数理解析研究所講究録
第312巻 1977年 59-77

Parameter estimation of a linear process

東工大 理　柴田 里程

Time series analysis において，よく想定される母集団の process は AR，MA，ARMA process 等であるが，それほど構造のはっきりしていない場合には，もっと一般に one sided linear stationary process であるとして，話を進める必要がある [3]，[4]，[7]．

ここでは，そのような process の無限個のパラメーターの推定，およびその推定されたパラメーターを使って予測を行なったときの予測誤差を問題にする．

§1 では best predictor の係数の最尤推定量は，実は，対応する order の AR を fit したときの Yule-Walker equation の解であることを示し，§2 ではその optimal な order を求め，§3 で efficient な selection method を与える．

§1. Prediction and parameter estimation

$\{x_t\}$ は平均 0 の one sided linear stationary Gaussian process;

$$x_t + a_1 x_{t-1} + \cdots = e_t \qquad (t = \cdots, -1, 0, 1, \cdots) \qquad (1-1)$$

であるとする。ただし，e_t は $N(0, \sigma^2)$ に従う i.i.d. で associated polynomial $A(z) = 1 + a_1 z + \cdots$ は $|z| \leqq 1$ で finite, nonzero であるとする。correlation を $r_\ell = E(x_t x_{t+\ell})$, $k \times k$ covariance matrix を $R(k)$ で表わす。

さらに，vector space $\{\alpha' = (\alpha_1, \alpha_2, \cdots)\}$ に

$$\|\alpha\|_R^2 = \sum_{i,j=1}^{\infty} \alpha_i \alpha_j r_{|i-j|}$$

なる norm を導入し，この norm によるパラメーター $a' = (a_1, a_2, \cdots)$ の $(h+k-1)$-dim. subspace

$$\{\alpha' = (0, \cdots, 0, \alpha_h, \alpha_{h+1}, \cdots, \alpha_{h+k-1}, 0, \cdots, 0)\}$$

への projection を

$$\underline{a}'(h, k) = (0, \cdots, 0, a_h(h, k), \cdots, a_{h+k-1}(h, k), 0, \cdots)$$

とすれば，よく知られているように，これは k step 前からの観測値 $x_{t-k+1}, \cdots, x_t$ にもとづく h step 先の x_{t+h} の best predictor $\hat{x}_{t+h}$ の係数を表わす vector になっている。すなわち

$$\hat{x}_{t+h} = E(x_{t+h} | x_{t-k+1}, \cdots, x_t) = -\sum_{i=h}^{h+k-1} a_i(h, k) x_{t+h-i}.$$

さて，n-samples $x_1, \cdots, x_n$ が与えられたときの $\underline{a}(h, k)$ の M.L.E. を求めてみよう。

Theorem 1.1

$\underline{a}(h, k)$ の一つの M.L.E. は次の equation

$$\sum_{j=h}^{h+k-1} \hat{r}_{|i-j|}\, \hat{a}_j(h, k) = -\hat{r}_i \quad (i = h, \cdots, h+k-1) \qquad (1-2)$$

の解

$$\hat{a}(h, k)' = (0, \cdots, 0, \hat{a}_h(h, k), \cdots \hat{a}_{h+k-1}(h, k), 0, \cdots)$$

で与えられる。ただし，$h+k \leqq n$， $\hat{r}_l = \frac{1}{n-l}\sum_{t=1}^{n-l} x_t x_{t+l}$.

Proof. $f(x_1, \cdots, x_n, \theta)$ で $x_1, \cdots, x_n$ の同時密度を表わせば，$r_0, r_1, \cdots, r_{n-1}$ は a と σ^2 の関数であるから

$$\max_{a, \sigma^2} f(x_1, \cdots, x_n, a, \sigma^2) \leqq \max_{r_0, \cdots, r_{n-1}} f(x_1, \cdots, x_n, r_0, \cdots, r_{n-1})$$
$$= f(x_1, \cdots, x_n, \hat{r}_0, \cdots, \hat{r}_{n-1})$$

ここで，$\{r_i\}_{i=0}^{\infty}$ と a の関係

$$\sum_{j=1}^{\infty} r_{|i-j|}\, a_j = -r_i \qquad (i = 1, 2, \cdots) \qquad (1-3)$$

より，$r_l^* = \hat{r}_l$ $(l = 0, 1, \cdots, n-1)$ であるような $\{r_i^*\}_{i=0}^{\infty}$ に対して (1-3) を満たし，$A(z)$ が $|z| \leqq 1$ で *finite*, *nonzero* であるような a^* が a の一つの M.L.E. になる。したがって，$\underline{a}(h, k)$ の M.L.E. は a^* の $\|\cdot\|_{R^*}$ による *projection* で与えられるから，結局 (1-2) の解となる。 ┘

ところで，theorem 1.1 の $\hat{a}(h,k)$ は，実は，$(h+k-1)$ th order AR model

$$x_{t+h}+\alpha_1 x_{t+h-1}+\cdots+\alpha_{h+k-1}x_{t-k+1}=\varepsilon_t$$

（ただし，$\alpha_1=\alpha_2=\cdots=\alpha_{h-1}=0$ ）

を fit させたときの

$$\alpha'=(0,\cdots,0,\alpha_h,\cdots,\alpha_{h+k-1},0,\cdots)$$

の approximate M.L.E.（Yule - Walker equation (1-2) の解）に他ならない。したがって

「AR model を fit させたときの Yule - Walker equation の解」

$\Longleftrightarrow$「対応する predictor の係数の M.L.E.」

であることがわかる。

さて，話を簡単にするために，$\{x_t\}$ と同じ構造をもちこれと独立なもう一つの process $\{y_t\}$ を考え，推定したパラメーター $\hat{a}(h,k)$ を使ってつくった predictor

$$\hat{y}_{t+h}=-\sum_{i=h}^{h+k-1}\hat{a}_i(h,k)\,y_{t+h-i}$$

の mean squared error を求めると

$$\begin{aligned}E^y(\hat{y}_{t+h}-y_{t+h})^2&=\|\hat{a}(h,k)-a\|_R^2+\sigma^2\\&=\|\hat{a}(h-k)-\underline{a}(h,k)\|_R^2+\|\underline{a}(h,k)-a(h,\infty)\|_R^2\\&\quad+\|\underline{a}(h,\infty)-a\|_R^2+\sigma^2\qquad(1-4)\end{aligned}$$

となる。ただし，$\underline{a}(h,\infty)$は$\{\alpha'=(0,\cdots,0,\alpha_h,\alpha_{h+1},\cdots)\}$への$a$の projection で$h=1$なら$\underline{a}(h,\infty)=a$。(1-4)は，また$\hat{x}_{n+h}-x_{n+h}$の漸近分散とも漸近的には等しく，第1項は推定量のバラツキ，第2項は bias，第3，4項は予測のバラツキを表わしている。さらに，$\|\hat{a}(h,k)-a\|_R^2$なる loss function は，h, kを fix したときの$\hat{a}(h,k)$の漸近分散行列(= Fisher information matrix)で normalize した quadratic loss function でもある。

(1-4)の第1項はkを fix したとき$\frac{1}{n}$の order で0に確率収束するが，そのとき第2項は$\{x_t\}$が$h+k-1$より低い order の AR でない限り残ってしまう。そこで，kをnに depend して決めていくことにより第1項，第2項を balance させながら，第1項+第2項を最小にもっていくことが必要になる。

§2. Efficiency of a selection method

$h=1$の場合だけを考えれば，あとは同様であるので，以後，この場合だけを考え簡単に$\underline{a}(k)=\underline{a}(1,k)$と表わす。また，$k$の上限$\bar{K}_n\ (<n)$を与え，$1\leqq k\leqq \bar{K}_n$の範囲で考える。

Notations.

$$X_j(k)'=(x_j,\ x_{j-1},\ \cdots,\ x_{j-k+1})$$

$$\hat{R}(k) = (\hat{r}_{\ell,m} \quad 1 \leqq \ell, m \leqq k)$$
$$= \frac{1}{n}\sum_{j=\bar{K}_n}^{n-1} X_{j+1}(k) X'_{j+1}(k)$$

$$\hat{r}(k)' = (\hat{r}_{0,1}, \hat{r}_{0,2}, \cdots, \hat{r}_{0,k})$$

$$\hat{a}(k) = (\hat{a}_1(k), \hat{a}_2(k), \cdots, \hat{a}_k(k), 0, \cdots)$$

: $\hat{R}(k)\,\hat{a}(k) = -\hat{r}(k)$ の解

($\hat{a}(k)$ を k-dim. vector とみて)

$$\underline{e}_{j+1,k} = x_{j+1} + \underline{a}_1(k) x_j + \cdots + \underline{a}_k(k) x_{j+1-k}$$

$$\underline{s}_k^2 = \frac{1}{n}\sum_{j=\bar{K}_n}^{n-1} \underline{e}_{j+1,k}^2$$

$$\sigma_k^2 = E(\underline{e}_{j+1,k}^2)$$

さらに，$\|\cdot\|$ は通常の Euclid norm を表わす.

Remark.

$\hat{a}(k)$ は §1 の $\hat{a}(1,k)$ とは厳密には違いがあるが，漸近的には同等である。ここでは，$\hat{a}(k)$ を採用し，loss $\|\hat{a}(k) - a\|_R^2$ を問題にする。

Assumptions

(A - 1) $\quad \sum_{j=1}^{\infty} |a_j| < \infty$

(A - 2) $\quad A(e^{i\lambda}) \neq 0 \quad \cdot (-\pi \leqq \lambda \lneqq \pi)$

(A - 3) $\quad \sum_{j=1}^{\infty} |r_j| < \infty$

(A - 4) $\quad \bar{K}_n = o(\sqrt{n})$

(A－5)　$\{x_t\}$ は finite order の AR ではない。

Lemma 2.1

(A－1)～(A－3) の仮定のもとで、ある absolute constants $c_1, c_2>0$ が存在して $\bar{K}_n \geqq {}^\forall k \geqq 1$ に対して

$$E\left(\frac{n}{k}\left\|\sum_{j=\bar{K}_n}^{n-1} X_j(k)e_{j+1}/n\right\|^2_{R(k)^{-1}}-\sigma^2\right)^2$$

$$\leqq\left(1-\frac{\bar{K}_n}{n}\right)\frac{c_1}{n}+\left(1-\frac{\bar{K}_n}{n}\right)^2\frac{c_2}{k}$$

Proof. Gaussian property を使って、$\{x_t\}$ の8次モーメントを評価すれば結果を得る。 ┘

Lemma 2.2

(A－1)～(A－3) の仮定のもとで、ある absolute constant $c_3>0$ が存在して $\bar{K}_n \geqq {}^\forall k \geqq 1$ に対して

$$E\left(\frac{n}{k}\left\|\sum_{j=\bar{K}_n}^{n-1} X_j(k)(\underline{e}_{j+1,k}-e_{j+1})/n\right\|^2\right)$$

$$\leqq c_3\left(1-\frac{\bar{K}_n}{n}\right)\|a-\underline{a}(k)\|^2$$

Proof. see [7]. ┘

Lemma 2.3

(A－1), (A－2), (A－4) の仮定のもとで

$$p-\lim_{n\to\infty}\left(\max_{k\leqq \bar{K}_n}\|\hat{R}(k)-R(k)\|\right)=0$$

$$p\text{-}\lim_{n\to\infty}\left(\max_{k\leq \bar{K}_n}\|\hat{R}(k)^{-1}-R(k)^{-1}\|\right)=0$$

ただし，matrix A の norm は $\|A\|=\sup_{\|\alpha\|\leq 1}\|A\alpha\|$ で定義される。

Proof. $\max_{k\leq \bar{K}_n}\|\hat{R}(k)-R(k)\|\leq\|\hat{R}(\bar{K}_n)-R(\bar{K}_n)\|$ であることと，[3] より従う。 ┘

$$\hat{a}(k)-\underline{a}(k)=-\hat{R}(k)^{-1}\left(\sum_{j=\bar{K}_n}^{n-1}X_j(k)\,e_{j+1,k}/n\right)$$

であるので，上の三つの lemma と $\|R(k)^{-1}\|$ が bounded であることより，次の proposition を得る。

Proposition 2.1

(A-1)～(A-4) の仮定のもとで，$1\leq k_n\leq\bar{K}_n$ で $\liminf_n k_n=\infty$ なる任意の positive integer の sequence $\{k_n\}$ について

$$p\text{-}\lim_{n\to\infty}\left|\frac{n}{k_n}\|\hat{a}(k_n)-\underline{a}(k_n)\|_R^2-\sigma^2\right|=0\text{。}$$

したがって，$\|\hat{a}(k_n)-a\|_R^2$ は漸近的に

$$L_n(k_n)=\frac{k_n}{n}\sigma^2+\|a-\underline{a}(k_n)\|_R^2$$

に等しくなる。

Definition 2.1

k_n^* は $L_n(k_n^*) = \min_{1 \leqq k \leqq \overline{K}_n} L_n(k)$ なる positive integer.

(A-5) の仮定のもとで，$\{k_n^*\}$ は n について monotone increasing で $\lim_n k_n^* = \infty$ なる sequence となる。

Theorem 2.1

(A-1) ～ (A-5) の仮定のもとで，任意の $1 \leqq k_n \leqq \overline{K}_n$ について

$$p - \lim_{n \to \infty} \frac{\| \hat{a}(k_n) - a \|_R^2}{\| \hat{a}(k_n^*) - a \|_R^2} \geqq 1 \qquad (2-1)$$

Proof.

$$\frac{\| \hat{a}(k_n) - a \|_R^2}{\| \hat{a}(k_n^*) - a \|_R^2} = \frac{L_n(k_n^*)}{\| \hat{a}(k_n^*) - a \|_R^2} \cdot \frac{\| \hat{a}(k_n) - a \|_R^2}{L_n(k_n)} \cdot \frac{L_n(k_n)}{L_n(k_n^*)}$$

で，proposition 2.1 より

$$p - \lim_{n \to \infty} \frac{L_n(k_n^*)}{\| \hat{a}(k_n^*) - a \|_R^2} = 1$$

また，

$$\frac{k}{(n L_n(k))^2} \leqq \frac{1}{n L_n(k) \sigma^2}$$

$$\leqq \frac{1}{n L_n(k_n^*) \sigma^2} \leqq \frac{1}{k_n^* \sigma^4}$$

であるから，lemma 2.1 〜 2.3 より proposition 2.1 と同様にして

$$p\text{-}\lim_{n\to\infty} \frac{\|\hat{a}(k_n) - a\|_R^2}{L_n(k_n)} = 1$$

┘

この theorem より k_n^* が漸近的に最も optimal な order であることがわかる。しかし，これはパラメーターに依存し，一般には unknown。そこで何らかの integer valued statistic $\tilde{k}_n$ で k_n^* を推定することを考えよう。

まず proposition 2.1 を拡張し，次の proposition を得る。

Proposition 2.2

(A-1) 〜 (A-5) の仮定のもとに，$1 \leqq l_n \leqq m_n \leqq \bar{K}_n$，$\lim_{n\to\infty} \frac{m_n}{l_n} = 1$ なる任意の integer の sequences $\{l_n\}$，$\{m_n\}$ に対して

$$p\text{-}\lim_{n\to\infty} \left(\max_{l_n \leqq k \leqq m_n} \left| \frac{\|\hat{a}(k) - a\|_R^2}{L_n(k)} - 1 \right| \right) = 0 .$$

Proof.

lemma 2.1 より

$$\sum_{k=l_n}^{m_n} E\left(\frac{\frac{n}{k} \left\| \sum_{j=\bar{K}_n}^{n-1} X_j(k)\, e_{j+1} / n \right\|_{R(k)^{-1}}^2 - \sigma^2}{\frac{n}{k} L_n(k)} \right)^2$$

$$\leqq c_1' \frac{\bar{K}_n^2}{n} + c_2' \sum_{k=l_n}^{m_n} \frac{k}{(n L_n(k))^2}$$

また，lemma 2.2 より

$$\sum_{k=l_n}^{m_n} E\left(\frac{\sqrt{\frac{n}{k}}\left\|\sum_{j=\overline{K}_n}^{n-1} X_j(k)e_{j+1,k}/n\right\|_{R(k)^{-1}} - \sqrt{\frac{n}{k}}\left\|\sum_{j=\overline{K}_n}^{n-1} X_j(k)e_{j+1}/n\right\|_{R(k)^{-1}}}{\sqrt{\frac{n}{k}L_n(k)}}\right)^2$$

$$\leqq c_3' \frac{\overline{K}_n^{\,2}}{n}$$

したがって，$l_n' = \min(l_n, \log k_n^*)$，$m_n' = \min(m_n, \log k_n^*)$ とすれば

$$\sum_{k=l_n'}^{m_n'} \frac{k}{(nL_n(k))^2} \leqq \text{const.}\frac{(\log k_n^*)^2}{k_n^{*2}},$$

$l_n'' = \max(l_n, \log k_n^*)$，$m_n'' = \max(m_n, \log k_n^*)$ とすれば

$$\sum_{k=l_n''}^{m_n''} \frac{k}{(nL_n(k))^2} \leqq \text{const.}\sum_{k=l_n''}^{m_n''} \frac{1}{k}$$

であるから，$\{l_n\}$，$\{m_n\}$ の条件より

$$p\text{-}\lim_{n\to\infty}\left(\max_{l_n \leqq k \leqq m_n}\left|\frac{\|\hat{a}(k) - \underline{a}(k)\|_R^2 - \frac{k}{n}\sigma^2}{L_n(k)}\right|\right) = 0. \quad \lrcorner$$

Theorem 2.2

(A－1)～(A－5) の仮定のもとに，ある sequence $\{k_n\}$ $(1 \leqq k_n \leqq \overline{K}_n)$ と positive constant α が存在して

$$p\text{-}\lim_{n\to\infty} \frac{\tilde{k}_n}{k_n} = \alpha$$

であるような order selection $1 \leqq \tilde{k}_n \leqq \bar{K}_n$ について，次の不等式が成り立つ。

$$p\text{-}\varliminf_{n\to\infty} \frac{\| \hat{a}(\tilde{k}_n) - a \|_R^2}{\| \hat{a}(k_n^*) - a \|_R^2} \geqq 1 \qquad (2\text{-}2)$$

Proof.

proposition 2.2 より $p\text{-}\lim \frac{\tilde{k}_n}{k_n} = \alpha$ の条件を使えば

$$p\text{-}\varliminf_{n\to\infty} \frac{\| \hat{a}(\tilde{k}_n) - a \|_R^2}{\| \hat{a}(k_n^*) - a \|_R^2} = p\text{-}\varliminf_{n\to\infty} \frac{L_n(\tilde{k}_n)}{L_n(k_n^*)}$$

で，k_n^* の定義より（2-2）が従う。 ┘

そこで，(2-2) の左辺の極限が存在し，等号が成り立つとき，$\hat{k}$ は efficient であるということにする。

§3. An efficient selection method

$S_n(k) = (n + 2k)\hat{\sigma}_k^2$ なる statistic を考える。ただし

$$n\hat{\sigma}_k^2 = \sum_{j=\bar{K}_n}^{n-1} (x_{j+1} + \hat{a}_1(k)x_j + \cdots + \hat{a}_k(k)x_{j+1-k})^2 .$$

すると $S_n(\hat{k}_n) = \min_{1 \leqq k \leqq \bar{K}_n} S_n(k)$ なる $\hat{k}_n$ が一つの efficient selection method となる。これは赤池の FPE 法，AIC 法 [1]，[6] や Mallows の Cp 法 [5] と同じような方法であるが，$\hat{\sigma}_k^2$ や $\hat{a}(k)$ の定義にいくぶんちがいがある。

$S_n(k)$ の behaviour はいくぶん critical で，今の段階では上の $\hat{k}_n$ でないと efficient であるとはいえない。
まず，$S_n(k)$ を次のように書きかえる。

$$
\begin{aligned}
S_n(k) &= n(\hat{\sigma}_k^2 - \underline{s}_k^2) + n(\underline{s}_k^2 - \sigma_k^2) \\
&\quad + n(\sigma_k^2 - \sigma^2) + 2k\sigma^2 + 2k(\hat{\sigma}_k^2 - \sigma^2) \\
&= nL_n(k) - (n\|\hat{a}(k) - \underline{a}(k)\|^2_{\hat{R}(k)} - k\sigma^2) \\
&\quad + 2k(\hat{\sigma}_k^2 - \sigma^2) + n(\underline{s}_k^2 - \sigma_k^2) \ .
\end{aligned}
$$

Lemma 3.1

(A-1)～(A-5) の仮定のもとで，任意の sequence $\{k_n\}$ ($1 \leqq k_n \leqq \bar{K}_n$) に対し

$$
p\text{-}\lim_{n\to\infty} \left| \frac{n\|\hat{a}(k_n) - \underline{a}(k_n)\|^2_{\hat{R}(k_n)} - k_n\sigma^2}{nL_n(k_n)} \right| = 0
$$

Proof.

$$
\frac{n\|\hat{a}(k_n) - \underline{a}(k_n)\|^2_{\hat{R}(k_n)} - k_n\sigma^2}{nL_n(k_n)}
= \frac{\|\hat{a}(k_n) - \underline{a}(k_n)\|^2_{\hat{R}(k_n)} + \|\underline{a}(k_n) - a\|^2_R}{L_n(k_n)} - 1
$$

であるから，proposition 2.2 と lemma 2.3 より結果を得る。

Lemma 3.2

lemma 3.1と同じ仮定のもとに

$$p\text{-}\lim_{n\to\infty}\frac{k_n|\hat{\sigma}_{k_n}^2-\sigma^2|}{nL_n(k_n)}=0$$

Proof.

$$\begin{aligned}|\hat{\sigma}_k^2-\sigma^2| &\leqq |\hat{\sigma}_k^2-\underline{s}_k^2|+|\underline{s}_k^2-\sigma_k^2|+|\sigma_k^2-\sigma^2|\\ &=\|\hat{a}(k)-\underline{a}(k)\|_{\hat{R}(k)}^2+|\underline{s}_k^2-\sigma_k^2|\\ &\quad+\|\underline{a}(k)-a\|_R^2\end{aligned}$$

であることと，ある k に depend しない constant $c\ (>0)$ が存在して

$$E(\underline{s}_k^2-\sigma_k^2)^2\leqq\frac{c}{n}$$

と評価されるから，lemma 3.1 より結果が従う。 ┘

以上の lemma より

$$S_n(k_n)=nL_n(k_n)(1+\delta_n)+n(\underline{s}_{k_n}^2-\sigma_{k_n}^2)$$

と表わされる。ただし，δ_n は $p\text{-}\lim_{n\to\infty}\delta_n=0$ なる r. v.

Lemma 3.3

lemma 3.1 と同じ仮定のもとに

$$p\text{-}\lim_{n\to\infty}\left|\frac{n(\underline{s}_{k_n^*}^2-\sigma_{k_n^*}^2)-n(\underline{s}_{k_n}^2-\sigma_{k_n}^2)}{nL_n(k_n)}\right|=0$$

Proof.

$k_n^* < k$ のとき

$$n(\underline{s}^2_{k_n^*} - \sigma^2_{k_n^*}) - n(\underline{s}^2_k - \sigma^2_k)$$

$$= 2n(\underline{a}(k_n^*) - \underline{a}(k))'(\hat{\mathbf{r}}(k_n) - \mathbf{r}(k_n))$$

$$+ n(\|\underline{a}(k_n^*)\|^2_{\hat{R}(k)} - \|\underline{a}(k_n^*)\|^2_R + \|\underline{a}(k)\|^2_R - \|\underline{a}(k)\|^2_{\hat{R}(k)})$$

で

$$E\left((n(\underline{a}(k_n^*) - \underline{a}(k)))'(\hat{\mathbf{r}}(k) - \mathbf{r}(k))\right)^2$$

$$\leqq \text{const.}\, n\|\underline{a}(k_n^*) - \underline{a}(k)\|^2_R ,$$

また，

$$|n(\|\underline{a}(k_n^*)\|^2_{\hat{R}(k)} - \|\underline{a}(k_n^*)\|^2_R + \|a(k)\|^2_R - \|\underline{a}(k)\|^2_{\hat{R}(k)})|$$

$$\leqq n\|\underline{a}(k_n^*) - \underline{a}(k)\|^2 \|\hat{R}(k) - R(k)\|$$

$$+ 2n|(\underline{a}(k_n^*))'(\hat{R}(k) - R(k))(\underline{a}(k))|$$

で

$$E\left(n(\underline{a}(k_n^*))'(\hat{R}(k) - R(k))(\underline{a}(k))\right)^2$$

$$\leqq \text{const.}\, n\|\underline{a}(k_n^*) - \underline{a}(k)\|_R \, |\underline{a}(k) - a(k_n^*)|$$

と評価され，$k_n^* > k$ のときも全く同じ評価になる。ただし

$$|\underline{a}(k) - \underline{a}(k_n^*)| = \sum_{i=1}^{\max(k, k_n^*)} |\underline{a}_i(k) - \underline{a}_i(k_n^*)| .$$

一方

$$\frac{n\|\underline{a}(k_n^*) - \underline{a}(k_n)\|_R \, |\underline{a}(k_n) - \underline{a}(k_n^*)|}{(nL_n(k_n))^2}$$

$$\leqq \frac{\| \underline{a}(k_n^*) - \underline{a}(k_n) \|_R^2}{L_n(k_n)} \frac{\sqrt{\max(k_n^*, k_n)}}{n L_n(k_n)}$$

$$\leqq \left| \frac{L_n(k_n) - L_n(k_n^*) + \frac{\sigma^2}{n}(k_n^* - k_n)}{L_n(k_n)} \right| \frac{\sqrt{\max(k_n^*, k_n)}}{n L_n(k_n)}$$

ここで，k_n^* の定義より

$$\frac{\sqrt{\max(k_n^*, k_n)}}{n L_n(k_n)} \longrightarrow 0$$

であるから，結果が従う。 ⌟

以上の結果より，次の proposition を得る。

Proposition 3.1

lemma 3.1 と同じ仮定のもとに

$$S(k_n^*) - S(k_n) = n L_n(k_n^*)(1 + \delta_n) - n L_n(k_n)(1 + \delta_n')$$

と表わされる。ただし，δ_n, δ_n' は

$$p\text{-}\lim_{n\to\infty} \delta_n = 0 \quad , \qquad p\text{-}\lim_{n\to\infty} \delta_n' = 0$$

なる r. v.

ここで，任意の $\varepsilon > 0$ に対して

$$\frac{L_n(k)}{L_n(k_n^*)} \leqq 1 + \varepsilon$$

なる最小の k を $\underline{k}_n^*(\varepsilon)$，最大の k を $\overline{k}_n^*(\varepsilon)$ とする。

Theorem 3.1

(A-1)〜(A-5)の仮定のもとに，任意の ε に対して

$$\lim_{n\to\infty} \Pr\left(\underline{k}_n^*(\varepsilon) \leqq \hat{k}_n \leqq \bar{k}_n^*(\varepsilon)\right) = 1$$

であり

$$\text{p-}\lim_{n\to\infty} \frac{L_n(\hat{k}_n)}{L_n(k_n^*)} = 1 \text{。}$$

Proof.

$\hat{k}_n > \bar{k}_n^*(\varepsilon)$ または $\hat{k}_n < \underline{k}_n^*(\varepsilon)$ ならば，$S_n(k)$ の convexity より

$$S_n(k_n^*) \geqq S_n(\bar{k}_n^*(\varepsilon)+1) \quad , \quad S_n(\underline{k}_n^*(\varepsilon)-1)$$

であり，proposition 3.1より

$$\lim_{n\to\infty} \frac{L_n(\underline{k}_n^*(\varepsilon)-1)}{L_n(k_n^*)} = 1 ,$$

$$\lim_{n\to\infty} \frac{L_n(\bar{k}_n^*(\varepsilon)+1)}{L_n(k_n^*)} = 1 .$$

これは，$\bar{k}_n^*(\varepsilon)$，$\underline{k}_n^*(\varepsilon)$ の定義に矛盾する。

また，$\underline{k}_n^*(\varepsilon) \leqq \hat{k}_n \leqq \bar{k}_n^*(\varepsilon)$ ならば

$$\frac{L_n(\hat{k}_n)}{L_n(k_n^*)} \leqq 1+\varepsilon$$

であるから

$$\lim_{n\to\infty} \Pr\left(1 \leqq \frac{L_n(\hat{k}_n)}{L_n(k_n^*)} \leqq 1+\varepsilon\right) = 1 .$$

Theorem 3. 2 ($\hat{k}$ の *asymptotic efficiency*)

(A-1)～(A-5) の仮定，および

$$\lim_{\varepsilon \to 0} \left(\lim_{n \to \infty} \frac{\bar{k}_n^*(\varepsilon)}{\underline{k}_n^*(\varepsilon)} \right) = 1 \qquad (3-1)$$

が成り立てば

$$p-\lim_{n \to \infty} \frac{\| \hat{a}(\hat{k}_n) - a \|_R^2}{\| \hat{a}(k_n^*) - a \|_R^2} = 1 \ .$$

Proof.

theorem 3. 1 の結果，および *proposition* 2. 2 より従う。 ┘

Example

$\{a_i\}$ が *exponentially decreasing sequence* ならば，

$$\| \underline{a}(k) - a \|_R^2$$

も同様であり，k_n^* は $\log n$ の *order* になる。このとき (3 − 1) が成り立つ。

References

1. Akaike, H. (1969). Fitting autoregressive models for prediction. Ann. Inst. Statist. Math. 21, 243-247.

2. Akaike, H. (1970). Statistical predictor identification. Ann. Inst. Statist. Math. 22, 203-217.

3. Berk, K. N. (1974). Consistent autoregressive spectral estimates. Ann. Statist. Vol. 2, No. 3, 489-502.

4. Huzii, M. (1976). On a spectral estimate obtained by an autoregressive model fitting. Tech. Rep. Dep. of Statist. Stanford Univ. No. 22.

5. Mallows, C. L. (1973). Some comments on Cp. Technometrics 12, 591-612.

6. Shibata, R. (1976). Selection of the order of an autoregressive model by Akaike's information criterion. Biometrika 63, 117-126.

7. Shibata, R. (1978). Convergence of least square estimates of autoregressive parameters. to be published in the Australian Journal of Statistics, Vol. 20, No. 1.