第28卷 第5期
2002年9月

自 动 化 学 报
ACTA AUTOMATICA SINICA

Vol. 28, No. 5
Sept., 2002

时变系统有限数据窗最小二乘辨识的有界收敛性[1)]

丁 锋 丁 韬 萧德云 杨家本
(清华大学自动化系 北京 100084)
(E-mail: dingf@mail.tsinghua.edu.cn)

摘 要 利用随机过程理论证明了有限数据窗最小二乘法的有界收敛性，给出了参数估计误差上界的计算公式，阐述了获得最小均方参数估计误差上界时数据窗长度的选择方法. 分析表明，对于时不变随机系统，数据窗长度越大，均方参数估计误差上界越小；对于确定性时变系统，数据窗长度越小，均方参数估计误差上界越小. 因此，对于时变随机系统，一个折中方案是寻求一个最佳数据窗长度，以使均方参数估计误差最小. 该文的研究成果对于提高辨识算法的实际应用效果有重要意义.

关键词 辨识，参数估计，最小二乘
中图分类号 TP273

BOUNDED CONVERGENCE OF FINITE DATA WINDOW LEAST SQUARES IDENTIFICATION FOR TIME-VARYING SYSTEMS

DING Feng DING Tao XIAO De-Yun YANG Jia-Ben
(*Department of Automation, Tsinghua University, Beijing* 100084)
(E-mail: dingf@mail.tsinghua.edu.cn)

Abstract In this paper the bounded convergence of finite data window least squares algorithm is proved by using stochastic process theory, and the formulae of computing the parameter estimation error are given. The way of choosing the data window length is stated so that the upper bound of the minimum mean square parameter estimation error is obtained. The analyses indicate that for time invariant stochastic systems, the smaller the data window length, the smaller the estimation error upper bound is, and that for deterministic time-varying systems, the larger the data window length, the smaller the estimation error upper bound is. So a compromise is to choose a best data window length for a minimum mean square parameter estimation error.

Key words Identification, parameter estimation, least squares

1) 国家自然科学基金(60074029, 69934010)和清华大学信息学院创新基金资助
收稿日期 2000-11-29 收修改稿日期 2001-08-15

1　引言

最小二乘法[1,2]、卡尔曼滤波算法[2]、辅助模型辨识方法[3,4]和递阶辨识算法[5]等可以给出时不变系统参数的一致估计,但是,它们没有跟踪时变参数和克服数据饱和的能力. 为了解决这一问题,时变参数辨识方法中一般要引入调节参量,如遗忘因子最小二乘法和遗忘梯度算法中的遗忘因子[6,7]、卡尔曼滤波算法或协方差修正算法中的修正项[8]、协方差复位最小二乘算法中的复位间隔[9]、多新息辨识方法中的迭代间隔和新息长度[10]、广义投影算法中记忆长度或数据窗长度、有限数据窗最小二乘法中的数据窗长度等,这使得其收敛性研究更为困难和复杂. 在众多的时变参数辨识方法中,谁的跟踪性能好,谁的估计误差小,谁优谁劣,如何评判. 作者认为,应在统一条件下(如持续激励条件、噪声方差相同等),对各种时变参数估计误差上界进行研究,并寻求调节参量的最佳值,以使均方参数估计误差上界最小. 通过比较各种方法的最小参数估计误差上界,从而作出选择. 因此,首要问题是研究各种时变系统辨识方法的估计误差上界. 本文研究有限数据窗最小二乘法的收敛性及其最小估计误差上界.

2　有限数据窗最小二乘法

设系统用多变量受控 AR(即 CAR)模型描述为

$$A(t,z)\boldsymbol{y}(t)=B(t,z)\boldsymbol{u}(t)+\boldsymbol{v}(t) \tag{1}$$

其中 $\boldsymbol{u}(t)\in R^r$ 为系统输入向量,$\boldsymbol{y}(t)\in R^m$ 为系统输出向量,$\boldsymbol{v}(t)\in R^m$ 为零均值随机噪声向量. 假设阶次 n_a,n_b 已知,$A(t,z)$和 $B(t,z)$均为单位后移算子 z^{-1}的多项式矩阵,且

$$A(t,z)=I+A_1(t)z^{-1}+A_2(t)z^{-2}+\cdots+A_{n_a}(t)z^{-n_a},$$

$$B(t,z)=B_1(t)z^{-1}+B_2(t)z^{-2}+\cdots+B_{n_b}(t)z^{-n_b}.$$

定义参数矩阵 $\Theta(t)$和信息向量 $\boldsymbol{\varphi}(t)$分别为

$\Theta^{\mathrm{T}}(t-1)=[A_1(t),A_2(t),\cdots,A_{n_a}(t),B_1(t),B_2(t),\cdots,B_{n_b}(t)]\in R^{m\times n}$, $n\triangleq mn_a+rn_b$,

$\boldsymbol{\varphi}(t)=[-\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots,-\boldsymbol{y}^{\mathrm{T}}(t-n_a),\boldsymbol{u}^{\mathrm{T}}(t-1),\cdots,\boldsymbol{u}^{\mathrm{T}}(t-n_b)]^{\mathrm{T}}\in R^n$, 则式(1)可等价写为

$$\boldsymbol{y}(t)=\Theta^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t)+\boldsymbol{v}(t) \tag{2}$$

或

$$\boldsymbol{y}^{\mathrm{T}}(t)=\boldsymbol{\varphi}^{\mathrm{T}}(t)\Theta(t-1)+\boldsymbol{v}^{\mathrm{T}}(t) \tag{3}$$

上式称为系统的辨识模型或辨识表达式.

常规最小二乘辨识方法是利用系统直到 t 时刻的输入输出数据$\{\boldsymbol{y}(i),\boldsymbol{\varphi}(i),\ 0\leqslant i\leqslant t\}$,通过极小化下列误差平方和准则函数 $J(\Theta)=\sum_{i=1}^{t}\|\boldsymbol{y}^{\mathrm{T}}(i)-\boldsymbol{\varphi}^{\mathrm{T}}(i)\Theta\|^2$,得到的参数估计算法,其中矩阵 X 的范数定义为 $\|X\|^2=\mathrm{tr}[XX^{\mathrm{T}}]$,$\mathrm{tr}[X]$表示矩阵 X 的迹. 常规最小二乘辨识算法如下:

$$\hat{\Theta}(t)=\hat{\Theta}(t-1)+P_0(t)\boldsymbol{\varphi}(t)[\boldsymbol{y}^{\mathrm{T}}(i)-\boldsymbol{\varphi}^{\mathrm{T}}(i)\hat{\Theta}(t-1)] \tag{4}$$

$$P_0^{-1}(t) = \sum_{i=1}^{t}\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t) = P_0^{-1}(t-1) + \boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t) \tag{5}$$

其中 $\hat{\Theta}(t)$ 为 $\Theta(t)$ 的估计.

随着时间推移和数据的积累，在持续激励条件下，式(5)协方差阵 $P_0(t)$ 将趋于零，故最小二乘算法没有跟踪时变参数的能力. 为了跟踪时变参数，而发展了有数据窗最小二乘法. 顾名思义，有限数据窗最小二乘法是为了克服数据饱和，采用数据窗长度固定的滚动数据窗内的有限数据，即从 $t=t-q+1$ 到 $t=t$ 的 q 组数据 $\{\boldsymbol{y}(i),\boldsymbol{\varphi}(i),\ t-q+1\leqslant i\leqslant t\}$ 进行辨识. 由此不难得到有限内数据窗最小二乘法

$$\hat{\Theta}(t) = \hat{\Theta}(t-1) + P(t)\boldsymbol{\varphi}(t)[\boldsymbol{y}^{\mathrm{T}}(i) - \boldsymbol{\varphi}^{\mathrm{T}}(i)\hat{\Theta}(t-1)] \tag{6}$$

$$P^{-1}(t) = \sum_{i=t-q+1}^{t}\boldsymbol{\varphi}(i)\boldsymbol{\varphi}^{\mathrm{T}}(i) = P^{-1}(t-1) + \boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t) - \boldsymbol{\varphi}(t-q)\boldsymbol{\varphi}^{\mathrm{T}}(t-q) \tag{7}$$

其中 q 称为数据窗长度. 令

$$\begin{aligned} P^{-1}(t) =& P^{-1}(t-1) + \boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t) - \boldsymbol{\varphi}(t-q)\boldsymbol{\varphi}^{\mathrm{T}}(t-q) = \\ & P^{-1}(t-1) + \boldsymbol{\Phi}(t)\boldsymbol{\Phi}^{\mathrm{T}}(t) \end{aligned} \tag{8}$$

$$\boldsymbol{\Phi}(t) = [\boldsymbol{\varphi}(t),\ \boldsymbol{\varphi}(t-q)] \tag{9}$$

则式(7)可以等价写为

$$P(t) = P(t-1) - P(t-1)\boldsymbol{\Phi}(t)[I + \boldsymbol{\Phi}^{\mathrm{T}}(t)P(t-1)\boldsymbol{\Phi}(t)]^{-1}\boldsymbol{\Phi}^{\mathrm{T}}(t)P(t-1) \tag{10}$$

本文的目标是，利用系统的输入输出数据 $\{y(i),\boldsymbol{\varphi}(i),\ t-q+1\leqslant i\leqslant t\}$，采用有限数据窗最小二乘算法，对系统的未知参数矩阵 $\Theta(t)$ 进行实时估计，并研究最小参数估计误差上界和最佳数据窗长度的选择方法.

3 基本引理

设观测噪声 $\{\boldsymbol{v}(t)\}$ 和参数变化率 $\{W(t)\triangleq\Theta(t)-\Theta(t-1)\}$ 是定义在概率空间 (Ω,F,P) 上的鞅差序列，且适应于由直到 t 时刻的观测生成的非降 σ 代数序列 $\{F_t,t\in N\}$，其中 $F_t=\sigma(\boldsymbol{y}(t),\boldsymbol{u}(t),\boldsymbol{y}(t-1),\boldsymbol{u}(t-1),\cdots,\boldsymbol{u}(0))$，$F_0$ 包含所有初始条件信息. $\{\boldsymbol{v}(t)\}$ 和 $\{W(t)\}$ 满足以下噪声假设.

(A1) $E[\boldsymbol{v}(t)\,|\,F_{t-1}]=0$, a. s. , $E[W(t)\,|\,F_{t-1}]=0$, a. s. , $E[\boldsymbol{v}(t)W^{\mathrm{T}}(t)\,|\,F_{t-1}]=0$, a. s. .

(A2) $E[\,\|\boldsymbol{v}(t)\|^2\,|\,F_{t-1}]=\sigma_v^2(t)\leqslant\sigma_v^2<\infty$, a. s. , $E[\,\|W(t)\|^2\,|\,F_{t-1}]=\sigma_W^2(t)\leqslant\sigma_W^2<\infty$, a. s. .

(A3) $\limsup\limits_{t\to\infty}\dfrac{1}{t}\sum\limits_{i=1}^{t}\|\boldsymbol{v}(i)\|^2\leqslant\sigma_v^2<\infty$, a. s. , $\limsup\limits_{t\to\infty}\dfrac{1}{t}\sum\limits_{i=1}^{t}\|W(i)\|^2\leqslant\sigma_W^2<\infty$, a. s. .

引理 1. 对于时变系统(2)和算法(6)～(7)，假设(A1)～(A3)成立，如果存在常数 $0<\alpha\leqslant\beta<\infty$ 和正整数 N，使下列持续激励条件成立

(A4) $\alpha I\leqslant\dfrac{1}{N}\sum\limits_{i=1}^{N}\boldsymbol{\varphi}(t-i+1)\boldsymbol{\varphi}^{\mathrm{T}}(t-i+1)\leqslant\beta I$, a. s. , $t\geqslant N$，则当 $q\geqslant N$ 时，协方差矩阵 $P(t)$ 满足 $(q-N+1)\alpha I\leqslant P^{-1}(t)\leqslant(q+N)\beta I$, a. s. .

证明. 证明很简单，这里从略.

引理 2. 对于时变系统(2)和算法(6)～(7)，假设(A1)～(A4)成立，那么当 $q\geqslant N+1$

时,式$\frac{\alpha}{2(N+1)\beta}\leqslant\frac{(q-N)\alpha}{(q+N)\beta}\leqslant 1-\boldsymbol{\varphi}^{\mathrm{T}}(t)P(t)\boldsymbol{\varphi}(t)\leqslant\frac{(q+N-1)\beta}{(q-N+1)\alpha}$, a.s. 成立.

证明. 利用引理 1,由式(7)可得 $P^{-1}(t)-\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)=\sum_{i=t-q+1}^{t-1}\boldsymbol{\varphi}(i)\boldsymbol{\varphi}^{\mathrm{T}}(i)$, $(q-N)\alpha I\leqslant P^{-1}(t)-\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)\leqslant(q+N-1)\beta I$. 上式左乘以 $\boldsymbol{\varphi}^{\mathrm{T}}(t)P(t)$,右乘以 $P(t)\boldsymbol{\varphi}(t)$,记 $x=\boldsymbol{\varphi}^{\mathrm{T}}(t)P(t)\boldsymbol{\varphi}(t)$,利用引理 1,依次可得

$$(q-N)\alpha\boldsymbol{\varphi}^{\mathrm{T}}(t)P^2(t)\boldsymbol{\varphi}(t)\leqslant x-x^2\leqslant(q+N-1)\beta\boldsymbol{\varphi}^{\mathrm{T}}(t)P^2(t)\boldsymbol{\varphi}(t),$$

$$\frac{(q-N)\alpha}{(q+N)\beta}\boldsymbol{\varphi}^{\mathrm{T}}(t)P(t)\boldsymbol{\varphi}(t)\leqslant x-x^2\leqslant\frac{(q+N-1)\beta}{(q-N+1)\alpha}\boldsymbol{\varphi}^{\mathrm{T}}(t)P(t)\boldsymbol{\varphi}(t),$$

或 $\frac{(q-N)\alpha}{(q+N)\beta}x\leqslant x-x^2\leqslant\frac{(q+N-1)\beta}{(q-N+1)\alpha}x$. 由于 $q\geqslant N+1$,我们有

$$\frac{q-N}{q+N}=\frac{1-\frac{N}{q}}{1+\frac{N}{q}}\geqslant\frac{1-\frac{N}{N+1}}{2}=\frac{1}{2(N+1)},$$

经过简单的推导可得引理 2 的结论. 证毕.

4 主要结果

定理 1. 对于时变系统(2)和算法(6)~(7),假设(A1)~(A4)成立,那么当 $q\geqslant N+1$ 时,有限数据窗最小二乘法给出的参数估计误差 $\|\hat{\Theta}(t)-\Theta(t)\|$ 满足

$$\lim_{t\to\infty}E[\|\hat{\Theta}(t)-\Theta(t)\|^2]\leqslant\frac{k_1\sigma_v^2}{q-N}+k_2(q+c_0)\sigma_w^2\triangleq f(q)\ ,$$

式中 $k_1=\left[\frac{2}{\alpha}+\frac{2N^3\beta}{\alpha^2}\right]\frac{2(N+1)M\beta}{\alpha^2}+\frac{2N^3\beta}{\alpha^2}$, $k_2=[\frac{2}{\alpha}+\frac{2N^3\beta}{\alpha^2}]\frac{2(N+1)\beta^2}{\alpha}$, $c_0=N+\frac{2N^3\beta}{k_2\alpha}$.

证明. 定义参数估计误差向量

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

利用式(6)和(3),有

$$\begin{aligned}\tilde{\Theta}(t)=&\hat{\Theta}(t)-\Theta(t)=\hat{\Theta}(t)-[\Theta(t-1)+W(t)]=\\&\tilde{\Theta}(t-1)+P(t)\boldsymbol{\varphi}(t)[-\boldsymbol{\varphi}^{\mathrm{T}}(t)\tilde{\Theta}(t-1)+\boldsymbol{v}^{\mathrm{T}}(t)]-W(t)=\\&\tilde{\Theta}(t-1)+P(t)\boldsymbol{\varphi}(t)[-\tilde{\boldsymbol{y}}^{\mathrm{T}}(t)+\boldsymbol{v}^{\mathrm{T}}(t)]-W(t)\end{aligned} \tag{12}$$

式中

$$\tilde{\boldsymbol{y}}^{\mathrm{T}}(t)=\boldsymbol{\varphi}^{\mathrm{T}}(t)\tilde{\Theta}(t-1) \tag{13}$$

于是有

$$\begin{aligned}&\tilde{\Theta}^{\mathrm{T}}(t)P^{-1}(t)\tilde{\Theta}(t)=\\&\qquad[\tilde{\Theta}^{\mathrm{T}}(t-1)+[-\tilde{\boldsymbol{y}}(t)+\boldsymbol{v}(t)]\boldsymbol{\varphi}^{\mathrm{T}}(t)P(t)-W^{\mathrm{T}}(t)]P^{-1}(t)\cdot\\&\qquad\{\tilde{\Theta}(t-1)+P(t)\boldsymbol{\varphi}(t)[-\tilde{\boldsymbol{y}}^{\mathrm{T}}(t)+\boldsymbol{v}^{\mathrm{T}}(t)]-W(t)\}=\\&\qquad\tilde{\Theta}^{\mathrm{T}}(t-1)P^{-1}(t)\tilde{\Theta}(t-1)+\tilde{\Theta}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t)[-\tilde{\boldsymbol{y}}^{\mathrm{T}}(t)+\boldsymbol{v}^{\mathrm{T}}(t)]-\\&\qquad\tilde{\Theta}^{\mathrm{T}}(t-1)P^{-1}(t)W(t)+[-\tilde{\boldsymbol{y}}(t)+\boldsymbol{v}(t)]\boldsymbol{\varphi}^{\mathrm{T}}(t)\tilde{\Theta}(t-1)+\end{aligned}$$

$$\begin{aligned}
&[-\tilde{\boldsymbol{y}}(t)+\boldsymbol{v}(t)]\boldsymbol{\varphi}^{\mathrm{T}}(t)P(t)\boldsymbol{\varphi}(t)[-\tilde{\boldsymbol{y}}^{\mathrm{T}}(t)+\boldsymbol{v}^{\mathrm{T}}(t)]-\\
&[-\tilde{\boldsymbol{y}}(t)+\boldsymbol{v}(t)]\boldsymbol{\varphi}^{\mathrm{T}}(t)W(t)-W^{\mathrm{T}}(t)P^{-1}(t)\tilde{\Theta}(t-1)-\\
&W^{\mathrm{T}}(t)\boldsymbol{\varphi}(t)[-\tilde{\boldsymbol{y}}^{\mathrm{T}}(t)+\boldsymbol{v}^{\mathrm{T}}(t)]+W^{\mathrm{T}}(t)P^{-1}(t)W(t)=\\
&\tilde{\Theta}^{\mathrm{T}}(t-1)[P^{-1}(t-1)+\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)-\boldsymbol{\varphi}(t-q)\boldsymbol{\varphi}^{\mathrm{T}}(t-q)]\tilde{\Theta}(t-1)+\\
&\tilde{\boldsymbol{y}}(t)[-\tilde{\boldsymbol{y}}^{\mathrm{T}}(t)+\boldsymbol{v}^{\mathrm{T}}(t)]-\tilde{\Theta}^{\mathrm{T}}(t-1)P^{-1}(t)W(t)+[-\tilde{\boldsymbol{y}}(t)+\boldsymbol{v}(t)]\tilde{\boldsymbol{y}}^{\mathrm{T}}(t)+\\
&\boldsymbol{\varphi}^{\mathrm{T}}(t)P(t)\boldsymbol{\varphi}(t)[-\tilde{\boldsymbol{y}}(t)+\boldsymbol{v}(t)][-\tilde{\boldsymbol{y}}^{\mathrm{T}}(t)+\boldsymbol{v}^{\mathrm{T}}(t)]-\\
&[-\tilde{\boldsymbol{y}}(t)+\boldsymbol{v}(t)]\boldsymbol{\varphi}^{\mathrm{T}}(t)W(t)-W^{\mathrm{T}}(t)P^{-1}(t)\tilde{\Theta}(t-1)-\\
&W^{\mathrm{T}}(t)\boldsymbol{\varphi}(t)[-\tilde{\boldsymbol{y}}^{\mathrm{T}}(t)+\boldsymbol{v}^{\mathrm{T}}(t)]+W^{\mathrm{T}}(t)P^{-1}(t)W(t)=\\
&\tilde{\Theta}^{\mathrm{T}}(t-1)P^{-1}(t-1)\tilde{\Theta}(t-1)-\boldsymbol{x}(t)\boldsymbol{x}^{\mathrm{T}}(t)-[1-\boldsymbol{\varphi}^{\mathrm{T}}(t)P(t)\boldsymbol{\varphi}(t)]\tilde{\boldsymbol{y}}(t)\tilde{\boldsymbol{y}}^{\mathrm{T}}(t)+\\
&\boldsymbol{\varphi}^{\mathrm{T}}(t)P(t)\boldsymbol{\varphi}(t)\boldsymbol{v}(t)\boldsymbol{v}^{\mathrm{T}}(t)+W^{\mathrm{T}}(t)P^{-1}(t)W(t)+\\
&[1-\boldsymbol{\varphi}^{\mathrm{T}}(t)P(t)\boldsymbol{\varphi}(t)][\tilde{\boldsymbol{y}}(t)\boldsymbol{v}^{\mathrm{T}}(t)+\boldsymbol{v}(t)\tilde{\boldsymbol{y}}^{\mathrm{T}}(t)]-\\
&\tilde{\Theta}^{\mathrm{T}}(t-1)P^{-1}(t)W(t)-W^{\mathrm{T}}(t)P^{-1}(t)\tilde{\Theta}(t-1)-\\
&[-\tilde{\boldsymbol{y}}(t)+\boldsymbol{v}(t)]\boldsymbol{\varphi}^{\mathrm{T}}(t)W(t)-W^{\mathrm{T}}(t)\boldsymbol{\varphi}(t)[-\tilde{\boldsymbol{y}}^{\mathrm{T}}(t)+\boldsymbol{v}^{\mathrm{T}}(t)]
\end{aligned}\tag{14}$$

式中

$$\boldsymbol{x}^{\mathrm{T}}(t)=\boldsymbol{\varphi}^{\mathrm{T}}(t-q)\tilde{\Theta}(t-1)\tag{15}$$

定义非负定函数

$$T(t)=\mathrm{tr}[\tilde{\Theta}^{\mathrm{T}}(t)P^{-1}(t)\tilde{\Theta}(t)]\tag{16}$$

条件(A4)取迹可得 $\|\boldsymbol{\varphi}(t)\|^2\leqslant M\triangleq nN\beta$, a. s. .

由于 $\boldsymbol{x}(t)$,$\boldsymbol{\varphi}^{\mathrm{T}}(t)P(t)\boldsymbol{\varphi}(t)$,$\tilde{\boldsymbol{y}}(t)$,$P(t)$,$\tilde{\Theta}^{\mathrm{T}}(t-1)$与 $W(t)$,$\boldsymbol{v}(t)$不相关,且是 F_{t-1}可测的,式(14)两边取迹后对 F_{t-1}取条件期望,利用条件(A1)~(A4)、引理 1 和关系 $\mathrm{tr}[AB]=\mathrm{tr}[BA]$, $\mathrm{tr}[A^{\mathrm{T}}]=\mathrm{tr}[A]$,可得

$$\begin{aligned}
E[T(t)|F_{t-1}]=T(t-1)-\|\boldsymbol{x}(t)\|^2-[1-\boldsymbol{\varphi}^{\mathrm{T}}(t)P(t)\boldsymbol{\varphi}(t)]\|\tilde{\boldsymbol{y}}(t)\|^2+\\
\boldsymbol{\varphi}^{\mathrm{T}}(t)P(t)\boldsymbol{\varphi}(t)\sigma_v^2(t)+E\{\mathrm{tr}[W^{\mathrm{T}}(t)P^{-1}(t)W(t)]|F_{t-1}\}\leqslant\\
T(t-1)-\frac{\alpha}{2(N+1)\beta}\|\tilde{\boldsymbol{y}}(t)\|^2+\frac{M\sigma_v^2}{(q-N+1)\alpha}+(q+N)\beta\sigma_W^2,
\end{aligned}$$

或

$$\begin{aligned}
&E[T(t)|F_{t-1}]-T(t-1)\triangleq\Delta T(t)\leqslant\\
&-\frac{\alpha}{2(N+1)\beta}\left[\|\tilde{\boldsymbol{y}}(t)\|^2-\frac{2(N+1)M\beta}{(q-N+1)\alpha^2}\sigma_v^2-\frac{2(N+1)(q+N)\beta^2}{\alpha}\sigma_W^2\right]\triangleq-b(t)
\end{aligned}\tag{17}$$

考虑集

$$R_t=\left[\tilde{\Theta}(t):\|\tilde{\boldsymbol{y}}(t)\|^2\leqslant\frac{2(N+1)M\beta}{(q-N+1)\alpha^2}\sigma_v^2+\frac{2(N+1)(q+N)\beta^2}{\alpha}\sigma_W^2,\ \text{a. s.}\right]$$

和集

$$N_\varepsilon(R_t)=\left[\tilde{\Theta}(t):\|\tilde{\boldsymbol{y}}(t)\|^2\leqslant\frac{2(N+1)M\beta}{(q-N+1)\alpha^2}\sigma_v^2+\frac{2(N+1)(q+N)\beta^2}{\alpha}{\sigma_W}^2+\varepsilon,\ \text{a. s.},\ \varepsilon>0\right],$$

注意到 $R_t\subset N_\varepsilon(R_t)$,利用式(17),对于 $\tilde{\Theta}(t)\in N_\varepsilon^c(R_t)$,有 $b(t)\geqslant\frac{\alpha\varepsilon}{2(N+1)\beta}>0$ 成立. 应用文

献[6]鞅超收敛定理,对于充分大 t,必有 $\tilde{\Theta}(t)\in N_\varepsilon(R_t)$,由于 ε 是任意的,故对于充分大 t,恒有 $\tilde{\Theta}(t)\in R_t$, a. s. ,即

$$\lim_{t\to\infty}\tilde{\Theta}(t)\in R_\infty=\lim_{t\to\infty}\left[\tilde{\Theta}(t):\|\tilde{\boldsymbol{y}}(t)\|^2\leqslant\frac{2(N+1)M\beta}{(q-N+1)\alpha^2}\sigma_v^2+\frac{2(N+1)(q+N)\beta^2}{\alpha}\sigma_W^2,\ \text{a.s.}\right].$$

或

$$\lim_{t\to\infty}\tilde{\Theta}(t)\in R_\infty=\lim_{t\to\infty}\left[\tilde{\Theta}(t):\|\tilde{\boldsymbol{y}}(t)\|^2\leqslant\frac{2(N+1)M\beta}{(q-N))\alpha^2}\sigma_v^2+\frac{2(N+1)(q+N)\beta^2}{\alpha}\sigma_W^2,\ \text{a.s.}\right].$$

有了上式,我们就可以得到参数估计误差均方有界的结论,其推导过程如下.令

$$\Gamma(t)=P(t)\boldsymbol{\varphi}(t)[-\tilde{\boldsymbol{y}}^{\mathrm{T}}(t)+\boldsymbol{v}^{\mathrm{T}}(t)]-W(t),$$

将上式代入式(12)可得

$$\tilde{\Theta}(t+i)=\tilde{\Theta}(t-1)+\sum_{k=0}^{i}\Gamma(t+k)\tag{18}$$

用$(t+i)$代替式(13)中 t 可得

$$\boldsymbol{\varphi}^{\mathrm{T}}(t+i)\tilde{\Theta}(t+i-1)=\tilde{\boldsymbol{y}}^{\mathrm{T}}(t+i),$$

使用式(18)可得

$$\boldsymbol{\varphi}^{\mathrm{T}}(t+i)\left[\tilde{\Theta}(t-1)+\sum_{k=0}^{i-1}\Gamma(t+k)\right]=\tilde{\boldsymbol{y}}^{\mathrm{T}}(t+i),$$

或

$$\boldsymbol{\varphi}^{\mathrm{T}}(t+i)\tilde{\Theta}(t-1)=\tilde{\boldsymbol{y}}^{\mathrm{T}}(t+i)-\boldsymbol{\varphi}^{\mathrm{T}}(t+i)\sum_{k=0}^{i-1}\Gamma(t+k)\tag{19}$$

上式两边取范数 $\|*\|^2$ 可得

$$\mathrm{tr}[\tilde{\Theta}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t+i)\boldsymbol{\varphi}^{\mathrm{T}}(t+i)\tilde{\Theta}(t-1)]=\left\|\tilde{\boldsymbol{y}}^{\mathrm{T}}(t+i)-\boldsymbol{\varphi}^{\mathrm{T}}(t+i)\sum_{k=0}^{i-1}\Gamma(t+k)\right\|^2\tag{20}$$

上式对 i 从 $i=0$ 到 $i=N-1$ 求和可得

$$\mathrm{tr}\{\tilde{\Theta}^{\mathrm{T}}(t-1)\left[\sum_{i=0}^{N-1}\boldsymbol{\varphi}(t+i)\boldsymbol{\varphi}^{\mathrm{T}}(t+i)\right]\tilde{\Theta}(t-1)=\sum_{i=0}^{N-1}\left\|\tilde{\boldsymbol{y}}^{\mathrm{T}}(t+i)-\boldsymbol{\varphi}^{\mathrm{T}}(t+i)\sum_{k=0}^{i-1}\Gamma(t+k)\right\|^2\tag{21}$$

使用条件(A4)可得

$$N\alpha\|\tilde{\Theta}(t-1)\|^2\leqslant2\left\{\sum_{i=0}^{N-1}\left[\|\tilde{\boldsymbol{y}}(t+i)\|^2+\left\|\boldsymbol{\varphi}^{\mathrm{T}}(t+i)\sum_{k=0}^{i-1}\Gamma(t+k)\right\|^2\right]\right\}\leqslant 2\left\{\sum_{i=0}^{N-1}\left[\|\tilde{\boldsymbol{y}}(t+i)\|^2+N\beta\left\|\sum_{k=0}^{i-1}\Gamma(t+k)\right\|^2\right]\right\},$$

或

$$\|\tilde{\Theta}(t-1)\|^2\leqslant\frac{2}{N\alpha}\left\{\sum_{i=0}^{N-1}\left[\|\tilde{\boldsymbol{y}}(t+i)\|^2+N^2\beta\sum_{k=0}^{N-1}\|\Gamma(t+k)\|^2\right]\right\}=$$

$$\frac{2}{N\alpha}\sum_{i=0}^{N-1}\|\tilde{\boldsymbol{y}}(t+i)\|^2+\frac{2N\beta}{\alpha}\sum_{i=0}^{N-1}\sum_{k=0}^{N-1}\|\Gamma(t+k)\|^2 \tag{22}$$

上式取期望后取极限可得

$$\begin{aligned}
&\lim_{t\to\infty}E[\|\tilde{\Theta}(t)\|^2]\leqslant\\
&\limsup_{t\to\infty}E\left\{\frac{2}{N\alpha}\sum_{i=0}^{N-1}\|\tilde{\boldsymbol{y}}(t+i)\|^2+\frac{2N\beta}{\alpha}\sum_{i=0}^{N-1}\sum_{k=0}^{N-1}\|\Gamma(t+k)\|^2\right\}\leqslant\\
&\limsup_{t\to\infty}E\left\{\frac{2}{\alpha}\|\tilde{\boldsymbol{y}}(t)\|^2+\frac{2N^3\beta}{\alpha}\|\Gamma(t)\|^2\right\}\leqslant\\
&\limsup_{t\to\infty}E\left\{\frac{2}{\alpha}\|\tilde{\boldsymbol{y}}(t)\|^2+\frac{2N^3\beta}{\alpha}\|P(t)\boldsymbol{\varphi}(t)[-\tilde{\boldsymbol{y}}^{\mathrm{T}}(t)+\boldsymbol{v}^{\mathrm{T}}(t)]-W(t)\|^2\right\}\leqslant\\
&\limsup_{t\to\infty}E\left\{\frac{2}{\alpha}\|\tilde{\boldsymbol{y}}(t)\|^2+\frac{2N^3\beta}{\alpha}[\boldsymbol{\varphi}^{\mathrm{T}}(t)P^2(t)\boldsymbol{\varphi}(t)(\|\tilde{\boldsymbol{y}}(t)\|^2+\|\boldsymbol{v}(t)\|^2)+\|W(t)\|^2]\right\}\leqslant\\
&\limsup_{t\to\infty}E\left\{\frac{2}{\alpha}\|\tilde{\boldsymbol{y}}(t)\|^2+\frac{2N^3\beta}{\alpha}\left[\frac{1}{(q-N+1)\alpha}(\|\tilde{\boldsymbol{y}}(t)\|^2+\sigma_v^2)+\sigma_W^2\right]\right\}\leqslant\\
&\limsup_{t\to\infty}\left\{\left[\frac{2}{\alpha}+\frac{2N^3\beta}{\alpha^2}\right]E[\|\tilde{\boldsymbol{y}}(t)\|^2]+\frac{2N^3\beta\sigma_v^2}{(q-N)\alpha^2}+\frac{2N^3\beta\sigma_W^2}{\alpha}\right\}\leqslant\\
&\left[\frac{2}{\alpha}+\frac{2N^3\beta}{\alpha^2}\right]\left[\frac{2(N+1)M\beta}{(q-N)\alpha^2}\sigma_v^2+\frac{2(N+1)(q+N)\beta^2}{\alpha}\sigma_W^2\right]+\\
&\frac{2N^3\beta\sigma_v^2}{(q-N)\alpha^2}+\frac{2N^3\beta\sigma_W^2}{\alpha}=\frac{k_1\sigma_v^2}{q-N}+k_2(q+c_0)\sigma_W^2.
\end{aligned}$$

证毕.

令 $f'(q)=-\frac{k_1\sigma_v^2}{(q-N)^2}+k_2\sigma_W^2=0$，解之得 $q=q_0=N+\sqrt{\frac{k_1}{k_2}}\frac{\sigma_v}{\sigma_W}$. 故最佳数据窗长度为 $q=[q_0]$或$q=[q_0]+1$，对应的最小参数估计误差上界为 $f([q_0])$或$f([q_0]+1)$，其中$[x]$表示不大于 x 的最大整数.

由定理 1，我们可以得到以下三个定理.

定理 2. 对于确定性时不变系统 $\boldsymbol{y}(t)=\boldsymbol{\varphi}(t)\Theta^{\mathrm{T}}$，定理 1 的条件成立，则有限数据窗最小二乘法给出的参数估计以指数速度一致收敛于真参数.

定理 3. 对于随机时不变系统 $\boldsymbol{y}(t)=\Theta^{\mathrm{T}}\boldsymbol{\varphi}(t)+\boldsymbol{v}(t)$，定理 1 的条件成立，则有限数据窗最小二乘法给出的参数估计误差满足

$$\lim_{t\to\infty}E[\|\hat{\Theta}(t)-\Theta\|^2]\leqslant\frac{k_1\sigma_v^2}{q-N}\triangleq f_1(q).$$

故当 q 越大，参数估计误差上界 $f_1(q)$越小. 也就是说，对于时不变系统，常规最小二乘法($q=t$)参数估计误差收敛于零.

定理 4. 对于确定性时变系统 $\boldsymbol{y}(t)=\Theta^{\mathrm{T}}(t)\boldsymbol{\varphi}(t)$，定理 1 的条件成立，则有限数据窗最小二乘法给出的参数跟踪误差满足

$$\lim_{t\to\infty}E[\|\hat{\Theta}(t)-\Theta(t)\|^2]\leqslant k_2(q+c_0)\sigma_W^2\triangleq f_2(q).$$

故当 q 越小，参数估计误差上界 $f_2(q)$越小. 所以最佳数据窗长度为 $q=N+1$，最小均方参数估计误差上界为 $f_2(N+1)=k_2(N+1+c_0)\sigma_W^2$.

参 考 文 献

1 冯纯伯，史 维. 自适应控制. 北京：电子工业出版社，1986

2 Goodwin G C, Sin K S. Adaptive Filtering Prediction and Control. New Jersey, Englewood Cliffs: Prentice-hall, Inc., 1984

3 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

4 丁 锋. 多变量系统的辅助模型辨识方法的收敛性分析. 控制理论与应用，1997，**14**(2):192～200

5 丁 锋，杨家本. 大系统的递阶辨识. 自动化学报，1999，**25**(5):647～654

6 丁 锋，杨家本. 关于鞅超收敛定理与遗忘因子最小二乘算法的收敛性分析. 控制理论与应用，1999，**16**(4):569～572

7 Ding Feng, Yang Jia-Ben, Xu Yong-Mao. Convergence analysis of forgetting gradient algorithm by using martingale hyperconvergence theorem. *Tsinghua Science and Technology*, 2000, **5**(2):188～193

8 王治祥，丁 锋. 一类时变系统参数跟踪估计. 见:1992 年中国控制与决策学术年会论文集(4th CDC)，哈尔滨: 1992. 52～56

9 Goodwin G C, Elliott H, Teoh E K. Deterministic convergence of a self-tuning regulator with covariance resetting. *IEE Proc. Pt. D*, 1983, **130**(1):6～8

10 丁 锋，谢新民，方崇智. 时变系统辨识的多新息方法. 自动化学报，1996，**22**(1):85～91

丁 锋 1984 年毕业于湖北工学院,之后在湖北制药厂工作 4 年,1990 年和 1994 年在清华大学自动化系分别获得硕士学位和博士学位,现任清华大学自动化系副教授. 研究兴趣为自适应辨识与控制及其应用.

丁 韬 1999 年毕业于清华大学自动化系,现为清华大学自动化系在读硕士生. 主要学术方向为自动控制与系统工程.

萧德云 1970 年毕业于清华大学,现任清华大学自动化系教授、博士生导师. 长期从事辨识建模、故障诊断、传感器信号融合、计算机应用和大型连续过程工业 CIMS 等领域的教学和科研.

杨家本 1959 年毕业于清华大学动力系,现任清华大学自动化系教授,博士生导师. 主要学术方向为系统工程和复杂系统的自组织理论与应用.