文章编号: 1001-0920(2002)01-0006-05

衰减激励条件下递阶最小二乘辨识的均方收敛性

丁 锋, 丁 韬, 杨家本, 徐用懋
(清华大学 自动化系, 北京 100084)

摘 要: 为减少递推辨识的计算量, 提出了递阶辨识原理, 它是将系统分解为多个维数较小的虚拟子系统进行辨识, 从而获得递阶最小二乘辨识方法。在衰减激励条件下, 针对时不变系统研究了递阶最小二乘法的收敛性, 得到了参数估计误差均方收敛于零时衰减指数应满足的条件。递阶最小二乘具有良好的性能, 其计算量比递推最小二乘辨识要小得多, 并具有容易实现等优点。

关键词: 辨识; 参数估计; 递阶辨识; 衰减激励
中图分类号: TP 273 **文献标识码**: A

Mean square convergence of hierarchical least square identification under the attenuating excitation

DING Feng, DING Tao, YANG Jia-ben, XU Yong-mao
(Department of Automation, Tsinghua University, Beijing 100084, China)

Abstract: In order to reduce the computational effort of existing identification algorithms, hierarchical identification principle is presented which is to decompose a system into some imaginary subsystems with smaller dimension and fewer variables. The parameters of each subsystem are estimated respectively, and hierarchical least squares (HLS) algorithm is obtained. For time invariant systems, the convergence of the HLS algorithm is studied under the attenuating excitation condition and the analysis shows that the parameter estimation error given by the HLS algorithm consistently converges to zero. The HLS algorithm has good properties and has less calculation than recursive least square identification and is easy to implement.

Key words: identification; parameter estimation; hierarchical identification; attenuating excitation

1 引 言

在以往辨识算法的收敛性分析中, 总是假设关于信息向量的强持续激励条件或弱激励条件成立[1~4]。然而, 实际中往往不允许长时间加入一持续激励信号对系统进行扰动试验。在这种情况下, 利用衰减激励信号得到的试验数据对系统进行辨识就显得格外重要。近年来, 在衰减激励条件下或不满足常规激励条件下, 辨识算法的收敛性研究受到普遍重视。从应用的角度说, 实际系统都存在各种干扰, 因此研究随机系统在衰减激励条件下辨识算法的收敛性, 更具理论价值和应用前景。

递阶最小二乘辨识是作者提出的估计(大)系统参数的一种新型辨识方法[5], 研究其收敛性比递推

收稿日期: 2000-08-28; **修回日期**: 2001-02-26
基金项目: 国家自然科学基金项目(60074029); 国家自然科学基金重点项目(69934010)
作者简介: 丁锋(1963—), 男, 湖北广水人, 副教授, 博士, 从事自适应辨识与控制等研究; 杨家本(1935—), 男, 北京人, 教授, 博士生导师, 从事自组织理论与应用等研究。

最小二乘的收敛性要困难得多。文献[5]在强持续激励条件下, 利用鞅超收敛定理近似地分析了它的收敛性。本文给出了实际中易于实现的衰减激励信号的定义, 从而引出衰减激励条件, 然后研究在此条件下递阶辨识的性能。

2　递阶最小二乘辨识算法

考虑用受控自回归(CAR) 模型描述的动态系统

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

其中

$$A(z) = 1 + a_1 z^{-1} + \cdots + a_{n_a} z^{-n_a}$$

$$B(z) = b_1 z^{-1} + \cdots + b_{n_b} z^{-n_b}$$

$\{u(t)\}$ 和 $\{y(t)\}$ 分别是系统的输入和输出序列, $\{v(t)\}$ 是零均值随机噪声序列, z^{-1} 为单位后移算子。设阶次 n_a 和 n_b 已知。

定义参数向量 θ 和信息向量 $\varphi(t)$ 为

$$\theta = [a_1, a_2, \cdots, a_{n_a}, b_1, b_2, \cdots, b_{n_b}]^{\mathrm{T}} \in R^n$$

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

其中, $\theta \in R^n$ 为待辨识的参数向量, 维数 $n \triangleq n_a + n_b$ 很大。则式(1) 可写成如下向量形式

$$y(t) = \varphi^{\mathrm{T}}(t)\theta + v(t) \tag{2}$$

递阶辨识的基本原理是: 首先将大系统分解为多个维数较小、变量数目较少的子系统, 然后分别估计每个子系统的参数。但是由于各子系统间存在关联变量, 即第 i 个子系统包含其它一些子系统的未知参数向量 $\theta_j\ (j \neq i)$, 使得迭代计算难以进行。为解决这一问题, 在计算 t 时刻 θ_i 的估计时, 其它子系统的未知参数 $\theta_j\ (j \neq i)$ 用它们在 $(t-1)$ 时刻的估计值 $\hat{\theta}_j(t-1)$ 代替。具体做法如下:

将大系统(2) 的信息向量和参数向量分解为 N 个维数为 n_i 的子信息向量和子参数向量, 即

$$\varphi(t) = \begin{bmatrix} \varphi_1(t) \\ \varphi_2(t) \\ \vdots \\ \varphi_N(t) \end{bmatrix}, \quad \theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_N \end{bmatrix} \tag{3}$$

$$\begin{cases} \varphi_i(t) \in R^{n_i}, \quad \theta_i \in R^{n_i} \\ n_1 + n_2 + \cdots + n_N = n \end{cases} \tag{4}$$

于是可将系统(2) 分解为 N 个虚拟子系统, 它们可表示为

$$y(t) - \alpha_i(t) = \varphi_i^{\mathrm{T}}(t)\theta_i + v(t), \quad i = 1, 2, \cdots, N \tag{5}$$

其中各子系统间的关联项

$$\alpha_i(t) = \sum_{j=1, j\neq i}^{N} \varphi_j^{\mathrm{T}}(t)\theta_j, \quad i = 1, 2, \cdots, N \tag{6}$$

递阶辨识就是用系统的输入输出数据 $(y(t), \varphi(t))$ 来估计每个子系统的参数。根据最小二乘原理, 可获得子系统(5) 参数 θ_i 的无偏估计, 算法如下

$$\hat{\theta}_i(t) = \hat{\theta}_i(t-1) + L_i(t)[y(t) - \alpha_i(t) - \varphi_i^{\mathrm{T}}(t)\hat{\theta}_i(t-1)] \tag{7}$$

$$L_i(t) = P_i(t)\varphi_i(t) = \frac{P_i(t-1)\varphi_i(t)}{1 + \varphi_i^{\mathrm{T}}(t)P_i(t-1)\varphi_i(t)} \tag{8}$$

$$P_i^{-1}(t) = P_i^{-1}(t-1) + \varphi_i(t)\varphi_i^{\mathrm{T}}(t) \tag{9}$$

或

$$P_i(t) = [I - L_i(t)\varphi_i^{\mathrm{T}}(t)]P_i(t-1)$$

$$i = 1, 2, \cdots, N$$

其中, $\hat{\theta}_i(t)$ 为 θ_i 在时刻 t 的估计, $\hat{\theta}_i(0)$ 是很小的实向量, $P_i(0) = \frac{1}{a}I,\ 0 < a \leq 1$。

式(7) 中关联变量 $\alpha_i(t)$ 包含了其它子系统的未知参数向量 $\theta_j\ (j \neq i)$, 为了进行递推计算, 可用它们在 $(t-1)$ 时刻的估计值 $\hat{\theta}_j(t-1)$ 代替。于是可得

$$\hat{\theta}_i(t) = \hat{\theta}_i(t-1) + L_i(t)[y(t) - \varphi^{\mathrm{T}}(t)\hat{\theta}(t-1)] \tag{10}$$

算法(8)～(10) 就是大系统(2) 的递阶最小二乘辨识算法(HLS)。若定义分块对角阵

$$P(t) = \mathrm{diag}[P_i(t), i = 1, 2, \cdots, N] \triangleq \begin{bmatrix} P_1(t) & 0 & \cdots & 0 \\ 0 & P_2(t) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & P_N(t) \end{bmatrix} \tag{11}$$

于是, 递阶最小二乘法(8)～(10) 又可等价表述为

$$\hat{\theta}(t) = \hat{\theta}(t-1) + P(t)\varphi(t)[y(t) - \varphi^{\mathrm{T}}(t)\hat{\theta}(t-1)] \tag{12}$$

$$P^{-1}(t) = P^{-1}(t-1) + \mathrm{diag}[\varphi_i(t)\varphi_i^{\mathrm{T}}(t), i = 1, 2, \cdots, N] \tag{13}$$

$$P(0) = P_0 = \frac{1}{a}I, \quad 0 < a \leq 1 \tag{14}$$

设观测噪声 $v(t)$ 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 时刻的观测生成的递增 σ 代数序列 $\{F_t, t \in N\}$。其中, $F_t = \sigma[y(t), u(t), y(t-1), u(t-1), \cdots, u(0)]$, F_0 包含所有初始条件信息, $\{v(t)\}$ 满足下列假设:

A1: $\quad E[v(t) | F_{t-1}] = 0, \quad$ a.s.

A2: $\quad E[v^2(t) | F_{t-1}] = \sigma_v^2(t) \leq \sigma_v^2 < \infty, \quad$ a.s.

A 3: $\lim\sup_{t\to\infty} \frac{1}{t}\sum_{i=1}^{t} v^2(i) \le \sigma_v^2 < \infty$, a. s.

3 衰减激励信号与衰减激励条件的定义

对于信号 $u_0(t)$, 若存在常数 $0 < \alpha_1 \le \beta_1 < \infty$ 和整数 $p > 0$, 使下式成立:

$$\text{A 4:}\quad \alpha_1 I \le \frac{1}{p}\sum_{i=1}^{p} U_0(t-i+1)U_0^{\mathrm T}(t-i+1) \le \beta_1 I$$

$$U_0(t) \triangleq [u_0(t), u_0(t-1), \cdots, u_0(t-n+1)]^{\mathrm T} \in R^n$$

则称 $u_0(t)$ 为 n 阶持续激励信号。

在先前辨识算法的收敛性分析中, 一般要求输入 $u(t)$ 至少为 n 阶持续激励信号, 或要求信息向量 $\varphi(t)$ 是充分丰富的, 即存在常数 $0 < \alpha \le \beta < \infty$ 和整数 $p > 0$, 满足

$$\text{A 5:}\quad \alpha I \le \frac{1}{p}\sum_{i=0}^{p-1} \varphi(t+i)\varphi^{\mathrm T}(t+i) \le \beta I, \quad \text{a. s.} \quad t > 0$$

条件 A 5 称为(n 阶) 强持续激励条件。相应地, 弱持续激励条件定义为

$$\text{A 6:}\quad \lim_{t\to\infty} \frac{1}{t}\sum_{i=1}^{t} \varphi(i)\varphi^{\mathrm T}(i) = R > 0$$

令

$$u(t) = u_0(t)/t^{\epsilon} + u_1(t), \quad \epsilon > 0 \tag{15}$$

其中, $u(t)$ 称为衰减激励信号, $u_0(t)$ 为满足 A 4 的持续激励信号, $u_1(t)$ 为非持续激励信号(最特别的情形是 $u_1(t) \equiv 0$)。对应的衰减激励条件可定义为

$$\text{A 7:}\quad \frac{\alpha}{(t+p-1)^{2\epsilon}} I \le \frac{1}{p}\sum_{i=0}^{p-1} \varphi(t+i)\varphi^{\mathrm T}(t+i) \le \beta I, \quad \text{a. s.}$$

$$t > 0, \quad \alpha > 0, \quad \beta > 0, \quad p > n$$

式中 $\epsilon > 0$ 称为衰减指数。当 $\epsilon = 0$ 时, 条件 A 7 等同于 A 5, 故强持续激励条件是衰减激励条件的一个特例。

值得指出的是, 衰减激励信号的定义可有多种形式, 如

$$u(t) = \frac{u_0(t)}{t^{\epsilon}(\ln t)^c + \cos \pi t + 2} + u_1(t)$$

$$c > 0, \quad \epsilon > 0 \tag{16}$$

对应的衰减激励条件也有多种形式。但是, 式(15) 的衰减激励信号和衰减激励条件 A 7 的表达式最简单, 在工程上也最易实现。因此, 本文针对条件 A 7 进行讨论。

4 主要结果

引理 1 若对称矩阵

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & \cdots & A_{NN} \end{bmatrix}$$

$$A_{ij} = A_{ji}^{\mathrm T} \in R^{n_i \times n_j}$$

满足 $0 < \alpha I \le A \le \beta I$, 其中 I 为单位阵, α 和 β 为常数。则下列两不等式成立。

$$\alpha I \le \mathrm{diag}[A_{ii}, i = 1, 2, \cdots, N] \le \beta I$$

$$\alpha I \le A_{ii} \le \beta I$$

证明较为简单, 这里从略。

引理 2 设非负序列 $\{x(t)\}$, $\{a_t\}$, $\{b_t\}$ 满足下列关系

$$x(t+1) \le (1-a_t)x(t) + b_t, \quad t \ge 0$$

其中 $a_t \in [0, 1)$, $\sum_{t=1}^{\infty} a_t = \infty$, $x(0) < \infty$, 则有

$$\lim_{t\to\infty} x(t) \le \lim_{t\to\infty} b_t/a_t$$

并且假设上式右端极限存在。

证明参见文献[6]。

引理 3 对于系统(2), 设 $H^{-1}(t) = \sum_{i=1}^{t} \varphi(i)\varphi^{\mathrm T}(i) + H^{-1}(0)$, $H(0) = H_0 = P_0 > 0$, 如果衰减激励条件 A 7 成立, 那么协方差矩阵 $H(t)$ 满足

$$(t+p)\beta I + H_0^{-1} \ge H^{-1}(t) \ge \begin{cases} \frac{\alpha}{1-2\epsilon}[t^{1-2\epsilon} - p^{1-2\epsilon}]I + H_0^{-1} & 0 \le \epsilon < 1/2 \\ \alpha(\ln t - \ln p)I + H_0^{-1}, & \epsilon = 1/2 \\ \frac{\alpha}{2\epsilon-1}\left[\frac{1}{p^{2\epsilon-1}} - \frac{1}{t^{2\epsilon-1}}\right]I + H_0^{-1} & \epsilon > 1/2 \end{cases}$$

证明 由于

$$H^{-1}(t) \le (t+p)\beta I + H_0^{-1}, \quad \text{a. s.}$$

$$pH^{-1}(t) \ge \sum_{i=1}^{t-p+1} \frac{p\alpha}{(i+p-1)^{2\epsilon}} I + pH_0^{-1}, \quad \text{a. s.}$$

将文献[7] 的方法用于上式, 即得引理的最后一个不等式。□

引理 4 在引理 3 的条件下, 有

$$\sum_{i=1}^{t}\varphi^{T}(i)H(i)\varphi(i)=O(\ln|H^{-1}(t)|)$$

其中 $f(t)=O(g(t))$ 表示 $g(t)\geq 0$, 且存在常数 $M>0$, 使得 $|f(t)|\leq Mg(t)$ 成立。

证明　由于

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

两边取行列式, 得

$$|H^{-1}(t-1)|=|H^{-1}(t)|[1-\varphi^{T}(t)H(t)\varphi(t)]$$

或

$$\varphi^{T}(t)H(t)\varphi(t)=\frac{|H^{-1}(t)|-|H^{-1}(t-1)|}{|H^{-1}(t)|}$$

因此

$$\sum_{i=1}^{t}\varphi^{T}(i)H(i)\varphi(i)\leq\ln|H^{-1}(t)|-\ln|H_0^{-1}|$$

同理

$$\sum_{i=1}^{t}\frac{|H^{-1}(i)|-|H^{-1}(i-1)|}{|H^{-1}(i-1)|}\geq \ln|H^{-1}(t)|-\ln|H_0^{-1}|$$

$$\sum_{i=1}^{t}\frac{|H^{-1}(i)|-|H^{-1}(i-1)|}{|H^{-1}(i)|}=\sum_{i=1}^{t}\frac{|H^{-1}(i)|-|H^{-1}(i-1)|}{O(|H^{-1}(i-1)|)}$$

由以上三式不难得知引理的结论成立[8]。□

定理 1　对于大系统(2) 和 HLS 算法(12)～(14), 假设 A1～A3 成立, 衰减激励条件 A7 成立, 且衰减指数 $\epsilon\leq 1/2$, 那么 HLS 算法给出的参数估计误差 $\hat{\theta}(t)-\theta$ 均方(m. s) 收敛于零, 即

$$\limsup_{t\to\infty}E[\|\hat{\theta}(t)-\theta\|^{2}]=0$$

或

$$\lim_{t\to\infty}\hat{\theta}(t)=\theta,\quad \text{m. s}$$

证明　鞅理论和随机过程理论是分析辨识方法收敛性的主要工具[8,9], 这里用随机过程理论来证明。式(13) 协方差阵 $P(t)$ 可表示为

$$P^{-1}(t)=P^{-1}(0)+\mathrm{diag}\left[\sum_{j=1}^{t}\varphi_i(j)\varphi_i^{T}(j),\quad i=1,2,\cdots,N\right]$$

由引理 1 知, 用 $P(t)$ 代替 $H(t)$, 引理 3 和引理 4 的结论亦成立。定义参数估计误差向量

$$\tilde{\theta}(t)=\hat{\theta}(t)-\theta \tag{17}$$

式(12) 两边减去 θ, 并将式(2) 代入, 得

$$\tilde{\theta}(t)=[I-P(t)\varphi(t)\varphi^{T}(t)]\times\tilde{\theta}(t-1)+P(t)\varphi(t)v(t) \tag{18}$$

式(18) 两边取范数, 得

$$\begin{aligned}\|\tilde{\theta}(t)\|^{2}=&\tilde{\theta}^{T}(t-1)[I-P(t)\varphi(t)\varphi^{T}(t)]^{T}[I-P(t)\varphi(t)\varphi^{T}(t)]\tilde{\theta}(t-1)+2\tilde{\theta}^{T}(t-1)[I-\\&P(t)\varphi(t)\varphi^{T}(t)]^{T}P(t)\varphi(t)v(t)+\\&\varphi^{T}(t)P^{2}(t)\varphi(t)v^{2}(t)\end{aligned} \tag{19}$$

利用假设 A2 和 A3, 由于 $\tilde{\theta}(t-1)$, $\varphi(t)$, $\tilde{y}(t)$ 和 $v(t)$ 独立, 式(19) 两边取数学期望得

$$\begin{aligned}E[\|\tilde{\theta}(t)\|^{2}]=&E[\tilde{\theta}^{T}(t-1)[I-P(t)\varphi(t)\varphi^{T}(t)]^{T}[I-\\&P(t)\varphi(t)\varphi^{T}(t)]\tilde{\theta}(t-1)]+\\&E[\varphi^{T}(t)P^{2}(t)\varphi(t)]\sigma_v^{2}(t)\end{aligned} \tag{20}$$

设 $[I-P(t)\varphi(t)\varphi^{T}(t)]$ 的最大特征值 $s_{\max}(t)$ 对应的单位特征向量为 x_0, 则有

$$[I-P(t)\varphi(t)\varphi^{T}(t)]x_0=s_{\max}(t)x_0$$

或

$$P(t)\varphi(t)\varphi^{T}(t)x_0=[1-s_{\max}(t)]x_0$$

上式两边同乘以 $\varphi^{T}(t)$, 得

$$\varphi^{T}(t)P(t)\varphi(t)\varphi^{T}(t)x_0=[1-s_{\max}(t)]\varphi^{T}(t)x_0$$

由于 $\varphi(t)$ 是强持续(或衰减) 激励的, 故从上式可求得

$$s_{\max}(t)=1-\varphi^{T}(t)P(t)\varphi(t)$$

由式(20) 得

$$\begin{aligned}E[\|\tilde{\theta}(t)\|^{2}]\leq&\\&E[(1-\varphi^{T}(t)P(t)\varphi(t))\|\tilde{\theta}(t-1)\|^{2}]+\\&E[\lambda_{\max}[P(t)]\varphi^{T}(t)P(t)\varphi(t)]\sigma_v^{2}\end{aligned} \tag{21}$$

式中 $\lambda_{\max}[X]$ 表示矩阵 X 的最大特征值。

利用引理 1, 引理 3 和引理 4, 有

$$\sum_{k=1}^{t}\varphi^{T}(k)P(k)\varphi(k)=O(\ln|P^{-1}(t)|)\to\infty$$

$$t\to\infty,\quad \epsilon\leq 1/2$$

$$\limsup_{t\to\infty}\lambda_{\max}[P(t)]=0,\quad \epsilon\leq 1/2$$

将引理 2 应用于式(21), 利用引理 3, 经过简单的归纳推导不难得出定理的结论。□

5　RLS 算法与 HLS 算法计算量比较

标准递推最小二乘算法(RLS) 和递阶最小二乘算法(HLS) 每迭代计算一步的计算量如表 1 所示。其中小括号内的数字表示 $n=100$, $N=10$, $n_i=10(i=1,2,\cdots,N)$ 时, 两种参数估计算法的计算

量。从表 1 可以看出, RLS 算法的计算量很大, 而 HLS 算法的计算量却小得多。

表 1 RLS 算法与 HLS 算法计算量比较

估计方法	乘法次数	加法次数
RLS	$2n^2 + 4n$ (20 400)	$2n^2 + 2n$ (20 200)
HLS	$\sum_{i=1}^{N}(2n_i^2 + 4n_i)$ (2 400)	$\sum_{i=1}^{N}(2n_i^2 + 2n_i)$ (2 200)

6 数字仿真

考虑下列仿真对象

$$(1 + \theta_1 z^{-1} + \theta_2 z^{-2} + \theta_3 z^{-3}) y(t) = (\theta_4 z^{-1} + \theta_5 z^{-2} + \theta_6 z^{-3}) u(t) + v(t)$$

其中 $\theta_1 = -2.050$, $\theta_2 = 1.375$, $\theta_3 = -0.300$, $\theta_4 = 1.000$, $\theta_5 = 2.500$, $\theta_6 = 1.000$。

仿真时 $\{u(t)\}$ 采用零均值单位方差的可测随机变量序列, $\{v(t)\}$ 采用零均值方差为 $\sigma_v^2 = 1^2$ 的不可测随机噪声序列。将系统分解为两个子系统, 即 $n = 6, N = 2, n_1 = 3, n_2 = 3$。用递阶最小二乘算法估计系统的参数, 不同数据长度下的参数估计 $\hat{\theta}(t)$ 和估计误差 δ 示于表 2。其中 $\delta = \frac{\|\hat{\theta}(t) - \theta\|}{\|\theta\|} \times 100\%$ 为相对参数估计误差范数。表 2 的参数估计精度是令人满意的。

表 2 HLS 算法参数估计及相对误差范数

t	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	δ
100	-2.207 5	1.660 3	-0.429 6	0.794 8	2.446 8	0.704 7	13.3%
300	-2.130 7	1.530 8	-0.375 5	0.895 8	2.316 7	0.778 5	9.5%
500	-2.114 3	1.501 5	-0.362 1	0.995 0	2.340 4	0.780 4	8.3%
1 000	-2.118 8	1.502 2	-0.360 7	0.995 7	2.368 6	0.846 3	0.7%
2 000	-2.023 5	1.321 2	-0.272 3	0.984 0	2.480 9	1.080 9	2.8%
真 值	-2.050	1.375	-0.300	1.000	2.500	1.000	

7 结 语

如何减少大系统辨识的计算量, 一直是辨识领域中有待解决的难题之一。本文提出的递阶最小二乘算法为系统辨识提供了新方法。关于其它模型(如输出误差(OE)模型、CARARMA 模型、CARAR 模型等)的递阶辨识及其收敛速率, 仍是今后有待攻克的研究难题。需要指出的是: 由于强持续激励条件是衰减激励条件下的一个特例, 故文中的结论在强持续激励条件下也成立。

参考文献(References):

[1] 冯纯伯, 史维. 自适应控制[M]. 北京: 电子工业出版社, 1986.

[2] Goodwin G C, Sin K S. Adaptive filtering prediction and control[M]. Englewood Cliffs: Prentice-hall Inc, 1984.

[3] 丁锋, 谢新民, 方崇智. 时变系统辨识的多新息方法[J]. 自动化学报(Acta Automat Sinica), 1996, 22(1): 85-91.

[4] Ding Feng, Xie Xinmin, Fang Chongzhi. Convergence of the forgetting factor algorithm for identifying time-varying systems[J]. Control Theory and Appl, 1994, 11(5): 634-638.

[5] 丁锋, 杨家本. 大系统的递阶辨识[J]. 自动化学报(Acta Automat Sinica), 1999, 25(5): 647-654.

[6] 郭雷. 时变随机系统——稳定性、估计与控制[M]. 长春: 吉林科学技术出版社, 1993.

[7] 丁锋, 杨家本. 衰减激励条件下确定性系统多新息算法的收敛性分析[J]. 清华大学学报(J of Tsinghua Univ), 1998, 38(9): 111-115.

[8] 丁锋. 多变量系统的辅助模型辨识方法的收敛性分析[J]. 控制理论与应用(Control Theory and Appl), 1997, 14(2): 192-200.

[9] 丁锋. 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[J]. 控制理论与应用(Control Theory and Appl), 1999, 16(4): 569-572.