ARMA 過程のパラメータ推定における RML 法の収束速度の改善†

田 淵 敏 明*・添 田 喬**

Improvement of Convergence Rate of RML Method in Parameter
Estimation of ARMA Process

Toshiaki TABUCHI* and Takashi SOEDA**

This paper is concerned with the improvement of convergence rate of RML (Recursive Maximum Likelihood) method in ARMA parameter estimation. A recursive parameter estimator is developed by using Gauss-Newton method. Main difference the proposed method with RML one is the estimation of the expectation of Hessian in Gauss-Newton method. Proposed method estimates this quantity so as to minimize the estimation error of Fisher's information matrix which is considered as the information processing ability of RML estimator. Resultant algorithm is constructed by linking RML estimator with extended Kalman filter. The feature of proposed method is that the stability monitoring is not required in intermediate computation.

It is shown in numerical experiment that the convergence rate of the parameter estimates is improved by using proposed method.

Key Words: ARMA process, parameter estimation, RML method, Fisher's information matrix, extended Kalman filter

1. ま え が き

ARMA 過程のパラメータ推定手法の代表的なものとして、逐次最尤推定法 (RML 法)¹⁾がある。これは漸近的に不偏、一致、有効推定量となるというすぐれた性質をもっている。RML 法は、未知パラメータに

関する最小化問題として扱われ、このとき最小化の方法として普通は Gauss-Newton 法²⁾が用いられる。したがって Hessian の期待値を決定することが重要になってくるが、ARMA 過程の場合、この量は未知であるのでなんらかの方法で推定しなければならない。このとき、Hessian の期待値は Fisher の情報行列と密接に関連しているので、この推定がパラメータの推定精度に大きな影響を及ぼすと考えられる。

本論文は、RML 推定器の情報処理能力の一つの指標として Fisher の情報行列に注目し、Hessian の期待値をこの情報行列の推定誤差ができるだけ小さくなるように推定することにより、パラメータ推定値の精度の向上を図り、収束速度を加速しようとするものである。このようにして得られる Hessian の期待値の推定値は、ARMA 過程の状態空間表現における状態量の平滑推定値に関係する。これを評価する場合、平滑器が必要になり、アルゴリズムが複雑になる。そこで複雑化をさけるために平滑器を使用しないように近似がなされる。その結果、本手法は Hessian の期待値の推定に拡張 Kalman フィルタを用いる形となり、最終的なアルゴリズムは RML 推定器と拡張 Kalman フィルタを連動させる形となる。したがって RML 法に比べ、拡張 Kalman フィルタの分だけ計算量は増えるが、本数値実験においては収束速度は大幅に改善されることが確認される。

2. 問題の記述および本手法の概略

定常 ARMA 過程

$$A(q^{-1})y(t) = B(q^{-1})e(t) \quad (1)$$

$$A(q^{-1}) = 1 - a_1q^{-1} - \dots - a_nq^{-n},$$

$$B(q^{-1}) = 1 + b_1q^{-1} + \dots + b_nq^{-n} \quad (2)$$

において、未知パラメータ $\alpha = [a_1 \dots a_n b_1 \dots b_n]^T$ を推

† 第 5 回適応制御シンポジウムで一部発表 (昭 60・1)

* 徳島工業短期大学 徳島県板野郡板野町

** 徳島大学本部 徳島市新蔵町 2-24

* Tokushima Technical College, Itano-gun, Tokushima

** The Head Office, University of Tokushima, Tokushima

(Received July 24, 1985)

(Revised April 25, 1986)

定することを考える。上式で q^{-1} は単位時間遅延演算子, $e(t)$ は $N[0, \sigma^2]$ に従う白色雑音, $y(t)$ は出力データである。また $A(q^{-1})$, $B(q^{-1})$ はそれぞれ定常条件, 可逆条件²⁾ を満足するものとする。このとき RML 法で α を推定する場合には, 評価関数

$$V_i(\alpha) = \frac{1}{2} \sum_{i=1}^t e^2(i) \quad (3)$$

の α に関する最小化が Gauss-Newton 法を用いて行われ, α の推定値 $\hat{\alpha}(t)$ は次式で与えられる⁷⁾。

$$\hat{\alpha}(t) = \hat{\alpha}(t-1) - [E\{V_i''(\hat{\alpha}(t-1))\}]^{-1} \cdot V_i'(\hat{\alpha}(t-1)) \quad (4)$$

ここで E は期待値を表わし, また $V_i'(\hat{\alpha}) = [\partial V_i(\alpha) / \partial \alpha]_{\hat{\alpha}}$, $V_i''(\hat{\alpha}) = [\partial^2 V_i(\alpha) / \partial \alpha \partial \alpha^T]_{\hat{\alpha}}$ であり, これらはつぎのようになる。

$$V_i'(\alpha) = - \sum_{i=1}^t \phi(i)e(i) \quad (5)$$

$$= V_{i-1}'(\alpha) - \phi(t)e(t) \quad (6)$$

$$\begin{aligned} \phi^T(t) &= - \frac{\partial e(t)}{\partial \alpha^T} \\ &= \frac{\phi^T(t)}{B(q^{-1})} \end{aligned} \quad (7)$$

$$V_i''(\alpha) \approx \sum_{i=1}^t \phi(i)(i)\phi^T(i)^T \quad (8)$$

$$\phi^T(t) = [y^T(t-1) : e^T(t-1)] \quad (9)$$

$$y^T(t-1) = [y(t-1) \cdots y(t-n)] \quad (10)$$

$$e^T(t-1) = [e(t-1) \cdots e(t-n)] \quad (10)$$

(4)において $E\{V_i'(\hat{\alpha}(t-1))\}$ は実現不可能であり, また $V_i'(\hat{\alpha}(t-1))$, $V_i''(\hat{\alpha}(t-1))$ は逐次計算ができない。そこで普通は, α の running estimate $\hat{\alpha}_r(t-1) \triangleq \{\hat{\alpha}(t-1), \dots, \hat{\alpha}(0)\}$ を用いて, (4)をつぎのように近似したものが用いられる。

$$\hat{\alpha}(t) = \hat{\alpha}(t-1) - [\hat{V}_i''(\hat{\alpha}_r(t-1))]^{-1} \cdot V_i'(\hat{\alpha}_r(t-1)) \quad (11)$$

ここで $\hat{V}_i''(\hat{\alpha}_r(t-1))$ は $E\{V_i''(\hat{\alpha}(t-1))\}$ の一つの推定値である。 $\hat{\alpha}(t)$ の収束を加速させるためには, $\hat{V}_i''(\hat{\alpha}_r(t-1))$ の選定が重要になってくるが, これを行うために $\hat{\alpha}(t)$ が α の近傍にあることを仮定した場合に成立するつぎの評価式⁸⁾を用いる。

$$V_i(\hat{\alpha}(t)) \leq \left[\frac{\lambda_{\max}(t-1) - \lambda_{\min}(t-1)}{\lambda_{\max}(t-1) + \lambda_{\min}(t-1)} \right]^2 V_{i-1}(\hat{\alpha}(t-1)) \quad (12)$$

ここで $\lambda_{\max}(t)$, $\lambda_{\min}(t)$ はそれぞれ $[\hat{V}_i''(\hat{\alpha}_r(t-1))]^{-1} \cdot V_i''(\alpha)$ の最大, 最小の固有値である。したがって $V_i''(\alpha) - \hat{V}_i''(\hat{\alpha}_r(t-1))$ が0に近くなるように $V_i''(\hat{\alpha}_r(t-1))$ を決定すれば, 収束速度は速くなる⁸⁾。しかし, $V_i''(\alpha) - \hat{V}_i''(\hat{\alpha}_r(t-1))$ はこのままでは扱えないので, この期待値をとったもの, すなわち

$$E\{\hat{V}_i''(\alpha)\} \triangleq E\{V_i''(\alpha) - \hat{V}_i''(\alpha)\} \quad (13)$$

を最小にするように $\hat{V}_i''(\alpha)$ を決定し, しかる後に α を $\hat{\alpha}_r(t-1)$ で置き換えたものを $\hat{V}_i''(\hat{\alpha}_r(t-1))$ とする。したがって全ての見本系列を考慮した場合の, 平均的な収束速度の加速を試みるものである。そこでまず $\hat{V}_i''(\alpha)$ を決定するために, $E\{V_i''(\alpha)\}$ をつぎのように書く。

$$E\{V_i''(\alpha)\} = \sum_{i=1}^t E\{\phi(i, \alpha)\phi^T(i, \alpha)\} \quad (\text{注1}), (\text{注2}) \quad (14)$$

$$\begin{aligned} &= \sum_{i=1}^t E\{E\{\phi(i, \alpha)\phi^T(i, \alpha) | y^{t-1}\}\} \\ &= \sum_{i=1}^t E\{\hat{\phi}(i, \alpha)\hat{\phi}^T(i, \alpha)\} \\ &\quad + \sum_{i=1}^t E\{\tilde{\phi}(i, \alpha)\tilde{\phi}^T(i, \alpha)\} \end{aligned} \quad (15)$$

ここで

$$\hat{\phi}(t, \alpha) = E\{\phi(t, \alpha) | y^{t-1}\} \quad (16)$$

であり, $\tilde{\phi}(t, \alpha) = \phi(t, \alpha) - \hat{\phi}(t, \alpha)$, $y^{t-1} = \{y(t-1), \dots, y(0)\}$ である。 $\hat{\phi}(t, \alpha)$ として(16)の条件付期待値を用いるとき, $E\{\tilde{\phi}(t, \alpha)\tilde{\phi}^T(t, \alpha)\}$ は最小になり⁹⁾, その結果(15)の右辺第2項全体も最小になる。そこで本手法では(15)を参考に

$$\hat{V}_i''(\alpha) = \sum_{i=1}^t \hat{\phi}(i, \alpha)\hat{\phi}^T(i, \alpha) \quad (17)$$

とする。このとき(13)は

$$E\{\hat{V}_i''(\alpha)\} = \sum_{i=1}^t E\{\tilde{\phi}(i, \alpha)\tilde{\phi}^T(i, \alpha)\} \quad (18)$$

となり最小になる。実際には(16), (17)は, α を running estimate で置き換えた次式を用いる。

$$\hat{\phi}(t, \hat{\alpha}_r(t-1)) = E\{\phi(t, \hat{\alpha}_r(t-1)) | y^{t-1}\} \quad (19)$$

$$\hat{V}_i''(\hat{\alpha}_r(t-1)) \triangleq \sum_{i=1}^t \hat{\phi}(i, \hat{\alpha}_r(i-1))\hat{\phi}^T(i, \hat{\alpha}_r(i-1)) \quad (20)$$

もし(16), (17)を実際に用いることができれば, 収束速度は加速されるが, これらの代わりに(19), (20)を用いるので収束速度の加速については期待事項となる。

なお, 実際のアルゴリズムは, $V_{i-1}'(\hat{\alpha}_r(t-1)) = 0$ を仮定して(6)より得られる $V_i'(\hat{\alpha}_r(t-1)) = -\hat{\phi}(t, \hat{\alpha}_r(t-1))e(t, \hat{\alpha}_r(t-1))$ および(20)より, 次式で与えられる。

(注1) $\phi(t, \hat{\alpha})$ は $\phi(t)$ を生成しているパラメータベクトル α を $\hat{\alpha}$ に置き換えた場合に得られるものである。したがって $\phi(t) = \phi(t, \alpha)$ と書ける。

(注2) $V_i''(\alpha)$ は(8)において左辺と右辺は等号で結ばれないが, $E\{V_i''(\alpha)\}$ については左辺と右辺は等号で結ばれる⁹⁾。

$$\left. \begin{aligned} \hat{\alpha}(t) &= \hat{\alpha}(t-1) + g(t)\varepsilon(t) \\ g(t) &= P(t-1)\hat{\phi}(t)/[\hat{\phi}^T(t)P(t-1)\hat{\phi}(t)+1] \\ P(t) &\triangleq [\hat{V}'(\hat{\alpha}_r(t-1))]^{-1} \\ &= [I - g(t)\hat{\phi}^T(t)]P(t-1) \end{aligned} \right\} \quad (21)$$

上式において $\hat{\phi}(t) \triangleq \hat{\phi}(t, \hat{\alpha}_r(t-1))$, $\varepsilon(t) \triangleq e(t, \hat{\alpha}_r(t-1))$ である。

次章では, $\hat{\phi}(t, \hat{\alpha}_r(t-1))$ を評価する。

3. $\hat{\phi}(t, \hat{\alpha}_r(t-1))$ の評価

(19)をもとに $\hat{\phi}(t, \hat{\alpha}_r(t-1))$ を評価するために, まず $\phi(t) = \phi(t, \alpha)$ を評価し, つぎに $\phi(t, \hat{\alpha}_r(t-1))$, $\hat{\phi}(t, \hat{\alpha}_r(t-1))$ を評価する。そこで(7)をつぎのように変形する。(7)の両辺に $2n \times 1$ ベクトル β をかけてスカラー化すると $\phi^T(t)\beta = \hat{\phi}^T(t)\beta/B(q^{-1})$ となり, これを状態空間で表わすと次式のようになる。

$$\left. \begin{aligned} z(t) &= B^T z(t-1) + c\phi^T(t)\beta \\ \phi^T(t)\beta &= c^T z(t) \end{aligned} \right\} \quad (22)$$

ここで

$$B^T = \begin{bmatrix} -b_1 & 1 & 0 & \cdots & 0 \\ -b_2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -b_{n-1} & 0 & 0 & \cdots & 1 \\ -b_n & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad c = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \quad (23)$$

である。(22)を convolution sum の形に書いて, B^T の安定性を考慮するとき, $t \gg 1$ で

$$\phi^T(t)\beta = c^T \sum_{i=1}^t B^{T^{t-i}} \phi^T(i)\beta \quad (24)$$

となり, 両辺を比較して(9)を用いると

$$\begin{aligned} \phi^T(t) &= c^T \sum_{i=1}^t B^{T^{t-i}} c \phi^T(i) \\ &= c^T \sum_{i=1}^t B^{T^{t-i}} c [y^T(i-1) : e^T(i-1)] \end{aligned} \quad (25)$$

となる。したがって

$$\begin{aligned} \phi(t, \hat{\alpha}_r(t-1)) &= e^T \sum_{i=1}^t \Phi(t, i) c [y^T(i-1) : e^T(i-1)] \\ \Phi(t, i) &\triangleq \hat{B}^T(t-1) \cdots \hat{B}^T(i), \quad \Phi(i, i) = I \end{aligned} \quad (26)$$

となる。このとき(26)の条件付期待値をとれば, $\hat{\phi}(t, \hat{\alpha}_r(t-1))$ が得られるが, $E\{e^T(i-1)|y^{t-1}\}$ が関係してくる。この評価を行いやすくするために, (1)をつぎのような状態空間で表現する。

$$\left. \begin{aligned} x(t) &= Bx(t-1) + cd^T y(t-1) \\ y(t) &= c^T x(t) + e(t) \end{aligned} \right\} \quad (28)$$

ここで

$$\begin{aligned} x(t) &= y(t) - e(t), \quad d^T = a^T + b^T, \quad a^T = [a_1 \cdots a_n], \\ b^T &= [b_1 \cdots b_n] \end{aligned} \quad (29)$$

である。上式より $e(i-1) = y(i-1) - x(i-1)$ を(26)に代入するとき

$$\begin{aligned} \phi^T(t, \hat{\alpha}_r(t-1)) &= c^T \sum_{i=1}^t \Phi(t, i) c [y^T(i-1) : \\ &\quad y^T(i-1) - x^T(i-1)] \end{aligned} \quad (30)$$

となり, したがって

$$\begin{aligned} \hat{\phi}^T(t, \hat{\alpha}_r(t-1)) &= c^T \sum_{i=1}^t \Phi(t, i) c [y^T(i-1) : \\ &\quad y^T(i-1) - \hat{x}^T(i-1|t-1)] \\ &= [c^T S_1(t) : c^T \{S_1(t) - S_2(t)\}] \end{aligned} \quad (31)$$

となる。ここで

$$\hat{x}(i-1|t-1) = E\{x(i-1)|y^{t-1}\} \quad (32)$$

$$\begin{aligned} S_1(t) &= \sum_{i=1}^t \Phi(t, i) c y^T(i-1) \\ &= \hat{B}^T(t-1) S_1(t-1) + c y^T(t-1), \quad S_1(0) = 0 \end{aligned} \quad (33)$$

$$S_2(t) = \sum_{i=1}^t \Phi(t, i) c \hat{x}^T(i-1|t-1) \quad (34)$$

である。したがって $E\{e^T(i-1)|y^{t-1}\}$ を評価する代わりに, $\hat{x}(i-1|t-1)$ を評価すればよい。これは(28)をもとに, B, σ^2 が既知であれば, 固定点平滑器を用いて $S_2(t)$ の繰返し関係を得ることができる。しかしその形は複雑になり, また固定点平滑器は漸近安定ではないので⁴⁾, B, σ^2 をその推定値で代用することを考えると, 固定点平滑器の使用はあまり好ましくない。そこで $\hat{x}(i-1|t-1) \approx \hat{x}(i-1|i-1)$ で近似するとき, $S_2(t)$ は次式で与えられる。

$$S_2(t) \approx \sum_{i=1}^t \Phi(t, i) c \hat{x}^T(i-1|i-1) \quad (35)$$

$$\begin{aligned} &= \hat{B}^T(t-1) S_2(t-1) + c \hat{x}^T(t-1|t-1), \\ S_2(0) &= 0 \end{aligned} \quad (36)$$

次章では, $\hat{x}(t|t)$ を評価する。

4. $\hat{x}(t|t)$ の評価

$\hat{x}(t|t)$ は, (28)のパラメータがすべて既知であれば, Kalman フィルタを用いればよいが, これらは未知であるので, 本手法では拡張 Kalman フィルタを用いる。

ところで, (28)に拡張 Kalman フィルタを用いる場合, 近似共分散行列 $Q_{xx}, Q_{xa}, Q_{ax}, Q_{aa}$ が関係するが, Q_{xa}, Q_{ax} については無視したほうがフィルタの性能が向上することが報告されている⁵⁾。そこで本手法もこれら無視する。その結果 $\hat{x}(t|t)$ は次式で与えられる。

$$\left. \begin{aligned} \hat{x}(t|t) &= \hat{x}(t|t-1) + k_x(t)\varepsilon(t) \\ \hat{x}(t|t-1) &= \hat{B}(t-1)\hat{x}(t-1|t-1) \\ &\quad + cd^T(t-1)y(t-1) \\ \varepsilon(t) &= y(t) - c^T \hat{x}(t|t-1) \\ k_x(t) &= Q_x(t|t-1)c/[c^T Q_x(t|t-1)c] \end{aligned} \right\}$$

$$\begin{aligned}
 & + \hat{\sigma}^2(t-1)] \\
 & \left. \begin{aligned}
 Q_x(t|t-1) &= \hat{B}(t-1)Q_x(t-1|t-1) \\
 & \times \hat{B}^T(t-1) + \hat{\phi}^T(t)Q_a(t-1)\hat{\phi}(t)cc^T \\
 Q_x(t|t) &= [I - k_x(t)c^T]Q_x(t|t-1) \\
 Q_a(t) &= [I - k_a(t)\hat{\phi}^T(t)]Q_a(t-1) \\
 k_a(t) &= Q_a(t-1)\hat{\phi}(t)/[c^T Q_x(t|t-1)c \\
 & + \hat{\sigma}^2(t-1)]
 \end{aligned} \right\} (37)
 \end{aligned}$$

上式で I は単位行列であり Q_x, Q_a はそれぞれ Q_{xx}, Q_{aa} のことである。また $Q_a(0) = P(0)$ に選ぶものとする。つぎに σ^2 の推定値 $\hat{\sigma}^2(t)$ は次式で評価する。

$$\hat{\sigma}^2(t) = \frac{1}{t} \sum_{i=1}^t E\{e^2(i)|y^i\}$$

このとき $\hat{\sigma}^2(t)$ は不偏推定量となり、Kalman フィルタをもとにつぎの繰返し関係が得られる。

$$\begin{aligned}
 \hat{\sigma}^2(t) &= \left(1 - \frac{1}{t}\right)\hat{\sigma}^2(t-1) \\
 & + \frac{1}{t} [y(t) - c^T \hat{x}(t|t)]^2 + c^T Q_x(t|t)c] \\
 & (38)
 \end{aligned}$$

しかし実際にはパラメータは未知であるので、上式は(37)と連動させて用いる。この結果 $\hat{\sigma}^2(t)$ は不偏推定量ではなくなる。

以上でパラメータ推定に必要な式は導出できたが、 $S_1(t), S_2(t)$ の遷移行列 $\hat{B}^T(t-1)$ が不安定になれば、 $\hat{\phi}(t), \hat{\alpha}(t-1)$ が発散して急速に $P(t) = 0, g(t) = 0$ となるので、 $\hat{\alpha}(t)$ は更新されなくなる。このことは RML 法の場合も起りうる問題であり、 $\hat{B}^T(t-1)$ の安定性についての監視が必要になる。本手法では、 $\hat{B}^T(t-1)$ が不安定になった場合も考えて、 $S_1(t), S_2(t)$ の遷移行列を形式的に

$$F(t-1) = \{[I - k_x(t)c^T]\hat{B}(t-1)\}^T \quad (39)$$

で置き換える。これは $\hat{B}^T(t-1)$ がたとえ不安定になっても、 $\hat{\alpha}(t), \hat{\sigma}^2(t)$ が有限であれば、 $(\hat{B}(t-1), c^T)$ が一様に完全可観測であるとき、(37)のフィルタは漸近安定となり⁶⁾、したがって $F(t-1)$ も安定になる。したがって本手法では、 $\hat{B}^T(t-1)$ の安定性に関する監視は必要がない。

(21)をもとに本手法をまとめると Table 1 のようになる。これが本論文の主要結果である。

5. 本手法と RML 法との比較

まず RML 法は、 $\hat{\alpha}(t)$ を次式で計算する¹⁾。

$$\left. \begin{aligned}
 \hat{\alpha}(t) &= \hat{\alpha}(t-1) + g(t)\hat{\ell}(t) \\
 \hat{\ell}(t) &= y(t) - \hat{\phi}^T(t)\hat{\alpha}(t-1) \\
 g(t) &= P(t-1)\hat{\phi}(t)/ \\
 & [\hat{\phi}^T(t)P(t-1)\hat{\phi}(t) + 1] \\
 P(t) &= [I - g(t)\hat{\phi}^T(t)]P(t-1)
 \end{aligned} \right\} (40)$$

Table 1 Proposed parameter estimation algorithm and RML algorithm

Proposed algorithm:

$$\begin{aligned}
 \hat{\alpha}(t) &= \hat{\alpha}(t-1) + g(t)\hat{\ell}(t) \\
 \hat{\ell}(t) &= y(t) - c^T \hat{x}(t|t-1) \\
 g(t) &= P(t-1)\hat{\phi}(t)/[\hat{\phi}^T(t)P(t-1)\hat{\phi}(t) + 1] \\
 P(t) &= [I - g(t)\hat{\phi}^T(t)]P(t-1) \\
 \hat{\phi}^T(t) &= [c^T S_1(t) : c^T \{S_1(t) - S_2(t)\}] \\
 S_1(t) &= F(t-1)S_1(t-1) + cy^T(t-1), S_1(0) = 0 \\
 S_2(t) &= F(t-1)S_2(t-1) + c\hat{x}^T(t-1|t-1), S_2(0) = 0 \\
 F(t-1) &= \{[I - k_x(t-1)c^T]\hat{B}(t-1)\}^T \\
 \hat{x}(t|t) &= \hat{x}(t|t-1) + k_x(t)\hat{\ell}(t) \\
 \hat{x}(t|t-1) &= \hat{B}(t-1)\hat{x}(t-1|t-1) + c\hat{\alpha}^T(t-1)y(t-1) \\
 k_x(t) &= Q_x(t|t-1)c/[c^T Q_x(t|t-1)c + \hat{\sigma}^2(t-1)] \\
 Q_x(t|t-1) &= \hat{B}(t-1)Q_x(t-1|t-1)\hat{B}^T(t-1) \\
 & + \hat{\phi}^T(t)Q_a(t-1)\hat{\phi}(t)cc^T \\
 Q_x(t|t) &= [I - k_x(t)c^T]Q_x(t|t-1) \\
 Q_a(t) &= [I - k_a(t)\hat{\phi}^T(t)]Q_a(t-1), Q_a(0) = P(0) \\
 k_a(t) &= Q_a(t-1)\hat{\phi}(t)/[c^T Q_x(t|t-1)c + \hat{\sigma}^2(t-1)] \\
 \hat{\sigma}^2(t) &= \left(1 - \frac{1}{t}\right)\hat{\sigma}^2(t-1) + \frac{1}{t} [y(t) - c^T \hat{x}(t|t)]^2 \\
 & + c^T Q_x(t|t)c] \\
 \hat{\phi}^T(t) &= [y^T(t-1) : y^T(t-1) - \hat{x}^T(t-1|t-1)] \\
 \hat{B}(t-1) &= \begin{bmatrix} -\hat{b}_1(t-1) & -\hat{b}_2(t-1) & \cdots & -\hat{b}_{n-1}(t-1) & -\hat{b}_n(t-1) \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}
 \end{aligned}$$

RML algorithm:

$$\begin{aligned}
 \hat{\alpha}(t) &= \hat{\alpha}(t-1) + g(t)\hat{\ell}(t) \\
 \hat{\ell}(t) &= y(t) - c^T \hat{x}(t) \\
 g(t) &= P(t-1)\hat{\phi}(t)/[\hat{\phi}^T(t)P(t-1)\hat{\phi}(t) + 1] \\
 P(t) &= [I - g(t)\hat{\phi}^T(t)]P(t-1) \\
 \hat{\phi}^T(t) &= [c^T S_1(t) : c^T \{S_1(t) - S_2(t)\}] \\
 S_1(t) &= \hat{B}^T(t-1)S_1(t-1) + cy^T(t-1), S_1(0) = 0 \\
 S_2(t) &= \hat{B}^T(t-1)S_2(t-1) + c\hat{x}^T(t-1), S_2(0) = 0 \\
 \hat{x}(t) &= \hat{B}(t-1)\hat{x}(t-1) + c\hat{\alpha}^T(t-1)y(t-1)
 \end{aligned}$$

RML algorithm (based on residual):

$$\begin{aligned}
 \hat{\alpha}(t) &= \hat{\alpha}(t-1) + g(t)\hat{\ell}(t) \\
 \hat{\ell}(t) &= y(t) - c^T \hat{x}(t) \\
 g(t) &= P(t-1)\hat{\phi}(t)/[\hat{\phi}^T(t)P(t-1)\hat{\phi}(t) + 1] \\
 P(t) &= [I - g(t)\hat{\phi}^T(t)]P(t-1) \\
 \hat{\phi}^T(t) &= [c^T S_1(t) : c^T \{S_1(t) - S_2(t)\}] \\
 S_1(t) &= \hat{B}^T(t-1)S_1(t-1) + cy^T(t-1), S_1(0) = 0 \\
 S_2(t) &= \hat{B}^T(t-1)S_2(t-1) + c\hat{x}^T(t-1), S_2(0) = 0 \\
 \hat{x}(t) &= \hat{B}(t-1)\hat{x}^*(t-1) + c\hat{\alpha}^T(t-1)y(t-1) \\
 \hat{x}^*(t) &= \hat{B}(t)\hat{x}^*(t-1) + c\hat{\alpha}^T(t)y(t-1)
 \end{aligned}$$

$$\begin{aligned}
 \hat{\phi}(t) &= \hat{\phi}(t)/\hat{B}_{t-1}(q^{-1}), \hat{B}_{t-1}(q^{-1}) \\
 & = 1 + \hat{b}_1(t-1)q^{-1} + \cdots + \hat{b}_n(t-1)q^{-n}
 \end{aligned}$$

アルゴリズムでの本手法と RML 法の違いは、 $\hat{\ell}(t), \hat{\phi}(t)$ において、本手法は

$$\left. \begin{aligned}
 \hat{\ell}(t) &= y(t) - c^T \hat{x}(t|t-1) \\
 \hat{\phi}^T(t) &= [y^T(t-1) : y^T(t-1) \\
 & - \hat{x}^T(t-1|t-1)]
 \end{aligned} \right\} (41)$$

としているのに対し、RML 法では

$$\left. \begin{aligned}
 \hat{\ell}(t) &= y(t) - \hat{\phi}^T(t)\hat{\alpha}(t-1) \\
 \hat{\phi}^T(t) &= [y^T(t-1) : \hat{\ell}(t-1) \cdots \hat{\ell}(t-n)]
 \end{aligned} \right\} (42)$$

としている点である。(41)を形式的に

$$\left. \begin{aligned} \varepsilon(t) &= y(t) - c^T \hat{x}(t) \\ \hat{\phi}^T(t) &= [y^T(t-1) : y^T(t-1) - \hat{x}^T(t-1)] \\ \hat{x}(t) &= \hat{B}(t-1)\hat{x}(t-1) \\ &\quad + c d^T(t-1)y(t-1) \end{aligned} \right\} \quad (43)$$

とすれば、RML 法になる。このことはつぎのようにしてわかる。いま、 $\hat{x}^T(t) = [\hat{x}_1(t) \cdots \hat{x}_n(t)]$ として (43) の第 3 式を書き直すつぎようになる。

$$\hat{B}_{t-1}(q^{-1})\{y(t) - \hat{x}_1(t)\} = \hat{A}_{t-1}(q^{-1})y(t) \quad (44)$$

また (42) の第 1 式を書き直すと

$$\hat{B}_{t-1}(q^{-1})\varepsilon(t) = \hat{A}_{t-1}(q^{-1})y(t) \quad (45)$$

となる。(44), (45) より $\hat{B}_{t-1}(q^{-1})\{y(t) - \hat{x}_1(t)\} = \hat{B}_{t-1}(q^{-1})\varepsilon(t)$ となり、これより

$$\hat{x}_1(t) = y(t) - \varepsilon(t) \quad (46)$$

すなわち、(42) の第 1 式を用いて

$$c^T \hat{x}(t) = \hat{\phi}^T(t) \hat{\alpha}(t-1) \quad (47)$$

となる。したがって (42), (43) の $\varepsilon(t)$ は同じになる。

また (43) の第 3 式を用いて、(46) を考慮すると

$$\begin{aligned} \hat{x}^T(t) &= [\hat{x}_1(t) \cdots \hat{x}_n(t)] \\ &= [\hat{x}_1(t) \cdots \hat{x}_1(t-n+1)] \\ &= [y(t) - \varepsilon(t) \cdots y(t-n+1) - \varepsilon(t-n+1)] \end{aligned} \quad (48)$$

となり、これを (43) の第 2 式の $\hat{x}^T(t-1)$ の部分に用いると (42) の $\hat{\phi}^T(t)$ が得られる。したがって、(42), (43) の $\hat{\phi}^T(t)$ も同じになる。本手法と RML 法の比較のために、Table 1 に (43) を用いた形の RML アルゴリズムを示す。また後の数値実験に関連して、residual を用いた RML アルゴリズムも示す。

ところで、Table 1 の本アルゴリズムと RML アルゴリズムを比較した場合、 $x(t)$ の推定式が異なっているが、これは $E\{\hat{V}_i^*(\alpha)\}$ の推定方式の違いによるものである。RML 法は、(14) をもとに

$$E\{\hat{V}_i^*(\alpha)\} \approx \sum_{i=1}^t \phi(i, \alpha) \phi^T(i, \alpha) \quad (49)$$

とおいて、 $\hat{V}_i^*(\alpha)$ を

$$\hat{V}_i^*(\alpha) = \sum_{i=1}^t \hat{\phi}(i, \alpha) \hat{\phi}^T(i, \alpha) \quad (50)$$

$$\left. \begin{aligned} \hat{\phi}^T(t, \alpha) &= \hat{\phi}^T(t, \alpha) / B(q^{-1}), \\ \hat{\phi}^T(t, \alpha) &= [y^T(t-1) : y^T(t-1) - \hat{x}^T(t-1)] \\ \hat{x}(t) &= B\hat{x}(t-1) + c d^T y(t-1) \end{aligned} \right\} \quad (51)$$

とするものである。この場合、(49) の近似が少し大胆な感じがするが、これに対し本手法では (13) が最小になるように $\hat{V}_i^*(\alpha)$ が決定される。その結果、 $\hat{\phi}(t, \alpha)$ は (16) のような期待値が関係してくる。これを評価するために拡張 Kalman フィルタが用いられ、その結果 RML 法に比べ $k_x, Q_x, Q_\alpha, k_\alpha, \hat{\sigma}^2$ の計算が余分に

入ってくる。

ところで (1), (3) より Fisher の情報行列 $J_i(\alpha)$ は

$$J_i(\alpha) = \sigma^{-2} \sum_{i=1}^t E\{\phi(i, \alpha) \phi^T(i, \alpha)\} \quad (52)$$

で与えられる²⁾。このとき (14) と上式を比較して

$$E\{\hat{V}_i^*(\alpha)\} = \sigma^2 J_i(\alpha) \quad (53)$$

を得る。したがって (13) を最小にするように $\hat{V}_i^*(\alpha)$ を決定することは

$$E\{\hat{V}_i^*(\alpha)\} = \sigma^2 J_i(\alpha) - E\{\hat{V}_i^*(\alpha)\} \quad (54)$$

を最小にするように $\hat{V}_i^*(\alpha)$ を決定することと同じである。そこで $\sigma^2 J_i(\alpha)$ を推定器が最大限に発揮できる情報処理能力と考えると、本手法の $\sigma^2 J_i(\alpha)$ の扱い方は $\hat{V}_i^*(\alpha)$ が (54) を最小にする意味で、 $\sigma^2 J_i(\alpha)$ の内の最大限を $\hat{V}_i^*(\alpha)$ に吸収させようとすることをめざしたものである。これに対し、RML 法では、 $\sigma^2 J_i(\alpha)$ を一挙に (49) の置き換えをもとに (50) で近似している。近似が粗く $\hat{V}_i^*(\alpha)$ は本手法に比べ情報処理能力の吸収度は低いと考えられる。しかし実際の計算においては、 α が関係する部分は、この running estimate で置き換えられるので、これらのことは期待事項である。

6. 数値実験例

つぎの 2 次の系⁴⁾ について実験を行った。

$$\begin{aligned} y(t) &= 1.5y(t-1) + 0.7y(t-2) + e(t) \\ &\quad - e(t-1) + 0.2e(t-2) \end{aligned} \quad (55)$$

$e(t)$ は $N[0, 1]$ に従う統計的に独立な乱数を用いた。実験は 15 の見本系列について行い、 α の推定精度の評価は推定誤差ベクトルの大きさを 15 本の見本系列について平均したもの、すなわち

$$d_\alpha(t) = \frac{1}{15} \sum_{j=1}^{15} \|\alpha - \hat{\alpha}^{(j)}(t)\| \quad (56)$$

を用いた。また σ^2 の推定精度の評価は

$$\hat{\sigma}^2(t) = \frac{1}{15} \sum_{j=1}^{15} \hat{\sigma}^{2(j)}(t) \quad (57)$$

を用いた。ここで、 $\hat{\alpha}^{(j)}(t)$ 、 $\hat{\sigma}^{2(j)}(t)$ はそれぞれ j 番目の見本系列に対する α 、 σ^2 の推定値である。本アルゴリズムの初期値はつぎのとおりである。

$$\left. \begin{aligned} \hat{x}(0|0) &= 0, \hat{\alpha}(0) = 0, P(0) = Q_\alpha(0) \\ &= 100I, Q_x(0|0) = 10I, \\ \hat{\sigma}^2(0) &= 0.1 \end{aligned} \right\} \quad (58)$$

また RML 法は Table 1 のアルゴリズムを用いた。

そして初期値はつぎのとおりである。

$$\hat{x}(0) = 0, \hat{\alpha}(0) = 0, P(0) = 100I \quad (59)$$

このとき、Fig. 1 は α の推定精度を示す。 $d_\alpha(t)$ は RML 法に比べ、本手法のほうが減少率が大きく、収

Fig. 1 Time evolution of parameter estimation errors

Fig. 2 Time evolution of noise variance estimates

束速度が加速されていることがわかる。Fig. 2 は σ^2 の推定精度を示すもので、●および○は RML 法をもとに

$$\hat{\sigma}^2(z) = \frac{1}{t} \sum_{i=1}^t [y(i) - \mathbf{c}^T \hat{\mathbf{x}}(i)]^2 \quad (60)$$

で推定した結果を $\hat{\sigma}^{2(j)}(z)$ として、(57)で評価したものである。この場合も、本手法の方が収束性が優れていることがわかる。

7. あとがき

本論文は、ARMA 過程のパラメータを RML 法で推定する場合の、推定値の収束速度の改善について述べた。その方法は、パラメータ推定器の情報処理能力を表わすと考えられる Fisher の情報行列をもとに、できるだけ多くこれを推定器に吸収させることであ

た。これは具体的には、Gauss-Newton 法による最小化アルゴリズムにおいて必要になる Hessian の期待値をできるだけ正確に推定することによりなされた。このことはパラメータ推定ゲイン $\mathbf{g}(t) = \mathbf{P}(t)\hat{\phi}(t)$ が (21) の $\mathbf{P}(t)$ の定義式をもとに

$$\mathbf{g}(t) = [\hat{\mathbf{V}}_t'(\hat{\mathbf{a}}_r(t-1))]^{-1}\hat{\phi}(t) \quad (61)$$

と書けるので、 $\hat{\mathbf{V}}_t'(\hat{\mathbf{a}}_r(t-1))$ をできるだけ精度よく求めることにより、 $\mathbf{g}(t)$ の精度を向上させようとする試みである。その結果、推定アルゴリズムは RML 推定器と拡張 Kalman フィルタを連動させる形になった。そして Kalman フィルタの安定性に関する理論を用いることにより、RML 法では必要とされる $\hat{\mathbf{B}}(t)$ の安定性の監視が必要になった。拡張 Kalman フィルタを用いることの得失については、計算量は RML 法に比べて増加するが、 $E\{\hat{\mathbf{V}}_t'(\hat{\mathbf{a}})\}$ をより正確に近似できることが期待されるので、その結果 $\hat{\mathbf{a}}(t)$ の収束速度の改善も期待できることである。本数値実験においては、収束速度は改善されることが確認された。

RML 法でも、 $\hat{\mathbf{B}}(t)$ もしくは $\hat{\mathbf{B}}_t(q^{-1})$ の安定性を監視する場合には、この分だけ計算量は増える。しかしこの場合には、計算量が増えた分だけ推定器の情報処理能力の向上に寄与するというようなことはないと思われる。

最後に、数値実験を行うにあたり多大の便宜を図ってくださった徳島大学教授河田照茂博士に謝意を表する次第である。

参 考 文 献

- 1) T. Söderström, L. Ljung and I. Gustavsson: A Theoretical Analysis of Recursive Identification Methods, *Automatica*, **14**, 231/244 (1978)
- 2) 得丸, 添田, 中溝, 秋月: 計数・測定, pp. 113, 128, 203/209, 培風館 (1982)
- 3) 植木, 添田, 中溝: 確率システム制御の基礎, pp. 92, 日新出版 (1975)
- 4) 片山: 応用カルマンフィルタ, pp. 119, 朝倉書店 (1984)
- 5) T. Yoshimura, K. Konishi, R. Kiyozumi and T. Soeda: Identification of Unknown Parameters in Linear Discrete-Time Systems by a Modified Extended Kalman Filter, *Int. J. Systems Sci.*, **11-1**, 97/105 (1980)
- 6) J.B. Moore and B.D.O. Anderson: Coping with Singular Transition Matrices in Estimation and Control Stability Theory, **31-3**, 571/586 (1980)
- 7) L. Ljung and T. Söderström: Theory and Practice of Recursive Identification, Chap. 2, MIT Press (1983)
- 8) N.K. Gupta and R.K. Mehra: Computational Aspects of Maximum Likelihood Estimation and Reduction in Sensitivity Function Calculations, *IEEE Trans. on Automatic Control*, **AC-19-6**, 774/783 (1974)