



# 输出误差系统的多新息辨识方法

## 摘要

研究了输出误差(OE)系统和输出误差自回归滑动平均(OEARMA)系统(即Box-Jenkins系统)的辅助模型随机梯度算法、辅助模型多新息随机梯度算法、辅助模型递推最小二乘算法、辅助模型多新息最小二乘算法、输出误差系统的修正辅助模型随机梯度算法、遗忘因子辅助模型随机梯度算法、变递推间隔辅助模型随机梯度算法、变递推间隔辅助模型多新息随机梯度算法、变递推间隔辅助模型递推最小二乘算法等,以及输出误差自回归(OEAR)系统的基于滤波的辅助模型多新息广义随机梯度算法和基于滤波的辅助模型多新息广义最小二乘算法。

## 关键词

参数估计;递推辨识;梯度搜索;最小二乘;滤波;分解;辅助模型辨识思想;多新息辨识理论;递阶辨识原理;输出误差系统;线性系统

中图分类号 TP273

文献标志码 A

收稿日期 2015-09-20

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

在这个辨识新思想、新理论、新原理、新概念相继问世的年代,为提高辨识精度、减小辨识算法的计算量,为解决含有不可测状态的状态空间系统、双线性参数系统乃至非线性系统辨识问题等,笔者等提炼和发展了滤波辨识技术、分解辨识技术、梯度迭代技术、最小二乘迭代技术、牛顿迭代技术,以及辅助模型辨识思想、多新息辨识理论、递阶辨识原理、耦合辨识概念等。这些辨识技术与辨识思想、辨识理论、辨识原理等相结合,便产生层出不穷的辨识方法,如基于数据滤波的最小二乘辨识方法、基于分解的辅助模型多新息辨识方法、基于分解的牛顿迭代辨识方法、基于滤波的耦合最小二乘迭代辨识方法、基于滤波的递阶最小二乘辨识方法、基于滤波的递阶梯度迭代辨识方法、基于分解的牛顿递推辨识方法和牛顿迭代辨识方法等。这些技术和理论原理的结合为系统辨识新的研究增添了无比绚丽的色彩。

最近出版的《系统辨识新论》<sup>[1]</sup>介绍了一些新颖的辨识方法,如辅助模型辨识方法、迭代辨识方法、多新息辨识方法、递阶辨识方法、耦合辨识方法等,一些典型递推辨识方法的收敛性分析汇聚在《系统辨识——辨识方法性能分析》<sup>[2]</sup>一书中,以及在《南京信息工程大学学报》上发表的关于“系统辨识”的连载论文<sup>[3-24]</sup>及其衍生出的辅助模型递推辨识方法、辅助模型迭代辨识方法、基于滤波的辅助模型递推辨识方法、基于滤波的辅助模型迭代辨识方法、递阶最小二乘辨识方法和递阶最小二乘迭代辨识方法、递阶随机梯度辨识方法和递阶梯度迭代辨识方法、基于滤波耦合辨识方法等的收敛性研究,更是给系统辨识研究的主旋律添上了玫瑰般色彩。

本文基于辅助模型辨识思想,研究输出误差系统和输出误差自回归滑动平均系统(即Box-Jenkins系统)的辅助模型多新息随机梯度算法和辅助模型多新息最小二乘算法;基于数据滤波技术,研究输出误差自回归(OEAR)系统的滤波多新息广义随机梯度算法和滤波多新息广义最小二乘算法等。这些方法只是国内外期刊和《南京信息工程大学学报》上连载多新息辨识论文中所讨论方法的冰山一角<sup>[8,14-16,18-21,23-59]</sup>为此,笔者将连载论文中2012年第1期<sup>[8]</sup>、2015年第1—2,5—6期<sup>[20-21,24]</sup>作为蓝本,并进行补充,加上“多新息辨识方法的性能分析”一章,便形成了《系统辨识——多新息辨识理论与方法》一书,即将由科学出版社出版。这些著作和连载论文中一些辨识方法的

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

收敛性分析为读者提供了广阔的研究空间.

## 1 输出误差系统

这里利用辅助模型辨识思想和多新息辨识理论,研究输出误差系统的辅助模型随机梯度(AM-SG)算法、辅助模型多新息随机梯度(AM-MISG)算法、变递推间隔辅助模型随机梯度(V-AM-SG)算法、变递推间隔辅助模型多新息随机梯度(V-AM-MISG)算法、辅助模型递推最小二乘算法、辅助模型多新息最小二乘算法等.其中 V-AM-MISG 算法以 Regular Paper 论文形式发表在国际期刊《Automatica》2011 年第 8 期<sup>[47]</sup>上,该文入选“2011 年中国百篇最具影响国际学术论文”.

关于等递推间隔辅助模型随机梯度算法(E-AM-SG 算法),即“双率采样数据系统的辅助模型随机梯度算法及其真实输出估计”论文发表在《IEEE Transactions on Automatic Control》2005 年第 9 期<sup>[60]</sup>上.关于等递推间隔辅助模型递推最小二乘算法(E-AM-RLS 算法),即“双率采样数据系统的辅助模型递推最小二乘算法及其真实输出估计”论文发表在《Automatica》2004 年第 10 期<sup>[61]</sup>上;双率系统的辅助模型辨识方法参见文献[62],变递推间隔辅助模型递推最小二乘算法参见文献[63].

### 1.1 系统描述与辨识模型

考虑下列输出误差模型(Output-Error model, OE 模型)描述的动态随机系统(参见图 1 上部分):

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

其中  $\{u(t)\}$  和  $\{y(t)\}$  分别是系统的输入和输出序列,  $\{v(t)\}$  是零均值方差为  $\sigma^2$  的随机白噪声序列,

$G(z) := \frac{B(z)}{A(z)}$  是系统模型的传递函数,  $z^{-1}$  为单位后移算子 [ $z^{-1}y(t) = y(t-1)$ ,  $zy(t) = y(t+1)$ ],  $A(z)$  和  $B(z)$  是单位后移算子  $z^{-1}$  的多项式:

$$A(z) := 1 + a_1z^{-1} + a_2z^{-2} + \dots + a_{n_a}z^{-n_a},$$

$$B(z) := b_1z^{-1} + b_2z^{-2} + \dots + b_{n_b}z^{-n_b}.$$

设阶次  $n_a$  和  $n_b$  已知,记  $n := n_a + n_b$ ,且  $t \leq 0$  时,  $y(t) = 0, u(t) = 0, v(t) = 0$ .图 1 中,  $x(t)$  是系统的真实输出,即不可测无噪输出,  $y(t)$  是  $x(t)$  的含噪量测.

定义系统真实输出

$$x(t) := \frac{B(z)}{A(z)}u(t) \in \mathbf{R}, \quad (2)$$

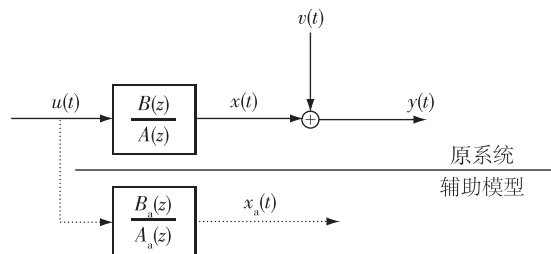


图 1 带辅助模型的输出误差系统

Fig. 1 The output-error system with the auxiliary model

定义参数向量  $\theta$  和信息向量  $\varphi(t)$  如下:

$$\begin{aligned} \theta &:= [a_1, a_2, \dots, a_{n_a}, b_1, b_2, \dots, b_{n_b}]^T \in \mathbf{R}^n, \\ \varphi(t) &:= [-x(t-1), -x(t-2), \dots, -x(t-n_a), \\ &\quad u(t-1), u(t-2), \dots, u(t-n_b)]^T \in \mathbf{R}^n. \end{aligned}$$

由式(2)可得

$$x(t) = [1 - A(z)]x(t) + B(z)u(t) = \varphi^T(t)\theta. \quad (3)$$

将式(2)代入式(1),使用式(3),可得输出误差系统的辨识模型:

$$y(t) = x(t) + v(t) = \varphi^T(t)\theta + v(t). \quad (4)$$

### 1.2 辅助模型随机梯度算法

#### 1.2.1 辅助模型随机梯度算法

设  $\hat{\theta}(t)$  是  $\theta$  在时刻  $t$  的估计,  $1/r(t)$  是收敛因子或步长.矩阵  $X$  的范数定义为  $\|X\|^2 := \text{tr}[XX^T]$ .

根据辨识模型(4),引入二次准则函数

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

使用负梯度搜索,极小化  $J_1(\theta)$ ,得到下列递推关系:

$$\begin{aligned} \hat{\theta}(t) &= \hat{\theta}(t-1) - \frac{1}{r(t)} \text{grad}[J_1(\hat{\theta}(t-1))] = \\ &= \hat{\theta}(t-1) + \frac{\varphi(t)}{r(t)} [y(t) - \varphi^T(t)\hat{\theta}(t-1)], \quad (5) \end{aligned}$$

$$r(t) = r(t-1) + \|\varphi(t)\|^2, \quad r(0) = 1. \quad (6)$$

根据信息向量  $\varphi(t)$  的定义可知,  $\varphi(t)$  是由系统可测输入  $u(t-i)$  和未知无噪输出  $x(t-i)$  构成的,这使得关系式(5)——(6)无法递推实现.解决的方法是利用辅助模型辨识思想,利用系统的可测信息(包括计算得到的信息)建立一个辅助模型,用辅助模型的输出  $x_a(t-i)$  代替信息向量  $\varphi(t)$  中的未知无噪输出项  $x(t-i)$ ,从而实现算法参数估计的递推计算,即基于辅助模型辨识思想<sup>[1]</sup>,构造一个辅助模型  $P_a(z)$ ,

如图 1 中下半部分,其中  $P_a(z) := \frac{B_a(z)}{A_a(z)}$  为辅助模型的传递函数,这里假定  $A_a(z)$  和  $B_a(z)$  是与  $A(z)$  和

$B(z)$ 结构和阶次相同的多项式,  $x_a(t)$ 为辅助模型的输出. 当然, 辅助模型  $P_a(z)$ 也可以取与  $G(z) = \frac{B(z)}{A(z)}$ 对应的有限脉冲响应模型(Finite Impulse Response model, FIR 模型)<sup>[62]</sup>.

用辅助模型的输出  $x_a(t-i)$ 和系统输入  $u(t-i)$ 定义  $\varphi(t)$ 的估计

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

由图1可知  $x_a(t) = \frac{B_a(z)}{A_a(z)}u(t)$ , 它与式(2)具有相同结构. 参照式(2)写作式(3)的形式, 把它写成向量形式为  $x_a(t) = \varphi_a^T(t)\theta_a$ , 其中  $\varphi_a(t)$ 和  $\theta_a$ 分别是辅助模型的信息向量和参数向量. 把  $\hat{\varphi}(t)$ 作为辅助模型的信息向量  $\varphi_a(t)$ , 把  $\hat{\theta}(t)$ 作为辅助模型的参数向量  $\theta_a$ , 得到  $x_a(t) = \hat{\varphi}^T(t)\hat{\theta}(t)$ . 式(5)–(6)中未知的  $\varphi(t)$ 用其估计  $\hat{\varphi}(t)$ 代替, 可以得到估计输出误差系统参数向量  $\theta$ 的辅助模型随机梯度算法(Auxiliary Model based Stochastic Gradient algorithm, AM-SG 算法)<sup>[1]</sup>:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\hat{\varphi}(t)}{r(t)}e(t), \quad (7)$$

$$e(t) = y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1), \quad (8)$$

$$r(t) = r(t-1) + \|\hat{\varphi}(t)\|^2, \quad (9)$$

$$\hat{\varphi}(t) = [-x_a(t-1), \dots, -x_a(t-n_a), u(t-1), \dots, u(t-n_b)]^T, \quad (10)$$

$$x_a(t) = \hat{\varphi}^T(t)\hat{\theta}(t), \quad (11)$$

$$\hat{\theta}(t) = [\hat{\alpha}_1(t), \hat{\alpha}_2(t), \dots, \hat{\alpha}_{n_a}(t), \hat{\beta}_1(t), \hat{\beta}_2(t), \dots, \hat{\beta}_{n_b}(t)]^T. \quad (12)$$

AM-SG 算法(7)–(12)随  $t$ 增加, 计算参数估计向量  $\hat{\theta}(t)$ 的步骤如下.

1) 初始化: 令  $t=1$ . 置初值  $\hat{\theta}(0) = \mathbf{1}_n/p_0, r(0) = 1, x_a(t-i) = 1/p_0, i=1, 2, \dots, n_a, p_0 = 10^6$ .

2) 采集输入输出数据  $u(t)$ 和  $y(t)$ , 由式(10)构造信息向量  $\hat{\varphi}(t)$ .

3) 由式(8)计算新息  $e(t)$ , 由式(9)计算  $r(t)$ .

4) 根据式(7)刷新参数估计向量  $\hat{\theta}(t)$ .

5) 由式(11)计算辅助模型的输出  $x_a(t)$ .

6)  $t$ 增1, 转到第2)步.

### 1.2.2 遗忘因子辅助模型随机梯度算法

为提高 AM-SG 算法的暂态性能和参数估计精度, 在式(9)中引入遗忘因子(Forgetting Factor, FF)  $\lambda$ , 就得到遗忘因子辅助模型随机梯度算法(Forgetting Factor AM-SG algorithm, FF-AM-SG 算法)

或辅助模型遗忘因子随机梯度算法(AM-FFSG 算法), 简称辅助模型遗忘梯度算法(AM-FG 算法):

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\hat{\varphi}(t)}{r(t)}e(t), \quad \hat{\theta}(0) = \mathbf{1}_n/p_0, \quad (13)$$

$$e(t) = y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1), \quad (14)$$

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

$$\hat{\varphi}(t) = [-x_a(t-1), \dots, -x_a(t-n_a), u(t-1), \dots, u(t-n_b)]^T, \quad (16)$$

$$x_a(t) = \hat{\varphi}^T(t)\hat{\theta}(t), \quad x_a(i) = 1/p_0, \quad i \leq 0, \quad (17)$$

$$\hat{\theta}(t) = [\hat{\alpha}_1(t), \hat{\alpha}_2(t), \dots, \hat{\alpha}_{n_a}(t), \hat{\beta}_1(t), \hat{\beta}_2(t), \dots, \hat{\beta}_{n_b}(t)]^T. \quad (18)$$

当遗忘因子  $\lambda = 1$ 时, AM-FG 算法退化为 AM-SG 算法.

### 1.2.3 修正辅助模型随机梯度算法

为提高 AM-SG 算法的暂态性能和稳态性能, 以及参数估计精度, 在式(7)中引入收敛指数  $\varepsilon$ , 就得到修正辅助模型随机梯度算法(Modified AM-SG algorithm, M-AM-SG 算法)<sup>[1]</sup>:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\hat{\varphi}(t)}{r^\varepsilon(t)}e(t), \quad \frac{1}{2} < \varepsilon \leq 1,$$

$$\hat{\theta}(0) = \mathbf{1}_n/p_0, \quad (19)$$

$$e(t) = y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1), \quad (20)$$

$$r(t) = r(t-1) + \|\hat{\varphi}(t)\|^2, \quad r(0) = 1, \quad (21)$$

$$\hat{\varphi}(t) = [-x_a(t-1), \dots, -x_a(t-n_a), u(t-1), \dots, u(t-n_b)]^T, \quad (22)$$

$$x_a(t) = \hat{\varphi}^T(t)\hat{\theta}(t), \quad x_a(i) = 1/p_0, \quad i \leq 0, \quad (23)$$

$$\hat{\theta}(t) = [\hat{\alpha}_1(t), \hat{\alpha}_2(t), \dots, \hat{\alpha}_{n_a}(t), \hat{\beta}_1(t), \hat{\beta}_2(t), \dots, \hat{\beta}_{n_b}(t)]^T. \quad (24)$$

当收敛指数  $\varepsilon = 1$ 时, M-AM-SG 算法退化为 AM-SG 算法.

**注1** 文献[60]证明了双率系统辅助模型随机梯度算法, 即等递推间隔辅助模型随机梯度算法(E-AM-SG 算法)的收敛性. 文献[64]研究了多变量输出误差系统辅助模型多变量随机梯度算法(AM-MSG 算法)的收敛性. 这里的 AM-SG 算法(7)–(12)是多变量的特殊情形, 其收敛性证明可以参考多变量系统的对应方法. 读者可以仿照其来证明 M-AM-SG 算法的收敛性. FF-AM-SG 算法参数估计的有界收敛性仍然是辨识领域的研究难题.

### 1.2.4 变递推间隔辅助模型随机梯度算法

假设输入数据  $\{u(t): t=0, 1, 2, \dots\}$ 都可得到, 输出数据存在损失, 定义可得到数据点的整数序列:

$$0 = t_0 < t_1 < t_2 < \dots < t_{s-1} < t_s < \dots, \quad t_s^* := t_s - t_{s-1} \geq 1,$$

即所有  $y(t_s), s=1, 2, \dots$ 都可量测得到. 对于这种情形,

可得到的观测数据为  $\{u(s), y(t_s); s=0, 1, 2, \dots\}$ .

用  $t_s$  代替式(4)中  $t$ , 得到变递推间隔方法的输出误差系统的辨识模型:

$$y(t_s) = \varphi^T(t_s)\theta + v(t_s). \quad (25)$$

根据辨识模型(25), 使用负梯度搜索, 极小化二次准则函数

$$J_2(\theta) := \frac{1}{2} [y(t_s) - \varphi^T(t_s)\theta]^2,$$

未知  $\varphi(t_s)$  用其估计  $\hat{\varphi}(t_s)$  代替, 可以得到估计输出误差系统参数向量  $\theta$  的变递推间隔辅助模型随机梯度算法 (interval-Varying AM-SG algorithm, V-AM-SG 算法) [1]:

$$\hat{\theta}(t_s) = \hat{\theta}(t_{s-1}) + \frac{\hat{\varphi}(t_s)}{r(t_s)} e(t_s), \quad \hat{\theta}(t_0) = \mathbf{1}_n/p_0, \quad (26)$$

$$\hat{\theta}(t) = \hat{\theta}(t_s), \quad t \in T_s = \{t_s, t_s+1, \dots, t_{s+1}-1\}, \quad (27)$$

$$e(t_s) = y(t_s) - \hat{\varphi}^T(t_s)\hat{\theta}(t_{s-1}), \quad (28)$$

$$r(t_s) = r(t_{s-1}) + \|\hat{\varphi}(t_s)\|^2, \quad r(t_0) = 1, \quad (29)$$

$$\hat{\varphi}(t_s) = [-x_a(t_s-1), \dots, -x_a(t_s-n_a), \\ u(t_s-1), \dots, u(t_s-n_b)]^T, \quad (30)$$

$$x_a(t_s+i) = \hat{\varphi}^T(t_s+i)\hat{\theta}(t_s), \quad i=0, 1, \dots, t_{s+1}^*-1, \quad (31)$$

$$\hat{\theta}(t_s) = [\hat{a}_1(t_s), \hat{a}_2(t_s), \dots, \hat{a}_{n_a}(t_s), \hat{b}_1(t_s), \\ \hat{b}_2(t_s), \dots, \hat{b}_{n_b}(t_s)]^T. \quad (32)$$

当  $t_s^* \equiv d$  (常数) 时, V-AM-SG 算法退化为等递推间隔辅助模型随机梯度算法 (E-AM-SG 算法).

### 1.3 辅助模型多新息随机梯度算法

辅助模型随机梯度辨识算法的收敛速度慢 (参见《系统辨识新论》[1] 中的仿真例子). 另一种改进参数估计精度的方法是借助于多新息辨识理论, 通过扩展新息维数, 推导出基于辅助模型的多新息随机梯度算法.

#### 1.3.1 辅助模型多新息随机梯度算法

设正整数  $p$  表示新息长度. 将 AM-SG 算法(7)~(12)中标量新息  $e(t) = y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1) \in \mathbf{R}$  扩展为新息向量

$$\mathbf{E}(p, t) := \begin{bmatrix} e(t) \\ e(t-1) \\ \vdots \\ e(t-p+1) \end{bmatrix} = \begin{bmatrix} y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1) \\ y(t-1) - \hat{\varphi}^T(t-1)\hat{\theta}(t-2) \\ \vdots \\ y(t-p+1) - \hat{\varphi}^T(t-p+1)\hat{\theta}(t-p) \end{bmatrix} \in \mathbf{R}^p.$$

通常可认为时刻  $(t-1)$  的估计  $\hat{\theta}(t-1)$  比时刻  $(t-i)$  ( $i \geq 2$ ) 的估计  $\hat{\theta}(t-i)$  更接近真值  $\theta$ . 因此, 将  $\mathbf{E}(p, t)$  表达式中  $\hat{\theta}(t-i)$  ( $i \geq 2$ ) 换为  $\hat{\theta}(t-1)$ , 则多新息向量可合理修改为

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

定义堆积输出向量  $\mathbf{Y}(p, t)$  和堆积信息矩阵  $\hat{\Phi}(p, t)$  如下:

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

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

多新息向量可以表达为

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

注意到  $\mathbf{Y}(1, t) = y(t)$ ,  $\mathbf{E}(1, t) = e(t)$ ,  $\hat{\Phi}(1, t) = \hat{\varphi}(t)$ , 式(7)可以等价表达为

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\hat{\Phi}(1, t)}{r(t)} \mathbf{E}(1, t).$$

上式中新息长度  $p=1$ . 这是新息长度为 1 的“多新息”随机梯度算法. 把上式中  $\hat{\Phi}(1, t)$  和  $\mathbf{E}(1, t)$  里的“1”换为  $p$ , 可以得到新息长度为  $p$  的, 估计输出误差系统参数向量  $\theta$  的辅助模型多新息随机梯度算法 (Auxiliary Model based Multi-Innovation Stochastic Gradient algorithm, AM-MISG 算法):

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\hat{\Phi}(p, t)}{r(t)} \mathbf{E}(p, t), \quad (33)$$

$$\mathbf{E}(p, t) = \mathbf{Y}(p, t) - \hat{\Phi}^T(p, t)\hat{\theta}(t-1), \quad (34)$$

$$r(t) = r(t-1) + \|\hat{\varphi}(t)\|^2, \quad (35)$$

$$\mathbf{Y}(p, t) = [y(t), y(t-1), \dots, y(t-p+1)]^T, \quad (36)$$

$$\hat{\Phi}(p, t) = [\hat{\varphi}(t), \hat{\varphi}(t-1), \dots, \hat{\varphi}(t-p+1)], \quad (37)$$

$$\hat{\varphi}(t) = [-x_a(t-1), \dots, -x_a(t-n_a), \\ u(t-1), \dots, u(t-n_b)]^T, \quad (38)$$

$$x_a(t) = \hat{\varphi}^T(t)\hat{\theta}(t), \quad (39)$$

$$\hat{\theta}(t) = [\hat{a}_1(t), \hat{a}_2(t), \dots, \hat{a}_{n_a}(t), \hat{b}_1(t), \hat{b}_2(t), \dots, \hat{b}_{n_b}(t)]^T. \quad (40)$$

当新息长度  $p=1$  时, AM-MISG 算法退化为 AM-SG 算法(7)~(12).

AM-MISG 算法(33)~(40)随  $t$  增加, 计算参数估计向量  $\hat{\theta}(t)$  的步骤如下:

1) 初始化: 令  $t=1$ , 给定新息长度  $p$ . 置初值  $\hat{\theta}(0) = \mathbf{1}_n/p_0$ ,  $r(0) = 1$ ,  $x_a(t-i) = 1/p_0$ ,  $i=1, 2, \dots, n_a$ ,  $p_0 = 10^6$ .

2) 采集输入输出数据  $u(t)$  和  $y(t)$ , 由式(38)



构造信息向量  $\hat{\varphi}(t)$ , 式(36)–(37)构造堆积输出向量  $Y(p, t)$  和堆积信息矩阵  $\hat{\Phi}(p, t)$ .

3) 由式(34)计算新息向量  $E(p, t)$ , 由式(35)计算  $r(t)$ .

4) 根据式(33)刷新参数估计向量  $\hat{\theta}(t)$ .

5) 由式(39)计算辅助模型的输出  $x_a(t)$ .

6)  $t$  增 1, 转到第 2) 步.

### 1.3.2 修正遗忘因子 AM-MISG 算法

在 AM-MISG 算法中引入了遗忘因子  $\lambda$  和收敛指数  $\varepsilon$ , 就得到修正遗忘因子 AM-MISG 算法:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\hat{\Phi}(p, t) E(p, t)}{r^\varepsilon(t)}, \quad \frac{1}{2} < \varepsilon \leq 1, \quad (41)$$

$$\hat{\theta}(0) = \mathbf{1}_n / p_0, \quad (41)$$

$$E(p, t) = Y(p, t) - \hat{\Phi}^T(p, t) \hat{\theta}(t-1), \quad (42)$$

$$r(t) = \lambda r(t-1) + \|\hat{\varphi}(t)\|^2, \quad 0 \leq \lambda \leq 1, \quad r(0) = 1, \quad (43)$$

$$Y(p, t) = [y(t), y(t-1), \dots, y(t-p+1)]^T, \quad (44)$$

$$\hat{\Phi}(p, t) = [\hat{\varphi}(t), \hat{\varphi}(t-1), \dots, \hat{\varphi}(t-p+1)], \quad (45)$$

$$\hat{\varphi}(t) = [-x_a(t-1), \dots, -x_a(t-n_a), u(t-1), \dots, u(t-n_b)]^T, \quad (46)$$

$$x_a(t) = \hat{\varphi}^T(t) \hat{\theta}(t), \quad x_a(i) = 1/p_0, \quad i \leq 0, \quad (47)$$

$$\hat{\theta}(t) = [\hat{\theta}_1(t), \hat{\theta}_2(t), \dots, \hat{\theta}_{n_a}(t), \hat{\theta}_1(t), \hat{\theta}_2(t), \dots, \hat{\theta}_{n_b}(t)]^T. \quad (48)$$

当遗忘因子  $\lambda = 1$ , 收敛指数  $\varepsilon = 1$  时, 修正遗忘因子 AM-MISG 算法退化为 AM-MISG 算法(33)–(40); 当遗忘因子收敛指数  $\varepsilon = 1$  时, 修正遗忘因子 AM-MISG 算法退化为遗忘因子 AM-MISG 算法(FF-AM-MISG 算法); 当遗忘因子  $\lambda = 1$  时, 修正遗忘因子 AM-MISG 算法退化为修正 AM-MISG 算法(M-AM-MISG 算法). 此外, 在收敛性分析中, 式(35)和(43)也可修改为

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

**注 2** 文献[42]证明了输出误差系统的 AM-MISG 算法的收敛性. 读者可以研究修正 AM-MISG 算法的收敛性. 遗忘因子 AM-MISG 算法的有界收敛性是辨识领域的研究难题.

### 1.3.3 变递推间隔辅助模型多新息随机梯度算法

根据多新息辨识理论, 基于 V-AM-SG 算法(26)–(32), 将输出  $y(t_s)$  和信息向量  $\hat{\varphi}(t_s)$  扩展为堆积输出向量  $Y(p, t_s)$  和堆积信息矩阵  $\hat{\Phi}(p, t_s)$ :

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

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

将式(26)中标量新息  $e(t_s) := y(t_s) - \hat{\varphi}^T(t_s) \hat{\theta}(t_{s-1}) \in \mathbf{R}$  扩展为新息向量

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

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

便获得估计参数向量  $\theta$  的变递推间隔辅助模型多新息随机梯度算法(interval-Varying AM-MISG algorithm, V-AM-MISG 算法):

$$\hat{\theta}(t_s) = \hat{\theta}(t_{s-1}) + \frac{\hat{\Phi}(p, t_s) E(p, t_s)}{r(t_s)}, \quad \hat{\theta}(t_0) = \mathbf{1}_n / p_0, \quad (49)$$

$$\hat{\theta}(t) = \hat{\theta}(t_{s-1}), \quad t \in [t_{s-1} + 1, t_s - 1], \quad (50)$$

$$E(p, t_s) = Y(p, t_s) - \hat{\Phi}^T(p, t_s) \hat{\theta}(t_{s-1}), \quad (51)$$

$$r(t_s) = r(t_{s-1}) + \|\hat{\varphi}(t_s)\|^2, \quad r(t_0) = 1, \quad (52)$$

$$Y(p, t_s) = [y(t_s), y(t_{s-1}), \dots, y(t_{s-p+1})]^T, \quad (53)$$

$$\hat{\Phi}(p, t_s) = [\hat{\varphi}(t_s), \hat{\varphi}(t_{s-1}), \dots, \hat{\varphi}(t_{s-p+1})], \quad (54)$$

$$\varphi(t_s) = [-x_a(t_s-1), \dots, -x_a(t_s-n_a), u(t_s-1), \dots, u(t_s-n_b)]^T, \quad (55)$$

$$x_a(t) = \hat{\varphi}^T(t) \hat{\theta}(t), \quad x_a(i) = 1/p_0, \quad i \leq 0. \quad (56)$$

当  $t_s^* \equiv d$  (常数) 时, V-AM-MISG 算法退化为等递推间隔辅助模型多新息随机梯度算法(E-AM-MISG 算法). 当  $t_s^* \equiv 1$  时, V-AM-MISG 算法退化为辅助模型多新息随机梯度算法(AM-MISG 算法). 文献[45]证明了 V-AM-MISG 算法参数估计和真实输出估计的收敛性.

## 1.4 辅助模型递推最小二乘算法

### 1.4.1 辅助模型递推最小二乘算法

对于输出误差系统辨识模型(4), 定义准则函数

$$J_3(\theta) := \sum_{j=1}^t [y(j) - \varphi^T(j) \theta]^2.$$

参照文献[1-2]中 RLS 算法的推导, 可以得到下列最小二乘递推关系:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t) [y(t) - \varphi^T(t) \hat{\theta}(t-1)], \quad (57)$$

$$L(t) = P(t) \varphi(t) =$$

$$P(t-1) \varphi(t) [1 + \varphi^T(t) P(t-1) \varphi(t)]^{-1}, \quad (58)$$

$$P(t) = [I_n - L(t) \varphi^T(t)] P(t-1). \quad (59)$$

由于信息向量  $\varphi(t)$  中包含了未知  $x(t-i)$ , 故是未知的, 使得关系(57)–(59)无法计算. 解决方法是用辅助模型的输出  $x_a(t-i)$  和系统输入  $u(t-i)$  定义  $\varphi(t)$  的估计

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

取计算  $x_a(t)$  的辅助模型如下:

$$x_a(t) := \hat{\varphi}^T(t) \hat{\theta}(t).$$

式(57)–(59)中未知向量  $\varphi(t)$  用其估计  $\hat{\varphi}(t)$  代替,可以得到估计输出误差系统参数向量  $\theta$  的辅助模型递推最小二乘算法(Auxiliary Model based Recursive Least Squares algorithm, AM-RLS 算法)<sup>[1]</sup>:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t) [y(t) - \hat{\varphi}^T(t) \hat{\theta}(t-1)], \quad (60)$$

$$L(t) = P(t) \hat{\varphi}(t) = P(t-1) \hat{\varphi}(t) [1 + \hat{\varphi}^T(t) P(t-1) \hat{\varphi}(t)]^{-1}, \quad (61)$$

$$P(t) = [I_n - L(t) \hat{\varphi}^T(t)] P(t-1), \quad (62)$$

$$\varphi(t) = [-x_a(t-1), \dots, -x_a(t-n_a), u(t-1), \dots, u(t-n_b)]^T, \quad (63)$$

$$x_a(t) = \hat{\varphi}^T(t) \hat{\theta}(t), \quad (64)$$

$$\hat{\theta}(t) = [\hat{\alpha}_1(t), \hat{\alpha}_2(t), \dots, \hat{\alpha}_{n_a}(t), \hat{b}_1(t), \hat{b}_2(t), \dots, \hat{b}_{n_b}(t)]^T. \quad (65)$$

AM-RLS 算法(60)–(65)随  $t$  增加,计算参数估计向量  $\hat{\theta}(t)$  的步骤如下:

1) 初始化:令  $t=1$ .置初值  $\hat{\theta}(0) = \mathbf{1}_n/p_0$ ,  $P(0) = p_0 I_n$ ,  $x_a(t-i) = 1/p_0$ ,  $i=1, 2, \dots, n_a$ ,  $p_0 = 10^6$ .

2) 采集输入输出数据  $u(t)$  和  $y(t)$ ,由式(63)构造信息向量  $\hat{\varphi}(t)$ .

3) 由式(61)计算增益向量  $L(t)$ ,由式(62)计算协方差阵  $P(t)$ .

4) 根据式(60)刷新参数估计向量  $\hat{\theta}(t)$ .

5) 由式(64)计算辅助模型的输出  $x_a(t)$ .

6)  $t$  增 1,转到第 2)步.

表 1 给出了 AM-RLS 算法每递推计算一步的计算量( $n=n_a+n_b$ ).

**注 3** 文献[61]证明了双率系统辅助模型递推最小二乘算法,即等递推间隔辅助模型递推最小二乘算法(E-AM-SG 算法)的收敛性.文献[65-66]研究了多变量输出误差系统辅助模型多变量递推最小二乘算法(AM-MRLS 算法)的收敛性.这里的 AM-RLS

算法(60)–(65)是它们的特殊情形,读者可以仿照加以证明.带遗忘因子 AM-RLS 算法参数估计的有界收敛性证明仍然是辨识领域的研究难题.

#### 1.4.2 变递推间隔辅助模型递推最小二乘算法

对于辨识模型(25),定义准则函数

$$J_4(\theta) := \sum_{j=1}^s [y(t_j) - \varphi^T(t_j) \theta]^2.$$

参照 AM-RLS 算法(60)–(65)的推导,由辅助模型的输出  $x_a(t_s-i)$  和系统输入  $u(t_s-i)$  来定义未知向量  $\varphi(t_s)$  的估计:

$$\hat{\varphi}(t_s) := [-x_a(t_s-1), -x_a(t_s-2), \dots, -x_a(t_s-n_a), u(t_s-1), u(t_s-2), \dots, u(t_s-n_b)]^T \in \mathbf{R}^n,$$

取计算  $x_a(t)$  的辅助模型如下:

$$x_a(t) := \hat{\varphi}^T(t) \hat{\theta}(t).$$

未知向量  $\varphi(t_s)$  用其估计  $\hat{\varphi}(t_s)$  代替,可以得到估计输出误差系统参数向量  $\theta$  的变递推间隔辅助模型递推最小二乘算法(interval-Varying AM-RLS algorithm, V-AM-RLS 算法)<sup>[63]</sup>:

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

$$\hat{\theta}(t_0) = \mathbf{1}_n/p_0, \quad (66)$$

$$\hat{\theta}(t) = \hat{\theta}(t_{s-1}), \quad t \in [t_{s-1}+1, t_s-1], \quad (67)$$

$$L(t_s) = P(t_s) \hat{\varphi}(t_s) = P(t_{s-1}) \hat{\varphi}(t_s) [1 + \hat{\varphi}^T(t_s) P(t_{s-1}) \hat{\varphi}(t_s)]^{-1}, \quad (68)$$

$$P(t_s) = [I_n - L(t_s) \hat{\varphi}^T(t_s)] P(t_{s-1}), \quad P(t_0) = p_0 I_n, \quad (69)$$

$$\hat{\varphi}(t_s) = [-x_a(t_s-1), \dots, -x_a(t_s-n_a), u(t_s-1), \dots, u(t_s-n_b)]^T, \quad (70)$$

$$x_a(t) = \hat{\varphi}^T(t) \hat{\theta}(t), \quad x_a(i) = 1/p_0, \quad i \leq 0, \quad (71)$$

$$\hat{\theta}(t) = [\hat{\alpha}_1(t), \hat{\alpha}_2(t), \dots, \hat{\alpha}_{n_a}(t), \hat{b}_1(t), \hat{b}_2(t), \dots, \hat{b}_{n_b}(t)]^T. \quad (72)$$

**注 4** 文献[63]证明了 V-AM-RLS 算法参数估计误差和真实输出估计误差的收敛性.对于信息向

表 1 AM-RLS 算法每步的计算量

Table 1 The computational efficiency of the AM-RLS algorithm

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

量包含未知量(如不可测真实输出、不可测噪声项)的线性时不变系统,所有带遗忘因子的辨识算法、有限数据窗辨识算法(即增益向量不趋于零的所有算法)的参数估计有界收敛性或估计误差上界都是开放的辨识难题,值得深入研究,如 CARMA 系统遗忘因子增广随机梯度算法、遗忘因子递推增广最小二乘算法、有限数据窗递推增广最小二乘算法,输出误差系统遗忘因子辅助模型随机梯度算法、遗忘因子辅助模型递推最小二乘算法等。

### 1.5 辅助模型多新息最小二乘算法

#### 1.5.1 辅助模型多新息最小二乘算法

根据多新息辨识理论,基于 AM-RLS 算法(60)—(65),将系统输出  $y(t)$  和信息向量  $\hat{\varphi}(t)$  扩展为堆积输出向量  $Y(p,t)$  和堆积信息矩阵  $\hat{\Phi}(p,t)$ :

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

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

将式(60)中标量新息  $e(t) := y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1) \in \mathbf{R}$  扩展为新息向量

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

可以得到估计输出误差系统参数向量  $\theta$  的辅助模型多新息最小二乘算法(Auxiliary Model based Multi-Innovation Least Squares algorithm, AM-MILS 算法):

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t) [Y(p,t) - \hat{\Phi}^T(p,t)\hat{\theta}(t-1)], \quad (73)$$

$$L(t) = P(t) \hat{\Phi}(p,t) =$$

$$P(t-1) \hat{\Phi}(p,t) [I_p + \hat{\Phi}^T(p,t)P(t-1)\hat{\Phi}(p,t)]^{-1}, \quad (74)$$

$$P(t) = P(t-1) - P(t-1) \hat{\Phi}(p,t) [I_p +$$

$$\hat{\Phi}^T(p,t)P(t-1)\hat{\Phi}(p,t)]^{-1} \hat{\Phi}^T(p,t)P(t-1) = [I_n - L(t)\hat{\Phi}^T(p,t)]P(t-1), \quad (75)$$

$$Y(p,t) = [y(t), y(t-1), \dots, y(t-p+1)]^T, \quad (76)$$

$$\hat{\Phi}(p,t) = [\hat{\varphi}(t), \hat{\varphi}(t-1), \dots, \hat{\varphi}(t-p+1)], \quad (77)$$

$$\hat{\varphi}(t) = [-x_a(t-1), \dots, -x_a(t-n_a), u(t-1), \dots, u(t-n_b)]^T, \quad (78)$$

$$x_a(t) = \hat{\varphi}^T(t)\hat{\theta}(t), \quad (79)$$

$$\hat{\theta}(t) = [\hat{a}_1(t), \hat{a}_2(t), \dots, \hat{a}_{n_a}(t), \hat{b}_1(t), \hat{b}_2(t), \dots, \hat{b}_{n_b}(t)]^T. \quad (80)$$

当新息长度  $p=1$  时,AM-MILS 算法退化为 AM-RLS 算法(60)—(65)。

AM-MILS 算法(73)—(80)随  $t$  增加,计算参数估计向量  $\hat{\theta}(t)$  的步骤如下:

1) 初始化:令  $t=1$ ,给定新息长度  $p$ .置  $\hat{\theta}(0) = \mathbf{1}_n/p_0, P(0) = p_0\mathbf{I}_n, x_a(t-i) = 1/p_0, i = 1, 2, \dots, n_a, p_0 = 10^6$ .

2) 采集输入输出数据  $u(t)$  和  $y(t)$ ,由式(78)

构造信息向量  $\hat{\varphi}(t)$ ,式(76)—(77)构造堆积输出向量  $Y(p,t)$  和堆积信息矩阵  $\hat{\Phi}(p,t)$ .

3) 由式(74)计算增益向量  $L(t)$ ,由式(75)计算协方差阵  $P(t)$ .

4) 根据式(73)刷新参数估计向量  $\hat{\theta}(t)$ .

5) 由式(79)计算辅助模型的输出  $x_a(t)$ .

6)  $t$  增 1,转到第 2)步。

#### 1.5.2 变递推间隔辅助模型多新息最小二乘算法

根据多新息辨识理论,基于 V-AM-RLS 算法(66)—(72),将系统输出  $y(t_s)$  和信息向量  $\hat{\varphi}(t_s)$  扩展为堆积输出向量  $Y(p,t_s)$  和堆积信息矩阵  $\hat{\Phi}(p,t_s)$ :

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

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

将式(66)中标量新息  $e(t_s) := y(t_s) - \hat{\varphi}^T(t_s)\hat{\theta}(t_{s-1}) \in \mathbf{R}$  扩展为新息向量

$$E(p,t_s) := Y(p,t_s) - \hat{\Phi}^T(p,t_s)\hat{\theta}(t_{s-1}) \in \mathbf{R}^p,$$

可以得到估计输出误差系统参数向量  $\theta$  的变递推间隔辅助模型多新息最小二乘算法(interval-Varying AM-MILS algorithm, V-AM-MILS 算法):

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

$$\hat{\theta}(t_0) = \mathbf{1}_n/p_0, \quad (81)$$

$$\hat{\theta}(t) = \hat{\theta}(t_{s-1}), \quad t \in [t_{s-1}+1, t_s-1], \quad (82)$$

$$L(t_s) = P(t_s) \hat{\Phi}(p,t_s) = P(t_{s-1}) \hat{\Phi}(p,t_s) [I_p + \hat{\Phi}^T(p,t_s)P(t_{s-1})\hat{\Phi}(p,t_s)]^{-1}, \quad (83)$$

$$P(t_s) = P(t_{s-1}) - P(t_{s-1}) \hat{\Phi}(p,t_s) [I_p + \hat{\Phi}^T(p,t_s)P(t_{s-1})\hat{\Phi}(p,t_s)]^{-1} \hat{\Phi}^T(p,t_s)P(t_{s-1}) = [I_n - L(t_s)\hat{\Phi}^T(p,t_s)]P(t_{s-1}), \quad P(t_0) = p_0\mathbf{I}_n, \quad (84)$$

$$Y(p,t_s) = [y(t_s), y(t_{s-1}), \dots, y(t_{s-p+1})]^T, \quad (85)$$

$$\hat{\Phi}(p,t_s) = [\hat{\varphi}(t_s), \hat{\varphi}(t_{s-1}), \dots, \hat{\varphi}(t_{s-p+1})], \quad (86)$$

$$\hat{\varphi}(t_s) = [-x_a(t_s-1), \dots, -x_a(t_s-n_a), u(t_s-1), \dots, u(t_s-n_b)]^T, \quad (87)$$

$$x_a(t) = \hat{\varphi}^T(t)\hat{\theta}(t), \quad x_a(i) = 1/p_0, i \leq 0, \quad (88)$$

$$\hat{\theta}(t_s) = [\hat{a}_1(t_s), \hat{a}_2(t_s), \dots, \hat{a}_{n_a}(t_s), \hat{b}_1(t_s), \hat{b}_2(t_s), \dots, \hat{b}_{n_b}(t_s)]^T. \quad (89)$$

注 5 当系统输入数据都可得到,只有输出数据存在缺失时,采用输出误差模型有其独到的优点,研究表明:在持续激励条件下,参数估计收敛于真参数,辅助模型的输出收敛于系统真实输出<sup>[47,63]</sup>。

## 2 Box-Jenkins 系统

本节研究 Box-Jenkins 系统的辅助模型辨识算

法,包括辅助模型广义增广随机梯度(AM-GESG)算法、辅助模型多新息广义增广随机梯度(AM-MI-GESG)算法、辅助模型递推广义增广最小二乘(AM-RGELS)算法、辅助模型多新息广义增广最小二乘(AM-MI-GELS)算法.

文献[67]研究了 Box-Jenkins 系统的辅助模型多新息广义增广随机梯度辨识算法的收敛性.文献[68]研究了多变量 Box-Jenkins 系统的辅助模型广义增广随机梯度辨识算法和辅助模型递推广义增广最小二乘算法的收敛性.Box-Jenkins 系统的“基于分解的辅助模型多新息广义增广随机梯度算法”和“基于分解的辅助模型递推广义增广最小二乘算法”发表在国际期刊《IET Signal Processing》2013 年第 8 期<sup>[51]</sup>上.该文获得《IET Journals》杂志的最佳论文奖“Premium (Best Paper) Awards”,该奖是《IET Journals》杂志每年从前两年发表论文中评选出的唯一一篇最佳论文(见 <http://digital-library.theiet.org/journals/premium-awards>).

## 2.1 系统描述与辨识模型

考虑下列 Box-Jenkins 模型(Box-Jenkins model, BJ 模型)描述的动态随机系统(图 2):

$$y(t) = \frac{B(z)}{A(z)}u(t) + \frac{D(z)}{C(z)}v(t), \quad (90)$$

其中  $\{u(t)\}$  和  $\{y(t)\}$  分别为系统的输入和输出序列,  $\{v(t)\}$  为零均值、不相关随机白噪声序列(不可测),  $A(z)$ ,  $B(z)$ ,  $C(z)$  和  $D(z)$  均为单位后移算子  $z^{-1}$  的多项式:

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

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

$$C(z) = 1 + c_1 z^{-1} + c_2 z^{-2} + \dots + c_{n_c} z^{-n_c},$$

$$D(z) = 1 + d_1 z^{-1} + d_2 z^{-2} + \dots + d_{n_d} z^{-n_d}.$$

设阶次  $n_a$ ,  $n_b$ ,  $n_c$  和  $n_d$  已知,记  $n := n_a + n_b + n_c + n_d$ ,且  $t \leq 0$  时,  $y(t) = 0$ ,  $u(t) = 0$ ,  $v(t) = 0$ .

Box-Jenkins 模型又称为输出误差自回归滑动平均模型(Output-Error AutoRegressive Moving Average model, OEARMA 模型).

图 2 中,  $G(z) := \frac{B(z)}{A(z)}$  是系统模型的传递函数,

$H(z) := \frac{D(z)}{C(z)}$  是噪声模型的传递函数,  $x(t)$  是系统的真实输出,即不可测无噪输出,  $w(t)$  是噪声模型的输出,也是不可测的,  $y(t)$  是  $x(t)$  的含噪量测.

定义未知中间变量  $x(t)$  和不可测噪声项  $w(t)$

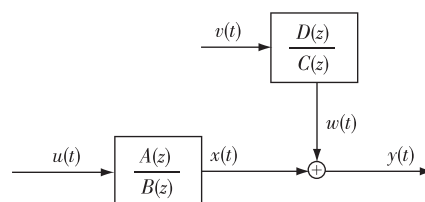


图 2 Box-Jenkins 系统

Fig. 2 The Box-Jenkins(BJ) system

分别为

$$x(t) := \frac{B(z)}{A(z)}u(t), \quad (91)$$

$$w(t) := \frac{D(z)}{C(z)}v(t). \quad (92)$$

Box-Jenkins 系统的干扰噪声  $w(t)$  是一个自回归滑动平均(ARMA)过程,是一个有色干扰噪声.

定义参数向量  $\theta$  和信息向量  $\varphi(t)$  如下:

$$\theta := \begin{bmatrix} \theta_s \\ \theta_n \end{bmatrix} \in \mathbf{R}^n,$$

$$\theta_s := [a_1, a_2, \dots, a_{n_a}, b_1, b_2, \dots, b_{n_b}]^T \in \mathbf{R}^{n_a+n_b},$$

$$\theta_n := [c_1, c_2, \dots, c_{n_c}, d_1, d_2, \dots, d_{n_d}]^T \in \mathbf{R}^{n_c+n_d},$$

$$\varphi(t) := \begin{bmatrix} \varphi_s(t) \\ \varphi_n(t) \end{bmatrix} \in \mathbf{R}^n,$$

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

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

由式(91)–(92)可得

$$x(t) = [1 - A(z)]x(t) + B(z)u(t) = \varphi_s^T(t)\theta_s, \quad (93)$$

$$w(t) = [1 - C(z)]w(t) + D(z)v(t) = \varphi_n^T(t)\theta_n + v(t), \quad (94)$$

将式(91)和(92)代入式(90)可得

$$y(t) = x(t) + w(t) = \varphi_s^T(t)\theta_s + w(t) = \quad (95)$$

$$\varphi_s^T(t)\theta_s + \varphi_n^T(t)\theta_n + v(t) = \varphi^T(t)\theta + v(t). \quad (96)$$

## 2.2 辅助模型广义增广随机梯度算法

设  $\hat{\theta}(t) := \begin{bmatrix} \hat{\theta}_s(t) \\ \hat{\theta}_n(t) \end{bmatrix} \in \mathbf{R}^n$  是  $\theta = \begin{bmatrix} \theta_s \\ \theta_n \end{bmatrix}$  在时刻  $t$  的

估计,  $1/r(t)$  是收敛因子或步长. 矩阵  $X$  的范数定义为  $\|X\|^2 := \text{tr}[XX^T]$ .

根据辨识模型(96), 引入二次准则函数

$$J_5(\theta) := \frac{1}{2} [y(t) - \varphi^T(t)\theta]^2,$$

使用负梯度搜索, 极小化  $J_5(\theta)$ , 得到下列递推关系:



$$\hat{\theta}(t) = \hat{\theta}(t-1) - \frac{1}{r(t)} \text{grad}[J_5(\hat{\theta}(t-1))] = \hat{\theta}(t-1) + \frac{\varphi(t)}{r(t)} [y(t) - \varphi^T(t) \hat{\theta}(t-1)], \quad (97)$$

$$r(t) = r(t-1) + \|\varphi(t)\|^2, \quad r(0) = 1. \quad (98)$$

根据信息向量  $\varphi(t)$  的定义可知,  $\varphi(t)$  由  $\varphi_s(t)$  和  $\varphi_n(t)$  两部分构成,  $\varphi_s(t)$  是由系统可测输入  $u(t-i)$  和未知无噪输出  $x(t-i)$  构成的,  $\varphi_n(t)$  是由不可测噪声项  $w(t-i)$  和  $v(t-i)$  构成的, 使得递推关系式 (97) — (98) 无法计算参数估计  $\hat{\theta}(t)$ , 即无法辨识 Box-Jenkins 系统 (90) 的参数. 解决方法是利用辅助模型辨识思想, 利用系统的可测信息 (包括计算得到的信息) 建立一个辅助模型, 用辅助模型的输出  $x_a(t-i)$  代替信息向量  $\varphi(t)$  中的未知无噪输出项  $x(t-i)$ , 未知噪声项  $w(t-i)$  和  $v(t-i)$  也通过两个辅助模型进行估算, 即用两个辅助模型的输出  $\hat{w}(t-i)$  和  $\hat{v}(t-i)$  代替信息向量  $\varphi_n(t)$  中的未知项, 从而实现算法参数估计的递推计算. 具体细节如下.

基于辅助模型辨识思想<sup>[1]</sup>, 构造一个辅助模型  $P_a(z)$ , 如图 3 的下部分所示, 其中  $P_a(z) := \frac{B_a(z)}{A_a(z)}$  为

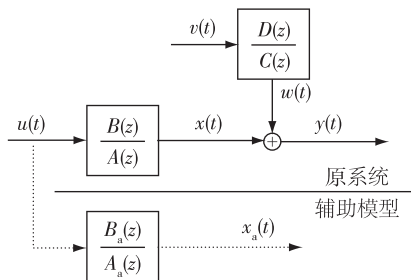


图 3 带辅助模型的 Box-Jenkins 模型描述的系统

Fig. 3 The Box-Jenkins systems with the auxiliary model

辅助模型的传递函数, 这里假定  $A_a(z)$  和  $B_a(z)$  是与  $A(z)$  和  $B(z)$  结构和阶次相同的多项式,  $x_a(t)$  为辅助模型的输出. 当然, 辅助模型  $P_a(z)$  也可以取与  $G(z) = \frac{B(z)}{A(z)}$  对应的有限脉冲响应模型<sup>[62]</sup>. 信息向量  $\varphi(t)$  里  $\varphi_s(t)$  中未知项  $x(t-i)$  用辅助模型的输出  $x_a(t-i)$  代替, 将代替后的  $\varphi_s(t)$  记作

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

它可看作  $\varphi_s(t)$  的估计. 又信息向量  $\varphi(t)$  里  $\varphi_n(t)$  中未知噪声项  $w(t-i)$  和  $v(t-i)$  用其对应的估计  $\hat{w}(t-i)$  和  $\hat{v}(t-i)$  代替, 将代替后的  $\varphi_n(t)$  记作

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

$\hat{\varphi}_n(t)$  可看作  $\varphi_n(t)$  的估计. 由图 3 可知  $x_a(t) = \frac{B_a(z)}{A_a(z)} u(t)$ , 它与式 (91) 具有相同的结构形式. 仿照式 (91) 写作式 (93) 的方式, 把它写成向量形式为  $x_a(t) = \varphi_a^T(t) \theta_a$ , 其中  $\varphi_a(t)$  和  $\theta_a$  分别是辅助模型的信息向量和参数向量. 当然, 辅助模型的信息向量和参数向量有多种选择方法<sup>[61-62]</sup>, 参见即将出版的《系统辨识——辅助模型辨识思想与方法》一书. 参照文献 [1, 47, 61, 63], 把  $\hat{\varphi}_s(t)$  作为辅助模型的信息向量  $\varphi_a(t)$ , 把  $\hat{\theta}_s(t)$  作为辅助模型的参数向量  $\theta_a$ , 得到

$$x_a(t) = \hat{\varphi}_s^T(t) \hat{\theta}_s(t).$$

用  $\hat{\varphi}_s(t)$  和  $\hat{\varphi}_n(t)$  代替  $\varphi(t)$  中 (未知的)  $\varphi_s(t)$  和  $\varphi_n(t)$ , 代替后记作

$$\hat{\varphi}(t) := \begin{bmatrix} \hat{\varphi}_s(t) \\ \hat{\varphi}_n(t) \end{bmatrix} \in \mathbf{R}^n,$$

它可看作  $\varphi(t)$  的估计. 由式 (95) 可得  $w(t) = y(t) - x(t)$ , 上式中未知变量  $x(t)$  用辅助模型的输出  $x_a(t)$  代替, 便得到计算  $\hat{w}(t)$  的辅助模型

$$\hat{w}(t) = y(t) - x_a(t) = y(t) - \hat{\varphi}_s^T(t) \hat{\theta}_s(t).$$

由式 (96) 可得  $v(t) = y(t) - \varphi^T(t) \theta$ , 上式中未知的  $\varphi(t)$  和  $\theta$  分别用其估计  $\hat{\varphi}(t)$  和  $\hat{\theta}(t)$  代替, 就得到计算  $\hat{v}(t)$  的辅助模型

$$\hat{v}(t) = y(t) - \hat{\varphi}_s^T(t) \hat{\theta}_s(t) - \hat{\varphi}_n^T(t) \hat{\theta}_n(t) = \hat{w}(t) - \hat{\varphi}_n^T(t) \hat{\theta}_n(t).$$

上式也可以直接根据式 (94), 未知量  $w(t)$ ,  $\varphi_n(t)$  和  $\theta_n$  分别用其估计  $\hat{w}(t)$ ,  $\hat{\varphi}_n(t)$  和  $\hat{\theta}_n(t)$  代替得到.

式 (97) — (98) 中未知的  $\varphi(t)$  用其估计  $\hat{\varphi}(t)$  代替, 可以总结出估计 Box-Jenkins 系统参数向量  $\theta$  的辅助模型广义增广随机梯度算法 (Auxiliary Model based Generalized Extended Stochastic Gradient algorithm, AM-GESG 算法)<sup>[32]</sup>:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\hat{\varphi}(t)}{r(t)} e(t), \quad (99)$$

$$e(t) = y(t) - \hat{\varphi}^T(t) \hat{\theta}(t-1), \quad (100)$$

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

$$\hat{\varphi}(t) = \begin{bmatrix} \hat{\varphi}_s(t) \\ \hat{\varphi}_n(t) \end{bmatrix}, \quad (102)$$

$$\hat{\varphi}_s(t) = [-x_a(t-1), \dots, -x_a(t-n_a), u(t-1), \dots, u(t-n_b)]^T, \quad (103)$$

$$\hat{\varphi}_n(t) = [-\hat{w}(t-1), \dots, -\hat{w}(t-n_c), \hat{v}(t-1), \dots, \hat{v}(t-n_d)]^T, \quad (104)$$

$$x_a(t) = \hat{\varphi}_s^T(t) \hat{\theta}_s(t), \quad (105)$$

$$\hat{w}(t) = y(t) - x_a(t), \quad (106)$$

$$v(t) = \hat{w}(t) - \hat{\varphi}_n^T(t) \hat{\theta}_n(t), \quad (107)$$

$$\theta(t) = \begin{bmatrix} \hat{\theta}_s(t) \\ \hat{\theta}_n(t) \end{bmatrix}. \quad (108)$$

AM-GESG 算法(99)——(108)随  $t$  增加, 计算参数估计向量  $\hat{\theta}(t)$  的步骤如下:

1) 初始化: 令  $t=1$ . 置初值  $\hat{\theta}(0) = \mathbf{1}_n/p_0, r(0) = 1, x_a(t-i) = 1/p_0, \hat{w}(t-i) = 1/p_0, \hat{v}(t-i) = 1/p_0, i=1, 2, \dots, \max[n_a, n_e, n_d], p_0 = 10^6, \mathbf{1}_n$  是一个元均为 1 的  $n$  维列向量.

2) 采集输入输出数据  $u(t)$  和  $y(t)$ , 由式(103)、(104)和(102)构造信息向量  $\hat{\varphi}_s(t), \hat{\varphi}_n(t)$  和  $\hat{\varphi}(t)$ .

3) 由式(100)计算新息  $e(t)$ , 由式(101)计算  $r(t)$ .

4) 根据式(99)刷新参数估计向量  $\hat{\theta}(t)$ .

5) 从式(108)的  $\hat{\theta}(t)$  中读取  $\hat{\theta}_s(t)$  和  $\hat{\theta}_n(t)$ . 由式(105)——(107)计算辅助模型的输出  $x_a(t), \hat{w}(t)$  和  $\hat{v}(t)$ .

6)  $t$  增 1, 转到第 2) 步.

**注 6** 为提高 AM-GESG 算法的暂态性能和参数估计精度, 可引入遗忘因子  $\lambda$ , 将式(101)修改为  $r(t) = \lambda r(t-1) + \|\hat{\varphi}(t)\|^2, 0 \leq \lambda \leq 1, r(0) = 1$ , 就得到遗忘因子辅助模型广义增广随机梯度算法 (Forgetting Factor AM-GESG algorithm, FF-AM-GESG 算法).

**注 7** 为提高 AM-GESG 算法的暂态性能和稳态性能, 可引入收敛指数  $\varepsilon$ , 将式(99)修改为

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\hat{\varphi}(t)}{r^\varepsilon(t)} e(t), \quad \frac{1}{2} < \varepsilon \leq 1,$$

就得到修正辅助模型广义增广随机梯度算法 (Modified AM-GESG algorithm, M-AM-GESG 算法).

### 2.3 辅助模型多新息广义增广随机梯度算法

尽管上述辅助模型广义增广随机梯度辨识算法(99)——(108)可以估计 Box-Jenkins 系统的参数, 但是算法的收敛速度很慢, 下面借助于多新息辨识理论, 通过扩展新息维数, 推导出基于辅助模型的多新息随机梯度算法, 从而提高算法的收敛速度和参数估计精度. 具体作法如下.

将 AM-GESG 算法(99)——(108)中标量新息  $e(t) = y(t) - \hat{\varphi}^T(t) \hat{\theta}(t-1) \in \mathbf{R}$  扩展为新息向量

$$E(p, t) := \begin{bmatrix} e(t) \\ e(t-1) \\ \vdots \\ e(t-p+1) \end{bmatrix} = \begin{bmatrix} y(t) - \hat{\varphi}^T(t) \hat{\theta}(t-1) \\ y(t-1) - \hat{\varphi}^T(t-1) \hat{\theta}(t-2) \\ \vdots \\ y(t-p+1) - \hat{\varphi}^T(t-p+1) \hat{\theta}(t-p) \end{bmatrix} \in \mathbf{R}^p,$$

其中正整数  $p$  表示新息长度.

通常可认为时刻  $(t-1)$  的估计  $\hat{\theta}(t-1)$  比时刻  $(t-i) (i \geq 2)$  的估计  $\hat{\theta}(t-i)$  更接近真值  $\theta$ . 因此, 将  $E(p, t)$  表达式中  $\hat{\theta}(t-i) (i \geq 2)$  换为  $\hat{\theta}(t-1)$ , 则多新息向量可合理修改为

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

定义堆积输出向量  $Y(p, t)$  和堆积信息矩阵  $\hat{\Phi}(p, t)$  如下:

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

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

多新息向量可以表达为

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

注意到  $Y(1, t) = y(t), E(1, t) = e(t), \hat{\Phi}(1, t) = \hat{\varphi}(t)$ , 式(99)可以等价表达为

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\hat{\Phi}(1, t)}{r(t)} E(1, t).$$

上式中新息长度  $p=1$ . 这是新息长度为 1 的“多新息”随机梯度算法. 把上式中  $\hat{\Phi}(1, t)$  和  $E(1, t)$  里的“1”换为  $p$ , 可以得到新息长度为  $p$  的, 估计 Box-Jenkins 系统参数向量  $\theta$  的辅助模型多新息广义增广随机梯度算法 (Auxiliary Model based Multi-Innovation Generalized Extended Stochastic Gradient algorithm, AM-MI-GESG 算法)<sup>[32]</sup>:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\hat{\Phi}(p, t)}{r^\varepsilon(t)} E(p, t), \quad \frac{1}{2} < \varepsilon \leq 1, \quad (109)$$

$$E(p, t) = Y(p, t) - \hat{\Phi}^T(p, t) \hat{\theta}(t-1), \quad (110)$$

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

$$Y(p, t) = [y(t), y(t-1), \dots, y(t-p+1)]^T, \quad (112)$$

$$\hat{\Phi}(p, t) = [\hat{\varphi}(t), \hat{\varphi}(t-1), \dots, \hat{\varphi}(t-p+1)], \quad (113)$$

$$\hat{\varphi}(t) = \begin{bmatrix} \hat{\varphi}_s(t) \\ \hat{\varphi}_n(t) \end{bmatrix}, \quad (114)$$

$$\hat{\varphi}_s(t) = [-x_a(t-1), \dots, -x_a(t-n_a), u(t-1), \dots, u(t-n_b)]^T, \quad (115)$$

$$\hat{\varphi}_n(t) = [-\hat{u}(t-1), \dots, -\hat{u}(t-n_c), \hat{v}(t-1), \dots, \hat{v}(t-n_d)]^T, \quad (116)$$

$$x_a(t) = \hat{\varphi}_s^T(t) \hat{\theta}_s(t), \quad (117)$$

$$\hat{w}(t) = y(t) - x_a(t), \quad (118)$$

$$\hat{v}(t) = \hat{w}(t) - \hat{\varphi}_n^T(t) \hat{\theta}_n(t), \quad (119)$$

$$\hat{\theta}(t) = \begin{bmatrix} \hat{\theta}_s(t) \\ \hat{\theta}_n(t) \end{bmatrix}. \quad (120)$$

当新息长度  $p=1$  时, AM-MI-GESG 算法退化为 AM-GESG 算法(90) — (108).

AM-MI-GESG 算法(109) — (120) 随  $t$  增加, 计算参数估计向量  $\hat{\theta}(t)$  的步骤如下:

1) 初始化: 令  $t=1$ , 给定新息长度  $p$ . 置  $\hat{\theta}(0) = \mathbf{1}_n / p_0, r(0) = 1, x_a(t-i) = 1/p_0, \hat{w}(t-i) = 1/p_0, \hat{v}(t-i) = 1/p_0, i=1, 2, \dots, \max[n_a, n_c, n_d], p_0 = 10^6$ . 给定遗忘因子  $\lambda$  和收敛指数  $\varepsilon$ .

2) 采集输入输出数据  $u(t)$  和  $y(t)$ , 由式(112) 构造堆积输出向量  $\mathbf{Y}(p, t)$ , 由式(115) — (116) 和 (114) 构造信息向量  $\hat{\varphi}_s(t), \hat{\varphi}_n(t)$  和  $\hat{\varphi}(t)$ , 由式(113) 构造堆积信息矩阵  $\hat{\Phi}(p, t)$ .

3) 由式(110) 计算新息向量  $\mathbf{E}(p, t)$ , 由式(111) 计算  $r(t)$ .

4) 根据式(109) 刷新参数估计向量  $\hat{\theta}(t)$ .

5) 从式(120) 的  $\hat{\theta}(t)$  中读取  $\hat{\theta}_s(t)$  和  $\hat{\theta}_n(t)$ . 由式(117) — (119) 计算辅助模型的输出  $x_a(t), \hat{w}(t)$  和  $\hat{v}(t)$ .

6)  $t$  增 1, 转到第 2) 步.

**注 8** 这个 AM-MI-GESG 算法(109) — (120) 中已经引入了  $\lambda$  和  $\varepsilon$ , 全称应为修正遗忘因子 AM-MI-GESG 算法. 此外, 式(111) 也可修改为

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

## 2.4 辅助模型递推广义增广最小二乘算法

对于 Box-Jenkins 系统辨识模型(96), 定义准则函数

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

参考文献[1-2] 中 RLS 算法的推导, 可以得到下列最小二乘递推关系:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \mathbf{L}(t) [y(t) - \varphi^T(t)\hat{\theta}(t-1)], \quad (121)$$

$$\mathbf{L}(t) = \mathbf{P}(t-1)\varphi(t) [1 + \varphi^T(t)\mathbf{P}(t-1)\varphi(t)]^{-1}, \quad (122)$$

$$\mathbf{P}(t) = [\mathbf{I}_n - \mathbf{L}(t)\varphi^T(t)]\mathbf{P}(t-1). \quad (123)$$

在式(121) — (123) 中, 由于信息向量  $\varphi(t)$  中的  $\varphi_s(t)$  和  $\varphi_n(t)$  都是未知的, 故用辅助模型的输出

$x_a(t-i)$  和系统输入  $u(t-i)$  定义  $\varphi_s(t)$  的估计

$$\hat{\varphi}_s(t) := [-x_a(t-1), -x_a(t-2), \dots, -x_a(t-n_a), u(t-1), u(t-2), \dots, u(t-n_b)]^T \in \mathbf{R}^{n_a+n_b},$$

用噪声  $w(t)$  和  $v(t)$  的估计  $\hat{w}(t)$  和  $\hat{v}(t)$  定义  $\varphi_n(t)$  的估计

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

用  $\hat{\varphi}_s(t)$  和  $\hat{\varphi}_n(t)$  定义  $\varphi(t)$  的估计

$$\hat{\varphi}(t) := \begin{bmatrix} \hat{\varphi}_s(t) \\ \hat{\varphi}_n(t) \end{bmatrix} \in \mathbf{R}^n.$$

取计算  $x_a(t), \hat{w}(t)$  和  $\hat{v}(t)$  的辅助模型如下:

$$x_a(t) := \hat{\varphi}_s^T(t) \hat{\theta}_s(t),$$

$$\hat{w}(t) := y(t) - x_a(t),$$

$$\hat{v}(t) := \hat{w}(t) - \hat{\varphi}_n^T(t) \hat{\theta}_n(t).$$

式(121) — (123) 中未知向量  $\varphi(t)$  用其估计  $\hat{\varphi}(t)$  代替, 可以得到估计 Box-Jenkins 系统参数向量  $\theta$  的辅助模型递推广义增广最小二乘算法 (Auxiliary Model based Recursive Generalized Extended Least Squares algorithm, AM-RGELS 算法):

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \mathbf{L}(t) [y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1)], \quad (124)$$

$$\mathbf{L}(t) = \mathbf{P}(t)\hat{\varphi}(t) =$$

$$\mathbf{P}(t-1)\hat{\varphi}(t) [1 + \hat{\varphi}^T(t)\mathbf{P}(t-1)\hat{\varphi}(t)]^{-1}, \quad (125)$$

$$\mathbf{P}(t) = [\mathbf{I}_n - \mathbf{L}(t)\hat{\varphi}^T(t)]\mathbf{P}(t-1), \quad (126)$$

$$\hat{\varphi}(t) = \begin{bmatrix} \hat{\varphi}_s(t) \\ \hat{\varphi}_n(t) \end{bmatrix}, \quad (127)$$

$$\hat{\varphi}_s(t) = [-x_a(t-1), \dots, -x_a(t-n_a), u(t-1), \dots, u(t-n_b)]^T, \quad (128)$$

$$\hat{\varphi}_n(t) = [-\hat{u}(t-1), \dots, -\hat{u}(t-n_c), \hat{v}(t-1), \dots, \hat{v}(t-n_d)]^T, \quad (129)$$

$$x_a(t) = \hat{\varphi}_s^T(t) \hat{\theta}_s(t), \quad (130)$$

$$\hat{w}(t) = y(t) - x_a(t), \quad (131)$$

$$\hat{v}(t) = \hat{w}(t) - \hat{\varphi}_n^T(t) \hat{\theta}_n(t), \quad (132)$$

$$\hat{\theta}(t) = \begin{bmatrix} \hat{\theta}_s(t) \\ \hat{\theta}_n(t) \end{bmatrix}. \quad (133)$$

AM-RGELS 算法(124) — (133) 随  $t$  增加, 计算参数估计向量  $\hat{\theta}(t)$  的步骤如下:

1) 初始化: 令  $t=1$ . 置初值  $\hat{\theta}(0) = \mathbf{1}_n/p_0, \mathbf{P}(0) = p_0\mathbf{I}_n, x_a(t-i) = 1/p_0, \hat{w}(t-i) = 1/p_0, \hat{v}(t-i) = 1/p_0, i=1, 2, \dots, \max[n_a, n_c, n_d], p_0 = 10^6$ .

2) 采集输入输出数据  $u(t)$  和  $y(t)$ , 由式(128) — (129) 和 (127) 构造信息向量  $\hat{\varphi}_s(t), \hat{\varphi}_n(t)$

和  $\hat{\varphi}(t)$ .

3) 由式(125)计算增益向量  $L(t)$ , 由式(126)计算协方差阵  $P(t)$ .

4) 根据式(124)刷新参数估计向量  $\hat{\theta}(t)$ .

5) 从式(133)的  $\hat{\theta}(t)$  中读取  $\hat{\theta}_s(t)$  和  $\hat{\theta}_n(t)$ . 由式(130)—(132)计算辅助模型的输出  $x_a(t)$ ,  $\hat{w}(t)$  和  $\hat{v}(t)$ .

6)  $t$  增 1, 转到第 2) 步.

表 2 给出了 AM-RGELS 算法每递推计算一步的计算量 ( $n = n_a + n_b + n_c + n_d$ ).

## 2.5 辅助模型多新息广义增广最小二乘算法

借助于多新息辨识理论, 基于 AM-RGELS 算法(124)—(133), 将系统输出  $y(t)$  和信息向量  $\hat{\varphi}(t)$  扩展为堆积输出向量  $Y(p, t)$  和堆积信息矩阵  $\hat{\Phi}(p, t)$ :

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

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

将式(124)中标量新息  $e(t) := y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1) \in \mathbf{R}$  扩展为新息向量

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

可以得到估计 Box-Jenkins 系统参数向量  $\theta$  的辅助模型多新息广义增广最小二乘算法 (Auxiliary Model based Multi-Innovation Generalized Extended Least Squares algorithm, AM-MI-GELS 算法):

$$\theta(t) = \hat{\theta}(t-1) + L(t)[Y(p, t) - \hat{\Phi}^T(p, t)\hat{\theta}(t-1)], \quad (134)$$

$$L(t) = P(t)\hat{\Phi}(p, t) = P(t-1)\hat{\Phi}(p, t)[I_p + \hat{\Phi}^T(p, t)P(t-1)\hat{\Phi}(p, t)]^{-1}, \quad (135)$$

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

$$Y(p, t) = [y(t), y(t-1), \dots, y(t-p+1)]^T, \quad (137)$$

$$\hat{\Phi}(p, t) = [\hat{\varphi}(t), \hat{\varphi}(t-1), \dots, \hat{\varphi}(t-p+1)], \quad (138)$$

$$\hat{\varphi}(t) = \begin{bmatrix} \hat{\varphi}_s(t) \\ \hat{\varphi}_n(t) \end{bmatrix}, \quad (139)$$

$$\hat{\varphi}_s(t) = [-x_a(t-1), \dots, -x_a(t-n_a), u(t-1), \dots, u(t-n_b)]^T, \quad (140)$$

$$\hat{\varphi}_n(t) = [-\hat{w}(t-1), \dots, -\hat{w}(t-n_c), \hat{v}(t-1), \dots, \hat{v}(t-n_d)]^T, \quad (141)$$

$$x_a(t) = \hat{\varphi}_s^T(t)\hat{\theta}_s(t), \quad (142)$$

$$\hat{w}(t) = y(t) - x_a(t), \quad (143)$$

$$\hat{v}(t) = \hat{w}(t) - \hat{\varphi}_n^T(t)\hat{\theta}_n(t), \quad (144)$$

$$\hat{\theta}(t) = \begin{bmatrix} \hat{\theta}_s(t) \\ \hat{\theta}_n(t) \end{bmatrix}. \quad (145)$$

当新息长度  $p = 1$  时, AM-MI-GELS 算法退化为 AM-RGELS 算法(124)—(133).

AM-MI-GELS 算法(134)—(145)随  $t$  增加, 计算参数估计向量  $\hat{\theta}(t)$  的步骤如下:

1) 初始化: 令  $t = 1$ , 给定新息长度  $p$ . 置初值  $\hat{\theta}(0) = \mathbf{1}_n/p_0$ ,  $P(0) = p_0 I_n$ ,  $x_a(t-i) = 1/p_0$ ,  $\hat{w}(t-i) = 1/p_0$ ,  $\hat{v}(t-i) = 1/p_0$ ,  $i = 1, 2, \dots, \max[n_a, n_c, n_d]$ ,  $p_0 = 10^6$ .

2) 采集输入输出数据  $u(t)$  和  $y(t)$ , 由式(140)—(141)和(139)构造信息向量  $\hat{\varphi}_s(t)$ ,  $\hat{\varphi}_n(t)$  和  $\hat{\varphi}(t)$ , 由式(137)—(138)构造堆积输出向量  $Y(p, t)$  和堆积信息矩阵  $\hat{\Phi}(p, t)$ .

3) 由式(135)计算增益向量  $L(t)$ , 由式(136)计算协方差阵  $P(t)$ .

4) 根据式(134)刷新参数估计向量  $\hat{\theta}(t)$ .

5) 从式(145)的  $\hat{\theta}(t)$  中读取  $\hat{\theta}_s(t)$  和  $\hat{\theta}_n(t)$ . 由式(142)—(144)计算辅助模型的输出  $x_a(t)$ ,  $\hat{w}(t)$  和  $\hat{v}(t)$ .

6)  $t$  增 1, 转到第 2) 步.

表 2 AM-RGELS 算法每步的计算量

Table 2 The computational efficiency of the AM-RGELS algorithm

变量	表达式	乘法次数	加法次数
$\hat{\theta}(t)$	$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t)e(t) \in \mathbf{R}^n$	$n$	$n$
	$e(t) := y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1) \in \mathbf{R}$	$n$	$n$
$L(t)$	$L(t) = \zeta(t) / [1 + \hat{\varphi}^T(t)\zeta(t)] \in \mathbf{R}^n$	$2n$	$n$
	$\zeta(t) := P(t-1)\hat{\varphi}(t) \in \mathbf{R}^n$	$n^2$	$n(n-1)$
$P(t)$	$P(t) = P(t-1) - L(t)\zeta^T(t) \in \mathbf{R}^{n \times n}$	$n^2$	$n^2$
$x_a(t)$	$x_a(t) = \hat{\varphi}_s^T(t)\hat{\theta}_s(t) \in \mathbf{R}$	$n_a + n_b$	$n_a + n_b - 1$
$\hat{w}(t)$	$\hat{w}(t) = y(t) - x_a(t) \in \mathbf{R}$	0	1
$\hat{v}(t)$	$\hat{v}(t) = \hat{w}(t) - \hat{\varphi}_n^T(t)\hat{\theta}_n(t) \in \mathbf{R}$	$n_c + n_d$	$n_c + n_d$
	总数	$2n^2 + 5n$	$2n^2 + 3n$
	总 flop 数	$N_2 := 4(n_a + n_b + n_c + n_d)^2 + 8(n_a + n_b + n_c + n_d)$	



**注9** 辨识 Box-Jenkins 系统的一些方法包括辅助模型多新息随机梯度算法<sup>[32]</sup>、辅助模型最小二乘迭代算法<sup>[69]</sup>和辅助模型梯度迭代算法<sup>[70]</sup>、非均匀采样数据系统的辅助模型广义增广随机梯度算法和辅助模型多新息广义增广随机梯度算法<sup>[43]</sup>。

### 3 OEAR 系统的滤波多新息辨识方法

文献[24]研究了方程误差自回归滑动平均系统(即 CARARMA 系统)的滤波辨识方法,本节研究输出误差自回归(OEAR)系统的基于输入输出数据滤波辨识方法,包括基于滤波的辅助模型广义随机梯度(F-AM-GSG)辨识方法、基于滤波的辅助模型多新息广义随机梯度(F-AM-MI-GSG)辨识方法、基于滤波的辅助模型递推广义最小二乘(F-AM-RGLS)辨识方法、基于滤波的辅助模型多新息广义最小二乘(F-AM-MI-GLS)辨识方法。最近,笔者提出了 OEAR 系统的辅助模型递推广义最小二乘方法、基于数据滤波的辅助模型递推最小二乘方法[71],王冬青研究了输出误差滑动平均(OEMA)系统的基于数据滤波的辅助模型递推最小二乘辨识方法和基于数据滤波的辅助模型最小二乘迭代辨识方法<sup>[72]</sup>。

#### 3.1 OEAR 系统描述与辨识模型

考虑输出误差自回归模型(Output-Error Autoregressive model, OEAR 模型)描述的动态随机系统:

$$y(t) = \frac{B(z)}{A(z)}u(t) + \frac{1}{C(z)}v(t), \quad (146)$$

其中 $\{u(t)\}$ 和 $\{y(t)\}$ 分别为系统的输入和输出序列, $\{v(t)\}$ 为零均值、不相关随机白噪声序列(不可测), $A(z)$ , $B(z)$ 和 $C(z)$ 均为单位后移算子 $z^{-1}$ 的多项式:

$$A(z) = 1 + a_1z^{-1} + a_2z^{-2} + \dots + a_{n_a}z^{-n_a},$$

$$B(z) = b_1z^{-1} + b_2z^{-2} + \dots + b_{n_b}z^{-n_b},$$

$$C(z) = 1 + c_1z^{-1} + c_2z^{-2} + \dots + c_{n_c}z^{-n_c}.$$

设阶次 $n_a$ , $n_b$ 和 $n_c$ 已知,记 $n := n_a + n_b + n_c$ ,且 $t \leq 0$ 时, $y(t) = 0, u(t) = 0, v(t) = 0$ 。

定义系统真实输出(即无噪输出) $x(t)$ 和噪声模型输出 $w(t)$ 分别为

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

$$w(t) := \frac{1}{C(z)}v(t) \in \mathbf{R}.$$

定义系统模型参数向量 $\theta$ 和噪声模型参数向量

$c$ ,系统模型信息向量 $\varphi(t)$ 和噪声模型信息向量 $\psi(t)$ 如下:

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

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

$$\varphi(t) := [-x(t-1), -x(t-2), \dots, -x(t-n_a),$$

$$u(t-1), u(t-2), \dots, u(t-n_b)]^T \in \mathbf{R}^{n_a+n_b},$$

$$\psi(t) := [-w(t-1), -w(t-2), \dots, -w(t-n_c)]^T \in \mathbf{R}^{n_c}.$$

系统的量测输入输出数据为 $\{u(t), y(t)\}$ ,系统的真实输出,即不可测无噪输出 $x(t)$ 是未知的,噪声模型的输出 $w(t)$ 是不可测的相关噪声或称为有色噪声, $y(t)$ 是 $x(t)$ 的含噪量测, $\varphi(t)$ 和 $\psi(t)$ 也是未知的。

由式(147)~(148)可得

$$x(t) = [1-A(z)]x(t) + B(z)u(t) = \varphi^T(t)\theta, \quad (149)$$

$$w(t) = [1-C(z)]w(t) + v(t) = \psi^T(t)c + v(t), \quad (150)$$

将式(147)和(148)代入式(146)可得

$$y(t) = x(t) + w(t) \quad (151)$$

$$= \varphi^T(t)\theta + \psi^T(t)c + v(t). \quad (152)$$

取滤波器 $C(z)$ ,定义滤波输入 $u_f(t)$ ,滤波输出 $y_f(t)$ ,滤波无噪中间变量 $x_f(t)$ ,滤波信息向量 $\varphi_f(t)$ 分别为

$$u_f(t) := C(z)u(t),$$

$$y_f(t) := C(z)y(t),$$

$$x_f(t) := C(z)x(t) = C(z)\frac{B(z)}{A(z)}u(t) = \frac{B(z)}{A(z)}u_f(t) =$$

$$[1-A(z)]x_f(t) + B(z)u_f(t) = \varphi_f^T(t)\theta,$$

$$\varphi_f(t) := C(z)\varphi(t) =$$

$$[-x_f(t-1), -x_f(t-2), \dots, -x_f(t-n_a),$$

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

式(146)两边乘以 $C(z)$ 得到

$$C(z)y(t) = \frac{B(z)}{A(z)}C(z)u(t) + v(t), \quad (153)$$

或

$$y_f(t) = \frac{B(z)}{A(z)}u_f(t) + v(t) = x_f(t) + v(t) =$$

$$\varphi_f^T(t)\theta + v(t). \quad (154)$$

这个滤波模型是一个白噪声干扰的输出误差模型(Output-Error model, OE 模型)。

#### 3.2 基于滤波的辅助模型广义随机梯度算法

对于 OEAR 系统的辨识模型(154)和(150),定义两个梯度准则函数:

$$J_7(\theta) := \frac{1}{2}[y_f(t) - \varphi_f^T(t)\theta]^2,$$

$$J_8(\mathbf{c}) := \frac{1}{2} [w(t) - \boldsymbol{\psi}^T(t) \mathbf{c}]^2. \quad (155)$$

设  $\hat{\boldsymbol{\theta}}(t)$  和  $\hat{\mathbf{c}}(t)$  分别是  $\boldsymbol{\theta}$  和  $\mathbf{c}$  在时刻  $t$  的估计,  $1/r_s(t)$  和  $1/r_n(t)$  为收敛因子或步长. 使用负梯度搜索, 极小化  $J_7(\boldsymbol{\theta})$  和  $J_8(\mathbf{c})$ , 可以得到下列递推关系:

$$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t-1) - \frac{1}{r_s(t)} \text{grad}[J_7(\hat{\boldsymbol{\theta}}(t-1))] = \hat{\boldsymbol{\theta}}(t-1) + \frac{\boldsymbol{\varphi}_f(t)}{r_s(t)} [y_f(t) - \boldsymbol{\varphi}_f^T(t) \hat{\boldsymbol{\theta}}(t-1)], \quad (156)$$

$$r_s(t) = r_s(t-1) + \|\boldsymbol{\varphi}_f(t)\|^2, \quad (157)$$

$$\mathbf{c}(t) = \hat{\mathbf{c}}(t-1) - \frac{1}{r_n(t)} \text{grad}[J_8(\hat{\mathbf{c}}(t-1))] = \hat{\mathbf{c}}(t-1) + \frac{\boldsymbol{\psi}(t)}{r_n(t)} [w(t) - \boldsymbol{\psi}^T(t) \hat{\mathbf{c}}(t-1)] = \hat{\mathbf{c}}(t-1) + \frac{\boldsymbol{\psi}(t)}{r_n(t)} [y(t) - \boldsymbol{\varphi}^T(t) \boldsymbol{\theta} - \boldsymbol{\psi}^T(t) \hat{\mathbf{c}}(t-1)], \quad (158)$$

$$r_n(t) = r_n(t-1) + \|\boldsymbol{\psi}(t)\|^2. \quad (159)$$

这些递推关系无法实现, 因为式 (156) — (159) 右边的  $y_f(t)$ ,  $\boldsymbol{\varphi}_f(t)$ ,  $\boldsymbol{\varphi}(t)$ ,  $\boldsymbol{\theta}$  和  $\boldsymbol{\psi}(t)$  都是未知的, 所以需要先构造它们的估计.

参照前文的辅助模型设计方法, 构造一个结构与  $G(z) = \frac{B(z)}{A(z)}$  相同的辅助模型, 用辅助模型的输出  $x_a(t-i)$  和系统输入  $u(t-i)$  定义  $\boldsymbol{\varphi}(t)$  的估计

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

用滤波辅助模型的输出  $x_{af}(t-i)$  和滤波输入  $u_f(t-i)$  定义  $\boldsymbol{\varphi}_f(t)$  的估计

$$\hat{\boldsymbol{\varphi}}_f(t) := [-x_{af}(t-1), -x_{af}(t-2), \dots, -x_{af}(t-n_a), u_f(t-1), u_f(t-2), \dots, u_f(t-n_b)]^T \in \mathbf{R}^{n_a+n_b},$$

用辅助模型的输出  $\hat{w}(t-i)$  (即  $w(t-i)$ ) 的估计定义  $\boldsymbol{\psi}(t)$  的估计

$$\hat{\boldsymbol{\psi}}(t) := [-\hat{w}(t-1), -\hat{w}(t-2), \dots, -\hat{w}(t-n_c)]^T \in \mathbf{R}^{n_c}.$$

根据式 (149), 取估算未知量  $x(t)$  的辅助模型为  $x_a(t) := \hat{\boldsymbol{\varphi}}^T(t) \hat{\boldsymbol{\theta}}(t)$ , 根据式 (153), 取估算未知量  $x_f(t)$  的辅助模型为  $x_{af}(t) := \hat{\boldsymbol{\varphi}}_f^T(t) \hat{\boldsymbol{\theta}}(t)$ . 由式 (151) 可得  $w(t) = y(t) - x(t)$ , 上式中未知变量  $x(t)$  用辅助模型的输出  $x_a(t)$  代替, 便得到计算  $\hat{w}(t)$  的辅助模型

$$\hat{w}(t) := y(t) - x_a(t) = y(t) - \hat{\boldsymbol{\varphi}}^T(t) \hat{\boldsymbol{\theta}}(t).$$

由噪声模型参数估计

$$\mathbf{c}(t) = [\hat{c}_1(t), \hat{c}_2(t), \dots, \hat{c}_{n_c}(t)]^T \in \mathbf{R}^{n_c}$$

构造  $C(z)$  的估计:

$$\hat{C}(t, z) := 1 + \hat{c}_1(t)z^{-1} + \hat{c}_2(t)z^{-2} + \dots + \hat{c}_{n_c}(t)z^{-n_c}.$$

用动态估计的滤波器  $\hat{C}(t, z)$  对  $u(t)$ ,  $y(t)$  和  $\hat{\boldsymbol{\varphi}}(t)$  进行滤波, 分别得到  $\hat{u}_f(t)$ ,  $\hat{y}_f(t)$  和  $\hat{\boldsymbol{\varphi}}_f(t)$  的估计:

$$\hat{u}_f(t) := \hat{C}(t, z)u(t),$$

$$\hat{y}_f(t) := \hat{C}(t, z)y(t),$$

$$\hat{\boldsymbol{\varphi}}_f(t) := \hat{C}(t, z)\hat{\boldsymbol{\varphi}}(t),$$

$\hat{u}_f(t)$ ,  $\hat{y}_f(t)$  和  $\hat{\boldsymbol{\varphi}}_f(t)$  也可按照下列递推式计算:

$$\hat{u}_f(t) = [1 + \hat{c}_1(t)z^{-1} + \hat{c}_2(t)z^{-2} + \dots + \hat{c}_{n_c}(t)z^{-n_c}]u(t) = u(t) + [u(t-1), u(t-2), \dots, u(t-n_c)]\hat{\mathbf{c}}(t),$$

$$\hat{y}_f(t) = [1 + \hat{c}_1(t)z^{-1} + \hat{c}_2(t)z^{-2} + \dots + \hat{c}_{n_c}(t)z^{-n_c}]y(t) = y(t) + [y(t-1), y(t-2), \dots, y(t-n_c)]\hat{\mathbf{c}}(t),$$

$$\hat{\boldsymbol{\varphi}}_f(t) = [1 + \hat{c}_1(t)z^{-1} + \hat{c}_2(t)z^{-2} + \dots + \hat{c}_{n_c}(t)z^{-n_c}]\hat{\boldsymbol{\varphi}}(t) = \hat{\boldsymbol{\varphi}}(t) + [\hat{\boldsymbol{\varphi}}(t-1), \hat{\boldsymbol{\varphi}}(t-2), \dots, \hat{\boldsymbol{\varphi}}(t-n_c)]\hat{\mathbf{c}}(t).$$

于是, 式 (156) — (159) 右边未知的  $y_f(t)$ ,  $\boldsymbol{\varphi}_f(t)$ ,  $\boldsymbol{\varphi}(t)$ ,  $\boldsymbol{\theta}$  和  $\boldsymbol{\psi}(t)$  分别用其估计  $\hat{y}_f(t)$ ,  $\hat{\boldsymbol{\varphi}}_f(t)$ ,  $\hat{\boldsymbol{\varphi}}(t)$ ,  $\hat{\boldsymbol{\theta}}(t-1)$  和  $\hat{\boldsymbol{\psi}}(t)$  代替, 可以得到估计 OEAR 系统参数向量  $\boldsymbol{\theta}$  和  $\mathbf{c}$  的基于滤波的辅助模型广义随机梯度算法 (F-AM-GSG 算法):

$$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t-1) + \frac{\hat{\boldsymbol{\varphi}}_f(t)}{r_s(t)} [\hat{y}_f(t) - \hat{\boldsymbol{\varphi}}_f^T(t) \hat{\boldsymbol{\theta}}(t-1)], \quad (160)$$

$$r_s(t) = r_s(t-1) + \|\hat{\boldsymbol{\varphi}}_f(t)\|^2, \quad (161)$$

$$\hat{\mathbf{c}}(t) = \hat{\mathbf{c}}(t-1) + \frac{\hat{\boldsymbol{\psi}}(t)}{r_n(t)} [y(t) - \hat{\boldsymbol{\varphi}}^T(t) \hat{\boldsymbol{\theta}}(t-1) - \hat{\boldsymbol{\psi}}^T(t) \hat{\mathbf{c}}(t-1)], \quad (162)$$

$$r_n(t) = r_n(t-1) + \|\hat{\boldsymbol{\psi}}(t)\|^2, \quad (163)$$

$$\hat{\boldsymbol{\varphi}}(t) = [-x_a(t-1), \dots, -x_a(t-n_a), u(t-1), u(t-2), \dots, u(t-n_b)]^T, \quad (164)$$

$$\hat{\boldsymbol{\psi}}(t) = [-\hat{w}(t-1), -\hat{w}(t-2), \dots, -\hat{w}(t-n_c)]^T, \quad (165)$$

$$\hat{\boldsymbol{\varphi}}_f(t) = [-x_{af}(t-1), \dots, -x_{af}(t-n_a), u_f(t-1), \dots, u_f(t-n_b)]^T, \quad (166)$$

$$\hat{u}_f(t) = u(t) + \hat{c}_1(t)u(t-1) + \hat{c}_2(t)u(t-2) + \dots + \hat{c}_{n_c}(t)u(t-n_c), \quad (167)$$

$$\hat{y}_f(t) = y(t) + \hat{c}_1(t)y(t-1) + \hat{c}_2(t)y(t-2) + \dots + \hat{c}_{n_c}(t)y(t-n_c), \quad (168)$$

$$x_a(t) = \hat{\boldsymbol{\varphi}}^T(t) \hat{\boldsymbol{\theta}}(t), \quad (169)$$

$$x_{af}(t) = \hat{\boldsymbol{\varphi}}_f^T(t) \hat{\boldsymbol{\theta}}(t), \quad (170)$$

$$\hat{w}(t) = y(t) - x_a(t), \quad (171)$$

$$\hat{\mathbf{c}}(t) = [\hat{c}_1(t), \hat{c}_2(t), \dots, \hat{c}_{n_c}(t)]^T, \quad (172)$$

$$\hat{\boldsymbol{\theta}}(t) = \begin{bmatrix} \hat{\boldsymbol{\theta}}(t) \\ \hat{\mathbf{c}}(t) \end{bmatrix}. \quad (173)$$

F-AM-GSG 算法 (160) — (173) 随  $t$  增加, 计算参数估计向量  $\hat{\theta}(t)$  和  $\hat{c}(t)$  的步骤如下:

1) 初始化: 令  $t = 1$ . 置初值  $\hat{\theta}(0) = \mathbf{1}_{n_a+n_b}/p_0$ ,  $\hat{c}(0) = \mathbf{1}_{n_c}/p_0$ ,  $r_s(0) = 1$ ,  $r_n(0) = 1$ ,  $x_a(t-i) = 1/p_0$ ,  $x_{af}(t-i) = 1/p_0$ ,  $\hat{w}(t-i) = 1/p_0$ ,  $\hat{y}_f(t-i) = 1/p_0$ ,  $\hat{u}_f(t-i) = 1/p_0$ ,  $i = 1, 2, \dots, \max[n_a, n_b, n_c]$ ,  $p_0 = 10^6$ . 给定小正数  $\varepsilon$ .

2) 采集输入输出数据  $u(t)$  和  $y(t)$ , 由式 (164) — (166) 构造信息向量  $\hat{\varphi}(t)$ ,  $\hat{\psi}(t)$  和  $\hat{\varphi}_f(t)$ .

3) 由式 (163) 计算  $r_n(t)$ .

4) 根据式 (162) 刷新参数估计向量  $\hat{c}(t)$ .

5) 从式 (172) 中  $\hat{c}(t)$  读取  $\hat{c}_i(t)$ . 由式 (167) — (168) 计算  $\hat{u}_f(t)$  和  $\hat{y}_f(t)$ .

6) 由式 (161) 计算  $r_s(t)$ .

7) 根据式 (160) 刷新参数估计向量  $\hat{\theta}(t)$ .

8) 由式 (169) — (171) 计算辅助模型的输出  $x_a(t)$ ,  $x_{af}(t)$  和  $\hat{w}(t)$ .

9) 将  $\hat{\theta}(t)$  与  $\hat{\theta}(t-1)$  进行比较, 将  $\hat{c}(t)$  与  $\hat{c}(t-1)$  进行比较, 如果它们满足  $\|\hat{\theta}(t) - \hat{\theta}(t-1)\| < \varepsilon$  和  $\|\hat{c}(t) - \hat{c}(t-1)\| < \varepsilon$ , 则终止递推计算过程, 得到满意精度的参数估计  $\hat{\theta}(t)$  和  $\hat{c}(t)$ ; 否则  $t$  增 1 转到第 2) 步, 进行递推计算.

**注 10** 细心的读者应该能看出式 (167) — (168) 应该修改为

$$\begin{aligned} \hat{u}_f(j) &= u(j) + \hat{c}_1(t)u(j-1) + \hat{c}_2(t)u(j-2) + \dots + \\ &\quad \hat{c}_{n_c}(t)u(j-n_c), \quad j = 1, 2, \dots, t, \\ \hat{y}_f(j) &= y(j) + \hat{c}_1(t)y(j-1) + \hat{c}_2(t)y(j-2) + \dots + \\ &\quad \hat{c}_{n_c}(t)y(j-n_c). \end{aligned}$$

### 3.3 基于滤波的辅助模型多新息广义随机梯度算法

下面借助于多新息辨识理论, 通过扩展新息维数, 推导出基于滤波的多新息广义随机梯度算法, 以改进算法的收敛速度, 设正整数  $p$  表示新息长度. 基于 F-AM-GSG 算法 (160) — (173), 将滤波输出  $\hat{y}_f(t)$ , 滤波信息向量  $\hat{\varphi}_f(t)$ , 系统输出  $y(t)$ , 信息向量  $\hat{\varphi}(t)$  和噪声信息向量  $\hat{\psi}(t)$  分别扩展为堆积滤波输出向量  $\hat{Y}_f(p, t)$ , 堆积滤波信息矩阵  $\hat{\Phi}_f(p, t)$ , 堆积输出向量  $Y(p, t)$ , 堆积信息矩阵  $\hat{\Phi}(p, t)$  和堆积噪声信息矩阵  $\hat{\Psi}(p, t)$ :

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

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

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

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

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

将式 (160) 和 (162) 中量新息

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

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

扩展为新息向量

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

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

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

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

可以总结出新息长度为  $p$  的, 估计 OEAR 系统参数向量  $\theta$  和  $c$  的基于滤波的辅助模型多新息广义随机梯度算法 (F-AM-MI-GSG 算法):

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\hat{\Phi}_f(p, t)}{r_s(t)} E_s(p, t), \quad (174)$$

$$E_s(p, t) = \hat{Y}_f(p, t) - \hat{\Phi}_f^T(p, t) \hat{\theta}(t-1), \quad (175)$$

$$r_s(t) = r_s(t-1) + \|\hat{\varphi}_f(t)\|^2, \quad (176)$$

$$\hat{c}(t) = \hat{c}(t-1) + \frac{\hat{\Psi}(p, t)}{r_n(t)} E_n(p, t), \quad (177)$$

$$E_n(p, t) = Y(p, t) - \hat{\Phi}^T(p, t) \hat{\theta}(t-1) - \hat{\Psi}^T(p, t) \hat{c}(t-1), \quad (178)$$

$$r_n(t) = r_n(t-1) + \|\hat{\psi}(t)\|^2, \quad (179)$$

$$\hat{Y}_f(p, t) = [\hat{y}_f(t), \hat{y}_f(t-1), \dots, \hat{y}_f(t-p+1)]^T, \quad (180)$$

$$\hat{\Phi}_f(p, t) = [\hat{\varphi}_f(t), \hat{\varphi}_f(t-1), \dots, \hat{\varphi}_f(t-p+1)], \quad (181)$$

$$Y(p, t) = [y(t), y(t-1), \dots, y(t-p+1)]^T, \quad (182)$$

$$\hat{\Phi}(p, t) = [\hat{\varphi}(t), \hat{\varphi}(t-1), \dots, \hat{\varphi}(t-p+1)], \quad (183)$$

$$\hat{\Psi}(p, t) = [\hat{\psi}(t), \hat{\psi}(t-1), \dots, \hat{\psi}(t-p+1)], \quad (184)$$

$$\hat{\varphi}(t) = [-x_a(t-1), \dots, -x_a(t-n_a), u(t-1), \dots, u(t-n_b)]^T, \quad (185)$$

$$\hat{\psi}(t) = [-\hat{w}(t-1), -\hat{w}(t-2), \dots, -\hat{w}(t-n_c)], \quad (186)$$

$$\hat{\varphi}_f(t) = [-x_{af}(t-1), \dots, -x_{af}(t-n_a), u_f(t-1), \dots, u_f(t-n_b)]^T, \quad (187)$$

$$\hat{u}_f(t) = u(t) + [u(t-1), u(t-2), \dots, u(t-n_c)] \hat{c}(t), \quad (188)$$

$$\hat{y}_f(t) = y(t) + [y(t-1), y(t-2), \dots, y(t-n_c)] \hat{c}(t), \quad (189)$$

$$x_a(t) = \hat{\varphi}^T(t) \hat{\theta}(t), \quad (190)$$

$$x_{af}(t) = \hat{\varphi}_f^T(t) \hat{\theta}(t), \quad (191)$$

$$\hat{w}(t) = y(t) - x_a(t), \quad (192)$$

$$\hat{\theta}(t) = \begin{bmatrix} \hat{\theta}(t) \\ \hat{c}(t) \end{bmatrix}. \quad (193)$$

F-AM-MI-GSG 算法(174)—(193)随  $t$  增加, 计算参数估计向量  $\hat{\theta}(t)$  和  $\hat{c}(t)$  的步骤如下:

1) 初始化: 令  $t = 1$ . 置初值  $\hat{\theta}(0) = \mathbf{1}_{n_a+n_b}/p_0$ ,  $\hat{c}(0) = \mathbf{1}_{n_c}/p_0$ ,  $r_s(0) = 1$ ,  $r_n(0) = 1$ ,  $x_a(t-i) = 1/p_0$ ,  $x_{af}(t-i) = 1/p_0$ ,  $\hat{w}(t-i) = 1/p_0$ ,  $\hat{y}_f(t-i) = 1/p_0$ ,  $\hat{u}_f(t-i) = 1/p_0$ ,  $i = 1, 2, \dots, \max[n_a, n_b, n_c]$ ,  $p_0 = 10^6$ . 给定小正数  $\varepsilon$ .

2) 采集输入输出数据  $u(t)$  和  $y(t)$ , 由式(185)—(187)构造信息向量  $\hat{\varphi}(t)$ ,  $\hat{\psi}(t)$  和  $\hat{\varphi}_f(t)$ .

3) 由式(182)—(184)构造堆积输出向量  $\mathbf{Y}(p, t)$ , 堆积信息矩阵  $\hat{\Phi}(p, t)$  和堆积噪声信息矩阵  $\hat{\Psi}(p, t)$ .

4) 由式(178)计算新息向量  $\mathbf{E}_n(p, t)$ , 由式(179)计算  $r_n(t)$ .

5) 根据式(177)刷新参数估计向量  $\hat{c}(t)$ .

6) 由式(188)—(189)计算  $\hat{u}_f(t)$  和  $\hat{y}_f(t)$ .

7) 由式(180)—(181)构造堆积滤波输出向量  $\hat{\mathbf{Y}}_f(p, t)$  和堆积滤波信息矩阵  $\hat{\Phi}_f(p, t)$ .

8) 由式(175)计算新息向量  $\mathbf{E}_s(p, t)$ , 由式(176)计算  $r_s(t)$ .

9) 根据式(174)刷新参数估计向量  $\hat{\theta}(t)$ .

10) 由式(190)—(192)计算辅助模型的输出  $x_a(t)$ ,  $x_{af}(t)$  和  $\hat{w}(t)$ .

11) 如果参数估计差满足  $\|\hat{\theta}(t) - \hat{\theta}(t-1)\| < \varepsilon$  和  $\|\hat{c}(t) - \hat{c}(t-1)\| < \varepsilon$ , 则终止递推计算过程, 得到满意精度的参数估计  $\hat{\theta}(t)$  和  $\hat{c}(t)$ ; 否则  $t$  增 1 转到第 2) 步, 进行递推计算.

### 3.4 基于滤波的辅助模型递推广义最小二乘算法

对于 OEAR 系统的辨识模型(154)和(150), 定义两个最小二乘准则函数:

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

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

参照文献[1-2]中 RLS 算法的推导, 极小化  $J_9(\theta)$  和  $J_{10}(c)$ , 可以得到下列最小二乘递推关系:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \mathbf{L}_s(t) [y_f(t) - \varphi_f^T(t)\hat{\theta}(t-1)], \quad (194)$$

$$\mathbf{L}_s(t) = \mathbf{P}_s(t-1)\varphi_f(t) [1 + \varphi_f^T(t)\mathbf{P}_s(t-1)\varphi_f(t)]^{-1}, \quad (195)$$

$$\mathbf{P}_s(t) = [\mathbf{I}_{n_a+n_b} - \mathbf{L}_s(t)\varphi_f^T(t)]\mathbf{P}_s(t-1), \quad (196)$$

$$\hat{c}(t) = \hat{c}(t-1) + \mathbf{L}_n(t) [w(t) - \psi^T(t)\hat{c}(t-1)] = \hat{c}(t-1) + \mathbf{L}_n(t) [y(t) - \varphi^T(t)\theta - \psi^T(t)\hat{c}(t-1)], \quad (197)$$

$$\mathbf{L}_n(t) = \mathbf{P}_n(t-1)\psi(t) [1 + \psi^T(t)\mathbf{P}_n(t-1)\psi(t)]^{-1}, \quad (198)$$

$$\mathbf{P}_n(t) = [\mathbf{I}_{n_c} - \mathbf{L}_n(t)\psi^T(t)]\mathbf{P}_n(t-1). \quad (199)$$

这些递推关系无法实现, 因为式(194)—(199)右边的  $y_f(t)$ ,  $\varphi_f(t)$ ,  $\varphi(t)$ ,  $\theta$  和  $\psi(t)$  都是未知的, 同样在辨识算法中使用它们的估计  $\hat{y}_f(t)$ ,  $\hat{\varphi}_f(t)$ ,  $\hat{\varphi}(t)$ ,  $\hat{\theta}(t-1)$  和  $\hat{\psi}(t)$  代替, 可以得到估计 OEAR 系统参数向量  $\theta$  和  $c$  的基于滤波的辅助模型递推广义最小二乘算法(F-AM-RGLS 算法):

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \mathbf{L}_s(t) [\hat{y}_f(t) - \hat{\varphi}_f^T(t)\hat{\theta}(t-1)], \quad (200)$$

$$\mathbf{L}_s(t) = \mathbf{P}_s(t-1)\hat{\varphi}_f(t) [1 + \hat{\varphi}_f^T(t)\mathbf{P}_s(t-1)\hat{\varphi}_f(t)]^{-1}, \quad (201)$$

$$\mathbf{P}_s(t) = [\mathbf{I}_{n_a+n_b} - \mathbf{L}_s(t)\hat{\varphi}_f^T(t)]\mathbf{P}_s(t-1), \quad (202)$$

$$\hat{c}(t) = \hat{c}(t-1) + \mathbf{L}_n(t) [y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1) - \hat{\psi}^T(t)\hat{c}(t-1)], \quad (203)$$

$$\mathbf{L}_n(t) = \mathbf{P}_n(t-1)\hat{\psi}(t) [1 + \hat{\psi}^T(t)\mathbf{P}_n(t-1)\hat{\psi}(t)]^{-1}, \quad (204)$$

$$\mathbf{P}_n(t) = [\mathbf{I}_{n_c} - \mathbf{L}_n(t)\hat{\psi}^T(t)]\mathbf{P}_n(t-1), \quad (205)$$

$$\hat{\varphi}(t) = [-x_a(t-1), \dots, -x_a(t-n_a), u(t-1), \dots, u(t-n_b)]^T, \quad (206)$$

$$\hat{\psi}(t) = [-\hat{w}(t-1), -\hat{w}(t-2), \dots, -\hat{w}(t-n_c)]^T, \quad (207)$$

$$\hat{y}_f(t) = y(t) + [y(t-1), y(t-2), \dots, y(t-n_c)]\hat{c}(t), \quad (208)$$

$$\hat{\varphi}_f(t) = \hat{\varphi}(t) + [\hat{\varphi}(t-1), \hat{\varphi}(t-2), \dots, \hat{\varphi}(t-n_c)]\hat{c}(t), \quad (209)$$

$$x_a(t) = \hat{\varphi}^T(t)\hat{\theta}(t), \quad (210)$$

$$\hat{w}(t) = y(t) - x_a(t), \quad (211)$$

$$\hat{\theta}(t) = \begin{bmatrix} \hat{\theta}(t) \\ \hat{c}(t) \end{bmatrix}. \quad (212)$$

F-AM-RGLS 算法(200)—(212)随  $t$  增加, 计算参数估计向量  $\hat{\theta}(t)$  和  $\hat{c}(t)$  的步骤如下:

1) 初始化: 令  $t = 1$ . 置初值  $\hat{\theta}(0) = \mathbf{1}_{n_a+n_b}/p_0$ ,  $\hat{c}(0) = \mathbf{1}_{n_c}/p_0$ ,  $\mathbf{P}_s(0) = p_0\mathbf{I}_{n_a+n_b}$ ,  $\mathbf{P}_n(0) = p_0\mathbf{I}_{n_c}$ ,  $x_a(t-i) = 1/p_0$ ,  $\hat{w}(t-i) = 1/p_0$ ,  $\hat{y}_f(t-i) = 1/p_0$ ,  $\hat{u}_f(t-i) = 1/p_0$ ,  $\hat{\varphi}_f(t-i) = \mathbf{1}_{n_a+n_b}/p_0$ ,  $i = 1, 2, \dots, \max[n_a, n_b, n_c]$ ,  $p_0 = 10^6$ . 给定小正数  $\varepsilon$ .

2) 采集输入输出数据  $u(t)$  和  $y(t)$ , 由式(206)—(207)构造信息向量  $\hat{\varphi}(t)$  和  $\hat{\psi}(t)$ .

3) 由式(204)—(205)计算增益向量  $\mathbf{L}_n(t)$  和协方差阵  $\mathbf{P}_n(t)$ .

4) 根据式(203)刷新参数估计向量  $\hat{c}(t)$ .

5) 由式(208)—(209)计算滤波输出  $\hat{y}_f(t)$  和滤



波信息向量  $\hat{\varphi}_f(t)$ .

6) 由式(201)—(202)计算增益向量  $L_s(t)$  和协方差阵  $P_s(t)$ .

7) 根据式(200)刷新参数估计向量  $\hat{\theta}(t)$ .

8) 由式(210)—(211)计算辅助模型的输出  $x_a(t)$  和  $\hat{w}(t)$ .

9) 如果参数估计差满足  $\|\hat{\vartheta}(t) - \hat{\vartheta}(t-1)\| < \varepsilon$ , 则终止递推计算过程, 得到满意精度的参数估计  $\hat{\theta}(t)$  和  $\hat{c}(t)$ ; 否则  $t$  增 1 转到第 2) 步, 进行递推计算.

### 3.5 基于滤波的辅助模型多新息广义最小二乘算法

基于 F-AM-RGLS 算法(200)—(212), 将滤波输出  $\hat{y}_f(t)$ , 滤波信息向量  $\hat{\varphi}_f(t)$ , 系统输出  $y(t)$ , 信息向量  $\hat{\varphi}(t)$  和噪声信息向量  $\hat{\psi}(t)$  分别扩展为堆积滤波输出向量  $\hat{Y}_f(p, t)$ , 堆积滤波信息矩阵  $\hat{\Phi}_f(p, t)$ , 堆积输出向量  $Y(p, t)$ , 堆积信息矩阵  $\hat{\Phi}(p, t)$  和堆积噪声信息矩阵  $\hat{\Psi}(p, t)$ :

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

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

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

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

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

将式(200)和(203)中标量新息

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

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

扩展为新息向量

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

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

可以得到估计 OEAR 系统参数向量  $\theta$  和  $c$  的基于滤波的辅助模型多新息广义最小二乘算法(F-AM-MI-GLS 算法):

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L_s(t) E_s(p, t), \quad (213)$$

$$E_s(p, t) = \hat{Y}_f(p, t) - \hat{\Phi}_f^T(p, t) \hat{\theta}(t-1), \quad (214)$$

$$L_s(t) = P_s(t-1) \hat{\Phi}_f(t) [I_p + \hat{\Phi}_f^T(t) P_s(t-1) \hat{\Phi}_f(t)]^{-1}, \quad (215)$$

$$P_s(t) = [I_{n_a+n_b} - L_s(t) \hat{\Phi}_f^T(t)] P_s(t-1), \quad (216)$$

$$\hat{c}(t) = \hat{c}(t-1) + L_n(t) E_n(p, t), \quad (217)$$

$$E_n(p, t) = Y(p, t) - \hat{\Phi}^T(p, t) \hat{\theta}(t-1) - \hat{\Psi}^T(p, t) \hat{c}(t-1), \quad (218)$$

$$L_n(t) = P_n(t-1) \hat{\Psi}(t) [I_p + \hat{\Psi}^T(t) P_n(t-1) \hat{\Psi}(t)]^{-1}, \quad (219)$$

$$P_n(t) = [I_{n_c} - L_n(t) \hat{\Psi}^T(t)] P_n(t-1), \quad (220)$$

$$\hat{Y}_f(p, t) = [\hat{y}_f(t), \hat{y}_f(t-1), \dots, \hat{y}_f(t-p+1)]^T, \quad (221)$$

$$\hat{\Phi}_f(p, t) = [\hat{\varphi}_f(t), \hat{\varphi}_f(t-1), \dots, \hat{\varphi}_f(t-p+1)], \quad (222)$$

$$Y(p, t) = [y(t), y(t-1), \dots, y(t-p+1)]^T, \quad (223)$$

$$\hat{\Phi}(p, t) = [\hat{\varphi}(t), \hat{\varphi}(t-1), \dots, \hat{\varphi}(t-p+1)], \quad (224)$$

$$\hat{\Psi}(p, t) = [\hat{\psi}(t), \hat{\psi}(t-1), \dots, \hat{\psi}(t-p+1)], \quad (225)$$

$$\hat{\varphi}(t) = [-x_a(t-1), \dots, -x_a(t-n_a), u(t-1), \dots, u(t-n_b)]^T, \quad (226)$$

$$\hat{\psi}(t) = [-\hat{w}(t-1), -\hat{w}(t-2), \dots, -\hat{w}(t-n_c)]^T, \quad (227)$$

$$\hat{y}_f(t) = y(t) + [y(t-1), y(t-2), \dots, y(t-n_c)] \hat{c}(t), \quad (228)$$

$$\hat{\varphi}_f(t) = \hat{\varphi}(t) + [\hat{\varphi}(t-1), \hat{\varphi}(t-2), \dots, \hat{\varphi}(t-n_c)] \hat{c}(t), \quad (229)$$

$$x_a(t) = \hat{\varphi}^T(t) \hat{\theta}(t), \quad (230)$$

$$\hat{w}(t) = y(t) - x_a(t), \quad (231)$$

$$\hat{\vartheta}(t) = \begin{bmatrix} \hat{\theta}(t) \\ \hat{c}(t) \end{bmatrix}. \quad (232)$$

F-AM-MI-GLS 算法(213)—(232)随  $t$  增加, 计算参数估计向量  $\hat{\theta}(t)$  和  $\hat{c}(t)$  的步骤如下:

1) 初始化: 令  $t=1$ , 给定新息长度  $p$ . 置  $\hat{\theta}(0) = \mathbf{1}_{n_a+n_b}/p_0$ ,  $\hat{c}(0) = \mathbf{1}_{n_c}/p_0$ , 和协方差阵初值  $P_s(0) = p_0 I_{n_a+n_b}$ ,  $P_n(0) = p_0 I_{n_c}$ . 置  $x_a(t-i) = 1/p_0$ ,  $\hat{w}(t-i) = 1/p_0$ ,  $\hat{y}_f(t-i) = 1/p_0$ ,  $\hat{u}_f(t-i) = 1/p_0$ ,  $\hat{\varphi}_f(t-i) = \mathbf{1}_{n_a+n_b}/p_0$ ,  $i=1, 2, \dots, \max[n_a, n_b, n_c]$ ,  $p_0 = 10^6$ . 给定小正数  $\varepsilon$ .

2) 采集输入输出数据  $u(t)$  和  $y(t)$ , 由式(226)—(227)构造信息向量  $\hat{\varphi}(t)$  和  $\hat{\psi}(t)$ .

3) 由式(223)—(225)构造堆积输出向量  $Y(p, t)$ , 堆积信息矩阵  $\hat{\Phi}(p, t)$  和堆积噪声信息矩阵  $\hat{\Psi}(p, t)$ .

4) 由式(218)计算新息向量  $E_n(p, t)$ , 由式(219)—(220)计算增益向量  $L_n(t)$  和协方差阵  $P_n(t)$ .

5) 根据式(217)刷新参数估计向量  $\hat{c}(t)$ .

6) 由式(228)—(229)计算滤波输出  $\hat{y}_f(t)$  和滤波信息向量  $\hat{\varphi}_f(t)$ .

7) 由式(221)—(222)构造堆积滤波输出向量  $\hat{Y}_f(p, t)$  和堆积滤波信息矩阵  $\hat{\Phi}_f(p, t)$ .

8) 由式(214)计算新息向量  $E_s(p, t)$ , 由式(215)—(216)计算增益向量  $L_s(t)$  和协方差阵  $P_s(t)$ .

9) 根据式(213)刷新参数估计向量  $\hat{\theta}(t)$ .

10) 由式(230)—(231)计算辅助模型的输出  $x_a(t)$  和  $\hat{w}(t)$ .

11) 如果参数估计差满足  $\|\hat{\vartheta}(t) - \hat{\vartheta}(t-1)\| < \varepsilon$ , 则终止递推计算过程, 得到满意精度的参数估计  $\hat{\theta}(t)$  和  $\hat{c}(t)$ ; 否则  $t$  增 1 转到第 2) 步, 进行递推计算.

## 4 结语

本文讨论了输出误差系统的辅助模型随机梯度 (AM-SG) 算法、辅助模型多新息随机梯度 (AM-MISG) 算法、辅助模型递推最小二乘 (AM-RLS) 算法、辅助模型多新息最小二乘 (AM-MILS) 算法, 研究了 Box-Jenkins 系统的辅助模型广义增广随机梯度 (AM-GESG) 算法、辅助模型多新息广义增广随机梯度 (AM-MI-GESG) 算法、辅助模型递推广义增广最小二乘 (AM-RGELS) 算法、辅助模型多新息广义增广最小二乘 (AM-MI-GELS) 算法, 研究了 OEAR 系统的基于滤波的辅助模型广义随机梯度 (F-AM-GSG) 算法、基于滤波的辅助模型多新息广义随机梯度 (F-AM-MI-GSG) 算法、基于滤波的辅助模型递推广义最小二乘 (F-AM-RGLS) 算法和基于滤波的辅助模型多新息广义最小二乘 (F-AM-MI-GLS) 算法. 这些方法可以推广到下列多输入单输出输出误差类系统、多输入多输出输出误差类系统、多变量输出误差类系统、多元输出误差类系统, 以及非线性输出误差类系统<sup>[1-2]</sup>.

1) 研究下列多输入 OEARMA 模型 (即多输入 Box-Jenkins 模型) 的辅助模型广义增广随机梯度算法、辅助模型多新息广义增广随机梯度算法、辅助模型递推广义增广随机梯度算法、辅助模型多新息广义增广最小二乘算法:

$$\textcircled{1} y(t) = \frac{1}{A(z)} \sum_{j=1}^r B_j(z) u_j(t) + \frac{D(z)}{C(z)} v(t),$$

$$\textcircled{2} y(t) = \sum_{j=1}^r \frac{B_j(z)}{F_j(z)} u_j(t) + \frac{D(z)}{C(z)} v(t),$$

$$B_j(z) := b_j(1)z^{-1} + b_j(2)z^{-2} + \dots + b_j(n_j)z^{-n_j},$$

$$F_j(z) := 1 + f_j(1)z^{-1} + f_j(2)z^{-2} + \dots + f_j(n_j)z^{-n_j}.$$

2) 对于多输入自回归输出误差 ARMA 模型 (AR-OEARMA 模型), 即多输入 AR-Box-Jenkins 模型描述的动态随机系统 (包括特殊的多输入 AR-OEMA 模型和多输入 AR-OEAR 模型):

$$\textcircled{1} A(z)y(t) = \frac{1}{F(z)} \sum_{j=1}^r B_j(z) u_j(t) + \frac{D(z)}{C(z)} v(t),$$

$$\textcircled{2} A(z)y(t) = \sum_{j=1}^r \frac{B_j(z)}{F_j(z)} u_j(t) + \frac{D(z)}{C(z)} v(t),$$

研究辅助模型广义增广随机梯度算法和辅助模型多新息广义增广随机梯度算法, 以及基于滤波的辅助模型广义增广随机梯度算法和基于滤波的辅助模型多新息广义增广随机梯度算法.

3) 研究下列非线性输出误差模型描述的非线性动态系统的辅助模型随机梯度辨识算法和辅助模

型多新息随机梯度辨识算法:

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

其中  $f(\ast)$  和  $g(\ast)$  为已知基函数. 当  $f(y) = y, g(u) = u$  时, 上式退化为线性输出误差系统. 其他非线性例子如下:

$$\textcircled{1} f(y(t)) = y^2(t), \quad g(u(t)) = \sin^2(u(t)).$$

$$\textcircled{2} f(y(t)) = |y(t)|, \quad g(u(t)) = \sqrt[3]{u^2(t)}.$$

4) 研究非线性输出误差 ARMA 模型描述的非线性动态系统的广义增广随机梯度 (GESG) 算法、多新息广义增广随机梯度 (MI-GESG) 算法、基于滤波的广义增广随机梯度 (F-GESG) 算法、基于滤波的多新息广义增广随机梯度 (F-MI-GESG) 算法, 以及基于分解的辨识算法和基于滤波的分解辨识算法:

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

其中  $f(\ast)$  和  $g(\ast)$  为已知基函数. 例子如下:

$$\textcircled{1} f(y(t)) = y^2(t), \quad g(u(t)) = \sin^2(u(t)).$$

$$\textcircled{2} f(y(t)) = |y(t)|, \quad g(u(t)) = \sqrt[3]{u^2(t)}.$$

5) 研究下列多变量输出误差 ARMA 模型 (即多变量 Box-Jenkins 模型) 描述的多变量系统的辅助模型广义增广随机梯度算法和辅助模型递推广义增广最小二乘算法<sup>[68]</sup>, 辅助模型多新息广义增广随机梯度算法和辅助模型多新息广义增广最小二乘算法:

$$\textcircled{1} \mathbf{y}(t) = \mathbf{A}^{-1}(z) \mathbf{B}(z) \mathbf{u}(t) + \frac{\mathbf{D}(z)}{\mathbf{C}(z)} \mathbf{v}(t),$$

$$\textcircled{2} \mathbf{y}(t) = \mathbf{A}^{-1}(z) \mathbf{B}(z) \mathbf{u}(t) + \mathbf{C}^{-1}(z) \mathbf{D}(z) \mathbf{v}(t),$$

其中  $\mathbf{u}(t) := [u_1(t), u_2(t), \dots, u_r(t)]^T \in \mathbf{R}^r$  为输入向量,  $\mathbf{y}(t) := [y_1(t), y_2(t), \dots, y_m(t)]^T \in \mathbf{R}^m$  为输出向量,  $\mathbf{v}(t) := [v_1(t), v_2(t), \dots, v_m(t)]^T \in \mathbf{R}^m$  为零均值白噪声向量,  $\mathbf{A}(z), \mathbf{B}(z), \mathbf{C}(z)$  和  $\mathbf{D}(z)$  是单位后移算子  $z^{-1}$  的多项式矩阵:

$$\mathbf{A}(z) := \mathbf{I} + \mathbf{A}_1 z^{-1} + \mathbf{A}_2 z^{-2} + \dots + \mathbf{A}_{n_a} z^{-n_a}, \quad \mathbf{A}_i \in \mathbf{R}^{m \times m},$$

$$\mathbf{B}(z) := \mathbf{B}_1 z^{-1} + \mathbf{B}_2 z^{-2} + \dots + \mathbf{B}_{n_b} z^{-n_b}, \quad \mathbf{B}_i \in \mathbf{R}^{m \times r},$$

$$\mathbf{C}(z) := \mathbf{I} + \mathbf{C}_1 z^{-1} + \mathbf{C}_2 z^{-2} + \dots + \mathbf{C}_{n_c} z^{-n_c}, \quad \mathbf{C}_i \in \mathbf{R}^{m \times m},$$

$$\mathbf{D}(z) := \mathbf{I} + \mathbf{D}_1 z^{-1} + \mathbf{D}_2 z^{-2} + \dots + \mathbf{D}_{n_d} z^{-n_d}, \quad \mathbf{D}_i \in \mathbf{R}^{m \times m}.$$

6) 对于多变量输出误差 ARMA 模型 (multivariable output-error ARMA model) 描述的多变量系统:

$$\textcircled{1} \mathbf{y}(t) = \mathbf{A}^{-1}(z) \mathbf{B}(z) \mathbf{u}(t) + \frac{\mathbf{D}(z)}{\mathbf{C}(z)} \mathbf{v}(t),$$

$$\textcircled{2} \mathbf{y}(t) = \mathbf{A}^{-1}(z) \mathbf{B}(z) \mathbf{u}(t) + \mathbf{C}^{-1}(z) \mathbf{D}(z) \mathbf{v}(t),$$

研究基于分解的辅助模型多新息广义增广随机梯度辨识算法和基于分解的辅助模型递推广义增广最小

二乘算法.

7) 对于多元输出误差 ARMA 模型 (multivariate output-error ARMA model) 描述的动态随机系统:

$$\begin{aligned} \textcircled{1} \mathbf{y}(t) &= \frac{\Phi(t)}{A(z)} \boldsymbol{\theta} + \frac{D(z)}{C(z)} \mathbf{v}(t), \\ \textcircled{2} \mathbf{y}(t) &= \frac{\Phi(t)}{A(z)} \boldsymbol{\theta} + C^{-1}(z) D(z) \mathbf{v}(t), \\ \textcircled{3} A(z) \mathbf{y}(t) &= \frac{\Phi(t)}{F(z)} \boldsymbol{\theta} + C^{-1}(z) D(z) \mathbf{v}(t), \end{aligned}$$

研究基于辅助模型的多新息广义增广随机梯度辨识算法和基于滤波的辅助模型多新息广义增广随机梯度辨识算法,其中  $\Phi(t) \in \mathbf{R}^{m \times n}$  是可测信息矩阵,  $\boldsymbol{\theta} \in \mathbf{R}^n$  为待辨识的参数向量.

8) 对于多元输出误差 ARMA 模型描述的动态随机系统:

$$\begin{aligned} \textcircled{1} \mathbf{y}(t) &= \frac{\Phi(t)}{A(z)} \boldsymbol{\theta} + \frac{D(z)}{C(z)} \mathbf{v}(t), \\ \textcircled{2} \mathbf{y}(t) &= \frac{\Phi(t)}{A(z)} \boldsymbol{\theta} + C^{-1}(z) D(z) \mathbf{v}(t), \\ \textcircled{3} A(z) \mathbf{y}(t) &= \frac{\Phi(t)}{F(z)} \boldsymbol{\theta} + C^{-1}(z) D(z) \mathbf{v}(t), \end{aligned}$$

研究基于分解递推广义增广最小二乘算法和基于分解的多新息广义增广最小二乘算法.

9) 研究下列多输入非线性自回归输出误差 ARMA 系统的辅助模型多新息广义增广随机梯度算法和辅助模型多新息广义增广最小二乘算法:

$$\begin{aligned} \textcircled{1} A(z) y(t) &= \sum_{j=1}^r \frac{B_j(z)}{F_j(z)} u_j^2(t) + \frac{D(z)}{C(z)} v(t). \\ \textcircled{2} A(z) y(t) &= \sum_{j=1}^r \frac{B_j(z)}{F_j(z)} u_j(t) u_{j+1}(t) + \frac{D(z)}{C(z)} v(t). \\ \textcircled{3} A(z) y(t) &= \sum_{j=1}^r \frac{B_j(z)}{F_j(z)} u_j(t) u_j(t-j) + \frac{D(z)}{C(z)} v(t). \\ \textcircled{4} A(z) y^2(t) &= \sum_{j=1}^r \frac{B_j(z)}{F_j(z)} u_j(t) u_{r-j+1}(t) + \frac{D(z)}{C(z)} v(t). \\ \textcircled{5} A(z) f(y(t)) &= \sum_{j=1}^r \frac{B_j(z)}{F_j(z)} g_j(u_j(t)) + \frac{D(z)}{C(z)} v(t). \\ \textcircled{6} A(z) f(y(t)) &= \sum_{j=1}^r \frac{B_j(z)}{F_j(z)} g(u_1(t), u_2(t), \dots, u_r(t)) + \frac{D(z)}{C(z)} v(t). \\ \textcircled{7} A(z) f(y(t)) &= \sum_{j=1}^r \frac{B_j(z)}{F_j(z)} g_j(u_1(t), u_2(t), \dots, u_r(t)) + \frac{D(z)}{C(z)} v(t). \end{aligned}$$

10) 研究多输入多输出输出误差 ARMA 模型 (多输入多输出 OEARMA 模型) 的辅助模型多新息广义增广辨识算法:

$$\textcircled{1} y_i(t) = \sum_{j=1}^m A_{ij}(z) y_j(t) + \frac{1}{F(z)} \sum_{j=1}^r B_{ij}(z) u_j(t) + \frac{D_i(z)}{C_i(z)} v_i(t),$$

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

$$\textcircled{2} y_i(t) = \sum_{j=1}^m A_{ij}(z) y_j(t) + \sum_{j=1}^r \frac{B_{ij}(z)}{F_{ij}(z)} u_j(t) + \frac{D_i(z)}{C_i(z)} v_i(t),$$

其中  $\mathbf{y}(t) = [y_1(t), y_2(t), \dots, y_m(t)]^T \in \mathbf{R}^m$  为系统输出向量,  $\mathbf{v}(t) = [v_1(t), v_2(t), \dots, v_m(t)]^T \in \mathbf{R}^m$  是零均值不相关白噪声向量,

$$\begin{aligned} A_{ij}(z) &:= a_{ij}(1)z^{-1} + a_{ij}(2)z^{-2} + \dots + a_{ij}(n_{ij})z^{-n_{ij}}, \\ B_{ij}(z) &:= b_{ij}(1)z^{-1} + b_{ij}(2)z^{-2} + \dots + b_{ij}(n_{ij})z^{-n_{ij}}, \\ F_{ij}(z) &:= 1 + f_{ij}(1)z^{-1} + f_{ij}(2)z^{-2} + \dots + f_{ij}(n_{ij})z^{-n_{ij}}, \\ C_i(z) &:= 1 + c_i(1)z^{-1} + c_i(2)z^{-2} + \dots + c_i(n_i)z^{-n_i}, \\ D_i(z) &:= 1 + d_i(1)z^{-1} + d_i(2)z^{-2} + \dots + d_i(n_i)z^{-n_i}. \end{aligned}$$

11) 研究下列多输入多输出非线性自回归输出误差 ARMA 模型 (非线性 AR-OEARMA 模型) 的辅助模型多新息广义增广辨识算法:

$$\textcircled{1} y_i(t) = \sum_{j=1}^m A_{ij}(z) y_j^2(t) + \sum_{j=1}^r \frac{B_{ij}(z)}{F_{ij}(z)} u_j(t) u_{j+1}(t) + \frac{D_i(z)}{C_i(z)} v_i(t),$$

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

$$\textcircled{2} y_i(t) = \sum_{j=1}^m A_{ij}(z) f_j(y_j(t)) + \sum_{j=1}^r \frac{B_{ij}(z)}{F_{ij}(z)} g_j(u_j(t)) + \frac{D_i(z)}{C_i(z)} v_i(t),$$

$$\textcircled{3} y_i(t) = \sum_{j=1}^m A_{ij}(z) f_i(y_j(t)) + \sum_{j=1}^r \frac{B_{ij}(z)}{F_{ij}(z)} g_i(u_j(t)) + \frac{D_i(z)}{C_i(z)} v_i(t),$$

$$\textcircled{4} y_i(t) = \sum_{j=1}^m A_{ij}(z) f_{ij}(y_1(t), y_2(t), \dots, y_m(t)) + \sum_{j=1}^r \frac{B_{ij}(z)}{F_{ij}(z)} g_{ij}(u_1(t), u_2(t), \dots, u_r(t)) + \frac{D_i(z)}{C_i(z)} v_i(t).$$

12) 研究一类多输入线性参数输出误差 ARMA 系统 (LP-OEARMA 模型) 的辅助模型广义增广随机梯度算法和辅助模型多新息广义增广随机梯度算法:

$$\textcircled{1} y(t) = \frac{1}{A(z)} \sum_{j=1}^r B_j(z) u_j(t) y(t-j) + \frac{D(z)}{C(z)} v(t),$$

$$\textcircled{2} y(t) = \sum_{j=1}^r \frac{B_j(z)}{A_j(z)} u_j(t) y(t-j) + \frac{D(z)}{C(z)} v(t).$$

进一步可以推广到下列非线性噪声情形

$$y(t) = \frac{1}{A(z)} \sum_{j=1}^r B_j(z) u_j(t) y(t-j) + \frac{v(t) + d_1 v(t-1) + d_2 v(t-2) + \dots + d_{m-1} v(t-m+1)}{C(z)}.$$

### 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 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, 马兴云. 规范状态空间系统辨识方法 [J]. 南京信息工程大学学报: 自然科学版, 2014, 6(6): 481-504  
DING Feng, MA Xingyun. Identification methods for canonical state space systems [J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2014, 6(6): 481-504
- [20] 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
- [21] 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 [J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2015, 7(2): 97-124
- [22] 丁锋, 陈启佳. 输出非线性方程误差类系统递推最小二乘辨识方法 [J]. 南京信息工程大学学报: 自然科学版, 2015, 7(3): 193-213  
DING Feng, CHEN Qijia. Recursive least squares identification methods for output nonlinear equation-error type systems. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2015, 7(3): 193-213
- [23] 丁锋, 郭兰杰. 线性参数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(4): 289-312  
DING Feng, GUO Lanjie. Multi-innovation identification methods for linear-parameter systems [J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2015, 7(4): 289-312
- [24] 丁锋. 方程误差类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 type systems [J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2015, 7(5): 385-407
- [25] 丁锋, 谢新民, 方崇智. 时变系统辨识的多新息方法 [J]. 自动化学报, 1996, 22(1): 85-91  
DING Feng, XIE Xinmin, FANG Chongzhi. Multi-innovation identification methods for time-varying systems [J]. Acta Automatica Sinica, 1996, 22(1): 85-91
- [26] 丁锋, 杨家本. 衰减激励条件下确定性系统多新息算法的收敛性分析 [J]. 清华大学学报: 自然科学版, 1998, 38(9): 111-115  
DING Feng, YANG Jiaben. Convergence analysis of multi-innovation identification under attenuating excitation conditions for deterministic systems [J]. Journal of Tsinghua University: Science and Technology, 1998, 38(9): 111-115
- [27] 丁锋, 丁韬, 杨家本, 等. 随机系统多新息辨识在衰减激励条件下的性能分析 [J]. 清华大学学报: 自然科学版, 2001, 41(9): 115-117  
DING Feng, DING Tao, YANG Jiaben, et al. Performance analysis of multi-innovation identification under attenuating excitation for stochastic systems [J]. Journal of Tsinghua University: Science and Technology, 2001, 41(9): 115-117
- [28] 丁锋, 丁韬. 时变多变量系统多新息投影算法的均方收敛性 [J]. 湖北工学院学报, 2001, 16(4): 16-20  
DING Feng, DING Tao. Mean square convergence of multi-innovation projection algorithms for time-varying multivariable systems [J]. Journal of Hubei University of Technology, 2001, 16(4): 16-20
- [29] 丁锋, 萧德云, 丁韬. 多新息随机梯度辨识方法 [J]. 控制理论与应用, 2003, 20(6): 870-874  
DING Feng, XIAO Deyun, DING Tao. Multi-innovation stochastic gradient identification methods [J]. Control Theory and Application, 2003, 20(6): 870-874
- [30] Ding F, Chen T. Performance analysis of multi-innovation gradient type identification methods [J]. Automatica, 2007, 43(1): 1-14
- [31] Ding F, Chen H B, Li M. Multi-innovation least squares identification methods based on the auxiliary model for MISO systems [J]. Applied Mathematics and Computation, 2007, 187(2): 658-668
- [32] 王冬青, 丁锋. Box-Jenkins 模型的基于辅助模型的多新息广义增广随机梯度算法 [J]. 控制与决策, 2008, 23(9): 999-1003+1010  
WANG Dongqing, DING Feng. Auxiliary model based multi-innovation generalized extended stochastic gradient algorithms for Box-Jenkins models [J]. Control and Decision, 2008, 23(9): 999-1003+1010
- [33] Ding F, Liu X P, Liu G. Auxiliary model based multi-innovation extended stochastic gradient parameter estimation with colored measurement noises [J]. Signal Processing, 2009, 89(10): 1883-1890
- [34] Zhang J B, Ding F, Shi Y. Self-tuning control based on multi-innovation stochastic gradient parameter estimation [J]. Systems & Control Letters, 2009, 58(1): 69-75
- [35] Han L L, Ding F. Multi-innovation stochastic gradient algorithms for multi-input multi-output systems [J]. Digital Signal Processing, 2009, 19(4): 545-554
- [36] Liu Y J, Xiao Y S, Zhao X L. Multi-innovation stochastic gradient algorithm for multiple-input single-output systems using the auxiliary model [J]. Applied Mathematics and Computation, 2009, 215(4): 1477-1483
- [37] Han L L, Ding F. Identification for multirate multi-input systems using the multi-innovation identification theory [J]. Computers and Mathematics with Applications, 2009, 57(9): 1438-1449
- [38] 于丽, 刘艳君, 丁锋. CARMA 模型多新息增广随机梯度参数估计算法的收敛性分析 [J]. 系统工程与电子技术, 2009, 31(6): 1446-1449  
YU Li, LIU Yanjun, DING Feng. Convergence of multi-innovation extended stochastic gradient parameter estimation for CARMA models [J]. Systems Engineering

- and Electronics, 2009, 31(6):1446-1449
- [39] Ding F. Several multi-innovation identification methods [J]. Digital Signal Processing, 2010, 20(4):1027-1039
- [40] 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
- [41] Liu Y J, Yu L, Ding F. Multi-innovation extended stochastic gradient algorithm and its performance analysis [J]. Circuits, Systems and Signal Processing, 2010, 29(4):649-667
- [42] Wang D Q, Ding F. Performance analysis of the auxiliary models based multi-innovation stochastic gradient estimation algorithm for output error systems [J]. Digital Signal Processing, 2010, 20(3):750-762
- [43] Xie L, Yang H Z, Ding F. Modeling and identification for non-uniformly periodically sampled-data systems [J]. IET Control Theory and Applications, 2010, 4(5):784-794
- [44] Wang D Q, Chu Y Y, Ding F. Auxiliary model-based RELS and MI-ELS algorithms for Hammerstein OEMA systems [J]. Computers and Mathematics with Applications, 2010, 59(9):3092-3098
- [45] Chen J, Zhang Y, Ding R F. Auxiliary model based multi-innovation algorithms for multivariable nonlinear systems [J]. Mathematical and Computer Modelling, 52(9/10):1428-1434
- [46] Han L L, Ding F. Parameter estimation for multirate multi-input systems using auxiliary model and multi-innovation [J]. Journal of Systems Engineering and Electronics, 2010, 21(6):1079-1083.
- [47] Ding F, Liu G, Liu X P. Parameter estimation with scarce measurements [J]. Automatica, 2011, 47(8):1646-1655
- [48] 丁洁, 谢莉, 丁锋. 非均匀采样系统多新息随机梯度辨识性能分析 [J]. 控制与决策, 2011, 26(9):1338-1342  
DING Jie, XIE Li, DING Feng. Performance analysis of multi-innovation stochastic gradient identification for non-uniformly sampled systems [J]. Control and Decision, 2011, 26(9):1338-1342
- [49] Chen J, Ding F. Least squares and stochastic gradient parameter estimation for multivariable nonlinear Box-Jenkins models based on the auxiliary model and multi-innovation identification theory [J]. Engineering Computations, 2012, 29(8):907-921
- [50] Xiao Y S, Song G L, Liao Y W, et al. Multi-innovation stochastic gradient parameter estimation for input nonlinear controlled autoregressive models [J]. International Journal of Control, Automation, and Systems, 2012, 10(3):639-643
- [51] Ding F, Duan H H. Two-stage parameter estimation algorithms for Box-Jenkins systems [J]. IET Signal Processing, 2013, 7(2):176-184
- [52] Ding F. Hierarchical multi-innovation stochastic gradient algorithm for Hammerstein nonlinear system modeling [J]. Applied Mathematical Modelling, 2013, 37(4):1694-1704
- [53] Chen J, Lv L X, Ding R F. Multi-innovation stochastic gradient algorithms for dual-rate sampled systems with preload nonlinearity [J]. Applied Mathematics Letters, 2014, 26(1):124-129
- [54] 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
- [55] Hu Y B, Liu B L, Zhou Q. A multi-innovation generalized extended stochastic gradient algorithm for output nonlinear autoregressive moving average systems [J]. Applied Mathematics and Computation, 2014, 247:218-224
- [56] 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
- [57] Mao Y W, Ding F. Data filtering-based multi-innovation stochastic gradient algorithm for nonlinear output error autoregressive systems [J]. Circuits, Systems and Signal Processing, 2015. DOI 10.1007/s00034-015-0064-y
- [58] Mao Y W, Ding F. A novel data filtering based multi-innovation stochastic gradient algorithm for Hammerstein nonlinear systems [J]. Digital Signal Processing, 2015, 46:215-225
- [59] Mao Y W, Ding F. Parameter estimation for nonlinear systems by using the data filtering and the multi-innovation identification theory [J]. International Journal of Computer Mathematics, 2015, 92(x). http://dx.doi.org/10.1080/00207160.2015.1077949
- [60] Ding F, Chen T. Parameter estimation of dual-rate stochastic systems by using an output error method [J]. IEEE Transactions on Automatic Control, 2005, 50(9):1436-1441
- [61] Ding F, Chen T. Combined parameter and output estimation of dual-rate systems using an auxiliary model [J]. Automatica, 2004, 40(10):1739-1748
- [62] Ding F, Chen T. Identification of dual-rate systems based on finite impulse response models [J]. International Journal of Adaptive Control and Signal Processing, 2004, 18(7):589-598
- [63] Ding F, Ding J. Least squares parameter estimation with irregularly missing data [J]. International Journal of Adaptive Control and Signal Processing, 2010, 24(7):540-553
- [64] Ding F, Liu X P. Auxiliary model based stochastic gradient algorithm for multivariable output error systems [J]. Acta Automatica Sinica, 2010, 36(7):993-998
- [65] 丁锋, 谢新民. 多变量系统的辅助模型辨识算法 [J]. 清华大学学报:自然科学版, 1992, 32(4):100-106  
DING Feng, XIE Xinmin. Auxiliary model identification method for multivariable systems [J]. Journal of Tsinghua University: Science and Technology, 1992, 32(4):100-106
- [66] 丁锋. 多变量系统的辅助模型辨识方法的收敛性分析 [J]. 控制理论与应用, 1997, 14(2):192-200  
DING Feng. Convergence analysis of auxiliary model identification algorithms for multivariable systems [J]. Control Theory and Applications, 1997, 14(2):192-200
- [67] Wang X H, Ding F. Convergence of the auxiliary model based multi-innovation generalized extended stochastic

- gradient algorithm for Box-Jenkins systems[J]. Nonlinear Dynamics 2015, 82(1/2): 269-280
- [68] Wang X H, Ding F. Performance analysis of the recursive parameter estimation algorithms for multivariable Box-Jenkins systems[J]. Journal of the Franklin Institute-Engineering and Applied Mathematics, 2014, 351(10): 4749-4764
- [69] Liu Y J, Wang D Q, Ding F. Least-squares based iterative algorithms for identifying Box-Jenkins models with finite measurement data[J]. Digital Signal Processing, 2010, 20(5): 1458-1467
- [70] Wang D Q, Yang G W, Ding F. Gradient-based iterative parameter estimation for Box-Jenkins systems with finite measurement data[J]. Computers and Mathematics with Applications, 2010, 60(5): 1200-1208
- [71] Ding F, Wang Y J, Ding J. Recursive least squares parameter identification for systems with colored noise using the filtering technique and the auxiliary model[J]. Digital Signal Processing, 2015, 37: 100-108
- [72] Wang D Q. Least squares-based recursive and iterative estimation for output error moving average systems using data filtering[J]. IET Control Theory and Applications, 2011, 5(14): 1648-1657

## Multi-innovation identification methods for output-error systems

DING Feng<sup>1,2,3</sup>

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** This paper studies the auxiliary model based stochastic gradient (AM-SG) algorithm, the auxiliary model based multi-innovation stochastic gradient (AM-MISG) algorithm, the auxiliary model based recursive least squares (AM-RLS) algorithm and the auxiliary model based multi-innovation least squares algorithm for output-error systems and output-error autoregressive moving average (OEARMA) systems (namely, Box-Jenkins systems), the modified AM-SG algorithm, the forgetting factor AM-SG algorithm, the interval-varying AM-SG algorithm, the interval-varying AM-MISG algorithm and the interval-varying AM-RLS algorithm for output-error systems, and presents the filtering based auxiliary model generalized stochastic gradient algorithm and the filtering based multi-innovation generalized least squares algorithm for output-error autoregressive systems (namely, OEAR systems).

**Key words** parameter estimation; recursive identification; gradient search; least squares; filtering; decomposition; auxiliary model identification idea; multi-innovation identification theory; hierarchical identification principle; output-error system; linear system