数理解析研究所講究録
第723巻 1990年 1-14

長期記憶モデルにおける二・三の問題

和歌山大・経　矢島美寛 (Yoshihiro Yajima)

1. 序

観測値間の系列相関の強い、いわゆる「長期記憶モデル」においては、従来のARMAモデル等、系列相関の弱いモデルについて成立した統計的性質をそのまま適用することはできない。ここではまず一例として、誤差項が長期記憶モデルにしたがう回帰モデルの最小二乗推定量(LSE)について論じる。後に今後解決すべきと思われる問題のうち、いくつかを紹介する。

2. 回帰モデル

モデル　$y_t = x_t' \beta + \varepsilon_t$

$x_t = (x_{t1}, \cdots, x_{tk})'$　説明変数

$\beta = (\beta_1, \cdots, \beta_k)'$　未知パラメータ

$\{\varepsilon_t\}$: 誤差項(定常過程)　$E\varepsilon_t = 0$

$$f(\lambda) = f^*(\lambda) / |1 - e^{i\lambda}|^{2d}, \quad 0 < d < 1/2$$

$$f^*(\lambda): \text{ positive continuous}$$

以下では $f(\lambda)$を $\{\varepsilon_t\}$ のスペクトル密度とする．$f(\lambda)$ は $\lambda \to 0$ のとき無限大に発散する．したがって低周波数の波が優勢な密度である．この性質に因んで、$\{\varepsilon_t\}$は「長期記憶モデル」にしたがうと言う。

<u>Grenander's conditions</u>

$$a_{ij}^T(h) \underset{d}{=} \sum_{t=1}^{T-h} x_{t+h,i}\, x_{tj}, \quad h=0,1,\cdots,$$

$$\underset{d}{=} \sum_{t=1-h}^{T} x_{t+h,i}\, x_{tj}, \quad h=0,-1,\cdots.$$

(G.1) $\qquad a_{ii}^T(0) \to \infty \ (T \to \infty), \quad i=1,\cdots,K.$

(G.2) $\qquad \lim_{T\to\infty} x_{T+1,i}^2 / a_{ii}^T(0) = 0, \quad i=1,\cdots,K.$

(G.3) $\qquad T \to \infty$ のとき

$$a_{ij}^T(h) / \{a_{ii}^T(0)\, a_{jj}^T(0)\}^{1/2} \ (\underset{d}{=} r_{ij}^T(h))$$

の極限が存在．$i,j=1,2,\cdots,K, \quad h=0,\pm 1,\cdots.$

(G.4) $\qquad R(0)$: nonsingular

ここで

$$R(h) \underset{d}{=} [\rho_{ij}(h)]_{K\times K}$$

$$\rho_{ij}(h) \underset{d}{=} \lim_{T\to\infty} r_{ij}^T(h)$$

このとき $R(h)$ は

$$R(h) = \int_{-\pi}^{\pi} e^{ih\lambda}\, dM(\lambda)$$

と表現できる。$M(\lambda)$は半正値な増分を持つエルミート行列

推定量

$$\hat{\beta}_{T,LSE} = (\tilde{X}_T' \tilde{X}_T)^{-1} \tilde{X}_T' Y_T \quad LSE$$

$$\hat{\beta}_{T,BLU} = (\tilde{X}_T' \Sigma_T^{-1} \tilde{X}_T)^{-1} \tilde{X}_T' \Sigma_T^{-1} Y_T \quad BLUE$$

ここで

$$Y_T = (y_1, y_2, \cdots, y_T)' \; T\times 1$$

$$\tilde{X}_T = (x_{ti})_{T\times K}$$

$$\Sigma_T = (\sigma_{ij})_{T\times T} \quad \sigma_{ij} = \gamma_{i-j} \; (= E\varepsilon_i \varepsilon_j)$$

LSEの漸近有効性

$\{\varepsilon_t\}$のスペクトル密度がpositive, continuousな場合については、BLUEに対してLSEが漸近有効になるための必要十分条件をGrenander (195※)が導出している。説明変数が多項式 $x_{ti} = t^{i-1}$, 三角関数 $x_{ti} = \sin \nu_i t$ or $\cos \nu_i t$ $(\nu_i \neq 0)$ (両者とも(G.1)〜(G.※)を満足する)などの代表的な例についてはLSEは漸近有効となる。しかしながら$\{\varepsilon_t\}$が長期記憶モデルにしたがうとき、$x_{ti} \equiv t^{i-1}$の場合については、LSEがもはや漸近有効でないことが示されている。(Beran & Künsch (1985), Samarov & Taqqu (1988), Yajima (1988)). ここでは他の一般の説明変数について、LSEの性質を議論する。比較を明瞭にする為、まずGrenanderの結果を示す。

Theorem (Grenander)

$$D_T \underset{d}{=} \mathrm{diag}(\|x_1\|_T, \|x_2\|_T, \cdots, \|x_K\|_T)$$

$$\|x_i\|_T \underset{d}{=} (a^T_{ii}(0))^{1/2} \quad i = 1, 2, \cdots, K$$

以下では(G.1)〜(G.4)を仮定する.

(i) $f(\lambda)$: continuous とする。

$$\lim_{T\to\infty} D_T E\{(\hat{\beta}_{T,LSE}-\beta)(\hat{\beta}_{T,LSE}-\beta)'\} D_T$$
$$= 2\pi R(0)^{-1} \int_{-\pi}^{\pi} f(\lambda)\, dM(\lambda)\, R(0)^{-1}$$

(ii) $f(\lambda)$: positive continuous とする

$$\lim_{T\to\infty} D_T E\{(\hat{\beta}_{T,BLU}-\beta)(\hat{\beta}_{T,BLU}-\beta)'\} D_T$$
$$= [(2\pi)^{-1} \int_{-\pi}^{\pi} f(\lambda)^{-1} dM(\lambda)]^{-1}$$

(iii) LSE が漸近有効になるための必要十分条件は、「$M(\lambda)$が $0 \leq \lambda \leq \pi$ の高々K個の点で増加し、各点での行列の階数の総和がKになる」ことである。但し、$f(\lambda)$は positive continuous とする。

次にいくつかの記号・仮定を導入する。

記号

$$m^T_{pq}(\lambda) \underset{d}{=} \frac{(\sum_{t=1}^{T} x_{tp} e^{-it\lambda})(\sum_{t=1}^{T} x_{tq} e^{it\lambda})}{2\pi \|x_p\|_T \|x_q\|_T}$$

$$M^T_{pq}(\lambda) \underset{d}{=} \int_{-\pi}^{\lambda} m^T_{pq}(\omega)\, d\omega$$

$$M^T(\lambda) \underset{a}{=} [\, M^T_{pq}(\lambda)\,]_{K\times K}$$

$M^T(\lambda)$ を complex matrix measure とみなせば、(G.3) は

$$M^T(\lambda) \xrightarrow[W]{} M(\lambda) \quad (T\to\infty)$$

を意味する。（ここで $\xrightarrow[W]{}$ は弱収束）。この事実が以下の結果を導く上で、key point になる。

<u>仮定.</u> 必要ならば説明変数の番号を入れ換えることにより、

$$M_{ii}(0+) - M_{ii}(0) > 0, \; i=1,\cdots,m, \; 0\le m\le K$$

$$M_{ii}(0+) - M_{ii}(0) = 0, \; i=m+1,\cdots,K.$$

とする。

<u>Theorem 1 (LSEの漸近分散)</u>

以下 (G.1)〜(G.4)を仮定する。また $\{\varepsilon_t\}$ は長期記憶モデルにしたがうとする。

(i) $m = 0$ とする。

Theorem (Grenander) の (i) の等式成立

$\Leftrightarrow$ $^\forall\delta\,(>0)\; ^\exists c$ s.t.

(*) $\int_{|\lambda|\le c} f(\lambda)\, dM^T_{ii}(\lambda) < \delta, \; i=m+1,\cdots,K, \; ^\forall T.$

(ii) $m > 0$ とする。(*)を仮定

(a) $\int_{-\pi}^{\pi} f(\lambda)\, dM_{ij}^{T}(\lambda) / T^{2d}$ の極限存在 $(T\to\infty)$, $1\le i,j\le m$

$\iff$ $\int_{-\pi}^{\pi} |1-e^{i\lambda}|^{-2d} dM_{ij}^{T}(\lambda)/T^{2d}$ の極限存在 $(T\to\infty)$, $1\le i,j\le m$

(b) $\widehat{D}_T \equiv diag(\|x_1\|_T \cdot T^{d}, \cdots, \|x_m\|_T T^{d}, \|x_{m+1}\|_T, \cdots, \|x_k\|_T)$

$b_{ij}^{(1)} \equiv \lim_{T\to\infty} \int_{-\pi}^{\pi} |1-e^{i\lambda}|^{-2d} dM_{ij}^{T}(\lambda)/T^{2d}$, $1\le i, j\le m$

このとき

$$\lim_{T\to\infty} \widehat{D}_T^{-1} (\widetilde{X}_T' \widetilde{X}_T) E\{(\widehat{\beta}_{T,LSE} - \beta)(\widehat{\beta}_{T,LSE} - \beta)'\}(\widetilde{X}_T' \widetilde{X}_T)\widehat{D}_T^{-1} = 2\pi B$$

ここで

$$B = \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix}$$

$$B_1 = [\, b_{ij}^{(1)} \,]_{m\times m} \quad B_2 = \left[\int_{-\pi}^{\pi} f(\lambda)\, dM_{i+m,j+m}(\lambda)\right]_{(k-m)\times(k-m)}$$

次にBLUEについて考える。 (G.1)〜(G.4)に加えて以下の条件を課す。

(G.5) $\max_{1\le t\le T} x_{ti}^2 / a_{ii}^T(0) = o(1/T^{\delta})$, ${}^{\exists}\delta > 1-2d$, $i=1,\cdots,k$.

<u>Theorem 2. (BLUEの漸近分散. I)</u>

次の仮定をおく.

$$0 < M_{ii}(0+) - M_{ii}(0) < 1 \quad , i=1,\cdots,m$$

$\int_{-\pi}^{\pi} f(\lambda)^{-1} dM(\lambda)$ が nonsingular

このとき (G.1) 〜 (G.5) の下で、Theorem (Grenander) の (ii)

の等式が成立する．

ところで

$$M_{\nu\nu}(0+)-M_{\nu\nu}(0)=1 \quad {}^{\exists}\nu\ (1\leq\nu\leq m)$$

のときは，$\int_{-\pi}^{\pi} f(\lambda)^{-1} dM(\lambda)$ は singular になる．一般的な場合は未解決である．ここでは典型的な例として多項式を考える．すなわち

$$x_{t\nu}=t^{\nu-1},\ \nu=1,\cdots,p.\quad 0\leq p\leq m$$

とする．したがって

$$M_{\nu\nu}(0+)-M_{\nu\nu}(0)=1\ ,\quad \nu=1,\cdots,p.$$

$$M_{\nu\nu}(0+)-M_{\nu\nu}(0)<1\ ,\quad \nu=p+1,\cdots,m.$$

とする．

<u>Theorem 3（BLUEの漸近分散 II）</u>

$\tilde{D}_T \equiv diag(\|x_1\|_T/T^{\alpha},\cdots,\|x_p\|_T/T^{\alpha},\|x_{p+1}\|_T,\cdots,\|x_k\|_T)$

上述の仮定および(G.1)〜(G.5)の下で．

$$\lim_{T\to\infty}\tilde{D}_T\, E\{(\hat{\beta}_{T,BLU}-\beta)(\hat{\beta}_{T,BLU}-\beta)'\}\tilde{D}_T$$
$$=2\pi W$$

ここで

$$W=\begin{bmatrix} W_1 & 0 \\ 0 & W_2 \end{bmatrix}$$

$$W_1 = [\,\omega^{(1)}_{ij}\,]_{p\times p} \qquad \omega^{(1)}_{ij} = \frac{\Gamma(i-\alpha)\Gamma(j-\alpha)\{(2i-1)(2j-1)\}^{1/2}}{f^*(0)\Gamma(i-2\alpha)\Gamma(j-2\alpha)(i+j-1-2\alpha)}$$

$$W_2 = \left[\int_{-\pi}^{\pi} f(\lambda)^{-1} dM_{i+p,\,j+p}(\lambda)\right]_{(k-p)\times(k-p)}$$

以上の結果よりLSEの漸近有効性について、Theorem (Grenander) の(iii)を一般化して、次の定理を得る。

<u>Theorem 4 (LSEの漸近有効性)</u>

次の仮定をおく。

$$M_{ii}(0+) - M_{ii}(0) = 0 \quad , \quad 1 \leq i \leq k$$

$$\forall \delta > 0 \ \exists c \ \ st$$

$$\int_{|\lambda|\leq c} f(\lambda)\, dM^{T}_{ii}(\lambda) < \delta, \quad 1 \leq i \leq k, \ \forall T.$$

このとき(G.1)〜(G.5)の下で、LSEが漸近有効になるための必要十分条件は、『$M(\lambda)$が$0<\lambda\leq\pi$の高々k個の点で増加し、各点での行列の階数の総和がkとなる』ことである。

Theorem 4の直観的解釈は次の様になる。$M(\lambda)$が$\lambda=0$で増加するときは、$f(\lambda)$も$\lambda=0$において発散するので、LSEは説明変数と誤差項のスペクトルを判別できず、漸近有効にはならない。しかし$M(\lambda)$が$0<\lambda\leq\pi$で増加する場

合には、$f(\lambda)$は半閉区間$(0,\pi]$で有限値を取るので、LSEは$M(\lambda)$と$f(\lambda)$を分離することができ、漸近有効となる.

ここでTheorem 1〜4の含意を明確にするため、ひとつの例をあげる。

Example

$$y_t = \beta_1 x_{t1} + \beta_2 x_{t2} + \varepsilon_t$$

ここで

$$x_{t1} = \tau_{10} + \tau_{11}\cos\nu_1 t + \tau_{12}\cos\nu_2 t$$

$$x_{t2} = \tau_{20} + \tau_{21}\cos\nu_1 t + \tau_{22}\cos\nu_2 t$$

$$\nu_i \neq 0,\ i=1,2\ ,\ \nu_1 \neq \nu_2.$$

このとき

$$R(h) = M_0 + \cos\nu_1 M_1 + \cos\nu_2 M_2$$

但し、$M_0 = \Gamma\tilde{\tau}_0\tilde{\tau}_0'\Gamma$, $M_i = \Gamma\tilde{\tau}_i\tilde{\tau}_i'\Gamma$, $i=1,2.$

$$\Gamma = diag\left(\frac{1}{\sqrt{\tau_{10}^2+(\tau_{11}^2+\tau_{12}^2)/2}},\ \frac{1}{\sqrt{\tau_{20}^2+(\tau_{21}^2+\tau_{22}^2)/2}}\right)$$

$$\tilde{\tau}_i = (\tau_{1i},\ \tau_{2i})',\ i=0,1,2.$$

いまLSEの漸近相対効率を

$$e(d) = \lim_{T\to\infty}\frac{\det[E(\hat{\beta}_{T,BLU}-\beta)(\hat{\beta}_{T,BLU}-\beta)']}{\det[E(\hat{\beta}_{T,LSE}-\beta)(\hat{\beta}_{T,LSE}-\beta)']}$$

とする。$\tilde{\tau}_i$について以下の三例を考える. 各々の例に

おって $e(\alpha)$ は互いに異なった値をとる.

(i) $\tilde{\tau}_0=(0,0)'$, $\tilde{\tau}_1=(1,0)'$, $\tilde{\tau}_2=(0,1)'$

Th. 1. (i), Th. 2, Th. 4 より

$$e(\alpha)=1, \quad {}^{\forall}\alpha.$$

(ii) $\tilde{\tau}_0=(1,0)'$, $\tilde{\tau}_1=(1,0)'$, $\tilde{\tau}_2=(0,1)'$

Th. 1. (ii), Th. 2 より

$$e(\alpha)=0, \quad {}^{\forall}\alpha.$$

(iii) $\tilde{\tau}_0=(1,0)'$, $\tilde{\tau}_1=(0,1)'$, $\tilde{\tau}_2=(0,0)'$

Th. 1. (ii), Th. 3 より

$$0<e(\alpha)=(1+2\alpha)\Gamma(1+\alpha)\Gamma(2-2\alpha)/\Gamma(1-\alpha)<1, \quad {}^{\forall}\alpha.$$

なお $f(\lambda)$ が positive continuos の場合は、漸近相対効率 e を $e(\alpha)$ と同様に定義すれば、Theorem (Grenander) より各例について

(i) $e=1$ (ii) $0<e<1$ (iii) $e=1$

を得る。

LSEの漸近分布

Theorem 5

次の仮定をおく

$\{\varepsilon_t\}$ 強定常過程 $E|\varepsilon_t|^k<\infty$, ${}^{\forall}k$

$\{\varepsilon_t\}$ の white noise processes の任意の次数のキュムラントは絶対総和可能

さらに Th. 1 (ii) と同様の仮定、および以下の二条件を導入する。

$$\max_{|\lambda|\le\delta}\left(m^T_{ii}(\lambda)\right)^{1/2}=o\left(1/T^{1/2+\alpha/2}\right),\ i=m+1,\cdots,k.\ {}^{\exists}\delta(>0)$$

$$\int_{-\pi}^{\pi}\left(m^T_{ii}(\lambda)\right)^{1/2}d\lambda=o\left(1/T^{1/2+\alpha/2}\right),\ i=1,\cdots,k.$$

このとき

$$\hat{D}_T^{-1}(\tilde{X}_T'\tilde{X}_T)(\hat{\beta}_T-\beta)\xrightarrow[L]{}N(0,2\pi B)\ (T\to\infty)$$

ここで $\xrightarrow[L]{}$ は法則収束とする。

<u>Remark</u>

具体的な例として

$$x_{ti}=t^{i-1},\ 1\le i\le m,$$

$$x_{ti}=\cos\nu_i t,\ \sin\nu_i t,\ \text{あるいは周期関数}\ m+1\le i\le k$$

ただし $\nu_i\neq 0$ とすれば

$$\max_{|\lambda|\le\delta}\left(m^T_{ii}(\lambda)\right)^{1/2}=O\left(1/T^{1/2}\right),\ i=m+1,\cdots,k\ {}^{\exists}\delta(>0)$$

$$\int_{-\pi}^{\pi}\left(m^T_{ii}(\lambda)\right)^{1/2}d\lambda=O(\log T/T),\ i=1,\cdots,k.$$

となり、Theorem 5 の仮定を満足する。

3. 今後の問題

回帰モデルを含め、今後解決すべき問題を二、三列挙する。

(1) 回帰モデルにおいてLSEが漸近有効でないとき、

LSEより相対効率の良い推定量は存在するか？（ただしBLUEは除外して）

例 $y_t = \beta_1 + \varepsilon_t$

(i) $\{\varepsilon_t\}$のスペクトル密度が $f(\lambda) = 1/|1-e^{i\lambda}|^{2d}$ つまり $f^*(\lambda) \equiv 1$ のときのBLUEは、真のスペクトル密度が $f(\lambda) = f^*(\lambda)/|1-e^{i\lambda}|^{2d}$, $f^*(\lambda)$は任意の正の連続関数であっても、漸近的には有効となる。（Adenstedt (1974)）．しかし実際にはdが未知なので何らかの推定量で代替しなくてはならない．

(ii) 任意のM-estimatorの漸近相対効率はLSEのそれに等しい．（Beran & künsch (1985)）従ってM-estimatorによってLSEを改善することはできない．

(iii) $h(t) : [0,1] \longrightarrow \mathbb{R}$ 有界変動な実関数

$h(t)$をtaperとして用いると，

$$\lim_{T\to\infty} \frac{T^{1-2d}\, Var\left(\sum_{t=1}^{T} h(t/T) y_t \big/ \sum_{t=1}^{T} h(t/T)\right)}{T^{1-2d}\, Var\left(\sum_{t=1}^{T} y_t / T\right)}$$

$$= \frac{\int_0^1 \int_0^1 h(x) h(y) |x-y|^{2d-1} dx\, dy}{\{\int_0^1 h(x) dx\}^2 \int_0^1 \int_0^1 |x-y|^{2d-1} dx\, dy}$$

上式は算術平均（この場合のLSE）の分散と、taperを用いた線形不偏推定量の分散の比の極限である．これを1より小さくする$h(t)$は存在するか？

(2) 検定に関しては local alternative などの問題がある

たとえば

$$H_0 : d = d_0 \qquad H_1 : d = d_0 + c/\sqrt{T}$$

とした場合、尤度比の性質、contiguity，LAN族に属するかが問題となる．特に $d_0 = 0$ のときは、短期記憶モデル vs 長期記憶モデルの検定となる．

(3) 最後に、真のスペクトル密度が想定するパラメトリックモデルに属さない incorrect model の問題がある．

例．

真のスペクトル密度 $f(\lambda) = f^*(\lambda)/|1-e^{i\lambda}|^{2\alpha}, 0<\alpha<1/2$

想定するモデル $f(\lambda) = |\theta(e^{i\lambda})|^2/|\phi(e^{i\lambda})|^2$

ARMAモデル

このときMLE等はどのような漸近的性質を持つか？

「以上」

References

Adenstedt, R.K. (1974). On large-sample estimation for the mean of a stationary random sequence. Ann. Statist. 2 1095-1107.

Beran, J. (1989). A test of location for data with slowly decaying serial correlations. Biometrika 76 261-269.

Beran, J. and Kunsch, H. (1985). Location estimators for processes with long range dependence. Res. Rep. 40 Seminar fur Statistik. ETH. Zurich.

Grenander, U. (1954). On estimation of regression coefficients in the case of an autocorrelated disturbance. Ann. Math. Statist. **25** 252-272.

Samarov, A. & Taqqu, M. S. (1988). On the efficiency of the sample mean in long-memory noise. J. Time Series Analysis **9** 191-200.

Yajima, Y. (1988). On estimation of a regression model with long-memory stationary errors. Ann. Statist. **16** 791-807.