

MISO 系统辅助模型多新息增广随机梯度算法

赵学良, 肖永松

(江南大学 控制科学与工程研究中心, 江苏 无锡 214122)

摘要:对于有色噪声干扰的输出误差多输入单输出(MISO)系统,常规的递推最小二乘辨识方法给出的参数估计是有偏的。为了提高随机梯度辨识方法的收敛精度和速度,用辅助模型的输出代替辨识模型信息向量中的未知不可测量变量,推导出其辅助模型增广随机梯度辨识算法;再引入新息长度扩展标量新息为新息向量,提出了基于辅助模型的MISO系统多新息增广随机梯度辨识算法。所得算法在每一次的迭代中不仅使用了当前数据和新息,而且使用了过去数据和新息,提高了参数估计精度和收敛速度。仿真例子验证了算法的有效性。

关键词:辅助模型;随机梯度;多新息辨识;MISO系统;输出误差模型

中图分类号:TP301.6

文献标识码:A

文章编号:1673-629X(2011)02-0039-04

Multi-innovation Extended Stochastic Gradient Algorithm Based on an Auxiliary Model for MISO Systems

ZHAO Xue-liang, XIAO Yong-song

(Control Science and Engineering Research Center, Jiangnan University, Wuxi 214122, China)

Abstract:For multiple-input single-output (MISO) output-error systems, estimates from the conventional recursive least squares parameter identification method are biased. In order to enhance accuracy and speed of convergence for stochastic gradient identification algorithm, an auxiliary model based multi-innovation extended stochastic gradient (AM-MI-ESG) algorithm is presented by replacing the unknown unmeasurable variables in the information vector with the outputs of the auxiliary model, and introducing the innovation length and expanding the scalar innovation to an innovation vector. The AM-MI-ESG algorithm uses not only the current data and innovation but also the past data and innovation at each iteration, thus the parameter estimation accuracy and convergence rate can be improved. The simulation results show that the proposed algorithm is effective.

Key words:auxiliary model; stochastic gradient; multi-innovation identification; MISO systems; output-error models

0 引言

对于有色噪声干扰下的输出误差系统的辨识问题^[1],常规的递推最小二乘辨识方法给出的参数估计是有偏的,因为它不是方程误差模型(ARX模型)。虽然偏差补偿递推最小二乘辨识方法^[2,3]和基于辅助模型的递推增广最小二乘辨识方法^[4],可以无偏估计这类系统参数,但它们都是针对单输入单输出系统提出来的,而且需要计算协方差矩阵,使得计算量增大。随机梯度算法因不涉及协方差矩阵而具有计算量小的优点,但收敛速度慢。

为了提高随机梯度辨识方法的收敛速度,可引入新息长度,从而导出多新息随机梯度算法^[5],多新息随机梯度算法就是在随机梯度算法与最小二乘算法的

收敛速度和计算量之间找到折中。文献[6]提出了基于辅助模型的Box-Jenkins模型多新息广义增广随机梯度算法,文献[7]将多新息辨识理论与前向神经网络相结合,提出了一种基于前向神经网络的多新息随机梯度算法。

实际系统中很多都可以抽象为多输入单输出系统,如通信系统^[8]、传感器系统^[9]等。模糊推理^[10]、人工神经网络^[11]、最小二乘支持向量机^[12]等已广泛应用于各类MISO系统的建模与辨识。

文中在文献[6,13]的基础上,借助辅助模型辨识思想,并结合多新息辨识理论,扩展标量新息到新息向量,推导出了基于辅助模型的多输入单输出系统多新息增广随机梯度辨识算法,并进行了仿真研究。提出的算法不仅克服了基于辅助模型的多新息最小二乘算法由于需要计算协方差阵而产生的计算量大的缺点,而且保留了基于辅助模型的随机梯度算法不需要矩阵求逆运算的优点,从而减少了计算量,提高了辨识精度。

收稿日期:2010-06-27;修回日期:2010-09-08

基金项目:国家自然科学基金(60574051)

作者简介:赵学良(1985-),男,江苏苏州人,硕士研究生,研究方向为系统建模与系统辨识。

1 系统描述与算法推导

考虑下列有色噪声干扰的输出误差多输入单输出系统:

$$y(t) = \sum_{i=1}^r \frac{B_i(z)}{A_i(z)} u_i(t) + D(z)v(t) \quad (1)$$

其中 $u_i(t) \in R^1 (i=1, 2, \dots, r)$ 为系统输入, $y(t) \in R^1$ 为系统输出, $v(t) \in R^1$ 为零均值随机白噪声, $A_i(z)$, $B_i(z)$ 和 $D(z)$ 均为单位后移算子 z^{-1} 的多项式 $[z^{-1}y(t) = y(t-1)]$, 且

$$A_i(z) := 1 + a_i(1)z^{-1} + a_i(2)z^{-2} + \dots + a_i(n_i)z^{-n_i}$$

$$B_i(z) := b_i(1)z^{-1} + b_i(2)z^{-2} + \dots + b_i(n_i)z^{-n_i}$$

$$D(z) := 1 + d_1z^{-1} + d_2z^{-2} + \dots + d_nz^{-n}$$

不妨假设阶次 n_i 和 n_d 已知, 且 $t \leq 0$ 时, $u_i(t) = 0$, $y(t) = 0$, $v(t) = 0$ 。文中的目标是: 利用系统的输入输出数据 $\{u_1(t), u_2(t), \dots, u_r(t), y(t)\}$ 来辨识系统 (1) 中未知参数 $a_i(j)$, $b_i(j)$ 和 d_i , 并进行仿真研究。

系统 (1) 的原理框图如图 1 所示, 中间变量 $x_i(t)$ 和 $x(t)$ 分别为虚拟子系统和整个系统的真实输出 (不可测), $w(t)$ 噪声模型的输出 (不可测), $y(t)$ 是 $x(t)$ 的含噪量测。

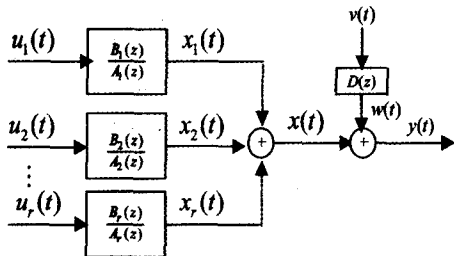


图1 有色噪声干扰的输出误差多输入单输出系统
定义中间变量:

$$x_i(t) := \frac{B_i(z)}{A_i(z)} u_i(t), \quad w(t) := D(z)v(t) \quad (2)$$

则式(1)可写为:

$$y(t) = \sum_{i=1}^r x_i(t) + w(t) = x(t) + w(t) \quad (3)$$

定义系统参数向量 θ , 虚拟子系统模型参数向量 θ_i , 系统模型参数向量 θ , 和噪声模型参数向量 θ_n 分别为:

$$\varphi(t) := [\theta_i^T, \theta_n^T]^T \in R^n, n := n_0 + n_d$$

$$\theta_i := [a_i(1), a_i(2), \dots, a_i(n_i), b_i(1), b_i(2), \dots, b_i(n_i)]^T \in R^{2n_i}$$

$$\theta := [\theta_1^T, \theta_2^T, \dots, \theta_r^T]^T \in R^{n_0}, n_0 := 2n_1 + 2n_2 + \dots + 2n_r$$

$$\theta_n := [d_1, d_2, \dots, d_n]^T \in R^{n_d}$$

定义系统信息向量:

$$\varphi(t) := [\varphi_i^T(t), \varphi_n^T(t)]^T \in R^n, n := n_0 + n_d$$

$$\varphi_i(t) := [-x_i(t-1), -x_i(t-2), \dots, -x_i(t-n_i), u_i(t-1), u_i(t-2), \dots, u_i(t-n_i)]^T \in R^{2n_i}$$

$$\varphi_n(t) := [\varphi_1^T(t), \varphi_2^T(t), \dots, \varphi_r^T(t)]^T \in R^{n_0}, n_0 := 2n_1 + 2n_2 + \dots + 2n_r$$

$$\varphi_n(t) := [v(t-1), v(t-2), \dots, v(t-n_d)]^T \in R^{n_d}$$

那么式(2)有如下向量形式:

$$x_i(t) = \varphi_i^T(t) \theta_i, \quad w(t) = \varphi_n^T(t) \theta_n + v(t) \quad (4)$$

把上两式代入式(3)可得:

$$y(t) = \sum_{i=1}^r \varphi_i^T(t) \theta_i + \varphi_n^T(t) \theta_n + v(t) = \varphi^T(t) \theta + v(t) \quad (5)$$

引入二次型准则函数:

$$J(\theta) = [\|y(t) - \varphi^T(t) \theta\|^2]$$

其中 $\|X\|^2 := \text{tr}[XX^T]$ 表示矩阵 X 的范数, 令 $\hat{\theta}(t)$ 表示 t 时刻参数向量 θ 的估计。根据梯度搜索原理, 极小化 $J(\theta)$ 可得估计 θ 的递推算法, 然而, 由于信息向量 $\varphi(t)$ 中含有未知中间变量 $x_i(t-i)$ 和可测噪声项 $v(t-i)$, 所以标准的递推算法不能实现, 这里的解决办法是利用辅助模型辨识思想, 分别为每个虚拟子系统建立一个辅助模型, 如图 2 所示, 其中

$$P_i(z) := \frac{\hat{B}_i(z)}{\hat{A}_i(z)} \text{ 为辅助模型的传递函数, } \hat{A}_i(z) \text{ 和 } \hat{B}_i(z) \text{ 是与 } A_i(z) \text{ 和 } B_i(z) \text{ 同阶次的多项式, } \hat{x}_i(t-i) \text{ 为辅助模型的输出。}$$

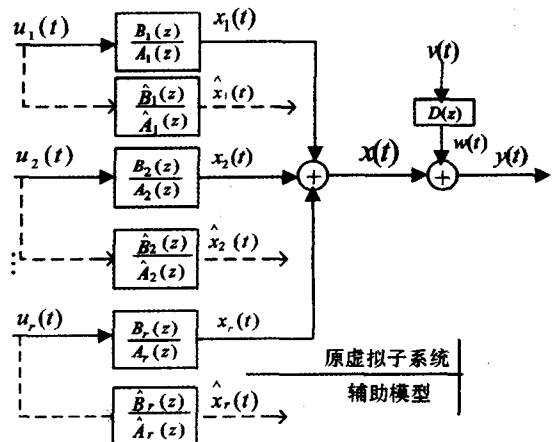


图2 带辅助模型的输出误差多输入单输出系统
信息向量 $\varphi(t)$ 里 $\varphi_i(t)$ 中未知项 $x_i(t-i)$ 用辅助模型的输出 $\hat{x}_i(t-i)$ 代替, 作了这种代替后的 $\varphi_i(t)$ 记作为:

$$\hat{\varphi}_i(t) = [\hat{x}_i(t-1), \hat{x}_i(t-2), \dots, \hat{x}_i(t-n_i), u_i(t-1), u_i(t-2), \dots, u_i(t-n_i)]^T$$

1), $u_i(t-2), \dots, u_i(t-n_i)]^T \in R^{2n_i}$

用 $\hat{\varphi}_i(t)$ 代替 $\varphi_i(t)$ 中未知的 $\varphi_i(t)$, 代替后的 $\varphi_i(t)$ 记作为:

$$\hat{\varphi}_i(t) = [\hat{\varphi}_i^T(t), \hat{\varphi}_i^{T_2}(t), \dots, \hat{\varphi}_i^{T_r}(t)]^T \in R^{n_i}$$

用估计残差 $\hat{v}(t-i)$ 代替不可测噪声项 $v(t-i)$, 代替后的 $\varphi_n(t)$ 记作为:

$$\hat{\varphi}_n(t) = [\hat{v}(t-1), \hat{v}(t-2), \dots, \hat{v}(t-n_d)]^T \in R^{n_d}$$

用 $\hat{\varphi}_i(t)$ 和 $\hat{\varphi}_n(t)$ 代替 $\varphi(t)$ 中未知的 $\varphi_i(t)$ 和 $\varphi_n(t)$, 代替后的 $\varphi(t)$ 记作为:

$$\hat{\varphi}(t) = [\hat{\varphi}_i^T(t), \hat{\varphi}_n^T(t)]^T \in R^n$$

由式(7)可得 $v(t) = y(t) - \varphi^T(t) \theta$, 式中未知的 $\varphi(t)$ 用其估计 $\hat{\varphi}(t)$ 代替, θ 用其估计 $\hat{\theta}(t)$ 代替, 那么残差 $\hat{v}(t)$ 可通过 $\hat{v}(t) = y(t) - \hat{\varphi}^T(t) \hat{\theta}(t)$ 计算, 照式(5), 每个虚拟子系统的辅助模型可取为 $\hat{x}_i(t) = \hat{\varphi}_i^T(t) \hat{\theta}_i(t)$, 其中 $\hat{\varphi}_i(t)$ 和 $\hat{\theta}_i(t)$ 分别是辅助模型的信息向量和参数向量。综合上述各式, 就可得到辨识参数向量 θ 的基于辅助模型的多新息增广随机梯度算法 (AM-ESG):

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\hat{\varphi}(t)}{r(t)} e(t)$$

$$e(t) = y(t) - \hat{\varphi}^T(t) \hat{\theta}(t-1)$$

$$r(t) = r(t-1) + \|\hat{\varphi}(t)\|^2, r(0) = 1$$

2 多新息辨识方法

AM-ESG 算法收敛速度慢 (参见后面的仿真例子), 为了提高算法收敛速度, 可以借助于多新息辨识理论, 将标量新息 $e(t) = y(t) - \hat{\varphi}^T(t) \hat{\theta}(t-1)$ 扩展为新息向量:

$$E(p, t) = [e(t), e(t-1), \dots, e(t-p+1)]^T \in R^p \quad (6)$$

其中正整数 p 表示新息长度, 且

$$e(t-i) = y(t-i) - \hat{\varphi}^T(t-i) \hat{\theta}(t-i-1) \in R^1 \quad (7)$$

一般认为, $t-1$ 时刻参数估计值 $\hat{\theta}(t-1)$ 比以前任何时刻的估计值 $\hat{\theta}(t-i)$ ($i=2, 3, \dots, p-1$) 都更接近于真参数, 所以新息向量可以写为:

$$E(p, t) =$$

$$\begin{bmatrix} y(t) - \hat{\varphi}^T(t) \hat{\theta}(t-1) \\ y(t-1) - \hat{\varphi}^T(t-1) \hat{\theta}(t-1) \\ \vdots \\ y(t-p+1) - \hat{\varphi}^T(t-p+1) \hat{\theta}(t-1) \end{bmatrix} \in R^p \quad (8)$$

定义堆积输出向量 $Y(p, t)$ 和堆积信息矩阵

$\hat{\Phi}(p, t)$ 分别为:

$$Y(p, t) := [y(t), y(t-1), \dots, y(t-p+1)]^T \in R^p \quad (9)$$

$$\hat{\Phi}(p, t) := [\hat{\varphi}(t), \hat{\varphi}(t-1), \dots, \hat{\varphi}(t-p+1)] \in R^{n \times p} \quad (10)$$

那么新息向量可表示为:

$$E(p, t) = Y(p, t) - \hat{\Phi}^T(p, t) \hat{\theta}(t-1) \quad (11)$$

因为 $E(1, t) = e(t)$, $Y(1, t) = y(t)$, $\hat{\Phi}(1, t) = \hat{\varphi}(t)$, 所以可以得到:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\hat{\Phi}^T(1, t)}{r(t)} [Y(1, t) - \hat{\Phi}^T(1, t) \hat{\theta}(t-1)] \quad (12)$$

把上式 $Y(1, t)$ 和 $\hat{\Phi}(1, t)$ 中的 1 换成 p , 就可得到辨识参数向量 θ 的基于辅助模型多新息增广随机梯度算法 (AM-MI-ESG):

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\hat{\Phi}^T(p, t)}{r(t)} E(p, t) \quad (13)$$

$$E(p, t) = Y(p, t) - \hat{\Phi}^T(p, t) \hat{\theta}(t-1) \quad (14)$$

$$r(t) = r(t-1) + \|\hat{\varphi}(t)\|^2, r(0) = 1 \quad (15)$$

$$Y(p, t) := [y(t), y(t-1), \dots, y(t-p+1)]^T \quad (16)$$

$$\hat{\Phi}(p, t) := [\hat{\varphi}(t), \hat{\varphi}(t-1), \dots, \hat{\varphi}(t-p+1)] \quad (17)$$

$$\hat{\varphi}(t) = [\hat{\varphi}_i^T(t), \hat{\varphi}_n^T(t)]^T, \hat{\theta}(t) = [\hat{\theta}_i^T(t), \hat{\theta}_n^T(t)]^T \quad (18)$$

$$\hat{\varphi}_i(t) = \begin{bmatrix} \hat{\varphi}_i(t) \\ \hat{\varphi}_i(t) \\ \vdots \\ \hat{\varphi}_i(t) \end{bmatrix}, \hat{\theta}_i(t) = \begin{bmatrix} \hat{\theta}_i(t) \\ \hat{\theta}_i(t) \\ \vdots \\ \hat{\theta}_i(t) \end{bmatrix} \quad (19)$$

$$\hat{\varphi}_i(t) = [\hat{x}_i(t-1), \hat{x}_i(t-2), \dots, \hat{x}_i(t-n_i), u_i(t-1), u_i(t-2), \dots, u_i(t-n_i)]^T \quad (20)$$

$$\hat{\varphi}_n(t) = [\hat{v}(t-1), \hat{v}(t-2), \dots, \hat{v}(t-n_d)]^T \quad (21)$$

$$\hat{x}_i(t) = \hat{\varphi}_i^T(t) \hat{\theta}_i(t) \quad (22)$$

$$\hat{v}(t) = y(t) - \hat{\varphi}^T(t) \hat{\theta}(t) \quad (23)$$

AM-MI-ESG 算法的计算步骤可归纳如下:

1) 初始化: 令 $t = 1$ 。置 $\hat{\theta}(0) = \frac{1_n}{p_0}$, $\hat{x}_i(t-j) = \frac{1}{p_0}$ ($i = 1, 2, \dots, r; j = 1, 2, \dots, n_i$), $\hat{v}(t-i) = \frac{1}{p_0}$, 其中 1_n 表示元素均为 1 的 n 维列向量, $p_0 = 10^6$; $t \leq 0$ 时, $u_i(t) = 0$, $y(t) = 0$, $v(t) = 0$ 。

2) 收集数据 $u_1(t)$, $u_2(t)$, \dots , $u_r(t)$ 和 $y(t)$, 用式 (20)、(21) 构造 $\hat{\varphi}_i(t)$ 和 $\hat{\varphi}_n(t)$, 用式 (16) ~ (18) 构造 $Y(p, t)$, $\hat{\varphi}(t)$ 和 $\hat{\Phi}(p, t)$ 。

3) 用式 (14) 计算 $E(p, t)$, 用式 (15) 计算 $r(t)$ 。

4) 用式 (13) 刷新 $\hat{\theta}(t)$ 。

5) 用式 (22) 计算 $\hat{x}_i(t)$, 用式 (23) 计算 $\hat{v}(t)$ 。

6) t 增 1, 转至步骤 2), 继续递推计算。

3 仿真例子

考虑如下的双输入单输出系统:

$$y(t) = \frac{B_1(z)}{A_1(z)} u_1(t) + \frac{B_2(z)}{A_2(z)} u_2(t) + D(z) v(t)$$

$$A_1(z) = 1 + a_1(1) z^{-1} = 1 - 0.16z^{-1}, A_2(z) = 1 + a_2(1) z^{-1} = 1 - 0.10z^{-1}$$

$$B_1(z) = b_1(1) z^{-1} = 0.85z^{-1}, B_2(z) = b_2(1) z^{-1} = 0.74z^{-1}$$

$$D(z) = 1 + d_1 z^{-1} = 1 + 0.24z^{-1}$$

$$\theta = [a_1(1), b_1(1), a_2(1), b_2(1), d_1]^T \\ = [-0.16, 0.85, -0.10, 0.74, 0.24]^T$$

仿真时, 输入 $u_1(t)$ 和 $u_2(t)$ 采用零均值单位方差互不相关可测的持续激励信号序列, $v(t)$ 采用零均值方差为 $\sigma^2 = 0.50^2$ 的白噪声序列, 对应的系统噪声比 $\delta_{\text{ns}} = 45.19\%$ 。在不同新息长度 $p = 1$ (即 AM-ESG 算法) 和 $p = 3, 5$ 下应用 AM-MI-ESG 算法估计这个系统的参数, 参数估计误差 $\delta = \|\hat{\theta}(t) - \theta\| / \|\theta\|$ 随 t 变化曲线如图 3 所示。

4 结束语

结合辅助模型辨识思想和多新息辨识理论, 推导出了辨识有色噪声干扰的输出误差多输入单输出系统的 AM-MI-ESG 算法。仿真表明所提出的算法可以有效估计系统参数。与 AM-ESG 算法相比, AM-MI-ESG 具有更小的参数估计误差和更高的收敛速度。

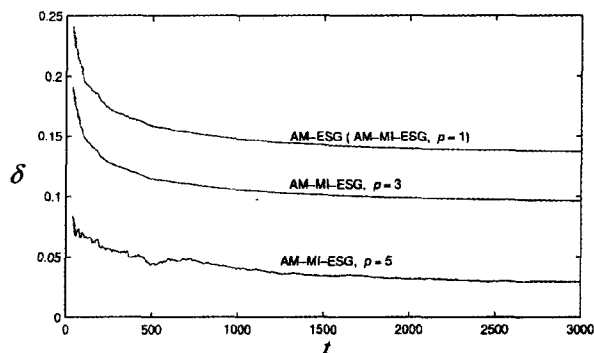


图 3 参数估计 δ 随数据长度 t 变化曲线

参考文献:

- [1] Ljung L. System Identification: Theory for the User[M]. 2nd Edition. Englewood Cliffs, New Jersey: Prentice-Hall, 1999.
- [2] 张勇, 杨慧中. 有色噪声干扰输出误差系统的偏差补偿递推最小二乘辨识方法[J]. 自动化学报, 2007, 33(10): 1053-1060.
- [3] 杨慧中, 张勇. Box-Jenkins 模型偏差补偿方法与其它辨识方法的比较[J]. 控制理论与应用, 2007, 24(2): 215-222.
- [4] 王冬青. 基于辅助模型的递推增广最小二乘辨识方法[J]. 控制理论与应用, 2009, 26(1): 51-56.
- [5] Ding F, Chen T. Performance analysis of multi-innovation gradient type identification methods[J]. Automatica, 2007, 43(1): 1-14.
- [6] 王冬青, 丁锋. Box-Jenkins 模型的基于辅助模型的多新息广义增广随机梯度算法[J]. 控制与决策, 2008, 23(9): 999-1003.
- [7] 刘英玉, 申东日, 陈义俊, 等. 基于前向神经网络的多新息随机梯度辨识算法[J]. 哈尔滨商业大学学报(自然科学版), 2006, 22(2): 83-86.
- [8] 岳殿武, 王谦. MISO 衰落信道中带有同信道干扰和噪声的空时分组编码系统容量分析[J]. 中国科学, 2009, 39(11): 1224-1230.
- [9] 魏国, 刘剑, 雷苗, 等. 基于 B 样条递推最小二乘的非线性 MISO 传感器系统建模方法[J]. 仪器仪表学报, 2009, 30(7): 1404-1409.
- [10] 伊学农, 周琪. 基于改进 BP 网络与 MISO 模型的污水系统建模[J]. 水处理技术, 2005, 20(1): 21-24.
- [11] 桑海峰, 王福利, 何大阔. 基于最小二乘支持向量机的发酵过程混合建模[J]. 仪器仪表学报, 2006, 27(6): 629-633.
- [12] 朱燕飞, 伍建平, 李琦. MISO 系统的混合核函数 LS-SVM 建模[J]. 控制与决策, 2005, 20(4): 417-420.
- [13] Ding F, Chen H B, Li M. Multi-innovation least squares identification methods based on the auxiliary model for MISO systems[J]. Applied Mathematics and Computation, 2007, 187(2): 658-668.