

マルコフ時系列情報の変数変換

石井直宏・鈴木宣夫

情報工学科

(1976年 9月 11日受理)

Transformation of Random Variables of Markov Process

Naohiro ISHII and Nobuo SUZUMURA

Department of Information Engineering

(Received September 11, 1976)

In recent years, there has been increasing interest in identification and prediction of systems and spectrum estimation by the autoregressive process. In this procedure, it is assumed that the time series is nearly Gaussian. However, physical or biological data are not necessarily Gaussian. Therefore it is desirable to develop some methods to represent non-Gaussian data exactly. In this paper, to tackle the above problems, a procedure is developed by the transformation of non-Gaussian variables.

First, it is proved that the order of Markov process is invariant by the transformation of diffeomorphism. Next, it is shown that the amount of Kullback information which estimates the order of Markov process, is preserved under the transformation. Thus there is no loss of information by diffeomorphism. By applying these results, non-Gaussian data are transformed into Gaussian ones. The computer simulation shows that the method developed here is applicable to the estimation of regression coefficients in the non-Gaussian case.

1. まえがき

一般に、システムを記述する変数が確率変数と見なされるとき、このシステムは確率的システムといわれる。この概念は広範なもので、通信システム、制御システム等の大部分のものはこの範ちゅうに入る確率システムの動作の状態を線形式で表現するものがある。これらはマルコフ過程の表現の問題として研究されている。また時系列的に生起する事象がランダムに生起するか、あるいは先行する事象に従属して生起するかという問題はマルコフ連鎖モデルとして取り扱うことができる。このモデルの変数が正規過程に従う場合、正規マルコフ過程と呼ばれ、線形式で表現されることが証明されている。このような線形式で表現される過程、すなわち線形過程の基本型として自己回帰方程式がある。

定常正規確率過程を表わす自己回帰方程式の研究の1つはシステムの同定、予測であり、他に新しいパワースペクトルの算出法にある。このような場合に自己回帰モ

デルの次数をいかに推定するかということが重要な問題となる。次数が推定されると、回帰係数の推定は最小2乗法によるパラメータの推定の問題となり変数が正規過程であれば、最尤推定にも一致している。自己回帰モデルの次数の推定法はいくつかが³⁾提案されているが、感度のよい次数の推定法が望まれる¹⁾。しかし従来の手法は、変数が正規過程に従うという前提条件にもとづいている。したがって、これらの手法を環境データや生物系の実験データにそのまま適用する際には注意を必要とする。その理由は、これらのデータは必ずしも正規過程に従わない大気汚染の実測データは決して正規過程と言えないので、実測データの3乗根、4乗根あるいは対数をとったものがほぼ正規分布に従うことから、非正規過程のデータ解析の手法が望まれてきている⁴⁾。

本研究は非正規過程のデータの取り扱いとして、微分同相写像による変数変換をとり上げ、この変換によるマルコフ過程の次数の不変性を示す。次に微分同相写像による Kullback 情報量の不変性を示す。この微分同相写

像により、非正規過程を正規過程へ変換した場合と非正規過程をそのままノンパラメトリック統計量でデータを解析した場合の簡単な比較を試みた。その結果、写像後のデータの解析結果がノンパラメトリックのデータ解析の結果とほとんど一致することを回帰分析により明らかにした。このことは非正規過程の場合、微分同相写像の意味が有効であることを示すものである。

2. 確率過程における微分同相写像について

確率変数の系列、すなわち確率過程 $\{X_n, n \in \mathbb{Z}\}$ を考える。確率変数のとる値は実数の連続値とする。 X_n は時刻 n での確率変数、 \mathbb{Z} は整数の集合を表わし、 $n \in \mathbb{Z}$ は n が $\dots, -1, 0, 1, 2, \dots$ をとることを示す。

確率過程 $\{X_n, n \in \mathbb{Z}\}$ が定常であるとは、任意の k 個の確率変数 $X_n, X_{n+1}, \dots, X_{n+k-1}$ の結合確率密度が等しいことである。次に定常確率過程の時間依存性を表わす多重マルコフ性を定義する。定常確率過程 $\{X_n, n \in \mathbb{Z}\}$ が m 重マルコフ過程であるとは生起確率において

$$P(X_{m+1} \in A | X_1, X_2, \dots, X_m) \\ \neq P(X_{m+1} \in A | X_2, X_3, \dots, X_m)$$

なる集合 $A \in \mathcal{B}$ (実数 R 上のボレル集合体) が存在し、かつ

$$P(X_{m+1} \in A | X_{-i}, X_{-i+1}, \dots, X_m) \\ = P(X_{m+1} \in A | X_{-i+1}, X_{-i+2}, \dots, X_m) \quad (1) \\ (i=0, 1, 2, \dots)$$

が $A \in \mathcal{B}$ について確率 1 で成立する場合をいう。以下本節ではマルコフ性の多重度がいかなる変換によって不変であるかを調べる。

n 次元ユークリッド空間 R^n から R^n への写像

$$\varphi: p \longrightarrow \varphi(p), \quad p \in R^n, \quad \varphi(p) \in R^n$$

が 1 対 1 onto 写像で、微分可能であり、逆写像も微分可能であるとする。このとき写像 φ を R^n から R^n への微分同相写像という。領域 R^n と値域 R^n が写像 φ によって微分同相であるための必要十分条件は $\varphi: p \rightarrow \varphi(p)$ が 1 対 1 onto 写像で微分可能であり、任意の点 $p \in R^n$ で $\varphi(p)$ のヤコビ行列式が 0 でないことである。これは逆写像の存在の言いかえである。

さて、定常確率過程 $\{X_n, n \in \mathbb{Z}\}$ の確率変数 X_n を変換して

$$T: X_n \rightarrow Y_n$$

すなわち $Y_n = T(X_n)$ となる写像が実数 R から R への微分同相写像とする。このとき微分同相写像による多重マルコフ性の不変な量として、次の命題が証明できる。

定常確率過程 $\{X_n, n \in \mathbb{Z}\}$ について、確率変数 X_n の微

分同相写像 T

$$T: X_n \rightarrow Y_n \quad \text{すなわち} \quad Y_n = T(X_n)$$

により得られた $\{Y_n, n \in \mathbb{Z}\}$ のマルコフ性の次数と $\{X_n, n \in \mathbb{Z}\}$ のマルコフ性の次数は等しい。

まず証明すべきことを定式化する。定常確率過程 $\{X_n, n \in \mathbb{Z}\}$ のマルコフ性の次数を m と仮定する。定義(1)式における多重マルコフ性の定義は条件付確率密度を用いて、いにかえることができ、一般に X_1, X_2, \dots, X_n の条件のもとで X_{n+1} の条件付確率密度を

$$f(X_{n+1} | X_1, X_2, \dots, X_n)$$

とすれば、 $\{X_n, n \in \mathbb{Z}\}$ が m 重マルコフ過程であるための必要十分条件は

$$f(X_{m+1} | X_1, X_2, \dots, X_m) \neq f(X_{m+1} | X_2, X_3, \dots, X_m)$$

かつ

$$f(X_{m+1} | X_{-i}, X_{-i+1}, \dots, X_m) \\ = f(X_{m+1} | X_{-i+1}, X_{-i+2}, \dots, X_m) \quad (2) \\ (i=0, 1, 2, \dots)$$

である。さらに Y_1, Y_2, \dots, Y_n の条件のもとでの Y_{n+1} の条件付確率密度を

$$g(Y_{n+1} | Y_1, Y_2, \dots, Y_n)$$

と記述すれば、命題の証明は(2)式の関係式が(3)式と同値の関係にあることを示せばよい。

$$g(Y_{m+1} | Y_1, Y_2, \dots, Y_m) \neq g(Y_{m+1} | Y_2, Y_3, \dots, Y_m)$$

かつ

$$g(Y_{m+1} | Y_{-i}, Y_{-i+1}, \dots, Y_m) \\ = g(Y_{m+1} | Y_{-i+1}, Y_{-i+2}, \dots, Y_m) \quad (3) \\ (i=0, 1, 2, \dots)$$

写像 T が微分同相写像であるから、微分同相な逆写像 T^{-1} 、すなわち

$$X_n = T^{-1}(Y_n)$$

が存在する、微分同相写像による確率密度の変換の関係は

$$g(Y_1, Y_2, \dots, Y_n) = f\{T^{-1}(Y_1), T^{-1}(Y_2), \dots, T^{-1}(Y_n)\} \\ \times \left| \frac{\partial (X_1, X_2, \dots, X_n)}{\partial (Y_1, Y_2, \dots, Y_n)} \right|$$

で与えられる。ここに $|\partial (X_1, X_2, \dots, X_n) / \partial (Y_1, Y_2, \dots, Y_n)|$ はヤコビ行列式であり。

$$\left| \frac{\partial (X_1, X_2, \dots, X_n)}{\partial (Y_1, Y_2, \dots, Y_n)} \right| = \begin{vmatrix} \frac{\partial X_1}{\partial Y_1} & \frac{\partial X_2}{\partial Y_1} & \dots & \frac{\partial X_n}{\partial Y_1} \\ \vdots & \vdots & & \vdots \\ \frac{\partial X_1}{\partial Y_n} & \frac{\partial X_2}{\partial Y_n} & \dots & \frac{\partial X_n}{\partial Y_n} \end{vmatrix} \equiv \Delta_n \quad (4)$$

である。さて(3)式の条件付確率密度は

$$\begin{aligned} g(Y_{m+1}|Y_{-i+1}, \dots, Y_m) &= \frac{g(Y_{-i}, Y_{-i+1}, \dots, Y_m, Y_{m+1})}{g(Y_{-i}, Y_{-i+1}, \dots, Y_m)} \\ &= \frac{f\{T^{-1}(Y_{-i}), T^{-1}(Y_{-i+1}), \dots, T^{-1}(Y_{m+1})\}}{f\{T^{-1}(Y_{-i}), T^{-1}(Y_{-i+1}), \dots, T^{-1}(Y_m)\}} \\ &\quad \cdot \frac{\Delta_{m+1}}{\Delta_m} \end{aligned}$$

ここで、微分同相写像の条件、1対1より

$$\frac{\partial X_i}{\partial Y_j} = 0 \quad (i \neq j)$$

$$\frac{\partial X_i}{\partial Y_j} \neq 0 \quad (i=j)$$

となる。したがって(4)式は

$$\Delta n \equiv \begin{vmatrix} \frac{\partial X_1}{\partial Y_1} & 0 \\ 0 & \frac{\partial X_2}{\partial Y_2} \\ 0 & \frac{\partial X_n}{\partial Y_n} \end{vmatrix} = \left| \frac{\partial X_1}{\partial Y_1} \right| \cdot \left| \frac{\partial X_2}{\partial Y_2} \right| \cdots \left| \frac{\partial X_n}{\partial Y_n} \right| \quad (5)$$

で与えられる。ここで $\partial X_i / \partial Y_i = J_i$ とおくと、(3)式の条件付確率密度は

$$\begin{aligned} g(Y_{m+1}|Y_{-i}, Y_{-i+1}, \dots, Y_m) &= \frac{g(Y_{-i}, Y_{-i+1}, \dots, Y_m, Y_{m+1})}{g(Y_{-i}, Y_{-i+1}, \dots, Y_m)} \\ &= \frac{f\{T^{-1}(Y_{-i}), T^{-1}(Y_{-i+1}), \dots, T^{-1}(Y_{m+1})\}}{f\{T^{-1}(Y_{-i}), T^{-1}(Y_{-i+1}), \dots, T^{-1}(Y_m)\}} \\ &\quad \times \frac{\begin{vmatrix} J_{-i}(Y_i) & 0 \\ 0 & J_{m+1}(Y_{m+1}) \end{vmatrix}}{\begin{vmatrix} J_{-i}(Y_i) & 0 \\ 0 & J_m(Y_m) \end{vmatrix}} \\ &= \frac{f\{T^{-1}(Y_{-i}), T^{-1}(Y_{-i+1}), \dots, T^{-1}(Y_{m+1})\}}{f\{T^{-1}(Y_{-i}), T^{-1}(Y_{-i+1}), \dots, T^{-1}(Y_m)\}} \\ &\quad \times \frac{|J_{-i}| |J_{-i+1}| \cdots |J_m| |J_{m+1}|}{|J_{-i}| |J_{-i+1}| \cdots |J_m|} \\ &= f\{T^{-1}(Y_{m+1}) | T^{-1}(Y_{-i}), T^{-1}(Y_{-i+1}), \dots, T^{-1}(Y_m)\} |J_{m+1}| \end{aligned}$$

ここで(2)式の第2番目の式より

$$\begin{aligned} &= \frac{f\{T^{-1}(Y_{-i+1}), T^{-1}(Y_{-i+2}), \dots, T^{-1}(Y_{m+1})\}}{f\{T^{-1}(Y_{-i+1}), T^{-1}(Y_{-i+2}), \dots, T^{-1}(Y_m)\}} \\ &\quad \times \frac{|J_{-i+1}| |J_{-i+2}| \cdots |J_m| |J_{m+1}|}{|J_{-i+1}| |J_{-i+2}| \cdots |J_m|} \\ &= \frac{g(Y_{-i+1}, Y_{-i+2}, \dots, Y_m, Y_{m+1})}{g(Y_{-i+1}, Y_{-i+2}, \dots, Y_m)} \\ &= g(Y_{m+1}|Y_{-i+1}, Y_{-i+2}, \dots, Y_m) \quad (6) \end{aligned}$$

となる。これで f の方から g の方への証明ができたが、

g の方から f の方への証明も、写像 T が微分同相写像であり、逆写像 T^{-1} もそうであることから、問題の条件の対称性により同様に示すことができる。上述の命題は生起確率で次のような命題でいいかえることもできる。

定常確率過程 $\{X_n, n \in \mathbb{Z}\}$ が m 重マルコフ過程であるための必要十分条件は、微分同相写像 $T: X_n \rightarrow Y_n$ によって変換された $Y_n, n=1, 2, 3, \dots$ について

$$P(Y_{m+1} \in A | Y_1, Y_2, \dots, Y_m) \neq P(Y_{m+1} \in A | Y_2, Y_3, \dots, Y_m)$$

を満足する A が存在し、同時に

$$\begin{aligned} P(Y_{m+1} \in A | Y_{-i}, Y_{-i+1}, \dots, Y_m) \\ = P(Y_{m+1} \in A | Y_{-i+1}, Y_{-i+2}, \dots, Y_m) \\ (i=0, 1, 2, \dots) \end{aligned} \quad (7)$$

が成立することである。ここに $A \in B$ で B は実数の集合 R 上のボレル集合体である。本節で重要なことは、写像 T によって実現する確率密度関数が多変量正規分布 (平均 μ , 共分散行列 A)

$$N(Y) = \frac{1}{(2\pi)^{n/2} |A|} \exp \left\{ -\frac{1}{2} (Y - \mu)^T A^{-1} (Y - \mu) \right\}$$

で与えることが出来る場合である。ここで

$$\begin{aligned} Y &= (Y_1, Y_2, \dots, Y_n) = (T(X_1), T(X_2), \dots, T(X_n)) \\ \mu &= (\mu_1, \mu_2, \dots, \mu_n) \end{aligned}$$

および

$$A = E[(Y - \mu)(Y - \mu)^T]$$

で与えられる。すなわち (X_1, X_2, \dots, X_n) の結合確率密度の写像による (Y_1, Y_2, \dots, Y_n) の確率密度が多変量正規分布となり

$$f(X_1, X_2, \dots, X_n) = \frac{N(Y_1, Y_2, \dots, Y_n)}{|J_1(Y_1)| |J_2(Y_2)| \cdots |J_n(Y_n)|}$$

の場合である

3. 写像による Kullback 情報量の不変性

微分同相写像 T による Kullback 情報量の不変性を示す。確率空間 $(X, S, \mu_i), i=1, 2$ が定義されている。測度 μ_1 と測度 μ_2 は絶対連続であり ($\mu_1 \equiv \mu_2$)、さらに λ がこれらの測度に絶対連続であれば、Radon-Nikodym の定理により、任意の $E \in S$ に対して

$$\mu_i(E) = \int_E f_i(X) d\lambda(X), \quad i=1, 2 \quad (8)$$

を満足する確率密度 $f_i(X)$ が存在する。いま $H_i, i=1, 2$ を各々変数 X が測度 $\mu_i, i=1, 2$ を有する各々母集団から抽出されたと仮定する。 H_1 と H_2 の類別のための X の情報量を

$$\log(f_1(X)/f_2(X)) \quad (9)$$

で定義する, また (1:2) を μ_1 でみたときの H_1 と H_2 を類別するための平均情報量とすれば,

$$I(1:2) = \int d\mu_1 \log_1 \frac{f_1(X)}{f_2(X)} = \int f_1(X) \log \frac{f_1(X)}{f_2(X)} d\lambda(X) \quad (10)$$

と定義される。確率空間 (X, S, μ_i) から (Y, U, ν_i) への変換を T とする。任意の $G \in U$ に対して

$$\nu_i(G) = \mu_i(T^{-1}G), \quad i=1, 2$$

と仮定すると $\nu_1 \equiv \nu_2 \equiv \gamma$ となる。(8)式と同様に

$$\nu_i(G) = \int_G g_i(Y) d\gamma(Y) \quad (11)$$

となる確率密度 $g_i(Y)$ が存在する。確率空間 (Y, U, ν_i) の上での平均情報量は

$$\begin{aligned} I'(1:2) &= \int d\nu_1(Y) \log \frac{g_1(Y)}{g_2(Y)} \\ &= \int d\mu_1(X) \log \frac{g_1 T(X)}{g_2 T(X)} \\ &= \int f_1(X) \log \frac{g_1 T(X)}{g_2 T(X)} d\lambda(X) \end{aligned} \quad (12)$$

となる, 次の命題は Kullback によって示されたものである。

平均情報量 $I(1:2)$ と $I'(1:2)$ において

$$I(1:2) \geq I'(1:2) \quad (13)$$

が成立し, 等号は T が十分統計量のときのみ成立する。

確率変数への作用素を導入し, マルコフ過程の次数の推定と Kullback 情報量との関連を明らかにする。確率変数 X_n へミキシング作用素を作用させることを

$$SX_n \equiv X_{(r \times N + 1)} \quad (14)$$

で表わす, ここで r は一様乱数, N は時系列データのサンプル総数, $[y]$ は実数 y のガウス記号である。 SX_n は時系列の部分系列 $\{X_n, X_{n+1}, X_{n+2}, \dots, X_{n+m}\}$ から新しい部分系列 $\{SX_n, X_{n+1}, X_{n+2}, \dots, X_{n+m}\}$ のように構成する, ミキシング作用素は部分系列 $\{X_1, X_2\}$ へ作用させると

$$\lim_{N \rightarrow \infty} E \left[E \left[P_N(X_2 | SX_1) - P(X_2) \right] \right] = 0 \quad (15)$$

の成立することが証明できる。ここで $P_N(Y)$ はサンプル個数 N の $P(Y)$ の推定値である。 $E[y]$ は変数 y の期待値であるが S はミキシングを意味する。まず

$$E_S \left[P_N(X_2 | SX_1) - P(X_2) \right] = E_S \left[P_N(X_2 | SX_1) \right] - P(X_2) \quad (16)$$

が成立する, (16)式の第1項は(17)式となることを証明する

$$E_S \left[P_N(X_2 | SX_1) \right] = P_N(X_2) \quad (17)$$

いま m_1 を変数 SX_1 のサンプルの数とする。原データ系列で $\{X_1, X_2\}$ の対の生起確率を $P_N(X_2)$ とする。 x_2 は x_1 の前の任意の変数を表わす。もし $(x_1, x_2) \in \{X_1, X_2\}$ と $(sx_1, x_2) \in \{SX_1, X_2\}$ とが等しい対であれば, そのような

対の数を k とする。このとき $P_N(X_2 | SX_1)$ の期待値は $P_N(X_2)$ で表わすことができる, まず

$$\begin{aligned} P \left[P_N(X_2 | SX_1) = \frac{k}{m_1} \right] &= P \left[P_N(SX_1, X_2) / P_N(SX_1) \right. \\ &= \frac{k}{m_1} \left. \right] = \left(\frac{m_1}{k} \right) P_N(X_2)^k (1 - P_N(X_2))^{m_1 - k} \end{aligned} \quad (18)$$

(18)式より

$$\begin{aligned} E_S \left[P_N(X_2 | SX_1) \right] &= \sum_{k=0}^{m_1} \frac{k}{m_1} \binom{m_1}{k} P_N(X_2)^k (1 - P_N(X_2))^{m_1 - k} \\ &= P_N(X_2) \end{aligned} \quad (19)$$

が成立する, 次に確率密度を推定する場合の Parzen⁵⁾の方法を適用すると次の式が導かれる。

$$\lim_{N \rightarrow \infty} E \left[E \left[P_N(X_2) (S_1) \right] \right] = \lim_{N \rightarrow \infty} E \left[P_N(X_2) \right] \quad (20)$$

となり

$$\begin{aligned} E_N \left[P_N(X_2) \right] &= E_N \left[\left(\frac{1}{h} \right) K \left[(X_2 - x_2) / h \right] \right] \\ &= \int_{-\infty}^{\infty} \left(\frac{1}{h} \right) k \left[(X_2 - \xi_2) / h \right] p(\xi_2) d\xi_2 \\ &= P(X_2) \int_{-\infty}^{\infty} K(Y/h) dY/h \end{aligned} \quad (21)$$

ここで

$$\int_{-\infty}^{\infty} K(Y/h) dY/h = \int_{-\infty}^{\infty} K(z) dz = 1 \quad (22)$$

が成立し, $(1/h) K[Y/h]$ は推定値の kernel を表わす(20)式, (21)式および(22)式より命題が証明された。なお同様に, $P_N(X_2 | SX_1)$ の分散は次式で与えられる

$$\begin{aligned} \text{Var} \left[P_N(X_2 | SX_1) \right] &= \sum_{k=0}^{m_1} \{ (k/m_1) - P_N(X_2) \}^2 \binom{m_1}{k} \\ &= P_N(X_2)^k (1 - P_N(X_2))^{m_1 - k} \\ &= P_N(X_2) (1 - P_N(X_2)) / m_1 = (1 - P_N(X_2)) / N \end{aligned} \quad (23)$$

Parzen の Kernel により, $P_N(X_2)$ は次式で表現される

$$P_N(X_2) = \frac{1}{N} \sum_{i=1}^N \frac{1}{h} K[(X_2 - x_i) / h] \quad (24)$$

ここで

$$K(Y/h) = \begin{cases} 1/2 & |Y/h| \leq 1 \text{ の時} \\ 0 & |Y/h| > 1 \text{ の時} \end{cases}$$

となる, (24)式は N 個の変数 $(1/h) K[(X_2 - x_2) / h]$ の和であり中心極限定理を適用すると $P_N(X_2)$ は漸近的に正規分布へ近づく。また十分大きなサンプル個数 N に対し, 二項分布の正規分布近似から $P_N(X_2 | SX_1)$ は平均(19)式, 分散(23)式の正規分布近似と見なすことができる, 次にミキシング作用素を導入した結合情報量 Z_i^{mix} および Z_i^{mix} を次式に定義する。

$$\begin{aligned} Z_i^{mix} &= - \sum_i \sum_{i=1} \dots \sum_1 P(X_i, X_{i-1}, \dots, X_2, SX_1) \\ &\quad \log P(X_i, X_{i-1}, \dots, X_2, SX_1) \end{aligned}$$

および

$$Z_i^{mix} = -\sum_i \sum_{i=1} \dots \sum_i \hat{P}(X_i, X_{i-1}, \dots, X_2, SX_1) \log \hat{P}(X_i, X_{i-1}, \dots, X_2, SX_1) \quad (25)$$

ここで

$$E_s[P(X_i, X_{i-1}, \dots, X_2, SX_1)] = \hat{P}(X_i, X_{i-1}, \dots, X_2, SX_1)$$

であり, $\hat{P}(X_i, X_{i-1}, \dots, X_2 | SX_1)$ は上述までの展開と同様にして十分大きなサンプル個数 N に対して $P(X_i, X_{i-1}, \dots, X_2)$ に近づく。単純マルコフ過程を条件付エントロピーで表現すれば,

$$H_0 \geq H_1 > H_2 = H_3 = \dots \quad (26)$$

で表わされる。ここで

$$H_i = -\sum_{i+1} \sum_i \dots \sum_i P(X_{i+1}, \dots, X_1) \cdot \log P(X_{i+1}, X_i, \dots, X_1)$$

である。図式を使って図式は,

$$H_0 \geq H_2^{mix} \neq H_2 = H_3 = \dots \quad (27)$$

のように記述できる, ここで

$$H_2^{mix} = Z_3^{mix} - Z_2^{mix} \quad (28)$$

上述の関係式は m 重マルコフ過程の必要十分条件の導出にも同様に適用できる。この条件は十分大きなサンプル個数に対して

$$H(X_{m+1} | X_m, X_{m-1}, \dots, X_2, X_1) \neq H(X_{m+1} | X_m, X_{m-1}, \dots, X_2, SX_1)$$

および

$$H(X_{m+1} | X_m, X_{m+1}, \dots, X_{-1+2}, X_{-i+1}) = H(X_{m+1} | X_m, X_{m-1}, \dots, X_{-i+2}, SX_{-i+1}) \quad (i=0, 1, 2, 3, \dots) \quad (29)$$

で与えられる, ここで $H(X_{m+1} | X_m, X_{m-1}, \dots, X_1)$ は X_1, X_2, \dots, X_m の条件のもとでの X_{m+1} の条件付情報量である $H(X_{m+1} | X_m, X_{m-1}, \dots, X_2, SX_1)$ は図式と同様な意味である, 図式の第2番目の式を分解すると

$$Z_{m+1} - Z_m = Z_{m+1}^{mix} - Z_m^{mix}$$

すなわち

$$Z_{m+1}^{mix} - Z_{n+1} = Z_m^{mix} - Z_m$$

で与えられる, ここで新たに

$$U \equiv Z_m^{mix} - Z_i$$

を定義する。次に $T: X_n \rightarrow Y_n$ の微分同相写像により U の値が不変であることを示す, 確率密度関数を p とすると,

$$U(X_1, X_2, \dots, X_i) = \int p(X_1, X_2, \dots, X_i) \log p(X_1, X_2, \dots, X_i) dX_1 dX_2 \dots dX_i \\ - \int p(SX_1, X_2, \dots, X_i) \log p(SX_1, X_2, \dots, X_i) dX_1 dX_2 \dots dX_i$$

において

$$dX_1 dX_2 \dots dX_i = \left| \frac{\partial(X_1 X_2 \dots X_i)}{\partial(Y_1 Y_2 \dots Y_i)} \right| dY_1 dY_2 \dots dY_i$$

$$p(X_1, X_2, \dots, X_i) = \left| \frac{\partial(Y_1 Y_2 \dots Y_i)}{\partial(X_1 X_2 \dots X_i)} \right| p'(Y_1, Y_2, \dots, Y_i)$$

したがって微分同相写像の条件より

$$U(X_1, X_2, \dots, X_i) = \int p'(Y_1, Y_2, \dots, Y_i) \log p'(Y_1, Y_2, \dots, Y_i) |J_1| |J_2| \dots |J_i| dY_1 dY_2 \dots dY_i \\ - \int p'(SY_1, Y_2, \dots, Y_i) \log p'(SY_1, Y_2, \dots, Y_i) |J_2| \dots |J_i| dY_1 dY_2 \dots dY_i \\ = \int p'(Y_1, Y_2, \dots, Y_i) \log p'(Y_1, Y_2, \dots, Y_i) dY_1 dY_2 \dots dY_i \\ - \int p'(SY_1, Y_2, \dots, Y_i) \log p'(SY_1, Y_2, \dots, Y_i) dY_1 dY_2 \dots dY_i \\ \equiv U'(Y_1, Y_2, \dots, Y_i) \quad (30)$$

となる, また図式の情報量 U は図式において

$$f_1(X_1, X_2, \dots, X_i) = p(X_1, X_2, \dots, X_i) \\ f_2(X_1, X_2, \dots, X_i) = p(SX_1) p(X_2, X_3, \dots, X_i) = p(X_1) p(X_2, X_3, \dots, X_i)$$

とおけば

$$U(X_1, X_2, \dots, X_i) = I(1:2) \quad (31)$$

が証明できる。すなわち $U(X_1, X_2, \dots, X_i)$ は図式で定義した。Kullback 情報量となる。図式, 図式より微分同相写像 T が図式の等号を満足することにより, T は十分統計量となる。したがって微分同相写像の変換によって情報量図式の U の損失はない。よって図式の写像による情報の損失はないことを示す。

以上により, 本節までの結論はマルコフ過程の微分同相写像による変換で, マルコフ次数が写像前と後で不変である。さらに次数を推定するための必要十分条件式は図式で与えられる。図式を分解して図式のように変換すると Kullback 情報量となり, 情報量 U は微分同相写像前と後でその値が不変である。したがって写像は十分統計量となり情報の損失がないことが明らかとなった。

4. 定常正規多重マルコフ過程の表現

定常正規確率過程 $\{X_n, n \in \mathbb{Z}\}$ の多重マルコフ過程の確率差分方程式による表現として, その基本型に自己回帰方程式がある。以下一般性を失うことなく

$$E(X_n) = 0, \quad E(X_n^2) = 1$$

と仮定する, 次の命題の成立することが証明されている 定常正規確率過程 $\{X_n, n \in \mathbb{Z}\}$ が, m 重マルコフ過程であるための必要十分条件は, X_n が次の確率差分方程式を満足することである。

$$X_n + a_1 X_{n-1} + \dots + a_m X_{n-m} = \xi_n \quad (32)$$

ただし条件として

(i) $\{\xi_n, n \in \mathbb{Z}\}$ は独立な定常正規過程である。

(ii) ξ_n と X_{n-i} ($i=1, 2, \dots, m$) は統計的独立である。

(iii) a_k ($k=1, 2, \dots, m$) は代数方程式

$$Z^m + a_1 z^{m-1} + \dots + a_{m-1} z + a_m = 0$$

の根が単位円の内部にあること

(iv) 定常正規過程は非特異 (non-singular) である, 正規過程が非特異であるとは, n 個の確率変数 $X = (X_1, X_2, \dots, X_n)$ の結合確率密度を

$$f(X) = \frac{1}{(2\pi)^{n/2} |A|^{1/2}} \exp \left\{ -\frac{1}{2} (X - \mu)^T A^{-1} (X - \mu) \right\}$$

とするとき A^{-1} (共分散行列の逆行列) が positive definite であることを意味する。

確率システムの同定の問題では, 時系列モデルとして③式の型の方程式が採用されている。しかしこの方程式が妥当であるのは, 観測時系列の変数が正規分布に従う場合である。③式の係数 a_k ($k=1, 2, \dots, m$) は自己回帰係数である。また③式は確率過程と差分方程式を結びつける重要な接点となっているが, 非正規過程に対しては, マルコフ性の次数や自己回帰係数など, どのように評価すべきか問題が付きまとう。

生体情報や大気汚染のデータは決して正規過程とは言えないので, 実測データの対数をとったり, 3乗根, 4乗根をとったものが正規分布に従う場合もあるので, データの前処理としてこの変換を行う。しかし実測データが任意の分布の場合, これらの変換により必ずしも正規分布になるとは限らない。

本節では必ずしも正規でない定常確率過程に対して③式の表現を一般化する。微分同相写像によるマルコフ次数が不変なことから次の命題が成立する。

定常 m 重マルコフ過程 $\{X_n, n \in \mathbb{Z}\}$ の微分同相写像 T により, 新しく得られた $\{T(X_n), n \in \mathbb{Z}\}$ が正規過程となるものとする。すなわち $\{X_n, n \in \mathbb{Z}\}$ の結合密度を $f(X_1, X_2, \dots, X_n)$ としたとき

$$N(Y_1, Y_2, \dots, Y_n) = f(X_1, X_2, \dots, X_n) |J_1| |J_2| \dots |J_n| \quad (3)$$

を満足するような写像

$$T: X_n \rightarrow Y_n$$

により, もとの m 重マルコフ過程 $\{X_n, n \in \mathbb{Z}\}$ は確率差分方程式

$$T(X_n) + \alpha_1 T(X_{n-1}) + \alpha_2 T(X_{n-2}) + \dots + \alpha_m T(X_{n-m}) = \xi_n \quad (3')$$

による表現が可能である。ここで $\{\xi_n, n \in \mathbb{Z}\}$ は独立な定常正規過程であり, ξ_n と X_{n-i} ($i=1, 2, \dots, m$) は独立である。③式の α_k ($k=1, 2, \dots, m$) は次の代数方程式

$$Z^m + \alpha_1 Z^{m-1} + \dots + \alpha_{m-1} Z + \alpha_m = 0$$

の根が単位円の内部に存在するように決定されなければならない。

5. 微分同相写像の実現

微分同相写像とマルコフ性の関係は前節までに明らか

にした。ここでは非正規過程が微分同相写像により, 正規過程へ変換するような一般性のある方法を示す。

与られた確率過程が非正規過程であり, その累積分布関数が Fig. 1 (A) のように表わされているものとする。

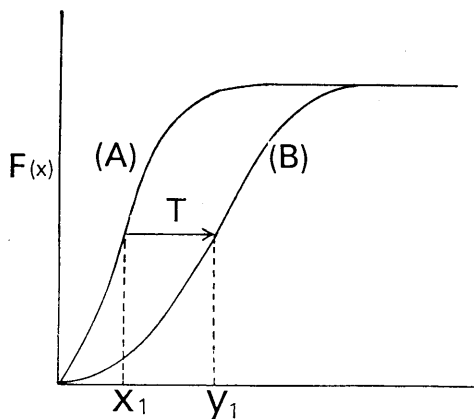


Fig. 1 A method of transformation by diffeomorphism.

このとき(B)のような累積分布関数が正規過程であるような同相写像 T による変換を考える。明らかに(A)の過程は(B)に従って正規過程となる。前節の命題に従って(A)から(B)への変数の変換 $T: X_1 \rightarrow Y_1$ において 1 対 1 で微係数が存在すれば, 微分同相写像となる。(A)の累積分布関数 $F(x)$ とおくと $T: X_1 \rightarrow Y_1$ により

$$\left. \frac{dF(x)}{dx} \right|_{x=x_1} = f(x_1)$$

から(B)の累積分布関数 $G(y)$ について

$$\left. \frac{dG(y)}{dy} \right|_{y=y_1} = g(y_1)$$

への写像を考えれば

$$g(y_1) = f(x_1) \left| \frac{dx_1}{dy_1} \right|$$

が成立する。いま(A)のデータ x_1 は

$$F(x_1) = \int_{-\infty}^{x_1} f(x) dx$$

を満足しているのだから(B)において

$$F(x_1) = \int_{-\infty}^{y_1} g(y) dy = G(y_1)$$

となるようにデータ y_1 を決定すればよい。

ここでもし(A)の過程の分布が指数分布であれば, Fig. 1 で累積分布関数の値 $F(x)$ で $x=0$ のとき連続性の条件を満足しない。したがって偏微係数の存在が保証されず, かつ変数 X_1 と Y_1 の間の 1 対 1 対応関係がない。よって微分同相写像は実現しえない, 指数分布を有するような過程から正規分布を有する過程への変数変換は図 1 の方法では実現できないことになる。

次に正規過程のアフィン同相写像を考える、ひとたび(A)の過程からの(B)過程へと変換された場合および(B)と異なる正規過程へ変換された場合との相違はどうかという問題が生ずる。まず(A)の過程の平均値 μ , 分散 σ^2 とする。この μ, σ^2 により(B)の正規過程 $N(\mu, \sigma^2)$ へ変換する、 $N(\mu, \sigma^2)$ の変数を新たにXとする。つぎにこのXを $ax+b$ とアフィン変換を行う、ここで a, b は常数とする、このときの変換後の分布も正規分布となり、 $N(a\mu+b, a^2\sigma^2)$ となる、したがって $b=-a\mu$, $a=1/\sigma$ とおけば、 $N(\mu, \sigma^2)$ から標準化された正規分布 $N(0, 1)$ へ変換することができる

$$x \rightarrow ax+b$$

の変換によって、まず $N(\mu, \sigma^2)$ から求められた系列相関係数 $\rho_j(X)$ から、変換後の系列相関係数 $\rho_j(aX+b)$ は次のように示される。

$$\rho_j(X) = \frac{\sum_{i=1}^{N-j} (x_i - \mu)(x_{i+j} - \mu)}{(N-j)\sigma^2} \longrightarrow$$

$$\rho_j(aX+b) = \frac{\sum_{i=1}^{N-j} \{ax_i+b - (a\mu+b)\} \{ax_{i+j}+b - (a\mu+b)\}}{(N-j)a^2\sigma^2}$$

$$= \sum_{i=1}^{N-j} (x_i - \mu)(x_{i+j} - \mu) / (N-j)\sigma^2 \quad (34)$$

ここで N はサンプル個数である。(34)式は系列相関係数値において

$$\rho_j(X) = \rho_j(aX+b) \quad (35)$$

を意味している。このことから(35)式の正規過程の表現において、(35)式の関係があれば同一の自己回帰方程式を満足することを示すことができる、以上のことから、図1の分布(A)から(B)への正規過程への変換は $N(0, 1)$ で代表させても、系列相関性の性質は保存され、(35)式の自己回帰方程式は唯一に表現可能である。

6. 計算機シミュレーションによる結果

計算機により非正規過程のデータを作成し、このデータ系列について微分同相写像を行ない、この写像のもつ意味がいかなるものであるかを確かめるため計算機シミュレーション実験を行なうまず本節で行なう計算機シミュレーションの計算方式を述べる。

結合情報量 $Z^{mix}(SX_{-n}, X_{-n+1}, \dots, X_1)$ と $Z(X_{-n}, X_{-n+1}, \dots, X_1)$ の差の情報量 U_x , すなわち

$$U_x = Z^{mix}(SX_{-n}, X_{-n+1}, \dots, X_1) - Z(X_{-n}, X_{-n+1}, \dots, X_1) \quad (36)$$

と図1の(A)から(B)への微分同相写像幅 T の後の上述の結合情報量の差の情報量 U_Y を

$$U_Y = Z^{mix}(SY_{-n}, Y_{-n+1}, \dots, Y_1) - Z(Y_{-n}, Y_{-n+1}, \dots, Y_1) \quad (37)$$

とおく、しかるに(37)式の T 幅写像後の過程が正規過程であれば、 U_Y は(36)式のように系列相関行列で表現するこ

とができる。このとき系列相関係数のみで記述するのでパラメトリック統計量という意味で $U_{Y,P}$ と記述することにする。

$$U_{Y,P} = \log (2\pi e)^{i/2\sigma^2} |S^i_{mix}| - \log (2\pi e)^{i/2\sigma^2} |S^i| \quad (38)$$

ここで

$$|S^i| = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{i-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{i-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{i-1} & \rho_{i-2} & \rho_{i-3} & \dots & 1 \end{bmatrix}, \quad |S^i_{mix}| = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{i-1}^{mix} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{i-2}^{mix} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{i-2} & \rho_{i-3} & \dots & \rho_{i-2}^{mix} & \dots \\ \rho_{i-1} & \rho_{i-2}^{mix} & \dots & \rho_{i-2}^{mix} & 1 \end{bmatrix}$$

であり、(38)式の行列式の行列の要素 ρ_j および ρ_j^{mix} は各々次式で与えられる。

$$\rho_j = \frac{\sum_{i=1}^{N-j} (y_i - \mu)(y_{i+j} - \mu)}{(N-j)\sigma^2}$$

および

$$\rho_j^{mix} = \frac{\sum_{i=1}^{N-j} (Sy_i - \mu)(y_{i+j} - \mu)}{(N-j)\sigma^2}$$

さて $\{X_n, n \in \mathbb{Z}\}$ が非正規過程の場合、(38)式で結果を求めることに不安がある。そこで厳密には、非正規過程の場合はノンパラメトリックな統計量による取り扱いが必要である。したがって、ここでは、ノンパラメトリック統計量 U_X の算出のため状態数 15×15 , 20×20 , および 25×25 の行列を構成し、計算を行った。次に微分同写像(Fig. 1の方法)により正規過程への変換後パラメトリック統計量 $U_{Y,P}$ の計算を行った。

次に微分同相写像の意味を明らかにするために $\{X_n, n \in \mathbb{Z}\}$ が非正規過程の場合でも、 $\{X_n, n \in \mathbb{Z}\}$ に対しパラメトリック統計量 $U_{X,P}$ を適用し、計算を行った。

計算機によりデータを作成する場合、与えられた分布を有するデータの変数が独立となる作成の方法はいくつか発表されている。これらの方法はデータの変数間に従属性を持たせるような生成法ではないので、ここでは一様乱数24個 r_1, r_2, \dots, r_{24} を発生し、これによりデータ X_1 を作成する。次に $r_k, r_{k+1}, \dots, r_{24+k}$ ($k < 24$) の一様乱数からデータ x_2 を作成する。このように一様乱数にオーバーラップを持たせて、従属性のあるデータ系列を作成した、この作成法による分布は正規分布、ガンマ分布、対数正規分布である。作成したデータについて、ノンパラメトリック統計量としての U_X , 微分同相写像後のパラメトリック統計量 $U_{Y,P}$, およびパラメトリック統計量 $U_{X,P}$ の計算を行った。次にこれらの計算値について線形回帰分析を行ない、Fig. 2に示すような結果を得た Fig. 2(a)は横軸がノンパラメトリック統計量 U_X , 縦軸がパラメトリック統計量 $U_{X,P}$ の回帰分析を示している Fig. 2(b)は横軸がノンパラメトリック統計量 U_X , 縦軸

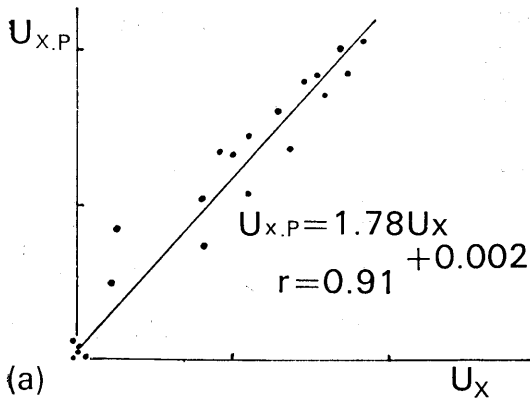


Fig. 2 (a) Relation between amounts of information calculated from nonparametric measure U_X and parametric measure $U_{X,P}$, respectively.

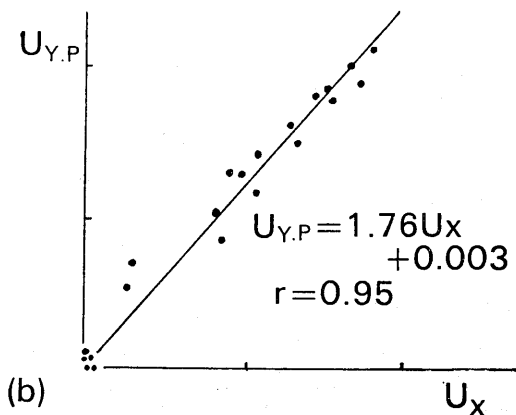


Fig. 2 (b) Relation between amounts of information calculated from nonparametric measure U_X and parametric measure $U_{Y,P}$, respectively.

がパメトリック統計量 $U_{Y,P}$ である。Fig. 2 (b)の相関係数値が0.95であり、Fig. 2 (a)の相関係数値が0.91である、このことは非正規過程のパラメトリック統計量による処理よりも、微分同相写像後のパラメトリック統計量による処理の方が、非正規過程データのノンパラメトリック統計量の処理結果にはば一致して来ることを示すものである。Fig. 2 (b)の $U_{Y,P}$ 縦軸の値は系列相関係数値

から計算していることから、写像後のデータによる系列相関係数値を使って式を展開することがより信頼性のある結果をうる。

7. むすび

正規・マルコフ過程の表現として自己回帰方程式がある。この方程式をもとにシステムの同定・予測およびパワースペクトルの算出が行なわれる。しかし環境データや生体情報では、必ずしも正規過程と見なされないものがある。実際大気汚染のデータは非正規過程により3乗根4乗根あるいは対数をとったデータ変換が行なわれている。本研究は非正規過程のデータの取り扱いとして、かなり一般的な変数変換の方法を提案した。この変換は微分同相写像による、はじめに微分同相写像によるマルコフ過程の次数の不変であることを証明した。次にこの写像による kullback 情報量の不変であることを示した。このことは、写像が十分統計量となることを示し、変換による情報の損失はないことを表わしている。最後に計算機シミュレーションによる微分同相写像の意味づけを明らかにした。

文 献

- 1) N. Ishii and N. Suzumura, Estimation of the Order of Autoregressive Process, International Journal of Systems Science (1976) accepted for publication
- 2) 石井直宏・鈴木宣夫, 自己回帰モデルの次数の推定について, 電子通信学会論文誌(A) (1977) 掲載
- 3) H. Akaike, Information Theory and Extension of the Maximum Likelihood Principle, 2nd International Symposium on Information Theory (1973) Akademia kiado, Budapest
- 4) 得丸英勝, 大気汚染濃度予測のモデリング, 数理科学, 153, (1976)
- 5) E. Parzen, An Estimation of a Probability Density Function and Mode, *Annals of Mathematical Statistics*, 33 (1962)
- 6) S. Kullback and R.A. Leibler, On Information and Sufficiency, *Annals of Mathematical Statistics*, 22 (1951)
- 7) M. Ogawara, a Note on the Test of Serial Correlation Coefficients, *Annals of Mathematical Statistics*, 22 (1951)