

情報量による自己回帰モデルの 評価と非正規過程への適用

正員 石井 直宏[†] 正員 岩田 彰[†]

正員 鈴木 宣夫[†]

Evaluation of the Autoregressive Model by Information Amount and Application to Non-Gaussian Process

Naohiro ISHII[†], Akira IWATA[†] and
Nobuo SUZUMURA[†], *Regular Members*

あらまし 定常正規確率過程を表す自己回帰方程式の研究はシステムの同定、予測であり、ほかに新しいバラスベクトルの算出などにある。自己回帰モデルへの適合の際の問題点はモデルの次数の推定である。従来、次数の推定法として最尤法や最尤法を拡張したFPE法があるが、いずれの方法も適合のための評価関数は線形予測式と実データの差の2乗平均値に係数を掛け合せたものである。本研究では、始めに情報量を自己回帰方程式へ導入することにより予測誤差の2乗の平均値を自己相関行列式の方程式で表すことができた。この方程式は従来の最小2乗法による回帰係数の算出法に代って見通しのよい次数推定の評価式を表している。次に非正規過程のデータを自己回帰モデルで表現できるように、データの前処理の方式について検討した。まずKullback-Leibler情報量を導入し、非正規過程を正規過程へ同相写像してもこの情報量が不変であることを示した。次に計算機シミュレーションにより、ここで述べたデータの前処理の方式を統計量の立場から検討した。

1. まえがき

一般に通信システム・制御システム・生体システムなどの大部分のシステムを記述する変量が確率的とみなされるとき、これらのシステムは確率システムといわれる。これらの確率システムの動作の状態を線形式で表現するものはマルコフ過程の表現の問題として研究されている。又、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る⁽¹⁾。このような線形式で表現される過程の基本形として自己回帰方程式がある。

定常正規確率過程を表す自己回帰方程式の研究はシステムの同定、予測であり、ほかに新しいバラスベクトルの算出法がある。このような場合、自己回帰モデルの次数の推定をいかに行うかが重要な問題となる。次数が決まれば、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、変数が正規過程に従えば、最尤推定法による結果とも一致する。自己回帰モデルの次数の推定はいろいろ提案されている。そのうち最尤推定法⁽²⁾による次数の推定および最尤推定法を拡張したFinal Prediction Error法(FPE法)⁽³⁾による次数の推定法がある。これらの方法の次数推定の評価式は予測誤差の平均2乗値にデータのサンプル個数とモデルの次数からなる係数を掛け合せたものとして表される。ここで予測残差の算出が重要となるが、一般には、最小2乗法による回帰係数と相関係数を計算する必要がある。

本論文では自己回帰方程式の表現に情報量を導入した。その結果、予測残差の2乗平均値が自己相関行列

[†]名古屋工業大学情報工学科, 名古屋市
Department of Information Engineering, Nagoya Institute of
Technology, Nagoya-shi, 446 Japan
論文番号: 昭 53-4[A-4]

の方程式で与えられることを示した。従って最尤推定法およびFPE法による自己回帰モデルの次数の推定の評価式を自己相関行列式の方程式で与えることができた。この評価式によって、ほかの次数推定法⁽⁷⁾との比較が可能となり、見通しのよい表現式となっている。次に自己回帰モデルは変数が正規過程に従うという前提条件の下で最適な表現となるが、環境汚染データ⁽⁴⁾や生物・医学データ⁽⁵⁾は必ずしも正規過程に従わない。そこで実測データの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ということから、非正規過程のデータを正規過程のデータへ変換する方法があるが、これらの方法は必ずしも一般性がない。そこで本論文では非正規過程のデータの取扱いとして、まずKullback-Leibler情報量をとりあげ、非正規過程の累積確率分布関数の正規過程のそれへの同相写像により、この情報量の不変性を証明した。次に統計量の立場から、この情報量を適用した写像の意味を明らかにするため計算機シミュレーションを行った。その結果、非正規過程のデータを正規過程へ変換することが、少ないサンプル個数でより信頼性のある結果を示すことを確かめた。

2. 情報量による自己回帰モデルの評価

自己回帰モデルの次数推定のため最尤法による残差2乗平均値の算出に代って、情報量を導入することによって次数の推定の方式を明らかにする。いま時刻を表す整数の集合を $I = \{\dots, -1, 0, 1, 2, \dots\}$ とし、確率過程 $\{X_t, t \in I\}$ の仮定として

$$E(X_t) = 0$$

とおいても、一般性を失わない。更に自己回帰モデルとして、 m 次の自己回帰過程を考える。これを

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_m X_{t-m} + Z_t \quad (1)$$

とおく。 $\{X_t\}$ が正規過程であれば、式(1)において $\{Z_t\}$ が独立な正規過程となり⁽¹⁾、以下に示すような展開が可能となる。まず

$$E(Z_t) = 0, \quad E(Z_t^2) = \sigma_z^2$$

とおく。 $\{Z_t\}$ の尤度関数は、 $\{Z_t\}$ の正規過程と独立性ということから、 Z_t の実現値を z_t 、 N をサンプル個数とするとき

$$p(z_{m+1}, z_{m+2}, \dots, z_N) = \frac{1}{(\sqrt{2\pi} \sigma_z)^{N-m}} \exp \left\{ -\frac{1}{2\sigma_z^2} \sum_{t=m+1}^N Z_t^2 \right\} \quad (2)$$

となる。式(2)は式(1)の変数変換により条件付確率

$$p(x_{m+1}, x_{m+2}, \dots, x_N | x_1, x_2, \dots, x_m) = \frac{1}{(\sqrt{2\pi} \sigma_z)^{N-m}} \exp \left\{ -\frac{1}{2\sigma_z^2} \sum_{t=m+1}^N (x_t - \alpha_1 x_{t-1} - \alpha_2 x_{t-2} - \dots - \alpha_m x_{t-m})^2 \right\} \quad (3)$$

となる。式(3)の対数尤度関数は

$$l(\alpha_1, \alpha_2, \dots, \alpha_m | x_1, x_2, \dots, x_m) = -(N-m) \log \sqrt{2\pi} - (N-m) \log \sigma_z - \frac{1}{2\sigma_z^2} \sum_{t=m+1}^N (x_t - \alpha_1 x_{t-1} - \dots - \alpha_m x_{t-m})^2 \quad (4)$$

自己回帰係数 $\alpha_1, \alpha_2, \dots, \alpha_m$ を推定するため式(4)の最後の和の項を最小にすることを考える。すなわち $S(\alpha_1, \alpha_2, \dots, \alpha_m | x_1, x_2, \dots, x_m)$

$$= \sum_{t=m+1}^N (x_t - \alpha_1 x_{t-1} - \dots - \alpha_m x_{t-m})^2 \quad (5)$$

を最小にするような $\alpha_1, \alpha_2, \dots, \alpha_m$ を見出すことである。例えば1次の自己回帰過程では、 α_1 の推定値 $\hat{\alpha}_1$ が式(5)を微分して正規方程式を解けば⁽²⁾

$$\hat{\alpha}_1 = \frac{\sum_{t=2}^N (x_{t-1} x_t)}{\sum_{t=2}^N x_t^2}$$

で与えられる。このとき予測残差の2乗和 $S(\hat{\alpha}_1)$ は

$$S(\hat{\alpha}_1) = (N-1) \sigma_x^2 (1 - \rho_1)$$

となる。ここで ρ_1 は1次の自己相関係数、 σ_x^2 は $\{X_t\}$ の分散を表す。又 Z_t の分散 s_z^2 は

$$s_z^2 = \{1 / (N-3)\} S(\hat{\alpha}_1)$$

で与えられる。同様にして m 次の自己回帰方程式の残差の2乗和は

$$S(\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_m) = (N-m) \{C_{xx}(0) - \hat{\alpha}_1 C_{xx}(1) - \dots - \hat{\alpha}_m C_{xx}(m)\} \quad (6)$$

で与えられる⁽²⁾。ここで $C_{xx}(0) [= \sigma_x^2]$ 、 $\{C_{xx}(i), i \geq 1\}$ は各々 $\{X_t\}$ の分散および共分散を表す。このとき Z_t の分散 $s_{z,m}^2$ は式(7)で与えられる。

$$s_{z,m}^2 = \frac{1}{N-2m-1} S(\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_m) \quad (7)$$

最尤法による自己回帰モデルの次数の推定は、整数 m に対して $s_{z,m}^2$ が最小値をとる数 m を次数とすることである。さていまままで述べた最尤法を情報理論的立場から考察する。

変数の組 $\{X_1, X_2, \dots, X_m\}$ を m 変量の正規過程とする。更にこの密度関数は、 R_m 及び $|R_m|$ を各々共分散行列および行列式とすると

$$p(X_1, X_2, \dots, X_m) = \frac{1}{(2\pi)^{m/2}} \frac{1}{|R_m|^{1/2}} \exp \left(-\frac{1}{2} \tilde{X}_{(1)} R_m^{-1} \tilde{X}_{(1)} \right) \quad (8)$$

で与えられる。ここで $\tilde{X}_{(1)} = (X_1, X_2, \dots, X_m)$ のベクトル表示であり、 $\tilde{X}_{(1)}^t$ は $\tilde{X}_{(1)}$ の転置を意味する。式(3)の密度関数と式(8)の定常性の条件より

$$\begin{aligned} & p(X_1, X_2, \dots, X_m, X_{m+1}, \dots, X_s) \\ &= p(X_1, X_2, \dots, X_m) \\ & \cdot p(X_{m+1}, X_{m+2}, \dots, X_s | X_1, X_2, \dots, X_m) \\ &= \frac{1}{(2\pi)^{m/2}} \frac{1}{|R_m|^{1/2}} \exp\left(-\frac{1}{2} \tilde{X}_{(1)} R_m^{-1} \tilde{X}_{(1)}^t\right) \\ & \cdot \frac{1}{(2\pi\sigma_x^2)^{\frac{s-m}{2}}} \exp\left(-\frac{1}{2\sigma_x^2} \sum_{t=m+1}^s Z_t^2\right) \end{aligned} \quad (9)$$

となる。

式(9)の平均情報量 $H(X_1, X_2, \dots, X_m, X_{m+1}, \dots, X_s)$ を計算すると次式のように展開される。

$$\begin{aligned} & H(X_1, X_2, X_3, \dots, X_m, X_{m+1}, \dots, X_s) \\ &= -\int \dots \int p(X_1, X_2, \dots, X_m, X_{m+1}, \dots, X_s) \\ & \cdot \log p(X_1, X_2, \dots, X_m, X_{m+1}, \dots, X_s) dx_1 dx_2 \dots dx_s \\ &= -\int \dots \int p(X_1, X_2, \dots, X_m) \log \frac{1}{(2\pi)^{m/2}} \frac{1}{|R_m|^{1/2}} \\ & \cdot \exp\left(-\frac{1}{2} \tilde{X}_{(1)} R_m^{-1} \tilde{X}_{(1)}^t\right) dx_1 dx_2 \dots dx_m \\ & -\int \dots \int p(X_1, X_2, \dots, X_m, X_{m+1}, \dots, X_s) \log \frac{1}{(2\pi\sigma_x^2)^{\frac{s-m}{2}}} \\ & \cdot \exp\left(-\frac{1}{2\sigma_x^2} \sum_{t=m+1}^s Z_t^2\right) dx_1 dx_2 \dots dx_s \\ &= \log(2\pi e)^{\frac{m}{2}} |R_m|^{\frac{1}{2}} + (s-m) \log(2\pi e)^{\frac{1}{2}} \\ & + (s-m) \log \sigma_x \end{aligned} \quad (10)$$

式(1)の Z_t を 2 乗し期待値をとり、 $\{X_t\}$ と $\{Z_t\}$ が独立であることを使えば式(10)の成立することが証明できる⁽²⁾。

$$\begin{aligned} \sigma_x^2 &= E(Z_t^2) = C_{xx}(0) - \hat{\alpha}_1 C_{xx}(1) - \dots - \hat{\alpha}_m C_{xx}(m) \\ &= (1 - \hat{\alpha}_1 \rho_1 - \dots - \hat{\alpha}_m \rho_m) \sigma_x^2 = V_m \sigma_x^2 \end{aligned} \quad (11)$$

ここで、 V_m は

$$V_m = 1 - \hat{\alpha}_1 \rho_1 - \dots - \hat{\alpha}_m \rho_m \quad (12)$$

とおいた。式(10)は式(12)を使って

$$\begin{aligned} & H(X_1, X_2, \dots, X_m, X_{m+1}, \dots, X_s) = \log(2\pi e)^{\frac{m}{2}} |R_m|^{\frac{1}{2}} \\ & + (s-m) \log(2\pi e)^{\frac{1}{2}} \sigma_x + \frac{s-m}{2} \log V_m \end{aligned} \quad (13)$$

で与えられる。いま $\{X_t\}$ が m 次の自己回帰過程の場合、 $S (\geq m+1)$ 個の変数の平均情報量の相隣り合うものについての関係式は式(13)より $\sigma_x > (2\pi e)^{\frac{1}{2}}$ のとき

$$\begin{aligned} & H(X_1, X_2, \dots, X_s) > \dots > H(X_1, X_2, \dots, X_m, X_{m+1}) \\ & > H(X_1, X_2, \dots, X_m) \end{aligned}$$

で与えられ、相隣り合う不等号間の平均情報量の差は

$$\log \sqrt{2\pi e} \sigma_x + (1/2) \log V_m \quad (14)$$

となる。式(14)は相隣り合う平均情報量の差が一定であることを示す。さて j 次の共分散行列式 $|R_j|$ と j 次の自己相関行列式 $|L_j|$ の間に

$$|R_j| = \sigma_x^{2j} |L_j|$$

の関係が成立する。 $s (\leq m)$ 個の変数の平均情報量は以下のように示される。相関行列 (L_j) は非負エルミート行列であるから固有値 λ_i はすべて非負であり、

$$|L_j| = \lambda_1 \lambda_2 \dots \lambda_j$$

が成立し、且つ $|S_j| \leq 1$ であることが示される。従って

$$|L_1| \geq |L_2| \geq \dots \geq |L_{s-1}| \geq |L_s|$$

よって m 次の自己回帰過程に対し、 $s (\leq m)$ 個の変数のとき

$$\begin{aligned} & H(X_1, X_2, \dots, X_s) \\ &= \log(2\pi e)^{\frac{s}{2}} |L_s|^{\frac{1}{2}} + s \log \sigma_x \end{aligned} \quad (15)$$

と表される。このとき

$$H(X_1, X_2, \dots, X_n) \geq H(X_1, X_2, \dots, X_{n-1}) \geq \dots \geq H(X_1)$$

が成立し、相隣り合う不等号間の平均情報量の差は

$$\log \sqrt{2\pi e} \sigma_x + \frac{1}{2} (\log |L_j| - \log |L_{j-1}|) \quad (16)$$

となる。以上が m 次自己回帰方程式の情報量による特徴づけである。さて最尤推定法による自己回帰モデルの次数の推定は、 Z_t の分散の推定値、すなわち式(7)の最小となる値 m をもって、その次数とするものである。式(6)、(7)及び式(11)より σ_x^2 を改めて $\sigma_{x,m}^2$ とおくと、

$$\sigma_{x,m}^2 = \frac{N-m}{N-2m-1} \sigma_{x,m}^2 \quad (17)$$

となる。式(13)より

$$\begin{aligned} \log \sigma_x^2 &= \left\{ H(X_1, X_2, \dots, X_m, X_{m+1}, \dots, X_s) \right. \\ & \left. - \log(2\pi e)^{\frac{m}{2}} |R_m|^{\frac{1}{2}} - (s-m) \log(2\pi e)^{\frac{1}{2}} \right\} \frac{2}{s-m} \\ &= \frac{1}{s-m} (\log |R_s| - \log |R_m|) \end{aligned} \quad (18)$$

式(7)と式(8)より

$$\begin{aligned} & \log \sigma_{s,m}^2 \\ &= \log \frac{N-m}{N-2m-1} + \frac{1}{s-m} (\log |R_s| - \log |R_m|) \end{aligned}$$

最尤推定法による自己回帰モデルの次数推定のアルゴリズムは

$$\min_{(i)} \log \{ \sigma_{s,i}^2 \} \quad (20)$$

となる整数 i を見出すことである。いいかえれば、このアルゴリズムは式(19)と式(20)より整数 s を固定して

$$\min_{(i)} \left\{ \log \frac{N-i}{N-2i-1} + \frac{1}{s-i} (\log |R_s| - \log |R_i|) \right\} \quad (21)$$

となる整数 i を見出すことである。予測残差の平均 2 乗値の対数の値を精度よく推定するために、サンプル個数 N が十分大でない場合を考慮して、変数の数 s は少ない方がよい。従って式(21)で

$$s = i + 1$$

とおけば

$$\begin{aligned} \min_{(i)} \left\{ \log \frac{N-i}{N-2i-1} + \log |R_{i+1}| - \log |R_i| \right\} \\ = \min_{(i)} \left\{ \log \frac{N-i}{N-2i-1} + \log |L_{i+1}| - \log |L_i| \right. \\ \left. + \log \sigma_i^2 \right\} \quad (22) \end{aligned}$$

となる。式(22)で最後の項は定常であれば定数とみなせることから、最尤法による次数推定のアルゴリズムは

$$\min_{(i)} \left\{ \log \frac{N-i}{N-2i-1} + \log |L_{i+1}| - \log |L_i| \right\} \quad (23)$$

を満たす整数 i を求めることになる。式(23)は自己相関行列式およびサンプル個数 N と次数 i からなる方程式を表す。

(例題) 2 次自己回帰方程式

$$X_t = X_{t-1} - 0.5 X_{t-2} + Z_t \quad (24)$$

を与える。ここで $\{Z_t\}$ は平均 0、分散 1 の独立な正規過程である。正規乱数を使って式(24)によりデータを生成し、サンプル個数 $N = 50$ 及び 250 のときの式(23)により次数を推定する。式(23)のかっこの中の値を U_i とおく。整数 i と U_i の値を図 1 に示す。

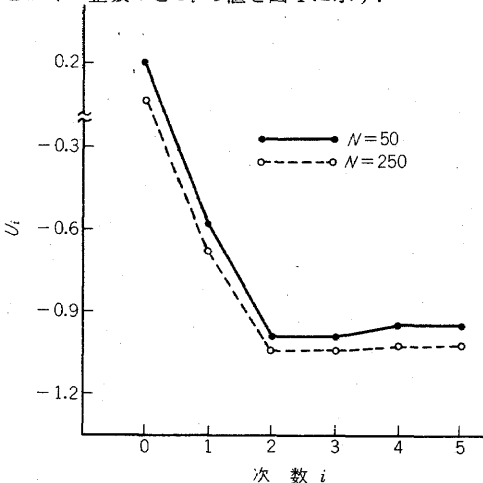


図 1 2 次自己回帰過程の例
Fig. 1-An example of the 2-nd order autoregressive process

図 1 よりこの自己回帰モデルの次数は U_i の値が $i = 2$ からほとんど変わらないことから $i = 2$ 次と推定できる。

FPE法⁽³⁾による自己回帰モデルの次数推定のアルゴリズムは式(23)と同じように残差の平均 2 乗値を自己相関行列式で置換えることができ、

$$\min_{(i)} \left\{ \log \frac{N+i}{N-i} + \log |L_{i+1}| - \log |L_i| \right\} \quad (25)$$

となる整数 i を求めることに帰着する。

3. 同相写像による Kullback 情報量の不変性

大多数のデータ、例えば生体情報や大気汚染のデータは決して正規過程とはいえない場合が多い。そこで通常、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、この種の変換が行われる。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。本論文では非正規過程のデータを累積分布関数の同相写像により、正規過程のデータへ変換するような一般性のある方法を示し、統計量の立場からこの写像の意味づけを行った。累積分布関数の同相写像は座標系の変換を引起し、その結果確率密度関数の変換を生じせしめる。このとき、確率密度関数の変換と情報量の関係は以下のように展開される⁽⁶⁾。

確率空間 (X, S, μ_i) $i = 1, 2$ が定義されている。 X は事象を表す変数、 S は事象の集合、測度 μ_1 と測度 μ_2 が絶対連続であり、 λ がこれらの測度に絶対連続ならラドン-ニコデムの定理⁽⁶⁾により、任意の $E \in S$ に対し

$$\mu_i(E) = \int_E f_i(X) d\lambda(X), \quad i = 1, 2 \quad (26)$$

を満足する確率密度関数 $f_i(X)$ が存在する。いま H_i , $i = 1, 2$ を変数 X を与えたとき $f_i(X)$, $i = 1, 2$ の各々の母集団から抽出される仮説とする。 X を与えたとき H_1 と H_2 を類別するための情報量を

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

と定義する。又、 $I(1:2)$ 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

$$I(1:2) = \int f_1(X) \log \frac{f_1(X)}{f_2(X)} d\lambda(X) \quad (27)$$

と定義される。累積分布関数の同相写像は座標変換を引起す。このことは確率空間の測度の変換とみなされる。いま確率空間 (X, S, μ_i) から (Y, U, V_i) への

変換を ξ とする. 任意の $G \in U$ に対して

$$\nu_i(G) = \int_G g_i(Y) d\tau(Y)$$

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

$$I'(1:2) = \int d\nu_i(Y) \log \frac{g_1(Y)}{g_2(Y)} = \int d\mu_i(X) \log \frac{g_1(\xi(X))}{g_2(\xi(X))} \\ = \int f_1(X) \log \frac{g_1(\xi(X))}{g_2(\xi(X))} d\lambda(X)$$

となる. 次の命題は Kullback によって示されたものである⁶⁾.

平均情報量 $I(1:2)$ と $I'(1:2)$ において

$$I(1:2) \geq I'(1:2) \quad (28)$$

が成立し, 等号は ξ が十分統計量のときのみ成立する.

確率変数へ作用素を導入し, 条件付情報量を構成する. 次に非正規過程から正規過程への同相写像により確率密度関数の変換を引起し, この情報量から

Kullback 情報量を作り出すことを示す. 更にこの変換に対し Kullback 情報量が不変となることを示す.

確率変数 X_n へミキシング作用素を作用させることを

$$MX_n \equiv X_{(r)}, \quad r = N \times \tau + 1 \quad (29)$$

ここで (\cdot) はガウス記号, N はサンプル数, τ は区間が $[0, 1]$ の一様乱数, よって $X_{(r)}$ は $\{\tau\}$ で指定される変数を示す. MX_n とおくことによってもとの時系列の部分系列 $\{X_n, X_{n+1}, X_{n+2}, \dots, X_{n+m}\}$ から新しい部分系列 $\{MX_n, X_{n+1}, X_{n+2}, \dots, X_{n+m}\}$ を構成する. ミキシング作用素を部分系列 $\{X_1, X_2, \dots, X_m\}$ へ作用させると

$$\lim_{N \rightarrow \infty} \frac{1}{N} E \left[E \left[\hat{p}_N(X_2, X_3, \dots, X_m | MX_1) - p(X_2 X_3 \dots X_m) \right] \right] \\ = 0 \quad (30)$$

の成立することが証明できる⁷⁾. ここで $\hat{p}_N(Y)$ は時系列をつくるサンプル個数 N の場合の $p(Y)$ の推定値である. $E(Y)$ はミキシングによる変数 Y の期待値である. 次にミキシング作用素を導入した結合情報量 Z_i^{mix} 及び $\underline{Z}_i^{\text{mix}}$ を次式に定義する.

$$Z_i^{\text{mix}} = - \sum_i \sum_{i-1} \dots \sum_1 p(X_i, X_{i-1}, \dots, X_2, MX_1) \\ \times \log p(X_i, X_{i-1}, \dots, X_2, MX_1) \quad (31)$$

$$\underline{Z}_i^{\text{mix}} = - \sum_i \sum_{i-1} \dots \sum_1 \hat{p}(X_i, X_{i-1}, \dots, X_2, MX_1) \\ \times \log \hat{p}(X_i, X_{i-1}, \dots, X_2, MX_1) \quad (32)$$

ここで

$$\hat{p}(X_i, X_{i-1}, \dots, X_2, MX_1) \\ \equiv E_N \left[E_M \left[\hat{p}(X_i, X_{i-1}, \dots, X_2, MX_1) \right] \right] \quad (33)$$

である. 単純マルコフ過程の必要十分条件は条件付情報量で表現すれば

$$H_0 \geq H_1 > H_2 = H_3 = \dots \quad (34)$$

で表される⁷⁾. ここで

$$H_i = - \sum_{i+1} \sum_i \dots \sum_1 p(X_{i+1}, X_i, \dots, X_1) \\ \times \log p(X_{i+1}, X_i, \dots, X_1) \quad (35)$$

である. 式(32)を使って, 式(34)は漸近的に

$$H_0 \geq \underline{H}_2^{\text{mix}} \approx H_2 = H_3 = \dots \quad (36)$$

のように記述できる. ここで

$$\underline{H}_2^{\text{mix}} = \underline{Z}_3^{\text{mix}} - \underline{Z}_2^{\text{mix}} \quad (37)$$

が成立する. 式(37)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 N に対して

$$H(X_{m+1} | X_m, X_{m-1}, \dots, X_2, X_1) \\ \approx \underline{H}(X_{m+1} | X_m, X_{m-1}, \dots, X_2, MX_1)$$

および

$$H(X_{m+1} | X_m, X_{m-1}, \dots, X_{-i+2}, X_{-i+1}) \\ = \underline{H}(X_{m+1} | X_m, X_{m-1}, \dots, X_{-i+2}, MX_{-i+1}) \\ (i = 0, 1, 2, \dots) \quad (38)$$

で与えられる⁷⁾. ここで $H(X_{m+1} | X_m, X_{m-1}, \dots, X_1)$ は X_1, X_2, \dots, X_m の条件の下での X_{m+1} の条件付情報量である. $\underline{H}(X_{m+1} | X_m, X_{m-1}, \dots, X_2, MX_1)$ は式(37)と同様な意味である. 式(38)の第 2 番目の式を分解すると

$$Z_{n+1} - Z_n = \underline{Z}_{n+1}^{\text{mix}} - \underline{Z}_n^{\text{mix}} \\ \text{書換えて} \\ \underline{Z}_{n+1}^{\text{mix}} - Z_{n+1} = \underline{Z}_n^{\text{mix}} - Z_n \quad (39)$$

で与えられる. ここで新たに

$$U \equiv \underline{Z}_i^{\text{mix}} - Z_i \quad (40)$$

と置く. 以下に $T: X_n \rightarrow Y_n$ の同相写像 T により式(40)の U の値が不変であることを示す. 式(40)は

$$U(X_1, X_2, \dots, X_i) = \int p(X_1, X_2, \dots, X_i) \log p(X_1, X_2, \dots, X_i) \\ \cdot dx_1 dx_2 \dots dx_i - \int \hat{p}(MX_1, X_2, \dots, X_i) \\ \cdot \log \hat{p}(MX_1, X_2, \dots, X_i) \cdot dx_1 dx_2 \dots dx_i \quad (41)$$

となるが, 同相写像は座標変換を引起し, その結果, ヤコビアンを $|\partial(X)/\partial(Y)|$ で表すと

$$dx_1 dx_2 \dots dx_i = \left| \frac{\partial(X_1, X_2, \dots, X_i)}{\partial(Y_1, Y_2, \dots, Y_i)} \right| dY_1 dY_2 \dots dY_i$$

が成立する。更に

$$P(X_1, X_2, \dots, X_i) = \frac{\partial(Y_1, Y_2, \dots, Y_i)}{\partial(X_1, X_2, \dots, X_i)} P'(Y_1, Y_2, \dots, Y_i) \quad (42)$$

が成立する。同相写像の条件より

$$\left| \frac{\partial(X_1, X_2, \dots, X_i)}{\partial(Y_1, Y_2, \dots, Y_i)} \right| = \begin{vmatrix} \frac{\partial X_1}{\partial Y_1} & & 0 \\ & \frac{\partial X_2}{\partial Y_2} & \\ 0 & & \frac{\partial X_i}{\partial Y_i} \end{vmatrix} = \left| \frac{\partial X_1}{\partial Y_1} \right| \left| \frac{\partial X_2}{\partial Y_2} \right| \dots \left| \frac{\partial X_i}{\partial Y_i} \right| \quad (43)$$

となる。ここで $|J_i| \equiv |\partial X_i / \partial Y_i|$ を表す。式(41), (42) 及び式(43)より

$$\begin{aligned} U(X_1, X_2, \dots, X_i) &= \int P'(Y_1, Y_2, \dots, Y_i) \log P'(Y_1, Y_2, \dots, Y_i) \\ &\cdot |J_1| |J_2| \dots |J_i| dY_1 dY_2 \dots dY_i \\ &- \int \widehat{P}(MY_1, Y_2, \dots, Y_i) \log \widehat{P}(MY_1, Y_2, \dots, Y_i) |J_1| \\ &\cdot |J_2| \dots |J_i| dY_1 dY_2 \dots dY_i \\ &= \int P'(Y_1, Y_2, \dots, Y_i) \log P'(Y_1, Y_2, \dots, Y_i) dY_1 dY_2 \dots dY_i \\ &- \int \widehat{P}(MY_1, Y_2, \dots, Y_i) \log \widehat{P}(MY_1, Y_2, \dots, Y_i) dY_1 dY_2 \dots dY_i \\ &\equiv U'(Y_1, Y_2, \dots, Y_i) \end{aligned} \quad (44)$$

となる。又式(44)の情報量 U は式(40)において

$$\left. \begin{aligned} f_1(X_1, X_2, \dots, X_i) &= P(X_1, X_2, \dots, X_i) \\ f_2(X_1, X_2, \dots, X_i) &= \widehat{P}(MX_1) \widehat{P}(X_2, X_3, \dots, X_i) \end{aligned} \right\} \quad (45)$$

とおけば、式(45)の第2式の右辺はサンプル個数が大なら漸近的に

$$P(X_1)P(X_2, X_3, \dots, X_i)$$

に近づく。従って

$$U(X_1, X_2, \dots, X_i) = I(1:2) \quad (46)$$

とおくことができる。すなわち $U(X_1, X_2, \dots, X_i)$ は式(40)で定義した Kullback 情報量となる。式(44)及び式(46)より同相写像 T が式(40)の等号を満足させるような変換 ξ を引起すことになり、この ξ は十分統計量となる。従って同相写像 T によって式(40)の左辺と右辺を変換しても情報の損失はないことを示している。このような同相写像による Kullback 情報量の不変性の性質はマルコフ過程の次数⁽⁷⁾ が写像前と写像後で不変であることを示している。

次に同相写像 T により正規過程のデータを作り出すような方式を述べる。与えられた時系列が非正規過程であり、その累積分布関数が図2(a)のように表されて

いるものとする。同相写像 T により図2(b)のような正規過程の累積分布関数へ変換する。図2(a)から(b)への同相写像 $T: X \rightarrow Y$ が一対一で微係数の存在が確かめられれば微分同相写像を引起す。図2(a)の累積分布関数 $F(x)$ とおくと、 $T: X \rightarrow Y$ により

$$\frac{dF(x)}{dx} \Big|_{x=x_1} = f(x_1)$$

から図2(b)の累積分布関数 $G(y)$ とすると、

$$\frac{dG(y)}{dy} \Big|_{y=y_1} = g(y_1)$$

への変換が引起され

$$g(y_1) = f(x_1) \left| \frac{\partial x}{\partial y} \right|_{\substack{x=x_1 \\ y=y_1}}$$

の関係式が成立する。

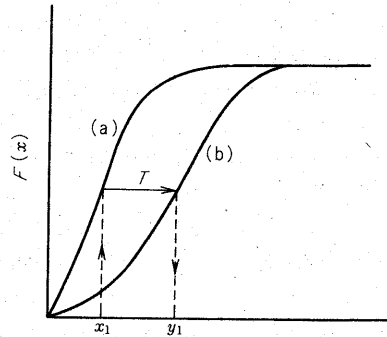


図2 同相写像の例

Fig. 2—An example of a homeomorphism.

いま図2(a)のデータ x_1 は

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

を満足しているので図2(b)において

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

となるようにデータ y_1 を決定すればよい。次に正規過程のアフィン同相写像を考える。まず図2(a)の過程の平均値 μ , 分散 σ^2 とする。この μ, σ^2 により図2(b)の正規過程 $N(\mu, \sigma^2)$ へ変換する。 $N(\mu, \sigma^2)$ のデータを新たに x とする。次にこの x を $ax + b$ とアフィン変換を行う。ここで a, b は定数とする。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。従って $b = -a\mu$, $a = 1/\sigma$ とおけば $N(\mu, \sigma^2)$ から標準化された正規過程 $N(0, 1)$ へ変換することができる。まず $N(\mu, \sigma^2)$ から求めた自己相関関数 $\rho_j(x)$

から、変換後の自己相関関数 $\rho_j(ax+b)$ は次のようになる。

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

及び

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

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

従って

$$\rho_j(x) = \rho_j(ax+b)$$

以上のことから図 2(a)の分布から図 2(b)への正規過程への変換は $N(0,1)$ への変換で代表させても、自己相関関数値は保存される。

4. 同相写像の計算機シミュレーション

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同相写像を行い、この写像のもつ意味を統計量の立場から明らかにするため、計算機シミュレーションを行った。まず 3 で定義した Kullback 情報量、すなわち結合情報量 $Z^{\text{mix}}(M_{-i}, X_{-i+1}, \dots, X_i)$ と $Z(X_{-i}, X_{-i+1}, \dots, X_i)$ の差の情報量 U_X 、

$$U_X \equiv Z^{\text{mix}}(MX_{-i}, X_{-i+1}, \dots, X_i) - Z(X_{-i}, X_{-i+1}, \dots, X_i) \quad (47)$$

と図 2(a)から図 2(b)への同相写像後の Kullback 情報量を

$$U_Y \equiv Z^{\text{mix}}(MY_{-i}, Y_{-i+1}, \dots, Y_i) - Z(Y_{-i}, Y_{-i+1}, \dots, Y_i) \quad (48)$$

とおく。同相写像後の過程が正規過程であれば、 U_Y は以下の式(49)に示すように自己相関行列式で表現することができる。このとき自己相関関数のみで表すので、パラメトリック統計量という意味で U_p と記述する。

$$U_p = \log(2\pi e)^{i/2} \sigma^{2i} |L^i_{\text{mix}}| - \log(2\pi e)^{i/2} \sigma^{2i} |L^i| \quad (49)$$

ここで

$$|L^i| = \begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{i-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{i-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{i-1} & \rho_{i-2} & \rho_{i-3} & \dots & 1 \end{vmatrix}, \quad |L^i_{\text{mix}}| = \begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{i-1}^{\text{mix}} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{i-2}^{\text{mix}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{i-1}^{\text{mix}} & \rho_{i-2}^{\text{mix}} & \rho_{i-3}^{\text{mix}} & \dots & 1 \end{vmatrix} \quad (50)$$

であり、式 50 の行列式の要素 ρ_j 及び ρ_j^{mix} は各々次式で与えられる。

$$\rho_j = \sum_{i=1}^{N-j} (Y_i - \mu)(Y_{i+j} - \mu) / (N-j)\sigma^2$$

及び

$$\rho_j^{\text{mix}} = \sum_{i=1}^{N-j} (MY_i - \mu)(Y_{i+j} - \mu) / (N-j)\sigma^2$$

さて $\{X_t, t \in I\}$ が非正規過程の場合、式(49)を用いて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、ノンパラメトリックな統計量による取扱が必要である。ここでは式(49)のノンパラメトリック統計量 U_X の算出のために状態数 15×15 、 20×20 及び 25×25 (ここで $i \times i$ は行列の行数の数 i と列の数 i を表す) の遷移確率行列を構成し、計算を行った。同相写像の統計量の意味を明らかにするため $\{X_t, t \in I\}$ が非正規過程の場合でも、 $\{X_t, t \in I\}$ に対しパラメトリック統計量 U_{Xp} 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ような方法は幾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、まずデータ x_1 を生成する。次に $r_k, r_{k+1}, \dots, r_{24+k} (k < 24)$ によりデータ x_2 を生成する。このように発生した一様乱数系列にオーバーラップを持たせて、従属性のあるデータ系列 $\{x_i\}$ 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、対数正規分布などである。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 、同相写像後のパラメトリック統計量 U_{Xp} 、及び同相写像前のパラメトリック統計量 U_{Xp} 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い、図 3 に示すような結果を得た。

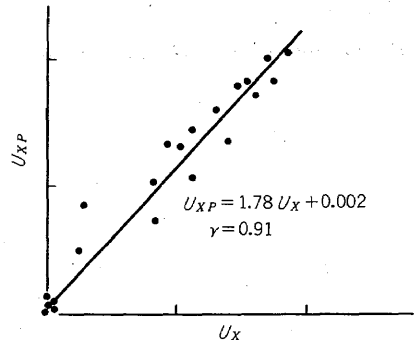


図 3 統計量 U_X と統計量 U_{Xp} による情報量の関係
Fig. 3-Relation between amounts of information calculated from statistics U_X and U_{Xp} .

図3は横軸がノンパラメトリック統計量、縦軸がパラメトリック統計量 U_{Xp} の回帰分析の結果を示している。ここで U_{Xp} は生成したデータに対し、そのままパラメトリック統計量式(49)を適用することを表す。

図4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_{Yp} である。ここで U_{Yp} は生成したデータを同相写像により正規過程に変換し、パラメトリック統計量式(49)を適用したものである。

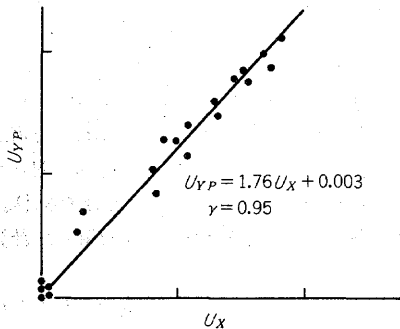


図4 統計量 U_X と統計量 U_{Yp} による情報量の関係
Fig. 4-Relation between amounts of information calculated from statistics U_X and U_{Xp} .

図3の相関係数値が0.91に対し図4の相関係数値が0.95である。このことは非正規過程のデータをそのままパラメトリック統計量を適用した処理よりも、非正規過程データを正規過程へ変換するような同相写像後のパラメトリック統計量を適用した処理の方がノンパラメトリック統計量による処理結果にはほぼ一致してくることを示す。更に図3の U_{Yp} 軸の値が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用いて式(49)を計算することが、より信頼性のある結果を示すことになる。

5. む す び

物理量とか測定値のデータ系列を自己回帰モデルへ

適合させる際の問題点は、モデルの次数の推定である。この問題に対して最尤法あるいは最尤法を拡張したFPE法が適用されている。これらの方法の主眼は予測残差の平均2乗値を最小にすることである。本研究では情報量を自己回帰方程式に導入し、その結果、予測残差の平均2乗値を自己相関行列式の方程式で表現することができた。この表現はほかの次数推定の方法と比較できる見通しのよい方式である。

次に非正規過程表現の基礎として、累積分布関数の同相写像による正規過程化の方式をとり上げ、このような写像による不変な情報量を求め、更に統計量の観点から写像の意味づけを行った。今後の問題として非正規過程データの同相写像後、自己回帰モデルを適合し、その予測値を使って非正規過程のデータ予測し、予測誤差の観点から同相写像の考察を加えることである。

文 献

- (1) M Ogawara: "A note on the test of serial correlation coefficients", *Annals of Math. Statistics*, 22, p. 115(1951).
- (2) G.M. Jenkins and D.G. Watts "Spectral analysis and its application", Holden Day (1969).
- (3) H. Akaike: "Information theory and an extension of the maximum likelihood principle", *Int. Symp. on Information Theory*, Akademia Kiado, Budapest (1973).
- (4) 得丸英勝: "環境汚染予測モデルと環境計画", *自動制御技術*, 22, p. 69(昭51).
- (5) 増山元三郎: "少数例のまとめ方1", 竹内書店新社(昭50).
- (6) S. Kullback and R.A. Leibler: "On information and sufficiency", *Annals of Math. Statistics*, 22, p. 79(1951).
- (7) 石井, 鈴木: "自己回帰モデルの次数の推定について", *信学論(A)*, J60-A, 6, p. 513(昭52-06).
(昭和52年4月22日受付)