

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、それは比構造のはつきりしていない場合には、もっと一般の one sided linear stationary process であるとして、話を進める必要がある [3], [4], [7].

こでは、そのような process の無限個のパラメータの推定、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う、たときの予測誤差を問題にする。

§1 では best predictor の係数の最大推定量は、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 to order を求め、§3 で efficient to selection method を与える。

§1. Prediction and parameter estimation

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

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

であるとする。ただし、 e_t は $N(0, \sigma^2)$ の従う i.i.d. な associated polynomial $A(z) = 1 + a_1 z + \dots$ は $|z| \leq 1$ で finite, nonzero でないとき $r_e = E(x_t x_{t+e})$, $k \times k$ covariance matrix で $R(k)$ を表わす。

この \mathbb{R} , vector space $\{\alpha' = (\alpha_1, \alpha_2, \dots)\}$ に

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

を norm を導入し, その norm による平行四辺形法則 $\|\alpha' - \beta'\|_R^2 = (\alpha_1, \alpha_2, \dots) \in (h+k-1)$ -dim. subspace

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

へ projection は

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

とすれば、よく知られる $\underline{\alpha}'$ は \mathbb{R} , \mathbb{R}^{h+k} step 前から α 観測値 x_{t-k+1}, \dots, x_t によって $\leq h$ step で α x_{t+h} の best predictor \hat{x}_{t+h} の係数を表わす vector である。

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

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

Theorem 1.1

$\underline{a}(h, k)$ の \rightarrow の M.L.E. は次の方程式

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

の解

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

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

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

$$\begin{aligned} \max_{\theta, \sigma^2} f(x_1, \dots, x_n, \theta, \sigma^2) &\leq \max_{r_0, \dots, r_{n-1}} f(x_1, \dots, x_n, r_0, \dots, r_{n-1}) \\ &= f(x_1, \dots, x_n, \hat{r}_0, \dots, \hat{r}_{n-1}) \end{aligned}$$

∴ て, $\{r_i\}_{i=0}^\infty$ と θ の関係

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

よし, $r_l^* = \hat{r}_l$ ($l=0, 1, \dots, n-1$) であるより $\{r_i^*\}_{i=0}^\infty$
は式 (1-3) を満たし, $A(z)$ が $|z| \leq 1$ で finite,
nonzero であるより θ^* が θ の \rightarrow の M.L.E. となる。したがって,
 $\underline{a}(h, k)$ の M.L.E. は θ^* の $\|\cdot\|_{R^k}$ による pro-
jection であるから, 結局 (1-2) の解となる。』

ところが、theorem 1.1 の $\hat{\alpha}(h, k)$ 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

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

≈ approximate M.L.E. (Yule-Walker equation (1-2))
の解) は他よりない。したがって

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

\Leftrightarrow 「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、 $\{x_t\}$ と同じ構造をもつ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 $- \hat{\alpha}(h, k)$ を使って \hat{y}_{t+h} predictor

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

mean squared error を求めると

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

となる。ただし、 $\underline{\alpha}(h, \infty)$ は $\{\alpha' = (0, \dots, 0, \alpha_h, \alpha_{h+1}, \dots)\}$ への α の projection で $h=1$ なら $\underline{\alpha}(h, \infty) = \alpha$ 。 $(1-4)$ は、また $\hat{x}_{n+h} - x_{n+h}$ の漸近分散とも漸近的には算しく、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 R , $\|\hat{\alpha}(h, k) - \alpha\|_R^2$ なる loss function は、 h, k を fix したときの $\hat{\alpha}(h, k)$ の漸近分散行列 (= Fisher information matrix) を 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項を最小化していくことが必要となる。

§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 $\alpha(k) = \underline{\alpha}(1, k)$ と表わす。また、 k の上限 \bar{K}_n ($< n$) を与え、 $1 \leq k \leq \bar{K}_n$ の範囲で考える。

Notations.

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

$$\hat{R}(k) = (\hat{r}_{\ell, m} \quad 1 \leq \ell, m \leq k)$$

$$= \frac{1}{n} \sum_{j=K_n}^{n-1} X_{j+1}(k) X'_{j+1}(k)$$

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

$$\hat{\alpha}(k) = (\hat{\alpha}_1(k), \hat{\alpha}_2(k), \dots, \hat{\alpha}_k(k), 0, \dots)$$

$$\therefore \hat{R}(k) \hat{\alpha}(k) = -\hat{r}'(k) \text{ a 解}$$

($\hat{\alpha}(k)$ は k -dim. vector とみなす)

$$e_{j+1,k} = x_{j+1} + \alpha_1(k)x_j + \dots + \alpha_k(k)x_{j+1-k}$$

$$S_k^2 = \frac{1}{n} \sum_{j=K_n}^{n-1} e_{j+1,k}^2$$

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

さうして, $\|\cdot\|$ は通常の Euclid norm を表す。

Remark.

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

Assumptions

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

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

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

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

(A-5) $\{x_t\}$ は finite order a 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 $\bar{K}_n \geq k \geq 1$ に対して

$$\begin{aligned} E\left(\frac{n}{k}\left\|\sum_{j=\bar{K}_n}^{n-1} X_j(k) e_{j+1}/n\right\|_{R(k)^{-1}}^2 - \sigma^2\right)^2 \\ \leq \left(1 - \frac{\bar{K}_n}{n}\right) \frac{c_1}{n} + \left(1 - \frac{\bar{K}_n}{n}\right)^2 \frac{c_2}{k} \end{aligned}$$

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。 \square

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 $\bar{K}_n \geq k \geq 1$ に対して

$$\begin{aligned} E\left(\frac{n}{k}\left\|\sum_{j=\bar{K}_n}^{n-1} X_j(k)(e_{j+1,k} - e_{j+1})/n\right\|^2\right) \\ \leq c_3 \left(1 - \frac{\bar{K}_n}{n}\right) \|a - \underline{a}(k)\|^2 \end{aligned}$$

Proof. see [7]. \square

Lemma 2.3

(A-1), (A-2), (A-4) の仮定のもとで

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

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

たゞたゞし, matrix A の norm は $\|A\| = \sup_{\|x\| \leq 1} \|Ax\|$ で定義される。

$$\text{Proof. } \max_{k \leq \bar{K}_n} \| \hat{R}(k) - R(k) \| \leq \| \hat{R}(\bar{K}_n) - R(\bar{K}_n) \|$$

であることと, [3] より従う。 \square

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

であるのを, 上の $\exists \rightarrow \forall$ lemma と $\|R(k)^{-1}\|$ が bounded であることをより, 次の proposition を得る。

Proposition 2.1

$(A-1) \sim (A-4)$ の仮定のもとで, $1 \leq k_n \leq \bar{K}_n$ で $\liminf_n k_n = \infty$ すなはち \exists a positive integer sequence $\{k_n\}$ $n \rightarrow \infty$

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

したがって, $\| \hat{\alpha}(k_n) - \alpha \|_R^2$ は漸近的

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

である。

Definition 2.1

k_n^* は $L_n(k_n^*) = \min_{1 \leq k \leq K_n} L_n(k)$ とし positive integer.

(A-5) 任意の α と $\tilde{\alpha}$, $\{k_n^*\}$ は $n \rightarrow \infty$ monotone increasing で $\lim_{n \rightarrow \infty} k_n^* = \infty$ とし sequence とする。

Theorem 2.1

(A-1) ~ (A-5) 任意の α と $\tilde{\alpha}$, 任意の $1 \leq k_m \leq K_m$ $K \rightarrow \infty$

$$p - \lim_{n \rightarrow \infty} \frac{\|\hat{\alpha}(k_m) - \alpha\|_R^2}{\|\hat{\alpha}(k_m^*) - \alpha\|_R^2} \geq 1 \quad (2-1)$$

Proof.

$$\frac{\|\hat{\alpha}(k_m) - \alpha\|_R^2}{\|\hat{\alpha}(k_m^*) - \alpha\|_R^2} = \frac{L_n(k_m^*)}{\|\hat{\alpha}(k_m^*) - \alpha\|_R^2} \frac{\|\hat{\alpha}(k_m) - \alpha\|_R^2}{L_n(k_m)} \frac{L_n(k_m)}{L_n(k_m^*)}$$

より proposition 2.1 と 9

$$p - \lim_{n \rightarrow \infty} \frac{L_n(k_m^*)}{\|\hat{\alpha}(k_m^*) - \alpha\|_R^2} = 1$$

よって

$$\begin{aligned} \frac{k}{(n L_n(k))^2} &\leq \frac{1}{n L_n(k) \sigma^2} \\ &\leq \frac{1}{n L_n(k_m^*) \sigma^2} \leq \frac{1}{k_m^* \sigma^4} \end{aligned}$$

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

$$p - \lim_{n \rightarrow \infty} \frac{\|\hat{\alpha}(kn) - \alpha\|_R^2}{L_n(kn)} = 1$$

】

この theorem は kn^* が漸近的に最も optimal の order であることがわかる。しかし、これはパラメータ θ 依存し、一般には unknown。そこで何らかの integer valued statistic $\tilde{kn} \geq kn^*$ 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) の仮定のもとで, $1 \leq l_n \leq m_n \leq \bar{K}_n$,

$\lim_{n \rightarrow \infty} \frac{m_n}{l_n} = 1$ ならば任意の integer の sequences $\{l_n\}$,

$\{m_n\}$ に対して

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

Proof.

lemma 2.1 より

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

$$\leq C'_1 \frac{\bar{K}_n^2}{n} + C'_2 \sum_{k=l_n}^{m_n} \frac{k}{(n L_n(k))^2}$$

よし, lemma 2.2 より

$$\sum_{k=l_n}^{m_n} E \left(\frac{\sqrt{\frac{n}{k}} \left\| \sum_{j=K_n}^{n-1} X_j(k) e_{j+1, k} / n \right\|_{R(k)^{-1}} - \sqrt{\frac{n}{k}} \left\| \sum_{j=K_n}^{n-1} X_j(k) e_{j+1} / n \right\|_{R(k)^{-1}} }{\sqrt{\frac{n}{k} L_n(k)}} \right)^2 \\ \leq C'_3 \frac{\bar{K}_n^2}{n}$$

$l \leq k \leq m$, $\bar{l}_n = \min(l_n, \log k_n^*)$, $m_n' = \min(m_n, \log k_n^*)$ より

$$\sum_{k=\bar{l}_n}^{m_n'} \frac{k}{(n L_n(k))^2} \leq \text{const.} \frac{(\log k_n^*)^2}{k_n^*},$$

$\bar{l}_n'' = \max(l_n, \log k_n^*)$, $m_n'' = \max(m_n, \log k_n^*)$ より

$$\sum_{k=\bar{l}_n''}^{m_n''} \frac{k}{(n L_n(k))^2} \leq \text{const.} \sum_{k=\bar{l}_n''}^{m_n''} \frac{1}{k}$$

よしより, $\{\bar{l}_n\}$, $\{m_n\}$ の条件より

$$p\text{-}\lim_{n \rightarrow \infty} \left(\max_{l_n \leq k \leq m_n} \left| \frac{\|\hat{\alpha}(k) - \underline{\alpha}(k)\|_R^2 - \frac{k}{n} \sigma^2}{L_n(k)} \right| \right) = 0. \quad \square$$

Theorem 2.2

(A-1) ~ (A-5) の仮定をとる, ある sequence $\{k_n\}$

($1 \leq k_n \leq \bar{K}_n$) より positive constant α が存在して

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

であるより order selection $1 \leq \hat{k}_n \leq \bar{K}_n$ $n \rightarrow \infty$ 时、
次の不等式が成立する。

$$p - \lim_{n \rightarrow \infty} \frac{\|\hat{\alpha}(\hat{k}_n) - \alpha\|_R^2}{\|\hat{\alpha}(k_n^*) - \alpha\|_R^2} \geq 1 \quad (2-2)$$

Proof.

proposition 2.2 より $p\text{-}\lim \frac{\hat{k}_n}{k_n} = \alpha$ の条件を用い
ば

$$p - \lim_{n \rightarrow \infty} \frac{\|\hat{\alpha}(\hat{k}_n) - \alpha\|_R^2}{\|\hat{\alpha}(k_n^*) - \alpha\|_R^2} = p - \lim_{n \rightarrow \infty} \frac{L_n(\hat{k}_n)}{L_n(k_n^*)}$$

で、 k_n^* の定義より (2-2) が従う。 \square

ここで、(2-2) の左辺の極限が存在し、等号が成立するとき、 \hat{k} は efficient であるといふこととする。

§3. An efficient selection method

$S_n(k) = (n+2k)\hat{\sigma}_k^2$ と statistic を考へる。ただし

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

すると $S_n(\hat{k}_n) = \min_{1 \leq k \leq \bar{K}_n} S_n(k)$ と \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 $\hat{\alpha}(k)$ の定義はいくぶんちがいがある。

$S_n(k)$ の behaviour は $\nu < \beta$ は critical で、今の段階では上の \hat{k}_n ではないと efficient であるとはいえない。
また、 $S_n(k)$ を R の式に書きかえる。

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

Lemma 3.1

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

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

Proof.

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

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 と同じ仮定のもとで

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

Proof.

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

であることに、ある k, n depend しない constant $c (> 0)$ が存在して

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

と評価されるから、lemma 3.1 より結果が従う。 □

以上で lemma 3.1

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

と表わされる。ただし、 δ_n は $p - \lim_{n \rightarrow \infty} \delta_n = 0$ である。

Lemma 3.3

lemma 3.1 と同じ仮定のもとで

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

Proof.

$k_m^* < k$ のとき

$$\begin{aligned} n(\underline{s}_{k_m^*}^2 - \sigma_{k_m^*}^2) - n(\underline{s}_k^2 - \sigma_k^2) \\ = 2n(\underline{\alpha}(k_m^*) - \underline{\alpha}(k))'(\hat{r}(k_m) - r(k_m)) \\ + n(\|\underline{\alpha}(k_m^*)\|_{\hat{R}(k)}^2 - \|\underline{\alpha}(k_m^*)\|_R^2 + \|\underline{\alpha}(k)\|_R^2 - \|\underline{\alpha}(k)\|_{\hat{R}(k)}^2) \end{aligned}$$

$$E((n(\underline{\alpha}(k_m^*) - \underline{\alpha}(k)))'(\hat{r}(k) - r(k)))^2$$

$$\leq \text{const. } n \|\underline{\alpha}(k_m^*) - \underline{\alpha}(k)\|_R^2,$$

また、

$$\begin{aligned} |n(\|\underline{\alpha}(k_m^*)\|_{\hat{R}(k)}^2 - \|\underline{\alpha}(k_m^*)\|_R^2 + \|\underline{\alpha}(k)\|_R^2 - \|\underline{\alpha}(k)\|_{\hat{R}(k)}^2)| \\ \leq n \|\underline{\alpha}(k_m^*) - \underline{\alpha}(k)\|^2 \|\hat{R}(k) - R(k)\| \\ + 2n |(\underline{\alpha}(k_m^*))'(\hat{R}(k) - R(k))(\underline{\alpha}(k))| \end{aligned}$$

また

$$\begin{aligned} E(n(\underline{\alpha}(k_m^*))'(\hat{R}(k) - R(k))(\underline{\alpha}(k)))^2 \\ \leq \text{const. } n \|\underline{\alpha}(k_m^*) - \underline{\alpha}(k)\|_R |\underline{\alpha}(k) - \underline{\alpha}(k_m^*)| \end{aligned}$$

と評価され、 $k_m^* > k$ のときも全く同じ評価がなる。次に

$$|\underline{\alpha}(k) - \underline{\alpha}(k_m^*)| = \sum_{i=1}^{\max(k, k_m^*)} |\alpha_i(k) - \alpha_i(k_m^*)|.$$

一方

$$\frac{n \|\underline{\alpha}(k_m^*) - \underline{\alpha}(k_m)\|_R |\underline{\alpha}(k_m) - \underline{\alpha}(k_m^*)|}{(n L_n(k_m))^2}$$

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

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

ここで、 k_n^* の定義より

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

であるから、結果が従う。 \blacksquare

以上の結果より、次の proposition を得る。

Proposition 3.1

lemma 3.1 と同じ仮定のもとで

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

と表わされる。ただし、 δ_n, δ_n' は

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

をもつ。

ここで、任意の $\varepsilon > 0$ に対して

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

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) の仮定のもとで、任意の $\varepsilon > 0$ に対して

$$\lim_{n \rightarrow \infty} \Pr(\underline{k}_n^*(\varepsilon) \leq \hat{k}_n \leq \bar{k}_n^*(\varepsilon)) = 1$$

である

$$p - \lim_{n \rightarrow \infty} \frac{L_n(\hat{k}_n)}{L_n(\underline{k}_n^*)} = 1.$$

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

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

である、 proposition 3.1 より

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

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

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

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

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

であるから

$$\lim_{n \rightarrow \infty} \Pr(1 \leq \frac{L_n(\hat{k}_n)}{L_n(\underline{k}_n^*)} \leq 1 + \varepsilon) = 1.$$

□

Theorem 3.2 (\hat{k} の asymptotic efficiency)

(A-1) ~ (A-5) の仮定, より $\hat{\alpha}$

$$\lim_{\varepsilon \rightarrow 0} \left(\lim_{n \rightarrow \infty} \frac{\bar{k}_n^*(\varepsilon)}{k_n^*(\varepsilon)} \right) = 1 \quad (3-1)$$

が成り立つば

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

Proof.

theorem 3.1 の結果, より proposition 2.2 より
従う。 \square

Example

$\{\alpha_i\}$ が exponentially decreasing sequence
ならば,

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

も同様であり, k_n^* は $\log n$ の order なつば。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.