Mario Faliva
Maria Grazia Zoia

Dynamic Model Analysis

Advanced Matrix Methods and Unit-Root Econometrics Representation Theorems

Second Edition

Springer

Dynamic Model Analysis

Second Edition

Mario Faliva • Maria Grazia Zoia

Dynamic Model Analysis

Advanced Matrix Methods and Unit-Root Econometrics Representation Theorems

Second Edition

Springer

Professor Mario Faliva
Professor Maria Grazia Zoia
Catholic University of Milan
Faculty of Economics
Department of Mathematics and Econometrics
Largo Gemelli, 1
20123 Milano
Italy
mario.faliva@unicatt.it
maria.zoia@unicatt.it

ISBN 978-3-540-85995-6 e-ISBN 978-3-540-85996-3

Library of Congress Control Number: 2008934857

Cover design: WMXDesign GmbH, Heidelberg, Germany

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Contents

Preface to the Second Edition .... **vii**

Preface to the First Edition .... **ix**

1 The Algebraic Framework of Unit- Root Econometrics **1**

1.1 Generalized Inverses 1
1.2 Orthogonal Complements 6
1.3 Empty Matrices 18
1.4 Partitioned Inversion: Classic and Newly Found Results 19
1.5 Useful Matrix Decompositions 30
1.6 Matrix Polynomial Functions: Zeroes, Roots and Poles 38
1.7 The Laurent Form of a Matrix-Polynomial Inverse about a Unit-Root 51
1.8 Matrix Polynomials and Difference Equation Systems 62
1.9 The Linear Matrix Polynomial 74
1.10 Index and Rank Properties of Matrix Coefficients vs. Pole Order in Matrix Polynomial Inversion 84
1.11 Closed-Forms of Laurent Expansion Coefficient Matrices. First Approach 97
1.12 Closed-Forms of Laurent Expansion Coefficient Matrices. Second Approach 111

2 The Statistical Setting **127**

2.1 Stochastic Processes: Preliminaries 127
2.2 Principal Multivariate Stationary Processes 130
2.3 The Source of Integration and the Seeds of Cointegration 142
2.4 Integrated and Cointegrated Processes 145
2.5 Casting a Glance at the Backstage of VAR Modelling 151
Appendix A Convenient Reparametrization of a VAR Model and Related Results 155
Appendix B Integrated Processes, Stochastic Trends and Rôle of Cointegration 159

3 Econometric Dynamic Models: from Classical Econometrics to Time Series Econometrics 161

3.1 Macroeconometric Structural Models versus VAR Models 161
3.2 Basic VAR Specifications and Engendered Processes 167
3.3 A Sequential Rank Criterion for the Integration Order of a VAR Solution.. 172
3.4 Representation Theorems for Processes I (1) 179
3.5 Representation Theorems for Processes I (2) 189
3.6 A Unified Representation Theorem 203

References 213

Notational Conventions, Symbols and Acronyms 215

List of Definitions..217

Preface to the Second Edition

This second edition sees the light three years after the first one: too short a time to feel seriously concerned to redesign the entire book, but sufficient to be challenged by the prospect of sharpening our investigation on the working of econometric dynamic models and to be inclined to change the title of the new edition by dropping the "Topics in" of the former edition.

After considerable soul searching we agreed to include several results related to topics already covered, as well as additional sections devoted to new and sophisticated techniques, which hinge mostly on the latest research work on linear matrix polynomials by the second author. This explains the growth of chapter one and the deeper insight into representation theorems in the last chapter of the book.

The rôle of the second chapter is that of providing a bridge between the mathematical techniques in the backstage and the econometric profiles in the forefront of dynamic modelling. For this purpose, we decided to add a new section where the reader can find the stochastic rationale of vector autoregressive specifications in econometrics.

The third (and last) chapter improves on that of the first edition by reaping the fruits of the thorough analytic equipment previously drawn up. As a result, the reappraisal of the representation theorem for second-order integrated processes sheds full light on the cointegration structure of the VAR model solution. Finally, a unified representation theorem of new conception is established: it provides a general frame of analysis for VAR models in the presence of unit roots and duly shapes the contours of the integration-cointegration features of the engendered processes, with first and second-order processes arising as special cases.

Milan, November 2008 Mario Faliva and Maria Grazia Zoia

Preface to the First Edition

Classical econometrics – which plunges its roots in economic theory with simultaneous equations models (SEM) as offshoots – and time series econometrics – which stems from economic data with vector autoregressive (VAR) models as offsprings – scour, like Janus's facing heads, the flowing of economic variables so as to bring to the fore their autonomous and non-autonomous dynamics. It is up to the so-called final form of a dynamic SEM, on the one hand, and to the so-called representation theorems of (unit-root) VAR models on the other, to provide informative closed form expressions for the trajectories, or time paths, of the economic variables of interest.

Should we look at the issues just put forward from a mathematical standpoint, the emblematic models of both classical and time series econometrics would turn out to be difference equation systems with ad hoc characteristics, whose solutions are attained via a final form or a representation theorem approach. The final solution – algebraic technicalities apart – arises in the wake of classical difference equation theory, displaying besides a transitory autonomous component, an exogenous one along with a stochastic nuisance term. This follows from a properly defined matrix function inversion admitting a Taylor expansion in the lag operator because of the assumptions regarding the roots of a determinant equation peculiar to SEM specifications.

Such was the state of the art when, after Granger's seminal work, time series econometrics came into the limelight and (co)integration burst onto the stage. While opening up new horizons to the modelling of economic dynamics, this nevertheless demanded a somewhat sophisticated analytical apparatus to bridge the unit-root gap between SEM and VAR models.

Over the past two decades econometric literature has by and large given preferential treatment to the role and content of time series econometrics as such and as compared with classical econometrics. Meanwhile, a fascinating – although at times cumbersome – algebraic tool kit has taken shape in a sort of osmotic relationship with (co)integration theory advancements.

The picture just outlined, where lights and shadows – although not explicitly mentioned – still share the scene, spurs us on to seek a deeper insight into several facets of dynamic model analysis, whence the idea of this monograph devoted to representation theorems and their analytical foundations.

The book is organised as follows.

Chapter 1 is designed to provide the reader with a self-contained treatment of matrix theory aimed at paving the way to a rigorous derivation of representation theorems later on. It brings together several results on generalized inverses, orthogonal complements, and partitioned inversion rules (some of them new) and investigates the issue of matrix polynomial inversion about a pole (in its relationships with difference equation theory) via Laurent expansions in matrix form, with the notion of Schur complement and a newly found partitioned inversion formula playing a crucial role in the determination of coefficients.

Chapter 2 deals with statistical setting problems tailored to the special needs of this monograph. In particular, it covers the basic concepts of stochastic processes – both stationary and integrated – with a glimpse at cointegration in view of a deeper insight to be provided in the next chapter.

Chapter 3, after outlining a common frame of reference for classical and time series econometrics bridging the unit-root gap between structural and vector autoregressive models, tackles the issue of VAR specification and resulting processes, with the integration orders of the latter drawn from the rank characteristics of the former. Having outlined the general setting, the central topic of representation theorems is dealt with, in the wake of time series econometrics tradition named after Granger and Johansen (to quote only the forerunner and the leading figure *par excellence*), and further developed along innovating directions, thanks to the effective analytical tool kit set forth in Chapter 1.

The book is obviously not free from external influences and acknowledgement must be given to the authors, quoted in the reference list, whose works have inspired and stimulated the writing of this book.

We express our gratitude to Siegfried Schaible for his encouragement regarding the publication of this monograph.

Our greatest debt is to Giorgio Pederzoli, who read the whole manuscript and made detailed comments and insightful suggestions.

We are also indebted to Wendy Farrar for her peerless checking of the text.

Finally, we thank Daniele Clarizia for his painstaking typing of the manuscript.

Milan, March 2005

Mario Faliva and Maria Grazia Zoia
Istituto di Econometria e Matematica
Università Cattolica, Milano

Chapter 1

The Algebraic Framework of Unit-Root Econometrics

Time series econometrics is centred around the representation theorems from which one can establish the integration and cointegration characteristics of the solutions for the vector autoregressive (VAR) models.

Such theorems, along the path established by Engle and Granger and by Johansen and his school, have brought about a parallel development of an ad hoc analytical implementation, although not yet quite satisfactory. The present chapter, by reworking and expanding some recent contributions due to Faliva and Zoia, systematically provides an algebraic setting based upon several interesting results on inversion by parts and on Laurent series expansion for the reciprocal of a matrix polynomial in a deleted neighbourhood of a unitary root. Rigorous and efficient, such a technique allows for a quick and new reformulation of the representation theorems as it will become clear in Chap. 3.

1.1 Generalized Inverses

We begin by giving some definitions and theorems on generalized inverses. For these and related results see Gantmacher (1959), Rao and Mitra (1971), Pringle and Rayner (1971), Campbell and Meyer (1979), Searle (1982).

Definition 1 – Generalized Inverse

A generalized inverse of a matrix $\boldsymbol{A}$ of order $m \times n$ is a matrix $\boldsymbol{A}^{-}$ of order $n \times m$ such that

$$\boldsymbol{A}\,\boldsymbol{A}^{-}\boldsymbol{A} = \boldsymbol{A} \tag{1.1}$$

The matrix $\boldsymbol{A}^{-}$ is not unique unless $\boldsymbol{A}$ is a square non-singular matrix.

We will adopt the following conventions

$$B = A^{-} \tag{1.2}$$

to indicate that B is a generalized inverse of A, and

$$A^{-} = B \tag{1.3}$$

to indicate that one possible choice for the generalized inverse of A is given by the matrix B.

Definition 2 – Reflexive Generalized Inverse

A reflexive generalized inverse of a matrix A of order $m \times n$ is a matrix A_{ρ}^{-} of order $n \times m$ such that

$$A A_{\rho}^{-} A = A \tag{1.4}$$

$$A_{\rho}^{-} A A_{\rho}^{-} = A_{\rho}^{-} \tag{1.4'}$$

The matrix A_{ρ}^{-} satisfies the rank condition

$$r(A_{\rho}^{-}) = r(A) \tag{1.5}$$

Definition 3 – Moore–Penrose Inverse

The Moore–Penrose generalized inverse of a matrix A of order $m \times n$ is a matrix A^{g} of order $n \times m$ such that

$$A A^{g} A = A \tag{1.6}$$

$$A^{g} A A^{g} = A^{g} \tag{1.7}$$

$$(A A^{g})' = A A^{g} \tag{1.8}$$

$$(A^{g} A)' = A^{g} A \tag{1.9}$$

where A' stands for the transpose of A. The matrix A^{g} exists and it is unique.

Definition 4 – Nilpotent Matrix

A square matrix $\boldsymbol{A}$ is called nilpotent, if a certain power of the matrix is the null matrix. The least exponent for which the power of the matrix vanishes is called the index of nilpotency.

Definition 5 – Index of a Matrix

Let $\boldsymbol{A}$ be a square matrix. The index of $\boldsymbol{A}$, written $ind(\boldsymbol{A})$, is the least non-negative integer υ for which

$$r(\boldsymbol{A}^{\upsilon}) = r(\boldsymbol{A}^{\upsilon+1}) \tag{1.10}$$

Note that (see, e.g., Campbell and Mayer, p 121)

(1) $ind(\boldsymbol{A}) = 0$ if $\boldsymbol{A}$ *is non-singular*,
(2) $ind(\boldsymbol{0}) = 1$,
(3) *if* $\boldsymbol{A}$ *is idempotent, then* $ind(\boldsymbol{A}) = 1$,
(4) *if* $ind(\boldsymbol{A}) = \upsilon$, *then* $ind(\boldsymbol{A}^{\upsilon}) = 1$,
(5) *if* $\boldsymbol{A}$ *is nilpotent with index of nilpotency* υ, *then* $ind(\boldsymbol{A}) = \upsilon$.

Definition 6 – Drazin Inverse

The Drazin inverse of a square matrix $\boldsymbol{A}$ is a square matrix $\boldsymbol{A}^D$ such that

$$\boldsymbol{A}\boldsymbol{A}^D = \boldsymbol{A}^D\boldsymbol{A} \tag{1.11}$$

$$\boldsymbol{A}^D\boldsymbol{A}\boldsymbol{A}^D = \boldsymbol{A}^D \tag{1.12}$$

$$\boldsymbol{A}^{\upsilon+1}\boldsymbol{A}^D = \boldsymbol{A}^{\upsilon} \tag{1.13}$$

where $\upsilon = ind(\boldsymbol{A})$.

The matrix $\boldsymbol{A}^D$ exists and is unique (Rao and Mitra p 96).

The following properties are worth mentioning

$$r(\boldsymbol{A}^D) = r(\boldsymbol{A}^{\upsilon}) \tag{1.14}$$

$$\boldsymbol{A}\boldsymbol{A}^D\boldsymbol{A}^-\boldsymbol{A}^D\boldsymbol{A} = \boldsymbol{A}^D \tag{1.14'}$$

The Drazin inverse is not a generalized inverse as defined in (1.1) above, unless the matrix $\boldsymbol{A}$ is of index 1. Should it be the case, the Drazin inverse will be denoted by $\boldsymbol{A}^{\neq}$ (Campbell and Meyer p 124).

Definition 7 – Right Inverse

A right inverse of a matrix $\boldsymbol{A}$ of order $m \times n$ and full row-rank is a matrix $\boldsymbol{A}_r^-$ of order $n \times m$ such that

$$\boldsymbol{A}\,\boldsymbol{A}_r^- = \boldsymbol{I} \tag{1.15}$$

Theorem 1

The general expression of $\boldsymbol{A}_r^-$ *is*

$$\boldsymbol{A}_r^- = \boldsymbol{H}'(\boldsymbol{A}\boldsymbol{H}')^{-1} \tag{1.16}$$

where $\boldsymbol{H}$ *is an arbitrary matrix of order* $m \times n$ *such that*

$$det(\boldsymbol{A}\boldsymbol{H}') \neq 0 \tag{1.17}$$

Proof

For a proof see Rao and Mitra (Theorem 2.1.1).

□

Remark 1

By taking $\boldsymbol{H} = \boldsymbol{A}$, we obtain

$$\boldsymbol{A}_r^- = \boldsymbol{A}'\,(\boldsymbol{A}\boldsymbol{A}')^{-1} = \boldsymbol{A}^g \tag{1.18}$$

a particularly useful form of right inverse.

Definition 8 – Left Inverse

A left inverse of a matrix $\boldsymbol{A}$ of order $m \times n$ and full column-rank is a matrix $\boldsymbol{A}_l^-$ of order $n \times m$ such that

$$\boldsymbol{A}_l^-\boldsymbol{A} = \boldsymbol{I} \tag{1.19}$$

Theorem 2

The general expression of $\boldsymbol{A}_l^-$ *is*

$$A_l^- = (\boldsymbol{K}'\boldsymbol{A})^{-1}\boldsymbol{K}' \tag{1.20}$$

where $\boldsymbol{K}$ *is an arbitrary matrix of order m × n such that*

$$det(\boldsymbol{K}'\boldsymbol{A}) \neq 0 \tag{1.21}$$

Proof

For a proof see Rao and Mitra (Theorem 2.1.1).

□

Remark 2

By letting $\boldsymbol{K} = \boldsymbol{A}$, we obtain

$$A_l^- = (\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}' = \boldsymbol{A}^g \tag{1.22}$$

a particularly useful form of left inverse.

Remark 3

A simple computation shows that

$$(A_l^-)' = (\boldsymbol{A}')_r^- \tag{1.23}$$

We will now introduce the notion of rank factorization.

Theorem 3

Any matrix $\boldsymbol{A}$ *of order m × n and rank* r *may be factored as follows*

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}' \tag{1.24}$$

where $\boldsymbol{B}$ *is of order* m × r, $\boldsymbol{C}$ *is of order* n × r, *and both* $\boldsymbol{B}$ *and* $\boldsymbol{C}$ *have rank equal to* r. *Such a representation is known as a rank factorization of* $\boldsymbol{A}$.

Proof

For a proof see Searle (p 194).

□

Remark 4

Trivially $\mathbf{B} = \mathbf{A}$ and $\mathbf{C} = \mathbf{I}$ if $\mathbf{A}$ is of full column-rank. Likewise, $\mathbf{B} = \mathbf{I}$ and $\mathbf{C} = \mathbf{A}$ if $\mathbf{A}$ is of full row-rank.

Theorem 4

Let $\mathbf{B}$ *and* $\mathbf{C}$ *be as in Theorem* 3 *and* $\boldsymbol{\Gamma}$ *be a square non-singular matrix of the appropriate order. Then the following factorizations prove to be true*

$$\mathbf{A}^g = (\mathbf{B}\mathbf{C}')^g = (\mathbf{C}')^g\mathbf{B}^g = \mathbf{C}\,(\mathbf{C}'\mathbf{C})^{-1}\,(\mathbf{B}'\mathbf{B})^{-1}\,\mathbf{B} \tag{1.25}$$

$$(\mathbf{B}\boldsymbol{\Gamma}\mathbf{C}')^g = (\mathbf{C}')^g\,\boldsymbol{\Gamma}^{-1}\,\mathbf{B}^g = \mathbf{C}\,(\mathbf{C}'\mathbf{C})^{-1}\,\boldsymbol{\Gamma}^{-1}\,(\mathbf{B}'\mathbf{B})^{-1}\,\mathbf{B}' \tag{1.25'}$$

$$\mathbf{A}_\rho^- = (\mathbf{B}\mathbf{C}')_\rho^- = (\mathbf{C}')_r^-\,\mathbf{B}_l^- = \mathbf{H}(\mathbf{C}'\mathbf{H})^{-1}\,(\mathbf{K}'\mathbf{B})^{-1}\,\mathbf{K}' \tag{1.26}$$

$$(\mathbf{B}\boldsymbol{\Gamma}\mathbf{C}')_\rho^- = (\mathbf{C}')_r^-\,\boldsymbol{\Gamma}^{-1}\,\mathbf{B}_l^- = \mathbf{H}(\mathbf{C}'\mathbf{H})^{-1}\,\boldsymbol{\Gamma}^{-1}\,(\mathbf{K}'\mathbf{B})^{-1}\,\mathbf{K}' \tag{1.26'}$$

$$\mathbf{A}^{\neq} = (\mathbf{B}\mathbf{C}')^{\neq} = \mathbf{B}(\mathbf{C}'\mathbf{B})^{-2}\,\mathbf{C}', \text{ under } ind(\mathbf{A}) = 1 \tag{1.27}$$

$$\mathbf{A}^D = (\mathbf{B}\mathbf{C}')^D = \mathbf{B}\mathbf{F}(\mathbf{G}'\mathbf{F})^{-3}\,\mathbf{G}'\mathbf{C}', \text{ under } ind(\mathbf{A}) = 2 \tag{1.27'}$$

where $\mathbf{H}$ *and* $\mathbf{K}$ *are as in Theorems* 1 *and* 2, *and* $\mathbf{F}$ *and* $\mathbf{G}$ *come from the rank factorization* $\mathbf{C}'\mathbf{B} = \mathbf{F}\mathbf{G}'$.

Proof

For proofs see Greville (1960), Pringle and Rayner (1971, p 31), Rao and Mitra (1971, p 28), and Campbell and Meyer (1979, p 149), respectively.

□

1.2 Orthogonal Complements

We shall now introduce some definitions and establish several results concerning orthogonal complements.

Definition 1 – Row Kernel

The row kernel, or null row space, of a matrix $\boldsymbol{A}$ of order $m \times n$ and rank r is the space of dimension $m - r$ of all solutions $\boldsymbol{x}$ of $\boldsymbol{x}' \boldsymbol{A} = \boldsymbol{0}'$.

Definition 2 – Orthogonal Complement

An orthogonal complement of a matrix $\boldsymbol{A}$ of order $m \times n$ and full column-rank is a matrix $\boldsymbol{A}_{\perp}$ of order $m \times (m - n)$ and full column-rank such that

$$\boldsymbol{A}'_{\perp} \boldsymbol{A} = \boldsymbol{0} \tag{1.28}$$

Remark 1

The matrix $\boldsymbol{A}_{\perp}$ is not unique. Indeed the columns of $\boldsymbol{A}_{\perp}$ form not only a spanning set, but even a basis for the row kernel of $\boldsymbol{A}$. In the light of the foregoing, a general representation for the orthogonal complement of a matrix $\boldsymbol{A}$ is given by

$$\boldsymbol{A}_{\perp} = \boldsymbol{\Lambda V} \tag{1.29}$$

where $\boldsymbol{\Lambda}$ is a particular orthogonal complement of $\boldsymbol{A}$ and $\boldsymbol{V}$ is an arbitrary square non-singular matrix connecting the reference basis (namely, the $m - n$ columns of $\boldsymbol{\Lambda}$) to another (namely, the $m - n$ columns of $\boldsymbol{\Lambda V}$). The matrix $\boldsymbol{V}$ is usually referred to as a transition matrix between bases (see Lancaster and Tismenetsky 1985, p 98).

We shall adopt the following conventions

$$\boldsymbol{\Lambda} = \boldsymbol{A}_{\perp} \tag{1.30}$$

to indicate that $\boldsymbol{\Lambda}$ is an orthogonal complement of $\boldsymbol{A}$, and

$$\boldsymbol{A}_{\perp} = \boldsymbol{\Lambda} \tag{1.31}$$

to indicate that one possible choice for the orthogonal complement of $\boldsymbol{A}$ is given by the matrix $\boldsymbol{\Lambda}$.

The equality

$$(\boldsymbol{A}_{\perp})_{\perp} = \boldsymbol{A} \tag{1.32}$$

reads accordingly.

We now prove the following invariance theorem

Theorem 1

The expressions

$$\boldsymbol{A}_{\perp}(\boldsymbol{H}'\boldsymbol{A}_{\perp})^{-1} \tag{1.33}$$

$$\boldsymbol{C}_{\perp}(\boldsymbol{B}'_{\perp}\boldsymbol{K}\boldsymbol{C}_{\perp})^{-1}\boldsymbol{B}'_{\perp} \tag{1.34}$$

and the rank of the partitioned matrix

$$\begin{bmatrix} \boldsymbol{J} & \boldsymbol{B}_{\perp} \\ \boldsymbol{C}_{\perp} & \boldsymbol{0} \end{bmatrix} \tag{1.35}$$

are invariant for any choice of $\boldsymbol{A}_{\perp}$, $\boldsymbol{B}_{\perp}$ *and* $\boldsymbol{C}_{\perp}$, *where* $\boldsymbol{A}$, $\boldsymbol{B}$ *and* $\boldsymbol{C}$ *are full column-rank matrices of order* m × n, $\boldsymbol{H}$ *is an arbitrary full column-rank matrix of order* m × (m – n) *such that*

$$det(\boldsymbol{H}'\boldsymbol{A}_{\perp}) \neq 0 \tag{1.36}$$

and both $\boldsymbol{J}$ *and* $\boldsymbol{K}$ *are arbitrary matrices of order* m *although one must have*

$$det(\boldsymbol{B}'_{\perp}\boldsymbol{K}\boldsymbol{C}_{\perp}) \neq 0 \tag{1.37}$$

Proof

To prove the invariance of the matrix (1.33) we must verify that

$$\boldsymbol{A}_{\perp 1}(\boldsymbol{H}'\boldsymbol{A}_{\perp 1})^{-1} - \boldsymbol{A}_{\perp 2}(\boldsymbol{H}'\boldsymbol{A}_{\perp 2})^{-1} = \boldsymbol{0} \tag{1.38}$$

where $\boldsymbol{A}_{\perp 1}$ and $\boldsymbol{A}_{\perp 2}$ are two choices of the orthogonal complement of $\boldsymbol{A}$. After the arguments advanced to arrive at (1.29), the matrices $\boldsymbol{A}_{\perp 1}$ and $\boldsymbol{A}_{\perp 2}$ are linked by the relation

$$\boldsymbol{A}_{\perp 2} = \boldsymbol{A}_{\perp 1}\boldsymbol{V} \tag{1.39}$$

for a suitable choice of the transition matrix $\boldsymbol{V}$.

Therefore, substituting $\boldsymbol{A}_{\perp 1}\ \boldsymbol{V}$ for $\boldsymbol{A}_{\perp 2}$ in the left-hand side of (1.38) yields

$$\begin{aligned} &\boldsymbol{A}_{\perp 1}(\boldsymbol{H}'\boldsymbol{A}_{\perp 1})^{-1} - \boldsymbol{A}_{\perp 1}\boldsymbol{V}(\boldsymbol{H}'\boldsymbol{A}_{\perp 1}\boldsymbol{V})^{-1} \\ &= \boldsymbol{A}_{\perp 1}(\boldsymbol{H}'\boldsymbol{A}_{\perp 1})^{-1} - \boldsymbol{A}_{\perp 1}\boldsymbol{V}\boldsymbol{V}^{-1}(\boldsymbol{H}'\boldsymbol{A}_{\perp 1})^{-1} = \boldsymbol{0} \end{aligned} \tag{1.40}$$

which proves the asserted invariance.

The proof of the invariance of the matrix (1.34) follows along the same lines as above, by repeating for $\boldsymbol{B}_\perp$ and $\boldsymbol{C}_\perp$ the reasoning used for $\boldsymbol{A}_\perp$.

The proof of the invariance of the rank of the matrix (1.35) follows upon noting that

$$\begin{aligned} r\left(\begin{bmatrix} \boldsymbol{J} & \boldsymbol{B}_{\perp 2} \\ \boldsymbol{C}'_{\perp 2} & \boldsymbol{0} \end{bmatrix}\right) &= r\left(\begin{bmatrix} \boldsymbol{J} & \boldsymbol{B}_{\perp 1}\boldsymbol{V}_1 \\ \boldsymbol{V}'_2\boldsymbol{C}'_{\perp 1} & \boldsymbol{0} \end{bmatrix}\right) \\ &= r\left(\begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{V}'_2 \end{bmatrix}\begin{bmatrix} \boldsymbol{J} & \boldsymbol{B}_{\perp 1} \\ \boldsymbol{C}'_{\perp 1} & \boldsymbol{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{V}_1 \end{bmatrix}\right) = r\left(\begin{bmatrix} \boldsymbol{J} & \boldsymbol{B}_{\perp 1} \\ \boldsymbol{C}'_{\perp 1} & \boldsymbol{0} \end{bmatrix}\right) \end{aligned} \tag{1.41}$$

where $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ are suitable choices of transition matrices.

□

The following theorem provides explicit expressions for the orthogonal complements of matrix products, which find considerable use in the text.

Theorem 2

Let $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *be full column-rank matrices of order* $1 \times \mathrm{m}$ *and* $m \times n$ *respectively. Then the orthogonal complement of the matrix product* $\boldsymbol{AB}$ *can be expressed as*

$$(\boldsymbol{AB})_\perp = [\boldsymbol{A}_\perp, (\boldsymbol{A}')^g \boldsymbol{B}_\perp] \tag{1.42}$$

In particular, if $1 = \mathrm{m}$ *then*

$$(\boldsymbol{AB})_\perp = (\boldsymbol{A}')^{-1} \boldsymbol{B}_\perp \tag{1.43}$$

while if $\mathrm{m} = \mathrm{n}$ *then*

$$(\boldsymbol{AB})_\perp = \boldsymbol{A}_\perp \tag{1.44}$$

Proof

It is enough to check that

$$(\boldsymbol{AB})' [\boldsymbol{A}_\perp, (\boldsymbol{A}')^g \boldsymbol{B}_\perp] = \boldsymbol{0} \tag{1.45}$$

and that the block matrix

$$[\boldsymbol{A}_\perp, (\boldsymbol{A}')^g \boldsymbol{B}_\perp, \boldsymbol{AB}] \tag{1.46}$$

is square and of full rank. Hence the right-hand side of (1.42) provides an explicit expression for the orthogonal complement of $\boldsymbol{AB}$ according to Definition 2 (see also Faliva and Zoia 2003). This proves (1.42).

The result (1.43) holds true by straightforward computation.

The result (1.44) is trivial in light of representation (1.29) for orthogonal complements.

□

Remark 2

The right-hand side of (1.42) provides a convenient expression of $(\boldsymbol{AB})_\perp$. Actually, a more general form of $(\boldsymbol{AB})_\perp$ is given by

$$(\boldsymbol{AB})_\perp = [\boldsymbol{A}_\perp \boldsymbol{\Psi}, (\boldsymbol{A}')_r^- \boldsymbol{B}_\perp \boldsymbol{\Omega}] \tag{1.47}$$

where $\boldsymbol{\Psi}$ and $\boldsymbol{\Omega}$ are arbitrary non-singular matrices.

Remark 3

The dual statement

$$[\boldsymbol{A}_\perp \boldsymbol{\Psi}, (\boldsymbol{A}')_r^- \boldsymbol{B}_\perp \boldsymbol{\Omega}]_\perp = \boldsymbol{AB} \tag{1.48}$$

holds true in the sense of equality (1.32).

The next two theorems provide representations for generalized and regular inverses of block matrices involving orthogonal complements.

Theorem 3

Suppose that $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *are as in Theorem* 2. *Then*

$$[\boldsymbol{A}_\perp, (\boldsymbol{A}')^g \boldsymbol{B}_\perp]^- = \begin{bmatrix} \boldsymbol{A}_\perp^g \\ \boldsymbol{B}_\perp^g \boldsymbol{A}' \end{bmatrix} \tag{1.49}$$

$$[\boldsymbol{A}_\perp, (\boldsymbol{A}')^g \boldsymbol{B}_\perp]^g = \begin{bmatrix} \boldsymbol{A}_\perp^g \\ \left((\boldsymbol{A}')^g \boldsymbol{B}_\perp\right)^g \end{bmatrix} \tag{1.50}$$

Proof

The results follow from Definitions 1 and 3 of Sect. 1.1 by applying Theorems 3.1 and 3.4, Corollary 4, in Pringle and Rayner (1971, p 38).

□

Theorem 4

The inverse of the composite matrix $[\boldsymbol{A},\ \boldsymbol{A}_{\perp}]$ *can be written as follows*

$$[\boldsymbol{A},\ \boldsymbol{A}_{\perp}]^{-1} = \begin{bmatrix} \boldsymbol{A}^{g} \\ \boldsymbol{A}_{\perp}^{g} \end{bmatrix} \tag{1.51}$$

which, in turns, leads to the noteworthy identity

$$\boldsymbol{A}\,\boldsymbol{A}^{g} + \boldsymbol{A}_{\perp}\,\boldsymbol{A}_{\perp}^{g} = \boldsymbol{I} \tag{1.52}$$

Proof

The proof is a by-product of Theorem 3.4, Corollary 4, in Pringle and Rayner, and the identity (1.52) ensues from the commutative property of the inverse.

□

The following theorem provides a useful generalization of the identity (1.52).

Theorem 5

Let $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *be full column-rank matrices of order* $m \times n$ *and* $m \times (m-n)$ *respectively, such that the composite matrix* $[\boldsymbol{A}, \boldsymbol{B}]$ *is non-singular. Then, the following identity*

$$\boldsymbol{A}\,(\boldsymbol{B}_{\perp}'\boldsymbol{A})^{-1}\ \boldsymbol{B}_{\perp}' + \boldsymbol{B}\,(\boldsymbol{A}_{\perp}'\boldsymbol{B})^{-1}\ \boldsymbol{A}_{\perp}' = \boldsymbol{I} \tag{1.53}$$

holds true.

Proof

Observe that, insofar as the square matrix $[\boldsymbol{A}, \boldsymbol{B}]$ is non-singular, both $\boldsymbol{B}_{\perp}'\ \boldsymbol{A}$ and $\boldsymbol{A}_{\perp}'\ \boldsymbol{B}$ are non-singular matrices too.
Furthermore, verify that

$$\begin{bmatrix} (\boldsymbol{B}_{\perp}'\boldsymbol{A})^{-1}\boldsymbol{B}_{\perp}' \\ (\boldsymbol{A}_{\perp}'\boldsymbol{B})^{-1}\boldsymbol{A}_{\perp}' \end{bmatrix} [\boldsymbol{A}, \boldsymbol{B}] = \begin{bmatrix} \boldsymbol{I}_n & 0 \\ 0 & \boldsymbol{I}_{m-n} \end{bmatrix}$$

This shows that $\begin{bmatrix} (B'_\perp A)^{-1} B'_\perp \\ (A'_\perp B)^{-1} A'_\perp \end{bmatrix}$ is the inverse of $[A, B]$. Hence the identity (1.53) follows from the commutative property of the inverse.

□

Remark 4

The matrix $(B'_\perp A)^{-1} B'_\perp$ is a left inverse of A whereas the matrix $B(A'_\perp B)^{-1}$ is a right inverse of $A'_\perp$. Together they form a pair of specular directional inverses denoted by A_s^- and $(A'_\perp)_s^-$, according to which identity (1.53) can be rewritten as

$$A A_s^- + (A'_\perp)_s^- A'_\perp = I \tag{1.54}$$

Theorem 6

With A denoting a square matrix of order n *and index $\upsilon \le 2$, let*

$$A^\upsilon = B_\upsilon C'_\upsilon \tag{1.55}$$

be a rank factorization of A^υ. Then, the following holds true

$$r(A^\upsilon) - r(A^{2\upsilon}) = r(C_{\upsilon\perp} B'_{\upsilon\perp}) - r(B'_{\upsilon\perp} C_{\upsilon\perp}) \tag{1.56}$$

Proof

Because of (1.55) observe that

$$r(A^\upsilon) = r(B_\upsilon C'_\upsilon) = r(B_\upsilon) \tag{1.57}$$

$$r(A^{2\upsilon}) = r(B_\upsilon C'_\upsilon B_\upsilon C'_\upsilon) = r(C'_\upsilon B_\upsilon) \tag{1.58}$$

Resorting to Theorem 19 of Marsaglia and Styan (1974) and bearing in mind the identity (1.52) the twin rank equalities

$$r([B_\upsilon, C_{\upsilon\perp}]) = r(B_\upsilon) + r((I - B_\upsilon B_\upsilon^g) C_{\upsilon\perp}) = r(B_\upsilon) + r(B'_{\upsilon\perp} C_{\upsilon\perp}) \tag{1.59}$$

$$\begin{aligned} r([\,\boldsymbol{B}_{\upsilon}\,, \boldsymbol{C}_{\upsilon\perp}\,]) &= r(\boldsymbol{C}_{\upsilon\perp}) + r([\boldsymbol{I} - \boldsymbol{C}'_{\upsilon\perp})^{g}\boldsymbol{C}'_{\upsilon\perp}]\boldsymbol{B}_{\upsilon}) \\ &= r(\boldsymbol{C}_{\upsilon\perp}) + r(\boldsymbol{C}'_{\upsilon}\boldsymbol{B}_{\upsilon}) \end{aligned} \quad (1.59')$$

are easily established. Equating the right-hand sides of (1.59) and (1.59') and bearing in mind (1.57) and (1.58) yields (1.56).

□

Corollary 6.1

The following statements are equivalent

(1) $ind(\boldsymbol{A}) = 1$

(2) $\boldsymbol{C}'\boldsymbol{B}$ *is a non-singular matrix*

(3) $r(\boldsymbol{B}) - r(\boldsymbol{C}'\boldsymbol{B}) = r(\boldsymbol{C}_{\perp}) - r(\boldsymbol{B}'_{\perp}\boldsymbol{C}_{\perp}) = 0$

(4) $\boldsymbol{B}'_{\perp}\boldsymbol{C}_{\perp}$ *is a non-singular matrix*

Proof

$Ind(\boldsymbol{A}) = 1$ is tantamount to saying that

$$r(\boldsymbol{B}\boldsymbol{C}') = r(\boldsymbol{A}) = r(\boldsymbol{A}^2) = r(\boldsymbol{B}\boldsymbol{C}'\ \boldsymbol{B}\boldsymbol{C}') = r(\boldsymbol{C}'\boldsymbol{B}) \quad (1.60)$$

which occurs iff $\boldsymbol{C}'\boldsymbol{B}$ is a non-singular matrix (see also Campbell and Meyer, Theorem 7.8.2)

Proofs of (3) and (4) follow from (1.56).

□

Theorem 7

Let $\boldsymbol{B}$ *and* $\boldsymbol{C}$ *be obtained by the rank factorization* $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}'$, $\boldsymbol{B}_2$ *and* $\boldsymbol{C}_2$ *obtained by the rank factorization* $\boldsymbol{A}^2 = \boldsymbol{B}_2\boldsymbol{C}'_2$. *Then the following identities prove true*

$$\boldsymbol{B}(\boldsymbol{C}'\boldsymbol{B})^{-1}\boldsymbol{C}' + \boldsymbol{C}_{\perp}(\boldsymbol{B}'_{\perp}\boldsymbol{C}_{\perp})^{-1}\boldsymbol{B}'_{\perp} = \boldsymbol{I}, \text{ if } ind(\boldsymbol{A}) = 1 \quad (1.61)$$

$$\boldsymbol{B}_2(\boldsymbol{C}'_2\boldsymbol{B}_2)^{-1}\boldsymbol{C}'_2 + \boldsymbol{C}_{2\perp}(\boldsymbol{B}'_{2\perp}\boldsymbol{C}_{2\perp})^{-1}\boldsymbol{B}'_{\perp} = \boldsymbol{I}, \text{ if } ind(\boldsymbol{A}) = 2 \quad (1.61')$$

Proof

Under *ind*($\boldsymbol{A}$) = 1, in light of (1.56), we have

$$r([\boldsymbol{B}'_{\perp}\boldsymbol{C}_{\perp}]) = r(\boldsymbol{B}_{\perp}) \tag{1.62}$$

so that, bearing in mind (1.59),

$$r([\boldsymbol{B}, \boldsymbol{C}_{\perp}]) = r(\boldsymbol{B}) + r(\boldsymbol{B}'_{\perp}) = n \tag{1.63}$$

and Theorem 5 applies accordingly.

The proof for *ind*($\boldsymbol{A}$) = 2 follows along the same lines.

□

Theorem 8

With $\boldsymbol{A}$, $\boldsymbol{B}$ *and* $\boldsymbol{C}$ *as in Theorem* 6, *let ind*($\boldsymbol{A}$) = 2 *and let* $\boldsymbol{F}$, $\boldsymbol{G}$, $\boldsymbol{R}$, $\boldsymbol{S}$, $\boldsymbol{B}_2$ *and* $\boldsymbol{C}_2$ *be full column-rank matrices defined according to the rank factorizations*

$$\boldsymbol{B}'_{\perp}\boldsymbol{C}_{\perp} = \boldsymbol{R}\boldsymbol{S}' \tag{1.64}$$

$$\boldsymbol{C}'\boldsymbol{B} = \boldsymbol{F}\boldsymbol{G}' \tag{1.65}$$

$$\boldsymbol{A}^2 = \boldsymbol{B}_2\boldsymbol{C}'_2 \tag{1.66}$$

Then, the following statements hold true

(1) $$r(\boldsymbol{R}_{\perp}) = r(\boldsymbol{S}_{\perp}) = r(\boldsymbol{F}_{\perp}) = r(\boldsymbol{G}_{\perp}) \tag{1.67}$$

(2) $$\boldsymbol{F}_{\perp} = \boldsymbol{C}_r^{-}\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp} \qquad \boldsymbol{G}_{\perp} = \boldsymbol{B}_l^{-}\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp} \tag{1.68}$$

$$= \boldsymbol{C}^{g}\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp} \qquad = \boldsymbol{B}^{g}\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp} \tag{1.68'}$$

(3) $$\boldsymbol{B}_{2\perp} = [\boldsymbol{B}_{\perp}, (\boldsymbol{A}')_{\rho}^{-}\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp}] \text{ and } \boldsymbol{C}_{2\perp} = [\boldsymbol{C}_{\perp}, \boldsymbol{A}_{\rho}^{-}\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp}] \tag{1.69}$$

where

$$\boldsymbol{A}_{\rho}^{-} = (\boldsymbol{C}')_r^{-}\boldsymbol{B}_l^{-} . \tag{1.70}$$

Proof

Proof of (1) Under *ind*($\boldsymbol{A}$) = 2, the following hold

$$r(\boldsymbol{A}^2) = r(\boldsymbol{C}'\boldsymbol{B}) = r(\boldsymbol{F}) = r(\boldsymbol{G}) \tag{1.71}$$

$$\begin{aligned} r(\boldsymbol{A}) - r(\boldsymbol{A}^2) &= r(\boldsymbol{F}_\perp) = r(\boldsymbol{G}_\perp) = r(\boldsymbol{C}_\perp) - r(\boldsymbol{B}'_\perp \boldsymbol{C}_\perp) \\ &= r(\boldsymbol{C}_\perp) - r(\boldsymbol{R}) = r(\boldsymbol{R}_\perp) = r(\boldsymbol{S}_\perp) \end{aligned} \tag{1.72}$$

according to (1.56) and (1.65).

Proof of (2) Bearing in mind that $\boldsymbol{F}^- \boldsymbol{F} = \boldsymbol{I}$ and resorting to Theorem 5, together with Remark 4, it is easy to check that the following hold

$$\boldsymbol{B}\boldsymbol{B}_l^- \boldsymbol{C}_\perp \boldsymbol{S}_\perp = \boldsymbol{B}\boldsymbol{B}_s^- \boldsymbol{C}_\perp \boldsymbol{S}_\perp = [\boldsymbol{I} - (\boldsymbol{B}'_\perp)_s^- \boldsymbol{B}'_\perp]\boldsymbol{C}_\perp \boldsymbol{S}_\perp = \boldsymbol{C}_\perp \boldsymbol{S}_\perp \tag{1.73}$$

$$\boldsymbol{G}'\boldsymbol{B}_l^- \boldsymbol{C}_\perp \boldsymbol{S}_\perp = \boldsymbol{F}^- \boldsymbol{F}\boldsymbol{G}'\boldsymbol{B}_l^- \boldsymbol{C}_\perp \boldsymbol{S}_\perp = \boldsymbol{F}^- \boldsymbol{C}'\boldsymbol{B}\boldsymbol{B}_s^- \boldsymbol{C}_\perp \boldsymbol{S}_\perp = \boldsymbol{F}^- \boldsymbol{C}'\boldsymbol{C}_\perp \boldsymbol{S}_\perp = \boldsymbol{0} \tag{1.74}$$

Further, taking into account (1.67), the conclusions that

$$r(\boldsymbol{B}_l^- \boldsymbol{C}_\perp \boldsymbol{S}_\perp) = r(\boldsymbol{B}\boldsymbol{B}_s^- \boldsymbol{C}_\perp \boldsymbol{S}_\perp) = r(\boldsymbol{C}_\perp \boldsymbol{S}_\perp) = r(\boldsymbol{S}_\perp) \tag{1.75}$$

$$r[\boldsymbol{G}, \boldsymbol{B}_l^- \boldsymbol{C}_\perp \boldsymbol{S}_\perp] = r(\boldsymbol{G}) + r(\boldsymbol{S}_\perp) = r(\boldsymbol{G}) + r(\boldsymbol{G}_\perp) \tag{1.76}$$

hold are easily drawn, and since both the orthogonality and the rank conditions of Definition 2 are satisfied, $\boldsymbol{B}_l^- \boldsymbol{C}_\perp \boldsymbol{S}_\perp$ provides one choice of $\boldsymbol{G}_\perp$.

Also, as a by-product of (1.72) it follows that

$$\boldsymbol{B}_l^- \boldsymbol{C}_\perp \boldsymbol{S}_\perp = \boldsymbol{B}^g \boldsymbol{B}\boldsymbol{B}_l^- \boldsymbol{C}_\perp \boldsymbol{S}_\perp = \boldsymbol{B}^g \boldsymbol{C}_\perp \boldsymbol{S}_\perp \tag{1.77}$$

The same conclusion about $\boldsymbol{C}_r^- \boldsymbol{B}_\perp \boldsymbol{R}_\perp$ with respect to $\boldsymbol{F}_\perp$ is then drawn likewise.

Proof of (3) Trivially

$$\boldsymbol{B}_2 = \boldsymbol{B}\boldsymbol{F}, \quad \boldsymbol{C}_2 = \boldsymbol{C}\boldsymbol{G} \tag{1.78}$$

whence (1.69) follows upon resorting to Theorem 2 along with (1.68), bearing in mind (1.26) of Sect. 1.1.

□

Corollary 8.1

The following statements are equivalent

(1) $ind(\boldsymbol{A}) = 2$

(2) $\boldsymbol{G}'\boldsymbol{F}$ *is a non-singular matrix*

(3) $ind(\boldsymbol{C}'\boldsymbol{B}) = 1$

(4) $$r(\boldsymbol{F}) - r(\boldsymbol{G}'\boldsymbol{F}) = r(\boldsymbol{G}_\perp) - r(\boldsymbol{F}_\perp'\boldsymbol{G}_\perp) = 0$$

(5) $$\boldsymbol{F}_\perp'\boldsymbol{G}_\perp = \boldsymbol{R}_\perp'\boldsymbol{B}_\perp'\boldsymbol{A}_\rho^-\boldsymbol{C}_\perp\boldsymbol{S}_\perp \text{ is a non-singular matrix}$$

Proof

$Ind(\boldsymbol{A}) = 2$ is tantamount to saying, on the one hand, that

$$r(\boldsymbol{BC}') = r(\boldsymbol{A}) > r(\boldsymbol{A}^2) = r(\boldsymbol{BC}'\,\boldsymbol{BC}') = r(\boldsymbol{BFG}'\boldsymbol{C}') \tag{1.79}$$

and, on the other, that

$$\begin{aligned} r(\boldsymbol{BFG}'\boldsymbol{C}') = r(\boldsymbol{A}^2) = r(\boldsymbol{B}_2\boldsymbol{C}_2') = r(\boldsymbol{A}^3) = r(\boldsymbol{BFG}'\boldsymbol{FG}'\boldsymbol{C}') \\ = r(\boldsymbol{B}_2\boldsymbol{G}'\boldsymbol{FC}_2') = r(\boldsymbol{G}'\boldsymbol{F}) \end{aligned} \tag{1.80}$$

which occurs iff $\boldsymbol{C}'\boldsymbol{B}$ is singular and $\boldsymbol{G}'\boldsymbol{F}$ is not (see also Campbell and Meyer, Theorem 7.8.2).

With this premise, it easy to check that

$$r(\boldsymbol{C}'\boldsymbol{B})^2 = r(\boldsymbol{FG}'\boldsymbol{FG}') = r(\boldsymbol{FG}') = r(\boldsymbol{C}'\boldsymbol{B}) \tag{1.81}$$

The rank relationship

$$r(\boldsymbol{F}) - r(\boldsymbol{G}'\boldsymbol{F}) = r(\boldsymbol{G}_\perp) - r(\boldsymbol{F}_\perp'\boldsymbol{G}_\perp) = 0 \tag{1.82}$$

(see also (1.58)) holds accordingly and $\boldsymbol{F}_\perp'\boldsymbol{G}_\perp$ is non-singular.

Finally, by resorting to representations (1.68) of $\boldsymbol{F}_\perp$ and $\boldsymbol{G}_\perp$, the product $\boldsymbol{F}_\perp'\boldsymbol{G}_\perp$ can be worked out in this fashion,

$$\boldsymbol{F}_\perp'\boldsymbol{G}_\perp = \boldsymbol{R}_\perp'\boldsymbol{B}_\perp'(\boldsymbol{C}')_r^-\boldsymbol{B}_l^-\boldsymbol{C}_\perp\boldsymbol{S}_\perp = \boldsymbol{R}_\perp'\boldsymbol{B}_\perp'\boldsymbol{A}_\rho^-\boldsymbol{C}_\perp\boldsymbol{S}_\perp \tag{1.83}$$

which completes the proof.

□

The following identities can easily be proved because of Theorem 4 of Sect. 1.1 and Theorem 4 of this section

$$\boldsymbol{AA}^g = \boldsymbol{BB}^g \tag{1.84}$$

$$\boldsymbol{A}^g\boldsymbol{A} = (\boldsymbol{C}')^g\,\boldsymbol{C}' \tag{1.85}$$

$$\boldsymbol{A}^{\neq}\boldsymbol{A} = \boldsymbol{AA}^{\neq} = \boldsymbol{B}(\boldsymbol{C}'\boldsymbol{B})^{-1}\,\boldsymbol{C}',\ \textit{under ind}(\boldsymbol{A}) = 1 \tag{1.86}$$

$$\boldsymbol{A}^D\boldsymbol{A} = \boldsymbol{AA}^D = \boldsymbol{B}_2(\boldsymbol{C}_2'\boldsymbol{B}_2)^{-1}\,\boldsymbol{C}_2',\ \textit{under ind}(\boldsymbol{A}) = 2 \tag{1.86'}$$

$$\boldsymbol{I}_m - \boldsymbol{A}\boldsymbol{A}^g = \boldsymbol{I}_m - \boldsymbol{B}\boldsymbol{B}^g = \boldsymbol{B}_\perp(\boldsymbol{B}_\perp)^g = (\boldsymbol{B}'_\perp)^g\boldsymbol{B}'_\perp \tag{1.87}$$

$$\boldsymbol{I}_n - \boldsymbol{A}^g\boldsymbol{A} = \boldsymbol{I}_n - (\boldsymbol{C}')^g\boldsymbol{C}' = (\boldsymbol{C}'_\perp)^g\boldsymbol{C}'_\perp = \boldsymbol{C}_\perp(\boldsymbol{C}_\perp)^g \tag{1.88}$$

$$\boldsymbol{I} - \boldsymbol{A}^{\neq}\boldsymbol{A} = \boldsymbol{I} - \boldsymbol{A}\boldsymbol{A}^{\neq} = \boldsymbol{C}_\perp(\boldsymbol{B}'_\perp\boldsymbol{C}_\perp)^{-1}\boldsymbol{B}'_\perp, \; under\; ind(\boldsymbol{A}) = 1 \tag{1.89}$$

$$\boldsymbol{I} - \boldsymbol{A}^D\boldsymbol{A} = \boldsymbol{I} - \boldsymbol{A}\boldsymbol{A}^D = \boldsymbol{C}_{2\perp}(\boldsymbol{B}'_{2\perp}\boldsymbol{C}_{2\perp})^{-1}\boldsymbol{B}'_{2\perp}, \; under\; ind(\boldsymbol{A}) = 2 \tag{1.89'}$$

where $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ are as in Theorem 3 of Sect. 1.1 and $\boldsymbol{B}_2$ and $\boldsymbol{C}_2$ are as in Theorem 8.

To conclude this section, we point out that an alternative definition of orthogonal complement – which differs slightly from that of Definition 2 – can be adopted for square singular matrices as indicated in the next definition.

Definition 3 – Left and Right Orthogonal Complements

Let $\boldsymbol{A}$ be a square matrix of order n and rank $r < n$. A left-orthogonal complement of $\boldsymbol{A}$ is a square matrix of order n and rank $n - r$, denoted by $\boldsymbol{A}_l^\perp$, such that

$$\boldsymbol{A}_l^\perp \boldsymbol{A} = \boldsymbol{0} \tag{1.90}$$

$$r([\boldsymbol{A}_l^\perp, \boldsymbol{A}]) = n \tag{1.91}$$

Similarly, a right-orthogonal complement of $\boldsymbol{A}$ is a square matrix of order n and rank $n - r$, denoted by $\boldsymbol{A}_r^\perp$, such that

$$\boldsymbol{A}\boldsymbol{A}_r^\perp = \boldsymbol{0} \tag{1.92}$$

$$r\left(\begin{bmatrix} \boldsymbol{A}_r^\perp \\ \boldsymbol{A} \end{bmatrix}\right) = n \tag{1.93}$$

Suitable choices for the matrices $\boldsymbol{A}_l^\perp$ and $\boldsymbol{A}_r^\perp$ turn out to be (see, e.g., Rao 1973)

$$\boldsymbol{A}_l^\perp = \boldsymbol{I} - \boldsymbol{A}\boldsymbol{A}^- \tag{1.94}$$
$$= \boldsymbol{I} - \boldsymbol{A}\boldsymbol{A}^g \tag{1.94'}$$

$$\boldsymbol{A}_r^\perp = \boldsymbol{I} - \boldsymbol{A}^-\boldsymbol{A} \tag{1.95}$$
$$= \boldsymbol{I} - \boldsymbol{A}^g\boldsymbol{A} \tag{1.95'}$$

1.3 Empty Matrices

By an empty matrix we mean a matrix whose number of rows or of columns is equal to zero.

Some formal rules prove useful when operating with empty matrices.

Let $\boldsymbol{B}$ and $\boldsymbol{C}$ be empty matrices of order $n \times 0$ and $0 \times p$, respectively, and $\boldsymbol{A}$ be a matrix of order $m \times n$.

We then assume the following formal rules of calculus

$$\boldsymbol{AB} = \underset{(m,\,0)}{\boldsymbol{D}}\text{, namely an empty matrix} \tag{1.96}$$

$$\boldsymbol{BC} = \underset{(n,\,p)}{\boldsymbol{0}}\text{, namely a null matrix} \tag{1.97}$$

Reading (1.97) backwards, the formula provides a rank factorization of a null matrix.

Moreover, note from Chipman and Rao (1964) that, since an empty square matrix $\boldsymbol{\Phi}$ of order zero – and rank zero, accordingly – is (trivially) of full rank – namely zero – its inverse can be formally defined as the empty square matrix of order zero as well. In other words

$$\boldsymbol{\Phi}^{-1} = \boldsymbol{\Phi} \tag{1.98}$$

The notion of empty matrix paves the way to some worthwhile extensions for the algebra of orthogonal complements.

Let $\boldsymbol{B}$ be an empty matrix of order $n \times 0$ and $\boldsymbol{A}$ be a non-singular matrix of order n. Then, we will agree upon the following formal relationships

$$\boldsymbol{B}_{\perp} = \boldsymbol{A}\text{, namely an arbitrary non-singular matrix} \tag{1.99}$$

$$\boldsymbol{A}_{\perp} = \boldsymbol{B}\text{, namely an empty matrix.} \tag{1.100}$$

As a by-product from these formal rules, the following equalities hold for $\boldsymbol{B}$ and $\boldsymbol{C}$ empty matrices of order $n \times 0$

$$\boldsymbol{B}(\boldsymbol{C}'\boldsymbol{B})^{-1}\,\boldsymbol{C}' = \boldsymbol{0}_n \tag{1.101}$$

$$\boldsymbol{C}_{\perp}(\boldsymbol{B}_{\perp}'\boldsymbol{C}_{\perp})^{-1}\,\boldsymbol{B}_{\perp}' = \boldsymbol{I}_n \tag{1.102}$$

1.4 Partitioned Inversion: Classic and Newly Found Results

This section, after reviewing some standard results on partitioned inversion, presents newly found inversion formulas which, like Pandora's box, provide the keys for establishing an elegant and rigorous approach to unit-root econometrics main theorems, as shown in Chap. 3.

To begin with we recall the following classic results.

Theorem 1

Let $\boldsymbol{A}$ and $\boldsymbol{D}$ be square matrices of order m and n, respectively, and let $\boldsymbol{B}$ and $\boldsymbol{C}$ be full column-rank matrices of order $m \times n$.

Consider the partitioned matrix, of order $m + n$,

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{D} \end{bmatrix} \tag{1.103}$$

Then any of the following sets of conditions is sufficient for the existence of $\boldsymbol{P}^{-1}$,

(a) *Both $\boldsymbol{A}$ and its Schur complement $\boldsymbol{E} = \boldsymbol{D} - \boldsymbol{C}'\boldsymbol{A}^{-1}\boldsymbol{B}$ are non-singular matrices*

(b) *Both $\boldsymbol{D}$ and its Schur complement $\boldsymbol{F} = \boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C}'$ are non-singular matrices*

Moreover the results listed below hold true, namely

(1) Under (a) *the partitioned inverse of $\boldsymbol{P}$ can be written as*

$$\boldsymbol{P}^{-1} = \begin{bmatrix} \boldsymbol{A}^{-1} + \boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{E}^{-1}\boldsymbol{C}'\boldsymbol{A}^{-1} & -\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{E}^{-1} \\ -\boldsymbol{E}^{-1}\boldsymbol{C}'\boldsymbol{A}^{-1} & \boldsymbol{E}^{-1} \end{bmatrix} \tag{1.104}$$

(2) Under (b) *the partitioned inverse of $\boldsymbol{P}$ can be written as*

$$\boldsymbol{P}^{-1} = \begin{bmatrix} \boldsymbol{F}^{-1} & -\boldsymbol{F}^{-1}\boldsymbol{B}\boldsymbol{D}^{-1} \\ -\boldsymbol{D}^{-1}\boldsymbol{C}'\boldsymbol{F}^{-1} & \boldsymbol{D}^{-1} + \boldsymbol{D}^{-1}\boldsymbol{C}'\boldsymbol{F}^{-1}\boldsymbol{B}\boldsymbol{D}^{-1} \end{bmatrix} \tag{1.105}$$

Proof

The matrix $\boldsymbol{P}^{-1}$ exists insofar as (see Rao 1973, p 32)

$$det(\boldsymbol{P}) = \begin{cases} det(\boldsymbol{A})det(\boldsymbol{E}) \neq 0, under \ \text{(a)} \\ det(\boldsymbol{D})det(\boldsymbol{F}) \neq 0, under \ \text{(b)} \end{cases} \tag{1.106}$$

The partitioned inversion formulas (1.104) and (1.105), under the assumptions (a) and (b), respectively, are standard results of the algebraic tool-kit of econometricians (see, e.g., Goldberger 1964; Theil 1971; Faliva 1987).

□

A weak version of the previous result – which provides a uniquely determined reflexive generalized inverse – is presented in the next theorem (see also Zoia 2006).

Theorem 2

Let $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$ and $\boldsymbol{P}$ be defined as in Theorem 1.

If $\boldsymbol{A}$ is a non-singular matrix, then the following expression holds true for a reflexive generalized inverse in partitioned form of $\boldsymbol{P}$,

$$\boldsymbol{P}_{\rho}^{-} = \begin{bmatrix} \boldsymbol{A}^{-1} + \boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{E}^{g}\boldsymbol{C}'\boldsymbol{A}^{-1} & -\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{E}^{g} \\ -\boldsymbol{E}^{g}\boldsymbol{C}'\boldsymbol{A}^{-1} & \boldsymbol{E}^{g} \end{bmatrix} \tag{1.107}$$

where $\boldsymbol{E}^{g}$ denotes the Moore–Penrose inverse of the Schur complement $\boldsymbol{E} = \boldsymbol{D} - \boldsymbol{C}'\boldsymbol{A}^{-1}\boldsymbol{B}$ of $\boldsymbol{A}$.

Likewise, if $\boldsymbol{D}$ is a non-singular matrix, then the following expression holds true for a reflexive generalized inverse in partitioned form of $\boldsymbol{P}$

$$\boldsymbol{P}_{\rho}^{-} = \begin{bmatrix} \boldsymbol{F}^{g} & -\boldsymbol{F}^{g}\boldsymbol{B}\boldsymbol{D}^{-1} \\ -\boldsymbol{D}^{-1}\boldsymbol{C}'\boldsymbol{F}^{g} & \boldsymbol{D}^{-1} + \boldsymbol{D}^{-1}\boldsymbol{C}'\boldsymbol{F}^{g}\boldsymbol{B}\boldsymbol{D}^{-1} \end{bmatrix} \tag{1.108}$$

where $\boldsymbol{F}^{g}$ denotes the Moore–Penrose inverse of the Schur complement $\boldsymbol{F} = \boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C}'$ of $\boldsymbol{D}$.

Proof

The proof is straightforward.

On the one hand, pre and post-multiplying the matrix on the right-hand side of (1.107) by $\boldsymbol{P}$, simple computations lead to find $\boldsymbol{P}$ itself.

On the other, pre and post-multiplying $\boldsymbol{P}$ by the matrix on the right-hand side of (1.107) simple computations lead to find this same matrix.

This proves (1.107).

Formula (1.108) is proved by a verbatim repetition of the proof of (1.107).

□

Let us now state an additional result.

Proposition

Let $\boldsymbol{\Lambda}$ be a full column-rank matrix. A general solution of the homogeneous equation

$$\boldsymbol{\Lambda}'\boldsymbol{X} = \boldsymbol{0} \tag{1.109}$$

is

$$\boldsymbol{X} = \boldsymbol{\Lambda}_{\perp}\boldsymbol{K} \tag{1.110}$$

where $\boldsymbol{K}$ is an arbitrary matrix.

Proof

According to Theorem 2.3.1 in Rao and Mitra, a general solution of (1.109) is

$$\boldsymbol{X} = (\boldsymbol{I} - (\boldsymbol{\Lambda}')^{-}\boldsymbol{\Lambda}')\boldsymbol{Z} \tag{1.111}$$

where $\boldsymbol{Z}$ is an arbitrary matrix.

Choosing $\boldsymbol{\Lambda}^{-} = \boldsymbol{\Lambda}^{g}$ and resorting to (1.52) of Sect. 1.2, such a solution can be rewritten as $\boldsymbol{X} = \boldsymbol{\Lambda}_{\perp}(\boldsymbol{\Lambda}'_{\perp}\boldsymbol{\Lambda}_{\perp})^{-1}\boldsymbol{\Lambda}'_{\perp}\boldsymbol{Z}$ and eventually as in (1.110) by putting $\boldsymbol{K} = (\boldsymbol{\Lambda}'_{\perp}\boldsymbol{\Lambda}_{\perp})^{-1}\boldsymbol{\Lambda}'_{\perp}\boldsymbol{Z}$

□

We shall now establish one of the main results (see also Faliva and Zoia 2002).

Theorem 3

Consider the block matrix

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix} \tag{1.112}$$

where $\boldsymbol{A}$, $\boldsymbol{B}$ *and* $\boldsymbol{C}$ *are as in Theorem 1.*

The condition

$$det(\boldsymbol{B}'_{\perp}\boldsymbol{A}\boldsymbol{C}_{\perp}) \neq 0 \leftrightarrow r(\boldsymbol{B}'_{\perp}\boldsymbol{A}\boldsymbol{C}_{\perp}) = m - n \tag{1.113}$$

is necessary and sufficient for the existence of $\boldsymbol{P}^{-1}$.

Further, the following representations of $\boldsymbol{P}^{-1}$ *hold*

$$\boldsymbol{P}^{-1} = \begin{bmatrix} \boldsymbol{H} & (\boldsymbol{I}-\boldsymbol{HA})(\boldsymbol{C}')^g \\ \boldsymbol{B}^g(\boldsymbol{I}-\boldsymbol{AH}) & \boldsymbol{B}^g(\boldsymbol{AHA}-\boldsymbol{A})(\boldsymbol{C}')^g \end{bmatrix} \tag{1.114}$$

$$= \begin{bmatrix} \boldsymbol{H} & \boldsymbol{K}(\boldsymbol{C}'\boldsymbol{K})^{-1} \\ (\tilde{\boldsymbol{K}}'\boldsymbol{B})^{-1}\tilde{\boldsymbol{K}}' & -(\tilde{\boldsymbol{K}}'\boldsymbol{B})^{-1}\tilde{\boldsymbol{K}}'\,\boldsymbol{AK}(\boldsymbol{C}'\boldsymbol{K})^{-1} \end{bmatrix} \tag{1.115}$$

where

$$\boldsymbol{H} = \boldsymbol{C}_{\perp}(\boldsymbol{B}'_{\perp}\boldsymbol{A}\boldsymbol{C}_{\perp})^{-1}\boldsymbol{B}'_{\perp} \tag{1.116}$$

$$\boldsymbol{K} = (\boldsymbol{A}'\boldsymbol{B}_{\perp})_{\perp},\ \tilde{\boldsymbol{K}} = (\boldsymbol{A}\boldsymbol{C}_{\perp})_{\perp} \tag{1.117}$$

Proof

Condition (1.113) follows from the rank identity (see Marsaglia and Styan 1974, Theorem 19)

$$\begin{aligned} r(\boldsymbol{P}) &= r(\boldsymbol{B}) + r(\boldsymbol{C}) + r\,[(\boldsymbol{I}-\boldsymbol{BB}^g)\boldsymbol{A}(\boldsymbol{I}-(\boldsymbol{C}')^g\boldsymbol{C}')] \\ &= n + n + r\,[(\boldsymbol{B}'_{\perp})^g\boldsymbol{B}'_{\perp}\boldsymbol{A}\ \boldsymbol{C}_{\perp}(\boldsymbol{C}_{\perp})^g] \\ &= 2n + r(\boldsymbol{B}'_{\perp}\boldsymbol{A}\boldsymbol{C}_{\perp}) = 2n + m - n = m + n \end{aligned} \tag{1.118}$$

where use has been made of the identities (1.87) and (1.88) of Sect. 1.2.

To prove (1.114), let the inverse of $\boldsymbol{P}$ be

$$\boldsymbol{P}^{-1} = \begin{bmatrix} \boldsymbol{P}_1 & \boldsymbol{P}_2 \\ \boldsymbol{P}_3 & \boldsymbol{P}_4 \end{bmatrix}$$

where the blocks in $\boldsymbol{P}^{-1}$ are of the same order as the corresponding blocks in $\boldsymbol{P}$. Then, in order to express the blocks of the former in terms of the blocks of the latter, write $\boldsymbol{P}^{-1}\boldsymbol{P} = \boldsymbol{I}$ and $\boldsymbol{PP}^{-1} = \boldsymbol{I}$ in partitioned form

$$\begin{bmatrix} \boldsymbol{P}_1 & \boldsymbol{P}_2 \\ \boldsymbol{P}_3 & \boldsymbol{P}_4 \end{bmatrix}\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}_m & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_n \end{bmatrix} \tag{1.119}$$

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{P}_1 & \boldsymbol{P}_2 \\ \boldsymbol{P}_3 & \boldsymbol{P}_4 \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}_m & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_n \end{bmatrix} \tag{1.119$'$}$$

and equate block to block as follows

$$\begin{cases} \boldsymbol{P}_1\boldsymbol{A}+\boldsymbol{P}_2\boldsymbol{C}'=\boldsymbol{I}_m & (1.120)\\ \boldsymbol{P}_1\boldsymbol{B}=\boldsymbol{0} & (1.121)\\ \boldsymbol{P}_3\boldsymbol{A}+\boldsymbol{P}_4\boldsymbol{C}'=\boldsymbol{0} & (1.122)\\ \boldsymbol{P}_3\boldsymbol{B}=\boldsymbol{I}_n & (1.123)\end{cases}$$

$$\begin{cases} \boldsymbol{A}\boldsymbol{P}_1+\boldsymbol{B}\boldsymbol{P}_3=\boldsymbol{I}_m & (1.120')\\ \boldsymbol{A}\boldsymbol{P}_2+\boldsymbol{B}\boldsymbol{P}_4=\boldsymbol{0} & (1.121')\\ \boldsymbol{C}'\boldsymbol{P}_1=\boldsymbol{0} & (1.122')\\ \boldsymbol{C}'\boldsymbol{P}_2=\boldsymbol{I}_n & (1.123')\end{cases}$$

From (1.120) and (1.120′) we get

$$\boldsymbol{P}_2=(\boldsymbol{C}')^g-\boldsymbol{P}_1\boldsymbol{A}\,(\boldsymbol{C}')^g=(\boldsymbol{I}-\boldsymbol{P}_1\boldsymbol{A})\,(\boldsymbol{C}')^g \tag{1.124}$$

$$\boldsymbol{P}_3=\boldsymbol{B}^g-\boldsymbol{B}^g\boldsymbol{A}\boldsymbol{P}_1=\boldsymbol{B}^g\,(\boldsymbol{I}-\boldsymbol{A}\boldsymbol{P}_1) \tag{1.125}$$

respectively.

From (1.121'), in light of (1.124), we can write

$$\boldsymbol{P}_4=-\boldsymbol{B}^g\boldsymbol{A}\boldsymbol{P}_2=-\boldsymbol{B}^g\boldsymbol{A}\,[(\boldsymbol{C}')^g-\boldsymbol{P}_1\boldsymbol{A}\,(\boldsymbol{C}')^g]=\boldsymbol{B}^g\,[\boldsymbol{A}\boldsymbol{P}_1\boldsymbol{A}-\boldsymbol{A}](\boldsymbol{C}')^g \tag{1.126}$$

Consider now the equation (121). Solving for $\boldsymbol{P}_1$ gives

$$\boldsymbol{P}_1=\boldsymbol{V}\boldsymbol{B}_\perp' \tag{1.127}$$

for some $\boldsymbol{V}$ (see Proposition above). Substituting the right-hand side of (1.127) for $\boldsymbol{P}_1$ in (1.120) and post-multiplying both sides by $\boldsymbol{C}_\perp$ we get

$$\boldsymbol{V}\boldsymbol{B}_\perp'\boldsymbol{A}\,\boldsymbol{C}_\perp=\boldsymbol{C}_\perp \tag{1.128}$$

which solved for $\boldsymbol{V}$ yields

$$\boldsymbol{V}=\boldsymbol{C}_\perp\,(\boldsymbol{B}_\perp'\boldsymbol{A}\,\boldsymbol{C}_\perp)^{-1} \tag{1.129}$$

in view of (1.113).

Substituting the right-hand side of (1.129) for $\boldsymbol{V}$ in (1.127) we obtain

$$\boldsymbol{P}_1=\boldsymbol{C}_\perp\,(\boldsymbol{B}_\perp'\boldsymbol{A}\,\boldsymbol{C}_\perp)^{-1}\boldsymbol{B}_\perp' \tag{1.130}$$

Hence, substituting the right-hand side of (1.130) for $\boldsymbol{P}_1$ in (1.124), (1.125) and (1.126), the expressions of the other blocks are easily found.

The proof of (1.115) follows as a by-product of (1.114), in light of identity (1.53) of Sect. 1.2, upon noticing that, on the one hand

$$\begin{aligned}\boldsymbol{I} - \boldsymbol{A}\boldsymbol{H} = \boldsymbol{I} - (\boldsymbol{A}\,\boldsymbol{C}_{\perp})\,(\boldsymbol{B}'_{\perp}\,(\boldsymbol{A}\,\boldsymbol{C}_{\perp}))^{-1}\,\boldsymbol{B}'_{\perp} = \boldsymbol{B}\,((\boldsymbol{A}\boldsymbol{C}_{\perp})'_{\perp}\;\boldsymbol{B})^{-1}\,(\boldsymbol{A}\boldsymbol{C}_{\perp})'_{\perp} \\ = \boldsymbol{B}\,(\widetilde{\boldsymbol{K}}'\,\boldsymbol{B})^{-1}\,\widetilde{\boldsymbol{K}}'\end{aligned} \tag{1.131}$$

whereas, on the other hand,

$$\boldsymbol{I} - \boldsymbol{H}\boldsymbol{A} = \boldsymbol{K}\,(\boldsymbol{C}'\boldsymbol{K})^{-1}\boldsymbol{C}' \tag{1.132}$$

□

The following corollaries provide further results whose usefulness will soon become clear.

Corollary 3.1

Should both assumptions (a) and (b) of Theorem 1 hold, then the following equality

$$(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C}')^{-1} = \boldsymbol{A}^{-1} + \boldsymbol{A}^{-1}\boldsymbol{B}\,(\boldsymbol{D} - \boldsymbol{C}'\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\boldsymbol{C}'\boldsymbol{A}^{-1} \tag{1.133}$$

would ensue.

Proof

Result (1.133) arises from equating the upper diagonal blocks of the right-hand sides of (1.104) and (1.105).

□

Corollary 3.2

Should both assumption (a) of Theorem 1 with $\boldsymbol{D} = \boldsymbol{0}$, and assumption (1.113) of Theorem 3 hold, then the equality

$$\boldsymbol{C}_{\perp}\,(\boldsymbol{B}'_{\perp}\boldsymbol{A}\,\boldsymbol{C}_{\perp})^{-1}\,\boldsymbol{B}'_{\perp} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}\,(\boldsymbol{C}'\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\boldsymbol{C}'\boldsymbol{A}^{-1} \tag{1.134}$$

would ensue.

Proof

Result (1.134) arises from equating the upper diagonal blocks of the right-hand sides of (1.114) and (1.104) for $\boldsymbol{D} = \boldsymbol{0}$

□

Corollary 3.3

By taking $\boldsymbol{D} = -\lambda\boldsymbol{I}$, let both assumption (b) of Theorem 1 in a deleted neighbourhood of $\lambda = 0$, and assumption (1.113) of Theorem 3 hold. Then the following equality

$$\boldsymbol{C}_{\perp}(\boldsymbol{B}'_{\perp}\boldsymbol{A}\,\boldsymbol{C}_{\perp})^{-1}\boldsymbol{B}'_{\perp} = \lim_{\lambda\to 0}\left\{\lambda\left(\lambda\boldsymbol{A}+\boldsymbol{B}\boldsymbol{C}'\right)^{-1}\right\} \tag{1.135}$$

ensues as $\lambda \to 0$.

Proof

To prove (1.135) observe that $\lambda^{-1}(\lambda\boldsymbol{A}+\boldsymbol{B}\boldsymbol{C}')$ plays the role of Schur complement of $\boldsymbol{D} = -\lambda\boldsymbol{I}$ in the partitioned matrix

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C}' & -\lambda\boldsymbol{I} \end{bmatrix} \tag{1.136}$$

whence

$$\lambda(\lambda\boldsymbol{A}+\boldsymbol{B}\boldsymbol{C}')^{-1} = [\boldsymbol{I},\ \ \boldsymbol{0}]\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C}' & -\lambda\boldsymbol{I} \end{bmatrix}^{-1}\begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} \tag{1.137}$$

Taking the limit as $\lambda \to 0$ of both sides of (1.137) yields

$$\lim_{\lambda\to 0}\left\{\lambda\left(\lambda\boldsymbol{A}+\boldsymbol{B}\boldsymbol{C}'\right)^{-1}\right\} = [\boldsymbol{I},\ \ \boldsymbol{0}]\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix}^{-1}\begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} \tag{1.138}$$

which eventually leads to (1.135) in view of (1.114).

□

Remark

Should $\boldsymbol{A} = \boldsymbol{I}$ and the product $\boldsymbol{B}\boldsymbol{C}' = \boldsymbol{\Lambda}$ be a matrix of index $\upsilon \le 2$, then Corollaries 7.6.2 and 7.6.4 in Campbell and Meyer would apply and the noteworthy results

$$\lim_{z\to 0} z(z\boldsymbol{I}+\boldsymbol{\Lambda})^{-1} = \boldsymbol{I} - \boldsymbol{\Lambda}^{\neq}\boldsymbol{\Lambda} = \boldsymbol{C}_{\perp}(\boldsymbol{B}'_{\perp}\,\boldsymbol{C}_{\perp})^{-1}\,\boldsymbol{B}'_{\perp} \quad \textit{under}\ \upsilon = 1, \tag{1.139}$$

$$\lim_{z\to 0} z^2(z\boldsymbol{I}+\boldsymbol{\Lambda})^{-1} = -(\boldsymbol{I} - \boldsymbol{\Lambda}^{D}\boldsymbol{\Lambda})\boldsymbol{\Lambda} = -\boldsymbol{C}_{2\perp}(\boldsymbol{B}'_{2\perp}\boldsymbol{C}_{2\perp})^{-1}\boldsymbol{B}'_{2\perp}\,\boldsymbol{\Lambda} \quad \textit{under}\ \upsilon = 2, \tag{1.139'}$$

would hold accordingly (see also (1.89) and (1.89′) of Sect. 1.2). Here $\boldsymbol{B}_2$ and $\boldsymbol{C}_2$ are obtained from the rank factorization $\boldsymbol{\Lambda}^2 = \boldsymbol{B}_2\boldsymbol{C}_2'$.

A mirror image of Theorem 3 is established in the next .

Theorem 4

Let $\boldsymbol{Я}$ and $\boldsymbol{K}$ be full column-rank matrices of order $m \times n$, $\boldsymbol{R}$ and $\boldsymbol{S}$ be full column-rank matrices of order $m \times (m-n)$, and $\boldsymbol{\aleph}$ be an arbitrary square matrix of order n.

Consider the partitioned matrix of order $m+n$,

$$\boldsymbol{Q} = \begin{bmatrix} \boldsymbol{RS}' & \boldsymbol{K} \\ \boldsymbol{Я}' & \boldsymbol{\aleph} \end{bmatrix} \tag{1.140}$$

If

$$det(\boldsymbol{R}_\perp'\boldsymbol{K}) \neq 0, \quad det(\boldsymbol{Я}'\boldsymbol{S}_\perp) \neq 0 \tag{1.141}$$

then the following hold true
(a) The matrix $\boldsymbol{Q}$ is non-singular
(b) Its inverse can be written in partitioned form as

$$\boldsymbol{Q}^{-1} = \begin{bmatrix} \boldsymbol{T} & \boldsymbol{S}_\perp(\boldsymbol{Я}'\boldsymbol{S}_\perp)^{-1} \\ (\boldsymbol{R}_\perp'\boldsymbol{K})^{-1}\boldsymbol{R}_\perp' & \boldsymbol{0} \end{bmatrix} \tag{1.142}$$

where

$$\boldsymbol{T} = \boldsymbol{Я}_\perp(\boldsymbol{S}'\boldsymbol{Я}_\perp)^{-1}(\boldsymbol{K}_\perp'\boldsymbol{R})^{-1}\boldsymbol{K}_\perp' - \boldsymbol{S}_\perp(\boldsymbol{Я}'\boldsymbol{S}_\perp)^{-1}\boldsymbol{\aleph}(\boldsymbol{R}_\perp'\boldsymbol{K})^{-1}\boldsymbol{R}_\perp' \tag{1.143}$$

Proof

Proof of (a) By applying Theorem 19 of Marsaglia anf Styan to the matrix $\boldsymbol{Q}$, we have the rank equality

$$r(\boldsymbol{Q}) = r(\boldsymbol{RS}') + r([\boldsymbol{I} - \boldsymbol{RS}'(\boldsymbol{RS}')^g]\boldsymbol{K})$$

$$+ r(\boldsymbol{Я}'[\boldsymbol{I} - (\boldsymbol{RS}')^g\boldsymbol{RS}']) + r(\boldsymbol{W}) \tag{1.144}$$

where

$$\boldsymbol{W} = (\boldsymbol{I} - \boldsymbol{YY}^g)[\boldsymbol{\aleph} - \boldsymbol{Я}'(\boldsymbol{RS}')^g\boldsymbol{K}](\boldsymbol{I} - \boldsymbol{X}^g\boldsymbol{X}) \tag{1.145}$$

$$\boldsymbol{X} = [\boldsymbol{I} - \boldsymbol{RS}'(\boldsymbol{RS}')^g]\boldsymbol{K} = (\boldsymbol{R}_\perp')^g\boldsymbol{R}_\perp'\boldsymbol{K}, \tag{1.146}$$

$$\boldsymbol{Y} = \boldsymbol{Я}'[\boldsymbol{I} - (\boldsymbol{RS}')^{g}\boldsymbol{RS}'] = \boldsymbol{Я}'\boldsymbol{S}_{\perp}(\boldsymbol{S}_{\perp})^{g}$$

From (1.25) of Sect. 1.1 along with (1.87) and (1.88) of Sect. 1.2 we get

$$\boldsymbol{I} - \boldsymbol{RS}'(\boldsymbol{RS}')^{g} = \boldsymbol{I} - \boldsymbol{RS}'(\boldsymbol{S}')^{g}(\boldsymbol{R})^{g} = (\boldsymbol{R}'_{\perp})^{g}\boldsymbol{R}'_{\perp} \tag{1.147}$$

$$\boldsymbol{I} - (\boldsymbol{RS}')^{g}\boldsymbol{RS}' = \boldsymbol{S}_{\perp}(\boldsymbol{S}_{\perp})^{g} \tag{1.148}$$

and $\boldsymbol{X}$ and $\boldsymbol{Y}$ can be accordingly written as

$$\boldsymbol{X} = (\boldsymbol{R}'_{\perp})^{g}\boldsymbol{R}'_{\perp}\boldsymbol{K}, \qquad \boldsymbol{Y} = \boldsymbol{Я}'\boldsymbol{S}_{\perp}(\boldsymbol{S}_{\perp})^{g} \tag{1.149}$$

By inspection of (1.149) both $\boldsymbol{X}$ and $\boldsymbol{Y}'$ turn out to be full column-rank matrices, so that the products $\boldsymbol{X}^{g}\boldsymbol{X}$ and $\boldsymbol{YY}^{g}$ are identity matrices and $\boldsymbol{W}$ is a null matrix. This together with (1.144) leads to conclude that

$$\begin{aligned} r(\boldsymbol{Q}) &= m - n + r(\boldsymbol{R}'_{\perp}\boldsymbol{K}) + r(\boldsymbol{Я}'\boldsymbol{S}_{\perp}) \\ &= m - n + n + n = m + n \rightarrow det(\boldsymbol{Q}) \neq 0 \end{aligned} \tag{1.150}$$

Proof of (b) Let the inverse of $\boldsymbol{Q}$ be

$$\boldsymbol{Q}^{-1} = \begin{bmatrix} \boldsymbol{T} & \boldsymbol{U} \\ \boldsymbol{W} & \boldsymbol{Z} \end{bmatrix} \tag{1.151}$$

where the blocks in $\boldsymbol{Q}^{-1}$ are of the same orders as the corresponding blocks in $\boldsymbol{Q}$. Then, in order to express the blocks of the former in terms of the blocks of the latter, write $\boldsymbol{Q}^{-1}\boldsymbol{Q} = \boldsymbol{I}$ and $\boldsymbol{QQ}^{-1} = \boldsymbol{I}$ in partitioned form as

$$\begin{bmatrix} \boldsymbol{T} & \boldsymbol{U} \\ \boldsymbol{W} & \boldsymbol{Z} \end{bmatrix}\begin{bmatrix} \boldsymbol{RS}' & \boldsymbol{K} \\ \boldsymbol{Я}' & \aleph \end{bmatrix} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \tag{1.152}$$

$$\begin{bmatrix} \boldsymbol{RS}' & \boldsymbol{K} \\ \boldsymbol{H}' & \aleph \end{bmatrix}\begin{bmatrix} \boldsymbol{T} & \boldsymbol{U} \\ \boldsymbol{W} & \boldsymbol{Z} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \tag{1.153}$$

and equate block to block as follows,

$$\begin{cases} \boldsymbol{TRS}' + \boldsymbol{UЯ}' = \boldsymbol{I} & (1.154) \\ \boldsymbol{WRS}' + \boldsymbol{ZЯ}' = \boldsymbol{0} & (1.155) \\ \boldsymbol{TK} + \boldsymbol{U}\aleph = \boldsymbol{0} & (1.156) \\ \boldsymbol{WK} + \boldsymbol{Z}\aleph = \boldsymbol{I} & (1.157) \end{cases}$$

$$\begin{cases} \boldsymbol{RS'T} + \boldsymbol{KW} = \boldsymbol{I} & (1.154') \\ \boldsymbol{RS'U} + \boldsymbol{KZ} = \boldsymbol{0} & (1.155') \\ \boldsymbol{Я}'\boldsymbol{T} + \boldsymbol{\aleph W} = \boldsymbol{0} & (1.156') \\ \boldsymbol{Я}'\boldsymbol{U} + \boldsymbol{\aleph Z} = \boldsymbol{I} & (1.157') \end{cases}$$

Post-multiplying (1.155) by $\boldsymbol{S}_\perp$ leads to

$$\boldsymbol{Z}\,\boldsymbol{Я}'\boldsymbol{S}_\perp = \boldsymbol{0} \tag{1.158}$$

This, together with (1.141), entails that

$$\boldsymbol{Z} = \boldsymbol{0} \tag{1.159}$$

Post and pre-multiplying (1.154) and (1.154′) by $\boldsymbol{S}_\perp$ and $\boldsymbol{R}'_\perp$, respectively, yields the twin equations

$$\boldsymbol{U}\,\boldsymbol{Я}'\boldsymbol{S}_\perp = \boldsymbol{S}_\perp \tag{1.160}$$

$$\boldsymbol{R}'_\perp \boldsymbol{K}\,\boldsymbol{W} = \boldsymbol{R}'_\perp \tag{1.161}$$

whence

$$\boldsymbol{U} = \boldsymbol{S}_\perp (\boldsymbol{Я}'\boldsymbol{S}_\perp)^{-1} \tag{1.162}$$

$$\boldsymbol{W} = (\boldsymbol{R}'_\perp \boldsymbol{K})^{-1} \boldsymbol{R}'_\perp \tag{1.163}$$

because of (1.141).

Post-multiplying (1.154) and (1.156) by $(\boldsymbol{S}')^g$ and $\boldsymbol{W}$, respectively, we get the pair of equations

$$\boldsymbol{TR} + \boldsymbol{UH}'(\boldsymbol{S}')^g = (\boldsymbol{S}')^g \tag{1.164}$$

$$\boldsymbol{T}\boldsymbol{K}\,\boldsymbol{W} + \boldsymbol{U}\boldsymbol{\aleph}\,\boldsymbol{W} = \boldsymbol{0} \tag{1.165}$$

Since

$$\boldsymbol{K}\,\boldsymbol{W} = \boldsymbol{K}(\boldsymbol{R}'_\perp \boldsymbol{K})^{-1} \boldsymbol{R}'_\perp = \boldsymbol{I} - \boldsymbol{R}(\boldsymbol{K}'_\perp \boldsymbol{R})^{-1} \boldsymbol{K}'_\perp \tag{1.166}$$

$$\boldsymbol{U}\,\boldsymbol{Я}' = \boldsymbol{S}_\perp (\boldsymbol{Я}'\boldsymbol{S}_\perp)^{-1} \boldsymbol{Я}' = \boldsymbol{I} - \boldsymbol{Я}_\perp (\boldsymbol{S}'\boldsymbol{Я}_\perp)^{-1} \boldsymbol{S}' \tag{1.167}$$

in light of (1.162) and (1.163) above and (1.53) of Sect. 1.2, equations (1.164) and (1.165) can be rewritten as

$$\boldsymbol{TR} = \boldsymbol{Я}_{\perp}(\boldsymbol{S}'\boldsymbol{Я}_{\perp})^{-1} \tag{1.168}$$

$$\boldsymbol{T} = \boldsymbol{TR}(\boldsymbol{K}'_{\perp}\boldsymbol{R})^{-1}\boldsymbol{K}'_{\perp} - \boldsymbol{U}\aleph\boldsymbol{W} \tag{1.169}$$

Then, replacing $\boldsymbol{TR}$ in (1.169) with the right-hand side of (1.168), we obtain

$$\boldsymbol{T} = \boldsymbol{Я}_{\perp}(\boldsymbol{S}'\boldsymbol{Я}_{\perp})^{-1}(\boldsymbol{K}'_{\perp}\boldsymbol{R})^{-1}\boldsymbol{K}'_{\perp} - \boldsymbol{U}\aleph\boldsymbol{W} \tag{1.170}$$

which tallies with (1.143) by making use of (1.162) and (1.163) above.

Simple computations show that (1.157) is satisfied, which completes the proof.

□

As a by-product of the foregoing theorem let us mention the following interesting application to orthogonal complement algebra.

Corollary 4.1

Let $\boldsymbol{A}$, $\boldsymbol{A}_{\rho}^{-}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{R}$, $\boldsymbol{S}$, $\boldsymbol{B}_2$, $\boldsymbol{C}_2$ be defined as in Theorem 8 of Sect. 1.2. Further, by introducing the matrices

$$\boldsymbol{Я} = \boldsymbol{C}'_{\perp}(\boldsymbol{A}_{\rho}^{-})'\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp} \tag{1.171}$$

$$\boldsymbol{K} = \boldsymbol{B}'_{\perp}\boldsymbol{A}_{\rho}^{-}\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp} \tag{1.172}$$

$$\boldsymbol{\Theta} = \boldsymbol{Я}_{\perp}(\boldsymbol{S}'\boldsymbol{Я}_{\perp})^{-1}(\boldsymbol{K}'_{\perp}\boldsymbol{R})^{-1}\boldsymbol{K}'_{\perp} \tag{1.173}$$

$$\boldsymbol{L} = \boldsymbol{R}'_{\perp}\boldsymbol{B}'_{\perp}\boldsymbol{A}_{\rho}^{-}\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp} = \boldsymbol{R}'_{\perp}\boldsymbol{K} = \boldsymbol{Я}'\boldsymbol{S}_{\perp} \tag{1.174}$$

$$\Delta = \boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp}\boldsymbol{L}^{-1}\boldsymbol{R}'_{\perp}\boldsymbol{B}'_{\perp} \tag{1.175}$$

$$\aleph = \boldsymbol{R}'_{\perp}\boldsymbol{B}'_{\perp}(\boldsymbol{A}_{\rho}^{-})^{2}\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp} \tag{1.176}$$

$$\tilde{\boldsymbol{T}} = \boldsymbol{\Theta} - \boldsymbol{S}_{\perp}\boldsymbol{L}^{-1}\aleph\boldsymbol{L}^{-1}\boldsymbol{R}'_{\perp} \tag{1.177}$$

we are able to establish the following equality

$$\boldsymbol{C}_{2\perp}(\boldsymbol{B}'_{2\perp}\boldsymbol{C}_{2\perp})^{-1}\boldsymbol{B}'_{2\perp} = \boldsymbol{C}_{\perp}\tilde{\boldsymbol{T}}\boldsymbol{B}'_{\perp} + \Delta\boldsymbol{A}_{\rho}^{-} + \boldsymbol{A}_{\rho}^{-}\Delta \tag{1.178}$$

Proof

In light of (1.69) of Sect. 1.2 and resorting to the partitioned inversion formula (1.142), the left-hand side of (1.178) can be worked out in this way

$$[\boldsymbol{C}_\perp,\ \boldsymbol{A}_\rho^- \boldsymbol{C}_\perp \boldsymbol{S}_\perp] \begin{bmatrix} \boldsymbol{R}\boldsymbol{S}' & \boldsymbol{K} \\ \boldsymbol{Я}' & \boldsymbol{\aleph} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{B}'_\perp \\ \boldsymbol{R}'_\perp \boldsymbol{B}'_\perp \boldsymbol{A}_\rho^- \end{bmatrix}$$

$$= [\boldsymbol{C}_\perp,\ \boldsymbol{A}_\rho^- \boldsymbol{C}_\perp \boldsymbol{S}_\perp] \begin{bmatrix} (\boldsymbol{\Theta} - \boldsymbol{S}_\perp \boldsymbol{L}^{-1} \boldsymbol{\aleph} \boldsymbol{L}^{-1} \boldsymbol{R}'_\perp) & \boldsymbol{S}_\perp \boldsymbol{L}^{-1} \\ \boldsymbol{L}^{-1} \boldsymbol{R}'_\perp & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{B}'_\perp \\ \boldsymbol{R}'_\perp \boldsymbol{B}'_\perp \boldsymbol{A}_\rho^- \end{bmatrix}$$

$$= \boldsymbol{C}_\perp \boldsymbol{\Theta}\ \boldsymbol{B}'_\perp - \boldsymbol{C}_\perp \boldsymbol{S}_\perp \boldsymbol{L}^{-1} \boldsymbol{\aleph} \boldsymbol{L}^{-1} \boldsymbol{R}'_\perp \boldsymbol{B}'_\perp + \boldsymbol{C}_\perp \boldsymbol{S}_\perp \boldsymbol{L}^{-1} \boldsymbol{R}'_\perp \boldsymbol{B}'_\perp \boldsymbol{A}_\rho^- + \boldsymbol{A}_\rho^- \boldsymbol{C}_\perp \boldsymbol{S}_\perp \boldsymbol{L}^{-1} \boldsymbol{R}'_\perp \boldsymbol{B}'_\perp$$

$$= \boldsymbol{C}_\perp \boldsymbol{\Theta}\ \boldsymbol{B}'_\perp - \boldsymbol{C}_\perp \boldsymbol{S}_\perp \boldsymbol{L}^{-1} \boldsymbol{\aleph} \boldsymbol{L}^{-1} \boldsymbol{R}'_\perp \boldsymbol{B}'_\perp + \boldsymbol{\Delta} \boldsymbol{A}_\rho^- + \boldsymbol{A}_\rho^- \boldsymbol{\Delta}$$

$$= \boldsymbol{C}_\perp (\boldsymbol{\Theta} - \boldsymbol{S}_\perp \boldsymbol{L}^{-1} \boldsymbol{\aleph} \boldsymbol{L}^{-1} \boldsymbol{R}'_\perp) \boldsymbol{B}'_\perp + \boldsymbol{\Delta} \boldsymbol{A}_\rho^- + \boldsymbol{A}_\rho^- \boldsymbol{\Delta} \tag{1.179}$$

□

1.5 Useful Matrix Decompositions

Following Rao and Mitra and Campbell and Meyer, we will now present a useful decomposition of a square matrix into a component of index one – henceforth called core component – and a nilpotent term.

Theorem 1

A singular non-null square matrix $\boldsymbol{A}$ *of index* $\upsilon \geq 1$ *has a unique decomposition*

$$\boldsymbol{A} = \boldsymbol{K} + \boldsymbol{H} \tag{1.180}$$

with the properties

(a) $$ind(\boldsymbol{K}) = 1 \tag{1.181}$$

(b) $$ind(\boldsymbol{H}) = \upsilon \tag{1.182}$$

(c) $$\boldsymbol{H}^\upsilon = \boldsymbol{0} \tag{1.183}$$

(d) $$\boldsymbol{HK} = \boldsymbol{KH} = \boldsymbol{0} \tag{1.184}$$

(e) $$r(\boldsymbol{A}^k) = r(\boldsymbol{H}^k) + r(\boldsymbol{K}), \quad k = 1, 2, ... \tag{1.185}$$

(f) $$\boldsymbol{A}^\upsilon = \boldsymbol{K}^\upsilon \tag{1.186}$$

(g) $$\boldsymbol{C}_{\upsilon\perp}(\boldsymbol{B}'_{\upsilon\perp}\boldsymbol{C}_{\upsilon\perp})^{-1}\boldsymbol{B}'_{\upsilon\perp} = \overline{\boldsymbol{C}}_{\perp}(\overline{\boldsymbol{B}}'_{\perp}\overline{\boldsymbol{C}}_{\perp})^{-1}\overline{\boldsymbol{B}}'_{\perp} \tag{1.187}$$

where the full column-rank matrices $\boldsymbol{B}_{\upsilon}$ *and* $\boldsymbol{C}_{\upsilon}$ *are obtained by rank factorization of* $\boldsymbol{A}^{\upsilon}$, *while* $\overline{\boldsymbol{B}}$ *and* $\overline{\boldsymbol{C}}$ *are obtained by rank factorization of* $\boldsymbol{K}$, *that is*

$$\boldsymbol{A}^{\upsilon} = \boldsymbol{B}_{\upsilon}\boldsymbol{C}'_{\upsilon} \tag{1.188}$$

$$\boldsymbol{K} = \overline{\boldsymbol{B}}\,\overline{\boldsymbol{C}}' \tag{1.189}$$

Proof

For proofs of (a)–(f) reference can be made to Rao and Mitra p 93, and Campbell and Meyer p 121. For what concerns (g) observe that

$$(\overline{\boldsymbol{B}}\,\overline{\boldsymbol{C}}')^{\upsilon} = \overline{\boldsymbol{B}}\;(\overline{\boldsymbol{C}}'\overline{\boldsymbol{B}})^{\upsilon-1}\;\overline{\boldsymbol{C}}' = \boldsymbol{B}_{\upsilon}\boldsymbol{C}'_{\upsilon} \tag{1.190}$$

in light of (1.186), (1.188) and (1.189). This implies that $\overline{\boldsymbol{B}}_{\perp}$ and $\overline{\boldsymbol{C}}_{\perp}$ act as orthogonal complements of $\boldsymbol{B}_{\upsilon}$ and $\boldsymbol{C}_{\upsilon}$, respectively, and (1.187) follows from Theorem 1 of Sect. 1.2.

□

The next result provides the so-called canonical form representations of a square matrix and of its Drazin inverse.

Theorem 2

With $\boldsymbol{A}$ *as in Theorem* 1, *there exist two non-singular matrices* $\boldsymbol{P}$ *and* $\boldsymbol{D}$ *and a nilpotent matrix* $\boldsymbol{\Upsilon}$ *of index* υ, *such that the following representations*

$$\boldsymbol{A} = \boldsymbol{P}\begin{bmatrix} \boldsymbol{D} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Upsilon} \end{bmatrix}\boldsymbol{P}^{-1} \tag{1.191}$$

$$\boldsymbol{A}^{D} = \boldsymbol{P}\begin{bmatrix} \boldsymbol{D}^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}\boldsymbol{P}^{-1} \tag{1.192}$$

hold true.

Proof

For a proof see Campbell and Meyer p 122.

□

Remark 1

Since $\boldsymbol{D}$ is non-singular, its eigenvalues are different from zero, and $\boldsymbol{I} - \boldsymbol{D}$ will accordingly have no unit roots.

Corollary 2.1

Partition $\boldsymbol{P}$ column-wise as

$$\boldsymbol{P} = [\boldsymbol{P}_1,\ \boldsymbol{P}_2] \tag{1.193}$$

and $\boldsymbol{P}^{-1}$ row-wise as

$$\boldsymbol{P}^{-1} = \begin{bmatrix} \boldsymbol{\Pi}_1 \\ \boldsymbol{\Pi}_2 \end{bmatrix} \tag{1.194}$$

so that $\boldsymbol{P}_1$, $\boldsymbol{D}$ and $\boldsymbol{\Pi}_1$ are conformable matrices, as well as $\boldsymbol{P}_2$, $\boldsymbol{\Upsilon}$, $\boldsymbol{\Pi}_2$.

Then, decompositions (1.191) and (1.192) can be rewritten as

$$\boldsymbol{A} = \boldsymbol{P}_1 \boldsymbol{D} \boldsymbol{\Pi}_1 + \boldsymbol{P}_2\ \boldsymbol{\Upsilon}\ \boldsymbol{\Pi}_2 \tag{1.195}$$

$$\boldsymbol{A}^D = \boldsymbol{P}_1 \boldsymbol{D}^{-1} \boldsymbol{\Pi}_1 \tag{1.196}$$

where $\boldsymbol{P}_1 \boldsymbol{D} \boldsymbol{\Pi}_1$ and $\boldsymbol{P}_2\ \boldsymbol{\Upsilon}\ \boldsymbol{\Pi}_2$ tally with the core and nilpotent parts of $\boldsymbol{A}$, respectively.

Proof

The proofs of (1.195) and (1.196) are straightforward.

By keeping in mind that $\boldsymbol{P}^{-1}\boldsymbol{P} = \boldsymbol{I}$, the equalities

$$\boldsymbol{\Pi}_1 \boldsymbol{P}_1 = \boldsymbol{I} \tag{1.197}$$

$$\boldsymbol{\Pi}_1 \boldsymbol{P}_2 = \boldsymbol{0} \tag{1.198}$$

$$\boldsymbol{\Pi}_2 \boldsymbol{P}_1 = \boldsymbol{0} \tag{1.199}$$

$$\boldsymbol{\Pi}_2 \boldsymbol{P}_2 = \boldsymbol{I} \tag{1.200}$$

are easily established. Then, it is easy to check that $\boldsymbol{P}_1 \boldsymbol{D} \boldsymbol{\Pi}_1$ and $\boldsymbol{P}_2 \boldsymbol{\Upsilon} \boldsymbol{\Pi}_2$ play the role of $\boldsymbol{K}$ and $\boldsymbol{H}$ in Theorem 1.

□

Remark 2

Even though the decomposition (1.191) is not unique, the quantities $\boldsymbol{P}_1 \boldsymbol{D} \boldsymbol{\Pi}_1$ and $\boldsymbol{P}_2 \boldsymbol{\Upsilon} \boldsymbol{\Pi}_2$ turn out to be unique, because of the uniqueness of the core-nilpotent decomposition (1.180).

The forthcoming theorems shed further light on the decompositions of matrices whose index is one or two.

Theorem 3

Let $\boldsymbol{A}$, $\boldsymbol{P}_1$, $\boldsymbol{P}_2$, $\boldsymbol{\Pi}_1$, $\boldsymbol{\Pi}_2$ *and* $\boldsymbol{D}$ *as in Corollary* 2.1, *and let*

$$ind(\boldsymbol{A}) = 1 \tag{1.201}$$

then, the following hold

(1)

$$\boldsymbol{A} = \boldsymbol{P}_1 \boldsymbol{D} \boldsymbol{\Pi}_1 \tag{1.202}$$

$$\boldsymbol{A}^D = \boldsymbol{P}_1 \boldsymbol{D}^{-1} \boldsymbol{\Pi}_1 \tag{1.203}$$

(2) $\boldsymbol{P}_1$, $\boldsymbol{P}_2$, $\boldsymbol{\Pi}_1$, $\boldsymbol{\Pi}_2$ *and* $\boldsymbol{D}$ *can be chosen as*

$$\boldsymbol{P}_1 = \boldsymbol{B} \tag{1.204}$$

$$\boldsymbol{P}_2 = \boldsymbol{C}_\perp \tag{1.205}$$

$$\boldsymbol{\Pi}_1 = (\boldsymbol{C}'\boldsymbol{B})^{-1} \boldsymbol{C}' \tag{1.206}$$

$$\boldsymbol{\Pi}_2 = (\boldsymbol{B}'_\perp \boldsymbol{C}_\perp)^{-1} \boldsymbol{B}'_\perp \tag{1.207}$$

$$\boldsymbol{D} = \boldsymbol{C}'\boldsymbol{B} \tag{1.208}$$

where $\boldsymbol{B}$ *and* $\boldsymbol{C}$ *are full column-rank matrices arising from rank factorization of* $\boldsymbol{A}$, *that is* $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}'$.

Proof

Observe first that the nilpotent matrix $\boldsymbol{\Upsilon}$ of representation (1.191) becomes a null matrix under (1.201), so that (1.195) of Corollary 2.1 reduces to (1.202).

Representation (1.203) corresponds to (1.196) with $\boldsymbol{A}^D$ replaced by $\boldsymbol{A}^{\neq}$ because $ind(\boldsymbol{A}) = 1$ (see Sect. 1.1, Definition 6).

Then, bearing in mind (1.202), (1.203) and (1.197), and resorting to identity (1.86) of Sect. 1.2, simple computations show that

$$\boldsymbol{P}_1\boldsymbol{\Pi}_1 = \boldsymbol{A}\boldsymbol{A}^{\neq} = \boldsymbol{B}(\boldsymbol{C}'\boldsymbol{B})^{-1}\boldsymbol{C}' \tag{1.209}$$

$$\boldsymbol{P}_1\boldsymbol{\Pi}_1\boldsymbol{P}_1 = \boldsymbol{P}_1 = \boldsymbol{B}(\boldsymbol{C}'\boldsymbol{B})^{-1}\boldsymbol{C}'\boldsymbol{P}_1 \tag{1.210}$$

whence we get the equation

$$\boldsymbol{B}'_{\perp}\boldsymbol{P}_1 = \boldsymbol{0} \tag{1.211}$$

The solution of (1.211) is given by

$$\boldsymbol{P}_1 = \boldsymbol{B}\boldsymbol{W} \tag{1.212}$$

for some $\boldsymbol{W}$, which can conveniently be chosen as $\boldsymbol{I}$. This eventually leads to (1.204) and to (1.206) in light of (1.209).

For what concerns $\boldsymbol{P}_2$ and $\boldsymbol{\Pi}_2$, notice how from $\boldsymbol{P}\boldsymbol{P}^{-1} = \boldsymbol{I}$ it follows that

$$\boldsymbol{P}_2\boldsymbol{\Pi}_2 = \boldsymbol{I} - \boldsymbol{P}_1\boldsymbol{\Pi}_1 = \boldsymbol{C}_{\perp}(\boldsymbol{B}'_{\perp}\boldsymbol{C}_{\perp})^{-1}\boldsymbol{B}'_{\perp} \tag{1.213}$$

where use has been made of Theorem 7 of Sect. 1.2.

Hereinafter, the proof of (1.205) and (1.207) can be obtained by a verbatim repetition of the arguments used to prove (1.204) and (1.206).

Finally, for what concerns $\boldsymbol{D}$ notice that

$$\boldsymbol{B}\boldsymbol{C}' = \boldsymbol{A} = \boldsymbol{P}_1\boldsymbol{D}\boldsymbol{\Pi}_1 = \boldsymbol{B}\boldsymbol{D}(\boldsymbol{C}'\boldsymbol{B})^{-1}\boldsymbol{C}' \tag{1.214}$$

in light of (1.202), (1.204) and (1.206).

Pre and post-multiplying (1.214) by $\boldsymbol{B}^g$ and $(\boldsymbol{C}')^g$ respectively, yields

$$\boldsymbol{I} = \boldsymbol{D}(\boldsymbol{C}'\boldsymbol{B})^{-1} \tag{1.215}$$

whose solution gives (1.208).

□

Remark 3

As by-products we have the equalities

$$\boldsymbol{P}_2\boldsymbol{\Pi}_2 = \boldsymbol{I} - \boldsymbol{A}\boldsymbol{A}^{\neq},\ \boldsymbol{P}_1\boldsymbol{\Pi}_1 = \boldsymbol{I} - \boldsymbol{P}_2\boldsymbol{\Pi}_2 = \boldsymbol{A}\boldsymbol{A}^{\neq} \tag{1.216}$$

$$\boldsymbol{K} = \boldsymbol{P}_1\boldsymbol{D}\boldsymbol{\Pi}_1 = \boldsymbol{A} \tag{1.217}$$

Theorem 4

With $\boldsymbol{A}$, $\boldsymbol{P}_1$, $\boldsymbol{P}_2$, $\boldsymbol{\Pi}_1$, $\boldsymbol{\Pi}_2$, $\boldsymbol{D}$ *and* $\boldsymbol{\Upsilon}$ *as in Corollary* 2.1, *let*

$$ind(\boldsymbol{A}) = 2 \tag{1.218}$$

Then, the following hold

(1)
$$\boldsymbol{A} = \boldsymbol{P}_1\boldsymbol{D}\boldsymbol{\Pi}_1 + \boldsymbol{P}_2\boldsymbol{\Upsilon}\boldsymbol{\Pi}_2 \tag{1.219}$$

$$\boldsymbol{A}^D = \boldsymbol{P}_1\boldsymbol{D}^{-1}\boldsymbol{\Pi}_1 \tag{1.220}$$

(2) $\boldsymbol{P}_1$, $\boldsymbol{P}_2$, $\boldsymbol{\Pi}_1$, $\boldsymbol{\Pi}_2$, $\boldsymbol{D}$ *and* $\boldsymbol{\Upsilon}$ *can be chosen as*

$$\boldsymbol{P}_1 = \boldsymbol{B}_2 \tag{1.221}$$

$$\boldsymbol{P}_2 = \boldsymbol{C}_{2\perp} \tag{1.222}$$

$$\boldsymbol{\Pi}_1 = (\boldsymbol{C}_2'\boldsymbol{B}_2)^{-1}\boldsymbol{C}_2' = (\boldsymbol{G}'\boldsymbol{F})^{-2}\boldsymbol{C}_2' \tag{1.223}$$

$$\boldsymbol{\Pi}_2 = (\boldsymbol{B}_{2\perp}'\boldsymbol{C}_{2\perp})^{-1}\boldsymbol{B}_{2\perp}' \tag{1.224}$$

$$\boldsymbol{D} = (\boldsymbol{C}_2'\boldsymbol{B}_2)^{-1}\boldsymbol{C}_2'\boldsymbol{B}\boldsymbol{C}'\boldsymbol{B}_2 = \boldsymbol{G}'\boldsymbol{F} \tag{1.225}$$

$$\boldsymbol{\Upsilon} = \boldsymbol{C}_{2\perp}^{g}\boldsymbol{B}\boldsymbol{C}'\boldsymbol{C}_{2\perp} \tag{1.226}$$

where $\boldsymbol{B}$ *and* $\boldsymbol{C}$ *are as in Theorem* 3, $\boldsymbol{B}_2$ *and* $\boldsymbol{C}_2$ *are full column-rank matrices arising from rank factorization of* $\boldsymbol{A}^2$, *that is* $\boldsymbol{A}^2 = \boldsymbol{B}_2\boldsymbol{C}_2'$, *while* $\boldsymbol{G}$ *and* $\boldsymbol{F}$ *are full column-rank matrices arising from rank factorization of* $\boldsymbol{C}'\boldsymbol{B}$, *that is* $\boldsymbol{C}'\boldsymbol{B} = \boldsymbol{F}\boldsymbol{G}'$.

Proof

Decompositions (1.219) and (1.220) mirror (1.195) and (1.196) of Corollary 2.1, subject to $\boldsymbol{\Upsilon}^2 = \boldsymbol{0}$ because of (1.218).

As $ind(\boldsymbol{A}) = 2$, then $ind(\boldsymbol{A}^2) = 1$ and the following hold accordingly

$$\boldsymbol{A}^2 = (\boldsymbol{P}_1\boldsymbol{D}\boldsymbol{\Pi}_1 + \boldsymbol{P}_2\boldsymbol{\Upsilon}\boldsymbol{\Pi}_2)\,(\boldsymbol{P}_1\boldsymbol{D}\boldsymbol{\Pi}_1 + \boldsymbol{P}_2\boldsymbol{\Upsilon}\boldsymbol{\Pi}_2) = \boldsymbol{P}_1\boldsymbol{D}^2\boldsymbol{\Pi}_1 \tag{1.227}$$

$$(\boldsymbol{A}^2)^{\neq} = \boldsymbol{P}_1\boldsymbol{D}^{-2}\boldsymbol{\Pi}_1 \tag{1.228}$$

as straightforward computations show, bearing in mind (1.197)–(1.200).

Then, Theorem 3 applies to $\boldsymbol{A}^2$ and $(\boldsymbol{A}^2)^{\neq}$ as specified in (1.227) and (1.228), and statements (1.221)–(1.224) prove to be true accordingly.

In order to find an expression for $\boldsymbol{D}$, it is enough to refer to decomposition (1.219) and notice that

$$\begin{aligned}\boldsymbol{B}\boldsymbol{C}' = \boldsymbol{A} = \boldsymbol{P}_1\boldsymbol{D}\boldsymbol{\Pi}_1 + \boldsymbol{P}_2\,\boldsymbol{\Upsilon}\boldsymbol{\Pi}_2 = \boldsymbol{B}_2\boldsymbol{D}(\boldsymbol{C}_2'\boldsymbol{B}_2)^{-1}\,\boldsymbol{C}_2' + \\ + \boldsymbol{C}_{2\perp}\,\boldsymbol{\Upsilon}(\boldsymbol{B}_{2\perp}'\boldsymbol{C}_{2\perp})^{-1}\boldsymbol{B}_{2\perp}'\end{aligned} \tag{1.229}$$

in light of (1.221)–(1.224).

Then, pre and post-multiplying (1.229) by $(\boldsymbol{C}_2'\boldsymbol{B}_2)^{-1}\boldsymbol{C}_2'$ and $\boldsymbol{B}_2$ respectively, we get

$$(\boldsymbol{C}_2'\boldsymbol{B}_2)^{-1}\boldsymbol{C}_2'\,\boldsymbol{B}\boldsymbol{C}'\boldsymbol{B}_2 = \boldsymbol{D} \tag{1.230}$$

which in turn, by resorting to (1.78) of Sect. 1.2, can be more simply written

$$\boldsymbol{D} = \boldsymbol{G}'\boldsymbol{F} \tag{1.231}$$

In light of (1.231), the non-singularity of $\boldsymbol{D}$ follows from Corollary 8.1 of Sect. 1.2.

For what concerns $\boldsymbol{\Upsilon}$, refer to (1.229) and pre and post-multiply by $(\boldsymbol{B}_{2\perp}'\boldsymbol{C}_{2\perp})^{-1}\boldsymbol{B}_{2\perp}'$ and $\boldsymbol{C}_{2\perp}$, respectively. This leads to

$$(\boldsymbol{B}_{2\perp}'\boldsymbol{C}_{2\perp})^{-1}\boldsymbol{B}_{2\perp}'\,\boldsymbol{B}\boldsymbol{C}'\boldsymbol{C}_{2\perp} = \boldsymbol{\Upsilon} \tag{1.232}$$

which, bearing in mind (1.61') of Sect. 1.2 can be written as

$$(\boldsymbol{B}_{2\perp}'\boldsymbol{C}_{2\perp})^{-1}\boldsymbol{B}_{2\perp}'\boldsymbol{B}\boldsymbol{C}'\boldsymbol{C}_{2\perp} = \boldsymbol{C}_{2\perp}^{g}[\boldsymbol{I} - \boldsymbol{B}_2(\boldsymbol{C}_2'\boldsymbol{B}_2)^{-1}\boldsymbol{C}_2']\boldsymbol{B}\boldsymbol{C}'\boldsymbol{C}_{2\perp} = \boldsymbol{C}_{2\perp}^{g}\boldsymbol{B}\boldsymbol{C}'\boldsymbol{C}_{2\perp} = \boldsymbol{\Upsilon} \tag{1.232}$$

since in light of (1.78) of the same section

$$\boldsymbol{C}_2'\boldsymbol{B}\boldsymbol{C}'\boldsymbol{C}_{2\perp} = \boldsymbol{G}'\boldsymbol{C}'\boldsymbol{B}\boldsymbol{C}'\boldsymbol{C}_{2\perp} = \boldsymbol{G}'\boldsymbol{F}\boldsymbol{C}'\boldsymbol{C}_{2\perp} = \boldsymbol{0} \tag{1.233}$$

The nilpotency $\boldsymbol{\Upsilon}$ can be easily verified by considering that

$$\begin{aligned}&\boldsymbol{C}_{2\perp}^{g}\boldsymbol{B}\boldsymbol{C}'\boldsymbol{C}_{2\perp}\boldsymbol{C}_{2\perp}^{g}\boldsymbol{B}\boldsymbol{C}'\boldsymbol{C}_{2\perp}=\boldsymbol{C}_{2\perp}^{g}\boldsymbol{B}\boldsymbol{C}'(\boldsymbol{I}-\boldsymbol{C}_{2}\boldsymbol{C}_{2\perp}^{g})\boldsymbol{B}\boldsymbol{C}'\boldsymbol{C}_{2\perp}=\\&\boldsymbol{C}_{2\perp}^{g}\boldsymbol{B}\boldsymbol{C}'\boldsymbol{B}\boldsymbol{C}'\boldsymbol{C}_{2\perp}=\boldsymbol{C}_{2\perp}^{g}\boldsymbol{B}\boldsymbol{F}\boldsymbol{G}'\boldsymbol{C}'\boldsymbol{C}_{2\perp}=\boldsymbol{C}_{2\perp}^{g}\boldsymbol{B}_{2}\boldsymbol{C}_{2}'\boldsymbol{C}_{2\perp}=\boldsymbol{0}\end{aligned} \tag{1.234}$$

□

Corollary 4.1

The following noteworthy relationship

$$\boldsymbol{A}^{D}=(\boldsymbol{A}^{2})^{\neq}\boldsymbol{A} \tag{1.235}$$

holds true.

Proof

The proof is straighforward, since

$$\boldsymbol{A}^{D}=\boldsymbol{P}_{1}\boldsymbol{D}^{-1}\boldsymbol{\Pi}_{1}=\boldsymbol{P}_{1}\boldsymbol{D}^{-2}\boldsymbol{\Pi}_{1}(\boldsymbol{P}_{1}\boldsymbol{D}\boldsymbol{\Pi}_{1}+\boldsymbol{P}_{2}\boldsymbol{\Upsilon}\boldsymbol{\Pi}_{2})=(\boldsymbol{A}^{2})^{\neq}\boldsymbol{A} \tag{1.236}$$

in light of (1.197) and (1.198).

□

Remark 4

The following equalities are easy to verify because of the above results and of (1.86’) and (1.89’) of Sect. 1.2, (see also Campbell and Meyer)

$$\boldsymbol{P}_{2}\boldsymbol{\Pi}_{2}=\boldsymbol{C}_{2\perp}(\boldsymbol{B}_{2\perp}'\boldsymbol{C}_{2\perp})^{-1}\boldsymbol{B}_{2\perp}'=\boldsymbol{I}-\boldsymbol{A}(\boldsymbol{A})^{D}=\boldsymbol{I}-\boldsymbol{A}^{2}(\boldsymbol{A}^{2})^{\neq} \tag{1.237}$$

$$\boldsymbol{P}_{1}\boldsymbol{\Pi}_{1}=\boldsymbol{B}_{2}(\boldsymbol{C}_{2}'\boldsymbol{B}_{2})^{-1}\boldsymbol{C}_{2}'=\boldsymbol{B}_{2}(\boldsymbol{G}'\boldsymbol{F})^{-2}\boldsymbol{C}_{2}'=\boldsymbol{A}(\boldsymbol{A})^{D}=\boldsymbol{A}^{2}(\boldsymbol{A}^{2})^{\neq} \tag{1.238}$$

$$\boldsymbol{K}=\boldsymbol{P}_{1}\boldsymbol{D}\boldsymbol{\Pi}_{1}=(\boldsymbol{P}_{1}\boldsymbol{D}^{2}\boldsymbol{\Pi}_{1})(\boldsymbol{P}_{1}\boldsymbol{D}^{-1}\boldsymbol{\Pi}_{1})_{2}=\boldsymbol{A}^{2}\boldsymbol{A}^{D}=\boldsymbol{A}^{D}\boldsymbol{A}^{2} \tag{1.239}$$

$$\begin{aligned}\boldsymbol{H}&=\boldsymbol{P}_{2}\boldsymbol{\Upsilon}\boldsymbol{\Pi}_{2}=\boldsymbol{P}_{2}\boldsymbol{\Pi}_{2}\boldsymbol{A}\boldsymbol{P}_{2}\boldsymbol{\Pi}_{2}=[\boldsymbol{I}-\boldsymbol{A}\boldsymbol{A}^{D}]\boldsymbol{A}[\boldsymbol{I}-\boldsymbol{A}\boldsymbol{A}^{D}]\\&=\boldsymbol{A}[\boldsymbol{I}-\boldsymbol{A}\boldsymbol{A}^{D}]=\boldsymbol{A}\boldsymbol{P}_{2}\boldsymbol{\Pi}_{2}=\boldsymbol{A}\boldsymbol{C}_{2\perp}(\boldsymbol{B}_{2\perp}'\boldsymbol{C}_{2\perp})^{-1}\boldsymbol{B}_{2\perp}'\\&=\boldsymbol{A}[\boldsymbol{I}-\boldsymbol{B}_{2}(\boldsymbol{C}_{2}'\boldsymbol{B}_{2})^{-1}\boldsymbol{C}_{2}']=\boldsymbol{A}[\boldsymbol{I}-\boldsymbol{B}_{2}(\boldsymbol{G}'\boldsymbol{F})^{-2}\boldsymbol{C}_{2}']\\&=\boldsymbol{B}(\boldsymbol{I}-\boldsymbol{F}\boldsymbol{G}'\boldsymbol{F}(\boldsymbol{G}'\boldsymbol{F})^{-2}\boldsymbol{G}')\boldsymbol{C}'=\boldsymbol{B}\boldsymbol{G}_{\perp}(\boldsymbol{F}_{\perp}'\boldsymbol{G}_{\perp})^{-1}\boldsymbol{F}_{\perp}'\boldsymbol{C}'\end{aligned} \tag{1.240}$$

1.6 Matrix Polynomial Functions: Zeroes, Roots and Poles

We begin by introducing the following definitions

Definition 1 – Matrix Polynomial

A matrix polynomial in the scalar argument z is an analytical matrix function of the form

$$\boldsymbol{A}(z) = \sum_{k=0}^{K} \boldsymbol{A}_k z^k, \qquad \boldsymbol{A}_K \neq \boldsymbol{0} \tag{1.241}$$

In the following we assume that, unless otherwise stated, the $\boldsymbol{A}'_k s$ are square matrices of order n with $\boldsymbol{A}_0 = \boldsymbol{I}$ as a normalization. The number n is called the order of the polynomial $\boldsymbol{A}(z)$.

The number K is called the degree of the polynomial, provided $\boldsymbol{A}_K = \boldsymbol{0}$. When $K = 1$, the matrix polynomial is said to be linear.

Definition 2 – Zero of a Matrix Polynomial

A point $z = \zeta$ is said to be a zero, or a root, of the matrix polynomial $\boldsymbol{A}(z)$ if

$$r(\boldsymbol{A}(\zeta)) \leq n - 1 \tag{1.242}$$

that is to say, if $\boldsymbol{A}(\zeta)$ is a singular matrix. The matrix $\boldsymbol{A}(\zeta)$ is said to be simply degenerate if (1.242) holds with the equality sign, otherwise multiply degenerate.

Should (1.242) hold true for $\zeta = 1$, then ζ will be referred to as a *unit root* of $\boldsymbol{A}(z)$.

Definition 3 – Nullity

The quantity

$$\omega = n - r(\boldsymbol{A}(\zeta)) \tag{1.243}$$

is called nullity, or degeneracy, of the matrix $\boldsymbol{A}(\zeta)$.

The inverse $\boldsymbol{A}^{-1}(z)$ is an analytic matrix function throughout the z-plane except for the zeros of $\boldsymbol{A}(z)$, which play the rôle of isolated singular points – actually, poles of some orders – of the function $\boldsymbol{A}^{-1}(z)$.

Definition 4 – Pole

An isolated point, $z = \zeta$, of a matrix function $\boldsymbol{A}^{-1}(z)$, such that the Euclidian norm $\left\|\boldsymbol{A}^{-1}(z)\right\| \to \infty$ as $z = \zeta$, is called a pole of $\boldsymbol{A}^{-1}(z)$.

Definition 5 – Order of Poles and Zeros

Let $z = \zeta$ be a pole of the matrix function $\boldsymbol{A}^{-1}(z)$, the positive integer H for which the following limit turns out to be a non-null matrix with finite entries

$$\lim_{z \to \zeta} (\zeta - z)^H \boldsymbol{A}^{-1}(z) \tag{1.244}$$

is called order of the pole of $\boldsymbol{A}^{-1}(z)$, as well as order of the zeros of the parent matrix polynomial $\boldsymbol{A}(z)$.

A matrix polynomial $\boldsymbol{A}(z)$ can be expanded about $z = \zeta$ in this way

$$\boldsymbol{A}(z) = \boldsymbol{A}(\zeta) + \sum_{k=1}^{K} (\zeta - z)^k (-1)^k \frac{1}{k!} \boldsymbol{A}^{(k)}(\zeta) \tag{1.245}$$

where

$$\boldsymbol{A}^{(k)}(\zeta) = \left(\frac{\partial^k \boldsymbol{A}(z)}{\partial z^k} \right)_{z=\zeta} = k! \sum_{j=k}^{K} \binom{j}{k} \boldsymbol{A}_j \zeta^{j-k} \tag{1.246}$$

as simple computations show.

Dividing $\boldsymbol{A}(z)$ by the linear polynomial $\boldsymbol{I} - z\boldsymbol{I}$ yields

$$\boldsymbol{A}(z) = (1 - z)\, \boldsymbol{Q}(z) + \boldsymbol{A} \tag{1.247}$$

where

$$\boldsymbol{Q}(z) = \sum_{k=1}^{K} (1 - z)^{k-1} (-1)^k \frac{1}{k!} \boldsymbol{A}^{(k)}(1) \tag{1.248}$$

is called the quotient, and

$$\boldsymbol{A} = \boldsymbol{A}(1) = \sum_{k=0}^{K} \boldsymbol{A}_k \tag{1.249}$$

is called the remainder (Gantmacher, vol I, p 78).

Assuming that $\boldsymbol{A}$ is a non-null singular matrix and defining $\boldsymbol{B}$ and $\boldsymbol{C}$ as per the rank factorization

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}' \tag{1.250}$$

we obtain the result stated in

Theorem 1

The product of the scalar $(1 - z)$ *by the Schur complement of the lower diagonal block of the partitioned matrix*

$$\begin{bmatrix} \boldsymbol{Q}(z) & \boldsymbol{B} \\ \boldsymbol{C}' & (z-1)\boldsymbol{I} \end{bmatrix} \tag{1.251}$$

coincides with the matrix polynomial

$$\boldsymbol{A}(z) = (1 - z)\,\boldsymbol{Q}(z) + \boldsymbol{A} \tag{1.252}$$

Proof

The Schur complement of the block $(z - 1)\boldsymbol{I}$ is the matrix

$$\boldsymbol{Q}(z) - \boldsymbol{B}[(z-1)\boldsymbol{I}]^{-1}\boldsymbol{C}' = \boldsymbol{Q}(z) + \frac{1}{(1-z)}\boldsymbol{A} \tag{1.253}$$

which multiplied by $(1 - z)$ yields (1.252) as claimed above.

□

Furthermore, dividing $\boldsymbol{Q}(z)$ by the linear polynomial $\boldsymbol{I} - z\boldsymbol{I}$ we get

$$\boldsymbol{Q}(z) = (1 - z)\,\boldsymbol{\Psi}(z) - \dot{\boldsymbol{A}} \tag{1.254}$$

where

$$\boldsymbol{\Psi}(z) = \sum_{k=2}^{K} (1-z)^{k-2}(-1)^k \frac{1}{k!}\boldsymbol{A}^{(k)}(1) \tag{1.255}$$

and where

$$\dot{\boldsymbol{A}} = \boldsymbol{A}^{(1)}(1) = \sum_{k=1}^{K} k\boldsymbol{A}_k \tag{1.256}$$

Substituting the right-hand side of (1.254) into (1.247) leads to the representation

$$\boldsymbol{A}(z) = (1-z)^2\,\boldsymbol{\Psi}(z) - (1-z)\,\dot{\boldsymbol{A}} + \boldsymbol{A} \tag{1.257}$$

□

A result mirroring the one already established in Theorem 1 holds for (1.257), as we will see in the proof of Theorem 5 of Sect. 1.10.

The following equalities are worth mentioning

$$\boldsymbol{Q}(1) = -\dot{\boldsymbol{A}} \tag{1.258}$$

$$\boldsymbol{\Psi}(1) = -\dot{\boldsymbol{Q}}(1) = \frac{1}{2}\ddot{\boldsymbol{A}} \tag{1.259}$$

where the number of dots indicates the order of the derivative.

Likewise $\boldsymbol{A}(z)$, its adjoint matrix $\boldsymbol{A}^{+}(z)$ can be expanded about $z = \zeta$ namely

$$\boldsymbol{A}^{+}(z) = \boldsymbol{A}^{+}(\zeta) + \sum_{j \geq 1} \frac{(-1)^{j}}{j!}(\zeta - z)^{j} \boldsymbol{A}^{+(j)}(\zeta) \tag{1.260}$$

where $\boldsymbol{A}^{+(j)}(\zeta)$ is written for $\left(\dfrac{\partial^{j} \boldsymbol{A}^{+}(z)}{\partial z^{j}} \right)_{z=\zeta}$.

Further results are given in

Theorem 2

The adjoint matrix $\boldsymbol{A}^{+}(\zeta)$ *enjoys the following properties*

(1) The rank of $\boldsymbol{A}^{+}(\zeta)$ *is* 1 *if the nullity of* $\boldsymbol{A}(\zeta)$ *is* 1,

(2) The adjoint $\boldsymbol{A}^{+}(\zeta)$ *and its derivatives up to and including the* $(\omega - 2)$*-th at least, all vanish as the nullity of* $\boldsymbol{A}(\zeta)$ *is* $\omega \geq 2$.

Proof

For a proof see Fraser et al. 1963, p 61–62 .

□

Remark 1

Under (1) the adjoint is expressible as follows

$$\boldsymbol{A}^{+}(\zeta) = \boldsymbol{C}_{\perp} \kappa \, \boldsymbol{B}'_{\perp} \tag{1.261}$$

for some scalar κ and where $\boldsymbol{B}_{\perp}$ and $\boldsymbol{C}_{\perp}$ are the eigenvectors corresponding to the (simple) root ζ .

Under (2) the following holds

$$\left.\frac{\partial^{\varphi} \boldsymbol{A}^{+}(z)}{\partial z^{\varphi}}\right|_{z=\zeta} = \boldsymbol{C}_{\perp}\boldsymbol{\Phi}\boldsymbol{B}'_{\perp} \tag{1.262}$$

where $\varphi \geq \omega - 1$, $\boldsymbol{\Phi}$ is a square matrix, and the columns of $\boldsymbol{B}_{\perp}$ and $\boldsymbol{C}_{\perp}$ are the eigenvectors corresponding to the (multiple) root ζ.

Corollary 2.1

The adjoint matrix $\boldsymbol{A}^{+}(z)$ can be factored as follows

$$\boldsymbol{A}^{+}(z) = (-1)^{\varphi}(\zeta - z)^{\varphi}\boldsymbol{\Gamma}(z) \tag{1.263}$$

where

$$\varphi \geq \omega - 1 \tag{1.264}$$

$$\begin{aligned}\boldsymbol{\Gamma}\ (z) = {} & \frac{1}{\varphi!}\boldsymbol{A}^{+(\varphi)}(\zeta) - \frac{1}{(\varphi+1)!}(\zeta - z)\boldsymbol{A}^{+(\varphi+1)}(\zeta) \\ & + \textit{terms of higher powers of } (\zeta - z)\end{aligned} \tag{1.265}$$

$$\boldsymbol{\Gamma}\ (\zeta) = \frac{1}{\varphi!}\boldsymbol{A}^{+(\varphi)}(\zeta) \neq 0 \tag{1.266}$$

Proof

The expansion (1.260) can be split into two parts in this way

$$\boldsymbol{A}^{+}(z) = \sum_{j=0}^{\varphi-1}\frac{(-1)^{j}}{j!}(\zeta - z)^{j}\boldsymbol{A}^{+(j)}(\zeta) + (-1)^{\varphi}(\zeta - z)^{\varphi}\boldsymbol{\Gamma}(z) \tag{1.267}$$

where the first part vanishes according to Theorem 2, whence (1.263).

□

The scalar φ will be henceforth referred to as the adjoint-multiplicity of the root ζ.

We will now give a definition, and establish some useful theorems.

Definition 6 – Characteristic Polynomial

The scalar polynomial $det\boldsymbol{A}(z)$ is referred to as the characteristic polynomial of $\boldsymbol{A}(z)$.

Theorem 3

The characteristic polynomial can be expanded about $z = \zeta$ as

$$det\boldsymbol{A}(z) = \sum_{j \geq 0} \frac{(-1)^j}{j!}(\zeta - z)^j \pi_j \tag{1.268}$$

where

$$\pi_j = \left(\frac{\partial^j det\boldsymbol{A}(z)}{\partial z^j} \right)_{z=\zeta} = \sum_{i=0}^{j-1} \binom{j-1}{i} tr\left\{ \boldsymbol{A}^{+(j-1-i)}(\zeta)\boldsymbol{A}^{(i+1)}(\zeta) \right\} \tag{1.269}$$

and the symbols $\boldsymbol{A}^{+(r)}(\zeta)$ and $\boldsymbol{A}^{(s)}(\zeta)$ stand for $\left(\frac{\partial^r \boldsymbol{A}^+(z)}{\partial z^r} \right)_{z=\zeta}$ and $\left(\frac{\partial^s \boldsymbol{A}(z)}{\partial z^s} \right)_{z=\zeta}$ respectively, adopting the conventions that $\pi_0 = det\boldsymbol{A}(\zeta)$, $\boldsymbol{A}^{(0)}(\zeta) = \boldsymbol{A}(\zeta)$ and $\boldsymbol{A}^{+(0)}(\zeta) = \boldsymbol{A}^+(\zeta)$.

The scalars π_j's up to and including the $(\omega - 1)$-th at least, all vanish as the nullity of $\boldsymbol{A}(\zeta)$ is $\omega \geq 1$.

Proof

Formula (1.268) is easily established by expanding in series the characteristic polynomial, and noting that

$$\begin{aligned} \frac{\partial det\boldsymbol{A}(z)}{\partial z} &= \left[\frac{\partial det\boldsymbol{A}(z)}{\partial det\boldsymbol{A}(z)} \right]' \frac{\partial det\boldsymbol{A}(z)}{\partial z} \\ &= \{vec(\boldsymbol{A}^+(z))'\}' vec\dot{\boldsymbol{A}}(z) = tr\{\boldsymbol{A}^+(z)\dot{\boldsymbol{A}}(z)\} \end{aligned} \tag{1.270}$$

by making use of the stacked form $vec\boldsymbol{A}$ of $\boldsymbol{A}$, because of matrix differential calculus and *vec-trace* relationships (Faliva 1987; Magnus and Neudecker 1999).

Higher order derivatives can be obtained by applying Leibniz formula to the right-hand side of (1.270). Evaluating $det\boldsymbol{A}(z)$ and its derivatives for

$z = \zeta$ yields the coefficients π_j, which turn out to vanish up to and including the $(\omega - 1)$-*th* at least, by the same arguments as those underlying Theorem 2 (Fraser, Duncan and Collar).

□

An important consequence is the next factorization, as stated in

Corollary 3.1

The characteristic polynomial can be factored as follows

$$det\boldsymbol{A}(z) = (-1)^{\mu}(\zeta - z)^{\mu}\,\theta(z) \tag{1.271}$$

where μ – to be understood as the multiplicity of the root $z = \zeta$ in $det\boldsymbol{A}(z)$ – is such that

$$\mu \gtrapprox \omega \tag{1.272}$$

$$\theta\,(z) = \frac{1}{\mu!}\pi_{\mu} - \frac{1}{(\mu+1)!}(\zeta - z)\pi_{\mu+1} + terms\ of\ higher\ powers\ of\ (\zeta - z) \tag{1.273}$$

$$\theta\,(\zeta) = \frac{1}{\mu!}\pi_{\mu} \neq 0 \tag{1.274}$$

Proof

The proof is a verbatim repetition of that of Corollary 2.1, by taking expansion (1.268) as a starting point.

□

The scalar μ will be henceforth referred to as the determinantal-multiplicity of the root ζ.

Remark 2

The following results establish some interesting relationship between the nullity of a matrix $\boldsymbol{A}(\zeta)$ and the multiplicities of the root ζ.

(1) determinantal-multiplicity of the root $\zeta \geq$ *nullity of* $\boldsymbol{A}(\zeta)$

(2) (*determinantal-multiplicity of the root* ζ) – 1 ≥

$\geq$ *adjoint-multiplicitiy of the root* $\zeta \geq$

$\geq$ *nullity of* $\boldsymbol{A}(\zeta) - 2$

Theorem 4

For a linear matrix polynomial the following holds

$$\mu = n - r(\boldsymbol{A})^{\upsilon} \tag{1.275}$$

where υ *denotes the index of the argument matrix.*

Proof

Consider the linear matrix polynomial

$$\boldsymbol{A}(z) = (1-z)\boldsymbol{I} + z\boldsymbol{A} = (1-z)\boldsymbol{I}_n + z\boldsymbol{B}\boldsymbol{C}' \tag{1.276}$$

along with the block matrix

$$\begin{bmatrix} (1-z)\boldsymbol{I}_n & \boldsymbol{B} \\ \boldsymbol{C}' & -\frac{1}{z}\boldsymbol{I}_r \end{bmatrix}$$

where r is written for $r(\boldsymbol{A})$.

Making use of the classic determinantal factorizations (see, e.g., Rao 1973, p 32), we can express the determinant of the block matrix above as

$$\begin{aligned} det\begin{bmatrix} (1-z)\boldsymbol{I}_n & \boldsymbol{B} \\ \boldsymbol{C}' & -\frac{1}{z}\boldsymbol{I}_r \end{bmatrix} &= det((1-z)\boldsymbol{I}_{\mathrm{n}})\, det(-\frac{1}{z}\boldsymbol{I}_r - \boldsymbol{C}'[(1-z)\boldsymbol{I}_{\mathrm{n}}]^{-1}\boldsymbol{B}) \\ &= (1-z)^n(-1)^r\, det\left(\frac{1}{z}\boldsymbol{I}_r + \left(\frac{1}{1-z}\right)\boldsymbol{C}'\boldsymbol{B}\right) \\ &= (-1)^r z^{-r}(1-z)^{n-r}\, det[(1-z)\boldsymbol{I}_r + z\boldsymbol{C}'\boldsymbol{B}] \end{aligned} \tag{1.277}$$

as well as

$$\begin{aligned} det\begin{bmatrix} (1-z)\boldsymbol{I}_n & \boldsymbol{B} \\ \boldsymbol{C}' & -\frac{1}{z}\boldsymbol{I}_r \end{bmatrix} &= \\ = det(-\frac{1}{z}\boldsymbol{I}_r)det\left((1-z)\boldsymbol{I}_n - \boldsymbol{B}\left(-\frac{1}{z}\boldsymbol{I}_r\right)^{-1}\boldsymbol{C}'\right) \end{aligned} \tag{1.278}$$

$$= (-1)^r z^{-r} \, det[(1-z)\boldsymbol{I}_n + z\boldsymbol{B}\boldsymbol{C}'] = (-1)^r z^{-r} \, det\boldsymbol{A}(z)$$

Equating the right-hand sides of (1.277) and (1.278) we can express $det\boldsymbol{A}(z)$ as

$$det\boldsymbol{A}(z) = (1-z)^{n-r} \, \lambda(z) \tag{1.279}$$

where

$$\lambda(z) = det[(1-z)\boldsymbol{I}_r + \boldsymbol{C}'\boldsymbol{B}] \tag{1.280}$$

If $ind(\boldsymbol{A}) = 1$, according to Corollary 6.1 of Sect. 1.2, $\boldsymbol{C}'\boldsymbol{B}$ is a non singular matrix and $\lambda(z) \neq 0$ in a neighborhood of $z = 1$. This proves (1.275) for $\upsilon = 1$.

Repeating the above argument, with respect to $\lambda(z)$, we can extend our conclusion to $\boldsymbol{A}$ of index two, and iterating the argument we can eventually establish (1.275).

□

The following results about unit roots are worth mentioning.

Proposition

The characteristic polynomial $det\boldsymbol{A}(z)$ has a possibly multiple unit-root $z = 1$ if and only if

$$det\mathbf{A} = 0 \tag{1.281}$$

where $\boldsymbol{A}$ is written for $\boldsymbol{A}(1)$.

Proof

According to (1.252) the so-called characteristic equation

$$det\boldsymbol{A}(z) = 0 \tag{1.282}$$

can be exhibited in the form

$$det[(1-z)\,\boldsymbol{Q}(z) + \boldsymbol{A}] = 0 \tag{1.283}$$

which for $z = 1$ entails

$$det(\boldsymbol{A}) = 0 \Rightarrow r(\boldsymbol{A}) < n \tag{1.284}$$

and vice versa.

□

Theorem 5

We consider two possibilities

(a) z = 1 *is a simple root of the characteristic polynomial det**A**(*z*) if and only if*

$$det\mathbf{A} = 0 \tag{1.285}$$

$$tr(\boldsymbol{A}^{+}(1)\dot{\boldsymbol{A}}) \neq 0 \tag{1.286}$$

where $\boldsymbol{A}^{+}(1)$ *denotes the adjoint matrix* $\boldsymbol{A}^{+}$(z) *of* $\boldsymbol{A}$(z) *evaluated at* z = 1

(b) $z = 1$ *is a root of multiplicity two of the characteristic polynomial det**A**(*z*) if and only if*

$$det\mathbf{A} = 0 \tag{1.287}$$

$$tr(\boldsymbol{A}^{+}(1)\dot{\boldsymbol{A}}) = 0 \tag{1.288}$$

$$tr(\dot{\boldsymbol{A}}^{+}(1)\dot{\boldsymbol{A}} + \boldsymbol{A}^{+}(1)\ddot{\boldsymbol{A}}) \neq 0 \tag{1.289}$$

where $\dot{\boldsymbol{A}}^{+}(1)$ *denotes the derivative of* $\boldsymbol{A}^{+}$ (z) *with respect to* z *evaluated at* z = *1*.

Proof

The proof is straightforward bearing in mind (1.270) and upon noting that, according to (1.269) the following holds

$$\begin{aligned}\pi_2 &= \left[\frac{\partial^2 det\boldsymbol{A}(z)}{\partial z^2}\right]_{z=1} = \left[\frac{\partial tr\{\boldsymbol{A}^{+}(z)\dot{\boldsymbol{A}}(z)\}}{\partial z}\right]_{z=1} \\ &= tr(\dot{\boldsymbol{A}}^{+}(1)\dot{\boldsymbol{A}} + \boldsymbol{A}^{+}(1)\ddot{\boldsymbol{A}})\end{aligned} \tag{1.290}$$

□

Let us now address the issue about the inversion of the matrix $\boldsymbol{A}(z)$ in a deleted neighbourhood of its zero $z = \zeta$. We will establish the following theorem.

Theorem 6

Let z = ζ *be a zero of the matrix polynomial* $\boldsymbol{A}$(z), μ *and* φ *be the determinantal and the adjoint multiplicity of the root* z = ζ, *respectively. Then the* *inverse matrix function $A^{-1}(z)$ exists in a deleted neighbourhood of $z = \zeta$ which represents a pole of order $\mu - \varphi$ of the given function.*

Proof

Insofar as $z = \zeta$ is an isolated zero of the matrix polynomial $\boldsymbol{A}(z)$, $\boldsymbol{A}^{-1}(z)$ is a meaningful expression in a deleted neighbourhood of $z = \zeta$, which can be expressed as

$$\boldsymbol{A}^{-1}(z) = \frac{1}{det\boldsymbol{A}(z)}\boldsymbol{A}^{+}(z) \tag{1.291}$$

Resorting to (1.263) and (1.271) above, the right-hand side can be worked out as follows

$$\begin{aligned}\frac{1}{det\boldsymbol{A}(z)}\boldsymbol{A}^{+}(z) &= \frac{1}{(-1)^{\mu}(\zeta - z)^{\mu}\theta(z)}(-1)^{\varphi}(\zeta - z)^{\varphi}\boldsymbol{\Gamma}(z) \\ &= (-1)^{\varphi-\mu}(\zeta - z)^{\varphi-\mu}\frac{1}{\theta(z)}\boldsymbol{\Gamma}(z)\end{aligned} \tag{1.292}$$

Now, observe that taking the limit of $(\zeta - z)^{\mu-\varphi}\ \boldsymbol{A}^{-1}(z)$ as z tends to ζ, yields

$$\lim_{z\to\zeta}(\zeta - z)^{\mu-\varphi}\boldsymbol{A}^{-1}(z) = \frac{(-1)^{\varphi-\mu}}{\theta(\zeta)}\boldsymbol{\Gamma}(\zeta) = \frac{\mu!}{\varphi!}\frac{(-1)^{\varphi-\mu}}{\pi_{\mu}}\boldsymbol{A}^{+(\varphi)}(\zeta) \neq \boldsymbol{0} \tag{1.293}$$

in light of (1.266) and (1.274).

Thus, according to Definition 5 the matrix function $\boldsymbol{A}^{-1}(z)$ turns out to exhibit a pole of order $\mu - \varphi$ about $z = \zeta$ (in this connection see also Franchi 2007).

□

The next remark widens the content of Remark 2.

Remark 3

The following inequalities establish some interesting relationships among nullity, determinantal and adjoint multiplicities, order of a pole and of a zero.

Nullity of $\boldsymbol{A}(\zeta) \geq$ (*determinantal-multiplicity of the root* ζ)

− (*adjoint-multiplicitiy of the root* ζ)

$$= \textit{order of the pole } \zeta \textit{ of } \boldsymbol{A}^{-1}(z)$$

$$= \textit{order of the zero } \zeta \textit{ of } \boldsymbol{A}(z)$$

Hitherto the discussion of the notion of zero, root and pole was carried out regarding a point ζ of the z plane.

In the wake of the foregoing analysis, we will conclude this section by establishing a key result on the behaviour of the matrix function $\boldsymbol{A}^{-1}(z)$ about a unit root of the parent matrix polynomial $\boldsymbol{A}(z)$, when $\boldsymbol{A}(z)$ is linear.

Theorem 7

Consider the linear matrix polynomial

$$\boldsymbol{A}(z) = \boldsymbol{I} + \boldsymbol{A}_1 z \tag{1.294}$$

and let $\boldsymbol{A} = \boldsymbol{A}(1) = \boldsymbol{I} + \boldsymbol{A}_1$ *be a singular matrix of index* $\upsilon \geq 1$.

Then the matrix function $\boldsymbol{A}^{-1}(z)$ *exhibits a pole of order* υ *at* $z = 1$.

Proof

Simple computations, by making use of the decomposition (1.180) of Theorem 1 in Sect. 1.5, lead to write $\boldsymbol{A}(z)$ as

$$\boldsymbol{A}(z) = (1 - z)\boldsymbol{I} + z(\boldsymbol{H} + \boldsymbol{K}) \tag{1.295}$$

where $\boldsymbol{H}$ and $\boldsymbol{K}$ are defined as in the said theorem.

The right-hand side of (1.295) can be factorized, throughout the z-plane except for the points $z = 1$ and $z = 0$, as

$$\frac{1}{\lambda + 1}(\boldsymbol{I} + \frac{1}{\lambda}\boldsymbol{H})(\lambda \boldsymbol{I} + \boldsymbol{K}) \tag{1.296}$$

in terms of the auxiliary variable $\lambda = \dfrac{1 - z}{z}$ which tends to zero as $z = 1$.

Accordingly, the inverse of the matrix polynomial $\boldsymbol{A}(z)$ can be expressed as

$$(\lambda + 1)\,(\lambda \boldsymbol{I} + \boldsymbol{K})^{-1}\,(\boldsymbol{I} + \frac{1}{\lambda}\boldsymbol{H})^{-1} \tag{1.297}$$

in a deleted neighbourhood of $\lambda = 0$ ($z = 1$).

Now, note that

(a) $$\lim_{\lambda \to 0} \lambda\,(\lambda \boldsymbol{I} + \boldsymbol{K})^{-1} = \boldsymbol{I} - \boldsymbol{K}^{\neq}\boldsymbol{K} \tag{1.298}$$

since $\boldsymbol{K}$ is of index 1 and Corollary 7.6.4, in Campbell and Meyer applies accordingly (see also (1.139) of Sect. 1.4)

(b) $$\lim_{\lambda\to 0} \lambda^{\upsilon-1}(\boldsymbol{I}+\frac{1}{\lambda}\boldsymbol{H})^{-1} = (-1)^{\upsilon-1}\boldsymbol{H}^{\upsilon-1} \neq 0 \qquad (1.299)$$

as $\boldsymbol{H}$ is nilpotent of index υ and the expansion

$$(\boldsymbol{I}+\frac{1}{\lambda}\boldsymbol{H})^{-1} = \boldsymbol{I}-\frac{1}{\lambda}\boldsymbol{H}+\ldots+(-1)^{\upsilon-1}\frac{1}{\lambda^{\upsilon-1}}\boldsymbol{H}^{\upsilon-1} \qquad (1.300)$$

holds accordingly.

(c) $$\lim_{\lambda\to 0}\left\{(1+\lambda)\lambda^{\upsilon}(\lambda\boldsymbol{I}+\boldsymbol{K})^{-1}(\boldsymbol{I}+\frac{1}{\lambda}\boldsymbol{H})^{-1}\right\} = (-1)^{\upsilon-1}(\boldsymbol{I}-\boldsymbol{K}^{\neq}\boldsymbol{K})\boldsymbol{H}^{\upsilon-1}$$
$$= \begin{cases}(\boldsymbol{I}-\boldsymbol{K}^{\neq}\boldsymbol{K}) & under\ \upsilon=1\\ (-1)^{\upsilon-1}\boldsymbol{H}^{\upsilon-1} & under\ \upsilon>1\end{cases} \qquad (1.301)$$

as a by-product of (1.298) and (1.299), bearing in mind that $\boldsymbol{K}$ and $\boldsymbol{H}$ are orthogonal, and adopting the convention that $\boldsymbol{H}^0=\boldsymbol{I}$.
Then, the following result

$$\lim_{z\to 1}\left\{(1-z)^{\upsilon}\boldsymbol{A}^{-1}\right\}$$
$$= \lim_{\lambda\to 0}\left\{(1+\lambda)\lambda^{\upsilon}(\lambda\boldsymbol{I}+\boldsymbol{K})^{-1}(\boldsymbol{I}+\frac{1}{\lambda}\boldsymbol{H})^{-1}\right\} = \begin{cases}(\boldsymbol{I}-\boldsymbol{K}^{\neq}\boldsymbol{K}) & under\ \upsilon=1\\ (-1)^{\upsilon-1}\boldsymbol{H}^{\upsilon-1} & under\ \upsilon>1\end{cases} \qquad (1.302)$$

holds true (as a by-product), leading to the conclusion that $z = 1$ plays the r ô le of a pole of order υ with respect to $\boldsymbol{A}^{-1}(z)$ according to Definition 5.

□

This final remark enriches – for linear polynomials – the picture outlined in Remarks 2 and 3.

Remark 4

The following inequalities specify the relationships between nullity, determinantal and adjoint-multiplicities of a unit root, and index of a matrix

$$\begin{aligned} nullity\ of\ \boldsymbol{A} \geq\ & (determinantal\text{-}multiplicity\ of\ the\ unit\ root)\\ & -(adjoint\text{-}multiplicitiy\ of\ the\ unit\ root)\\ & = index\ of\ \boldsymbol{A}\end{aligned}$$

1.7 The Laurent Form of a Matrix-Polynomial Inverse about a Unit-Root

In this section the reader will find the essentials of matrix-polynomial inversion about a pole, a topic whose technicalities will extend over the forthcoming sections, to cover duly analytical demands of dynamic-model econometrics in Chap. 3.

We begin by enunciating the following basic result

Let

$$\boldsymbol{A}(z) = \sum_{k=0}^{K} \boldsymbol{A}_k \, z^k, \qquad \boldsymbol{A}_K \neq \boldsymbol{0}$$

be a matrix polynomial whose characteristic polynomial

$$\pi(z) = det\boldsymbol{A}(z) \tag{1.303}$$

has all roots either lying outside the unit circle or being equal to one. Then the inverse of the matrix polynomial $\boldsymbol{A}(z)$ can be expressed by a Laurent expansion

$$\boldsymbol{A}^{-1}(z) = \underbrace{\sum_{j=1}^{H} \frac{1}{(1-z)^j} \boldsymbol{N}_j}_{\text{Principal part}} + \underbrace{\sum_{i=0}^{\infty} \boldsymbol{M}_i z^i}_{\text{Regular part}} \tag{1.304}$$

in a deleted neighbourhood of $z = 1$, where H is a non-negative integer, the coefficient matrices $\boldsymbol{M}_i$ of the regular part exhibit exponentially decreasing entries, and the coefficient matrices $\boldsymbol{N}_j$ of the principal part vanish if $\boldsymbol{A}$ is of full rank. The matrices $\boldsymbol{N}_H$ and $\boldsymbol{N}_1$ are referred to as leading coefficient and residue, respectively. As usual $\boldsymbol{A}$ and $\dot{\boldsymbol{A}}$ will be written for $\boldsymbol{A}(1)$ and $\left(\frac{\partial \boldsymbol{A}(z)}{\partial z}\right)_{z=1}$, respectively.

This statement provides a matrix extension of classical results of Laurent series theory (see, e.g., Jeffrey 1992; Markuscevich 1965). A deeper insight into the subject will be provided in the course of this section.

The forthcoming definition about the order of a pole is the mirror image, within a Laurent expansion framework, of the one already formulated.

Definition 1 – Order of a Pole in a Laurent Expansion

The point z_0 is a pole of order H of the (matrix) function $\boldsymbol{A}^{-1}(z)$ if and only if the principal part of the Laurent expansion of $\boldsymbol{A}^{-1}(z)$ about z_0 contains a finite number of terms forming a polynomial of degree H in $(z_0 - z)^{-1}$, i.e. if and only if $\boldsymbol{A}^{-1}(z)$ admits the Laurent expansion

$$\boldsymbol{A}^{-1}(z) = \sum_{j=1}^{H} \frac{1}{(z_0 - z)^j} \boldsymbol{N}_j + \sum_{i=0}^{\infty} z^i \boldsymbol{M}_i, \;\; \boldsymbol{N}_H \neq \boldsymbol{0} \tag{1.305}$$

in a deleted neighbourhood of z_0.

When $H = 1$ the pole located at z_0 is referred to as a simple pole.

Observe that, if (1.305) holds true, then both the matrix function $(z_0 - z)^H \boldsymbol{A}^{-1}(z)$ and its derivatives have finite limits as z tends to z_0, the former $\boldsymbol{N}_H$ being a non null matrix.

The simplest form of the Laurent expansion (1.304) is

$$\boldsymbol{A}^{-1}(z) = \frac{1}{(1-z)} \boldsymbol{N}_1 + \boldsymbol{M}(z) \tag{1.306}$$

which corresponds to the case of a simple pole at $z = 1$ where

$$\boldsymbol{N}_1 = \lim_{z \to 1} [(1-z)] \boldsymbol{A}^{-1}(z)] \tag{1.307}$$

$$\boldsymbol{M}(1) = -\lim_{z \to 1} \frac{\partial[(1-z)\boldsymbol{A}^{-1}(z)]}{\partial z} \tag{1.308}$$

and $\boldsymbol{M}(z)$ stands for $\sum_{i=0}^{\infty} z^i \boldsymbol{M}_i$.

Theorem 1

The residue $\boldsymbol{N}_1$ is a singular matrix which has the following representation

$$\boldsymbol{N}_1 = \boldsymbol{C}_\perp \boldsymbol{E} \boldsymbol{B}'_\perp \tag{1.309}$$

for some $\boldsymbol{E}$. Here $\boldsymbol{B}$ and $\boldsymbol{C}$ are full column-matrices obtained by a rank factorization of the matrix $\boldsymbol{A}$, that is

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}' \tag{1.310}$$

Proof

Keeping in mind (1.247) of Sect. 1.6, since the equalities

$$\boldsymbol{A}(z)\,\boldsymbol{A}^{-1}(z) = \boldsymbol{I} \Leftrightarrow [(1-z)\,\boldsymbol{Q}\,(z) + \boldsymbol{A}][\frac{1}{(1-z)}\boldsymbol{N}_1 + \boldsymbol{M}(z)] = \boldsymbol{I} \tag{1.311}$$

$$\boldsymbol{A}^{-1}(z)\,\boldsymbol{A}(z) = \boldsymbol{I} \Leftrightarrow [\frac{1}{(1-z)}\boldsymbol{N}_1 + \boldsymbol{M}(z)]\,[(1-z)\,\boldsymbol{Q}\,(z) + \boldsymbol{A}] = \boldsymbol{I} \tag{1.312}$$

hold true in a deleted neighbourhood of $z = 1$, the term containing the negative power of $(1 - z)$ in the left-hand sides of (1.311) and (1.312) must vanish. This occurs as long as $\boldsymbol{N}_1$ satisfies the twin conditions

$$\boldsymbol{A}\boldsymbol{N}_1 = \boldsymbol{0} \tag{1.313}$$

$$\boldsymbol{N}_1\boldsymbol{A} = \boldsymbol{0} \tag{1.314}$$

which, in turn, entails the singularity of $\boldsymbol{N}_1$ (we rule out the case of a null $\boldsymbol{A}$).

According to Lemma 2.3.1 in Rao and Mitra (1971), the equations (1.313) and (1.314) have the common solution

$$\boldsymbol{N}_1 = [\boldsymbol{I} - \boldsymbol{A}^g\boldsymbol{A}]\boldsymbol{\Theta}[\boldsymbol{I} - \boldsymbol{A}\boldsymbol{A}^g] \tag{1.315}$$

where $\boldsymbol{\Theta}$ is arbitrary. Resorting to identities (1.87) and (1.88) of Sect. 1.2, the solution can be given in the form (1.309) where $\boldsymbol{E} = \boldsymbol{C}_\perp^g\boldsymbol{\Theta}(\boldsymbol{B}'_\perp)^g$

□

Corollary 1.1

The residue $\boldsymbol{N}_1$ satisfies the homogeneous equation

$$\boldsymbol{C}'\boldsymbol{N}_1 = \boldsymbol{0} \tag{1.316}$$

Proof

Proof is simple and is omitted.

□

Corollary 1.2

The columns of $\boldsymbol{C}$ belong to the null row space of $\boldsymbol{N}_1$. Should $\boldsymbol{E}$ be a non-singular matrix, then they would form a basis for the same space.

Proof

Proof is simple and is omitted.

□

Corollary 1.3

The matrix function $\boldsymbol{C}'\boldsymbol{A}^{-1}(z)$ is analytic at $z = 1$.

Proof

Premultiplying both sides of (1.306) by $\boldsymbol{C}'$ and bearing in mind (1.316), the term involving the singularity located at $z = 1$ disappears and we eventually get

$$\boldsymbol{C}'\boldsymbol{A}^{-1}(z) = \boldsymbol{C}'\boldsymbol{M}(z) \tag{1.317}$$

which is an analytic function at $z = 1$.

□

Another case of the Laurent expansion (1.304) which turns out to be of prominent interest is

$$\boldsymbol{A}^{-1}(z) = \sum_{j=1}^{2} \frac{1}{(1-z)^j} \boldsymbol{N}_j + \boldsymbol{M}(z) \tag{1.318}$$

corresponding to a second order pole at $z = 1$, where

$$\boldsymbol{N}_2 = \lim_{z \to 1} [(1-z)^2 \boldsymbol{A}^{-1}(z)] \tag{1.319}$$

$$\boldsymbol{N}_1 = -\lim_{z \to 1} \frac{\partial [(1-z)^2 \boldsymbol{A}^{-1}(z)]}{\partial z} \tag{1.320}$$

$$\boldsymbol{M}(1) = \frac{1}{2} \lim_{z \to 1} \frac{\partial^2 [(1-z)^2 \boldsymbol{A}^{-1}(z)]}{\partial z^2} \tag{1.321}$$

In this connection we have the following

Theorem 2

The leading coefficient $\boldsymbol{N}_2$ *is a singular matrix which has the following representation*

$$\boldsymbol{N}_2 = \boldsymbol{C}_\perp \boldsymbol{S}_\perp \boldsymbol{W} \boldsymbol{R}'_\perp \boldsymbol{B}'_\perp \tag{1.322}$$

for some $\boldsymbol{W}$*. Here* $\boldsymbol{B}$ *and* $\boldsymbol{C}$ *are as in Theorem 1, whereas* $\boldsymbol{R}$ *and* $\boldsymbol{S}$ *are full column-rank matrices obtained by a rank factorization of* $\boldsymbol{B}'_\perp \dot{\boldsymbol{A}} \boldsymbol{C}_\perp$*, that is*

$$\boldsymbol{B}'_\perp \dot{\boldsymbol{A}} \boldsymbol{C}_\perp = \boldsymbol{R}\boldsymbol{S}' \tag{1.323}$$

Proof

Keeping in mind (1.257) of Sect. 1.6, as the equalities

$$\begin{gathered}\boldsymbol{A}(z)\,\boldsymbol{A}^{-1}(z) = \boldsymbol{I} \Leftrightarrow \\ \Leftrightarrow [(1-z)^2\,\boldsymbol{\Psi}(z) - (1-z)\,\dot{\boldsymbol{A}} + \boldsymbol{A}][\frac{1}{(1-z)^2}\boldsymbol{N}_2 + \frac{1}{(1-z)}\boldsymbol{N}_1 + \boldsymbol{M}(z)] = \boldsymbol{I}\end{gathered} \tag{1.324}$$

$$\begin{gathered}\boldsymbol{A}^{-1}(z)\,\boldsymbol{A}(z) = \boldsymbol{I} \Leftrightarrow \\ \Leftrightarrow [\frac{1}{(1-z)^2}\boldsymbol{N}_2 + \frac{1}{(1-z)}\boldsymbol{N}_1 + \boldsymbol{M}(z)]\,[(1-z)^2\,\boldsymbol{\Psi}(z) - (1-z)\,\dot{\boldsymbol{A}} + \boldsymbol{A}] = \boldsymbol{I}\end{gathered} \tag{1.325}$$

hold true in a deleted neighbourhood of $z = 1$, the terms containing the negative powers of $(1-z)$ in the left-hand sides of (1.324) and (1.325) must vanish. This occurs provided $\boldsymbol{N}_2$ and $\boldsymbol{N}_1$ satisfy the following set of conditions

$$\boldsymbol{A}\boldsymbol{N}_2 = \boldsymbol{0}\,, \qquad \boldsymbol{N}_2\boldsymbol{A} = \boldsymbol{0} \tag{1.326}$$

$$\dot{\boldsymbol{A}}\boldsymbol{N}_2 = \boldsymbol{A}\boldsymbol{N}_1, \qquad \boldsymbol{N}_2\dot{\boldsymbol{A}} = \boldsymbol{N}_1\boldsymbol{A} \tag{1.327}$$

Equalities (1.326), in turn, imply the singularity of $\boldsymbol{N}_2$.

Using the same arguments as in the proof of (1.309), $\boldsymbol{N}_2$ can be written as

$$\boldsymbol{N}_2 = \boldsymbol{C}_\perp \boldsymbol{\Omega} \boldsymbol{B}'_\perp \tag{1.328}$$

where $\boldsymbol{\Omega}$ is arbitrary.

Substituting (1.328) into (1.327) and pre and post-multiplying by $\boldsymbol{B}'_{\perp}$ and $(\boldsymbol{B}'_{\perp})^{g}$ the former equality, we arrive at

$$\boldsymbol{B}'_{\perp}\dot{\boldsymbol{A}}\boldsymbol{C}_{\perp}\boldsymbol{\Omega} = \boldsymbol{0} \tag{1.329}$$

By inspection of (1.328) and (1.329) the conclusion

$$\boldsymbol{N}_2 \neq \boldsymbol{0} \rightarrow \boldsymbol{\Omega} \neq \boldsymbol{0} \rightarrow det(\boldsymbol{B}'_{\perp}\dot{\boldsymbol{A}}\boldsymbol{C}_{\perp}) = 0 \tag{1.330}$$

is easily drawn.

The result

$$\boldsymbol{\Omega}\boldsymbol{B}'_{\perp}\dot{\boldsymbol{A}}\boldsymbol{C}_{\perp} = \boldsymbol{0} \tag{1.331}$$

which is specular to (1.329) can be derived similarly.

Bearing in mind the rank factorization (1.323), equations (1.331) and (1.329) can be more conveniently restated as

$$\begin{cases} \boldsymbol{S}'\boldsymbol{\Omega} = \boldsymbol{0} \\ \boldsymbol{\Omega}\boldsymbol{R} = \boldsymbol{0} \end{cases} \tag{1.332}$$

and, in view of the usual reasoning, $\boldsymbol{\Omega}$ can be written as

$$\boldsymbol{\Omega} = \boldsymbol{S}_{\perp}\boldsymbol{W}\boldsymbol{R}'_{\perp} \tag{1.333}$$

where $\boldsymbol{W}$ is arbitrary.

Substituting (1.333) into (1.328) eventually leads to (1.322)

□

Remark

Notice that, because of (1.323), (1.327) above, and (1.52) of Sect. 1.2, the following hold

$$\boldsymbol{A}\boldsymbol{A}^{g}\boldsymbol{A}\boldsymbol{N}_1 = \boldsymbol{A}\boldsymbol{N}_1 \rightarrow \boldsymbol{A}\boldsymbol{A}^{g}\dot{\boldsymbol{A}}\boldsymbol{N}_2 = \dot{\boldsymbol{A}}\boldsymbol{N}_2 \tag{1.334}$$

$$\begin{aligned} r(\boldsymbol{B}^{g}\dot{\boldsymbol{A}}\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp}) &= r(\boldsymbol{B}\boldsymbol{B}^{g}\dot{\boldsymbol{A}}\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp}) = r\{[\boldsymbol{I} - (\boldsymbol{B}'_{\perp})^{g}\boldsymbol{B}'_{\perp}]\dot{\boldsymbol{A}}\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp}\} \\ &= r(\dot{\boldsymbol{A}}\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp}) \end{aligned} \tag{1.335}$$

Corollary 2.1

The leading coefficient $\boldsymbol{N}_2$ satisfies the homogeneous equation

$$(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp}\boldsymbol{N}_2 = \boldsymbol{0} \tag{1.336}$$

Proof

The proof is simple and is omitted.

□

Corollary 2.2

The columns of $(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})_{\perp}$ belong to the null row space of $\boldsymbol{N}_2$. Should $\boldsymbol{W}$ be a non-singular matrix, then they would form a basis for the same space.

Proof

The proof is simple and is omitted.

□

Corollary 2.3

The matrix function $(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp}\boldsymbol{A}^{-1}(z)$ exhibits a simple pole at $z = 1$.

Proof

Premultiplying both sides of (1.318) by $(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp}$ and bearing in mind (1.336), the term in $(1 - z)^{-2}$ disappears and we eventually obtain

$$(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp}\boldsymbol{A}^{-1}(z) = \frac{1}{(1-z)}(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp}\boldsymbol{N}_1 + (\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp}\boldsymbol{M}(z) \tag{1.337}$$

which exhibits a simple pole located at $z = 1$.

□

Theorem 3

The residue $\boldsymbol{N}_1$ *has the representation*

$$\boldsymbol{N}_1 = \boldsymbol{A}^g \dot{\boldsymbol{A}} \boldsymbol{N}_2 + \boldsymbol{N}_2 \dot{\boldsymbol{A}} \boldsymbol{A}^g + \boldsymbol{C}_\perp \boldsymbol{T} \boldsymbol{B}'_\perp \tag{1.338}$$

for some $\boldsymbol{T}$.

Proof

According to Theorem 2.33 in Rao and Mitra 1971, the equations (1.327) have the common solution

$$\boldsymbol{N}_1 = \boldsymbol{A}^g \dot{\boldsymbol{A}} \boldsymbol{N}_2 + \boldsymbol{N}_2 \dot{\boldsymbol{A}} \boldsymbol{A}^g - \boldsymbol{A}^g \boldsymbol{A} \boldsymbol{N}_2 \dot{\boldsymbol{A}} \boldsymbol{A}^g + [\boldsymbol{I} - \boldsymbol{A}^g \boldsymbol{A}] \Phi [\boldsymbol{I} - \boldsymbol{A} \boldsymbol{A}^g] \tag{1.339}$$

where Φ is arbitrary. In view of (1.326) and resorting to the identities (1.87) and (1.88) of Sect. 1.2, the solution can be given the form (1.338) where $\boldsymbol{T} = \boldsymbol{C}_\perp^g \Phi (\boldsymbol{B}'_\perp)^g$.

□

Corollary 3.1

The block matrix $[\boldsymbol{N}_1, \ \boldsymbol{N}_2]$ satisfies the homogeneous equation

$$(\boldsymbol{B}^g \boldsymbol{V})'_\perp \boldsymbol{C}' [\boldsymbol{N}_1, \boldsymbol{N}_2] = \boldsymbol{0} \tag{1.340}$$

Here $\boldsymbol{V}$ is a full column-rank matrix obtained by a rank factorization of $\dot{\boldsymbol{A}} \boldsymbol{C}_\perp \boldsymbol{S}_\perp$, that is

$$\dot{\boldsymbol{A}} \boldsymbol{C}_\perp \boldsymbol{S}_\perp = \boldsymbol{V} \boldsymbol{\Lambda}' \tag{1.341}$$

Proof

Bearing in mind (1.322), (1.338), (1.341), along with (1.25) of Sect. 1.1, consider the product

$$\begin{aligned} &\boldsymbol{C}' [\boldsymbol{N}_1, \boldsymbol{N}_2] = \boldsymbol{C}' [\boldsymbol{A}^g \dot{\boldsymbol{A}} \boldsymbol{N}_2 + \boldsymbol{N}_2 \dot{\boldsymbol{A}} \boldsymbol{A}^g + \boldsymbol{C}_\perp \boldsymbol{T} \boldsymbol{B}'_\perp, \boldsymbol{N}_2] \\ &\boldsymbol{C}' [(\boldsymbol{C}')^g \boldsymbol{B}^g \dot{\boldsymbol{A}} \boldsymbol{C}_\perp \boldsymbol{S}_\perp \boldsymbol{W} \boldsymbol{R}'_\perp \boldsymbol{B}'_\perp + \boldsymbol{C}_\perp \boldsymbol{S}_\perp \boldsymbol{W} \boldsymbol{R}'_\perp \boldsymbol{B}'_\perp \dot{\boldsymbol{A}} \boldsymbol{A}^g + \boldsymbol{C}_\perp \boldsymbol{T} \boldsymbol{B}'_\perp, \\ &\boldsymbol{C}_\perp \boldsymbol{S}_\perp \boldsymbol{W} \boldsymbol{R}'_\perp \boldsymbol{B}'_\perp] = [\boldsymbol{B}^g \dot{\boldsymbol{A}} \boldsymbol{C}_\perp \boldsymbol{S}_\perp \boldsymbol{W} \boldsymbol{R}'_\perp \boldsymbol{B}'_\perp, \boldsymbol{0}] = [\boldsymbol{B}^g \boldsymbol{V} \boldsymbol{\Lambda}' \boldsymbol{W} \boldsymbol{R}'_\perp \boldsymbol{B}'_\perp, \boldsymbol{0}] \end{aligned} \tag{1.342}$$

Pre-multiplying by $(\boldsymbol{B}^g \boldsymbol{V})'_\perp$ eventually gives (1.340).

□

Corollary 3.2

The columns of $\boldsymbol{C}(\boldsymbol{B}^g\boldsymbol{V})_\perp$ belong to the null row space of $[\boldsymbol{N}_1, \boldsymbol{N}_2]$. Should $\boldsymbol{W}$ be a non-singular matrix and $\begin{bmatrix} \boldsymbol{\Lambda}'\boldsymbol{W}\boldsymbol{R}'_\perp \\ \boldsymbol{S}'\boldsymbol{T} \end{bmatrix}$ be a full row-rank matrix, then they would form a basis for the same space.

Proof

The fact that the columns of $\boldsymbol{C}(\boldsymbol{B}^g\boldsymbol{V})_\perp$ belong to the null row space of $[\boldsymbol{N}_1, \boldsymbol{N}_2]$ follows from (1.340).

In order to get a basis, the rank of $\boldsymbol{C}(\boldsymbol{B}^g\boldsymbol{V})_\perp$ and the dimension of the null row space of $[\boldsymbol{N}_1, \boldsymbol{N}_2]$ must be the same, that is

$$r(\boldsymbol{C}(\boldsymbol{B}^g\boldsymbol{V})_\perp) = n - r([\boldsymbol{N}_1, \boldsymbol{N}_2]) \tag{1.343}$$

For what concerns $r(\boldsymbol{C}(\boldsymbol{B}^g\boldsymbol{V})_\perp)$, simple computations show that

$$\begin{aligned} r(\boldsymbol{C}(\boldsymbol{B}^g\boldsymbol{V})_\perp) &= r((\boldsymbol{B}^g\boldsymbol{V})_\perp) = r(\boldsymbol{A}) - r(\boldsymbol{B}^g\boldsymbol{V}) = r(\boldsymbol{A}) - r(\boldsymbol{V}) \\ &= r(\boldsymbol{A}) - r(\dot{\boldsymbol{A}}\boldsymbol{C}_\perp\boldsymbol{S}_\perp) \end{aligned} \tag{1.344}$$

As a first step to determine $r[\boldsymbol{N}_1, \boldsymbol{N}_2]$, notice that, if $\boldsymbol{W}$ is non-singular, the following holds

$$\begin{aligned} &(\boldsymbol{I} - \boldsymbol{N}_2\boldsymbol{N}_2^g) = \\ &[\boldsymbol{I} - (\boldsymbol{C}_\perp\boldsymbol{S}_\perp\boldsymbol{W}\boldsymbol{R}'_\perp\boldsymbol{B}'_\perp)(\boldsymbol{R}'_\perp\boldsymbol{B}'_\perp)^g\boldsymbol{W}^{-1}(\boldsymbol{C}_\perp\boldsymbol{S}_\perp)^g] = [\boldsymbol{I} - \boldsymbol{C}_\perp\boldsymbol{S}_\perp(\boldsymbol{C}_\perp\boldsymbol{S}_\perp)^g] \end{aligned} \tag{1.345}$$

in light of (1.25′) of Sect. 1.1.

Hence, resorting to Theorem 19 of Marsaglia and Styan and to (1.87) of Sect. 1.2 yields

$$\begin{aligned} r[\boldsymbol{N}_1, \boldsymbol{N}_2] &= r(\boldsymbol{N}_2) + r[(\boldsymbol{I} - \boldsymbol{N}_2\boldsymbol{N}_2^g)\boldsymbol{N}_1] = r(\boldsymbol{S}_\perp) + r[(\boldsymbol{C}_\perp\boldsymbol{S}_\perp)'_\perp\boldsymbol{N}_1] \\ &= r(\boldsymbol{S}_\perp) + r\left\{\begin{bmatrix} \boldsymbol{C}' \\ \boldsymbol{S}'\boldsymbol{C}_\perp^g \end{bmatrix}(\boldsymbol{A}^g\dot{\boldsymbol{A}}\boldsymbol{N}_2 + \boldsymbol{N}_2\dot{\boldsymbol{A}}\boldsymbol{A}^g + \boldsymbol{C}_\perp\boldsymbol{T}\boldsymbol{B}'_\perp)\right\} \\ &= r(\boldsymbol{S}_\perp) + r\begin{pmatrix} \boldsymbol{B}^g\dot{\boldsymbol{A}}\boldsymbol{N}_2 \\ \boldsymbol{S}'\boldsymbol{T}\boldsymbol{B}'_\perp \end{pmatrix} = r(\boldsymbol{S}_\perp) + r\begin{pmatrix} \boldsymbol{B}^g\boldsymbol{V}\boldsymbol{\Lambda}'\boldsymbol{W}\boldsymbol{R}'_\perp\boldsymbol{B}'_\perp \\ \boldsymbol{S}'\boldsymbol{T}\boldsymbol{B}'_\perp \end{pmatrix} \end{aligned} \tag{1.346}$$

$$= r(\boldsymbol{S}_\perp) + r\left\{\begin{bmatrix} \boldsymbol{B}^g\boldsymbol{V} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix}\begin{bmatrix} \boldsymbol{\Lambda}'\boldsymbol{W}\boldsymbol{R}'_\perp \\ \boldsymbol{S}'\boldsymbol{T} \end{bmatrix}\boldsymbol{B}'_\perp\right\} = r(\boldsymbol{S}_\perp) + r\begin{bmatrix} \boldsymbol{\Lambda}'\boldsymbol{W}\boldsymbol{R}'_\perp \\ \boldsymbol{S}'\boldsymbol{T} \end{bmatrix}$$

This, if $\begin{bmatrix} \boldsymbol{\Lambda}'\boldsymbol{W}\boldsymbol{R}'_\perp \\ \boldsymbol{S}'\boldsymbol{T} \end{bmatrix}$ is of full row-rank, entails that

$$\begin{aligned} r[\boldsymbol{N}_1, \boldsymbol{N}_2] &= r(\boldsymbol{S}_\perp) + r(\boldsymbol{\Lambda}'\boldsymbol{W}\boldsymbol{R}'_\perp) + r(\boldsymbol{S}'\boldsymbol{T}) \\ &= r(\boldsymbol{S}_\perp) + r(\boldsymbol{\Lambda}') + r(\boldsymbol{S}) = r(\boldsymbol{C}_\perp) + r(\dot{\boldsymbol{A}}\boldsymbol{C}_\perp\boldsymbol{S}_\perp) = n - r(\boldsymbol{A}) + r(\dot{\boldsymbol{A}}\boldsymbol{C}_\perp\boldsymbol{S}_\perp) \end{aligned} \tag{1.347}$$

which, in turn entails that the dimension of the null row space of [$\boldsymbol{N}_1$, $\boldsymbol{N}_2$] is equal to $r(\boldsymbol{A}) + r(\dot{\boldsymbol{A}}\boldsymbol{C}_\perp\boldsymbol{S}_\perp)$.

This together with (1.344) leads to conclude that, under the assumptions made about $\boldsymbol{W}$ and $\begin{bmatrix} \boldsymbol{\Lambda}'\boldsymbol{W}\boldsymbol{R}'_\perp \\ \boldsymbol{S}'\boldsymbol{T} \end{bmatrix}$, the columns of $\boldsymbol{C}(\boldsymbol{B}^g\boldsymbol{V})_\perp$ provide a basis for the null row space of [$\boldsymbol{N}_1$, $\boldsymbol{N}_2$] as claimed above.

□

Corollary 3.3

The matrix function $(\boldsymbol{B}^g\boldsymbol{V})'_\perp\boldsymbol{C}'\boldsymbol{A}^{-1}(z)$ is analytic at $z = 1$

Proof

Pre-multiplying both sides of (1.318) by $(\boldsymbol{B}^g\boldsymbol{V})'_\perp\boldsymbol{C}'$ and bearing in mind (1.340), both the terms in $(1-z)^{-2}$ and $(1-z)^{-1}$ involving the singularity located at $z = 1$ disappear, and we eventually get

$$(\boldsymbol{B}^g\boldsymbol{V})'_\perp\boldsymbol{C}'\boldsymbol{A}^{-1}(z) = (\boldsymbol{B}^g\boldsymbol{V})'_\perp\boldsymbol{C}'\boldsymbol{M}(z) \tag{1.348}$$

which is an analytic function at $z = 1$

□

Finally, the next result provides a better understanding of the algebraic rationale of expansion (1.304).

Theorem 4

*Under the assumptions that the roots of the characteristic polynomial, det**A**(z), lie either outside or on the unit circle and, in the latter case, be equal to one, the matrix function $\boldsymbol{A}^{-1}(z)$ admits the expansion*

$$\boldsymbol{A}^{-1}(z) = \sum_{j=1}^{H} \frac{1}{(1-z)^j} \boldsymbol{N}_j + \boldsymbol{M}(z) \tag{1.349}$$

where H *is a non negative integer and*

$$\boldsymbol{M}(z) = \sum_{i=0}^{\infty} z^i \boldsymbol{M}_i \tag{1.350}$$

with the entries of the coefficient matrices $\boldsymbol{M}_i$ decreasing exponentially.

Proof

First of all observe that, on the one hand, the factorization

$$det\boldsymbol{A}(z) = k\,(1-z)^a \, \Pi_j \, (1 - \frac{z}{z_j}) \tag{1.351}$$

holds for $det\,\boldsymbol{A}(z)$, where a is a non negative integer, the $z_j' s$ denote the roots lying outside the unit circle ($\left| z_j \right| > 1$) and k is a suitably chosen scalar. On the other hand, the partial fraction expansion

$$\{det\,\boldsymbol{A}\,(z)\}^{-1} = \sum_{i=1}^{a} \lambda_i \, \frac{1}{(1-z)^i} + \sum_j \mu_j \, \frac{1}{1 - \dfrac{z}{z_j}} \tag{1.352}$$

holds for the reciprocal of $det\boldsymbol{A}(z)$, where the $\lambda_i' s$ and the $\mu_j' s$ are properly chosen coefficients, under the assumption that the roots $z_j' s$ are real and simple for algebraic convenience. Should some roots be complex and/or repeated, the expansion still holds with the addition of rational terms whose numerators are linear in z whereas the denominators are higher order polynomials in z (see, e.g. Jeffrey 1992, p 382). This, apart from algebraic burdening, does not ultimately affect the conclusions stated in the theorem.

As long as $\left| z_j \right| > 1$, a power expansion of the form

$$\left(1-\frac{z}{z_j}\right)^{-1} = \sum_{k=0}^{\infty}(z_j)^{-k} z^k \tag{1.353}$$

holds for $|z| \leq 1$.

This and (1.352) lead to the conclusion that $\{det\boldsymbol{A}(z)\}^{-1}$ can be written in the form

$$\begin{aligned}\{det\,\boldsymbol{A}(z)\}^{-1} &= \sum_{i=1}^{a}\lambda_i \frac{1}{(1-z)^i} + \sum_j \mu_j \sum_{k=0}^{\infty}(z_j)^{-k} z^k \\ &= \sum_{i=1}^{a}\lambda_i \frac{1}{(1-z)^i} + \sum_{k=0}^{\infty}\eta_k z^k\end{aligned} \tag{1.354}$$

where the $\eta_k' s$ are exponentially decreasing weights depending on the $\mu_j' s$ and the $z_j' s$.

Now, provided $\boldsymbol{A}^{-1}(z)$ exists in a deleted neighbourhood of $z = 1$, it can be expressed in the form

$$\boldsymbol{A}^{-1}(z) = \{det\,\boldsymbol{A}(z)\}^{-1}\,\boldsymbol{A}^{+}(z) \tag{1.355}$$

where the adjoint matrix $\boldsymbol{A}^{+}(z)$ can be expanded about $z = 1$ yielding

$$\boldsymbol{A}^{+}(z) = \boldsymbol{A}^{+}(1) - \dot{\boldsymbol{A}}^{+}(1)\,(1-z) + \textit{terms of higher powers of}\,(1-z) \tag{1.356}$$

Substituting the right-hand sides of (1.354) and (1.356) for $\{det\,\boldsymbol{A}(z)\}^{-1}$ and $\boldsymbol{A}^{+}(z)$ respectively, into (1.355), we can eventually express $\boldsymbol{A}^{-1}(z)$ in the form (1.349), where the exponentially decay property of the regular part matrices $\boldsymbol{M}_i$ is a by-product of the aforesaid property of the coefficients $\eta_k' s$.

□

1.8 Matrix Polynomials and Difference Equation Systems

Insofar as the algebra of polynomial functions of the complex variable z and the algebra of polynomial functions of the lag operator L are isomorphic (see, e.g., Dhrymes 1971, p 23), the techniques developed in the previous sections provide an analytical tool-kit paving the way to elegant closed-form solutions to finite difference equation systems which are of prominent interest in econometrics.

Indeed, a non- homogeneous linear system of difference equations with constant coefficients can be conveniently written in operator form as follows

$$\boldsymbol{A}(L)\,\boldsymbol{y}_t = \boldsymbol{g}_t \tag{1.357}$$

where $\boldsymbol{y}_t$ is an n x 1 vector of unknown functions, $\boldsymbol{g}_t$ an n x 1 vector of given real valued functions commonly called forcing functions in mathematical physics (see, e.g., Vladimirov 1984, p 38), L is the lag operator defined by the relations

$$L\boldsymbol{y}_t = \boldsymbol{y}_{t-1}, \quad L^0\boldsymbol{y}_t = \boldsymbol{y}_t, \quad L^K\boldsymbol{y}_t = \boldsymbol{y}_{t-K} \tag{1.358}$$

Here K stands for an arbitrary integer, and $\boldsymbol{A}(L)$ is a matrix polynomial of L, defined as

$$\boldsymbol{A}(z) = \sum_{k=0}^{K} \boldsymbol{A}_k\, L^k \tag{1.359}$$

where $\boldsymbol{A}_0, \boldsymbol{A}_1, ..., \boldsymbol{A}_K$ are square matrices of constant coefficients, with $\boldsymbol{A}_0 = \boldsymbol{I}$ as a normalization rule, and $\boldsymbol{A}_K \neq \boldsymbol{0}$. The number K determines the order of the difference equation (1.357).

Remark 1

The so-called companion-form reparametrization allows to rewrite a K-order difference equation as a first order system, (see, e.g., Banierjee et al., p 142), namely

$$\underset{(\breve{n},1)}{\breve{\boldsymbol{y}}_t} + \underset{(\breve{n},\breve{n})}{\breve{\boldsymbol{A}}_1}\, \underset{(\breve{n},1)}{\breve{\boldsymbol{y}}_{t-1}} = \underset{(\breve{n},1)}{\breve{\boldsymbol{g}}_t} \tag{1.360}$$

by taking

$$\underset{(\breve{n},1)}{\breve{\boldsymbol{y}}_t} = \begin{bmatrix} \boldsymbol{y}_t \\ \boldsymbol{y}_{t-1} \\ \vdots \\ \boldsymbol{y}_{t-K+1} \end{bmatrix} \tag{1.361}$$

$$\underset{(\breve{n},\breve{n})}{\breve{A}_1} = \begin{bmatrix} A_1 & \vdots & A_2 & A_3 & \cdot\cdot & A_K \\ \cdots & & \cdots & \cdots & \cdots & \cdots \\ -I_n & \vdots & 0 & 0 & 0 & 0 \\ 0 & \vdots & -I_n & 0 & 0 & 0 \\ \cdot\cdot & \vdots & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \\ 0 & \vdots & 0 & 0 & -I_n & 0 \end{bmatrix} \tag{1.362}$$

$$\underset{(\breve{n},1)}{\breve{g}_t} = J g_t \tag{1.363}$$

$$\underset{(\breve{n},n)}{J} = \begin{bmatrix} I_n \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{1.364}$$

where the K^2 blocks of $\breve{A}_1$ and the K blocks of J are square matrices of order n, and $\breve{n} = nK$.

The parent K-order difference system can be recovered from its companion form (1.360) by pre-multiplying both sides by the selection matrix J'. This gives

$$J' \breve{y}_t + J' \breve{A}_1 \breve{y}_{t-1} = J' \breve{g}_t \tag{1.365}$$

whence

$$y_t + [A_1, A_2, \dots A_K] \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-K} \end{bmatrix} = J' J g_t \longleftrightarrow \sum_{k=0}^{K} A_k y_{t-k} = g_t \tag{1.366}$$

ensues as a by-product.

The isomorphic nature of the transformation at stake enables to attain the solution of a parent K-order difference system as a by-product of the solution of the engendered companion-form first order system.

By replacing g_t with 0, we obtain the homogeneous equation corresponding to (1.357), otherwise known as reduced equation.

Any solution of the non-homogeneous equation (1.357) will be referred to as a particular solution, whereas the general solution of the reduced equation will be referred to as the complementary solution. The latter turns out to depend on the roots z_j of the characteristic equation

$$det\, \boldsymbol{A}(z) = \boldsymbol{0} \tag{1.367}$$

via the solutions $\boldsymbol{h}_j$ of the generalized eigenvector problem

$$\boldsymbol{A}(z_j)\, \boldsymbol{h}_j = \boldsymbol{0} \tag{1.368}$$

Before further investigating the issue of how to handle equation (1.357), some special-purpose analytical tooling is required.

As pointed out in Sect. 1.7, the following Laurent expansions hold for the matrix function $\boldsymbol{A}^{-1}(z)$ in a deleted neighbourhood of $z = 1$

$$\boldsymbol{A}^{-1}(z) = \frac{1}{(1-z)}\boldsymbol{N}_1 + \boldsymbol{M}(z) \tag{1.369}$$

$$\boldsymbol{A}^{-1}(z) = \frac{1}{(1-z)^2}\boldsymbol{N}_2 + \frac{1}{(1-z)}\boldsymbol{N}_1 + \boldsymbol{M}(z) \tag{1.370}$$

under the case of a simple pole and a second order pole, located at $z = 1$, respectively.

Thanks to the aforementioned isomorphism, by replacing 1 by the identity operator $\boldsymbol{I}$ and z by the lag operator L, we obtain the counterparts of the expansions (1.369) and (1.370) in operator form, namely

$$\boldsymbol{A}^{-1}(L) = \frac{1}{(I-L)}\boldsymbol{N}_1 + \boldsymbol{M}(L) \tag{1.371}$$

$$\boldsymbol{A}^{-1}(L) = \frac{1}{(I-L)^2}\boldsymbol{N}_2 + \frac{1}{(I-L)}\boldsymbol{N}_1 + \boldsymbol{M}(L) \tag{1.372}$$

Let us now introduce a few operators related to L which play a crucial role in the study of those difference equations we are primarily interested in. For these and related results see Elaydi (1996) and Mickens (1990).

Definition 1 – Backward Difference Operator

The backward difference operator, denoted by ∇, is defined by the relation

$$\nabla = I - L \tag{1.373}$$

Higher order operators, ∇^K, are defined as follows

$$\nabla^K = (I - L)^K, \;\; K = 2, 3 \ldots \tag{1.374}$$

whereas $\nabla^0 = I$.

Definition 2 – Antidifference Operator

The antidifference operator, denoted by ∇^{-1}, – otherwise known as *indefinite sum operator* and written as Σ – is defined as the operator, such that the identity

$$(I-L)\,\nabla^{-1}\boldsymbol{x}_t = \boldsymbol{x}_t \tag{1.375}$$

holds true for arbitrary $\boldsymbol{x}_t$. In other words ∇^{-1} acts as a right inverse of $I-L$.

Higher order operators, ∇^{-K}, are defined accordingly, i.e.

$$(I-L)^K\,\nabla^{-K} = I \tag{1.376}$$

In light of the identities (1.375) and (1.376), insofar as a K-order difference operator annihilates a $(K-1)$-degree polynomial, the following hold true

$$\nabla^{-1}\boldsymbol{0} = \boldsymbol{c} \tag{1.377}$$

$$\nabla^{-2}\boldsymbol{0} = \boldsymbol{c}t + \boldsymbol{d} \tag{1.378}$$

where $\boldsymbol{c}$ and $\boldsymbol{d}$ are arbitrary.

We now state without proof the well-known result of

Theorem 1

The general solution of the non-homogeneous equation (1.357) *consists of the sum of any particular solution of the given equation and of the complementary solution.*

□

Remark 2

Should we solve a companion-form first order system, the solution of the parent K-order system would be recovered by algebraic methods, because of the considerations put forward in Remark 1.

Thanks to the previous arguments, we are able to establish the following elegant results.

Theorem 2

A particular solution of the non-homogeneous equation (1.357) *can be expressed in operator form as*

$$\overline{\boldsymbol{y}}_t = \boldsymbol{A}^{-1}(L)\,\boldsymbol{g}_t \tag{1.379}$$

In particular, the following hold true

(1) $$\overline{\boldsymbol{y}}_t = \boldsymbol{N}_1 \nabla^{-1}\boldsymbol{g}_t + \boldsymbol{M}(L)\boldsymbol{g}_t = \boldsymbol{N}_1 \sum_{\tau \le t} \boldsymbol{g}_\tau + \boldsymbol{M}(L)\,\boldsymbol{g}_t \tag{1.380}$$

when $z = 1$ *is a simple pole of* $\boldsymbol{A}^{-1}(z)$

(2) $$\overline{\boldsymbol{y}}_t = \boldsymbol{N}_2 \nabla^{-2}\boldsymbol{g}_t + \boldsymbol{N}_1 \nabla^{-1}\boldsymbol{g}_t + \boldsymbol{M}(L)\boldsymbol{g}_t = \boldsymbol{N}_2 \sum_{\vartheta \le t}\sum_{\tau \le \vartheta} \boldsymbol{g}_\tau + \boldsymbol{N}_1 \sum_{\tau \le t} \boldsymbol{g}_\tau + \boldsymbol{M}(L)\,\boldsymbol{g}_t$$
$$= \boldsymbol{N}_2 \sum_{\tau \le t} (t+1-\tau)\boldsymbol{g}_\tau + \boldsymbol{N}_1 \sum_{\tau \le t} \boldsymbol{g}_\tau + \boldsymbol{M}(L)\,\boldsymbol{g}_t \tag{1.381}$$

when $z = 1$ *is a second order pole of* $\boldsymbol{A}^{-1}(z)$.

Proof

Clearly, the right-hand side of (1.379) is a solution, provided $\boldsymbol{A}^{-1}(L)$ is a meaningful operator. Indeed, this is the case for $\boldsymbol{A}^{-1}(L)$ as defined in (1.371) and in (1.372) for a simple and a second order pole at $z = 1$, respectively.

To prove the second part of the theorem, observe first that in view of Definitions 1 and 2, the following operator identities hold

$$\frac{1}{1-L} = \nabla^{-1} = \sum \rightarrow \frac{1}{1-L}\boldsymbol{x}_t = \sum_{\tau \le t} \boldsymbol{x}_\tau \tag{1.382}$$

$$\frac{1}{(1-L)^2} = \nabla^{-2} = \sum\sum \rightarrow \frac{1}{(1-L)^2}\boldsymbol{x}_t = \sum_{\vartheta \le t}\sum_{\tau \le \vartheta} \boldsymbol{x}_\tau \tag{1.383}$$

where $\boldsymbol{x}_t$ is arbitrary. Further, simple sum-calculus rules show that

$$\sum_{\vartheta \le t}\sum_{\tau \le \vartheta} \boldsymbol{x}_\tau = \sum_{\tau \le t} (t+1-\tau)\boldsymbol{x}_\tau = \sum_{\tau \ge 0} (t+1)\boldsymbol{x}_{t-\tau} \tag{1.383'}$$

Thus, in view of expansions (1.369) and (1.370) and the foregoing identities, statements (1) and (2) are easily established.

□

Remark 3

If $\boldsymbol{x}_t$ is a vector of constants, say $\boldsymbol{\eta}$, then as particular cases of (1.382) and (1.383') one obtains

$$\nabla^{-1}\boldsymbol{\eta} = \boldsymbol{\eta}t \quad , \nabla^{-2}\boldsymbol{\eta} = \boldsymbol{\eta}\frac{1}{2}(t+1)t \tag{1.384}$$

Theorem 2 bis

Consider the companion-form first order system

$$\breve{\boldsymbol{A}}(L)\breve{\boldsymbol{y}}_t = \breve{\boldsymbol{g}}_t \tag{1.385}$$

where $\breve{\boldsymbol{A}}(L) = \boldsymbol{I} + \breve{\boldsymbol{A}}_1 L$, *and* $\breve{\boldsymbol{y}}_t, \breve{\boldsymbol{A}}_1, \breve{\boldsymbol{g}}_t$ *are defined as in (1.361)–(1.363) above.*

A particular solution of the non-homogeneous system at stake is given by

(1)
$$\bar{\breve{\boldsymbol{y}}}_t = \breve{\boldsymbol{N}}_1 \sum_{\tau \le t} \breve{\boldsymbol{g}}_\tau + \breve{\boldsymbol{M}}(L)\breve{\boldsymbol{g}}_t \tag{1.386}$$

when z = *1 is a simple pole of* $\breve{\boldsymbol{A}}^{-1}(z)$

(2)
$$\bar{\breve{\boldsymbol{y}}}_t = \breve{\boldsymbol{N}}_2 \sum_{\tau \le t} (t+1-\tau)\breve{\boldsymbol{g}}_\tau + \breve{\boldsymbol{N}}_1 \sum_{\tau \le t} \breve{\boldsymbol{g}}_\tau + \breve{\boldsymbol{M}}(L)\breve{\boldsymbol{g}}_t \tag{1.387}$$

when z = *1 is a second order pole of* $\breve{\boldsymbol{A}}^{-1}(z)$

Here the $\breve{\boldsymbol{N}}_j's$ *and* $\breve{\boldsymbol{M}}$ *play the same role as the* $\boldsymbol{N}_j$ *and* $\boldsymbol{M}$ *in Theorem* 2. *As far as the subvector* $\boldsymbol{y}_t$ *is concerned, the corresponding solution is given by*

(a)
$$\bar{\boldsymbol{y}}_t = \boldsymbol{J}'\breve{\boldsymbol{N}}_1\boldsymbol{J} \sum_{\tau \le t} \boldsymbol{g}_\tau + \boldsymbol{J}'\breve{\boldsymbol{M}}(L)\boldsymbol{J}\boldsymbol{g}_t \tag{1.388}$$

when z = *1 is a simple pole of* $\breve{\boldsymbol{A}}^{-1}(z)$

(b)
$$\bar{\boldsymbol{y}}_t = \boldsymbol{J}'\breve{\boldsymbol{N}}_2\boldsymbol{J} \sum_{\tau \le t} (t+1-\tau)\boldsymbol{g}_\tau + \boldsymbol{J}'\breve{\boldsymbol{N}}_1\boldsymbol{J} \sum_{\tau \le t} \boldsymbol{g}_\tau + \boldsymbol{J}'\breve{\boldsymbol{M}}(L)\boldsymbol{J}\boldsymbol{g}_t \tag{1.389}$$

when z = *1 is a second order pole of* $\breve{\boldsymbol{A}}^{-1}(z)$.

Proof

The proofs of (1.386) and (1.387) can be carried out by repeating step by step that of Theorem 2.

The proofs of (1.388) and (1.389) ensue from the arguments pointed out in Remark 1.

□

Theorem 3

The solution of the reduced equation

$$\boldsymbol{A}(L)\,\overline{\boldsymbol{\xi}}_t = \boldsymbol{0} \tag{1.390}$$

corresponding to the unit root z *= 1 can be written in operator form as*

$$\overline{\boldsymbol{\xi}}_t = \boldsymbol{A}^{-1}(L)\,\boldsymbol{0} \tag{1.391}$$

where the operator $\boldsymbol{A}^{-1}(L)$ *is defined as in* (1.371) *or in* (1.372)*, depending upon the order (first or second, respectively) of the pole of* $\boldsymbol{A}^{-1}(z)$ *at* z *= 1.*

The following closed-form expressions of the solution hold

$$\overline{\boldsymbol{\xi}}_t = \boldsymbol{N}_1\boldsymbol{c} \tag{1.392}$$

$$\overline{\boldsymbol{\xi}}_t = \boldsymbol{N}_2\boldsymbol{c}t + \boldsymbol{N}_2\boldsymbol{d} + \boldsymbol{N}_1\boldsymbol{c} \tag{1.393}$$

for a first and a second order pole respectively, with ***c*** *and* ***d*** *arbitrary vectors.*

Proof

The proof follows from arguments similar to those of Theorem 2 by making use of results (1.377) and (1.378) above.

□

Theorem 3 bis

Consider the companion-form first order system given by formula (1.385) *of Theorem* 2 *bis. The solution of the reduced equation*

$$\breve{\boldsymbol{A}}(L)\overline{\overline{\boldsymbol{\xi}}}_t = \boldsymbol{0} \tag{1.394}$$

corresponding to the unit root $z = 1$ *is given by*

$$\overline{\overline{\boldsymbol{\xi}}}_t = \breve{\boldsymbol{N}}_1\breve{\boldsymbol{c}} \tag{1.395}$$

$$\overline{\overline{\boldsymbol{\xi}}}_t = \breve{\boldsymbol{N}}_2\breve{\boldsymbol{c}}t + \breve{\boldsymbol{N}}_2\breve{\boldsymbol{d}} + \breve{\boldsymbol{N}}_1\breve{\boldsymbol{c}} \tag{1.396}$$

for a first and a second order pole respectively, with $\breve{\boldsymbol{c}}$, $\breve{\boldsymbol{d}}$ *arbitrary vectors and where the* $\breve{\boldsymbol{N}}_j$'s, $\breve{\boldsymbol{c}}$ *and* $\breve{\boldsymbol{d}}$ *play the same role as the* $\boldsymbol{N}_j$'s, $\boldsymbol{c}$ *and* $\boldsymbol{d}$ *in Theorem 3.*

As far as the subvector $\bar{\boldsymbol{\xi}}_t$ *is concerned, the corresponding solution is given by*

$$\bar{\boldsymbol{\xi}}_t = \boldsymbol{J}'\breve{\boldsymbol{N}}_1\breve{\boldsymbol{c}} \tag{1.397}$$

$$\bar{\boldsymbol{\xi}}_t = \boldsymbol{J}'\breve{\boldsymbol{N}}_2\breve{\boldsymbol{c}}t + \boldsymbol{J}'\breve{\boldsymbol{N}}_2\breve{\boldsymbol{d}} + \boldsymbol{J}'\breve{\boldsymbol{N}}_1\breve{\boldsymbol{c}} \tag{1.398}$$

for a first and a second order pole respectively.

Proof

The proofs of (1.395) and (1.396) are made by a verbatim repetition of the proof of Theorem 3.

The proofs of (1.397) and (1.398) ensue from the arguments advanced in Remark 1.

□

Theorem 4

The solution of the reduced equation

$$\boldsymbol{A}(L)\ \bar{\boldsymbol{\xi}}_t = \boldsymbol{0} \tag{1.399}$$

corresponding to unit roots is a polynomial of the same degree as the order, reduced by one, of the pole of $\boldsymbol{A}^{-1}(z)$ *at* $z = 1$.

Proof

Should $z = 1$ be either a simple or a second order pole of $\boldsymbol{A}^{-1}(z)$, then Theorem 3 would apply. The proof for a higher order pole follows along the same lines.

□

Two additional results, whose role will become clear later on, are worth stating.

Theorem 5

Let $z = 1$ *be a simple pole of* $\boldsymbol{A}^{-1}(z)$ *and*

$$\boldsymbol{N}_1 = \boldsymbol{F}\boldsymbol{G}' \tag{1.400}$$

be a rank factorization of $\boldsymbol{N}_1$*. Then the following hold*

$$\boldsymbol{F}_{\perp}' \, \overline{\boldsymbol{y}}_t = \boldsymbol{F}_{\perp}' \boldsymbol{M}(L)\, \boldsymbol{g}_t \tag{1.401}$$

$$\boldsymbol{F}_{\perp}' \, \overline{\boldsymbol{\xi}}_t = \boldsymbol{0} \tag{1.402}$$

where $\overline{\boldsymbol{y}}_t$ *and* $\overline{\boldsymbol{\xi}}_t$ *are determined as in* (1.380) *and* (1.392) *of Theorems* 2 *and* 3, *respectively.*

Proof

The proof is simple and is omitted.

□

Theorem 5 bis

Consider the companion-form first order system given by formula (1.385) *of Theorem* 2 *bis*. *Let* $z = 1$ *be a simple pole of* $\breve{\boldsymbol{A}}^{-1}(z)$ *and*

$$\breve{\boldsymbol{N}}_1 = \breve{\boldsymbol{F}}\breve{\boldsymbol{G}}' \tag{1.403}$$

be a rank factorization of $\breve{\boldsymbol{N}}_1$. *Then the following hold*

$$\breve{\boldsymbol{F}}_{\perp}' \overline{\breve{\boldsymbol{y}}}_t = \breve{\boldsymbol{F}}_{\perp}' \breve{\boldsymbol{M}}(L) \breve{\boldsymbol{g}}_t \tag{1.404}$$

$$\breve{\boldsymbol{F}}_{\perp}' \overline{\breve{\boldsymbol{\xi}}}_t = \boldsymbol{0} \tag{1.405}$$

where $\overline{\breve{\boldsymbol{y}}}_t$ *and* $\overline{\breve{\boldsymbol{\xi}}}_t$ *are determined as in* (1.386) *and* (1.395) *of Theorems* 2*bis and* 3*bis*, *respectively.*

As far as the subvectors $\overline{\boldsymbol{y}}_t$ *and* $\overline{\boldsymbol{\xi}}_t$ *are concerned, the corresponding solutions are given by*

$$(\boldsymbol{J}'\breve{\boldsymbol{F}})_{\perp}' \boldsymbol{J}\overline{\breve{\boldsymbol{y}}}_t = (\boldsymbol{J}'\breve{\boldsymbol{F}})_{\perp}' \boldsymbol{J}'\breve{\boldsymbol{M}}(L)\boldsymbol{J}\boldsymbol{g}_t \tag{1.406}$$

$$(\boldsymbol{J}'\breve{\boldsymbol{F}})_{\perp}' \boldsymbol{J}'\overline{\breve{\boldsymbol{\xi}}}_t = \boldsymbol{0} \tag{1.407}$$

Proof

The proofs of (1.404),(1.405) are made by a verbatim repetition of the proof of Theorem 5; the proofs of (1.406), (1.407) ensue from the arguments advanced in Remark 1.

□

Theorem 6

Let z = *1 be a second order pole of* $\boldsymbol{A}^{-1}(z)$ *and*

$$\boldsymbol{N}_2 = \boldsymbol{H}\boldsymbol{K}' \tag{1.408}$$

a rank factorization of $\boldsymbol{N}_2$. *Then the following hold*

$$\boldsymbol{H}'_{\perp}\,\overline{\boldsymbol{y}}_t = \boldsymbol{H}'_{\perp}\boldsymbol{N}_1 \sum_{\tau \le t} \boldsymbol{g}_\tau + \boldsymbol{H}'_{\perp}\boldsymbol{M}(L)\,\boldsymbol{g}_t \tag{1.409}$$

$$\boldsymbol{H}'_{\perp}\,\overline{\boldsymbol{\xi}}_t = \boldsymbol{H}'_{\perp}\boldsymbol{N}_1\boldsymbol{c} \tag{1.410}$$

where $\overline{\boldsymbol{y}}_t$ *and* $\overline{\boldsymbol{\xi}}_t$ *are determined as in* (1.381) *and* (1.393) *of Theorems* 2 *and* 3, *respectively.*

Besides, should $[\boldsymbol{N}_2, \boldsymbol{N}_1]$ *not be of full row-rank and*

$$[\boldsymbol{N}_2, \boldsymbol{N}_1] = \boldsymbol{G}\boldsymbol{L}' \tag{1.411}$$

represent a rank factorization of the same, then the following would hold

$$\boldsymbol{G}'_{\perp}\,\overline{\boldsymbol{y}}_t = \boldsymbol{G}'_{\perp}\boldsymbol{M}(L)\boldsymbol{g}_t \tag{1.412}$$

$$\boldsymbol{G}'_{\perp}\,\overline{\boldsymbol{\xi}}_t = 0 \tag{1.413}$$

where $\overline{\boldsymbol{y}}_t$ *and* $\overline{\boldsymbol{\xi}}_t$ *are as above.*

Proof

The proof is simple and is omitted.

□

Theorem 6 bis

Consider the companion-form first order system given by formula (1.385) *of Theorem* 2 *bis. Let* z = 1 *be a second order pole of* $\breve{\boldsymbol{A}}^{-1}(z)$ *and*

$$\breve{N}_2 = \breve{H}\breve{K}' \tag{1.414}$$

be a rank factorization of $\breve{N}_2$. *Then the following hold*

$$\breve{H}'_{\perp}\overline{\breve{y}}_t = \breve{H}'_{\perp}\breve{N}_1\sum_{\tau\le t}\boldsymbol{\varepsilon}_\tau + \breve{H}'_{\perp}\breve{M}(L)\breve{g}_t \tag{1.415}$$

$$\breve{H}'_{\perp}\overline{\breve{\xi}}_t = \mathbf{0} \tag{1.416}$$

where $\overline{\breve{y}}_t$ *and* $\overline{\breve{\xi}}_t$ *are determined as in* (1.387) *and* (1.396) *of Theorems* 2*bis and* 3*bis, respectively.*

Besides, should $[\breve{N}_2, \breve{N}_1]$ *not be of full row-rank and*

$$[\breve{N}_2, \breve{N}_1] = \breve{G}\breve{L}' \tag{1.417}$$

represent a rank factorization of the same, then the following would hold

$$\breve{G}'_{\perp}\overline{\breve{y}}_t = \breve{G}'_{\perp}\breve{M}(L)\breve{g}_t \tag{1.418}$$

$$\breve{G}'_{\perp}\overline{\breve{\xi}}_t = \mathbf{0} \tag{1.419}$$

where $\overline{\breve{y}}_t$ *and* $\overline{\breve{\xi}}_t$ *are as above.*

As far as the subvectors $\overline{y}_t$ *and* $\overline{\xi}_t$ *are concerned, the corresponding solutions can be obtained from* (1.415) *and* (1.416) *in this way*

$$(J'\breve{H})'_{\perp}J'\,\overline{y}_t = (J'\breve{H})'_{\perp}J'\,\breve{N}_1\sum_{\tau\le t}Jg_\tau + (J'\breve{H})'_{\perp}J'\breve{M}(L)Jg_t \tag{1.420}$$

$$(J'\breve{H})'_{\perp}J'\,\overline{\xi}_t = \mathbf{0} \tag{1.421}$$

and from (1.418) *and* (1.419) *as follows*

$$(J'\breve{G})'_{\perp}J'\,\overline{y}_t = (J'\breve{G})'_{\perp}J'\breve{M}(L)Jg_t \tag{1.422}$$

$$(J'\breve{G})'_{\perp}J'\,\overline{\xi}_t = \mathbf{0} \tag{1.423}$$

Proof

The proof is simple and it is omitted

□

1.9 The Linear Matrix Polynomial

This section is devoted to the inversion of a linear matrix polynomial about a unit root. Such a topic, whose algebraic foundation is provided by the decompositions of Sect. 1.5, is worth considering not only in itself, but also as a bridgehead for matrix-polynomial inversion in general, via a companion-form isomorphic transformation as we will see afterwards.

Theorem 1

With $\boldsymbol{A}$, $\boldsymbol{H}$ *and* $\boldsymbol{K}$ *as in Theorem* 1 *of Sect.* 1.5, *consider a linear matrix polynomial specified as*

$$\boldsymbol{A}(z) = (1-z)\boldsymbol{I} + z\,\boldsymbol{A} = (1-z)\boldsymbol{I} + z(\boldsymbol{H}+\boldsymbol{K}) \tag{1.424}$$

and let det$\boldsymbol{A}(z)$ *have all its roots outside the unit circle, except for a possibly multiple unit root.*

Then, the following Laurent expansion of $\boldsymbol{A}^{-1}(z)$,

$$\boldsymbol{A}^{-1}(z) = \sum_{i=1}^{\upsilon}\frac{1}{(1-z)^i}\boldsymbol{N}_i + \sum_{i=0}^{\infty}\widetilde{\boldsymbol{M}}_i z^i \ , \tag{1.425}$$

holds about z = 1, with the $\widetilde{\boldsymbol{M}}_i's$ *decaying at an exponential rate and with* $\boldsymbol{N}_\upsilon$ *and* $\boldsymbol{N}_{\upsilon-1}$ *given by*

$$\boldsymbol{N}_\upsilon = (-1)^{\upsilon-1}\boldsymbol{N}\boldsymbol{H}^{\upsilon-1} \tag{1.426}$$

$$\boldsymbol{N}_{\upsilon-1} = (-1)^{\upsilon-2}\boldsymbol{N}\boldsymbol{H}^{\upsilon-2} + (1-\upsilon)\boldsymbol{N}_\upsilon \tag{1.427}$$

where

$$\boldsymbol{N} = \boldsymbol{C}_{\upsilon\perp}(\boldsymbol{B}'_{\upsilon\perp}\boldsymbol{C}_{\upsilon\perp})^{-1}\boldsymbol{B}'_{\upsilon\perp} \tag{1.428}$$

$\boldsymbol{B}_\upsilon$ *and* $\boldsymbol{C}_\upsilon$ *are full column-rank matrices obtained by rank factorization of* $\boldsymbol{A}^\upsilon$, *and* $\boldsymbol{H}^0 = \boldsymbol{I}$ *by convention.*

Proof

Following an argument mirroring that of the proof of Theorem 7 in Sect. 1.6, let us factorize $\boldsymbol{A}(z)$ – in a deleted neighbourhood of $z = 1$ – as

$$\boldsymbol{A}(z) = (\boldsymbol{I}-\boldsymbol{H})\boldsymbol{H}(z)\boldsymbol{K}(z) \tag{1.429}$$

where $\boldsymbol{H}$ and $\boldsymbol{K}$ denote the nilpotent and core components of $\boldsymbol{A}$ as specified in Theorem 1 of Sect. 1.5, and $\boldsymbol{K}(z)$ and $\boldsymbol{H}(z)$ are specified as follows

$$\boldsymbol{K}(z) = (1-z)(\boldsymbol{I}-\boldsymbol{K}) + \boldsymbol{K} = (1-z)\,\boldsymbol{I} + z\boldsymbol{K} \tag{1.430}$$

$$\boldsymbol{H}(z) = \boldsymbol{I} + \overline{\boldsymbol{H}}\frac{1}{(1-z)} \tag{1.431}$$

with

$$\overline{\boldsymbol{H}} = (\boldsymbol{I}-\boldsymbol{H})^{-1}\boldsymbol{H} \tag{1.432}$$

The inverse of $\boldsymbol{A}(z)$ can be expressed as a product as well, namely

$$\boldsymbol{A}^{-1}(z) = \boldsymbol{K}^{-1}(z)\boldsymbol{H}^{-1}(z)(\boldsymbol{I}-\boldsymbol{H})^{-1} \tag{1.433}$$

provided the inverses in the right hand-side exist.

Now, notice that the non-unit roots of $\boldsymbol{K}(z)$ are the same as those of $\boldsymbol{A}(z)$ whereas, for what concerns the unit root of $\boldsymbol{K}(z)$, reference can be made to Theorem 7 of Sect. 1.6 with $\upsilon = 1$ as $ind(\boldsymbol{K}) = 1$. Therefore, according to Theorem 4 of Sect. 1.7 we can expand $\boldsymbol{K}^{-1}(z)$ about $z = 1$ as

$$\boldsymbol{K}^{-1}(z) = \frac{1}{(1-z)}\boldsymbol{N} + \widetilde{\boldsymbol{M}}(z) \tag{1.434}$$

where $\widetilde{\boldsymbol{M}}(z)$ is written for $\sum_{i=0}^{\infty}\widetilde{\boldsymbol{M}}_i z^i$, the $\widetilde{\boldsymbol{M}}_i's$ decay at an exponential rate, and $\boldsymbol{N}$ is given by

$$\begin{aligned}\boldsymbol{N} = \boldsymbol{I} - (\boldsymbol{K})^{\neq}\boldsymbol{K} = \overline{\boldsymbol{C}}_{\perp}(\overline{\boldsymbol{B}}'_{\perp}\overline{\boldsymbol{C}}_{\perp})^{-1}\overline{\boldsymbol{B}}'_{\perp} = \overline{\boldsymbol{C}}_{\upsilon\perp}(\overline{\boldsymbol{B}}'_{\upsilon\perp}\overline{\boldsymbol{C}}_{\upsilon\perp})^{-1}\overline{\boldsymbol{B}}'_{\upsilon\perp} \\ = \boldsymbol{I} - (\boldsymbol{A}^{\upsilon})^{\neq}\boldsymbol{A}^{\upsilon}\end{aligned} \tag{1.435}$$

where $\overline{\boldsymbol{B}}$ and $\overline{\boldsymbol{C}}$ are defined by the rank factorization $\boldsymbol{K} = \overline{\boldsymbol{B}}\overline{\boldsymbol{C}}'$. Straightforward computation shows that $\boldsymbol{N}$ is an idempotent matrix.

Formula (1.435) is established by resorting to (1.139) of Sect. 1.4 along with (1.187) of Sect. 1.5 and (1.89) of Sect. 1.2, upon noticing that

$$\lim_{z\to 1}(1-z)\boldsymbol{K}^{-1}(z) = \lim_{\lambda\to 0}\lambda(\lambda\boldsymbol{I}+\boldsymbol{K})^{-1} \tag{1.436}$$

where $\boldsymbol{K}(z)$ is specified as in (1.430) and λ stands for $\frac{1-z}{z}$.

Besides, upon noticing that $\boldsymbol{H}$ and $(\boldsymbol{I}-\boldsymbol{H})^{-1}$ commute, the matrix $\overline{\boldsymbol{H}}$ enjoys the same nilpotency property as $\boldsymbol{H}$, and we can accordingly expand $\boldsymbol{H}^{-1}(z)$ about $z = 1$ as follows

$$\boldsymbol{H}^{-1}(z) = \boldsymbol{I} + \sum_{i=1}^{\upsilon-1} \frac{(-1)^i}{(1-z)^i} \overline{\boldsymbol{H}}^i \tag{1.437}$$

Then, replacing (1.434) and (1.437) into (1.433) eventually gives (1.425).

Seeking for the expressions of $\boldsymbol{N}_\upsilon$ and $\boldsymbol{N}_{\upsilon-1}$, substitute first the right-hand sides of (1.434) and (1.437) into the right-hand side of (1.433), that is

$$\begin{aligned} \boldsymbol{A}^{-1}(z) &= \left(\frac{1}{(1-z)} \boldsymbol{N} + \tilde{\boldsymbol{M}}(z) \right) \left(\boldsymbol{I} + \sum_{i=1}^{\upsilon-1} \frac{(-1)^i}{(1-z)^i} \overline{\boldsymbol{H}}^i \right) (\boldsymbol{I}-\boldsymbol{H})^{-1} \\ &= \frac{(-1)^{\upsilon-1}}{(1-z)^{\upsilon}} \boldsymbol{N}\boldsymbol{H}^{\upsilon-1}(\boldsymbol{I}-\boldsymbol{H})^{-\upsilon} + \\ &+ \frac{1}{(1-z)^{\upsilon-1}} \{(-1)^{\upsilon-2} \boldsymbol{N}\boldsymbol{H}^{\upsilon-2}(\boldsymbol{I}-\boldsymbol{H})^{-\upsilon+1} \\ &+ (-1)^{\upsilon-1} \tilde{\boldsymbol{M}}(z) \boldsymbol{H}^{\upsilon-1}(\boldsymbol{I}-\boldsymbol{H})^{-\upsilon}\} + \dots \end{aligned} \tag{1.438}$$

Now, compare the right-hand sides of (1.438) and (1.425), and conclude accordingly that

$$\boldsymbol{N}_\upsilon = (-1)^{\upsilon-1} \boldsymbol{N}\boldsymbol{H}^{\upsilon-1}(\boldsymbol{I}-\boldsymbol{H})^{-\upsilon} \tag{1.439}$$

$$\boldsymbol{N}_{\upsilon-1} = (-1)^{\upsilon-2} \boldsymbol{N}\boldsymbol{H}^{\upsilon-2}(\boldsymbol{I}-\boldsymbol{H})^{-\upsilon+1} + (-1)^{\upsilon-1} \tilde{\boldsymbol{M}}(1) \boldsymbol{H}^{\upsilon-1}(\boldsymbol{I}-\boldsymbol{H})^{-\upsilon} \tag{1.440}$$

Formulas (1.439) and (1.440) turn out to be equal to (1.426) and (1.427) upon applying the binomial theorem to $(\boldsymbol{I}-\boldsymbol{H})^{-\upsilon}$ and $(\boldsymbol{I}-\boldsymbol{H})^{-\upsilon+1}$, bearing in mind the nilpotency of $\boldsymbol{H}$ and making use of the orthogonality relationship

$$\tilde{\boldsymbol{M}}(1)\boldsymbol{H} = \boldsymbol{0} \tag{1.441}$$

This latter relationship proves true upon resorting to the following arguments. Insofar as the equalities

$$\boldsymbol{K}^{-1}(z)\boldsymbol{K}(z) = \boldsymbol{I} \tag{1.442}$$

$$\boldsymbol{K}(z)\boldsymbol{K}^{-1}(z) = \boldsymbol{I} \tag{1.443}$$

hold true in a deleted neighborhood of $z = 1$, the terms containing the negative power of $(1 - z)$ in the left-hand sides of (1.442) and (1.443) must vanish, which occurs if both

$$\boldsymbol{NK} = \boldsymbol{0} \tag{1.444}$$

and

$$\boldsymbol{KN} = \boldsymbol{0} \tag{1.445}$$

When (1.444) applies, bearing in mind (1.430) and (1.434), equality (1.442) becomes

$$\boldsymbol{N} + \widetilde{\boldsymbol{M}}(z)\boldsymbol{K}(z) = \boldsymbol{I} \tag{1.446}$$

whence

$$\widetilde{\boldsymbol{M}}(z) = (\boldsymbol{I} - \boldsymbol{N})\boldsymbol{K}^{-1}(z) = (\boldsymbol{I} - \boldsymbol{N})\widetilde{\boldsymbol{M}}(z) \tag{1.447}$$

which, by taking the limit as $z = 1$, entails

$$\boldsymbol{N}\widetilde{\boldsymbol{M}}(1) = \boldsymbol{0} \tag{1.448}$$

Likewise, replacing (1.442) by (1.443), we find that

$$\widetilde{\boldsymbol{M}}(1)\,\boldsymbol{N} = \boldsymbol{0} \tag{1.449}$$

Solving (1.448) and (1.449), by making use of Lemma 2.31 in Rao and Mitra (1971), yields

$$\widetilde{\boldsymbol{M}}(1) = (\boldsymbol{I} - \boldsymbol{N})\boldsymbol{\Theta}(\boldsymbol{I} - \boldsymbol{N}) = \overline{\boldsymbol{B}}(\overline{\boldsymbol{C}}'\overline{\boldsymbol{B}})^{-1}\overline{\boldsymbol{C}}'\boldsymbol{\Theta}\,\overline{\boldsymbol{B}}(\overline{\boldsymbol{C}}'\overline{\boldsymbol{B}})^{-1}\overline{\boldsymbol{C}}' \tag{1.450}$$

for some $\boldsymbol{\Theta}$. This together with the orthogonality property

$$\boldsymbol{KH} = \boldsymbol{0} \rightarrow \overline{\boldsymbol{C}}'\boldsymbol{H} = \boldsymbol{0} \tag{1.451}$$

eventually proves (1.441).

□

Linear matrix polynomials whose inverse have either a simple or a second order pole at $z = 1$, play a crucial role in dynamic modelling. As we will soon realize, new insights on the subject can be gained from an analysis resting on the canonical form of the matrix $\boldsymbol{A}(1) = \boldsymbol{A}$ of Theorem 2 of Sect. 1.5.

Before tackling this issue, let us first make the following

Remark

With

$$\boldsymbol{A}(z) = \boldsymbol{P}\begin{bmatrix} \boldsymbol{I} - (\boldsymbol{I} - \boldsymbol{D})z & \boldsymbol{0} \\ \boldsymbol{0} & (1-z)\boldsymbol{I} + \boldsymbol{\Upsilon} z \end{bmatrix}\boldsymbol{P}^{-1} \tag{1.452}$$

where $\boldsymbol{D}$, $\boldsymbol{P}$, and $\boldsymbol{\Upsilon}$ have the same meaning as in Theorem 2 of Sect. 1.5, we can see that

$$det\boldsymbol{A}(z) = det(\boldsymbol{I} - (\boldsymbol{I} - \boldsymbol{D})\, z)det((1 - z)\boldsymbol{I} + \boldsymbol{\Upsilon} z) \tag{1.453}$$

where (see, e.g., Bellman 1970, p 219)

$$det((1 - z)\boldsymbol{I} + \boldsymbol{\Upsilon} z) = z^{\gamma}\, det\left(\frac{(1-z)}{z}\boldsymbol{I} + \boldsymbol{\Upsilon}\right) = z^{\gamma}\, det\left(\frac{(1-z)}{z}\boldsymbol{I}\right) \tag{1.454}$$

and γ represents the dimension of $\boldsymbol{\Upsilon}$. In light of (1.453) and (1.454), the roots of $det(\boldsymbol{I} - (\boldsymbol{I} - \boldsymbol{D})z)$ correspond to the non-unit roots of $det\boldsymbol{A}(z)$, which lie outside the unit circle by assumption. Their reciprocals, that is the eigenvalues of $(\boldsymbol{I} - \boldsymbol{D})$ will therefore lie inside the unit circle (see also Remark 1, of Sect. 1.5), and the matrix $(\boldsymbol{I} - \boldsymbol{D})$ will be stable.

Theorem 2

Consider a linear matrix polynomial specified as in Theorem 1 and let $\boldsymbol{A}$ *be of index* $\upsilon = 1$. *Then, the inverse of the matrix polynomial* $\boldsymbol{A}(z)$ *has a simple pole about* $z = 1$. *Indeed, the following holds*

$$\begin{aligned} \boldsymbol{A}^{-1}(z) &= \boldsymbol{P}\begin{bmatrix} \boldsymbol{I} + \sum_{j=1}^{\infty} (\boldsymbol{I} - \boldsymbol{D})^j z^j & \boldsymbol{0} \\ \boldsymbol{0} & (1 - z)^{-1}\boldsymbol{I} \end{bmatrix}\boldsymbol{P}^{-1} \\ &= \boldsymbol{P}_1\left(\boldsymbol{I} + \sum_{j=1}^{\infty}(\boldsymbol{I} - \boldsymbol{D})^j z^j\right)\boldsymbol{\Pi}_1 + \frac{1}{(1-z)}\boldsymbol{P}_2\boldsymbol{\Pi}_2 \end{aligned} \tag{1.455}$$

where $\boldsymbol{D}$, $\boldsymbol{P}$, $\boldsymbol{P}_1$, $\boldsymbol{P}_2$, $\boldsymbol{\Pi}_1$, $\boldsymbol{\Pi}_2$ *are defined as in Corollary* 2.1 *and Theorem* 3 *of Sect*. 1.5.

Proof

Under the assumption that $ind(\boldsymbol{A}) = 1$, $\boldsymbol{A} = \boldsymbol{K}$ (see Theorem 1 of Sect. 1.5) and according to Theorem 2 of Sect. 1.5, the following holds

$$\boldsymbol{A} = \boldsymbol{P}\begin{bmatrix} \boldsymbol{D} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}\boldsymbol{P}^{-1} \tag{1.456}$$

This, in turn, entails that

$$\begin{aligned} \boldsymbol{A}(z) &= (1-z)(\boldsymbol{I}-\boldsymbol{A}) + \boldsymbol{A} \\ &= (1-z)\left[\boldsymbol{P}\boldsymbol{P}^{-1} - \boldsymbol{P}\begin{bmatrix} \boldsymbol{D} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}\boldsymbol{P}^{-1}\right] + \boldsymbol{P}\begin{bmatrix} \boldsymbol{D} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}\boldsymbol{P}^{-1} \\ &= \boldsymbol{P}\begin{bmatrix} \boldsymbol{I}-(\boldsymbol{I}-\boldsymbol{D})z & \boldsymbol{0} \\ \boldsymbol{0} & (1-z)\boldsymbol{I} \end{bmatrix}\boldsymbol{\Pi} \end{aligned} \tag{1.457}$$

where $\boldsymbol{\Pi} = \boldsymbol{P}^{-1}$.

Then by inverting the right-hand side of (1.457), and recalling that $(\boldsymbol{I}-\boldsymbol{D})$ is a stable matrix (see the Remark above), we can easily derive (1.455) and conclude that $\boldsymbol{A}^{-1}(z)$ has a simple pole at $z = 1$.

□

Corollary 2.1

Let $\boldsymbol{A}(z)$ be a linear polynomial defined as in Theorem 2. Then, the following hold

(a)
$$\boldsymbol{P}_2\boldsymbol{\Pi}_2 = \boldsymbol{N}_1 = \boldsymbol{C}_\perp(\boldsymbol{B}'_\perp\boldsymbol{C}_\perp)^{-1}\boldsymbol{B}'_\perp = \boldsymbol{I} - \boldsymbol{A}^{\neq}\boldsymbol{A} \tag{1.458}$$

(b)
$$\boldsymbol{P}_1\sum_{j=0}^{\infty}(\boldsymbol{I}-\boldsymbol{D})^j\boldsymbol{\Pi}_1 z^j = \sum_{j=0}^{\infty}\boldsymbol{M}_j z^j = \boldsymbol{M}(z) \tag{1.459}$$

(c)
$$\boldsymbol{M}(1) = \boldsymbol{A}^{\neq} \tag{1.460}$$

$$r(\boldsymbol{M}(1)) = r(\boldsymbol{A}) \tag{1.461}$$

$$\boldsymbol{M}(1)\boldsymbol{N}_1 = \boldsymbol{0} \tag{1.462}$$

$$\boldsymbol{M}_0 = \boldsymbol{A}^{\neq}\boldsymbol{A} = (\boldsymbol{I}-\boldsymbol{N}_1) \tag{1.463}$$

$$\boldsymbol{M}_j = (\boldsymbol{I}-\boldsymbol{A})^j - \boldsymbol{N}_1,\ \ j \geq 1 \tag{1.464}$$

Proof

Proof of (a) It is sufficient observe that, in light of (1.455) and by virtue of (1.307) of Sect. 1.7 we have

$$\lim_{z \to 1} (1 - z)\, \boldsymbol{A}^{-1}(z) = \boldsymbol{P}_2 \boldsymbol{\Pi}_2 = \boldsymbol{N}_1 \tag{1.465}$$

and that, according to (1.205), (1.207) and (1.216) of Sect. 1.5 the following proves to be true (see also the Remark in Sect. 1.4)

$$\boldsymbol{P}_2 \boldsymbol{\Pi}_2 = \boldsymbol{C}_\perp (\boldsymbol{B}'_\perp \boldsymbol{C}_\perp)^{-1} \boldsymbol{B}'_\perp = \boldsymbol{I} - \boldsymbol{A}^{\neq} \boldsymbol{A} \tag{1.466}$$

Proof of (b) By comparing (1.425) and (1.455) we derive that the term $\boldsymbol{P}_1 \sum_{j=0}^{\infty} (\boldsymbol{I} - \boldsymbol{D})^j \boldsymbol{\Pi}_1 z^j$ plays the rôle of regular component $\sum_{j=0}^{\infty} \boldsymbol{M}_j z^j$ in the Laurent expansion of $\boldsymbol{A}^{-1}(z)$ about $z = 1$. Hence, bearing in mind (1.308) of Sect. 1.7 and (1.203) of Sect. 1.5 we get

$$\begin{aligned} \boldsymbol{M}(1) &= -\lim_{z \to 1} \frac{\partial [(1-z)\boldsymbol{A}^{-1}(z)]}{\partial z} = \\ &\lim_{z \to 1} \boldsymbol{P}_1 \sum_{j=0}^{\infty} (\boldsymbol{I} - \boldsymbol{D})^j \boldsymbol{\Pi}_1 z^j = \boldsymbol{P}_1 \boldsymbol{D}^{-1} \boldsymbol{\Pi}_1 = \boldsymbol{A}^{\neq} \end{aligned} \tag{1.467}$$

with (1.461) and (1.462) which follow from the Drazin inverse properties (see Definition 6 in Sect. 1.1) because of (1.458).

Proof of (c) Finally, by comparing term to term the right-hand sides of (1.425) and (1.455), we get

$$\boldsymbol{M}_j = \boldsymbol{P}_1 (\boldsymbol{I} - \boldsymbol{D})^j \boldsymbol{\Pi}_1, \tag{1.468}$$

which for $j = 0$ specializes into

$$\boldsymbol{M}_0 = \boldsymbol{P}_1 \boldsymbol{\Pi}_1 = \boldsymbol{I} - \boldsymbol{P}_2 \boldsymbol{\Pi}_2 = \boldsymbol{A}^{\neq} \boldsymbol{A} = \boldsymbol{I} - \boldsymbol{N}_1 \tag{1.469}$$

because of (1.213) of Sect. 1.5 and of (1.458) above.

Moreover, upon resorting to (1.204), (1.206), (1.208) and (1.209) of Sect. 1.5 and (1.61) of Sect. 1.2, we can write

$$\begin{aligned} \boldsymbol{M}_j &= \boldsymbol{P}_1 (\boldsymbol{I} - \boldsymbol{D})^j \boldsymbol{\Pi}_1 \\ &= \boldsymbol{B} \sum_{\gamma=0}^{j} (-1)^\gamma \binom{j}{\gamma} (\boldsymbol{C}'\boldsymbol{B})^\gamma (\boldsymbol{C}'\boldsymbol{B})^{-1} \boldsymbol{C}' = \sum_{\gamma=0}^{j} (-1)^\gamma \binom{j}{\gamma} \boldsymbol{B} (\boldsymbol{C}'\boldsymbol{B})^{\gamma-1} \boldsymbol{C}' \\ &= \boldsymbol{B}(\boldsymbol{C}'\boldsymbol{B})^{-1}\boldsymbol{C}' + \sum_{\gamma=1}^{j} (-1)^\gamma \binom{j}{\gamma} (\boldsymbol{B}\boldsymbol{C}')^\gamma = \end{aligned} \tag{1.470}$$

$$= \boldsymbol{B}(\boldsymbol{C}'\boldsymbol{B})^{-1}\boldsymbol{C}' - \boldsymbol{I} + \sum_{\gamma=0}^{j} (-1)^{\gamma} \binom{j}{\gamma} (\boldsymbol{B}\boldsymbol{C}')^{\gamma}$$

$$= -\boldsymbol{N}_1 + \sum_{\gamma=0}^{j} (-1)^{\gamma} \binom{j}{\gamma} (\boldsymbol{B}\boldsymbol{C}')^{\gamma} = (\boldsymbol{I} - \boldsymbol{A})^{j} - \boldsymbol{N}_1$$

□

Theorem 3

Consider a linear matrix polynomial specified as in Theorem 1 *and let* $\boldsymbol{A}$ *be of index* $\upsilon = 2$. *Then the inverse of the matrix polynomial* $\boldsymbol{A}(z)$ *has a pole of second order about* $z = 1$. *Indeed the following holds*

$$\boldsymbol{A}^{-1}(z) = \boldsymbol{P} \begin{bmatrix} \boldsymbol{I} + \sum_{j=1}^{\infty} (\boldsymbol{I} - \boldsymbol{D})^{j} z^{j} & \boldsymbol{0} \\ \boldsymbol{0} & (1-z)^{-1}\boldsymbol{I} - (1-z)^{-2}\boldsymbol{\Upsilon} z \end{bmatrix} \boldsymbol{P}^{-1}$$
$$= \boldsymbol{P}_1 \left(\boldsymbol{I} + \sum_{j=1}^{\infty} (\boldsymbol{I} - \boldsymbol{D})^{j} z^{j} \right) \boldsymbol{\Pi}_1 + \frac{1}{(1-z)} \boldsymbol{P}_2 (\boldsymbol{I} + \boldsymbol{\Upsilon}) \boldsymbol{\Pi}_2 - \frac{1}{(1-z)^2} \boldsymbol{P}_2 \boldsymbol{\Upsilon} \boldsymbol{\Pi}_2 \tag{1.471}$$

where $\boldsymbol{D}, \boldsymbol{P}, \boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{\Pi}_1, \boldsymbol{\Pi}_2$ *and* $\boldsymbol{\Upsilon}$ *are defined as in Corollary* 2.1 *and Theorem* 4 *of Sect*. 1.5.

Proof

Under the assumption that $ind(\boldsymbol{A}) = 2$, $\boldsymbol{A} = \boldsymbol{H} + \boldsymbol{K}$ (see Theorem 1 of Sect. 1.5), and according to Theorem 2 of Sect. 1.5, the following holds

$$\boldsymbol{A} = \boldsymbol{P} \begin{bmatrix} \boldsymbol{D} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Upsilon} \end{bmatrix} \boldsymbol{P}^{-1} \tag{1.472}$$

This, in turn, implies that

$$\boldsymbol{A}(z) = (1-z)(\boldsymbol{I} - \boldsymbol{A}) + \boldsymbol{A}$$
$$= (1-z)\left[\boldsymbol{P}\boldsymbol{P}^{-1} - \boldsymbol{P} \begin{bmatrix} \boldsymbol{D} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Upsilon} \end{bmatrix} \boldsymbol{P}^{-1} \right] + \boldsymbol{P} \begin{bmatrix} \boldsymbol{D} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Upsilon} \end{bmatrix} \boldsymbol{P}^{-1} \tag{1.473}$$

$$= \boldsymbol{P}\begin{bmatrix} \boldsymbol{I}-(\boldsymbol{I}-\boldsymbol{D})z & \boldsymbol{0} \\ \boldsymbol{0} & (1-z)\boldsymbol{I}+\boldsymbol{\Upsilon}z \end{bmatrix}\boldsymbol{\Pi}$$

where $\boldsymbol{\Pi} = \boldsymbol{P}^{-1}$.

By inverting the right-hand side of (1.473), taking into account that $(\boldsymbol{I}-\boldsymbol{D})$ is a stable matrix (see the Remark prior to Theorem 2), bearing in mind the nilpotency of $\boldsymbol{\Upsilon}$ and the relation

$$[(1-z)\boldsymbol{I}+\boldsymbol{\Upsilon}z]^{-1} = (1-z)^{-1}\boldsymbol{I} - (1-z)^{-2}\boldsymbol{\Upsilon}z \tag{1.474}$$

we can easily obtain formula (1.471) and conclude that $\boldsymbol{A}^{-1}(z)$ has a second-order pole at $z = 1$.

□

Corollary 3.1

Let $\boldsymbol{A}(z)$ be a linear polynomial defined as in Theorem 3. Then, the following hold

(a) $$-\boldsymbol{P}_2\boldsymbol{\Upsilon}\boldsymbol{\Pi}_2 = \boldsymbol{N}_2 = -\boldsymbol{H} = -\boldsymbol{A}(\boldsymbol{I}-\boldsymbol{A}\boldsymbol{A}^D) = -\boldsymbol{A}\,\boldsymbol{C}_{2\perp}(\boldsymbol{B}_{2\perp}'\boldsymbol{C}_{2\perp})^{-1}\boldsymbol{B}_{2\perp}' \tag{1.475}$$

$$\boldsymbol{P}_2(\boldsymbol{I}+\boldsymbol{\Upsilon})\boldsymbol{\Pi}_2 = \boldsymbol{N}_1 = (\boldsymbol{I}+\boldsymbol{A})(\boldsymbol{I}-\boldsymbol{A}\boldsymbol{A}^D) = (\boldsymbol{I}+\boldsymbol{A})\,\boldsymbol{C}_{2\perp}(\boldsymbol{B}_{2\perp}'\boldsymbol{C}_{2\perp})^{-1}\boldsymbol{B}_{2\perp}' \tag{1.476}$$

$$\boldsymbol{P}_2\boldsymbol{\Pi}_2 = \boldsymbol{N}_2 + \boldsymbol{N}_1 = \boldsymbol{C}_{2\perp}(\boldsymbol{B}_{2\perp}'\boldsymbol{C}_{2\perp})^{-1}\boldsymbol{B}_{2\perp}' \tag{1.477}$$

(b) $$\boldsymbol{P}_1\sum_{j=0}^{\infty}(\boldsymbol{I}-\boldsymbol{D})^j\boldsymbol{\Pi}_1 z^j = \sum_{j=0}^{\infty}\boldsymbol{M}_j z^j = \boldsymbol{M}(z) \tag{1.478}$$

(c) $$\boldsymbol{M}(1) = \boldsymbol{A}^{\boldsymbol{D}} \tag{1.479}$$

$$r(\boldsymbol{M}(1)) = r(\boldsymbol{A}^2) \tag{1.480}$$

$$\boldsymbol{M}(1)\boldsymbol{N}_2 = \boldsymbol{0}, \boldsymbol{M}(1)\boldsymbol{N}_1 = \boldsymbol{0} \tag{1.481}$$

$$\boldsymbol{M}_0 = \boldsymbol{A}\boldsymbol{A}^D = \boldsymbol{A}^D\boldsymbol{A} \tag{1.482}$$

$$\boldsymbol{M}_j = (\boldsymbol{I}-\boldsymbol{A})^j - \boldsymbol{N} + j\boldsymbol{A}\boldsymbol{N} \quad, j \geq 1 \tag{1.483}$$

where $\boldsymbol{N}$ stands for $\boldsymbol{N}_1 + \boldsymbol{N}_2$.

Proof

Proof of (a) The proof of (1.475) is straightforward because, in light of (1.471) and by virtue of (1.319) of Sect. 1.7 we have

$$\lim_{z\to 1}(1-z)^2 \boldsymbol{A}^{-1}(z) = -\boldsymbol{P}_2 \boldsymbol{\Upsilon\Pi}_2 = \boldsymbol{N}_2 \tag{1.484}$$

and then according to (1.240) of Sect. 1.5 we obtain (see also the Remark of Sect. 1.4)

$$-\boldsymbol{P}_2 \boldsymbol{\Upsilon\Pi}_2 = -\boldsymbol{H} = -\boldsymbol{A}(\boldsymbol{I}-\boldsymbol{A}\boldsymbol{A}^D) = -\boldsymbol{A}\,\boldsymbol{C}_{2\perp}(\boldsymbol{B}'_{2\perp}\boldsymbol{C}_{2\perp})^{-1}\boldsymbol{B}'_{2\perp} \tag{1.485}$$

To prove (1.476) observe that, in light of (1.471) and taking into account (1.320) of Sect. 1.7 we have

$$\lim_{z\to 1}\frac{\partial(1-z)^2 \boldsymbol{A}^{-1}(z)}{\partial z} = \boldsymbol{P}_2(\boldsymbol{I}+\boldsymbol{\Upsilon})\boldsymbol{\Pi}_2 = \boldsymbol{N}_1 \tag{1.486}$$

This, by virtue of (1.237) and (1.240) of Sect. 1.5, can be rewritten as

$$\boldsymbol{P}_2(\boldsymbol{I}+\boldsymbol{\Upsilon})\boldsymbol{\Pi}_2 = (\boldsymbol{I}+\boldsymbol{A})(\boldsymbol{I}-\boldsymbol{A}\boldsymbol{A}^D) = (\boldsymbol{I}+\boldsymbol{A})(\boldsymbol{C}_{2\perp}(\boldsymbol{B}'_{2\perp}\boldsymbol{C}_{2\perp})^{-1}\boldsymbol{B}'_{2\perp}) \tag{1.487}$$

The proof of (1.477) is straightforward in light of (1.475) and (1.476) above (see (1.237) of Sect. 1.5).

Proof of (b) By comparing (1.425) and (1.471) we derive that the term $\boldsymbol{P}_1\sum_{j=0}^{\infty}(\boldsymbol{I}-\boldsymbol{D})^j\boldsymbol{\Pi}_1 z^j$ plays the rôle of regular component $\sum_{j=0}^{\infty}\boldsymbol{M}_j z^j$ in the Laurent expansion of $\boldsymbol{A}^{-1}(z)$ about $z=1$. Hence, from (1.321) of Sect. 1.7 and (1.220) of Sect. 1.5 we have

$$\begin{aligned}\boldsymbol{M}(1) &= \frac{1}{2}\lim_{z\to 1}\frac{\partial^2[(1-z)^2\boldsymbol{A}^{-1}(z)]}{\partial z^2} = \\ &= \lim_{z\to 1}\boldsymbol{P}_1\Big(\sum_{j=0}^{\infty}(\boldsymbol{I}-\boldsymbol{D})^j z^j\Big)\boldsymbol{\Pi}_1 = \boldsymbol{P}_1\boldsymbol{D}^{-1}\boldsymbol{\Pi}_1 = \boldsymbol{A}^D\end{aligned} \tag{1.488}$$

The relations (1.480) and (1.481) follow from the Drazin inverse properties (see Definition 6 in Sect. 1.1) bearing in mind (1.475) and (1.476).

Proof of (c) Finally, by comparing term to term the right-hand sides of (1.425) and (1.471), we get

$$M_j = P_1(I - D)^j \Pi_1, \tag{1.489}$$

which for $j = 0$ specializes into

$$M_0 = P_1\Pi_1 = I - P_2\Pi_2 = I - N \tag{1.490}$$

because of (1.477).

Moreover, upon recalling (1.221), (1.223), (1.225), (1.240), as well as Theorem 5 of Sect. 1.2 we find out that

$$\begin{aligned}
& P_1(I - D)^j \Pi_1 = \\
& B_2 \sum_{\gamma=0}^{j} (-1)^\gamma \binom{j}{\gamma} (G'F)^\gamma (G'F)^{-2} C_2' = \sum_{\gamma=0}^{j} (-1)^\gamma \binom{j}{\gamma} BF(G'F)^{\gamma-2} G'C' \\
& = BF(G'F)^{-2}G'C' - jBF(G'F)^{-1}G'C' + \sum_{\gamma=2}^{j} (-1)^\gamma \binom{j}{\gamma} (BC')^\gamma \\
& = (BF(G'F)^{-2}G'C') - jB[I - G_\perp(F_\perp' G_\perp)^{-1} F_\perp']C' + \sum_{\gamma=2}^{j} (-1)^\gamma \binom{j}{\gamma} (BC')^\gamma \\
& = (BF(G'F)^{-2}G'C' - jBC' + jBG_\perp(F_\perp' G_\perp)^{-1} F_\perp' C' + \sum_{\gamma=2}^{j} (-1)^\gamma \binom{j}{\gamma} (BC')^\gamma \\
& = -N + j\,BG_\perp(F_\perp' G_\perp)^{-1} F_\perp' C' + \sum_{\gamma=0}^{j} (-1)^\gamma \binom{j}{\gamma} (BC')^\gamma = (I - A)^j - N + jAN
\end{aligned} \tag{1.491}$$

□

The extension of the above results to a multi-lag dynamic specification can be easily obtained through a companion-form representation.

1.10 Index and Rank Properties of Matrix Coefficients vs. Pole Order in Matrix Polynomial Inversion

This section will be devoted to presenting several relationships between either index or rank characteristics of matrices, in the Taylor expansion of a matrix polynomial $A(z)$ about $z = 1$, and the order of the poles inherent in the Laurent expansion of its inverse, $A^{-1}(z)$, in a deleted neighbourhood of $z = 1$. Basically, references will be made to previous sections for notational purposes as well as for relevant results.

As far as notation is concerned, A will denote $A(z)$ evaluated at $z = 1$ and the dot notation will be adopted for derivatives, e.g. $\dot{A} = \left(\frac{\partial A(z)}{\partial z}\right)_{z=1}$.

In addressing the issue of establishing the pole order of a unit-root, we will avail ourselves of a twofold approach.

On the one hand, we will assume a general K-order matrix polynomial, that is

$$\boldsymbol{A}(z) = \sum_{k=0}^{K} \boldsymbol{A}_k z^k, \ \boldsymbol{A}_0 = \boldsymbol{I}_n, \ \boldsymbol{A}_K \neq \boldsymbol{0}_n \tag{1.492}$$

as a reference frame, and determine the pole order in terms of ranks of (functions of) the coefficient matrices.

On the other hand, we will tackle the problem from the standpoint of a linear polynomial specification – actually a first-degree mirror image of a higher-degree polynomial in companion form – that is

$$\begin{aligned} \breve{\boldsymbol{A}}(z) &= \boldsymbol{I}_{\breve{n}} + z\breve{\boldsymbol{A}}_1 \\ &= (1-z)\boldsymbol{I}_{\breve{n}} + z\breve{\boldsymbol{A}}, \quad \breve{\boldsymbol{A}} = \boldsymbol{I}_{\breve{n}} + \breve{\boldsymbol{A}}_1 \end{aligned} \tag{1.493}$$

where $\breve{\boldsymbol{A}}_1$ is specified as in (1.362) of Sect. 1.8.

Let us now establish several theorems which characterize the occurrence of simple and second-order poles associated with a unit root.

Theorem 1

Each one of the following conditions is necessary and sufficient for the inverse of a matrix polynomial $\boldsymbol{A}(z)$ to be analytical at $z = 1$,

(a) $\qquad det\boldsymbol{A} \neq 0 \qquad (1.494)$

(b) $\qquad ind(\boldsymbol{A}) = 0 \qquad (1.495)$

Under the stated condition, $z = 1$ is neither a zero of $\boldsymbol{A}(z)$ nor a pole of $\boldsymbol{A}^{-1}(z)$.

Proof

The proof rests on the arguments put forward while introducing the Laurent expansion in Sect. 1.7, as well as on Definition 5 of Sect. 1.1. Reference can also be made to the Proposition prior to Theorem 5 in Sect. 1.6.

□

The following lemma paves the way to the subsequent theorems and corollaries bridging the gap between first and higher – order polynomials by means of companion – form arguments.

Lemma 2

Let $\breve{A}$ be a singular square matrix of order $\breve{n}$ partitioned in this way

$$\underset{(\breve{n},\breve{n})}{\breve{A}} = \begin{bmatrix} \breve{A}_{11} & \breve{A}_{12} \\ \breve{A}_{21} & \breve{A}_{22} \end{bmatrix} = \left[\begin{array}{c|ccccc} I_n + A_1 & A_2 & A_3 & .. & A_K \\ \hline -I_n & I_n & 0 & 0 & 0 \\ 0 & -I_n & I_n & 0 & 0 \\ .. & .. & .. & .. & .. \\ 0 & 0 & 0 & -I_n & 0 \end{array}\right] \tag{1.496}$$

where $\breve{n} = nK$, A_1, A_2, ... , A_K are square matrices of order *n,* and let A denote the Schur complement of $\breve{A}_{22}$, namely

$$A = \breve{A}_{11} - \breve{A}_{12}\breve{A}_{22}^{-1}\breve{A}_{21} = I + A_1 + A_2 + ... + A_K \tag{1.497}$$

Further, denote by $\dot{A}$ the matrix

$$\dot{A} = \sum_{k=1}^{K} kA_k \tag{1.498}$$

and by B and C full column-rank matrices obtained by a rank factorization of A, that is

$$A = BC' \tag{1.499}$$

Then, the following statements are equivalent

$$ind(\breve{A}) = 1 \tag{1.500}$$

$$det(B'_{\perp}\dot{A}C_{\perp}) \neq 0 \tag{1.501}$$

Proof

First of all observe that A can be factorized as follows

$$\breve{A} = \begin{bmatrix} I & \breve{A}_{12}\breve{A}_{22}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} A & 0 \\ \breve{A}_{21} & \breve{A}_{22} \end{bmatrix} = \begin{bmatrix} A & \breve{A}_{12} \\ 0 & \breve{A}_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ \breve{A}_{22}^{-1}\breve{A}_{12} & I \end{bmatrix} \tag{1.502}$$

and that in light of the rank factorization $\breve{A} = \breve{B}\breve{C}'$, the conclusion that

$$\breve{B}'_{\perp}\breve{A} = 0 \tag{1.503}$$

$$\breve{A}\breve{C}_{\perp} = \boldsymbol{0} \tag{1.504}$$

is easily drawn.

Now, equations (1.503) and (1.504) can be more conveniently rewritten in partitioned form as

$$[\boldsymbol{Y}_1, \boldsymbol{Y}_2]\begin{bmatrix} \boldsymbol{BC}' & \breve{\boldsymbol{A}}_{12} \\ \boldsymbol{0} & \breve{\boldsymbol{A}}_{22} \end{bmatrix} = [\boldsymbol{0}', \boldsymbol{0}'] \tag{1.505}$$

$$\begin{bmatrix} \boldsymbol{BC}' & \boldsymbol{0} \\ \breve{\boldsymbol{A}}_{21} & \breve{\boldsymbol{A}}_{22} \end{bmatrix}\begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix} \tag{1.506}$$

by resorting to (1.502) and introducing the partitioning $\breve{\boldsymbol{B}}'_{\perp} = [\boldsymbol{Y}_1, \boldsymbol{Y}_2]$ and $\breve{\boldsymbol{C}}'_{\perp} = \begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{bmatrix}$.

Solving (1.505) yields

$$\boldsymbol{Y}_1' = \boldsymbol{B}'_{\perp}, \boldsymbol{Y}_2' = [-\boldsymbol{B}'_{\perp}(\boldsymbol{A}_2 + \boldsymbol{A}_3 + \ldots \boldsymbol{A}_K), -\boldsymbol{B}'_{\perp}(\boldsymbol{A}_3 + \ldots \boldsymbol{A}_K), \ldots - \boldsymbol{B}'_{\perp}\boldsymbol{A}_K] \tag{1.507}$$

and solving (1.506) yields

$$\boldsymbol{X}_1 = \boldsymbol{C}_{\perp}, \qquad \boldsymbol{X}_2 = \boldsymbol{u}_{K-1} \otimes \boldsymbol{C}_{\perp} \tag{1.508}$$

where the symbol $\otimes$ denotes the Kronecker product (see, e.g., Neudecker 1968) and $\boldsymbol{u}$ is the vector of 1's.

By making use of (1.507) and (1.508), some computations give

$$\begin{aligned}
\breve{\boldsymbol{B}}'_{\perp}\breve{\boldsymbol{C}}_{\perp} &= \boldsymbol{B}'_{\perp}\boldsymbol{C}_{\perp} - \boldsymbol{B}'_{\perp}(\boldsymbol{A}_2 + \boldsymbol{A}_3 + \ldots \boldsymbol{A}_K + \boldsymbol{A}_3 + \ldots + \boldsymbol{A}_K + \ldots + \boldsymbol{A}_K)\boldsymbol{C}_{\perp} \\
&= \boldsymbol{B}'_{\perp}\boldsymbol{C}_{\perp} - \boldsymbol{B}'_{\perp}\sum_{k=2}^{K}(k-1)\boldsymbol{A}_k\boldsymbol{C}_{\perp} = \boldsymbol{B}'_{\perp}\boldsymbol{C}_{\perp} - \boldsymbol{B}'_{\perp}\left(\sum_{k=1}^{K} k\boldsymbol{A}_k - \sum_{k=1}^{K}\boldsymbol{A}_k\right)\boldsymbol{C}_{\perp} \\
&= \boldsymbol{B}'_{\perp}(\boldsymbol{A} - \dot{\boldsymbol{A}})\boldsymbol{C}_{\perp} = -\boldsymbol{B}'_{\perp}\dot{\boldsymbol{A}}\boldsymbol{C}_{\perp}
\end{aligned} \tag{1.509}$$

which, in light of Corollary 6.1 in Sect. 1.2, proves (1.501).

□

Theorem 3

*Each of the following statements is equivalent for the inverse of a matrix polynomial **A**(z) to exhibit a simple pole at z = 1,*

(a1)
$$\begin{cases} det\, \boldsymbol{A} = 0, \quad \boldsymbol{A} \neq \boldsymbol{0} & (1.510) \\ det \begin{bmatrix} -\dot{\boldsymbol{A}} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix} \neq 0 & (1.510') \end{cases}$$

*where **B** and **C** are defined as per a rank factorization of **A**, that is*

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}' \tag{1.511}$$

(b1)
$$ind(\breve{\boldsymbol{A}}) = 1 \tag{1.512}$$

Proof

From (1.252) of Sect. 1.6 it follows that

$$\frac{1}{(1-z)}\boldsymbol{A}(z) = \frac{1}{(1-z)}[(1-z)\,\boldsymbol{Q}(z) + \boldsymbol{B}\boldsymbol{C}'] \tag{1.513}$$

where $\boldsymbol{Q}(z)$ is as defined in (1.248) of Sect. 1.6.

According to Theorem 1 of Sect. 1.6 the right-hand side of (1.513) corresponds to the Schur complement of the lower diagonal block, $(z-1)\,\boldsymbol{I}$, in the partitioned matrix

$$\boldsymbol{P}(z) = \begin{bmatrix} \boldsymbol{Q}(z) & \boldsymbol{B} \\ \boldsymbol{C}' & (z-1)\boldsymbol{I} \end{bmatrix} \tag{1.514}$$

Therefore, by (1.105) of Theorem 1 of Sect. 1.4, we know that

$$(1-z)\,\boldsymbol{A}^{-1}(z) = [\boldsymbol{I} \quad \boldsymbol{0}] \begin{bmatrix} \boldsymbol{Q}(z) & \boldsymbol{B} \\ \boldsymbol{C}' & (z-1)\boldsymbol{I} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} \tag{1.515}$$

provided $det\, \boldsymbol{P}(z) \neq 0$.

By taking the limit of both sides of (1.515) as z tends to 1, bearing in mind Theorem 1 of Sect. 1.6, the outcome is

$$\begin{aligned} &\lim_{z\to 1} [(1-z)\,\boldsymbol{A}^{-1}(z)] = \\ &= [\boldsymbol{I} \;\; \boldsymbol{0}] \begin{bmatrix} \boldsymbol{Q}(1) & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} = [\boldsymbol{I} \;\; \boldsymbol{0}] \begin{bmatrix} -\dot{\boldsymbol{A}} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} \end{aligned} \tag{1.516}$$

which, in view of Definition 5 of Sect. 1.6, leads to conclude that $z = 1$ is a simple pole of $\boldsymbol{A}^{-1}(z)$.

The proof of (b1) rests on Theorem 7 of Sect. 1.6.

□

Corollary 3.1

Each of the following conditions is equivalent to (1.510’), under (1.510),

(a2) $$det(\boldsymbol{B}'_{\perp}\dot{\boldsymbol{A}}\boldsymbol{C}_{\perp}) \neq 0 \tag{1.517}$$

(a3) $$r\left(\begin{bmatrix} \dot{\boldsymbol{A}} & \boldsymbol{A} \\ \boldsymbol{A} & \boldsymbol{0} \end{bmatrix}\right) = n + r(\boldsymbol{A}) \tag{1.518}$$

Condition (a2), assuming the singularity of $\boldsymbol{A}$, is equivalent to statement (b1) of the last theorem.

Proof

Equivalence of (a2) and (a1) follows from Theorem 3 of Sect. 1.4.

Equivalence of (a2) and (b1) follows from Lemma 2.

Equivalence of (a3) and (a1) is easily proved upon noticing that

$$\begin{aligned} & det\begin{bmatrix} -\dot{\boldsymbol{A}} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix} \neq 0 \Leftrightarrow r\left(\begin{bmatrix} -\dot{\boldsymbol{A}} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix}\right) \\ & = r\left(\begin{bmatrix} -\boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B} \end{bmatrix}\begin{bmatrix} -\dot{\boldsymbol{A}} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{C}' \end{bmatrix}\right) = r\left(\begin{bmatrix} \dot{\boldsymbol{A}} & \boldsymbol{A} \\ \boldsymbol{A} & \boldsymbol{0} \end{bmatrix}\right) \\ & = r(\boldsymbol{A}) + r(\boldsymbol{A}) + r((\boldsymbol{I} - \boldsymbol{A}\boldsymbol{A}^{g})\dot{\boldsymbol{A}}(\boldsymbol{I} - \boldsymbol{A}^{g}\boldsymbol{A})) \\ & = r(\boldsymbol{A}) + r(\boldsymbol{B}) + r(\boldsymbol{B}'_{\perp}\dot{\boldsymbol{A}}\boldsymbol{C}_{\perp}) = r(\boldsymbol{A}) + r(\boldsymbol{B}) + r(\boldsymbol{B}_{\perp}) = r(\boldsymbol{A}) + n \end{aligned} \tag{1.519}$$

in light of Theorem 19 in Marsaglia and Styan 1974, and of identities (1.87) and (1.88) of Sect. 1.2.

□

Lemma 4

Let $\breve{A}$, A, $\dot{A}$, $\breve{B}$, $\breve{C}$, B and C be defined as in Lemma 2. Furthermore, let R and S be obtained by a rank factorization of $B'_{\perp}\dot{A}C_{\perp}$, that is

$$B'_{\perp}\dot{A}C_{\perp} = RS' \tag{1.520}$$

Then, the following statements are equivalent

$$ind(\breve{A}) = 2 \tag{1.521}$$

$$det[R'_{\perp}B'_{\perp}\widetilde{A}C_{\perp}S_{\perp}] \neq 0 \tag{1.522}$$

where

$$\widetilde{A} = \frac{1}{2}\ddot{A} - \dot{A}A^{g}\dot{A} \tag{1.523}$$

$$\ddot{A} = \sum_{J=2}^{K} j(j-1)\, A_{j} \tag{1.524}$$

and

$$A^{g} = (C')^{g}B^{g} \tag{1.525}$$

is the Moore-Penrose inverse of A.

Proof

According to (1.509) of Lemma 2, any rank factorization of $\breve{B}'_{\perp}\breve{C}_{\perp}$ is also a rank factorization of $-B'_{\perp}\dot{A}C_{\perp}$ and vice-versa, which entails in particular that

$$R = \breve{R},\ R_{\perp} = \breve{R}_{\perp},\ S = -\breve{S},\ S_{\perp} = -\breve{S}_{\perp} \tag{1.526}$$

Additionally, by applying Theorem 2 of Sect. 1.4 to the matrix $\breve{A}$ and making use of (1.496)–(1.498), (1.507) and (1.508) the following results are easily obtained

$$\breve{A}_{\rho}^{-} = \begin{bmatrix} A^{g} & -A^{g}\breve{A}_{12}\breve{A}_{22}^{-1} \\ -\breve{A}_{22}^{-1}\breve{A}_{21}A^{g} & \breve{A}_{22}^{-1} + \breve{A}_{22}^{-1}\breve{A}_{21}A^{g}\breve{A}_{12}\breve{A}_{22}^{-1} \end{bmatrix} \tag{1.527}$$

$$\begin{aligned}
\breve{A}_{12}\breve{A}_{22}^{-1}X_2 &= \breve{A}_{12}\breve{A}_{22}^{-1}(u_{K-1}\otimes C_\perp) \\
&= [A_2, A_3, \dots, A_K]\begin{bmatrix} I & 0 & \cdots & 0 \\ I & I & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ I & I & I & I \end{bmatrix}\begin{bmatrix} C_\perp \\ C_\perp \\ \cdots \\ C_\perp \end{bmatrix} \\
&= \sum_{k=2}^{K}(k-1)A_k C_\perp = \left(\sum_{k=2}^{K}kA_k - \sum_{k=2}^{K}A_k\right)C_\perp = \left(\sum_{k=1}^{K}kA_k - \sum_{k=1}^{K}A_k\right)C_\perp \\
&= (\dot{A} - A + I)C_\perp = (\dot{A} + I)C_\perp
\end{aligned} \tag{1.528}$$

$$Y_2\breve{A}_{22}^{-1}\breve{A}_{21} = B'_\perp(\dot{A} + I) \tag{1.528'}$$

$$Y_2\breve{A}_{22}^{-1}X_2 = Y_2\breve{A}_{22}^{-1}(u_{K-1}\otimes C_\perp) = -\frac{1}{2}B'_\perp\ddot{A}C_\perp \tag{1.529}$$

Moreover, the following proves true

$$\begin{aligned}
\breve{B}'_\perp\breve{A}_\rho^-\breve{C}_\perp &= [Y_1, Y_2]\begin{bmatrix} A^g & -A^g\breve{A}_{12}\breve{A}_{22}^{-1} \\ -\breve{A}_{22}^{-1}\breve{A}_{21}A^g & \breve{A}_{22}^{-1} + \breve{A}_{22}^{-1}\breve{A}_{21}A^g\breve{A}_{12}\breve{A}_{22}^{-1} \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \\
&= Y_1A^gX_1 - Y_2\breve{A}_{22}^{-1}\breve{A}_{21}A^gX_1 - Y_1A^g\breve{A}_{12}\breve{A}_{22}^{-1}X_2 + Y_2\breve{A}_{22}^{-1}X_2 \\
&\quad + Y_2\breve{A}_{22}^{-1}\breve{A}_{21}A^g\breve{A}_{12}\breve{A}_{22}^{-1}X_2 = B'A^gC_\perp - B'_\perp(\dot{A}+I)A^gC_\perp \\
&\quad - B'_\perp A^g(\dot{A}+I)C_\perp - \frac{1}{2}B'_\perp\ddot{A}C_\perp + \frac{1}{2}B'_\perp(\dot{A}+I)A^g(\dot{A}+I)C_\perp \\
&= -B'_\perp\left(\frac{1}{2}\ddot{A} - \dot{A}A^g\dot{A}\right)C_\perp = -B'_\perp\tilde{A}C_\perp
\end{aligned} \tag{1.530}$$

Then, pre and post-multiplying (1.530) by $R'_\perp$ and $S_\perp$ respectively, and bearing in mind Corollary 8.1 of Sect. 1.2, inequality (1.522) is eventually dimostrated.

□

Remark

If reference is made to $\breve{A}$ and to the related matrices of Lemma 4, those introduced in Corollary 4.1 of Sect. 1.4 become

$$\boldsymbol{\mathcal{R}} = \breve{C}'_{\perp}(\breve{A}_{\rho}^{-})'\breve{B}_{\perp}\breve{R}_{\perp} = -C'_{\perp}\tilde{A}'B_{\perp}R_{\perp} \tag{1.531}$$

$$K = \breve{B}'_{\perp}\breve{A}_{\rho}^{-}\breve{C}_{\perp}\breve{S}_{\perp} = -B'_{\perp}\widetilde{A}_{\perp}C_{\perp}S_{\perp} \tag{1.532}$$

$$\Theta = \boldsymbol{\mathcal{R}}_{\perp}(\breve{S}'\boldsymbol{\mathcal{R}}_{\perp})^{-1}(K'_{\perp}\breve{R})^{-1}K'_{\perp} = \boldsymbol{\mathcal{R}}_{\perp}(S'\boldsymbol{\mathcal{R}}_{\perp})^{-1}(K'_{\perp}R)^{-1}K'_{\perp} \tag{1.533}$$

$$L = \breve{R}'_{\perp}\breve{B}'_{\perp}\breve{A}_{\rho}^{-}\breve{C}'_{\perp}\breve{S}_{\perp} = -R'_{\perp}B'_{\perp}\tilde{A}C_{\perp}S_{\perp} = R'_{\perp}K = \boldsymbol{\mathcal{R}}'S_{\perp} \tag{1.534}$$

$$\aleph = \breve{R}'_{\perp}\breve{B}'_{\perp}(\breve{A}_{\rho}^{-})^{2}\breve{C}_{\perp}\breve{S}_{\perp} = -R'_{\perp}\breve{B}'_{\perp}(\breve{A}_{\rho}^{-})^{2}\breve{C}_{\perp}S_{\perp} \tag{1.535}$$

$$\tilde{T} = \Theta - \breve{S}_{\perp}L^{-1}\aleph L^{-1}\breve{R}'_{\perp} = \Theta - S_{\perp}L^{-1}\aleph L^{-1}R'_{\perp} \tag{1.536}$$

Theorem 5

Each of the following set of conditions is equivalent for the inverse of a matrix polynomial $A(z)$ *to exhibit a pole of second order*

(ã1)

$$\begin{cases} det A = 0, \quad A \neq 0 & (1.537) \\ det(B'_{\perp}\dot{A}C_{\perp}) = 0, \quad B'_{\perp}\dot{A}C_{\perp} \neq 0 & (1.537') \\ det\begin{bmatrix} \widetilde{A} & (B_{\perp}R_{\perp})_{\perp} \\ (C_{\perp}S_{\perp})'_{\perp} & 0 \end{bmatrix} \neq 0 & (1.537'') \end{cases}$$

where B *and* C *are full column-rank matrices obtained by rank factorization of* A, R *and* S *are full column-rank matrices obtained by rank factorization of* $B'_{\perp}\dot{A}C_{\perp}$, *i.e.*

$$B'_{\perp}\dot{A}C_{\perp} = RS' \tag{1.538}$$

and $\widetilde{A}$ *is defined as in formula* (1.523)

$(\widetilde{b}1)$

$$ind(\breve{A}) = 2 \tag{1.539}$$

Proof

To prove (ã1) observe that from (1.257) of Sect. 1.6, it follows that

$$\frac{1}{(1-z)^2}\boldsymbol{A}(z) = \frac{1}{(1-z)^2}[(1-z)^2\ \boldsymbol{\Psi}(z) - (1-z)\dot{\boldsymbol{A}} + \boldsymbol{BC}'] \tag{1.540}$$

where $\boldsymbol{\Psi}(z)$ is as defined in (1.255) of Sect. 1.6.

Pre and post-multiplying $\dot{\boldsymbol{A}}$ by $(\boldsymbol{B}')^g\boldsymbol{B}' + (\boldsymbol{B}'_\perp)^g\boldsymbol{B}'_\perp = \boldsymbol{I}$ and by $\boldsymbol{C}\boldsymbol{C}^g + \boldsymbol{C}_\perp\boldsymbol{C}_\perp^g = \boldsymbol{I}$ respectively, and making use of (1.532) we get

$$\begin{aligned}\dot{\boldsymbol{A}} &= (\boldsymbol{B}'_\perp)^g\boldsymbol{B}'_\perp\dot{\boldsymbol{A}}\boldsymbol{C}_\perp\boldsymbol{C}_\perp^g + (\boldsymbol{B}'_\perp)^g\boldsymbol{B}'_\perp\dot{\boldsymbol{A}}(\boldsymbol{C}')^g\boldsymbol{C}' + \boldsymbol{BB}^g\dot{\boldsymbol{A}}\\ &= (\boldsymbol{B}'_\perp)^g\boldsymbol{R}\boldsymbol{S}'\boldsymbol{C}_\perp^g + (\mathbf{I} - \boldsymbol{BB}^g)\dot{\boldsymbol{A}}(\boldsymbol{C}')^g\boldsymbol{C}' + \boldsymbol{BB}^g\dot{\boldsymbol{A}}\end{aligned} \tag{1.541}$$

and therefore the equality

$$\begin{aligned}(1-z)\dot{\boldsymbol{A}} - \boldsymbol{A} &= (1-z)(\boldsymbol{B}'_\perp)^g\boldsymbol{R}\boldsymbol{S}'\boldsymbol{C}_\perp^g(1-z)\dot{\boldsymbol{A}}(\boldsymbol{C}')^g\boldsymbol{C}'\\ &\quad + (1-z)\boldsymbol{BB}^g\dot{\boldsymbol{A}} + \boldsymbol{B}[(1-z)\boldsymbol{B}^g\dot{\boldsymbol{A}}(\boldsymbol{C}')^g + \boldsymbol{I}]\boldsymbol{C}'\end{aligned} \tag{1.542}$$

follows as a by-product.

Substituting the right-hand side of (1.542) into (1.540) and putting

$$\boldsymbol{F} = [\boldsymbol{B}, (\boldsymbol{B}'_\perp)^g\boldsymbol{R},\quad \dot{\boldsymbol{A}}(\boldsymbol{C}')^g],\quad \boldsymbol{G}' = \begin{bmatrix}\boldsymbol{C}'\\ \boldsymbol{S}'\boldsymbol{C}_\perp^g\\ \boldsymbol{B}^g\dot{\boldsymbol{A}}\end{bmatrix} \tag{1.543}$$

$$\boldsymbol{V}(z) = \begin{bmatrix} -\boldsymbol{I} - (1-z)\boldsymbol{B}^g\dot{\boldsymbol{A}}(\boldsymbol{C}')^g & \boldsymbol{0} & (1-z)\boldsymbol{I}\\ \boldsymbol{0} & (1-z)\boldsymbol{I} & \boldsymbol{0}\\ (1-z)\boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0}\end{bmatrix} \tag{1.544}$$

$$\begin{aligned}\boldsymbol{\Lambda}(z) &= (1-z)^2\boldsymbol{V}^{-1}(z)\\ &= \begin{bmatrix}\boldsymbol{0} & \boldsymbol{0} & (1-z)\boldsymbol{I}\\ \boldsymbol{0} & (1-z)\boldsymbol{I} & \boldsymbol{0}\\ (1-z)\boldsymbol{I} & \boldsymbol{0} & (1-z)\boldsymbol{B}^g\dot{\boldsymbol{A}}(\boldsymbol{C}')^g + \boldsymbol{I}\end{bmatrix}\end{aligned} \tag{1.545}$$

we can rewrite (1.540) in the form

$$\frac{1}{(1-z)^2}\boldsymbol{A}(z) = \frac{1}{(1-z)^2}[(1-z)^2\boldsymbol{\Psi}(z) - \boldsymbol{F}\boldsymbol{V}(z)\boldsymbol{G}'] \tag{1.546}$$

We note that the right-hand side of (1.546) corresponds to the Schur complement of the lower diagonal block, $\boldsymbol{\Lambda}(z)$, of the partitioned matrix

$$\boldsymbol{P}(z) = \begin{bmatrix} \boldsymbol{\Psi}(z) & \boldsymbol{F} \\ \boldsymbol{G}' & \boldsymbol{\Lambda}(z) \end{bmatrix} \tag{1.547}$$

Hence by (1.105) of Theorem 1 in Sect. 1.4, the following holds

$$(1-z)^2 \boldsymbol{A}^{-1}(z) = [\boldsymbol{I} \;\; \boldsymbol{0}] \begin{bmatrix} \boldsymbol{\Psi}(z) & \boldsymbol{F} \\ \boldsymbol{G}' & \boldsymbol{\Lambda}(z) \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} \tag{1.548}$$

provided $det\, \boldsymbol{P}(z) \neq 0$.

The right-hand side of (1.548) can be rewritten in the more convenient partitioned form

$$[\boldsymbol{I} \quad \boldsymbol{0} \;\vdots\; \boldsymbol{0}] \left[\begin{array}{cc:c} \boldsymbol{\Psi}(z) & (\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp} & \dot{\boldsymbol{A}}(\boldsymbol{C}')^{g} \\ (\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp} & (1-z)\boldsymbol{U}_1\boldsymbol{U}'_1 & (1-z)\boldsymbol{U}_2 \\ \cdots & \cdots\cdots\cdots & \cdots \\ \boldsymbol{B}^{g}\dot{\boldsymbol{A}} & (1-z)\boldsymbol{U}'_2 & \boldsymbol{\Theta}(z) \end{array}\right]^{-1} \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \\ \cdots \\ \boldsymbol{0} \end{bmatrix} \tag{1.549}$$

where

$$\Theta(z) = (1-z)\, \boldsymbol{B}^{g}\, \dot{\boldsymbol{A}}\, (\boldsymbol{C}')^{g} + \boldsymbol{I}, \tag{1.550}$$

and

$$\boldsymbol{U}_1 = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{I} \end{bmatrix}, \qquad \boldsymbol{U}_2 = \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} = \boldsymbol{U}_{1\perp} \tag{1.551}$$

denote selection matrices such that

$$(\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp}\boldsymbol{U}_1 = (\boldsymbol{B}'_{\perp})^{g}\boldsymbol{R},\; (\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp}\boldsymbol{U}_2 = \boldsymbol{B} \tag{1.552}$$

$$(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})_{\perp}\boldsymbol{U}_1 = (\boldsymbol{C}'_{\perp})^{g}\boldsymbol{S},\; (\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})_{\perp}\boldsymbol{U}_2 = \boldsymbol{C} \tag{1.553}$$

Here, use has been made of the formulas

$$(\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp} = [\boldsymbol{B},\; (\boldsymbol{B}'_{\perp})^{g}\boldsymbol{R}],\, (\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})_{\perp} = [\boldsymbol{C},\; (\boldsymbol{C}'_{\perp})^{g}\boldsymbol{S}] \tag{1.554}$$

established in Theorem 2 of Sect. 1.2.

Furthermore, observe that

$$P(1)=\begin{bmatrix}\Psi(1) & F\\ G' & A(1)\end{bmatrix}=\left[\begin{array}{cc:c}\frac{1}{2}\ddot{A} & (B_\perp R_\perp)_\perp & \dot{A}(C')^g\\ (C_\perp S_\perp)'_\perp & 0 & 0\\ \cdots & \cdots & \cdots\\ B^g\dot{A} & 0 & I\end{array}\right] \tag{1.555}$$

in light of (1.259) of Sect. 1.6.

Now, since the matrix

$$\begin{aligned}J&=\begin{bmatrix}\frac{1}{2}\ddot{A} & (B_\perp R_\perp)_\perp\\ (C_\perp S_\perp)'_\perp & 0\end{bmatrix}-\begin{bmatrix}\dot{A}(C')^g\\ 0\end{bmatrix}[B^g\dot{A}\quad 0]\\ &=\begin{bmatrix}\frac{1}{2}\ddot{A}-\dot{A}A^g\dot{A} & (B_\perp R_\perp)_\perp\\ (C_\perp S_\perp)'_\perp & 0\end{bmatrix}\end{aligned} \tag{1.556}$$

corresponds to the Schur complement of the lower diagonal block $\boldsymbol{I}$ of the partitioned matrix on the right-hand side of (1.555), it follows that

$$det(P(1))=det(J)\,det(I)=det\,(J) \tag{1.557}$$

which, in turn, entails that

$$det(P(1))\neq 0 \tag{1.558}$$

by virtue of (1.537″).

In view of the foregoing, should we take the limit of both sides of (1.548) as z tends to 1, we would obtain

$$\begin{aligned}\lim_{z\to 1}[(1-z)^2A^{-1}(z)]&=[I\quad 0]\begin{bmatrix}\frac{1}{2}\ddot{A} & F\\ G' & A(1)\end{bmatrix}^{-1}\begin{bmatrix}I\\ 0\end{bmatrix}\\ &=[I\quad 0\ \vdots\ 0]\left[\begin{array}{cc:c}\frac{1}{2}\ddot{A} & (B_\perp R_\perp)_\perp & \dot{A}(C')^g\\ (C_\perp S_\perp)'_\perp & 0 & 0\\ \cdots & \cdots & \cdots\\ B^g\dot{A} & 0 & I\end{array}\right]^{-1}\begin{bmatrix}I\\ 0\\ \cdots\\ 0\end{bmatrix}\end{aligned} \tag{1.559}$$

which, in view of (1.105) of Theorem 1 in Sect. 1.4, implies

$$\lim_{z\to 1}[(1-z)^2A^{-1}(z)]=[I\ \ 0]\,J^{-1}\begin{bmatrix}I\\ 0\end{bmatrix}= \tag{1.560}$$

$$= [\boldsymbol{I} \;\; \boldsymbol{0}]\begin{bmatrix} \frac{1}{2}\ddot{\boldsymbol{A}} - \dot{\boldsymbol{A}}\boldsymbol{A}^{g}\dot{\boldsymbol{A}} & (\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp} \\ (\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp} & \boldsymbol{0} \end{bmatrix}^{-1}\begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} = [\boldsymbol{I} \;\; \boldsymbol{0}]\begin{bmatrix} \boldsymbol{J}_1 & \boldsymbol{J}_2 \\ \boldsymbol{J}_3 & \boldsymbol{J}_4 \end{bmatrix}\begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} = \boldsymbol{J}_1$$

This, in view of Definition 5 of Sect. 1.6, leads to conclude that $z = 1$ is a second order pole of $\boldsymbol{A}^{-1}(z)$.

The proof of ($\tilde{\text{b}}$1) rests on Theorem 7 in Sect. 1.6.

□

Corollary 5.1

Each of the following statements are equivalent to (1.537″), under (1.537) and (1.537′),

ã2) $$det(\boldsymbol{R}'_{\perp}\boldsymbol{B}'_{\perp}\widetilde{\boldsymbol{A}}\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp}) \neq 0 \tag{1.561}$$

ã3) $$r\left(\begin{bmatrix} \widetilde{\boldsymbol{A}} & \boldsymbol{A} + \boldsymbol{A}_l^{\perp}\dot{\boldsymbol{A}}\boldsymbol{A}_r^{\perp} \\ \boldsymbol{A} + \boldsymbol{A}_l^{\perp}\dot{\boldsymbol{A}}\boldsymbol{A}_r^{\perp} & \boldsymbol{0} \end{bmatrix}\right) = n + r(\boldsymbol{A}) + r(\boldsymbol{A}_l^{\perp}\dot{\boldsymbol{A}}\boldsymbol{A}_r^{\perp}) \tag{1.562}$$

where $\boldsymbol{A}_l^{\perp}$ and $\boldsymbol{A}_r^{\perp}$ are defined as in (1.94) and (1.95) of Sect. 1.2.

Proof

Equivalence of (ã2) and (1.537″) follows from Theorem 3 of Sect. 1.4, given that Theorem 1 of Sect. 1.2 applies, bearing in mind (1.32) of Sect. 1.2.

Equivalence of ã3) and ã1) can be easily proved following an argument similar to that of Corollary 3.1. First observe that

$$\begin{aligned} det&\left(\begin{bmatrix} \widetilde{\boldsymbol{A}} & (\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp} \\ (\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp} & \boldsymbol{0} \end{bmatrix}\right) \neq 0 \Leftrightarrow r\left(\begin{bmatrix} \widetilde{\boldsymbol{A}} & (\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp} \\ (\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp} & \boldsymbol{0} \end{bmatrix}\right) \\ &= r\,(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})_{\perp} + r\,(\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp} \\ &+ r\,\{[\boldsymbol{I} - (\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp}\,(\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp}^{-}]\widetilde{\boldsymbol{A}}\,[\boldsymbol{I} - ((\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp})^{-}\,(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp}]\} \\ &= r\,(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})_{\perp} + r\,(\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp} + r\,(\boldsymbol{R}'_{\perp}\boldsymbol{B}'_{\perp}\widetilde{\boldsymbol{A}}\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp}) \\ &= r\,(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})_{\perp} + r\,(\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp} + r\,(\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp}) = r\,(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})_{\perp} + n \end{aligned} \tag{1.563}$$

in light of Theorem 19 in Marsaglia and Styan, and of identities (1.87) and (1.88) of Sect. 1.2. Then, notice that

$$
\begin{aligned}
r\,(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})_{\perp} &= r\,((\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp}\,(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp}) = r\,(\boldsymbol{A} + (\boldsymbol{B}'_{\perp})^{g}\,\boldsymbol{RS}'\,\boldsymbol{C}^{g}_{\perp}) \\
&= r\,(\boldsymbol{A} + \boldsymbol{A}^{\perp}_{l}\dot{\boldsymbol{A}}\boldsymbol{A}^{\perp}_{r}) = r\,(\boldsymbol{A}) + r\,(\boldsymbol{A}^{\perp}_{l}\dot{\boldsymbol{A}}\boldsymbol{A}^{\perp}_{r})
\end{aligned}
\tag{1.564}
$$

in light of (1.538), (1.554) above, (1.94) and (1.95) of Sect. 1.2 together with Theorem 14 in Marsaglia and Styan, and also that

$$
\begin{aligned}
& r\left(\begin{bmatrix} \widetilde{\boldsymbol{A}} & (\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp} \\ (\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp} & \boldsymbol{0} \end{bmatrix}\right) \\
&= r\left\{\begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & (\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp} \end{bmatrix}\begin{bmatrix} \widetilde{\boldsymbol{A}} & (\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp} \\ (\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp} & \boldsymbol{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & (\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp} \end{bmatrix}\right\} \\
&= r\left(\begin{bmatrix} \widetilde{\boldsymbol{A}} & (\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp}(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp} \\ (\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp}(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp} & \boldsymbol{0} \end{bmatrix}\right) \\
&= r\left(\begin{bmatrix} \widetilde{\boldsymbol{A}} & \boldsymbol{A} + (\boldsymbol{B}'_{\perp})^{g}\boldsymbol{RS}'\boldsymbol{C}^{g}_{\perp} \\ \boldsymbol{A} + (\boldsymbol{B}'_{\perp})^{g}\boldsymbol{RS}'\boldsymbol{C}^{g}_{\perp} & \boldsymbol{0} \end{bmatrix}\right) \\
&= r\left(\begin{bmatrix} \widetilde{\boldsymbol{A}} & \boldsymbol{A} + \boldsymbol{A}^{\perp}_{l}\dot{\boldsymbol{A}}\boldsymbol{A}^{\perp}_{r} \\ \boldsymbol{A} + \boldsymbol{A}^{\perp}_{l}\dot{\boldsymbol{A}}\boldsymbol{A}^{\perp}_{r} & \boldsymbol{0} \end{bmatrix}\right)
\end{aligned}
\tag{1.565}
$$

This, together with (1.563) proves the claimed equivalence.

□

1.11 Closed-Forms of Laurent Expansion Coefficient Matrices. First Approach

In this section closed-form expressions for the matrices of Laurent expansions of matrix-polynomial inverse $\boldsymbol{A}^{-1}(z)$ about a simple and a second order pole located at $z = 1$, are derived from the arguments proposed in Sects. 1.7 and 1.10.

We also present a collection of useful properties and worthwhile relationships, as by-products of the main results, which pave the way to obtaining special expansions with either truncated or annihilated principal parts via pole order-reduction or removal.

Notation and matrix qualifications of the Sect. 1.10 apply unless otherwise stated.

The simple pole case is dealt with in the following

Theorem 1

Let the inverse, $A^{-1}(z)$*, of the matrix polynomial*

$$\boldsymbol{A}(z) = (1 - z)\,\boldsymbol{Q}(z) + \boldsymbol{A} \tag{1.566}$$

have a simple pole located at z = 1, *so that the Laurent expansion*

$$\boldsymbol{A}^{-1}(z) = \frac{1}{(1-z)}\boldsymbol{N}_1 + \boldsymbol{M}(z) \tag{1.567}$$

applies in a deleted neighbourhood of z = 1.

Then the following closed-form representations hold for $\boldsymbol{N}_1$ *and* $\boldsymbol{M}(1)$ *respectively*

$$\boldsymbol{N}_1 = \mathfrak{I}_1 = -\,\boldsymbol{C}_\perp(\boldsymbol{B}'_\perp\,\dot{\boldsymbol{A}}\,\boldsymbol{C}_\perp)^{-1}\boldsymbol{B}'_\perp \tag{1.568}$$

$$\begin{aligned} \boldsymbol{M}(1) &= -\frac{1}{2}\,\mathfrak{I}_1\,\ddot{\boldsymbol{A}}\,\mathfrak{I}_1 + \mathfrak{I}_2\,\mathfrak{I}_3 \\ &= -\frac{1}{2}\boldsymbol{N}_1\,\ddot{\boldsymbol{A}}\,\boldsymbol{N}_1 + (\boldsymbol{I} + \boldsymbol{N}_1\dot{\boldsymbol{A}})\,\boldsymbol{A}^g(\boldsymbol{I} + \dot{\boldsymbol{A}}\boldsymbol{N}_1) \end{aligned} \tag{1.569}$$

where

$$\begin{bmatrix} \mathfrak{I}_1 & \mathfrak{I}_2 \\ \mathfrak{I}_3 & \mathfrak{I}_4 \end{bmatrix} = \begin{bmatrix} -\dot{\boldsymbol{A}} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix}^{-1} \tag{1.570}$$

Proof

In view of (1.307) of Sect. 1.7, the matrix $\boldsymbol{N}_1$ is given by

$$\boldsymbol{N}_1 = \lim_{z\to 1}\,[(1 - z)\,\boldsymbol{A}^{-1}(z)] \tag{1.571}$$

which, by virtue of (1.516) of Sect. 1.10, can be written as

$$\boldsymbol{N}_1 = [\boldsymbol{I}\;\;\boldsymbol{0}]\begin{bmatrix} -\dot{\boldsymbol{A}} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix}^{-1}\begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} = \mathfrak{I}_1 \tag{1.572}$$

whence, according to the partitioned inversion (1.114) of Theorem 3, Sect. 1.4, the elegant closed-form solution

$$\boldsymbol{N}_1 = -\,\boldsymbol{C}_\perp(\boldsymbol{B}'_\perp\,\dot{\boldsymbol{A}}\,\boldsymbol{C}_\perp)^{-1}\boldsymbol{B}'_\perp \tag{1.573}$$

ensues.

In view of (1.308) of Sect. 1.7 the matrix $\boldsymbol{M}(1)$ is given by

$$\boldsymbol{M}(1) = -\lim_{z\to 1} \frac{\partial[(1-z)\boldsymbol{A}^{-1}(z)]}{\partial z} \tag{1.574}$$

which, in light of (1.515) of Sect. 1.10, can be rewritten as

$$\boldsymbol{M}(1) = -\lim_{z\to 1} \frac{\partial}{\partial z} \begin{bmatrix} \boldsymbol{Q}(z) & \boldsymbol{B} \\ \boldsymbol{C}' & (z-1)\boldsymbol{I} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} \tag{1.575}$$

Differentiating, taking the limit as z tends to 1, making use of the aforementioned partitioned inversion formula and bearing in mind (1.258) and (1.259) of Sect. 1.6, simple computations lead to

$$\begin{aligned} &\boldsymbol{M}(1) = \\ &= \lim_{z\to 1}\left([\boldsymbol{I} \quad \boldsymbol{0}] \begin{bmatrix} \boldsymbol{Q}(z) & \boldsymbol{B} \\ \boldsymbol{C}' & (z-1)\boldsymbol{I} \end{bmatrix}^{-1} \begin{bmatrix} \dot{\boldsymbol{Q}}(z) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{Q}(z) & \boldsymbol{B} \\ \boldsymbol{C}' & (z-1)\boldsymbol{I} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} \right) \\ &= [\boldsymbol{I} \quad \boldsymbol{0}] \begin{bmatrix} -\dot{\boldsymbol{A}} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix}^{-1} \begin{bmatrix} -\frac{1}{2}\ddot{\boldsymbol{A}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} -\dot{\boldsymbol{A}} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} \\ &= [\mathfrak{I}_1, \ \mathfrak{I}_2] \begin{bmatrix} -\frac{1}{2}\ddot{\boldsymbol{A}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \mathfrak{I}_1 \\ \mathfrak{I}_3 \end{bmatrix} \\ &= [\boldsymbol{N}_1, \ (\boldsymbol{I} + \boldsymbol{N}_1\dot{\boldsymbol{A}})(\boldsymbol{C}')^g] \begin{bmatrix} -\frac{1}{2}\ddot{\boldsymbol{A}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{N}_1 \\ \boldsymbol{B}^g(\boldsymbol{I} + \dot{\boldsymbol{A}}\boldsymbol{N}_1) \end{bmatrix} \\ &= -\frac{1}{2}\boldsymbol{N}_1\ddot{\boldsymbol{A}}\boldsymbol{N}_1 + (\boldsymbol{I} + \boldsymbol{N}_1\dot{\boldsymbol{A}})\boldsymbol{A}^g(\boldsymbol{I} + \dot{\boldsymbol{A}}\boldsymbol{N}_1) \end{aligned} \tag{1.576}$$

which completes the proof.

□

Corollary 1.1

The following hold true

(1) $$r(\boldsymbol{N}_1) = n - r(\boldsymbol{A}) \tag{1.577}$$

(2) The null row-space of $\boldsymbol{N}_1$ is spanned by the columns of an arbitrary orthogonal complement of $\boldsymbol{C}_\perp$, which can be conveniently chosen as $\boldsymbol{C}$

(3) The matrix function

$$\boldsymbol{C}'\boldsymbol{A}^{-1}(z) = \boldsymbol{C}'\boldsymbol{M}(z) \tag{1.578}$$

is analytic at $z = 1$

Proof

The proof of (1) rests on the rank equality

$$r(\boldsymbol{N}_1) = r(\boldsymbol{C}_\perp) = n - r(\boldsymbol{C}) = n - r(\boldsymbol{A}) \tag{1.579}$$

which is implied by the rank factorization and orthogonal complement rank properties.

As far as (2) is concerned, observe that the columns of $\boldsymbol{C}$ form a basis for the null row-space of $\boldsymbol{N}_1$ as per Corollary 1.2 of Sect. 1.7 upon noticing that $\boldsymbol{E} = (\boldsymbol{B}'_\perp \dot{\boldsymbol{A}} \boldsymbol{C}_\perp)^{-1}$ in light of (1.568) above.

Statement (3) is a verbatim repetition of Corollary 1.3 of Sect. 1.7.

□

Corollary 1.2

The following statements hold

(1)
$$tr(\boldsymbol{N}_1 \dot{\boldsymbol{A}}) = r(\boldsymbol{A}) - n \tag{1.580}$$

(2)
$$\boldsymbol{N}_1 \dot{\boldsymbol{A}} \boldsymbol{N}_1 = -\boldsymbol{N}_1 \Rightarrow \dot{\boldsymbol{A}} = -\boldsymbol{N}_1^- \tag{1.581}$$

(3)
$$\begin{cases} \boldsymbol{M}(1)\boldsymbol{B} = (\boldsymbol{I} + \boldsymbol{N}_1\dot{\boldsymbol{A}})(\boldsymbol{C}')^g = \mathfrak{I}_2 \\ \boldsymbol{C}'\boldsymbol{M}(1) = \boldsymbol{B}^g(\boldsymbol{I} + \dot{\boldsymbol{A}}\boldsymbol{N}_1) = \mathfrak{I}_3 \end{cases} \tag{1.582}$$

(4)
$$\boldsymbol{A}\boldsymbol{M}(1)\boldsymbol{A} = \boldsymbol{A} \Rightarrow \boldsymbol{M}(1) = \boldsymbol{A}^- \tag{1.583}$$

Proof

Proof of (1) follows by checking that

$$\begin{aligned} tr(\boldsymbol{N}_1 \dot{\boldsymbol{A}}) &= -\,tr(\boldsymbol{C}_\perp (\boldsymbol{B}'_\perp \dot{\boldsymbol{A}} \boldsymbol{C}_\perp)^{-1} \boldsymbol{B}'_\perp \dot{\boldsymbol{A}}) \\ &= -\,tr((\boldsymbol{B}'_\perp \dot{\boldsymbol{A}} \boldsymbol{C}_\perp)^{-1} (\boldsymbol{B}'_\perp \dot{\boldsymbol{A}} \boldsymbol{C}_\perp)) = -\,tr(\boldsymbol{I}_{n-r}) = r - n \end{aligned} \tag{1.584}$$

where r is written for $r(\boldsymbol{A})$.

Proof of (2) follows from (1.568) by simple computation and from Definition 1 of Sect. 1.1.

Proof of (3) follows from (1.569) by simple computation in view of (1.25) of Sect. 1.1.

Proof of (4) follows from (1.576), in view of (1.568), by straightforward computations.

□

The next theorem deals with the case of a second order pole.

Theorem 2

Let the inverse, $\boldsymbol{A}^{-1}(z)$, of the matrix polynomial

$$\boldsymbol{A}(z) = (1-z)^2\,\boldsymbol{\Psi}(z) - (1-z)\,\dot{\boldsymbol{A}} + \boldsymbol{A} \tag{1.585}$$

have a second order pole located at $z = 1$*, so that the Laurent expansion*

$$\boldsymbol{A}^{-1}(z) = \frac{1}{(1-z)^2}\boldsymbol{N}_2 + \frac{1}{(1-z)}\boldsymbol{N}_1 + \boldsymbol{M}(z) \tag{1.586}$$

applies in a deleted neighbourhood of $z = 1$.

Then the following closed-form representations hold for $\boldsymbol{N}_2$ and $\boldsymbol{N}_1$

$$\boldsymbol{N}_2 = \boldsymbol{J}_1 = \boldsymbol{C}_\perp\boldsymbol{S}_\perp(\boldsymbol{R}'_\perp\,\boldsymbol{B}'_\perp\,\widetilde{\boldsymbol{A}}\,\boldsymbol{C}_\perp\boldsymbol{S}_\perp)^{-1}\boldsymbol{R}'_\perp\,\boldsymbol{B}'_\perp \tag{1.587}$$

$$\boldsymbol{N}_1 = \boldsymbol{J}_1\,\widetilde{\widetilde{\boldsymbol{A}}}\,\boldsymbol{J}_1 + \boldsymbol{J}_2\boldsymbol{U}_2\boldsymbol{B}^g\,\dot{\boldsymbol{A}}\,\boldsymbol{J}_1 + \boldsymbol{J}_1\,\dot{\boldsymbol{A}}\,(\boldsymbol{C}')^g\,\boldsymbol{U}'_2\,\boldsymbol{J}_3 - \boldsymbol{J}_2\boldsymbol{U}_1\,\boldsymbol{U}'_1\,\boldsymbol{J}_3 \tag{1.588}$$

$$= [\boldsymbol{N}_2,\ \ \boldsymbol{I} - \boldsymbol{N}_2\widetilde{\boldsymbol{A}}]\begin{bmatrix} \widetilde{\widetilde{\boldsymbol{A}}} & \dot{\boldsymbol{A}}\boldsymbol{A}^g \\ \boldsymbol{A}^g\dot{\boldsymbol{A}} & -\boldsymbol{C}_\perp(\boldsymbol{B}'_\perp\dot{\boldsymbol{A}}\boldsymbol{C}_\perp)^g\boldsymbol{B}'_\perp \end{bmatrix}\begin{bmatrix} \boldsymbol{N}_2 \\ \boldsymbol{I} - \widetilde{\boldsymbol{A}}\boldsymbol{N}_2 \end{bmatrix} \tag{1.588$'$}$$

where

$$\widetilde{\boldsymbol{A}} = \frac{1}{2}\ddot{\boldsymbol{A}} - \dot{\boldsymbol{A}}\,\boldsymbol{A}^g\,\dot{\boldsymbol{A}} \tag{1.589}$$

$$\widetilde{\widetilde{\boldsymbol{A}}} = \frac{1}{6}\dddot{\boldsymbol{A}} - \dot{\boldsymbol{A}}\,\boldsymbol{A}^g\,\dot{\boldsymbol{A}}\,\boldsymbol{A}^g\,\dot{\boldsymbol{A}} \tag{1.590}$$

$$\begin{bmatrix} J_1 & J_2 \\ J_3 & J_4 \end{bmatrix} = \begin{bmatrix} \widetilde{A} & (B_\perp R_\perp)_\perp \\ (C_\perp S_\perp)'_\perp & 0 \end{bmatrix}^{-1} \tag{1.591}$$

Here, according to Theorem 2 of Sect. 1.2, $(B_\perp R_\perp)_\perp = [B,\ (B'_\perp)^g R]$ *and* $(C_\perp S_\perp)_\perp = [C,\ (C'_\perp)^g S]$*, and*

$$U_1 = \begin{bmatrix} 0 \\ I \end{bmatrix}, \quad U_2 = \begin{bmatrix} I \\ 0 \end{bmatrix} = U_{1\perp} \tag{1.592}$$

are selection matrices such that

$$(B_\perp R_\perp)_\perp U_2 = B, \ \ (B_\perp R_\perp)_\perp U_1 = (B'_\perp)^g R, \tag{1.593}$$

Proof

In view of (1.319) of Sect. 1.7, the matrix N_2 is given by

$$N_2 = \lim_{z\to 1} [(1-z)^2 A^{-1}(z)] \tag{1.594}$$

which, in light of (1.560) of Sect. 1.10, can be expressed as

$$N_2 = [I \quad 0] \begin{bmatrix} \frac{1}{2}\ddot{A} - \dot{A}A^g\dot{A} & (B_\perp R_\perp)_\perp \\ (C_\perp S_\perp)'_\perp & 0 \end{bmatrix}^{-1} \begin{bmatrix} I \\ 0 \end{bmatrix} \tag{1.595}$$

whence, according to the partitioned inversion formula (1.114) of Theorem 3, Sect. 1.4, the elegant closed form

$$N_2 = C_\perp S_\perp (R'_\perp B'_\perp \widetilde{A} C_\perp S_\perp)^{-1} R'_\perp B'_\perp \tag{1.596}$$

ensues.

In view of (1.320) of Sect. 1.7, the matrix N_1 is given by

$$N_1 = -\lim_{z\to 1} \frac{\partial [(1-z)^2 A^{-1}(z)]}{\partial z} \tag{1.597}$$

which, in light of (1.546) and (1.548) of Sect. 1.10, can be written as

$$N_1 = -\lim_{z\to 1} \frac{\partial [(1-z)^2 \Psi(z) - FV(z)G']^{-1}}{\partial z} \tag{1.598}$$

$$= -\lim_{z \to 1} \frac{\partial}{\partial z} \left\{ [\boldsymbol{I} \;\; \boldsymbol{0}] \begin{bmatrix} \boldsymbol{\Psi}(z) & \boldsymbol{F} \\ \boldsymbol{G}' & \boldsymbol{\Lambda}(z) \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} \right\}$$

Before differentiating and taking the limit, we can resort to (1.549) of Sect. 1.10 and consequently write

$$[\boldsymbol{I} \;\; \boldsymbol{0}] \begin{bmatrix} \boldsymbol{\Psi}(z) & \boldsymbol{F} \\ \boldsymbol{G}' & \boldsymbol{\Lambda}(z) \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix}$$

$$= [\boldsymbol{I} \quad \boldsymbol{0} \;\; \vdots \;\; \boldsymbol{0}] \begin{bmatrix} \boldsymbol{\Psi}(z) & (\boldsymbol{B}_\perp \boldsymbol{R}_\perp)_\perp & \vdots & \dot{\boldsymbol{A}}(\boldsymbol{C}')^g \\ & & \vdots & \\ (\boldsymbol{C}_\perp \boldsymbol{S}_\perp)'_\perp & (1-z)\boldsymbol{U}_1\boldsymbol{U}'_1 & \vdots & (1-z)\boldsymbol{U}_2 \\ \cdots & \cdots\;\cdots\cdots\cdots & \cdots & \\ \boldsymbol{B}^g\dot{\boldsymbol{A}} & (1-z)\boldsymbol{U}'_2 & \vdots & \boldsymbol{\Theta}(z) \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \\ \cdots \\ \boldsymbol{0} \end{bmatrix} \tag{1.599}$$

where

$$\Theta(z) = (1-z)\, \boldsymbol{B}^g \dot{\boldsymbol{A}}\, (\boldsymbol{C}')^g + \boldsymbol{I} \tag{1.600}$$

Since $\Theta(z)$ is non-singular in a neighbourhood of $z = 1$, partitioned inversion formula (1.105) of Theorem 1, Sect. 1.4, applies and therefore

$$\begin{aligned} (1-z)^2\boldsymbol{A}^{-1}(z) &= [\boldsymbol{I} \quad \boldsymbol{0}] \left\{ \begin{bmatrix} \boldsymbol{\Psi}(z) & (\boldsymbol{B}_\perp \boldsymbol{R}_\perp)_\perp \\ (\boldsymbol{C}_\perp \boldsymbol{S}_\perp)'_\perp & (1-z)\boldsymbol{U}_1\boldsymbol{U}'_1 \end{bmatrix} + \right. \\ &\left. - \begin{bmatrix} \dot{\boldsymbol{A}}(\boldsymbol{C}')^g \\ (1-z)\boldsymbol{U}_2 \end{bmatrix} \boldsymbol{\Theta}^{-1}(z) [\boldsymbol{B}^g \dot{\boldsymbol{A}}, (1-z)\boldsymbol{U}'_2] \right\}^{-1} \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} \\ &= [\boldsymbol{I} \quad \boldsymbol{0}]\, \{\boldsymbol{\Omega}_0(z) + (1-z)\, \boldsymbol{\Omega}_1(z) + (1-z)^2\, \boldsymbol{\Omega}_2(z)\}^{-1} \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} \\ &= [\boldsymbol{I} \quad \boldsymbol{0}]\, \boldsymbol{\Omega}^{-1}(z) \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} \end{aligned} \tag{1.601}$$

where

$$\boldsymbol{\Omega}(z) = \boldsymbol{\Omega}_0(z) + (1-z)\, \boldsymbol{\Omega}_1(z) + (1-z)^2\, \boldsymbol{\Omega}_2(z) \tag{1.602}$$

$$\boldsymbol{\Omega}_0(z) = \begin{bmatrix} \boldsymbol{\Psi}(z) - \dot{\boldsymbol{A}}(\boldsymbol{C}')^g \boldsymbol{\Theta}^{-1}(z) \boldsymbol{B}^g \dot{\boldsymbol{A}} & (\boldsymbol{B}_\perp \boldsymbol{R}_\perp)_\perp \\ (\boldsymbol{C}_\perp \boldsymbol{S}_\perp)'_\perp & \boldsymbol{0} \end{bmatrix} \tag{1.603}$$

$$\boldsymbol{\Omega}_1(z) = \begin{bmatrix} \boldsymbol{0} & -\dot{\boldsymbol{A}}(\boldsymbol{C}')^g \boldsymbol{\Theta}^{-1}(z)\boldsymbol{U}_2' \\ -\boldsymbol{U}_2\boldsymbol{\Theta}^{-1}(z)\boldsymbol{B}^g\dot{\boldsymbol{A}} & \boldsymbol{U}_1\boldsymbol{U}_1' \end{bmatrix} \tag{1.604}$$

$$\boldsymbol{\Omega}_2(z) = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{U}_2\boldsymbol{\Theta}^{-1}(z)\boldsymbol{U}_2' \end{bmatrix} \tag{1.605}$$

In particular, we have

$$\boldsymbol{\Omega}(1) = \boldsymbol{\Omega}_0(1) = \begin{bmatrix} \tilde{\boldsymbol{A}} & (\boldsymbol{B}_\perp \boldsymbol{R}_\perp)_\perp \\ (\boldsymbol{C}_\perp \boldsymbol{S}_\perp)_\perp' & \boldsymbol{0} \end{bmatrix} \tag{1.606}$$

Differentiating both sides of (1.601) yields

$$\frac{\partial[(1-z)^2 \boldsymbol{A}^{-1}(z)}{\partial z} = [\boldsymbol{I} \quad \boldsymbol{0}]\, \boldsymbol{\Omega}^{-1}(z)\, \dot{\boldsymbol{\Omega}}(z)\, \boldsymbol{\Omega}^{-1}(z) \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} \tag{1.607}$$

Now, observe that

$$\dot{\boldsymbol{\Omega}}(z) = \dot{\boldsymbol{\Omega}}_0(z) - \boldsymbol{\Omega}_1(z) + \textit{terms in } (1-z) \textit{ and } (1-z)^2 \tag{1.608}$$

and therefore by simple computations it follows that

$$\dot{\boldsymbol{\Omega}}(1) = \dot{\boldsymbol{\Omega}}_0(1) - \boldsymbol{\Omega}_1(1) = \begin{bmatrix} \tilde{\tilde{\boldsymbol{A}}} & \dot{\boldsymbol{A}}(\boldsymbol{C}')^g \boldsymbol{U}_2' \\ \boldsymbol{U}_2\boldsymbol{B}^g\dot{\boldsymbol{A}} & -\boldsymbol{U}_1\boldsymbol{U}_1' \end{bmatrix} \tag{1.609}$$

because of

$$\dot{\boldsymbol{\Psi}}(1) = \frac{1}{6}\dddot{\boldsymbol{A}} \tag{1.610}$$

$$\dot{\boldsymbol{\Theta}}^{-1}(1) = -\boldsymbol{\Theta}^{-1}(1)\, \dot{\boldsymbol{\Theta}}(1)\, \boldsymbol{\Theta}^{-1}(1) = -\dot{\boldsymbol{\Theta}}(1) = \boldsymbol{B}^g \dot{\boldsymbol{A}}(\boldsymbol{C}')^g \tag{1.611}$$

where the symbols $\dot{\boldsymbol{\Omega}}_0(1)$, $\dot{\boldsymbol{\Psi}}(1)$, $\dot{\boldsymbol{\Theta}}(1)$ indicate the derivatives of the matrices $\boldsymbol{\Omega}_0(z)$, $\boldsymbol{\Psi}(z)$ and $\boldsymbol{\Theta}(z)$ evaluated at $z = 1$.

Combining (1.597) with (1.607) and making use of (1.606) and (1.609), the matrix $\boldsymbol{N}_1$ turns out to be expressible in the following form

$$\boldsymbol{N}_1 = [\boldsymbol{I} \quad \boldsymbol{0}] \begin{bmatrix} \tilde{\boldsymbol{A}} & (\boldsymbol{B}_\perp \boldsymbol{R}_\perp)_\perp \\ (\boldsymbol{C}_\perp \boldsymbol{S}_\perp)_\perp' & \boldsymbol{0} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\tilde{\boldsymbol{A}}} & \dot{\boldsymbol{A}}(\boldsymbol{C}')^g \boldsymbol{U}_2' \\ \boldsymbol{U}_2\boldsymbol{B}^g\dot{\boldsymbol{A}} & -\boldsymbol{U}_1\boldsymbol{U}_1' \end{bmatrix}. \tag{1.612}$$

$$\cdot \begin{bmatrix} \widetilde{A} & (B_\perp R_\perp)_\perp \\ (C_\perp S_\perp)'_\perp & 0 \end{bmatrix}^{-1} \begin{bmatrix} I \\ 0 \end{bmatrix}$$

Now, one can verify that

$$\begin{aligned}[I \quad 0] \begin{bmatrix} \widetilde{A} & (B_\perp R_\perp)_\perp \\ (C_\perp S_\perp)'_\perp & 0 \end{bmatrix}^{-1} &= [N_2 \,, (I - N_2 \widetilde{A})(C_\perp S_\perp)'^{g}_\perp] \\ &= [N_2 \;, I - N_2 \widetilde{A}] \begin{bmatrix} I & 0 \\ 0 & (C_\perp S_\perp)'^{g}_\perp \end{bmatrix} \end{aligned} \tag{1.613}$$

$$\begin{aligned}\begin{bmatrix} \widetilde{A} & (B_\perp R_\perp)_\perp \\ (C_\perp S_\perp)'_\perp & 0 \end{bmatrix}^{-1} \begin{bmatrix} I \\ 0 \end{bmatrix} &= \begin{bmatrix} N_2 \\ (B_\perp R_\perp)^{g}_\perp (I - \widetilde{A} N_2) \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ 0 & (B_\perp R_\perp)^{g}_\perp \end{bmatrix} \begin{bmatrix} N_2 \\ I - \widetilde{A} N_2 \end{bmatrix} \end{aligned} \tag{1.614}$$

in view of the partitioned inversion formula (1.114) of Theorem 3, Sect. 1.4.

Besides, resorting to Theorem 3 of Sect. 1.2 the following results are easy to verify

$$((C_\perp S_\perp)'_\perp)^g U_2 = (C')^g, \quad U'_2 (B_\perp R_\perp)^g_\perp = B^g \tag{1.615}$$

$$((C_\perp S_\perp)'_\perp)^g U_1 = (S' C^g_\perp)^g \tag{1.616}$$

$$U'_1 \, (B_\perp R_\perp)^g_\perp \; = \; ((B'_\perp)^g R)^g \tag{1.617}$$

Now define

$$\Pi \; = \; R'_\perp B'_\perp \widetilde{A} C_\perp \tag{1.618}$$

$$\Sigma = (C'_\perp C_\perp)^{-1} S [S' (C'_\perp C_\perp)^{-1} S]^{-1} \tag{1.619}$$

Then, bearing in mind (1.616) above and making use of (1.53) of Sect. 1.2, simple computations show that

$$\begin{aligned}(I - N_2 \widetilde{A})((C_\perp S_\perp)'_\perp)^g U_1 &= (I - N_2 \widetilde{A})\, C_\perp \Sigma \\ = C_\perp [I - S_\perp (\Pi S_\perp)^{-1} \Pi] \Sigma &= C_\perp \Pi_\perp (S' \Pi_\perp)^{-1} S' (S')^g S' \,\Sigma \end{aligned} \tag{1.620}$$

$$= \boldsymbol{C}_{\perp}\boldsymbol{\Pi}_{\perp}(\boldsymbol{S}'\boldsymbol{\Pi}_{\perp})^{-1}\boldsymbol{S}'(\boldsymbol{S}')^{g} = \boldsymbol{C}_{\perp}[\boldsymbol{I}-\boldsymbol{S}_{\perp}(\boldsymbol{\Pi}\boldsymbol{S}_{\perp})^{-1}\boldsymbol{\Pi}](\boldsymbol{S}')^{g}$$
$$= (\boldsymbol{I}-\boldsymbol{N}_2\widetilde{\boldsymbol{A}})\,\boldsymbol{C}_{\perp}(\boldsymbol{S}')^{g}$$

The equality

$$\boldsymbol{U}_1'\,(\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp}^{g}\,(\boldsymbol{I}-\widetilde{\boldsymbol{A}}\boldsymbol{N}_2) = \boldsymbol{R}^{g}\,\boldsymbol{B}_{\perp}'\,(\boldsymbol{I}-\widetilde{\boldsymbol{A}}\boldsymbol{N}_2) \tag{1.621}$$

proves true likewise.

In view of the foregoing, after proper substitutions and some computations, the following closed-form expression for $\boldsymbol{N}_1$ is obtained

$$\boldsymbol{N}_1 = [\boldsymbol{N}_2,\ \ \boldsymbol{I}-\boldsymbol{N}_2\widetilde{\boldsymbol{A}}]\begin{bmatrix} \widetilde{\widetilde{\boldsymbol{A}}} & \dot{\boldsymbol{A}}\boldsymbol{A}^{g} \\ \boldsymbol{A}^{g}\dot{\boldsymbol{A}} & -\boldsymbol{C}_{\perp}(\boldsymbol{S}')^{g}\boldsymbol{R}^{g}\boldsymbol{B}_{\perp}' \end{bmatrix}\begin{bmatrix} \boldsymbol{N}_2 \\ \boldsymbol{I}-\widetilde{\boldsymbol{A}}\boldsymbol{N}_2 \end{bmatrix} \tag{1.622}$$

which eventually leads to (1.588′), upon noting that

$$\boldsymbol{C}_{\perp}(\boldsymbol{S}')^{g}\boldsymbol{R}^{g}\boldsymbol{B}_{\perp}' = \boldsymbol{C}_{\perp}(\boldsymbol{R}\boldsymbol{S}')^{g}\,\boldsymbol{B}_{\perp}' = \boldsymbol{C}_{\perp}\,(\boldsymbol{B}_{\perp}'\dot{\boldsymbol{A}}\boldsymbol{C}_{\perp})^{g}\boldsymbol{B}_{\perp}' \tag{1.623}$$

and bearing in mind (1.538) of Sect. 1.10. The closed-form (1.588) of $\boldsymbol{N}_1$ can then be drawn from (1.612) in light of (1.591).

□

Corollary 2.1

Under υ = 2 the following hold

(1)
$$r\,(\boldsymbol{N}_2) = n - r\,(\boldsymbol{A}) - r\,(\boldsymbol{B}_{\perp}'\dot{\boldsymbol{A}}\boldsymbol{C}_{\perp}) \tag{1.624}$$

(2)
$$r\,[\boldsymbol{N}_2, \boldsymbol{N}_1] = n - r\,(\boldsymbol{A}) + r\,(\dot{\boldsymbol{A}}\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp}) \tag{1.625}$$

(3) The null row-space of the block matrix $[\boldsymbol{N}_2, \boldsymbol{N}_1]$ is spanned by the columns of the matrix $\boldsymbol{C}(\boldsymbol{B}^{g}\boldsymbol{V})_{\perp}$, where $\boldsymbol{V}$ is defined by the rank factorization (1.341) of Sect. 1.7.

(4) The null row-space of $\boldsymbol{N}_2$ is spanned by the columns of an arbitrary orthogonal complement of $\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp}$ which can conveniently be chosen as

$$(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})_{\perp} = [\boldsymbol{C}(\boldsymbol{B}^{g}\boldsymbol{V})_{\perp},\ \boldsymbol{C}\boldsymbol{B}^{g}\boldsymbol{V},\ (\boldsymbol{C}_{\perp}')^{g}\boldsymbol{S}] \tag{1.626}$$

The former block in the right-hand side provides a spanning set for the null row-space of $[\boldsymbol{N}_2, \boldsymbol{N}_1]$, as stated above, whereas the latter blocks provide the completion spanning set for the null row-space of $\boldsymbol{N}_2$.

(5) The matrix function

$$(\boldsymbol{B}^g \boldsymbol{V})'_{\perp} \boldsymbol{C}' \boldsymbol{A}^{-1}(z) = (\boldsymbol{B}^g \boldsymbol{V})'_{\perp} \boldsymbol{C}' \boldsymbol{M}(z) \tag{1.627}$$

is analytic at $z = 1$.

(6) The matrix function

$$\begin{aligned} &[\boldsymbol{C}\boldsymbol{B}^g \boldsymbol{V},\ (\boldsymbol{C}'_{\perp})^g \boldsymbol{S}]' \boldsymbol{A}^{-1}(z) \\ &= \frac{1}{(1-z)}[\boldsymbol{C}\boldsymbol{B}^g \boldsymbol{V},\ (\boldsymbol{C}'_{\perp})^g \boldsymbol{S}]' \boldsymbol{N}_1 + [\boldsymbol{C}\boldsymbol{B}^g \boldsymbol{V},\ (\boldsymbol{C}'_{\perp})^g \boldsymbol{S}]' \boldsymbol{M}(z) \end{aligned} \tag{1.628}$$

exhibits a simple pole at $z = 1$.

(7) The matrix function

$$\boldsymbol{S}'_{\perp} \boldsymbol{C}'_{\perp} \boldsymbol{A}^{-1}(z) = \frac{1}{(1-z)^2} \boldsymbol{S}'_{\perp} \boldsymbol{C}'_{\perp} \boldsymbol{N}_2 + \frac{1}{(1-z)} \boldsymbol{S}'_{\perp} \boldsymbol{C}'_{\perp} \boldsymbol{N}_1 + \boldsymbol{S}'_{\perp} \boldsymbol{C}'_{\perp} \boldsymbol{M}(z) \tag{1.629}$$

exhibits a second order pole at $z = 1$.

Proof

Proof of (1) It is easy to check that

$$\begin{aligned} r(\boldsymbol{N}_2) &= r(\boldsymbol{R}'_{\perp} \boldsymbol{B}'_{\perp}) = r(\boldsymbol{R}_{\perp}) = r(\boldsymbol{B}_{\perp}) - r(\boldsymbol{B}'_{\perp} \dot{\boldsymbol{A}} \boldsymbol{C}_{\perp}) \\ &= n - r(\boldsymbol{A}) - r(\boldsymbol{B}'_{\perp} \dot{\boldsymbol{A}} \boldsymbol{C}_{\perp}) \end{aligned} \tag{1.630}$$

Proof of (2) Pre and post-multiplying (1.338) of Sect. 1.7 and (1.622) above by $\boldsymbol{S}'(\boldsymbol{C}_{\perp})^g$ and $(\boldsymbol{B}'_{\perp})^g$ respectively, we get the twin results

$$\begin{aligned} \boldsymbol{S}'(\boldsymbol{C}_{\perp})^g \boldsymbol{N}_1 (\boldsymbol{B}'_{\perp})^g &= \boldsymbol{S}'(\boldsymbol{C}_{\perp})^g (\boldsymbol{A}^g \dot{\boldsymbol{A}} \boldsymbol{N}_2 + \boldsymbol{N}_2 \dot{\boldsymbol{A}} \boldsymbol{A}^g + \boldsymbol{C}_{\perp} \boldsymbol{T} \boldsymbol{B}'_{\perp})(\boldsymbol{B}'_{\perp})^g = \\ &= \boldsymbol{S}' \boldsymbol{T} \end{aligned} \tag{1.631}$$

$$\begin{aligned} &\boldsymbol{S}'(\boldsymbol{C}_{\perp})^g \boldsymbol{N}_1 (\boldsymbol{B}'_{\perp})^g = \\ &= \boldsymbol{S}'(\boldsymbol{C}_{\perp})^g [\boldsymbol{N}_2\ ,\ \boldsymbol{I} - \boldsymbol{N}_2] \begin{bmatrix} \tilde{\tilde{\boldsymbol{A}}} & \dot{\boldsymbol{A}} \boldsymbol{A}^g \\ \boldsymbol{A}^g \dot{\boldsymbol{A}} & -\boldsymbol{C}_{\perp} (\boldsymbol{S}')^g \boldsymbol{R}^g \boldsymbol{B}'_{\perp} \end{bmatrix}. \end{aligned} \tag{1.632}$$

$$\cdot \begin{bmatrix} N_2 \\ I - \tilde{A} N_2 \end{bmatrix} (B'_\perp)^g = -R^g [I - B'_\perp \tilde{A} N_2 (B'_\perp)^g] = -(K'_\perp R)^{-1} K'_\perp$$

where K is written for $B'_\perp \tilde{A} C_\perp S_\perp$, and use has been made of Theorem 5 of Sect. 1.2 and of the orthogonality relationships

$$S' C_\perp^g N_2 = 0 \tag{1.633}$$

$$C_\perp^g A^g = 0 \tag{1.634}$$

Equating the right-hand sides of (1.631) and (1.632) yields

$$S'T = -(K'_\perp R)^{-1} K'_\perp \tag{1.635}$$

Then, in order to establish (1.625) we can follow the argument used in the proof of Corollary 3.2 of Sect. 1.7, with $\begin{bmatrix} \Lambda' W R'_\perp \\ S'T \end{bmatrix}$ replaced by $\begin{bmatrix} \Lambda'(R'_\perp K)^{-1} R'_\perp \\ -(K'_\perp R)^{-1} K'_\perp \end{bmatrix}$, and verify that the latter matrix is of full row-rank.

Indeed the following holds

$$\begin{aligned} & r\left(\begin{bmatrix} \Lambda'(R'_\perp K)^{-1} R'_\perp \\ -(K'_\perp R)^{-1} K'_\perp \end{bmatrix}\right) = \\ & r((K'_\perp R)^{-1} K'_\perp) + r(\Lambda'(R'_\perp K)^{-1} R'_\perp \{I - [(K'_\perp R)^{-1} K'_\perp]^g (K'_\perp R)^{-1} K'_\perp\}) \\ & = r(R) + r(\Lambda'(R'_\perp K)^{-1} R'_\perp K K^g) = r(R) + r(\Lambda' K^g) = r(R) + r(\Lambda') \end{aligned} \tag{1.636}$$

thanks to Theorem 19 of Marsaglia and Styan and to the equalities (see Theorem 4 of Sect. 1.1 and Theorem 4 of Sect. 1.2),

$$[(K'_\perp R)^{-1} K'_\perp]^g = (K'_\perp)^g K'_\perp R \tag{1.637}$$

$$\{I - [(K'_\perp R)^{-1} K'_\perp]^g (K'_\perp R)^{-1} K'_\perp\} = I - (K'_\perp)^g K'_\perp = K K^g \tag{1.638}$$

Given (1.636), formula (1.347) of Sect. 1.7 applies yielding (1.625).

Proof of (3) The columns of $C(B^g V)_\perp$ form a basis for the null row-space of $[N_2, N_1]$ according to Corollary 3.2 of Sect. 1.7 in light of the foregoing result.

Proof of (4) The columns of $(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})_{\perp}$ form a basis for the null row-space of $\boldsymbol{N}_2$ according to Corollary 2.2 of Sect. 1.7 as $\boldsymbol{W} = (\boldsymbol{R}'_{\perp}\boldsymbol{B}'_{\perp}\widetilde{\boldsymbol{A}}\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})^{-1}$ by (1.587) above.

As far as the matrix $(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})_{\perp}$ is concerned, it can be split as shown in (1.626) upon resorting to (1.47) of Sect. 1.2 and replacing $\boldsymbol{A}_{\perp}$, $\boldsymbol{\Psi}$, $\boldsymbol{B}_{\perp}$ and $\boldsymbol{\Omega}$ by $\boldsymbol{C}$, $[(\boldsymbol{B}^g\boldsymbol{V})_{\perp}, \boldsymbol{B}^g\boldsymbol{V}]$, $\boldsymbol{S}$ and $\boldsymbol{I}$, respectively.

The columns of the block $\boldsymbol{C}[(\boldsymbol{B}^g\boldsymbol{V})_{\perp}$form a basis for the null row-space of $[\boldsymbol{N}_2, \boldsymbol{N}_1]$ according to (3) above, whereas all the columns $(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})_{\perp}$ are involved to form a basis for the null row-space of $\boldsymbol{N}_2$, whence the completion role played by the columns of $[\boldsymbol{C}\boldsymbol{B}^g\boldsymbol{V}, (\boldsymbol{C}'_{\perp})^g\boldsymbol{S}]$.

Proof of (5) The statement is a verbatim repetition of Corollary 3.3 of Sect. 1.7.

Proof of (6) The proof mirrors that of Corollary 2.3 of Sect. 1.7.

Indeed, in light of (1.587) and (1.626), by premultiplying the right-hand side of (1.586) by $[\boldsymbol{C}\boldsymbol{B}^g\boldsymbol{V}, (\boldsymbol{C}'_{\perp})^g\boldsymbol{S}]'$, the term in $(1-z)^{-2}$ vanishes, thus reducing the order of the pole located at $z=1$ and making $[\boldsymbol{C}\boldsymbol{B}^g\boldsymbol{V}, (\boldsymbol{C}'_{\perp})^g\boldsymbol{S}]'\boldsymbol{A}^{-1}(z)$ to be a function with a simple pole at $z = 1$.

Proof of (7) In light of (1.587) by premultiplying the right-hand side of (1.586) by $(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'$, the term in $(1-z)^{-2}$ does not vanish and the function $(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'\boldsymbol{A}^{-1}(z)$ is characterized by a second order pole at $z = 1$ accordingly.

□

Corollary 2.2

The following statements hold true

(1) $$tr\,(\boldsymbol{N}_2\dot{\boldsymbol{A}}) = 0 \tag{1.639}$$

(2) $$\boldsymbol{N}_2\widetilde{\boldsymbol{A}}\boldsymbol{N}_2 = \boldsymbol{N}_2 \;\Rightarrow\; \widetilde{\boldsymbol{A}} = \boldsymbol{N}_2^- \tag{1.640}$$

(3) $$(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp}\boldsymbol{N}_1(\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp} = -\boldsymbol{U}_1\boldsymbol{U}'_1 \tag{1.641}$$

where $\boldsymbol{U}_1$ is defined as in (1.592) while $(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})_{\perp}$ and $(\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp}$ are written for as $[\boldsymbol{C}, (\boldsymbol{C}'_{\perp})^g\boldsymbol{S}]$ and $[\boldsymbol{B}, (\boldsymbol{B}'_{\perp})^g\boldsymbol{R}]$, respectively.

(4) $$\boldsymbol{A}\boldsymbol{N}_1\boldsymbol{A} = \boldsymbol{0} \tag{1.642}$$

(5) $$\boldsymbol{A}\boldsymbol{M}(1)\,\boldsymbol{A} = \boldsymbol{A} + \dot{\boldsymbol{A}}\boldsymbol{N}_2\dot{\boldsymbol{A}} \tag{1.643}$$

(6) $$\boldsymbol{A}(\boldsymbol{M}(1) - \boldsymbol{N}_1 \boldsymbol{N}_2^g \boldsymbol{N}_1)\boldsymbol{A} = \boldsymbol{A} \Rightarrow \boldsymbol{M}(1) - \boldsymbol{N}_1 \boldsymbol{N}_2^g \boldsymbol{N}_1 = \boldsymbol{A}^- \tag{1.644}$$

Proof

Point (1): the proof follows by checking that

$$\begin{aligned} tr\,(\boldsymbol{N}_2 \dot{\boldsymbol{A}}) &= tr\,(\boldsymbol{C}_\perp \boldsymbol{S}_\perp (\boldsymbol{R}'_\perp \boldsymbol{B}'_\perp \widetilde{\boldsymbol{A}}\, \boldsymbol{C}_\perp \boldsymbol{S}_\perp)^{-1} \boldsymbol{R}'_\perp \boldsymbol{B}'_\perp \dot{\boldsymbol{A}}) \\ &= tr\,((\boldsymbol{R}'_\perp \boldsymbol{B}'_\perp \widetilde{\boldsymbol{A}}\, \boldsymbol{C}_\perp \boldsymbol{S}_\perp)^{-1} \boldsymbol{R}'_\perp \boldsymbol{B}'_\perp \dot{\boldsymbol{A}}\, \boldsymbol{C}_\perp \boldsymbol{S}_\perp) \\ &= tr\,((\boldsymbol{R}'_\perp \boldsymbol{B}'_\perp \widetilde{\boldsymbol{A}}\, \boldsymbol{C}_\perp \boldsymbol{S}_\perp)^{-1} \boldsymbol{R}'_\perp \boldsymbol{R}\boldsymbol{S}' \boldsymbol{S}_\perp) = tr\,(\boldsymbol{0}) = 0 \end{aligned} \tag{1.645}$$

in light of (1.520) of Sect. 1.10 (see also Theorem 5 of Sect. 1.6).

Point (2): the proof ensues from (1.596) after elementary computations. The implication concerning the generalized inverse is trivial.

Point (3): the result follows from (1.612) upon noting that

$$(\boldsymbol{C}_\perp \boldsymbol{S}_\perp)'_\perp [\boldsymbol{I} \quad \boldsymbol{0}] = [\boldsymbol{0} \;\; \boldsymbol{I}] \begin{bmatrix} \widetilde{\boldsymbol{A}} & (\boldsymbol{B}_\perp \boldsymbol{R}_\perp)_\perp \\ (\boldsymbol{C}_\perp \boldsymbol{S}_\perp)'_\perp & \boldsymbol{0} \end{bmatrix} \tag{1.646}$$

$$\begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} (\boldsymbol{B}_\perp \boldsymbol{R}_\perp)_\perp = \begin{bmatrix} \widetilde{\boldsymbol{A}} & (\boldsymbol{B}_\perp \boldsymbol{R}_\perp)_\perp \\ (\boldsymbol{C}_\perp \boldsymbol{S}_\perp)'_\perp & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{I} \end{bmatrix} \tag{1.647}$$

Point (4): the result is obvious in view of (1.326) and (1.327) of Sect. 1.7.

To prove the result (5), first observe that the following identity

$$\boldsymbol{A}(z)\boldsymbol{A}^{-1}(z)\boldsymbol{A}(z) = \boldsymbol{A}(z) \tag{1.648}$$

is trivially satisfied in a deleted neighbourhood of $z = 1$. Then, substituting the right-hand side of (1.585) and (1.586) for $\boldsymbol{A}(z)$ and $\boldsymbol{A}^{-1}(z)$ in (1.648), making use of (1.326), (1.327) of Sect. 1.7 together with (1.642) above, after simple computations we obtain

$$\begin{aligned} -\dot{\boldsymbol{A}} \boldsymbol{N}_2 \dot{\boldsymbol{A}} + \boldsymbol{A}\boldsymbol{M}(z)\boldsymbol{A} &+ \textit{terms in positive powers of}\,(1-z) \\ &= \boldsymbol{A} - (1-z)\dot{\boldsymbol{A}} + (1-z)^2 \boldsymbol{\Psi}(z) \end{aligned} \tag{1.649}$$

Expanding $\boldsymbol{M}(z)$ about $z = 1$, that is to say

$$\boldsymbol{M}(z) = \boldsymbol{M}(1) + \textit{terms in positive powers of}\,(1-z) \tag{1.650}$$

and substituting the right-hand side of (1.650) for $\boldsymbol{M}(z)$ in (1.649), collecting equal powers and equating term by term, we at the end get for the constant terms the equality

$$-\dot{A}\,N_2\,\dot{A}\;+AM\,(1)\,A=A \tag{1.651}$$

Point (6): because of (1.326) and (1.327) of Sect. 1.7, identity (1.643) can be restated as follows

$$AM\,(1)\,A=A+\;\dot{A}N_2\;N_2^gN_2\;\dot{A}\;=A+A\,N_1N_2^gN_1\,A \tag{1.652}$$

which, rearranging the terms, eventually leads to (1.644).

□

1.12 Closed-Forms of Laurent Expansion Coefficient Matrices. Second Approach

In this section closed-form expressions for the coefficients of a matrix-polynomial inverse $A^{-1}(z)$ about a pole located at $z = 1$, are obtained in the wake of the analytical framework of Sect. 1.9.

Let us first establish the following results which provide explicit solutions for the principal-part matrices of the inverse of a linear polynomial about a unit root.

Lemma 1

Consider a linear matrix polynomial

$$A(z)=I+z\,A_1 \tag{1.653}$$

such that $A = I + A_1$ is as in Theorem 1 of Sect. 1.5, with the non-unit roots of the characteristic polynomial $detA(z)$ all lying outside the unit circle.

The inverse function $A^{-1}(z)$ can be expanded in a deleted neighborhood of $z = 1$ as

$$A^{-1}(z)=\sum_{i=1}^{\upsilon}\frac{1}{(1-z)^i}N_i+M(z) \tag{1.654}$$

where the matrix polynomial $M(z)=\sum_{i=0}^{\infty}M_iz^i$ is characterized by coefficient matrices M_i with exponentially decreasing entries, and where the principal-part coefficient matrices N_υ and $N_{\upsilon-1}$ can be expressed as

$$N_\upsilon=(-1)^{\upsilon-1}\,NH^{\upsilon-1} \tag{1.655}$$

$$N_{\upsilon-1} = (-1)^{\upsilon-2} NH^{\upsilon-2} + (1-\upsilon)N_{\upsilon} \tag{1.656}$$

in terms of the idempotent matrix

$$N = I - (A^{\upsilon})^{\neq} A^{\upsilon} = C_{\upsilon\perp}(B'_{\upsilon\perp}C_{\upsilon\perp})^{-1}B'_{\upsilon\perp} \tag{1.657}$$

and of the nilpotent matrix $\boldsymbol{H}$ of Theorem 1 of Sect. 1.5.

Proof

For the proof see Theorem 1 of Sect. 1.9.

□

Theorem 2

With $A(z)$ *as in Lemma* 1, *let*

$$ind(A) = 1 \tag{1.658}$$

Then

$$A^{-1}(z) = \frac{1}{(1-z)} N_1 + M(z) \tag{1.659}$$

with

$$N_1 = N = I - A^{\neq} A \tag{1.660}$$

$$= C_{\perp}(B'_{\perp}C_{\perp})^{-1}B'_{\perp} \tag{1.661}$$

$$M(1) = A^{\neq} \tag{1.662}$$

$$= (I - N)A^{-}_{\rho}(I - N) \tag{1.663}$$

$$= B(C'B)^{-2}C' \tag{1.664}$$

where $\boldsymbol{B}$ *and* $\boldsymbol{C}$ *are obtained from the rank factorization* $A = BC'$.

Proof

If $ind(A) = 1$, then $A^{-1}(z) = K^{-1}(z)$ from Theorem 1 of Sect. 1.9 with $H = 0$. As such $A^{-1}(z)$ turns out to have a simple pole at $z = 1$, the expansion (1.659) holds true with matrix residue N_1 which can be given the forms

(1.660) and (1.661) by straightforward application of formulas (1.655) and (1.657), upon taking $\boldsymbol{H}^0 = \boldsymbol{I}$ by convention.

As far as the closed forms (1.662)–(1.664) of $\boldsymbol{M}(1)$ are concerned, reference can be made to formulas (1.460) of Sect. 1.9, as well as (1.14′) and (1.27) of Sect. 1.1.

□

Corollary 2.1

Under the same assumptions as in Theorem 2 the following can be shown to be true

(1) $$r(\boldsymbol{N}_1) = n - r(\boldsymbol{A}), \tag{1.665}$$

(2) The null row-space of $\boldsymbol{N}_1$ is spanned by the columns of an arbitrary orthogonal complement of $\boldsymbol{C}_\perp$, which can conveniently be chosen as $\boldsymbol{C}$,

(3) The matrix function

$$\boldsymbol{C}'\boldsymbol{A}^{-1}(z) = \boldsymbol{C}'\boldsymbol{M}(z) \tag{1.666}$$

is analytic at $z = 1$

Proof

The proof is the same as that of Corollary 1.1 in the previous section.

□

Theorem 3

With $\boldsymbol{A}(z)$ as in Lemma 1, let

$$ind(\boldsymbol{A}) = 2 \tag{1.667}$$

Then

$$\boldsymbol{A}^{-1}(z) = \sum_{i=1}^{2} \frac{1}{(1-z)^i}\boldsymbol{N}_i + \boldsymbol{M}(z) \tag{1.668}$$

with

$$\boldsymbol{N}_2 = -\boldsymbol{N}\boldsymbol{A} \tag{1.669}$$

$$= -C_{\perp} S_{\perp} (R'_{\perp} B'_{\perp} A_{\rho}^{-} C_{\perp} S_{\perp})^{-1} R'_{\perp} B'_{\perp} \tag{1 669'}$$

$$= -BG_{\perp} (F'_{\perp} G_{\perp})^{-1} F'_{\perp} C \tag{1.669''}$$

$$N_1 = N - N_2 \tag{1.670}$$

$$= C_{\perp} T B'_{\perp} - A_{\rho}^{-} N_2 - N_2 A_{\rho}^{-} \tag{1.670'}$$

where

$$N = I - (A^2)^{\neq} A^2 \tag{1.671}$$

$$= C_{2\perp} (B'_{2\perp} C_{2\perp})^{-1} B'_{2\perp} \tag{1.671'}$$

$$T = \widetilde{T} - S_{\perp} L^{-1} R'_{\perp} \tag{1.672}$$

and the matrices $\boldsymbol{F}$, $\boldsymbol{G}$, $\boldsymbol{R}$, $\boldsymbol{S}$ *and the matrices* $\widetilde{\boldsymbol{T}}$, $\boldsymbol{L}$ *are defined as in Theorem* 8 *of Sect.* 1.2 *and in Corollary* 4.1 *of Sect.* 1.4, *respectively.*

Proof

If $ind(A) = 2$, then from Theorem 1 of Sect. 1.9 $A^{-1}(z)$ exhibits a second order pole at $z = 1$, whence the expansion (1.668), whose principal part matrices N_2 and N_1 can be written as

$$N_2 = -NH = -NA \tag{1.673}$$

$$N_1 = N - N_2 = N(I + H) = N(I + A) \tag{1.674}$$

by straightforward application of formulas (1.655) and (1.656), upon taking $H^0 = I$ and bearing in mind (1.184), (1.187) and (1.189) of Theorem 1 of Sect. 1.5. This proves (1.669) and (1.670).

For what concerns (1.669′) and (1.669″) observe that, upon resorting to (1.657) above and (1.240) of Sect. 1.5, (1.68) and (1.68') of Theorem 8 in Sect. 1.2, we obtain

$$\begin{aligned} N_2 &= -NA = -C_{2\perp} (B'_{2\perp} C_{2\perp})^{-1} B'_{2\perp} A = = -BG_{\perp} (F'_{\perp} G_{\perp})^{-1} F'_{\perp} C' \\ &= -C_{\perp} S_{\perp} (R'_{\perp} B'_{\perp} A_{\rho}^{-} C_{\perp} S_{\perp})^{-1} R'_{\perp} B'_{\perp} \end{aligned} \tag{1.675}$$

For what concerns (1.670) observe that, in light of (1.657) above, of Corollary 4.1 of Sect. 1.4 and upon noting that $\Delta = -\boldsymbol{N}_2$, we obtain

$$\begin{aligned}\boldsymbol{N}_1 = \boldsymbol{N} - \boldsymbol{N}_2 &= \boldsymbol{C}_\perp \tilde{\boldsymbol{T}} \boldsymbol{B}'_\perp - \boldsymbol{N}_2 \boldsymbol{A}_\rho^- - \boldsymbol{A}_\rho^- \boldsymbol{N}_2 - \boldsymbol{C}_\perp \boldsymbol{S}_\perp \boldsymbol{L}^{-1} \boldsymbol{R}'_\perp \boldsymbol{B}'_\perp \\ &= \boldsymbol{C}_\perp (\tilde{\boldsymbol{T}} - \boldsymbol{S}_\perp \boldsymbol{L}^{-1} \boldsymbol{R}'_\perp) \boldsymbol{B}'_\perp - \boldsymbol{N}_2 \boldsymbol{A}_\rho^- - \boldsymbol{A}_\rho^- \boldsymbol{N}_2 \\ &= \boldsymbol{C}_\perp \boldsymbol{T} \boldsymbol{B}'_\perp - \boldsymbol{N}_2 \boldsymbol{A}_\rho^- - \boldsymbol{A}_\rho^- \boldsymbol{N}_2\end{aligned} \tag{1.676}$$

□

Corollary 3.1

Under the same assumptions as in Theorem 3 the following prove to hold true

(1) $$r(\boldsymbol{N}_2) = r(\boldsymbol{A}) - r(\boldsymbol{A}^2) \tag{1.677}$$

(2) $$r([\boldsymbol{N}_2, \boldsymbol{N}_1]) = r(\boldsymbol{N}) \tag{1.678}$$

$$= r(\boldsymbol{N}_2) + n - r(\boldsymbol{A}), \tag{1.679}$$

(3) the null row-space of the block matrix $[\boldsymbol{N}_2, \boldsymbol{N}_1]$ is spanned by the columns of the matrix $\boldsymbol{C}(\boldsymbol{B}^g \boldsymbol{C}_\perp \boldsymbol{S}_\perp)_\perp$,

(4) The null row-space of $\boldsymbol{N}_2$ is spanned by the columns of an arbitrary orthogonal complement of $\boldsymbol{C}_\perp \boldsymbol{S}_\perp$ which can be conveniently chosen as

$$(\boldsymbol{C}_\perp \boldsymbol{S}_\perp)_\perp = [-\boldsymbol{C}(\boldsymbol{B}^g \boldsymbol{C}_\perp \boldsymbol{S}_\perp)_\perp,\ -\boldsymbol{C}\boldsymbol{B}^g \boldsymbol{C}_\perp \boldsymbol{S}_\perp,\ (\boldsymbol{C}'_\perp)^g \boldsymbol{S}] \tag{1.680}$$

The former block in the right-hand side provides a spanning set for the null row-space of $[\boldsymbol{N}_2, \boldsymbol{N}_1]$ as stated above, whereas the latter blocks provide the completion spanning set for the null row-space of $\boldsymbol{N}_2$.

(5) The matrix function

$$(\boldsymbol{B}^g \boldsymbol{C}_\perp \boldsymbol{S}_\perp)'_\perp \boldsymbol{C}' \boldsymbol{A}^{-1}(z) = (\boldsymbol{B}^g \boldsymbol{C}_\perp \boldsymbol{S}_\perp)'_\perp \boldsymbol{C}' \boldsymbol{M}(z) \tag{1.681}$$

is analytic at $z = 1$.

(6) The matrix function

$$[-\boldsymbol{C}\boldsymbol{B}^g \boldsymbol{C}_\perp \boldsymbol{S}_\perp, (\boldsymbol{C}'_\perp)^g \boldsymbol{S}]' \boldsymbol{A}^{-1}(z) = \tag{1.682}$$

$$= \frac{1}{(1-z)}[-\boldsymbol{C}\boldsymbol{B}^g\boldsymbol{C}_\perp\boldsymbol{S}_\perp, (\boldsymbol{C}'_\perp)^g\boldsymbol{S}]'\boldsymbol{N}_1 + [-\boldsymbol{C}\boldsymbol{B}^g\boldsymbol{C}_\perp\boldsymbol{S}_\perp, (\boldsymbol{C}'_\perp)^g\boldsymbol{S}]'\boldsymbol{M}(z)$$

exhibits a simple pole at $z = 1$.

(7) The matrix function

$$\boldsymbol{S}'_\perp\boldsymbol{C}'_\perp\boldsymbol{A}^{-1}(z) = \frac{1}{(1-z)^2}\boldsymbol{S}'_\perp\boldsymbol{C}'_\perp\boldsymbol{N}_2 + \frac{1}{(1-z)}\boldsymbol{S}'_\perp\boldsymbol{C}'_\perp\boldsymbol{N}_1 + \boldsymbol{S}'_\perp\boldsymbol{C}'_\perp\boldsymbol{M}(z) \tag{1.683}$$

exhibits a second order pole at $z = 1$.

Proof

Proof of (1) Observe that

$$r(\boldsymbol{N}_2) = r(\boldsymbol{R}'_\perp\boldsymbol{B}'_\perp) = r(\boldsymbol{R}_\perp) = r(\boldsymbol{B}_\perp) - r(\boldsymbol{B}'_\perp\boldsymbol{C}_\perp) = r(\boldsymbol{A}) - r(\boldsymbol{A}^2) \tag{1.684}$$

where the last identity ensues from Theorem 6 of Sect. 1.2.

Proof of (2). From (1.669) and (1.674) it follows that

$$[\boldsymbol{N}_2, \boldsymbol{N}_1] = [-\boldsymbol{N}\boldsymbol{A}, \boldsymbol{N}(\boldsymbol{I}+\boldsymbol{A})] = \boldsymbol{N}[-\boldsymbol{A}, (\boldsymbol{I}+\boldsymbol{A})] \tag{1.685}$$

which in turn entails the rank equality

$$r(\boldsymbol{N}[-\boldsymbol{A}, (\boldsymbol{I}+\boldsymbol{A})]) = r(\boldsymbol{N}) = n - r(\boldsymbol{A}^2) \tag{1.686}$$

in light of (1.671) above and (1.86’) of Sect. 1.2, upon noting that the matrix $[-\boldsymbol{A}, (\boldsymbol{I}+\boldsymbol{A})]$ is of full row-rank as straightforward application of Theorem 19 in Marsaglia and Styan, shows. Combining (1.686) with (1.677) eventually leads to (1.679).

Proof of (3) The null row-space of $[\boldsymbol{N}_2, \boldsymbol{N}_1]$ is the same as that of $\boldsymbol{N}$ in light of the proof of (2) and the latter space is spanned by the columns of $(\boldsymbol{C}_{2\perp})_\perp$ in light of (1.671′). By resorting to Theorem 8 together with Remark 3 of Sect. 1.2 the following representations of $\boldsymbol{C}_{2\perp}$ and $(\boldsymbol{C}_{2\perp})_\perp$ are easily established

$$\boldsymbol{C}_{2\perp} = [\boldsymbol{C}_\perp, \boldsymbol{A}_\rho^-\boldsymbol{C}_\perp\boldsymbol{S}_\perp] = [\boldsymbol{C}_\perp, (\boldsymbol{C}')^g\boldsymbol{B}^g\boldsymbol{C}_\perp\boldsymbol{S}_\perp] \tag{1.687}$$

$$(\boldsymbol{C}_{2\perp})_\perp = \boldsymbol{C}(\boldsymbol{B}^g\boldsymbol{C}_\perp\boldsymbol{S}_\perp)_\perp \tag{1.688}$$

with the right-hand side of (1.688) providing the required basis.

The proofs of (4)–(7) can be obtained by following essentially the same reasoning used to prove the analogous points of Corollary 2.1 of the

Sect. 1.11, with $-\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp}$ playing the role of $\boldsymbol{V}$ insofar as $\dot{\boldsymbol{A}} = (\boldsymbol{A} - \boldsymbol{I})$ and $\dot{\boldsymbol{A}}\,\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp} = -\,\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp} = \boldsymbol{V}$ (see formula (1.341) of Sect. 1.7).

□

Until now we have been concerned with linear matrix polynomials, however since higher-order matrix polynomials happen to be the rule in dynamic modelling, there is a need to provide a bridge between first and higher order polynomial analysis.

Fortunately, the linear polynomial mirror-image of an arbitrary order polynomial can be attained via companion-form and Schur complement algebraic arguments, and the road to provide an effective analytic support to dynamic model building, is already paved.

In this connection let us state the following.

Lemma 4

Consider a linear matrix polynomial specified as

$$\boldsymbol{A}(z) = \sum_{k=0}^{K} \boldsymbol{A}_k z^k, \qquad \boldsymbol{A}_0 = \boldsymbol{I}, \qquad \boldsymbol{A}_K \neq \boldsymbol{0} \tag{1.689}$$

where $\boldsymbol{A} = \sum_{k=0}^{K} \boldsymbol{A}_k$ is a non-null singular matrix of order n. Let the characteristic polynomial $det\boldsymbol{A}(z)$ have all its roots lying outside the unit circle, except for a possibly multiple unit-root, and let the companion matrix

$$\underset{(nK,nK)}{\breve{\boldsymbol{A}}} = \begin{bmatrix} \boldsymbol{I}_n + \boldsymbol{A}_1 & \vdots & \boldsymbol{A}_2 & \boldsymbol{A}_3 & .. & \boldsymbol{A}_K \\ \cdots & & \cdots & \cdots & \cdots & \cdots \\ -\boldsymbol{I}_n & \vdots & \boldsymbol{I}_n & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \vdots & -\boldsymbol{I}_n & \boldsymbol{I}_n & \boldsymbol{0} & \boldsymbol{0} \\ .. & \vdots & .. & .. & .. & .. \\ \boldsymbol{0} & \vdots & \boldsymbol{0} & \boldsymbol{0} & -\boldsymbol{I}_n & \boldsymbol{I}_n \end{bmatrix} \tag{1.690}$$

be of index υ.

Then, the inverse function $\boldsymbol{A}^{-1}(z)$ can be expanded in a deleted neighborhood of $z = 1$ as

$$\boldsymbol{A}^{-1}(z) = \sum_{i=1}^{\upsilon} \frac{1}{(1-z)^i} \boldsymbol{N}_i + \boldsymbol{M}(z) \tag{1.691}$$

Here the matrix polynomial $\boldsymbol{M}(z) = \sum_{i=0}^{\infty} \boldsymbol{M}_i z^i$ is characterized by coefficient matrices $\boldsymbol{M}_i$ with exponentially decreasing entries. The principal part coefficient matrices $\boldsymbol{N}_\upsilon$ and $\boldsymbol{N}_{\upsilon-1}$ can be expressed as

$$\boldsymbol{N}_\upsilon = (-1)^{\upsilon-1} \boldsymbol{J}'\breve{\boldsymbol{N}}\breve{\boldsymbol{H}}^{\upsilon-1}\boldsymbol{J} \tag{1.692}$$

$$\boldsymbol{N}_{\upsilon-1} = (-1)^{\upsilon-2} \boldsymbol{J}'\breve{\boldsymbol{N}}\breve{\boldsymbol{H}}^{\upsilon-2}\boldsymbol{J} + (1-\upsilon)\boldsymbol{J}'\breve{\boldsymbol{N}}_\upsilon\boldsymbol{J} \tag{1.693}$$

in terms of the nilpotent matrix $\breve{\boldsymbol{H}}$ in the decomposition of $\breve{\boldsymbol{A}}$ of Theorem 1 in Sect. 1.5, of the idempotent matrix

$$\breve{\boldsymbol{N}} = \boldsymbol{I} - (\breve{\boldsymbol{A}}^\upsilon)^{\neq} \breve{\boldsymbol{A}}^\upsilon \tag{1.694}$$

and of the selection matrix

$$\underset{(nK,n)}{\boldsymbol{J}} = \begin{bmatrix} \boldsymbol{I}_n \\ \boldsymbol{0} \\ \vdots \\ \boldsymbol{0} \end{bmatrix} \tag{1.695}$$

Proof

Consider the companion linear polynomial (see also (1.493) of Sect. 1.10),

$$\breve{\boldsymbol{A}}(z) = (1-z)\boldsymbol{I}_{nK} + z\breve{\boldsymbol{A}} = \boldsymbol{I}_{nK} + z(\breve{\boldsymbol{A}} - \boldsymbol{I}_{nK}) \tag{1.696}$$

which can be conveniently rewritten in partitioned form as follows

$$\breve{\boldsymbol{A}}(z) = \begin{bmatrix} \breve{\boldsymbol{A}}_{11}(z) & \breve{\boldsymbol{A}}_{12}(z) \\ \breve{\boldsymbol{A}}_{21}(z) & \breve{\boldsymbol{A}}_{22}(z) \end{bmatrix} =$$
$$= \left[\begin{array}{c:cccccc} \boldsymbol{I}_n + \boldsymbol{A}_1 z & \boldsymbol{A}_2 z & \boldsymbol{A}_3 z & \cdots\cdots & \boldsymbol{A}_{K-1} z & \boldsymbol{A}_K z \\ \cdots & \cdots & \cdots & & \cdots & \cdots \\ -\boldsymbol{I}_n z & \boldsymbol{I}_n & \boldsymbol{0} & & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I}_n z & \boldsymbol{I}_n & & \boldsymbol{0} & \boldsymbol{0} \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & & \cdots & \cdot\cdot \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & & -\boldsymbol{I}_n z & \boldsymbol{I}_n \end{array}\right] \tag{1.697}$$

Now consider the Schur complement of the lower diagonal block $\breve{\boldsymbol{A}}_{22}(z)$, that is

$$\breve{A}_{11}(z) - \breve{A}_{12}(z)\breve{A}_{22}^{-1}(z)\breve{A}_{21}(z) \tag{1.698}$$

and notice that

$$\breve{A}_{22}^{-1}(z) = \begin{bmatrix} I & 0 & \dots & 0 & 0 \\ Iz & I & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ Iz^{K-2} & Iz^{K-3} & \dots & Iz & I \end{bmatrix} \tag{1.699}$$

Now, bearing in mind (1.106) of Sect. 1.4, simple computations yield

$$\breve{A}_{11}(z) - \breve{A}_{12}(z)\breve{A}_{22}^{-1}(z)\breve{A}_{21}(z) = I + A_1(z) + A_2 z^2 + \dots + A_K z^K = A(z) \tag{1.700}$$

$$det\breve{A}(z) = det\breve{A}_{22}(z) det A(z) = det A(z) \tag{1.701}$$

and, straightforward application of partitioned inversion formula (1.105) of Theorem 1 in Sect. 1.4 leads to the result

$$J'\breve{A}^{-1}(z)J = A^{-1}(z) \tag{1.702}$$

Because of the assumptions made and of (1.701), Lemma 1 applies to the matrix polynomial $\breve{A}(z)$, and an expansion of $\breve{A}^{-1}(z)$ about $z = 1$ such as

$$\breve{A}^{-1}(z) = \sum_{i=1}^{\upsilon} \frac{1}{(1-z)^i} \breve{N}_i + \breve{M}(z) \tag{1.703}$$

holds true, with $\breve{M}(z)$, $\breve{N}_\upsilon$ and $\breve{N}_{\upsilon-1}$ specified accordingly.

This together with (1.702) leads to the conclusion that

$$\begin{aligned} A^{-1}(z) &= \sum_{i=1}^{\upsilon} \frac{1}{(1-z)^i} N_i + M(z) = J'\breve{A}^{-1}(z)J \\ &= \sum_{i=1}^{\upsilon} \frac{1}{(1-z)^i} J'\breve{N}_i J + J'\breve{M}(z)J \end{aligned} \tag{1.704}$$

which completes the proof.

□

Theorem 5

With A(z) as in Lemma 4*, let*

$$ind(\breve{\boldsymbol{A}}) = 1 \tag{1.705}$$

Then

$$\boldsymbol{A}^{-1}(z) = \frac{1}{(1-z)} \boldsymbol{N}_1 + \boldsymbol{M}(z) \tag{1.706}$$

with

$$\boldsymbol{N}_1 = -\boldsymbol{C}_\perp (\boldsymbol{B}'_\perp \dot{\boldsymbol{A}} \boldsymbol{C}_\perp)^{-1} \boldsymbol{B}'_\perp \tag{1.707}$$

$$\boldsymbol{M}(1) = -\frac{1}{2} \boldsymbol{N}_1 \ddot{\boldsymbol{A}} \boldsymbol{N}_1 + (\boldsymbol{I} + \boldsymbol{N}_1 \dot{\boldsymbol{A}}) \boldsymbol{A}^g (\boldsymbol{I} + \dot{\boldsymbol{A}} \boldsymbol{N}_1) \tag{1.708}$$

where ***B*** *and* ***C*** *are derived from the rank factorization of* ***A****, that is* $\boldsymbol{A} = \boldsymbol{BC}'$*.*

Proof

If $ind(\breve{\boldsymbol{A}}) = 1$, then Theorem 2 applies to $\breve{\boldsymbol{A}}^{-1}(z)$ defined as in (1.696), and Lemma 4 applies to $\boldsymbol{A}^{-1}(z)$ with $\upsilon = 1$ and $\boldsymbol{H} = \boldsymbol{0}$, which eventually entails that

$$\boldsymbol{N}_1 = \boldsymbol{J}'\breve{\boldsymbol{N}}_1\boldsymbol{J} = \boldsymbol{J}'\breve{\boldsymbol{N}}\boldsymbol{J} = \boldsymbol{J}'(\boldsymbol{I} - \breve{\boldsymbol{A}}^D \breve{\boldsymbol{A}})\boldsymbol{J} = \boldsymbol{J}'\breve{\boldsymbol{C}}_\perp (\breve{\boldsymbol{B}}'_\perp \breve{\boldsymbol{C}}_\perp)^{-1} \breve{\boldsymbol{B}}'_\perp \boldsymbol{J} \tag{1.709}$$

$$\begin{aligned} \boldsymbol{M}(1) &= \boldsymbol{J}'(\breve{\boldsymbol{I}} - \breve{\boldsymbol{N}}_1)\breve{\boldsymbol{A}}_\rho^-(\breve{\boldsymbol{I}} - \breve{\boldsymbol{N}}_1)\boldsymbol{J} \\ &= \boldsymbol{J}'\breve{\boldsymbol{A}}_\rho^-\boldsymbol{J} - \boldsymbol{J}'\breve{\boldsymbol{A}}_\rho^-\breve{\boldsymbol{N}}_1\boldsymbol{J} - \boldsymbol{J}'\breve{\boldsymbol{N}}_1\breve{\boldsymbol{A}}_\rho^-\boldsymbol{J} + \boldsymbol{J}'\breve{\boldsymbol{N}}_1\breve{\boldsymbol{A}}_\rho^-\breve{\boldsymbol{N}}_1\boldsymbol{J} \end{aligned} \tag{1.710}$$

where $\breve{\boldsymbol{B}}$ and $\breve{\boldsymbol{C}}$ come from the rank factorization of $\breve{\boldsymbol{A}}$, that is $\breve{\boldsymbol{A}} = \breve{\boldsymbol{B}}\breve{\boldsymbol{C}}'$. Then, making use of formulas (1.507)–(1.509) of Sect. 1.10, we get the closed form (1.707) for $\boldsymbol{N}_1$.

Furthermore, keeping in mind also formulas (1.527), (1.528), (1.528′), (1.529) and (1.530) of the same section, some computations show that

$$\boldsymbol{J}'\breve{\boldsymbol{A}}_\rho^-\boldsymbol{J} = \boldsymbol{A}^g, \tag{1.711}$$

$$\begin{aligned} \boldsymbol{J}'\breve{\boldsymbol{A}}_\rho^-\breve{\boldsymbol{N}}_1\boldsymbol{J} &= -[\boldsymbol{A}^g,\ -\boldsymbol{A}^g\breve{\boldsymbol{A}}_{12}\breve{\boldsymbol{A}}_{22}^{-1}] \begin{bmatrix} \boldsymbol{C}_\perp \\ (\boldsymbol{u}_{K-1} \otimes \boldsymbol{C}_\perp) \end{bmatrix} (\boldsymbol{B}'_\perp \dot{\boldsymbol{A}} \boldsymbol{C}_\perp)^{-1} \boldsymbol{B}'_\perp \\ &= -[\boldsymbol{A}^g\boldsymbol{C}_\perp, -\boldsymbol{A}^g(\dot{\boldsymbol{A}} + \boldsymbol{I})\boldsymbol{C}_\perp](\boldsymbol{B}'_\perp \dot{\boldsymbol{A}} \boldsymbol{C}_\perp)^{-1} \boldsymbol{B}'_\perp = \boldsymbol{A}^g \dot{\boldsymbol{A}} \boldsymbol{C}_\perp (\boldsymbol{B}'_\perp \dot{\boldsymbol{A}} \boldsymbol{C}_\perp)^{-1} \boldsymbol{B}'_\perp \\ &= -\boldsymbol{A}^g \dot{\boldsymbol{A}} \boldsymbol{N}_1 \end{aligned} \tag{1.712}$$

$$
\begin{aligned}
\boldsymbol{J}'\breve{\boldsymbol{N}}_1\breve{\boldsymbol{A}}_\rho^- \boldsymbol{J} &= -\boldsymbol{C}_\perp(\boldsymbol{B}'_\perp\dot{\boldsymbol{A}}\boldsymbol{C}_\perp)^{-1}[\boldsymbol{Y}_1, \boldsymbol{Y}_2]\begin{bmatrix} \boldsymbol{A}^g \\ -\breve{\boldsymbol{A}}_{22}^{-1}\breve{\boldsymbol{A}}_{21}\boldsymbol{A}^g \end{bmatrix} \\
&= -\boldsymbol{C}_\perp(\boldsymbol{B}'_\perp\dot{\boldsymbol{A}}\boldsymbol{C}_\perp)^{-1}[\breve{\boldsymbol{B}}'_\perp\boldsymbol{A}^g,\ -\breve{\boldsymbol{B}}'_\perp(\dot{\boldsymbol{A}}+\boldsymbol{I})\boldsymbol{A}^g] \\
&= \boldsymbol{C}_\perp(\boldsymbol{B}'_\perp\dot{\boldsymbol{A}}\boldsymbol{C}_\perp)^{-1}\boldsymbol{B}'_\perp\dot{\boldsymbol{A}}\boldsymbol{A}^g = -\boldsymbol{N}_1\dot{\boldsymbol{A}}\boldsymbol{A}^g
\end{aligned}
\tag{1.713}
$$

$$
\begin{aligned}
\boldsymbol{J}'\breve{\boldsymbol{N}}_1\breve{\boldsymbol{A}}_\rho^-\breve{\boldsymbol{N}}_1\boldsymbol{J} &= \boldsymbol{C}_\perp(\boldsymbol{B}'_\perp\dot{\boldsymbol{A}}\boldsymbol{C}_\perp)^{-1}\breve{\boldsymbol{B}}'_\perp\breve{\boldsymbol{A}}_\rho^-\breve{\boldsymbol{C}}_\perp(\boldsymbol{B}'_\perp\dot{\boldsymbol{A}}\boldsymbol{C}_\perp)^{-1}\boldsymbol{B}'_\perp \\
&= -\frac{1}{2}\boldsymbol{N}_1\ddot{\boldsymbol{A}}\boldsymbol{N}_1 + \boldsymbol{N}_1\dot{\boldsymbol{A}}\boldsymbol{A}^g\dot{\boldsymbol{A}}\boldsymbol{N}_1
\end{aligned}
\tag{1.714}
$$

Substituting the right hand-sides of (1.711)–(1.714) into the right-hand side of (1.710) eventually yields (1.708).

□

Corollary 5.1

Under the assumptions of Theorem 5 the following relations can be shown to be true

(1)
$$ r(\boldsymbol{N}_1) = n - r(\boldsymbol{A}), \tag{1.715} $$

(2) The null row-space of $\boldsymbol{N}_1$ is spanned by the columns of an arbitrary orthogonal complement of $\boldsymbol{C}_\perp$, which can conveniently be chosen as $\boldsymbol{C}$,

(3) The matrix function

$$ \boldsymbol{C}'\boldsymbol{A}^{-1}(z) = \boldsymbol{C}'M(z) \tag{1.716} $$

is analytic at $z = 1$

Proof

The proof is the same as that of the proof of Corollary 1.1 in the Sect. 1.11.

□

Theorem 6

With ***A****(z) as in Lemma* 4*, let*

$$ind(\breve{\boldsymbol{A}}) = 2 \tag{1.717}$$

Then

$$\boldsymbol{A}^{-1}(z) = \sum_{i=1}^{2} \frac{1}{(1-z)^i} \boldsymbol{N}_i + \boldsymbol{M}(z) \tag{1.718}$$

with

$$\boldsymbol{N}_2 = \boldsymbol{C}_\perp \boldsymbol{S}_\perp (\boldsymbol{R}'_\perp \boldsymbol{B}'_\perp \tilde{\boldsymbol{A}} \boldsymbol{C}_\perp \boldsymbol{S}_\perp)^{-1} \boldsymbol{R}'_\perp \boldsymbol{B}'_\perp \tag{1.719}$$

$$\boldsymbol{N}_1 = \boldsymbol{C}_\perp \boldsymbol{T} \boldsymbol{B}'_\perp + \boldsymbol{A}^g \dot{\boldsymbol{A}} \boldsymbol{N}_2 + \boldsymbol{N}_2 \dot{\boldsymbol{A}} \boldsymbol{A}^g \tag{1.720}$$

where

$$\tilde{\boldsymbol{A}} = \left(\frac{1}{2} \ddot{\boldsymbol{A}} - \dot{\boldsymbol{A}} \boldsymbol{A}^g \dot{\boldsymbol{A}} \right) \tag{1.721}$$

$$\boldsymbol{T} = \boldsymbol{\Theta} - \boldsymbol{S}_\perp \boldsymbol{L}^{-1} (\boldsymbol{L} + \boldsymbol{\aleph}) \boldsymbol{L}^{-1} \boldsymbol{R}'_\perp \tag{1.722}$$

R*,* ***S*** *are obtained by the rank factorization* $\boldsymbol{B}'_\perp \dot{\boldsymbol{A}} \boldsymbol{C}_\perp = \boldsymbol{R}\boldsymbol{S}'$*, and* $\boldsymbol{\Theta}$*,* ***L****, and* $\boldsymbol{\aleph}$ *are specified as in* (1.533)–(1.535) *of Sect.* 1.10.

Proof

If $ind(\breve{\boldsymbol{A}}) = 2$, then Theorem 3 applies to $\breve{\boldsymbol{A}}^{-1}(z)$ defined as in (1.696), and Lemma 4 applies to $\boldsymbol{A}^{-1}(z)$ with $\upsilon = 2$ and $\boldsymbol{H}^2 = \boldsymbol{0}$, which entails that

$$\boldsymbol{N}_2 = \boldsymbol{J}'\breve{\boldsymbol{N}}_2\boldsymbol{J} = -\boldsymbol{J}'\breve{\boldsymbol{C}}_\perp \breve{\boldsymbol{S}}_\perp (\breve{\boldsymbol{R}}'_\perp \breve{\boldsymbol{B}}'_\perp \breve{\boldsymbol{A}}_\rho^- \breve{\boldsymbol{C}}_\perp \breve{\boldsymbol{S}}_\perp)^{-1} \breve{\boldsymbol{R}}'_\perp \breve{\boldsymbol{B}}'_\perp \boldsymbol{J} \tag{1.723}$$

$$\boldsymbol{N}_1 = \boldsymbol{J}'\breve{\boldsymbol{N}}_1\boldsymbol{J} = \boldsymbol{J}'(\breve{\boldsymbol{C}}_\perp \boldsymbol{T} \breve{\boldsymbol{B}}_\perp - \breve{\boldsymbol{A}}_\rho^- \breve{\boldsymbol{N}}_2 - \breve{\boldsymbol{N}}_2 \breve{\boldsymbol{A}}_\rho^-)\boldsymbol{J} \tag{1.724}$$

where $\breve{\boldsymbol{N}}_2$ and $\breve{\boldsymbol{N}}_1$ are defined in accordance with Theorem 3 when applied to $\breve{\boldsymbol{A}}^{-1}(z)$, $\breve{\boldsymbol{R}}$ and $\breve{\boldsymbol{S}}$ are obtained from the rank factorization of $\breve{\boldsymbol{B}}'_\perp \breve{\boldsymbol{C}}_\perp = \breve{\boldsymbol{R}}\breve{\boldsymbol{S}}'$.

By making use of (1.507), (1.508), (1.526) and (1.534) of Sect. 1.10, formula (1.723) can be rewritten as in (1.719).

Moreover, bearing in mind (1.526), (1.527), (1.528) and (1.528') of Sect. 1.10, some computations show that

$$J'\breve{A}_{\rho}^{-}\breve{N}_2 J = -[A^g, \ -A^g \breve{A}_{12}\breve{A}_{22}^{-1}]\breve{C}_{\perp}\breve{S}_{\perp}(\breve{R}_{\perp}'\breve{B}_{\perp}'\breve{A}_{\rho}^{-}\breve{C}_{\perp}\breve{S}_{\perp})^{-1}\breve{R}_{\perp}'\breve{B}_{\perp}'J$$

$$= [A^g, \ -A^g \breve{A}_{12}\breve{A}_{22}^{-1}]\begin{bmatrix} C_{\perp} \\ (u_{K-1} \otimes C_{\perp}) \end{bmatrix} S_{\perp}(R_{\perp}'B_{\perp}'\widetilde{A}C_{\perp}S_{\perp})^{-1}R_{\perp}'B_{\perp}' \quad (1.725)$$

$$= [A^g C_{\perp}, -A^g(\dot{A}+I)C_{\perp}]S_{\perp}(R_{\perp}'B_{\perp}'\tilde{A}C_{\perp}S_{\perp})^{-1}R_{\perp}'B_{\perp}' = -A^g\dot{A}N_2$$

$$J'\breve{N}_1\breve{A}_{\rho}^{-}J = -C_{\perp}\breve{S}_{\perp}(\breve{R}_{\perp}'B_{\perp}'\breve{A}_{\rho}^{-}\breve{C}_{\perp}\breve{S}_{\perp})^{-1}\breve{R}_{\perp}'B_{\perp}'\begin{bmatrix} A^g \\ -\breve{A}_{22}^{-1}\breve{A}_{21}A^g \end{bmatrix} \quad (1.726)$$

$$= C_{\perp}S_{\perp}(R_{\perp}'B_{\perp}'\tilde{A}C_{\perp}S_{\perp})^{-1}R_{\perp}'[B_{\perp}'A^g, -B_{\perp}'(\dot{A}+I)A^g] = -N_2\dot{A}A^g$$

Substituting the right hand-sides of (1.725) and (1.726) into the right hand-side of (1.724) eventually yields (1.720).

□

Corollary 6.1

Under υ = 2 the following hold

(1) $$r(N_2) = n - r(A) - r(B_{\perp}'\dot{A}C_{\perp}) \quad (1.727)$$

(2) $$r([N_2, N_1]) = n - r(A) + r(\dot{A}C_{\perp}S_{\perp}) \quad (1.728)$$

(3) The null row-space of the block matrix $[N_2, N_1]$ is spanned by the columns of the matrix $C(B^g V)_{\perp}$, where V is defined by the rank factorization (1.341) of Sect. 1.7.

(4) The null row-space of N_2 is spanned by the columns of an arbitrary orthogonal complement of $C_{\perp}S_{\perp}$ which can conveniently be chosen as

$$(C_{\perp}S_{\perp})_{\perp} = [C(B^g V)_{\perp}, \ CB^g V, \ (C_{\perp}')^g S] \quad (1.729)$$

The former block in the right-hand side provides a spanning set for the null row-space of $[N_2, N_1]$, as stated above, whereas the latter blocks provide the completion spanning set for the null row-space of N_2.

(5) The matrix function

$$(\boldsymbol{B}^g \boldsymbol{V})'_{\perp} \boldsymbol{C}' \boldsymbol{A}^{-1}(z) = (\boldsymbol{B}^g \boldsymbol{V})'_{\perp} \boldsymbol{C}' \boldsymbol{M}(z) \tag{1.730}$$

is analytic at $z = 1$.

(6) The matrix function

$$\begin{aligned} &[\boldsymbol{C}\boldsymbol{B}^g \boldsymbol{V}, (\boldsymbol{C}'_{\perp})^g \boldsymbol{S}]' \boldsymbol{A}^{-1}(z) \\ &= \frac{1}{(1-z)} [\boldsymbol{C}\boldsymbol{B}^g \boldsymbol{V}, (\boldsymbol{C}'_{\perp})^g \boldsymbol{S}]' \boldsymbol{N}_1 + [\boldsymbol{C}\boldsymbol{B}^g \boldsymbol{V}, (\boldsymbol{C}'_{\perp})^g \boldsymbol{S}]' \boldsymbol{M}(z) \end{aligned} \tag{1.731}$$

exhibits a simple pole at $z = 1$.

(7) The matrix function

$$\boldsymbol{S}'_{\perp} \boldsymbol{C}'_{\perp} \boldsymbol{A}^{-1}(z) = \frac{1}{(1-z)^2} \boldsymbol{S}'_{\perp} \boldsymbol{C}'_{\perp} \boldsymbol{N}_2 + \frac{1}{(1-z)} \boldsymbol{S}'_{\perp} \boldsymbol{C}'_{\perp} \boldsymbol{N}_1 + \boldsymbol{S}'_{\perp} \boldsymbol{C}'_{\perp} \boldsymbol{M}(z) \tag{1.732}$$

exhibits a second order pole at $z = 1$.

Proof

Proof of (1). The proof is a verbatim repetition of that of point (1) in Corollary 2.1 of the Sect. 1.11.

Proof of (2). According to formula (1.346) of Sect. 1.7, the following holds

$$r[(\boldsymbol{N}_1, \boldsymbol{N}_2)] = r(\boldsymbol{S}_{\perp}) + r\left(\begin{bmatrix} \boldsymbol{\Lambda}' \boldsymbol{W} \boldsymbol{R}'_{\perp} \\ \boldsymbol{S}' \boldsymbol{T} \end{bmatrix}\right) \tag{1.733}$$

where

$$\boldsymbol{W} = (\boldsymbol{R}'_{\perp} \boldsymbol{B}'_{\perp} \widetilde{\boldsymbol{A}} \boldsymbol{C}_{\perp} \boldsymbol{S}_{\perp})^{-1} \tag{1.734}$$

the matrix $\boldsymbol{\Lambda}$ is obtained from the rank factorization

$$\dot{\boldsymbol{A}} \boldsymbol{C}_{\perp} \boldsymbol{S}_{\perp} = \boldsymbol{V} \boldsymbol{\Lambda}' \tag{1.735}$$

and

$$\boldsymbol{S}' \boldsymbol{T} = \boldsymbol{S}'[\boldsymbol{\Theta} - \boldsymbol{S}_{\perp} \boldsymbol{L}^{-1} (\boldsymbol{L} + \boldsymbol{\aleph}) \boldsymbol{L}^{-1} \boldsymbol{R}'_{\perp}] = \boldsymbol{S}' \boldsymbol{\Theta} = (\boldsymbol{K}'_{\perp} \boldsymbol{R})^{-1} \boldsymbol{K}'_{\perp} \tag{1.736}$$

with $\boldsymbol{\Theta}$ defined in (1.533) of Sect. 1.10.

Following the same argument as in the proof of (2) in Corollary 2.1 of Sect. 1.11, we can establish that $\begin{bmatrix} \boldsymbol{\Lambda}'\boldsymbol{W}\boldsymbol{R}'_{\perp} \\ \boldsymbol{S}'\boldsymbol{T} \end{bmatrix}$ is a full row-rank matrix and eventually conclude that

$$\begin{gathered} r[\boldsymbol{N}_1, \boldsymbol{N}_2] = r(\boldsymbol{S}_{\perp}) + r(\boldsymbol{\Lambda}'\boldsymbol{W}\boldsymbol{R}'_{\perp}) + r(\boldsymbol{S}'\boldsymbol{T}) \\ r(\boldsymbol{S}_{\perp}) + r(\boldsymbol{\Lambda}') + r(\boldsymbol{S}) = r(\boldsymbol{C}_{\perp}) + r(\dot{\boldsymbol{A}}\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp}) = n - r(\boldsymbol{A}) + r(\dot{\boldsymbol{A}}\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp}) \end{gathered} \tag{1.737}$$

Proof of (3)–(7). The proofs are verbatim repetitions of the proofs of the corresponding points of Corollary 2.1 of the Sect. 1.11.

□

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

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

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 $\text{WN}_{(n)}$, 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)

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

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

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

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

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

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

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

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

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

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

$$\boldsymbol{y}_t = \boldsymbol{\mu} + [\boldsymbol{M}_0, \boldsymbol{M}_1, ..., \boldsymbol{M}_q] \begin{bmatrix} \boldsymbol{\varepsilon}_t \\ \boldsymbol{\varepsilon}_{t-1} \\ \vdots \\ \boldsymbol{\varepsilon}_{t-q} \end{bmatrix} \tag{2.21}$$

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

$$\begin{aligned} \boldsymbol{\Gamma}(h) &= E\left\{ [\boldsymbol{M}_0, \boldsymbol{M}_1, ..., \boldsymbol{M}_q] \begin{bmatrix} \boldsymbol{\varepsilon}_t \\ \boldsymbol{\varepsilon}_{t-1} \\ \vdots \\ \boldsymbol{\varepsilon}_{t-q} \end{bmatrix} [\boldsymbol{\varepsilon}'_{t+h}, \boldsymbol{\varepsilon}'_{t-1+h}, ..., \boldsymbol{\varepsilon}'_{t-q+h}] \begin{bmatrix} \boldsymbol{M}'_0 \\ \boldsymbol{M}'_1 \\ \vdots \\ \boldsymbol{M}'_q \end{bmatrix} \right\} \\ &= [\boldsymbol{M}_0, \boldsymbol{M}_1, ..., \boldsymbol{M}_q]\, E\left\{ \begin{bmatrix} \boldsymbol{\varepsilon}_t \\ \boldsymbol{\varepsilon}_{t-1} \\ \vdots \\ \boldsymbol{\varepsilon}_{t-q} \end{bmatrix} [\boldsymbol{\varepsilon}'_{t+h}, \boldsymbol{\varepsilon}'_{t-1+h}, ..., \boldsymbol{\varepsilon}'_{t-q+h}] \right\} \begin{bmatrix} \boldsymbol{M}'_0 \\ \boldsymbol{M}'_1 \\ \vdots \\ \boldsymbol{M}'_q \end{bmatrix} \\ &= [\boldsymbol{M}_0, \boldsymbol{M}_1, ..., \boldsymbol{M}_q]\, (\boldsymbol{D}_h \otimes \boldsymbol{\Sigma}) \begin{bmatrix} \boldsymbol{M}'_0 \\ \boldsymbol{M}'_1 \\ \vdots \\ \boldsymbol{M}'_q \end{bmatrix} \end{aligned} \tag{2.22}$$

whence (2.20).

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

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

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

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

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 $\boldsymbol{y}_t$

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

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

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

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

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 $\boldsymbol{y}_t$

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

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

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 $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_p$, 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

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

which implies the relationships

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

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

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

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

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

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

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*

$$A\,(L)\, y_t = \boldsymbol{\eta} + \boldsymbol{\varepsilon}_t \tag{2.94}$$

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

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

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 $\delta \le d - 1$.

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

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

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

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

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

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

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

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

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

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

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

$$\boldsymbol{\xi}_t = \sum_{i\ge0} \boldsymbol{M}_i L^i \boldsymbol{\varepsilon}_t + \sum_{i\ge0} \boldsymbol{M}_i \boldsymbol{\eta} + \boldsymbol{N}_1 \nabla^{-1} \boldsymbol{\varepsilon}_t + \boldsymbol{N}_1 \nabla^{-1} \boldsymbol{\eta} + \boldsymbol{N}_2 \nabla^{-2} \boldsymbol{\varepsilon}_t + \boldsymbol{N}_2 \nabla^{2} \boldsymbol{\eta} + \cdots \quad (2.143)$$

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,

$$\boldsymbol{\xi}_t = \underset{\substack{i\ge0 \\ VMA\ process}}{\sum \boldsymbol{M}_i L^i \boldsymbol{\varepsilon}_t} + \underset{\substack{constant \\ term}}{\boldsymbol{a}_0} + \underset{\substack{1st-order \\ random\ walk}}{\boldsymbol{N}_1 \boldsymbol{\rho}_t(1)} + \underset{\substack{linear\ trend \\ in\ t}}{\boldsymbol{a}_1 t} + \underset{\substack{2nd-order \\ random\ walk}}{\boldsymbol{N}_2 \boldsymbol{\rho}_t(2)} + \underset{\substack{quadratic\ trend \\ in\ t}}{\boldsymbol{a}_2 t^2} + \ldots \quad (2.144)$$

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

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

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)

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

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

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

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

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

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

where

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

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

$$A(L) \underset{(n,1)}{y_t} = \mu + \varepsilon_t \,, \; A(L) = I_n + \tilde{A} L, \; n = mK \tag{2.152}$$

by resorting to the auxiliary vectors and (companion) matrix

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

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

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

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

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

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

$$= \begin{bmatrix} (-\tilde{A}_1 - \lambda I) & \vdots & -\tilde{A}_2 & -\tilde{A}_3 & ...+ & \tilde{A}_{K-1} & -\tilde{A}_K \\ \cdots & & \cdots & \cdots & & \cdots & \cdots \\ I_m & \vdots & -\lambda I_m & 0 & & 0 & 0 \\ 0 & \vdots & I_m & -\lambda I_m & & 0 & 0 \\ .. & \vdots & .. & .. & & \cdots & .. \\ 0 & \vdots & 0 & 0 & & I_m & -\lambda I_m \end{bmatrix} \tag{2.161}$$

and making use of the formula of Schur (see, e.g, Gantmacher vol I p 46) $det(A_1 - \lambda I)$ can be rewritten (provided $\lambda \neq 0$) as

$$det(A_1 - \lambda I) = detA_{22}(\lambda) det(A_{11}(\lambda) - A_{12}(\lambda) A_{22}^{-1}(\lambda) A_{21}(\lambda))$$

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

upon noting that

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

This, together with the fact that

$$detA_{22} = (det(-\lambda I_m))^{K-1} \propto \lambda^{m(K-1)} \tag{2.164}$$

eventually leads to express $det(A_1 - \lambda I)$ in this manner

$$det(A_1 - \lambda I) \propto \lambda^{m(K-1)}\ det(\tilde{A}_1 + \lambda I + \frac{\tilde{A}_2}{\lambda} + \frac{\tilde{A}_3}{\lambda^2} + ... + \frac{\tilde{A}_K}{\lambda^{K-1}}) \tag{2.165}$$

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.

Chapter 3

Econometric Dynamic Models: From Classical Econometrics to Time Series Econometrics

3.1 Macroeconometric Structural Models Versus VAR Models

According to the so-called time series econometrics, the typical assumption of classical econometrics about the determinant role played by economic theory in model specification is refuted. Therefore, the core of econometric modelling rests crucially on VAR specifications (after the premises of Sect. 2.5) with the addition of integration and cointegration analysis to overcome the problem of non stationary variables and detect possibly stable economic relationships from available data.

This implies that the conceptual frame based upon the interaction among economic theory, mathematics and empirical evidence – provided with the pertinent statistical reading key – which characterizes classical econometrics, leads to a mirror reinterpretation within the time series econometrics. The implication is essentially an overturning between the *operative role* of empirical evidence and the *guide role* of economic theory.

Thus, whereas the empirical evidence plays a confirmatory role in comparison with economic theory within classical econometrics – about which a *iuris tantum* presumption of *a priori* reliability does indeed exist, although not explicitly expressed – in time series econometrics the perspective is in a certain way overturned. Here are the data – that is the empirical evidence – to outline the frame of reference, while economic theory intervenes with a confirmatory role to validate *a posteriori* the coherence of the results obtained through statistical methods, according to principles accepted by economic theory.

In light of these brief considerations, it is possible to single out the common aspects as well as the distinctive features of the above mentioned approaches to econometric modelling. One could then understand the *modus operandi* of econometric research within both perspectives, when the common denominators are provided by economic theory and by empirical evidence, even though with different hierarchical roles.

Given these preliminaries, the reader will find in this section a comparison – restricted to the essential features – between the dynamic specification in the context of structural econometrics and in that of VAR modelling. In particular, the proposed characterization rests on the assumption that the roots of the characteristic polynomial associated with the model lie outside or on the unitary circle, with the reduced form acting as a unifying frame of reference.

On the one hand, the reference to the reduced form could lead to a restrictive reading key of VAR models in subordinate terms with respect to structural models. This ranking, on the other hand, turns out to be only apparent when the nature and the role of the (unit) roots of the parent characteristic polynomial are taken in due account.

As a starting point it may be convenient to consider the following general primary form for the model

$$\underset{(n,1)}{\boldsymbol{y}_t} = \boldsymbol{\Gamma}\boldsymbol{y}_t + \boldsymbol{\Gamma}^*(L)\,\boldsymbol{y}_t + \boldsymbol{A}^*(L)\,\boldsymbol{x}_t + \boldsymbol{\varepsilon}_t \tag{3.1}$$

where

$$\boldsymbol{\Gamma}^*(L) = \sum_{k=1}^{K} \boldsymbol{\Gamma}_k\, L^k \tag{3.2}$$

and

$$\boldsymbol{A}^*(L) = \sum_{r=0}^{R} \boldsymbol{A}_r\, L^r \tag{3.3}$$

The notation reflects the one currently used in econometric literature (see, e.g., Faliva 1987). Here the vectors $\boldsymbol{y}$ and $\boldsymbol{x}$ denote the endogenous and the exogenous variables respectively, $\boldsymbol{\varepsilon}$ represents a white noise vector of disturbances, whereas $\boldsymbol{\Gamma}$, $\boldsymbol{\Gamma}_s$ and $\boldsymbol{A}_r$ stand for matrices of parameters.

Next, we consider first the point of view of classical econometrics and then that of time series econometrics.

The distinctive features of structural models are

(1) $$\boldsymbol{\Gamma} * \boldsymbol{I} = \boldsymbol{0}, \tag{3.4}$$

where the symbol $*$ denotes the Hadamard product (see, e.g., Styan 1973) .

(2) $\boldsymbol{\Gamma}$, $\boldsymbol{\Gamma}_k$ (k = 1, 2, ..., K) and $\boldsymbol{A}_r$ (r = 0, 1, ..., R) are sparse matrices, specified according to the economic theory, *ex ante* with respect to model estimation and validation.

While formula (3.1) expresses the so-called structural form of the model, which arises from the transposition of the economic theory into a model, the secondary (reduced) form of the model is given by

$$\boldsymbol{y}_t = \boldsymbol{P}(L)\,\boldsymbol{y}_t + \boldsymbol{\Pi}(L)\,\boldsymbol{x}_t + \boldsymbol{\mu}_t \tag{3.5}$$

where

$$\boldsymbol{\mu}_t = (\boldsymbol{I} - \boldsymbol{\Gamma})^{-1}\boldsymbol{\varepsilon}_t \tag{3.6}$$

$$\boldsymbol{P}(L) = \sum_{k=1}^{K} \boldsymbol{P}_k L^k \tag{3.7}$$

$$\boldsymbol{\Pi}(L) = \sum_{r=0}^{R} \boldsymbol{\Pi}_r L^r \tag{3.8}$$

with

$$\boldsymbol{P}_k = (\boldsymbol{I}_L - \boldsymbol{\Gamma})^{-1}\,\boldsymbol{\Gamma}_k,\; k = 1, 2, ..., K \tag{3.9}$$

$$\boldsymbol{\Pi}_r = (\boldsymbol{I}_L - \boldsymbol{\Gamma})^{-1}\,\boldsymbol{A}_r,\; r = 0, 1, ..., R \tag{3.10}$$

In a more compact form model (3.5) may be written as follows

$$\boldsymbol{A}(L)\,\boldsymbol{y}_t = \boldsymbol{\Pi}(L)\,\boldsymbol{x}_t + \boldsymbol{\mu}_t \tag{3.11}$$

with $\boldsymbol{A}(L)$ defined as

$$\boldsymbol{A}(L) = \boldsymbol{I} - \boldsymbol{P}(L) \tag{3.12}$$

The spectrum of the characteristic polynomial

$$\left|\boldsymbol{A}(z)\right| = det\,[\boldsymbol{I} - \boldsymbol{P}(z)] \tag{3.13}$$

plays a crucial role in the analysis. As a matter of fact, the assumption that all its roots lie outside the unitary circle is indeed a main feature of structural models.

Starting from the reduced form in formula (3.11), it is possible to obtain, through suitable computations (cf. Faliva 1987, p 167), the so–called final form of the model, namely

$$\boldsymbol{y}_t = \boldsymbol{H}\lambda^{(t)} + [\boldsymbol{A}(L)]^{-1}\,\boldsymbol{\Pi}(L)\,\boldsymbol{x}_t + [\boldsymbol{A}(L)]^{-1}\,\boldsymbol{\mu}_t \tag{3.14}$$

Here $\lambda^{(t)}$ denotes the vector

$$\underset{(nK,1)}{\lambda^{(t)}} = \begin{bmatrix} \lambda_1^t \\ \lambda_2^t \\ \vdots \\ \lambda_{nK}^t \end{bmatrix} \tag{3.15}$$

whose elements are the t-th powers of the solutions $\lambda_1, \lambda_2, \ldots, \lambda_{nK}$ (which are all assumed to be distinct, in order to simplify the formulas) of the equation

$$det\,(\lambda^K \boldsymbol{I}_n - \sum_{k=1}^{K} \lambda^{K-k} \boldsymbol{P}_k) = 0 \tag{3.16}$$

and $\boldsymbol{H}$ is a matrix whose columns $\boldsymbol{h}_i$ are the non trivial solutions of the homogeneous systems

$$(\lambda_i^K \boldsymbol{I} - \lambda_i^{K-1} \boldsymbol{P}_1 - \ldots - \lambda_i \boldsymbol{P}_{K-1} - \boldsymbol{P}_K)\, \boldsymbol{h}_i = \boldsymbol{0}_n \quad i = 1, 2, \ldots, nK \tag{3.17}$$

The term $\boldsymbol{H}\lambda^{(t)}$, on the right side of (3.14), reflects the dynamics of the endogenous variables of inside origin (so–called autonomous component) not due to exogenous or casual factors, which corresponds to the general solution of the homogeneous equation

$$\boldsymbol{A}\,(L)\,\boldsymbol{y}_t = 0 \tag{3.18}$$

The last two terms in the second member of (3.14) represent a particular solution of the non-homogeneous equation (3.11).

The term

$$[\boldsymbol{A}\,(L)]^{-1}\,\boldsymbol{\Pi}\,(L)\,\boldsymbol{x}_t = \sum_{\tau=0}^{\infty} \boldsymbol{K}_\tau \boldsymbol{x}_{t-\tau} \tag{3.19}$$

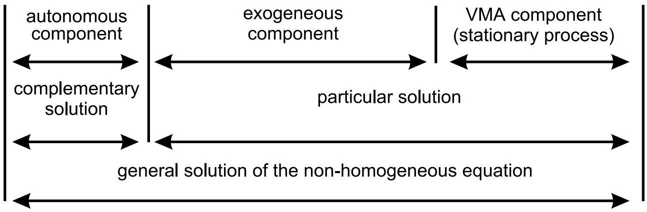
reflects the deterministic dynamics, due to exogenous factors (so-called exogenous component), while the term

$$[\boldsymbol{A}\,(L)]^{-1}\,\boldsymbol{\mu}_t = \sum_{\tau=0}^{\infty} \boldsymbol{C}_\tau \boldsymbol{\mu}_{t-\tau} \tag{3.20}$$

reflects the dynamics induced by casual factors (so-called stochastic component), which assumes the form of a causal moving average *VMA* (∞) of a multivariate white noise, namely a stationary process.

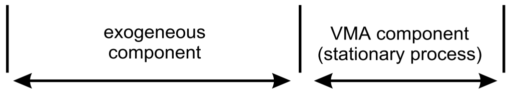
The complete reading key of (3.14) is illustrated in the following scheme

$$\boldsymbol{y}_t = \boldsymbol{H}\lambda^{(t)} + [\boldsymbol{I} - \boldsymbol{P}(L)]^{-1}\,\Pi(L)\,\boldsymbol{x}_t + [\boldsymbol{I} - \boldsymbol{P}(L)]^{-1}\mu_t \tag{3.21}$$



It is worth mentioning that, in this context, the autonomous component assumes a transitory character which is uninfluential in the long run. This is because the scalars $\lambda_1, \lambda_2, \ldots, \lambda_{nK}$ are the reciprocals of the roots of the characteristic polynomial (3.13) and as such lie inside the unit circle. As a result, the component $\boldsymbol{H}\lambda^{(t)}$ may be neglected when t is quite high, leading to a more concise representation

$$\boldsymbol{y}_t = [\boldsymbol{I} - \boldsymbol{P}(L)]^{-1}\,\Pi(L)\,\boldsymbol{x}_t + [\boldsymbol{I} - \boldsymbol{P}(L)]^{-1}\mu_t \tag{3.22}$$



What we have seen so far are the salient points in the analysis of linear dynamic models from the viewpoint of classical econometrics.

As far as time series econometrics is concerned, the distinctive features of the VAR model are

(1) $\boldsymbol{\Gamma} = \boldsymbol{0} \Rightarrow \boldsymbol{\Gamma}_k = \boldsymbol{P}_k\,, \quad k = 1, 2, ..., K$ (3.23)

(2) $\boldsymbol{\Gamma}_k$ ($k = 1, 2, ..., K$), are full matrices, in absence of an economic informative theory, *ex ante* with respect to model estimation and validation.

(3) $\boldsymbol{A}^*(L) = \boldsymbol{0} \Rightarrow \boldsymbol{A}_r = \boldsymbol{\Pi}_r = \boldsymbol{0}, \quad r = 0, 1, ..., R$ (3.24)

As long as the distinction between endogenous and exogenous variables is no longer drawn, all relevant variables *de facto* turn out to be treated as endogenous.

Here the primary and secondary forms are the same: the model is in fact automatically specified by the reduced form, which in light of (3.23) and of (3.24) reads as

$$\underset{(n,1)}{\boldsymbol{y}_t} = \boldsymbol{P}(L)\,\boldsymbol{y}_t + \boldsymbol{\varepsilon}_t \tag{3.25}$$

where

$$\boldsymbol{P}(L) = \boldsymbol{\Gamma}^*(L) \tag{3.26}$$

with the further qualification

$$\boldsymbol{\Pi}(L) = \boldsymbol{A}^*(L) = \boldsymbol{0} \tag{3.27}$$

The specification of the VAR model, according to (3.25) and in view of (3.7), assumes the form

$$\boldsymbol{y}_t = \sum_{k=1}^{K} \boldsymbol{P}_k \boldsymbol{y}_{t-k} + \boldsymbol{\varepsilon}_t \tag{3.28}$$

According to (3.12) the following representation holds

$$\boldsymbol{A}(L)\boldsymbol{y}_t = \boldsymbol{\varepsilon}_t \tag{3.29}$$

The solution of the operational equation (3.29) – which is the counterpart to the notion of final form of classical econometrics – is the object of the so-called representation theorems, and can be given a form such as

$$\boldsymbol{y}_t = \boldsymbol{H\lambda}^{(t)} + \boldsymbol{k}_0 + \boldsymbol{k}_1 t + \boldsymbol{N}_2 \sum_{\tau \le t} (t+1-\tau)\,\boldsymbol{\varepsilon}_\tau + \boldsymbol{N}_1 \sum_{\tau \le t} \boldsymbol{\varepsilon}_\tau + \sum_{i=0}^{\infty} \boldsymbol{M}_i \boldsymbol{\varepsilon}_{t-i} \tag{3.30}$$

whose rationale will become clear from the subsequent Sects. 3.4 and 3.5, which are concerned with specifications of prominent interest for econometrics involving processes integrated up to the second order.

In formula (3.30) the term $\boldsymbol{H\lambda}^{(t)}$, analogously to what was pointed out for (3.14), represents the autonomous component which is of transitory character and corresponds to the solution of the homogeneous equation

$$\boldsymbol{A}(L)\,\boldsymbol{y}_t = \boldsymbol{0} \tag{3.31}$$

inherent to the roots of the characteristic polynomial which lie outside the unitary circle. Conversely, the term $\boldsymbol{k}_0 + \boldsymbol{k}_1 t$ represents the autonomous component (so-called deterministic trend) which has permanent character insofar as it is inherent to unitary roots.

The other terms represent a particular solution of the non-homogeneous equation (3.29). Specifically, the term $\sum_{i=0}^{\infty} \boldsymbol{M}_i \boldsymbol{\varepsilon}_{t-i}$ is a causal moving-average process – whose analogy with (3.20) is evident–associated with the regular part of Laurent expansion of $\boldsymbol{A}^{-1}(z)$ in a deleted neighbourhood of $z = 1$.

The term

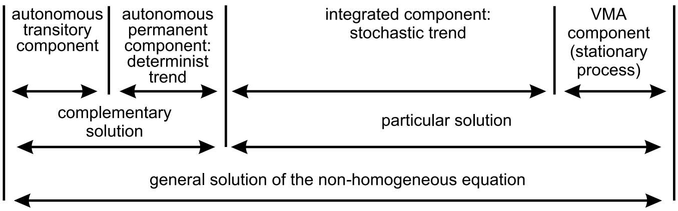
$$\boldsymbol{N}_2 \sum_{\tau \le t} (t + 1 - \tau)\, \boldsymbol{\varepsilon}_\tau + \boldsymbol{N}_1 \sum_{\tau \le t} \boldsymbol{\varepsilon}_\tau \tag{3.32}$$

on the other hand, reflects the (random) dynamics, i.e. the stochastic trend or integrated component associated with the principal part of the Laurent expansion of $\boldsymbol{A}^{-1}(z)$ in a deleted neighbourhood of $z = 1$, where $z = 1$ is meant to be a second order pole of $\boldsymbol{A}^{-1}(z)$.

As it will become evident in Sects. 3.4 and 3.5, the cointegration relations of the model turn out to be associated with the left eigenvectors corresponding to the null eigenvalues of $\boldsymbol{N}_2$ and $\boldsymbol{N}_1$.

The complete reading key of (3.30) is illustrated in the following scheme

$$\boldsymbol{y}_t = \boldsymbol{H}\boldsymbol{\lambda}^{(t)} + \boldsymbol{k}_0 + \boldsymbol{k}_1 t + \boldsymbol{N}_1 \sum_{\tau \le t} \boldsymbol{\varepsilon}_\tau + \boldsymbol{N}_2 \sum_{\tau \le t} (t + 1 - \tau)\, \boldsymbol{\varepsilon}_\tau + \sum_{i=0}^{\infty} \boldsymbol{M}_i \boldsymbol{\varepsilon}_{t-i}$$



(3.33)

3.2 Basic VAR Specifications and Engendered Processes

The usefulness of VAR specifications to actually grasp (as anticipated in Sect. 2.5) the dynamics of economic variables rests on *ad hoc* rank qualifications of the parameter matrices in the reference model.

Before going into the matter in due depth and eventually tackling the major issues of representation theorems content and meaning, let us first gain an insight into the general setting of unit-root econometrics.

Definition 1 – Basic VAR Model

A vector autoregressive (VAR) model

$$\underset{(n,n)}{\boldsymbol{A}(L)}\ \underset{(n,1)}{\boldsymbol{y}_t} = \underset{(n,1)}{\boldsymbol{\varepsilon}_t} + \underset{(n,1)}{\boldsymbol{\eta}}, \quad \boldsymbol{\varepsilon}_t \sim WN_{(n)} \tag{3.34}$$

where $\boldsymbol{\eta}$ is a vector of constants (drift vector) and

$$\boldsymbol{A}(L) = \sum_{k=0}^{K} \boldsymbol{A}_k L^k, \; \boldsymbol{A}_0 = \boldsymbol{I}, \boldsymbol{A}_K \neq \boldsymbol{0} \tag{3.35}$$

is a matrix polynomial whose characteristic polynomial

$$\pi(z) = det\, \boldsymbol{A}(z) \tag{3.36}$$

can be factored as

$$\pi(z) = (1-z)^{\alpha}\, \tilde{\pi}(z) \tag{3.37}$$

Here, $\alpha \geq 0$ is a non negative integer and $\tilde{\pi}(z)$ has all roots outside the unit circle, will be referred to as a basic VAR model of order K and dimension n.

Definition 2 – Error Correction Model

VAR models can also be specified in terms of both levels and differences by resorting to representations such as

$$\boldsymbol{Q}(L)\, \nabla \boldsymbol{y}_t + \boldsymbol{A}\boldsymbol{y}_t = \boldsymbol{\varepsilon}_t + \boldsymbol{\eta} \tag{3.38}$$

$$\boldsymbol{\Psi}(L)\, \nabla^2 \boldsymbol{y}_t - \dot{\boldsymbol{A}}\, \nabla \boldsymbol{y}_t + \boldsymbol{A}\boldsymbol{y}_t = \boldsymbol{\varepsilon}_t + \boldsymbol{\eta} \tag{3.39}$$

where the symbols have the same meaning as in (1.252) and (1.257) of Sect. 1.6. Such representations are referred to as error-correction models (ECM).

The following propositions summarize the fundamental features of VAR-based econometric modelling. Here the proofs, when not derived as by-products from the results presented in Chaps. 1 and 2, are justified by developments to be found in later sections.

Proposition 1

A basic VAR model, as per Definition 1, engenders a stationary process, i.e.

$$\boldsymbol{y}_t \sim I(0) \tag{3.40}$$

whenever

$$r(\boldsymbol{A}) = n \tag{3.41}$$

or, otherwise stated, whenever

$$\alpha = 0 \tag{3.42}$$

Proof

The proposition ensues from Theorem 1 of Sect. 1.10 together with Theorem 1 of Sect. 2.3, after the arguments developed therein.

□

Thereafter the matrix $\boldsymbol{A}$, even if singular, will always be assumed to be non null.

Proposition 2

A basic VAR model, as per Definition 1, engenders an integrated process, i.e.

$$\boldsymbol{y}_t \sim I\,(d) \tag{3.43}$$

where d is a positive integer, whenever

$$r\,(\boldsymbol{A}) < n \tag{3.44}$$

or, otherwise stated, whenever

$$\alpha > 0 \tag{3.45}$$

Proof

The proof is the same as that of Proposition 1.

□

Proposition 3

A basic VAR model, as per Definition 1, engenders a first order integrated process, i.e.

$$\boldsymbol{y}_t \sim I\,(1) \tag{3.46}$$

if

$$det\,(\boldsymbol{A}) = 0 \tag{3.47}$$

$$det\,(\boldsymbol{B}'_{\perp}\,\dot{\boldsymbol{A}}\,\boldsymbol{C}_{\perp}) \neq 0 \tag{3.48}$$

where $\boldsymbol{B}_{\perp}$ and $\boldsymbol{C}_{\perp}$ denote the orthogonal complements of the matrices $\boldsymbol{B}$ and $\boldsymbol{C}$ of the rank factorization

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}' \tag{3.49}$$

Proof

The statement ensues from Theorem 3, in view of Corollary 3.1, of Sect. 1.10, together with Theorem 1 of Sect. 2.3, after the considerations made therein.

□

Proposition 4

Under the assumptions of Proposition 3, the twin processes $\boldsymbol{C}'_{\perp}\boldsymbol{y}_t$ and $\boldsymbol{C}'\boldsymbol{y}_t$ are integrated of first order and stationary respectively, i.e.

$$\boldsymbol{C}'_{\perp}\boldsymbol{y}_t \sim I(1) \tag{3.50}$$

$$\boldsymbol{C}'\boldsymbol{y}_t \sim I(0) \tag{3.51}$$

which, in turn, entails the cointegration property

$$\boldsymbol{y}_t \sim CI(1, 1) \tag{3.52}$$

to hold true for the process $\boldsymbol{y}_t$.

Proof

The proposition mirrors the twin statements (2.97) and (2.98) of Theorem 2 in Sect. 2.3 once we take $d = 1$ and $\delta = d - 1 = 0$. Indeed, writing the VAR model in the ECM form (3.38), making use of (3.49) and rearranging the terms, we obtain

$$\boldsymbol{B}\boldsymbol{C}'\boldsymbol{y}_t = \boldsymbol{\varepsilon}_t + \boldsymbol{\eta} - \boldsymbol{Q}(L)\nabla\boldsymbol{y}_t \tag{3.53}$$

As long as $\boldsymbol{y}_t \sim I(1)$, the following holds true

$$\boldsymbol{Q}(L)\nabla\boldsymbol{y}_t \sim I(0) \tag{3.54}$$

which, in turn, entails

$$\boldsymbol{B}\boldsymbol{C}'\boldsymbol{y}_t \sim I(0) \qquad \boldsymbol{C}'\boldsymbol{y}_t \sim I(0) \tag{3.55}$$

in view of (3.53).

For what concerns (3.50), the result is an outcome of the representation theorem for $I(1)$ processes to which Sect. 3.4 will thereinafter be devoted. After (3.50) and (3.51), the conclusion (3.52), in tune with Definition 6 of Sect. 2.4, about the cointegrated nature of $\boldsymbol{y}_t$, is accordingly drawn.

□

Proposition 5

A basic VAR model, as per Definition 1, engenders a second order integrated process, i.e.

$$y_t \sim I(2) \tag{3.56}$$

if

$$det(A) = 0 \tag{3.57}$$

$$det(B'_\perp \dot{A} C_\perp) = 0, \quad B'_\perp \dot{A} C_\perp \neq 0 \tag{3.58}$$

$$det(R'_\perp B'_\perp \tilde{A} C_\perp S_\perp) \neq 0 \tag{3.59}$$

where the matrices $B_\perp$ and $C_\perp$ have the same meaning as in Proposition 3, the matrices $R_\perp$ and $S_\perp$ denote the orthogonal complements of R and S in the rank factorization

$$B'_\perp \dot{A} C_\perp = RS' \tag{3.60}$$

and the matrix $\tilde{A}$ is given by

$$\tilde{A} = \frac{1}{2}\ddot{A} - \dot{A} A^g \dot{A} \tag{3.61}$$

Proof

The result follows from Theorem 5, in view of Corollary 5.1, of Sect. 1.10 together with Theorem 1 of Sect. 2.3, in accordance with the reasoning followed therein.

□

Proposition 6

Under the assumptions of Proposition 5, the twin processes $S'_\perp C'_\perp y_t$ and $(C_\perp S_\perp)'_\perp y_t$ are integrated of the second and of a lower order, respectively, i.e.

$$S'_\perp C'_\perp y_t \sim I(2) \tag{3.62}$$

$$(C_\perp S_\perp)'_\perp y_t \sim I(\gamma), \ \gamma < 2 \tag{3.63}$$

which in turn entails the cointegration properties

$$\boldsymbol{y}_t \sim CI\,(2, 2\text{–}\gamma) \tag{3.64}$$

to hold true for the process $\boldsymbol{y}_t$.

Proof

A link– although partial – between this proposition and Theorem 2 in Sect. 2.3 can be established by resorting to the representation (1.626) of Sect. 1.11 for $(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})_{\perp}$.

Actually, results (3.62) and (3.63) are spin-offs of the representation theorem for $I\,(2)$ processes to which Sect. 3.5 will thereinafter be devoted.

After (3.62) and (3.63), the conclusion (3.64), in tune with Definition 6 of Sect. 2.4, about the cointegrated nature of $\boldsymbol{y}_t$, is accordingly drawn.

□

3.3 A Sequential Rank Criterion for the Integration Order of a VAR Solution

The following theorem provides a chain rule for the integration order of a process generated by a VAR model on the basis of the rank characteristics of its matrix coefficients.

Theorem 1

Consider a basic VAR model, as per Definition 1 of Sect. 3.2

$$\boldsymbol{A}\,(L)\,\boldsymbol{y}_t = \boldsymbol{\varepsilon}_t \;\; \boldsymbol{\varepsilon}_t \sim WN_{(n)} \tag{3.65}$$

where the symbols have the same meaning as in that section, the matrices $\boldsymbol{\Gamma}$, $\tilde{\boldsymbol{\Gamma}}$ *and* $\boldsymbol{\Lambda}$ *are defined as follows*

$$\boldsymbol{\Gamma} = [\,\boldsymbol{A}_l^{\perp}\,\dot{\boldsymbol{A}}\,\boldsymbol{A}_r^{\perp}\,, \boldsymbol{A}],\;\; \tilde{\boldsymbol{\Gamma}} = \begin{bmatrix} \boldsymbol{A}_l^{\perp}\dot{\boldsymbol{A}}\boldsymbol{A}_r^{\perp} \\ \boldsymbol{A} \end{bmatrix} \tag{3.66}$$

$$\boldsymbol{\Lambda} = (\boldsymbol{I} - \boldsymbol{\Gamma}\boldsymbol{\Gamma}^{g})\,\tilde{\boldsymbol{A}}\,(\boldsymbol{I} - \tilde{\boldsymbol{\Gamma}}^{g}\,\tilde{\boldsymbol{\Gamma}}\,) \tag{3.67}$$

and $\boldsymbol{A}_l^{\perp}$ *and* $\boldsymbol{A}_r^{\perp}$ *are as in Definition 3 of Sect. 1.2.*

The following results hold true.

(1) If

$$r\,(\boldsymbol{A}) = n \tag{3.68}$$

then

$$\boldsymbol{y}_t \sim I\,(0) \tag{3.69}$$

whereas if

$$r\,(\boldsymbol{A}) < n \tag{3.70}$$

then

$$\boldsymbol{y}_t \sim I\,(d),\;\; d > 0 \tag{3.71}$$

(2) Under rank condition (3.70), if

$$r\,([\,\dot{\boldsymbol{A}}\,, \boldsymbol{A}]) = n \tag{3.72}$$

then

$$\boldsymbol{y}_t \sim I\,(d),\;\; d \geq 1 \tag{3.73}$$

whereas if

$$r\,([\,\dot{\boldsymbol{A}}\,, \boldsymbol{A}]) < n \tag{3.74}$$

then

$$\boldsymbol{y}_t \sim I\,(d),\;\; d \geq 2 \tag{3.75}$$

(3) Under rank condition (3.72), if

$$r\,(\boldsymbol{\Gamma}) = n \tag{3.76}$$

then

$$\boldsymbol{y}_t \sim I\,(1) \tag{3.77}$$

whereas if

$$r\,(\boldsymbol{\Gamma}) < n \tag{3.78}$$

then

$$\boldsymbol{y}_t \sim I\,(d),\;\; d > 1 \tag{3.79}$$

(4) Under rank condition (3.78), if

$$r\,([\,\tilde{\boldsymbol{A}}\,, \boldsymbol{\Gamma}]) = n \tag{3.80}$$

then

$$\boldsymbol{y}_t \sim I(d), \;\; d \geq 2 \tag{3.81}$$

whereas if

$$r\,([\tilde{\boldsymbol{A}}, \boldsymbol{\Gamma}]) < n \tag{3.82}$$

then

$$\boldsymbol{y}_t \sim I(d), \;\; d \geq 3 \tag{3.83}$$

(5) Under rank condition (3.80) if

$$r\,([\boldsymbol{\Lambda}, \Gamma]) = n \tag{3.84}$$

then

$$\boldsymbol{y}_t \sim I\,(2) \tag{3.85}$$

whereas if

$$r\,([\boldsymbol{\Lambda}, \;\Gamma]) < n \tag{3.86}$$

then

$$\boldsymbol{y}_t \sim I(d), \;\; d > 2 \tag{3.87}$$

Proof

To prove (1) refer back to Theorem 4 of Sect. 1.7 and Theorem 1 of Sect. 2.3 (compare also with Propositions 1 and 2 of Sect. 3.2).

To prove point (2) refer back to Theorems 3 and 5, as well as to Corollary 3.1 of Sect. 1.10, and also to Theorem 1 of Sect. 2.3 (compare also with Propositions 3 and 5 of Sect. 3.2). Then, observe that

$$r\left(\begin{bmatrix} -\dot{\boldsymbol{A}} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix}\right) = r\left[\begin{bmatrix} -\boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B} \end{bmatrix}\begin{bmatrix} -\dot{\boldsymbol{A}} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{C}' \end{bmatrix}\right] = r\left(\begin{bmatrix} \dot{\boldsymbol{A}} & \boldsymbol{A} \\ \boldsymbol{A} & \boldsymbol{0} \end{bmatrix}\right) \tag{3.88}$$

whence

$$r\left(\begin{bmatrix} -\dot{\boldsymbol{A}} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix}\right) = n + r\,(\boldsymbol{A}) \Leftrightarrow r\left(\begin{bmatrix} \dot{\boldsymbol{A}} & \boldsymbol{A} \\ \boldsymbol{A} & \boldsymbol{0} \end{bmatrix}\right) = n + r\,(\boldsymbol{A}) \tag{3.89}$$

$$= r\,([\dot{\boldsymbol{A}}, \boldsymbol{A}]) = n$$

but not necessarily the other way around.

Hence, d is possibly equal to one as per (3.73) under (3.72), whereas this is no longer possible as per (3.75) under (3.74).

The proof of (3) rests on Theorems 14 and 19 in Marsaglia and Styan, and Definition 3 of Sect. 1.2, yielding the rank equalities

$$\begin{aligned} & r\left(\begin{bmatrix} -\dot{\boldsymbol{A}} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix}\right) = r\left(\begin{bmatrix} \dot{\boldsymbol{A}} & \boldsymbol{A} \\ \boldsymbol{A} & \boldsymbol{0} \end{bmatrix}\right) \\ & = r\,(\boldsymbol{A}) + r\,(\boldsymbol{A}) + r\,((\boldsymbol{I} - \boldsymbol{A}\boldsymbol{A}^g)\,\dot{\boldsymbol{A}}\,(\boldsymbol{I} - \boldsymbol{A}^g\boldsymbol{A})) \\ & = r\,(\boldsymbol{A}) + r\,(\boldsymbol{A}) + r\,(\boldsymbol{A}_l^{\perp}\dot{\boldsymbol{A}}\boldsymbol{A}_r^{\perp}) = r\,(\boldsymbol{A}) + r\,(\boldsymbol{\Gamma}) \end{aligned} \tag{3.90}$$

Hence, after (3.76) it is true that

$$r\left(\begin{bmatrix} -\dot{\boldsymbol{A}} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix}\right) = n + r\,(\boldsymbol{A}) \tag{3.91}$$

and (3.77) follows accordingly, in view of the theorems quoted in proving (2), whereas the circumstance (3.79) occurs under (3.78).

To prove (4) refer, on the one hand, back to Theorem 3 – along with its corollary – of Sect. 1.10 and Theorem 1 of Sect. 2.3 (compare also with Proposition 5 of the Sect. 3.2) and, on the other hand, to the proof of (2), by replacing

$$\begin{bmatrix} -\boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B} \end{bmatrix}, \begin{bmatrix} -\dot{\boldsymbol{A}} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix}, \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{C}' \end{bmatrix} \tag{3.92}$$

with

$$\begin{aligned} & \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & (\boldsymbol{B}_{\perp}')^g \boldsymbol{R} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{B} \end{bmatrix}, \begin{bmatrix} \tilde{\boldsymbol{A}} & (\boldsymbol{B}_{\perp}')^g \boldsymbol{R} & \boldsymbol{B} \\ \boldsymbol{S}'(\boldsymbol{C}_{\perp})^g & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{C}' & \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}, \\ & \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{S}'(\boldsymbol{C}_{\perp})^g & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{C}' \end{bmatrix} \end{aligned} \tag{3.93}$$

respectively, after the equalities

$$(\boldsymbol{B}_\perp \boldsymbol{R}_\perp)_\perp = [(\boldsymbol{B}'_\perp)^g \boldsymbol{R}, \boldsymbol{B}] \tag{3.94}$$

$$(\boldsymbol{C}_\perp \boldsymbol{S}_\perp)_\perp = [(\boldsymbol{C}'_\perp)^g \boldsymbol{S}, \boldsymbol{C}] \tag{3.95}$$

$$(\boldsymbol{B}'_\perp)^g \boldsymbol{R}\boldsymbol{S}'(\boldsymbol{C}_\perp)^g = \boldsymbol{A}_l^\perp \dot{\boldsymbol{A}} \boldsymbol{A}_r^\perp \tag{3.96}$$

because of (1.554) of Sect. 1.10 and the pairs (1.87)–(1.94) and (1.88)–(1.95) of Sect. 1.2.

Next observe that

$$r\left(\begin{bmatrix} \tilde{\boldsymbol{A}} & (\boldsymbol{B}_\perp \boldsymbol{R}_\perp)_\perp \\ (\boldsymbol{C}_\perp \boldsymbol{S}_\perp)'_\perp & \boldsymbol{0} \end{bmatrix}\right)$$

$$= r\left(\begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & (\boldsymbol{B}'_\perp)^g \boldsymbol{R} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & B \end{bmatrix} \begin{bmatrix} \tilde{\boldsymbol{A}} & (\boldsymbol{B}'_\perp)^g \boldsymbol{R} & \boldsymbol{B} \\ \boldsymbol{S}'(\boldsymbol{C}_\perp)^g & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{C}' & \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{S}'(\boldsymbol{C}_\perp)^g & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{C}' \end{bmatrix}\right) \tag{3.97}$$

$$= r\begin{bmatrix} \tilde{\boldsymbol{A}} & \boldsymbol{\Gamma} \\ \tilde{\boldsymbol{\Gamma}} & \boldsymbol{0} \end{bmatrix}$$

whence

$$\begin{aligned} & r\left(\begin{bmatrix} \tilde{\boldsymbol{A}} & (\boldsymbol{B}_\perp \boldsymbol{R}_\perp)_\perp \\ (\boldsymbol{C}_\perp \boldsymbol{S}_\perp)'_\perp & \boldsymbol{0} \end{bmatrix}\right) = n + r\,((\boldsymbol{C}_\perp \boldsymbol{S}_\perp)_\perp) \Leftrightarrow \\ & \Leftrightarrow r\begin{bmatrix} \tilde{\boldsymbol{A}} & \boldsymbol{\Gamma} \\ \tilde{\boldsymbol{\Gamma}} & \boldsymbol{0} \end{bmatrix} = n + r\,((\boldsymbol{C}_\perp \boldsymbol{S}_\perp)_\perp) \quad\; r\,([\tilde{\boldsymbol{A}},\ \boldsymbol{\Gamma}]) = n \end{aligned} \tag{3.98}$$

but not necessarily the other way around, given that

$$r\,(\boldsymbol{\Gamma}) = r\,(\tilde{\boldsymbol{\Gamma}}) = r\,(\boldsymbol{C}_\perp \boldsymbol{S}_\perp)_\perp \tag{3.99}$$

Hence d is possibly equal to two as per (3.81) under (3.80), whereas this is no longer possible as per (3.83) under (3.82).

The proof of point (5) proceeds along the same lines as that of point (3) by replacing $\boldsymbol{A}_r^{\perp}$, $\dot{\boldsymbol{A}}$, $\boldsymbol{A}_l^{\perp}$, $\boldsymbol{A}$ respectively, with $(\boldsymbol{I} - \tilde{\boldsymbol{\Gamma}}^g \tilde{\boldsymbol{\Gamma}})$, $\tilde{\boldsymbol{A}}$, $(\boldsymbol{I} - \boldsymbol{\Gamma}\boldsymbol{\Gamma}^g)$, $\boldsymbol{\Gamma}$.

In this connection, observe that

$$r\left(\begin{bmatrix} \tilde{\boldsymbol{A}} & (\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp} \\ (\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp} & \boldsymbol{0} \end{bmatrix}\right) = r\left(\begin{bmatrix} \tilde{\boldsymbol{A}} & \boldsymbol{\Gamma} \\ \tilde{\boldsymbol{\Gamma}} & \boldsymbol{0} \end{bmatrix}\right)$$
$$= r(\boldsymbol{\Gamma}) + r(\boldsymbol{\Gamma}) + r((\boldsymbol{I} - \boldsymbol{\Gamma}\boldsymbol{\Gamma}^g)\tilde{\boldsymbol{A}}(\boldsymbol{I} - \tilde{\boldsymbol{\Gamma}}^g \tilde{\boldsymbol{\Gamma}}))$$
$$= r(\boldsymbol{\Gamma}) + r([\boldsymbol{\Lambda}, \ \boldsymbol{\Gamma}]) \tag{3.100}$$

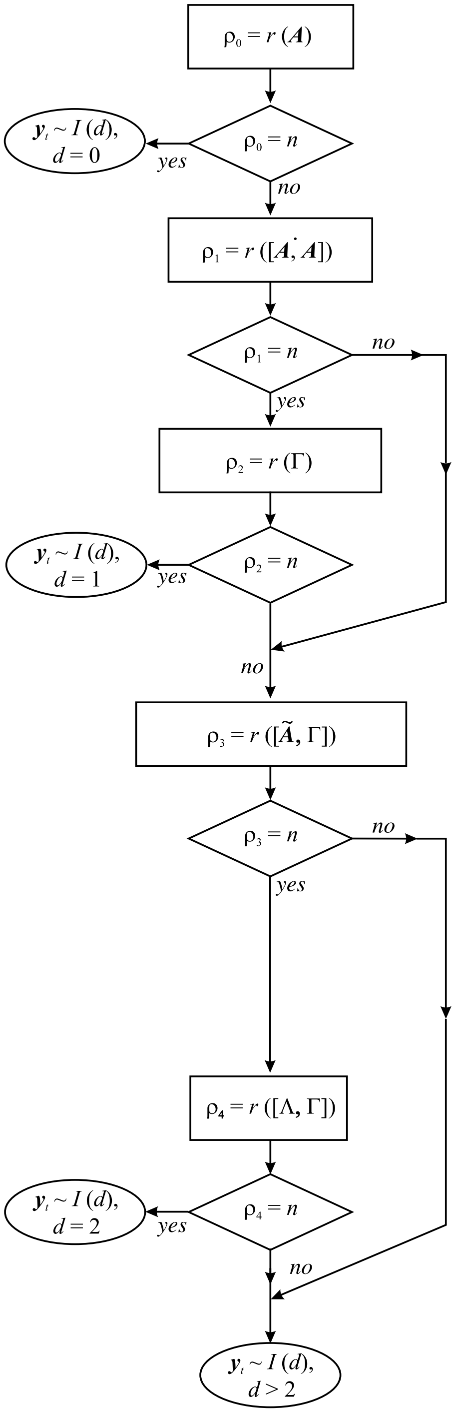
Hence, after (3.84) the following holds

$$r\left(\begin{bmatrix} \tilde{\boldsymbol{A}} & (\boldsymbol{B}_{\perp}\boldsymbol{R}_{\perp})_{\perp} \\ (\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp} & \boldsymbol{0} \end{bmatrix}\right) = n + r(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})_{\perp} \tag{3.101}$$

and (3.85) ensues accordingly, in view of the theorems quoted in proving (4), whereas the circumstances (3.87) occurs under (3.86).

□

The sequential procedure for integration order identification inherent in the theorem can be given an enlightening visualization by the decision chart of the next page.


ρ₀ = r (A)
ρ₀ = n
yes
no
y_t ~ I (d), d = 0
ρ₁ = r ([A˙, A])
ρ₁ = n
no
yes
ρ₂ = r (Γ)
ρ₂ = n
yes
y_t ~ I (d), d = 1
no
ρ₃ = r ([Ã, Γ])
ρ₃ = n
no
yes
ρ₄ = r ([Λ, Γ])
ρ₄ = n
yes
y_t ~ I (d), d = 2
no
y_t ~ I (d), d > 2

3.4 Representation Theorems for Processes I (1)

Representation theorems – whose role in unit-root econometrics mirrors that of the final form for the dynamic models of structural econometrics – are concerned with the closed form solutions of VAR models in presence of unit-roots, with the inherent reading keys in terms of integrated vs. stationary components of the solutions and cointegration effects as possible offspring. Such theorems, after Granger's seminal work and the major contributions due to the school named after Johansen, stand as a milestone of the so-called time series econometrics.

Although the stage is by and large already set up, the underlying analytical details still present some subtle facets, which have actually hindered, in some respects, a fully satisfactory treatment of the whole matter.

The remaining of this chapter will be specifically devoted to representation theorems with the aim of shedding proper light on the subject after the considerations developed so far. The clarity of the statements and the fluent structure of the proofs are indebted to the innovative as well as rigorous algebraic foundations drawn up in the first chapter. Hence, an elegant reappraisal of classical results is combined with original contributions, widening and enriching both the content and the significance of the theorems presented.

This section deals with $I(1)$ processes, while the next will cover $I(2)$ processes.

Theorem 1

Consider a VAR model specified as follows

$$\underset{(n,\,n)}{\boldsymbol{A}(L)}\,\boldsymbol{y}_t = \boldsymbol{\varepsilon}_t + \boldsymbol{\eta},\;\; \boldsymbol{\varepsilon}_t \sim WN \tag{3.102}$$

where $\boldsymbol{\eta}$ *is a vector of constant terms and*

$$\boldsymbol{A}(L) = \sum_{j=0}^{K} \boldsymbol{A}_j L^j,\;\; \boldsymbol{A}_0 = \boldsymbol{I},\;\; \boldsymbol{A}_K \neq \boldsymbol{0} \tag{3.103}$$

is a matrix polynomial whose characteristic polynomial det $\boldsymbol{A}(z)$ *is assumed to have a possibly repeated unit-root with all other roots lying outside the unit circle.*

Let

$$det \begin{bmatrix} -\dot{\boldsymbol{A}} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix} \neq 0 \tag{3.104}$$

*where **B** and **C** are defined according to the rank factorization*

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}' \tag{3.105}$$

of the singular matrix $\boldsymbol{A}(1) = \boldsymbol{A} \neq \boldsymbol{0}$.

Finally, define

$$\begin{bmatrix} \mathfrak{I}_1 & \mathfrak{I}_2 \\ \mathfrak{I}_3 & \mathfrak{I}_4 \end{bmatrix} = \begin{bmatrix} -\dot{\boldsymbol{A}} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix}^{-1} \tag{3.106}$$

Then, the following representation holds for the process engendered by the model (3.102) above

$$\boldsymbol{y}_t = \boldsymbol{k}_0 + \boldsymbol{k}_1 t + \boldsymbol{N}_1 \sum_{\tau \leq t} \boldsymbol{\varepsilon}_\tau + \sum_{i=0}^{\infty} \boldsymbol{M}_i \boldsymbol{\varepsilon}_{t-i} \tag{3.107}$$

with

$$\boldsymbol{N}_1 = \mathfrak{I}_1 = -\boldsymbol{C}_\perp (\boldsymbol{B}'_\perp \dot{\boldsymbol{A}} \boldsymbol{C}_\perp)^{-1} \boldsymbol{B}'_\perp \tag{3.108}$$

$$\boldsymbol{k}_0 = \boldsymbol{N}_1 \boldsymbol{c} + \boldsymbol{M} \boldsymbol{\eta} \tag{3.109}$$

$$\boldsymbol{k}_1 = \boldsymbol{N}_1 \boldsymbol{\eta} \tag{3.110}$$

*where **c** is arbitrary, $\boldsymbol{M}_0, \boldsymbol{M}_1, \ldots,$ are coefficient matrices whose entries decrease at an exponential rate, and **M** is a short form for $\sum_{i=0}^{\infty} \boldsymbol{M}_i$.*

Note that solution (3.107) represents an integrated process which turns out to be cointegrated, that is

(1) $$\boldsymbol{y}_t \sim I(1) \Rightarrow \nabla \boldsymbol{y}_t \sim I(0) \tag{3.111}$$

(2) $$\boldsymbol{C}'\boldsymbol{y}_t \sim I(0) \Rightarrow \boldsymbol{y}_t \sim CI(1, 1), \; cir = r(\boldsymbol{A}) \tag{3.112}$$

where cir stands for cointegration rank.

Proof

The relationship (3.102) is nothing but a linear difference equation whose general solution can be represented as (see Theorem 1, Sect. 1.8)

$$\boldsymbol{y}_t = \left\{\begin{array}{l} \textit{complementary} \\ \textit{solution} \end{array}\right\} + \left\{\begin{array}{l} \textit{particular solution of the} \\ \textit{non-homogeneous equation} \end{array}\right\} \tag{3.113}$$

As far as the complementary solution, i.e. the general solution of the reduced equation

$$\boldsymbol{A}\,(L)\,\boldsymbol{y}_t = \boldsymbol{0} \tag{3.114}$$

is concerned, we have to distinguish between a permanent component associated with a (possibly repeated) unit-root and a transitory component associated with the roots of the characteristic polynomial $det\,\boldsymbol{A}\,(z)$ lying outside the unit circle.

By referring back to Theorem 3 of Sect. 1.8, the permanent component can be expressed as

$$\bar{\boldsymbol{\xi}}_t = \boldsymbol{N}_1\boldsymbol{c} \tag{3.115}$$

or, in view of (3.108), also as

$$\bar{\boldsymbol{\xi}}_t = \boldsymbol{C}_{\perp}\tilde{\boldsymbol{c}} \tag{3.116}$$

by taking

$$\tilde{\boldsymbol{c}} = -(\boldsymbol{B}'_{\perp}\,\dot{\boldsymbol{A}}\,\boldsymbol{C}_{\perp})^{-1}\,\boldsymbol{B}'_{\perp}\boldsymbol{c} \tag{3.117}$$

The transitory component of the complementary solution can be expressed as

$$\tilde{\boldsymbol{\xi}}_t = \boldsymbol{H}\boldsymbol{\lambda}^{(t)} \tag{3.118}$$

where the symbols have the same meaning as in (3.14) – (3.15) of Sect. 3.1 and reference should be made to all roots of the characteristic equation

$$det\,\boldsymbol{A}\,(z) = 0 \tag{3.119}$$

except for the unit-roots.

Given that the elements of $\boldsymbol{\lambda}^{(t)}$ decrease at an exponential rate, the contribution of component $\tilde{\boldsymbol{\xi}}_t$ turns out to be ultimately immaterial, and as such it is ignored by the closed-form solution of equation (3.102) in the right-hand side of (3.107).

As far as the search for a particular solution of the non-homogeneous equation (3.102) in the statement of the theorem is concerned, we can refer back to the proof of Theorem 2 of Sect. 1.8 and write

$$\boldsymbol{y}_t = \boldsymbol{A}^{-1}(L)(\boldsymbol{\eta} + \boldsymbol{\varepsilon}_t) \tag{3.120}$$

accordingly.

Because of (3.104), $\boldsymbol{A}^{-1}(z)$ has a simple pole located at $z = 1$ (cf. Theorem 3 and Corollary 3.1 of Sect. 1.10) and expansion (1.306) of Sect. 1.7 holds true. This together with the isomorphism of algebras of polynomials in complex-variable and lag-operator arguments, leads to the formal Laurent expansion

$$\boldsymbol{A}^{-1}(L) = \frac{1}{(I-L)}\boldsymbol{N}_1 + \boldsymbol{M}(L) \tag{3.121}$$

where $\boldsymbol{M}(L) = \sum_{i=0}^{\infty} \boldsymbol{M}_i L^i$, and $\boldsymbol{N}_1$ is given by (1.568) of Sect. 1.11.

Combining (3.120) and (3.121) yields the solution

$$\boldsymbol{y}_t = \frac{1}{(I-L)}\boldsymbol{N}_1(\boldsymbol{\eta} + \boldsymbol{\varepsilon}_t) + \boldsymbol{M}(L)(\boldsymbol{\eta} + \boldsymbol{\varepsilon}_t) \tag{3.122}$$

Thanks to sum-calculus operator identities and rules (see (1.382) and (1.384) of Sect. 1.8), namely

$$\frac{1}{(I-L)} = \nabla^{-1} = \sum_{\tau \le t}, \quad \nabla^{-1}\boldsymbol{\eta} = \boldsymbol{\eta} t \tag{3.123}$$

the closed-form expression

$$\boldsymbol{y}_t = \boldsymbol{M}(1)\boldsymbol{\eta} + \boldsymbol{N}_1\boldsymbol{\eta} t + \boldsymbol{N}_1\sum_{\tau \le t}\boldsymbol{\varepsilon}_\tau + \sum_{j=0}^{\infty}\boldsymbol{M}_j\,\boldsymbol{\varepsilon}_{t-j} \tag{3.124}$$

can be derived.

Combining the particular solution on the right-hand side of (3.122) with the permanent component on the right-hand side of (3.115), we eventually obtain for the process $\boldsymbol{y}_t$ the representation

$$\boldsymbol{y}_t = \bar{\boldsymbol{\xi}}_t + \boldsymbol{M}\boldsymbol{\eta} + \boldsymbol{N}_1\boldsymbol{\eta} t + \boldsymbol{N}_1\sum_{\tau \le t}\boldsymbol{\varepsilon}_\tau + \sum_{j=0}^{\infty}\boldsymbol{M}_j\,\boldsymbol{\varepsilon}_{t-j} \tag{3.125}$$

which tallies with (3.107), in light of (3.109) and (3.110).

With respect to results (1) and (2), their proofs rest on the following considerations.

Result (1) – By inspection of (3.107) we deduce that $\boldsymbol{y}_t$ is the resultant of a drift component $\boldsymbol{k}_0$, of a deterministic linear trend component $\boldsymbol{k}_1 t$, of a first order stochastic trend component $\boldsymbol{N}_1 \sum_{\tau \le t} \boldsymbol{\varepsilon}_\tau$, and of a VMA (∞) component in the white noise $\boldsymbol{\varepsilon}_t$. Therefore, the solution $\boldsymbol{y}_t$ displays the connotation of a first order integrated process, and consequently $\nabla \boldsymbol{y}_t$ qualifies as a stationary process.

Result (2) – It ensues from (3.107), in view of (3.108), through premultiplication of both sides by $\boldsymbol{C}'$. Because of the orthogonality of $\boldsymbol{C}'$ with $\boldsymbol{N}_1$ and in view of (3.110), the terms of $\boldsymbol{C}'\boldsymbol{y}_t$ involving both deterministic and stochastic trends, namely $\boldsymbol{C}'\boldsymbol{k}_1 t$ and $\boldsymbol{C}'\boldsymbol{N}_1 \sum_{\tau \le t} \boldsymbol{\varepsilon}_\tau$, disappear. The non stationary terms being annihilated, the resulting process $\boldsymbol{C}'\boldsymbol{y}_t$ turns out to be stationary.

As long as $\boldsymbol{y}_t \sim I(1)$ and also $\boldsymbol{C}'\boldsymbol{y}_t \sim I(0)$, the solution process $\boldsymbol{y}_t$ turns out to be cointegrated and therefore we can write $\boldsymbol{y}_t \sim CI(1, 1)$.

□

The following corollary highlights some interesting features about the stationary process obtained by differencing the solution of the VAR model (3.102). To prepare the road toward deriving the intended results, let us first define the partitioned matrix

$$\boldsymbol{\Pi}(z) = \begin{bmatrix} \boldsymbol{\Pi}_1(z) & \boldsymbol{\Pi}_2(z) \\ \boldsymbol{\Pi}_3(z) & \boldsymbol{\Pi}_4(z) \end{bmatrix} = \begin{bmatrix} \boldsymbol{Q}(z) & \boldsymbol{B} \\ \boldsymbol{C}' & (z-1)\boldsymbol{I} \end{bmatrix}^{-1} \tag{3.126}$$

Formula (3.126) is a meaningful expression and the matrix function $\boldsymbol{\Pi}(z)$ is a matrix polynomial because of the Cayley–Hamilton theorem (see, e.g., Rao and Mitra), provided the inverse in the right-hand side exists, which actually does occur in a neighbourhood of $z = 1$ under the assumptions of Theorem 1 above.

Corollary 1.1

Alternative VMA representations of the stationary process $\nabla \boldsymbol{y}_t$ are

$$\nabla \boldsymbol{y}_t = \boldsymbol{\Pi}_1(L)(\boldsymbol{\eta} + \boldsymbol{\varepsilon}_t) = \boldsymbol{N}_1\boldsymbol{\eta} + \boldsymbol{\Pi}_1(L)\,\boldsymbol{\varepsilon}_t \tag{3.127}$$

$$= \boldsymbol{\delta} + \boldsymbol{\Xi}(L)\,\boldsymbol{\varepsilon}_t \tag{3.128}$$

where $\boldsymbol{\Pi}_1(L)$ is obtained from the leading diagonal block of (3.126) by replacing z with $L,$ while $\boldsymbol{\delta}$ and $\boldsymbol{\Xi}(L)$ are given by

$$\boldsymbol{\delta} = \boldsymbol{N}_1\boldsymbol{\eta} \tag{3.129}$$

$$\boldsymbol{\Xi}(L) = \sum_{j=0}^{\infty}\boldsymbol{\Xi}_j L^j = \boldsymbol{M}(L)\,\nabla + \boldsymbol{N}_1 \tag{3.130}$$

The operator relationship

$$\boldsymbol{M}(L)\nabla = \boldsymbol{\Pi}_1(L) - \boldsymbol{N}_1 \tag{3.131}$$

is also fulfilled.

Furthermore the following statements are true

(1) The matrix polynomial $\boldsymbol{\Xi}(z)$ has a simple zero at $z = 1$

(2) $\nabla \boldsymbol{y}_t$ is a non invertible VMA process

(3) $E(\nabla \boldsymbol{y}_t) = \boldsymbol{N}_1\boldsymbol{\eta}$ (3.132)

Proof

From (1.515) of Sect. 1.10, because of the isomorphism between matrix polynomials in a complex variables z and in the lag operator L, this useful relationship holds true

$$\boldsymbol{A}^{-1}(L)\nabla = [\boldsymbol{I},\ \boldsymbol{0}]\begin{bmatrix}\boldsymbol{Q}(L) & \boldsymbol{B}\\ \boldsymbol{C}' & -\nabla\boldsymbol{I}\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{I}\\ \boldsymbol{0}\end{bmatrix} \tag{3.133}$$

and therefore, in view of (3.120) and by virtue of (3.126), the following VMA representation of $\nabla\boldsymbol{y}_t$

$$\begin{aligned}\nabla\boldsymbol{y}_t &= [\boldsymbol{I},\ \boldsymbol{0}]\begin{bmatrix}\boldsymbol{Q}(L) & \boldsymbol{B}\\ \boldsymbol{C}' & -\nabla\boldsymbol{I}\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{I}\\ \boldsymbol{0}\end{bmatrix}(\boldsymbol{\eta} + \boldsymbol{\varepsilon}_t) = \boldsymbol{\Pi}_1(L)\,(\boldsymbol{\eta} + \boldsymbol{\varepsilon}_t)\\ &= \boldsymbol{N}_1\boldsymbol{\eta} + \boldsymbol{\Pi}_1(L)\,\boldsymbol{\varepsilon}_t\end{aligned} \tag{3.134}$$

is obtained in a straightforward manner, upon noting that

$$\boldsymbol{\Pi}_1(1) = \boldsymbol{N}_1 \tag{3.135}$$

by bearing in mind (1.307) of Sect. 1.7, (1.515) of Sect. 1.10, and (3.126) above.

The VMA representation (3.128) follows from (3.107) by elementary computations.

The equality (3.131) is shown to be true by comparing the right-hand sides of (3.127) and (3.128) in light of (3.130).

For what concerns results (1)–(3), their proofs rest on the following considerations:

Result (1) – The matrix polynomial

$$\Xi\,(z) = (1 - z)\,\boldsymbol{M}\,(z) + \boldsymbol{N}_1 \tag{3.136}$$

has a simple zero at $z = 1$, according to Definition 5 of Sect. 1.6 and by virtue of Theorem 3, along with Corollary 3.1, of Sect. 1.10. Indeed the following hold

$$\Xi\,(1) = \boldsymbol{N}_1 \Rightarrow \; det\,\Xi\,(1) = 0 \tag{3.137}$$

$$\dot{\Xi}\,(1) = -\,\boldsymbol{M}\,(1) \tag{3.138}$$

$$det\,(\boldsymbol{D}'_\perp\,\dot{\Xi}\,(1)\,\boldsymbol{E}_\perp) \neq 0 \Leftrightarrow det \begin{bmatrix} -\dot{\Xi}(1) & \boldsymbol{D} \\ \boldsymbol{E}' & \boldsymbol{0} \end{bmatrix} \neq 0 \tag{3.139}$$

recalling (1.510’) and (1.517) of Sect. 1.10, and taking

$$\boldsymbol{D}_\perp = -\,\boldsymbol{C},\;\; \boldsymbol{E}_\perp = \boldsymbol{B} \tag{3.140}$$

where $\boldsymbol{D}$ and $\boldsymbol{E}$ are defined as per a rank factorization of $\boldsymbol{N}_1$, namely

$$\boldsymbol{N}_1 = \boldsymbol{D}\boldsymbol{E}' \tag{3.141}$$

$$\boldsymbol{D} = -\,\boldsymbol{C}_\perp\,(\boldsymbol{B}'_\perp\,\dot{\boldsymbol{A}}\,\boldsymbol{C}_\perp)^{-1},\;\; \boldsymbol{E} = \boldsymbol{B}_\perp \tag{3.142}$$

Result (2) – The conclusion is easily drawn by inspection of (3.128) because of (1) above.

Result (3) – The proof is straightforward by taking the expected value of both sides of (3.128) in light of (3.129).

□

What is claimed in Theorem 1 and in its Corollary not only reflects but also extends the content of the basic representation theorem of time series econometrics.

This theorem can likewise be given a dual version (see, e.g., Banjeree et al. 1993; Johansen 1995), which originates from a VMA model for the difference process $\nabla \boldsymbol{y}_t$. This in turn underlies a VAR model for the parent process $\boldsymbol{y}_t$ whose integration and cointegration properties can eventually be established (see, in this connection, the considerations made in Sect. 2.5).

Theorem 2

Consider two stochastic processes $\boldsymbol{\xi}_t$ and $\boldsymbol{y}_t$, the former being defined as the finite difference of the latter, namely as

$$\underset{(n,1)}{\boldsymbol{\xi}_t} = \nabla \boldsymbol{y}_t \tag{3.143}$$

and let $\boldsymbol{\xi}_t$ be stationary with the VMA(∞) representation

$$\boldsymbol{\xi}_t = \boldsymbol{\Xi}(L)(\boldsymbol{\eta} + \boldsymbol{\varepsilon}_t) \tag{3.144}$$

Moreover, assume that the parent matrix polynomial

$$\boldsymbol{\Xi}(z) = \boldsymbol{\Xi}_0 + \sum_{i=1}^{\infty} \boldsymbol{\Xi}_i z^i \tag{3.145}$$

has a first-order zero at $z = 1$, and that the coefficient matrices $\boldsymbol{\Xi}_i$ exhibit exponentially decreasing entries.

Then the companion process $\boldsymbol{y}_t$ admits a VAR generating model

$$\boldsymbol{A}(L)\boldsymbol{y}_t = \boldsymbol{\eta} + \boldsymbol{\varepsilon}_t \tag{3.146}$$

such that its parent matrix polynomial $\boldsymbol{A}(z)$ has a first-order zero at $z = 1$, and its characteristic polynomial $\det \boldsymbol{A}(z)$ has, besides a (possibly multiple) unit-root, all other roots lying outside the unit circle.

The engendered process $\boldsymbol{y}_t$ enjoys the integration and cointegration properties

$$\left.\begin{array}{r} \boldsymbol{y}_t \sim I(1) \\ \boldsymbol{C}'\boldsymbol{y}_t \sim I(0) \end{array}\right\} \Rightarrow \boldsymbol{y}_t \sim CI(1, 1) \tag{3.147}$$

where $\boldsymbol{C}'$ is defined as per a rank factorization of $\boldsymbol{A} = \boldsymbol{A}(1)$, that is

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}' \tag{3.148}$$

Proof

In view of (1.247), (1.248) and (1.258) of Sect. 1.6 and of the isomorphism between polynomials in a complex variable z and in the lag operator L, we obtain the paired expansions

$$\Xi(z) = \tilde{\Phi}(z)(1-z) + \Xi(1) \Leftrightarrow \Xi(L) = \tilde{\Phi}(L)\nabla + \Xi(1) \tag{3.149}$$

where

$$\tilde{\Phi}(z) = \sum_{k \geq 1}(-1)^k \frac{1}{k!}\Xi^{(k)}(1)(1-z)^{k-1} \Leftrightarrow$$
$$\Leftrightarrow \tilde{\Phi}(L) = \sum_{k \geq 1}(-1)^k \frac{1}{k!}\Xi^{(k)}(1)\nabla^{k-1} \tag{3.150}$$

$$\tilde{\Phi}(1) = -\dot{\Xi}(1) \tag{3.151}$$

Next, in light of Definition 5 of Sect. 1.6 and by virtue of Theorem 3 and Corollary 3.1 of Sect. 1.10, the following hold true

$$det\, \Xi(1) = 0 \tag{3.152}$$

$$det\left(\begin{bmatrix} -\dot{\Xi}(1) & \tilde{\boldsymbol{B}} \\ \tilde{\boldsymbol{C}}' & \boldsymbol{0} \end{bmatrix}\right) \neq 0 \Leftrightarrow det\,(\tilde{\boldsymbol{B}}'_{\perp}\dot{\Xi}(1)\tilde{\boldsymbol{C}}_{\perp}) \neq 0 \tag{3.153}$$

where $\tilde{\boldsymbol{B}}$ and $\tilde{\boldsymbol{C}}$ are defined by a rank factorization of $\Xi(1)$, such as

$$\Xi(1) = \tilde{\boldsymbol{B}}\tilde{\boldsymbol{C}}' \tag{3.154}$$

Given these premises, expansion (1.306) of Sect. 1.7 holds for $\Xi^{-1}(z)$ in a deleted neighbourhood of the simple pole $z = 1$, and we can then write

$$\Xi^{-1}(z) = \frac{1}{(1-z)}\tilde{\boldsymbol{N}}_1 + \tilde{\boldsymbol{M}}(z) \Leftrightarrow \Xi^{-1}(L) = \tilde{\boldsymbol{N}}_1\nabla^{-1} + \tilde{\boldsymbol{M}}(L) \tag{3.155}$$

where

$$\tilde{\boldsymbol{N}}_1 = -\tilde{\boldsymbol{C}}_{\perp}(\tilde{\boldsymbol{B}}'_{\perp}\dot{\Xi}(1)\tilde{\boldsymbol{C}}_{\perp})^{-1}\tilde{\boldsymbol{B}}'_{\perp} \tag{3.156}$$

$$\tilde{\boldsymbol{M}}(1) = -\frac{1}{2}\tilde{\boldsymbol{N}}_1\ddot{\Xi}(1)\tilde{\boldsymbol{N}}_1 + (\boldsymbol{I} + \tilde{\boldsymbol{N}}_1\dot{\Xi}(1))\,\Xi^{g}(1)\,(\boldsymbol{I} + \dot{\Xi}(1)\tilde{\boldsymbol{N}}_1) \tag{3.157}$$

in view of Theorem 1 of Sect. 1.11.

Applying the operator $\boldsymbol{\Xi}^{-1}(L)$ to both sides of (3.144) yields

$$\boldsymbol{\Xi}^{-1}(L)\,\boldsymbol{\xi}_t = \boldsymbol{\eta} + \boldsymbol{\varepsilon}_t \Rightarrow \boldsymbol{A}\,(L)\,\boldsymbol{y}_t = \boldsymbol{\eta} + \boldsymbol{\varepsilon}_t \tag{3.158}$$

namely the VAR representation (3.146), whose parent matrix polynomial

$$\boldsymbol{A}\,(z) = \boldsymbol{\Xi}^{-1}(z)\,(1-z) = (1-z)\,\tilde{\boldsymbol{M}}\,(z) + \tilde{\boldsymbol{N}}_1 \tag{3.159}$$

turns out to have a first order zero at $z = 1$.

This is due to the following reasons:

(1)
$$\boldsymbol{A} = \tilde{\boldsymbol{N}}_1 \Rightarrow det\,\boldsymbol{A} = 0 \tag{3.160}$$

whence the rank factorization

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}' \tag{3.161}$$

where $\boldsymbol{B}$ and $\boldsymbol{C}$ are conveniently chosen as

$$\boldsymbol{B} = -\,\tilde{\boldsymbol{C}}_\perp\,(\tilde{\boldsymbol{B}}'_\perp\dot{\boldsymbol{\Xi}}\,(1)\,\tilde{\boldsymbol{C}}_\perp)^{-1},\ \ \boldsymbol{C} = \tilde{\boldsymbol{B}}_\perp \tag{3.162}$$

(2) A possible choice for matrices $\boldsymbol{B}_\perp$ and $\boldsymbol{C}_\perp$, in light of (3.162), will therefore be

$$\boldsymbol{B}_\perp = \tilde{\boldsymbol{C}},\ \ \boldsymbol{C}_\perp = \tilde{\boldsymbol{B}} \tag{3.163}$$

(3)
$$\dot{\boldsymbol{A}} = -\,\tilde{\boldsymbol{M}}\,(1) \tag{3.164}$$

(4)
$$det\,(\boldsymbol{B}'_\perp\,\dot{\boldsymbol{A}}\,\boldsymbol{C}_\perp) \neq 0 \Leftrightarrow det\left(\begin{bmatrix} -\dot{\boldsymbol{A}} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix}\right) \neq 0 \tag{3.165}$$

upon noting that, according to (3.163) and (3.164) above and (1.583) of Sect. 1.11, the following identity holds

$$\boldsymbol{B}'_\perp\,\dot{\boldsymbol{A}}\,\boldsymbol{C}_\perp = -\tilde{\boldsymbol{C}}'\,\tilde{\boldsymbol{M}}\,(1)\,\tilde{\boldsymbol{B}} = -\boldsymbol{I} \tag{3.166}$$

The proof of what has been claimed about the roots of the characteristic polynomial $det\,\boldsymbol{A}(z)$ rests on what was shown in Theorem 4 of Sect. 1.7.

Finally, insofar as $\boldsymbol{A}\,(z)$ has a first order zero at $z = 1$ in light of (3.160) and (3.165) above, $\boldsymbol{A}^{-1}(z)$ has a first order pole at $z = 1$.

Hence, by Theorem 1 of Sect. 2.3 the VAR model (3.146) engenders an integrated process $\boldsymbol{y}_t$ of the first order, i.e. $\boldsymbol{y}_t \sim I\,(1)$.

Indeed, according to (3.159), the following Laurent expansions holds

$$\begin{aligned} \boldsymbol{A}(L)^{-1} &= \frac{1}{(I-L)}\boldsymbol{N}_1 + \boldsymbol{M}(L) = \Xi(L)\nabla^{-1} \\ &= \frac{1}{(I-L)}\Xi(1) + \tilde{\Phi}(L) \end{aligned} \tag{3.167}$$

Then, in light of (3.146), (3.154) and (3.167) the following holds true

$$\begin{aligned} \boldsymbol{y}_t &= \boldsymbol{A}(L)^{-1}(\boldsymbol{\eta} + \boldsymbol{\varepsilon}_t) = [\Xi(1)\nabla^{-1} + \tilde{\Phi}(L)](\boldsymbol{\eta} + \boldsymbol{\varepsilon}_t) \\ &= \sum_{i=0}^{\infty} \tilde{\Phi}_i\, \boldsymbol{\varepsilon}_{t-i} + \tilde{\boldsymbol{B}}\,\tilde{\boldsymbol{C}}' \sum_{\tau \le t} \boldsymbol{\varepsilon}_\tau + \tilde{\boldsymbol{B}}\,\tilde{\boldsymbol{C}}'\boldsymbol{\eta} t + \boldsymbol{k} \end{aligned} \tag{3.168}$$

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

Furthermore, because of representation (3.168), $\tilde{\boldsymbol{B}}_\perp = \boldsymbol{C}$ plays the role of the matrix of the cointegration vectors since

$$\tilde{\boldsymbol{B}}'_\perp \boldsymbol{y}_t = \boldsymbol{C}'\boldsymbol{y}_t \sim I(0) \tag{3.169}$$

Finally, the cointegration property, $\boldsymbol{y}_t \sim CI(1, 1)$, holds as a by-product.

□

3.5 Representation Theorems for Processes *I* (2)

Following the same path of reasoning as in the Sect. 3.4, we will provide a neat and rigorous formulation of the representation theorem for $I(2)$ processes, to be followed – as corollaries – by some interesting related results.

To conclude we will show how to derive a dual form of this theorem.

Theorem 1

Consider a VAR model specified as

$$\underset{(n,\,n)}{\boldsymbol{A}(L)}\,\boldsymbol{y}_t = \boldsymbol{\varepsilon}_t + \boldsymbol{\eta}, \;\; \boldsymbol{\varepsilon}_t \sim WN \tag{3.170}$$

where $\boldsymbol{\eta}$ *is a vector of constant terms and*

$$\boldsymbol{A}(L) = \sum_{j=0}^{K} \boldsymbol{A}_j L^j, \;\; \boldsymbol{A}_0 = \boldsymbol{I}, \boldsymbol{A}_K \neq \boldsymbol{0} \tag{3.171}$$

is a matrix polynomial whose characteristic polynomial det $\boldsymbol{A}(z)$ *is assumed to have a multiple unit-root with all other roots lying outside the unit circle.*

Let

(a) $$det \begin{bmatrix} -\dot{\boldsymbol{A}} & \boldsymbol{B} \\ \boldsymbol{C}' & \boldsymbol{0} \end{bmatrix} = 0 \tag{3.172}$$

where $\boldsymbol{B}$ *and* $\boldsymbol{C}$ *are defined as per the rank factorization*

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}' \tag{3.173}$$

of the singular matrix $\boldsymbol{A}(1) = \boldsymbol{A} \neq \boldsymbol{0}$*;*

(b) $$det \begin{bmatrix} \tilde{\boldsymbol{A}} & (\boldsymbol{B}_\perp \boldsymbol{R}_\perp)_\perp \\ (\boldsymbol{C}_\perp \boldsymbol{S}_\perp)_\perp & \boldsymbol{0} \end{bmatrix} \neq 0 \tag{3.174}$$

where

$$\tilde{\boldsymbol{A}} = \frac{1}{2}\ddot{\boldsymbol{A}} - \dot{\boldsymbol{A}}\boldsymbol{A}^g\dot{\boldsymbol{A}} \tag{3.175}$$

and $\boldsymbol{R}$ *and* $\boldsymbol{S}$ *are defined as per the rank factorization*

$$\tilde{\boldsymbol{B}}'_\perp \dot{\boldsymbol{A}} \boldsymbol{C}_\perp = \boldsymbol{R}\boldsymbol{S}' \tag{3.176}$$

of the singular matrix $\tilde{\boldsymbol{B}}'_\perp \dot{\boldsymbol{A}} \boldsymbol{C}_\perp \neq \boldsymbol{0}$ *with* $(\boldsymbol{B}_\perp \boldsymbol{R}_\perp)_\perp$ *and* $(\boldsymbol{C}_\perp \boldsymbol{S}_\perp)_\perp$ *standing for* $[\boldsymbol{B}, (\boldsymbol{B}'_\perp)^g\boldsymbol{R}]$ *and* $[\boldsymbol{C}, (\boldsymbol{C}'_\perp)^g\boldsymbol{S}]$*, respectively.*

Moreover, define

$$\begin{bmatrix} \boldsymbol{J}_1 & \boldsymbol{J}_2 \\ \boldsymbol{J}_3 & \boldsymbol{J}_4 \end{bmatrix} = \begin{bmatrix} \tilde{\boldsymbol{A}} & (\boldsymbol{B}_\perp \boldsymbol{R}_\perp)_\perp \\ (\boldsymbol{C}_\perp \boldsymbol{S}_\perp)'_\perp & \boldsymbol{0} \end{bmatrix}^{-1} \tag{3.177}$$

Then, the following representation holds for the solution of (3.170) above

$$\boldsymbol{y}_t = \boldsymbol{k}_0 + \boldsymbol{k}_1 t + \boldsymbol{k}_2 t^2 + \boldsymbol{N}_1 \sum_{\tau \le t} \boldsymbol{\varepsilon}_\tau + \boldsymbol{N}_2 \sum_{\tau \le t} (t + 1 - \tau)\boldsymbol{\varepsilon}_\tau + \sum_{i=0}^{\infty} \boldsymbol{M}_i\, \boldsymbol{\varepsilon}_{t-i} \tag{3.178}$$

with

$$\boldsymbol{N}_2 = \boldsymbol{J}_1 = \boldsymbol{C}_\perp \boldsymbol{S}_\perp (\boldsymbol{R}'_\perp \boldsymbol{B}'_\perp \tilde{\boldsymbol{A}} \boldsymbol{C}_\perp \boldsymbol{S}_\perp)^{-1} \boldsymbol{R}'_\perp \boldsymbol{B}'_\perp \tag{3.179}$$

$$N_1 = [N_2,\ I - N_2 \tilde{A}] \begin{bmatrix} \tilde{\tilde{A}} & \dot{A}A^g \\ A^g\dot{A} & -C_\perp (B'_\perp \dot{A} C_\perp)^g B'_\perp \end{bmatrix} \begin{bmatrix} N_2 \\ I - \tilde{A}N_2 \end{bmatrix} \tag{3.180}$$

$$k_0 = N_1 c + N_2 d + M \eta \tag{3.181}$$

$$k_1 = N_1 \eta + N_2 (c + \frac{1}{2}\eta) \tag{3.182}$$

$$k_2 = \frac{1}{2} N_2 \eta \tag{3.183}$$

where $\boldsymbol{c}$ *and* $\boldsymbol{d}$ *are arbitrary vectors,* $M_0, M_1, \ldots,$ *are coefficient matrices whose entries decrease at an exponential rate,* $\boldsymbol{M}$ *is a short form for* $\sum_{i=0}^{\infty} M_i$ *and* $\tilde{\tilde{A}}$ *is defined as*

$$\tilde{\tilde{A}} = \frac{1}{6} \dddot{A} - \dot{A} A^g \dot{A} A^g \dot{A} \tag{3.184}$$

The following properties of the solution y_t *are satisfied*

(1) *The vector* y_t *is an integrated process, that is*

$$y_t \sim I(2) \tag{3.185}$$

(2) *The vector* y_t *is a cointegrated process, in accordance with the relationships*

$$(C_\perp S_\perp)'_\perp y_t \sim I(\gamma),\ \gamma < 2,\ \ y_t \sim CI(2, 2-\gamma),\ cir = r(A) + r(B'_\perp \dot{A} C_\perp) \tag{3.186}$$

$$(B^g V)'_\perp C' y_t \sim I(0) \rightarrow y_t \sim CI(2,2),\ \ cir = r(A) - r(V) \tag{3.187}$$

$$\left[CB^g V, (C'_\perp)^g S]\right]' y_t \sim I(1) \rightarrow y_t \sim CI(2,1),\ \ cir = r(B'_\perp \dot{A} C_\perp) + r(V) \tag{3.188}$$

where cir stands for cointegration rank, and $\boldsymbol{V}$ *is defined according to the rank factorization*

$$\dot{A}C_{\perp}S_{\perp} = VW' \tag{3.189}$$

(3) The vector y_t is a polynomially cointegrated process, as long as

$$V'(B^g)'C'y_t - V'(BB')^{\neq}\dot{A}\nabla y_t \sim I(0) \rightarrow y_t \sim PCI(2,2) \tag{3.190}$$

where $(BB')^{\neq}$ is the Drazin inverse of BB'.

Proof

The structure of the proof mirrors that of the representation theorem stated in the Sect. 3.3.

Once again relationship (3.170) reads as a linear difference equation, whose general solution can be partitioned as (see Theorem 1 of Sect. 1.8).

$$y_t = \begin{Bmatrix} complementary \\ solution \end{Bmatrix} + \begin{Bmatrix} particular\ solution\ of\ the \\ non\text{-}homogeneous\ equation \end{Bmatrix} \tag{3.191}$$

As usual, for what concerns the complementary solution, i.e. the (general) solution of the reduced equation

$$A(L)\, y_t = 0 \tag{3.192}$$

we have to distinguish between a permanent component associated with the unit-roots and a transitory component associated with the roots of the characteristic polynomial $det\, A(z)$ lying outside the unit circle.

According to Theorem 3 of Sect. 1.8, the permanent component turns out to take the form

$$\bar{\xi}_t = N_1 c + N_2 d + N_2 ct \tag{3.193}$$

where c and d are arbitrary, whereas the transitory component $\tilde{\xi}_t$ takes the same form as in formula (3.118) of Sect. 3.4 and is likewise ignored by the closed-form solution of equation (3.170) on the right-hand side of (3.178).

When searching for a particular solution for the non-homogeneous equation (3.170) in the statement of the theorem, by resorting to Theorem 2 of Sect. 1.8, we get

$$y_t = A^{-1}(L)\,(\eta + \varepsilon_t) \tag{3.194}$$

Owing to (3.172) and (3.174), $A^{-1}(z)$ has a second-order pole at $z = 1$ (cf. Theorem 5 and Corollary 5.1 of Sect. 1.10) and expansion (1.318) of

Sect. 1.7 holds accordingly. This, together with the isomorphism of algebras of polynomials in a complex variable and in the lag-operator L, leads to the formal Laurent expansion,

$$\boldsymbol{A}^{-1}(L) = \frac{1}{(I-L)^2}\boldsymbol{N}_2 + \frac{1}{(I-L)}\boldsymbol{N}_1 + \boldsymbol{M}(L) \tag{3.195}$$

where $\boldsymbol{M}(L) = \sum_{i=0}^{\infty} \boldsymbol{M}_i L^i$, and $\boldsymbol{N}_2$ and $\boldsymbol{N}_1$ are given by (1.587) and (1.588') of Sect. 1.11, respectively.

Combining (3.194) and (3.195) yields the solution

$$\boldsymbol{y}_t = \frac{1}{(I-L)^2}\boldsymbol{N}_2(\boldsymbol{\eta} + \boldsymbol{\varepsilon}_t) + \frac{1}{(I-L)}\boldsymbol{N}_1(\boldsymbol{\eta} + \boldsymbol{\varepsilon}_t) + \boldsymbol{M}(L)(\boldsymbol{\eta} + \boldsymbol{\varepsilon}_t) \tag{3.196}$$

Thanks to sum-calculus operator identities and rules (see (1.382), (1.383), (1.383') and (1.384) of Sect. 1.8), namely

$$\frac{1}{(I-L)} = \nabla^{-1} = \sum_{\tau \le t}, \quad \nabla^{-1}\boldsymbol{\eta} = \boldsymbol{\eta} t \tag{3.197}$$

$$\frac{1}{(I-L)^2} = \nabla^{-2} = \sum_{\vartheta \le t}\sum_{\tau \le \vartheta} = \sum_{\tau \le t}(t+1-\tau), \; \nabla^{-2}\boldsymbol{\eta} = \frac{1}{2}\boldsymbol{\eta}(t+1)t \tag{3.197'}$$

the closed-form expression

$$\begin{aligned}\boldsymbol{y}_t = \boldsymbol{M}\boldsymbol{\eta} + (\boldsymbol{N}_1 + \frac{1}{2}\boldsymbol{N}_2)\boldsymbol{\eta} t + \frac{1}{2}\boldsymbol{N}_2\boldsymbol{\eta} t^2 + \boldsymbol{N}_1\sum_{\tau \le t}\boldsymbol{\varepsilon}_\tau \\ + \boldsymbol{N}_2\sum_{\tau \le t}(t+1-\tau)\boldsymbol{\varepsilon}_\tau + \sum_{i=0}^{\infty}\boldsymbol{M}_i\boldsymbol{\varepsilon}_{t-i}\end{aligned} \tag{3.198}$$

is derived.

A closed-form expression of $\boldsymbol{M}(1)$ can be derived following the reasoning of Theorem 2 of Sect. 1.11.

Combining the particular solution in the right-hand side of (3.198) of the non-homogeneous equation (3.170), with the permanent component in the right-hand side of (3.193) of the complementary solution, we eventually derive for the process $\boldsymbol{y}_t$ the representation

$$\begin{aligned}\boldsymbol{y}_t = \overline{\xi}_t + \boldsymbol{M}\boldsymbol{\eta} + (\boldsymbol{N}_1 + \frac{1}{2}\boldsymbol{N}_2)\boldsymbol{\eta} t + \frac{1}{2}\boldsymbol{N}_2\boldsymbol{\eta} t^2 + \boldsymbol{N}_1\sum_{\tau \le t}\boldsymbol{\varepsilon}_\tau \\ + \boldsymbol{N}_2\sum_{\tau \le t}(t+1-\tau)\boldsymbol{\varepsilon}_\tau + \sum_{i=0}^{\infty}\boldsymbol{M}_i\boldsymbol{\varepsilon}_{t-i}\end{aligned} \tag{3.199}$$

which tallies with (3.178), in light of (3.181), (3.182) and (3.183).

As far as results (1)–(3) are concerned, their proofs rest on the following considerations.

Result (1) – By inspection of (3.178) we deduce that $\boldsymbol{y}_t$ is the resultant of a drift component $\boldsymbol{k}_0$, of linear and quadratic deterministic trends, $\boldsymbol{k}_1 t$ and $\boldsymbol{k}_2 t^2$ respectively, of first and second order stochastic trend components, $\boldsymbol{N}_1 \sum_{\tau \le t} \boldsymbol{\varepsilon}_\tau$ and $\boldsymbol{N}_2 \sum_{\tau \le t} (t+1-\tau)\,\boldsymbol{\varepsilon}_\tau$ respectively, and of a VMA(∞) component in the white noise argument $\boldsymbol{\varepsilon}_t$. The overall effect is that of a second order integrated process $\boldsymbol{y}_t$ whence $\nabla^2 \boldsymbol{y}_t$ qualifies as a stationary process.

Result (2) – First of all observe that the matrix $(\boldsymbol{C}_\perp \boldsymbol{S}_\perp)'_\perp$ is orthogonal to $\boldsymbol{N}_2$ as specified in (3.179) and related to $\boldsymbol{k}_2$ as in (3.183). Next, keep in mind that according to Corollary 2.1 of Sect. 1.11, this very matrix can be split into three blocks in this way

$$(\boldsymbol{C}_\perp \boldsymbol{S}_\perp)'_\perp = \begin{bmatrix} (\boldsymbol{B}^g \boldsymbol{V})'_\perp \boldsymbol{C}' \\ \boldsymbol{V}'(\boldsymbol{B}^g)'\boldsymbol{C}' \\ \boldsymbol{S}'\boldsymbol{C}^g_\perp \end{bmatrix} \tag{3.200}$$

where the upper block is orthogonal to $[\boldsymbol{N}_1, \boldsymbol{N}_2]$ and to $\boldsymbol{k}_1$.

Pre-multiplying (3.178) by the partitioned matrix (3.200) yields the composite process

$$(\boldsymbol{C}_\perp \boldsymbol{S}_\perp)'_\perp \boldsymbol{y}_t = \begin{bmatrix} (\boldsymbol{B}^g \boldsymbol{V})'_\perp \boldsymbol{C}' \boldsymbol{y}_t \\ \boldsymbol{V}'(\boldsymbol{B}^g)'\boldsymbol{C}' \boldsymbol{y}_t \\ \boldsymbol{S}'\boldsymbol{C}^g_\perp \boldsymbol{y}_t \end{bmatrix}$$

$$= \begin{bmatrix} (\boldsymbol{B}^g \boldsymbol{V})'_\perp \boldsymbol{C}' \boldsymbol{M}\boldsymbol{\eta} + (\boldsymbol{B}^g \boldsymbol{V})'_\perp \boldsymbol{C}' \sum_{i=0}^{\infty} \boldsymbol{M}_i \boldsymbol{\varepsilon}_{t-i} \\ \boldsymbol{V}'(\boldsymbol{B}^g)'\boldsymbol{C}'(\boldsymbol{M}\boldsymbol{\eta} + \boldsymbol{N}_1 \boldsymbol{c})\boldsymbol{\eta} + \boldsymbol{V}'(\boldsymbol{B}^g)'\boldsymbol{C}'\boldsymbol{N}_1(\boldsymbol{\eta} t + \sum_{\tau \le t} \boldsymbol{\varepsilon}_\tau) + \boldsymbol{V}'(\boldsymbol{B}^g)'\boldsymbol{C}' \sum_{i=0}^{\infty} \boldsymbol{M}_i \boldsymbol{\varepsilon}_{t-i} \\ \boldsymbol{S}'\boldsymbol{C}^g_\perp(\boldsymbol{M}\boldsymbol{\eta} + \boldsymbol{N}_1 \boldsymbol{c}) + \boldsymbol{S}'\boldsymbol{C}^g_\perp \boldsymbol{N}_1(\boldsymbol{\eta} t + \sum_{\tau \le t} \boldsymbol{\varepsilon}_\tau) + \boldsymbol{S}'\boldsymbol{C}^g_\perp \sum_{i=0}^{\infty} \boldsymbol{M}_i \boldsymbol{\varepsilon}_{t-i} \end{bmatrix} \tag{3.201}$$

where the second-order (stochastic and deterministic) trends are no longer present, whereas the first-order trends are still present in the middle and bottom blocks, but not in the upper one.

The integration order of $(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp}\boldsymbol{y}_t$ decreases accordingly, dropping to zero in the upper block.

Finally, $\boldsymbol{y}_t$ turns out to be a cointegrated processes, as claimed in (3.186), whose cointegration rank equals the rank of $(\boldsymbol{C}_{\perp}\boldsymbol{S}_{\perp})'_{\perp}$, namely, $r(\boldsymbol{A}) + r(\boldsymbol{B}'_{\perp}\dot{\boldsymbol{A}}\boldsymbol{C}_{\perp})$ (see (1.624) of Sect. 1.11).

Moving to (3.187) and (3.188), the results stated above follow from (3.201) and from (1.624) and (1.625) of Sect. 1.11.

Result (3) – To prove (3.190) we may proceed by pointing out that

(a) On the one hand, by resorting to (3.201) it is easy to check that

$$\boldsymbol{V}'(\boldsymbol{B}^g)'\boldsymbol{C}'\boldsymbol{y}_t = \boldsymbol{V}'(\boldsymbol{B}^g)'\boldsymbol{C}'\boldsymbol{N}_1(\boldsymbol{\eta}t + \sum_{\tau \le t}\boldsymbol{\varepsilon}_\tau) + \textit{stationary components} \tag{3.202}$$

(b) On the other hand, differencing both sides of (3.178) gives

$$\nabla \boldsymbol{y}_t = \boldsymbol{N}_2\,(\boldsymbol{\eta}t + \sum_{\tau \le t}\boldsymbol{\varepsilon}_\tau) + \textit{stationary components} \tag{3.203}$$

By inspection of the right-hand sides of (3.202) and (3.203) it is apparent that both $\boldsymbol{V}'(\boldsymbol{B}^g)'\boldsymbol{C}'\boldsymbol{y}_t$ and $\nabla\boldsymbol{y}_t$ are $I(1)$ processes, since both expressions exhibit (the same) first-order stochastic and deterministic trends. This suggests the possibility of arriving at a stationary process by a linear combination of these processes.

Now, let $\boldsymbol{D}$ be a matrix with the same dimension as $\boldsymbol{V}'(\boldsymbol{B}^g)'\boldsymbol{C}'$, and consider the linear form

$$\boldsymbol{V}'(\boldsymbol{B}^g)'\boldsymbol{C}'\boldsymbol{y}_t + \boldsymbol{D}\nabla\boldsymbol{y}_t = [\boldsymbol{V}'(\boldsymbol{B}^g)'\boldsymbol{C}'\boldsymbol{N}_1 + \boldsymbol{D}\boldsymbol{N}_2]\,(\boldsymbol{\eta}t + \sum_{\tau \le t}\boldsymbol{\varepsilon}_\tau) + \textit{stationary components} \tag{3.204}$$

The deterministic and stochastic trends on the right-hand side of (3.204) will vanish provided the following equality holds

$$\boldsymbol{V}'(\boldsymbol{B}^g)'\boldsymbol{C}'\boldsymbol{N}_1 + \boldsymbol{D}\boldsymbol{N}_2 = \boldsymbol{0} \tag{3.205}$$

which occurs when

$$\boldsymbol{D} = -\boldsymbol{V}'(\boldsymbol{B}\boldsymbol{B}')^{\neq}\dot{\boldsymbol{A}} \tag{3.206}$$

as simple computations show by resorting to (1.327) of Sect. 1.7, bearing in mind (1.22) and (1.27) of Sect. 1.1.

Hence the process $\boldsymbol{V}'(\boldsymbol{B}^g)'\boldsymbol{C}'\boldsymbol{y}_t - \boldsymbol{V}'(\boldsymbol{B}\boldsymbol{B}')^{\neq}\dot{\boldsymbol{A}}\nabla\boldsymbol{y}_t$ is free from non stationary components and (3) holds accordingly.

□

The following corollary provides an insight into the stationary process obtained by differencing twice the solution of the VAR model (3.170).

In this connection, let us first introduce the partitioned matrix

$$\boldsymbol{\Pi}(z) = \begin{bmatrix} \boldsymbol{\Pi}_1(z) & \boldsymbol{\Pi}_2(z) \\ \boldsymbol{\Pi}_3(z) & \boldsymbol{\Pi}_4(z) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Psi}(z) & \boldsymbol{F} \\ \boldsymbol{G}' & \boldsymbol{\Lambda}(z) \end{bmatrix}^{-1} \tag{3.207}$$

where $\boldsymbol{\Psi}(z)$, $\boldsymbol{F}$, $\boldsymbol{G}$ and $\boldsymbol{\Lambda}(z)$ are as defined by (1.255) of Sect. 1.6 and by (1.543) and (1.545) of Sect. 1.10.

Formula (3.207) is a meaningful expression and the matrix function $\boldsymbol{\Pi}(z)$ is a matrix polynomial because of the Cayley–Hamilton theorem (see, e.g., Rao and Mitra), provided the inverse in the right-hand side exists, which actually occurs in a neighbourhood of $z = 1$ under the assumptions of Theorem 1 above.

Corollary 1.1

Alternative VMA representations of the stationary process $\nabla^2\boldsymbol{y}_t$ are

$$\nabla^2\boldsymbol{y}_t = \boldsymbol{\Pi}_1(L)(\boldsymbol{\eta} + \boldsymbol{\varepsilon}_t) = \boldsymbol{N}_2\boldsymbol{\eta} + \boldsymbol{\Pi}_1(L)\boldsymbol{\varepsilon}_t \tag{3.208}$$

$$= \boldsymbol{\delta} + \boldsymbol{\Xi}(L)\boldsymbol{\varepsilon}_t \tag{3.209}$$

where $\boldsymbol{\Pi}_1(L)$ is obtained from the leading diagonal block of (3.207) by replacing z with L, while $\boldsymbol{\delta}$ and $\boldsymbol{\Xi}(L)$ are given by

$$\boldsymbol{\delta} = \boldsymbol{N}_2\boldsymbol{\eta} \tag{3.210}$$

$$\boldsymbol{\Xi}(L) = \sum_{j=0}^{\infty}\boldsymbol{\Xi}_j L^j = \boldsymbol{M}(L)\nabla^2 + \boldsymbol{N}_1\nabla + \boldsymbol{N}_2 \tag{3.211}$$

Moreover, the operator relationship

$$\boldsymbol{M}(L)\nabla^2 = \boldsymbol{\Pi}_1(L) - \boldsymbol{N}_1\nabla - \boldsymbol{N}_2 \tag{3.212}$$

holds.

Furthermore, the following statements are true

(1) the matrix polynomial $\boldsymbol{\Xi}$ (L) has a second order zero at $z = 1$;
(2) $\nabla^2 \boldsymbol{y}_t$ is a non invertible VMA process;
(3) $E\,(\nabla^2 \boldsymbol{y}_t) = \boldsymbol{N}_2 \boldsymbol{\eta}$.

Proof

The proof is similar to that of Corollary 1.1 of Sect. 3.4.

Thanks to (1.548) of Sect. 1.10 and by virtue of the isomorphism between polynomials in a complex variable z and in the lag operator L, the following equality proves to be true

$$A^{-1}(L)\nabla^2 = [\boldsymbol{I} \;\; \boldsymbol{0}] \begin{bmatrix} \boldsymbol{\Psi}(L) & \boldsymbol{F} \\ \boldsymbol{G}' & \Lambda(L) \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} \tag{3.213}$$

This, in view of (3.194) and by virtue of (3.207), leads to the intended representation for $\nabla^2 \boldsymbol{y}_t$, namely

$$\begin{aligned} \nabla^2 \boldsymbol{y}_t &= [\boldsymbol{I} \;\; \boldsymbol{0}] \begin{bmatrix} \boldsymbol{\Psi}(L) & \boldsymbol{F} \\ \boldsymbol{G}' & \Lambda(L) \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0} \end{bmatrix} (\boldsymbol{\eta} + \boldsymbol{\varepsilon}_t) = \boldsymbol{\Pi}_1(L)\,(\boldsymbol{\eta} + \boldsymbol{\varepsilon}_t) \\ &= \boldsymbol{N}_2 \boldsymbol{\eta} + \boldsymbol{\Pi}_1(L)\,\boldsymbol{\varepsilon}_t \end{aligned} \tag{3.214}$$

upon noting that

$$\boldsymbol{\Pi}_1\,(1) = \boldsymbol{N}_2 \tag{3.215}$$

by bearing in mind (1.319) of Sect. 1.7, (1.548) of Sect. 1.10, and (3.207) above.

The VMA representation (3.209), as well as (3.210) and (3.211), follows from (3.178) by elementary computations.

The equality (3.212) proves true by comparing the right-hand sides of (3.208) and (3.209) in view of (3.211).

As far as statements (1)–(3) are concerned, their proofs rest on the reasoning developed here below.

Result (1) – The matrix polynomial

$$\boldsymbol{\Xi}\,(z) = (1-z)^2\,\boldsymbol{M}\,(z) + (1-z)\,\boldsymbol{N}_1 + \boldsymbol{N}_2 \tag{3.216}$$

has a second order zero at $z = 1$, according to Definition 5 of Sect. 1.6 and by virtue of Theorem 5, along with Corollary 5.1, of Sect. 1.10. Indeed the following hold

$$\Xi(1) = N_2 \Rightarrow det\, \Xi(1) = 0 \tag{3.217}$$

$$\dot{\Xi}(1) = -N_1, \quad \ddot{\Xi}(1) = 2M_1 \tag{3.218}$$

$$det\,(D'_\perp\, \dot{\Xi}(1)\, E_\perp) = 0 \Leftrightarrow det \begin{bmatrix} -\dot{\Xi}(1) & D \\ E' & 0 \end{bmatrix} = 0 \tag{3.219}$$

$$det\,(U'_\perp\, D'_\perp\, \tilde{\ddot{\Xi}}(1)\, E_\perp U_\perp) \neq 0 \tag{3.220}$$

where

(a) the matrices D and E are defined as per a rank factorization of $\Xi(1)$, namely

$$\Xi(1) = N_2 = DE', \; D = C_\perp S_\perp (R'_\perp B'_\perp \tilde{A}\, C_\perp S_\perp)^{-1}, \; E = B_\perp R_\perp \tag{3.221}$$

and thus the matrices $D_\perp$ and $E_\perp$ can be chosen as

$$D_\perp = (C_\perp S_\perp)_\perp, \qquad E_\perp = (B_\perp R_\perp)_\perp \tag{3.222}$$

(b) The matrix U is defined as per a rank factorization of $D'_\perp\, \dot{\Xi}(1)\, E_\perp$, namely

$$D'_\perp\, \dot{\Xi}(1)\, E_\perp = UU' \tag{3.223}$$

In light of (1.641) of Sect. 1.11, the matrix U coincides with the selection matrix U_1 of (1.592) of the same section, whence the following holds

$$D_\perp U_{1\perp} = C, \qquad E_\perp U_{1\perp} = B \tag{3.224}$$

by virtue of (1.593) of Sect. 1.11.

(c) The matrix $\tilde{\ddot{\Xi}}$ is defined as follows

$$\tilde{\ddot{\Xi}} = \frac{1}{2}\ddot{\Xi}(1) - \dot{\Xi}(1)\, \Xi^g(1)\, \dot{\Xi}(1) = M(1) - N_1 N_2^g N_1 \tag{3.225}$$

and consequently

$$D'_\perp\, U'_\perp\, \tilde{\ddot{\Xi}}(1)\, E_\perp U_\perp = C'(M(1) - N_1 N_2^g N_1)\, B = I \tag{3.226}$$

by virtue of (1.644) of Sect. 1.11.

Result (2) – The assertion simply follows from representation (40), taking into account (1) above.

□

Result (3) – The proof is straightforward, taking the expected value of both sides of (3.209) in light of (3.210).

□

Also for the representation theorem established in this section there exists a dual version. It originates from a specification of a stationary second difference process $\nabla^2 \boldsymbol{y}_t$ through a VMA model characterized so as to be the mirror image of a VAR model, whose solution enjoys particular integration and cointegration properties.

Theorem 2

Consider two stochastic processes $\boldsymbol{\xi}_t$ *and* $\boldsymbol{y}_t$*, the former being defined as the second difference of the latter, namely as*

$$\underset{(n,1)}{\boldsymbol{\xi}_t} = \nabla^2 \boldsymbol{y}_t \tag{3.227}$$

and let $\boldsymbol{\xi}_t$ *be stationary with the VMA(∞) representation*

$$\boldsymbol{\xi}_t = \boldsymbol{\Xi}\,(L)\,(\boldsymbol{\eta} + \boldsymbol{\varepsilon}_t) \tag{3.228}$$

Moreover, assume that the parent matrix polynomial

$$\boldsymbol{\Xi}\,(z) = \boldsymbol{\Xi}_0 + \sum_{i=1}^{\infty} \boldsymbol{\Xi}_i\, z^i \tag{3.229}$$

has a second-order zero at z = 1 and that the coefficient matrices $\boldsymbol{\Xi}_i$ *exhibit exponentially decreasing entries.*

Then, the companion process $\boldsymbol{y}_t$ *admits a VAR generating model*

$$\boldsymbol{A}\,(L)\,\boldsymbol{y}_t = \boldsymbol{\eta} + \boldsymbol{\varepsilon}_t \tag{3.230}$$

whose parent matrix polynomial $\boldsymbol{A}(z)$ *has a second-order zero at z = 1 and whose characteristic polynomial det*$\boldsymbol{A}(z)$ *has, besides a repeated unit-root, all other roots lying outside the unit circle.*

The engendered process $\boldsymbol{y}_t$ *enjoys the following major integration and cointegration properties*

$$\left.\begin{array}{r} \boldsymbol{y}_t \sim I(2) \\ (\boldsymbol{C}_\perp \boldsymbol{S}_\perp)'_\perp \boldsymbol{y}_t \sim I(1) \end{array}\right\} \Rightarrow \boldsymbol{y}_t \sim CI\,(2,1) \tag{3.231}$$

where $\boldsymbol{C}$ *and* $\boldsymbol{S}$ *are defined through rank factorizations of* $\boldsymbol{A}(1) = \boldsymbol{A}$ *and* $\boldsymbol{B}'_\perp\, \dot{\boldsymbol{A}}\, \boldsymbol{C}_\perp$ *respectively, namely*

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}' \tag{3.232}$$

$$\boldsymbol{B}'_{\perp}\, \dot{\boldsymbol{A}}\, \boldsymbol{C}_{\perp} = \boldsymbol{R}\boldsymbol{S}' \tag{3.233}$$

Proof

In light of (1.258) and (1.259) of Sect. 1.6 and of the isomorphism between polynomials in a complex variable z and in the lag operator L, we obtain the paired expansions

$$\begin{gathered}\boldsymbol{\Xi}(z) = \tilde{\boldsymbol{\Psi}}(z)(1-z)^2 - \dot{\boldsymbol{\Xi}}(1)(1-z) + \boldsymbol{\Xi}(1) \Leftrightarrow \\ \Leftrightarrow \boldsymbol{\Xi}(L) = \tilde{\boldsymbol{\Psi}}(L)\,\nabla^2 - \dot{\boldsymbol{\Xi}}(1)\,\nabla + \boldsymbol{\Xi}(1)\end{gathered} \tag{3.234}$$

where

$$\begin{gathered}\tilde{\boldsymbol{\Psi}}(z) = \sum_{k\geq 2}(-1)^k\,(1-z)^{k-2}\frac{1}{k!}\boldsymbol{\Xi}^{(k)}(1) \Leftrightarrow \\ \Leftrightarrow \tilde{\boldsymbol{\Psi}}(L) = \sum_{k\geq 2}(-1)^k\,\frac{1}{k!}\boldsymbol{\Xi}^{(k)}(1)\,\nabla^{k-2}\end{gathered} \tag{3.235}$$

with

$$\tilde{\boldsymbol{\Psi}}(1) = \frac{1}{2}\ddot{\boldsymbol{\Xi}}(1) \tag{3.236}$$

Then, in view of Definition 5 of Sect. 1.6 and by virtue of Theorem 5 and Corollary 5.1 of Sect. 1.10, the following hold true

$$det\,\boldsymbol{\Xi}(1) = 0 \tag{3.237}$$

$$det\left(\begin{bmatrix} -\dot{\boldsymbol{\Xi}}(1) & \tilde{\boldsymbol{B}} \\ \tilde{\boldsymbol{C}}' & \boldsymbol{0}\end{bmatrix}\right) = 0 \Leftrightarrow det\,(\tilde{\boldsymbol{B}}'_{\perp}\,\dot{\boldsymbol{\Xi}}(1)\,\tilde{\boldsymbol{C}}_{\perp}) = 0 \tag{3.238}$$

$$det\left(\begin{bmatrix} \tilde{\boldsymbol{\Xi}} & (\tilde{\boldsymbol{B}}_{\perp}\tilde{\boldsymbol{R}}_{\perp})_{\perp} \\ (\tilde{\boldsymbol{C}}_{\perp}\tilde{\boldsymbol{S}}_{\perp})'_{\perp} & \boldsymbol{0}\end{bmatrix}\right) \neq 0 \Leftrightarrow det\,(\tilde{\boldsymbol{R}}'_{\perp}\,\tilde{\boldsymbol{B}}'_{\perp}\,\tilde{\boldsymbol{\Xi}}\,\tilde{\boldsymbol{C}}_{\perp}\,\tilde{\boldsymbol{S}}_{\perp}) \neq 0 \tag{3.239}$$

where $\tilde{\boldsymbol{B}}$ and $\tilde{\boldsymbol{C}}$, on the one hand, and $\tilde{\boldsymbol{R}}$ and $\tilde{\boldsymbol{S}}$ on the other, are defined through rank factorizations of $\Xi(1)$ and $\tilde{\boldsymbol{B}}'_{\perp}\dot{\Xi}(1)\tilde{\boldsymbol{C}}_{\perp}$, respectively, that is to say

$$\Xi(1) = \tilde{\boldsymbol{B}}\tilde{\boldsymbol{C}}' \tag{3.240}$$

$$\tilde{\boldsymbol{B}}'_{\perp}\dot{\Xi}(1)\tilde{\boldsymbol{C}}_{\perp} = \tilde{\boldsymbol{R}}\tilde{\boldsymbol{S}}' \tag{3.241}$$

and where

$$\tilde{\Xi} = \frac{1}{2}\ddot{\Xi}(1) - \dot{\Xi}(1)\,\Xi^{g}(1)\,\dot{\Xi}(1) \tag{3.242}$$

Given these premises, expansion (1.318) of Sect. 1.7 holds for $\Xi^{-1}(z)$ in a deleted neighbourhood of the second order pole $z = 1$, and hence we can write

$$\begin{gathered}\Xi^{-1}(z) = \frac{1}{(1-z)^2}\tilde{\boldsymbol{N}}_2 + \frac{1}{(1-z)}\tilde{\boldsymbol{N}}_1 + \tilde{\boldsymbol{M}}(z) \Leftrightarrow \\ \Leftrightarrow \Xi^{-1}(L) = \tilde{\boldsymbol{N}}_2\nabla^{-2} + \tilde{\boldsymbol{N}}_1\nabla^{-1} + \tilde{\boldsymbol{M}}(L)\end{gathered} \tag{3.243}$$

where

$$\tilde{\boldsymbol{N}}_2 = \tilde{\boldsymbol{C}}_{\perp}\tilde{\boldsymbol{S}}_{\perp}\{\tilde{\boldsymbol{R}}'_{\perp}\tilde{\boldsymbol{B}}'_{\perp}\tilde{\Xi}\tilde{\boldsymbol{C}}_{\perp}\tilde{\boldsymbol{S}}_{\perp}\}^{-1}\tilde{\boldsymbol{R}}'_{\perp}\tilde{\boldsymbol{B}}'_{\perp} \tag{3.244}$$

$$\begin{gathered}\tilde{\boldsymbol{N}}_1 = [\tilde{\boldsymbol{N}}_2\ , \boldsymbol{I} - \tilde{\boldsymbol{N}}_2\tilde{\Xi}]\cdot \\ \cdot\begin{bmatrix}\tilde{\tilde{\Xi}} & \dot{\Xi}(1)\,\Xi^{g}(1) \\ \Xi^{g}(1)\,\dot{\Xi}(1) & -\tilde{\boldsymbol{C}}_{\perp}(\tilde{\boldsymbol{B}}'_{\perp}\dot{\Xi}(1)\tilde{\boldsymbol{C}}_{\perp})^{g}\tilde{\boldsymbol{B}}'_{\perp}\end{bmatrix}\begin{bmatrix}\tilde{\boldsymbol{N}}_2 \\ \boldsymbol{I} - \tilde{\Xi}\tilde{\boldsymbol{N}}_2\end{bmatrix}\end{gathered} \tag{3.245}$$

$$\tilde{\tilde{\Xi}} = \frac{1}{6}\dddot{\Xi}(1) - [\dot{\Xi}(1)\,\Xi^{g}(1)]^2\,\dot{\Xi}(1) \tag{3.246}$$

in view of Theorem 2 of Sect. 1.11.

Applying the operator $\Xi^{-1}(L)$ to both sides of (3.228) yields

$$\Xi^{-1}(L)\,\boldsymbol{\xi}_t = \boldsymbol{\eta} + \boldsymbol{\varepsilon}_t \Rightarrow \boldsymbol{A}(L)\,\boldsymbol{y}_t = \boldsymbol{\eta} + \boldsymbol{\varepsilon}_t \tag{3.247}$$

that is the VAR representation (3.230), whose parent matrix polynomial

$$\boldsymbol{A}(z) = \Xi^{-1}(z)\,(1-z)^2 = (1-z)^2\,\tilde{\boldsymbol{M}}(z) + (1-z)\,\tilde{\boldsymbol{N}}_1 + \tilde{\boldsymbol{N}}_2 \tag{3.248}$$

is characterized by a second-order zero at $z = 1$.

Indeed, the following results hold true:

(1) $$\boldsymbol{A} = \tilde{\boldsymbol{N}}_2 \Rightarrow det\, \boldsymbol{A} = 0 \tag{3.249}$$

whence the rank factorization (3.232), with $\boldsymbol{B}$ and $\boldsymbol{C}$ which are conveniently chosen as

$$\boldsymbol{B} = \tilde{\boldsymbol{C}}_\perp \tilde{\boldsymbol{S}}_\perp \{\tilde{\boldsymbol{R}}'_\perp \tilde{\boldsymbol{B}}'_\perp \tilde{\boldsymbol{\Xi}} \tilde{\boldsymbol{C}}_\perp \tilde{\boldsymbol{S}}_\perp\}^{-1},\ \boldsymbol{C} = \tilde{\boldsymbol{B}}_\perp \tilde{\boldsymbol{R}}_\perp \tag{3.250}$$

(2) A possible choice for matrices $\boldsymbol{B}_\perp$ and $\boldsymbol{C}_\perp$ will therefore be

$$\begin{aligned} \boldsymbol{B}_\perp &= (\tilde{\boldsymbol{C}}_\perp \tilde{\boldsymbol{S}}_\perp)_\perp = [(\tilde{\boldsymbol{C}}'_\perp)^g \tilde{\boldsymbol{S}},\ \ \tilde{\boldsymbol{C}}], \\ \boldsymbol{C}_\perp &= (\tilde{\boldsymbol{B}}_\perp \tilde{\boldsymbol{R}}_\perp)_\perp = [(\tilde{\boldsymbol{B}}'_\perp)^g \tilde{\boldsymbol{R}},\ \ \tilde{\boldsymbol{B}}] \end{aligned} \tag{3.251}$$

(3) $$\dot{\boldsymbol{A}} = -\,\tilde{\boldsymbol{N}}_1 \tag{3.252}$$

whence, by virtue of (1.641) of Sect. 1.11, the following proves to be true

$$\boldsymbol{B}'_\perp \dot{\boldsymbol{A}} \boldsymbol{C}_\perp = -\,(\tilde{\boldsymbol{C}}_\perp \tilde{\boldsymbol{S}}_\perp)'_\perp \tilde{\boldsymbol{N}}_1 (\tilde{\boldsymbol{B}}_\perp \tilde{\boldsymbol{R}}_\perp)_\perp = \boldsymbol{U}_1 \boldsymbol{U}'_1 \tag{3.253}$$

where $\boldsymbol{U}_1$ is as defined in (1.592) of Sect. 1.11.

(4) In light of (3.253) the matrix $\boldsymbol{B}'_\perp \dot{\boldsymbol{A}} \boldsymbol{C}_\perp$ is singular, whence the rank factorization (3.233) occurs with $\boldsymbol{R} = \boldsymbol{S} = \boldsymbol{U}_1$ and we can choose $\boldsymbol{R}_\perp = \boldsymbol{S}_\perp = \boldsymbol{U}_2$ accordingly, where $\boldsymbol{U}_2$ is as defined in (1.592) of Sect. 1.11.

(5) $$\ddot{\boldsymbol{A}} = 2\,\tilde{\boldsymbol{M}}(1) \tag{3.254}$$

$$\tilde{\boldsymbol{A}} = \frac{1}{2}\ddot{\boldsymbol{A}} - \dot{\boldsymbol{A}} \boldsymbol{A}^g \dot{\boldsymbol{A}} = \tilde{\boldsymbol{M}}(1) - \tilde{\boldsymbol{N}}_1 \tilde{\boldsymbol{N}}_2^g \tilde{\boldsymbol{N}}_1 \tag{3.255}$$

because of (3.248), (3.249) and (3.252).

(6) The matrix $\boldsymbol{R}'_\perp \boldsymbol{B}'_\perp \tilde{\boldsymbol{A}} \boldsymbol{C}_\perp \boldsymbol{S}_\perp$ is non-singular. In fact the following holds

$$\begin{aligned} &\boldsymbol{R}'_\perp \boldsymbol{B}'_\perp \tilde{\boldsymbol{A}} \boldsymbol{C}_\perp \boldsymbol{S}_\perp = \\ &= \boldsymbol{U}'_2 (\tilde{\boldsymbol{C}}_\perp \tilde{\boldsymbol{S}}_\perp)_\perp (\tilde{\boldsymbol{M}}(1) - \tilde{\boldsymbol{N}}_1 \tilde{\boldsymbol{N}}_2^g \tilde{\boldsymbol{N}}_1)(\tilde{\boldsymbol{B}}_\perp \tilde{\boldsymbol{R}}_\perp)_\perp \boldsymbol{U}_2 \\ &= \tilde{\boldsymbol{C}}' (\tilde{\boldsymbol{M}}(1) - \tilde{\boldsymbol{N}}_1 \tilde{\boldsymbol{N}}_2^g \tilde{\boldsymbol{N}}_1) \tilde{\boldsymbol{B}} = \boldsymbol{I} \end{aligned} \tag{3.256}$$

in view of (1.593) and (1.644) of Sect. 1.11.

The proof of what has been claimed about the roots of the characteristic polynomial $det\, \boldsymbol{A}(z)$ relies upon the reasoning of Theorem 4 of Sect. 1.7.

Now, note that, insofar as $\boldsymbol{A}(z)$ has a second order zero at $z = 1$, $\boldsymbol{A}^{-1}(z)$ has a second order pole at $z = 1$. According to (3.248) the following Laurent expansions holds

$$\begin{aligned} \boldsymbol{A}^{-1}(L) &= \sum_{j=1}^{2} \frac{1}{(I-L)^j} \boldsymbol{N}_j + \boldsymbol{M}(L) = \boldsymbol{\Xi}(L)\nabla^{-2} \\ &= \tilde{\boldsymbol{\Psi}}(L) - \dot{\boldsymbol{\Xi}}(1)\nabla^{-1} + \boldsymbol{\Xi}(1)\nabla^{-2} \end{aligned} \tag{3.257}$$

Then, from (3.230) and (3.257) it follows that

$$\begin{aligned} \boldsymbol{y}_t &= \boldsymbol{A}(L)^{-1}(\boldsymbol{\varepsilon}_t + \boldsymbol{\eta}) = (\tilde{\boldsymbol{\Psi}}(L) - \dot{\boldsymbol{\Xi}}(1)\nabla^{-1} + \boldsymbol{\Xi}(1)\nabla^{-2})(\boldsymbol{\varepsilon}_t + \boldsymbol{\eta}) \\ &= \tilde{\boldsymbol{k}}_0 + \tilde{\boldsymbol{k}}_1 t + \tilde{\boldsymbol{B}}\tilde{\boldsymbol{C}}'\tilde{\boldsymbol{k}}_2 t^2 - \dot{\boldsymbol{\Xi}}(1)\sum_{\tau \le t} \boldsymbol{\varepsilon}_\tau \\ &\quad + \tilde{\boldsymbol{B}}\tilde{\boldsymbol{C}}' \sum_{\tau \le t}(t+1-\tau)\boldsymbol{\varepsilon}_\tau + \sum_{i=0}^{\infty} \tilde{\boldsymbol{\Psi}}_i \boldsymbol{\varepsilon}_{t-i} \end{aligned} \tag{3.258}$$

where $\tilde{\boldsymbol{k}}_0$, $\tilde{\boldsymbol{k}}_1$ and $\tilde{\boldsymbol{k}}_2$ are vectors of constants.

Looking at (3.258) it is easy to realize that pre-multiplication by $\tilde{\boldsymbol{B}}'_\perp$ leads to the elimination of both quadratic deterministic trends and cumulated random walks, which is equivalent to claim that

$$\tilde{\boldsymbol{B}}'_\perp \boldsymbol{y}_t \sim I(1) \Rightarrow \boldsymbol{y}_t \sim CI(2, 1) \tag{3.259}$$

with $\tilde{\boldsymbol{B}}_\perp = (\boldsymbol{C}_\perp \boldsymbol{U}_2)_\perp = (\boldsymbol{C}_\perp \boldsymbol{S}_\perp)_\perp$ playing the role of the matrix of the cointegrating vectors.

□

3.6 A Unified Representation Theorem

This last section is devoted to a unified representation theorem which – by resorting to the somewhat sophisticated analytic tool-kit set forth in the first chapter – shapes the contours of a general approach to VAR models with unit roots.

Otherwise stated, this theorem sets a framework, consistent with the operating path already traced in the preceding sections, which will eventually lead to shed effective light on VAR-modelling solutions and offspring, and will as well create a natural bridge between VAR specifications once tailored on $I(1)$ rather than on $I(2)$ processes.

Theorem

Consider a VAR model specified as

$$\underset{(n,1)}{\boldsymbol{A}(L)\ \boldsymbol{y}_t} = \boldsymbol{\eta} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \text{WN}(n) \tag{3.260}$$

where $\boldsymbol{\eta}$ *is a vector of constant terms and*

$$\boldsymbol{A}(L) = \sum_{k=0}^{K} \boldsymbol{A}_k L^k, \quad \boldsymbol{A}_0 = \boldsymbol{I}_n, \quad \boldsymbol{A}_K \neq \boldsymbol{0}_n \tag{3.261}$$

Let the characteristic polynomial det$\boldsymbol{A}(z)$ have all roots lying outside the unit circle, except for one or more unit roots, and let the specification be accomplished through an assumption about the index, say υ, of the companion matrix

$$\underset{(\breve{n},\breve{n})}{\breve{\boldsymbol{A}}} = \begin{bmatrix} \boldsymbol{I}+\boldsymbol{A}_1 & \vdots & \boldsymbol{A}_2 & \boldsymbol{A}_3 & .. & \boldsymbol{A}_K \\ \cdots & & \cdots & \cdots & \cdots & \cdots \\ -\boldsymbol{I}_n & \vdots & \boldsymbol{I}_n & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \vdots & -\boldsymbol{I}_n & \boldsymbol{I}_n & \boldsymbol{0} & \boldsymbol{0} \\ .. & \vdots & .. & .. & .. & .. \\ \boldsymbol{0} & \vdots & \boldsymbol{0} & \boldsymbol{0} & -\boldsymbol{I}_n & \boldsymbol{I}_n \end{bmatrix} \tag{3.262}$$

Then, the following representation holds for the solution of (3.260) above

$$\boldsymbol{y}_t = \sum_{j=1}^{\upsilon} \boldsymbol{N}_j \boldsymbol{\rho}_t(j) + \sum_{j=0}^{\upsilon} \boldsymbol{k}_j t^j + \sum_{i=0}^{\infty} \boldsymbol{M}_i \boldsymbol{\varepsilon}_{t-i}, \quad \boldsymbol{N}_\upsilon \neq \boldsymbol{0}, \quad \boldsymbol{k}_\upsilon \neq \boldsymbol{0} \tag{3.263}$$

where the matrices $\boldsymbol{N}_j$ *are algebraic functions of the Drazin inverse of* $\breve{\boldsymbol{A}}^{\boldsymbol{\upsilon}}$*, as well as of the nilpotent component of* $\breve{\boldsymbol{A}}$ *and of the coefficients of the expansion of* $\boldsymbol{M}(z) = \sum_{i=0}^{\infty} \boldsymbol{M} z^i$ *about z =1, the vectors* $\boldsymbol{k}_j$ *are linear functions of* $\boldsymbol{N}_{j+l}\ (l \geq 0)$ *depending on arbitrary constants, the matrices* $\boldsymbol{M}_i$ *are characterized by exponentially decreasing entries, and the vectors* $\boldsymbol{\rho}_t(j)$ *are random walks of consecutive order defined recursively by the relations*

$$\boldsymbol{\rho}_t(1) = \sum_{\tau \leq t} \boldsymbol{\varepsilon}_\tau \tag{3.264}$$

$$\boldsymbol{\rho}_t(j)=\sum_{\tau\le t}\boldsymbol{\rho}_{\tau}(j-1)\ \ ,for\, j>1 \tag{3.265}$$

The solution (3.263) is a an integrated and cointegrated process, such that

(1)
$$\boldsymbol{y}_t\sim I(\upsilon) \tag{3.266}$$

(2)
$$\begin{cases} \boldsymbol{D}_{\upsilon}'\boldsymbol{y}_t\sim I(0)\rightarrow & \boldsymbol{y}_t\sim CI(\upsilon,\upsilon)\\ \boldsymbol{D}_{\upsilon-1}'\boldsymbol{y}_t\sim I(1)\rightarrow & \boldsymbol{y}_t\sim CI(\upsilon,\upsilon-1)\\ \vdots & \\ \boldsymbol{D}_1'\boldsymbol{y}_t\sim I(\upsilon-1)\rightarrow & \boldsymbol{y}_t\sim CI(\upsilon,1)\end{cases} \tag{3.267}$$

where $\boldsymbol{D}_{\upsilon},\ \boldsymbol{D}_{\upsilon-1},\ \ldots,\boldsymbol{D}_1$ *are full column-rank matrices such that the columns of* $\boldsymbol{D}_{\upsilon}$ *span the null row-space of* $[\boldsymbol{N}_{\upsilon},\ \ \boldsymbol{N}_{\upsilon-1},\ \ ...,\ \ \boldsymbol{N}_1]$, *the columns of* $[\boldsymbol{D}_{\upsilon},\boldsymbol{D}_{\upsilon-1}]$ *span the null row-space of* $[\boldsymbol{N}_{\upsilon},\ \ \boldsymbol{N}_{\upsilon-1},\ \ ...,\ \ \boldsymbol{N}_2]$,…, *and the columns of* $[\boldsymbol{D}_{\upsilon},\ \boldsymbol{D}_{\upsilon-1},\ \ldots,\boldsymbol{D}_1]$ *span the null row-space of* $\boldsymbol{N}_{\upsilon}$. *The matrix* $\boldsymbol{D}_{\upsilon}$ *as well as* $\boldsymbol{D}_{\upsilon-1}$, ..., $\boldsymbol{D}_2$ *may happen to be empty matrices. The overall cointegration rank is equal to the nullity of* $\boldsymbol{N}_{\upsilon}$.

In particular, the following statements hold true

(a) *if*

$$det\,\boldsymbol{A}=0 \tag{3.268}$$

$$det\,\text{Л}\neq 0 \tag{3.269}$$

then,

$$\upsilon=1\rightarrow\boldsymbol{y}_t\sim I(1) \tag{3.270}$$

$$\boldsymbol{N}_1=-\boldsymbol{C}_{\perp}\,\text{Л}^{-1}\boldsymbol{B}_{\perp}' \tag{3.271}$$

$$\boldsymbol{k}_1=\boldsymbol{N}\boldsymbol{\eta} \tag{3.272}$$

$$\boldsymbol{k}_0=\boldsymbol{N}_1\boldsymbol{c}+\boldsymbol{M}\boldsymbol{\eta} \tag{3.273}$$

$$\boldsymbol{D}_1=\boldsymbol{C} \tag{3.274}$$

$$\boldsymbol{C}'\,\boldsymbol{y}_t\sim I(0)\quad\rightarrow\quad\boldsymbol{y}_t\sim CI(1,1),\quad cir=n-r(\boldsymbol{A}) \tag{3.275}$$

where

$$\text{Л}=\boldsymbol{B}_{\perp}'\ \dot{\boldsymbol{A}}\ \boldsymbol{C}_{\perp} \tag{3.276}$$

Here $\boldsymbol{A}$, $\dot{\boldsymbol{A}}$ *and* $\boldsymbol{M}$ *are short forms for* $\sum_{k=0}^{K}\boldsymbol{A}_k$, $\sum_{k=1}^{K}k\boldsymbol{A}_k$ and $\sum_{i=0}^{\infty}\boldsymbol{M}_i$, $\boldsymbol{BC}'$ *is a rank factorization of* $\boldsymbol{A}$, $\boldsymbol{c}$ *is arbitrary and* cir *denotes the cointegration rank.*

(b) *If*

$$det\, \boldsymbol{A} = 0 \tag{3.277}$$

$$det\, \text{Л} = 0 \tag{3.278}$$

$$det(\boldsymbol{R}'_\perp\, \boldsymbol{B}'_\perp\, \tilde{\boldsymbol{A}}\, \boldsymbol{C}_\perp \boldsymbol{S}_\perp) \neq 0 \tag{3.279}$$

then

$$\upsilon = 2 \rightarrow \boldsymbol{y}_t \sim I(2) \tag{3.280}$$

$$\boldsymbol{N}_2 = \boldsymbol{C}_\perp \boldsymbol{S}_\perp (\boldsymbol{R}'_\perp \boldsymbol{B}'_\perp \tilde{\boldsymbol{A}} \boldsymbol{C}_\perp \boldsymbol{S}_\perp)^{-1} \boldsymbol{R}'_\perp \boldsymbol{B}'_\perp \tag{3.281}$$

$$\boldsymbol{N}_1 = -\boldsymbol{C}_\perp \text{Л}^g \boldsymbol{B}'_\perp + \boldsymbol{N}_2 \boldsymbol{X} \boldsymbol{N}_2 + \boldsymbol{Y} \boldsymbol{N}_2 + \boldsymbol{N}_2 \boldsymbol{Z} \tag{3.282}$$

$$\boldsymbol{k}_2 = \frac{1}{2} \boldsymbol{N}_2 \boldsymbol{\eta} \tag{3.283}$$

$$\boldsymbol{k}_1 = \boldsymbol{N}_1 \boldsymbol{\eta} + \boldsymbol{N}_2 (\boldsymbol{c} + \frac{1}{2} \boldsymbol{\eta}) \tag{3.284}$$

$$\boldsymbol{k}_0 = \boldsymbol{N}_1 \boldsymbol{c} + \boldsymbol{N}_2 \boldsymbol{d} + \boldsymbol{M} \boldsymbol{\eta} \tag{3.285}$$

$$\boldsymbol{D}_2 = \boldsymbol{C}(\boldsymbol{B}^g \boldsymbol{V})_\perp \tag{3.286}$$

$$\boldsymbol{D}_1 = \left[\boldsymbol{C} \boldsymbol{B}^g \boldsymbol{V}, (\boldsymbol{C}'_\perp)^g \boldsymbol{S} \right] \tag{3.287}$$

$$(\boldsymbol{B}^g \boldsymbol{V})'_\perp \boldsymbol{C}' \boldsymbol{y}_t \sim I(0) \quad \rightarrow \quad \boldsymbol{y}_t \sim CI(2,2), \quad cir = r(\boldsymbol{A}) - r(\boldsymbol{V}) \tag{3.288}$$

$$\left[\boldsymbol{C} \boldsymbol{B}^g \boldsymbol{V}, (\boldsymbol{C}'_\perp)^g \boldsymbol{S} \right]' \boldsymbol{y}_t \sim I(1) \;\rightarrow\; \boldsymbol{y}_t \sim CI(2,1),\; cir = r(\text{Л}) + r(\boldsymbol{V}) \tag{3.289}$$

where

$$\tilde{\boldsymbol{A}} = (\frac{1}{2} \ddot{\boldsymbol{A}} - \dot{\boldsymbol{A}} \boldsymbol{A}^g \dot{\boldsymbol{A}}) \tag{3.290}$$

$$\tilde{\tilde{\boldsymbol{A}}} = \frac{1}{6} \dddot{\boldsymbol{A}} - \dot{\boldsymbol{A}} \boldsymbol{A}^g \dot{\boldsymbol{A}} \boldsymbol{A}^g \dot{\boldsymbol{A}} \tag{3.291}$$

$$\boldsymbol{X} = \tilde{\tilde{\boldsymbol{A}}} - \tilde{\boldsymbol{A}} \boldsymbol{Y} - \boldsymbol{Z} \tilde{\boldsymbol{A}} + \tilde{\boldsymbol{A}} \boldsymbol{C}_\perp \text{Л}^g \boldsymbol{B}'_\perp \tilde{\boldsymbol{A}} \tag{3.292}$$

$$\boldsymbol{Y} = \boldsymbol{A}^g \dot{\boldsymbol{A}} + \boldsymbol{C}_\perp \text{Л}^g \boldsymbol{B}'_\perp \tilde{\boldsymbol{A}} \tag{3.293}$$

$$\boldsymbol{Z} = \dot{\boldsymbol{A}} \boldsymbol{A}^g + \tilde{\boldsymbol{A}} \boldsymbol{C}_\perp \text{Л}^g \boldsymbol{B}'_\perp \tag{3.294}$$

Here $\ddot{\boldsymbol{A}}$ *and* $\dddot{\boldsymbol{A}}$ *are short forms for* $\sum_{k=2}^{K} k(k-1)\boldsymbol{A}_k$ *and* $\sum_{k=3}^{K} k(k-1)(k-2)\boldsymbol{A}_k$, *and* $\boldsymbol{R}\boldsymbol{S}'$ *and* $\boldsymbol{V}\boldsymbol{\Lambda}'$ *are rank factorizations of* $\boldsymbol{B}'_\perp\, \dot{\boldsymbol{A}}\, \boldsymbol{C}_\perp$ and $\dot{\boldsymbol{A}} \boldsymbol{C}_\perp \boldsymbol{S}_\perp$, *while* $\boldsymbol{d}$ *is arbitrary.*

Proof

Since the VAR model (3.260) is a linear difference equation, its general solution can be represented (see Theorem 1, Sect. 1.8) in this manner

$$y_t = \begin{Bmatrix} complementary \\ solution \end{Bmatrix} + \begin{Bmatrix} particular\ solution\ of\ the \\ non\text{-}homogeneous\ equation \end{Bmatrix} \tag{3.295}$$

As far as the complementary solution, i.e. the (general) solution of the reduced equation

$$\boldsymbol{A}\,(L)\,\boldsymbol{y}_t = \boldsymbol{0} \tag{3.296}$$

is concerned, we have to consider both a transitory component $\tilde{\boldsymbol{\xi}}_t$, associated with the roots of the characteristic polynomial $det\boldsymbol{A}(z)$ lying outside the unit circle, and a permanent component $\bar{\boldsymbol{\xi}}_t$, due to unit roots.

On the one hand, the transitory component of the complementary solution can be expressed as

$$\tilde{\boldsymbol{\xi}}_t = \boldsymbol{H}\boldsymbol{\lambda}^{(t)} \tag{3.297}$$

where the symbols have the same meaning as in (3.14)–(3.17) of Sect. 3.1 and reference is made to the non-unit roots of the characteristic polynomial. Since the elements of $\boldsymbol{\lambda}^{(t)}$ decrease at an exponential rate, the contribution of component $\tilde{\boldsymbol{\xi}}_t$ becomes immaterial when t increases and as such it is ignored on the right-hand side of (3.263).

The permanent component $\bar{\boldsymbol{\xi}}_t$, on the other hand, can be expressed – by resorting to Theorem 3 of Sect. 1.8 and extending result (3.296) therein to higher-order poles – as a polynomial in t of degree υ-1 with coefficients $\boldsymbol{a}_j$ depending linearly on the matrices $\boldsymbol{N}_{j+l}$ $(l \geq 0)$, as well as on arbitrary constants.

When searching for a particular solution of the complete (non-homogeneous) equation (3.260), according to Theorem 2 of Sect. 1.8 and to Theorem 1 of Sect. 2.3, we have

$$\bar{\boldsymbol{y}}_t = \boldsymbol{A}^{-1}(\boldsymbol{L})\,(\boldsymbol{\eta} + \boldsymbol{\varepsilon}_t) \tag{3.298}$$

Bearing in mind the isomorphism of algebras of matrix polynomials in a complex variable and in the lag operator, reference to Theorem 7 of Sect. 1.6 together with Remark 1 of Sect. 1.8 and to Theorem 1 of Sect. 2.3, leads to the formal Laurent expansion

$$\boldsymbol{A}^{-1}(L)=\boldsymbol{N}_{\upsilon}\nabla^{-\upsilon}+....+\boldsymbol{N}_{1}\nabla^{-1}+\sum_{i=0}^{\infty}\boldsymbol{M}_{i}L^{i} \tag{3.299}$$

Here, the matrices $\boldsymbol{M}_i$ have exponentially decreasing entries by Theorem 4 of Sect. 1.7. The matrices $\boldsymbol{N}_j$ can be obtained by resorting to Theorem 1 of Sect. 1.9 to derive the expressions of the parent matrices of the companion form, and to Theorem 2 bis of Sect. 1.8 to recover the intended matrices from the latter ones by post and pre-multiplication by a selection matrix and its transpose. In particular, $\boldsymbol{N}_{\upsilon}$ and $\boldsymbol{N}_{\upsilon-1}$ turn out to be

$$\boldsymbol{N}_{\upsilon}=(-1)^{\upsilon-1}\boldsymbol{J}'\boldsymbol{N}\boldsymbol{H}^{\upsilon-1}\boldsymbol{J} \tag{3.300}$$

$$\boldsymbol{N}_{\upsilon-1}=(-1)^{\upsilon-2}\boldsymbol{J}'\boldsymbol{N}\boldsymbol{H}^{\upsilon-2}\boldsymbol{J}+(1-\upsilon)\boldsymbol{N}_{\upsilon} \tag{3.301}$$

where

$$\underset{(\breve{n},\breve{n})}{\boldsymbol{N}}=\boldsymbol{I}-(\breve{\boldsymbol{A}}^{\upsilon})^{\neq}\breve{\boldsymbol{A}}^{\upsilon} \tag{3.302}$$

$$\underset{(\breve{n},n)}{\boldsymbol{J}}=\begin{bmatrix}\boldsymbol{I}_n\\ \boldsymbol{0}_n\\ \vdots\\ \boldsymbol{0}_n\end{bmatrix} \tag{3.303}$$

and $\boldsymbol{H}$ is the nilpotent component of $\breve{\boldsymbol{A}}$ as specified by Theorem 1 of Sect. 1. 5

Substituting (3.299) into (3.298) yields

$$\begin{aligned}\bar{\boldsymbol{y}}_t&=(\boldsymbol{N}_{\upsilon}\nabla^{-\upsilon}+...+\boldsymbol{N}_{1}\nabla^{-1}+\sum_{i=0}^{\infty}\boldsymbol{M}_{i}L^{i})(\boldsymbol{\eta}+\boldsymbol{\varepsilon}_t)\\&=(\boldsymbol{N}_{\upsilon}\nabla^{-\upsilon}+...+\boldsymbol{N}_{1}\nabla^{-1}+\sum_{i=0}^{\infty}\boldsymbol{M}_{i}L^{i})\boldsymbol{\eta}+\\&\quad+\boldsymbol{N}_{\upsilon}\nabla^{-\upsilon}\boldsymbol{\varepsilon}_t+...+\boldsymbol{N}_{1}\nabla^{-1}\boldsymbol{\varepsilon}_t+\sum_{i=0}^{\infty}\boldsymbol{M}_{i}\boldsymbol{\varepsilon}_{t-i}\end{aligned}$$

$$=(\boldsymbol{N}_{\upsilon}\nabla^{-\upsilon}+...+\boldsymbol{N}_{1}\nabla^{-1}+\boldsymbol{M})\boldsymbol{\eta}+\boldsymbol{N}_{\upsilon}\boldsymbol{\rho}_t(\upsilon)+....+\boldsymbol{N}_{1}\boldsymbol{\rho}_t(1)+\sum_{i=0}^{\infty}\boldsymbol{M}_{i}\boldsymbol{\varepsilon}_{t-i} \tag{3.304}$$

where $\boldsymbol{M}$ is a short form for $\sum_{i=0}^{\infty}\boldsymbol{M}_i$, $\sum_{i=0}^{\infty}\boldsymbol{M}_i\boldsymbol{\varepsilon}_{t-i}$ is a VMA process and as such it is stationary, and the vectors $\boldsymbol{\rho}_t(j)$ are random walks of progressive order (see (3.264) and (3.265) above) which originate the stochastic trends of the solution.

From sum-calculus rules quoted in (1.384) of Sect. 1.8 – duly extended to higher-order summations – it follows that the expression $\nabla^{-j}\boldsymbol{\eta}$ is a *j-th* degree polynomial of *t*. This in turn implies that the term

$$\boldsymbol{\zeta}_t = (\boldsymbol{N}_\upsilon \nabla^{-\upsilon} + ... + \boldsymbol{N}_1 \nabla^{-1} + \boldsymbol{M})\boldsymbol{\eta} \tag{3.305}$$

– likewise the permanent component $\overline{\boldsymbol{\xi}}_t$ – is itself a υ-*th* degree polynomial of *t*, with coefficient $\boldsymbol{b}_j$ depending linearly upon the matrices $\boldsymbol{N}_{j+l}$ ($l \geq 0$), and $\boldsymbol{M}$.

Altogether the term (3.305) and the permanent component $\overline{\boldsymbol{\xi}}_t$ of the complementary solution form the polynomial

$$\boldsymbol{\kappa}_t = \overline{\boldsymbol{\xi}}_t + \boldsymbol{\zeta}_t = \sum_{j=0}^{\upsilon} \boldsymbol{k}_j t^j \tag{3.306}$$

which originates the deterministic trends of the solution.

Combining the permanent component $\overline{\boldsymbol{\xi}}_t$ of the complementary solution with the particular solution (3.304) above, rearranging terms and collecting identical powers, we eventually arrive at the representation (3.263) for the stochastic process generated by the VAR model (3.260).

The integration property of $\boldsymbol{y}_t$, as claimed in (1), is inherent in the representation (3.263), in accordance with Theorem 1 of Sect. 2.3, because of the presence in the solution $\boldsymbol{y}_t$ of both stochastic and deterministic trends up to the υ-*th* order.

As far as the cointegration properties as claimed in (2) are concerned, they are justified on the basis of the following considerations.

Following an argument similar to that used in the proofs of Theorems 1 and 2 of Sect. 1.7, it is easy to show that $\boldsymbol{N}_\upsilon$ is a singular matrix, with the rank factorization

$$\boldsymbol{N}_\upsilon = \boldsymbol{Б}\boldsymbol{Ч}' \tag{3.307}$$

as a by-product result.

Conversely, no general assertion can be made on the rank of the composite matrices $[\boldsymbol{N}_\upsilon, \boldsymbol{N}_{\upsilon-1}], \ldots, [\boldsymbol{N}_\upsilon, \boldsymbol{N}_{\upsilon-1}, \ldots, \boldsymbol{N}_1]$.

Since $\boldsymbol{N}_{\upsilon}$ is not of full rank there will exits a non-empty matrix − to be identified with $\boldsymbol{Б}_{\perp}$ − whose columns will form a basis for the null row-space of $\boldsymbol{N}_{\upsilon}$.

In a similar fashion as in (1.626) of Sect. 1.11, it is possible to make a partitioning of the matrix $\boldsymbol{Б}_{\perp}$ of the form

$$\boldsymbol{Б}_{\perp} = [\,\boldsymbol{D}_{\upsilon},\ \boldsymbol{D}_{\upsilon-1},\ \ldots, \boldsymbol{D}_{1}\,] \tag{3.308}$$

where $\boldsymbol{D}_{\upsilon}$ is either an empty matrix when $[\,\boldsymbol{N}_{\upsilon}, \boldsymbol{N}_{\upsilon-1}, \ldots, \boldsymbol{N}_{1}\,]$ is of full row-rank, or a full column-rank matrix whose columns span the null row-space of the latter matrix otherwise; $[\,\boldsymbol{D}_{\upsilon},\ \boldsymbol{D}_{\upsilon-1}\,]$ is either an empty matrix when $[\,\boldsymbol{N}_{\upsilon}, \boldsymbol{N}_{\upsilon-1}, \ldots, \boldsymbol{N}_{2}\,]$ is of full row-rank, or a full column-rank matrix whose columns span the null row-space of the latter matrix otherwise; and so on.

More precisely, as long as $\boldsymbol{D}_{\upsilon}$ is not empty $\boldsymbol{D}_{\upsilon}'$ is orthogonal to $\boldsymbol{N}_{\upsilon}, \boldsymbol{N}_{\upsilon-1}, \ldots, \boldsymbol{N}_{1}$ as well as to $\boldsymbol{k}_{\upsilon}, \boldsymbol{k}_{\upsilon-1}, \ldots, \boldsymbol{k}_{1}$ (see (3.306) above); as long as $[\,\boldsymbol{D}_{\upsilon},\ \boldsymbol{D}_{\upsilon-1}\,]$ is not empty, $\boldsymbol{D}_{\upsilon-1}'$ is orthogonal to $\boldsymbol{N}_{\upsilon}, \boldsymbol{N}_{\upsilon-1}, \ldots, \boldsymbol{N}_{2}$, as well as to $\boldsymbol{k}_{\upsilon},\ \boldsymbol{k}_{\upsilon-1}, \ldots, \boldsymbol{k}_{2}$; and so on.

Hence, pre-multiplying both sides of (3.263) by the matrices $\boldsymbol{D}_{\upsilon}'$, $\boldsymbol{D}_{\upsilon-1}'$ $\ldots, \boldsymbol{D}_{1}'$, one at a time, eventually provides the cointegration relationships inherent in model (3.260) as claimed in (3.267).

Let us now consider point (a). Condition (3.268) and (3.269) entail (3.270) and the other way around, according to Proposition 3 of Sect. 3.2, which in turn rests on Theorem 3 and Corollary 3.1 of Sect. 1.10.

Formula (3.271) is the simplest form of (3.302) and can be obtained by the latter as shown in Theorem 5 of Sect. 1.12.

The proof of (3.272) and (3.273) rests on the reasoning underlying (3.306) above and takes advantage of (1.384) and (1.392) of Sect. 1.8.

Results (3.274) and (3.275), which mirror the upper statement of (3.267), rest on Corollary 5.1 of Sect. 1.10, upon resorting to the isomorphism of the algebras of polynomials in a complex variable and in the lag operator (see also, in this connection, Proposition 4 of Sect. 3.2 and Theorem 1 of Sect. 3.4).

Finally, we will deal with point (b). Conditions (3.277), (3.278), and (3.279) entail (3.280) and the other way around. According to Proposition 5 of Sect. 3.2, which in turn rests on Theorem 5 and Corollary 5.1 of Sect. 1.10.

Formulas (3.281) and (3.282) are special cases of (3.300) and (3.301), and can be obtained from the latter two as shown in Theorem 6 of Sect. 1.12 with the addition of some algebraic steps to pass from (1.720) of the theorem to (3.282) here above.

The proof of (3.283), (3.284) and (3.285) is based upon the reasoning underlying (3.306) above and takes advantage of (1.384) and (1.393) of Sect. 1.8

Results (3.286)–(3.289) are special cases of (3.267) above, resting on Corollary 6.1 of Sect. 1.12, upon resorting to the isomorphism of algebras of polynomials in a complex variables and in the lag operator (see also, in this connection, Proposition 6 of Sect. 3.2 and Theorem 1 of Sect. 3.5).

□

References

Banerjee A, Dolado J, Galbraith JW, Hendry DF (1993) Co-integration, error-correction and the econometric analysis of non stationary data. Oxford University Press, New York

Bellman R (1970) Introduction to matrix analysis, 2nd edn. McGraw Hill, New York

Beveridge S, Nelson CR (1981) A new approach to the decomposition of economic time series into permanent and transient components with particular attention to measurement of the business cycle. J. Monet Econ 7:151–174

Blanc-Lapierre A, Fortet R (1953) Théorie des fonctions aléatoires. Masson, Paris

Campbell SL, Meyer CD (1979) Generalized inverse of linear transformations. Pitman, London

Charenza WW, Deadman DF (1992) New directions in econometric practise. Edgar, Aldershot, UK

Chipman JS, Rao MM (1964) The treatment of linear restrictions in regression analysis. Econometrica 32:198–209

Cramer H, Leadbetter MR (1967) Stationary and related stochastic processes. Wiley, New York

Dhrymes PJ (1971) Distributed lags: problems of estimation and formulation. Holden Day, San Francisco

Elaydi SN (1996) An introduction to difference equations. Springer, New York

Faliva M (1987) Econometria: principi e metodi. UTET, Torino

Faliva M, Zoia MG (1999) Econometria delle serie storiche. Giappichelli, Torino

Faliva M, Zoia MG (2002) On a partitioned inversion formula having useful applications in econometrics. Econom Theory 18:525–530

Faliva M, Zoia MG (2003) A new proof of the representation theorem for I(2) processes. J Interdiscipl Math 6:331–347

Franchi M (2007) The integration order of vector autoregressive processes. Econom Theory 23:546–553

Fraser RA, Duncan WJ, Collar AR (1963) Elementary matrices and some applications to dynamics and differential equations. Cambridge University Press, Cambridge

Gantmacher FR (1959) The theory of matrices. Chelsea, New York

Goldberger AS (1964) Econometric theory. Wiley, New York

Greville TNE (1960) Some applications of the pseudoinverse of a matrix. Siam Rev 2:15–22

Hansen PR, Johansen S (1998) Workbook on cointegration. Oxford University Press, New York

Hatanaka M (1996) Time-series-based econometrics. Oxford University Press, New York

Jeffrey A (1992) Complex analysis and applications. CRC, London

Johansen S (1995) Likelihood-based inference in cointegrated vector autoregressive models. Oxford University Press, New York

Lancaster P, Tismenetsky M (1985) The theory of matrices, 2nd edn. Academic Press, New York

Lütkepohl H (1991) Introduction to multiple time series analysis. Springer, Berlin

Magnus JR, Neudecker H (1999) Matrix differential calculus with applications to statistics and econometrics. Wiley, New York

Markuscevich AI (1965) Theory of a complex variable. Prentice Hall, London

Marsaglia G, Styan GPH (1974) Equalities and inequalities for ranks of matrices. Linear Multilinear Algebra 2:269–292

Mickens RE (1990) Difference equations. Theory and applications, 2nd edn. Van Nostrand Reinhold, New York

Neudecker H (1968) The Kronecker matrix product and some of its applications in econometrics. Statistica Neerlandica 22:69–82

Papoulis A (1965) Probability, random variables and stochastic processes. McGraw Hill, New York

Pringle RM, Rayner AA (1971) Generalized inverse matrices with applications to statistics. Griffin, London

Rao CR (1973) Linear statistical inference and its applications, 2nd edn. Wiley, New York

Rao CR, Mitra SK (1971) Generalized inverse of matrices and its applications. Wiley, New York

Searle SR (1982) Matrix algebra useful for statistics. Wiley, New York

Stock JH, Watson MW (1988) Variable trends in economic time series. J Econ Persp 2:147–174

Styan GPH (1973) Hadamard products and multivariate statistical analysis. Linear Algebra Appl 6:217–240

Theil H (1971) Principles of econometrics. North Holland, Amsterdam

Vladimirov VS (1984) Equations of mathematical physics. MIR, Moscow

Yaglom AM (1962) Stationary random functions. Prentice Hall, Englewood Cliffs, NJ

Wold H (1938) A study in the analysis of stationary time series (2nd edn. 1954) Almqvist & Wiksells, Uppsala

Zoia MG (2006) A new algebraic approach to representation theorems for (co) integrated processes up to the second order. Cahiers du Département d'économétrie, Université de Genève 6:1–20

Notational Conventions, Symbols and Acronyms

The following notational conventions will be used throughout the text:

– Bold lower case letters indicate vectors.
– Bold upper case letters indicate matrices.
– Both notations $[\boldsymbol{A} \quad \boldsymbol{B}]$ and $[\boldsymbol{A}, \ \boldsymbol{B}]$ will be used, depending on convenience, for column-wise partitioned matrices
– Both notations $\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix}$ and $\begin{bmatrix} \boldsymbol{A} & \vdots & \boldsymbol{B} \\ \cdots & & \cdots \\ \boldsymbol{C} & \vdots & \boldsymbol{D} \end{bmatrix}$ will be used, depending on convenience, for block matrices

Symbols and Acronyms

	Meaning	*Section*
$\boldsymbol{A}^-$	generalized inverse of $\boldsymbol{A}$	1.1 (Definition 1)
$r(\boldsymbol{A})$	rank of $\boldsymbol{A}$	1.1
$\boldsymbol{A}_\rho^-$	reflexive generalized inverse	1.1 (Definition 2)
$\boldsymbol{A}'$	transpose of $\boldsymbol{A}$	1.1
$\boldsymbol{A}^g$	Moore-Penrose inverse of $\boldsymbol{A}$	1.1 (Definition 3)
$ind(\boldsymbol{A})$	index of $\boldsymbol{A}$	1.1 (Definition 5)
$\boldsymbol{A}^D, \boldsymbol{A}^{\neq}$	Drazin inverse of $\boldsymbol{A}$	1.1 (Definition 6)
$\boldsymbol{A}_r^-$	right inverse of $\boldsymbol{A}$	1.1 (Definition 7)
$\boldsymbol{A}_l^-$	left inverse of $\boldsymbol{A}$	1.1 (Definition 8)
$det(\boldsymbol{A})$	determinant of $\boldsymbol{A}$	1.1
$\boldsymbol{A}_\perp$	orthogonal complement of $\boldsymbol{A}$	1.2 (Definition 2)
$\boldsymbol{A}_s^-, (\boldsymbol{A}_\perp')_s^-$	specular directional inverses	1.2 (Remark 4)
$\boldsymbol{A}_l^\perp$	left orthogonal complement of $\boldsymbol{A}$	1.2 (Definition 3)
$\boldsymbol{A}_r^\perp$	right orthogonal complement of $\boldsymbol{A}$	1.2 (ibid)

$\boldsymbol{A}(z)$	matrix polynomial of z	1.6 (Definition 1)
$\dot{\boldsymbol{A}}(z), \ddot{\boldsymbol{A}}(z), \dddot{\boldsymbol{A}}(z)$	dot notation for derivatives	1.6
$\boldsymbol{A}, \dot{\boldsymbol{A}}, \ddot{\boldsymbol{A}}, \dddot{\boldsymbol{A}}$	short notation for $\boldsymbol{A}(1), \dot{\boldsymbol{A}}(1), \ddot{\boldsymbol{A}}(1), \dddot{\boldsymbol{A}}(1)$	1.6
$\boldsymbol{A}^{+}$	adjoint of $\boldsymbol{A}$	1.6
$tr\, \boldsymbol{A}$	trace of $\boldsymbol{A}$	1.6
$vec\, \boldsymbol{A}$	staked form of $\boldsymbol{A}$	1.6
L	lag operator	1.8
∇	backward difference operator	1.8 (Definition 1)
∇^{-1}	antidifference operator	1.8 (Definition 2)
Σ	indefinite sum operator	1.8 (ibidem)
$\boldsymbol{A}\otimes\boldsymbol{B}$	Kronecker product of $\boldsymbol{A}$ and $\boldsymbol{B}$	1.10
$\boldsymbol{u}$	vector of 1's	1.10
E	expectation operator	2.1
$\boldsymbol{\Gamma}(h)$	autocovariance matrix of order h	2.1
$I(d)$	integrated process of order d	2.1 (Definition 5)
$I(0)$	stationary process	2.1
$WN_{(n)}$	n-dimensional white noise	2.2 (Definition 1)
δ_v	discrete unitary function	2.2
VMA (q)	vector moving average process of order q	2.2 (Definition 2)
VAR (p)	vector autoregressive process of order p	2.2 (Definition 5)
VARMA (p, q)	vector autoregressive moving average process of order (p, q)	2.2 (Definition 7)
$CI(d, b)$	cointegrated system of order (d, b)	2.4 (Definition 6)
$PCI(d, b)$	polynomially cointegrated system of order (d, b)	2.4 (Definition 7)
$\boldsymbol{A} * \boldsymbol{B}$	Hadamard product of $\boldsymbol{A}$ and $\boldsymbol{B}$	3.1

List of Definitions

Section 1.1

Definition 1 *Generalized Inverse* 1
Definition 2 *Reflexive Generalized Inverse* 2
Definition 3 *Moore-Penrose Inverse* 2
Definition 4 *Nilpotent Matrix* 3
Definition 5 *Index of a Matrix* 3
Definition 6 *Drazin Inverse* 3
Definition 7 *Right Inverse* 4
Definition 8 *Left Inverse* 4

Section 1.2

Definition 1 *Row Kernel* 7
Definition 2 *Orthogonal Complement* 7
Definition 3 *Left and Right Orthogonal Complements* 17

Section 1.6

Definition 1 *Matrix Polynomial* 38
Definition 2 *Zero of a Matrix Polynomial* 38
Definition 3 *Nullity* 38
Definition 4 *Pole* 39
Definition 5 *Order of Poles and Zeros* 39
Definition 6 *Characteristic Polynomial* 43

Section 1.7

Definition 1 *Order of a Pole in a Laurent Expansion* 52

Section 1.8

Definition 1 *Backward Difference Operator* 65
Definition 2 *Antidifference Operator* 66

Section 2.1

Definition 1 *Stationary Processes* 128
Definition 2 *Stationarity in Mean* 128
Definition 3 *Covariance Stationarity* 129
Definition 4 *Stationarity in the Wide Sense* 129
Definition 5 *Integrated Processes* 129

Section 2.2

Definition 1 *White Noise* 130
Definition 2 *Vector Moving-Average Processes* 131
Definition 3 *First Difference of a White Noise* 133
Definition 4 *Second Difference of a White Noise* 133
Definition 5 *Vector Autoregressive Processes* 134
Definition 6 *Invertible Processes* 137
Definition 7 *Vector Autoregressive Moving-Average Processes* ... 139

Section 2.4

Definition 1 *Random Walk* 145
Definition 2 *Random Walk With Drift* 146
Definition 3 *Cumulated Random Walk* 147
Definition 4 *Deterministic Trends* 147
Definition 5 *Stochastic Trends* 148
Definition 6 *Cointegrated Systems* 149
Definition 7 *Polynomially Cointegrated Systems* 150

Section 3.2

Definition 1 *Basic VAR Model* 167
Definition 2 *Error Correction Model* 168