



规范状态空间系统辨识方法

摘要

因为状态空间模型既包含了未知状态,又包含了未知参数,且二者是非线性乘积关系,使得辨识问题变得复杂.针对这一问题,详细研究了规范状态空间系统的状态与参数联合估计方法.采用交互估计理论,即采用递推方法或迭代方法实现系统状态与参数的交互估计.基本思路是在计算参数估计时,辨识算法信息向量中的未知状态用其估计值代替,然后利用获得的参数估计,设计基于参数估计的状态观测器或基于参数估计的 Kalman 滤波算法估计系统的状态,二者形成一个交互计算过程(递阶计算过程).沿着这条思路,分别从递推方案和迭代方案,研究和提出了基于状态观测器和基于 Kalman 滤波状态估计的随机梯度辨识算法、递推最小二乘辨识算法、多新息随机梯度辨识算法、多新息最小二乘辨识算法,以及模型分解的辨识算法,并给出了几个典型算法的计算步骤、流程图和计算量.

关键词

参数估计;递推辨识;迭代辨识;最小二乘;梯度搜索;状态观测器;Kalman 滤波;状态估计;模型分解;状态空间系统

中图分类号 TP273

文献标志码 A

收稿日期 2014-12-08

资助项目 国家自然科学基金(61273194);江苏省自然科学基金(BK2012549);高等学校学科创新引智“111计划”(B12018)

作者简介

丁锋,男,博士,教授,博士生导师,主要从事系统辨识、过程建模、自适应控制方面的研究.fding@jiangnan.edu.cn

0 引言

系统建模(system modeling)和系统辨识(system identification)或系统参数估计是自动化控制理论的基础^[1-2].2011—2012年本文第一作者在《南京信息工程大学学报》连载了系统辨识方面的论文^[3-13],其部分内容汇集成书,涉及著作《系统辨识新论》^[1]与《系统辨识——辨识方法性能分析》^[2].从2014年起,笔者等又在该刊连载系统辨识方面的论文,涉及多元系统耦合多新息随机梯度类辨识方法^[14]、多元伪线性回归系统部分耦合多新息随机梯度类辨识方法^[15]、类多变量输出误差系统的耦合多新息辨识方法^[16]、多变量方程误差类系统的部分耦合迭代辨识方法^[17]、类多变量方程误差类系统的递阶多新息辨识方法^[18]等.本文主要研究规范型状态空间系统的状态估计与参数辨识问题.

状态空间模型不仅能方便地描述系统和外界输入的作用关系,也能揭示系统的内在结构及特性.状态空间模型在系统分析、系统控制、系统优化等方面具有独特的优点,如状态反馈极点配置、观测器设计、最优控制等.状态空间系统辨识一直是控制领域的研究热点.当系统状态可量测时,就很容易利用最小二乘方法辨识状态空间系统的参数;当系统状态不可测时,状态空间系统的辨识要复杂得多,因为既要辨识系统未知参数矩阵,还要估计未知系统状态,且辨识模型涉及未知状态与未知参数乘积的非线性关系,这使得研究状态空间系统的辨识方法更为困难^[19-20].

递阶辨识原理是研究复杂多变量系统辨识^[21-22]、非线性系统辨识^[23-24],以及多变量状态空间系统参数和状态联合估计的有效方法^[19-20].在状态空间系统辨识领域,文献[19]提出了多变量系统状态空间模型的递阶辨识方法,文献[20,25]提出了一般双率采样数据系统状态空间模型递阶辨识方法,以及能观测性规范型模型的递阶辨识方法,文献[26-27]研究了非均匀周期采样数据状态空间模型的递阶辨识及其连续系统的重构问题.最近,本文第一作者利用递阶辨识原理,提出了稀少量测数据系统的联合状态与参数估计方法^[28],推导了 Hammerstein 非线性系统的梯度迭代辨识方法与最小二乘迭代辨识方法^[23],提出了基于模型分解的 Hammerstein 系统递阶多新息随机梯度辨识方法(该文入选“2013年中国百篇最具影响国际学术论文”)^[24].

1 江南大学 物联网工程学院,无锡,214122

2 江南大学 控制科学与工程研究中心,无锡,214122

3 江南大学 教育部轻工过程先进控制重点实验室,无锡,214122

系统辨识算法可以用递推方案实现,也可以用迭代方案实现.对于实际系统,随着时间的推移,采集数据不断增加,不断利用新数据更新参数估计,即下一时刻的参数估计等于前一时刻的参数估计加上修正项,便是一种递推辨识算法.递推方案中的递推变量一般与真实世界中的时间有关,且每一步递推计算中都纳入了新量测的数据来更新参数估计.递推辨识算法的每一步计算量比较小,特别适合在线辨识 (online identification), 即实时参数估计 (real-time parameter estimation)^[1].

在迭代方案中,通常是先采集一批数据,利用批数据来刷新参数估计,特别适用于系统信息向量包含未知变量或者算法中含有未知参变量情形.迭代方案中迭代变量是自然数,与真实世界中的时间无关,且每一次迭代计算中没有纳入新数据,都是使用原来的一批数据,不断从这批量测数据中提取有用的信息来更新参数估计.迭代辨识算法只能用于离线辨识 (offline identification), 不适用于在线辨识^[1].

由于状态空间系统既包含未知参数,又包含未知状态,故需要利用辅助模型辨识思想,针对系统中的未知中间变量(这里指系统的状态),利用系统中可测到的输入输出数据,设计一个辅助模型,用辅助模型估算系统的状态.文中的辅助模型采用基于参数估计的观测器和基于参数估计的 Kalman 滤波状态估计算法.根据这种思想,针对观测器规范型状态空间系统,提出基于观测器或基于 Kalman 滤波状态估计的随机梯度辨识方法、多新息随机梯度辨识方法、递推最小二乘辨识方法、多新息最小二乘辨识方法.针对能观测性规范型状态空间系统,提出了基于观测器或基于 Kalman 滤波状态估计的梯度迭代辨识方法、最小二乘迭代辨识方法.当状态空间系统的维数很大时,最小二乘辨识算法的计算负担便体现出来.

由于最小二乘辨识算法涉及参数估计误差协方差阵,其计算量都是很大的,特别当参数数目很多时,算法的计算量更大.在这种情况下,可以采用递阶辨识原理,将分解的思想引入到辨识中,研究基于分解的辨识算法,减小计算量.辨识中的分解有两种,一种是基于矩阵分解的辨识算法,它是针对辨识算法中涉及的求逆矩阵,对在每步计算过程中求逆矩阵中不变的子矩阵与变化的子矩阵进行分解,利用分块矩阵求逆引理,来降低辨识算法的计算量,如

输出误差系统的最小二乘迭代辨识方法^[29];另一种是基于辨识模型分解的辨识算法,它是将辨识模型分解为一些子辨识模型,应用最小二乘优化原理,分别推导每个子模型的辨识算法,主要是通过分解,减小参数估计协方差阵的维数来达到降低计算量的目的,如 Box-Jenkins 系统两阶段递推最小二乘参数辨识方法^[30].

为减小状态空间系统最小二乘辨识算法的计算量,采用递阶辨识原理,将状态空间系统对应的辨识模型分解为两个子系统,提出了基于观测器或基于 Kalman 滤波状态估计的模型分解递推最小二乘辨识算法、模型分解多新息最小二乘辨识算法、模型分解最小二乘迭代辨识方法(即递阶最小二乘迭代辨识算法),顺便也讨论了基于观测器或基于 Kalman 滤波状态估计的模型分解随机梯度辨识算法、模型分解多新息随机梯度辨识算法、模型分解梯度迭代辨识算法.

1 基于状态估计的递推参数辨识算法

考虑观测器规范型状态空间系统^[31-32]:

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t), \quad (1)$$

$$y(t) = \mathbf{c}\mathbf{x}(t) + v(t), \quad (2)$$

其中 $\mathbf{x}(t) := [x_1(t), x_2(t), \dots, x_n(t)]^T \in \mathbf{R}^n$ 为系统状态向量, $u(t) \in \mathbf{R}$ 和 $y(t) \in \mathbf{R}$ 分别为系统的输入和输出, $v(t) \in \mathbf{R}$ 是均值为零的白噪声, $\mathbf{A} \in \mathbf{R}^{n \times n}$, $\mathbf{b} \in \mathbf{R}^n$, $\mathbf{c} \in \mathbf{R}^{1 \times n}$ 是系统的参数矩阵(向量):

$$\mathbf{A} := \begin{bmatrix} -a_1 & 1 & 0 & \cdots & 0 & 0 \\ -a_2 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ -a_{n-1} & 0 & 0 & \cdots & 0 & 1 \\ -a_n & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \in \mathbf{R}^{n \times n},$$

$$\mathbf{b} := \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \\ b_n \end{bmatrix} \in \mathbf{R}^n,$$

$$\mathbf{c} := [1, 0, 0, \dots, 0] \in \mathbf{R}^{1 \times n}.$$

由式(1)–(2)可得:

$$x_i(t+1) = -a_i x_1(t) + x_{i+1}(t) + b_i u(t), \quad (3)$$

$$i = 1, 2, \dots, n-1,$$

$$x_n(t+1) = -a_n x_1(t) + b_n u(t), \quad (4)$$

$$y(t) = x_1(t) + v(t). \quad (5)$$

式(3)两边同乘 z^{-i} , 得到

$$x_i(t-i+1) = -a_i x_i(t-i) + x_{i+1}(t-i) + b_i u(t-i),$$

$$i = 1, 2, \dots, n-1.$$

上式两边对 i 进行求和 $\sum_{i=1}^{n-1}$, 整理得到

$$x_1(t) = - \sum_{i=1}^{n-1} a_i x_1(t-i) + x_n(t-n+1) + \sum_{i=1}^{n-1} b_i u(t-i). \quad (6)$$

式(4)两边同乘以 z^{-n} , 得到

$$x_n(t-n+1) = -a_n x_n(t-n) + b_n u(t-n).$$

将上式代入式(6)得

$$x_1(t) = - \sum_{i=1}^n a_i x_1(t-i) + \sum_{i=1}^n b_i u(t-i). \quad (7)$$

定义参数向量 θ 和信息向量 $\varphi(t)$ 如下:

$$\theta := [a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_n]^T \in \mathbf{R}^{2n},$$

$$\varphi(t) := [-x_1(t-1), -x_1(t-2), \dots, -x_1(t-n),$$

$$u(t-1), u(t-2), \dots, u(t-n)]^T \in \mathbf{R}^{2n}.$$

由式(5)和式(7), 可以得到观测器规范型状态空间系统的辨识模型:

$$y(t) = x_1(t) + v(t) = \varphi^T(t)\theta + v(t). \quad (8)$$

从信息向量 $\varphi(t)$ 的定义可以看出, 随机梯度算法、递推最小二乘算法等无法用于状态不可测的状态空间系统参数向量 θ 的辨识, 因为信息向量 $\varphi(t)$ 中包含未知状态变量 $x_1(t-i)$. 解决方法是利用系统输入输出数据 $\{u(t), y(t): t=1, 2, 3, \dots\}$, 基于辅助模型辨识思想, 使用估计的状态向量 $\hat{x}(t)$ 代替信息向量 $\varphi(t)$ 中的未知状态变量, 用代替后的信息向量 $\hat{\varphi}(t)$ 研究和提出辨识状态空间系统的随机梯度算法、递推最小二乘算法、多新息随机梯度算法和多新息最小二乘算法等.

设 $\hat{x}(t) := [\hat{x}_1(t), \hat{x}_2(t), \dots, \hat{x}_n(t)]^T \in \mathbf{R}^n$ 是状态向量 $x(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T \in \mathbf{R}^n$ 的估计, 定义估计的信息向量, 即 $\varphi(t)$ 的估计:

$$\hat{\varphi}(t) := [-\hat{x}_1(t-1), -\hat{x}_1(t-2), \dots, -\hat{x}_1(t-n),$$

$$u(t-1), u(t-2), \dots, u(t-n)]^T \in \mathbf{R}^{2n}. \quad (9)$$

基于状态估计的参数辨识算法(或参数与状态联合估计算法)包含参数估计和状态估计两部分. 在参数估计算法中, 使用了估计的信息向量 $\hat{\varphi}(t)$, 而 $\hat{\varphi}(t)$ 涉及估计的状态 $\hat{x}(t)$; 在状态估计算法中, 即在状态观测器和 Kalman 滤波状态估计算法中, 是使用参数估计计算状态估计. 二者形成一个交互的计算过程. 下面分别讨论递推参数辨识算法和状态估计算法. 但要指出的是: 在参数辨识算法中使用了状

态估计, 同样在状态估计算法中使用了参数估计.

1.1 随机梯度辨识算法

令 $\hat{\theta}(t)$ 为参数向量 θ 在时刻 t 的估计, $\|X\|^2 := \text{tr}[XX^T]$ 表示矩阵(或向量) X 的范数. 定义梯度准则函数(gradient criterion function):

$$J_1(\theta) := \|y(t) - \varphi^T(t)\theta\|^2.$$

参考文献[1-2], 利用负梯度搜索, 极小化梯度准则函数 $J_1(\theta)$, 可以得到估计参数向量 θ 的随机梯度算法:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\varphi(t)}{r(t)} [y(t) - \varphi^T(t)\hat{\theta}(t-1)], \quad (10)$$

$$r(t) = r(t-1) + \|\varphi(t)\|^2, \quad r(0) = 1. \quad (11)$$

由于信息向量 $\varphi(t)$ 中存在不可测的状态变量 $x_1(t-i)$, 这使得算法(10)~(11)不可实现. 解决方法是用状态变量 $x_1(t-i)$ 的估计 $\hat{x}_1(t-i)$ 构造信息向量 $\varphi(t)$ 的估计, 参考式(9), 将式(10)~(11)中的 $\varphi(t)$ 用 $\hat{\varphi}(t)$ 替换, 可以得到估计参数向量 θ 的随机梯度算法(Stochastic Gradient algorithm, SG 算法):

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\hat{\varphi}(t)}{r(t)} [y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1)],$$

$$\hat{\theta}(0) = \mathbf{1}_{2n}/p_0, \quad (12)$$

$$r(t) = r(t-1) + \|\hat{\varphi}(t)\|^2, \quad r(0) = 1, \quad (13)$$

$$\hat{\varphi}(t) = [-\hat{x}_1(t-1), -\hat{x}_1(t-2), \dots, -\hat{x}_1(t-n),$$

$$u(t-1), u(t-2), \dots, u(t-n)]^T, \quad (14)$$

$$\hat{\theta}(t) = [\hat{a}_1(t), \hat{a}_2(t), \dots, \hat{a}_n(t), \hat{b}_1(t),$$

$$\hat{b}_2(t), \dots, \hat{b}_n(t)]^T. \quad (15)$$

这是参数估计算法, 算法中状态估计可采用观测器(46)~(49)或 Kalman 滤波状态估计算法(66)~(72)计算. 因此

1) 式(12)~(15)与式(46)~(49)构成了基于状态观测器的随机梯度参数辨识算法(O-SG 算法);

2) 式(12)~(15)与式(66)~(72)构成了基于 Kalman 滤波状态估计的随机梯度参数辨识算法(KF-SG 算法).

随机梯度辨识算法步骤简单、计算量小, 但存在参数估计收敛速度慢的问题, 这就引出了下面的多新息随机梯度辨识算法.

1.2 多新息随机梯度辨识算法

在辨识算法中, 式(12)中

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

称为新息(innovation). 这里 $e(t)$ 是标量, 即标量新息. 设新息长度为 p . 定义堆积输出向量 $Y(p, t)$ 和堆

积信息矩阵 $\hat{\Phi}(p, t)$ 如下:

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

$$\hat{\Phi}(p, t) := [\hat{\varphi}(t), \hat{\varphi}(t-1), \dots, \hat{\varphi}(t-p+1)] \in \mathbf{R}^{(2n) \times p}.$$

为提高随机梯度算法的收敛速度,将标量新息 $e(t)$ 扩展为新息向量:

$$E(p, t) := \begin{bmatrix} y(t) - \hat{\varphi}^T(t) \hat{\theta}(t-1) \\ y(t-1) - \hat{\varphi}^T(t-1) \hat{\theta}(t-1) \\ \vdots \\ y(t-p+1) - \hat{\varphi}^T(t-p+1) \hat{\theta}(t-1) \end{bmatrix} =$$

$$Y(p, t) - \hat{\Phi}^T(p, t) \hat{\theta}(t-1) \in \mathbf{R}^p.$$

参考文献[1, 33-34]中多新息随机梯度辨识算法的推导,从SG算法(12)–(15),可以得到多新息随机梯度算法(Multi-Innovation Stochastic Gradient algorithm, MISG算法):

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\hat{\Phi}(p, t)}{r(t)} E(p, t), \quad \hat{\theta}(0) = \mathbf{1}_{2n}/p_0, \quad (16)$$

$$E(p, t) = Y(p, t) - \hat{\Phi}^T(p, t) \hat{\theta}(t-1), \quad (17)$$

$$r(t) = r(t-1) + \|\hat{\varphi}(t)\|^2, \quad r(0) = 1, \quad (18)$$

$$Y(p, t) = [y(t), y(t-1), \dots, y(t-p+1)]^T, \quad (19)$$

$$\hat{\Phi}(p, t) = [\hat{\varphi}(t), \hat{\varphi}(t-1), \dots, \hat{\varphi}(t-p+1)], \quad (20)$$

$$\hat{\varphi}(t) = [-\hat{x}_1(t-1), -\hat{x}_1(t-2), \dots, -\hat{x}_1(t-n),$$

$$u(t-1), u(t-2), \dots, u(t-n)]^T, \quad (21)$$

$$\hat{\theta}(t) = [\hat{a}_1(t), \hat{a}_2(t), \dots, \hat{a}_n(t), \hat{b}_1(t),$$

$$\hat{b}_2(t), \dots, \hat{b}_n(t)]^T. \quad (22)$$

同样,算法中状态估计可采用观测器(46)–(49)或者Kalman滤波状态估计算法(66)–(72)计算.因此

1) 式(16)–(22)与式(46)–(49)构成了基于状态观测器的多新息随机梯度参数辨识算法(O-MISG算法);

2) 式(16)–(22)与式(66)–(72)构成了基于Kalman滤波状态估计的多新息随机梯度参数辨识算法(KF-MISG算法).

注1 式(18)中 $r(t)$ 使得算法有更快的收敛速度, $r(t)$ 也可修改为

$$r(t) = r(t-1) + \|\hat{\Phi}(p, t)\|^2, \quad r(0) = 1.$$

通过扩展新息长度,多新息随机梯度辨识算法能够改进随机梯度辨识算法的参数估计精度,理论分析表明当新息长度 $p=t$ 时,多新息参数估计逼近最小二乘参数估计.因为数据量是随时间 t 增加的,算法运行过程中的新息长度 p 一般是不变的(尽管可以动态调整),实践中 p 不可能不断增加以致趋于无穷,因此多新息随机梯度算法改进的参数估计精

度是有限的.在收敛速度与计算量方面,多新息随机梯度算法是随机梯度算法与最小二乘算法的折中.

下面介绍收敛速度快的递推最小二乘算法.

1.3 递推最小二乘辨识算法

定义输出向量 $Y(t)$ 和信息矩阵 $H(t)$ 如下:

$$Y(t) := \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(t) \end{bmatrix} \in \mathbf{R}^t,$$

$$H(t) := \begin{bmatrix} \varphi^T(1) \\ \varphi^T(2) \\ \vdots \\ \varphi^T(t) \end{bmatrix} \in \mathbf{R}^{t \times (2n)}.$$

定义和极小化最小二乘准则函数:

$$J_2(\theta) := \|Y(t) - H(t)\theta\|^2 =$$

$$[Y(t) - H(t)\theta]^T [Y(t) - H(t)\theta] =$$

$$\sum_{j=1}^t [y(j) - \varphi^T(j)\theta]^2. \quad (23)$$

假设 $[H^T(t)H(t)]$ 非奇异,可以得到估计参数向量 θ 的递推最小二乘算法:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + P(t)\varphi(t)[y(t) - \varphi^T(t)\hat{\theta}(t-1)], \quad (24)$$

$$P^{-1}(t) = P^{-1}(t-1) + \varphi(t)\varphi^T(t). \quad (25)$$

为避免求逆运算,将矩阵求逆引理^[1-2]

$$(A+BC)^{-1} = A^{-1} - A^{-1}B(I+CA^{-1}B)^{-1}CA^{-1}$$

应用于式(25)得到

$$P(t) = P(t-1) - \frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)}{1 + \varphi^T(t)P(t-1)\varphi(t)}. \quad (26)$$

定义增益向量 $L(t) := P(t)\varphi(t) \in \mathbf{R}^{2n}$, 得到

$$L(t) = P(t-1)\varphi(t) - \frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)\varphi(t)}{1 + \varphi^T(t)P(t-1)\varphi(t)} =$$

$$\frac{P(t-1)\varphi(t)}{1 + \varphi^T(t)P(t-1)\varphi(t)}. \quad (27)$$

由式(26)和式(27)求得:

$$P(t) = [I - L(t)\varphi^T(t)]P(t-1), \quad (28)$$

从而得到与式(24)–(25)等价的递推最小二乘辨识算法:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t)[y(t) - \varphi^T(t)\hat{\theta}(t-1)], \quad (29)$$

$$L(t) = \frac{P(t-1)\varphi(t)}{1 + \varphi^T(t)P(t-1)\varphi(t)}, \quad (30)$$

$$P(t) = [I - L(t)\varphi^T(t)]P(t-1). \quad (31)$$

针对信息向量 $\varphi(t)$ 包含未知状态变量的问题,同样采用其估计 $\hat{\varphi}(t)$ 代替,即用 $\hat{\varphi}(t)$ 替换式(29)–(31)中的 $\varphi(t)$,得到可实现的估计参数向量

θ 的递推最小二乘算法(Recursive Least Squares algorithm, RLS 算法):

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

$$\hat{\theta}(0) = \mathbf{1}_{2n}/p_0, \quad (32)$$

$$L(t) = P(t-1) \hat{\varphi}(t) [1 + \hat{\varphi}^T(t) P(t-1) \hat{\varphi}(t)]^{-1}, \quad (33)$$

$$P(t) = [I - L(t) \hat{\varphi}^T(t)] P(t-1), \quad P(0) = p_0 I_{2n}, \quad (34)$$

$$\hat{\varphi}(t) = [-\hat{x}_1(t-1), -\hat{x}_1(t-2), \dots, -\hat{x}_1(t-n),$$

$$u(t-1), u(t-2), \dots, u(t-n)]^T, \quad (35)$$

$$\hat{\theta}(t) = [\hat{a}_1(t), \hat{a}_2(t), \dots, \hat{a}_n(t), \hat{b}_1(t),$$

$$\hat{b}_2(t), \dots, \hat{b}_n(t)]^T. \quad (36)$$

算法中状态估计可采用观测器(46)—(49)或者 Kalman 滤波状态估计算法(66)—(72)计算. 因此

1) 式(32)—(36)与式(46)—(49)构成了基于状态观测器的递推最小二乘参数辨识算法(O-RLS 算法);

2) 式(32)—(36)与式(66)—(72)构成了基于 Kalman 滤波状态估计的递推最小二乘参数辨识算法(KF-RLS 算法).

1.4 多新息最小二乘辨识算法

借助于多新息辨识理论^[1,34-35], 扩展标量新息 $e(t) := y(t) - \hat{\varphi}^T(t) \hat{\theta}(t-1)$ 为新息向量:

$$E(p, t) := \begin{bmatrix} y(t) - \hat{\varphi}^T(t) \hat{\theta}(t-1) \\ y(t-1) - \hat{\varphi}^T(t-1) \hat{\theta}(t-1) \\ \vdots \\ y(t-p+1) - \hat{\varphi}^T(t-p+1) \hat{\theta}(t-1) \end{bmatrix} =$$

$$Y(p, t) - \hat{\Phi}^T(p, t) \hat{\theta}(t-1) \in \mathbf{R}^p,$$

其中堆积输出向量 $Y(p, t)$ 和堆积信息矩阵 $\hat{\Phi}(p, t)$ 定义为

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

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

因为 $E(1, t) = e(t)$, $Y(1, t) = y(t)$, $\hat{\Phi}(1, t) = \hat{\varphi}(t)$, 所以式(32)—(34)可以等价写为

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t) E(1, t),$$

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

$$L(t) = P(t-1) \hat{\Phi}(1, t) [1 + \hat{\Phi}^T(1, t) P(t-1) \hat{\Phi}(1, t)]^{-1},$$

$$P(t) = P(t-1) - L(t) \hat{\Phi}^T(1, t) P(t-1).$$

这就是新息长度 $p=1$ 的“多”新息最小二乘辨识算法. 参考文献[1,35], 把上式 $E(1, t)$, $Y(1, t)$ 和 $\hat{\Phi}(1, t)$ 中的“1”替换为 p , 可以得到新息长度为 p 、估计参数向量 θ 的多新息最小二乘算法(Multi-innovation Least Squares algorithm, MILS 算法):

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t) E(p, t), \quad \hat{\theta}(0) = \mathbf{1}_{2n}/p_0, \quad (37)$$

$$E(p, t) = Y(p, t) - \hat{\Phi}^T(p, t) \hat{\theta}(t-1), \quad (38)$$

$$L(t) = P(t-1) \hat{\Phi}(p, t) [I_p + \hat{\Phi}^T(p, t) P(t-1) \hat{\Phi}(p, t)]^{-1}, \quad (39)$$

$$P(t) = P(t-1) - L(t) \hat{\Phi}^T(p, t) P(t-1),$$

$$P(0) = p_0 I_{2n}, \quad (40)$$

$$Y(p, t) = [y(t), y(t-1), \dots, y(t-p+1)]^T, \quad (41)$$

$$\hat{\Phi}(p, t) = [\hat{\varphi}(t), \hat{\varphi}(t-1), \dots, \hat{\varphi}(t-p+1)], \quad (42)$$

$$\hat{\varphi}(t) = [-\hat{x}_1(t-1), -\hat{x}_1(t-2), \dots, -\hat{x}_1(t-n),$$

$$u(t-1), u(t-2), \dots, u(t-n)]^T, \quad (43)$$

$$\hat{\theta}(t) = [\hat{a}_1(t), \hat{a}_2(t), \dots, \hat{a}_n(t), \hat{b}_1(t),$$

$$\hat{b}_2(t), \dots, \hat{b}_n(t)]^T. \quad (44)$$

算法中状态估计可采用观测器(46)—(49)或者 Kalman 滤波状态估计算法(66)—(72)计算. 因此

1) 式(37)—(44)与式(46)—(49)构成了基于状态观测器的多新息最小二乘参数辨识算法(O-MILS 算法);

2) 式(37)—(44)与式(66)—(72)构成了基于 Kalman 滤波状态估计的多新息最小二乘参数辨识算法(KF-MILS 算法).

2 基于递推参数估计的状态估计算法

状态空间系统既包含未知参数, 又包含未知状态. 针对其辨识模型信息向量中存在不可测状态变量问题, 在前述的随机梯度辨识算法、递推最小二乘辨识算法、多新息随机梯度辨识算法和多新息最小二乘辨识算法中, 本文均采取了状态估计代替真实状态. 下面讨论基于递推参数估计的状态观测器和基于递推参数估计的 Kalman 滤波状态估计算法, 来估计系统的状态变量.

2.1 基于递推参数估计的状态观测器

状态观测器是根据系统的输入输出数据重构系统的状态, 对系统状态变量进行估计. 状态观测器的出现, 为实际中状态不可测系统的状态反馈提供了可能性, 能够应用于控制工程的许多方面.

为估计系统(1)—(2)的状态, 当参数已知时, 可以设计下列开环观测器估计系统状态:

$$\hat{\mathbf{x}}(t+1) = \mathbf{A} \hat{\mathbf{x}}(t) + \mathbf{b} u(t). \quad (45)$$

当系统参数 $\mathbf{A} \in \mathbf{R}^{n \times n}$, $\mathbf{b} \in \mathbf{R}^n$ 未知时, 分别用其时刻 t 估计值 $\hat{\mathbf{A}}(t)$ 和 $\hat{\mathbf{b}}(t)$ 代替, 得到基于递推参数估计的开环状态观测器(recursive parameter estimation based open-loop state observer):

$$\hat{\mathbf{x}}(t+1) = \hat{\mathbf{A}}(t) \hat{\mathbf{x}}(t) + \hat{\mathbf{b}}(t) u(t), \quad \hat{\mathbf{x}}(1) = \mathbf{1}_n/p_0, \quad (46)$$

$$\hat{\mathbf{A}}(t) = \begin{bmatrix} -\hat{a}_1(t) & 1 & 0 & \cdots & 0 \\ -\hat{a}_2 & 0 & 1 & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ -\hat{a}_{n-1}(t) & 0 & \cdots & 0 & 1 \\ -\hat{a}_n(t) & 0 & \cdots & 0 & 0 \end{bmatrix}, \quad (47)$$

$$\hat{\mathbf{b}}(t) = [\hat{b}_1(t), \hat{b}_2(t), \dots, \hat{b}_n(t)]^T, \quad (48)$$

$$\hat{\mathbf{x}}(t) = [\hat{x}_1(t), \hat{x}_2(t), \dots, \hat{x}_n(t)]^T. \quad (49)$$

当然,也可设计基于递推参数估计的闭环状态观测器(recursive parameter estimation based closed-loop state observer)^[28]:

$$\hat{\mathbf{x}}(t+1) = \hat{\mathbf{A}}(t)\hat{\mathbf{x}}(t) + \hat{\mathbf{b}}(t)u(t) + \mathbf{K}(t)[y(t) - \mathbf{c}\hat{\mathbf{x}}(t)],$$

$$\hat{\mathbf{x}}(1) = \mathbf{1}_n/p_0, \quad (50)$$

$$\hat{\mathbf{A}}(t) = \begin{bmatrix} -\hat{a}_1(t) & 1 & 0 & \cdots & 0 \\ -\hat{a}_2(t) & 0 & 1 & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ -\hat{a}_{n-1}(t) & 0 & \cdots & 0 & 1 \\ -\hat{a}_n(t) & 0 & \cdots & 0 & 0 \end{bmatrix}, \quad (51)$$

$$\hat{\mathbf{b}}(t) = [\hat{b}_1(t), \hat{b}_2(t), \dots, \hat{b}_n(t)]^T, \quad (52)$$

$$\hat{\mathbf{x}}(t) = [\hat{x}_1(t), \hat{x}_2(t), \dots, \hat{x}_n(t)]^T, \quad (53)$$

其中 $\mathbf{K}(t) \in \mathbf{R}^n$ 是观测器的增益向量,应该根据 $[\hat{\mathbf{A}}(t) - \mathbf{K}(t)\mathbf{c}]$ 的特征值在单位圆内进行选择.

将开环观测器或闭环状态观测器分别与前述的随机梯度辨识算法、多新息随机梯度辨识算法、递推最小二乘辨识算法和多新息最小二乘辨识算法相结合,构成联合参数与状态估计算法,可以同时估计系统的参数和状态,得到基于状态观测器的随机梯度辨识算法、基于状态观测器的多新息随机梯度辨识算法、基于状态观测器的递推最小二乘辨识算法以及基于状态观测器的多新息最小二乘辨识算法.

下面针对基于状态观测器的递推最小二乘辨识算法,给出其计算步骤、流程图和计算量.

2.2 基于状态观测器的递推最小二乘辨识算法

RLS 算法(32)~(36)和开环状态观测器(46)~(49)构成了基于状态观测器的递推最小二乘算法(Observer based Recursive Least Squares algorithm, O-RLS 算法),重写如下:

$$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t-1) + \mathbf{L}(t)[y(t) - \hat{\boldsymbol{\varphi}}^T(t)\hat{\boldsymbol{\theta}}(t-1)],$$

$$\hat{\boldsymbol{\theta}}(0) = \mathbf{1}_{2n}/p_0, \quad (54)$$

$$\mathbf{L}(t) = \mathbf{P}(t-1)\hat{\boldsymbol{\varphi}}(t)[1 + \hat{\boldsymbol{\varphi}}^T(t)\mathbf{P}(t-1)\hat{\boldsymbol{\varphi}}(t)]^{-1}, \quad (55)$$

$$\mathbf{P}(t) = \mathbf{P}(t-1) - \mathbf{L}(t)[\hat{\boldsymbol{\varphi}}^T(t)\mathbf{P}(t-1)],$$

$$\mathbf{P}(0) = p_0\mathbf{I}_{2n}, \quad (56)$$

$$\hat{\boldsymbol{\varphi}}(t) = [-\hat{x}_1(t-1), -\hat{x}_1(t-2), \dots, -\hat{x}_1(t-n),$$

$$u(t-1), u(t-2), \dots, u(t-n)]^T, \quad (57)$$

$$\hat{\mathbf{x}}(t+1) = \hat{\mathbf{A}}(t)\hat{\mathbf{x}}(t) + \hat{\mathbf{b}}(t)u(t), \quad \hat{\mathbf{x}}(1) = \mathbf{1}_n/p_0, \quad (58)$$

$$\hat{\mathbf{A}}(t) = \begin{bmatrix} -\hat{a}_1(t) & 1 & 0 & \cdots & 0 \\ -\hat{a}_2(t) & 0 & 1 & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ -\hat{a}_{n-1}(t) & 0 & \cdots & 0 & 1 \\ -\hat{a}_n(t) & 0 & \cdots & 0 & 0 \end{bmatrix}, \quad (59)$$

$$\hat{\mathbf{b}}(t) = [\hat{b}_1(t), \hat{b}_2(t), \dots, \hat{b}_n(t)]^T, \quad (60)$$

$$\hat{\mathbf{x}}(t) = [\hat{x}_1(t), \hat{x}_2(t), \dots, \hat{x}_n(t)]^T, \quad (61)$$

$$\hat{\boldsymbol{\theta}}(t) = [\hat{a}_1(t), \hat{a}_2(t), \dots, \hat{a}_n(t), \hat{b}_1(t),$$

$$\hat{b}_2(t), \dots, \hat{b}_n(t)]^T. \quad (62)$$

从 O-RLS 算法的时标(即递推时间次序)看,对于每一步递推计算,先计算参数估计,后计算状态估计,初值的设置也是这样.O-RLS 算法计算参数估计和状态估计的流程如图 1 所示.基于状态观测器的递推最小二乘辨识算法(54)~(62)计算步骤如下:

1) 初始化:给定参数估计精度 ε ,令 $t = 1$,
 $\hat{\boldsymbol{\theta}}(0) = \mathbf{1}_{2n}/p_0$, $\mathbf{P}(0) = p_0\mathbf{I}_{2n}$, $\hat{\mathbf{x}}(1) = \mathbf{1}_n/p_0$, $\hat{x}_1(i) = 0$,
 $i \leq 0$, $p_0 = 10^6$.

2) 采集输入数据 $u(t)$ 和输出数据 $y(t)$,用式(57)构造信息向量 $\hat{\boldsymbol{\varphi}}(t)$.

3) 用式(55)计算增益向量 $\mathbf{L}(t)$,用式(56)计算协方差矩阵 $\mathbf{P}(t)$.

4) 用式(54)刷新参数估计 $\hat{\boldsymbol{\theta}}(t)$.

5) 从式(62)的 $\hat{\boldsymbol{\theta}}(t)$ 中读出 $\hat{a}_i(t)$ 和 $\hat{b}_i(t)$,用式(59)和式(60)构造参数估计矩阵 $\hat{\mathbf{A}}(t)$ 和参数估计向量 $\hat{\mathbf{b}}(t)$.

6) 用式(58)计算 $\hat{\mathbf{x}}(t+1)$.

7) 比较 $\hat{\boldsymbol{\theta}}(t)$ 和 $\hat{\boldsymbol{\theta}}(t-1)$,若 $\|\hat{\boldsymbol{\theta}}(t) - \hat{\boldsymbol{\theta}}(t-1)\| \leq \varepsilon$,结束计算,得到参数估计 $\hat{\boldsymbol{\theta}}(t)$;否则 t 增 1,转到第 2)步.

辨识算法的计算量是评价计算效率的一个重要指标.辨识算法的计算量可用其乘法运算次数和加法运算次数表示.一次加法运算称为一个 flop,一次乘法运算也称为一个 flop.除法作为乘法对待,减法作为加法对待.这样就可以用 flop 数,即浮点运算数来表示计算量的大小.表 1 列出了基于状态观测器的递推最小二乘辨识算法的计算量.

2.3 基于递推参数估计的 Kalman 滤波状态估计算法

Kalman 滤波算法是状态观测器的扩展,它是假设系统参数已知时,估计随机系统状态变量的有效

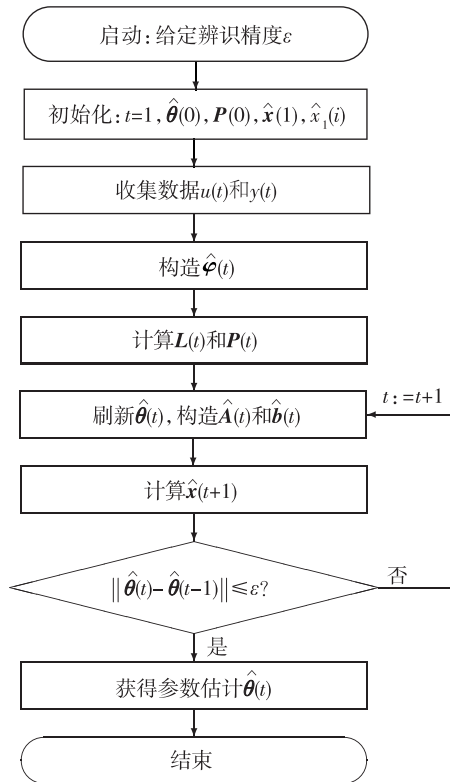


图1 计算 O-RLS 算法参数估计 $\hat{\theta}(t)$ 和状态估计 $\hat{x}(t+1)$

Fig. 1 The flowchart of computing the O-RLS parameter estimate $\hat{\theta}(t)$ and state estimate $\hat{x}(t+1)$

方法.对于系统参数辨识而言,系统参数是未知的,无法利用 Kalman 滤波方法估计系统状态.这里基于辨识算法提供的参数估计,利用 Kalman 滤波方法估计系统状态变量.

当系数矩阵(向量)已知时,估计系统(1)–(2)状态向量 $x(t)$ 的 Kalman 滤波状态估计算法可表示为

$$\begin{aligned} \hat{x}(t+1) &= A\hat{x}(t) + bu(t) + L(t)[y(t) - c\hat{x}(t)], \\ \hat{x}(1) &= \mathbf{1}_n/p_0, \end{aligned} \quad (63)$$

$$L(t) = AP(t)c^T [1 + cP(t)c^T]^{-1}, \quad (64)$$

$$P(t+1) = AP(t)A^T - L(t)cP(t)A^T, \quad P(1) = I_n. \quad (65)$$

当系统参数矩阵 $A \in \mathbf{R}^{n \times n}$ 和参数向量 $b \in \mathbf{R}^n$ 未知时,为了获得系统的状态估计,可用时刻 t 的递推参数估计 $\hat{A}(t)$ 和 $\hat{b}(t)$ 代替式(1)–(2)中 A 和 b ,得到基于递推参数估计的 Kalman 滤波状态估计算法^[28,36]:

$$\begin{aligned} \hat{x}(t+1) &= \hat{A}(t)\hat{x}(t) + \hat{b}(t)u(t) + L(t)[y(t) - c\hat{x}(t)], \\ \hat{x}(1) &= \mathbf{1}_n/p_0, \end{aligned} \quad (66)$$

$$L(t) = \hat{A}(t)P(t)c^T [1 + cP(t)c^T]^{-1}, \quad (67)$$

$$\begin{aligned} P(t+1) &= \hat{A}(t)P(t)\hat{A}^T(t) - L(t)cP(t)\hat{A}^T(t), \\ P(1) &= I_n, \end{aligned} \quad (68)$$

$$\hat{A}(t) = \begin{bmatrix} -\hat{a}_1(t) & 1 & 0 & \cdots & 0 \\ -\hat{a}_2(t) & 0 & 1 & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ -\hat{a}_{n-1}(t) & 0 & \cdots & 0 & 1 \\ -\hat{a}_n(t) & 0 & \cdots & 0 & 0 \end{bmatrix}, \quad (69)$$

$$\hat{b}(t) = [\hat{b}_1(t), \hat{b}_2(t), \dots, \hat{b}_n(t)]^T, \quad (70)$$

$$\hat{x}(t) = [\hat{x}_1(t), \hat{x}_2(t), \dots, \hat{x}_n(t)]^T, \quad (71)$$

$$\hat{\theta}(t) = [\hat{a}_1(t), \hat{a}_2(t), \dots, \hat{a}_n(t), \hat{b}_1(t), \hat{b}_2(t), \dots, \hat{b}_n(t)]^T. \quad (72)$$

递推参数估计算法可采用前述的一些辨识算法,构成基于递推参数估计的 Kalman 滤波状态估计算法.例如:

1) 式(12)–(15)与式(66)–(72)构成了基于随机梯度参数估计的 Kalman 滤波状态估计算法(SG-KF 算法);

2) 式(16)–(22)与式(66)–(72)构成了基于多新息随机梯度参数估计的 Kalman 滤波状态估计算法(MISG-KF 算法);

3) 式(32)–(36)与式(66)–(72)构成了基于

表1 O-RLS 算法的计算量

Table 1 The computational efficiency of the O-RLS algorithm

变量	表达式	乘法次数	加法次数
$\hat{\theta}(t)$	$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t)e(t) \in \mathbf{R}^{2n}$	$2n$	$2n$
	$e(t) := y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1) \in \mathbf{R}$	$2n$	$2n$
$L(t)$	$L(t) = \zeta(t) / [1 + \hat{\varphi}^T(t)\zeta(t)] \in \mathbf{R}^{2n}$	$4n$	$2n$
	$\zeta(t) := P(t-1)\hat{\varphi}(t) \in \mathbf{R}^{2n}$	$4n^2$	$4n^2 - 2n$
$P(t)$	$P(t) = P(t-1) - L(t)\zeta^T(t) \in \mathbf{R}^{2n \times 2n}$	$4n^2$	$4n^2$
$\hat{x}(t+1)$	$\hat{x}(t+1) = \hat{A}(t)\hat{x}(t) + \hat{b}u(t) \in \mathbf{R}^n$	$2n$	$2n-1$
总数		$8n^2 + 10n$	$8n^2 + 6n - 1$
总 flop 数		$16n^2 + 16n - 1$	

递推最小二乘参数估计的 Kalman 滤波状态估计算法(RLS-KF 算法);

4) 式(37)—(44)与式(66)—(72)构成了基于多新息最小二乘参数估计的 Kalman 滤波状态估计算法(MILS-KF 算法).

下面针对基于 Kalman 滤波状态估计的多新息最小二乘辨识算法,给出其计算步骤、流程图和计算量.

2.4 基于 Kalman 滤波状态估计的 MILS 辨识算法

MILS 算法(37)—(44)与 Kalman 滤波状态估计算法(66)—(72)构成了基于 Kalman 滤波状态估计的多新息最小二乘算法(Kalman Filtering based Multi-Innovation Least Squares algorithm, KF-MILS 算法),重写如下:

$$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t-1) + \mathbf{L}(t)\mathbf{E}(p, t), \quad \hat{\boldsymbol{\theta}}(0) = \mathbf{1}_{2n}/p_0, \quad (73)$$

$$\mathbf{E}(p, t) = \mathbf{Y}(p, t) - \hat{\boldsymbol{\Phi}}^T(p, t)\hat{\boldsymbol{\theta}}(t-1), \quad (74)$$

$$\mathbf{L}(t) = \mathbf{P}(t-1)\hat{\boldsymbol{\Phi}}(p, t)[\mathbf{I}_p + \hat{\boldsymbol{\Phi}}^T(p, t)\mathbf{P}(t-1)\hat{\boldsymbol{\Phi}}(p, t)]^{-1}, \quad (75)$$

$$\mathbf{P}(t) = \mathbf{P}(t-1) - \mathbf{L}(t)\hat{\boldsymbol{\Phi}}^T(p, t)\mathbf{P}(t-1), \quad \mathbf{P}(0) = p_0\mathbf{I}_{2n}, \quad (76)$$

$$\mathbf{Y}(p, t) = [y(t), y(t-1), \dots, y(t-p+1)]^T, \quad (77)$$

$$\hat{\boldsymbol{\Phi}}(p, t) = [\hat{\boldsymbol{\varphi}}(t), \hat{\boldsymbol{\varphi}}(t-1), \dots, \hat{\boldsymbol{\varphi}}(t-p+1)], \quad (78)$$

$$\hat{\boldsymbol{\varphi}}(t) = [-\hat{x}_1(t-1), -\hat{x}_1(t-2), \dots, -\hat{x}_1(t-n), u(t-1), u(t-2), \dots, u(t-n)]^T, \quad (79)$$

$$\hat{\mathbf{x}}(t+1) = \hat{\mathbf{A}}(t)\hat{\mathbf{x}}(t) + \hat{\mathbf{b}}(t)u(t) + \mathbf{L}_1(t)[y(t) - \mathbf{c}\hat{\mathbf{x}}(t)], \quad \hat{\mathbf{x}}(1) = \mathbf{1}_n/p_0, \quad (80)$$

$$\mathbf{L}_1(t) = \hat{\mathbf{A}}(t)\mathbf{P}_1(t)\mathbf{c}^T[1 + \mathbf{c}\mathbf{P}_1(t)\mathbf{c}^T]^{-1}, \quad (81)$$

$$\mathbf{P}_1(t+1) = \hat{\mathbf{A}}(t)\mathbf{P}_1(t)\hat{\mathbf{A}}^T(t) - \mathbf{L}_1(t)\mathbf{c}\mathbf{P}_1(t)\hat{\mathbf{A}}^T(t), \quad \mathbf{P}_1(1) = \mathbf{I}_n, \quad (82)$$

$$\hat{\mathbf{A}}(t) = \begin{bmatrix} -\hat{a}_1(t) & 1 & 0 & \cdots & 0 \\ -\hat{a}_2(t) & 0 & 1 & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ -\hat{a}_{n-1}(t) & 0 & \cdots & 0 & 1 \\ -\hat{a}_n(t) & 0 & \cdots & 0 & 0 \end{bmatrix}, \quad (83)$$

$$\hat{\mathbf{b}}(t) = [\hat{b}_1(t), \hat{b}_2(t), \dots, \hat{b}_n(t)]^T, \quad (84)$$

$$\hat{\mathbf{x}}(t) = [\hat{x}_1(t), \hat{x}_2(t), \dots, \hat{x}_n(t)]^T, \quad (85)$$

$$\hat{\boldsymbol{\theta}}(t) = [\hat{a}_1(t), \hat{a}_2(t), \dots, \hat{a}_n(t), \hat{b}_1(t), \hat{b}_2(t), \dots, \hat{b}_n(t)]^T. \quad (86)$$

基于 Kalman 滤波状态估计的多新息最小二乘辨识算法(73)—(86)计算步骤如下:

1) 确定新息长度 p , 给定参数估计精度 ε .

2) 置初值: 令 $t=1$, $\hat{\boldsymbol{\theta}}(0) = \mathbf{1}_{2n}/p_0$, $\mathbf{P}(0) = p_0\mathbf{I}_{2n}$, $\hat{\mathbf{x}}(1) = \mathbf{1}_n/p_0$, $\mathbf{P}_1(1) = \mathbf{I}_n$, $\hat{x}_i(i) = 0, i \leq 0, p_0 = 10^6$.

3) 收集输入输出数据 $u(t)$ 和 $y(t)$, 用式(77)构造 $\mathbf{Y}(p, t)$.

4) 根据式(79)构造 $\hat{\boldsymbol{\varphi}}(t)$, 式(78)构造 $\hat{\boldsymbol{\Phi}}(p, t)$.

5) 由式(75)计算 $\mathbf{L}(t)$, 用式(76)计算 $\mathbf{P}(t)$.

6) 用式(74)计算 $\mathbf{E}(p, t)$, 用式(73)刷新参数估计 $\hat{\boldsymbol{\theta}}(t)$.

7) 从式(86)的 $\hat{\boldsymbol{\theta}}(t)$ 中读出 $\hat{a}_i(t)$ 和 $\hat{b}_i(t)$, 根据式(83)和(84)分别构造 $\hat{\mathbf{A}}(t)$ 和 $\hat{\mathbf{b}}(t)$.

8) 用式(81)计算 $\mathbf{L}_1(t)$, 用式(82)计算 $\mathbf{P}_1(t+1)$, 用式(80)计算 $\hat{\mathbf{x}}(t+1)$.

9) 比较 $\hat{\boldsymbol{\theta}}(t)$ 与 $\hat{\boldsymbol{\theta}}(t-1)$: 如果 $\|\hat{\boldsymbol{\theta}}(t) - \hat{\boldsymbol{\theta}}(t-1)\| \leq \varepsilon$, 就结束循环, 获得估计参数 $\hat{\boldsymbol{\theta}}(t)$; 反之, 则 t 增 1, 转到第 3) 步.

基于 Kalman 滤波状态估计的多新息最小二乘辨识算法(73)—(86)计算参数估计和状态估计流程如图 2 所示, 计算量如表 2 所示. 值得指出的是, 与多新息随机梯度算法相比, 多新息最小二乘算法改进参数估计的能力是极其有限的, 因为最小二乘算

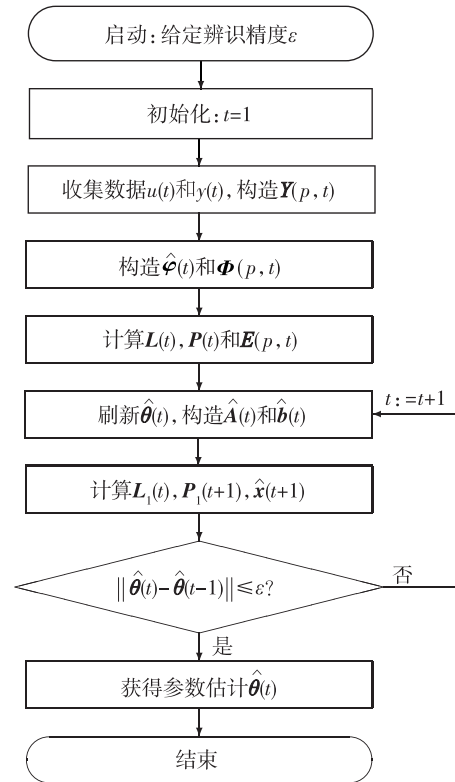


图 2 计算 KF-MILS 参数估计 $\hat{\boldsymbol{\theta}}(t)$ 和状态估计 $\hat{\mathbf{x}}(t+1)$ 流程
Fig. 2 The flowchart of computing the KF-MILS parameter estimate $\hat{\boldsymbol{\theta}}(t)$ and state estimate $\hat{\mathbf{x}}(t+1)$

表 2 KF-MILS 算法的计算量

Table 2 The computational efficiency of the KF-MILS algorithm

表达式	乘法次数	加法次数
$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t-1) + \mathbf{L}(t)\mathbf{E}(p, t) \in \mathbf{R}^{2n}$	$2np$	$2np$
$\mathbf{E}(p, t) = \mathbf{Y}(p, t) - \hat{\boldsymbol{\Phi}}^T(p, t)\hat{\boldsymbol{\theta}}(t-1) \in \mathbf{R}^p$	$2np$	$2np$
$\mathbf{L}(t) = \boldsymbol{\zeta}(t) [\mathbf{I}_p + \hat{\boldsymbol{\Phi}}^T(p, t)\boldsymbol{\zeta}(t)]^{-1} \in \mathbf{R}^{2n \times p}$	$4np^2 + p^3$	$4np^2 + p^3 - 2p^2 + p - 2np$
$\boldsymbol{\zeta}(t) := \mathbf{P}(t-1)\hat{\boldsymbol{\Phi}}(p, t) \in \mathbf{R}^{2n \times p}$	$4n^2p$	$4n^2p - 2np$
$\mathbf{P}(t) = \mathbf{P}(t-1) - \mathbf{L}(t)\boldsymbol{\zeta}^T(t) \in \mathbf{R}^{2n \times 2n}$	$4n^2p$	$4n^2p$
$\hat{\mathbf{x}}(t+1) = \hat{\mathbf{A}}(t)\hat{\mathbf{x}}(t) + \hat{\mathbf{b}}(t)u(t) + \mathbf{L}_1(t)[y(t) - \mathbf{c}\hat{\mathbf{x}}(t)] \in \mathbf{R}^n$	$2n$	$2n-1$
$\mathbf{L}_1(t) = \hat{\mathbf{A}}(t)[\mathbf{P}_1(t)\mathbf{c}^T] / [1 + \mathbf{c}\mathbf{P}_1(t)\mathbf{c}^T] \in \mathbf{R}^n$	n	$n+1$
$\mathbf{L}_1(t) = \hat{\mathbf{A}}(t)[\mathbf{P}_1(t)\mathbf{c}^T] / [1 + \mathbf{c}\mathbf{P}_1(t)\mathbf{c}^T] \in \mathbf{R}^n$	$2n$	n
$\mathbf{P}_1(t+1) = \hat{\mathbf{A}}(t)\boldsymbol{\zeta}_1(t) - \mathbf{L}_1(t)\mathbf{c}\boldsymbol{\zeta}_1(t) \in \mathbf{R}^{n \times n}$	$n^2 + 2n$	$n^2 + n - 1$
$\boldsymbol{\zeta}_1(t) := \mathbf{P}_1(t)\hat{\mathbf{A}}^T(t) \in \mathbf{R}^{n \times n}$	n^2	$n^2 - n$
总数	$8n^2p + 4np^2 + 4np + p^3 + 2n^2 + 7n$	$8n^2p + 4np^2 + p^3 - 2p^2 + p + 2n^2 + 3n - 1$
总 flop 数	$16n^2p + 8np^2 + 4np + 2p^3 - 2p^2 + p + 4n^2 + 10n - 1$	

法从数据提取有用信息的能力极其强大,但是变递推间隔多新息最小二乘算法 (V-MILS) 能提高稀疏数据系统和损失数据系统的参数估计精度^[1,35].

3 基于模型分解和状态估计的递推辨识算法

对于高维数状态空间系统,辨识算法中无论是矩阵乘积运算还是求逆运算,特别对于最小二乘辨识算法涉及的参数估计误差协方差阵,其计算量都是很大的,严重影响计算效率.对于高维系统,即对于参数数目多的大规模线性系统和结构复杂的非线性系统,可以采用递阶辨识原理,将分解的思想引入到辨识中,研究基于分解的辨识算法,减小辨识算法的计算量.特别是与递推最小二乘算法相比,递阶最小二乘算法的计算量要小得多.因此,有必要在保持辨识精度或辨识精度不至于牺牲太大的条件下,采用递阶辨识原理,研究计算量小的辨识方法,提高计算效率.

下面分别讨论递推参数辨识算法和状态估计算法.但要指出的是:在参数辨识算法中使用了状态估计,同样在状态估计算法中使用了参数估计.

3.1 基于模型分解的随机梯度辨识算法

基于模型分解的随机梯度辨识算法也称为递阶随机梯度辨识算法.定义参数向量 $\boldsymbol{\theta}$ 和信息向量 $\boldsymbol{\varphi}(t)$ 如下:

$$\boldsymbol{\theta} := \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} \in \mathbf{R}^{2n}, \quad \boldsymbol{\varphi}(t) := \begin{bmatrix} \boldsymbol{\phi}(t) \\ \boldsymbol{\psi}(t) \end{bmatrix} \in \mathbf{R}^{2n},$$

$$\mathbf{a} := [a_1, a_2, \dots, a_n]^T \in \mathbf{R}^n,$$

$$\mathbf{b} := [b_1, b_2, \dots, b_n]^T \in \mathbf{R}^n,$$

$$\boldsymbol{\phi}(t) := [-x_1(t-1), -x_1(t-2), \dots, -x_1(t-n)]^T \in \mathbf{R}^n,$$

$$\boldsymbol{\psi}(t) := [u(t-1), u(t-2), \dots, u(t-n)]^T \in \mathbf{R}^n,$$

则系统(1)–(2)的辨识模型(8)可改写为

$$y(t) = \boldsymbol{\phi}^T(t)\mathbf{a} + \boldsymbol{\psi}^T(t)\mathbf{b} + v(t) = \boldsymbol{\varphi}^T(t)\boldsymbol{\theta} + v(t). \quad (87)$$

定义中间变量:

$$y_1(t) := y(t) - \boldsymbol{\psi}^T(t)\mathbf{b} \in \mathbf{R}, \quad (88)$$

$$y_2(t) := y(t) - \boldsymbol{\phi}^T(t)\mathbf{a} \in \mathbf{R}. \quad (89)$$

利用式(88)和式(89),可以将辨识模型(87)分解为如下两个子模型:

$$y_1(t) = \boldsymbol{\phi}^T(t)\mathbf{a} + v(t), \quad (90)$$

$$y_2(t) = \boldsymbol{\psi}^T(t)\mathbf{b} + v(t). \quad (91)$$

定义准则函数:

$$J_3(\mathbf{a}) := \|y_1(t) - \boldsymbol{\phi}^T(t)\mathbf{a}\|^2,$$

$$J_4(\mathbf{b}) := \|y_2(t) - \boldsymbol{\psi}^T(t)\mathbf{b}\|^2.$$

设参数向量 \mathbf{a} 和 \mathbf{b} 在时刻 t 的估计是 $\hat{\mathbf{a}}(t)$ 和 $\hat{\mathbf{b}}(t)$.使用负梯度搜索,极小化准则函数 $J_3(\mathbf{a})$ 和 $J_4(\mathbf{b})$,可以得到下列递推关系:

$$\hat{\mathbf{a}}(t) = \hat{\mathbf{a}}(t-1) + \frac{\boldsymbol{\phi}(t)}{r_1(t)} [y_1(t) - \boldsymbol{\phi}^T(t)\hat{\mathbf{a}}(t-1)] =$$

$$\hat{\mathbf{a}}(t-1) + \frac{\boldsymbol{\phi}(t)}{r_1(t)} [y(t) - \boldsymbol{\psi}^T(t)\mathbf{b} - \boldsymbol{\phi}^T(t)\hat{\mathbf{a}}(t-1)], \quad (92)$$

$$\hat{\mathbf{b}}(t) = \hat{\mathbf{b}}(t-1) + \frac{\boldsymbol{\psi}(t)}{r_2(t)} [y_2(t) - \boldsymbol{\psi}^T(t)\hat{\mathbf{b}}(t-1)] =$$

$$\hat{\mathbf{b}}(t-1) + \frac{\boldsymbol{\psi}(t)}{r_2(t)} [y(t) - \boldsymbol{\phi}^T(t)\mathbf{a} - \boldsymbol{\psi}^T(t)\hat{\mathbf{b}}(t-1)], \quad (93)$$

$$r_1(t) = r_1(t-1) + \|\boldsymbol{\phi}(t)\|^2, \quad r_1(0) = 1, \quad (94)$$

$$r_2(t) = r_2(t-1) + \|\boldsymbol{\psi}(t)\|^2, \quad r_2(0) = 1. \quad (95)$$

这个算法不能实现,因为式(92)–(95)右边包含了未知信息向量 $\boldsymbol{\phi}(t)$ 和未知参数向量 \mathbf{b} 和 \mathbf{a} . 参考随机梯度辨识算法(12)–(15)的推导,利用递阶辨识原理,未知信息向量 $\boldsymbol{\phi}(t)$ 用其估计 $\hat{\boldsymbol{\phi}}(t)$ 代替,未知参数向量 \mathbf{b} 和 \mathbf{a} 用其在前一时刻的估计 $\hat{\mathbf{b}}(t-1)$ 和 $\hat{\mathbf{a}}(t-1)$ 代替,便可以得到计算参数估计 $\hat{\mathbf{a}}(t)$ 和 $\hat{\mathbf{b}}(t)$ 的基于模型分解的随机梯度算法(model Decomposition based Stochastic Gradient algorithm, D-SG 算法):

$$\hat{\mathbf{a}}(t) = \hat{\mathbf{a}}(t-1) + \frac{\hat{\boldsymbol{\phi}}(t)}{r_1(t)} [y(t) - \hat{\boldsymbol{\phi}}^T(t)\hat{\mathbf{a}}(t-1) - \boldsymbol{\psi}^T(t)\hat{\mathbf{b}}(t-1)],$$

$$\hat{\mathbf{a}}(0) = \mathbf{1}_n/p_0, \quad (96)$$

$$\hat{\mathbf{b}}(t) = \hat{\mathbf{b}}(t-1) + \frac{\boldsymbol{\psi}(t)}{r_2(t)} [y(t) - \hat{\boldsymbol{\phi}}^T(t)\hat{\mathbf{a}}(t-1) - \boldsymbol{\psi}^T(t)\hat{\mathbf{b}}(t-1)],$$

$$\hat{\mathbf{b}}(0) = \mathbf{1}_n/p_0, \quad (97)$$

$$r_1(t) = r_1(t-1) + \|\hat{\boldsymbol{\phi}}(t)\|^2, \quad r_1(0) = 1, \quad (98)$$

$$r_2(t) = r_2(t-1) + \|\boldsymbol{\psi}(t)\|^2, \quad r_2(0) = 1, \quad (99)$$

$$\hat{\boldsymbol{\phi}}(t) = [-\hat{x}_1(t-1), -\hat{x}_1(t-2), \dots, -\hat{x}_1(t-n)]^T, \quad (100)$$

$$\boldsymbol{\psi}(t) = [u(t-1), u(t-2), \dots, u(t-n)]^T, \quad (101)$$

$$\hat{\mathbf{a}}(t) = [\hat{a}_1(t), \hat{a}_2(t), \dots, \hat{a}_n(t)]^T, \quad (102)$$

$$\hat{\mathbf{b}}(t) = [\hat{b}_1(t), \hat{b}_2(t), \dots, \hat{b}_n(t)]^T. \quad (103)$$

这是参数估计算法,算法中状态估计可采用观测器(46)–(49)或 Kalman 滤波状态估计算法(66)–(72)计算.因此

1) 式(96)–(103)与式(46)–(49)构成了基于状态观测器的模型分解随机梯度参数辨识算法(O-D-SG 算法);

2) 式(96)–(103)与式(66)–(72)构成了基于 Kalman 滤波状态估计的模型分解随机梯度参数辨识算法(KF-D-SG 算法).

模型分解使得信息向量和参数向量的维数均从 $2n$ 降为 n ,有效减少的是矩阵范数运算量,但因为随机梯度辨识算法本身计算量较小,在这里模型分解的优势并未得到充分体现.下面讨论基于模型分解的多新息随机梯度辨识算法、基于模型分解的递推最小二乘辨识算法、基于模型分解的多新息最小二乘辨识算法.

3.2 基于模型分解的多新息随机梯度辨识算法

基于模型分解的多新息随机梯度辨识算法也称为递阶多新息随机梯度辨识算法^[24,30].令新息长度

为 p ,定义堆积信息矩阵 $\hat{\boldsymbol{\Phi}}(p,t)$ 和 $\boldsymbol{\Psi}(p,t)$ 如下:

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

$$\boldsymbol{\Psi}(p,t) := [\boldsymbol{\psi}(t), \boldsymbol{\psi}(t-1), \dots, \boldsymbol{\psi}(t-p+1)] \in \mathbf{R}^{n \times p}.$$

将式(96)或(97)中辨识新息(标量新息)

$$e(t) := y(t) - \hat{\boldsymbol{\phi}}^T(t)\hat{\mathbf{a}}(t-1) - \boldsymbol{\psi}^T(t)\hat{\mathbf{b}}(t-1) \in \mathbf{R}$$

推广为新息向量

$$\mathbf{E}(p,t) = \mathbf{Y}(p,t) - \hat{\boldsymbol{\Phi}}^T(p,t)\hat{\mathbf{a}}(t-1) - \boldsymbol{\Psi}^T(p,t)\hat{\mathbf{b}}(t-1) \in \mathbf{R}^p.$$

参考多新息随机梯度辨识算法(16)–(22)的推导,可以得到计算参数估计向量 $\hat{\mathbf{a}}(t)$ 和 $\hat{\mathbf{b}}(t)$ 的基于模型分解的多新息随机梯度算法(model Decomposition based Multi-Innovation Stochastic Gradient algorithm, D-MISG 算法):

$$\hat{\mathbf{a}}(t) = \hat{\mathbf{a}}(t-1) + \frac{\hat{\boldsymbol{\Phi}}(p,t)}{r_1(t)} \mathbf{E}(p,t), \quad \hat{\mathbf{a}}(0) = \mathbf{1}_n/p_0, \quad (104)$$

$$\hat{\mathbf{b}}(t) = \hat{\mathbf{b}}(t-1) + \frac{\boldsymbol{\Psi}(p,t)}{r_2(t)} \mathbf{E}(p,t), \quad \hat{\mathbf{b}}(0) = \mathbf{1}_n/p_0, \quad (105)$$

$$\mathbf{E}(p,t) = \mathbf{Y}(p,t) - \hat{\boldsymbol{\Phi}}^T(p,t)\hat{\mathbf{a}}(t-1) - \boldsymbol{\Psi}^T(p,t)\hat{\mathbf{b}}(t-1), \quad (106)$$

$$r_1(t) = r_1(t-1) + \|\hat{\boldsymbol{\phi}}(t)\|^2, \quad r_1(0) = 1, \quad (107)$$

$$r_2(t) = r_2(t-1) + \|\boldsymbol{\psi}(t)\|^2, \quad r_2(0) = 1, \quad (108)$$

$$\mathbf{Y}(p,t) = [y(t), y(t-1), \dots, y(t-p+1)]^T, \quad (109)$$

$$\hat{\boldsymbol{\Phi}}(p,t) = [\hat{\boldsymbol{\phi}}(t), \hat{\boldsymbol{\phi}}(t-1), \dots, \hat{\boldsymbol{\phi}}(t-p+1)], \quad (110)$$

$$\boldsymbol{\Psi}(p,t) = [\boldsymbol{\psi}(t), \boldsymbol{\psi}(t-1), \dots, \boldsymbol{\psi}(t-p+1)], \quad (111)$$

$$\hat{\boldsymbol{\phi}}(t) = [-\hat{x}_1(t-1), -\hat{x}_1(t-2), \dots, -\hat{x}_1(t-n)]^T, \quad (112)$$

$$\boldsymbol{\psi}(t) = [u(t-1), u(t-2), \dots, u(t-n)]^T, \quad (113)$$

$$\hat{\mathbf{a}}(t) = [\hat{a}_1(t), \hat{a}_2(t), \dots, \hat{a}_n(t)]^T, \quad (114)$$

$$\hat{\mathbf{b}}(t) = [\hat{b}_1(t), \hat{b}_2(t), \dots, \hat{b}_n(t)]^T. \quad (115)$$

式(107)–(108)可以修改为

$$r_1(t) = r_1(t-1) + \|\hat{\boldsymbol{\Phi}}(p,t)\|^2, \quad r_1(0) = 1, \quad (116)$$

$$r_2(t) = r_2(t-1) + \|\boldsymbol{\Psi}(p,t)\|^2, \quad r_2(0) = 1. \quad (117)$$

算法中使用的状态估计可采用观测器(46)–(49)或者 Kalman 滤波状态估计算法(66)–(72)计算.因此

1) 式(104)–(115)与式(46)–(49)构成了基于状态观测器的 D-MISG 参数辨识算法(O-D-MISG 算法);

2) 式(104)–(115)与式(66)–(72)构成了基于 Kalman 滤波状态估计的 D-MISG 参数辨识算法(KF-D-MISG 算法).

3.3 基于模型分解的递推最小二乘辨识算法

基于模型分解的递推最小二乘辨识算法也称为递阶最小二乘辨识算法^[37].定义子系统的输出向量 $\mathbf{Y}_1(t)$ 和 $\mathbf{Y}_2(t)$,以及信息矩阵 $\boldsymbol{\Phi}(t)$ 和 $\boldsymbol{\Psi}(t)$ 如下:

$$Y_1(t) := [y_1(1), y_1(2), \dots, y_1(t)]^T \in \mathbf{R}^t,$$

$$Y_2(t) := [y_2(1), y_2(2), \dots, y_2(t)]^T \in \mathbf{R}^t,$$

$$\Phi(t) := [\phi(1), \phi(2), \dots, \phi(t)]^T \in \mathbf{R}^{t \times n},$$

$$\Psi(t) := [\psi(1), \psi(2), \dots, \psi(t)]^T \in \mathbf{R}^{t \times n},$$

定义和极小化两个子模型(90)–(91)的最小二乘准则函数:

$$J_5(\mathbf{a}) := \|Y_1(t) - \Phi(t)\mathbf{a}\|^2,$$

$$J_6(\mathbf{b}) := \|Y_2(t) - \Psi(t)\mathbf{b}\|^2.$$

参考递推最小二乘辨识算法(32)–(36),利用式(88)–(89),可以得到计算参数估计向量 $\hat{\mathbf{a}}(t)$ 和 $\hat{\mathbf{b}}(t)$ 的递推最小二乘辨识算法:

$$\hat{\mathbf{a}}(t) = \hat{\mathbf{a}}(t-1) + L_1(t) [y_1(t) - \phi^T(t)\hat{\mathbf{a}}(t-1)] = \hat{\mathbf{a}}(t-1) + L_1(t) [y_1(t) - \psi^T(t)\mathbf{b} - \phi^T(t)\hat{\mathbf{a}}(t-1)], \quad (118)$$

$$L_1(t) = P_1(t-1)\phi(t) [1 + \phi^T(t)P_1(t-1)\phi(t)]^{-1}, \quad (119)$$

$$P_1(t) = [I - L_1(t)\phi^T(t)]P_1(t-1), \quad (120)$$

$$\hat{\mathbf{b}}(t) = \hat{\mathbf{b}}(t-1) + L_2(t) [y_2(t) - \psi^T(t)\hat{\mathbf{b}}(t-1)] = \hat{\mathbf{b}}(t-1) + L_2(t) [y_2(t) - \phi^T(t)\mathbf{a} - \psi^T(t)\hat{\mathbf{b}}(t-1)], \quad (121)$$

$$L_2(t) = P_2(t-1)\psi(t) [1 + \psi^T(t)P_2(t-1)\psi(t)]^{-1}, \quad (122)$$

$$P_2(t) = [I - L_2(t)\psi^T(t)]P_2(t-1). \quad (123)$$

同样这个算法的右边包含了未知信息向量 $\phi(t)$, 未知参数向量 \mathbf{b} 和 \mathbf{a} , 利用递阶辨识原理, 未知信息向量 $\phi(t)$ 用其估计 $\hat{\phi}(t)$ 代替, 未知参数向量 \mathbf{b} 和 \mathbf{a} 用其在前一时刻的估计 $\hat{\mathbf{b}}(t-1)$ 和 $\hat{\mathbf{a}}(t-1)$ 代替, 可以得到计算参数估计向量 $\hat{\mathbf{a}}(t)$ 和 $\hat{\mathbf{b}}(t)$ 的基于模型分解的递推最小二乘算法 (model Decomposition based Recursive Least Squares algorithm, D-RLS 算法):

$$\hat{\mathbf{a}}(t) = \hat{\mathbf{a}}(t-1) + L_1(t) [y_1(t) - \psi^T(t)\hat{\mathbf{b}}(t-1) - \hat{\phi}^T(t)\hat{\mathbf{a}}(t-1)], \quad \hat{\mathbf{a}}(0) = \mathbf{1}_n/p_0, \quad (124)$$

$$L_1(t) = P_1(t-1)\hat{\phi}(t) [1 + \hat{\phi}^T(t)P_1(t-1)\hat{\phi}(t)]^{-1}, \quad (125)$$

$$P_1(t) = [I - L_1(t)\hat{\phi}^T(t)]P_1(t-1), \quad P_1(0) = p_0\mathbf{I}_n, \quad (126)$$

$$\hat{\mathbf{b}}(t) = \hat{\mathbf{b}}(t-1) + L_2(t) [y_2(t) - \hat{\phi}^T(t)\hat{\mathbf{a}}(t-1) - \psi^T(t)\hat{\mathbf{b}}(t-1)], \quad \hat{\mathbf{b}}(0) = \mathbf{1}_n/p_0, \quad (127)$$

$$L_2(t) = P_2(t-1)\psi(t) [1 + \psi^T(t)P_2(t-1)\psi(t)]^{-1}, \quad (128)$$

$$P_2(t) = [I - L_2(t)\psi^T(t)]P_2(t-1), \quad P_2(0) = p_0\mathbf{I}_n, \quad (129)$$

$$\hat{\phi}(t) = [-\hat{x}_1(t-1), -\hat{x}_1(t-2), \dots, -\hat{x}_1(t-n)]^T, \quad (130)$$

$$\psi(t) = [u(t-1), u(t-2), \dots, u(t-n)]^T, \quad (131)$$

$$\hat{\mathbf{a}}(t) = [\hat{a}_1(t), \hat{a}_2(t), \dots, \hat{a}_n(t)]^T, \quad (132)$$

$$\hat{\mathbf{b}}(t) = [\hat{b}_1(t), \hat{b}_2(t), \dots, \hat{b}_n(t)]^T. \quad (133)$$

算法中使用的状态估计可采用观测器(46)–(49)或者 Kalman 滤波状态估计算法(66)–(72)计算. 因此

1) 式(124)–(133)与式(46)–(49)构成了基于状态观测器的模型分解递推最小二乘参数辨识算法(O-D-RLS 算法);

2) 式(124)–(133)与式(66)–(72)构成了基于 Kalman 滤波状态估计的模型分解递推最小二乘参数辨识算法(KF-D-RLS 算法).

表3列出了基于模型分解的递推最小二乘辨识算法(124)–(133)的计算量.

3.4 基于模型分解的多新息最小二乘辨识算法

基于模型分解的多新息最小二乘辨识算法也称为递阶多新息最小二乘辨识算法^[18]. 令新息长度为

表3 基于模型分解的递推最小二乘算法的计算量

Table 3 The computational efficiency of the D-RLS algorithm

变量	表达式	乘法次数	加法次数
$\hat{\mathbf{a}}(t)$	$\hat{\mathbf{a}}(t) = \hat{\mathbf{a}}(t-1) + L_1(t)e(t) \in \mathbf{R}^n$	n	n
	$e(t) := y_1(t) - \hat{\phi}^T(t)\hat{\mathbf{a}}(t-1) - \psi^T(t)\hat{\mathbf{b}}(t-1) \in \mathbf{R}$	$2n$	$2n$
$L_1(t)$	$L_1(t) = \zeta_1(t) / [1 + \hat{\phi}^T(t)\zeta_1(t)] \in \mathbf{R}^n$	$2n$	n
	$\zeta_1(t) := P_1(t-1)\hat{\phi}(t) \in \mathbf{R}^n$	n^2	$n^2 - n$
$P_1(t)$	$P_1(t) = P_1(t-1) - L_1(t)\zeta_1^T(t) \in \mathbf{R}^{n \times n}$	n^2	n^2
$\hat{\mathbf{b}}(t)$	$\hat{\mathbf{b}}(t) = \hat{\mathbf{b}}(t-1) + L_2(t)e(t) \in \mathbf{R}^n$	n	n
	$L_2(t) = \zeta_2(t) / [1 + \psi^T(t)\zeta_2(t)] \in \mathbf{R}^n$	$2n$	n
$P_2(t)$	$\zeta_2(t) := P_2(t-1)\psi(t) \in \mathbf{R}^n$	n^2	$n^2 - n$
	$P_2(t) = P_2(t-1) - L_2(t)\zeta_2^T(t) \in \mathbf{R}^{n \times n}$	n^2	n^2
总数		$4n^2 + 8n$	$4n^2 + 4n$
总 flop 数		$8n^2 + 12n$	

p , 定义堆积信息矩阵 $\hat{\Phi}(p, t)$ 和 $\Psi(p, t)$ 如下:

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

$$\Psi(p, t) := [\psi(t), \psi(t-1), \dots, \psi(t-p+1)] \in \mathbf{R}^{n \times p}.$$

将式(124)或(127)中辨识新息(标量新息)

$$e(t) := y(t) - \hat{\phi}^T(t) \hat{a}(t-1) - \psi^T(t) \hat{b}(t-1) \in \mathbf{R}$$

推广为新息向量

$$E(p, t) = Y(p, t) - \hat{\Phi}^T(p, t) \hat{a}(t-1) - \Psi^T(p, t) \hat{b}(t-1) \in \mathbf{R}^p.$$

参考多新息最小二乘辨识算法(37)–(44)的推导,可以得到计算参数估计向量 $\hat{a}(t)$ 和 $\hat{b}(t)$ 的基于模型分解的多新息最小二乘算法(model Decomposition based Multi-Innovation Least Squares algorithm, D-MILS 算法):

$$\hat{a}(t) = \hat{a}(t-1) + L_1(t) E(p, t), \quad \hat{a}(0) = \mathbf{1}_n / p_0, \quad (134)$$

$$E(p, t) = Y(p, t) - \hat{\Phi}^T(p, t) \hat{a}(t-1) - \Psi^T(p, t) \hat{b}(t-1), \quad (135)$$

$$L_1(t) = P_1(t-1) \hat{\Phi}(p, t) [I_p + \hat{\Phi}^T(p, t) P_1(t-1) \hat{\Phi}(p, t)]^{-1}, \quad (136)$$

$$P_1(t) = P_1(t-1) - L_1(t) \hat{\Phi}^T(p, t) P_1(t-1), \\ P_1(0) = p_0 I_n, \quad (137)$$

$$\hat{b}(t) = \hat{b}(t-1) + L_2(t) E(p, t), \quad \hat{b}(0) = \mathbf{1}_n / p_0, \quad (138)$$

$$L_2(t) = P_2(t-1) \Psi(p, t) [I_p + \Psi^T(p, t) P_2(t-1) \Psi(p, t)]^{-1}, \quad (139)$$

$$P_2(t) = P_2(t-1) - L_2(t) \Psi^T(p, t) P_2(t-1), \\ P_2(0) = p_0 I_n, \quad (140)$$

$$Y(p, t) = [y(t), y(t-1), \dots, y(t-p+1)]^T, \quad (141)$$

$$\hat{\Phi}(p, t) = [\hat{\phi}(t), \hat{\phi}(t-1), \dots, \hat{\phi}(t-p+1)], \quad (142)$$

$$\Psi(p, t) = [\psi(t), \psi(t-1), \dots, \psi(t-p+1)], \quad (143)$$

$$\hat{\phi}(t) = [-\hat{x}_1(t-1), -\hat{x}_1(t-2), \dots, -\hat{x}_1(t-n)]^T, \quad (144)$$

$$\psi(t) = [u(t-1), u(t-2), \dots, u(t-n)]^T. \quad (145)$$

算法中使用的状态估计可采用观测器(46)–(49)或者 Kalman 滤波状态估计算法(66)–(72)计算.因此

1) 式(134)–(145)与式(46)–(49)构成了基于状态观测器的模型分解多新息最小二乘参数辨识算法(O-D-MILS 算法);

2) 式(134)–(145)与式(66)–(72)构成了基于 Kalman 滤波状态估计的模型分解多新息最小二乘参数辨识算法(KF-D-MILS 算法).

表4列出了基于模型分解的多新息最小二乘辨识算法(134)–(145)的计算量.

4 状态空间系统的迭代辨识算法

迭代辨识方法与递推辨识方法构成了两大类重要的参数估计方法.递推辨识的递推变量与时间有关,因而可以用于在线估计参数;迭代辨识的迭代变量是自然数,与时间无关,通常用于离线估计系统参数.迭代辨识方法主要使用梯度搜索、最小二乘搜索、牛顿搜索等原理实现.本节主要研究状态空间系统的梯度迭代辨识算法和最小二乘迭代辨识算法,以及基于模型分解的梯度迭代辨识算法和基于模型分解的最小二乘迭代辨识算法.

为了与上述观测器规范型状态空间系统区别,这里考虑能观测性规范型状态空间系统^[38]:

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t), \quad (146)$$

$$y(t) = \mathbf{c}\mathbf{x}(t) + v(t), \quad (147)$$

表4 基于模型分解的多新息最小二乘算法的计算量

Table 4 The computational efficiency of the D-MILS algorithm

表达式	乘法次数	加法次数
$\hat{a}(t) = \hat{a}(t-1) + L_1(t) E(p, t) \in \mathbf{R}^n$	np	np
$E(p, t) = Y(p, t) - \hat{\Phi}^T(p, t) \hat{a}(t-1) - \Psi^T(p, t) \hat{b}(t-1) \in \mathbf{R}^p$	$2np$	$2np$
$L_1(t) = \zeta_1(t) [I_p + \hat{\Phi}^T(p, t) \zeta_1(t)]^{-1} \in \mathbf{R}^{n \times p}$	$2np^2 + p^3$	$2np^2 + p^3 - np - 2p^2$
$\zeta_1(t) := P_1(t-1) \hat{\Phi}(p, t) \in \mathbf{R}^{n \times p}$	$n^2 p$	$n^2 p - np$
$P_1(t) = P_1(t-1) - L_1(t) \zeta_1^T(t) \in \mathbf{R}^{n \times n}$	$n^2 p$	$n^2 p$
$\hat{b}(t) = \hat{b}(t-1) + L_2(t) E(p, t) \in \mathbf{R}^n$	np	np
$L_2(t) = \zeta_2(t) [I_p + \Psi^T(p, t) \zeta_2(t)]^{-1} \in \mathbf{R}^{n \times p}$	$2np^2 + p^3$	$2np^2 + p^3 - np - 2p^2$
$\zeta_2(t) := P_2(t-1) \Psi(p, t) \in \mathbf{R}^{n \times p}$	$n^2 p$	$n^2 p - np$
$P_2(t) = P_2(t-1) - L_2(t) \zeta_2^T(t) \in \mathbf{R}^{n \times n}$	$n^2 p$	$n^2 p$
总数	$4n^2 p + 4np + 4np^2 + 2p^3$	$4n^2 p + 4np^2 + 2p^3 - 4p^2$
总 flop 数	$8n^2 p + 8np^2 + 4p^3 + 4np - 4p^2$	

其中 $\mathbf{x}(t) := [x_1(t), x_2(t), \dots, x_n(t)]^T \in \mathbf{R}^n$ 为系统状态向量, $u(t) \in \mathbf{R}$ 和 $y(t) \in \mathbf{R}$ 为系统的输入和输出, $v(t) \in \mathbf{R}$ 是均值为零的白噪声, $\mathbf{A} \in \mathbf{R}^{n \times n}$, $\mathbf{b} \in \mathbf{R}^n$, $\mathbf{c} \in \mathbf{R}^{1 \times n}$ 是系统的参数矩阵(向量):

$$\mathbf{A} := \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix} \in \mathbf{R}^{n \times n},$$

$$\mathbf{b} := \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \\ b_n \end{bmatrix} \in \mathbf{R}^n,$$

$$\mathbf{c} := [1, 0, 0, \dots, 0] \in \mathbf{R}^{1 \times n}.$$

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

$$\boldsymbol{\vartheta} := \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} \in \mathbf{R}^{2n},$$

$$\mathbf{a} := [a_1, a_2, \dots, a_n]^T \in \mathbf{R}^n,$$

$$\mathbf{b} := [b_1, b_2, \dots, b_n]^T \in \mathbf{R}^n,$$

$$\boldsymbol{\phi}(t) := \begin{bmatrix} \boldsymbol{\phi}(t) \\ \boldsymbol{\psi}(t) \end{bmatrix} \in \mathbf{R}^{2n},$$

$$\boldsymbol{\phi}(t) := \begin{bmatrix} -x_n(t-n) \\ -x_{n-1}(t-n) \\ \vdots \\ -x_1(t-n) \end{bmatrix} \in \mathbf{R}^n,$$

$$\boldsymbol{\psi}(t) := \begin{bmatrix} u(t-1) \\ u(t-2) \\ \vdots \\ u(t-n) \end{bmatrix} \in \mathbf{R}^n.$$

向量 $\boldsymbol{\phi}(t)$ 由未知状态变量 $x_i(t-n)$ 构成, 向量 $\boldsymbol{\psi}(t)$ 则由系统输入 $u(t-i)$ 构成. 由式(146)–(147)可得:

$$x_i(t+1) = x_{i+1}(t) + b_i u(t), \quad i = 1, 2, \dots, n-1, \quad (148)$$

$$x_n(t+1) = -\sum_{i=1}^n a_i x_{n-i+1}(t) + b_n u(t). \quad (149)$$

式(148)两边同乘 z^{-i} , 得到

$$x_i(t-i+1) = x_{i+1}(t-i) + b_i u(t-i), \quad i = 1, 2, \dots, n-1.$$

对上式求和 $\sum_{i=1}^{n-1}$ 得到

$$x_1(t) = x_n(t-n+1) + \sum_{i=1}^{n-1} b_i u(t-i). \quad (150)$$

式(149)两边同乘 z^{-n} 得到

$$x_n(t-n+1) = -\sum_{i=1}^n a_i x_{n-i+1}(t-n) + b_n u(t-n). \quad (151)$$

将式(151)代入式(150), 可以得到

$$x_1(t) = -\sum_{i=1}^n a_i x_{n-i+1}(t-n) + b_n u(t-n) + \sum_{i=1}^{n-1} b_i u(t-i) = -\sum_{i=1}^n a_i x_{n-i+1}(t-n) + \sum_{i=1}^n b_i u(t-i). \quad (152)$$

将式(152)代入式(147), 可以得到能观测性规范型状态空间系统(146)–(147)的辨识模型:

$$y(t) = x_1(t) + v(t) = \boldsymbol{\varphi}^T(t) \boldsymbol{\vartheta} + v(t). \quad (153)$$

从式(153)可以看出, 辨识的难点在于信息向量中存在着未知状态变量 $x_i(t-n)$. 下面基于辅助模型辨识思想, 利用系统输入输出数据 $\{u(t), y(t) : t=1, 2, 3, \dots, L\}$ ($L \gg n$ 为数据长度) 研究基于状态估计的状态空间系统的梯度迭代辨识算法和最小二乘迭代辨识算法, 来估计系统参数, 并针对高维数系统存在矩阵计算量过大影响计算效率的问题, 研究计算量较小的基于模型分解的梯度迭代辨识算法和基于模型分解的最小二乘迭代辨识算法.

4.1 梯度迭代辨识算法

考虑从 $i=t-p+1$ 到 $i=t$ 的 p 组数据, 定义堆积输出向量 $\mathbf{Y}(t)$ 和堆积信息矩阵 $\mathbf{H}(t)$ 如下:

$$\mathbf{Y}(t) := [y(t-p+1), y(t-p+2), \dots, y(t)]^T \in \mathbf{R}^p, \quad (154)$$

$$\mathbf{H}(t) := [\boldsymbol{\varphi}(t-p+1), \boldsymbol{\varphi}(t-p+2), \dots, \boldsymbol{\varphi}(t)]^T \in \mathbf{R}^{p \times (2n)}. \quad (155)$$

利用移动数据窗口的数据, 定义一个动态二次准则函数:

$$J_7(\boldsymbol{\vartheta}) := \|\mathbf{Y}(t) - \mathbf{H}(t) \boldsymbol{\vartheta}\|^2.$$

假设 $\hat{\boldsymbol{\vartheta}}_k(t)$ 表示 $\boldsymbol{\vartheta}$ 在时刻 t 的第 k 次迭代估计, $\mu_k(t) \geq 0$ 表示时变迭代步长 (time-varying step-size) 或者时变收敛因子 (time-varying convergence factor). 根据梯度搜索原理, 参考文献[1], 极小化准则函数 $J_7(\boldsymbol{\vartheta})$ 可以得到下面的递推关系式:

$$\hat{\boldsymbol{\vartheta}}_k(t) = \hat{\boldsymbol{\vartheta}}_{k-1}(t) - \frac{\mu_k(t)}{2} \text{grad}[J_7(\hat{\boldsymbol{\vartheta}}_{k-1}(t))] = \hat{\boldsymbol{\vartheta}}_{k-1}(t) + \mu_k(t) \mathbf{H}^T(t) [\mathbf{Y}(t) - \mathbf{H}(t) \hat{\boldsymbol{\vartheta}}_{k-1}(t)]. \quad (156)$$

由于信息向量 $\boldsymbol{\varphi}(t)$ 包含未知中间状态变量 $x_i(t-n)$, 为了实现参数估计的递推计算, 可以用其第 $k-1$ 次迭代估计值 $\hat{x}_{i,k-1}(t-n)$ 代替, 构造 $\boldsymbol{\varphi}(t)$ 的第 k 次迭代估计 $\hat{\boldsymbol{\varphi}}_k(t)$ 和堆积信息矩阵 $\mathbf{H}(t)$ 的第 k 次迭代估计 $\hat{\mathbf{H}}_k(t)$ 如下:

$$\hat{\boldsymbol{\varphi}}_k(t) := [-\hat{x}_{n,k-1}(t-n), -\hat{x}_{n-1,k-1}(t-n), \dots, -\hat{x}_{1,k-1}(t-n), u(t-1), u(t-2), \dots, u(t-n)]^T \in \mathbf{R}^{2n},$$

$$\hat{\mathbf{H}}_k(t) := [\hat{\boldsymbol{\varphi}}_k(t-p+1), \hat{\boldsymbol{\varphi}}_k(t-p+2), \dots, \hat{\boldsymbol{\varphi}}_k(t)]^T \in \mathbf{R}^{p \times (2n)}.$$

式(156)中信息矩阵 $\mathbf{H}(t)$ 用其估计值 $\hat{\mathbf{H}}_k(t)$ 代替,得到

$$\hat{\boldsymbol{\theta}}_k(t) = \hat{\boldsymbol{\theta}}_{k-1}(t) + \mu_k(t) \hat{\mathbf{H}}_k^T(t) [\mathbf{Y}(t) - \hat{\mathbf{H}}_k(t) \hat{\boldsymbol{\theta}}_{k-1}(t)],$$

$$k=1, 2, 3, \dots, \quad (157)$$

亦可写作

$$\hat{\boldsymbol{\theta}}_k(t) = [\mathbf{I} - \mu_k(t) \hat{\mathbf{H}}_k^T(t) \hat{\mathbf{H}}_k(t)] \hat{\boldsymbol{\theta}}_{k-1}(t) + \mu_k(t) \hat{\mathbf{H}}_k^T(t) \mathbf{Y}(t).$$

假设 $\lambda_{\max}[\mathbf{X}]$ 为实对称阵 \mathbf{X} 的最大的特征值.上式可看成 k 的离散时间系统.为保证“状态” $\hat{\boldsymbol{\theta}}_k(t)$ 收敛,要求矩阵 $[\mathbf{I} - \mu_k(t) \hat{\mathbf{H}}_k^T(t) \hat{\mathbf{H}}_k(t)]$ 的特征值均不在单位圆外,且单位圆上没有重特征值,即

$$\lambda_{\max}[\mathbf{I} - \mu_k(t) \hat{\mathbf{H}}_k^T(t) \hat{\mathbf{H}}_k(t)] \leq 1,$$

故收敛因子 μ_k 的一个保守选择为

$$\mu_k(t) \leq \frac{2}{\lambda_{\max}[\hat{\mathbf{H}}_k^T(t) \hat{\mathbf{H}}_k(t)]}.$$

综合以上各式,可以得到估计参数向量 $\boldsymbol{\theta}$ 的梯度迭代算法(Gradient based Iterative algorithm, GI 算法):

$$\hat{\boldsymbol{\theta}}_k(t) = \hat{\boldsymbol{\theta}}_{k-1}(t) + \mu_k(t) \hat{\mathbf{H}}_k^T(t) [\mathbf{Y}(t) - \hat{\mathbf{H}}_k(t) \hat{\boldsymbol{\theta}}_{k-1}(t)],$$

$$\hat{\boldsymbol{\theta}}_0(0) = \mathbf{1}_{2n}/p_0, \quad (158)$$

$$\mathbf{Y}(t) = [y(t-p+1), y(t-p+2), \dots, y(t)]^T, \quad (159)$$

$$\hat{\mathbf{H}}_k(t) = [\hat{\boldsymbol{\varphi}}_k(t-p+1), \hat{\boldsymbol{\varphi}}_k(t-p+2), \dots, \hat{\boldsymbol{\varphi}}_k(t)]^T, \quad (160)$$

$$\boldsymbol{\varphi}_k(t) = [-\hat{x}_{n,k-1}(t-n), -\hat{x}_{n-1,k-1}(t-n), \dots, -\hat{x}_{1,k-1}(t-n),$$

$$u(t-1), u(t-2), \dots, u(t-n)]^T, \quad (161)$$

$$\mu_k(t) \leq 2\lambda_{\max}^{-1}[\hat{\mathbf{H}}_k^T(t) \hat{\mathbf{H}}_k(t)], \quad (162)$$

$$\hat{\boldsymbol{\theta}}_k(t) = [\hat{a}_{1,k}(t), \hat{a}_{2,k}(t), \dots, \hat{a}_{n,k}(t), \hat{b}_{1,k}(t),$$

$$\hat{b}_{2,k}(t), \dots, \hat{b}_{n,k}(t)]^T. \quad (163)$$

这个算法中 k 是迭代变量, t 是时间.当时间 t 固定,就是利用一个固定数据窗口的数据(数据窗长度是 p)进行迭代计算参数估计向量 $\hat{\boldsymbol{\theta}}_k(t)$;数据窗长度应该足够大时(对于慢时变参数系统,长度可以适当大一些;对于快时变参数系统,长度可以适当小一些),迭代到一定步数,参数估计到达一定精度,就可以增加 t ,将数据窗向前移动一步,利用新数据窗的数据,重新开始计算迭代参数估计,其计算步骤和流程图可参见《系统辨识新论》^[1]中有关迭代辨识的章节.

下面给出非动态数据窗情况下的梯度迭代算法.在式(154)—(155)中取定 $p=L, t=L$,令

$$\mathbf{Y} := [y(1), y(2), \dots, y(L)]^T \in \mathbf{R}^L,$$

$$\mathbf{H} := [\boldsymbol{\varphi}(1), \boldsymbol{\varphi}(2), \dots, \boldsymbol{\varphi}(L)]^T \in \mathbf{R}^{L \times (2n)}.$$

令 $\mu_k \geq 0$ 为收敛因子. \mathbf{Y} 和 \mathbf{H} 就包含了所有量测输

入输出数据 $\{u(t), y(t) : t=1, 2, \dots, L\}$.

令 $\hat{\boldsymbol{\theta}}_k$ 表示 $\boldsymbol{\theta}$ 的第 k 次迭代估计, $\mu_k \geq 0$ 表示迭代步长或者收敛因子.定义准则函数:

$$J_g(\boldsymbol{\theta}) := \|\mathbf{Y}(t) - \mathbf{H}(t) \boldsymbol{\theta}\|^2.$$

使用负梯度搜索,极小化 $J_g(\boldsymbol{\theta})$,采用与算法(158)—(163)类似的推导,可以得到有限量测数据 GI 算法:

$$\hat{\boldsymbol{\theta}}_k = \hat{\boldsymbol{\theta}}_{k-1} + \mu_k \hat{\mathbf{H}}_k^T [\mathbf{Y} - \hat{\mathbf{H}}_k \hat{\boldsymbol{\theta}}_{k-1}], \quad \hat{\boldsymbol{\theta}}_0 = \mathbf{1}_{2n}/p_0, \quad (164)$$

$$\mathbf{Y} = [y(1), y(2), \dots, y(L)]^T, \quad (165)$$

$$\hat{\mathbf{H}}_k = [\hat{\boldsymbol{\varphi}}_k(1), \hat{\boldsymbol{\varphi}}_k(2), \dots, \hat{\boldsymbol{\varphi}}_k(L)]^T, \quad (166)$$

$$\hat{\boldsymbol{\varphi}}_k(t) = [-\hat{x}_{n,k-1}(t-n), -\hat{x}_{n-1,k-1}(t-n), \dots, -\hat{x}_{1,k-1}(t-n),$$

$$u(t-1), u(t-2), \dots, u(t-n)]^T, \quad (167)$$

$$\mu_k \leq 2\lambda_{\max}^{-1}[\hat{\mathbf{H}}_k^T \hat{\mathbf{H}}_k], \quad (168)$$

$$\hat{\boldsymbol{\theta}}_k = [\hat{a}_{1,k}, \hat{a}_{2,k}, \dots, \hat{a}_{n,k}, \hat{b}_{1,k}, \hat{b}_{2,k}, \dots, \hat{b}_{n,k}]^T. \quad (169)$$

算法中使用的状态估计可采用基于迭代参数估计的状态观测器(238)—(241)或者基于迭代参数估计的 Kalman 滤波状态估计算法(258)—(263)计算.因此

1) 式(164)—(169)与式(238)—(241)构成了基于状态观测器的梯度迭代参数辨识算法(O-GI 算法);

2) 式(164)—(169)与式(258)—(263)构成了基于 Kalman 滤波状态估计的梯度迭代参数辨识算法(KF-GI 算法).

4.2 基于模型分解的梯度迭代辨识算法

基于模型分解的梯度迭代辨识算法也称为递阶梯度迭代辨识算法^[39].重写式(153)如下:

$$y(t) = \boldsymbol{\phi}^T(t) \mathbf{a} + \boldsymbol{\psi}^T(t) \mathbf{b} + v(t) = \boldsymbol{\varphi}^T(t) \boldsymbol{\theta} + v(t), \quad (170)$$

其中各变量定义同上.

定义两个中间变量:

$$y_1(t) := y(t) - \boldsymbol{\psi}^T(t) \mathbf{b}, \quad (171)$$

$$y_2(t) := y(t) - \boldsymbol{\phi}^T(t) \mathbf{a}. \quad (172)$$

于是系统(170)可分解为两个子系统:

$$y_1(t) = \boldsymbol{\phi}^T(t) \mathbf{a} + v(t), \quad (173)$$

$$y_2(t) = \boldsymbol{\psi}^T(t) \mathbf{b} + v(t). \quad (174)$$

同样考虑从 $i=t-p+1$ 到 $i=t$ 的 p 组数据,定义两个子系统堆积输出向量 $\mathbf{Y}_1(t)$ 和 $\mathbf{Y}_2(t)$,堆积信息矩阵 $\boldsymbol{\Phi}(t)$ 和 $\boldsymbol{\Psi}(t)$ 如下:

$$\mathbf{Y}_1(t) := [y_1(t-p+1), y_1(t-p+2), \dots, y_1(t)]^T \in \mathbf{R}^p,$$

$$\mathbf{Y}_2(t) := [y_2(t-p+1), y_2(t-p+2), \dots, y_2(t)]^T \in \mathbf{R}^p,$$

$$\boldsymbol{\Phi}(t) := [\boldsymbol{\phi}(t-p+1), \boldsymbol{\phi}(t-p+2), \dots, \boldsymbol{\phi}(t)]^T \in \mathbf{R}^{p \times n},$$

$$\boldsymbol{\Psi}(t) := [\boldsymbol{\psi}(t-p+1), \boldsymbol{\psi}(t-p+2), \dots, \boldsymbol{\psi}(t)]^T \in \mathbf{R}^{p \times n},$$

则从式(171)和式(172)可得

$$\begin{aligned} Y_1(t) &= Y(t) - \Psi(t)b, \\ Y_2(t) &= Y(t) - \Phi(t)a. \end{aligned}$$

定义两个子系统的准则函数:

$$\begin{aligned} J_9(a) &:= \|Y_1(t) - \Phi(t)a\|^2, \\ J_{10}(b) &:= \|Y_2(t) - \Psi(t)b\|^2. \end{aligned}$$

设 $\hat{a}_k(t)$ 表示 a 在时刻 t 的第 k 次迭代估计, $\hat{b}_k(t)$ 表示 b 在时刻 t 的第 k 次迭代估计, $\mu_{a,k}(t) \geq 0$ 和 $\mu_{b,k}(t) \geq 0$ 分别表示两个子系统的时变迭代步长或时变收敛因子. 根据梯度搜索原理, 极小化准则函数 $J_9(a)$ 和 $J_{10}(b)$ 可以得到下面的递推关系式:

$$\begin{aligned} \hat{a}_k(t) &= \hat{a}_{k-1}(t) - \frac{\mu_{a,k}(t)}{2} \text{grad}[J_9(\hat{a}_{k-1}(t))] = \\ &\hat{a}_{k-1}(t) + \mu_{a,k}(t) \Phi^T(t) [Y_1(t) - \Phi(t)\hat{a}_{k-1}(t)] = \\ &\hat{a}_{k-1}(t) + \mu_{a,k}(t) \Phi^T(t) [Y(t) - \Psi(t)b - \Phi(t)\hat{a}_{k-1}(t)], \end{aligned} \quad (175)$$

$$\begin{aligned} \hat{b}_k(t) &= \hat{b}_{k-1}(t) - \frac{\mu_{b,k}(t)}{2} \text{grad}[J_{10}(\hat{b}_{k-1}(t))] = \\ &\hat{b}_{k-1}(t) + \mu_{b,k}(t) \Psi^T(t) [Y_2(t) - \Psi(t)\hat{b}_{k-1}(t)] = \\ &\hat{b}_{k-1}(t) + \mu_{b,k}(t) \Psi^T(t) [Y(t) - \Phi(t)a - \Psi(t)\hat{b}_{k-1}(t)]. \end{aligned} \quad (176)$$

信息向量 $\phi(t)$ 包含的不可测状态 $x_i(t-n)$, 用其第 $k-1$ 次迭代估计值 $\hat{x}_{i,k-1}(t-n)$ 代替, 构造 $\hat{\phi}(t)$ 的第 k 次迭代 $\hat{\phi}_k(t)$, 构造堆积信息矩阵 $\Phi(t)$ 的估计 $\hat{\Phi}_k(t)$ 如下:

$$\hat{\phi}_k(t) := [-\hat{x}_{n,k-1}(t-n), -\hat{x}_{n-1,k-1}(t-n), \dots, -\hat{x}_{1,k-1}(t-n)]^T \in \mathbf{R}^n,$$

$$\hat{\Phi}_k(t) := [\hat{\phi}_k(t-p+1), \hat{\phi}_k(t-p+2), \dots, \hat{\phi}_k(t)]^T \in \mathbf{R}^{p \times n}.$$

为保证参数估计收敛, $\mu_{a,k}(t)$ 和 $\mu_{b,k}(t)$ 的保守取值为

$$\mu_{a,k}(t) \leq \frac{2}{\lambda_{\max}[\hat{\Phi}_k^T(t)\hat{\Phi}_k(t)]}, \quad (177)$$

$$\mu_{b,k}(t) \leq \frac{2}{\lambda_{\max}[\Psi^T(t)\Psi(t)]}. \quad (178)$$

为了实现参数估计的递推计算, 式(175) — (176) 中未知 $H(t)$ 用其估计 $\hat{H}_k(t)$ 代替, 未知参数向量 b 和 a 分别用其前一次迭代估计 $\hat{b}_{k-1}(t)$ 和 $\hat{a}_{k-1}(t)$ 代替, 可以得到计算参数估计向量 $\hat{a}_k(t)$ 和 $\hat{b}_k(t)$ 的基于模型分解的梯度迭代算法 (model Decomposition based Gradient Iterative algorithm, D-GI 算法):

$$\hat{a}_k(t) = \hat{a}_{k-1}(t) + \mu_{a,k}(t) \hat{\Phi}_k^T(t) [Y(t) - \hat{\Phi}_k(t)\hat{a}_{k-1}(t) - \Psi(t)\hat{b}_{k-1}(t)], \quad (179)$$

$$\hat{b}_k(t) = \hat{b}_{k-1}(t) + \mu_{b,k}(t) \Psi^T(t) [Y(t) - \hat{\Phi}_k(t)\hat{a}_{k-1}(t) - \Psi(t)\hat{b}_{k-1}(t)], \quad (180)$$

$$Y(t) = [y(t-p+1), y(t-p+2), \dots, y(t)]^T, \quad (181)$$

$$\hat{\Phi}_k(t) = [\hat{\phi}_k(t-p+1), \hat{\phi}_k(t-p+2), \dots, \hat{\phi}_k(t)]^T, \quad (182)$$

$$\Psi(t) = [\psi(t-p+1), \psi(t-p+2), \dots, \psi(t)]^T, \quad (183)$$

$$\hat{\phi}_k(t) = [-\hat{x}_{n,k-1}(t-n), -\hat{x}_{n-1,k-1}(t-n), \dots, -\hat{x}_{1,k-1}(t-n)]^T, \quad (184)$$

$$\psi(t) = [u(t-1), u(t-2), \dots, u(t-n)]^T, \quad (185)$$

$$\mu_{a,k}(t) \leq 2\lambda_{\max}^{-1}[\hat{\Phi}_k^T(t)\hat{\Phi}_k(t)], \quad (186)$$

$$\mu_{b,k}(t) \leq 2\lambda_{\max}^{-1}[\Psi^T(t)\Psi(t)], \quad (187)$$

$$\hat{a}_k(t) = [\hat{a}_{1,k}(t), \hat{a}_{2,k}(t), \dots, \hat{a}_{n,k}(t)]^T, \quad (188)$$

$$\hat{b}_k(t) = [\hat{b}_{1,k}(t), \hat{b}_{2,k}(t), \dots, \hat{b}_{n,k}(t)]^T. \quad (189)$$

令 $t=L, p=L, \hat{a}_k$ 和 \hat{b}_k 是有限量测数据下参数向量 a 和 b 第 k 次迭代的估计, 便有估计参数向量 a 和 b 的有限量测数据 D-GI 算法:

$$\hat{a}_k = \hat{a}_{k-1} + \mu_{a,k} \hat{\Phi}_k^T [Y - \hat{\Phi}_k \hat{a}_{k-1} - \Psi \hat{b}_{k-1}], \quad (190)$$

$$\hat{b}_k = \hat{b}_{k-1} + \mu_{b,k} \Psi^T [Y - \hat{\Phi}_k \hat{a}_{k-1} - \Psi \hat{b}_{k-1}], \quad (191)$$

$$Y = [y(1), y(2), \dots, y(L)]^T, \quad (192)$$

$$\hat{\Phi}_k = [\hat{\phi}_k(1), \hat{\phi}_k(2), \dots, \hat{\phi}_k(L)]^T, \quad (193)$$

$$\Psi = [\psi(1), \psi(2), \dots, \psi(L)]^T, \quad (194)$$

$$\hat{\phi}_k(t) = [-\hat{x}_{n,k-1}(t-n), -\hat{x}_{n-1,k-1}(t-n), \dots, -\hat{x}_{1,k-1}(t-n)]^T, \quad (195)$$

$$\psi(t) = [u(t-1), u(t-2), \dots, u(t-n)]^T, \quad (196)$$

$$\mu_{a,k} \leq 2\lambda_{\max}^{-1}[\hat{\Phi}_k^T \hat{\Phi}_k], \quad (197)$$

$$\mu_{b,k} \leq 2\lambda_{\max}^{-1}[\Psi^T \Psi], \quad (198)$$

$$\hat{a}_k = [\hat{a}_{1,k}, \hat{a}_{2,k}, \dots, \hat{a}_{n,k}]^T, \quad (199)$$

$$\hat{b}_k = [\hat{b}_{1,k}, \hat{b}_{2,k}, \dots, \hat{b}_{n,k}]^T. \quad (200)$$

算法中使用的状态估计可采用基于迭代参数估计的状态观测器(238) — (241) 或者基于迭代参数估计的 Kalman 滤波状态估计算法(258) — (263) 计算. 因此

1) 式(190) — (200) 与式(238) — (241) 构成了基于状态观测器的模型分解梯度迭代参数辨识算法(O-D-GI 算法);

2) 式(190) — (200) 与式(258) — (263) 构成了基于 Kalman 滤波状态估计的模型分解梯度迭代参数辨识算法(KF-D-GI 算法).

4.3 最小二乘迭代辨识算法

重写式(153)如下:

$$y(t) = \phi^T(t)a + \psi^T(t)b + v(t) = \varphi^T(t)\vartheta + v(t), \quad (201)$$

其中各变量定义同上.

考虑从 $i=t-p+1$ 到 $i=t$ 的 p 组数据,定义堆积输出向量 $\mathbf{Y}(t)$,堆积信息矩阵 $\mathbf{H}(t)$ 如下:

$$\mathbf{Y}(t) := \begin{bmatrix} y(t-p+1) \\ y(t-p+2) \\ \vdots \\ y(t) \end{bmatrix} \in \mathbf{R}^p,$$

$$\mathbf{H}(t) := \begin{bmatrix} \boldsymbol{\varphi}^T(t-p+1) \\ \boldsymbol{\varphi}^T(t-p+2) \\ \vdots \\ \boldsymbol{\varphi}^T(t) \end{bmatrix} \in \mathbf{R}^{p \times (2n)}.$$

定义最小二乘准则函数:

$$J_{11}(\boldsymbol{\vartheta}) := \|\mathbf{Y}(t) - \mathbf{H}(t)\boldsymbol{\vartheta}\|^2.$$

令 $J_{11}(\boldsymbol{\vartheta})$ 对 $\boldsymbol{\vartheta}$ 一阶偏导为零:

$$\frac{\partial J_{11}(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}} = -2\mathbf{H}^T(t) [\mathbf{Y}(t) - \mathbf{H}(t)\boldsymbol{\vartheta}] \Big|_{\boldsymbol{\vartheta}=\hat{\boldsymbol{\vartheta}}(t)} = \mathbf{0}.$$

假设信息向量 $\boldsymbol{\varphi}(t)$ 是持续激励的,即 $[\mathbf{H}^T(t) \cdot \mathbf{H}(t)]$ 是可逆矩阵,故而可求得参数向量 $\boldsymbol{\vartheta}$ 的最小二乘估计为

$$\hat{\boldsymbol{\vartheta}}(t) = [\mathbf{H}^T(t)\mathbf{H}(t)]^{-1}\mathbf{H}^T(t)\mathbf{Y}(t). \quad (202)$$

设 $k=1,2,3, \dots$ 是迭代变量, $\hat{\boldsymbol{\vartheta}}_k(t)$ 为 $\boldsymbol{\vartheta}$ 在时刻 t 的第 k 次迭代估计.由于信息向量 $\boldsymbol{\varphi}(t)$ 包含不可测中间状态变量 $x_i(t-n)$,用其第 $k-1$ 次迭代估计值 $\hat{x}_{i,k-1}(t-n)$ 构造 $\boldsymbol{\varphi}(t)$ 的第 k 次迭代估计:

$$\hat{\boldsymbol{\varphi}}_k(t) := \begin{bmatrix} \hat{\boldsymbol{\varphi}}_k(t) \\ \boldsymbol{\psi}(t) \end{bmatrix} \in \mathbf{R}^{2n},$$

$$\hat{\boldsymbol{\phi}}_k(t) := \begin{bmatrix} -\hat{x}_{n,k-1}(t-n) \\ -\hat{x}_{n-1,k-1}(t-n) \\ \vdots \\ -\hat{x}_{1,k-1}(t-n) \end{bmatrix} \in \mathbf{R}^n.$$

用 $\hat{\boldsymbol{\varphi}}_k(t)$ 构造堆积信息矩阵 $\mathbf{H}(t)$ 的第 k 次迭代估计:

$$\hat{\mathbf{H}}_k(t) := [\hat{\boldsymbol{\varphi}}_k(t-p+1), \hat{\boldsymbol{\varphi}}_k(t-p+2), \dots, \hat{\boldsymbol{\varphi}}_k(t)]^T \in \mathbf{R}^{p \times (2n)}.$$

参考文献[1,40],用 $\hat{\mathbf{H}}_k(t)$ 代替式(202)中的 $\mathbf{H}(t)$,可以得到计算参数估计向量 $\hat{\boldsymbol{\vartheta}}_k(t)$ 的最小二乘迭代算法(Least Squares based Iterative algorithm, LSI 算法):

$$\hat{\boldsymbol{\vartheta}}_k(t) = [\hat{\mathbf{H}}_k^T(t)\hat{\mathbf{H}}_k(t)]^{-1}\hat{\mathbf{H}}_k^T(t)\mathbf{Y}(t),$$

$$k=1,2,3, \dots, \quad (203)$$

$$\hat{\mathbf{H}}_k(t) = [\hat{\boldsymbol{\varphi}}_k(t-p+1), \hat{\boldsymbol{\varphi}}_k(t-p+2), \dots, \hat{\boldsymbol{\varphi}}_k(t)]^T, \quad (204)$$

$$\mathbf{Y}(t) = [y(t-p+1), y(t-p+2), \dots, y(t)]^T, \quad (205)$$

$$\hat{\boldsymbol{\varphi}}_k(t) = [-\hat{x}_{n,k-1}(t-n), -\hat{x}_{n-1,k-1}(t-n), \dots, -\hat{x}_{1,k-1}(t-n),$$

$$u(t-1), u(t-2), \dots, u(t-n)]^T, \quad (206)$$

$$\hat{\boldsymbol{\theta}}_k(t) = [\hat{a}_{1,k}(t), \hat{a}_{2,k}(t), \dots, \hat{a}_{n,k}(t), \hat{b}_{1,k}(t),$$

$$\hat{b}_{2,k}(t), \dots, \hat{b}_{n,k}(t)]^T. \quad (207)$$

令 $t=L, p=L$,可以得到估计参数向量 $\boldsymbol{\vartheta}$ 的有限量测数据 LSI 算法:

$$\hat{\boldsymbol{\vartheta}}_k = (\hat{\mathbf{H}}_k^T \hat{\mathbf{H}}_k)^{-1} \hat{\mathbf{H}}_k^T \mathbf{Y}, \quad k=1,2,3, \dots, \quad (208)$$

$$\hat{\mathbf{H}}_k = [\hat{\boldsymbol{\phi}}_k(1), \hat{\boldsymbol{\phi}}_k(2), \dots, \hat{\boldsymbol{\phi}}_k(L)]^T, \quad (209)$$

$$\mathbf{Y} = [y(1), y(2), \dots, y(L)]^T, \quad (210)$$

$$\hat{\boldsymbol{\phi}}_k(t) = [-\hat{x}_{n,k-1}(t-n), -\hat{x}_{n-1,k-1}(t-n), \dots, -\hat{x}_{1,k-1}(t-n),$$

$$u(t-1), u(t-2), \dots, u(t-n)]^T, \quad (211)$$

$$\hat{\boldsymbol{\theta}}_k = [\hat{a}_{1,k}, \hat{a}_{2,k}, \dots, \hat{a}_{n,k}, \hat{b}_{1,k}, \hat{b}_{2,k}, \dots, \hat{b}_{n,k}]^T. \quad (212)$$

算法中使用的状态估计可采用基于迭代参数估计的状态观测器(238)——(241)或者基于迭代参数估计的 Kalman 滤波状态估计算法(258)——(263)计算.因此

1) 式(208)——(212)与式(238)——(241)构成了基于状态观测器的 LSI 参数辨识算法(O-LSI 算法);

2) 式(208)——(212)与式(258)——(263)构成了基于 Kalman 滤波状态估计的 LSI 参数辨识算法(KF-LSI 算法).

表 5 给出了有限量测数据最小二乘迭代辨识算法的计算量.

表 5 有限量测数据最小二乘迭代算法的计算量

Table 5 The computational efficiency of the LSI algorithm

表达式	乘法次数	加法次数
$\hat{\boldsymbol{\vartheta}}_k = \mathbf{S}_k \boldsymbol{\beta}_k \in \mathbf{R}^{2n}$	$4n^2$	$4n^2 - 2n$
$\mathbf{S}_k := (\mathbf{S}'_k)^{-1} \in \mathbf{R}^{2n \times 2n}$	$8n^3$	$8n^3 - 4n^2$
$\mathbf{S}'_k := \mathbf{H}^T \mathbf{H} \in \mathbf{R}^{2n \times 2n}$	$4n^2 L$	$4n^2 L - 4n^2$
$\boldsymbol{\beta}_k = \mathbf{H}^T \mathbf{Y} \in \mathbf{R}^{2n}$	$2nL$	$2nL - 2n$
总数	$8n^3 + 4n^2 + 4n^2 L + 2nL$	$8n^3 + 4n^2 L + 2nL - 4n^2 - 4n$
总 flop 数	$16n^3 + 8n^2 L + 4nL - 4n$	

4.4 基于模型分解的最小二乘迭代辨识算法

基于模型分解的最小二乘迭代辨识算法也称为递阶最小二乘迭代辨识算法^[39].重写式(153)如下:

$$y(t) = \boldsymbol{\phi}^T(t)\mathbf{a} + \boldsymbol{\psi}^T(t)\mathbf{b} + v(t) = \boldsymbol{\varphi}^T(t)\boldsymbol{\vartheta} + v(t), \quad (213)$$

其中各变量定义同上.

定义两个中间变量:

$$y_1(t) := y(t) - \boldsymbol{\psi}^T(t)\mathbf{b}, \quad (214)$$

$$y_2(t) := y(t) - \boldsymbol{\phi}^T(t)\mathbf{a}. \quad (215)$$

利用式(214)和式(215),可以将系统(213)分解为两个子系统:

$$y_1(t) = \boldsymbol{\phi}^T(t)\mathbf{a} + v(t), \quad (216)$$

$$y_2(t) = \boldsymbol{\psi}^T(t) \mathbf{b} + v(t). \quad (217)$$

考虑从 $i=t-p+1$ 到 $i=t$ 的 p 组数据,定义两个子系统堆积输出向量 $\mathbf{Y}_1(t)$ 和 $\mathbf{Y}_2(t)$, 和堆积信息矩阵 $\boldsymbol{\Phi}(t)$ 和 $\boldsymbol{\Psi}(t)$ 如下:

$$\mathbf{Y}_1(t) := [y_1(t-p+1), y_1(t-p+2), \dots, y_1(t)]^T \in \mathbf{R}^p,$$

$$\mathbf{Y}_2(t) := [y_2(t-p+1), y_2(t-p+2), \dots, y_2(t)]^T \in \mathbf{R}^p,$$

$$\boldsymbol{\Phi}(t) := [\boldsymbol{\phi}(t-p+1), \boldsymbol{\phi}(t-p+2), \dots, \boldsymbol{\phi}(t)]^T \in \mathbf{R}^{p \times n},$$

$$\boldsymbol{\Psi}(t) := [\boldsymbol{\psi}(t-p+1), \boldsymbol{\psi}(t-p+2), \dots, \boldsymbol{\psi}(t)]^T \in \mathbf{R}^{p \times n},$$

则有

$$\mathbf{Y}_1(t) = \mathbf{Y}(t) - \boldsymbol{\Psi}(t) \mathbf{b},$$

$$\mathbf{Y}_2(t) = \mathbf{Y}(t) - \boldsymbol{\Phi}(t) \mathbf{a}.$$

定义两个子系统的准则函数:

$$J_{12}(\mathbf{a}) := \|\mathbf{Y}_1(t) - \boldsymbol{\Phi}(t) \mathbf{a}\|^2,$$

$$J_{13}(\mathbf{b}) := \|\mathbf{Y}_2(t) - \boldsymbol{\Psi}(t) \mathbf{b}\|^2.$$

假设 $[\boldsymbol{\Phi}^T(t) \boldsymbol{\Phi}(t)]$ 和 $[\boldsymbol{\Psi}^T(t) \boldsymbol{\Psi}(t)]$ 均非奇异, 极小化 $J_{12}(\mathbf{a})$ 和 $J_{13}(\mathbf{b})$ 可得子系统参数向量 \mathbf{a} 和 \mathbf{b} 的最小二乘估计:

$$\hat{\mathbf{a}}(t) = [\boldsymbol{\Phi}^T(t) \boldsymbol{\Phi}(t)]^{-1} \boldsymbol{\Phi}^T(t) \mathbf{Y}_1(t) = [\boldsymbol{\Phi}^T(t) \boldsymbol{\Phi}(t)]^{-1} \boldsymbol{\Phi}^T(t) [\mathbf{Y}(t) - \boldsymbol{\Psi}(t) \mathbf{b}], \quad (218)$$

$$\hat{\mathbf{b}}(t) = [\boldsymbol{\Psi}^T(t) \boldsymbol{\Psi}(t)]^{-1} \boldsymbol{\Psi}^T(t) \mathbf{Y}_2(t) = [\boldsymbol{\Psi}^T(t) \boldsymbol{\Psi}(t)]^{-1} \boldsymbol{\Psi}^T(t) [\mathbf{Y}(t) - \boldsymbol{\Phi}(t) \mathbf{a}]. \quad (219)$$

因为 $\boldsymbol{\Phi}(t)$ 中的 $\boldsymbol{\phi}(t)$ 包含了未知状态变量 $x_i(t-n)$, 可以用其第 $k-1$ 次迭代估计值 $\hat{x}_{i,k-1}(t-n)$ 构造 $\boldsymbol{\phi}(t)$ 的第 k 次迭代估计, 记作:

$$\hat{\boldsymbol{\phi}}_k(t) := \begin{bmatrix} \hat{\boldsymbol{\phi}}_k(t) \\ \boldsymbol{\psi}(t) \end{bmatrix} \in \mathbf{R}^{2n},$$

$$\hat{\boldsymbol{\phi}}_k(t) := \begin{bmatrix} -\hat{x}_{n,k-1}(t-n) \\ -\hat{x}_{n-1,k-1}(t-n) \\ \vdots \\ -\hat{x}_{1,k-1}(t-n) \end{bmatrix} \in \mathbf{R}^n.$$

用 $\hat{\boldsymbol{\phi}}_k(t)$ 定义 $\boldsymbol{\Phi}(t)$ 的估计:

$$\hat{\boldsymbol{\Phi}}_k(t) := [\hat{\boldsymbol{\phi}}_k(t-p+1), \hat{\boldsymbol{\phi}}_k(t-p+2), \dots, \hat{\boldsymbol{\phi}}_k(t)]^T \in \mathbf{R}^{p \times 2n}.$$

将式(218)–(219)右边未知 $\boldsymbol{\Phi}(t)$ 用其迭代估计 $\hat{\boldsymbol{\Phi}}_k(t)$ 代替, 未知的参数向量 \mathbf{b} 和 \mathbf{a} 用其前一次迭代估计 $\hat{\mathbf{b}}_{k-1}(t)$ 和 $\hat{\mathbf{a}}_{k-1}(t)$ 代替, 从而得到计算参数估计向量 $\hat{\mathbf{a}}_k(t)$ 和 $\hat{\mathbf{b}}_k(t)$ 的基于模型分解的最小二乘迭代算法 (model Decomposition based Least Squares Iterative algorithm, D-LSI 算法):

$$\hat{\mathbf{a}}_k(t) = [\hat{\boldsymbol{\Phi}}_k^T(t) \hat{\boldsymbol{\Phi}}_k(t)]^{-1} \hat{\boldsymbol{\Phi}}_k^T(t) [\mathbf{Y}(t) - \boldsymbol{\Psi}(t) \hat{\mathbf{b}}_{k-1}(t)], \quad (220)$$

$$\hat{\mathbf{b}}_k(t) = [\boldsymbol{\Psi}^T(t) \boldsymbol{\Psi}(t)]^{-1} \boldsymbol{\Psi}^T(t) [\mathbf{Y}(t) -$$

$$\hat{\boldsymbol{\Phi}}_k(t) \hat{\mathbf{a}}_{k-1}(t)], \quad (221)$$

$$\mathbf{Y}(t) = [y(t-p+1), y(t-p+2), \dots, y(t)]^T, \quad (222)$$

$$\hat{\boldsymbol{\Phi}}_k(t) = [\hat{\boldsymbol{\phi}}_k(t-p+1), \hat{\boldsymbol{\phi}}_k(t-p+2), \dots, \hat{\boldsymbol{\phi}}_k(t)]^T, \quad (223)$$

$$\boldsymbol{\Psi}(t) = [\boldsymbol{\psi}(t-p+1), \boldsymbol{\psi}(t-p+2), \dots, \boldsymbol{\psi}(t)]^T, \quad (224)$$

$$\hat{\boldsymbol{\phi}}_k(t) = [-\hat{x}_{n,k-1}(t-n), -\hat{x}_{n-1,k-1}(t-n), \dots, -\hat{x}_{1,k-1}(t-n)]^T, \quad (225)$$

$$\boldsymbol{\psi}(t) = [u(t-1), u(t-2), \dots, u(t-n)]^T, \quad (226)$$

$$\hat{\mathbf{a}}_k(t) = [\hat{a}_{1,k}(t), \hat{a}_{2,k}(t), \dots, \hat{a}_{n,k}(t)]^T, \quad (227)$$

$$\hat{\mathbf{b}}_k(t) = [\hat{b}_{1,k}(t), \hat{b}_{2,k}(t), \dots, \hat{b}_{n,k}(t)]^T. \quad (228)$$

令 $t=L, p=L$, 从而得到估计参数向量 \mathbf{a} 和 \mathbf{b} 的有限量测数据 D-LSI 算法:

$$\hat{\mathbf{a}}_k = [\hat{\boldsymbol{\Phi}}_k^T \hat{\boldsymbol{\Phi}}_k]^{-1} \hat{\boldsymbol{\Phi}}_k^T [\mathbf{Y} - \boldsymbol{\Psi} \hat{\mathbf{b}}_{k-1}], \quad (229)$$

$$\hat{\mathbf{b}}_k = [\boldsymbol{\Psi}^T \boldsymbol{\Psi}]^{-1} \boldsymbol{\Psi}^T [\mathbf{Y} - \hat{\boldsymbol{\Phi}}_k \hat{\mathbf{a}}_{k-1}], \quad (230)$$

$$\mathbf{Y} = [y(1), y(2), \dots, y(L)]^T, \quad (231)$$

$$\hat{\boldsymbol{\Phi}}_k = [\hat{\boldsymbol{\phi}}_k(1), \hat{\boldsymbol{\phi}}_k(2), \dots, \hat{\boldsymbol{\phi}}_k(L)]^T, \quad (232)$$

$$\boldsymbol{\Psi} = [\boldsymbol{\psi}(1), \boldsymbol{\psi}(2), \dots, \boldsymbol{\psi}(L)]^T, \quad (233)$$

$$\hat{\boldsymbol{\phi}}_k(t) = [-\hat{x}_{n,k-1}(t-n), -\hat{x}_{n-1,k-1}(t-n), \dots, -\hat{x}_{1,k-1}(t-n)]^T, \quad (234)$$

$$\boldsymbol{\psi}(t) = [u(t-1), u(t-2), \dots, u(t-n)]^T, \quad (235)$$

$$\hat{\mathbf{a}}_k = [\hat{a}_{1,k}, \hat{a}_{2,k}, \dots, \hat{a}_{n,k}]^T, \quad (236)$$

$$\hat{\mathbf{b}}_k = [\hat{b}_{1,k}, \hat{b}_{2,k}, \dots, \hat{b}_{n,k}]^T. \quad (237)$$

算法中使用的状态估计可采用基于迭代参数估计的状态观测器 (238)–(241) 或基于迭代参数估计的 Kalman 滤波状态估计算法 (258)–(263) 计算. 因此

1) 式(229)–(237)与式(238)–(241)构成了基于状态观测器的模型分解最小二乘迭代参数辨识算法 (O-D-LSI 算法);

2) 式(229)–(237)与式(258)–(263)构成了基于 Kalman 滤波状态估计的模型分解最小二乘迭代参数辨识算法 (KF-D-LSI 算法).

表 6 给出了有限量测数据 D-LSI 辨识算法的计算量.

5 基于迭代参数估计的状态估计算法

在上节讨论的状态空间系统的参数迭代辨识算法中, 信息向量中的未知状态变量用的是状态估计, 而状态估计是基于迭代参数估计计算的. 本节研究基于迭代参数估计的状态观测器和 Kalman 滤波状态估计算法.

5.1 基于迭代参数估计的状态观测器

如果能观测性规范型状态空间系统 (146)–

表 6 有限量测数据 D-LSI 辨识算法的计算量
Table 6 The computational efficiency of the D-LSI algorithm

表达式	乘法次数	加法次数
$\hat{a}_k = S_{1,k} \beta_{1,k} \in \mathbf{R}^n$	n^2	$n^2 - n$
$S_{1,k} := (S'_{1,k})^{-1} \in \mathbf{R}^{n \times n}$	n^3	$n^3 - n^2$
$S'_{1,k} := \hat{\Phi}_k^T \hat{\Phi}_k \in \mathbf{R}^{n \times n}$	$n^2 L$	$n^2 L - n^2$
$\beta_{1,k} := \hat{\Phi}_k^T E_1 \in \mathbf{R}^n$	nL	$nL - n$
$E_1 := Y - \Psi \hat{b}_{k-1} \in \mathbf{R}^L$	nL	nL
$\hat{b}_k = S_{2,k} \beta_{2,k} \in \mathbf{R}^n$	n^2	$n^2 - n$
$S_{2,k} := (S'_{2,k})^{-1} \in \mathbf{R}^{n \times n}$	n^3	$n^3 - n^2$
$S'_{2,k} := \Psi^T \Psi \in \mathbf{R}^{n \times n}$	$n^2 L$	$n^2 L - n^2$
$\beta_{2,k} := \Psi^T E_2 \in \mathbf{R}^n$	nL	$nL - n$
$E_2 := Y - \hat{\Phi}_k \hat{a}_{k-1} \in \mathbf{R}^L$	nL	nL
总数	$2n^3 + 2n^2 L + 2n^2 + 4nL$	$2n^3 + 2n^2 L + 4nL - 4n^2 - 4n$
总 flop 数	$4n^3 + 4n^2 L + 8nL - 2n^2 - 4n$	

(147)的参数已知时,可以采用观测器估计其状态.对于参数未知情形,虽然可以像在 2.1 节那样设计基于递推参数估计的状态观测器(46)–(49),但是这里不这样做,而是设计基于迭代参数估计的状态观测器.

参考基于递推参数估计的状态观测器算法(46)–(49),采用前述迭代算法获得的参数估计,可以得到基于迭代参数估计的状态观测器(iterative parameter estimation based state observer):

$$\hat{\mathbf{x}}_k(t+1) = \hat{\mathbf{A}}_k \hat{\mathbf{x}}_k(t) + \hat{\mathbf{b}}_k u(t), \quad \hat{\mathbf{x}}_k(1) = \mathbf{1}_n / p_0, \quad (238)$$

$$\hat{\mathbf{A}}_k = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -\hat{a}_{n,k} & -\hat{a}_{n-1,k} & -\hat{a}_{n-2,k} & \cdots & -\hat{a}_{1,k} \end{bmatrix}, \quad (239)$$

$$\hat{\mathbf{b}}_k = [\hat{b}_{1,k}, \hat{b}_{2,k}, \hat{b}_{3,k}, \cdots, \hat{b}_{n,k}]^T, \quad (240)$$

$$\hat{\mathbf{x}}_k(t) = [\hat{x}_{1,k}(t), \hat{x}_{2,k}(t), \cdots, \hat{x}_{n,k}(t)]^T. \quad (241)$$

类似地,可以给出基于迭代参数估计的闭环状态观测器(iterative parameter estimation based closed-loop state observer):

$$\hat{\mathbf{x}}_k(t+1) = \hat{\mathbf{A}}_k \hat{\mathbf{x}}_k(t) + \hat{\mathbf{b}}_k u(t) + \mathbf{K}_k [y(t) - \mathbf{c} \hat{\mathbf{x}}_k(t)], \quad \hat{\mathbf{x}}_k(1) = \mathbf{1}_n / p_0, \quad (242)$$

$$\hat{\mathbf{A}}_k = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -\hat{a}_{n,k} & -\hat{a}_{n-1,k} & -\hat{a}_{n-2,k} & \cdots & -\hat{a}_{1,k} \end{bmatrix}, \quad (243)$$

$$\hat{\mathbf{b}}_k = [\hat{b}_{1,k}, \hat{b}_{2,k}, \hat{b}_{3,k}, \cdots, \hat{b}_{n,k}]^T, \quad (244)$$

$$\hat{\mathbf{x}}_k(t) = [\hat{x}_{1,k}(t), \hat{x}_{2,k}(t), \cdots, \hat{x}_{n,k}(t)]^T. \quad (245)$$

其中 $\mathbf{K}_k \in \mathbf{R}^n$ 是观测器的增益向量,应该根据 $[\hat{\mathbf{A}}_k - \mathbf{K}_k \mathbf{c}]$ 的特征值在单位圆内进行选择.

将状态观测器分别与上面的梯度迭代辨识算法、基于模型分解的梯度迭代辨识算法、最小二乘迭代辨识算法和基于模型分解的最小二乘迭代辨识算法相结合,构成联合迭代参数与状态估计算法,可以同时估计系统的参数和状态,得到基于状态观测器的梯度迭代辨识算法、基于状态观测器的模型分解梯度迭代辨识算法、基于状态观测器的最小二乘迭代辨识算法和基于状态观测器的模型分解最小二乘迭代辨识算法.

下面针对基于状态观测器的最小二乘迭代辨识算法,给出其计算步骤、流程图和计算量.

5.2 基于状态观测器的最小二乘迭代辨识算法

有限量测数据 LSI 算法(208)–(212)和开环状态观测器(238)–(241)构成了计算参数估计向量 $\hat{\boldsymbol{\theta}}_k$ 的基于状态观测器的最小二乘迭代算法(Observer based Least Squares Iterative algorithm, O-LSI 算法),重写如下:

$$\hat{\boldsymbol{\theta}}_k = (\hat{\mathbf{H}}_k^T \hat{\mathbf{H}}_k)^{-1} \hat{\mathbf{H}}_k^T \mathbf{Y}, \quad k=1,2,3,\cdots, \quad (246)$$

$$\hat{\mathbf{H}}_k = [\hat{\boldsymbol{\varphi}}_k(1), \hat{\boldsymbol{\varphi}}_k(2), \cdots, \hat{\boldsymbol{\varphi}}_k(L)]^T, \quad (247)$$

$$\mathbf{Y} = [y(1), y(2), \cdots, y(L)]^T, \quad (248)$$

$$\hat{\boldsymbol{\varphi}}_k(t) = [-\hat{x}_{n,k-1}(t-n), -\hat{x}_{n-1,k-1}(t-n), \cdots, -\hat{x}_{1,k-1}(t-n), u(t-1), u(t-2), \cdots, u(t-n)]^T, \quad (249)$$

$$\hat{\mathbf{x}}_k(t+1) = \hat{\mathbf{A}}_k \hat{\mathbf{x}}_k(t) + \hat{\mathbf{b}}_k u(t), \quad \hat{\mathbf{x}}_k(1) = \mathbf{1}_n / p_0, \quad (250)$$

$$\hat{\mathbf{A}}_k = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -\hat{a}_{n,k} & -\hat{a}_{n-1,k} & -\hat{a}_{n-2,k} & \cdots & -\hat{a}_{1,k} \end{bmatrix}, \quad (251)$$

$$\hat{\mathbf{b}}_k = [\hat{b}_{1,k}, \hat{b}_{2,k}, \hat{b}_{3,k}, \cdots, \hat{b}_{n,k}]^T, \quad (252)$$

$$\hat{\boldsymbol{\theta}}_k = [\hat{a}_{1,k}, \hat{a}_{2,k}, \cdots, \hat{a}_{n,k}, \hat{b}_{1,k}, \hat{b}_{2,k}, \cdots, \hat{b}_{n,k}]^T, \quad (253)$$

$$\hat{\mathbf{x}}_k(t) = [\hat{x}_{1,k}(t), \hat{x}_{2,k}(t), \cdots, \hat{x}_{n,k}(t)]^T. \quad (254)$$

基于状态观测器的有限量测数据最小二乘迭代辨识算法(246)–(254)的计算步骤如下:

1) 置初值:令 $k=1, \hat{\mathbf{x}}_0(j) =$ 随机向量(保证 $\hat{\mathbf{H}}_1$ 可逆), $j = -(n-1), -(n-2), \cdots, L$, 给定辨识精度 $\varepsilon > 0$.

2) 收集输入输出数据 $\{u(t), y(t) : t=1, 2, \cdots, L\}$, 用式(248)构造 \mathbf{Y} .

- 3) 用式(249)构造 $\hat{\varphi}_k(t)$, 用式(247)构造 \hat{H}_k .
- 4) 用式(246)刷新参数估计 $\hat{\vartheta}_k$, 从式(253)的 $\hat{\vartheta}_k$ 中读取 $\hat{a}_{i,k}$ 和 $\hat{b}_{i,k}$.
- 5) 利用式(251) — (252)构造 \hat{A}_k 和 \hat{b}_k .
- 6) 令 $t=1, \hat{x}_k(1) = \mathbf{1}_n/p_0, p_0 = 10^6$.
- 7) 用式(250)计算 $\hat{x}_k(t+1)$.
- 8) 若 $t \leq L, t$ 增加 1, 转到第 7) 步.
- 9) 比较 $\hat{\vartheta}_k$ 和 $\hat{\vartheta}_{k-1}$, 若 $\|\hat{\vartheta}_k - \hat{\vartheta}_{k-1}\| \leq \varepsilon$, 计算结束, 获取迭代次数 k 和参数估计 $\hat{\vartheta}_k$; 反之, k 增加 1, 转到第 3) 步.

基于状态观测器的有限量测数据最小二乘迭代辨识算法的计算流程如图 3 所示, 计算量如表 7 所示.

5.3 基于迭代参数估计的 Kalman 滤波状态估计算法

对于能观测性状态空间系统(146) — (147), 当系统参数已知时, 也可采用与式(63) — (65) 类似的 Kalman 滤波算法来估计系统状态:

$$\hat{\mathbf{x}}(t+1) = \mathbf{A}\hat{\mathbf{x}}(t) + \mathbf{b}u(t) + \mathbf{L}(t)[y(t) - \mathbf{c}\hat{\mathbf{x}}(t)],$$

$$\hat{\mathbf{x}}(1) = \mathbf{1}_n/p_0, \quad (255)$$

$$\mathbf{L}(t) = \mathbf{A}\mathbf{P}(t)\mathbf{c}^T[1 + \mathbf{c}\mathbf{P}(t)\mathbf{c}^T]^{-1}, \quad (256)$$

$$\mathbf{P}(t+1) = \mathbf{A}\mathbf{P}(t)\mathbf{A}^T - \mathbf{L}(t)\mathbf{c}\mathbf{P}(t)\mathbf{A}^T, \quad \mathbf{P}(1) = \mathbf{I}_n. \quad (257)$$

当能观测性状态空间系统的参数未知时, 依然可以利用递推参数估计建立基于参数估计的 Kalman 滤波状态估计算法, 同样这里不这么做, 而是设计基于前述迭代参数估计的 Kalman 滤波状态估计算法.

当系统参数矩阵 $\mathbf{A} \in \mathbf{R}^{n \times n}$ 和参数向量 $\mathbf{b} \in \mathbf{R}^n$ 未知时, 为了获得系统的状态估计, 可用第 k 次迭代参数估计 $\hat{\mathbf{A}}_k$ 和 $\hat{\mathbf{b}}_k$ 代替式(255) — (257) 中 \mathbf{A} 和 \mathbf{b} , 便得到状态空间系统(146) — (147) 的基于迭代参数估计的 Kalman 滤波状态估计算法:

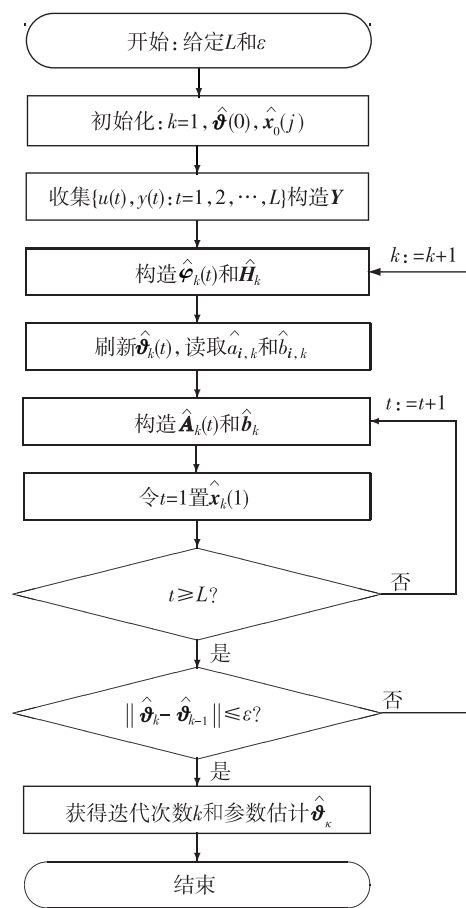


图 3 计算基于状态观测器的最小二乘迭代参数估计 $\hat{\vartheta}_k$ 的流程

Fig. 3 The flowchart of computing the state observer based LSI estimate $\hat{\vartheta}_k$

$$\hat{\mathbf{x}}_k(t+1) = \hat{\mathbf{A}}_k\hat{\mathbf{x}}_k(t) + \hat{\mathbf{b}}_k u(t) + \mathbf{L}_k(t)[y(t) - \mathbf{c}\hat{\mathbf{x}}_k(t)],$$

$$\hat{\mathbf{x}}_k(1) = \mathbf{1}_n/p_0, \quad (258)$$

$$\mathbf{L}_k(t) = \hat{\mathbf{A}}_k\mathbf{P}_k(t)\mathbf{c}^T[1 + \mathbf{c}\mathbf{P}_k(t)\mathbf{c}^T]^{-1}, \quad (259)$$

$$\mathbf{P}_k(t+1) = \hat{\mathbf{A}}_k\mathbf{P}_k(t)\hat{\mathbf{A}}_k^T - \mathbf{L}_k(t)\mathbf{c}\mathbf{P}_k(t)\hat{\mathbf{A}}_k^T,$$

表 7 基于状态观测器的最小二乘迭代算法的计算量

Table 7 The computational efficiency of the O-LSI algorithm

表达式	乘法次数	加法次数
$\hat{\vartheta}_k = \mathbf{S}_k \boldsymbol{\beta}_k \in \mathbf{R}^{2n}$	$4n^2$	$4n^2 - 2n$
$\mathbf{S}_k := (\mathbf{S}'_k)^{-1} \in \mathbf{R}^{2n \times 2n}$	$8n^3$	$8n^3 - 4n^2$
$\mathbf{S}'_k := \mathbf{H}_k^T \mathbf{H}_k \in \mathbf{R}^{2n \times 2n}$	$4n^2 L$	$4n^2 L - 4n^2$
$\boldsymbol{\beta}_k = \mathbf{H}_k^T \mathbf{Y} \in \mathbf{R}^{2n}$	$2nL$	$2nL - 2n$
$\hat{\mathbf{x}}_k(t+1) = \hat{\mathbf{A}}_k \hat{\mathbf{x}}_k(t) + \hat{\mathbf{b}}_k u(t) \in \mathbf{R}^n$	$2nL$	$(2n-1)L$
总数	$8n^3 + 4n^2 L + 4nL + 4n^2$	$8n^3 + 4n^2 L + 4nL - 4n^2 - 4n - L$
总 flop 数	$16n^3 + 8n^2 L + 8nL - 4n - L$	

$$P_k(1) = I_n, \quad (260)$$

$$\hat{A}_k = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -\hat{a}_{n,k} & -\hat{a}_{n-1,k} & -\hat{a}_{n-2,k} & \cdots & -\hat{a}_{1,k} \end{bmatrix}, \quad (261)$$

$$\hat{b}_k = [\hat{b}_{1,k}, \hat{b}_{2,k}, \hat{b}_{3,k}, \dots, \hat{b}_{n,k}]^T, \quad (262)$$

$$\hat{x}_k(t) = [\hat{x}_{1,k}(t), \hat{x}_{2,k}(t), \dots, \hat{x}_{n,k}(t)]^T. \quad (263)$$

式(263)中 $\hat{x}_k(t)$ 是对应于第 k 次迭代参数估计参数 \hat{A}_k 和 \hat{b}_k 时的状态估计.这个基于迭代参数估计的Kalman滤波状态估计算法计算量极大,因为对每次迭代获得的参数估计,状态估计都要从 $t=1$ 到 $t=L$ 计算一遍(L 为数据长度).由于最小二乘迭代参数估计算法比梯度迭代参数估计算法收敛快得多,所以基于最小二乘迭代参数估计的Kalman滤波状态估计算法的计算量比基于梯度迭代参数估计的Kalman滤波状态估计算法的计算量要小得多.

Kalman滤波状态估计算法中的参数估计可采用前述的一些迭代辨识算法计算.例如:

1) 式(164)—(169)与式(258)—(263)构成了基于梯度迭代参数估计的Kalman滤波状态估计算法(GI-KF算法);

2) 式(190)—(200)与式(258)—(263)构成了基于模型分解的梯度迭代参数估计的Kalman滤波状态估计算法(D-GI-KF算法);

3) 式(208)—(212)与式(258)—(263)构成了基于最小二乘迭代参数估计的Kalman滤波状态估计算法(LSI-KF算法);

4) 式(229)—(237)与式(258)—(263)构成了基于模型分解的最小二乘迭代参数估计的Kalman滤波状态估计算法(D-LSI-KF算法).

下面给出基于Kalman滤波状态估计的D-LSI参数辨识算法的计算步骤、流程图(图4)和计算量.

5.4 基于Kalman滤波状态估计的D-LSI辨识算法

基于模型分解的有限量测数据最小二乘迭代辨识算法(229)—(237)与Kalman滤波状态估计算法(258)—(263)相结合,构成了基于Kalman滤波状态估计的模型分解最小二乘迭代算法(Kalman Filtering based model Decomposition Least Squares Iterative algorithm, KF-D-LSI算法):

$$\hat{a}_k = [\hat{\Phi}_k^T \hat{\Phi}_k]^{-1} \hat{\Phi}_k^T [Y - \Psi \hat{b}_{k-1}], \quad (264)$$

$$\hat{b}_k = [\Psi^T \Psi]^{-1} \Psi^T [Y - \hat{\Phi}_k \hat{a}_{k-1}], \quad (265)$$

$$Y = [y(1), y(2), \dots, y(L)]^T, \quad (266)$$

$$\hat{\Phi}_k = [\hat{\phi}_k(1), \hat{\phi}_k(2), \dots, \hat{\phi}_k(L)]^T, \quad (267)$$

$$\Psi = [\psi(1), \psi(2), \dots, \psi(L)]^T, \quad (268)$$

$$\hat{\phi}_k(t) = [-\hat{x}_{n,k-1}(t-n), -\hat{x}_{n-1,k-1}(t-n), \dots, -\hat{x}_{1,k-1}(t-n)]^T, \quad (269)$$

$$\psi(t) = [u(t-1), u(t-2), \dots, u(t-n)]^T, \quad (270)$$

$$\hat{x}_k(t+1) = \hat{A}_k \hat{x}_k(t) + \hat{b}_k u(t) + L_k(t) [y(t) - \mathbf{c} \hat{x}_k(t)], \quad (271)$$

$$\hat{x}_k(1) = \mathbf{1}_n / p_0, \quad (271)$$

$$L_k(t) = \hat{A}_k P_k(t) \mathbf{c}^T [1 + \mathbf{c} P_k(t) \mathbf{c}^T]^{-1}, \quad (272)$$

$$P_k(t+1) = \hat{A}_k P_k(t) \hat{A}_k^T - L_k(t) \mathbf{c} P_k(t) \hat{A}_k^T, \quad (273)$$

$$P_k(1) = I_n, \quad (273)$$

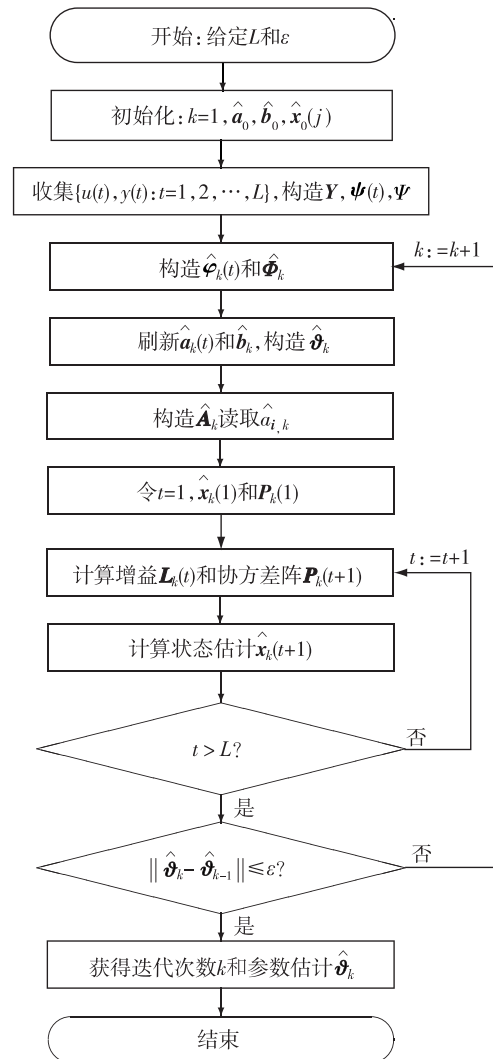


图4 基于Kalman滤波状态估计的模型分解最小二乘迭代参数估计 $\hat{\theta}_k$ 的流程

Fig. 4 The flowchart of computing the Kalman filtering based D-LSI estimate $\hat{\theta}_k$

$$\hat{\mathbf{A}}_k = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -\hat{a}_{n,k} & -\hat{a}_{n-1,k} & -\hat{a}_{n-2,k} & \cdots & -\hat{a}_{1,k} \end{bmatrix}, \quad (274)$$

$$\hat{\mathbf{x}}_k(t) = [\hat{x}_{1,k}(t), \hat{x}_{2,k}(t), \dots, \hat{x}_{n,k}(t)]^T, \quad (275)$$

$$\hat{\mathbf{a}}_k = [\hat{a}_{1,k}, \hat{a}_{2,k}, \dots, \hat{a}_{n,k}]^T, \quad (276)$$

$$\hat{\mathbf{b}}_k = [\hat{b}_{1,k}, \hat{b}_{2,k}, \dots, \hat{b}_{n,k}]^T, \quad (277)$$

$$\hat{\boldsymbol{\theta}}_k = \begin{bmatrix} \hat{\mathbf{a}}_k \\ \hat{\mathbf{b}}_k \end{bmatrix}. \quad (278)$$

有限量测数据下,基于 Kalman 滤波状态估计的模型分解最小二乘迭代辨识算法(264)–(278) 计算流程如图 4 所示,计算量如表 8 所示.计算参数估计与状态估计的步骤如下:

1) 置初值: 令 $k = 1$, $\hat{\mathbf{a}}_0 = \mathbf{1}_n/p_0$, $\hat{\mathbf{b}}_0 = \mathbf{1}_n/p_0$, $\hat{\mathbf{x}}_0(j)$ 是随机向量(保证 $\hat{\boldsymbol{\Phi}}_1$ 可逆), $j = -(n-1), -(n-2), \dots, L$, 给定辨识精度 $\varepsilon > 0$.

2) 收集输入输出数据 $\{u(t), y(t) : t = 1, 2, \dots, L\}$, 用式(266)构造 \mathbf{Y} .

3) 用式(270)构造 $\boldsymbol{\psi}(t)$, 用式(268)构造 $\boldsymbol{\Psi}$.

4) 用式(269)构造 $\hat{\boldsymbol{\Phi}}_k(t)$, 用式(267)构造 $\hat{\boldsymbol{\Phi}}_k$.

5) 根据式(264)和(265)刷新参数估计 $\hat{\mathbf{a}}_k$ 和 $\hat{\mathbf{b}}_k$.

6) 从式(276)的 $\hat{\mathbf{a}}_k$ 中读取 $\hat{a}_{i,k}$, 利用式(274)构造 $\hat{\mathbf{A}}_k$.

7) 令 $t = 1$, $\hat{\mathbf{x}}_k(1) = \mathbf{1}_n/p_0$, $\mathbf{P}_k(1) = \mathbf{I}_n$.

8) 利用式(272)计算 $\mathbf{L}_k(t)$, 用式(273)计算 $\mathbf{P}_k(t+1)$, 用式(271)计算 $\hat{\mathbf{x}}_k(t+1)$.

9) 若 $t \leq L$, t 增加 1, 转到第 8) 步.

10) 用式(278)构成 $\hat{\boldsymbol{\theta}}_k$. 比较 $\hat{\boldsymbol{\theta}}_k$ 和 $\hat{\boldsymbol{\theta}}_{k-1}$, 若 $\|\hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_{k-1}\| \leq \varepsilon$, 终止循环(在仿真中可以把 $\|\hat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}\| \leq \varepsilon$ 作为终止循环条件), 获取迭代次数 k 和参数估计 $\hat{\boldsymbol{\theta}}_k$; 反之, k 增加 1, 转到第 3) 步.

6 结语

针对观测器规范型状态空间系统, 研究和提出了基于状态观测器(Kalman 滤波状态估计)的随机梯度辨识方法、递推最小二乘辨识方法、多新息随机梯度辨识方法和多新息最小二乘辨识方法, 以及基于状态观测器(Kalman 滤波状态估计)的模型分解

表 8 基于 Kalman 滤波状态估计的模型分解最小二乘迭代辨识算法的计算量
Table 8 The computational efficiency of the Kalman Filtering based D-LSI algorithm

表达式	乘法次数	加法次数
$\hat{\mathbf{a}}_k = \mathbf{S}_{1,k} \boldsymbol{\beta}_{1,k} \in \mathbf{R}^n$	n^2	$n^2 - n$
$\mathbf{S}_{1,k} := (\mathbf{S}'_{1,k})^{-1} \in \mathbf{R}^{n \times n}$	n^3	$n^3 - n^2$
$\mathbf{S}'_{1,k} := \hat{\boldsymbol{\Phi}}_k^T \hat{\boldsymbol{\Phi}}_k \in \mathbf{R}^{n \times n}$	$n^2 L$	$n^2 L - n^2$
$\boldsymbol{\beta}_{1,k} := \hat{\boldsymbol{\Phi}}_k^T \mathbf{E}_1 \in \mathbf{R}^n$	nL	$nL - n$
$\mathbf{E}_1 := \mathbf{Y} - \boldsymbol{\Psi} \hat{\boldsymbol{\theta}}_{k-1} \in \mathbf{R}^L$	nL	nL
$\hat{\mathbf{b}}_k = \mathbf{S}_{2,k} \boldsymbol{\beta}_{2,k} \in \mathbf{R}^n$	n^2	$n^2 - n$
$\mathbf{S}_{2,k} := (\mathbf{S}'_{2,k})^{-1} \in \mathbf{R}^{n \times n}$	n^3	$n^3 - n^2$
$\mathbf{S}'_{2,k} := \boldsymbol{\Psi}^T \boldsymbol{\Psi} \in \mathbf{R}^{n \times n}$	$n^2 L$	$n^2 L - n^2$
$\boldsymbol{\beta}_{2,k} := \boldsymbol{\Psi}^T \mathbf{E}_2 \in \mathbf{R}^n$	nL	$nL - n$
$\mathbf{E}_2 := \mathbf{Y} - \hat{\boldsymbol{\Phi}}_k \hat{\mathbf{a}}_{k-1} \in \mathbf{R}^L$	nL	nL
$\mathbf{x}_k(t+1) = \hat{\mathbf{A}}_k \hat{\mathbf{x}}_k(t) + \hat{\mathbf{b}}_k u(t) + \mathbf{L}_k(t) e(t) [y(t) - \mathbf{c} \hat{\mathbf{x}}_k(t)] \in \mathbf{R}^n$	$2nL$	$(2n-1)L$
$\mathbf{L}_k(t) = \hat{\mathbf{A}}_k [\mathbf{P}_k(t) \mathbf{c}^T] / [1 + \mathbf{c} \mathbf{P}_k(t) \mathbf{c}^T] \in \mathbf{R}^n$	nL	$(n+1)L$
$\mathbf{L}_k(t) = \hat{\mathbf{A}}_k [\mathbf{P}_k(t) \mathbf{c}^T] / [1 + \mathbf{c} \mathbf{P}_k(t) \mathbf{c}^T] \in \mathbf{R}^n$	$2nL$	nL
$\mathbf{P}_k(t+1) = \hat{\mathbf{A}}_k \boldsymbol{\zeta}(t) - \mathbf{L}_k(t) \mathbf{c} \boldsymbol{\zeta}(t) \in \mathbf{R}^{n \times n}$	$(n^2 + 2n)L$	$(n^2 + n - 1)L$
$\boldsymbol{\zeta}(t) := \mathbf{P}_k(t) \hat{\mathbf{A}}_k^T \in \mathbf{R}^{n \times n}$	$n^2 L$	$(n^2 - n)L$
总数	$4n^2 L + 11nL + 2n^3 + 2n^2$	$4n^2 L + 8nL - L + 2n^3 - 2n^2 - 4n$
总 flop 数	$4n^3 + 8n^2 L + 19nL - L - 4n$	

随机梯度辨识方法、模型分解递推最小二乘辨识方法、模型分解多新息随机梯度辨识方法和模型分解多新息最小二乘辨识方法,给出了基于状态观测器的递推最小二乘参数辨识算法和基于 Kalman 滤波状态估计的多新息最小二乘参数辨识算法的计算步骤、流程图和计算量。

针对能观测性规范型状态空间系统,研究和提出了基于状态观测器(Kalman 滤波状态估计)的梯度迭代辨识方法、模型分解梯度迭代辨识方法、最小二乘迭代辨识方法和模型分解最小二乘迭代辨识方法,给出了基于状态观测器的最小二乘迭代参数辨识算法和基于 Kalman 滤波状态估计的模型分解最小二乘迭代参数辨识算法的计算步骤、流程图和计算量。

值得指出的是上述方法中,两种规范型状态空间系统模型互换,其联合状态估计和参数估计算法仍然适用。本文针对输出方程存在白噪声干扰、状态不含干扰时的状态空间系统,研究了联合状态和参数估计,但输出存在相关噪声干扰和(或)状态存在干扰时的状态空间系统的联合参数与状态估计问题尚未提及,实际应用中干扰噪声极为复杂,这是尚需解决的课题。

参考文献

References

- [1] 丁锋.系统辨识新论[M].北京:科学出版社,2013
DING Feng. System identification: New theory and methods[M]. Beijing: Science Press, 2013
- [2] 丁锋.系统辨识:辨识方法性能分析[M].北京:科学出版社,2014
DING Feng. System identification: Performance analysis for identification methods [M]. Beijing: Science Press, 2014
- [3] 丁锋.系统辨识(1):辨识导引[J].南京信息工程大学学报:自然科学版,2011,3(1):1-22
DING Feng. System identification. Part A: Introduction to the identification[J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2011, 3(1): 1-22
- [4] 丁锋.系统辨识(2):系统描述的基本模型[J].南京信息工程大学学报:自然科学版,2011,3(2):97-117
DING Feng. System identification. Part B: Basic models for system description [J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2011, 3(2): 97-117
- [5] 丁锋.系统辨识(3):辨识精度与辨识基本问题[J].南京信息工程大学学报:自然科学版,2011,3(3):193-226
DING Feng. System identification. Part C: Identification

丁锋,等.规范状态空间系统辨识方法.

DING Feng, et al. Identification methods for canonical state space systems.

accuracy and basic problems [J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2011, 3(3): 193-226

- [6] 丁锋.系统辨识(4):辅助模型辨识思想与方法[J].南京信息工程大学学报:自然科学版,2011,3(4):289-318
DING Feng. System identification. Part D: Auxiliary model identification idea and methods [J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2011, 3(4): 289-318
- [7] 丁锋.系统辨识(5):迭代搜索原理与辨识方法[J].南京信息工程大学学报:自然科学版,2011,3(6):481-510
DING Feng. System identification. Part E: Iterative search principle and identification methods [J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2011, 3(6): 481-510
- [8] 丁锋.系统辨识(6):多新息辨识理论与方法[J].南京信息工程大学学报:自然科学版,2012,4(1):1-28
DING Feng. System identification. Part F: Multi-innovation identification theory and methods [J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2012, 4(1): 1-28
- [9] 丁锋.系统辨识(7):递阶辨识原理与方法[J].南京信息工程大学学报:自然科学版,2012,4(2):97-124
DING Feng. System identification. Part G: Hierarchical identification principle and methods [J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2012, 4(2): 97-124
- [10] 丁锋.系统辨识(8):耦合辨识概念与方法[J].南京信息工程大学学报:自然科学版,2012,4(3):193-212
DING Feng. System identification. Part H: Coupling identification concept and methods [J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2012, 4(3): 193-212
- [11] 丁锋.辨识方法的计算效率(1):递推算法[J].南京信息工程大学学报:自然科学版,2012,4(4):289-300
DING Feng. Computational efficiency of the identification methods. Part A: Recursive algorithms [J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2012, 4(4): 289-300
- [12] 丁锋.辨识方法的计算效率(2):迭代算法[J].南京信息工程大学学报:自然科学版,2012,4(5):385-401
DING Feng. Computational efficiency of the identification methods. Part B: Iterative algorithms [J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2012, 4(5): 385-401
- [13] 丁锋.辨识方法的计算效率(3):信息向量耦合算法[J].南京信息工程大学学报:自然科学版,2012,4(6):481-495
DING Feng. Computational efficiency of the identification methods. Part C: Coupled information vector algorithms [J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2012, 4(6): 481-495
- [14] 丁锋,汪菲菲.多元系统耦合多新息随机梯度类辨识

- 方法[J].南京信息工程大学学报:自然科学版,2014,6(1):1-16
DING Feng, WANG Feifei. Coupled multi-innovation stochastic gradient type identification methods for multivariate systems[J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2014, 6(1): 1-16
- [15] 丁锋,汪菲菲,汪学海.多元伪线性回归系统部分耦合多新息随机梯度类辨识方法[J].南京信息工程大学学报:自然科学版,2014,6(2):97-112
DING Feng, WANG Feifei, WANG Xuehai. Partially Coupled multi-innovation stochastic gradient type identification methods for multivariate pseudo-linear regressive systems[J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2014, 6(2): 97-112
- [16] 丁锋,汪菲菲,汪学海.类多变量输出误差系统的耦合多新息辨识方法[J].南京信息工程大学学报:自然科学版,2014,6(3):193-210
DING Feng, WANG Feifei, WANG Xuehai. Coupled multi-innovation identification methods for multivariable output-error-like systems [J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2014, 6(3): 193-210
- [17] 丁锋,汪菲菲,汪学海.多变量方程误差类系统的部分耦合迭代辨识方法[J].南京信息工程大学学报:自然科学版,2014,6(4):289-305
DING Feng, WANG Feifei, WANG Xuehai. Partially coupled iterative identification methods for multivariable equation error type systems [J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2014, 6(4): 289-305
- [18] 丁锋,王艳娇.类多变量方程误差类系统的递阶多新息辨识方法[J].南京信息工程大学学报:自然科学版,2014,6(5):385-404
DING Feng, WANG Yanjiao. Hierarchical multi-innovation identification methods for multivariable equation-error-like type systems [J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2014, 6(5): 385-404
- [19] 丁锋,萧德云.多变量系统状态空间模型的递阶辨识[J].控制与决策,2005,20(8):848-853,859
DING Feng, XIAO Deyun. Hierarchical identification of state space models for multivariable systems [J]. Control and Decision, 2005, 20(8): 848-853, 859
- [20] Ding F, Chen T. Hierarchical identification of lifted state-space models for general dual-rate systems [J]. IEEE Transactions on Circuits and Systems. I: Regular Papers, 2005, 52(6): 1179-1187
- [21] Ding F, Chen T. Hierarchical gradient-based identification of multivariable discrete-time systems [J]. Automatica, 2005, 41(2): 315-325
- [22] Ding F, Chen T. Hierarchical least squares identification methods for multivariable systems [J]. IEEE Transactions on Automatic Control, 2005, 50(3): 397-402
- [23] Ding F, Liu X G, Chu J. Gradient based and least squares based iterative algorithms for Hammerstein systems using the hierarchical identification principle [J]. IET Control Theory and Applications, 2013, 7(2): 176-184.
- [24] Ding F. Hierarchical multi-innovation stochastic gradient algorithm for Hammerstein nonlinear system modeling [J]. Applied Mathematical Modelling, 2013, 37(4): 1694-1704
- [25] 丁锋,陈通文,萧德云.一般双率系统状态空间模型及其递阶辨识[J].自动化学报,2004,30(5):652-663
DING Feng, CHEN Tongwen, XIAO Deyun. State-space modeling and hierarchical identification for general dual-rate stochastic systems [J]. Acta Automatica Sinica, 2004, 30(5): 652-663
- [26] 丁锋,陈通文,萧德云.非均匀周期采样多率系统的一种递阶辨识方法[J].电子学报,2004,32(9):1414-1420
DING Feng, CHEN Tongwen, XIAO Deyun. Hierarchical identification of non-uniformly periodically sampled multirate systems [J]. Acta Electronica Sinica, 2004, 32(9): 1414-1420
- [27] 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
- [28] Ding F. State filtering and parameter identification for state space systems with scarce measurements [J]. Signal Processing, 2014, 104: 369-380
- [29] Ding F. Decomposition based fast least squares algorithm for output error systems [J]. Signal Processing, 2013, 93(5): 1235-1242
- [30] Ding F, Duan H H. Two-stage parameter estimation algorithms for Box-Jenkins systems [J]. IET Signal Processing, 2013, 7(2): 176-184
- [31] Ding F. Combined state and least squares parameter estimation algorithms for dynamic systems [J]. Applied Mathematical Modelling, 2014, 38(1): 403-412
- [32] Ma X Y, Ding F. Gradient-based parameter identification algorithms for observer canonical state space systems using state estimates [J]. Circuits, Systems & Signal Processing, 2015, doi: 10. 1007/s00034-014-9911-5
- [33] Ding F, Chen T. Performance analysis of multi-innovation gradient type identification methods [J]. Automatica, 2007, 43(1): 1-14
- [34] Ding F. Several multi-innovation identification methods [J]. Digital Signal Processing, 2010, 20(4): 1027-1039
- [35] Ding F, Liu X P, Liu G. Multi-innovation least squares identification for system modeling [J]. IEEE Transactions on Systems, Man, and Cybernetics. Part B: Cybernetics, 2010, 40(3): 767-778
- [36] 丁锋,刘艳君,于丽.基于卡尔曼滤波思想的时变增益最优观测器设计[J].科学技术与工程,2008,8(15):4346-4348
DING Feng, LIU Yanjun, YU Li. Time-varying gain observers based on the Kalman filtering [J]. Science Technology and Engineering, 2008, 8(15): 4346-4348
- [37] 丁锋,杨家本.大系统的递阶辨识[J].自动化学报,1999,25(5):647-654
DING Feng, YANG Jiaben. Hierarchical identification of large scale systems [J]. Acta Automatica Sinica, 1999, 25(5): 647-654
- [38] Ma X Y, Ding F. Recursive and iterative least squares pa-

- parameter estimation algorithms for observability canonical state space systems[J]. Journal of the Franklin Institute, 2015, doi:10.1016/j.jfranklin.2014.10.024
- [39] Ding F, Liu X M, Chen H B, et al. Hierarchical gradient based and hierarchical least squares based iterative parameter identification for CARARMA systems[J]. Signal Processing, 2014, 97:31-39
- [40] Ding F, Liu X P, Liu G. Gradient based and least-squares based iterative identification methods for OE and OEMA systems[J]. Digital Signal Processing, 2010, 20(3): 664-677

Identification methods for canonical state space systems

DING Feng^{1,2,3} MA Xingyun¹

1 School of Internet of Things Engineering, Jiangnan University, Wuxi 214122

2 Control Science and Engineering Research Center, Jiangnan University, Wuxi 214122

3 Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi 214122

Abstract Because the state space model contains both the unknown states and the unknown parameters, and they involve the nonlinear product relations, which makes the identification problem more complicated. In order to solve this problem, this paper studies the combined state and parameter estimation methods for canonical state space systems. The interactive estimation theory is used to derive the combined state and parameter estimation algorithms by means of the recursive or iterative scheme. When computing the parameter estimates, the unknown states in the information vector of the identification algorithms are replaced with their estimates, the obtained parameter estimates are used to design the parameter estimates based observer or the parameter estimates based Kalman filtering algorithm to estimate the states of the systems. They form an interactive estimation process (a hierarchical estimation process). Along this line, from the recursive scheme or the iterative scheme, this paper presents the observer based or the Kalman filtering based stochastic gradient (SG) identification algorithm, recursive least squares (LS) identification algorithm, multi-innovation SG algorithm, multi-innovation LS identification algorithm, and the model decomposition based identification methods. Finally, the computational efficiency, the computational steps and the flowcharts of some typical algorithms are discussed.

Key words parameter estimation; recursive identification; iterative identification; least squares; gradient search; state observer; Kalman filter; state estimation; model decomposition; state space system