

A Characterization Of
Cointegration

By Phoebus J. Dhrymes, Columbia University

Working Paper No. 659

8/7/93

dp-9293-6
pgs-12
(Double Sided)

A CHARACTERIZATION OF COINTEGRATION

Phoebus J. Dhrymes*

Columbia University

08/07/93

Abstract

In this paper we revisit the definition and characterization of cointegration given in Engle and Granger (1987) (EG). In addition to correcting a number of errors and misstatements in that paper, we redefine the meaning of cointegration in the context of a $MAR(\infty)$ (multivariate AR) rendition of the stochastic sequence, rather than the $MMA(\infty)$ rendition of **differenced sequence** as in EG.

1 Introduction

The important paper by EG (1987) is often quoted and utilized extensively. Its usefulness, however, is marred by a number of misstatements that mislead practitioners and may lead to serious error. The plan of the paper is to offer an alternative proof of the representation theorem, given therein, and in the process point out the difficulties in the application of that framework; finally, we give a new characterization of cointegration, which obviates the need for the representation theorem above.

* This is a preliminary version and is not to be quoted, except by permission of the author. Comments, however, are welcome.

2 Cointegrated Sequences and their Properties

2.1 Restatement of EG Results

The problem as set forth in the paper by EG is roughly speaking as follows.¹

¹ As pointed out above, and not meaning any derogation of the seminal nature of that paper, Engle and Granger (1987) contains several misstatements, errors and incongruities that ought not to exist in so widely quoted a source. It is our purpose to provide a restatement of the EG results that do not suffer from what we had noted above. For example, the definition of integrated processes, in EG p. 252, is too restrictive. If, in the scalar case, $(I - L)^d x_t = \sum_{j=0}^{\infty} \alpha_j \epsilon_{t-j}$, such that $\sum_{j=0}^{\infty} |\alpha_j| < \infty$, the right hand side represents a regular (purely nondeterministic) process, a formulation found very frequently in applications. On the other hand, it is not always possible for such processes to have a rational representation. This would require the α_j above, to contain only $n + m$ free parameters, where n is the order of the autoregression, and m the order of the moving average. Also the terminology “..with no deterministic component ..” in the definition is confusing. In fact, the representation in Eq. (3.1) of EG p. 255, has nothing to do with the Wold decomposition; it has to do with the requirement that the spectral density of $(I - L)x_t$, in the example above, be positive a.e., and have a unilateral Fourier series representation. (see Proposition 3 of Ch. 1, in Dhrymes, (1993) mimeo.)

The statement in EG Eq. (3.3) p. 255, is in error; it should be $A(L)(I - L)x_t' = d(L)\epsilon_t'$. Moreover, the statement “.. $d(I)$ is finite ..”, is correct but incomplete. In fact $d(I) = 0$.

Lemma 1 is totally unnecessary.

On p. 258 the statement “Since $C(B)$ has full rank and equals I_N at $B = 0$, its *inverse* (italics added) [at $B = 0$] is $A(0)$ which is also I_N ,” is in error; $C(L)$ cannot be of full rank, and does not have an inverse. This is obvious since $|C(L)| = d(L)$ and moreover, $|C(I)| = 0$ by the cointegration assumption. For example, take the simple case

$$C(L) = [B(L)]^{-1}A(L), \quad B(L) = I_q - B_1L, \quad A(L) = I_q - A_1L.$$

For $C(L)$ to be well defined, we require the characteristic roots of B_1 , which are the inverse of the roots of $|I_q - B_1z| = 0$, to be less than one in modulus; for $C(L)$ to be invertible, we further require the characteristic roots of A_1 , which are the inverse of the roots of $|I_q - A_1z| = 0$, to be less than one in modulus as well. Notice that, in either case, **no characteristic root of A_1 , or B_1 is zero or unity.** Thus, there exist nonsingular matrices Q_i , $i = 1, 2$, such that

$$A_1 = Q_1\Lambda_1Q_1^{-1}, \quad B_1 = Q_2\Lambda_2Q_2^{-1}, \quad C(I) = Q_2[I_q - \Lambda]^{-1}Q_2^{-1}Q_1[I_q - \Lambda_1]Q_1^{-1},$$

and $C(I)$ cannot be of rank $q - r$, $r > 0$, as required by the cointegration assumption. What this shows is that for a series operator, $C(L)$, containing a finite number of free parameters, we cannot, in general, have both cointegration and invertibility. Whether this is possible in the limit, i.e. when $C(L)$ has (countably) infinitely many free parameter needs to be demonstrated, which is not, by the authors.

Let $X = \{X_t : t \in \mathcal{N}\}$ be a stochastic sequence defined on the probability space $(\Omega, \mathcal{A}, \mathcal{P})$, where X_t is a q -element **row vector** and \mathcal{N} is the integer lattice on R , and let

$$(I - L)X'_t = C(L)\epsilon'_t, \quad C(L) = \sum_{j=0}^{\infty} C_j \epsilon'_{t-j}, \quad (1)$$

where $\epsilon = \{\epsilon_t : t \in \mathcal{N}\}$ a $MWN(\Sigma)$, (multivariate white noise process with covariance matrix Σ). We remind the reader that by convention $C_0 = I_q$. For the right member of Eq. (1) to have meaning, we require

$$\sum_{j=0}^{\infty} \|C_j\| < \infty, \quad (2)$$

in which case the right member converges absolutely with probability one. In this context, we probe the question of what are the implications of asserting that $X \sim CI(1, 1, r)$, $r < q$. We have

Theorem 1 (Engle and Granger). Let $X = \{X'_t : t \in \mathcal{N}\}$ be a stochastic sequence defined on the probability space $(\Omega, \mathcal{A}, \mathcal{P})$, and suppose it is of the form²

$$(I - L)X'_t = \sum_{j=0}^{\infty} C_j \epsilon'_{t-j}, \quad C_0 = I_q, \quad \sum_{j=0}^{\infty} \|C_j\| < \infty. \quad (3)$$

Moreover, let $X \sim CI(1, 1, r)$, and β be a $q \times r$ matrix containing the r linearly independent cointegrating vectors; the following statements are true:

- i. there exists a representation

$$C(L) = C(I) + (I - L)C^*(L);$$

- ii. $\text{rank}[C(I)] = q - r$, $q > r$;

- iii. X has a representation as $MARMA(\infty, \infty)$; specifically let $d(L) = |C(L)|$, and $H(L)$ be the adjoint of³ $C(L) = [c_{rs}(L)]$, $c_{rs}(L) = \sum_{j=0}^{\infty} c_{rs}^{(j)} L^j$, where $c_{rs}^{(j)}$ is the r, s element of the matrix C_j . Then

$$H(L)(I - L)X'_t = d(L)I_q \epsilon'_t, \quad d(I) = 0;$$

² The only restriction on the generality of the results here, is that the stationarity of the differenced process $(I - L)X_t$, is rendered in the form of a **general linear process**, or $MMA(\infty)$. This property may be ensured by asserting that the spectral densities of the components of X_t are strictly positive and have a **unilateral Fourier series representation**.

³ An easily accessible, and not very technical treatment of polynomial lag operators in a scalar and matrix form may be found in the author's Dhrymes (1970) chapter 12; Dhrymes (1971), chapter 2; Dhrymes (1978), ssecond ed. (1984), chapter 5.

iv. the operator $C^*(L)$ is invertible;

v. there exist matrices β , Γ of dimension $q \times r$ and rank r such that

$$C(I)'\beta = 0, \quad C(I)\Gamma = 0, \quad H(I) = \Gamma\beta',$$

and moreover, $H(0) = I_q$;

vi. there exists an error correction representation,

$$A(L)(I - L)X'_t = -\Gamma Z'_{t-1} + b(L)\epsilon'_t, \quad (4)$$

Proof: By long division (of the type one learns in elementary school), with $C_{(n)}(L) = \sum_{j=0}^n C_{n-j}L^{n-j}$ as the **dividend**, and $L - I$ as the **divisor**, we obtain

$$C_{(n)}(I) = \sum_{j=0}^n C_j, \text{ as the } \mathbf{remainder},$$

$$-C_{(n)}^*(L) = \sum_{j=0}^{n-1} \left(\sum_{i=0}^{n-1-j} C_{i+j+1} \right) L^j, \text{ as the } \mathbf{quotient}, \text{ so that}$$

$$C_{(n)}(L) = C_{(n)}(I) + (I - L)C_{(n)}^*(L).$$

In the preceding, $C_{(n)}^*(L)$ must be **invertible** for otherwise we can repeat the process, thus obtaining cointegration of order higher than one. By a limiting process, i.e. letting $n \rightarrow \infty$, part i. is proved with

$$C(I) = \sum_{j=0}^{\infty} C_j, \quad C^*(L) = -\sum_{j=0}^{\infty} \left(\sum_{i=0}^{\infty} C_{i+j+1} \right) L^j.$$

To prove ii., consider the representation

$$(I - L)X'_t = C(I)\epsilon'_t + (I - L)C^*(L)\epsilon'_t. \quad (5)$$

premultiply by β' , and define $X_t\beta = Z_t$. By the cointegrating assumption, Z_t is an r -element row vector whose elements are jointly **stationary**. Thus, in

$$(I - L)Z'_t = \beta' C(I)\epsilon'_t + (I - L)\beta' C^*(L)\epsilon'_t, \quad (6)$$

⁴ The statement of this in EG Eq. (3.4) p. 256, is mistaken; the error process cannot possibly be the same as in EG Eq. (3.3) p. 255, which itself contains a misprint by omitting the factor $(1 - B)$ in the notation of that paper.

we conclude that $\beta' C(I) = 0$, and that $\beta' C^*(L)\epsilon'_t$ represents a **stationary process**; moreover since $\text{rank}(\beta) = r$, we conclude that $\text{rank}[C(I)] = q - r$, completing the proof of ii.

To prove iii. we note that $H(L)$ has elements which are determinants of $q - 1$ -dimensioned submatrices of $C(L)$; since

$$\sum_{j=0}^{\infty} \|C_j\| < \infty, \quad \text{and} \quad \|C_s\| \|C_r\| \leq (1/2)[\|C_s\|^2 + \|C_r\|^2]$$

we conclude that the elements of $H(L)$ are well defined. Since $H(L)$ is the **adjoint** of $C(L)$, we have $H(L)C(L) = d(L)I_q$. Premultiplying Eq. (4.11) by $H(L)$ we find

$$H(L)(I - L)X'_t = d(L)I_q\epsilon_t = d(L)\epsilon'_t. \quad (7)$$

which is seen to represent X as a *MARMA*(∞, ∞) process, and thus concludes the proof of iii.

The proof of iv., follows immediately from the argument above, but we may amplify as follows. Note that, if we employ the long division of part i. **with the divisor** $(I - L)^2$, we should find

$$C(L) = R(L) + (I - L)^2 C^{**}(L),$$

where $R(L) = R_0 + R_1 L$. What the reader may not realize is that we can attain the same result by **dividing** $C^*(L)$, as defined in i., again by $(I - L)$, thus obtaining

$$C(L) = C(I) + (I - L)C^*(I) + (I - L)^2 C^{**}(L).$$

Comparing with the previous result we have

$$R_0 = C(I) + C^*(I), \quad R_1 = -C^*(I).$$

If $X \sim CI(2, 2, r)$, $\beta' C^*(I) = 0$, as well as $\beta' C(I) = 0$. Since, **in fact**, $X \sim I(1)$, we must conclude that $C^*(I)$ is of **full rank**, which implies that the inverse of $C^*(L)$, $[C^*(L)]^{-1} = H^*(L)/d^*(L)$, is well defined.

To prove v. we note that, by the arguments in the proof of iii. and iv., the dimension of the row, as well as column, null space of $C(I)$ is r ; by the cointegration assumption β spans (is a basis for) the row null space. Moreover, there must exist r linearly independent vectors in the **column null space**, which thus, span (form a basis for) that space. Let these vectors be denoted by the $q \times r$ matrix Γ . Finally, $C(I)H(I) = d(I) = 0$. and we see that $H(I)$ is **in the column null space** of $C(I)$; moreover $C(L)H(L) = H(L)C(L)$, so that we also have

$H(I)C(I) = 0$ and thus, $H(I)$ is in the **row null space** of $C(I)$ as well. We conclude therefore that $H(I) = \Gamma\beta'$. In view of the standard conventions regarding normalization, $d(0) = 1$ and consequently, from $C(L)H(L) = d(L)I_q$, we obtain $C(0)H(0) = d(0)I_q$; since by convention $C(0) = I_q$ we conclude $H(0) = I_q$, completing the proof of v.

To prove vi. we begin with the result of part iii., so that we have $H(L)(I-L)X'_t = d(L)\epsilon'_t$. Adding $H(I)(I-L)X'_t$ to both sides we find

$$\begin{aligned} H(L) + H(I) &= A(L), & b(L) &= d(L)I_q + \Gamma\beta' C^*(L), \\ A(L)(I-L)X_t &= -\Gamma Z'_{t-1} + b(L)\epsilon'_t. \end{aligned} \quad (8)$$

q.e.d.

2.2 Empirical Implementation

EG essentially recommend an implementation of the preceding through the covariance matrix

$$M_T = \frac{1}{T^2} X' X, \quad X = (X_t), \quad t = 1, 2, \dots, T. \quad (9)$$

Specifically, they recommend that the cointegrating vectors be determined as the characteristic vectors of M_T (ultimately) corresponding to the r zero (?) roots of that matrix. To probe into these issues let us examine more closely what we are dealing with. To this effect put

$$\eta'_t = \sum_{j=0}^{\infty} C_j \epsilon'_{t-j}, \quad (10)$$

note that η is a zero mean covariance stationary process with

$$\psi(0) = E\eta'_t \eta_t = \sum_{j=0}^{\infty} C_j \Sigma C'_j \quad (11)$$

$$\psi(\tau) = E\eta'_{t+\tau} \eta_t = \sum_{j=0}^{\infty} C_{j+|\tau|} \Sigma C'_j, \quad \psi(-\tau) = \psi'(\tau),$$

and consequently,

$$\text{Cov}(X'_t) = \sum_{\tau=-(t-1)}^{(t-1)} (t-|\tau|)\psi(\tau), \quad (12)$$

since on the assumption $X_0 = 0$, $X_t = \sum_{j=1}^t \eta_j$. Thus,

$$M_T = \frac{1}{T^2} \sum_{t=1}^T \sum_{\tau=-(t-1)}^{(t-1)} (t - |\tau|) \psi(\tau) \quad (13)$$

$$\begin{aligned} \|EM_T\| &\leq \frac{1}{T^2} \sum_{t=1}^T t \sum_{\tau=-(t-1)}^{t-1} \left(1 - \frac{|\tau|}{t}\right) \|\psi(\tau)\| \\ &\leq \frac{T(T+1)}{2T^2} \sum_{\tau=-\infty}^{\infty} \|\psi(\tau)\|, \end{aligned}$$

and EM_T/T^2 converges with T , provided we can show that

$$\sum_{\tau=-\infty}^{\infty} \|\psi(\tau)\| < \infty.$$

But this is easily established from Eq. (3), since

$$\sum_{\tau=-\infty}^{\infty} \|\psi(\tau)\| \leq \left(\sum_{j=0}^{\infty} \|C_j\| \right) \Sigma \left(\sum_{j=0}^{\infty} \|C_j\| \right) < \infty. \quad (14)$$

Thus

$$\lim_{T \rightarrow \infty} EM_T = \psi_{\infty}^* = \sum_{\tau=-\infty}^{\infty} \psi(\tau) = \psi(0) + \sum_{\tau=1}^{\infty} \sum_{s=0}^{\infty} (C_{s=|\tau|} \Sigma C'_s + C_s \Sigma C_{s+|\tau|}).$$

Remark 1. What is established by the preceding discussion is that M_T is not a matrix that has fixed expectation, in the sense that in the stationary case $(1/T) \sum_{t=1}^T EX'_t X_t = \Phi$. Thus, if the cointegrating matrix β is taken to be a subset of the characteristic vectors of M_T it cannot be taken to be the estimator of a fixed matrix, in the same way as in the stationary case. Moreover, it is not sufficient that $\lim_{T \rightarrow \infty} M_T = \psi_{\infty}^*$, one must also show, at least, that $M_T \xrightarrow{P \text{ or a.c.}} \psi_{\infty}^*$ in order for the procedure to, possibly, yield estimators which have as their limit some submatrix of the matrix of characteristic vectors of ψ_{∞}^* .

We now ask the question: Using the definition of cointegration alone, can we determine what properties a matrix β must have in order that

$$\text{Cov}(X'_t) = \sum_{\tau=-(t-1)}^{t-1} (t - |\tau|) \psi(\tau), \quad \text{Cov}(Z'_t) = \Phi, \quad \forall t, \quad Z_t = X_t \beta.$$

Using the results of Eqs. (11) and (12), we find that

$$\text{Cov}(X'_t) = \psi(0) + \sum_{\tau=1}^{t-1} (t - \tau) \left(\sum_{s=0}^{\infty} C_{s+\tau} \Sigma C'_s + C_s \Sigma C'_{s+\tau} \right). \quad (15)$$

Consequently, by the definition of Z_t , we have

$$\begin{aligned} \text{Cov}(Z'_t) &= \beta' \text{Cov}(X'_t) \beta = t \beta' \psi(0) \beta \\ &+ \sum_{s=1}^{t-1} (t-s) \left(\sum_{j=0}^{\infty} \beta' C_{s+j} \Sigma C'_j \beta + \beta' C_s \Sigma C'_{s+j} \beta \right). \end{aligned} \quad (16)$$

If we require $\beta' \psi(0) \beta = 0$, we must have $\text{Cov}(Z'_t) = 0$, since $\|\psi(\tau)\| \leq \|\psi(0)\|$. Thus, this approach does not yield any constructive information on the relation of β to the constituent elements of the problem, and casts some doubt on the usefulness of the suggestion of EG that β be obtained as a submatrix of the characteristic vectors of M_T .

Let us now look at the problem in the context of

$$(I - L)X'_t = [C(I) + F(L)]\epsilon'_t, \quad F(L) = \sum_{j=0}^{\infty} F_j L^j = (I - L)C^*(L).$$

Since

$$X_t = X_{t-1} + \eta_t,$$

we obtain the recursive relation

$$\text{Cov}(X'_t) = \text{Cov}(X'_{t-1}) + \sum_{\tau=-(t-1)}^{t-1} \psi(\tau), \quad (17)$$

whence we again obtain

$$\text{Cov}(X'_t) = \sum_{\tau=-(t-1)}^{t-1} (t - |\tau|) \psi(\tau).$$

Further, from

$$\psi(\tau) = E \eta'_{t+\tau} \eta_t.$$

we find

$$\begin{aligned} \psi(\tau) &= \delta_0(\tau) C(I) \Sigma C(I)' + C(I) \sum_{j=0}^{\infty} (E \epsilon'_{t+\tau} \epsilon_{t-j}) F'_j \\ &+ \sum_{j=0}^{\infty} F_j (E \epsilon'_{t+\tau-j} \epsilon_t) C(I)' + \sum_{j=0}^{\infty} \sum_{s=0}^{\infty} F_j (E \epsilon'_{t+\tau-j} \epsilon_{t-s}) F'_s \end{aligned}$$

$$\psi(0) = C(I) \Sigma C(I)' + C(I) \Sigma F'_0 + F_0 \Sigma C(I)' + \sum_{j=0}^{\infty} F_j \Sigma F'_j$$

$$\psi(\tau) = F_{|\tau|} \Sigma C(I)' + \sum_{j=0}^{\infty} F_{j+|\tau|} \Sigma F'_j, \quad \text{for } \tau > 0,$$

$$\psi(\tau) = C(I) \Sigma F_{|\tau|} F'_0 + \sum_{j=0}^{\infty} F'_{j+|\tau|} \Sigma F'_j, \quad \text{for } \tau < 0.$$

Bearing in mind that $F_0 = I_q - C(I)$, $F_j = C_j$, $j \geq 1$, we see that utilizing the decomposition $C(L) = C(I) + F(L)$ does not shed any light on the matter since we still obtain

$$\text{Cov}(Z'_t) = t\beta' \psi(0)\beta + \sum_{\tau=1}^{t-1} (t - |\tau|) \left[\sum_{j=0}^{\infty} (\beta' C_{j+|\tau|} \Sigma C'_j \beta + \beta' C_j \Sigma C'_{j+|\tau|} \beta) \right],$$

and the manner in which the cointegrating matrix forces stationarity, i.e. independence of the right member above from t , still eludes us!

3 An Alternative Definition of Cointegration

Definition 2. Let X be a zero mean square integrable stochastic sequence defined on the probability space $(\Omega, \mathcal{A}, \mathcal{P})$; suppose $(I - L)^d X'_t$ is covariance stationary, and X is representable as a $MAR(\infty)$ process, i.e.

$$D(L)X_t = \epsilon'_t, \quad (18)$$

where $\epsilon = \{\epsilon'_t : t \in \mathcal{N}\}$ is $MWN(\Sigma)$. The sequence X is said to be cointegrated of order b and cointegrating rank r if and only if there exists a matrix β of maximal rank r such that $X_t \beta$ is $I(d - b)$.

Remark 2. In the particular case, $d = b = 1$, which is found extensively in applications, the definition merely states that the operator $D(L)$ “almost has a unit root”. If it has a unit root then we can write $D(L) = (I - L)\bar{D}(L)$, and $\bar{D}(L)$ is **invertible**, in which case we may write

$$(I - L)X'_t = [\bar{D}(L)]^{-1} \epsilon'_t. \quad (19)$$

The operator $\bar{D}(L)$ **must be invertible** for if not, the sequence is integrated of a higher order than one.⁵ Eq. (19) looks very much like Eq. (3.1) in EG, p. 255, but of course it is not; moreover, it makes quite clear why the statement in EG p. 258 that “... $C(B)$ has full rank .. and its inverse...” is in error, since in that case there is no possibility of cointegration.

Thus, $D(L)$ must be of the form

$$D(L) = D_* + (I - L)\bar{D}(L); \quad (20)$$

⁵ Note that in all such discussions the existence of roots less than unity in modulus is ruled out.

however, D_* cannot be of full rank since, if it were we would have the representation

$$D_*X'_t = \epsilon'_t - \bar{D}(L)[(I - L)X'_t] \quad (21)$$

which entails a **contradiction** since a nonsingular transformation of an $I(1)$ process is also $I(1)$, but the right member is clearly stationary! Thus, D_* must be singular and its rank, r , determines the **cointegrating rank**.

This definition of cointegration also disposes of the problem raised earlier, viz. how to convince ourselves that Z_t has a time-invariant covariance matrix when the only knowledge we have is of the covariance matrix of X_t .

4 Empirical Implementation

From Eq. (19) we easily determine

$$D(I) = D_* + (I - I)\bar{D}(I) = D_*, \quad D(0) = D_* + \bar{D}(0), \quad \text{or} \quad \bar{D}_0 = I_q - D(I). \quad (22)$$

Since the matrix $D(I)$ is of order q and rank r , by the singular value decomposition theorem, see Dhrymes (1984) p. 78, there exist matrices Γ , β , of dimension $q \times r$ and rank r such that

$$D(I) = \Gamma\beta'. \quad (23)$$

Now, combining the results above we have

$$D(I)X'_t + \bar{D}_0(I - L)X'_t = \epsilon'_t - \sum_{j=1}^{\infty} \bar{D}_j \Delta X_{t-j}.$$

Simplifying the left member we obtain

$$\Delta X'_t - \Gamma\beta'X'_{t-1} = \epsilon'_t - \sum_{j=1}^{\infty} \bar{D}_j \Delta X_{t-j}, \quad (24)$$

which produces a relationship from which the cointegrating vectors may be obtained, see for example Johansen (1988).

Finally, we note that a representation theorem, like the one given in EG, is superfluous, since all such relationships are transparent in this context.

References

- Dhrymes, P. J. (1970), *Econometrics: Statistical Foundations and Applications*, New York: Harper and Row, corrected edition (1974) New York: Springer Verlag.
- (1971), *Distributed Lags: Problems of Estimation and Formulation* (1971), San Francisco: Holden-Day; second edition, Amsterdam: North Holland (1981).
- (1978), *Mathematics for Econometrics*, New York: Springer Verlag; second edition (1984). ————— (1993), “Topics in Time Series”, mimeo.
- Engle, R. F. and C.W.J. Granger (1987), “Cointegration and Error Correction: Representation, Estimation and Testing”, *Econometrica*, vol. 55, pp. 251-276.
- Johansen, S. (1988), “Statistical Analysis of Cointegration Vectors”, *Journal of Economic Dynamics and Control*, vol. 12, pp. 231-254.