短 文

时变系统辨识的多新息方法

丁 锋　谢新民　方崇智

(清华大学自动化系　100084)

摘　　要

推广了估计时不变参数的单新息修正技术，提出了多新息辨识方法. 该方法可以抑制坏数据对参数估计的影响，具有较强的鲁棒性. 分析表明多新息方法可以跟踪时变参数，计算量也较遗忘因子最小二乘法和卡尔曼(Kalman)滤波算法要小. 仿真结果说明多新息算法估计系统参数是有效的.

关键词: 参数估计，时变系统，收敛性，多新息技术.

1　引言

现有的一些参数估计方法，如最小二乘法、卡尔曼滤波(KF)算法、最小均方算法(LMS)等都是使用单新息修正技术的单新息(single innovation)辨识方法，即 t 时刻的参数估计 $\hat{\theta}(t)$ 是在 $\hat{\theta}(t-1)$ 的基础上依靠增益向量 $\boldsymbol{L}(t)$ 与标量新息(即单新息) $e(t)$ 的乘积来修正，亦即[1]

$$\hat{\theta}(t)=\hat{\theta}(t-1)+\boldsymbol{L}(t)e(t).$$

本文将单新息修正技术加以推广，从新息修正角度提出了多新息辨识算法. 对标量系统而言，多新息修正算法的参数估计是通过新息向量(多新息)来修正的；对多变量系统而言，多新息算法的参数估计是通过新息矩阵(多新息)来修正的.

对于时变参数系统，Goodwin 等[2] 建议用协方差复位(resetting)技术来跟踪时变参数，其缺点是使协方差阵失去了以前的信息，参数估计不断出现过渡过程. 同样，投影算法完全没有利用过去的估计误差信息，其参数估计性能也难以保证. 而本文的多新息算法和遗忘因子最小二乘法(FFLS)一样，都利用了以前的部分信息，因而有较好的辨识效果.

2　多新息辨识算法及其收敛性分析

考虑下列时变系统

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

其中 $\theta(t)\in R^n$ 为系统的时变参数向量，$y(t)\in R^1$ 为系统输出，$u(t)\in R^1$ 为系统输入，

本文于1993年7月26日收到

$v(t)\in R^1$ 为零均值随机噪声，$\varphi(t)\in R^n$ 是由系统输出输入数据 $(y(t-1), u(t-1), y(t-2), u(t-2), \cdots)$ 构成的信息向量，上标 T 表示矩阵转置.

估计系统(1)时变参数 $\theta(t)$ 的多新息辨识算法可表达为

$$\hat{\theta}(t_{s+1})=\hat{\theta}(t_s)+\frac{\Gamma(p,\ t_s)}{r(q,\ t_s)}\boldsymbol{E}(p,\ t_s), \tag{2a}$$

$$\boldsymbol{E}(p,\ t_s)=Y(p,\ t_s)-\Gamma^{\mathrm{T}}(p,\ t_s)\hat{\theta}(t_s), \tag{2b}$$

$$\hat{\theta}(t)=\hat{\theta}(t_s),\ t\in T_s=\{t_s,\ t_s+1,\ t_s+2,\ \cdots,\ t_{s+1}-1\}, \tag{2c}$$

$$r(q,\ t_s)=1+\mathrm{tr}[\Gamma(q,\ t_s)\Gamma^{\mathrm{T}}(q,\ t_s)],\ q\geqslant p, \tag{2d}$$

$$\Gamma(p,\ t_s)=[\varphi(t_{s+1}-1),\ \varphi(t_{s+1}-2),\ \cdots,\ \varphi(t_{s+1}-p)]\in R^{n\times p}\ \text{为增益矩阵},$$

$$\boldsymbol{E}(p,\ t_s)=[e(t_{s+1}-1),\ e(t_{s+1}-2),\ \cdots,\ e(t_{s+1}-p)]^{\mathrm{T}}\in R^p\ \text{为新息向量},$$

$$Y(p,\ t_s)=[y(t_{s+1}-1),\ y(t_{s+1}-2),\ \cdots,\ y(t_{s+1}-p)]^{\mathrm{T}}\in R^p, \tag{2e}$$

$$e(t_s)=y(t_s)-\hat{y}(t_s),\ \hat{y}(t_s)=\varphi^{\mathrm{T}}(t_s)\theta(t_s). \tag{2f}$$

其中 $\hat{\theta}(t)$ 为 $\theta(t)$ 的估计，tr 表示迹. 整数序列 $\{t_s\}$ 满足 $1=t_0<t_1<t_2<\cdots$，$1\leqslant t_s^*=t_{s+1}-t_s<\infty$.

工程上，t_s^* 可认为是控制周期，但在这一个周期 t_s^* 秒内，每 $\frac{1}{N}$ 秒(每 $\frac{2}{N}$ 秒，…)都可获得一组可用的数据信息，可利用其中 p 组数据信息产生新息向量 $\boldsymbol{E}(p,\ t_s)$ 来对 $\hat{\theta}(t_s)$ 进行修正，多新息修正算法就是因此而得名. 当 $t_s^*=p=1$ 时，多新息算法就是常规单新息的投影算法; 当 $t_s^*<p$ 时，多新息算法利用了过去的部分信息.

关于多新息辨识算法的收敛性有定理.

定理 1. 对于时变系统(1)，假设 1)系统噪声 $\{v(t)\}$ 均方有界，即

(A1) $$\sigma_v^2\triangleq\sup_t E[v^2(t)]<\infty$$

2) 参数变化率 $w(t)=\theta(t)-\theta(t-1)$ 均方有界，即

(A2) $$\sigma_w^2\triangleq\sup_t E\|w(t)\|^2<\infty$$

其中范数定义为 $\|X\|^2=\mathrm{tr}(XX^{\mathrm{T}})$.

3) 持续激励条件成立:

(A3) $$\frac{1}{p}\sum_{i=1}^{p}\varphi(t-i+1)\varphi^{\mathrm{T}}(t-i+1)\geqslant\alpha I,\ \alpha>0,\ \text{a.s.},\ \forall\ \ t\geqslant 0$$

$$0\leqslant m\leqslant\|\varphi(t)\|^2=\varphi^{\mathrm{T}}(t)\varphi(t)\leqslant M<\infty,\ \text{a.s.},\ \forall\ \ t\geqslant 0$$

那么多新息算法给出的参数估计误差 $\|\hat{\theta}(t)-\theta(t)\|$ 有界，即

$$\lim_{t\to\infty}E\|\hat{\theta}(t)-\theta(t)\|\leqslant A\cdot\{\sup E[v^2(t)]\}^{\frac{1}{2}}+B\cdot\{\sup E\|w(t)\|^2\}^{\frac{1}{2}},$$

$$\begin{cases}A=\dfrac{pM\cdot\max(t_{\max}^*-1,\ p-t_{\min}^*)}{(qm+1)(1-\rho)}+\dfrac{t_{\max}^*}{1-\rho},\ B=\dfrac{p\sqrt{M}}{(qm+1)(1-\rho)},\\ t_{\max}^*=\max(t_1^*,\ t_2^*,\ t_3^*,\ \cdots),\ t_{\min}^*=\min(t_1^*,\ t_2^*,\ t_3^*,\ \cdots),\ 0<\rho<1.\end{cases}$$

证明. 由于 $\theta(t+1)=\theta(t)+w(t+1)$，$\theta(t_s+i)=\theta(t_s)+\sum_{j=1}^{i}w(t_s+j)$ 当 $i\leqslant t_s^*$ 时，有

$$\theta(t_{s+1}-i)=\theta(t_s+t_s^*-i)=\theta(t_s)+\sum_{j=1}^{t_s^*-i}w(t_s+j), \tag{3}$$

而 $\theta(t-1)=\theta(t)-w(t)$ $\theta(t_s-j)=\theta(t_s)-\sum_{l=1}^{j}w(t_s-l+1)$. 当 $i>t_s^*$时，有

$$\theta(t_{s+1}-i)=\theta[t_s-(i-t_s^*)]=\theta(t_s)-\sum_{l=1}^{i-t_s^*}w(t_s-l+1). \tag{4}$$

定义函数 $\delta(x)=\begin{cases}1, & x\geqslant 0,\\ 0, & x<0,\end{cases}$ 则对于任意 $i\geqslant 0$,，(3)式和(4)式可以合并成

$$\theta(t_{s+1}-i)=\theta(t_s)+\sum_{j=1}^{|t_s^*-i|}[\delta(t_s^*-i)w(t_s+j)-\delta(i-t_s^*)w(t_s-j+1)]. \tag{5}$$

由(2e)式和(1)式可得

$$Y(p,, t_s)=\Gamma^{\mathrm{T}}(p,, t_s)\theta(t_s)+W(p, t_s)+V(p, t_s). \tag{6}$$

其中

$$W(p, t_s)=\begin{bmatrix}\varphi^{\mathrm{T}}(t_{s+1}-1)\sum_{j=1}^{|t_s^*-1|}[\delta(t_s^*-1)w(t_s+j)-\delta(1-t_s^*)w(t_s-j+1)]\\ \vdots\\ \varphi^{\mathrm{T}}(t_{s+1}-p)\sum_{j=1}^{|t_s^*-p|}[\delta(t_s^*-p)w(t_s+j)-\delta(p-t_s^*)w(t_s-j+1)]\end{bmatrix},$$

$$V(p, t_s)=\begin{bmatrix}v(t_{s+1}-1)\\ \vdots\\ v(t_{s+1}-p)\end{bmatrix}.$$

定义参数估计误差向量 $\tilde{\theta}(t)=\hat{\theta}(t)-\theta(t)$，利用(2a)、(2b)式和(6)式，有

$$\begin{aligned}\tilde{\theta}(t_{s+1})&=\hat{\theta}(t_{s+1})-\theta(t_{s+1})\\ &=\left[I-\frac{\Gamma(p, t_s)\Gamma^{\mathrm{T}}(p, t_s)}{r(q, t_s)}\right]\tilde{\theta}(t_s)+\psi(p, t_s).\end{aligned} \tag{7}$$

其中

$$\psi(p\ t_s)=\frac{\Gamma(p, t_s)[W(p, t_s)+V(p, t_s)]}{r(q, t_s)}-\sum_{i=1}^{t_s^*}w(t_s+i). \tag{8}$$

利用条件(A1)—(A3)，有

1) $$qm+1\leqslant r(q, t_s)\leqslant qM+1,$$

2) $$\left[I-\frac{\Gamma(p, t_s)\Gamma^{\mathrm{T}}(p, t_s)}{r(q, t_s)}\right]\leqslant\rho I,\ 0<\rho<1,$$

3) $$E\|\Gamma(p, t_s)W(p, t_s)\|\leqslant pM\cdot\max(t_s^*-1, p-t_s^*)\sigma_w,$$

4) $$E\|\Gamma(p, t_s)V(p, t_s)\|\leqslant p\sqrt{M}\ \sigma_v.$$

令 $T(s)=E\|\tilde{\theta}(t_s)\|$，(7)式两边取范数，并利用以上诸式，可得

$$T(s+1)\leqslant\rho T(s)+\left[\frac{pM\cdot\max(t_s^*-1,\ p-t_s^*)}{(qm+1)}+t_s^*\right]\sigma_w+\frac{p\sqrt{M}}{(qm+1)}\sigma_v$$

$$\leqslant\rho^s T(1)+\frac{1-\rho^s}{1-\rho}\left[\left(\frac{pM\cdot\max(t_{\max}^*-1,\ p-t_{\min}^*)}{(qm+1)}+t_{\max}^*\right)\sigma_w+\frac{p\sqrt{M}}{(qm+1)}\sigma_v\right]. \tag{9}$$

(9)式两边取极限($s\to\infty$)得到

$$\lim_{t\to\infty}E\|\hat{\theta}(t)-\theta)\|=\lim_{t\to\infty}E\|\tilde{\theta}(t)\|$$
$$=\lim_{s\to\infty}E\|\tilde{\theta}(t_{s+1})\|=\lim_{s\to\infty}T(s+1)=A\sigma_w+B\sigma_v,$$

这就证明了定理1的论断.

3 数字仿真

例1. 仿真下列时不变系统

$$A(z)y(t)=B(z)u(t)+v(t),$$
$$A(z)=1+a_1z^{-1}+a_2z^{-2}=1-0.75z^{-1}+0.5z^{-2},$$
$$B(z)=b_1z^{-1}+b_2z^{-2}=z^{-1}+0.3z^{-2},$$
$$\theta(t)=\theta=[a_1,\ a_2,\ b_1,\ b_2]^{\mathrm{T}},$$
$$\varphi(t)=[-y(t-1),\ -y(t-2),\ u(t-1),\ u(t-2)]^{\mathrm{T}},$$

其中 z^{-1} 为单位后移算子.

仿真时$\{u(t)\}$采用零均值单位方差不相关可测随机变量序列，$\{v(t)\}$采用零均值方差为 $\sigma_v^2=0.3^2$ 不相关随机噪声序列. 利用多新息算法(2)估计这个系统的参数(取 $t_s^*=p=n=4$，$q=10$)，不同迭代次数(t/t_s^*)下的参数估计如表1所示，其中 δ_{ns} 为噪信比，

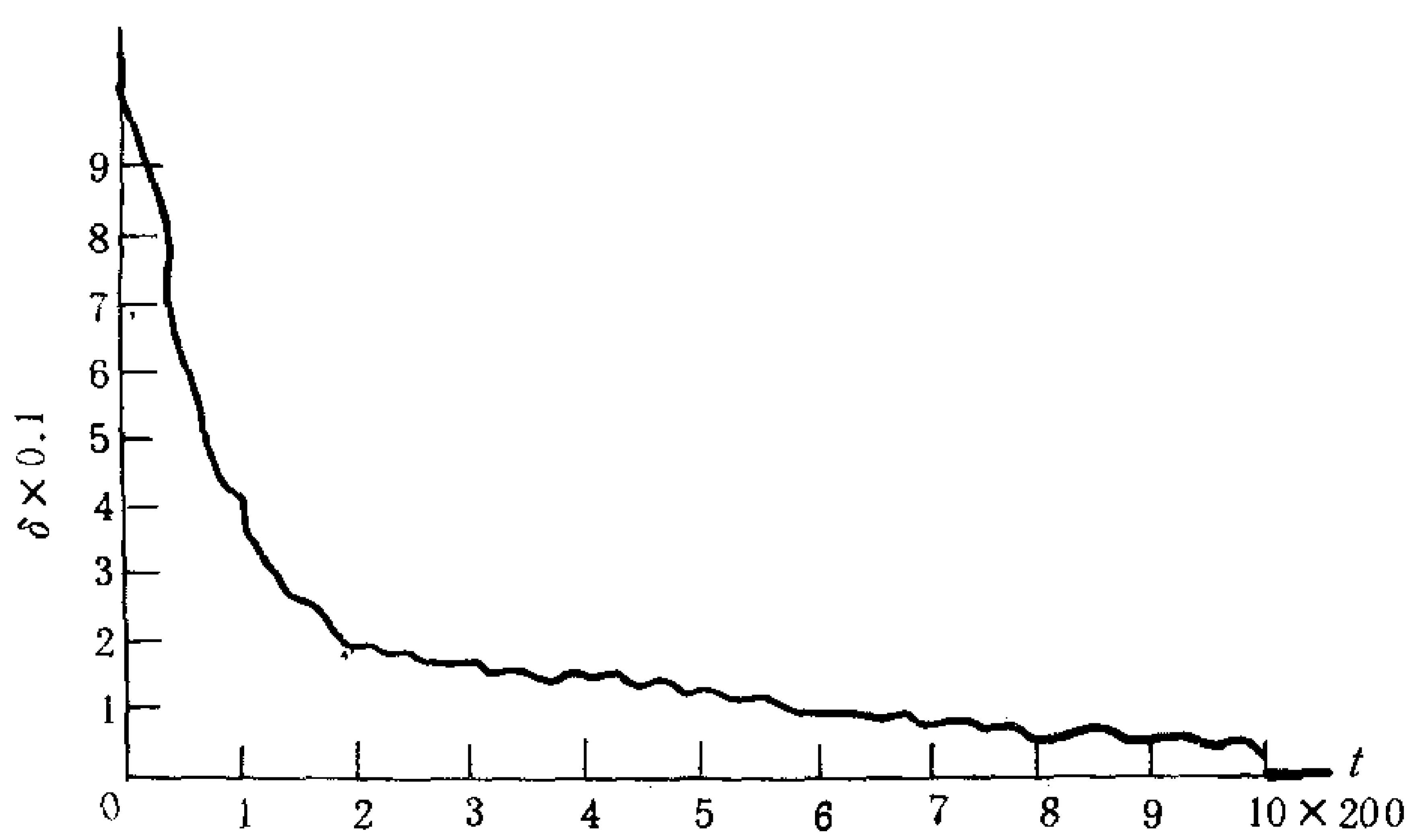

图1 例1的 δ 随 t 变化曲线

参数估计相对误差范数定义为 $\delta=\|\hat{\theta}(t)-\theta(t)\|/\|\theta(t)\|$，$\delta$ 随 t 变化曲线如图 1 所示. 从表 1 和图 1 可知，δ 随 t 增大而下降，说明多新息算法估计时不变参数也是有效的.

表 1　例1 的参数及其估计

	$\sigma_v^2=0.30^2$		$\delta_{ns}=25.45\%$		
t/t_s^*	a_1 -0.7500	a_2 0.5000	b_1 1.0000	b_2 0.3000	δ (%)
$300/t_s^*$	−0.6165	0.4182	0.9634	0.4603	16.461
$500/t_s^*$	−0.6485	0.4336	0.9755	0.4123	12.114
$1000/t_s^*$	−0.7677	0.4867	1.0033	0.3805	6.055
$2000/t_s^*$	−0.7055	0.4997	1.0343	0.2709	4.586

例 2.　仿真下列时变系统

$$y(t)+a_1(t)y(t-1)=b_1(t)u(t-1)+v(t),$$

$$\theta(t)=\begin{bmatrix} a_1(t) \\ b_1(t) \end{bmatrix}=\begin{bmatrix} 0.5 \\ 0.5+0.5\sqrt[4]{t} \end{bmatrix},\quad \varphi(t)=\begin{bmatrix} -y(t-1) \\ u(t-1) \end{bmatrix}.$$

仿真条件同例 1. 参数及其估计如图 2(a)所示（取 $t_s^*=p=n=2$，$q=4$），δ 随 t 变化曲线如图 2(b)所示. 从图 2 可知，多新息算法具有跟踪时变参数的能力.

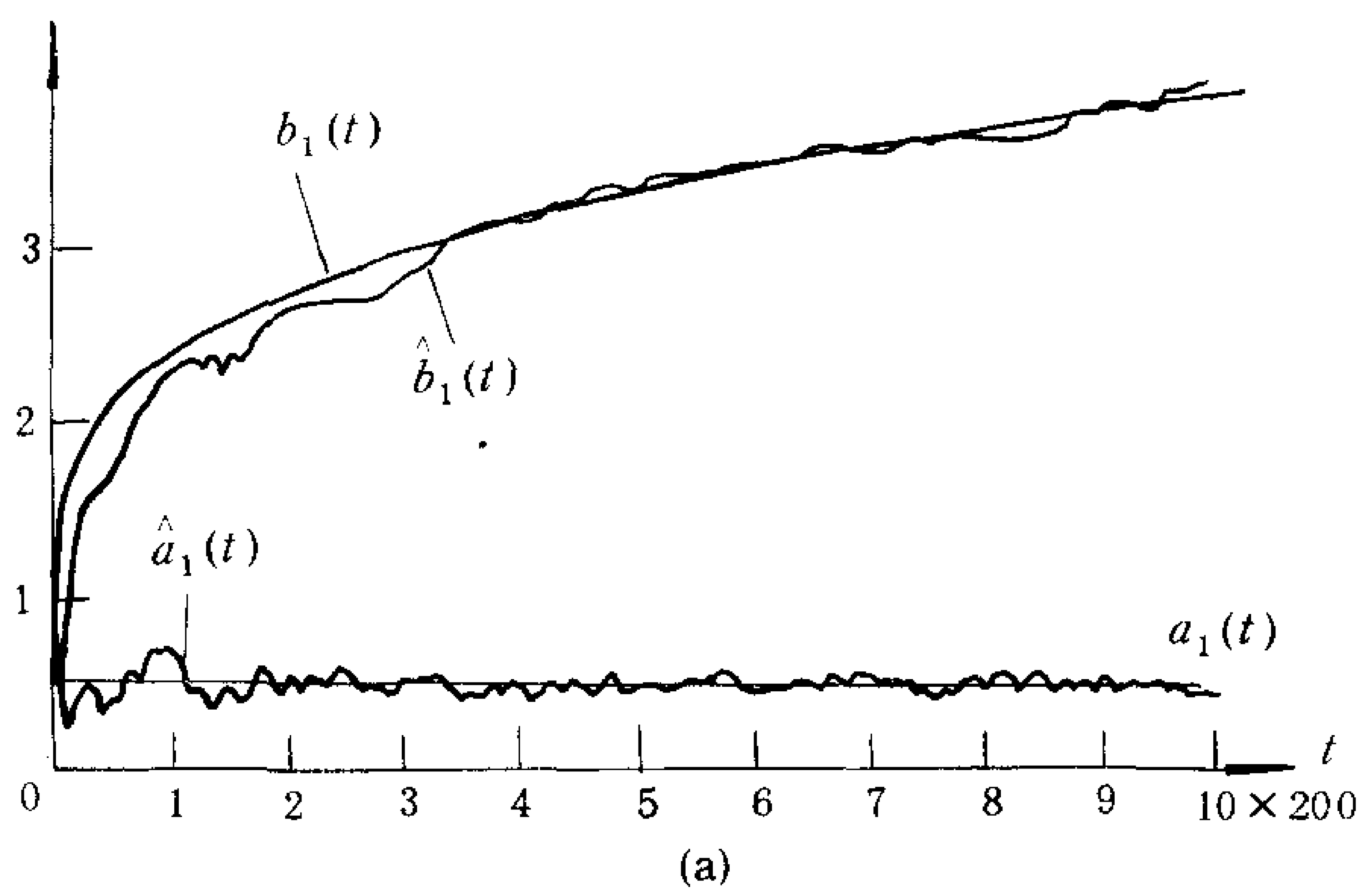

图 2(a)　例 2 的参数及其估计变化曲线

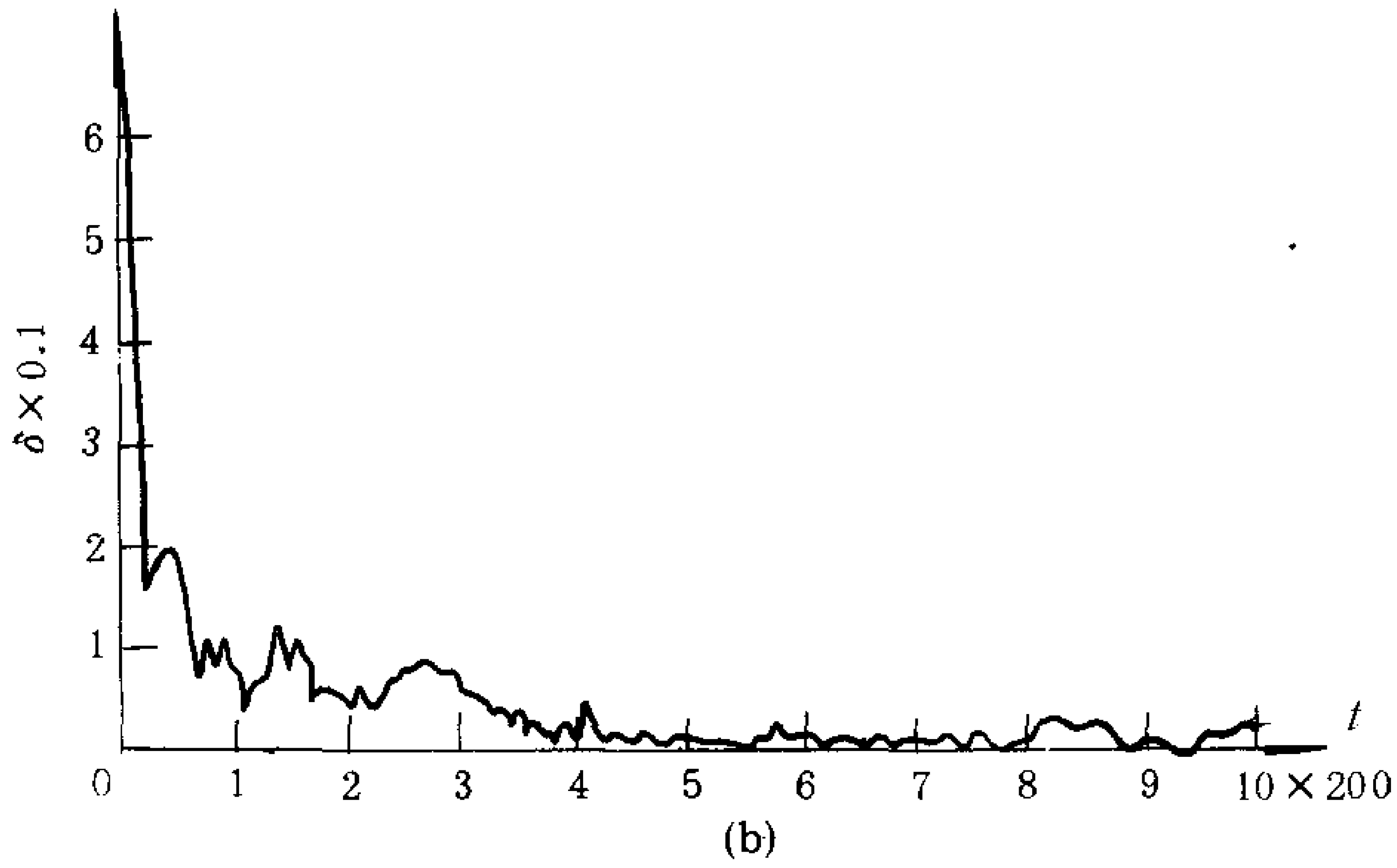

图 2(b)　例 2 的 δ 随 t 变化曲线

4 结语

多新息辨识算法有以下特点:

1) 可以通过适当选择 q 值来减小参数估计误差，提高算法的辨识精度.

2) 当 $t_s^*<p$ 时，多新息算法充分(重复)利用了系统的数据信息，在一定程度上可以提高算法的收敛速度.

3) 多新息算法可以抑制坏数据对参数估计的影响，计算量也较最小二乘法和卡尔曼滤波估计算法小.

4) 多新息算法的收敛性分析中，对噪声没有平稳性假设，但是定理1的结果说明数据的平稳性有利于提高辨识精度.

5) 多新息算法可以作为系统辨识和参数估计的基本方法之一，它与最小二乘法，卡尔曼滤波算法和随机逼近算法一样，可以用于各种模型的参数估计.

6) 多新息算法是单新息的投影算法的推广.

参 考 文 献

[1] 方崇智，肖德云. 过程辨识. 北京: 清华大学出版社，1988.

[2] 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.

[3] Goodwin G C, Sin K S. Adaptive filtering prediction and control. Prentice-hall, INC., Englewood Cliffs, New Jersey, 1984.

MULTI-INNOVATION IDENTIFICATION METHOD FOR TIME-VARYING SYSTEMS

DING FENG　XIE XINMIN　FANG CHONGZHI

(*Department of Automation, Tsinghua University, Beijing* 100084)

ABSTRACT

In this paper, the single innovation modification technique of estimating time-invariant parameters is extended, and the multi-innovation identification method is presented. This method may overcome the effect of bad data on the parameter estimation. It has stronger robustness and may track time-varying parameters. Its computational burden is less than forgetting factor least squares algorithm and Kalman filter algorithm. The simulation results indicate that the multi-innovation algorithm works quite well.

Key words: Parameter estimation, time-varying system, convergence, multi-innovation technique.