldsymbol{C}'_\perp)^g \boldsymbol{C}'_\perp + (\boldsymbol{C}')^g \boldsymbol{C}' \tag{2.103}$$

holds true and, in turn, leads us to split $\boldsymbol{y}_t$ into two components, as shown in (2.96).

Since the following integration properties

$$\boldsymbol{y}_t \sim I(d) \tag{2.104}$$

$$(\boldsymbol{C}')^g \boldsymbol{C}'\boldsymbol{y}_t \sim I(\delta) \tag{2.105}$$

hold in light of the foregoing, the conclusion that the component $(\boldsymbol{C}'_\perp)^g \boldsymbol{C}'_\perp \boldsymbol{y}_t$ maintains the integration order inherent in the parent process $\boldsymbol{y}_t$, that is to say

$$(\boldsymbol{C}'_\perp)^g \boldsymbol{C}'_\perp \boldsymbol{y}_t \sim I(d) \tag{2.106}$$

is eventually drawn.

Finally, in light of (2.105) and (2.106), with (2.104) as a benchmark, the seeds of the concept of cointegration – whose notion and role will receive considerable attention in Sect. 2.4 and in Chap. 3 – are sown.

□

2.4 Integrated and Cointegrated Processes

We start with introducing the basic notions concerning both integrated and cointegrated processes along with some related results.

Definition 1– Random Walk

A n-dimensional random-walk is a multivariate $I(1)$ process $\boldsymbol{\xi}_t$ such that

$$\underset{(n,1)}{\nabla \boldsymbol{\xi}_t} = \boldsymbol{\varepsilon}_t, \boldsymbol{\varepsilon}_t \sim WN_{(n)} \tag{2.107}$$

The following representations

$$\boldsymbol{\xi}_t = \sum_{\tau \le t} \boldsymbol{\varepsilon}_\tau \tag{2.108}$$

$$= \boldsymbol{\xi}_0 + \sum_{\tau = 1}^{t} \boldsymbol{\varepsilon}_\tau \tag{2.108'}$$

hold accordingly, where $\boldsymbol{\xi}_0$ stands for an initial condition vector, independent from $\boldsymbol{\varepsilon}_t$, $t > 0$, and assumed to have zero mean and finite second moments (see, e.g., Hatanaka 1996).

The process, while stationary in mean, namely

$$E(\boldsymbol{\xi}_t) = 0 \tag{2.109}$$

is not covariance stationary, because of

$$E(\boldsymbol{\xi}_t \boldsymbol{\xi}_t') = E(\boldsymbol{\xi}_0 \boldsymbol{\xi}_0') + \boldsymbol{\Gamma}_\varepsilon(0)\, t \tag{2.110}$$

as a simple computation shows.

Definition 2 – Random Walk with Drift

A random walk with drift is a multivariate $I(1)$ process $\boldsymbol{\xi}_t$ defined as follows

$$\nabla \boldsymbol{\xi}_t = \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t, \ \boldsymbol{\varepsilon}_t \sim WN_{(n)} \tag{2.111}$$

where $\boldsymbol{\mu}$ is a drift vector.

The representation

$$\boldsymbol{\xi}_t = \boldsymbol{\xi}_0 + \boldsymbol{\mu} t + \sum_{\tau = 1}^{t} \boldsymbol{\varepsilon}_\tau \tag{2.112}$$

holds true, where $\boldsymbol{\xi}_0$ is a random vector depending on the initial conditions and independent from $\boldsymbol{\varepsilon}_t$, $t > 0$. Moreover, $\boldsymbol{\xi}_0$ is assumed to have first and second moments both finite.

A process of this nature is neither stationary in mean nor in covariance, as simple computations show. In fact

$$E(\boldsymbol{\xi}_t) = E(\boldsymbol{\xi}_0) + \boldsymbol{\mu} t \tag{2.113}$$

$$\boldsymbol{V}(\boldsymbol{\xi}_t) = \boldsymbol{V}(\boldsymbol{\xi}_0) + \boldsymbol{\Gamma}_\varepsilon(0)\, t \tag{2.114}$$

where $\boldsymbol{V}$ stands for covariance matrix.

The notion of random walk can be generalized to cover processes whose k-order difference, $k > 1$, leads back to a white noise process.

In this connection, we give the following definition (see also Hansen and Johansen 1998, p 110).

Definition 3 – Cumulated Random Walk

By a cumulated random walk we mean a multivariate $I\,(2)$ process defined after the property

$$\nabla^2 \boldsymbol{\xi}_t = \boldsymbol{\varepsilon}_t,\ \boldsymbol{\varepsilon}_t \sim WN_{(n)} \tag{2.115}$$

The following representations

$$\boldsymbol{\xi}_t = \sum_{\vartheta \le t} \sum_{\tau \le \vartheta} \boldsymbol{\varepsilon}_\tau \tag{2.116}$$

$$= \sum_{\tau \le t} (t + 1 - \tau)\, \boldsymbol{\varepsilon}_\tau \tag{2.116'}$$

$$= \sum_{\tau \le 0} (\tau + 1)\, \boldsymbol{\varepsilon}_{t-\tau} \tag{2.116''}$$

hold true, and the analysis of the process can be carried out along the same line as in Definition 1.

Cumulated random walks with drift can be likewise defined along the lines traced in Definition 2.

Inasmuch as an analogue signal vs. noise (in system theory), and trend vs. disturbances (in time series analysis) is established, and noise as well as disturbances stand for non systematic nuisance components, the term signal or trend fits in with any component which exhibits either a regular time path or evolving stochastic swings. Whence the notions of deterministic and stochastic trends which follow.

Definition 4 – Deterministic Trends

The term deterministic trend will be henceforth used to indicate polynomial functions in the time variable, namely

$$f_t = at + bt^2 + \dots + dt^r \tag{2.117}$$

where r is a positive integer and a, b ..., d denote parameters.

Linear (first-order) and quadratic (second-order) deterministic trends turn out to be of major interest for time series econometrics owing to their connection with random walks with drifts.

Definition 5 – Stochastic Trends

By a stochastic trend we mean a vector $\boldsymbol{\varphi}_t$ defined as

$$\boldsymbol{\varphi}_t = \sum_{\tau=1}^{t} \boldsymbol{\varepsilon}_\tau , \quad \boldsymbol{\varepsilon}_t \sim WN_{(n)} \tag{2.118}$$

Upon noting that

$$\nabla \boldsymbol{\varphi}_t = \boldsymbol{\varepsilon}_t \tag{2.119}$$

the notion of stochastic trend turns out to mirror that of random walk.

Remark

If reference is made to a cumulated random walk, as specified by (2.115), we can analogously define a second order stochastic trend in this manner

$$\boldsymbol{\varphi}_t = \sum_{\vartheta=1}^{t} \sum_{\tau=1}^{\vartheta} \boldsymbol{\varepsilon}_\tau , \quad \boldsymbol{\varepsilon}_t \sim WN_{(n)} \tag{2.120}$$

Should a drift enter the underlying random walk specification, a trend mixing stochastic and deterministic features would occur.

The material covered so far has provided the reader with a first glance at integrated processes.

A deeper insight into the subject matter, in connection with the analysis of VAR models with unit roots, will be gained in Chap. 3.

When dealing with several integrated processes, the question may be raised as to whether it would be possible to recover stationarity – besides trivially differencing the given processes – by some sort of a clearing-house mechanism, capable to lead non-stationarities to balance each others out (at least to some degree).

This idea is at the root of cointegration theory that looks for those linear forms of stochastic processes with preassigned integration orders which turn out to be more stationary – possibly, stationary *tout court* – than the original ones.

Here below we will give a few basic notions about cointegration, postponing a closer scrutiny of this fascinating topic to Chap. 3.

Definition 6 – Cointegrated Systems

The components of a multivariate integrated process $\boldsymbol{y}_t$ form a cointegrated system of order (d, b), with d and b non negative integer numbers such that $d \geq b$, and we write

$$\boldsymbol{y}_t \sim CI\,(d, b) \tag{2.121}$$

if the following conditions are fulfilled

(1) The n scalar random processes which represent the elements of the vector $\boldsymbol{y}_t$ are integrated of order d, that is to say

$$\underset{(n,1)}{\boldsymbol{y}_t} \sim I\,(d) \tag{2.122}$$

(2) There exist one or more (linearly independent) vectors $\boldsymbol{\alpha}$ neither null nor proportional to an elementary vector, such that the linear form

$$\underset{(1,1)}{x_t} = \boldsymbol{\alpha}'\boldsymbol{y}_t \tag{2.123}$$

is integrated of order $d - b$, i.e.

$$x_t \sim I\,(d - b) \tag{2.124}$$

The vectors $\boldsymbol{\alpha}$ are called cointegration vectors. The number of cointegration vectors, which are linearly independent, identifies the so-called cointegration rank for the process $\boldsymbol{y}_t$.

The basic purpose of cointegration is that of describing the stable relations of the economy through linear relations which are more stationary than the variables under consideration.

Observe, in particular, that the class of $CI\,(1, 1)$ processes is that of $I\,(1)$ processes which by cointegration give rise to stationary processes.

Definition (6) can be extended to the case of a possibly different order of integration for the components of the vector $\boldsymbol{y}_t$ (see, e.g., Charenza and Deadman 1992).

In practice, conditions (1) and (2) can be reformulated in this way

(1) The variables $y_{t1}, y_{t2}, \ldots, y_{tn}$, which represent the elements of the vector $\boldsymbol{y}_t$, are integrated of (possibly) different orders d_h $(h = 1, 2, ..., K)$, with $d_1 > d_2, \ldots, > d_K \geq b$, and these orders are, at least, equal pairwise. By defining the integration order of a vector as the highest integration order of its components, we will simply write

$$\boldsymbol{y}_t \sim I\,(d_1) \tag{2.125}$$

(2) For every subset of (two or more) elements of the vector $\boldsymbol{y}_t$, integrated of the same order, there exists al least one cointegration vector by which we obtain – through a linear combination of the previous ones – a variable that is integrated of an order corresponding to that of another subset of (two or more) elements of $\boldsymbol{y}_t$.

As a result there will exist one or more linearly independent vectors $\boldsymbol{\alpha}$ (encompassing the weights of the said linear combinations), neither null nor proportional to an elementary vector, such that the linear form

$$\underset{(1,1)}{x_t} = \boldsymbol{\alpha}'\boldsymbol{y}_t \tag{2.126}$$

is integrated of order $d_1 - b$, i.e.

$$x_t \sim I(d_1 - b) \tag{2.127}$$

We are now ready to introduce the notion of polynomial cointegration (see, e.g., Johansen 1995).

Definition 7 – Polynomially Cointegrated Systems

The components of a multivariate stochastic process $\boldsymbol{y}_t$ integrated of order $d \geq 2$ form a polynomially cointegrated system of order (d, b), where b is a non negative integer satisfying the condition $b \leq d$, and we write

$$\boldsymbol{y}_t \sim PCI(d, b) \tag{2.128}$$

if there exist vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}_k$ $(1 \leq k \leq d - b + 1)$ – at least one of them, besides $\boldsymbol{\alpha}$, neither null nor proportional to an elementary vector – such that the linear form in levels and differences

$$\underset{(1,1)}{z_t} = \boldsymbol{\alpha}'\boldsymbol{y}_t + \sum_{k=1}^{d-b+1} \boldsymbol{\beta}'_k \nabla^k \boldsymbol{y}_t \tag{2.129}$$

is an integrated process of order $d - b$, i.e.

$$z_t \sim I(d - b) \tag{2.130}$$

Observe, in particular, that the class of $PCI(2, 2)$ processes is that of $I(2)$ processes which by polynomial cointegration give rise to stationary processes.

Cointegration is actually a cornerstone of time series econometrics as Chap. 3 will show. A quick glance at the role of cointegration, in connection with the notion of stochastic trends, will be taken in Appendix B.

2.5 Casting a Glance at the Backstage of VAR Modelling

Wold's theorem (Wold 1938, p 89) states that, in the absence of singular components, a weakly stationary process has a representation as a one-sided moving average in a white noise argument, such as

$$\boldsymbol{\xi}_t = \sum_{i \geq 0} \boldsymbol{M}_i \boldsymbol{\varepsilon}_{t-i}, \quad \boldsymbol{\varepsilon}_t \sim WN \tag{2.131}$$

which, in operational form, can be written as

$$\boldsymbol{\xi}_t = (\sum_{i \geq 0} \boldsymbol{M}_i L^i) \boldsymbol{\varepsilon}_t \tag{2.132}$$

Indeed, as long as the algebra of analytic functions in a complex variable and in the lag operator are isomorphic, the operator in brackets on the right hand-side of (2.132) can be thought of as a formal Taylor expansion in the lag operator L of a parent analytical function $\boldsymbol{\Phi}(z)$, that is

$$\boldsymbol{\Phi}(z) = \sum_{i \geq 0} \boldsymbol{M}_i z^i \rightarrow \boldsymbol{\Phi}(L) = \sum_{i \geq 0} \boldsymbol{M}_i L^i \tag{2.133}$$

Beveridge and Nelson (1981) as well as Stock and Watson (1988) establish a bridgehead beyond Wold's theorem, pioneering a route to evolutive processes through a representation of a first-order integrated processes as the sum of a one-sided moving average (as before) and a random walk process, namely

$$\boldsymbol{\xi}_t = \sum_{i \geq 0} \boldsymbol{M}_i \boldsymbol{\varepsilon}_{t-i} + \boldsymbol{N} \sum_{\tau \leq t} \boldsymbol{\varepsilon}_\tau \tag{2.134}$$

which, in operational form, can be written as

$$\boldsymbol{\xi}_t = (\sum_{i \geq 0} \boldsymbol{M}_i L^i + \boldsymbol{N} \nabla^{-1}) \boldsymbol{\varepsilon}_t \tag{2.135}$$

Still, due to the above mentioned isomorphism, the operator in brackets in the right hand-side of (2.132) can be thought of as a Laurent expansion, about a simple pole, in the lag operator of a parent function $\boldsymbol{\Phi}(z)$ having an isolated singularity, namely a first-order pole, located at $z = 1$, that is

$$\boldsymbol{\Phi}(z) = (\sum_{i \geq 0} \boldsymbol{M}_i z^i + \boldsymbol{N} \frac{1}{(1-z)}) \quad \rightarrow \quad \boldsymbol{\Phi}(L) = (\sum_{i \geq 0} \boldsymbol{M}_i L^i + \boldsymbol{N} \nabla^{-1}) \tag{2.136}$$

A breakthrough in this direction leads to an extensive class of stochastic processes, duly shaping the contours of economic time series investigations, which can be written in the form

$$\xi_t = \boldsymbol{\Phi}(L)(\boldsymbol{\varepsilon}_t + \boldsymbol{\eta}) \tag{2.137}$$

Here $\boldsymbol{\eta}$ is an n-vector of constant terms, $\boldsymbol{\Phi}(L)$ is a matrix function of the lag operator L, isomorphic to its mirror image $\boldsymbol{\Phi}(z)$ in the complex argument z, which has a possibly removable isolated singularity located at $z = 1$. On the foregoing premise, a Laurent expansion such as

$$\boldsymbol{\Phi}(z) = \underbrace{\sum_{i \ge 0} \boldsymbol{M}_i z_i}_{regular\ \ part} + \underbrace{\sum_{j \ge 0} \boldsymbol{N}_j \frac{1}{(1-z)^j}}_{principal\ \ part} \tag{2.138}$$

holds true for $\boldsymbol{\Phi}(z)$ in a deleted neighbourhood of $z = 1$, which in turn entails the specular expansion in operator form

$$\boldsymbol{\Phi}(L) = \underbrace{\sum_{i \ge 0} \boldsymbol{M}_i L^i}_{regular\ \ part} + \underbrace{\sum_{j \ge 0} \boldsymbol{N}_j \nabla^{-j}}_{principal\ \ part} \tag{2.139}$$

to be a meaningful expression, upon the isomorphic argument previously put forward.

Should the said singularity be removable, the principle part of the Laurent expansion would vanish, leading to a Taylor expansion such as

$$\boldsymbol{\Phi}(z) = \underbrace{\sum_{i \ge 0} \boldsymbol{M}_i z_i}_{regular\ \ part} \rightarrow \boldsymbol{\Phi}(L) = \underbrace{\sum_{i \ge 0} \boldsymbol{M}_i L^i}_{regular\ \ part} \tag{2.140}$$

Should it not be the case, the principal part would no longer vanish and reference to the twofold – principal vs. regular part – expansions (2.138) and (2.139) becomes mandatory.

Before putting expansion (2.139) into (2.137) in order to gain a better and deeper understanding of its meaning and implications, we must take some preliminary steps by introducing a suitable notational apparatus. To do so, define a *k-th* order random walk $\boldsymbol{\rho}_t(k)$ as

$$\boldsymbol{\rho}_t(k) = \sum_{\tau \le t} \boldsymbol{\rho}_\tau(k-1) \tag{2.141}$$

where k is a positive integer, with

$$\boldsymbol{\rho}_t(1) = \sum_{\tau \le t} \boldsymbol{\varepsilon}_\tau \tag{2.142}$$

In this way, $\boldsymbol{\rho}_t(2)$ tallies with the notion of a cumulated random walk as per Definition 3 of Sect. 2.4.

Moreover, let us denote by $\boldsymbol{i}_t(k)$ a *k-th* order integrated process, corresponding to a *k-th* order stochastic trend – to be identified with a *k-th* order random walk $\boldsymbol{\rho}_t(k)$, a scale factor apart – and/or a *k-th* order deterministic trend to be identified with the *k-th* power of *t*, a scale factor apart. As to $\boldsymbol{i}_t(0)$, it is meant to be a stationary process, corresponding to a moving average process and/or a constant term.

Moving (2.139) into (2.137) yields a model specification such as

$$\boldsymbol{\xi}_t = \sum_{i\geq 0} \boldsymbol{M}_i L^i \boldsymbol{\varepsilon}_t + \sum_{i\geq 0} \boldsymbol{M}_i \boldsymbol{\eta} + \boldsymbol{N}_1 \nabla^{-1} \boldsymbol{\varepsilon}_t + \boldsymbol{N}_1 \nabla^{-1} \boldsymbol{\eta} + \boldsymbol{N}_2 \nabla^{-2} \boldsymbol{\varepsilon}_t + \boldsymbol{N}_2 \nabla^{2} \boldsymbol{\eta} + \cdots \tag{2.143}$$

which, taking advantage of the new notation, leads to the following expansion into a stationary component, random-walks and powers of *t*, namely stochastic and deterministic trend series,

$$\boldsymbol{\xi}_t = \underset{VMA\ process}{\sum_{i\geq 0} \boldsymbol{M}_i L^i \boldsymbol{\varepsilon}_t} + \underset{\substack{constant\\ term}}{\boldsymbol{a}_0} + \underset{\substack{1st-order\\ random\ walk}}{\boldsymbol{N}_1 \boldsymbol{\rho}_t(1)} + \underset{\substack{linear\ trend\\ in\ t}}{\boldsymbol{a}_1 t} + \underset{\substack{2nd-order\\ random\ walk}}{\boldsymbol{N}_2 \boldsymbol{\rho}_t(2)} + \underset{\substack{quadratic\ trend\\ in\ t}}{\boldsymbol{a}_2 t^2} + \ldots \tag{2.144}$$

Eventually, the foregoing expression can be rewritten as a formal expansion in integrated-processes, namely

$$\boldsymbol{\xi}_t = \underset{stationary\ process}{\boldsymbol{i}_t(0)} + \underset{\substack{1^{st}\ order\ integrated\\ process}}{\boldsymbol{i}_t(1)} + \underset{\substack{2^{nd}\ order\ integrated\\ process}}{\boldsymbol{i}_t(2)} + \cdots \tag{2.145}$$

Each representation of this type brings elucidatory contributions to the understanding of the model content and meaning.

Let us now emphasize the fact that there is a considerable empirical evidence showing that the dynamics inherent in economic variables mirror mostly those of integrated processes of first and second-order. Stationary variables are not frequent at all and third-order integrated ones are even hard to find.

The reasoning just advanced leads to select as reference models the following truncated forms of a parent specification such as (2.145)

$$\boldsymbol{\xi}_t = \begin{cases} \boldsymbol{i}_t(0) \\ \boldsymbol{i}_t(0) + \boldsymbol{i}_t(1) \\ \boldsymbol{i}_t(0) + \boldsymbol{i}_t(1) + \boldsymbol{i}_t(2) \end{cases} \tag{2.146}$$

depending on the nature of the dynamics brought up on empirical basis.

This eventually leads to focus on the specifications

(a)
$$\boldsymbol{\xi}_t = \underbrace{\sum_{i\geq 0} \boldsymbol{M}_i \boldsymbol{\varepsilon}_{t-i}}_{VMA\ \ process} + \underbrace{\boldsymbol{a}_0}_{constant} \tag{2.147}$$

if there is evidence of a stationary generating model for the economic variable under study,

(b)
$$\boldsymbol{\xi}_t = \underbrace{\sum_{i\geq 0} \boldsymbol{M}_i \boldsymbol{\varepsilon}_{t-i}}_{VMA\ \ process} + \underbrace{\boldsymbol{a}_0}_{constant} + \underbrace{\boldsymbol{N}_1 \boldsymbol{\rho}_t(1)}_{\substack{1st-order \\ random\ \ walk}} + \underbrace{\boldsymbol{a}_1 t}_{\substack{linear\ \ trend \\ in\ \ t}} \tag{2.148}$$

if there is evidence of a first-order integrated generating model for the economic variables under study

(c)
$$\boldsymbol{\xi}_t = \underbrace{\sum_{i\geq 0} \boldsymbol{M}_i \boldsymbol{\varepsilon}_{t-i}}_{VMA\ \ process} + \underbrace{\boldsymbol{a}_0}_{constant} + \underbrace{\boldsymbol{N}_1 \boldsymbol{\rho}_t(1)}_{\substack{1st-order \\ random\ \ walk}} + \underbrace{\boldsymbol{a}_1 t}_{\substack{linear\ \ trend \\ in\ \ t}} + \underbrace{\boldsymbol{N}_2 \boldsymbol{\rho}_t(2)}_{\substack{2nd-order \\ random\ \ walk}} + \underbrace{\boldsymbol{a}_2 t^2}_{\substack{quadratic\ \ trend \\ in\ \ t}} \tag{2.149}$$

if there is evidence of a second-order integrated generating model for the economic variables under study.

Under suitable restrictions, which go from the invertibility of the VMA process for what concerns model (2.147) (see Definition 6, Sect. 2.2) to more sophisticated regularity conditions for what concerns models (2.148) and (2.149) (see dual representation theorems in Sects. 3.4 and 3.5), these specifications can be considered as the offsprings of underlying VAR models in the wake of the arguments put forward in Sects. 2 and 4.

Appendix A: A Convenient Reparametrization of a VAR Model and Related Results

An m-dimensional VAR(K) model

$$\tilde{\boldsymbol{A}}(L) \underset{(m,1)}{\tilde{\boldsymbol{y}}_t} = \tilde{\boldsymbol{\mu}} + \tilde{\boldsymbol{\varepsilon}}_t \,, \; \tilde{\boldsymbol{\varepsilon}}_t \sim WN_{(m)} \tag{2.150}$$

where

$$\tilde{\boldsymbol{A}}(L) = \sum_{k=0}^{K} \tilde{\boldsymbol{A}}_k L^k \,, \quad \tilde{\boldsymbol{A}}_0 = \boldsymbol{I}, \quad \tilde{\boldsymbol{A}}_K \neq \boldsymbol{0} \tag{2.151}$$

can be conveniently rewritten as an mK-dimensional VAR(1) model, such as

$$\boldsymbol{A}(L) \underset{(n,1)}{\boldsymbol{y}_t} = \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t \,, \; \boldsymbol{A}(L) = \boldsymbol{I}_n + \tilde{\boldsymbol{A}} L \boldsymbol{,} \; n = mK \tag{2.152}$$

by resorting to the auxiliary vectors and (companion) matrix

$$\underset{(n,1)}{\boldsymbol{y}_t} = \begin{bmatrix} \tilde{\boldsymbol{y}}_t \\ \tilde{\boldsymbol{y}}_{t-1} \\ \vdots \\ \tilde{\boldsymbol{y}}_{t-K+1} \end{bmatrix} \tag{2.153}$$

$$\boldsymbol{\mu} = \begin{bmatrix} \tilde{\boldsymbol{\mu}} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{2.154}$$

$$\boldsymbol{\varepsilon}_t = \begin{bmatrix} \tilde{\boldsymbol{\varepsilon}}_t \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{2.155}$$

$$\tilde{A} = -A_1 = \begin{bmatrix} \tilde{A}_1 & \vdots & \tilde{A}_2 & \tilde{A}_3 & .. & \tilde{A}_K \\ \cdots & & \cdots & \cdots & \cdots & \cdots \\ -I & \vdots & 0 & 0 & 0 & 0 \\ 0 & \vdots & -I & 0 & 0 & 0 \\ .. & \vdots & .. & .. & .. & .. \\ 0 & \vdots & 0 & 0 & -I & 0 \end{bmatrix} \tag{2.156}$$

Introducing the selection matrix

$$\underset{(n,m)}{J} = \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{2.157}$$

whose K blocks are square matrices of order m, the solution of the parent model (2.150) can be recovered from the solution of its companion form (2.152), by pre-multiplying the latter by J', namely

$$\tilde{y}_t = J' y_t \tag{2.158}$$

The following theorem establishes the relationships linking the eigenvalues of the companion matrix A_1 to the roots of the characteristic polynomial $det\tilde{A}(z)$.

Theorem

The non-null eigenvalues λ_i of the companion matrix A_1 are the reciprocals of the roots z_j of the characteristic polynomial $det\tilde{A}(z)$, that is

$$\lambda_i = z_j^{-1} \tag{2.159}$$

Proof

The eigenvalues of the matrix A_1 are the solutions with respect to λ of the determinant equation

$$det(A_1 - \lambda I) = 0 \tag{2.160}$$

Partitioning $A_1 - \lambda I$ in this way

$$\begin{bmatrix} A_{11}(\lambda) & A_{12}(\lambda) \\ A_{21}(\lambda) & A_{22}(\lambda) \end{bmatrix} =$$

$$= \left[\begin{array}{c:ccccc} (-\tilde{A}_1 - \lambda I) & -\tilde{A}_2 & -\tilde{A}_3 & ...+\ \tilde{A}_{K-1} & -\tilde{A}_K \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \boldsymbol{I}_m & -\lambda \boldsymbol{I}_m & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_m & -\lambda \boldsymbol{I}_m & \boldsymbol{0} & \boldsymbol{0} \\ .. & .. & .. & ... & .. \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}_m & -\lambda \boldsymbol{I}_m \end{array}\right] \tag{2.161}$$

and making use of the formula of Schur (see, e.g, Gantmacher vol I p 46) $det(\boldsymbol{A}_1 - \lambda \boldsymbol{I})$ can be rewritten (provided $\lambda \neq 0$) as

$$det(\boldsymbol{A}_1 - \lambda \boldsymbol{I}) = det\boldsymbol{A}_{22}(\lambda) det(\boldsymbol{A}_{11}(\lambda) - \boldsymbol{A}_{12}(\lambda)\boldsymbol{A}_{22}^{-1}(\lambda)\boldsymbol{A}_{21}(\lambda))$$

$$\propto det\boldsymbol{A}_{22}(\lambda)\ det(\tilde{\boldsymbol{A}}_1 + \lambda \boldsymbol{I} + \frac{\tilde{\boldsymbol{A}}_2}{\lambda} + \frac{\tilde{\boldsymbol{A}}_3}{\lambda^2} + ... + \frac{\tilde{\boldsymbol{A}}_K}{\lambda^{K-1}}) \tag{2.162}$$

upon noting that

$$\boldsymbol{A}_{22}^{-1}(z) = \begin{bmatrix} -\frac{1}{\lambda}\boldsymbol{I} & \boldsymbol{0} & ..... & \boldsymbol{0} & \boldsymbol{0} \\ -\frac{1}{\lambda^2}\boldsymbol{I} & -\frac{1}{\lambda}\boldsymbol{I} & ..... & \boldsymbol{0} & \boldsymbol{0} \\ .... & .... & ..... & ..... & .... \\ -\frac{1}{\lambda^{K-1}}\boldsymbol{I} & -\frac{1}{\lambda^{K-2}}\boldsymbol{I} & ..... & -\frac{1}{\lambda^2}\boldsymbol{I} & -\frac{1}{\lambda}\boldsymbol{I} \end{bmatrix} \tag{2.163}$$

This, together with the fact that

$$det\boldsymbol{A}_{22} = (det(-\lambda \boldsymbol{I}_m))^{K-1} \propto \lambda^{m(K-1)} \tag{2.164}$$

eventually leads to express $det(\boldsymbol{A}_1 - \lambda I)$ in this manner

$$det(\boldsymbol{A}_1 - \lambda I) \propto \lambda^{m(K-1)}\ det(\tilde{\boldsymbol{A}}_1 + \lambda \boldsymbol{I} + \frac{\tilde{\boldsymbol{A}}_2}{\lambda} + \frac{\tilde{\boldsymbol{A}}_3}{\lambda^2} + ... + \frac{\tilde{\boldsymbol{A}}_K}{\lambda^{K-1}}) \tag{2.165}$$

This, with simple computations, can be rewritten as

$$det(\boldsymbol{A}_1 - \lambda I) \propto det(\lambda^K \boldsymbol{I} + \lambda^{K-1}\tilde{\boldsymbol{A}}_1 + \lambda^{K-2}\tilde{\boldsymbol{A}}_2 +,......,+\tilde{\boldsymbol{A}}_K) \tag{2.166}$$

Equating to zero and replacing λ ($\lambda \neq 0$) by $\frac{1}{z}$ yields the equation

$$det(\boldsymbol{I} + \tilde{\boldsymbol{A}}_1 z + \tilde{\boldsymbol{A}}_2 z^2 +,...,+\tilde{\boldsymbol{A}}_K z^K) = 0 \tag{2.167}$$

which tallies with the characteristic equation associated with the VAR model (2.150).

Hence, the claimed relationship between non-null eigenvalues of $\boldsymbol{A}_1$ and roots of $det\tilde{\boldsymbol{A}}(z)$ is proved.

□

As a by-product of the previous theorem, whenever the roots of $det\tilde{\boldsymbol{A}}(z)$ are located outside the unit circle (and the solution of the VAR model (2.150) is stationary) the companion matrix $\boldsymbol{A}_1$ is stable and the other way around.

Appendix B: Integrated Processes, Stochastic Trends and Role of Cointegration

Let

$$\underset{(2,\,1)}{\boldsymbol{\xi}_t} = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} \sim I\,(1) \tag{2.168}$$

be a bivariate process integrated of order 1, specified as follows

$$\boldsymbol{\xi}_t = \boldsymbol{A}\boldsymbol{\vartheta}_t + \boldsymbol{\eta}_t \tag{2.169}$$

where $\boldsymbol{\vartheta}_t$ is a vector of stochastic trends

$$\underset{(2,\,1)}{\boldsymbol{\vartheta}_t} = \sum_{\tau=1}^{t} \boldsymbol{\varepsilon}_\tau\,,\;\; \boldsymbol{\varepsilon}_t \sim WN_{(2)} \tag{2.170}$$

and $\boldsymbol{\eta}_t$ is a bivariate process which is covariance stationary with a null mean, which is tantamount to saying that

$$\underset{(2,\,1)}{\boldsymbol{\eta}_t} \sim I\,(0) \tag{2.171}$$

Let us suppose that the matrix $\boldsymbol{A} = [a_{ij}]$ is such that $a_{ij} \neq 0$ for $i, j = 1, 2$ and assume this matrix to be singular, i.e.

$$r\,(\boldsymbol{A}) = 1 \tag{2.172}$$

Then it follows that

(1) The matrix has a null eigenvalue associated with a (left) eigenvector $\boldsymbol{p}'$ such that

$$\boldsymbol{p}'\boldsymbol{A} = \boldsymbol{0}' \tag{2.173}$$

(2) The matrix can be factored into two non–null vectors, in terms of the representation

$$\boldsymbol{A} = \underset{(2,\,1)}{\boldsymbol{b}}\;\underset{(1,\,2)}{\boldsymbol{c}'}\;,\; \boldsymbol{b}'\boldsymbol{b} \neq 0,\; \boldsymbol{c}'\boldsymbol{c} \neq 0 \tag{2.174}$$

Now, according to (2.174), formula (2.169) can be rewritten as

$$\boldsymbol{\xi}_t = \boldsymbol{b}\boldsymbol{c}'\boldsymbol{\vartheta}_t + \boldsymbol{\eta}_t \tag{2.175}$$

where

$$\boldsymbol{c}'\boldsymbol{\vartheta}_t \sim I\,(1) \tag{2.176}$$

Then, by premultiplying both sides of (2.175) by $\boldsymbol{p}'$ we get

$$\boldsymbol{p}'\xi_t = \boldsymbol{p}'\boldsymbol{b}\boldsymbol{c}'\vartheta_t + \boldsymbol{p}'\eta_t = \boldsymbol{p}'\eta_t \sim I\,(0) \tag{2.177}$$

since from formulas (2.173) and (2.174) it follows that

$$\boldsymbol{p}'\boldsymbol{b} = 0 \Rightarrow \boldsymbol{p}'\boldsymbol{b}\boldsymbol{c}'\vartheta_t = 0 \tag{2.178}$$

Finally, by virtue of (2.168) and (2.177) the conclusion that

$$\xi_t \sim CI\,(1,\,1) \tag{2.179}$$

is easily drawn.

□

Considering the above results we realize that

(1) The process ξ_t is integrated of first order owing to the presence of a stochastic trend via the process $\boldsymbol{c}'\vartheta_t$, which plays the role of a common trend (cf. Stock and Watson 1988) and turns out to influence both components of ξ_t through the (non-null) elements of $\boldsymbol{b}$
(2) The vector $\boldsymbol{p}$ (left eigenvector associated with the null eigenvalue of the matrix $\boldsymbol{A}$) is a cointegration vector for ξ_t since $\boldsymbol{p}'\xi_t$ is stationary
(3) The cointegrability of ξ_t relies crucially on the annihilation of (common) trends, according to (2.178) above

The very meaning of cointegration is thus that of making immaterial or at least weakening the role of the non stationary components.

http://www.springer.com/978-3-540-85995-6

Dynamic Model Analysis
Advanced Matrix Methods and Unit-Root Econometrics
Representation Theorems
Faliva, M.; Zoia, M.G.
2009, X, 218 p., Hardcover
ISBN: 978-3-540-85995-6