

マルコフ切り換え自己回帰モデル

統計数理研究所 北 川 源 四 郎

(1994 年 2 月 受付)

1. ま え が き

脳波には α 波, β 波などと呼ばれるいくつかの特徴的な波形があり, それらが睡眠の状態や環境の変化などに依存して交互に現れることはよく知られている. また, マクロ経済学においては経済変動の結果, 発展期と後退期が交互に現れたり, 動物の個体数や黒点数のデータでは上昇期と下降期が見られるように, いくつかの異なった変動パターンからなる時系列は数多く見られる.

このような, 何種類かの特徴的なパターンからなるデータの解析のために, 確率構造が切り換わるモデルが考えられており, 非線形時系列 (Tong (1990)), 生物 (Churchill (1989), Katzoff and Shumway (1993)), 医学 (Gorden and Smith (1988)) など応用されている. とくに, マクロ経済学への応用 (Goldfeld and Quandt (1973), McCulloch and Tsay (1993)) では, 経済が発展期か収縮期かの判定が興味ある問題とされている. また最近, 工学の分野では音声認識などへの応用から隠れマルコフモデル (hidden Markov model) と呼ばれ, 注目をあびている.

本稿では, 状態に応じて切り換わる自己回帰モデルを考え, その簡単な性質と状態空間モデルを用いて状態の推定やパラメータの推定を行う方法を示す. 4 章の一部は Kitagawa (1988) において口頭発表したものである. マルコフ切り換えモデルへの状態空間モデルの利用は Shumway and Stoffer (1991) や Komaki (1993) でも行われている. 本稿のモデルの場合には, 自己回帰モデルを用いているので, 非ガウス型フィルタおよび平滑化がギブスサンプラーや数値積分を用いるまでもなく簡単に実現できることが特長である.

2. マルコフ切り換え自己回帰モデル

2.1 モデル

時系列 y_n は未知の“状態” s_n に依存してモデルを切り換えるものとする. すなわち, “状態” s_n がとり得る値の集合を $S = \{1, \dots, k\}$ とし, $s_n = j$ のとき時系列 y_n は自己回帰モデル

$$(2.1) \quad y_n = \sum_{i=1}^{m_j} a_{ij} y_{n-i} + b_j \varepsilon_n, \quad \varepsilon_n \sim N(0, \sigma^2)$$

に従うものとする. ただし, ε_n は y_{n-j} , $j > 0$ とは独立な白色雑音とする. 本稿では Komaki (1993) にしたがって“状態” s_n をモードと呼ぶことにするが, 英語の文献では regime と呼ぶことが多いようである. モードの切り換えについては以下のような簡単なマルコフチェーンモデル

$$(2.2) \quad \Pr(s_n = i \mid s_{n-1} = j) = p_{ij}$$

を仮定する. 当然 $p_{1j} + \dots + p_{kj} = 1$ となりたつ. このとき (2.1), (2.2) に従うモデルをマルコフ

切り換え自己回帰モデル (Markov Switching AutoRegressive model, MSAR モデル) と呼ぶことにする。MSAR モデルは、切り換えが観測する対象自体の内生的なメカニズムではなく、他の要因によって生じる場合を考えその切り換えをマルコフ推移確率でモデル化している。

通常のマルコフ切り換えモデルは時系列 y_n の値に依存して切り換えが生じる構造のものが多いが MSAR モデルでは、時系列の値とは独立に切り換えが生じるという単純な構造になっており、 y_n だけに注目すればむしろランダム切り換えというべきかもしれない。

2.2 MSAR モデルの性質

MSAR モデルは切り換えが時系列の値と独立に生じるために各モードの確率は簡単に求めることができる。まず、モードが $s_n=i$ となる確率を q_{in} と表すことにすると

$$(2.3) \quad \begin{bmatrix} q_{1n} \\ \vdots \\ q_{kn} \end{bmatrix} = \begin{bmatrix} p_{11} & \cdots & p_{1k} \\ \vdots & \ddots & \vdots \\ p_{k1} & \cdots & p_{kk} \end{bmatrix} \begin{bmatrix} q_{1,n-1} \\ \vdots \\ q_{k,n-1} \end{bmatrix}$$

がなりたつ。ただし、 $q_{1j} + \cdots + q_{kj} = 1$ 。したがって、モードの定常確率 $\Pr(s_n=i) = q_i$ を求めるためには $q_i = q_{in} = q_{i,n-1}$ とおいて上式を解けばよい。

また、多重推移確率 $\Pr(s_n=i | s_{n-l}=j) = p_{ij}^{(l)}$ についても

$$(2.4) \quad \begin{aligned} p_{ij}^{(1)} &= p_{ij} \\ p_{ij}^{(l)} &= \sum_{s=1}^l p_{is} p_{sj}^{(l-1)}, \quad l=2, 3, \dots \end{aligned}$$

により簡単に求めることができる。さらに、逆の推移確率 $\beta_{ij}^{(l)} = \Pr(s_{n-l}=i | s_n=j)$ も

$$(2.5) \quad \Pr(s_{n-l}=i | s_n=j) = \frac{\Pr(s_n=j | s_{n-l}=i) \Pr(s_{n-l}=i)}{\Pr(s_n=j)} = \frac{p_{ji}^{(l)} q_i}{q_j}$$

により求めることができる。

次に時系列 y_n の性質について考えてみる。以下の議論は一般の場合についても同様にできるが、式がかなり複雑になるので、AR 次数がすべて 1 ($m_j=1$) の場合の結果だけを示しておく。まず明らかに平均値は 0 となる。次に条件付きの自己共分散関数を

$$(2.6) \quad C_l^{(j)} = E[y_n y_{n-l} | s_n=j], \quad j=1, \dots, k; \quad l=0, 1, \dots$$

と定義する。すなわち、 $C_l^{(j)}$ は $s_n=j$ という条件の下でのラグ l の自己共分散である。 $C_l^{(j)}$ については次の式が成り立つ。

$$(2.7) \quad \begin{aligned} C_0^{(j)} &= E[y_n^2 | s_n=j] = E[(a_{1j} y_{n-1} + b_j \varepsilon_n) y_n | s_n=j] \\ &= a_{1j} C_1^{(j)} + b_j^2 \sigma^2 \end{aligned}$$

$$(2.8) \quad \begin{aligned} C_l^{(j)} &= E[y_n y_{n-l} | s_n=j] = E[(a_{1j} y_{n-1} + b_j \varepsilon_n) y_{n-l} | s_n=j] \\ &= a_{1j} E[y_{n-1} y_{n-l} | s_n=j] \\ &= a_{1j} \sum_{i=1}^k \frac{p_{ji} q_i}{q_j} E[y_{n-1} y_{n-l} | s_{n-1}=i] \\ &= a_{1j} \sum_{i=1}^k \frac{p_{ji} q_i}{q_j} C_{l-1}^{(i)} \end{aligned}$$

これを MSAR モデルの Yule-Walker 方程式と呼ぶことにする。 a_{1j} , p_{ji} , b_j ($i, j=1, \dots, k$), σ^2 が与えられると、この方程式を解くことにより、 $C_l^{(j)}$ を求めることができる。

実際の時系列 y_n の自己共分散関数 C_l は状態 j の定常確率 q_j を用いて

$$(2.9) \quad C_t = \sum_{j=1}^k q_j C_t^{(j)}$$

によって求めることができる。

図1に $k=2$,

$$(2.10) \quad y_n = 0.8y_{n-1} + \varepsilon_n, \quad s_n=1 \text{ のとき}$$

$$(2.11) \quad y_n = -0.8y_{n-1} + \varepsilon_n, \quad s_n=2 \text{ のとき}$$

$\varepsilon_n \sim N(0, 1)$ の場合の実現値の例と自己相関関数 $R_t = C_t/C_0$ を示す。(a) と (b) はそれぞれ $p_{11}=1.0, p_{22}=0.0$ および $p_{11}=0.0, p_{22}=1.0$ の場合を示す。すなわち、(a) は (2.10), (b) は (2.11) に従う通常の AR モデルの場合である。

一方、(c) は $p_{11}=0.5, p_{22}=0.5$ の場合、(d) は $p_{11}=0.95, p_{22}=0.95$ の場合である。いずれの

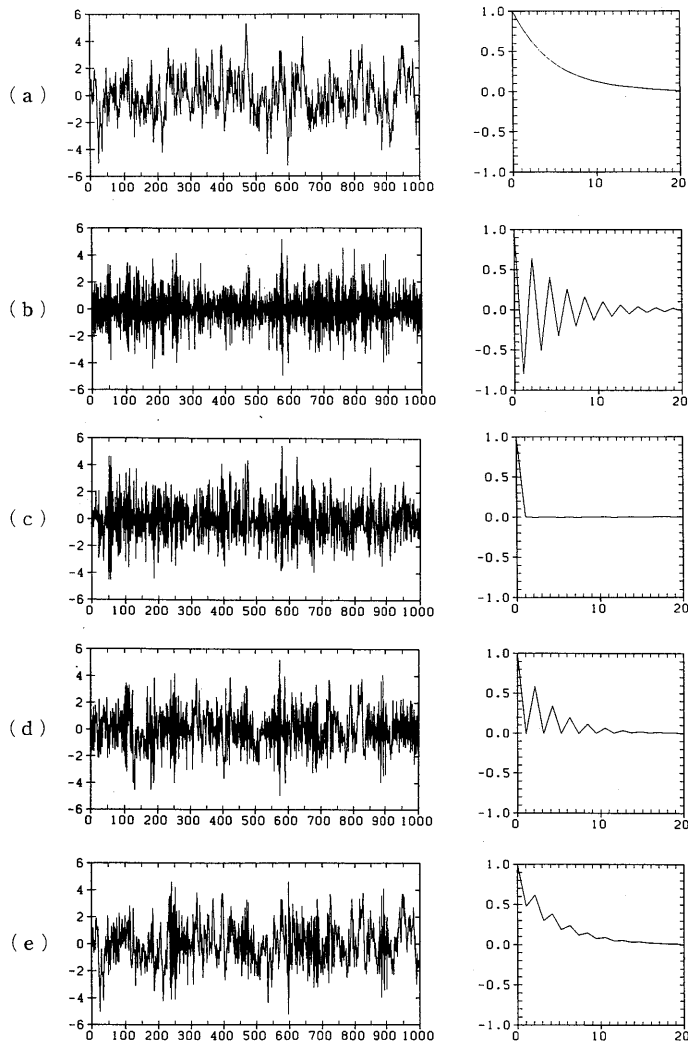


図1. MSAR モデルのシミュレーションによる実現値と自己相関関数。

場合も各モードの定常確率は $q_1 = q_2 = 0.5$ であるが、自己相関関数は全く異なっている。(c) のときには MSAR モデルによって生成される時系列は無相関となる。これは、複数の自己回帰過程のマルコフ切り換えにより白色雑音が生成できるというおもしろい結果を意味している。一方、(d) から推察されるように $p_{11} = p_{22} \rightarrow 1$ の場合には自己相関関数は (2.10) と (2.11) の二つのモデルの自己相関関数の平均となる。(e) は $p_{11} = 0.95, p_{22} = 0.80$ の場合である。このとき、 $q_1 = 0.8, q_2 = 0.2$ となり、自己相関関数も (a) と (b) の q_1 と q_2 による加重平均に近い。

3. 状態推定とパラメータ推定

ここで、 s_n を状態と考えると (2.2) をシステムモデル、(2.1) を観測モデルとみなすことにより離散状態-状態空間モデルが得られる。

時刻 n までの観測値を $Y_n = \{y_1, \dots, y_n\}$ と表すことにする。このとき、状態 s_n の予測分布およびフィルタは非ガウス型フィルタ (Kitagawa (1987)) と同様に以下の式により求めることができる。

[予測]

$$\begin{aligned}
 (3.1) \quad \Pr(s_n = i | Y_{n-1}) &= \sum_{j=1}^k \Pr(s_n = i, s_{n-1} = j | Y_{n-1}) \\
 &= \sum_{j=1}^k \Pr(s_n = i | s_{n-1} = j) \Pr(s_{n-1} = j | Y_{n-1}) \\
 &= \sum_{j=1}^k p_{ij} \Pr(s_{n-1} = j | Y_{n-1})
 \end{aligned}$$

[フィルタ]

$$\begin{aligned}
 (3.2) \quad \Pr(s_n = i | Y_n) &= \Pr(s_n = i | Y_{n-1}, y_n) \\
 &= \frac{p(y_n | s_n = i, Y_{n-1}) \Pr(s_n = i | Y_{n-1})}{p(y_n | Y_{n-1})}
 \end{aligned}$$

ここで $p(y_n | s_n = i, Y_{n-1})$ と $p(y_n | Y_{n-1})$ はそれぞれ

$$\begin{aligned}
 (3.3) \quad p(y_n | s_n = i, Y_{n-1}) &= \frac{1}{\sqrt{2\pi b_i^2 \sigma^2}} \exp\left\{-\frac{1}{2b_i^2 \sigma^2} \left(y_n - \sum_{j=1}^{m_i} a_{ji} y_{n-j}\right)^2\right\}, \\
 p(y_n | Y_{n-1}) &= \sum_{j=1}^k p(y_n | s_n = j, Y_{n-1}) \Pr(s_n = j | Y_{n-1})
 \end{aligned}$$

で与えられる。

同様に平滑化の公式も

$$\begin{aligned}
 (3.4) \quad \Pr(s_n = i, s_{n-1} = j | Y_n) \\
 = \Pr(s_n = i | Y_n) \frac{\Pr(s_{n+1} = j | Y_n) \Pr(s_{n+1} = j | s_n = i)}{\Pr(s_{n+1} = j | Y_n)}
 \end{aligned}$$

がなりたつので

$$\begin{aligned}
 (3.5) \quad \Pr(s_n = i | Y_n) \\
 = \Pr(s_n = i | Y_n) \sum_{j=1}^k \frac{\Pr(s_{n+1} = j | Y_n) \Pr(s_{n+1} = j | s_n = i)}{\Pr(s_{n+1} = j | Y_n)}
 \end{aligned}$$

により簡単に与えられる。

一般の連続状態-状態空間モデルに非ガウス型フィルタおよび平滑化の公式を適用するため

には、状態空間を多数の点で表現し、積分を和で近似する必要がある。しかし、離散状態-状態空間モデルの場合には状態空間は k 個の状態 $\{1, \dots, k\}$ だけで完全に表現でき、フィルタおよび平滑化後の確率も (3.1), (3.2), (3.5) のように簡単に計算できる。

k 状態 MSAR モデルにはパラメータ $\theta = \{\sigma^2, p_{ij}, a_{ij}, i, j = 1, \dots, k; l = 1, \dots, m_j\}$ が含まれている。ただし、 $\sum_{i=1}^k p_{ij} = 1$ が必要だが、これまでは、これらは与えられたものとしていたが、実際にはこれらは未知である。 Y_n が与えられたとき θ の対数尤度は

$$(3.6) \quad \ell(\theta | y) = \sum_{n=1}^N \log p(y_n | Y_{n-1})$$

で求められる。ここで、 $p(y_n | Y_{n-1})$ は (3.3) 式で与えられ、また簡単のために $y_0, \dots, y_{1-\max(m_1, \dots, m_k)}$ は与えられているものとする。パラメータ θ の最尤推定値はこの対数尤度を最大化することによって求めることができる。

4. 例

本章では例として 4 状態のマルコフ切り換え AR モデルのパラメータ推定と状態推定を考える。

$P = (q_{ij})$ が次のように与えられているものとする。

$$(4.1) \quad P = \begin{bmatrix} 1-3\alpha & \beta & \beta & \beta \\ \alpha & 1-\beta-2\gamma & \gamma & \gamma \\ \alpha & \gamma & 1-\beta-2\gamma & \gamma \\ \alpha & \gamma & \gamma & 1-\beta-2\gamma \end{bmatrix}$$

このとき、各モードの定常確率は $q_1 = \beta/(\beta+3\alpha)$, $q_2 = q_3 = q_4 = \alpha/(\beta+3\alpha)$ である。

また、各時点 n では時系列 y_n はモード s_n に依存して以下のように 2 次の AR モデルにしたがって生成されるものとする。

$$(4.2) \quad y_n = \begin{cases} 1.785y_{n-1} - 0.903y_{n-2} + \varepsilon_n, & s_n = 1 \text{ のとき} \\ 1.344y_{n-1} - 0.903y_{n-2} + \varepsilon_n, & s_n = 2 \text{ のとき} \\ 1.386y_{n-1} - 0.640y_{n-2} + \varepsilon_n, & s_n = 3 \text{ のとき} \\ 0.800y_{n-1} - 0.640y_{n-2} + \varepsilon_n, & s_n = 4 \text{ のとき} \end{cases}$$

ただし、 ε_n は平均 0、分散 1 の正規白色雑音とする。これらの AR モデルの固有根はそれぞれ以下のようになっている。

$$(4.3) \quad \begin{array}{lll} r=0.95, & \phi=20^\circ, & s_n=1 \text{ のとき} \\ r=0.95, & \phi=45^\circ, & s_n=2 \text{ のとき} \\ r=0.80, & \phi=30^\circ, & s_n=3 \text{ のとき} \\ r=0.80, & \phi=60^\circ, & s_n=4 \text{ のとき} \end{array}$$

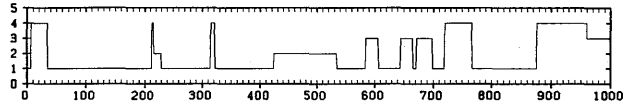
図 2 に $\alpha=0.0033$, $\beta=0.016$, $\gamma=0.002$ の場合の 4 状態 MSAR モデルの実現値 ($N=1000$) を示す。(a) はモード s_n , (b) は時系列 y_n である。

パラメータの最尤推定値は $\hat{\sigma}^2=0.9888$, $\hat{\alpha}=0.005614$, $\hat{\beta}=0.02267$, $\hat{\gamma}=0.132 \times 10^{-5}$, また各モードの AR モデルの係数は表 1 のようであった。

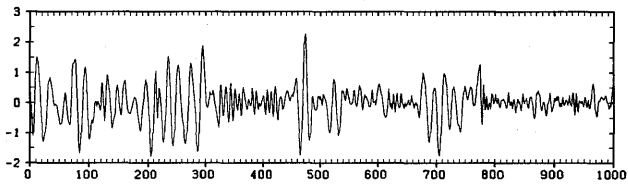
図 3 にこれらのパラメータを用いてフィルタおよび平滑化により求めた各モードの事後確率を示す。図 2(a) に示された各モードをかなりよく推定していることがわかる。

表1.

	AR ₁	AR ₂	AR ₃	AR ₄
a_{i1}	1.8095	1.2848	1.2848	0.8219
a_{i2}	-0.9272	-0.8866	-0.5970	-0.6387

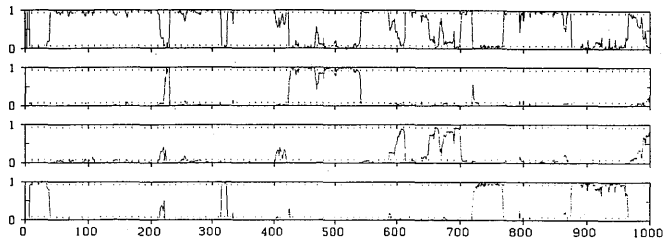


(a)

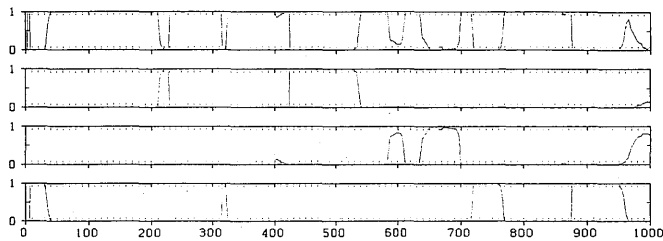


(b)

図2. 4状態MSARモデルの実現値。(a)モード, (b)時系列.



(a)



(b)

図3. MSARモデルの各モードの事後確率。(a)フィルタ, (b)平滑化.

5. あとがき

マルコフチェーンに従ってモードが切り換わる自己回帰モデル (MSAR モデル) を考え、その簡単な性質と非ガウス型フィルタにもとづくパラメータおよびモードの推定法を示した。

前章では最尤法により極めてよい再現性が得られることを示したが、モード 2, 3, 4 に関しては交換可能であることを考えるとむしろ意外な結果ともいえる。上記の推定結果は AR 係数に関しては真の値を初期値として非線形最適化を行った場合であるが、それが良い結果を得られた原因ではない。乱数を用いて初期値を変動させてもモードの順番が換わることはあってもほぼ同様の推定値が得られた。

4 つの AR モデルの固有根が前章の場合より接近している場合には、事後確率は図 3 のように常に 0 または 1 に近いとは限らず中間的な値が得られることもある。しかし、むしろこれは当然の結果と考えられる。

本稿では AR 次数は既知としたが実際にはデータにもとづいて決定すべきものである。AIC を用いて決定した場合、どの程度よいモデルが得られるかは慎重に検討してみる必要がある。より本質的な問題としてはモードの数の決定がある。モードの数 k に関しても原理的には AIC による推定が可能であるが、かなり困難な問題であることが予想される。

参 考 文 献

- Churchill, G.A. (1989). Stochastic models for heterogeneous DNA sequences, *Bull. Math. Biol.*, **51**, 4451-4460.
- Goldfeld, S.M. and Quandt, R.E. (1973). A Markov model for switching regression, *J. Econometrics*, **1**, 3-16.
- Gorden, K. and Smith, A.F.M. (1988). Modeling and monitoring discontinuous changes in time series, *Bayesian Analysis of Time Series and Dynamic Linear Models* (ed. J.C. Spall), 359-392, Marcel Dekker, New York.
- Katzoff, M.J. and Shumway, R.H. (1993). Nonlinear structural models for mortality series, *Proceedings of the 2nd U.S.-Japan Joint Seminar on Statistical Time Series Analysis*, January 25-29, 1993, Hawaii, 319-348.
- Kitagawa, G. (1987). Non-Gaussian state-space modeling of nonstationary time series, *J. Amer. Statist. Assoc.*, **82**, 1032-1063.
- Kitagawa, G. (1988). Numerical approach to non-Gaussian smoothing and its applications, *Computing Science and Statistics, Proceedings of the 20th Symposium on the Interface* (eds. E.J. Wegman, D.T. Gantz and J.J. Miller), Fairfax, Virginia, April 1988, 379-388.
- Komaki, F. (1993). State-space modeling of time series sampled from continuous processes with pulses, *Biometrika*, **80**, 417-429.
- McCulloch, R.E. and Tsay, R.S. (1993). Statistical analysis of macroeconomic time series via Markov switching models, *Proceedings of the 2nd U.S.-Japan Joint Seminar on Statistical Time Series Analysis*, January 25-29, 1993, Hawaii, 195-230.
- Shumway, R.H. and Stoffer, D.S. (1991). Dynamic linear models with switching, *J. Amer. Statist. Assoc.*, **86**, 763-769.
- Tong, H. (1990). *Nonlinear Time Series Analysis: A Dynamical System Approach*, Oxford University Press, London.

Markov Switching Autoregressive Model

Genshiro Kitagawa

(The Institute of Statistical Mathematics)

For the analysis of time series that consists of several different patterns, we consider a Markov switching autoregressive (MSAR) model. In this modeling an autoregressive model is selected depending on the state of the hidden Markov chain. Some properties of the MSAR model are shown and non-Gaussian filtering and smoothing formulae are shown. These formulae can be applied to the estimation of the parameters of the MSAR model and to the computation of the posterior probability of each state. Some numerical results are shown to illustrate the characteristics of the MSAR process and the estimation of the posterior probability.