

文章编号: 1005—8893 (2005) 03—0034—04

# 一类线性过程中的经验分布函数的收敛速度<sup>\*</sup>

阮宏顺

(江苏工业学院 信息科学系, 江苏 常州 213016)

摘要: 涉及到一类平稳的线性过程中分布函数的一致收敛速度。考虑线性过程:  $X(n) = \sum_{i=0}^{\infty} \delta(i) Z(n-i)$ , 在如下条件下: ①  $Z(n)$  为 i.i.d.r.v's 且  $E|Z(n)| < +\infty$ ; ②  $Z(n)$  的分布函数  $F$  具有有界密度; ③ 参数  $\delta(i)$  满足  $|\delta(i)| < g(i)$ , 其中函数  $g$  满足  $\sum_{i=1}^{\infty} ig(i) < \infty$ , 给出了  $\sup_{x \in R} |F_n(x) - F(x)| \rightarrow 0$  的速度为  $N^{-\frac{1}{2}} \left( \lg N \left[ \lg \lg \frac{N}{\log N} \right] \right)^{\frac{1}{2}}$ , 在相同的条件下, 比文献 [1] 的速度  $N^{-\frac{1}{2}} (\log N)$  快。

关键词: 线性过程; 分布函数; 收敛速度

中图分类号: O 211

文献标识码: A

本文涉及到一类线性过程:

$$X(n) = \sum_{i=0}^{\infty} \delta(i) Z(n-i) \quad (1)$$

其中参数  $\delta(i)$  满足适当的条件,  $Z(n)$  独立同分布, 其方差不一定有限。象便于观测的时间序列 (如股票价格的变化, 电话噪音等) 作为该过程的一个重要分支。

研究  $X(n) = \sum_{i=0}^{h(m)} \delta(i) Z(n-i)$  ( $n=1, \dots, m$ ) 的经验分布函数的收敛渐近性。对于独立同分布 (i.i.d.) 情形, 经验分布函数的收敛速度及其极限分布已经得到了广泛的研究, Chung<sup>[2]</sup>, Smirnov<sup>[3]</sup> 各自证明了经验分布函数一致收敛的重对数律, 而文献 [4] 中讨论了该过程中核估计的强一致相合性。

## 1 预备知识

本文的所有结果都是在类似的条件下, 为避免重复, 先给出下面条件: ① 设  $Z(n)$  为 i.i.d.r.v's, 且  $E|Z(n)| < +\infty$ ; ②  $X(n)$  的分布函数  $F$  具有有界密度; ③ 参数  $|\delta(i)| \leq g(i)$ , 其中实值函数  $g: N^+ \rightarrow N^+$  (这里或以后出现的  $N^+$  都表示自然数集), 且  $\sum_{i=1}^{\infty} ig(i) < \infty$ ; ④

参数  $\delta(i)$  满足:  $|\delta(i)| < g(i) \equiv c^{\rho^i}$ ,

$0 < \rho < 1$ ,  $c > 0$  (此处或下文所出现的正常数  $c$  不一定相同)。

因为  $E \left| \sum_{i=0}^{\infty} \delta(i) Z(n-i) \right| \leq \sum_{i=0}^{\infty} g(i)$   $E|Z(0)| < \infty$ , 在条件①、②、③或①、②、④下, 在式 (1) 中的无穷和以概率 1 地绝对收敛。于是, 对于整数  $j, k$ ,  $(X(1), \dots, X(k))$  的联合分布与  $(X(1+j), \dots, X(k+j))$  的联合分布相同, 因此线性过程  $X(n)$  几乎处处存在, 且是严平稳的。

定义整值非降函数  $h: N^+ \rightarrow N^+$ , 当  $m \rightarrow \infty$  时,  $h(m) \rightarrow \infty$ , 对充分大的  $m$ , 有  $h(m) < m$ , 且  $\sum_{i=1}^{\infty} \frac{ig(h(m)+i)}{g(h(m))} = O(1)$ , 记

$$\begin{cases} V_m(n) \triangleq \sum_{i=0}^{h(m)-1} \delta(i) Z(n-i), & n=1, \dots, m \\ W_m(n) \triangleq \sum_{i=h(m)}^{\infty} \delta(i) Z(n-i), & n=1, \dots, m \end{cases} \quad (2)$$

显然, 对于所有的  $m$  及  $n=1, \dots, m$ , 有  $X(n) = V_m(n) + W_m(n)$

现考察随机变量组:

\* 收稿日期: 2005—05—08

作者简介: 阮宏顺 (1955—), 男, 安徽安庆人, 副教授, 硕士, 主要从事概率极限理论研究。

$$\left\{ \begin{array}{l} V_1(1) \\ V_2(1), V_2(2) \\ \dots \\ V_M(1), V_M(2), \dots, V_M(M) \\ M=1, 2, \dots, m (m \in N^+) \end{array} \right. \quad (3)$$

每1行包含1组同分布的随机变量。记  $F_{V_M}(x) = P\{V_M(1) \leq x\}$  为第  $M$  行随机变量的分布函数。对于每个  $M$ , 当  $|i-j| \geq h(M)$  时,  $V_M(i)$  与  $V_M(j)$  相互独立。因此集合

$$S_M^k = \{V_M(k), V_M(k+h(M)), \dots, V_M(k+l_M^k h(M))\}, k=1, 2, \dots, h(M)$$

中的随机变量相互独立, 且有

$$\bigcap_k S_M^k = \emptyset \quad (\emptyset \text{ 为空集})$$

$$\bigcup_k S_M^k = \{V_M(i); i=1, \dots, M\}$$

其中常数  $l_M^k = [M/h(M)]$  或  $[M/h(M)] - 1$ ,  $[x]$  表示  $x$  的整数部分。

## 2 主要结果及证明

引理 1<sup>[1]</sup>: 对于 (2) 式中的  $W_m(n)$  和  $h$ , 在条件①和③下存在常数  $c > 0$ , 对于所有的  $n \leq m$  有  $|W_m(n)| \leq O(mg(h(m)))$  a. s.

(4)

引理 2<sup>[4]</sup>: 若满足条件①, ②, ③, 则  $m \rightarrow \infty$  时  $\sup_{x \in \mathbb{R}} |F(x) - F_{V_m}(x)| = O((g(h(m)))^{1/2})$

(5)

其中  $F, F_{V_m}$  分别为  $X(n), V_m(n)$  的分布函数。

[注]: 如果取  $g(i) = c_0^i$ ,  $h(N) = [c_0 \log N]$ , 那么式 (4)、式 (5) 中的速度为  $O(N^{c_0 \log^{c_0+1}})$ , 可选择足够大的  $c_0$ , 使收敛速度加快。

引理 3<sup>[6]</sup>: (Hartman-Wintner 重对数律) 设  $\{X_n: n \geq 1\}$  是 i. i. d. r. v 序列,  $EX_n = 0, EX_n^2 = 1$ 。则

$$\lim_{n \rightarrow \infty} \sup \frac{\sum_{i=1}^n X_i}{\sqrt{2n \log \log n}} = 1 \quad \text{a. s.}$$

定理: 设线性过程 (1) 满足条件①, ②, ④则有

$$P\left\{\lim_{N \rightarrow \infty} \sup N^{\frac{1}{2}} ((\log N) \log \log \frac{N}{\log N})^{-\frac{1}{2}} \sup_{x \in \mathbb{R}} |F(x) - F(x)| < \infty\right\} = 1$$

$$\sup_{x \in \mathbb{R}} |F(x) - F(x)| < \infty = 1$$

其中  $F(x)$  是关于  $X(1), \dots, X(N)$  的经验分布函数。

证明:

$$|F(x) - F(x)| = |F(F^{-1}(t)) - t|$$

a. s.

其中  $x = F^{-1}(t) = \inf\{y: F(y) \geq t\}$ , 于是

$$\sup_{x \in \mathbb{R}} |F(x) - F(x)| = \sup_{0 \leq t \leq 1} |H_N(t) - t|$$

其中  $H_N(t)$  是随机变量  $F(X(n))$  的经验分布函数。而  $F(X(n))$  服从区间  $[0, 1]$  上的均匀分布。因此只要证明

$$P\left\{\lim_{N \rightarrow \infty} \sup N^{\frac{1}{2}} ((\log N) \log \log \frac{N}{\log N})^{-\frac{1}{2}} \sup_{0 \leq t \leq 1} |H_N(t) - t| < \infty\right\} = 1 \quad (6)$$

这里和下面要用以下记号:  $F, H, H^{(N)}$  分别是  $X(n), F(X(n)), F(V_N(n))$  的分布函数。其带有下标  $N$  总是表示相应的经验分布函数, 如  $H_N^{(N)}$  是关于  $F(V_N(n))$  ( $n=1, \dots, N$ ) 的经验分布函数。显然

$$\begin{aligned} \sup_{0 \leq t \leq 1} |H_N(t) - t| &= \sup_{0 \leq t \leq 1} |H_N(t) - H_N^{(N)}(t) + H_N^{(N)}(t) - H^{(N)}(t) + H^{(N)}(t) - t| \leq \\ &\sup_{0 \leq t \leq 1} |H_N(t) - H_N^{(N)}(t)| + \sup_{0 \leq t \leq 1} |H_N^{(N)}(t) - H^{(N)}(t)| + \sup_{0 \leq t \leq 1} |H^{(N)}(t) - t| \triangleq \text{I} + \text{II} + \text{III} \end{aligned} \quad (7)$$

下面分别对它们进行论证。

为了对 II 进行估值, 有必要将区间  $[0, 1]$  进行分割。令  $t_{N,i} = i/N$  ( $i=0, 1, \dots, N$ ), 那么, 对于  $t_{N,i} \leq t \leq t_{N,i+1}$ , 有

$$H_N^{(N)}(t_{N,i}) - H^{(N)}(t_{N,i+1}) \leq H_N^{(N)}(t) - H^{(N)}(t) \leq H_N^{(N)}(t_{N,i+1}) - H^{(N)}(t_{N,i})$$

$$\begin{aligned} \text{如果 } |H_N^{(N)}(t_{N,i}) - H^{(N)}(t_{N,i+1})| \geq |H_N^{(N)}(t_{N,i+1}) - H^{(N)}(t_{N,i})|, \text{ 那么, 对于 } t_{N,i} \leq t \leq t_{N,i+1}, \\ |H_N^{(N)}(t) - H^{(N)}(t)| \leq |H_N^{(N)}(t_{N,i}) - H^{(N)}(t_{N,i+1})| \leq |H_N^{(N)}(t_{N,i}) - H^{(N)}(t_{N,i})| + (H^{(N)}(t_{N,i+1}) - H^{(N)}(t_{N,i})). \end{aligned}$$

$$\begin{aligned} \text{如果 } |H_N^{(N)}(t_{N,i}) - H^{(N)}(t_{N,i+1})| < |H_N^{(N)}(t_{N,i+1}) - H^{(N)}(t_{N,i})|, \text{ 那么 } |H_N^{(N)}(t) - H^{(N)}(t)| \leq |H_N^{(N)}(t_{N,i+1}) - H^{(N)}(t_{N,i})| \\ |H_N^{(N)}(t_{N,i})| \leq |H_N^{(N)}(t_{N,i+1}) - H^{(N)}(t_{N,i+1})| + (H^{(N)}(t_{N,i+1}) - H^{(N)}(t_{N,i})) \end{aligned}$$

因此

$$\begin{aligned} \sup_{0 \leq t \leq 1} |H_N^{(N)}(t) - H^{(N)}(t)| \leq \max_{0 \leq i \leq N} |H_N^{(N)}(t_{N,i}) - H^{(N)}(t_{N,i})| + \max_{0 \leq i \leq N} (H^{(N)}(t_{N,i+1}) - H^{(N)}(t_{N,i})) \triangleq \text{A} + \text{B} \end{aligned}$$

为了估计  $B$ , 考察

$$\begin{aligned} H^{(N)}(t_{N,i+1}) - H^{(N)}(t_{N,i}) &= P(V_N(n) \leq F^{-1}(t_{N,i+1})) - P(V_N(n) \leq F^{-1}(t_{N,i})) \leq \\ &F(F^{-1}(t_{N,i+1}) + O(N^{1/2(c_0 \log \rho + 1)})) - \\ &F(F^{-1}(t_{N,i}) - O(N^{1/2(c_0 \log \rho + 1)})) \leq \\ &O(N^{-1}) + O(N^{1/2(c_0 \log \rho + 1)}) \end{aligned}$$

这里  $c_0$  足够大, 并注意  $F$  的密度有界及引理 1, 当  $N \rightarrow \infty$  时, 进一步有

$$\max_{0 \leq i \leq N} (H^{(N)}(t_{N,i+1}) - H^{(N)}(t_{N,i})) = O(N^{-1})$$

为了估计  $A$ , 记  $H_N^{(N)}(t_{N,i}) \triangleq \frac{1}{N} \sum_{k=1}^N U_{k,i}$ ,

其中  $U_{k,i} = \begin{cases} 1 & F(V_N(k)) \leq t_{N,i} \\ 0 & \text{otherwise} \end{cases}$ , 对于每个

$k$  和  $i$ , 随机变量  $U_{k,i}$  与  $N$  有关。从 (3) 式可以看出, 对于每个  $N$  和  $i$ , 当  $|r_1 - r_2| \geq h(N)$  时, 随机变量  $U_{r_1,i}$  和  $U_{r_2,i}$  相互独立。根据条件 ④, 取  $h(N) = [c_0 \log N]$ , 在下文中都是采用这种取法。记

$$S_j \triangleq \sum_{i=0}^{N_j-1} U_{h(N) \cdot j + i} \quad (8)$$

$$\text{其中 } N_j = \left[ \frac{N}{h(N)} \right] + 1 = O\left(\frac{N}{\log N}\right)$$

由 Hartman-Wintner 重对数律得, 存在常数  $G$ , 使

$$\limsup_{N \rightarrow \infty} \frac{S_j - ES_j}{\sqrt{2 \frac{N}{\log N} \log \log \frac{N}{\log N}}} < G \quad \text{a. s.}$$

于是当  $N$  充分大时, 有

$$\frac{S_j - ES_j}{\sqrt{2 \frac{N}{\log N} \log \log \frac{N}{\log N}}} < G \quad \text{a. s.}$$

由于平稳性, 从而有

$$\frac{\sum_{j=1}^{h(N)} (S_j - ES_j)}{\sqrt{2 \frac{N}{\log N} \log \log \frac{N}{\log N}}} < G \circ h(N) \quad \text{a. s.}$$

即存在常数  $G'$ , 使

$$\frac{N^{\frac{1}{2}}}{\left[ \log N \log \log \frac{N}{\log N} \right]^{\frac{1}{2}}} (H_N^{(N)}(t_{N,i}) - H^{(N)}(t_{N,i})) < G' \quad \text{a. s.}$$

进一步有

$$\frac{N^{\frac{1}{2}}}{\left[ \log N \log \log \frac{N}{\log N} \right]^{\frac{1}{2}}} \max_{0 \leq i \leq N} (H_N^{(N)}(t_{N,i}) - H^{(N)}(t_{N,i})) < G' \quad \text{a. s.}$$

这表明 (7) 式中的第 2 项 (II) 具有所需要的阶。

(7) 式中的第 3 项 (III) 比较容易估计。由于  $F$  具有有界导数, 显然有

$$\max_{0 \leq n \leq N} |F(V_N(n)) - F(X(n))| \leq \max_{0 \leq n \leq N} |W_N(n)|$$

当  $N \rightarrow \infty$  时, 结合引理 2 得

$$\sup_{0 \leq t \leq 1} |H^{(N)}(t) - t| = O(Ng(h(N))) = O(N^{-\lambda})$$

由 [注] 知,  $\lambda$  可以任意大。

最后考察 (7) 式中的第 1 项 (I), 与前面的论证相类似。

$$\begin{aligned} \sup_{0 \leq t \leq 1} |H_N(t) - H_N^{(N)}(t)| &\leq \\ \max_{0 \leq i \leq N} |H_N(t_{N,i}) - H_N^{(N)}(t_{N,i})| &+ \\ \max_{0 \leq i \leq N} (H_N^{(N)}(t_{N,i+1}) - H_N^{(N)}(t_{N,i})) &\triangleq C + D \end{aligned}$$

下面先讨论第 1 项 (C)。

由贝努里-马尔可夫不等式得  $\sum_N P$

$$\left\{ \frac{\max_{0 \leq i \leq N} N^{\frac{1}{2}}}{\left[ \log N \log \log \frac{N}{\log N} \right]^{\frac{1}{2}}} |H_N(t_{N,i}) - H_N^{(N)}(t_{N,i})| > c \right\} \leq \sum_N \sum_i E \frac{|H_N(t_{N,i}) - H_N^{(N)}(t_{N,i})|}{c} \frac{N^{\frac{1}{2}}}{\left[ \log N \log \log \frac{N}{\log N} \right]^{\frac{1}{2}}}$$

由于

$$E |H_N(t_{N,i}) - H_N^{(N)}(t_{N,i})| \leq N^{-1} \sum_{i=1}^N E |\chi(t_{N,i}, V_N(n), X(n))|$$

其中

$$\chi(t_{N,i}, V_N(n), X(n)) = \begin{cases} -1, & F(V_N(n)) \leq t_{N,i} \text{ 且 } F(X(n)) > t_{N,i} \\ 1, & F(V_N(n)) > t_{N,i} \text{ 且 } F(X(n)) \leq t_{N,i} \\ 0, & \text{其它} \end{cases}$$

由全数学期望得

$$E |\chi(t_{N,i}, V_N(n), X(n))| \leq \int_0^{N^{-\eta}} \{P(t_{N,i} - x \leq F(V_N(n)) \leq t_{N,i}) + P(t_{N,i} \leq F(V_N(n)) \leq t_{N,i} + x)\} d\Gamma(x) + P(|F(X(n)) - F(V_N(n))| \geq N^{-\eta})$$

其中  $\eta$  某一正常数,  $\Gamma(x)$  是  $|F(X(n)) - F(V_N(n))|$  的分布函数, 即

$$\Gamma(x) = P(|F(X(n)) - F(V_N(n))| \leq x)$$

记  $f$  是  $F$  的导数, 由平稳的分布具有有界导数及引理 1 得

$$P\{|F(X(n)) - F(V_N(n))| \geq N^{-\eta}\} \leq N^\eta E \left\{ \sup_{x \in \mathbb{R}} |W_N(n)| \right\} \leq N^{\eta-\lambda}$$

由于上面涉及到的积分的阶为  $O(N^{-\eta})$ , 故当  $N \rightarrow \infty$  时, 有

$$E|\chi(t_{N,i}, V_N(n), X(n))| \leq O(N^{-\eta}) + O(N^{\eta-\lambda})$$

由于取  $\eta = \frac{\lambda}{2}$ , 其阶为  $O(N^{-\frac{\lambda}{2}})$  以及通过选择  $c_0$  可使  $\lambda$  任意大, 使

$$\sum_N \sum_i \frac{c^{-1} N^{\frac{1}{2}}}{\left[ \log N \log \log \frac{N}{\log N} \right]^{\frac{1}{2}}} \cdot O(N^{-\frac{\lambda}{2}}) < \infty$$

由 Borel-Cantelli 定理, 项 (C) 具有所需要的阶。

最后, 考虑项 (D), 只要证明

$$\sum_N P \left\{ \frac{N^{\frac{1}{2}}}{\left[ \log N \log \log \frac{N}{\log N} \right]^{\frac{1}{2}}} \max_{0 \leq i \leq N} (H_N^{(N)}(t_{N,i+1}) - H_N^{(N)}(t_{N,i})) > c \right\} < \infty$$

记

$$\Psi \triangleq P \left\{ \frac{N^{\frac{1}{2}}}{\left[ \log N \log \log \frac{N}{\log N} \right]^{\frac{1}{2}}} (H_N^{(N)}(t_{N,i+1}) - H_N^{(N)}(t_{N,i})) > c \right\} = P \left\{ N^{-1} \sum_{i=1}^N (U_{k,i} - E U_{k,i}) > c N^{-\frac{1}{2}} \left[ \log N \log \log \frac{N}{\log N} \right]^{\frac{1}{2}} - E U_{k,i} \right\} \quad (9)$$

$$\text{其中 } U_{k,i} = \begin{cases} 1, & t_{N,i} \leq F(V_N(n)) \leq t_{N,i+1} \\ 0, & \text{otherwise} \end{cases}$$

显然  $E U_{k,i} = H^{(N)}(t_{N,i+1}) - H^{(N)}(t_{N,i})$

前面已证明它的阶为  $O(N^{-1})$ , 类似于 (8)

式运用到 (9) 式, 记  $S_j \triangleq \sum_{i=0}^{N_j-1} U_{h(N)l+j,i}$

$$\Psi \leq \sum_{j=1}^{h(N)} P(S_j - E S_j > N_j (c N^{-\frac{1}{2}} \left[ \log N \log \log \frac{N}{\log N} \right]^{\frac{1}{2}} - E U_{k,i})) \leq$$

$$\sum_{j=1}^{[c_0 \log N]} \exp \left[ -c N^{\frac{1}{2}} \left( \frac{1}{\log N} \log \log \frac{N}{\log N} \right)^{\frac{1}{2}} \right] E \exp(S_j - E S_j) \leq c \log N \exp \left[ -c N^{\frac{1}{2}} \left( \frac{1}{\log N} \log \log \frac{N}{\log N} \right)^{\frac{1}{2}} \right]$$

从而

$$\sum_N P \left\{ N^{\frac{1}{2}} \left[ \log N \log \log \frac{N}{\log N} \right]^{-\frac{1}{2}} \max_{0 \leq i \leq N} (H_N^{(N)}(t_{N,i+1}) - H_N^{(N)}(t_{N,i})) > c \right\} \leq \sum_N N c \log N \exp \left[ -c N^{\frac{1}{2}} \left( \frac{1}{\log N} \log \log \frac{N}{\log N} \right)^{\frac{1}{2}} \right]$$

这是有限值。证毕。

### 参考文献:

- [1] Hesses C H. Rates of Convergence for the Empirical Distribution Function and the Empirical Characteristic Function of a Broad Class of Linear Processes [J]. J Mult Anal. 1990, 35: 186-220.
- [2] Chung K L. An Estimate Concerning the Kolmogorov Limit Distribution [J]. Trans Amer Math Soc. 1949, 67: 36-50.
- [3] Smirnov N V. Approximate Laws of Distribution Random Variables from Empirical Data [J]. Uspehi Mat Nauk. 1944, 10: 179-206.
- [4] 阮宏顺, 许波. 线性过程中核估计的强一致相容性 [J]. 江苏石油化工学院学报, 1998, 10 (1): 38-40.
- [5] 阮宏顺. 一类线性过程中的分布函数的收敛速度 [J]. 江苏工业学院学报, 2003, 15 (1): 47-49.
- [6] 陆传荣, 林正炎, 陆传贵. 概率论极限理论引论 [M]. 北京: 高等教育出版社, 1989.

## Rate of Convergence for the Empirical Distribution Function of a Broad Class of Linear Processes

RUAN Hong-shun

(Department of Information Science, Jiangsu Polytechnic University, Changzhou 213016, China)

**Abstract:** This paper deals with uniform rate of convergence for a broad class of stationary linear processes. In particular, the class is considered under the condition: ① a finite first absolute moment, ② the distribution function has bounded density, and ③ the parameters are bounded in absolute value by some function which satisfies. It is proved that the empirical distribution function converges to, uniformly in, at a rate which is faster than that of [1] under the same condition.

**Key words:** linear process; distribution function; rate of convergence