Article ID: 1000 - 8152(2002)03 - 0423 - 05

Bounded Convergence of Forgetting Factor Least Square Algorithm for Time-Varying Systems *

DING Feng, XIAO Deyun and DING Tao
(Department of Automation, Tsinghua University · Beijing, 100084, P. R. China)

Abstract: Based on stochastic process theory, the bounded convergence of forgetting factor least square algorithm (FFLS for short) is studied and the upper bound of the parameter tracking error is given. The analyses indicate that: i) for time-invariant deterministic systems, the estimates given by the FFLS algorithm converge to their true values at exponential rate; ii) for time-invariant stochastic systems, the FFLS algorithm can give a bounded mean square parameter estimation error; iii) for time-varying stochastic systems, the FFLS algorithm may track the time-varying parameters and its parameter tracking error is bounded (that is, the parameter tracking error is small when the parameter change rate is small).

Key words: time-varying system; identification; parameter estimation; least squares; bounded convergence
Document code: A

时变系统遗忘因子最小二乘法的有界收敛性

丁 锋 萧德云 丁 韬
(清华大学自动化系·北京, 100084)

摘要: 利用随机过程理论研究了遗忘因子最小二乘法(FFLS)的有界收敛性,给出了参数估计误差的上界.分析表明:i) 对于时不变确定性系统,FFLS 算法产生的参数估计以指数速度收敛于真参数; ii) 对于时不变随机系统,FFLS 算法给出有界均方估计误差; iii) 对于时变随机系统,FFLS 算法可以跟踪时变参数,且跟踪误差有界.

关键词: 时变系统; 辨识; 参数估计; 最小二乘; 有界收敛性

1 Problem formulation

Consider the following time-varying system[1,2]

$$y(t) = \varphi^{\mathrm{T}}(t)\theta(t-1) + v(t), \tag{1}$$

where $y(t)$ is the output of the system, $\theta(t) \in \mathbb{R}^n$ is the time-varying parameter vector of the system to be identified, $\varphi(t) \in \mathbb{R}^n$ is the regressive information vector consisting of the observations up to time $(t-1)$, and $\{v(t)\}$ is a stochastic noise sequence with zero mean, the superscript T denotes a matrix transpose.

The forgetting factor least square algorithm is an important method to identify the parameters of time-varying systems, and it has good tracking performance. Its convergence has been attracted much attention[1~8]. However, most references only studied the limit behavior of the estimation error[5~7]. Based on stochastic process theory, the convergence of the FFLS algorithm is studied and the upper bound of the parameter tracking error is given in this paper.

The objective of this paper is, by means of the FFLS algorithm, to obtain the real-time estimation of the time-varying parameter vector $\theta(t)$ by utilizing the observations $(y(i), \varphi(i), i \leqslant t)$ up to and including time t, and the upper bound of the parameter estimation error is found out, which results in quite good tracking performance in the simulation examples.

The least square algorithm with the forgetting factor for identifying the time-varying parameter vector of model (1) can be described as

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

$$P^{-1}(t) = \lambda P^{-1}(t-1) + \varphi(t)\varphi^{\mathrm{T}}(t),\ 0 < \lambda < 1,\ P(0) = P_0 > 0, \tag{3}$$

* Foundation item: supported by National Natural Science Foundation of China (60074029, 69934010) and the Foundation of Information School of Tsinghua University (985 - 50 - 50).
Received date: 2000 - 07 - 25; Revised date: 2001 - 02 - 26.

where $\hat{\theta}(t)$ denotes the estimates of $\theta(t)$, λ is the forgetting factor, $\hat{\theta}(0)$ is a random variable with $\mathrm{E}[\hat{\theta}^{\mathrm{T}}(0)\hat{\theta}(0)] \leqslant M_0 < \infty$, $\hat{\theta}(0)$ and $\{v(t)\}$ are independent.

2 Bounded convergence of the FFLS algorithm

Lemma 1 For system (1) and algorithms (2), (3), assume that the following strong persistent excitation condition holds

A1) $\alpha I \leqslant \frac{1}{N}\sum_{i=1}^{N}\varphi(t+i)\varphi^{\mathrm{T}}(t+i) \leqslant \beta I$, a.s., for any $t > 0$, $0 < \alpha \leqslant \beta < \infty$, $N \geqslant n$, then the covariance matrix $P(t)$ satisfies

$$\frac{\lambda^{N-1}}{1-\lambda}\alpha I + \lambda^t[P_0^{-1} - \frac{\alpha}{1-\lambda}I] \leqslant$$
$$P^{-1}(t) \leqslant \frac{N\beta}{1-\lambda}I + \lambda^t[P_0^{-1} - \frac{N\beta}{1-\lambda}I].$$

Proof See Ref. [6], we have

$$P^{-1}(t) = \lambda P^{-1}(t-1) + \varphi(t)\varphi^{\mathrm{T}}(t) \leqslant \frac{N\beta}{1-\lambda}I + \lambda^t[P_0^{-1} - \frac{N\beta}{1-\lambda}I],$$

$$NP^{-1}(t) = N\sum_{i=1}^{t}\lambda^{t-i}\varphi(i)\varphi^{\mathrm{T}}(i) + N\lambda^t P^{-1}(0) \geqslant \frac{\lambda^{N-1}-\lambda^t}{1-\lambda}N\alpha I + N\lambda^t P_0^{-1} = \frac{\lambda^{N-1}}{1-\lambda}N\alpha I + N\lambda^t[P_0^{-1} - \frac{\alpha}{1-\lambda}I].$$

This completes the proof of Lemma 1. As $t \to \infty$, we have

$$\frac{\lambda^{N-1}}{1-\lambda}\alpha I \leqslant \lim_{t\to\infty}P^{-1}(t) \leqslant \frac{N\beta}{1-\lambda}I,\ 0 < \lambda < 1.$$

For convenience, let P_0^{-1} satisfy

$$\frac{\alpha}{1-\lambda}I \leqslant P_0^{-1} \leqslant \frac{N\beta}{1-\lambda}I,$$

then

$$\frac{\lambda^{N-1}}{1-\lambda}\alpha I \leqslant P^{-1}(t) \leqslant \frac{N\beta}{1-\lambda}I,\ 0 < \lambda < 1. \quad (4)$$

Theorem 1 For system (1) and algorithms (2), (3), assume that Condition A1) holds, $\{v(t)\}$ is an independent random variable sequence with zero mean and mean square bounded, i.e.

A2)
$$\mathrm{E}[v^2(t)] = \sigma_v^2(t) \leqslant \sigma_v^2 < \infty,$$
$$\mathrm{E}[v(t)v(i)] = 0,\ t \neq i,$$
$$\mathrm{E}[\varphi(t-i)v(t)] = 0,\ i \geqslant 0.$$

The parameter changing rate $w(t) \triangleq \theta(t) - \theta(t-1)$ is bounded, $\{w(t)\}$ and $\{v(t)\}$ are independent, i.e.

A3)
$$\mathrm{E}[\|w(t)\|^2] = \sigma_w^2(t) \leqslant \sigma_w^2 < \infty,$$
$$\mathrm{E}[w(t)w^{\mathrm{T}}(i)] = 0,\ t \neq i,$$
$$\mathrm{E}[w(t)v(i)] = 0,$$

then as $t \to \infty$, the estimation error given by the FFLS algorithm is uniformly bounded, i.e.

$$\mathrm{E}[\|\hat{\theta}(t) - \theta(t)\|^2] \leqslant 3\alpha^{-2}\lambda^{2(t-N+1)}(1-\lambda)^2\|P_0^{-1}\|^2 M_0 + \frac{3n(1-\lambda)}{\alpha\lambda^{N-1}}\sup_t \mathrm{E}[v^2(t)] + \frac{3N^2\beta^2}{\alpha^2\lambda^{2(N-1)}(1-\lambda)^2}\sup_t \mathrm{E}[\|w(t)\|^2] \triangleq f(\lambda, t),$$

where the norm of the matrix X is defined by $\|X\|^2 = \mathrm{tr}[XX^{\mathrm{T}}]$, $\mathrm{tr}[X]$ denotes the trace of the matrix X.

Proof Define the parameter tracking error as

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

Using relations (1) ~ (3), we get

$$\tilde{\theta}(t) = \hat{\theta}(t) - [\theta(t-1) + w(t)] = [I - P(t)\varphi(t)\varphi^{\mathrm{T}}(t)]\tilde{\theta}(t-1) + P(t)\varphi(t)v(t) - w(t) =$$
$$\lambda P(t)P^{-1}(t-1)\tilde{\theta}(t-1) + P(t)\varphi(t)v(t) - w(t) =$$
$$\lambda^t P(t)P^{-1}(0)\tilde{\theta}(0) + P(t)\sum_{i=1}^{t}\lambda^{t-i}\varphi(i)v(i) - P(t)\sum_{i=1}^{t}\lambda^{t-i}P^{-1}(i)w(i) \triangleq$$
$$\gamma_1(t) + \gamma_2(t) + \gamma_3(t), \quad (5)$$

where

$$\gamma_1(t) = \lambda^t P(t)P_0^{-1}\tilde{\theta}(0),$$
$$\gamma_2(t) = P(t)\sum_{i=1}^{t}\lambda^{t-i}\varphi(i)v(i) = P(t)H_t^{\mathrm{T}}v_t,$$
$$\gamma_3(t) = -P(t)\sum_{i=1}^{t}\lambda^{t-i}P^{-1}(i)w(i),$$
$$H_t^{\mathrm{T}} = [\mu^{t-1}\varphi(1), \mu^{t-2}\varphi(2), \cdots, \mu\varphi(t-1), \varphi(t)],$$
$$v_t = [\mu^{t-1}v(1), \mu^{t-2}v(2), \cdots, \mu v(t-1), v(t)]^{\mathrm{T}},$$
$$\mu \triangleq \sqrt{\lambda}.$$

$P(t)$ in (3) may be expressed as

$$P^{-1}(t) = H_t^{\mathrm{T}}H_t + \lambda^t P_0^{-1}.$$

Since

$$0 \leqslant \mathrm{E}[\|\gamma_1(t)\|^2] = \lambda^{2t}\mathrm{E}[\|P(t)P_0^{-1}\tilde{\theta}(0)\|^2] \leqslant$$

$$\frac{\lambda^{2t}(1-\lambda)^2}{\alpha^2\lambda^{2(N-1)}}\|P_0^{-1}\|^2\|\tilde{\theta}(0)\|^2 \leqslant$$
$$\alpha^{-2}\lambda^{2(t-N+1)}(1-\lambda)^2\|P_0^{-1}\|^2 M_0. \tag{6}$$

Since v_t^{T} and H_t are statistically independent, we have[8]

$$0 \leqslant \mathrm{E}[\|\gamma_2(t)\|^2] =$$
$$\mathrm{E}[\|P(t)H_t^{\mathrm{T}}v_t\|^2] = \mathrm{E}\{\mathrm{tr}[P(t)H_t^{\mathrm{T}}v_t v_t^{\mathrm{T}}H_tP(t)]\} \leqslant$$
$$\mathrm{E}\{\mathrm{tr}[P(t)H_t^{\mathrm{T}}H_tP(t)]\sigma_v^2 = \mathrm{tr}\{\mathrm{E}[H_tP^2(t)H_t^{\mathrm{T}}]\}\sigma_v^2 \leqslant$$
$$\mathrm{tr}\{\mathrm{E}[P(t)H_t^{\mathrm{T}}H_t]\}\frac{(1-\lambda)\sigma_v^2}{\alpha\lambda^{N-1}} \leqslant \frac{n(1-\lambda)}{\alpha\lambda^{N-1}}\sigma_v^2, \text{ for large } t, \tag{7}$$

$$0 \leqslant \mathrm{E}[\|\gamma_3(t)\|^2] =$$
$$\mathrm{E}[\|P(t)\sum_{i=1}^{t}\lambda^{t-i}P^{-1}(i)w(i)\|^2] \leqslant$$
$$\frac{(1-\lambda)^2}{\alpha^2\lambda^{2(N-1)}}\mathrm{E}[\|\sum_{i=1}^{t}\lambda^{t-i}P^{-1}(i)w(i)\|^2] =$$
$$\frac{(1-\lambda)^2}{\alpha^2\lambda^{2(N-1)}}\mathrm{E}[\sum_{i=1}^{t}\sum_{j=1}^{t}\lambda^{2t-i-j}w^{\mathrm{T}}(i)P^{-1}(i)P^{-1}(j)w(j)].$$

Using $2x^{\mathrm{T}}Qy \leqslant x^{\mathrm{T}}Qx + y^{\mathrm{T}}Qy, (Q \geqslant 0)$, we have

$$0 \leqslant \mathrm{E}[\|\gamma_3(t)\|^2] \leqslant$$
$$\frac{(1-\lambda)^2}{\alpha^2\lambda^{2(N-1)}}\sum_{i=1}^{t}\sum_{j=1}^{t}\lambda^{2t-i-j}\mathrm{E}[w^{\mathrm{T}}(i)P^{-2}(i)w(i) +$$
$$w^{\mathrm{T}}(j)P^{-2}(j)w(j)]/2 \leqslant$$
$$\frac{(1-\lambda)^2}{\alpha^2\lambda^{2(N-1)}}\sum_{i=1}^{t}\sum_{j=1}^{t}\lambda^{2t-i-j}\frac{N^2\beta^2\sigma_w^2}{(1-\lambda)^2} =$$
$$\frac{N^2\beta^2\sigma_w^2}{\alpha^2\lambda^{2(N-1)}}\frac{(1-\lambda^t)^2}{(1-\lambda)^2} \leqslant \frac{N^2\beta^2}{\alpha^2\lambda^{2(N-1)}(1-\lambda)^2}\sigma_w^2. \tag{8}$$

Taking the norm $\|\cdot\|^2$ of both sides of (5) gives

$$\mathrm{E}[\|\hat{\theta}(t)-\theta(t)\|^2] =$$
$$\mathrm{E}[\|\gamma_1(t)+\gamma_2(t)+\gamma_3(t)\|^2] \leqslant$$
$$3\{\mathrm{E}[\|\gamma_1(t)\|^2] + \mathrm{E}[\|\gamma_2(t)\|^2] + \mathrm{E}[\|\gamma_3(t)\|^2]\}. \tag{9}$$

Substituting (6 ~ 8) into (9) will lead to the conclusion of Theorem 1. This proves the assertion of Theorem 1.

Theorem 2 For time invariant deterministic systems

$$y(t) = \varphi^{\mathrm{T}}(t)\theta,$$

Condition A1) holds, then for $\{\hat{\theta}(t)\}$ given by the FFLS algorithm, as $t \to \infty$, $\mathrm{E}[\|\hat{\theta}(t)-\theta\|^2] \to 0$ exponentially fast.

Proof Since $v(t) \equiv 0$ and $w(t) \equiv 0$, from (9) it is easy to obtain

$$\mathrm{E}[\|\tilde{\theta}(t)\|^2] = \mathrm{E}[\|\gamma_1(t)\|^2] \leqslant$$
$$\alpha^{-2}\lambda^{2(t-N+1)}(1-\lambda)^2\|P_0^{-1}\|^2 M_0 \triangleq f_1(\lambda,t),$$
$$\mathrm{E}[\|\hat{\theta}(t)-\theta\|^2] = O(\lambda^{2t}) \to 0,\ 0 < \lambda < 1.$$

This proves Theorem 2.

Theorem 3 For time invariant stochastic systems

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

Assumptions A1) ~ A3) hold, then $\{\hat{\theta}(t)\}$ given by the FFLS algorithm, as $t \to \infty$, satisfies

$$\mathrm{E}[\|\hat{\theta}(t)-\theta\|^2] \leqslant$$
$$2\alpha^{-2}\lambda^{2(t-N+1)}(1-\lambda)^2\|P_0^{-1}\|^2 M_0 +$$
$$\frac{2n(1-\lambda)}{\alpha\lambda^{N-1}}\sigma_v^2 \triangleq f_2(\lambda,t).$$

Proof Since $w(t) \equiv 0$, from (9) we have

$$\mathrm{E}[\|\hat{\theta}(t)-\theta\|^2] =$$
$$\mathrm{E}[\|\gamma_1(t)+\gamma_2(t)\|^2] \leqslant$$
$$2\{\mathrm{E}[\|\gamma_1(t)\|^2] + \mathrm{E}[\|\gamma_2(t)\|^2]\}. \tag{10}$$

Substituting (6) and (7) into (10) will lead to the conclusion of Theorem 3. This completes the proof of Theorem 3.

Theorem 4 For time-varying systems

$$y(t) = \varphi^{\mathrm{T}}(t)\theta(t),$$

Assumptions A1) ~ A3) hold, then $\{\hat{\theta}(t)\}$ given by the FFLS algorithm, as $t \to \infty$, satisfies

$$\mathrm{E}[\|\hat{\theta}(t)-\theta(t)\|^2] \leqslant$$
$$2\alpha^{-2}\lambda^{2(t-N+1)}(1-\lambda)^2\|P_0^{-1}\|^2 M_0 +$$
$$\frac{2N^2\beta^2}{\alpha^2\lambda^{2(N-1)}(1-\lambda)^2}\sigma_w^2 \triangleq f_3(\lambda,t).$$

Proof Since $v(t) \equiv 0$, from (9) we have

$$\mathrm{E}[\|\hat{\theta}(t)-\theta(t)\|^2] =$$
$$\mathrm{E}[\|\gamma_1(t)+\gamma_3(t)\|^2] \leqslant$$
$$2\{\mathrm{E}[\|\gamma_1(t)\|^2 + \mathrm{E}[\|\gamma_3(t)\|^2]\}. \tag{11}$$

Substituting (6) and (8) into (11) will lead to the conclusion of Theorem 4. This completes the proof of Theorem 4.

Theorem 5 For system (1) and algorithms (2), (3), assume that the conditions of Theorem 1 hold, then for the variable forgetting factor $0 < \lambda_{\min} \leqslant \lambda_t \leqslant \lambda_{\max} < 1$, we have

$$\mathrm{E}[\|\hat{\theta}(t)-\theta(t)\|^2] \leqslant$$
$$3\alpha^{-2}\lambda_{\max}^{2(t-N+1)}(1-\lambda_{\min})^2\|P_0^{-1}\|^2 M_0 +$$
$$\frac{3n(1-\lambda_{\min})}{\alpha\lambda_{\min}^{N-1}}\sigma_v^2 + \frac{3N^2\beta^2}{\alpha^2\lambda_{\min}^{2(N-1)}(1-\lambda_{\max})^2}\sigma_w^2.$$

Proof To be omitted.

Theorems 1 ~ 5 may be used to compute the upper bounds of the estimation errors given by the FFLS algorithm.

3 Simulation studies

Example 1 Consider the following time-invariant system

$A(z)y(t) = B(z)u(t) + v(t)$,

$A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} = 1 - 1.2z^{-1} + 0.52z^{-2}$,

$B(z) = b_1 z^{-1} + b_2 z^{-2} = 1.2z^{-1} - 0.52z^{-2}$,

$\theta^T = [-1.2, 0.52, 1.2, -0.52]$,

$\varphi^T(t) = [-y(t-1), -y(t-2), u(t-1), u(t-2)]$,

where $\{u(t)\}$ is taken as a zero mean and unit variance random variable sequence, and $\{v(t)\}$ as a white noise sequence with zero mean and variance $\sigma_v^2 = 0.2^2$. The FFLS algorithm is applied to estimate the parameters of this system. The results of simulation are shown in Table 1, Table 2 and in Fig.1, where the noise-signal ratio of the system is $\delta_{ns} = 23.49\%$, the forgetting factor is $\lambda = 0.95$.

Table 1 The parameter estimates of Example 1

t	a_1	a_2	b_1	b_2	δ
100	-1.1732	0.5090	1.2169	-0.4620	3.623%
300	-1.3074	0.5864	1.2182	-0.5837	7.711%
500	-1.1865	0.5023	1.2046	-0.4847	2.268%
1000	-1.1649	0.5225	1.1582	-0.4902	3.366%
2000	-1.1438	0.4849	1.1924	-0.4652	4.670%
true values	-1.2000	0.5200	1.2000	-0.5200	

Table 2 The estimation error and its upper bound of Example 1

t	$\Vert\tilde{\theta}(t)\Vert^2$	$f_2(\lambda, t)$
100	0.004489363	0.325792432
300	0.020337632	0.340827316
500	0.001760129	0.312141001
1000	0.003876060	0.296362013
2000	0.007459215	0.286299318

Example 2 Consider the following time-varying system

$y(t) + a_1(t)y(t-1) + a_2(t)y(t-2) =$

$b_1(t)u(t-1) + b_2(t)u(t-2) + v(t)$,

$a_1(t) = -1.20 + 0.01\sin(0.01t)$, $a_2(t) = 0.52$,

$b_1(t) = 1.20$, $b_2(t) = 1.2 + 0.001\sqrt{t + 100}$.

Simulation conditions are the same as those of Example 1, $\lambda = 0.75$. The estimation error and its upper bound are shown in Table 3 and in Fig.2.

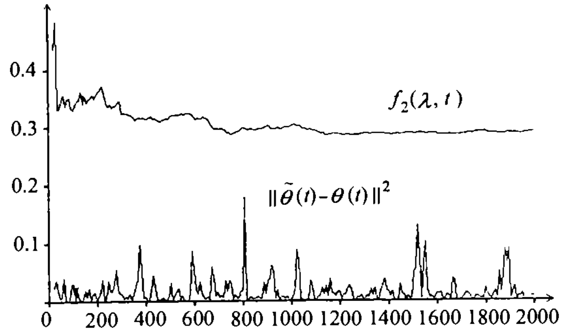

Fig. 1 The estimation error and upper bound of Example 1

Table 3 The estimation error and its upper bound of Example 2

t	$\Vert\tilde{\theta}(t)\Vert^2$	$f(\lambda, t)$
100	0.014390509	0.726138771
300	0.023099916	0.562118948
500	0.016471678	0.581226349
1000	0.019195369	0.556566477
2000	0.010161544	0.561459243

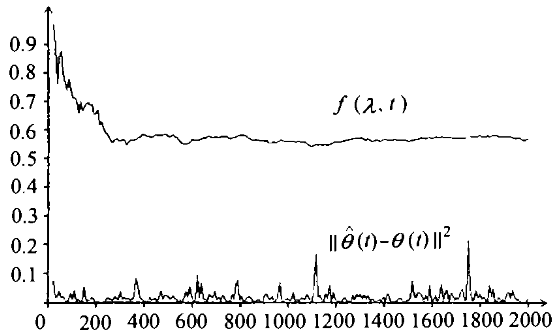

Fig. 2 The estimation error and upper bound of Example 2

From Tables 1 ~ 3 and Figs.1,2, the estimation error is less than its upper bound, the conclusions are correct.

4 Conclusions

From the convergence analyses the following conclusions are reached: i) for time-invariant deterministic systems, PEE given by the FFLS algorithm converges to zero at an exponential rate, but PEE given by the RLS algorithm converges to zero at the rate of $(1/t)$; ii) for the time-invariant stochastic systems, the FFLS algorithm gives a bounded mean square PEE, but PEE given by the RLS algorithm converges to zero under the mean square sense, and its convergence rate is of $(1/t)$; iii) for the time-varying stochastic systems, the mean square PEE given by the FFLS algorithm is bounded, but the mean square PEE given by the RLS algorithm is unbounded, so the RLS algorithm has no performance to track the time-varying parameters.

References

[1] Wittenmark B. A two-level estimator for time-varying parameters [J]. Automatica, 1979, 15(1):85 – 89

[2] Moustafa K A F. Identification of stochastic time-varying systems [J]. IEE Proc. Part-D, 1983, 130(4):137 – 142

[3] Lozano L R. Convergence analysis of recursive identification algorithms with forgetting factor [J]. Automatica, 1983, 19(1):95 – 97

[4] Canetti R M and M D Espana. Convergence analysis of the least-square identification algorithm with a variable forgetting factor for time-varying linear systems [J]. Automatica, 1989, 25(4):609 – 612

[5] Ding Feng. Martingale hyperconvergence theorem and the convergence of forgetting factor least square algorithm [J]. Control Theory and Applications, 1997, 14(1):90 – 95 (in Chinese)

[6] Ding Feng and Yang Jiaben. Comments on martingale hyperconvergence theorem and the convergence of forgetting factor least square algorithm [J]. Control Theory and Applications, 1999, 16(4):569 – 572 (in Chinese)

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

[8] Fang Chongzhi and Xiao Deyun. Process Identification [M]. Beijing: Tsinghua University Press, 1988

本文作者简介

DING Feng Born in 1963. He received the B.S. degree in electrical engineering from Hubei University of Technology in 1984, and the M.S. and the Ph. D degrees in automatic control all from Department of Automation, Tsinghua University in 1990 and 1994, respectively. Since 1988, he has been with Department of Automation, Tsinghua University. He is currently associate professor at Tsinghua University, Beijing. He has authored and co-authored over sixty papers. His research interests are in stochastic systems, identification and parameter estimation, time-varying systems. Email: dingf@mail.tsinghua.edu.cn

XIAO Deyun Born in 1945. He graduated from Tsinghua University in 1970. Now he is a professor and Ph. D advisor in Department of Automation, Tsinghua University. His main area of interest and expertise is identification modeling, fault diagnosis, multi-sensor fusion, computer control, CIMS in process industry and so on. Email: xiaody@mail.tsinghua.edu.cn

DING Tao Born in 1977. He received the B.S. degree in automatic control from Department of Automation, Tsinghua University, Beijing, P.R. China in 1999. He is currently a graduate at Tsinghua University, Beijing. His research interests are in identification, parameter estimation and systems engineering. Email: dingtao@mails.tsinghua.edu.cn