



辅助模型辨识方法(2):输入非线性输出误差系统

摘要

针对具有已知基的输入非线性输出误差系统,提出了基于过参数化模型的辅助模型递推辨识方法和辅助模型递阶辨识方法,提出了基于关键项分离的辅助模型递推辨识方法、基于关键项分离的辅助模型两阶段辨识方法和辅助模型三阶段辨识方法,提出了基于双线性参数模型分解的辅助模型随机梯度算法和基于双线性参数模型分解的辅助模型递推最小二乘算法,并给出了几个典型辨识算法的计算量、计算步骤.这些算法的收敛性分析都是需要研究的辨识课题.

关键词

参数估计;递推辨识;梯度搜索;最小二乘;过参数化模型;关键项分离;模型分解;辅助模型辨识思想;递阶辨识原理;输入非线性系统

中图分类号 TP273

文献标志码 A

收稿日期 2016-03-30

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

系统的数学模型是研究系统动态特性和行为品质的前提,是一切控制问题的基础.获取系统数学模型的一种有效手段就是系统辨识.近几十年来,系统辨识及其参数估计得到了广泛研究,形成了诸如最小二乘类、梯度搜索类、牛顿方法类、极大似然类等递推和迭代辨识方法,这些方法源于严格的理论推导,历经仿真和实践的双重检验,已被人们奉为经典^[1-3].

近年来,辅助模型辨识思想、多新息辨识理论、递阶辨识原理、滤波辨识理念、耦合辨识概念等新的辨识思想、辨识理论、辨识概念相继问世,为辨识领域的研究注入了新的活力,并与传统的辨识方法相互交融,促使辨识新方法层出不穷,精彩纷呈,因而也引发了笔者对辨识的重新思考,于是自2011年起在《南京信息工程大学学报》上连载了一系列的文章,就系统辨识的一些基本问题:辨识的定义、方法、目的、步骤^[4],基本模型类的描述^[5],辨识精度与可辨识性^[6]以及辨识算法的计算效率^[7-9]等进行了深刻而独到的阐述;就人们熟知而又容易混淆的两个概念:递推算法和迭代算法从本质上进行了剖析、区分和定义^[10];就辅助模型辨识思想^[11]、多新息辨识理论^[12]、递阶辨识原理^[13]、耦合辨识概念^[14]、滤波辨识理念^[15-16]进行了精心的提炼与总结.《系统辨识新论》一书就是这些科学思考的结晶^[1].一些典型递推辨识方法的收敛性分析则汇聚在《系统辨识学术专著丛书》第3分册《系统辨识——辨识方法性能分析》^[2]一书中.多新息辨识方法则在《系统辨识学术专著丛书》第6分册《系统辨识——多新息辨识理论与方法》^[3]一书中进行了专门讨论.后继将陆续出版《系统辨识学术专著丛书》第2分册《系统辨识——系统辨识方法论文》、第4分册《系统辨识——辅助模型辨识思想与方法》、第5分册《系统辨识——迭代辨识原理与搜索方法》、第7分册《系统辨识——递阶辨识原理与方法》、第8分册《系统辨识——耦合辨识概念与方法》等.

针对一些典型非线性系统,本文第一作者及其所带领的研究生团队已经做了大量研究,提出了一些有价值的研究成果^[17-25],其中多篇论文入选“ESI全球高被引论文”、“中国百篇最具影响国际学术论文”^[23-25],文献[26]获得2015年《IET Control Theory and Applications》杂志的最佳论文奖“Premium (Best Paper) Awards”.

文献[27]详细研究了输入非线性方程误差自回归系统的多新息

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

辨识方法,基于过参数化辨识模型和关键项分离辨识模型,研究了输入非线性方程误差系统的随机梯度算法、递推最小二乘算法、多新息辨识方法等^[28]. 本文基于辅助模型辨识思想和递阶辨识原理,分别使用过参数化、关键项分离、双线性模型分解,研究输入非线性输出误差系统辨识问题,提出了基于过参数化的辅助模型辨识方法、基于关键项分离的辅助模型辨识方法、基于双线性参数模型分解的辅助模型辨识方法等.

1 基于过参数化模型的辅助模型辨识方法

本节研究输入非线性输出误差系统的基于过参数化模型的辅助模型递推辨识方法,包括基于过参数化模型的辅助模型随机梯度算法和基于过参数化模型的辅助模型递推最小二乘算法.

1.1 系统描述与过参数化辨识模型

考虑输入非线性输出误差模型(Input Nonlinear Output-Error model, IN-OE 模型)描述的非线性随机系统,系统的框图如图 1 所示.该系统由一个静态非线性环节串联一个线性输出误差子系统构成,其线性输出误差子系统为

$$y(t) = \frac{B(z)}{A(z)} \bar{u}(t) + v(t), \quad (1)$$

非线性环节的输出 $\bar{u}(t)$ 是一组系数为 c_j 的已知基函数 $f_j(u(t))$ 的线性组合,即

$$\bar{u}(t) = \sum_{j=1}^m c_j f_j(u(t)) = \mathbf{f}^T(u(t)) \mathbf{c} = \mathbf{c}^T \mathbf{f}(u(t)), \quad (2)$$

$$\mathbf{f}(t) := \mathbf{f}(u(t)) = [f_1(u(t)), f_2(u(t)), \dots, f_m(u(t))]^T \in \mathbf{R}^m, \quad (3)$$

$$\mathbf{c} := [c_1, c_2, \dots, c_m]^T \in \mathbf{R}^m, \quad (4)$$

其中 $u(t) \in \mathbf{R}$ 和 $y(t) \in \mathbf{R}$ 分别是系统的输入和输出, $v(t) \in \mathbf{R}$ 是零均值白噪声, $\bar{u}(t) \in \mathbf{R}$ 是非线性环节的输出,也是线性环节的输入,是不可测的内部变量, $x(t)$ 是系统的真实输出(true output)或无噪输出(noise-free output),其测量值是含噪输出 $y(t)$, $A(z)$ 和 $B(z)$ 是单位后移算子 z^{-1} [$z^{-1}y(t) = y(t-1)$] 的多项式:

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

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

非线性基函数 $f_j(\cdot)$ 可以是多项式、三角函数、指数函数、分段线性函数等, $\mathbf{f}(t) \in \mathbf{R}^m$ 和 $\mathbf{c} \in \mathbf{R}^m$ 是非线性环节的信息向量和参数向量, a_i 和 b_i 是线性动态子系统的参数. 辨识的目的是基于量测数据 $u(t)$ 和

$y(t)$, 研究估计参数 a_i , b_i 和 c_i 的辨识方法.

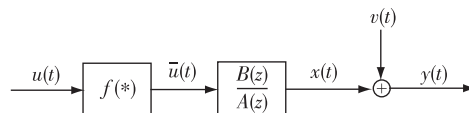


图 1 输入非线性输出误差系统

Fig. 1 The IN-OE system

从图 1 中可以看出,非线性函数 $f(\cdot)$ 与线性动态子系统 $G(z) := \frac{B(z)}{A(z)}$ 串联, $A(z)$ 是首 1 多项式,如果中间变量 $\bar{u}(t)$ 不可测,那么 $f(\cdot)$ 与 $B(z)$ 的参数不能由 $u(t)$ 和 $y(t)$ 唯一确定,因为对于任意非零常数 α , 对于非线性函数 $f'(\cdot) := \alpha f(\cdot)$ 和多项式 $B'(z) := B(z)/\alpha$, 式(1)–(2)具有相同的输入输出关系,即从 $u(t)$ 到 $y(t)$ 的映射关系是不变的. 因此,为了保证参数辨识的唯一性,必须对 $f(\cdot)$ 或 $B(z)$ 的系数进行规范化或归一化处理. 典型的归一化假设有:

(A1) 假设 $f(\cdot)$ 或 $B(z)$ 的某个非零系数为 1, 如第 1 个系数为 1, 即 $b_1 = 1$ 或 $c_1 = 1$, 这相当于取 $\alpha = b_1$ 或 $\alpha = c_1$;

(A2) 假设线性动态子系统的增益 $G(z)$ 为 1, 即

$$G(1) := \frac{B(1)}{A(1)} = \frac{b_1 + b_2 + \dots + b_{n_b}}{1 + a_1 + a_2 + \dots + a_{n_a}} = 1;$$

(A3) 假设多项式 $B(z)$ 系数的模为 1, 即 $b_1^2 + b_2^2 + \dots + b_{n_b}^2 = 1$, 且 $b_1 > 0$, 或非线性函数 $f(\cdot)$ 系数的模为 1, 即 $c_1^2 + c_2^2 + \dots + c_m^2 = 1$, 且 $c_1 > 0$.

对于基于过参数化模型的辨识方法,这里采用第(A1)个假设 $b_1 = 1$. 读者可以采用假设 $c_1 = 1$, 推导相应的辨识方法. 定义参数向量 $\boldsymbol{\theta}$ 和信息向量 $\boldsymbol{\phi}(t)$ 如下:

$$\boldsymbol{\theta} := [\boldsymbol{\theta}_0^T, \boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T, \dots, \boldsymbol{\theta}_{n_b}^T]^T = \begin{bmatrix} \boldsymbol{\theta}_0 \\ \mathbf{b} \otimes \mathbf{c} \end{bmatrix} \in \mathbf{R}^n,$$

$$n := n_a + mn_b, \quad (5)$$

$$\boldsymbol{\phi}(t) := [\boldsymbol{\phi}_0^T(t), \boldsymbol{\phi}_1^T(t), \boldsymbol{\phi}_2^T(t), \dots, \boldsymbol{\phi}_{n_b}^T(t)]^T \in \mathbf{R}^n, \quad (6)$$

其中

$$\boldsymbol{\theta}_0 := \mathbf{a} = [a_1, a_2, \dots, a_{n_a}]^T \in \mathbf{R}^{n_a}, \quad (7)$$

$$\mathbf{b} := [b_1, b_2, \dots, b_{n_b}]^T \in \mathbf{R}^{n_b}, \quad (8)$$

$$\boldsymbol{\theta}_i := b_i \mathbf{c} = [b_i c_1, b_i c_2, \dots, b_i c_m]^T \in \mathbf{R}^m, \quad i = 1, 2, \dots, n_b, \quad (9)$$

$$\boldsymbol{\phi}_0(t) := [-x(t-1), -x(t-2), \dots, -x(t-n_a)]^T \in \mathbf{R}^{n_a}, \quad (10)$$

$$\boldsymbol{\phi}_i(t) := \mathbf{f}(t-i) = [f_1(u(t-i)), f_2(u(t-i)), \dots, f_m(u(t-i))]^T \in \mathbf{R}^m. \quad (11)$$

根据图 1,系统的无噪输出可表示为

$$x(t) := \frac{B(z)}{A(z)} \bar{u}(t). \quad (12)$$

利用移位算子 z^{-1} 的性质,使用式(2),可得

$$x(t) = [1 - A(z)]x(t) + B(z)\bar{u}(t) = \boldsymbol{\phi}^T(t)\boldsymbol{\theta}, \quad (13)$$

由式(1),(12)和(13)可得

$$y(t) = x(t) + v(t) = \boldsymbol{\phi}^T(t)\boldsymbol{\theta} + v(t). \quad (14)$$

式(14)称为输入非线性输出误差系统的过参数化辨识模型(over-parameterization identification model).

注 1 因为参数向量 $\boldsymbol{\theta}$ 包含了系统线性块参数 b_i 与非线性块参数 c_j 的交叉乘积项,使得 $\boldsymbol{\theta}$ 的维数 $n = n_a + mn_b$ 大于系统的实际参数数目 $n_a + n_b + m$ ($m, n_b \geq 2$ 时),产生很多冗余估计.特别地,当 m 和 n_b 很大时,导致辨识算法的计算量很大.

注 2 文献[3]研究了输入非线性方程误差系统的过参数化多新息辨识方法,与其不同的是此处辨识模型的信息向量 $\boldsymbol{\phi}_0(t)$ 含有未知变量 $x(t-i)$,这是辨识的难点,本文的解决办法是借助辅助模型辨识思想,构造一个辅助模型,用辅助模型的输出代替未知变量,来研究系统参数的辨识方法.

注 3 本文第一作者等的早期工作针对输入非线性输出误差系统,提出了基于过参数化模型的辅助模型递推最小二乘辨识方法,并利用鞅收敛定理建立了算法参数估计的一致收敛性^[29].

下面针对输入非线性输出误差系统,推导出基于过参数化模型的辅助模型随机梯度辨识算法以及辅助模型递推最小二乘辨识算法,用以获取过参数化辨识模型(14)中参数向量 $\boldsymbol{\theta}$ 的估计 $\hat{\boldsymbol{\theta}}(t)$,进而从中提取出原系统参数 a_i, b_i 和 c_i 的估计.

1.2 基于过参数化模型的辅助模型随机梯度辨识方法

令 $\hat{\boldsymbol{\theta}}(t)$ 为参数向量 $\boldsymbol{\theta}$ 在时刻 t 的估计, $\|\mathbf{X}\|^2 := \text{tr}[\mathbf{X}\mathbf{X}^T]$ 表示矩阵(或向量) \mathbf{X} 的范数, $\mathbf{1}_n$ 是元均为 1 的 n 维列向量, p_0 是一个很大的正数,如 $p_0 = 10^6$.

针对 IN-OE 系统的过参数化辨识模型(14),定义梯度准则函数(gradient criterion function)

$$J_1(\boldsymbol{\theta}) := \frac{1}{2} [y(t) - \boldsymbol{\phi}^T(t)\boldsymbol{\theta}]^2,$$

设 $1/r(t)$ 为步长.参考文献[1-3],利用负梯度搜索优化准则函数 $J_1(\boldsymbol{\theta})$,可得下列递推关系:

$$\begin{aligned} \hat{\boldsymbol{\theta}}(t) &= \hat{\boldsymbol{\theta}}(t-1) - \frac{1}{r(t)} \text{grad}[J_1(\hat{\boldsymbol{\theta}}(t-1))] = \\ & \hat{\boldsymbol{\theta}}(t-1) + \frac{\boldsymbol{\phi}(t)}{r(t)} [y(t) - \boldsymbol{\phi}^T(t)\hat{\boldsymbol{\theta}}(t-1)], \end{aligned} \quad (15)$$

$$r(t) = r(t-1) + \|\boldsymbol{\phi}(t)\|^2, \quad r(0) = 1. \quad (16)$$

注意到信息向量 $\boldsymbol{\phi}(t)$ 中含有未知变量 $x(t-i)$,因此算法(15)一(16)不可实现.解决的方案是借助辅助模型辨识思想,利用 $u(t)$ 作为输入,建立一个辅助模型,如图 2 所示,其中 $\frac{B_a(z)}{A_a(z)}$ 是与 $\frac{B(z)}{A(z)}$ 结构相同的有理传递函数, $f_a(u(t))$ 与非线性 $f(u(t))$ 解析式相同,辅助模型可以表示为

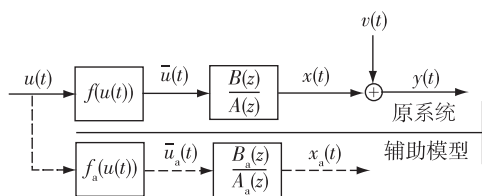


图 2 带辅助模型的 IN-OE 系统

Fig. 2 The IN-OE system with an auxiliary model

$$x_a(t) = \frac{B_a(z)}{A_a(z)} \bar{u}_a(t), \quad \bar{u}_a(t) = f_a(u(t)).$$

参考式(12)写成式(13)的形式,上式的 $x_a(t)$ 可以表示为

$$x_a(t) := \boldsymbol{\phi}_a^T(t)\boldsymbol{\theta}_a(t), \quad (17)$$

其中 $\boldsymbol{\phi}_a(t)$ 和 $\boldsymbol{\theta}_a(t)$ 分别为辅助模型的信息向量和参数向量.

这里将系统参数估计向量 $\hat{\boldsymbol{\theta}}(t)$ 作为辅助模型的信息向量 $\boldsymbol{\theta}_a(t)$,即 $\boldsymbol{\theta}_a(t) := \hat{\boldsymbol{\theta}}(t)$,用辅助模型的输出 $x_a(t-i)$ 定义 $\hat{\boldsymbol{\phi}}_0(t)$ 的估计:

$$\hat{\boldsymbol{\phi}}_0(t) := [-x_a(t-1), -x_a(t-2), \dots, -x_a(t-n_a)]^T \in \mathbf{R}^{n_a}, \quad (18)$$

用 $\hat{\boldsymbol{\phi}}_0(t)$ 和 $\boldsymbol{\phi}_i(t)$ 定义 $\boldsymbol{\phi}(t)$ 的估计:

$$\hat{\boldsymbol{\phi}}(t) := [\hat{\boldsymbol{\phi}}_0^T(t), \boldsymbol{\phi}_1^T(t), \boldsymbol{\phi}_2^T(t), \dots, \boldsymbol{\phi}_{n_b}^T(t)]^T \in \mathbf{R}^n, \quad (19)$$

$$\boldsymbol{\phi}_i(t) := [f_1(u(t-i)), f_2(u(t-i)), \dots, f_m(u(t-i))]^T \in \mathbf{R}^m, \quad i = 1, 2, \dots, n_b. \quad (20)$$

用 $\hat{\boldsymbol{\phi}}(t)$ 作为辅助模型的信息向量 $\boldsymbol{\phi}_a(t)$,即 $\boldsymbol{\phi}_a(t) := \hat{\boldsymbol{\phi}}(t)$,于是辅助模型(17)可以表示为

$$x_a(t) := \hat{\boldsymbol{\phi}}^T(t)\hat{\boldsymbol{\theta}}(t). \quad (21)$$

因此,用 $\hat{\boldsymbol{\phi}}(t)$ 代替式(15)一(16)中的 $\boldsymbol{\phi}(t)$,联立式(18)一(21)可得辨识参数向量 $\boldsymbol{\theta}$ 的过参数化辅助模型随机梯度算法(O-AM-SG算法):

$$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t-1) + \frac{\hat{\boldsymbol{\phi}}(t)}{r(t)} [y(t) - \hat{\boldsymbol{\phi}}^T(t)\hat{\boldsymbol{\theta}}(t-1)], \quad (22)$$

$$r(t) = r(t-1) + \|\hat{\boldsymbol{\phi}}(t)\|^2, \quad (23)$$

$$\hat{\boldsymbol{\phi}}(t) = [\hat{\boldsymbol{\phi}}_0^T(t), \boldsymbol{\phi}_1^T(t), \boldsymbol{\phi}_2^T(t), \dots, \boldsymbol{\phi}_{n_b}^T(t)]^T, \quad (24)$$

$$\hat{\boldsymbol{\phi}}_0(t) = [-x_a(t-1), -x_a(t-2), \dots, -x_a(t-n_a)]^T, \quad (25)$$

$$\boldsymbol{\phi}_i(t) = [f_1(u(t-i)), f_2(u(t-i)), \dots, f_m(u(t-i))]^T, \quad i=1, 2, \dots, n_b, \quad (26)$$

$$x_a(t) = \hat{\boldsymbol{\phi}}^T(t) \hat{\boldsymbol{\theta}}(t). \quad (27)$$

注 4 为提高随机梯度辨识算法的收敛速度^[1-3],在 O-AM-SG 算法中引入遗忘因子 λ ,将式 (23) 修改为

$$r(t) = \lambda r(t-1) + \|\hat{\boldsymbol{\phi}}(t)\|^2, \quad 0 \leq \lambda \leq 1, \quad r(0) = 1, \quad (28)$$

就得到遗忘因子 O-AM-SG 算法.

值得指出的是,所有的随机梯度类算法,包括辅助模型递阶随机梯度算法、辅助模型两阶段随机梯度算法和辅助模型三阶段随机梯度算法等,都可以引入遗忘因子来改进算法收敛速度,提高参数估计精度.

O-AM-SG 算法可以计算出时刻 t 的参数估计向量 $\hat{\boldsymbol{\theta}}(t)$. 根据参数向量 $\boldsymbol{\theta}$ 和 $\boldsymbol{\theta}_i$ 的定义式,可知 $\hat{\boldsymbol{\theta}}(t)$ 与系统的参数估计 $\hat{a}_i(t), \hat{b}_i(t), \hat{c}_i(t)$ 的关系如下:

$$\hat{\boldsymbol{\theta}}(t) := [\hat{\boldsymbol{\theta}}_0^T(t), \hat{\boldsymbol{\theta}}_1^T(t), \hat{\boldsymbol{\theta}}_2^T(t), \dots, \hat{\boldsymbol{\theta}}_{n_b}^T(t)]^T = \begin{bmatrix} \hat{\boldsymbol{\theta}}_0(t) \\ \widehat{\mathbf{b} \otimes \mathbf{c}}(t) \end{bmatrix} \in \mathbf{R}^n, \quad \hat{\boldsymbol{\theta}}_0(t) := \hat{\mathbf{a}}(t) = [\hat{a}_1(t), \hat{a}_2(t), \dots, \hat{a}_{n_a}(t)]^T \in \mathbf{R}^{n_a}, \quad \hat{\mathbf{b}}(t) := [\hat{b}_1(t), \hat{b}_2(t), \dots, \hat{b}_{n_b}(t)]^T \in \mathbf{R}^{n_b}, \quad \hat{\mathbf{c}}(t) := [\hat{c}_1(t), \hat{c}_2(t), \dots, \hat{c}_m(t)]^T \in \mathbf{R}^m, \quad \hat{\boldsymbol{\theta}}_i(t) := \widehat{\mathbf{b}_i \mathbf{c}}(t) = [b_i \widehat{c}_1(t), b_i \widehat{c}_2(t), \dots, b_i \widehat{c}_m(t)]^T \in \mathbf{R}^m, \quad i=1, 2, \dots, n_b.$$

因此可以根据 $\hat{\boldsymbol{\theta}}(t)$ 确定系统的参数估计 $\hat{a}_i(t), \hat{b}_i(t), \hat{c}_i(t)$. 下面介绍确定系统参数估计的平均值方法.

令 $\hat{\theta}_i(t)$ 代表向量 $\hat{\boldsymbol{\theta}}(t)$ 的第 i 个元,由于假设 $b_1=1$,参数 a_i 和 c_i 的估计可从 $\hat{\boldsymbol{\theta}}(t)$ 中读取出来,即

$$\hat{a}_i(t) = \hat{\theta}_i(t), \quad i=1, 2, \dots, n_a, \quad (29)$$

$$\hat{c}_j(t) = \hat{\theta}_{n_a+j}(t), \quad j=1, 2, 3, \dots, m, \quad (30)$$

$$\hat{b}_i(t) = \frac{\hat{\theta}_{im-m+n_a+j}(t)}{\hat{c}_j(t)}, \quad i=2, 3, \dots, n_b.$$

由于 $j=1, 2, \dots, m$,对任意 $i(1 < i \leq n_b)$, b_i 有 m 个估计值,即冗余估计.这里取这 m 个估计值的平均值作为 b_i 的估计^[3,17,18,29-30],即

$$\hat{b}_i(t) = \frac{1}{m} \sum_{j=1}^m \frac{\hat{\theta}_{im-m+n_a+j}(t)}{\hat{c}_j(t)}, \quad i=2, 3, \dots, n_b. \quad (31)$$

式(22)—(31)构成了基于平均值的 O-AM-SG 算法.

基于过参数化模型的辨识算法包含两步.第 1 步

辨识出过参数向量 $\boldsymbol{\theta}$;第 2 步是从获得参数估计 $\hat{\boldsymbol{\theta}}(t)$ 中提取系统的参数.提取参数可采用上面的平均值法,以及排列组合法、矩阵奇异值分解等方法^[30].

利用基于过参数化的辅助模型随机梯度算法计算系统参数估计 $\hat{a}_i(t), \hat{b}_i(t)$ 和 $\hat{c}_i(t)$ 的步骤如下:

1) 初始化:令 $t=1$,给定数据长度 L_e 和基函数 $f_j(\ast)$;置初值 $\hat{\boldsymbol{\theta}}(0) = \mathbf{1}_n/p_0, r(0) = 1, x_a(t-i) = 1/p_0, i=0, 1, 2, \dots, n_a, p_0 = 10^6$.

2) 采集数据 $u(t)$ 和 $y(t)$,用式(25)构造 $\hat{\boldsymbol{\phi}}_0(t)$;根据式(26)构造 $\boldsymbol{\phi}_1(t), \boldsymbol{\phi}_2(t), \dots, \boldsymbol{\phi}_{n_b}(t)$;用式(24)构造 $\hat{\boldsymbol{\phi}}(t)$.

3) 用式(23)计算 $r(t)$.

4) 用式(22)刷新参数估计向量 $\hat{\boldsymbol{\theta}}(t)$.根据式(29)—(31),从 $\hat{\boldsymbol{\theta}}(t)$ 中提取或计算参数估计 $\hat{a}_i(t), \hat{b}_i(t)$ 和 $\hat{c}_i(t)$.

5) 用式(27)计算辅助模型输出 $x_a(t)$.

6) 如果 $t < L_e, t$ 增加 1,转到第 2)步;否则,终止计算,获得参数估计向量 $\hat{\boldsymbol{\theta}}(L_e)$.

7) 获取参数估计 $\hat{a}_i(L_e), \hat{b}_i(L_e)$ 和 $\hat{c}_i(L_e)$.

如果辅助模型的输出 $x_a(t)$ 逼近系统的真实输出 $x(t)$,就可以把它作为 $x(t)$ 的估计 $\hat{x}(t)$,即 $\hat{x}(t) = x_a(t)$.在这种情况下,式(25)和(27)可修改为

$$\hat{\boldsymbol{\phi}}_0(t) = [-\hat{x}(t-1), -\hat{x}(t-2), \dots, -\hat{x}(t-n_a)]^T, \quad (32)$$

$$\hat{x}(t) = \hat{\boldsymbol{\phi}}^T(t) \hat{\boldsymbol{\theta}}(t). \quad (33)$$

1.3 基于过参数化模型的辅助模型最小二乘辨识方法 考虑式(14)的过参数化辨识模型,即

$$y(t) = \boldsymbol{\phi}^T(t) \boldsymbol{\theta} + v(t), \quad (34)$$

其中各变量定义同上.

定义输出向量 \mathbf{Y}_t ,信息矩阵 \mathbf{H}_t 如下:

$$\mathbf{Y}_t := \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(t) \end{bmatrix} \in \mathbf{R}^t, \quad \mathbf{H}_t := \begin{bmatrix} \boldsymbol{\phi}^T(1) \\ \boldsymbol{\phi}^T(2) \\ \vdots \\ \boldsymbol{\phi}^T(t) \end{bmatrix} \in \mathbf{R}^{t \times n}.$$

定义二次准则函数(quadratic criterion function)

$$J_2(\boldsymbol{\theta}) := \frac{1}{2} (\mathbf{Y}_t - \mathbf{H}_t \boldsymbol{\theta})^T (\mathbf{Y}_t - \mathbf{H}_t \boldsymbol{\theta}).$$

极小化这个二次准则函数,令 $J_2(\boldsymbol{\theta})$ 对 $\boldsymbol{\theta}$ 的偏导数为零,用式(37)中 $\hat{\mathbf{H}}_t$ 代替未知 \mathbf{H}_t ,联立式(18)—(21),得到过参数化辅助模型最小二乘算法(O-AM-LS算法):

$$\hat{\boldsymbol{\theta}}(t) = (\hat{\mathbf{H}}_t^T \hat{\mathbf{H}}_t)^{-1} \hat{\mathbf{H}}_t^T \mathbf{Y}_t, \quad (35)$$

$$Y_i = [y(1), y(2), \dots, y(t)]^T, \quad (36)$$

$$\hat{H}_i = [\hat{\phi}(1), \hat{\phi}(2), \dots, \hat{\phi}(t)]^T, \quad (37)$$

$$\hat{\phi}(t) = [\hat{\phi}_0^T(t), \phi_1^T(t), \phi_2^T(t), \dots, \phi_{n_b}^T(t)]^T, \quad (38)$$

$$\hat{\phi}_0(t) = [-x_a(t-1), -x_a(t-2), \dots, -x_a(t-n_a)]^T, \quad (39)$$

$$\phi_i(t) = [f_1(u(t-i)), f_2(u(t-i)), \dots, f_m(u(t-i))]^T, \quad (40)$$

$$i = 1, 2, \dots, n_b,$$

$$x_a(t) = \hat{\phi}^T(t)\hat{\theta}(t), \quad x_a(-i) = 1/p_0, \quad i = 0, 1, \dots, n_a. \quad (41)$$

定义协方差阵 $P(t)$ 和向量 $\xi(t)$ 如下:

$$P^{-1}(t) := \hat{H}_i^T \hat{H}_i = P^{-1}(t-1) + \hat{\phi}(t)\hat{\phi}^T(t),$$

$$P(0) = p_0 I_n, \quad (42)$$

$$\xi(t) := \hat{H}_i^T Y_i = \xi(t-1) + \hat{\phi}(t)y(t), \quad \xi(0) = 0. \quad (43)$$

那么辨识参数向量 θ 的过参数化辅助模型最小二乘算法(O-AM-LS 算法)^[1-3]可以等价写为

$$\hat{\theta}(t) = P(t)\xi(t), \quad (44)$$

$$\xi(t) = \xi(t-1) + \hat{\phi}(t)y(t), \quad \xi(0) = 0, \quad (45)$$

$$P(t) = P(t-1) - \frac{P(t-1)\hat{\phi}(t)\hat{\phi}^T(t)P(t-1)}{1 + \hat{\phi}^T(t)P(t-1)\hat{\phi}(t)},$$

$$P(0) = p_0 I_n, \quad (46)$$

$$\hat{\phi}(t) = [\hat{\phi}_0^T(t), \phi_1^T(t), \phi_2^T(t), \dots, \phi_{n_b}^T(t)]^T, \quad (47)$$

$$\hat{\phi}_0(t) = [-x_a(t-1), -x_a(t-2), \dots, -x_a(t-n_a)]^T, \quad (48)$$

$$\phi_i(t) = [f_1(u(t-i)), f_2(u(t-i)), \dots, f_m(u(t-i))]^T, \quad (49)$$

$$i = 1, 2, \dots, n_b,$$

$$x_a(t) = \hat{\phi}^T(t)\hat{\theta}(t), \quad x_a(-i) = 1/p_0, \quad i = 0, 1, \dots, n_a. \quad (50)$$

1.4 基于过参数化模型的辅助模型递推最小二乘辨识方法

根据式(44)–(45),使用式(42),能够得到辨识参数向量 θ 的过参数化辅助模型递推最小二乘算法(O-AM-RLS 算法):

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t)[y(t) - \hat{\phi}^T(t)\hat{\theta}(t-1)], \quad (51)$$

$$\hat{\theta}(0) = \mathbf{1}_n/p_0,$$

$$L(t) = P(t-1)\hat{\phi}(t)[1 + \hat{\phi}^T(t)P(t-1)\hat{\phi}(t)]^{-1}, \quad (52)$$

$$P(t) = [I_n - L(t)\hat{\phi}^T(t)]P(t-1), \quad P(0) = p_0 I_n, \quad (53)$$

$$\hat{\phi}(t) = [\hat{\phi}_0^T(t), \phi_1^T(t), \phi_2^T(t), \dots, \phi_{n_b}^T(t)]^T, \quad (54)$$

$$\hat{\phi}_0(t) = [-x_a(t-1), -x_a(t-2), \dots, -x_a(t-n_a)]^T, \quad (55)$$

$$\phi_i(t) = [f_1(u(t-i)), f_2(u(t-i)), \dots, f_m(u(t-i))]^T, \quad (56)$$

$$i = 1, 2, \dots, n_b,$$

$$x_a(t) = \hat{\phi}^T(t)\hat{\theta}(t), \quad x_a(-i) = 1/p_0, \quad i = 0, 1, \dots, n_a. \quad (57)$$

注5 根据 $P^{-1}(t)$ 的定义式(42),应该取初值 $P^{-1}(0) = \mathbf{0}$,但这会导致其逆不存在^[1],故通常将其设置为一个接近于零的正定矩阵,如 $P^{-1}(0) = I_n/p_0$,即 $P(0) = p_0 I_n$ 是一个很大的正定矩阵,其中 I_n 是一个 n 阶单位阵,取 $p_0 = 10^6$.

注6 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 数来表示计算量的大小^[8-9].表 1 列出了 O-AM-RLS 算法(51)–(57)的计算量.

从表 1 可以看出,最小二乘算法每步的计算量是 $N_1 := 4n^2 + 8n - 1$.简单说最小二乘算法的计算量是 $N_1 = O[(n_a + mn_b)^2]$.

2 基于过参数化模型的辅助模型递阶辨识方法

输入非线性输出误差(IN-OE)系统的过参数化辨识模型参数多、维数高,导致 O-AM-RLS 算法的计算量大.为此,利用递阶辨识原理研究计算量小的递阶辨识方法是十分必要的^[13].本节讨论 IN-OE 系统基于过参数化的辅助模型递阶辨识方法,其思想是借助递阶辨识原理将系统过参数化辨识模型分解成若干个维数低、参数少的子辨识模型(子系统),然

表 1 过参数化辅助模型递推最小二乘算法的计算量

Table 1 The computational efficiency of the O-AM-RLS algorithm

变量	表达式	乘法次数	加法次数
$\hat{\theta}(t)$	$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t)e(t) \in \mathbf{R}^{n_a + mn_b}$	$n_a + mn_b$	$n_a + mn_b$
	$e(t) := y(t) - \hat{\phi}^T(t)\hat{\theta}(t-1) \in \mathbf{R}$	$n_a + mn_b$	$n_a + mn_b$
$L(t)$	$L(t) = \zeta(t) / [1 + \hat{\phi}^T(t)\zeta(t)] \in \mathbf{R}^{n_a + mn_b}$	$2(n_a + mn_b)$	$n_a + mn_b$
	$\zeta(t) := P(t-1)\hat{\phi}(t) \in \mathbf{R}^{n_a + mn_b}$	$(n_a + mn_b)^2$	$(n_a + mn_b - 1)(n_a + mn_b)$
$P(t)$	$P(t) = P(t-1) - L(t)\zeta^T(t) \in \mathbf{R}^{(n_a + mn_b) \times (n_a + mn_b)}$	$(n_a + mn_b)^2$	$(n_a + mn_b)^2$
$x_a(t)$	$x_a(t) = \hat{\phi}^T(t)\hat{\theta}(t) \in \mathbf{R}$	$n_a + mn_b$	$n_a + mn_b - 1$
总数		$2(n_a + mn_b)^2 + 5(n_a + mn_b)$	$2(n_a + mn_b)^2 + 3(n_a + mn_b) - 1$
总 flop 数		$N_1 := 4(n_a + mn_b)^2 + 8(n_a + mn_b) - 1$	

后推导每个子系统的参数辨识算法,并协调辨识算法间的关联项.

2.1 基于过参数化的递阶辨识模型

考虑输入非线性输出误差系统(1)–(6)的过参数化辨识模型(14),根据式(7)–(11)参数向量和信息向量的定义,模型(14)可改写为

$$y(t) = \boldsymbol{\phi}^T(t)\boldsymbol{\theta} + v(t) = \boldsymbol{\phi}_0^T(t)\boldsymbol{\theta}_0 + \boldsymbol{\phi}_1^T(t)\boldsymbol{\theta}_1 + \cdots + \boldsymbol{\phi}_{n_b}^T(t)\boldsymbol{\theta}_{n_b} + v(t). \quad (58)$$

其中

$$\boldsymbol{\theta}_0 := [a_1, a_2, \dots, a_{n_a}]^T \in \mathbf{R}^{n_a}, \quad (59)$$

$$\boldsymbol{\theta}_i := [b_i c_1, b_i c_2, \dots, b_i c_m]^T = b_i \mathbf{c} \in \mathbf{R}^m, \quad (60)$$

$$i = 1, 2, \dots, n_b,$$

$$\boldsymbol{\phi}_0(t) := [-x(t-1), -x(t-2), \dots, -x(t-n_a)]^T \in \mathbf{R}^{n_a}, \quad (61)$$

$$\boldsymbol{\phi}_i(t) := [f_1(u(t-i)), f_2(u(t-i)), \dots, f_m(u(t-i))]^T \in \mathbf{R}^m. \quad (62)$$

定义中间变量

$$y_i(t) := y(t) - \sum_{j=0, j \neq i}^{n_b} \boldsymbol{\phi}_j^T(t)\boldsymbol{\theta}_j, \quad i = 0, 1, \dots, n_b. \quad (63)$$

系统(58)可以分解为 $n_b + 1$ 个虚拟子系统:

$$y_i(t) = \boldsymbol{\varphi}_i^T(t)\boldsymbol{\theta}_i + v(t), \quad i = 0, 1, \dots, n_b. \quad (64)$$

下面用梯度搜索和最小二乘原理来研究估计这 $n_b + 1$ 个子系统参数向量 $\boldsymbol{\theta}_i$ 的辨识方法.

2.2 基于过参数化模型的辅助模型递阶随机梯度辨识方法

针对第 i 个虚拟子系统,定义准则函数

$$J_{3i}(\boldsymbol{\theta}_i) := \frac{1}{2} [y_i(t) - \boldsymbol{\phi}_i^T(t)\boldsymbol{\theta}_i]^2.$$

令 $\hat{\boldsymbol{\theta}}_i(t)$ 为第 i 个虚拟子系统参数向量 $\boldsymbol{\theta}_i$ 在时刻 t 的估计.利用负梯度搜索,极小化准则函数 $J_{3i}(\boldsymbol{\theta}_i)$,根据递阶辨识原理,未知变量用辅助模型的输出代替,可以得到辨识参数向量 $\boldsymbol{\theta}_i$ 的过参数化辅助模型递阶随机梯度算法(O-AM-HSG算法):

$$\hat{\boldsymbol{\theta}}_0(t) = \hat{\boldsymbol{\theta}}_0(t-1) + \frac{\hat{\boldsymbol{\phi}}_0(t)}{r_0(t)} e(t), \quad (65)$$

$$e(t) = y(t) - \hat{\boldsymbol{\phi}}_0^T(t)\hat{\boldsymbol{\theta}}_0(t-1) - \sum_{j=1}^{n_b} \boldsymbol{\phi}_j^T(t)\hat{\boldsymbol{\theta}}_j(t-1), \quad (66)$$

$$r_0(t) = r_0(t-1) + \|\hat{\boldsymbol{\phi}}_0(t)\|^2, \quad r_0(0) = 1, \quad (67)$$

$$\hat{\boldsymbol{\theta}}_i(t) = \hat{\boldsymbol{\theta}}_i(t-1) + \frac{\boldsymbol{\phi}_i(t)}{r_i(t)} e(t), \quad i = 1, 2, \dots, n_b, \quad (68)$$

$$r_i(t) = r_i(t-1) + \|\boldsymbol{\phi}_i(t)\|^2, \quad r_i(0) = 1, \quad (69)$$

$$\hat{\boldsymbol{\phi}}_0(t) = [-x_a(t-1), -x_a(t-2), \dots, -x_a(t-n_a)]^T, \quad (70)$$

$$\boldsymbol{\phi}_i(t) = [f_1(u(t-i)), f_2(u(t-i)), \dots, f_m(u(t-i))]^T, \quad (71)$$

$$x_a(t) = \hat{\boldsymbol{\phi}}_0^T(t)\hat{\boldsymbol{\theta}}_0(t-1) + \sum_{j=1}^{n_b} \boldsymbol{\phi}_j^T(t)\hat{\boldsymbol{\theta}}_j(t-1). \quad (72)$$

注7 O-AM-HSG算法比O-AM-SG算法的收敛速度略有提高.为进一步提高收敛速度,可引入遗忘因子 λ ,将式(67)和(69)修改为

$$r_0(t) = \lambda r_0(t-1) + \|\hat{\boldsymbol{\phi}}_0(t)\|^2, \quad 0 \leq \lambda \leq 1, \quad r_0(0) = 1, \quad (73)$$

$$r_i(t) = \lambda r_i(t-1) + \|\boldsymbol{\phi}_i(t)\|^2, \quad r_i(0) = 1, \quad (74)$$

就得到遗忘因子O-AM-HSG算法(65)–(66), (68), (70)–(72).

对于O-AM-HSG算法(65)–(72),如果将式(65)中的 $r_0(t)$ 和式(68)中的 $r_i(t)$ 都修改为同一 $r(t)$,并取

$$r(t) = r(t-1) + \|\hat{\boldsymbol{\phi}}_0(t)\|^2 + \sum_{i=1}^{n_b} \|\boldsymbol{\phi}_i(t)\|^2, \quad (75)$$

$$r(0) = 1,$$

则O-AM-HSG算法就等同于O-AM-SG算法(22)–(27).

基于递阶辨识原理的递阶随机梯度类辨识算法,包括基于辅助模型的两阶段随机梯度算法、基于辅助模型的三阶段随机梯度算法、基于辅助模型的递阶随机梯度算法等与基于辅助模型的随机梯度算法间的细微区别就在于 $r(t)$ 的选取.当取相同的 $r(t)$ 时,这些算法就等价.

利用O-AM-HSG算法(65)–(72)计算出所有子参数向量 $\boldsymbol{\theta}_i$ 的估计 $\hat{\boldsymbol{\theta}}_i(t)$ ($i = 0, 1, 2, \dots, n_b$) 后,就可从中提取计算出系统的参数估计 $\hat{a}_i(t)$, $\hat{b}_i(t)$ 和 $\hat{c}_i(t)$,细节如下.

令 $\hat{\boldsymbol{\theta}}_i(j, t)$ 表示参数估计向量 $\hat{\boldsymbol{\theta}}_i(t)$ 的第 j 个元.根据 $b_1 = 1$ 的假设,以及 $\boldsymbol{\theta}_i$ 的定义式(59)–(60),可知

$$\hat{a}_j(t) = \hat{\boldsymbol{\theta}}_0(j, t), \quad j = 1, 2, \dots, n_a, \quad (76)$$

$$\hat{c}_j(t) = \hat{\boldsymbol{\theta}}_1(j, t), \quad j = 1, 2, \dots, m. \quad (77)$$

在获得估计 $\hat{c}_j(t)$ 后,可用 $\frac{\hat{\boldsymbol{\theta}}_i(j, t)}{\hat{c}_j(t)}$ 计算 b_i 的估计 ($i = 2, 3, \dots, n_b$).注意到 $j = 1, 2, \dots, m$, 故 b_i 有 m 个不同的估计 $\hat{b}_i(t)$, 这里仍取它们的平均值作为 b_i 的估计,即

$$\hat{b}_i(t) = \frac{1}{m} \sum_{j=1}^m \frac{\hat{\boldsymbol{\theta}}_i(j, t)}{\hat{c}_j(t)}, \quad i = 2, 3, \dots, n_b. \quad (78)$$

式(65)–(78)构成了基于平均值的O-AM-HSG算法,其实现步骤如下:

1) 初始化:令 $t = 1$, 给定数据窗长度 L_e 和基函

数 $f_j(\ast)$, 置初值 $\hat{\theta}_0(0) = \mathbf{1}_{n_a}/p_0$, $\hat{\theta}_i(0) = \mathbf{1}_m/p_0$, $x_a(t-j) = 1/p_0, j=0, 1, 2, \dots, n_a, p_0 = 10^6$.

2) 采集输入输出数据 $u(t)$ 和 $y(t)$, 用式(70)构造 $\hat{\phi}_0(t)$, 按照式(71)依次构造 $\phi_1(t), \phi_2(t), \dots, \phi_{n_b}(t)$.

3) 用式(66)计算新息 $e(t)$, 按照式(67)和(69)依次计算 $r_0(t), r_1(t), r_2(t), \dots, r_{n_b}(t)$.

4) 用式(65)和(68)依次刷新参数估计向量 $\hat{\theta}_0(t), \hat{\theta}_1(t), \hat{\theta}_2(t), \dots, \hat{\theta}_{n_b}(t)$.

5) 利用式(76)——(78)计算参数估计 $\hat{a}_i(t), \hat{b}_i(t)$ 和 $\hat{c}_i(t)$.

6) 用式(72)计算辅助模型输出 $x_a(t)$.

7) 如果 $t < L_e, t$ 增加 1, 转到第 2) 步; 否则, 终止计算, 获得参数估计向量 $\hat{\theta}(L_e)$.

8) 获取参数估计 $\hat{a}_i(L_e), \hat{b}_i(L_e)$ 和 $\hat{c}_i(L_e)$.

2.3 基于过参数化模型的辅助模型递阶最小二乘辨识方法

根据递阶辨识原理, 分解 IN-OE 系统过参数模型(58)可以得到 n_b+1 个虚拟子系统

$$y_i(t) = \phi_i^T(t)\theta_i + v(t), \quad i=0, 1, \dots, n_b, \quad (79)$$

其中

$$\begin{aligned} y_i(t) &:= y(t) - \sum_{j=0, j \neq i}^{n_b} \phi_j^T(t)\theta_j, \\ \phi_0(t) &:= \begin{bmatrix} -x(t-1) \\ -x(t-2) \\ \vdots \\ -x(t-n_a) \end{bmatrix} \in \mathbf{R}^{n_a}, \\ \phi_i(t) &:= \begin{bmatrix} f_1(u(t-i)) \\ f_2(u(t-i)) \\ \vdots \\ f_m(u(t-i)) \end{bmatrix} \in \mathbf{R}^m, \quad i=1, 2, \dots, n_b, \end{aligned}$$

根据辨识模型(79)的第 i 个子系统, 定义准则函数 (criterion function)

$$J_{4i}(\theta_i) := \frac{1}{2} \sum_{k=1}^t [y_i(k) - \phi_i^T(k)\theta_i]^2.$$

令 $\hat{\theta}_i(t)$ 表示第 i 个子系统参数向量 θ_i 在时刻 t 的估计. 极小化 $J_{4i}(\theta_i)$ 或令 $J_{4i}(\theta_i)$ 对 θ_i 的偏导数为零, 算法涉及的未知量用其估计代替, 联立辅助模型(70)——(72), 可以得到辨识参数向量 θ_i 的过参数化辅助模型递阶最小二乘算法 (O-AM-HLS 算法):

$$\hat{\theta}_0(t) = \hat{\theta}_0(t-1) + L_0(t)e(t), \quad (80)$$

$$e(t) = y(t) - \hat{\phi}_0^T(t)\hat{\theta}_0(t-1) - \sum_{j=1}^{n_b} \phi_j^T(t)\hat{\theta}_j(t-1), \quad (81)$$

$$L_0(t) = P_0(t-1)\hat{\phi}_0(t) [1 + \hat{\phi}_0^T(t)P_0(t-1)\hat{\phi}_0(t)]^{-1}, \quad (82)$$

$$P_0(t) = P_0(t-1) - L_0(t)\hat{\phi}_0^T(t)P_0(t-1), \quad (83)$$

$$\hat{\theta}_i(t) = \hat{\theta}_i(t-1) + L_i(t)e(t), \quad i=1, 2, \dots, n_b, \quad (84)$$

$$L_i(t) = P_i(t-1)\phi_i(t) [1 + \phi_i^T(t)P_i(t-1)\phi_i(t)]^{-1}, \quad (85)$$

$$P_i(t) = P_i(t-1) - L_i(t)\phi_i^T(t)P_i(t-1), \quad (86)$$

$$\hat{\phi}_0(t) = [-x_a(t-1), -x_a(t-2), \dots, -x_a(t-n_a)]^T, \quad (87)$$

$$\phi_i(t) = [f_1(u(t-i)), f_2(u(t-i)), \dots, f_m(u(t-i))]^T, \quad (88)$$

$$x_a(t) = \hat{\phi}_0^T(t)\hat{\theta}_0(t) + \sum_{j=1}^{n_b} \phi_j^T(t)\hat{\theta}_j(t). \quad (89)$$

注 8 O-AM-HLS 算法中, 参数估计向量 $\hat{\theta}_0(t)$ 和信息向量 $\hat{\phi}_0(t)$ 是 n_a 维, 其他 n_b 个参数估计向量 $\hat{\theta}_i(t)$ 和与其对应的信息向量 $\phi_i(t)$ 都是 m 维的. 因此, O-AM-HLS 算法的初始值可设置如下:

$$\hat{\theta}_0(0) = \mathbf{1}_{n_a}/p_0, \quad P_0(0) = p_0 I_{n_a}, \quad p_0 = 10^6, \quad (90)$$

$$\hat{\theta}_i(0) = \mathbf{1}_m/p_0, \quad P_i(0) = p_0 I_m, \quad i=1, 2, \dots, n_b, \quad (91)$$

$$x_a(-i) = 1/p_0, \quad i=0, 1, \dots, n_a. \quad (92)$$

算法(80)——(89)计算出参数估计 $\hat{\theta}_0(t), \hat{\theta}_1(t), \dots, \hat{\theta}_{n_b}(t)$ 后, 可利用式(76)——(78)提取计算出系统参数 a_i, b_i 和 c_i 的估计. 因此, 式(80)——(89)和式(76)——(78)构成了基于平均值的 O-AM-HLS 算法, 其实现步骤如下:

1) 初始化: 令 $t=1$, 给定数据窗长度 L_e 和基函数 $f_j(\ast)$, 按照式(90)——(92)置初值.

2) 采集输入输出数据 $u(t)$ 和 $y(t)$, 用式(87)构造 $\hat{\phi}_0(t)$, 根据式(88)依次构造信息向量 $\phi_1(t), \phi_2(t), \dots, \phi_{n_b}(t)$.

3) 用式(81)计算新息 $e(t)$, 用式(82)——(83)计算 $L_0(t)$ 和 $P_0(t)$, 根据式(85)——(86)计算 $L_i(t)$ 和 $P_i(t)$.

4) 用式(80)刷新参数估计向量 $\hat{\theta}_0(t)$, 用式(84)刷新参数估计向量 $\hat{\theta}_1(t), \hat{\theta}_2(t), \dots, \hat{\theta}_{n_b}(t)$.

5) 根据式(89)计算辅助模型的输出 $x_a(t)$.

6) 利用式(76)——(78)提取参数估计 $\hat{a}_i(t), \hat{b}_i(t)$ 和 $\hat{c}_i(t)$.

7) 如果 $t < L_e, t$ 增加 1, 转到第 2) 步; 否则, 终止计算, 获得参数估计向量 $\hat{\theta}(L_e)$.

8) 输出参数估计 $\hat{a}_i(L_e), \hat{b}_i(L_e)$ 和 $\hat{c}_i(L_e)$.

O-AM-HLS 算法每递推计算一步的计算量如表 2 所示.

注 9 利用递阶辨识原理, 将 IN-OE 系统过参数化辨识模型(58)分解成 n_b+1 个子系统, 使得信息

向量和参数向量的维数从 $n := n_a + mn_b$ 降为 n_a 或 m , 因为最小二乘算法的协方差阵的计算量是维数的平方, 且 $(n_1 + n_2 + \dots + n_r)^2 > n_1^2 + n_2^2 + \dots + n_r^2$, 所以 O-AM-HLS 算法比 O-AM-RLS 算法的计算量小.

3 基于关键项分离的辅助模型辨识方法

输入非线性输出误差 (IN-OE) 系统的过参数化辨识模型 (14) 的参数向量 θ 包含了参数乘积 (b_i 与 c_j 的交叉乘积), 导致冗余参数, 其参数个数为 $n_a + mn_b$ 个, 大于系统实际参数数目 $n_a + n_b + m$ 个 ($m, n_b > 2$), 使得最小二乘辨识方法计算量增加. 而且在辨识出过参数化向量后, 还需从中计算系统的参数估计, 增加额外计算开支. 为避免出现冗余参数估计和辨识方法计算量大的问题 (指同类算法间的比较), 下面基于关键项分离原理^[31-33] 研究随机梯度辨识方法和最小二乘辨识方法.

3.1 基于关键项分离的辨识模型

考虑输入非线性输出误差 (IN-OE) 系统 (1) — (2), 重写如下:

$$y(t) = \frac{B(z)}{A(z)} \bar{u}(t) + v(t), \quad (93)$$

$$\bar{u}(t) = \sum_{j=1}^m c_j f_j(u(t)), \quad (94)$$

各变量的定义同上, 不同的是, (为了描述方便) 这里将多项式 $B(z)$ 修改为

$$B(z) := b_0 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_{n_b} z^{-n_b}.$$

考虑到系统参数的可辨识性, 归一化假设也相应调

整为 $b_0 = 1$, 随后讨论的基于关键项分离的两阶段辨识方法、三阶段辨识方法都采用这种假设.

定义中间变量 (系统无噪输出) 为

$$x(t) := \frac{B(z)}{A(z)} \bar{u}(t) \in \mathbf{R},$$

或

$$x(t) = [1 - A(z)]x(t) + [B(z) - 1]\bar{u}(t) + b_0 \bar{u}(t). \quad (95)$$

选择 $\bar{u}(t)$ 作为关键项, 利用移位算子 z^{-1} 的性质, 联合式 (94) — (95) 可得

$$\begin{aligned} x(t) = & - \sum_{i=1}^{n_a} a_i x(t-i) + \sum_{i=1}^{n_b} b_i \bar{u}(t-i) + b_0 \bar{u}(t) = \\ & - \sum_{i=1}^{n_a} a_i x(t-i) + \sum_{i=1}^{n_b} b_i \bar{u}(t-i) + \\ & \sum_{j=1}^m c_j f_j(u(t)). \end{aligned} \quad (96)$$

定义线性块的参数向量 \mathbf{a} 和 \mathbf{b} , 以及相应的信息向量 $\varphi_a(t)$ 和 $\varphi_b(t)$ 如下:

$$\begin{aligned} \mathbf{a} := \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n_a} \end{bmatrix} \in \mathbf{R}^{n_a}, \quad \varphi_a(t) := \begin{bmatrix} -x(t-1) \\ -x(t-2) \\ \vdots \\ -x(t-n_a) \end{bmatrix} \in \mathbf{R}^{n_a}, \\ \mathbf{b} := \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n_b} \end{bmatrix} \in \mathbf{R}^{n_b}, \quad \varphi_b(t) := \begin{bmatrix} \bar{u}(t-1) \\ \bar{u}(t-2) \\ \vdots \\ \bar{u}(t-n_b) \end{bmatrix} \in \mathbf{R}^{n_b}, \end{aligned}$$

定义非线性块的参数向量 \mathbf{c} , 以及相应的信息向量

表 2 基于过参数化的辅助模型递阶最小二乘算法的计算量

Table 2 The computational efficiency of the O-AM-HLS algorithm

变量	表达式	乘法次数	加法次数
$\hat{\theta}_0(t)$	$\hat{\theta}_0(t) = \hat{\theta}_0(t-1) + L_0(t)e(t) \in \mathbf{R}^{n_a + n_b}$	n_a	n_a
$e(t)$	$e(t) = y(t) - \hat{\phi}_0^T(t)\hat{\theta}_0(t-1) - \sum_{i=1}^{n_b} \phi_i^T(t)\hat{\theta}_i(t-1) \in \mathbf{R}$	$n_a + mn_b$	$n_a + mn_b$
$L_0(t)$	$L_0(t) = \zeta_0(t) / [1 + \phi_0^T(t)\zeta_0(t)] \in \mathbf{R}^{n_a}$	$2n_a$	n_a
	$\zeta_0(t) := P_0(t-1)\phi_0(t) \in \mathbf{R}^{n_a}$	n_a^2	$(n_a - 1)n_a$
$P_0(t)$	$P_0(t) = P_0(t-1) - L_0(t)\zeta_0^T(t) \in \mathbf{R}^{n_a \times n_a}$	n_a^2	n_a^2
$x_a(t)$	$x_a(t) = \hat{\phi}_0^T(t)\hat{\theta}_0(t) + \sum_{j=1}^{n_b} \phi_j^T(t)\hat{\theta}_j(t) \in \mathbf{R}$	$n_a + mn_b$	$n_a + mn_b - 1$
$\hat{\theta}_i(t)$	$\hat{\theta}_i(t) = \hat{\theta}_i(t-1) + L_i(t)e(t) \in \mathbf{R}^m$	m	m
$L_i(t)$	$L_i(t) = \zeta_i(t) / [1 + \phi_i^T(t)\zeta_i(t)] \in \mathbf{R}^m$	$2m$	m
	$\zeta_i(t) := P_i(t-1)\phi_i(t) \in \mathbf{R}^m$	m^2	$m(m-1)$
$P_i(t)$	$P_i(t) = P_i(t-1) - L_i(t)\zeta_i^T(t) \in \mathbf{R}^{m \times m}$	m^2	m^2
	总数	$2m^2 n_b + 5mn_b + 2n_a^2 + 5n_a$	$2m^2 n_b + 3mn_b + 2n_a^2 + 3n_a - 1$
	总 flop 数	$N_2 := 4m^2 n_b + 8mn_b + 4n_a^2 + 8n_a - 1$	

$f(t)$ 如下:

$$\mathbf{c} := \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix} \in \mathbf{R}^m, \quad \mathbf{f}(t) := \mathbf{f}(u(t)) = \begin{bmatrix} f_1(u(t)) \\ f_2(u(t)) \\ \vdots \\ f_m(u(t)) \end{bmatrix} \in \mathbf{R}^m.$$

进一步,定义信息向量

$$\boldsymbol{\varphi}(t) := [\boldsymbol{\varphi}_a^T(t), \boldsymbol{\varphi}_b^T(t), \mathbf{f}^T(t)]^T \in \mathbf{R}^{n_0}, \quad n_0 := n_a + n_b + m,$$

令参数向量

$$\boldsymbol{\vartheta} := [\mathbf{a}^T, \mathbf{b}^T, \mathbf{c}^T]^T \in \mathbf{R}^{n_0},$$

则式(96)和(93)可表示为

$$x(t) = \boldsymbol{\varphi}_a^T(t) \mathbf{a} + \boldsymbol{\varphi}_b^T(t) \mathbf{b} + \bar{u}(t) \quad (97)$$

$$= \boldsymbol{\varphi}^T(t) \boldsymbol{\vartheta}, \quad (98)$$

$$y(t) = \boldsymbol{\varphi}^T(t) \boldsymbol{\vartheta} + v(t). \quad (99)$$

式(99)称为 IN-OE 系统(93)—(94)的关键项分离辨识模型(key term separation based identification model).

注 10 IN-OE 系统关键项分离辨识模型(99)的输出 $y(t)$ 是系统待辨识参数的线性组合.与过参数化辨识模型(14)相比,该辨识模型参数向量 $\boldsymbol{\vartheta}$ 包含了系统参数 a_i, b_i 和 c_i , 不涉及参数乘积的冗余参数.

注 11 关键项分离原理只解决了分离问题,其分离出来的关键项通常是一个中间变量,是未知的.例如在辨识模型(99)中,信息向量 $\boldsymbol{\varphi}(t)$ 中作为关键项的 $\bar{u}(t-i)$ 是未知的,需要借助于辅助模型辨识思想实现系统未知参数和关键项的联合估计.关键项分离与辅助模型的完美结合,便产生基于关键项分离的辅助模型辨识方法.

下面讨论 IN-OE 系统的基于关键项分离的辅助模型随机梯度辨识方法和基于关键项分离的辅助模型递推最小二乘辨识方法.

3.2 基于关键项分离的辅助模型随机梯度辨识方法

$$\text{令 } \hat{\boldsymbol{\vartheta}}(t) := \begin{bmatrix} \hat{\mathbf{a}}(t) \\ \hat{\mathbf{b}}(t) \\ \hat{\mathbf{c}}(t) \end{bmatrix} \in \mathbf{R}^{n_0} \text{ 是参数向量 } \boldsymbol{\vartheta} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{c} \end{bmatrix} \text{ 在}$$

时刻 t 的估计.对于辨识模型(99),定义梯度准则函数

$$J_s(\boldsymbol{\vartheta}) := \frac{1}{2} [y(t) - \boldsymbol{\varphi}^T(t) \boldsymbol{\vartheta}]^2.$$

参考文献[1-2],利用负梯度搜索,极小化梯度准则函数 $J_s(\boldsymbol{\vartheta})$,可以得到下列梯度递推关系:

$$\hat{\boldsymbol{\vartheta}}(t) = \hat{\boldsymbol{\vartheta}}(t-1) + \frac{\boldsymbol{\varphi}(t)}{r(t)} [y(t) - \boldsymbol{\varphi}^T(t) \hat{\boldsymbol{\vartheta}}(t-1)], \quad (100)$$

$$r(t) = r(t-1) + \|\boldsymbol{\varphi}(t)\|^2, \quad r(0) = 1. \quad (101)$$

由于信息向量 $\boldsymbol{\varphi}(t)$ 中的子信息向量 $\boldsymbol{\varphi}_a(t)$ 和 $\boldsymbol{\varphi}_b(t)$ 中分别含有未知内部变量 $x(t-i)$ 和 $\bar{u}(t-i)$,因此算法(100)—(101)不可实现.解决思路仍然是借助辅助模型辨识思想,根据式(94)中 $\bar{u}(t)$ 表达式的结构,用系统输入 $u(t)$ 和参数估计 $\hat{\boldsymbol{\vartheta}}(t)$ 构造估算 $\bar{u}(t)$ 的非线性辅助模型:

$$\bar{u}_a(t) = f_a(u(t)) = \sum_{j=1}^m \hat{c}_j(t) f_j(u(t)) = \mathbf{f}^T(t) \hat{\mathbf{c}}(t).$$

用非线性辅助模型的输出 $\bar{u}_a(t-i)$ 构造 $\boldsymbol{\varphi}_b(t)$ 的估计

$$\hat{\boldsymbol{\varphi}}_b(t) := [\bar{u}_a(t-1), \bar{u}_a(t-2), \dots, \bar{u}_a(t-n_b)]^T \in \mathbf{R}^{n_b},$$

用辅助模型的输出 $x_a(t-i)$ 构造 $\boldsymbol{\varphi}_a(t)$ 的估计

$$\hat{\boldsymbol{\varphi}}_a(t) := [-x_a(t-1), -x_a(t-2), \dots, -x_a(t-n_a)]^T \in \mathbf{R}^{n_a}.$$

进而构造信息向量 $\boldsymbol{\varphi}(t)$ 的估计

$$\hat{\boldsymbol{\varphi}}(t) := [\hat{\boldsymbol{\varphi}}_a^T(t), \hat{\boldsymbol{\varphi}}_b^T(t), \mathbf{f}^T(t)]^T \in \mathbf{R}^{n_0}.$$

根据式(97),用 $\hat{\boldsymbol{\varphi}}(t)$ 和 $\hat{\boldsymbol{\vartheta}}(t)$ 定义估算 $x(t)$ 的辅助模型

$$x_a(t) := \hat{\boldsymbol{\varphi}}_a^T(t) \hat{\mathbf{a}}(t) + \hat{\boldsymbol{\varphi}}_b^T(t) \hat{\mathbf{b}}(t) + \bar{u}_a(t) = \hat{\boldsymbol{\varphi}}^T(t) \hat{\boldsymbol{\vartheta}}(t).$$

辅助模型的输出 $x_a(t)$ 可作为 $x(t)$ 的估计 $\hat{x}(t)$,即 $\hat{x}(t) = x_a(t)$;非线性辅助模型的输出 $\bar{u}_a(t)$ 可作为 $\bar{u}(t)$ 的估计 $\hat{\bar{u}}(t)$,即 $\hat{\bar{u}}(t) = \bar{u}_a(t)$.

将式(100)—(101)中未知信息向量 $\boldsymbol{\varphi}(t)$ 用其估计 $\hat{\boldsymbol{\varphi}}(t)$ 代替,就得到了辨识参数向量 $\boldsymbol{\vartheta}$ 的基于关键项分离的辅助模型随机梯度算法(KT-AM-SG算法):

$$\hat{\boldsymbol{\vartheta}}(t) = \hat{\boldsymbol{\vartheta}}(t-1) + \frac{\hat{\boldsymbol{\varphi}}(t)}{r(t)} [y(t) - \hat{\boldsymbol{\varphi}}^T(t) \hat{\boldsymbol{\vartheta}}(t-1)], \quad (102)$$

$$r(t) = r(t-1) + \|\hat{\boldsymbol{\varphi}}(t)\|^2, \quad (103)$$

$$\hat{\boldsymbol{\varphi}}(t) = [\hat{\boldsymbol{\varphi}}_a^T(t), \hat{\boldsymbol{\varphi}}_b^T(t), \mathbf{f}^T(t)]^T, \quad (104)$$

$$\hat{\boldsymbol{\varphi}}_a(t) = [-x_a(t-1), -x_a(t-2), \dots, -x_a(t-n_a)]^T, \quad (105)$$

$$\hat{\boldsymbol{\varphi}}_b(t) = [\bar{u}_a(t-1), \bar{u}_a(t-2), \dots, \bar{u}_a(t-n_b)]^T, \quad (106)$$

$$\mathbf{f}(t) = [f_1(u(t)), f_2(u(t)), \dots, f_m(u(t))]^T, \quad (107)$$

$$\bar{u}_a(t) = \mathbf{f}^T(t) \hat{\mathbf{c}}(t), \quad (108)$$

$$x_a(t) = \hat{\boldsymbol{\varphi}}^T(t) \hat{\boldsymbol{\vartheta}}(t), \quad (109)$$

$$\hat{\boldsymbol{\vartheta}}(t) = [\hat{\mathbf{a}}^T(t), \hat{\mathbf{b}}^T(t), \hat{\mathbf{c}}^T(t)]^T. \quad (110)$$

利用 KT-AM-SG 算法(102)—(110)计算参数估计向量的步骤如下:

1) 初始化:令 $t = 1$, 给定数据窗长度 L_e 和基函数 $f_j(\cdot)$, 置初值 $\hat{\boldsymbol{\vartheta}}(0) = \mathbf{1}_{n_0}/p_0, r(0) = 1, x_a(t-i) = 1/p_0, \bar{u}_a(t-i) = 1/p_0, i = 0, 1, 2, \dots, \max[n_a, n_b], p_0 = 10^6$.

2) 采集输入输出数据 $u(t)$ 和 $y(t)$, 根据式

(105)–(107)构造 $\hat{\varphi}_a(t)$, $\hat{\varphi}_b(t)$ 和 $f(t)$.

3) 用式(104)构造 $\hat{\varphi}(t)$, 用式(103)计算 $r(t)$.

4) 用式(102)刷新参数估计向量 $\hat{\boldsymbol{\theta}}(t)$.

5) 从式(110)的参数估计向量 $\hat{\boldsymbol{\theta}}(t)$ 中读出 $\hat{c}(t)$.

6) 用式(108)–(109)计算辅助模型的输出 $\bar{u}_a(t)$ 和 $x_a(t)$.

7) 如果 $t < L_e$, t 增加 1, 转到第 2) 步; 否则, 终止计算, 获得参数估计向量 $\hat{\boldsymbol{\theta}}(L_e)$.

8) 提取参数估计向量 $\hat{\boldsymbol{a}}(L_e)$, $\hat{\boldsymbol{b}}(L_e)$ 和 $\hat{\boldsymbol{c}}(L_e)$.

注 12 为提高参数估计精度, 可在式(103)中引入遗忘因子 λ , 即

$$r(t) = \lambda r(t-1) + \|\hat{\varphi}(t)\|^2, \quad 0 \leq \lambda \leq 1, \quad r(0) = 1. \quad (111)$$

就得到遗忘因子 KT-AM-SG 算法.

注 13 与 O-AM-SG 算法(22)–(27)相比, KT-AM-SG 算法(102)–(110)参数估计向量的维数小, 故计算量小. 读者可以对两个算法每递推计算一步需要的乘法次数和加法次数进行比较.

3.3 基于关键项分离的辅助模型递推最小二乘辨识方法

考虑式(99)的关键项分离辨识模型, 即

$$y(t) = \boldsymbol{\varphi}^T(t) \boldsymbol{\theta} + v(t), \quad (112)$$

其中各变量定义同上. 定义二次准则函数

$$J_6(\boldsymbol{\theta}) := \frac{1}{2} \sum_{j=1}^t [y(j) - \boldsymbol{\varphi}^T(j) \boldsymbol{\theta}]^2.$$

极小化这个二次准则函数或令 $J_6(\boldsymbol{\theta})$ 对 $\boldsymbol{\theta}$ 的偏导数为零, 算法中涉及的未知变量用其估计代替或辅助模型的输出代替, 能够得到辨识参数向量 $\boldsymbol{\theta}$ 的基于关键项分离的辅助模型递推最小二乘算法(KT-AM-RLS 算法):

$$\begin{aligned} \hat{\boldsymbol{\theta}}(t) &= \hat{\boldsymbol{\theta}}(t-1) + \mathbf{L}(t) [y(t) - \hat{\boldsymbol{\varphi}}^T(t) \hat{\boldsymbol{\theta}}(t-1)], \\ \hat{\boldsymbol{\theta}}(0) &= \mathbf{1}_{n_0}/p_0, \end{aligned} \quad (113)$$

$$\mathbf{L}(t) = \mathbf{P}(t-1) \hat{\boldsymbol{\varphi}}(t) [1 + \hat{\boldsymbol{\varphi}}^T(t) \mathbf{P}(t-1) \hat{\boldsymbol{\varphi}}(t)]^{-1}, \quad (114)$$

$$\mathbf{P}(t) = [\mathbf{I}_{n_0} - \mathbf{L}(t) \hat{\boldsymbol{\varphi}}^T(t)] \mathbf{P}(t-1), \quad \mathbf{P}_0(0) = p_0 \mathbf{I}_{n_0}, \quad (115)$$

$$\hat{\boldsymbol{\varphi}}(t) = [\hat{\boldsymbol{\varphi}}_a^T(t), \hat{\boldsymbol{\varphi}}_b^T(t), \mathbf{f}^T(t)]^T, \quad (116)$$

$$\hat{\boldsymbol{\varphi}}_a(t) = [-x_a(t-1), -x_a(t-2), \dots, -x_a(t-n_a)]^T, \quad (117)$$

$$\hat{\boldsymbol{\varphi}}_b(t) = [\bar{u}_a(t-1), \bar{u}_a(t-2), \dots, \bar{u}_a(t-n_b)]^T, \quad (118)$$

$$\mathbf{f}(t) = [f_1(u(t)), f_2(u(t)), \dots, f_m(u(t))]^T, \quad (119)$$

$$\bar{u}_a(t) = \mathbf{f}^T(t) \hat{\boldsymbol{c}}(t), \quad (120)$$

$$x_a(t) = \hat{\boldsymbol{\varphi}}^T(t) \hat{\boldsymbol{\theta}}(t), \quad (121)$$

$$\hat{\boldsymbol{\theta}}(t) = [\hat{\boldsymbol{a}}^T(t), \hat{\boldsymbol{b}}^T(t), \hat{\boldsymbol{c}}^T(t)]^T. \quad (122)$$

利用 KT-AM-RLS 算法(113)–(122)计算 IN-OE 系统参数估计的步骤如下, 每一步的计算量如表 3 所示 ($n_0 = n_a + n_b + m$).

1) 初始化: 令 $t = 1$, 置初值 $\hat{\boldsymbol{\theta}}(0) = \mathbf{1}_{n_0}/p_0$, $x_a(t-i) = 1/p_0$, $\bar{u}_a(t-i) = 1/p_0$, $i = 0, 1, 2, \dots, \max[n_a, n_b]$, $p_0 = 10^6$.

2) 采集输入输出数据 $u(t)$ 和 $y(t)$, 用式(117)–(119)构造 $\hat{\boldsymbol{\varphi}}_a(t)$, $\hat{\boldsymbol{\varphi}}_b(t)$ 和 $f(t)$.

3) 用式(116)构造 $\hat{\boldsymbol{\varphi}}(t)$.

4) 用式(114)计算增益矩阵 $\mathbf{L}(t)$, 用式(115)计算协方差阵 $\mathbf{P}(t)$.

5) 用式(113)刷新参数估计向量 $\hat{\boldsymbol{\theta}}(t)$.

6) 根据式(120)–(121)计算辅助模型的输出 $\bar{u}_a(t)$ 和 $x_a(t)$.

7) 如果 $t < L_e$, t 增加 1, 转到第 2) 步; 否则, 终止计算, 获得参数估计向量 $\hat{\boldsymbol{\theta}}(L_e)$.

8) 输出参数估计向量 $\hat{\boldsymbol{a}}(L_e)$, $\hat{\boldsymbol{b}}(L_e)$ 和 $\hat{\boldsymbol{c}}(L_e)$.

4 基于关键项分离的辅助模型两阶段辨识方法

输入非线性输出误差系统关键项分离辨识模型(99)参数向量含 $n_a + n_b + m$ 个参数, 相对于过参数化

表 3 基于关键项分离的辅助模型递推最小二乘算法的计算量

Table 3 The computational efficiency of the KT-AM-RLS algorithm

变量	表达式	乘法次数	加法次数
$\hat{\boldsymbol{\theta}}(t)$	$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t-1) + \mathbf{L}(t) e(t) \in \mathbf{R}^{n_0}$	n_0	n_0
	$e(t) := y(t) - \hat{\boldsymbol{\varphi}}^T(t) \hat{\boldsymbol{\theta}}(t-1) \in \mathbf{R}$	n_0	n_0
$\mathbf{L}(t)$	$\mathbf{L}(t) = \boldsymbol{\zeta}(t) / [1 + \hat{\boldsymbol{\varphi}}^T(t) \boldsymbol{\zeta}(t)] \in \mathbf{R}^{n_0}$	$2n_0$	n_0
	$\boldsymbol{\zeta}(t) := \mathbf{P}(t-1) \hat{\boldsymbol{\varphi}}(t) \in \mathbf{R}^{n_0}$	n_0^2	$(n_0 - 1)n_0$
$\mathbf{P}(t)$	$\mathbf{P}(t) = \mathbf{P}(t-1) - \mathbf{L}(t) \boldsymbol{\zeta}^T(t) \in \mathbf{R}^{n_0 \times n_0}$	n_0^2	n_0^2
$\bar{u}_a(t)$	$\bar{u}_a(t) = \mathbf{f}^T(t) \hat{\boldsymbol{c}}(t) \in \mathbf{R}$	m	$m-1$
$x_a(t)$	$x_a(t) = \hat{\boldsymbol{\varphi}}^T(t) \hat{\boldsymbol{\theta}}(t) \in \mathbf{R}$	n_0	n_0-1
总数		$2n_0^2 + 5n_0 + m$	$2n_0^2 + 3n_0 + m - 2$
总 flop 数		$N_3 := 4(n_a + n_b + m)^2 + 8(n_a + n_b + m) + 2m - 2$	

方法来说,待辨识参数个数减少了,算法计算量减小了,辨识效率提高了.若线性块和非线性块的参数过多,计算量还是很大的.为进一步提高辨识效率,利用递阶辨识原理,将关键项分离辨识模型(99)分解为两个或三个子系统,研究相应的辨识方法.

4.1 基于关键项分离的两阶段辨识模型

考虑输入非线性输出误差(IN-OE)系统,其基于关键项分离的辨识模型(99),(97)和(94)重写如下:

$$y(t) = \boldsymbol{\varphi}^T(t) \boldsymbol{\vartheta} + v(t), \quad (123)$$

$$x(t) = \boldsymbol{\varphi}^T(t) \boldsymbol{\vartheta} = \boldsymbol{\varphi}_a^T(t) \boldsymbol{a} + \boldsymbol{\varphi}_b^T(t) \boldsymbol{b} + \bar{u}(t), \quad (124)$$

$$\bar{u}(t) = \boldsymbol{f}^T(u(t)) \boldsymbol{c}, \quad (125)$$

其中参数向量和信息向量定义为

$$\boldsymbol{\vartheta} := [\boldsymbol{a}^T, \boldsymbol{b}^T, \boldsymbol{c}^T]^T \in \mathbf{R}^{n_0}, \quad n_0 := n_a + n_b + m,$$

$$\boldsymbol{a} := [a_1, a_2, \dots, a_{n_a}]^T \in \mathbf{R}^{n_a},$$

$$\boldsymbol{b} := [b_1, b_2, \dots, b_{n_b}]^T \in \mathbf{R}^{n_b},$$

$$\boldsymbol{c} := [c_1, c_2, \dots, c_m]^T \in \mathbf{R}^m,$$

$$\boldsymbol{\varphi}(t) := [\boldsymbol{\varphi}_a^T(t), \boldsymbol{\varphi}_b^T(t), \boldsymbol{f}^T(t)]^T \in \mathbf{R}^{n_0},$$

$$\boldsymbol{\varphi}_a(t) := [-x(t-1), -x(t-2), \dots, -x(t-n_a)]^T \in \mathbf{R}^{n_a},$$

$$\boldsymbol{\varphi}_b(t) := [\bar{u}(t-1), \bar{u}(t-2), \dots, \bar{u}(t-n_b)]^T \in \mathbf{R}^{n_b},$$

$$\boldsymbol{f}(t) := \boldsymbol{f}(u(t)) = [f_1(u(t)), f_2(u(t)), \dots, f_m(u(t))]^T \in \mathbf{R}^m.$$

根据参数向量 $\boldsymbol{\vartheta}$ 和信息向量 $\boldsymbol{\varphi}(t)$ 的定义,式(123)可等价表示为

$$y(t) = \boldsymbol{\varphi}_a^T(t) \boldsymbol{a} + \boldsymbol{\varphi}_b^T(t) \boldsymbol{b} + \boldsymbol{f}^T(t) \boldsymbol{c} + v(t). \quad (126)$$

式(126)包含了3个子参数向量 \boldsymbol{a} , \boldsymbol{b} 和 \boldsymbol{c} , 可以把任意两个参数向量合在一起,形成一个两阶段辨识模型.这里把线性部分参数向量 \boldsymbol{a} 和 \boldsymbol{b} 合在一起,则式(123)可等价表示为

$$y(t) = \boldsymbol{\varphi}_1^T(t) \boldsymbol{\theta} + \boldsymbol{f}^T(t) \boldsymbol{c} + v(t), \quad (127)$$

$$\boldsymbol{\theta} := \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{bmatrix} \in \mathbf{R}^{n_a+n_b}, \quad \boldsymbol{\varphi}_1(t) := \begin{bmatrix} \boldsymbol{\varphi}_a(t) \\ \boldsymbol{\varphi}_b(t) \end{bmatrix} \in \mathbf{R}^{n_a+n_b}.$$

针对基于关键项分离的两阶段辨识模型(127),定义两个中间变量:

$$y_1(t) := y(t) - \boldsymbol{f}^T(t) \boldsymbol{c} \in \mathbf{R}, \quad (128)$$

$$y_2(t) := y(t) - \boldsymbol{\varphi}_1^T(t) \boldsymbol{\theta} \in \mathbf{R}, \quad (129)$$

则式(127)可分解为两个子模型:

$$y_1(t) = \boldsymbol{\varphi}_1^T(t) \boldsymbol{\theta} + v(t), \quad (130)$$

$$y_2(t) = \boldsymbol{f}^T(t) \boldsymbol{c} + v(t). \quad (131)$$

式(130)—(131)构成了 IN-OE 系统的两阶段辨识模型.下面研究其参数向量 $\boldsymbol{\theta}$ 和 \boldsymbol{c} 的辅助模型随机梯度辨识算法和辅助模型递推最小二乘辨识算法.

4.2 基于关键项分离的辅助模型两阶段随机梯度辨识方法

针对两阶段辨识模型(130)和(131),定义两个准则函数:

$$J_7(\boldsymbol{\theta}) := \frac{1}{2} [y_1(t) - \boldsymbol{\varphi}_1^T(t) \boldsymbol{\theta}]^2,$$

$$J_8(\boldsymbol{c}) := \frac{1}{2} [y_2(t) - \boldsymbol{f}^T(t) \boldsymbol{c}]^2.$$

设 $\hat{\boldsymbol{\theta}}(t) := \begin{bmatrix} \hat{\boldsymbol{a}}(t) \\ \hat{\boldsymbol{b}}(t) \end{bmatrix} \in \mathbf{R}^{n_a+n_b}$ 是参数向量 $\boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{bmatrix}$ 在时刻 t 的估计, $\hat{\boldsymbol{c}}(t) \in \mathbf{R}^m$ 是参数向量 \boldsymbol{c} 在时刻 t 的估计.利用负梯度搜索,极小化这两个准则函数,算法中涉及的未知变量用其估计代替或辅助模型的输出代替,能够得到辨识系统参数向量 $\boldsymbol{\theta}$ 和 \boldsymbol{c} 的基于关键项分离的辅助模型两阶段随机梯度算法(KT-AM-2S-SG算法):

$$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t-1) + \frac{\hat{\boldsymbol{\varphi}}_1(t)}{r_1(t)} [y(t) - \hat{\boldsymbol{\varphi}}_1^T(t) \hat{\boldsymbol{\theta}}(t-1) - \boldsymbol{f}^T(t) \hat{\boldsymbol{c}}(t-1)], \quad (132)$$

$$r_1(t) = r_1(t-1) + \|\hat{\boldsymbol{\varphi}}_1(t)\|^2, \quad r_1(0) = 1, \quad (133)$$

$$\hat{\boldsymbol{c}}(t) = \hat{\boldsymbol{c}}(t-1) + \frac{\boldsymbol{f}(t)}{r_2(t)} [y(t) - \hat{\boldsymbol{\varphi}}_1^T(t) \hat{\boldsymbol{\theta}}(t-1) - \boldsymbol{f}^T(t) \hat{\boldsymbol{c}}(t-1)], \quad (134)$$

$$r_2(t) = r_2(t-1) + \|\boldsymbol{f}(t)\|^2, \quad r_2(0) = 1, \quad (135)$$

$$\hat{\boldsymbol{\varphi}}_1(t) = [-x_a(t-1), -x_a(t-2), \dots, -x_a(t-n_a), \bar{u}_a(t-1), \bar{u}_a(t-2), \dots, \bar{u}_a(t-n_b)]^T, \quad (136)$$

$$\boldsymbol{f}(t) = [f_1(u(t)), f_2(u(t)), \dots, f_m(u(t))]^T, \quad (137)$$

$$\bar{u}_a(t) = \boldsymbol{f}^T(t) \hat{\boldsymbol{c}}(t), \quad (138)$$

$$x_a(t) = \hat{\boldsymbol{\varphi}}_1^T(t) \hat{\boldsymbol{\theta}}(t) + \bar{u}_a(t). \quad (139)$$

对于上述的两阶段随机梯度算法,可在式(133)和(135)中引入遗忘因子,即

$$r_1(t) = \lambda_1 r_1(t-1) + \|\hat{\boldsymbol{\varphi}}_1(t)\|^2, \quad 0 < \lambda_1 \leq 1,$$

$$r_2(t) = \lambda_2 r_2(t-1) + \|\boldsymbol{f}(t)\|^2, \quad 0 < \lambda_2 \leq 1.$$

注14 如果式(132)和(134)中的 $r_1(t)$ 和 $r_2(t)$ 采用一个共同的收敛因子 $r(t)$, 并取

$$r(t) = r(t-1) + \|\hat{\boldsymbol{\varphi}}_1(t)\|^2 + \|\boldsymbol{f}(t)\|^2,$$

则 KT-AM-2S-SG 算法(132)—(139)完全等同于 KT-AM-SG 算法(102)—(110).

利用 KT-2S-SG 算法(132)—(139)计算 IN-OE 系统参数估计向量的步骤如下:

1) 初始化: 令 $t = 1$, 给定数据长度 L_e 和非线性函数 $f_j(\cdot)$, 置初值 $\hat{\boldsymbol{\theta}}(0) = \mathbf{1}_{n_a+n_b}/p_0$, $\hat{\boldsymbol{c}}(0) = \mathbf{1}_m/p_0$,

$x_a(t-i) = 1/p_0, \bar{u}_a(t-i) = 1/p_0, i=0, 1, 2, \dots, \max[n_a, n_b], p_0 = 10^6$.

2) 采集输入输出数据 $u(t)$ 和 $y(t)$, 根据式 (136) — (137) 构造 $\hat{\varphi}_1(t)$ 和 $f(t)$.

3) 用式 (133) 计算 $r_1(t)$, 用式 (135) 计算 $r_2(t)$.

4) 用式 (132) 刷新参数估计向量 $\hat{\theta}(t)$, 用式 (134) 刷新参数估计向量 $\hat{c}(t)$.

5) 用式 (138) — (139) 计算辅助模型的输出 $\bar{u}_a(t)$ 和 $x_a(t)$.

6) 如果 $t < L_e$, t 增加 1, 转到第 2) 步; 否则, 终止计算, 获得参数估计向量 $\hat{\theta}(L_e)$ 和 $\hat{c}(L_e)$.

4.3 基于关键项分离的辅助模型两阶段递推最小二乘辨识方法

考虑两阶段辨识模型 (130) 和 (131), 重写如下:

$$y_1(t) = \varphi_1^T(t)\theta + v(t), \quad (140)$$

$$y_2(t) = f^T(t)c + v(t), \quad (141)$$

其中

$$y_1(t) := y(t) - f^T(t)c,$$

$$y_2(t) := y(t) - \varphi_1^T(t)\theta.$$

根据辨识模型 (140) — (141), 定义两个准则函数

$$J_9(\theta) := \frac{1}{2} \sum_{j=1}^t [y_1(j) - \varphi_1^T(j)\theta]^2,$$

$$J_{10}(c) := \frac{1}{2} \sum_{j=1}^t [y_2(j) - f^T(j)c]^2.$$

极小化准则函数 $J_9(\theta)$ 和 $J_{10}(c)$, 参考递推最小二乘算法的推导, 算法中涉及的未知变量用其估计代替或辅助模型的输出代替, 联立辅助模型 (136) — (139), 可得辨识参数向量的基于关键项分离的辅助模型两阶段递推最小二乘算法 (KT-AM-2S-RLS 算法):

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L_1(t) [y(t) - \hat{\varphi}_1^T(t)\hat{\theta}(t-1) - f^T(t)\hat{c}(t-1)], \quad (142)$$

$$L_1(t) = P_1(t-1)\hat{\varphi}_1(t) [1 + \hat{\varphi}_1^T(t)P_1(t-1)\hat{\varphi}_1(t)]^{-1}, \quad (143)$$

$$P_1(t) = [I_{n_a+n_b} - L_1(t)\hat{\varphi}_1^T(t)]P_1(t-1), \quad (144)$$

$$\hat{c}(t) = \hat{c}(t-1) + L_2(t) [y(t) - \hat{\varphi}_1^T(t)\hat{\theta}(t-1) - f^T(t)\hat{c}(t-1)], \quad (145)$$

$$L_2(t) = P_2(t-1)f(t) [1 + f^T(t)P_2(t-1)f(t)]^{-1}, \quad (146)$$

$$P_2(t) = [I_m - L_2(t)f^T(t)]P_2(t-1), \quad (147)$$

$$\hat{\varphi}_1(t) = [-x_a(t-1), -x_a(t-2), \dots, -x_a(t-n_a), \bar{u}_a(t-1), \dots, \bar{u}_a(t-n_b)]^T, \quad (148)$$

$$f(t) = [f_1(u(t)), f_2(u(t)), \dots, f_m(u(t))]^T, \quad (149)$$

$$\bar{u}_a(t) = f^T(t)\hat{c}(t), \quad (150)$$

$$x_a(t) = \hat{\varphi}_1^T(t)\hat{\theta}(t) + \bar{u}_a(t). \quad (151)$$

算法中参数估计向量和协方差矩阵, 以及相关变量初始值可设置为

$$\hat{\theta}(0) = \mathbf{1}_{n_a+n_b}/p_0, \quad P_1(0) = p_0 I_{n_a+n_b}, \quad p_0 = 10^6, \quad (152)$$

$$\hat{c}(0) = \mathbf{1}_m/p_0, \quad P_2(0) = p_0 I_m, \quad (153)$$

$$x_a(i) = 1/p_0, \quad \bar{u}_a(-i) = 1/p_0, \quad i=0, 1, 2, \dots, \max[n_a, n_b]. \quad (154)$$

利用 KT-AM-2S-RLS 算法 (142) — (153) 计算 IN-OE 系统参数估计向量的步骤如下, 每步的计算量如表 4 所示.

1) 初始化: 令 $t=1$, 给定数据长度 L_e 和非线性函数 $f_j(\cdot)$, 按照式 (152) — (154) 置初值.

2) 采集输入输出数据 $u(t)$ 和 $y(t)$, 根据式 (148) — (149) 构造 $\hat{\varphi}_1(t)$ 和 $f(t)$.

3) 用式 (143) — (144) 计算 $L_1(t)$ 和 $P_1(t)$, 用式 (146) — (147) 计算 $L_2(t)$ 和 $P_2(t)$.

4) 用式 (142) 刷新参数估计向量 $\hat{\theta}(t)$, 用式 (145) 刷新参数估计向量 $\hat{c}(t)$.

5) 根据式 (150) — (151) 计算辅助模型的输出 $\bar{u}_a(t)$ 和 $x_a(t)$.

6) 如果 $t < L_e$, t 增加 1, 转到第 2) 步; 否则, 终止计算, 获得估计 $\hat{\theta}(L_e)$ 和 $\hat{c}(L_e)$.

注 15 最小二乘算法的计算量是 $O(n^2)$ (n 是参数向量的维数, 即协方差阵的维数). KT-AM-RLS 算法的计算量为 $N_3 = O[(n_a+n_b+m)^2]$, 而 KT-AM-2S-RLS 算法通过分解, 使得参数向量的维数分别为 n_a+n_b 和 n_c , 尽管要计算两个协方差阵, 其计算量 $N_4 = O[(n_a+n_b)^2+m^2]$ 也比 KT-AM-RLS 算法 (113) — (122) 小, 因为 $(n_a+n_b)^2+m^2 \leq (n_a+n_b+m)^2$.

5 基于关键项分离的辅助模型三阶段辨识方法

将关键项分离辨识模型分解为两个子系统进行辨识, 在一定程度上减小了辨识算法的计算量, 提高了辨识效率. 要进一步提高计算效率, 可以利用递阶辨识原理将系统分解成三个子系统进行辨识. 下面研究基于关键项分离的辅助模型三阶段辨识方法.

5.1 基于关键项分离的三阶段辨识模型

考虑输入非线性输出误差 (IN-OE) 系统, 其基于关键项分离的辨识模型 (99)、(97) 和 (94) 重写如下:

表 4 基于关键项分离的辅助模型两阶段递推最小二乘算法的计算量

Table 4 The computational efficiency of the KT-AM-2S-RLS algorithm

变量	表达式	乘法次数	加法次数
$\hat{\theta}(t)$	$\hat{\theta}(t) = \hat{\theta}(t-1) + L_1(t)e(t) \in \mathbf{R}^{n_a+n_b}$	n_a+n_b	n_a+n_b
$e(t)$	$e(t) := y(t) - \hat{\varphi}_1^T(t)\hat{\theta}(t-1) - f^T(t)\hat{c}(t-1) \in \mathbf{R}$	n_a+n_b+m	n_a+n_b+m
$L_1(t)$	$L_1(t) = \zeta_1(t) / [1 + \hat{\varphi}_1^T(t)\zeta_1(t)] \in \mathbf{R}^{n_a+n_b}$	$2(n_a+n_b)$	n_a+n_b
$\zeta_1(t)$	$\zeta_1(t) := P_1(t-1)\hat{\varphi}_1(t) \in \mathbf{R}^{n_a+n_b}$	$(n_a+n_b)^2$	$(n_a+n_b)^2 - (n_a+n_b)$
$P_1(t)$	$P_1(t) = P_1(t-1) - L_1(t)\zeta_1^T(t) \in \mathbf{R}^{(n_a+n_b) \times (n_a+n_b)}$	$(n_a+n_b)^2$	$(n_a+n_b)^2$
$\hat{c}(t)$	$\hat{c}(t) = \hat{c}(t-1) + L_2(t)e(t) \in \mathbf{R}^m$	m	m
$L_2(t)$	$L_2(t) = \zeta_2(t) / [1 + f^T(t)\zeta_2(t)] \in \mathbf{R}^m$	$2m$	m
$\zeta_2(t)$	$\zeta_2(t) := P_2(t-1)f(t) \in \mathbf{R}^m$	m^2	$m^2 - m$
$P_2(t)$	$P_2(t) = P_2(t-1) - L_2(t)\zeta_2^T(t) \in \mathbf{R}^{m \times m}$	m^2	m^2
$\bar{u}_a(t)$	$\bar{u}_a(t) = f^T(t)\hat{c}(t) \in \mathbf{R}$	m	$m-1$
$x_a(t)$	$x_a(t) = \hat{\varphi}_1^T(t)\hat{\theta}(t) + \bar{u}_a(t) \in \mathbf{R}$	n_a+n_b	n_a+n_b
总数		$2(n_a+n_b)^2 + 2m^2 + 5(n_a+n_b+m)$	$2(n_a+n_b)^2 + 2m^2 + 3(n_a+n_b+m) - 1$
总 flop 数		$N_4 := 4(n_a+n_b)^2 + 4m^2 + 8(n_a+n_b+m) - 1$	

$$y(t) = \varphi^T(t)\vartheta + v(t) \quad (155)$$

$$= \varphi_a^T(t)\mathbf{a} + \varphi_b^T(t)\mathbf{b} + f^T(t)\mathbf{c} + v(t), \quad (156)$$

$$x(t) = \varphi^T(t)\vartheta = \varphi_a^T(t)\mathbf{a} + \varphi_b^T(t)\mathbf{b} + \bar{u}(t), \quad (157)$$

$$\bar{u}(t) = f^T(t)\mathbf{c}, \quad (158)$$

其中参数向量和信息向量定义如下:

$$\vartheta := [\mathbf{a}^T, \mathbf{b}^T, \mathbf{c}^T]^T \in \mathbf{R}^{n_0}, \quad n_0 := n_a + n_b + m,$$

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

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

$$\mathbf{c} := [c_1, c_2, \dots, c_m]^T \in \mathbf{R}^m,$$

$$\varphi(t) := [\varphi_a^T(t), \varphi_b^T(t), f^T(t)]^T \in \mathbf{R}^{n_0},$$

$$\varphi_a(t) := [-x(t-1), -x(t-2), \dots, -x(t-n_a)]^T \in \mathbf{R}^{n_a},$$

$$\varphi_b(t) := [\bar{u}(t-1), \bar{u}(t-2), \dots, \bar{u}(t-n_b)]^T \in \mathbf{R}^{n_b},$$

$$f(t) := [f_1(u(t)), f_2(u(t)), \dots, f_m(u(t))]^T \in \mathbf{R}^m.$$

式(156)中输出 $y(t)$ 是 3 个参数向量 \mathbf{a} , \mathbf{b} 和 \mathbf{c} 的线性组合, 因此, 可利用递阶辨识原理, 将该系统分解成 3 个虚拟子系统.

定义 3 个中间变量:

$$y_a(t) := y(t) - \varphi_b^T(t)\mathbf{b} - f^T(t)\mathbf{c}, \quad (159)$$

$$y_b(t) := y(t) - \varphi_a^T(t)\mathbf{a} - f^T(t)\mathbf{c}, \quad (160)$$

$$y_c(t) := y(t) - \varphi_a^T(t)\mathbf{a} - \varphi_b^T(t)\mathbf{b}. \quad (161)$$

于是, 系统(156)可以分解为 3 个子系统,

$$y_a(t) = \varphi_a^T(t)\mathbf{a} + v(t), \quad (162)$$

$$y_b(t) = \varphi_b^T(t)\mathbf{b} + v(t), \quad (163)$$

$$y_c(t) = f^T(t)\mathbf{c} + v(t). \quad (164)$$

此即基于关键项分离的三阶段辨识模型. 下面讨论该辨识模型的辅助模型三阶段随机梯度辨识方法和辅助模型三阶段最小二乘辨识方法.

5.2 基于关键项分离的辅助模型三阶段随机梯度辨识方法

针对 3 个子系统(162) — (164), 定义 3 个梯度准则函数:

$$J_a(\mathbf{a}) := \frac{1}{2} [y_a(t) - \varphi_a^T(t)\mathbf{a}]^2,$$

$$J_b(\mathbf{b}) := \frac{1}{2} [y_b(t) - \varphi_b^T(t)\mathbf{b}]^2,$$

$$J_c(\mathbf{c}) := \frac{1}{2} [y_c(t) - f^T(t)\mathbf{c}]^2.$$

设 $\hat{\mathbf{a}}(t) \in \mathbf{R}^{n_a}$, $\hat{\mathbf{b}}(t) \in \mathbf{R}^{n_b}$ 和 $\hat{\mathbf{c}}(t) \in \mathbf{R}^m$ 分别是参数向量 \mathbf{a} , \mathbf{b} 和 \mathbf{c} 在时刻 t 的估计. 利用负梯度搜索极小化准则函数 $J_a(\mathbf{a})$, $J_b(\mathbf{b})$ 和 $J_c(\mathbf{c})$, 算法中涉及的未知变量用其估计代替或辅助模型的输出代替, 可得辨识参数向量的基于关键项分离的辅助模型三阶段随机梯度算法 (KT-AM-3S-SG 算法):

$$\hat{\mathbf{a}}(t) = \hat{\mathbf{a}}(t-1) + \frac{\hat{\varphi}_a(t)}{r_1(t)} e(t), \quad (165)$$

$$e(t) = y(t) - \hat{\varphi}_a^T(t)\hat{\mathbf{a}}(t-1) - \hat{\varphi}_b^T(t)\hat{\mathbf{b}}(t-1) - f^T(t)\hat{\mathbf{c}}(t-1), \quad (166)$$

$$r_1(t) = r_1(t-1) + \|\hat{\varphi}_a(t)\|^2, \quad r_1(0) = 1, \quad (167)$$

$$\hat{\mathbf{b}}(t) = \hat{\mathbf{b}}(t-1) + \frac{\hat{\varphi}_b(t)}{r_2(t)} e(t), \quad (168)$$

$$r_2(t) = r_2(t-1) + \|\hat{\varphi}_b(t)\|^2, \quad r_2(0) = 1, \quad (169)$$

$$\hat{\mathbf{c}}(t) = \hat{\mathbf{c}}(t-1) + \frac{\mathbf{f}(t)}{r_3(t)} e(t), \quad (170)$$

$$r_3(t) = r_3(t-1) + \|\mathbf{f}(t)\|^2, \quad r_3(0) = 1, \quad (171)$$

$$\hat{\varphi}_a(t) = [-x_a(t-1), -x_a(t-2), \dots, -x_a(t-n_a)]^T, \quad (172)$$

$$\hat{\boldsymbol{\varphi}}_b(t) = [\bar{u}_a(t-1), \bar{u}_a(t-2), \dots, \bar{u}_a(t-n_b)]^T, \quad (173)$$

$$\mathbf{f}(t) = [f_1(u(t)), f_2(u(t)), \dots, f_m(u(t))]^T, \quad (174)$$

$$\bar{u}_a(t) = \mathbf{f}^T(t) \hat{\mathbf{c}}(t), \quad (175)$$

$$x_a(t) = \hat{\boldsymbol{\varphi}}_a^T(t) \hat{\mathbf{a}}(t) + \hat{\boldsymbol{\varphi}}_b^T(t) \hat{\mathbf{b}}(t) + \bar{u}_a(t). \quad (176)$$

5.3 基于关键项分离的辅助模型三阶段递推最小二乘辨识方法

针对输入非线性输出误差系统的三阶段辨识模型(162)–(164),定义3个二次准则函数:

$$J_a(\mathbf{a}) := \frac{1}{2} \sum_{j=1}^l [y_a(j) - \boldsymbol{\varphi}_a^T(j) \mathbf{a}]^2,$$

$$J_b(\mathbf{b}) := \frac{1}{2} \sum_{j=1}^l [y_b(j) - \boldsymbol{\varphi}_b^T(j) \mathbf{b}]^2,$$

$$J_c(\mathbf{c}) := \frac{1}{2} \sum_{j=1}^l [y_c(j) - \mathbf{f}^T(j) \mathbf{c}]^2.$$

设 $\hat{\mathbf{a}}(t) \in \mathbf{R}^{n_a}$, $\hat{\mathbf{b}}(t) \in \mathbf{R}^{n_b}$ 和 $\hat{\mathbf{c}}(t) \in \mathbf{R}^m$ 分别是参数向量 \mathbf{a} , \mathbf{b} 和 \mathbf{c} 在时刻 t 的估计. 根据递推最小二乘算法的推导, 极小化准则函数 $J_a(\mathbf{a})$, $J_b(\mathbf{b})$ 和 $J_c(\mathbf{c})$, 算法中涉及的未知变量用其估计代替或辅助模型的输出代替, 联立辅助模型(172)–(176), 可以得到辨识参数向量的基于关键项分离的辅助模型三阶段递推最小二乘算法(KT-AM-3S-RLS算法):

$$\hat{\mathbf{a}}(t) = \hat{\mathbf{a}}(t-1) + \mathbf{L}_a(t) e(t), \quad \hat{\mathbf{a}}(0) = \mathbf{1}_{n_a}/p_0, \quad (177)$$

$$e(t) = y(t) - \hat{\boldsymbol{\varphi}}_a^T(t) \hat{\mathbf{a}}(t-1) - \hat{\boldsymbol{\varphi}}_b^T(t) \hat{\mathbf{b}}(t-1) - \mathbf{f}^T(t) \hat{\mathbf{c}}(t-1), \quad (178)$$

$$\mathbf{L}_a(t) = \mathbf{P}_a(t-1) \hat{\boldsymbol{\varphi}}_a(t) [1 + \hat{\boldsymbol{\varphi}}_a^T(t) \mathbf{P}_a(t-1) \hat{\boldsymbol{\varphi}}_a(t)]^{-1}, \quad (179)$$

$$\mathbf{P}_a(t) = [\mathbf{I}_{n_a} - \mathbf{L}_a(t) \hat{\boldsymbol{\varphi}}_a^T(t)] \mathbf{P}_a(t-1), \quad \mathbf{P}_a(0) = p_0 \mathbf{I}_{n_a}, \quad (180)$$

$$\hat{\mathbf{b}}(t) = \hat{\mathbf{b}}(t-1) + \mathbf{L}_b(t) e(t), \quad \hat{\mathbf{b}}(0) = \mathbf{1}_{n_b}/p_0, \quad (181)$$

$$\mathbf{L}_b(t) = \mathbf{P}_b(t-1) \hat{\boldsymbol{\varphi}}_b(t) [1 + \hat{\boldsymbol{\varphi}}_b^T(t) \mathbf{P}_b(t-1) \hat{\boldsymbol{\varphi}}_b(t)]^{-1}, \quad (182)$$

$$\mathbf{P}_b(t) = [\mathbf{I}_{n_b} - \mathbf{L}_b(t) \hat{\boldsymbol{\varphi}}_b^T(t)] \mathbf{P}_b(t-1), \quad \mathbf{P}_b(0) = p_0 \mathbf{I}_{n_b}, \quad (183)$$

$$\hat{\mathbf{c}}(t) = \hat{\mathbf{c}}(t-1) + \mathbf{L}_c(t) e(t), \quad \hat{\mathbf{c}}(0) = \mathbf{1}_m/p_0, \quad (184)$$

$$\mathbf{L}_c(t) = \mathbf{P}_c(t-1) \mathbf{f}(t) [1 + \mathbf{f}^T(t) \mathbf{P}_c(t-1) \mathbf{f}(t)]^{-1}, \quad (185)$$

$$\mathbf{P}_c(t) = [\mathbf{I}_m - \mathbf{L}_c(t) \mathbf{f}^T(t)] \mathbf{P}_c(t-1), \quad \mathbf{P}_c(0) = p_0 \mathbf{I}_m, \quad (186)$$

$$\hat{\boldsymbol{\varphi}}_a(t) = [-x_a(t-1), -x_a(t-2), \dots, -x_a(t-n_a)]^T, \quad (187)$$

$$\hat{\boldsymbol{\varphi}}_b(t) = [\bar{u}_a(t-1), \bar{u}_a(t-2), \dots, \bar{u}_a(t-n_b)]^T, \quad (188)$$

$$\mathbf{f}(t) = [f_1(u(t)), f_2(u(t)), \dots, f_m(u(t))]^T, \quad (189)$$

$$\bar{u}_a(t) = \mathbf{f}^T(t) \hat{\mathbf{c}}(t), \quad (190)$$

$$x_a(t) = \hat{\boldsymbol{\varphi}}_a^T(t) \hat{\mathbf{a}}(t) + \hat{\boldsymbol{\varphi}}_b^T(t) \hat{\mathbf{b}}(t) + \bar{u}_a(t), \quad x_a(-i) = 1/p_0,$$

$$\bar{u}_a(-i) = 1/p_0, \quad i = 0, 1, 2, \dots, \max[n_a, n_b]. \quad (191)$$

KT-AM-3S-RLS算法(177)–(191)每步的计算量如表5所示.

6 基于双线性参数模型分解的辅助模型辨识方法

前述基于分解的辨识方法都是将线性回归模型分解为多个子系统, 如将过参数化辨识模型分解为 $n_b + 1$ 个子系统的辅助模型递阶辨识方法, 将基于关键项分离的辨识模型分解为两个子系统的辅助模型两阶段辨识方法, 将基于关键项分离的辨识模型分解为三个子系统的辅助模型三阶段辨识方法, 而本节的方法直接对输入非线性输出误差系统的双线性参数模型进行分解, 研究相应的辨识方法, 包括基于双线性模型分解的辅助模型随机梯度算法和基于双线性模型分解的辅助模型递推最小二乘算法.

6.1 基于双线性参数模型分解的辨识模型

考虑输入非线性输出误差(IN-OE)系统(1)–(2), 重写如下:

$$y(t) = \frac{B(z)}{A(z)} \bar{u}(t) + v(t), \quad (192)$$

$$\bar{u}(t) = \sum_{j=1}^m c_j f_j(u(t)), \quad (193)$$

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

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

式中 $u(t) \in \mathbf{R}$ 和 $y(t) \in \mathbf{R}$ 分别为系统的输入和输出, $v(t) \in \mathbf{R}$ 是零均值白噪声, 非线性环节的输出 $\bar{u}(t) \in \mathbf{R}$ 是基函数 $\mathbf{f}(t) := [f_1(u(t)), f_2(u(t)), \dots, f_m(u(t))]$ 的线性组合, a_i , b_i 和 c_i 是未知待辨识参数.

定义参数向量 \mathbf{a} , \mathbf{b} 和 \mathbf{c} 如下:

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

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

$$\mathbf{c} := [c_1, c_2, \dots, c_m]^T \in \mathbf{R}^m,$$

定义信息向量 $\boldsymbol{\varphi}_a(t)$ 和信息矩阵 $\mathbf{F}(t)$ 如下:

$$\boldsymbol{\varphi}_a(t) := [-x(t-1), -x(t-2), \dots, -x(t-n_a)]^T \in \mathbf{R}^{n_a},$$

$$\mathbf{F}(t) := \begin{bmatrix} \mathbf{f}(u(t-1)) \\ \mathbf{f}(u(t-2)) \\ \vdots \\ \mathbf{f}(u(t-n_b)) \end{bmatrix} \in \mathbf{R}^{n_b \times m},$$

定义系统无噪声输出

$$x(t) := \frac{B(z)}{A(z)} \bar{u}(t).$$

利用移位算子的性质 $z^{-1}x(t) = x(t-1)$, 有

$$x(t) = [1 - A(z)]x(t) + B(z)\bar{u}(t) = -\sum_{i=1}^{n_a} a_i x(t-i) + \sum_{i=1}^{n_b} b_i \mathbf{f}(u(t-i)) \mathbf{c} =$$

表 5 基于关键项分离的辅助模型三阶段递推最小二乘算法的计算量

Table 5 The computational efficiency of the KT-AM-3S-RLS algorithm

变量	表达式	乘法次数	加法次数
$\hat{\mathbf{a}}(t)$	$\hat{\mathbf{a}}(t) = \hat{\mathbf{a}}(t-1) + \mathbf{L}_a(t)e(t) \in \mathbf{R}^{n_a}$	n_a	n_a
$e(t)$	$e(t) = y(t) - \hat{\boldsymbol{\varphi}}_a^T(t)\hat{\mathbf{a}}(t-1) - \mathbf{f}^T(t)\hat{\mathbf{c}}(t-1) - \hat{\boldsymbol{\varphi}}_b^T(t)\hat{\mathbf{b}}(t-1) \in n\mathbf{R} + n_b + m$		$n_a + n_b + m$
$\mathbf{L}_a(t)$	$\mathbf{L}_a(t) = \boldsymbol{\zeta}_1(t)/[1 + \hat{\boldsymbol{\varphi}}_a^T(t)\boldsymbol{\zeta}_1(t)] \in \mathbf{R}^{n_a}$	$2n_a$	n_a
$\boldsymbol{\zeta}_1(t)$	$\boldsymbol{\zeta}_1(t) := \mathbf{P}_a(t-1)\boldsymbol{\varphi}_a(t) \in \mathbf{R}^{n_a}$	n_a^2	$(n_a - 1)n_a$
$\mathbf{P}_a(t)$	$\mathbf{P}_a(t) = \mathbf{P}_a(t-1) - \mathbf{L}_a(t)\boldsymbol{\zeta}_1^T(t) \in \mathbf{R}^{n_a \times n_a}$	n_a^2	n_a^2
$\hat{\mathbf{b}}(t)$	$\hat{\mathbf{b}}(t) = \hat{\mathbf{b}}(t-1) + \mathbf{L}_b(t)e(t) \in \mathbf{R}^{n_b}$	n_b	n_b
$\mathbf{L}_b(t)$	$\mathbf{L}_b(t) = \boldsymbol{\zeta}_2(t)/[1 + \hat{\boldsymbol{\varphi}}_b^T(t)\boldsymbol{\zeta}_2(t)] \in \mathbf{R}^{n_b}$	$2n_b$	n_b
$\boldsymbol{\zeta}_2(t)$	$\boldsymbol{\zeta}_2(t) := \mathbf{P}_b(t-1)\hat{\boldsymbol{\varphi}}_b(t) \in \mathbf{R}^{n_b}$	n_b^2	$(n_b - 1)n_b$
$\mathbf{P}_b(t)$	$\mathbf{P}_b(t) = \mathbf{P}_b(t-1) - \mathbf{L}_b(t)\boldsymbol{\zeta}_2^T(t) \in \mathbf{R}^{n_b \times n_b}$	n_b^2	n_b^2
$\hat{\mathbf{c}}(t)$	$\hat{\mathbf{c}}(t) = \hat{\mathbf{c}}(t-1) + \mathbf{L}_c(t)e(t) \in \mathbf{R}^m$	m	m
$\mathbf{L}_c(t)$	$\mathbf{L}_c(t) = \boldsymbol{\zeta}_3(t)/[1 + \mathbf{f}^T(t)\boldsymbol{\zeta}_3(t)] \in \mathbf{R}^m$	$2m$	m
$\boldsymbol{\zeta}_3(t)$	$\boldsymbol{\zeta}_3(t) := \mathbf{P}_c(t-1)\mathbf{f}(t) \in \mathbf{R}^m$	m^2	$(m - 1)m$
$\mathbf{P}_c(t)$	$\mathbf{P}_c(t) = \mathbf{P}_c(t-1) - \mathbf{L}_c(t)\boldsymbol{\zeta}_3^T(t) \in \mathbf{R}^{m \times m}$	m^2	m^2
$\bar{\mathbf{u}}_a(t)$	$\bar{\mathbf{u}}_a(t) = \mathbf{f}^T(t)\hat{\mathbf{c}}(t) \in \mathbf{R}$	m	$m - 1$
$\mathbf{x}_a(t)$	$\mathbf{x}_a(t) = \hat{\boldsymbol{\varphi}}_a^T(t)\hat{\mathbf{a}}(t) + \hat{\boldsymbol{\varphi}}_b^T(t)\hat{\mathbf{b}}(t) + \bar{\mathbf{u}}_a(t) \in \mathbf{R}$	$n_a + n_b$	$n_a + n_b$
总数		$2(n_a^2 + n_b^2 + m^2) + 5(n_a + n_b)(n_a^2 + n_b^2 + m^2) + 3(n_a + n_b + m) - 1$	
总 flop 数		$N_5 := 4(n_a^2 + n_b^2 + m^2) + 8(n_a + n_b + m) - 1$	

$$\boldsymbol{\varphi}_a^T(t)\mathbf{a} + \mathbf{b}^T\mathbf{F}(t)\mathbf{c}, \quad (194)$$

于是,IN-OE 系统(192) — (193)可以表示为

$$y(t) = x(t) + v(t) = \boldsymbol{\varphi}_a^T(t)\mathbf{a} + \mathbf{b}^T\mathbf{F}(t)\mathbf{c} + v(t). \quad (195)$$

此即输入非线性输出误差系统的双线性参数辨识模型(bilinear-parameter identification model),因为该模型包含了系统参数向量 \mathbf{b} 和 \mathbf{c} 的乘积.

注 16 双线性参数辨识模型(195)包含了参数向量 \mathbf{b} 与 \mathbf{c} 的乘积,故是参数不可辨识的.为了使得参数可辨识,需要对其进行归一化.如

(A4) 假设 $\|\mathbf{c}\| = 1$,且 \mathbf{c} 的第一个元为正(当然也可以假设任意元为正);

(A5) 假设 $\|\mathbf{b}\| = 1$,且 \mathbf{b} 的第一个元为正(当然也可以假设任意元为正);

(A6) 假设线性动态子系统的增益 1,即

$$G(1) := \frac{B(1)}{A(1)} = \frac{b_1 + b_2 + \dots + b_{n_b}}{1 + a_1 + a_2 + \dots + a_{n_a}} = 1.$$

(A7) 当然也可以假设 b_i 或者 c_j 中的一个是定常数.如果假设 $b_1 = 1$,则本节的方法退化为基于关键项分离的辨识方法.

定义信息向量 $\boldsymbol{\varphi}(\mathbf{c}, t)$ 和 $\boldsymbol{\psi}(\mathbf{b}, t)$,以及虚拟输出 $y_1(t)$ 为

$$\boldsymbol{\varphi}(\mathbf{c}, t) := \begin{bmatrix} \boldsymbol{\varphi}_a(t) \\ \mathbf{F}(t)\mathbf{c} \end{bmatrix} \in \mathbf{R}^{n_a + n_b},$$

$$\boldsymbol{\psi}(\mathbf{b}, t) := \mathbf{F}^T(t)\mathbf{b} \in \mathbf{R}^m,$$

$$y_1(t) := y(t) - \boldsymbol{\varphi}_a^T(t)\mathbf{a}.$$

把双线性参数辨识模型(195)分成两个虚拟子系统,

一个含线性子系统参数向量 $\boldsymbol{\theta} := \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} \in \mathbf{R}^{n_a + n_b}$,另一个

个非线性部分的参数向量 $\boldsymbol{\vartheta} := \mathbf{c} \in \mathbf{R}^m$,即

$$y(t) = \boldsymbol{\varphi}^T(\mathbf{c}, t)\boldsymbol{\theta} + v(t), \quad (196)$$

$$y_1(t) = \boldsymbol{\psi}^T(\mathbf{b}, t)\boldsymbol{\vartheta} + v(t). \quad (197)$$

6.2 基于双线性参数模型分解的辅助模型随机梯度辨识方法

设 $\hat{\boldsymbol{\theta}}(t) := \begin{bmatrix} \hat{\mathbf{a}}_t \\ \hat{\mathbf{b}}_t \end{bmatrix}$ 和 $\hat{\boldsymbol{\vartheta}}(t) := \hat{c}_t$ 分别是参数向量 $\boldsymbol{\theta}$

和 $\boldsymbol{\vartheta}$ 在时刻 t 的估计.对于模型(196)和(197),使用负梯度搜索,极小化准则函数

$$J_{11}(\boldsymbol{\theta}) := \frac{1}{2}[y(t) - \boldsymbol{\varphi}^T(\mathbf{c}, t)\boldsymbol{\theta}]^2,$$

$$J_{12}(\boldsymbol{\vartheta}) := \frac{1}{2}[y_1(t) - \boldsymbol{\psi}^T(\mathbf{b}, t)\boldsymbol{\vartheta}]^2,$$

算法中涉及的未知变量用其估计代替或辅助模型的输出代替,可得辨识参数向量的基于双线性参数模型分解的辅助模型随机梯度算法(D-AM-SG算法):

$$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t-1) + \frac{\hat{\boldsymbol{\varphi}}(\hat{\mathbf{c}}_t, t)}{r_1(t)}[y(t) - \hat{\boldsymbol{\varphi}}^T(\hat{\mathbf{c}}_t, t)\hat{\boldsymbol{\theta}}(t-1)], \quad (198)$$

$$r_1(t) = r_1(t-1) + \|\hat{\boldsymbol{\varphi}}(\hat{\mathbf{c}}_t, t)\|^2, \quad r_1(0) = 1, \quad (199)$$

$$\hat{\boldsymbol{\varphi}}(\hat{\mathbf{c}}_t, t) = \begin{bmatrix} \hat{\boldsymbol{\varphi}}_a(t) \\ \mathbf{F}(t) \hat{\mathbf{c}}_t \end{bmatrix}, \quad (200)$$

$$\hat{\boldsymbol{\varphi}}_a(t) := [-x_a(t-1), -x_a(t-2), \dots, -x_a(t-n_a)]^T, \quad (201)$$

$$\hat{\boldsymbol{\vartheta}}(t) = \hat{\boldsymbol{\vartheta}}(t-1) + \frac{\boldsymbol{\psi}(\hat{\mathbf{b}}_{t-1}, t)}{r_2(t)} [y(t) - \hat{\boldsymbol{\varphi}}_a^T(t) \hat{\mathbf{a}}_{t-1} - \boldsymbol{\psi}^T(\hat{\mathbf{b}}_{t-1}, t) \hat{\boldsymbol{\vartheta}}(t-1)], \quad (202)$$

$$r_2(t) = r_2(t-1) + \|\boldsymbol{\psi}(\hat{\mathbf{b}}_{t-1}, t)\|^2, \quad r_2(0) = 1, \quad (203)$$

$$\boldsymbol{\psi}(\hat{\mathbf{b}}_{t-1}, t) = \mathbf{F}^T(t) \hat{\mathbf{b}}_{t-1}, \quad (204)$$

$$\mathbf{F}(t) = \begin{bmatrix} f_1(u(t-1)) & f_2(u(t-1)) & \dots & f_m(u(t-1)) \\ f_1(u(t-2)) & f_2(u(t-2)) & \dots & f_m(u(t-2)) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(u(t-n_b)) & f_2(u(t-n_b)) & \dots & f_m(u(t-n_b)) \end{bmatrix}, \quad (205)$$

$$x_a(t) = \hat{\boldsymbol{\varphi}}_a^T(t) \hat{\mathbf{a}}_t + \hat{\mathbf{b}}_t^T \mathbf{F}(t) \hat{\mathbf{c}}_t, \quad (206)$$

$$\hat{\boldsymbol{\theta}}(t) = [\hat{\mathbf{a}}_t^T, \hat{\mathbf{b}}_t^T]^T. \quad (207)$$

6.3 基于双线性参数模型分解的辅助模型递推最小二乘辨识方法

考虑 IN-OE 系统的双线性参数模型分解得到的两个子模型 (196) 和 (197), 即

$$y(t) = \boldsymbol{\varphi}^T(\mathbf{c}, t) \boldsymbol{\theta} + v(t), \quad (208)$$

$$y_1(t) = \boldsymbol{\psi}^T(\mathbf{b}, t) \boldsymbol{\vartheta} + v(t), \quad (209)$$

其中

$$\boldsymbol{\varphi}(\mathbf{c}, t) := \begin{bmatrix} \boldsymbol{\varphi}_a(t) \\ \mathbf{F}(t) \mathbf{c} \end{bmatrix} \in \mathbf{R}^{n_a+n_b},$$

$$\boldsymbol{\psi}(\mathbf{b}, t) := \mathbf{F}^T(t) \mathbf{b} \in \mathbf{R}^m,$$

$$y_1(t) := y(t) - \boldsymbol{\varphi}_a^T(t) \mathbf{a},$$

参数向量 $\mathbf{a}, \mathbf{b}, \mathbf{c}$, 信息向量 $\boldsymbol{\varphi}_a(t)$ 及信息矩阵 $\mathbf{F}(t)$ 的定义同上.

对于模型 (208) 和 (209), 极小化最小二乘准则函数

$$J_{13}(\boldsymbol{\theta}) := \frac{1}{2} \sum_{j=1}^l [y(j) - \boldsymbol{\varphi}^T(\mathbf{c}, j) \boldsymbol{\theta}]^2,$$

$$J_{14}(\boldsymbol{\vartheta}) := \frac{1}{2} \sum_{j=1}^l [y_1(j) - \boldsymbol{\psi}^T(\mathbf{b}, j) \boldsymbol{\vartheta}]^2,$$

参考递推最小二乘算法的推导, 极小化准则函数 $J_{13}(\boldsymbol{\theta})$ 和 $J_{14}(\boldsymbol{\vartheta})$, 算法中涉及的未知变量用其估计代替或辅助模型的输出代替, 可得辨识参数向量的基于双线性参数模型分解的辅助模型递推最小二乘算法 (D-AM-RLS 算法):

$$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t-1) + \mathbf{L}_1(t) [y(t) - \hat{\boldsymbol{\varphi}}^T(\hat{\mathbf{c}}_t, t) \hat{\boldsymbol{\theta}}(t-1)], \quad (210)$$

$$\mathbf{L}_1(t) = \mathbf{P}_1(t-1) \hat{\boldsymbol{\varphi}}(\hat{\mathbf{c}}_t, t) [1 +$$

$$\hat{\boldsymbol{\varphi}}^T(\hat{\mathbf{c}}_t, t) \mathbf{P}_1(t-1) \hat{\boldsymbol{\varphi}}(\hat{\mathbf{c}}_t, t)]^{-1}, \quad (211)$$

$$\mathbf{P}_1(t) = [\mathbf{I}_{n_a+n_b} - \mathbf{L}_1(t) \hat{\boldsymbol{\varphi}}^T(\hat{\mathbf{c}}_t, t)] \mathbf{P}_1(t-1), \quad (212)$$

$$\boldsymbol{\varphi}(\hat{\mathbf{c}}_t, t) = \begin{bmatrix} \hat{\boldsymbol{\varphi}}_a(t) \\ \mathbf{F}(t) \hat{\mathbf{c}}_t \end{bmatrix}, \quad (213)$$

$$\hat{\boldsymbol{\varphi}}_a(t) = [-x_a(t-1), -x_a(t-2), \dots, -x_a(t-n_a)]^T, \quad (214)$$

$$\hat{\boldsymbol{\vartheta}}(t) = \hat{\boldsymbol{\vartheta}}(t-1) + \mathbf{L}_2(t) [y(t) - \hat{\boldsymbol{\varphi}}_a^T(t) \hat{\mathbf{a}}_{t-1} - \boldsymbol{\psi}^T(\hat{\mathbf{b}}_{t-1}, t) \hat{\boldsymbol{\vartheta}}(t-1)], \quad (215)$$

$$\mathbf{L}_2(t) = \mathbf{P}_2(t-1) \boldsymbol{\psi}(\hat{\mathbf{b}}_{t-1}, t) [1 + \boldsymbol{\psi}^T(\hat{\mathbf{b}}_{t-1}, t) \mathbf{P}_2(t-1) \boldsymbol{\psi}(\hat{\mathbf{b}}_{t-1}, t)]^{-1}, \quad (216)$$

$$\mathbf{P}_2(t) = [\mathbf{I}_m - \mathbf{L}_2(t) \boldsymbol{\psi}^T(\hat{\mathbf{b}}_{t-1}, t)] \mathbf{P}_2(t-1), \quad (217)$$

$$\boldsymbol{\psi}(\hat{\mathbf{b}}_{t-1}, t) = \mathbf{F}^T(t) \hat{\mathbf{b}}_{t-1}, \quad (218)$$

$$\mathbf{F}(t) = \begin{bmatrix} f_1(u(t-1)) & f_2(u(t-1)) & \dots & f_m(u(t-1)) \\ f_1(u(t-2)) & f_2(u(t-2)) & \dots & f_m(u(t-2)) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(u(t-n_b)) & f_2(u(t-n_b)) & \dots & f_m(u(t-n_b)) \end{bmatrix}, \quad (219)$$

$$x_a(t) = \hat{\boldsymbol{\varphi}}_a^T(t) \hat{\mathbf{a}}_t + \hat{\mathbf{b}}_t^T \mathbf{F}(t) \hat{\mathbf{c}}_t, \quad (220)$$

$$\hat{\boldsymbol{\theta}}(t) = [\hat{\mathbf{a}}_t^T, \hat{\mathbf{b}}_t^T]^T. \quad (221)$$

D-AM-RLS 算法 (210) — (221) 的初值可设置为

$$\hat{\boldsymbol{\theta}}(0) = \begin{bmatrix} \hat{\mathbf{a}}_0 \\ \hat{\mathbf{b}}_0 \end{bmatrix} = \mathbf{1}_{n_a+n_b} / p_0, \quad \hat{\boldsymbol{\vartheta}}(0) = \hat{\mathbf{c}}_0 \text{ 是随机向量, 满足}$$

$\|\hat{\mathbf{c}}_0\| = 1$, 且第一个元为正, $\mathbf{P}_1(0) = p_0 \mathbf{I}_{n_a+n_b}$, $\mathbf{P}_2(0) = p_0 \mathbf{I}_m$, $x_a(t-i) = 1/p_0$, $i = 0, 1, 2, \dots, n_a$, $p_0 = 10^6$. 采用假设 (A4) 规范化 $\hat{\mathbf{c}}_t$:

$$\hat{\mathbf{c}}_t := \text{sgn}[\hat{\boldsymbol{\vartheta}}_1(t)] \frac{\hat{\boldsymbol{\vartheta}}(t)}{\|\hat{\boldsymbol{\vartheta}}(t)\|}, \quad \hat{\boldsymbol{\vartheta}}(t) := \hat{\mathbf{c}}_t.$$

7 算法的计算量比较

辨识算法的计算量可以用乘法次数与加法次数衡量. 除法作为乘法对待, 减法作为加法对待. 一次乘法运算是一个 flop (floating point operation), 一次加法运算也是一个 flop (floating point operation). 乘法运算与加法运算之和就是一个算法的计算量. 如果同为递推算法或迭代算法, 可以比较每一步的计算量.

表 6 给出了 O-AM-RLS 算法、O-AM-HLS 算法、KT-AM-RLS 算法、KT-AM-2S-RLS 算法、KT-AM-3S-RLS 算法、D-AM-RLS 算法每步的计算量 (总 flop 数). 为了更直观地看出算法计算量的大小, 表 6 中给出了 $n_a = 10, n_b = 10, m = 10$ 时的 flop 数.

表 6 算法每步计算量的比较

Table 6 The computational efficiency of the algorithms

算法	总 flop 数	flop 数($n_a = 10, n_b = 10, m = 10$)
O-AM-RLS	$N_1 = 4(n_a + mn_b)^2 + 8(n_a + mn_b) - 1$	49 279
O-AM-HLS	$N_2 = 4m^2n_b + 8mn_b + 4n_a^2 + 8n_a - 1$	5 279
KT-AM-RLS	$N_3 = 4(n_a + n_b + m)^2 + 8(n_a + n_b + m) + 2m - 2$	3 858
KT-AM-2S-RLS	$N_4 = 4(n_a + n_b)^2 + 4m^2 + 8(n_a + n_b + m) - 1$	2 239
KT-AM-3S-RLS	$N_5 = 4(n_a^2 + n_b^2 + m^2) + 8(n_a + n_b + m) - 1$	1 439
D-AM-RLS	$N_6 = 4(n_a + n_b)^2 + 4m^2 + 6mn_b + 6(n_a + n_b) + 4n_a + 4m - 1$	2 779

从表 6 可以看出一些辨识算法计算量的差异。借用递阶辨识原理的分解辨识算法可以减小算法的计算量,提高辨识效率,尤其是 AM-HLS 算法比 AM-RLS 算法计算效率高很多。

8 结语

针对具有已知基的输入非线性输出误差系统,分别基于过参数化辨识模型、关键项分离辨识模型、双线性参数模型分解辨识模型,研究了一些辅助模型递推辨识方法,这些辨识算法的收敛性都是有待研究的课题。

尽管这些方法是针对白噪声干扰输入非线性方程误差系统提出的,但是可以推广到下列有色噪声干扰的输入非线性输出误差系统。

1) 输入非线性输出误差滑动平均模型 (IN-OEMA 模型):

$$y(t) = \frac{B(z)}{A(z)}\bar{u}(t) + D(z)v(t),$$

$$\bar{u}(t) = f(u(t)) = \gamma_1 f_1(u(t)) + \gamma_2 f_2(u(t)) + \dots + \gamma_m f_m(u(t)).$$

2) 输入非线性输出误差自回归模型 (IN-OEAR 模型):

$$y(t) = \frac{B(z)}{A(z)}\bar{u}(t) + \frac{1}{C(z)}v(t),$$

$$\bar{u}(t) = f(u(t)) = \gamma_1 f_1(u(t)) + \gamma_2 f_2(u(t)) + \dots + \gamma_m f_m(u(t)).$$

3) 输入非线性输出误差自回归滑动平均模型 (IN-OEARMA 模型), 即输入非线性 Box-Jenkins 模型 (IN-BJ 模型):

$$y(t) = \frac{B(z)}{A(z)}\bar{u}(t) + \frac{D(z)}{C(z)}v(t),$$

$$\bar{u}(t) = f(u(t)) = \gamma_1 f_1(u(t)) + \gamma_2 f_2(u(t)) + \dots + \gamma_m f_m(u(t)).$$

本文的方法还可以推广到多输入非线性单输出输出误差类系统、多输入非线性多输出输出误差类系统、非线性多变量输出误差类系统,以及反馈非线性系统。

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: 多新息辨识理论与方法 [M]. 北京: 科学出版社, 2016
DING Feng. System identification: Multi-Innovation identification theory and methods [M]. Beijing: Science Press, 2016
- [4] 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
- [5] 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
- [6] 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 accuracy and basic problems [J]. Journal of Nanjing University of Information Science and Technology (Natural Science Edition), 2011, 3(3): 193-226
- [7] 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
- [8] 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
- [9] 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
- [10] 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
- [11] 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
- [12] 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
- [13] 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
- [14] 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
- [15] 丁锋. 方程误差系统的多新息辨识方法[J]. 南京信息工程大学学报(自然科学版), 2015, 7(5): 385-407
DING Feng. Multi-innovation identification methods for equation-error systems [J]. Journal of Nanjing University of Information Science and Technology (Natural Science Edition), 2015, 7(5): 385-407
- [16] 丁锋. 输出误差系统的多新息辨识方法[J]. 南京信息工程大学学报(自然科学版), 2015, 7(6): 481-503
DING Feng. Multi-innovation identification methods for output-error systems [J]. Journal of Nanjing University of Information Science and Technology (Natural Science Edition), 2015, 7(6): 481-503
- [17] Ding F, Chen T. Identification of Hammerstein nonlinear ARMAX systems [J]. Automatica, 2005, 41(9): 1479-1489
- [18] Ding F, Shi Y, Chen T. Gradient-based identification methods for Hammerstein nonlinear ARMAX models [J]. Nonlinear Dynamics, 2006, 45(1/2): 31-43
- [19] Chen H B, Xiao Y S, Ding F. Hierarchical gradient parameter estimation algorithm for Hammerstein nonlinear systems using the key term separation principle [J]. Applied Mathematics and Computation, 2014, 247: 1202-1210
- [20] Chen H B, Ding F. Hierarchical least squares identification for Hammerstein nonlinear controlled autoregressive systems [J]. Circuit, System and Signal Processing, 2015, 34(1): 61-75
- [21] Chen H B, Ding F, Xiao Y S. Decomposition based parameter estimation algorithm for input nonlinear systems using the key term separation technique [J]. Nonlinear Dynamics, 2015, 79(3): 2027-2035
- [22] Mao Y W, Ding F. Multi-innovation stochastic gradient identification for Hammerstein controlled autoregressive systems based on the filtering technique [J]. Nonlinear Dynamics, 2015, 79(3): 1745-1755
- [23] Ding F, Liu X P, Liu G. Identification methods for Hammerstein nonlinear systems [J]. Digital Signal Processing, 2011, 21(2): 215-238
- [24] Wang D Q, Ding F. Least squares based and gradient based iterative identification for Wiener nonlinear systems [J]. Signal Processing, 2011, 91(5): 1182-1189
- [25] Ding F. Hierarchical multi-innovation stochastic gradient algorithm for Hammerstein nonlinear system modeling [J]. Applied Mathematical Modelling, 2013, 37(4): 1694-1704
- [26] 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
- [27] 丁锋, 毛亚文. 输入非线性方程误差自回归系统的多新息辨识方法[J]. 南京信息工程大学学报(自然科学版), 2015, 7(1): 1-23
DING Feng, MAO Yawen. Multi-innovation identification methods for input nonlinear equation-error autoregressive systems [J]. Journal of Nanjing University of Information Science and Technology (Natural Science Edition), 2015, 7(1): 1-23
- [28] 丁锋, 陈慧波. 输入非线性方程误差系统的多新息辨识方法[J]. 南京信息工程大学学报(自然科学版), 2015, 7(2): 97-124
DING Feng, CHEN Huibo. Multi-innovation identification methods for input nonlinear equation-error systems. Journal of Nanjing University of Information Science and Technology (Natural Science Edition), 2015, 7(2): 97-124
- [29] Ding F, Shi Y, Chen T. Auxiliary model based least-squares identification methods for Hammerstein output-error systems [J]. Systems & Control Letters, 2007, 56(5): 373-380
- [30] 范伟, 丁锋. Hammerstein 非线性系统参数估计分离的三种方法 [J]. 控制科学与工程, 2008, 8(6): 1586-1589
FAN Wei, DING Feng. Three methods of separating pa-

- rameters for Hammerstein nonlinear systems [J]. Science Technology and Engineering, 2008, 8(6): 1586-1589
- [31] Vörös J. Iterative algorithm for parameter identification of Hammerstein systems with two-segment nonlinearities [J]. IEEE Transactions on Automatic Control, 1999, 44(11): 2145-2149
- [32] Vörös J. Modeling and parameter identification of systems with multi-segment piecewise-linear characteristics [J]. IEEE Transactions on Automatic Control, 2002, 47(1): 184-188
- [33] Vörös J. Identification of Hammerstein systems with time-varying piecewise-linear characteristics [J]. IEEE Transactions on Circuits and Systems. II: Express Briefs, 2005, 52(12): 865-869

Auxiliary model identification methods. Part B: Input nonlinear output-error systems

DING Feng^{1,2,3} CHEN Huibo¹

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 For input nonlinear output-error systems with known bases, this paper presents the over-parameterization model based auxiliary model (AM) recursive identification methods, the over-parameterization model based AM hierarchical identification methods, the key term separation based AM recursive identification methods, the key term separation based AM two-stage recursive identification methods, the key term separation based AM three-stage recursive identification methods, the bilinear-in-parameter model decomposition based AM stochastic gradient identification methods and the bilinear-in-parameter model decomposition based AM recursive least squares identification methods. Finally, the computational efficiency and the computational steps of several typical identification algorithms are discussed. The convergence of the proposed algorithms needs further study.

Key words parameter estimation; recursive identification; gradient search; least squares; over-parameterization model; key term separation; model decomposition; auxiliary model identification ideal; hierarchical identification principle; input nonlinear system