

摘要

多率系统在实际系统中广泛存在,如通信系统,化学工业过程,生物医学等.多率系统中存在两种或两种以上的操作频率,其数学模型要比传统的离散时间系统复杂.鉴于此,多率系统辨识的研究受到了国内外学者的关注和重视.本文基于多新息辨识理论,辅助模型辨识思想,递阶辨识原理,最小二乘辨识原理和梯度搜索原理研究多率多输入系统的辨识问题,并用数值仿真例子验证提出算法的有效性.论文的工作主要有如下几个方面:

1. 研究了多输入多输出系统的辨识问题.借助于多新息辨识理论,引入新息长度,将单新息向量扩展为多新息矩阵,提出了多新息随机梯度算法.应用鞅收敛定理证明了多新息随机梯度算法的收敛性.
2. 研究了方程误差多率系统的辨识问题.为了提高算法的收敛速度和估计精度,从多新息角度,引入新息长度,将单新息标量扩展为多新息向量,提出了多新息随机梯度算法,应用鞅定理证明多新息随机梯度算法的收敛性.由于多率系统离散化后得到高维、参数数目多的辨识模型,算法的计算量大.借助递阶辨识原理,提出了可以降低计算量的递阶最小二乘算法.
3. 研究了输出误差多率系统的辨识问题.结合辅助模型辨识思想和多新息辨识理论,提出了辅助模型多新息随机梯度算法.利用辅助模型辨识思想,提出了辅助模型最小二乘算法,应用鞅收敛定理证明辅助模型最小二乘算法的收敛性.
4. 研究了多率系统子模型的辨识问题.基于辅助模型辨识思想,对多率系统的每个子模型分别建立辅助模型,提出了辅助子模型最小二乘算法.与辅助模型最小二乘算法比较,需要辨识的参数数目较少,计算量较小.
5. 研究了 Box-Jenkins 多率系统的辨识问题.结合辅助模型辨识思想和多新息辨识理论,提出了辅助模型多新息广义增广随机梯度算法.辅助模型多新息广义增广随机梯度算法充分利用多率输入输出当前和过去时刻的新息,提高了辅助广义增广随机梯度算法的收敛速度和估计精度;辅助模型多新息广义增广随机梯度算法不需要计算协方差矩阵,计算量要小于辅助模型广义增广最小二乘算法.

关键词: 多率系统, 多新息, 辅助模型, 递阶辨识, 最小二乘, 随机梯度

Abstract

Multirate systems abundantly exist in actual areas, such as communication, chemical process and biomedical systems. Multirate systems have two or more sampling and updating rates. The mathematical description of multirate systems is more complicated than that of conventional discrete-time systems. Thus, the identification of multirate systems has been attracted more and more attention. In this dissertation, based on the multi-innovation identification theory, the auxiliary model identification idea, the hierarchical identification principle, the least squares identification principle and the gradient search principle, the identification of multirate multiple-input systems has been studied. The numerical examples are given to test the effectiveness of the proposed algorithms. The main contributions of this dissertation are as follows:

1. The identification problem of multiple-input multiple-output systems is studied. By means of the multi-innovation identification theory, introducing an innovation length and expanding a single innovation vector into an multi-innovation matrix, the multi-innovation stochastic gradient algorithm is derived. Using the Martingale convergence theorem, the convergence properties of the multi-innovation stochastic gradient algorithm are analyzed.
2. The identification problem of error equation multirate systems is studied. In order to improve the convergence speed and enhance the estimates accuracy of algorithms, from the viewpoint of multi-innovation, introducing an innovation length and expanding the single scalar to an multi-innovation vector, the multi-innovation stochastic gradient algorithm is proposed. The convergence properties of the proposed algorithm are analyzed by using the Martingale convergence theorem. Since the state-space model of multirate systems obtained by using the discretization technique has a high demension and a large number of parameters needed to be identified, this case leads to a large computational load for the algorithms. For reducing the computational load, by using the hierarchical identification principle, the hierarchical least squares algorithm is presented.
3. The identification problem of output error multirate systems is considered. Combining the auxiliary model identification idea and the multi-innovation identification theory, the multi-innovation stochastic gradient based auxiliary model algorithm is derived. By using the auxiliary model identification idea, the least squares based auxiliary model algorithm is presented. Using the Martingale convergence theorem, the convergence properties of the least squares based auxiliary model algorithm are analyzed.

4. The identification problem of multirate systems with sub-models is considered. Based on the auxiliary model identification idea, the auxiliary sub-models are set up. The least squares based auxiliary sub-models algorithm is proposed. Compared with the least squares based auxiliary model algorithm, the proposed algorithm has less parameters needed to identified and less computational load.

5. The identification problem of Box-Jenkins multirate systems is considered. Combining the auxiliary model identification idea and the multi-innovation identification theory, the auxiliary model multi-innovation generalized extended stochastic gradient algorithm is proposed. The auxiliary model multi-innovation generalized extended stochastic gradient algorithm repeatedly uses the multirate input-output data. Thus, compared with the auxiliary model generalized extended stochastic gradient algorithm, it has faster convergence speed and better parameter estimates accuracy. The auxiliary model multi-innovation generalized extended stochastic gradient algorithm does not compute the covariance matrix. Compared with the auxiliary model generalized extended least squares algorithm, it has less computational load.

Keywords: multirate system, multi-innovation, auxiliary model, hierarchical identification, least squares, stochastic gradient

目 录

摘 要	I
Abstract	III
第一章 绪 论	1
1.1 论文研究的目的和意义	1
1.2 国内外研究现状	2
1.2.1 系统辨识概述	2
1.2.2 多率系统辨识综述	3
1.3 论文的主要内容	7
第二章 多变量系统多新息随机梯度辨识方法	9
2.1 引言	9
2.2 模型描述	10
2.3 多新息随机梯度算法	10
2.3.1 算法的提出	10
2.3.2 算法收敛性	12
2.3.3 仿真试验	19
2.4 本章小结	22
第三章 方程误差多率系统	23
3.1 引言	23
3.2 模型推导	23
3.3 随机梯度算法	26
3.3.1 算法的提出	26
3.3.2 仿真试验	29
3.4 多新息随机梯度算法	32
3.4.1 算法的提出	32
3.4.2 算法收敛性	33
3.4.3 仿真试验	40
3.5 递阶最小二乘算法	44
3.5.1 算法的提出	44
3.5.2 仿真试验	46

3.6 本章小结	48
第四章 输出误差多率系统	49
4.1 引言	49
4.2 模型推导	49
4.3 辅助模型多新息随机梯度算法	52
4.3.1 算法的提出	52
4.3.2 仿真试验	56
4.4 辅助模型最小二乘算法	60
4.4.1 算法的提出	60
4.4.2 算法收敛性	61
4.4.3 仿真试验	65
4.5 本章小结	71
第五章 多率系统子模型辨识方法	73
5.1 引言	73
5.2 模型描述	73
5.3 辅助模型最小二乘算法	74
5.3.1 算法的提出	74
5.3.2 仿真试验	80
5.4 本章小结	85
第六章 Box-Jenkins 多率系统	87
6.1 引言	87
6.2 模型描述	87
6.3 辅助模型辨识算法	89
6.3.1 辅助模型广义增广最小二乘算法	89
6.3.2 辅助模型广义增广随机梯度算法	92
6.3.3 仿真试验	94
6.4 本章小结	99
第七章 结论与展望	101
7.1 结论	101
7.2 展望	102
致 谢	103

参考文献	105
攻读博士学位期间发表和录用论文目录	113

第一章 绪论

1.1 论文研究的目的和意义

传统的离散采样系统,输入刷新周期和输出采样周期是相同的,只存在一种操作频率,称为单率系统.多率系统则不然,存在两种或两种以上的操作频率^[1].例如,对于 r 输入单输出离散时间系统,如果有 r 个刷新周期和采样周期,存在 $r+1$ 种操作频率,就得到多率系统.

多率系统在实际领域中有着广泛的应用,如航空飞行系统模拟^[2],通信系统领域^[3],网络控制系统领域^[4,5],数字信号过程领域^[2,6],化学工业过程^[7-10],生物医学领域^[11]等.在航天领域中,由于雷达系统的雷达天线是间断性的给出目标,这使得雷达跟踪器和火力控制的设计是在多采样数据的基础上,得到了多率系统^[2].在数字用户线路(DSL)系统中,多个服务器之间产生串音(Crosstalk)是不可避免的,然而,由于产生串音的信号来自不同的服务器,因此各个信号的频率不同,使得不同频率的输入信号同时存在于数字用户线路系统中,即串音系统为多率多输入系统.从频谱管理的角度来看,人们需要知道产生串音的服务器之间的关系,即需要对串音系统进行辨识^[3].在数字控制系统中,许多被控对象内部各处信号的变化速率可能相差很大,例如温度信号与电信号的变化速率不同,在这种情况下,最好的方法是在系统各处针对不同变化速率的信号采用不同的采样周期,从而可采用不同转换速度的A/D和D/A转换器,在花费较小成本的前提下,提高计算机控制系统的控制品质,此时的数字控制系统为多率控制系统^[12].在信号处理领域,对于助听器的设计,输入信号被分解为不同的次波段,每个次波段承载着不同的权重,以能够读出来自不同波段的信号范围,因此多率滤波器组的设计是这个设计中的关键^[6].在化学工业过程中,如工业漂白过程,利用纤维材料增加物质的亮度,在这个过程中,某些输出变量的获得,如亮度,经过实验分析需要的时间长,然而,输入相对要快一些,得到的是快速率输入慢速率输出的多率系统^[8,9].在生物医学领域,医学参数成像技术可以提供生物系统的生理学和生物化学方程的图像量化,并能够实现相应的解剖结构的方程的图像可视化,这种技术的实现需要估计生物系统的参数,所使用的数据是通过不同动态过程得到的不同采样数据^[11].因此研究多率系统有强大的实际研究背景和重要的研究意义.

系统辨识,控制理论和状态估计是现代控制理论中三个密切相关的研究领域.系统辨识是研究建立系统数学模型的理论和方法.辨识先于控制,使得系统辨识在控制问题中显得尤为重要.目前,在研究多率系统控制问题的文献中,大多数控制算法都是假设多率系统的参数是已知的,即假设多率系统对应的连续系统参数是已知的,或多率系统对应的单率模型的参数是已知的,或多率系统的状态是已知的.然而,实际系统中要事先知道系统的参数和状态几乎是不可能的.因此,多率系统的辨识在其控制问题中占有重要的位置.

1.2 国内外研究现状

1.2.1 系统辨识概述

根据辨识专家瑞典学者 Ljung. L 对系统辨识的定义: 系统辨识有三要素 - 数据, 模型类和准则. 系统辨识是按照一个准则, 在模型类中选择一个与数据拟合得最好的模型^[15]. 三个要素中数据是辨识的基础, 准则是辨识的优化目标, 模型类是所寻找的模型范围. 优化的方法也是辨识的关键之一. 辨识也就是从系统的输入输出数据提取数学模型的过程, 数据的采集一般是通过实验设计得到, 所得到的数据能包含系统的充分信息. 模型类主要是数学模型和辨识模型. 数学模型分为很多类别, 如参数模型和非参数模型, 连续模型和离散模型, 确定性模型和随机模型等. 辨识模型是把系统的参数写成参数向量的形式, 输入输出写成一个信息向量形式, 也称为辨识表达式. 辨识模型分为线性系统辨识模型 (可分为线性系统脉冲响应模型, 时变系数差分方程描述的线性系统, 线性动态系统等) 和非线性系统辨识模型 (可分为静态非线性系统, 非线性动态系统, 状态方程等). 辨识的基本步骤包括实验设计, 数据采集, 数据预处理, 结构辨识, 参数估计, 模型验证.

当控制输入刷新周期和输出采样周期相等, 则得到传统的离散时间采样系统^[1]. 离散时间采样系统是多率系统的一个特例. 目前, 对于离散时间系统的辨识研究已经较为成熟, 并得到了一定的成果, 如澳大利亚学者 Goodwin 和 Sin 的《Adaptive Filtering, Prediction and Control》^[13], 瑞典学者 Söderström 和 Stoica 的《System Identification》^[14], 瑞典学者 Ljung 的《System Identification: Theory for the Users》^[15], 方崇智和萧德云的《过程辨识》^[16]等.

辨识方法有很多类别, 按其计算方式可分为一次完成算法, 递推辨识和迭代估计辨识方法; 按其实效性可分为在线和离线辨识方法; 按其属性特征可分为最小二乘辨识方法, 随机梯度辨识方法, 输出误差辨识方法, 多新息辨识方法, 递阶辨识方法, 极大似然估计方法和贝叶斯辨识方法等. 这里简单介绍本文将要用到的递推辨识方法, 最小二乘辨识方法, 随机梯度辨识方法, 辅助模型辨识方法, 多新息辨识方法和递阶辨识方法.

递推辨识方法是一种在线辨识, 随着时间推移, 采集数据的增加, 不断更新参数估计, 即下一时刻的参数估计等于前一刻的参数估计加上修正项, 这种推算过程称为递推算法.

最小二乘辨识方法^[15, 16]是基本的参数估计方法之一, 是极小化模型和系统输出的误差平方和得到参数估计值. 它可以用于递推算法, 也可用于迭代算法. 针对不同的模型辨识, 对最小二乘法进行修正, 得到各种相应的最小二乘类辨识算法. 如辨识方程误差类的递推最小二乘算法等.

梯度辨识方法^[15, 16]是对准则函数利用最速下降法搜索参数估计. 其特点不需要计算协方差矩阵, 因此, 与最小二乘算法比较, 其计算量小.

辅助模型辨识方法来源于自适应模型参考系统技术^[17, 18], 主要的思想是借助于一个参

考模型 (也称为辅助模型), 用参考模型的输出估计代替系统不可测的无噪输出, 使得参考模型输出逼近这些不可测的无噪输出, 从而获得系统参数的一致估计^[19-23]. 这种方法是针对于辨识模型信息向量中存在不可测的变量而提出的.

多新息辨识方法^[24-28]是近年来发展起来的辨识方法, 多新息辨识理论的主要思想是将单新息扩展为多新息, 使得算法可以充分利用可测的输入输出数据, 以提高算法的收敛速度和估计精度.

递阶辨识方法^[29-31]可以用来解决维数大, 变量数多的大系统辨识, 并对具有复杂结构耦合多变量系统的辨识也很有用. 递阶辨识原理首先将辨识模型分解为一些子模型, 其次采用最小二乘算法辨识各个子模型, 最后协调处理各个子辨识算法的关联项.

本文将利用上述五种辨识方法研究多率系统的辨识问题, 提出算法估计系统的参数.

对于辨识算法的收敛性, 在文献 [13], [14], [15] 和 [16] 中, 对各种算法的收敛性进行了分析. 在文献 [32] 和 [33] 中, 应用随机过程和鞅收敛定理, 对给出的算法进行了收敛分析, 部分算法可以得到参数估计误差的上界.

1.2.2 多率系统辨识综述

自 20 世纪 50 年代, 人们开始研究多率系统控制的问题. 研究者主要采用古典采样控制系统的方法^[34], Sklansky 和 Ragazzini 提出频域分解 (Frequency decomposition) 技术研究误差采样控制系统^[35], 之后, Kranc 提出切换分解 (Switch decomposition) 技术把多率控制结构转化为等价的单率控制结构, 将多率控制问题转化为单率控制问题^[36]. 1959 年, Kalman 和 Bertram 用状态空间稳定分析技术研究多率系统的控制 (分析与综合), 开启了多率系统在时间域中的研究^[37]. 随着计算机控制系统的发展, 多率控制系统的研究也得到了迅速的发展, 从 20 世纪 80 年代以来, 多率系统的 H_2/H_∞ 最优控制^[38,39], 预测控制^[40-42], 自适应控制^[43,44]等一系列控制问题成为国际控制的活跃领域. 由于辨识是控制问题的基础, 因此多率系统的辨识备受国内外学者的关注, 成为国际控制领域的热点, 得到了一定的成果.

r 输入 m 输出的多率系统如图 1.1 所示. P_c 是连续时间过程, 离散输入信号 $u_j(kT_j)$

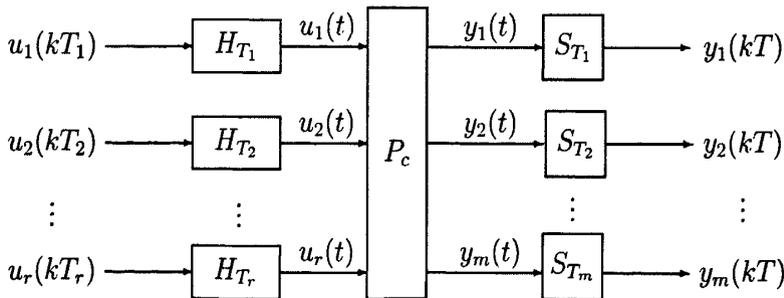


图 1.1 多率系统

Fig. 1.1 The multirate systems

($j = 1, 2, \dots, r$) 经过零阶保持器 H_{T_j} 得到连续信号 $u_j(t)$, 连续输出信号 $y_i(t)$ ($i = 1, 2, \dots, m$) 经过采样器 S_{T_i} 得到离散信号 $y_j(kT)$. $T_j := p_j h$ ($j = 1, 2, \dots, r$) 为输入刷新周期, $S_{T_i} = q_i h$ ($i = 1, 2, \dots, m$) 为输出采样周期, h 为基周期. 可测的输入输出数据为 $\{u_1(kT_1), u_2(kT_2), \dots, u_r(kT_r), y_1(kT), y_2(kT), \dots, y_m(kT) : k = 0, 1, 2, \dots\}$. 即使 P_c 为连续时不变过程, 如图 1.1 所示的多率系统是时变周期系统.

多率系统的数学模型主要有频域模型, 时域模型和小波模型^[45]. 在文 [45] 详细综述了这三种模型. 本文主要介绍时域模型中的有理分式传递函数模型和状态空间模型.

对于双率系统, 存在两种操作频率的多率系统, 其控制刷新周期为 ph , 输出采样周期为 qh , 可获得的输入输出数据为 $\{\mathbf{u}(kph), \mathbf{y}(kqh) : k = 0, 1, 2, \dots\}$, 其中 $\mathbf{u}(kph) = [u_1(kph), u_2(kph), \dots, u_r(kph)]^T$, $\mathbf{y}(kqh) = [y_1(kqh), y_2(kqh), \dots, y_r(kqh)]^T$. 假设 P_c 是一个线性时不变系统, 用采样周期 h 离散化 P_c 得到单率系统 $P_1 = S_h P_c H_h$, P_1 具有如下有理分式的传递函数,

$$\mathbf{y}(kqh) = P_1(z)\mathbf{u}(kph) = \frac{B(z)}{A(z)}\mathbf{u}(kph),$$

其中 $A(z)$ 和 $B(z)$ 为维数 n 的多项式函数, z^{-1} 是后移算子: $z^{-1}\mathbf{y}(kqh) = \mathbf{y}(kqh - h)$. 然而, 上述所得模型不适于双率系统的辨识. 例如, 当 $p = 2, q = 3$ 时, 从上式得到递归方程为,

$$\begin{aligned} \mathbf{y}(kqh) + a_1\mathbf{y}(kqh - h) + a_2\mathbf{y}(kqh - 2h) + \dots + a_n\mathbf{y}(kqh - nh) \\ = \mathbf{b}_0\mathbf{u}(kph) + \mathbf{b}_1(kph - h) + \mathbf{b}_2(kph - 2h) + \dots + \mathbf{b}_n(kph - nh). \end{aligned}$$

当 k 是 3 的倍数时, $\mathbf{y}(kqh)$ 是可得到的, 而 $\mathbf{y}(kqh - h)$ 和 $\mathbf{y}(kqh - 2h)$ 是不可得到的; 当 k 是 2 的倍数时, $\mathbf{u}(kph)$ 是可得到的, 而 $\mathbf{u}(kph - h)$ 是不可得到的. 可以采用多项式变换技术 (Polynomial transformation technique)^[1] 推导得到双率系统的辨识模型,

$$\mathbf{y}(kqh)P_2(z) = \frac{B(z)\phi(z)}{A(z)\phi(z)}\mathbf{u}(kph) = \frac{\beta(z^{pq})}{\alpha(z^{pq})}\mathbf{u}(kph),$$

其中 $\phi(z)$ 为 pq 维的多项式. 此时, 基于双率输入输出数据, 模型 $P_2(z)$ 是可用于辨识的双率系统.

对于多率系统的状态空间模型的推导, 一般采用提升技术 (Lifting technique). 1957 年, Kranc 提出的切换分解技术^[36] 是提升技术的雏形, 后来由 Friedland 和 Khargonekar 等人引入提升技术^[46, 47]. 提升技术是引入提升算子, 对输入和输出进行提升, 进而提升系统, 得到多率系统的离散时间状态间模型. 提升技术已经是处理多率系统的标准工具. 这里给出双率系统提升后的离散时间系统. 定义提升因子 L_p 和 L_q ,

$$\mathbf{u}_l(kT) = L_p\mathbf{u}(kT) = \begin{bmatrix} \mathbf{u}(kT) \\ \mathbf{u}(kT + T_1) \\ \vdots \\ \mathbf{u}(kT + (q-1)T_1) \end{bmatrix}, \quad \mathbf{y}_l(kT) = L_q\mathbf{y}(kT) = \begin{bmatrix} \mathbf{y}(kT) \\ \mathbf{y}(kT + T_2) \\ \vdots \\ \mathbf{y}(kT + (p-1)T_2) \end{bmatrix}.$$

则提升后的双率系统为

$$\begin{cases} \mathbf{x}(kT + T) = \mathbf{A}_l \mathbf{x}(kT) + \mathbf{B}_l \mathbf{u}_l(kT), \\ \mathbf{y}_l(kT) = \mathbf{C}_l \mathbf{x}(kT) + \mathbf{D}_l \mathbf{u}_l(kT), \end{cases}$$

这里 $T_1 = ph$ 为刷新周期, $T_2 = qh$ 为采样周期, $T = pqh$ 为框架周期. $\mathbf{A}_l, \mathbf{B}_l, \mathbf{C}_l$ 和 \mathbf{D}_l 为具有适当维数的矩阵. 在文献 [31] 有详细的推导, 这里不再赘述.

Chen 和 Francis 将离散化技术引入到研究多率系统的控制问题中 [48], 将多率系统在一个框架周期内进行离散化, 得到多率系统的离散时间状态空间模型. 在本文的第三章和第四章中有详细的描述.

在频率域中, 对于多率系统的辨识, 主要研究方法是通过频率辨识模型或利用频域工具得到时域模型, 得到的大多是系统的频谱, 为非参数模型 [49,50]. 小波模型的辨识, 主要特点是数据重采样, 小波模型的使用并不普遍, 但小波辨识模型将会对系统的频域特性有比较好的反映 [51]. 基于均匀和非均匀频域响应数据的离散时不变系统, 在噪声协方差函数是先验的条件下, Akçay 等利用子空间辨识方法得到系统的参数估计, 验证了算法的收敛性 [52,53]. Akçay 和 Türkay 将子空间辨识方法推广至离散光谱率的辨识 [54]. 基于非均匀频域响应数据, Akçay 研究了辅助量子空间辨识方法辨识多输入多输出离散时不变系统, 在不需要噪声协方差函数是先验的条件下, 得到系统的参数估计, 验证了算法的收敛性 [55]. Gillberg 和 Ljung 基于均匀和非均匀采样数据, 在频率域中研究了输出误差连续系统的辨识, 主要应用多项式插值和均方最小二乘算法 [56,57].

对于多率系统的有理分式传递函数模型, Lu 和 Fisher 利用最小二乘辨识方法得到了双率 DARMA 模型的参数估计 [58,59], 接着他们的工作, Ding 等考虑了双率 CAR 模型的辨识, 并且分析了提出的最小二乘算法的收敛性 [60]. Ding 和 Chen 针对确定性双率系统, 提出了最小二乘算法和带有遗忘因子最小二乘算法估计系统的参数, 针对带有色噪声双率系统, 利用辅助变量最小二乘算法得到系统的参数估计, 并对所提出的算法分析了收敛性, 同时得到参数估计误差上界 [61]; 对于双率 FIR 系统, 提出辅助模型最小二乘算法得到系统的参数估计 [62]; 针对输出误差双率系统, 提出辅助模型最小二乘算法和辅助模型随机梯度算法, 分析了参数估计和输出估计的收敛性 [22,23]. Ding 等得到修正的随机梯度算法辨识双率系统的参数 [63]. Sahebsara 等将多输入多输出多率系统分解为若干个双率系统, 用最小二乘算法辨识双率系统的有理分式传递函数模型 [9], 得到系统的参数估计. 上述文献都是基于多项式变化技术, 得到与双率系统辨识模型等价的一个模型, 直接利用可测的双率输入输出数据辨识系统参数.

对于快速率慢输出非均匀采样系统, 基于传递函数模型, Zhu 等应用输出误差辨识方法估计系统的参数 [64]. 接着他们的工作, Mo 等研究了快输入慢输出双率系统的辨识问题, 提出了两步算法, 第一步结合最小二乘算法和 Newton-Raphson 算法得到双率系统 FIR 模型的参数估计, 第二步为了解决 Newton-Raphson 算法带来的初始化和极小点问题, 利用输出误差辨识方法得到系统的参数估计 [10]. Ding 和 Ding 利用多项式转换技术变双率系统为可辨识系

统,得到了基于残差的增广最小二乘算法估计系统参数^[65].他们又分析了输出误差模型的不规则损失输出数据的最小二乘参数估计^[66].Ding等研究了基于多项式变换的损失观测时间序列AR建模问题^[67].有理分式传递函数模型主要针对双率系统,能否推广到多率系统,是可以进一步考虑的问题.

对于多率系统状态空间模型,基于提升技术,Li等得到了多率系统离散时间状态空间模型,在假设系统的状态是已知的情形下,估计系统的参数^[68];Li等提出一种新颖的子空间辨识方法辨识非均匀多率采样系统的故障检测和隔离^[69];在 H_2/H_∞ 性能指标设定的条件下,Sheng等通过设计最优滤波器估计多率采样系统的状态^[70];Li等通过得到卡尔曼滤波器,估计系统的状态,研究非均匀多率采样系统的故障检测问题^[71];Yu等在 l_2/l_∞ 意义下,通过设计多率系统滤波器,得到系统的状态估计^[72];Ding等应用递阶辨识原理,利用子空间状态空间辨识(4SID)方法得到双率系统的状态和参数估计^[31],同时也解决了由提升技术带来的因果约束问题.Liu和Lu研究了一类多率多输入系统的辨识,利用迭代最小二乘算法得到系统参数估计^[73].通信领域中申音系统是一类多输入单输出多率系统,Shi等针对这类系统,基于提升技术,推导系统的状态空间模型,进而得到传递函数模型,利用递推最小二乘算法得到系统的参数估计^[3].Raghavan等对化学工业过程中出现的非均匀多率采样系统,基于状态空间模型,利用期望最大化方法得到了系统参数的极大似然估计^[7].值得一提的是,由于提升后的系统存在因果约束问题,因此,基于提升技术辨识多率系统时,需要考虑这一问题^[74].

最近,针对多率多输入系统,Han等利用离散化技术,推导多率系统的状态空间模型,基于其传递函数模型,提出多新息随机梯度算法和辅助模型最小二乘算法估计系统参数^[75,76].

经过多项式变换技术,提升技术和离散化技术得到的多率系统的辨识模型,总会增加待辨识参数的数目,使得提出的多率系统的辨识算法的计算量大.因此,研究能够减少计算量的辨识算法是多率系统辨识方向之一.

对于多率系统辨识算法的收敛性问题,主要应用随机过程理论和鞅收敛定理来进行分析. Shi等针对多率FIR系统提出的辅助模型最小二乘算法,给出了参数估计误差一致有界的条件^[3].Ding等分析了辅助模型最小二乘算法和辅助模型随机梯度算法的收敛性,给出了参数估计一致收敛于真值的充分条件,同时得到了采样间输出估计也一致逼近真值^[22,23].Ding等在弱持续激励条件下,给出了最小二乘类辨识算法参数估计一致收敛于真值的充分条件,同时得到了参数估计误差一致有界的条件^[60-62].Ding等给出了双率随机梯度算法,给出了参数估计和采样间输出估计的一致收敛的充分条件^[77].Han等提出了多新息随机梯度算法和多率辅助模型最小二乘算法,证明了算法的收敛性^[75,76],但还没有得到参数估计误差一致有界的条件,是可以进一步考虑的问题.辨识算法的收敛性分析和理论证明一直是系统辨识领域的难点之一,对于多率系统的辨识算法也不例外.因此,对于多率系统辨识算法的收敛性分析需要继续探索.

1.3 论文的主要内容

本文主要研究多率多输入系统的辨识问题. 利用离散化技术推导多率系统的离散时间状态空间模型, 得到其传递函数模型. 由于离散化后的多率系统为多变量系统, 因此本文也研究了多变量系统的辨识. 结合多新息辨识理论、辅助模型思想、递阶辨识原理、最小二乘辨识原理和梯度搜索原理, 研究多变量系统、方程误差多率系统、输出误差多率系统、子模型多率系统和 Box-Jenkins 多率系统的辨识问题. 应用随机过程理论和鞅收敛定理分析提出算法的收敛性, 利用数值仿真例子验证算法的有效性. 本文主要内容和章节安排如下:

1. 第一章为绪论. 介绍了多率系统的实际应用背景, 说明了多率系统辨识的实际应用价值和理论研究意义. 对系统辨识进行了概述, 介绍了本文要用到的辨识方法. 对多率系统辨识的研究进行了综述.
2. 第二章研究了多变量系统的辨识问题. 基于多新息辨识理论, 引入新息长度, 扩展单新息向量为多新息矩阵, 提出了多新息随机梯度算法. 用鞅收敛定理分析算法的收敛性.
3. 第三章研究了方程误差多率系统的辨识问题. 为了提高算法的收敛速度和估计精度, 借助多新息辨识理论, 提出了多新息随机梯度算法, 应用鞅收敛定理证明算法的收敛性. 应用递阶辨识原理, 提出了计算量较小的递阶最小二乘算法.
4. 第四章研究了输出误差多率系统的辨识问题. 借助辅助模型辨识思想和多新息辨识理论, 提出了辅助模型多新息随机梯度算法. 利用最小二乘辨识原理, 提出辅助模型最小二乘算法, 用鞅收敛定理分析了算法的收敛性.
5. 第五章研究了带有子模型多率系统的辨识问题. 基于辅助模型辨识思想, 提出了辅助子模型最小二乘算法. 与辅助模型最小二乘算法比较, 提出的算法有较小的计算量.
6. 第六章研究了 Box-Jenkins 多率系统的辨识问题. 结合辅助模型辨识思想和多新息辨识理论, 提出了辅助模型多新息广义增广随机梯度算法.
7. 第七章为结论与展望. 对本文工作进行了总结, 并对多率系统辨识的研究作了展望.

第二章 多变量系统多新息随机梯度辨识方法

2.1 引言

随着工程实践的需要和控制理论的发展,许多工业系统已不再局限于单变量系统,而是结构复杂、信息不完全、模型不确定的多变量 (Multiple-input and multiple-output: MIMO) 系统^[78]. 多率系统就是 MIMO 系统中各个输入和输出通道的刷新和采用周期不同而产生的. 因此,有必要深入研究 MIMO 系统辨识方法来适应现代工业系统的发展.

对于 MIMO 系统的辨识,由于 MIMO 系统维数高、参数数目多,使得辨识算法的计算量大、参数估计精度低 (在相同数据长度下),成为 MIMO 系统辨识要解决的关键问题之一. 较早的研究方法是首先将多输入多输出系统分解为几个子系统,再利用最小二乘算法估计每个子系统的参数,从而得到系统的参数估计,但是该方法需要计算多个协方差矩阵,导致算法计算量很大^[79,80]. 自上世纪 80 年代以来,人们开始致力于状态空间辨识方法的研究,并提出了子空间辨识 (4SID) 方法来研究 MIMO 系统的辨识^[81,82]. 4SID 方法的主要思想是,通过计算输入输出数据空间的一个特别结构的子空间 (数据压缩阵),直接辨识系统状态空间模型的参数矩阵,所使用的方法是 RQ 分解或 QR 分解和奇异值 SVD 分解技术,随着信息矩阵维数的增加,算法的计算量也会增加^[83,84]. Ding 和 Chen 基于递阶辨识方法,将 MIMO 系统分解为两个子系统,利用随机梯度算法和最小二乘算法得到系统的参数估计^[23,29].

众所周知,随机梯度算法的计算量较小,但是收敛速度较慢. 为了提高随机梯度算法的收敛速度, Ding 和 Chen 基于多新息辨识理论^[85],引入新息长度,将单新息标量扩展为多新息向量,充分利用输入输出信息,提出多新息随机梯度算法和多新息遗忘梯度算法辨识线性回归系统,分析了算法的收敛性,得到了参数估计误差的上界,估算得到适当的新息长度^[26]. Liu 等针对多输入单输出输出误差模型系统,提出了基于辅助模型多新息随机梯度算法^[86]. Ding 等对于多输入单输出的 ARX 系统,提出了多新息最小二乘算法,并分析了其收敛性^[87]. Ding 等对于输出误差滑动平均系统,分别提出了辅助模型多新息增广随机梯度算法^[88]和多新息增广最小二乘算法^[89]. Wang 和 Ding 解决了辅助模型多新息增广随机梯度算法的收敛性的证明问题^[90]. Zhang 等将多新息辨识方法应用到离散时间系统的自校正控制问题^[91].

本章基于多新息辨识理论,将多新息随机梯度算法推广至 MIMO 系统辨识. 组织如下: 第 2.2 节是对 MIMO 系统的描述,推导辨识模型. 第 2.3 节借助多新息辨识理论,引入新息长度,推导多新息随机梯度算法. 第 2.4 节应用鞅收敛定理,分析算法的收敛性,用数值仿真例子验证算法的有效性. 第 2.5 节为本章总结.

2.2 模型描述

考虑如下多输入多输出线性自回归模型

$$\mathbf{A}(z)\mathbf{y}(t) = \mathbf{B}(z)\mathbf{u}(t) + \mathbf{v}(t), \quad (2.2.1)$$

其中 $\mathbf{u}(t) \in \mathbb{R}^r$ 是系统的控制输入, $\mathbf{y}(t) \in \mathbb{R}^m$ 是系统的输出, $\mathbf{v}(t) \in \mathbb{R}^m$ 是零均值的随机白噪声, $\mathbf{A}(z)$ 和 $\mathbf{B}(z)$ 是带有后移算子 z^{-1} 的矩阵多项式, 其中 $z^{-1} : z^{-1}\mathbf{y}(t) = \mathbf{y}(t-1)$,

$$\mathbf{A}(z) = \mathbf{I} + \mathbf{A}_1 z^{-1} + \cdots + \mathbf{A}_{n_a} z^{-n_a}, \quad \mathbf{A}_i \in \mathbb{R}^{m \times m},$$

$$\mathbf{B}(z) = \mathbf{B}_1 z^{-1} + \cdots + \mathbf{B}_{n_b} z^{-n_b}, \quad \mathbf{B}_j \in \mathbb{R}^{m \times r}.$$

定义参数矩阵 $\boldsymbol{\theta}$ 和信息向量 $\boldsymbol{\varphi}(t)$ 如下:

$$\boldsymbol{\theta}^T := [\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_{n_a}, \mathbf{B}_1, \mathbf{B}_2, \cdots, \mathbf{B}_{n_b}] \in \mathbb{R}^{m \times n}, \quad n := mn_a + rn_b,$$

$$\boldsymbol{\varphi}(t) := [-\mathbf{y}^T(t-1), -\mathbf{y}^T(t-2), \cdots, -\mathbf{y}^T(t-n_a), \\ \mathbf{u}^T(t-1), \mathbf{u}^T(t-2), \cdots, \mathbf{u}^T(t-n_b)]^T \in \mathbb{R}^n.$$

则 (2.2.1) 式可表达为,

$$\mathbf{y}(t) = \boldsymbol{\theta}^T \boldsymbol{\varphi}(t) + \mathbf{v}(t), \quad (2.2.2)$$

其中 \mathbf{A}_i ($i = 1, 2, \cdots, n_a$) 和 \mathbf{B}_j ($j = 1, 2, \cdots, n_b$) 是待辨识的参数.

假设: 当 $t \leq 0$, $\mathbf{u}(t) = \mathbf{0}$, $\mathbf{y}(t) = \mathbf{0}$ 和 $\mathbf{v}(t) = \mathbf{0}$. $\{\mathbf{u}(t), \mathbf{y}(t) : t = 0, 1, 2, \cdots\}$ 或者 $\{\mathbf{y}(t), \boldsymbol{\varphi}(t) : t = 0, 1, 2, \cdots\}$ 是可测量的输入输出信息. 为了研究方便, 假定 t 为当前时刻, 则 $\mathbf{y}(t)$ 和 $\boldsymbol{\varphi}(t)$ 称为当前的信息, $\{\mathbf{y}(t-i), \boldsymbol{\varphi}(t-i) : i = 1, 2, \cdots, p-1\}$ 称为过去的信息.

2.3 多新息随机梯度算法

2.3.1 算法的提出

假定 \mathbf{I} 代表具有适当维数的单位矩阵; $\mathbf{1}_{n \times m}$ 代表元素为 1 的 $n \times m$ 维的矩阵; \mathbf{T} 是指矩阵的转置; \mathbf{E} 是指期望算子; 矩阵 \mathbf{X} 的范数定义为 $\|\mathbf{X}\|^2 = \text{tr}[\mathbf{X}\mathbf{X}^T] = \text{tr}[\mathbf{X}^T\mathbf{X}]$.

定义准则函数

$$J(\boldsymbol{\theta}) = \|\mathbf{y}(t) - \boldsymbol{\varphi}(t)\boldsymbol{\theta}\|^2,$$

参照文 [13, 15], 根据梯度搜索原理极小化 $J(\boldsymbol{\theta})$ 得到如下估计系统 (2.2.2) 的参数 $\boldsymbol{\theta}$ 的随机梯度 (SG) 算法:

$$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t-1) + \frac{\boldsymbol{\varphi}(t)}{r(t)} \mathbf{e}(t), \quad (2.3.1)$$

$$\mathbf{e}(t) = \mathbf{y}^T(t) - \boldsymbol{\varphi}^T(t)\hat{\boldsymbol{\theta}}(t-1), \quad (2.3.2)$$

$$r(t) = r(t-1) + \|\boldsymbol{\varphi}(t)\|^2, \quad r(0) = 1, \quad (2.3.3)$$

其中 $\mathbf{e}(t) \in \mathbb{R}^{1 \times m}$ 是单新息向量, 其中的每个元素代表当前时刻的标量新息.

SG 算法仅用到了当前时刻的信息, 导致算法的收敛速度慢, 参数精度不高. 为了提高 SG 算法的收敛速度和估计精度, 本文根据多新息辨识理论, 充分利用可测量的输入输出信息, 推导多新息随机梯度算法. 引入新息长度 p , 扩展单新息向量 $\mathbf{e}(t) \in \mathbb{R}^{1 \times m}$ 为多新息矩阵

$$\mathbf{E}(p, t) = \begin{bmatrix} \mathbf{y}^T(t) - \boldsymbol{\varphi}^T(t)\hat{\boldsymbol{\theta}}(t-1) \\ \mathbf{y}^T(t-1) - \boldsymbol{\varphi}^T(t-1)\hat{\boldsymbol{\theta}}(t-1) \\ \vdots \\ \mathbf{y}^T(t-p+1) - \boldsymbol{\varphi}^T(t-p+1)\hat{\boldsymbol{\theta}}(t-1) \end{bmatrix} \in \mathbb{R}^{p \times m},$$

其中 $\{\mathbf{y}(t-i), \boldsymbol{\varphi}(t-i): i = 1, 2, \dots, p-1\}$ 为过去的信息, 这里 p 为正整数.

定义信息矩阵 $\Phi(p, t)$ 和输出矩阵 $\mathbf{Y}(p, t)$ 如下:

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

$$\mathbf{Y}(p, t) = [\mathbf{y}(t), \mathbf{y}(t-1), \dots, \mathbf{y}(t-p+1)]^T \in \mathbb{R}^{p \times m},$$

则新息矩阵 $\mathbf{E}(p, t)$ 可表达为

$$\mathbf{E}(p, t) = \mathbf{Y}(p, t) - \Phi^T(p, t)\hat{\boldsymbol{\theta}}(t-1).$$

参照文 [26], 得到如下估计系统 (2.2.2) 参数 $\boldsymbol{\theta}$ 的以新息长度为 p 的多新息随机梯度 (MISG) 算法:

$$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t-1) + \frac{\Phi(p, t)}{r(t)} \mathbf{E}(p, t), \quad (2.3.4)$$

$$\mathbf{E}(p, t) = \mathbf{Y}(p, t) - \Phi^T(p, t)\hat{\boldsymbol{\theta}}(t-1), \quad (2.3.5)$$

$$r(t) = r(t-1) + \|\boldsymbol{\varphi}(t)\|^2, \quad r(0) = 1, \quad (2.3.6)$$

$$\Phi(p, t) = [\boldsymbol{\varphi}(t), \boldsymbol{\varphi}(t-1), \dots, \boldsymbol{\varphi}(t-p+1)], \quad (2.3.7)$$

$$\mathbf{Y}(p, t) = [\mathbf{y}(t), \mathbf{y}(t-1), \dots, \mathbf{y}(t-p+1)]^T, \quad (2.3.8)$$

$$\boldsymbol{\varphi}(t) = [-\mathbf{y}^T(t-1), \dots, -\mathbf{y}^T(t-n_a), \mathbf{u}^T(t-1), \dots, \mathbf{u}^T(t-n_b)]^T. \quad (2.3.9)$$

由于 $\mathbf{E}(p, t) \in \mathbb{R}^{p \times m}$ 是多新息矩阵, 因此算法 (2.3.4)-(2.3.8) 被称为多新息随机梯度算法. 当 $p=1$ 时, MISG 算法 (2.3.4)-(2.3.8) 就变为 SG 算法 (2.3.1)-(2.3.3).

MISG 算法有如下收敛性质 [26]:

- 与 SG 算法 (2.3.1)-(2.3.3) 比较, SG 算法仅利用了当前时刻的信息 $\{\mathbf{y}(t), \boldsymbol{\varphi}(t)\}$, MISG 算法 (2.3.4)-(2.3.8) 不仅利用了当前时刻的信息 $\{\mathbf{y}(t), \boldsymbol{\varphi}(t)\}$, 还利用了过去时刻的可测信息 $\{\mathbf{y}(t-i), \boldsymbol{\varphi}(t-i): i = 1, 2, \dots, p-1\}$, 因此 MISG 算法有更好的参数估计收敛特性.

- MISG 算法重复利用可测的信息. 例如, 在 t 时刻, MISG 算法用到了 $\{\mathbf{y}(t-i), \boldsymbol{\varphi}(t-i) : i = 0, 1, \dots, p-1\}$, 在 $t+1$ 时刻, MISG 算法用到了 $\{\mathbf{y}(t+1-i), \boldsymbol{\varphi}(t+1-i) : i = 0, 1, \dots, p-1\}$. 在相邻的两个迭代中, 重复利用了信息 $\{\mathbf{y}(t-i), \boldsymbol{\varphi}(t-i) : i = 0, 1, \dots, p-2\}$. 因此, MISG 算法可以提高参数估计的精度.
- 随着新息长度 p 增大, MISG 算法参数估计的效果越来越接近递推最小二乘 (RLS) 算法. 与 RLS 算法比较, MISG 算法不需要计算协方差矩阵, 所以计算量较小.

初值 $\hat{\boldsymbol{\theta}}(0)$ 一般为 0 矩阵或很小的矩阵, 即 $\hat{\boldsymbol{\theta}}(0) = 10^{-6} \mathbf{1}_{n \times m}$. 利用 MISG 算法计算参数估计的步骤如下:

1. 令 $t = 1$, $\hat{\boldsymbol{\theta}}(0) = 10^{-6} \mathbf{1}_{n \times m}$, $r(0) = 1$, 选取数据长度 L_e 和新息长度 p .
2. 搜集输入输出信息 $\mathbf{u}(t)$ 和 $\mathbf{y}(t)$, 由式 (2.3.9) 构成信息向量 $\boldsymbol{\varphi}(t)$.
3. 由 (2.3.7) 式构成信息矩阵 $\Phi(p, t)$, 以及由式 (2.3.8) 构成输出矩阵 $\mathbf{Y}(p, t)$.
4. 由 (2.3.6) 式计算 $r(t)$, 以及由式 (2.3.5) 计算 $\mathbf{E}(p, t)$.
5. 由 (2.3.4) 式刷新参数估计 $\hat{\boldsymbol{\theta}}(t)$.
6. 增加 t 至 $t+1$, 返回第二步继续计算. 当 $t = L_e$ 时, 算法计算停止.

用 MISG 算法计算参数估计 $\hat{\boldsymbol{\theta}}(kT)$ 的流程图如图 2.1 所示.

2.3.2 算法收敛性

鞅 (Martingale) 的概念在现代概率论中起重要作用, 它已成为当今辨识与自适应控制领域内必不可少的研究工具之一. 鞅理论中最常用的部分是不等式与收敛定理, 其中鞅收敛定理是研究时不变随机系统递推辨识收敛性的数学分析工具. 本小节应用鞅收敛定理来分析 MISG 算法的收敛性.

假设 $\{\mathbf{v}(t), \mathcal{F}_t\}$ 是定义在概率空间 $\{\Omega, \mathcal{F}_t, P\}$ 鞅差向量序列, $\{\mathcal{F}_t\}$ 是由 $\mathbf{v}(t)$ 生成的 σ 代数序列, 即, $\mathcal{F}_t = \sigma(\mathbf{v}(t), \mathbf{v}(t-1), \mathbf{v}(t-2), \dots)$ [13]. 噪声向量序列 $\{\mathbf{v}(t)\}$ 满足

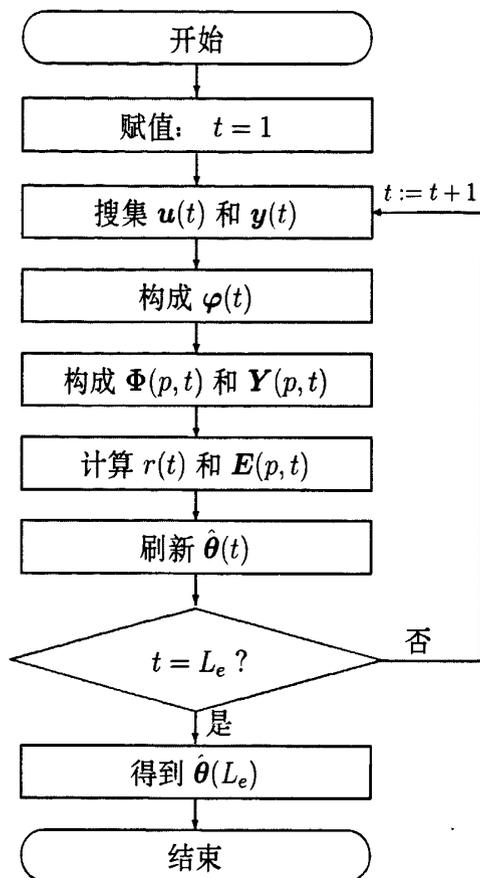
$$(A1) \quad E[\mathbf{v}(t) | \mathcal{F}_{t-1}] = \mathbf{0}, \text{ a.s.},$$

$$(A2) \quad E[\|\mathbf{v}(t)\|^2 | \mathcal{F}_{t-1}] = \sigma^2, \text{ a.s.},$$

$$(A3) \quad \limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{i=1}^t \|\mathbf{v}(i)\|^2 \leq \sigma^2 < \infty, \text{ a.s.}$$

引理 2.3.1 (Kronecker 引理)[13] 若

$$\mathbf{S}_t = \sum_{k=1}^t \mathbf{A}_k, \quad \mathbf{y}_t = \sum_{k=1}^t \mathbf{S}_k^{-1} \mathbf{x}_k,$$

图 2.1 计算参数估计 $\hat{\theta}(t)$ 流程图Fig. 2.1 The flowchart of computing the parameter estimate $\hat{\theta}(t)$

其中 A_k 是 $p \times p$ 的非负定阵, x_k 为 p 维的向量序列, 则由

$$\lim_{t \rightarrow \infty} y_t = y, \|y\| < \infty \text{ 和 } \lim_{t \rightarrow \infty} S_t^{-1} x_k = 0,$$

可导出

$$\lim_{t \rightarrow \infty} S_t^{-1} \sum_{k=1}^t x_k = 0.$$

Kronecker 引理的证明可参考谢新民和丁锋编著的《自适应控制系统》^[33].

引理 2.3.2 对于系统 (2.2.2) 和 MISG 算法 (2.3.4)–(2.3.8), 有如下结论

$$\sum_{t=1}^{\infty} \sum_{i=0}^{p-1} \frac{\|\varphi(t-i)\|^2}{r^2(t)} < \infty, \text{ a.s., } t \geq p.$$

证明: 根据 $r(t)$ 的定义式 (2.3.6), 有

$$\sum_{i=1}^{\infty} \frac{\|\varphi(i)\|^2}{r(i-1)r(i)} = \sum_{i=1}^{\infty} \frac{r(i) - r(i-1)}{r(i-1)r(i)} = \sum_{i=1}^{\infty} \left[\frac{1}{r(i-1)} - \frac{1}{r(i)} \right]$$

$$= \frac{1}{r(0)} - \frac{1}{r(\infty)} < \infty, \text{ a.s.}$$

因此,

$$\begin{aligned} \sum_{t=1}^{\infty} \sum_{i=0}^{p-1} \frac{\|\varphi(t-i)\|^2}{r^2(t)} &\leq \sum_{t=1}^{\infty} \sum_{i=0}^{p-1} \frac{\|\varphi(t-i)\|^2}{r(t-i-1)r(t-i)} \\ &= \sum_{i=0}^{p-1} \sum_{t=1}^{\infty} \frac{\|\varphi(t-i)\|^2}{r(t-i-1)r(t-i)} < \infty, \text{ a.s.} \end{aligned}$$

其中 $r(t) = 1$, 当 $t < 0$. 引理 2.3.2 证明完毕. □

引理 2.3.3 如果存在函数 $f(t) \geq 0$ 使得 $\lim_{t \rightarrow \infty} f(t) = f_0 < \infty$, 则有

$$\lim_{t \rightarrow \infty} \frac{1}{t} [f(1) + f(2) + \cdots + f(t)] = f_0.$$

证明: 证明简单, 这里从略. □

定理 2.3.1 (鞅收敛定理) 设 $\{T(t)\}, \{\varepsilon(t)\}, \{\varpi(t)\}$ 均为非负随机变量序列, 它们适应于递增 σ 代数序列 \mathcal{F}_t (即它们是 \mathcal{F}_t 可测的), 且满足

$$E[T(t+1)|\mathcal{F}_t] \leq T(t) - \varepsilon(t) + \varpi(t).$$

若 $\sum_{t=1}^{\infty} \varpi(t) < \infty$, a.s. 则 $T(t)$ 几乎肯定 (a.s.) 地收敛于一有限的随机变量 T_0 , 即 $T(t) \rightarrow T_0 < \infty$,

并且有 $\sum_{t=1}^{\infty} \varepsilon(t) < \infty$, a.s.

鞅收敛定理的证明可参考郭雷编著的《时变随机系统——稳定性、估计与控制》^[92].

定理 2.3.2 对于系统 (2.2.2) 和 MISG 算法 (2.3.4)–(2.3.8), 假设满足 (A1)–(A3), 信息向量是持续激励的, 即, 存在常数 α 和 $\beta, 0 < \alpha \leq \beta < \infty$ 和整数 N 使得满足下面的强持续激励条件

$$(A4) \quad \alpha \mathbf{I} \leq \frac{1}{N} \sum_{i=0}^{N-1} \varphi(t+i) \varphi^T(t+i) \leq \beta \mathbf{I}, \text{ a.s., } t > 0.$$

则参数估计 $\hat{\theta}(t)$ 一致收敛于真值 θ , 即, $\lim_{t \rightarrow \infty} \hat{\theta}(t) = \theta$, a.s.

证明: 定义噪声矩阵

$$\mathbf{V}(p, t) := [\mathbf{v}(t), \mathbf{v}(t-1), \cdots, \mathbf{v}(t-p+1)]^T \in \mathbb{R}^{p \times m}$$

和参数估计误差矩阵

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

将 (2.2.2) 式和 (2.3.4) 式代入上式中, 由 (2.3.5) 式得到

$$\begin{aligned}\tilde{\theta}(t) &= \tilde{\theta}(t-1) + \frac{\Phi(p,t)}{r(t)} [\mathbf{Y}(p,t) - \Phi^T(p,t)\tilde{\theta}(t-1)] \\ &= \tilde{\theta}(t-1) + \frac{\Phi(p,t)}{r(t)} [-\Phi^T(p,t)\tilde{\theta}(t-1) + \mathbf{V}(p,t)] \\ &=: \tilde{\theta}(t-1) + \frac{\Phi(p,t)}{r(t)} [-\tilde{\mathbf{Y}}(p,t) + \mathbf{V}(p,t)],\end{aligned}\quad (2.3.10)$$

其中

$$\tilde{\mathbf{Y}}(p,t) := \Phi^T(p,t)\tilde{\theta}(t-1). \quad (2.3.11)$$

由 (2.3.10) 式和 (2.3.11) 式, 得到

$$\begin{aligned}\tilde{\theta}^T(t)\tilde{\theta}(t) &= \tilde{\theta}^T(t-1)\tilde{\theta}(t-1) + \frac{1}{r(t)}\tilde{\theta}^T(t-1)\Phi(p,t)[- \tilde{\mathbf{Y}}(p,t) + \mathbf{V}(p,t)] \\ &\quad + \frac{1}{r(t)}[-\tilde{\mathbf{Y}}^T(p,t) + \mathbf{V}^T(p,t)]\Phi^T(p,t)\tilde{\theta}(t-1) \\ &\quad + \frac{1}{r^2(t)}[-\tilde{\mathbf{Y}}^T(p,t) + \mathbf{V}^T(p,t)]\Phi^T(p,t)\Phi(p,t)[- \tilde{\mathbf{Y}}(p,t) + \mathbf{V}(p,t)] \\ &= \tilde{\theta}^T(t-1)\tilde{\theta}(t-1) - \frac{2}{r(t)}\tilde{\mathbf{Y}}^T(p,t)\tilde{\mathbf{Y}}(p,t) + \frac{1}{r(t)}[\tilde{\mathbf{Y}}^T(p,t)\mathbf{V}(p,t) \\ &\quad + \mathbf{V}^T(p,t)\tilde{\mathbf{Y}}(p,t)] + \frac{1}{r^2(t)}[\tilde{\mathbf{Y}}^T(p,t)\Phi^T(p,t)\Phi(p,t)\tilde{\mathbf{Y}}(p,t) \\ &\quad - \tilde{\mathbf{Y}}^T(p,t)\Phi^T(p,t)\Phi(p,t)\mathbf{V}(p,t) - \mathbf{V}^T(p,t)\Phi^T(p,t)\Phi(p,t)\tilde{\mathbf{Y}}(p,t) \\ &\quad + \mathbf{V}^T(p,t)\Phi^T(p,t)\Phi(p,t)\mathbf{V}(p,t)].\end{aligned}\quad (2.3.12)$$

定义随机 Lyapunov 函数

$$T(t) := \|\tilde{\theta}(t)\|^2 = \text{tr}[\tilde{\theta}^T(t)\tilde{\theta}(t)].$$

由 (2.3.12) 式得到

$$\begin{aligned}T(t) &\leq T(t-1) - \frac{2}{r(t)}\|\tilde{\mathbf{Y}}(p,t)\|^2 + \frac{2}{r(t)}\text{tr}[\tilde{\mathbf{Y}}^T(p,t)\mathbf{V}(p,t)] \\ &\quad + \frac{1}{r^2(t)}\{\|\Phi(p,t)\tilde{\mathbf{Y}}(p,t)\|^2 + \|\Phi(p,t)\mathbf{V}(p,t)\|^2 \\ &\quad - 2\text{tr}[\tilde{\mathbf{Y}}^T(p,t)\Phi^T(p,t)\Phi(p,t)\mathbf{V}(p,t)]\} \\ &= T(t-1) - \left[\frac{2}{r(t)} - \frac{1}{r^2(t)}\|\Phi(p,t)\|^2 \right] \|\tilde{\mathbf{Y}}(p,t)\|^2 \\ &\quad + \frac{1}{r^2(t)}\|\Phi(p,t)\mathbf{V}(p,t)\|^2 + \frac{2}{r(t)}\text{tr}[\tilde{\mathbf{Y}}^T(p,t)\mathbf{V}(p,t)] \\ &\quad - \frac{2}{r^2(t)}\text{tr}[\tilde{\mathbf{Y}}^T(p,t)\Phi^T(p,t)\Phi(p,t)\mathbf{V}(p,t)]\end{aligned}$$

$$\begin{aligned} &\leq T(t-1) - \frac{1}{r(t)} \|\tilde{\mathbf{Y}}(p, t)\|^2 + \frac{1}{r^2(t)} \left(\sum_{i=0}^{p-1} \|\boldsymbol{\varphi}(t-i)\|^2 \right) \left(\sum_{i=0}^{p-1} \|\mathbf{v}(t-i)\|^2 \right) \\ &\quad + \frac{2}{r(t)} \text{tr}[\tilde{\mathbf{Y}}^T(p, t) \mathbf{V}(p, t)] - \frac{2}{r^2(t)} \text{tr}[\tilde{\mathbf{Y}}^T(p, t) \boldsymbol{\Phi}^T(p, t) \boldsymbol{\Phi}(p, t) \mathbf{V}(p, t)]. \end{aligned} \quad (2.3.13)$$

由于 $T(t-1)$, $\tilde{\mathbf{Y}}(p, t)$, $\boldsymbol{\Phi}(p, t)$ 和 $\boldsymbol{\varphi}(t)$ 与 $\mathbf{V}(p, t)$ 独立并且 \mathcal{F}_{t-1} 是可测的, (2.3.13) 式两边对 \mathcal{F}_{t-1} 取条件期望, 再利用假设 (A1)-(A2) 得到

$$\mathbb{E}[T(t)|\mathcal{F}_{t-1}] \leq T(t-1) - \frac{1}{r(t)} \|\tilde{\mathbf{Y}}(p, t)\|^2 + \frac{p}{r^2(t)} \sum_{i=0}^{p-1} \|\boldsymbol{\varphi}(t-i)\|^2 \sigma^2. \quad (2.3.14)$$

从 $t=1$ 到 $t=\infty$ 累加 (2.3.14) 式右边最后一项, 应用引理 2.3.1 得到

$$\sum_{t=1}^{\infty} \frac{p}{r^2(t)} \sum_{i=0}^{p-1} \|\boldsymbol{\varphi}(t-i)\|^2 \sigma^2 = p\sigma^2 \sum_{t=1}^{\infty} \sum_{i=0}^{p-1} \frac{\|\boldsymbol{\varphi}(t-i)\|^2}{r^2(t)} < \infty.$$

对 (2.3.14) 式应用鞅收敛定理, 则 $T(t)$ 一致收敛于一个有限的随机变量 T_0 , 即,

$$\mathbb{E}[\|\tilde{\boldsymbol{\theta}}(t)\|^2] \rightarrow T_0, \quad t \rightarrow \infty, \quad (2.3.15)$$

和

$$\sum_{t=1}^{\infty} \frac{1}{r(t)} \|\tilde{\mathbf{Y}}(p, t)\|^2 < \infty, \text{ a.s.} \quad \text{或者} \quad \lim_{t \rightarrow \infty} \sum_{i=1}^t \frac{1}{r(i)} \|\tilde{\mathbf{Y}}(p, i)\|^2 < \infty, \text{ a.s.} \quad (2.3.16)$$

应用 Kronecker 引理, 有

$$\frac{1}{r(t)} \sum_{i=0}^t \|\tilde{\mathbf{Y}}(p, i)\|^2 \rightarrow 0, \text{ a.s.}, \quad \text{当 } t \rightarrow \infty. \quad (2.3.17)$$

定义

$$\Delta \tilde{\boldsymbol{\theta}}(t) := \frac{\boldsymbol{\Phi}(p, t)}{r(t)} [-\tilde{\mathbf{Y}}(p, t) + \mathbf{V}(p, t)]. \quad (2.3.18)$$

将 (2.3.18) 式代入 (2.3.10) 式, 有

$$\tilde{\boldsymbol{\theta}}(t+i) = \tilde{\boldsymbol{\theta}}(t-1) + \sum_{k=0}^i \Delta \tilde{\boldsymbol{\theta}}(t+k). \quad (2.3.19)$$

对 (2.3.11) 式两边左乘 $\boldsymbol{\Phi}(p, t)$, 并用 $t+i$ 代替 t 得到

$$\boldsymbol{\Phi}(p, t+i) \boldsymbol{\Phi}^T(p, t+i) \tilde{\boldsymbol{\theta}}(t+i-1) = \boldsymbol{\Phi}(p, t+i) \tilde{\mathbf{Y}}(p, t+i). \quad (2.3.20)$$

将 (2.3.19) 式代入 (2.3.20) 式得到

$$\boldsymbol{\Phi}(p, t+i) \boldsymbol{\Phi}^T(p, t+i) \tilde{\boldsymbol{\theta}}(t-1)$$

$$= \Phi(p, t+i) \tilde{Y}(p, t+i) - \Phi(p, t+i) \Phi^T(p, t+i) \sum_{k=0}^{i-1} \Delta \tilde{\theta}(t+k).$$

从 $i=0$ 到 $i=N-1$ 累加上式得到

$$\begin{aligned} & \left[\sum_{i=0}^{N-1} \Phi(p, t+i) \Phi^T(p, t+i) \right] \tilde{\theta}(t-1) \\ &= \sum_{i=0}^{N-1} \Phi(p, t+i) \tilde{Y}(p, t+i) - \sum_{i=0}^{N-1} \Phi(p, t+i) \Phi^T(p, t+i) \sum_{k=0}^{i-1} \Delta \tilde{\theta}(t+k). \end{aligned} \quad (2.3.21)$$

由 (A4), 可以得到

$$\alpha \mathbf{I} \leq \frac{1}{N} \sum_{i=0}^{N-1} \varphi(t+i-k) \varphi^T(t+i-k) \leq \beta \mathbf{I}, \text{ a.s.},$$

从 $k=0$ 到 $k=p-1$ 累加得到

$$p\alpha \mathbf{I} \leq \frac{1}{N} \sum_{k=0}^{p-1} \sum_{i=0}^{N-1} \varphi(t+i-k) \varphi^T(t+i-k) \leq p\beta \mathbf{I}, \text{ a.s.}$$

上式两边取迹, 有

$$pn\alpha N \leq \sum_{i=0}^{N-1} \sum_{k=0}^{p-1} \|\varphi(t+i-k)\|^2 = \sum_{i=0}^{N-1} \|\Phi(p, t+i)\|^2 \leq pn\beta N, \quad (2.3.22)$$

$$\|\Phi(p, t+i)\|^2 \leq pn\beta N. \quad (2.3.23)$$

这里应用了一个事实

$$\frac{1}{N} \sum_{k=0}^{p-1} \sum_{i=0}^{N-1} \varphi(t+i-k) \varphi^T(t+i-k) = \frac{1}{N} \sum_{i=0}^{N-1} \Phi(p, t+i) \Phi^T(p, t+i).$$

对 (2.3.21) 式两边取范数并利用 (2.3.22) 式和 (2.3.23) 式, 有

$$\begin{aligned} \|\tilde{\theta}(t-1)\| &\leq \frac{1}{Np\alpha n} \left\| \sum_{i=0}^{N-1} \Phi(p, t+i) \tilde{Y}(p, t+i) \right\| \\ &\quad + \frac{1}{Np\alpha n} \left\| \sum_{i=0}^{N-1} \Phi(p, t+i) \Phi^T(p, t+i) \sum_{k=0}^{i-1} \Delta \tilde{\theta}(t+k) \right\| \\ &\leq \frac{1}{Np\alpha n} \left\| \sum_{i=0}^{N-1} \Phi(p, t+i) \tilde{Y}(p, t+i) \right\| + \frac{\beta}{\alpha} \sum_{i=0}^{N-1} \sum_{k=0}^{i-1} \|\Delta \tilde{\theta}(t+k)\|. \end{aligned}$$

对上面不等式两边取平方, 并应用公式 $(a_1 + a_2 + \cdots + a_n)^2 \leq n(a_1^2 + a_2^2 + \cdots + a_n^2)$, 得到

$$\|\tilde{\theta}(t-1)\|^2 \leq \frac{2}{(Np\alpha n)^2} \left\| \sum_{i=0}^{N-1} \Phi(p, t+i) \tilde{Y}(p, t+i) \right\|^2$$

$$\begin{aligned}
& +2 \left(\frac{\beta}{\alpha} \right)^2 \left[\sum_{i=0}^{N-1} \sum_{k=0}^{i-1} \|\Delta \tilde{\theta}(t+k)\| \right]^2 \\
& \leq \frac{2}{(Np\alpha n)^2} N \sum_{i=0}^{N-1} \|\Phi(p, t+i)\|^2 \|\tilde{Y}(p, t+i)\|^2 \\
& \quad +2 \left(\frac{\beta}{\alpha} \right)^2 N^2 \sum_{i=0}^{N-1} \sum_{k=0}^{i-1} \|\Delta \tilde{\theta}(t+k)\|^2 \\
& \leq \frac{2}{(Np\alpha n)^2} N \sum_{i=0}^{N-1} Nnp\beta \|\tilde{Y}(p, t+i)\|^2 + 2 \left(\frac{N\beta}{\alpha} \right)^2 \sum_{i=0}^{N-1} \sum_{k=0}^{i-1} \|\Delta \tilde{\theta}(t+k)\|^2 \\
& =: k_1 \sum_{i=0}^{N-1} \|\tilde{Y}(p, t+i)\|^2 + k_2 \sum_{i=0}^{N-1} \sum_{k=0}^{i-1} \|\Delta \tilde{\theta}(t+k)\|^2, \tag{2.3.24}
\end{aligned}$$

其中

$$k_1 := \frac{2\beta}{np\alpha^2} < \infty, \quad k_2 := 2 \left(\frac{N\beta}{\alpha} \right)^2 < \infty.$$

用 l 代替 (2.3.24) 式中的 t , 两边同时除以 $r(l)$, 并从 $l = l_0$ ($l_0 < \infty$) 到 $l = t$ 累加得到

$$\sum_{l=l_0}^t \frac{\|\tilde{\theta}(l-1)\|^2}{r(l)} \leq k_1 \sum_{i=0}^{N-1} \sum_{l=l_0}^t \frac{\|\tilde{Y}(p, l+i)\|^2}{r(l)} + k_2 \sum_{i=0}^{N-1} \sum_{k=0}^{i-1} \sum_{l=l_0}^t \frac{\|\Delta \tilde{\theta}(l+k)\|^2}{r(l)}. \tag{2.3.25}$$

根据 $r(t)$ 的定义如 (2.3.6) 式, 有

$$r(t) = r(t-p) + \sum_{i=0}^{p-1} \|\varphi(t-i)\|^2 = r(t-p) + \|\Phi(p, t)\|^2,$$

$$\frac{\|\Phi(p, t)\|^2}{r(t)} = \sum_{i=0}^{p-1} \frac{\|\varphi(t-i)\|^2}{r(t)} \leq 1. \tag{2.3.26}$$

利用 (2.3.18) 式和 (2.3.26) 式, 有

$$\begin{aligned}
\sum_{l=l_0}^t \frac{\|\Delta \tilde{\theta}(l+k)\|^2}{r(l)} & = \sum_{l=l_0}^t \left\| \frac{\Phi(p, l+k)}{r(l+k)} [-\tilde{Y}(p, l+k) + \mathbf{V}(p, l+k)] \right\|^2 \\
& \leq \sum_{l=l_0}^t \frac{2}{r(l)} \frac{\|\Phi(p, l+k)\|^2}{r^2(l+k)} \left[\|\tilde{Y}(p, l+k)\|^2 + \|\mathbf{V}(p, l+k)\|^2 \right] \\
& \leq \sum_{l=l_0}^t \frac{2}{r(l)r(l+k)} \left[\|\tilde{Y}(p, l+k)\|^2 + \|\mathbf{V}(p, l+k)\|^2 \right] \\
& \leq \sum_{l=l_0}^t \frac{2\|\tilde{Y}(p, l+k)\|^2}{r(l+k)} + \sum_{l=l_0}^t \frac{2\|\mathbf{V}(p, l+k)\|^2}{r(l+k)}.
\end{aligned}$$

应用假设 (A3) 和 (2.3.16) 式得到

$$\lim_{t \rightarrow \infty} \sum_{l=l_0}^t \frac{\|\Delta \tilde{\theta}(l+k)\|^2}{r(l)} < \infty, \text{ a.s.} \quad (2.3.27)$$

由 (2.3.25) 式, (2.3.16) 式和 (2.3.27) 式, 有

$$\lim_{t \rightarrow \infty} \sum_{l=l_0}^t \frac{\|\tilde{\theta}(l-1)\|^2}{r(l)} < \infty, \text{ a.s.}$$

应用 Kronecker 引理得到

$$\lim_{t \rightarrow \infty} \frac{1}{r(t)} \sum_{l=l_0}^t \|\tilde{\theta}(l-1)\|^2 = 0, \text{ a.s.}$$

由于 $T(t) = \|\tilde{\theta}(t)\|^2$ 一致收敛于有限的随机变量 T_0 , 应用引理 2.3.2 得到

$$\lim_{t \rightarrow \infty} \|\tilde{\theta}(t)\|^2 = 0, \text{ a.s.}$$

定理 2.3.2 证明完毕. □

2.3.3 仿真试验

例 2.3.1 考虑双输入双输出仿真系统,

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} + \begin{bmatrix} 0.70 & 0.40 \\ -0.30 & 0.80 \end{bmatrix} \begin{bmatrix} y_1(t-1) \\ y_2(t-1) \end{bmatrix} = \begin{bmatrix} 1.00 & -0.50 \\ -0.40 & 1.20 \end{bmatrix} \begin{bmatrix} u_1(t-1) \\ u_2(t-1) \end{bmatrix} + \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix},$$

其中参数矩阵和信息矩阵为

$$\begin{aligned} \theta^T &= \begin{bmatrix} a_{11} & a_{12} & b_{11} & b_{12} \\ a_{21} & a_{22} & b_{21} & b_{22} \end{bmatrix} \\ &= \begin{bmatrix} 0.70 & 0.40 & 1.00 & -0.50 \\ -0.30 & 0.80 & -0.40 & 1.20 \end{bmatrix}, \\ \varphi(t) &= [-y_1(t-1), -y_2(t-1), u_1(t-1), u_2(t-1)]^T. \end{aligned}$$

仿真时, $\{u_1(t)\}$ 和 $\{u_2(t)\}$ 采用零均值单位方差不相关可测的持续激励信号序列, $\{v_1(t)\}$ 和 $\{v_2(t)\}$ 采用零均值方差分别为 $\sigma_1^2 = 1.00^2$ 和 $\sigma_2^2 = 1.00^2$ 的白噪声序列. 用 SG 算法和 MISG 算法估计这个系统的参数. 不同新息长度下的参数估计及参数估计误差如表 2.1-2.4 所示, 其中参数估计误差 $\delta := \|\hat{\theta}(t) - \theta\| / \|\theta\|$ 随 t 变化曲线如图 2.2 所示, 取 $p = 1, 2, 5$ 和 10. 对应两个输出通道的噪信比分别为 $\delta_{\text{ns}}(1) = 91.26\%$ 和 $\delta_{\text{ns}}(2) = 73.42\%$.

表 2.1 SG 参数估计及其误差
Table 2.1 The SG parameter estimates and errors

t	100	500	1000	2000	3000	4000	5000
$a_{11} = 0.70000$	0.52633	0.61067	0.62313	0.63777	0.64511	0.65123	0.65524
$a_{12} = 0.40000$	0.42564	0.39304	0.40959	0.41103	0.40656	0.40572	0.40542
$b_{11} = 1.00000$	0.15459	0.23636	0.26667	0.29385	0.30984	0.32119	0.33121
$b_{12} = -0.50000$	-0.18140	-0.19657	-0.20972	-0.22003	-0.22619	-0.23070	-0.23377
$a_{21} = -0.30000$	-0.19429	-0.25152	-0.25752	-0.26941	-0.27553	-0.27742	-0.27840
$a_{22} = 0.80000$	0.72265	0.75444	0.75463	0.77146	0.77680	0.77877	0.77912
$b_{21} = -0.40000$	0.03085	-0.00839	-0.02471	-0.03839	-0.04622	-0.05117	-0.05610
$b_{22} = 1.20000$	0.42725	0.48015	0.50806	0.53931	0.55292	0.56301	0.57031
δ (%)	62.39826	56.68550	54.39377	52.14216	50.98014	50.15610	49.48603

表 2.2 MISG 参数估计及其误差 ($p = 2$)
Table 2.2 The MISG parameter estimates and errors ($p = 2$)

t	100	500	1000	2000	3000	4000	5000
$a_{11} = 0.70000$	0.65433	0.69131	0.67982	0.68070	0.68281	0.68643	0.68841
$a_{12} = 0.40000$	0.45138	0.39330	0.41301	0.41211	0.40341	0.40179	0.40232
$b_{11} = 1.00000$	0.38175	0.49810	0.53746	0.57129	0.59032	0.60390	0.61641
$b_{12} = -0.50000$	-0.41931	-0.41006	-0.42011	-0.42290	-0.42565	-0.42835	-0.42955
$a_{21} = -0.30000$	-0.24792	-0.29440	-0.28929	-0.29531	-0.29866	-0.29800	-0.29661
$a_{22} = 0.80000$	0.81266	0.78446	0.78102	0.79415	0.79742	0.79663	0.79372
$b_{21} = -0.40000$	-0.10448	-0.15506	-0.17655	-0.19259	-0.20096	-0.20593	-0.21160
$b_{22} = 1.20000$	0.80307	0.84531	0.87243	0.90343	0.91367	0.92229	0.92804
δ (%)	38.93459	32.47791	29.91736	27.55976	26.42100	25.59078	24.87554

表 2.3 MISG 参数估计及其误差 ($p = 5$)
Table 2.3 The MISG parameter estimates and errors ($p = 5$)

t	100	500	1000	2000	3000	4000	5000
$a_{11} = 0.70000$	0.69481	0.69586	0.67022	0.68588	0.68177	0.69430	0.69621
$a_{12} = 0.40000$	0.43367	0.39604	0.41442	0.41091	0.40184	0.39450	0.39914
$b_{11} = 1.00000$	0.65165	0.80570	0.84165	0.86861	0.87690	0.89259	0.90469
$b_{12} = -0.50000$	-0.58537	-0.50666	-0.51324	-0.50268	-0.50244	-0.50109	-0.49958
$a_{21} = -0.30000$	-0.23400	-0.31131	-0.28870	-0.29840	-0.30553	-0.29929	-0.29647
$a_{22} = 0.80000$	0.80981	0.76704	0.79004	0.79851	0.79235	0.80043	0.79151
$b_{21} = -0.40000$	-0.29001	-0.33597	-0.35292	-0.36014	-0.35856	-0.36160	-0.36542
$b_{22} = 1.20000$	1.25323	1.19841	1.19813	1.20816	1.20314	1.19824	1.19583
δ (%)	18.77942	10.09957	8.24996	6.74593	6.39756	5.56046	4.95748

表 2.4 MISG 参数估计及其误差 ($p = 10$)
Table 2.4 The MISG parameter estimates and errors ($p = 10$)

t	100	500	1000	2000	3000	4000	5000
$a_{11} = 0.70000$	0.70017	0.68675	0.65505	0.69749	0.69897	0.69999	0.70019
$a_{12} = 0.40000$	0.41471	0.41294	0.41929	0.40698	0.38927	0.39159	0.40043
$b_{11} = 1.00000$	0.65459	0.90240	0.93338	0.95274	0.96123	0.96967	0.98336
$b_{12} = -0.50000$	-0.59420	-0.45994	-0.48727	-0.47762	-0.47764	-0.48270	-0.48195
$a_{21} = -0.30000$	-0.21280	-0.31309	-0.27209	-0.29870	-0.30002	-0.29730	-0.29448
$a_{22} = 0.80000$	0.78866	0.75660	0.80406	0.79557	0.81290	0.80198	0.78474
$b_{21} = -0.40000$	-0.34456	-0.38765	-0.40179	-0.39796	-0.39407	-0.38861	-0.39258
$b_{22} = 1.20000$	1.40696	1.23480	1.22409	1.23296	1.21068	1.20744	1.20156
δ (%)	20.74499	5.93343	4.44864	3.03682	2.39902	1.87428	1.47766

表 2.5 RLS 算法参数估计及其误差
Table 2.5 The RLS parameter estimates and errors

t	100	500	1000	2000	3000	4000	5000
$a_{11} = 0.70000$	0.68567	0.69456	0.67988	0.68394	0.68672	0.69013	0.69207
$a_{12} = 0.40000$	0.44900	0.39513	0.40616	0.40804	0.40090	0.39890	0.40064
$b_{11} = 1.00000$	0.90345	1.02848	1.01228	1.00473	1.00129	1.00723	1.02257
$b_{12} = -0.50000$	-0.44788	-0.38020	-0.46659	-0.46210	-0.46625	-0.47665	-0.47767
$a_{21} = -0.30000$	-0.25753	-0.29303	-0.29060	-0.29664	-0.29950	-0.29901	-0.29727
$a_{22} = 0.80000$	0.79673	0.77308	0.78300	0.79379	0.79739	0.79690	0.79334
$b_{21} = -0.40000$	-0.33106	-0.41144	-0.41960	-0.40297	-0.39511	-0.38560	-0.39345
$b_{22} = 1.20000$	1.38781	1.16856	1.18120	1.22071	1.18928	1.18886	1.18309
δ (%)	11.55553	6.35976	2.58414	2.31598	1.86039	1.56626	1.85258

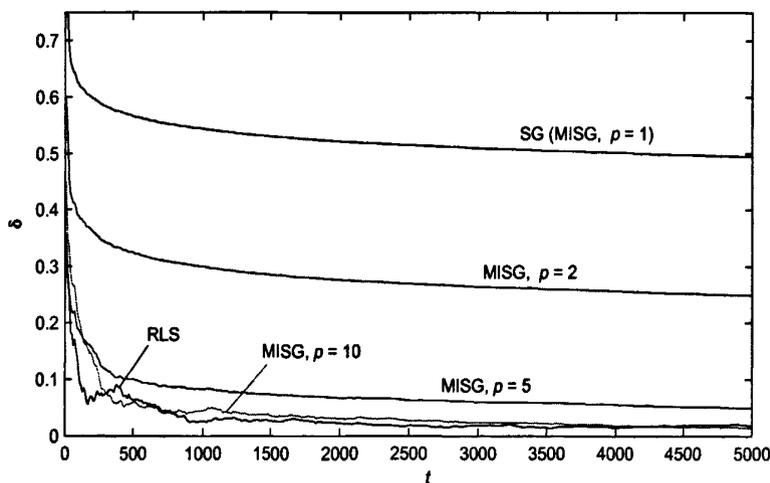


图 2.2 RLS 和 MISG 参数估计误差 δ 随 t 变化曲线
Fig. 2.2 The RLS and MISG parameter estimation errors δ versus t

为了比较, 图 2.2 给出了 RLS 算法的曲线, 表 2.5 给出了 RLS 算法的参数估计和误差. 从表 2.1-2.5 以及图 2.2, 可以得到如下结论:

- 随着 t 的增加, MISG 算法的参数估计误差越来越小并趋于零.
- 当 $p \geq 2$ 时, MISG 算法比 SG 算法有更快的收敛速度和更高的估计精度.
- 随着新息长度 p 的增加, MISG 算法的参数估计越来越接近 RLS 算法的参数估计.

2.4 本章小结

由于多率多输入系统经过提升或离散化后得到多变量系统, 因此本章研究了多输入多输出系统的辨识问题. 基于多新息辨识理论, 引入新息长度, 将单新息向量扩展为多新息矩阵, 推导了多新息随机梯度算法辨识多输入多输出系统. 由于多新息随机梯度算法可以充分利用输入输出数据, 因此提高了随机梯度算法的收敛速度, 同时提高了参数估计精度. 应用鞅收敛定理分析了多新息随机梯度算法的收敛性, 并用仿真例子验证了算法的有效性.

本章的内容发表在国际刊物《Digital Signal Processing》2009 年第 4 期上.

第三章 方程误差多率系统

3.1 引言

对于多变量系统,当每个输入通道或输出通道的采样周期不相同,就得到了一个多率系统.多率系统广泛存在于过程工业中,如在化工过程控制中,某些化工产品的质量数据(如炼油厂汽车辛烷值)不能直接在线量测,需要试验室化验分析得到,通常需要几分钟,甚至是十几分钟,而控制信号的刷新频率相对较快.如这类化工控制中的许多软测量问题,都可以归结为多率系统的建模,参数辨识或状态估计问题^[1].

对于多率系统的辨识,目前人们主要研究双率系统的辨识.一般的多率系统由于存在两种以上的操作频率,使得其结构要比双率系统的结构复杂得多,其辨识问题的复杂度高,计算量大. Shi 等将多率多输入串音系统转化为多率多输入 FIR 模型,基于辅助模型辨识思想,利用 RLS 算法辨识系统的参数^[9]. Sahebsara 等将多输入多输出多率系统分解为几个双率系统,将多输入多输出多率系统的辨识问题转化为双率系统的辨识问题,应用多项式分解技术,利用 RLS 算法得到系统的参数估计^[3].然而,对于随机框架下的多率多输入系统的辨识问题还需进一步的研究.

本章研究多率多输入方程误差系统的辨识问题.组织如下:第 3.2 节采用离散化技术,建立多率输入与输出数据间的映射关系,推导多率系统的辨识模型.第 3.3 节推导随机梯度算法,为了提高算法的收敛速度和参数估计精度,引入遗忘因子,得到遗忘梯度算法.第 3.4 节基于多新息辨识理论,提出多新息随机梯度算法.应用鞅收敛定理分析了多新息随机梯度算法的收敛性.第 3.5 节应用递阶辨识原理,提出递阶最小二乘算法估计系统的参数.第 3.6 节为本章总结.

3.2 模型推导

本章研究一类多率多输入系统如图 3.1 所示,离散时间输入信号 $u_j(kT_j)$ 经过零阶保持器

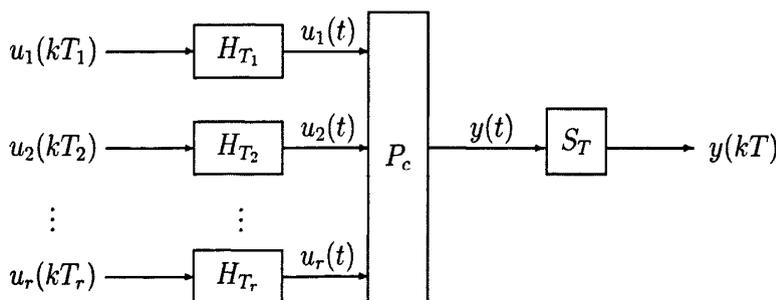


图 3.1 多率多输入系统

Fig. 3.1 The multirate multiple-input systems

H_{T_j} 得到连续信号 $u_j(t)$, 经过连续时间过程 P_c 产生系统的输出 $y(t)$, 在经过输出采样器 S_T 得到离散信号 $y(kT)$, 其中 $T_j := p_j h$ 为第 j 个输入刷新周期. 为了研究方便, 假定 p_1, p_2, \dots, p_r 是互质的正整数, q 是 p_1, p_2, \dots, p_r 的最小公倍数. 其中, h 为基周期, $T := qh$ 为框架周期^[31].

在第 k 个框架周期上零阶保持器 H_{T_j} ($j = 1, 2, \dots, r$) 对控制输入 u_j 均匀的刷新, 则零阶保持器 H_{T_j} 的特性可表示为

$$u_j(t) = u_j(kT_j), \quad kT_j \leq t < (k+1)T_j, \quad j = 1, 2, \dots, r. \quad (3.2.1)$$

输出采样器 S_T 的采样周期等于输入的框架周期. 这样的刷新和采样方案如图 3.2 所示.

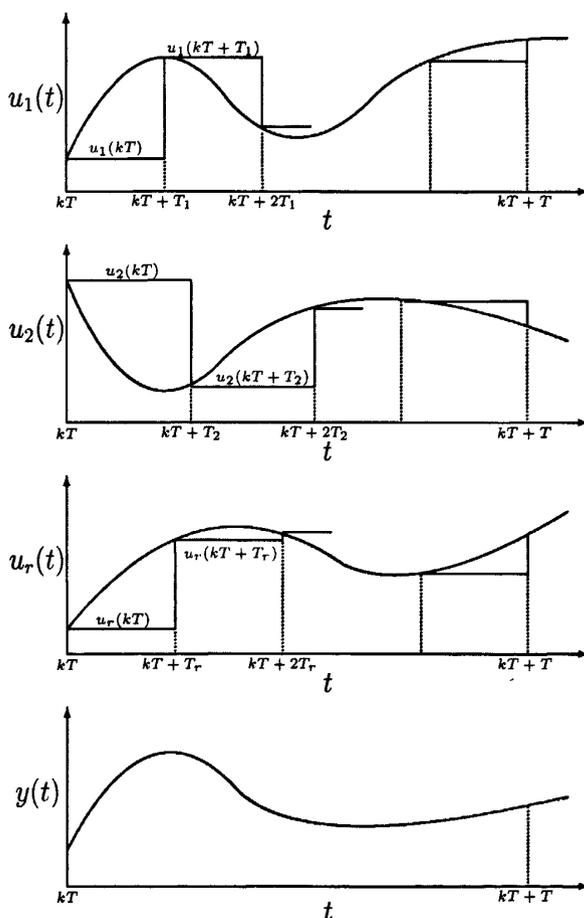


图 3.2 输入周期刷新和输出周期采样方案

Fig. 3.2 The input updating and output sampling patterns

对于如图 3.1 所示的多率系统, 可测的输入输出数据为

$$\{u_1(kT_1), u_2(kT_2), \dots, u_r(kT_r), y(kT) : k = 1, 2, \dots\}$$

例如, 对于双输入单输出的多率系统, 如果 $T_1 = 2$ s, $T_2 = 3$ s, 则其可测的输入输出数据为

$$\{u_1(0), u_1(2), u_1(4), u_1(6), \dots, u_2(0), u_2(3), u_2(6), \dots, y(0), y(6), y(12), \dots\}.$$

图 3.1 所示的多率系统是一个时变系统. 即使连续系统 P_c 是一个时不变系统, 由于不同的刷新周期和采样周期, 使得 $\{u_1(kT_1), u_2(kT_2), \dots, u_r(kT_r)\}$ 到 $y(kT)$ 是时变系统 [38, 39, 93]. 对于图 3.1 所示的时变多率系统, 这里采用离散化技术, 在框架周期 T 上得到相应的一个时不变多输入离散时间状态空间模型.

设图 3.1 中的连续过程 P_c 具有下列状态空间模型,

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}_c \mathbf{x}(t) + \sum_{j=1}^r \mathbf{B}_{c_j} u_j(t), \\ y(t) = \mathbf{C} \mathbf{x}(t), \end{cases} \quad (3.2.2)$$

其中 $\mathbf{x}(t) \in \mathbb{R}^n$ 为状态向量, $u_j(t)$ ($j = 1, 2, \dots, r$) 为第 j 个输入通道的控制输入, $y(t) \in \mathbb{R}^1$ 为系统的输出, \mathbf{A}_c , \mathbf{B}_{c_j} 和 \mathbf{C} 为适当维数的矩阵. 参考文 [48], 以周期 h 离散化 P_c 可得

$$\begin{cases} \mathbf{x}(kh + h) = \mathbf{A}_d \mathbf{x}(kh) + \sum_{j=1}^r \mathbf{B}_{d_j} u_j(kh), \\ y(kh) = \mathbf{C} \mathbf{x}(kh), \end{cases} \quad (3.2.3)$$

其中矩阵 \mathbf{A}_d 和 \mathbf{B}_{d_j} 有如下表达式,

$$\mathbf{A}_d := e^{\mathbf{A}_c h}, \quad \mathbf{B}_{d_j} := \int_0^h e^{\mathbf{A}_c t} dt \mathbf{B}_{c_j}.$$

模型 (3.2.3) 中存在不可测的输入输出数据 $u_j(kh)$ ($k \neq p_j$) 和 $y(kh)$ ($k \neq q$), 因此不能作为辨识模型. 基于模型 (3.2.3), 利用离散化技术, 可以建立可测量输入 $\{u_1(kT_1), u_2(kT_2), \dots, u_r(kT_r)\}$ 到可测量输出 $y(kT)$ 的数学模型.

定理 3.2.1 对于图 3.1 所示的多率多输入系统, 令 $\nu_j := q/p_j$, 它的状态空间模型可表达为如下

$$\begin{cases} \mathbf{x}(kT + T) = \mathbf{A} \mathbf{x}(kT) + \sum_{j=1}^r \sum_{s=1}^{\nu_j} \mathbf{B}_{j_s} u_j(kT + (s-1)T_j), \\ y(kT) = \mathbf{C} \mathbf{x}(kT), \end{cases} \quad (3.2.4)$$

其中

$$\mathbf{A} := e^{\mathbf{A}_c T} = \mathbf{A}_d^q \in \mathbb{R}^{n \times n},$$

$$\mathbf{B}_{j_s} := \sum_{i=1}^{p_j} \mathbf{A}_d^{q-(s-1)p_j-i} \mathbf{B}_{d_j} \in \mathbb{R}^n, \quad j = 1, 2, \dots, r, \quad s = 1, 2, \dots, \nu_j.$$

证明: 将 (3.2.3) 中的 k 用 kq 来代替, 有

$$\begin{cases} \mathbf{x}(kT+h) = \mathbf{A}_d \mathbf{x}(kT) + \sum_{j=1}^r \mathbf{B}_{dj} u_j(kT), \\ y(kT) = \mathbf{C} \mathbf{x}(kT). \end{cases}$$

则

$$\begin{aligned} \mathbf{x}(kT+sh) &= \mathbf{A}_d^s \mathbf{x}(kT) + \sum_{j=1}^r \sum_{i=1}^s \mathbf{A}_d^{s-i} \mathbf{B}_{dj} u_j(kT+(i-1)h), \\ \mathbf{x}(kT+T) &= \mathbf{A}_d^q \mathbf{x}(kT) + \sum_{j=1}^r \sum_{s=0}^{\nu_j-1} \sum_{i=1}^{p_j} \mathbf{A}_d^{q-sp_j-i} \mathbf{B}_{dj} u_j(kT+sT_j+(i-1)h). \end{aligned}$$

应用 (3.2.1) 式可得

$$\begin{aligned} u_j(kT+sT_j) &= u_j(kT+sT_j+h) = \cdots = u_j(kT+sT_j+(p_j-1)h), \\ k &= 0, 1, 2, \cdots, s = 1, 2, \cdots, \nu_j-1 \text{ 和 } j = 1, 2, \cdots, r. \end{aligned}$$

因此,

$$\begin{aligned} \mathbf{x}(kT+T) &= \mathbf{A}_d^q \mathbf{x}(kT) + \sum_{j=1}^r \sum_{s=0}^{\nu_j-1} \sum_{i=1}^{p_j} \mathbf{A}_d^{q-sp_j-i} \mathbf{B}_{dj} u_j(kT+sT_j) \\ &= \mathbf{A} \mathbf{x}(kT) + \sum_{j=1}^r \sum_{s=1}^{\nu_j} \sum_{i=1}^{p_j} \mathbf{A}_d^{q-(s-1)p_j-i} \mathbf{B}_{dj} u_j(kT+(s-1)T_j) \\ &= \mathbf{A} \mathbf{x}(kT) + \sum_{j=1}^r \sum_{s=1}^{\nu_j} \mathbf{B}_{js} u_j(kT+(s-1)T_j). \end{aligned}$$

定理 3.2.1 证明完毕. □

3.3 随机梯度算法

3.3.1 算法的提出

引入单位后移算子 z^{-1} [即 $z^{-1}y(kT) = y(kT-T)$]. 从 (3.2.4) 式可得

$$\begin{aligned} y(kT) &= \sum_{j=1}^r \sum_{s=1}^{\nu_j} \mathbf{C} (z\mathbf{I} - \mathbf{A})^{-1} \mathbf{B}_{js} u_j(kT+(s-1)T_j) \\ &= \sum_{j=1}^r \sum_{s=1}^{\nu_j} \frac{\mathbf{C} \operatorname{adj}[z\mathbf{I} - \mathbf{A}] \mathbf{B}_{js}}{\det[z\mathbf{I} - \mathbf{A}]} u_j(kT+(s-1)T_j) \\ &=: \frac{1}{\alpha(z)} \sum_{j=1}^r \sum_{s=1}^{\nu_j} \beta_{js}(z) u_j(kT+(s-1)T_j) \end{aligned}$$

或者

$$\alpha(z)y(kT) = \sum_{j=1}^r \sum_{s=1}^{\nu_j} \beta_{js}(z)u_j(kT + (s-1)T_j), \quad (3.3.1)$$

其中 $\alpha(z)$ 和 $\beta_{js}(z)$ 为带有后移算子 z^{-1} 的多项式,

$$\begin{aligned} \alpha(z) &= z^{-n} \det[z\mathbf{I} - \mathbf{A}] = 1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \cdots + \alpha_n z^{-n}, \quad \alpha_i \in \mathbb{R}^1, \quad i = 1, 2, \cdots, n, \\ \beta_{js}(z) &= z^{-n} \mathbf{C} \operatorname{adj}[z\mathbf{I} - \mathbf{A}] \mathbf{B}_{js} = \beta_{js}(1)z^{-1} + \beta_{js}(2)z^{-2} + \cdots + \beta_{js}(n)z^{-n}, \quad \beta_{js}(l) \in \mathbb{R}^1, \\ & \quad j = 1, 2, \cdots, r, \quad s = 1, 2, \cdots, \nu_j, \quad l = 1, 2, \cdots, n, \end{aligned}$$

其中 α_i 和 $\beta_{js}(l)$ 为待辨识的参数.

一般地, 在实际系统中总会存在干扰, 因此在系统 (3.3.1) 中引入一个随机噪声 $v(t)$, 得到

$$\alpha(z)y(kT) = \sum_{j=1}^r \sum_{s=1}^{\nu_j} \beta_{js}(z)u_j(kT + (s-1)T_j) + v(kT), \quad (3.3.2)$$

其中 $v(kT)$ 为零均值的随机白噪声.

定义参数向量 $\boldsymbol{\theta}$ 和信息向量 $\boldsymbol{\varphi}(kT)$ 如下:

$$\begin{aligned} \boldsymbol{\theta} &:= [\alpha_1, \cdots, \alpha_n, \beta_{11}(1), \cdots, \beta_{11}(n), \cdots, \beta_{1\nu_1}(1), \cdots, \beta_{1\nu_1}(n), \cdots, \\ & \quad \beta_{r1}(1), \cdots, \beta_{r1}(n), \cdots, \beta_{r\nu_r}(1), \cdots, \beta_{r\nu_r}(n)]^T \in \mathbb{R}^{n_0}, \\ n_0 &:= n + n(\nu_1 + \nu_2 + \cdots + \nu_r), \\ \boldsymbol{\varphi}(kT) &:= [-y(kT - T), \cdots, -y(kT - nT), \phi_1^T(kT), \cdots, \phi_r^T(kT)]^T \in \mathbb{R}^{n_0}, \\ \phi_j(kT) &:= [u_j(kT - T), \cdots, u_j(kT - nT), \cdots, u_j(kT + (\nu_j - 1)T_j - T), \\ & \quad \cdots, u_j(kT + (\nu_j - 1)T_j - nT)]^T \in \mathbb{R}^{n\nu_j}, \quad j = 1, 2, \cdots, r. \end{aligned}$$

则 (3.3.2) 式可以等价的写为一个线性回归模型

$$y(kT) = \boldsymbol{\varphi}^T(kT)\boldsymbol{\theta} + v(kT). \quad (3.3.3)$$

定义准则函数

$$J(\boldsymbol{\theta}) = \sum_{i=1}^k [y(kT) - \boldsymbol{\varphi}^T(kT)\boldsymbol{\theta}]^2,$$

应用梯度搜索原理极小化 $J(\boldsymbol{\theta})$, 得到如下估计参数 $\boldsymbol{\theta}$ 的随机梯度 (MR-SG) 算法:

$$\hat{\boldsymbol{\theta}}(kT) = \hat{\boldsymbol{\theta}}(kT - T) + \frac{\boldsymbol{\varphi}(kT)}{r(kT)} e(kT), \quad (3.3.4)$$

$$e(kT) = y(kT) - \boldsymbol{\varphi}^T(kT)\hat{\boldsymbol{\theta}}(kT - T), \quad (3.3.5)$$

$$r(kT) = r(kT - T) + \|\varphi(kT)\|^2, \quad r(0) = 1, \quad (3.3.6)$$

其中 $e(kT)$ 是标量单新息.

取初始值 $\hat{\theta}(0) = 10^{-6}\mathbf{1}_{n_0}$, 利用 MR-SG 算法计算参数估计 $\hat{\theta}(kT)$ 的步骤如下:

1. 令 $k = 1$, $\hat{\theta}(0) = 10^{-6}\mathbf{1}_{n_0}$, $r(0) = 1$, 并选取数据长度 L_e .
2. 搜集可测量的输入输出数据 $\{u_1(kT_1), u_2(kT_2), \dots, u_r(kT_r), y(kT)\}$, 并由式 (3.3.3) 构成信息向量 $\varphi(kT)$.
3. 由式 (3.3.6) 计算 $r(kT)$, 由式 (3.3.5) 计算 $e(kT)$.
4. 由式 (3.3.4) 刷新 $\hat{\theta}(kT)$.
5. 如果 $k < L_e$, k 增至 $k + 1$, 返回第 2 步继续计算. 否则, 得到参数估计 $\hat{\theta}(L_e T)$.

用 MR-SG 算法计算参数估计 $\hat{\theta}(kT)$ 的流程图如图 3.3 所示.

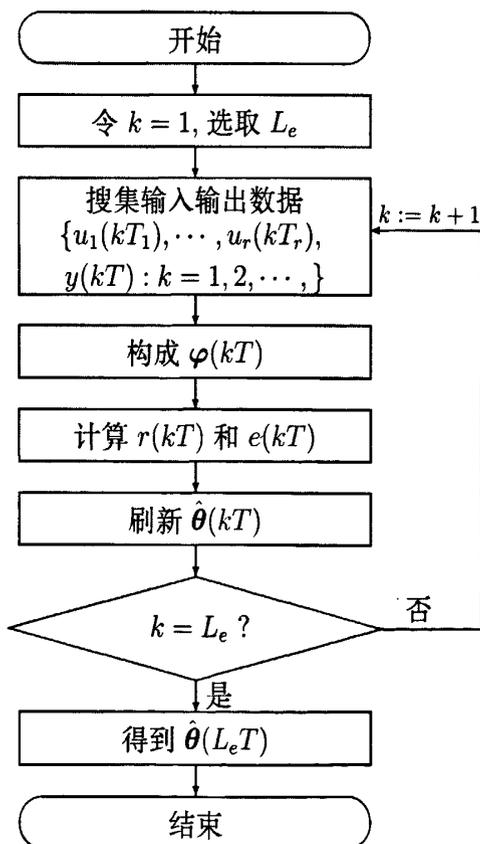


图 3.3 计算参数估计 $\hat{\theta}(kT)$ 的流程图

Fig. 3.3 The flowchart of computing the parameter estimate $\hat{\theta}(kT)$

MR-SG 算法有较小的计算量, 但是其参数估计精度和收敛速度相对较低. 为了提高 MR-SG 算法的跟踪特性, 引入一个遗忘因子 λ , 得到如下的遗忘随机梯度 (MR-FG) 算法:

$$\hat{\theta}(kT) = \hat{\theta}(kT - T) + \frac{\varphi(kT)}{r(kT)} e(kT), \quad (3.3.7)$$

$$e(kT) = y(kT) - \varphi^T(kT) \hat{\theta}(kT - T), \quad (3.3.8)$$

$$r(kT) = \lambda r(kT - T) + \|\varphi(kT)\|^2, \quad 0 \leq \lambda \leq 1, \quad r(0) = 1. \quad (3.3.9)$$

当 $\lambda = 1$ 时, MR-FG 算法就是 MR-SG 算法.

3.3.2 仿真试验

例 3.3.1 考虑双输入单输出多率系统, 刷新周期分别为 $T_1 = 2h$ 和 $T_2 = 3h$, 输出采样周期为 $T = 6h$, 取基周期为 $h = 2$, 相对应的传递函数模型为

$$\begin{aligned} \alpha(z)y(kT) = & \beta_{11}(z)u_1(kT) + \beta_{12}(z)u_1(kT + T_1) + \beta_{13}(z)u_1(kT + 2T_1) \\ & + \beta_{21}(z)u_2(kT) + \beta_{22}(z)u_2(kT + T_2) + v(kT), \end{aligned}$$

其中

$$\begin{aligned} \alpha(z) &= 1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} = 1 + 0.58z^{-1} + 0.36z^{-2}, \\ \beta_{11}(z) &= \beta_{11}(1)z^{-1} + \beta_{11}(2)z^{-2} = -0.20z^{-1} + 0.50z^{-2}, \\ \beta_{12}(z) &= \beta_{12}(1)z^{-1} + \beta_{12}(2)z^{-2} = -0.65z^{-1} + 0.60z^{-2}, \\ \beta_{13}(z) &= \beta_{13}(1)z^{-1} + \beta_{13}(2)z^{-2} = 1.00z^{-1} - 0.50z^{-2}, \\ \beta_{21}(z) &= \beta_{21}(1)z^{-1} + \beta_{21}(2)z^{-2} = -0.30z^{-1} + 0.80z^{-2}, \\ \beta_{22}(z) &= \beta_{22}(1)z^{-1} + \beta_{22}(2)z^{-2} = 0.60z^{-1} + 0.20z^{-2}, \end{aligned}$$

参数向量和信息向量为

$$\begin{aligned} \theta &= [\alpha_1, \alpha_2, \beta_{11}(1), \beta_{11}(2), \beta_{12}(1), \beta_{12}(2), \beta_{13}(1), \beta_{13}(2), \\ & \quad \beta_{21}(1), \beta_{21}(2), \beta_{22}(1), \beta_{22}(2)]^T \\ &= [0.60, 0.40, -0.20, 0.50, -0.65, 0.60, 1.00, -0.50, \\ & \quad -0.40, 0.80, 0.70, 0.20]^T, \\ \varphi(kT) &= [-y(kT - T), -y(kT - 2T), \phi_1^T(kT), \phi_2^T(kT)]^T, \\ \phi_1(kT) &= [u_1(kT - T), u_1(kT - 2T), u_1(kT + T_1 - T), u_1(kT + T_1 - 2T), \\ & \quad u_1(kT + 2T_1 - T), u_1(kT + 2T_1 - 2T)]^T, \\ \phi_2(kT) &= [u_2(kT - T), u_2(kT - 2T), u_2(kT + T_2 - T), u_2(kT + T_2 - 2T)]^T. \end{aligned}$$

仿真时, $\{u_1(kT)\}$ 和 $\{u_2(kT)\}$ 采用零均值单位方差不相关可测的持续激励信号序列, $\{v(kT)\}$ 采用零均值方差为 $\sigma^2 = 0.10^2$ 的白噪声序列. 用 MR-SG 算法和 MR-FG 算法估计这个系统的参数, 在不同遗忘因子下参数估计及参数估计误差如表 3.1-3.3 所示, 其中参数估计误差 $\delta := \|\hat{\theta}(t) - \theta\| / \|\theta\|$ 随 k 变化曲线如图 3.4 所示. 系统的噪信比为 $\delta_{ns} = 28.60\%$.

表 3.1 随机梯度算法的参数估计及其误差
Table 3.1 The MR-SG parameter estimates and their errors

k	100	500	1000	2000	3000	4000	5000
$\alpha_1 = 0.58000$	0.43979	0.46766	0.47871	0.49771	0.50587	0.50925	0.51499
$\alpha_2 = 0.36000$	0.02631	0.11787	0.15795	0.17748	0.19538	0.20565	0.21098
$\beta_{11}(1) = -0.20000$	-0.09215	-0.10721	-0.11069	-0.11461	-0.11589	-0.11629	-0.11740
$\beta_{11}(2) = 0.50000$	0.19005	0.21464	0.22333	0.23141	0.23652	0.23926	0.24271
$\beta_{12}(1) = -0.65000$	-0.32234	-0.34278	-0.35156	-0.36190	-0.36740	-0.37072	-0.37341
$\beta_{12}(2) = 0.60000$	0.33088	0.35379	0.36336	0.37276	0.37935	0.38230	0.38479
$\beta_{13}(1) = 1.00000$	0.21700	0.26992	0.29356	0.31505	0.32831	0.33644	0.34244
$\beta_{13}(2) = -0.50000$	-0.15138	-0.19067	-0.20511	-0.21789	-0.22448	-0.22820	-0.23162
$\beta_{21}(1) = -0.30000$	-0.02224	-0.03158	-0.03802	-0.04707	-0.05125	-0.05388	-0.05682
$\beta_{21}(2) = 0.80000$	0.04129	0.09297	0.11543	0.13550	0.14773	0.15618	0.16245
$\beta_{22}(1) = 0.60000$	0.12640	0.17073	0.18431	0.19594	0.20268	0.20760	0.21126
$\beta_{22}(2) = 0.20000$	-0.01596	0.00278	0.00827	0.01187	0.01404	0.01571	0.01675
δ (%)	72.80197	66.63867	64.16917	62.01685	60.71533	59.90995	59.27976

表 3.2 多率遗忘梯度算法的参数估计及其误差 ($\lambda=0.990$)
Table 3.2 The MR-FG parameter estimates and their errors with $\lambda=0.990$

k	100	500	1000	2000	3000	4000	5000
$\alpha_1 = 0.58000$	0.44980	0.53777	0.58209	0.63734	0.63152	0.61675	0.62172
$\alpha_2 = 0.36000$	0.05003	0.21058	0.33948	0.37246	0.40037	0.39410	0.38129
$\beta_{11}(1) = -0.20000$	-0.09623	-0.12832	-0.14910	-0.17404	-0.18325	-0.18788	-0.19279
$\beta_{11}(2) = 0.50000$	0.19722	0.25269	0.30246	0.36556	0.40826	0.43597	0.45920
$\beta_{12}(1) = -0.65000$	-0.33405	-0.38534	-0.43311	-0.52514	-0.57161	-0.59764	-0.61691
$\beta_{12}(2) = 0.60000$	0.33786	0.39723	0.43920	0.48665	0.52426	0.54018	0.55180
$\beta_{13}(1) = 1.00000$	0.23171	0.37039	0.50262	0.67975	0.79705	0.87399	0.91656
$\beta_{13}(2) = -0.50000$	-0.16014	-0.25304	-0.31510	-0.37208	-0.39757	-0.41304	-0.43160
$\beta_{21}(1) = -0.30000$	-0.02746	-0.06189	-0.10158	-0.17929	-0.22301	-0.24713	-0.26698
$\beta_{21}(2) = 0.80000$	0.05971	0.19201	0.31376	0.47078	0.58011	0.65637	0.70149
$\beta_{22}(1) = 0.60000$	0.13636	0.24122	0.31435	0.41155	0.47675	0.52168	0.54663
$\beta_{22}(2) = 0.20000$	-0.01135	0.03493	0.07352	0.12201	0.15892	0.18401	0.19605
δ (%)	70.93573	56.28265	44.32610	29.19973	19.41479	13.01109	9.12504

表 3.3 多率遗忘梯度算法的参数估计及其误差 ($\lambda=0.970$)

Table 3.3 The MR-FG parameter estimates and their errors with $\lambda=0.970$

k	100	500	1000	2000	3000	4000	5000
$\alpha_1 = 0.58000$	0.47449	0.62489	0.62300	0.60987	0.59677	0.58484	0.58518
$\alpha_2 = 0.36000$	0.10240	0.31563	0.39891	0.37702	0.37177	0.36651	0.36464
$\beta_{11}(1) = -0.20000$	-0.10629	-0.16688	-0.19288	-0.20176	-0.20052	-0.20353	-0.19772
$\beta_{11}(2) = 0.50000$	0.21468	0.32745	0.40642	0.46782	0.48530	0.49867	0.49735
$\beta_{12}(1) = -0.65000$	-0.36180	-0.46767	-0.55074	-0.63725	-0.65020	-0.64182	-0.64748
$\beta_{12}(2) = 0.60000$	0.35372	0.47046	0.51993	0.56258	0.57964	0.59496	0.59583
$\beta_{13}(1) = 1.00000$	0.26875	0.58407	0.79193	0.93960	0.98497	0.99994	0.99937
$\beta_{13}(2) = -0.50000$	-0.18062	-0.34756	-0.40831	-0.44695	-0.47160	-0.48176	-0.49147
$\beta_{21}(1) = -0.30000$	-0.03996	-0.13073	-0.20311	-0.27744	-0.29629	-0.29446	-0.29606
$\beta_{21}(2) = 0.80000$	0.10614	0.38937	0.58238	0.72396	0.77533	0.78971	0.79303
$\beta_{22}(1) = 0.60000$	0.15977	0.37711	0.48808	0.56670	0.59024	0.59625	0.59501
$\beta_{22}(2) = 0.20000$	0.00066	0.11122	0.17988	0.20713	0.20877	0.20736	0.20535
δ (%)	66.45910	36.98295	19.55295	6.72007	2.70979	1.34611	0.83792

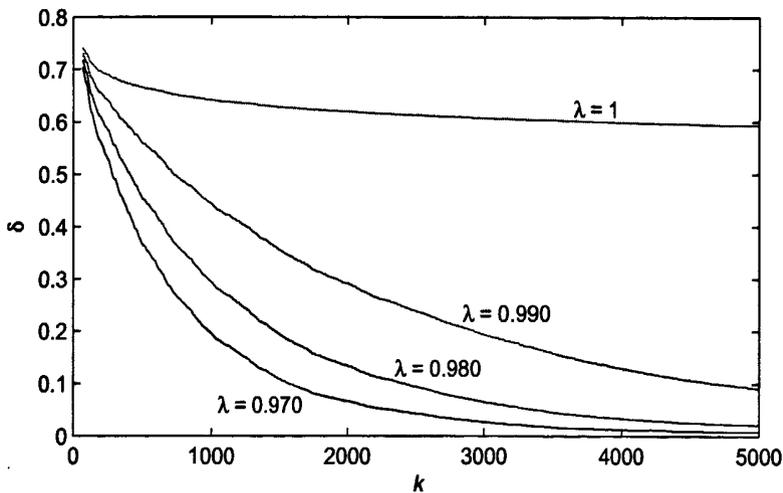


图 3.4 不同遗忘因子 λ 下参数估计误差 δ 随数据长度 k 变化曲线
 Fig. 3.4 The parameter estimation errors δ versus k with different λ

从表 3.1-3.3 和图 3.4 得到如下结论:

- 由 MR-SG 和 MR-FG 算法得到的参数估计误差随着 k 的逐渐增加会越来越小, 并趋于零. 验证了 MR-SG 和 MR-FG 算法是有效的.
- 当 λ 减小时, MR-FG 算法加速了 MR-SG 算法的收敛速度, 但是也会增大参数误差的波动. 因此选取合适的遗忘因子 λ 是得到快速收敛的参数估计和较小的参数估计误差的关键.

3.4 多新息随机梯度算法

在上一节中, 为了提高随机梯度算法的收敛速度和估计精度, 引入遗忘因子, 提出了遗忘梯度算法. 对于遗忘梯度算法, 如果选取的遗忘因子不合适, 会导致参数估计收敛的波动很大. 因此本节应用多新息辨识理论, 将单新息标量扩展为多新息向量, 提出多新息随机梯度算法辨识多率方程误差系统. 并用鞅收敛定理从理论上证明算法的收敛性, 同时用仿真例子验证算法的有效性. 多新息随机梯度算法充分利用可测的多率输入输出信息, 因此, 可以提高随机梯度算法的收敛速度和参数估计精度, 同时避免了参数估计收敛波动大的问题.

3.4.1 算法的提出

为了提高 MR-SG 算法 (3.3.4)–(3.3.6) 的收敛速度和参数估计精度, 基于多新息辨识理论, 引入新息长度, 将单新息标量 $e(kT)$ 扩展为多新息向量

$$\mathbf{E}(p, kT) = \begin{bmatrix} e(kT) \\ e(kT - T) \\ \vdots \\ e(kT - pT + T) \end{bmatrix} \in \mathbb{R}^p,$$

其中 p 是正整数, 且有

$$e(kT - iT) = y(kT - iT) - \boldsymbol{\varphi}^T(kT - iT)\hat{\boldsymbol{\theta}}(kT - iT - T).$$

一般地, 人们认为前一时刻 $kT - T$ 的参数估计 $\hat{\boldsymbol{\theta}}(kT - T)$ 更靠近参数真值 $\boldsymbol{\theta}$, 与更前时刻的参数估计 $\hat{\boldsymbol{\theta}}(kT - iT)$ ($i = 2, 3, \dots, p-1$) 比较, 因此, 将新息向量取作为

$$\mathbf{E}(p, kT) = \begin{bmatrix} y(kT) - \boldsymbol{\varphi}^T(kT)\hat{\boldsymbol{\theta}}(kT - T) \\ y(kT - T) - \boldsymbol{\varphi}^T(kT - T)\hat{\boldsymbol{\theta}}(kT - T) \\ \vdots \\ y(kT - pT + T) - \boldsymbol{\varphi}^T(kT - pT + T)\hat{\boldsymbol{\theta}}(kT - T) \end{bmatrix} \in \mathbb{R}^p.$$

定义信息矩阵 $\Phi(p, kT)$ 和输出向量 $\mathbf{Y}(p, kT)$ 如下:

$$\Phi(p, kT) = [\boldsymbol{\varphi}(kT), \boldsymbol{\varphi}(kT - T), \dots, \boldsymbol{\varphi}(kT - pT + T)] \in \mathbb{R}^{n_0 \times p},$$

$$\mathbf{Y}(p, kT) = [y(kT), y(kT - T), \dots, y(kT - pT + T)]^T \in \mathbb{R}^p.$$

则新息向量 $\mathbf{E}(p, kT)$ 可以表达为

$$\mathbf{E}(p, kT) = \mathbf{Y}(p, kT) - \Phi^T(p, kT)\hat{\boldsymbol{\theta}}(kT - T).$$

参考文献 [26], 得到如下估计多率系统 (3.3.3) 的参数 $\boldsymbol{\theta}$ 的多新息随机梯度 (MR-MISG) 算法:

$$\hat{\boldsymbol{\theta}}(kT) = \hat{\boldsymbol{\theta}}(kT - T) + \frac{\Phi(p, kT)}{r(kT)}\mathbf{E}(p, kT), \quad (3.4.1)$$

$$E(p, kT) = Y(p, kT) - \Phi^T(p, kT)\hat{\theta}(kT - T), \quad (3.4.2)$$

$$r(kT) = r(kT - T) + \|\varphi(kT)\|^2, \quad r(0) = 1, \quad (3.4.3)$$

$$\Phi(p, kT) = [\varphi(kT), \varphi(kT - T), \dots, \varphi(kT - pT + T)], \quad (3.4.4)$$

$$Y(p, kT) = [y(kT), y(kT - T), \dots, y(kT - pT + T)]^T. \quad (3.4.5)$$

当 $p = 1$ 时, MR-MISG 算法 (3.4.1)–(3.4.5) 变为 MR-SG 算法 (3.3.4)–(3.3.6).

与 MR-SG 算法比较, MR-MISG 算法由于充分利用输入输出数据, 因此有更好的收敛性. 初始化 MR-MISG 算法, 取 $\hat{\theta}(0) = 10^{-6}\mathbf{1}_{n_0}$. 利用 MR-MISG 算法计算参数估计 $\hat{\theta}(kT)$ 的步骤如下:

1. 令 $k = 1$, $\hat{\theta}(0) = 10^{-6}\mathbf{1}_{n_0}$, $r(0) = 1$, 选取数据长度 L_e 和新息长度 p .
2. 搜集可测量的输入输出数据 $\{u_1(kT_1), u_2(kT_2), \dots, u_r(kT_r), y(kT) : k = 0, 1, 2, \dots\}$, 并由 (3.3.3) 式构成 $\varphi(kT)$.
3. 由 (3.4.4) 式构成 $\Phi(p, kT)$, 以及由 (3.4.5) 式构成 $Y(p, kT)$.
4. 由 (3.4.3) 式计算 $r(kT)$, 由 (3.4.2) 式计算 $E(p, kT)$.
5. 由 (3.4.1) 式刷新 $\hat{\theta}(kT)$.
6. 如果 $k < L_e$, k 增至 $k + 1$, 返回第 2 步继续计算. 否则, 得到参数估计 $\hat{\theta}(L_e T)$.

用 MR-MISG 算法计算参数估计 $\hat{\theta}(kT)$ 的流程图如图 3.5 所示.

3.4.2 算法收敛性

假设 $\{v(kT), \mathcal{F}_k\}$ 是定义在概率空间 $\{\Omega, \mathcal{F}_k, P\}$ 鞅差向量序列, 其中 $\{\mathcal{F}_k\}$ 是由 $v(kT)$ 生成的 σ 代数序列, 即, $\mathcal{F}_k = \sigma(v(kT), v(kT - T), v(kT - 2T), \dots)$ [13]. 噪声序列 $\{v(kT)\}$ 满足

$$(A1) \quad E[v(kT)|\mathcal{F}_{k-1}] = 0, \text{ a.s.},$$

$$(A2) \quad E[v^2(kT)|\mathcal{F}_{k-1}] \leq \sigma^2, \text{ a.s.},$$

$$(A3) \quad \limsup_{k \rightarrow \infty} \frac{1}{k} \sum_{i=1}^k v^2(iT) \leq \sigma^2 < \infty, \text{ a.s.}$$

引理 3.4.1 对于多率多输入系统 (3.3.3) 和 MR-MISG 算法 (3.4.1)–(3.4.5), 有下面的不等式成立

$$\sum_{k=1}^{\infty} \sum_{i=0}^{p-1} \frac{\|\varphi(kT - iT)\|^2}{r^2(kT)} < \infty, \text{ a.s.}$$

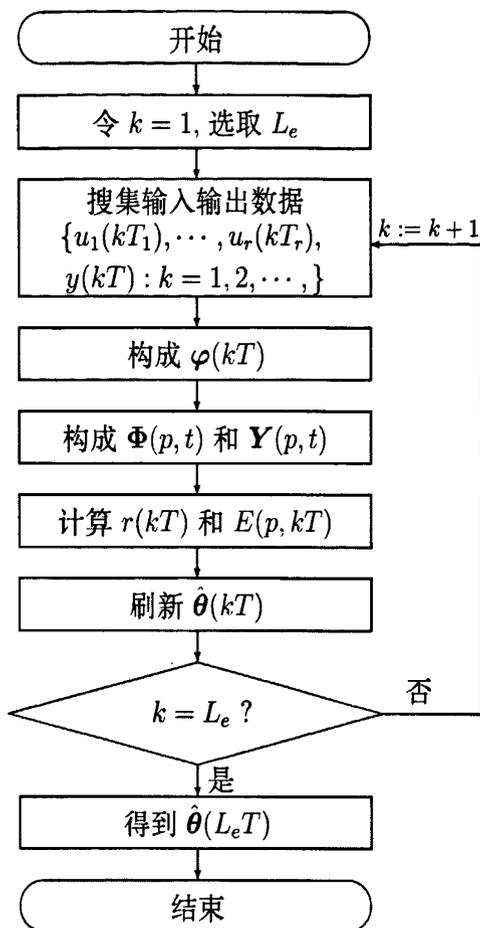


图 3.5 计算参数估计 $\hat{\theta}(kT)$ 的流程图

Fig. 3.5 The flowchart of computing the parameter estimate $\hat{\theta}(kT)$

证明: 由 $r(kT)$ 的定义 (3.4.3) 式, 得到

$$\begin{aligned} \sum_{i=1}^{\infty} \frac{\|\varphi(iT)\|^2}{r(iT-T)r(iT)} &= \sum_{i=1}^{\infty} \frac{r(iT) - r(iT-T)}{r(iT-T)r(iT)} \\ &= \sum_{i=1}^{\infty} \left[\frac{1}{r(iT-T)} - \frac{1}{r(iT)} \right] \\ &= \frac{1}{r(0)} - \frac{1}{r(\infty)} < \infty. \end{aligned}$$

因此,

$$\begin{aligned} \sum_{k=1}^{\infty} \sum_{i=0}^{p-1} \frac{\|\varphi(kT-iT)\|^2}{r^2(kT)} &\leq \sum_{k=1}^{\infty} \sum_{i=0}^{p-1} \frac{\|\varphi(kT-iT)\|^2}{r(kT-iT+T)r(kT-iT)} \\ &= \sum_{i=0}^{p-1} \sum_{k=1}^{\infty} \frac{\|\varphi(kT-iT)\|^2}{r(kT-iT+T)r(kT-iT)} < \infty, \text{ a.s.} \end{aligned}$$

其中 $r(kT) = 1$, 当 $k < 0$ 时. 引理 3.5.1 证明完毕. □

定义参数估计误差向量 $\tilde{\theta}(kT)$ 和噪声向量 $\mathbf{V}(p, kT)$ 如下:

$$\tilde{\theta}(kT) := \hat{\theta}(kT) - \theta, \quad (3.4.6)$$

$$\mathbf{V}(p, kT) := [v(kT), v(kT - T), \dots, v(kT - pT + T)]^T \in \mathbb{R}^p.$$

定理 3.5.1 对于多率多输入系统 (3.3.3) 和 MR-MISG 算法 (3.4.1)–(3.4.5), 假设 (A1)–(A3) 成立, 并且信息向量 $\varphi(kT)$ 是持续激励的, 即, 存在实数 a 和 b 满足 $0 < a \leq b < \infty$ 和整数 N 使得强持续激励条件成立

$$(A4) \quad a\mathbf{I} \leq \frac{1}{N} \sum_{i=0}^{N-1} \varphi(kT + iT)\varphi^T(kT + iT) \leq b\mathbf{I}, \quad \text{a.s.}, \quad k > 0.$$

则参数估计误差 $\tilde{\theta}(kT)$ 一致趋于零, 即,

$$\lim_{k \rightarrow \infty} \tilde{\theta}(kT) = \mathbf{0}, \quad \text{a.s.}$$

证明: 定义随机 Lyapunov 函数

$$Q(kT) := \|\tilde{\theta}(kT)\|^2 = \tilde{\theta}^T(kT)\tilde{\theta}(kT).$$

将 (3.3.3) 式和 (3.4.1) 式代入 (3.5.3) 式中, 再利用 (3.4.2) 式得到

$$\begin{aligned} \tilde{\theta}(kT) &= \tilde{\theta}(kT - T) + \frac{\Phi(p, kT)}{r(kT)} [\mathbf{Y}(p, kT) - \Phi^T(p, kT)\tilde{\theta}(kT)] \\ &= \tilde{\theta}(kT - T) + \frac{\Phi(p, kT)}{r(kT)} [-\Phi^T(p, kT)\tilde{\theta}(kT - T) + \mathbf{V}(p, kT)] \\ &=: \tilde{\theta}(kT - T) + \frac{\Phi(p, kT)}{r(kT)} [-\tilde{\mathbf{y}}(p, kT) + \mathbf{V}(p, kT)], \end{aligned} \quad (3.4.7)$$

其中

$$\tilde{\mathbf{y}}(p, kT) := \Phi^T(p, kT)\tilde{\theta}(kT - T). \quad (3.4.8)$$

根据 $Q(kT)$ 的定义, 由 (3.5.4) 式得到

$$\begin{aligned} Q(kT) &= \tilde{\theta}^T(kT - T)\tilde{\theta}(kT - T) + \frac{2}{r(kT)} \tilde{\mathbf{y}}^T(p, kT)[- \tilde{\mathbf{y}}(p, kT) \\ &\quad + \mathbf{V}(p, kT)] + \frac{1}{r^2(kT)} [\tilde{\mathbf{y}}^T(p, kT)\Phi^T(p, kT)\Phi(p, kT)\tilde{\mathbf{y}}(p, kT) \\ &\quad - 2\tilde{\mathbf{y}}^T(p, kT)\Phi^T(p, kT)\Phi(p, kT)\mathbf{V}(p, kT) + \|\Phi(p, kT)\mathbf{V}(p, kT)\|^2] \\ &\leq Q(kT - T) - \left[\frac{2}{r(kT)} - \frac{\|\Phi(p, kT)\|^2}{r^2(kT)} \right] \|\tilde{\mathbf{y}}(p, kT)\|^2 \end{aligned}$$

$$\begin{aligned}
 & + \frac{2}{r(kT)} \tilde{\mathbf{y}}^T(p, kT) \mathbf{V}(p, kT) - \frac{2}{r^2(kT)} \tilde{\mathbf{y}}^T(p, kT) \Phi^T(p, kT) \Phi(p, kT) \mathbf{V}(p, kT) \\
 & + \frac{1}{r^2(kT)} \sum_{i=0}^{p-1} v^2(kT - iT) \sum_{i=0}^{p-1} \|\varphi(kT - iT)\|^2 \\
 \leq & Q(kT - T) - \frac{1}{r(kT)} \|\tilde{\mathbf{y}}(p, kT)\|^2 + \frac{2}{r(kT)} \tilde{\mathbf{y}}^T(p, kT) \mathbf{V}(p, kT) \\
 & - \frac{2}{r^2(kT)} \tilde{\mathbf{y}}^T(p, kT) \Phi^T(p, kT) \Phi(p, kT) \mathbf{V}(p, kT) \\
 & + \frac{p}{r^2(kT)} \sum_{i=0}^{p-1} \|\varphi(kT - iT)\|^2 v^2(kT - iT). \tag{3.4.9}
 \end{aligned}$$

由于 $Q(kT - T)$, $\tilde{\mathbf{y}}(p, kT)$, $\Phi(p, kT)$ 和 $\varphi(kT)$ 与 $\mathbf{V}(p, kT)$ 独立, 并且 \mathcal{F}_{k-1} 是可测的, 对 (3.4.9) 式两边 \mathcal{F}_{k-1} 取条件期望, 并利用假设 (A1)-(A2) 可以得到

$$\mathbb{E}[Q(kT) | \mathcal{F}_{k-1}] \leq Q(kT - T) - \frac{1}{r(kT)} \|\tilde{\mathbf{y}}(p, kT)\|^2 + \frac{p}{r^2(kT)} \sum_{i=0}^{p-1} \|\varphi(kT - iT)\|^2 \sigma^2. \tag{3.4.10}$$

从 $k = 1$ 到 $k = \infty$ 累加 (3.4.10) 式右边最后一项, 应用引理 3.5.1 得到

$$\sum_{k=1}^{\infty} \frac{p}{r^2(kT)} \sum_{i=0}^{p-1} \|\varphi(kT - iT)\|^2 \sigma^2 = p\sigma^2 \sum_{k=1}^{\infty} \sum_{i=0}^{p-1} \frac{\|\varphi(kT - iT)\|^2}{r^2(kT)} < \infty.$$

对 (3.4.10) 式应用鞅收敛定理, 可得 $Q(kT)$ 一致趋于一个有限的随机变量 Q_0 , 即,

$$\mathbb{E}[\|\tilde{\boldsymbol{\theta}}(kT)\|^2] \rightarrow Q_0, \quad k \rightarrow \infty, \tag{3.4.11}$$

和

$$\lim_{k \rightarrow \infty} \sum_{i=0}^k \frac{1}{r(iT)} \|\tilde{\mathbf{y}}(p, iT)\|^2 < \infty, \quad \text{a.s.} \tag{3.4.12}$$

(3.4.11) 式表明 MR-MISG 参数估计误差是一致有界的.

定义

$$\Delta \tilde{\boldsymbol{\theta}}(kT) := \frac{\Phi(p, kT)}{r(kT)} [-\tilde{\mathbf{y}}(p, kT) + \mathbf{V}(p, kT)]. \tag{3.4.13}$$

将 (3.4.13) 式代入 (3.5.4) 式得到

$$\begin{aligned}
 \tilde{\boldsymbol{\theta}}(kT) &= \tilde{\boldsymbol{\theta}}(kT - T) + \Delta \tilde{\boldsymbol{\theta}}(kT), \\
 \tilde{\boldsymbol{\theta}}(kT + T) &= \tilde{\boldsymbol{\theta}}(kT - T) + \Delta \tilde{\boldsymbol{\theta}}(kT) + \Delta \tilde{\boldsymbol{\theta}}(kT + T), \\
 &\vdots
 \end{aligned}$$

$$\tilde{\theta}(kT + iT) = \tilde{\theta}(kT - T) + \sum_{s=0}^i \Delta \tilde{\theta}(kT + sT). \quad (3.4.14)$$

用 $\Phi(p, kT)$ 左乘 (3.4.8) 式两边并用 $k+i$ 代替 k 得到

$$\Phi(p, kT + iT)\Phi^T(p, kT + iT)\tilde{\theta}(kT + iT - T) = \Phi(p, kT + iT)\tilde{y}(p, kT + iT). \quad (3.4.15)$$

根据 (3.4.14) 式和 (3.4.15) 式得到

$$\begin{aligned} & \Phi(p, kT + iT)\Phi^T(p, kT + iT)\tilde{\theta}(kT - T) \\ &= \Phi(p, kT + iT)\tilde{y}(p, kT + iT) - \Phi(p, kT + iT)\Phi^T(p, kT + iT) \sum_{s=0}^{i-1} \Delta \tilde{\theta}(kT + sT). \end{aligned}$$

将上式两边从 $i=0$ 到 $i=N-1$ 累加得到

$$\begin{aligned} & \left[\sum_{i=0}^{N-1} \Phi(p, kT + iT)\Phi^T(p, kT + iT) \right] \tilde{\theta}(kT - T) \\ &= \sum_{i=0}^{N-1} \Phi(p, kT + iT)\tilde{y}(p, kT + iT) \\ & \quad - \sum_{i=0}^{N-1} \Phi(p, kT + iT)\Phi^T(p, kT + iT) \sum_{s=0}^{i-1} \Delta \tilde{\theta}(kT + sT). \end{aligned} \quad (3.4.16)$$

由于 (A4) 对所有 $k > 0$ 成立, 因此我们有

$$\begin{aligned} a\mathbf{I} &\leq \frac{1}{N} \sum_{i=0}^{N-1} \varphi(kT + iT)\varphi^T(kT + iT) \leq b\mathbf{I}, \text{ a.s.} \\ a\mathbf{I} &\leq \frac{1}{N} \sum_{i=0}^{N-1} \varphi(kT + iT - T)\varphi^T(kT + iT - T) \leq b\mathbf{I}, \text{ a.s.} \\ &\vdots \\ a\mathbf{I} &\leq \frac{1}{N} \sum_{i=0}^{N-1} \varphi(kT + iT - pT + T)\varphi^T(kT + iT - pT + T) \leq b\mathbf{I}, \text{ a.s.} \end{aligned}$$

累加上述不等式, 可得

$$pa\mathbf{I} \leq \frac{1}{N} \sum_{s=0}^{p-1} \sum_{i=0}^{N-1} \varphi(kT + iT - sT)\varphi^T(kT + iT - sT) \leq pb\mathbf{I}, \text{ a.s.},$$

即

$$pa\mathbf{I} \leq \frac{1}{N} \sum_{i=0}^{N-1} \Phi(p, kT + iT)\Phi^T(p, kT + iT) \leq pb\mathbf{I}, \text{ a.s.}$$

对上面不等式的两边取迹, 得到

$$pn_0aN \leq \sum_{i=0}^{N-1} \sum_{s=0}^{p-1} \|\varphi(kT+iT-sT)\|^2 \leq pn_0bN,$$

$$\|\Phi(p, kT+iT)\|^2 \leq pn_0bN,$$

其中

$$\sum_{i=0}^{N-1} \sum_{s=0}^{p-1} \|\varphi(kT+iT-sT)\|^2 = \sum_{i=0}^{N-1} \|\Phi(p, kT+iT)\|^2.$$

对 (3.4.16) 式取范数并利用上面的不等式, 得到

$$\begin{aligned} & \|\tilde{\theta}(kT-T)\| \\ & \leq \frac{1}{paNn_0} \left\| \sum_{i=0}^{N-1} \Phi(p, kT+iT) \tilde{y}(p, kT+iT) \right\| \\ & \quad + \frac{1}{paNn_0} \left\| \sum_{i=0}^{N-1} \Phi(p, kT+iT) \Phi^T(p, kT+iT) \sum_{s=0}^{i-1} \Delta \tilde{\theta}(kT+sT) \right\| \\ & \leq \frac{1}{paNn_0} \left\| \sum_{i=0}^{N-1} \Phi(p, kT+iT) \tilde{y}(p, kT+iT) \right\| + \frac{b}{a} \sum_{i=0}^{N-1} \sum_{s=0}^{i-1} \|\Delta \tilde{\theta}(kT+sT)\|. \end{aligned}$$

上面不等式平方后得到

$$\begin{aligned} \|\tilde{\theta}(kT-T)\|^2 & \leq \frac{2}{(paNn_0)^2} \left\| \sum_{i=0}^{N-1} \Phi(p, kT+iT) \tilde{y}(p, kT+iT) \right\|^2 \\ & \quad + 2 \left(\frac{b}{a} \right)^2 \left[\sum_{i=0}^{N-1} \sum_{s=0}^{i-1} \|\Delta \tilde{\theta}(kT+sT)\| \right]^2 \\ & \leq \frac{2N}{(paNn_0)^2} \sum_{i=0}^{N-1} \|\Phi(p, kT+iT)\|^2 \|\tilde{y}(p, kT+iT)\|^2 \\ & \quad + 2 \left(\frac{bN}{a} \right)^2 \sum_{i=0}^{N-1} \sum_{s=0}^{i-1} \|\Delta \tilde{\theta}(kT+sT)\|^2 \\ & \leq \frac{2N}{(paNn_0)^2} \sum_{i=0}^{N-1} Nn_0pb \|\tilde{y}(p, kT+iT)\|^2 \\ & \quad + 2 \left(\frac{bN}{a} \right)^2 \sum_{i=0}^{N-1} \sum_{s=0}^{i-1} \|\Delta \tilde{\theta}(kT+sT)\|^2 \\ & =: d_1 \sum_{i=0}^{N-1} \|\tilde{y}(p, kT+iT)\|^2 + d_2 \sum_{i=0}^{N-1} \sum_{s=0}^{i-1} \|\Delta \tilde{\theta}(kT+sT)\|^2, \end{aligned}$$

其中

$$d_1 := \frac{2b}{n_0 p a^2} < \infty, \quad d_2 := 2 \left(\frac{bN}{a} \right)^2 < \infty.$$

在上式中用 l 代替 k , 两边同时除以 $r(lT)$ 并从 $l = l_0$ ($l_0 < \infty$) 到 $l = k$ 累加, 有

$$\sum_{l=l_0}^k \frac{\|\tilde{\theta}(lT - T)\|^2}{r(lT)} \leq d_1 \sum_{i=0}^{N-1} \sum_{l=l_0}^k \frac{\|\tilde{y}(p, lT + iT)\|^2}{r(lT)} + d_2 \sum_{i=0}^{N-1} \sum_{s=0}^{i-1} \sum_{l=l_0}^k \frac{\|\Delta\tilde{\theta}(lT + sT)\|^2}{r(lT)}.$$

根据 $r(kT)$ 的定义 (3.4.3) 式, 有

$$\frac{\|\Phi(p, kT)\|^2}{r(kT)} = \sum_{i=0}^{p-1} \frac{\|\varphi(kT - iT)\|^2}{r(kT)} \leq 1. \quad (3.4.17)$$

利用 (3.4.13) 式和 (3.4.17) 式, 得到

$$\begin{aligned} & \sum_{l=l_0}^k \frac{\|\Delta\tilde{\theta}(lT + sT)\|^2}{r(lT)} \\ &= \sum_{l=l_0}^k \left\| \frac{\Phi(p, lT + sT)}{r(lT + sT)} [-\tilde{y}(p, lT + sT) + \mathbf{V}(p, lT + sT)] \right\|^2 \\ &\leq \sum_{l=l_0}^k \frac{2}{r(lT)} \frac{\|\Phi(p, lT + sT)\|^2}{r^2(lT + sT)} (\|\tilde{y}(p, lT + sT)\|^2 + \|\mathbf{V}(p, lT + sT)\|^2) \\ &\leq \sum_{l=l_0}^k \frac{2}{r(lT)r(lT + sT)} (\|\tilde{y}(p, lT + sT)\|^2 + \sum_{j=0}^{p-1} \|v(lT + sT - jT)\|^2) \\ &\leq \sum_{l=l_0}^k \frac{2\|\tilde{y}(p, lT + sT)\|^2}{r(lT + sT)} + \sum_{l=l_0}^k \sum_{j=0}^{p-1} \frac{2\|v(lT + sT - jT)\|^2}{r(lT + sT)}. \end{aligned}$$

利用 (3.4.12) 式和 (A3), 得到

$$\sum_{l=l_0}^k \frac{\|\Delta\tilde{\theta}(lT + sT)\|^2}{r(lT)} < \infty, \quad \text{a.s.} \quad (3.4.18)$$

联合 (3.4.12) 式和 (3.4.18) 式, 得到

$$\sum_{l=l_0}^k \frac{\|\tilde{\theta}(lT - T)\|^2}{r(lT)} < \infty, \quad \text{a.s.}$$

应用 Kronecker 引理得到

$$\frac{1}{r(kT)} \sum_{l=l_0}^k \|\tilde{\theta}(lT - T)\|^2 = 0, \quad \text{a.s.}, \quad k \rightarrow \infty.$$

由于 $Q(kT) = \|\hat{\theta}(kT)\|^2$ 一致收敛于一个有限的随机变量 Q_0 , 应用引理 2.3.3 可得

$$\lim_{k \rightarrow \infty} \|\hat{\theta}(kT)\|^2 = 0, \text{ a.s.}$$

这表明参数估计 $\hat{\theta}(kT)$ 一致收敛于参数真值 θ , 当 k 趋于无穷时.

定理 3.5.1 证明完毕. □

3.4.3 仿真试验

例 3.4.1 对于双输入单输出多率系统, 其刷新周期分别为 $T_1 = 2h$ 和 $T_2 = 3h$, 输出采样周期 $T = 6h$, 取基周期 $h = 2$, 相应的传递函数模型为

$$\begin{aligned} \alpha(z)y(kT) = & \beta_{11}(z)u_1(kT) + \beta_{12}(z)u_1(kT + T_1) + \beta_{13}(z)u_1(kT + 2T_1) \\ & + \beta_{21}(z)u_2(kT) + \beta_{22}(z)u_2(kT + T_2) + v(kT), \end{aligned}$$

其中

$$\begin{aligned} \alpha(z) &= 1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} = 1 + 0.60z^{-1} + 0.40z^{-2}, \\ \beta_{11}(z) &= \beta_{11}(1)z^{-1} + \beta_{11}(2)z^{-2} = -0.20z^{-1} + 0.50z^{-2}, \\ \beta_{12}(z) &= \beta_{12}(1)z^{-1} + \beta_{12}(2)z^{-2} = -0.65z^{-1} + 0.60z^{-2}, \\ \beta_{13}(z) &= \beta_{13}(1)z^{-1} + \beta_{13}(2)z^{-2} = 1.00z^{-1} - 0.50z^{-2}, \\ \beta_{21}(z) &= \beta_{21}(1)z^{-1} + \beta_{21}(2)z^{-2} = -0.40z^{-1} + 0.80z^{-2}, \\ \beta_{22}(z) &= \beta_{22}(1)z^{-1} + \beta_{22}(2)z^{-2} = 0.70z^{-1} + 0.20z^{-2}, \end{aligned}$$

参数向量和信息向量为

$$\begin{aligned} \theta &= [\alpha_1, \alpha_2, \beta_{11}(1), \beta_{11}(2), \beta_{12}(1), \beta_{12}(2), \beta_{13}(1), \beta_{13}(2), \\ & \quad \beta_{21}(1), \beta_{21}(2), \beta_{22}(1), \beta_{22}(2)]^T \\ &= [0.60, 0.40, -0.20, 0.50, -0.65, 0.60, 1.00, -0.50, \\ & \quad -0.40, 0.80, 0.70, 0.20]^T, \\ \varphi(kT) &= [-y(kT - T), -y(kT - 2T), \phi_1^T(kT), \phi_2^T(kT)]^T, \\ \phi_1^T(kT) &= [u_1(kT - T), u_1(kT - 2T), u_1(kT + T_1 - T), u_1(kT + T_1 - 2T), \\ & \quad u_1(kT + 2T_1 - T), u_1(kT + 2T_1 - 2T)], \\ \phi_2^T(kT) &= [u_2(kT - T), u_2(kT - 2T), u_2(kT + T_2 - T), u_2(kT + T_2 - 2T)]. \end{aligned}$$

仿真时, $\{u_1(kT_1)\}$ 和 $\{u_2(kT_2)\}$ 采用零均值单位方差不相关可测的持续激励信号序列. $\{v(kT)\}$ 采用零均值方差为 $\sigma^2 = 0.10^2$. 应用 MR-SG 和 MR-MISG 算法估计这个系统的参数, 其参数估计和参数估计误差如表 3.4-3.7 所示, 不同新息长度下的参数估计误差 $\delta := \|\hat{\theta}(t) - \theta\|/\|\theta\|$ 随 k 变化曲线如图 3.6 所示. 系统的噪信比为 $\delta_{\text{ns}} = 28.71\%$.

表 3.4 MR-SG 参数估计和参数估计误差
Table 3.4 The MR-SG parameter estimates and their errors

k	1000	2000	3000	4000	5000	6000	7000
$\alpha_1 = 0.60000$	0.53517	0.55146	0.55871	0.56594	0.57329	0.57832	0.58174
$\alpha_2 = 0.40000$	0.29300	0.31106	0.32205	0.32752	0.32807	0.32888	0.33114
$\beta_{11}(1) = -0.20000$	-0.04823	-0.05270	-0.05468	-0.05578	-0.05758	-0.05875	-0.05944
$\beta_{11}(2) = 0.50000$	0.27403	0.27879	0.28211	0.28538	0.28784	0.28956	0.29049
$\beta_{12}(1) = -0.65000$	-0.07521	-0.09030	-0.09999	-0.10704	-0.11229	-0.11664	-0.12011
$\beta_{12}(2) = 0.60000$	0.20006	0.21268	0.22060	0.22581	0.23032	0.23311	0.23590
$\beta_{13}(1) = 1.00000$	0.32575	0.34391	0.35528	0.36276	0.36860	0.37345	0.37756
$\beta_{13}(2) = -0.50000$	-0.25273	-0.26029	-0.26434	-0.26723	-0.26925	-0.27112	-0.27294
$\beta_{21}(1) = -0.40000$	-0.01315	-0.02442	-0.03100	-0.03563	-0.03945	-0.04224	-0.04461
$\beta_{21}(2) = 0.80000$	0.20257	0.21950	0.23023	0.23756	0.24306	0.24784	0.25132
$\beta_{22}(1) = 0.70000$	0.12113	0.13572	0.14606	0.15248	0.15690	0.16122	0.16481
$\beta_{22}(2) = 0.20000$	0.02017	0.02246	0.02495	0.02748	0.02871	0.03029	0.03108
δ (%)	68.70980	66.76752	65.54056	64.70667	64.08048	63.56211	63.13930

表 3.5 MR-MISG 参数估计和参数估计误差 ($p = 2$)
Table 3.5 The MR-MISG parameter estimates and their errors ($p = 2$)

k	1000	2000	3000	4000	5000	6000	7000
$\alpha_1 = 0.60000$	0.64417	0.64935	0.64971	0.65400	0.65964	0.66293	0.66433
$\alpha_2 = 0.40000$	0.40100	0.41688	0.42465	0.42606	0.42349	0.42174	0.42289
$\beta_{11}(1) = -0.20000$	-0.13622	-0.14021	-0.14153	-0.14176	-0.14362	-0.14452	-0.14486
$\beta_{11}(2) = 0.50000$	0.44203	0.44197	0.44319	0.44526	0.44658	0.44742	0.44723
$\beta_{12}(1) = -0.65000$	-0.07764	-0.10843	-0.12698	-0.14011	-0.14982	-0.15775	-0.16398
$\beta_{12}(2) = 0.60000$	0.35005	0.36207	0.36981	0.37527	0.37981	0.38248	0.38521
$\beta_{13}(1) = 1.00000$	0.54046	0.56485	0.58015	0.58966	0.59720	0.60350	0.60885
$\beta_{13}(2) = -0.50000$	-0.32313	-0.32882	-0.33179	-0.33386	-0.33554	-0.33700	-0.33863
$\beta_{21}(1) = -0.40000$	-0.10320	-0.12004	-0.13019	-0.13670	-0.14222	-0.14603	-0.14952
$\beta_{21}(2) = 0.80000$	0.36150	0.38421	0.39795	0.40734	0.41417	0.42042	0.42478
$\beta_{22}(1) = 0.70000$	0.24851	0.27103	0.28646	0.29607	0.30272	0.30902	0.31417
$\beta_{22}(2) = 0.20000$	0.12849	0.13165	0.13518	0.13871	0.13999	0.14207	0.14317
δ (%)	52.06904	49.39824	47.72820	46.62058	45.78970	45.09440	44.53354

表 3.6 MR-MISG 参数估计和参数估计误差 ($p = 5$)
Table 3.6 The MR-MISG parameter estimates and their errors ($p = 5$)

k	1000	2000	3000	4000	5000	6000	7000
$\alpha_1 = 0.60000$	0.67586	0.66717	0.66009	0.66099	0.66319	0.66318	0.66215
$\alpha_2 = 0.40000$	0.44668	0.44888	0.44807	0.44357	0.43885	0.43603	0.43582
$\beta_{11}(1) = -0.20000$	-0.10876	-0.12076	-0.12610	-0.12891	-0.13335	-0.13576	-0.13760
$\beta_{11}(2) = 0.50000$	0.45346	0.45512	0.45795	0.46061	0.46257	0.46374	0.46382
$\beta_{12}(1) = -0.65000$	-0.32599	-0.36863	-0.39193	-0.40781	-0.41943	-0.42854	-0.43520
$\beta_{12}(2) = 0.60000$	0.46641	0.47849	0.48509	0.49031	0.49482	0.49736	0.49995
$\beta_{13}(1) = 1.00000$	0.87866	0.89343	0.90340	0.90875	0.91269	0.91631	0.91920
$\beta_{13}(2) = -0.50000$	-0.38574	-0.39004	-0.39277	-0.39457	-0.39651	-0.39809	-0.39963
$\beta_{21}(1) = -0.40000$	-0.21620	-0.24032	-0.25470	-0.26373	-0.27092	-0.27562	-0.27982
$\beta_{21}(2) = 0.80000$	0.54452	0.57435	0.59040	0.60141	0.60949	0.61649	0.62134
$\beta_{22}(1) = 0.70000$	0.57521	0.58941	0.59913	0.60499	0.60835	0.61185	0.61463
$\beta_{22}(2) = 0.20000$	0.16601	0.17489	0.18022	0.18504	0.18721	0.18991	0.19139
δ (%)	26.12643	23.05862	21.30515	20.16205	19.30051	18.62362	18.10855

表 3.7 MR-MISG 参数估计和参数估计误差 ($p = 12$)
Table 3.7 The MR-MISG parameter estimates and their errors ($p = 12$)

k	1000	2000	3000	4000	5000	6000	7000
$\alpha_1 = 0.60000$	0.60608	0.60581	0.60438	0.60341	0.60322	0.60346	0.60371
$\alpha_2 = 0.40000$	0.40510	0.40566	0.40443	0.40309	0.40347	0.40288	0.40284
$\beta_{11}(1) = -0.20000$	-0.17118	-0.17911	-0.18277	-0.18532	-0.18670	-0.18791	-0.18878
$\beta_{11}(2) = 0.50000$	0.48566	0.48975	0.49071	0.49254	0.49346	0.49405	0.49422
$\beta_{12}(1) = -0.65000$	-0.61144	-0.62226	-0.62733	-0.63105	-0.63327	-0.63510	-0.63596
$\beta_{12}(2) = 0.60000$	0.57165	0.57907	0.58220	0.58388	0.58576	0.58697	0.58789
$\beta_{13}(1) = 1.00000$	0.99474	0.99548	0.99718	0.99770	0.99804	0.99829	0.99828
$\beta_{13}(2) = -0.50000$	-0.49636	-0.49494	-0.49468	-0.49422	-0.49450	-0.49485	-0.49506
$\beta_{21}(1) = -0.40000$	-0.38480	-0.38846	-0.39132	-0.39265	-0.39313	-0.39380	-0.39448
$\beta_{21}(2) = 0.80000$	0.79094	0.79487	0.79536	0.79545	0.79528	0.79592	0.79612
$\beta_{22}(1) = 0.70000$	0.68492	0.68885	0.69110	0.69303	0.69386	0.69452	0.69475
$\beta_{22}(2) = 0.20000$	0.20282	0.20272	0.20256	0.20317	0.20380	0.20388	0.20362
δ (%)	3.08808	2.27116	1.87802	1.61116	1.45258	1.31651	1.23680

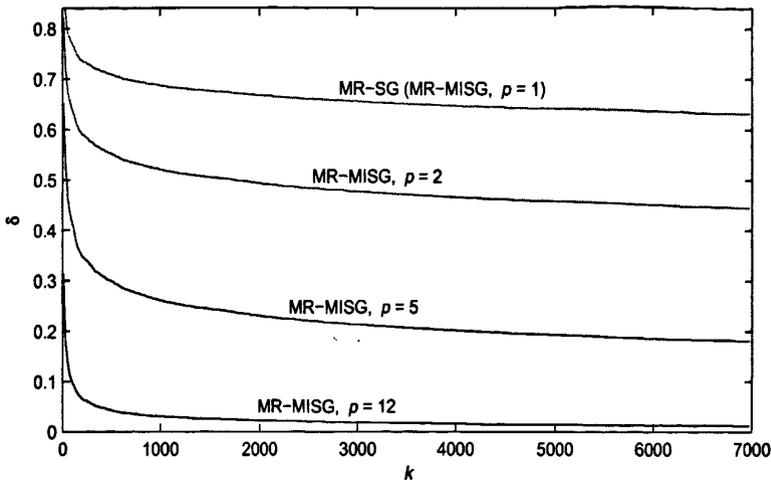


图 3.6 不同新息长度 p 下参数估计误差 δ 随数据长度 k 变化曲线
 Fig. 3.6 The parameter estimation errors δ versus k with different p

从表 3.4-3.7 与图 3.6, 可得到下面的结论:

- MR-SG 和 MR-MISG 算法的参数估计误差随着 k 的逐渐增加变得越来越小, 并趋于零. 验证了 MR-SG 和 MR-MISG 算法的有效性.
- 由 MR-MISG ($p \geq 2$) 算法得到的参数估计的精度要高于 MR-SG 算法, 且收敛速度要快于 MR-SG 算法.

本节的内容发表在国际期刊《Computers & Mathematics with Applications》2009 年 57 卷第 9 期上. 参加了于 2008 年 5 月由南京理工大学主办的 2008 年江苏省博士论坛.

3.5 递阶最小二乘法

对于 r 输入的多率系统如图 3.1 所示, 经过离散化后得到时不变离散状态空间模型, 推导出辨识模型 (3.3.3), 需要辨识的参数个数为 $n + n(\nu_1 + \nu_2 + \cdots + \nu_r)$. 由此可知, 如图 3.1 所示的多率系统, 待辨识参数数目较多, 常规的辨识算法 (如最小二乘法) 的计算量很大. 在辨识领域中, 递阶辨识原理可用于解决维数大、变量数多的系统辨识的计算量大问题. 本节应用递阶辨识原理, 提出递阶最小二乘法, 得到如图 3.1 所示的多率系统的参数估计.

3.5.1 算法的提出

对于多率多输入系统如图 3.1 所示, 其辨识模型 (3.3.3), 极小化准则函数

$$J(\boldsymbol{\theta}) = \sum_{i=1}^k [y(kT) - \boldsymbol{\varphi}^T(kT)\boldsymbol{\theta}]^2,$$

可以得到递推最小二乘 (MR-RLS) 算法, 如下:

$$\hat{\boldsymbol{\theta}}(kT) = \hat{\boldsymbol{\theta}}(kT - T) + \mathbf{P}(kT)\boldsymbol{\varphi}(kT)[y(kT) - \boldsymbol{\varphi}^T(kT)\hat{\boldsymbol{\theta}}(kT - T)], \quad (3.5.1)$$

$$\mathbf{P}(kT) = \mathbf{P}(kT - T) - \frac{\mathbf{P}(kT - T)\boldsymbol{\varphi}(kT)\boldsymbol{\varphi}^T(kT)\mathbf{P}(kT - T)}{1 + \boldsymbol{\varphi}^T(kT)\mathbf{P}(kT - T)\boldsymbol{\varphi}(kT)}, \quad (3.5.2)$$

其中 $\mathbf{P}(kT)$ 是协方差矩阵. 取初始值 $\hat{\boldsymbol{\theta}}(0) = \mathbf{1}_n/p_0$ 和 $\mathbf{P}(0) = p_0\mathbf{I}_n$, 一般地, $p_0 > 0$ 取很大的数, e.g., $p_0 = 10^6$.

事实上, 离散化后的多率系统是高维的, 用 MR-RLS 算法 (3.5.1)–(3.5.2) 计算参数估计时, 每一次的迭代计算中都要计算协方差矩阵的逆矩阵, 这使得算法的计算量很大.

递阶辨识原理源于大系统递阶控制的分解协调原理 (Decomposition-coordination principle). 其步骤分为三步: 第一步, 辨识模型分解为一些子辨识模型; 第二步, 采用最小二乘原理分别辨识一些子模型; 第三步, 采用协调处理各子辨识算法间的关联项. 下面应用递阶辨识原理, 以解决 MR-RLS 算法 (3.5.1)–(3.5.2) 计算量大的问题.

首先将信息向量 $\boldsymbol{\varphi}(kT)$ 分解为 N 子信息向量, 将参数向量 $\boldsymbol{\theta}$ 分解为 N 个子参数向量, 即,

$$\boldsymbol{\varphi} = \begin{bmatrix} \boldsymbol{\varphi}_1(kT) \\ \boldsymbol{\varphi}_2(kT) \\ \vdots \\ \boldsymbol{\varphi}_N(kT) \end{bmatrix} \in \mathbb{R}^{n_0}, \quad \boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\theta}_1(kT) \\ \boldsymbol{\theta}_2(kT) \\ \vdots \\ \boldsymbol{\theta}_N(kT) \end{bmatrix} \in \mathbb{R}^{n_0},$$

其中 $\boldsymbol{\varphi}_i(kT) \in \mathbb{R}^{n_i}$ ($i = 1, 2, \dots, N$), $\boldsymbol{\theta}_i \in \mathbb{R}^{n_i}$ 且 $n_1 + n_2 + \cdots + n_N = n_0$. 从而系统 (3.3.3) 可以分解为 N 个 (虚拟的) 子系统

$$y_i(kT) = \boldsymbol{\varphi}_i^T(kT)\boldsymbol{\theta}_i + v_i(kT), \quad (3.5.3)$$

$$y_i(kT) = y(kT) - \sum_{j=1, j \neq i}^N \varphi_j(kT) \theta_j, \quad i = 1, 2, \dots, N, \quad (3.5.4)$$

这里假定 $v_i(kT)$ 是零均值的白噪声。

其次用最小二乘算法辨识每一个子系统。令 $\hat{\theta}_i(kT)$ 是第 i 个参数向量 θ_i 的估计，对于子系统 (3.5.3)，极小化

$$J(\theta_i) = \sum_{i=1}^k [y_i(kT) - \varphi_i^T(kT) \theta_i]^2, \quad i = 1, 2, \dots, N,$$

得到子系统 (3.5.3) 的最小二乘参数估计 $\hat{\theta}_i(kT)$,

$$\hat{\theta}_i(kT) = \hat{\theta}_i(kT - T) + P_i(kT) \varphi_i(kT) [y_i(kT) - \varphi_i^T(kT) \hat{\theta}_i(kT - T)], \quad (3.5.5)$$

$$P_i(kT) = P_i(kT - T) - \frac{P_i(kT - T) \varphi_i(kT) \varphi_i^T(kT) P_i(kT - T)}{1 + \varphi_i^T(kT) P_i(kT - T) \varphi_i(kT)}, \quad (3.5.6)$$

其中 $P_i(kT)$ 是第 i 个子系统的协方差矩阵。然而，算法 (3.5.5)–(3.5.6) 却不能实现。其原因是：将 (3.5.3) 式代入 (3.5.5) 式得到

$$\hat{\theta}_i(kT) = \hat{\theta}_i(kT - T) + P_i(kT) \varphi_i(kT) [y(kT) - \sum_{j \neq i}^N \varphi_j(kT) \theta_j - \varphi_i^T(kT) \hat{\theta}_i(kT - T)],$$

上式中存在未知的参数向量 θ_j 。

最后，协调子辨识算法间的关联项。在 (3.5.5) 式中，将未知的参数向量 θ_j 用它们前一时刻 $kT - T$ 的估计值 $\hat{\theta}_j(kT - T)$ 来代替，则有

$$\begin{aligned} \hat{\theta}_i(kT) &= \hat{\theta}_i(kT - T) \\ &+ P_i(kT) \varphi_i(kT) [y(kT) - \sum_{j \neq i}^N \varphi_j(kT) \theta_j(kT - T) - \varphi_i^T(kT) \hat{\theta}_i(kT - T)] \\ &= \hat{\theta}_i(kT - T) + P_i(kT) \varphi_i(kT) [y(kT) - \varphi^T(kT) \hat{\theta}(kT - T)], \quad i = 1, 2, \dots, N. \end{aligned}$$

至此，可以得到估计多率系统 (3.3.3) 的参数 θ 的递阶最小二乘辨识 (MR-HLS) 算法，如下：

$$\hat{\theta}_i(kT) = \hat{\theta}_i(kT - T) + P_i(kT) \varphi_i(kT) [y(kT) - \varphi^T(kT) \hat{\theta}(kT - T)], \quad (3.5.7)$$

$$P_i(kT) = P_i(kT - T) - \frac{P_i(kT - T) \varphi_i(kT) \varphi_i^T(kT) P_i(kT - T)}{1 + \varphi_i^T(kT) P_i(kT - T) \varphi_i(kT)}. \quad (3.5.8)$$

取初始值 $\hat{\theta}_i(0) = \mathbf{1}_{n_i}/p_0$ 和 $P(0) = p_0 \mathbf{I}_{n_i}$ (e.g., $p_0 = 10^6$)。

MR-HLS 算法的特性：

表 3.8 计算量的比较
Table 3.8 The comparison of computational load

算法	乘法次数	加法次数
MR-RLS	$2n_0^2 + 4n_0 = 336$	$2n_0^2 + 2n_0 = 312$
MR-HLS	$\sum_{i=1}^N 2n_i^2 + 4n_i = 144$	$\sum_{i=1}^N 2n_i^2 + 2n_i = 120$

- 与 RLS 算法比较, MR-HLS 算法有较小的计算量. 例如, 对于一个双输入单输出的多率系统, 如果 $p_1 = 2, p_2 = 3, q = 6$ 和 $n = 2$, 则 $n_0 = n + n(\nu_1 + \nu_2) = 12$. 将辨识模型分解为 3 个子系统 ($N=3$), 每个子系统的维数 $n_i = 4, i = 1, 2, 3$. MR-RLS 算法和 MR-HLS 算法计算量的比较如表 3.8 所示.
- 当 $n_1 = n_2 = \dots = n_N = 1$, MR-HLS 算法 (3.5.7)–(3.5.8) 变为随机梯度算法. 此时, MR-HLS 算法具有最小的计算量.
- 子系统的个数 N 取决于计算的复杂性和参数收敛的精确性. 因此, N 的取值不会很大也不会很小. 这也是一个需要进一步研究的问题.

3.5.2 仿真试验

例 3.5.1 考虑双输入单输出多率系统, 刷新周期分别为 $T_1 = 2h$ 和 $T_2 = 3h$, 输出采样周期 $T = 6h$, 取基周期 $h = 2$, 相对应的传递函数模型为

$$\alpha(z)y(kT) = \beta_{11}(z)u_1(kT) + \beta_{12}(z)u_1(kT + T_1) + \beta_{13}(z)u_1(kT + 2T_1) \\ + \beta_{21}(z)u_2(kT) + \beta_{22}(z)u_2(kT + T_2) + v(kT),$$

其中

$$\alpha(z) = 1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} = 1 + 0.58z^{-1} + 0.52z^{-2}, \\ \beta_{11}(z) = \beta_{11}(1)z^{-1} + \beta_{11}(2)z^{-2} = -0.20z^{-1} + 0.50z^{-2}, \\ \beta_{12}(z) = \beta_{12}(1)z^{-1} + \beta_{12}(2)z^{-2} = -0.64z^{-1} + 3.30z^{-2}, \\ \beta_{13}(z) = \beta_{13}(1)z^{-1} + \beta_{13}(2)z^{-2} = 1.20z^{-1} - 0.50z^{-2}, \\ \beta_{21}(z) = \beta_{21}(1)z^{-1} + \beta_{21}(2)z^{-2} = -0.38z^{-1} + 0.76z^{-2}, \\ \beta_{22}(z) = \beta_{22}(1)z^{-1} + \beta_{22}(2)z^{-2} = 0.68z^{-1} + 0.20z^{-2},$$

参数向量和信息向量为

$$\theta = [\alpha_1, \alpha_2, \beta_{11}(1), \beta_{11}(2), \beta_{12}(1), \beta_{12}(2), \beta_{13}(1), \beta_{13}(2), \beta_{21}(1), \beta_{21}(2), \\ \beta_{22}(1), \beta_{22}(2)]^T \\ = [0.58, 0.52, -0.20, 0.50, -0.64, 3.30, 1.20, -0.50, -0.38, 0.76,$$

$$\begin{aligned}
 & 0.68, 0.20]^T, \\
 \varphi(kT) &= [-y(kT - T), -y(kT - 2T), \phi_1^T(kT), \phi_2^T(kT)]^T, \\
 \phi_1(kT) &= [u_1(kT - T), u_1(kT - 2T), u_1(kT + T_1 - T), u_1(kT + T_1 - 2T), \\
 & \quad u_1(kT + 2T_1 - T), u_1(kT + 2T_1 - 2T)]^T, \\
 \phi_2(kT) &= [u_2(kT - T), u_2(kT - 2T), u_2(kT + T_2 - T), u_2(kT + T_2 - 2T)]^T.
 \end{aligned}$$

仿真时, $\{u_1(kT)\}$ 和 $\{u_2(kT)\}$ 采用零均值单位方差不相关可测的持续激励信号序列, $\{v(kT)\}$ 采用零均值方差为 $\sigma^2 = 1.00^2$ 的白噪声序列. 用 MR-HLS 算法估计这个系统的参数, 将辨识模型分为两个子系统. 其参数估计及参数估计误差如表 3.9 所示, 其中参数估计误差 $\delta := \|\hat{\theta}(t) - \theta\|/\|\theta\|$ 随 k 变化曲线如图 3.7 所示. 系统的噪信比分别为 $\delta_{ns} = 7.27\%$.

表 3.9 MR-HLS 算法参数估计及其误差
Table 3.9 The MR-HLS parameter estimates and their errors

k	100	200	500	1000	2000	3000	4000	5000
$\alpha_1 = 0.58000$	0.65281	0.63632	0.59369	0.58561	0.58329	0.58092	0.58110	0.58242
$\alpha_2 = 0.52000$	0.68780	0.60918	0.54331	0.54279	0.53837	0.53325	0.53024	0.53033
$\beta_{11}(1) = -0.20000$	-0.29475	-0.26845	-0.25340	-0.19235	-0.18936	-0.19137	-0.20286	-0.19877
$\beta_{11}(2) = 0.50000$	0.10707	0.26763	0.39075	0.42658	0.48323	0.46764	0.49030	0.49407
$\beta_{12}(1) = -0.64000$	-0.09494	-0.39826	-0.58471	-0.59684	-0.64587	-0.65625	-0.63376	-0.63238
$\beta_{12}(2) = 3.30000$	3.34616	3.28354	3.32370	3.33162	3.33850	3.30978	3.32547	3.31727
$\beta_{13}(1) = 1.20000$	1.24717	1.11060	1.21325	1.20561	1.14602	1.17802	1.19252	1.18997
$\beta_{13}(2) = -0.50000$	-0.48013	-0.45870	-0.40755	-0.42643	-0.44327	-0.46157	-0.46950	-0.47814
$\beta_{21}(1) = -0.38000$	-0.70652	-0.57098	-0.45746	-0.41360	-0.40726	-0.39901	-0.38850	-0.37508
$\beta_{21}(2) = 0.76000$	0.59904	0.73893	0.64630	0.74559	0.73060	0.73907	0.73889	0.73702
$\beta_{22}(1) = 0.68000$	0.87602	0.88066	0.78880	0.74807	0.71160	0.70060	0.69126	0.68047
$\beta_{22}(2) = 0.20000$	0.24566	0.16880	0.16077	0.21395	0.19407	0.18119	0.19077	0.19832
δ (%)	21.08370	11.96617	6.30986	3.68074	2.70206	1.85424	1.31470	1.04156

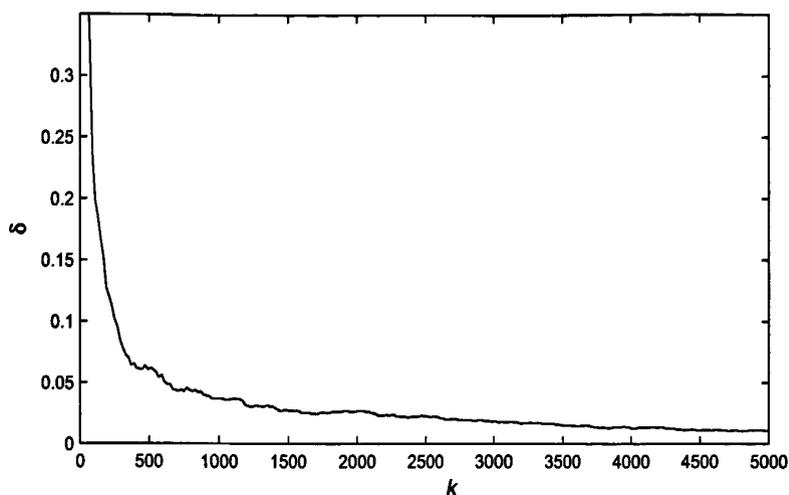


图 3.7 参数估计误差 δ 随 k 变化曲线
Fig. 3.7 The parameter estimation errors δ versus k

从表 3.9 和图 3.7 可以看出, 随着数据长度的增加, 算法的参数估计误差越来越小, 并趋于零. 验证了 MR-HLS 算法的有效性.

3.6 本章小结

本章针对方程误差多率系统, 用离散化技术推导其相对应的时不变离散状态空间模型, 得到其辨识模型. 为了提高随机梯度算法的收敛速度和参数估计精度, 引入了遗忘因子, 得到遗忘梯度算法估计系统的参数. 然而, 遗忘因子的选取是遗忘梯度算法的关键且不容易选取, 选取的不合适, 会导致算法的收敛波动性大. 本章基于多新息辨识理论, 提出多新息随机梯度算法估计系统的参数, 应用鞅收敛定理, 从理论上证明了多新息随机梯度算法的收敛性. 由于多新息随机梯度算法充分利用当前和过去时刻的信息, 因此可以提高随机梯度算法的收敛速度和参数估计精度. 利用递阶辨识原理, 提出计算量小的递阶最小二乘算法估计系统的参数.

第四章 输出误差多率系统

4.1 引言

辅助模型辨识方法^[94]主要解决系统中存在状态或输出不可测量的辨识问题. 最初, 研究多变量系统传递函数阵的子模型的辨识问题^[95,96], 经过发展和推广, 成为研究输出误差类模型辨识的主要研究方法. 基于辅助模型辨识思想的辨识理论和方法已经形成一大类辨识方法族, 如辅助模型最小二乘辨识方法, 辅助模型随机梯度辨识方法等, 应用于多种系统辨识问题中, 如双率采样系统^[22,23], 损失数据系统^[67], 非线性系统^[97]等.

辅助模型辨识方法的基本思想: 借助于一个辅助模型, 用系统的可测信息建立一个辅助模型, 用辅助模型的输出代替系统的不可测变量(如系统的真实输出或系统的状态). 通过选择辅助模型的参数, 使辅助模型的输出逼近这些不可测变量, 从而获得系统参数的一致估计.

本文将辅助模型辨识方法推广至输出误差多率系统的辨识. 组织如下: 第 4.2 节利用离散化技术, 得到多率系统离散时间状态空间模型. 第 4.3 节提出辅助模型多新息随机梯度算法估计系统参数, 用数值仿真例子验证算法有效性. 第 4.4 节借助辅助模型辨识思想, 提出辅助模型最小二乘算法估计系统参数, 应用鞅收敛定理分析算法的收敛性, 用数值仿真例子验证算法的有效性. 第 4.5 节为本章总结.

4.2 模型推导

对于一类多率多输入系统图 4.1 所示, 其中连续过程 P_c 有如下的状态空间模型:

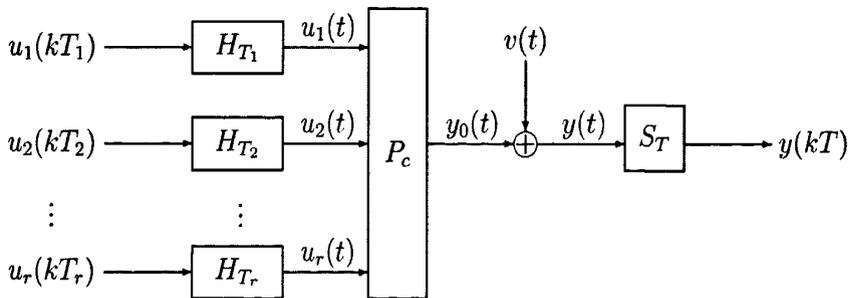


图 4.1 带噪声的多率多输入系统

Fig. 4.1 The multirate multiple-input system with noises

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}_c \mathbf{x}(t) + \sum_{j=1}^r \mathbf{B}_{c_j} u_j(t), \\ y(t) = \mathbf{C} \mathbf{x}(t) + v(t), \end{cases} \quad (4.2.1)$$

其中 $\mathbf{x}(t) \in \mathbb{R}^n$ 为状态向量, $u_j(t)$ ($j = 1, 2, \dots, r$) 为第 j 个输入通道的控制输入; $y(t) \in \mathbb{R}^1$ 为系统的输出, $v(t) \in \mathbb{R}^1$ 是一个零均值的随机噪声, \mathbf{A}_c , \mathbf{B}_{c_j} 和 \mathbf{C} 为有适当维数的矩阵.

S_T 是周期为 T 的采样器, H_{T_j} 是周期为 T_j 的零阶保持器. 离散时间输入信号 $u_j(kT_j)$ 经过零阶保持器 H_{T_j} 得到连续信号 $u_j(t)$, 经过连续时间过程 P_c 产生系统的无噪输出 $y_0(t)$ (不可测), $y(t)$ 为系统受到外部干扰 $v(t)$ 后的输出, 经过输出采样器 S_T 得到离散信号 $y(kT)$. 这里 $T_j := p_j h$ 为第 j 个输入刷新周期. 为了研究方便, 假定 p_1, p_2, \dots, p_r 是互质的正整数, q 是 p_1, p_2, \dots, p_r 的最小公倍数. 其中, h 被称为基周期, $T := qh$ 为框架周期^[31].

在第 k 个框架周期上零阶保持器 H_{T_j} ($j = 1, 2, \dots, r$) 对控制输入 u_j 均匀的刷新, 则零阶保持器 H_{T_j} 的特性可表示为

$$u_j(t) = u_j(kT_j), \quad kT_j \leq t < (k+1)T_j, \quad j = 1, 2, \dots, r. \quad (4.2.2)$$

输出采样器 S_T 的采样周期等于输入的框架周期. 这样的刷新和采样方案如第三章图 3.2 所示. 对于多率系统图 4.1 的可测的输入输出数据为 $\{u_1(kT_1), u_2(kT_2), \dots, u_r(kT_r), y(kT) : k = 1, 2, \dots\}$, 即, 多率输入输出数据.

下面定理采用离散化的方法, 建立多率多输入系统如图 4.1 所示多率输入输出数据的映射关系, 即推导其相对应的时不变离散时间状态空间模型.

定理 2.4.1 对于多率多输入系统如图 4.1 所示, $\nu_j := q/p_j, j = 1, 2, \dots, r$, 则可以得到多率系统的状态空间表达式为

$$\begin{cases} \mathbf{x}(kT + T) = \mathbf{A}\mathbf{x}(kT) + \sum_{j=1}^r \sum_{i=1}^{\nu_j} \mathbf{B}_{ji} u_j(kT + (i-1)T_j), \\ y(kT) = \mathbf{C}\mathbf{x}(kT) + v(kT), \end{cases} \quad (4.2.3)$$

其中 $\mathbf{x}(kT) := \mathbf{x}(t)|_{t=kT}$, $u_j(kT + iT_j) := u_j(t)|_{t=kT+iT_j}$, $y(kT) := y(t)|_{t=kT}$, $v(kT) := v(t)|_{t=kT}$, 和

$$\begin{aligned} \mathbf{A}_h &:= e^{\mathbf{A}_c h} \in \mathbb{R}^{n \times n}, \quad \mathbf{A} := e^{\mathbf{A}_c qh}, \\ \mathbf{B}_{p_j} &:= \int_0^{T_j} e^{\mathbf{A}_c t} dt \mathbf{B}_{c_j} \in \mathbb{R}^n, \\ \mathbf{B}_{ji} &:= \mathbf{A}_h^{q-ip_j} \mathbf{B}_{p_j} \in \mathbb{R}^n, \quad j = 1, 2, \dots, r, \quad i = 1, 2, \dots, \nu_j. \end{aligned}$$

证明: 根据线性系统知识, (4.2.1) 式的状态解为

$$\mathbf{x}(kT + T) = e^{\mathbf{A}_c T} \mathbf{x}(kT) + \sum_{j=1}^r \int_{kT}^{kT+T} e^{\mathbf{A}_c (kT+T-\tau)} \mathbf{B}_{c_j} u_j(\tau) d\tau.$$

利用零阶保持器的特性 (4.2.2), 得到

$$\begin{aligned} \mathbf{x}(kT + T) &= e^{\mathbf{A}_c qh} \mathbf{x}(kT) + \sum_{j=1}^r \left[\sum_{i=1}^{\nu_j} \int_{kT+(i-1)T_j}^{kT+iT_j} e^{\mathbf{A}_c (kT+T-\tau)} d\tau \mathbf{B}_{c_j} u_j(kT + (i-1)T_j) \right] \\ &= \mathbf{A}\mathbf{x}(kT) + \sum_{j=1}^r \left[\sum_{i=1}^{\nu_j} e^{\mathbf{A}_c (T-iT_j)} \int_{-T_j}^0 e^{-\mathbf{A}_c t} dt \mathbf{B}_{c_j} u_j(kT + (i-1)T_j) \right] \end{aligned}$$

$$\begin{aligned}
&= \mathbf{A}\mathbf{x}(kT) + \sum_{j=1}^r \left[\sum_{i=1}^{\nu_j} e^{\mathbf{A}_c(T-iT_j)} \int_0^{T_j} e^{\mathbf{A}_c t} dt \mathbf{B}_{c_j} u_j(kT + (i-1)T_j) \right] \\
&= \mathbf{A}\mathbf{x}(kT) + \sum_{j=1}^r \left[\sum_{i=1}^{\nu_j} \mathbf{A}_h^{q-ip_j} \mathbf{B}_{p_j} u_j(kT + (i-1)T_j) \right] \\
&= \mathbf{A}\mathbf{x}(kT) + \sum_{j=1}^r \sum_{i=1}^{\nu_j} \mathbf{B}_{j_i} u_j(kT - (i-1)T_j). \tag{4.2.4}
\end{aligned}$$

令 (4.2.1) 式中的 $t = kT$ 得到离散化后的输出方程

$$y(kT) = \mathbf{C}\mathbf{x}(kT) + v(kT). \tag{4.2.5}$$

联合 (4.2.4) 式和 (4.2.5) 式得到定理的结论. \square

例 4.2.1 对于双输入单输出连续系统

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}_c \mathbf{x}(t) + \mathbf{B}_{c1} u_1(t) + \mathbf{B}_{c2} u_2(t), \\ y(t) = \mathbf{C}\mathbf{x}(t), \end{cases} \tag{4.2.6}$$

其中

$$\mathbf{A}_c = \begin{bmatrix} -1 & 0 \\ 1 & -2 \end{bmatrix}, \quad \mathbf{B}_{c1} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \mathbf{B}_{c2} = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \quad \text{和} \quad \mathbf{C} = [3 \quad -1].$$

取 $h = 0.1s$, $p_1 = 2$, $p_2 = 3$ 和 $q = 6$, 因此 $\nu_1 = 3$, $\nu_2 = 2$, $T_1 = 2s$, $T_2 = 3s$ 和 $T = 6s$.

首先, 计算 $e^{\mathbf{A}_c t}$,

$$[z\mathbf{I} - \mathbf{A}_c]^{-1} = \begin{bmatrix} \frac{1}{z+1} & 0 \\ \frac{1}{z+1} - \frac{1}{z+2} & \frac{1}{z+2} \end{bmatrix},$$

因此, 有

$$e^{\mathbf{A}_c t} = L^{-1}\{[z\mathbf{I} - \mathbf{A}_c]^{-1}\} = \begin{bmatrix} e^{-t} & 0 \\ e^{-t} - e^{-2t} & e^{-2t} \end{bmatrix}.$$

则, 得到

$$\mathbf{A}_h = e^{\mathbf{A}_c h} = \begin{bmatrix} e^{-0.1} & 0 \\ e^{-0.1} - e^{-0.2} & e^{-0.2} \end{bmatrix}.$$

其次, 计算 \mathbf{B}_{p1} 和 \mathbf{B}_{p2} ,

$$\begin{aligned}
\mathbf{B}_{p1} &= \int_0^{T_1} e^{\mathbf{A}_c t} dt \mathbf{B}_{c1} = \begin{bmatrix} 1 - e^{-0.2} \\ \frac{3}{2} - e^{-0.2} - \frac{1}{2} e^{-0.4} \end{bmatrix}, \\
\mathbf{B}_{p2} &= \int_0^{T_2} e^{\mathbf{A}_c t} dt \mathbf{B}_{c2} = \begin{bmatrix} -2 + 2e^{-0.3} \\ -\frac{1}{2} + 2e^{-0.3} - \frac{3}{2} e^{-0.6} \end{bmatrix}.
\end{aligned}$$

最后, 得到

$$\begin{aligned}
 \mathbf{A} &= e^{\mathbf{A}_c T} = \begin{bmatrix} e^{-0.6} & 0 \\ e^{-0.6} - e^{-1.2} & e^{-1.2} \end{bmatrix} = \begin{bmatrix} 0.5488 & 0 \\ 0.2476 & 0.3012 \end{bmatrix}, \\
 \mathbf{B}_{11} &= \mathbf{A}_h^{q-p_1} \mathbf{B}_{p_1} = \begin{bmatrix} e^{-0.4} - e^{-0.6} \\ e^{-0.4} - 2e^{-0.6} + \frac{5}{2}e^{-0.8} - e^{-1.0} - \frac{1}{2}e^{-1.2} \end{bmatrix} = \begin{bmatrix} 0.1215 \\ -0.3594 \end{bmatrix}, \\
 \mathbf{B}_{12} &= \mathbf{A}_h^{q-2p_1} \mathbf{B}_{p_1} = \begin{bmatrix} e^{-0.2} - e^{-0.4} \\ e^{-0.2} - \frac{1}{2}e^{-0.4} - \frac{1}{2}e^{-0.8} \end{bmatrix} = \begin{bmatrix} 0.1484 \\ 0.2589 \end{bmatrix}, \\
 \mathbf{B}_{13} &= \mathbf{A}_h^{q-3p_1} \mathbf{B}_{p_1} = \begin{bmatrix} 1 - e^{-0.2} \\ \frac{3}{2} - e^{-0.2} - \frac{1}{2}e^{-0.4} \end{bmatrix} = \begin{bmatrix} 0.1813 \\ 0.3461 \end{bmatrix}, \\
 \mathbf{B}_{21} &= \mathbf{A}_h^{q-p_2} \mathbf{B}_{p_2} = \begin{bmatrix} -2e^{-0.3} + 2e^{-0.6} \\ -2e^{-0.3} + \frac{7}{2}e^{-0.6} - \frac{3}{2}e^{-1.2} \end{bmatrix} = \begin{bmatrix} -0.3840 \\ -1.1102 \end{bmatrix}, \\
 \mathbf{B}_{22} &= \mathbf{A}_h^{q-2p_2} \mathbf{B}_{p_2} = \begin{bmatrix} -2 + 2e^{-0.3} \\ -\frac{1}{2} + 2e^{-0.3} - \frac{3}{2}e^{-0.6} \end{bmatrix} = \begin{bmatrix} -0.5184 \\ 0.1584 \end{bmatrix}.
 \end{aligned}$$

进一步地, 可以通过计算得到这个系统的传递函数模型

$$G(z) = \begin{bmatrix} \frac{0.7239z - 0.3371}{z^2 - 0.8500z + 0.1653} & \frac{0.1863z - 0.02875}{z^2 - 0.8500z + 0.1653} & \frac{0.1978z - 0.01877}{z^2 - 0.8500z + 0.1653} \\ -\frac{0.0418z - 0.1672}{z^2 - 0.8500z + 0.1653} & \frac{-1.397z + 0.5099}{z^2 - 0.8500z + 0.1653} & \end{bmatrix}.$$

4.3 辅助模型多新息随机梯度算法

4.3.1 算法的提出

令 z^{-1} 为单位后移算子: $z^{-1}y(kT) = y(kT - T)$. 参照文 [75], 根据 (4.2.3) 式, 有

$$\begin{aligned}
 y(kT) &= \sum_{j=1}^r \sum_{i=1}^{\nu_j} \mathbf{C}(z\mathbf{I} - \mathbf{A})^{-1} \mathbf{B}_{ji} u_j(kT + (i-1)T_j) + v(kT) \\
 &=: \frac{1}{\alpha(z)} \sum_{j=1}^r \sum_{i=1}^{\nu_j} \beta_{ji}(z) u_j(kT + (i-1)T_j) + v(kT), \tag{4.3.1}
 \end{aligned}$$

其中 $\alpha(z)$ 和 $\beta_{ji}(z)$ 是 z^{-1} 的多项式,

$$\begin{aligned}
 \alpha(z) &:= z^{-n} \det[z\mathbf{I} - \mathbf{A}] 1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \cdots + \alpha_n z^{-n}, \quad \alpha_i \in \mathbb{R}^1, \quad i = 1, 2, \dots, n, \\
 \beta_{ji}(z) &:= z^{-n} \mathbf{C} \operatorname{adj}[z\mathbf{I} - \mathbf{A}] \mathbf{B}_{ji} \beta_{ji}(1) z^{-1} + \beta_{ji}(2) z^{-2} + \cdots + \beta_{ji}(n) z^{-n}, \quad \beta_{ji}(l) \in \mathbb{R}^1, \\
 & \quad j = 1, 2, \dots, r, \quad i = 1, 2, \dots, \nu_j, \quad l = 1, 2, \dots, n,
 \end{aligned}$$

其中 α_i 和 $\beta_{ji}(1)$ 是待辨识参数.

定义中间不可测变量 (为未知的无噪输出或真实输出) 如下:

$$y_0(kT) := \frac{1}{\alpha(z)} \sum_{j=1}^r \sum_{i=1}^{\nu_j} \beta_{ji}(z) u_j(kT + (i-1)T_j). \tag{4.3.2}$$

则 (4.3.1) 式可以表达为

$$y(kT) = y_0(kT) + v(kT). \quad (4.3.3)$$

定义参数向量 θ 和信息向量 $\varphi_0(kT)$ 如下:

$$\begin{aligned} \theta &:= [\alpha_1, \alpha_2, \dots, \alpha_n, \beta_{11}(1), \beta_{11}(2), \dots, \beta_{11}(n), \dots, \beta_{1\nu_1}(1), \beta_{1\nu_1}(2), \dots, \\ &\quad \beta_{1\nu_1}(n), \dots, \beta_{r1}(1), \beta_{r1}(2), \dots, \beta_{r1}(n), \dots, \beta_{r\nu_r}(1), \beta_{r\nu_r}(2), \dots, \\ &\quad \beta_{r\nu_r}(n)]^T \in \mathbb{R}^{n_0}, \quad n_0 := n + n(\nu_1 + \nu_2 + \dots + \nu_r), \\ \varphi_0(kT) &:= [\phi_0^T(kT), \phi_1^T(kT), \dots, \phi_r^T(kT)]^T \in \mathbb{R}^{n_0}, \\ \phi_0(kT) &:= [-y_0(kT - T), -y_0(kT - 2T), \dots, -y_0(kT - nT)]^T \in \mathbb{R}^n, \\ \phi_j(kT) &:= [u_j(kT - T), u_j(kT - 2T), \dots, u_j(kT - nT), \dots, \\ &\quad u_j(kT + (\nu_j - 1)T_j - T), u_j(kT + (\nu_j - 1)T_j - 2T), \dots, \\ &\quad u_j(kT + (\nu_j - 1)T_j - nT)]^T \in \mathbb{R}^{n\nu_j}, \quad j = 1, 2, \dots, r. \end{aligned}$$

(4.3.2) 式和 (4.3.3) 式可以等价的写为

$$y_0(kT) = \varphi_0^T(kT)\theta \quad \text{和} \quad y(kT) = \varphi_0^T(kT)\theta + v(kT). \quad (4.3.4)$$

极小化准则函数

$$J(\theta) = \sum_{i=1}^k [y_i(kT) - \varphi_0^T(kT)\theta]^2,$$

然而由于信息向量 $\varphi_0(kT)$ 中含有未知中间变量 $y_0(kT - iT)$ ($i = 1, 2, \dots, n$), 因此用常规的 SG 算法不能得到多率系统 (4.3.4) 的参数 θ 的估计.

利用辅助模型辨识思想^[22], 建立如图 4.2 所示的辅助模型,

$$\hat{y}_0(kT) = \frac{\hat{\beta}(z)}{\hat{\alpha}(z)} \mathbf{u}(kT) \quad \text{或} \quad \hat{y}_0(kT) = \varphi^T(kT)\hat{\theta}(kT),$$

其中

$$\begin{aligned} \hat{\beta}(z) &:= [\beta_{11}(z), \beta_{12}(z), \dots, \beta_{1\nu_1}(z), \dots, \beta_{r1}(z), \beta_{r2}(z), \dots, \beta_{r\nu_r}(z)], \\ \mathbf{u}(kT) &:= [u_1(kT), u_1(kT + T_1), \dots, u_1(kT + (\nu_1 - 1)T_1), \dots, \\ &\quad u_r(kT), u_r(kT + T_r), \dots, u_r(kT + (\nu_r - 1)T_r)]^T, \\ \varphi(kT) &:= [\phi^T(kT), \phi_1^T(kT), \dots, \phi_r^T(kT)]^T, \\ \phi(kT) &:= [-\hat{y}_0(kT - T), -\hat{y}_0(kT - 2T) \dots, -\hat{y}_0(kT - nT)]^T, \end{aligned}$$

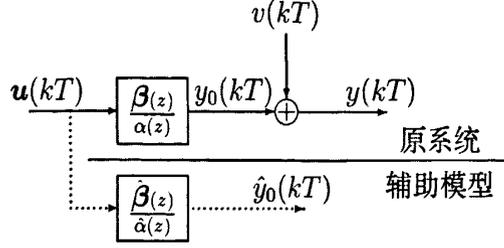


图 4.2 带辅助模型的输出误差系统

Figure 4.2 The output error system with the auxiliary model

其中 $\hat{y}_0(kT)$, $\hat{\varphi}(kT)$ 和 $\hat{\theta}(kT)$ 分别是 $y_0(kT)$, $\varphi_0(kT)$ 和 $\theta(kT)$ 在 kT 时刻的估计值, $\hat{\beta}(z)$ 和 $\hat{\alpha}(z)$ 分别是 $\beta(z)$ 和 $\alpha(z)$ 的估计. 信息向量 $\varphi_0(kT)$ 中不可测的数据 $y_0(kT - iT)$ 用辅助模型的输出 $\hat{y}_0(kT - iT)$ 代替.

参考文 [22], 将信息向量 $\varphi_0(kT)$ 中的未知中间变量 $y_0(kT - iT)$ 用其估计值 $\hat{y}_0(kT - iT)$ 来代替, 进一步地, $\varphi_0(kT)$ 用其估计 $\hat{\varphi}(kT)$ 代替, 得到估计多率系统 (4.3.2) 的参数 θ 的辅助模型随机梯度 (AM-SG) 算法, 如下:

$$\hat{\theta}(kT) = \hat{\theta}(kT - T) + \frac{\varphi(kT)}{r(kT)} e(kT), \quad (4.3.5)$$

$$e(kT) = y(kT) - \varphi(kT)^T(kT) \hat{\theta}(kT - T), \quad (4.3.6)$$

$$r(kT) = r(kT - T) + \|\varphi(kT)\|^2, \quad r(0) = 1, \quad (4.3.7)$$

$$\varphi(kT) = [\phi^T(kT), \phi_1^T(kT), \phi_2^T(kT), \dots, \phi_r^T(kT)]^T, \quad (4.3.8)$$

$$\phi(kT) = [-\hat{y}_0(kT - T), -\hat{y}_0(kT - 2T), \dots, -\hat{y}_0(kT - nT)]^T, \quad (4.3.9)$$

$$\begin{aligned} \phi_j(kT) = & [u_j(kT - T), \dots, u_j(kT - nT), \dots, u_j(kT + (q_j - 1)T_j - T), \\ & \dots, u_j(kT + (q_j - 1)T_j - nT)]^T, \end{aligned} \quad (4.3.10)$$

$$\hat{y}_0(kT) = \varphi^T(kT) \hat{\theta}(kT), \quad (4.3.11)$$

其中 $e(kT)$ 为单新息标量.

为了提高 AM-SG 算法的收敛速度和估计精度, 基于多新息辨识理论, 引入新息长度 p , 将单新息 $e(kT)$ 扩展为多新息向量

$$\mathbf{E}(p, kT) = \begin{bmatrix} y(kT) - \varphi^T(kT)(kT) \hat{\theta}(kT - T) \\ y(kT - T) - \varphi^T(kT)(kT - T) \hat{\theta}(kT - T) \\ \vdots \\ y(kT - pT + T) - \varphi^T(kT)(kT - pT + T) \hat{\theta}(kT - T) \end{bmatrix} \in \mathbb{R}^p,$$

其中 p 是正整数.

定义信息矩阵 $\Phi(p, kT)$ 和输出向量 $\mathbf{Y}(p, kT)$ 如下:

$$\Phi(p, kT) = [\varphi(kT), \varphi(kT - T), \dots, \varphi(kT - pT + T)] \in \mathbb{R}^{n_0 \times p},$$

$$\mathbf{Y}(p, kT) = [y(kT), y(kT - T), \dots, y(kT - pT + T)]^T \in \mathbb{R}^p,$$

则多新息向量可以表达为

$$\mathbf{E}(p, kT) = \mathbf{Y}(p, kT) - \Phi^T(p, kT)\hat{\boldsymbol{\theta}}(kT - T).$$

参照文 [22, 26], 得到估计多率系统 (4.3.2) 的参数 $\boldsymbol{\theta}$ 的辅助模型多新息随机梯度 (AM-MISG) 算法, 如下:

$$\hat{\boldsymbol{\theta}}(kT) = \hat{\boldsymbol{\theta}}(kT - T) + \frac{\boldsymbol{\varphi}(kT)}{r(kT)} \mathbf{E}(p, kT), \quad (4.3.12)$$

$$\mathbf{E}(p, kT) = \mathbf{Y}(p, kT) - \Phi^T(p, kT)\hat{\boldsymbol{\theta}}(kT - T), \quad (4.3.13)$$

$$r(kT) = r(kT - T) + \|\boldsymbol{\varphi}(kT)\|^2, \quad r(0) = 1, \quad (4.3.14)$$

$$\mathbf{Y}(p, kT) = [y(kT), y(kT - T), \dots, y(kT - pT + T)]^T, \quad (4.3.15)$$

$$\Phi(p, kT) = [\boldsymbol{\varphi}(kT), \boldsymbol{\varphi}(kT - T), \dots, \boldsymbol{\varphi}(kT - pT + T)], \quad (4.3.16)$$

$$\boldsymbol{\varphi}(kT) = [\phi^T(kT), \phi_1^T(kT), \phi_2^T(kT), \dots, \phi_r^T(kT)]^T, \quad (4.3.17)$$

$$\boldsymbol{\phi}(kT) = [-\hat{y}_0(kT - T), -\hat{y}_0(kT - 2T), \dots, -\hat{y}_0(kT - nT)]^T, \quad (4.3.18)$$

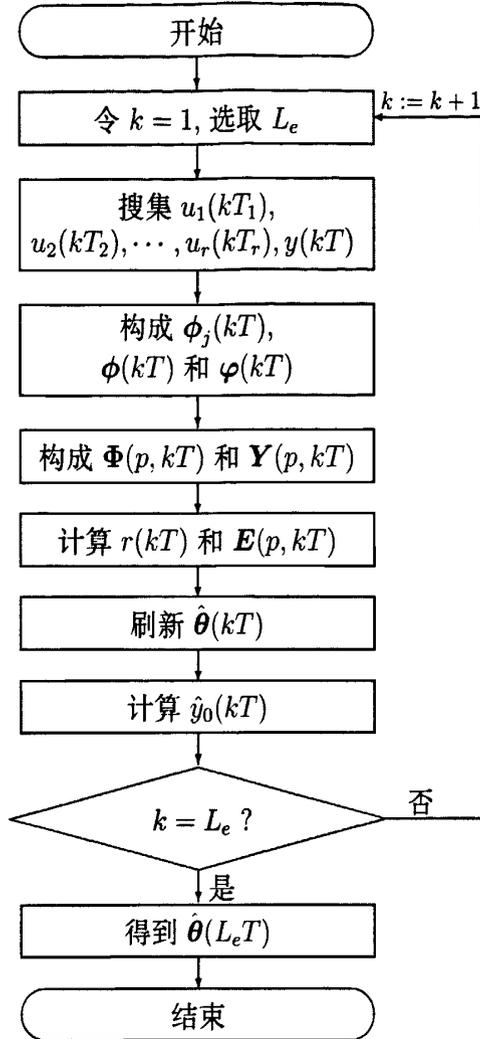
$$\begin{aligned} \boldsymbol{\phi}_j(kT) = & [u_j(kT - T), \dots, u_j(kT - nT), \dots, u_j(kT + (q_j - 1)T_j - T), \\ & \dots, u_j(kT + (q_j - 1)T_j - nT)]^T, \end{aligned} \quad (4.3.19)$$

$$\hat{y}_0(kT) = \boldsymbol{\varphi}^T(kT)\hat{\boldsymbol{\theta}}(kT). \quad (4.3.20)$$

当 $p = 1$ 时, AM-MISG 算法 (4.3.12)–(4.3.20) 变为 AM-SG 算法 (4.3.5)–(4.3.11).

初始化 AM-MISG 算法, 一般地, $\hat{\boldsymbol{\theta}}(0)$ 取为零向量或者很小的实向量, 如 $\hat{\boldsymbol{\theta}}(0) = 10^{-6}\mathbf{1}_{n_0}$. 用 AM-MISG 算法计算参数估计 $\hat{\boldsymbol{\theta}}(kT)$ 的步骤如下:

1. 令 $k = 1$, $\hat{\boldsymbol{\theta}}(0) = \mathbf{1}_{n_0}/10^{-6}$, 并选取数据长度 L_e .
2. 搜集多率输入输出数据 $\{u_1(kT_1), u_2(kT_2), \dots, u_r(kT_r), y(kT)\}$.
3. 由 (4.3.19) 式构成信息向量 $\boldsymbol{\phi}_j(kT)$, 由 (4.3.18) 式构成 $\boldsymbol{\phi}(kT)$, 进一步地, 由 (4.3.17) 式构成 $\boldsymbol{\varphi}(kT)$.
4. 由 (4.3.16) 式构成信息矩阵 $\Phi(p, kT)$, 由 (4.3.15) 式构成输出向量 $\mathbf{Y}(p, kT)$.
5. 由 (4.3.14) 计算 $r(kT)$, 由 (4.3.13) 式计算 $\mathbf{E}(p, kT)$, 由 (4.3.12) 式刷新 $\hat{\boldsymbol{\theta}}(kT)$, 以及由 (4.3.20) 式计算 $\hat{y}_0(kT)$.
6. 如果 $k < L_e$, k 增至 $k + 1$, 返回第 2 步继续计算. 否则, 得到参数估计 $\hat{\boldsymbol{\theta}}(L_e T)$.

图 4.3 计算参数估计 $\hat{\theta}(kT)$ 的流程图Fig. 4.3 The flowchart of computing the parameter estimate $\hat{\theta}(kT)$

计算参数估计 $\hat{\theta}(kT)$ 的流程图如图 4.3 所示.

与 AM-SG 算法 (4.3.5)-(4.3.11) 比较, AM-MISG 算法 (4.3.12)-(4.3.20) 充分利用输入输出数据, 因此其收敛速度快且估计精度高.

4.3.2 仿真试验

例 4.3.1 对于双输入单输出多率系统, 其输入刷新周期分别为 $T_1 = 2h$ 和 $T_2 = 3h$, 输出采样周期 $T = 6h$, 取基周期 $h = 2$, 相应的传递函数模型为

$$y(kT) = \frac{1}{\alpha(z)} [\beta_{11}(z)u_1(kT) + \beta_{12}(z)u_1(kT + T_1) + \beta_{13}(z)u_1(kT + 2T_1) + \beta_{21}(z)u_2(kT) + \beta_{22}(z)u_2(kT + T_2)] + v(kT),$$

其中

$$\begin{aligned}\alpha(z) &= 1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} = 1 + 0.58z^{-1} + 0.36z^{-2}, \\ \beta_{11}(z) &= \beta_{11}(1)z^{-1} + \beta_{11}(2)z^{-2} = -0.30z^{-1} + 0.36z^{-2}, \\ \beta_{12}(z) &= \beta_{12}(1)z^{-1} + \beta_{12}(2)z^{-2} = -0.65z^{-1} + 0.57z^{-2}, \\ \beta_{13}(z) &= \beta_{13}(1)z^{-1} + \beta_{13}(2)z^{-2} = 0.57z^{-1} - 0.82z^{-2}, \\ \beta_{21}(z) &= \beta_{21}(1)z^{-1} + \beta_{21}(2)z^{-2} = -0.40z^{-1} + 0.22z^{-2}, \\ \beta_{22}(z) &= \beta_{22}(1)z^{-1} + \beta_{22}(2)z^{-2} = 0.31z^{-1} + 1.20z^{-2}.\end{aligned}$$

参数向量和信息向量为

$$\begin{aligned}\boldsymbol{\theta} &= [\alpha_1, \alpha_2, \beta_{11}(1), \beta_{11}(2), \beta_{12}(1), \beta_{12}(2), \beta_{13}(1), \beta_{13}(2), \beta_{21}(1), \beta_{21}(2), \\ &\quad \beta_{22}(1), \beta_{22}(2)]^T \\ &= [0.58, 0.36, -0.30, 0.36, -0.65, 0.57, 0.57, -0.82, -0.40, 0.22, \\ &\quad 0.31, 1.20]^T, \\ \boldsymbol{\varphi}(kT) &= [\boldsymbol{\phi}^T(kT), \boldsymbol{\phi}_1^T(kT), \boldsymbol{\phi}_2^T(kT)]^T, \\ \boldsymbol{\phi}(kT) &= [-y_0(kT - T), -y_0(kT - 2T)]^T, \\ \boldsymbol{\phi}_1(kT) &= [u_1(kT - T), u_1(kT - 2T), u_1(kT + T_1 - T), u_1(kT + T_1 - 2T), \\ &\quad u_1(kT + 2T_1 - T), u_1(kT + 2T_1 - 2T)]^T, \\ \boldsymbol{\phi}_2(kT) &= [u_2(kT - T), u_2(kT - 2T), u_2(kT + T_2 - T), u_2(kT + T_2 - 2T)]^T.\end{aligned}$$

则

$$y_0(kT) = \boldsymbol{\varphi}^T(kT)\boldsymbol{\theta}, \quad y(kT) = y_0(kT) + v(kT),$$

仿真时, $\{u_1(kT_1)\}$ 和 $\{u_2(kT_2)\}$ 采用零均值单位方差不相关可测的持续激励信号序列. $\{v(kT)\}$ 采用零均值方差为 $\sigma^2 = 0.50^2$ 的白噪声序列. 应用 AM-SG 算法 (4.3.5)-(4.3.11) 和 AM-MISG 算法 (4.3.12)-(4.3.20) 估计这个系统的参数. 参数估计和参数估计误差如表 4.1-4.4 所示, 参数估计误差 $\delta := \|\hat{\boldsymbol{\theta}}(kT) - \boldsymbol{\theta}\|/\|\boldsymbol{\theta}\|$ 随 k 变化曲线如图 4.4 所示. 系统的噪信比为 $\delta_{\text{ns}} = 12.61\%$.

从表 4.1-4.4 和图 4.4 可以得到, 由 AM-MISG 算法估计得到的参数误差随着 k 的逐渐增加会越来越小, 且趋于零. 与 AM-SG 算法比较, AM-MISG 算法 $p \geq 2$ 有更快的收敛速度和高的估计精度. 这个仿真例子验证了 AM-SG 算法和 AM-MISG 算法的有效性.

表 4.1 AM-SG 参数估计和参数估计误差
Table 4.1 The AM-SG parameter estimates and their errors

k	100	500	1000	2000	3000	4000	5000
$\alpha_1 = 0.58000$	0.46974	0.52407	0.53616	0.55456	0.56098	0.56620	0.56985
$\alpha_2 = 0.36000$	0.04264	0.07133	0.10340	0.12504	0.13632	0.14632	0.15382
$\beta_{11}(1) = -0.30000$	-0.13253	-0.14593	-0.15812	-0.16305	-0.16733	-0.17077	-0.17342
$\beta_{11}(2) = 0.36000$	0.37355	0.37839	0.38136	0.38442	0.38610	0.38815	0.38963
$\beta_{12}(1) = -0.65000$	-0.48219	-0.49038	-0.49557	-0.50049	-0.50311	-0.50486	-0.50581
$\beta_{12}(2) = 0.57000$	0.01697	0.06993	0.09183	0.11094	0.12115	0.12810	0.13363
$\beta_{13}(1) = 0.57000$	0.36669	0.38466	0.39272	0.39946	0.40306	0.40589	0.40842
$\beta_{13}(2) = -0.82000$	-0.58643	-0.65834	-0.68725	-0.71227	-0.72190	-0.73081	-0.73663
$\beta_{21}(1) = -0.40000$	-0.57070	-0.55063	-0.54094	-0.53676	-0.53313	-0.52974	-0.52741
$\beta_{21}(2) = 0.22000$	-0.17258	-0.13550	-0.12551	-0.11074	-0.10413	-0.09921	-0.09578
$\beta_{22}(1) = 0.31000$	0.48801	0.45427	0.45220	0.44149	0.43374	0.43192	0.42996
$\beta_{22}(2) = 1.20000$	0.24414	0.33689	0.37358	0.41017	0.43105	0.44518	0.45468
δ (%)	63.80501	57.10521	54.33061	51.75382	50.35554	49.37843	48.68755

表 4.2 AM-MISG 参数估计和参数估计误差 ($p = 2$)
Table 4.2 The AM-MISG parameter estimates and their errors ($p = 2$)

k	100	500	1000	2000	3000	4000	5000
$\alpha_1 = 0.58000$	0.50885	0.58679	0.59277	0.61317	0.61712	0.62039	0.62209
$\alpha_2 = 0.36000$	0.13332	0.18858	0.23683	0.26215	0.27558	0.28580	0.29324
$\beta_{11}(1) = -0.30000$	-0.16635	-0.18451	-0.19963	-0.20530	-0.21075	-0.21500	-0.21798
$\beta_{11}(2) = 0.36000$	0.37550	0.37959	0.38107	0.38368	0.38475	0.38667	0.38780
$\beta_{12}(1) = -0.65000$	-0.50588	-0.51446	-0.51986	-0.52499	-0.52779	-0.52929	-0.53029
$\beta_{12}(2) = 0.57000$	0.07892	0.16142	0.19214	0.21790	0.23089	0.23986	0.24724
$\beta_{13}(1) = 0.57000$	0.39568	0.42033	0.43104	0.43834	0.44206	0.44454	0.44683
$\beta_{13}(2) = -0.82000$	-0.65124	-0.75202	-0.78409	-0.80912	-0.81566	-0.82276	-0.82652
$\beta_{21}(1) = -0.40000$	-0.56778	-0.53286	-0.51756	-0.51142	-0.50636	-0.50099	-0.49777
$\beta_{21}(2) = 0.22000$	-0.09402	-0.04600	-0.03532	-0.01838	-0.01185	-0.00631	-0.00317
$\beta_{22}(1) = 0.31000$	0.49934	0.44205	0.43609	0.41875	0.40733	0.40438	0.40145
$\beta_{22}(2) = 1.20000$	0.33645	0.49157	0.55057	0.60498	0.63529	0.65561	0.66870
δ (%)	55.88134	45.47126	41.61051	38.29935	36.52138	35.32884	34.51899

表 4.3 AM-MISG 参数估计和参数估计误差 ($p = 6$)
 Table 4.3 The AM-MISG parameter estimates and their errors ($p = 6$)

k	100	500	1000	2000	3000	4000	5000
$\alpha_1 = 0.58000$	0.54765	0.63453	0.61929	0.62246	0.61328	0.61239	0.60868
$\alpha_2 = 0.36000$	0.28103	0.37997	0.38429	0.37687	0.38083	0.37683	0.37764
$\beta_{11}(1) = -0.30000$	-0.22068	-0.25384	-0.25931	-0.26287	-0.26740	-0.27196	-0.27354
$\beta_{11}(2) = 0.36000$	0.34319	0.35129	0.34799	0.35395	0.35333	0.35670	0.35767
$\beta_{12}(1) = -0.65000$	-0.50902	-0.53914	-0.55654	-0.56652	-0.57265	-0.57599	-0.57914
$\beta_{12}(2) = 0.57000$	0.23430	0.38129	0.42651	0.45327	0.46316	0.47173	0.47910
$\beta_{13}(1) = 0.57000$	0.42110	0.45767	0.47505	0.48672	0.49365	0.49619	0.49888
$\beta_{13}(2) = -0.82000$	-0.73649	-0.84498	-0.85856	-0.86851	-0.86457	-0.86522	-0.86339
$\beta_{21}(1) = -0.40000$	-0.46195	-0.43231	-0.42472	-0.42429	-0.42328	-0.41737	-0.41587
$\beta_{21}(2) = 0.22000$	0.10478	0.14482	0.15319	0.16340	0.16386	0.16893	0.16960
$\beta_{22}(1) = 0.31000$	0.47413	0.38270	0.37455	0.35567	0.34609	0.34444	0.34192
$\beta_{22}(2) = 1.20000$	0.63470	0.88510	0.96029	1.01211	1.03761	1.05614	1.06479
δ (%)	35.93795	20.64063	16.26615	13.36463	11.91119	10.89418	10.28458

表 4.4 AM-MISG 参数估计和参数估计误差 ($p = 12$)
 Table 4.4 The AM-MISG parameter estimates and their errors ($p = 12$)

k	100	500	1000	2000	3000	4000	5000
$\alpha_1 = 0.58000$	0.61771	0.61588	0.59055	0.59044	0.57969	0.58509	0.58144
$\alpha_2 = 0.36000$	0.36878	0.41065	0.37581	0.36371	0.36770	0.36190	0.36595
$\beta_{11}(1) = -0.30000$	-0.29222	-0.30741	-0.29244	-0.29049	-0.29180	-0.29558	-0.29414
$\beta_{11}(2) = 0.36000$	0.31647	0.32523	0.32591	0.34159	0.34064	0.34785	0.35005
$\beta_{12}(1) = -0.65000$	-0.47948	-0.56484	-0.59742	-0.60984	-0.61746	-0.62054	-0.62385
$\beta_{12}(2) = 0.57000$	0.38008	0.53158	0.56189	0.56763	0.56309	0.56415	0.56611
$\beta_{13}(1) = 0.57000$	0.46859	0.51132	0.54162	0.55436	0.56070	0.56042	0.56086
$\beta_{13}(2) = -0.82000$	-0.73855	-0.82113	-0.82956	-0.83887	-0.83384	-0.83452	-0.83138
$\beta_{21}(1) = -0.40000$	-0.37991	-0.39882	-0.39971	-0.40449	-0.40763	-0.40066	-0.40081
$\beta_{21}(2) = 0.22000$	0.20092	0.21614	0.22105	0.22506	0.21676	0.22240	0.22090
$\beta_{22}(1) = 0.31000$	0.42555	0.35153	0.34951	0.33395	0.32739	0.32838	0.32547
$\beta_{22}(2) = 1.20000$	0.89173	1.10912	1.15259	1.16587	1.17366	1.18323	1.18254
δ (%)	21.59399	8.08506	4.67345	3.30350	2.65871	2.20320	1.97512

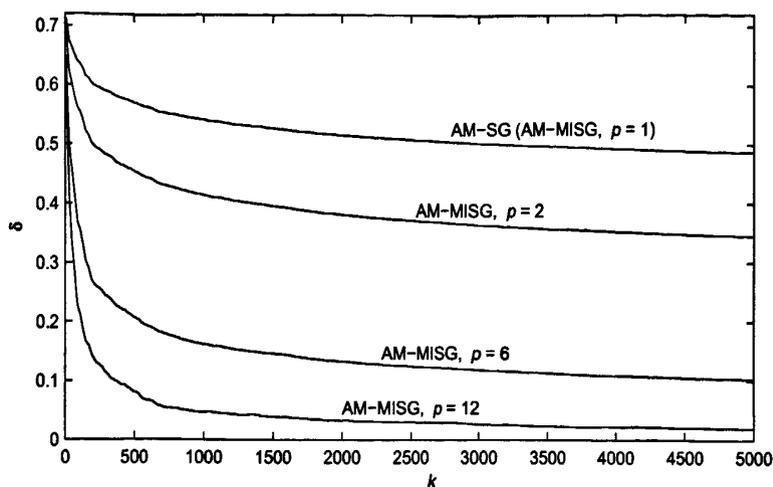


图 4.4 参数估计误差 δ 随数据长度 k 变化曲线
Fig. 4.4 The parameter estimation errors δ versus k

本节的内容已于 2010 年被国际期刊《Journal of Systems Engineering and Electronics》录用，以及 2010 年 5 月在徐州召开的 22nd Chinese Control and Decision Conference 国际会议录用。

4.4 辅助模型最小二乘算法

本节利用最小二乘原理和辅助模型辨识思想研究输出误差多率系统的辨识问题。提出辅助模型最小二乘算法估计系统的参数，应用鞅收敛定理，从理论上证明提出算法的收敛性。最后用数值仿真例子验证算法的有效性。

4.4.1 算法的提出

对于多率多输入系统的模型 (4.3.4)，由于信息向量 $\varphi_0(kT)$ 中含有未知中间变量 $y_0(kT - iT)$ ($i = 1, 2, \dots, n$)，因此用常规的 RLS 算法不能得到多率系统 (4.3.4) 的参数 θ 的估计。

同上一节，基于辅助模型辨识思想^[22]，建立辅助模型

$$\hat{y}_0(kT) = \varphi^T(kT)\hat{\theta}(kT),$$

将信息向量 $\varphi_0(kT)$ 中的未知中间变量 $y_0(kT - iT)$ 辅助模型输出 $\hat{y}_0(kT - iT)$ 来代替，进一步地， $\varphi_0(kT)$ 用其估计 $\varphi(kT)$ 代替，得到估计多率系统 (4.3.4) 的参数 θ 的辅助模型最小二乘 (AM-RLS) 算法，如下：

$$\hat{\theta}(kT) = \hat{\theta}(kT - T) + \mathbf{P}(kT)\varphi(kT)[y(kT) - \varphi^T(kT)\hat{\theta}(kT - T)], \quad (4.4.1)$$

$$\mathbf{P}^{-1}(kT) = \mathbf{P}^{-1}(kT - T) + \varphi(kT)\varphi^T(kT), \quad (4.4.2)$$

$$\boldsymbol{\varphi}(kT) = [\boldsymbol{\phi}^T(kT), \boldsymbol{\phi}_1^T(kT), \dots, \boldsymbol{\phi}_r^T(kT)]^T, \quad (4.4.3)$$

$$\boldsymbol{\phi}(kT) = [-\hat{y}_0(kT - T), -\hat{y}_0(kT - 2T), \dots, -\hat{y}_0(kT - nT)]^T, \quad (4.4.4)$$

$$\begin{aligned} \boldsymbol{\phi}_j(kT) = & [u_j(kT - T), \dots, u_j(kT - nT), \dots, u_j(kT + (\nu_j - 1)T_j - T), \\ & \dots, u_j(kT + (\nu_j - 1)T_j - nT)]^T \in \mathbb{R}^{n\nu_j}, j = 1, 2, \dots, r. \end{aligned} \quad (4.4.5)$$

$$\hat{y}_0(kT) = \boldsymbol{\varphi}^T(kT)\hat{\boldsymbol{\theta}}(kT). \quad (4.4.6)$$

其中 $\boldsymbol{P}(kT)$ 是参数估计误差协方差矩阵.

初始化 AM-RLS 算法, 令 $\boldsymbol{P}(0) = p_0 \boldsymbol{I}_{n_0 \times n_0}$, 一般地, p_0 取很大的正数, 如 $p_0 = 10^6$; $\hat{\boldsymbol{\theta}}(0)$ 取零向量或者很小的正实向量, 如 $\hat{\boldsymbol{\theta}}(0) = \mathbf{1}_{n_0}/p_0$. 用 AM-RLS 算法计算参数估计 $\hat{\boldsymbol{\theta}}(kT)$ 的步骤如下:

1. 令 $k = 1$, $\hat{\boldsymbol{\theta}}(0) = 10^{-6} \mathbf{1}_{n_0}$, $\boldsymbol{P}(0) = p_0 \boldsymbol{I}_{n_0 \times n_0}$, $\hat{y}_0(kT - iT) = 1/p_0$, 并选取数据长度 L_e .
2. 搜集可测量的多率输入输出数据 $\{u_1(kT_1), u_2(kT_2), \dots, u_r(kT_r), y(kT) : k = 0, 1, 2, \dots\}$.
3. 由 (4.4.4) 式构成信息向量 $\boldsymbol{\phi}(kT)$, 由 (4.4.5) 式构成 $\boldsymbol{\phi}_j(kT)$ ($j = 1, 2, \dots, r$), 以及由 (4.4.3) 式构成 $\boldsymbol{\varphi}(kT)$.
4. 由 (4.4.2) 计算 $\boldsymbol{P}(kT)$, 由 (4.4.1) 式刷新 $\hat{\boldsymbol{\theta}}(kT)$, 以及由式 (4.4.6) 计算 $\hat{y}_0(kT)$.
5. 如果 $k < L_e$, k 增至 $k + 1$, 返回第 2 步继续计算. 否则, 得到参数估计 $\hat{\boldsymbol{\theta}}(L_e T)$.

用 AM-RLS 算法 (4.4.1)–(4.4.1) 计算参数估计 $\hat{\boldsymbol{\theta}}(kT)$ 的流程图如图 4.5 所示.

4.4.2 算法收敛性

令 \boldsymbol{X} 是方阵, $\lambda_{\max}[\boldsymbol{X}]$ 和 $\lambda_{\min}[\boldsymbol{X}]$ 分别代表 \boldsymbol{X} 的最大和最小特征值. 如果 $f(k) \rightarrow 0$ 和 $g(k) \rightarrow 0$, 当 $k \rightarrow \infty$, $f(k) = O(g(k))$ 是指存在常数 $\delta_1 > 0$ 和 k_0 使得 $\left| \frac{f(k)}{g(k)} \right| \leq \delta_1, k \geq k_0$, $f(k) = o(g(k))$ 是指 $\frac{f(k)}{g(k)} \rightarrow 0$, 当 $k \rightarrow \infty$.

设 $\{v(kT), \mathcal{F}_k\}$ 是定义在概率空间 $\{\Omega, \mathcal{F}_k, P\}$ 的鞅差向量序列, $\{\mathcal{F}_k\}$ 是由 $v(kT)$ 生成的 σ 代数序列, 即, $\mathcal{F}_k = \sigma(v(kT), v(kT - T), v(kT - 2T), \dots)$. 噪声序列 $\{v(kT)\}$ 满足

$$(A1) \quad E[v(kT)|\mathcal{F}_{k-1}] = 0, \text{ a.s.}, \quad E[v^2(kT)|\mathcal{F}_{k-1}] \leq \sigma_v^2, \text{ a.s.}$$

定义参数估计误差向量 $\tilde{\boldsymbol{\theta}}(kT)$ 和一个非负有限方程 $V(kT)$ 如下:

$$\begin{aligned} \tilde{\boldsymbol{\theta}}(kT) & := \hat{\boldsymbol{\theta}}(kT) - \boldsymbol{\theta}, \\ V(kT) & := \tilde{\boldsymbol{\theta}}^T(kT) \boldsymbol{P}^{-1}(kT) \tilde{\boldsymbol{\theta}}(kT). \end{aligned} \quad (4.4.7)$$

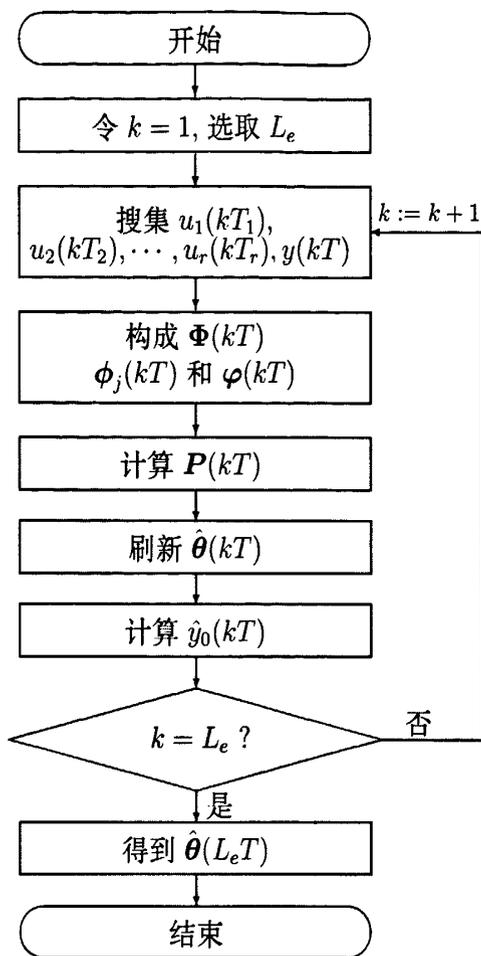


图 4.5 计算参数估计 $\hat{\theta}(kT)$ 的流程图

Fig. 4.5 The flowchart of computing the parameter estimate $\hat{\theta}(kT)$

引理 4.3.1 对于多率系统 (4.3.4) 和 AM-RLS 算法 (4.4.1)–(4.4.6), 对任意 $c > 1$, 有下面不等式成立

$$\sum_{i=1}^{\infty} \frac{\varphi^T(iT)P(iT)\varphi(iT)}{\{\ln|\mathbf{P}^{-1}(iT)|\}^c} < \infty, \text{ a.s.}$$

证明: 根据 $P(kT)$ 的定义, 有

$$\begin{aligned} \mathbf{P}^{-1}(iT - T) &= \mathbf{P}^{-1}(iT) - \varphi(iT)\varphi^T(iT) \\ &= \mathbf{P}^{-1}(iT)[\mathbf{I} - \mathbf{P}(iT)\varphi(iT)\varphi^T(iT)]. \end{aligned}$$

对上式两边取行列式, 并利用公式 $|\mathbf{I} + \mathbf{AB}| = |\mathbf{I} + \mathbf{BA}|$ 得到

$$\begin{aligned} |\mathbf{P}^{-1}(iT - T)| &= |\mathbf{P}^{-1}(iT)| |\mathbf{I} - \mathbf{P}(iT)\varphi(iT)\varphi^T(iT)| \\ &= |\mathbf{P}^{-1}(iT)| [1 - \varphi^T(iT)\mathbf{P}(iT)\varphi(iT)]. \end{aligned}$$

因此,

$$\boldsymbol{\varphi}^T(iT)\mathbf{P}(iT)\boldsymbol{\varphi}(iT) = \frac{\mathbf{P}^{-1}(iT) - \mathbf{P}^{-1}(iT - T)}{\mathbf{P}^{-1}(iT)}.$$

对上式从 $i = 1$ 到 $i = \infty$ 累加得到

$$\begin{aligned} \sum_{i=1}^{\infty} \frac{\boldsymbol{\varphi}^T(iT)\mathbf{P}(iT)\boldsymbol{\varphi}(iT)}{\{\ln|\mathbf{P}^{-1}(iT)|\}^c} &= \sum_{i=1}^{\infty} \frac{|\mathbf{P}^{-1}(iT)| - |\mathbf{P}^{-1}(iT - T)|}{\{|\mathbf{P}^{-1}(iT)| \ln|\mathbf{P}^{-1}(iT)|\}^c} \\ &\leq \int_{|\mathbf{P}^{-1}(0)|}^{|\mathbf{P}^{-1}(\infty)|} \frac{dx}{x[\ln x]^c} \\ &= \frac{1}{1-c} \left[\frac{1}{[\ln|\mathbf{P}^{-1}(\infty)|]^{1-c}} - \frac{1}{[\ln|\mathbf{P}^{-1}(0)|]^{1-c}} \right] < \infty. \quad (4.4.8) \end{aligned}$$

其中 $\mathbf{P}^{-1}(iT)$ 是 i 的非减函数.

引理 4.3.1 证明完毕. □

引理 4.3.2 对于多率系统 (4.3.4) 和 AM-RLS 算法 (4.4.1)–(4.4.6), 假设 (A1) 成立且 $\frac{1}{\alpha(z)} - \frac{1}{2}$ 是严格正实函数. 则下面的不等式成立

$$E[V(kT) + S(kT)|\mathcal{F}_{k-1}] \leq V(kT - T) + S(kT - T) + 2\boldsymbol{\varphi}^T(kT)\mathbf{P}(kT)\boldsymbol{\varphi}(kT)\sigma_v^2$$

和 $S(kT) \geq 0$, 其中

$$\begin{aligned} S(kT) &:= 2 \sum_{i=1}^k \tilde{u}(iT)\tilde{y}(iT), \quad \tilde{u}(iT) := -\boldsymbol{\varphi}^T(iT)\tilde{\boldsymbol{\theta}}(iT), \\ \tilde{y}(iT) &:= \frac{1}{2}\boldsymbol{\varphi}^T(iT)\tilde{\boldsymbol{\theta}}(iT) + [y(iT) - \hat{y}_0(iT) - v(iT)]. \end{aligned}$$

证明: 定义新息 $e(kT)$ 和残差 $\eta(kT)$ 如下:

$$e(kT) := y(kT) - \boldsymbol{\varphi}^T(kT)\hat{\boldsymbol{\theta}}(kT - T), \quad (4.4.9)$$

$$\eta(kT) := y(kT) - \boldsymbol{\varphi}^T(kT)\hat{\boldsymbol{\theta}}(kT). \quad (4.4.10)$$

从 (4.3.4) 式, (4.4.1) 式和 (4.4.6) 式, 得到

$$\eta(kT) = y(kT) - \hat{y}_0(kT) = y_0(kT) - \hat{y}_0(kT) + v(kT), \quad (4.4.11)$$

$$\begin{aligned} \eta(kT) &= [1 - \boldsymbol{\varphi}^T(kT)\mathbf{P}(kT)\boldsymbol{\varphi}(kT)]e(kT) \\ &= [1 + \boldsymbol{\varphi}^T(kT)\mathbf{P}(kT - T)\boldsymbol{\varphi}(kT)]^{-1}e(kT). \end{aligned} \quad (4.4.12)$$

将 (4.4.1) 式代入 (4.4.7) 式, 并利用 (4.4.9)–(4.4.12) 式得到

$$\tilde{\boldsymbol{\theta}}(kT) = \tilde{\boldsymbol{\theta}}(kT - T) + \mathbf{P}(kT)\boldsymbol{\varphi}(kT)e(kT)$$

$$= \tilde{\theta}(kT - T) + \mathbf{P}(kT - T)\varphi(kT)\eta(kT). \quad (4.4.13)$$

利用 (4.4.2) 式和 (4.4.9)-(4.4.13) 式, 以及参照 $\tilde{u}(kT)$ 和 $\tilde{y}(kT)$ 的定义, 由 (4.4.7) 式可以得到

$$\begin{aligned} V(kT) &= V(kT - T) + [\varphi^T(kT)\tilde{\theta}(kT)]^2 + 2\eta(kT)\varphi^T(kT)\tilde{\theta}(kT) \\ &\quad - \varphi^T(kT)\mathbf{P}^{-1}(kT)\varphi(kT)[1 - \varphi^T(kT)\mathbf{P}^{-1}(kT)\varphi(kT)]e^2(kT) \\ &\leq V(kT - T) + [\varphi^T(kT)\tilde{\theta}(kT)]^2 + 2\eta(kT)\varphi^T(kT)\tilde{\theta}(kT) \\ &= V(kT - T) + 2\varphi^T(kT)\tilde{\theta}(kT)v(kT) \\ &\quad + 2\varphi^T(kT)\tilde{\theta}(kT)\left[\frac{1}{2}\varphi^T(kT)\tilde{\theta}(kT) + \eta(kT) - v(kT)\right] \\ &\leq V(kT - T) - 2\tilde{u}^T(kT)\tilde{y}(kT) \\ &\quad + 2\varphi^T(kT)[\tilde{\theta}(kT - T) + \mathbf{P}(kT)\varphi(kT)e(kT)]v(kT) \\ &= V(kT - T) - 2\tilde{u}^T(kT)\tilde{y}(kT) + 2\varphi^T(kT)\tilde{\theta}(kT - T)v(kT) \\ &\quad + 2\varphi^T(kT)\mathbf{P}(kT)\varphi(kT)\{[e(kT) - v(kT)]^T v(kT) + v^2(kT)\}. \end{aligned} \quad (4.4.14)$$

这里应用了一个事实^[22]

$$1 - \varphi^T(kT)\mathbf{P}^{-1}(kT)\varphi(kT) = [1 + \varphi^T(kT)\mathbf{P}^{-1}(kT - T)\varphi(kT)]^{-1} \geq 0.$$

由于 $V(kT - T)$, $\varphi^T(kT)\tilde{\theta}(kT - T)$, $e(kT) - v(kT)$ 和 $\varphi^T(kT)\mathbf{P}(kT)\varphi(kT)$ 与 $v(kT)$ 独立, 且 \mathcal{F}_{k-1} 是可测的, 对 (4.4.14) 式 \mathcal{F}_{k-1} 取条件期望, 应用 (A1) 得到

$$\mathbb{E}[V(kT)|\mathcal{F}_{k-1}] \leq V(kT - T) - 2\mathbb{E}[\tilde{u}^T(kT)\tilde{y}(kT)|\mathcal{F}_{k-1}] + 2\varphi^T(kT)\mathbf{P}(kT)\varphi(kT)\sigma_v^2. \quad (4.4.15)$$

从 (4.4.11) 式和 (4.3.2) 式得到,

$$\begin{aligned} \alpha(z)[\eta(kT) - v(kT)] &= \alpha(z)[y_0(kT) - \hat{y}_0(kT)] \\ &= \sum_{j=1}^r \sum_{i=1}^{\nu_j} \beta_{ji}(z)u_{ji}(kT + (i-1)T_j) - \alpha(z)\hat{y}_0(kT) \\ &= -\hat{y}_0(kT) + \varphi^T(kT)\theta = -\varphi^T(kT)\hat{\theta}(kT) + \varphi^T(kT)\theta \\ &= -\varphi^T(kT)\tilde{\theta}(kT) = \tilde{u}(kT). \end{aligned}$$

由于 $\frac{1}{\alpha(z)} - \frac{1}{2}$ 是严格正实函数, 参照文 [13] Appendix C, 有下面不等式成立

$$S(kT) \geq 0, \text{ a.s.}$$

(4.4.15) 式两边同时加上 $S(kT)$ 即可得到引理 4.3.2. □

定理 4.3.1 对于多率系统 (4.3.4) 和 AM-RLS 算法 (4.4.1)-(4.4.6), 设 (A1) 和引理 4.3.2 的条件成立, 则参数估计误差满足

$$\|\hat{\theta}(kT) - \theta\|^2 = O\left(\frac{\{\ln \lambda_{\max}[\mathbf{P}^{-1}(kT)]\}^c}{\lambda_{\min}[\mathbf{P}^{-1}(kT)]}\right), \text{ a.s., } c > 1.$$

证明: 根据 $V(kT)$ 的定义 (4.4.7) 式, 有

$$\|\tilde{\theta}(kT)\|^2 \leq \frac{\tilde{\theta}^T(kT)\mathbf{P}^{-1}(kT)\tilde{\theta}(kT)}{\lambda_{\min}[\mathbf{P}^{-1}(kT)]} = \frac{V(kT)}{\lambda_{\min}[\mathbf{P}^{-1}(kT)]}. \quad (4.4.16)$$

令

$$W(kT) := \frac{V(kT) + S(kT)}{[\ln |\mathbf{P}^{-1}(kT)|]^c}.$$

由于 $\ln |\mathbf{P}^{-1}(kT)|$ 是非减的, 由引理 4.3.3, 有

$$\begin{aligned} E[W(kT)|\mathcal{F}_{k-1}] &\leq \frac{V(kT - T) + S(kT - T)}{[\ln |\mathbf{P}^{-1}(kT)|]^c} + \frac{2\varphi^T(kT)\mathbf{P}(kT)\varphi(kT)}{[\ln |\mathbf{P}^{-1}(kT)|]^c} \sigma_v^2 \\ &\leq W(kT - T) + \frac{2\varphi^T(kT)\mathbf{P}(kT)\varphi(kT)}{[\ln |\mathbf{P}^{-1}(kT)|]^c} \sigma_v^2. \end{aligned} \quad (4.4.17)$$

利用引理 4.3.2, 从 $k = 1$ 到 $k = \infty$ 累加 (4.4.17) 式右边的最后一项, 并对 (4.4.17) 应用鞅收敛定理, 可以得到 $W(kT)$ 一致收敛于一个有限随机变量 W_0 , 即,

$$W(kT) = \frac{V(kT) + S(kT)}{[\ln |\mathbf{P}^{-1}(kT)|]^c} \rightarrow W_0 < \infty, \text{ a.s.}$$

由引理 4.3.3, $S(kT) \geq 0$, 可以得到

$$V(kT) = O([\ln |\mathbf{P}^{-1}(kT)|]^c), \text{ a.s.}$$

利用 (4.4.16) 式得到

$$\|\tilde{\theta}(kT)\|^2 = O\left(\frac{\{\ln \lambda_{\max}[\mathbf{P}^{-1}(kT)]\}^c}{\lambda_{\min}[\mathbf{P}^{-1}(kT)]}\right), \text{ a.s., } c > 1.$$

定理 4.3.1 证明完毕. □

4.4.3 仿真试验

例 4.4.1 对于 3 个输入单输出多率系统, 3 个输入刷新周期分别为 $T_1 = 2h$, $T_2 = 3h$ 和 $T_3 = h$, 输出采样周期 $T = 6h$, 取基周期 $h = 2$, 相应的传递函数模型为

$$\begin{aligned} y(kT) = \frac{1}{\alpha(z)} &[\beta_{11}(z)u_1(kT) + \beta_{12}(z)u_1(kT + T_1) + \beta_{13}(z)u_1(kT + 2T_1) \\ &+ \beta_{21}(z)u_2(kT) + \beta_{22}(z)u_2(kT + T_2) + \beta_{31}(z)u_3(kT) + \beta_{32}(z)u_3(kT + T_3)] \end{aligned}$$

$$+\beta_{33}(z)u_3(kT+2T_3)+\beta_{34}(z)u_3(kT+3T_3)+\beta_{35}(z)u_3(kT+4T_3) \\ +\beta_{36}(z)u_3(kT+5T_3)]+v(kT),$$

其中

$$\begin{aligned} \alpha(z) &= 1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \alpha_3 z^{-3} = 1 + 0.600z^{-1} + 0.400z^{-2} + 0.300z^{-3}, \\ \beta_{11}(z) &= \beta_{11}(1)z^{-1} + \beta_{11}(2)z^{-2} + \beta_{11}(3)z^{-3} = -0.200z^{-1} + 0.500z^{-2} - 0.400z^{-3}, \\ \beta_{12}(z) &= \beta_{12}(1)z^{-1} + \beta_{12}(2)z^{-2} + \beta_{12}(3)z^{-3} = -0.650z^{-1} + 0.600z^{-2} + 0.100z^{-3}, \\ \beta_{13}(z) &= \beta_{13}(1)z^{-1} + \beta_{13}(2)z^{-2} + \beta_{13}(3)z^{-3} = 1.000z^{-1} - 0.500z^{-2} + 0.400z^{-3}, \\ \beta_{21}(z) &= \beta_{21}(1)z^{-1} + \beta_{21}(2)z^{-2} + \beta_{21}(3)z^{-3} = -0.400z^{-1} + 0.800z^{-2} - 0.200z^{-3}, \\ \beta_{22}(z) &= \beta_{22}(1)z^{-1} + \beta_{22}(2)z^{-2} + \beta_{22}(3)z^{-3} = 0.700z^{-1} + 0.200z^{-2} + 0.300z^{-3}, \\ \beta_{31}(z) &= \beta_{31}(1)z^{-1} + \beta_{31}(2)z^{-2} + \beta_{31}(3)z^{-3} = 1.000z^{-1} - 0.900z^{-2} + 0.200z^{-3}, \\ \beta_{32}(z) &= \beta_{32}(1)z^{-1} + \beta_{32}(2)z^{-2} + \beta_{32}(3)z^{-3} = 1.000z^{-1} - 0.750z^{-2} + 0.125z^{-3}, \\ \beta_{33}(z) &= \beta_{33}(1)z^{-1} + \beta_{33}(2)z^{-2} + \beta_{33}(3)z^{-3} = 1.200z^{-1} - 1.080z^{-2} + 0.240z^{-3}, \\ \beta_{34}(z) &= \beta_{34}(1)z^{-1} + \beta_{34}(2)z^{-2} + \beta_{34}(3)z^{-3} = 1.200z^{-1} - 0.780z^{-2} + 0.120z^{-3}, \\ \beta_{35}(z) &= \beta_{35}(1)z^{-1} + \beta_{35}(2)z^{-2} + \beta_{35}(3)z^{-3} = 0.800z^{-1} - 0.880z^{-2} + 0.200z^{-3}, \\ \beta_{36}(z) &= \beta_{36}(1)z^{-1} + \beta_{36}(2)z^{-2} + \beta_{36}(3)z^{-3} = 0.600z^{-1} - 0.700z^{-2} + 0.400z^{-3}. \end{aligned}$$

系统的辨识模型为

$$w(kT) = \varphi_0^T(kT)\theta, \quad y(kT) = w(kT) + v(kT),$$

其中

$$\begin{aligned} \theta &= [\alpha_1, \alpha_2, \alpha_3, \beta_{11}(1), \beta_{11}(2), \beta_{11}(3), \beta_{12}(1), \beta_{12}(2), \beta_{12}(3), \\ &\quad \beta_{13}(1), \beta_{13}(2), \beta_{13}(3), \beta_{21}(1), \beta_{21}(2), \beta_{21}(3), \beta_{22}(1), \beta_{22}(2), \beta_{22}(3), \\ &\quad \beta_{31}(1), \beta_{31}(2), \beta_{31}(3), \beta_{32}(1), \beta_{32}(2), \beta_{32}(3), \beta_{33}(1), \beta_{33}(2), \beta_{33}(3), \\ &\quad \beta_{34}(1), \beta_{34}(2), \beta_{34}(3), \beta_{35}(1), \beta_{35}(2), \beta_{35}(3), \beta_{36}(1), \beta_{36}(2), \beta_{36}(3)]^T \\ &= [0.600, 0.400, 0.300, -0.200, 0.500, -0.400, -0.650, 0.600, 0.100, \\ &\quad 1.000, -0.500, 0.400, -0.400, 0.800, -0.200, 0.700, 0.200, 0.300, \\ &\quad 1.000, -0.900, 0.200, 1.000, -0.750, 0.125, 1.200, -1.080, 0.240, \\ &\quad 1.200, 0.780, 0.120, 0.800, -0.880, 0.200, 0.600, -0.700, 0.400]^T, \\ \varphi_0(kT) &= [-w(kT-T), -w(kT-2T), -w(kT-3T), \\ &\quad u_1(kT-T), u_1(kT-2T), u_1(kT-3T), \\ &\quad u_1(kT+T_1-T), u_1(kT+T_1-2T), u_1(kT+T_1-3T), \\ &\quad u_1(kT+2T_1-T), u_1(kT+2T_1-2T), u_1(kT+2T_1-3T), \end{aligned}$$

$$\begin{aligned}
& u_2(kT - T), u_2(kT - 2T), u_2(kT - 3T), \\
& u_2(kT + T_2 - T), u_2(kT + T_2 - 2T), u_2(kT + T_2 - 3T), \\
& u_3(kT - T), u_3(kT - 2T), u_3(kT - 3T), \\
& u_3(kT + T_3 - T), u_3(kT + T_3 - 2T), u_3(kT + T_3 - 3T), \\
& u_3(kT + 2T_3 - T), u_3(kT + 2T_3 - 2T), u_3(kT + 2T_3 - 3T), \\
& u_3(kT + 3T_3 - T), u_3(kT + 3T_3 - 2T), u_3(kT + 3T_3 - 3T), \\
& u_3(kT + 4T_3 - T), u_3(kT + 4T_3 - 2T), u_3(kT + 4T_3 - 3T), \\
& u_3(kT + 5T_3 - T), u_3(kT + 5T_3 - 2T), u_3(kT + 5T_3 - 3T)]^T.
\end{aligned}$$

仿真时, 输入 $\{u_1(kT_1)\}$, $\{u_2(kT_2)\}$ 和 $\{u_3(kT_3)\}$ 采用零均值单位方差不相关可测的持续激励信号序列. $\{v(kT)\}$ 采用零均值方差为 $\sigma^2 = 0.10^2$ 的白噪声序列. 应用 AM-RLS 算法估计这个系统的参数. 参数估计和参数估计误差如表 4.5 所示, 参数估计误差 $\delta := \|\hat{\theta}(k) - \theta\|/\|\theta\|$ 随 k 变化曲线如图 4.6 所示.

从表 4.5 和图 4.6 可以得到, 由 AM-RLS 算法估计得到的参数误差随着 k 的逐渐增加会越来越小, 且趋于零. 这个仿真例子验证了算法的有效性.

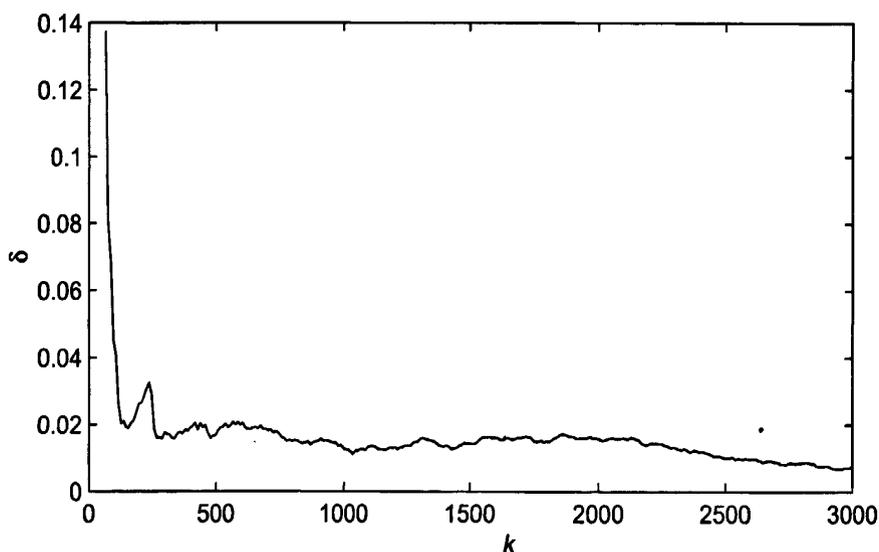


图 4.6 参数估计误差 δ 随数据长度 k 变化曲线
Fig. 4.6 The parameter estimation errors δ versus k

表 4.5 参数估计和参数估计误差
Table 4.5 The parameter estimates and their errors

k	100	500	1000	2000	3000	4000	5000
$\alpha_1 = 0.60000$	0.63877	0.59233	0.59233	0.59064	0.58818	0.59262	0.59493
$\alpha_2 = 0.40000$	0.44409	0.39913	0.39925	0.39931	0.39644	0.39789	0.39848
$\alpha_3 = 0.30000$	0.31463	0.30610	0.30385	0.30487	0.30330	0.30198	0.30179
$\beta_{11}(1) = -0.20000$	-0.21610	-0.20412	-0.19802	-0.19898	-0.19841	-0.19791	-0.19877
$\beta_{11}(2) = 0.50000$	0.49364	0.50004	0.49823	0.50297	0.50312	0.50009	0.49983
$\beta_{11}(3) = -0.40000$	-0.36752	-0.39777	-0.40079	-0.40172	-0.40262	-0.39989	-0.39985
$\beta_{12}(1) = -0.65000$	-0.65068	-0.64008	-0.64225	-0.64184	-0.64186	-0.64477	-0.64779
$\beta_{12}(2) = 0.60000$	0.53519	0.60706	0.61147	0.61093	0.61267	0.60790	0.60724
$\beta_{12}(3) = 0.10000$	0.09437	0.08324	0.08707	0.08422	0.08560	0.09156	0.09446
$\beta_{13}(1) = 1.00000$	1.00843	0.98729	0.99184	0.99335	0.99108	0.99323	0.99509
$\beta_{13}(2) = -0.50000$	-0.45575	-0.50104	-0.50474	-0.50696	-0.51159	-0.50810	-0.50530
$\beta_{13}(3) = 0.40000$	0.41215	0.41374	0.40809	0.40967	0.41117	0.40503	0.40309
$\beta_{21}(1) = -0.40000$	-0.36993	-0.39777	-0.39866	-0.39902	-0.40161	-0.40015	-0.40174
$\beta_{21}(2) = 0.80000$	0.81764	0.81688	0.81125	0.81099	0.81161	0.80724	0.80464
$\beta_{21}(3) = -0.20000$	-0.15287	-0.19673	-0.19441	-0.20072	-0.20330	-0.20158	-0.19968
$\beta_{22}(1) = 0.70000$	0.71133	0.70324	0.70637	0.70476	0.70407	0.70305	0.70230
$\beta_{22}(2) = 0.20000$	0.19498	0.19212	0.19924	0.19982	0.19719	0.19812	0.19968
$\beta_{22}(3) = 0.30000$	0.29294	0.28025	0.29131	0.29450	0.29641	0.29823	0.29758
$\beta_{31}(1) = 1.00000$	0.99672	0.99788	1.00041	0.99737	0.99508	0.99650	0.99780
$\beta_{31}(2) = -0.90000$	-0.85260	-0.90432	-0.90444	-0.90722	-0.90927	-0.90505	-0.90375
$\beta_{31}(3) = 0.20000$	0.20293	0.21384	0.21077	0.21178	0.21229	0.20770	0.20616
$\beta_{32}(1) = 1.00000$	0.99127	1.00545	1.00124	0.99885	0.99920	0.99996	1.00055
$\beta_{32}(2) = -0.75000$	-0.70381	-0.75525	-0.75784	-0.76305	-0.76497	-0.75834	-0.75482
$\beta_{32}(3) = 0.12500$	0.13945	0.14888	0.14350	0.14226	0.13890	0.13414	0.13198
$\beta_{33}(1) = 1.20000$	1.18218	1.21511	1.20524	1.20561	1.20748	1.20611	1.20480
$\beta_{33}(2) = -1.08000$	-1.04016	-1.09637	-1.09271	-1.09781	-1.10057	-1.09229	-1.08900
$\beta_{33}(3) = 0.24000$	0.20107	0.24734	0.24874	0.25458	0.25389	0.24895	0.24637
$\beta_{34}(1) = 1.20000$	1.15523	1.19779	1.19662	1.19517	1.19757	1.19846	1.19775
$\beta_{34}(2) = -0.78000$	-0.74178	-0.80712	-0.79912	-0.80143	-0.80418	-0.79638	-0.79370
$\beta_{34}(3) = 0.12000$	0.14225	0.12481	0.12320	0.12734	0.12791	0.12484	0.12267
$\beta_{35}(1) = 0.80000$	0.78083	0.79753	0.79760	0.79868	0.79929	0.79714	0.79743
$\beta_{35}(2) = -0.88000$	-0.85950	-0.89721	-0.88978	-0.89246	-0.89369	-0.89024	-0.88831
$\beta_{35}(3) = 0.20000$	0.17777	0.21088	0.21062	0.21041	0.21226	0.20882	0.20580
$\beta_{36}(1) = 0.60000$	0.55837	0.59753	0.59635	0.59680	0.60080	0.60179	0.60149
$\beta_{36}(2) = -0.70000$	-0.66682	-0.71112	-0.71246	-0.71227	-0.71322	-0.70787	-0.70643
$\beta_{36}(3) = 0.40000$	0.40310	0.41663	0.41042	0.41039	0.41212	0.40573	0.40255
δ (%)	4.51365	1.74012	1.29096	1.44891	1.57504	1.00093	0.74490

例 4.4.2 对于双输入单输出多率系统, 其输入刷新周期分别为 $T_1 = 2h$ 和 $T_2 = 3h$, 输出采样周期为 $T = 6h$, 取基周期为 $h = 2$, 相应的传递函数模型为

$$y(kT) = \frac{1}{\alpha(z)} [\beta_{11}(z)u_1(kT) + \beta_{12}(z)u_1(kT + T_1) + \beta_{13}(z)u_1(kT + 2T_1) + \beta_{21}(z)u_2(kT) + \beta_{22}(z)u_2(kT + T_2)] + v(kT),$$

其中

$$\begin{aligned} \alpha(z) &= 1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} = 1 + 0.60z^{-1} + 0.40z^{-2}, \\ \beta_{11}(z) &= \beta_{11}(1)z^{-1} + \beta_{11}(2)z^{-2} = -0.20z^{-1} + 0.50z^{-2}, \\ \beta_{12}(z) &= \beta_{12}(1)z^{-1} + \beta_{12}(2)z^{-2} = -0.65z^{-1} + 0.60z^{-2}, \\ \beta_{13}(z) &= \beta_{13}(1)z^{-1} + \beta_{13}(2)z^{-2} = 1.00z^{-1} - 0.50z^{-2}, \\ \beta_{21}(z) &= \beta_{21}(1)z^{-1} + \beta_{21}(2)z^{-2} = -0.40z^{-1} + 0.80z^{-2}, \\ \beta_{22}(z) &= \beta_{22}(1)z^{-1} + \beta_{22}(2)z^{-2} = 0.70z^{-1} + 0.20z^{-2}. \end{aligned}$$

辨识模型为

$$y_0(kT) = \varphi_0^T(kT)\theta, \quad y(kT) = y_0(kT) + v(kT),$$

参数向量和信息向量为

$$\begin{aligned} \theta &= [\alpha_1, \alpha_2, \beta_{11}(1), \beta_{11}(2), \beta_{12}(1), \beta_{12}(2), \beta_{13}(1), \beta_{13}(2), \beta_{21}(1), \beta_{21}(2), \\ &\quad \beta_{22}(1), \beta_{22}(2)]^T \\ &= [0.60, 0.40, -0.20, 0.50, -0.65, 0.60, 1.00, -0.50, -0.40, 0.80, 0.70, 0.20]^T, \\ \varphi_0(kT) &= [\phi_0^T(kT), \phi_1^T(kT), \phi_2^T(kT)]^T, \\ \phi_0(kT) &= [-y_0(kT - T), -y_0(kT - 2T)]^T, \\ \phi_1(kT) &= [u_1(kT - T), u_1(kT - 2T), u_1(kT + T_1 - T), u_1(kT + T_1 - 2T), \\ &\quad u_1(kT + 2T_1 - T), u_1(kT + 2T_1 - 2T)]^T, \\ \phi_2(kT) &= [u_2(kT - T), u_2(kT - 2T), u_2(kT + T_2 - T), u_2(kT + T_2 - 2T)]^T. \end{aligned}$$

仿真时, $\{u_1(kT_1)\}$ 和 $\{u_2(kT_2)\}$ 采用零均值单位方差不相关可测的持续激励信号序列。 $\{v(kT)\}$ 采用零均值的白噪声序列, 方差分别为 $\sigma^2 = 0.50^2$ 和 $\sigma^2 = 1.00^2$. 应用 AM-RLS 算法估计这个系统的参数, 在不同噪声方差下的参数估计和参数估计误差如表 4.6-4.7 所示, 参数估计误差 $\delta := \|\hat{\theta}(kT) - \theta\|/\|\theta\|$ 随 k 变化的曲线如图 4.7 所示.

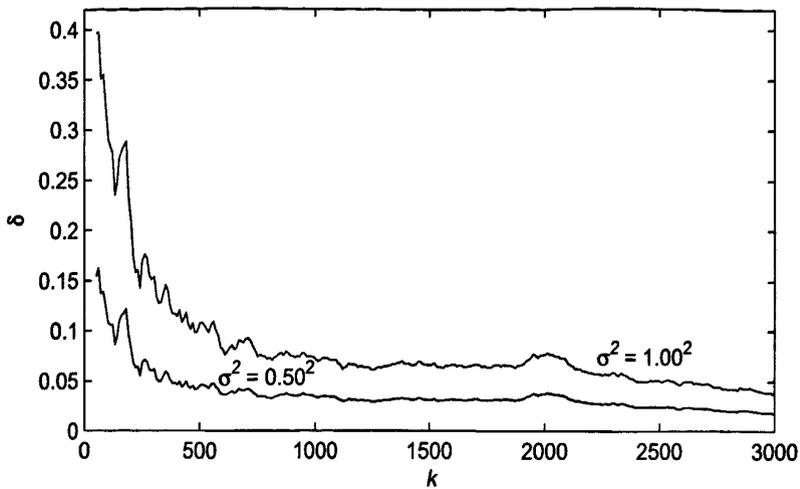
从表 4.6-4.7 和图 4.7 可以知道, 随着数据长度 k 的增加, AM-RLS 算法的参数估计误差越来越小, 且趋于零. 噪声方差越小, 算法的收敛速度越快且估计精度越高.

表 4.6 参数估计和参数估计误差 ($\sigma^2 = 0.50^2$)
 Table 4.6 The parameter estimates and their errors ($\sigma^2 = 0.50^2$)

k	100	500	1000	1500	2000	2500	3000
$\alpha_1 = 0.60000$	0.52804	0.57601	0.58832	0.58984	0.57900	0.58664	0.58879
$\alpha_2 = 0.40000$	0.35983	0.39506	0.39766	0.39520	0.38569	0.39175	0.39327
$\beta_{11}(1) = -0.20000$	-0.25131	-0.21684	-0.19302	-0.19113	-0.19355	-0.19025	-0.19490
$\beta_{11}(2) = 0.50000$	0.48526	0.48678	0.48255	0.50086	0.50496	0.49541	0.49586
$\beta_{12}(1) = -0.65000$	-0.62235	-0.64773	-0.64456	-0.63707	-0.64500	-0.63895	-0.64251
$\beta_{12}(2) = 0.60000$	0.64043	0.63942	0.62308	0.61331	0.61835	0.60827	0.59895
$\beta_{13}(1) = 1.00000$	0.92740	0.98532	1.01948	1.01745	1.01140	1.00837	1.01296
$\beta_{13}(2) = -0.50000$	-0.55227	-0.53622	-0.52411	-0.53557	-0.54402	-0.52669	-0.51761
$\beta_{21}(1) = -0.40000$	-0.30988	-0.39445	-0.39279	-0.39175	-0.40431	-0.40018	-0.41020
$\beta_{21}(2) = 0.80000$	0.89136	0.85732	0.84014	0.83448	0.83905	0.82473	0.81392
$\beta_{22}(1) = 0.70000$	0.66411	0.70456	0.73195	0.71914	0.70967	0.70779	0.70904
$\beta_{22}(2) = 0.20000$	0.09769	0.16731	0.19006	0.17844	0.16822	0.17778	0.18324
δ (%)	10.75583	4.52681	3.39213	3.17580	3.72441	2.44856	1.83827

表 4.7 参数估计和参数估计误差 ($\sigma^2 = 1.00^2$)
 Table 4.7 The parameter estimates and their errors ($\sigma^2 = 1.00^2$)

k	100	500	1000	1500	2000	2500	3000
$\alpha_1 = 0.60000$	0.37787	0.52894	0.56917	0.57603	0.55680	0.57175	0.57641
$\alpha_2 = 0.40000$	0.27866	0.36826	0.38476	0.38384	0.36791	0.38040	0.38414
$\beta_{11}(1) = -0.20000$	-0.30577	-0.23011	-0.18498	-0.18154	-0.18634	-0.18005	-0.18937
$\beta_{11}(2) = 0.50000$	0.48994	0.48268	0.46819	0.50359	0.51092	0.49179	0.49249
$\beta_{12}(1) = -0.65000$	-0.63444	-0.65679	-0.64047	-0.62443	-0.63913	-0.62758	-0.63471
$\beta_{12}(2) = 0.60000$	0.73697	0.68328	0.64432	0.62453	0.63464	0.61524	0.59689
$\beta_{13}(1) = 1.00000$	0.82331	0.97237	1.04130	1.03712	1.02376	1.01741	1.02641
$\beta_{13}(2) = -0.50000$	-0.72963	-0.60602	-0.56143	-0.57888	-0.59166	-0.55691	-0.53798
$\beta_{21}(1) = -0.40000$	-0.21699	-0.38896	-0.38413	-0.38277	-0.40819	-0.39998	-0.42006
$\beta_{21}(2) = 0.80000$	1.02163	0.91989	0.88166	0.86971	0.87798	0.84976	0.82822
$\beta_{22}(1) = 0.70000$	0.55171	0.69541	0.76048	0.73721	0.71835	0.71474	0.71742
$\beta_{22}(2) = 0.20000$	-0.05965	0.11361	0.17170	0.15174	0.13378	0.15311	0.16468
δ (%)	28.97605	10.75110	7.15040	6.73308	7.64008	5.09158	3.82674

图 4.7 参数估计误差 δ 随数据长度 k 变化曲线Fig. 4.7 The parameter estimation errors δ versus k

本节内容已发表在国际期刊《Mathematical and Computer Modelling》2009年第50卷第7-8期上，参加了于2009年6月在美国召开的28th American Control Conference国际顶级会议上。

4.5 本章小结

本章研究了输出误差多率系统的辨识问题。利用离散化技术，推导了多率系统的离散状态空间模型，得到辨识模型。基于辅助模型辨识思想和多新息辨识理论，提出了辅助模型多新息随机梯度算法估计多率系统的参数，提出的算法充分利用可测的多率输入输出数据，提高了辅助模型随机梯度算法的收敛速度和参数估计精度。同时提出了辅助模型最小二乘算法估计系统的参数，应用鞅收敛定理，从理论上证明了算法的收敛性。

第五章 多率系统子模型辨识方法

5.1 引言

多率多输入系统在实际中有广泛的应用,如无线通信领域中,对于多用户 CDMA(Code division multiple access)系统,每个用户被指定用不同的码序列在不同的信道发送控制信号,每个控制信号通过的转换器或滤波器下得到输出信号,输出信号可在同一个信道^[98,99],其数学模型为多率多输入系统.

目前,针对多输入单输出误差模型,输出误差滑动平均模型和输出误差 Box-Jenkins 模型,文 [87], [100], [101] 中,分别提出了辅助模型多新息最小二乘算法,辅助模型递推增广最小二乘算法和辅助模型多新息广义增广随机梯度算法得到了系统的参数估计. Ding 等利用辅助模型辨识思想,研究了 Hammerstein 输出误差系统的辨识^[97],将其推广至研究非线性系统的辨识. 对于双率系统, Ding 和 Chen 提出辅助模型最小二乘算法和辅助模型随机梯度算法估计系统参数和损失输出,分析了提出算法的收敛性^[22,23]. Zhu 等应用输出误差辨识方法研究了快速率慢输出非均匀采样系统辨识问题^[64]. Mo 等结合输出误差辨识方法和 Newton-Raphson 算法估计快输入慢输出双率系统的参数^[10]. Liu 等针对非均匀采样系统的辨识,提出辅助模型最小二乘算法估计系统的参数^[102]. 对于多率多输入系统,利用辅助模型辨识方法, Shi 等解决了 FIR 模型的辨识问题^[3], Han 和 Ding 研究了 ARX 模型的辨识,分析了提出算法的收敛性^[76].

本章研究带有子模型多率系统的辨识问题. 组织如下: 第 5.2 节利用离散化技术,推导多率系统的离散时间状态空间模型. 第 5.3 节基于辅助模型辨识思想,提出辅助子模型最小二乘算法估计系统参数,用数值仿真例子验证提出算法的有效性. 第 5.4 节为本章总结.

5.2 模型描述

对于多率多输入系统图 5.1 所示, P_{c_j} 是一个连续时间过程,有如下的状态空间模型

$$\begin{cases} \dot{\mathbf{x}}_j(t) = \mathbf{A}_{c_j}\mathbf{x}_j(t) + \mathbf{B}_{c_j}u_j(t), & j = 1, 2, \dots, r, \\ y_j(t) = \mathbf{C}_j\mathbf{x}_j(t) + D_ju_j(t), \end{cases} \quad (5.2.1)$$

其中 $\mathbf{x}_j(t) \in \mathbb{R}^{n_j}$ 为状态向量, $u_j(t) \in \mathbb{R}^1$ 为第 j 个通道的控制输入, $y_j(t) \in \mathbb{R}^1$ 为第 j 个通道的输出, $y_0(t) \in \mathbb{R}^1$ 为未知的无噪输出, $y(t) \in \mathbb{R}^1$ 为系统输出, $v(t) \in \mathbb{R}^1$ 为外部干扰,是白噪声. \mathbf{A}_{c_j} , \mathbf{B}_{c_j} , \mathbf{C}_j 和 D_j 是具有适当维数的的矩阵. H_{T_j} 为周期 $T_j := p_j h$ 的零阶保持器, S_T 为周期 $T := qh$ 的采样器. 其中, 令 p_1, p_2, \dots, p_r 是正整数且互质, q 是 (p_1, p_2, \dots, p_r) 的最小公倍数^[68], h 是基周期, T 是框架周期^[31]. 对于图 5.1 所示的多率系统, 可测的多率输入输出数据为 $\{u_1(kT_1), u_2(kT_2), \dots, u_r(kT_r), y(kT): k = 0, 1, 2, \dots\}$.

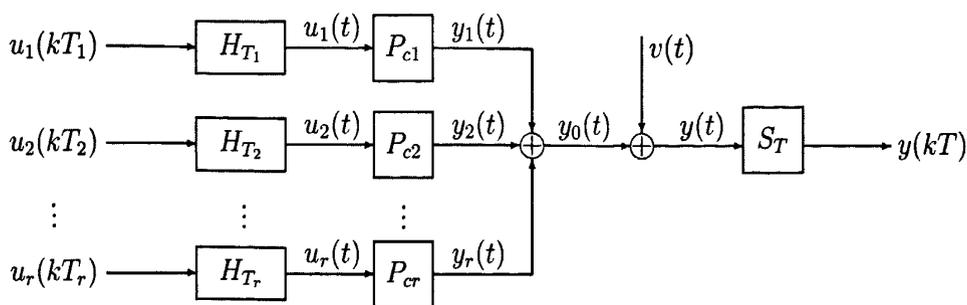


图 5.1 多率多输入系统

Fig. 5.1 The multirate multiple-input system

仿照第 3.2 节和第 4.2 节, 应用离散化技术, 图 5.1 所示的多率系统的状态空间模型为

$$\begin{cases} \mathbf{x}_j(kT + T) = \mathbf{A}_j \mathbf{x}_j(kT) + \sum_{i=1}^{\nu_j} \mathbf{B}_{ji} u_j(kT - (i-1)T_j), & j = 1, 2, \dots, r \\ y_j(kT) = \mathbf{C}_j \mathbf{x}_j(kT) + D_j u_j(kT), \\ y(kT) = \sum_{j=1}^r y_j(kT) + v(kT), \end{cases} \quad (5.2.2)$$

其中 $\nu_j := q/p_j$, $\mathbf{x}_j(kT) := \mathbf{x}_j(t)|_{t=kT}$, $u_j(kT + iT_j) := u_j(t)|_{t=kT+iT_j}$, $y(kT) := y(t)|_{t=kT}$, 和

$$\begin{aligned} \mathbf{A}_h &:= e^{\mathbf{A}_{cj}h} \in \mathbb{R}^{n_j \times n_j}, & \mathbf{A}_j &:= e^{\mathbf{A}_{cj}qh}, \\ \mathbf{B}_{pj} &:= \int_0^{T_j} e^{\mathbf{A}_{cj}t} dt \mathbf{B}_{cj} \in \mathbb{R}^{n_j}, \\ \mathbf{B}_{ji} &:= \mathbf{A}_h^{q-ip_j} \mathbf{B}_{pj}, & j &= 1, 2, \dots, r, \quad i = 1, 2, \dots, \nu_j. \end{aligned}$$

5.3 辅助模型最小二乘算法

5.3.1 算法的提出

z^{-1} 为后移算子: $z^{-1}y(kT) = y(kT - T)$. 从状态空间模型 (5.2.2), 输出方程可以表达为如下:

$$\begin{aligned} y(kT) &= \sum_{i=1}^r y_j(kT) + v(kT), \\ y_j(kT) &= \sum_{i=1}^{\nu_j} [\mathbf{C}_j (z\mathbf{I} - \mathbf{A}_j)^{-1} \mathbf{B}_{ji} u_j(kT - (i-1)T_j) + D_j u_j(kT)] \\ &= \sum_{i=1}^{\nu_j} \left[\frac{\mathbf{C}_j \text{adj}[z\mathbf{I} - \mathbf{A}_j] \mathbf{B}_{ji}}{\det[z\mathbf{I} - \mathbf{A}_j]} u_j(kT - (i-1)T_j) + D_j u_j(kT) \right] \\ &= \sum_{i=1}^{\nu_j} \left[\frac{z^{-n_j} \mathbf{C}_j \text{adj}[z\mathbf{I} - \mathbf{A}_j] \mathbf{B}_{ji}}{z^{-n_j} \det[z\mathbf{I} - \mathbf{A}_j]} u_j(kT - (i-1)T_j) + D_j u_j(kT) \right] \end{aligned}$$

$$=: \sum_{i=1}^{\nu_j} \frac{\beta_{ji}(z)}{\alpha_j(z)} u_j(kT - (i-1)T_j), \quad j = 1, 2, \dots, r, \quad (5.3.1)$$

其中 $\alpha_j(z)$ 和 $\beta_{ji}(z)$ 是 z^{-1} 的多项式,

$$\begin{aligned} \alpha_j(z) &:= z^{-n_j} \det[z\mathbf{I} - \mathbf{A}_j] = 1 + \alpha_j(1)z^{-1} + \alpha_j(2)z^{-2} + \dots + \alpha_j(n_j)z^{-n_j}, \\ \beta_{j1}(z) &:= z^{-n_j} \mathbf{C}_j \operatorname{adj}[z\mathbf{I} - \mathbf{A}_j] \mathbf{B}_{j1} + z^{-n_j} \det[z\mathbf{I} - \mathbf{A}_j] D_j \\ &= \beta_{j1}(0) + \beta_{j1}(1)z^{-1} + \beta_{j1}(2)z^{-2} + \dots + \beta_{j1}(n_j)z^{-n_j}, \\ \beta_{ji}(z) &:= z^{-n} \mathbf{C}_j \operatorname{adj}[z\mathbf{I} - \mathbf{A}_j] \mathbf{B}_{ji} \\ &= \beta_{ji}(1)z^{-1} + \beta_{ji}(2)z^{-2} + \dots + \beta_{ji}(n_j)z^{-n_j}, \quad i = 2, 3, \dots, \nu_j. \end{aligned}$$

定义参数向量 $\boldsymbol{\theta}_j$ 和信息向量 $\boldsymbol{\varphi}_j(kT)$ 如下:

$$\begin{aligned} \boldsymbol{\theta}_j &:= [\alpha_j(1), \alpha_j(2), \dots, \alpha_j(n_j), \beta_{j1}(0), \beta_{j1}(1), \beta_{j1}(2), \dots, \beta_{j1}(n_j), \beta_{j2}(1), \\ &\quad \beta_{j2}(2), \dots, \beta_{j2}(n_j), \dots, \beta_{j\nu_j}(1), \beta_{j\nu_j}(2), \dots, \beta_{j\nu_j}(n_j)]^T \in \mathbb{R}^{n_j + \nu_j n_j + 1}, \\ \boldsymbol{\varphi}_j(kT) &:= [\boldsymbol{\zeta}_j^T(kT), \boldsymbol{\phi}_j^T(kT)]^T \in \mathbb{R}^{n_j + \nu_j n_j + 1} \\ \boldsymbol{\zeta}_j(kT) &:= [-y_j(kT - T), -y_j(kT - 2T), \dots, -y_j(kT - n_j T)] \in \mathbb{R}^{n_j}, \\ \boldsymbol{\phi}_j(kT) &:= [u_j(kT), u_j(kT - T), u_j(kT - 2T), \dots, u_j(kT - n_j T), \\ &\quad u_j(kT + T_j - T), u_j(kT + T_j - 2T), \dots, u_j(kT + T_j - n_j T), \dots, \\ &\quad u_j(kT + (\nu_j - 1)T_j - T), u_j(kT + (\nu_j - 1)T_j - 2T), \dots, \\ &\quad u_j(kT + (\nu_j - 1)T_j - n_j T)]^T \in \mathbb{R}^{\nu_j n_j + 1}, \quad j = 1, 2, \dots, r, \end{aligned}$$

则 (5.3.1) 式可以写为

$$\begin{aligned} y_j(kT) &= \boldsymbol{\varphi}_j^T(kT) \boldsymbol{\theta}_j, \quad j = 1, 2, \dots, r, \\ \mathbf{y}(kT) &= \sum_{j=1}^r \boldsymbol{\varphi}_j^T(kT) \boldsymbol{\theta}_j + \mathbf{v}(kT) = \boldsymbol{\varphi}^T(kT) \boldsymbol{\theta} + \mathbf{v}(kT), \end{aligned} \quad (5.3.2)$$

其中

$$\boldsymbol{\theta} := \begin{bmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \\ \vdots \\ \boldsymbol{\theta}_r \end{bmatrix} \in \mathbb{R}^{n_0}, \quad \boldsymbol{\varphi}(kT) := \begin{bmatrix} \boldsymbol{\varphi}_1(kT) \\ \boldsymbol{\varphi}_2(kT) \\ \vdots \\ \boldsymbol{\varphi}_r(kT) \end{bmatrix} \in \mathbb{R}^{n_0}, \quad n_0 := \sum_{j=1}^r (n_j + \nu_j n_j) + r,$$

其中参数向量 $\boldsymbol{\theta}$ 为待辨识参数.

注意到信息向量 $\boldsymbol{\varphi}(kT)$ 含有未知的中间变量 $y_j(kT - iT)$ ($i = 1, 2, \dots, n_j$), 因此常规的 RLS 算法不能用于辨识多率系统 (5.3.2) 得到参数估计.

基于辅助模型辨识思想，建立辅助子模型如图 5.2 所示，

$$\hat{y}_j(kT) = \frac{\hat{\beta}_j(z)}{\hat{\alpha}_j(z)} u_j(kT) \quad \text{或} \quad \hat{y}_j(kT) = \hat{\varphi}_j^T(kT) \hat{\theta}_j(kT),$$

其中

$$\beta_j(z) := [\beta_{j1}(z), \beta_{j2}(z), \dots, \beta_{j\nu_j}(z)],$$

$$u_j(kT) := [u_j(kT), u_j(kT + T_j), \dots, u_j(kT + (\nu_j - 1)T_j)]^T,$$

$$\hat{\varphi}(kT) := [\hat{\zeta}_j^T(kT), \hat{\phi}_j^T(kT)]^T,$$

$$\hat{\zeta}_j(kT) := [-\hat{y}_j(kT - T), -\hat{y}_j(kT - 2T) \dots, -\hat{y}_j(kT - n_jT)]^T,$$

其中 $\hat{y}_j(kT)$, $\hat{\zeta}_j(kT)$ 和 $\hat{\theta}_j(kT)$ 分别是 $y_j(kT)$, $\zeta_j(kT)$ 和 θ_j 在 kT 时刻的估计值。信息向量 $\varphi(kT)$ 中不可测的变量 $y_j(kT - iT)$ 用辅助模型的输出 $\hat{y}_j(kT - iT)$ 代替。

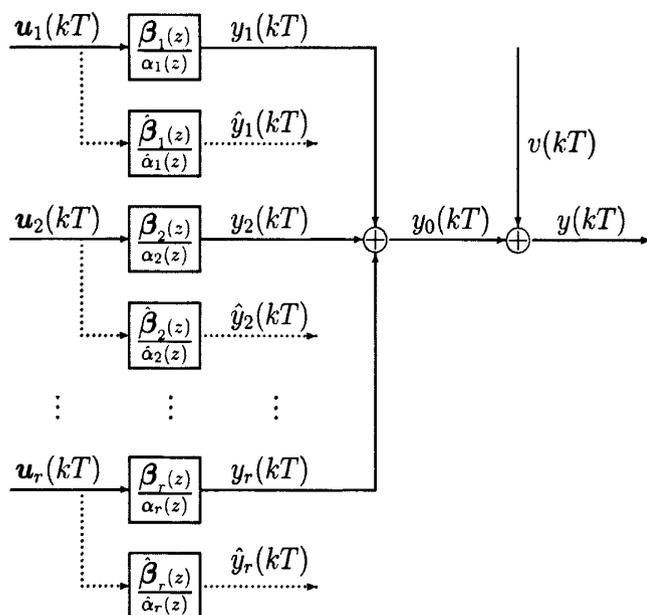


图 5.2 带辅助子模型的输出误差系统

Fig. 5.2 The output error system with auxiliary sub-models

参照文 [22]，信息向量 $\varphi(kT)$ 中不可测的变量 $y_j(kT - iT)$ 用其估计值 $\hat{y}_j(kT - iT)$ 代替， $\varphi_j(kT)$ 用其估计值 $\hat{\varphi}_j(kT)$ 来代替，进一步地， $\varphi(kT)$ 用其估计值 $\hat{\varphi}(kT)$ 来代替。推导得到估计多率系统 (5.3.2) 的参数 θ 的基于辅助子模型最小二乘 (ASM-RLS) 算法，如下：

$$\hat{\theta}(kT) = \hat{\theta}(kT - T) + L(kT)[y(kT) - \hat{\varphi}^T(kT)\hat{\theta}(kT - T)], \quad (5.3.3)$$

$$L(kT) = P(kT - T)\hat{\varphi}(kT)[1 + \hat{\varphi}^T(kT)P(kT - T)\hat{\varphi}(kT)]^{-1}, \quad (5.3.4)$$

$$P(kT) = [I - L(kT)\hat{\varphi}^T(kT)]P(kT - T), \quad (5.3.5)$$

$$\hat{\varphi}(kT) = \begin{bmatrix} \hat{\varphi}_1(kT) \\ \hat{\varphi}_2(kT) \\ \vdots \\ \hat{\varphi}_r(kT) \end{bmatrix}, \quad \hat{\theta}(kT) = \begin{bmatrix} \hat{\theta}_1(kT) \\ \hat{\theta}_2(kT) \\ \vdots \\ \hat{\theta}_r(kT) \end{bmatrix}, \quad (5.3.6)$$

$$\hat{\varphi}_j(kT) = [\hat{\zeta}_j^T(kT), \phi_j^T(kT)]^T, \quad (5.3.7)$$

$$\hat{\zeta}_j(kT) = [-\hat{y}_j(kT - T), -\hat{y}_j(kT - 2T), \dots, -\hat{y}_j(kT - n_jT)]^T, \quad (5.3.8)$$

$$\begin{aligned} \phi_j(kT) = & [u_j(kT), u_j(kT - T), u_j(kT - 2T), \dots, u_j(kT - n_jT), \\ & u_j(kT + T_j - T), u_j(kT + T_j - 2T), \dots, u_j(kT + T_j - n_jT), \dots, \\ & u_j(kT + (\nu_j - 1)T_j - T), u_j(kT + (\nu_j - 1)T_j - 2T), \dots, \\ & u_j(kT + (\nu_j - 1)T_j - n_jT)]^T, \end{aligned} \quad (5.3.9)$$

$$\hat{y}_j(kT) = \hat{\varphi}_j^T(kT)\hat{\theta}_j(kT), \quad j = 1, 2, \dots, r, \quad (5.3.10)$$

其中 $P(kT)$ 是参数估计误差的协方差阵。

设 $u_j(kT_j) = 0$ 和 $y_j(kT) = 0$ ($j = 1, 2, \dots, r$), 当 $k \leq 0$ 时. 初始化 ASM-RLS 算法, 一般地, $P(0) = p_0 \mathbf{1}_{n_0 \times n_0}$, p_0 为很大的正实数, 如 $p_0 = 10^6$, $\hat{\theta}(0)$ 为很小的正实向量, 如 $\hat{\theta}(0) = \mathbf{1}_{n_0}/p_0$. 用 ASM-RLS 算法计算参数估计 $\hat{\theta}(kT)$ 的步骤如下:

1. 令 $k = 1$, $\hat{\theta}(0) = \mathbf{1}_{n_0}/p_0$, $P(0) = p_0 \mathbf{1}_{n_0 \times n_0}$, 选取数据长度 L_e .
2. 搜集多率输入输出数据 $\{u_1(kT_1), u_2(kT_2), \dots, u_r(kT_r), y(kT)\}$.
3. 由 (5.3.8) 式构成信息向量 $\hat{\zeta}_j(kT)$, 由 (5.3.9) 式构成 $\phi_j(kT)$.
4. 由 (5.3.7) 式构成信息向量 $\hat{\varphi}_j(kT)$, 进一步地, 由 (5.3.6) 式构成 $\hat{\varphi}(kT)$.
5. 由 (5.3.4) 式计算 $P(kT)$, 由 (5.3.3) 式刷新 $\hat{\theta}(kT)$, 并由 (5.3.10) 式计算 $\hat{y}_j(kT)$.
6. 如果 $k < L_e$, k 增至 $k + 1$, 返回第 2 步继续计算. 否则, 得到参数估计 $\hat{\theta}(L_e T)$.

计算参数估计 $\hat{\theta}(kT)$ 的流程图如图 5.3 所示.

为了比较, 利用另外一种算法得到多率系统 (5.3.2) 的参数 θ 的估计, 仍然是基于辅助模型辨识思想.

定义 $\alpha^*(z) := \alpha_1(z)\alpha_2(z)\cdots\alpha_r(z)$ 和 $\beta_{ji}^*(z) := \frac{\alpha_j^*(z)}{\alpha_j(z)}\beta_{ji}(z)$. 从 (5.3.1) 式, 得到

$$y(kT) = \frac{1}{\alpha^*(z)} \sum_{j=1}^r \sum_{i=1}^{\nu_j} \beta_{ji}^*(z) u_j(kT - (i-1)T_j) + v(kT), \quad (5.3.11)$$

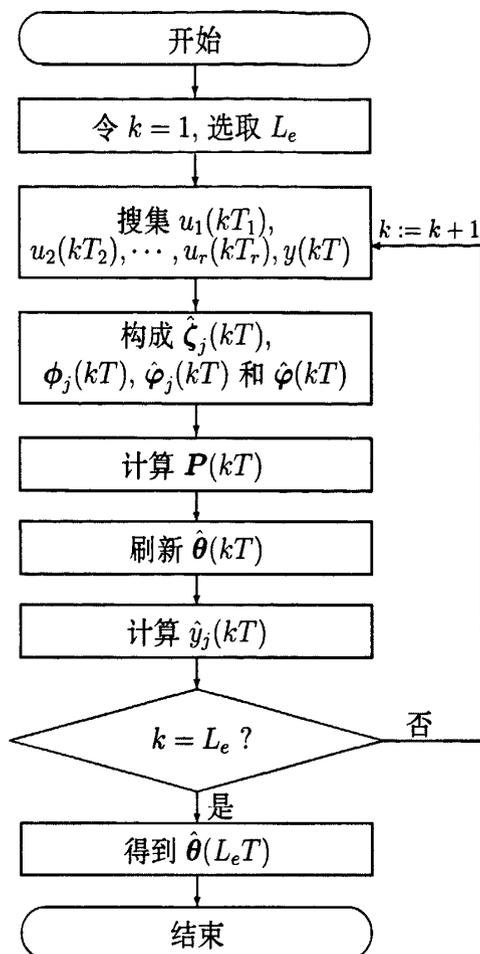


图 5.3 计算参数估计 $\hat{\theta}(kT)$ 的流程图

Fig. 5.3 The flowchart of computing the parameter estimate $\hat{\theta}(kT)$

其中

$$\alpha^*(z) := 1 + \alpha^*(1)z^{-1} + \alpha^*(2)z^{-2} + \cdots + \alpha^*(n^*)z^{-n^*}, \quad n^* = n_1 + n_2 + \cdots + n_r,$$

$$\beta_{j1}^*(z) := \beta_{j1}^*(0) + \beta_{j1}^*(1)z^{-1} + \beta_{j1}^*(2)z^{-2} + \cdots + \beta_{j1}^*(n^*)z^{-n^*},$$

$$\beta_{ji}^*(z) := \beta_{ji}^*(1)z^{-1} + \beta_{ji}^*(2)z^{-2} + \cdots + \beta_{ji}^*(n^*)z^{-n^*}, \quad j = 1, 2, \cdots, r, \quad i = 2, 3, \cdots, \nu_j.$$

根据图 5.1 所示, 未知的无噪输出可以表达为如下:

$$y_0(kT) := \frac{1}{\alpha^*(z)} \sum_{j=1}^r \sum_{i=1}^{\nu_j} \beta_{ji}^*(z) u_j(kT - (i-1)T_j). \quad (5.3.12)$$

定义参数向量 ϑ 和信息向量 $\psi(kT)$ 如下:

$$\begin{aligned} \vartheta := & [\alpha^*(1), \alpha^*(2), \cdots, \alpha^*(n^*), \beta_{11}^*(0), \beta_{11}^*(1), \beta_{11}^*(2), \cdots, \beta_{11}^*(n), \beta_{12}^*(1), \\ & \beta_{12}^*(2), \cdots, \beta_{12}^*(n), \cdots, \beta_{1\nu_1}^*(1), \beta_{1\nu_1}^*(2), \cdots, \beta_{1\nu_1}^*(n), \cdots, \beta_{r1}^*(0), \end{aligned}$$

$$\begin{aligned}
& \beta_{r_1}^*(1), \beta_{r_1}^*(2), \dots, \beta_{r_1}^*(n), \beta_{r_2}^*(1), \beta_{r_2}^*(2), \dots, \beta_{r_2}^*(n), \dots, \beta_{r_{\nu_r}}^*(1), \\
& \beta_{r_{\nu_r}}^*(2), \dots, \beta_{r_{\nu_r}}^*(n)]^T \in \mathbb{R}^m, \quad m := n^* + r + n^*(\nu_1 + \nu_2 + \dots + \nu_r), \\
\psi(kT) & := [\psi_0^T(kT), \psi_1^T(kT), \dots, \psi_r^T(kT)]^T \in \mathbb{R}^m, \\
\psi_0(kT) & := [-y_0(kT - T), -y_0(kT - 2T), \dots, -y_0(kT - n^*T)]^T \in \mathbb{R}^{n^*}, \\
\psi_j(kT) & := [u_j(kT), u_j(kT - T), u_j(kT - 2T), \dots, u_j(kT - n^*T), \\
& \quad u_j(kT + T_j - T), u_j(kT + T_j - 2T), \dots, u_j(kT + T_j - n^*T), \dots, \\
& \quad u_j(kT + (\nu_j - 1)T_j - T), u_j(kT + (\nu_j - 1)T_j - 2T), \dots, \\
& \quad u_j(kT + (\nu_j - 1)T_j - n^*T)]^T \in \mathbb{R}^{n^*\nu_j + 1}, \quad j = 1, 2, \dots, r.
\end{aligned}$$

则 (5.3.11) 式可以写为

$$y(kT) = \psi^T(kT)\theta + v(kT). \quad (5.3.13)$$

根据辅助模型辨识思想, 建立如下辅助模型,

$$\hat{y}_0(kT) = \hat{\psi}^T(kT)\hat{\vartheta}(kT),$$

其中 $\hat{y}_0(kT)$ 和 $\hat{\vartheta}(kT)$ 分别是 $y_0(kT)$ 和 ϑ 在时刻 kT 时的估计值. 信息向量 $\psi(kT)$ 中的未知变量 $y_0(kT - iT)$ 用辅助模型输出 $\hat{y}_0(kT - iT)$ 代替.

参照文 [22], 信息向量中 $\psi(kT)$ 中的未知变量 $y_0(kT - iT)$ 用其估计值 $\hat{y}_0(kT - iT)$ 代替, $\psi_0(kT)$ 用其估计值 $\hat{\psi}_0(kT)$ 代替, 进一步, $\psi(kT)$ 用其估计值 $\hat{\psi}(kT)$ 代替, 可以得到估计多率系统 (5.3.2) 的参数 θ 的辅助模型最小二乘 (AM-RLS) 算法, 如下:

$$\hat{\vartheta}(kT) = \hat{\vartheta}(kT - T) + \mathbf{L}(kT)[y(kT) - \hat{\psi}^T(kT)\hat{\vartheta}(kT - T)], \quad (5.3.14)$$

$$\mathbf{L}(kT) = \mathbf{P}(kT - T)\hat{\psi}(kT)[1 + \hat{\psi}^T(kT)\mathbf{P}(kT - T)\hat{\psi}(kT)]^{-1}, \quad (5.3.15)$$

$$\mathbf{P}(kT) = [\mathbf{I} - \mathbf{L}(kT)\hat{\psi}^T(kT)]\mathbf{P}(kT - T), \quad (5.3.16)$$

$$\hat{\psi}(kT) = [\hat{\psi}_0^T(kT), \psi_1^T(kT), \psi_2^T(kT), \dots, \psi_r^T(kT)]^T, \quad (5.3.17)$$

$$\hat{\psi}_0(kT) = [-\hat{y}_0(kT - T), -\hat{y}_0(kT - 2T), \dots, -\hat{y}_0(kT - n^*T)]^T, \quad (5.3.18)$$

$$\begin{aligned}
\psi_j(kT) & = [u_j(kT), u_j(kT - T), u_j(kT - 2T), \dots, u_j(kT - n^*T), \\
& \quad u_j(kT + T_j - T), u_j(kT + T_j - 2T), \dots, u_j(kT + T_j - n^*T), \dots, \\
& \quad u_j(kT + (\nu_j - 1)T_j - T), u_j(kT + (\nu_j - 1)T_j - 2T), \dots, \\
& \quad u_j(kT + (\nu_j - 1)T_j - n^*T)]^T, \quad (5.3.19)
\end{aligned}$$

$$\hat{y}_0(kT) = \hat{\psi}^T(kT)\hat{\vartheta}(kT). \quad (5.3.20)$$

与 AM-RLS 算法进行比较, ASM-RLS 算法不需要估计冗余的参数, 因此有较小的计算量. 例如, 对于一个双输入单输出多率系统, 其两个输入数刷新周期分别为 $T_1 = 2h$ 和 $T_2 = 3h$,

采样周期 $T = 6h$, $n_1 = n_2 = 2$, 则由 ASM-RLS 算法估计的参数数目为 16, 由 AM-RLS 算法估计的参数数目为 28. 显然, AM-RLS 算法计算量要小于 ASM-RLS 算法.

5.3.2 仿真试验

例 5.3.1 对于双输入单输出多率系统, 其输入刷新周期分别为 $T_1 = 3h$ 和 $T_2 = 2h$, 输出采样周期为 $T = 6h$, 取基周期 $h = 1$, 相应的传递函数模型为

$$\begin{aligned} y(kT) &= y_1(kT) + y_2(kT) + v(kT) \\ &= \frac{1}{\alpha_1(z)} [\beta_{11}(z)u_1(kT) + \beta_{12}(z)u_1(kT + T_1)] + \frac{1}{\alpha_2(z)} [\beta_{21}(z)u_2(kT) \\ &\quad + \beta_{22}(z)u_2(kT + T_2) + \beta_{23}(z)u_2(kT + 2T_2)] + v(kT), \end{aligned}$$

其中

$$\begin{aligned} \alpha_1(z) &= 1 + \alpha_1(1)z^{-1} = 1 + 0.30z^{-1}, \\ \alpha_2(z) &= 1 + \alpha_2(1)z^{-1} = 1 + 0.01z^{-1}, \\ \beta_{11}(z) &= \beta_{11}(1)z^{-1} + \beta_{11}(2)z^{-2} = -0.40z^{-1} + 1.50z^{-2}, \\ \beta_{12}(z) &= \beta_{12}(1)z^{-1} + \beta_{12}(2)z^{-2} = -0.80z^{-1} + 0.60z^{-2}, \\ \beta_{21}(z) &= \beta_{21}(1)z^{-1} + \beta_{21}(2)z^{-2} = 2.00z^{-1} + 0.20z^{-2}, \\ \beta_{22}(z) &= \beta_{22}(1)z^{-1} + \beta_{22}(2)z^{-2} = -0.40z^{-1} + 0.80z^{-2}, \\ \beta_{23}(z) &= \beta_{23}(1)z^{-1} + \beta_{23}(2)z^{-2} = 1.70z^{-1} + 2.80z^{-2}, \end{aligned}$$

参数向量和信息向量为

$$\begin{aligned} \theta_1 &= [\alpha_1(1), \beta_{11}(1), \beta_{11}(2), \beta_{12}(1), \beta_{12}(2)]^T, \\ \theta_2 &= [\alpha_2(1), \beta_{21}(1), \beta_{21}(2), \beta_{22}(1), \beta_{22}(2), \beta_{23}(1), \beta_{23}(2)]^T, \\ \theta &= [\theta_1^T, \theta_2^T]^T \\ &= [0.30, -0.40, 1.50, -0.80, 0.60, 0.01, 2.00, 0.20, -0.40, 0.80, 1.70, 2.80]^T, \\ \varphi(kT) &= [\varphi_1^T(kT), \varphi_2^T(kT)]^T, \\ \varphi_1(kT) &= [-y_1(kT - T), \phi_1^T(kT)]^T, \\ \varphi_2(kT) &= [-y_2(kT - T), \phi_2^T(kT)]^T, \\ \phi_1(kT) &= [u_1(kT - T), u_1(kT - 2T), u_1(kT + T_1 - T), u_1(kT + T_1 - 2T)]^T, \\ \phi_2(kT) &= [u_2(kT - T), u_2(kT - 2T), u_2(kT + T_2 - T), u_2(kT + T_2 - 2T), \\ &\quad u_2(kT + 2T_2 - T), u_2(kT + 2T_2 - 2T)]^T. \end{aligned}$$

仿真时, $\{u_1(kT_1)\}$ 和 $\{u_2(kT_2)\}$ 采用零均值单位方差不相关可测的持续激励信号序列. $\{v(kT)\}$ 采用零均值方差分别为 $\sigma^2 = 1.00^2$ 和 $\sigma^2 = 0.20^2$ 的白噪声序列. 应用 ASM-RLS 算法估计这个系统的参数. 参数估计和参数估计误差如表 5.1-5.2 所示, 参数估计误差 $\delta := \|\hat{\theta}(kT) - \theta\|/\|\theta\|$ 随 k 变化曲线如图 5.4 所示.

表 5.1 参数估计和参数估计误差 ($\sigma^2 = 1.00^2$)
 Table 5.1 The parameters and their estimates ($\sigma^2 = 1.00^2$)

k	100	500	1000	1500	2000	2500	3000
$\alpha_1(1) = 0.30000$	0.32220	0.30451	0.25873	0.26213	0.27462	0.28438	0.27842
$\beta_{11}(1) = -0.40000$	-0.43734	-0.42718	-0.38153	-0.37992	-0.38586	-0.38043	-0.38857
$\beta_{11}(2) = 1.50000$	1.51740	1.49726	1.49121	1.51909	1.51724	1.49560	1.49801
$\beta_{12}(1) = -0.80000$	-0.60871	-0.71906	-0.81057	-0.80164	-0.79382	-0.77722	-0.79890
$\beta_{12}(2) = 0.60000$	0.68748	0.64352	0.69170	0.67187	0.65648	0.62934	0.61964
$\alpha_2(1) = 0.01000$	-0.02428	-0.01478	-0.01689	-0.01102	-0.01145	-0.00904	-0.00858
$\beta_{21}(1) = 2.00000$	1.95818	1.93753	2.00719	2.01894	2.02355	2.01757	2.02332
$\beta_{21}(2) = 0.20000$	0.16968	0.23445	0.20949	0.24558	0.21574	0.22928	0.23409
$\beta_{22}(1) = -0.40000$	-0.52342	-0.45447	-0.40885	-0.40924	-0.40059	-0.41109	-0.40388
$\beta_{22}(2) = 0.80000$	0.63932	0.80047	0.78747	0.77650	0.80953	0.81446	0.81262
$\beta_{23}(1) = 1.70000$	1.59585	1.76313	1.71215	1.71228	1.69950	1.69680	1.68706
$\beta_{23}(2) = 2.80000$	2.80171	2.85249	2.81653	2.81326	2.83064	2.83378	2.82224
δ (%)	7.31986	3.59338	2.52545	2.42469	1.86910	1.62536	1.42547

表 5.2 参数估计和参数估计误差 ($\sigma^2 = 0.20^2$)
 Table 5.2 The parameters and their estimates ($\sigma^2 = 0.20^2$)

k	100	500	1000	1500	2000	2500	3000
$\alpha_1(1) = 0.30000$	0.28894	0.29484	0.28793	0.28979	0.29308	0.29560	0.29463
$\beta_{11}(1) = -0.40000$	-0.42712	-0.41038	-0.39868	-0.39767	-0.39850	-0.39716	-0.39862
$\beta_{11}(2) = 1.50000$	1.50958	1.50080	1.49856	1.50394	1.50340	1.49897	1.49944
$\beta_{12}(1) = -0.80000$	-0.78852	-0.79355	-0.80836	-0.80469	-0.80184	-0.79762	-0.80160
$\beta_{12}(2) = 0.60000$	0.64182	0.61735	0.62430	0.61860	0.61427	0.60795	0.60573
$\alpha_2(1) = 0.01000$	-0.00830	-0.00986	-0.01083	-0.00982	-0.01003	-0.00960	-0.00953
$\beta_{21}(1) = 2.00000$	1.99498	1.98883	2.00192	2.00422	2.00500	2.00371	2.00481
$\beta_{21}(2) = 0.20000$	0.21684	0.21035	0.20359	0.21026	0.20390	0.20649	0.20737
$\beta_{22}(1) = -0.40000$	-0.44500	-0.41442	-0.40366	-0.40304	-0.40098	-0.40289	-0.40134
$\beta_{22}(2) = 0.80000$	0.76717	0.80124	0.79822	0.79580	0.80218	0.80312	0.80274
$\beta_{23}(1) = 1.70000$	1.66628	1.70963	1.70110	1.70151	1.69928	1.69891	1.69703
$\beta_{23}(2) = 2.80000$	2.83986	2.81770	2.80718	2.80519	2.80792	2.80816	2.80564
δ (%)	2.20864	0.94951	0.83493	0.74728	0.63918	0.57169	0.54928

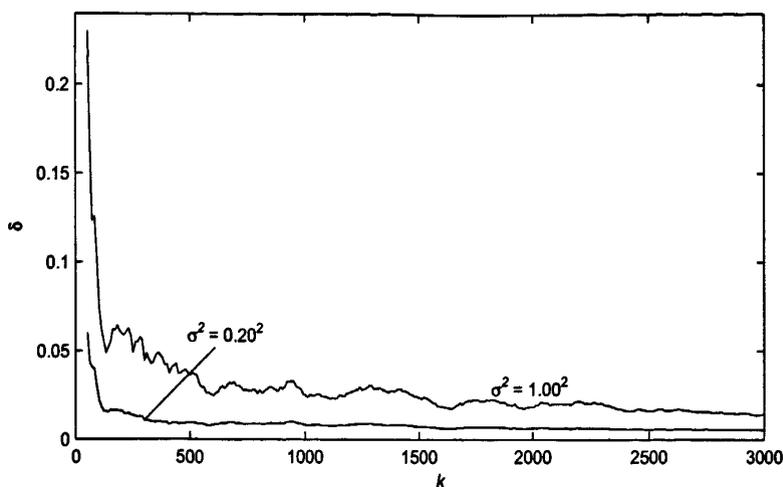


图 5.4 参数估计误差 δ 随数据长度 k 变化曲线
 Fig. 5.4 The parameter estimation errors δ versus k

从表 5.1-5.2 和图 5.4 可以得到, 由 ASM-RLS 算法估计得到的参数误差随着 k 的逐渐增加会越来越小, 且趋于零。噪声方差越小, 参数估计收敛于真值的速度越快, 参数精度越高。这个仿真例子验证了 ASM-RLS 算法的有效性。

例 5.3.2 对于例 5.3.1 所述的双输入单输出多率系统, 其传递函数模型为

$$y(kT) = \frac{1}{\alpha^*(z)} [\beta_{11}^*(z)u_1(kT) + \beta_{12}^*(z)u_1(kT + T_1) + \beta_{21}^*(z)u_2(kT) + \beta_{22}^*(z)u_2(kT + T_2) + \beta_{23}^*(z)u_2(kT + 2T_2)] + v(kT),$$

其中

$$\begin{aligned} \alpha^*(z) &= \alpha_1(z)\alpha_2(z) = 1 + \alpha^*(1)z^{-1} + \alpha^*(2)z^{-2} = 1 + 0.301z^{-1} + 0.003z^{-2}, \\ \beta_{11}^*(z) &= \alpha_2(z)\beta_{11}(z) = \beta_{11}^*(1)z^{-1} + \beta_{11}^*(2)z^{-2} + \beta_{11}^*(3)z^{-3} \\ &= -0.400z^{-1} + 1.496z^{-2} + 0.015z^{-3}, \\ \beta_{12}^*(z) &= \alpha_2(z)\beta_{12}(z) = \beta_{12}^*(1)z^{-1} + \beta_{12}^*(2)z^{-2} + \beta_{12}^*(3)z^{-3} \\ &= -0.800z^{-1} + 0.592z^{-2} + 0.006z^{-3}, \\ \beta_{21}^*(z) &= \alpha_1(z)\beta_{21}(z) = \beta_{21}^*(1)z^{-1} + \beta_{21}^*(2)z^{-2} + \beta_{21}^*(3)z^{-3} \\ &= 2.000z^{-1} + 0.120z^{-2} + 0.060z^{-3}, \\ \beta_{22}^*(z) &= \alpha_1(z)\beta_{22}(z) = \beta_{22}^*(1)z^{-1} + \beta_{22}^*(2)z^{-2} + \beta_{22}^*(3)z^{-3} \\ &= -0.400z^{-1} + 0.680z^{-2} + 0.240z^{-3}, \\ \beta_{23}^*(z) &= \alpha_1(z)\beta_{23}(z) = \beta_{23}^*(1)z^{-1} + \beta_{23}^*(2)z^{-2} + \beta_{23}^*(3)z^{-3} \\ &= 1.700z^{-1} + 3.310z^{-2} + 0.840z^{-3}. \end{aligned}$$

参数向量和信息向量为

$$\begin{aligned} \boldsymbol{\vartheta} &= [\alpha^*(1), \alpha^*(2), \beta_{11}^*(1), \beta_{11}^*(2), \beta_{11}^*(3), \beta_{12}^*(1), \beta_{12}^*(2), \beta_{12}^*(3), \\ &\quad \beta_{21}^*(1), \beta_{21}^*(2), \beta_{21}^*(3), \beta_{22}^*(1), \beta_{22}^*(2), \beta_{22}^*(3), \beta_{23}^*(1), \beta_{23}^*(2), \beta_{23}^*(3)]^T, \\ &= [0.301, 0.003, -0.400, 1.496, 0.015, -0.800, 0.592, 0.006, \\ &\quad 2.000, 0.120, 0.060, -0.400, 0.680, 0.240, 1.700, 3.310, 0.840]^T \\ \boldsymbol{\psi}(kT) &= [-y(kT - T), -y(kT - 2T), \boldsymbol{\psi}_1^T(kT), \boldsymbol{\psi}_2^T(kT)]^T, \\ \boldsymbol{\psi}_1(kT) &= [u_1(kT - T), u_1(kT - 2T), u_1(kT - 3T), u_1(kT + T_1 - T), \\ &\quad u_1(kT + T_1 - 2T), u_1(kT + T_1 - 3T)]^T, \\ \boldsymbol{\psi}_2(kT) &= [u_2(kT - T), u_2(kT - 2T), u_2(kT - 3T), \\ &\quad u_2(kT + T_2 - T), u_2(kT + T_2 - 2T), u_2(kT + T_2 - 3T), \\ &\quad u_2(kT + 2T_2 - T), u_2(kT + 2T_2 - 2T), u_2(kT + 2T_2 - 3T)]^T. \end{aligned}$$

仿真条件与例 5.3.1 得相同. 应用 AM-RLS 算法估计这个系统的参数. 参数估计和参数估计误差如表 5.3-5.4 所示, 参数估计误差 $\delta := \|\hat{\boldsymbol{\theta}}(kT) - \boldsymbol{\theta}\|/\|\boldsymbol{\theta}\|$ 随 k 变化曲线如图 5.5 所示.

表 5.3 参数估计和参数估计误差 ($\sigma^2 = 1.00^2$)
Table 5.3 The parameters and their estimates ($\sigma^2 = 1.00^2$)

k	100	500	1000	1500	2000	2500	3000
$\alpha^*(1) = 0.30100$	0.22777	0.16955	0.17558	0.19515	0.20973	0.21433	0.23200
$\alpha^*(2) = 0.00300$	0.05996	0.02159	0.00237	-0.00382	0.00411	0.00683	0.00824
$\beta_{11}^*(1) = -0.40000$	-0.45964	-0.42304	-0.37854	-0.37857	-0.38255	-0.37824	-0.38719
$\beta_{11}^*(2) = 1.49600$	1.60709	1.55432	1.52618	1.54212	1.54127	1.52042	1.51305
$\beta_{11}^*(3) = 0.01500$	0.16132	-0.12162	-0.12493	-0.08987	-0.08239	-0.07491	-0.06107
$\beta_{12}^*(1) = -0.80000$	-0.60695	-0.66626	-0.70875	-0.73885	-0.73450	-0.74578	-0.76933
$\beta_{12}^*(2) = 0.59200$	0.43293	0.62930	0.69491	0.67942	0.66116	0.65215	0.64951
$\beta_{12}^*(3) = 0.00600$	-0.06457	-0.13659	-0.12535	-0.12492	-0.10786	-0.09205	-0.06939
$\beta_{21}^*(1) = 2.00000$	1.88058	1.92423	2.00445	2.01752	2.02168	2.01618	2.02107
$\beta_{21}^*(2) = 0.12000$	-0.12621	-0.12146	-0.12022	-0.05594	-0.04978	-0.03020	0.00986
$\beta_{21}^*(3) = 0.06000$	0.09383	0.11156	0.11722	0.09587	0.11156	0.10245	0.10354
$\beta_{22}^*(1) = -0.40000$	-0.55369	-0.44497	-0.44495	-0.44843	-0.41307	-0.40708	-0.40751
$\beta_{22}^*(2) = 0.68000$	0.67316	0.81949	0.75203	0.73819	0.72250	0.71800	0.69744
$\beta_{22}^*(3) = 0.24000$	0.08056	0.19686	0.18192	0.18472	0.20577	0.20471	0.20544
$\beta_{23}^*(1) = 1.70000$	1.78537	1.72273	1.71329	1.69417	1.70756	1.70475	1.70380
$\beta_{23}^*(2) = 3.31000$	3.20842	3.11312	3.15383	3.18004	3.18209	3.18794	3.21333
$\beta_{23}^*(3) = 0.84000$	0.61660	0.50075	0.47202	0.54498	0.60812	0.61235	0.67868
δ (%)	11.60489	11.94387	11.57281	9.34889	8.05563	7.47378	5.64271

表 5.4 参数估计和参数估计误差 ($\sigma^2 = 0.20^2$)
Table 5.4 The parameters and their estimates ($\sigma^2 = 0.20^2$)

k	100	500	1000	1500	2000	2500	3000
$\alpha^*(1) = 0.30100$	0.26011	0.27560	0.27070	0.27827	0.28147	0.28513	0.29073
$\alpha^*(2) = 0.00300$	0.01014	0.00648	0.00262	0.00159	0.00319	0.00381	0.00413
$\beta_{11}^*(1) = -0.40000$	-0.41087	-0.40473	-0.39574	-0.39574	-0.39651	-0.39566	-0.39747
$\beta_{11}^*(2) = 1.49600$	1.52953	1.50698	1.50407	1.50598	1.50571	1.50046	1.49814
$\beta_{11}^*(3) = 0.01500$	0.00223	-0.01038	-0.02140	-0.00863	-0.00662	-0.00071	0.00536
$\beta_{12}^*(1) = -0.80000$	-0.75421	-0.77120	-0.77782	-0.78559	-0.78517	-0.78869	-0.79454
$\beta_{12}^*(2) = 0.59200$	0.58089	0.60008	0.61737	0.61147	0.60738	0.60356	0.60149
$\beta_{12}^*(3) = 0.00600$	-0.02913	-0.02379	-0.02504	-0.02168	-0.01787	-0.01252	-0.00629
$\beta_{21}^*(1) = 2.00000$	1.97398	1.98444	2.00074	2.00339	2.00419	2.00308	2.00410
$\beta_{21}^*(2) = 0.12000$	0.02135	0.07103	0.06022	0.08091	0.08306	0.09259	0.10481
$\beta_{21}^*(3) = 0.06000$	0.07394	0.07140	0.07430	0.06839	0.07121	0.06819	0.06739
$\beta_{22}^*(1) = -0.40000$	-0.43656	-0.41028	-0.41054	-0.41028	-0.40305	-0.40130	-0.40093
$\beta_{22}^*(2) = 0.68000$	0.68579	0.70664	0.69614	0.69169	0.68862	0.68658	0.68162
$\beta_{22}^*(3) = 0.24000$	0.19240	0.23340	0.22518	0.22822	0.23245	0.23430	0.23598
$\beta_{23}^*(1) = 1.70000$	1.71283	1.70384	1.70255	1.69915	1.70185	1.70141	1.70120
$\beta_{23}^*(2) = 3.31000$	3.24452	3.27189	3.26971	3.28123	3.28211	3.28791	3.29652
$\beta_{23}^*(3) = 0.84000$	0.71687	0.77570	0.75238	0.77751	0.79074	0.79937	0.81834
δ (%)	4.21788	2.34101	2.82214	2.01975	1.73454	1.35756	0.80031

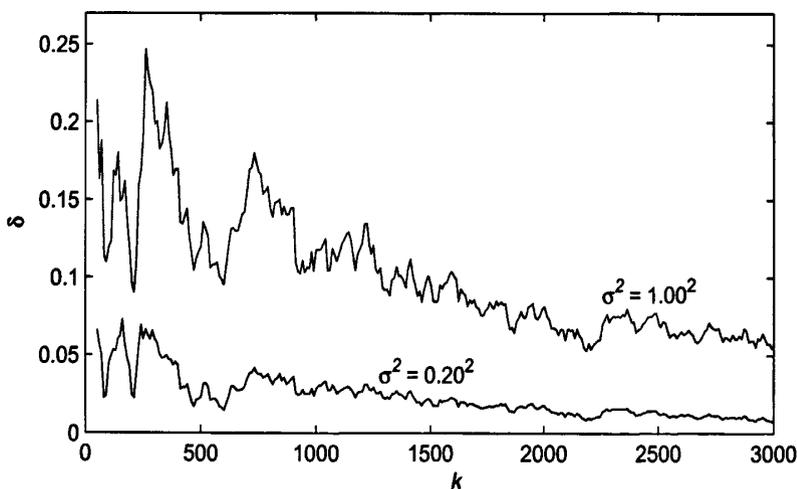


图 5.5 参数估计误差 δ 随数据长度 k 变化曲线

Fig. 5.5 The parameter estimation errors δ versus k

从表 5.3-5.4 和图 5.5, 可以得到, 由 AM-RLS 算法估计得到的参数误差随着 k 的逐渐增加会越来越小, 且趋于零. 噪声方差越小, 参数估计收敛于真值的速度越快, 参数精度越高. 这个仿真例子验证了 AM-RLS 算法的有效性.

从例 5.3.1 和例 5.3.2 看出, 由 ASM-RLS 算法估计的参数数目为 12, 由 AM-RLS 算法估计的参数数目为 17. 显然, AM-RLS 算法计算量要小于 ASM-RLS 算法.

5.4 本章小结

本章研究了多率输出误差系统的辨识. 应用离散化技术, 推导多率系统的离散状态空间模型, 得到其辨识模型. 基于辅助模型辨识思想, 提出了辅助子模型最小二乘算法估计系统的参数. 与辅助模型最小二乘算法比较, 提出的算法需要辨识的参数数目少, 有更小的计算量. 用数值仿真例子验证了提出算法的有效性, 并比较了辨识参数数目.

本章内容已投到国际期刊《Mathematics and Computers in Simulation》上, 已退修再

第六章 Box-Jenkins 多率系统

6.1 引言

在实际系统中,外部的干扰不总是白噪声,对于带有色噪声系统的辨识,利用常规的最小二乘算法得到的参数估计是有偏的.针对有色噪声单变量 CAR 系统, Feng 和 Zheng 提出消偏最小二乘 (BELS) 算法,得到参数的无偏估计^[103], Zheng 将 BELS 算法推广至多变量系统的辨识^[104]. Ding 等和 Zhang 和 Yang 借助偏差补偿原理,提出了偏差补偿最小二乘 (BCLS) 算法估计输出误差模型^[105,106],文献 [107] 将 BCLS 算法推广至 Box-Jenkins 模型.针对带有色噪声的自回归 (AR) 模型, Zheng 提出改进的最小二乘算法辨识系统参数^[108], Mahmoudi 和 Karimi 结合低阶和高阶 Yule-Walker 方程和改进的最小二乘算法估计系统参数^[109], Hasan 等利用极大似然估计方法得到系统的参数估计^[110].对于基于 Damped sinusoidal 模型噪声的 AR 系统, Hasan 等利用自相关分析法和最小二乘算法得到系统参数估计,同时得到噪声系统的参数估计^[111,112].对于基于 Ramp-Cepstrum 模型噪声的 AR 系统, Fattah 等利用自相关分析法和基于残差的最小二乘算法得到系统的参数估计^[113].对于滑动输出误差模型, Ding 等结合辅助模型辨识方法和多新息辨识方法提出了辅助模型多新息增广随机梯度算法辨识系统的参数^[88].对于输入输出都存在有色噪声的线性系统, Zhang 和 Feng 通过滤波的设计,用最小二乘辨识方法得到了系统参数的估计^[114].

上面所述文献都是带有色噪声单率系统的辨识,对于带有色噪声多率系统的辨识研究并不多见,基于此,本章研究 Box-Jenkins 多率系统的辨识问题.组织如下:第 6.2 节利用离散化技术,推导 Box-Jenkins 多率系统的状态空间模型.第 6.3 节结合辅助模型辨识思想和多新息辨识理论,辅助模型广义增广最小二乘算法和辅助模型多新息广义增广随机梯度算法,用数值仿真例子验证算法的有效性.第 6.4 节为本章总结.

6.2 模型描述

对于带有色噪声的多率多输入系统如图 6.1 所示, P_{cj} 是一个连续时间过程,有如下的状态空间模型

$$\begin{cases} \dot{\mathbf{x}}_j(t) = \mathbf{A}_{cj}\mathbf{x}_j(t) + \mathbf{B}_{cj}u_j(t), & j = 1, 2, \dots, r, \\ y_j(t) = \mathbf{C}_j\mathbf{x}_j(t) + D_ju_j(t), \end{cases}$$

其中 $\mathbf{x}_j(t) \in \mathbb{R}^{n_j}$ 为状态向量, $u_j(t) \in \mathbb{R}^1$ 为第 j 个输入通道的控制输入, $y_j(t) \in \mathbb{R}^1$ 为第 j 个通道的输出, $y_0(t) \in \mathbb{R}^1$ 为未知的无噪输出, $y(t) \in \mathbb{R}^1$ 为系统的输出. $e(t) \in \mathbb{R}^1$ 为外部干扰,有色噪声.系统 P_n 的状态空间模型为

$$\begin{cases} \dot{\mathbf{x}}_{cn}(t) = \mathbf{A}_{cn}\mathbf{x}_{cn}(t) + \mathbf{B}_{cn}v(t), \\ e(t) = \mathbf{C}_n\mathbf{x}_{cn}(t) + D_nv(t), \end{cases}$$

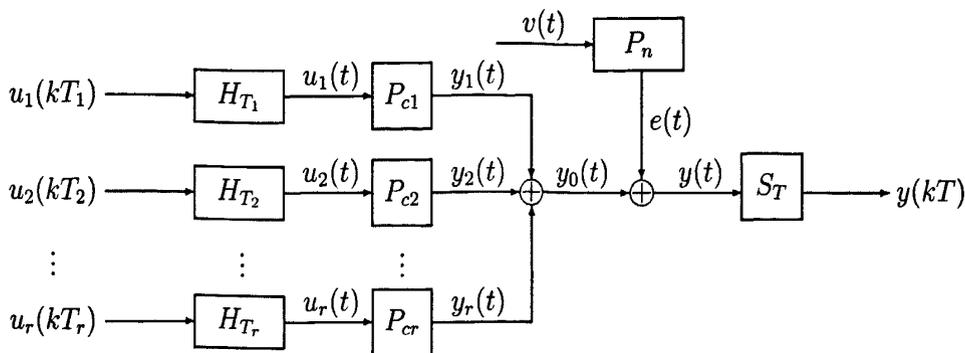


图 6.1 带有色噪声的多率多输入系统

Fig. 6.1 The multirate multiple-input systems with colored noises

其中 $\mathbf{x}_{cn}(t) \in \mathbb{R}^{n_e}$ 为状态, $v(t) \in \mathbb{R}^1$ 是控制输入, 为白噪声, $\mathbf{A}_{cj}, \mathbf{B}_{cj}, \mathbf{C}_j, \mathbf{D}_j, \mathbf{A}_{cn}, \mathbf{B}_{cn}, \mathbf{C}_n$ 和 \mathbf{D}_n 是具有适当维数的矩阵. H_{T_j} 为周期 $T_j := p_j h$ 的零阶保持器, S_T 为周期 $T := qh$ 的采样器, 其中 h 为基周期, T 为框架周期^[31]. $u_j(kT_j)$ 和 $y(kT)$ 是离散时间输入和输出信号. 对于多率系统图 6.1 所示的多率输入输出数据为 $\{u_1(kT_1), u_2(kT_2), \dots, u_r(kT_r), y(kT): k = 0, 1, 2, \dots\}$. 为了研究方便, 令 p_1, p_2, \dots, p_r 是正整数且互质, q 是 (p_1, p_2, \dots, p_r) 的最小公倍数.

仿照第 3.2 节和第 4.2 节, 应用离散化技术, 推导图 6.1 所示的多率系统的离散状态空间模型. 在框架周期 T 上离散化 P_{cj} ($j = 1, 2, \dots, r$), 得到

$$\begin{cases} \mathbf{x}_j(kT + T) = \mathbf{A}_j \mathbf{x}_j(kT) + \sum_{i=1}^{p_j} \mathbf{B}_{ji} u_j(kT - (i-1)T_j), \\ y_j(kT) = \mathbf{C}_j \mathbf{x}_j(kT) + \mathbf{D}_j u_j(kT), \quad j = 1, 2, \dots, r, \end{cases} \quad (6.2.1)$$

其中 $\mathbf{x}_j(kT) := \mathbf{x}_j(t)|_{t=kT}$, $u_j(kT + iT_j) := u_j(t)|_{t=kT+iT_j}$, $y_j(kT) := y_j(t)|_{t=kT}$ 和

$$\begin{aligned} \mathbf{A}_h &:= e^{\mathbf{A}_{cj}h} \in \mathbb{R}^{n_j \times n_j}, \quad \mathbf{A}_j := e^{\mathbf{A}_{cj}T}, \\ \mathbf{B}_{pj} &:= \int_0^{T_j} e^{\mathbf{A}_{cj}t} dt \mathbf{B}_{cj} \in \mathbb{R}^{n_j}, \\ \mathbf{B}_{ji} &:= \mathbf{A}_h^{q-ip_j} \mathbf{B}_{pj} \in \mathbb{R}^{n_j}, \quad j = 1, 2, \dots, r, i = 1, 2, \dots, q_j. \end{aligned}$$

在框架周期 T 上离散化 P_n , 得到

$$\begin{cases} \mathbf{x}_n(kT + T) = \mathbf{A}_n \mathbf{x}_n(kT) + \mathbf{B}_n v(kT), \\ e(kT) = \mathbf{C}_n \mathbf{x}_n(kT) + \mathbf{D}_n v(kT), \end{cases} \quad (6.2.2)$$

其中 $\mathbf{x}_n(kT) := \mathbf{x}_n(t)|_{t=kT}$, $v(kT) := v(t)|_{t=kT}$, $e(kT) := e(t)|_{t=kT}$ 和

$$\mathbf{A}_n := e^{\mathbf{A}_{cn}T} \in \mathbb{R}^{n_e \times n_e}, \quad \mathbf{B}_n := \int_0^T e^{\mathbf{A}_{cn}t} dt \mathbf{B}_{cn} \in \mathbb{R}^{n_e}.$$

参照图 6.1 所示, 输出方程可以表达为

$$y(kT) = y_0(kT) + e(kT)$$

$$= \sum_{j=1}^r y_j(kT) + e(kT). \quad (6.2.3)$$

方程 (6.2.1)–(6.2.3) 就是图 6.1 所示多率系统的离散状态空间模型。

6.3 辅助模型辨识算法

6.3.1 辅助模型广义增广最小二乘算法

令 z^{-1} 为单位后移算子: $z^{-1}u_j(kT + T_j) = u_j(kT + T_j - T)$. 从 (6.2.1)–(6.2.3) 式可得输出方程的表达式为

$$\begin{aligned} y(kT) &= \sum_{i=1}^r y_i(kT) + e(kT), \\ y_j(kT) &= \sum_{i=1}^{q_j} [C_j(z\mathbf{I} - \mathbf{A}_j)^{-1} \mathbf{B}_{ji} u_j(kT - (i-1)T_j) + D_j u_j(kT)] \\ &=: \sum_{i=1}^{q_j} \frac{\beta_{ji}(z)}{\alpha_j(z)} u_j(kT - (i-1)T_j), \quad j = 1, 2, \dots, r, \\ e(kT) &= \mathbf{C}_n(z\mathbf{I} - \mathbf{A}_n)^{-1} \mathbf{B}_n v(kT) + D_n v(kT) =: \frac{\eta(z)}{\gamma(z)} v(kT), \end{aligned}$$

其中 $\alpha_j(z)$, $\beta_{ji}(z)$, $\eta(z)$ 和 $\gamma(z)$ 为 z^{-1} 的多项式,

$$\alpha_j(z) := z^{-n_j} \det[z\mathbf{I} - \mathbf{A}_j] = 1 + \alpha_j(1)z^{-1} + \dots + \alpha_j(n_j)z^{-n_j},$$

$$\begin{aligned} \beta_{j1}(z) &:= z^{-n_j} \mathbf{C}_j \operatorname{adj}[z\mathbf{I} - \mathbf{A}_j] \mathbf{B}_{j1} + z^{-n_j} \det[z\mathbf{I} - \mathbf{A}_j] D_j \\ &= \beta_{j1}(0) + \beta_{j1}(1)z^{-1} + \dots + \beta_{j1}(n_j)z^{-n_j}, \end{aligned}$$

$$\beta_{ji}(z) := z^{-n_j} \mathbf{C}_j \operatorname{adj}[z\mathbf{I} - \mathbf{A}_j] \mathbf{B}_{ji} = \beta_{ji}(1)z^{-1} + \dots + \beta_{ji}(n_j)z^{-n_j}, \quad i = 2, 3, \dots, q_j,$$

$$\gamma(z) := z^{-n_e} \det[z\mathbf{I} - \mathbf{A}_n] = 1 + \gamma_1 z^{-1} + \gamma_2 z^{-2} + \dots + \gamma_{n_e} z^{-n_e},$$

$$\eta(z) := z^{-n_e} \mathbf{C}_n \operatorname{adj}[z\mathbf{I} - \mathbf{A}_n] \mathbf{B}_n = \eta_1 z^{-1} + \eta_2 z^{-2} + \dots + \eta_{n_e} z^{-n_e}.$$

定义参数向量 θ_j 和信息向量 $\phi_j(kT)$ 如下:

$$\begin{aligned} \theta_j &:= [\alpha_j(1), \dots, \alpha_j(n_j), \beta_{j1}(0), \beta_{j1}(1), \dots, \beta_{j1}(n_j), \beta_{j2}(1), \dots, \beta_{j2}(n_j), \dots, \\ &\quad \beta_{jq_j}(1), \dots, \beta_{jq_j}(n_j)]^T \in \mathbb{R}^{n_j + q_j n_j + 1}, \\ \phi_j(kT) &:= [-y_j(kT - T), \dots, -y_j(kT - n_j T), u_j(kT), u_j(kT - T), \dots, \\ &\quad u_j(kT - n_j T), \dots, u_j(kT + (q_j - 1)T_j - T), \dots, \\ &\quad u_j(kT + (q_j - 1)T_j - n_j T)]^T \in \mathbb{R}^{n_j + q_j n_j + 1}. \end{aligned}$$

则无噪输出方程可以表达为

$$y_0(kT) = \boldsymbol{\varphi}_s^T(kT)\boldsymbol{\theta}_s,$$

其中

$$\boldsymbol{\theta}_s := \begin{bmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \\ \vdots \\ \boldsymbol{\theta}_r \end{bmatrix} \in \mathbb{R}^{n_s}, \quad \boldsymbol{\varphi}_s(kT) := \begin{bmatrix} \boldsymbol{\phi}_1(kT) \\ \boldsymbol{\phi}_2(kT) \\ \vdots \\ \boldsymbol{\phi}_n(kT) \end{bmatrix} \in \mathbb{R}^{n_s}, \quad n_s := \sum_{j=1}^r (q_j + 1)n_j + r.$$

定义噪声系统的参数向量 $\boldsymbol{\theta}_n$ 和信息向量 $\boldsymbol{\varphi}_n(kT)$ 如下:

$$\begin{aligned} \boldsymbol{\theta}_n &:= [\gamma_1, \dots, \gamma_{n_e}, \eta_1, \dots, \eta_{n_e}] \in \mathbb{R}^{2n_e}, \\ \boldsymbol{\varphi}_n(kT) &:= [-e(kT - T), \dots, -e(kT - n_e T), v(kT - T), \dots, v(kT - n_e T)]^T \in \mathbb{R}^{2n_e}, \end{aligned}$$

则

$$e(kT) = \boldsymbol{\varphi}_n^T(kT)\boldsymbol{\theta}_n + v(kT).$$

因此, 图 6.1 所示多率系统的输出方程 (或辨识模型) 为

$$\begin{aligned} y(kT) &= \boldsymbol{\varphi}_s^T(kT)\boldsymbol{\theta}_s + \boldsymbol{\varphi}_n^T(kT)\boldsymbol{\theta}_n + v(kT) \\ &= \boldsymbol{\varphi}^T(kT)\boldsymbol{\theta} + v(kT), \end{aligned} \quad (6.3.1)$$

其中

$$\boldsymbol{\theta} := \begin{bmatrix} \boldsymbol{\theta}_s \\ \boldsymbol{\theta}_n \end{bmatrix} \in \mathbb{R}^{n_0}, \quad \boldsymbol{\varphi}(kT) := \begin{bmatrix} \boldsymbol{\varphi}_s(kT) \\ \boldsymbol{\varphi}_n(kT) \end{bmatrix} \in \mathbb{R}^{n_0}, \quad n_0 := n_s + 2n_e.$$

然而, 对于多率系统 6.3.1 的辨识模型, 不能用常规的最小二乘算法得到系统的参数估计.

原因: 在信息向量 $\boldsymbol{\varphi}(kT)$ 中含有未知的变量 $y_j(kT - iT)$ ($j = 1, 2, \dots, r, i = 1, 2, \dots, n_j$), $e(kT - lT)$ 和 $v(kT - lT)$ ($l = 1, 2, \dots, n_e$).

基于辅助模型辨识思想, 建立辅助模型

$$\hat{y}_j(kT) = \hat{\boldsymbol{\varphi}}_j^T(kT)\hat{\boldsymbol{\theta}}_j(kT),$$

其中 $\hat{y}_j(kT)$, $\hat{\boldsymbol{\varphi}}_j(kT)$ 和 $\hat{\boldsymbol{\theta}}_j(kT)$ 分别是 $y_j(kT)$, $\boldsymbol{\varphi}_j(kT)$ 和 $\boldsymbol{\theta}_j$ 在时刻 kT 的估计值. 信息向量 $\boldsymbol{\varphi}(kT)$ 中的不可测数据 $y_j(kT - iT)$ 用辅助模型输出 $\hat{y}_j(kT - iT)$ 代替.

$e(kT - lT)$ ($l = 1, 2, \dots, n_e$) 的估计值可由下式得到

$$\hat{e}(kT) = y(kT) - \hat{\boldsymbol{\varphi}}_s^T(kT)\hat{\boldsymbol{\theta}}_s(kT),$$

其中 $\hat{e}(kT)$, $\hat{\varphi}_s(kT)$ 和 $\hat{\theta}_s(kT)$ 分别是 $e(kT)$, $\varphi_s(kT)$ 和 θ_s 在时刻 kT 的估计值.

进一步地, $v(kT - lT)$ 的估计值可由下式得到

$$\hat{v}(kT) = y(kT) - \hat{\varphi}^T(kT)\hat{\theta}(kT),$$

其中 $\hat{v}(kT)$, $\hat{\varphi}(kT)$ 和 $\hat{\theta}(kT)$ 分别是 $v(kT)$, $\varphi(kT)$ 和 θ 在时刻 kT 的估计值.

在信息向量 $\varphi_j(kT)$ 中的 $y_j(kT)$ 用其估计值 $\hat{y}_j(kT)$ 代替, 在信息向量 $\varphi_n(kT)$ 中的 $e(kT)$ 和 $v(kT)$ 用其估计值 $\hat{e}(kT)$ 和 $\hat{v}(kT)$ 代替, 进一步地, $\varphi(kT)$ 用其估计值 $\hat{\varphi}(kT)$ 代替. 则根据辅助模型辨识思想, 得到估计多率系统 (6.3.1) 的参数 θ 的辅助模型广义增广最小二乘 (AM-GELS) 算法, 如下:

$$\hat{\theta}(kT) = \hat{\theta}(kT - T) + \mathbf{L}(kT)[y(kT) - \hat{\varphi}^T(kT)\hat{\theta}(kT - T)], \quad (6.3.2)$$

$$\mathbf{L}(kT) = \mathbf{P}(kT - T)\hat{\varphi}(kT)[1 + \hat{\varphi}^T(kT)\mathbf{P}(kT - T)\hat{\varphi}(kT)]^{-1}, \quad (6.3.3)$$

$$\mathbf{P}(kT) = [\mathbf{I} - \mathbf{L}(kT)\hat{\varphi}^T(kT)]\mathbf{P}(kT - T), \quad (6.3.4)$$

$$\hat{\varphi}(kT) = \begin{bmatrix} \hat{\varphi}_s(kT) \\ \hat{\varphi}_n(kT) \end{bmatrix}, \quad \hat{\theta}(kT) = \begin{bmatrix} \hat{\theta}_s(kT) \\ \hat{\theta}_n(kT) \end{bmatrix}, \quad (6.3.5)$$

$$\hat{\varphi}_s(kT) = \begin{bmatrix} \hat{\phi}_1(kT) \\ \hat{\phi}_2(kT) \\ \vdots \\ \hat{\phi}_n(kT) \end{bmatrix}, \quad \hat{\theta}_s(kT) = \begin{bmatrix} \hat{\theta}_1(kT) \\ \hat{\theta}_2(kT) \\ \vdots \\ \hat{\theta}_r(kT) \end{bmatrix}, \quad (6.3.6)$$

$$\hat{\phi}_j(kT) = [-\hat{y}_j(kT - T), \dots, -\hat{y}_j(kT - n_jT), u_j(kT), u_j(kT - T), \dots, u_j(kT - n_jT), \dots, u_j(kT + (q_j - 1)T_j - T), \dots, u_j(kT + (q_j - 1)T_j - n_jT)]^T, \quad (6.3.7)$$

$$\hat{\varphi}_n(kT) = [-\hat{e}(kT - T), \dots, -\hat{e}(kT - n_eT), \hat{v}(kT - T), \dots, \hat{v}(kT - n_eT)]^T, \quad (6.3.8)$$

$$\hat{y}_j(kT) = \hat{\varphi}_j^T(kT)\hat{\theta}_j(kT), \quad j = 1, 2, \dots, r, \quad (6.3.9)$$

$$\hat{e}(kT) = y(kT) - \hat{\varphi}_s^T(kT)\hat{\theta}_s(kT), \quad (6.3.10)$$

$$\hat{v}(kT) = y(kT) - \hat{\varphi}^T(kT)\hat{\theta}(kT), \quad (6.3.11)$$

其中 $\mathbf{P}(kT)$ 为参数估计误差的协方差阵.

初始化 AM-GELS 算法, 一般地, $\mathbf{P}(0) = p_0 \mathbf{1}_{n_0 \times n_0}$, p_0 为很大的正实数, 如 $p_0 = 10^6$, $\hat{\theta}(0)$ 为很小的正实向量, 如 $\hat{\theta}(0) = \mathbf{1}_{n_0}/p_0$. 用 AM-GELS 算法计算参数估计 $\hat{\theta}(kT)$ 的步骤如下:

1. 令 $k = 1$, $\hat{\theta}(0) = \mathbf{1}_{n_0}/p_0$, 选取数据长度 L_e .
2. 搜集多率输入输出数据 $\{u_1(kT_1), u_2(kT_2), \dots, u_r(kT_r), y(kT)\}$.
3. 由 (6.3.9) 式构成 $\hat{\phi}_j(kT)$, 由 (6.3.6) 式构成 $\hat{\varphi}_s(kT)$, 由 (6.3.8) 式构成 $\hat{\varphi}_n(kT)$, 进一步地, 由 (6.3.5) 式构成 $\hat{\varphi}(kT)$.

4. 由 (6.3.4) 式计算 $\mathbf{P}(kT)$, 由 (6.3.3) 式计算 $\mathbf{L}(kT)$, 由 (6.3.2) 式刷新 $\hat{\boldsymbol{\theta}}(kT)$, 由 (6.3.9) 式计算 $\hat{y}_j(kT)$, 由 (6.3.10) 式计算 $\hat{e}(kT)$, 以及由 (6.3.11) 式计算 $\hat{v}(kT)$.
5. 如果 $k < L_e$, k 增至 $k + 1$, 返回第 2 步继续计算. 否则, 得到参数估计 $\hat{\boldsymbol{\theta}}(L_e T)$.

6.3.2 辅助模型广义增广随机梯度算法

对于多率系统 (6.3.1), 基于辅助模型辨识思想, 得到估计多率系统 (6.3.1) 的参数 $\boldsymbol{\theta}$ 的辅助模型广义增广随机梯度 (AM-GESG) 算法, 如下:

$$\hat{\boldsymbol{\theta}}(kT) = \hat{\boldsymbol{\theta}}(kT - T) + \frac{\hat{\boldsymbol{\varphi}}(kT)}{r(kT)} f(kT), \quad (6.3.12)$$

$$f(kT) = y(kT) - \hat{\boldsymbol{\varphi}}^T(kT) \hat{\boldsymbol{\theta}}(kT - T), \quad (6.3.13)$$

$$r(kT) = r(kT - T) + \|\hat{\boldsymbol{\varphi}}(kT)\|^2, \quad r(0) = 1, \quad (6.3.14)$$

其中 $f(kT)$ 是单新息标量.

为了提高 AM-GESG 算法的收敛速度和估计精度, 基于多新息辨识理论, 提出辅助模型多新息广义增广随机梯度算法. 引入新息长度 p , 扩展单新息 $f(kT)$ 标量为多新息向量

$$\mathbf{F}(p, kT) = \begin{bmatrix} y(kT) - \hat{\boldsymbol{\varphi}}^T(kT) \hat{\boldsymbol{\theta}}(kT - T) \\ y(kT - T) - \hat{\boldsymbol{\varphi}}^T(kT - T) \hat{\boldsymbol{\theta}}(kT - T) \\ \vdots \\ y(kT - pT + T) - \hat{\boldsymbol{\varphi}}^T(kT - pT + T) \hat{\boldsymbol{\theta}}(kT - T) \end{bmatrix} \in \mathbb{R}^p,$$

其中 p 是正整数.

定义信息矩阵 $\Phi(p, kT)$ 和输出向量 $\mathbf{Y}(p, kT)$ 如下:

$$\Phi(p, kT) := [\hat{\boldsymbol{\varphi}}(kT), \hat{\boldsymbol{\varphi}}(kT - T), \dots, \hat{\boldsymbol{\varphi}}(kT - pT + T)] \in \mathbb{R}^{n_0 \times p},$$

$$\mathbf{Y}(p, kT) := [y(kT), y(kT - T), \dots, y(kT - pT + T)]^T \in \mathbb{R}^p.$$

则多新息向量可以表达为

$$\mathbf{F}(p, kT) = \mathbf{Y}(p, kT) - \Phi^T(p, kT) \hat{\boldsymbol{\theta}}(kT - T).$$

参照文 [75], 得到估计多率系统 (6.3.1) 的参数 $\boldsymbol{\theta}$ 的辅助模型多新息广义增广随机梯度 (AM-MI-GESG) 算法, 如下:

$$\hat{\boldsymbol{\theta}}(kT) = \hat{\boldsymbol{\theta}}(kT - T) + \frac{\hat{\boldsymbol{\varphi}}(kT)}{r(kT)} \mathbf{F}(p, kT), \quad (6.3.15)$$

$$\mathbf{F}(p, kT) = \mathbf{Y}(p, kT) - \Phi^T(p, kT) \hat{\boldsymbol{\theta}}(kT - T), \quad (6.3.16)$$

$$r(kT) = r(kT - T) + \|\hat{\boldsymbol{\varphi}}(kT)\|^2, \quad r(0) = 1, \quad (6.3.17)$$

$$\mathbf{Y}(p, kT) = [y(kT), \dots, y(kT - pT + T)]^T, \quad (6.3.18)$$

$$\Phi(p, kT) = [\hat{\varphi}(kT), \dots, \hat{\varphi}(kT - pT + T)], \quad (6.3.19)$$

$$\hat{\varphi}(kT) = \begin{bmatrix} \hat{\varphi}_s(kT) \\ \hat{\varphi}_n(kT) \end{bmatrix}, \quad \hat{\theta}(kT) = \begin{bmatrix} \hat{\theta}_s(kT) \\ \hat{\theta}_n(kT) \end{bmatrix}, \quad (6.3.20)$$

$$\hat{\varphi}_s(kT) = \begin{bmatrix} \hat{\varphi}_1(kT) \\ \hat{\varphi}_2(kT) \\ \vdots \\ \hat{\varphi}_n(kT) \end{bmatrix}, \quad \hat{\theta}_s(kT) = \begin{bmatrix} \hat{\theta}_1(kT) \\ \hat{\theta}_2(kT) \\ \vdots \\ \hat{\theta}_r(kT) \end{bmatrix}, \quad (6.3.21)$$

$$\hat{\varphi}_j(kT) = [-\hat{y}_j(kT - T), \dots, -\hat{y}_j(kT - n_jT), u_j(kT), u_j(kT - T), \dots, u_j(kT - n_jT), \dots, u_j(kT + (q_j - 1)T_j - T), \dots, u_j(kT + (q_j - 1)T_j - n_jT)]^T, \quad (6.3.22)$$

$$\hat{\varphi}_n(kT) = [-\hat{e}(kT - T), \dots, -\hat{e}(kT - n_eT), \hat{v}(kT - T), \dots, \hat{v}(kT - n_eT)]^T, \quad (6.3.23)$$

$$\hat{y}_j(kT) = \hat{\varphi}_j^T(kT) \hat{\theta}_j(kT), \quad j = 1, 2, \dots, r, \quad (6.3.24)$$

$$\hat{e}(kT) = y(kT) - \hat{\varphi}_s^T(kT) \hat{\theta}_s(kT), \quad (6.3.25)$$

$$\hat{v}(kT) = y(kT) - \hat{\varphi}^T(kT) \hat{\theta}(kT), \quad (6.3.26)$$

当 $p = 1$ 时, AM-MI-GESG 算法 (6.3.15)–(6.3.26) 是 AM-GESG 算法 (6.3.12)–(6.3.14).

初始化 AM-MI-GESG 算法, 一般地, $\mathbf{P}(0) = p_0 \mathbf{1}_{n_0 \times n_0}$, p_0 为很大的正实数, 如 $p_0 = 10^6$, $\hat{\theta}(0)$ 为很小的正实向量, 如 $\hat{\theta}(0) = \mathbf{1}_{n_0}/p_0$. 用 AM-MI-GESG 算法计算参数估计 $\hat{\theta}(kT)$ 的步骤如下:

1. 令 $k = 1$, $\hat{\theta}(0) = \mathbf{1}_{n_0}/p_0$, 选取数据长度 L_e .
2. 搜集多率输入输出数据 $\{u_1(kT_1), u_2(kT_2), \dots, u_r(kT_r), y(kT)\}$.
3. 由 (6.3.22) 式构成 $\hat{\varphi}_j(kT)$, 由 (6.3.21) 式构成 $\hat{\varphi}_s(kT)$, 由 (6.3.23) 式构成 $\hat{\varphi}_n(kT)$, 进一步地, 由 (6.3.20) 式构成 $\hat{\varphi}(kT)$.
4. 由 (6.3.19) 式构成信息矩阵 $\Phi(p, kT)$ 以及由 (6.3.18) 式构成输出向量 $\mathbf{Y}(p, kT)$.
5. 由 (6.3.17) 式计算 $r(kT)$, 由 (6.3.16) 式得到 $\mathbf{F}(p, kT)$, 由 (6.3.15) 式刷新 $\hat{\theta}(kT)$, 由 (6.3.24) 式计算 $\hat{y}_j(kT)$, 由 (6.3.25) 式计算 $\hat{e}(kT)$, 以及由 (6.3.26) 式计算 $\hat{v}(kT)$.
6. 如果 $k < L_e$, k 增至 $k + 1$, 返回第 2 步继续计算. 否则, 得到参数估计 $\hat{\theta}(L_e T)$.

计算参数估计 $\hat{\theta}(kT)$ 的流程图如图 6.2 所示.

与 AM-GESG 算法比较, AM-MI-GESG 算法充分应用了可测的输入输出数据, 因此有更快的收敛速度和更高的估计精度. 与 AM-GELS 算法比较, AM-MI-GESG 算法不需要计算参数估计误差协方差矩阵, 因此有较小的计算量.

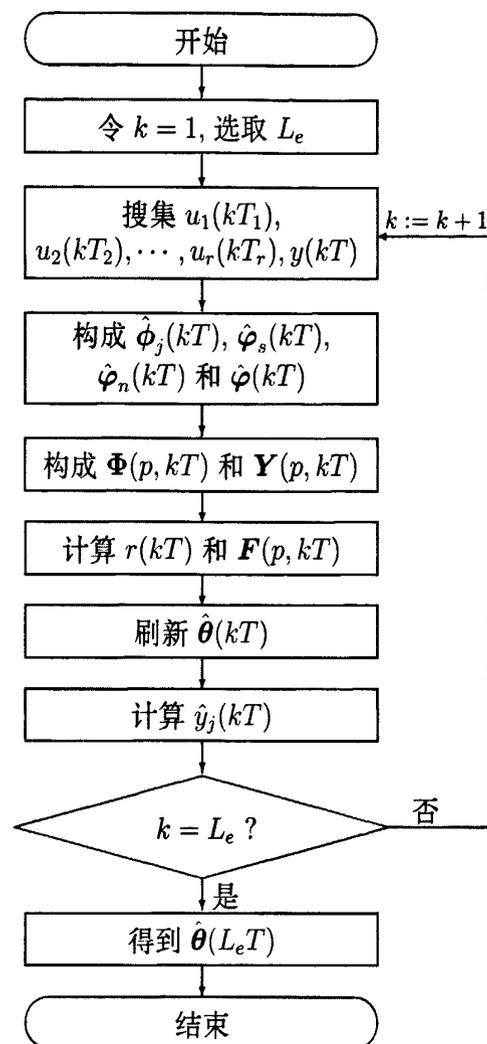


图 6.2 计算参数估计 $\hat{\theta}(kT)$ 的流程图

Fig. 6.2 The flowchart of computing the parameter estimate $\hat{\theta}(kT)$

6.3.3 仿真试验

例 6.3.1 对于双输入单输出多率系统，其输入刷新周期分别为 $T_1 = 3h$ 和 $T_2 = 2h$ ，输出采样周期为 $T = 6h$ ，取基周期为 $h = 1$ ，相应的传递函数模型为

$$\begin{aligned}
 y(kT) &= y_1(kT) + y_2(kT) + e(kT) \\
 &= \frac{1}{\alpha_1(z)} [\beta_{11}(z)u_1(kT) + \beta_{12}(z)u_1(kT + T_1)] + \frac{1}{\alpha_2(z)} [\beta_{21}(z)u_2(kT) \\
 &\quad + \beta_{22}(z)u_2(kT + T_2) + \beta_{23}(z)u_2(kT + 2T_2)] + \frac{\eta(z)}{\gamma(z)}(kT)v(kT),
 \end{aligned}$$

其中

$$\alpha_1(z) = 1 + \alpha_1(1)z^{-1} = 1 + 0.800z^{-1},$$

$$\begin{aligned}
\alpha_2(z) &= 1 + \alpha_2(1)z^{-1} = 1 + 0.008z^{-1}, \\
\beta_{11}(z) &= \beta_{11}(1)z^{-1} + \beta_{11}(2)z^{-2} = -0.160z^{-1} + 0.800z^{-2}, \\
\beta_{12}(z) &= \beta_{12}(1)z^{-1} + \beta_{12}(2)z^{-2} = -0.010z^{-1} + 0.240z^{-2}, \\
\beta_{21}(z) &= \beta_{21}(1)z^{-1} + \beta_{21}(2)z^{-2} = 0.400z^{-1} + 0.400z^{-2}, \\
\beta_{22}(z) &= \beta_{22}(1)z^{-1} + \beta_{22}(2)z^{-2} = 0.100z^{-1} + 1.000z^{-2}, \\
\beta_{23}(z) &= \beta_{23}(1)z^{-1} + \beta_{23}(2)z^{-2} = 1.270z^{-1} + 0.240z^{-2}, \\
\eta(z) &= \eta_1 z^{-1} + \eta_2 z^{-2} = -0.140z^{-1} + 1.200z^{-2}, \\
\gamma(z) &= \gamma_1 z^{-1} + \gamma_2 z^{-2} = -0.300z^{-1} + 0.031z^{-2}.
\end{aligned}$$

参数向量和信息向量为

$$\begin{aligned}
\theta_s &= [\alpha_1(1), \beta_{11}(1), \beta_{11}(2), \beta_{12}(1), \beta_{12}(2), \alpha_2(1), \beta_{21}(1), \beta_{21}(2), \\
&\quad \beta_{22}(1), \beta_{22}(2), \beta_{23}(1), \beta_{23}(2)]^T, \\
\theta_n &= [\gamma_1, \gamma_2, \eta_1, \eta_2]^T, \\
\theta &= [\theta_s^T, \theta_n^T]^T \\
&= [0.800, -0.160, 0.800, -0.010, 0.240, 0.008, 0.400, 0.400, \\
&\quad 0.100, 1.000, 1.270, 0.240, -0.300, 0.031, -0.014, 1.200]^T, \\
\varphi(kT) &= [\varphi_s^T(kT), \varphi_n^T(kT)]^T, \\
\varphi_s(kT) &= [-y_1(kT - T), u_1(kT - T), u_1(kT - 2T), u_1(kT + T_1 - T), \\
&\quad u_1(kT + T_1 - 2T), -y_2(kT - T), u_2(kT - T), u_2(kT - 2T), \\
&\quad u_2(kT + T_2 - T), u_2(kT + T_2 - 2T), u_2(kT + 2T_2 - T), \\
&\quad u_2(kT + 2T_2 - 2T)]^T, \\
\varphi_n(kT) &= [-e(kT - T), -e(kT - 2T), v(kT - T), v(kT - 2T)]^T,
\end{aligned}$$

仿真时, $\{u_1(kT_1)\}$ 和 $\{u_2(kT_2)\}$ 采用零均值单位方差不相关可测的持续激励信号序列. $\{v(kT)\}$ 采用零均值方差为 $\sigma^2 = 0.80^2$ 的白噪声序列. 应用 AM-GESG, AM-MI-GESG 和 AM-GELS 算法估计这个系统的参数. 参数估计和参数估计误差如表 6.1-6.5 所示, 参数估计误差 $\delta := \|\hat{\theta}(kT) - \theta\|/\|\theta\|$ 随 k 变化曲线如图 6.3 所示.

表 6.1 AM-GESG 参数估计和参数估计误差
Table 6.1 The AM-GESG parameter estimates and their errors

k	100	500	1000	1500	2000	2500	3000
$\alpha_1(1)= 0.80000$	0.37322	0.44854	0.49487	0.51520	0.52244	0.53092	0.53871
$\beta_{11}(1)= -0.16000$	0.11874	0.10249	0.09187	0.08745	0.08420	0.08057	0.07823
$\beta_{11}(2)= 0.80000$	0.52685	0.54913	0.55893	0.56563	0.56870	0.57201	0.57433
$\beta_{12}(1)= -0.01000$	-0.22684	-0.23008	-0.23575	-0.23961	-0.23918	-0.23875	-0.23937
$\beta_{12}(2)= 0.24000$	0.15039	0.17865	0.19034	0.19646	0.20010	0.20225	0.20469
$\alpha_2(1)= 0.00800$	-0.17599	-0.22785	-0.24458	-0.25577	-0.25946	-0.26407	-0.26771
$\beta_{21}(1)= 0.40000$	0.35711	0.35839	0.35676	0.35568	0.35599	0.35618	0.35712
$\beta_{21}(2)= -0.10000$	-0.11990	-0.12753	-0.12602	-0.12706	-0.12923	-0.13075	-0.13137
$\beta_{22}(1)= 0.10000$	0.23120	0.22140	0.21945	0.21869	0.21773	0.21640	0.21608
$\beta_{22}(2)= 1.00000$	0.23455	0.29545	0.31717	0.32983	0.33839	0.34527	0.35052
$\beta_{23}(1)= 1.27000$	0.39860	0.45451	0.47148	0.48216	0.48909	0.49519	0.50037
$\beta_{23}(2)= 0.24000$	0.01582	0.01115	0.00882	0.00869	0.00622	0.00504	0.00378
$\gamma_1= -0.30000$	-0.46891	-0.45813	-0.45562	-0.45349	-0.45146	-0.44926	-0.44812
$\gamma_2= 0.03100$	-0.15548	-0.14298	-0.13897	-0.13752	-0.13659	-0.13350	-0.13294
$\eta_1= -0.01400$	0.22436	0.21219	0.20226	0.19822	0.19435	0.19323	0.19109
$\eta_2= 1.20000$	0.60216	0.63957	0.65914	0.67078	0.67717	0.68224	0.68665
δ (%)	63.49053	59.21890	57.48481	56.57211	56.01281	55.52226	55.13605

表 6.2 AM-MI-GESG 参数估计和参数估计误差 ($p = 2$)
Table 6.2 The AM-MI-GESG parameter estimates and their errors ($p = 2$)

k	100	500	1000	1500	2000	2500	3000
$\alpha_1(1)= 0.80000$	0.49947	0.58505	0.63464	0.65385	0.65786	0.66436	0.66930
$\beta_{11}(1)= -0.16000$	0.12767	0.09384	0.07497	0.06627	0.06055	0.05358	0.04952
$\beta_{11}(2)= 0.80000$	0.63348	0.65940	0.66899	0.67742	0.68033	0.68394	0.68600
$\beta_{12}(1)= -0.01000$	-0.22229	-0.21834	-0.21973	-0.22143	-0.21840	-0.21552	-0.21475
$\beta_{12}(2)= 0.24000$	0.20403	0.23719	0.24902	0.25522	0.25841	0.25983	0.26098
$\alpha_2(1)= 0.00800$	-0.40025	-0.43923	-0.44315	-0.44647	-0.44464	-0.44657	-0.44833
$\beta_{21}(1)= 0.40000$	0.37801	0.36884	0.35992	0.35482	0.35272	0.35187	0.35238
$\beta_{21}(2)= -0.10000$	-0.12289	-0.13215	-0.12838	-0.12938	-0.13139	-0.13310	-0.13322
$\beta_{22}(1)= 0.10000$	0.26066	0.23720	0.22934	0.22534	0.22259	0.21934	0.21794
$\beta_{22}(2)= 1.00000$	0.34837	0.41851	0.44104	0.45291	0.46110	0.46784	0.47241
$\beta_{23}(1)= 1.27000$	0.49709	0.54832	0.56037	0.56802	0.57296	0.57792	0.58251
$\beta_{23}(2)= 0.24000$	-0.00826	-0.01681	-0.02199	-0.02238	-0.02601	-0.02764	-0.02898
$\gamma_1= -0.30000$	-0.44341	-0.42786	-0.42352	-0.41916	-0.41504	-0.41248	-0.41069
$\gamma_2= 0.03100$	-0.16710	-0.14121	-0.13264	-0.12823	-0.12581	-0.12131	-0.11951
$\eta_1= -0.01400$	0.22870	0.19885	0.18253	0.17624	0.16976	0.16740	0.16395
$\eta_2= 1.20000$	0.77300	0.82513	0.85220	0.86720	0.87639	0.88286	0.88781
δ (%)	56.36715	51.96559	50.28596	49.45464	48.89421	48.43123	48.10120

表 6.3 AM-MI-GESG 参数估计和参数估计误差 ($p = 6$)
 Table 6.3 The AM-MI-GESG parameter estimates and their errors ($p = 6$)

k	100	500	1000	1500	2000	2500	3000
$\alpha_1(1) = 0.80000$	0.73096	0.74454	0.77537	0.78387	0.77947	0.78160	0.77731
$\beta_{11}(1) = -0.16000$	0.08133	-0.00747	-0.03921	-0.05382	-0.06109	-0.07384	-0.07856
$\beta_{11}(2) = 0.80000$	0.82011	0.81370	0.80394	0.81284	0.81030	0.81165	0.80968
$\beta_{12}(1) = -0.01000$	-0.09997	-0.07632	-0.07402	-0.07591	-0.07185	-0.06744	-0.06726
$\beta_{12}(2) = 0.24000$	0.26320	0.28115	0.27916	0.28232	0.28298	0.28140	0.27650
$\alpha_2(1) = 0.00800$	-0.12622	-0.10822	-0.10464	-0.09594	-0.08847	-0.08421	-0.08466
$\beta_{21}(1) = 0.40000$	0.34989	0.35961	0.34980	0.34084	0.33979	0.34222	0.34634
$\beta_{21}(2) = -0.10000$	-0.19373	-0.19196	-0.17453	-0.17017	-0.16998	-0.17082	-0.16737
$\beta_{22}(1) = 0.10000$	0.18079	0.12791	0.11558	0.11100	0.10622	0.10163	0.10013
$\beta_{22}(2) = 1.00000$	0.65229	0.77179	0.81184	0.83112	0.84148	0.85110	0.85568
$\beta_{23}(1) = 1.27000$	0.92570	1.05361	1.08078	1.09745	1.10860	1.11725	1.12663
$\beta_{23}(2) = 0.24000$	0.10885	0.15043	0.15770	0.16795	0.17015	0.17213	0.17510
$\gamma_1 = -0.30000$	-0.33554	-0.33005	-0.32725	-0.32657	-0.32293	-0.32209	-0.32225
$\gamma_2 = 0.03100$	-0.16277	-0.09351	-0.07783	-0.07072	-0.06548	-0.05758	-0.05368
$\eta_1 = -0.01400$	0.14249	0.08142	0.05939	0.05653	0.04881	0.04643	0.04151
$\eta_2 = 1.20000$	1.00878	1.07332	1.10287	1.11680	1.12574	1.12912	1.13107
δ (%)	28.43141	18.91206	16.00909	14.66862	13.78723	12.98661	12.44310

表 6.4 AM-MI-GESG 参数估计和参数估计误差 ($p = 12$)
 Table 6.4 The AM-MI-GESG parameter estimates and their errors ($p = 12$)

k	100	500	1000	1500	2000	2500	3000
$\alpha_1(1) = 0.80000$	0.87009	0.78893	0.81091	0.81808	0.81074	0.81353	0.80294
$\beta_{11}(1) = -0.16000$	0.00464	-0.09916	-0.12388	-0.13610	-0.13613	-0.15212	-0.15336
$\beta_{11}(2) = 0.80000$	0.85611	0.82521	0.79856	0.81581	0.81116	0.81232	0.80838
$\beta_{12}(1) = -0.01000$	0.00748	0.01492	0.00788	-0.00266	-0.00152	0.00076	-0.00237
$\beta_{12}(2) = 0.24000$	0.25334	0.25035	0.24598	0.25422	0.25560	0.25365	0.24394
$\alpha_2(1) = 0.00800$	-0.02006	-0.01378	-0.01090	-0.00159	0.00114	0.00509	0.00313
$\beta_{21}(1) = 0.40000$	0.38524	0.40821	0.39435	0.37923	0.37730	0.38351	0.39178
$\beta_{21}(2) = -0.10000$	-0.15669	-0.14671	-0.11813	-0.11168	-0.11192	-0.11474	-0.10864
$\beta_{22}(1) = 0.10000$	0.18314	0.13460	0.12002	0.11682	0.11175	0.10629	0.10599
$\beta_{22}(2) = 1.00000$	0.86812	0.96278	0.99081	0.99853	0.99988	1.00439	1.00157
$\beta_{23}(1) = 1.27000$	1.16706	1.25576	1.25539	1.26015	1.26321	1.26472	1.27221
$\beta_{23}(2) = 0.24000$	0.19568	0.22838	0.22361	0.23064	0.23080	0.22966	0.23239
$\gamma_1 = -0.30000$	-0.25810	-0.30233	-0.30403	-0.30697	-0.30145	-0.30356	-0.30563
$\gamma_2 = 0.03100$	-0.09320	-0.01412	-0.00680	-0.00554	-0.00151	0.00278	0.00606
$\eta_1 = -0.01400$	0.09364	0.01336	-0.00304	0.00453	-0.00167	-0.00136	-0.00611
$\eta_2 = 1.20000$	1.11462	1.16279	1.18760	1.19705	1.20586	1.20272	1.19991
δ (%)	14.01159	5.11074	3.00427	2.72628	2.41251	2.02976	1.43290

表 6.5 AM-GELS 参数估计和参数估计误差
Table 6.5 The AM-GELS parameter estimates and their errors

k	100	500	1000	1500	2000	2500	3000
$\alpha_1(1) = 0.80000$	0.85588	0.81526	0.81830	0.80789	0.80676	0.81233	0.80874
$\beta_{11}(1) = -0.16000$	-0.19918	-0.16448	-0.17093	-0.18193	-0.16813	-0.18851	-0.18357
$\beta_{11}(2) = 0.80000$	0.81208	0.81536	0.77838	0.81461	0.80785	0.80917	0.80536
$\beta_{12}(1) = -0.01000$	-0.02323	0.02898	0.01734	-0.00855	-0.00899	-0.00083	-0.00542
$\beta_{12}(2) = 0.24000$	0.31845	0.25671	0.24167	0.26300	0.26612	0.25779	0.24106
$\alpha_2(1) = 0.00800$	-0.19212	-0.06400	-0.03168	-0.01085	-0.03891	-0.00328	-0.00375
$\beta_{21}(1) = 0.40000$	0.42143	0.41046	0.39240	0.36963	0.37322	0.38359	0.39460
$\beta_{21}(2) = -0.10000$	-0.22404	-0.15844	-0.10994	-0.09919	-0.11537	-0.10760	-0.09794
$\beta_{22}(1) = 0.10000$	0.23214	0.12681	0.10983	0.10746	0.10031	0.09228	0.09726
$\beta_{22}(2) = 1.00000$	1.00155	1.03277	1.03960	1.03487	1.02842	1.03177	1.02023
$\beta_{23}(1) = 1.27000$	1.15039	1.24842	1.25503	1.27634	1.24821	1.28193	1.28755
$\beta_{23}(2) = 0.24000$	-0.05372	0.16581	0.19318	0.21906	0.18734	0.22799	0.23052
$\gamma_1 = -0.30000$	-0.24253	-0.31588	-0.31638	-0.31139	-0.30276	-0.30730	-0.30838
$\gamma_2 = 0.03100$	0.02914	0.03450	0.02860	0.01903	0.02193	0.02185	0.02490
$\eta_1 = -0.01400$	0.05451	-0.02204	-0.02395	-0.00656	-0.00953	-0.00724	-0.01804
$\eta_2 = 1.20000$	1.15946	1.19907	1.21784	1.21987	1.22864	1.22282	1.21384
δ (%)	18.40981	5.76416	3.77236	2.96163	3.95715	2.62634	1.85835

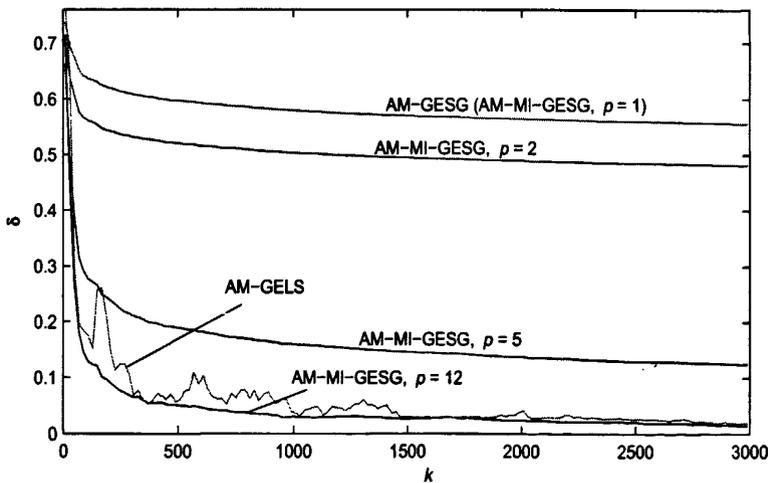


图 6.3 参数估计误差 δ 随数据长度 k 变化曲线
Fig. 6.3 The parameter estimation errors δ versus k

从表 6.1-6.5 和图 6.3 可以得到, 由 AM-MI-GESG 算法估计得到的参数误差随着 k 的逐渐增加会越来越小, 且趋于零. 与 AM-GESG 算法比较, AM-MI-GESG 算法 $p \geq 2$ 有更快的收敛速度和更高的估计精度. 当 $p \geq 2$ 时, AM-MI-GESG 算法得到的参数估计越来越接近 AM-GELS 算法的参数估计. 这个仿真例子验证了 AM-GESG 算法, AM-MI-GESG 和 AM-GELS 算法的有效性, 并比较了三种算法的收敛性质.

6.4 本章小结

本章研究了 Box-Jenkins 多率系统的辨识问题. 应用离散化技术, 推导多率系统的离散状态空间模型, 得到辨识模型. 为了提高算法的收敛速度和估计精度, 基于多新息辨识理论, 提出了辅助模型多新息广义增广随机梯度算法估计系统的参数. 借助于辅助模型辨识思想, 提出辅助模型广义增广最小二乘算法估计系统的参数. 与辅助模型广义增广最小二乘算法比较, 辅助模型多新息广义增广随机梯度算法不需要计算参数估计误差的协方差矩阵, 因此计算量较小. 辅助模型多新息广义增广随机梯度算法是辅助模型广义增广随机梯度算法和辅助模型广义增广最小二乘算法的折衷算法. 用数值仿真例子验证了这三种算法的有效性, 并比较三种算法的收敛性.

第七章 结论与展望

7.1 结论

随着社会的不断发展,在实际控制系统中,被控制对象一般为多变量系统,且结构复杂,对于这样的系统,由于各种原因,如控制信号变化速率不同,检测装置的采样周期不同,此时,用传统的离散采样系统描述系统是不现实的.在这种情形下,需要用多率系统描述系统.因此,多率系统有很强的应用背景和理论研究意义,但是对于多率系统辨识研究仍有一定的空白.

论文基于多新息辨识理论,辅助模型辨识思想,递阶辨识原理,最小二乘辨识原理和随机梯度辨识理论,研究了多变量系统和多率多输入系统的辨识问题.本文有如下的结论和创新:

1. 对于多变量系统,将多新息随机梯度算法推广至多输入多输出系统的辨识.由于多新息随机梯度算法充分利用输入输出数据,因此收敛速度和估计精度要高于随机梯度算法,又因为多新息随机梯度算法不需要计算协方差矩阵,所以计算量要小于最小二乘算法.应用鞅收敛定理分析了算法的收敛性.
2. 对于方程误差多率系统,为了提高算法的收敛速度和估计精度,基于多新息辨识理论,提出多新息随机梯度算法估计系统的参数,应用鞅定理分析了算法的收敛性.应用递阶辨识原理,提出计算量较小的递阶最小二乘算法估计系统的参数.
3. 对于输出误差多率系统,结合辅助模型辨识思想和多新息辨识理论,提出了辅助模型多新息随机梯度算法估计系统的参数.利用最小二乘原理,提出了辅助模型最小二乘算法估计系统的参数,用鞅收敛定理分析了辅助模型最小二乘算法的收敛性.
4. 对于带有子模型的多率系统,利用辅助模型辨识思想,对多率系统的子模型分别建立辅助子模型,提出基于辅助子模型最小二乘算法估计系统参数.与辅助模型最小二乘算法比较,需要辨识的参数数目少,有较小的计算量.
5. 对于 Box-Jenkins 多率系统,提出辅助模型广义增广最小二乘算法和辅助模型多新息广义增广随机梯度算法估计系统的参数.

7.2 展望

对于多率系统的研究已有半个世纪,多率系统在实际系统中有广泛的应用,所以对于多率系统辨识的研究有理论和应用价值.由于多率系统的结构复杂,其辨识问题还需进一步地研究和探索.

1. 多率系统一般采用提升或离散化技术得到其状态空间模型,得到维数大的 MIMO 系统,要考虑约束问题,其辨识算法的计算量较大.如何将有理分式传递函数模型推广至多率系统,得到与多率系统等价的辨识模型.
2. 带有色噪声多率系统辨识,能否用极大似然估计或期望估计方法得到系统的参数估计.
3. 多率系统的可辨识性问题.
4. 多率系统的收敛证明问题.收敛性证明一直是辨识问题的难点,对于多率系统也不例外.如何在弱收敛条件下,得到参数估计的一致收敛等问题.
5. 在通信领域和生物医学领域有很广泛的应用.如何将现有的多率系统的研究理论应用于实际领域,以适应实际系统的发展.

致 谢

清晰记得三年前，校园阳光暖暖，参加博士入学考试开启了我在江南大学学习的序幕。这三年里，学到的，懂得的很多。在即将离开美丽校园的时候，心中感慨万分。

感谢我的导师丁锋教授。我是从地域偏僻，学术较为不发达的地方而来，不知道的，不懂的很多，感谢老师的悉心指导，让我逐渐融入学术氛围，开阔了眼界，拓宽了思路，让我懂得学术的涵义。学生清晰记得第一篇文章录用时，老师的笑逐颜开。谢谢老师对学生的耐心和期待。三年来，老师从学习到生活都给了我无微不至的关怀。在此谨向我的导师丁锋教授表示诚挚的敬意和深深的感谢。

感谢 Dr. Yang Shi. Dr. Shi 是我在加拿大 University of Saskatchewan 联合培养时的导师。感谢 Dr. Shi 一直以来的支持。在学术上，Dr. Shi 让我了解到学术的另一片天空；敬佩 Dr. Shi 教育和管理 NCS Group 的方式；感谢 Dr. Shi 的关心，让我在异国他乡感受到来自 NCS Group 的暖意。

感谢实验室肖永松、于丽、陈慧波、丁洁、刘艳君、谢莉、张彩霞、陆静等同学的给予热情鼓励和帮助。

感谢 NCS Group 张俏、林杨、方华臻、于波、张辉、黄继、吴健、余显昊以及 Group 外尹武涛，张晓雨等同学给予的帮助和关心，珍惜在加国的美好记忆。

感谢 C205 邱芳、籍艳、王珍珍、楼旭阳、朱会宾等同学给予的帮助和关心，珍惜在江大校园的美好记忆。

感谢江南大学信控学院所有的老师在我读博士期间给予的关心。

深深感谢我的父亲韩那德木德，我的母亲华敖敦格日乐以及我的弟弟韩阿斯冷，这三年来一如既往的给予的支持和鼓励，使我顺利完成我的学业。我的家人一直是我前进的动力。

参考文献

- [1] Ding F, Chen T. Modeling and identification for multirate systems [J]. *Acta Automatica Sinica*, 2005, 31(1): 105-122
- [2] Glasson D P. Development and applications of multirate digital control [J]. *Control Systems Magazine*, 1983, 3(4): 2-8
- [3] Shi Y, Ding F, Chen T. Multirate crosstalk identification in xDSL systems [J]. *IEEE Transactions on Communication*, 2006, 54(10): 1878-1886
- [4] Lee S H. Multirate digital control system design and its application to computer disk drives [J]. *IEEE Transactions on Control Systems Technology*, 2006, 14(1): 124-133
- [5] Kadu S C, Bhushan M, Gudi R. Optimal sensor network design for multirate systems [J]. *Journal of Process Control*, 2008, 18(6): 954-609
- [6] Atlas L, Duhamel P. Recent developments in the core of digital signal processing [J]. *IEEE Signal Processing Magazine*, 1999, 16(1): 16-31
- [7] Raghavan H, Tangirala A K, Gopaluni R B, et al. Identification of chemical processes with irregular output sampling [J]. *Control Engineering Practice*, 2006, 14(5): 467-480
- [8] Han Z G, Shah S L, Robson C, et al. Development of softsensors for a pulp bleaching process using partial least squares [C]. In *proceedings of IEEE Advanced Process Control Applications for Industry Workshop*, April 28-30, 2003 Vancouver, Canada
- [9] Sahebsara M, Chen T, Shah S L. Frequency-domain parameter estimation of general multi-rate systems [J]. *Computers and Chemical Engineering*, 2006, 30(5): 838-849
- [10] Mo S Y, Chen X, Zhao J, et al. A two-stage method for identification of dual-rate systems with fast input and very slow output [J]. *Industrial & Engineering Chemistry Research*, 2009, 48(4): 1980-1988
- [11] Feng D G, Huang S C, Wang Z Z, et al. An unbiased parametric imaging algorithm for nonuniformly sampled biomedical system parameter estimation [J]. *IEEE Transactions on Medical Imaging*, 1996, 15(4): 512-518
- [12] 肖建. 多采样率数字控制系统 [M]. 北京: 科学出版社, 2003
- [13] Goodwin G C, Sin K S. *Adaptive Filtering Prediction and Control* [M]. Englewood Cliffs, New Jersey: Prentice-hall, 1984
- [14] Söderström T, Stoica P. *System Identification* [M]. Englewood Cliffs, New Jersey: Prentice-Hall, 1989
- [15] Ljung L. *System Identification: Theory for the User*, 2nd ed [M]. Englewood Cliffs, New Jersey: Prentice-Hall, 1999
- [16] 方崇智, 萧德云. 过程辨识 [M]. 北京: 清华大学出版社, 1988
- [17] Landau I D. Unbiased recursive identification using model reference adaptive techniques [J]. *IEEE Transactions on Automatic Control*, 1976, AC-21(2): 194-202

- [18] Landau I D. Elimination of the real positivity condition in the design of parallel MRAS [J]. IEEE Transactions on Automatic Control, 1978, AC-23(6): 1015-1020
- [19] Dugard L, Landau I D. Recursive output error identification algorithms theory and evaluation [J]. Automatica, 1980, 16(5): 443-462
- [20] Stoica P, Söderström T. Analysis of an output error identification algorithm [J]. Automatica, 1981, 17(6): 861-863
- [21] Söderström T, Stoica P. Some properties of the output error method [J]. Automatica, 1982, 18(1): 93-99
- [22] Ding F, Chen T. Combined parameter and output estimation of dual-rate systems using an auxiliary model [J]. Automatica, 2004, 40(10): 1739-1748
- [23] Ding F, Chen T. Parameter estimation of dual-rate stochastic systems by using an output error method [J]. IEEE Transactions on Automatic Control, 2005, 50(9): 1436-1441
- [24] 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
- [25] 丁锋, 萧德云, 丁韬. 多新息随机梯度辨识方法 [J]. 控制理论与应用, 2003, 20(6): 870-874
- [26] Ding F, Chen T. Performance analysis of multi-innovation gradient type identification methods [J]. Automatica, 2007, 43(1): 1-14
- [27] Ding F. Several multi-innovation identification methods [J]. Digital Signal Processing, 2009, doi:10.1016/j.dsp.2009.10.030
- [28] Han L L, Ding F. Multi-innovation stochastic gradient algorithms for multi-input multi-output ARX systems [J]. Digital Signal Processing, 2009, 19(4): 545-554
- [29] Ding F, Chen T. Hierarchical gradient-based identification of multivariable discrete-time systems [J]. Automatica, 2005, 41(2): 315-325
- [30] Ding F, Chen T. Hierarchical least squares identification methods for multivariable systems [J]. IEEE Transactions on Automatic Control, 2005, 50(3): 397-402
- [31] Ding F, Chen T. Hierarchical identification of lifted state-space models for general dual-rate systems [J]. IEEE Transactions on Circuits and Systems, 2005, 52(6): 1179-1187
- [32] Chen H F, Guo L. Identification and Stochastic Adaptive Control [M]. Boston: Birkhauser, 1991
- [33] 谢新民, 丁锋. 自适应控制系统 [M]. 北京: 清华大学出版社, 2002
- [34] 潘嵘嵘, 项林英, 郭戈. 多采样率数字控制系统的进展与走向 [J]. 兰州理工大学学报, 2005, 31(5): 78-82
- [35] Sklansky J, Ragazzini J R. Analysis of errors in sampled-data feedback systems [J]. Transactions of the AIEE, 1955, 74(PartII): 65-71
- [36] Kranc G M. Input-Output analysis of multirate feedback systems [J]. IRE Transactions Automatic Control, 1957, 3(1): 21-28

- [37] Kalman R E, Bertram J E. A unified approach to the theory of sampling systems [J]. Journal of Franklin Institute, 1959, 267(5): 405-436
- [38] Chen T, Qiu L. H_∞ design of general multirate sampled-data control systems [J]. Automatica, 1994, 30(7): 1139-1152
- [39] Qiu L, Chen T. H_2 optimal design of multirate sampled-data systems [J]. IEEE Transactions on Automatic Control, 1994, 39(12):2506-2511
- [40] Lee J H, Gelormino M S, Morari M. Model predictive control of multirate sampled-data systems: A state-space approach [J]. International Journal of Control, 1992, 55(1): 153-191
- [41] Scattolini R, Schiavonoi N. A multirate model-based predictive controller [J]. IEEE Transactions on Automatic Control, 1995, 40(6): 1093-1097
- [42] Sheng J, Chen T, Shah S L. Generalized predictive control for non-uniformly sampled systems [J]. Journal of Process Control, 2002, 12(8): 875-885
- [43] Albertos P, Salt J, Tornero J. Dual-rate adaptive control [J]. Automatica, 1996, 32(7): 1027-1030
- [44] Mizumoto I, Chen T, Ohdaira S, et al. Adaptive output feedback control of general MIMO systems using multirate sampling and its application to a cart-crane system [J]. Automatica, 2007, 43(12): 2077-2085
- [45] 倪博溢, 萧德云. 多采样率系统的辨识问题综述 [J]. 控制理论与应用, 2009, 26(1): 62-68
- [46] Friedland B. Sampled-data control systems containing periodically varying members [C]. Proceeding of the first IFAC Congress, 1960
- [47] Khargonekar P P, Poolla K, Tannenbaum A. Robust control of linear time-invariant plants using periodic compensation [J]. IEEE Transactions on Automatic Control, 1985, 30(11): 1088-1096
- [48] Chen T, Francis B. Optimal Sampled-Data Control Systems [M]. London: Springer, 1995
- [49] Gustafsson F, Gunnarsson S, Ljung L. Shaping frequency-dependent time resolution when estimating spectral properties with parametric methods [J]. IEEE Transactions on Signal Processing, 1997, 45(4): 1025-1035
- [50] Lahalle E, Perez P, Fleury G. Simplified LMS algorithm in the case of non-uniformly sampled signals [C]. Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing. HongKong: IEEE Press, 2003, 81-84
- [51] 吕立华, 宋执环, 李平. 一种基于小波多分辨率分析的多采样率系统辨识方法 [J]. 控制理论与应用, 2002, 19(2): 225-228
- [52] McKelvey T, Akçay H, Ljung L. Subspace-based multivariable system identification from frequency response data [J]. IEEE Transactions on Automatic Control, 1996, 41(7): 960-979
- [53] McKelvey T, Akçay H, Ljung L. Subspace-based identification of infinite-dimensional multivariable systems from frequency-response data [J]. Automatica, 1996, 32(6): 885-902
- [54] Akçay H, Türkay S. Frequency domain subspace-based identification of discrete-time power spectra from non uniformly spaced measurements [J]. Automatica, 2004, 40(8): 1333-1347

- [55] Akçay H. An insight into instrumental variable frequency-domain subspace identification [J]. *Automatica*, 2010, 46(2): 375-382
- [56] Gillberg J, Ljung L. Frequency domain identification of continuous-time output error models, Part I: Uniformly sampled data and frequency function approximation [J]. *Automatica*, 2010, 46(1): 1-10
- [57] Gillberg J, Ljung L. Frequency domain identification of continuous-time output error models, Part II: Non-uniformly sampled data and B-spline output approximation [J]. *Automatica*, 2010, 46(1): 11-18
- [58] Lu W P, Fisher D G. Output estimation with multi-rate sampling [J]. *International Journal of Control*, 1988, 48(1): 149-160
- [59] Lu W P, Fisher D G. Least-squares output estimation with multirate sampling [J]. *IEEE Transactions on Automatic Control*, 1989, 34(6): 669-672
- [60] Ding F, Liu P X, Shi Y. Convergence analysis of estimation algorithms for dual-rate stochastic systems [J]. *Applied Mathematics and Computation*, 2006, 176(1): 245-261
- [61] Ding F, Chen T. Parameter estimation for dual-rate systems with finite measurement data [J]. *Dynamics of Continuous, Discret and Impulsive Systems, Series B: Applications & Algorithms*, 2004, 11(1-2): 101-121
- [62] Ding F, Chen T. Identification of dual-rate systems based on finite impulse response models [J]. *International Journal of Adaptive Control and Signal Processing*, 2004, 18(7): 589-598
- [63] Ding J, Shi Y, Wang H G, et al. A modified stochastic gradient based parameter estimation algorithm for dual-rate sampled-data systems [J]. *Digital Signal Processing*, 2009, doi:10.1016/j.dsp.2009.10.023
- [64] Zhu Y C, Telkamp H, Wang J H, et al. System identification using slow and irregular output samples [J]. *Journal of Processing Control*, 2009, 19(1): 58-67
- [65] Ding J, Ding F. The residual based extended least squares identification method for dual-rate systems [J]. *Computers & Mathematics with Applications*, 2008, 56(6): 1479-1487
- [66] Ding F, Ding J. Least squares parameter estimation for systems with irregularly missing data [J]. *International Journal of Adaptive Control and Signal Processing*, 2010, 24, doi 10.1002/acs.1141
- [67] Ding J, Han L L, Chen X M. Time series AR modeling with missing observations based on the polynomial transformation [J]. *Mathematical and Computer Modelling*, 2010, 51(5-6): 527-536
- [68] Li D, Shah S L, Chen T. Identification of fast-rate models from multirate data [J]. *International Journal of Control*, 2001, 74(7): 680-689
- [69] Li W H, Han Z G, Shah S L. Subspace identification for FDI in systems with non-uniformly sampled multirate data [J]. *Automatica*, 2006, 42(4): 619-627
- [70] Sheng J, Chen T, Shah S L. Optimal filtering for multirate systems [J]. *IEEE Transactions on Circuits and Systems-II: Express Briefs*, 2005, 52(4): 228-232
- [71] Li W H, Shah S L, Xiao D Y. Kalman filters in non-uniformly sampled multirate systems: For FDI and beyond [J]. *Automatica*, 2008, 44(1): 199-208

- [72] Yu B, Shi Y, Huang H. $l_2 - l_\infty$ filtering for multirate systems using lifted models [J]. Circuits, Systems, and Signal Processing, 2008, 27(5): 699-711
- [73] Liu X G, Lu J. Least squares based iterative identification for a class of multirate systems [J]. Automatica, 2010, 46(3): 549-554
- [74] Ding F, Qiu L, Chen T. Reconstruction of continuous-time systems from their non-uniformly sampled discrete-time systems [J]. Automatica, 2009, 45(2): 324-332
- [75] Han L L, Ding F. Identification for multirate multi-input systems using the multi-innovation identification theory [J]. Computers & Mathematics with Applications, 2009, 57(9): 1438-1449
- [76] Han L L, Sheng J, Ding F, et al. Auxiliary model identification method for multirate multi-input systems based on least squares [J]. Mathematical and Computer Modelling, 2009, 50(7-8): 1100-1106
- [77] Ding F, Liu P X, Yang H Z. Parameter identification and intersample output estimation for dual-rate systems [J]. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 2008, 38(4): 966-975
- [78] 袁平. 多变量系统辨识方法比较研究 (I) [D]. 无锡: 江南大学通信与控制工程硕士论文, 2008
- [79] Sinha N K, Kwong Y H. Recursive estimation of the parameters of linear multivariable systems [J]. Automatica, 1979, 15(4): 471-475
- [80] El-Sherief H. Parameters identification of a state-space model of multivariable systems using the extended least-squares method [J]. IEEE Transactions on Systems, Man, and Cybernetics, 1981, 11(3): 223-227
- [81] Overshee P V, De Moor B. N4SID: Subspace algorithm for the identification of combined determine-stochastic systems [J]. Automatica, 1994, 30(1): 75-93
- [82] Verhaegen M. Identification of the deterministic part of MIMO state space models given in innovations form input-output data [J]. Automatica, 1994, 30(1): 61-71
- [83] Chou C T, Verhaegen M. Subspace algorithms for the identification of multivariable dynamic error-in-variable models [J]. Automatica, 1997, 33(10): 1857-1869
- [84] Zhang C J, Bitmead R R. Subspace system identification for training-based MIMO channel estimation [J]. Automatica, 2005, 41(9): 1623-1632
- [85] 丁锋. 时变参数系统辨识及其应用 [D]. 北京: 清华大学自动化系博士论文, 1994
- [86] Liu Y J, Xiao Y S, Zhao X L. Multi-innovation stochastic gradient algorithms for multi-input single-output systems using the auxiliary model [J]. Applied Mathematics and Computation, 2009, 215(4): 1477-1483
- [87] 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

- [88] Ding F, Liu P X, Liu G J. Auxiliary model based multi-innovation extended stochastic gradient parameter estimation with colored measurement noises [J]. *Signal Processing*, 2009, 89(10): 1883-1890
- [89] Ding F, Liu P X, Liu G. Multi-innovation least squares identification for system modelling [J]. *IEEE Transactions on Systems, Man, and Cybernet, Part B: Cybernetics*, 2010, doi:10.1109/TSMCB.2009.2028871
- [90] Wang D Q, Ding F. Performance analysis of the auxiliary models based multi-innovation stochastic gradient estimation algorithm for output error systems [J]. *Digital Signal Processing*, 2010, 20(3): 750-762
- [91] Zhang J B, Ding F, Shi Y. Self-tuning control based on multi-innovation stochastic gradient parameter estimation [J]. *Systems & Control Letters*, 2009, 58(1): 69-75
- [92] 郭雷. 时变随机系统 — 稳定性、估计与控制 [M]. 长春: 吉林科学技术出版社, 1993
- [93] Qiu L, Chen T. Multirate sampled-data systems: all H_∞ suboptimal controllers and the minimum entropy controller [J]. *IEEE Transactions on Automatic Control*, 1999, 44(3): 537-550
- [94] 丁锋. 多变量系统辨识 [D]. 北京: 清华大学自动化系硕士论文, 1990
- [95] 丁锋, 谢新民. 传递函数阵子模型参数递推估计: 辅助模型方法 [J]. *控制与决策*, 1991, 6(6): 447-452
- [96] 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
- [97] Ding F, Shi Y, Chen T. Auxiliary model based least-squares identification methods for Hammerstein output-error systems [J]. *Systems & Control Letters*, 2007, 56(5): 373-380
- [98] Xu Z Y, Tsatsanis M K. Blind channel estimation for long code multiuser CDMA systems [J]. *IEEE Transactions on Signal Processing*, 2000, 48(4): 988-1001
- [99] Tsatsanis M K, Giannakis G B. Optimal decorrelating receivers for DS-CDMA systems: A signal processing framework [J]. *IEEE Transactions on Signal Processing*, 1996, 44(12): 3044-3055
- [100] 王冬青, 丁锋. 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+1010
- [101] 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
- [102] Liu Y J, Xie L, Ding F. An auxiliary model recursive least squares algorithm and its convergence for non-uniformly sampled multirate systems [J]. *Proceedings of the Institution of Mechanical Engineers, Part I - Journal of Systems and Control Engineering*, 2009, 223(4): 445-454
- [103] Feng C B, Zheng W X. Robust identification of stochastic linear systems with correlated noise [J]. *IEE Proceedings - Control Theory & Applications*, 1991, 138(5): 484-492
- [104] Zheng W X. Least-squares identification of a class of multivariable systems with correlated disturbances [J]. *Journal of the Franklin Institute*, 1999, 336(8): 1309-1324

- [105] Ding F, Chen T, Qiu L. Bias compensation based recursive least squares identification algorithm for MISO systems [J]. IEEE Transactions on Circuits and Systems – II: Express Briefs, 2006, 53(5): 349-353
- [106] 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算法 [J]. 自动化学报, 2007, 33(10): 1054-1060
- [107] 杨慧中, 张勇. Box-Jenkins 模型偏差补偿方法与其他辨识方法的比较 [J]. 控制理论与应用, 2007, 24(2): 215-222
- [108] Zheng W X. Estimation of the parameters of autoregressive signals from colored noise-corrupted measurements [J]. IEEE Signal Processing Letters, 2000, 7(7): 201-204
- [109] Mahmoudi A, Karimi M. Parameter estimation of autoregressive signals from observations corrupted with colored noise [J]. Signal Processing, 2010, 90(1): 157-164
- [110] Hasan M K, Mollah M M, Yahagi T. Maximum likelihood estimation of autoregressive systems degraded by colored noise [J]. Journal of Signal Processing, 1997, 1(2): 125-133
- [111] Hasan M K, Rahim Chowdhury A K M Z, Khan M R. Identification of autoregressive signals in colored noise using damped sinusoidal model [J]. IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications, 2003, 50(7): 966-969
- [112] Hasan M K, Fattah S A, Khan M R. Identification of noisy AR systems using damped sinusoidal model of autocorrelation function [J]. IEEE Signal Processing Letters, 2003, 10(6): 157-160
- [113] Fattah S A, Zhu W P, Ahmad M O. Identification of autoregressive systems in noise based on a Ramp-Cepstrum model [J]. IEEE Transactions on Circuits and Systems – II: Express Briefs, 2008, 55(10): 1051-1055
- [114] Zhang Y, Feng C B. Unbiased parameter estimation of linear systems with colored noises [J]. Automatica, 1997, 33(5): 969-973

攻读博士学位期间发表和录用论文目录

1. Han L L, Ding F. Identification for multirate multi-input systems using the multi-innovation identification theory. *Computers & Mathematics with Applications*. 2009, 57(9): 1438-1449. (SCI: 441HQ; EI: 20091412008131)
2. Han L L, Ding F. Multi-innovation stochastic gradient algorithms for multi-input multi-output ARX systems. *Digital Signal Processing*, 2009, 19(4): 545-554. (SCI: 444RI; EI: 20091812066794)
3. Han L L, Sheng J, Ding F, Shi Y. Auxiliary model identification method for multirate multi-input systems based on least squares. *Mathematical and Computer Modelling*, 2009, 50(7-8): 1100-1106. (SCIE: 490CP; EI: 20093612288166)
4. Han L L, Ding F. Parameters estimation for multirate multi-input systems by using the least squares methods. Submitted to *Mathematics and Computers in Simulation*. 已修改再审.
5. Han L L, Ding F. Parameter estimation of multirate multivariable systems by using auxiliary model and multi-innovation identification techniques. Accepted by *Journal of Systems Engineering and Electronics*. (SCIE)
6. Ding J, Han L L, Ding F. Time series AR modeling with missing observations based on the polynomial transformation. *Mathematical and Computer Modelling*, 2010, 51(5-6): 527-536. (SCIE: 544NP; EI: 20100312640023)
7. Han L L, Sheng J, Ding F, Shi Y. Recursive least squares identification for multirate multi-input single-output systems. 28th American Control Conference (ACC 2009), St. Louis, Missouri, USA, June 10-12, 2009, 5604-5609. (EI: 20094812497633)
8. Han L L, Ding F. Convergence analysis of MISG algorithms for MIMO systems. 2008 Chinese Control and Decision Conference (CCDC 2008), Yantai, Shandong, China, July 2-4, 2008. 4534-4539. (EI: 084011610563)
9. Han L L, Ding J, Ding F. Auxiliary model based multi-innovation stochastic gradient identification for multirate multi-input systems. 2010 Chinese Control and Decision Conference (CCDC 2010), Xuzhou, Jiangsu, China. May 26-28, 2010. (EI)
10. Han L L, Ding F. Identification of the multirate MISO systems based on MISG algorithm. 2008 Jiangsu Doctoral Forum on Control Science and Engineering. Nanjing University of Science & Technology, Nanjing, China, May 23-25, 2008
11. Han L L, Ding F. Parameter estimation for a class of multirate systems with colored noises using the auxiliary model and multi-innovation methods. 待投.

