Hokkaido Mathematical Journal Vol. 20 (1991) p. 45-90

The theory of KM_2O-Langevin equations and its applications to data analysis (I): Stationary analysis

Dedicated to Professor Nobuyuki Ikeda on his sixtieth birthday

Yasunori OKABE and Yuji NAKANO
(Received May 23, 1989, Revised October 5, 1990)

§ 1. Introduction

One of the authors has in a series of papers([20]-[34]) developed the theory of KMO-Langevin equations describing the time evolution of stationary Gaussian processes with reflection positivity in the discrete as well as continuous time case, and in [35] established the theory of KM_2O-Langevin equations for general weakly stationary time series. His original aim was two-fold:

1) Deeper understanding of the mathematical structure behind significant Kubo's fluctuation-dissipation theorem in non-equilibrium statistical physics([11]);

2) Applications of this theory to various fields of science through the universal and versatile nature of pure mathematics.

The purpose of this paper is to refine the results of [35] and create the more appropriate theory of KM_2O-Langevin equations for applications to data analysis; in fact, we discuss multi-dimensional weakly stationary time series whose time paramerter space is a finite interval of $\boldsymbol{Z}$. Further, we analyze a finite number of actual data representating such time series and propose a **Test**$(\boldsymbol{S})$ which is expected to be effective to verify the weak stationarity for them.

Various phenomena that generate random changes with the passing of time are observed and studied in natural science, engineering, economics, medical science and the like. In such phenomena, we obtain a finite number of actual observations, and as an important subject of scientific research, we wish to analyze them, to find a certain law behind them, to study their internal structure, and to forecast and control their future movements.

The theory of stochastic processes in pure mathematics provides us mathematical models with a certain law governing random phenomena and clarifies their universal structure. In the field of applied mathematics, on the other hand, almost all researchers in time series analysis use simplified models such as autoregressive(AR) or autoregressive and moving average(ARMA) models in the model fitting for random phenomena ([1]-[6], [12], [17], [39]-[42]). From the viewpoint of the theory of stochastic processes, AR(resp. ARMA) models can be characterized as time series with time parameter space $\boldsymbol{Z}$ that have two qualitative characters — the weak stationarity and the finite multiple Markovian property in the narrow(resp. wide) sense.

It turns out that we cannot conclude, through the analysis of a finite set of actual data observed in a random phenomenon with a discrete time parameter, that the time series representing the phenomenon posseses the weak stationarity as well as the finite multiple Markovian property; indeed, we need an infinite set of data for checking such properties. For this reason, the up-to-date time series analysis based upon AR or ARMA models has encountered a good deal of criticism saying "*just measurement without theory*", from econometricians who prefer traditional simultaneous equation models([38]).

Having in mind a true exchange between pure and applied sciences, we should refrain from assuming, in the model fitting problem, conditions beyond one's ability to verify. We are convinced that it is important for pure mathematicians to try to discover a certain essential law behind random phenomena and establish computer algorithms that are rooted in the appropriate mathematical theory.

The outline of the present paper is as follows: As the first part, we build in § 2 a theory of KM_2O-Langevin equations associated with multidimensional weakly stationary time series whose time parameter space is a finite interval of $\boldsymbol{Z}$. As an application of this theory to data analysis, the second part is divided into five steps from § 3 through § 7. At first, we introduce a sample forward KM_2O-Langevin equation(resp. data and force) associated with a d-dimensional data in § 3. Next in § 4, by using the KM_2O-Langevin force, we state a criterion that a d-dimensional data is a realization of a local and weakly stationary time series. And one more main thema in § 4 is to introduce a Test(S) whose effectiveness is proved by actual examples in § 4-§ 5. We finish this paper in § 6-§ 7 with a predictor formula depending upon the theory of KM_2O-Langevin equations.

Now let us state the detailed contents of this paper. Let $\boldsymbol{X}=(X(n)\,;\,|n|\leq N)$ be a d-dimensional weakly stationary process with mean vector zero and covariance function R:

$$(1.1) \qquad R(n)=E(X(n)\,{}^tX(0)) \quad (|n|\leq N),$$

where d and N are any fixed natural numbers. We call such a process $\boldsymbol{X}$ a local and weakly stationary time series. In subsection [2.1] of § 2, we extract two kinds of d-dimensional orthogonal time series $\boldsymbol{\nu}_+=(\nu_+(n);0\le n\le N)$ and $\boldsymbol{\nu}_-=(\nu_-(-n);0\le n\le N)$ from the original data $\boldsymbol{X}$, by taking the innovation approach([13]). It is noted that either of condition (2.5) and (2.6) holds for block Toeplitz matrices S_n composed of R.

Under the non-degenerate condition (2.5), we introduce in subsection [2.2] a system $\{\gamma_+(n,k), \gamma_-(n,k), \delta_+(m), \delta_-(m); 1\le k<n\le N, 1\le m\le N\}$ of members in $M(d;\boldsymbol{R})$ and establish the equations

$$(1.2)\qquad \begin{cases} X(0)=\nu_+(0) \\ X(n)=-\sum_{k=1}^{n-1}\gamma_+(n,k)X(k)-\delta_+(n)X(0)+\nu_+(n) \end{cases}$$

$$(1.3)\qquad \begin{cases} X(0)\ \ =\nu_-(0) \\ X(-n)=-\sum_{k=1}^{n-1}\gamma_-(n,k)X(-k)-\delta_-(n)X(0)+\nu_-(-n) \end{cases}$$

for any $n\in\{1,\cdots,N\}$. The time evolution of $\boldsymbol{X}$ is thus governed by the forward(resp. backward) equation (1.2)(resp. (1.3)) with dissipative(or deterministic) and fluctuating(or random) parts. The covariance matrix $V_+(n)$ (resp. $V_-(n)$) of the random force $\nu_+(n)$ (resp. $\nu_-(-n)$) depends upon n, because $\boldsymbol{\nu}_+$(resp. $\boldsymbol{\nu}_-$) is not always a white noise. We prove relations that determine the system $\{\gamma_+(n,k), \gamma_-(n,k), V_+(m), V_-(m); 1\le k<n\le N, 1\le m\le N\}$ in terms of $R(0)$ and $\{\delta_+(n), \delta_-(n); 1\le n\le N\}$(Theorem 2.2). These can be regarded as a kind of the **fluctuation-dissipation theorem** investigated in [20]-[35]. Furthermore, the latter quantities $\delta_+(n)$ and $\delta_-(n)$ can be calculated inductively from the covariance function R of $\boldsymbol{X}$(Theorem 2.3). We designate equation (1.2) (resp. (1.3)) and the random force $\boldsymbol{\nu}_+$(resp. $\boldsymbol{\nu}_-$) a **forward** (resp. **backward**) **KM_2O-Langevin equation** and **force** associated with **X**, respectively. The system $\{\gamma_+(n,k), \gamma_-(n,k), \delta_+(m), \delta_-(m), V_+(l), V_-(l); 1\le k<n\le N, 1\le m\le N, 0\le l\le N\}$ is called a **KM_2O-Langevin data** associated with R. We note that $\delta_+(m)$ and $\delta_-(m)$ $(1\le m\le N)$ correspond to the partial autocorrelation coefficients used in the fitting of AR models([6], [12], [40], [41]).

Conversely, suppose that we are given a system $\{\gamma_+(n,k), \gamma_-(n,k), \delta_+(m), \delta_-(m), V_+(l), V_-(l); 1\le k<n\le N, 1\le m\le N, 0\le l\le N\}$ of members in $M(d;\boldsymbol{R})$ satisfying the relations in Theorem 2.2 and a d-dimensional orthogonal time series $\boldsymbol{\nu}_+=(\nu_+(n);0\le n\le N)$ such that $E(\nu_+(n))=0$ and $E(\nu_+(n)^t\nu_+(n))=V_+(n)$. Then the forward KM_2O-Langevin equation (1.2) has a unique solution, denoted by $\boldsymbol{X}_+=(X(n);0\le n\le N)$. In subsection [2.4], this $\boldsymbol{X}_+$ is proved to be weakly stationary(Theorem 2.5).

The problem of non-linearity for one-dimensional time series is discussed in the final two subsections. We introduce in [2.6] two kinds of KM_2O-Langevin equations of non-linear type 2 and 3(Theorem 2.9 and Corollary 2.1). The final subsection [2.7] treats two examples of strictly stationary time series induced by the logistic and tent transformations in chaotic dynamical systems([8], [15], [18]).

As the first step in data analysis, we define in subsection [3.1] a sample mean vector $\mu^{\mathfrak{X}}$ and a sample covariance function $R^{\mathfrak{X}}$ for a given d-dimensional data $\mathfrak{X}=(\mathfrak{X}(n)\,;0\leq n\leq N)$. Considering a standardized data $\mathscr{X}=(\mathscr{X}(n)\,;0\leq n\leq N)$ of $\mathfrak{X}$, we introduce a **sample forward KM_2O-Langevin equation**(resp. **data** and **force**) associated with the original data $\mathfrak{X}$. In subsection [3.2], by taking the first difference data $\tilde{\mathfrak{X}}=(\mathfrak{X}(n)-\mathfrak{X}(n-1)\,;0\leq n\leq N)$ of a given d-dimensional data $\mathfrak{X}=(\mathfrak{X}(n)\,;-1\leq n\leq N)$, we apply the result in [3.1] to the data $\tilde{\mathfrak{X}}$ to form a sample first difference forward KM_2O-Langevin (resp. data and force) associated with $\tilde{\mathfrak{X}}$. We note that it is useful in data analysis to take the first difference of the original data([5]).

In subsection [3.3](resp. [3.4]), we treat a one-dimensional data $\mathscr{Y}=(\mathscr{Y}(n)\,;0\leq n\leq N)$(resp. $\mathscr{Y}_{-1}=(\mathscr{Y}(n)\,;-1\leq n\leq N)$) and apply the results in [3.1] and [2.6](resp. [3.1] and [3.2]) to $\mathscr{Y}$(resp. $\mathscr{Y}_{-1}$). Then, we get two kinds of sample(resp. sample first difference) forward KM_2O-Langevin equations of non-linear type 2 and 3 associated with $\mathscr{Y}$(resp. $\mathscr{Y}_{-1}$).

To compress the abnormal values of the original data, we may consider Arct tranformation. Actually, three kinds of transformations are introduced in the final three subsections [3.5]-[3.7]. For a d-dimensional data $\mathfrak{X}=(\mathfrak{X}(n)\,;0\leq n\leq N)$ such that all components $\mathfrak{X}_j(n)$ of $\mathfrak{X}(n)(1\leq j\leq d)$ are positive, we put $\mathrm{Log}\,\mathfrak{X}=({}^t(\log\mathfrak{X}_1(n),\cdots,\log\mathfrak{X}_d(n))\,;0\leq n\leq N)$ in [3.5]. This transformation is often used in the analysis of economic data. For a two-dimensional data $\mathfrak{X}=({}^t(\mathfrak{X}_1(n),\mathfrak{X}_2(n))\,;0\leq n\leq N)$, we define in [3.6] and [3.7] $\mathscr{X}_w=({}^t(\mathscr{X}_1(n),\mathscr{X}_2(n)+w\xi_u(n))\,;0\leq n\leq N)$ and $\mathrm{Arct}\,\mathscr{X}=({}^t(\arctan(\mathscr{X}_1(n)),\arctan(\mathscr{X}_2(n)))\,;0\leq n\leq N)$, respectively. Here $({}^t(\mathscr{X}_1(n),\mathscr{X}_2(n))\,;0\leq n\leq N)$(resp. $(\xi_u(n)\,;0\leq n\leq N)$) is the standardized data of the original data $\mathfrak{X}$(resp. the random uniform numbers in $(0,1)$). The value w in $\mathscr{X}_w$ is chosen from the unit interval$(0,1)$ and called a weight. The point is that the weak stationarity for the transformed data $\mathscr{X}_w$ implies the same property for the original data $\mathfrak{X}$. This procedure is necessary when condition (2.5) does not hold. The second transformation $\mathrm{Arct}\,\mathscr{X}$ has the advantage of compressing abnormal values in the original data $\mathfrak{X}$ and reproducing the weak stationarity. This is useful in causal analysis, which will be studied as a development of the present approach to data analysis([36]).

The second step of our data analysis is discussed in § 4. We first set up a criterion to decide whether or not any given d-dimensional data $\mathfrak{X}$ can be regarded as a realization of a local and weakly stationary time series that has the sample covariance function $R^{\mathfrak{X}}$ as its covariance function, by making repeated experiments for strictly stationary time series in the chaotic dynamical system argued in § 2. As in subsection [3.1], we consider in subsection [4.1] the sample forward KM_2O-Langevin data $\{\gamma_+(n, k), V_+(m); 0\leq k<n\leq N, 0\leq m\leq N\}$ associated with the given data $\mathfrak{X}=(\mathfrak{X}(n); 0\leq n\leq N)$. But an experiance rule in data analysis([3]) tells us that the number

$$(1.4) \qquad M+1=[3\sqrt{N+1}/d]$$

is a maximum effective length of the sample covariance function $R^{\mathscr{X}}$ of the standardized data $\mathscr{X}$ of $\mathfrak{X}$. Therefore, we have to use only the subsystem $\{\gamma_+(n, k), V_+(m); 0\leq k<n\leq M, 0\leq m\leq M\}$ as a reliable source in our data analysis. Further, we need so many copies of the original data that we construct, for each $i\in\{0, \cdots, N-M\}$, the part $\mathscr{X}_i=(\mathscr{X}(i+n); 0\leq n\leq M)$ with data number $M+1$. By the method in [3.1], we get thus d-dimensional data $\boldsymbol{\nu}_{+i}=(\nu_{+i}(n); 0\leq n\leq M)$ such that for any $n\in\{1, \cdots, M\}$,

$$(1.5) \qquad \begin{cases} \mathscr{X}(i) = \nu_{+i}(0) \\ \mathscr{X}(i+n)=-\sum_{k=1}^{n-1}\gamma_+(n, k)\mathscr{X}(i+k)-\delta_+(n)\mathscr{X}(i)+\nu_{+i}(n). \end{cases}$$

Now our problem is to decide which $\mathscr{X}_i$ can be regarded as a realization of a local and weakly stationary time series with $R^{\mathscr{X}}$ as its covariance function, and it is reduced to the same problem for the standardized data $\boldsymbol{\xi}_{+i}$ of $\boldsymbol{\nu}_{+i}$. Thus our test consists of three criteria given for $\boldsymbol{\xi}_{+i}$; $(M)_i$, $(V)_i$ and $(O)_i$ for checking mean zero, variance one and the orthogonality, respectively. Having done repeated experiments for several types of data obtained from random normal numbers, random uniform numbers, logistic and tent transformations, we are in a position to propose the following **Test**(S) : if the rates of $i\in\{0, \cdots, N\text{-}M\}$ such that each of $(M)_i$, $(V)_i$ and $(O)_i$ holds are over 80%, 70% and 80%, respectively, then we conclude that the local and weak stationarity is valid for the data $\mathscr{X}$ as well as for the original data $\mathfrak{X}$. The same procedure works in each situation stated in subsections [3.2]-[3.6]. In particular, Test(S) is called Test$(S)_{\mathrm{Log}}$, Test$(S)_w$ and Test$(S)_{\mathrm{Arct}}$, according as we perform it for the transformed data Log$\mathfrak{X}$, $\mathscr{X}_w$ and Arct$\mathscr{X}$. We show in Tables 4.

1-4.12 the results of these Test(S), Test$(S)_w$ and Test$(S)_{\mathrm{Arct}}$ for the concrete data stated above.

As the third step, we take up in § 5 three concrete data such as Wolfer's sunspot numbers, Lynx in MacKenzie River in Canada and NEC's stock prices in Japan and apply the procedure in § 4 to decide the validity of the local and weak stationarity for them. The results are illustrated in Tables 5.1-5.7. We then find that a two-dimensional time series composing of Wolfer's sunspot numbers and Lynx in MacKenzie River in Canada over the period of 114 years from 1821 to 1934 passes both Test(S) and Test$(S)_{\mathrm{Arct}}$.

The fourth step in § 6 treats the data $\mathscr{X}=(\mathscr{X}(n)\,;\,0\le n\le N)$ that passed our stationary Test(S) as well as $(M)_{N-M}$, $(V)_{N-M}$ and $(O)_{N-M}$ in § 4, and we construct a simulation $\hat{\mathscr{X}}_{N-M}=(\hat{\mathscr{X}}(N-M+n)\,;\,0\le n\le M)$ of the part $\mathscr{X}_{N-M}=(\mathscr{X}(N-M+n)\,;0\le n\le M)$ in each setting in subsections of § 3 and § 4.

The final fifth step in § 7 is to give some prediction formulae for the values in finite-step future of the data in each setting of subsections from § 3 to § 6. In particular, we consider the data of Wolfer's sunspot numbers from 1880 to 1979. It does not pass Test(S), but its first difference data does. Thus we can get in subsection [7. 5] KM_2O-predictors for nine years from 1980 to 1988, based upon the first difference forward KM_2O-Langevin equation, and compare them with the hidden actual observations, which are already known at present (1989). Further, by using the data of Wolfer's sunspot numbers from 1888 to 1988, we get KM_2O-predictors of Wolfer's sunspot numbers for nine years from 1989 to 1997, which are not yet known. The results are shown in Tables 7.1-7.4 and Figures 7.1-7.2.

In a forthcoming paper([36]), we will study the so-called causal relation between two given sets of data; our method is based upon the proposed Test(S). We believe that this will convince you the effectiveness of our approach to causal analysis.

The authors would like to thank the referees for their valuable and constructive advices.

§ 2. KM_2O-Langevin equations

In order to perform a data analysis based upon the concept of local and weak stationarity, we begin with describing a refinement of the theory of KM_2O-Langevin equations developed in [35]. Let d and N be any fixed natural numbers.

[2. 1] Let $\boldsymbol{X}=(X(n)\,;|n|\le N)$ be any d-dimensional local and weakly stationary time series on a probability space $(\Omega, \mathscr{B}, P)$ with covariance function R:

$$(2.1)\qquad R(n)=E(X(n)\,{}^tX(0)) \quad (|n|\leq N).$$

It is noted that

$$(2.2)\qquad {}^tR(n)=R(-n) \qquad (|n|\leq N).$$

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

$$(2.3)\qquad S_n=\begin{pmatrix} R(0) & R(1) & R(2) & \cdots\cdots & R(n-1) \\ {}^tR(1) & R(0) & R(1) & \cdots\cdots & \cdots \\ & \cdots & & \cdots\cdots & \cdots \\ & \cdots & & \cdots\cdots & \cdots \\ & \cdots & & \cdots\cdots & \cdots \\ {}^tR(n-2) & & & \cdots\cdots & R(0) \quad R(1) \\ {}^tR(n-1) & & & \cdots\cdots & {}^tR(1) \quad R(0) \end{pmatrix}.$$

In this section, we assume

$$(2.4)\qquad R(0)\in GL(d\,;\boldsymbol{R}).$$

Then we can see that either of the following (2.5) and (2.6) holds:

(2.5) $\quad S_n\in GL(nd\,;\boldsymbol{R})$ for any $n\in\{1,\cdots,N\}$.

(2.6) There exists $n\in\{1,\cdots,N-1\}$ such that $\det(S_n)=0$. In this case,
$S_n\in GL(nd\,;\boldsymbol{R})$ for any $n\in\{1,\cdots,N_0\}$,
$S_n\notin GL(nd\,;\boldsymbol{R})$ for any $n\in\{N_0+1,\cdots,N\}$,

where $N_0=\max\{n\in\{1,\cdots,N-1\}\,;\det(S_n)\neq 0\}$.

Let $\boldsymbol{M}$, $\boldsymbol{M}_0^+(n)$ and $\boldsymbol{M}_0^-(n)\,(0\leq n\leq N)$ be the closed linear subspaces of $L^2(\Omega, \mathscr{B}, P)$ defined by

$$(2.7)\qquad \boldsymbol{M}=\text{the closed linear hull of } \{X_j(m)\,;1\leq j\leq d,\ |m|\leq N\}$$
$$(2.8)\qquad \boldsymbol{M}_0^+(n)=\text{the closed linear hull of } \{X_j(m)\,;1\leq j\leq d,\ 0\leq m\leq n\}$$
$$(2.9)\qquad \boldsymbol{M}_0^-(n)=\text{the closed linear hull of } \{X_j(-m)\,;1\leq j\leq d,\ 0\leq m\leq n\},$$

where $X(m)={}^t(X_1(m),\cdots,X_d(m))\,(|m|\leq N)$. Then we introduce two d-dimensional time series $\boldsymbol{\nu}_+=(\nu_+(n)\,;0\leq n\leq N)$ and $\boldsymbol{\nu}_-=(\nu_-(-n)\,;0\leq n\leq N)$ by

$$(2.10)\qquad \nu_+(n)=X(n)-P_{\boldsymbol{M}_0^+(n-1)}X(n)$$
$$(2.11)\qquad \nu_-(-n)=X(-n)-P_{\boldsymbol{M}_0^-(n-1)}X(-n),$$

where $\boldsymbol{M}_0^+(-1)=\boldsymbol{M}_0^-(-1)=\{0\}$ and $P_{\boldsymbol{M}_0^+(n-1)}$ (resp. $P_{\boldsymbol{M}_0^-(n-1)}$) stands for the orthogonal projection on the space $M_0^+(n-1)$(resp. $M_0^-(n-1)$). It is immediate to see the following:

(2.12) $\nu_+(0)=\nu_-(0)=X(0)$

(2.13) $\boldsymbol{\nu}_+$ and $\boldsymbol{\nu}_-$ are both orthogonal time series with mean vector zero

(2.14) $M_0^+(n)$=the closed linear hull of $\{\nu_{+j}(m)\,;1\leq j\leq d,\,0\leq m\leq n\}$

(2.15) $M_0^-(n)$=the closed linear hull of $\{\nu_{-j}(-m)\,;1\leq j\leq d,\,0\leq m\leq n\}$,

where $\nu_+(m)={}^t(\nu_{+1}(m),\cdots,\nu_{+d}(m))$ and $\nu_-(-m)={}^t(\nu_{-1}(-m),\cdots,\nu_{-d}(-m))$. We denote by $V_+(n)$(resp. $V_-(n)$) the covariance matrix of $\nu_+(n)$(resp. $\nu_-(-n)$)$(0\leq n\leq N)$:

(2.16) $V_+(n)=E(\nu_+(n){}^t\nu_+(n))$

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

[2.2] This subsection treats the case where condition (2.5) holds. Similarly to (2.16) and (2.17) in [35], we have

THEOREM 2.1. *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* $M(d\,;\boldsymbol{R})$ *such that for any* $n\in\{1,\cdots,N\}$,

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

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

We call equation (2.18)(resp. (2.19)) a **forward** (resp. **backward**) **KM$_2$O-Langevin equation** for $\boldsymbol{X}$. Further, the random force $\boldsymbol{\nu}_+$(resp. $\boldsymbol{\nu}_-$) is said to be a **forward** (resp. **backward**) **KM$_2$O-Langevin force** associated with $\boldsymbol{X}$. Moreover, we designate the system $\{\gamma_+(n,k),\,\gamma_-(n,k),\,\delta_+(m),\,\delta_-(m),\,V_+(l),\,V_-(l)\,;\,1\leq k<n\leq N,\,1\leq m\leq N,\,0\leq l\leq N\}$ a **KM$_2$O-Langevin data** associated with the covariance function R of $\boldsymbol{X}$.

Since the proofs in Theorems 3.1 and 4.1 of [35] can be applied to our local time series $\boldsymbol{X}$, we obtain the fundamental recursive relations among the KM$_2$O-Langevin data associated with R.

THEOREM 2.2. *For any* $n,\,k\in\boldsymbol{N}$, $1\leq k<n\leq N$,

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

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

(2.22) $V_+(n)=(\mathrm{I}-\delta_+(n)\delta_-(n))V_+(n-1)$

(2.23) $V_-(n)=(\mathrm{I}-\delta_-(n)\delta_+(n))V_-(n-1)$

(2.24) $\delta_-(n)V_+(n-1)=V_-(n-1){}^t\delta_+(n)$

(2.25) $\delta_-(n)V_+(n)=V_-(n){}^t\delta_+(n)$, *where*

(2.26) $\gamma_+(n, 0)=\delta_+(n)$ *and* $\gamma_-(n, 0)=\delta_-(n)$.

The relations (2.20)-(2.23) imply that $\gamma_+(\cdot,*)$, $\gamma_-(\cdot,*)$, $V_+(\cdot)$ and $V_-(\cdot)$ can be determined by $\delta_+(\cdot)$, $\delta_-(\cdot)$ and $R(0)$. In particular, we note that the relations (2.22) and (2.23) correspond to the **generalized second fluctuation-dissipation theorem** based upon the first KMO-Langevin equation ([21], [27], [32], [33]).

By Lemma 4.1(i) in [35],

(2.27) $$\det S_n=\prod_{k=0}^{n-1}\det V_+(k) \quad (1\le n\le N),$$

hence it follows from condition (2.5) that

(2.28) $$V_+(n)\in GL(d\,;\boldsymbol{R}) \qquad (0\le n\le N-1).$$

Similarly to Lemma 4.2 in [35], we can get an algorithm for calculating the fundamental quantities $\delta_+(\cdot)$ and $\delta_-(\cdot)$ from the covariance function R.

THEOREM 2.3. *For any* $n\in\boldsymbol{N}$, $1\le n\le N$,

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

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

REMARK 2.1.

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

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

REMARK 2.2. It follows from (2.24), (2.29) and (2.30) that

(2.33) $$\sum_{k=0}^{n-1}\gamma_+(n, k)R(k+1)=\sum_{k=0}^{n-1}R(k+1){}^t\gamma_-(n, k) \quad (1\le n\le N).$$

In fact, this relation (2.33) can be proved similarly to Lemma 4.3 in [35], which played an important role in the proof of (2.24).

REMARK 2.3. When $d=1$, we can see that

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

[2.3] This subsection treats any one-dimensional weakly stationary time series $\boldsymbol{X}=(X(n);|n|\leq N)$ for which condition (2.6) holds. Then we show

THEOREM 2.4. *There exists a unique system* $\{\gamma(n, k), \delta(m); 1\leq k< n\leq N_0, 1\leq m\leq N_0\}$ *of real numbers such that*

(i) *for any* $n\in\{1, \cdots, N_0-1\}$,

$$X(n)=-\sum_{k=1}^{n-1}\gamma(n, k)X(k)-\delta(n)X(0)+\nu_+(n) \tag{2.35}$$

$$X(-n)=-\sum_{k=1}^{n-1}\gamma(n, k)X(-k)-\delta(n)X(0)+\nu_-(-n) \tag{2.36}$$

(ii) *for any* $n\in\{N_0, \cdots, N\}$,

$$X(n)=-\sum_{k=1}^{n-1}\gamma(N_0, N_0-n+k)X(k)-\delta(N_0)X(n-N_0) \tag{2.37}$$

$$X(-n)=-\sum_{k=1}^{n-1}\gamma(N_0, N_0-n+k)X(-k)-\delta(N_0)X(-n+N_0) \tag{2.38}$$

(iii) *for any* $n, k\in\boldsymbol{N}$, $1\leq k<n\leq N_0$,

$$\gamma(n, k)=\gamma(n-1, k-1)+\delta(n)\gamma(n-1, n-k-1) \tag{2.39}$$

$$V(n)=(1-\delta(n)^2)V(n-1) \tag{2.40}$$

$$\delta(n)=-(R(n)+\sum_{k=0}^{n-2}\gamma(n-1, k)R(k+1))V(n-1)^{-1} \tag{2.41}$$

$$|\delta(n-1)|<1 \tag{2.42}$$

$$|\delta(N_0)|=1, \tag{2.43}$$

where $V(n)=V_+(n)=V_-(n)$ *and* $\gamma(n, 0)=\delta(n)$.

PROOF. Since we can apply the results in subsections [2.1] and [2.2] up to time N_0, it is sufficient to prove

$$V(N_0)=0. \tag{2.44}$$

Suppose that $V(N_0)\neq 0$. We then claim that

(2.45) $\{X(n); 0\leq n\leq N_0\}$ is linearly independent in $\boldsymbol{M}$.

Let $c_n\in\boldsymbol{R}(0\leq n\leq N_0)$ such that $\sum_{n=0}^{N_0}c_nX(n)=0$. Since

$$X(N_0)=-\sum_{k=0}^{N_0-1}\gamma(N_0, k)X(k)+\nu_+(N_0),$$

we have

$$\sum_{k=0}^{N_0-1}(c_k-c_{N_0}\gamma(N_0, k))X(k)+c_{N_0}\nu_+(N_0)=0.$$

Multiplying both hand sides by $\nu_+(N_0)$ and then taking the expectation with respect to P, we see from (2.13) and (2.14) that $c_{N_0}V_+(N_0)=0$ and so $c_{N_0}=0$. Similarly, $c_n=0(0\leq n\leq N_0-1)$. Thus, (2.45) was proved.

Since

$$S_{N_0+1}=E((X(N_0),\cdots,X(0))^t(X(N_0),\cdots,X(0))), \tag{2.46}$$

it follows from (2.45) that $S_{N_0+1}\in GL(d(N_0+1);\boldsymbol{R})$, which contradicts condition (2.6). Therefore, we have proved (2.44). (Q. E. D.)

REMARK 2.4. In the one-dimensional case, we find from Theorem 2.4 that once δ takes the value 1 or -1 at some time N_0, the time evolution of $\boldsymbol{X}$ becomes deterministic after the time N_0.

[2.4] Conversely to subsection [2.1], assume that we are given any system $\{V,\delta_+(n);1\leq n\leq N\}$ of members in $M(d;\boldsymbol{R})$ such that V is symmetric and positive definite. Then we can construct a triple $(V_+(1),\delta_-(1),V_-(1))$ by

$$\begin{cases} V_+(1)=V-\delta_+(1)V{}^t\delta_+(1) \\ \delta_-(1)=V{}^t\delta_+(1)V^{-1} \\ V_-(1)=V-\delta_-(1)V{}^t\delta_-(1). \end{cases} \tag{2.47}$$

In order to continue the following construction of $(V_+(n),\delta_-(n),V_-(n))$ from $(V_+(n-1),\delta_-(n-1),V_-(n-1))(2\leq n\leq N)$:

$$\begin{cases} V_+(n)=V_+(n-1)-\delta_+(n)V_-(n-1){}^t\delta_+(n) \\ \delta_-(n)V_+(n-1)=V_-(n-1){}^t\delta_+(n) \\ V_-(n)=V_-(n-1)-\delta_-(n)V_+(n-1){}^t\delta_-(n), \end{cases} \tag{2.48}$$

we suppose that

$$V_+(n-1)\in GL(d;\boldsymbol{R})\quad(1\leq n\leq N), \tag{2.49}$$

where $V_+(0)=V$. Furthermore, we assume that

$$V_+(n)\text{ are non-negative definite}\quad(1\leq n\leq N). \tag{2.50}$$

REMARK 2.5. When $d=1$, we can start with a system $\{V,\delta(n);1\leq n\leq N\}$ such that

$$V>0 \tag{2.51}$$
$$\delta(n)\in[-1,1]\quad(1\leq n\leq N). \tag{2.52}$$

Then we define $V(n)$ by

$$V(n)=\left\{\prod_{k=1}^{n}(1-\delta(k)^2)\right\}V\quad(1\leq n\leq N), \tag{2.53}$$

which satisfies conditions (2.49) and (2.50) if $|\delta(n)|<1(1\leq n\leq N-1)$.

Next we construct a system $\{\gamma_+(m, n), \gamma_-(m, n); 0\leq n<m\leq N\}$ of members in $M(d;\boldsymbol{R})$ according to the algorithm (2.20) and (2.21) with (2.26).

Finally, for any d-dimensional time series $\boldsymbol{\nu}_+=(\nu_+(n); 0\leq n\leq N)$ on a probability space $(\Omega, \mathscr{B}, P)$ such that for any $m, n\in \boldsymbol{N}^*$, $0\leq m, n\leq N$,

$$(2.54)\quad E(\nu_+(n))=0$$

$$(2.55)\quad E(\nu_+(m)^t\nu_+(n))=\delta_{mn}V_+(n),$$

we construct a d-dimensional time series $\boldsymbol{X}_+=(X(n); 0\leq n\leq N)$ by the following recursive relation:

$$(2.56)\quad X(0)=\nu_+(0)$$

$$(2.57)\quad X(n)=-\sum_{k=1}^{n-1}\gamma_+(n, k)X(k)-\delta_+(n)X(0)+\nu_+(n)\quad (1\leq n\leq N).$$

Then, similarly to Theorem 6.1 in [35], we have

THEOREM 2.5. *X_+ is a weakly stationary time series with a given KM_2O-Langevin data.*

[2.5] In this subsection we obtain a prediction formula based upon the forward KM_2O-Langevin equation (2.57).

THEOREM 2.6. *For any $m, n\in \boldsymbol{N}^*$, $0\leq n<m\leq N$,*

$$(2.58)\quad P_{\boldsymbol{M}_0^+(n)}X(m)=\sum_{k=0}^{n}Q_+(m, n; k)X(k),$$

where the coefficient matrices $Q_+(m, n; k)$ are given by the following recursive relation$(0\leq n<n+1<m\leq N)$:

$$(2.59)\quad \begin{cases} Q_+(n+1, n; k)=-\gamma_+(n+1, k) \\ Q_+(m, n; k)=\sum_{l=n+1}^{m-1}\gamma_+(m, l)Q_+(l, n; k)-\gamma_+(m, k). \end{cases}$$

PROOF. It is noted from (2.56) and (2.57) that (2.14) holds. Hence, we see from (2.55) that (2.58) holds for $m=n+1$. Next, by mathematical induction, we assume that (2.58) holds for any $m\in\{n+1, \cdots, m_0\}$. Then by (2.57),

$$\begin{aligned} P_{\boldsymbol{M}_0^+(n)}X(m_0+1) &= -\sum_{k=0}^{n}\gamma_+(m_0+1, k)X(k) \\ &\qquad -\sum_{l=n+1}^{m_0}\gamma_+(m_0+1, l)P_{\boldsymbol{M}_0^+(n)}X(l) \\ &= \sum_{k=0}^{n}(\sum_{l=n+1}^{m_0}\gamma_+(m_0+1, l)Q_+(l, n; k) \\ &\qquad -\gamma_+(m_0+1, k)))X(k), \end{aligned}$$

which implies that (2.58) holds also for $m=m_0+1$. Thus, we complete the proof of Theorem 2.6. (Q. E. D.)

Then, for any $m, n\in N^*$, $0\leq n<m\leq N$, we define a prediction error matrix $e_+(m, n)$ by

(2.60) $$e_+(m, n)=E((X(m)-P_{M_0^+(n)}X(m))^t(X(m)-P_{M_0^+(n)}X(m))).$$

Similarly to Theorem 5.1(iii) in [35], we have

THEOREM 2.7. *For any* $m, n\in N^*$, $0\leq n<m\leq N$,

(2.61) $$e_+(m, n)=\sum_{k=n+1}^{m}P_+(m, k)^tP_+(m, k),$$

where matrices $P_+(m, k)(1\leq k<m\leq N)$ *are given by the following recursive relation*:

(2.62) $$\begin{cases} P_+(k, k)=V_+(k)^{1/2} \\ P_+(m, k)=-\sum_{l=k}^{m-1}\gamma_+(m, l)P_+(l, k). \end{cases}$$

As a consequence of (2.48) and (2.61), it is easy to get

THEOREM 2.8. *For any* $n\in\{0, \cdots, N-1\}$,

$$e_+(n+1, n)=(I-\delta_+(n+1)\delta_-(n+1))\cdots(I-\delta_+(1)\delta_-(1))R(0).$$

[2.6] This subsection aims to introduce a class of non-linear Langevin equations. For any one-dimensional stochastic process $\boldsymbol{Y}=(Y(n); 0\leq n\leq N)$ on a probability space $(\Omega, \mathscr{B}, P)$ such that

(2.63) $$Y(n)\in L^6(\Omega, \mathscr{B}, P) \quad (0\leq n\leq N),$$

we define a one-dimensional time series $\boldsymbol{X}^{(1)}=(X^{(1)}(n); 0\leq n\leq N)$ and two-dimensional time series $\boldsymbol{X}^{(p)}=(X^{(p)}(n); 0\leq n\leq N)(2\leq p\leq 3)$ by

(2.64) $$X^{(1)}(n)=Y(n)-E(Y(n))$$

(2.65)$_p$ $$X^{(p)}(n)=\begin{pmatrix} Y(n)-E(Y(n)) \\ Y(n)^p-E(Y(n)^p) \end{pmatrix}.$$

THEOREM 2.9. *If* $\boldsymbol{X}^{(p)}(1\leq p\leq 3)$ *are all weakly stationary time series with condition* (2.5), *then there exist uniquely three kinds of systems* $\{\gamma_+^{(1)}(n, k); 0\leq k<n\leq N\}$ *and* $\{\gamma_{+i}^{(p)}(n, k); 1\leq i\leq 2,\ 0\leq k<n\leq N\}(2\leq p\leq 3)$ *consisting of real numbers and three kinds of one-dimensional time series* $\boldsymbol{\nu}_+^{(p)}=(\nu_+^{(p)}(n); 0\leq n\leq N)(1\leq p\leq 3)$ *such that*

(i) *for any* $n\in\{1, \cdots, N\}$,

$$(2.66)\quad \begin{cases} Y(0)-E(Y(0))=\nu_+^{(1)}(0) \\ Y(n)-E(Y(n))=-\sum_{k=0}^{n-1}\gamma_+^{(1)}(n,k)(Y(k)-E(Y(k)))+\nu_+^{(1)}(n) \end{cases}$$

(2.67) $\boldsymbol{\nu}_+^{(1)}$ *is an orthogonal time series with mean zero*

(2.68) $E(Y(m)\nu_+^{(1)}(n))=0 \quad (0\leq m\leq n-1)$

(2.69) *the closed linear hull of* $\{Y(m)-E(Y(m))\,;0\leq m\leq n\}$
$=$ *the closed linear hull of* $\{\nu_+^{(1)}(m)\,;0\leq m\leq n\}$

(ii) *for each* $p\in\{2,3\}$ *and any* $n\in\{1,\cdots,N\}$,

$$(2.70)_p\quad \begin{cases} Y(0)-E(Y(0))=\nu_+^{(p)}(0) \\ Y(n)-E(Y(n))=-\sum_{k=0}^{n-1}\gamma_{+1}^{(p)}(n,k)(Y(k)-E(Y(k))) \\ \qquad\qquad -\sum_{k=0}^{n-1}\gamma_{+2}^{(p)}(n,k)(Y(k)^p-E(Y(k)^p))+\nu_+^{(p)}(n) \end{cases}$$

(2.71) $\boldsymbol{\nu}_+^{(p)}$ *is an orthogonal time series with mean zero*

(2.72) $E(Y(m)\nu_+^{(p)}(n))=E(Y(m)^p\nu_+^{(p)}(n))=0 \quad (0\leq m\leq n-1)$

(2.73) $\sigma(Y(m)\,;0\leq m\leq n)=\sigma(\nu_+^{(p)}(m)\,;0\leq m\leq n)$.

PROOF. We first note from (2.6) in [35] that the condition (2.5) for $\boldsymbol{X}^{(p)}$ is equivalent to

(2.74) $\{Y(m)-E(Y(m)),\ Y(m)^p-E(Y(m)^p)\,;0\leq m\leq N-1\}$ is linearly independent in $L^2(\Omega,\mathscr{B},P)$.

Hence, the proof of (i) is immediate.

For the proof of (ii), let us fix any $p\in\{2,3\}$. Concerning the existence, Theorem 2.1 assures us that there exist a system $\{\gamma_+^{(p)}(n,k)\,;\ 0\leq k<n\leq N\}$ of members in $M(2\,;\boldsymbol{R})$ and a two-dimensional orthogonal time series $({}^t(\nu_{+1}^{(p)}(n),\ \nu_{+2}^{(p)}(n))\,;0\leq n\leq N)$ such that for any $n\in\{1,\cdots,N\}$,

$$(2.75)\quad \begin{cases} X^{(p)}(0)={}^t(\nu_{+1}^{(p)}(0),\ \nu_{+2}^{(p)}(0)) \\ X^{(p)}(n)=-\sum_{k=0}^{n-1}\gamma_+^{(p)}(n,k)X^{(p)}(k)+{}^t(\nu_{+1}^{(p)}(n),\ \nu_{+2}^{(p)}(n)). \end{cases}$$

We set

$$\gamma_{+1}^{(p)}(n,k)=\gamma_{+11}^{(p)}(n,k),\ \gamma_{+2}^{(p)}(n,k)=\gamma_{+12}^{(p)}(n,k) \text{ and } \nu_+^{(p)}(n)=\nu_{+1}^{(p)}(n).$$

Taking the first component of both hand sides in (2.75), we see that the system $\{\gamma_{+i}^{(p)}(n,k)\,;1\leq i\leq 2,\,0\leq k<n\leq N\}$ and the time series $\boldsymbol{\nu}_+^{(p)}=(\nu_+^{(p)}(n)\,;0\leq n\leq N)$ satisfy $(2.70)_p$ and (2.71).

Since (2.72) can be shown from (2.13), (2.14) and $(2.65)_p$, we turn to the proof of (2.73). It is clear to see from $(2.70)_p$ that the right hand side of (2.73) is contained in the left hand side of (2.73). Since $Y(0)=$

$E(Y(0))+\nu_+^{(p)}(0)$, it holds that $Y(0)$ and $Y(0)^p$ are $\sigma(\nu_+^{(1)}(0))$-measurable. Hence, we see from $(2.70)_p$ that $Y(1)$ is $\sigma(\nu_+^{(1)}(0), \nu_+^{(1)}(1))$-measurable and $Y(1)^2$ is so. In this way, a repeated use of $(2.70)_p$ implies that $Y(n)$ is $\sigma(\nu_+^{(1)}(m); 0\le m\le n)$-measurable for any $n\in\{0, \cdots, N\}$. Thus we have (2.73).

Concerning the uniqueness, let $\{\tilde{\gamma}_{+i}^{(p)}(n, k); 1\le i\le 2, 0\le k<n\le N\}$ and $\tilde{\boldsymbol{\nu}}_+^{(p)}=(\tilde{\nu}_+^{(p)}(n); 0\le n\le N)$ be another objects satisfying $(2.70)_p$–(2.73). By taking the second component of (2.75),

$$Y(n)^p-E(Y(n)^p)=-\sum_{k=0}^{n-1}\gamma_{+21}^{(p)}(n, k)(Y(k)-E(Y(k))) -\sum_{k=0}^{n-1}\gamma_{+22}^{(p)}(n, k)(Y(k)^p-E(Y(k)^p))+\nu_{+2}^{(p)}(n).$$

We set

$$\tilde{\gamma}_+^{(p)}(n, k)=\begin{pmatrix}\tilde{\gamma}_{+1}^{(p)}(n, k) & \tilde{\gamma}_{+2}^{(p)}(n, k)\\ \gamma_{+21}^{(p)}(n, k) & \gamma_{+22}^{(p)}(n, k)\end{pmatrix} \text{ and } \mu_+^{(p)}(n)={}^t(\tilde{\nu}_+^{(p)}(n), \nu_{+2}^{(p)}(n)).$$

Then it can be seen that for any $n\in\{1, \cdots, N\}$,

$$(2.76)\quad \begin{cases} X^{(p)}(0)=\mu_+^{(p)}(0)\\ X^{(p)}(n)=-\sum_{k=0}^{n-1}\tilde{\gamma}_+^{(p)}(n, k)X^{(p)}(k)+\mu_+^{(p)}(n). \end{cases}$$

Since for any $m, n\in\{0, \cdots, N\}$, $0\le m<n\le N$, $E(X^{(p)}(m)\,{}^t\mu_+^{(p)}(n))=0$, it follows from (2.76) that for any $n\in\{1, \cdots, N\}$,

$$(2.77)\quad \mu_+^{(p)}(n)=X^{(p)}(n)-P_{\boldsymbol{M}_0^{+(p)}(n-1)}X^{(p)}(n),$$

where

$$(2.78)\quad \boldsymbol{M}_0^{+(p)}(n-1)=\text{the closed linear hull of } \{Y(m)-E(Y(m)), Y(m)^p-E(Y(m)^p); 0\le m\le n-1\}.$$

Hence, by (2.10) and (2.77), we find that

$$(2.79)\quad \mu_+^{(p)}(n)=\nu_+^{(p)}(n)\quad (0\le n\le N).$$

Furthermore, it follows from (2.75), (2.76) and (2.79) that

$$\sum_{k=0}^{n-1}\gamma_+^{(p)}(n, k)X^{(p)}(k)=\sum_{k=0}^{n-1}\tilde{\gamma}_+^{(p)}(n, k)X^{(p)}(k)\quad (1\le n\le N)$$

and so by (2.74)

$$\gamma_+^{(p)}(n, k)=\tilde{\gamma}_+^{(p)}(n, k)\quad (0\le k<n\le N),$$

which implies that $\gamma_{+i}^{(p)}(n, k)=\tilde{\gamma}_{+i}^{(p)}(n, k)\quad (1\le i\le 2, 0\le k<n\le N)$.

Thus we have completed the proof of Theorem 2.9. (Q. E. D.)

COROLLARY 2.1. *For each* $p\in\{1,2,3\}$,

$$\nu_+^{(p)}(n)=Y(n)-P_{M_0^{+(p)}(n-1)}Y(n) \quad (1\leq n\leq N),$$

where $M_0^{+(p)}(n-1)$ *is defined by* (2.78).

REMARK 2.6. For each $p\in\{2,3\}$, the coefficients $\gamma_+^{(1)}(n,k)$(resp. $\gamma_{+i}^{(p)}(n,k)$, $1\leq i\leq 2$)$(0\leq k<n\leq N)$ in equation (2.66)(resp. $(2.70)_p$) can be calculated inductively from the covariance function of $X^{(1)}$(resp. $X^{(p)}$), according to the algorithm in Theorems 2.2 and 2.3.

The equation (2.66)(resp. $\{\gamma_+^{(1)}(n,k);0\leq k<n\leq N\}$ and $\nu_+^{(1)}$) is nothing but the forward KM_2O-Langevin equation(resp. data and force) associated with $X^{(1)}$. For each $p\in\{2,3\}$, equation $(2.70)_p$(resp. $\{\gamma_{+i}^{(p)}(n,k);1\leq i\leq 2, 0\leq k<n\leq N\}$ and $\nu_+^{(p)}$) is called a **forward KM_2O-Langevin equation** (resp. **data** and **force**) of **non-linear type p** associated with $X^{(1)}$.

REMARK 2.7. In the same situation as in Theorem 2.9, we define three kinds of prediction errors $e_+^{(p)}(m,n)$$(0\leq n\leq m\leq N, 1\leq p\leq 3)$ by

$$(2.80) \qquad e_+^{(p)}(m,n)=E(|Y(m)-P_{M_0^{+(p)}(n)}Y(m)|^2).$$

Then we have

$$(2.81) \qquad e_+^{(2)}(m,n)\leq e_+^{(1)}(m,n) \text{ and } e_+^{(3)}(m,n)\leq e_+^{(1)}(m,n).$$

It would be a serious task to get a relation between $e_+^{(2)}(m,n)$ and $e_+^{(3)}(m,n)$, unless we assume a further structure for the time series $\boldsymbol{Y}$.

REMARK 2.8. By treating the non-linear type other than $(2.65)_p$, we can derive several kinds of non-linear Langevin equations different from $(2.70)_p$, associated with $X^{(1)}$, which will be used in [36].

[2.7] Among various one-dimensional strictly stationary time series induced by transformations in chaotic dynamical systems([15], [18]), we consider the following two special cases:

EXAMPLE 2.1. Let φ_l be the logistic transformation on $[0,1]$, i.e.,

$$(2.82) \qquad \varphi_l(x)=4x(1-x).$$

It is known([18]) that φ_l is a Kolmogorov transformation with the unique invariant probability measure P_l given by

$$(2.83) \qquad P_l(dx)=\frac{1}{\pi\sqrt{x(1-x)}}dx.$$

This transformation φ_l has been studied as a difference model of some ecological system.

We define a strictly stationary time series $\boldsymbol{Y}_l=(Y_l(n)\,;0\leq n<\infty)$ on the probability space $([0,1],\mathscr{B}[0,1],P_l)$ by

$$(2.84)\qquad Y_l(n)(x)=(\overbrace{\varphi_l\circ\cdots\circ\varphi_l}^{n\text{ times}})(x).$$

Then we know([18]) that

$$(2.85)\qquad \begin{aligned} R_l(n)&\equiv E((Y_l(n)-E(Y_l(n)))(Y_l(0)-E(Y_l(0))))\\ &=8^{-1}\delta_{0,n}, \end{aligned}$$

which implies that $(Y_l(n)-E(Y_l(n))\,;0\leq n<\infty)$ is a white noise.

EXAMPLE 2.2. By changing the roundish shape of the curve φ_l into the tented curve and then shifting the position of the peak from 1/2 to some $p\in(0,1)$, we define a mapping $\varphi_{t,p}$ on $[0,1]$ by

$$(2.86)\qquad \varphi_{t,p}(x)=\begin{cases} p^{-1}x & \text{if } x\in[0,p]\\ (1-p)^{-1}(1-x) & \text{if } x\in[p,1]. \end{cases}$$

It is known([18]) that $\varphi_{t,p}$ is mixing and the unique invariant probability measure coincides with the Lebesgue measure. This $\varphi_{t,p}$ is called a tent transformation.

On the analogy of (2.84), we can define a strictly stationary time series $\boldsymbol{Y}_{t,p}=(Y_{t,p}(n)\,;0\leq n<\infty)$ on the probability space $([0,1],\mathscr{B}([0,1]),\mathrm{d}x)$, which possesses the covariance function

$$(2.87)\qquad \begin{aligned} R_{t,p}(n)&\equiv E((Y_{t,p}(n)-E(Y_{t,p}(n)))(Y_{t,p}(0)-E(Y_{t,p}(0))))\\ &=\begin{cases} (-1)^n 12^{-1}|2p-1|^n & \text{for } p\in(0,1/2)\\ 12^{-1}\delta_{0,n} & \text{for } p=1/2\\ 12^{-1}|2p-1|^n & \text{for } p\in(1/2,1) \end{cases} \end{aligned}$$

([18]). This implies that $(Y_{t,p}(n)-E(Y_{t,p}(n))\,;0\leq n<\infty)$ is a white noise or a simple Markov process, according as $p=1/2$ or $p\neq 1/2$.

§ 3. Stationary analysis

Let d and N be any fixed natural numbers.

[3.1] For any given $N+1$ vectors $\mathscr{X}(n)$ in $\boldsymbol{R}^d(0\leq n\leq N)$, we denote by $\mu^{\mathscr{X}}$ and $R^{\mathscr{X}}=(R^{\mathscr{X}}_{jk})_{1\leq j,k\leq d}$ the **sample mean vector** and the **sample covariance function** of the data $\mathscr{X}=(\mathscr{X}(n)\,;0\leq n\leq N)$, respectively :

$$(3.1)\qquad \mu^{\mathcal{X}}\equiv\frac{1}{N+1}\sum_{n=0}^{N}\mathcal{X}(n)$$

$$(3.2)\qquad \begin{cases} R_{jk}^{\mathcal{X}}(n) \equiv\dfrac{1}{N+1}\displaystyle\sum_{m=0}^{N-n}(\mathcal{X}_j(n+m)-\mu_j^{\mathcal{X}})(\mathcal{X}_k(m)-\mu_k^{\mathcal{X}}) \\ R_{jk}^{\mathcal{X}}(-n)\equiv R_{kj}^{\mathcal{X}}(n), \end{cases}$$

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

$$(3.3)\qquad \mathscr{X}(n)=\begin{bmatrix}\sqrt{R_{11}^{\mathcal{X}}(0)^{-1}} & & 0\\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\mathcal{X}}(0)^{-1}}\end{bmatrix}(\mathcal{X}(n)-\mu^{\mathcal{X}}).$$

We call this procedure a **standardization** of $\mathcal{X}$. Let $R^{\mathscr{X}}$ be the sample covariance function of the standardized data $\mathscr{X}=(\mathscr{X}(n)\,;0\leq n\leq N)$. It is noted that

$$(3.4)\qquad R_{jk}^{\mathscr{X}}(\cdot)=\frac{R_{jk}^{\mathcal{X}}(\cdot)}{\sqrt{R_{jj}^{\mathcal{X}}(0)R_{kk}^{\mathcal{X}}(0)}}\quad(1\leq j,\ k\leq d).$$

According to the algorithm (2.20)-(2.23), (2.26), (2.29) and (2.30) with R in (2.29)-(2.30) replaced by $R^{\mathscr{X}}$ in (3.4), 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 $M(d\,;\boldsymbol{R})$, under the assumption

$$(3.5)\qquad V_+(n)\in GL(d\,;\boldsymbol{R})\quad(0\leq n\leq N-1).$$

REMARK 3.1. When $d=1$, it follows from (2.52) and (2.53) in Remark 2.5 that condition (3.5) is equivalent to

$$(3.6)\qquad |\delta(n)|<1\quad(1\leq n\leq N-1).$$

Now we define $N+1$ vectors $\nu_+(n)$ in $\boldsymbol{R}^d(0\leq n\leq N)$ by

$$(3.7)\qquad \begin{cases}\nu_+(0)=\mathscr{X}(0)\\ \nu_+(n)=\mathscr{X}(n)+\displaystyle\sum_{k=0}^{n-1}\gamma_+(n,k)\mathscr{X}(k)\quad(1\leq n\leq N),\end{cases}$$

where $\gamma_+(n,0)=\delta_+(n)(1\leq n\leq N)$. It is convenient to write the equivalent form of (3.7):

$$\begin{cases}\mathscr{X}(0)=\nu_+(0)\\ \mathscr{X}(n)=-\displaystyle\sum_{k=1}^{n-1}\gamma_+(n,k)\mathscr{X}(k)-\delta_+(n)\mathscr{X}(0)+\nu_+(n)\quad(1\leq n\leq N).\end{cases}$$

Furthermore it can be seen that

$$
(3.8)\qquad
\begin{cases}
\mathscr{X}(0)-\mu^{\mathscr{X}}=\begin{bmatrix}\sqrt{R_{11}^{\mathscr{X}}(0)} & & 0\\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\mathscr{X}}(0)}\end{bmatrix}\nu_+(0)\\[2ex]
\mathscr{X}(n)-\mu^{\mathscr{X}}=-\sum_{k=0}^{n-1}\begin{bmatrix}\sqrt{R_{11}^{\mathscr{X}}(0)} & & 0\\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\mathscr{X}}(0)}\end{bmatrix}\gamma_+(n,k)\cdot\\
\qquad\cdot\begin{bmatrix}\sqrt{R_{11}^{\mathscr{X}}(0)^{-1}} & & 0\\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\mathscr{X}}(0)^{-1}}\end{bmatrix}(\mathscr{X}(k)-\mu^{\mathscr{X}})\\
\qquad+\begin{bmatrix}\sqrt{R_{11}^{\mathscr{X}}(0)} & & 0\\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\mathscr{X}}(0)}\end{bmatrix}\nu_+(n)\qquad(1\le n\le N).
\end{cases}
$$

DEFINITION 3.1. We call equation (3.8)(resp. $\{\gamma_+(n,k);0\le k<n\le N\}$ and $\boldsymbol{\nu}_+=(\nu_+(n);0\le n\le N)$) a **sample forward KM₂O-Langevin equation**(resp. **data** and **force**) associated with the original data $\mathscr{X}=(\mathscr{X}(n);0\le n\le N)$.

[3.2] For any given $N+2$ vectors $\mathscr{X}(n)$ in $\boldsymbol{R}^d(-1\le n\le N)$, we define, as a new data, the first difference $\tilde{\mathscr{X}}(n)$ of $\mathscr{X}(n)$ by

$$
(3.9)\qquad \tilde{\mathscr{X}}(n)=\mathscr{X}(n)-\mathscr{X}(n-1)\quad(0\le n\le N),
$$

which is often used in the analysis of economic data([5]). Let $\mu^{\tilde{\mathscr{X}}}$ and $R^{\tilde{\mathscr{X}}}$ be the sample mean vector and the sample covariance function of the data $\tilde{\mathscr{X}}=(\tilde{\mathscr{X}}(n);0\le n\le N)$, respectively.

The procedure in [3.1] applied to this data $\tilde{\mathscr{X}}$ gives us the sample forward KM₂O-Langevin equation associated with $\tilde{\mathscr{X}}$:

$$
(3.10)\qquad
\begin{cases}
\tilde{\mathscr{X}}(0)-\mu^{\tilde{\mathscr{X}}}=\begin{bmatrix}\sqrt{R_{11}^{\tilde{\mathscr{X}}}(0)} & & 0\\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\tilde{\mathscr{X}}}(0)}\end{bmatrix}\tilde{\nu}_+(0)\\[2ex]
\tilde{\mathscr{X}}(n)-\mu^{\tilde{\mathscr{X}}}=-\sum_{k=0}^{n-1}\begin{bmatrix}\sqrt{R_{11}^{\tilde{\mathscr{X}}}(0)} & & 0\\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\tilde{\mathscr{X}}}(0)}\end{bmatrix}\tilde{\gamma}_+(n,k)\cdot\\
\qquad\cdot\begin{bmatrix}\sqrt{R_{11}^{\tilde{\mathscr{X}}}(0)^{-1}} & & 0\\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\tilde{\mathscr{X}}}(0)^{-1}}\end{bmatrix}(\tilde{\mathscr{X}}(k)-\mu^{\tilde{\mathscr{X}}})\\
\qquad+\begin{bmatrix}\sqrt{R_{11}^{\tilde{\mathscr{X}}}(0)} & & 0\\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\tilde{\mathscr{X}}}(0)}\end{bmatrix}\tilde{\nu}_+(n)\qquad(1\le n\le N),
\end{cases}
$$

where $\{\tilde{\gamma}_+(n,k);0\le k<n\le N\}$ (resp. $\tilde{\nu}_+=(\tilde{\nu}_+(n);0\le n\le N)$) is the sample forward KM_2O-Langevin data(resp. force) associated with $\tilde{\mathscr{X}}$ in the sense of Definition 3.1. Taking account of (3.9), we give

DEFINITION 3.2. We designate equation (3.10)(resp. $\{\tilde{\gamma}_+(n,k);0\le k<n\le N\}$ and $\tilde{\nu}_+$) a **sample first difference forward KM_2O-Langevin equation**(resp. **data** and **force**) associated with the original data $\mathscr{X}=(\mathscr{X}(n);-1\le n\le N)$.

REMARK 3.2. The standardized data $\tilde{\tilde{\mathscr{X}}}=(\tilde{\tilde{\mathscr{X}}}(n);0\le n\le N)$ of $\tilde{\mathscr{X}}$ is defined by

$$(3.11)\qquad \tilde{\tilde{\mathscr{X}}}(n)=\begin{bmatrix}\sqrt{R_{11}^{\tilde{\mathscr{X}}}(0)^{-1}} & & 0\\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\tilde{\mathscr{X}}}(0)^{-1}}\end{bmatrix}(\tilde{\mathscr{X}}(n)-\mu^{\tilde{\mathscr{X}}}).$$

[3.3] For any given one-dimensional data $\mathscr{Y}(n)(0\le n\le N)$, we construct two-dimensional data $\mathscr{X}^{(p)}=(\mathscr{X}^{(p)}(n);0\le n\le N)(2\le p\le 3)$ by

$$(3.12)_p\qquad \mathscr{X}^{(p)}(n)={}^t(\mathscr{Y}(n),\mathscr{Y}(n)^p).$$

Applying the procedure in [3.1] to these $\mathscr{X}^{(p)}(2\le p\le 3)$ and then using the idea in Theorem 2.9, we have

$$(3.13)_p\qquad \begin{cases}\mathscr{Y}(0)-\mu_1=\alpha_1\nu_+^{(p)}(0)\\ \mathscr{Y}(n)-\mu_1=-\sum_{k=0}^{n-1}\gamma_{+1}^{(p)}(n,k)(\mathscr{Y}(k)-\mu_1)\\ \qquad -\sum_{k=0}^{n-1}(\alpha_1/\alpha_p)\gamma_{+2}^{(p)}(n,k)(\mathscr{Y}(k)^p-\mu_p)+\alpha_1\nu_+^{(p)}(n)\quad(1\le n\le N),\end{cases}$$

where

$$(3.14)\qquad \mu_q=\frac{1}{N+1}\sum_{n=0}^{N}\mathscr{Y}(n)^q$$

$$(3.15)\qquad \alpha_q=\left(\frac{1}{N+1}\sum_{n=0}^{N}(\mathscr{Y}(n)^q-\mu_q)^2\right)^{1/2}\quad(1\le q\le 3).$$

Here we note that for each $p\in\{2,3\}$ the coefficients $\gamma_{+1}^{(p)}(\cdot,*)$ and $\gamma_{+2}^{(p)}(\cdot,*)$ (resp. the random force $\nu_+^{(p)}$) in the equation $(3.13)_p$ are the (1,1)-and (1,2)-components(resp. the first component) of the sample forward KM_2O-Langevin data(resp. force) associated with $\mathscr{X}^{(p)}$ in the sense of Definition 3.1.

Following the nomenclature of equation $(2.70)_p$, we give

DEFINITION 3.3. For each $p\in\{2,3\}$ we call equation $(3.13)_p$ (resp.

$\{\gamma_{+i}^{(p)}(n, k) ; 1\leq i\leq 2, 0\leq k<n\leq N\}$ and $\nu_{+}^{(p)}=(\nu_{+}^{(p)}(n) ; 0\leq n\leq N))$ a **sample forward KM_2O-Langevin equation** (resp. **data** and **force**) of **non-linear type p** associated with the original data $\mathscr{Y}=(\mathscr{Y}(n) ; 0\leq n\leq N)$.

[3.4] In this subsection, we define for any given $N+2$ values $\mathscr{Y}(n)$ in $\boldsymbol{R}^1(-1\leq n\leq N)$ its first difference $\tilde{\mathscr{Y}}(n)$ by

$$(3.16)\quad \tilde{\mathscr{Y}}(n)=\mathscr{Y}(n)-\mathscr{Y}(n-1)\quad (0\leq n\leq N).$$

By applying the procedure in [3.3] to this data $\tilde{\mathscr{Y}}=(\tilde{\mathscr{Y}}(n) ; 0\leq n\leq N)$, we have, for each $p\in\{2, 3\}$,

$$(3.17)_p\quad \begin{cases}\tilde{\mathscr{Y}}(0)-\tilde{\mu}_1=\tilde{\alpha}_1\tilde{\nu}_{+}^{(p)}(0)\\ \tilde{\mathscr{Y}}(n)-\tilde{\mu}_1=-\sum_{k=0}^{n-1}\tilde{\gamma}_{+1}^{(p)}(n, k)(\tilde{\mathscr{Y}}(k)-\tilde{\mu}_1)\\ \qquad -\sum_{k=0}^{n-1}(\tilde{\alpha}_1/\tilde{\alpha}_p)\tilde{\gamma}_{+2}^{(p)}(n, k)(\tilde{\mathscr{Y}}(k)^p-\tilde{\mu}_p)+\tilde{\alpha}_1\tilde{\nu}_{+}^{(p)}(n)\quad (1\leq n\leq N),\end{cases}$$

where

$$(3.18)\quad \tilde{\mu}_q=\frac{1}{N+1}\sum_{n=0}^{N}\tilde{\mathscr{Y}}(n)^q$$

$$(3.19)\quad \tilde{\alpha}_q=\left(\frac{1}{N+1}\sum_{n=0}^{N}(\tilde{\mathscr{Y}}(n)^q-\tilde{\mu}_q)^2\right)^{1/2}\quad (1\leq q\leq 3).$$

As in Definitions 3.2 and 3.3, we give a name to equation $(3.17)_p$.

DEFINITION 3.4. For each $p\in\{2, 3\}$, we designate equation $(3.17)_p$ (resp. $\{\tilde{\gamma}_{+i}^{(p)}(n, k) ; 1\leq i\leq 2, 0\leq k<n\leq N\}$ and $\tilde{\nu}_{+}^{(p)}=(\tilde{\nu}_{+}^{(p)}(n) ; 0\leq n\leq N))$ a **sample first difference forward KM_2O-Langevin equation** (resp. **data** and **force**) of **non-linear type p** associated with the original data $\mathscr{Y}$.

REMARK 3.3. For each $p\in\{2, 3\}$, the standardized data $\tilde{\mathscr{X}}^{(p)}=(\tilde{\mathscr{X}}^{(p)}(n) ; 0\leq n\leq N)$ of $\tilde{\mathscr{Z}}^{(p)}=({}^t\tilde{\mathscr{Y}}(n), (\tilde{\mathscr{Y}}(n))^p) ; 0\leq n\leq N)$ is given by

$$(3.20)\quad \tilde{\mathscr{X}}^{(p)}(n)=\begin{bmatrix}\sqrt{R_{11}^{\tilde{\mathscr{Z}}}(0)^{-1}} & 0\\ 0 & \sqrt{R_{22}^{\tilde{\mathscr{Z}}}(0)^{-1}}\end{bmatrix}(\tilde{\mathscr{Z}}^{(p)}(n)-\mu^{\tilde{\mathscr{Z}}^{(p)}}).$$

[3.5] For a d-dimensional data $\mathscr{Z}=({}^t(\mathscr{Z}_1(n), \cdots, \mathscr{Z}_d(n)) ; 0\leq n\leq N)$ such that

$$(3.21)\quad \mathscr{Z}_j(n)>0\quad (1\leq j\leq d),$$

we define a d-dimensional data $\mathrm{Log}\,\mathscr{Z}=((\mathrm{Log}\,\mathscr{Z})(n) ; 0\leq n\leq N)$ by

$$(3.22)\quad (\mathrm{Log}\,\mathscr{Z})_j(n)=\log(\mathscr{Z}_j(n))\quad (1\leq j\leq d).$$

Similarly to the first difference in (3.9), this transformation (3.22) is often used in the analysis of the economic data.

[3.6] We return to subsection [3.1] for $d=2$. Choosing a positive number $w\in(0,1)$ and a standardized random uniform numbers $\boldsymbol{\xi}_u=(\xi_u(n);0\leq n\leq N)$, we define a two-dimensional data $\mathscr{X}_w=(\mathscr{X}_w(n);0\leq n\leq N)$ by

$$(3.23)\qquad \mathscr{X}_w(n)={}^t(\mathscr{X}_1(n),\ \mathscr{X}_2(n)+w\ \xi_u(n)).$$

It deserves mention that the independence of $\boldsymbol{\xi}_u$ and $\mathscr{X}$ guarantees the condition (2.5) for this new data $\mathscr{X}_w$ and that the local and weak stationarity for $\mathscr{X}_w$ implies the same property for $\mathscr{X}$.

[3.7] Under the same situation as in [3.6], we introduce another two-dimensional data $\mathrm{Arct}\mathscr{X}=((\mathrm{Arct}\mathscr{X})(n);0\leq n\leq N)$ by

$$(3.24)\qquad (\mathrm{Arct}\mathscr{X})(n)={}^t(\arctan(\mathscr{X}_1(n)),\ \arctan(\mathscr{X}_2(n))).$$

This transformation is effective in compressing abnormal values in the original data $\mathscr{X}$ and reproducing the local and weak stationarity, as will be seen in [4.4] of § 4.

§ 4. Test (S) for local and weak stationarity

[4.1] Let us return to the same setting as in [3.1]. For any given data $\mathscr{X}=(\mathscr{X}(n);0\leq n\leq N)$ in $\boldsymbol{R}^d$, we constructed the sample forward $\mathrm{KM_2O}$-Langevin data $\{\gamma_+(n,k);0\leq k<n\leq N\}$(resp. force $\boldsymbol{\nu}_+=(\nu_+(n);0\leq n\leq N)$) associated with $\mathscr{X}$.

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

$$(4.1)\qquad V_+(n)=W_+(n){}^tW_+(n)\quad(0\leq n\leq N),$$

we define a d-dimensional data $\boldsymbol{\xi}_+=(\xi_+(n);0\leq n\leq N)$ by

$$(4.2)\qquad \xi_+(n)=W_+(n)^{-1}\nu_+(n).$$

The results in subsections [2.1]-[2.4] assure us that

(4.3) *$\mathscr{X}$ is a realization of a local and weakly stationary time series with $R^{\mathscr{X}}$ in (3.4) as its covariance function*

if and only if

(4.4) *$\boldsymbol{\xi}_+$ realizes a d-dimensional standardized white noise.*

Further, letting $\xi_+(n)={}^t(\xi_{+1}(n),\cdots,\xi_{+d}(n))$, we construct a one-dimensional data $\boldsymbol{\xi}=(\xi(n);0\leq n\leq d(N+1)-1)$ by

$$(4.5)\qquad \xi(n)=\xi_{+p}(m),\ \ n=dm+p-1(1\leq p\leq d,\ 0\leq m\leq N).$$

We then note that (4. 4) is equivalent to

(4. 6) $\quad$ *$\boldsymbol{\xi}$ realizes a one-dimensional standardized white noise.*

An experiance rule in data analysis([3]), however, tells us that we should not use the whole series $\{R^{\mathscr{X}}(n)\,;0\leq n\leq N\}$, because an effective number of the sample covariance function $R^{\mathscr{X}}$ is rather smaller than N and considered to be between $[2\sqrt{N+1}/d]$ and $[3\sqrt{N+1}/d]$. Here we choose the maximum value

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

Thus, in what follows, we are going to make use of the system $\{\gamma_+(n, k)\,;0\leq k<n\leq M\}$, the sample forward KM_2O-Langevin data associated with the reliable part $\{R^{\mathscr{X}}(n)\,;0\leq n\leq M\}$ of $R^{\mathscr{X}}$.

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

(4. 8) $\quad \mathscr{X}_i=(\mathscr{X}(i+n)\,;0\leq n\leq M).$

Similarly to (3. 7), we define $\boldsymbol{\nu}_{+i}=(\nu_{+i}(n)\,;0\leq n\leq M)$ by

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

which can be rewritten into

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

This is the sample forward KM_2O-Langevin equation associated with $\mathscr{X}_i$. Noting (4. 2) and (4. 5), we define a d-dimensional data $\boldsymbol{\xi}_{+i}=(\xi_{+i}(n)\,;0\leq n\leq M)=({}^t(\xi_{+i1}(n), \cdots, \xi_{+id}(n))\,;0\leq n\leq M)$ and a one-dimensional data $\boldsymbol{\xi}_i=(\xi_i(n)\,;0\leq n\leq d(M+1)-1)$ by

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

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

For the same reason as in the assertion of equivalence among (4. 3), (4. 4) and (4. 6), the following assertions are equivalent: for each $i\in\{0, \cdots, N-M\}$,

$(4. 12)_i$ $\quad$ *$\mathscr{X}_i$ realizes a d-dimensional local and weakly stationary time series with $R^{\mathscr{X}}$ in* (3. 4) *as its covariance function.*

$(4.13)_i$ $\boldsymbol{\xi}_{+i}$ realizes a d-dimensional standardized white noise.
$(4.14)_i$ $\boldsymbol{\xi}_i$ realizes a one-dimensional standardized white noise.

We then define the sample mean $\mu^{\boldsymbol{\xi}_i}$, the sample variance $v^{\boldsymbol{\xi}_i}$ and the sample covariance function $R^{\boldsymbol{\xi}_i}(n, m)$ of $\boldsymbol{\xi}_i (1 \le n \le L, 0 \le m \le L-n)$ by

$$(4.15) \quad \mu^{\boldsymbol{\xi}_i} = \frac{1}{d(M+1)} \sum_{k=0}^{d(M+1)-1} \xi_i(k)$$

$$(4.16) \quad v^{\boldsymbol{\xi}_i} = \frac{1}{d(M+1)} \sum_{k=0}^{d(M+1)-1} \xi_i(k)^2$$

$$(4.17) \quad R^{\boldsymbol{\xi}_i}(n, m) = \frac{1}{d(M+1)} \sum_{k=m}^{d(M+1)-1-n} \xi_i(k)\xi_i(n+k),$$

where the effective length L of $R^{\boldsymbol{\xi}_i}$, in this case, is taken to be the minimum one, i. e.,

$$(4.18) \quad L = [2\sqrt{d(M+1)}] - 1.$$

In order to check condition $(4.14)_i$ based upon suitable statistical properties of white noise, we need statistical estimates which assert that $\mu^{\boldsymbol{\xi}_i}$, $v^{\boldsymbol{\xi}_i}-1$ and $R^{\boldsymbol{\xi}_i}(n, m)$ are all sufficiently close to zero for every m, n, $1 \le n \le L$, $0 \le m \le L-n$. For that purpose, for each $i \in \{0, \cdots, N-M\}$, we rewrite (4.15)-(4.17) into

$$(4.19) \quad \mu^{\boldsymbol{\xi}_i} = \frac{1}{\sqrt{d(M+1)}} \left(\frac{1}{\sqrt{d(M+1)}} \sum_{k=0}^{d(M+1)-1} \xi_i(k) \right)$$

$$(4.20) \quad v^{\boldsymbol{\xi}_i} - 1 = \frac{1}{d(M+1)} \sum_{k=0}^{d(M+1)-1} (\xi_i(k)^2 - 1)$$

$$(4.21) \quad R^{\boldsymbol{\xi}_i}(n, m) = \sum_{j=1}^{2} \frac{(L_{n,m}^{(j)})^{1/2}}{d(M+1)} ((L_{n,m}^{(j)})^{-1/2} R_j^{\boldsymbol{\xi}_i}(n, m)).$$

Here the decomposition of $R^{\boldsymbol{\xi}_i}$ into two parts $R_1^{\boldsymbol{\xi}_i}$ and $R_2^{\boldsymbol{\xi}_i}$ in (4.21) is defined as follows: For any fixed m, n, $1 \le n \le L$, $0 \le m \le L-n$, we devide $d(M+1)$ and m by $2n$ and n, respectively;

$$(4.22) \quad d(M+1) = q(2n) + r \quad (0 \le r \le 2n-1)$$

$$(4.23) \quad m = sn + t \quad (0 \le t \le n-1).$$

And if $r \in \{0, \cdots, n\}$, then

$$(4.24) \quad R_1^{\boldsymbol{\xi}_i}(n, m) = \begin{cases} \displaystyle\sum_{k=0}^{n-t-1} \xi_i(m+k)\xi_i(m+n+k) & (s \text{ is even}) \\ \displaystyle\quad + \sum_{j=(s+2)/2}^{q-1} \left(\sum_{k=0}^{n-1} \xi_i(2jn+k)\xi_i((2j+1)n+k) \right) & \\ \displaystyle\sum_{j=(s+1)/2}^{q-1} \left(\sum_{k=0}^{n-1} \xi_i(2jn+k)\xi_i((2j+1)n+k) \right) & (s \text{ is odd}) \end{cases}$$

$$(4.25)\qquad R_2^{\xi_i}(n,m)=\begin{cases}\displaystyle\sum_{j=s/2}^{q-2}\Big(\sum_{k=0}^{n-1}\xi_i((2j+1)n+k)\xi_i(2(j+1)n+k)\Big) & \\ \displaystyle\quad+\sum_{k=0}^{r-1}\xi_i((2q-1)n+k)\xi_i(2qn+k) & (s\text{ is even})\\ \displaystyle\sum_{k=0}^{n-t-1}\xi_i(m+k)\xi_i(m+n+k) & (s\text{ is odd})\\ \displaystyle\quad+\sum_{j=(s+1)/2}^{q-2}\Big(\sum_{k=0}^{n-1}\xi_i((2j+1)n+k)\xi_i(2(j+1)n+k)\Big) & \\ \displaystyle\quad+\sum_{k=0}^{r-1}\xi_i((2q-1)n+k)\xi_i(2qn+k) & \end{cases}$$

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

$$(4.26)\qquad R_1^{\xi_i}(n,m)=\begin{cases}\displaystyle\sum_{k=0}^{n-t-1}\xi_i(m+k)\xi_i(m+n+k) & (s\text{ is even})\\ \displaystyle\quad+\sum_{j=(s+2)/2}^{q-1}\Big(\sum_{k=0}^{n-1}\xi_i(2jn+k)\xi_i((2j+1)n+k)\Big) & \\ \displaystyle\quad+\sum_{k=0}^{r-n-1}\xi_i(2qn+k)\xi_i((2q+1)n+k) & \\ \displaystyle\sum_{j=(s+1)/2}^{q-1}\Big(\sum_{k=0}^{n-1}\xi_i(2jn+k)\xi_i((2j+1)n+k)\Big) & (s\text{ is odd})\\ \displaystyle\quad+\sum_{k=0}^{r-n-1}\xi_i(2qn+k)\xi_i((2q+1)n+k) & \end{cases}$$

$$(4.27)\qquad R_2^{\xi_i}(n,m)=\begin{cases}\displaystyle\sum_{j=s/2}^{q-1}\Big(\sum_{k=0}^{n-1}\xi_i((2j+1)n+k)\xi_i(2(j+1)n+k)\Big) & (s\text{ is even})\\ \displaystyle\sum_{k=0}^{n-t-1}\xi_i(m+k)\xi_i(m+n+k) & (s\text{ is odd})\\ \displaystyle\quad+\sum_{j=(s+1)/2}^{q-1}\Big(\sum_{k=0}^{n-1}\xi_i((2j+1)n+k)\xi_i(2(j+1)n+k)\Big). & \end{cases}$$

Furthermore, $L_{n,m}^{(j)}$ stand for the number of terms in $R_j^{\xi_i}(n,m)(1\le j\le 2)$; if $r\in\{0,\cdots,n\}$, then

$$(4.28)\qquad \begin{cases}L_{n,m}^{(1)}=\begin{cases}n(q+(s/2))-m & (s\text{ is even})\\ n(q-(s+1)/2) & (s\text{ is odd})\end{cases}\\ L_{n,m}^{(2)}=\begin{cases}n(q-1-s/2)+r & (s\text{ is even})\\ n(q-1+(s+1)/2)+r-m & (s\text{ is odd})\end{cases}\end{cases}$$

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

$$(4.29)\qquad \begin{cases}L_{n,m}^{(1)}=\begin{cases}n(q-1+(s/2))+r-m & (s\text{ is even})\\ n(q-1-(s+1)/2)+r & (s\text{ is odd})\end{cases}\\ L_{n,m}^{(2)}=\begin{cases}n(q-s/2) & (s\text{ is even})\\ n(q+(s+1)/2)-m & (s\text{ is odd}).\end{cases}\end{cases}$$

We note that

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

We are now in a position to give a criterion for condition $(4.14)_i$. If we substitute, for the usual orthogonality of white noise, the stronger property of independence, we find that for each $i\in\{0, \cdots, N-M\}$ and $j\in\{1, 2\}$, μ^{ξ_i} and $R_j^{\xi_i}(n, m)$ consist of sums of $d(M+1)$ and $L_{n,m}^{(j)}$ independent random variables with mean zero and variance one, respectively. Hence, we can infer from the central limit theorem that for each (j, m, n), $1\leq j\leq 2$, $1\leq n\leq L$, $0\leq m\leq L-n$,

$$\left|\frac{1}{\sqrt{d(M+1)}}\sum_{k=0}^{d(M+1)-1}\xi_i(k)\right|<1.96 \qquad \text{with probability } 0.95$$
$$(L_{n,m}^{(j)})^{-1/2}|R_j^{\xi_i}(n, m)|<1.96 \qquad \text{with probability } 0.95$$

and so by (4.15), (4.19) and (4.21),

$$(4.31)_i \quad \sqrt{d(M+1)}\,|\mu^{\xi_i}|<1.96 \qquad \text{with probability } 0.95$$
$$(4.32)_i \quad d(M+1)\Big(\sum_{j=1}^{2}(L_{n,m}^{(j)})^{1/2}\Big)^{-1}|R^{\xi_i}(n, m)|<1.96 \qquad \text{with probability } 0.90.$$

Moreover, we want to derive a similar rate at which the quantity $v^{\xi_i}-1$ in (4.20) is sufficiently close to zero. Since we cannot get any useful information about the fourth moment of the white noise ξ_i without extra assumptions in addition to the local and weak stationarity of $\mathscr{X}$, we replace the quantity $v^{\xi_i}-1$ by the following:

$$(4.33)_i \quad (v^{\xi_i}-1)^{\sim}=\Big(\sum_{k=0}^{d(M+1)-1}(\xi_i(k)^2-1)\Big)\Big(\sum_{k=0}^{d(M+1)-1}(\xi_i(k)^2-1)^2\Big)^{-1/2}.$$

Applying both the central limit theorem and the law of large numbers, we can use the Student-t-distribution and come to the concludion: for each $i\in\{0, \cdots, N-M\}$,

$$(4.34)_i \quad |(v^{\xi_i}-1)^{\sim}|<2.2414 \qquad \text{with probability } 0.975$$

(see Remarks 4.1 and 4.2).

Thus, for each $i\in\{0, \cdots, N-M\}$, we have obtained the following criteria $(M)_i$, $(V)_i$ and $(O)_i$ in order to check $(4.14)_i$;

$$(4.35)_i \quad \begin{cases} (M)_i: \textit{the inequality } (4.31)_i \textit{ holds.} \\ (V)_i: \textit{the inequality } (4.34)_i \textit{ holds.} \\ (O)_i: \textit{the inequality } (4.32)_i \textit{ holds.} \end{cases}$$

Concerning the main problem of testing the local and weak stationarity of the original data $\mathscr{X}$, we would like to propose:

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

The tests applied to the transformed data Log$\mathscr{X}$ in [3.5], $\mathscr{X}_w$ in [3.6] and Arct$\mathscr{X}$ in [3.7] are called **Test**$(S)_{\mathbf{Log}}$, **Test**$(S)_{\mathbf{w}}$ and **Test**$(S)_{\mathbf{Arct}}$, respectively.

[4.2] To reach the final form of Test(S) above, we made repeated experiments and observed the validity of Test(S) for various concrete data such as random normal numbers, random uniform numbers, tent transformation($p=1/2, 2/3$) and the logistic transformation as well as for the transformed data obtained by taking the first difference, by multiplying or adding the above data $\mathscr{X}(n)$ by the scalar n (this is expected to destroy the stationarity), and by taking the square or cube. Our results of one hundred experiments are illustrated in Table 4.1 which shows the rate of the numbers of data passing Test(S). In these experiments, we used random normal numbers(resp. random uniform numbers) with 100 prime seed numbers from 2 to 541, tent transformations ($p=1/2, 2/3$) with 100 initial values $(100 \cdot m)/(2 \cdot 13799)$, $1 \leq m \leq 100$, and the logistic transformation with 100 initial values $0.005 \cdot m$, $1 \leq m \leq 100$, where the two initial values 0.250 and 0.500 are, in particular, replaced by 0.249 and 0.499, respectively.

j	Random normal numbers	Random uniform numbers	Tent transformation ($p=1/2$)	Tent transformation ($p=2/3$)	Logistic transformation
1	0.99	1.00	0.96	1.00	0.99
2	0.86	1.00	0.99	1.00	1.00
3	0.66	0.99	0.98	1.00	1.00
4	0.23	0.02	0.03	0.02	0.01
5	0.00	0.00	0.00	0.00	0.00
6	0.98	1.00	0.97	1.00	1.00
7	0.73	0.93	0.95	0.97	1.00
8	0.58	0.82	0.95	0.95	0.98
9	0.27	0.06	0.19	0.13	0.04
10	0.00	0.00	0.00	0.00	0.00
11	0.72	0.97	0.00 *	0.83	0.00 *
12	0.63 *	0.98	0.52 *	0.50 *	0.90
13	0.73	0.96	0.99	1.00	1.00
14	0.58 *	0.93	0.94	1.00	1.00

The first row in Table 4.1 indicates the results for the original data $\mathscr{X}(n)$ $(0\le n\le 100)$ and the jth row$(1\le j\le 14)$ for the transformed data $\mathscr{X}_j=(\mathscr{X}_j(n)\,;0\le n\le 99)$ given by

$$(4.36)\quad \begin{cases} \mathscr{X}_1(n)=\mathscr{X}(n),\ \mathscr{X}_2(n)=\mathscr{X}(n)^2,\ \mathscr{X}_3(n)=\mathscr{X}(n)^3, \\ \mathscr{X}_4(n)=n\mathscr{X}(n),\ \mathscr{X}_5(n)=\mathscr{X}(n)+n,\ \mathscr{X}_6(n)=\mathscr{X}(n)-\mathscr{X}(n-1), \\ \mathscr{X}_7(n)=(\mathscr{X}(n)-\mathscr{X}(n-1))^2,\ \mathscr{X}_8(n)=(\mathscr{X}(n)-\mathscr{X}(n-1))^3 \\ \mathscr{X}_9(n)=n(\mathscr{X}(n)-\mathscr{X}(n-1)),\ \mathscr{X}_{10}(n)=\mathscr{X}(n)-\mathscr{X}(n-1)+n, \\ \mathscr{X}_{11}(n)={}^t(\mathscr{X}_1(n),\ \mathscr{X}_2(n)),\ \mathscr{X}_{12}(n)={}^t(\mathscr{X}_1(n),\ \mathscr{X}_3(n)), \\ \mathscr{X}_{13}(n)={}^t(\mathscr{X}_6(n),\ \mathscr{X}_7(n)),\ \mathscr{X}_{14}(n)={}^t(\mathscr{X}_6(n),\ \mathscr{X}_8(n)). \end{cases}$$

[4.3] Let us illustrate in the following Tables 4.2-4.6 the details of our experiments for each type of data in Table 4.1:

j	(M)	(V)	(O)	(S)
1	0.930	0.944	1.000	S
2	0.958	0.915	1.000	S
3	0.887	0.915	1.000	S
4	0.887	0.690	0.958	NS
5	1.000	0.000	1.000	NS
6	1.000	0.958	1.000	S
7	0.930	0.803	0.930	S
8	0.972	0.859	0.803	S
9	1.000	0.690	0.859	NS
10	1.000	0.000	1.000	NS
11	0.965	0.824	0.942	S
12	0.895	0.791	0.919	S
13	0.977	0.860	0.965	S
14	0.953	0.779	0.826	S

Table 4.2 Random normal numbers with seed number 353

j	(M)	(V)	(O)	(S)
1	0.972	1.000	1.000	S
2	0.944	1.000	1.000	S
3	0.944	1.000	1.000	S
4	1.000	0.507	0.901	NS
5	1.000	0.000	1.000	NS
6	0.958	1.000	1.000	S
7	0.958	1.000	1.000	S
8	1.000	1.000	0.972	S
9	0.972	0.606	0.831	NS
10	1.000	0.000	1.000	NS
11	0.953	0.977	0.977	S
12	0.965	0.965	1.000	S
13	0.965	0.965	0.965	S
14	0.953	0.953	1.000	S

Table 4.3 Random uniform numbers with seed number 131

j	(M)	(V)	(O)	(S)
1	1.000	1.000	0.944	S
2	1.000	1.000	0.915	S
3	1.000	1.000	0.873	S
4	0.958	0.535	0.930	NS
5	1.000	0.000	1.000	NS
6	1.000	1.000	0.958	S
7	1.000	1.000	0.803	S
8	1.000	1.000	0.803	S
9	0.972	0.592	0.958	NS
10	1.000	0.000	1.000	NS
11	0.972	0.268	1.000	NS
12	0.972	0.915	1.000	S
13	1.000	1.000	1.000	S
14	0.958	1.000	0.972	S

Table 4.4 Tent($p=1/2$) with initial value 0.076093

j	(M)	(V)	(O)	(S)
1	1.000	1.000	1.000	S
2	1.000	1.000	1.000	S
3	0.986	1.000	1.000	S
4	1.000	0.592	0.986	NS
5	1.000	0.000	1.000	NS
6	1.000	1.000	1.000	S
7	0.986	1.000	0.986	S
8	0.986	1.000	1.000	S
9	1.000	0.606	1.000	NS
10	1.000	0.000	1.000	NS
11	0.953	0.965	0.860	S
12	0.977	0.872	1.000	S
13	0.977	1.000	0.953	S
14	0.977	1.000	0.942	S

Table 4.5 Tent($p=2/3$) with initial value 0.326111

j	(M)	(V)	(O)	(S)
1	0.958	1.000	1.000	S
2	0.958	1.000	1.000	S
3	0.986	1.000	1.000	S
4	1.000	0.620	0.958	NS
5	1.000	0.000	1.000	NS
6	1.000	0.930	1.000	S
7	1.000	1.000	0.930	S
8	1.000	1.000	0.972	S
9	1.000	0.648	1.000	NS
10	1.000	0.000	1.000	NS
11	1.000	0.256	0.988	NS
12	0.988	0.721	1.000	S
13	0.965	1.000	0.965	S
14	0.988	1.000	1.000	S

Table 4.6 Logistic with initial

Tables 4.2-4.6 denote the rate of i such that each of $(M)_i$, $(V)_i$ and $(O)_i$ holds for random normal numbers with seed number 353, random uniform numbers with seed number 131, tent transformation($p=1/2$) with initial value 0.076093, tent transformation($p=2/3$) with initial value 0.326111 and the logistic transformation with initial value 0.02. Here "S" and "NS" stand for stationarity and non-stationarity, respectively.

[4.4] We can say from [4.2] and [4.3] that the experimental results of the above tables are in agreement with the expected ones from the theory in all rows except for $j=11$, 12, 14; but we cannot for $j=11$, 12, 14. It seems that the disagreement in the 14th row in Tables 4.1-4.2 comes from the occurrence of abnormal values in random normal numbers. On the other hand, for $j=11$, 12 in Tables 4.1, 4.4 and 4.6, it lies in the strong dependence between components of the two-dimensional data. In order to overcome these difficulties, we adopted the modified Test$(S)_{\mathrm{Arct}}$ in the case of the 14th row in Table 4.2; Test$(S)_{0.07}$ with weight 0.07 in the case of the 11-12th rows in Tables 4.4-4.6. The results are ilustrated in Tables 4.7-4.12, respectively; Good accordance with the theory.

j	Random normal numbers
11	0.99
12	0.97
14	0.99

Table 4.7 Test$(S)_{\mathrm{Arct}}$

j	Tent transformation ($p=1/2$)	Tent transformation ($p=2/3$)	Logistic transformation
11	0.98	1.00	0.98
12	0.88	0.87	0.97
14	0.93	1.00	1.00

Table 4.8 Test$(S)_{0.07}$

Tables 4.7-4.8 denote the results based on Test$(S)_{\mathrm{Arct}}$ and Test$(S)_{0.07}$ for the same data as in Table 4.1.

j	(M)	(V)	(O)	(S)
11	0.977	0.942	0.966	S
12	0.930	0.988	0.988	S
14	1.000	0.884	0.965	S

Table 4.9 Test$(S)_{\mathrm{Arct}}$ for random normal numbers with seed number 353

j	(M)	(V)	(O)	(S)
11	0.988	0.919	1.000	S
12	0.977	0.965	1.000	S
14	0.965	1.000	0.977	S

Table 4.10 Test$(S)_{0.07}$ for tent transformation($p=1/2$) with initial value 0.076093

j	(M)	(V)	(O)	(S)
11	0.907	0.988	0.884	S
12	0.953	0.965	1.000	S
14	0.977	1.000	0.930	S

Table 4.11 Test$(S)_{0.07}$ for tent transformation($p=2/3$) with initial value 0.326111

j	(M)	(V)	(O)	(S)
11	0.953	0.953	0.988	S
12	0.977	0.872	0.988	S
14	0.988	1.000	0.988	S

Table 4.12 Test$(S)_{0.07}$ for logistic transformation with initial value 0.02

Table 4.9(resp. 4.10-4.12) denotes the results based on Test$(S)_{\mathrm{Arct}}$ (resp. Test$(S)_{0.07}$) for the same data as in Table 4.2(resp. 4.4-4.6).

[4.5] We give two remarks concerning the criterion $(V)_i$.

REMARK 4.1. In place of $(V)_i$, we made lots of experiments for the following criterion $(\tilde{V})_i$ based upon the Student-t-distribution:

$(\tilde{V})_i$ $\quad |(v^{\xi_i}-1)^{\approx}|<t_M(0.025)$ $\quad$ with probability 0.95,

where

$$(4.35)_i \quad (v^{\xi_i}-1)^{\approx}=\{\sum_{k=0}^{d(M+1)-1}(\xi_i(k)^2-1)\}\cdot \cdot\{\sum_{k=0}^{d(M+1)-1}(\xi_i(k)^2-\frac{1}{d(M+1)}\sum_{m=0}^{d(M+1)-1}\xi_i(m)^2)^2\}^{-1/2}.$$

The differernce between $(V)_i$ and $(\tilde{V})_i$ lies in that we replace, in $(V)_i$, the value $(1/d(M+1))\sum_{m=0}^{d(M+1)-1}\xi_i(m)^2$ and $t_M(0.025)$ by 1 and $t_\infty(0.025)=2.2414$, respectively, from the law of large numbers. However we could not find any marked difference among them in our repeated experiments. Hence we adopted $(V)_i$, because it is simpler and we would not like to put the further assumptions to the fourth moments of $\xi_i(n)$.

REMARK 4.2. Concerning the $(V)_i$ part, we required that the rate of $i\in\{0,\cdots,N-M\}$ for which $(V)_i$ holds is over 70%, which is rather low compared with 80% for $(M)_i$ and $(O)_i$. The reason of this difference comes from the fact that we adopted a lenient standard in the inquality $(4.34)_i$, different from the one in inqualities $(4.31)_i$ and $(4.32)_i$, because the accuracy rate of approximations in $(V)_i$ is far worse than the one in $(M)_i$ and $(O)_i$.

§ 5. Wolfer's sunspot numbers, Lynx in Canada and NEC's stock prices in Japan

There exist two fundamental works by W. S. Jevons([10]) and by C. G. Mata and F. I. Schaffner([14]) about the theory of sunspots from the economic point of view. On the other hand, from the viewpoint of mathematical statistics, G. U. Yule([42]) studied the problem of periodicity of sunspots by using the AR(2)-model. The same problem for Canadian Lynx cycle was investigated by P. A. P. Moran ([17]) where the AR(2)-model was fitted to the new data obtained by taking the logarithmic transformation.

We can find many statistical studies looking into an outstanding observation that both time series of Wolfer's sunspot numbers and of Lynx in Canada would have periodicity of about 11 years([16], [37]). However, we do not know any researches trying to answer this serious question: Do these two time series have the local and weak stationarity? Indeed, such stationarity has often been assumed explicitly or implicitly.

This section aims to investigate this problem of local and weak stationarity for Wolfer's sunspot numbers, Lynx in Canada and NEC's stock prices in Japan, based upon our theory developed in § 2-§ 4.

[5.1]. We illustrate the results of Test(S) for Wolfer's sunspot numbers in Tables 5.1-5.2.

year	j	(M)	(V)	(O)	(S)
1821-1934	1	1.000	0.819	1.000	S
	2	1.000	0.819	0.880	S
	3	1.000	0.602	0.843	NS
	4	0.988	0.819	1.000	S
	5	1.000	0.771	1.000	S
	6	1.000	0.554	1.000	NS
	7	1.000	1.000	1.000	S
	8	1.000	1.000	1.000	S
	9	1.000	0.843	1.000	S
	10	0.988	0.964	1.000	S
	11	0.880	0.952	1.000	S
	12	0.952	0.843	0.928	S
	13	0.986	1.000	1.000	S

Table 5.1 Wolfer's sunspot numbers

Here the data $\mathscr{X}_1^{(j)}=(\mathscr{X}_1^{(j)}(n)\,;0\leq n\leq 113)$ in the jth row in Table 5.1$(1\leq j\leq 13)$ are defined by

$$(5.1)\qquad \mathscr{X}_1^{(j)}(n)=\begin{cases}(\mathscr{X}_1^{(1)}(n))^j & (1\leq j\leq 3)\\ (\mathscr{X}_1^{(1)}(n+1)-\mathscr{X}_1^{(1)}(n))^{j-3} & (4\leq j\leq 6)\\ \arctan(\mathscr{X}_1^{(j-6)}(n)) & (7\leq j\leq 12)\\ \log(\mathscr{X}_1^{(1)}(n)) & (j=13),\end{cases}$$

where $\mathscr{X}_1^{(1)}$ denotes the observed data of Wolfer's sunspot numbers for 114 years from 1821 to 1934. It seems that the original data $\mathscr{X}_1^{(1)}$ and its first difference $\mathscr{X}_1^{(4)}$ as well as their squares $\mathscr{X}_1^{(2)}$ and $\mathscr{X}_1^{(5)}$ have all the local and weak stationarity. In addition, it brings us a better result to operate the arctangent transform and/or the logarithmic one.

Following the same notation as in Table 5.1, we show in Table 5.2 the results of Test(S) for Wolfer's sunspot numbers $\mathscr{X}_2^{(1)}$ for 100 years from 1880 to 1979. We note that the original data $\mathscr{X}_2^{(1)}$, its square $\mathscr{X}_2^{(2)}$ and cube $\mathscr{X}_2^{(3)}$ do not pass Test(S), but its first difference $\mathscr{X}_2^{(4)}$ does. As in Table 5.1, we have a good result if we operate the arctangent and logarithmic transformations.

year	j	(M)	(V)	(O)	(S)
1880-1979	1	0.972	0.380	1.000	NS
	2	0.944	0.465	0.817	NS
	3	0.958	0.465	0.690	NS
	4	1.000	0.859	1.000	S
	5	1.000	0.338	0.887	NS
	6	1.000	0.338	1.000	NS
	7	1.000	0.789	1.000	S
	8	0.986	0.873	1.000	S
	9	0.972	0.873	0.944	S
	10	0.944	0.887	0.986	S
	11	1.000	0.873	1.000	S
	12	1.000	0.732	1.000	S
	13	0.986	0.845	1.000	S

Table 5.2 Wolfer's sunspot numbers

[5.2] Table 5.3 states for the results of Test(S) for a time series $\mathscr{X}_3^{(1)}$ denoting the amount of capture of Lynx in MacKenzie River in Canada whose data is known only for 114 years from 1821 to 1934([7]).

year	j	(M)	(V)	(O)	(S)
1821-1934	1	0.940	0.938	1.000	S
	2	0.976	0.867	0.928	S
	3	1.000	0.855	0.687	NS
	4	0.952	0.867	1.000	S
	5	0.964	0.831	0.928	S
	6	0.976	0.831	0.976	S
	7	0.964	0.988	1.000	S
	8	1.000	0.928	0.952	S
	9	0.976	0.940	0.952	S
	10	0.976	0.940	1.000	S
	11	0.940	0.952	0.928	S
	12	0.976	0.855	0.916	S
	13	0.958	0.986	1.000	S

Table 5.3 Lynx in MacKenzie River in Canada

It seems that the original data $\mathscr{X}_3^{(1)}$, its square $\mathscr{X}_3^{(2)}$ and the first difference data $\mathscr{X}_3^{(4)}$, its square $\mathscr{X}_3^{(5)}$ and cube $\mathscr{X}_3^{(6)}$ have the local and weak stationarity; in addition, the arctangent and logarithmic transforms bring us the high-level stationarity.

[5.3] We illustrate in Table 5.4 the results of Test(S) for a time series $\mathscr{X}_4^{(1)}$ consisting of the data of length 108 of NEC's stock prices in Japan from April 1, 1987 to August 31, 1987.

year	j	(M)	(V)	(O)	(S)
1987.4.1 — 1987.8.31	1	1.000	0.734	1.000	S
	2	1.000	0.670	1.000	NS
	3	1.000	0.671	0.975	NS
	4	0.987	0.974	0.962	S
	5	0.974	0.821	0.769	NS
	6	0.949	0.821	0.769	NS
	7	1.000	0.911	0.962	S
	8	1.000	0.911	0.886	S
	9	1.000	0.835	0.822	S
	10	0.962	1.000	1.000	S
	11	1.000	1.000	1.000	S
	12	0.987	0.821	0.962	S

Table 5. 4 NEC's stock prices in Japan

It seems that the original data $\mathscr{X}_4^{(1)}$ and its first difference $\mathscr{X}_4^{(4)}$ in the above period have the local and weak stationarity. The arctangent transfom brings us the high-level stationarity.

On the other hand, Table 5.5 indicates the results for Test(S) for a time series $\mathscr{X}_5^{(1)}$ consisting of the data of length 119 of NEC's stock prices in Japan in the period from August 31, 1987 to February 10, 1988 which contains the so-called " *black Monday* (*October* 19, 1987)". With respect to the stationarity, this is far worse than any other data we discuss here.

year	j	(M)	(V)	(O)	(S)
1987.8.31 — 1988.2.10	1	1.000	0.787	0.686	NS
	2	0.955	0.753	0.663	NS
	3	0.966	0.730	0.640	NS
	4	1.000	0.136	0.989	NS
	5	0.977	0.375	0.989	NS
	6	1.000	0.136	1.000	NS
	7	0.933	0.697	0.640	NS
	8	0.944	0.697	0.652	NS
	9	0.944	0.685	0.663	NS
	10	0.989	0.966	0.841	S
	11	0.989	0.818	0.716	NS
	12	0.898	0.716	0.864	S

Table 5. 5 NEC's stock prices in Japan

[5.4] Table 5.6 shows the results of Test(S) for five two-dimensional data $\mathscr{X}_6^{(j)}=(\mathscr{X}_6^{(j)}(n)\,;0\leq n\leq 113)(1\leq j\leq 5)$ consisting of Wolfer's sunspot numbers and Lynx in MacKenzie River in Canada from 1821 to 1934. Here we define

$$
(5.2)\qquad \mathscr{X}_6^{(j)}(n)=\begin{cases} {}^t(\mathscr{X}_1^{(1)}(n),\ \mathscr{X}_3^{(1)}(n)) & (j=1)\\ {}^t(\mathscr{X}_1^{(4)}(n),\ \mathscr{X}_3^{(4)}(n)) & (j=2)\\ {}^t(\arctan(\mathscr{X}_1^{(1)}(n)),\ \arctan(\mathscr{X}_3^{(1)}(n))) & (j=3)\\ {}^t(\arctan(\mathscr{X}_1^{(4)}(n)),\ \arctan(\mathscr{X}_3^{(4)}(n))) & (j=4)\\ {}^t(\log(\mathscr{X}_1^{(1)}(n)),\ \log(\mathscr{X}_3^{(1)}(n))) & (j=5). \end{cases}
$$

It seems that the two-dimensional data ^{t}(Wolfer's sunspot numbers, Lynx in MacKenzie River in Canada) as well as its first difference data passes three kinds of tests, Test(S), Test$(S)_{\mathrm{Arct}}$ and Test$(S)_{\mathrm{Log}}$.

year	j	(M)	(V)	(O)	(S)
1821-1934	1	0.980	0.949	0.919	S
	2	0.949	0.939	0.919	S
	3	0.980	0.939	0.990	S
	4	0.970	0.939	0.939	S
	5	0.960	0.929	1.000	S

Table 5.6 ^{t}(Wolfer's sunspot numbers, Lynx in Canada)

[5.5] Finally, Table 5.7 indicates the results of Test$(S)_{\mathrm{Arct}}$ for NEC's stock prices in Japan, based upon the sample first difference forward KM_2O-Langevin equations of non-linear type p, $2\leq p\leq 3$, treated in the subsection [3.4].

year	j	(M)	(V)	(O)	(S)
1987.4.1-	1	0.979	0.851	0.904	S
1987.8.31	2	0.947	0.840	0.894	S
1987.9.1-	3	1.000	0.654	0.712	NS
1988.2,10	4	0.885	0.606	0.837	NS

Table 5.7 NEC's stock prices in Japan

Here the data $\mathscr{X}_7^{(j)}=(\mathscr{X}_7^{(j)}(n);0\leq n\leq 107)(1\leq j\leq 2)$ and $\mathscr{X}_7^{(j)}=(\mathscr{X}_7^{(j)}(n);0\leq n\leq 118)(3\leq j\leq 4)$ in the jth row in Table 5.7 are defined by

$$(5.3)\qquad \mathscr{X}_7^{(j)}(n)=\begin{cases} {}^t(\mathrm{arct}(\mathscr{X}_4^{(4)}(n)),\ \mathrm{arct}(\mathscr{X}_4^{(5)}(n))) & (j=1)\\ {}^t(\mathrm{arct}(\mathscr{X}_4^{(4)}(n)),\ \mathrm{arct}(\mathscr{X}_4^{(6)}(n))) & (j=2)\\ {}^t(\mathrm{arct}(\mathscr{X}_5^{(4)}(n)),\ \mathrm{arct}(\mathscr{X}_5^{(5)}(n))) & (j=3)\\ {}^t(\mathrm{arct}(\mathscr{X}_5^{(4)}(n)),\ \mathrm{arct}(\mathscr{X}_5^{(6)}(n))) & (j=4).\end{cases}$$

We can say that the two-dimensional data which contains that " *black Monday* " does not have the local and weak stationarity even if we take the first difference and then the arctangent transform. Note that the one-dimensional data obtained by the same transformation passed Test(S) as shown in Table 5.5.

REMARK 5.1. We will in [36] develop the present stationary Test(S) by studying in more details non-linear KM₂O-Langevin equations of higher order. Furthermore, we will investigate the problem of causal relation among Wolfer's sunspot numbers and some meteorological data.

§ 6. Simulation

[6.1] Returning to the same setting as in [3.1] and [4.1], we treat any d-dimensional data $\mathscr{X}=(\mathscr{X}(n);0\leq n\leq N)$ that passed Test(S) together with $(M)_{N-M}$, $(V)_{N-M}$ and $(O)_{N-M}$ in § 4. Hence the data $\mathscr{X}_{N-M}=(\mathscr{X}(N-M+n);0\leq n\leq M)$ in (4.8) can be regarded as a realization of the weakly stationary time series $\boldsymbol{X}_{N-M}=(X_{N-M}(n);0\leq n\leq M)$.

By applying the prediction formula of one-step future in Theorem 2.6 to $\boldsymbol{X}_{N-M}$, we see from (3.8) that it would be reasonable to define a simulation $\hat{\mathscr{X}}_{N-M}=(\hat{\mathscr{X}}_{N-M}(n);0\leq n\leq M)$ of $\mathscr{X}_{N-M}=(\mathscr{X}(N-M+n);0\leq n\leq M)$ by

$$(6.1)\qquad \begin{cases} \hat{\mathscr{X}}_{N-M}(0)=\mathscr{X}(N-M)\\ \hat{\mathscr{X}}_{N-M}(n)=\mu^{\mathscr{X}}-\sum_{k=0}^{n-1}\begin{bmatrix}\sqrt{R_{11}^{\mathscr{X}}(0)} & & 0\\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\mathscr{X}}(0)}\end{bmatrix}\gamma_+(n,k)\cdot\\ \qquad\cdot\begin{bmatrix}\sqrt{R_{11}^{\mathscr{X}}(0)^{-1}} & & 0\\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\mathscr{X}}(0)^{-1}}\end{bmatrix}(\mathscr{X}(N-M+k)-\mu^{\mathscr{X}}),\end{cases}$$

for every $n\in\{1,\cdots,M\}$.

[6.2] For the original data $\mathscr{X}=(\mathscr{X}(n);-1\leq n\leq N)$ treated in [3.2] and [4.3], similarly to [6.1], we consider the case where the data $\tilde{\mathscr{X}}$ passes Test(S) together with $(M)_{N-M}$, $(V)_{N-M}$ and $(O)_{N-M}$ in § 4.

Taking the same consideration as in [6.1] and noting (3.10), we define a simulation $\hat{\hat{\mathscr{X}}}_{N-M}=(\hat{\hat{\mathscr{X}}}_{N-M}(n)\,;0\leq n\leq M)$ of $\mathscr{X}_{N-M}=(\mathscr{X}(N-M+n)\,;0\leq n\leq M)$ by

$$(6.2)\qquad \begin{cases} \hat{\hat{\mathscr{X}}}_{N-M}(0)=\mathscr{X}(N-M) \\ \hat{\hat{\mathscr{X}}}_{N-M}(n)=\mathscr{X}(N-M+n-1)+\mu^{\tilde{\mathscr{X}}} \end{cases}$$

$$-\sum_{k=0}^{n-1}\begin{bmatrix}\sqrt{R_{11}^{\tilde{\mathscr{X}}}(0)} & & 0 \\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\tilde{\mathscr{X}}}(0)}\end{bmatrix}\tilde{\gamma}_{+}(n,k)\cdot$$

$$\cdot\begin{bmatrix}\sqrt{R_{11}^{\tilde{\mathscr{X}}}(0)^{-1}} & & 0 \\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\tilde{\mathscr{X}}}(0)^{-1}}\end{bmatrix}(\tilde{\mathscr{X}}(N-M+k)-\mu^{\tilde{\mathscr{X}}})$$

for every $n\in\{1,\cdots,M\}$.

[6.3] On the analogy of [6.1] and [6.2], we consider the original data $\mathscr{Y}=(\mathscr{Y}(n)\,;0\leq n\leq N)$ treated in [3.3] and [4.3] such that for each $p\in\{2,3\}$ the standardized data $\mathscr{X}^{(p)}$ passes Test(S) together with $(M)_{N-M_p}$, $(V)_{N-M_p}$ and $(O)_{N-M_p}(2\leq p\leq 3)$, where $M_p(2\leq p\leq 3)$ are given by

$$(6.3)\qquad M_2=M_3=[3\sqrt{N+1}/2]-1.$$

Taking the same consideration as in [6.1] and then noting $(3.13)_p(2\leq p\leq 3)$, we define two kinds of simulations $\hat{\mathscr{Y}}_{N-M_p}=(\hat{\mathscr{Y}}^{(p)}_{N-M_p}(n)\,;0\leq n\leq M_p)$ of $\mathscr{Y}_{N-M_p}=(\mathscr{Y}(N-M_p+n)\,;0\leq n\leq M_p)$ by

$$(6.4)_p\qquad \begin{cases} \hat{\mathscr{Y}}^{(p)}_{N-M_p}(0)=\mathscr{Y}(N-M_p) \\ \hat{\mathscr{Y}}^{(p)}_{N-M_p}(n)=\mu_1-\sum_{k=0}^{n-1}\gamma^{(p)}_{+1}(n,k)(\mathscr{Y}(N-M_p+k)-\mu_1) \end{cases}$$

$$-\sum_{k=0}^{n-1}(\alpha_1/\alpha_p)\gamma^{(p)}_{+2}(n,k)(\mathscr{Y}(N-M_p+k)^p-\mu_p)$$

for every $p\in\{2,3\}$ and $n\in\{1,\cdots,M_p\}$.

[6.4] Finally, for the original data $\mathscr{Y}_{-1}=(\mathscr{Y}(n)\,;-1\leq n\leq N)$ argued in [3.4] and [4.3], we consider the case where Test(S) holds together with $(M)_{N-M_p}$, $(V)_{N-M_p}$ and $(O)_{N-M_p}$ in § 4 for the standardized data $\tilde{\mathscr{X}}^{(p)}$ in (3.20).

By taking account of $(3.17)_p(2\leq p\leq 3)$, the same procedure allows us to define two kinds of simulations $\hat{\hat{\mathscr{Y}}}^{(p)}_{N-M_p}=(\hat{\hat{\mathscr{Y}}}^{(p)}_{N-M_p}(n)\,;0\leq n\leq M_p)$ of $\mathscr{Y}_{N-M_p}=(\mathscr{Y}(N-M_p+n)\,;0\leq n\leq M_p)$ by

$$
(6.5)_p \quad \begin{cases} \hat{\tilde{\mathscr{Y}}}^{(p)}_{N-M_p}(0) = \mathscr{Y}(N-M_p) \\ \hat{\tilde{\mathscr{Y}}}^{(p)}_{N-M_p}(n) = \mathscr{Y}(N-M_p+n-1)+\tilde{\mu}_1 \\ \qquad -\sum_{k=0}^{n-1}\tilde{\gamma}^{(p)}_{+1}(n,k)(\tilde{\mathscr{Y}}(N-M_p+k)-\tilde{\mu}_1) \\ \qquad -\sum_{k=0}^{n-1}(\tilde{\alpha}_1/\tilde{\alpha}_p)\tilde{\gamma}^{(p)}_{+2}(n,k)(\tilde{\mathscr{Y}}(N-M_p+k)^p-\tilde{\mu}_p), \end{cases}
$$

for every $p\in\{2,3\}$ and $n\in\{1,\cdots,M_p\}$.

§ 7. Prediction

The simulations obtained in § 6 are based upon the so-called backward prediction formulae. In this section we give forward prediction formulae for finite-step future $\hat{\mathscr{X}}(N+m)(1\le m\le M-1)$ of the data $\mathscr{X}=(\mathscr{X}(n);0\le n\le N)$ that passed Test(S) together with $(M)_{N-M}$, $(V)_{N-M}$ and $(O)_{N-M}$ in § 4.

[7.1] We consider the same situation as in [6.1]. Since $\mathscr{X}_{N-M}=(\mathscr{X}(N-M+n);0\le n\le M)$ in (4.8) can be regarded as a realization of the weakly stationary time series $\boldsymbol{X}_{N-M}=(X_{N-M}(n);0\le n\le M)$, the system $\{\gamma_+(n,k);0\le k<n\le M\}$ can be regarded as a candidate for the forward KM_2O-Langevin data associated with $\boldsymbol{X}_{N-M}$. Therefore, it follows from (3.8) that it would be reasonable to give a prediction formula for one-step future $\hat{\mathscr{X}}(N+1)$ of the data $\mathscr{X}$ by

$$
(7.1) \quad \hat{\mathscr{X}}(N+1)=\mu^{\mathscr{X}}-\sum_{k=0}^{M-1}\begin{bmatrix}\sqrt{R^{\mathscr{X}}_{11}(0)} & & 0\\ & \ddots & \\ 0 & & \sqrt{R^{\mathscr{X}}_{dd}(0)}\end{bmatrix}\gamma_+(M,k)\cdot \begin{bmatrix}\sqrt{R^{\mathscr{X}}_{11}(0)^{-1}} & & 0\\ & \ddots & \\ 0 & & \sqrt{R^{\mathscr{X}}_{dd}(0)^{-1}}\end{bmatrix}(\mathscr{X}(N-M+1+k)-\mu^{\mathscr{X}}).
$$

In fact, if there exists an $\boldsymbol{R}^d$-valued random variable $X_{N-M}(M+1)$ such that the extended time series $(X_{N-M}(n);0\le n\le M+1)$ is still weakly stationary, then the forward KM_2O-Langevin data associated with it must be equal to the one associated with $\boldsymbol{X}_{N-M}$. Hence, (7.1) comes from (6.1).

Furthemore, if the prediction formulae for finite step-future $\hat{\mathscr{X}}(N+m)$ $(1\le m\le m_0)$ until some time $m_0\in\{1,\cdots,M-2\}$ are obtained, then a prediction formula for m_0+1-step future $\hat{\mathscr{X}}(N+m_0+1)$ of $\mathscr{X}$ is given by

$$
\begin{aligned}
(7.2)\qquad & \hat{\mathscr{X}}(N+m_0+1) \\
&= \mu^{\mathscr{X}} - \sum_{k=0}^{M-m_0-1} \begin{bmatrix} \sqrt{R_{11}^{\mathscr{X}}(0)} & & 0 \\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\mathscr{X}}(0)} \end{bmatrix} \gamma_+(M, k) \cdot \\
&\qquad \cdot \begin{bmatrix} \sqrt{R_{11}^{\mathscr{X}}(0)^{-1}} & & 0 \\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\mathscr{X}}(0)^{-1}} \end{bmatrix} (\mathscr{X}(N-M+m_0+1+k)-\mu^{\mathscr{X}}) \\
&\quad - \sum_{m=0}^{m_0} \begin{bmatrix} \sqrt{R_{11}^{\mathscr{X}}(0)} & & 0 \\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\mathscr{X}}(0)} \end{bmatrix} \gamma_+(M, M-m_0-1+m) \cdot \\
&\qquad\qquad \cdot \begin{bmatrix} \sqrt{R_{11}^{\mathscr{X}}(0)^{-1}} & & 0 \\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\mathscr{X}}(0)^{-1}} \end{bmatrix} (\hat{\mathscr{X}}(N+m)-\mu^{\mathscr{X}}).
\end{aligned}
$$

Definition 7.1. We call (7.2) with (7.1) **KM_2O-predictors.**

[7.2] For the data $\mathscr{X}=(\mathscr{X}(n)\,;\,-1\le n\le N)$ treated in [6.2], the KM_2O-predictors for m-step future $\hat{\tilde{\mathscr{X}}}(N+m)$ of $\mathscr{X}(1\le m\le M-1)$ are given inductively by

$$
\begin{aligned}
(7.3)\qquad & \hat{\tilde{\mathscr{X}}}(N+1) = \mathscr{X}(N)+\mu^{\mathscr{X}} \\
&\quad - \sum_{k=0}^{M-1} \begin{bmatrix} \sqrt{R_{11}^{\tilde{\mathscr{X}}}(0)} & & 0 \\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\tilde{\mathscr{X}}}(0)} \end{bmatrix} \tilde{\gamma}_+(M, k) \cdot \\
&\qquad \cdot \begin{bmatrix} \sqrt{R_{11}^{\tilde{\mathscr{X}}}(0)^{-1}} & & 0 \\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\tilde{\mathscr{X}}}(0)^{-1}} \end{bmatrix} (\tilde{\bar{\mathscr{X}}}(N-M+k+1)-\mu^{\tilde{\mathscr{X}}})
\end{aligned}
$$

and

$$
\begin{aligned}
(7.4)\qquad & \hat{\tilde{\mathscr{X}}}(N+m_0+1) = \hat{\tilde{\mathscr{X}}}(N+m_0)+\mu^{\tilde{\mathscr{X}}} \\
&\quad - \sum_{k=0}^{M-m_0-1} \begin{bmatrix} \sqrt{R_{11}^{\tilde{\mathscr{X}}}(0)} & & 0 \\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\tilde{\mathscr{X}}}(0)} \end{bmatrix} \tilde{\gamma}_+(M, k) \cdot \\
&\qquad \cdot \begin{bmatrix} \sqrt{R_{11}^{\tilde{\mathscr{X}}}(0)^{-1}} & & 0 \\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\tilde{\mathscr{X}}}(0)^{-1}} \end{bmatrix} (\tilde{\bar{\mathscr{X}}}(N-M_0+m_0+k+1)-\mu^{\tilde{\bar{\mathscr{X}}}})
\end{aligned}
$$

$$- \begin{bmatrix} \sqrt{R_{11}^{\tilde{\mathscr{X}}}(0)} & & 0 \\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\tilde{\mathscr{X}}}(0)} \end{bmatrix} \tilde{\gamma}_+(M, M-m_0) \cdot$$

$$\cdot \begin{bmatrix} \sqrt{R_{11}^{\tilde{\mathscr{X}}}(0)^{-1}} & & 0 \\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\tilde{\mathscr{X}}}(0)^{-1}} \end{bmatrix} (\hat{\tilde{\mathscr{X}}}(N+1) - \mathscr{X}(n) - \mu^{\tilde{\mathscr{X}}})$$

$$- \sum_{m=2}^{m_0} \begin{bmatrix} \sqrt{R_{11}^{\tilde{\mathscr{X}}}(0)} & & 0 \\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\tilde{\mathscr{X}}}(0)} \end{bmatrix} \tilde{\gamma}_+(M, M-m_0-1+m) \cdot$$

$$\cdot \begin{bmatrix} \sqrt{R_{11}^{\tilde{\mathscr{X}}}(0)^{-1}} & & 0 \\ & \ddots & \\ 0 & & \sqrt{R_{dd}^{\tilde{\mathscr{X}}}(0)^{-1}} \end{bmatrix} (\hat{\tilde{\mathscr{X}}}(N+m) - \hat{\tilde{\mathscr{X}}}(N+m-1) - \mu^{\tilde{\mathscr{X}}}).$$

[7.3] For the data $\mathscr{Y} = (\mathscr{Y}(n); 0 \leq n \leq N)$ treated in [6.3], we can give two kinds of KM₂O-predictors for finite-step future $\hat{\mathscr{Y}}^{(p)}(N+m)$ of $\mathscr{Y}$ $(2 \leq p \leq 3,\ 1 \leq m \leq M_p - 1)$ as follows:

$$(7.5)_p \quad \hat{\mathscr{Y}}^{(p)}(N+1) = \mu_1 - \sum_{k=0}^{M_p-1} \gamma_{+1}^{(p)}(M_p, k)(\mathscr{Y}(n-M_p+1+k) - \mu_1) - \sum_{k=0}^{M_p-1} (a_1/a_p)\gamma_{+2}^{(p)}(M_p, k)(\mathscr{Y}(N-M_p+1+k)^p - \mu_p)$$

and

$$(7.6)_p \quad \hat{\mathscr{Y}}^{(p)}(N+m_0+1) = \mu_1 - \sum_{k=0}^{M_p-1} \gamma_{+1}^{(p)}(M_p, k)(\mathscr{Y}(N-M_p+1+k) - \mu_1)$$
$$- \sum_{m=1}^{m_0} \gamma_{+2}^{(p)}(M_p, M_p-1-m_0+m)(\hat{\mathscr{Y}}^{(p)}(N+m) - \mu_p)$$
$$- \sum_{k=0}^{M_p-m_0-1} (a_1/a_p)\gamma_{+2}^{(p)}(M_p, k)(\mathscr{Y}(n-M_p+m_0+1+k)^p - \mu_p)$$
$$-(a_1/a_p)\gamma_{+2}^{(p)}(M_p, M_p-m_0)(\hat{\mathscr{Y}}^{(p)}(N+1)^p - \mu_p)$$
$$- \sum_{m=2}^{m_0} \gamma_{+2}^{(p)}(M_p, M_p-m_0-1+m)(\hat{\mathscr{Y}}^{(p)}(N+m)^p - \mu_p),$$

where $m_0 \in \{1, \cdots, M_p-2\}$ and μ_q and $a_q (1 \leq q \leq 3)$ are given in (3.14) and (3.15), respectively.

[7.4] Finally, also for the data $\mathscr{Y}_{-1} = (\mathscr{Y}(n); -1 \leq n \leq N)$ treated in [6.4], we can give two kinds of KM₂O-predictors for finite-step future $\hat{\hat{\mathscr{Y}}}^{(p)}(N+m)$ of $\mathscr{Y}_{-1} (2 \leq p \leq 3,\ 1 \leq m \leq M_p - 1)$ as follows:

$$(7.7)_p \quad \hat{\tilde{\mathscr{Y}}}^{(p)}(N+1)=\mathscr{Y}(N)+\tilde{\mu}_1-\sum_{k=0}^{M_p-1}\tilde{\gamma}^{(p)}_{+1}(M_p,k)(\tilde{\mathscr{Y}}(N-M_p+1+k)-\tilde{\mu}_1)$$
$$-\sum_{k=0}^{M_p-1}(\tilde{\alpha}_1/\tilde{\alpha}_p)\tilde{\gamma}^{(p)}_{+2}(M_p,k)(\tilde{\mathscr{Y}}(N-M_p+1+k)^p-\tilde{\mu}_p)$$

and

$$(7.8)_p \quad \hat{\tilde{\mathscr{Y}}}^{(p)}(N+m_0+1)=\hat{\tilde{\mathscr{Y}}}^{(p)}(N+m_0)+\tilde{\mu}_1$$
$$-\sum_{k=0}^{M_p-m_0-1}\tilde{\gamma}^{(p)}_{+1}(M_p,k)(\hat{\tilde{\mathscr{Y}}}^{(p)}(N-M_p+m_0+1+k)-\tilde{\mu}_1)$$
$$-\tilde{\gamma}^{(p)}_{+1}(M_p,M_p-m_0)(\hat{\tilde{\mathscr{Y}}}^{(p)}(N+1)-\mathscr{Y}(N)-\tilde{\mu}_1)$$
$$-\sum_{m=2}^{m_0}\tilde{\gamma}^{(p)}_{+1}(M_p,M_p-m_0-1+m)(\hat{\tilde{\mathscr{Y}}}^{(p)}(N+m)-\hat{\tilde{\mathscr{Y}}}^{(p)}(N+m-1)-\tilde{\mu}_1)$$
$$-\sum_{k=0}^{M_p-m_0-1}(\tilde{\alpha}_1/\tilde{\alpha}_p)\tilde{\gamma}^{(p)}_{+2}(M_p,k)(\tilde{\mathscr{Y}}(N-M_p+m_0+1+k)^p-\tilde{\mu}_p)$$
$$-(\tilde{\alpha}_1/\tilde{\alpha}_p)\tilde{\gamma}^{(p)}_{+2}(M_p,M_p-m_0)((\hat{\tilde{\mathscr{Y}}}^{(p)}(N+1)-\mathscr{Y}(N))^p-\tilde{\mu}_p)$$
$$-\sum_{m=2}^{m_0}(\tilde{\alpha}_1/\tilde{\alpha}_p)\tilde{\gamma}^{(p)}_{+2}(M_p,M_p-m_0-1+m)((\hat{\tilde{\mathscr{Y}}}^{(p)}(N+m)$$
$$-\hat{\tilde{\mathscr{Y}}}^{(p)}(N+m-1))^p-\tilde{\mu}_p),$$

where $m_0\in\{1,\cdots,M_p-2\}$ and $\tilde{\mu}_q$ and $\tilde{\alpha}_q(1\le q\le 3)$ are given by (3.18) and (3.19), respectively.

[7.5] Table 7.1(resp. Table 7.2) shows the results of Test(S) for Wolfer's sunspot numbers for 100 years from 1880 to 1979(resp. from 1889 to 1988). We find that the original data do not have the local and weak stationarity, but their first differences do have.

	(M)	(V)	(O)	(S)
1880-1979 (no difference)	0.972	0.380	1.000	NS
1879-1979 (difference)	0.986	0.873	1.000	S

Table 7.1 Test(S) for Wolfer's sunspot numbers

	(M)	(V)	(O)	(S)
1889-1988 (no difference)	1.000	0.465	1.000	NS
1888-1988 (difference)	0.944	0.713	1.000	S

Table 7.2 Test(S) for Wolfer's sunspot numbers

Therefore, by virtue of formulae (7. 3) and (7. 4), we get the results in Table 7.3 and Figure 7.1(resp. Table 7.4 and Figure 7.2), by applying the sample first difference forward KM_2O-Langevin equation in [3. 2] to the data from 1879 to 1979(resp. from 1888 to 1988).

KM_2O-predictor for Wolfer's sunspot numbers

year	observation	KM_2O-predictor
1980	154.6	163.9
1981	140.9	154.1
1982	115.9	133.1
1983	66.6	107.5
1984	45.9	87.0
1985	17.9	73.0
1986	13.4	69.3
1987	29.2	73.6
1988	100.2	110.2

Table 7. 3 (1980-1988)

year	KM_2O-predictor
1989	142.6
1990	162.5
1991	153.1
1992	124.9
1993	90.3
1994	64.2
1995	54.5
1996	37.8
1997	50.3

Table 7. 4 (1989-1997)

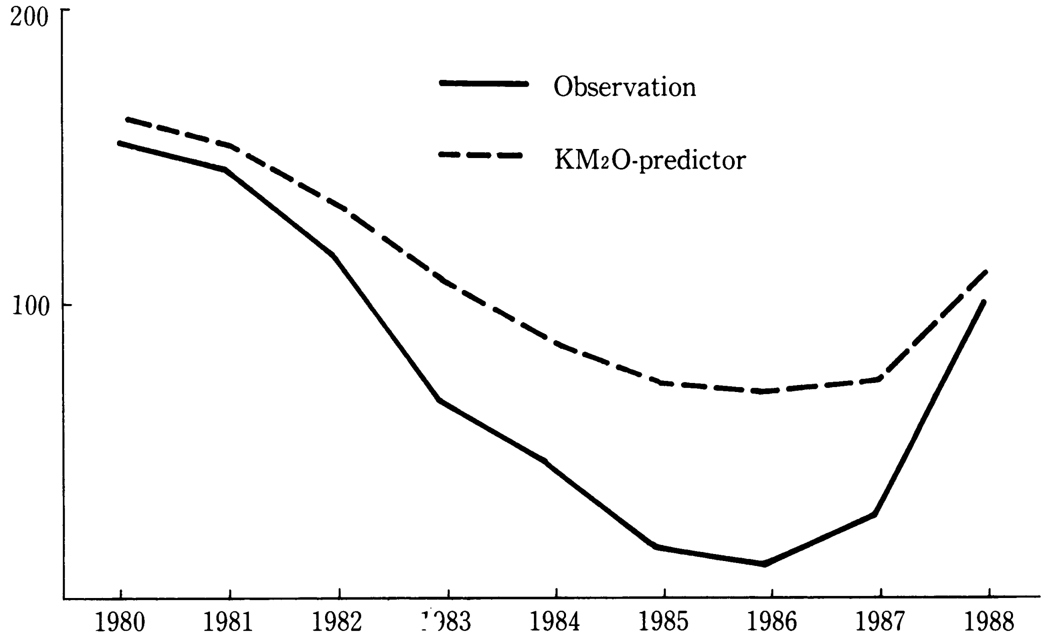

Figure 7. 1 Prediction for Wolfer's sunspot numbers

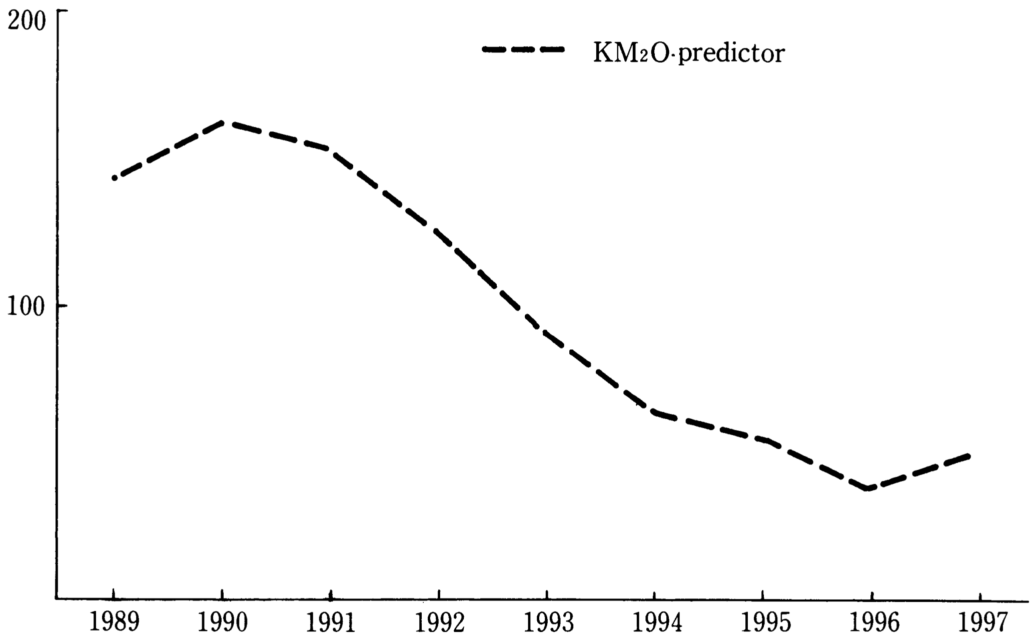

Figure 7.2 Prediction for Wolfer's sunspot numbers

References

[1] AKAIKE, H., Fitting autoregressive models for prediction, Ann. Inst. Statist. Math., 21 (1969), 243-247.

[2] AKAIKE, H., Autoregressive model fitting for control, Ann. Inst. Statist. Math., 23 (1971), 163-180.

[3] AKAIKE, H. and T. NAKAGAWA, Statistical analysis and control for a dynamical system, Science Company, 1973 (in Japanese).

[4] AKAIKE, H., Information theory and an extension of the maximum likelihood principle, 2nd Inter. Symp. on Information Theory (Petrov, B. W. and Csaki, F. eds.), Akademiaikiado, Budapest, 1973, 261-281.

[5] BOX, G. E. P. and G. M. JENKINS, Time series analysis, Forcasting and Control, Holden-Day, 1976.

[6] DURBIN, J., The fitting of time series model, Rev. Int. Stat., 28 (1960), 233-244.

[7] ELTON, C. and M. NICHOLSON, The ten-year cycle in numbers of the Lynx in Canada, J. Animal Ecology, 11 (1942), 215-244.

[8] FUJISAKA, H. and T. YAMADA, Theoretical study of time correlation functions in a discrete chaotic process, Z. Naturgorsch, 33a (1978), 1455-1460.

[9] GRANGER, C. W. J., Investigating causal relations by econometric models and cross-spectral methods, Econometrica, 37 (1969), 424-438.

[10] JEVONS, W. S., Commercial crises and sun-spots, Nature, 15 (1878), 33-37.

[11] 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.

[12] LEVINSON, N., The Wiener RMS error criterion in filter design and prediction, J. Math. Phys., 25 (1947), 261-278.

[13] MAKAGON, A. and V. MANDREKAR, On complex operator-valued O-U processes, *T*-positivity, and innovations of Okabe and Masani, J. Multivar. Anal., 29 (1989), 68-83.

[14] MATA, C. G. and F. I. SHAFFNER, Solar and economic relationships; a preliminary report, The Qurterly Journal of Economics, 49 (1935), 1-51.

[15] MAY, R. M., Simple mathematical models with very complicated dynamics, Nature, 261 (1976), 459-467.

[16] MORAN, P. A. P., The statistical analysis of the sunspot and lynx cycles, J. Animal Ecology, 18 (1949), 115-116.

[17] MORAN, P. A. P., The statistical analysis of the Canadian lynx cycle I. structure and prediction, J. Animal Ecology, 22 (1951), 163-173.

[18] MORI, H., B.-C. SO and T. OSE, Time correlation functions of one-dimensional transformations, Prog. Theor. Phys., 66 (1981), 1266-1283.

[19] NAKANO, Y. and Y. OKABE, On a multi-dimensional $[\alpha, \beta, \gamma]$-Langevin equation, Proc. Japan Acad., 59 (1983), 171-173.

[20] 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. 26 (1979), 115-165.

[21] OKABE, Y., 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, 28 (1981), 169-213.

[22] OKABE, Y., 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. 28 (1982), 793-804.

[23] OKABE, Y., On a Langevin equation, Sugaku, 33 (1981), 306-324 (in Japanese).

[24] OKABE, Y., On a wave equation associated with prediction errors for a stationary Gaussian process, Leccture Notes in Control and Information Sciences, 49 (1983), 215-226.

[25] OKABE, Y., A generalized fluctuation-dissipation theorem for the one-dimensional diffusion process, Commun. Math. Phys., 98 (1985), 449-468.

[26] OKABE, Y., On KMO-Langevin equations for stationary Gaussian process with *T*-positivity, J. Fac. Sci. Univ. Tokyo, Sect. IA. 33 (1986), 1-56.

[27] OKABE, Y., On the theory of Brownian motion with the Alder-Wainwright effect, J. Stat. Phys., 45 (1986), 953-981.

[28] OKABE, Y., KMO-Langevin equation and fluctuation-dissipation theorem (I), Hokkaido Math. J., 15 (1986), 163-216.

[29] OKABE, Y., KMO-Langevin equation and fluctuation-dissipation theorem (II), Hokkaido Math. J., 15 (1986), 317-355.

[30] OKABE, Y., 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.

[31] OKABE, Y., On long time tails of correlation functions for KMO-Langevin equations, Proceeding of Fourth Japan-USSR symposium on Probability theory, Kyoto, July, 1986, Lecture Notes in Mathematics, Springer, vol. 1299, 391-397.

[32] OKABE, Y., On the theory of discrete KMO-Langevin equations with reflection positivity (I), Hokkaido Math. J., 16 (1987), 315-341.

[33] OKABE, Y., On the theory of discrete KMO-Langevin equations with reflection

positivity (II), Hokkaido Math. J., 17 (1988), 1-44.
[34] OKABE, Y., On the theory of discrete KMO-Langevin equations with reflection positivity (III), Hokkaido Math. J., 18 (1989), 149-174.
[35] 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, 1988, vol. 2, 601-645.
[36] OKABE, Y. and A. INOUE, The theory of KM_2O-Langevin equations and its applications to data analysis (II) : Causal analysis, to be submitted to Nagoya Math. J.
[37] PRIESTLEY, M. B., Non-linear and non-stationary time series, Academic Press, 1988.
[38] 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).
[39] WALKER, G., On periodicity in series of related terms, Proc. Roy. Soc., Ser. A, 131 (1931), 518-532.
[40] WHITTLE, P., On the fitting of multivariate autoregressions, and the approximate canonical factorization of a spectral density matrix, Biometrika, 50 (1963), 129-134.
[41] WIGGINS, R. A. and E. A. ROBINSON, Recursive solution to the multichannel fitting problem, J. Geophys. Res., 70 (1965), 1885-1891.
[42] YULE, G. U., On a method of investigating periodicities in disturbed series, with special reference 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

Faculty of Economics
Shiga University
Hikone
522 Japan