第6卷 第2期
1989年6月

控制理论与应用
CONTROL THEORY AND APPLICATIONS

Vol.6, No.2
June, 1989

线性系统初值最小二乘估计的收敛性问题*

程兆林
（山东大学数学系，济南）

赵克友
（青岛大学数学系，青岛）

郭 雷
（中国科学院系统科学研究所，北京）

摘 要

本文是文献〔1〕的发展，讨论带随机量测噪声线性离散时间系统的初始状态最小二乘估计的均方收敛性与几乎处处收敛性，给出估计误差均方收敛于零或几乎处处收敛于零的充要条件，并给出收敛速率估计.

一、引 言

当输入确定时，带量测噪声的线性系统的状态估计问题实质上是一个初值估计问题. 注意到系统的零初值响应不含有关状态初值的任何信息，因此，初值估计问题只与系统的零输入响应有关. 文献〔1〕讨论了带量测噪声的零输入线性系统初值的马尔柯夫估计的均方收敛性. 本文则进一步讨论初值的最小二乘估计的均方收敛性及几乎处处收敛性.

设零输入线性定常系统

$$\Sigma:\ \begin{cases} x_{k+1}=Ax_k \\ y_k=Cx_k+v_k \end{cases} \tag{1}$$

其中 $A\in R^{n\times n}$，$C\in R^{m\times n}$，$x_k\in R^n$，$y_k\in R^m$，$v_k\in R^m$。v_k 为量测噪声，其统计假设为

（A.1） $v_k=[v_{k1}\ v_{k2}\ \cdots\ v_{km}]^\tau$，$v_{kj}$，$k=0,1,\cdots$，$j=1,2,\cdots,m$，相互独立，满足：

$$Ev_{kj}=0,\ k=0,1,\cdots,\ j=1,2,\cdots,m; \tag{2}$$

$$0<\sigma_1=\inf_{k,j}Ev_{kj}^2\leqslant\sup_{k,j}Ev_{kj}^2=\sigma_2<\infty,\ k=0,1,\cdots,\ j=1,2,\cdots,m \tag{3}$$

（A.2）存在实数 $\mu\geqslant1$，使对任何 $\varepsilon>0$ 及 $j=1,2,\cdots,m$，有

$$\sum_{k=1}^{\infty}P(|v_{kj}|\geqslant k^{\frac{1}{2\mu}}\varepsilon)<\infty \tag{4}$$

*国家自然科学基金资助的项目.
本文于1987年5月13日收到. 1989年2月1日收到修改稿.

熟知，若系统（1）按 Kalman 意义完全能观测，则对系统（1）的初值 x_0 的加权最小二乘估计为[2],[3]

$$\hat{x}_{LS}(N)=[H^{\tau}(N)W(N)H(N)]^{-1}H^{\tau}(N)W(N)Y(N) \tag{5}$$

估计误差为

$$\hat{x}_{LS}(N)-x_0=[H^{\tau}(N)W(N)H(N)]^{-1}H^{\tau}(N)W(N)V(N) \tag{6}$$

式中，

$$H(N)=[C^{\tau}(CA)^{\tau}\cdots(CA^{N-1})^{\tau}]^{\tau} \tag{7}$$

$$Y(N)=[y_0^{\tau}\ y_1^{\tau}\cdots y_{N-1}^{\tau}]^{\tau} \tag{8}$$

$$V(N)=[v_0^{\tau}\ v_1^{\tau}\cdots v_{N-1}^{\tau}]^{\tau} \tag{9}$$

权矩阵 $W(N)$ 对称正定，为讨论方便起见，本文还规定 $W(N)$ 为对角阵，且一致正定，一致有界，即存在二正数 α，β，使对一切自然数 N 有

$$\alpha I_{mN}\leqslant W(N)\leqslant\beta I_{mN} \tag{10}$$

本文讨论：1°估计误差(6)均方收敛于零或几乎处处收敛于零的充要条件，2°给出收敛速率估计。

二、主 要 结 果

引理 1[1] 设 G 为 $n\times n$ 复方阵，令

$$S(N)=\sum_{k=0}^{N-1}(\overline{G^k})^{\tau}G^k \tag{11}$$

式中 $\overline{G}^{\tau}$ 为 G 的共轭转置矩阵，则如下结论成立：

1° $S^{-1}(N)\to 0(N\to\infty)$ 的充要条件为 G 的任一特征根落在单位圆外或单位圆周，即 $|\lambda_j(G)|\geqslant 1$，$j=1,2,\cdots,n$。

2° 若1°中充要条件被满足，则

$$\mathrm{trace}S^{-1}(N)=O\left(\frac{1}{N}\right),\quad N\to\infty \tag{12}$$

引理 2 设对称正定阵 W 满足

$$\alpha I\leqslant W\leqslant\beta I \tag{13}$$

则

$$\alpha W\leqslant W^2\leqslant\beta W \tag{14}$$

式中 α，β 为二正数。

引理 3[4] 设随机变量 e_k，$k=1,2,\cdots$，相互独立，满足

1° $Ee_k=0$，　$\sup Ee_k^2<\infty$； (15)

2° 存在正数 μ，使对任何 $\varepsilon>0$ 有

$$\sum_{k=1}^{\infty}P(|e_k|\geqslant k^{\frac{1}{2\mu}}\varepsilon)<\infty \tag{16}$$

并设常数组 $\{a_{N,k},\ N=1,2,\cdots,\ k=1,2,\cdots,N\}$ 满足

$$\sum_{k=1}^{N} a_{N,k}^{2} \leqslant 1 \tag{17}$$

则

$$\lim_{N\to\infty} N^{-\frac{1}{2\mu}} \sum_{k=1}^{N} a_{N,k} e_k = 0, \quad a.s. \tag{18}$$

定理 设系统（1）按Kalman意义完全能观测，则如下结论成立：

1° 若噪声序列 $\{v_k\}$ 满足统计假设$(A.1)$，则估计误差（6）均方收敛于零的充要条件为 A 的任一特征根落在单位圆外或单位圆周，即 $|\lambda_j(A)|\geqslant 1$，$j=1,2,\cdots,n$；收敛速率为

$$\|E[\hat{x}_{LS}(N)-x_0][\hat{x}_{LS}(N)-x_0]^{\tau}\| = O\left(\frac{1}{N}\right), \quad N\to\infty \tag{19}$$

并且，这一收敛速率估计为精确估计，即（19）式右端不能再作任何改进。

2° 若噪声序列 $\{v_k\}$ 满足统计假设$(A.1)$及$(A.2)$，则估计误差（6）几乎处处收敛于零的充要条件为 A 的任一特征根落在单位圆外或单位圆周，即 $|\lambda_j(A)|\geqslant 1$，$j=1,2,\cdots,n$；收敛速率为

$$\|\hat{x}_{LS}(N)-x_0\| = o\left(N^{-\frac{1}{2}\left(1-\frac{1}{\mu}\right)}\right), \quad a.s. \quad N\to\infty \tag{20}$$

并且，这一收敛速率不能改进为$O(N^{-\frac{1}{2}})$。

证 证1°。熟知，估计误差（6）的均方差为

$$P(N)=[H^{\tau}(N)W(N)H(N)]^{-1}H^{\tau}(N)W(N)R(N)W(N)H(N)[H^{\tau}(N)\cdot W(N)H(N)]^{-1} \tag{21}$$

式中，$R(N)=EV(N)V^{\tau}(N)$。由（$A.1$），$R(N)$满足

$$\sigma_1 I_{mN} \leqslant R(N) \leqslant \sigma_2 I_{mN} \tag{22}$$

并且，由（21）—（22），（10）及引理2得

$$\sigma_1 \frac{\alpha}{\beta}[H^{\tau}(N)H(N)]^{-1} \leqslant P(N) \leqslant \sigma_2 \frac{\beta}{\alpha}[H^{\tau}(N)H(N)]^{-1} \tag{23}$$

记 $S_0=H^{\tau}(n)H(n)$，由系统（1）按Kalman意义完全能观测，知 R_0 对称正定。记 $p=[N/n]$，$[\cdot]$表整数部分，由（7）得

$$[H^{\tau}((p+1)n)H((p+1)n)]^{-1} \leqslant [H^{\tau}(N)H(N)]^{-1} \leqslant [H^{\tau}(pn)H(pn)]^{-1} \tag{24}$$

记

$$S_1(p)=\sum_{k=0}^{p-1}(A^{kn})^{\tau}A^{kn} \tag{25}$$

由

$$H^{\tau}(pn)H(pn)=\sum_{k=0}^{p-1}(A^{kn})^{\tau}S_0A^{kn} \tag{26}$$

推得
$$\frac{1}{\lambda_2}S_1^{-1}(p)\leqslant[H^\tau(pn)\,H(pn)]^{-1}\leqslant\frac{1}{\lambda_1}S_1^{-1}(p) \tag{27}$$

式中 λ_1，λ_2 分别为 S_0 的最小，最大特征根，并且 λ_1，$\lambda_2>0$。联合（23）—（24），（27），对 $S_1(p)$应用引理1，并注意到 p 与 N 同阶，即得所欲证。

证2°。注意到几乎处处收敛必导致依概率收敛，故定理的必要性部分的证明可仿文献[5]之定理5.6.1，从略。下证充分性。令

$$J_1(N)=[H^\tau(N)W(N)H(N)]^{-\frac{1}{2}} \tag{28}$$
$$J_2(N)=Q(N)V(N) \tag{29}$$
$$Q(N)=[H^\tau(N)W(N)H(N)]^{-\frac{1}{2}}H^\tau(N)W(N) \tag{30}$$

（i）估计$\|J_1(N)\|$。仍记 $p=[N/n]$，联合（10），（24）、（27）得

$$J_1^2(N)\leqslant\frac{1}{\alpha\lambda_1}S_1^{-1}(p) \tag{31}$$

对 $S_1(p)$应用引理1，并注意到 p 与 N 同阶，推得

$$\|J_1(N)\|=(\mathrm{trace}J_1^2(N))^{\frac{1}{2}}\leqslant\left(\frac{1}{\alpha\lambda_1}\mathrm{trace}S_1^{-1}(p)\right)^{\frac{1}{2}}=O(N^{-\frac{1}{2}}),N\to\infty \tag{32}$$

（ii）估计$\|J_2(N)\|$。注意到 $Q^\tau(N)Q(N)$是秩为 n 的对称非负定阵，故存在正交阵 $T(N)$，使

$$Q^\tau(N)Q(N)=T^\tau(N)\mathrm{diag}(\lambda_1^{(N)},\lambda_2^{(N)},\ \cdots,\ \lambda_n^{(N)},\ 0,\ \cdots,\ 0)T(N) \tag{33}$$

再注意

$$Q(N)Q^\tau(N)\leqslant\beta I_n \tag{34}$$

故
$$\begin{aligned}O<\lambda_j^{(N)}&\leqslant\mathrm{trace}Q^\tau(N)Q(N)=\mathrm{trace}Q(N)Q^\tau(N)\\&\leqslant n\beta,\ N=1,\ 2,\ \cdots,\ j=1,\ 2,\ \cdots,\ n\end{aligned} \tag{35}$$

$T(N)$为 $mN\times mN$ 阵，记 $T(N)$的$(i,\ j)$元为 $t_{mN,i,j}$。改记（9）式中之 $V(N)$为

$$V(N)=[e_1\quad e_2\cdots e_{mN}]^\tau \tag{36}$$

联合（29），（33），（35），推得

$$\begin{aligned}\|J_2(N)\|^2=V^\tau(N)Q^\tau(N)Q(N)V(N)&=\sum_{j=1}^{n}\lambda_j^{(N)}\left(\sum_{k=1}^{mN}t_{mN,j,k}e_k\right)^2\\&\leqslant n^2\beta\left(\sum_{j=1}^{mN}t_{mN,j,k}e_k\right)^2\end{aligned} \tag{37}$$

$T(N)$还为正交阵，故

$$\sum_{k=1}^{mN}t_{mN,j,k}^2=1,\ N=1,\ 2,\ \cdots,\ j=1,\ 2,\ \cdots,\ n \tag{38}$$

随机序列 $\{e_k,\ k=1,\ 2,\ \cdots\}$ 满足统计假设（A.1），(A.2)数列 $\{t_{mN,j,k},\ N=1,\ 2,$

$\cdots$, $k=1, 2, \cdots, mN$, $j=1, 2, \cdots, n\}$ 满足（38）式，因此，$\{e_k\}$，$\{t_{mN,j,k}\}$ 满足引理3的全部条件，故

$$\sum_{k=1}^{mN} t_{mN,j,k} e_k = o\left(N^{\frac{1}{2\mu}}\right), \quad a.s. \qquad N\to\infty \tag{39}$$

联系到（37）式，得

$$\|J_2(N)\| = o\left(N^{\frac{1}{2\mu}}\right), \quad a.s. \qquad N\to\infty \tag{40}$$

于是，联合（6），（28）—（30），（32），（40）即得（20）。

（19）式的右端不能再作任何改进。这是因为很容易找到系统，其初值的最小二乘估计的均方误差达到$\frac{1}{N}$。例如系统

$$\Sigma_0: \begin{cases} x_{k+1} = x_k \\ y_k = x_k + v_k \end{cases} \tag{41}$$

式中 v_k, $k=0, 1, \cdots$，相互独立，$v_k \in N(0, 1)$。当取权矩阵 $W(N)=I_N$，$N=1, 2, \cdots$，作初值最小二乘估计，即有

$$E(\hat{x}_{LS}(N) - x_0)^2 = E\left(\sum_{k=0}^{N-1} \frac{v_k}{N}\right)^2 = \frac{1}{N} \tag{42}$$

（20）式的右端亦不能改进为$O(N^{-\frac{1}{2}})$。如若不然，仍以系统Σ_0为例，取$W(N)=I_N$，$N=1, 2, \cdots$，作初值最小二乘估计，其估计误差即有

$$|\hat{x}_{LS}(N) - x_0| = \left|\sum_{k=0}^{N-1} \frac{v_k}{N}\right| = O(N^{-\frac{1}{2}}), \quad a.s.$$

而这与重对数律

$$\overline{\lim_{N\to\infty}} \frac{\left|\sum_{k=0}^{N-1} v_k\right|}{\sqrt{N \log\log N}} = \sqrt{2}, \quad a.s.$$

矛盾，故（20）式右端不能改进为$O(N^{-\frac{1}{2}})$。证毕。

评述。熟知，正态随机序列 v_k, $k=0, 1, \cdots$，若满足统计假设（A.1）则必满足（A.2）。因此，上述定理可处理一大类量测噪声。至于统计假设（A.2）是否可以减弱抑或删去，作者拟在另一篇文章中进行讨论。

参考文献

〔1〕陈兆宽、程兆林，带随机量测噪声的线性系统的初始状态马尔柯夫估计的均方收敛性，系统科学与数学，5:1,（1985），63—72。

〔2〕 中国科学院数学研究所概率组编，离散时间系统滤波的数学方法，国防工业出版社，北京，（1975），57—60.
〔3〕 关肇直、陈翰馥，线性控制系统的能控性和能观测性，科学出版社，北京，（1975），30—39.
〔4〕 陈希孺、陈桂景、吴启光、赵林诚，线性模型参数的估计理论，科学出版社，北京，（1985），114—119.
〔6〕 陈希孺，数理统计引论，科学出版社，北京，（1981），518—522.

The Convergence Problem of the Least Squares Estimation of the Initial State of Linear Systems

Cheng Zhaolin
(Department of Mathematics, Shandong University, Jinan)
Zhao keyou
(Department of Mathematics, Qingdao University, Qingdao)
Guo Lei
(Institute of Systems Science, Academia Sinica, Beijing)

Abstract

This paper deals with the least squares estimation of the initial state of linear systems with zero input and measurement noise, and provides some sufficient and necessary conditions in order to guarantee the least squares estimation error converging to zero (almost surely or in the mean square sense). The convergence rates are also obtained. And an example is given to show that the convergence rate in the mean square sense can not be improved any more.