解　説

時系列解析—時間領域でのモデリングとその応用[†]

北　川　　源四郎[††]

1. まえがき

時間と共に不規則な変動を示す対象の観測値の系列を時系列と呼ぶことにする．航空機や船舶の振動，工業的な生産プロセス，地震，気象や経済現象などの多くはこのような対象と考えられるであろう．このような不規則な変動の中に一定の規則性，すなわち確率的な構造を見い出し，ひいてはその予測や制御を実現しようとするのが時系列解析の目的である．

良く知られたスペクトル解析はこれらの不規則変動を周波数成分ごとに分解し，各周波数帯ごとにどのように成分が分布しているかを表示することにより解析を行おうとする周波数領域からの接近法である．これに対し本稿で示す方法は時系列モデルのあてはめに基礎を置く時間領域からの接近法である．スペクトルと線形時系列モデルの間には一定の対応関係が存在することから両者の接近法は同等であるような印象を与える．しかしながら，限られた個数の観測値にもとづいて意味のある結果を抽き出そうとする場合には両者の著しい相違が明らかとなる．このような場合には，適切なモデルを導入することによって初めて有効な情報を抽き出すことができる．たとえば，従来のスペクトル解析ではフィードバックループの存在の下で得られたデータからシステムの動得性を推定することは不可能であった．しかし，このような場合にも適切なモデルの導入によって動得性の推定，解析が可能となる．さらに，時系列モデルは直ちに予測や制御へ利用できるという特長を持っている．このように時間領域においてパラメトリックなモデルを利用する事により従来のスペクトル解析では困難であった問題が容易に取り扱えるようになった．

しかしながらこれらの成果は，それぞれの対象，目的に応じて適切なモデルが利用されることを前提としている．「モデルを通して情報を測る」方法は，適切なモデルの利用によってはじめてその威力を発揮するからである．こうして，いくつかのモデルが想定されるときモデルの選択が合理的に行えるか否かがこの方法の成功の鍵となる．

情報量規準 AIC はこの問に対する一つの有力な解答である．AIC は一般化エントロピーの −2 倍の不偏推定を目指して得られたものである．エントロピーは我々が想定したモデルが規定する確率分布の真の（観測値を生成した）分布の近似としての良さを評価する規準とみなすことができる．したがって，AIC を最小とするモデルを選択する AIC 最小化によって客観的かつ合理的なモデル選択が実現される．ただし，AIC の値がほぼ等しく優劣をつけ難いモデルがいくつか存在する場合には，AIC 最小化法はそのうちの 1 つのモデルだけを選択することによるリスクを伴う．このような場合には，むしろこれらのモデルを平均することによって，より安定したモデルが得られると期待される．本稿で簡単に紹介するベイズモデルはこの目的のために導入されたものであり，ベイズモデルの利用もまたエントロピーの最大化の具体化の 1 つであることが明らかにされている．

この稿では，まず線形定常時系列モデルをとりあげ AIC 最小化法とベイズ法によるあてはめ方とスペクトル推定，フィードバック系の解析，予測，制御への応用について解説する．適切なモデル利用の必要性の認識とモデルの良し悪しを比較する規準の存在は我々により適切なモデルを探し続けることを要求する．第 6 章で示される局所定常 AR モデル，非定常モデル，非線形モデルはこうした試みの一例であり今後の発展が望まれる分野である．最後に時系列の解析，予測，制御を目的として開発されたプログラムパッケージ TIMSAC について，その特徴や構成などを簡単に解説する．

† Time Series Analysis—A Time Domain Approach by Genshiro KITAGAWA (The Institute of Statistical Mathematics).

†† 統計数理研究所

2. AIC 最小化法とベイズ法[6),10),20)]

統計的なモデルを決定することはそれに対応する確率分布を定めていることと同等である．観測値 $x_1, \cdots, x_N$ を生成する真の確率密度関数を $g(y)$ と書き表わすことにすると，我々が想定するモデルの良さはそのモデルが規定する確率密度 $f(y)$ の $g(y)$ の近似としての良さとして評価することができる．したがってこの分布の近似の良さを評価する自然な規準を導入することにより統計的モデリングのための統一的な方法が確立することになる．

この規準として我々は一般化されたエントロピー[10)]

$$B(g\,;\,f)=\int \log\left\{\frac{f(y)}{g(y)}\right\}g(y)dy$$
$$=E_Y\log\left\{\frac{f(Y)}{g(Y)}\right\} \tag{2.1}$$

を用いることにする．$B(g\,;\,f)$ の符号を反転したものはカルバックの情報量として知られている．$B(g\,;\,f)$ は0また負の値をとり，分布 $g(y)$ と $f(y)$ が一致する場合に限り0となる．ボルツマンによるエントロピーの確率論的な解釈に従えば，一般化エントロピーは我々が想定したモデル，すなわち分布 $f(y)$ から $g(y)$ のような分布が得られる確率の対数という解釈が得られる．したがって，$B(g\,;\,f)$ が大であるほど $f(y)$ は $g(y)$ を良く近似しているとみなすことができる．

統計的なモデルのあてはめは，データ x にもとづいて真の分布 $g(y)$ を良く近似する確率分布 $f(y|x)$ を推定することとみなせることから，$B(g\,;\,f)$ をモデルあてはめの規準とすることができる．統計的なモデリングにおいて $B(g\,;\,f)$ を最大にするようにモデルを選択しようとする立場はエントロピー最大化の原理と呼ばれている．

次に，このエントロピー最大化の原理が実際のモデリングの場面にどのように具体化されているかを見てみよう．まず，観測値 $x_1, x_2, \cdots, x_N$ はパラメータ θ によって規定される確率分布 $f(x|\theta)$ に従うというモデルを考え，そのパラメータ θ の推定を行うことにする．エントロピー $B(g(y)\,;\,f(y\,;\,\theta))$ は

$$B(g\,;\,f)=E_Y\log f(Y|\theta)-E_Y\log g(Y) \tag{2.2}$$

と書け，右辺の第2項は θ に依存しない定数であることから，エントロピーの最大化は $E_Y\log f(Y|\theta)$ の最大化と同値であることが分かる．一方，観測値が独立に得られる場合尤度 (likelihood)

$$L(\theta)=\prod_{i=1}^{N} f(x_i|\theta) \tag{2.3}$$

の対数である対数尤度

$$l(\theta)=\sum_{i=1}^{N}\log f(x_i|\theta) \tag{2.4}$$

は N が無限大に近づくとき，確率1で

$$\frac{1}{N}l(\theta)\longrightarrow E_Y\log f(Y|\theta) \tag{2.5}$$

となる．このことから，尤度あるいは対数尤度を最大にするようにパラメータ θ を決める最尤法は近似的にエントロピーの最大化を実現しようとしていることがわかる．

エントロピー最大化の原理をさらに進めると，最尤推定値 $\hat{\theta}$ で定められる各種のモデルの相互比較をそれぞれの最大対数尤度 $l(\hat{\theta})$ の比較で実現しようという考えに導かれる．しかし，回帰モデルの場合に次数の増加と共に最大対数尤度は単調に増加し，常に最高次数が選択されることからも分かるように，最大対数尤度はそのままの形ではモデルの相互比較の役に立たない．それは，最大対数尤度 $l(\hat{\theta})$ が最尤推定値 $\hat{\theta}$ で規定されるモデルの $E_Y f(Y|\theta)$ の推定量として偏差を持っていることによる．この偏りの修正を施すことによってはじめて各種のモデルの良さの相互比較が可能となる．

簡単のためにモデル $f(y|\theta)$ は θ を適当に選ぶことによって真の分布を含む，すなわち $f(y|\theta_0)=g(y)$ となるような θ_0 が存在する場合を考えることにする．$E_Y\log f(Y|\hat{\theta})$ を θ_0 の周辺で展開することにより

$$E_Y\log f(Y|\hat{\theta})=\int\log f(y|\hat{\theta})g(y)dy$$
$$\sim\int\log f(y|\theta_0)g(y)dy$$
$$+\int\frac{\partial}{\partial\theta}\log f(y|\theta_0)g(y)dy(\hat{\theta}-\theta_0)$$
$$+\frac{1}{2}(\hat{\theta}-\theta_0)'\int\frac{\partial^2}{\partial\theta\partial\theta'}\log f(y|\theta_0)g(y)dy(\hat{\theta}-\theta_0)] \tag{2.6}$$

となる．θ_0 は $E_Y\log f(Y|\theta)$ の最大値を与えることから右辺の第2項は0となる．したがって

$$J(\theta_0)=-E\frac{\partial^2}{\partial\theta\partial\theta'}\log f(Y|\theta_0)$$

とおくと

$$E_Y\log f(Y|\hat{\theta})\sim E_Y\log f(Y|\theta_0)-\frac{1}{2}(\hat{\theta}-\theta_0)'J(\theta_0)(\hat{\theta}-\theta_0) \tag{2.7}$$

となる．一方，$E_Y \log f(Y|\theta_0)$ は

$$E_Y \log f(Y|\theta_0) = \{E_Y \log f(Y|\theta_0) - \frac{1}{N}\sum_{i=1}^{N} \log f(x_i|\theta_0)\} + \frac{1}{N}\sum_{i=1}^{N} \{\log f(x_i|\theta_0) - \log f(x_i|\hat{\theta})\} + \frac{1}{N}\sum_{i=1}^{N} \log f(x_i|\hat{\theta}) \quad (2.8)$$

となる．ここで $\frac{1}{N}\sum_{i=1}^{N} \log f(x_i|\theta_0)$ を展開して

$$\frac{1}{N}\sum_{i=1}^{N} \log f(x_i|\theta_0) \sim \frac{1}{N}\sum_{i=1}^{N} \log f(x_i|\hat{\theta}) + \frac{1}{N}\sum_{i=1}^{N} \frac{\partial}{\partial\theta} \log f(x_i|\hat{\theta})(\theta_0-\hat{\theta}) + \frac{1}{2}(\theta_0-\hat{\theta})\frac{1}{N}\sum \frac{\partial^2}{\partial\theta\partial\theta'} \log f(x_i|\theta_0)(\theta_0-\hat{\theta}) \quad (2.9)$$

となるが，$\hat{\theta}$ は最尤推定値であることから，右辺第2項は0となる．さらに $\frac{1}{N}\sum_{i=1}^{N} \log f(x_i|\hat{\theta}) \equiv l(\hat{\theta})$ を使うと (2.7) 式は

$$E_Y \log f(Y|\hat{\theta}) \sim l(\hat{\theta}) - (\theta_0-\hat{\theta})^t(\theta_0)(\theta_0-\hat{\theta}) + E_Y \log f(Y|\theta_0) - \frac{1}{N}\sum_{i=1}^{N} \log f(x_i|\theta_0) \quad (2.10)$$

ここで

$$E_X\{E_Y \log f(Y|\theta_0) - \frac{1}{N}\sum_{i=1}^{N} \log f(x_i|\theta_0)\} = 0$$

$$E_X\{(\theta_0-\hat{\theta})^t J(\theta_0)(\theta_0-\hat{\theta})\} = \mathrm{trace}\, J^{-1}(\theta_0) I(\theta_0) = k \quad (2.11)$$

ただし，k はパラメータ θ の次元，$I(\theta_0)$ は Fisher の情報行列

$$E_X\left\{\frac{\partial}{\partial\theta_i} \log f(X|\theta_0) \frac{\partial}{\partial\theta_j} \log f(X|\theta_0)\right\}$$

であることを考慮すると

$$E_X\{E_Y \log f(Y|\hat{\theta}) - l(\hat{\theta})\} \sim -k \quad (2.12)$$

となる．このようにして $l(\hat{\theta})-k$ はほぼ $E_Y \log f(Y|\hat{\theta})$ の不偏推定量を与えることがわかる．

赤池によって提案された AIC[6),10),20)] はこの値を -2 倍して

$$-2l(\hat{\theta}) + 2k \quad (2.13)$$

と定義される．以上のように，AIC を最小にすることとは，エントロピー $B(g;f)$ を近似的に最大にすることに相当する．したがって AIC を最小にするモデルを採用する AIC 最小化法に従えば合理的かつほぼ自動的なモデル選択が可能となる．AIC 最小化法もまたエントロピー最大化原理の一つの実現とみなすことができる．

しかしながら，ほとんど等しい AIC の値を持つモデルがいくつか存在する場合には AIC 最小化法の適用には多少のリスクが伴う．このような場合にはベイズモデルの利用が有効である．ベイズモデルとエントロピー最大化原理の関係などは参考文献10) に譲ってここでは特殊な利用法についてのみ示すことにする．

AIC は最尤推定値 $\hat{\theta}$ で規定されるモデルの平均対数尤度の -2 倍の漸近的不偏推定量として得られたものである．このことから，

$$\exp\{-\frac{1}{2}\mathrm{AIC}\}$$

がこのモデルの“尤度”の定義として適当であることがわかる[7)]．ベイズの定理によれば

$$(\text{事後確率}) \propto \exp\{-\frac{1}{2}\mathrm{AIC}\} \times (\text{事前確率})$$

となる．したがって，適当な事前確率を定め，パラメータを最尤法で定めた種々のモデルをそれぞれの事後確率にもとづいて平均することによって一つの実用的なベイズ推定値が得られることになる．

3. 線形定常時系列モデル

線形定常な過程は一般に白色雑音 ε_n を入力とする線形システムの出力として

$$x_n = \varepsilon_n + \sum_{i=1}^{\infty} b_i \varepsilon_{n-i} \quad (3.1)$$

と表わすことができる．あるいは，過去の履歴の線形和として

$$x_n = \sum_{i=1}^{\infty} a_i x_{n-i} + \varepsilon_n \quad (3.2)$$

と表現することができる．しかしながら，有限個の観測値にもとづいて推測を行う時系列解析では無限個のパラメータを持つこれらのモデルは有効ではない．そこで有限個のパラメータによって規定されるモデルを考えることにする．

(3.1), (3.2) において無限和を有限和で置き換えた

$$x_n = \varepsilon_n + b_1\varepsilon_{n-1} + \cdots + b_l\varepsilon_{n-l} \quad (3.3)$$

$$x_n = a_1 x_{n-1} + \cdots + a_m x_{n-m} + \varepsilon_n \quad (3.4)$$

(ε_n は平均0，分散 σ^2 の白色雑音)

はそれぞれ，l 次の移動平均 (MA) モデル，m 次の自己回帰 (AR) モデルと呼ばれ，時系列解析に使われる代表的なモデルである．ただし，一次の AR 過程を表現するには無限次の MA モデルが必要であり，その逆も同様であることから，有限のパラメータを持つ

モデルで一般の線形時系列をよく表現するためには自己回帰と移動平均の両方の項を備えたモデルの方が都合がよいことが分かる．

$$x_n+\sum_{i=1}^{m}a_ix_{n-i}=\varepsilon_n+\sum_{i=1}^{l}b_i\varepsilon_{n-i} \tag{3.5}$$

は (m,l) 次の自己回帰移動平均 (ARMA) モデルと呼ばれる．

Backshift operator $Bx_n=x_{n-1}$ を導入すると (3.5) は

$$(1+\sum_{i=1}^{m}a_iB^i)x_n=(1+\sum_{i=1}^{l}b_iB^i)\varepsilon_n$$

と書ける．したがって，x_n は白色雑音 ε_n を入力とし伝達関数が $(1+\sum_{i=1}^{m}a_iB^i)^{-1}(1+\sum_{i=1}^{l}b_iB^i)$ であるようなシステムの出力とみなすことができる．時系列 x_n が定常であるためには $1+\sum_{i=1}^{m}a_iB^i=0$ の根がすべて単位円の外になければならない．

ARMA 過程の自己共分散関数 $\gamma_k=Ex_nx_{n+k}$ は

$$\gamma_k+\sum_{i=1}^{m}a_i\gamma_{|k-i|}=\lambda_{-k}+\sum_{i=1}^{l}b_i\lambda_{i-k} \tag{3.6}$$
$$(k=0,1,\cdots)$$

によって与えられる．ただし，λ_j は ε_{n-j} と x_n の共分散 $E\varepsilon_{n-j}x_n$ であり

$$\lambda_j+\sum_{i=1}^{m}a_i\lambda_{j-i}=\sigma^2(\delta_j+\sum_{i=1}^{l}b_i\delta_{j-i}) \tag{3.7}$$

の解として得られる．ただし，δ_n はクロネッカーのデルタ ($\delta_0=1$, $\delta_i=0$ $i\neq0$) である．このことから，自己共分散 γ_k は，$k>l$ では

$$\gamma_k+a_1\gamma_{k-1}+\cdots+a_m\gamma_{k-m}=0$$

に従って減衰することが分かる．特に l 次の MA モデルの場合 $k>l$ では $\gamma_k=0$ となる．

ARMA 過程のスペクトル[1]は γ_k のフーリエ変換として

$$p(f)=\frac{|1+b_1e^{-2\pi if}+\cdots+b_le^{-2\pi ilf}|^2}{|1+a_1e^{-2\pi if}+\cdots+a_me^{-2\pi imf}|^2}\sigma^2 \tag{3.8}$$
$$\left(-\frac{1}{2}\leqslant f\leqslant\frac{1}{2}\right)$$

で与えられる．AR 過程，MA 過程のスペクトルは (3.8) で，それぞれ分子および分母を 1 とおいたものである．(3.8) において分母が小となる周波数でスペクトルはピークを持ち，分子が小となる所では谷をもつことになる．直観的には，AR モデルはスペクトルのピークの，また MA モデルは谷の表現に適したモデルであるといえる．この事からも一般のスペクトルを少ないパラメータで表現するには ARMA モデルの利用が有効であることが理解できる．

線形システムの表現として状態空間表現が知られているが，白色雑音を入力とする線形システムの表現

$$z_n=Fz_{n-1}+G\varepsilon_n$$
$$x_n=Hz_n \tag{3.9}$$

が実は ARMA モデルと同等であることが知られている[2]．ARMA モデル (3.5) が得られているとき w_k をインパルス応答関数 ($w_0=1$, $w_k=b_k-\sum_{i=1}^{k}a_iw_{k-i}$) とするとき，

$$F=\begin{pmatrix} & 1 & & \\ & & \ddots & \\ & & & 1 \\ -a_m & -a_{m-1} & \cdots & -a_1\end{pmatrix},\quad G=\begin{pmatrix}w_0\\ w_1\\ \vdots\\ w_{m-1}\end{pmatrix} \tag{3.10}$$
$$H=(1\ 0\ \cdots\ 0)$$

とおくことによって状態空間表現 (3.9) が得られる．

時系列モデルは直ちに予測への応用が可能である．自己回帰モデル (3.4) と観測値 $x_{n-m},\cdots,x_{n-1}$ が与えられている場合には x_n の予測値は

$$\hat{x}_n=\sum_{i=1}^{m}a_ix_{n-i} \tag{3.11}$$

で与えられる．より一般に時系列モデルの状態空間表現 (3.9) を用いればカルマンフィルタによる予測ができる．すなわち $i=1,\cdots,n$ について次の 2 つのステップを繰り返せばよい．

時間更新

$$\hat{z}_{n|n-1}=F\hat{z}_{n-1|n-1}$$
$$P_{n|n-1}=FP_{n-1|n-1}F^t+\sigma^2GG^t \tag{3.12}$$

観測による修正

$$\hat{z}_{n|n}=\hat{z}_{n|n-1}+K_n(x_n-H\hat{z}_{n|n-1})$$
$$P_{n|n}=P_{n|n-1}-K_nHP_{n|n-1}$$
$$K_n=P_{n|n-1}H^t(HP_{n|n-1}H^t)^{-1} \tag{3.13}$$

(3.6) で求められる共分散関数の理論値 γ_k を使って初期共分散行列を $P_{0|0}=(\gamma_{|i-j|})$，初期ベクトル $\hat{z}_{0|0}$ として零ベクトルを使い上の 2 つのステップを繰り返すとき $\hat{x}_{n+1|n}=H\hat{z}_{n+1|n}$ は観測値 $x_1,\cdots,x_n$ が得られた下での x_{n+1} の予測値となり，$r_{n+1}=HP_{n+1|n}H^t$ はその予測誤差の分散を与える．より正確には，x_{n+1} は平均 $\hat{x}_{n+1|n}$，分散 r_{n+1} の正規分布に従うことになる．

4. 時系列モデルのあてはめ

次に，与えられた観測値 $x_1,\cdots,x_N$ に時系列モデルをあてはめる方法を示す．まず，一般の ARMA モデルに対する AIC 最小化法を示し，次に AR モデルの

場合にはこれが最小二乗法などによって近似できることを示す．さらに AR モデルのベイズ推定値を与える．

ARMA モデル

前章で示したように時系列モデルの状態空間表現を用いればカルマンフィルタによる逐次予測が可能であった．これを利用して正規分布を仮定した下での ARMA モデルの尤度，AIC の計算ができる．

(m, l) 次の ARMA モデルの尤度は次のように書くことができる．

$$\begin{aligned} &L(x_1, \cdots, x_N \mid \sigma^2, a_1, \cdots, a_m, b_1, \cdots, b_l) \\ &= p(x_1, x_2, \cdots, x_N) \\ &= p(x_1) \prod_{i=2}^{N} p(x_i \mid x_1, \cdots, x_{i-1}) \end{aligned} \tag{4.1}$$

ここで，$p(x_1, \cdots, x_N)$ は $x_1, \cdots, x_N$ の同時分布，$p(x_i \mid x_1, \cdots, x_{i-1})$ は観測値 $x_1, \cdots, x_{i-1}$ が得られた下での x_i の条件付確率分布を表わす．カルマンフィルタ (3.12), (3.13) の結果を使って，$p(x_n \mid x_1, \cdots, x_{n-1})$ の平均は $\hat{x}_{n|n-1}$，分散は $r_n = H p_{n|n-1} H'$ で与えられる．したがって

$$\begin{aligned} &p(x_n \mid x_1, \cdots, x_{n-1}) \\ &= (2\pi r_n)^{-1/2} \exp\{-\frac{1}{2r_n}(x_n - \hat{x}_{n|n-1})^2\} \end{aligned}$$

となるから，尤度 (4.1) は

$$L' = \prod_{n=1}^{N} \left(\frac{1}{2\pi r_n}\right)^{1/2} \exp\left\{-\frac{1}{2r_n}(x_n - \hat{x}_{n|n-1})^2\right\} \tag{4.2}$$

で求められる[4]．

このように，パラメータの値を定めるとモデルの尤度が計算できるのでこの尤度を数値的に最大化することによって (m, l) 次の ARMA モデルの最尤推定値 $\hat{a}_1, \cdots, \hat{a}_m$, $\hat{b}_1, \cdots, \hat{b}_l, \hat{\sigma}^2$ が求められる．このモデルの AIC は尤度 (4.1) の最大値の対数を $\hat{l}$ として

$$\mathrm{AIC}(m, l) = -2\hat{l} + 2(m + l + 1) \tag{4.3}$$

となる．AIC 最小化法に従えば，(m, l) のさまざまな組み合わせに対し AIC を計算し，これを最小とするモデルを採用すればよい．

実際に (m, l) のすべての組み合わせに対しモデルを求めるのはかなりの計算量を要するが，正準相関解析によってほぼ最適な次数が分かるので，その近辺を探索するようにすると効率がよい[2]．

AR モデル

AR モデルの尤度は上に示した計算法で $l=0$ とおくことによって計算できるが，以下に示す Yule-Walker 法や最小二乗法を用いればきわめて簡単に推定値が求められ，多くの場合，実用上十分な近似が得られる．

(3.4) より

$$\varepsilon_n = x_n - a_1 x_{n-1} - \cdots - a_m x_{n-m}$$

となる．したがって，予測誤差の二乗平均

$$\begin{aligned} E\varepsilon_n^2 &= E(x_n - a_1 x_{n-1} - \cdots - a_m x_{n-m})^2 \\ &= \gamma_0 - 2\sum_{i=1}^{m} a_i \gamma_i + \sum_{i=1}^{m} \sum_{j=1}^{m} a_i a_j \gamma_{|i-j|} \end{aligned}$$

を最小とする $a_1, \cdots, a_m$ は Yule-Walker 方程式[1]

$$\gamma_i = \sum_{j=1}^{m} a_j \gamma_{|i-j|} \quad (i = 1, \cdots, m) \tag{4.4}$$

の解として得られ，$E\varepsilon_n^2$ の最小値は

$$d(m) = \gamma_0 - \sum_{i=1}^{m} a_i \gamma_i \tag{4.5}$$

で与えられる．データ $x_1, \cdots, x_N$ が得られる場合には，自己共分散関数の推定値を

$$\gamma_k = \frac{1}{N} \sum_{i=1}^{N-k} x_i x_{i+k} \qquad (k = 0, 1, \cdots, m)$$

によって求め，これを (4.4), (4.5) に代入することによって $\hat{a}_1, \cdots, \hat{a}_m, \hat{\sigma}^2$ が求められる．これが Yule-Walker 法である．

次に，最小二乗解は次のようにして得られる．尤度 (4.1) は

$$L = p(x_1, \cdots, x_M) \prod_{i=M+1}^{N} p(x_i \mid x_{i-1}, \cdots, x_{i-m})$$

と書ける．したがって，対数尤度 $\log L$ は

$$\begin{aligned} l &= \log p(x_1, \cdots, x_M) \\ &\quad + \sum_{i=M+1}^{N} \log p(x_i \mid x_{i-1}, \cdots, x_{i-m}) \end{aligned}$$

となるが，$\log p(x_1, \cdots, x_M)$ の部分を無視すると，l の最大化は

$$d(m) = \frac{1}{N-M} \sum_{i=M+1}^{N} (x_i - a_1 x_{i-1} - \cdots - a_M x_{i-m})^2 \tag{4.6}$$

の最小化と同値になる．この最小二乗法の解法としては直交変換にもとづく方法がきわめて精度がよく，また取り扱いが便利である．

$$Z = \begin{pmatrix} x_M & x_{M-1} \cdots & x_1 \\ x_{M+1} & x_M & \cdots x_2 \\ \vdots & \vdots & \vdots \\ x_{N-1} & x_{N-2} \cdots & x_{N-M} \end{pmatrix},$$

$$y = \begin{pmatrix} x_{M+1} \\ x_{M+2} \\ \vdots \\ x_N \end{pmatrix}, \quad a = \begin{pmatrix} a_1 \\ \vdots \\ a_m \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{4.7}$$

とおくと (4.6) は

$$d(m)=\frac{1}{N-M}\|\boldsymbol{y}-Z\boldsymbol{a}\|^2$$

と書ける．ただし $\|\cdot\|$ はユークリッドノルムである．ここで行列 Z を三角化する直交行列 U を考えることにする．すなわち，

$$UZ=\begin{pmatrix} S_{11} & \cdots & S_{1M} \\ & \ddots & \vdots \\ & & S_{MM} \\ 0 & & \end{pmatrix} \tag{4.8}$$

$$UU^t=U^tU=I$$

とする．このUを構成する代表的なものとしてハウスホルダー変換[12]がある．このとき，$U\boldsymbol{y}=\boldsymbol{z}$ とおくと

$$\begin{aligned} d(m)&=\frac{1}{N-M}\|U\boldsymbol{y}-UZ\boldsymbol{a}\|^2 \\ &=\frac{1}{N-M}\left\|\begin{pmatrix} z_1 \\ \vdots \\ z_{N-M}\end{pmatrix} -\begin{pmatrix} S_{11} & \cdots & S_{1M} \\ & \ddots & \vdots \\ & & S_{MM} \\ 0 & & \end{pmatrix}\begin{pmatrix} a_1 \\ \vdots \\ a_m \\ 0\end{pmatrix}\right\|^2 \\ &=\frac{1}{N-M}\left\{\left\|\begin{pmatrix} z_1 \\ \vdots \\ z_m\end{pmatrix} -\begin{pmatrix} S_{11} & \cdots & S_{1m} \\ & \ddots & \vdots \\ 0 & & S_{mm}\end{pmatrix}\begin{pmatrix} a_1 \\ \vdots \\ a_m\end{pmatrix}\right\|^2+\sum_{i=m+1}^{N-M} z^2_i\right\} \end{aligned} \tag{4.9}$$

となるから，$a_1,\cdots,a_m$ を

$$\begin{pmatrix} a_1 \\ \vdots \\ a_m\end{pmatrix}=\begin{pmatrix} S_{11} & \cdots & S_{1m} \\ & \ddots & \vdots \\ & & S_{mm}\end{pmatrix}^{-1}\begin{pmatrix} z_1 \\ \vdots \\ z_m\end{pmatrix} \tag{4.10}$$

とするとき $d(m)$ は最小値 $\frac{1}{N-M}\sum z_i^2$ をとり，そのモデルの AIC は近似的に

$$\mathrm{AIC}(m)=(N-M)\log d(m)+2(m+1) \tag{4.11}$$

で与えられる．いったん三角行列 UZ およびベクトル $\boldsymbol{z}$ が求められると (4.11) によって $\mathrm{AIC}(m)(m=0,1,\cdots,M)$ は簡単に求められる．したがって，実際には，まず $\mathrm{AIC}(m)$ の最小値を探し，その次数 m に対して (4.10) によって係数 $\hat{a}_1,\cdots,\hat{a}_m$ を求めればよい．直交変換にもとづくこの解法は一変量の AR モデルだけでなく，多変量 AR モデル，非線形モデル，局所定常 AR モデルなどへも適用できる．

ベイズ推定値[7),13)]

次に自己回帰型のベイズモデルを示す．まず，第2章の議論に従って m 次の AR モデルの“尤度”を

$$p(x\,|\,m)=\exp\left\{-\frac{1}{2}\mathrm{AIC}(m)\right\}$$

と定義する．また m 次のモデルの事前確率を

$$\pi(m)\propto\frac{1}{m+1}$$

とする．このとき m 次の AR モデルの事後確率は

$$\pi(m\,|\,x)\propto p(x\,|\,m)\pi(m)$$

で与えられる．我々のベイズ推定値は各種次数のモデルの事後確率による加重平均として得られる．このモデルの平均化は

$$b_m=\mathrm{sign}\,(S_{mm})\sqrt{\frac{z_{m+1}}{\sum_{i=m+1}^{N-M} z_i^2}}$$

によって推定される偏自己相関係数を使って行われる．まず，偏自己相関係数のベイズ推定値が

$$\hat{b}_m=b_m\sum_{k=m}^{M}\pi(m\,|\,x)$$

で求められる．次に AR 係数のベイズ推定値は，次数に関する漸化式 ($m=1,\cdots,M$)

$$a_m{}^m=b_m$$

$$a_i{}^m=a_i{}^{m-1}-b_m a_{m-i}{}^{m-1}\qquad (i=1,\cdots,m-1)$$

によって得られる $a_1{}^M,\cdots,a_M{}^M$ として与えられる．このように，ベイズモデルは常に最高次数 M 次のモデルとなる．

5. 線形定常モデルの応用

線形時系列モデルは第3章に示したように直ちに予測に利用できる．本章ではそのほかの応用について簡単に示すことにする．

スペクトル推定

一変量の AR モデルや ARMA モデルが得られると (3.8) によって対応するパワースペクトルが得られる．したがって，AIC 最小化法またはベイズ法によってモデルを推定することにすれば，ウィンドウの選択などの主観的な判断に頼ることなくスペクトルの推定ができる．

また，多変量 AR モデル

$$\boldsymbol{x}_n=\sum_{i=1}^{m} A_i\boldsymbol{x}_{n-i}+\boldsymbol{\varepsilon}_n \tag{5.1}$$

$$\boldsymbol{\varepsilon}_n\sim N(O,S)$$

が得られると，パワーおよびクロススペクトルの行列の $P(f)$ の推定値は

$$P(f)=(A(f)^*S^{-1}A(f))^{-1}\left(-\frac{1}{2}\leq f\leqq\frac{1}{2}\right) \tag{5.2}$$

で得られる[1]. ただし, $A(f)=\sum_{j=1}^{m} A_j \exp(-2\pi i j f)$, $*$は複素共役を表わす. また x_n の第 i 成分 $x_n(i)$ を入力, 第 j 成分 $x_n(j)$ を出力とする系の周波数応答関数を $G_{ij}(f)$ と表わすことにすると

$$G_{ij}(f)=-\frac{q_{ij}(f)}{q_{ii}(f)} \qquad (5.3)$$

となる. ただし, $q_{ij}(f)$ は $P(f)$ の逆行列の (i, j) 成分を示す.

フィードバック系の解析[1]

フィードバック系の解析のために, 各変数 $x_n(i)$ は自分自身以外 $x_n(j)$ $(i \neq j)$ からの影響とその変数固有の雑音 $u_n(i)$ の和として表現するモデル

$$x_n(i)=\sum_{m=1}^{M}\sum_{j=1}^{k} a_m(i, j)x_{n-m}(j)+u_n(i) \qquad (5.4)$$

を考えることにする. $a_m(i, j)$ は入力 $x_n(j)$ に対する $x_n(i)$ のインパルス応答関数であり $a_m(i, i)=0$ と仮定する. ここで, 雑音 $u_n(i)$ AR モデル

$$u_n(i)=\sum_{l=1}^{L} c_l(i)u_{n-l}(i)+\varepsilon_n(i) \qquad (5.5)$$

を考え, これに (5.4) を代入すると

$$x_n(i)=\sum_{l=1}^{L} c_l(i)x_{n-l}(i) + \sum_{m=1}^{M+L}\sum_{\substack{j=1\\ j\neq i}}^{k} A_m(i, j)x_{n-m}(j)+\varepsilon_n(i) \qquad (5.6)$$

となる. ただし $A_m(i, j)=a_m(i, j)-\sum_{l=1}^{m-1} c_l(i)a_{m-l}(i, j)$ である. (5.6) は白色雑音間の無相関 $E\varepsilon_n(i)\varepsilon_n(j)=0$ を仮定した $M+L$ 次の k 変量 AR とみなすことができる. したがって, k 変量 AR モデル (5.1) が得られ, $\varepsilon_n(i)$ 間の無相関の仮定がほぼ満たされる場合には

$$c_m(i)=A_m(i, i)$$
$$a_1(i, j)=A_1(i, j)$$
$$a_m(i, j)=A_m(i, j)+\sum_{l=1}^{m-1} c_l(i)a_{m-l}(i, j) \qquad (5.7)$$

によってインパルス応答関数と雑音のモデルが得られる. このとき, $x_n(j)$ から $x_n(i)$ への周波数応答関数は

$$a_{ij}(f)=\sum_{m=1}^{M} a_m(i, j) \exp(-2\pi\sqrt{-1}mf) \qquad (5.8)$$

で求められる.

モデル (5.4) は $x_n(i)$ の変動は結局のところ各変数 $x_n(j)$ への固有の雑音 $u_n(j)$ からの影響の和として表わせるものと仮定している. このとき

$$\gamma_{ij}=\frac{|b_{ij}(f)|^2 p_{u(j)}(f)}{\sum_{l=1}^{k} |b_{il}(f)|^2 p_{u(l)}(f)} \qquad (5.9)$$

は $x_n(i)$ のパワースペクトル $p_{ii}(f)$ のうち雑音源 $u_n(j)$ に起因する部分の割合を表わす. ただし $P_{u(i)}(f)$ は雑音 $u(i)$ のパワースペクトルである. r_{ij} はノイズ寄与率と呼ばれ, 制御系の設計などにきわめて重要な役をはたす.

統計的制御系の設計[1]

k 変量 AR モデル (5.1) において,

$$x_n=\begin{bmatrix} z_n \\ y_n \end{bmatrix}$$

とし, z_n は被制御変数, y_n は操作変数とする. (5.1) において被制御変数 z_n の部分だけに注目することにすると

$$z_n=\sum_{m=1}^{M} a_m z_{n-m}+\sum_{m=1}^{M} b_m y_{n-m}+\varepsilon_n \qquad (5.10)$$

という表現が得られる. この制御用のモデルは

$$F=\begin{pmatrix} a(1) & I & & \\ a(2) & & \ddots & \\ \vdots & & & I \\ a(M) & & & \end{pmatrix},\quad G=\begin{pmatrix} b(1) \\ b(2) \\ \vdots \\ b(M) \end{pmatrix},\quad V_n=\begin{pmatrix} \varepsilon_n \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$
$$H=[I\ 0 \cdots 0] \qquad (5.11)$$

とおくとき

$$Z_n=FZ_{n-1}+Gy_{n-1}+V_n$$
$$z_n=HZ_n \qquad (5.12)$$

という状態空間表現を持つ. この制御系に対し, 制御入力の良さを評価するために二次評価基準

$$J=E\{Z_n{}^tQZ_n+y_n{}^tRy_n\} \qquad (5.13)$$

を導入することにする. ただし, Q, R はそれぞれ非負, 正定行列であり, Eは期待値を示すものとする. このとき Jを最小とする最適制御は

$$y_n=DZ_n \qquad (5.14)$$

で実現される. ただし最適制御ゲイン Dは $P_0=Q$ とおき

$$M_i=P_{i-1}-P_{i-1}G(R+G^tP_{i-1}G)^{-1}G^tP_{i-1}$$
$$P_i=F^tM_iF+Q \qquad (5.15)$$

を P_i がある Pに収束するまで繰り返すとき

$$D=-(R+G^tPG)^{-1}G^tPF \qquad (5.16)$$

によって得られる.

このように, 制御系の観測データに多変量 AR モデルをあてはめ, それにもとづいて最適制御系を構成することができる. この簡単な制御方式によってすでに

(1) セメントキルンの制御[1]

(2) 火力発電所ボイラの制御[16]

(3) 船舶のオートパイロット[17]

が実用化されている.

6. 非定常，非線形モデル

これまで時系列は定常，線形であるものとしてモデル化を行ってきたが，実際には時間と共に徐々に確率構造が変化したり顕著な非線形性がみられる場合がある．これらの非定常あるいは非線形時系列の解析，予測のために種々のモデルが開発されつつある．ここでは，そのうちのいくつかをごく簡単に紹介する．

局所定常 AR モデル[18),12)]

時間と共に徐々に構造が変化していく時系列は時間軸を適当に分割することによりそれぞれの区間ではほぼ定常であるとみなせる．したがって，分割の仕方と各区間上の AR モデルを適切に選択することにより時間と共に徐々にその構造が変化していく時系列を表現できるモデルが得られる．このモデルを局所定常 AR モデルと呼ぶ．

長さ N の非定常時系列 $x_1, x_2, \cdots, x_N$ が得られているとする．これを，それぞれ長さ $n_i(i=1, \cdots, k)$ の k 個の小区間に分割し，各小区間では時系列は定常とみなすことにし，AR モデル

$$x_n=\sum_{j=1}^{m_i} a_j{}^i x_{n-j}+\varepsilon_n{}^i \tag{6.1}$$

をあてはめる．このモデルのあてはまりの良さは

$$\mathrm{AIC}=\sum_{i=1}^{k}\{n_i \log \hat{\sigma}_i{}^2+2(m_i+1)\} \tag{6.2}$$

によって評価できる．ここで，$\hat{\sigma}_i{}^2$ は i 番目の小区間にあてはめられたモデルの予測誤差の分散の推定値である．MAICE 法に従えば，種々の k, n_i の組のうち AIC を最小とするものを選べばよい．実際には可能なすべての組合せについて調べるのは不可能に近いが，最適な分割を近似的に求める計算法が開発されている．さらに局所定常 AR モデルあてはめのためのベイズ手法も開発されている．(3.8) によって各小区間上のスペクトルが得られる．したがって，このようなモデルの導入によりスペクトルの時間変化を自動的にとらえることができる．

スペクトルの時間変化がもっと急激な場合を想定したモデルもいくつか開発されている[9),15)].

季節調整

季節成分とトレンド成分からなる時系列 x_n を想定し，この時系列をトレンド成分 T_n，季節成分 S_n と不規則成分 I_n の和

$$x_n=T_n+S_n+I_n \tag{6.3}$$

に分解するベイズモデルが提案されている[8)]．このモデルは N 個の観測値に対し $2N$ 個以上のパラメータ $\{T_n, S_n\}$ を想定したモデルであるが，T_n と S_n の時間的な動きに対し事前分布の形で制約を課すことにより (6.3) の合理的な分解が可能となる．このベイズモデルによる季節調整用の計算プログラムがすでに開発されている[11)].

また (6.3) の分解は，状態空間表現とフィルタの計算法を利用してより効率的に実現できることも知られている[14)].

非線形モデル[19)]

実際の時系列の中には，リミットサイクルの存在を示す強い周期性を持ったもの，振幅と周期との間に関連があるものなど非線形過程に特有な現象を示すものがある．非線形時系列モデルとしては，非線形自己回帰モデル，非線形移動平均モデル，両線形モデルなどが知られているが，尾崎は[19)]

$$x_n=\sum_{i=1}^{m}(a_i+b_i e^{-x^2_{n-i}})x_{n-i}+\varepsilon_n \tag{6.4}$$

がこれらの非線形現象の解析に適していることを示している．これらの非線形モデルも線形モデルと同様に AIC 最小化法であてはめることができる．

7. プログラムパッケージ TIMSAC

これまで解説してきた時系列の時間領域でのモデリングとその応用のための計算プログラムが開発されている．赤池を中心に開発されたプログラムパッケージ TIMSAC (Time Series Analysis and Control) はその名の通り時系列の解析，予測，制御を目的としたものであり，モデリングに AIC 最小化法，ベイズ法を一貫して採用していることにその特徴がある．

TIMSAC[1)] の第1版は 1972 年に出版され，**表-1** に示すように，従来のスペクトル解析のためのプログラムと，AR モデルによる時系列の解析と制御のためのプログラムで構成されている．1975 年には TIMSAC-74[3)] が出版されたがこのパッケージは，ARMA モデル，局所定常 AR モデル，バイスペクトル計算などに関連したプログラムから成っている．1979 年に出版された TIMSAC-78[5)] には最小二乗法やベイズ法による AR モデル，局所定常 AR モデル，非線形モデルなどのあてはめ，最尤法による AR モデルや ARMA モデルのあてはめ，予測や適応制御のためのプログラムなどが採用されている．TIMSAC シリーズは今後

表-1

パッケージ名	プログラム名	概　要
TIMSAC	AUTCOR	自己共分散関数計算
	MULCOR	相互共分散関数計算（直接法）
	FFTCOR	〃 (FFT 法)
	AUSPEC	パワースペクトル推定
	MULSPE	クロススペクトル推定
	SGLFRE	周波数応答関数推定（一入力）
	MULFRE	〃 （多入力）
	FPEAUT	一変量 AR モデルのあてはめ (Yule-Walker 法)
	FPEC	多変量 AR モデルのあてはめ (Yule-Walker 法)
	MULNOS	ノイズ寄与率計算
	DECONV	インパルス応答計算
	RASPEC	時系列モデルによるスペクトル推定（一変量）
	MULRSP	時系列モデルによるスペクトル推定（多変量）
	OPTDES	最適制御系設計
	OPTSIM	制御系シミュレーション
	WNOISE	ホワイトノイズ生成
TIMSAC-74	CANARM	正準相関解析（一変量）
	AUTARM	ARMA モデルあてはめ（一変量）
	COVGEN	与えられたスペクトルに対応する共分散関数の生成
	CANOCA	正準相関解析（多変量）
	MARKOV	ARMA モデルのあてはめ（多変量）
	PRDCTR	ARMA モデルによる予測
	SIMCON	ARMA モデルによる最適制御系の設計，シミュレーション
	NONST	局所定常 AR モデル(Yule-Walker法)
	PWDPLY	スペクトル表示
	FRDPLY	周波数応答関数表示
	THIRMO	三次モーメント計算
	BISPEC	バイスペクトル計算
TIMSAC-78	UNIMAR	AR モデル（一変量，最小二乗法）
	UNIBAR	〃 （ 〃 ，ベイズ）
	BSUBST	非線形モデル（一変量，ベイズ）
	MULMAR	AR モデル（多変量，最小二乗法）
	MULBAR	〃 （ 〃 ，ベイズ）
	PERARS	周期的ARモデル
	MLOCAR	局所定常 AR モデル（一変量，最小二乗法）
	BLOCAR	〃 （ 〃 ，ベイズ）
	MLOMAR	〃 （多変量，最小二乗法）
	BLOMAR	〃 （ 〃 ，ベイズ）
	NADCON	適応制御系
	EXSAR	AR モデル（一変量，最尤法）
	XSARMA	ARMA モデル（〃， 〃 ）
	TSSBST	Subset regression
	TSWIND	データウィンドウによる一次処理
	TSCHCK	予測誤差によるモデルのチェック
	TSROOT	固有根の計算
	TSTIMS	TIMSAC-78 の編集版
	TSCANC	正準相関解析および ARMA モデルのあてはめ

とも，新しいモデルを加えて順次開発されていく予定である．

8. むすび

時系列解析における時間領域でのモデリングにもとづく方法について解説した．AIC 最小化法やベイズ法による統計的モデルのあてはめの方法を紹介し，とくに線形定常時系列モデルのあてはめ方とそのいくつかの応用を示した．非定常，非線形モデルはごく簡単に紹介するにとどめた．この分野は現在，盛んに研究が行われており，今後の発展が期待される．最後に，時系列モデルのあてはめと解析，予測，制御への応用のためのプログラムパッケージ TIMSAC について，その構成を示した．

参考文献

1) 赤池弘次，中川東一郎：ダイナミックシステムの統計的解析と制御，p. 190，サイエンス社，(TIMSAC のリストを含む) (1972).
2) Akaike, H.: Stochastic theory of Minimal Realization, IEEE Trans. Automatic Control, AC-19, No. 6, pp. 667-674 (1974).
3) Akaike, H., Arahata, E. and Ozaki, T.: TIMSAC-74, A time series analysis and control program package-(1) & (2), Computer Science Monographs No. 5 & No. 6, The Institute of Statistical Mathematics, Tokyo (1975).
4) Akaike, H.: Covariance matrix computation of the state variable of a stationary Gaussian process, Ann. Inst. Statist. Math., 30, B, pp. 479-504 (1978).
5) Akaike, H., Kitagawa, G., Arahata, E. and Tada, F.: TIMSAC-78, Computer Science Monographs No. 11, The Institute of Statistical Mathematics, Tokyo (1979).
6) 赤池弘次：統計的検定の新しい考え方，数理科学，No. 198, pp. 51-57 (1979).
7) Akaike, H.: A Bayesian extension of the minimum AIC procedure of autoregressive model fitting, Biometrika, Vol. 66, No. 2, pp. 237-242 (1979).
8) Akaike, H.: Likelihood and the Bayesian procedure, A paper presented at the Int. Meeting on Bayesian Statistics, Valencia. Spain, (May 28-June 2, 1979).
9) Akaike, H.: On the construction of composite time series models, A paper presented at congress of ISI meeting, Manila, Philippines (1979).
10) 赤池弘次：一物理学周辺の確率統計一 エントロピーとモデルの尤度，日本物理学会誌，第35巻，第7号，pp. 608-614 (1980).
11) Akaike, H. and Ishiguro, M.: BAYSEA, A

Bayesian Seasonal Adjustment Program, Computer Science Monographs, The Institute of Statistical Mathematics, Tokyo (1980).

12) Kitagawa, G. and Akaike, H.: A procedure for the modeling of non-stationary time series, Ann. Inst. Statist. Math., Vol. 30, No. 2, B, pp. 351-363 (1978).

13) 北川源四郎，赤池弘次，大津皓平：局所定常自己回帰による適応制御系の実現，統計数理研究所彙報，第27巻，第1号，pp. 96-105 (1980).

14) Kitagawa, G.: A Nonstationary Time Series Model and its Fitting by a Recursive Technique, Research Memo. No. 184, The Institute of Statistical Mathematics (1980).

15) Kitagawa, G.: Changing Spectrum Estimation, Research Memo. No. 194, The Inst. Statist. Math. (1980).

16) Nakamura, H. and Akaike, H.: Use of statistical identification for optional control of a supercritical thermal power plant, Identification and System parameter identification, Pergamon Press, pp. 221-232 (1979).

17) Ohtsu, K., Horigome, M. and Kitagawa, G.: A new ship's auto pilot design through a stochastic model, AUTOMATICA, Vol. 15, No. 3, pp. 255-268 (1979).

18) 尾崎　統：局所定常時系列のスペクトル解析，応用物理，第46巻，第1号，pp. 61-69 (1977).

19) Ozaki, T.: Statistical Analysis of Perturbed Limit Cycle Processes through Nonlinear Time Series Models, Research Memo. No. 158, The Institute of Statistical Mathematics (1979).

20) 数理科学，特集情報量規準 No. 153，サイエンス社 (1976).

(昭和55年8月21日受付)