2000-17

UNIVERSITY OF CALIFORNIA, SAN DIEGO

DEPARTMENT OF ECONOMICS

THE JOHANSEN-GRANGER REPRESENTATION THEOREM:
AN EXPLICIT EXPRESSION FOR *I*(1) PROCESSES

BY

PETER REINHARD HANSEN

DISCUSSION PAPER 2000-17
JULY 2000

The Johansen-Granger Representation Theorem: An Explicit Expression for $I(1)$ Processes

Peter Reinhard Hansen †

UC San Diego and Brown University

Email: reinhard@ucsd.edu

Phone: (401) 863-3836

July, 2000

Abstract

The Johansen-Granger representation theorem for the cointegrated vector autoregressive process is derived using the companion form. This approach yields an explicit representation of all coefficients and initial values.

This result is useful for impulse response analysis, common feature analysis and asymptotic analysis of cointegrated processes.

†I thank Graham Elliott, James D. Hamilton, Hans Christian Kongsted, Anders Rahbek, and Halbert White for valuable comments. All errors remain my responsibility. Financial support from the Danish Social Science Research Council and the Danish Research Academy is gratefully acknowledged.

1. Introduction

The Johansen-Granger representation theorem[1] states that a vector autoregressive process $A(L)X_t = \varepsilon_t$, integrated of order one, has the representation $X_t = C\sum_{i=1}^{t}\varepsilon_i + C(L)\varepsilon_t + A_0$, where $\{C(L)\varepsilon_t\}$ is stationary if $\{\varepsilon_t\}$ is stationary and where A_0 depends on initial values $(X_0, X_{-1}, \ldots)$, (see Johansen (1991, 1996)). Johansen's result gives explicit values of C whereas the coefficients of the lag polynomial, $C(L)$, and the initial value, A_0, are given implicitly.

This representation of cointegrated processes is known as the Granger representation and is synonymous with the Wold representation for stationary processes. Because the representation divides X_t into a random walk and a stationary process, it can be viewed as multivariate Beveridge-Nelson decomposition where the labels are permanent and transitory components, (see Beveridge and Nelson (1981)).

The Granger representation is valuable in the asymptotic analysis of cointegrated processes, where typically only an explicit expression for C is needed. Explicit values for the coefficients in $C(L)$ are useful in common feature analysis, (see Engle and Kozicki (1993)), and in impulse response analysis, (see Lütkepohl and Reimers (1992), Warne (1993), and Lütkepohl and Saikkonen (1997)), where the coefficients of $C(L)$ are interpreted as the transitory effects of the shocks ε_t. Similarly, in asymptotic analysis of the model with structural breaks, see Hansen (2000), it is valuable to have an explicit value for A_0.

In this paper, explicit values of coefficients as well as initial values are found using the companion form, making use of the algebraic structure that characterizes this model.

From Johansen (1996) we adopt the following definitions: for an $m \times n$ matrix a with full column rank n, we define $\bar{a} = a(a'a)^{-1}$ and let the orthogonal complement of a, be the full rank $m \times (m-n)$ matrix $a_\perp$ that has $a_\perp' a = 0$.

In Section 2 the explicit representation is derived. In Section 3 we consider deterministic aspects of the representation. Section 4 contains concluding remarks and the appendix contains relevant algebra.

[1] The original Granger representation theorem, given by Engle and Granger (1987), asserts the existence of an error correction representation of X_t, under the assumptions that ΔX_t and $\beta' X_t$ have stationary and invertible VARMA representations, for some matrix β. The Johansen-Granger representation theorem, of Johansen (1991, 1996), makes assumptions on the autoregressive parameters, that precisely characterizes $I(1)$ processes, and states results on the moving average representation of X_t.

2. The Granger Representation for Autoregressive Processes Integrated of Order One

We consider the p-dimensional vector autoregressive process of order k

$$X_t = \Pi_1 X_{t-1} + \Pi_2 X_{t-2} + \cdots + \Pi_k X_{t-k} + \Phi D_t + \varepsilon_t, \quad t = 1, \ldots, T,$$

where the process' deterministic terms are contained in D_t and where ε_t, $t = 1, \ldots, T$ is a sequence of independent identically distributed stochastic variables with mean zero[2].

The process can be re-written in error correction form:

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Phi D_t + \varepsilon_t, \quad t = 1, \ldots, T$$

where $\Pi = -I + \sum_{i=1}^{k} \Pi_i$ and $\Gamma_i = -\sum_{j=i+1}^{k} \Pi_j$. The conditions that ensure that X_t is integrated of order one, referred to as X_t being $I(1)$, are stated in the following assumption:

Assumption 2.1. *The assumptions of the Johansen-Granger representation theorem are:*

(i) *The roots of the characteristic polynomial*

$$\det(A(z)) = \det(I - \Pi_1 z - \Pi_2 z^2 - \cdots \Pi_k z^k)$$

are either outside the unit circle or equal to one.

(ii) *The matrix Π has reduced rank $r < p$, and can therefore be expressed as the product $\Pi = \alpha\beta'$ where α and β are $p \times r$ matrices of full column rank r.*

(iii) *The matrix $\alpha_\perp' \Gamma \beta_\perp$ has full rank, where $\Gamma = I - \sum_{i=1}^{k-1} \Gamma_i$ and where $\alpha_\perp$ and $\beta_\perp$ are the orthogonal complements to α and β.*

The first assumption ensures that the process is not explosive (roots in the unit circle) or seasonally cointegrated (roots on the boundary of the unit circle different from $z = 1$), (see Hylleberg, Engle, Granger, and Yoo (1990) or Johansen and Schaumburg (1998)). The second ensures that there are at least $p - r$ unit roots and induces cointegration whenever $r \geq 1$. The

[2] The Granger representation is not relying on the assumptions on ε_t, since it is entirely an algebraic derivation. However the i.i.d. assumption is important for some of the interpretations of the representation.

third assumption restricts the process from being $I(2)$, because (iii) together with (ii) ensures that the number of unit roots is exactly $p-r$.

Under these assumptions, Johansen (1991) showed that X_t has the representation $X_t = C\sum_{i=1}^{t}(\varepsilon_i + \Phi D_i) + C(L)(\varepsilon_t + \Phi D_t) + A$, where $C = \beta_\perp \left(\alpha_\perp' \Gamma \beta_\perp\right)^{-1} \alpha_\perp'$. By using the companion form of the process, it is possible to obtain explicit values for the coefficients of the lag polynomial $C(L) = C_0 + C_1 L + C_2 L^2 + \cdots$, and the initial values contained in A, as I show below.

The following lemma will be useful.

Lemma 2.2. *Let a and b be $m \times n$ matrices, $m \geq n$ with full column rank n, and let $a_\perp$ and $b_\perp$ be their orthogonal complements, respectively.*

The following five statements are equivalent.

- (i) *The matrix $(I + b'a)$ does not have 1 as an eigenvalue.*
- (ii) *Let v be a vector in $\mathbf{R}^m$. Then $(b'a)v = 0$ implies $v = 0$.*
- (iii) *The matrix $b'a$ has full rank.*
- (iv) *The $m \times m$ matrix $(b, a_\perp)$ has full rank.*
- (v) *The matrix $b_\perp' a_\perp$ has full rank.*

Proof. The equivalence of (i), (ii) and (iii) is straightforward, and the identity

$$|(a, a_\perp)|\,|(b, a_\perp)| = |(a, a_\perp)'(b, a_\perp)| = |\begin{pmatrix} a'b & 0 \\ a_\perp' b & a_\perp' a_\perp \end{pmatrix}| = |a'b|\,|a_\perp' a_\perp|$$

proves that (iii) holds if and only if (iv) holds. Finally, the identity

$$|(b, b_\perp)|\,|(b, a_\perp)| = |(b, b_\perp)'(b, a_\perp)| = |\begin{pmatrix} b'b & 0 \\ b_\perp' b & b_\perp' a_\perp \end{pmatrix}| = |b'b|\,|b_\perp' a_\perp|$$

completes the proof. ■

2.1. The Companion Form

We transform the process into the companion form, by defining

$$X_t^* = \left(X_t', X_{t-1}', \ldots, X_{t-k+1}'\right)'$$

so that with suitable definitions

$$\begin{aligned}\Delta X_t^* &= \Pi^* X_{t-1}^* + \Phi_t^* + \varepsilon_t^* \\ &= \alpha^* \beta^{*\prime} X_{t-1}^* + \Phi_t^* + \varepsilon_t^*,\end{aligned}$$

which converts the process to a vector autoregressive process of order one. The needed definitions are

$$\Pi^* = \begin{pmatrix} \alpha\beta' + \Gamma_1 & \Gamma_2 - \Gamma_1 & \cdots & \Gamma_{k-1} - \Gamma_{k-2} & -\Gamma_{k-1} \\ I & -I & & & 0 \\ & & \ddots & & \vdots \\ & & & -I & 0 \\ 0 & 0 & & I & -I \end{pmatrix},$$

$$\alpha^* = \begin{pmatrix} \alpha & \Gamma_1 & \cdots & \Gamma_{k-1} \\ 0 & I & & 0 \\ \vdots & & \ddots & \\ 0 & & & I \end{pmatrix}, \beta^* = \begin{pmatrix} \beta & I & 0 & \cdots & 0 \\ 0 & -I & I & & \\ \vdots & & & \ddots & \\ & & & & I \\ 0 & \cdots & 0 & & -I \end{pmatrix},$$

$$\varepsilon_t^* = \begin{pmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \Phi_t^* = \begin{pmatrix} \Phi D_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

It is easily verified that the orthogonal complements of α^* and β^* are given by

$$\alpha_\perp^* = \begin{pmatrix} \alpha_\perp \\ -\Gamma_1'\alpha_\perp \\ \vdots \\ -\Gamma_{k-1}'\alpha_\perp \end{pmatrix}, \quad \beta_\perp^* = \begin{pmatrix} \beta_\perp \\ \vdots \\ \beta_\perp \end{pmatrix}.$$

Lemma 2.3. *Let α, β, α^* and β^* be defined as above, and assume that Assumption 1.2.1 holds. Then the eigenvalues of the matrix $(I + \beta^{*\prime}\alpha^*)$ are all less than one in absolute value.*

Proof. By Assumption 1.2.1 (iii), the identity

$$\alpha_\perp^{*\prime}\beta_\perp^* = \alpha_\perp'(I - \Gamma_1 - \cdots - \Gamma_{k-1})\beta_\perp$$

shows that $\alpha_\perp^{*\prime}\beta_\perp^*$ has full rank, and by Lemma 1.2.2, we have that 1 is not an eigenvalue of $(I+\beta^{*\prime}\alpha^*)$. However we need to show that the eigenvalues are smaller than one in absolute value. Therefore consider an eigenvector $v = (v_1', \ldots, v_k')' \neq 0$ of $(I+\beta^{*\prime}\alpha^*)$, e.g. $(I+\beta^{*\prime}\alpha^*)v = \lambda v$. The upper $r+p$ rows of $(I+\beta^{*\prime}\alpha^*)v$ yields

$$\begin{aligned} v_1 + \beta'(\alpha v_1 + \Gamma_1 v_2 + \cdots + \Gamma_{k-1}v_k) &= \lambda v_1 \\ (\alpha v_1 + \Gamma_1 v_2 + \cdots + \Gamma_{k-1}v_k) &= \lambda v_2 \end{aligned}$$

which implies $\lambda\beta' v_2 = (\lambda - 1)v_1$, and the remaining part implies $v_2 = \lambda v_3 = \cdots = \lambda^{k-2}v_k$. The case $\lambda = 0$ clearly fulfills $|\lambda| < 1$ so assume $\lambda \neq 0$. Multiply the second set of equations by $(\lambda-1)/\lambda^k$ and substitute $z = 1/\lambda$ to obtain

$$[I(1-z) - \alpha\beta' z - \Gamma_1(1-z)z - \cdots - \Gamma_{k-1}(1-z)z^{k-1}]v_k = 0.$$

This is equivalent to

$$\left| I(1-z) - \alpha\beta' z - \sum_{i=1}^{k-1}\Gamma_i(1-z)z^i \right| = 0,$$

and since Assumption 1.2.1 has $|z| > 1$ we conclude that $|\lambda| < 1$. ■

The result has the implication that under Assumption 1.2.1 the sum $\sum_{i=0}^{\infty}(1+\beta^{*\prime}\alpha^*)^i$ is convergent with limit $(\beta^{*\prime}\alpha^*)^{-1}$, such that a process defined by $Y_t = \sum_{i=0}^{\infty}(1+\beta^{*\prime}\alpha^*)^i u_{t-i}$ is stationary whenever u_t is stationary.

Lemma 2.4. *With the definitions above we have the identities:*

$$\begin{aligned} (I - C\Gamma) &= (I - C\Gamma)\bar{\beta}\beta' \\ I &= (I - C\Gamma)\bar{\beta}\beta' + C(\Gamma - I) + C\bar{\alpha}_\perp\alpha_\perp'. \end{aligned}$$

Proof. Since $I = \beta(\beta'\beta)^{-1}\beta' + \beta_\perp(\beta_\perp'\beta_\perp)^{-1}\beta_\perp' = \bar{\beta}\beta' + \beta_\perp\bar{\beta}_\perp'$, the first identity follows from

$$(I - C\Gamma) = (I - C\Gamma)(\bar{\beta}\beta' + \beta_\perp\bar{\beta}_\perp')$$

$$
\begin{aligned}
&= (I - C\Gamma)\bar{\beta}\beta' + \beta_\perp \bar{\beta}_\perp' - \beta_\perp(\alpha_\perp'\Gamma\beta_\perp)^{-1}\alpha_\perp'\Gamma\beta_\perp\bar{\beta}_\perp' \\
&= (I - C\Gamma)\bar{\beta}\beta',
\end{aligned}
$$

and the second follows by applying the first identity and that $C = C\bar{\alpha}_\perp\alpha_\perp'$. ■

We are now ready to formulate the main result.

Theorem 2.5 (The Johansen-Granger representation theorem). *Let a process be given by the equation*

$$
\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Phi D_t + \varepsilon_t,
$$

and assume that Assumption 1.2.1 holds. Then the process has the representation

$$
X_t = C\sum_{i=1}^{t}(\varepsilon_i + \Phi D_i) + C(L)(\varepsilon_t + \Phi D_t) + C(X_0 - \Gamma_1 X_{-1} - \cdots - \Gamma_{k-1}X_{-k+1})
$$

where $C = \beta_\perp \left(\alpha_\perp'\Gamma\beta_\perp\right)^{-1}\alpha_\perp'$ and where the coefficients of $C(L)$ are given by

$$
C_i = GQ^i E_{1,2}
$$

where

$$
\begin{aligned}
G &= ((I - C\Gamma), -C\Gamma_1, \ldots, -C\Gamma_{k-1}) \\
Q &= \begin{pmatrix}
I + \Pi & \Gamma_1 & \cdots & \Gamma_{k-2} & \Gamma_{k-1} \\
\Pi & \Gamma_1 & \cdots & \Gamma_{k-2} & \Gamma_{k-1} \\
0 & I & & 0 & 0 \\
\vdots & & \ddots & & \vdots \\
0 & & & I & 0
\end{pmatrix} \\
E_{1,2} &= (I_p, I_p, 0, \cdots, 0)'.
\end{aligned}
$$

Proof. Under Assumption 1.2.1 the $pk \times pk$ matrix $(\beta^*, \alpha_\perp^*)$ has full rank. We can therefore obtain the Granger representation for X_t^* by finding the moving average representation for the processes $\beta^{*\prime}X_t^*$ and $\alpha_\perp^{*\prime}X_t^*$ individually and then stacking them and multiplying by $(\beta^*, \alpha_\perp^*)^{\prime -1}$.

First, consider the process

$$
\beta^{*\prime}X_t^* = (I + \beta^{*\prime}\alpha^*)\beta^{*\prime}X_{t-1}^* + \beta^{*\prime}(\varepsilon_t^* + \Phi_t^*).
$$

Since all the eigenvalues of $(I+\beta^{*\prime}\alpha^*)$, according to Lemma 2.3, are smaller than one in absolute value, the process has the stationary representation

$$\beta^{*\prime}X_t^* = C^*(L)(\varepsilon_t^* + \Phi_t^*)$$

where $C_i^* = (I+\beta^{*\prime}\alpha^*)^i\beta^{*\prime}$, and where by stationary we mean that $\beta^{*\prime}X_t^* - E\left(\beta^{*\prime}X_t^*\right)$ is stationary.

Next consider the random walk

$$\begin{aligned}\alpha_\perp^{*\prime}X_t^* &= \alpha_\perp^{*\prime}X_{t-1}^* + \alpha_\perp^{*\prime}(\varepsilon_t^* + \Phi_t^*) \\ &= \alpha_\perp^{*\prime}X_0^* + \sum_{i=1}^t \alpha_\perp^{*\prime}(\varepsilon_i^* + \Phi_i^*).\end{aligned}$$

A representation for X_t^* is now obtained as

$$X_t^* = (\beta^*, \alpha_\perp^*)^{\prime -1}\begin{pmatrix} C^*(L)(\varepsilon_t^* + \Phi_t^*) \\ \sum_{i=1}^t \alpha_\perp^{*\prime}(\varepsilon_i^* + \Phi_i^*) + \alpha_\perp^{*\prime}X_0^* \end{pmatrix}.$$

The entire matrix $(\beta^*, \alpha_\perp^*)^{\prime -1}$ is given in the Appendix, for our purposes we only need its upper p rows that define the equation for X_t. These rows are given by

$$\left((I - C\Gamma)\bar{\beta}, -C\Gamma_1^s, \ldots, -C\Gamma_{k-1}^s, C\bar{\alpha}_\perp\right)$$

with the definition $\Gamma_i^s = \Gamma_i + \cdots + \Gamma_{k-1}$. For simplicity, we define

$$F = \left((I - C\Gamma)\bar{\beta}, -C\Gamma_1^s, \ldots, -C\Gamma_{k-1}^s\right)$$

and obtain the representation for X_t:

$$\begin{aligned}X_t &= (F, C\bar{\alpha}_\perp)\begin{pmatrix} C^*(L)(\varepsilon_t^* + \Phi_t^*) \\ \sum_{i=1}^t \alpha_\perp^{*\prime}(\varepsilon_i^* + \Phi_i^*) + \alpha_\perp^{*\prime}X_0^* \end{pmatrix} \\ &= FC^*(L)(\varepsilon_t^* + \Phi_t^*) + C\bar{\alpha}_\perp\sum_{i=1}^t \alpha_\perp^{*\prime}(\varepsilon_i^* + \Phi_i^*) + C\bar{\alpha}_\perp\alpha_\perp^{*\prime}X_0^* \\ &= C(L)(\varepsilon_t + \Phi D_t) + C\sum_{i=1}^t(\varepsilon_i + \Phi D_i) + A,\end{aligned}$$

where the initial value is explicitly given by

$$A = C\bar{\alpha}_\perp \alpha_\perp^{*\prime} X_0^* = C(X_0 - \Gamma_1 X_{-1} - \cdots - \Gamma_{k-1} X_{-(k-1)}),$$

and the coefficients of the polynomial $C(L)$ are given by

$$\begin{aligned} C_i &= F(I + \beta^{*\prime}\alpha^*)^i \beta^{*\prime} E_1 \\ &= F D^i B' E_{1,2}, \end{aligned}$$

with the additional definitions

$$\begin{aligned} D &= (I + \beta^{*\prime}\alpha^*) \\ B &= \begin{pmatrix} \beta & 0 & 0 & \cdots & 0 \\ 0 & I & 0 & & 0 \\ 0 & 0 & 0 & & 0 \\ \vdots & & & \ddots & \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, \; E_1 = \begin{pmatrix} I_p \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \text{ and } E_{1,2} = \begin{pmatrix} I_p \\ I_p \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \end{aligned}$$

Because $(I + \beta'\alpha)\beta' = \beta'(I + \alpha\beta')$ we have that

$$DB' = B'Q$$

where Q is as given in the theorem. Thus, the coefficients can be written as

$$C_i = F D^i B' E_{1,2} = F B' Q^i E = G Q^i E_{1,2}$$

where

$$G = FB' = \left((I - C\Gamma), -C\Gamma_1^s, \ldots, -C\Gamma_{k-1}^s\right),$$

where we applied the identity $(I - C\Gamma)\bar{\beta}\beta' = (I - C\Gamma)$ of Lemma 2.4. This completes the proof. ■

Corollary 2.6. *The coefficients of $C(L)$ can be obtained recursively from the formula*

$$C_i = C_{i-1} + \sum_{j=1}^{i} (\Pi + \Gamma_j)\Delta C_{i-j}, \quad i = 1, 2, \ldots,$$

where $C_0 = I - C$ and $\Delta C_0 = I$ and where we set $\Gamma_j = 0$ for $j \geq k$.

Proof. From the proof of the Johansen-Granger representation theorem we have that $C_i = GQ^i E_{1,2}$. So by defining

$$A_i = \begin{pmatrix} A_{1,i} \\ A_{2,i} \\ \vdots \\ A_{k,i} \end{pmatrix} = QA_{i-1} = Q^i E_{1,2},$$

tedious algebra (given in the Appendix) leads to the relation

$$C_i = C_{i-1} + A_{2,i}, \quad C_{-1} = -C, \quad i = 0, 2, \ldots$$

and the structure of Q yields the equation

$$A_{2,i} = \sum_{j=1}^{i} (\Pi + \Gamma_j) A_{2,i-j}, \quad A_{2,0} = I \quad i = 1, 2, \ldots .$$

By inserting $A_{2,i-j} = C_{i-j} - C_{i-j-1}$ we find the equation of the corollary. ■

As a special case we formulate the representation for the vector autoregressive process of order one.

Corollary 2.7. *Let $\Delta X_t = \alpha\beta' X_{t-1} + \varepsilon_t$ be a process fulfilling Assumption 1.2.1. Then we have the representation*

$$X_t = C \sum_{i=1}^{t} \varepsilon_i + (1 - C) \sum_{i=0}^{\infty} \left(I + \alpha\beta'\right)^i \varepsilon_{t-i} + CX_0$$

where $C = \beta_\perp \left(\alpha'_\perp \beta_\perp\right) \alpha'_\perp$.

The result of Corollary 2.7 is derived directly in Johansen (1996) by dividing the process into its stationary and non-stationary part with the identity $I = \alpha \left(\beta'\alpha\right) \beta' + \beta_\perp \left(\alpha'_\perp \beta_\perp\right) \alpha'_\perp$. The proof of Theorem 2.5 made use of the more general identity $I = (I - C\Gamma)\bar{\beta}\beta' + C(\Gamma - I) + C\bar{\alpha}_\perp \alpha'_\perp$ of Lemma 2.4, which simplifies to the identity in Johansen (1996) when $\Gamma = I$, as is the case for a VAR(1) process.

3. Deterministic Terms

In this section we study the stationary polynomial's role for the deterministic term. The deterministic part plays an important role for the asymptotic analysis of this model, because the limits of some test statistics depend on the deterministic term. The literature has developed a notation for models with different deterministic terms which we shall adopt.

First we analyze the model H_1. This model contains only a constant $\Phi D_t = \mu_0$, which in general will give rise to a linear trend in the process X_t. Next, we also analyze its sub-model H_1^*, which has the deterministic term $\Phi D_t = \alpha\rho_0$. This is equivalent to the restriction on the constant $C\mu = 0$, which is precisely what is needed for X_t not to have a linear trend. We also analyze the models H and H^*. Model H has a linear deterministic trend $\Phi D_t = \mu_0 + \mu_1 t$, which gives rise to a quadratic trend in the process X_t, and the sub-model H^*, has the deterministic trend restricted to $\Phi D_t = \mu_0 + \alpha\rho_1 t$, which prevents the X_t from having a quadratic trend.

3.1. The Models H_1 and H_1^*

When the deterministic term is simply a constant $\mu_0 = \Phi D_t$, the Granger representation is given by

$$X_t = C\sum_{i=1}^{t}\varepsilon_{t-i} + C(L)\varepsilon_t + C(1)\mu_0 + C\mu_0 t + A.$$

So unless $C\mu_0 = 0$, the constant μ_0 leads to a deterministic linear trend in the process X_t. The matrix $C(1)$ is calculated in the appendix and is found to be

$$\begin{aligned} C(1) &= -(I - C\Gamma)\bar{\beta}\bar{\alpha}'\left(I - \Gamma C\right) - C\left(\sum_{i=1}^{k-1} i\Gamma_i\right)C \\ &= -BA' - C\Psi C, \end{aligned}$$

where $B = (I - C\Gamma)\bar{\beta}$, $A' = \bar{\alpha}'\left(\Gamma C - I\right)$ and $\Psi = \sum_{i=1}^{k-1} i\Gamma_i = \sum_{i=1}^{k-1}\sum_{j=i}^{k-1}\Gamma_j$.

This result encompasses two findings from Hansen and Johansen (1998). The first is that

$$E(\beta' X_t) = \beta' C(1)\mu_0 = \bar{\alpha}'(\Gamma C - I)\mu_0,$$

and the second is that in H_1^*, where $\mu_0 = \alpha\rho_0$, the linear trend vanishes while the constant in the process is given by $C(1)\mu = -(I - C\Gamma)\bar{\beta}\rho$.

3.2. Models H and H^*

When the deterministic term contains a linear trend, $\Phi D_t = \mu_0 + \mu_1 t$, the deterministic part of the Granger representation is given by

$$\tfrac{1}{2}C\mu_1 t^2 + C\left(\mu_0 + \tfrac{1}{2}\mu_1\right)t + C(L)\left(\mu_0 + \mu_1 t\right),$$

(see Hansen and Johansen (1998)). This can be re-written as

$$\tfrac{1}{2}C\mu_1 t^2 + \left(C\mu_0 + \left(\tfrac{1}{2}C + C(1)\right)\mu_1\right)t + \left(C(1)\mu_0 - \sum_{i=0}^{\infty} iC_i\mu_1\right). \tag{3.1}$$

So unless $\alpha_\perp'\mu_1 = 0$ the linear trend μ_1 leads to a quadratic deterministic trend in the process X_t. The only term of (3.1) not derived previously, is $\sum_{i=0}^{\infty} iC_i$. This term is derived in the appendix and is given by

$$BA' + C\Psi C + BA'\Gamma BA' - BA'\Psi C - C\Psi BA' - C\Psi C\Psi C - \frac{k(k-1)}{2}C\Psi C.$$

In model H^* where the linear trend is restricted to $\mu_1 = \alpha\rho_1$, (3.1) reduces to

$$\left(C\mu_0 - (I - C\Gamma)\bar{\beta}\rho_1\right)t + C(1)\mu_0 + \left((C\Gamma - I)\bar{\beta} - C(1)\right)\rho_1$$

which encompasses a result from Johansen (1996, equation 5.20), because the expression for τ_1, in Johansen (1996, equation 5.20), equals $\left(C\mu_0 - (I - C\Gamma)\bar{\beta}\rho_1\right)$.

4. Conclusion

We gave an explicit expression of the moving average representation for processes integrated of order one using the companion form for the process. The explicit expression is useful to have in studies of impulse response functions and in common features analysis. As a side benefit the approach gives a new proof of the Johansen-Granger representation theorem, a proof that some might find more intuitive and easy to follow than previous proofs.

Appendix A: Proofs

A.1. The Inverse of $(\beta^*, \alpha_\perp^*)$

In the proof of the Johansen-Granger representation theorem we need an explicit expression for the first p rows of $(\beta^*, \alpha_\perp^*)'^{-1}$. The entire matrix is given by

$$(\beta^*, \alpha_\perp^*)'^{-1} = \begin{pmatrix} (I - C\Gamma)\bar{\beta} & -C\Gamma_1^s & \cdots & & -C\Gamma_{k-1}^s & C\bar{\alpha}_\perp \\ (I - C\Gamma)\bar{\beta} & -C\Gamma_1^s - I & \ddots & & -C\Gamma_{k-1}^s & C\bar{\alpha}_\perp \\ (I - C\Gamma)\bar{\beta} & -C\Gamma_1^s & \ddots & \ddots & -C\Gamma_{k-1}^s & C\bar{\alpha}_\perp \\ \vdots & \vdots & & \ddots & \ddots \quad \vdots & \vdots \\ (I - C\Gamma)\bar{\beta} & -C\Gamma_1^s & & & \ddots \quad -C\Gamma_{k-1}^s & C\bar{\alpha}_\perp \\ (I - C\Gamma)\bar{\beta} & -C\Gamma_1^s & & & -C\Gamma_{k-1}^s - I & C\bar{\alpha}_\perp \end{pmatrix},$$

which is verified by multiplying it by $(\beta^*, \alpha_\perp^*)'$ and using the identity $(I - C\Gamma)\bar{\beta}\beta' = (I - C\Gamma)$ from Lemma 2.4.

A.2. The Expression for C_i

In Corollary 2.6 we asserted the relation $C_i = C_{i-1} + A_{2,i},\ i = 1, 2, \ldots$. This relation is proved as follows. First, notice from the equation for A_i given by

$$A_i = \begin{pmatrix} A_{1,i} \\ A_{2,i} \\ \vdots \\ A_{k,i} \end{pmatrix} = \begin{pmatrix} I + \Pi & \Gamma_1 & \cdots & \Gamma_{k-2} & \Gamma_{k-1} \\ \Pi & \Gamma_1 & \cdots & \Gamma_{k-2} & \Gamma_{k-1} \\ 0 & I & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & & I & 0 \end{pmatrix} A_{i-1}, \quad A_0 = E_{1,2}$$

that $A_{k,i} = A_{2,i-k+2}$ $k \geq 2$, and that $A_{1,i} = A_{1,i-1} + A_{2,i}$, why $A_{1,i} = \sum_{j=1}^{i} A_{2,j}$. So that

$$A_{2,i} = \sum_{j=1}^{i} (\Pi + \Gamma_j) A_{2,i-j}, \quad A_{2,0} = I \quad i = 1, 2, \ldots,$$

and note that $CA_{2,i} = C\Gamma_1 A_{2,i-1} + \cdots + C\Gamma_{k-1} A_{k,i-1}$.

Next consider

$$C_i \;=\; GA_i = (I - C\Gamma)\, A_{1,i} - C\Gamma_1^s A_{2,i} - \cdots - C\Gamma_{k-1}^s A_{k,i}$$

$$
\begin{aligned}
&= (I - C\Gamma)\left(A_{2,i} + A_{1,i-1}\right) + C(\Gamma - I)A_{2,i} - C\Gamma_2^s A_{3,i} - \cdots - C\Gamma_{k-1}^s A_{k,i} \\
&= (I - C)\,A_{2,i} + (I - C\Gamma)\,A_{1,i-1} - C\Gamma_2^s A_{3,i} - \cdots - C\Gamma_{k-1}^s A_{k,i} \\
&= (I - C)\,A_{2,i} + (I - C\Gamma)\,A_{1,i-1} \\
&\quad - C(\Gamma_1^s - \Gamma_1)A_{2,i-1} - \cdots - C(\Gamma_{k-2}^s - \Gamma_{k-2})A_{k-1,i-1} - C(\Gamma_{k-1} - \Gamma_{k-1})A_{k,i-1} \\
&= GA_{i-1} + A_{2,i} \\
&= C_{i-1} + A_{2,i}.
\end{aligned}
$$

Which completes the proof.

A.3. An Expression for $C(1)$

In the analysis of the deterministic terms we need to calculate

$$
C(1) = F\sum_{i=0}^{\infty}\left(I + \beta^{*\prime}\alpha^*\right)^i \beta^{*\prime}E_1 = -F\left(\beta^{*\prime}\alpha^*\right)^{-1}\beta^{*\prime}E_1.
$$

The inverse of

$$
\beta^{*\prime}\alpha^* = \begin{pmatrix}
\beta'\alpha & \beta'\Gamma_1 & \beta'\Gamma_2 & \cdots & \beta'\Gamma_{k-2} & \beta'\Gamma_{k-1} \\
\alpha & \Gamma_1 - I & \Gamma_2 & \cdots & \Gamma_{k-2} & \Gamma_{k-1} \\
0 & I & -I & & & 0 \\
\vdots & & \ddots & \ddots & & \vdots \\
 & & & \ddots & -I & 0 \\
0 & & & & I & -I
\end{pmatrix},
$$

is given by

$$
\left(\beta^{*\prime}\alpha^*\right)^{-1} = \begin{pmatrix}
\bar{\alpha}'\left(I - \Gamma C\right)\Gamma\bar{\beta} & \bar{\alpha}'\left(I - \Gamma C\right)\Gamma_1^s & \cdots & \bar{\alpha}'\left(I - \Gamma C\right)\Gamma_{k-1}^s \\
\left(I - C\Gamma\right)\bar{\beta} & -C\Gamma_1^s - I & & -C\Gamma_{k-1}^s \\
\vdots & & \ddots & \\
\left(I - C\Gamma\right)\bar{\beta} & -C\Gamma_1^s - I & \cdots & -C\Gamma_{k-1}^s - I
\end{pmatrix}.
$$

So

$$\left(\beta^{*\prime}\alpha^{*}\right)^{-1}\beta^{*\prime} = \begin{pmatrix} \bar{\alpha}'\left(I-\Gamma C\right) & -\bar{\alpha}'\left(I-\Gamma C\right)\Gamma_1^s & \cdots & -\bar{\alpha}'\left(I-\Gamma C\right)\Gamma_{k-1}^s \\ -C & C\Gamma_1^s + I & & C\Gamma_{k-1}^s \\ \vdots & & \ddots & \\ -C & C\Gamma_1^s & & C\Gamma_{k-1}^s + I \end{pmatrix}$$

and therefore we find

$$\left(\beta^{*\prime}\alpha^{*}\right)^{-1}\beta^{*\prime}E_1 = \begin{pmatrix} \bar{\alpha}'\left(I-\Gamma C\right) \\ -C \\ \vdots \\ -C \end{pmatrix}, \tag{A.1}$$

and finally that

$$\begin{aligned} C(1) &= -\left((I-C\Gamma)\bar{\beta}, -C\Gamma_1^s, \ldots, -C\Gamma_{k-1}^s\right)\begin{pmatrix} \bar{\alpha}'\left(I-\Gamma C\right) \\ -C \\ -C \\ -C \end{pmatrix} \\ &= (I-C\Gamma)\bar{\beta}\bar{\alpha}'\left(\Gamma C - I\right) - C\left(\sum_{i=1}^{k-1}\sum_{j=i}^{k-1}\Gamma_j\right)C. \\ &= BA' - C\Psi C \end{aligned}$$

where $B = (I-C\Gamma)\bar{\beta}$, $A' = \bar{\alpha}'\left(\Gamma C - I\right)$ and $\Psi = \sum_{i=1}^{k-1}\sum_{j=i}^{k-1}\Gamma_j = \sum_{i=1}^{k-1} i\Gamma_i$.

A.4. An Expression for $\sum_{i=0}^{\infty} iC_i$

In the case where the deterministic term is given by $\Phi D_t = \mu_0 + \mu_1 t$ we make use of

$$\sum_{i=0}^{\infty}\left(I+\beta^{*\prime}\alpha^{*}\right)^{i} i = \left(\left(\beta^{*\prime}\alpha^{*}\right)^{-2} + \left(\beta^{*\prime}\alpha^{*}\right)^{-1}\right).$$

The second term is calculated in the case with $\Phi D_t = \mu$, and the term we need to add is given by

$$\left(\beta^{*\prime}\alpha^{*}\right)^{-2}\beta^{*\prime}E_1 = \begin{pmatrix} A'\Gamma BA' - A'\Psi C \\ BA' + C\Psi C + C \\ \vdots \\ BA' + C\Psi C + (k-1)\,C \end{pmatrix}$$

thus

$$\begin{aligned} F\left(\beta^{*\prime}\alpha^{*}\right)^{-2}\beta^{*\prime}E_1 &= BA'\Gamma BA' - BA'\Psi C \\ &\quad - C\Psi BA' - C\Psi C\Psi C - \frac{k(k-1)}{2}C\Psi C \end{aligned}$$

so that

$$\begin{aligned} \sum_{i=0}^{\infty} C_i i &= F\left(\left(\beta^{*\prime}\alpha^{*}\right)^{-1} + \left(\beta^{*\prime}\alpha^{*}\right)^{-2}\right)\beta^{*\prime}E_1 \\ &= BA' + C\Psi C + BA'\Gamma BA' - BA'\Psi C \\ &\quad - C\Psi BA' - C\Psi C\Psi C - \frac{k(k-1)}{2}C\Psi C. \end{aligned}$$

References

Beveridge, S., and C. R. Nelson (1981): "A New Approach to Decompositions of Time Series Into Permanent and Transitory Components with Particular Attentions to Measurement of the 'Business Cycle'," *Journal of Monetary Economics*, 7, 151–174.

Engle, R. F., and C. W. J. Granger (1987): "Co-Integration and Error Correction: Representation, Estimation and Testing," *Econometrica*, 55, 251–276.

Engle, R. F., and S. Kozicki (1993): "Testing for Common Features," *Journal of Business and Economic Statistics*, 11, 369–380.

Hansen, P., and S. Johansen (1998): *Workbook on Cointegration*. Oxford University Press, Oxford.

Hansen, P. R. (2000): "Structural Changes in Cointegrated Processes," Ph.D. thesis, University of California at San Diego.

Hylleberg, S., R. F. Engle, C. W. J. Granger, and S. Yoo (1990): "Seasonal Integration and Cointegration," *Journal of Econometrics*, 44, 215–238.

Johansen, S. (1991): "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica*, 59, 1551–1580.

——— (1996): *Likelihood Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press, Oxford, 2nd edn.

Johansen, S., and E. Schaumburg (1998): "Likelihood Analysis of Seasonal Cointegration," *Journal of Econometrics*, 88, 301–339.

Lütkepohl, H., and H.-E. Reimers (1992): "Impulse Response Analysis of Cointegrated Systems," *Journal of Enonomic Dynamics and Control*, 16, 53–78.

Lütkepohl, H., and P. Saikkonen (1997): "Impulse Response Analysis in Infinite Order Cointegrated Vector Autoregressive Processes," *Journal of Econometrics*, 81, 127–157.

Warne, A. (1993): "A Common Trends Model: Identification, Estimation and Inference," *Seminar Paper No. 555, IIES, Stockholm University.*