

**THE THEORY OF KM_2O -LANGEVIN EQUATIONS
AND APPLICATIONS TO DATA ANALYSIS (II):
CAUSAL ANALYSIS (1)**

YASUNORI OKABE AND AKIHIKO INOUE

Dedicated to Professor H. P. McKean, Jr. on his sixtieth birthday

§1. Introduction

It is not too much to say that the problem of finding a cause-and-effect relationship is a fascinating and eternal theme in both natural and social sciences. It is often difficult to decide whether one is the cause of another in two related phenomena, but it is an important problem. It is related to the internal structure of phenomena which generate deterministic or random changes as time passes. We note that the phenomena to be considered are often not deterministic but random. For example, in physical systems such as quantum mechanics or chaotic classical mechanics, it is well known that certain probabilistic reasonings are indispensable.

In several areas in natural science, particularly in physics, causal relations are often discovered by experiments and later explained in theory. However, also in the fields where such methods can not be applied, there sometimes occurs the necessity of deciding a causal direction, or deciding which is the cause and which is the effect in two related phenomena. For example, it has been asserted that Wolfer's sunspot numbers are positively related to the numbers of Lynx and to the economy in Canada ([6], [9], [10], [11], [35], [39]).

Almost all researches in time series analysis have been done on the basis of AR -time series, which has both weak stationarity and finite multiple Markovian property. In particular, such properties have been assumed explicitly or implicitly. However, since we wish to find true intercourse between pure and applied sciences, we should not assume such conditions as those are not yet checked by data analysis, particularly in the problem of finding causal directions. It is desirable

Received December 18, 1990.

and even necessary to decide the causal direction in two related phenomena on the basis of computation algorithms of mathematical model-free theory.

As a continuation to the previous papers ([28], [31]), this paper aims to propose in §1 **Test(CS)** for deciding the causal direction between given two phenomena, based upon the theory of KM_2O -Langevin equations. Hence our theory consists of three parts. The first is a theoretical part ([28]) where we developed the theory of KM_2O -Langevin equations describing the time evolution of multi-dimensional local time series with weak stationarity. By a local time series, we mean a time series whose time parameter space consists of finite interval in $\mathbf{Z}$. The second is a practical part ([31]) where we proposed **Test(S)** for checking weak stationarity of random phenomena. As the third part, we shall propose in this paper **Test(CS)**, which determines the causal direction between two kinds of random phenomena. We would like to note that in the background of both tests, there have been attempts to understand mathematically the philosophy behind the so-called fluctuation-dissipation theorem in statistical physics ([7], [12]-[29]).

Thus, our strategy in the causal analysis is as follows: Given two kinds of data which represent two different random phenomena, respectively.

- (i) We construct a multi-dimensional time series with the two given data as its components and then apply **Test(S)** to it.
- (ii) If it passes **Test(S)**, then we apply **Test(CS)** to the two given data.
- (iii) If it does not pass **Test(S)**, then we have to find a bijective transformation such that the new transformed time series passes **Test(S)**. If it can be done, we apply **Test(CS)** to the new time series.

Now let us state the contents of the present paper. In §2, we recall briefly the theory of KM_2O -Langevin equation developed in [28] and [31]. From the point of view of canonical representation for stationary processes, we characterize in §3 the class of AR -time series in the framework of generalized forward AR -Langevin equation, which can be obtained by rewriting the forward KM_2O -Langevin equation. This generalized forward AR -Langevin equation will play an important role in the prediction problem for stationary time series ([30], [32]). Another reason why we take up AR -time series is a practical one; we would like to reinforce the experiments for **Test(S)** proposed in [31], where we have used only the data with white noise property. To this end, we construct in §4 some one-dimensional AR -time series according to the characterization theorem in §3 and then apply **Test(S)** to some dynamical systems obtained by transforming them in several ways.

Following the concept of causal representation for stationary processes, we

give in §5 the definition of causality which is quantified by introducing a causality function, and then propose **Test(CS)** for determining the causal direction between two kinds of random phenomena.

In the final two sections §6 and §7, **Test(CS)** is applied to concrete data. We consider in §6 the dynamical systems constructed in §4 whose weak stationarity are checked in §4. In §7, at first, we treat the two-dimensional data composed of Wolfer's sunspot numbers and trapped numbers of Canadian Lynx in MacKenzie River from 1821 to 1934 whose weak stationarity has been already checked in [31]. Finally we investigate the mutual causal relation among Wolfer's sunspot numbers, the rainfall averages in a year in Sapporo and the temperature averages in a year in Sapporo from 1889 to 1988.

The authors would like to thank the referee for his valuable and constructive advices.

§2. **KM₂O-Langevin equations**

For any fixed natural numbers d and N , let $\mathbf{X} = (X(n); |n| \leq N)$ be an $\mathbf{R}^d$ -valued weakly stationary local time series on a probability space $(\Omega, \mathcal{B}, P)$ with expectation vector zero and covariance matrix function R :

$$(2.1) \quad R(m-n) \equiv E(X(m)^t X(n)) \quad (m, n \in \mathbf{Z}, |m| \leq N, |n| \leq N).$$

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

$$(2.2) \quad S_n = \begin{pmatrix} R(0) & R(1) & \cdots & R(n-1) \\ R(-1) & R(0) & \cdots & R(n-2) \\ \vdots & \vdots & \ddots & \vdots \\ R(-(n-1)) & R(-(n-2)) & \cdots & R(0) \end{pmatrix}.$$

In the sequel we treat the case where the following condition (2.3) holds:

$$(2.3) \quad S_n \in GL(nd; \mathbf{R}) \text{ for any } n \in \{1, \dots, N\}.$$

For any d -dimensional time series $\mathbf{Y} = ({}^t(Y_1(n), \dots, Y_d(n)); l \leq n \leq r)$ on the probability space $(\Omega, \mathcal{B}, P) (-\infty \leq l < r \leq \infty)$, we define, for each n_1, n_2 , $l \leq n_1 \leq n_2 \leq r$, the closed subspace $\mathbf{M}_{n_1}^{n_2}(\mathbf{Y})$ of $L^2(\Omega, \mathcal{B}, P)$ by

$$(2.4) \quad \mathbf{M}_{n_1}^{n_2}(\mathbf{Y}) = \text{the closed linear hull of } \{Y_j(m); 1 \leq j \leq d, n_1 \leq m \leq n_2\}.$$

Then we introduce two d -dimensional time series $\nu_+ = (\nu_+(n); 0 \leq n \leq N)$ and $\nu_- = (\nu_-(-n); 0 \leq n \leq N)$ by

$$(2.5) \quad \nu_+(n) = X(n) - P_{\mathbf{M}_0^{n-1}(\mathbf{X})}X(n) \text{ and } \nu_-(-n) = X(-n) - P_{\mathbf{M}_{-n+1}^0(\mathbf{X})}X(-n),$$

where $\mathbf{M}_0^{-1}(\mathbf{Y}) = \mathbf{M}_1^0(\mathbf{X}) = \{0\}$ and $P_{\mathbf{M}_0^{n-1}(\mathbf{X})}$ (resp. $P_{\mathbf{M}_{-n+1}^0(\mathbf{X})}$) stands for the orthogonal projection down to the space $\mathbf{M}_0^{n-1}(\mathbf{X})$ (resp. $\mathbf{M}_{-n+1}^0(\mathbf{X})$). In particular, it holds that

$$(2.6) \quad \nu_+(0) = \nu_-(0) = X(0)$$

$$(2.7) \quad \nu_+ \text{ and } \nu_- \text{ are both orthogonal time series with mean vector zero}$$

$$(2.8) \quad \mathbf{M}_0^n(\mathbf{X}) = \mathbf{M}_0^n(\nu_+) \text{ and } \mathbf{M}_{-n}^0(\mathbf{X}) = \mathbf{M}_{-n}^0(\nu_-) \quad (0 \leq n \leq N).$$

The covariance matrices of $\nu_\pm(\pm n)$ are denoted by $V_\pm(n)$ ($0 \leq n \leq N$):

$$(2.9) \quad V_+(n) = E(\nu_+(n)^t \nu_+(n)) \text{ and } V_-(n) = E(\nu_-(-n)^t \nu_-(-n)).$$

Then we have

DECOMPOSITION THEOREM ([28], [31]). *There exists a unique system $\{\gamma_+(n, k), \gamma_-(n, k), \delta_+(m), \delta_-(m); 1 \leq k < n \leq N, 1 \leq m \leq N\}$ of members in $M(d; \mathbf{R})$ such that for any $n \in \{1, \dots, N\}$,*

$$(2.10_\pm) \quad X(\pm n) = - \sum_{k=1}^{n-1} \gamma_\pm(n, k)X(\pm k) - \delta_\pm(n)X(0) + \nu_\pm(\pm n).$$

We call equation (2.10₊) (resp. (2.10₋)) the *forward* (resp. *backward*) KM_2O -Langevin equation for $\mathbf{X}$. The random force ν_+ (resp. ν_-) is said to be the *forward* (resp. *backward*) KM_2O -Langevin force associated with $\mathbf{X}$. Moreover, we designate the system $\{\gamma_+(n, k), \gamma_-(n, k), \delta_+(m), \delta_-(m), V_+(l), V_-(l); 1 \leq k < n \leq N, 1 \leq m \leq N, 0 \leq l \leq N\}$ as the KM_2O -Langevin data associated with the covariance matrix function R of $\mathbf{X}$.

We have the following fluctuation-dissipation theorem which relates the fluctuating (or random) part in (2.10_±) with their dissipating (or deterministic) part in (2.10_±):

FLUCTUATION-DISSIPATION THEOREM ([8], [3], [37], [38], [28], [31]). *For any $n, k \in \mathbf{N}, 1 \leq k < n \leq N$,*

$$(2.11_\pm) \quad \gamma_\pm(n, k) = \gamma_\pm(n-1, k-1) + \delta_\pm(n)\gamma_\mp(n-1, n-k-1)$$

$$(2.12_\pm) \quad V_\pm(n) = (I - \delta_\pm(n)\delta_\mp(n))V_\pm(n-1)$$

$$(2.13) \quad \delta_-(n)V_+(n-1) = V_-(n-1)^t \delta_+(n)$$

$$(2.14) \quad \delta_-(n)V_+(n) = V_-(n)^t \delta_+(n),$$

where

$$(2.15) \quad \gamma_+(n, 0) = \delta_+(n) \quad \text{and} \quad \gamma_-(n, 0) = \delta_-(n).$$

Since $V_{\pm}(n)$ are positive definite, we have the following algorithm by which the fundamental quantities $\delta_{\pm}(\cdot)$ can be calculated from the covariance matrix function R :

ALGORITHM ([8], [3], [37], [38], [28], [31]). For any $n \in \mathbf{N}$, $1 \leq n \leq N$,

$$(2.16_{\pm}) \quad \delta_{\pm}(n) = - (R(\pm n) + \sum_{k=0}^{n-2} \gamma_{\pm}(n-1, k) R(\pm(k+1))) V_{\mp}(n-1)^{-1}.$$

Conversely, let us be given a symmetric positive definite matrix $V \in GL(d; \mathbf{R})$ and N elements $\delta_+(n)$ ($1 \leq n \leq N$) in $M(d; \mathbf{R})$. Then, with the relations (2.12_±)–(2.13) in the fluctuation-dissipation theorem in mind, we can construct a sequence of triples $(V_+(n), \delta_-(n), V_-(n))$ ($1 \leq n \leq N$) by

$$(2.17) \quad \begin{cases} V_+(0) = V_-(0) = V \\ V_+(n) = V_+(n-1) - \delta_+(n) V_-(n-1)^t \delta_+(n) \\ \delta_-(n) V_+(n-1) = V_-(n-1)^t \delta_+(n) \\ V_-(n) = V_-(n-1) - \delta_-(n) V_+(n-1)^t \delta_-(n). \end{cases}$$

In so doing, we assume that

$$(2.18) \quad V_+(n) \text{ and } V_-(n) \text{ are positive definite } (1 \leq n \leq N-1).$$

Furthermore, we construct a system $\{\gamma_+(m, n), \gamma_-(m, n); 0 \leq n < m \leq N\}$ of members in $M(d; \mathbf{R})$ according to the relations (2.11_±) in the fluctuation-dissipation theorem. Finally, we prepare a d -dimensional time series $\nu_+ = (\nu_+(n); 0 \leq n \leq N)$ on a probability space $(\Omega, \mathcal{B}, P)$ such that for any $m, n \in \mathbf{N}^*$, $0 \leq m, n \leq N$,

$$(2.19) \quad E(\nu_+(n)) = 0 \quad \text{and} \quad E(\nu_+(m)^t \nu_+(n)) = \delta_{mn} V_+(n)$$

and construct a d -dimensional time series $\mathbf{X}_+ = (X(n); 0 \leq n \leq N)$ by the forward KM_2O -Langevin equation (2.10₊) with (2.6). Then we have

RECONSTRUCTION THEOREM ([28], [31]). $\mathbf{X}_+$ is a weakly stationary local time series.

§3. A generalized AR -Langevin equation

Let $\mathbf{X} = (X(n) ; n \in \mathbf{Z})$ be an $\mathbf{R}^d$ -valued weakly stationary time series on a probability space $(\mathcal{Q}, \mathcal{B}, P)$ with mean vector zero and covariance matrix function R satisfying

$$(3.1) \quad R(0) \in GL(d ; \mathbf{R}).$$

3.1. For a fixed $K \in \mathbf{N}$, we say that $\mathbf{X}$ is an $AR(K)$ -time series if there exist a system $\{A(k), \sigma ; 1 \leq k \leq K\}$ of members in $M(d ; \mathbf{R})$ and an $\mathbf{R}^d$ -valued time series $\xi_+ = (\xi_+(n) ; n \in \mathbf{Z})$ such that

$$(3.2) \quad X(n) = \sum_{k=1}^K A(k)X(n-k) + \sigma \xi_+(n) \quad (n \in \mathbf{Z})$$

$$(3.3) \quad \sigma \in GL(d ; \mathbf{R})$$

$$(3.4) \quad E(\xi_+(n)) = 0 \quad \text{and} \quad E(\xi_+(m)^t \xi_+(n)) = \delta_{mn} I \quad (m, n \in \mathbf{Z})$$

$$(3.5) \quad \mathbf{M}_{-\infty}^n(\mathbf{X}) = \mathbf{M}_{-\infty}^n(\xi_+) \quad (n \in \mathbf{Z}).$$

We call equation (3.2) a *forward $AR(K)$ -Langevin equation* ([2]). The condition (3.4) shows that ξ_+ is a standard white noise in the wide sense. Furthermore, condition (3.5) means that there exists a causal relation between $\mathbf{X}$ and ξ_+ , which will be discussed in §5.

We treat the case where $\mathbf{X}$ is an $AR(K)$ -time series for some $K \in \mathbf{N}$.

LEMMA 3.1 (Yule-Walker equation).

$$(3.6) \quad R(0) = \sum_{k=1}^K A(k)^t R(k) + \sigma^t \sigma$$

$$(3.7) \quad R(n) = \sum_{k=1}^K A(k) R(n-k) \quad \text{for any } n \in \mathbf{N}.$$

Proof. Multiplying both sides of equation (3.2) by ${}^t \xi_+(n)$ and then taking an expectation with respect to the probability P , we see from (3.4) and (3.5) that

$$(3.8) \quad E(X(n)^t \xi_+(n)) = E(\xi_+(n)^t \xi_+(n)) = \sigma^t \sigma \quad (n \in \mathbf{Z}).$$

Therefore, multiplying both sides of equation (3.2) by ${}^t X(n)$ and ${}^t X(0)$ and then taking an expectation with respect to P , we see from (3.4) and (3.5) that (3.6) and (3.7) hold. Q.E.D.

Concerning Toeplitz condition (2.3), we have

LEMMA 3.2. **X** satisfies condition (2.3) for any $N \in \mathbf{N}$, that is,

$$(3.9) \quad S_n \in GL(d; \mathbf{R}) \quad \text{for any } n \in \mathbf{N}.$$

Proof. By putting $A(n) = 0$ ($n \geq K + 1$), it follows from (3.6) and (3.7) that for any $n \geq K$,

$$\begin{pmatrix} I & -A(1) & \cdots & -A(n) \\ & I & & \\ & & \ddots & 0 \\ & 0 & & I \end{pmatrix} S_{n+1} = \begin{pmatrix} \sigma^t \sigma & 0 & \cdots & 0 \\ {}^t R(1) & & & \\ \vdots & S_n & & \\ {}^t R(n) & & & \end{pmatrix},$$

which, together with (3.1) and (3.3), yields (3.9).

Q.E.D.

3.2. The purpose of this subsection is to characterize the $AR(K)$ -property within the framework of the theory of KM_2O -Langevin equations. For that purpose, we consider the case where **X** satisfies condition (3.9). We then can apply the results in §2 (cf. [28]) to obtain the forward KM_2O -Langevin equation

$$(3.10) \quad X(0) = \nu_+(0)$$

$$(3.11) \quad X(m) = - \sum_{k=0}^{m-1} \gamma_+(m, k) X(k) + \nu_+(m) \quad (m \in \mathbf{N}).$$

By using the unitary group $(U(n); n \in \mathbf{Z})$ on the Hilbert space $\mathbf{M}_{-\infty}^{+\infty}(\mathbf{X})$ such that

$$(3.12) \quad U(n)(X_j(m)) = X_j(m+n) \quad (1 \leq j \leq d, m \in \mathbf{Z}),$$

we define, for each $K \in \mathbf{N}$, an $\mathbf{R}^d$ -valued time series $\xi_+^{(K)} = (\xi_+^{(K)}(n); n \in \mathbf{Z})$ by

$$(3.13) \quad \xi_+^{(K)}(n) = V_+(K)^{-1/2} U(n-K) \nu_+(K).$$

Immediately from (2.9) and (3.13), we obtain

LEMMA 3.3. For each $K \in \mathbf{N}$,

- (i) $\xi_+^{(K)}$ is weakly stationary
- (ii) $E\{\xi_+^{(K)}(n) {}^t \xi_+^{(K)}(n)\} = I$ ($n \in \mathbf{Z}$)
- (iii) $\xi_+^{(K)}(K) = V_+(K)^{-1/2} \nu_+(K)$.

Moreover, for each fixed $K \in \mathbf{N}$, operating $U(n - K)$ to both sides of equation (3.11) with $\mathbf{m}$ replaced by K , we obtain

$$(3.14) \quad X(n) = - \sum_{k=1}^K \gamma_+(K, K - k) X(n - k) + V_+(K)^{1/2} \xi_+^{(K)}(n) \quad (n \in \mathbf{Z}).$$

We note that $\xi_+^{(K)}$ does not always satisfy both the orthogonal condition (3.4) and the causal condition (3.5). By taking (3.2) and (3.14) into account we call equation (3.14) a *generalized forward AR(K)-Langevin equation* associated with $\mathbf{X}$.

We are now in a position to show

THEOREM 3.1. *Let $\mathbf{X}$ be a weakly stationary time series satisfying condition (3.1). Then, for any fixed $K \in \mathbf{N}$, the following three conditions are equivalent:*

- (i) $\mathbf{X}$ is an $AR(K)$ -time series.
- (ii) $\mathbf{X}$ satisfies condition (3.9) and $\delta_+(n) = 0$ for any $n \geq K + 1$.
- (iii) $\mathbf{X}$ satisfies condition (3.9) and $\xi_+^{(K)}$ is a standard white noise in the wide sense satisfying the causal relation;

$$(3.15) \quad \mathbf{M}_{-\infty}^n(\mathbf{X}) = \mathbf{M}_{-\infty}^n(\xi_+^{(K)}) \quad \text{for any } n \in \mathbf{Z}.$$

Proof. (i) $\rightarrow$ (ii); Since $\mathbf{X}$ can be regarded as an $AR(K + 1)$ -time series with $A(K + 1) = 0$, it follows from Lemma 3.1 that

$$(A(1) \cdots A(K)0) = (R(K + 1) \cdots R(1)) \begin{pmatrix} R(K) & R(K - 1) & \cdots & R(0) \\ R(K - 1) & R(K - 2) & \cdots & R(-1) \\ \vdots & & \ddots & \vdots \\ R(0) & R(-1) & \cdots & R(-K) \end{pmatrix}^{-1}.$$

On the other hand, multiplying both sides of equation (3.11) with $\mathbf{m} = K + 1$ by $X(k)$ ($0 \leq k \leq K + 1$) and then taking the expectation with respect to P , we find from (2.7) and (2.8) that

$$\begin{aligned} & (\gamma_+(K + 1, K) \cdots \gamma_+(K + 1, 0)) \\ &= (R(K + 1) \cdots R(1)) \begin{pmatrix} R(K) & R(K - 1) & \cdots & R(0) \\ R(K - 1) & R(K - 2) & \cdots & R(-1) \\ \vdots & & \ddots & \vdots \\ R(0) & R(-1) & \cdots & R(-K) \end{pmatrix}^{-1}. \end{aligned}$$

Hence, it holds that $\delta_+(K + 1) \equiv \gamma_+(K + 1, 0) = 0$. Similarly, we can show (ii).

In order to prove the part: (ii) $\rightarrow$ (iii), set

$$(3.16) \quad A(k) = -\gamma_+(K, K-k) \quad \text{for } 1 \leq k \leq K.$$

Then it can be shown from (2.11₊) and (ii) that

$$(3.17) \quad \gamma_+(n, n-k) = \begin{cases} -A(k) & \text{for any } n \geq K+1, 1 \leq k \leq K \\ 0 & \text{for any } n \geq K+1, K+1 \leq k \leq n. \end{cases}$$

Next we shall show that $\xi_+^{(K)}$ is a white noise in the wide sense. By Lemma 3.3, it suffices to prove

$$E\{\xi_+^{(K)}(n_1) {}^t\xi_+^{(K)}(n_2)\} = 0 \quad \text{for any } n_1, n_2 \in \mathbf{Z}, n_1 > n_2.$$

Furthermore, it can be seen from (2.7) and (3.13) that we have only to show

$$(3.18) \quad U(p)\nu_+(K) = \nu_+(K+p) \quad \text{for any } p \in \mathbf{N}.$$

By operating the unitary operator $U(p)$ to both sides of equation (3.11) with m replaced by K , we have

$$U(p)\nu_+(K) = X(K+p) + \sum_{k=1}^K \gamma_+(K, K-k)X(K-k+p).$$

Since it follows from (2.11₊), (3.11) and (3.17) that

$$X(n) = \sum_{k=1}^K A(k)X(n-k) + \nu_+(n) \quad \text{for any } n \geq K+1,$$

(3.18) holds.

Finally we shall show (3.15). It is easy to see from (3.14) that

$$(3.19) \quad \mathbf{M}_{-\infty}^n(\mathbf{X}) \supset \mathbf{M}_{-\infty}^n(\xi_+^{(K)}) \quad \text{for any } n \in \mathbf{Z}.$$

By using a resolution of the identity $\{E(\theta); \theta \in (-\pi, \pi]\}$ such that

$$(3.20) \quad U(n) = \int_{(-\pi, \pi]} e^{-in\theta} dE(\theta) \quad (n \in \mathbf{Z}),$$

we see from (3.12) and (3.13) that

$$(3.21) \quad X(n) = \int_{(-\pi, \pi]} e^{-in\theta} dE(\theta)X(0)$$

$$(3.22) \quad \xi_+^{(K)}(n) = \int_{(-\pi, \pi]} e^{-in\theta} dE(\theta)\xi_+^{(K)}(0).$$

Set

$$(3.23) \quad P(z) = I - \sum_{k=1}^K A(k)z^k \quad (z \in \mathbf{C}).$$

Denoting by $\Delta(d\theta)$ the spectral measure matrix of the weakly stationary time series $\mathbf{X}$, we can see from (3.13), (3.14) and (3.21)–(3.23) that

$$(3.24) \quad P(e^{i\theta})\Delta(d\theta)^t \overline{P(e^{i\theta})} = V_+(K)d\theta.$$

By using Lebesgue's decomposition theorem for bounded signed measures, we can see that

$$(3.25) \quad P(e^{i\theta}) \in GL(d; \mathbf{C}) \quad \text{a.e. } \theta \in (-\pi, \pi]$$

$$(3.26) \quad \Delta(d\theta) = P(e^{i\theta})^{-1} V_+(K) \overline{P(e^{i\theta})}^{-1} d\theta.$$

Furthermore, we claim

$$(3.27) \quad \mathbf{M}_{-\infty}^{+\infty}(\mathbf{X}) = \mathbf{M}_{-\infty}^{+\infty}(\xi_+^{(K)}).$$

By virtue of (3.19), it suffices to show that $Y = 0$ for any $Y \in \mathbf{M}_{-\infty}^{+\infty}(\mathbf{X}) \ominus \mathbf{M}_{-\infty}^{+\infty}(\xi_+^{(K)})$. There exists $f = (f_1, \dots, f_d) \in L^2((-\pi, \pi]; \Delta(d\theta))$ such that $Y = \int_{(-\pi, \pi]} f(\theta) dE(\theta) X(0)$. Since Y is orthogonal to all vectors $\xi_+^{(K)}(n)$ ($n \in \mathbf{Z}$), we have

$$\int_{(-\pi, \pi]} e^{-in\theta} V_+(K)^{-1/2} P(e^{i\theta}) \Delta(d\theta)^t \overline{f(\theta)} = 0 \quad \text{for any } n \in \mathbf{Z}.$$

By the uniqueness of Fourier transformation, we find that

$$V_+(K)^{-1/2} P(e^{i\theta}) \Delta(d\theta)^t \overline{f(\theta)} = 0,$$

which, together with (3.25) and (3.26), implies that $f = 0$ and so $Y = 0$.

Moreover, we claim

$$(3.28) \quad \mathbf{M}_{-\infty}^n(\mathbf{X}) \text{ is orthogonal to } \{\xi_{+j}^{(K)}(l); 1 \leq j \leq d, l > n\} \text{ for any } n \in \mathbf{Z}.$$

It suffices to prove that for any $k \leq n < l$, $E\{X(k)^t \xi_+^{(K)}(l)\} = 0$. By (3.13),

$$E\{X(k)^t \xi_+^{(K)}(l)\} = E\{X(K+k-l)^t \nu_+(K)\} V_+(K)^{-1/2}.$$

By virtue of (3.18), we can see from (2.7) and (2.8) that $E\{X(K+k-l)^t \nu_+(K)\} = 0$ and so (3.28) holds.

We are now in a position to prove (3.15). By virtue of (3.19), we have only to show that $Y = 0$ for any $Y \in \mathbf{M}_{-\infty}^n(\mathbf{X}) \ominus \mathbf{M}_{-\infty}^n(\xi_+^{(K)})$. By (3.28), Y is orthogonal to $\{\xi_{+j}^{(K)}(l); 1 \leq j \leq d, l > n\}$ and so orthogonal to the space $\mathbf{M}_{-\infty}^{+\infty}(\xi_+^{(K)})$. Therefore, we see from (3.27) that $Y = 0$.

Thus, we have proved that part (ii) $\rightarrow$ (iii). The converse statement (iii) $\rightarrow$ (i)

follows immediately from (3.14).

Q.E.D.

COROLLARY 3.1. *Let $\mathbf{X}$ be any d -dimensional $AR(K)$ -time series. Then $\mathbf{X}$ has the spectral measure matrix $\Delta(d\theta)$ such that*

$$(3.29) \quad \Delta(d\theta) = P(e^{i\theta})^{-1} V_+(K)^t \overline{P(e^{i\theta})}^{-1} d\theta,$$

where the matrix polynomial $P(z)$ ($z \in \mathbf{C}$) is given by (3.23).

§4. Test(S) for weak stationarity

Let d and N be any fixed natural numbers.

4.1. For any given $N+1$ vectors $\mathcal{Z}(n)$ in $\mathbf{R}^d$ ($0 \leq n \leq N$), we denote by $\mu^{\mathcal{Z}}$ and $R^{\mathcal{Z}} = (R_{jk}^{\mathcal{Z}}(\cdot))_{1 \leq j,k \leq d}$ the *sample mean vector* and the *sample covariance matrix function* of the data $\mathcal{Z} = (\mathcal{Z}(n); 0 \leq n \leq N)$, respectively:

$$(4.1) \quad \mu^{\mathcal{Z}} \equiv \frac{1}{N+1} \sum_{n=0}^N \mathcal{Z}(n)$$

$$(4.2) \quad \begin{cases} R_{jk}^{\mathcal{Z}}(n) \equiv \frac{1}{N+1} \sum_{m=0}^{N-n} (\mathcal{Z}_j(n+m) - \mu_j^{\mathcal{Z}})(\mathcal{Z}_k(m) - \mu_k^{\mathcal{Z}}) \\ R_{jk}^{\mathcal{Z}}(-n) \equiv R_{kj}^{\mathcal{Z}}(n), \end{cases}$$

where $\mu^{\mathcal{Z}} = {}^t(\mu_1^{\mathcal{Z}}, \dots, \mu_d^{\mathcal{Z}})$ and $\mathcal{Z}(n) = {}^t(\mathcal{Z}_1(n), \dots, \mathcal{Z}_d(n))$ ($0 \leq n \leq N$).

Let $\mathcal{X} = (\mathcal{X}(n); 0 \leq n \leq N)$ be the standardized data of $\mathcal{Z}$:

$$(4.3) \quad \mathcal{X}(n) = \begin{pmatrix} R_{11}^{\mathcal{Z}}(0)^{-1/2} & & 0 \\ & \ddots & \\ 0 & & R_{dd}^{\mathcal{Z}}(0)^{-1/2} \end{pmatrix} (\mathcal{Z}(n) - \mu^{\mathcal{Z}}) \quad (0 \leq n \leq N).$$

Since the sample covariance matrix function $R^{\mathcal{X}}(\cdot)$ of $\mathcal{X}$ can be extended to a non-negative definite matrix function, we can construct a system $\{\gamma_+(n, k), \gamma_-(n, k), \delta_+(m), \delta_-(m), V_+(l), V_-(l); 1 \leq k < n \leq N, 1 \leq m \leq N, 0 \leq l \leq N\}$ of members in $M(d; \mathbf{R})$, under the assumption

$$(4.4) \quad V_+(n) \in GL(d; \mathbf{R}) \quad (0 \leq n \leq N-1),$$

according to the relations (2.11_±)–(2.13) and (2.16_±) with R in (2.16_±) replaced by $R^{\mathcal{X}}$.

An experience rule in data analysis tells us that an effective number of the whole series $(R^{\mathcal{X}}(n); 0 \leq n \leq N)$ of the matrix function $R^{\mathcal{X}}$ is considered to be

between $[2\sqrt{N+1}/d]$ and $[3\sqrt{N+1}/d]$ ([1]). Here we choose the maximum value and put

$$(4.5) \quad M = [3\sqrt{N+1}/d] - 1.$$

Thus, in what follows, we intend to make use of the subsystem $\{\gamma_+(n, k), V_+(l); 0 \leq k < n \leq M, 0 \leq l \leq M\}$, which is the sample forward KM_2O -Langevin data associated with the reliable part $(R^{\mathcal{X}}(n); 0 \leq n \leq M)$ of $R^{\mathcal{X}}$.

In order to analyze the internal structure of $\mathcal{X}$, we consider, for each $i \in \{0, \dots, N-M\}$, the shifted data $\mathcal{X}_i$ with $\mathcal{X}(i)$ as its initial point:

$$(4.6) \quad \mathcal{X}_i = (\mathcal{X}(i+n); 0 \leq n \leq M).$$

According to (2.10₊) with $\mathbf{X}$ replaced by $\mathcal{X}_i$ we define $\nu_{+i} = (\nu_{+i}(n); 0 \leq n \leq M)$ by

$$(4.7) \quad \begin{cases} \nu_{+i}(0) = \mathcal{X}(i) \\ \nu_{+i}(n) = \mathcal{X}(i+n) + \sum_{k=0}^{n-1} \gamma_+(n, k) \mathcal{X}(i+k) \quad (1 \leq n \leq M). \end{cases}$$

This is the sample forward KM_2O -Langevin force associated with $\mathcal{X}_i$.

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

$$(4.8) \quad V_+(n) = W_+(n)^t W_+(n) \quad (0 \leq n \leq M),$$

we define, for each $i \in \{0, \dots, N-M\}$, a d -dimensional data $\xi_{+i} = (\xi_{+i}(n); 0 \leq n \leq M) = ({}^t(\xi_{+i1}(n), \dots, \xi_{+id}(n)); 0 \leq n \leq M)$ and a one-dimensional data $\xi_i = (\xi_i(n); 0 \leq n \leq d(M+1)-1)$ by

$$(4.9) \quad \xi_{+i}(n) = W_+(n)^{-1} \nu_{+i}(n)$$

$$(4.10) \quad \xi_i(n) = \xi_{+ij}(m), \quad n = dm + j - 1 \quad (1 \leq j \leq d, 0 \leq m \leq M),$$

where $\xi_{+i}(n) = {}^t(\xi_{+i1}(n), \dots, \xi_{+id}(n))$.

By virtue of the fundamental principle—the decomposition theorem, the fluctuation-dissipation theorem and the reconstruction theorem in §2, we can assert that for each $i \in \{0, \dots, N-M\}$,

$$(4.11) \quad \mathcal{X}_i \text{ is a realization of a weakly stationary local time series with the covariance matrix function } R^{\mathcal{X}}$$

if and only if

$$(4.12) \quad \xi_i \text{ is a realization of a standard white noise in the wide sense.}$$

In order to check whether ξ_i satisfies (4.12) or not, we put ξ_i to the following criteria $(M)_i$, $(V)_i$ and $(O)_i$ for checking mean zero, variance one and the orthogonality, respectively:

$$(M)_i \quad \sqrt{d(M+1)} |\mu^{\xi_i}| < 1.96$$

$$(V)_i \quad |(v^{\xi_i} - 1)^{\sim}| < 2.2414$$

$$(O)_i \quad d(M+1) \left(\sum_{j=1}^2 (L_{n,m}^{(j)})^{1/2} \right)^{-1} |R^{\xi_i}(n, m)| < 1.96,$$

where μ^{ξ_i} , $(v^{\xi_i} - 1)^{\sim}$ and $R^{\xi_i}(n, m)$ ($1 \leq n \leq M$, $0 \leq m \leq M - n$) are defined by

$$(4.13) \quad \mu^{\xi_i} = \frac{1}{d(M+1)} \sum_{k=0}^{d(M+1)-1} \xi_i(k)$$

$$(4.14) \quad (v^{\xi_i} - 1)^{\sim} = \left\{ \sum_{k=0}^{d(M+1)-1} (\xi_i(k)^2 - 1) \right\} \left\{ \sum_{k=0}^{d(M+1)-1} (\xi_i(k)^2 - 1)^2 \right\}^{-1/2}$$

$$(4.15) \quad R^{\xi_i}(n, m) = \frac{1}{d(M+1)} \sum_{k=m}^{d(M+1)-1-n} \xi_i(k) \xi_i(n+k).$$

Furthermore, $L_{n,m}^{(j)}$ ($j = 1, 2$) are defined as follows: We divide $d(M+1)$ and m by $2n$ and n and denote their remainders by r and t , respectively;

$$(4.16) \quad \begin{cases} d(M+1) = q(2n) + r & (0 \leq r \leq 2n-1) \\ m = sn + t & (0 \leq t \leq n-1), \end{cases}$$

if $0 \leq r \leq n$, then

$$(4.17) \quad \begin{cases} L_{n,m}^{(1)} = \begin{cases} n(q + (s/2)) - m & (s \text{ is even}) \\ n(q - (s+1)/2) & (s \text{ is odd}) \end{cases} \\ L_{n,m}^{(2)} = \begin{cases} n(q - 1 - s/2) + r & (s \text{ is even}) \\ n(q - 1 + (s+1)/2) + r - m & (s \text{ is odd}) \end{cases} \end{cases}$$

and if $n+1 \leq r \leq 2n-1$, then

$$(4.18) \quad \begin{cases} L_{n,m}^{(1)} = \begin{cases} n(q - 1 + (s/2)) + r - m & (s \text{ is even}) \\ n(q - 1 - (s+1)/2) + r & (s \text{ is odd}) \end{cases} \\ L_{n,m}^{(2)} = \begin{cases} n(q - s/2) & (s \text{ is even}) \\ n(q + (s+1)/2) - m & (s \text{ is odd}). \end{cases} \end{cases}$$

We note that

$$(4.19) \quad d(M+1) - n - m = L_{n,m}^{(1)} + L_{n,m}^{(2)}.$$

Now **Test(S)** for the weak stationarity of a given data $\mathcal{Z} = (\mathcal{Z}(n); 0 \leq n \leq N)$ is proposed as follows:

Test(S): The rates of successful $i \in \{0, \dots, N-M\}$ are over 80%, 70% and 80% for $(M)_i$, $(V)_i$ and $(O)_i$, respectively.

4.2. To attain **Test(S)** above, we carried out in [31] repeated experiments for various concrete data such as random normal numbers, random uniform numbers, logistic and tent transformations as well as for the data transformed in four ways: taking the first difference, multiplying or adding the above data $\mathcal{Z}(n)$ by time n (this is expected to switch from stationarity to non-stationarity), and taking the square or cube.

In order to reinforce the above experiments for **Test(S)**, we construct some one-dimensional AR -time series according to the reconstruction theorem in §2 and the characterization theorem for AR -time series in §3 (Theorem 3.1). **Test(S)** is then applied to the data obtained by transforming them in several ways.

For a positive number V and any K numbers $\delta(n)$, $-1 < \delta(n) < 1$ ($1 \leq n \leq K$), we get a KM_2O -Langevin data $\{\gamma(n, k), V(l); 0 \leq k < n < \infty, 0 \leq l < \infty\}$ by the same procedure as in the reconstruction theorem in §2, where we put $\delta(n) = 0$ for any $n \geq K+1$. For any standardized random uniform numbers $\xi(n)$ ($0 \leq n \leq N$), $K \leq N$, we define an $AR(K)$ -time series $\mathcal{Z} = (\mathcal{Z}(n); 0 \leq n \leq N)$ by

$$(4.20) \quad \mathcal{Z}(n) = \begin{cases} -\sum_{k=0}^{n-1} \gamma(n, k) \mathcal{Z}(k) + V(n)^{1/2} \xi(n) & (0 \leq n \leq K) \\ -\sum_{k=1}^K \gamma(K, K-k) \mathcal{Z}(n-K) + V(K)^{1/2} \xi(n) & (K+1 \leq n \leq N). \end{cases}$$

In particular, we consider the $AR(2)$ -time series $\mathcal{Z} = (\mathcal{Z}(n); 0 \leq n \leq 104)$ such that

$$(4.21) \quad V = 1 \quad \text{and} \quad \delta(1) = 0.6, \quad \delta(2) = -0.3$$

and put $\mathcal{Y} = (\mathcal{Y}(n); -5 \leq n \leq 99)$ as

$$(4.22) \quad \mathcal{Y}(n) = \mathcal{Z}(n+5).$$

Then we define five kinds of time series $\mathcal{Z}_j = (\mathcal{Z}_j(n); 0 \leq n \leq 99)$ ($1 \leq j \leq 5$) by

$$(4.23) \quad \begin{cases} \mathcal{Z}_1(n) = \mathcal{Y}(n), & \mathcal{Z}_2(n) = \mathcal{Y}(n)^2, \mathcal{Z}_3(n) = \mathcal{Y}(n)^3, \\ \mathcal{Z}_4(n) = \mathcal{Y}(n-3) - 2\mathcal{Y}(n-5), & \mathcal{Z}_5(n) = \mathcal{Y}(n-3) - 2\mathcal{Y}(n-5)^2. \end{cases}$$

We made one-hundred experiments of Test(S) for the following two-dimensional data $({}^t(\mathcal{Z}_j(n), \mathcal{Z}_k(n)) ; 0 \leq n \leq 99) (1 \leq j < k \leq 5)$, by taking standardized random uniform numbers $\xi(n)$ ($0 \leq n \leq 104$) with prime seed numbers from 2 to 541. The results are illustrated in Table 4.1, which denotes the rate of the numbers of data passing Test(S).

Data	$\mathcal{Z}_2$	$\mathcal{Z}_3$	$\mathcal{Z}_4$	$\mathcal{Z}_5$
$\mathcal{Z}_1$	0.93	0.74	0.70	0.97
$\mathcal{Z}_2$		0.23*	0.99	0.58*
$\mathcal{Z}_3$			0.93	0.54*
$\mathcal{Z}_4$				0.97

Table 4.1. Test(S) for *AR*-dynamics

Table 4.1 shows that the results for the data ${}^t(\mathcal{Z}_2, \mathcal{Z}_3)$, ${}^t(\mathcal{Z}_2, \mathcal{Z}_5)$ and ${}^t(\mathcal{Z}_3, \mathcal{Z}_5)$ do not work out according to theory. By taking account of the possibility of arctangent transformation that compresses the abnormal values or the multiplication effect between two data $\mathcal{Z}_2$ and $\mathcal{Z}_3$ obtained by taking the square and cube of the data $\mathcal{Y}$, we applied Test(S)_{Arct} to them and got a good result as shown in Table 4.2, where $\mathcal{X}_j$ is the standardized data of $\mathcal{Z}_j$ ($j = 2, 3, 5$).

Data	Arct $\mathcal{X}_3$	Arct $\mathcal{X}_5$
Arct $\mathcal{X}_2$	0.83	0.97
Arct $\mathcal{X}_3$		0.98

Table 4.2. Test(S)_{Arct} for *AR*-dynamics

4.3. For one hundred standardized random uniform numbers $\xi = (\xi(n) ; -4 \leq n \leq 98)$ with prime seed numbers from 2 to 541, we construct MA(3)-time series $\mathcal{Z} = (\mathcal{Z}(n) ; -1 \leq n \leq 99)$ by

$$(4.24) \quad \mathcal{Z}(n) = 7\xi(n-1) - 3\xi(n-2) - 5\xi(n-3)$$

and then define $\mathcal{Z}_j = (\mathcal{Z}_j(n) ; 0 \leq n \leq 99) (1 \leq j \leq 8)$ by

$$(4.25) \quad \begin{cases} \mathcal{Z}_1(n) = \mathcal{Z}(n), & \mathcal{Z}_2(n) = \mathcal{Z}(n)^2, & \mathcal{Z}_3(n) = n\mathcal{Z}(n), \\ \mathcal{Z}_4(n) = \mathcal{Z}(n) - \mathcal{Z}(n-1), & \mathcal{Z}_5(n) = \mathcal{Z}_4(n)^2, & \mathcal{Z}_6(n) = n\mathcal{Z}_4(n), \\ \mathcal{Z}_7(n) = {}^t(\mathcal{Z}_1(n), \mathcal{Z}_2(n)), & \mathcal{Z}_8(n) = {}^t(\mathcal{Z}_4(n), \mathcal{Z}_5(n)). \end{cases}$$

j	1	2	3	4	5	6	7	8
rate	1.00	0.93	0.15	1.00	0.84	0.08	0.98	0.94

Table 4.3. Test(S) for $MA(3)$ -time series

Table 4.3 shows the rate of the numbers passing Test(S). The results work out according to theory.

§5. Test(CS) for causality

Let $\mathbf{X} = (X(n) ; n \in \mathbf{Z})$ and $\mathbf{Y} = (Y(n) ; n \in \mathbf{Z})$ be $\mathbf{R}^{d_1}$ - and $\mathbf{R}^{d_2}$ -valued time series defined on a probability space $(\Omega, \mathcal{B}, P)$, respectively. We say that $\mathbf{X}$ *causes* $\mathbf{Y}$ if for each $n \in \mathbf{Z}$, there exists a measurable mapping F_n from the infinite-dimensional space $(\mathbf{R}^{d_1})^{\mathbf{N}^*}$ to the finite-dimensional space $\mathbf{R}^{d_2}$ such that

$$(5.1) \quad Y(n) = F_n(X(n), X(n-1), X(n-2), \dots).$$

In particular, we say that $\mathbf{X}$ *causes* $\mathbf{Y}$ *linearly* or *the causality is linear* if the function F_n is linear. Moreover, for $p \in \mathbf{N}$, *the causality is said to be non-linear of degree p* if the function F_n is a polynomial of degree p .

EXAMPLE 5.1. Let $d_1 = d_2 = 1$ and

$$(5.2) \quad Y(n) = c_1 X(n - s_1)^{p_1} + c_2 X(n - s_2)^{p_2} \quad (n \in \mathbf{Z}),$$

where $c_j \in \mathbf{R}$ and $s_j \in \mathbf{N}^*$, $p_j \in \mathbf{N}$ ($j = 1, 2$). It follows that $\mathbf{X}$ causes $\mathbf{Y}$, and if $p_1 = p_2 = 1$, then the causality is linear.

In what follows, we treat the case $d_1 = d$, $d_2 = 1$ and assume that the $d + 1$ -dimensional time series $\mathbf{U} = ({}^t(X(n), Y(n)) ; n \in \mathbf{Z})$ is weakly stationary with mean vector zero. So we have three kinds of covariance matrix functions R_j ($j = 1, 2, 3$):

$$(5.3) \quad R_1(n) = E(X(n){}^t X(0)) \in M(d, d ; \mathbf{R})$$

$$(5.4) \quad R_2(n) = E(Y(n){}^t X(0)) \in M(1, d ; \mathbf{R})$$

$$(5.5) \quad R_3(n) = E(Y(n)Y(0)) \in M(1, 1, ; \mathbf{R}) \equiv \mathbf{R}.$$

5.1. In this subsection, we are concerned with an observation to see whether $\mathbf{X}$ causes $\mathbf{Y}$ linearly or not. At first, note that $\mathbf{X}$ causes $\mathbf{Y}$ linearly if and only if

$$(5.6) \quad \|Y(0) - P_{\mathbf{M}_{\infty}^{\circ}(\mathbf{X})} Y(0)\| = 0.$$

Since $\mathbf{U}$ is weakly stationary, we have

$$\|Y(0) - P_{\mathbf{M}_0^*(\mathbf{X})} Y(0)\| = \lim_{n \rightarrow \infty} \|Y(n) - P_{\mathbf{M}_0^*(\mathbf{X})} Y(n)\|.$$

Using the weak stationarity of $\mathbf{U}$ again, the norm

$$\|Y(n) - P_{\mathbf{M}_0^*(\mathbf{X})} Y(n)\| = (R_3(0) - \|P_{\mathbf{M}_0^*(\mathbf{X})} Y(n)\|^2)^{1/2}$$

is decreasing in n and so we see that (5.6) is equivalent to

$$(5.7) \quad C_n(\mathbf{Y} | \mathbf{X}) \equiv \|P_{\mathbf{M}_0^*(\mathbf{X})} Y(n)\| \text{ increases to } R_3(0)^{1/2} \text{ as } n \rightarrow \infty.$$

We call $C_n(\mathbf{Y} | \mathbf{X})$ the *causality function from $\mathbf{X}$ to $\mathbf{Y}$* . Our next goal is to give an algorithm to calculate the quantity $C_n(\mathbf{Y} | \mathbf{X})$. To this end, we further assume that condition (2.3) holds for the process $\mathbf{X}$. Note that $\mathbf{X} = (X(n); n \in \mathbf{Z})$ is a d -dimensional weakly stationary time series with covariance matrix function R_1 . As in §2, ν_+ (resp. $\{\nu_+(n, k), V_+(l); 0 \leq k < n < \infty, l \in \mathbf{N}^*\}$) denotes the forward KM_2O -Langevin force (resp. data) associated with $\mathbf{X}$. Then it follows from (2.7) and (2.8) that there exists a unique system $\{C(n, k); 0 \leq k \leq n < \infty\}$ of members in $M(1, d; \mathbf{R})$ such that

$$(5.8) \quad P_{\mathbf{M}_0^*(\mathbf{X})} Y(n) = \sum_{k=0}^n C(n, k) \nu_+(k) \quad (n \in \mathbf{N}^*).$$

THEOREM 5.1. (i) For each $n \in \mathbf{N}^*$,

$$(5.9) \quad C_n(\mathbf{Y} | \mathbf{X}) = \left\{ \sum_{k=0}^n C(n, k) V_+(k)^t C(n, k) \right\}^{1/2}.$$

(ii) For any $n, k \in \mathbf{N}^*, 0 \leq k \leq n$,

$$(5.10) \quad C(n, k) = \begin{cases} R_2(n) R_1(0)^{-1} & (k = 0) \\ \{R_2(n - k) + \sum_{l=0}^{k-1} R_2(n - l)^t \gamma_+(k, l)\} V_+(k)^{-1} & (0 < k \leq n) \end{cases}.$$

Proof. (i) follows from (2.7), (2.9) and (5.8). For any $k, 0 < k \leq n$, we see from (2.7) and (2.8) that

$$E(Y(n)^t \nu_+(k)) = E\{(P_{\mathbf{M}_0^*(\mathbf{X})} Y(n))^t \nu_+(k)\} = C(n, k) V_+(k)$$

and so by (2.10₊)

$$\begin{aligned} C(n, k) &= E\{Y(n)^t \nu_+(k)\} V_+(k)^{-1} \\ &= E\{Y(n)^t (X(k) + \sum_{l=0}^{k-1} \gamma_+(k, l) X(l))\} V_+(k)^{-1} \end{aligned}$$

$$= \{R_2(n-k) + \sum_{l=0}^{k-1} R_2(n-l)^t \gamma_+(k, l)\} V_+(k)^{-1}.$$

The formula for the case $k = 0$ follows in the same manner.

Q.E.D.

5.2. Concerning the non-linear causality, we define a concept of a little restrictive non-linearity. Let $\mathbf{U} = ({}^t(X(n), Y(n)) ; n \in \mathbf{Z})$ be a two-dimensional time series such that the $p+1$ -dimensional time series $\mathbf{V} = ({}^t(X(n), X(n)^2, \dots, X(n)^p, Y(n)) ; n \in \mathbf{Z})$ is weakly stationary. Then we say that $\mathbf{X} = (X(n) ; n \in \mathbf{Z})$ *causes* $\mathbf{Y}$ with non-linearity of degree p in the narrow sense if the following holds:

$$(5.11) \quad \lim_{n \rightarrow \infty} \|Y(n) - P_{\mathbf{M}_0^n(\mathbf{X}; p)} Y(n)\| = 0,$$

where

$$(5.12) \quad \mathbf{M}_0^n(\mathbf{X}; p) = \text{the closed linear hull of } \{X(m)^q ; 0 \leq m \leq n, 1 \leq q \leq p\}.$$

This is nothing but saying that the p -dimensional time series $\mathbf{W} = ({}^t(X(n), X(n)^2, \dots, X(n)^p) ; n \in \mathbf{Z})$ causes $\mathbf{Y}$ linearly.

§6. Test(CS) for AR-dynamics

Let d and N be any natural numbers. For any given $N+1$ data $\mathcal{Z}(n)$ in $\mathbf{R}^d$ (resp. $\mathcal{W}(n)$ in $\mathbf{R}$) ($0 \leq n \leq N$), we denote by $\mu^{\mathcal{Z}}$ (resp. $\mu^{\mathcal{W}}$) the sample mean vector of the data $\mathcal{Z}$ (resp. $\mathcal{W}$) and by $R^{\mathcal{Z}}$ (resp. $R^{\mathcal{W}}$) the sample covariance matrix function of the data $\mathcal{Z}$ (resp. $\mathcal{W}$). Furthermore, we define the *sample mutual covariance matrix function* $R^{\mathcal{W}\mathcal{Z}}(n) = (R_1^{\mathcal{W}\mathcal{Z}}(n), \dots, R_d^{\mathcal{W}\mathcal{Z}}(n))$ of the data $\mathcal{W}$ and $\mathcal{Z}$ by

$$(6.1) \quad R_j^{\mathcal{W}\mathcal{Z}}(n) = \frac{1}{N+1} \sum_{k=0}^{N-n} (\mathcal{W}(n+k) - \mu^{\mathcal{W}})(\mathcal{Z}_j(k) - \mu_j^{\mathcal{Z}}) \\ (1 \leq j \leq d, 0 \leq n \leq M),$$

where M is the effective number given by

$$(6.2) \quad M = [3\sqrt{N+1} / (d+1)] - 1.$$

Let $\mathcal{X}$ and $\mathcal{Y}$ be the standardization of the data $\mathcal{Z}$ and $\mathcal{W}$, respectively. By using (5.9) and (5.10), we define the *sample causality function* $C_n(\mathcal{W} | \mathcal{Z})$ from $\mathcal{Z}$ to $\mathcal{W}$ by

$$(6.3) \quad C_n(\mathcal{W} | \mathcal{Z}) = \left\{ \sum_{k=0}^n C(n, k) V_+(k)^t C(n, k) \right\}^{1/2} \quad (0 \leq n \leq M),$$

where

$$(6.4) \quad C(n, k) = \begin{cases} R^{\mathcal{Y}\mathcal{X}}(n)R^{\mathcal{X}}(0)^{-1} & (k=0) \\ \{R^{\mathcal{Y}\mathcal{X}}(n-k) + \sum_{l=0}^{k-1} R^{\mathcal{Y}\mathcal{X}}(n-l)' \gamma_+(k, l) V_+(k)^{-1} & (0 < k \leq n) \end{cases}.$$

In what follows, we shall illustrate the graphs of sample causality functions for various kinds of models.

EXAMPLE 6.1 (Independent *AR*-time series). Let $\mathcal{X}_1 = \{\mathcal{X}_1(n); 0 \leq n \leq 102\}$ be an *AR*(2)-time series with characteristics $(V, \delta(1), \delta(2)) = (1, 0.6, -0.3)$ in (4.21) of the subsection [4.2] and $\mathcal{X}_2 = \{\mathcal{X}_2(n); 0 \leq n \leq 102\}$ be an *AR*(3)-time series with characteristics $(V, \delta(1), \delta(2), \delta(3)) = (1, 0.9, -0.4, 0.5)$, where the seed numbers of the standardized random uniform numbers $\xi(n)$ ($0 \leq n \leq 102$) are 229 and 127 for $\mathcal{X}_1$ and $\mathcal{X}_2$, respectively. We illustrate in Figure 6.1 the graphs of $C_n(\mathcal{X}_1|\mathcal{X}_2)$ and $C_n(\mathcal{X}_2|\mathcal{X}_1)$ as a line graph and a short-dashed graph, respectively.

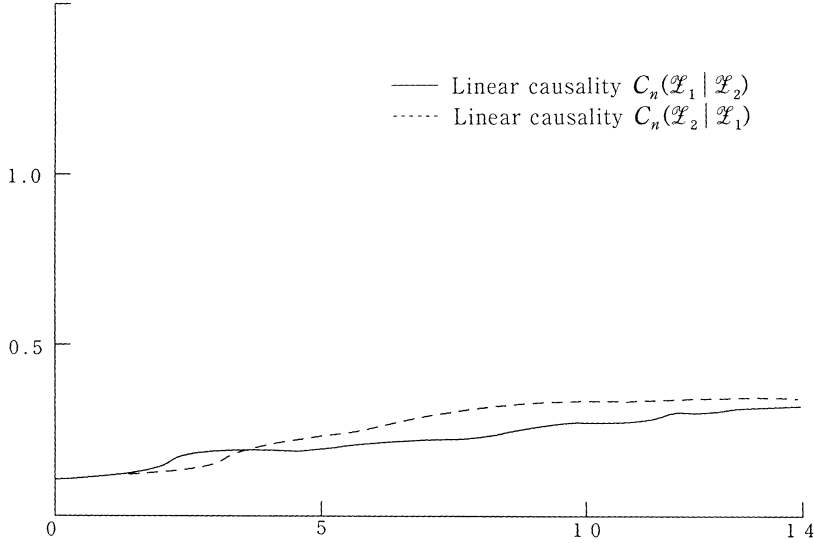
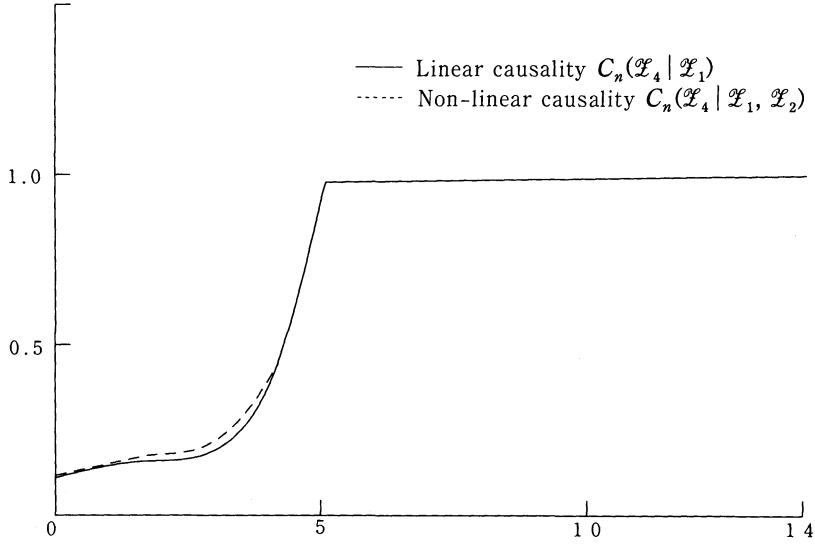
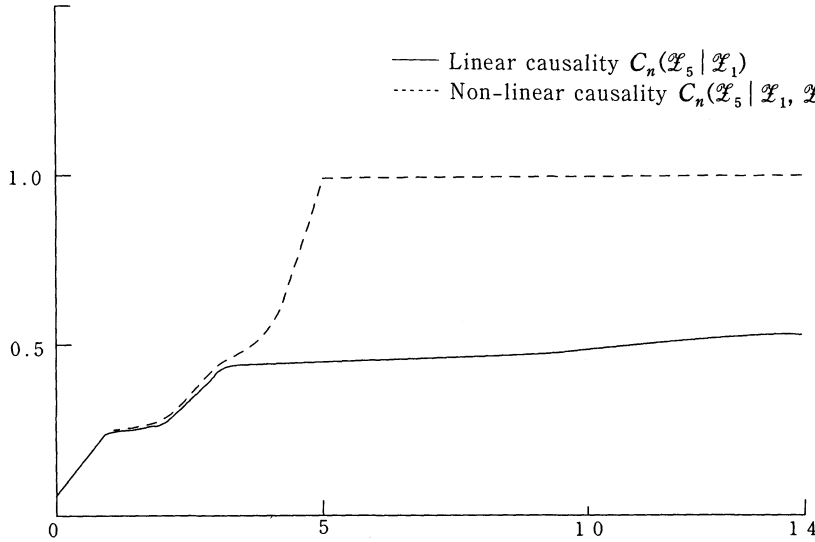


Figure 6.1. Independent *AR*-time series

EXAMPLE 6.2. Using the *AR*(2)-time series $\mathcal{X}_1$ in Example 6.1, we construct five kinds of time series $\mathcal{X}_j$ ($1 \leq j \leq 5$) according to the dynamics (4.23). The graphs of $C_n(\mathcal{X}_4|\mathcal{X}_1)$ and $C_n(\mathcal{X}_4|\mathcal{X}_1, \mathcal{X}_2)$ are drawn in Figure 6.2 as a line graph and a short-dashed line, respectively. This implies that the causality for the linear system $\mathcal{X}_4$ in (4.23) is covered by the function $C_n(\mathcal{X}_4|\mathcal{X}_1)$.

Figure 6.2. *AR*-dynamics

EXAMPLE 6.3. In the previous Example 6.2, we draw in Figure 6.3 the graphs of $C_n(\mathcal{Z}_5 | \mathcal{Z}_1)$ and $C_n(\mathcal{Z}_5 | \mathcal{Z}_1, \mathcal{Z}_2)$ as a line graph and a short-dashed line, respectively. We find that the function $C_n(\mathcal{Z}_5 | \mathcal{Z}_1)$ is insufficient, but the function $C_n(\mathcal{Z}_5 | \mathcal{Z}_1, \mathcal{Z}_2)$ works out effectively for the non-linear system $\mathcal{Z}_5$ in (4.23).

Figure 6.3. *AR*-dynamics

EXAMPLE 6.4. Considering Example 6.2 again, this time, we draw in Figure 6.4 the graphs of $C_n(\mathcal{L}_5|\mathcal{L}_1)$ and $C_n(\mathcal{L}_5|\mathcal{L}_1, \mathcal{L}_3)$ as a line graph and a short-dashed line, respectively. We see that the function $C_n(\mathcal{L}_5|\mathcal{L}_1, \mathcal{L}_3)$ with degree 3 is not effective for the non-linear system $\mathcal{L}_5$ with degree 2 in (4.23).

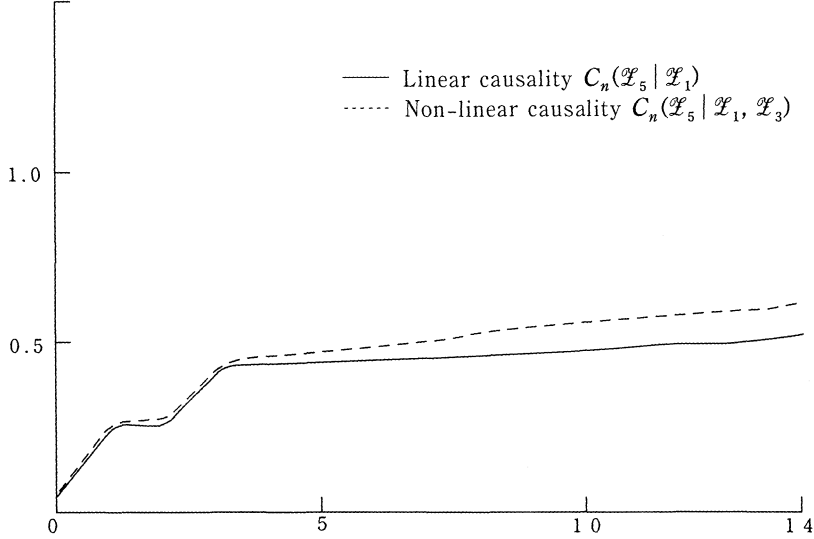


Figure 6.4. AR-dynamics

§7. Test(CS) for meteorological data

7.1. Let $\mathcal{X}_1 = (\mathcal{X}_1(n) ; 0 \leq n \leq 113)$ and $\mathcal{X}_2 = (\mathcal{X}_2(n) ; 0 \leq n \leq 113)$ be Wolfer's sunspot numbers and the trapped numbers of Canadian Lynx in MacKenzie River from 1821 to 1934, respectively. Trapping records of Lynx are found in Elton and Nicholson's paper ([4]). We denote by $\mathcal{X}_j$ the standardized data of $\mathcal{X}_j$ ($j = 1, 2$). We define four kinds of one-dimensional time series $\text{Arct } \mathcal{X}_j = ((\text{Arct } \mathcal{X}_j)(n) ; 0 \leq n \leq 113)$ and $\text{Log } \mathcal{X}_j = ((\text{Log } \mathcal{X}_j)(n) ; 0 \leq n \leq 113)$ ($j = 1, 2$) by

$$(7.1) \quad (\text{Arct } \mathcal{X}_j)(n) = \arct(\mathcal{X}_j(n))$$

$$(7.2) \quad (\text{Log } \mathcal{X}_j)(n) = \log(\mathcal{X}_j(n))$$

and then three kinds of two-dimensional time series $\mathcal{X} = (\mathcal{X}(n) ; 0 \leq n \leq 113)$, $\text{Arct } \mathcal{X} = ((\text{Arct } \mathcal{X})(n) ; 0 \leq n \leq 113)$ and $\text{Log } \mathcal{X} = ((\text{Log } \mathcal{X})(n) ; 0 \leq n \leq 113)$ by

$$(7.3) \quad \mathcal{X}(n) = {}^t(\mathcal{X}_1(n), \mathcal{X}_2(n))$$

$$(7.4) \quad (\text{Arct } \mathcal{X})(n) = {}^t((\text{Arct } \mathcal{X}_1)(n), (\text{Arct } \mathcal{X}_2)(n))$$

$$(7.5) \quad (\text{Log } \mathcal{Z})(n) = '((\text{Log } \mathcal{Z}_1)(n), \text{Log } \mathcal{Z}_2)(n)).$$

We have already in [31] checked that all the time series $\mathcal{Z}$, $\text{Arct } \mathcal{X}$ and $\text{Log } \mathcal{Z}$ have weak stationarity. Hence, we can apply our $\text{Test}(\text{CS})$ to them. The graphs of the sample causality functions for the data $\mathcal{Z}$, $\text{Arct } \mathcal{X}$ and $\text{Log } \mathcal{Z}$ are drawn in Figures 7.1, 7.2, and 7.3, respectively.

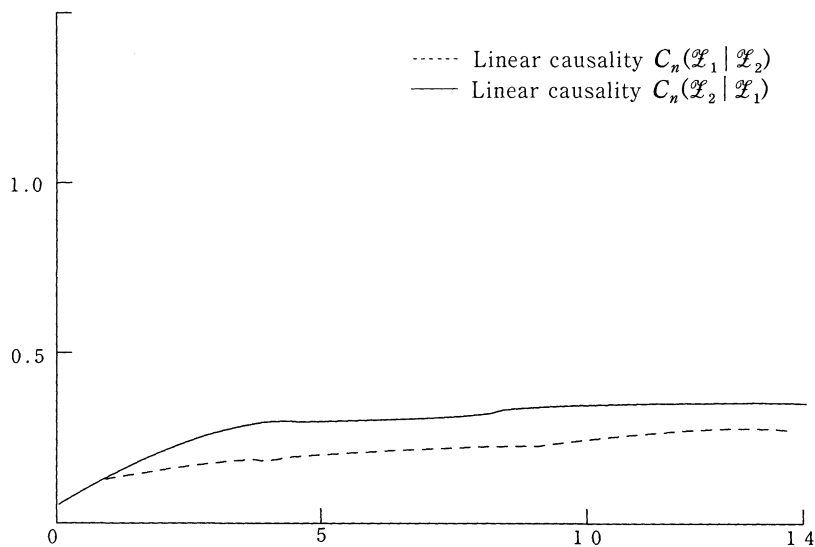


Figure 7.1. (Sunspot, Lynx in MacKenzie River)

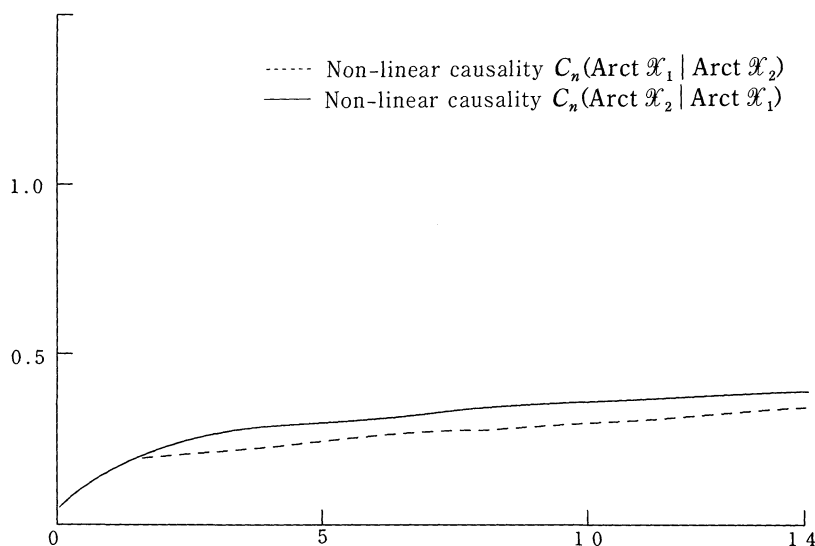


Figure 7.2. Arct (Sunspot, Lynx in MacKenzie River)

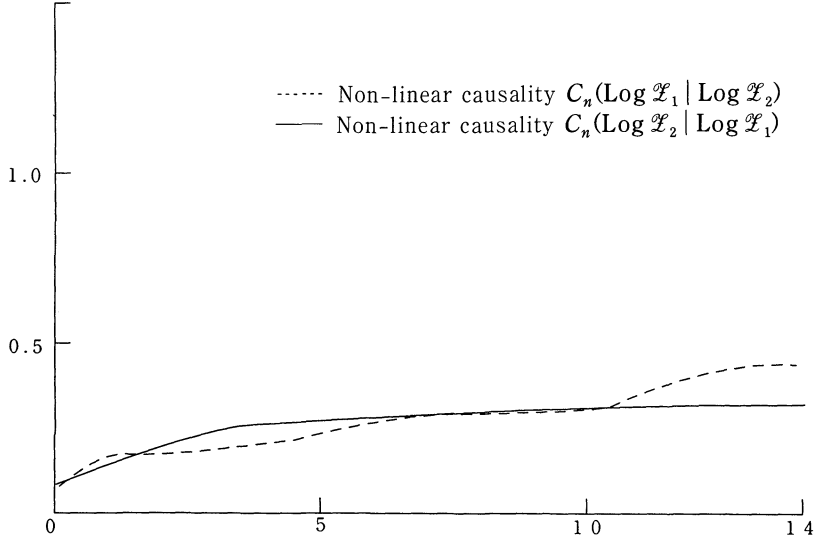


Figure 7.3. Log (Sunspot, Lynx in MacKenzie River)

Having taken the results in §6 into account, we can not say that there exists certain marked causal relation between Wolfer's sunspot numbers and the trapped numbers of Canadian Lynx in MacKenzie River.

7.2. We study the problem of causal direction among three data $\mathcal{L}_j = (\mathcal{L}_j(n); 0 \leq n \leq 99)$, $j = 3, 4, 5$, of Wolfer's sunspot numbers, the rainfall averages in a year in Sapporo and the temperature averages in a year in Sapporo, respectively. All of them are taken from 1889 to 1988. Let $\mathcal{X}_j$ be the standardized data of $\mathcal{L}_j$ ($j = 3, 4, 5$).

We apply $\text{Test}(S)$ and $\text{Test}(S)_{\text{Arct}}$ to three kinds of two-dimensional time series ${}^t(\mathcal{L}_3, \mathcal{L}_4)$, ${}^t(\mathcal{L}_3, \mathcal{L}_5)$ and ${}^t(\mathcal{L}_4, \mathcal{L}_5)$. The results are shown in Table 7.1 and Table 7.2, which show that the three-dimensional time series ${}^t(\mathcal{L}_3, \mathcal{L}_4, \mathcal{L}_5)$ does not have weak stationarity, but the three-dimensional time series ${}^t(\text{Arct } \mathcal{X}_3, \text{Arct } \mathcal{X}_4, \text{Arct } \mathcal{X}_5)$ does have.

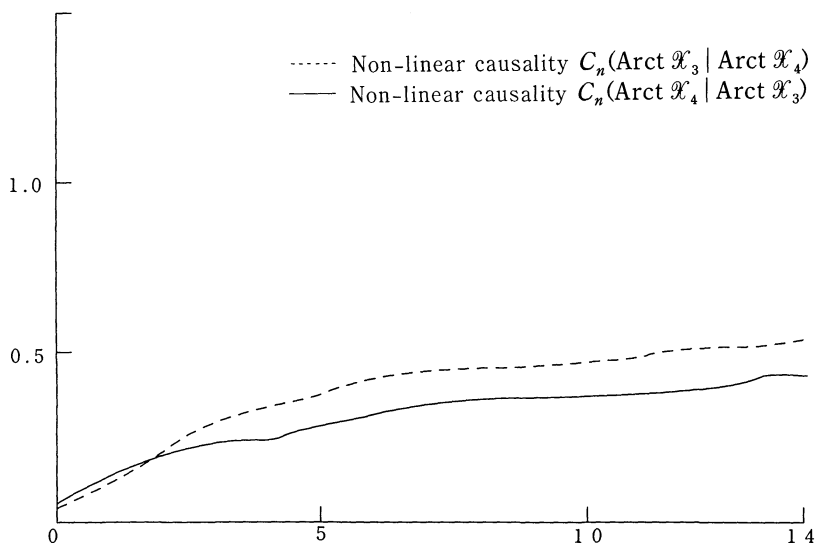
	(M)	(V)	(O)	(S)
${}^t(\mathcal{L}_3, \mathcal{L}_4)$	0.977	0.860	1.000	S
${}^t(\mathcal{L}_3, \mathcal{L}_5)$	0.988	0.674	0.988	NS
${}^t(\mathcal{L}_4, \mathcal{L}_5)$	1.000	0.872	0.988	S

Table 7.1. Test(S)

	(M)	(V)	(O)	(S)
${}^t(\text{Arct } \mathcal{X}_3, \text{Arct } \mathcal{X}_4)$	1.000	1.000	1.000	S
${}^t(\text{Arct } \mathcal{X}_3, \text{Arct } \mathcal{X}_5)$	1.000	0.953	0.965	S
${}^t(\text{Arct } \mathcal{X}_4, \text{Arct } \mathcal{X}_5)$	0.988	0.907	0.977	S

Table 7.2. $\text{Test}(S)_{\text{Arct}}$

Therefore, we can apply $\text{Test}(CS)$ to the data ${}^t(\text{Arct } \mathcal{X}_3, \text{Arct } \mathcal{X}_4, \text{Arct } \mathcal{X}_5)$. The graphs of the sample causality functions for the data ${}^t(\text{Arct } \mathcal{X}_3, \text{Arct } \mathcal{X}_4)$ and ${}^t(\text{Arct } \mathcal{X}_4, \text{Arct } \mathcal{X}_3)$ are drawn in Table 7.4 as a short-dashed line and a line graph, respectively. We can not say that there exists certain notable causal relation between Wolfer's sunspot numbers and the rainfall averages in a year in Sapporo.

Figure 7.4. Arct (Sunspot, Rain in Sapporo)

The short-dashed line and the line graph in Figure 7.5 stand for the graphs of the functions $C_n(\text{Arct } \mathcal{X}_3 | \text{Arct } \mathcal{X}_5)$ and $C_n(\text{Arct } \mathcal{X}_5 | \text{Arct } \mathcal{X}_3)$, respectively. It is likely that Wolfer's sunspot numbers cause the temperature averages in a year in Sapporo.

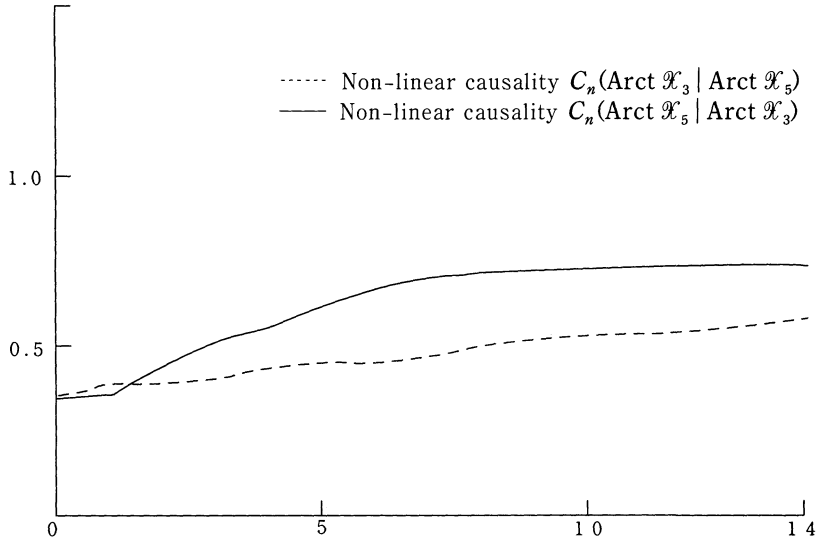


Figure 7.5. Arct (Sunspot, Temperature in Sapporo)

Moreover, Figure 7.6 indicates the graphs of the functions $C_n(\text{Arct } \mathcal{X}_4 | \text{Arct } \mathcal{X}_5)$ and $C_n(\text{Arct } \mathcal{X}_5 | \text{Arct } \mathcal{X}_4)$ as a short-dashed line and a line graph, respectively. It is likely that the rainfall averages in a year cause the temperature averages in a year in Sapporo.

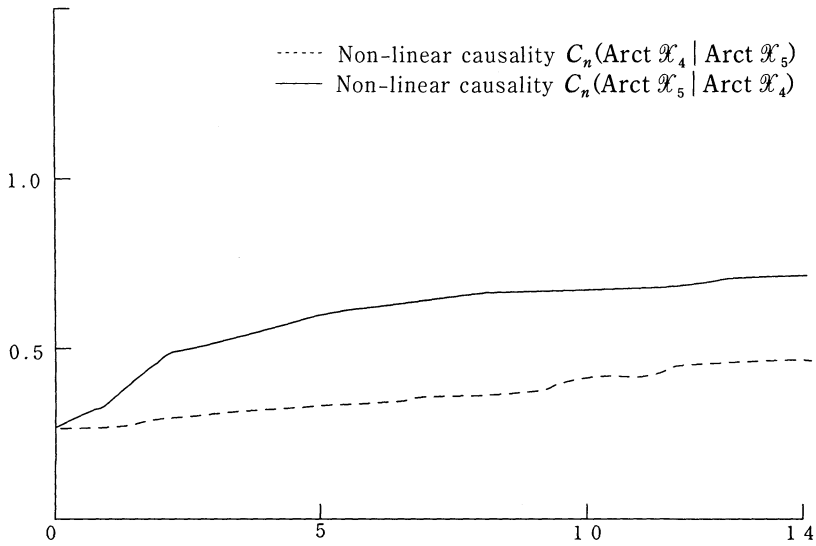


Figure 7.6. Arct (Rain in Sapporo, Temperature in Sapporo)

Finally, Figure 7.7 shows the graphs of the sample causality functions $C_n(\text{Arct } \mathcal{X}_3 | \text{Arct } \mathcal{X}_4, \text{Arct } \mathcal{X}_5)$, $C_n(\text{Arct } \mathcal{X}_4 | \text{Arct } \mathcal{X}_3, \text{Arct } \mathcal{X}_5)$ and $C_n(\text{Arct } \mathcal{X}_5 | \text{Arct } \mathcal{X}_3, \text{Arct } \mathcal{X}_4)$ as a long-dashed line, a short-dashed line and a line graph, respectively.

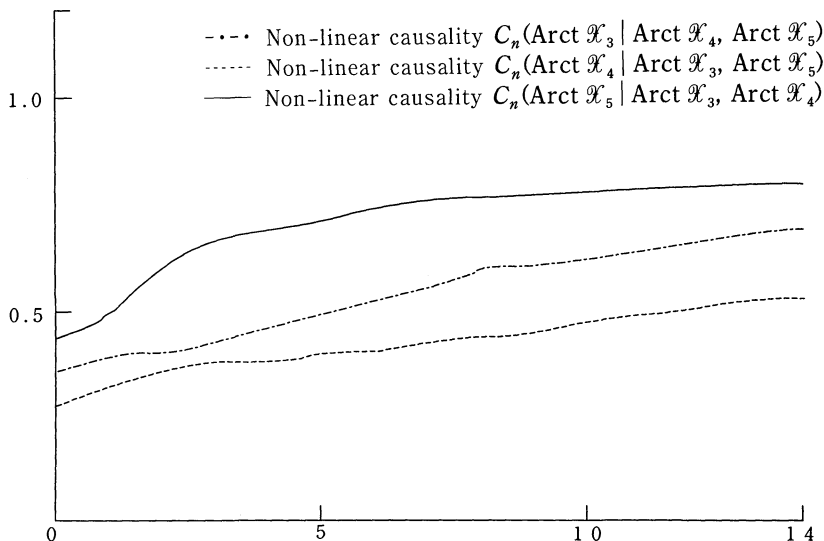


Figure 7.7. Arct (Sunspot, Rain in Sapporo, Temperature in Sapporo)

In order to take a positive and objective view of the existence of causal relation from the graph of the sample causality function, we need to obtain certain quantitative criterion for our **Test(CS)**, which will be proposed in a forthcoming paper ([34]), by using the same statistical reasoning as **Test(S)** in [31].

REFERENCES

- [1] Akaike, H. and T. Nakagawa, Statistical analysis and control for a dynamical system, Science Company, 1973 (in Japanese).
- [2] Box, G.E.P. and G.M. Jenkins, Time series analysis, Forecasting and Control, Holden-Day, 1976.
- [3] Durbin, J., The fitting of time series model, Rev. Int. Stat., **28** (1960), 233–244.
- [4] Elton, C. and M. Nicholson, The ten-year cycle in numbers of the Lynx in Canada, J. Animal Ecology, **11** (1942), 215–244.
- [5] Granger, C. W. J., Investigating causal relations by econometric models and cross-spectral methods, Econometrica, **37** (1969), 424–438.
- [6] Jevons, W. S., Commercial crises and sun-spots, Nature, **15** (1878), 33–37.
- [7] Kubo, R., Statistical mechanical theory of irreversible processes I, general theory and simple applications to magnetic and conduction problem, J. Phys. Soc. Japan, **12**

- (1957), 570–586.
- [8] Levinson, N., The Wiener RMS error criterion in filter design and prediction, *J. Math. Phys.*, **25** (1947), 261–278.
 - [9] Mata, C.G. and F.I. Shaffner, Solar and economic relationships; a preliminary report, *The Quarterly Journal of Economics*, Nov. (1934).
 - [10] Moran, P.A.P., The statistical analysis of the sunspot and Lynx cycles, *J. Animal Ecology*, **18** (1949), 115–116.
 - [11] —, The statistical analysis of the Canadian Lynx cycle I, structure and prediction, *J. Animal Ecology*, **22** (1951), 163–173.
 - [12] Nakano, Y. and Y. Okabe, On a multi-dimensional $[\alpha, \beta, \gamma]$ -Langevin equation, *Proc. Japan Acad.*, **59** (1983), 171–173.
 - [13] Okabe, Y., On a stationary Gaussian process with T -positivity and its associated Langevin equation and S -matrix, *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, **26** (1979), 115–165.
 - [14] —, On a stochastic differential equation for a stationary Gaussian process with T -positivity and the fluctuation-dissipation theorem, *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, **28** (1981), 169–213.
 - [15] —, On a stochastic differential equation for a stationary Gaussian process with finite multiple Markovian property and the fluctuation-dissipation theorem, *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, **28** (1982), 793–804.
 - [16] —, On a Langevin equation, *Sūgaku*, **33** (1981), 306–324 (in Japanese).
 - [17] —, On a wave equation associated with prediction errors for a stationary Gaussian process, *Lecture Notes in Control and Information Sciences*, **49** (1983), 215–226.
 - [18] —, A generalized fluctuation-dissipation theorem for the one-dimensional diffusion process, *Commun. Math. Phys.*, **98** (1985), 449–468.
 - [19] —, On KMO -Langevin equations for stationary Gaussian process with T -positivity, *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, **33** (1986), 1–56.
 - [20] —, On the theory of Brownian motion with the Alder-Wainwright effect, *J. Stat. Phys.*, **45** (1986), 953–981.
 - [21] —, KMO -Langevin equation and fluctuation-dissipation theorem (I), *Hokkaido Math. J.*, **15** (1986), 163–216.
 - [22] —, KMO -Langevin equation and fluctuation-dissipation theorem (II), *Hokkaido Math. J.*, **15** (1986), 317–355.
 - [23] —, Stokes-Boussinesq-Langevin equation and fluctuation-dissipation theorem, *Prob. Theory and Math. Stat.* (ed. by Prohorov et al.), VNU Science Press, 1986, vol. 2, 431–436.
 - [24] —, On long time tails of correlation functions for KMO -Langevin equations, *Proceeding of the Fourth Japan-USSR symposium on Probability theory*, Kyoto, July, 1986, *Lecture Notes in Mathematics*, Springer Vol. 1299, 391–397.
 - [25] —, On the theory of discrete KMO -Langevin equations with reflection positivity (I), *Hokkaido Math. J.*, **16** (1987), 315–341.
 - [26] —, On the theory of discrete KMO -Langevin equations with reflection positivity (II), *Hokkaido Math. J.*, **17** (1988), 1–44.
 - [27] —, On the theory of discrete KMO -Langevin equations with reflection positivity (III), *Hokkaido Math. J.*, **18** (1989), 149–174.
 - [28] —, On stochastic difference equations for the multi-dimentional weakly station-

- ary time series, Prospect of Algebraic Analysis (ed. by M. Kashiwara and T. Kawai), Academic Press, 1988, vol. 2, 601–645.
- [29] —, Langevin equation and Fluctuation-dissipation theorem, Stochastic Processes and their Applications (ed. by S. Albeverio et al.), Kluwer Academic Publishers, 1990, 275–299.
 - [30] —, 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. of Japan, **45** (1993), 277–294.
 - [31] Okabe, Y. and Y. Nakano, The theory of KM_2O -Langevin equations and its applications to data analysis (I): Stationary analysis, Hokkaido Math. J., **20** (1991), 45–90.
 - [32] Okabe, Y. and T. Ootsuka, 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. of Japan.
 - [33] —, The theory of KM_2O -Langevin equations and its applications to data analysis (III): Prediction analysis (in preparation).
 - [34] Y. Okabe, The theory of KM_2O -Langevin equations and its applications to data (IV): Causal analysis (2) (in preparation).
 - [35] Priestly, M. B., Non-linear and non-stationary time series, Academic Press, 1988.
 - [36] Walker, G., On periodicity in series of related terms, Proc. Roy. Soc. Ser. A, **131** (1931), 518–532.
 - [37] Whittle, P., On the fitting of multivariate autoregressions, and the approximate canonical factorization of a spectral density matrix, Biometrika, **50** (1963), 129–134.
 - [38] Wiggins, R.A. and E.A. Robinson, Recursive solution to the multichannel fitting problem, J. Geophys. Res., **70** (1965), 1885–1891.
 - [39] Yule, G.U., On a method of investigating periodicities in disturbed series, with special references to Wolfer's sunspot numbers, Phil. Trans. Roy. Soc. London Ser. A, **226** (1927), 268–298.

Department of Mathematics
Faculty of Science
Hokkaido University
Sapporo 060
Japan