

# 最尤法を適用した自己回帰モデルによる同定

石井直宏・岩田 彰・鈴木宣夫

情 報 工 学 科

(1977年9月10日受理)

## Identification by Autoregressive Model Using Maximum Likelihood Method

Naohiro ISHII, Akira IWATA, Nobuo SUZUMURA

*Department of Information Engineering*

(Received September 10, 1977)

In this paper, the autoregressive process is evaluated by the amount of information measure. The residual variance which is applied to the estimation of the order of the autoregressive model, is substituted for the determinant of the correlation matrix of the process. The equation of the determinant of the correlation matrix is an useful one to estimate the order of the autoregressive process. Furthermore it was shown that the maximum likelihood method to estimate the order is related intimately with the method derived from the probabilistic definition of the  $m$ -th order Gaussian Markov process. Next, for the estimation of the parameter of the time variant system, we applied the AR-MA model only whose correlations of AR are dependent on time. To identify the system adaptively, Kalman filter method is employed to the model. Then we developed a short-time likelihood function which evaluates the degree of the identification and of the nonstationarity of the observed output sequence of the system. Finally the short-time likelihood function was applied to the EEG data.

### 1. まえがき

確率的システムの動作を解析するため、しばしばストカスティックなモデルが採用される。これらのモデルの基礎としてガウス・マルコフ過程が中心的な役割を果している。ガウス・マルコフ過程は自己回帰過程で表現されることが証明されている。実際の確率的システムの同定や予測および観測時系列のパワースペクトルの算出に自己回帰過程がよく使用されている<sup>1)</sup>。システムの観測データから自己回帰過程としてのモデル、すなわち自己回帰モデルを適合させる際の問題点はモデルの次数の決定である。従来の研究から最尤法<sup>2)</sup>や最尤法から導かれた偏相関関数法、および最尤法の拡張として Final Prediction Error法(FPE)<sup>3)</sup>がよく適用されているが、これらの諸法について経験的な議論はなされているが統一的な評価は見当たらない。本論文は自己回帰過程の解析で大きな役割を果す最尤法に焦点をあて、情報理論的考察を加えたものである。上記の諸法とは別にガウス・マル

コフ過程<sup>3),4)</sup>の確率論的立場からマルコフ過程の次数が定義される。この確率的定義式より出発し、情報量を導入するとガウス・マルコフ過程の必要十分条件を導くことができる。したがっていずれの方法も情報量を共通の概念として比較検討できるようになり、統一的に評価ができる。モデルの次数が決定されると、次にシステムの構造を反映するパラメータの推定が問題となる。そこで本論文では確率システムの例として生体システムに注目した。生体システムは時間的に変動する系であり、その挙動を解析する場合、定常時系列モデルでは十分でない。このため本論文では時間的に可変な時系列モデルを検討した。時間的に可変なモデルとして、自己回帰係数部分が時間的に変化する自己回帰移動平均型モデル<sup>5)</sup>を採用した。このモデルによるシステムの適応的同定としてカルマン・フィルタ法を示した<sup>10),11)</sup>。システムの同定の評価および時系列の中で局所的非常区間の検出のため、短時間尤度関数を導入し考察を加えた。この尤度関数を

生体の光刺激による脳波反応の時系列へ適用し、非定常区間の検出を行ない、脳波時系列のパワ・スペクトルとの対応関係を明らかにした。

## 2. 最尤法の情報量による評価

本節では自己回帰過程を情報量尺度によって特徴づけを行なう。自己回帰過程の残差の分散が情報量尺度でどのように表現されるかを明らかにする。

いま仮定として  $\{X_t, t \in I\}$  は  $X_t$  を確率変数とする確率過程とする。  $t$  は時刻を表わし、  $I$  は時刻を表わす整数の集合である。一般性を失うことなく

$$E(X_t) = 0$$

と仮定する。このとき  $\{X_t\}$  が  $m$  次の自己回帰過程なら

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + \dots + a_m X_{t-m} + Z_t \quad (1)$$

で表わされる。ここで  $\{a_i\}$  は回帰係数であり、  $\{Z_t\}$  は独立な正規過程で次の条件を満足するものとする。

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

ここで係数  $\{a_i\}$  より構成した

$$\sum_{j=0}^m a_{m-j} y^{m-j} = 0, \quad a_0 = 1$$

の代数方程式の根が単位円の内部に存在するとき、方程式(1)は  $m$  重正規マルコフ過程であるための必要十分条件となっている<sup>3), 4)</sup>。

$\{Z_t\}$  の独立正規過程の条件より、  $Z_t$  の実現値を小文字の  $z_t$  で表示すれば、  $\{z_t\}$  の結合確率密度関数は

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

で与えられる。ここで  $N$  はサンプル個数を表わす。(2)式は(1)式の変数変換により次式のように書き換えられる。

$X_t$  の実現値を小文字の  $x_t$  で表示すれば、結合密度関数は

$$p(x_{m+1}, x_{m+2}, \dots, x_N | x_1, x_2, \dots, x_m) = \frac{1}{(\sqrt{2\pi}\sigma_z)^{N-m}} \exp\left\{-\frac{1}{2\sigma_z^2} \sum_{t=m+1}^N (x_t - a_1 x_{t-1} - a_2 x_{t-2} - \dots - a_m x_{t-m})^2\right\} \quad (3)$$

(3)式の対数尤度関数は

$$L(a_1, a_2, \dots, a_m | x_1, x_2, \dots, x_m) = -(N-m) \log \sqrt{2\pi} - (N-m) \log \sigma_z - \frac{1}{2\sigma_z^2} \sum_{t=m+1}^N (x_t - a_1 x_{t-1} - a_2 x_{t-2} - \dots - a_m x_{t-m})^2 \quad (4)$$

回帰係数  $\{a_i\}$  を求めるには(4)式の最後の項

$$S(a_1, a_2, \dots, a_m | x_1, x_2, \dots, x_m) = \sum_{t=m+1}^N (x_t - a_1 x_{t-1} - \dots - a_m x_{t-m})^2 \quad (5)$$

を最小にする必要がある。(5)式は残差の二乗和と呼ばれ

ている。たとえば一次の自己回帰モデルの係数  $a_1$  は(5)式を  $a_1$  で微分することにより

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

と求められる。ここで  $\hat{a}_1$  は  $a_1$  の推定値である。また残差の二乗和  $S(\hat{a}_1)$  は

$$S(\hat{a}_1) \simeq (N-1) \sigma_z^2 (1 - \rho_1)$$

で与えられる。ここで  $\rho_1$  は時刻 1 遅れの自己相関関数であり、  $\sigma_z^2$  は  $\{X_t\}$  の分散を表わす。このとき  $\{Z_t\}$  の分散の推定値  $S_z^2$  は次式で与えられる。

$$S_z^2 = \left(\frac{1}{N-3}\right) S(\hat{a}_1)$$

一般に  $m$  次自己回帰モデルの残差の二乗和は

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

ここで  $C_{xx}(0)$  と  $C_{xx}(i)$ ,  $i \geq 1$  は  $\{X_t\}$  の分散 ( $\equiv \sigma_x^2$ ) と共分散を各々表わす。  $\{Z_t\}$  の分散の推定値は

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

で与えられる<sup>1)</sup>。ここで  $S_z^2$  を  $S_{z,m}^2$  とおいた。

次に情報論的な考察から自己回帰過程のエントロピー (平均情報量) を求める。まず  $m$  変量の正規確率密度関数を導入する。

$$P(X_1, X_2, \dots, X_m) = \frac{1}{(2\pi)^{m/2}} \cdot \frac{1}{|R_m|^{1/2}} \exp\left(-\frac{1}{2} X_{(1)} R_m^{-1} X_{(1)}\right) \quad (8)$$

ここで  $X_{(1)}$  はベクトル表示で  $X_{(1)} = (X_1, X_2, \dots, X_m)$  である。  $X_{(1)}'$  は  $X_{(1)}$  の転置を表わす。  $|R_m|$  は共分散行列 ( $R_m$ ) の行列式である。さて  $m$  変量の正規過程  $\{X_1, X_2, \dots, X_m\}$  から 1 変量の正規過程  $\{X_t\}$  への写像を次の様に考える。

変量  $X_1$  から  $X_t$ , 変量  $X_2$  から  $X_{t+1}$ へ、以下同様に変量  $X_m$  から  $X_{t+m-1}$  への写像とする。このような  $\{X_1, X_2, \dots, X_m\}$  から  $\{X_t\}$  への写像により  $\{X_t\}$  の確率密度関数は次のように表わされる。

$$P(X_1 X_2 \dots X_m X_{m+1} \dots X_t) \simeq P(X_1 X_2 \dots X_m) \cdot P(X_{m+1} X_{m+2} \dots X_t | X_1 X_2 \dots X_m) \\ = \frac{1}{(2\pi)^{m/2}} \cdot \frac{1}{|R_m|^{1/2}} \exp\left(-\frac{1}{2} X_{(1)} R_m^{-1} X_{(1)}\right) \cdot \frac{1}{(2\pi\sigma_z^2)^{\frac{t-m}{2}}} \exp\left(-\frac{1}{2\sigma_z^2} \sum_{i=m+1}^t Z_i^2\right) \quad (9)$$

ここで  $X_{(1)} = (X_1 X_2 \dots X_m)$  である。さて  $\{X_t\}$  が  $m$  次の自己回帰過程とすると過程  $\{X_1 X_2 \dots X_m X_{m+1} \dots X_t\}$  のエントロピー  $H(X_1 X_2 \dots X_m X_{m+1} \dots X_t)$  は(9)式を使って次のように算出できる。

$$\begin{aligned}
H(X_1 X_2 \cdots X_m X_{m+1} \cdots X_S) &= - \int \cdots \int p(X_1 X_2 \cdots X_m X_{m+1} \cdots X_S) \log p(X_1 X_2 \cdots X_m X_{m+1} \cdots X_S) \\
&\quad dX_1 dX_2 \cdots dX_S = - \int \cdots \int p(X_1 X_2 \cdots X_m) \log \\
&\quad \frac{1}{(2\pi)^{m/2}} \cdot \frac{1}{|R_m|^{1/2}} \exp\left(-\frac{1}{2} X_{(1)} R_m^{-1} X_{(1)}\right) \\
&\quad \cdot dX_1 dX_2 \cdots dX_m - \int \cdots \int p(X_1 X_2 \cdots X_m X_{m+1} \cdots X_S) \\
&\quad \log \frac{1}{(2\pi\sigma_x^2)^{\frac{S-m}{2}}} \exp\left(-\frac{1}{2\sigma_x^2} \sum_{i=m+1}^S Z_i^2\right) \cdot dX_1 dX_2 \cdots dX_S \\
&= \log(2\pi e)^{\frac{1}{2}} |R_m|^{\frac{1}{2}} + (S-m) \log(2\pi e)^{\frac{1}{2}} \\
&\quad + (S-m) \log \sigma_x
\end{aligned} \quad (10)$$

一方(1), (5)および(6)式より

$$\sigma_x^2 = C_{xx}(0) - \hat{a}_1 C_{xx}(1) - \cdots - \hat{a}_m C_{xx}(m) \quad (11)$$

が成立する。(11)式は次式のように分解される。

$$(1 - \hat{a}_1 \rho_1 - \cdots - \hat{a}_m \rho_m) \sigma_x^2 \quad (12)$$

ここで $\sigma_x^2$ は $\{X_i\}$ の分散で $\rho_j$ は時刻遅れ $j$ の自己相関関数である。いま

$$V_m \equiv 1 - \hat{a}_1 \rho_1 - \hat{a}_2 \rho_2 - \cdots - \hat{a}_m \rho_m \quad (13)$$

とおけば(10)式は(13)式を使って

$$\begin{aligned}
H(X_1 X_2 \cdots X_m X_{m+1} \cdots X_S) &= \log(2\pi e)^{\frac{m}{2}} |R_m|^{\frac{1}{2}} \\
&\quad + (S-m) \log(2\pi e)^{\frac{1}{2}} \sigma_x + \frac{(S-m)}{2} \log V_m
\end{aligned} \quad (14)$$

となる。したがって $\{X_i\}$ の $m$ 次自己回帰過程の条件と $\sigma_x^2 > (2\pi e)^{-1}$ のとき

$$\begin{aligned}
H(X_1 X_2 \cdots X_S) &> \cdots > H(X_1 X_2 \cdots X_m X_{m+1}) \\
&> H(X_1 X_2 \cdots X_m)
\end{aligned} \quad (15)$$

が成立する。隣り合わせのエントロピーの差は

$$\log \sqrt{2\pi e} \sigma_x + \frac{1}{2} \log V_m \quad (16)$$

(16)式は $V_m$ および $\sigma_x$ が常数と見なされるので、 $m$ 変数以上のエントロピーの隣り合わせの差の値は一定であることを示している。次に $m$ 変数以下の場合について検討してみる。

共分散行列と相関行列の間には次のような関係式が成立する。

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

ここで $(S_j)$ は相関行列であり、前述の写像によりその行列式は

$$|S_j| \equiv \begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{j-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{j-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{j-1} & \rho_{j-2} & \rho_{j-3} & \cdots & 1 \end{vmatrix}$$

である。 $(S_j)$ は非負のエルミート行列であることから、すべての固有値は非負であり、

$$|S_j| = \lambda_1 \lambda_2 \cdots \lambda_j$$

が成立する。ここで $\lambda_j$ は固有値を表わす、相関行列式の

間には次式のような関係式が成立する。

$$1 \geq |S_1| \geq |S_2| \geq \cdots |S_{S-1}| \geq |S_S|$$

$\{X_i\}$ が $m$ 次の自己回帰過程とするとき過程 $\{X_1 X_2 \cdots X_S\}$  (ただし $S \leq m$ )のエントロピーは

$$H(X_1 X_2 \cdots X_S) = \log(2\pi e)^{\frac{1}{2}} |S_S|^{\frac{1}{2}} + S \log \sigma_x \quad (17)$$

一方情報理論より

$$H(X_1 X_2 \cdots X_m) \geq H(X_1 X_2 \cdots X_{m-1}) \geq \cdots H(X_1)$$

が成立する。隣り合わせのエントロピーの差は

$$\log \sqrt{2\pi e} \sigma_x + \frac{1}{2} (\log |S_j| - \log |S_{j-1}|) \quad (18)$$

が成立する。ただし $j \leq m$ である。

最尤推定法による自己回帰モデルの次数の推定は $\{Z_i\}$ の分散の推定値(7)式の最小値をとる整数 $m$ でもって、その次数とするものである。

いま $\sigma_x^2$ と $S_S^2$ を $m$ 次の自己回帰過程の場合ということ

から各々 $\sigma_{x,m}^2$ と $S_{S,m}^2$ に書き換える、このとき推定値

$S_{S,m}^2$ は

$$S_{S,m}^2 = \frac{N-m}{N-2m-1} \sigma_{x,m}^2 \quad (19)$$

で与えられる。(14)式から、次の式が導かれる。

$$\begin{aligned}
\log \sigma_{x,m}^2 &= \{H(X_1 X_2 \cdots X_m X_{m+1} \cdots X_S) \\
&\quad - \log(2\pi e)^{\frac{1}{2}} |R_m|^{\frac{1}{2}} - (S-m) (2\pi e)^{\frac{1}{2}}\} \frac{2}{S-m} \\
&= \frac{1}{S-m} (\log |R_S| - \log |R_m|)
\end{aligned} \quad (20)$$

(19)式と(20)式から

$$\begin{aligned}
\log S_{S,m}^2 &= \log \frac{N-m}{N-2m-1} + \frac{1}{S-m} (\log |R_S| \\
&\quad - \log |R_m|)
\end{aligned} \quad (21)$$

となる。

最尤推定法による自己回帰モデルの次数の推定のアルゴリズムは

$$\text{Min}_{(i)} \log \{S_{S,i}^2\} \quad (22)$$

を満足する整数でもってその次数とすることと等しい。

(2)式と(22)式より、このアルゴリズムは書き換えられて次のようになる。

$$\text{Min}_{(i)} \left\{ \log \frac{N-i}{N-2i-1} + \frac{1}{S-i} (\log |R_S| - \log |R_i|) \right\} \quad (23)$$

少ないサンプル個数で精度よく次数を推定するためには、行列式 $|R_S|$ の次元 $S$ を小さくおくことが望ましい。そこで

(23)式の $|R_S|$ の $S$ を

$$S = i+1$$

と設定すると、(23)式は(24)式のようにになる。

$$\begin{aligned} & \min_{(i)} \left\{ \log \frac{N-i}{N-2i-1} + \log |R_{i+1}| - \log |R_i| \right\} \\ &= \min_{(i)} \left\{ \log \frac{N-i}{N-2i-1} + \log |S_{i+1}| - \log |S_i| \right. \\ & \quad \left. + \log \sigma_x^2 \right\} \quad (24) \end{aligned}$$

24式から、最大推定法による次数推定のアルゴリズムは、25式のようになる。

$$\min_{(i)} \left\{ \log \frac{N-i}{N-2i-1} + \log |S_{i+1}| - \log |S_i| \right\} \quad (25)$$

同様にFinal Prediction Error法(FPE)<sup>2)</sup>における残差の二乗和を自己相関行列式でおきかえることができる。この場合の次数推定のアルゴリズムは次式のようになる。

$$\min_{(i)} \left\{ \log \frac{N-i}{N-1} + \log |S_{i+1}| - \log |S_i| \right\}$$

最尤法による自己回帰モデルの次数の推定の例題を計算機シミュレーションにより作りそのテストを行なった。シミュレーションとして2次の自己回帰過程

$$X_t = X_{t-1} - 0.5X_{t-1} + Z_t \quad (26)$$

を生成した。ここで $\{Z_t\}$ は平均0で分散1の独立な正規過程である。生成したサンプル個数として $N=50$ と250の場合である。25式を

$$U_i = \log \frac{N-i}{N-2i-1} + \log |S_{i+1}| - \log |S_i|$$

のような形で用い、 $i=1, 2, \dots, 5$ に対する $U_i$ の値をプロットしたものをFig. 1に示す。Fig. 1のたて軸が $U_i$ の値を示す。横軸が次数  $i=1, 2, 3, \dots, 5$  である。

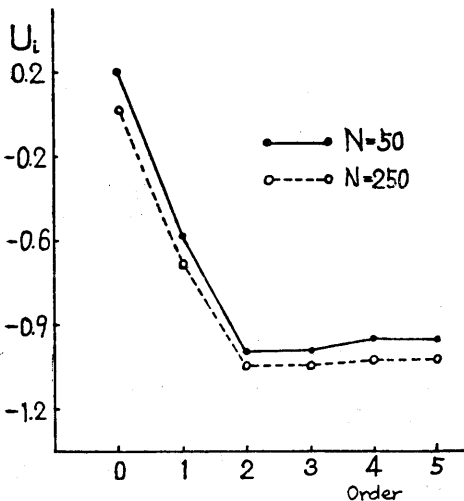


Fig. 1 Estimation of the order of the autoregressive model.

### 3. 最尤法とマルコフ過程の次数

ここではマルコフ過程の確率的定義から導出されたがウス・マルコフ過程の次数と最尤法<sup>1)</sup>および偏相関関

数法<sup>2)</sup>によって導出された次数の比較を行なう。

$\{X_i\}$ が $m$ 重マルコフ過程は条件付確率により次のように定義される。

$$\begin{aligned} & P(X_{m+1} \leq \lambda | X_1, X_2, \dots, X_{m-1}, X_m) \\ &= P(X_{m+1} \leq \lambda | X_2, \dots, X_{m-1}, X_m) \end{aligned}$$

でかつ

$$\begin{aligned} & P(X_{m+1} \leq \lambda | X_{-i}, X_{-i+1}, \dots, X_{m-1}, X_m) \\ &= P(X_{m+1} \leq \lambda | X_{-i+1}, X_{-i+2}, \dots, X_{m-1}, X_m) \\ & \quad (i=0, 1, 2, 3, \dots) \quad (27) \end{aligned}$$

が成立することである。ここで $P(a|b)$ は、事象 $b$ の生じた後に事象 $a$ の生起する確率を示す。 $\lambda$ は実数を表わす。条件付情報量と27式を使って $m$ 重マルコフ過程の必要十分条件は次の式で表わされる<sup>5)</sup>。

$$H_0 \geq H_1 \geq H_2 \geq \dots \geq H_{m-1} > H_m = H_{m+1} = \dots \quad (28)$$

ここで

$$H_0 = -\sum_i P(X_i) \log P(X_i),$$

$$H_i = -\sum_{i+1} \sum_1 \dots \sum_1 P(X_i X_2 \dots X_i X_{i+1}) \log$$

$$P(X_{i+1} | X_1 X_2 \dots X_{i-1} X_i) \quad (i=1, 2, 3, \dots)$$

の条件付エントロピーを表わす。

$\{X_i\}$ がガウス・マルコフ過程の場合、28式から $m$ 次自己回帰過程であるための必要十分条件が次の式で与えられる<sup>5)</sup>。

$m$ より小さいすべての正整数 $n$ に対して

$$|S_n| \geq |S_{n-1}|^{\frac{1}{2}} \cdot |S_{n+1}|^{\frac{1}{2}}$$

$m$ と等しい整数 $n$ に対して

$$|S_n| > |S_{n-1}|^{\frac{1}{2}} |S_{n+1}|^{\frac{1}{2}}$$

$m$ より大きいすべての正整数 $n$ に対して

$$|S_n| = |S_{n-1}|^{\frac{1}{2}} |S_{n+1}|^{\frac{1}{2}} \quad (29)$$

が成立することである。

一方相関行列 $(S_{n+1})$ と偏相関関数 $(\pi_i)$ の間に次式のような関係が成立する<sup>7)</sup>。

$$\begin{aligned} |S_{n+1}| &= (1-\pi_1^2)^n (1-\pi_2^2)^{n-1} \dots (1-\pi_n^2) \\ & \quad (n=1, 2, 3, \dots) \quad (30) \end{aligned}$$

したがって29式は30式によって書き換えられ次のように表現される。すなわち $m$ 次自己回帰過程であるための必要十分条件は次の式で与えられる。

$m$ より小さいすべての正整数 $n$ に対して

$$\pi_n^2 \geq 0$$

$m$ と等しい正と数 $n$ に対して

$$\pi_n^2 > 0$$

$m$ より大きいすべての正整数 $n$ に対して

$$\pi_n^2 = 0 \quad (31)$$

が成立することである。

ここで偏相関関数 $\pi_k$ は共分散関数と次式の関係にある<sup>1)</sup>。

$$\pi_k = \frac{\|R_{k+1}\|}{\|R_k\|}$$

上式で  $|R_k|$  は  $\{X_i\}$  の共分散行列の行列式を表わす。  
 $\|R_{k+1}\|$  は共分散行列  $(R_{k+1})$  の第1列と第  $(K+1)$  行目の要素の余因子行列の行列式である。

③式は  $m$  次自己回帰モデルの次数を推定するために最尤法から展開された偏相関関数法の結果と一致するものである。このことは偏相関関数法による自己回帰モデルの次数の推定の結果がガウス・マルコフ過程の確率的定義による次数の結果と一致することを示している。

さて④式の両辺の対数をとれば、その基本型として

$$\log|S_i| - \log|S_{i-1}| \geq \log|S_{i+1}| - \log|S_i| \quad (32)$$

で示することができる。一方、サンプル個数  $N$  が次数  $m$  に比べて十分大であれば、最尤法による自己回帰モデルの次数推定のアルゴリズムは④式より次式のように表わされる。

$$\min_{(i)} \{ \log|S_{i+1}| - \log|S_i| \} \quad (33)$$

③式はガウス・マルコフ過程のマルコフ次数の確率的定義式を出発点として導出されたものであり、③式はガウス・マルコフ過程の最尤法から出発したものである。③式と③式を比較すると同じ式の型であり、いずれの方法でもほぼ等しい次数の結果となることが考えられる。ここで注意しておきたいことは、③式と③式の比較から、わかるように、通常よく使われる最尤法やFPE法により推定された次数は確率的な定義にもとずいた次数より少々大きな値をとることが予想される。モデルの次数が決定されると次にシステムの構造を反映するパラメータの推定が問題となる。次節では時間的に変動するシステムを取り上げる。

#### 4. 時間的に可変な自己回帰モデル

時間に関して可変な時系列モデルとして自己回帰移動平均 (AR—MA) 型のモデルを拡張した次のようなモデルが提案されている<sup>9)</sup>。

$$\begin{aligned} Y(t) + a_1(t)Y(t-1) + \dots + a_n(t)Y(t-n) \\ = \lambda\{e_0(t) + C_1e_0(t-1) + \dots + C_me_0(t-m)\} \\ + D'(t) \end{aligned} \quad (34)$$

ここで  $a_i(t)$  ( $i=1, 2, \dots, n$ ) は自己回帰 (AR部) の係数であり、時刻  $t$  に依存する。 $C_j$  ( $j=1, 2, \dots, m$ ) は移動平均 (MA部) の係数である。 $\lambda$  は雑音  $\{e_0(t)\}$  の振幅を表すパラメータである。 $D'(t)$  は低周波の外乱を表わす。 $\{Y(t)\}$  はこのモデルの出力系列である。出力系列  $\{Y(t)\}$  に  $Z$  変換を適用し

$$Y(t-k) = Z^{-k}Y(t)$$

および

$$A(Z^{-1}|t) = 1 + a_1(t)Z^{-1} + \dots + a_n(t)Z^{-n} \quad (35)$$

$$C(Z^{-1}|t) = 1 + C_1Z^{-1} + \dots + C_mZ^{-m} \quad (36)$$

とおくと③式は③式のように表わされる。Fig. 2 に、この方程式のブロック表示する。

$$A(Z^{-1}|t)Y(t) = \lambda C(Z^{-1})e_0(t) + D'(t) \quad (37)$$

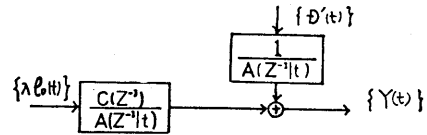


Fig. 2 Block diagram of the time variant model

いま補助系列  $\{X(t)\}$  を導入し③式を次のように書き換えてMA部をprefilterとして分離する。

$$Y(t) = C(Z^{-1})X(t)$$

$$X(t) = A^{-1}(Z^{-1}|t)\{\lambda e_0(t) + D(t)\}$$

$$D(t) = C^{-1}(Z^{-1})D'(t)$$

以上の変換により、モデルの動的な主要部は、Fig. 3 の破線内で表現できるAR型の時系列モデルとなる。

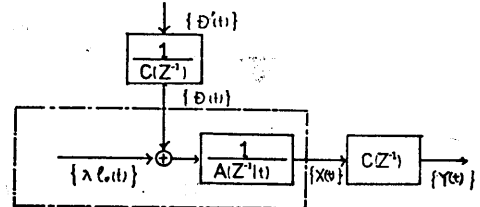


Fig. 3 Block diagram of the time variant model of AR-type

ここで時間に関して可変な係数  $a_i(t)$  および  $D(t)$  が簡単な一次の差分方程式に従うマルコフ過程であると仮定して、次のモデルを脳波の解析に用いる。このモデルはBohlinのモデル<sup>9)</sup>と等しい。

$$\begin{cases} Y(t) = X(t) + C_1X(t-1) + \dots + C_mX(t-m) \\ X(t) + a_1(t)X(t-1) + \dots + a_n(t)X(t-n) \\ = \lambda e_0(t) + D(t) \\ a_i(t) = a_i(t-1) + q_i e_i(t-1), \quad i=1, 2, \dots, n \\ D(t) = D(t-1) + q_{n+1} e_{n+1}(t) \end{cases} \quad (38)$$

ここで  $\{e_i(t)\}$  ( $i=1, 2, \dots, n$ ) はそれぞれ  $N(0, 1)$  (平均0, 分散1の正規過程) に従う独立で無相関な正規白色雑音である。 $q_i$  は係数  $a_i(t)$  の変化の速さを決定するパラメータであり、とくに

$$q_1 = q_2 = q_n = q_{n+1} = 0$$

の場合は定常な時系列モデルとなる。

次に確率変数のベクトルを次のように定義する。

$$\theta(t) = \begin{pmatrix} a_1(t) \\ a_2(t) \\ \vdots \\ a_n(t) \\ D(t) \end{pmatrix}, \quad v(t) = \begin{pmatrix} e_1(t) \\ e_2(t) \\ \vdots \\ e_n(t) \\ e_{n+1}(t) \end{pmatrix}$$

$$Q = \begin{pmatrix} q_1^2 & & & \\ & q_2^2 & & 0 \\ & & \ddots & \\ 0 & & & q_{n+1}^2 \end{pmatrix}$$

および

$$\begin{aligned} \mathbf{x}(t) &= [-X(t-1), -X(t-2), \dots, -X(t-n), 1] \\ \mathbf{e}(t) &= \lambda \mathbf{e}_0(t) \end{aligned}$$

上述の定義によって(38)式は次のようなシステムを記述する状態方程式を得る。

$$X(t) = Y(t) - C_1 X(t-1) - \dots - C_m X(t-m) \quad (39)$$

$$\begin{cases} \theta(t+1) = \theta(t) + v(t) \\ X(t) = \mathbf{x}(t) \theta(t) + \mathbf{e}(t) \end{cases} \quad (40)$$

ここで  $\{v(t)\}$  と  $\{\mathbf{e}(t)\}$  は相互に独立な白色正規過程で、それぞれ  $N(0, Q)$  と  $N(0, \lambda^2)$  の分布に従う。時刻  $t=1, 2, \dots, N$  に対して観測値系列  $\{Y(t)\}$  が(39), (40)式のモデルを満たすとすれば、 $\theta(t)$  の最小二乗平均の意味で最良の推定値  $\hat{\theta}(t)$  (この場合、雑音の正規過程の仮定から最尤推定値ともなる) はカルマン・フィルタによる推定の問題に帰着され、次の方程式系によって各時刻  $t$  ごとに  $\theta(t)$  が算出される。

$$\begin{cases} \gamma^2(t) = \lambda^2 + \mathbf{x}(t) \sigma(t) \mathbf{x}^t(t) \\ S(t) = \gamma^{-1}(t) \sigma(t) \mathbf{x}^t(t) \\ \xi(t) = \gamma^{-1}(t) \{X(t) - \mathbf{x}(t) \hat{\theta}(t)\} \\ \hat{\theta}(t+1) = \hat{\theta}(t) + \xi(t) S(t) \\ \sigma(t+1) = \sigma(t) + Q - S(t) S^t(t) \end{cases} \quad (41)$$

ここで  $\hat{X}(t) = \mathbf{x}(t) \theta(t)$  は  $X(t)$  の推定値であり

$$\tilde{X}(t) = \hat{X}(t) - X(t)$$

は推定誤差を表わす。 $\gamma^2(t)$  は推定された  $\hat{X}(t)$  の分散

$\text{cov}[\tilde{X}(t), \tilde{X}(t)]$  である。 $\tilde{X}(t)$  を  $\gamma(t)$  で正規化した残差  $\xi(t)$  はフィルタの理想状態において、 $N(0, 1)$  に従う正規過程となる。また  $S(t)$  はカルマン・フィルタの利得であり、 $\xi(t)$  との積により各サンプル毎に推定値  $\hat{\theta}$  を更新する。 $\sigma(t)$  は状態ベクトル  $\hat{\theta}(t)$  の推定誤差の共分散であり

$$\sigma(t) = E[(\theta(t) - \hat{\theta}(t))(\theta(t) - \hat{\theta}(t))^t]$$

である。

## 5. 短時間尤度関数の導入と非定常区間の推定

モデルを表現する(38)式において、 $\{q_i\}$  はこのモデルの本質を決定する重要なパラメータであり、係数  $\{a_i(t)\}$  の変化の速さを表わす定数である。 $\{q_i\}$  の設定値によってフィルタの同定する状態が大きく変わる為、対象となる信号の非定常性の度合い<sup>12)</sup>によって適切な値を設定しなければならない。

長いサンプル系列  $\{Y(t) = y(t) | t=1, 2, \dots, N\}$  を与えて

フィルタリング、すなわち(40)式を実行した結果得られる残差系列  $\{\xi(t) | t=1, 2, \dots, N\}$  あるいはイノベーション系列  $\{\tilde{X}(t) | t=1, 2, \dots, N\}$  に対して最尤法を適用し、より適切な  $q_i$  の値を設定定することが行なわれている。

$N$  まいノベーション系列  $\{\tilde{X}(t) | t=1, 2, \dots, N\}$  の結合密度関数は  $\{\tilde{X}(t)\}$  の独立な正規過程の条件により

$$p_N = \prod_{t=1}^N \frac{1}{\sqrt{2\pi} \gamma(t)} \exp \left[ -\frac{1}{2} (\xi(t) \gamma(t) - 0) \gamma(t)^{-2} (\xi(t) \gamma(t) - 0) \right] \quad (42)$$

となるから尤度関数  $L(\lambda, Q | Y_N)$  の対数をとると<sup>9)</sup>,

$$\log L(\lambda, Q | Y_N) = C - \frac{1}{2} \sum_{t=1}^N \log \gamma^2(t) - \frac{1}{2} \sum_{t=1}^N \xi^2(t) \quad (43)$$

ここで  $Y_N = \{y(t) | t=1, 2, \dots, N\}$  で  $C$  は常数、 $\gamma^2(t)$  および  $\xi(t)$  は差分方程式(40)式によって計算される。パラメータ  $\lambda, Q$  の最尤推定値は(43)式の最大値を与える値となる。これによって得られるパラメータ  $q_i$  の値は、全サンプル ( $t=1, 2, \dots, N$ ) を通しての平均的な値である。ここで入力信号の一部の周波数成分が大きく変化する非定常部分に興味の中心があるときのパラメータの設定をいかに行なうかという問題がある。さらに設定されたパラメータの値に対して時間の経過とともにカルマン・フィルタによる同定の状態がどのように変化するかを評価するために、次のような尤度関数を考慮した指標を用いる。

$$\begin{aligned} KT(t) &= \frac{1}{M} \sum_{s=t-M+1}^t \log \gamma(S) \\ VT(t) &= \frac{1}{M} \sum_{s=t-M+1}^t \xi^2(t) \end{aligned} \quad (44)$$

ここで  $M$  は興味ある変化をとらえるのに適当な時間幅で本論文ではこれを短時間と呼んだ。このときの短時間尤度関数を

$$LLT(t | \lambda, Q) = -\frac{1}{2} [2KT(t) + \log VT(t)] \quad (45)$$

と定義する。時刻  $t \geq M$  に対して  $LLT(t)$  は各サンプルごとに、直前の  $M$  個の系列

$$\{\xi(t-M+1), \xi(t-M+2), \dots, \xi(t)\}$$

によって計算される。ここで(40)式の  $\{v(t)\}$ ,  $\{\mathbf{e}(t)\}$  は正規性の雑音系列と仮定されていて、 $LLT(t)$  は系列  $\{\xi(t-M+1), \xi(t-M+2), \dots, \xi(t)\}$  に対する尤度関数になっているが、 $\{v(t)\}$ ,  $\{\mathbf{e}(t)\}$  の正規性が成立しないときでも、線形推定法則にもとづいて、推定誤差ベクトル  $\hat{\theta}(t) = \theta(t) - \hat{\theta}(t)$  の分散行列  $\sigma(t)$  を2次形式の意味で最小にするようなカルマン・フィルタが(40)式と同じになることが知られている。したがって残差系列  $\{\xi(t)\}$  あるいはイノベーション系列  $\{\tilde{X}(t)\}$  に従う分布が何んでも(40)式の  $VT(t)$  は残差の平均二乗値を最小にするという意味でパラメータ  $\{q_i\}$  の設定およびフィルタの同定

状態の時間的変化の評価に対する指標となる。

次に時系列の非定常区間の推定のために(45)式で定義した短時間尤度関数を適用する。時系列モデルを観測時系列へより正確に近づけることはイノベーション系列の尤度関数を最大にすることを意味している。時系列の定常区間については(45)式のモデルのパラメータ  $q_i^2$  が小さい値でよく、したがって  $LLT(t)$  の値が大となる。すなわち尤度関数の値が高い値を示す。このとき  $q_i^2$  の値を大きな値で設定しておく(45)式の係数  $a_i(t)$  の変動が大きくなり  $\xi^2(t)$  の値も大きくなる。したがって  $LLT(t)$  の値が小さくなる。非定常区間については、もし  $q_i^2$  の値を小さく設定すると  $a_i(t)$  の大きな変動が要求されたにもかかわらず追従性が悪くなり  $\xi^2(t)$  の値も大きくなる。したがって  $LLT(t)$  の値が小さくなる。このとき  $q_i^2$  の値を大きく設定すると  $\xi^2(t)$  は小さい値をとり、したがって  $LLT(t)$  の値は大きくなる。以上のことから、 $q_i^2$  の値を2つ設定しておいて観測時系列からフィルタを構成し、定常区間および非定常区間の分類が可能となる。

上述の短時間尤度関数を光刺激による脳波反応の時系列へ適用する。安静閉眼状態で眼から30cmのところにストロボの光源をおき6Hzの光刺激を与える。Fig. 4にはこのときの脳波時系列を示している。約16秒までは安静閉眼時の $\alpha$ 波(10.5Hz)が見られるが6Hzの光刺激後に脳波の振幅が減少し、光駆動による脳波反応が明確に示されている。

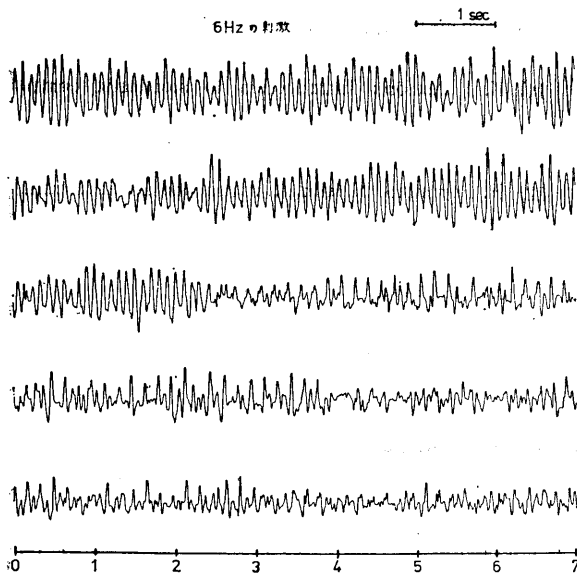


Fig. 4 EEG driven by photic stimulation

次に(45)式のパラメータ  $q_i^2=2^{-15}$  と  $q_i^2=2^{-20}$  を設定し、(45)式の短時間尤度関数  $LLT(t)$  の値を計算する。この結

果を Fig. 5 に示す。Fig. 5 で実線がパラメータ  $q_i^2=2^{-20}$  の場合であり、点線が  $q_i^2=2^{-16}$  の場合の結果である。もし Fig. 4 の脳波時系列が定常であれば  $q_i^2=2^{-20}$  とおいたときの  $LLT(t)$  の値が  $q_i^2=2^{-16}$  とおいたときの  $LLT(t)$  の値よりもつねに大きくなるはずである。しかるに Fig. 5 の①、②および③区間はパラメータ  $q_i^2=2^{-16}$  の  $LLT(t)$  の値が  $q_i^2=2^{-20}$  の  $LLT(t)$  の値よりも大きな値を示している。これらの区間では他の区間に比べ時系列の変化が早いことを示すものである。

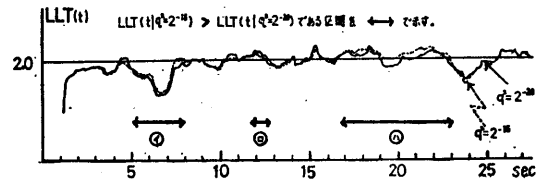


Fig. 5 Short-time likelihood function of EEG

さて(46)式の AR-MA 型モデルによる観測時系列のパワースペクトルは

$$\varphi(f|t) = 2^2 h \left| \frac{C(Z^{-1})}{A(Z^{-1}|t)} \right|^2 \quad (46)$$

で与えられる。(46)式で  $h$  はサンプリング周期で 10msec である。Fig. 6 に Fig. 3 の脳波時系列のパワースペクトルを示す。たて軸の  $y^2$  は観測時系列の分散である。変数となる。AR部の次数は  $n=14$ 、 $q_i^2=2^{-16}$  としてパワースペクトルを計算したものである。

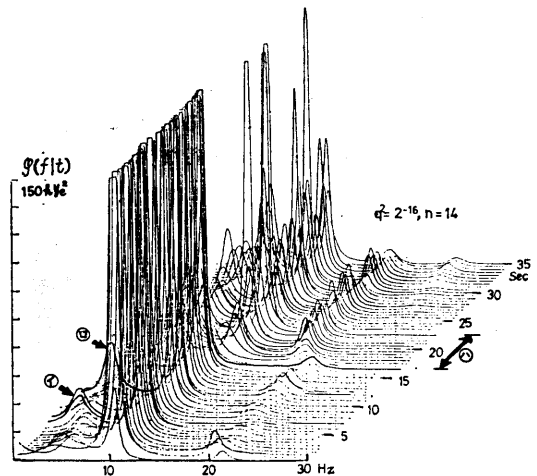


Fig. 6 Power spectrum of EEG

Fig. 6 では時刻  $t=16.5$  秒では  $\alpha$  波の鋭いピークが 10Hz 付近にあり、その高調波が 20Hz 付近に見られる。6Hz の光刺激が加わると刺激周波数および高調波 6, 12, 18, 24 Hz の成分と  $\alpha$  波の成分が不安定な現われ方をするようになる。刺激の加えられ始めの④の部分は特に強い非定常

性を示す。①および②の非定常性の部分は支配的な $\alpha$ 波の中で $t=5$ 秒付近より始まる低周波成分の変化と対応している。なお短時間尤度関数の値が小さい場合も非定常性を示す。Fig. 5では④の部分と23, 24, 25秒の部分である。

## 6. むすび

システムの同定および予測などでガウス・マルコフ過程は重要な役割を果たしている。この過程は自己回帰過程で表現される。したがって実際のシステムの記述はどは自己回帰過程による解析となる。この解析で中心的な役割を果たして来ているのが最尤法を主とするものである。本論文では自己回帰モデルの次数推定の方法として適用される最尤法、およびこの方法から導かれた偏相関関数法および拡張した最尤法としてよく使われているFPE法を統一的に評価するため、最尤法に情報理論的考察を加えた。さらに確率的立場から定義したマルコフ過程の次数推定の情報理論的考察から、これらの諸法が情報量を共通の概念として統一的に評価ができた。次にシステムの次数が推定されると、システムのパラメータをどのように把握するかという問題が生ずる。

たとえば生体系などのシステム自体が時間的に変動する現象には時間的に可変な時系列モデルが望まれる。本論文では時間的に可変な自己回帰部分を有する時系列モデルをとり上げ、そのシステムへの適応的同定としてカルマンフィルタ法を示した。ここで短時間尤度関数を導入し、観測時系列における追従性の評価および非定常区間の検出について考察を加えた。最後にこの検出法を脳波時系列へ適用した。今後の問題として任意の非定常な時系列へこれらの検出法を適用する場合、パラメータ $q^2$ の設定をどのような手続きを経て行なうかについて検討を必要とする。

## 文 献

- 1) G.M. Jenkins and D.G. Watts: Spectral analysis and its Application, Holden Day (1969)
- 2) H. Akaike: "Information Theory and Extension of the Maximum Likelihood Principle," 2nd International Symposium on Information Theory, Akademia Kiado, Budapest (1973)
- 3) J.L. Doob: Stochastic Processes, John Wiley (1953)
- 4) M. Ogawara: "Time series analysis and stochastic prediction" Bull. Mafh. Stat., vol 8 (1958)
- 5) N. Ishii and N. Suzumura: "Estimation of the order of autoregressive process," Int. Journal of Systems Science vol. 8. No. 8 (1977)
- 6) N. Ishii, N. Suzumura and S. Nishino: "Invariants of Markov process by the transformation of variables," Int. Journal of Systems Science vol. 8. No. 12 (1977)
- 7) O. Barndorff-Nielsen and G. Shou: "On the parametrization of autoregressive models by partial autocorrelations" Journal of Multivariate Analysis, vol. 3 (1973)
- 8) 石井, 岩田, 鈴村: "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". 電子通信学会論文誌 (A) J 61-A, No.1 (1978)
- 9) T. Bohlin "Four cases of identification of changing systems". System Identification: Advances and Case Studies. Academic Press (1976)
- 10) K. J. Åström: Introduction to stochastic control theory, Academic Press (1970)
- 11) 有本卓: カルマン・フィルタ, 産業図書 (1977).
- 12) H. Sugimoto, N. Ishii, A. Iwata and N. Suzumura: "Stationarity and normality test for biomedical data". Computer programs in biomedicine vol. 7 No. 4 (1977)