*Hokkaido Mathematical Journal Vol. 24 (1995) p. 179–213*

# On a causal analysis of economic time series

Yuji NAKANO

(Received December 15, 1993)

**Abstract.** This paper describes a testing methodology for causal relations between time series. The concept of the local causality and the instantaneous local causality is introduced. The mathematical structure of the local causality is shown. The data of GNP and Money Supply are analized by the proposed test.

*Key words*: Local and weak stationality. $KM_2O$-Langevin equation. Test(S), Local causality, Instantaneous local causality, Real GNP, Money Supply, Local Causal Test, Instantaneous Local Causal Test.

## 1. Introduction

A weakly stationary process, whose time parameter space is a finite interval of $\mathbf{T}$, is called a local and weakly stationary process. The letter $\mathbf{T}$ denotes $\{0, \pm 1, \pm 2, \ldots\}$. In the present paper, we propose a concept of causality, which we call local causality, in local and weakly stationary processes. Local causality is defined from the predictional point of view. We propose a method, which we call the Local Causal Test, how to test local causal relations in local and weakly stationary processes. As an application, time series of Money Supply and Real Gross National Product (RGNP), which are known as the most important time series in economics, are analyzed to find local causal relations.

The well-known Granger's causality (Granger [7]) was defined for stochastic processes whose time parameter space is $\mathbf{T}$. Up to this day, many works on causal analysis in Granger's sense (eg., [7], Sims [33], Sargent [31], Ram [30], Komura [10]) are known. Among them a test which is called the Granger and Sargent Test is often used. Including these works, almost all studies in time series analysis use simplified models such as autoregressive (AR) or autoregressive moving average (ARMA) models in the model fitting for given data. As economists emphasize, the methods above, assuming "weak stationarity" of given data, have a contradiction from the viewpoint of the theory of stochastic processes (e.g., Sawa [32]).

---

*1991 Mathematics Subject Classification* : 62M10, 90A20.

On the other hand, we do not take the position that given data have "weak stationarity". Okabe and Nakano [26] constructed Test(S) which states a criterion that multi-dimensional data are a realization of a local and weakly stationary process. We apply this Test(S) to given data, including some transformed data. If the data pass Test(S), viz. that they are accepted to be a realization of a local and weakly stationary time series, we proceed to further analysis. In like manner, Okabe [19] defined causality in local and weakly stationary processes from the viewpoint of the prediction and proposed a method how to test it. Okabe and Inoue [25] developed this analysis further. As we mentioned at the beginning of this section, we shall propose a concept of causality from the predictional point of view, and develop its analysis such as the Local Causal Test in this paper. We compare the Local Causal Test with the Granger-Sargent Test when both tests are applied to time series of Money Supply and RGNP.

The outline of this paper is as follows: Through this paper the theory of $\mathrm{KM_2O}$-Langevin equations, which are associated with local and weakly stationary processes, plays a crucial role. Therefore we overview in §2 briefly the theory of $\mathrm{KM_2O}$-Langevin equations. As an application of this theory, it was proposed to apply Test(S) to the question of whether given data are a realization of a local and weakly stationary process or not (see [26]). In §3 we summarize the deduced process of Test(S). We applied Test(S) to quarterly data of Money Supply and RGNP in three periods. [30] and [10] analyzed the causal relation in Granger's sense between these data in the period from 1955-I to 1971-II. They assumed that the first order differences of log-transformed data are a realization of an AR model. However the result of Table 3.1 shows that they are not a realization of a local and weakly stationary process. In the periods from 1965 to 1987 and from 1965 to 1990, we accept from Table 3.2 and Table 3.3 that the second order differences of the original data are a realization of a local and weakly stationary process.

§4 introduces Granger's causality for stochastic processes whose time parameter space is $\mathbf{T}$ and summarizes briefly the Granger-Sargent Test which is a representative test for Granger's causality. As above mentioned, the Granger-Sargent Test also assumes that given data are a realization of an AR model. This test is applied to quarterly data of Money Supply and RGNP in the periods from 1965 to 1987 and from 1965 to 1990. The results shown in Table 4.1 and Table 4.2 report that RGNP causes, in Granger's sense, Money Supply in both periods. However converse relations are not accepted . These results are compared with the results of the Local Causal Test which is developed in §6.

We define in §5 the local causality and the instantaneous local causality between local and weakly stationary processes, and characterize them. On the basis of the theory of $KM_2O$-Langevin equations, we investigate in Theorem 5.2 a mathematical structure of the causal relation between local and weakly stationary processes. Theorem 5.1 gives a theory which judges causal relations between local and weakly stationary processes. Moreover, a theory which judges instantaneous causal relations between local and weakly stationary processes is given by Theorem 5.5.

In §6 we apply the theory developed in §5 to data analysis. As an application of Theorem 5.1, we propose the Local Causal Test which tests the local causality between given time series, which are accepted by Test(S) as a realization of a local and weakly stationary process. In the same way, the Instantaneous Local Causal Test is proposed as an application of Theorem 5.5. Both tests are applied to quarterly data of Money Supply and RGNP in the periods from 1965 to 1987 and from 1965 to 1990. Table 5.1 and Table 5.2 report that RGNP locally causes Money Supply in both periods and Money Supply locally causes RGNP in the period from 1965 to 1987. Now, it is an established theory that Money Supply and RGNP are mutually related. On the other hand, we can not accept that Money Supply locally causes RGNP in the period from 1965 to 1990. However, this phenomenon is explicable from the point of view of economics. Since the Granger-Sargent Test does not accept in both periods that Money Supply causes RGNP in Granger's sense, we can assert the efficiency of the Local Causal Test.

## 2. $KM_2O$-Langevin equations

The theory of $KM_2O$-Langevin equations was introduced by Okabe [17]. Following the notation and terminology of [17] and Okabe-Nakano [26], we overview it in this section. Let $d, N \in \mathbf{N}$. Let $\mathbf{X} = (X(n); |n| \le N)$ be any $d$-dimensional stochastic process on a probability space $(\Omega, \mathcal{B}, \mathcal{P})$.

**Definition 2.1** $\mathbf{X}$ is called *a local and weakly stationary process with covariance function* $R$ if it holds that for any $n, m \in \mathbf{T}, |n| \le N, |m| \le N$,

$$E(X) = E[X(n)] = \mu \tag{2.1}$$

$$E[(X(n)-\mu)^t(X(m)-\mu)] = R(n-m). \tag{2.2}$$

Without loss of generality, we assume that $\mu = \mathbf{0}$.

For any $n \in \{1, \cdots, N\}$, we define a block Toeplitz matrix $S_n \in M(nd; \mathbf{R})$ by

$$S_n = \begin{pmatrix} R(0) & R(1) & \cdots & R(n-1) \\ {}^tR(1) & R(0) & \cdots & R(n-2) \\ \vdots & \vdots & \ddots & \vdots \\ {}^tR(n-2) & {}^tR(n-3) & \cdots & R(1) \\ {}^tR(n-1) & {}^tR(n-2) & \cdots & R(0) \end{pmatrix} \tag{2.3}$$

In this paper, we assume

$$R(0) \in GL(d; \mathbf{R}) \tag{A-1}$$

and

$$S_n \in GL(nd; \mathbf{R}). \tag{A-2}$$

We set

$$X(n) = \begin{pmatrix} X_1(n) \\ X_2(n) \\ \vdots \\ X_d(n) \end{pmatrix} \qquad (|n| \le N). \tag{2.4}$$

For $n_1 < n_2$, $n_1$, $n_2 \in \{-N, \cdots, N\}$, we define $\mathbf{M}_{n_1}^{n_2}(\mathbf{X})$, which is a closed linear subspace of $L^2(\Omega, \mathcal{B}, P)$ by

$$\begin{aligned} \mathbf{M}_{n_1}^{n_2}(\mathbf{X}) \ &= \ \text{the closed linear hull of} \\ &\qquad \{X_j(m); 1 \le j \le d, \ n_1 \le m \le n_2\}. \end{aligned} \tag{2.5}$$

Especially, we define that

$$\mathbf{M}_0^{-1}(\mathbf{X}) = \mathbf{M}_1^0(\mathbf{X}) = \mathbf{0}. \tag{2.6}$$

For $n \in \{0, \cdots, N\}$, $P_{\mathbf{M}_0^{n-1}(\mathbf{X})}$ is a projection operator on $\mathbf{M}_0^{n-1}(\mathbf{X})$, and $P_{\mathbf{M}_{-n+1}^0(\mathbf{X})}$ projection operator on $\mathbf{M}_{-n+1}^0(\mathbf{X})$. Now, the random forces of $\mathbf{X}$, which we call, $\nu_+ = (\nu_+(n); 0 \le n \le N)$, $\nu_- = (\nu_-(-n); 0 \le n \le N)$ are introduced:

$$\nu_+(n) = X(n) - P_{\mathbf{M}_0^{n-1}(\mathbf{X})}X(n) \tag{2.7}$$

$$\nu_-(-n) = X(-n) - P_{\mathbf{M}^0_{-n+1}(\mathbf{X})} X(-n). \tag{2.8}$$

It holds that

$$\nu_+(0) = \nu_-(0) = X(0). \tag{2.9}$$

$P_{\mathbf{M}_0^{n-1}(\mathbf{X})} X(n)$ and $P_{\mathbf{M}^0_{-n+1}(\mathbf{X})} X(-n)$ are called fluctuation parts of $\mathbf{X}$.

For any $n \in \{1, \cdots, N\}$, $k \in \{0, \cdots, n-1\}$, there exist $\gamma_+(n,k)$, $\gamma_-(n,k) \in M(d; \mathbf{R})$ such that

$$P_{\mathbf{M}_0^{n-1}(\mathbf{X})} X(n) = -\sum_{k=0}^{n-1} \gamma_+(n,k) X(k) \tag{2.10}$$

$$P_{\mathbf{M}^0_{-n+1}(\mathbf{X})} X(-n) = -\sum_{k=0}^{n-1} \gamma_-(n,k) X(-k). \tag{2.11}$$

It holds that

$$\mathbf{M}_0^n(\mathbf{X}) = \mathbf{M}_0^n(\nu_+) \tag{2.12}$$

$$\mathbf{M}^0_{-n}(\mathbf{X}) = \mathbf{M}^0_{-n}(\nu_-). \tag{2.13}$$

The following Theorem 2.1 is known as an expression formula of $\mathbf{X}$.

**Theorem 2.2** *There exists the unique system*

$$\{\gamma_+(n,k),\ \gamma_-(n,k) \in M(d; \mathbf{R});\ 0 \le k < n \le N\}$$

*such that*

$$X(n) = -\sum_{k=0}^{n-1} \gamma_+(n,k) X(k) + \nu_+(n) \tag{2.14}$$

$$X(-n) = -\sum_{k=0}^{n-1} \gamma_-(n,k) X(-k) + \nu_-(-n). \tag{2.15}$$

Here, $\delta_+(n)$, $\delta_-(n)$ $(1 \le n \le N)$ which are known as partial correlation functions, are defined as

$$\delta_+(n) = \gamma_+(n,0),\ \delta_-(n) = \gamma_-(n,0). \tag{2.16}$$

The equations (2.14) and (2.15) are called KM$_2$O-Langevin equations associated with $\mathbf{X}$. There exist interactions between fluctuation parts and dissipation parts. These interactions are called the Fluctuation-Dissipation Theorem (FDT). For $n \in \{0, \cdots, N\}$, we set

$$E(\nu_+(n)^t\nu_+(n)) = V_+(n) \text{ and } E(\nu_-(-n)^t\nu_-(-n)) = V_-(n). \quad (2.17)$$

FDT is described as follows:

**Theorem 2.3** (FDT). *For* $1 \le k < n \le N$,

(i) $\gamma_+(n,k) = \gamma_+(n-1,k-1) + \delta_+(n)\gamma_-(n-1,n-1-k)$

(ii) $\gamma_-(n,k) = \gamma_-(n-1,k-1) + \delta_-(n)\gamma_+(n-1,n-1-k)$

(iii) $\delta_+(n) = -(R(n) + \sum_{m=0}^{n-2}\gamma_+(n-1,m)R(m+1))V_-(n-1)^{-1}$

(iv) $\delta_-(n) = -({}^tR(n) + \sum_{m=0}^{n-2}\gamma_-(n-1,m){}^tR(m+1))V_+(n-1)^{-1}$.

*For* $1 \le n \le N$,

(v) $V_+(n) = (I - \delta_+(n)\delta_-(n))V_+(n-1)$

(vi) $V_-(n) = (I - \delta_-(n)\delta_+(n))V_-(n-1)$.

*For the special case* $n = 0$, *we get*

(vii) $V_+(0) = V_-(0) = R(0)$

(viii) $\delta_+(1) = -R(1)R(0)^{-1}$

(ix) $\delta_-(1) = -{}^tR(1)R(0)^{-1}$.

*When* $d = 1$, $R(n) = {}^tR(-n)$. *Therefore, we can see that*

$$\begin{cases} \delta_+(*) = \delta_-(*) \\ \gamma_+(*,\cdot) = \gamma_-(*,\cdot) \\ V_+(*) = V_-(*). \end{cases} \quad (2.18)$$

The system

$$\{\gamma_+(n,k),\ \gamma_-(n,k),\ V_+(l),\ V_-(l);\ 0 \le k < n \le N,\ 0 \le l \le N\}$$

is called a KM$_2$O-Langevin data associated with covariance function $R$.

## 3. Test(S)

Test(S) was proposed by Okabe-Nakano [26] to test whether a given time series is a realization of a local and weakly stationary process or not. Following [26], we summarize the deduced processes of Test(S).

Any $d, N \in \mathbf{N}$ be fixed. We are given any $N+1$ vectors $\mathcal{Z}(n) \in \mathbf{R}^d$ $(0 \le n \le N)$. $\mathcal{Z} = (\mathcal{Z}(n);\ 0 \le n \le N)$ is called data. The sample mean vector

$\mu^{\mathcal{Z}}$ of $\mathcal{Z}$ and the sample covariance matrix function $R^{\mathcal{Z}} = (R^{\mathcal{Z}}_{jk})_{1\le j,k\le d}$ of $\mathcal{Z}$ are defined as follows:

$$\mu^{\mathcal{Z}} \equiv \frac{1}{N+1}\sum_{m=0}^{N}\mathcal{Z}(m) \tag{3.1}$$

$$R^{\mathcal{Z}}_{jk}(n) \equiv \frac{1}{N+1}\sum_{m=0}^{N-n}(\mathcal{Z}_j(n+m)-\mu^{\mathcal{Z}}_j)(\mathcal{Z}_k(m)-\mu^{\mathcal{Z}}_k) \tag{3.2}$$

$$R^{\mathcal{Z}}_{jk}(-n) \equiv R^{\mathcal{Z}}_{kj}(n), \tag{3.3}$$

where

$$\mu^{\mathcal{Z}} = \begin{pmatrix}\mu^{\mathcal{Z}}_1\\ \vdots\\ \mu^{\mathcal{Z}}_d\end{pmatrix},\ \mathcal{Z}(n) = \begin{pmatrix}\mathcal{Z}_1(n)\\ \vdots\\ \mathcal{Z}_d(n)\end{pmatrix} \qquad (0\le n\le N). \tag{3.4}$$

The standardized data $\mathcal{X} = (\mathcal{X}(n);\ 0\le n\le N)$ of $\mathcal{Z}$ is defined as follows:

$$\mathcal{X}(n) = \begin{pmatrix}\sqrt{R^{\mathcal{Z}}_{11}(0)^{-1}} & & \mathbf{0}\\ & \ddots & \\ \mathbf{0} & & \sqrt{R^{\mathcal{Z}}_{dd}(0)^{-1}}\end{pmatrix}(\mathcal{Z}(n)-\mu^{\mathcal{Z}}). \tag{3.5}$$

Let $R^{\mathcal{X}} = (R^{\mathcal{X}}_{jk})_{1\le j,k\le d}$ be the sample covariance matrix function of $\mathcal{X}$ defined similarly to (3.1), (3.2) and (3.3). We can define the sample block Toeplitz matrix $S^{\mathcal{X}}_n$ $(1\le n\le N)$ similarly to (2.3). Here, it is assumed that

$$S^{\mathcal{X}}_n \in GL(nd;\mathbf{R}) \qquad (1\le n\le N). \tag{3.6}$$

Replacing $R$ by $R^{\mathcal{X}}$ in the algorithm from (i) to (ix) in §2, we get the sample KM$_2$O-Langevin data

$$\{\gamma_+(n,k),\ \gamma_-(n,k),\ V_+(l),\ V_-(l);\ 0\le k<n\le N,\ 0\le l\le N\}.$$

Then $\nu_+ = (\nu_+(n);\ 0\le n\le N)$ which is called the sample random force of data $\mathcal{X}$ is introduced by

$$\begin{cases}\nu_+(0) = \mathcal{X}(0)\\ \nu_+(n) = \mathcal{X}(n) + \displaystyle\sum_{k=0}^{n-1}\gamma_+(n,k)\mathcal{X}(k) \qquad (1\le n\le N).\end{cases} \tag{3.7}$$

We choose lower triangular matrices $W_+(n) \in GL(d;\mathbf{R})$ such that

$$V_+(n) = W_+(n)^t W_+(n) \qquad (0 \le n \le N). \tag{3.8}$$

We define the $d$-dimensional data $\xi_+ = (\xi_+(n);\ 0 \le n \le N)$ by

$$\xi_+(n) = W_+(n)^{-1}\nu_+(n) \qquad (0 \le n \le N). \tag{3.9}$$

Set

$$\xi_+(n) = \begin{pmatrix} \xi_{+1}(n) \\ \vdots \\ \xi_{+d}(n) \end{pmatrix} \qquad (0 \le n \le N). \tag{3.10}$$

Rearranging (3.10), we can construct the one-dimensional data $\xi = (\xi(n);$ $0 \le n \le d(N+1)-1)$ as follows: For $n = 0, \cdots, d(N+1)-1$,

$$\xi(n) = \xi_{+p}(m),\ n = dm + p - 1 \quad (1 \le p \le d,\ 0 \le m \le N). \tag{3.11}$$

Then, the Construction Theorem of Okabe [17] suggests that (S.1) and (S.2) below are equivalent to each other.

(S.1) *$\mathcal{X}$ is a realization of a local and weakly stationary time series with $R^{\mathcal{X}}$ as its covariance function.*

(S.2) *$\xi$ realizes an one-dimensional standardized white noise.*

To test (S.2), we introduce

$$\mu^{\xi},\ (v^{\xi}-1)^{\sim} \text{ and } R^{\xi}(n,m)\ (1 \le n \le L_N,\ 0 \le m \le L_N - n)$$

by

$$\mu^{\xi} = \frac{1}{d(N+1)} \sum_{k=0}^{d(N+1)-1} \xi(k) \tag{3.12}$$

$$\begin{aligned}(v^{\xi}-1)^{\sim} = {} & \frac{1}{d(N+1)} \left( \sum_{k=0}^{d(N+1)-1} \xi(k)^2 \right) \\ & \times \left( \sum_{k=0}^{d(N+1)-1} (\xi(k)^2 - 1)^2 \right)^{-1/2}\end{aligned} \tag{3.13}$$

$$R^{\xi}(n,m) = \frac{1}{d(N+1)} \sum_{k=m}^{d(N+1)-1-n} \xi(k)\xi(n+k). \tag{3.14}$$

Here $L_N$ is an effective length of $R^{\xi}$, in this case, is taken to be $L_N = [2\sqrt{d(N+1)}] - 1$.

We institute the following criterion (M), (V), and (O) for checking whether $\xi$ satisfies (S.2) or not.

(M) $\sqrt{d(N+1)}|\mu^{\xi}| < 1.96$

(V) $|(v^{\xi} - 1)^{\sim}| < 2.2414$

(O) for any $n, m$ $(1 \leq n \leq L_N,\ 0 \leq m \leq L_N - n)$

$$d(N+1)(\sum_{j=1}^{2}(L_{n,m}^{(j)})^{1/2})^{-1}|R^{\xi}(n,m)| < 1.96.$$

Here $L_{n,m}^{(j)}$ $(1 \leq j \leq 2)$ are defined as follows: Dividing $d(N+1)$ and $m$ by $2n$ and $n$ respectively, we get the following expression form.

$$d(N+1) = q(2n) + r \quad (0 \leq r \leq 2n-1) \tag{3.15}$$

$$m = sn + t \qquad (0 \leq t \leq n-1). \tag{3.16}$$

If $r \in \{0, \cdots, n\}$, then

$$\begin{cases} L_{n,m}^{(1)} = \begin{cases} n(q + (s/2)) - m & (s \text{ is even}) \\ n(q - (s+1)/2) & (s \text{ is odd}) \end{cases} \\ L_{n,m}^{(2)} = \begin{cases} n(q - 1 - (s/2)) + r & (s \text{ is even}) \\ n(q - 1 + (s+1)/2) + r - m & (s \text{ is odd}) \end{cases} \end{cases} \tag{3.17}$$

and if $r \in \{n+1, \cdots, 2n-1\}$,

$$\begin{cases} L_{n,m}^{(1)} = \begin{cases} n(q - 1 + (s/2)) + r - m & (s \text{ is even}) \\ n(q - 1 - (s+1)/2) + r & (s \text{ is odd}) \end{cases} \\ L_{n,m}^{(2)} = \begin{cases} n(q - (s/2)) & (s \text{ is even}) \\ n(q + (s+1)/2) - m & (s \text{ is odd}). \end{cases} \end{cases} \tag{3.18}$$

Now, it is known that the estimator $R(n)$ has a poor performance when $n$ comes close to $N$. A rule of experience concerning data analysis tells us that an effective number of the sample covariance matrix function $R^{\mathcal{X}}$ is considered to be at most $[3\sqrt{N+1}/d]$. Therefore, we set

$$M = [3\sqrt{N+1}/d] - 1. \tag{3.19}$$

Making use of the reliable $\{R(n);\ 0 \leq n \leq M\}$ and the reliable subsystem

$\{\gamma_+(n,k),\ \gamma_-(n,k),\ V_+(l), V_-(l);\ 0 \le k < n \le M,\ 0 \le l \le M\}$, we restate the new criterion alternative to (M), (V) and (O).

For each $i \in \{0, \cdots, N-M\}$, we consider the shifted data $\mathcal{X}_i$ with $\mathcal{X}(i)$ as its initial point $\mathcal{X}_i(0)$:

$$\mathcal{X}_i = (\mathcal{X}(i+n); 0 \le n \le M). \tag{3.20}$$

Similarly to (3.7), the sample random force $\nu_{+i} = (\nu_{+i}(n);\ 0 \le n \le M)$ of data $\mathcal{X}_i$ is defined by

$$\begin{cases} \nu_{+i}(0) = \mathcal{X}(i) \\ \nu_{+i}(n) = \mathcal{X}(i+n) + \sum_{k=0}^{n-1} \gamma_+(n,k)\mathcal{X}(i+k) \quad (1 \le n \le M). \end{cases} \tag{3.21}$$

In (3.10) replacing $\xi(n)$ by $\xi_i(n)$, $\xi_{+j}(n)$ $(1 \le j \le d)$ by $\xi_{+ij}(n)$ $(1 \le j \le d)$ and $N$ by $M$ respectively, the one-dimensional data $\xi_i = (\xi_i(n);\ 0 \le n \le d(M+1)-1)$ is constructed similarly to (3.11). Moreover, we replace $\xi(n)$ by $\xi_i(n)$ and $N$ by $M$ from (3.12) to (3.18). Then, we get the criterion $(\mathrm{M})_i$, $(\mathrm{V})_i$ and $(\mathrm{O})_i$ which checks that $\xi_i$ is a realization of a normalized white noise. Concerning the main problem of testing the local and weak stationarity of the original data $\mathcal{Z}$, [26] proposed:

**Test(S)** : *the rate of* $i \in \{0, \cdots, N-M\}$ *for which* $(\mathrm{M})_i$ (resp. $(\mathrm{V})_i$ *and* $(\mathrm{O})_i$ ) *holds is over* 80 *percent* (resp. 70 *percent and* 80 *percent*).

We say that data $\mathcal{Z}$ is a realization of a local and weakly stationary process if Test(S) is accepted. Also we say simply that $\mathcal{Z}$ has the local and weak stationarity.

The efficiency of Test(S) was certified in [26].

Money Supply and Gross National Product (GNP) are known as representative economic indices. Money Supply is the stock of money consisting of coin, currency, and bank demand deposits. There are several ways to define Money Supply. The bank of Japan has used for some time $\mathrm{M}_1$ as Money Supply, but now mainly uses $\mathrm{M}_2$ +CD. Gross National Product is the the total value of the goods and services produced in a nation during a specific period. There are two kinds. One is called Real Gross National Product (RGNP) and another is called Nominal Gross National Product (NGNP).

Ram [30] and Komura [10] discussed the Granger's causal relations between quarterly time series, which are Money Supply ($\mathrm{M}_1$) and RGNP of Japan from 1955-I to 1971-II. They assumed that given data have weak stationarity in a wide sense.

Now, we apply Test(S) to these data. Let data $\mathcal{Z}_1 = ({}^t(\mathcal{Z}_{11}(n), \mathcal{Z}_{12}(n));\ 0 \le n \le 65)$ be $\mathrm{M}_1$ and RGNP above. Transforming $\mathcal{Z}_\infty$, we introduce seven two-dimensional data $\mathcal{Z}_1^{(j)} = (\mathcal{Z}_1^{(j)}(n);\ 0 \le n \le N_j^{(1)})$ $(0 \le j \le 6$, $N_0^{(1)} = N_3^{(1)} = N_4^{(1)} = 65$, $N_1^{(1)} = N_5^{(1)} = N_6^{(1)} = 64$, $N_2^{(1)} = 63)$ by

$$\mathcal{Z}_1^{(j)}(n) = {}^t(\mathcal{Z}_{11}^{(j)}(n), \mathcal{Z}_{12}^{(j)}(n)) \tag{3.22}$$

$$= \begin{cases} {}^t(\mathcal{Z}_{11}(n), \mathcal{Z}_{12}(n)) & (j=0) \\ {}^t(\mathcal{Z}_{11}(n+1) - \mathcal{Z}_{11}(n), \mathcal{Z}_{12}(n+1) - \mathcal{Z}_{12}(n)) & (j=1) \\ {}^t(\mathcal{Z}_{11}^{(1)}(n+1) - \mathcal{Z}_{11}^{(1)}(n), \mathcal{Z}_{12}^{(1)}(n+1) - \mathcal{Z}_{12}^{(1)}(n)) & (j=2) \\ {}^t(\log \mathcal{Z}_{11}(n), \log \mathcal{Z}_{12}(n)) & (j=3) \\ {}^t(\arctan \mathcal{Z}_{11}(n), \arctan \mathcal{Z}_{12}(n)) & (j=4) \\ {}^t(\mathcal{Z}_{11}^{(3)}(n+1) - \mathcal{Z}_{11}^{(3)}(n), \mathcal{Z}_{12}^{(3)}(n+1) - \mathcal{Z}_{12}^{(3)}(n)) & (j=5) \\ {}^t(\arctan \mathcal{Z}_{11}^{(1)}(n), \arctan \mathcal{Z}_{12}^{(1)}(n)) & (j=6). \end{cases}$$

Table 3.1 shows the results of Test(S) for these data.

Table 3.1 Test(S) for ${}^t(\mathrm{M}_1$, RGNP) from 1955-I to 1971-II.

| $j$ | (M) | (V) | (O) | (S) |
|---|---|---|---|---|
| 0 | 1.000 | 0.163 | 0.981 | NS |
| 1 | 0.962 | 0.203 | 1.000 | NS |
| 2 | 1.000 | 0.188 | 1.000 | NS |
| 3 | 0.981 | 0.381 | 1.000 | NS |
| 4 | 1.000 | 0.527 | 1.000 | NS |
| 5 | 0.962 | 0.500 | 0.925 | NS |
| 6 | 0.981 | 0.444 | 1.000 | NS |

Here, (M), (V) and (O) denote the rate of $i$ such that $(\mathrm{M})_i$, $(\mathrm{V})_i$ and $(\mathrm{O})_i$ hold respectively. "S" and "NS" indicate for stationarity and non-stationarity respectively. We could not get "stationary data" as far as we tried.

Let $\mathcal{Z}_2 = ({}^t(\mathcal{Z}_{21}(n), \mathcal{Z}_{22}(n);\ 0 \le n \le 91)$ and $\mathcal{Z}_3 = ({}^t(\mathcal{Z}_{31}(n), \mathcal{Z}_{32}(n);\ 0 \le n \le 103)$ be quarterly time series of $\mathrm{M}_2$+CD and RGNP from 1965 to 1987 and from 1965 to 1990 respectively. Let $j \in \{0, 1, 2, 5\}$. Table 3.2 and Table 3.3 report the results of Test(S) for transformed data $\mathcal{Z}_2^{(j)} =$

$(\mathcal{Z}_2^{(j)}(n);\ 0 \le n \le N_j^{(2)})$ $(N_0^{(2)} = 91,\ N_1^{(2)} = N_5^{(2)} = 90,\ N_2^{(2)} = 89)$ and $\mathcal{Z}_3^{(j)} = (\mathcal{Z}_3^{(j)}(n);\ 0 \le n \le N_j^{(3)})$ $(N_0^{(3)} = 103, N_1^{(3)} = N_5^{(3)} = 102,$ $N_2^{(3)} = 101)$ respectively. Here the number $j$ of $\mathcal{Z}_2^{(j)}$ and $\mathcal{Z}_3^{(j)}$ corresponds to transformations of (3.22). We can find that the second order differences in the original data in both periods have local and weak stationarity. We apply these results in §4 and §6.

Table 3.2 Test(S) for ${}^t(\mathrm{M}_2+\mathrm{CD}, \mathrm{RGNP})$ from 1965 to 1987.

| $j$ | (M) | (V) | (O) | (S) |
|---|---|---|---|---|
| 0 | 1.000 | 0.000 | 1.000 | NS |
| 1 | 0.961 | 0.628 | 0.974 | NS |
| 2 | 0.974 | 0.701 | 0.974 | S |
| 5 | 1.000 | 0.628 | 0.871 | NS |

Table 3.3 Test(S) for ${}^t(\mathrm{M}_2 + \mathrm{CD}, \mathrm{RGNP})$ from 1965 to 1990.

| $j$ | (M) | (V) | (O) | (S) |
|---|---|---|---|---|
| 0 | 1.000 | 0.088 | 1.000 | NS |
| 1 | 0.955 | 0.674 | 0.898 | NS |
| 2 | 0.977 | 0.704 | 0.943 | S |
| 5 | 0.988 | 0.528 | 0.865 | NS |

## 4. Granger's causality and the Granger-Sargent Test

It has long been recognized that high correlation among a set of variables does not in any necessary sense establish that they are causally related (Pierce and Haugh [29]). Wold [34] emphasized the importance of causal analysis in science, explaining the examples of economic time series. Under the circumstances, Granger [7] introduced the definitions of causal relation in stochastic processes whose time parameter spaces are $\mathbf{T}$ in view of the predictability as follows:

Let $\mathbf{X} = (X(n);\ n \in \mathbf{T})$, and $\mathbf{Y} = (Y(n);\ n \in \mathbf{T})$ be $d_1$ and $d_2$-dimensional stochastic processes. For each $n \in \mathbf{T}$, $I_n$ is an information set, including at least $\{X(n), Y(n)\}$. Let $\widetilde{I}(n) = \{I(m), m < n\}$, $\overline{\widetilde{I}}(n) = \{I(m),$ $m \le n\}$. $\widetilde{X}(n)$, $\overline{\widetilde{X}}(n)$, $\widetilde{Y}(n)$, $\overline{\widetilde{Y}}(n)$ are defined similarly. $\widetilde{I}(n) - \widetilde{Y}(n)$ is equal to the set of elements of $\widetilde{I}(n)$ without the elements of $\widetilde{Y}(n)$. Denote by $\sigma^2(X(n)|\widetilde{I}(n))$, $\sigma^2(X(n)|\widetilde{I}(n) - \widetilde{Y}(n))$ the mean square prediction error of $X(n)$ given information set $\widetilde{I}(n)$, $\widetilde{I}(n) - \widetilde{Y}(n)$ respectively. Granger's definitions of causality are:

**Definition 4.1** (Granger's causality) If $\sigma^2(X(n)|\widetilde{I}(n)) < \sigma^2(X(n)|\,\widetilde{I}(n) - \widetilde{Y}(n))$, it is said that $Y(n)$ *causes* $X(n)$ *in the sense of Granger*, denoted by $Y(n) \overset{\mathrm{GC}}{\Longrightarrow} X(n)$. Otherwise, we say that $Y(n)$ *does not cause* $X(n)$ *in the sense of Granger*, denoted by $Y(n) \overset{\mathrm{GC}}{\not\Longrightarrow} X(n)$.

**Definition 4.2** (Granger's instantaneous causality) If $\sigma^2(X(n)|\widetilde{I}(n), \widetilde{\widetilde{Y}}(n)) < \sigma^2(X(n)|\widetilde{I}(n))$, it is said that $Y(n)$ *causes* $X(n)$ *instantaneously in the sense of Granger.*

In the stationary case, $\sigma^2(X(n)|\widetilde{I}(n))$, $\sigma^2(X(n)|\widetilde{I}(n) - \widetilde{Y}(n))$, and $\sigma^2(X(n)|\widetilde{I}(n), \widetilde{\widetilde{Y}}(n))$ are independent of $n$. Then, we denote simply $Y(n) \overset{\mathrm{GC}}{\Longrightarrow} X(n)$ by $\mathbf{Y} \overset{\mathrm{GC}}{\Longrightarrow} \mathbf{X}$.

Under the assumption that given data are a realization of an AR model, some tests (e.g., [7], Sargent [31], Sims [33], etc.) are proposed to test Granger's causal relations among them. Now, one of such tests which is called the Granger-Sargent Test is well known and applied to economic time series (e.g., [30], [10]).

The Granger-Sargent Test is as follows: Let $\mathbf{Z} = (\begin{pmatrix} X(n) \\ Y(n) \end{pmatrix}; n \in \mathbf{T})$ be a bivariate AR($m$)-model with mean $\mathbf{0}$ such that

$$X(n) = \sum_{i=1}^{m} a_i X(n-i) + \sum_{i=1}^{m} b_i Y(n-i) + u_1(n) \tag{4.1}$$

$$Y(n) = \sum_{i=1}^{m} c_i Y(n-i) + \sum_{i=1}^{m} d_i X(n-i) + u_2(n) \tag{4.2}$$

where $m \in \mathbf{N}$. It is assumed that $I_n = \{X(n), Y(n)\}$, $n \in \mathbf{T}$. To judge $\mathbf{Y} \overset{\mathrm{GC}}{\Longrightarrow} \mathbf{X}$ or not, we test the null hypothesis

$\mathrm{H}_0$ : $b_1 = \cdots = b_m = 0$

against an alternative hypothesis

$\mathrm{H}_1$ : exists $j \in \{1, \cdots, m\}$ such that $b_j \neq 0$.

The testing procedure is as follows: Given data $\mathcal{Z} = (\begin{pmatrix} \mathcal{X}(n) \\ \mathcal{Y}(n) \end{pmatrix};\ 0 \leq n \leq N)$, we estimate the coefficients $\{a_i, b_j;\ 1 \leq i,\ j \leq m\}$ by least square estimation. Then, the coefficient of determination of (4.1) is defined by

$$R^2 = \frac{\sum_{n=m}^{N}(\sum_{i=1}^{m} a_i\mathcal{X}(n-i) + \sum_{i=1}^{m} b_i\mathcal{Y}(n-i))^2}{\sum_{n=m}^{N} \mathcal{X}(n)^2}. \tag{4.3}$$

Secondly, it is assumed that $\mathcal{X} = (\mathcal{X}(n);\ 0 \leq n \leq N)$ is a realization of the following one-dimensional AR($m$)-model

$$X(n) = \sum_{i=1}^{m} e_i X(n-i) + u_3(n). \tag{4.4}$$

Using data $\mathcal{X}$, the coefficients $\{e_i;\ 1 \leq i \leq m\}$ are estimated by least square estimation. The coefficient of determination of (4.4) is given by

$$R_1^2 = \frac{\sum_{n=m}^{N}(\sum_{i=1}^{m} e_i\mathcal{X}(n-i))^2}{\sum_{n=m}^{N} \mathcal{X}(n)^2}. \tag{4.5}$$

Now, the test statistic $F_1$ is defined by

$$F_1 = \frac{(R^2 - R_1^2)/m}{(1-R^2)/(N-3m+1)}. \tag{4.6}$$

At the $\alpha$ significant level,

$$\begin{aligned} &\text{If} F_1 > F(m, N-3m+1)_\alpha \text{ then reject } \mathrm{H}_0 \\ &\text{If} F_1 \leq F(m, N-3m+1)_\alpha \text{ then accept } \mathrm{H}_0 \end{aligned}$$

where $F(m, N-3m+1)_\alpha$ is the critical value at the $\alpha$ level of $F$ distribution with $m$ and $N-3m+1$ degrees of freedom.

To test $\mathbf{X} \overset{\mathrm{GC}}{\Longrightarrow} \mathbf{Y}$ or not, exchanging of $X$ for $Y$, we can define $F_2$ similarly to $F_1$.

*Remark* 4.1 Coefficients of (4.1), (4.2) and (4.4) are conveniently estimated by the sample covariance matrix function of $\mathcal{Z}$ (e.g., Akaike and Nakagawa [1], [10]), alternative to least square estimation.

We apply the Granger-Sargent Test to the quarterly data of ${}^t$(M$_2$+CD, RGNP) in two periods from 1965 to 1987 and from 1965 to 1990. In §3, the second order differences in the original data were accepted to be stationary. It is here assumed that the order $m$ of the AR model is at most $\sqrt{(N+1)}/2$.

Table 4.1 The Granger-Sargent Test for ${}^{t}(\mathrm{M}_2 + \mathrm{CD}, \mathrm{RGNP})$ from 1965 to 1987.

| $m$ | $N - 3m + 1$ | $F_1$ | $F_2$ |
|---|---|---|---|
| 1 | 87 | 0.0820 | 0.029 |
| 2 | 84 | 4.010* | −0.0232 |
| 3 | 81 | 3.580* | 0.204 |
| 4 | 78 | 2.803* | 0.293 |
| 5 | 75 | 2.157* | 0.343 |
| 6 | 72 | 2.279* | 0.381 |
| 7 | 69 | 3.653* | 0.333 |
| 8 | 66 | 3.271* | 0.469 |
| 9 | 63 | 2.917* | 0.747 |

Table 4.2 The Granger-Sargent Test for ${}^{t}(\mathrm{M}_2 + \mathrm{CD}, \mathrm{RGNP})$ from 1965 to 1990.

| $m$ | $N - 3m + 1$ | $F_1$ | $F_2$ |
|---|---|---|---|
| 1 | 99 | 2.870* | 0.475 |
| 2 | 96 | 2.318 | 0.634 |
| 3 | 93 | 1.937 | 0.295 |
| 4 | 90 | 2.545* | 0.688 |
| 5 | 87 | 1.782 | 1.520 |
| 6 | 84 | 1.608 | 1.209 |
| 7 | 81 | 1.213 | 1.091 |
| 8 | 78 | 0.948 | 0.996 |
| 9 | 75 | 0.956 | 0.969 |
| 10 | 72 | 0.826 | 0.898 |

Table 4.1 and Table 4.2 report the results of the Granger-Sargent Test for the second order differences in the original data in two periods. Since we choose the second order differences in the data, $N = 89$ in the Table 4.1 and $N = 101$ in the Table 4.2 respectively. Here we choose that $\alpha = 0.10$. The symbol "$*$" indicates the cases in which the $F_1$ value or $F_2$ value exceeds the critical value at the 0.10 level. In both periods, there are some cases where $F_1$ exceeds the critical point at the 10 percent level. On the other hand, there are no cases where the $F_2$ exceeds the critical point at the 10

percent level. Therefore we accept in both periods that

$$\mathrm{M}_2 + \mathrm{CD} \overset{\mathrm{GC}}{\not\Longrightarrow} \mathrm{RGNP}, \ \mathrm{RGNP} \overset{\mathrm{GC}}{\Longrightarrow} \mathrm{M}_2 + \mathrm{CD}. \tag{4.7}$$

In §6, these results are compared with the results by the Local Causal Test.

## 5. A local causality and its characterization

We develop, in this section, a causal analysis of local and weakly stationary processes. Granger's causal analysis introduced in §4 assumes that given data are a realization of a bivariate AR model. We do not assume that given data are a realization of a specified process. After accepting by Test(S) that the data are a realization of a local and weak stationary process, we proceed to further analysis. In like manner, Okabe [19] defined a causality in local and weakly stationary processes from the viewpoint of the prediction and proposed a method how to test it. Okabe and Inoue [25] developed this analysis further. The definition of causality by [25] is as follows: Let $\mathbf{X} = (X(n); n \in \mathbf{T})$, $\mathbf{Y} = (Y(n); n \in \mathbf{T})$ be $d_1$ and $d_2$-dimensional stochastic processes respectively. It is said that $\mathbf{Y}$ causes $\mathbf{X}$ in the sense of Okabe-Inoue if for each $n \in \mathbf{T}$ there exists a measurable mapping $F_n$ from the infinite-dimensional space $(\mathbf{R}^{d_2})^{\mathbf{N}^*}$ to the finite dimensional space $\mathbf{R}^{d_1}$ such that $X(n) = F_n(Y(n), Y(n-1), Y(n-2), \cdots)$.

Under the assumption that $\mathbf{Z} = (\begin{pmatrix} X(n) \\ Y(n) \end{pmatrix}; n \in \mathbf{T})$ is a $(d_1 + d_2)$-dimensional weakly stationary process, the necessary and sufficient condition in which $\mathbf{Y}$ causes $\mathbf{X}$ in the sense of Okabe-Inoue was investigated and applied to data analysis in [25]. Moreover, Okabe-Ootsuka [27] and [28] investigated the non-linear prediction problem.

Now, let us introduce the definition of local causality in local and weakly stationary processes and investigate its characterization. Let $\mathbf{X} = (X(n); |n| \leq N)$ and $\mathbf{Y} = (Y(n); |n| \leq N)$ be $d_1$ and $d_2$-dimensional local and weakly stationary processes respectively. Moreover, we assume that $\mathbf{Z} = (Z(n); |n| \leq N)$,

$$Z(n) = \begin{pmatrix} X(n) \\ Y(n) \end{pmatrix} \tag{5.1}$$

is a $(d_1 + d_2)$-dimensional local and weakly stationary process.

We set

$$X(n) = \begin{pmatrix} X_1(n) \\ \vdots \\ X_{d_1}(n) \end{pmatrix}, \qquad Y(n) = \begin{pmatrix} Y_1(n) \\ \vdots \\ Y_{d_2}(n) \end{pmatrix}. \tag{5.2}$$

For each $n \in \{0, \cdots, N\}$, $I(n)$ is an information set at time $n$, including at least $X_i(n)$, $Y_j(n)$ $(1 \le i \le d_1,\ 1 \le j \le d_2)$. The available set at time $n$ is defined by

$$I_0^n = \{I_m; 0 \le m \le n\}. \tag{5.3}$$

Let $J$ be an information set and $\hat{X}_J(n)$ be the linear prediction $X(n)$ by $J$. Then, the prediction error of $X(n)$ by $J$ and its variance are defined :

$$\epsilon(X(n)|J) = X(n) - \hat{X}_J(n) \tag{5.4}$$

$$\sigma^2(X(n)|J) = \|\epsilon(X(n)|J)\|^2. \tag{5.5}$$

$I_0^n - X_0^n$ denotes all the elements of $I_0^n$ eliminated $X_i(m)$ $(0 \le m \le n,$ $1 \le i \le d_1)$. Similarly, $I_0^n - Y_0^n$ is defined. Here we set $I_0^{-1} = I_0^{-1} - X_0^{-1} = I_0^{-1} - Y_0^{-1} = \phi$. Local causality between $\mathbf{X}$ and $\mathbf{Y}$ is defined as follows:

**Definition 5.1** (local causality) If there exits $n \in \{0, \cdots, N\}$ such that

$$\sigma(X(n)|I_0^{n-1}) < \sigma(X(n)|I_0^{n-1} - Y_0^{n-1}), \tag{5.6}$$

we say that $\mathbf{Y}$ *causes* $\mathbf{X}$ *locally*, denoted by $\mathbf{Y} \overset{\mathrm{LC}}{\Longrightarrow} \mathbf{X}$. Otherwise, we say that $\mathbf{Y}$ *does not cause* $\mathbf{X}$ *locally*, denoted by $\mathbf{Y} \overset{\mathrm{LC}}{\not\Longrightarrow} \mathbf{X}$.

**Definition 5.2** (instantaneous local causality) If there exits $n \in \{0, \cdots,$ $N\}$ such that

$$\sigma(X(n)|I_0^{n-1}, Y(n)) < \sigma(X(n)|I_0^{n-1}), \tag{5.7}$$

we say that *instantaneous local causality of* $\mathbf{Y}$ *to* $\mathbf{X}$ *occurs*, denoted by $\mathbf{Y} \overset{\mathrm{ILC}}{\Longrightarrow} \mathbf{X}$. Otherwise, we say that *instantaneous local causality of* $\mathbf{Y}$ *to* $\mathbf{X}$ *does not occur*, denoted by $\mathbf{Y} \overset{\mathrm{ILC}}{\not\Longrightarrow} \mathbf{X}$.

Here, we discuss only the case when $I_n = \{X_i(n), Y_j(n);\ 1 \le i \le d_1,$ $1 \le j \le d_2\}$ $(0 \le n \le N)$. Hence, it follows that $I_0^{n-1} - Y_0^{n-1} = \{X_i(m);$ $1 \le i \le d_1,\ 0 \le m \le n-1\}$ $(0 \le n \le N)$.

Let $n \in \{0, \cdots, N\}$. We get the following $\mathrm{KM_2O}$-Langevin equations:

$$X(n) = P_{\mathbf{M}_0^{n-1}(\mathbf{X})} X(n) + \nu_{+X}(n) \tag{5.8}$$

$$Z(n) = P_{\mathbf{M}_0^{n-1}(\mathbf{Z})} Z(n) + \nu_{+Z}(n). \tag{5.9}$$

Here

$$\nu_{+Z}(n) = \begin{pmatrix} \nu_{+Z,X}(n) \\ \nu_{+Z,Y}(n) \end{pmatrix} \tag{5.10}$$

is defined by

$$\nu_{+Z,X}(n) = X(n) - P_{\mathbf{M}_0^{n-1}(\mathbf{Z})} X(n) \tag{5.11}$$

$$\nu_{+Z,Y}(n) = Y(n) - P_{\mathbf{M}_0^{n-1}(\mathbf{Z})} Y(n). \tag{5.12}$$

The following Lemma 5.1 is clear.

**Lemma 5.1**

$$\epsilon(X(n)|I_0^{n-1} - Y_0^{n-1}) = \nu_{+X}(n) \tag{5.13}$$

$$\epsilon(X(n)|I_0^{n-1}) = \nu_{+Z,X}(n). \tag{5.14}$$

Now, we have

**Lemma 5.2** $\mathbf{Y} \overset{\mathrm{LC}}{\not\Longrightarrow} \mathbf{X}$ *if and only if*

$$\nu_{+X}(n) = \nu_{+Z,X}(n) \qquad (0 \le n \le N). \tag{5.15}$$

*Proof.* We assume that $\mathbf{Y} \overset{\mathrm{LC}}{\not\Longrightarrow} \mathbf{X}$. Then,

$$\begin{aligned} \|\nu_{+X}(n) \ - \ & \nu_{+Z,X}(n)\|^2 \\ &= \langle \nu_{+X}(n) - \nu_{+Z,X}(n),\ \nu_{+X}(n) - \nu_{+Z,X}(n)\rangle \\ &= \langle \nu_{+X}(n),\ \nu_{+X}(n)\rangle + \langle \nu_{+Z,X}(n),\ \nu_{+Z,X}(n)\rangle \\ &\quad -2\langle \nu_{+Z,X}(n),\ \nu_{+X}(n)\rangle. \end{aligned} \tag{5.16}$$

From (5.13) and (5.14),

$$\|\nu_{+X}(n)\|^2 = \|\nu_{+Z,X}(n)\|^2. \tag{5.17}$$

Since $\nu_{+X}(n) - \nu_{+Z,X}(n) \in \mathbf{M}_0^{n-1}(\mathbf{Z})$,

$$\langle \nu_{+Z,X}(n),\ \nu_{+X}(n) - \nu_{+Z,X}(n)\rangle = 0. \tag{5.18}$$

Hence

$$\begin{aligned}\|\nu_{+X}(n) \ - \ \nu_{+Z,X}(n)\|^2 &= 2\|\nu_{+Z,X}(n)\|^2 \\ &\quad - \ 2\langle \nu_{+Z,X}(n),\ \nu_{+X}(n) - \nu_{+Z,X}(n) + \nu_{+Z,X}(n)\rangle \\ &= 2\|\nu_{+Z,X}(n)\|^2 - 2\|\nu_{+Z,X}(n)\|^2 \\ &= 0.\end{aligned} \tag{5.19}$$

Therefore, we get $\nu_{+X}(n) = \nu_{+Z,X}(n)$. It is clear to prove the sufficient condition. $\square$

Let us consider how to characterize (5.15). For $n \in \{0, \cdots, N\}$, we set

$$V_{+X}(n) = E\nu_{+X}(n)^t\nu_{+X}(n) \tag{5.20}$$

$$V_{+Z,X}(n) = E\nu_{+Z,X}(n)^t\nu_{+Z,X}(n) \tag{5.21}$$

$$V_{+Z}(n) = E\nu_{+Z}(n)^t\nu_{+Z}(n). \tag{5.22}$$

$V_{+X}(n)$, $V_{+Z,X}(n)$ and $V_{+Z}(n)$ are covariance matrices of $\nu_{+X}(n)$, $\nu_{+Z,X}(n)$ and $\nu_{+Z}(n)$ respectively. Furthermore, let $W_{+X}(n)$, $W_{+Z,X}(n)$ and $W_{+Z}(n)$ be lower triangular matrices such that

$$V_{+X}(n) = W_{+X}(n)^tW_{+X}(n) \tag{5.23}$$

$$V_{+Z,X}(n) = W_{+Z,X}(n)^tW_{+Z,X}(n) \tag{5.24}$$

$$V_{+Z}(n) = W_{+Z}(n)^tW_{+Z}(n). \tag{5.25}$$

Then we get the following Theorem 5.1.

**Theorem 5.1** *The necessary and sufficient condition of* $\mathbf{Y} \overset{\mathrm{LC}}{\not\Longrightarrow} \mathbf{X}$ *is that either of the following* (L-1), (L-2) *or* (L-3) *holds. Here,* $(I)_d$ *denotes the d-dimensional identity matrix.*

(L-1) $W_{+Z,X}(n)^{-1}\nu_{+X}(n)$, $0 \le n \le N$ *is a* $d_1$*-dimensional white noise with mean* $\mathbf{0}$ *and covariance matrix* $(I)_{d_1}$.

(L-2) $W_{+X}(n)^{-1}\nu_{+Z,X}(n)$, $0 \le n \le N$ *is a* $d_1$*-dimensional white noise with mean* $\mathbf{0}$ *and covariance matrix* $(I)_{d_1}$.

(L-3) $W_{+Z}(n)^{-1}\begin{pmatrix} \nu_{+X}(n) \\ \nu_{+Z,Y}(n) \end{pmatrix}$, $0 \le n \le N$ *is a* $(d_1+d_2)$*-dimensional white noise with mean* $\mathbf{0}$ *and covariance matrix* $(I)_{(d_1+d_2)}$.

*Proof.* We prove that (L-3) is equivalent to $\mathbf{Y} \overset{\text{LC}}{\not\Longrightarrow} \mathbf{X}$. We assume that (L-3) holds. Then, we obtain

$$\begin{aligned} &W_{+Z}(n)^{-1}E\begin{pmatrix} \nu_{+X}(n) \\ \nu_{+Z,Y}(n) \end{pmatrix}^t \begin{pmatrix} \nu_{+X}(n) \\ \nu_{+Z,Y}(n) \end{pmatrix}^t W_{+Z}(n)^{-1} \\ &\qquad = (I)_{(d_1+d_2)}. \end{aligned} \tag{5.26}$$

Therefore

$$\begin{aligned} & E\begin{pmatrix} \nu_{+X}(n)^t\nu_{+X}(n) & \nu_{+X}(n)^t\nu_{+Z,Y}(n) \\ \nu_{+Z,Y}(n)^t\nu_{+X}(n) & \nu_{+Z,Y}(n)^t\nu_{+Z,Y}(n) \end{pmatrix} \\ = \; & E\begin{pmatrix} \nu_{+Z,X}(n)^t\nu_{+Z,X}(n) & \nu_{+Z,X}(n)^t\nu_{+Z,Y}(n) \\ \nu_{+Z,Y}(n)^t\nu_{+Z,X}(n) & \nu_{+Z,Y}(n)^t\nu_{+Z,Y}(n) \end{pmatrix}. \end{aligned} \tag{5.27}$$

From the special case of (5.27),

$$E\nu_{+X}(n)^t\nu_{+X}(n) = E\nu_{+Z,X}(n)^t\nu_{+Z,X}(n). \tag{5.28}$$

Setting

$$\nu_{+X}(n) = \begin{pmatrix} \nu_{+X,1}(n) \\ \vdots \\ \nu_{+X,d_1}(n) \end{pmatrix}, \quad \nu_{+Z,X}(n) = \begin{pmatrix} \nu_{+Z,X,1}(n) \\ \vdots \\ \nu_{+Z,X,d_1}(n) \end{pmatrix}, \tag{5.29}$$

we have $E\nu_{+X,i}(n)^2 = E\nu_{+Z,X,i}(n)^2$ $(1 \le i \le d_1)$. As shown in Lemma 5.2, we get $\nu_{+X,i}(n) = \nu_{+Z,X,i}(n)$ $(1 \le i \le d_1)$. The proof of the converse is clear. □

Sims [33] introduced a distributed lag model as a model of Granger's causal relation between AR models. Here we get the following Theorem 5.2 which shows the structure of $\mathbf{Y} \overset{\text{LC}}{\not\Longrightarrow} \mathbf{X}$.

**Theorem 5.2** *The necessary and sufficient condition of* $\mathbf{Y} \overset{\text{LC}}{\not\Longrightarrow} \mathbf{X}$ *is as follows: there exist the unique matrices* $\{A(n,k) \in M(d_2 \times d_1; \mathbf{R}),\ 0 \le k \le n \le N\}$, *and appropriate matrices* $\{B(n,k) \in M(d_2 \times d_2; \mathbf{R}),\ 0 \le k \le n \le$

$N\}$, *such that for any* $n \in \{0, \cdots, N\}$, $Y(n)$ *is expressed as*

$$Y(n) = \sum_{k=0}^{n} A(n,k)X(k) + \sum_{k=0}^{n} B(n,k)\nu^{\star}(k). \tag{5.30}$$

*Here,*

$$B(n,n) \in GL(d_2; \mathbf{R}) \tag{5.31}$$

(5.32) $(\nu^{\star}(n); 0 \le n \le N)$ *is orthogonal to* $\mathbf{M}_0^N(\mathbf{X})$, *and is a* $d_2$*-dimensional white noise with mean* $\mathbf{0}$ *and covariance matrix* $(I)_{d_2}$

(5.33) $\nu^{\star}(n)$ *is orthogonal to* $\mathbf{M}_0^{n-1}(\mathbf{Y})$, *for* $n = 1, \cdots, n-1$.

*Proof.* We assume $\mathbf{Y} \overset{\mathrm{LC}}{\not\Longrightarrow} \mathbf{X}$. For $n \in \{0, \cdots, N\}$, let $W_{+Z}(n)$ be a lower triangular matrix defined by (5.25). We set

$$\begin{aligned} \nu^{*}(n) &= W_{+Z}(n)^{-1}\nu_{+Z}(n) \\ &= W_{+Z}(n)^{-1}\begin{pmatrix} \nu_{+Z,X}(n) \\ \nu_{+Z,Y}(n) \end{pmatrix} = \begin{pmatrix} \nu_1^{*}(n) \\ \nu_2^{*}(n) \end{pmatrix}. \end{aligned} \tag{5.34}$$

Let $d = d_1 + d_2$. Then, $\nu^{*}(n)$, $0 \le n \le N$ is a $d$-dimensional white noise with mean $\mathbf{0}$ and covariance matrix $(I)_d$. We define $\mathrm{KM_2O}$-Langevin equations of $\mathbf{X}$ and $\mathbf{Z}$ as follows:

$$X(n) = -\sum_{k=0}^{n-1} \gamma_{+X}(n,k)X(k) + \nu_{+X}(n) \tag{5.35}$$

$$\begin{aligned} \begin{pmatrix} X(n) \\ Y(n) \end{pmatrix} &= -\sum_{k=0}^{n-1} \gamma_{+Z}(n,k) \begin{pmatrix} X(k) \\ Y(k) \end{pmatrix} \\ &\quad + \begin{pmatrix} \nu_{+Z,X}(n) \\ \nu_{+Z,Y}(n) \end{pmatrix}. \end{aligned} \tag{5.36}$$

Multiplying $W_{+Z}(n)^{-1}$ from the left-hand side of (5.36), we get

$$\begin{aligned} W_{+Z}(n)^{-1}\begin{pmatrix} X(n) \\ Y(n) \end{pmatrix} &= -\sum_{k=0}^{n-1} W_{+Z}(n)^{-1} \\ &\times \gamma_{+Z}(n,k)\begin{pmatrix} X(k) \\ Y(k) \end{pmatrix} + \begin{pmatrix} \nu_1^{*}(n) \\ \nu_2^{*}(n) \end{pmatrix}. \end{aligned} \tag{5.37}$$

We set a lower triangular matrix $W_{+Z}(n)$ by

$$W_{+Z}(n)=\begin{pmatrix} B_{11}(n) & \mathbf{0} \\ B_{21}(n) & B_{22}(n) \end{pmatrix}, \tag{5.38}$$

where $B_{ij}(n)\in M(d_i\times d_j;\mathbf{R})$ $(i,j=1,2)$, $B_{12}(n)=\mathbf{0}$, and $B_{ii}(n)\in GL(d_i;\mathbf{R})$ $(i=1,2)$. $W_{+Z}(n)^{-1}$ is also a lower triangular matrix defined by

$$W_{+Z}(n)^{-1}=\begin{pmatrix} C_{11}(n) & \mathbf{0} \\ C_{21}(n) & C_{22}(n) \end{pmatrix}, \tag{5.39}$$

where $C_{ij}(n)\in M(d_i\times d_j;\mathbf{R})$ $(i,j=1,2)$, $C_{12}(n)=\mathbf{0}$, and $C_{ii}(n)\in GL(d_i;\mathbf{R})$ $(i=1,2)$. Moreover, we set

$$W_{+Z}(n)^{-1}\gamma_{+Z}(n,k)=\begin{pmatrix} \Gamma_{11}(n,k) & \Gamma_{12}(n,k) \\ \Gamma_{21}(n,k) & \Gamma_{22}(n,k) \end{pmatrix}. \tag{5.40}$$

Then, (5.37) leads to

$$\begin{aligned} &C_{21}(n)X(n)+C_{22}(n)Y(n) \\ &=-\sum_{k=0}^{n-1}\Gamma_{21}(n,k)X(k)-\sum_{k=0}^{n-1}\Gamma_{22}(n,k)Y(k)+\nu_2^*(n). \end{aligned} \tag{5.41}$$

Since $C_{22}(n)\in GL(d_2;\mathbf{R})$,

$$\begin{aligned} Y(n)&=-C_{22}(n)^{-1}C_{21}(n)X(n) \\ &-\sum_{k=0}^{n-1}C_{22}(n)^{-1}\Gamma_{21}(n,k)X(k)-\sum_{k=0}^{n-1}C_{22}(n)^{-1}\Gamma_{22}(n,k)Y(k) \\ &+C_{22}(n)^{-1}\nu_2^*(n). \end{aligned} \tag{5.42}$$

Any component of $\nu_2^*(n)$ is a linear combination of components of $\nu_{+Z}(n)$. Therefore,

$$EX(k)^t\nu_2^*(n)=\mathbf{0} \qquad (0\le k\le n-1). \tag{5.43}$$

Moreover, any component of $\nu_{+Z,X}(n)$ is expressed as a linear combination of components of $\nu_1^*(n)$. This shows

$$E\nu_{+Z,X}(n)^t\nu_2^*(n)=\mathbf{0}. \tag{5.44}$$

From the definition of $\mathbf{M}_0^{n-1}(\mathbf{Z})$, $E\{P_{\mathbf{M}_0^{n-1}(\mathbf{Z})}X(n)\}^t\nu_2^*(n)=\mathbf{0}$.

The KM$_2$O-Langevin equation shows

$$X(n) = P_{\mathbf{M}_0^{n-1}(\mathbf{Z})}X(n) + \nu_{+Z,X}(n). \tag{5.45}$$

Hence, we get

$$EX(n)^t\nu_2^*(n) = \mathbf{0}. \tag{5.46}$$

Similarly to (5.45),

$$X(n+1) = P_{\mathbf{M}_0^{n}(\mathbf{X})}X(n+1) + \nu_{+X}(n+1). \tag{5.47}$$

We have $\nu_{+X}(n+1) = \nu_{+Z,X}(n+1)$ from the assumption. Hence $E\nu_{+X}(n+1)^t\nu_2^*(n) = \mathbf{0}$. Therefore we have

$$EX(n+1)^t\nu_2^*(n) = \mathbf{0}. \tag{5.48}$$

Inductively, we get

$$EX(k)^t\nu_2^*(n) = \mathbf{0} \qquad (0 \le k \le N). \tag{5.49}$$

Substituting $Y(k)$, $0 \le k \le n-1$ in (5.41) inductively, we have the expression of (5.30).

Let us show that $\{A(n,k),\ 0 \le k \le n \le N\}$ is unique. If $Y(n)$ has another expression form such as

$$Y(n) = \sum_{k=0}^{n} \widetilde{A}(n,k)X(k) + \sum_{k=0}^{n} \widetilde{B}(n,k)\eta^\star(k), \tag{5.50}$$

we multiply ${}^tX(m)$ $(0 \le m \le n)$ to (5.30) and (5.50) from the right-hand side. Taking their expectations, we get

$$\sum_{k=0}^{n} A(n,k)R(k-m) = \sum_{k=0}^{n} \widetilde{A}(n,k)R(k-m). \tag{5.51}$$

The Toeplitz condition of §2 suggests

$$A(n,k) = \widetilde{A}(n,k) \qquad (0 \le k \le n). \tag{5.52}$$

Let us prove the converse. We set

$$\gamma_{+Z}(n,k) = \begin{pmatrix} \gamma_{11}(n,k) & \gamma_{12}(n,k) \\ \gamma_{21}(n,k) & \gamma_{22}(n,k) \end{pmatrix} \qquad (0 \le k < n \le N), \tag{5.53}$$

where $\gamma_{ij}(n,k) \in M(d_i \times d_j;\, \mathbf{R})$ $(i,j=1,2)$. From (5.36), we have

$$\begin{aligned} X(n) &= -\sum_{k=0}^{n-1} \gamma_{11}(n,k)X(k) \\ &\quad -\sum_{k=0}^{n-1} \gamma_{12}(n,k)Y(k) + \nu_{+Z,X}(n). \end{aligned} \tag{5.54}$$

We substitute $Y(k)$, $0 \le k \le n-1$ of (5.30) in (5.54). Then there exist appropriate matrices $C(n,k) \in M(d_1;\mathbf{R})$ and $D(n,k) \in M(d_1;\mathbf{R})$ such that $X(n)$ is expressed as

$$\begin{aligned} X(n) &= -\sum_{k=0}^{n-1} C(n,k)X(k) \\ &\quad -\sum_{k=0}^{n-1} D(n,k)\nu^\star(k) + \nu_{+Z,X}(n). \end{aligned} \tag{5.55}$$

Since $B(k,k) \in GL(d_2;\mathbf{R})$, any components of $\nu^\star(k)$ belong to $\mathbf{M}_0^k(\mathbf{Z})$. Hence

$$E\nu_{+Z,X}(n)^t\nu^\star(k) = \mathbf{0} \qquad (0 \le k \le n-1). \tag{5.56}$$

The assumption leads to

$$EX(l)^t\nu^\star(k) = 0 \qquad (0 \le l \le N). \tag{5.57}$$

Multiplying ${}^t\nu^\star(k)$ $(0 \le k \le n-1)$ to the right-hand side of (5.55) and taking its expectation, we get

$$D(n,k) = \mathbf{0} \qquad (0 \le k \le n-1). \tag{5.58}$$

Hence, we have

$$\nu_{+Z,X}(n) = \nu_{+X}(n). \tag{5.59}$$

This shows $\mathbf{Y} \overset{\mathrm{LC}}{\not\Longrightarrow} \mathbf{X}$. □

*Remark* 5.7 In (5.30) and (5.51), the following holds.

$$B(n,k)\nu^\star(k) = \widetilde{B}(n,k)\eta^\star(k),\ 0 \le k \le n \le N \tag{5.60}$$

$$B(n,k)^tB(n,k) = \widetilde{B}(n,k)^t\widetilde{B}(n,k),\ 0 \le k \le n \le N. \tag{5.61}$$

We can show (5.60) as follows: Since $\sum_{k=0}^{n} B(n,k)\nu^\star(k)$ $= \sum_{k=0}^{n} \widetilde{B}(n,k)\eta^\star(k)$,

$$\sum_{k=0}^{n-1} B(n,k)\nu^\star(k) - \sum_{k=0}^{n-1} \widetilde{B}(n,k)\eta^\star(k) = B(n,n)\nu^\star(n) - \widetilde{B}(n,n)\eta^\star(n). \tag{5.62}$$

Components of the left-hand side of the above equation belong to $\mathbf{M}_0^{n-1}(\mathbf{Z})$. On the other hand, $B(n,n)\nu^\star(n) - \widetilde{B}(n,n)\eta^\star(n)$ is orthogonal to $\mathbf{M}_0^{n-1}(\mathbf{Z})$. Hence we get (5.60). It is easy to get (5.61) from (5.60).

Similarly to (5.35), we get that

$$Y(n) = -\sum_{k=0}^{n-1} \gamma_{+Y}(n,k)Y(k) + \nu_{+Y}(n). \tag{5.63}$$

*Example* 5.1 Let $d_1 = d_2 = 1$ and $Y(n) = aX(n) + w(n)$ $(a \neq 0)$. Here $w(n)$, $0 \le n \le N$ is a white noise with variance $\sigma^2$, and independent of $X(k)$, $0 \le n \le N$. Let the covariance function of $\mathbf{X}$ be $\{R(n); |n| \le N\}$ such that $R(1) \neq 0$. From Theorem 5.2, $\mathbf{Y} \overset{\mathrm{LC}}{\not\Longrightarrow} \mathbf{X}$. We get that $\nu_{+Z,Y}(1) = Y(1) - aR(1)X(0)/R(0)$ and $\nu_{+Y}(1) = Y(1) - a^2R(1)Y(0)/(a^2R(0) + \sigma^2)$. Therefore, we have $\nu_{+Z,Y}(1) \not\equiv \nu_{+Y}(1)$. This leads to $\mathbf{X} \overset{\mathrm{LC}}{\Longrightarrow} \mathbf{Y}$.

**Theorem 5.3** *The necessary and sufficient condition of* $\mathbf{Y} \overset{\mathrm{LC}}{\not\Longrightarrow} \mathbf{X}$ *and* $\mathbf{X} \overset{\mathrm{LC}}{\not\Longrightarrow} \mathbf{Y}$ *is*

$$E\nu_{+X}(n)^t\nu_{+Y}(m) = \mathbf{0} \qquad n \neq m,\ n,m \in \{0,\cdots,N\}. \tag{5.64}$$

*Proof.* If $\mathbf{Y} \overset{\mathrm{LC}}{\not\Longrightarrow} \mathbf{X}$ and $\mathbf{X} \overset{\mathrm{LC}}{\not\Longrightarrow} \mathbf{Y}$, $\nu_{+X}(n) = \nu_{+Z,X}(n), \nu_{+Y}(n) = \nu_{+Z,Y}(n)$, for each $n \in \{0,\cdots,N\}$. Hence, (5.64) holds.

We show the converse. For each $m \in \{0,\cdots,N\}$, $Y(m)$ is a linear combination of $\nu_{+Y}(0),\cdots,\nu_{+Y}(m)$. Therefore we have

$$E\nu_{+X}(n)^tY(m) = \mathbf{0} \qquad (0 \le m \le n-1). \tag{5.65}$$

It is clear that $E\nu_{+Z,X}(n)^tY(m) = \mathbf{0}$ $(0 \le m \le n-1)$. On the other hand, $\nu_{+X}(n) - \nu_{+Z,X}(n)$ is a linear combination of $X(0)$, $\cdots$, $X(n-1)$, $Y(0)$,

$\cdots$, $Y(n-1)$. Hence, we get $\|\nu_{+X}(n) - \nu_{+Z,X}(n)\|^2 = 0$. This shows that $\nu_{+X}(n) - \nu_{+Z,X}(n) = \mathbf{0}$. Similarly, we have $\nu_{+Y}(n) - \nu_{+Z,Y}(n) = \mathbf{0}$. □

It is easy to get the following Corollary 5.1.

**Corollary 5.1** *If $EX(0)^tY(0) = \mathbf{0}$, the necessary and sufficient condition of $\mathbf{Y} \overset{\text{LC}}{\not\Longrightarrow} \mathbf{X}$ and $\mathbf{X} \overset{\text{LC}}{\not\Longrightarrow} \mathbf{Y}$ is*

$$EX(n)^tY(m) = \mathbf{0} \qquad n, m \in \{0, \cdots, N\}. \tag{5.66}$$

Corollary 5.1 shows that the conception of causality is more universal than the conception of correlation.

Let us now consider the instantaneous local causality.

**Theorem 5.4** *The necessary and sufficient condition that instantaneous local causality of $\mathbf{X}$ to $\mathbf{Y}$ does not occur is*

$$C_{21}(n) = \mathbf{0} \qquad (0 \leq n \leq N). \tag{5.67}$$

*where $C_{21}(n)$ is defined by* (5.39).

*Proof.* (5.36) and (5.41) show that the necessary and sufficient condition of $\mathbf{X} \overset{\text{ILC}}{\not\Longrightarrow} \mathbf{Y}$ is

$$\|\nu_{+Z,Y}(n)\|^2 = \|C_{22}(n)^{-1}\nu_2^*(n)\|^2 \qquad (0 \leq n \leq N). \tag{5.68}$$

From (5.37), we get

$$\begin{aligned}\begin{pmatrix} \nu_1^*(n) \\ \nu_2^*(n) \end{pmatrix} &= \begin{pmatrix} C_{11}(n) & \mathbf{0} \\ C_{21}(n) & C_{22}(n) \end{pmatrix} \begin{pmatrix} \nu_{+Z,X}(n) \\ \nu_{+Z,Y}(n) \end{pmatrix} \\ &= \begin{pmatrix} C_{11}(n)\nu_{+Z,X}(n) \\ C_{21}(n)\nu_{+Z,X}(n) + C_{22}(n)\nu_{+Z,Y}(n) \end{pmatrix}. \end{aligned} \tag{5.69}$$

Hence

$$\begin{aligned} &\|C_{22}(n)^{-1}\nu_2^*(n)\|^2 \\ &\quad = \|C_{22}(n)^{-1}C_{21}(n)\nu_{+Z,X}(n) + \nu_{+Z,Y}(n)\|^2 \\ &\quad = \|C_{22}(n)^{-1}C_{21}(n)\nu_{+Z,X}(n)\|^2 + \|\nu_{+Z,Y}(n)\|^2 \\ &\quad + 2E^t(C_{22}(n)^{-1}C_{21}(n)\nu_{+Z,X}(n))\nu_{+Z,Y}(n). \end{aligned} \tag{5.70}$$

Since $C_{11}(n) \in GL(d_1; \mathbf{R})$, we obtain from (5.69)

$$E\nu_{+Z,X}(n)^t(C_{21}(n)\nu_{+Z,X}(n) + C_{22}(n)\nu_{+Z,Y}(n)) = \mathbf{0}. \tag{5.71}$$

We shall indicate by $trA$ the trace of a matrix $A$. Now, we have from (5.71)

$$\begin{aligned} & E^t(C_{22}(n)^{-1}C_{21}(n)\nu_{+Z,X}(n))\nu_{+Z,Y}(n) \\ & = trE\nu_{+Z,Y}(n)^t(C_{22}(n)^{-1}C_{21}(n)\nu_{+Z,X}(n)) \\ & = trEC_{22}(n)^{-1}(C_{22}(n)\nu_{+Z,Y}(n) + C_{21}(n)\nu_{+Z,X}(n) \\ & \quad -C_{21}(n)\nu_{+Z,X}(n)) \times {}^t(C_{22}(n)^{-1}C_{21}(n)\nu_{+Z,X}(n)) \\ & = -trE(C_{22}(n)^{-1}C_{21}(n)\nu_{+Z,X}(n))^t(C_{22}(n)^{-1}C_{21}(n)\nu_{+Z,X}(n)) \\ & = -E^t(C_{22}(n)^{-1}C_{21}(n)\nu_{+Z,X}(n))(C_{22}(n)^{-1}C_{21}(n)\nu_{+Z,X}(n)) \\ & = -\|C_{22}(n)^{-1}C_{21}(n)\nu_{+Z,X}(n)\|^2. \end{aligned} \tag{5.72}$$

Hence,we get

$$\begin{aligned} \|C_{22}(n)^{-1}\nu_2^*(n)\|^2 &= \|\nu_{+Z,Y}(n)\|^2 \\ & \quad -\|C_{22}(n)^{-1}C_{21}(n)\nu_{+Z,X}(n)\|^2. \end{aligned} \tag{5.73}$$

This shows that the necessary and sufficient condition of $\mathbf{X} \stackrel{\text{ILC}}{\not\Longrightarrow} \mathbf{Y}$ is

$$\|C_{22}(n)^{-1}C_{21}(n)\nu_{+Z,X}(n)\|^2 = 0. \tag{5.74}$$

Therefore, we have

$$C_{22}(n)^{-1}C_{21}(n)\nu_{+Z,X}(n) = \mathbf{0}. \tag{5.75}$$

This leads to

$$C_{22}(n)^{-1}C_{21}(n)V_{+Z,X}(n) = \mathbf{0}. \tag{5.76}$$

Since $C_{22}(n) \in GL(d_2; \mathbf{R})$, and $V_{+Z,X}(n) \in GL(d_1; \mathbf{R})$, we have $C_{21}(n) = \mathbf{0}$. □

For $n \in \{0, \cdots, N\}$ and $i, j \in \{1, \cdots, d\}$, let $V_{+Z,ij}(n)$ be the $(i, j)$ component of covariance matrix $V_{+Z}(n)$. Then we obtain:

**Corollary 5.2** *The necessary and sufficient condition that instantaneous local causality of* $\mathbf{X}$ *to* $\mathbf{Y}$ *does not occur is*

$$\begin{aligned} & V_{+Z,ij}(n) = 0 \\ & \quad (d_1 + 1 \le i \le d_1 + d_2,\ 1 \le j \le d_1,\ 0 \le n \le N). \end{aligned} \tag{5.77}$$

*Proof.* The necessary and sufficient condition that (5.67) holds is

$$B_{21}(n) = \mathbf{0} \qquad (0 \le n \le N). \tag{5.78}$$

It is clear that (5.78) is equivalent to (5.77). □

It is easy to get the following Corollary 5.3.

**Corollary 5.3** *The necessary and sufficient condition that instantaneous local causality of* $\mathbf{X}$ *to* $\mathbf{Y}$ *does not occur is*

$$E\nu_{+Z,Y}(n)^t\nu_{+Z,X}(n) = \mathbf{0} \qquad (0 \le n \le N). \tag{5.79}$$

Corollary 5.3 shows that $\mathbf{Y} \overset{\text{ILC}}{\not\Longrightarrow} \mathbf{X}$ and $\mathbf{X} \overset{\text{ILC}}{\not\Longrightarrow} \mathbf{Y}$ are equivalent to each other.

Now we characterize $\mathbf{X} \overset{\text{ILC}}{\not\Longrightarrow} \mathbf{Y}$.

**Theorem 5.5** *The necessary and sufficient condition that instantaneous local causality of* $\mathbf{X}$ *to* $\mathbf{Y}$ *does not occur is that*

$$\begin{pmatrix} C_{11}(n) & \mathbf{0} \\ \mathbf{0} & C_{22}(n) \end{pmatrix} \nu_{+Z}(n), \quad 0 \le n \le N \tag{5.80}$$

*is a white noise with mean* $\mathbf{0}$ *and covariance matrix* $(I)_d$.

*Proof.* The necessity is clear from Corollary 5.2. We show the sufficiency. We set

$$C(n) = \begin{pmatrix} C_{11}(n) & \mathbf{0} \\ \mathbf{0} & C_{22}(n) \end{pmatrix}. \tag{5.81}$$

Since

$$V_{+Z}(n) = C(n)^{-1} \times {}^tC(n)^{-1}, \tag{5.82}$$

we obtain (5.77). □

## 6. Data analysis of local causality

We introduce the Local Causal Test for the analysis of local causality and the Instantaneous Local Causal Test for the analysis of instantaneous local causality. These tests are applied to the data of ${}^t(M_2 + \text{CD, RGNP})$.

Let $\mathcal{X} = (\mathcal{X}(n); 0 \le n \le N)$ be $d_1$-dimensional data, and $\mathcal{Y} = (\mathcal{Y}(n);$

$0 \leq n \leq N)$ $d_2$-dimensional data. Set $d = d_1 + d_2$. We construct the $d$-dimensional data $\mathcal{Z} = (\begin{pmatrix} \mathcal{X}(n) \\ \mathcal{Y}(n) \end{pmatrix}; 0 \leq n \leq N)$. Following §3, we construct the sample KM$_2$O-Langevin equations and other related quantities. These quantities are represented by replacing $\mathbf{X}$ with $\mathcal{X}$, $\mathbf{Y}$ with $\mathcal{Y}$, and $\mathbf{Z}$ with $\mathcal{Z}$ respectively in the variables defined in §5. At first we apply Test(S) to data $\mathcal{Z}$. If Test(S) for $\mathcal{Z}$ is accepted, we proceed to test whether $\mathbf{Y} \stackrel{\mathrm{LC}}{\Longrightarrow} \mathbf{X}$ or not.

Let $M$ be defined by (3.19). For each $i \in \{0, \cdots, N-M\}$, we introduce the shifted data $\mathcal{Z}_i$, $\mathcal{X}_i$ and $\mathcal{Y}_i$ similarly to (3.20). Then we get the sample KM$_2$O-Langevin fluctuation force $\begin{pmatrix} \nu_{+i,\mathcal{Z},\mathcal{X}}(n) \\ \nu_{+i,\mathcal{Z},\mathcal{Y}}(n) \end{pmatrix}$ $(0 \leq n \leq M)$ of $\mathcal{Z}_i$, the sample KM$_2$O-Langevin fluctuation force $\nu_{+i,\mathcal{X}}(n)$ $(0 \leq n \leq M)$ of $\mathcal{X}_i$, the sample KM$_2$O-Langevin fluctuation force $\nu_{+i,\mathcal{Y}}(n)(0 \leq n \leq M)$ of $\mathcal{Y}_i$ respectively. In Test(S) for $\mathcal{Z}$, by replacing the component $\nu_{+i,\mathcal{Z},\mathcal{X}}(n)$ of $\begin{pmatrix} \nu_{+i,\mathcal{Z},\mathcal{X}}(n) \\ \nu_{+i,\mathcal{Z},\mathcal{Y}}(n) \end{pmatrix}$ by $\nu_{+i,\mathcal{X}}(n)$ for all $n \in \{0, \cdots, M\}$, we get

$$W_{+\mathcal{Z}}(n)^{-1} \begin{pmatrix} \nu_{+i,\mathcal{X}}(n) \\ \nu_{+i,\mathcal{Z},\mathcal{Y}}(n) \end{pmatrix}, \quad 0 \leq n \leq M. \tag{6.1$_i$}$$

Test(S) for the one-dimensional data constructed from $(6.1)_i$ is called the LC$_1$ Test. If the LC$_1$ Test is accepted (resp. not accepted), we can find from (L-3) of Theorem 5.1 that $\mathbf{Y} \stackrel{\mathrm{LC}}{\not\Longrightarrow} \mathbf{X}$ (or $\mathbf{Y} \stackrel{\mathrm{LC}}{\Longrightarrow} \mathbf{X}$). Similarly, Test(S) for the one-dimensional data constructed from

$$W_{+\mathcal{Z}}(n)^{-1} \begin{pmatrix} \nu_{+i,\mathcal{Z},\mathcal{X}}(n) \\ \nu_{+i,\mathcal{Y}}(n) \end{pmatrix}, \quad 0 \leq n \leq M \tag{6.2$_i$}$$

is called the LC$_2$ Test. If the LC$_2$ Test is accepted (or not accepted), $\mathbf{X} \stackrel{\mathrm{LC}}{\not\Longrightarrow} \mathbf{Y}$ (or $\mathbf{X} \stackrel{\mathrm{LC}}{\Longrightarrow} \mathbf{Y}$). The Local Causal Test (LC Test) is the general term for the LC$_1$ Test and the LC$_2$ Test.

Similarly to §4, we apply these tests to the second order differences in the quarterly data of ${}^t(M_2 + \mathrm{CD}, \mathrm{RGNP})$ in two periods from 1965 to 1987 and from 1965 to 1990.

"S" (or "NS") indicates that the LC Test is accepted (or not accepted). Table 6.1 and Table 6.2 report that RGNP locally causes Money Supply in both periods and Money Supply locally causes RGNP in the period from 1965 to 1987. Now, it is an established theory that Money Supply and

Table 6.1 The LC Test for ${}^t(\mathrm{M_2{+}CD},\ \mathrm{RGNP})$ from 1965 to 1987.

| $j$ | LC Test | (M) | (V) | (O) | (S) |
|---|---|---|---|---|---|
| 2 | $\mathrm{LC}_1$ | 0.974 | 0.649 | 0.948 | NS |
| 2 | $\mathrm{LC}_2$ | 0.987 | 0.675 | 0.987 | NS |

Table 6.2 The LC Test for ${}^t(\mathrm{M_2{+}CD},\ \mathrm{RGNP})$ from 1965 to 1990.

| $j$ | LC Test | (M) | (V) | (O) | (S) |
|---|---|---|---|---|---|
| 2 | $\mathrm{LC}_1$ | 0.977 | 0.693 | 0.931 | NS |
| 2 | $\mathrm{LC}_2$ | 0.954 | 0.736 | 0.886 | S |

RGNP are mutually related. On the other hand, we can not accept that Money Supply locally causes RGNP in the period from 1965 to 1990. We can explain these phenomena as follows: During the three years from 1987 to 1989, Japan went through a so-called bubble economy. The anomalous increase in the Money Supply is an example of this phenomenon. However this increase was nominal and failed substantially to boost RGNP.

Figure 6.1 and Figure 6.2 illustrate the local causal relations between Money Supply and RGNP.

Let us now compare the LC Test with the Granger-Sargent Test. (4.7) shows that the Granger-Sargent Test could not accept $\mathrm{M_2 + CD} \overset{\mathrm{GC}}{\Longrightarrow} \mathrm{RGNP}$ in the period from 1965 to 1987. On the other hand, the LC Test accepts that $\mathrm{M_2 + CD} \overset{\mathrm{LC}}{\Longrightarrow} \mathrm{RGNP}$ in the same period. As can be seen from the data analysis above, we can assert the efficiency of the LC Test.

Secondly, we consider how to test $\mathbf{X} \overset{\mathrm{ILC}}{\not\Longleftrightarrow} \mathbf{Y}$. We set

$$W_{+\mathcal{Z}}(n)^{-1} = \begin{pmatrix} C_{11}(n) & \mathbf{0} \\ C_{12}(n) & C_{22}(n) \end{pmatrix}, \quad 0 \le n \le M. \tag{6.3}$$

Let $i \in \{0, \cdots, N-M\}$. Similarly to the LC Test, Test(S) for the one-

$$\boxed{\mathrm{M_2 + CD}} \overset{\Longleftarrow}{\Longrightarrow} \boxed{\mathrm{RGNP}}$$

Figure 6.1 Local causality from 1965 to 1987.

$$\boxed{\mathrm{M_2 + CD}} \overset{\Longleftarrow}{\not\Longrightarrow} \boxed{\mathrm{RGNP}}$$

Figure 6.2 Local causality from 1965 to 1990.

Table 6.3 The ILC Test for ${}^t(\mathrm{M}_2 + \mathrm{CD}, \mathrm{RGNP})$.

| period | $j$ | Test | (M) | (V) | (O) | (S) |
|---|---|---|---|---|---|---|
| 1965–1987 | 2 | ILC | 0.974 | 0.662 | 0.961 | NS |
| 1965–1990 | 2 | ILC | 0.954 | 0.704 | 0.909 | S |

$$\boxed{\mathrm{M}_2 + \mathrm{CD}} \overset{\Leftarrow}{\Longrightarrow} \boxed{\mathrm{RGNP}}$$

Figure 6.3 Instantaneous local causality from 1965 to 1987.

$$\boxed{\mathrm{M}_2 + \mathrm{CD}} \overset{\not\Leftarrow}{\not\Longrightarrow} \boxed{\mathrm{RGNP}}$$

Figure 6.4 Instantaneous local causality from 1965 to 1990.

dimensional data constructed from

$$\begin{pmatrix} C_{11}(n) & \mathbf{0} \\ \mathbf{0} & C_{22}(n) \end{pmatrix} \begin{pmatrix} \nu_{+i,\mathcal{Z},\mathcal{X}}(n) \\ \nu_{+i,\mathcal{Z},\mathcal{Y}}(n) \end{pmatrix}, \quad 0 \le n \le M \tag{6.4}$$

is called the Instantaneous Local Causal Test (ILC Test). If the ILC Test is accepted (or not accepted), we can find from Theorem 5.5 that $\mathbf{X} \overset{\text{ILC}}{\not\Longleftrightarrow} \mathbf{Y}$ (or $\mathbf{X} \overset{\text{ILC}}{\Longrightarrow} \mathbf{Y}$). Applications of the ILC Test to the second differences of quarterly data of ${}^t(\mathrm{M}_2 + \mathrm{CD}, \mathrm{RGNP})$ are shown in Table 6.3, Figure 6.3 and Figure 6.4.

Similarly to Table 6.1 and Table 6.2, "S" (or "NS") indicates that the test is accepted (or not accepted). These results are illustrated in Figure 6.3 and Figure 6.4.

## Appendix

Table A Quarterly data of Japan Money Supply $M_2 + CD$ from 1965 to 1990, (unit: one billion yen, source: Databank, Toyokeizai Company, 1976, 1991).

| Year | I | II | III | IV |
|---|---|---|---|---|
| 1965 | 21678 | 22398 | 23376 | 25394 |
| 1966 | 25687 | 26354 | 27501 | 29522 |
| 1967 | 29731 | 30528 | 31660 | 34097 |
| 1968 | 34169 | 35482 | 36301 | 39153 |
| 1969 | 39435 | 41346 | 42817 | 46399 |
| 1970 | 46612 | 48810 | 50285 | 54237 |
| 1971 | 55002 | 58845 | 61908 | 67398 |
| 1972 | 68224 | 72260 | 75533 | 84040 |
| 1973 | 85346 | 90134 | 92831 | 98188 |
| 1974 | 98235 | 102159 | 102908 | 109494 |
| 1975 | 109374 | 113823 | 116458 | 125330 |
| 1976 | 126234 | 132192 | 134881 | 142248 |
| 1977 | 142350 | 147143 | 148910 | 158033 |
| 1978 | 157331 | 165076 | 167461 | 178720 |
| 1979 | 177587 | 184497 | 187794 | 195012 |
| 1980 | 194734 | 200250 | 199238 | 208985 |
| 1981 | 208097 | 217791 | 219200 | 232041 |
| 1982 | 230485 | 238028 | 240229 | 250466 |
| 1983 | 247926 | 255752 | 257236 | 268692 |
| 1984 | 267172 | 274751 | 279321 | 289714 |
| 1985 | 291609 | 299756 | 298904 | 314938 |
| 1986 | 313893 | 323076 | 323060 | 343887 |
| 1987 | 341860 | 355366 | 358173 | 380867 |
| 1988 | 380995 | 394455 | 398179 | 419732 |
| 1989 | 419615 | 433606 | 438292 | 470020 |
| 1990 | 472477 | 488055 | 495901 | 504972 |

Table B Quarterly data of RGNP from 1965 to 1990,
(unit: one billion yen, source: Databank, Toyokeizai Company, 1991).

| Year | I | II | III | IV |
|---|---|---|---|---|
| 1965 | 98209.69 | 99735.81 | 101896.29 | 103098.64 |
| 1966 | 106076.01 | 110631.11 | 113185.28 | 115268.23 |
| 1967 | 118371.29 | 121498.60 | 125714.28 | 128152.62 |
| 1968 | 131073.26 | 135839.40 | 138949.09 | 147732.80 |
| 1969 | 148577.10 | 152773.07 | 156197.76 | 162995.67 |
| 1970 | 167713.04 | 169192.18 | 174021.33 | 173813.68 |
| 1971 | 174559.62 | 177414.66 | 180425.90 | 182278.55 |
| 1972 | 187336.36 | 191269.92 | 195155.95 | 200124.06 |
| 1973 | 206472.57 | 208376.62 | 208501.38 | 210108.22 |
| 1974 | 204711.41 | 206395.56 | 208692.68 | 207357.88 |
| 1975 | 207093.75 | 212430.81 | 214653.93 | 217128.78 |
| 1976 | 218960.03 | 220760.27 | 223468.66 | 224092.79 |
| 1977 | 229443.96 | 231294.76 | 232687.24 | 236124.81 |
| 1978 | 239969.33 | 241920.21 | 245332.34 | 248396.84 |
| 1979 | 252587.19 | 256803.72 | 259272.22 | 261244.60 |
| 1980 | 264987.60 | 264164.02 | 266706.46 | 269937.72 |
| 1981 | 273731.49 | 273943.75 | 276643.88 | 277837.72 |
| 1982 | 280889.91 | 284347.12 | 285907.10 | 288785.51 |
| 1983 | 289804.67 | 290870.27 | 295221.30 | 296067.13 |
| 1984 | 301134.08 | 305081.63 | 306776.56 | 309458.94 |
| 1985 | 314832.56 | 320582.12 | 322546.73 | 327069.65 |
| 1986 | 324254.63 | 329177.45 | 330809.87 | 335300.05 |
| 1987 | 337484.68 | 339459.30 | 345743.11 | 353281.28 |
| 1988 | 359179.54 | 361860.14 | 368643.16 | 373095.36 |
| 1989 | 376977.24 | 376976.68 | 386060.58 | 391148.04 |
| 1990 | 397447.82 | 402877.40 | 407432.80 | 409539.20 |

**Acknowledgments** The author wishes to express his sincere thanks to Professor Okabe of Hokkaido University. Without his advice, guidance and encouragement, I could not have completed this paper. The author is grateful to Professor Katayama of Shiga University for his comments on economics. The author would like to express his hearty thanks to the referee for his kind comments.

## References


[1] Akaike H. and Nakagawa T., *Statistical analysis and control for a dynamical system.* Science Company, 1973 (in Japanese).

[2] Box G.E.P. and Jenkins G.M., *Time series analysis, forecasting and control.* Holden-Day, 1970.

[3] Caines P.E., Keng C.W. and Sethi S.P., *Causality analysis and multivariate autoregressive modelling with an application to super-market sales analysis.* J. Eco. Dynam. Control, **3** (1981), 267–298.

[4] Deniau C.I., Fiori G. and Mathias A., *A V.A.R. model for the public dept: The french case.* Document de travail GREQE. **8807** (1988).

[5] Fuller W.A., *Introduction to statistical time series.* John Wily and Sons, 1976.

[6] Geweke J., *Comparing alternative tests of causality in temporal systems: Analytic results and experimental evidence.* J. Econometrics, Feb., (1983), 161–194.

[7] Granger C.W.J., *Investigating causal relations by econometric models and cross-spectral methods.* Econometrica, **37**, No. 3 (1969), 424–438.

[8] Granger C.W.J. and Newbold P., *Spurious regressions in econometrics.* J. Econometrics, **2** (1974), 111–120.

[9] Hsiao H., *Autoregressive modelling and money-income causality detection.* Journal of Monetary Economics, **7** (1981), 85–106.

[10] Komura T., *Japanese Economics and Financial Policy.* Toyokeizai Company, 1986 (in Japanese).

[11] Levinson N., *The Wiener RMS error criterion in filter design and prediction.* J. Math. Phys., **25** (1947), 261–278.

[12] Ljung G.M. and Box G.E.P., *On a measure of lack of fit in time series models.* Biometrica, **65**, 2 (1978), 297–303.

[13] Mata C.G. and Shaffner F.I., *Solar and economic relationships; a preliminary report.* The Quarterly Journal of Economics, **49** (1935), 1–51.

[14] Mehra Y.P., *Money wages, prices, and causality.* Journal of Political Economy, **85**, No. 6 (1977), 1227–1244.

[15] Nakano Y. and Okabe Y., *On a multi-dimensional $[\alpha,\beta,\gamma]$-Langevin equation.* Proc. Japan Acad., **59** (1983), 171–173.

[16] Nakano Y., *Statistical study of economic time series: Okabe's theory and its applications.* Shigadaigaku Sosyo, **20**, Koomura Press, 1991 (in Japanese).

[17] Okabe Y., *On stochastic difference equations for the multi-dimensional weakly stationary time series.* Prospect of Algebraic Analysis (ed. by M. Kashiwara and T. Kawai), Academic Press, **2** (1988), 601–645.

[18] Okabe Y., *On the stationarity, causality and anomality in random phenomena.* Preprint of special lecture in Japanese Mathematical Society, 1988 (in Japanese).

[19] Okabe,Y., *On the theory of $KM_2O$-Langevin equations and quantum statistics.* Hokkaido Univ. Tech. Repo. Ser. No. 9 (1988), 58–64.

[20] Okabe Y., *The theory of non-linear prediction and causal analysis.* System / Control / Information, Vol. **33**, No. 9 (1989), 478–487 (in Japanese).

[21] Okabe Y., *Natural science and stationary stochastic processes.* Surikagaku, Science Company, No. 340 (1991), 21–28 (in Japanese).

[22] Okabe Y., *On a Langevin equation.* Sugaku, Iwanami, **33** (1991), 306–324 (in Japanese).

[23] Okabe Y., *A new algorithm driven from the view-point of the fluctuation-dissipation theorem in the theory of $KM_2O$-Langevin equations.* Hokkaido University Preprint Series in Mathematics, No. 166 (1992).

[24] Okabe Y., *Application of the theory of $KM_2O$-Langevin equations to the linear prediction problem for the multi-dimensional weakly stationary time series.* J. Math. Soc. Japan, **45**, No. 2 (1993), 277–294.

[25] Okabe Y. and Inoue A., *The theory of $KM_2O$-Langevin equations and its applications to data analysis (II): Causal analysis.* Nagoya Math. J. **134** (1994), 1–28.

[26] Okabe Y. and Nakano Y., *The theory of $KM_2O$-Langevin equations and its applications to data analysis (I) : Stationary analysis.* Hokkaido Mathematical Journal, **20** (1991), 45–90.

[27] Okabe Y. and Ootsuka T., *Application of the theory of $KM_2O$-Langevin equations to the non-linear prediction problem for the one-dimensional strictly stationary time series.* to appear in J. Math. Soc. Japan.

[28] Okabe Y. and Ootsuka T., *The theory of $KM_2O$-Langevin equations and its applications to data analysis (III): Prediction analysis.* in preparation.

[29] Pierce D.A. and Haugh L.D., *Causality in temporal systems: characterizations and a survey.* Journal of Econometrics, **5** (1977), 265–293.

[30] Ram R., *Money, income and causality in Japan-supplementary evidence: comment.* Southern Economic Journal **50** (April 1984), 1214–1218.

[31] Sargent T.J., *A classical macroeconometric model for the United States.* Journal of Political Economy, **84**, No. 2 (1976), 207–237.

[32] Sawa T., *On an effectiveness of macro econometric models.* New development of econometrics (ed. by H. Takeuchi), Tokyo University Publisher, 1983, 265–278 (in Japanese).

[33] Sims C.A., *Money, income, and causality.* American Economic Review, **62** (1972), 540–552.

[34] Wold H., *Causality and econometrics.* Econometrica, **22** (1954), 162–177.

Faculty of Economics
Shiga University
1-1-1 Banba, Hikone
Shiga 522, Japan
E-mail: nakano@biwakoshiga-u.ac.jp