第 14 卷第 2 期
1997 年 4 月

控制理论与应用
CONTROL THEORY AND APPLICATIONS

Vol. 14, No. 2
Apr., 1997

多变量系统的辅助模型辨识方法的收敛性分析*

丁 锋
(清华大学自动化系・北京,100084)

摘要: 丁锋,谢新民[1]假设协方差阵 $P_0^{-1}(t)=\sum_{s=1}^{t}\varphi_0(s)\varphi_0^{\mathrm{T}}(s)$ 的最大特征值与最小特征值之比(即条件数)有界,和系统噪声 $\{w(t)\}$ 方差有界,证明了辅助模型辨识算法参数估计的一致收敛性. 本文将分别放松这两个条件,即 i) 条件数无界, ii) 系统噪声为非平稳噪声,且方差无界,探讨了辅助模型算法的收敛性. 分析表明在这种较弱的假设下,辅助模型算法保持很强的鲁棒性,参数估计是一致收敛的.

关键词: 多变量系统; 参数估计; 条件数; 辅助模型方法; 鞅收敛定理

1 引 言

对于输出误差 (OE) 模型,目前还没有一种有效算法估计其参数. 常规最小二乘法不可能给出 OE 模型参数的一致估计;而用递推增广最小二乘法 (RELS) 估计 OE 模型的参数,不仅增加了系统待辨识参数数目,而且加大了算法的计算量.

先前的多变量系统辨识算法,均把多变量系统按照输出的数目 (m) 分解为 m 个子系统进行辨识,这可称为子系统辨识算法 (SSIA)[2~5];SSIA 算法增加了许多不必要的计算,使得计算量大大增加.

E1-Sherief[5]采用 RELS 分别估计 OE 模型描述的多变量系统的每个子系统的参数,该方法不仅使系统待估计参数的数目增多,计算量增大,而且还难以保证参数估计的收敛性.

Solo[6]在条件数有界和噪声方差有界假设下,用一个错误引理证明了 RELS 或 AML 算法的收敛性[7],丁锋和谢新民[8]在同一假设下证明了多变量系统的 RELS 的收敛性. Chen, Guo[9]在噪声方差有界假设下分析了多变量系统的 RELS 的收敛速率.

辅助模型辨识算法是一种估计系统参数的有效方法,它是借助于辅助模型的输出代替系统的真实输出的一种辨识方法. 丁锋和谢新民分别用辅助模型方法估计多变量系统传递函数阵子子模型[10]和 OE 模型[1]的参数. 但是以上两篇文章都没有讨论参数估计收敛的速率,且对噪声和条件数有较严格的限制. 本文将分别放松这两个条件,即在噪声方差无界或条件数无界下,来探讨辅助模型方法参数估计的收敛速率. 其分析方法同文[9]一样,是借助于鞅 (Martingale) 理论.

2 辅助模型辨识算法

状态空间模型可以通过代数运算化为如下输出误差模型 (OE)[1,5]

$$y(t)=P_1^{-1}(z)Q(z)u(t)+w(t) \tag{1a}$$

或

$$P_1(z)y(t)=Q(z)u(t)+P_1(z)w(t), \tag{1b}$$

$$P_1(z)=[P_{ij}(z)]\in\mathbb{R}^{m\times m},\quad Q(z)=[q_{ij}(z)]\in\mathbb{R}^{m\times r},$$

---

* 国家自然科学基金资助项目.
本文于 1995 年 3 月 12 日收到. 1996 年 9 月 24 日收到修改稿.

$$P_{ij}(z) = \delta_{ij} + a_{ij}(1)z^{-1} + a_{ij}(2)z^{-2} + \cdots + a_{ij}(v_j)z^{-v_j},$$
$$q_{ij}(z) = b_{ij}(1)z^{-1} + b_{ij}(2)z^{-2} + \cdots + b_{ij}(v)z^{-v},$$
$$v = \max(v_1, v_2, \cdots, v_m), \quad \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

其中 $u(t) \in \mathbb{R}^r$ 和 $y(t) \in \mathbb{R}^m$ 分别为系统输入和含噪输出，$w(t) \in \mathbb{R}^m$ 为零均值不相关随机噪声. z^{-1} 为单位后移算子，$[a_{ij}(s), b_{ij}(s)]$ 为系统参数. $v_i, i = 1, 2, \cdots, m$ 为系统的结构指数.

系统(1)又可化为如下形式

$$y(t) = A^{-1}(z)B(z)u(t) + w(t). \tag{2}$$

其中

$$A(z) = I + A_1 z^{-1} + A_2 z^{-2} + \cdots + A_v z^{-v},$$
$$B(z) = B_1 z^{-1} + B_2 z^{-2} + \cdots + B_v z^{-v},$$

将(2)式改写为

$$A(z)x(t) = B(z)u(t), \tag{3a}$$
$$y(t) = x(t) + w(t). \tag{3b}$$

其中 $x(t) \in \mathbb{R}^m$ 为系统的真实输出(不可测)，$y(t)$ 是 $x(t)$ 的含噪量测(可测).

如果 $x(t)$ 可测，可以很方便地利用 $\{u(t), x(t)\}$ 来估计(3)式的参数，由于 $x(t)$ 不可测，但是可以通过辅助模型

$$\hat{x}(t) = \theta_a^{\mathrm{T}}(t)\varphi_a(t) \tag{4}$$

算出它的近似值. 其中 $\varphi_a(t)$ 和 $\theta_a(t)$ 分别为辅助模型在 t 时刻的信息向量和参数矩阵. $\varphi_a(t)$ 和 $\theta_a(t)$ 的选择方式决定了算法的具体形式和收敛性.

置

$$\theta^{\mathrm{T}} = [A_1, A_2, \cdots, A_v, B_1, B_2, \cdots, B_v] \in \mathbb{R}^{m \times n},$$
$$\varphi_0^{\mathrm{T}}(t) = [-x^{\mathrm{T}}(t-1), \cdots, -x^{\mathrm{T}}(t-v)\ u^{\mathrm{T}}(t-1), \cdots, u^{\mathrm{T}}(t-v)] \in \mathbb{R}^{1 \times n},$$
$$n = (m + r)v. \tag{5}$$

则(3)式可写为

$$x(t) = \theta^{\mathrm{T}}\varphi_0(t), \tag{6a}$$
$$y(t) = \theta^{\mathrm{T}}\varphi_0(t) + w(t). \tag{6b}$$

这样，估计系统(6)参数 θ 的递推辅助模型最小二乘法 (RAMLS) 可以表示为

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t)[y^{\mathrm{T}}(t) - \varphi^{\mathrm{T}}(t)\hat{\theta}(t-1)], \tag{7a}$$
$$L(t) = P(t)\varphi(t) = \frac{P(t-1)\varphi(t)}{1 + \varphi^{\mathrm{T}}(t)P(t-1)\varphi(t)}, \tag{7b}$$
$$P^{-1}(t) = P^{-1}(t-1) + \varphi(t)\varphi^{\mathrm{T}}(t) \tag{7c}$$

或

$$P(t) = [I - L(t)\varphi^{\mathrm{T}}(t)]P(t-1),$$
$$\varphi^{\mathrm{T}}(t) = [-\hat{x}^{\mathrm{T}}(t-1), \cdots, -\hat{x}^{\mathrm{T}}(t-v)u^{\mathrm{T}}(t-1), \cdots, u^{\mathrm{T}}(t-v)],$$
$$\hat{x}(t) = \hat{\theta}^{\mathrm{T}}(t)\varphi(t), \tag{7d}$$

$P(0) = P_0 = aI$, $a > 0$, $\theta(0)$ 是任意有界随机矩阵，I 为单位阵. $\hat{\theta}(t)$ 为 θ 的估计.

很明显

$$P^{-1}(t) = \sum_{s=1}^{t}\varphi(s)\varphi^{\mathrm{T}}(s) + \frac{1}{a}I, \quad P(t) = [\sum_{s=1}^{t}\varphi(s)\varphi^{\mathrm{T}}(s) + \frac{1}{a}I]^{-1}, \tag{8}$$

定义 $P_0(t)=[\sum_{s=1}^{t}\varphi_0(s)\varphi_0^{\mathrm{T}}(s)+\frac{1}{a}I]^{-1}$，$\lambda_{\min}[X]$，$\lambda_{\max}[X]$，$\mathrm{tr}[X]$ 分别表示 X 的最小特征值，最大特征值和迹. 并且令 $\|X\|^2=\mathrm{tr}[XX^{\mathrm{T}}]$，有

$$r(t)=\mathrm{tr}[P^{-1}(t)]=\frac{n}{a}+\sum_{s=1}^{t}\|\varphi(s)\|^2=r(t-1)+\|\varphi(t)\|^2, \tag{9}$$

$$r_0(t)=\mathrm{tr}[P_0^{-1}(t)]=\frac{n}{a}+\sum_{s=1}^{t}\|\varphi_0(s)\|^2=r_0(t-1)+\|\varphi_0(t)\|^2. \tag{10}$$

3　多变量系统辅助模型辨识方法的收敛性分析

设 $\{w(t)\}$ 是定义在概率空间 (Ω,F,P) 上的鞅 (Martingale) 差序列，且适应于递增 σ-代数序列 $(F_t,t\in N)$，其中 F_t是由直到 t 时刻的观测生成的 σ-代数，F_0 包含所有初始条件信息.

关于 RAMLS 算法(7)的收敛性有定理.

定理 1　对于系统(6)，作以下假设

A1)　$\mathrm{E}[w(t)|F_{t-1}]=0$，　a.s.；

A2)　$\mathrm{E}[\|w(t)\|^2|F_{t-1}]=\sigma_w^2(t)\leqslant\sigma_w^2<\infty$，　a.s.；

A3)　$H(z)=A^{-1}(z)-\frac{1}{2}I$ 严格正实，

那么辅助模型算法(7)给出的参数估计误差 $\|\hat{\theta}(t)-\theta\|^2$ 满足

i) $\|\hat{\theta}(t)-\theta\|^2=O\left(\frac{[\ln r_0(t)]^c}{\lambda_{\min}[P_0^{-1}(t)]}\right)$，　a.s.，　对任意 $c>1$；

ii) $\|\hat{\theta}(t)-\theta\|^2=O\left(\frac{\ln r_0(t)[\ln\ln r_0(t)]^c}{\lambda_{\min}[P_0^{-1}(t)]}\right)$，　a.s.，　对任意 $c>1$；

iii) ….

当持续激励条件(条件数为无穷大)

$$\alpha_1 I\leqslant\frac{1}{t}\sum_{s=1}^{t}\varphi_0(s)\varphi_0^{\mathrm{T}}(s)\leqslant\beta_1 t^{c_0}I,\quad \alpha_1>0,\quad \beta_1>0,\quad \text{a.s.},\quad \text{对任意 } c>1$$

成立时，定理 1 的结论可表示为

i) $\|\hat{\theta}(t)-\theta\|^2=O\left(\frac{[\ln t]^c}{t}\right)\to 0$，　a.s.，　对任意 $c>1$；

ii) $\|\hat{\theta}(t)-\theta\|^2=O\left(\frac{\ln t[\ln\ln t]^c}{t}\right)\to 0$，　a.s.，　对任意 $c>1$；

iii) ….

本文中 $O(*)$，$o(*)$ 分别定义为 i) $f(t)=O(g(t))$ 表示当 t 充分大时，$|\frac{f(t)}{g(t)}|\leqslant M<\infty$；ii) $f(t)=o(g(t))$ 表示 $\lim_{t\to\infty}\frac{f(t)}{g(t)}=0$.

定理 1 说明在噪声 $\{w(t)\}$ 的方差有界情况下，辅助模型辨识算法给出参数估计的收敛速度为 $P_0^{-1}(t)$ 的最大特征值(或迹)的对数与最小特征值之比，在持续激励条件下(条件数无界)，参数估计一致收敛于零. 下面的定理在考虑噪声 $\{w(t)\}$ 方差无界条件下，如何获得一致参数估计.

定理 2　对于系统(6)，作以下假设

A1)　$\mathrm{E}[w(t)|F_{t-1}]=0$；　a.s.；

A2)　$\mathrm{E}[\|w(t)\|^2|F_{t-1}]\leqslant k_1 r^{\varepsilon}(t)$，　a.s.，　$0\leqslant k_1<\infty$，　$0\leqslant\varepsilon<1$；

A3)　$H(z) = A^{-1}(z) - \frac{1}{2}I$ 严格证实，

那么辅助模型辨识算法(7)给出的参数估计满足

$$\| \hat{\theta}(t) - \theta \|^2 = o\left(\frac{\lambda_{\max}[P_0^{-1}(t)]}{\lambda_{\min}[P_0^{-1}(t)]}\right), \quad \text{a.s.}$$

当持续激励条件(条件数有界)

$$\alpha_2 I \leqslant \frac{1}{t}\sum_{s=1}^{t}\varphi_0(s)\varphi_0^{\mathrm{T}}(s) \leqslant \beta_2 I, \quad 0 < \alpha_2 \leqslant \beta_2 < \infty, \quad \text{a.s.}$$

成立时，则有

$$\lim_{t\to\infty} \| \theta(t) - \theta \|^2 = o\left(\frac{\beta_2}{\alpha_2}\right) = o(1) = 0, \quad \text{a.s.}$$

定理 1 的证明　定义参数估计误差矩阵

$$\tilde{\theta}(t) = \theta(t) - \theta \tag{11}$$

和非负定函数

$$T(t) = \mathrm{tr}[\tilde{\theta}^{\mathrm{T}}(t)P^{-1}(t)\tilde{\theta}(t)], \tag{12}$$

很明显

$$\| \tilde{\theta}(t) \|^2 \leqslant \frac{1}{\lambda_{\min}[P^{-1}(t)]}\mathrm{tr}[\tilde{\theta}^{\mathrm{T}}(t)P^{-1}(t)\tilde{\theta}(t)] = \frac{1}{\lambda_{\min}[P^{-1}(t)]}T(t). \tag{13}$$

证明的思路是先拟合一个鞅过程，然后利用鞅收敛定理得证. 这可以化分为四个命题(引理).

引理 1　在定理 1 的假设条件下，下式成立

$$\mathrm{E}[T(t) + S(t) | F_{t-1}] \leqslant T(t-1) + S(t-1) + 2\varphi^{\mathrm{T}}(t)P(t)\varphi(t)\sigma_w^2.$$

其中

$$S(t) = 2\sum_{s=1}^{t}\tilde{u}^{\mathrm{T}}(s)\tilde{y}(s) + c_1 > 0, \quad c_1 < \infty, \tag{14}$$

$$\tilde{y}(t) = \frac{1}{2}\tilde{\theta}^{\mathrm{T}}(t)\varphi(t) + [y(t) - \hat{\theta}^{\mathrm{T}}(t)\varphi(t) - w(t)], \tag{15}$$

$$\tilde{u}(t) = -\tilde{\theta}^{\mathrm{T}}(t)\varphi(t). \tag{16}$$

引理 2　设 $D \in \mathbb{R}^{m\times n}$，$E \in \mathbb{R}^{n\times m}$ 是两个矩阵，则下列等式成立

$\det[I_m + DE] = \det[I_n + ED]$，$\det[X] = |X|$ 为 X 的行列式.

引理 3　试证下列不等式

i) $\sum_{s=1}^{t}\varphi^{\mathrm{T}}(s)P(s)\varphi(s) \leqslant \ln|P^{-1}(t)| + n\ln a$；　a.s.

ii) $\sum_{s=1}^{\infty}\frac{\varphi^{\mathrm{T}}(s)P(s)\varphi(s)}{[\ln|P^{-1}(s)|]^c} < \infty$，　a.s.　对任意 $c > 1$

iii) $\sum_{s=1}^{\infty}\frac{\varphi^{\mathrm{T}}(s)P(s)\varphi(s)}{\ln|P^{-1}(s)|[\ln\ln|P^{-1}(s)|]^c} < \infty$，　a.s.　对任意 $c > 1$；

iv) …

引理 4　$\ln|P^{-1}(t)| = O(\ln r(t))$　或　$\ln|P_0^{-1}(t)| = O(\ln r_0(t))$.

以上四个引理的证明见附录.

现在回到证明定理 1.

i) 令

$$V(t)=\frac{T(t)+S(t)}{[\ln P^{-1}(t)|]^{c}},\quad c>1,\tag{17}$$

利用引理 1 得到

$$\begin{aligned}\mathrm{E}[V(t)|F_{t-1}]&\leqslant\frac{T(t-1)+S(t-1)}{[\ln|P^{-1}(t)|]^{c}}+\frac{2\varphi^{\mathrm{T}}(t)P(t)\varphi(t)}{[\ln|P^{-1}(t)|]^{c}}\sigma_{w}^{2}\\&\leqslant V(t-1)+\frac{2\varphi^{\mathrm{T}}(t)P(t)\varphi(t)}{[\ln|P^{-1}(t)|]^{c}}\sigma_{w}^{2}\quad[\text{由于}\ln|P^{-1}(t)|\text{ 非降}],\end{aligned}\tag{18}$$

根据引理 3-ii) 知上式右边第 2 项的和 $\sum\limits_{t=1}^{\infty}$ 有界，应用鞅收敛定理（文[11]第 4 章）知 $V(t)$ 几乎肯定收敛于一有界随机变量，即

$$V(t)=\frac{T(t)+S(t)}{[\ln|P^{-1}(t)|]^{c}}\rightarrow C<\infty,\quad \text{a.s.}\tag{19}$$

由 $H(z)$ 是严格正实的假设，不难得到

$$\sum_{i=1}^{t}\|\tilde{u}(i)\|^{2}=O([\ln|P^{-1}(t)|]^{c}).\tag{20}$$

由式(13)式、(19)式和引理 4，可得

$$\|\hat{\theta}(t)-\theta\|^{2}=O\left(\frac{[\ln|P^{-1}(t)|]^{c}}{\lambda_{\min}[P^{-1}(t)]}\right)=O\left(\frac{[\ln r(t)]^{c}}{\lambda_{\min}[P^{-1}(t)]}\right),\quad \text{a.s.}\quad c>1.\tag{21}$$

由于假设 $A(z)$ 稳定，由(20)式和文[11]第 3 章定理 3.2.4，可知

$$\begin{aligned}\sum_{s=1}^{t}\|\eta(s)-w(s)\|^{2}&=\sum_{s=1}^{t}\|x(s)-\hat{x}(s)\|^{2}\quad(\eta(s)\text{ 见附录})\\&\leqslant K_{1}\sum_{s=1}^{t}\|\tilde{u}(s)\|^{2}+K_{2},\quad(0\leqslant K_{1}<\infty,\quad 0\leqslant K_{2}<\infty)\\&=O([\ln|P^{-1}(t)|]^{c})=O([\ln r(t)]^{c}).\end{aligned}\tag{22}$$

下面证明 $r(t)=O(r_{0}(t))$，$\lambda_{\min}[P^{-1}(t)]=O(\lambda_{\min}[P_{0}^{-1}(t)]$.

定义向量差

$$\begin{aligned}\tilde{\varphi}(t)&=\varphi(t)-\varphi_{0}(t)\quad[\text{参见 }\varphi(t)\text{ 和 }\varphi_{0}(t)\text{ 的定义}]\\&=[x^{\mathrm{T}}(t-1)-\hat{x}^{\mathrm{T}}(t-1),\cdots,x^{\mathrm{T}}(t-v)-\hat{x}^{\mathrm{T}}(t-v);0,\cdots,0]^{\mathrm{T}}.\end{aligned}\tag{23}$$

因此

$$\begin{aligned}\sum_{s=1}^{t}\|\tilde{\varphi}(s)\|^{2}&=\sum_{s=1}^{t}\sum_{j=1}^{v}\|x(s-j)-\hat{x}(s-j)\|^{2}\\&=O\Big(\sum_{s=1}^{t}\|x(s)-\hat{x}(s)\|^{2}\Big)=O([\ln r(t)]^{c}),\end{aligned}\tag{24}$$

$$r(t)\leqslant 2r_{0}(t)+2\sum_{s=1}^{t}\|\tilde{\varphi}(s)\|^{2}=2r_{0}(t)+O([\ln r(t)]^{c}),\tag{25}$$

$$r(t)=O(r_{0}(t)).\tag{26}$$

对于任意向量 $\omega\in\mathbb{R}^{n}$，且 $\|\omega\|=1$，有

$$\begin{aligned}\sum_{s=1}^{t}[\omega^{\mathrm{T}}\varphi(s)]^{2}&=\sum_{s=1}^{t}[\omega^{\mathrm{T}}\varphi_{0}(s)+\omega^{\mathrm{T}}\tilde{\varphi}(s)]^{2}\\&\leqslant 2\sum_{s=1}^{t}[\omega^{\mathrm{T}}\varphi_{0}(s)]^{2}+2\sum_{s=1}^{t}\|\tilde{\varphi}(s)\|^{2}\end{aligned}$$

$$=2\sum_{s=1}^{t}[\omega^{\mathrm{T}}\varphi_0(s)]^2+O([\ln r_0(t)]^c),\tag{27}$$

于是

$$\lambda_{\min}[P^{-1}(t)]\leqslant 2\lambda_{\min}[P_0^{-1}(t)]+O(\lambda_{\min}[P^{-1}(t)]),$$
$$\lambda_{\min}[P^{-1}(t)]=O(\lambda_{\min}[P_0^{-1}(t)]).\tag{28}$$

联立(21)式,(26)式和(28)式,可得

$$\|\hat{\theta}(t)-\theta\|^2=O\left(\frac{[\ln r_0(t)]^c}{\lambda_{\min}[P_0^{-1}(t)]}\right),\quad \text{a.s.}\quad c>1.\tag{29}$$

用类似的方法可以证明定理 1 的 ii),iii),….

定理 2 的证明　运用定理 2 的假设,按照引理 1 证明的推导,可得

$$\mathrm{E}[T(t)+S(t)|F_{t-1}]\leqslant T(t-1)+S(t-1)+2k_1\varphi^{\mathrm{T}}(t)P(t)\varphi(t)r^{\varepsilon}(t)\tag{30}$$

其中 $S(t)$ 的定义同引理 1. 令

$$W(t)=\frac{T(t)+S(t)}{r(t)}.\tag{31}$$

利用(30)式,可得

$$\mathrm{E}[W(t)|F_{t-1}]\leqslant W(t-1)-W(t-1)\frac{r(t)-r(t-1)}{r(t)}+2k_1\frac{\varphi^{\mathrm{T}}(t)P(t)\varphi(t)}{[r(t)]^{1-\varepsilon}}.\tag{32}$$

利用附录的 (A12) 式,从引理 4 的证明知 $\dfrac{1}{r(t)}\leqslant\dfrac{1}{|P^{-1}(t)|^{1/n}}$,现在考虑(32)式右边最后一项的和

$$\sum_{t=1}^{\infty}\frac{\varphi^{\mathrm{T}}(t)P(t)\varphi(t)}{[r(t)]^{1-\varepsilon}}\leqslant\sum_{t=1}^{\infty}\frac{|P^{-1}(t)|-|P^{-1}(t-1)|}{|P^{-1}(t)|^{1+\mu}},\quad \left(\mu=\frac{1-\varepsilon}{n}>0\right)$$
$$\leqslant\int_{|P^{-1}(0)|}^{|P^{-1}(\infty)|}\frac{\mathrm{d}x}{x^{1+\mu}}=\frac{1}{\mu}\left[\frac{1}{|P^{-1}(0)|^{\mu}}-\frac{1}{|P^{-1}(\infty)|^{\mu}}\right]<\infty.\tag{33}$$

将鞅收敛定理应用于(32)式,可知 $W(t)$ 渐近地收敛于一有限随机变量,即

$$W(t)=\frac{T(t)+S(t)}{r(t)}\to W<\infty,\quad \text{a.s.}\tag{34}$$

且

$$\sum_{t=1}^{\infty}W(t-1)\frac{r(t)-r(t-1)}{r(t)}<\infty,\quad \text{a.s.}\tag{35}$$

根据 $r(t)$ 的定义和文[11]定理 4.6.2 的证明过程可知

$$\sum_{t=1}^{\infty}\frac{r(t)-r(t-1)}{r(t)}=\sum_{t=1}^{\infty}\frac{\|\varphi(t)\|^2}{r(t)}\to\infty,\ \text{a.s.}$$

故由(34)式和(35)式知 $W=0$, a.s. 而且

$$\frac{T(t)}{r(t)}\to 0,\quad \text{a.s.}\tag{36}$$

参见(12)式,由于 $r(t)\leqslant n\cdot\lambda_{\max}[P^{-1}(t)]$,由(36)式可得

$$\frac{\lambda_{\min}[P^{-1}(t)]\|\tilde{\theta}(t)\|^2}{n\cdot\lambda_{\max}[P^{-1}(t)]}\leqslant\frac{T(t)}{r(t)}\to 0,\quad \text{a.s.}$$

类似于定理 1 中的推导,可知

$$r(t)=O(r_0(t)),\quad \lambda_{\min}[P^{-1}(t)]=O(\lambda_{\min}[P_0^{-1}(t)]),$$
$$\lambda_{\max}[P^{-1}(t)]=O(\lambda_{\max}[P_0^{-1}(t)])=O(r_0(t)),$$

于是

$$\|\tilde{\theta}(t)\|^2 = O\left(\frac{\lambda_{\max}[P_0^{-1}(t)]}{\lambda_{\min}[P_0^{-1}(t)]}\right) = O\left(\frac{r_0(t)}{\lambda_{\min}[P_0^{-1}(t)]}\right).$$

定理 2 证毕.

参考文献

1 丁锋,谢新民.多变量系统的辅助模型辨识算法.清华大学学报(自然科学版),1992,32(4):100—106

2 Gauthier, A. and Landau, I. D.. On the recursive identification of multi-input multi-output systems. Automatica, 1978, 14:609—614

3 Sinha, N. K. and Kwong, Y. H.. Recursive Estimation of the parameters of linear multivariable systems. Automatica, 1979, 15:471—475

4 El-Sherief, H. and Sinha, N. K.. Choices of models for the identification of linear multivariable discrete-time systems. IEE Proc. Pt. D, 1979, 126(12):1326—1330

5 El-Sherief, H.. Parametric identification of a state-space model of multivariable systems using the extended least-squares method. IEEE Trans., SMC, 1981, 11(3):223—227

6 Solo, V.. The convergence of AML. IEEE Trans. Automat. Contr., 1989, AC-24(6):958—962

7 Lai, T. L. and Ying, Z. L.. Recursive identification and adaptive prediction in linear stochastic systems. SIAM J. Control and Optimization, 1991, 29(5):1061—1090

8 丁锋,谢新民.多变量系统递推增广最小二乘法收敛性分析.控制与决策,1992,7(6):443—447

9 Chen, H. F. and Guo, L.. Convergence rate of least-squares identification and adaptive Control for stochastic systems. Int. J. Control, 1986, 44(5):1459—1476

10 丁锋,谢新民.传递函数阵子子模型参数递推估计:辅助模型方法.控制与决策,1991,6(6):447—452

11 冯纯伯,史维.自适应控制.电子工业出版社,1986,第3,4章

12 丁锋.时变参数系统辨识及其应用.清华大学自动化系博士学位论文,1994年附录F

附 录

引理 1 的证明 分别定义新息向量 $e(t)$ 和残量向量 $\eta(t)$ 为

$$e(t) = y(t) - \hat{\theta}^{\mathrm{T}}(t-1)\varphi(t), \tag{A1}$$

$$\eta(t) = y(t) - \hat{\theta}^{\mathrm{T}}(t)\varphi(t) = y(t) - \hat{x}(t). \tag{A2}$$

利用(7a)式,知残差 $\eta(t)$ 和新息 $e(t)$ 有如下关系

$$\eta(t) = [1 - \varphi^{\mathrm{T}}(t)P(t)\varphi(t)]e(t) = \frac{e(t)}{1 + \varphi^{\mathrm{T}}(t)P(t-1)\varphi(t)}, \tag{A3}$$

(7a)式两边同减去 θ,并利用(A1)~(A3)式,可得

$$\begin{aligned}\tilde{\theta}(t) &= \hat{\theta}(t) - \theta = \tilde{\theta}(t-1) + P(t)\varphi(t)e^{\mathrm{T}}(t)\\ &= \tilde{\theta}(t-1) + P(t-1)\varphi(t)\eta^{\mathrm{T}}(t),\end{aligned} \tag{A4}$$

利用(7c)式和(A4)式,可得

$$\begin{aligned}\tilde{\theta}^{\mathrm{T}}(t)P^{-1}(t)\tilde{\theta}(t) &= \tilde{\theta}^{\mathrm{T}}(t-1)P^{-1}(t-1)\tilde{\theta}(t-1) + \tilde{\theta}^{\mathrm{T}}(t)\varphi(t)\varphi^{\mathrm{T}}(t)\tilde{\theta}(t) + \tilde{\theta}^{\mathrm{T}}(t)\varphi(t)\eta^{\mathrm{T}}(t)\\ &\quad + e(t)\varphi^{\mathrm{T}}(t)\tilde{\theta}(t) + e(t)\varphi^{\mathrm{T}}(t)P(t)\varphi(t)\varphi^{\mathrm{T}}(t)\tilde{\theta}(t) - e(t)\varphi^{\mathrm{T}}(t)P(t)\varphi(t)\eta^{\mathrm{T}}(t),\end{aligned} \tag{A5}$$

上式两边取迹 tr[*],并运用 $\mathrm{tr}[AB] = \mathrm{tr}[BA]$, $\mathrm{tr}[A^{\mathrm{T}}] = \mathrm{tr}[A]$,有

$$\begin{aligned}T(t) &= T(t-1) + \|\varphi^{\mathrm{T}}(t)\tilde{\theta}(t)\|^2 + 2\varphi^{\mathrm{T}}(t)\tilde{\theta}(t)\eta(t)\\ &\quad - [1 - \varphi^{\mathrm{T}}(t)P(t)\varphi(t)]\varphi^{\mathrm{T}}(t)P(t)\varphi(t)\|e(t)\|^2\\ &= T(t-1) + 2\varphi^{\mathrm{T}}(t)\tilde{\theta}(t)\left[\frac{1}{2}\tilde{\theta}^{\mathrm{T}}(t)\varphi(t) + [\eta(t) - w(t)]\right]\\ &\quad + 2\varphi^{\mathrm{T}}(t)\tilde{\theta}(t)w(t) - [1 - \varphi^{\mathrm{T}}(t)P(t)\varphi(t)]\varphi^{\mathrm{T}}(t)P(t)\varphi(t)\|e(t)\|^2,\end{aligned} \tag{A6}$$

利用(15),(16),(A2)和(A4)式,注意到 $0 < \varphi^{\mathrm{T}}(t)P(t)\varphi(t) < 1$,有

$$T(t) \leqslant T(t-1) - 2\tilde{u}^{\mathrm{T}}(t)\tilde{y}(t) + 2\varphi^{\mathrm{T}}(t)[\tilde{\theta}(t-1) + P(t)\varphi(t)e^{\mathrm{T}}(t)]w(t)$$
$$= T(t-1) - 2\tilde{u}^{\mathrm{T}}(t)\tilde{y}(t) + 2\varphi^{\mathrm{T}}(t)\tilde{\theta}(t-1)w(t)$$
$$+ 2\varphi^{\mathrm{T}}(t)P(t)\varphi(t)\{[e(t) - w(t)]^{\mathrm{T}}w(t) + \|w(t)\|^2\}, \tag{A7}$$

考虑到 $\varphi^{\mathrm{T}}(t)\tilde{\theta}(t-1), e(t) - w(t), \varphi^{\mathrm{T}}(t)P(t)\varphi(t)$ 与 $w(t)$ 不相关，且是 F_{t-1} 可测的，(A7) 式两边对 F_{t-1} 取条件期望，运用假设 A1) 和 A2) 可得

$$\mathrm{E}[T(t)|F_{t-1}] \leqslant T(t-1) - 2E[\tilde{u}^{\mathrm{T}}(t)\tilde{y}(t)|F_{t-1}] + 2\varphi^{\mathrm{T}}(t)P(t)\varphi(t)\sigma_w^2, \tag{A8}$$

由于

$$A(z)[\eta(t) - w(t)] = A(z)\eta(t) - A(z)y(t) + B(z)u(t)$$
$$= A(z)[-\hat{\theta}^{\mathrm{T}}(t)\varphi(t)] + B(z)u(t) = -A(z)\hat{x}(t) + B(z)u(t)$$
$$= -\hat{x}(t) + \theta^{\mathrm{T}}\varphi(t) = -\hat{\theta}^{\mathrm{T}}(t)\varphi(t) + \theta^{\mathrm{T}}\varphi(t)$$
$$= -\tilde{\theta}^{\mathrm{T}}(t)\varphi(t) = \tilde{u}(t), \tag{A9}$$

由(15)式，(16)式，(A2)式和(A9)式可得

$$\tilde{y}(t) = \frac{1}{2}\tilde{\theta}^{\mathrm{T}}(t)\varphi(t) + [\eta(t) - w(t)]$$
$$= -\frac{1}{2}\tilde{u}(t) + A^{-1}(z)\tilde{u}(t) = [A^{-1}(z) - \frac{1}{2}I]\tilde{u}(t) = H(z)\tilde{u}(t)$$
$$= [A^{-1}(z) - \frac{1+\rho}{2}]\tilde{u}(t) + \frac{\rho}{2}\tilde{u}(t) = \tilde{y}_1(t) + \frac{\rho}{2}\tilde{u}(t) \tag{A10}$$

式中 $\tilde{y}_1(t) = H_1(z)\tilde{u}(t), \quad H_1(z) = [A^{-1}(z) - \frac{1+\rho}{2}].$

由于 $H(z)$ 为严格正实传递函数阵，故存在 $\rho > 0$ 使得 $H_1(z)$ 亦为严格正实传递函数阵，其中 $\tilde{y}_1(t)$ 可认为是 $\tilde{u}(t)$ 作用严格正实传递函数阵 $H_1(z)$ 的输出，故有

$$2\sum_{s=1}^{t}\tilde{u}^{\mathrm{T}}(s)\tilde{y}_1(s) + c_1 > 0, \quad c_1 > 0,$$

$$S(t) = 2\sum_{s=1}^{t}\tilde{u}^{\mathrm{T}}(s)\tilde{y}_1(s) + c_1 + \rho\sum_{s=1}^{t}\|\tilde{u}(s)\|^2 > 0, \quad c_1 > 0. \tag{A11}$$

(A8)式两边加上 $S(t)$ 即得引理 1 的结论.

引理 2 的证明　对下列矩阵恒等式两边取行列式立即得证

$$\begin{bmatrix} I_m & -D \\ 0 & ED + I_n \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ -E & I_n \end{bmatrix}\begin{bmatrix} I_m + DE & -D \\ 0 & I_n \end{bmatrix}\begin{bmatrix} I_m & 0 \\ E & I_n \end{bmatrix}.$$

引理 3 的证明　由于

$$P^{-1}(t-1) = P^{-1}(t) - \varphi(t)\varphi^{\mathrm{T}}(t) = P^{-1}(t)[I - P(t)\varphi(t)\varphi^{\mathrm{T}}(t)],$$

上式两边取行列式，运用引理 2 有

$$|P^{-1}(t-1)| = |P^{-1}(t)\|I - P(t)\varphi(t)\varphi^{\mathrm{T}}(t)| = |P^{-1}(t)|[1 - \varphi^{\mathrm{T}}(t)P(t)\varphi(t)],$$
$$\varphi^{\mathrm{T}}(t)P(t)\varphi(t) = \frac{|P^{-1}(t)| - |P^{-1}(t-1)|}{|P^{-1}(t)|}. \tag{A12}$$

i) $$\sum_{s=1}^{t}\varphi^{\mathrm{T}}(s)P(s)\varphi(s) = \sum_{s=1}^{t}\frac{|P^{-1}(s)| - |P^{-1}(s-1)|}{|P^{-1}(s)|}$$
$$= \sum_{s=1}^{t}\int_{|P^{-1}(s-1)|}^{|P^{-1}(s)|}\frac{\mathrm{d}x}{|P^{-1}(s)|} \leqslant \int_{|P^{-1}(0)|}^{|P^{-1}(t)|}\frac{\mathrm{d}x}{x} = \ln|P^{-1}(t)| - \ln|P^{-1}(0)|$$
$$= \ln|P^{-1}(t)| - \ln|\frac{1}{a}I_n| = \ln|P^{-1}(t)| + n\ln a.$$

ii) $$\sum_{s=1}^{\infty}\frac{\varphi^{\mathrm{T}}(s)P(s)\varphi(s)}{[\ln|P^{-1}(s)|]^c} = \sum_{s=1}^{\infty}\frac{|P^{-1}(s)| - |P^{-1}(s-1)|}{|P^{-1}(s)|[\ln|P^{-1}(s)|]^c}$$
$$\leqslant \int_{|P^{-1}(0)|}^{|P^{-1}(\infty)|}\frac{\mathrm{d}x}{x[\ln x]^c} = \frac{-1}{c-1}\frac{1}{[\ln x]^{c-1}}\Bigg|_{|P^{-1}(0)|}^{|P^{-1}(\infty)|}, \quad (c > 1)$$
$$= \frac{1}{c-1}\left[\frac{1}{[\ln|P^{-1}(0)|]^{c-1}} - \frac{1}{[\ln|P^{-1}(\infty)|]^{c-1}}\right] < \infty.$$

按照上述类似的计算,可得

$$\text{iii)}\quad \sum_{s=1}^{\infty}\frac{\varphi^{\mathrm{T}}(s)P(s)\varphi(s)}{\ln|P^{-1}(s)|[\ln\ln|P^{-1}(s)|]^{c}}$$

$$\leqslant\frac{1}{c-1}\left[\frac{1}{[\ln\ln|P^{-1}(0)|]^{c-1}}-\frac{1}{[\ln\ln|P^{-1}(\infty)|]^{c-1}}\right]<\infty.$$

注 若 $|P^{-1}(0)|$ 太小,上述求和下限 s 可从 s_0开始,$0<s_0<\infty$.

引理 4 的证明 由于

$$r(t)=O(\lambda_{\max}[P^{-1}(t)]),$$

$$r^{n}(t)\geqslant|P^{-1}(t)|\geqslant\lambda_{\max}[P^{-1}(t)]\left(\frac{1}{a}\right)^{n-1}\geqslant r(t)\frac{1}{n}\left(\frac{1}{a}\right)^{n-1},$$

上式两边取对数得

$$\ln r(t)-\ln n-(n-1)\ln a\leqslant\ln|P^{-1}(t)|\leqslant n\ln r(t).$$

此即引理 4.

Convergence Analysis of the Auxiliary Model Identification Algorithm for Multivariable Systems

DING Feng

(Department of Automation, Tsinghua University · Beijing, 100084, PRC)

Abstract: The convergence of the recursive auxiliary model (RAM) identification algorithm was proved provided that both ratio of the maximum to minimum eigenvalues of the covariance matrix $P_0^{-1}(t)=\sum_{s=1}^{s=t}\varphi_0(s)\varphi_0^{\mathrm{T}}(s)$ [i. e. conditional number] and the variance of the observation noise $\{w(t)\}$ are bounded in reference[1]. In this paper, the convergence rate of RAM algorithm is studied under either unbounded conditional number or unbounded noise variance. The results show that the parameter estimates given by RAM algorithm are consistently convergent and that this algorithm is robust.

Key words: Multivariable system; parameter estimation; conditional number; auxiliary model identification algorithm; martingale convergence theorem

本文作者简介

丁 锋 见本刊 1997 年第 1 期第 95 页.