INTRODUCING NON-LINEARITY INTO COINTEGRATION*

Clive W. J. Granger**

Resumo

Uma classe abrangente de processos é estudada, na qual estes são compostos por um par de séries observáveis, que por sua vez são compostos por uma ponderação de duas séries não observáveis - uma persistente e outra não-persistente. Desse arcabouço, segue que as séries observáveis são persistentes, porém têm uma combinação linear que não o é, portanto são cointegradas. Por construção, as séries consideradas não são necessariamente $I(1)$ e outros aspectos lineares da teoria usual podem ser descartados, apesar de existirem modelos de correção-de-erro na maioria dos casos. Discute-se alguns exemplos teóricos posteriormente.

Abstract

A wide class of processes is considered in which initially a persistent series and a transient series are generated separately, as unobserved components, and from them a pair of observable series are generated as a weighted combination of the components. It follows that the observable series will be persistent but have a linear combination that is transient, and so can be considered cointegrated. From the construction, the series need not be $I(1)$ and other linear aspects of standard theory is removable, but often error-correction models are available. Some theoretical examples are provided.

Palavras-Chave: Persistent, transient components, cointegration, non-linear models

Código JEL: C22, C51

* Invited Lecture to the Latin American Meeting of the Econometric Society, Rio de Janeiro, July 1996. This work was supported by NSF Grant SES 93-08295.

** Department of Economics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0508, USA.

1. Introduction.

The usual form of cointegration is very linear in form and yet much of economic modeling is becoming increasingly non-linear. This paper proposes a structure within which a variety of non-linear forms of models can be generated which can be considered to be cointegrated, and which may have a discoverable error-correction mechanism. The structure proposed is not completely general but is broad enough to provide a wide variety of possible forms to capture many interesting properties of actual data.

The standard form of cointegration, in the bivariate case, considers a pair of $I(1)$ series X_t , Y_t so that ΔX_t , ΔY_t are both $I(0)$ or stationary, but there exists a linear combination $Z_t = X_t - AY_t$ which is $I(0)$. It is known that cointegration will occur if and only if there is a decomposition

$$\left. \begin{aligned} X_t &= AW_t + \tilde{X}_t \\ Y_t &= W_t + \tilde{Y}_t \end{aligned} \right\} \quad (1.1)$$

where W_t is $I(1)$ and $\tilde{X}_t$, $\tilde{Y}_t$ are both $I(0)$. W_t is known as the "common stochastic trend" or "common $I(1)$ factor." There will be a corresponding error-correction model, which in vector form is:

$$\Delta \underline{X}_t = \underline{\gamma} Z_{t-1} + \text{lags } \Delta X_t, \Delta Y_t + \underline{\varepsilon}_t \quad (1.2)$$

where

$$\underline{X}_t = \begin{pmatrix} X_t \\ Y_t \end{pmatrix}, \quad \underline{\gamma} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}$$

and $|\gamma_1| + |\gamma_2| \neq 0$. An estimate of the common stochastic trend with useful interpretation and properties is

$$\widehat{\underline{W}}_t = \underline{\gamma}'_{\perp} \underline{X}_t \quad (1.3)$$

where $\underline{\gamma}_{\perp}$ is a vector orthogonal to $\underline{\gamma}$ such that $\underline{\gamma}'_{\perp} \underline{\gamma} = \underline{0}$.

This estimate is discussed by Gonzalo and Granger (1995) and general accounts of cointegration can be found in Engle and Granger (1991), Banerjee et. al. (1993), Watson (1994), Johansen (1995) and Hatanaka (1995), for example. The concept of cointegration seems to have been a useful one and has been widely applied. However, there also seems to be a wide belief that the economy contains some important non-linear aspects. In this paper the question of how non-linearity can be introduced into cointegration is discussed, expanding on Granger and Swanson (1996).

There have already been a number of papers that introduced some aspects of non-linearity into cointegration. For example:

- (a) Granger and Hallman (1991) keep the $I(1)$ and $I(0)$ classification and consider series X_t, Y_t that are $I(1)$, not linearly cointegrated but there exists functions F_1, F_2 such that $F_1(X_t), F_2(Y_t)$ are both $I(1)$ and

$$Z_t = F_1(X_t) - F_2(Y_t)$$

is $I(0)$. The functions are chosen non-parametrically.

- (b) Corradi, Swanson, and White (1995) consider the construction

$$\left. \begin{aligned} \Delta X_t &= \gamma_1 Z_{t-1} + \phi_1 g(Z_{t-1}) + \varepsilon_{x,t} \\ \Delta Y_t &= \gamma_2 Z_{t-1} + \phi_2 g(Z_{t-1}) + \varepsilon_{y,t} \\ Z_t &= X_t - AY_t \quad \varepsilon' \text{'s i.i.d} \end{aligned} \right\} \quad (1.5)$$

so that

$$Z_t = \theta Z_{t-1} + \phi g(Z_{t-1}) + \eta_t \quad (1.6)$$

$\theta = 1 + \gamma_1 - A\gamma_2$ is assumed to obey $|\theta| < 1$ and $g(Z)$ is taken to be bounded. With these constraints it follows that X_t, Y_t are not ergodic (in fact are $I(1)$) but $\Delta X_t, \Delta Y_t, Z_t$ are all ergodic (actually $I(0)$).

This is an example of processes having cointegration but with a non-linear error correction model. An early example discussion of such a process was given by Escribano (1987).

- (c) Granger (1995) defines a general class of processes which have “extended memory in mean” if

$$E [X_{t+h} | I_t] \not\rightarrow \text{constant as } h \uparrow$$

where I_t is an extensive memory set available at time t . A pair of such series may be called cointegrated if $Z_t = X_t - AY_t$ is short memory in mean.

Thus X_t, Y_t are not necessarily $I(1)$, so this linear property is replaced by the non-linear idea of extended memory in mean, but cointegration is still defined in terms of a linear combination.

- (d) Escribano and Mira (1996) attempt a complete generalization. They define X_t as being $NI(1)$ if it is α -mixing but $\sum_{j=1}^t X_{t-j}$ is not α -mixing. Then, if X_t, Y_t are both $NI(1)$ and there exists a function $g(X_t, Y_t, \theta)$ such that this function is α -mixing for most θ , but not for $\theta = \theta^*$, then X_t, Y_t are called non-linear cointegrated.

Many of these definitions are analytically difficult to handle and test and estimation may be very difficult. In the following section, a template is presented whereby a wide variety of different types of processes can be considered which are linearly cointegrated. Essentially a “build-it-yourself” kit is presented where non-linearity can be inserted into cointegration.

2. Generating Cointegration.

Suppose that initially one generates two “unobserved components” having distinctly different temporal properties, W_t is “persistent” or “has long memory” however one cares to define such qualities, and Z_t is “transitory” or “short-memory”. Thus, W_t need not be $I(1)$ and Z_t need to be $I(0)$, however $I(1), I(0)$ are formally defined.

Now suppose that a pair of observable series X_t, Y_t are defined from W_t, Z_t by

$$X_t = AW_t + C_2Z_t \quad (2.1)$$

$$Y_t = W_t - C_1Z_t \quad (2.2)$$

with the constraints $C_1A + C_2 = 1, A \neq 0$. Thus, solving the equations gives

$$\left. \begin{array}{l} Z_t = X_t - AY_t \text{ is transitory} \\ W_t = C_1X_t + C_2Y_t \text{ is persistent} \end{array} \right\}. \quad (2.3)$$

As the sum of a persistent and a transitory series will be persistent (presumably), it follows that X_t, Y_t will both be persistent, even though there is a linear combination of them, Z_t , that is transitory so they can be given a name such as "cointegrated".

As a simple example, let Z_t, W_t be generated by

$$Z_t = \lambda Z_{t-1} + e_{zt} \quad |\lambda| < 1 \quad (2.4)$$

$$W_t = \phi W_{t-1} + e_{wt} \quad |\phi| \geq 1 \quad (2.5)$$

with e_{zt}, e_{wt} each white noises, so that Z_t is stationary (i.e. $I(0)$) but W_t is not necessarily $I(1)$ as it may be explosive, with $(1 - \phi B)W_t$ being $I(0)$. X_t, Y_t will be given by (2.1), (2.2).

Applying the filter $(1 - \phi B)$ to (2.1) gives

$$\begin{aligned} (1 - \phi B)X_t &= C_2(1 - \phi B)Z_t + A(1 - \phi B)W_t \\ &= C_2(\lambda - \phi)Z_{t-1} + Ae_{wt} + C_2e_{zt} \end{aligned} \quad (2.6)$$

using (2.4) and (2.5). This can be written as

$$(1 - \phi B)X_t = \gamma_2 Z_{t-1} + e_{xt} \quad (2.7)$$

and similarly (2.2) leads to

$$(1 - \phi B) Y_t = \gamma_2 Z_{t-1} + e_{xt} \quad (2.8)$$

where $\gamma_1 = C_2(\lambda - \phi)$ and $\gamma_2 = -C_1(\lambda - \phi)$. Equations (2.7) and (2.8) thus provide the generalized error-correction model, using $(1 - \phi B)$ as the filter that takes X_t to $I(0)$.

A more generalized example has time varying parameters, so that

$$Z_t = \lambda_t Z_{t-1} + e_{zt} \quad |\lambda_t| < 1 \quad (2.9)$$

$$\begin{aligned} W_t &= \phi_t W_{t-1} + e_{wt} \quad |\phi_t| \geq 1 \\ \text{or } E[\phi_t] &= 1 \end{aligned} \quad (2.10)$$

replace (2.4) and (2.5). Z_t is still transitory and W_t persistent, with λ_t, ϕ_t being time varying parameters, either stochastic perhaps, depending on some other variable, such as an indicator of the business cycle, or deterministic. Note the TVP filter $(1 - \phi_t B)$ makes W_t equal to the white noise e_{wt} and thus $I(0)$. Let X_t, Y_t be generated by

$$X_t = C_{2t} Z_t + A_t W_t \quad (2.11)$$

$$Y_t = -C_1 Z_t + W_t \quad (2.12)$$

with constraints $A_t > 0, C_1 A_t + C_{2t} = 1$. Applying filter $(1 - \phi_t B)$ to (2.11) and (2.12) and using (2.9) and (2.10) produces the error-correction model

$$\left. \begin{aligned} (1 - \phi_t B) X_t &= \gamma_{1t} Z_{t-1} + \Delta A_t \cdot \phi_t W_{t-1} + \text{white noise} \\ (1 - \phi_t B) Y_t &= \gamma_{2t} Z_{t-1} + \text{white noise} \end{aligned} \right\} \quad (2.13)$$

where

$$\left. \begin{aligned} \gamma_{1t} &= C_{2t} \lambda_t - \phi_t C_{2t-1} \\ \gamma_{2t} &= -C_1 (\lambda_t - \phi_t) \end{aligned} \right\}. \quad (2.14)$$

For (2.13) to be a true error-correction model it needs to be balanced, and for this ΔA_t needs to be $I(0)$ and independent of W_t , so that the term $\Delta A_t \cdot \phi_t W_{t-1}$, which is just $\Delta A_t \cdot W_t$, can be $I(0)$ in mean. [In general an $I(0) \times I(1)$, is $I(0)$, for example.] It is seen that a fairly major generalization has been achieved with little effort. X_t , Y_t are persistent, but $Z_t = X_t - A_t Y_t$ is transient and the common persistent term is found as $W_t = C_1 X_t + C_2 Y_t$. A form of error correction model is obtained in which if it is found that γ_{1t} , γ_{2t} do vary with time then from (2.14) it is seen that this time variation can arise from various places in the original specification.

W_t generated by (2.10) with ϕ_t a stochastic process having an expected value of one, has been called a stochastic unit root process by Granger and Swanson (1997). They take $\phi_t = \exp \alpha_t$, where α_t is an $I(0)$ process with moments such that $E[\phi_t] = 1$. Thus, on occasions W_t is stationary, other times it is nearly a random walk and occasionally is explosive. Such processes are difficult to distinguish from pure unit root processes using standard unit root tests, as simulations show and have variance that increases exponentially in time not the usual linear growth. They provide an example of a process that is not strictly $I(1)$, which is plausible to occur in an actual economy, which can be used to produce superior forecasts because of their flexibility and which can certainly be used as possible common stochastic trends.

The next section considers an alternative method of generating W_t .

3. Non-Linear Stochastic Trends.

A wide class of class of stochastic trends can be generated by

$$\Delta W_{t-1} = g(W_t) + \varepsilon_{t+1} \quad (3.1)$$

where $g(W) \geq 0$ which has been studied by Granger, Inoue, and Morin (1996) (henceforth GIM). If g is a positive constant, one gets

a random walk with drift, and so the process contains a linear deterministic trend if $g(W_t)$ declines in size and W_t increases, one gets a sub-linear stochastic trend, and so forth. There is an associated deterministic process

$$\Delta a_{t-1} = g(a_t) \quad (3.2)$$

with $W_0 = a_0$ and it is found that W_t/a_t tends to a constant or to a random variable but $W_t - a_t$ will be $I(1)$ or $I(2)$ plus a deterministic trend. GIM particularly consider the subclass

$$\Delta W_{t+1} = cW_t^\alpha + \gamma W_t^\beta \varepsilon_{t+1}$$

and provide simulation and examples for using actual data. If W_t is generated by (3.1) and Z_t by the stationary mechanism (1.6), with $(\theta < 1)$ and g bounded taking $\phi = 1$ for convenience, then with X_t , Y_t constructed from (2.1) and (2.2) one still gets (2.3). It follows that X_t , Y_t are persistent, as then contain the stochastic trend W_t but the linear combination Z_t is stationary.

The strict error-correction equations in the case are

$$\left. \begin{aligned} \Delta X_t - Ag(W_{t-1}) &= C_2 [(\theta - 1) Z_{t-1} + g(Z_{t-1})] + w.n. \\ \Delta Y_t - g(W_{t-1}) &= -C_1 [(\theta - 1) Z_{t-1} + g(Z_{t-1})] + w.n. \end{aligned} \right\} \quad (3.3)$$

the term $g(W_{t-1})$ can be replaced to a close approximation by $g(X_{t-1})$ and $g(Y_{t-1})$ in the two equations respectively. It should be noted that if (3.3) is used, the Gonzalo-Granger procedure for estimating the common trend W_t works perfectly, due to the way that the series are constructed.

4. Comments and Conclusions.

It may be noted that the equations (2.1) and (2.2) seem to be more constrained than the traditional decomposition (1.1) and so it can be asked whether the approach proposed in Section 2 is sufficiently general. To answer the question consider (2.1) and (2.2)

extended by the addition of “measurement noise”, which is taken to be $I(0)$ although not necessarily white noise, so that

$$\left. \begin{aligned} X_t &= AW_t + C_2Z_t + \varepsilon_{1t} \\ Y_t &= W_t - C_1Z_t + \varepsilon_{2t} \end{aligned} \right\} \quad (4.1)$$

with the constraint $C_2 + AC_1 = 1$ as usual. Now, decompose the noise vector by

$$\left. \begin{aligned} \varepsilon_{1t} &= A\eta_{1t} + C_2\eta_{2t} \\ \varepsilon_{2t} &= \eta_{1t} - C_1\eta_{2t} \end{aligned} \right\} \quad (4.2)$$

which, solving, gives

$$\left. \begin{aligned} \eta_{1t} &= C_1\varepsilon_{1t} + C_2\varepsilon_{2t} \\ \eta_{2t} &= \varepsilon_{1t} - A\varepsilon_{2t} \end{aligned} \right\} \quad (4.3)$$

The existence of this solution shows that the decomposition exists. One can now write (4.1) in the form

$$\begin{aligned} X_t &= A\overline{W}_t + C_2\overline{Z}_t \\ Y_t &= \overline{W}_t - C_1\overline{Z}_t \end{aligned}$$

where $\overline{W}_t = W_t + \eta_{1t}$, $\overline{Z}_t = Z_t + \eta_{2t}$. $\overline{W}_t$ will still be persistent and $\overline{Z}_t$ transitory and X_t , Y_t will be cointegrated. It is thus seen that the starting form for the analysis, equations (2.1) and (2.2) is general enough. It allows for the possibility of introducing wide classes of non-linearity into cointegration, the basic step of starting with $I(1)$ processes can be replaced with any form of persistence, and a form of cointegration maintained. The one generalization from linearity that is not attempted here is a non-linear cointegration relationship. If Z_t is required not only to be transient but also to have mean zero, it will be mean (i.e. zero) reverting. It follows that the pair of series

(X_t, Y_t) will have a form of equilibrium as with current cointegration, so that if there are no shocks to the economy from time T , then the process would converge from the point (X_T, Y_T) to some point on the line $Z = 0$. Note that for the system (2.4) and (2.12) this line is $X = A_t Y$ and so is not constant.

The model considered in this paper have all been bivariate, there is no difficulty in writing down the multivariate equivalents, although finding the properties of such models may be more difficult. An example of the mathematics involved is given in the first section of Granger and Swanson (1996).

It is hoped that the framework proposed here will provide researchers with a rich group of possible models, with development only limited by their own imaginations. Conceptually, if a few years time, a wide variety of non-linear models will be available and will have found useful applications.

Submetido em Setembro de 1996. Revisado em Outubro de 1996.

References

- Banerjee, A.; Dolado, J.; Galbraith, J.W. & Hendry, D.F. 1993. *Cointegration, Error-Correction and the Econometrics Analysis of Non-Stationary Data*, Oxford University Press.
- Corradi, V.; Swanson, N.R. & White, H. 1996. "Testing for Stationary-Ergodicity and for Comovements Between Nonlinear Discrete Time Markov Processes." Working Paper, Department of Economics, University of Pennsylvania.
- Engle, R.F. & Granger, C.W.J. 1991. *Long-run Economic Relationships: Readings in Cointegration*. Oxford University Press.
- Escribano, A. 1987. "Cointegration, Time Co-Trends and Error-Correction Systems: An Alternative Approach." CORE discussion paper.

- Escribano, A. & Mira, S. 1996. "Non-linear Cointegration and Non-linear Error-Correction Models." Working Paper, Department of Economics, Carlos III University of Madrid.
- Gonzalo, J. & Granger, C.W.J. 1995. "Estimation of Common Long Memory Components in Cointegrated Systems." *Journal of Business and Economic Statistics* **13**: 27-35.
- Granger, C.W.J. 1995. "Modeling Non-linear Relationships Between Extended-Memory Variables." *Econometrica* **63**: 265-279.
- Granger, C.W.J.; Inoue, T. & Morin, N. 1996. "Nonlinear Stochastic Trends." *Journal of Econometrics* to appear.
- Granger, C.W.J. & Hallman, J. 1991. "Long Memory Processes with Attractor." *Oxford Bulletin of Economics and Statistics* **53**: 11-26.
- Granger, C.W.J. & Swanson, N. 1996. "Further Developments in the Study of Cointegrated Variables." *Oxford Bulletin of Economics and Statistics* **58**: 537-553.
- Granger, C.W.J. & Swanson, N. 1997. "An Introduction to Stochastic Unit Root Processes." *Journal of Econometrics* to appear.
- Hatanaka, M. 1995. *Times Series Based Econometrics: Unit Roots and Cointegrations*. Oxford University Press.
- Johansen, S. 1995. *Likelihood Based Inference on Cointegration in the Vector Autoregressive Model*. Oxford University Press.
- Watson, M. 1994. "Unit Roots, Structural Breaks and Trends." In Engle, R.F. and McFadden, D.L. eds., Vol. IV. *Handbook of Econometrics*. Elsevier-North Holland.