反転不能な移動平均過程における推定の問題について

小　瀧　光　博

1. はしがき

本稿では，真の係数が1あるいは−1である様な1階の移動平均過程の係数の推定とそれに関連した統計的諸問題が扱われる．従来，移動平均過程の統計的解析，特にその係数パラメータの推定は，その特性方程式の根の絶対値が1よりも大になる様なケースにのみ限定されていた．すなわち，1階の移動平均過程を例にとると，その係数の絶対値が1よりも小になる様なケースである．その様な移動平均過程は反転可能といわれ，“意味のある”無限次数の自己回帰表現を得ることが可能となる．パラメータの推定量としては，通常最尤推定量が用いられるが，その統計的性質については，適当な仮定の下で，漸近的正規性に基いた標準的結果が成立する．一方，特性方程式の根の1つが絶対値にといて1に等しくなっている様なケースの移動平均過程（その次数が1の時には，その係数が1あるいは−1となる様なもの）は，反転不能といわれ，そこから得られる自己回帰表現は“収束”するものとはならない．Plosser and Schwert (1977) 等においても指摘されている様に，この様な反転不能な移動平均過程は，主として時系列解析における過度の“差分操作 (overdifferencing)”によって生じてくる場合が多い．多くの時系列，とりわけマクロ経済時系列については，その生のデータ系列は非定常的なものとなっている．Box-Jenkins流のアプローチにおいては，元の系列を何回か差分することによって定常

時系列に変換することができると想定される．ところが，この差分操作が“過度”になされた時，変換された系列は定常ではあるが反転不能なものとなってしまうのである．

ところで，反転不能な移動平均過程における統計的解析上の問題点は，単に“有意味な”無限次数の自己回帰表現が得られないということだけではなく，そのことによって係数パラメータの推定が非常に困難なものとなっているということにもある．反転不能な場合にも係数パラメータの推定量としては，通常最尤推定量が考えられるが，後で詳細に論じられる様に，その尤度関数は係数パラメータの真の値（1階の反転不能移動平均過程の場合，1ないしは－1）の近傍で非常に“たちの悪い”不連続的な関数となっている．また，移動平均過程においてはパラメータの識別条件が考慮されなければならず，そのためパラメータ空間がある閉空間（1階の移動平均過程の場合，[－1，1] 区間）に制限されることになり，さらにそのことはパラメータの真の値がパラメータ空間の境界上に存在することを意味する．これらのことによって，反転不能な移動平均過程における最尤推定量の漸近的性質を評価することは，きわめて厄介なものとなっているのである．それ故，反転不能な移動平均過程における係数パラメータの最尤推定量の漸近的性質については，まだごく一部しか確立されておらず，過去の研究においてもすべて1階の移動平均過程での分析に限定されている．この論文においても1階の過程が取り上げられることになるが，恐らく過程の次数が高くなっていくにつれて漸近的性質は複雑な様相を示してくるものと思われる．反転可能な場合の様な統一的な結果は得られないであろうと思われる．

この分野における過去の研究結果について触れておくと，Davidson (1981), Plosser and Schwert (1977) といった人達は，反転不能な1階の移動平均過程における係数の最尤推定量には，漸近的正規性に基いた，通常の標準的漸近理論は適用できないことを示した．さらに，Sargan and Bhargava (1983), Anderson and Takemura (1986), Tanaka and Sachell (1989) といった人達によって，最尤推定量の漸近的性質がかなり明瞭に

なってきた．まず，係数の真の値が1あるいは－1である時には，その尤度関数において，反転可能なケースでは“無視”できた項が無視し得ないことが指摘される．これらの項の中には“初期値”の効果を示す項も含まれることも示される．そうして，これらの項を無視するか否かによって3つのタイプの（最尤）推定量が提示される．これらの人達の論文において，これら3つのタイプの推定量の真値への収束の速度が通常よりも著しく速くなっていることが非統一的に示されたが，それらの漸近分布を評価することはできなかった．また，なぜこういった結果が得られるのか，あるいは得られなかったのかということについての説明も十分になされなかった．

本稿の主たる目的は，過去の論文において得られた，反転不能な1階の移動平均過程の最尤推定量についての漸近的結果を，過去の論文のそれとは異なるアプローチによって導出し，且つ前述の3つのタイプの推定量を統一的に分析することである．既に述べた最尤推定量の漸近的性質の評価を難しいものとしている2つの要因を詳細に吟味することによって，なぜ最尤推定量の真の値への収束の速度が通常よりも速いのか，また，なぜ最尤推定量の漸近分布を評価するのが困難なものとなっているのかが明らかにされる．特に，尤度関数がパラメータの真値の近傍で非常に“たちの悪い”不連続的なものになっている点を解明するために，過去の諸文献においては明瞭に述べられてこなかった，（1階の）移動平均過程における最適線形予測と反転可能性との関連が分析される．最適線形予測量を構成する表現と反転によってもたらされる自己回帰表現とを比較考察することによって，反転可能性と尤度関数の性質との関連性が論じられる．3つのタイプの推定量を導出するための尤度方程式の漸近的性質についても，本稿において強調される事項である．尤度方程式はパラメータの複雑な非線形関数となっているが，一般に，線形近似し，3次以降の項を無視することによって最尤推定量が一義的に求められ，且つその分布が評価できる．このことは，反転可能な移動平均過程をはじめとする多くの標準的漸近く理論が成立する様な時系列に対して，漸近的に可能となる．しかしながら，反転不能な移動平均過程においては，その尤度方程式は漸近的にもパラ

メータに対して高度に複雑な非線形関数となっており，従ってそこから一義的な解を得てその分布を評価することはきわめて困難であることが指摘される．

この論文のもう一つの目的は，その漸近分布が反転不能であっても正規分布になる様な，最尤推定量以外の推定量を得ることである．そのことは，パラメータ空間の制限を取り除くことによって，また，定数項部分が相対的に大きくなる様な，尤度関数とは別の評価関数を構築することによって達成されることが示される．

2.　1階の移動平均過程とパラメータの識別条件 ——パラメータ空間の制限より生ずる問題をめぐって

1階の移動平均過程

$$y_t=\varepsilon_t-a\varepsilon_{t-1}\ ;\ \varepsilon_t\sim i.i.d.(0,\ \sigma^2) \tag{1}$$

を考察しよう．ここで $\varepsilon_t\sim i.i.d.(0,\ \sigma^2)$ は，ε_t が平均０，分散 σ^2 の独立同一分布に従うことを意味する．(1)から，次の様なもう一つ別の（１階の）移動平均過程が得られる．すなわち

$$y_t=e_t-a^{-1}e_{t-1} \tag{2}$$

ここで $e_t=\varepsilon_t-(1-a^{-2})a\varepsilon_{t-1}-(1-a^{-2})\Sigma_{k=0}^{\infty}a^{-k}\varepsilon_{t-2-k}$. また，(2)が移動平均過程になっていること，すなわち，e_t が平均０，分散 $a^2\sigma^2$ 且つ系列的に無相関であることは，容易に確かめられる．すなわち

$$Ee_t^e=[1+(1-a^{-2})^2a^2+(1-a^{-2})^2(1-a^{-2})^{-1}]\sigma^2$$
$$=a^2\sigma^2$$

$$Ee_te_{t-j}=[-(1-a^{-2})a^{-j+2}+(1-a^{-2})^2a^{-j+2}+(1-a^{-2})^2a^{-j}+\cdots]\sigma^2$$
$$=[-(1-a^{-2})a^{-j+2}+(1-a^{-2})^2a^{-j+2}\sum_{m=0}^{\infty}a^{-2m}]\sigma^2$$
$$=[-(1-a^{-2}a^{-j+2}+(1-a^{-2})a^{-j+2}]\sigma^2$$
$$=0,\ \ j=1,\ 2,\ 3,\ \cdots$$

(1)における a について，$|a|\neq1$ である時(1)によって生成される y_t は反転可能といわれ．(1)と(2)は y_t の異なる２つの移動平均過程となる．また，$|a|=1$ である時 y_t は反転不能といわれ，(1) と(2)は同一の表

現となる．(1)において $|a|>1$ である時，(2)はその係数の絶対値が 1 より小となる様な 1 階の移動平均過程を表わすものとなっている．すなわち，(1)に対して(2)の表現が常に可能てあるということは，y_t が 1 階の移動平均過程によんて生成されるものである時，常に $|a|\leq 1$ である様なものが存在することを意味する．y_t の自己相関係数が

$$Ey_ty_{t-1}/Ey_t^2=\frac{-a}{1+a^2}=\frac{-(1/a)}{1+(1/a)^2}$$

$$Ey_ty_{t-j}/Ey_t^2=0,\quad j=2,\ 3,\ \cdots$$

であることより，(1)と(2)は観測的に同等 (observationally equivalent) といわれる．すなわち，1 階の移動平均過程には反転可能である場合，観測的に同等の 2 つの異なるものが存在する．

観測的に同等な確率過程の存在は，(1)におけるパラメータ $(a,\ \sigma^2)$ の識別の問題が考察される必要があることを示唆する．しかしながら，(2)の存在は必らずしも直ちに $(a,\ \sigma^2)$ が識別されないことを意味するわけではない．注意しなければならないことは，パラメータ $(a,\ \sigma^2)$ の識別か問題になってくるのは ε_t が正規分布する（その時，e_t も正規分布することになり，従って独立同一分布に従うことになる）場合の尤度関数に対してであって，ε_t が正規分布しない場合には，ε_t の 3 次以上のモーメント次第で，$(a,\ \sigma^2)$ は尤度関数に関して識別されるかもしれないし，また，たとえ ε_t が正規分布したとしても，$(a,\ \sigma^2)$ を識別する様な尤度関数以外の評価関数が存在し得るかもしれないのである．実際，本稿においても 5 節で，その様な評価関数を提示する．

最尤推定を考察するために ε_t が正規分布に従うものと仮定すると，$(a,\ \sigma^2)$ は尤度関数に関して識別されないことになる．そのことは，(1)と(2)の異なるパラメータ（$(a,\ \sigma^2)$ と $(a^{-1},\ a^2\sigma^2)$）に対して同一の自己共分散ないしは自己相関係数が対応していることより明白である．(1)か(2)の一方において係数パラメータは絶対値において 1 を越えないので，$(a,\ \sigma^2)$ が識別可能となる様に，a のとり得る領域，すなわち a のパラメータ空間を R^1（1 次元実数空間）から［－1，1］区間に制限する．この制限

によって (a, σ^2) は識別可能となるけれども，同時にこの制限は反転不能な場合の移動平均仮定の最尤推定に一つの困難性をもたらすことになる．すなわち，この場合係数パラメータの真の値（1あるいは−1）がパラメータ空間の境界上にあることになるのである．このことは，例えば a のある関数 $M(a)$ を閉区間［−1，1］上で最小にする問題を考えた時，$a=1$ で $M(a)$ が極小になるためには

$$\left.\frac{dM}{da}\right|_{a=1}=0,\quad \left.\frac{d^2M}{da^2}\right|_{a=1}>0 \tag{3}$$

かあるいは

$$\left.\frac{dM}{da}\right|_{a=1}<0 \tag{4}$$

のいずれか一方が成立すればよいということに端的に示されよう．$-M(a)$ が何らかの尤度関数である時，最尤推定の標準理論に従えば $M(a)$ を最小にする様な a を $\hat{a}^*$ と表わし，パラメータ a の真の値 a_0 がパラメータ空間の内点であるならば，適当な仮定の下で尤度方程式

$$\left.\frac{dM}{da}\right|_{a=\hat{a}^*}=0 \tag{5}$$

より

$$\sqrt{T}(\hat{a}^*-a_0)=\left[\frac{1}{T}\left.\frac{d^2M}{da^2}\right|_{a=a_0}\right]^{-1}\left[-1/\sqrt{T}\left.\frac{dM}{da}\right|_{a=a_0}\right]+O_p(1/\sqrt{T}) \tag{6}$$

が得られる．ここで T はデータ数を表わす．$x\leq y$ となる任意の実数 x, y に対して，(6)は

$$\begin{aligned}&\lim_{T\to\infty}P[a_0+x/\sqrt{T}\leq\hat{a}^*\leq a_0+y/\sqrt{T}]\\&=\lim_{T\to\infty}P\left[x\leq\left(1/T\left.\frac{d^2M}{da^2}\right|_{a=a_0}\right)^{-1}\left(-1/\sqrt{T}\left.\frac{dM}{da}\right|_{a=a_0}\right)\leq y\right]\end{aligned} \tag{7}$$

とも書ける．ここで $P[A]$ は事象 A が起きる確率を表わす．ところで，もし a のとり得る領域が $\{a\leq a_0\}$ に制限されるならば，制限されたパラメータ空間 $\{a\leq a_0\}$ の上で $M(a)$ を最小にする様な a を $\hat{a}^{**}$ と表わす時，

$$\lim_{T\to\infty}P[\hat{a}^{**}=a_0]$$

$$=\lim_{T\to\infty} P\left[1/\sqrt{T}\ \frac{dM}{da}\bigg|_{a=a_0}<0\right]+\lim_{T\to\infty} P\left[1/\sqrt{T}\ \frac{dM}{da}\bigg|_{a=a_0}=0,\right.$$

$$\left.\frac{1}{\mathrm{T}}\ \frac{\mathrm{d}^2\mathrm{M}}{\mathrm{da}^2}\bigg|_{\mathrm{a}=\mathrm{a}_0}>0\right] \tag{8}$$

が成立することが，(3)(4)を考慮することにより直ちに示される．関係式(8)は，$T\to\infty$ の時のみならず任意の T に対して成立するものであることに注意せよ．一方，$\bar{a}<a_0$ なる任意の $\bar{a}$ に対しては

$$\lim_{T\to\infty} P[\hat{a}^{**}=\bar{a}]$$

$$=\lim_{T\to\infty} P\left[1/\sqrt{T}\ \frac{dM}{da}\bigg|_{a=\bar{a}}=0,\ \ \frac{1}{T}\ \frac{d^2M}{da^2}\bigg|_{a=\bar{a}}>0\right] \tag{9}$$

が成立しなければならないことにより，注意の $z>0$ に対して

$$\lim_{T\to\infty} P[a_0-z/\sqrt{T}\leq\hat{a}^{**}\leq a_0]$$

$$=\lim_{T\to\infty} P\left[-z\leq\left(1/T\ \frac{d^2M}{da^2}\bigg|_{a=a_0}\right)^{-1}\left(-1/\sqrt{T}\ \frac{dM}{da}\bigg|_{a=a_0}\right)\leq 0\right] \tag{10}$$

であることが示される．a_0 がパラメータ空間の内点であるケースでは，1点確率は測度ゼロであることより，$\hat{a}^*$ が a_0 に等しくなる確率は（小標本においても大標本においても）ゼロとなる．しかしながら，明らかに $\left\{1/\sqrt{T}\ \frac{dM}{da}\Big|_{a=a_0}<0\right\}$ が起きる確率は測度ゼロではなく，そのことは a_0 がパラメータ空間の境界上にあるケースでは，$\hat{a}^{**}$ が a_0 である確率，すなわち $M(a)$ が a_0 で極小となる確率もゼロではないことを意味する((8)を見よ)．従って a_0 がパラメータ空間の境界上にあるケースでは，$\hat{a}^{**}$ が a_0 に等しくなる確率，とりわけその極限値を評価することが1つの目標となるであろう．但し(10)は，$\hat{a}^{**}$ の漸近分布が，適当な仮定の下で $\hat{a}^*$ のそれ((7)で与えられる）と同様に評価できるであろうことを示唆している．

反転不能な1階の移動平均過程(1)の最尤推定の問題に戻って，$-M(a)$ は ε_t の正規性が仮定された時の尤度関数としよう．また，パラメータ a のとり得る領域は識別条件が課された結果，[－1，1]に制限されているものとする．(1)と(2)が y_t について同一の分散・自己共分散を生み出

すことより，

$$M(a)=M(1/a) \tag{11}$$

であることが直ちに示される．(11)は

$$\left.\frac{dM}{da}\right|_{a=1}=0,\ \text{あるいは}\ \left.\frac{dM}{da}\right|_{a=-1}=0 \tag{12}$$

を意味する．(12)は y_t が反転可能である場合（a の真の値 a_0 が $|a_0|<1$ である場合）においても成立する．このことは，(y_t が反転可能であろうがなかろうが) $a=1$ あるいは $a=-1$ で $M(a)$ の1階の微係数が常に0となっていることを意味する．(12)を考慮すると，$|a_0|=1$ である時(8)は

$$\begin{aligned}&\lim_{T\to\infty}P[\hat{a}^{**}=a_0]\\&=\lim_{T\to\infty}P\left[\frac{1}{T}\left.\frac{d^2M}{da^2}\right|_{a=a_0}>0\right]\end{aligned} \tag{13}$$

と書きかえられるであろう．

以上考察してきた様に，反転不能な移動平均過程の最尤推定問題は，移動平均過程に伴う識別性の問題やパラメータ空間の制限とパラメータの真値がパラメータ空間の境界上にくることによって，著しく“標準理論”を逸脱したものになっている．

3. 最適線形予測量と反転可能性——尤度関数に与える影響

(1)によって生成される y_t について，n 次の自己回帰モデルに適合させてみよう．(1)より直ちに

$$\begin{aligned}&y_t+ay_{t-1}+a^2y_{t-2}+\cdots+a^ny_{t-n}\\&=\varepsilon_t-a^{n+1}\varepsilon_{t-n-1}\end{aligned} \tag{14}$$

が得られる．y_t が反転可能である時，すなわち $|a|<1$ である時，(14)において $n\to\infty$ とすると

$$y_t+ay_{t-1}+a^2y_{t-2}+\cdots=\varepsilon_t \tag{15}$$

となる．(15)は y_t についての無限次数の自己回帰表現であり，(15)が(1)からもたらされるということが“反転”という名称の由来になっている．

すなわち，　$|\lambda| \leq 1$ なる任意の複素数λに対して

$$(1-a\lambda)^{-1}=1+a\lambda+a^2\lambda^2+\cdots$$

一方，y_t が反転不能である時，例えば $a=1$ である時(14)は

$$y_t+y_{t-1}+y_{t-2}+\cdots+y_{t-n}=\varepsilon_t-\varepsilon_{t-n-1} \tag{16}$$

(16)において $n\to\infty$ としても ε_{t-n-1} の項は決して消えない．すなわち，(1)において $a=1$ である時，“意味”のある無限次数の自己回帰表現は得られない．

ところで，時系列が無限次数の（有限次数でもかまわないが…）自己回帰過程によって表現できるという性質が何故“望まれる”ものとみなされるのかというと，それが線形で表わされるものの中で“最良”な予測量に対応するものになっているからである．$|a|<1$ の時，(15)において

$$R(y_t \mid y_{t-1}, y_{t-2}, \cdots)=-(ay_{t-1}+a^2y_{t-2}+\cdots) \tag{17}$$

は，$(y_{t-1}, y_{t-2}, \cdots)$ を用いての y_t の最適線形予測量となっている．最適線形予測量は，$(y_{t-1}, y_{t-2}, \cdots)$ の任意の線形結合 $b_1y_{t-1}+b_2y_{t-2}+\cdots$ について，y_t との平均２乗誤差

$$E\{y_t-(b_1y_{t-1}+b_2y_{t-2}+\cdots)\}^2$$

を最小にする様なものであるが，ある線形結合 $b_1y_{t-1}+b_2y_{t-2}+\cdots$ が最適線形予測量となっているための必要十分条件は

$$E\{[y_t-(b_1y_{t-1}+b_2y_{t-2}+\cdots)]y_{t-j}\}=0,\quad j=1, 2, \cdots \tag{18}$$

で与えられる．$b_1=-a$，$b_2=-a^2$，… の時，$y_t-(b_1y_{t-1}+b_2y_{t-2}+\cdots)=\varepsilon_t$ であるので，(17)で与えられる $R(y_t \mid y_{t-1}, y_{t-2}, \cdots)$ が $(y_{t-1}, y_{t-2}, \cdots)$ を用いての y_t の最適線形予測量となっていることは明らかである．

しかしながら，反転不能のケースにおいて“意味”のある無限数の自己回帰表現が得られないことが，反転不能のケースで $(y_{t-1}, y_{t-2}, \cdots)$ を用いての y_t の最適線形予測量が見い出せないことを意味するものではない．次の命題では反転可能性があるか否かにかかわらず，$(y_{t-1}, y_{t-2}, \cdots, y_{t-n})$ を用いての y_t の最適線形予測量の表現が与えられる．

命題1.　(1)によって生成される y_t に対して，$(y_{t-1}, y_{t-2}, \cdots y_{t-n})$ を用い

ての y_t の最適線形予測量を $P(y_t \mid y_{t-1}, y_{t-2}, \cdots, y_{t-n})$ と書くことにすると

$$P(y_t \mid y_{t-1}, y_{t-2}, \cdots, y_{t-n}) = -\bar{b}_n^{-1}\{(a+a^3+\cdots+a^{2n-1})y_{t-1} + (a^2+a^4+\cdots+a^{2n-2})y_{t-2}+\cdots+(a^{k+1}+\cdots+a^{2n-k-1})y_{t-k-1} + \cdots + (a^{n-1}+a^{n+1})y_{t-n+1}+a^n y_{t-n}\} \tag{19}$$

及び

$$y_t - P(y_t \mid y_{t-1}, y_{t-2}, \cdots, y_{t-n}) = \varepsilon_t - \bar{b}_n^{-1} a^{n+1}\left\{\sum_{j=0}^{n} a^{n-j}\varepsilon_{t-1-j}\right\} \tag{20}$$

であることが示される．但し $\bar{b}_n = 1+a^2+\cdots+a^{2n}$.

この命題は，関係式(20)を用いて $j=1, 2, \cdots, n$ について(18)が成立することを示すことによって，また，(19)の y_{t-j} に $\varepsilon_{t-j}-a\varepsilon_{t-j-1}$ を代入して(20)を導出することによって容易に証明できる．

$$R(y_t \mid y_{t-1}, y_{t-2}, \cdots, y_{t-n}) = -(ay_{t-1}+a^2y_{t-2}+\cdots+a^n y_{t-n}) \tag{21}$$

と置くと

$$E\{y_t - R(y_t \mid y_{t-1}, y_{t-2}, \cdots, y_{t-n})\}^2 = \sigma^2(1+a^{2(n+1)}) > E\{y_t - P(y_t \mid y_{t-1}, y_{t-2}, \cdots, y_{t-n})\}^2 \tag{22}$$

ここで(22)における不等号は，$E\{[y_t - R(y_t \mid y_{t-1}, y_{t-2}, \cdots, y_{t-n})]y_{t-n}\} = E\{[\varepsilon_t - a^{n+1}\varepsilon_{t-n-1}]y_{t-n}\} = a^{n+2}\sigma^2 \neq 0$ であることより $R(y_t \mid y_{t-1}, y_{t-2}, \cdots, y_{t-n})$ が最適線形予測量でないことからでてくる．$a=1$ の時，(19)(20)はそれぞれ

$$P(y_t \mid y_{t-1}, y_{t-2}, \cdots, y_{t-n}) = -\left\{\left(\frac{n}{n+1}\right)y_{t-1} + \left(\frac{n-1}{n+1}\right)y_{t-2} + \cdots + \left(\frac{2}{n+1}\right)y_{t-n+1} + \left(\frac{1}{n+1}\right)y_{t-n}\right\} \tag{23}$$

$$y_t - P(y_t \mid y_{t-1}, y_{t-2}, \cdots, y_{t-n}) = \varepsilon_t - \left(\frac{1}{n+1}\right)\left\{\sum_{j=0}^{n} \varepsilon_{t-1-j}\right\} \tag{24}$$

と書かれる．n が充分大である時，(24)式右辺第２項は $O_p(1/\sqrt{n})$ であるので，$n\to\infty$ とすることによって $a=1$ の場合においても $P(y_t \mid y_{t-1},$

y_{t-2}, …) あるいは $y_t-P(y_t \mid y_{t-1}, y_{t-2}, \cdots)$ に関する表現がもたらされることになる．ところで

$$E\{y_t-P(y_t \mid y_{t-1}, y_{t-2}, \cdots, y_{t-n})\}^2 = \begin{cases} \sigma^2[1+(1-a^2)(1-a^{2(n+1)})^{-1}a^{2(n+1)}] \cdots \mid a \mid <1 \text{ の時} \\ \sigma^2\left[1+\left(\frac{1}{n+1}\right)\right] \quad \cdots a=1 \text{ の時} \end{cases} \tag{25}$$

は，$|a|<1$ の場合には $y_t-P(y_t \mid y_{t-1}, y_{t-2}, \cdots, y_{t-n})$ は（そして $y_t-R(y_t \mid y_{t-1}, y_{t-2}, \cdots, y_{t-n})$ も）$a^{2(n+1)}$ という幾何級数的なスピードで $\varepsilon_t=y_t-P(y_t \mid y_{t-1}, y_{t-2}, \cdots)$ に収束していくのに対して，$a=1$ の場合には $(1/n+1)$ という非常に遅いスピードで収束していくことを示す．同様なことは y_{t-n} の係数についても成り立つ．また，$|a|<1$ の時には n が大の時，$y_t-P(y_t \mid y_{t-1}, y_{t-2}, \cdots, y_{t-n})$ と $y_t-R(y_t \mid y_{t-1}, y_{t-2}, \cdots, y_{t-n})$ との差は全んどないのに対して，$a=1$ の時には n が大であっても明瞭な差が存在する．

ところで，充分大きな n に対して a が1の近傍にある時，それがどの程度1の近くにあるかで(19)や(21)の右辺の表現式の係数，とりわけ y_{t-n} の係数は著しく異なってくる．すなわち，a が1に近づくにつれて，(19)や(21)の表現は著しく変化していく．例えば，$n\to\infty$ の時 $P_n\to\infty$ となる様な正数列 $\{P_n\}$ に対して $a=1-C/P_n$ とした時，a^n がどの様に変化するか考察してみよう．但し C は適当な正数とする．任意の実数 $\lambda>0$ に対して

$$\lim_{n\to\infty}[1-C/n]^{n\lambda}=(1/e^c)^{\lambda} \tag{26}$$

であることを考慮して，$P_n \doteqdot n\lambda_n^{-1}$ で，$n\to\infty$ の時 $\lambda_n^{-1}\to 0$ 且つ $n\lambda_n^{-1}\to\infty$ となるくらい P_n が“大きくない”時，

$$[1-C/P_n]^n \doteqdot [[1-C/P_n]^{P_n}]^{\lambda_n} \tag{27}$$

であることより，a^n は0に収束する．つまり，この場合の(19)及び(20)（そして(21)）は，反転可能な場合（a が明らかに1から離れている場合）と同じ動きをする．しかしながら，任意の実数 $\lambda>0$ に対して（例えば $\lambda^{-1}=1/4,\ 1/2,\ 1,\ 2,\ 4$ とか）$P_n \doteqdot n\lambda^{-1}$ である時，(26)によって a^n は $(1/e^c)^{\lambda}$ に収束する．$(1/e^c)^{\lambda}$ は1と0の間にある実数であり，λ^{-1} が大き

くなるにつれて，すなわち P_n が大きくなるにつれて大きくなっていく．また，$P_n \doteqdot n\lambda_n^{-1}$ で，$n\to\infty$ の時 $\lambda_n^{-1}\to\infty$ となるくらい P_n が“大きい”時，(26)及び(27)より a^n は1に収束する．また，この場合には

$$[1-C/P_n]^k \doteqdot [[1-C/P_n]^{P_n}]^{\lambda_n\left(\frac{k}{n}\right)} \geq [[1-C/P_n]^{P_n}]^{\lambda_n}, \quad k=1,\ 2,\ \cdots,\ n$$

であることより，すべての $k=1, 2, \cdots, n$ について a_k が1に収束することが示される．そして，そのことは，$n\to\infty$ の時，この様な a に対しては(19)や(20)は，(23)や(24)に収束していくことを意味する．しかしながら，例えば $a=1-C/_n$ である場合には $n\to\infty$ としても，(19)や(20)は(23)(24)には収束していかない．以上の様に，a^n の収束値は P_n をどう与えるか次第で著しく異なったものとなっているが，このことは(19)や(20)の表現を a の関数とみなした時，それは $a=1$ の近傍で不連続な“たちの悪い”関数となっていることを意味する．

ε_t の正規性を仮定し，$(y_1, y_2, \cdots, y_T)$ が観測値として与えられる時の尤度関数について考察してみよう．$f(y_t)$ は y_t の周辺分布を表わす確率密度関数とし，$f(y_t \mid y_{t-1}, \cdots, y_1)$ は $(y_{t-1}, \cdots, y_1)$ が与えられた時の y_t の条件付分布を表わす確率密度関数とすると，$(y_1, y_2, \cdots, y_T)$ の対数尤度関数 $g_T(a, \sigma^2)$ は，

$$g_T(a,\ \sigma^2)=\log f(y_1)+\sum_{t=2}^{T}\log f(y_t \mid y_{t-1},\ \cdots,\ y_1) \tag{28}$$

と表わされる．また，ε_t の正規性は

$$\log f(y_1)=-\frac{1}{2}\log 2\pi-\frac{1}{2}\log \sigma^2(1+a^2)-\frac{1}{2}\{\sigma^2(1+a^2)\}^{-1}y_1^2 \tag{29}$$

$$\begin{aligned}\log f(y_t \mid y_{t-1},\ \cdots,\ y_1)=&-\frac{1}{2}\log 2\pi-\frac{1}{2}\log[E\{y_t-P(y_t \mid y_{t-1},\ \cdots,\ y_1)\}^2]\\&-\frac{1}{2}[E\{y_t-P(y_t \mid y_{t-1},\ \cdots,\ y_1)\}^2]^{-1}\\&[y_t-P(y_t \mid y_{t-1},\ \cdots,\ y_1)]^2,\quad t=2,\ \cdots,\ T\end{aligned} \tag{30}$$

をもたらす．ここで $P(y_t \mid y_{t-1}, \cdots, y_1)$ は(19)において定義された $(y_{t-1}, \cdots, y_1)$ を用いての y_t の最適線形予測量である．(28) (30)は，(対数) 尤

度関数 $g_T(a, \sigma^2)$ が $P(y_t \mid y_{t-1}, \cdots, y_1)$, $t=2, \cdots, T$ の連続関数であることを意味する．よって，上で論じてきたことより $g_T(a, \sigma^2)$ は $a=1$ の近傍で不連続な“たちの悪い”関数となっていることが示される．

ところで，Tanaka and Sachell (1989) をはじめとして，反転不能移動平均過程の推定を扱っている文献においては，しばしば ε_t についての初期条件 $\varepsilon_0=0$ が課された場合の分析についても触れられている．その場合，(1)によって生成される y_t は $t\geq 1$ についてのみ定義されることになる．非定常時系列の場合と異なって，反転不能移動平均過程において $\varepsilon_0=0$ を仮定することの意味はあまり明瞭ではないが，$\varepsilon_0=0$ が仮定される時（$n=t-1$ とした時の）(14)からもわかる様に，$|a|<1$ の場合と異なり $a=1$ の場合には“初期値の効果” $a^t\varepsilon_0$ は無視できないものとなっている．よく知られている様に，反転可能移動平均仮定の最尤推定では，$\varepsilon_0=0$ が仮定される場合とされない場合とでは漸近的には差異がない．しかしながら，反転不能の場合，$\varepsilon_0=0$ の仮定の有無は最尤推定に実質的な影響を与えるものと思われる．$\varepsilon_0=0$ が仮定されると，a の値にかかわらず，$t\geq 1$ に対して

$$y_t+ay_{t-1}+a^2y_{t-2}+\cdots+a^{t-1}y_1=\varepsilon_t \tag{31}$$

が成立しているのでいわゆる反転可能性の問題は生じなくなる．この場合

$$P(y_t \mid y_{t-1}, \cdots, y_1)=-(ay_{t-1}+\cdots+a^{t-1}y_1), \quad t=2, \cdots, T \tag{32}$$

と解釈される．しかしながら，$P(y_t \mid y_{t-1}, \cdots, y_1)$ を a の関数とみる時 $P(y_T \mid y_{T-1}, \cdots, y_1)$ は，$T\to\infty$ である時 $a=1$ の近傍で不連続的な“たちの悪い”関数になっていることは，初期条件が仮定されていない場合と同様である．

以上考察してきた尤度関数の不連続性や“たちの悪さ”は，尤度関数の1階の微係数が(7)での様に $\sqrt{T}$ で標準化されないかもしれないことを，従って最尤推定量のパラメータ真値への収束の測度が $1/\sqrt{T}$ でないかもしれないことを，また，尤度方程式の解が漸近的にも複数個存在する可能性をも示唆する．

4. 評価関数の諸特性と各種最尤推定量の導出

(1)において観測値 $(y_1, y_2, \cdots, y_T)$ が与えられた時，(1)は

$$\underset{\sim}{y}(T)=A\underset{\sim}{\varepsilon}(T)+\underset{\sim}{b}\varepsilon_0 \tag{33}$$

と表わされる．ここで $\underset{\sim}{y}(T)=(y_1, y_2, \cdots, y_T)'$，$\underset{\sim}{\varepsilon}(T)=(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_T)'$ 及び $\underset{\sim}{b}=(-a, 0, \cdots, 0)'$ は $T\times1$ ベクトル，また $A=\{\alpha_{ij}\}$ は $\alpha_{ii}=1$, $\alpha_{i+1,i}=-a$ そしてその他の i, j について $\alpha_{ij}=0$ となる様な $T\times T$ 行列となっている．ε_t が正規分布する場合の $(y_1, y_2, \cdots, y_T)$ の対数尤度関数 $g_T(a, \sigma^2)$ は，(33)より

$$g_T(a, \sigma^2)=-\frac{T}{2}\log 2\pi-\frac{T}{2}\log\sigma^2-\frac{1}{2}\log|AA'+\underset{\sim}{b}\underset{\sim}{b}'|$$

$$-\frac{1}{2\sigma^2}\underset{\sim}{y}(T)'[AA'+\underset{\sim}{b}\underset{\sim}{b}']^{-1}\underset{\sim}{y}(T) \tag{34}$$

と表わされることになる．また，関係

$$[AA'+\underset{\sim}{b}\underset{\sim}{b}']^{-1}=(A[I-A^{-1}\underset{\sim}{b}(-1)\underset{\sim}{b}'A'^{-1}]A')^{-1}$$
$$=A'^{-1}(I-A^{-1}\underset{\sim}{b}\underset{\sim}{b}'A'^{-1}/(1+\underset{\sim}{b}'A'^{-1}A^{-1}\underset{\sim}{b}))A^{-1}$$

及び

$$1+\underset{\sim}{b}'A'^{-1}A^{-1}\underset{\sim}{b}=\sum_{j=0}^{T}a^{2j}=|AA'+\underset{\sim}{b}\underset{\sim}{b}'|$$

を用いて

$$\underset{\sim}{y}(T)'[AA'+\underset{\sim}{b}\underset{\sim}{b}']^{-1}\underset{\sim}{y}(T)$$
$$=\underset{\sim}{y}(T)'A'^{-1}A^{-1}\underset{\sim}{y}(T)-\underset{\sim}{y}(T)'(-1)A'^{-1}A^{-1}\underset{\sim}{b}\underset{\sim}{b}'A'^{-1}A^{-1}$$
$$(-1)\underset{\sim}{y}(T)\Big/\left(\sum_{j=0}^{T}a^{2j}\right) \tag{35}$$

が得られる．さらに

$$A^{-1}=\begin{bmatrix} 1 & & & & & \\ a & 1 & & 0 & & \\ a^2 & a & 1 & & & \\ \vdots & \vdots & \vdots & & & \\ a^{T-1} & a^{T-2} & a^{T-3} & \cdots & a & 1 \end{bmatrix}$$

$\underset{\sim}{b}' A'^{-1} A^{-1}(-1) = \left(a \sum_{j=0}^{T-1} a^{2j},\ a^2 \sum_{j=0}^{T-2} a^{2j},\ \cdots,\ a^{T-1} \sum_{j=0}^{1} a^{2j},\ a^T\right)$ を用いて

$$\underset{\sim}{y}(T)'[AA'+\underset{\sim}{b}\underset{\sim}{b}']^{-1}\underset{\sim}{y}(T)$$
$$=\sum_{t=1}^{T-1}\left(y_{t+1}+\sum_{j=1}^{t} a^j y_{t+1-j}\right)^2 + y_1^2 - \left[\sum_{t=1}^{T} a^t\left(\sum_{j=0}^{T-t} a^{2j} y_t\right)\right]^2 \Big/ \left(\sum_{j=0}^{T} a^{2j}\right) \tag{36}$$

であることが示される．(1)の係数パラメータ a の真値を a_0 とする時，Anderson and Takemura (1986) に従って，a_0 を推定するための２つの評価関数 $M_1(a)$ 及び $M_2(a)$ を紹介する．

$$M_1(a)=\sum_{t=1}^{T-1}\left(y_{t+1}+\sum_{j=1}^{t} a^j y_{t+1-j}\right)^2 \Big/ T + y_1^2 \Big/ T - \left[\sum_{t=1}^{T} a^t\left(\sum_{j=0}^{T-t} a^{2j} y_t\right)\right]^2$$
$$\Big/ T\left(\sum_{j=0}^{T} a^{2j}\right) \tag{37}$$

$$M_2(a)=M_1(a)\left\{\sum_{j=0}^{T} a^{2j}\right\}^{1/T} \tag{38}$$

ここで，$M_2(a)$ は実は尤度関数に対応するものとなっており，マイナスの対数尤度関数 $-g_T(a, \sigma^2)$ に σ^2 の最尤推定量 $\hat{\sigma}^2=\underset{\sim}{y}(T)'[AA'+\underset{\sim}{b}\underset{\sim}{b}']^{-1}\underset{\sim}{y}(T)/T$ を σ^2 の代わりに代入して，a に依存しない定数項部分を無視したものが $\log M_2(a)$ となっている．それ故，a のパラメータ空間 $[-1, 1]$ において，$g_T(a, \sigma^2)$ を最大にする様な a と $M_2(a)$ を最小にする a とは同値なものであり，$M_2(a)$ を最小にする a を $\hat{a}_{ML}$ と記すことにすればそれは a_0 の最尤推定量ということになる．一方，$M_1(a)$ は $M_2(a)$ から $|AA'+\underset{\sim}{b}\underset{\sim}{b}'|^{1/T}=\{\Sigma_{j=0}^{T} a^{2j}\}^{1/T}$ 部分を取り除いたものに相当する．$M_1(a)$ を最小にする a を $\hat{a}_{LS}$ と記すことにすれば，(36)から明らかになる様に，それは $\underset{\sim}{y}(T)'[AA'+\underset{\sim}{b}\underset{\sim}{b}']^{-1}\underset{\sim}{y}(T)/T$ を最小にする a ということになるので，Anderson and Takemura (1986) では最小２乗推定量と呼ばれている．

ところで

$$a^2\underset{\sim}{y}(T)'[AA'+\underset{\sim}{b}\underset{\sim}{b}']^{-1}\underset{\sim}{y}(T)=\underset{\sim}{y}(T)'[HH'+\underset{\sim}{d}\underset{\sim}{d}']^{-1}\underset{\sim}{y}(T)$$

$$a^{-2}\left\{\sum_{j=0}^{T}a^{2j}\right\}^{1/T}$$

ここで $\underset{\sim}{d}=(-a^{-1},\ 0,\ \cdots,\ 0)'$ は $T\times 1$ ベクトル，また $H=\{\eta_{ij}\}$ は $\eta_{ii}=1,\ \eta_{i+1,\,i}=-a^{-1}$ そしてその他の i, j について $\eta_{ij}=0$ となる様な $T\times T$ 行列，であることに注意すると，

$$M_1(a)=a^{-2}M_1(a^{-1}),\quad \text{注意の } a\in[-1,\ 1] \text{ に対して} \tag{39}$$

$$M_2(a)=M_2(a^{-1}),\quad \text{任意の } a\in[-1,\ 1] \text{ に対して} \tag{40}$$

であることが示される．(40)については，2節(11)式において既に示されているともいえる．2節(11)における $M(a)$ がマイナスの尤度数関数であったことを思い出せばよい．(39)は

$$\left.\frac{dM_1}{da}\right|_{a=1}<0,\quad \left.\frac{dM_1}{da}\right|_{a=-1}>0 \tag{41}$$

を意味し，(40)は

$$\left.\frac{dM_2}{da}\right|_{a=1}=0,\quad \left.\frac{dM_2}{da}\right|_{a=-1}=0 \tag{42}$$

を意味する．a のパラメータ空間は［－1，1］であることを思い出すならば，(41)は $M_1(a)$ が $a=1$ あるいは $a=-1$ で常に極小になっていること，言い換えると $M_1(a)$ が $a=1$ あるいは $a=-1$ で極小になる確率は 1 であることを意味する．また，(42)は $M_2(a)$ が $a=1$ あるいは $a=-1$ で常に極小か極大（すなわち，極値）になっていることを，そしてその確率が 1 であることを意味している．$M_2(a)$ が $a=1$ あるいは $a=-1$ で極小となるためには，その 2 次微係数 $\dfrac{d^2M_2}{da^2}$ の $a=1$ または $a=-1$ での符号が正となる必要があることは明らかであるが，(1)の y_t が反転不能である場合（a の真値 a_0 が $|a_0|=1$ である場合）でのその確率（とりわけ，極限確率）の導出は，この分野における一つの目標とされてきた．$\dfrac{d^2M_2}{da^2}$ が $a=1$ あるいは $a=-1$ で正となる確率（の極限）は，Sargan and Bhargava (1983), Anderson and Takamura (1986) 及び Tanaka and Sachell (1989) によって独立に評価されている．しかしながら，ここで注意しなければならない点は，$M_1(a)$ が確率 1 で $a=1$ あるい

は $a=-1$ で極小値になっていることが $\hat{a}_{LS}$ が1あるいは−1であることを意味するものではなく，［−1，1］区間のある内点（それは，一致性の観点から−1や1の近傍）に別の極小値が存在する可能性を排除するものではないということである．その場合，その極小値はパラメータ空間の内点であるので，その点において M_1 の1階の微係数はゼロとなり，いわゆる尤度方程式の様な形態の表現式が得られ，それが a に関して一意的に解けるならば，解くことによって $\hat{a}_{LS}$ を求めたり評価したりすることができる．$M_2(a)$ についても同様で，［−1，1］の内点に別の極小値が存在し得る．しかしながら，$M_1(a)$ や $M_2(a)$ が1や−1で極小値になる確率が0でないという性質は，関数 $M_1(a)$ 及び $M_2(a)$ のある種の“たちの悪さ”を予感させるものがある．

3節において既に言及した様に，反転不能移動平均過程を扱っている場合，しばしば ε_t についての初期条件 $\varepsilon_0=0$ が課される．その場合，次の様な評価関数が与えられることになる．

$$M_3(a)=\sum_{t=1}^{T-1}\left(y_{t+1}+\sum_{j=1}^{t}a^j y_{t+1-j}\right)^2\Big/T \tag{43}$$

$M_3(a)$ は(37)で与えられている $M_1(a)$ の第1項に相当し，$\varepsilon_0=0$ が仮定された時の尤度関数に対応するものになっている．すなわち，マイナスの対数尤度関数において，σ^2 にその最尤推定量 $\tilde{\sigma}^2=\underline{y}(T)'A'^{-1}A^{-1}\underline{y}(T)/T$ を代入し，a に関係しない部分を除いたものが $\log M_3(a)$ となっている．Anderson and Takemura (1986) に従って，$M_3(a)$ を最小にする様な a を $\hat{a}_{CLS}$ と記し，条件付最小2乗推定量と呼ぶことにする．$\left.\frac{dM_3}{da}\right|_{a=1}<0$ である漸近確率が，$a_0=1$ 及び $\varepsilon_0=0$ が仮定された時に $\chi_1^2<1$ である様な確率に等しくなることが容易に示される．ここで χ_1^2 は自由度1の χ^2 分布に従う確率変数を表わしている，また，この証明については，例えば Tanaka and Sachell (1989) を見よ．

5．各種最尤推定量の漸近的性質と新たな推定量の構築

前節において導入された各種の最尤推定量 $\hat{a}_{LS}$, $\hat{a}_{ML}$ 及び $\hat{a}_{CLS}$ の何らか

の漸近的性質を得るために，以下の様な量を定義しよう．

$$X_1(a)=\sum_{t=1}^{T-1}\varepsilon_{t+1}\left(\sum_{j=0}^{t-1}a^j\varepsilon_{t-j}\right),\ X_2(a)=\sum_{t=2}^{T-1}\varepsilon_{t+1}\left(\sum_{j=0}^{t-2}(j+1)a^j\varepsilon_{t-1-j}\right),$$

$$X_3(a)=\sum_{t=3}^{T-1}\varepsilon_{t+1}\left(\sum_{j=0}^{t-3}(j+1)(j+2)a^j\varepsilon_{t-2-j}\right),\ Y(a)=\sum_{t=1}^{T-1}\left(\sum_{j=0}^{t-1}a^j\varepsilon_{t-j}\right)^2,$$

$$Z(a)=\sum_{t=2}^{T-1}\left(\sum_{j=0}^{t-1}a^j\varepsilon_{t-j}\right)\left(\sum_{j=0}^{t-2}(j+1)a^j\varepsilon_{t-1-j}\right),\ W_0(a)=a^2\left(\sum_{t=0}^{T-1}a^t\varepsilon_t\right)^2\Big/\left(\sum_{j=0}^{T}a^{2j}\right),$$

$$W_1(a)=\left\{\sum_{t=1}^{T-1}a^{t+1}\left(\sum_{j=0}^{t-1}a^j\varepsilon_{t-j}\right)\right\}^2\Big/\left(\sum_{j=0}^{T}a^{2j}\right),\ W_2(a)=\left\{\sum_{t=1}^{T-1}a^{t+1}\left(\sum_{j=0}^{t-1}a^j\varepsilon_{t-j}\right)\right\}$$

$$\left\{\sum_{t=1}^{T}a^t\varepsilon_t\right\}\Big/\left(\sum_{j=0}^{T}a^{2j}\right)$$

また，(43)の y_t に $\varepsilon_t-a_0\varepsilon_{t-1}$ を代入することによって

$$M_3(a)=\sum_{t=1}^{T-1}\varepsilon_{t+1}^2/T+2(a-a_0)X_1(a)/T+(a-a_0)^2Y(a)/T \tag{44}$$

が得られる．同様にして

$$M_1(a)=M_3(a)-W_0(a)/T-(a-a_0)^2W_1(a)/T+2(a-a_0)$$
$$W_2(a)/T+f(a)/T \tag{45}$$

ここで $f(a)=a_0\Big[\varepsilon_0^2\left\{\sum_{t=0}^{T-1}a^{2t}\right\}-2\varepsilon_0(a-a_0)\left\{\sum_{t=1}^{T-1}a^t\left(\sum_{j=0}^{t-1}a^j\varepsilon_{t-j}\right)\right\}$

$$-2\varepsilon_0\left\{\sum_{t=0}^{T-1}a^t\varepsilon_{t+1}\right\}\Big]\Big/\left(\sum_{t=0}^{T}a^{2t}\right).$$

閉区間［－1，1］における任意の a に対して，

$$a=\frac{C}{P_T}\quad あるいは\quad a=-1+\frac{C}{P_T} \tag{46}$$

と置くことができる．但し C は適当な非負の実数，また $\{P_T\}$ は任意の T について $P_T=P\geq\frac{1}{2}C$ か $T\to\infty$ の時 $P_T=\infty$ となる様な正の整数列とする．その時，その様な P_T とともに，$\hat{a}_{LS}$, $\hat{a}_{ML}$ 及び $\hat{a}_{CLS}$ の（一致性を中心とした）漸近的結果を得るために必要な補題を３つばかり述べることができる．

補題1. (46)で与えられる a, P_T に対して

(i) $\lim_{T\to\infty} P_T=\infty$ ならば, $\lim_{T\to\infty} a^{2P_T}=e^{-2c}$

(ii) $\lim_{T\to\infty} S_T/P_T=0$ となる様な任意の正の整数列 $\{S_T\}$ に対して,
$a^{2S_T}=1+O(S_T P_T^{-1})$

(iii) $\lim_{T\to\infty} P_T/S_T=0$ となる様な任意の正の整数列 $\{S_T\}$ に対して,
$\lim_{T\to\infty} a^{2S_T}=0$

この補題の結果はきわめて基本的なものであり, 証明を与える必要性は全くないが, 本稿では3節(26) (27)式において同様の結果がより詳細に論じられている. この補題を用いて次の結果が得られる.

補題2. (46)で与えられる a, P_T に対して, $Q_T=\min(P_T, T)$ とすると

(i) $X_1(a)=O_p(T^{1/2}Q_T^{1/2})$

(ii) $X_2(a)=O_p(T^{1/2}Q_T^{3/2})$

(iii) $X_3(a)=O_p(T^{1/2}Q_T^{5/2})$

(iv) $Y(a)=O_p(TQ_T)$

(v) $Z(a)=O_p(TQ_T^2)$

証明： $Y_i(a)=\sum_{t=1}^{T-1} d_{i,t}(a)$, $i=1, 2, 3$ と置くことにしよう.

ここで $d_{1,t}(a)=\Sigma_{j=0}^{t-1}a^{2j}$, $d_{2,t}(a)=\Sigma_{j=0}^{t-1}(j+1)^2a^{2j}$
及び $d_{3,t}(a)=\Sigma_{j=0}^{t-3}(j+1)^2(j+2)^2a^{2j}$.

補題1より

$$d_{i,t}(a)=O(\min(t^{1+2(i-1)}\ Q_T^{1+2(i-1)})),\quad t=1, 2,\cdots, T,\ i=1, 2, 3$$

であることが示されるが, それは

$$Y_i(a)=O(TQ_T^{1+2(i-1)}),\quad i=1, 2, 3$$

を意味する. $EX_i(a)=0$ 及び $E\{X_i(a)\}^2=Y_i(a)/\sigma^4$, $i=1, 2, 3$ であることに注意すれば, (i)(ii)(iii)が直ちに得られる. また, $EY(a)=Y_1(a)/\sigma^2$ 及び $Y(a)$ は常に非負であることは, (iv)を意味する. さらに,

$$\left(\sum_{j=0}^{t-1} a^j \varepsilon_{t-j}\right)^2 = O_p(d_{1,t}(a)),\quad t=2,\ \cdots,\ T$$

$$\left(\sum_{j=0}^{t-2} (j+1) a^j \varepsilon_{t-1-j}\right)^2 = O_p(d_{2,t}(a)),\quad t=3,\ \cdots,\ T$$

であることより，

$$\left(\sum_{j=0}^{t-1} a^j \varepsilon_{t-j}\right)\left(\sum_{j=0}^{t-2} (j+1) a^j E_{t-1-j}\right) = O_p([d_{1,t}(a)]^{1/2}[d_{2,t}(a)]^{1/2}),$$

$t=3,\ \cdots,\ T$

が得られるが，それは(V)を意味する．　　　　　　　　　　　　　証明終わり

補題3.　(46)で与えられる a, P_T に対して，$Q_T = \min(P_T,\ T)$ とすると，

(i)　$|a_0| = 1$ である時，$f(a) = O_p(Q_T^{-1/2})$

(ii)　$W_0(a) = O_p(1)$

(iii)　$W_1(a) = O_p(Q_T)$

(iv)　$W_2(a) = O_p(Q_T^{1/2})$

(v)　$\left\{\sum_{t=0}^{T} a^{2+}\right\}^{1/T} = 1 + O(\log Q_T / T)$

証明：補題2の証明において用いられた論法によって

$$\sum_{t=0}^{T} a^{2+} = O(Q_T),\quad E\left\{\sum_{t=1}^{T} a^t \varepsilon_t\right\}^2 = O(Q_T)$$

$$E\left\{\sum_{t=1}^{T-1} a^t \left(\sum_{j=0}^{t-1} a^j \varepsilon_{t-j}\right)\right\}^2 = O(Q_T^2)$$

が得られる．それ故，

$$\left(\sum_{t=1}^{T} a^t \varepsilon_t\right)^2 = O_p(Q_T),\quad \left\{\sum_{t=1}^{T-1} a^t \left(\sum_{j=0}^{t-1} a^j \varepsilon_{t-j}\right)\right\}^2 = O(Q_T^2)$$

$$\left(\sum_{t=1}^{T} a^t \varepsilon_t\right)\left\{\sum_{t=1}^{T-1} a^t \left(\sum_{j=0}^{t-1} a^j \varepsilon_{t-j}\right)\right\} = O_p(Q_T^{3/2})$$

(ii)(iii)(iv)はこれらの結果より直ちに得られる．$|a_0| = 1$ の時には $a - a_0 = O(P_T^{-1})$ であるので，(i)も示されることになる．また，

$$\left\{\sum_{t=0}^{T} a^{2t}\right\}^{1/T}=1+O\left(\frac{\log\{\Sigma_{t=0}^{T}a^{2t}\}}{T}\right)$$

であることを用いて，(iv)と(v)も示される． <u>証明終わり</u>

これらの補題から，以下の二つの定理が打ち立てられる．

<u>**定理1.**</u>　(1)において $|a_0|=1$ を仮定すると，

(i)　$\hat{a}_{LS}-a_0=O_p(T^{-1})$ (47)

$\hat{a}_{ML}-a_0=O_p(T^{-1})$ (48)

(ii)　$\varepsilon_0=0$ が仮定される時

$\hat{a}_{CLS}-a_0=O_p(T^{-1})$ (49)

<u>**証明：**</u>最初に(ii)を示すことにする．(46)から明らかな様に，$|a_0|=1$ の時 $|a-a_0|=C/p_T$ と書けることより，$a_0=1$ の時

$$M_3(a)=\sum_{t=1}^{T-1}\varepsilon_{t+1}^2/\mathrm{T}-\frac{2\mathrm{C}}{\mathrm{P_T}}\frac{\mathrm{Q_T^{1/2}}}{\mathrm{T^{1/2}}}\frac{\mathrm{X_1(a)}}{\mathrm{T^{1/2}Q_T^{1/2}}}+\frac{\mathrm{C^2}}{\mathrm{P_T}}\frac{\mathrm{Q_T}}{\mathrm{P_T}}\frac{\mathrm{Y(a)}}{\mathrm{TQ_T}}$$

$a_0=-1$ の時

$$M_3(a)=\sum_{t=1}^{T-1}\varepsilon_{t+1}^2/T+\frac{2C}{P_T}\frac{Q_T^{1/2}}{T^{1/2}}\frac{X_1(a)}{T^{1/2}Q_T^{1/2}}+\frac{C^2}{P_T}\frac{Q_T}{P_T}\frac{Y(a)}{TQ_T}$$

となることが示される．ここで補題2によって

$$\frac{X_1(a)}{T^{1/2}Q_T^{1/2}}=O_p(1),\quad \frac{Y(a)}{TQ_T}=O_p(1)$$

であることに注意せよ．$\lim_{T\to\infty}P_T/T=0$ 及び $\mathrm{P_T}>\mathrm{C}$ となる様な P_T に対して，$M_3(a)$ の第3項は $O_P(P_T^{-1})$，第2項は $O_P(P_T^{-1/2}T^{-1/2})$ であり，従って第3項の方のオーダーが大きくなる．第3項は常に正であるので，$M_3(a)>M_3(a_0)$ が成立していることになり，$a-a_0=O(P_T^{-1})$ となる様などんな a も決して $M_3(a)$ を最小にしないことになる．他方，$\lim_{T\to\infty}T/P_T=0$ となる様な $|a|=1-C/P_T$ に対しては，$M_3(a)$ の第3項は $O_P(P_T^{-1})$，第2項は $O_P(P_T^{-2}T)$ であり，第2項のオーダーの方が第3項より大きくなる．補題1は，

$$\frac{X_1(a)}{T^{1/2}Q^{1/2}}=\frac{X_1(a_0)}{T^{1/2}Q^{1/2}}+O_p(T/P_T)$$

を意味するが，そのことより $M_3(a)$ の中で第1項を除けば，$\dfrac{2C}{P_T}\dfrac{X_1(a_0)}{T}$ のオーダーが最も大きくなることがわかる．$M_3(a)$ は $|C|$ が増加するにつれて減少していくので，この様な a についても $M_3(a)$ は決して最小にはならないことになる．この様にして

$$\hat{a}_{CLS}=1-C/T \quad \text{あるいは} \quad -1+C/T$$

が示されることになる．

(i)については，$a_0=1$ の時

$$M_1(a)=M_3(a)-W_0(a)/T-\frac{2C}{P_T}\frac{Q_T^{1/2}}{T}\frac{W_2(a)}{Q_T^{1/2}}-\frac{C^2}{P_T}\frac{1}{T}\frac{Q_T}{P_T}\frac{W_1(a)}{Q_T}$$
$$+O_p(T^{-1}Q_T^{-1/2})$$

という表現が得られ，また，$a_0=-1$ の時の表現は $a_0=1$ の時の表現において C を $-C$ と取り換えることによって得られる．$\lim_{T\to\infty}P_T/T=0$ 及び $P_T>C$ なる任意の P_T，$a=1-\dfrac{C}{P_T}$ に対して，補題2及び3より，$M_1(a)$ の中で $M_3(a)$ の第1項を除けば最もオーダーが大きくなる項は，

$$\frac{C^2}{P_T}\frac{Q_T}{P_T}\frac{Y(a)}{TQ_T}=O_p(P_T^{-1})$$

であることが示されるが，この項は常に正であるのでそのことは $M_1(a)>M_1(a_0)$ であることを意味する．他方，$\lim_{T\to\infty}T/P_T=0$ となる様な任意の $a=1-C/P_T$ に対して，$M_1(a)$ の中で最もオーダーが大きくなる項は，$M_3(a)$ の第1項を除けば $W_0(a)/T$ であり，それは a^2 が1に近づくにつれて増加するものであるので，$M_1(a)>M_1(a_0)$ が示されることになる．よって $\hat{a}_{LS}-a_0=O_p(T^{-1})$．同様の論法によって，$\hat{a}_{ML}-a_0=O_p(T^{-1})$ も示される．

証明終わり

定理2. (1)において $|a_0|=1$ を仮定すると

(i) $a-a_0=C/T$ なる様な任意の a に対して

$$\operatorname*{Plim}_{T\to\infty}\left|\frac{dM_i}{da}\Big|_a-\left\{\frac{dM_i}{da}\Big|_{a_0}-C\frac{d^2M_i}{da^2}\Big|_{a_0}\right\}\right|\neq 0,\quad i=1,\ 2 \tag{50}$$

ここで C は適当な定数とする.

(ii) $\varepsilon_0=0$ が仮定される時, $a-a_0=C/T$ なる任意の a に対して

$$\operatorname*{Plim}_{T\to\infty}\left|\frac{dM_3}{da}\Big|_a-\left\{\frac{dM_3}{da}\Big|_{a_0}-C\frac{d^2M_3}{da^2}\Big|_{a_0}\right\}\right|\neq 0 \tag{51}$$

ここで C は適当な定数とする.

証明:(ii)のみ示す.(i)も(ii)と同様にして示される.補題1より, $a=1-C/T$ である様な任意の a に対して

$$\lim_{T\to\infty} a^T=e^{-c}\neq 1$$

であることより

$$\operatorname*{Plim}_{T\to\infty}\left|\frac{X_i(a)}{T^i}-\frac{X_i(a_0)}{T^i}\right|\neq 0,\quad i=1,\ 2,\ 3$$

$$\operatorname*{Plim}_{T\to\infty}\left|\frac{Y(a)}{T^2}-\frac{Y(a_0)}{T^2}\right|\neq 0,\quad \operatorname*{Plim}_{T\to\infty}\left|\frac{Z(a)}{T^3}-\frac{Z(a_0)}{T^3}\right|\neq 0$$

が得られるが,これらの結果は(51)をもたらす. 証明終わり

これらの定理は, a_0 の3つの推定量 $\hat{a}_{LS}$, $\hat{a}_{ML}$ 及び $\hat{a}_{CLS}$ には,標準的漸近理論は適用できないことを結論づけるものである.定理1では,3つの推定量は一致性を持っているけれども,それらの a_0 への収束の速度は標準的漸近理論におけるもの(それは $T^{-1/2}$)より速く, T^{-1} の速さで収束していくことが述べられている.また,定理2は,推定量を導出・評価するための方程式(尤度方程式に相当する)

$$\frac{dM_i}{da}\Big|_a=0,\quad i=1,\ 2,\ 3$$

が a に関して,漸近的にも(勿論,小標本においても)高度に非線形であり,それ故 T が充分大の時でさえ複数個の解が存在し得ることを意味している.そのことは,これらの推定量の漸近分布を導出することがきわめて困難であることを示唆している.

a_0 の新たな推定量を構成するために，始めに，y_t を

$$z_t=y_t+(-1)^t u_0+v_0 \tag{52}$$

に変換する．ここで，$u_0\neq0$ 及び $v_0\neq0$ は適当に与えられる定数とする．次に，以下の様な評価関数 $R(a, u, v, w)$ を導入しよう．

$$R(a, u, v, w)=\sum_{t-1}^{T-1}\left[\sum_{j=0}^{t} a^j z_{t+1-j}-(t+1)(-1)^{t+1}u-(t+1)v-w\right]^2\Big/$$

$$T+z_1^2/T-\left[\sum_{t=1}^{T} a^t\left(\sum_{j=0}^{T-t} a^{2j}\right)(z_t-t(-1)^t u-tv-w)\right]\Big/$$

$$T\left(\sum_{j=0}^{T} a^{2j}\right) \tag{53}$$

また，$\varepsilon_0=0$ が仮定される場合には以下の様なものが提案されることになる。

$$S(a, u, v, w)=\sum_{t=1}^{T-1}\left[\sum_{j=0}^{t} a^j z_{t+1-j}-(t+1)(-1)^{t+1}u-(t+1)v-w\right]^2\Big/T \tag{54}$$

評価関数(53)(54)においては，それぞれ関数値を最小にする様なパラメータ (a, u, v, w) の値が求められることになるが，(a, u, v, w) のパラメータ空間は R^4 とされる．すなわち，(1)の係数 a も含めてパラメータ空間に対しての何らの制限もなされないことになる．ε_t の正規性の仮定を取り除き，そのかわりに $E\varepsilon_t^4<\infty$ を仮定することによって，$R(a, u, v, w)$ を最小にする様な a である $\hat{a}_R$ と $S(a, u, v, w)$ を最小にする a である $\hat{a}_S$ の漸近的性質について，次の様な結果が得られることになる．

定理3. (1)において $E\varepsilon_t^4<\infty$ を仮定しよう．その時

(i) $a_0=1$ である時，$T^{5/2}(\hat{a}_R-a_0)$ の漸近分布は平均0，分散 σ^2/v_0^2 の正規分布となる．また，$\varepsilon_0=0$ が仮定された時，$a_0=1$ について，$T^{5/2}(\hat{a}_S-a_0)$ は漸近的に平均0，分散 σ^2/v_0^2 の正規分布に従う．

(ii) $a_0=-1$ である時，$T^{5/2}(\hat{a}_R-a_0)$ の漸近分布は平均0，分散 σ^2/u_0^2 の正規分布となる．また，$a_0=-1$ で $\varepsilon_0=0$ が仮定された時，$T^{5/2}(\hat{a}_S-a_0)$ の漸近分布は $T^{5/2}(\hat{a}_R-a_0)$ のそれと等しくなる．

(iii)　$|a_0|<1$ である時，$T^{1/2}(\hat{a}_R-a_0)$ 漸近分布は $T^{1/2}(\hat{a}_{ML}-a_0)$ のそれと等しくなり，また，$|a_0|<1$ で $\varepsilon_0=0$ が仮定された時，$T^{1/2}(\hat{a}_S-a_0)$ の漸近分布は $T^{1/2}(\hat{a}_{CLS}-a_0)$ のそれと等しくなる．

証明：(i)の $\varepsilon_0=0$ が仮定された時の $\hat{a}_S$ についての結果のみ示す．(i)におけるその他の結果や(ii)については，それと同様の方法で示すことができる．また，(iii)の結果は標準的漸近理論の適用によって得られることができる．

(52)における z_t の定義より

$$\sum_{j=0}^{t} a^j z_{t+1-j}-(t+1)\ [(-1)^{t+1}\mathrm{u}+\mathrm{v}]-\mathrm{w}$$

$$=\varepsilon_{t+1}+(a-a_0)\left(\sum_{j=0}^{t-1} a^j\varepsilon_{t-j}\right)+F_t(a),\quad t=1,\ \cdots,\ T-1$$

が得られる．ここで $F_t(a)=(\Sigma_{j=0}^{t}a^j)v_0+(-1)^{t+1}[\Sigma_{j=0}^{t}(-1)^j a^j]u_0-(t+1)$ $[(-1)^{t+1}u+v]-wF_t(a)$ はさらに

$$F_t(a)=\left[\left(\sum_{j=0}^{t} a^j\right)-(t+1)\right]v-\left(\sum_{j=0}^{t} a^j\right)(v-v_0)$$

$$+(-1)^{t+1}\left[\sum_{j=0}^{t}(-1)^j a^j\right]u_0-(t+1)[(-1)^{t+1}u+v]-w$$

$$=(a-a_0)\left\{\sum_{j=0}^{t-1}(t-j)a^j\right\}v-\left\{\sum_{j=0}^{t} a^j\right\}(v-v_0)$$

$$+\left\{(-1)^{t+1}\left[\sum_{j=0}^{t}(-1)^t a^t\right]u_0-w\right\}-(t+1)(-1)^{t+1}u$$

と表されることになるので，そのことは，$\sum_{t=1}^{T-1}\varepsilon_{t+1}^2/T$ を除けば

$$2(a-a_0)v\sum_{t=1}^{T-1}\left\{\sum_{j=0}^{t-1}(t-j)a^j\right\}\varepsilon_{t+1}/T$$

と $(a-a_0)^2v^2\sum_{t=1}^{T-1}\left\{\sum_{j=0}^{t-1}(t-j)a^j\right\}^2\Big/T$ のオーダーが $S(a,\ u,\ v,\ w)$ の項の中で

最も大きくなっていることを意味する．それ故，$\hat{a}_S-a_0=O_p(T^{-5/2})$ が直ちにでてくる．また，これらの2つの項を評価すれば，$\sqrt{T}(\hat{a}_S-a_0)$ が漸

近的に平均0，分散 σ^2/v_0^2 の正規分布に従うことは容易に示すことができる．
証明終わり

この定理によって，y_t が反転不能であっても反転可能であってもその漸近分布が常に正規分布となる様な推定量の存在が示されたことになる．

6. まとめ並びに課題

本稿では，反転不能な1階の移動平均過程における係数パラメータの推定の問題が論じられてきた．係数パラメータの最尤推定量が通常より“速く”その真の値に収束していくこと，また，その漸近分布を評価することがきわめて困難であることが，過去の文献とは異なるアプローチによって説明されてきた．本稿でのアプローチによって明らかとなったその主たる原因は，パラメータの識別のために課されることになるパラメータ空間の制限によって，係数パラメータの真の値がパラメータ空間の境界上にくること，最適線形予測における過去の系列の係数値が幾何級数的に減少していかないことに関連して，尤度関数がパラメータの真の値の近傍で非常に“たちの悪い”不連続な関数になっていることに集約されるものであった．

既に述べてきた様に，最尤推定量については，一致性並びにその収束の速度が通常より“速い”ということだけがわずかに証明されているに過ぎず，その漸近分布の評価もなされていないし，到底その漸近的性質の全貌が明らかにされているとはいい難い．漸近的正規性に基いた標準的統計理論が適用できず，広い範囲の確率過程に対して統一的・画一的な理論を組み立てていくことが困難であることは，非定常過程の推定の問題と似通っているが，反転不能移動平均過程の場合の方がはるかに厄介であり，確立された結果もごくわずかなものに過ぎない．

ところで，標準的漸近理論が適用できないということは，（最尤推定量の）漸近的有効性は何ら証明されてていことを意味する．我々がパラメータの推定量として最尤推定量を求める理由は，何といってもその漸近的有効性が保証されているということにあるのであるが，この場合それが保証

されていないということは，最尤推定量に固執しなければならない必然性は何らないということを示唆する．他に望ましい推定量を考え出そうとする試みも，この場合正当化されることになるであろう．

本稿の5節の最後に，我々はパラメータ空間の制限を取り除くことによって，また，定数項部分を考慮することによって，尤度関数に関するものとは別の評価関数を構築し，それを最小にすることによって得られる新たな推定量を提案しその漸近的正規性を証明した．しかしながら，最尤推定量とは異なり，この推定量は1階の移動平均過程でのみ定義可能であり，高次の過程やベクトル過程に拡張定義することはできない．そういう意味で，もっと一般的に広い範囲で適用可能である様な推定量・推定手法を考えていくことが必要であろう．といっても，それよりも前に高次の反転不能移動平均過程や反転不能ベクトル過程自体の統計的性質や最尤推定量の性質がもう少し調査されなければならないと思われる．

はしがきにおいても触れた様に，本稿も含め他の文献においても高次の反転不能な移動平均過程の統計的分析は，全くといってもいいくらい扱われていない．また，ベクトル次元の反転不能移動平均過程についても，co-integration との深い関連性についてはよく知られているが，反転可能性や移動平均過程の推定という観点から論じた文献は全んど存在していない．恐らく，反転不能な移動平均過程の推定については，反転可能な場合の様な統一的・画一的な結果はとても期待できないであろうし，過程の次数や次元が高くなればなるほど複雑なものとなり，また，パラメータに強く依存したものとなっているであろうと思われる．以上述べてきた様に，反転不能な移動平均過程の推定問題では，得られている結果はごくわずかなものに過ぎず，今後取り組み解決していかなければならない厄介な課題は数多く存在しているのである．

参　考　文　献

Anderson, T. W. and Takemura, A. (1986) Why Do Noninvertible Estimated Moving Average Occur?, Journal of the Time Series Analysis, Vol. 7, 235-254.

Davidson, J. E. H. (1981) Problems with the Estimation of Moving Average Processes, Journal of Econometrics, Vol. 16, 295-310.

Plosser, C. I. and Schwert, G. W. (1977) Estimation of a Non-invertible Moving Average Processes: The Case of Overdifferencing, Journal of Econometrics, Vol. 6, 199-224.

Sargan, J. D. and Bhargawa, A. (1983) Maximum-likelihood Estimation of Regression Models with First-order Moving Average Errors When the Root Lies on the Unit Circle, Econometrica, Vol. 51, 799-820.

Tanaka, K. and Sachell, S. E. (1989) Asymptotic Properties of the Maximum-likelihood and Nonlinear Least-squares Estimators for Noninvertible Moving Average Models, Econometric Theory, Vol. 5, 333-353.