

Article

The Geometric Meaning of the Notion of Joint Unpredictability of a Bivariate VAR(1) Stochastic Process

Umberto Triacca

Department of Computer Engineering, Computer Science and Mathematics, University of L'Aquila, Via Vetoio I-67010 Coppito, L'Aquila, I-67100, Italy; E-Mail: umberto.triacca@ec.univaq.it; Tel.: +39-086-2 43-4849, Fax.: +39-086-243-4407

Received: 21 August 2013; in revised form: 5 November 2013 / Accepted: 5 November 2013 / Published: 14 November 2013

Abstract: This paper investigates, in a particular parametric framework, the geometric meaning of joint unpredictability for a bivariate discrete process. In particular, the paper provides a characterization of the joint unpredictability in terms of distance between information sets in an Hilbert space.

Keywords: Hilbert spaces; predictability; stochastic process

JEL Classification: C18; C32

1. Introduction

Let (Ω, \mathcal{F}, P) be a probability space and $\{\mathbf{y}_t = [y_{1,t}, y_{2,t}]'; t = 0, 1, \dots\}$ a bivariate stochastic process defined on (Ω, \mathcal{F}, P) . We consider the differenced process $\{\Delta \mathbf{y}_t = [\Delta y_{1,t}, \Delta y_{2,t}]'; t = 1, \dots\}$, where Δ is the first-difference operator. Following Caporale and Pittis [1] and Hassapis *et al.* [2], we say that the process $\{\Delta \mathbf{y}_t; t = 1, \dots\}$ is jointly unpredictable if

$$E(\Delta \mathbf{y}_{t+1} | \sigma(\mathbf{y}_t, \dots, \mathbf{y}_0)) = \mathbf{0} \quad \forall t \quad (1)$$

where $\sigma(\mathbf{y}_t, \dots, \mathbf{y}_0)$ is the σ -field generated by past vectors \mathbf{y}_i $i = 0, \dots, t$.

The goal of this paper is to show that the notion of joint unpredictability, in a particular parametric framework, can be characterized by a geometric condition. This characterization is given in terms of distance between information sets in an Hilbert space. In particular, we will show that the process $\{\Delta \mathbf{y}_t; t = 1, \dots\}$ is jointly unpredictable if and only if the information contained in its past is 'much

distant' from the information contained in its future. Even if our result is not as general as might seem desirable, we think that the intuition gained from this characterization makes the notion of joint unpredictability more clear.

The rest of the paper is organized as follows. Section 2 presents the utilized mathematical framework. Sections 3 presents the geometric characterization. Section 4 concludes.

2. Preliminaries

Definitions, notation, and preliminary results from Hilbert space theory will be presented prior to establish the main result. An excellent overviews of the applications of Hilbert space methods to time series analysis can be found in Brockwell and Davis [3].

We use the following notations and symbols. Let (Ω, \mathcal{F}, P) be a probability space. We consider the Hilbert space $L^2(\Omega, \mathcal{F}, P)$ of all real square integrable random variables on (Ω, \mathcal{F}, P) . The inner product in $L^2(\Omega, \mathcal{F}, P)$ is defined by $\langle z, w \rangle = E(zw)$ for any $z, w \in L^2(\Omega, \mathcal{F}, P)$. The space $L^2(\Omega, \mathcal{F}, P)$ is a normed space and the norm is given by $\|w\| = [E(w^2)]^{1/2}$. The distance between $z, w \in L^2(\Omega, \mathcal{F}, P)$ is $d(z, w) = \|z - w\|$. A sequence $\{z_n\} \subset L^2(\Omega, \mathcal{F}, P)$ is said to converge to a limit point $z \in L^2(\Omega, \mathcal{F}, P)$ if $d(z_n, z) \rightarrow 0$ as $n \rightarrow \infty$. A point $z \in L^2(\Omega, \mathcal{F}, P)$ is a limit point of a set M (subset of $L^2(\Omega, \mathcal{F}, P)$) if it is a limit point of a sequence from M . In particular, M is said to be closed if it contains all its limit points. If S is a arbitrary subset of $L^2(\Omega, \mathcal{F}, P)$, then the set of all $\alpha_1 z_1 + \dots + \alpha_h z_h$ ($h = 1, 2, \dots$; $\alpha_1, \dots, \alpha_h$ arbitrary real numbers; z_1, \dots, z_h arbitrary elements of S) is called a linear manifold spanned by S and is symbolized by $sp(S)$. If we add to $sp(S)$ all its limit points we obtain a closed set that we call the closed linear manifold or subspace spanned by S , symbolized by $\overline{sp}(S)$. Two elements $z, w \in L^2$ are called orthogonal, and we write $z \perp w$, if $\langle z, w \rangle = 0$. If S is any subset of $L^2(\Omega, \mathcal{F}, P)$, then we write $x \perp S$ if $x \perp s$ for all $s \in S$; similarly, the notation $S \perp T$, for two subsets S and T of $L^2(\Omega, \mathcal{F}, P)$, indicates that all elements of S are orthogonal to all elements of T . For a given $z \in L^2(\Omega, \mathcal{F}, P)$ and a closed subspace M of $L^2(\Omega, \mathcal{F}, P)$, we define the orthogonal projection of z on M , denoted by $P(z|M)$, as the unique element of M such that $\|z - P(z|M)\| \leq \|z - w\|$ for any $w \in M$. We remember that if $z \perp M$, then $P(z|M) = 0$.

If M and N are two arbitrary subsets of $L^2(\Omega, \mathcal{F}, P)$, then the quantity

$$d(M, N) = \inf \{ \|m - n\| ; m \in M, n \in N \}$$

is called distance between M and N .

We close this section introducing some further definitions, concerning discrete stochastic processes in $L^2(\Omega, \mathcal{F}, P)$.

Let $\{x_t\}$ be a univariate stochastic process. We say that $\{x_t\}$ is integrated of order one (denoted $x_t \sim I(1)$) if the process $\{\Delta x_t = x_t - x_{t-1}\}$ is stationary whereas $\{x_t\}$ is not stationary. We say that the bivariate stochastic process $\{y_t = [y_{1,t}, y_{2,t}]'\}$ is integrated of order one if $y_{1,t} \sim I(1)$ and $y_{2,t} \sim I(1)$.

A stochastic process $\{y_t\}$ Granger causes another stochastic process $\{x_t\}$, with respect to a given information set I_t that contains at least $x_{t-j}, y_{t-j}, j > 0$, if x_t can be better predicted by using past values of y than by not doing so, all other information in I_t (including the past of x) being

used in either case. More formally, we say that $\{y_t\}$ is Granger causal for $\{x_t\}$ with respect to $H_{xy}(t) = \overline{sp} \{x_t, y_t, x_{t-1}, y_{t-1}, \dots\}$ if

$$\|x_{t+1} - P(x_{t+1}|H_{xy}(t))\|^2 < \|x_{t+1} - P(x_{t+1}|H_x(t))\|^2$$

where $H_x(t) = \overline{sp} \{x_t, x_{t-1}, \dots\}$.

Two stochastic processes, $\{x_t\}$, and $\{y_t\}$, both of which are individually $I(1)$, are said to be cointegrated if there exists a non-zero constant β such that $\{z_t = x_t - \beta y_t\}$ is a stationary ($I(0)$) process.

It is important to note that cointegration between two variables implies the existence of causality (in the Granger sense) between them in at least one direction (see Granger [4]).

3. A Geometric Characterization

In this section we assume that $\{\mathbf{y}_t = [y_{1,t}, y_{2,t}]'; t = 0, 1, \dots\}$ be a bivariate stochastic process defined on (Ω, \mathcal{F}, P) , integrated of order one, with $y_{1,0} = y_{2,0} = 0$, that has a VAR(1) representation

$$\mathbf{y}_t = \mathbf{A}\mathbf{y}_{t-1} + \mathbf{u}_t \tag{2}$$

where

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

is a fixed (2×2) coefficient matrix and $\mathbf{u}_t = [u_{1,t}, u_{2,t}]'$ is i.i.d. with $E(\mathbf{u}_t) = \mathbf{0}$ and $E(\mathbf{u}_t \mathbf{u}_t') = \Sigma = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$ for all t and $E(\mathbf{u}_t \mathbf{u}_s') = \mathbf{0}$ for $s \neq t$.

In this framework we have that $\{y_{2,t}\}$ does not Granger cause $\{y_{1,t}\}$ if and only if $a_{12} = 0$. Similarly, $\{y_{1,t}\}$ does not Granger cause $\{y_{2,t}\}$ if and only if $a_{21} = 0$.

We observe that the VAR residuals are usually correlated and hence the covariance matrix Σ is seldom a diagonal matrix. However, because the main aim of this study is pedagogical, we assume that Σ is diagonal for analytical convenience.

We consider the following information sets: $I_{\Delta y_1}(t+) = \{\Delta y_{1,t+1}, \Delta y_{1,t+2}, \dots\}$, $I_{\Delta y_2}(t+) = \{\Delta y_{2,t+1}, \Delta y_{2,t+2}, \dots\}$, $H_{\Delta y_1}(t) = \overline{sp} \{\Delta y_{1,t}, \Delta y_{1,t-1}, \dots\}$ and $H_{\Delta y_2}(t) = \overline{sp} \{\Delta y_{2,t}, \Delta y_{2,t-1}, \dots\}$.

Theorem 3.1. *Let \mathbf{y}_t be a VAR(1) process defined as in (2). The differenced process $\{\Delta \mathbf{y}_t; t = 1, \dots\}$ is jointly unpredictable if and only if*

$$d(I_{\Delta y_1}(t+), H_{\Delta y_2}(t)) = \sigma_{\Delta y_1} \text{ and } d(I_{\Delta y_2}(t+), H_{\Delta y_1}(t)) = \sigma_{\Delta y_2}$$

Theorem 1 provides a geometric characterization of the notion of joint unpredictability of a bivariate process in term of distance between information sets. It is important to note that

$$d(I_{\Delta y_1}(t+), H_{\Delta y_2}(t)) \leq \sigma_{\Delta y_1} \text{ and } d(I_{\Delta y_2}(t+), H_{\Delta y_1}(t)) \leq \sigma_{\Delta y_2}$$

Thus we have that the process $\{\Delta \mathbf{y}_t = [\Delta y_{1,t}, \Delta y_{2,t}]'; t = 1, \dots\}$ is jointly unpredictable if and only if the distances $d(I_{\Delta y_1}(t+), H_{\Delta y_2}(t))$ and $d(I_{\Delta y_2}(t+), H_{\Delta y_1}(t))$ achieve their maximum value, respectively.

It is intuitive to think that if these distances achieve their maximum value, then $\sigma(\mathbf{y}_t, \dots, \mathbf{y}_0)$ does not contain any valuable information about the future of the differenced series, $\Delta \mathbf{y}_t = [\Delta y_{1,t}, \Delta y_{2,t}]'$ and hence these are jointly unpredictable with respect to the information set $\sigma(\mathbf{y}_t, \dots, \mathbf{y}_0)$, that is $E(\Delta \mathbf{y}_{t+1} | \sigma(\mathbf{y}_t, \dots, \mathbf{y}_0)) = \mathbf{0}$.

We recall that Theorem 1 holds only in a bivariate setting.

3.1. Lemmas

In order to prove Theorem 1, we need the following lemmas.

Lemma 3.2. *Let V be a closed subspace of $L^2(\Omega, \mathcal{F}, P)$ and $G \neq \emptyset$ a subset of $L^2(\Omega, \mathcal{F}, P)$ such that $\|g\| = \eta \in R, \forall g \in G$. $G \perp V$ if and only if $d(G, V) = \eta$.*

Proof. Focker and Triacca ([5], p. 767).

Lemma 1 establishes a relationship between the orthogonality of sets/spaces in the Hilbert space $L^2(\Omega, \mathcal{F}, \mathbb{P})$ and their distance. We note that the orthogonality between G and V holds if and only if the distance $d(G, V)$ achieves the maximum value. In fact, $d(G, V)$ can not be greater than η since $0 \in V$.

Lemma 3.3. *The processes $\{y_{1,t}\}$ and $\{y_{2,t}\}$ are not cointegrated if and only if $\mathbf{A} = \mathbf{I}$.*

Proof. By (2) we have

$$\begin{bmatrix} \Delta y_{1,t} \\ \Delta y_{2,t} \end{bmatrix} = \begin{bmatrix} a_{11} - 1 & a_{12} \\ a_{21} & a_{22} - 1 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix}$$

These equations must be balanced, that is the order of integration of $(a_{11} - 1)y_{1,t-1} + a_{12}y_{2,t-1}$ and $a_{21}y_{1,t-1} + (a_{22} - 1)y_{2,t-1}$ must be zero.

(\Rightarrow) If $\mathbf{A} \neq \mathbf{I}$, since $(a_{11} - 1)y_{1,t-1} + a_{12}y_{2,t-1} \sim I(0)$ and $a_{21}y_{1,t-1} + (a_{22} - 1)y_{2,t-1} \sim I(0)$, we can have three cases.

Case (1) $\mathbf{A} = [a_{ij}]$, with $a_{ij} \neq 0 \ i, j = 1, 2, i \neq j$ and $a_{ii} \neq 1 \ i = 1, 2$.

Case (2)

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ 0 & 1 \end{bmatrix}$$

with $a_{11} \neq 1$ and $a_{21} \neq 0$.

Case (3)

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ a_{21} & a_{22} \end{bmatrix}$$

with $a_{21} \neq 0$ and $a_{22} \neq 1$.

In all three cases, there exists at least a not trivial linear combination of the processes $\{y_{1,t}\}$ and $\{y_{2,t}\}$ that is stationary. Thus we can conclude that $\{y_{1,t}\}$ and $\{y_{2,t}\}$ are cointegrated.

(\Leftarrow) If $\mathbf{A} = \mathbf{I}$, then $a_{12} = a_{21} = 0$ and so $\{y_{1,t}\}$ does not Granger cause $\{y_{2,t}\}$ and $\{y_{2,t}\}$ does not Granger cause $\{y_{1,t}\}$. It follows that $\{y_{1,t}\}$ and $\{y_{2,t}\}$ are not cointegrated.

Lemma 3.4. *If $\{y_{1,t}\}$ and $\{y_{2,t}\}$ are cointegrated, then $|a_{11} + a_{22} - 1| < 1$.*

Proof. We subtract $[y_{1,t-1}, y_{2,t-1}]'$ from both sides of Equation (2) by obtaining

$$\begin{bmatrix} \Delta y_{1,t} \\ \Delta y_{2,t} \end{bmatrix} = \begin{bmatrix} a_{11} - 1 & a_{12} \\ a_{21} & a_{22} - 1 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix}$$

If $\{y_{1,t}\}$ and $\{y_{2,t}\}$ are cointegrated, we have

$$\begin{aligned} \begin{bmatrix} \Delta y_{1,t} \\ \Delta y_{2,t} \end{bmatrix} &= \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \begin{bmatrix} \beta_1 & \beta_2 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} \\ &= \beta_1 \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \begin{bmatrix} 1 & \beta_2/\beta_1 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} \\ &= \begin{bmatrix} \vartheta_1 \\ \vartheta_2 \end{bmatrix} \begin{bmatrix} 1 & -\beta \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} \\ &= \begin{bmatrix} \vartheta_1 \\ \vartheta_2 \end{bmatrix} (y_{1,t-1} - \beta y_{2,t-1}) + \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} \end{aligned}$$

where $\beta = -(\beta_2/\beta_1)$ is the cointegration coefficient and $\vartheta_1 = \beta_1\alpha_1$ and $\vartheta_2 = \beta_1\alpha_2$ are the speed of adjustment coefficients.

We observe that

$$\Delta y_{1,t} - \beta \Delta y_{2,t} = \vartheta_1 y_{1,t-1} - \beta \vartheta_2 y_{1,t-1} - \beta \vartheta_1 y_{2,t-1} + \beta^2 \vartheta_2 y_{2,t-1} + u_{1,t} - \beta u_{2,t} \tag{3}$$

By rearranging Equation (3) we obtain an AR(1) model for $y_{1,t} - \beta y_{2,t}$:

$$y_{1,t} - \beta y_{2,t} = \delta (y_{1,t-1} - \beta y_{2,t-1}) + u_{1,t} - \beta u_{2,t}$$

where $\delta = 1 + \vartheta_1 - \beta \vartheta_2 = a_{11} + a_{22} - 1$. Since $\{y_{1,t}\}$ and $\{y_{2,t}\}$ are cointegrated, $\{y_{1,t} - \beta y_{2,t}\}$ is a stationary process and so

$$|a_{11} + a_{22} - 1| < 1$$

Lemma 3.5. *The process $\{\Delta \mathbf{y}_t = [\Delta y_{1,t}, \Delta y_{2,t}]'; t = 0, 1, \dots\}$ is jointly unpredictable if and only if*

$$\mathbf{A} = \mathbf{I}$$

Proof. (\Rightarrow) process $\{\Delta \mathbf{y}_t = [\Delta y_{1,t}, \Delta y_{2,t}]'; t = 0, 1, \dots\}$ is jointly unpredictable, then

$$E(\mathbf{y}_t | \sigma(\mathbf{y}_{t-1}, \dots, \mathbf{y}_1)) = \mathbf{y}_{t-1}$$

On the other hand, since

$$\mathbf{y}_t = \mathbf{A} \mathbf{y}_{t-1} + \mathbf{u}_t$$

with $E(\mathbf{u}_t) = \mathbf{0}$ and $E(\mathbf{u}_t \mathbf{u}'_t) = \Sigma$ for all t and $E(\mathbf{u}_t \mathbf{u}'_s) = \mathbf{0}$ for $s \neq t$, we have that

$$E(\mathbf{y}_t | \sigma(\mathbf{y}_{t-1}, \dots, \mathbf{y}_1)) = \mathbf{A} \mathbf{y}_{t-1}$$

Hence we have

$$A\mathbf{y}_{t-1} = \mathbf{y}_{t-1}$$

and so

$$A = I$$

(\Leftarrow) If

$$A = I$$

then $\mathbf{y}_t = \mathbf{y}_{t-1} + \mathbf{u}_t$ with $E(\mathbf{u}_t) = \mathbf{0}$ and $E(\mathbf{u}_t\mathbf{u}'_t) = \Sigma$ for all t and $E(\mathbf{u}_t\mathbf{u}'_s) = \mathbf{0}$ for $s \neq t$, and hence we have

$$E(\mathbf{y}_t | \sigma(\mathbf{y}_{t-1}, \dots, \mathbf{y}_0)) = \mathbf{y}_{t-1}$$

Thus we can conclude that the process $\{\Delta\mathbf{y}_t = [\Delta y_{1,t}, \Delta y_{2,t}]'; t = 1, \dots\}$ is jointly unpredictable.

Before to conclude this subsection we observe that Equation (2) can be written in lag operator notation. The lag operator L is defined such that $L\mathbf{y}_t = \mathbf{y}_{t-1}$. We have that

$$(I - AL)\mathbf{y}_t = \mathbf{u}_t$$

or

$$\begin{bmatrix} 1 - a_{11}L & -a_{12}L \\ -a_{21}L & 1 - a_{22}L \end{bmatrix} \begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix}$$

3.2. Proof of Theorem 1

Sufficiency. If

$$d(I_{\Delta y_1}(t+), H_{\Delta y_2}(t)) = \sigma_{\Delta y_1} \text{ and } d(I_{\Delta y_2}(t+), H_{\Delta y_1}(t)) = \sigma_{\Delta y_2}$$

then, by Lemma 1, we have

$$I_{\Delta y_1}(t+) \perp H_{\Delta y_2}(t)$$

and

$$I_{\Delta y_2}(t+) \perp H_{\Delta y_1}(t)$$

Now we assume that a_{12} and a_{21} are not both equal to zero. We can have three cases.

Case (1) $a_{12} \neq 0$ and $a_{21} = 0$. This implies that

$$r_{1,t} = (a_{11} - 1)y_{1,t} + a_{12}y_{2,t} + u_{1,t}$$

and

$$\Delta y_{2,t} = u_{2,t}$$

Thus

$$\begin{aligned} < \Delta y_{1,t+1}, \Delta y_{2,t} > &= E(\Delta y_{1,t+1} \Delta y_{2,t}) \\ &= (a_{11} - 1)E(y_{1,t}u_{2,t}) + a_{12}E(y_{2,t}u_{2,t}) + E(u_{1,t+1}u_{2,t}) \\ &= (a_{11} - 1)E(y_{1,t}u_{2,t}) + a_{12}E(u_{2,t} \sum_{s=1}^t u_{2,s}) \\ &= (a_{11} - 1)E(y_{1,t}u_{2,t}) + a_{12}\sigma_2^2 \end{aligned}$$

Now, we note that

$$E(y_{1,t}u_{2,t}) = a_{11}^t E(y_{1,0}u_{2,t}) = 0$$

Thus

$$\langle \Delta y_{1,t+1}, \Delta y_{2,t} \rangle = a_{12}\sigma_2^2 \neq 0$$

but this is absurd since

$$I_{\Delta y_1}(t+) \perp H_{\Delta y_2}(t)$$

Case (2) $a_{12} = 0$ and $a_{21} \neq 0$. In this case we have

$$\langle \Delta y_{2,t+1}, \Delta y_{1,t} \rangle = a_{21}\sigma_1^2 \neq 0$$

Again this is absurd since

$$I_{\Delta y_2}(t+) \perp H_{\Delta y_1}(t)$$

Case (3) $a_{12} \neq 0$ and $a_{21} \neq 0$. We note that

$$\begin{bmatrix} \Delta y_{1,t} \\ \Delta y_{2,t} \end{bmatrix} = \begin{bmatrix} (1 - a_{22}L)\gamma(L) & a_{12}L\gamma(L) \\ a_{21}L\gamma(L) & (1 - a_{11}L)\gamma(L) \end{bmatrix} \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix}$$

where

$$\gamma(L) = \frac{1 - L}{(1 - a_{11}L)(1 - a_{22}L) - a_{12}a_{21}L^2}$$

By Lemma 2, we have that $\{y_{1t}\}$ and $\{y_{2t}\}$ are cointegrated and hence the matrix

$$\mathbf{A} - \mathbf{I} = \begin{bmatrix} a_{11} - 1 & a_{12} \\ a_{21} & a_{22} - 1 \end{bmatrix}$$

has rank 1. It follows that

$$a_{12}a_{21} = (1 - a_{11})(1 - a_{22})$$

Thus

$$\begin{aligned} \gamma(L) &= \frac{1 - L}{(1 - a_{11}L)(1 - a_{22}L) - (1 - a_{11})(1 - a_{22})L^2} \\ &= \frac{1 - L}{(1 - L)(1 + L) - (1 - L)(a_{11} + a_{22})L} \\ &= \frac{1}{1 + L - (a_{11} + a_{22})L} \\ &= \frac{1}{1 - (a_{11} + a_{22} - 1)L} \\ &= \frac{1}{1 - \delta L} \end{aligned}$$

where $\delta = a_{11} + a_{22} - 1$.

Since $\{y_{1t}\}$ and $\{y_{2t}\}$ are cointegrated, by Lemma 3 we have that $|\delta| < 1$ and hence

$$\gamma(L) = 1 + \delta L + \delta^2 L^2 + \dots$$

Now, we can have two cases.

Case (a) $\delta = 0$. In this case we have

$$\Delta y_{1,t} = u_{1,t} - a_{22}u_{1,t-1} + a_{12}u_{2,t-1}$$

and

$$\Delta y_{2,t} = a_{21}u_{1,t-1} + u_{2,t} - a_{11}u_{2,t-1}$$

Thus

$$\langle \Delta y_{1,t+1}, \Delta y_{2,t} \rangle = a_{12}\sigma_2^2 \neq 0$$

and

$$\langle \Delta y_{2,t+1}, \Delta y_{1,t} \rangle = a_{21}\sigma_1^2 \neq 0$$

but this is absurd since

$$I_{\Delta y_1}(t+) \perp H_{\Delta y_2}(t)$$

and

$$I_{\Delta y_2}(t+) \perp H_{\Delta y_1}(t)$$

Case (b) $\delta \neq 0$. In this case we have

$$\begin{aligned} \Delta y_{1,t} &= u_{1,t} + (a_{11} - 1)u_{1,t-1} + \delta(a_{11} - 1)u_{1,t-2} + \dots + a_{12}u_{2,t-1} + a_{12}\delta u_{2,t-2} + \dots \\ &= u_{1,t} + (a_{11} - 1) \sum_{i=0}^{\infty} \delta^i u_{1,t-1-i} + a_{12} \sum_{i=0}^{\infty} \delta^i u_{2,t-1-i} \end{aligned}$$

and

$$\begin{aligned} \Delta y_{2,t} &= u_{2,t} + (a_{22} - 1)u_{2,t-1} + \delta(a_{22} - 1)u_{2,t-2} + \dots + a_{21}u_{1,t-1} + a_{21}\delta u_{1,t-2} + \dots \\ &= u_{2,t} + (a_{22} - 1) \sum_{i=0}^{\infty} \delta^i u_{2,t-1-i} + a_{21} \sum_{i=0}^{\infty} \delta^i u_{1,t-1-i} \end{aligned}$$

Thus

$$\langle \Delta y_{1,t+1}, \Delta y_{2,t} \rangle = (a_{11} - 1)\sigma_1^2 \frac{\delta}{1 - \delta^2} a_{21} + \left[1 + (a_{22} - 1) \frac{\delta}{1 - \delta^2} \right] \sigma_2^2 a_{12}$$

and

$$\langle \Delta y_{2,t+1}, \Delta y_{1,t} \rangle = (a_{22} - 1)\sigma_2^2 \frac{\delta}{1 - \delta^2} a_{12} + \left[1 + (a_{11} - 1) \frac{\delta}{1 - \delta^2} \right] \sigma_1^2 a_{21}$$

Now, we consider the system

$$\begin{bmatrix} \left[1 + (a_{22} - 1) \frac{\delta}{1 - \delta^2} \right] \sigma_2^2 & (a_{11} - 1)\sigma_1^2 \frac{\delta}{1 - \delta^2} \\ (a_{22} - 1)\sigma_2^2 \frac{\delta}{1 - \delta^2} & \left[1 + (a_{11} - 1) \frac{\delta}{1 - \delta^2} \right] \sigma_1^2 \end{bmatrix} \begin{bmatrix} a_{12} \\ a_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The determinant of the matrix

$$\begin{bmatrix} \left[1 + (a_{22} - 1) \frac{\delta}{1 - \delta^2} \right] \sigma_2^2 & (a_{11} - 1)\sigma_1^2 \frac{\delta}{1 - \delta^2} \\ (a_{22} - 1)\sigma_2^2 \frac{\delta}{1 - \delta^2} & \left[1 + (a_{11} - 1) \frac{\delta}{1 - \delta^2} \right] \sigma_1^2 \end{bmatrix}$$

is

$$\sigma_1^2 \sigma_2^2 \left(1 - \frac{\delta}{1 + \delta} \right)$$

Since $\sigma_1^2 \sigma_2^2 > 0$ and $\frac{\delta}{1+\delta} \neq 1$, we have that

$$\sigma_1^2 \sigma_2^2 \left(1 - \frac{\delta}{1+\delta}\right) \neq 0$$

Thus $a_{12} \neq 0, a_{21} \neq 0$ implies that $\langle \Delta y_{1,t+1}, \Delta y_{2,t} \rangle \neq 0$ or $\langle \Delta y_{2,t+1}, \Delta y_{1,t} \rangle \neq 0$, but this is absurd since

$$I_{\Delta y_1}(t+) \perp H_{\Delta y_2}(t)$$

and

$$I_{\Delta y_2}(t+) \perp H_{\Delta y_1}(t)$$

In all Cases (1–3) we obtain an absurd conclusion, thus we can state that

$$a_{12} = 0, \quad a_{21} = 0$$

Now, we prove that $a_{11} = a_{22} = 1$. We have that

$$\Delta y_{i,t} = (a_{ii} - 1)y_{i,t} + u_{it} \quad i = 1, 2$$

Since the error term $u_t = [u_{1,t}, u_{2,t}]'$ is stationary these equations must be balanced, that is the order of integration of $\Delta y_{i,t}$ and $(a_{ii} - 1)y_{i,t}$ must be the same. By the hypothesis that $y_{i,t} \sim I(1)$, it follows that $\Delta y_{it} \sim I(0)$ (i.e., stationary) and $(a_{ii} - 1)y_{i,t}$ is $I(1)$, hence $\Delta y_{i,t} = (a_{ii} - 1)y_{i,t} + u_{i,t}$ $i = 1, 2$ implies that $a_{11} = a_{22} = 1$. Thus $\mathbf{A} = \mathbf{I}$ and hence, by Lemma 4, it follows that the process $\{\Delta \mathbf{y}_t; t = 1, \dots\}$ is jointly unpredictable.

Necessity. If the process $\{\Delta \mathbf{y}_t; t = 1, \dots\}$ is jointly unpredictable, then by Lemma 4 it follows that $\mathbf{A} = \mathbf{I}$ and hence $\Delta y_{1,t} = u_{1,t}$ and $\Delta y_{2,t} = u_{2,t} \forall t$. This implies that $P(\Delta y_{1,t+h} | H_{\Delta y_2}(t)) = 0$ and $P(\Delta y_{2,t+h} | H_{\Delta y_1}(t)) = 0 \forall h > 0$. Therefore we have that $\Delta y_{1,t+h} \perp H_{\Delta y_2}(t)$ and $\Delta y_{2,t+h} \perp H_{\Delta y_1}(t) \forall h > 0$. Thus, by Lemma 1, it follows that

$$d(I_{\Delta y_1}(t+), H_{\Delta y_2}(t)) = \sigma_{\Delta y_1} \text{ and } d(I_{\Delta y_2}(t+), H_{\Delta y_1}(t)) = \sigma_{\Delta y_2}$$

Theorem 1 is proved.

4. Conclusions

In this paper we have considered the following geometric condition concerning the distance between information sets

$$d(I_{\Delta y_1}(t+), H_{\Delta y_2}(t)) = \sigma_{\Delta y_1} \text{ and } d(I_{\Delta y_2}(t+), H_{\Delta y_1}(t)) = \sigma_{\Delta y_2} \tag{4}$$

It says that the distances $d(I_{\Delta y_1}(t+), H_{\Delta y_2}(t))$ and $d(I_{\Delta y_2}(t+), H_{\Delta y_1}(t))$ achieve their maximum value, respectively. Theorem 1 tells us that, under the hypothesis that the process \mathbf{y}_t follows a bivariate VAR(1) model, the condition Equation (4) represents a geometric characterization of the notion of joint unpredictability. If this condition holds, the processes Δy_1 and Δy_2 are jointly unpredictable since the past of the bivariate process \mathbf{y}_t does not contain any valuable information about the future of the differenced series. The information in the past is too far from the future information.

Even if the bivariate VAR(1) assumption is far from general, we think that this geometric characterization is useful in order to throw light on the concept of joint unpredictability of a stochastic process.

Acknowledgments

We thank two anonymous referees for their helpful and constructive comments, and Fulvia Focker for helpful comments.

Conflicts of Interest

The author declares no conflict of interest.

References

1. Caporale, G.M.; Pittis, N. Cointegration and predictability of asset prices. *J. Int. Money Financ.* **1998**, *17*, 441–453.
2. Hassapis C.; Kalyvitis, S.; Pittis, N. Cointegration and joint efficiency of international commodity markets. *Q. Rev. Econ. Financ.* **1999**, *39*, 213–231.
3. Brockwell, P.J.; Davis, R.A. *Time Series: Theory and Methods*, 2nd ed.; Springer-Verlag: New York, NY, USA, 1991.
4. Granger, C.W.J. Developments in the study of cointegrated economic variables. *Oxf. Bull. Econ. Stat.* **1986**, *48*, 213–228.
5. Focker, F.; Triacca, U. Interpreting the concept of joint unpredictability of asset returns: A distance approach. *Physica A* **2006**, *2*, 765–770.

© 2013 by the author; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).