DOI:10.13878/j.cnki.jnuist.2016.01.001

丁锋[1,2,3]

辅助模型辨识方法(1):自回归输出误差系统

摘要

研究了自回归输出误差(AR-OE)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,自回归输出误差自回归滑动平均(AR-OEARMA)系统(即AR-Box-Jenkins系统)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最小二乘算法,以及AR-Box-Jenkins系统的基于滤波的辅助模型广义增广随机梯度算法、基于滤波的辅助模型多新息广义增广随机梯度算法、基于滤波的辅助模型递推广义增广最小二乘算法等.

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;方程误差系统;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 2015-11-21

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

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

0 引言

辨识算法有两种基本形式——递推辨识与迭代辨识.递推辨识可以用于在线辨识,实时估计系统的参数,它在每一次递推计算中都使用了实时采集的输入输出数据,也就是每一次递推计算都使用了新的数据,通过使用实时数据刷新参数估计,递推变量(recursive index)是客观世界的时间.迭代辨识只能用于离线辨识系统参数,它是采集一批数据进行辨识,在每一步迭代计算参数估计中,并没引入新的观测数据,依旧是使用同一批数据刷新参数估计,迭代变量(iterative index)是自然数,与客观世界的时间无关.虽然从静态数学观点看,"递推"、"递归"、"迭代"等具有相近、相同的意思,但是从动态系统的角度看,优化的准则函数往往是与时间有关的.为区别静态优化问题与动态优化问题,本文只使用"递推"和"迭代"两个词,不使用"递归"一词,且约定"递推"用于动态优化问题,"迭代"用于静态优化问题[1-3].对系统辨识而言,就有递推辨识方法(递推参数估计方法)和迭代辨识方法(迭代参数估计方法)之分.

辨识方法一般是通过优化系统输出与模型输出误差平方(和)得到的算法.一般的寻优方法有梯度搜索方法、最小二乘搜索方法、牛顿搜索方法,以及一些变形的搜索方法,如共轭梯度方法、拟牛顿方法等.最小二乘一般用于线性参数系统辨识问题,其扩展可以用于双线性参数系统辨识问题等.对于线性参数系统辨识问题(系统可以是非线性的,但关于参数空间是线性的),牛顿辨识方法就退化为最小二乘辨识方法[4].对于非线性参数系统,辨识的准则函数(也称为目标函数、损失函数)依旧是系统输出与模型输出误差平方(和),有的文献称之为预报误差,这是不妥的(实际上是模型输出误差,不是预报误差),更有甚者把极小化这一准则函数得到的方法称为预报误差辨识方法,就更不妥了,因为这是辨识的定义,而不是辨识方法的定义,如果按照此不妥的定义,那么所有辨识方法都是预报误差辨识方法,名称"预报误差辨识方法"也就失去意义了.

对于同一准则函数,采用不同的优化手段将产生不同的辨识方法,如随机梯度辨识方法、梯度迭代辨识方法、递推最小二乘辨识方法、最小二乘迭代辨识方法、牛顿递推辨识方法、牛顿迭代辨识方法等.从这个层面讲,辨识方法的优化手段也极其重要,这就是笔者提出辨识四要素的缘由.辨识四要素包括输入输出数据、模型、准则函数、

优化方法,简称为数据、模型、准则和优化方法.系统辨识是通过设计适当的输入信号,利用实验的输入输出数据,选择一类模型,构造一个误差准则函数,用一种优化方法确定一个与数据拟合得最好的模型[1].

辨识的四要素是根本,而辨识中每一种思想、理论、原理、概念、理念从孕育到诞生都经历着一个奇妙的过程,随之产生的每一个辨识方法都有其特点和适用范围.例如,多新息辨识方法可以改进随机梯度类辨识算法参数估计精度;递阶辨识方法可以解决复杂结构多变量系统、双线性参数系统、非线性系统的辨识问题,能够减小算法的计算量;耦合辨识方法可以解决多变量系统各子系统间存在耦合参数时的辨识问题,从而降低算法的计算量;滤波辨识方法能够提高有色噪声干扰系统参数估计的精度;辅助模型辨识方法可以解决存在不可测变量的系统的辨识问题.

在这个辨识新思想、新理论、新原理、新概念、新理念相继问世的年代,笔者提炼和发展了辅助模型辨识思想、多新息辨识理论、递阶辨识原理、耦合辨识概念、滤波辨识理念等,它们与梯度搜索、最小二乘搜索、牛顿搜索等相结合,以及其在有色噪声干扰系统、非线性系统中的应用,便产生层出不穷的辨识方法,这为系统辨识的研究增添了无比繁荣的景象.最近出版的《系统辨识学术专著丛书》第1分册《系统辨识新论》[1]、第3分册《系统辨识——辨识方法性能分析》[2]、第6分册《系统辨识——多新息辨识理论与方法》,《南京信息工程大学学报》上2011年至今的系统辨识连载论文中的辨识方法,及其派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等[5-59]的收敛性研究,更是给系统辨识研究的主旋律添上了迷人的色彩.

文献[3,7-8]利用辅助模型辨识思想,研究了输出误差系统和线性参数系统的辅助模型辨识方法.自回归输出误差系统是输出误差系统的推广,本文将文献[3,8]中输出误差系统的辅助模型辨识方法推广到自回归输出误差系统,研究自回归输出误差系统的辅助模型随机梯度辨识方法和基于滤波技术的辅助模型辨识方法等.

1 自回归输出误差系统

利用辅助模型辨识思想,将文献[3,8]中输出误差(OE)系统的辅助模型辨识方法推广到自回归输出误差(AR-OE)系统,研究AR-OE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)算法等.输出误差系统的变递推间隔辅助模型多新息随机梯度算法以“Regular Paper”形式发表在国际期刊《Automatica》2011年第8期上[36],该文入选“2011年中国百篇最具影响国际学术论文”.

1.1 AR-OE 系统描述与辨识模型

考虑下列自回归输出误差模型(AutoRegressive Output-Error model, AR-OE 模型)描述的动态随机系统(参见图1上半部分):

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

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σ^2的随机白噪声序列,z^{-1}为单位后移算子[$z^{-1}y(t)=y(t-1)$, $zy(t)=y(t+1)$],$A(z)$, $F(z)$和$B(z)$是单位后移算子z^{-1}的多项式:

$$A(z):=1+a_1z^{-1}+a_2z^{-2}+\cdots+a_{n_a}z^{-n_a},$$
$$F(z):=1+f_1z^{-1}+f_2z^{-2}+\cdots+f_{n_f}z^{-n_f},$$
$$B(z):=b_1z^{-1}+b_2z^{-2}+\cdots+b_{n_b}z^{-n_b}.$$

设阶次n_a, n_f和n_b已知,记$n:=n_a+n_f+n_b$,且$t\leqslant 0$时,$y(t)=0$, $u(t)=0$, $v(t)=0$.图1中,$x(t)$是不可测中间变量.

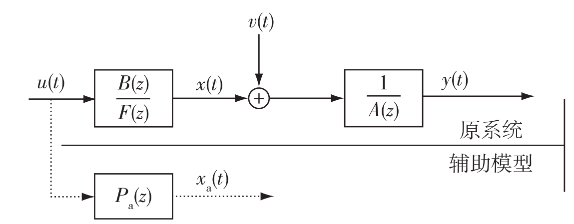

图1 带辅助模型的自回归输出误差(AR-OE)系统
Fig. 1 The autoregressive output-error system with the auxiliary model

定义系统中间变量(intermediate variable)

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

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

$$\boldsymbol{\vartheta}:=\begin{bmatrix}\boldsymbol{a}\\ \boldsymbol{\rho}\end{bmatrix}\in\mathbf{R}^n,$$
$$\boldsymbol{a}:=[a_1,a_2,\cdots,a_{n_a}]^{\mathrm{T}}\in\mathbf{R}^{n_a},$$
$$\boldsymbol{\rho}:=[f_1,f_2,\cdots,f_{n_f},b_1,b_2,\cdots,b_{n_b}]^{\mathrm{T}}\in\mathbf{R}^{n_f+n_b},$$

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

$$\boldsymbol{\varphi}_y(t):=[-y(t-1),-y(t-2),\cdots,-y(t-n_a)]^{\mathrm{T}}\in\mathbf{R}^{n_a},$$

$$\boldsymbol{\phi}(t):=[-x(t-1),-x(t-2),\cdots,-x(t-n_f),\ u(t-1),u(t-2),\cdots,u(t-n_b)]^{\mathrm{T}}\in\mathbf{R}^{n_f+n_b}.$$

由式(2)可得

$$x(t)=[1-F(z)]x(t)+B(z)u(t)=\boldsymbol{\phi}^{\mathrm{T}}(t)\boldsymbol{\rho}. \quad (3)$$

将式(2)代入式(1),使用式(3),可得自回归输出误差系统的辨识模型(identification model):

$$y(t)=[1-A(z)]y(t)+x(t)+v(t)=\boldsymbol{\varphi}_y^{\mathrm{T}}(t)\boldsymbol{a}+\boldsymbol{\phi}^{\mathrm{T}}(t)\boldsymbol{\rho}+v(t)=\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\vartheta}+v(t). \quad (4)$$

设 $\hat{X}(t)$ 为 X 在时刻 t 的估计.这意味着 $\hat{\boldsymbol{a}}(t)$ 是 $\boldsymbol{a}$ 在时刻 t 的估计,$\hat{\boldsymbol{\vartheta}}(t)$ 是 $\boldsymbol{\vartheta}$ 在时刻 t 的估计,$\hat{\boldsymbol{\vartheta}}(t):=\begin{bmatrix}\hat{\boldsymbol{a}}(t)\\ \hat{\boldsymbol{\rho}}(t)\end{bmatrix}\in\mathbf{R}^n$ 是 $\boldsymbol{\vartheta}=\begin{bmatrix}\boldsymbol{a}\\ \boldsymbol{\rho}\end{bmatrix}$ 在时刻 t 的估计,$\hat{v}(t)$ 是 $v(t)$ 的估计,等等.

1.2 辅助模型的建立

辨识模型(4)中信息向量 $\boldsymbol{\varphi}(t)$ 是未知的,因为其除了包含系统可测输入输出数据 $u(t-i)$ 和 $y(t-i)$,还包含了未知内部变量 $x(t-i)$,这是输出误差类系统辨识的困难.解决这一困难的方案是使用辅助模型辨识思想,用系统的可测信息(包括计算获得的信息)建立一个辅助模型,用辅助模型的输出代替这些未知变量,从而获得基于辅助模型的辨识方法.具体思路如下.

构造一个辅助模型 $P_{\mathrm{a}}(z)$,如图1中下半部分所示,$x_{\mathrm{a}}(t)$ 为辅助模型的输出.用辅助模型的输出 $x_{\mathrm{a}}(t-i)$ 和系统输入 $u(t-i)$ 定义 $\boldsymbol{\phi}(t)$ 的估计

$$\hat{\boldsymbol{\phi}}(t):=[-x_{\mathrm{a}}(t-1),-x_{\mathrm{a}}(t-2),\cdots,-x_{\mathrm{a}}(t-n_f),\ u(t-1),u(t-2),\cdots,u(t-n_b)]^{\mathrm{T}}\in\mathbf{R}^{n_f+n_b}.$$

用 $\boldsymbol{\varphi}_y(t)$ 和 $\hat{\boldsymbol{\phi}}(t)$ 构造 $\boldsymbol{\varphi}(t)$ 的估计

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

参考文献[8],取估算 $x(t)$ 的辅助模型为

$$x_{\mathrm{a}}(t)=\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\rho}}(t).$$

1.3 辅助模型随机梯度辨识方法

定义矩阵 $\boldsymbol{X}$ 的范数为 $\|\boldsymbol{X}\|^2:=\mathrm{tr}[\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}]$.设 $1/r(t)$ 是收敛因子(convergence factor)或步长(step-size).根据辨识模型(4),引入梯度准则函数(gradient criterion function):

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

使用负梯度搜索和辅助模型辨识思想,极小化 $J_1(\boldsymbol{\vartheta})$,未知的 $\boldsymbol{\varphi}(t)$ 用其估计 $\hat{\boldsymbol{\varphi}}(t)$ 代替,可以得到估计 AR-OE 系统参数向量 $\boldsymbol{\vartheta}$ 的辅助模型随机梯度算法(Auxiliary Model based Stochastic Gradient algorithm,AM-SG 算法)[1,8]:

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\frac{\hat{\boldsymbol{\varphi}}(t)}{r(t)}e(t), \quad (5)$$

$$e(t)=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t-1), \quad (6)$$

$$r(t)=r(t-1)+\|\hat{\boldsymbol{\varphi}}(t)\|^2, \quad (7)$$

$$\hat{\boldsymbol{\varphi}}(t)=\begin{bmatrix}\boldsymbol{\varphi}_y(t)\\ \hat{\boldsymbol{\phi}}(t)\end{bmatrix}, \quad (8)$$

$$\boldsymbol{\varphi}_y(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a)]^{\mathrm{T}}, \quad (9)$$

$$\hat{\boldsymbol{\phi}}(t)=[-x_{\mathrm{a}}(t-1),\cdots,-x_{\mathrm{a}}(t-n_f),u(t-1),\cdots,u(t-n_b)]^{\mathrm{T}}, \quad (10)$$

$$x_{\mathrm{a}}(t)=\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\rho}}(t), \quad (11)$$

$$\hat{\boldsymbol{\vartheta}}(t)=[\hat{a}_1(t),\hat{a}_2(t),\cdots,\hat{a}_{n_a}(t),\hat{\boldsymbol{\rho}}^{\mathrm{T}}(t)]^{\mathrm{T}}, \quad (12)$$

$$\hat{\boldsymbol{\rho}}(t)=[\hat{f}_1(t),\hat{f}_2(t),\cdots,\hat{f}_{n_f}(t),\hat{b}_1(t),\hat{b}_2(t),\cdots,\hat{b}_{n_b}(t)]^{\mathrm{T}}. \quad (13)$$

AM-SG 算法(5)—(13)计算参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$ 的步骤如下:

1)初始化:令 $t=1$.置初值 $\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_n/p_0$,$r(0)=1$,$x_{\mathrm{a}}(t-i)=1/p_0$,$i=1,2,\cdots,n_f$,$p_0=10^6$,$\mathbf{1}_n$ 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(9)—(10)和(8)构造信息向量 $\boldsymbol{\varphi}_y(t)$,$\hat{\boldsymbol{\phi}}(t)$ 和 $\hat{\boldsymbol{\varphi}}(t)$.

3)由式(6)计算新息 $e(t)$,由式(7)计算 $r(t)$.

4)根据式(5)刷新参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$.

5)从式(12)中 $\hat{\boldsymbol{\vartheta}}(t)$ 读出 $\hat{\boldsymbol{\rho}}(t)$,由式(11)计算辅助模型的输出 $x_{\mathrm{a}}(t)$.

6)t 增1,转到第2)步.

注1 设 q 为数据窗长度.将式(7)修改为

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

就得到辅助模型广义投影算法(AM-GP 算法).

注2 为提高 AM-SG 算法的暂态性能和参数估计精度,可在式(7)中引入遗忘因子(Forgetting Factor,FF)λ,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).

注3 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,可在式(5)中引入收敛指数(convergence index)ε,就得到修正辅助模型随机梯

度算法(Modified AM-SG algorithm, M-AM-SG 算法)[1,8].

注 4 随机梯度类算法和多新息随机梯度类算法,如增广随机梯度算法、广义增广随机梯度算法、辅助模型随机梯度算法、辅助模型增广随机梯度算法、多新息随机梯度算法、辅助模型多新息广义增广随机梯度算法等,都可以引入遗忘因子和(或)收敛指数,来改进算法的参数估计性能.

注 5 AR-OE 系统的 AM-SG 算法和 M-AM-SG 算法参数估计的一致收敛性,FF-AM-SG 算法和 FF-M-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.

1.4 辅助模型多新息随机梯度辨识方法

下面借助于多新息辨识理论[1,3],通过扩展新息维数,可推导出基于辅助模型的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(5)—(13)中标量新息(innovation) $e(t)=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\cdot\hat{\boldsymbol{\vartheta}}(t-1)\in\mathbf{R}$ 扩展为新息向量(innovation vector):

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

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

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

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

于是,可以得到新息长度为 p 的,估计输出误差系统参数向量 $\boldsymbol{\vartheta}$ 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):

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\frac{\hat{\boldsymbol{\Phi}}(p,t)}{r(t)}\boldsymbol{E}(p,t),\tag{14}$$

$$\boldsymbol{E}(p,t)=\boldsymbol{Y}(p,t)-\hat{\boldsymbol{\Phi}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\vartheta}}(t-1),\tag{15}$$

$$r(t)=r(t-1)+\|\hat{\boldsymbol{\varphi}}(t)\|^{2},\tag{16}$$

$$\boldsymbol{Y}(p,t)=[y(t),y(t-1),\cdots,y(t-p+1)]^{\mathrm{T}},\tag{17}$$

$$\hat{\boldsymbol{\Phi}}(p,t)=[\hat{\boldsymbol{\varphi}}(t),\hat{\boldsymbol{\varphi}}(t-1),\cdots,\hat{\boldsymbol{\varphi}}(t-p+1)],\tag{18}$$

$$\hat{\boldsymbol{\varphi}}(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a),\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t)]^{\mathrm{T}},\tag{19}$$

$$\hat{\boldsymbol{\phi}}(t)=[-x_{\mathrm{a}}(t-1),\cdots,-x_{\mathrm{a}}(t-n_f),u(t-1),\cdots,u(t-n_b)]^{\mathrm{T}},\tag{20}$$

$$x_{\mathrm{a}}(t)=\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\rho}}(t),\tag{21}$$

$$\hat{\boldsymbol{\vartheta}}(t)=[\hat{a}_1(t),\hat{a}_2(t),\cdots,\hat{a}_{n_a}(t),\hat{\boldsymbol{\rho}}^{\mathrm{T}}(t)]^{\mathrm{T}}.\tag{22}$$

当新息长度 $p=1$ 时,AM-MISG 算法退化为 AM-SG 算法(5)—(13).

AM-MISG 算法(14)—(22)计算参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$ 的步骤如下:

1) 初始化:令 $t=1$,给定新息长度 p.置初值 $\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_n/p_0$, $r(0)=1$, $x_{\mathrm{a}}(t-i)=1/p_0$, $i=1,2,\cdots,n_f$, $p_0=10^6$.

2) 采集输入输出数据 $u(t)$ 和 $y(t)$,由式(20)和(19)构造信息向量 $\hat{\boldsymbol{\phi}}(t)$ 和 $\hat{\boldsymbol{\varphi}}(t)$,由式(17)—(18)构造堆积输出向量 $\boldsymbol{Y}(p,t)$ 和堆积信息矩阵 $\hat{\boldsymbol{\Phi}}(p,t)$.

3) 由式(15)计算新息向量 $\boldsymbol{E}(p,t)$,由式(16)计算 $r(t)$.

4) 根据式(14)刷新参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$.

5) 从式(22)的 $\hat{\boldsymbol{\vartheta}}(t)$ 中读出 $\hat{\boldsymbol{\rho}}(t)$,由式(21)计算辅助模型的输出 $x_{\mathrm{a}}(t)$.

6) t 增 1,转到第 2)步.

注 6 文献[28]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.

1.5 辅助模型递推最小二乘辨识方法

对于输出误差系统辨识模型(4),定义准则函数(criterion function):

$$J_2(\boldsymbol{\vartheta}):=\sum_{j=1}^{t}[y(j)-\boldsymbol{\varphi}^{\mathrm{T}}(j)\boldsymbol{\vartheta}]^{2}.$$

参照文献[1-2]算法的推导,使用辅助模型辨识思想,未知向量 $\boldsymbol{\varphi}(t)$ 用其估计 $\hat{\boldsymbol{\varphi}}(t)$ 代替,可以得到估计输出误差系统参数向量 $\boldsymbol{\vartheta}$ 的辅助模型递推最小二乘算法(Auxiliary Model based Recursive Least Squares algorithm, AM-RLS 算法)[1]:

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\boldsymbol{L}(t)[y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t-1)],\tag{23}$$

$$\boldsymbol{L}(t)=\boldsymbol{P}(t-1)\hat{\boldsymbol{\varphi}}(t)[1+\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\boldsymbol{P}(t-1)\hat{\boldsymbol{\varphi}}(t)]^{-1},\tag{24}$$

$$\boldsymbol{P}(t)=[\boldsymbol{I}_n-\boldsymbol{L}(t)\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)]\boldsymbol{P}(t-1),\tag{25}$$

$$\hat{\boldsymbol{\varphi}}(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a),\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t)]^{\mathrm{T}},\tag{26}$$

$$\hat{\boldsymbol{\phi}}(t)=[-x_{\mathrm{a}}(t-1),\cdots,-x_{\mathrm{a}}(t-n_f),u(t-1),\cdots,u(t-n_b)]^{\mathrm{T}},\tag{27}$$

$$x_{\mathrm{a}}(t)=\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\rho}}(t),\tag{28}$$

$$\hat{\boldsymbol{\vartheta}}(t)=[\hat{a}_1(t),\hat{a}_2(t),\cdots,\hat{a}_{n_a}(t),\hat{\boldsymbol{\rho}}^{\mathrm{T}}(t)]^{\mathrm{T}}.\tag{29}$$

AM-RLS 算法(23)—(29)计算参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$ 的步骤如下:

1) 初始化:令 $t=1$.置初值 $\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_n/p_0$, $\boldsymbol{P}(0)=p_0\boldsymbol{I}_n$, $x_{\mathrm{a}}(t-i)=1/p_0$, $i=1,2,\cdots,n_{\mathrm{f}}$, $p_0=10^6$, $\boldsymbol{I}_n$ 是一个 n 阶单位阵.

2) 采集输入输出数据 $u(t)$ 和 $y(t)$,由式(27)和(26)构造信息向量 $\hat{\boldsymbol{\phi}}(t)$ 和 $\hat{\boldsymbol{\varphi}}(t)$.

3) 由式(24)计算增益向量 $\boldsymbol{L}(t)$,由式(25)计

算协方差阵 $\boldsymbol{P}(t)$.

4) 根据式(23)刷新参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$.

5) 从式(29)的 $\hat{\boldsymbol{\vartheta}}(t)$ 中读出 $\hat{\boldsymbol{\rho}}(t)$,由式(28)计算辅助模型的输出 $x_{\mathrm{a}}(t)$.

6) t 增 1,转到第 2)步.

注 7 文献[13]证明了输出误差双率采样数据系统的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[10,12]研究了多变量输出误差系统的辅助模型递推最小二乘算法(AM-MRLS 算法)的收敛性.读者可尝试证明 AM-RLS 算法(23)—(29)的收敛性.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.

注 8 本节的辨识方法可以推广到多输入 AR-OE 模型描述的动态随机系统:

1) $A(z)y(t)=\dfrac{1}{F(z)}\sum\limits_{j=1}^{r}B_j(z)u_j(t)+v(t)$,

2) $A(z)y(t)=\sum\limits_{j=1}^{r}\dfrac{B_j(z)}{F_j(z)}u_j(t)+v(t)$,

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

和非线性自回归输出误差模型(N-AR-OE 模型)描述的非线性动态随机系统:

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

其中 $f(*)$ 和 $g(*)$ 为已知基函数,例子如下:

① $f(y(t))=y^2(t)$,
$g(u(t))=\sin^2(u(t))+\mathrm{e}^{\cos(t/\pi)}$.

② $f(y(t))=|y(t)|$,
$g(u(t))=\sqrt[3]{u^2(t)}+\ln[u^2(t-1)+1]$.

2 自回归输出误差滑动平均系统

《系统辨识——多新息辨识理论与方法》[3]详细介绍了输出误差滑动平均(OEMA)系统的辅助模型辨识方法,这里简单给出自回归 OEMA 系统的辅助模型辨识方法.

2.1 AR-OEMA 系统描述与辨识模型

考虑下列自回归输出误差滑动平均模型(AutoRegressive Output-Error Moving Average model,AR-OEMA 模型)描述的动态随机系统(参见图 2):

$$A(z)y(t)=\frac{B(z)}{F(z)}u(t)+D(z)v(t),\tag{30}$$

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),$A(z)$,$B(z)$,$D(z)$ 和 $F(z)$ 均为单位后移算子 z^{-1} 的多项式:

$$A(z)=1+a_1z^{-1}+a_2z^{-2}+\cdots+a_{n_a}z^{-n_a},$$
$$B(z)=b_1z^{-1}+b_2z^{-2}+\cdots+b_{n_b}z^{-n_b},$$
$$D(z)=1+d_1z^{-1}+d_2z^{-2}+\cdots+d_{n_d}z^{-n_d},$$
$$F(z):=1+f_1z^{-1}+f_2z^{-2}+\cdots+f_{n_f}z^{-n_f}.$$

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

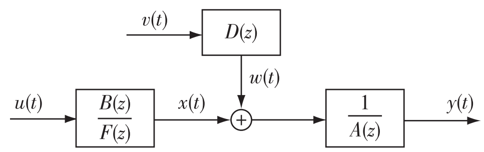

图 2 自回归输出误差滑动平均系统

Fig. 2 The AR-OEMA system

定义未知中间变量(intermediate variable)

$$x(t):=\frac{B(z)}{F(z)}u(t),\tag{31}$$

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

$$\boldsymbol{\vartheta}:=\begin{bmatrix}\boldsymbol{a}\\\boldsymbol{\rho}\\\boldsymbol{d}\end{bmatrix}\in\mathbf{R}^n,$$
$$\boldsymbol{a}:=[a_1,a_2,\cdots,a_{n_a}]^{\mathrm{T}}\in\mathbf{R}^{n_a},$$
$$\boldsymbol{\rho}:=[f_1,f_2,\cdots,f_{n_f},b_1,b_2,\cdots,b_{n_b}]^{\mathrm{T}}\in\mathbf{R}^{n_f+n_b},$$
$$\boldsymbol{d}:=[d_1,d_2,\cdots,d_{n_d}]^{\mathrm{T}}\in\mathbf{R}^{n_d},$$
$$\boldsymbol{\varphi}(t):=[\boldsymbol{\varphi}_y^{\mathrm{T}}(t),\boldsymbol{\phi}^{\mathrm{T}}(t),v(t-1),v(t-2),\cdots,v(t-n_d)]^{\mathrm{T}}\in\mathbf{R}^n,$$
$$\boldsymbol{\varphi}_y(t):=[-y(t-1),-y(t-2),\cdots,-y(t-n_a)]^{\mathrm{T}}\in\mathbf{R}^{n_a},$$
$$\boldsymbol{\phi}(t):=[-x(t-1),-x(t-2),\cdots,-x(t-n_f),$$
$$u(t-1),u(t-2),\cdots,u(t-n_b)]^{\mathrm{T}}\in\mathbf{R}^{n_f+n_b}.$$

设 $\hat{\boldsymbol{\vartheta}}(t):=\begin{bmatrix}\hat{\boldsymbol{a}}(t)\\\hat{\boldsymbol{\rho}}(t)\\\hat{\boldsymbol{d}}(t)\end{bmatrix}\in\mathbf{R}^n$ 是 $\boldsymbol{\vartheta}=\begin{bmatrix}\boldsymbol{a}\\\boldsymbol{\rho}\\\boldsymbol{d}\end{bmatrix}$ 在时刻 t 的估计.由式(31)可得

$$x(t)=[1-F(z)]x(t)+B(z)u(t)=\boldsymbol{\phi}^{\mathrm{T}}(t)\boldsymbol{\rho}.\tag{32}$$

将式(31)代入式(30)可得

$$y(t)=[1-A(z)]y(t)+x(t)+[D(z)-1]v(t)+v(t)=\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\vartheta}+v(t).\tag{33}$$

系统的可测输入输出数据为$\{u(t),y(t)\}$,中间变量 $x(t)$ 和白噪声 $v(t)$ 是不可测的.因此,信息向量 $\boldsymbol{\phi}(t)$ 和 $\boldsymbol{\varphi}(t)$ 是未知的.根据辅助模型辨识思想:用输出信息向量 $\boldsymbol{\varphi}_y(t)$,辅助模型的输出 $x_{\mathrm{a}}(t-i)$ 和系

统输入 $u(t-i)$,以及噪声 $v(t-i)$ 的估计 $\hat{v}(t-i)$ 构造 $\boldsymbol{\varphi}(t)$ 和 $\boldsymbol{\phi}(t)$ 的估计:

$$\hat{\boldsymbol{\varphi}}(t)=[\boldsymbol{\varphi}_y^{\mathrm{T}}(t),\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t),\hat{v}(t-1),\hat{v}(t-2),\cdots,\hat{v}(t-n_d)]^{\mathrm{T}}\in\mathbf{R}^n,\tag{34}$$

$$\hat{\boldsymbol{\phi}}(t):=[-x_{\mathrm{a}}(t-1),-x_{\mathrm{a}}(t-2),\cdots,-x_{\mathrm{a}}(t-n_f),u(t-1),u(t-2),\cdots,u(t-n_b)]^{\mathrm{T}}\in\mathbf{R}^{n_f+n_b}.\tag{35}$$

用 $\hat{\boldsymbol{\phi}}(t)$ 和 $\hat{\boldsymbol{\rho}}(t)$ 构造估算 $x(t)$ 的辅助模型

$$x_{\mathrm{a}}(t)=\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\rho}}(t).\tag{36}$$

$x_{\mathrm{a}}(t)$ 可作为 $x(t)$ 的估计,即 $\hat{x}(t)=x_{\mathrm{a}}(t)$.根据式(33)可得计算 $v(t)$ 的估计 $\hat{v}(t)$ 的辅助模型:

$$\hat{v}(t)=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t).\tag{37}$$

2.2 辅助模型增广随机梯度算法

根据辨识模型(33),可得下列梯度递推关系:

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\frac{\boldsymbol{\varphi}(t)}{r(t)}[y(t)-\boldsymbol{\varphi}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t-1)],\tag{38}$$

$$r(t)=r(t-1)+\|\boldsymbol{\varphi}(t)\|^2,\quad r(0)=1.\tag{39}$$

式(38)—(39)中未知的 $\boldsymbol{\varphi}(t)$ 用其估计 $\hat{\boldsymbol{\varphi}}(t)$ 代替,可以得到估计 AR-OEMA 系统参数向量 $\boldsymbol{\vartheta}$ 的辅助模型增广随机梯度算法(Auxiliary Model based Extended Stochastic Gradient algorithm,AM-ESG 算法):

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\frac{\hat{\boldsymbol{\varphi}}(t)}{r(t)}e(t),\tag{40}$$

$$e(t)=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t-1),\tag{41}$$

$$r(t)=r(t-1)+\|\hat{\boldsymbol{\varphi}}(t)\|^2,\tag{42}$$

$$\hat{\boldsymbol{\varphi}}(t)=[\boldsymbol{\varphi}_y^{\mathrm{T}}(t),\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t),\hat{v}(t-1),\hat{v}(t-2),\cdots,\hat{v}(t-n_d)]^{\mathrm{T}},\tag{43}$$

$$\boldsymbol{\varphi}_y(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a)]^{\mathrm{T}},\tag{44}$$

$$\hat{\boldsymbol{\phi}}(t)=[-x_{\mathrm{a}}(t-1),\cdots,-x_{\mathrm{a}}(t-n_f),u(t-1),\cdots,u(t-n_b)]^{\mathrm{T}},\tag{45}$$

$$x_{\mathrm{a}}(t)=\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\rho}}(t),\tag{46}$$

$$\hat{v}(t)=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t),\tag{47}$$

$$\hat{\boldsymbol{\vartheta}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{\rho}}^{\mathrm{T}}(t),\hat{\boldsymbol{d}}^{\mathrm{T}}(t)]^{\mathrm{T}}.\tag{48}$$

AM-ESG 算法(40)—(48)计算参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$ 的步骤如下:

1) 初始化:令 $t=1$,置初值 $\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_n/p_0$,$r(0)=1$,$x_{\mathrm{a}}(t-i)=1/p_0$,$\hat{v}(t-i)=1/p_0$,$i=1,2,\cdots,\max[n_f,n_d]$,$p_0=10^6$.

2) 采集输入输出数据 $u(t)$ 和 $y(t)$,由式(44)—(45)和(43)构造信息向量 $\boldsymbol{\varphi}_y(t)$,$\hat{\boldsymbol{\phi}}(t)$ 和 $\hat{\boldsymbol{\varphi}}(t)$.

3) 由式(41)计算新息 $e(t)$,由式(42)计算 $r(t)$.

4) 根据式(40)刷新参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$.

5) 从式(48)的 $\hat{\boldsymbol{\vartheta}}(t)$ 中读取 $\hat{\boldsymbol{a}}(t)$,$\hat{\boldsymbol{\rho}}(t)$ 和 $\hat{\boldsymbol{d}}(t)$.由式(46)—(47)计算辅助模型的输出 $x_{\mathrm{a}}(t)$ 和 $\hat{v}(t)$.

6) t 增 1,转到第 2)步.

注 9 可在式(49)中引入遗忘因子 λ,得到遗忘因子辅助模型增广随机梯度算法(Forgetting Factor AM-ESG algorithm,FF-AM-ESG 算法).可在式(40)中引入收敛指数 ε,得到修正辅助模型增广随机梯度算法(Modified AM-ESG algorithm,M-AM-ESG 算法).

2.3 辅助模型多新息增广随机梯度算法

借助于多新息辨识理论,基于 AM-ESG 算法(40)—(48),将系统输出 $y(t)$ 和信息向量 $\hat{\boldsymbol{\varphi}}(t)$ 扩展为堆积输出向量 $\boldsymbol{y}(p,t)$ 和堆积信息矩阵 $\hat{\boldsymbol{\Phi}}(p,t)$,将标量新息 $e(t)\in\mathbf{R}$ 扩展为新息向量,就得到估计 AR-OEMA 系统参数向量 $\boldsymbol{\vartheta}$ 的辅助模型多新息增广随机梯度算法(Auxiliary Model based Multi-Innovation Extended Stochastic Gradient algorithm,AM-MI-ESG 算法):

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\frac{\hat{\boldsymbol{\Phi}}(p,t)}{r(t)}\boldsymbol{E}(p,t),\tag{49}$$

$$r(t)=r(t-1)+\|\hat{\boldsymbol{\varphi}}(t)\|^2,\tag{50}$$

$$\boldsymbol{E}(p,t)=\boldsymbol{Y}(p,t)-\hat{\boldsymbol{\Phi}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\vartheta}}(t-1),\tag{51}$$

$$\boldsymbol{Y}(p,t)=[y(t),y(t-1),\cdots,y(t-p+1)]^{\mathrm{T}},\tag{52}$$

$$\hat{\boldsymbol{\Phi}}(p,t)=[\hat{\boldsymbol{\varphi}}(t),\hat{\boldsymbol{\varphi}}(t-1),\cdots,\hat{\boldsymbol{\varphi}}(t-p+1)],\tag{53}$$

$$\hat{\boldsymbol{\varphi}}(t)=[\boldsymbol{\varphi}_y^{\mathrm{T}}(t),\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t),\hat{v}(t-1),\hat{v}(t-2),\cdots,\hat{v}(t-n_d)]^{\mathrm{T}},\tag{54}$$

$$\boldsymbol{\varphi}_y(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a)]^{\mathrm{T}},\tag{55}$$

$$\hat{\boldsymbol{\phi}}(t)=[-x_{\mathrm{a}}(t-1),\cdots,-x_{\mathrm{a}}(t-n_f),u(t-1),\cdots,u(t-n_b)]^{\mathrm{T}},\tag{56}$$

$$x_{\mathrm{a}}(t)=\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\rho}}(t),\tag{57}$$

$$\hat{v}(t)=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t),\tag{58}$$

$$\hat{\boldsymbol{\vartheta}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{\rho}}^{\mathrm{T}}(t),\hat{\boldsymbol{d}}^{\mathrm{T}}(t)]^{\mathrm{T}}.\tag{59}$$

当新息长度 $p=1$ 时,AM-MI-ESG 算法退化为 AM-ESG 算法(40)—(48).

AM-MI-ESG 算法(49)—(59)计算参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$ 的步骤如下:

1) 初始化:令 $t=1$,给定新息长度 p.置 $\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_n/p_0$,$r(0)=1$,$x_{\mathrm{a}}(t-i)=1/p_0$,$\hat{v}(t-i)=1/p_0$,$i=1,2,\cdots,\max[n_f,n_d]$,$p_0=10^6$.

2) 采集输入输出数据 $u(t)$ 和 $y(t)$,由式(52)构造堆积输出向量 $\boldsymbol{Y}(p,t)$,由式(55)—(56)和(54)构造信息向量 $\boldsymbol{\varphi}_y(t)$,$\hat{\boldsymbol{\phi}}(t)$ 和 $\hat{\boldsymbol{\varphi}}(t)$,由式(53)构造堆积信息矩阵 $\hat{\boldsymbol{\Phi}}(p,t)$.

3) 由式(51)计算新息向量 $\boldsymbol{E}(p,t)$,由式(50)计算 $r(t)$.

4) 根据式(49)刷新参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$.

5) 从式(59)的 $\hat{\boldsymbol{\vartheta}}(t)$ 中读取 $\hat{\boldsymbol{a}}(t)$,$\hat{\boldsymbol{\rho}}(t)$ 和 $\hat{\boldsymbol{d}}(t)$. 由式(57)—(58)计算辅助模型的输出 $x_{\mathrm{a}}(t)$ 和 $\hat{v}(t)$.

6) t 增 1,转到第 2)步.

注 10 可在式(49)和(50)中分别引入收敛指数 ε 和遗忘因子 λ,就得到修正遗忘因子 AM-MI-ESG 算法.

2.4 辅助模型递推增广最小二乘算法

对于 AR-OEMA 系统辨识模型(33),参照文献[1-3]中 RLS 算法的推导,可以得到估计 AR-OEMA 系统参数向量 $\boldsymbol{\vartheta}$ 的辅助模型递推增广最小二乘算法(Auxiliary Model based Recursive Extended Least Squares algorithm,AM-RELS 算法):

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

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

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

$$\hat{\boldsymbol{\varphi}}(t)=[\boldsymbol{\varphi}_y^{\mathrm{T}}(t),\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t),\hat{v}(t-1),\hat{v}(t-2),\cdots,\hat{v}(t-n_d)]^{\mathrm{T}}, \quad (63)$$

$$\boldsymbol{\varphi}_y(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a)]^{\mathrm{T}}, \quad (64)$$

$$\hat{\boldsymbol{\phi}}(t)=[-x_{\mathrm{a}}(t-1),\cdots,-x_{\mathrm{a}}(t-n_f),u(t-1),\cdots,u(t-n_b)]^{\mathrm{T}}, \quad (65)$$

$$x_{\mathrm{a}}(t)=\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\rho}}(t), \quad (66)$$

$$\hat{v}(t)=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t), \quad (67)$$

$$\hat{\boldsymbol{\vartheta}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{\rho}}^{\mathrm{T}}(t),\hat{\boldsymbol{d}}^{\mathrm{T}}(t)]^{\mathrm{T}}. \quad (68)$$

AM-RELS 算法(60)—(68)计算参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$ 的步骤如下:

1) 初始化:令 $t=1$,置初值 $\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_n/p_0$,$\boldsymbol{P}(0)=p_0\boldsymbol{I}_n$,$x_{\mathrm{a}}(t-i)=1/p_0$,$\hat{v}(t-i)=1/p_0$,$i=1,2,\cdots,\max[n_f,n_d]$,$p_0=10^6$.

2) 采集输入输出数据 $u(t)$ 和 $y(t)$,由式(64)—(65)和(63)构造信息向量 $\boldsymbol{\varphi}_y(t)$,$\hat{\boldsymbol{\phi}}(t)$ 和 $\hat{\boldsymbol{\varphi}}(t)$.

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

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

5) 从式(68)的 $\hat{\boldsymbol{\vartheta}}(t)$ 中读取 $\hat{\boldsymbol{a}}(t)$,$\hat{\boldsymbol{\rho}}(t)$ 和 $\hat{\boldsymbol{d}}(t)$. 由式(66)—(67)计算辅助模型的输出 $x_{\mathrm{a}}(t)$ 和 $\hat{v}(t)$.

6) t 增 1,转到第 2)步.

注 11 本节的辨识方法可以推广到多输入 AR-OEMA 模型描述的动态随机系统:

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

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

和非线性自回归输出误差滑动平均模型(N-AR-OEMA 模型)描述的非线性动态随机系统:

3) $A(z)f(y(t))=\dfrac{B(z)}{F(z)}g(u(t))+D(z)v(t)$.

3 自回归输出误差自回归系统

《系统辨识——多新息辨识理论与方法》[3] 介绍了输出误差滑自回归(OEAR)系统的辅助模型辨识方法,这里简单给出自回归 OEAR 系统的辅助模型辨识方法.

3.1 AR-OEAR 系统描述与辨识模型

考虑下列自回归输出误差自回归模型(AutoRegressive Output-Error AutoRegressive model,AR-OEAR 模型)描述的动态随机系统(参见图 3):

$$A(z)y(t)=\frac{B(z)}{F(z)}u(t)+\frac{1}{C(z)}v(t), \quad (69)$$

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 $F(z)$ 均为单位后移算子 z^{-1} 的多项式:

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

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

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

$$F(z):=1+f_1z^{-1}+f_2z^{-2}+\cdots+f_{n_f}z^{-n_f}.$$

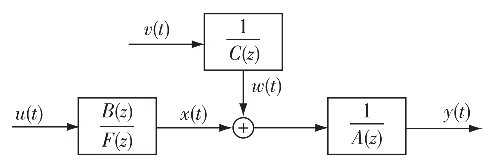

图 3 自回归输出误差自回归系统

Fig. 3 The AR-OEAR system

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

定义未知中间变量(intermediate variable)$x(t)$ 和不可测噪声项 $w(t)$ 分别为

$$x(t):=\frac{B(z)}{F(z)}u(t), \quad (70)$$

$$w(t):=\frac{1}{C(z)}v(t). \quad (71)$$

AR-OEAR 系统的干扰噪声 $w(t)$ 是一个自回归(AR)过程,是一个有色干扰噪声.

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

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

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

$\boldsymbol{\rho}:=[f_1,f_2,\cdots,f_{n_f},b_1,b_2,\cdots,b_{n_b}]^{\mathrm{T}}\in\mathbf{R}^{n_f+n_b}$,

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

$\boldsymbol{\varphi}(t):=[\boldsymbol{\varphi}_y^{\mathrm{T}}(t),\boldsymbol{\phi}^{\mathrm{T}}(t),-w(t-1),-w(t-2),\cdots,-w(t-n_c)]^{\mathrm{T}}\in\mathbf{R}^{n}$,

$\boldsymbol{\varphi}_y(t):=[-y(t-1),-y(t-2),\cdots,-y(t-n_a)]^{\mathrm{T}}\in\mathbf{R}^{n_a}$,

$\boldsymbol{\phi}(t):=[-x(t-1),-x(t-2),\cdots,-x(t-n_f),u(t-1),u(t-2),\cdots,u(t-n_b)]^{\mathrm{T}}\in\mathbf{R}^{n_f+n_b}$.

设 $\hat{\boldsymbol{\vartheta}}(t):=\begin{bmatrix}\hat{\boldsymbol{a}}(t)\\ \hat{\boldsymbol{\rho}}(t)\\ \hat{\boldsymbol{c}}(t)\end{bmatrix}\in\mathbf{R}^{n}$ 是 $\boldsymbol{\vartheta}=\begin{bmatrix}\boldsymbol{a}\\ \boldsymbol{\rho}\\ \boldsymbol{c}\end{bmatrix}$在时刻 t 的估计.

由式(70)—(71)可得

$$x(t)=[1-F(z)]x(t)+B(z)u(t)=\boldsymbol{\phi}^{\mathrm{T}}(t)\boldsymbol{\rho},\tag{72}$$

$$w(t)=[1-C(z)]w(t)+v(t).\tag{73}$$

将式(70)和(71)代入式(69)可得

$$\begin{aligned}y(t)&=[1-A(z)]y(t)+x(t)+w(t)\\&=\boldsymbol{\varphi}_y^{\mathrm{T}}(t)\boldsymbol{a}+x(t)+w(t)&(74)\\&=\boldsymbol{\varphi}_y^{\mathrm{T}}(t)\boldsymbol{a}+\boldsymbol{\phi}^{\mathrm{T}}(t)\boldsymbol{\rho}+[1-C(z)]w(t)+v(t)\\&=\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\vartheta}+v(t).&(75)\end{aligned}$$

系统的可测输入输出数据为$\{u(t),y(t)\}$,中间变量 $x(t)$ 和噪声 $w(t)$ 是不可测的.因此,信息向量 $\boldsymbol{\varphi}(t)$ 是未知的.按照第 2 节的辅助模型辨识思想:用输出信息向量 $\boldsymbol{\varphi}_y(t)$,辅助模型的输出 $x_{\mathrm{a}}(t-i)$ 和系统输入 $u(t-i)$,以及噪声 $w(t-i)$ 的估计 $\hat{w}(t-i)$ 构造 $\boldsymbol{\varphi}(t)$ 和 $\boldsymbol{\phi}(t)$ 的估计:

$$\hat{\boldsymbol{\varphi}}(t)=[\boldsymbol{\varphi}_y^{\mathrm{T}}(t),\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t),-\hat{w}(t-1),-\hat{w}(t-2),\cdots,-\hat{w}(t-n_c)]^{\mathrm{T}}\in\mathbf{R}^{n},\tag{76}$$

$$\hat{\boldsymbol{\phi}}(t):=[-x_{\mathrm{a}}(t-1),-x_{\mathrm{a}}(t-2),\cdots,-x_{\mathrm{a}}(t-n_f),u(t-1),u(t-2),\cdots,u(t-n_b)]^{\mathrm{T}}\in\mathbf{R}^{n_f+n_b}.\tag{77}$$

用 $\hat{\boldsymbol{\phi}}(t)$ 和 $\hat{\boldsymbol{\rho}}(t)$ 构造估算 $x(t)$ 的辅助模型

$$x_{\mathrm{a}}(t)=\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\rho}}(t).\tag{78}$$

$x_{\mathrm{a}}(t)$ 可作为 $x(t)$ 的估计,即 $\hat{x}(t)=x_{\mathrm{a}}(t)$.根据式(74)可得计算 $w(t)$ 的估计 $\hat{w}(t)$ 的辅助模型:

$$\hat{w}(t)=y(t)-\boldsymbol{\varphi}_y^{\mathrm{T}}(t)\hat{\boldsymbol{a}}(t)-x_{\mathrm{a}}(t).\tag{79}$$

3.2 辅助模型广义随机梯度算法

根据辨识模型(75),可以得到下列梯度递推关系:

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\frac{\boldsymbol{\varphi}(t)}{r(t)}[y(t)-\boldsymbol{\varphi}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t-1)],\tag{80}$$

$$r(t)=r(t-1)+\|\boldsymbol{\varphi}(t)\|^2,\quad r(0)=1.\tag{81}$$

式(80)—(81)中未知的 $\boldsymbol{\varphi}(t)$ 用其估计 $\hat{\boldsymbol{\varphi}}(t)$ 代替,可以总结出估计 AR-OEAR 系统参数向量 $\boldsymbol{\vartheta}$ 的辅助模型广义随机梯度算法(Auxiliary Model based Generalized Stochastic Gradient algorithm,AM-GSG 算法):

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\frac{\hat{\boldsymbol{\varphi}}(t)}{r(t)}e(t),\tag{82}$$

$$e(t)=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t-1),\tag{83}$$

$$r(t)=r(t-1)+\|\hat{\boldsymbol{\varphi}}(t)\|^2,\tag{84}$$

$$\hat{\boldsymbol{\varphi}}(t)=[\boldsymbol{\varphi}_y^{\mathrm{T}}(t),\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t),-\hat{w}(t-1),-\hat{w}(t-2),\cdots,-\hat{w}(t-n_c)]^{\mathrm{T}},\tag{85}$$

$$\boldsymbol{\varphi}_y(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a)]^{\mathrm{T}},\tag{86}$$

$$\hat{\boldsymbol{\phi}}(t)=[-x_{\mathrm{a}}(t-1),\cdots,-x_{\mathrm{a}}(t-n_f),u(t-1),\cdots,u(t-n_b)]^{\mathrm{T}},\tag{87}$$

$$x_{\mathrm{a}}(t)=\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\rho}}(t),\tag{88}$$

$$\hat{w}(t)=y(t)-\boldsymbol{\varphi}_y^{\mathrm{T}}(t)\hat{\boldsymbol{a}}(t)-x_{\mathrm{a}}(t),\tag{89}$$

$$\hat{\boldsymbol{\vartheta}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\rho}^{\mathrm{T}}(t),\hat{\boldsymbol{c}}^{\mathrm{T}}(t)]^{\mathrm{T}}.\tag{90}$$

AM-GSG 算法(82)—(90)计算参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$ 的步骤如下:

1) 初始化:令 $t=1$.置初值 $\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_n/p_0$,$r(0)=1$,$x_{\mathrm{a}}(t-i)=1/p_0$,$\hat{w}(t-i)=1/p_0$,$i=1,2,\cdots,\max[n_f,n_c]$,$p_0=10^6$.

2) 采集输入输出数据 $u(t)$ 和 $y(t)$,由式(86)—(87)和(85)构造信息向量 $\boldsymbol{\varphi}_y(t)$,$\hat{\boldsymbol{\phi}}(t)$ 和 $\hat{\boldsymbol{\varphi}}(t)$.

3) 由式(83)计算新息 $e(t)$,由式(84)计算 $r(t)$.

4) 根据式(82)刷新参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$.

5) 从式(90)的 $\hat{\boldsymbol{\vartheta}}(t)$ 中读取 $\hat{\boldsymbol{a}}(t)$,$\hat{\rho}(t)$ 和 $\hat{\boldsymbol{c}}(t)$.由式(88)—(89)计算辅助模型的输出 $x_{\mathrm{a}}(t)$ 和 $\hat{w}(t)$.

6) t 增 1,转到第 2)步.

注 12 可在式(84)中引入遗忘因子 λ,得到遗忘因子辅助模型广义随机梯度算法(Forgetting Factor AM-GSG algorithm,FF-AM-GSG 算法).可在式(82)中引入收敛指数 ε,得到修正辅助模型广义随机梯度算法(Modified AM-GSG algorithm,M-AM-GSG 算法).

3.3 辅助模型多新息广义随机梯度算法

借助于多新息辨识理论,基于 AM-GSG 算法(82)—(90),将系统输出 $y(t)$ 和信息向量 $\hat{\boldsymbol{\varphi}}(t)$ 扩展为堆积输出向量 $\boldsymbol{Y}(p,t)$ 和堆积信息矩阵 $\hat{\boldsymbol{\Phi}}(p,t)$,将标量新息 $e(t)\in\mathbf{R}$ 扩展为新息向量,就可以得到估计 AR-OEAR 系统参数向量 $\boldsymbol{\vartheta}$ 的辅助模型多新息广义随机梯度算法(Auxiliary Model based Multi-Innovation Generalized Stochastic Gradient algorithm,AM-MI-GSG 算法):

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\frac{\hat{\boldsymbol{\Phi}}(p,t)}{r(t)}\boldsymbol{E}(p,t),\tag{91}$$

$$r(t)=r(t-1)+\|\hat{\boldsymbol{\varphi}}(t)\|^2,\tag{92}$$

$$\boldsymbol{E}(p,t)=\boldsymbol{y}(p,t)-\hat{\boldsymbol{\Phi}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\vartheta}}(t-1), \tag{93}$$
$$\boldsymbol{Y}(p,t)=[y(t),y(t-1),\cdots,y(t-p+1)]^{\mathrm{T}}, \tag{94}$$
$$\hat{\boldsymbol{\Phi}}(p,t)=[\hat{\boldsymbol{\varphi}}(t),\hat{\boldsymbol{\varphi}}(t-1),\cdots,\hat{\boldsymbol{\varphi}}(t-p+1)], \tag{95}$$
$$\hat{\boldsymbol{\varphi}}(t)=[\boldsymbol{\varphi}_y^{\mathrm{T}}(t),\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t),-\hat{w}(t-1),-\hat{w}(t-2),\cdots,-\hat{w}(t-n_c)]^{\mathrm{T}}, \tag{96}$$
$$\boldsymbol{\varphi}_y(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a)]^{\mathrm{T}}, \tag{97}$$
$$\hat{\boldsymbol{\phi}}(t)=[-x_{\mathrm{a}}(t-1),\cdots,-x_{\mathrm{a}}(t-n_f),u(t-1),\cdots,u(t-n_b)]^{\mathrm{T}}, \tag{98}$$
$$x_{\mathrm{a}}(t)=\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\rho}}(t), \tag{99}$$
$$\hat{w}(t)=y(t)-\boldsymbol{\varphi}_y^{\mathrm{T}}(t)\hat{\boldsymbol{a}}(t)-x_{\mathrm{a}}(t), \tag{100}$$
$$\hat{\boldsymbol{\vartheta}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{\rho}}^{\mathrm{T}}(t),\hat{\boldsymbol{c}}^{\mathrm{T}}(t)]^{\mathrm{T}}. \tag{101}$$

当新息长度 $p=1$ 时,AM-MI-GSG 算法退化为 AM-GSG 算法(82)—(90).

AM-MI-GSG 算法(91)—(101)计算参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$ 的步骤如下:

1) 初始化:令 $t=1$,给定新息长度 p.置 $\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_n/p_0$,$r(0)=1$,$x_{\mathrm{a}}(t-i)=1/p_0$,$\hat{w}(t-i)=1/p_0$,$i=1,2,\cdots,\max[n_f,n_c]$,$p_0=10^6$.

2) 采集输入输出数据 $u(t)$ 和 $y(t)$,由式(94)构造堆积输出向量 $\boldsymbol{Y}(p,t)$,由式(97)—(98)和(96)构造信息向量 $\boldsymbol{\varphi}_y(t)$,$\hat{\boldsymbol{\phi}}(t)$ 和 $\hat{\boldsymbol{\varphi}}(t)$,由式(95)构造堆积信息矩阵 $\hat{\boldsymbol{\Phi}}(p,t)$.

3) 由式(93)计算新息向量 $\boldsymbol{E}(p,t)$,由式(92)计算 $r(t)$.

4) 根据式(91)刷新参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$.

5) 从式(101)的 $\hat{\boldsymbol{\vartheta}}(t)$ 中读取 $\hat{\boldsymbol{a}}(t)$,$\hat{\boldsymbol{\rho}}(t)$ 和 $\hat{\boldsymbol{c}}(t)$.由式(99)—(100)计算辅助模型的输出 $x_{\mathrm{a}}(t)$ 和 $\hat{w}(t)$.

6) t 增 1,转到第 2)步.

注 13 可在式(91)和(92)中分别引入收敛指数 ε 和遗忘因子 λ,就得到修正遗忘因子 AM-MI-GSG 算法.

3.4 辅助模型递推广义最小二乘算法

对于 AR-OEAR 系统辨识模型(75),可以得到估计 AR-OEAR 系统参数向量 $\boldsymbol{\vartheta}$ 的辅助模型递推广义最小二乘算法(Auxiliary Model based Recursive Generalized Least Squares algorithm, AM-RGLS 算法):

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\boldsymbol{L}(t)[y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t-1)], \tag{102}$$
$$\boldsymbol{L}(t)=\boldsymbol{P}(t-1)\hat{\boldsymbol{\varphi}}(t)[1+\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\boldsymbol{P}(t-1)\hat{\boldsymbol{\varphi}}(t)]^{-1}, \tag{103}$$
$$\boldsymbol{P}(t)=[\boldsymbol{I}_n-\boldsymbol{L}(t)\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)]\boldsymbol{P}(t-1), \tag{104}$$
$$\hat{\boldsymbol{\varphi}}(t)=[\boldsymbol{\varphi}_y^{\mathrm{T}}(t),\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t),-\hat{w}(t-1),-\hat{w}(t-2),\cdots,-\hat{w}(t-n_c)]^{\mathrm{T}}, \tag{105}$$
$$\boldsymbol{\varphi}_y(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a)]^{\mathrm{T}}, \tag{106}$$
$$\hat{\boldsymbol{\phi}}(t)=[-x_{\mathrm{a}}(t-1),\cdots,-x_{\mathrm{a}}(t-n_f),u(t-1),\cdots,u(t-n_b)]^{\mathrm{T}}, \tag{107}$$
$$x_{\mathrm{a}}(t)=\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t)\hat{\rho}(t), \tag{108}$$
$$\hat{w}(t)=y(t)-\boldsymbol{\varphi}_y^{\mathrm{T}}(t)\hat{\boldsymbol{a}}(t)-x_{\mathrm{a}}(t), \tag{109}$$
$$\hat{\boldsymbol{\vartheta}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{\rho}}^{\mathrm{T}}(t),\hat{\boldsymbol{c}}^{\mathrm{T}}(t)]^{\mathrm{T}}. \tag{110}$$

AM-RGLS 算法(102)—(110)计算参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$ 的步骤如下.

1) 初始化:令 $t=1$.置初值 $\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_n/p_0$,$\boldsymbol{P}(0)=p_0\boldsymbol{I}_n$,$x_{\mathrm{a}}(t-i)=1/p_0$,$\hat{w}(t-i)=1/p_0$,$i=1,2,\cdots,\max[n_f,n_c]$,$p_0=10^6$.

2) 采集输入输出数据 $u(t)$ 和 $y(t)$,由式(106)—(107)和(105)构造信息向量 $\boldsymbol{\varphi}_y(t)$,$\hat{\boldsymbol{\phi}}(t)$ 和 $\hat{\boldsymbol{\varphi}}(t)$.

3) 由式(103)计算增益向量 $\boldsymbol{L}(t)$,由式(104)计算协方差阵 $\boldsymbol{P}(t)$.

4) 根据式(102)刷新参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$.

5) 从式(110)的 $\hat{\boldsymbol{\vartheta}}(t)$ 中读取 $\hat{\boldsymbol{a}}(t)$,$\hat{\boldsymbol{\rho}}(t)$ 和 $\hat{\boldsymbol{c}}(t)$.由式(108)—(109)计算辅助模型的输出 $x_{\mathrm{a}}(t)$ 和 $\hat{w}(t)$.

6) t 增 1,转到第 2)步.

注 14 本节的辨识方法可以推广到多输入 AR-OEAR 模型描述的动态随机系统:

1) $A(z)y(t)=\dfrac{1}{F(z)}\sum\limits_{j=1}^{r}B_j(z)u_j(t)+\dfrac{1}{C(z)}v(t)$,

2) $A(z)y(t)=\sum\limits_{j=1}^{r}\dfrac{B_j(z)}{F_j(z)}u_j(t)+\dfrac{1}{C(z)}v(t)$

和非线性自回归输出误差自回归模型(N-AR-OEAR 模型)描述的非线性动态随机系统:

3) $A(z)f(y(t))=\dfrac{B(z)}{F(z)}g(u(t))+\dfrac{1}{C(z)}v(t)$.

4 回归输出误差自回归滑动平均系统

自回归 Box-Jenkins 系统是 Box-Jenkins 系统的推广,是单输入单输出随机系统的一般形式.文献[3,8]详细介绍了输出误差自回归滑动平均(OEARMA)系统(即 Box-Jenkins 系统)的辅助模型辨识方法,这里简单介绍自回归 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)算法等.

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 期上[49].该文获得《IET Journals》的最佳论文奖"Premium (Best Paper)

Awards”,该奖是《IET Journals》每年从前两年发表论文中评选出的唯一一篇最佳论文(http://digital-library.theiet.org/journals/premium-awards).

4.1 AR-BJ 系统描述与辨识模型

考虑下列自回归 Box-Jenkins 模型(AutoRegressive Box-Jenkins model,AR-BJ 模型)描述的动态随机系统(参见图 4):

$$A(z)y(t)=\frac{B(z)}{F(z)}u(t)+\frac{D(z)}{C(z)}v(t), \tag{111}$$

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),$A(z)$,$B(z)$,$C(z)$,$D(z)$和$F(z)$均为单位后移算子z^{-1}的多项式:

$$A(z)=1+a_1z^{-1}+a_2z^{-2}+\cdots+a_{n_a}z^{-n_a},$$
$$B(z)=b_1z^{-1}+b_2z^{-2}+\cdots+b_{n_b}z^{-n_b},$$
$$C(z)=1+c_1z^{-1}+c_2z^{-2}+\cdots+c_{n_c}z^{-n_c},$$
$$D(z)=1+d_1z^{-1}+d_2z^{-2}+\cdots+d_{n_d}z^{-n_d},$$
$$F(z):=1+f_1z^{-1}+f_2z^{-2}+\cdots+f_{n_f}z^{-n_f}.$$

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

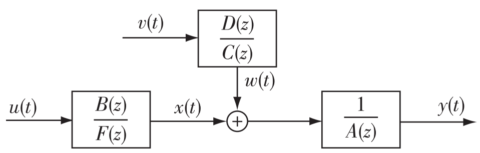

图 4 AR-BJ 系统

Fig. 4 The AR-BJ system

AR-BJ 模型又称为自回归输出误差自回归滑动平均模型(AutoRegressive Output-Error AutoRegressive Moving Average model,AR-OEARMA 模型).

定义未知中间变量(intermediate variable)$x(t)$和不可测噪声项$w(t)$分别为

$$x(t):=\frac{B(z)}{F(z)}u(t), \tag{112}$$

$$w(t):=\frac{D(z)}{C(z)}v(t). \tag{113}$$

AR-BJ 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$\boldsymbol{\vartheta}$和信息向量$\boldsymbol{\varphi}(t)$如下:

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

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

$$\boldsymbol{\rho}:=[f_1,f_2,\cdots,f_{n_f},b_1,b_2,\cdots,b_{n_b}]^{\mathrm{T}}\in\mathbf{R}^{n_f+n_b},$$

$$\boldsymbol{\theta}:=[c_1,c_2,\cdots,c_{n_c},d_1,d_2,\cdots,d_{n_d}]^{\mathrm{T}}\in\mathbf{R}^{n_c+n_d},$$

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

$$\boldsymbol{\varphi}_y(t):=[-y(t-1),-y(t-2),\cdots,-y(t-n_a)]^{\mathrm{T}}\in\mathbf{R}^{n_a},$$

$$\boldsymbol{\phi}(t):=[-x(t-1),-x(t-2),\cdots,-x(t-n_f),$$
$$u(t-1),u(t-2),\cdots,u(t-n_b)]^{\mathrm{T}}\in\mathbf{R}^{n_f+n_b},$$

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

系统的可测输入输出数据为$\{u(t),y(t)\}$,中间变量$x(t)$,$w(t)$和白噪声$v(t)$是不可测的,是未知的.因此,信息向量$\boldsymbol{\varphi}_y(t)$是可测的,信息向量$\boldsymbol{\phi}(t)$和$\boldsymbol{\psi}(t)$是未知的.

由式(112)—(113)可得

$$x(t)=[1-F(z)]x(t)+B(z)u(t)=\boldsymbol{\phi}^{\mathrm{T}}(t)\boldsymbol{\rho}, \tag{114}$$

$$w(t)=[1-C(z)]w(t)+D(z)v(t)=\boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\theta}+v(t), \tag{115}$$

将式(112)和(113)代入式(111)可得

$$\begin{aligned}y(t)&=[1-A(z)]y(t)+x(t)+w(t)=\\ &\boldsymbol{\varphi}_y^{\mathrm{T}}(t)\boldsymbol{a}+x(t)+w(t)= \qquad (116)\\ &\boldsymbol{\varphi}_y^{\mathrm{T}}(t)\boldsymbol{a}+\boldsymbol{\phi}^{\mathrm{T}}(t)\boldsymbol{\rho}+\boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\theta}+v(t)=\\ &[\boldsymbol{\varphi}_y^{\mathrm{T}}(t),\boldsymbol{\phi}^{\mathrm{T}}(t),\boldsymbol{\psi}^{\mathrm{T}}(t)]\begin{bmatrix}\boldsymbol{a}\\ \boldsymbol{\rho}\\ \boldsymbol{\theta}\end{bmatrix}+v(t)=\\ &\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\vartheta}+v(t). \qquad (117)\end{aligned}$$

由于信息向量$\boldsymbol{\varphi}(t)$包含了未知的$\boldsymbol{\phi}(t)$和$\boldsymbol{\psi}(t)$,所以必须借助于辅助模型辨识思想,用辅助模型的输出来构造这些未知向量的估计,从而提出基于辅助模型的递推辨识方法.

4.2 辅助模型的建立

设$\hat{\boldsymbol{\vartheta}}(t):=\begin{bmatrix}\hat{\boldsymbol{a}}(t)\\ \hat{\boldsymbol{\rho}}(t)\\ \hat{\boldsymbol{\theta}}(t)\end{bmatrix}\in\mathbf{R}^n$ 是 $\boldsymbol{\vartheta}=\begin{bmatrix}\boldsymbol{a}\\ \boldsymbol{\rho}\\ \boldsymbol{\theta}\end{bmatrix}$在时刻$t$的估计.

辨识模型(117)中信息向量$\boldsymbol{\varphi}(t)$包含未知变量$x(t-i)$和不可测噪声项$w(t-i)$和$v(t-i)$,这是辨识的困难.一种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)建立辅助模型,用辅助模型的输出$x_a(t-i)$,$\hat{w}(t-i)$和$\hat{v}(t-i)$代替信息向量$\boldsymbol{\varphi}(t)$中的未知项$x(t-i)$,$w(t-i)$和$v(t-i)$.具体而言,构造一个辅助模型$\boldsymbol{P}_{\mathrm{a}}(z)$,用系统输入$u(t-i)$和辅助模型的输出$x_{\mathrm{a}}(t-i)$构造$\boldsymbol{\phi}(t)$的估计

$$\hat{\boldsymbol{\phi}}(t):=[-x_{\mathrm{a}}(t-1),-x_{\mathrm{a}}(t-2),\cdots,-x_{\mathrm{a}}(t-n_f),u(t-1),u(t-2),\cdots,u(t-n_b)]^{\mathrm{T}}\in\mathbf{R}^{n_f+n_b}.$$

取估算 $x(t)$ 的辅助模型为

$$x_{\mathrm{a}}(t)=\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\rho}}(t).$$

$x_{\mathrm{a}}(t)$ 可作为 $x(t)$ 的估计,即 $\hat{x}(t)=x_{\mathrm{a}}(t)$.用 $\boldsymbol{\varphi}_y(t)$ 和 $\hat{\boldsymbol{\phi}}(t)$,以及辅助模型的输出 $\hat{w}(t-i)$ 和 $\hat{v}(t-i)$ 构造 $\boldsymbol{\varphi}(t)$ 的估计

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

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

根据式(116)和(117),可以得到计算 $w(t)$ 和 $v(t)$ 的估计 $\hat{w}(t)$ 和 $\hat{v}(t)$ 的辅助模型:

$$\hat{w}(t)=y(t)-\boldsymbol{\varphi}_y^{\mathrm{T}}(t)\hat{\boldsymbol{a}}(t)-x_{\mathrm{a}}(t)=y(t)-\boldsymbol{\varphi}_y^{\mathrm{T}}(t)\hat{\boldsymbol{a}}(t)-\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\rho}}(t),$$

$$\hat{v}(t)=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t)=\hat{w}(t)-\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t).$$

上式也可以直接根据式(115),未知量 $w(t)$,$\boldsymbol{\psi}(t)$ 和 $\boldsymbol{\theta}$ 分别用其估计 $\hat{w}(t)$,$\hat{\boldsymbol{\psi}}(t)$ 和 $\hat{\boldsymbol{\theta}}(t)$ 代替得到.

4.3 辅助模型广义增广随机梯度算法

根据辨识模型(117)和辅助模型辨识思想,可以总结出估计 AR-BJ 系统参数向量 $\boldsymbol{\vartheta}$ 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):

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\frac{\hat{\boldsymbol{\varphi}}(t)}{r(t)}e(t),\tag{118}$$

$$e(t)=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t-1),\tag{119}$$

$$r(t)=r(t-1)+\|\hat{\boldsymbol{\varphi}}(t)\|^2,\tag{120}$$

$$\hat{\boldsymbol{\varphi}}(t)=[\boldsymbol{\varphi}_y^{\mathrm{T}}(t),\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t),\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)]^{\mathrm{T}},\tag{121}$$

$$\boldsymbol{\varphi}_y(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a)]^{\mathrm{T}},\tag{122}$$

$$\hat{\boldsymbol{\phi}}(t)=[-x_{\mathrm{a}}(t-1),\cdots,-x_{\mathrm{a}}(t-n_f),u(t-1),\cdots,u(t-n_b)]^{\mathrm{T}},\tag{123}$$

$$\hat{\boldsymbol{\psi}}(t)=[-\hat{w}(t-1),\cdots,-\hat{w}(t-n_c),\hat{v}(t-1),\cdots,\hat{v}(t-n_d)]^{\mathrm{T}},\tag{124}$$

$$x_{\mathrm{a}}(t)=\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\rho}}(t),\tag{125}$$

$$\hat{w}(t)=y(t)-\boldsymbol{\varphi}_y^{\mathrm{T}}(t)\hat{\boldsymbol{a}}(t)-x_{\mathrm{a}}(t),\tag{126}$$

$$\hat{v}(t)=\hat{w}(t)-\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t),\tag{127}$$

$$\hat{\boldsymbol{\vartheta}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{\rho}}^{\mathrm{T}}(t),\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t)]^{\mathrm{T}}.\tag{128}$$

AM-GESG 算法(118)—(128)计算参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$ 的步骤如下:

1) 初始化:令 $t=1$.置初值 $\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_n/p_0$,$r(0)=1$,$x_{\mathrm{a}}(t-i)=1/p_0$,$\hat{w}(t-i)=1/p_0$,$\hat{v}(t-i)=1/p_0$,$i=1,2,\cdots,\max[n_f,n_c,n_d]$,$p_0=10^6$,$\mathbf{1}_n$ 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(122)—(124)和(121)构造信息向量 $\boldsymbol{\varphi}_y(t)$,$\hat{\boldsymbol{\phi}}(t)$,$\hat{\boldsymbol{\psi}}(t)$ 和 $\hat{\boldsymbol{\varphi}}(t)$.

3) 由式(119)计算新息 $e(t)$,由式(120)计算 $r(t)$.

4) 根据式(118)刷新参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$.

5) 从式(128)的 $\hat{\boldsymbol{\vartheta}}(t)$ 中读取 $\hat{\boldsymbol{a}}(t)$,$\hat{\boldsymbol{\rho}}(t)$ 和 $\hat{\boldsymbol{\theta}}(t)$.由式(125)—(127)计算辅助模型的输出 $x_{\mathrm{a}}(t)$,$\hat{w}(t)$ 和 $\hat{v}(t)$.

6) t 增 1,转到第 2)步.

注 15 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(forgetting factor)λ,将式(120)修改为

$$r(t)=\lambda\, r(t-1)+\|\hat{\boldsymbol{\varphi}}(t)\|^2,\quad 0\leqslant\lambda\leqslant1,r(0)=1,$$

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 16 为提高 AM-GESG 算法的暂态性能和稳态性能,可引入收敛指数(convergence index)ε,将式(118)修改为

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\frac{\hat{\boldsymbol{\varphi}}(t)}{r^{\varepsilon}(t)}e(t),\quad \frac{1}{2}<\varepsilon\leqslant1,$$

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).

4.4 辅助模型多新息广义增广随机梯度算法

下面借助于多新息辨识理论,将 AM-GESG 算法(118)—(128)中标量新息 $e(t)\in\mathbf{R}$ 扩展为新息向量 $\boldsymbol{E}(p,t)$,将系统输出 $y(t)$ 和信息向量 $\hat{\boldsymbol{\varphi}}(t)$ 扩展为堆积输出向量 $\boldsymbol{y}(p,t)$ 和堆积信息矩阵 $\hat{\boldsymbol{\Phi}}(p,t)$,可以得到估计 AR-BJ 系统参数向量 $\boldsymbol{\theta}$ 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):

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\frac{\hat{\boldsymbol{\Phi}}(p,t)}{r(t)}\boldsymbol{E}(p,t),\tag{129}$$

$$r(t)=r(t-1)+\|\hat{\boldsymbol{\varphi}}(t)\|^2,\tag{130}$$

$$\boldsymbol{E}(p,t)=\boldsymbol{Y}(p,t)-\hat{\boldsymbol{\Phi}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\vartheta}}(t-1),\tag{131}$$

$$\boldsymbol{Y}(p,t)=[y(t),y(t-1),\cdots,y(t-p+1)]^{\mathrm{T}},\tag{132}$$

$$\hat{\boldsymbol{\Phi}}(p,t)=[\hat{\boldsymbol{\varphi}}(t),\hat{\boldsymbol{\varphi}}(t-1),\cdots,\hat{\boldsymbol{\varphi}}(t-p+1)],\tag{133}$$

$$\hat{\boldsymbol{\varphi}}(t)=[\boldsymbol{\varphi}_y^{\mathrm{T}}(t),\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t),\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)]^{\mathrm{T}},\tag{134}$$

$$\boldsymbol{\varphi}_y(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a)]^{\mathrm{T}},\tag{135}$$

$$\hat{\boldsymbol{\phi}}(t)=[-x_a(t-1),\cdots,-x_a(t-n_f),u(t-1),\cdots,u(t-n_b)]^{\mathrm{T}},\tag{136}$$
$$\hat{\boldsymbol{\psi}}(t)=[-\hat{w}(t-1),\cdots,-\hat{w}(t-n_c),\hat{v}(t-1),\cdots,\hat{v}(t-n_d)]^{\mathrm{T}},\tag{137}$$
$$x_a(t)=\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\rho}}(t),\tag{138}$$
$$\hat{w}(t)=y(t)-\boldsymbol{\phi}_y^{\mathrm{T}}(t)\hat{\boldsymbol{a}}(t)-x_a(t),\tag{139}$$
$$\hat{v}(t)=\hat{w}(t)-\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t),\tag{140}$$
$$\hat{\boldsymbol{\vartheta}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{\rho}}^{\mathrm{T}}(t),\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t)]^{\mathrm{T}}.\tag{141}$$

当新息长度 $p=1$ 时,AM-MI-GESG 算法退化为 AM-GESG 算法(118)—(128).

AM-MI-GESG 算法(129)—(141)计算参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$ 的步骤如下:

1) 初始化:令 $t=1$,给定新息长度 p.置 $\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_n/p_0$,$r(0)=1$,$x_a(t-i)=1/p_0$,$\hat{w}(t-i)=1/p_0$,$\hat{v}(t-i)=1/p_0$,$i=1,2,\cdots,\max[n_f,n_c,n_d]$,$p_0=10^6$.

2) 采集输入输出数据 $u(t)$ 和 $y(t)$,由式(132)构造堆积输出向量 $\boldsymbol{Y}(p,t)$,由式(135)—(137)和(134)构造信息向量 $\boldsymbol{\varphi}_y(t)$,$\hat{\boldsymbol{\phi}}(t)$,$\hat{\boldsymbol{\psi}}(t)$ 和 $\hat{\boldsymbol{\varphi}}(t)$,由式(133)构造堆积信息矩阵 $\hat{\boldsymbol{\Phi}}(p,t)$.

3) 由式(131)计算新息向量 $\boldsymbol{E}(p,t)$,由式(130)计算 $r(t)$.

4) 根据式(129)刷新参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$.

5) 从式(141)的 $\hat{\boldsymbol{\vartheta}}(t)$ 中读取 $\hat{\boldsymbol{a}}(t)$,$\hat{\boldsymbol{\rho}}(t)$ 和 $\hat{\boldsymbol{\theta}}(t)$.由式(138)—(140)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 17 在式(129)和(130)中分别引入收敛指数 ε 和遗忘因子 λ,即

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\frac{\hat{\boldsymbol{\Phi}}(p,t)}{r^{\varepsilon}(t)}\boldsymbol{E}(p,t),\quad \frac{1}{2}<\varepsilon\leqslant 1,\tag{142}$$
$$r(t)=\lambda r(t-1)+\|\hat{\boldsymbol{\varphi}}(t)\|^2,\quad 0\leqslant\lambda\leqslant 1,\tag{143}$$

就得到修正遗忘因子 AM-MI-GESG 算法(131)—(143).式(143)也可修改为

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

4.5 辅助模型递推广义增广最小二乘算法

对于 AR-BJ 系统辨识模型(117),定义准则函数(criterion function):

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

参照文献[1-3]中 RLS 算法的推导,使用辅助模型辨识思想,可以得到估计 AR-BJ 系统参数向量 $\boldsymbol{\vartheta}$ 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):

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\boldsymbol{L}(t)[y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t-1)],\tag{144}$$
$$\boldsymbol{L}(t)=\boldsymbol{P}(t-1)\hat{\boldsymbol{\varphi}}(t)[1+\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\boldsymbol{P}(t-1)\hat{\boldsymbol{\varphi}}(t)]^{-1},\tag{145}$$
$$\boldsymbol{P}(t)=[\boldsymbol{I}_n-\boldsymbol{L}(t)\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)]\boldsymbol{P}(t-1),\tag{146}$$
$$\hat{\boldsymbol{\varphi}}(t)=[\boldsymbol{\varphi}_y^{\mathrm{T}}(t),\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t),\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)]^{\mathrm{T}},\tag{147}$$
$$\boldsymbol{\varphi}_y(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a)]^{\mathrm{T}},\tag{148}$$
$$\hat{\boldsymbol{\phi}}(t)=[-x_a(t-1),\cdots,-x_a(t-n_f),u(t-1),\cdots,u(t-n_b)]^{\mathrm{T}},\tag{149}$$
$$\hat{\boldsymbol{\psi}}(t)=[-\hat{w}(t-1),\cdots,-\hat{w}(t-n_c),\hat{v}(t-1),\cdots,\hat{v}(t-n_d)]^{\mathrm{T}},\tag{150}$$
$$x_a(t)=\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\rho}}(t),\tag{151}$$
$$\hat{w}(t)=y(t)-\boldsymbol{\varphi}_y^{\mathrm{T}}(t)\hat{\boldsymbol{a}}(t)-x_a(t),\tag{152}$$
$$\hat{v}(t)=\hat{w}(t)-\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t),\tag{153}$$
$$\hat{\boldsymbol{\vartheta}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{\rho}}^{\mathrm{T}}(t),\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t)]^{\mathrm{T}}.\tag{154}$$

AM-RGELS 算法(144)—(154)计算参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$ 的步骤如下:

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

2) 采集输入输出数据 $u(t)$ 和 $y(t)$,由式(148)—(150)和(147)构造信息向量 $\boldsymbol{\varphi}_y(t)$,$\hat{\boldsymbol{\phi}}(t)$,$\hat{\boldsymbol{\psi}}(t)$ 和 $\hat{\boldsymbol{\varphi}}(t)$.

3) 由式(145)计算增益向量 $\boldsymbol{L}(t)$,由式(146)计算协方差阵 $\boldsymbol{P}(t)$.

4) 根据式(144)刷新参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$.

5) 从式(154)的 $\hat{\boldsymbol{\vartheta}}(t)$ 中读取 $\hat{\boldsymbol{a}}(t)$,$\hat{\boldsymbol{\rho}}(t)$ 和 $\hat{\boldsymbol{\theta}}(t)$.由式(151)—(153)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 18 Box-Jenkins 系统是 AR-BJ 系统的特例.在 AR-BJ 系统辨识算法中令 $n_a=0$,就得到 Box-Jenkins 系统的辨识方法,如 Box-Jenkins 系统的辅助模型多新息广义增广随机梯度算法[18]、辅助模型最小二乘迭代算法[32]和辅助模型梯度迭代算法[33],非均匀采样数据 Box-Jenkins 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[31].

本节的辨识方法可以推广到多输入 AR-BJ 模型描述的动态随机系统:

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

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

和非线性自回归输出误差 ARMA 模型(N-AR-OE-ARMA 模型),即 N-AR-BJ 模型描述的非线性动态随

机系统：

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

5 AR-BJ 系统的滤波辅助模型辨识方法

《系统辨识——多新息辨识理论与方法》[3]详细介绍了 Box-Jenkins 系统的滤波辅助模型辨识方法，这里介绍自回归 Box-Jenkins 系统的基于输入输出数据滤波的辨识方法，包括基于滤波的辅助模型广义增广随机梯度(F-AM-GESG)辨识方法、基于滤波的辅助模型多新息广义增广随机梯度(F-AM-MI-GESG)辨识方法、基于滤波的辅助模型递推广义增广最小二乘(F-AM-RGELS)辨识方法.

5.1 AR-BJ 系统描述与滤波辨识模型

考虑第 4 节的 AR-BJ 模型描述的动态随机系统，重写如下：

$$A(z)y(t)=\frac{B(z)}{F(z)}u(t)+\frac{D(z)}{C(z)}v(t), \tag{155}$$

其中各变量和定义同上.

定义系统真实输出 $x(t)$ 和噪声模型输出 $w(t)$ 分别为

$$x(t):=\frac{B(z)}{F(z)}u(t)\in\mathbf{R}, \tag{156}$$

$$w(t):=\frac{D(z)}{C(z)}v(t)\in\mathbf{R}. \tag{157}$$

定义系统模型的参数向量 $\boldsymbol{\vartheta}$ 和噪声模型的参数向量 $\boldsymbol{\theta}$，系统模型信息向量 $\boldsymbol{\varphi}(t)$ 和噪声模型信息向量 $\boldsymbol{\psi}(t)$ 如下：

$$\boldsymbol{\vartheta}:=\begin{bmatrix}\boldsymbol{a}\\ \boldsymbol{\rho}\end{bmatrix}\in\mathbf{R}^{n_0},\quad n_0:=n_a+n_f+n_b,\quad n:=n_0+n_c+n_d,$$

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

$$\boldsymbol{\rho}:=[f_1,f_2,\cdots,f_{n_f},b_1,b_2,\cdots,b_{n_b}]^{\mathrm{T}}\in\mathbf{R}^{n_f+n_b},$$

$$\boldsymbol{\theta}:=[c_1,c_2,\cdots,c_{n_c},d_1,d_2,\cdots,d_{n_d}]^{\mathrm{T}}\in\mathbf{R}^{n_c+n_d},$$

$$\boldsymbol{\varphi}(t):=\begin{bmatrix}\boldsymbol{\varphi}_y(t)\\ \boldsymbol{\phi}(t)\end{bmatrix}\in\mathbf{R}^{n_0},$$

$$\boldsymbol{\varphi}_y(t):=[-y(t-1),-y(t-2),\cdots,-y(t-n_a)]^{\mathrm{T}}\in\mathbf{R}^{n_a},$$

$$\boldsymbol{\phi}(t):=[-x(t-1),-x(t-2),\cdots,-x(t-n_f),\ u(t-1),u(t-2),\cdots,u(t-n_b)]^{\mathrm{T}}\in\mathbf{R}^{n_f+n_b},$$

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

系统的量测输入输出数据为$\{u(t),y(t)\}$，中间变量 $x(t)$，噪声项 $w(t)$ 和 $v(t)$ 都是不可测的未知量，因此子信息向量 $\boldsymbol{\phi}(t)$ 和 $\boldsymbol{\psi}(t)$ 是未知的，导致 $\boldsymbol{\varphi}(t)$ 也是未知的.

由式(156)—(157)可得

$$x(t)=[1-F(z)]x(t)+B(z)u(t)=\boldsymbol{\phi}^{\mathrm{T}}(t)\boldsymbol{\rho}, \tag{158}$$

$$w(t)=[1-C(z)]w(t)+D(z)v(t)=\boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\theta}+v(t), \tag{159}$$

把式(156)和(157)代入式(155)可得

$$y(t)=[1-A(z)]y(t)+x(t)+w(t)=$$

$$\boldsymbol{\varphi}_y^{\mathrm{T}}(t)\boldsymbol{a}+x(t)+w(t)= \tag{160}$$

$$\boldsymbol{\varphi}_y^{\mathrm{T}}(t)\boldsymbol{a}+\boldsymbol{\phi}^{\mathrm{T}}(t)\boldsymbol{\rho}+w(t)=\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\vartheta}+w(t)= \tag{161}$$

$$\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\vartheta}+\boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\theta}+v(t). \tag{162}$$

取滤波器 $L(z):=H^{-1}(z)=\frac{C(z)}{D(z)}$，它是噪声模型传递函数 $H(z)$ 的倒数.定义滤波输出 $y_{\mathrm{f}}(t)$ 和滤波信息向量 $\boldsymbol{\varphi}_{\mathrm{f}}(t)$ 分别为

$$y_{\mathrm{f}}(t):=L(z)y(t)=\frac{C(z)}{D(z)}y(t)\in\mathbf{R}, \tag{163}$$

$$\boldsymbol{\varphi}_{\mathrm{f}}(t):=L(z)\boldsymbol{\varphi}(t)=\frac{C(z)}{D(z)}\boldsymbol{\varphi}(t)\in\mathbf{R}^{n_0}. \tag{164}$$

它们可以写成递推形式

$$y_f(t)=C(z)y(t)+[1-D(z)]y_f(t)=$$

$$y(t)+[y(t-1),y(t-2),\cdots,y(t-n_c),-y_{\mathrm{f}}(t-1),-y_f(t-2),\cdots,-y_f(t-n_d)]\boldsymbol{\theta}, \tag{165}$$

$$\boldsymbol{\varphi}_{\mathrm{f}}(t)=C(z)\boldsymbol{\varphi}(t)+[1-D(z)]\boldsymbol{\varphi}_f(t)=$$

$$\boldsymbol{\varphi}(t)+[\boldsymbol{\varphi}(t-1),\boldsymbol{\varphi}(t-2),\cdots,\boldsymbol{\varphi}(t-n_c),-\boldsymbol{\varphi}_{\mathrm{f}}(t-1),-\boldsymbol{\varphi}_{\mathrm{f}}(t-2),\cdots,-\boldsymbol{\varphi}_f(t-n_d)]\boldsymbol{\theta}. \tag{166}$$

式(161)两边乘以 $L(z)=\frac{C(z)}{D(z)}$ 得到

$$L(z)y(t)=L(z)\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\vartheta}+v(t),$$

或

$$y_{\mathrm{f}}(t)=\boldsymbol{\varphi}_{\mathrm{f}}^{\mathrm{T}}(t)\boldsymbol{\vartheta}+v(t). \tag{167}$$

式(167)和(159)构成了滤波辨识模型(filtered identification model)，其干扰 $v(t)$ 是一个白噪声.它们包含了系统的参数向量 $\boldsymbol{\vartheta}$ 和 $\boldsymbol{\theta}$.由于多项式 $C(z)$ 和 $D(z)$ 是未知的(即 $L(z)$ 是未知的)，故 $y_{\mathrm{f}}(t)$ 和 $\boldsymbol{\varphi}_{\mathrm{f}}(t)$ 是未知的，由噪声 $w(t-i)$ 和 $v(t-i)$ 构成的信息向量 $\boldsymbol{\psi}(t)$ 也是未知的，故须采用递推方案或迭代方案来研究新的辨识方法，交互估计参数和滤波变量.

5.2 基于滤波的辅助模型广义增广随机梯度算法

对于 AR-BJ 系统的辨识模型(167)和(159)，定义两个梯度准则函数：

$$J_4(\boldsymbol{\vartheta}):=\frac{1}{2}[y_{\mathrm{f}}(t)-\boldsymbol{\varphi}_{\mathrm{f}}^{\mathrm{T}}(t)\boldsymbol{\vartheta}]^2,$$

$$J_5(\boldsymbol{\theta}):=\frac{1}{2}[w(t)-\boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\theta}]^2. \tag{168}$$

这两个准则函数实际上是相等的,即 $J_4(\boldsymbol{\vartheta})=J_5(\boldsymbol{\theta})$,因为辨识算法都是通过极小化白噪声得到,这是辨识的定义.每个准则函数推导一个参数向量的辨识算法.

设 $\hat{\boldsymbol{\vartheta}}(t):=\begin{bmatrix}\hat{\boldsymbol{a}}(t)\\ \hat{\boldsymbol{\rho}}(t)\end{bmatrix}\in\mathbf{R}^{n_0}$ 和 $\hat{\boldsymbol{\theta}}(t)$ 分别是 $\boldsymbol{\vartheta}=\begin{bmatrix}\boldsymbol{a}\\ \boldsymbol{\rho}\end{bmatrix}$ 和 $\boldsymbol{\theta}$ 在时刻 t 的估计.使用负梯度搜索,极小化 $J_4(\boldsymbol{\vartheta})$ 和 $J_5(\boldsymbol{\theta})$,可以得到下列递推关系:

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\frac{\boldsymbol{\varphi}_{\rm f}(t)}{r_1(t)}[y_{\rm f}(t)-\boldsymbol{\varphi}_{\rm f}^{\rm T}(t)\hat{\boldsymbol{\vartheta}}(t-1)],\tag{169}$$

$$r_1(t)=r_1(t-1)+\|\boldsymbol{\varphi}_{\rm f}(t)\|^2,\tag{170}$$

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\boldsymbol{\psi}(t)}{r_2(t)}[w(t)-\boldsymbol{\psi}^{\rm T}(t)\hat{\boldsymbol{\theta}}(t-1)]=$$

$$\hat{\boldsymbol{\theta}}(t-1)+\frac{\boldsymbol{\psi}(t)}{r_2(t)}[y(t)-\boldsymbol{\varphi}^{\rm T}(t)\boldsymbol{\vartheta}-\boldsymbol{\psi}^{\rm T}(t)\hat{\boldsymbol{\theta}}(t-1)],\tag{171}$$

$$r_2(t)=r_2(t-1)+\|\boldsymbol{\psi}(t)\|^2.\tag{172}$$

参照辅助模型辨识思想,式(169)—(172)右边的 $y_{\rm f}(t)$,$\boldsymbol{\varphi}_{\rm f}(t)$,$\boldsymbol{\varphi}(t)$,$\boldsymbol{\vartheta}$ 和 $\boldsymbol{\psi}(t)$ 分别用其估计 $\hat{y}_{\rm f}(t)$,$\hat{\boldsymbol{\varphi}}_{\rm f}(t)$,$\hat{\boldsymbol{\varphi}}(t)$,$\hat{\boldsymbol{\vartheta}}(t-1)$ 和 $\hat{\boldsymbol{\psi}}(t)$ 代替,并引入遗忘因子 λ_1 和 λ_2,可以得到估计 AR-BJ 系统参数向量 $\boldsymbol{\vartheta}$ 和 $\boldsymbol{\theta}$ 的带遗忘因子的基于滤波的辅助模型广义增广随机梯度算法(FF-F-AM-GESG 算法),也可称为基于滤波的遗忘因子辅助模型广义增广随机梯度算法(F-FF-AM-GESG 算法):

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\frac{\hat{\boldsymbol{\varphi}}_{\rm f}(t)}{r_1(t)}[\hat{y}_{\rm f}(t)-\hat{\boldsymbol{\varphi}}_{\rm f}^{\rm T}(t)\hat{\boldsymbol{\vartheta}}(t-1)],\tag{173}$$

$$r_1(t)=\lambda_1r_1(t-1)+\|\hat{\boldsymbol{\varphi}}_{\rm f}(t)\|^2,\quad 0\leqslant\lambda_1\leqslant1,\tag{174}$$

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\hat{\boldsymbol{\psi}}(t)}{r_2(t)}[y(t)-\hat{\boldsymbol{\varphi}}^{\rm T}(t)\hat{\boldsymbol{\vartheta}}(t-1)-\hat{\boldsymbol{\psi}}^{\rm T}(t)\hat{\boldsymbol{\theta}}(t-1)],\tag{175}$$

$$r_2(t)=\lambda_2r_2(t-1)+\|\hat{\boldsymbol{\psi}}(t)\|^2,\quad 0\leqslant\lambda_2\leqslant1,\tag{176}$$

$$\hat{y}_{\rm f}(t)=y(t)+[y(t-1),y(t-2),\cdots,y(t-n_c),-\hat{y}_{\rm f}(t-1),-\hat{y}_{\rm f}(t-2),\cdots,-\hat{y}_{\rm f}(t-n_d)]\hat{\boldsymbol{\theta}}(t),\tag{177}$$

$$\hat{\boldsymbol{\varphi}}_{\rm f}(t)=\hat{\boldsymbol{\varphi}}(t)+[\hat{\boldsymbol{\varphi}}(t-1),\hat{\boldsymbol{\varphi}}(t-2),\cdots,\hat{\boldsymbol{\varphi}}(t-n_c),-\hat{\boldsymbol{\varphi}}_{\rm f}(t-1),-\hat{\boldsymbol{\varphi}}_{\rm f}(t-2),\cdots,-\hat{\boldsymbol{\varphi}}_{\rm f}(t-n_d)]\hat{\boldsymbol{\theta}}(t),\tag{178}$$

$$\hat{\boldsymbol{\varphi}}(t)=\begin{bmatrix}\boldsymbol{\varphi}_y(t)\\ \hat{\boldsymbol{\phi}}(t)\end{bmatrix},\tag{179}$$

$$\boldsymbol{\varphi}_y(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a)]^{\rm T},\tag{180}$$

$$\hat{\boldsymbol{\phi}}(t)=[-x_{\rm a}(t-1),-x_{\rm a}(t-2),\cdots,-x_{\rm a}(t-n_f),u(t-1),u(t-2),\cdots,u(t-n_b)]^{\rm T},\tag{181}$$

$$\hat{\boldsymbol{\psi}}(t)=[-\hat{w}(t-1),-\hat{w}(t-2),\cdots,-\hat{w}(t-n_c),\hat{v}(t-1),\hat{v}(t-2),\cdots,\hat{v}(t-n_d)]^{\rm T},\tag{182}$$

$$x_{\rm a}(t)=\hat{\boldsymbol{\phi}}^{\rm T}(t)\hat{\boldsymbol{\rho}}(t),\tag{183}$$

$$\hat{w}(t)=y(t)-\boldsymbol{\varphi}_y^{\rm T}(t)\hat{\boldsymbol{a}}(t)-x_{\rm a}(t),\tag{184}$$

$$\hat{v}(t)=\hat{w}(t)-\hat{\boldsymbol{\psi}}^{\rm T}(t)\hat{\boldsymbol{\theta}}(t),\tag{185}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\begin{bmatrix}\hat{\boldsymbol{a}}(t)\\ \hat{\boldsymbol{\rho}}(t)\end{bmatrix}.\tag{186}$$

当遗忘因子 $\lambda_1=\lambda_2=1$ 时,F-FF-AM-GESG 算法(173)—(186)退化为基于滤波的辅助模型广义增广随机梯度算法(F-AM-GESG 算法).F-FF-AM-GESG 算法(173)—(186)计算参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$ 和 $\hat{\boldsymbol{\theta}}(t)$ 的步骤如下:

1) 初始化:令 $t=1$.置初值 $\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_{n_0}/p_0$,$\hat{\boldsymbol{\theta}}(0)=\mathbf{1}_{n_c+n_d}/p_0$,$r_1(0)=1$,$r_2(0)=1$,$x_{\rm a}(t-i)=1/p_0$,$\hat{w}(t-i)=1/p_0$,$\hat{v}(t-i)=1/p_0$,$\hat{y}_{\rm f}(t-i)=1/p_0$,$\hat{\boldsymbol{\varphi}}_{\rm f}(t-i)=\mathbf{1}_{n_0}/p_0$,$\hat{\boldsymbol{\varphi}}(t-i)=\mathbf{1}_{n_0}/p_0$,$i=1,2,\cdots,\max[n_f,n_c,n_d]$,$p_0=10^6$.给定遗忘因子 λ_1,λ_2 和小正数 ε.

2) 采集输入输出数据 $u(t)$ 和 $y(t)$,由式(180)—(182)和(179)构造信息向量 $\boldsymbol{\varphi}_y(t)$,$\hat{\boldsymbol{\phi}}(t)$,$\hat{\boldsymbol{\psi}}(t)$ 和 $\hat{\boldsymbol{\varphi}}(t)$.

3) 由式(176)计算 $r_2(t)$.

4) 根据式(175)刷新参数估计向量 $\hat{\boldsymbol{\theta}}(t)$.

5) 由式(177)—(178)计算 $\hat{y}_{\rm f}(t)$ 和 $\hat{\boldsymbol{\varphi}}_{\rm f}(t)$.

6) 由式(174)计算 $r_1(t)$.

7) 根据式(173)刷新参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$.

8) 从式(186)的 $\hat{\boldsymbol{\vartheta}}(t)$ 中读出 $\hat{\boldsymbol{a}}(t)$ 和 $\hat{\boldsymbol{\rho}}(t)$,根据式(183)—(185)计算辅助模型的输出 $x_{\rm a}(t)$,$\hat{w}(t)$ 和 $\hat{v}(t)$.

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

上述估计 $\boldsymbol{\vartheta}$ 和 $\boldsymbol{\theta}$ 的 F-AM-RGELS 算法,等价于极小化下列两个梯度乘准则函数:

$$J_6(\boldsymbol{\vartheta})=\frac{1}{2}\left\{\frac{\hat{C}(t,z)}{\hat{D}(t,z)}\left[A(z)y(t)-\frac{B(z)}{F(z)}u(t)\right]\right\}^2,$$

$$J_7(\boldsymbol{\theta})=\frac{1}{2}\left\{\hat{A}(t-1,z)y(t)-\frac{\hat{B}(t-1,z)}{\hat{F}(t-1,z)}u(t)-[1-C(z)]\hat{w}(t)-[D(z)-1]\hat{v}(t)\right\}^2.$$

5.3 基于滤波的辅助模型多新息广义增广随机梯度算法

设正整数 p 表示新息长度.基于 F-AM-GESG 算法，将滤波输出 $\hat{y}_{\mathrm{f}}(t)$、滤波信息向量 $\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)$、系统输出 $y(t)$、信息向量 $\hat{\boldsymbol{\varphi}}(t)$ 和噪声信息向量 $\hat{\boldsymbol{\psi}}(t)$ 分别扩展为堆积滤波输出向量 $\hat{\boldsymbol{Y}}_{\mathrm{f}}(p,t)$，堆积滤波信息矩阵 $\hat{\boldsymbol{\Phi}}_{\mathrm{f}}(p,t)$，堆积输出向量 $\boldsymbol{Y}(p,t)$、堆积信息矩阵 $\hat{\boldsymbol{\Phi}}(p,t)$ 和堆积噪声信息矩阵 $\hat{\boldsymbol{\Psi}}(p,t)$，将式（173）和（175）中标量新息

$$e_1(t):=\hat{y}_{\mathrm{f}}(t)-\hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t-1)\in\mathbf{R},$$

$$e_2(t):=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t-1)-\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)\in\mathbf{R} \tag{187}$$

扩展为新息向量 $\boldsymbol{E}_1(p,t)$ 和 $\boldsymbol{E}_2(p,t)$，可以得到估计 AR-BJ 系统参数向量 $\boldsymbol{\vartheta}$ 和 $\boldsymbol{\theta}$ 的基于滤波的辅助模型多新息广义增广随机梯度算法（Filtering based AM-MI-GESG algorithm，F-AM-MI-GESG 算法）：

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\frac{\hat{\boldsymbol{\Phi}}_{\mathrm{f}}(p,t)}{r_1(t)}\boldsymbol{E}_1(p,t), \tag{188}$$

$$\boldsymbol{E}_1(p,t)=\hat{\boldsymbol{Y}}_{\mathrm{f}}(p,t)-\hat{\boldsymbol{\Phi}}_{\mathrm{f}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\vartheta}}(t-1), \tag{189}$$

$$r_1(t)=r_1(t-1)+\|\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)\|^2, \tag{190}$$

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\hat{\boldsymbol{\Phi}}(p,t)}{r_2(t)}\boldsymbol{E}_2(p,t), \tag{191}$$

$$\boldsymbol{E}_2(p,t)=\boldsymbol{Y}(p,t)-\hat{\boldsymbol{\Phi}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\vartheta}}(t-1)-\hat{\boldsymbol{\Psi}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\theta}}(t-1), \tag{192}$$

$$r_2(t)=r_2(t-1)+\|\hat{\boldsymbol{\psi}}(t)\|^2, \tag{193}$$

$$\hat{\boldsymbol{Y}}_{\mathrm{f}}(p,t)=[\hat{y}_{\mathrm{f}}(t),\hat{y}_{\mathrm{f}}(t-1),\cdots,\hat{y}_{\mathrm{f}}(t-p+1)]^{\mathrm{T}}, \tag{194}$$

$$\hat{\boldsymbol{\Phi}}_{\mathrm{f}}(p,t)=[\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t),\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-1),\cdots,\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-p+1)], \tag{195}$$

$$\hat{\boldsymbol{Y}}(p,t)=[y(t),y(t-1),\cdots,y(t-p+1)]^{\mathrm{T}}, \tag{196}$$

$$\hat{\boldsymbol{\Phi}}(p,t)=[\hat{\boldsymbol{\varphi}}(t),\hat{\boldsymbol{\varphi}}(t-1),\cdots,\hat{\boldsymbol{\varphi}}(t-p+1)], \tag{197}$$

$$\hat{\boldsymbol{\Psi}}(p,t)=[\hat{\boldsymbol{\psi}}(t),\hat{\boldsymbol{\psi}}(t-1),\cdots,\hat{\boldsymbol{\psi}}(t-p+1)], \tag{198}$$

$$\begin{aligned}\hat{y}_{\mathrm{f}}(t)=y(t)+[&y(t-1),y(t-2),\cdots,y(t-n_c),\\&-\hat{y}_{\mathrm{f}}(t-1),-\hat{y}_{\mathrm{f}}(t-2),\cdots,-\hat{y}_{\mathrm{f}}(t-n_d)]\hat{\boldsymbol{\theta}}(t),\end{aligned} \tag{199}$$

$$\begin{aligned}\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)=\hat{\boldsymbol{\varphi}}(t)+[&\hat{\boldsymbol{\varphi}}(t-1),\hat{\boldsymbol{\varphi}}(t-2),\cdots,\hat{\boldsymbol{\varphi}}(t-n_c),\\&-\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-1),-\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-2),\cdots,-\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-n_d)]\hat{\boldsymbol{\theta}}(t),\end{aligned} \tag{200}$$

$$\hat{\boldsymbol{\varphi}}(t)=\begin{bmatrix}\boldsymbol{\varphi}_y(t)\\ \hat{\boldsymbol{\phi}}(t)\end{bmatrix}, \tag{201}$$

$$\boldsymbol{\varphi}_y(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a)]^{\mathrm{T}}, \tag{202}$$

$$\begin{aligned}\hat{\boldsymbol{\phi}}(t)=[&-x_{\mathrm{a}}(t-1),-x_{\mathrm{a}}(t-2),\cdots,-x_{\mathrm{a}}(t-n_f),\\&u(t-1),u(t-2),\cdots,u(t-n_b)]^{\mathrm{T}},\end{aligned} \tag{203}$$

$$\begin{aligned}\hat{\boldsymbol{\psi}}(t)=[&-\hat{w}(t-1),-\hat{w}(t-2),\cdots,-\hat{w}(t-n_c),\\&\hat{v}(t-1),\hat{v}(t-2),\cdots,\hat{v}(t-n_d)]^{\mathrm{T}},\end{aligned} \tag{204}$$

$$x_{\mathrm{a}}(t)=\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\rho}}(t), \tag{205}$$

$$\hat{w}(t)=y(t)-\boldsymbol{\varphi}_y^{\mathrm{T}}(t)\hat{\boldsymbol{a}}(t)-x_{\mathrm{a}}(t), \tag{206}$$

$$\hat{v}(t)=\hat{w}(t)-\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t), \tag{207}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\begin{bmatrix}\hat{\boldsymbol{a}}(t)\\ \hat{\boldsymbol{\rho}}(t)\end{bmatrix}. \tag{208}$$

F-AM-MI-GESG 算法（188）—（208）计算参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$ 和 $\hat{\boldsymbol{\theta}}(t)$ 的步骤如下：

1）初始化：令 $t=1$，给定新息长度 p.置初值 $\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_{n_0}/p_0$，$\hat{\boldsymbol{\theta}}(0)=\mathbf{1}_{n_c+n_d}/p_0$，$r_1(0)=1$，$r_2(0)=1$，$x_{\mathrm{a}}(t-i)=1/p_0$，$\hat{w}(t-i)=1/p_0$，$\hat{v}(t-i)=1/p_0$，$\hat{y}_{\mathrm{f}}(t-i)=1/p_0$，$\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-i)=\mathbf{1}_{n_0}/p_0$，$\hat{\boldsymbol{\varphi}}(t-i)=\mathbf{1}_{n_0}/p_0$，$i=1,2,\cdots,\max[n_f,n_c,n_d]$，$p_0=10^6$.给定小正数 ε.

2）由式（202）—（204）和（201）构造信息向量 $\boldsymbol{\varphi}_y(t)$，$\hat{\boldsymbol{\phi}}(t)$，$\hat{\boldsymbol{\psi}}(t)$ 和 $\hat{\boldsymbol{\varphi}}(t)$.

3）由式（196）—（198）构造堆积输出向量 $\boldsymbol{Y}(p,t)$，堆积信息矩阵 $\hat{\boldsymbol{\Phi}}(p,t)$ 和堆积噪声信息矩阵 $\hat{\boldsymbol{\Psi}}(p,t)$.

4）由式（192）计算新息向量 $\boldsymbol{E}_2(p,t)$，由式（193）计算 $r_2(t)$.

5）根据式（191）刷新参数估计向量 $\hat{\boldsymbol{\theta}}(t)$.

6）由式（199）—（200）计算 $\hat{y}_{\mathrm{f}}(t)$ 和 $\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)$.

7）由式（194）—（195）构造堆积滤波输出向量 $\hat{\boldsymbol{Y}}_{\mathrm{f}}(p,t)$ 和堆积滤波信息矩阵 $\hat{\boldsymbol{\Phi}}_{\mathrm{f}}(p,t)$.

8）由式（189）计算新息向量 $\boldsymbol{E}_1(p,t)$，由式（190）计算 $r_1(t)$.

9）根据式（188）刷新参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$.

10）从式（208）的 $\hat{\boldsymbol{\vartheta}}(t)$ 中读出 $\hat{\boldsymbol{a}}(t)$ 和 $\hat{\boldsymbol{\rho}}(t)$，根据式（205）—（207）计算辅助模型的输出 $x_{\mathrm{a}}(t)$，$\hat{w}(t)$ 和 $\hat{v}(t)$.

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

5.4 基于滤波的辅助模型广义增广递推最小二乘算法

对于 AR-BJ 系统的辨识模型（167）和（159），定义两个最小二乘准则函数：

$$J_8(\boldsymbol{\vartheta}):=\sum_{j=1}^{t}[y_{\mathrm{f}}(j)-\boldsymbol{\varphi}_{\mathrm{f}}^{\mathrm{T}}(j)\boldsymbol{\vartheta}]^2,$$

$$J_9(\boldsymbol{\theta}):=\sum_{j=1}^{t}[w(j)-\psi^{\mathrm{T}}(j)\boldsymbol{\theta}]^2.$$

参照文献[1-3]中 RLS 算法的推导，根据辅助模型辨识思想和滤波辨识理念，极小化 $J_8(\boldsymbol{\vartheta})$ 和

$J_9(\boldsymbol{\theta})$,未知量 $y_{\mathrm{f}}(t)$,$\boldsymbol{\varphi}_{\mathrm{f}}(t)$,$\boldsymbol{\varphi}(t)$,$\boldsymbol{\vartheta}$ 和 $\boldsymbol{\psi}(t)$ 分别用其估计 $\hat{y}_{\mathrm{f}}(t)$,$\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)$,$\hat{\boldsymbol{\varphi}}(t)$,$\hat{\boldsymbol{\vartheta}}(t-1)$ 和 $\hat{\boldsymbol{\psi}}(t)$ 代替,可以得到估计 AR-BJ 系统参数向量 $\boldsymbol{\vartheta}$ 和 $\boldsymbol{\theta}$ 的基于滤波的辅助模型递推广义增广最小二乘算法(Filtering based AM-RGELS algorithm,F-AM-RGELS 算法):

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\boldsymbol{L}_1(t)[\hat{y}_{\mathrm{f}}(t)-\hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t-1)],\tag{209}$$

$$\boldsymbol{L}_1(t)=\boldsymbol{P}_1(t-1)\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)[1+\hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\mathrm{T}}(t)\boldsymbol{P}_1(t-1)\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)]^{-1},\tag{210}$$

$$\boldsymbol{P}_1(t)=[\boldsymbol{I}_{n_0}-\boldsymbol{L}_1(t)\hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\mathrm{T}}(t)]\boldsymbol{P}_1(t-1),\tag{211}$$

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{L}_2(t)[y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t-1)-\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)],\tag{212}$$

$$\boldsymbol{L}_2(t)=\boldsymbol{P}_2(t-1)\hat{\boldsymbol{\psi}}(t)[1+\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)\boldsymbol{P}_2(t-1)\hat{\boldsymbol{\psi}}(t)]^{-1},\tag{213}$$

$$\boldsymbol{P}_2(t)=[\boldsymbol{I}_{n_c+n_d}-\boldsymbol{L}_2(t)\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)]\boldsymbol{P}_2(t-1),\tag{214}$$

$$\hat{y}_{\mathrm{f}}(t)=y(t)+[y(t-1),y(t-2),\cdots,y(t-n_c),-\hat{y}_{\mathrm{f}}(t-1),-\hat{y}_{\mathrm{f}}(t-2),\cdots,-\hat{y}_{\mathrm{f}}(t-n_d)]\hat{\boldsymbol{\theta}}(t),\tag{215}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)=\hat{\boldsymbol{\varphi}}(t)+[\hat{\boldsymbol{\varphi}}(t-1),\hat{\boldsymbol{\varphi}}(t-2),\cdots,\hat{\boldsymbol{\varphi}}(t-n_c),-\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-1),-\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-2),\cdots,-\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-n_d)]\hat{\boldsymbol{\theta}}(t),\tag{216}$$

$$\hat{\boldsymbol{\varphi}}(t)=\begin{bmatrix}\boldsymbol{\varphi}_y(t)\\ \hat{\boldsymbol{\phi}}(t)\end{bmatrix},\tag{217}$$

$$\boldsymbol{\varphi}_y(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a)]^{\mathrm{T}},\tag{218}$$

$$\hat{\boldsymbol{\phi}}(t)=[-x_{\mathrm{a}}(t-1),-x_{\mathrm{a}}(t-2),\cdots,-x_{\mathrm{a}}(t-n_f),u(t-1),u(t-2),\cdots,u(t-n_b)]^{\mathrm{T}},\tag{219}$$

$$\hat{\boldsymbol{\psi}}(t)=[-\hat{w}(t-1),-\hat{w}(t-2),\cdots,-\hat{w}(t-n_c),\hat{v}(t-1),\hat{v}(t-2),\cdots,\hat{v}(t-n_d)]^{\mathrm{T}},\tag{220}$$

$$x_{\mathrm{a}}(t)=\hat{\boldsymbol{\phi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\rho}}(t),\tag{221}$$

$$\hat{w}(t)=y(t)-\boldsymbol{\varphi}_y^{\mathrm{T}}(t)\hat{\boldsymbol{a}}(t)-x_{\mathrm{a}}(t),\tag{222}$$

$$\hat{v}(t)=\hat{w}(t)-\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t),\tag{223}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\begin{bmatrix}\hat{\boldsymbol{a}}(t)\\ \hat{\boldsymbol{\rho}}(t)\end{bmatrix}.\tag{224}$$

F-AM-RGELS 算法(209)—(224)计算参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$ 和 $\hat{\boldsymbol{\theta}}(t)$ 的步骤如下:

1) 初始化:令 $t=1$.置初值 $\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_{n_0}/p_0$,$\hat{\boldsymbol{\theta}}(0)=\mathbf{1}_{n_c+n_d}/p_0$,$\boldsymbol{P}_1(0)=p_0\boldsymbol{I}_{n_0}$,$\boldsymbol{P}_2(0)=p_0\boldsymbol{I}_{n_c+n_d}$,$x_{\mathrm{a}}(t-i)=1/p_0$,$\hat{w}(t-i)=1/p_0$,$\hat{v}(t-i)=1/p_0$,$\hat{y}_{\mathrm{f}}(t-i)=1/p_0$,$\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-i)=\mathbf{1}_{n_0}/p_0$,$\hat{\boldsymbol{\varphi}}(t-i)=\mathbf{1}_{n_0}/p_0$,$i=1,2,\cdots,\max[n_f,n_c,n_d]$,$p_0=10^6$.给定小正数 ε.

2) 由式(218)—(220)和(217)构造信息向量 $\boldsymbol{\varphi}_y(t)$,$\hat{\boldsymbol{\phi}}(t)$,$\hat{\boldsymbol{\psi}}(t)$ 和 $\hat{\boldsymbol{\varphi}}(t)$.

3) 由式(213)—(214)计算增益向量 $\boldsymbol{L}_2(t)$ 和协方差阵 $\boldsymbol{P}_2(t)$.

4) 根据式(212)刷新参数估计向量 $\hat{\boldsymbol{\theta}}(t)$.

5) 由式(215)—(216)计算 $\hat{y}_{\mathrm{f}}(t)$ 和 $\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)$.

6) 由式(210)—(211)计算增益向量 $\boldsymbol{L}_1(t)$ 和协方差阵 $\boldsymbol{P}_1(t)$.

7) 根据式(209)刷新参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$.

8) 从式(224)的 $\hat{\boldsymbol{\vartheta}}(t)$ 中读出 $\hat{\boldsymbol{a}}(t)$ 和 $\hat{\boldsymbol{\rho}}(t)$,根据式(221)—(223)计算辅助模型的输出 $x_{\mathrm{a}}(t)$,$\hat{w}(t)$ 和 $\hat{v}(t)$.

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

上述估计 $\boldsymbol{\vartheta}$ 和 $\boldsymbol{\theta}$ 的 F-AM-RGELS 算法,等价于极小化下列两个最小二乘准则函数:

$$J_{10}(\boldsymbol{\vartheta})=\sum_{j=1}^{t}\left\{\frac{\hat{C}(t,z)}{\hat{D}(t,z)}\left[A(z)y(j)-\frac{B(z)}{F(z)}u(j)\right]\right\}^2,$$

$$J_{11}(\boldsymbol{\theta})=\sum_{j=1}^{t}\left\{\hat{A}(t-1,z)y(j)-\frac{\hat{B}(t-1,z)}{\hat{F}(t-1,z)}u(j)-[1-C(z)]\hat{w}(j)-[D(z)-1]\hat{v}(j)\right\}^2.$$

注 19 本节的辨识方法可以推广到多输入 AR-BJ 模型描述的动态随机系统:

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

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

和非线性自回归输出误差 ARMA 模型(N-AR-OE-ARMA 模型),即 N-AR-BJ 模型描述的非线性动态随机系统:

3) $A(z)f(y(t))=\dfrac{B(z)}{F(z)}g(u(t))+\dfrac{D(z)}{C(z)}v(t)$.

6 结语

本文讨论了自回归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,研究了自回归 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,研究了自回归 Box-Jenkins 系统的滤波辅助模型广义增广随机梯度(F-AM-GESG)算法、滤波辅助模型多新息广义增广随机梯度(F-AM-MI-GESG)算法、滤波辅助模型递推广义增广最小二乘(F-AM-RGELS)

算法等.这些方法可以推广到多输入单输出自回归输出误差类系统、多输入多输出自回归输出误差类系统、多变量自回归输出误差类系统、多元自回归输出误差类系统，以及非线性自回归输出误差类系统[1-3].

1）研究下列多输入 AR-OEARMA 模型（即多输入 AR-BJ 模型）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：

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

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

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

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

2）对于多输入自回归输出误差 ARMA 模型（即 AR-OEARMA 模型）描述的动态随机系统：

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

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

研究基于分解的辅助模型广义增广随机梯度（D-AM-GESG）算法、基于分解的辅助模型多新息广义增广随机梯度（D-AM-MI-GESG）算法，基于分解的辅助模型递推广义增广最小二乘（D-AM-RGELS）算法[3].

3）对于多输入自回归输出误差 ARMA 模型描述的动态随机系统：

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

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

研究基于滤波的辅助模型广义增广随机梯度算法、基于滤波的辅助模型多新息广义增广随机梯度算法，基于滤波的辅助模型递推广义增广最小二乘算法.

4）研究下列非线性自回归输出误差模型（Nonlinear AR-OE model，N-AR-OE 模型）描述的非线性动态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：

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

其中 $f(*)$ 和 $g(*)$ 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：

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

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

5）研究下列非线性自回归输出误差 ARMA 模型（N-AR-OEARMA 模型），即 N-AR-BJ 模型描述的非线性动态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：

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

其中 $f(*)$ 和 $g(*)$ 为已知基函数，例子如下：

$$①\ f(y(t))=y^2(t),\quad g(u(t))=\sin^2(u(t))+e^{\cos(t/\pi)}.$$

$$②\ f(y(t))=|y(t)|,\quad g(u(t))=\sqrt[3]{u^2(t)}+\ln[u^2(t-1)+1].$$

6）针对下列非线性自回归输出误差 ARMA 模型描述的非线性动态系统：

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

研究基于分解的辅助模型广义增广随机梯度（D-AM-GESG）算法、基于分解的辅助模型多新息广义增广随机梯度（D-AM-MI-GESG）算法、基于分解的辅助模型递推广义增广最小二乘（D-AM-RGELS）算法[3].

7）针对下列非线性自回归输出误差 ARMA 模型描述的非线性动态系统：

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

研究基于滤波的辅助模型广义增广随机梯度（F-AM-GESG）算法、基于滤波的辅助模型多新息广义增广随机梯度（F-AM-MI-GESG）算法、基于滤波的辅助模型递推广义增广最小二乘（F-AM-RGELS）算法.

8）研究下列多变量自回归输出误差 ARMA 模型（即多变量 AR-BJ 模型}）描述的多变量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[53]：

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

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

③ $A(z)\boldsymbol{y}(t)=\boldsymbol{F}^{-1}(z)\boldsymbol{B}(z)\boldsymbol{u}(t)+\boldsymbol{C}^{-1}(z)\boldsymbol{D}(z)\boldsymbol{v}(t)$,

④ $\boldsymbol{A}(z)\boldsymbol{y}(t)=\boldsymbol{F}^{-1}(z)\boldsymbol{B}(z)\boldsymbol{u}(t)+\boldsymbol{C}^{-1}(z)\boldsymbol{D}(z)\boldsymbol{v}(t)$,

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

$\boldsymbol{A}(z):=\boldsymbol{I}+\boldsymbol{A}_1z^{-1}+\boldsymbol{A}_2z^{-2}+\cdots+\boldsymbol{A}_{n_a}z^{-n_a},\boldsymbol{A}_i\in\mathbf{R}^{m\times m}$,

$\boldsymbol{B}(z):=\boldsymbol{B}_1z^{-1}+\boldsymbol{B}_2z^{-2}+\cdots+\boldsymbol{B}_{n_b}z^{-n_b},\boldsymbol{B}_i\in\mathbf{R}^{m\times r}$,

$\boldsymbol{C}(z):=\boldsymbol{I}+\boldsymbol{C}_1z^{-1}+\boldsymbol{C}_2z^{-2}+\cdots+\boldsymbol{C}_{n_c}z^{-n_c},\boldsymbol{C}_i\in\mathbf{R}^{m\times m}$,

$\boldsymbol{D}(z):=\boldsymbol{I}+\boldsymbol{D}_1z^{-1}+\boldsymbol{D}_2z^{-2}+\cdots+\boldsymbol{D}_{n_d}z^{-n_d},\boldsymbol{D}_i\in\mathbf{R}^{m\times m}$,

$\boldsymbol{F}(z):=\boldsymbol{I}+\boldsymbol{F}_1z^{-1}+\boldsymbol{F}_2z^{-2}+\cdots+\boldsymbol{F}_{n_f}z^{-n_f},\boldsymbol{F}_i\in\mathbf{R}^{m\times m}$.

9) 对于多变量自回归输出误差 ARMA 模型(multivariable AR-OE ARMA model)描述的多变量系统:

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

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

③ $\boldsymbol{A}(z)\boldsymbol{y}(t)=\frac{\boldsymbol{B}(z)}{F(z)}\boldsymbol{u}(t)+\frac{D(z)}{C(z)}\boldsymbol{v}(t)$,

④ $A(z)\boldsymbol{y}(t)=\boldsymbol{F}^{-1}(z)\boldsymbol{B}(z)\boldsymbol{u}(t)+\boldsymbol{C}^{-1}(z)\boldsymbol{D}(z)\boldsymbol{v}(t)$,

⑤ $\boldsymbol{A}(z)\boldsymbol{y}(t)=\boldsymbol{F}^{-1}(z)\boldsymbol{B}(z)\boldsymbol{u}(t)+\boldsymbol{C}^{-1}(z)\boldsymbol{D}(z)\boldsymbol{v}(t)$,

⑥ $A(z)\boldsymbol{y}(t)=\frac{\boldsymbol{B}(z)}{F(z)}\boldsymbol{u}(t)+\boldsymbol{C}^{-1}(z)\boldsymbol{D}(z)\boldsymbol{v}(t)$,

⑦ $\boldsymbol{A}(z)\boldsymbol{y}(t)=\frac{\boldsymbol{B}(z)}{F(z)}\boldsymbol{u}(t)+\boldsymbol{C}^{-1}(z)\boldsymbol{D}(z)\boldsymbol{v}(t)$,

研究基于分解的辅助模型广义增广随机梯度(D-AM-GESG)算法、基于分解的辅助模型多新息广义增广随机梯度(D-AM-MI-GESG)算法、基于分解的辅助模型递推广义增广最小二乘(D-AM-RGELS)算法[3].

10) 对于多变量自回归输出误差 ARMA 模型描述的多变量系统:

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

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

③ $\boldsymbol{A}(z)\boldsymbol{y}(t)=\frac{\boldsymbol{B}(z)}{F(z)}\boldsymbol{u}(t)+\frac{D(z)}{C(z)}\boldsymbol{v}(t)$,

研究基于滤波的辅助模型广义增广随机梯度(F-AM-GESG)算法、基于滤波的辅助模型多新息广义增广随机梯度(F-AM-MI-GESG)算法、基于滤波的辅助模型递推广义增广最小二乘(F-AM-RGELS)算法.

11) 对于多元自回归输出误差模型(multivariate AR-OE model)描述的动态随机系统:

① $A(z)\boldsymbol{y}(t)=\frac{\boldsymbol{\Phi}(t)}{F(z)}\boldsymbol{\theta}+\boldsymbol{v}(t)$,

② $\boldsymbol{A}(z)\boldsymbol{y}(t)=\frac{\boldsymbol{\Phi}(t)}{F(z)}\boldsymbol{\theta}+\boldsymbol{v}(t)$,

③ $A(z)\boldsymbol{y}(t)=\boldsymbol{F}^{-1}(z)\boldsymbol{\Phi}(t)\boldsymbol{\theta}+\boldsymbol{v}(t)$,

④ $\boldsymbol{A}(z)\boldsymbol{y}(t)=\boldsymbol{F}^{-1}(z)\boldsymbol{\Phi}(t)\boldsymbol{\theta}+\boldsymbol{v}(t)$,

其中 $\boldsymbol{\Phi}(t)\in\mathbf{R}^{m\times n}$是可测信息矩阵,$\boldsymbol{\theta}\in\mathbf{R}^n$ 为待辨识的参数向量,研究(基于分解的)辅助模型随机梯度算法、(基于分解的)辅助模型多新息随机梯度算法、(基于分解的)辅助模型递推最小二乘算法.

12) 对于多元自回归输出误差 ARMA 模型(multivariate AR-OE ARMA model)描述的动态随机系统:

① $A(z)\boldsymbol{y}(t)=\frac{\boldsymbol{\Phi}(t)}{F(z)}\boldsymbol{\theta}+\frac{D(z)}{C(z)}\boldsymbol{v}(t)$,

② $\boldsymbol{A}(z)\boldsymbol{y}(t)=\frac{\boldsymbol{\Phi}(t)}{F(z)}\boldsymbol{\theta}+\frac{D(z)}{C(z)}\boldsymbol{v}(t)$,

③ $A(z)\boldsymbol{y}(t)=\boldsymbol{F}^{-1}(z)\boldsymbol{\Phi}(t)\boldsymbol{\theta}+\frac{D(z)}{C(z)}\boldsymbol{v}(t)$,

④ $\boldsymbol{A}(z)\boldsymbol{y}(t)=\boldsymbol{F}^{-1}(z)\boldsymbol{\Phi}(t)\boldsymbol{\theta}+\frac{D(z)}{C(z)}\boldsymbol{v}(t)$,

⑤ $A(z)\boldsymbol{y}(t)=\frac{\boldsymbol{\Phi}(t)}{F(z)}\boldsymbol{\theta}+\boldsymbol{C}^{-1}(z)\boldsymbol{D}(z)\boldsymbol{v}(t)$,

⑥ $\boldsymbol{A}(z)\boldsymbol{y}(t)=\frac{\boldsymbol{\Phi}(t)}{F(z)}\boldsymbol{\theta}+\boldsymbol{C}^{-1}(z)\boldsymbol{D}(z)\boldsymbol{v}(t)$,

⑦ $A(z)\boldsymbol{y}(t)=\boldsymbol{F}^{-1}(z)\boldsymbol{\Phi}(t)\boldsymbol{\theta}+\boldsymbol{C}^{-1}(z)\boldsymbol{D}(z)\boldsymbol{v}(t)$,

⑧ $\boldsymbol{A}(z)\boldsymbol{y}(t)=\boldsymbol{F}^{-1}(z)\boldsymbol{\Phi}(t)\boldsymbol{\theta}+\boldsymbol{C}^{-1}(z)\boldsymbol{D}(z)\boldsymbol{v}(t)$,

研究基于辅助模型的广义增广随机梯度(AM-GESG)算法、基于辅助模型的递推广义增广最小二乘(AM-RGELS)算法、基于分解的辅助模型广义增广随机梯度(D-AM-GESG)算法、基于分解的辅助模型递推广义增广最小二乘(D-AM-RGELS)算法.

13) 对于多元自回归输出误差 ARMA 模型描述的动态随机系统:

① $A(z)\boldsymbol{y}(t)=\frac{\boldsymbol{\Phi}(t)}{F(z)}\boldsymbol{\theta}+\frac{D(z)}{C(z)}\boldsymbol{v}(t)$,

② $\boldsymbol{A}(z)\boldsymbol{y}(t)=\frac{\boldsymbol{\Phi}(t)}{F(z)}\boldsymbol{\theta}+\frac{D(z)}{C(z)}\boldsymbol{v}(t)$,

③ $A(z)\boldsymbol{y}(t)=\boldsymbol{F}^{-1}(z)\boldsymbol{\Phi}(t)\boldsymbol{\theta}+\frac{D(z)}{C(z)}\boldsymbol{v}(t)$, ④ $\boldsymbol{A}(z)\boldsymbol{y}(t)=\boldsymbol{F}^{-1}(z)\boldsymbol{\Phi}(t)\boldsymbol{\theta}+\dfrac{D(z)}{C(z)}\boldsymbol{v}(t)$，

⑤ $A(z)\boldsymbol{y}(t)=\dfrac{\boldsymbol{\Phi}(t)}{F(z)}\boldsymbol{\theta}+\boldsymbol{C}^{-1}(z)\boldsymbol{D}(z)\boldsymbol{v}(t)$，

研究基于滤波（分解）的辅助模型广义增广随机梯度（F-AM-GESG）算法、基于滤波（分解）的辅助模型递推广义增广最小二乘（F-AM-RGELS）算法.

14）研究下列多输入非线性自回归输出误差ARMA系统的（基于分解的、基于滤波的）辅助模型广义增广随机梯度算法和（基于分解的、基于滤波的）辅助模型递推广义增广最小二乘算法：

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

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

③ $A(z)y(t)=\sum_{j=1}^{r}\dfrac{B_j(z)}{F_j(z)}u_j(t)u_j(t-j)+\dfrac{D(z)}{C(z)}v(t)$.

④ $A(z)y^2(t)=\sum_{j=1}^{r}\dfrac{B_j(z)}{F_j(z)}u_j(t)u_{r-j+1}(t)+\dfrac{D(z)}{C(z)}v(t)$.

⑤ $A(z)f(y(t))=\sum_{j=1}^{r}\dfrac{B_j(z)}{F_j(z)}g_j(u_j(t))+\dfrac{D(z)}{C(z)}v(t)$.

⑥ $A(z)f(y(t))=\sum_{j=1}^{r}\dfrac{B_j(z)}{F_j(z)}g(u_1(t),u_2(t),\cdots,u_r(t))+\dfrac{D(z)}{C(z)}v(t)$.

⑦ $A(z)f(y(t))=\sum_{j=1}^{r}\dfrac{B_j(z)}{F_j(z)}g_j(u_1(t),u_2(t),\cdots,u_r(t))+\dfrac{D(z)}{C(z)}v(t)$.

15）研究下列多输入多输出自回归输出误差ARMA模型的（基于分解的、基于滤波的）辅助模型广义增广随机梯度算法和（基于分解的、基于滤波的）辅助模型递推广义增广最小二乘算法：

① $y_i(t)=\sum_{j=1}^{m}A_{ij}(z)y_j(t)+\dfrac{1}{F(z)}\sum_{j=1}^{r}B_{ij}(z)u_j(t)+\dfrac{D_i(z)}{C_i(z)}v_i(t)$，
$i=1,2,\cdots,m$，

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

$A_{ij}(z):=a_{ij}(1)z^{-1}+a_{ij}(2)z^{-2}+\cdots+a_{ij}(n_{ij})z^{-n_{ij}}$，

$B_{ij}(z):=b_{ij}(1)z^{-1}+b_{ij}(2)z^{-2}+\cdots+b_{ij}(n_{ij})z^{-n_{ij}}$，

$F_{ij}(z):=1+f_{ij}(1)z^{-1}+f_{ij}(2)z^{-2}+\cdots+f_{ij}(n_{ij})z^{-n_{ij}}$，

$C_i(z):=1+c_i(1)z^{-1}+c_i(2)z^{-2}+\cdots+c_i(n_i)z^{-n_i}$，

$D_i(z):=1+d_i(1)z^{-1}+d_i(2)z^{-2}+\cdots+d_i(n_i)z^{-n_i}$.

16）研究下列多输入多输出非线性自回归输出误差ARMA模型的（基于分解的、基于滤波的）辅助模型广义增广随机梯度算法和（基于分解的、基于滤波的）辅助模型递推广义增广最小二乘算法：

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

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

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

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

17）研究下列多输入线性参数自回归输出误差ARMA系统（LP-AR-OE ARMA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：

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

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

进一步可以推广到下列非线性噪声情形：

③ $A(z)y(t)=\dfrac{1}{F(z)}\sum_{j=1}^{r}B_j(z)u_j(t)y(t-j)+\dfrac{v(t)+d_1v(t-1)+d_2v(t-2)v(t-3)}{C(z)}$.

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：多新息辨识理论与方法[M].北京：科学出版社，2016
DING Feng.System identification：Multi-Innovation identification theory and methods[M]. Beijing：Science Press，2016

[4] Ding F，Liu X P，Liu G.Identification methods for Hammerstein nonlinear systems[J].Digital Signal Processing，2011，21(2)：215-238

[5] 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

[6] 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

[7] 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

[8] 丁锋.输出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(6):481-503
DING Feng. Multi-innovation identification methods for output-error type systems.Journal of Nanjing University of Information Science and Technology (Natural Science Edition),2015,7(6):481-503

[9] 丁锋,谢新民.传递函数阵子子模型参数递推估计:辅助模型方法[J].控制与决策,1991,6(6):447-452
DING Feng, XIE Xinmin. Recursive parameter estimation of transfer matrix sub-submodels: An auxiliary model method[J].Control and Decision,1991,6(6):447-452

[10] 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

[11] 王治祥,丁锋.*z-s* 变换及其应用[J].控制与决策.1995,10(1):89-92
WANG Zhixiang, DING Feng.*z-s* transform and its applications[J].Control and Decision,1995,10(1):89-92

[12] 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

[13] Ding F, Chen T. Combined parameter and output estimation of dual-rate systems using an auxiliary model [J].Automatica,2004,40(10):1739-1748

[14] 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

[15] 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

[16] Ding F, Shi Y, Chen T. Auxiliary model based least-squares identification methods for Hammerstein output-error systems[J].Systems and Control Letters,2007,56 (5):373-380

[17] 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

[18] 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

[19] 谢莉,王冬青,丁锋.随机干扰系统的辅助模型递推广义增广最小二乘辨识方法[J].科学技术与工程, 2008,8(14):3944-3945,3965
XIE Li, WANG Dongqing, DING Feng. AM-RGELS algorithms for general stochastic systems[J].Science Technology and Engineering,2008,8(14):3944-3945,3965

[20] 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

[21] 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

[22] Liu Y J, Xie L, Ding F. An auxiliary model recursive least squares parameter estimation algorithm for non-uniformly sampled multirate systems[J].Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering,2009,223(4):445-454

[23] Han L L, Sheng J, Ding F, et al. Auxiliary models based recursive least squares identification for multirate multi-input systems [J]. Mathematical and Computer Modelling,2009,50(7-8):1100-1106

[24] 刘艳君,谢莉,丁锋.非均匀采样数据系统的 AM-RLS 辨识方法及仿真研究[J].系统仿真学报,2009,21 (19):6186-6189
LIU Yanjun, XIE Li, DING Feng. AM-RLS identification and simulation studies for non-uniformly sampled-data systems[J].Journal of System Simulation,2009,21(19): 6186-6189

[25] 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

[26] Liu X G, Lu J. Least squares based iterative identification for a class of multirate systems[J].Automatica,2010,46 (3):549-554

[27] 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

[28] 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

[29] 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

［30］ Wang D Q，Chu Y Y，Yang G W，et al. Auxiliary model-based recursive generalized least squares parameter estimation for Hammerstein OEAR systems［J］. Mathematical and Computer Modelling，2010，52（1/2）：309-317

［31］ 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

［32］ 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

［33］ 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

［34］ Xiang L L，Xie L B，Ding R F. Hierarchical least squares algorithms for single-input multiple-output systems based on the auxiliary model［J］. Mathematical and Computer Modelling，2010，52（5/6）：918-924

［35］ 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

［36］ Ding F，Liu G，Liu X P. Parameter estimation with scarce measurements［J］. Automatica，2011，47（8）：1646-1655

［37］ Ding F，Liu X P，Liu G. Gradient based and least-squares based iterative identification methods for OE and OEMA systems［J］. Digital Signal Processing，2010，20（3）：664-677

［38］ 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

［39］ Han H Q，Song G L，Xiao Y S，et al. Performance analysis of the AM-SG parameter estimation for multivariable systems［J］. Applied Mathematics and Computation，2011，217（12）：5566-5572

［40］ Xie L，Yang H Z，Ding F. Recursive least squares parameter estimation for non-uniformly sampled systems based on the data filtering［J］. Mathematical and Computer Modelling，2011，54（1/2）：315-324

［41］ Xie L，Yang H Z. Gradient based iterative identification for non-uniform sampled output error systems［J］. Journal of Vibration and Control，2011，17（3）：471-478

［42］ Zhang Z N，Ding F，Liu X G. Hierarchical gradient based iterative parameter estimation algorithm for multivariable output error moving average systems［J］. Computers and Mathematics with Applications，2011，61（3）：672-682

［43］ Ding F，Gu Y. Performance analysis of the auxiliary model-based least-squares identification algorithm for one-step state-delay systems［J］. International Journal of Computer Mathematics，2012，89（15）：2019-2028

［44］ Wang D Q，Ding F. Hierarchical least squares estimation algorithm for Hammerstein-Wiener systems［J］. IEEE Signal Processing Letters，2012，19（12）：825-828

［45］ Gu Y，Ding F. Auxiliary model based least squares identification method for a state space model with a unit time-delay［J］. Applied Mathematical Modelling，2012，36（12）：5773-5779

［46］ Han L L，Wu F X，Sheng J，et al. Two recursive least squares parameter estimation algorithms for multirate multiple-input systems by using the auxiliary model［J］. Mathematics and Computers in Simulation，2012，82（5）：777-789

［47］ Duan H H，Jia J，Ding R F. Two-stage recursive least squares parameter estimation algorithm for output error models［J］. Mathematical and Computer Modelling，2012，55（3/4）：1151-1159

［48］ Zhang Z N，Jia J，Ding R F. Hierarchical least squares based iterative estimation algorithm for multivariable Box-Jenkins-like systems using the auxiliary model［J］. Applied Mathematical Computation，2012，218（9）：5580-5587

［49］ Ding F，Duan H H. Two-stage parameter estimation algorithms for Box-Jenkins systems［J］. IET Signal Processing，2013，7（2）：176-184.

［50］ Ding F，Gu Y. Performance analysis of the auxiliary model-based stochastic gradient parameter estimation algorithm for state space systems with one-step state delay［J］. Circuits，Systems and Signal Processing，2013，32（2）：585-599.

［51］ Hu P P，Ding F，Shen J. Auxiliary model based least squares parameter estimation algorithm for feedback nonlinear systems using the hierarchical identification principle［J］. Journal of the Franklin Institute：Engineering and Applied Mathematics，2013，350（10）：3248-3259

［52］ Ding F. Hierarchical parameter estimation algorithms for multivariable systems using measurement information［J］. Information Sciences，2014，277：396-405

［53］ 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

［54］ Wang C，Tang T. Recursive least squares estimation algorithm applied to a class of linear-in-parameters output error moving average systems［J］. Applied Mathematics Letters，2014，29：36-41

［55］ Wang C，Tang T. Several gradient-based iterative estimation algorithms for a class of nonlinear systems using the filtering technique［J］. Nonlinear Dynamics，2014，77（3）：769-780

［56］ 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

［57］ Guo L J，Ding F. Least squares based iterative algorithm for pseudo-linear autoregressive moving average systems using the data filtering technique［J］. Journal of the Franklin Institute：Engineering and Applied Mathematics，2015，352（10）：4339-4353

［58］ 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

［59］ Wang C，Zhu L. Parameter identification of a class of nonlinear systems based on the multi-innovation identification theory[J].Journal of the Franklin Institute: Engineering and Applied Mathematics,2015,352(10):4624-4637

Auxiliary model identification methods.Part A:Autoregressive output-error systems

DING Feng[1,2,3]

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 presents an auxiliary model (AM) based stochastic gradient (SG) algorithm,an AM multi-innovation SG algorithm and an AM recursive least squares algorithm for autoregressive output-error systems and presents a filtering based AM generalized extended SG algorithm,a filtering based AM multi-innovation generalized extended SG algorithm and a filtering based AM recursive generalized extended least squares algorithm for autoregressive output-error autoregressive moving average (AR-OEARMA) systems,namely autoregressive Box-Jenkins systems.

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