文章编号: 1000 - 8152(2003)01 - 0109 - 04

衰减激励条件下最小均方算法的收敛性

丁 锋, 萧德云, 丁 韬
(清华大学 自动化系,北京 100084)

摘要: 给出了衰减激励信号的定义,并在衰减激励条件下,利用随机过程理论,研究了随机系统最小均方算法的收敛速率,阐述了参数估计误差收敛时,衰减指数和算法中设计参变量(收敛因子或步长)的选择方法.分析表明:在衰减激励条件下,最小均方算法也具有良好的性能:当衰减指数和设计参变量满足一定条件时,则参数估计误差一致收敛于零.

关键词: 参数估计; 辨识; 衰减激励; 最小均方算法
中图分类号: TP273 **文献标识码**: A

Convergence of least mean squares algorithm under attenuating excitation conditions

DING Feng, XIAO De-yun, DING Tao
(Department of Automation, Tsinghua University, Beijing 100084, China)

Abstract: The definition of attenuating excitation signal is given, and the convergence rate of the least mean square algorithm, using stochastic process theory, is studied for stochastic systems under attenuating excitation. The way to choose the attenuating index and design variable (convergent factor or stepsize) is stated for guaranteeing the convergence of the parameter estimates, and the analysis indicates that under attenting excitation the least mean square algorithm also has a good performance: i.e., the parameter estimation error given by the least mean square algorithm uniformly converges to zero when the attenuating index and design variable satisfy some proper conditions.

Key words: parameter estimation; identification; attenuating excitation; least mean square algorithm

1 引言(Introduction)

辨识算法的收敛性往往是考虑时间 $t \to \infty$ 时参数估计的行为,然而在实际中,不容许长时间加入一持续激励信号对系统进行扰动试验,因此,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了普遍重视.文[1,2]分别研究了不完全激励条件和不充分激励条件下辨识算法的收敛性.文[3,4]分别讨论了衰减激励条件下,确定系统投影算法和多信息辨识算法的收敛性.本文将这一研究推广到随机系统,研究最小均方算法在衰减激励条件下的收敛性.分析表明在适当条件下,LMS 参数估计均方误差收敛于零.因为衰减激励信号对系统的影响只是暂时的(短时间的),而又可保证参数估计的收敛性,这对于提高辨识算法的应用效果具有重要意义.在条件衰减激励条件下,研究最小均方算法的收敛性仍是今后需攻克的难题.

2 系统描述与衰减激励条件的定义(System description and the definitions of attenuating excitation conditions)

考虑下列随机系统的辨识问题

$$\begin{cases} A(z)y(t) = B(z)u(t) + v(t), \\ A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_{n_a} z^{-n_a}, \\ B(z) = b_1 z^{-1} + b_2 z^{-2} + \cdots + b_{n_b} z^{-n_b}. \end{cases} \quad (1)$$

其中 $\{u(t)\}$ 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 $\varphi(t)$ 分别为

$$\theta = [a_1, a_2, \cdots, a_{n_a}, b_1, b_2, \cdots, b_{n_b}]^{\mathrm{T}} \in \mathbb{R}^n,$$
$$n = n_a + n_b,$$
$$\varphi(t) = [-y(t-1), \cdots, -y(t-n_a), u(t-1), \cdots, u(t-n_b)]^{\mathrm{T}} \in \mathbb{R}^n,$$

收稿日期:2000 - 08 - 28; 收修改稿日期:2001 - 05 - 21.
基金项目:国家自然科学基金(60074029);国家自然科学基金重点项目(69934010);清华大学信息学院创新基金资助项目.

则式(1)可以写成如下向量形式

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

其中上标 T 表示矩阵转置.

估计系统(2)参数 θ 的最小均方算法有如下形式

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

其中 μ_t 为收敛因子或步长.理论分析表明只要 $0<\mu_t\varphi^{\mathrm{T}}(t)\varphi(t)<2$ 时,LMS 算法就收敛,但并不保证参数估计精度.为了研究方便,通常将 LMS 算法修改为

$$\hat{\theta}(t)=\hat{\theta}(t-1)+\mu_t\frac{\varphi(t)}{1+\|\varphi(t)\|^2}[y(t)-\varphi^{\mathrm{T}}(t)\hat{\theta}(t-1)],\ 0<\mu_t\leqslant 1. \tag{3}$$

其中矩阵 X 的范数定义为 $\|X\|^2=\mathrm{tr}[XX^{\mathrm{T}}]$.

设噪声 $v(t)$ 是定义在概率空间 (Ω,F,P) 上的鞅差序列,且适应于由直到且包含 t 时刻的观测生成的递增 σ 代数序列 $\{F_t,t\in\mathbb{N}\}$,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[5]:

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

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

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

定义 持续激励信号:对于信号 $u_0(t)$,若存在常数 $0<\alpha_1\leqslant\beta_1<\infty$,正整数 $N\geqslant n$ 使下式成立

A4)

$$\alpha_1 I\leqslant\frac{1}{N}\sum_{i=1}^{N}U_0(t-i+1)U_0^{\mathrm{T}}(t-i+1)\leqslant\beta_1 I,\ t>0,$$

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

则称 $u_0(t)$ 为 n 阶持续激励信号.

在先前的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[6],即存在常数 $0<\alpha\leqslant\beta<\infty$ 和正整数 $N\geqslant n$ 满足

A5) $\alpha I\leqslant\dfrac{1}{N-1}\sum\limits_{i=1}^{N-1}\varphi(t+i)\varphi^{\mathrm{T}}(t+i)\leqslant\beta I$, a.s., $t>0$.

条件 A5)称为(n 阶)强持续激励条件.相应地,弱持续激励条件定义为

A6) $\lim\limits_{t\to\infty}\dfrac{1}{t}\sum\limits_{i=1}^{t}\varphi(i)\varphi^{\mathrm{T}}(i)=R>0$.

条件激励条件[6]定义为

A7) $\alpha I\leqslant\mathrm{E}[\dfrac{1}{N}\sum\limits_{i=0}^{N-1}\varphi(t+i)\varphi^{\mathrm{T}}(t+i)\mid F_{t-1}]\leqslant\beta I$, a.s., $t>0$, $0<\alpha\leqslant\beta<\infty$.

令

$$u(t)=\frac{u_0(t)}{t^{\varepsilon}}+u_1(t),\ \varepsilon>0. \tag{4}$$

上式中 $u(t)$ 称为衰减激励信号,ε 称为衰减指数,$u_0(t)$ 为满足 A4)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以定义为

A8) $\dfrac{\alpha}{(t+N-1)^{2\varepsilon}}I\leqslant\dfrac{1}{N}\sum\limits_{i=0}^{N-1}\varphi(t+i)\varphi^{\mathrm{T}}(t+i)\leqslant\beta I$, a.s., $t>0,\alpha,\beta>0,N\geqslant n$.

当 $\varepsilon=0$ 时,条件 A8)等同如 A5),故强持续激励条件是衰减激励条件的一个特例.

同样,条件衰减激励条件定义为

A9) $\dfrac{\alpha}{(t+N-1)^{2\varepsilon}}I\leqslant\mathrm{E}[\dfrac{1}{N}\sum\limits_{i=0}^{N-1}\varphi(t+i)\varphi^{\mathrm{T}}(t+i)\mid F_{t-1}]\leqslant\beta I$, a.s., $t>0,\alpha,\beta>0,N\geqslant n$.

值得指出的是,衰减激励信号的定义可以有多种形式,如

$$u(t)=\frac{u_0(t)}{t^{\varepsilon}(\ln t)^{c}+\cos\pi t+2}+u_1(t),\ c>0,\varepsilon>0, \tag{5}$$

对应的衰减激励条件也有多种形式.但是,式(4)的衰减激励信号和衰减激励条件 A8)的表达式最简单,在工程上也最容易实现.要验证以条件期望形式给出的激励条件 A7)和 A9)都是不容易的,而条件 A5)和 A8)极易验证.因此,本文的讨论是针对条件 A8)进行的.

3 基本引理(Basic lemmas)

引理 1 设非负序列 $\{x(t)\},\{a_t\},\{b_t\}$ 满足下列关系

$$x(t+1)\leqslant(1-a_t)x(t)+b_t,\ t\geqslant 0,$$

而 $a_t\in[0,1)$,$\sum\limits_{t=1}^{\infty}a_t=\infty$,$x(0)<\infty$,则有

$$\lim_{t\to\infty}x(t)\leqslant\lim_{t\to\infty}\frac{b_t}{a_t}.$$

其中假设上式右端极限存在.

证 参见文[6].

引理 2 对于系统(2)和算法(3),定义转移矩阵,即

$$L(t+1,i)=[I-\mu_t\frac{\varphi(t)\varphi^{\mathrm{T}}(t)}{1+\|\varphi(t)\|^2}]L(t,i),\quad L(i,i)=I. \tag{6}$$

如果衰减激励条件 A8)成立,μ_t 是非增的,则有

$$\rho_t\triangleq\lambda_{\max}[L^{\mathrm{T}}(t+N,t)L(t+N,t)]\leqslant 1-\frac{N\alpha\mu_{t+N-1}}{2(1+M)(N^2+1)(t+N-1)^{2\varepsilon}},\ a.s.,$$

$M\triangleq nN\beta$.

其中 $\lambda_{\max}(X)$ 表示 X 的最大特征值.

证　设 v_0 是矩阵 $L^{\mathrm T}(t+N,t)L(t+N,t)$ 的最大特征值 ρ_t 对应的单位特征向量,构造差分方程[4]

$$\begin{cases} x_{i+1} = [I - \mu_t \dfrac{\varphi(i)\varphi^{\mathrm T}(i)}{1+\|\varphi(i)\|^2}]x_i = \\ \qquad L(i+1,i)x_i, \\ x_t = v_0. \end{cases} \tag{7}$$

利用转移矩阵 $L(t,i)$ 的性质 $L(t,i)L(i,s) = L(t,s)$,有

$$x_{t+N} = L(t+N,t)x_t = L(t+N,t)v_0,$$

$$\|x_{t+N}\|^2 = v_0^{\mathrm T}L^{\mathrm T}(t+N,t)L(t+N,t)v_0 = \rho_t,$$

$$\begin{aligned} x_{i+1}^{\mathrm T}x_{i+1} &= x_i^{\mathrm T}[I-\mu_i \frac{\varphi(i)\varphi^{\mathrm T}(i)}{1+\|\varphi(i)\|^2}]^2 x_i = \\ &x_i^{\mathrm T}[I - 2\mu_i \frac{\varphi(i)\varphi^{\mathrm T}(i)}{1+\|\varphi(i)\|^2} + \\ &\mu_i^2 \frac{\|\varphi(i)\|^2\varphi(i)\varphi^{\mathrm T}(i)}{(1+\|\varphi(i)\|^2)^2}]x_i \leqslant \\ &x_i^{\mathrm T}[I - \frac{\mu_i(2-\mu_i)\varphi(i)\varphi^{\mathrm T}(i)}{1+\|\varphi(i)\|^2}]x_i = \\ &x_i^{\mathrm T}x_i - \mu_i(2-\mu_i)\frac{\|\varphi^{\mathrm T}(i)x_i\|^2}{1+\|\varphi(i)\|^2}, \end{aligned}$$

或

$$\mu_i(2-\mu_i)\frac{\|\varphi^{\mathrm T}(i)x_i\|^2}{1+\|\varphi(i)\|^2} \leqslant \|x_i\|^2 - \|x_{i+1}\|^2.$$

由于 $0 < \mu_t \leqslant 1$, 于是有

$$\begin{aligned} &\sum_{i=0}^{N-1}\mu_{t+i}\frac{\|\varphi^{\mathrm T}(t+i)x_{t+i}\|^2}{1+\|\varphi(t+i)\|^2} \leqslant \\ &\sum_{i=0}^{N-1}\mu_{t+i}(2-\mu_{t+i})\frac{\|\varphi^{\mathrm T}(t+i)x_{t+i}\|^2}{1+\|\varphi(t+i)\|^2} \leqslant \\ &\|x_t\|^2 - \|x_{t+N}\|^2 = 1 - \rho_t. \end{aligned} \tag{8}$$

对任意 $i \in [0,N-1]$, 利用公式 $(\sum a_ib_i)^2 \leqslant (\sum a_2^2)(\sum b_i^2)$, 由式(7)和式(8)有

$$\begin{aligned} &\|x_{t+i} - v_0\| = \\ &\left\|\sum_{j=0}^{i-1}\mu_{t+j}\frac{\varphi(t+j)\varphi^{\mathrm T}(t+j)}{1+\|\varphi(t+j)\|^2}x_{t+j}\right\| \leqslant \\ &\sum_{j=0}^{i-1}\mu_{t+j}\frac{\|\varphi(t+j)\|\,\|\varphi^{\mathrm T}(t+j)x_{t+j}\|}{1+\|\varphi(t+j)\|^2} \leqslant \\ &[i\sum_{j=0}^{i-1}\mu_{t+j}^2\frac{\|\varphi^{\mathrm T}(t+j)x_{t+j}\|^2}{1+\|\varphi(t+j)\|^2}]^{1/2} \leqslant \\ &[i\sum_{j=0}^{i-1}\mu_{t+j}\frac{\|\varphi^{\mathrm T}(t+j)x_{t+j}\|^2}{1+\|\varphi(t+j)\|^2}]^{1/2} \leqslant \end{aligned} \tag{9}$$

$$\sqrt{i(1-\rho t)} \leqslant \sqrt{N(1-\rho t)}. \tag{10}$$

衰减激励条件 A8)取迹,可得

$$\|\varphi(t)\|^2 \leqslant M = nN\beta,\ \text{a.s.},\ t > 0. \tag{11}$$

在条件 A8)两边左乘 $v_0^{\mathrm T}$,右乘 v_0,考虑到 μ_t 是非增的,利用式(8)和式(9),有

$$\begin{aligned} &\frac{N\alpha}{(t+N-1)^{2\epsilon}} \leqslant \\ &v_0^{\mathrm T}\sum_{i=0}^{N-1}\varphi(t+i)\varphi^{\mathrm T}(t+i)v_0 \leqslant \\ &\frac{1+M}{\mu_{t+N-1}}v_0^{\mathrm T}\sum_{i=0}^{N-1}\mu_{t+i}\frac{\varphi(t+i)\varphi^{\mathrm T}(t+i)}{1+\|\varphi(t+i)\|^2}v_0 \leqslant \\ &\frac{1+M}{\mu_{t+N-1}}\sum_{i=0}^{N-1}\mu_{t+i}\frac{\|\varphi^{\mathrm T}(t+i)(v_0 - x_{t+i} + x_{t+i})\|^2}{1+\|\varphi(t+i)\|^2} \leqslant \\ &\frac{2(1+M)}{\mu_{t+N-1}}[\sum_{i=0}^{N-1}\mu_{t+i}\frac{\|\varphi^{\mathrm T}(t+i)(x_{t+i} - v_0)\|^2}{1+\|\varphi(t+i)\|^2} + \\ &\sum_{i=0}^{N-1}\mu_{t+i}\frac{\|\varphi^{\mathrm T}(t+i)x_{t+i}\|^2}{1+\|\varphi(t+i)\|^2}] \leqslant \\ &\frac{2(1+M)}{\mu_{t+N-1}}[\sum_{i=0}^{N-1}\|x_{t+i} - v_0\|^2 + (1-\rho t)] \leqslant \\ &\frac{2(1+M)}{\mu_{t+N-1}}[NN(1-\rho t) + (1-\rho t)] = \\ &\frac{2(1+M)}{\mu_{t+N-1}}(N^2+1)(1-\rho t),\ \text{a.s.}. \end{aligned}$$

从上式求得 ρ_t 即得引理的结论.

如果将衰减激励条件 A8)改为

A10) $\dfrac{1}{N}\displaystyle\sum_{i=0}^{N-1}\frac{\varphi(t+i)\varphi^{\mathrm T}(t+i)}{1+\|\varphi(t+i)\|^2} \geqslant \dfrac{\alpha I}{(t+N-1)^{2\epsilon}} > 0$, a.s., $t > 0$, $N \geqslant n$,

则有

$$\rho_t \leqslant 1 - \frac{N\alpha\mu_{t+N-1}}{2(N^2+1)(t+N-1)^{2\epsilon}},\ \text{a.s.}. \tag{12}$$

4　主要结果(Main results)

关于最小均方算法(3)的收敛性有定理.

定理 1　对于系统(2)和算法(3),假设 A1)~A3)成立,噪声 $\{v(t)\}$ 与输入 $\{u(t)\}$ 不相关,衰减激励条件 A8)成立,那么 LMS 算法给出的参数估计误差 $\mathrm E[\|\hat\theta(t) - \theta\|^2]$ 收敛于零充分条件是:非增收敛因子 $\mu_t \in (0,1]$ 满足

$$\sum_{t=1}^{\infty}\frac{\mu_t}{t^{2\epsilon}} = \infty,$$

$$\lim_{t\to\infty}\frac{\mu_{t-N+1}^2(t+N-1)^{2\epsilon}}{\mu_{t+N-1}} = 0.$$

鞅理论和随机过程理论是分析常规辨识方法收敛的主

要工具[7~10],下面用随机过程理论来证明这个定理.

证 定义参数估计误差向量

$$\tilde{\theta}(t)=\hat{\theta}(t)-\theta(t), \tag{13}$$

并假定 $\tilde{\theta}(0)$ 与 $\{v(t)\}$ 无关,且 $\mathrm{E}[\|\tilde{\theta}(0)\|^2]<\infty$,利用式(2)和式(3)可得

$$\begin{aligned}
&\tilde{\theta}(t)=\hat{\theta}(t)-\theta=\\
&\tilde{\theta}(t-1)+\mu_t\frac{\varphi(t)}{1+\|\varphi(t)\|^2}[-\varphi^{\mathrm{T}}(t)\tilde{\theta}(t-1)+v(t)]=\\
&[I-\mu_t\frac{\varphi(t)\varphi^{\mathrm{T}}(t)}{1+\|\varphi(t)\|^2}]\tilde{\theta}(t-1)+\frac{\mu_t\varphi(t)}{1+\|\varphi(t)\|^2}v(t)=\\
&L(t+1,t)\tilde{\theta}(t-1)+\frac{\mu_t\varphi(t)}{1+\|\varphi(t)\|^2}v(t)=\\
&L(t+1,t-N+1)\tilde{\theta}(t-N)+\\
&\sum_{i=0}^{N-1}L(t+1,t-i+1)\frac{\mu_{t-i}\varphi(t-i)}{1+\|\varphi(t-i)\|^2}v(t-i).
\end{aligned} \tag{14}$$

式(14)两边取范数 $\|*\|^2$ 得到

$$\begin{aligned}
&\|\tilde{\theta}(t)\|^2=\\
&\tilde{\theta}^{\mathrm{T}}(t-N)L^{\mathrm{T}}(t+1,t-N+1)\cdot\\
&L(t+1,t-N+1)\tilde{\theta}(t-N)+\\
&2\tilde{\theta}^{\mathrm{T}}(t-N)L^{\mathrm{T}}(t+1,t-N+1)\cdot\\
&\sum_{i=0}^{N-1}L(t+1,t-i+1)\frac{\mu_{t-i}\varphi(t-i)}{1+\|\varphi(t-i)\|^2}v(t-i)+\\
&\|\sum_{i=0}^{N-1}L(t+1,t-i+1)\frac{\mu_{t-i}\varphi(t-i)}{1+\|\varphi(t-i)\|^2}v(t-i)\|^2\leqslant\\
&\tilde{\theta}^{\mathrm{T}}(t-N)L^{\mathrm{T}}(t+1,t-N+1)\cdot\\
&L(t+1,t-N+1)\tilde{\theta}(t-N)+\\
&2\tilde{\theta}^{\mathrm{T}}(t-N)L^{\mathrm{T}}(t+1,t-N+1)\cdot\\
&\sum_{i=0}^{N-1}L(t+1,t-i+1)\frac{\mu_{t-i}\varphi(t-i)}{1+\|\varphi(t-i)\|^2}v(t-i)+\\
&N\sum_{i=0}^{N-1}\|L(t+1,t-i+1)\frac{\mu_{t-i}\varphi(t-i)}{1+\|\varphi(t-i)\|^2}v(t-i)\|^2.
\end{aligned} \tag{15}$$

由于对任意 $i\geqslant 1$,$L^{\mathrm{T}}(t+1,t-i+1)L(t+1,t-i+1)$ 的最大特征值小于或等于 1. 令 $T(t)=\mathrm{E}[\|\tilde{\theta}(t)\|^2]$,式(15)两边取期望,并利用条件 A2)~A5)得到

$$\begin{aligned}
&T(t)\leqslant\\
&\rho_{t-N+1}T(t-N)+0+\\
&N\sum_{i=0}^{N-1}\mathrm{E}\{\frac{\mu_{t-i}\varphi(t-i)}{1+\|\varphi(t-i)\|^2}v(t-i)\|^2\}\leqslant\\
&\rho_{t-N+1}T(t-N)+2N\sum_{i=0}^{N-1}\mu_{t-i}^2\sigma_v^2(t-i)\leqslant\\
&\rho_{t-N+1}T(t-N)+2N\sigma_v^2\sum_{i=0}^{N-1}\mu_{t-i}^2.
\end{aligned} \tag{16}$$

由于 μ_t 是非增的,利用式(12)可得

$$\begin{aligned}
&T(t)\leqslant\\
&\rho_{t-N+1}T(t-N)+N^2\mu_{t-N+1}^2\sigma_v^2\leqslant\\
&(1-\frac{N\alpha\mu_{t+N-1}}{2(N^2+1)(t+N-1)^{2\epsilon}})T(t-\\
&N)+N^2\sigma_v^2\mu_{t-N+1}^2.
\end{aligned} \tag{17}$$

为了使 $T(t)=\mathrm{E}[\|\tilde{\theta}(t)\|^2]\to 0$,根据引理 1,收敛因子(设计参变量)必须满足

$$\sum_{t=1}^{\infty}\frac{N\alpha\mu_{t+N-1}}{2(N^2+1)(t+N-1)^{2\epsilon}}=\infty, \tag{18}$$

$$\begin{aligned}
&\lim_{t\to\infty}T(t)\leqslant\\
&\lim_{t\to\infty}N^2\sigma_v^2\mu_{t-N+1}^2\frac{2(N^2+1)(t+N-1)^{2\epsilon}}{N\alpha\mu_{t+N-1}}=0.
\end{aligned} \tag{19}$$

定理 1 证毕.

定理 2 在定理 1 的假设下:

i) 如果收敛因子 μ_t 满足

$$\mu_t=\frac{c_1}{(t+N-1)^{2\epsilon}[\ln(t+N-1)]^c},\quad c>0,\ c_1>0,\ 0\leqslant 4\epsilon<1,$$

则均方参数估计误差以 $O\left(\frac{1}{(\ln t)^c}\right)$ 速度收敛于零.

ii) 如果收敛因子 μ_t 满足

$$\mu_t=\frac{c_1}{(t+N-1)^{2\epsilon}\ln(t+N-1)},\quad c_1>0,\ 0\leqslant 4\epsilon\leqslant 1,$$

则均方参数估计误差以 $O\left(\frac{1}{\ln t}\right)$ 速度收敛于零.

iii) 如果收敛因子 μ_t 满足

$$\mu_t=\frac{c_1}{(t+N-1)^{2\epsilon+c}},\ c_1>0,\quad c>0,\ 0\leqslant 4\epsilon+c\leqslant 1,$$

则均方参数估计误差以 $O\left(\frac{1}{t^c}\right)$ 速度收敛于零.

证 根据极限和级数理论,满足式(18)和式(19)的 μ_t 无穷多,将 μ_t 代入式(19)不难得到定理 2 的结论.

上述定理说明:若设计参变量选择为 $\mu_t=O(t^{-2\epsilon}/\ln t)$,衰减指数满足 $0\leqslant\epsilon<1/4$,则 LMS 参数估计均方误差以 $O\left(\frac{1}{\ln t}\right)$ 速度收敛于零,即 $\mathrm{E}[\|\tilde{\theta}(t)\|^2]=O\left(\frac{1}{\ln t}\right)$.

(下转第 116 页)

梯度下降学习方法[1]相比所需CPU时间最短.这也是本文提出方法的优点之一.图3中实线是实际系统,虚线是模型输出.

5 结束语(Conclusion)

基于一种新的模糊模型,用三角形隶属函数代替模糊聚类隶属函数,用加权递推最小二乘算法代替启发式误差反馈学习算法,从而大大缩短了非线性系统模糊模型的建模时间,适用于在线辨识与控制.仿真结果表明了这种方法的有效性与实用性.

参考文献(References):

[1] KIM E, PARK M, JI S, et al. A new approach to fuzzy modeling [J]. *IEEE Trans on Fuzzy System*, 1997, 5(3):328 - 337.

[2] PEDRYCZ W. An identification algorithm in fuzzy relational systems [J]. *Fuzzy Sets and System*, 1984, 13(2): 153 - 167.

[3] SUGENO M, YASUKAWA T. A fuzzy-logic-based approach to qualitative modeling [J]. *IEEE Trans on Fuzzy System*, 1993, 1(1): 7 - 31.

[4] TAKAGI T, SUGENO M. Fuzzy identification of systems and its application to modeling and control [J]. *IEEE Trans Systems, Man, and Cybernetics*, 1985, 15(1):116 - 132.

[5] NOZAKI K, TANAKA H. A simple but powerful heuristic method for generating fuzzy rules from numerical data [J]. *Fuzzy Sets and System*, 1997, 86(3): 251 - 270.

[6] LO Ji-Chang, YANG Chien-Hsing. A heuristic error-feedback learning algorithm for fuzzy modeling [J]. *IEEE Trans Systems, Man, and Cybernetics*, 1999, 29(6): 686 - 691.

[7] GOMEZ-SKARMETA A F, DELGADO M, VILA M A, et al. About the of fuzzy clustering techniques for fuzzy model identification [J]. *Fuzzy Sets and Systems*, 1999, 106(2):179 - 188.

作者简介:

刘福才 (1966 —),男,1989年和1994年于东北重型机械学院自动控制系分别获学士学位和硕士学位.现为燕山大学电气工程学院自动化系副教授,哈尔滨工业大学控制工程系在读博士生.目前参加国家杰出青年基金和省自然科学基金2项,获省市级以上奖励3项,发表论文60余篇.研究方向为模糊辨识与预测控制,电力拖动及其计算机控制. Email:lfc_xb@263.net;

关新平 (1963 —),男,博士,燕山大学电气工程学院院长,教授,博士生导师.主要研究方向为鲁棒控制,时滞系统,混沌系统,神经网络,ATM网络控制.在国内外刊物发表论文80余篇.

裴润 (1939 —),男,现为哈尔滨工业大学控制工程系教授,博士生导师.主要研究方向为非线性预测控制,机器人控制,计算机控制.

(上接第112页)

5 结束语(Conclusion)

在衰减激励条件下,对随机系统最小均方算法的收敛性分析表明:只要衰减指数 ε 满足 $0 \leqslant \varepsilon < 1/4$ 时,最小均方算法给出的参数估计误差就收敛于零,而条件数可以是无界的,这放宽了最小二乘收敛的条件.因为衰减激励信号对系统的影响只是暂时的(短时间的),而又可保证参数估计的收敛性,这对于提高辨识算法的应用效果具有重要意义.

参考文献(References):

[1] BITTANTI S, BOLZERN P, CAMPI M. RLS identification algorithms with incomplete excitation: convergence analysis and application to adaptive control [J]. *IEEE Trans Automatic Control*, 1990, 35(12):1371 - 1373.

[2] VIANO M C. Partial convergence of extended least squares for insufficiently excited linear systems [J]. *Int J Adaptive Control and Signal Processing*, 1990, 4(3):207 - 217.

[3] CHEN H F, GUO L. Adaptive control via consistent estimation for deterministic systems [J]. *Int J Control*, 1987, 45(6):2183 - 2202.

[4] DING Feng, YANG Jiaben. Convergence analysis of multi-innovation identifiction under attenuating excitation condition for deterministic systems [J]. *J of Tsinghua University*, 1998, 38(9): 111 - 115 (in Chinese).

[5] GOODWIN G C, SIN K S. *Adaptive filtering prediction and control* [M]. Englewood Cliffs, NJ: Prentice-Hall, 1984.

[6] GUO Lei. *Time-Varying Stochastic Systems: Stability, Estimation and Control* [M]. Changchun: Jilin Science and Technology Press, 1993 (in Chinese).

[7] DING Feng, XIE Xinmin, FANG Chongzhi. Multi-innovation identification method for time-varying systems [J]. *Acta Automatica Sinica*, 1996, 22(1):85 - 91 (in Chinese).

[8] DING Feng, YANG Jiaben. Comment on martingale hyperconvergence theorem and the convergence analysis of the forgetting factor least squares algorithms [J]. *Control Theory & Applications*, 1999, 16(4):569 - 572 (in Chinese).

[9] DING Feng, XIE Xinmin. Auxiliary model identification algorithms for multivariable systems [J]. *J of Tsinghua University*, 1992, 32(4):100 - 106 (in Chinese).

[10] DING Feng, XIE Xinmin. Convergence analysis of recursive extended least square algorithm for multivariable systems [J]. *Control and Decision*, 1992, 7(6):443 - 447 (in Chinese).

作者简介:

丁锋 (1963 —),男,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,1997年任清华大学自动化系副教授,现在加拿大Alberta大学作访问教授.研究兴趣为自适应辨识与控制及其应用.以第1作者发表学术论文70余篇. Email:dingf@mail.tsinghua.edu.cn;

萧德云 (1945 —),男,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, Email:xiaody@mail.tsinghua.edu.cn;

丁韬 (1977 —),男,1999年毕业于清华大学自动化系,2002年获得清华大学自动化系硕士学位,现留学美国攻读博士学位.主要学术方向为自动控制与系统工程. Email:dingtao@mails.tsinghua.edu.cn