第3卷 第2期
1983年4月

数学研究与评论
JOURNAL OF MATHEMATICAL RESEARCH AND EXPOSITION

Vol.3 No.2
Apr. 1983

线性模型中误差方差估计的指数收敛速度*

苏淳

（中国科学技术大学）

考虑线性回归模型

$$y_j = x_j'\beta + \mu_j, \quad j = 1, 2, \cdots \tag{1}$$

其中 $\{\mu_j\}$ 为独立的试验误差序列，满足条件：

$$E\mu_j = 0, \ \mathrm{Var}\,\mu_j = \sigma^2, \ 0 < \sigma^2 < \infty, \ j = 1, 2, \cdots \tag{2}$$

根据最小二乘法规定和模型自身特点，通常，在考虑极限性质时，给 σ^2 以估计．

$$\hat{\sigma}_n^2 = \frac{1}{n-r}\left\{\sum_{j=1}^{n}\mu_j^2 - \sum_{k=1}^{r}\left(\sum_{j=1}^{n} a_{nkj}\mu_j\right)^2\right\}. \tag{3}$$

其中 r 是一个与设计矩阵 $X_n = (x_1 \vdots \cdots \vdots x_n)$ 有关的常正整数，数列 $\{a_{nkj}\}$ 满足等式 $\sum_{j=1}^{n} a_{nkj}a_{nlj} = \delta_{kl}$，这里 δ_{kl} 为 Kronecker 符号，$n = 1, 2, \cdots$．本文旨在讨论 $\hat{\sigma}_n^2 \xrightarrow{\text{in pr.}} \sigma^2$ 的指数收敛速度，主要结果是：

定理 如果 $\mu_1, \mu_2 \cdots$ 独立，满足（2）式，对每个 $\eta > 0$，存在仅依赖于 η 的常数 $M_\eta > 0$，$t_\eta > 0$，使得对一切 $|t| \leqslant t_\eta$，一切 n，有

$$\prod_{j=1}^{n} Ee^{t(\mu_j^2 - \sigma^2)} = Ee^{t\sum_{j=1}^{n}(\mu_j^2 - \sigma^2)} \leqslant M_\eta e^{n\eta|t|} \tag{4}$$

又如果对一切 n, k, j，存在常数 $0 < L < \infty$，使得 $0 \leqslant a_{nkj}^2 \leqslant \frac{L}{n}$．则对任何充分小的 $\varepsilon > 0$，存在仅依赖于 ε 的常数 $A = A(\varepsilon) > 0$，和 $0 < \rho = \rho(\varepsilon) < 1$，使得

$$P\{|\hat{\sigma}_n^2 - \sigma^2| \geqslant \varepsilon\} \leqslant A\rho^r. \tag{5}$$

为证定理，先证一个引理．

引理 设 $\mu_1, \mu_2, \cdots$ 独立、对称，满足条件（2）、（4），又存在 $0 < L < \infty$，使 $0 < a_{nj}^2 \leqslant \frac{L}{n}$，且 $\sum_{j=1}^{n} a_{nj}^2 = 1$，则对任何充分小的 $\varepsilon > 0$ 和充分大的 n，存在仅依赖于 ε 的常数 $A > 0, 0 < \rho < 1$，使得

$$P\left\{\left|\left(\sum_{j=1}^{n} a_{nj}\mu_j\right)^2 - \sigma^2\right| \geqslant \frac{n-r}{r}\varepsilon\right\} \leqslant A\rho^n; \tag{6}$$

证 易知，只须证明

$$J_n = P\left\{\left(\sum_{j=1}^{n} a_{nj}\mu_j\right)^2 \geqslant \frac{n-r}{r}\varepsilon\right\} \leqslant A\rho^n. \tag{7}$$

为方便计，记 $a_{nj} = a_j$，并设 $\sigma^2 = \frac{1}{L}$．若不然，可令 $\tilde{\mu}_j = \frac{\mu_j}{\sigma\sqrt{L}}$，以 $\varepsilon_1 = \frac{\varepsilon}{\sigma^2 L}$ 代 ε，由条件(4)中 $\eta > 0$ 的任意性，知存在 $\widetilde{M}_\eta > 0, \tilde{t}_\eta > 0$，使当 $|t| \leqslant \tilde{t}_\eta$ 时对一切 n，仍有 $\prod_{j=1}^{n} Ee^{t(\tilde{\mu}_j^2 - \frac{1}{L})} = \prod_{j=1}^{n} Ee^{\frac{t}{\sigma^2 L}(\mu_j^2 - \sigma^2)} \leqslant \widetilde{M}_\eta e^{|t|n\eta}$．事实上，只要令 $\tilde{\eta} = \sigma^2 L\eta$，取 $\widetilde{M}_\eta = M_{\tilde{\eta}}$，并使 $\left|\frac{t}{\sigma^2 L}\right| \leqslant |t_{\tilde{\eta}}|$ 即可．

*1981年6月29日收到．推荐者：陈希孺（中国科技大学）．

由 Чебышев 不等式，对一切 $t>0$，有

$$J_n \leqslant e^{t\left(1-\frac{n}{r}\right)\varepsilon} E e^{t\left(\sum_{j=1}^{n} a_j \mu_j\right)^2} = e^{t\left(1-\frac{n}{r}\right)\varepsilon} \sum_{k=0}^{\infty} \frac{t^k E\left(\sum_{j=1}^{n} a_j \mu_j\right)^{2k}}{k!}$$

$$= e^{t\left(1-\frac{n}{r}\right)\varepsilon} \left\{\left(\sum_{0<k<\frac{\varepsilon}{r}n} + \sum_{\frac{\varepsilon}{r}n<k<\frac{\ln 2}{2}n} + \sum_{k>\frac{\ln 2}{2}n}\right) \frac{t^k E\left(\sum_{j=1}^{n} a_j \mu_j\right)^{2k}}{k!}\right\}$$

$$\triangleq J_{1n} + J_{2n} + J_{3n} \tag{8}$$

由 $\mu_1, \mu_2, \cdots$ 独立、对称，知 $E\left(\sum_{j=1}^{n} a_j \mu_j\right)^{2k}$ 中的非零项皆具形式 $\frac{(2k)!}{(2i_1)!(2i_2)!\cdots(2i_l)!}$ $E(a_{j_1}\mu_{j_1})^{2i_1} E(a_{j_2}\mu_{j_2})^{2i_2} \cdots E(a_{j_l}\mu_{j_l})^{2i_l}$，其中 $i_1 + \cdots + i_l = k$。由 Stirling 公式知

$$\frac{\dfrac{(2k)!}{(2i_1)!(2i_2)!\cdots(2i_l)!}}{\dfrac{k!}{i_1! i_2! \cdots i_l!}} \leqslant \frac{(2k)^{2k+\frac{1}{2}} i_1^{i_1+\frac{1}{2}} i_2^{i_2+\frac{1}{2}} \cdots i_l^{i_l+\frac{1}{2}} e^{\frac{1}{12}\left[\frac{1}{2k}+\sum_{j=1}^{l}\frac{1}{i_j}\right]}}{k^{k+\frac{1}{2}} (2i_1)^{2i_1+\frac{1}{2}} (2i_2)^{2i_2+\frac{1}{2}} \cdots (2i_l)^{2i_l+\frac{1}{2}}}$$

$$\leqslant \frac{k^k}{i_1^{i_1} i_2^{i_2} \cdots i_l^{i_l}} 2^{\frac{1}{2}(1-l)} e^{\frac{1}{12}\min(k,n)} \leqslant l^k 2^{\frac{1}{2}(1-l)} e^{\frac{1}{12}\min(k,n)} \tag{9}$$

这里，不等式 $\frac{k^k}{i_1^{i_1} i_2^{i_2} \cdots i_l^{i_l}} \leqslant l^k$ 可由 $f(t) = t\ln t$ 在 $(0, +\infty)$ 中的凸性得出。又通过对 $g_1(l) = l^k 2^{\frac{1}{2}(1-l)}$ 求导，知

$$l^k 2^{\frac{1}{2}(1-l)} \leqslant k^k 2^{\frac{1}{2}(1-k)} \qquad (1 \leqslant l \leqslant k) \tag{10}$$

综合 (9)(10) 二式，知当 $\frac{\varepsilon}{r}n < k < \frac{\ln 2}{2}n$ 时，有

$$E\left(\sum_{j=1}^{n} a_j \mu_j\right)^{2k} \leqslant \sqrt{2}\left(\frac{\ln 2}{2}\right)^{\frac{\varepsilon}{r}n} \cdot n^k \left(\frac{e^{\frac{1}{12}}}{\sqrt{2}}\right)^k E\left(\sum_{j=1}^{n} a_j^2 \mu_j^2\right)^k \leqslant \sqrt{2}\left(\frac{\ln 2}{2}\right)^{\frac{\varepsilon}{r}n} \cdot L^k E\left(\sum_{j=1}^{n} \mu_j^2\right)^k$$

所以

$$J_{2n} \leqslant \sqrt{2} e^{t\left(1-\frac{n}{r}\right)\varepsilon} \left(\frac{\ln 2}{2}\right)^{\frac{\varepsilon}{r}n} \sum_{\frac{\varepsilon n}{r}<k<\frac{\ln 2}{2}n} \frac{t^k L^k E\left(\sum_{j=1}^{n} \mu_j^2\right)^k}{k!} \leqslant \sqrt{2} e^{t\left(1-\frac{n}{r}\right)\varepsilon + nt} \left(\frac{\ln 2}{2}\right)^{\frac{\varepsilon n}{r}} E e^{Lt\sum_{j=1}^{n}\left(\mu_j^2 - \frac{1}{L}\right)}$$

命 $\eta = \frac{1}{L} > 0$，取 $0 < t < \frac{t_\eta}{L}$ 足够小，使 $0 < \rho_1 \triangleq e^{t\left(2-\frac{\varepsilon}{r}\right)} \left(\frac{\ln 2}{2}\right)^{\frac{\varepsilon}{r}} < 1$，由条件(4)知：存在仅依赖于 ε 的常数 $A_1 > 0$，使

$$J_{2n} \leqslant A_1 \rho_1^n \tag{11}$$

又通过对 $g_2(l) = \left(\frac{1}{n}\right)^{\frac{\ln 2}{2}n} 2^{\frac{1}{2}(n-l)}$ 取导数，可以证得：当 $k \geqslant \frac{\ln 2}{2}n$ 且 $1 \leqslant l \leqslant n$ 时，有 $\left(\frac{1}{n}\right)^k 2^{\frac{n-l}{2}} \leqslant 1$ 记 $\rho_2 = \frac{e^{\frac{1}{12}}}{\sqrt{2}} < 1$，我们就有

$$l^k 2^{\frac{1}{2}(1-l)} e^{\frac{1}{12}\min(n,k)} = \sqrt{2} n^k 2^{-\frac{n}{2}} e^{\frac{1}{12}\min(k,n)} \cdot \left(\frac{1}{n}\right)^k 2^{\frac{1}{2}(n-l)} \leqslant \sqrt{2} n^k \rho_2^n.$$

结合(9)式，立刻得到

$$J_{3n} \leqslant e^{t\left(1-\frac{n}{r}\right)\varepsilon} \sqrt{2} \rho_2^n \sum_{k>\frac{\ln 2}{2}n} \frac{(nt)^k E\left(\sum_{j=1}^{n} a_j^2 \mu_j^2\right)^k}{k!} \leqslant$$

$$e^{t(1-\frac{n}{r})\varepsilon}\sqrt{2}\rho_2^n\sum_{k\geq\frac{\ln 2}{2}n}\frac{(Lt)^k E\left(\sum_{j=1}^n\mu_j^2\right)^k}{k!}\leqslant\sqrt{2}\rho_2^n e^{t(1-\frac{n}{r})\varepsilon+nt}Ee^{Lt\sum_{j=1}^n(\mu_j^2-\frac{1}{L})}$$

仍命 $\eta=\frac{1}{L}>0$，取 $0<t<\frac{t_\eta}{L}$ 足够小，使 $0<\rho_3\triangleq e^{t(2-\frac{\varepsilon}{r})}\rho_2<1$，由条件 (4) 得

$$J_{3n}\leqslant\sqrt{2}M_\eta\rho_2^n e^{t(1-\frac{n}{r})\varepsilon+2tn}\triangleq A_2\rho_3^n \tag{12}$$

取定 $0<t_0<\frac{t_\eta}{L}$，使其满足 J_{2n}，J_{3n} 中的要求，我们来估计 J_{1n}。取 θ：$0<\theta<1-\frac{e^{\frac{1}{12}}}{\sqrt{2}}$，取 $\eta_1=\frac{1}{L}\theta t_0$，由条件（4）知：存在 $M_{\eta_1}>0$，$t_{\eta_1}>0$，使当

$$0<\frac{\varepsilon L}{r}<t_{\eta_1}=t_{\frac{\theta}{L}t_0} \tag{13}$$

时，就有

$$Ee^{\frac{\varepsilon L}{r}\left(\sum_{j=1}^n(\mu_j^2-\frac{1}{L})\right)}\leqslant M_{\eta_1}e^{\frac{\varepsilon L}{r}\eta_1 n}=M_\theta e^{\frac{\varepsilon}{r}\theta n t_0} \tag{14}$$

由于可取 $\varepsilon>0$ 充分小，所以 (13)(14) 的成立是没有问题的。不难证明：

$$\lim_{n\to\infty}e^{-\frac{\varepsilon}{r}n}\sum_{k>[\frac{\varepsilon}{r}n]+1}\frac{\left(\frac{\varepsilon}{r}n\right)^k}{k!}=\frac{1}{2}$$

所以当 n 充分大，就有 $\sum_{k>\frac{\varepsilon}{r}n}\frac{\left(\frac{\varepsilon}{r}n\right)^k}{k!}\geqslant\frac{1}{4}e^{\frac{\varepsilon}{r}n}$。于是

$$E\left(\sum_{j=1}^n a_j^2\mu_j^2\right)^{[\frac{\varepsilon}{r}n]}\frac{e^{\frac{\varepsilon}{r}n}}{4}\leqslant E\left(\sum_{j=1}^n a_j^2\mu_j^2\right)^{[\frac{\varepsilon}{r}n]}\sum_{k>\frac{\varepsilon}{r}n}\frac{\left(\frac{\varepsilon}{r}n\right)^k}{k!}$$

$$\leqslant\sum_{k>\frac{\varepsilon}{r}n}\frac{\left(\frac{\varepsilon}{r}n\right)^k}{k!}\left\{1+E\left(\sum_{j=1}^n a_j^2\mu_j^2\right)^k\right\}\leqslant e^{\frac{\varepsilon}{r}n}+Ee^{\frac{\varepsilon L}{r}\left(\sum_{j=1}^n\mu_j^2\right)},$$

从而

$$E\left(\sum_{j=1}^n a_j^2\mu_j^2\right)^{[\frac{\varepsilon}{r}n]}\leqslant 4+4Ee^{\frac{\varepsilon L}{r}\left(\sum_{j=1}^n(\mu_j^2-\frac{1}{L})\right)}\leqslant 4+4M_\theta e^{\frac{\varepsilon}{r}\theta t_0 n}. \tag{15}$$

于是，当 $1\leqslant k\leqslant\frac{\varepsilon}{r}n$ 时，由（9）(10) 可知

$$E\left(\sum_{j=1}^n a_j\mu_j\right)^{2k}\leqslant e^{\frac{k}{12}}k^k 2^{\frac{1}{2}(1-k)}E\left(\sum_{j=1}^n a_j^2\mu_j^2\right)^k\leqslant\sqrt{2}\left(\frac{e^{\frac{1}{12}}\varepsilon}{\sqrt{2}r}\right)^n n^k E\left(\sum_{j=1}^n a_j^2\mu_j^2\right)^k$$

$$\leqslant\sqrt{2}\left(\frac{e^{\frac{1}{12}}\varepsilon}{\sqrt{2}r}\right)^k n^k\left\{1+E\left(\sum_{j=1}^n a_j^2\mu_j^2\right)^{[\frac{\varepsilon}{r}n]}\right\}\leqslant\sqrt{2}\left(\frac{e^{\frac{1}{12}}\varepsilon}{\sqrt{2}r}\right)^k n^k\left\{5+4M_\theta e^{\frac{\varepsilon}{r}\theta t_0 n}\right\},$$

所以

$$J_{1n}\leqslant\sqrt{2}e^{t(1-\frac{n}{r})\varepsilon}\left\{5+4M_\theta e^{\frac{\varepsilon}{r}\theta t_0 n}\right\}\sum_{k=0}^{[\frac{\varepsilon}{r}n]}\frac{1}{k!}\left(\frac{e^{\frac{1}{12}}\varepsilon nt}{\sqrt{2}r}\right)^k$$

$$\leqslant\sqrt{2}\{5+4M_\theta e^{\frac{\varepsilon}{r}\theta t_0 n}\}e^{t(1-\frac{n}{r})\varepsilon+tn\frac{e^{\frac{1}{12}}\varepsilon}{\sqrt{2}r}}.$$

取 $t=t_0$，注意到 θ 的取法，立知 $J_{1n}\leqslant A_3\rho_4^n$。由此式结合 (11)(12) 知引理获证。下面来

证定理本身。

众所周知，若随机变量X有有限方差，则其中位数 $m(X)$ 满足 $E(X)-\sqrt{2\mathrm{Var}(X)}\leqslant m(X)\leqslant E(X)+\sqrt{2\mathrm{Var}(X)}$. 我们有

$$\left.\begin{aligned} E\Big(\sum_{j=1}^{n} a_{nkj}\mu_j\Big) &= \sum_{j=1}^{n} a_{nkj}E\mu_j = 0 \\ \mathrm{Var}\Big(\sum_{j=1}^{n} a_{nkj}\mu_j\Big) &= \sum_{j=1}^{n} a_{nkj}^2 \mathrm{Var}(\mu_j) = \sigma^2 \end{aligned}\right\} \text{对一切} n、k,$$

故有 $\left|m\Big(\sum_{j=1}^{n} a_{nkj}\mu_j\Big)\right|\leqslant\sqrt{2}\sigma$. 记 μ_j 的对称化随机变量为 $\tilde{\mu}_j$，由引理及弱对称不等式知：

$$P\left\{\left|\sum_{j=1}^{n} a_{nkj}\mu_j - m\Big(\sum_{j=1}^{n} a_{nkj}\mu_j\Big)\right|\geqslant\sqrt{\sigma^2+\frac{n-r}{r}\varepsilon}\right\}\leqslant 2P\left\{\left|\sum_{j=1}^{n} a_{nkj}\tilde{\mu}_j\right|\geqslant\sqrt{\sigma^2+\frac{n-r}{r}\varepsilon}\right\}\leqslant A\rho^n.$$

注意到 $m\Big(\sum_{j=1}^{n} a_{nkj}\mu_j\Big)\leqslant\sqrt{2}\sigma$, 即知 $P\left\{\left|\sum_{j=1}^{n} a_{nkj}\mu_j\right|\geqslant\sqrt{\sigma^2+\frac{n-r}{r}\varepsilon}\right\}\leqslant A\rho^n$.　　(16)

我们有 $\hat{\sigma}_n^2-\sigma^2=\frac{1}{n-r}\left\{\sum_{j=1}^{n}(\mu_j^2-\sigma^2)-\sum_{k=1}^{r}\left[\Big(\sum_{j=1}^{n} a_{nkj}\mu_j\Big)^2-\sigma^2\right]\right\}$. 由(16)式可知

$$\begin{aligned} &P\left\{\frac{1}{n-r}\left|\sum_{k=1}^{r}\left[\Big(\sum_{j=1}^{n} a_{nkj}\mu_j\Big)^2-\sigma^2\right]\right|\geqslant\varepsilon\right\} \\ &\leqslant\sum_{k=1}^{r} P\left\{\left|\Big(\sum_{j=1}^{n} a_{nkj}\mu_j\Big)^2-\sigma^2\right|\geqslant\frac{n-r}{r}\varepsilon\right\}\leqslant A\rho^n. \end{aligned}\tag{17}$$

取 $0<\eta<\varepsilon$, $0<t<t_\eta$, 由条件（4）知

$$P\left\{\frac{1}{n-r}\sum_{j=1}^{n}(\mu_j^2-\sigma^2)\geqslant\varepsilon\right\}\leqslant e^{-(n-r)\varepsilon t}Ee^{t\sum_{j=1}^{n}(\mu_j^2-\sigma^2)}\leqslant M_\eta e^{r\varepsilon t}e^{t(\eta-\varepsilon)n}\triangleq c_1\tilde{\rho}_1^n. \tag{18}$$

同理可知

$$P\left\{\frac{1}{n-r}\sum_{j=1}^{n}(\mu_j^2-\sigma^2)\leqslant-\varepsilon\right\}\leqslant c_2\tilde{\rho}_2^n. \tag{19}$$

综合（17）—（19）知定理获证。

感谢陈希孺教授对作者的指导。

参　考　文　献

[1] Baum, L. E., Katz, M., and Read, R. R., Trans. Amer. Math. Soc., 102 (1962) №2, 187-199.

Exponential Convergence Rates of Error-variance Estimates in Linear Models

Sn Chun

Abstract

Suppose given a linear model $y_j=x_j'\beta+\mu_j$, $j=1,2,\cdots$, The random errors all have a mean zero and unknown variance σ^2, $0<\sigma^2<\infty$. Let $\hat{\sigma}_n^2$ be the estimate of σ^2 based on the residual sum of squares and calculated from (x_j, y_j), $j=1,\cdots,n$. In this paper we show that if $\mu_1,\mu_2,\cdots$, are independent but not necessarily identically distributed, and some further conditions on $\{\mu_j\}$ and $(x_1|\cdots|x_n)$ are satisfied, then for any $\varepsilon>0$ there exist constant ρ_ε, $0<\rho_\varepsilon<1$, Such that

$$P(|\hat{\sigma}_n^2-\sigma^2|\geqslant\varepsilon)=O(\rho_\varepsilon^n).$$