

多新息随机梯度辨识方法

丁 锋, 萧德云, 丁 韬
(清华大学 自动化系, 北京 100084)

摘要: 多新息随机梯度辨识方法是系统辨识和参数估计的一种基本方法. 该方法由于采用了间断迭代, 因此可以克服坏数据对参数估计的影响, 且具有较强的鲁棒性, 又可以跟踪时变参数. 作者从理论上给出了多新息随机梯度辨识方法的推导过程, 同时列出多新息随机梯度辨识方法的各种变形. 数字仿真实验表明多新息随机梯度辨识方法具有良好的性能.

关键词: 辨识; 参数估计; 多新息辨识; 最小均方算法

中图分类号: TP273 **文献标识码:** A

Multi-innovation stochastic gradient identification method

DING Feng, XIAO De-yun, DING Tao

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

Abstract: Multi-innovation stochastic gradient identification algorithm is one of the basic methods in the area of system identification and parameter estimation. It can overcome the effect of bad data on parameter estimation, and has strong robustness, and can track time-varying parameters because the discontinuous recursive computation was applied. By using a dimension search, the multi-innovation stochastic gradient identification was derived via minimizing the criteria, and some derivation algorithms from the multi-innovation stochastic gradient identification were given. The numeric simulation experiments indicated that the proposed algorithm has good performance.

Key words: identification; parameter estimation; multi-innovation identification; least mean square algorithm

1 引言(Introduction)

辨识领域中具有挑战性的研究课题有: 1) 提出新的辨识方法^[1~4]; 2) 辨识方法的(有界)收敛性、收敛速率、估计误差上界的研究^[5~7]; 3) 辨识试验中输入信号的设计等^[8]. 第 1 个问题是辨识的基石; 第 2 个是辨识中的理论问题, 受到了较多的关注, 已有的研究成果说明辨识算法收敛性的前提只要求输入信号是持续激励的, 除此之外对输入信号无特别的限制; 第 3 个是辨识实践中输入信号的选择问题. 近年来, 丁锋等人提出了辅助模型辨识方法^[1,2]、递阶辨识方法^[3]、多新息辨识方法^[4]等. 与最小二乘法、卡尔曼滤波算法、最小均方算法一样, 这些方法可用于解决多种模型的参数估计、自适应滤波、自适应信号处理和自适应控制等问题.

丁锋等人提出的多新息辨识方法^[4]是受文^[9]算法间断迭代的启发, 直接给出了多新息辨识算法

的数学表达式. 由于采用了间断迭代, 可以克服坏数据对参数估计的影响. 本文的目标是利用系统的输入输出数据, 从理论上详细推导多新息辨识算法, 对系统的未知参数进行实时估计, 并给出多新息辨识算法的一些变形.

2 系统描述(System description)

考虑下列线性参数模型的辨识问题

$$y(t) = \varphi^T(t)\theta + v(t). \quad (1)$$

其中, $y(t) \in \mathbb{R}^1$ 为系统输出, $\theta \in \mathbb{R}^n$ 为待辨识的未知参数向量, $\varphi(t) \in \mathbb{R}^n$ 为信息向量, $v(t) \in \mathbb{R}^1$ 为零均值随机噪声.

式(1)称为线性参数模型的辨识模型或辨识表达式. 线性参数模型(1)可以是非线性系统. 对于受控 AR(即 CAR)模型描述的系统

$$A(z)y(t) = B(z)u(t) + v(t). \quad (2)$$

其中, $u(t) \in \mathbb{R}^1$ 为系统输入, $A(z)$ 和 $B(z)$ 均为单

位后移算子 z^{-1} 的多项式,且

$$A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_{n_a} z^{-n_a},$$

$$B(z) = b_1 z^{-1} + b_2 z^{-2} + \dots + b_{n_b} z^{-n_b}.$$

令

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

$$n \triangleq n_a + n_b,$$

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

式(2)就可化为式(1)的形式.

3 多新息辨识算法 (Multi-innovation identification algorithm)

与投影算法的推导一样^[10-12],多新息辨识算法可以用负梯度搜索得到,也可通过求解一个优化问题得到,下面给出推导多新息辨识算法的三种方法.

考虑 $t = t - p + 1$ 到 $t = t$ 时共 p 组数据,令

$$Y(p, t) = \begin{bmatrix} y(t) \\ y(t-1) \\ \vdots \\ y(t-p+1) \end{bmatrix} \in \mathbb{R}^p,$$

$$V(p, t) = \begin{bmatrix} v(t) \\ v(t-1) \\ \vdots \\ v(t-p+1) \end{bmatrix} \in \mathbb{R}^p,$$

$$\Phi(p, t) = [\varphi(t), \varphi(t-1), \dots, \varphi(t-p+1)] \in \mathbb{R}^{n \times p}.$$

由式(1)可得到矩阵方程

$$Y(p, t) = \Phi^T(p, t)\theta + V(p, t). \quad (3)$$

定义正整数序列 $\{t_s, s = 0, 1, 2, \dots\}$ 满足

$$0 = t_0 < t_1 < t_2 < \dots,$$

$$1 \leq t_s^* = t_s - t_{s-1} < \infty.$$

用 t_s 代替式(3)中 t 可得

$$Y(p, t_s) = \Phi^T(p, t_s)\theta + V(p, t_s). \quad (4)$$

提出时变或时不变系统辨识方法通常有两种途径,其一是根据某种优化准则或原理推导出来,例如,最小二乘法和本文的多新息梯度算法;其二是采用某种近似、或进行类比直接给出辨识算法,例如文[4]给出的多新息辨识方法,这样提出的方法是无法用严密的数学语言推导出来的,倘若能证明这些方法的收敛性,或仿真结果可行,或实际应用成功,那么不管用什么途径提出的方法都是有效的.下面利用最速下降法或梯度法近似推导多新息辨识方法.

1) 最速下降法或梯度法.

目标函数为

$$J(\theta) = \frac{1}{2} \| V(p, t_s) \|^2 =$$

$$\frac{1}{2} V^T(p, t_s) V(p, t_s) = \frac{1}{2} \| Y(p, t_s) - \Phi^T(p, t_s)\theta \|^2. \quad (5)$$

式中 $\| X \|^2 = \text{tr} [XX^T]$ 表示矩阵 X 的范数.假设步长为 μ_i , 则梯度迭代算法可表示为

$$\begin{aligned} \hat{\theta}(t_s) &= \hat{\theta}(t_{s-1}) - \mu_i \text{grad} J(\hat{\theta}(t_{s-1})) = \\ &= \hat{\theta}(t_{s-1}) + \mu_i \Phi(p, t_s) [Y(p, t_s) - \Phi^T(p, t_s)\hat{\theta}(t_{s-1})]. \end{aligned} \quad (6)$$

将上式代入式(5),可求得近似最佳步长为

$$\mu_i = \frac{1}{\| \Phi(p, t_s) \|^2}.$$

将上式代入式(6),即得变迭代间隔多新息投影辨识算法(V-MIP):

$$\begin{aligned} \hat{\theta}(t_s) &= \\ \hat{\theta}(t_{s-1}) &+ \frac{\Phi(p, t_s)}{\| \Phi(p, t_s) \|^2} [Y(p, t_s) - \Phi^T(p, t_s)\hat{\theta}(t_{s-1})], \end{aligned} \quad (7)$$

$$\hat{\theta}(t) = \hat{\theta}(t_{s-1}), t \in T_s = \{t_{s-1}, t_{s-1}+1, \dots, t_s-1\}, \quad (8)$$

$$\Phi(p, t_s) = [\varphi(t_s), \varphi(t_s-1), \dots, \varphi(t_s-p+1)] \in \mathbb{R}^{n \times p}, \quad (9)$$

$$Y(p, t_s) = [y(t_s), y(t_s-1), \dots, y(t_s-p+1)]^T \in \mathbb{R}^p. \quad (10)$$

其中, p 为新息长度, $\hat{\theta}(t)$ 为 θ 的估计, $\hat{\theta}(0)$ 取很小实向量.

2) 准则函数法(一).

目标函数为

$$\begin{aligned} J(\theta) &= \| V(p, t_s) \|^2 = V^T(p, t_s) V(p, t_s) = \\ &= \| Y(p, t_s) - \Phi^T(p, t_s)\theta \|^2. \end{aligned} \quad (11)$$

假设 $\theta = \hat{\theta}(t_{s-1})$ 使 $J(\theta) = \min$, 令

$$\theta = \hat{\theta}(t_s) = \hat{\theta}(t_{s-1}) + \Phi(p, t_s)\mu_i, \quad (12)$$

那么问题就转化为求 μ_i , 把式(12)代入式(11),可得

$$\begin{aligned} J(\hat{\theta}(t_s)) &= \\ &= \| Y(p, t_s) - \Phi^T(p, t_s)(\hat{\theta}(t_{s-1}) + \Phi(p, t_s)\mu_i) \|^2. \end{aligned}$$

令 $J(\hat{\theta}(t_s))$ 关于 μ_i 的一阶偏导数为零,即

$$\frac{\partial J(\hat{\theta}(t_s))}{\partial \mu_i} = 0,$$

则有

$$\begin{aligned} Y(p, t_s) - \Phi^T(p, t_s)\hat{\theta}(t_{s-1}) - \\ \Phi^T(p, t_s)\Phi(p, t_s)\mu_i = 0, \end{aligned}$$

可近似解得

$$\mu_i = \frac{Y(p, t_s) - \Phi^T(p, t_s)\hat{\theta}(t_{s-1})}{\| \Phi(p, t_s) \|^2}.$$

将上式代入式(12),即得变迭代间隔多新息投影辨识算法(7)~(10).

3) 准则函数法(二).

仿照确定性系统投影算法的推导方法^[10-12],也可导出多新息辨识方法,过程如下:

由于 $\hat{\theta}(t_s)$ 是参数估计,则 $\Phi^T(p, t_s)\hat{\theta}(t_s)$ 是输出向量的预报,认为 $Y(p, t_s) = \Phi^T(p, t_s)\hat{\theta}(t_s)$. 给定 $\hat{\theta}(t_{s-1})$, $\Phi(p, t_s)$ 和 $Y(p, t_s)$,当约束条件为

$$Y(p, t_s) = \Phi^T(p, t_s)\hat{\theta}(t_s)$$

时,使准则函数

$$J(\hat{\theta}(t_s)) = \frac{1}{2} \|\hat{\theta}(t_s) - \hat{\theta}(t_{s-1})\|^2$$

极小.

证 证明的方法类似于文献[11,12](略).

在上面的推导过程中,都假设 $\|\Phi(p, t_s)\|^2 \neq 0$. 否则,为了防止式(7)右边第二项分母为零,解决的方法之一是:当 $\|\Phi(p, t_s)\|^2 = 0$ 时,令 $\hat{\theta}(t_s) = \hat{\theta}(t_{s-1})$,或者在式(7)右边第二项分母上加上一个正常数 $c > 0$. 也可将式(7)修改为

$$\hat{\theta}(t_s) = \hat{\theta}(t_{s-1}) + \frac{\Phi(p, t_s)}{1 + \|\Phi(p, t_s)\|^2} [Y(p, t_s) - \Phi^T(p, t_s)\hat{\theta}(t_{s-1})].$$

当持续激励条件满足时,对于确定性系统,投影算法和遗忘因子最小二乘法给出的参数估计以指数速度收敛于真参数,而最小二乘法的收敛速度为 $(1/t)$. 对于随机系统,投影算法和遗忘因子最小二乘法给出的参数估计很快收敛到真参数附近波动,参数估计误差方差不会趋于零. 而无论是确定性系统还是随机系统,最小二乘法的收敛速度仍为 $(1/t)$,参数估计误差方差趋于零. 多新息辨识算法的性能类似于投影算法和遗忘因子最小二乘法.

4 多新息辨识算法的几种形式(Some special forms of multi-innovation identification)

比较投影算法和随机梯度算法的区别^[10-12],容易将变迭代间隔多新息投影辨识算法(7)~(10)推广为变迭代间隔多新息广义投影辨识算法(V-MIGP):

$$\begin{aligned} \hat{\theta}(t_s) &= \\ \hat{\theta}(t_{s-1}) &+ \frac{\Phi(p, t_s)}{r(q, t_s)} [Y(p, t_s) - \Phi^T(p, t_s)\hat{\theta}(t_{s-1})], \\ \hat{\theta}(t) &= \hat{\theta}(t_s), \\ t \in T_s &= \{t_s, t_s + 1, \dots, t_{s+1} - 1\}, \\ r(q, t_s) &= \text{tr} [\Phi(q, t_s)\Phi^T(q, t_s)], \quad q \geq p, \\ \Phi(p, t_s) &= \end{aligned}$$

$$[\varphi(t_s), \varphi(t_s-1), \dots, \varphi(t_s-p+1)] \in \mathbb{R}^{n \times p}.$$

其中, q 为记忆长度或数据窗长度.进一步可推广为变迭代间隔多新息随机梯度辨识算法(V-MISG):

$$\begin{aligned} \hat{\theta}(t_s) &= \\ \hat{\theta}(t_{s-1}) &+ \frac{\Phi(p, t_s)}{r(t_s)} [Y(p, t_s) - \Phi^T(p, t_s)\hat{\theta}(t_{s-1})], \\ \hat{\theta}(t) &= \hat{\theta}(t_s), \\ t \in T_s &= \{t_s, t_s + 1, \dots, t_{s+1} - 1\}, \\ r(t_s) &= \|\Phi(s, t_s)\|^2 = \sum_{i=0}^{s-1} \|\varphi(t_s - i)\|^2, \\ \Phi(p, t_s) &= [\varphi(t_s), \varphi(t_s-1), \dots, \varphi(t_s-p+1)], \\ Y(p, t_s) &= [y(t_s), y(t_s-1), \dots, y(t_s-p+1)]^T. \end{aligned}$$

① 当迭代间隔 $t_s^* = 1$ 时,从V-MIGP算法得到多新息广义投影辨识算法(MIGP):

$$\begin{aligned} \hat{\theta}(t) &= \\ \hat{\theta}(t-1) &+ \frac{\Phi(p, t)}{r(q, t)} [Y(p, t) - \Phi^T(p, t)\hat{\theta}(t-1)], \\ r(q, t) &= \text{tr} [\Phi(q, t)\Phi^T(q, t)], \quad q \geq p, \\ \Phi(p, t) &= [\varphi(t), \varphi(t-1), \dots, \varphi(t-p+1)], \\ Y(p, t) &= [y(t), y(t-1), \dots, y(t-p+1)]^T. \end{aligned}$$

当记忆长度 $q = p$ 时,MIGP算法退化为多新息投影辨识算法(MIP).

② 当迭代间隔 $t_s^* = 1$ 时,从V-MISG算法得到多新息随机梯度辨识算法(MISG):

$$\begin{aligned} \hat{\theta}(t) &= \\ \hat{\theta}(t-1) &+ \frac{\Phi(p, t)}{r(t)} [Y(p, t) - \Phi^T(p, t)\hat{\theta}(t-1)], \\ r(t) &= \sum_{i=0}^{t-1} \|\varphi(t-i)\|^2 = \\ &r(t-1) + \|\varphi(t)\|^2, \quad r(0) = 1, \\ \Phi(p, t) &= [\varphi(t), \varphi(t-1), \dots, \varphi(t-p+1)], \\ Y(p, t) &= [y(t), y(t-1), \dots, y(t-p+1)]^T. \end{aligned}$$

③ 当新息长度 $p = 1$ 时,从V-MIGP算法得到变迭代间隔广义投影辨识算法(V-GP):

$$\begin{aligned} \hat{\theta}(t_s) &= \hat{\theta}(t_{s-1}) + \frac{\varphi(t_s)}{r(q, t_s)} [y(t_s) - \varphi^T(t_s)\hat{\theta}(t_{s-1})], \\ \hat{\theta}(t) &= \hat{\theta}(t_s), \quad t \in T_s = \{t_s, t_s + 1, \dots, t_{s+1} - 1\}, \\ r(q, t_s) &= \sum_{i=0}^{q-1} \|\varphi(t_s - i)\|^2, \quad q \geq 1. \end{aligned}$$

当记忆长度 $q = p = 1$ 时,V-GP算法退化为变迭代间隔投影辨识算法(V-P).

④ 当新息长度 $p = 1$ 时,从V-MISG算法得到变迭代间隔随机梯度辨识算法(V-SG):

$$\begin{aligned} \hat{\theta}(t_s) &= \hat{\theta}(t_{s-1}) + \frac{\varphi(t_s)}{r(t_s)} [y(t_s) - \varphi^T(t_s) \hat{\theta}(t_{s-1})], \\ \hat{\theta}(t) &= \hat{\theta}(t_s), t \in T_s = \{t_s, t_s + 1, \dots, t_{s+1} - 1\}, \\ r(t_s) &= \sum_{i=0}^{s-1} \|\varphi(t_s - i)\|^2, q \geq 1. \end{aligned}$$

⑤ 当迭代间隔 $t_s^* = k, p = 1$ 时,从 V-MIGP 算法得到等迭代间隔广义投影算法(E-GP):

$$\begin{aligned} \hat{\theta}(t) &= \hat{\theta}(t-k) + \frac{\varphi(t)}{r(q,t)} [y(t) - \varphi^T(t) \hat{\theta}(t-k)], \\ r(q,t) &= \sum_{i=0}^{q-1} \|\varphi(t-i)\|^2, q \geq p = 1, \end{aligned}$$

或

$$\begin{aligned} \hat{\theta}(ik) &= \hat{\theta}((i-1)k) + \frac{\varphi(ik)}{r(q,ik)} [y(ik) - \varphi^T(ik) \hat{\theta}((i-1)k)], \\ \hat{\theta}(t) &= \hat{\theta}((i-1)k), \\ t \in T_i &= \{(i-1)k, (i-1)k + 1, \dots, ik - 1\}, \\ r(q,ik) &= \sum_{j=0}^{q-1} \|\varphi(ik-j)\|^2, q \geq 1. \end{aligned}$$

当记忆长度 $q = p = 1$ 时,E-GP 算法退化为等迭代间隔投影辨识算法(E-P).

⑥ 当迭代间隔 $t_s^* = k, p = 1$ 时,从 V-MISG 算法得到等迭代间隔随机梯度辨识算法(E-SG):

$$\begin{aligned} \hat{\theta}(t) &= \hat{\theta}(t-k) + \frac{\varphi(t)}{r(t)} [y(t) - \varphi^T(t) \hat{\theta}(t-k)], \\ r(t) &= \sum_{i=0}^{t-1} \|\varphi(t-i)\|^2 = r(t-1) + \|\varphi(t)\|^2, \\ r(0) &= r_0 > 0. \end{aligned}$$

⑦ 当迭代间隔 $t_s^* = k(s = 1, 2, \dots)$ 时,从 V-MIGP 算法得到等迭代间隔多新息广义投影辨识算法(E-MIGP):

$$\begin{aligned} \hat{\theta}(t) &= \hat{\theta}(t-k) + \frac{\Phi(p,t)}{r(q,t)} [Y(p,t) - \Phi^T(p,t) \hat{\theta}(t-k)], \\ r(q,t) &= \text{tr} [\Phi(q,t) \Phi^T(q,t)] = \\ &= \sum_{i=0}^{q-1} \|\varphi(t-i)\|^2, q \geq p, \end{aligned}$$

$$\begin{aligned} \Phi(p,t) &= [\varphi(t), \varphi(t-1), \dots, \varphi(t-p+1)] \in \mathbb{R}^{n \times p}, \\ Y(p,t) &= [y(t), y(t-1), \dots, y(t-p+1)]^T \in \mathbb{R}^p, \end{aligned}$$

或

$$\begin{aligned} \hat{\theta}(ik) &= \hat{\theta}((i-1)k) + \frac{\Phi(p,ik)}{r(q,ik)} [Y(p,ik) - \Phi^T(p,ik) \hat{\theta}((i-1)k)], \\ \hat{\theta}(t) &= \hat{\theta}((i-1)k), \\ t \in T_i &= \{(i-1)k, (i-1)k + 1, \dots, ik - 1\}, \\ \Phi(p,ik) &= [\varphi(ik), \varphi(ik-1), \dots, \varphi(ik-p+1)], \end{aligned}$$

$$\begin{aligned} r(q,ik) &= \text{tr} [\Phi(q,ik) \Phi^T(q,ik)], q \geq p, \\ Y(p,ik) &= [y(ik), y(ik-1), \dots, y(ik-p+1)]^T. \end{aligned}$$

当记忆长度 $q = p$ 时,E-MIGP 算法退化为等迭代间隔多新息投影辨识算法(E-MP).

⑧ 当迭代间隔 $t_s^* = k$ 时,从 V-MISG 算法得到等迭代间隔多新息随机梯度辨识算法(E-MISG):

$$\begin{aligned} \hat{\theta}(t) &= \\ \hat{\theta}(t-k) &+ \frac{\Phi(p,t)}{r(t)} [Y(p,t) - \Phi^T(p,t) \hat{\theta}(t-k)], \\ r(t) &= r(t-1) + \|\varphi(t)\|^2, r(0) = 1, \\ \Phi(p,t) &= [\varphi(t), \varphi(t-1), \dots, \varphi(t-p+1)], \\ Y(p,t) &= [y(t), y(t-1), \dots, y(t-p+1)]^T. \end{aligned}$$

⑨ 在 MISG 算法中引入遗忘因子 λ , 得到多新息遗忘梯度辨识算法(MIFG):

$$\begin{aligned} \hat{\theta}(t) &= \\ \hat{\theta}(t-1) &+ \frac{\Phi(p,t)}{r(t)} [Y(p,t) - \Phi^T(p,t) \hat{\theta}(t-1)], \\ r(t) &= \lambda r(t-1) + \|\varphi(t)\|^2, \\ 0 < \lambda &\leq 1, r(0) = 1, \\ \Phi(p,t) &= [\varphi(t), \varphi(t-1), \dots, \varphi(t-p+1)], \\ Y(p,t) &= [y(t), y(t-1), \dots, y(t-p+1)]^T. \end{aligned}$$

⑩ 在 E-MISG 算法中引入遗忘因子 λ , 得到等迭代间隔多新息遗忘梯度辨识算法(E-MIFG):

$$\begin{aligned} \hat{\theta}(t) &= \\ \hat{\theta}(t-k) &+ \frac{\Phi(p,t)}{r(t)} [Y(p,t) - \Phi^T(p,t) \hat{\theta}(t-k)], \\ r(t) &= \lambda r(t-1) + \|\varphi(t)\|^2, \\ 0 \leq \lambda &\leq 1, r(0) = r_0 > 0, \\ \Phi(p,t) &= [\varphi(t), \varphi(t-1), \dots, \varphi(t-p+1)], \\ Y(p,t) &= [y(t), y(t-1), \dots, y(t-p+1)]^T. \end{aligned}$$

此外,当迭代间隔 $t_s^* = 1$,新息长度 $p = 1$ 时,V-MIGP 算法退化为广义投影辨识算法(GP),V-MISG 算法退化为随机梯度辨识算法(SG).

5 仿真试验 (Simulation experiment)

例 1 考虑下列系统

$$\begin{aligned} A(z)y(t) &= B(z)u(t) + v(t), \\ A(z) &= 1 + a_1 z^{-1} + a_2 z^{-2} = \\ &= 1 - 0.412z^{-1} + 0.309z^{-2}, \\ B(z) &= b_1 z^{-1} + b_2 z^{-2} = 1.968z^{-1} + 1.963z^{-2}, \\ \theta(t) &= \theta = [a_1, a_2, b_1, b_2]^T, \\ \varphi(t) &= [-y(t-1), -y(t-2), u(t-1), u(t-2)]^T. \end{aligned}$$

$\{u(t)\}$ 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 $\sigma_v^2 = 1.0^2$ 不相关随

机变量序列.取 $t_s^* = 5, q = 16$, 利用多新息广义投影算法估计这个系统的参数, 不同新息长度 $p = 1, 2, 3, 5$ 下的参数估计误差 δ 随迭代步数 s 的变化曲线如图 1 ($t_s^* = 5, q = 16; \sigma_v^2 = 1.0^2, \delta_{ns} = 31.38\%$) 所示, 参数估计相对误差范数定义为 $\delta = \|\hat{\theta}(t_s) - \theta\| / \|\theta\|$. 从图 1 可知, δ 随迭代步数 s 增大 (或数据长度加大) 而减小, 随新息长度 p 的增大而减小. 由此可见, 多新息辨识方法, 由于增加了新息长度, 加快了参数估计收敛速度, 减小了算法对噪声的敏感性, 增强了算法的鲁棒性, 故多新息算法估计系统参数是有效的.

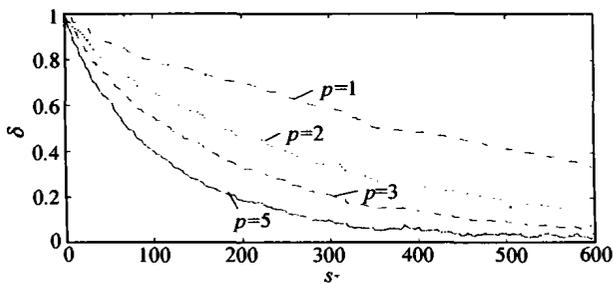


图 1 不同新息长度 p 下例 1 估计误差 δ 随迭代步数 s 变化曲线

Fig. 1 δ vs s with different p

6 结论 (Conclusions)

为了提高算法跟踪时变参数的能力, 或克服数据饱和现象, 通常在算法中引入一些参量, 如遗忘因子最小二乘法中的遗忘因子、有限数据窗辨识算法中的数据窗长度、遗忘梯度算法中的遗忘因子、加权最小二乘法中的加权矩阵, 以及多新息辨识算法中的新息长度 p 、迭代间隔 t_s^* 、记忆长度 q 、遗忘因子 λ 等. 研究这些算法的估计误差上界, 以及如何选择这些参量, 使参数估计误差上界最小是辨识领域的又一研究课题.

参考文献 (References):

- [1] 丁锋, 谢新民. 多变量系统的辅助模型辨识算法 [J]. 清华大学学报 (自然科学版), 1992, 32(4): 100 - 106.
(DING Feng, XIE Xinmin. Auxiliary model identification algorithm for multivariable systems [J]. *J of Tsinghua University*, 1992, 32(4): 100 - 106.)
- [2] 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, 1997, 14(2): 192 - 200.
(DING Feng. Convergence analysis of auxiliary model identification algorithm for multivariable systems [J]. *Control Theory & Applications*, 1997, 14(2): 192 - 200.)
- [3] 丁锋, 杨家本. 大系统的递阶辨识 [J]. 自动化学报, 1999, 25(5): 647 - 654.
(DING Feng, YANG Jia-ben. Hierarchical identification for large-

scale systems [J]. *Acta Automatica Sinica*, 1999, 25(5): 647 - 654.)

- [4] 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, 1996, 22(1): 85 - 91.
(DING Feng, XIE Xinmin, FANG Chongzhi. Multi-innovation identification method for time-varying systems [J]. *Acta Automatica Sinica*, 1996, 22(1): 85 - 91.)
- [5] 丁锋. 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, 1997, 14(1): 90 - 95.
(DING Feng. Martingale hyperconvergence theorem and the convergence of forgetting factor least squares algorithm [J]. *Control Theory & Applications*, 1997, 14(1): 90 - 95.)
- [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 [J]. 控制理论与应用, 1999, 16(4): 569 - 572.
(DING Feng, YANG Jiaben. Comments on martingale hyperconvergence theorem and the convergence of forgetting factor least squares algorithm [J]. *Control Theory & Applications*, 1999, 16(4): 569 - 572.)
- [7] 丁锋, 谢新民, 方崇智. 辨识时变系统遗忘因子算法的收敛性 [J]. 控制理论与应用, 1994, 11(5): 634 - 638.
(DING Feng, XIE Xinmin, FANG Chongzhi. Convergence of the forgetting factor algorithm for identifying time-varying systems [J]. *Control Theory & Applications*, 1994, 11(5): 634 - 638.)
- [8] 方崇智, 萧德云. 过程辨识 [M]. 北京: 清华大学出版社, 1988.
(FANG Chongzhi, XIAO Deyun. *Process Identification* [M]. Beijing: Tsinghua University Press, 1988.)
- [9] MOUSTAFA K A F. Identification of stochastic time-varying systems [J]. *IEE Proc, Part D*, 1983, 130(4): 137 - 142.
- [10] GOODWIN G C, SIN K S. Adaptive filtering prediction and control [J]. Englewood Cliffs, New Jersey: Prentice-hall, Inc., 1984.
- [11] 冯纯伯, 史维. 自适应控制 [M]. 北京: 电子工业出版社, 1986.
(FENG Chunbo, SHI Wei. *Adaptive Control* [M]. Beijing: Electronics Industry Press, 1986.)
- [12] 谢新民, 丁锋. 自适应控制系统 [M]. 北京: 清华大学出版社, 2002.
(XIE Xinmin, DING Feng. *Adaptive Control Systems* [M]. Beijing: Tsinghua University Press, 2002.)

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 (优秀博士论文), 1997年任清华大学自动化系副教授, 现在加拿大 Alberta 大学做访问学者. 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. E-mail: 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. E-mail: 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.