文章编号:1674-7070(2014)05-0385-20

丁锋[1,2,3]　王艳娇[1]

类多变量方程误差类系统的递阶多新息辨识方法

摘要

根据递阶辨识原理,研究了类多变量方程误差系统和类多变量方程误差ARMA系统递阶随机梯度方法和递阶梯度迭代方法、递阶最小二乘方法和递阶最小二乘迭代方法.进一步利用多新息辨识理论,推导了递阶多新息梯度辨识方法和递阶多新息最小二乘辨识方法.为减小计算量,推导了基于滤波的类多变量方程误差ARMA系统递阶辨识方法和递阶多新息辨识方法.讨论了几个典型辨识算法的计算量,并给出了计算参数估计的步骤.

关键词

参数估计;递推辨识;梯度搜索;最小二乘搜索;多新息辨识理论;递阶辨识原理;类多变量系统;数据滤波技术

中图分类号　TP273

文献标志码　A

收稿日期　2014-10-05

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

在现代大规模工业控制领域,广泛存在着结构复杂的多输入多输出系统.单输入单输出系统简称标量系统,多输入多输出系统简称多变量系统.不同于标量系统,多变量系统在于它的多个输入输出共同影响,呈现结构复杂性和耦合性,其规模性呈现维数高、参数数目多,干扰是各种各样的,呈现相关性和随机性.正是多变量系统结构的复杂性、干扰的多样性,使得多变量系统辨识具有诱人的魅力,也诞生出许多丰富多彩的辨识方法.本文第一作者最近出版的《系统辨识新论》和《系统辨识——辨识方法性能分析》详细介绍了一些辨识新方法,研究了辨识方法参数估计的收敛性问题[1-5],包括基于辅助模型辨识思想的辅助模型辨识方法[6]、基于梯度搜索和最小二乘搜索或牛顿搜索的迭代辨识方法[7]、基于多新息辨识理论的多新息辨识方法[8]、基于递阶辨识原理的递阶辨识方法[9]、基于耦合辨识概念的耦合辨识方法[10]等,其他一些辨识方法可参见笔者等在《南京信息工程大学学报》上刊登的连载论文[11-17].

随着自动化控制科学和计算技术的发展,尽管人们处理复杂问题的能力越来越强大,但是计算量往往是问题规模的非线性函数,甚至是平方、立方、多项式或指数函数.在这种情况下,人们自然会想到分解,将大规模问题分解为一些规模较小的子问题.例如,假设计算量是问题规模 n 的平方(即 n^2),分解为2个子问题,其规模分别为 n_1 和 n_2,2个子问题的计算量之和 $n_1^2+n_2^2$ 小于 n^2(注意 $n_1+n_2=n$),即使加上由于分解产生的附加计算量,一般子问题的计算量之和小于整体问题的计算量.分解可以减小问题的计算量,这种方法已经广泛用在各个领域,例如大系统的递阶控制是借助于"分解-协调原理".

本文第一作者主要研究系统辨识问题,1996年给清华大学硕士生、博士生讲授《大系统理论及应用》课程时,受大系统递阶控制的"分解-协调原理"的启发,决定将"分解"的思想引入到系统辨识中,经过深入长久的思考,终于取得了突破性进展,提出了基于分解的"递阶辨识原理"(hierarchical identification principle),开辟了递阶辨识研究领域.基于递阶辨识原理的标量系统和多变量系统的递阶辨识方法相继发表在国内外的著名学术期刊上,如《自动化学报》1999年第5期[18]、《Automatica》2005年第2期[19]、《IEEE Transactions on Automatic Control》2005年第3期[20].递阶辨识原理不仅能够解决大规模

多变量系统辨识方法计算量大的问题,而且能解决结构复杂非线性系统的辨识问题,还能用于复杂耦合矩阵方程的迭代解研究,如矩阵方程的递阶梯度迭代算法和递阶最小二乘迭代算法[21-25],类多变量系统的递阶梯度迭代算法和递阶最小二乘迭代算法[1,2,19-20,26-27],其他递阶辨识论文可参见文献[28-34].

众所周知,最小二乘辨识算法收敛速度快,但需要计算协方差阵,因而计算量大.为减小计算量,人们寻求别的方法,如随机梯度辨识算法,虽然其减小了计算量,但收敛速度慢,相同递推步数下参数估计精度低.为在计算量与辨识精度之间找到折衷,为改进随机梯度辨识算法的收敛速度,又不致于计算量大幅度增加,本文第一作者进行了细致深入的思考,通过借鉴标量—向量—矩阵的发展历程,引入新息长度,原创性地提出了多新息辨识理论(multi-innovation identification theory)和多新息辨识方法[1,35],包括多新息随机梯度辨识算法和多新息最小二乘辨识方法.新息是能够改变参数估计精度的有用信息.多新息随机梯度算法比随机梯度算法收敛速度快,参数估计精度高.最近,笔者的多新息随机梯度类辨识方法和多新息最小二乘辨识方法 Regular Paper 分别发表在控制领域国际著名期刊《Automatica》2007年第1期[36]和《IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics》2010 年第 3 期[37]上,基于辅助模型的稀少量测数据系统的多新息随机梯度辨识方法 Regular Paper 发表在《Automatica》2011 年第 8 期[38]上.多新息辨识理论可以用于研究线性系统和非线性系统的辨识问题,相关论文可参见文献[33,39-43].

本文针对类多变量方程误差系统和类多变量方程误差 ARMA 系统,利用递阶辨识原理、多新息辨识理论、数据滤波技术,研究和提出了递阶梯度类辨识方法、递阶最小二乘类辨识方法、递阶迭代辨识方法、递阶最小二乘迭代辨识方法,递阶多新息梯度类辨识方法、递阶多新息最小二乘类辨识方法,以及基于数据滤波的递阶辨识方法和基于滤波的递阶多新息辨识方法,并讨论了几个典型辨识算法的计算量,给出了计算步骤.由于篇幅限制,本文只给出一些辨识算法,省略了具体推导过程,其细节可参见即将出版的《系统辨识——多新息辨识方法》一书(科学出版社,2015).

1 类多变量方程误差系统描述

考虑类多变量方程误差系统(multivariable equation-error-like system),即类多变量受控自回归系统(multivariable controlled autoregressive-like system,多变量 CAR-like 系统),或称为类多变量 ARX 系统(multivariable ARX-like system)[1,2,19-20,30]:

$$\alpha(z)\boldsymbol{y}(t)=\boldsymbol{Q}(z)\boldsymbol{u}(t)+\boldsymbol{v}(t), \tag{1}$$

其中 $\boldsymbol{y}(t):=[y_1(t),y_2(t),\cdots,y_m(t)]^{\mathrm{T}}\in\mathbf{R}^m$ 为 m 维观测输出向量,$\boldsymbol{v}(t):=[v_1(t),v_2(t),\cdots,v_m(t)]^{\mathrm{T}}\in\mathbf{R}^m$ 为 r 维观测输入向量,$\boldsymbol{v}(t):=[v_1(t),v_2(t),\cdots,v_m(t)]^{\mathrm{T}}\in\mathbf{R}^m$ 为 m 维白噪声向量,$\alpha(z)$ 为单位后移算子 z^{-1} 的特征多项式,$\boldsymbol{Q}(z)$ 为单位后移算子 z^{-1} 的多项式矩阵,它们可以表示为

$$\alpha(z):=1+\alpha_1z^{-1}+\alpha_2z^{-2}+\cdots+\alpha_nz^{-n},\quad \alpha_i\in\mathbf{R}, \tag{2}$$

$$\boldsymbol{Q}(z):=\boldsymbol{Q}_1z^{-1}+\boldsymbol{Q}_2z^{-2}+\cdots+\boldsymbol{Q}_nz^{-n},\quad \boldsymbol{Q}_i\in\mathbf{R}^{m\times r}. \tag{3}$$

将式(2)—(3)代入式(1)得到

$$\boldsymbol{y}(t)+\alpha_1\boldsymbol{y}(t-1)+\alpha_2\boldsymbol{y}(t-2)+\cdots+\alpha_n\boldsymbol{y}(t-n)=\boldsymbol{Q}_1\boldsymbol{u}(t-1)+\boldsymbol{Q}_2\boldsymbol{u}(t-2)+\cdots+\boldsymbol{Q}_n\boldsymbol{u}(t-n)+\boldsymbol{v}(t).$$

定义模型参数向量 $\boldsymbol{\alpha}$,参数矩阵 $\boldsymbol{\theta}$,输入信息向量 $\boldsymbol{\varphi}(t)$ 和输出信息矩阵 $\boldsymbol{\psi}(t)$ 如下:

$$\boldsymbol{\alpha}:=[\alpha_1,\alpha_2,\cdots,\alpha_n]^{\mathrm{T}}\in\mathbf{R}^n,$$

$$\boldsymbol{\theta}^{\mathrm{T}}:=[\boldsymbol{Q}_1,\boldsymbol{Q}_2,\cdots,\boldsymbol{Q}_n]\in\mathbf{R}^{m\times(nr)},$$

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

$$\boldsymbol{\psi}(t):=[\boldsymbol{y}(t-1),\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(t-n)]\in\mathbf{R}^{m\times n}.$$

于是,可以得到类多变量方程误差系统(1)的递阶辨识模型(hierarchical identification model):

$$\boldsymbol{y}(t)+\boldsymbol{\psi}(t)\boldsymbol{\alpha}=\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}(t)+\boldsymbol{v}(t). \tag{4}$$

类多变量方程误差系统的递阶辨识模型(4)既包含参数向量 $\boldsymbol{\alpha}$,又包含参数矩阵 $\boldsymbol{\theta}$,这是辨识的困难.解决这一辨识困难办法之一是用 Kronecker 积将这个模型中的参数向量 $\boldsymbol{\alpha}$ 和参数矩阵 $\boldsymbol{\theta}$ 化为一个大参数向量的多元辨识模型(multivariate identification model).下面介绍类多变量方程误差系统的多元辨识方法.

2 类多变量方程误差系统多元辨识方法

为了便于辨识,使用 Kronecker 积,将模型(4)中的信息向量 $\boldsymbol{\varphi}(t)$ 和信息矩阵 $\boldsymbol{\psi}(t)$ 化为一个大信息矩阵 $\boldsymbol{\Phi}(t)$,将参数向量 $\boldsymbol{\alpha}$ 和参数矩阵 $\boldsymbol{\theta}$ 化为一个大参数向量 $\boldsymbol{\vartheta}$.定义信息矩阵 $\boldsymbol{\Phi}(t)$ 和参数向量 $\boldsymbol{\vartheta}$ 如下:

$$\boldsymbol{\Phi}(t):=[-\boldsymbol{\psi}(t),\boldsymbol{I}_m\otimes\boldsymbol{\varphi}^{\mathrm{T}}(t)]\in\mathbf{R}^{m\times n_0},\quad n_0:=n+mnr,$$

$$\boldsymbol{\vartheta}:=\boldsymbol{v}\begin{bmatrix}\boldsymbol{\alpha}\\ \mathrm{col}[\boldsymbol{\theta}]\end{bmatrix}\in\mathbf{R}^{n_0},$$

或

$$\boldsymbol{\Phi}(t):=[-\boldsymbol{\psi}(t),\boldsymbol{\varphi}^{\mathrm{T}}(t)\otimes\boldsymbol{I}_m]\in\mathbf{R}^{m\times n_0},\quad n_0:=n+mnr,$$

$$\boldsymbol{\vartheta}:=\boldsymbol{\nu}\begin{bmatrix}\boldsymbol{\alpha}\\ \mathrm{col}[\boldsymbol{\theta}^{\mathrm{T}}]\end{bmatrix}\in\mathbf{R}^{n_0},$$

则递阶辨识模型(4)可以写为下列多元线性回归模型(multivariate linear regressive model)或多元辨识模型(multivariate identification model):

$$\boldsymbol{y}(t)=\boldsymbol{\Phi}(t)\boldsymbol{\vartheta}+\boldsymbol{v}(t).\tag{5}$$

假设阶次 m,n 和 r 已知,且当 $t\leqslant 0$ 时,$\boldsymbol{y}(t)=\mathbf{0}$,$\boldsymbol{\Phi}(t)=\mathbf{0}$,$\boldsymbol{v}(t)=\mathbf{0}$.辨识的目标就是利用梯度搜索、最小二乘原理、多新息辨识理论,研究和提出辨识方法,从系统的观测数据$\{\boldsymbol{y}(t),\boldsymbol{\Phi}(t):t=1,2,3,\cdots\}$中提炼出系统的参数向量 $\boldsymbol{\vartheta}$ 信息.

2.1 多元辨识方法

2.1.1 多元随机梯度辨识算法

令 $\|\boldsymbol{X}\|^2=\mathrm{tr}[\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}]$ 表示矩阵 $\boldsymbol{X}$ 的范数.令 $\hat{\boldsymbol{\vartheta}}(t)$ 为参数向量 $\boldsymbol{\vartheta}$ 在时刻 t 的估计.参考文献[1,14],能够获得估计多元线性回归辨识模型(5)参数向量 $\boldsymbol{\vartheta}$ 的多元随机梯度算法(Multivariate Stochastic Gradient algorithm,M-SG 算法)[2]:

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

$$\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_{n_0}/p_0,\tag{6}$$

$$r(t)=r(t-1)+\|\boldsymbol{\Phi}(t)\|^2,\quad r(0)=1,\tag{7}$$

$$\boldsymbol{\Phi}(t)=[-\boldsymbol{\psi}(t),\boldsymbol{I}_m\otimes\boldsymbol{\varphi}^{\mathrm{T}}(t)],\tag{8}$$

$$\boldsymbol{\varphi}(t)=[\boldsymbol{u}^{\mathrm{T}}(t-1),\boldsymbol{u}^{\mathrm{T}}(t-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(t-n)]^{\mathrm{T}},\tag{9}$$

$$\boldsymbol{\psi}(t)=[\boldsymbol{y}(t-1),\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(t-n)].\tag{10}$$

2.1.2 多元递推最小二乘辨识算法

根据式(5),能够得到估计参数向量 $\boldsymbol{\vartheta}$ 的多元递推最小二乘算法(Multivariate Recursive Least Squares algorithm,M-RLS 算法)[2]:

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\boldsymbol{L}(t)[\boldsymbol{y}(t)-\boldsymbol{\Phi}(t)\hat{\boldsymbol{\vartheta}}(t-1)],$$

$$\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_{n_0}/p_0,\tag{11}$$

$$\boldsymbol{L}(t)=\boldsymbol{P}(t-1)\boldsymbol{\Phi}^{\mathrm{T}}(t)[\boldsymbol{I}+\boldsymbol{\Phi}(t)\boldsymbol{P}(t-1)\boldsymbol{\Phi}^{\mathrm{T}}(t)]^{-1},\tag{12}$$

$$\boldsymbol{P}(t)=\boldsymbol{P}(t-1)-\boldsymbol{L}(t)\boldsymbol{\Phi}(t)\boldsymbol{P}(t-1),\boldsymbol{P}(0)=p_0\boldsymbol{I}_{n_0},\tag{13}$$

$$\boldsymbol{\Phi}(t)=[-\boldsymbol{\psi}(t),\boldsymbol{I}_m\otimes\boldsymbol{\varphi}^{\mathrm{T}}(t)],\tag{14}$$

$$\boldsymbol{\varphi}(t)=[\boldsymbol{u}^{\mathrm{T}}(t-1),\boldsymbol{u}^{\mathrm{T}}(t-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(t-n)]^{\mathrm{T}},\tag{15}$$

$$\boldsymbol{\psi}(t)=[\boldsymbol{y}(t-1),\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(t-n)].\tag{16}$$

2.2 多元多新息辨识方法

与多元递推最小二乘算法相比,多元随机梯度算法的计算量小、收敛速度慢.为了改进多元随机梯度辨识算法的收敛速度,根据多新息辨识理论,引入新息长度,将新息向量扩展为一个大新息向量,得到多元多新息辨识方法.

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

$$\boldsymbol{\Gamma}(p,t):=[\boldsymbol{\Phi}^{\mathrm{T}}(t),\boldsymbol{\Phi}^{\mathrm{T}}(t-1),\cdots,\boldsymbol{\Phi}^{\mathrm{T}}(t-p+1)]\in\mathbf{R}^{n_0\times(mp)},$$

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

令新息长度 $p\geqslant 1$.应用多新息辨识理论,将多元辨识算法(6)中新息向量 $\boldsymbol{e}(t):=\boldsymbol{y}(t)-\boldsymbol{\Phi}(t)\hat{\boldsymbol{\vartheta}}(t-1)\in\mathbf{R}^m$ 扩展为一个大新息向量

$$\boldsymbol{E}(p,t):=\begin{bmatrix}\boldsymbol{y}(t)-\boldsymbol{\Phi}(t)\hat{\boldsymbol{\vartheta}}(t-1)\\ \boldsymbol{y}(t-1)-\boldsymbol{\Phi}(t-1)\hat{\boldsymbol{\vartheta}}(t-1)\\ \vdots\\ \boldsymbol{y}(t-p+1)-\boldsymbol{\Phi}(t-p+1)\hat{\boldsymbol{\vartheta}}(t-1)\end{bmatrix}=$$

$$\boldsymbol{Y}(p,t)-\boldsymbol{\Gamma}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\vartheta}}(t-1)\in\mathbf{R}^{mp}.\tag{17}$$

2.2.1 多元多新息随机梯度辨识算法

借助于多新息辨识理论[1,16,36,38,42],推广多元随机梯度辨识算法(6)—(10),可以得到估计多元系统(5)参数向量 $\boldsymbol{\vartheta}$ 的新息长度为 p 的多元多新息随机梯度算法(Multivariate Multi-Innovation Stochastic Gradient algorithm,M-MISG 算法)[2]:

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\frac{\boldsymbol{\Gamma}(p,t)}{r(t)}\boldsymbol{E}(p,t),\quad\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_{n_0}/p_0,\tag{18}$$

$$\boldsymbol{E}(p,t)=\boldsymbol{Y}(p,t)-\boldsymbol{\Gamma}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\vartheta}}(t-1),\tag{19}$$

$$r(t)=r(t-1)+\|\boldsymbol{\Phi}(t)\|^2,\quad r(0)=1,\tag{20}$$

$$\boldsymbol{Y}(p,t)=[\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots,\boldsymbol{y}^{\mathrm{T}}(t-p+1)]^{\mathrm{T}},\tag{21}$$

$$\boldsymbol{\Gamma}(p,t)=[\boldsymbol{\Phi}^{\mathrm{T}}(t),\boldsymbol{\Phi}^{\mathrm{T}}(t-1),\cdots,\boldsymbol{\Phi}^{\mathrm{T}}(t-p+1)],\tag{22}$$

$$\boldsymbol{\Phi}(t)=[-\boldsymbol{\psi}(t),\boldsymbol{I}_m\otimes\boldsymbol{\varphi}^{\mathrm{T}}(t)],\tag{23}$$

$$\boldsymbol{\varphi}(t)=[\boldsymbol{u}^{\mathrm{T}}(t-1),\boldsymbol{u}^{\mathrm{T}}(t-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(t-n)]^{\mathrm{T}},\tag{24}$$

$$\boldsymbol{\psi}(t)=[\boldsymbol{y}(t-1),\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(t-n)].\tag{25}$$

当新息长度 $p=1$ 时,多元多新息随机梯度算法退化为多元随机梯度算法.

2.2.2 多元多新息递推最小二乘辨识算法

借助于多新息辨识理论[1,16,37,42],根据多元递推最小二乘辨识算法(11)—(16),可以得到新息长度为 p 的多元多新息递推最小二乘算法(Multivariate Multi-Innovation Least Squares algorithm,M-MILS 算法)[1]:

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\boldsymbol{L}(t)\boldsymbol{E}(p,t),\quad\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_{n_0}/p_0,\tag{26}$$

$$\boldsymbol{E}(p,t)=\boldsymbol{Y}(p,t)-\boldsymbol{\Gamma}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\vartheta}}(t-1),\tag{27}$$

$$\boldsymbol{L}(t)=\boldsymbol{P}(t-1)\boldsymbol{\Gamma}(p,t)[\boldsymbol{I}+\boldsymbol{\Gamma}^{\mathrm{T}}(p,t)\boldsymbol{P}(t-1)\boldsymbol{\Gamma}(p,t)]^{-1},\tag{28}$$

$$\boldsymbol{P}(t)=\boldsymbol{P}(t-1)-\boldsymbol{L}(t)\boldsymbol{\Gamma}^{\mathrm{T}}(p,t)\boldsymbol{P}(t-1),$$

$$\boldsymbol{P}(0)=p_0\boldsymbol{I}_{n_0}, \tag{29}$$

$$\boldsymbol{Y}(p,t)=[\boldsymbol{y}^{\mathrm{T}}(t),\boldsymbol{y}^{\mathrm{T}}(t-1),\cdots,\boldsymbol{y}^{\mathrm{T}}(t-p+1)]^{\mathrm{T}}, \tag{30}$$

$$\boldsymbol{\Gamma}(p,t)=[\boldsymbol{\Phi}^{\mathrm{T}}(t),\boldsymbol{\Phi}^{\mathrm{T}}(t-1),\cdots,\boldsymbol{\Phi}^{\mathrm{T}}(t-p+1)], \tag{31}$$

$$\boldsymbol{\Phi}(t)=[-\boldsymbol{\psi}(t),\boldsymbol{I}_m\otimes\boldsymbol{\varphi}^{\mathrm{T}}(t)], \tag{32}$$

$$\boldsymbol{\varphi}(t)=[\boldsymbol{u}^{\mathrm{T}}(t-1),\boldsymbol{u}^{\mathrm{T}}(t-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(t-n)]^{\mathrm{T}}, \tag{33}$$

$$\boldsymbol{\psi}(t)=[\boldsymbol{y}(t-1),\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(t-n)]. \tag{34}$$

2.3 变递推间隔多元多新息辨识方法

为了处理数据丢失的情况,定义一个整数序列(integer sequence)$\{t_s,s=0,1,2,\cdots\}$满足

$$0=t_0<t_1<t_2<t_3<\cdots<t_{s-1}<t_s<\cdots,$$

且 $t_s^*:=t_s-t_{s-1}\geqslant 1$,使得 $t=t_s(s=1,2,3,\cdots)$ 时,$\boldsymbol{y}(t)$ 可得到,即对任意 $s=1,2,3,\cdots,\boldsymbol{y}(t_s)$ 都可得到.

2.3.1 变递推间隔多元多新息投影算法

参考文献[1,16,42],能够得到变递推间隔多元多新息投影算法(interval-Varying Multivariate Multi-Innovation Projection algorithm, V-M-MI-Proj 算法)[1,42]:

$$\hat{\boldsymbol{\vartheta}}(t_s)=\hat{\boldsymbol{\vartheta}}(t_{s-1})+\mu(t_s)\boldsymbol{\Gamma}(p,t_s)\boldsymbol{E}(p,t_s),\quad s=1,2,3,\cdots, \tag{35}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t_s),\quad t\in T_s:=\{t_s,t_s+1,\cdots,t_{s+1}-1\}, \tag{36}$$

$$\mu(t_s)=\frac{\|\boldsymbol{\Gamma}(p,t_s)\boldsymbol{E}(p,t_s)\|^2}{\|\boldsymbol{\Gamma}^{\mathrm{T}}(p,t_s)\boldsymbol{\Gamma}(p,t_s)\boldsymbol{E}(p,t_s)\|^2}, \tag{37}$$

$$\boldsymbol{E}(p,t_s)=\boldsymbol{Y}(p,t_s)-\boldsymbol{\Gamma}^{\mathrm{T}}(p,t_s)\hat{\boldsymbol{\vartheta}}(t_{s-1}), \tag{38}$$

$$\boldsymbol{\Gamma}(p,t_s)=[\boldsymbol{\Phi}^{\mathrm{T}}(t_s),\boldsymbol{\Phi}^{\mathrm{T}}(t_s-1),\cdots,\boldsymbol{\Phi}^{\mathrm{T}}(t_s-p+1)], \tag{39}$$

$$\boldsymbol{Y}(p,t_s)=[\boldsymbol{y}^{\mathrm{T}}(t_s),\boldsymbol{y}^{\mathrm{T}}(t_s-1),\cdots,\boldsymbol{y}^{\mathrm{T}}(t_s-p+1)]^{\mathrm{T}}, \tag{40}$$

$$\boldsymbol{\Phi}(t_s)=[-\boldsymbol{\psi}(t_s),\boldsymbol{I}_m\otimes\boldsymbol{\varphi}^{\mathrm{T}}(t_s)], \tag{41}$$

$$\boldsymbol{\varphi}(t_s)=[\boldsymbol{u}^{\mathrm{T}}(t_s-1),\boldsymbol{u}^{\mathrm{T}}(t_s-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(t_s-n)]^{\mathrm{T}}, \tag{42}$$

$$\boldsymbol{\psi}(t_s)=[\boldsymbol{y}(t_s-1),\boldsymbol{y}(t_s-2),\cdots,\boldsymbol{y}(t_s-n)]. \tag{43}$$

当 $t_s^*=p=1$ 时,多元多新息算法就是常规的单新息多元投影算法.式(36)表示在数据丢失的区间内,参数估计保持不变.这个算法计算收敛因子(步长)$\mu(t_s)$比较复杂,下面给出简化的变递推间隔多元多新息投影算法(V-M-MI-Proj 算法)[1,42]:

$$\hat{\boldsymbol{\vartheta}}(t_s)=\hat{\boldsymbol{\vartheta}}(t_{s-1})+\frac{\boldsymbol{\Gamma}(p,t_s)}{\|\boldsymbol{\Gamma}(p,t_s)\|^2}\boldsymbol{E}(p,t_s),\quad s=1,2,3,\cdots, \tag{44}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t_s),\quad t\in T_s:=\{t_s,t_s+1,\cdots,t_{s+1}-1\}, \tag{45}$$

$$\boldsymbol{E}(p,t_s)=\boldsymbol{Y}(p,t_s)-\boldsymbol{\Gamma}^{\mathrm{T}}(p,t_s)\hat{\boldsymbol{\vartheta}}(t_{s-1}), \tag{46}$$

$$\boldsymbol{\Gamma}(p,t_s)=[\boldsymbol{\Phi}^{\mathrm{T}}(t_s),\boldsymbol{\Phi}^{\mathrm{T}}(t_s-1),\cdots,\boldsymbol{\Phi}^{\mathrm{T}}(t_s-p+1)], \tag{47}$$

$$\boldsymbol{Y}(p,t_s)=[\boldsymbol{y}^{\mathrm{T}}(t_s),\boldsymbol{y}^{\mathrm{T}}(t_s-1),\cdots,\boldsymbol{y}^{\mathrm{T}}(t_s-p+1)]^{\mathrm{T}}, \tag{48}$$

$$\boldsymbol{\Phi}(t_s)=[-\boldsymbol{\psi}(t_s),\boldsymbol{I}_m\otimes\boldsymbol{\varphi}^{\mathrm{T}}(t_s)], \tag{49}$$

$$\boldsymbol{\varphi}(t_s)=[\boldsymbol{u}^{\mathrm{T}}(t_s-1),\boldsymbol{u}^{\mathrm{T}}(t_s-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(t_s-n)]^{\mathrm{T}}, \tag{50}$$

$$\boldsymbol{\psi}(t_s)=[\boldsymbol{y}(t_s-1),\boldsymbol{y}(t_s-2),\cdots,\boldsymbol{y}(t_s-n)]. \tag{51}$$

2.3.2 变递推间隔多元多新息广义投影算法

多元多新息投影辨识算法(26)—(34)可以推广为变递推间隔多元多新息广义投影算法(interval-Varying Multivariate Multi-Innovation Generalized Projection algorithm, V-M-MI-GP 算法)[1,35,38]:

$$\hat{\boldsymbol{\vartheta}}(t_s)=\hat{\boldsymbol{\vartheta}}(t_{s-1})+\frac{\boldsymbol{\Gamma}(p,t_s)}{r(q,t_s)}[\boldsymbol{Y}(p,t_s)-\boldsymbol{\Gamma}^{\mathrm{T}}(p,t_s)\hat{\boldsymbol{\vartheta}}(t_{s-1})],\quad s=1,2,3,\cdots, \tag{52}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t_s),\quad t\in T_s=\{t_s,t_s+1,\cdots,t_{s+1}-1\}, \tag{53}$$

$$r(q,t_s)=Tr[\boldsymbol{\Gamma}(q,t_s)\boldsymbol{\Gamma}^{\mathrm{T}}(q,t_s)],\quad q\geqslant p, \tag{54}$$

$$\boldsymbol{\Gamma}(p,t_s)=[\boldsymbol{\Phi}^{\mathrm{T}}(t_s),\boldsymbol{\Phi}^{\mathrm{T}}(t_s-1),\cdots,\boldsymbol{\Phi}^{\mathrm{T}}(t_s-p+1)], \tag{55}$$

$$\boldsymbol{Y}(p,t_s)=[\boldsymbol{y}^{\mathrm{T}}(t_s),\boldsymbol{y}^{\mathrm{T}}(t_s-1),\cdots,\boldsymbol{y}^{\mathrm{T}}(t_s-p+1)]^{\mathrm{T}}, \tag{56}$$

$$\boldsymbol{\Phi}(t_s)=[-\boldsymbol{\psi}(t_s),\boldsymbol{I}_m\otimes\boldsymbol{\varphi}^{\mathrm{T}}(t_s)], \tag{57}$$

$$\boldsymbol{\varphi}(t_s)=[\boldsymbol{u}^{\mathrm{T}}(t_s-1),\boldsymbol{u}^{\mathrm{T}}(t_s-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(t_s-n)]^{\mathrm{T}}, \tag{58}$$

$$\boldsymbol{\psi}(t_s)=[\boldsymbol{y}(t_s-1),\boldsymbol{y}(t_s-2),\cdots,\boldsymbol{y}(t_s-n)]. \tag{59}$$

变递推间隔多新息广义投影辨识算法可以通过加大 q 来克服噪声敏感性,适当选择 q 可以减小时变多元系统参数估计误差上界.

2.3.3 变递推间隔多元多新息随机梯度算法

进一步可推广为变递推间隔多元多新息随机梯度算法(interval-Varying Multivariate Multi-Innovation Stochastic Gradient algorithm, V-M-MISG 算法)[1,35,38,42]:

$$\hat{\boldsymbol{\vartheta}}(t_s)=\hat{\boldsymbol{\vartheta}}(t_{s-1})+\frac{\boldsymbol{\Gamma}(p,t_s)}{r(t_s)}[\boldsymbol{Y}(p,t_s)-\boldsymbol{\Gamma}^{\mathrm{T}}(p,t_s)\hat{\boldsymbol{\vartheta}}(t_{s-1})],\quad s=1,2,3,\cdots, \tag{60}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t_s),\quad t\in T_s=\{t_s,t_s+1,\cdots,t_{s+1}-1\}, \tag{61}$$

$$r(t_s)=\|\boldsymbol{\Gamma}(s,t_s)\|^2=\sum_{i=0}^{s-1}\|\boldsymbol{\Phi}(t_s-i)\|^2, \tag{62}$$

$$\boldsymbol{\Gamma}(p,t_s)=[\boldsymbol{\Phi}^{\mathrm{T}}(t_s),\boldsymbol{\Phi}^{\mathrm{T}}(t_s-1),\cdots,\boldsymbol{\Phi}^{\mathrm{T}}(t_s-p+1)], \tag{63}$$

$$\boldsymbol{Y}(p,t_s)=[\boldsymbol{y}^{\mathrm{T}}(t_s),\boldsymbol{y}^{\mathrm{T}}(t_s-1),\cdots,\boldsymbol{y}^{\mathrm{T}}(t_s-p+1)]^{\mathrm{T}}, \tag{64}$$

$$\boldsymbol{\Phi}(t_s)=[-\boldsymbol{\psi}(t_s),\boldsymbol{I}_m\otimes\boldsymbol{\varphi}^{\mathrm{T}}(t_s)], \tag{65}$$

$$\boldsymbol{\varphi}(t_s)=[\boldsymbol{u}^{\mathrm{T}}(t_s-1),\boldsymbol{u}^{\mathrm{T}}(t_s-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(t_s-n)]^{\mathrm{T}}, \tag{66}$$

$$\boldsymbol{\psi}(t_s)=[\boldsymbol{y}(t_s-1),\boldsymbol{y}(t_s-2),\cdots,\boldsymbol{y}(t_s-n)]. \tag{67}$$

当递推间隔 $t_s^*\equiv 1$ 时,V-M-MISG 算法退化为 M-MISG 算法.随着递推步数 s 的增加,$r(t_s)\to\infty$,算法增益矩阵 $\boldsymbol{\Gamma}(p,t_s)/r(t_s)$ 趋于零,故变递推间隔多新息随机梯度辨识算法没有跟踪时变参数的能力.当然,也可以推导变递推间隔修正多元多新息随机梯度辨识算法和变递推间隔遗忘因子多元多新息随机梯度辨识算法.

2.3.4 变递推间隔多元多新息最小二乘算法

借助于最小二乘原理,将 V-M-MI-GP 算法进一

步推广,能够得到变递推间隔多元多新息递推最小二乘算法(interval-Varying Multivariate Multi-Innovation Least Squares algorithm, V-M-MILS 算法)[1,38,42]:

$$\hat{\boldsymbol{\vartheta}}(t_s)=\hat{\boldsymbol{\vartheta}}(t_{s-1})+\boldsymbol{L}(t_s)[\boldsymbol{Y}(p,t_s)-\boldsymbol{\Gamma}^{\mathrm{T}}(p,t_s)\hat{\boldsymbol{\vartheta}}(t_{s-1})],\quad s=1,2,3,\cdots, \tag{68}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t_s),\quad t\in T_s=\{t_s,t_s+1,\cdots,t_{s+1}-1\}, \tag{69}$$

$$\boldsymbol{L}(t_s)=\boldsymbol{P}(t_s)\boldsymbol{\Gamma}(p,t_s)=\boldsymbol{P}(t_{s-1})\boldsymbol{\Gamma}(p,t_s)[\boldsymbol{I}_{mp}+\boldsymbol{\Gamma}^{\mathrm{T}}(p,t_s)\boldsymbol{P}(t_{s-1})\boldsymbol{\Gamma}(p,t_s)]^{-1}, \tag{70}$$

$$\boldsymbol{P}(t_s)=\boldsymbol{P}(t_{s-1})-\boldsymbol{L}(t_s)\boldsymbol{\Gamma}^{\mathrm{T}}(p,t_s)\boldsymbol{P}(t_{s-1}),\quad \boldsymbol{P}(0)=p_0\boldsymbol{I}_{n_0}, \tag{71}$$

$$\boldsymbol{\Gamma}(p,t_s)=[\boldsymbol{\Phi}^{\mathrm{T}}(t_s),\boldsymbol{\Phi}^{\mathrm{T}}(t_s-1),\cdots,\boldsymbol{\Phi}^{\mathrm{T}}(t_s-p+1)], \tag{72}$$

$$\boldsymbol{Y}(p,t_s)=[\boldsymbol{y}^{\mathrm{T}}(t_s),\boldsymbol{y}^{\mathrm{T}}(t_s-1),\cdots,\boldsymbol{y}^{\mathrm{T}}(t_s-p+1)]^{\mathrm{T}}, \tag{73}$$

$$\boldsymbol{\Phi}(t_s)=[-\boldsymbol{\psi}(t_s),\boldsymbol{I}_m\otimes\boldsymbol{\varphi}^{\mathrm{T}}(t_s)], \tag{74}$$

$$\boldsymbol{\varphi}(t_s)=[\boldsymbol{u}^{\mathrm{T}}(t_s-1),\boldsymbol{u}^{\mathrm{T}}(t_s-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(t_s-n)]^{\mathrm{T}}, \tag{75}$$

$$\boldsymbol{\psi}(t_s)=[\boldsymbol{y}(t_s-1),\boldsymbol{y}(t_s-2),\cdots,\boldsymbol{y}(t_s-n)]. \tag{76}$$

在这些变递推间隔的辨识算法中,递推间隔 t_s 不必取连续的自然数,可以是变化的.当遇到坏数据或不可信数据时,可以跳过这部分数据,因而具有克服损失数据的能力.

3 类多变量方程误差系统的递阶辨识方法

考虑类多变量输出误差系统(1)对应的递阶辨识模型(5),重写如下:

$$\boldsymbol{y}(t)+\boldsymbol{\psi}(t)\boldsymbol{\alpha}=\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}(t)+\boldsymbol{v}(t), \tag{77}$$

其中各变量的定义同上.

辨识模型(77)中包含了一个参数向量 $\boldsymbol{\alpha}$ 和一个参数矩阵 $\boldsymbol{\theta}$,辨识的目标是基于递阶辨识原理,利用系统的观测数据$\{\boldsymbol{u}(t),\boldsymbol{y}(t):t=1,2,3,\cdots\}$,提出递阶梯度辨识方法和递阶最小二乘辨识方法,估计系统模型参数向量 $\boldsymbol{\alpha}$ 和参数矩阵 $\boldsymbol{\theta}$.

3.1 递阶随机梯度辨识方法

根据式(77),定义准则函数(criterion function)

$$J_1(\boldsymbol{\alpha},\boldsymbol{\theta}):=\|\boldsymbol{y}(t)+\boldsymbol{\psi}(t)\boldsymbol{\alpha}-\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}(t)\|^2.$$

令 $\hat{\boldsymbol{\alpha}}(t)$ 和 $\hat{\boldsymbol{\theta}}(t)$ 分别为 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\theta}$ 在时刻 t 的估计.根据递阶辨识原理和梯度搜索,极小化准则函数 $J_1(\boldsymbol{\alpha},\boldsymbol{\theta})$,可以得到估计参数向量 $\boldsymbol{\alpha}$ 和参数矩阵 $\boldsymbol{\theta}$ 的递阶随机梯度算法(Hierarchical Stochastic Gradient algorithm, HSG 算法)[1,2,19]:

$$\hat{\boldsymbol{\alpha}}(t)=\hat{\boldsymbol{\alpha}}(t-1)-\frac{\boldsymbol{\psi}^{\mathrm{T}}(t)}{r(t)}[\boldsymbol{y}(t)+\boldsymbol{\psi}(t)\hat{\boldsymbol{\alpha}}(t-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t)],\quad t=1,2,3,\cdots, \tag{78}$$

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\boldsymbol{\varphi}(t)}{r(t)}[\boldsymbol{y}(t)+\boldsymbol{\psi}(t)\hat{\boldsymbol{\alpha}}(t-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t)]^{\mathrm{T}}, \tag{79}$$

$$r(t)=r(t-1)+\|\boldsymbol{\psi}(t)\|^2+\|\boldsymbol{\varphi}(t)\|^2,\quad r(0)=1, \tag{80}$$

$$\boldsymbol{\varphi}(t)=[\boldsymbol{u}^{\mathrm{T}}(t-1),\boldsymbol{u}^{\mathrm{T}}(t-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(t-n)]^{\mathrm{T}}, \tag{81}$$

$$\boldsymbol{\psi}(t)=[\boldsymbol{y}(t-1),\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(t-n)]. \tag{82}$$

式(80)也可以修改为

$$r(t)=r(t-1)+\max[\|\boldsymbol{\psi}(t)\|^2,\|\boldsymbol{\varphi}(t)\|^2],\quad r(0)=1. \tag{83}$$

3.2 递阶最小二乘辨识方法

定义和极小化最小二乘准则函数

$$J_2(\boldsymbol{\alpha},\boldsymbol{\theta}):=\sum_{j=1}^{t}\|\boldsymbol{y}(j)+\boldsymbol{\psi}(j)\boldsymbol{\alpha}-\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}(j)\|^2,$$

根据递阶辨识原理和最小二乘原理,极小化准则函数 $J_2(\boldsymbol{\alpha},\boldsymbol{\theta})$,可以推导出估计参数向量 $\boldsymbol{\alpha}$ 和参数矩阵 $\boldsymbol{\theta}$ 的递阶最小二乘算法(Hierarchical Least Squares algorithm, HLS 算法)[1,2,20,30]:

$$\hat{\boldsymbol{\alpha}}(t)=\hat{\boldsymbol{\alpha}}(t-1)+\boldsymbol{L}_1(t)[\boldsymbol{y}(t)+\boldsymbol{\psi}(t)\hat{\boldsymbol{\alpha}}(t-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t)], \tag{84}$$

$$\boldsymbol{L}_1(t)=-\boldsymbol{P}_1(t-1)\boldsymbol{\psi}^{\mathrm{T}}(t)[\boldsymbol{I}_m+\boldsymbol{\psi}(t)\boldsymbol{P}_1(t-1)\boldsymbol{\psi}^{\mathrm{T}}(t)]^{-1}, \tag{85}$$

$$\boldsymbol{P}_1(t)=[\boldsymbol{I}_n+\boldsymbol{L}_1(t)\boldsymbol{\psi}(t)]\boldsymbol{P}_1(t-1), \tag{86}$$

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{L}_2(t)[\boldsymbol{y}(t)+\boldsymbol{\psi}(t)\hat{\boldsymbol{\alpha}}(t-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t)]^{\mathrm{T}}, \tag{87}$$

$$\boldsymbol{L}_2(t)=\frac{\boldsymbol{P}_2(t-1)\boldsymbol{\varphi}(t)}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{P}_2(t-1)\boldsymbol{\varphi}(t)}, \tag{88}$$

$$\boldsymbol{P}_2(t)=[\boldsymbol{I}_{nr}-\boldsymbol{L}_2(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)]\boldsymbol{P}_2(t-1), \tag{89}$$

$$\boldsymbol{\varphi}(t)=[\boldsymbol{u}^{\mathrm{T}}(t-1),\boldsymbol{u}^{\mathrm{T}}(t-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(t-n)]^{\mathrm{T}}, \tag{90}$$

$$\boldsymbol{\psi}(t)=[\boldsymbol{y}(t-1),\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(t-n)]. \tag{91}$$

3.3 递阶梯度迭代辨识方法

在相同数据长度下,递阶梯度迭代辨识算法比 HSG 算法能产生更高的参数估计精度,下面给出递阶梯度迭代辨识方法.

设 L 为数据长度($L\gg n+nr$),定义准则函数

$$J_3(\boldsymbol{\alpha},\boldsymbol{\theta}):=\sum_{t=1}^{L}\|\boldsymbol{y}(t)+\boldsymbol{\psi}(t)\boldsymbol{\alpha}-\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}(t)\|^2,$$

令 $k=1,2,\cdots$为迭代变量,$\hat{\boldsymbol{\alpha}}_k$ 和 $\hat{\boldsymbol{\theta}}_k$ 分别为参数向量 $\boldsymbol{\alpha}$ 和参数矩阵 $\boldsymbol{\theta}$ 在第 k 次迭代时的参数估计,$\mu_k\geqslant 0$ 为收敛因子(convergence factor).使用负梯度搜索,极小化准则函数 $J_3(\boldsymbol{\alpha},\boldsymbol{\theta})$,可以得到估计参数向量 $\boldsymbol{\alpha}$ 和参数矩阵 $\boldsymbol{\theta}$ 的递阶梯度迭代算法(Hierarchical Gradient based Iterative algorithm, HGI 算法)[2,19]:

$$\hat{\boldsymbol{\alpha}}_k=\hat{\boldsymbol{\alpha}}_{k-1}-\mu_k\sum_{t=1}^{L}\boldsymbol{\psi}^{\mathrm{T}}(t)[\boldsymbol{y}(t)+\boldsymbol{\psi}(t)\hat{\boldsymbol{\alpha}}_{k-1}-$$

$$\hat{\boldsymbol{\theta}}_{k-1}^{\mathrm{T}}\boldsymbol{\varphi}(t)], \quad k=1,2,3,\cdots, \tag{92}$$

$$\hat{\boldsymbol{\theta}}_k=\hat{\boldsymbol{\theta}}_{k-1}+\mu_k\sum_{t=1}^{L}\boldsymbol{\varphi}(t)[\boldsymbol{y}(t)+\boldsymbol{\psi}(t)\hat{\boldsymbol{\alpha}}_{k-1}-\hat{\boldsymbol{\theta}}_{k-1}^{\mathrm{T}}\boldsymbol{\varphi}(t)]^{\mathrm{T}}, \tag{93}$$

$$\boldsymbol{\varphi}(t)=[\boldsymbol{u}^{\mathrm{T}}(t-1),\boldsymbol{u}^{\mathrm{T}}(t-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(t-n)]^{\mathrm{T}}, \tag{94}$$

$$\boldsymbol{\psi}(t)=[\boldsymbol{y}(t-1),\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(t-n)], \tag{95}$$

$$\mu_k\leqslant 2\Big(\sum_{t=1}^{L}[\|\boldsymbol{\psi}(t)\|^2+\|\boldsymbol{\varphi}(t)\|^2]\Big)^{-1}. \tag{96}$$

3.4 递阶最小二乘迭代辨识方法

使用最小二乘搜索原理,极小化准则函数 $J_3(\boldsymbol{\alpha},\boldsymbol{\theta})$,取收敛因子 $\mu_k=1$,可以得到递阶最小二乘迭代算法(Hierarchical Least Squares based Iterative algorithm, HLSI 算法)[2,20]:

$$\hat{\boldsymbol{\alpha}}_k=-\Big[\sum_{t=1}^{L}\boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\psi}(t)\Big]^{-1}\sum_{t=1}^{L}\boldsymbol{\psi}^{\mathrm{T}}(t)[\boldsymbol{y}(t)-\hat{\boldsymbol{\theta}}_{k-1}^{\mathrm{T}}\boldsymbol{\varphi}(t)], \tag{97}$$

$$\hat{\boldsymbol{\theta}}_k=\Big[\sum_{t=1}^{L}\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)\Big]^{-1}\sum_{t=1}^{L}\boldsymbol{\varphi}(t)[\boldsymbol{y}(t)+\boldsymbol{\psi}(t)\hat{\boldsymbol{\alpha}}_{k-1}], \tag{98}$$

$$\boldsymbol{\varphi}(t)=[\boldsymbol{u}^{\mathrm{T}}(t-1),\boldsymbol{u}^{\mathrm{T}}(t-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(t-n)]^{\mathrm{T}}, \tag{99}$$

$$\boldsymbol{\psi}(t)=[\boldsymbol{y}(t-1),\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(t-n)]. \tag{100}$$

同样,在相同数据长度下,由于充分利用了所采集的数据,递阶最小二乘迭代辨识方法比 HLS 算法有更好的参数辨识精度.

4 类多变量方程误差系统的递阶多新息辨识方法

考虑类多变量方程误差系统(1)的递阶辨识模型(4),重写如下:

$$\boldsymbol{y}(t)+\boldsymbol{\psi}(t)\boldsymbol{\alpha}=\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}(t)+\boldsymbol{v}(t), \tag{101}$$

其中各变量的定义同上.这里基于递阶辨识原理和多新息辨识理论[1],利用系统的观测数据 $\{\boldsymbol{u}(t),\boldsymbol{y}(t):t=1,2,3,\cdots\}$,推导递阶多新息随机梯度辨识方法和递阶多新息最小二乘辨识方法,估计系统参数向量 $\boldsymbol{\alpha}$ 和参数矩阵 $\boldsymbol{\theta}$.

4.1 递阶多新息随机梯度辨识方法

递阶随机梯度算法的收敛速度慢,为了提高参数估计的收敛速度,下面讨论递阶多新息随机梯度辨识方法.

设 $p\geqslant 1$ 为新息长度(innovation length).定义堆积输出向量 $\boldsymbol{Y}(p,t)$ 和 $\boldsymbol{Y}_1(p,t)$,堆积输出矩阵 $\boldsymbol{Z}(p,t)$ 和 $\boldsymbol{Y}_2(p,t)$,堆积信息向量 $\boldsymbol{\Phi}_1(p,t)$,堆积信息矩阵 $\boldsymbol{\Psi}_1(p,t)$,$\boldsymbol{\Phi}_2(p,t)$ 和 $\boldsymbol{\Psi}_2(p,t)$ 分别为

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

$$\boldsymbol{Y}_1(p,t):=\begin{bmatrix}\boldsymbol{y}(t)\\ \boldsymbol{y}(t-1)\\ \vdots\\ \boldsymbol{y}(t-p+1)\end{bmatrix}\in\mathbf{R}^{mp},$$

$$\boldsymbol{Z}(p,t):=[\boldsymbol{Z}(t),\boldsymbol{Z}(t-1),\cdots,\boldsymbol{Z}(t-p+1)]\in\mathbf{R}^{m\times p},$$

$$\boldsymbol{Y}_2(p,t):=[\boldsymbol{y}(t),\boldsymbol{y}(t-1),\cdots,\boldsymbol{y}(t-p+1)]\in\mathbf{R}^{m\times p},$$

$$\boldsymbol{\Phi}_1(p,t):=\begin{bmatrix}\boldsymbol{\varphi}(t)\\ \boldsymbol{\varphi}(t-1)\\ \vdots\\ \boldsymbol{\varphi}(t-p+1)\end{bmatrix}\in\mathbf{R}^{nrp},$$

$$\boldsymbol{\Psi}_1(p,t):=\begin{bmatrix}\boldsymbol{\psi}(t)\\ \boldsymbol{\psi}(t-1)\\ \vdots\\ \boldsymbol{\psi}(t-p+1)\end{bmatrix}\in\mathbf{R}^{mp\times n},$$

$$\boldsymbol{\Phi}_2(p,t):=[\boldsymbol{\varphi}(t),\boldsymbol{\varphi}(t-1),\cdots,\boldsymbol{\varphi}(t-p+1)]\in\mathbf{R}^{nr\times p},$$

$$\boldsymbol{\Psi}_2(p,t):=[\boldsymbol{\psi}(t),\boldsymbol{\psi}(t-1),\cdots,\boldsymbol{\psi}(t-p+1)]^{\mathrm{T}}\in\mathbf{R}^{(np)\times m}.$$

式(78)和(79)的新息向量分别为

$$\boldsymbol{e}_1(t):=\boldsymbol{y}(t)+\boldsymbol{\psi}(t)\hat{\boldsymbol{\alpha}}(t-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t)\in\mathbf{R}^m,$$

$$\boldsymbol{e}_2(t):=\boldsymbol{y}(t)+\boldsymbol{\psi}(t)\hat{\boldsymbol{\alpha}}(t-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t)\in\mathbf{R}^m.$$

根据多新息辨识理论,将向量新息 $\boldsymbol{e}_1(t)\in\mathbf{R}^m$ 扩展为一个大新息向量:

$$\boldsymbol{E}_1(p,t):=\begin{bmatrix}\boldsymbol{y}(t)+\boldsymbol{\psi}(t)\hat{\boldsymbol{\alpha}}(t-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t-1)\\ \boldsymbol{y}(t-1)+\boldsymbol{\psi}(t-1)\hat{\boldsymbol{\alpha}}(t-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t-2)\\ \vdots\\ \boldsymbol{y}(t-p+1)+\boldsymbol{\psi}(t-p+1)\hat{\boldsymbol{\alpha}}(t-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t-p+1)\end{bmatrix}=$$

$$\boldsymbol{Y}_1(p,t)+\boldsymbol{\Psi}_1(p,t)\hat{\boldsymbol{\alpha}}(t-1)-(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1))\boldsymbol{\Phi}_1(p,t)\in\mathbf{R}^{mp}.$$

将向量新息 $\boldsymbol{e}_2(t)\in\mathbf{R}^m$ 扩展为新息矩阵:

$$\begin{aligned}\boldsymbol{E}_2(p,t):=[&\boldsymbol{y}(t)+\boldsymbol{\psi}(t-1)\hat{\boldsymbol{\alpha}}(t-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t),\\ &\boldsymbol{y}(t-1)+\boldsymbol{\psi}(t-2)\hat{\boldsymbol{\alpha}}(t-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t-1),\cdots,\\ &\boldsymbol{y}(t-p+1)+\boldsymbol{\psi}(t-p+1)\hat{\boldsymbol{\alpha}}(t-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t-p+1)]=\\ &\boldsymbol{Y}_2(p,t)+\boldsymbol{\Psi}_2^{\mathrm{T}}(p,t)(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\alpha}}(t-1))-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\Phi}_2(p,t)\in\mathbf{R}^{m\times p}.\end{aligned}$$

根据多新息辨识理论[1,36,42],可以得到辨识系统(101)的递阶多新息随机梯度算法(Hierarchical Multi-Innovation Stochastic Gradient algorithm, HMISG 算法):

$$\hat{\boldsymbol{\alpha}}(t)=\hat{\boldsymbol{\alpha}}(t-1)-\frac{\boldsymbol{\Psi}_1^{\mathrm{T}}(p,t)}{r(t)}[\boldsymbol{Y}_1(p,t)+\boldsymbol{\Psi}_1(p,t)\hat{\boldsymbol{\alpha}}(t-1)-$$

$$(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1))\boldsymbol{\Phi}_1(p,t)], \tag{104}$$

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\boldsymbol{\Phi}_2(p,t)}{r(t)}[\boldsymbol{Y}_2(p,t)+\boldsymbol{\Psi}_2^{\mathrm{T}}(p,t)(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\alpha}}(t-1))-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\Phi}_2(p,t)]^{\mathrm{T}}, \tag{105}$$

$$r(t)=r(t-1)+\|\boldsymbol{\Psi}_1(p,t)\|^2+\|\boldsymbol{\Phi}_1(p,t)\|^2,\quad r(0)=1, \tag{106}$$

$$\boldsymbol{Y}_1(p,t)=\begin{bmatrix}\boldsymbol{y}(t)\\ \boldsymbol{y}(t-1)\\ \vdots\\ \boldsymbol{y}(t-p+1)\end{bmatrix}, \tag{107}$$

$$\boldsymbol{\Psi}_1(p,t)=\begin{bmatrix}\boldsymbol{\psi}(t)\\ \boldsymbol{\psi}(t-1)\\ \vdots\\ \boldsymbol{\psi}(t-p+1)\end{bmatrix}, \tag{108}$$

$$\boldsymbol{\Phi}_1(p,t)=\begin{bmatrix}\boldsymbol{\varphi}(t)\\ \boldsymbol{\varphi}(t-1)\\ \vdots\\ \boldsymbol{\varphi}(t-p+1)\end{bmatrix}, \tag{109}$$

$$\boldsymbol{Y}_2(p,t)=[\boldsymbol{y}(t),\boldsymbol{y}(t-1),\cdots,\boldsymbol{y}(t-p+1)], \tag{110}$$

$$\boldsymbol{\Psi}_2(p,t)=[\boldsymbol{\psi}(t),\boldsymbol{\psi}(t-1),\cdots,\boldsymbol{\psi}(t-p+1)]^{\mathrm{T}}, \tag{111}$$

$$\boldsymbol{\Phi}_2(p,t)=[\boldsymbol{\varphi}(t),\boldsymbol{\varphi}(t-1),\cdots,\boldsymbol{\varphi}(t-p+1)]. \tag{112}$$

当新息长度 $p=1$ 时,递阶多新息随机梯度算法退化为递阶随机梯度算法.式(106)的等价表达为

$$r(t)=r(t-1)+\|\boldsymbol{\Psi}_1(p,t)\|^2+\|\boldsymbol{\Phi}_2(p,t)\|^2=r(t-1)+\|\boldsymbol{\Psi}_2(p,t)\|^2+\|\boldsymbol{\Phi}_1(p,t)\|^2=r(t-1)+\|\boldsymbol{\Psi}_2(p,t)\|^2+\|\boldsymbol{\Phi}_2(p,t)\|^2,\quad r(0)=1 \tag{113}$$

式(106)也可以修改为

$$r(t)=r(t-1)+\|\boldsymbol{\psi}_2(t)\|^2+\|\boldsymbol{\varphi}_2(t)\|^2,\quad r(0)=1, \tag{114}$$

或

$$r(t)=r(t-1)+\max[\|\boldsymbol{\psi}_2(t)\|^2,\|\boldsymbol{\varphi}_2(t)\|^2],\quad r(0)=1. \tag{115}$$

表1列出了 HMISG 算法(104)—(112)中各变量的维数,表2列出了 HMISG 算法的计算量.

表1 HMISG 算法各变量的维数

Table 1 The dimensions of the variables in the HMISG algorithm

变量名称	维数
参数估计向量	$\hat{\boldsymbol{\alpha}}(t)\in\mathbf{R}^n$
参数估计矩阵	$\hat{\boldsymbol{\theta}}(t)\in\mathbf{R}^{(nr)\times m}$
输入向量、输出向量	$\boldsymbol{u}(t)\in\mathbf{R}^r$, $\boldsymbol{y}(t)\in\mathbf{R}^m$
输入信息向量	$\boldsymbol{\varphi}(t)\in\mathbf{R}^{nr}$
输出信息矩阵	$\boldsymbol{\psi}(t)\in\mathbf{R}^{m\times n}$
堆积输出向量	$\boldsymbol{Y}_1(p,t)\in\mathbf{R}^{mp}$
堆积输出矩阵	$\boldsymbol{Y}_2(p,t)\in\mathbf{R}^{m\times p}$
输入堆积信息向量	$\boldsymbol{\Phi}_1(p,t)\in\mathbf{R}^{nrp}$
输入堆积信息矩阵	$\boldsymbol{\Phi}_2(p,t)\in\mathbf{R}^{(nr)\times p}$
输出堆积信息矩阵	$\boldsymbol{\Psi}_1(p,t)\in\mathbf{R}^{(mp)\times n}$, $\boldsymbol{\Psi}_2(p,t)\in\mathbf{R}^{(np)\times m}$

4.2 递阶多新息最小二乘辨识方法

下面推导递阶多新息最小二乘辨识算法,实现对系统的未知参数向量 $\boldsymbol{\alpha}$ 和参数矩阵 $\boldsymbol{\theta}$ 的实时估计.考虑 $t-p+1$ 到 t 时共 p 组数据,$\boldsymbol{Y}(p,t)$,$\boldsymbol{Z}(p,t)$,$\boldsymbol{Y}_1(p,t)$,$\boldsymbol{Y}_2(p,t)$,$\boldsymbol{\Phi}_1(p,t)$,$\boldsymbol{\Phi}_2(p,t)$,$\boldsymbol{\Psi}_1(p,t)$和 $\boldsymbol{\Psi}_2(p,t)$定义同上.根据递阶辨识原理与多新息辨识理论[1-2],可以得到估计参数向量 $\boldsymbol{\alpha}$ 和参数矩阵 $\boldsymbol{\theta}$ 的递阶多新息最小二乘辨识算法:

$$\hat{\boldsymbol{\alpha}}(t)=\hat{\boldsymbol{\alpha}}(t-1)-\boldsymbol{P}_1(t)\boldsymbol{\Psi}_1^{\mathrm{T}}(p,t)[\boldsymbol{Y}_1(p,t)+\boldsymbol{\Psi}_1^{\mathrm{T}}(p,t)\boldsymbol{\alpha}(t-1)-(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1))\boldsymbol{\Phi}_1(p,t)], \tag{116}$$

$$\boldsymbol{P}_1^{-1}(t):=\sum_{j=1}^{t}\boldsymbol{\Psi}_1^{\mathrm{T}}(p,j)\boldsymbol{\Psi}_1(p,j)=\boldsymbol{P}_1^{-1}(t-1)+\boldsymbol{\Psi}_1^{\mathrm{T}}(p,t)\boldsymbol{\Psi}_1(p,t), \tag{117}$$

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{P}_2(t)\boldsymbol{\Phi}_2(p,t)[\boldsymbol{Y}_2(p,t)+\boldsymbol{\Psi}_2^{\mathrm{T}}(p,t)(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\alpha}}(t-1))-\boldsymbol{\theta}^{\mathrm{T}}(t-1)\boldsymbol{\Phi}_2(p,t)]^{\mathrm{T}}, \tag{118}$$

表2 HMISG 算法的计算量

Table 2 The computational efficiency of the HMISG algorithm

表达式	乘法次数	加法次数
$\hat{\boldsymbol{\alpha}}(t)=\hat{\boldsymbol{\alpha}}(t-1)-\frac{\boldsymbol{\Psi}_1^{\mathrm{T}}(p,t)}{r(t)}\boldsymbol{E}_1(p,t)\in\mathbf{R}^n$	$mpn+n$	mpn
$\boldsymbol{E}_1(p,t)=[\boldsymbol{Y}_1(p,t)+\boldsymbol{\Psi}_1(p,t)\hat{\boldsymbol{\alpha}}(t-1)-(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1))\boldsymbol{\Phi}_1(p,t)]\in\mathbf{R}^{mp}$	$mp(n+nr)$	$mp(n+nr)$
$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\boldsymbol{\Phi}_2(p,t)}{r(t)}\boldsymbol{E}_2^{\mathrm{T}}(p,t)\in\mathbf{R}^{(nr)\times m}$	$mpnr+nrm$	$mpnr$
$\boldsymbol{E}_2(p,t)=[\boldsymbol{Y}_2(p,t)+\boldsymbol{\Psi}_2^{\mathrm{T}}(p,t)(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\alpha}}(t-1))-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\Phi}_2(p,t)]\in\mathbf{R}^{m\times p}$	$mp(n+nr)$	$mp(n+nr)$
$r(t)=r(t-1)+\|\boldsymbol{\Psi}_1(p,t)\|^2+\|\boldsymbol{\Phi}_1(p,t)\|^2\in\mathbf{R}$	$mpn+nrp$	$mpn+nrp$
总数	$mp(4n+3nr)+nrp+nrm+n$	$mp(4n+3nr)+nrp$
总 flop 数	$2mp(4n+3nr)+2nrp+n+nrm$	

$$\boldsymbol{P}_2^{-1}(t):=\sum_{j=1}^{t}\boldsymbol{\Phi}_2(p,j)\boldsymbol{\Phi}_2^{\mathrm{T}}(p,j)=\boldsymbol{P}_2^{-1}(t-1)+\boldsymbol{\Phi}_2(p,t)\boldsymbol{\Phi}_2^{\mathrm{T}}(p,t). \tag{119}$$

$\boldsymbol{P}_1(t)\in\mathbf{R}^{n\times n}$ 和 $\boldsymbol{P}_2(t)\in\mathbf{R}^{(nr)\times(nr)}$ 为 2 个协方差矩阵. 应用矩阵求逆引理[1-2]

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

于式(117)和(119)得到

$$\boldsymbol{P}_1(t)=\boldsymbol{P}_1(t-1)-\boldsymbol{P}_1(t-1)\boldsymbol{\Psi}_1^{\mathrm{T}}(p,t)[\boldsymbol{I}_{mp}+\boldsymbol{\Psi}_1(p,t)\boldsymbol{P}_1(t-1)\boldsymbol{\Psi}_1^{\mathrm{T}}(p,t)]^{-1}\boldsymbol{\Psi}_1(p,t)\boldsymbol{P}_1(t-1), \tag{120}$$

$$\boldsymbol{P}_2(t)=\boldsymbol{P}_2(t-1)-\boldsymbol{P}_2(t-1)\boldsymbol{\Phi}_2(p,t)[\boldsymbol{I}_p+\boldsymbol{\Phi}_2^{\mathrm{T}}(p,t)\boldsymbol{P}_2(t-1)\boldsymbol{\Phi}_2(p,t)]^{-1}\boldsymbol{\Phi}_2^{\mathrm{T}}(p,t)\boldsymbol{P}_2(t-1). \tag{121}$$

并引入增益矩阵 $\boldsymbol{L}_1(t):=-\boldsymbol{P}_1(t)\boldsymbol{\Psi}_1^{\mathrm{T}}(p,t)\in\mathbf{R}^{n\times(mp)}$ 和增益向量 $\boldsymbol{L}_2(t):=\boldsymbol{P}_2(t)\boldsymbol{\Phi}_2(p,t)\in\mathbf{R}^{nr\times p}$. 式(120)两边右乘 $\boldsymbol{\Psi}_1^{\mathrm{T}}(p,t)$, 式(121)两边右乘 $\boldsymbol{\Phi}_2(p,t)$ 得到

$$\boldsymbol{L}_1(t)=-\boldsymbol{P}_1(t-1)\boldsymbol{\Psi}_1^{\mathrm{T}}(p,t)[\boldsymbol{I}_{mp}+\boldsymbol{\Psi}_1(p,t)\boldsymbol{P}_1(t-1)\boldsymbol{\Psi}_1^{\mathrm{T}}(p,t)]^{-1},$$

$$\boldsymbol{L}_2(t)=\boldsymbol{P}_2(t-1)\boldsymbol{\Phi}_2(p,t)[\boldsymbol{I}_p+\boldsymbol{\Phi}_2^{\mathrm{T}}(p,t)\boldsymbol{P}_2(t-1)\boldsymbol{\Phi}_2(p,t)]^{-1}.$$

综合以上各式,容易得到估计参数向量 $\boldsymbol{\alpha}$ 和参数矩阵 $\boldsymbol{\theta}$ 的递阶多新息最小二乘算法(Hierarchical Multi-Innovation Least Squares algorithm, HMILS 算法)[1,37]:

$$\hat{\boldsymbol{\alpha}}(t)=\hat{\boldsymbol{\alpha}}(t-1)+\boldsymbol{L}_1(t)[\boldsymbol{Y}_1(p,t)+\boldsymbol{\Psi}_1(p,t)\hat{\boldsymbol{\alpha}}(t-1)-(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1))\boldsymbol{\Phi}_1(p,t)], \tag{122}$$

$$\boldsymbol{L}_1(t)=-\boldsymbol{P}_1(t-1)\boldsymbol{\Psi}_1^{\mathrm{T}}(p,t)[\boldsymbol{I}_{mp}+\boldsymbol{\Psi}_1(p,t)\boldsymbol{P}_1(t-1)\boldsymbol{\Psi}_1^{\mathrm{T}}(p,t)]^{-1}, \tag{123}$$

$$\boldsymbol{P}_1(t)=\boldsymbol{P}_1(t-1)-\boldsymbol{P}_1(t-1)\boldsymbol{\Psi}_1^{\mathrm{T}}(p,t)[\boldsymbol{I}_{mp}+\boldsymbol{\Psi}_1(p,t)\boldsymbol{P}_1(t-1)\boldsymbol{\Psi}_1^{\mathrm{T}}(p,t)]^{-1}\boldsymbol{\Psi}_1(p,t)\boldsymbol{P}_1(t-1)=[\boldsymbol{I}_n+\boldsymbol{L}_1(t)\boldsymbol{\Psi}_1(p,t)]\boldsymbol{P}_1(t-1), \tag{124}$$

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{L}_2(t)[\boldsymbol{Y}_2(p,t)+\boldsymbol{\Psi}_2^{\mathrm{T}}(p,t)(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\alpha}}(t-1))-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\Phi}_2(p,t)]^{\mathrm{T}}, \tag{125}$$

$$\boldsymbol{L}_2(t)=\boldsymbol{P}_2(t-1)\boldsymbol{\Phi}_2(p,t)[\boldsymbol{I}_p+\boldsymbol{\Phi}_2^{\mathrm{T}}(p,t)\boldsymbol{P}_2(t-1)\boldsymbol{\Phi}_2(p,t)]^{-1}, \tag{126}$$

$$\boldsymbol{P}_2(t)=\boldsymbol{P}_2(t-1)-\boldsymbol{P}_2(t-1)\boldsymbol{\Phi}_2(p,t)[\boldsymbol{I}_p+\boldsymbol{\Phi}_2^{\mathrm{T}}(p,t)\boldsymbol{P}_2(t-1)\boldsymbol{\Phi}_2(p,t)]^{-1}\boldsymbol{\Phi}_2^{\mathrm{T}}(p,t)\boldsymbol{P}_2(t-1)=[\boldsymbol{I}_{nr}-\boldsymbol{L}_2(t)\boldsymbol{\Phi}_2^{\mathrm{T}}(p,t)]\boldsymbol{P}_2(t-1), \tag{127}$$

$$\boldsymbol{Y}_1(p,t)=\begin{bmatrix}\boldsymbol{y}(t)\\ \boldsymbol{y}(t-1)\\ \vdots\\ \boldsymbol{y}(t-p+1)\end{bmatrix}, \tag{128}$$

$$\boldsymbol{\Psi}_1(p,t)=\begin{bmatrix}\boldsymbol{\psi}(t)\\ \boldsymbol{\psi}(t-1)\\ \vdots\\ \boldsymbol{\psi}(t-p+1)\end{bmatrix}, \tag{129}$$

$$\boldsymbol{\Phi}_1(p,t)=\begin{bmatrix}\boldsymbol{\varphi}(t)\\ \boldsymbol{\varphi}(t-1)\\ \vdots\\ \boldsymbol{\varphi}(t-p+1)\end{bmatrix}, \tag{130}$$

$$\boldsymbol{Y}_2(p,t)=[\boldsymbol{y}(t),\boldsymbol{y}(t-1),\cdots,\boldsymbol{y}(t-p+1)], \tag{131}$$

$$\boldsymbol{\Psi}_2(p,t)=[\boldsymbol{\psi}(t),\boldsymbol{\psi}(t-1),\cdots,\boldsymbol{\psi}(t-p+1)]^{\mathrm{T}}, \tag{132}$$

$$\boldsymbol{\Phi}_2(p,t)=[\boldsymbol{\varphi}(t),\boldsymbol{\varphi}(t-1),\cdots,\boldsymbol{\varphi}(t-p+1)]. \tag{133}$$

HMILS 算法(122)—(133)计算参数估计的步骤如下:

1) 初始化:令 $t=1$, $\boldsymbol{P}_1(0)=p_0\boldsymbol{I}_n$, $\boldsymbol{P}_2(0)=p_0\boldsymbol{I}_{nr}$, $\hat{\boldsymbol{\alpha}}(0)=\mathbf{1}_n/p_0$, $\hat{\boldsymbol{\theta}}(0)=\mathbf{1}_{nr\times m}/p_0$.

2) 采集输入输出数据 $\boldsymbol{u}(t)$ 和 $\boldsymbol{y}(t)$, 由式(128)和式(131)构造堆积输出向量 $\boldsymbol{Y}_1(p,t)$ 和堆积输出矩阵 $\boldsymbol{Y}_2(p,t)$, 由式(128)—(133)构造堆积信息向量 $\boldsymbol{\Phi}_1(p,t)$ 和堆积信息矩阵 $\boldsymbol{\Psi}_1(p,t)$, $\boldsymbol{\Phi}_2(p,t)$ 和 $\boldsymbol{\Psi}_2(p,t)$.

3) 用式(123)和式(124)计算 $\boldsymbol{L}_1(t)$ 和 $\boldsymbol{P}_1(t)$, 用式(126)和式(127)计算 $\boldsymbol{L}_2(t)$ 和 $\boldsymbol{P}_2(t)$.

4) 用式(122)和(125)刷新参数估计向量 $\hat{\boldsymbol{\alpha}}(t)$ 和参数估计矩阵 $\hat{\boldsymbol{\theta}}(t)$.

5) t 增 1,转到第 2)步.

表 3 列出了 HMILS 算法(122)—(133)中各变量的维数,表 4 列出了 HMILS 算法的计算量.

表 3　HMILS 算法中各变量维数

Table 3　The dimensions of the variables in the HMILS algorithm

变量名称	维数
参数估计向量	$\hat{\boldsymbol{\alpha}}(t)\in\mathbf{R}^{n}$
参数估计矩阵	$\hat{\boldsymbol{\theta}}(t)\in\mathbf{R}^{(nr)\times m}$
输入向量、输出向量	$\boldsymbol{u}(t)\in\mathbf{R}^{r}$,　$\boldsymbol{y}(t)\in\mathbf{R}^{m}$
输出信息矩阵	$\boldsymbol{\psi}(t)\in\mathbf{R}^{m\times n}$
输入信息向量	$\boldsymbol{\varphi}(t)\in\mathbf{R}^{nr}$
堆积输出向量	$\boldsymbol{Y}_1(p,t)\in\mathbf{R}^{mp}$
堆积输出矩阵	$\boldsymbol{Y}_2(p,t)\in\mathbf{R}^{m\times p}$
输入堆积信息向量	$\boldsymbol{\Phi}_1(p,t)\in\mathbf{R}^{nrp}$
输入堆积信息矩阵	$\boldsymbol{\Phi}_2(p,t)\in\mathbf{R}^{(nr)\times p}$
输出堆积信息矩阵	$\boldsymbol{\Psi}_1(p,t)\in\mathbf{R}^{(mp)\times n}$,　$\boldsymbol{\Psi}_2(p,t)\in\mathbf{R}^{(np)\times m}$
增益矩阵	$\boldsymbol{L}_1(t)\in\mathbf{R}^{n\times(mp)}$,　$\boldsymbol{L}_2(t)\in\mathbf{R}^{(nr)\times p}$
输出数据协方差矩阵	$\boldsymbol{P}_1(t)\in\mathbf{R}^{n\times n}$
输入数据协方差矩阵	$\boldsymbol{P}_2(t)\in\mathbf{R}^{(nr)\times(nr)}$

5 类多变量方程误差 ARMA 系统的递阶辨识方法

考虑类多变量方程误差 ARMA 系统(multivariable equation error ARMA-like system):

$$\alpha(z)\boldsymbol{y}(t)=\boldsymbol{Q}(z)\boldsymbol{u}(t)+\frac{D(z)}{C(z)}\boldsymbol{v}(t), \tag{134}$$

其中 $\boldsymbol{y}(t):=[y_1(t),y_2(t),\cdots,y_m(t)]^{\mathrm{T}}\in\mathbf{R}^m$ 为 m 维观测输出向量,$\boldsymbol{u}(t)\in\mathbf{R}^r$ 为 r 维观测输入向量,$\boldsymbol{v}(t):=[v_1(t),v_2(t),\cdots,v_m(t)]^{\mathrm{T}}\in\mathbf{R}^m$ 为 m 维白噪声向量,$\alpha(z)$,$C(z)$和 $D(z)$为单位后移算子 z^{-1}的多项式,$\boldsymbol{Q}(z)$为单位后移算子 z^{-1}的多项式矩阵,它们定义为

$\alpha(z):=1+\alpha_1z^{-1}+\alpha_2z^{-2}+\cdots+\alpha_nz^{-n},\quad \alpha_i\in\mathbf{R}$,

$\boldsymbol{Q}(z):=\boldsymbol{Q}_1z^{-1}+\boldsymbol{Q}_2z^{-2}+\cdots+\boldsymbol{Q}_nz^{-n},\quad \boldsymbol{Q}_i\in\mathbf{R}^{m\times r}$,

$C(z):=1+c_1z^{-1}+c_2z^{-2}+\cdots+c_{n_c}z^{-n_c},\quad c_i\in\mathbf{R}$,

$D(z):=1+d_1z^{-1}+d_2z^{-2}+\cdots+d_{n_d}z^{-n_d},\quad d_i\in\mathbf{R}$.

定义中间变量 $\boldsymbol{w}(t):=\frac{D(z)}{C(z)}\boldsymbol{v}(t)$,代入式(134)可得

$$\alpha(z)\boldsymbol{y}(t)=\boldsymbol{Q}(z)\boldsymbol{u}(t)+\boldsymbol{w}(t). \tag{135}$$

由此可得

$$\boldsymbol{w}(t)=-\sum_{i=1}^{n_c}c_i\boldsymbol{w}(t-i)+\boldsymbol{v}(t)+\sum_{i=1}^{n_d}d_i\boldsymbol{v}(t-i). \tag{136}$$

将式(136)代入式(135)可得

$$\boldsymbol{y}(t)+\sum_{i=1}^{n}\alpha_i\boldsymbol{y}(t-i)+\sum_{i=1}^{n_c}c_i\boldsymbol{w}(t-i)-\sum_{i=1}^{n_d}d_i\boldsymbol{v}(t-i)=\sum_{i=1}^{n}\boldsymbol{Q}_i\boldsymbol{u}(t-i)+\boldsymbol{v}(t). \tag{137}$$

定义模型参数向量 $\boldsymbol{\vartheta}$,参数矩阵 $\boldsymbol{\theta}$,输入信息向量 $\boldsymbol{\varphi}(t)$和输出信息矩阵 $\boldsymbol{\psi}(t)$如下:

$\boldsymbol{\vartheta}:=\begin{bmatrix}\boldsymbol{\alpha}\\ \boldsymbol{\vartheta}_{\mathrm{n}}\end{bmatrix}\in\mathbf{R}^{n+n_c+n_d},\ n_0:=n+n_c+n_d$,

$\boldsymbol{\alpha}:=[\alpha_1,\alpha_2,\cdots,\alpha_n]^{\mathrm{T}}\in\mathbf{R}^n$,

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

$\boldsymbol{\theta}^{\mathrm{T}}:=[\boldsymbol{Q}_1,\boldsymbol{Q}_2,\cdots,\boldsymbol{Q}_n]\in\mathbf{R}^{m\times(nr)}$,

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

$\boldsymbol{\psi}(t):=[\boldsymbol{\psi}_{\mathrm{s}}(t),\boldsymbol{\psi}_{\mathrm{n}}(t)]\in\mathbf{R}^{m\times n_0}$,

$\boldsymbol{\psi}_{\mathrm{s}}(t):=[\boldsymbol{y}(t-1),\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(t-n)]\in\mathbf{R}^{m\times n}$,

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

于是,由式(138)可以得到

$$\boldsymbol{y}(t)+\boldsymbol{\psi}(t)\boldsymbol{\vartheta}=\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}(t)+\boldsymbol{v}(t). \tag{138}$$

式(138)即为类多变量方程误差 ARMA 系统(134)的递阶辨识模型(hierarchical identification model).与类多变量方程误差系统(1)的递阶辨识模

表 4 递阶多新息最小二乘算法

Table 4 The computational efficiency of the HMILS algorithm

表达式	乘法次数	加法次数
$\hat{\boldsymbol{\alpha}}(t)=\hat{\boldsymbol{\alpha}}(t-1)+\boldsymbol{L}_1(t)\boldsymbol{E}_1(t)\in\mathbf{R}^n$	mpn	mpn
$\boldsymbol{E}_1(p,t):=[\boldsymbol{Y}_1(p,t)+\boldsymbol{\Psi}_1(p,t)\hat{\boldsymbol{\alpha}}(t-1)-(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1))\boldsymbol{\Phi}_1(p,t)]\in\mathbf{R}^{mp}$	$mp(n+nr)$	$mp(n+nr)$
$\boldsymbol{L}_1(t):=-\boldsymbol{R}_1(t)\boldsymbol{\Lambda}'_1(t)\in\mathbf{R}^{n\times(mp)}$	$(mp)^2n$	$(mp)^2n-mpn$
$\boldsymbol{R}_1(t):=\boldsymbol{P}_1(t-1)\boldsymbol{\Psi}_1^{\mathrm{T}}(t)\in\mathbf{R}^{n\times(mp)}$	mpn^2	mpn^2-mpn
$\boldsymbol{\Lambda}_1(t):=\boldsymbol{I}_{mp}+\boldsymbol{\Psi}_1(t)\boldsymbol{R}_1(t)\in\mathbf{R}^{(mp)\times(mp)}$	$(mp)^2n$	$(mp)^2n$
$\boldsymbol{\Lambda}'_1(t):=\boldsymbol{\Lambda}_1^{-1}(t)\in\mathbf{R}^{(mp)\times(mp)}$	$(mp)^3$	$(mp)^3-(mp)^2$
$\boldsymbol{P}_1(t)=\boldsymbol{P}_1(t-1)+\boldsymbol{L}_1(t)\boldsymbol{R}_1^{\mathrm{T}}(t)\in\mathbf{R}^{n\times n}$	mpn^2	mpn^2
$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{L}_2(t)\boldsymbol{E}_2^{\mathrm{T}}(t)\in\mathbf{R}^{(nr)\times m}$	$mpnr$	$mpnr$
$\boldsymbol{E}_2(p,t):=[\boldsymbol{Y}_2(p,t)+\boldsymbol{\Psi}_2^{\mathrm{T}}(p,t)(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\alpha}}(t-1))-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\Phi}_2(p,t)]\in\mathbf{R}^{m\times p}$	$mp(n+nr)$	$mp(n+nr)$
$\boldsymbol{L}_2(t):=\boldsymbol{R}_2(t)\boldsymbol{\Lambda}'_2(t)\in\mathbf{R}^{(nr)\times p}$	p^2nr	$p^2nr-pnr$
$\boldsymbol{R}_2(t):=\boldsymbol{P}_2(t-1)\boldsymbol{\Phi}_2(t)\in\mathbf{R}^{(nr)\times p}$	$p(nr)^2$	$p(nr)^2-pnr$
$\boldsymbol{\Lambda}_2(t):=\boldsymbol{I}_p+\boldsymbol{\Phi}_2^{\mathrm{T}}(t)\boldsymbol{R}_2(t)\in\mathbf{R}^{p\times p}$	p^2nr	p^2nr
$\boldsymbol{\Lambda}'_2(t):=\boldsymbol{\Lambda}_2^{-1}(t)\in\mathbf{R}^{p\times p}$	p^3	p^3-p^2
$\boldsymbol{P}_2(t)=\boldsymbol{P}_2(t-1)-\boldsymbol{L}_2(t)\boldsymbol{R}_2^{\mathrm{T}}(t)\in\mathbf{R}^{(nr)\times nr}$	$p(nr)^2$	$p(nr)^2$
总数	$p^3(m^3+1)+p^2(2m^2n+2nr)+p(3mn+3mnr+2mn^2+2n^2r^2)$	$p^3(m^3+1)+p^2(2m^2n+2nr-m^2-1)+p(mn+3mnr+2mn^2+2n^2r^2-2nr)$
总 flop 数	$2p^3(m^3+1)+2p^2(2m^2n+2nr)-p^2(m^2+1)+2p(2mn+3mnr+2mn^2+2n^2r^2)-2pnr$	

型(4)相比,不同的是这里的信息矩阵 $\boldsymbol{\psi}(t)$ 包含了未知噪声项 $\boldsymbol{w}(t-i)$ 和 $\boldsymbol{v}(t-i)$.

下面基于递阶辨识原理利用系统的观测数据 $\{\boldsymbol{u}(t),\boldsymbol{y}(t):t=1,2,3,\cdots\}$,给出递阶广义增广随机梯度辨识方法和递阶广义增广最小二乘辨识方法,估计系统参数向量 $\boldsymbol{\vartheta}$ 和参数矩阵 $\boldsymbol{\theta}$.

5.1 递阶广义增广随机梯度辨识方法

定义准则函数

$$J_4(\boldsymbol{\vartheta},\boldsymbol{\theta}):=\|\boldsymbol{y}(t)+\boldsymbol{\psi}(t)\boldsymbol{\vartheta}-\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}(t)\|^2,$$

令 $\hat{\boldsymbol{\vartheta}}(t):=\begin{bmatrix}\hat{\boldsymbol{\alpha}}(t)\\ \hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(t)\end{bmatrix}$ 为 $\boldsymbol{\vartheta}$ 在时刻 t 的估计,$\hat{\boldsymbol{\theta}}(t)$ 为 $\boldsymbol{\theta}$ 在时刻 t 的估计.使用负梯度搜索,极小化准则函数 $J_4(\boldsymbol{\vartheta},\boldsymbol{\theta})$,可以得到类多变量方程误差类 ARMA 系统(134)的递阶广义增广随机梯度辨识方法(Hierarchical Generalized Extended Stochastic Gradient algorithm,HGESG 算法):

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)-\frac{\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)}{r(t)}[\boldsymbol{y}(t)+\hat{\boldsymbol{\psi}}(t)\hat{\boldsymbol{\vartheta}}(t-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t)],\tag{139}$$

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\boldsymbol{\varphi}(t)}{r(t)}[\boldsymbol{y}(t)+\hat{\boldsymbol{\psi}}(t)\hat{\boldsymbol{\vartheta}}(t-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t)]^{\mathrm{T}},\tag{140}$$

$$r(t)=r(t-1)+\|\hat{\boldsymbol{\psi}}(t)\|^2+\|\boldsymbol{\varphi}(t)\|^2,r(0)=1,\tag{141}$$

$$\boldsymbol{\varphi}(t)=[\boldsymbol{u}^{\mathrm{T}}(t-1),\boldsymbol{u}^{\mathrm{T}}(t-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(t-n)]^{\mathrm{T}},\tag{142}$$

$$\hat{\boldsymbol{\psi}}(t)=[\boldsymbol{\psi}_{\mathrm{s}}(t),\hat{\boldsymbol{\psi}}_{\mathrm{n}}(t)],\tag{143}$$

$$\boldsymbol{\psi}_{\mathrm{s}}(t)=[\boldsymbol{y}(t-1),\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(t-n)],\tag{144}$$

$$\hat{\boldsymbol{\psi}}_{\mathrm{n}}(t)=[\hat{\boldsymbol{w}}(t-1),\hat{\boldsymbol{w}}(t-2),\cdots,\hat{\boldsymbol{w}}(t-n_c),-\hat{\boldsymbol{v}}(t-1),-\hat{\boldsymbol{v}}(t-2),\cdots,-\hat{\boldsymbol{v}}(t-n_d)],\tag{145}$$

$$\hat{\boldsymbol{w}}(t)=\boldsymbol{y}(t)+\boldsymbol{\psi}_{\mathrm{s}}(t)\hat{\boldsymbol{\alpha}}(t)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t)\boldsymbol{\varphi}(t),\tag{146}$$

$$\hat{\boldsymbol{v}}(t)=\boldsymbol{y}(t)+\hat{\boldsymbol{\psi}}(t)\hat{\boldsymbol{\vartheta}}(t)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t)\boldsymbol{\varphi}(t),\tag{147}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\begin{bmatrix}\hat{\boldsymbol{\alpha}}(t)\\ \hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(t)\end{bmatrix},\tag{148}$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\alpha}_1(t),\hat{\alpha}_2(t),\cdots,\hat{\alpha}_n(t)]^{\mathrm{T}},\tag{149}$$

$$\hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(t)=[\hat{c}_1(t),\hat{c}_2(t),\cdots,\hat{c}_{n_c}(t),\hat{d}_1(t),\hat{d}_2(t),\cdots,\hat{d}_{n_d}(t)]^{\mathrm{T}}.\tag{150}$$

5.2 递阶广义增广最小二乘辨识方法

定义准则函数

$$J_5(\boldsymbol{\vartheta},\boldsymbol{\theta}):=\sum_{j=1}^{t}\|\boldsymbol{y}(j)+\boldsymbol{\psi}(j)\boldsymbol{\vartheta}-\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}(j)\|^2,$$

根据递阶辨识原理,令 $J_5(\boldsymbol{\vartheta},\boldsymbol{\theta})$ 对 $\boldsymbol{\vartheta}$ 和 $\boldsymbol{\theta}$ 的偏导数为零,可以得到递推最小二乘算法[2,20]:

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)-\boldsymbol{L}_1(t)[\boldsymbol{y}(t)+\boldsymbol{\psi}(t)\hat{\boldsymbol{\vartheta}}(t-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t)],\tag{151}$$

$$\boldsymbol{L}_1(t)=-\boldsymbol{P}_1(t-1)\boldsymbol{\psi}^{\mathrm{T}}(t)[\boldsymbol{I}_m+\boldsymbol{\psi}(t)\boldsymbol{P}_1(t-1)\boldsymbol{\psi}^{\mathrm{T}}(t)]^{-1},\tag{152}$$

$$\boldsymbol{P}_1(t)=[\boldsymbol{I}_{n_0}+\boldsymbol{L}_1(t)\boldsymbol{\psi}(t)]\boldsymbol{P}_1(t-1),\tag{153}$$

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{L}_2(t)[\boldsymbol{y}(t)+\boldsymbol{\psi}(t)\hat{\boldsymbol{\vartheta}}(t-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t)]^{\mathrm{T}},\tag{154}$$

$$\boldsymbol{L}_2(t)=\frac{\boldsymbol{P}_2(t-1)\boldsymbol{\varphi}(t)}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{P}_2(t-1)\boldsymbol{\varphi}(t)},\tag{155}$$

$$\boldsymbol{P}_2(t)=[\boldsymbol{I}_{nr}-\boldsymbol{L}_2(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)]\boldsymbol{P}_2(t-1).\tag{156}$$

与递阶广义增广梯度迭代辨识算法(139)—(150)类似,式(151)—(154)中 $\boldsymbol{\psi}(t)$ 包含的不可测噪声项 $\boldsymbol{w}(t-i)$ 和 $\boldsymbol{v}(t-i)$ 分别用其估计值 $\hat{\boldsymbol{w}}(t-i)$ 和 $\hat{\boldsymbol{v}}(t-i)$ 代替,便得到类多变量方程误差类 ARMA 系统(134)的递阶广义增广最小二乘算法(Hierarchical Generalized Extended Least Squares algorithm,HGELS 算法):

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\boldsymbol{L}_1(t)[\boldsymbol{y}(t)+\hat{\boldsymbol{\psi}}(t)\hat{\boldsymbol{\vartheta}}(t-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t)],\tag{157}$$

$$\boldsymbol{L}_1(t)=-\boldsymbol{P}_1(t-1)\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)[\boldsymbol{I}_m+\hat{\boldsymbol{\psi}}(t)\boldsymbol{P}_1(t-1)\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)]^{-1},\tag{158}$$

$$\boldsymbol{P}_1(t)=[\boldsymbol{I}_{n_0}+\boldsymbol{L}_1(t)\hat{\boldsymbol{\psi}}(t)]\boldsymbol{P}_1(t-1),\tag{159}$$

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{L}_2(t)[\boldsymbol{y}(t)+\hat{\boldsymbol{\psi}}(t)\hat{\boldsymbol{\vartheta}}(t-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t)]^{\mathrm{T}},\tag{160}$$

$$\boldsymbol{L}_2(t)=\frac{\boldsymbol{P}_2(t-1)\boldsymbol{\varphi}(t)}{1+\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{P}_2(t-1)\boldsymbol{\varphi}(t)},\tag{161}$$

$$\boldsymbol{P}_2(t)=[\boldsymbol{I}_{nr}-\boldsymbol{L}_2(t)\boldsymbol{\varphi}^{\mathrm{T}}(t)]\boldsymbol{P}_2(t-1),\tag{162}$$

$$\boldsymbol{\varphi}(t)=[\boldsymbol{u}^{\mathrm{T}}(t-1),\boldsymbol{u}^{\mathrm{T}}(t-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(t-n)]^{\mathrm{T}},\tag{163}$$

$$\boldsymbol{\psi}_{\mathrm{s}}(t)=[\boldsymbol{y}(t-1),\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(t-n)],\tag{164}$$

$$\hat{\boldsymbol{\psi}}_{\mathrm{n}}(t)=[\hat{\boldsymbol{w}}(t-1),\hat{\boldsymbol{w}}(t-2),\cdots,\hat{\boldsymbol{w}}(t-n_c),-\hat{\boldsymbol{v}}(t-1),-\hat{\boldsymbol{v}}(t-2),\cdots,-\hat{\boldsymbol{v}}(t-n_d)],\tag{165}$$

$$\hat{\boldsymbol{\psi}}(t)=[\boldsymbol{\psi}_{\mathrm{s}}(t),\hat{\boldsymbol{\psi}}_{\mathrm{n}}(t)],\tag{166}$$

$$\hat{\boldsymbol{w}}(t)=\boldsymbol{y}(t)+\boldsymbol{\psi}_{\mathrm{s}}(t)\hat{\boldsymbol{\alpha}}(t)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t)\boldsymbol{\varphi}(t),\tag{167}$$

$$\hat{\boldsymbol{v}}(t)=\boldsymbol{y}(t)+\hat{\boldsymbol{\psi}}(t)\hat{\boldsymbol{\vartheta}}(t)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t)\boldsymbol{\varphi}(t),\tag{168}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\begin{bmatrix}\hat{\boldsymbol{\alpha}}(t)\\ \hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(t)\end{bmatrix},\tag{169}$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\alpha}_1(t),\hat{\alpha}_2(t),\cdots,\hat{\alpha}_n(t)]^{\mathrm{T}},\tag{170}$$

$$\hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(t)=[\hat{c}_1(t),\hat{c}_2(t),\cdots,\hat{c}_{n_c}(t),\hat{d}_1(t),\hat{d}_2(t),\cdots,\hat{d}_{n_d}(t)]^{\mathrm{T}}.\tag{171}$$

HGELS 算法(157)—(171)计算参数估计的步骤如下:

1) 初始化:令 $t=1$,$\boldsymbol{P}_1(0)=p_0\boldsymbol{I}_{n_0}$,$\boldsymbol{P}_2(0)=p_0\boldsymbol{I}_{nr}$,

$\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_{n_0}/p_0$,$\hat{\boldsymbol{\theta}}(0)=\mathbf{1}_{nr\times m}/p_0$,$\boldsymbol{w}(i)=\mathbf{1}_m/p_0$,$\boldsymbol{v}(i)=\mathbf{1}_m/p_0$,$\boldsymbol{u}(i)=0$,$\boldsymbol{y}(i)=0$,$i\leqslant 0$,$p_0=10^6$.

2) 采集输入输出数据 $\boldsymbol{u}(t)$ 和 $\boldsymbol{y}(t)$,由式(163)—(166)构造 $\boldsymbol{\varphi}(t)$,$\boldsymbol{\psi}_{\rm s}(t)$,$\hat{\boldsymbol{\psi}}_{\rm n}(t)$ 和 $\hat{\boldsymbol{\psi}}(t)$.

3) 用式(158)和式(159)计算 $\boldsymbol{L}_1(t)$ 和 $\boldsymbol{P}_1(t)$,用式(161)和式(162)计算 $\boldsymbol{L}_2(t)$ 和 $\boldsymbol{P}_2(t)$.

4) 用式(157)和式(160)刷新参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$ 和参数估计矩阵 $\hat{\boldsymbol{\theta}}(t)$.

5) 用式(167)和式(168)分别计算 $\hat{\boldsymbol{w}}(t)$ 和 $\hat{\boldsymbol{v}}(t)$.

6) t 增 1,转到第 2)步.

5.3 递阶梯度迭代辨识方法

考虑数据长度为 L 的有限测量数据,即从 $t=1$ 到 $t=L$ 的一组数据,定义准则函数

$$J_6(\boldsymbol{\vartheta},\boldsymbol{\theta}):=\sum_{t=1}^{L}\|\boldsymbol{y}(t)+\boldsymbol{\psi}(t)\boldsymbol{\vartheta}-\boldsymbol{\theta}^{\rm T}\boldsymbol{\varphi}(t)\|^2,$$

令 $\hat{\boldsymbol{\alpha}}_k$ 和 $\hat{\boldsymbol{\theta}}_k$ 分别为参数向量 $\boldsymbol{\alpha}$ 和参数矩阵 $\boldsymbol{\theta}$ 第 k 次迭代的参数估计,$\mu_k\geqslant 0$ 为迭代步长.使用负梯度搜索,极小化准则函数 $J_6(\boldsymbol{\vartheta},\boldsymbol{\theta})$,未知矩阵 $\boldsymbol{\psi}(t)$ 用其估计 $\hat{\boldsymbol{\psi}}_k(t)$,可以得到估计参数向量 $\boldsymbol{\alpha}$ 和参数矩阵 $\boldsymbol{\theta}$ 的递阶梯度迭代算法(Hierarchical Gradient based Iterative algorithm,HGI 算法)[1,26]:

$$\hat{\boldsymbol{\vartheta}}_k=\hat{\boldsymbol{\vartheta}}_{k-1}-\mu_k\sum_{t=1}^{L}\hat{\boldsymbol{\psi}}_k^{\rm T}(t)[\boldsymbol{y}(t)+\hat{\boldsymbol{\psi}}_k(t)\hat{\boldsymbol{\vartheta}}_{k-1}-\hat{\boldsymbol{\theta}}_{k-1}^{\rm T}\boldsymbol{\varphi}(t)],\tag{172}$$

$$\hat{\boldsymbol{\theta}}_k=\hat{\boldsymbol{\theta}}_{k-1}+\mu_k(t)\sum_{t=1}^{L}\boldsymbol{\varphi}(t)[\boldsymbol{y}(t)+\hat{\boldsymbol{\psi}}_k(t)\hat{\boldsymbol{\vartheta}}_{k-1}-\hat{\boldsymbol{\theta}}_{k-1}^{\rm T}\boldsymbol{\varphi}(t)]^{\rm T},\tag{173}$$

$$\boldsymbol{\varphi}(t)=[\boldsymbol{u}^{\rm T}(t-1),\boldsymbol{u}^{\rm T}(t-2),\cdots,\boldsymbol{u}^{\rm T}(t-n)]^{\rm T},\tag{174}$$

$$\boldsymbol{\psi}_{\rm s}(t)=[\boldsymbol{y}(t-1),\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(t-n)],\tag{175}$$

$$\hat{\boldsymbol{\psi}}_{{\rm n},k}(t)=[\hat{\boldsymbol{w}}_{k-1}(t-1),\cdots,\hat{\boldsymbol{w}}_{k-1}(t-n_c),-\hat{\boldsymbol{v}}_{k-1}(t-1),\cdots,-\hat{\boldsymbol{v}}_{k-1}(t-n_d)],\tag{176}$$

$$\hat{\boldsymbol{\psi}}_k(t)=[\boldsymbol{\psi}_{\rm s}(t),\hat{\boldsymbol{\psi}}_{{\rm n},k}(t)],\tag{177}$$

$$\hat{\boldsymbol{w}}_k(t)=\boldsymbol{y}(t)+\boldsymbol{\psi}_{\rm s}(t)\hat{\boldsymbol{\alpha}}_k-\hat{\boldsymbol{\theta}}_k^{\rm T}\boldsymbol{\varphi}(t),\tag{178}$$

$$\hat{\boldsymbol{v}}_k(t)=\boldsymbol{y}(t)+\hat{\boldsymbol{\psi}}_k(t)\hat{\boldsymbol{\vartheta}}_k-\hat{\boldsymbol{\theta}}_k^{\rm T}\boldsymbol{\varphi}(t),\tag{179}$$

$$\mu_k\leqslant 2\Big(\sum_{t=1}^{L}[\|\hat{\boldsymbol{\psi}}_k(t)\|^2+\|\boldsymbol{\varphi}(t)\|^2]\Big)^{-1},\tag{180}$$

$$\hat{\boldsymbol{\vartheta}}_k=\begin{bmatrix}\hat{\boldsymbol{\alpha}}_k\\ \hat{\boldsymbol{\vartheta}}_{{\rm n},k}\end{bmatrix},\tag{181}$$

$$\hat{\boldsymbol{\alpha}}_k=[\hat{\alpha}_{1,k},\hat{\alpha}_{2,k},\cdots,\hat{\alpha}_{n,k}]^{\rm T},\tag{182}$$

$$\hat{\boldsymbol{\vartheta}}_{{\rm n},k}=[\hat{c}_{1,k},\hat{c}_{2,k},\cdots,\hat{c}_{n_c,k},\hat{d}_{1,k},\hat{d}_{2,k},\cdots,\hat{d}_{n_d,k}]^{\rm T}.\tag{183}$$

5.4 递阶最小二乘迭代辨识方法

根据参考文献[1,23]的最小二乘搜索原理,极小化准则函数 $J_6(\boldsymbol{\vartheta},\boldsymbol{\theta})$,未知信息矩阵 $\boldsymbol{\psi}(t)$ 用其估计 $\hat{\boldsymbol{\psi}}_k(t)$ 代替,可以得到估计 $\boldsymbol{\vartheta}$ 和 $\boldsymbol{\theta}$ 的递阶最小二乘迭代算法(Hierarchical Least Squares based Iterative algorithm,HLSI 算法)(取 $\mu_k=1$ 时)[1,26]:

$$\hat{\boldsymbol{\vartheta}}_k=-\Big[\sum_{t=1}^{L}\hat{\boldsymbol{\psi}}_k^{\rm T}(t)\hat{\boldsymbol{\psi}}_k(t)\Big]^{-1}\sum_{t=1}^{L}\hat{\boldsymbol{\psi}}_k^{\rm T}(t)[\boldsymbol{y}(t)-\hat{\boldsymbol{\theta}}_{k-1}^{\rm T}\boldsymbol{\varphi}(t)],\tag{184}$$

$$\hat{\boldsymbol{\theta}}_k=\Big[\sum_{t=1}^{L}\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^{\rm T}(t)\Big]^{-1}\sum_{t=1}^{L}\boldsymbol{\varphi}(t)[\boldsymbol{y}(t)+\hat{\boldsymbol{\psi}}_k(t)\hat{\boldsymbol{\vartheta}}_{k-1}],\tag{185}$$

$$\boldsymbol{\varphi}(t)=[\boldsymbol{u}^{\rm T}(t-1),\boldsymbol{u}^{\rm T}(t-2),\cdots,\boldsymbol{u}^{\rm T}(t-n)]^{\rm T},\tag{186}$$

$$\boldsymbol{\psi}_{\rm s}(t)=[\boldsymbol{y}(t-1),\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(t-n)],\tag{187}$$

$$\hat{\boldsymbol{\psi}}_{{\rm n},k}(t)=[\hat{\boldsymbol{w}}_{k-1}(t-1),\cdots,\hat{\boldsymbol{w}}_{k-1}(t-n_c),-\hat{\boldsymbol{v}}_{k-1}(t-1),\cdots,-\hat{\boldsymbol{v}}_{k-1}(t-n_d)],\tag{188}$$

$$\hat{\boldsymbol{\psi}}_k(t)=[\boldsymbol{\psi}_{\rm s}(t),\hat{\boldsymbol{\psi}}_{{\rm n},k}(t)],\tag{189}$$

$$\hat{\boldsymbol{w}}_k(t)=\boldsymbol{y}(t)+\boldsymbol{\psi}_{\rm s}(t)\hat{\boldsymbol{\alpha}}_k-\hat{\boldsymbol{\theta}}_k^{\rm T}\boldsymbol{\varphi}(t),\tag{190}$$

$$\hat{\boldsymbol{v}}_k(t)=\boldsymbol{y}(t)+\hat{\boldsymbol{\psi}}_k(t)\hat{\boldsymbol{\vartheta}}_k-\hat{\boldsymbol{\theta}}_k^{\rm T}\boldsymbol{\varphi}(t),\tag{191}$$

$$\hat{\boldsymbol{\vartheta}}_k=\begin{bmatrix}\hat{\boldsymbol{\alpha}}_k\\ \hat{\boldsymbol{\vartheta}}_{{\rm n},k}\end{bmatrix},\tag{192}$$

$$\hat{\boldsymbol{\alpha}}_k=[\hat{\alpha}_{1,k},\hat{\alpha}_{2,k},\cdots,\hat{\alpha}_{n,k}]^{\rm T},\tag{193}$$

$$\hat{\boldsymbol{\vartheta}}_{{\rm n},k}=[\hat{c}_{1,k},\hat{c}_{2,k},\cdots,\hat{c}_{n_c,k},\hat{d}_{1,k},\hat{d}_{2,k},\cdots,\hat{d}_{n_d,k}]^{\rm T}.\tag{194}$$

6 类多变量方程误差 ARMA 系统的递阶多新息辨识方法

考虑类多变量方程误差 ARMA 系统(134)对应的辨识模型(138),重写如下:

$$\boldsymbol{y}(t)+\boldsymbol{\psi}(t)\boldsymbol{\vartheta}=\boldsymbol{\theta}^{\rm T}\boldsymbol{\varphi}(t)+\boldsymbol{v}(t),\tag{195}$$

其中各变量定义同上.下面基于递阶辨识原理和多新息辨识理论,利用系统的观测数据$\{\boldsymbol{u}(t),\boldsymbol{y}(t):t=1,2,3,\cdots\}$,推导递阶多新息广义增广随机梯度辨识方法和递阶多新息广义增广最小二乘辨识方法,估计系统模型参数向量 $\boldsymbol{\vartheta}$ 和参数矩阵 $\boldsymbol{\theta}$.

6.1 递阶多新息广义增广随机梯度辨识方法

参考类多变量误差系统的递阶多新息随机梯度辨识算法(104)—(112)的推导,可以得到类多变量方程误差 ARMA 系统(195)的递阶多新息广义增广随机梯度算法(Hierarchical Multi-Innovation Generalized Extended Stochastic Gradient algorithm,H-MI-GESG 算法):

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)-\frac{\hat{\boldsymbol{\varPsi}}_1^{\rm T}(p,t)}{r(t)}[\boldsymbol{Y}_1(p,t)+\hat{\boldsymbol{\varPsi}}_1(p,t)\hat{\boldsymbol{\vartheta}}(t-1)-$$

$$(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1))\boldsymbol{\Phi}_1(p,t)], \tag{196}$$

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\boldsymbol{\Phi}_2(p,t)}{r(t)}[\boldsymbol{Y}_2(p,t)+\hat{\boldsymbol{\Psi}}_2^{\mathrm{T}}(p,t)$$

$$(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\vartheta}}(t-1))-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\Phi}_2(p,t)]^{\mathrm{T}}, \tag{197}$$

$$r(t)=r(t-1)+\|\hat{\boldsymbol{\Psi}}_1(p,t)\|^2+\|\boldsymbol{\Phi}_1(p,t)\|^2,$$

$$r(0)=1, \tag{198}$$

$$\boldsymbol{Y}_1(p,t)=\begin{bmatrix}\boldsymbol{y}(t)\\ \boldsymbol{y}(t-1)\\ \vdots\\ \boldsymbol{y}(t-p+1)\end{bmatrix}, \tag{199}$$

$$\hat{\boldsymbol{\Psi}}_1(p,t)=\begin{bmatrix}\hat{\boldsymbol{\psi}}(t)\\ \hat{\boldsymbol{\psi}}(t-1)\\ \vdots\\ \hat{\boldsymbol{\psi}}(t-p+1)\end{bmatrix}, \tag{200}$$

$$\boldsymbol{\Phi}_1(p,t)=\begin{bmatrix}\boldsymbol{\varphi}(t)\\ \boldsymbol{\varphi}(t-1)\\ \vdots\\ \boldsymbol{\varphi}(t-p+1)\end{bmatrix}, \tag{201}$$

$$\boldsymbol{Y}_2(p,t)=[\boldsymbol{y}(t),\boldsymbol{y}(t-1),\cdots,\boldsymbol{y}(t-p+1)], \tag{202}$$

$$\hat{\boldsymbol{\Psi}}_2(p,t)=[\hat{\boldsymbol{\psi}}(t),\hat{\boldsymbol{\psi}}(t-1),\cdots,\hat{\boldsymbol{\psi}}(t-p+1)]^{\mathrm{T}}, \tag{203}$$

$$\boldsymbol{\Phi}_2(p,t)=[\boldsymbol{\varphi}(t),\boldsymbol{\varphi}(t-1),\cdots,\boldsymbol{\varphi}(t-p+1)], \tag{204}$$

$$\boldsymbol{\varphi}(t)=[\boldsymbol{u}^{\mathrm{T}}(t-1),\boldsymbol{u}^{\mathrm{T}}(t-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(t-n)]^{\mathrm{T}}, \tag{205}$$

$$\hat{\boldsymbol{\psi}}(t)=[\boldsymbol{\psi}_{\mathrm{s}}(t),\hat{\boldsymbol{\psi}}_{\mathrm{n}}(t)], \tag{206}$$

$$\boldsymbol{\psi}_{\mathrm{s}}(t)=[\boldsymbol{y}(t-1),\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(t-n)], \tag{207}$$

$$\hat{\boldsymbol{\psi}}_{\mathrm{n}}(t)=[\hat{\boldsymbol{w}}(t-1),\hat{\boldsymbol{w}}(t-2),\cdots,\hat{\boldsymbol{w}}(t-n_c),$$

$$-\hat{\boldsymbol{v}}(t-1),-\hat{\boldsymbol{v}}(t-2),\cdots,-\hat{\boldsymbol{v}}(t-n_d)], \tag{208}$$

$$\hat{\boldsymbol{w}}(t)=\boldsymbol{y}(t)+\boldsymbol{\psi}_{\mathrm{s}}(t)\hat{\boldsymbol{\alpha}}(t)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t)\boldsymbol{\varphi}(t), \tag{209}$$

$$\hat{\boldsymbol{v}}(t)=\boldsymbol{y}(t)+\hat{\boldsymbol{\psi}}(t)\hat{\boldsymbol{\vartheta}}(t)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t)\boldsymbol{\varphi}(t), \tag{210}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\begin{bmatrix}\hat{\boldsymbol{\alpha}}(t)\\ \hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(t)\end{bmatrix}, \tag{211}$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\alpha}_1(t),\hat{\alpha}_2(t),\cdots,\hat{\alpha}_n(t)]^{\mathrm{T}}, \tag{212}$$

$$\hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(t)=[\hat{c}_1(t),\hat{c}_2(t),\cdots,\hat{c}_{n_c}(t),\hat{d}_1(t),$$

$$\hat{d}_2(t),\cdots,\hat{d}_{n_d}(t)]^{\mathrm{T}}. \tag{213}$$

H-MI-GESG 算法(196)—(212)估计参数 $\hat{\boldsymbol{\alpha}}(t)$ 和 $\hat{\boldsymbol{\theta}}(t)$ 的计算步骤如下:

1) 初始化:令 $t=1$,$\hat{\boldsymbol{\vartheta}}(0)=\boldsymbol{1}_{n_0}/p_0$,$\hat{\boldsymbol{\theta}}(0)=\boldsymbol{1}_{nr\times m}/p_0$,$\hat{\boldsymbol{w}}(i)=\boldsymbol{1}_m/p_0$,$\hat{\boldsymbol{v}}(i)=\boldsymbol{1}_m/p_0$,$\boldsymbol{u}(i)=0$,$\boldsymbol{y}(i)=0$,$i\leqslant 0$,$p_0=10^6$.

2) 采集输入输出数据 $\boldsymbol{u}(t)$ 和 $\boldsymbol{y}(t)$,用式(199)和(202)构造堆积输出向量 $\boldsymbol{Y}_1(p,t)$ 和堆积输出矩阵 $\boldsymbol{Y}_2(p,t)$,用式(201)和(204)构造堆积信息向量 $\boldsymbol{\Phi}_1(p,t)$ 和堆积信息矩阵 $\boldsymbol{\Phi}_2(p,t)$,用式(205)—(208)构造 $\boldsymbol{\varphi}(t)$,$\boldsymbol{\psi}_{\mathrm{s}}(t)$,$\hat{\boldsymbol{\psi}}_{\mathrm{n}}(t)$ 和 $\hat{\boldsymbol{\psi}}(t)$,用式(205)和式(208)构造堆积信息矩阵 $\hat{\boldsymbol{\Psi}}_1(p,t)$ 和 $\hat{\boldsymbol{\Psi}}_2(p,t)$.

3) 用式(198)计算 $r(t)$,用式(196)和式(197)刷新参数估计 $\hat{\boldsymbol{\vartheta}}(t)$ 和 $\hat{\boldsymbol{\theta}}(t)$.

4) 从式(211)$\hat{\boldsymbol{\vartheta}}(t)$ 中读出 $\hat{\boldsymbol{\alpha}}(t)$,用式(209)和式(210)计算 $\hat{\boldsymbol{w}}(t)$ 和 $\hat{\boldsymbol{v}}(t)$.

5) t 增 1,转到第 2)步.

6.2 递阶多新息广义增广最小二乘辨识方法

借鉴递阶多新息最小二乘辨识算法(122)—(133)的推导过程,未知矩阵 $\boldsymbol{\Psi}_1(p,t)$ 和 $\boldsymbol{\Psi}_2(p,t)$ 用其估计 $\hat{\boldsymbol{\Psi}}_1(p,t)$ 和 $\hat{\boldsymbol{\Psi}}_2(p,t)$ 代替,可得到类多变量方程误差 ARMA 系统(195)参数向量 $\boldsymbol{\vartheta}$ 和参数矩阵 $\boldsymbol{\theta}$ 的递阶多新息广义增广最小二乘算法(Hierarchical Multi-Innovation Generalized Extended Least Squares algorithm,H-MI-GELS 算法)[37]如下:

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\boldsymbol{L}_1(t)[\boldsymbol{Y}_1(p,t)+\hat{\boldsymbol{\Psi}}_1(p,t)\hat{\boldsymbol{\vartheta}}(t-1)-$$

$$(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1))\boldsymbol{\Phi}_1(p,t)], \tag{214}$$

$$\boldsymbol{L}_1(t)=-\boldsymbol{P}_1(t-1)\hat{\boldsymbol{\Psi}}_1^{\mathrm{T}}(p,t)[\boldsymbol{I}_{mp}+$$

$$\hat{\boldsymbol{\Psi}}_1(p,t)\boldsymbol{P}_1(t-1)\hat{\boldsymbol{\Psi}}_1^{\mathrm{T}}(p,t)]^{-1}, \tag{215}$$

$$\boldsymbol{P}_1(t)=\boldsymbol{P}_1(t-1)-\boldsymbol{P}_1(t-1)\hat{\boldsymbol{\Psi}}_1^{\mathrm{T}}(p,t)[\boldsymbol{I}_{mp}+$$

$$\hat{\boldsymbol{\Psi}}_1(p,t)\boldsymbol{P}_1(t-1)\hat{\boldsymbol{\Psi}}_1^{\mathrm{T}}(p,t)]^{-1}\hat{\boldsymbol{\Psi}}_1(p,t)\boldsymbol{P}_1(t-1)=$$

$$[\boldsymbol{I}_{n_0}+\boldsymbol{L}_1(t)\hat{\boldsymbol{\Psi}}_1(p,t)]\boldsymbol{P}_1(t-1), \tag{216}$$

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{L}_2(t)[\boldsymbol{Y}_2(p,t)+$$

$$\hat{\boldsymbol{\Psi}}_2^{\mathrm{T}}(p,t)(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\vartheta}}(t-1))-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\Phi}_2(p,t)]^{\mathrm{T}}, \tag{217}$$

$$\boldsymbol{L}_2(t)=\boldsymbol{P}_2(t-1)\boldsymbol{\Phi}_2(p,t)[\boldsymbol{I}_p+$$

$$\boldsymbol{\Phi}_2^{\mathrm{T}}(p,t)\boldsymbol{P}_2(t-1)\boldsymbol{\Phi}_2(p,t)]^{-1}, \tag{218}$$

$$\boldsymbol{P}_2(t)=\boldsymbol{P}_2(t-1)-\boldsymbol{P}_2(t-1)\boldsymbol{\Phi}_2(p,t)[\boldsymbol{I}_p+$$

$$\boldsymbol{\Phi}_2^{\mathrm{T}}(p,t)\boldsymbol{P}_2(t-1)\boldsymbol{\Phi}_2(p,t)]^{-1}\boldsymbol{\Phi}_2^{\mathrm{T}}(p,t)\boldsymbol{P}_2(t-1)=$$

$$[\boldsymbol{I}_{nr}-\boldsymbol{L}_2(t)\boldsymbol{\Phi}_2^{\mathrm{T}}(p,t)]\boldsymbol{P}_2(t-1), \tag{219}$$

$$\boldsymbol{Y}_1(p,t)=\begin{bmatrix}\boldsymbol{y}(t)\\ \boldsymbol{y}(t-1)\\ \vdots\\ \boldsymbol{y}(t-p+1)\end{bmatrix}, \tag{220}$$

$$\hat{\boldsymbol{\Psi}}_1(p,t)=\begin{bmatrix}\hat{\boldsymbol{\psi}}(t)\\ \hat{\boldsymbol{\psi}}(t-1)\\ \vdots\\ \hat{\boldsymbol{\psi}}(t-p+1)\end{bmatrix}, \tag{221}$$

$$\boldsymbol{\Phi}_1(p,t)=\begin{bmatrix}\boldsymbol{\varphi}(t)\\ \boldsymbol{\varphi}(t-1)\\ \vdots\\ \boldsymbol{\varphi}(t-p+1)\end{bmatrix}, \tag{222}$$

$$\boldsymbol{Y}_2(p,t)=[\boldsymbol{y}(t),\boldsymbol{y}(t-1),\cdots,\boldsymbol{y}(t-p+1)], \tag{223}$$

$$\hat{\boldsymbol{\Psi}}_2(p,t)=[\hat{\boldsymbol{\psi}}(t),\hat{\boldsymbol{\psi}}(t-1),\cdots,\hat{\boldsymbol{\psi}}(t-p+1)]^{\mathrm{T}}, \tag{224}$$

$$\boldsymbol{\Phi}_2(p,t)=[\boldsymbol{\varphi}(t),\boldsymbol{\varphi}(t-1),\cdots,\boldsymbol{\varphi}(t-p+1)], \tag{225}$$

$$\boldsymbol{\varphi}(t)=[\boldsymbol{u}^{\mathrm{T}}(t-1),\boldsymbol{u}^{\mathrm{T}}(t-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(t-n)]^{\mathrm{T}}, \tag{226}$$

$$\hat{\boldsymbol{\psi}}(t)=[\boldsymbol{\psi}_{\mathrm{s}}(t),\hat{\boldsymbol{\psi}}_{\mathrm{n}}(t)], \tag{227}$$

$$\boldsymbol{\psi}_{\mathrm{s}}(t)=[\boldsymbol{y}(t-1),\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(t-n)], \tag{228}$$

$$\hat{\boldsymbol{\psi}}_{\mathrm{n}}(t)=[\hat{\boldsymbol{w}}(t-1),\hat{\boldsymbol{w}}(t-2),\cdots,\hat{\boldsymbol{w}}(t-n_c),\ -\hat{\boldsymbol{v}}(t-1),-\hat{\boldsymbol{v}}(t-2),\cdots,-\hat{\boldsymbol{v}}(t-n_d)], \tag{229}$$

$$\hat{\boldsymbol{w}}(t)=\boldsymbol{y}(t)+\boldsymbol{\psi}_{\mathrm{s}}(t)\hat{\boldsymbol{\alpha}}(t)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t)\boldsymbol{\varphi}(t), \tag{230}$$

$$\hat{\boldsymbol{v}}(t)=\boldsymbol{y}(t)+\hat{\boldsymbol{\psi}}(t)\hat{\boldsymbol{\vartheta}}(t)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t)\boldsymbol{\varphi}(t), \tag{231}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\begin{bmatrix}\hat{\boldsymbol{\alpha}}(t)\\ \hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(t)\end{bmatrix}, \tag{232}$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\alpha}_1(t),\hat{\alpha}_2(t),\cdots,\hat{\alpha}_n(t)]^{\mathrm{T}}, \tag{233}$$

$$\hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(t)=[\hat{c}_1(t),\hat{c}_2(t),\cdots,\hat{c}_{n_c}(t),\hat{d}_1(t),\ \hat{d}_2(t),\cdots,\hat{d}_{n_d}(t)]^{\mathrm{T}}. \tag{234}$$

7 基于滤波的类多变量方程误差 ARMA 系统的递阶辨识方法

根据递阶辨识原理,结合滤波技术,进一步研究类多变量方程误差 ARMA 系统的递阶辨识方法.考虑类多变量方程误差 ARMA 系统(134),重写如下,

$$\alpha(z)\boldsymbol{y}(t)=\boldsymbol{Q}(z)\boldsymbol{u}(t)+\frac{D(z)}{C(z)}\boldsymbol{v}(t), \tag{235}$$

其中各变量定义同上,且重新定义中间变量如下,

$$\boldsymbol{w}(t):=\frac{D(z)}{C(z)}\boldsymbol{v}(t)\in\mathbf{R}^m. \tag{236}$$

将式(235)两边同时乘以$\frac{C(z)}{D(z)}$得到

$$\alpha(z)\frac{C(z)}{D(z)}\boldsymbol{y}(t)=\boldsymbol{Q}(z)\frac{C(z)}{D(z)}\boldsymbol{u}(t)+\boldsymbol{v}(t), \tag{237}$$

定义滤波输入向量 $\boldsymbol{u}_{\mathrm{f}}(t)$ 和滤波输出向量 $\boldsymbol{y}_{\mathrm{f}}(t)$ 为

$$\boldsymbol{u}_{\mathrm{f}}(t):=\frac{C(z)}{D(z)}\boldsymbol{u}(t)\in\mathbf{R}^r,$$

$$\boldsymbol{y}_{\mathrm{f}}(t):=\frac{C(z)}{D(z)}\boldsymbol{y}(t)\in\mathbf{R}^m.$$

代入式(237)可得

$$\alpha(z)\boldsymbol{y}_{\mathrm{f}}(t)=\boldsymbol{Q}(z)\boldsymbol{u}_{\mathrm{f}}(t)+\boldsymbol{v}(t). \tag{238}$$

定义滤波新息向量 $\boldsymbol{\varphi}_{\mathrm{f}}(t)$ 和滤波新息矩阵 $\boldsymbol{\psi}_{\mathrm{f}}(t)$ 如下:

$$\boldsymbol{\varphi}_{\mathrm{f}}(t):=[\boldsymbol{u}_{\mathrm{f}}^{\mathrm{T}}(t-1),\boldsymbol{u}_{\mathrm{f}}^{\mathrm{T}}(t-2),\cdots,\boldsymbol{u}_{\mathrm{f}}^{\mathrm{T}}(t-n)]^{\mathrm{T}}\in\mathbf{R}^{nr},$$

$$\boldsymbol{\psi}_{\mathrm{f}}(t):=[\boldsymbol{y}_{\mathrm{f}}(t-1),\boldsymbol{y}_{\mathrm{f}}(t-2),\cdots,\boldsymbol{y}_{\mathrm{f}}(t-n)]\in\mathbf{R}^{m\times n},$$

参数向量 $\boldsymbol{\alpha}$, $\boldsymbol{\vartheta}_{\mathrm{n}}$,参数矩阵 $\boldsymbol{\theta}$,信息矩阵 $\boldsymbol{\psi}_{\mathrm{n}}(t)$ 定义同上.

式(238)和(236)可以重写如下:

$$\boldsymbol{y}_{\mathrm{f}}(t)+\boldsymbol{\psi}_{\mathrm{f}}(t)\boldsymbol{\alpha}=\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}_{\mathrm{f}}(t)+\boldsymbol{v}(t), \tag{239}$$

$$\boldsymbol{w}(t)=-\boldsymbol{\psi}_{\mathrm{n}}(t)\boldsymbol{\vartheta}_{\mathrm{n}}+\boldsymbol{v}(t). \tag{240}$$

式(239)、(240)即为类多变量方程误差 ARMA 系统(235)的基于滤波的子系统递阶辨识模型和噪声模型.

7.1 基于滤波的递阶随机梯度辨识方法

根据式(239)和(240),定义 2 个梯度准则函数(gradient criterion function):

$$J_7(\boldsymbol{\alpha},\boldsymbol{\theta}):=\|\boldsymbol{y}_{\mathrm{f}}(t)+\boldsymbol{\psi}_{\mathrm{f}}(t)\boldsymbol{\alpha}-\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}_{\mathrm{f}}(t)\|^2,$$

$$J_8(\boldsymbol{\vartheta}_{\mathrm{n}}):=\|\boldsymbol{w}(t)+\boldsymbol{\psi}_{\mathrm{n}}(t)\boldsymbol{\vartheta}_{\mathrm{n}}\|^2.$$

使用负梯度搜索,极小化准则函数 $J_7(\boldsymbol{\alpha},\boldsymbol{\theta})$ 和 $J_8(\boldsymbol{\vartheta}_{\mathrm{n}})$,未知向量和矩阵 $\boldsymbol{\varphi}_{\mathrm{f}}(t)$, $\boldsymbol{\psi}_{\mathrm{f}}(t)$, $\boldsymbol{y}_{\mathrm{f}}(t)$, $\boldsymbol{\psi}_{\mathrm{n}}(t)$,用其估计值 $\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)$, $\hat{\boldsymbol{\psi}}_{\mathrm{f}}(t)$, $\hat{\boldsymbol{y}}_{\mathrm{f}}(t)$, $\hat{\boldsymbol{\psi}}_{\mathrm{n}}(t)$ 代替,可以得到类多变量方程误差 ARMA 系统(235)的基于滤波的递阶随机梯度算法(Filtering based Hierarchical Stochastic Gradient algorithm,F-HSG 算法):

$$\hat{\boldsymbol{\alpha}}(t)=\hat{\boldsymbol{\alpha}}(t-1)-\frac{\hat{\boldsymbol{\psi}}_{\mathrm{f}}^{\mathrm{T}}(t)}{r_1(t)}[\hat{\boldsymbol{y}}_{\mathrm{f}}(t)+\hat{\boldsymbol{\psi}}_{\mathrm{f}}(t)\hat{\boldsymbol{\alpha}}(t-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)], \tag{241}$$

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)}{r_1(t)}[\hat{\boldsymbol{y}}_{\mathrm{f}}(t)+\hat{\boldsymbol{\psi}}_{\mathrm{f}}(t)\hat{\boldsymbol{\alpha}}(t-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)]^{\mathrm{T}}, \tag{242}$$

$$r_1(t)=r_1(t-1)+\|\hat{\boldsymbol{\psi}}_{\mathrm{f}}(t)\|^2+\|\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)\|^2,\quad r_1(0)=1, \tag{243}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)=[\hat{\boldsymbol{u}}_{\mathrm{f}}^{\mathrm{T}}(t-1),\hat{\boldsymbol{u}}_{\mathrm{f}}^{\mathrm{T}}(t-2),\cdots,\hat{\boldsymbol{u}}_{\mathrm{f}}^{\mathrm{T}}(t-n)]^{\mathrm{T}}, \tag{244}$$

$$\hat{\boldsymbol{\psi}}_{\mathrm{f}}(t)=[\hat{\boldsymbol{y}}_{\mathrm{f}}(t-1),\hat{\boldsymbol{y}}_{\mathrm{f}}(t-2),\cdots,\hat{\boldsymbol{y}}_{\mathrm{f}}(t-n)], \tag{245}$$

$$\hat{\boldsymbol{u}}_{\mathrm{f}}(t)=-\hat{d}_1(t)\hat{\boldsymbol{u}}_{\mathrm{f}}(t-1)-\hat{d}_2(t)\hat{\boldsymbol{u}}_{\mathrm{f}}(t-2)-\cdots-\hat{d}_{n_d}(t)\hat{\boldsymbol{u}}_{\mathrm{f}}(t-n_d)+\boldsymbol{u}(t)+\hat{c}_1(t)\boldsymbol{u}(t-1)+\hat{c}_2(t)\boldsymbol{u}(t-2)+\cdots+\hat{c}_{n_c}(t)\boldsymbol{u}(t-n_c), \tag{246}$$

$$\hat{\boldsymbol{y}}_{\mathrm{f}}(t)=-\hat{d}_1(t)\hat{\boldsymbol{y}}_{\mathrm{f}}(t-1)-\hat{d}_2(t)\hat{\boldsymbol{y}}_{\mathrm{f}}(t-2)-\cdots-\hat{d}_{n_d}(t)\hat{\boldsymbol{y}}_{\mathrm{f}}(t-n_d)+\boldsymbol{y}(t)+\hat{c}_1(t)\boldsymbol{y}(t-1)+\hat{c}_2(t)\boldsymbol{y}(t-2)+\cdots+\hat{c}_{n_c}(t)\boldsymbol{y}(t-n_c), \tag{247}$$

$$\hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(t)=\hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(t-1)-\frac{\hat{\boldsymbol{\psi}}_{\mathrm{n}}^{\mathrm{T}}(t)}{r_2(t)}[\hat{\boldsymbol{w}}(t)+\hat{\boldsymbol{\psi}}_{\mathrm{n}}(t)\hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(t-1)], \tag{248}$$

$$r_2(t)=r_2(t-1)+\|\hat{\boldsymbol{\psi}}_{\mathrm{n}}(t)\|^2,\quad r_2(0)=1, \tag{249}$$

$$\boldsymbol{\varphi}(t)=[\boldsymbol{u}^{\mathrm{T}}(t-1),\boldsymbol{u}^{\mathrm{T}}(t-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(t-n)]^{\mathrm{T}}, \tag{250}$$

$$\boldsymbol{\psi}_{\mathrm{s}}(t)=[\boldsymbol{y}(t-1),\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(t-n)], \tag{251}$$

$$\hat{\boldsymbol{\psi}}_{\mathrm{n}}(t)=[\hat{\boldsymbol{w}}(t-1),\hat{\boldsymbol{w}}(t-2),\cdots,\hat{\boldsymbol{w}}(t-n_c),$$

$$-\hat{\boldsymbol{v}}(t-1),-\hat{\boldsymbol{v}}(t-2),\cdots,-\hat{\boldsymbol{v}}(t-n_d)],\tag{252}$$

$$\hat{\boldsymbol{w}}(t)=\boldsymbol{y}(t)+\boldsymbol{\psi}_{\rm s}(t)\hat{\boldsymbol{\alpha}}(t-1)-\hat{\boldsymbol{\theta}}^{\rm T}(t-1)\boldsymbol{\varphi}(t),\tag{253}$$

$$\hat{\boldsymbol{v}}(t)=\hat{\boldsymbol{y}}_{\rm f}(t)+\hat{\boldsymbol{\psi}}_{\rm f}(t)\hat{\boldsymbol{\alpha}}(t)-\hat{\boldsymbol{\theta}}^{\rm T}(t)\hat{\boldsymbol{\varphi}}_{\rm f}(t),\tag{254}$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\alpha}_1(t),\hat{\alpha}_2(t),\cdots,\hat{\alpha}_n(t)]^{\rm T},\tag{255}$$

$$\hat{\boldsymbol{\vartheta}}_{\rm n}(t)=[\hat{c}_1(t),\hat{c}_2(t),\cdots,\hat{c}_{n_c}(t),\hat{d}_1(t),\hat{d}_2(t),\cdots,\hat{d}_{n_d}(t)]^{\rm T}.\tag{256}$$

7.2 基于滤波的递阶最小二乘辨识方法

根据式(239)和(240),定义2个最小二乘准则函数(least squares criterion function):

$$J_9(\boldsymbol{\alpha},\boldsymbol{\theta}):=\sum_{j=1}^{t}\|\boldsymbol{y}_{\rm f}(j)+\boldsymbol{\psi}_{\rm f}(j)\boldsymbol{\alpha}-\boldsymbol{\theta}^{\rm T}\boldsymbol{\varphi}_{\rm f}(j)\|^2,$$

$$J_{10}(\boldsymbol{\vartheta}_{\rm n}):=\sum_{j=1}^{t}\|\boldsymbol{w}(j)+\boldsymbol{\psi}_{\rm n}(j)\boldsymbol{\vartheta}_{\rm n}\|^2.$$

根据递阶辨识原理和最小二乘原理,极小化准则函数 $J_9(\boldsymbol{\alpha},\boldsymbol{\theta})$ 和 $J_{10}(\boldsymbol{\vartheta}_{\rm n})$,可以得到类多变量方程误差ARMA系统(235)的基于滤波的递阶最小二乘辨识方法(Filtering based Hierarchical Least Squares algorithm, F-HLS算法):

$$\hat{\boldsymbol{\alpha}}(t)=\hat{\boldsymbol{\alpha}}(t-1)+\boldsymbol{L}_1(t)[\hat{\boldsymbol{y}}_{\rm f}(t)+\hat{\boldsymbol{\psi}}_{\rm f}(t)\hat{\boldsymbol{\alpha}}(t-1)-\hat{\boldsymbol{\theta}}^{\rm T}(t-1)\hat{\boldsymbol{\varphi}}_{\rm f}(t)],\tag{257}$$

$$\boldsymbol{L}_1(t)=-\boldsymbol{P}_1(t-1)\hat{\boldsymbol{\psi}}_{\rm f}^{\rm T}(t)[\boldsymbol{I}_m+\hat{\boldsymbol{\psi}}_{\rm f}(t)\boldsymbol{P}_1(t-1)\hat{\boldsymbol{\psi}}_{\rm f}^{\rm T}(t)]^{-1},\tag{258}$$

$$\boldsymbol{P}_1(t)=[\boldsymbol{I}_n+\boldsymbol{L}_1(t)\hat{\boldsymbol{\psi}}_{\rm f}(t)]\boldsymbol{P}_1(t-1),\tag{259}$$

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{L}_2(t)[\hat{\boldsymbol{y}}_{\rm f}(t)+\hat{\boldsymbol{\psi}}_{\rm f}(t)\hat{\boldsymbol{\alpha}}(t-1)-\hat{\boldsymbol{\theta}}^{\rm T}(t-1)\hat{\boldsymbol{\varphi}}_{\rm f}(t)]^{\rm T},\tag{260}$$

$$\boldsymbol{L}_2(t)=\frac{\boldsymbol{P}_2(t-1)\hat{\boldsymbol{\varphi}}_{\rm f}(t)}{1+\hat{\boldsymbol{\varphi}}_{\rm f}^{\rm T}(t)\boldsymbol{P}_2(t-1)\hat{\boldsymbol{\varphi}}_{\rm f}(t)},\tag{261}$$

$$\boldsymbol{P}_2(t)=[\boldsymbol{I}_{nr}-\boldsymbol{L}_2(t)\hat{\boldsymbol{\varphi}}_{\rm f}^{\rm T}(t)]\boldsymbol{P}_2(t-1),\tag{262}$$

$$\hat{\boldsymbol{\varphi}}_{\rm f}(t)=[\hat{\boldsymbol{u}}_{\rm f}^{\rm T}(t-1),\hat{\boldsymbol{u}}_{\rm f}^{\rm T}(t-2),\cdots,\hat{\boldsymbol{u}}_{\rm f}^{\rm T}(t-n)]^{\rm T},\tag{263}$$

$$\hat{\boldsymbol{\psi}}_{\rm f}(t)=[\hat{\boldsymbol{y}}_{\rm f}(t-1),\hat{\boldsymbol{y}}_{\rm f}(t-2),\cdots,\hat{\boldsymbol{y}}_{\rm f}(t-n)],\tag{264}$$

$$\hat{\boldsymbol{u}}_{\rm f}(t)=-\hat{d}_1(t)\hat{\boldsymbol{u}}_{\rm f}(t-1)-\hat{d}_2(t)\hat{\boldsymbol{u}}_{\rm f}(t-2)-\cdots-\hat{d}_{n_d}(t)\hat{\boldsymbol{u}}_{\rm f}(t-n_d)+\boldsymbol{u}(t)+\hat{c}_1(t)\boldsymbol{u}(t-1)+\hat{c}_2(t)\boldsymbol{u}(t-2)+\cdots+\hat{c}_{n_c}(t)\boldsymbol{u}(t-n_c),\tag{265}$$

$$\hat{\boldsymbol{y}}_{\rm f}(t)=-\hat{d}_1(t)\hat{\boldsymbol{y}}_{\rm f}(t-1)-\hat{d}_2(t)\hat{\boldsymbol{y}}_{\rm f}(t-2)-\cdots-\hat{d}_{n_d}(t)\hat{\boldsymbol{y}}_{\rm f}(t-n_d)+\boldsymbol{y}(t)+\hat{c}_1(t)\boldsymbol{y}(t-1)+\hat{c}_2(t)\boldsymbol{y}(t-2)+\cdots+\hat{c}_{n_c}(t)\boldsymbol{y}(t-n_c),\tag{266}$$

$$\hat{\boldsymbol{\vartheta}}_{\rm n}(t)=\hat{\boldsymbol{\vartheta}}_{\rm n}(t-1)+\boldsymbol{L}_3(t)[\hat{\boldsymbol{w}}(t)+\hat{\boldsymbol{\psi}}_{\rm n}(t)\hat{\boldsymbol{\vartheta}}_{\rm n}(t-1)],\tag{267}$$

$$\boldsymbol{L}_3(t)=-\boldsymbol{P}_3(t-1)\hat{\boldsymbol{\psi}}_{\rm n}^{\rm T}(t)[\boldsymbol{I}_m+\hat{\boldsymbol{\psi}}_{\rm n}(t)\boldsymbol{P}_3(t-1)\hat{\boldsymbol{\psi}}_{\rm n}^{\rm T}(t)]^{-1},\tag{268}$$

$$\boldsymbol{P}_3(t)=[\boldsymbol{I}_{n_c+n_d}+\boldsymbol{L}_3(t)\hat{\boldsymbol{\psi}}_{\rm n}(t)]\boldsymbol{P}_3(t-1),\tag{269}$$

$$\boldsymbol{\varphi}(t)=[\boldsymbol{u}^{\rm T}(t-1),\boldsymbol{u}^{\rm T}(t-2),\cdots,\boldsymbol{u}^{\rm T}(t-n)]^{\rm T},\tag{270}$$

$$\boldsymbol{\psi}_{\rm s}(t)=[\boldsymbol{y}(t-1),\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(t-n)],\tag{271}$$

$$\hat{\boldsymbol{\psi}}_{\rm n}(t)=[\hat{\boldsymbol{w}}(t-1),\hat{\boldsymbol{w}}(t-2),\cdots,\hat{\boldsymbol{w}}(t-n_c),-\hat{\boldsymbol{v}}(t-1),-\hat{\boldsymbol{v}}(t-2),\cdots,-\hat{\boldsymbol{v}}(t-n_d)],\tag{272}$$

$$\hat{\boldsymbol{w}}(t)=\boldsymbol{y}(t)+\boldsymbol{\psi}_{\rm s}(t)\hat{\boldsymbol{\alpha}}(t-1)-\hat{\boldsymbol{\theta}}^{\rm T}(t-1)\boldsymbol{\varphi}(t),\tag{273}$$

$$\hat{\boldsymbol{v}}(t)=\hat{\boldsymbol{y}}_{\rm f}(t)+\hat{\boldsymbol{\psi}}_{\rm f}(t)\hat{\boldsymbol{\alpha}}(t)-\hat{\boldsymbol{\theta}}^{\rm T}(t)\hat{\boldsymbol{\varphi}}_{\rm f}(t),\tag{274}$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\alpha}_1(t),\hat{\alpha}_2(t),\cdots,\hat{\alpha}_n(t)]^{\rm T},\tag{275}$$

$$\hat{\boldsymbol{\vartheta}}_{\rm n}(t)=[\hat{c}_1(t),\hat{c}_2(t),\cdots,\hat{c}_{n_c}(t),\hat{d}_1(t),\hat{d}_2(t),\cdots,\hat{d}_{n_d}(t)]^{\rm T}.\tag{276}$$

7.3 基于滤波的递阶梯度迭代辨识方法

考虑数据长度为 L 的有限测量数据,即从 $t=1$ 到 $t=L$ 的一组数据,定义准则函数

$$J_{11}(\boldsymbol{\alpha},\boldsymbol{\theta}):=\sum_{t=1}^{L}\|\boldsymbol{y}_{\rm f}(t)+\boldsymbol{\psi}_{\rm f}(t)\boldsymbol{\alpha}-\boldsymbol{\theta}^{\rm T}\boldsymbol{\varphi}_{\rm f}(t)\|^2,$$

$$J_{12}(\boldsymbol{\vartheta}_{\rm n}):=\sum_{t=1}^{L}\|\boldsymbol{w}(t)+\boldsymbol{\psi}_{\rm n}(t)\boldsymbol{\vartheta}_{\rm n}\|^2.$$

令 $\hat{\boldsymbol{\alpha}}_k$,$\hat{\boldsymbol{\vartheta}}_{{\rm n},k}$ 和 $\hat{\boldsymbol{\theta}}_k$ 分别为参数向量 $\boldsymbol{\alpha}$,$\boldsymbol{\vartheta}_{\rm n}$ 和 $\boldsymbol{\theta}$ 第 k 次迭代的参数估计,$\mu_{1,k}\geqslant 0$ 和 $\mu_{2,k}\geqslant 0$ 为迭代步长.使用负梯度搜索,极小化准则函数 $J_{11}(\boldsymbol{\alpha},\boldsymbol{\theta})$ 和 $J_{12}(\boldsymbol{\vartheta}_{\rm n})$,未知向量和矩阵 $\boldsymbol{\psi}_{\rm f}(t)$,$\boldsymbol{\psi}_{\rm n}(t)$,$\boldsymbol{\varphi}_{\rm f}(t)$,$\boldsymbol{y}_{\rm f}(t)$ 和 $\boldsymbol{w}(t)$,分别用其估计 $\hat{\boldsymbol{\psi}}_{{\rm f},k}(t)$,$\hat{\boldsymbol{\psi}}_{{\rm n},k}(t)$,$\hat{\boldsymbol{\varphi}}_{{\rm f},k}(t)$,$\hat{\boldsymbol{y}}_{{\rm f},k}(t)$ 和 $\hat{\boldsymbol{w}}_k(t)$ 代替,可以得到类多变量方程误差ARMA系统(235)的基于滤波的递阶梯度迭代算法(Filtering based Hierarchical Gradient Iterative algorithm, F-HGI算法):

$$\hat{\boldsymbol{\alpha}}_k=\hat{\boldsymbol{\alpha}}_{k-1}-\mu_{1,k}\sum_{t=1}^{L}\hat{\boldsymbol{\psi}}_{{\rm f},k}^{\rm T}(t)[\hat{\boldsymbol{y}}_{{\rm f},k}(t)+\hat{\boldsymbol{\psi}}_{{\rm f},k}(t)\hat{\boldsymbol{\alpha}}_{k-1}-\hat{\boldsymbol{\theta}}_{k-1}^{\rm T}\hat{\boldsymbol{\varphi}}_{{\rm f},k}(t)],\tag{277}$$

$$\hat{\boldsymbol{\theta}}_k=\hat{\boldsymbol{\theta}}_{k-1}+\mu_{1,k}\sum_{t=1}^{L}\hat{\boldsymbol{\varphi}}_{{\rm f},k}(t)[\hat{\boldsymbol{y}}_{{\rm f},k}(t)+\hat{\boldsymbol{\psi}}_{{\rm f},k}(t)\hat{\boldsymbol{\alpha}}_{k-1}-\hat{\boldsymbol{\theta}}_{k-1}^{\rm T}\hat{\boldsymbol{\varphi}}_{{\rm f},k}(t)]^{\rm T},\tag{278}$$

$$\hat{\boldsymbol{\psi}}_{{\rm f},k}(t)=[\hat{\boldsymbol{y}}_{{\rm f},k-1}(t-1),\hat{\boldsymbol{y}}_{{\rm f},k-1}(t-2),\cdots,\hat{\boldsymbol{y}}_{{\rm f},k-1}(t-n)],\tag{279}$$

$$\hat{\boldsymbol{\varphi}}_{{\rm f},k}(t)=[\hat{\boldsymbol{u}}_{{\rm f},k-1}^{\rm T}(t-1),\hat{\boldsymbol{u}}_{{\rm f},k-1}^{\rm T}(t-2),\cdots,\hat{\boldsymbol{u}}_{{\rm f},k-1}^{\rm T}(t-n)]^{\rm T},\tag{280}$$

$$\hat{\boldsymbol{u}}_{{\rm f},k}(t)=-\hat{d}_{1,k}\hat{\boldsymbol{u}}_{{\rm f},k}(t-1)-\hat{d}_{2,k}\hat{\boldsymbol{u}}_{{\rm f},k}(t-2)-\cdots-\hat{d}_{n_d,k}\hat{\boldsymbol{u}}_{{\rm f},k}(t-n_d)+\boldsymbol{u}(t)+\hat{c}_{1,k}\boldsymbol{u}(t-1)+\hat{c}_{2,k}\boldsymbol{u}(t-2)+\cdots+\hat{c}_{n_c,k}\boldsymbol{u}(t-n_c),\tag{281}$$

$$\hat{\boldsymbol{y}}_{{\rm f},k}(t)=-\hat{d}_{1,k}\hat{\boldsymbol{y}}_{{\rm f},k}(t-1)-\hat{d}_{2,k}\hat{\boldsymbol{y}}_{{\rm f},k}(t-2)-\cdots-\hat{d}_{n_d,k}\hat{\boldsymbol{y}}_{{\rm f},k}(t-n_d)+\boldsymbol{y}(t)+\hat{c}_{1,k}\boldsymbol{y}(t-1)+\hat{c}_{2,k}\boldsymbol{y}(t-2)+\cdots+\hat{c}_{n_c,k}\boldsymbol{y}(t-n_c),\tag{282}$$

$$\hat{\boldsymbol{\vartheta}}_{{\rm n},k}=\hat{\boldsymbol{\vartheta}}_{{\rm n},k-1}+\mu_{2,k}\sum_{t=1}^{L}\hat{\boldsymbol{\psi}}_{{\rm n},k}^{\rm T}(t)[\hat{\boldsymbol{w}}_k(t)+\hat{\boldsymbol{\psi}}_{{\rm n},k}(t)\hat{\boldsymbol{\vartheta}}_{{\rm n},k-1}],\tag{283}$$

$$\boldsymbol{\varphi}(t)=[\boldsymbol{u}^{\mathrm{T}}(t-1),\boldsymbol{u}^{\mathrm{T}}(t-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(t-n)]^{\mathrm{T}}, \tag{284}$$

$$\boldsymbol{\psi}_{\mathrm{s}}(t)=[\boldsymbol{y}(t-1),\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(t-n)], \tag{285}$$

$$\hat{\boldsymbol{\psi}}_{\mathrm{n},k}(t)=[\hat{\boldsymbol{w}}_{k-1}(t-1),\cdots,\hat{\boldsymbol{w}}_{k-1}(t-n_c),\ -\hat{\boldsymbol{v}}_{k-1}(t-1),\cdots,-\hat{\boldsymbol{v}}_{k-1}(t-n_d)], \tag{286}$$

$$\hat{\boldsymbol{w}}_k(t)=\boldsymbol{y}(t)+\boldsymbol{\psi}_{\mathrm{s}}(t)\hat{\boldsymbol{\alpha}}_{k-1}-\hat{\boldsymbol{\theta}}_{k-1}^{\mathrm{T}}\boldsymbol{\varphi}(t), \tag{287}$$

$$\hat{\boldsymbol{v}}_k(t)=\hat{\boldsymbol{y}}_{\mathrm{f},k}(t)+\hat{\boldsymbol{\psi}}_{\mathrm{f},k}(t)\hat{\boldsymbol{\alpha}}_k-\hat{\boldsymbol{\theta}}_k^{\mathrm{T}}\hat{\boldsymbol{\varphi}}_{\mathrm{f},k}(t), \tag{288}$$

$$\mu_{1,k}\leqslant 2\Big(\sum_{t=1}^{L}[\|\hat{\boldsymbol{\psi}}_{\mathrm{f},k}(t)\|^2+\|\hat{\boldsymbol{\varphi}}_{\mathrm{f},k}(t)\|^2]\Big)^{-1}, \tag{289}$$

$$\mu_{2,k}\leqslant 2\Big(\sum_{t=1}^{L}\|\hat{\boldsymbol{\psi}}_{\mathrm{n},k}(t)\|^2\Big)^{-1}, \tag{290}$$

$$\hat{\boldsymbol{\alpha}}_k=[\hat{\alpha}_{1,k},\hat{\alpha}_{2,k},\cdots,\hat{\alpha}_{n,k}]^{\mathrm{T}}, \tag{291}$$

$$\hat{\boldsymbol{\vartheta}}_{\mathrm{n},k}=[\hat{c}_{1,k},\hat{c}_{2,k},\cdots,\hat{c}_{n_c,k},\hat{d}_{1,k},\hat{d}_{2,k},\cdots,\hat{d}_{n_d,k}]^{\mathrm{T}}. \tag{292}$$

F-HGI 算法(277)—(292)计算参数估计的步骤如下:

1) 置初值:令 $k=1$, $\hat{\boldsymbol{\alpha}}_0=\mathbf{1}_n/p_0$, $\hat{\boldsymbol{\vartheta}}_{\mathrm{n},0}=\mathbf{1}_{n_c+n_d}/p_0$, $\hat{\boldsymbol{\theta}}_0=\mathbf{1}_{nr\times m}/p_0$, $\hat{\boldsymbol{w}}_0(t)=\mathbf{1}_m/p_0$, $\hat{\boldsymbol{v}}_0(t)=\mathbf{1}_m/p_0$, $\hat{\boldsymbol{u}}_{\mathrm{f},0}(t)=\mathbf{1}_m/p_0$, $\hat{\boldsymbol{y}}_{\mathrm{f},0}(t)=\mathbf{1}_m/p_0$, $p_0=10^6$.

2) 确定数据长度 L,收集输入输出数据$\{\boldsymbol{u}(t),\boldsymbol{y}(t):t=1,2,\cdots,L\}$,给定参数估计精度 ε,用式(284)构造 $\boldsymbol{\varphi}(t)$.

3) 用式(285)和式(286)构造 $\boldsymbol{\psi}_{\mathrm{s}}(t)$ 和 $\hat{\boldsymbol{\psi}}_{\mathrm{n},k}(t)$,用式(279)和式(280)构造 $\hat{\boldsymbol{\psi}}_{\mathrm{f},k}(t)$ 和 $\hat{\boldsymbol{\psi}}_{\mathrm{f},k}(t)$.

4) 用式(287)计算 $\hat{\boldsymbol{w}}_k(t)$,根据式(289)选择最大的 $\mu_{1,k}$,用式(283)刷新 $\hat{\boldsymbol{\vartheta}}_{\mathrm{n},k}$.

5) 从式(292)中读出 $\hat{c}_{i,k}$ 和 $\hat{d}_{i,k}$,由 式(281)和式(282)计算 $\hat{\boldsymbol{u}}_{\mathrm{f},k}(t)$ 和 $\hat{\boldsymbol{y}}_{\mathrm{f},k}(t)$.

6) 根据式(290)选择最大的 $\mu_{2,k}$,用式(277)和式(278)刷新参数估计 $\hat{\boldsymbol{\alpha}}_k$ 和 $\hat{\boldsymbol{\theta}}_k$,用式(288)计算 $\hat{\boldsymbol{v}}_k(t)$.

7) 比较 $\hat{\boldsymbol{\alpha}}_k$ 和 $\hat{\boldsymbol{\alpha}}_{k-1}$, $\hat{\boldsymbol{\vartheta}}_{\mathrm{n},k}$ 和 $\hat{\boldsymbol{\vartheta}}_{\mathrm{n},k-1}$, $\hat{\boldsymbol{\theta}}_k$ 和 $\hat{\boldsymbol{\theta}}_{k-1}$,如果 $\|\hat{\boldsymbol{\alpha}}_k-\hat{\boldsymbol{\alpha}}_{k-1}\|+\|\hat{\boldsymbol{\vartheta}}_{\mathrm{n},k}-\hat{\boldsymbol{\vartheta}}_{\mathrm{n},k-1}\|+\|\hat{\boldsymbol{\theta}}_k-\hat{\boldsymbol{\theta}}_k-1\|\leqslant\varepsilon$ 就结束计算,获得迭代次数 k 和参数估计 $\hat{\boldsymbol{\alpha}}_k$, $\hat{\boldsymbol{\vartheta}}_{\mathrm{n},k}$ 和 $\hat{\boldsymbol{\theta}}_k$;否则,$k$ 增 1,转到第 3)步.

7.4 基于滤波的递阶最小二乘迭代辨识方法

参考基于滤波的递阶梯度迭代辨识算法(277)—(292)的推导,令 $J_{11}(\boldsymbol{\alpha},\boldsymbol{\theta})$ 和 $J_{12}(\boldsymbol{\vartheta}_{\mathrm{n}})$ 对 $\boldsymbol{\alpha},\boldsymbol{\theta}$ 和 $\boldsymbol{\vartheta}_{\mathrm{n}}$ 的偏导数为零,可以推导出类多变量方程误差类 ARMA 系统(235)的基于滤波的最小二乘迭代算法(Filtering based Hierarchical Least Squares Iterative algorithm,F-HLSI 算法):

$$\hat{\boldsymbol{\alpha}}_k=\hat{\boldsymbol{\alpha}}_{k-1}-\Big[\sum_{t=1}^{L}\hat{\boldsymbol{\psi}}_{\mathrm{f},k}^{\mathrm{T}}(t)\hat{\boldsymbol{\psi}}_{\mathrm{f},k}(t)\Big]^{-1}\sum_{t=1}^{L}\hat{\boldsymbol{\psi}}_{\mathrm{f},k}^{\mathrm{T}}(t)[\hat{\boldsymbol{y}}_{\mathrm{f},k}(t)-\hat{\boldsymbol{\psi}}_{\mathrm{f},k}(t)\hat{\boldsymbol{\alpha}}_{k-1}-\hat{\boldsymbol{\theta}}_{k-1}^{\mathrm{T}}\hat{\boldsymbol{\varphi}}_{\mathrm{f},k}(t)], \tag{293}$$

$$\hat{\boldsymbol{\theta}}_k=\hat{\boldsymbol{\theta}}_{k-1}+\Big[\sum_{t=1}^{L}\hat{\boldsymbol{\varphi}}_{\mathrm{f},k}(t)\hat{\boldsymbol{\varphi}}_{\mathrm{f},k}^{\mathrm{T}}(t)\Big]^{-1}\sum_{t=1}^{L}\hat{\boldsymbol{\varphi}}_{\mathrm{f},k}(t)[\hat{\boldsymbol{y}}_{\mathrm{f},k}(t)-\hat{\boldsymbol{\psi}}_{\mathrm{f},k}(t)\hat{\boldsymbol{\alpha}}_{k-1}-\hat{\boldsymbol{\theta}}_{k-1}^{\mathrm{T}}\hat{\boldsymbol{\varphi}}_{\mathrm{f},k}(t)]^{\mathrm{T}}, \tag{294}$$

$$\hat{\boldsymbol{\psi}}_{\mathrm{f},k}(t)=[\hat{\boldsymbol{y}}_{\mathrm{f},k-1}(t-1),\hat{\boldsymbol{y}}_{\mathrm{f},k-1}(t-2),\cdots,\hat{\boldsymbol{y}}_{\mathrm{f},k-1}(t-n)], \tag{295}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{f},k}(t)=[\hat{\boldsymbol{u}}_{\mathrm{f},k-1}^{\mathrm{T}}(t-1),\hat{\boldsymbol{u}}_{\mathrm{f},k-1}^{\mathrm{T}}(t-2),\cdots,\hat{\boldsymbol{u}}_{\mathrm{f},k-1}^{\mathrm{T}}(t-n)]^{\mathrm{T}}, \tag{296}$$

$$\hat{\boldsymbol{u}}_{\mathrm{f},k}(t)=-\hat{d}_{1,k}\hat{\boldsymbol{u}}_{\mathrm{f},k}(t-1)-\hat{d}_{2,k}\hat{\boldsymbol{u}}_{\mathrm{f},k}(t-2)-\cdots-\hat{d}_{n_d,k}\hat{\boldsymbol{u}}_{\mathrm{f},k}(t-n_d)+\boldsymbol{u}(t)+\hat{c}_{1,k}\boldsymbol{u}(t-1)+\hat{c}_{2,k}\boldsymbol{u}(t-2)+\cdots+\hat{c}_{n_c,k}\boldsymbol{u}(t-n_c), \tag{297}$$

$$\hat{\boldsymbol{y}}_{\mathrm{f},k}(t)=-\hat{d}_{1,k}\hat{\boldsymbol{y}}_{\mathrm{f},k}(t-1)-\hat{d}_{2,k}\hat{\boldsymbol{y}}_{\mathrm{f},k}(t-2)-\cdots-\hat{d}_{n_d,k}\hat{\boldsymbol{y}}_{\mathrm{f},k}(t-n_d)+\boldsymbol{y}(t)+\hat{c}_{1,k}\boldsymbol{y}(t-1)+\hat{c}_{2,k}\boldsymbol{y}(t-2)+\cdots+\hat{c}_{n_c,k}\boldsymbol{y}(t-n_c), \tag{298}$$

$$\hat{\boldsymbol{\vartheta}}_{\mathrm{n},k}=\hat{\boldsymbol{\vartheta}}_{\mathrm{n},k-1}+\Big[\sum_{t=1}^{L}\hat{\boldsymbol{\psi}}_{\mathrm{n},k}^{\mathrm{T}}(t)\hat{\boldsymbol{\psi}}_{\mathrm{n},k}(t)\Big]^{-1}\sum_{t=1}^{L}\hat{\boldsymbol{\psi}}_{\mathrm{n},k}^{\mathrm{T}}(t)[\hat{\boldsymbol{w}}_k(t)+\hat{\boldsymbol{\psi}}_{\mathrm{n},k}(t)\hat{\boldsymbol{\vartheta}}_{\mathrm{n},k-1}], \tag{299}$$

$$\boldsymbol{\varphi}(t)=[\boldsymbol{u}^{\mathrm{T}}(t-1),\boldsymbol{u}^{\mathrm{T}}(t-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(t-n)]^{\mathrm{T}}, \tag{300}$$

$$\boldsymbol{\psi}_{\mathrm{s}}(t)=[\boldsymbol{y}(t-1),\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(t-n)], \tag{301}$$

$$\hat{\boldsymbol{\psi}}_{\mathrm{n},k}(t)=[\hat{\boldsymbol{w}}_{k-1}(t-1),\cdots,\hat{\boldsymbol{w}}_{k-1}(t-n_c),\ -\hat{\boldsymbol{v}}_{k-1}(t-1),\cdots,-\hat{\boldsymbol{v}}_{k-1}(t-n_d)], \tag{302}$$

$$\hat{\boldsymbol{w}}_k(t)=\boldsymbol{y}(t)+\boldsymbol{\psi}_{\mathrm{s}}(t)\hat{\boldsymbol{\alpha}}_{k-1}-\hat{\boldsymbol{\theta}}_{k-1}^{\mathrm{T}}\boldsymbol{\varphi}(t), \tag{303}$$

$$\hat{\boldsymbol{v}}_k(t)=\hat{\boldsymbol{y}}_{\mathrm{f},k}(t)+\hat{\boldsymbol{\psi}}_{\mathrm{f},k}(t)\hat{\boldsymbol{\alpha}}_k-\hat{\boldsymbol{\theta}}_k^{\mathrm{T}}\hat{\boldsymbol{\varphi}}_{\mathrm{f},k}(t), \tag{304}$$

$$\hat{\boldsymbol{\alpha}}_k=[\hat{\alpha}_{1,k},\hat{\alpha}_{2,k},\cdots,\hat{\alpha}_{n,k}]^{\mathrm{T}}, \tag{305}$$

$$\hat{\boldsymbol{\vartheta}}_{\mathrm{n},k}=[\hat{c}_{1,k},\hat{c}_{2,k},\cdots,\hat{c}_{n_c,k},\hat{d}_{1,k},\hat{d}_{2,k},\cdots,\hat{d}_{n_d,k}]^{\mathrm{T}}. \tag{306}$$

8 基于滤波的类多变量方程误差 ARMA 系统的递阶多新息辨识方法

利用滤波技术、递阶辨识原理,研究和提出类多变量方程误差 ARMA 系统的基于滤波的递阶多新息辨识方法.

8.1 基于滤波的递阶多新息随机梯度算法

考虑类多变量方程误差 ARMA 系统(235)对应的基于滤波的子系统递阶辨识模型(239)和噪声模型(240),重写如下:

$$\boldsymbol{y}_{\mathrm{f}}(t)+\boldsymbol{\psi}_{\mathrm{f}}(t)\boldsymbol{\alpha}=\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}_{\mathrm{f}}(t)+\boldsymbol{v}(t), \tag{307}$$

$$\boldsymbol{w}(t)=-\boldsymbol{\psi}_{\mathrm{n}}(t)\boldsymbol{\vartheta}_{\mathrm{n}}+\boldsymbol{v}(t), \tag{308}$$

其中各变量定义同上.参考 4.1 节递阶多新息随机梯度辨识方法(104)—(112)和基于滤波的递阶随机梯度算法(241)—(256)的推导,可以得到估计 $\boldsymbol{\alpha}$, $\boldsymbol{\theta}$ 的基于滤波的递阶多新息随机梯度算法(Filtering based Hierarchical Multi-Innovation Stochastic Gradient algorithm,F-HMISG 算法):

$$\hat{\boldsymbol{\alpha}}(t)=\hat{\boldsymbol{\alpha}}(t-1)-\frac{\hat{\boldsymbol{\Psi}}_{\mathrm{f1}}^{\mathrm{T}}(p,t)}{r_1(t)}[\hat{\boldsymbol{Y}}_{\mathrm{f1}}(p,t)+\hat{\boldsymbol{\Psi}}_{\mathrm{f1}}(p,t)\hat{\boldsymbol{\alpha}}(t-1)-(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1))\hat{\boldsymbol{\Phi}}_{\mathrm{f1}}(p,t)], \tag{309}$$

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\hat{\boldsymbol{\Phi}}_{\mathrm{f2}}(p,t)}{r_1(t)}[\hat{\boldsymbol{Y}}_{\mathrm{f2}}(p,t)+\hat{\boldsymbol{\Psi}}_{\mathrm{f2}}^{\mathrm{T}}(p,t)(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\alpha}}(t-1))-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\hat{\boldsymbol{\Phi}}_{\mathrm{f2}}(p,t)]^{\mathrm{T}}, \tag{310}$$

$$r_1(t)=r_1(t-1)+\|\hat{\boldsymbol{\Psi}}_{\mathrm{f1}}(p,t)\|^2+\|\hat{\boldsymbol{\Phi}}_{\mathrm{f1}}(p,t)\|^2,\quad r_1(0)=1, \tag{311}$$

$$\hat{\boldsymbol{Y}}_{\mathrm{f1}}(p,t)=\begin{bmatrix}\hat{\boldsymbol{y}}_{\mathrm{f}}(t)\\ \hat{\boldsymbol{y}}_{\mathrm{f}}(t-1)\\ \vdots\\ \hat{\boldsymbol{y}}_{\mathrm{f}}(t-p+1)\end{bmatrix}, \tag{312}$$

$$\hat{\boldsymbol{\Psi}}_{\mathrm{f1}}(p,t)=\begin{bmatrix}\hat{\boldsymbol{\psi}}_{\mathrm{f}}(t)\\ \hat{\boldsymbol{\psi}}_{\mathrm{f}}(t-1)\\ \vdots\\ \hat{\boldsymbol{\psi}}_{\mathrm{f}}(t-p+1)\end{bmatrix}, \tag{313}$$

$$\hat{\boldsymbol{\Phi}}_{\mathrm{f1}}(p,t)=\begin{bmatrix}\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)\\ \hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-1)\\ \vdots\\ \hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-p+1)\end{bmatrix}, \tag{314}$$

$$\hat{\boldsymbol{Y}}_{\mathrm{f2}}(p,t)=[\hat{\boldsymbol{y}}_{\mathrm{f}}(t),\hat{\boldsymbol{y}}_{\mathrm{f}}(t-1),\cdots,\hat{\boldsymbol{y}}_{\mathrm{f}}(t-p+1)], \tag{315}$$

$$\hat{\boldsymbol{\Psi}}_{\mathrm{f2}}(p,t)=[\hat{\boldsymbol{\psi}}_{\mathrm{f}}(t),\hat{\boldsymbol{\psi}}_{\mathrm{f}}(t-1),\cdots,\hat{\boldsymbol{\psi}}_{\mathrm{f}}(t-p+1)]^{\mathrm{T}}, \tag{316}$$

$$\hat{\boldsymbol{\Phi}}_{\mathrm{f2}}(p,t)=[\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t),\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-1),\cdots,\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-p+1)], \tag{317}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)=[\hat{\boldsymbol{u}}_{\mathrm{f}}^{\mathrm{T}}(t-1),\hat{\boldsymbol{u}}_{\mathrm{f}}^{\mathrm{T}}(t-2),\cdots,\hat{\boldsymbol{u}}_{\mathrm{f}}^{\mathrm{T}}(t-n)], \tag{318}$$

$$\hat{\boldsymbol{\psi}}_{\mathrm{f}}(t)=[\hat{\boldsymbol{y}}_{\mathrm{f}}(t-1),\hat{\boldsymbol{y}}_{\mathrm{f}}(t-2),\cdots,\hat{\boldsymbol{y}}_{\mathrm{f}}(t-n)], \tag{319}$$

$$\hat{\boldsymbol{u}}_{\mathrm{f}}(t)=-\hat{d}_1(t)\hat{\boldsymbol{u}}_{\mathrm{f}}(t-1)-\hat{d}_2(t)\hat{\boldsymbol{u}}_{\mathrm{f}}(t-2)-\cdots-\hat{d}_{n_d}(t)\hat{\boldsymbol{u}}_{\mathrm{f}}(t-n_d)+\boldsymbol{u}(t)+\hat{c}_1(t)\boldsymbol{u}(t-1)+\hat{c}_2(t)\boldsymbol{u}(t-2)+\cdots+\hat{c}_{n_c}(t)\boldsymbol{u}(t-n_c), \tag{320}$$

$$\hat{\boldsymbol{y}}_{\mathrm{f}}(t)=-\hat{d}_1(t)\hat{\boldsymbol{y}}_{\mathrm{f}}(t-1)-\hat{d}_2(t)\hat{\boldsymbol{y}}_{\mathrm{f}}(t-2)-\cdots-\hat{d}_{n_d}(t)\hat{\boldsymbol{y}}_{\mathrm{f}}(t-n_d)+\boldsymbol{y}(t)+\hat{c}_1(t)\boldsymbol{y}(t-1)+\hat{c}_2(t)\boldsymbol{y}(t-2)+\cdots+\hat{c}_{n_c}(t)\boldsymbol{y}(t-n_c), \tag{321}$$

$$\hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(t)=\hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(t-1)-\frac{\hat{\boldsymbol{\Psi}}_{\mathrm{n}}^{\mathrm{T}}(p,t)}{r_2(t)}[\hat{\boldsymbol{W}}(p,t)+\hat{\boldsymbol{\Psi}}_{\mathrm{n}}(p,t)\hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(t-1)], \tag{322}$$

$$r_2(t)=r_2(t-1)+\|\hat{\boldsymbol{\psi}}_{\mathrm{n}}(t)\|^2,\quad r_2(0)=1, \tag{323}$$

$$\hat{\boldsymbol{W}}(p,t)=\begin{bmatrix}\hat{\boldsymbol{w}}(t)\\ \hat{\boldsymbol{w}}(t-1)\\ \vdots\\ \hat{\boldsymbol{w}}(t-p+1)\end{bmatrix}, \tag{324}$$

$$\hat{\boldsymbol{\Psi}}_{\mathrm{n}}(p,t)=\begin{bmatrix}\hat{\boldsymbol{\psi}}_{\mathrm{n}}(t)\\ \hat{\boldsymbol{\psi}}_{\mathrm{n}}(t-1)\\ \vdots\\ \hat{\boldsymbol{\psi}}_{\mathrm{n}}(t-p+1)\end{bmatrix}, \tag{325}$$

$$\boldsymbol{\varphi}(t)=[\boldsymbol{u}^{\mathrm{T}}(t-1),\boldsymbol{u}^{\mathrm{T}}(t-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(t-n)]^{\mathrm{T}}, \tag{326}$$

$$\boldsymbol{\psi}_{\mathrm{s}}(t)=[\boldsymbol{y}(t-1),\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(t-n)], \tag{327}$$

$$\hat{\boldsymbol{\psi}}_{\mathrm{n}}(t)=[\hat{\boldsymbol{w}}(t-1),\hat{\boldsymbol{w}}(t-2),\cdots,\hat{\boldsymbol{w}}(t-n_c),-\hat{\boldsymbol{v}}(t-1),-\hat{\boldsymbol{v}}(t-2),\cdots,-\hat{\boldsymbol{v}}(t-n_d)], \tag{328}$$

$$\hat{\boldsymbol{w}}(t)=\boldsymbol{y}(t)+\boldsymbol{\psi}_{\mathrm{s}}(t)\hat{\boldsymbol{\alpha}}(t-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t), \tag{329}$$

$$\hat{\boldsymbol{v}}(t)=\hat{\boldsymbol{y}}_{\mathrm{f}}(t)+\hat{\boldsymbol{\psi}}_{\mathrm{f}}(t)\hat{\boldsymbol{\alpha}}(t)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t)\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t), \tag{330}$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\alpha}_1(t),\hat{\alpha}_2(t),\cdots,\hat{\alpha}_n(t)]^{\mathrm{T}}, \tag{331}$$

$$\hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(t)=[\hat{c}_1(t),\hat{c}_2(t),\cdots,\hat{c}_{n_c}(t),\hat{d}_1(t),\hat{d}_2(t),\cdots,\hat{d}_{n_d}(t)]^{\mathrm{T}}. \tag{332}$$

F-HMISG 算法的计算步骤如下.

1) 令 $t=1$, $\hat{\boldsymbol{\alpha}}(0)=\mathbf{1}_n/p_0$, $\hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(0)=\mathbf{1}_{n_c+n_d}/p_0$, $\hat{\boldsymbol{\theta}}(0)=\mathbf{1}_{nr\times m}/p_0$, $\hat{\boldsymbol{u}}_{\mathrm{f}}(i)=\mathbf{1}_m/p_0$, $\hat{\boldsymbol{y}}_{\mathrm{f}}(i)=\mathbf{1}_m/p_0$, $\hat{\boldsymbol{w}}(i)=\mathbf{1}_m/p_0$, $\hat{\boldsymbol{v}}(i)=\mathbf{1}_m/p_0$, $\boldsymbol{u}(i)=0$, $\boldsymbol{y}(i)=0$, $i\leqslant 0$, $p_0=10^6$.

2) 采集输入输出数据 $\boldsymbol{u}(t)$ 和 $\boldsymbol{y}(t)$,根据式(324)—(328)构造 $\boldsymbol{\varphi}(t)$, $\boldsymbol{\psi}_{\mathrm{s}}(t)$, $\hat{\boldsymbol{\psi}}_{\mathrm{n}}(t)$ 和 $\hat{\boldsymbol{\Psi}}_{\mathrm{n}}(p,t)$, $\hat{\boldsymbol{W}}(p,t)$,用式(312)—(319)构造 $\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)$, $\hat{\boldsymbol{\psi}}_{\mathrm{f}}(t)$, $\hat{\boldsymbol{Y}}_{\mathrm{f1}}(p,t)$, $\hat{\boldsymbol{\Psi}}_{\mathrm{f1}}(p,t)$, $\hat{\boldsymbol{\Phi}}_{\mathrm{f1}}(p,t)$ 和 $\hat{\boldsymbol{Y}}_{\mathrm{f2}}(p,t)$, $\hat{\boldsymbol{\Psi}}_{\mathrm{f2}}(p,t)$, $\hat{\boldsymbol{\Phi}}_{\mathrm{f2}}(p,t)$.

3) 根据式(329)计算 $\hat{\boldsymbol{w}}(t)$,用式(323)计算 $r_2(t)$,用式(322)刷新 $\hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(t)$.

4) 从式(332)中读出 $\hat{c}_i(t)$ 和 $\hat{d}_i(t)$,并用式(320)和(321)计算 $\hat{\boldsymbol{u}}_{\mathrm{f}}(t)$ 和 $\hat{\boldsymbol{y}}_{\mathrm{f}}(t)$.

5) 由式(311)计算 $r_1(t)$,用式(309)和(310)刷新 $\hat{\boldsymbol{\alpha}}(t)$ 和 $\hat{\boldsymbol{\theta}}(t)$.

6) 由式(330)计算 $\hat{\boldsymbol{v}}(t)$.

7) t 增1,转到第2)步.

8.2 基于滤波的递阶多新息最小二乘算法

参考递阶多新息最小二乘算法算法(122)—(133)的推导,结合滤波技术和多新息辨识理论,可以推导出类多变量方程误差 ARMA 系统(235)的基于滤波的递阶多新息最小二乘算法(Filtering based Hierarchical Multi-Innovation Least Squares algorithm, F-HMILS 算法):

$$\hat{\boldsymbol{\alpha}}(t)=\hat{\boldsymbol{\alpha}}(t-1)+\boldsymbol{L}_1(t)[\hat{\boldsymbol{Y}}_{\mathrm{f1}}(p,t)+\hat{\boldsymbol{\Psi}}_{\mathrm{f1}}(p,t)\hat{\boldsymbol{\alpha}}(t-1)-(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1))\hat{\boldsymbol{\Phi}}_{\mathrm{f1}}(p,t)], \tag{333}$$

$$\boldsymbol{L}_1(t)=-\boldsymbol{P}_1(t-1)\hat{\boldsymbol{\Psi}}_{\mathrm{f1}}^{\mathrm{T}}(p,t)[\boldsymbol{I}_{mp}+$$

$$\hat{\boldsymbol{\Psi}}_{\mathrm{f1}}(p,t)\boldsymbol{P}_1(t-1)\hat{\boldsymbol{\Psi}}_{\mathrm{f1}}^{\mathrm{T}}(p,t)]^{-1},\qquad(334)$$

$$\boldsymbol{P}_1(t)=\boldsymbol{P}_1(t-1)-\boldsymbol{P}_1(t-1)\hat{\boldsymbol{\Psi}}_{\mathrm{f1}}^{\mathrm{T}}(p,t)[\boldsymbol{I}_p+\hat{\boldsymbol{\Psi}}_{\mathrm{f1}}(p,t)\boldsymbol{P}_1(t-1)\hat{\boldsymbol{\Psi}}_{\mathrm{f1}}^{\mathrm{T}}(p,t)]^{-1}\hat{\boldsymbol{\Psi}}_{\mathrm{f1}}(p,t)\boldsymbol{P}_1(t-1)=[\boldsymbol{I}_n+\boldsymbol{L}_1(t)\hat{\boldsymbol{\Psi}}_{\mathrm{f1}}(p,t)]\boldsymbol{P}_1(t-1),\qquad(335)$$

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{L}_2(t)[\hat{\boldsymbol{Y}}_{\mathrm{f2}}(p,t)+\hat{\boldsymbol{\Psi}}_{\mathrm{f2}}^{\mathrm{T}}(p,t)(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\alpha}}(t-1))-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\hat{\boldsymbol{\Phi}}_{\mathrm{f2}}(p,t)]^{\mathrm{T}},\qquad(336)$$

$$\boldsymbol{L}_2(t)=\boldsymbol{P}_2(t-1)\hat{\boldsymbol{\Phi}}_{\mathrm{f2}}(p,t)[\boldsymbol{I}_p+\hat{\boldsymbol{\Phi}}_{\mathrm{f2}}^{\mathrm{T}}(p,t)\boldsymbol{P}_2(t-1)\hat{\boldsymbol{\Phi}}_{\mathrm{f2}}(p,t)]^{-1},\qquad(337)$$

$$\boldsymbol{P}_2^{-1}(t)=\boldsymbol{P}_2(t-1)-\boldsymbol{P}_2(t-1)\hat{\boldsymbol{\Phi}}_{\mathrm{f2}}(p,t)[\boldsymbol{I}_p+\hat{\boldsymbol{\Phi}}_{\mathrm{f2}}^{\mathrm{T}}(p,t)\boldsymbol{P}_2(t-1)\hat{\boldsymbol{\Phi}}_{\mathrm{f2}}(p,t)]^{-1}\hat{\boldsymbol{\Phi}}_{\mathrm{f2}}^{\mathrm{T}}(p,t)\boldsymbol{P}_2(t-1)=[\boldsymbol{I}_{nr}-\boldsymbol{L}_2(t)\hat{\boldsymbol{\Phi}}_{\mathrm{f2}}^{\mathrm{T}}(p,t)]\boldsymbol{P}_2(t-1),\qquad(338)$$

$$\hat{\boldsymbol{Y}}_{\mathrm{f1}}(p,t)=\begin{bmatrix}\hat{\boldsymbol{y}}_{\mathrm{f}}(t)\\\hat{\boldsymbol{y}}_{\mathrm{f}}(t-1)\\\vdots\\\hat{\boldsymbol{y}}_{\mathrm{f}}(t-p+1)\end{bmatrix},\qquad(339)$$

$$\hat{\boldsymbol{\Psi}}_{\mathrm{f1}}(p,t)=\begin{bmatrix}\hat{\boldsymbol{\psi}}_{\mathrm{f}}(t)\\\hat{\boldsymbol{\psi}}_{\mathrm{f}}(t-1)\\\vdots\\\hat{\boldsymbol{\psi}}_{\mathrm{f}}(t-p+1)\end{bmatrix},\qquad(340)$$

$$\hat{\boldsymbol{\Phi}}_{\mathrm{f1}}(p,t)=\begin{bmatrix}\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)\\\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-1)\\\vdots\\\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-p+1)\end{bmatrix},\qquad(341)$$

$$\hat{\boldsymbol{Y}}_{\mathrm{f2}}(p,t)=[\hat{\boldsymbol{y}}_{\mathrm{f}}(t),\hat{\boldsymbol{y}}_{\mathrm{f}}(t-1),\cdots,\hat{\boldsymbol{y}}_{\mathrm{f}}(t-p+1)],\qquad(342)$$

$$\hat{\boldsymbol{\Psi}}_{\mathrm{f2}}(p,t)=[\hat{\boldsymbol{\psi}}_{\mathrm{f}}(t),\hat{\boldsymbol{\psi}}_{\mathrm{f}}(t-1),\cdots,\hat{\boldsymbol{\psi}}_{\mathrm{f}}(t-p+1)]^{\mathrm{T}},\qquad(343)$$

$$\hat{\boldsymbol{\Phi}}_{\mathrm{f2}}(p,t)=[\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t),\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-1),\cdots,\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-p+1)],\qquad(344)$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)=[\hat{\boldsymbol{u}}_{\mathrm{f}}^{\mathrm{T}}(t-1),\hat{\boldsymbol{u}}_{\mathrm{f}}^{\mathrm{T}}(t-2),\cdots,\hat{\boldsymbol{u}}_{\mathrm{f}}^{\mathrm{T}}(t-n)],\qquad(345)$$

$$\hat{\boldsymbol{\psi}}_{\mathrm{f}}(t)=[\hat{\boldsymbol{y}}_{\mathrm{f}}(t-1),\hat{\boldsymbol{y}}_{\mathrm{f}}(t-2),\cdots,\hat{\boldsymbol{y}}_{\mathrm{f}}(t-n)],\qquad(346)$$

$$\hat{\boldsymbol{u}}_{\mathrm{f}}(t)=-\hat{d}_1(t)\hat{\boldsymbol{u}}_{\mathrm{f}}(t-1)-\hat{d}_2(t)\hat{\boldsymbol{u}}_{\mathrm{f}}(t-2)-\cdots-\hat{d}_{n_d}(t)\hat{\boldsymbol{u}}_{\mathrm{f}}(t-n_d)+\boldsymbol{u}(t)+\hat{c}_1(t)\boldsymbol{u}(t-1)+\hat{c}_2(t)\boldsymbol{u}(t-2)+\cdots+\hat{c}_{n_c}(t)\boldsymbol{u}(t-n_c),\qquad(347)$$

$$\hat{\boldsymbol{y}}_{\mathrm{f}}(t)=-\hat{d}_1(t)\hat{\boldsymbol{y}}_{\mathrm{f}}(t-1)-\hat{d}_2(t)\hat{\boldsymbol{y}}_{\mathrm{f}}(t-2)-\cdots-\hat{d}_{n_d}(t)\hat{\boldsymbol{y}}_{\mathrm{f}}(t-n_d)+\boldsymbol{y}(t)+\hat{c}_1(t)\boldsymbol{y}(t-1)+\hat{c}_2(t)\boldsymbol{y}(t-2)+\cdots+\hat{c}_{n_c}(t)\boldsymbol{y}(t-n_c),\qquad(348)$$

$$\hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(t)=\hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(t-1)+\boldsymbol{P}_3(t)\hat{\boldsymbol{\Psi}}_{\mathrm{n}}^{\mathrm{T}}(p,t)[\hat{\boldsymbol{W}}(p,t)+\hat{\boldsymbol{\Psi}}_{\mathrm{n}}(p,t)\hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(t-1)],\qquad(349)$$

$$\boldsymbol{L}_3(t)=-\boldsymbol{P}_3(t-1)\hat{\boldsymbol{\Psi}}_{\mathrm{n}}^{\mathrm{T}}(p,t)[\boldsymbol{I}_{mp}+\hat{\boldsymbol{\Psi}}_{\mathrm{n}}(p,t)\boldsymbol{P}_3(t-1)\hat{\boldsymbol{\Psi}}_{\mathrm{n}}^{\mathrm{T}}(p,t)]^{-1},\qquad(350)$$

$$\boldsymbol{P}_3(t)=\boldsymbol{P}_3(t-1)-\boldsymbol{P}_3(t-1)\hat{\boldsymbol{\Psi}}_{\mathrm{n}}^{\mathrm{T}}(p,t)[\boldsymbol{I}_{mp}+\hat{\boldsymbol{\Psi}}_{\mathrm{n}}(p,t)\boldsymbol{P}_3(t-1)\hat{\boldsymbol{\Psi}}_{\mathrm{n}}^{\mathrm{T}}(p,t)]^{-1}\hat{\boldsymbol{\Psi}}_{\mathrm{n}}(p,t)\boldsymbol{P}_3(t-1)=[\boldsymbol{I}_{n_c+n_d}+\boldsymbol{L}_3(t)\hat{\boldsymbol{\Psi}}_{\mathrm{n}}(p,t)]\boldsymbol{P}_3(t-1),\qquad(351)$$

$$\hat{\boldsymbol{W}}(p,t)=\begin{bmatrix}\hat{\boldsymbol{w}}(t)\\\hat{\boldsymbol{w}}(t-1)\\\vdots\\\hat{\boldsymbol{w}}(t-p+1)\end{bmatrix},\qquad(352)$$

$$\hat{\boldsymbol{\Psi}}_{\mathrm{n}}(p,t)=\begin{bmatrix}\hat{\boldsymbol{\psi}}_{\mathrm{n}}(t)\\\hat{\boldsymbol{\psi}}_{\mathrm{n}}(t-1)\\\vdots\\\hat{\boldsymbol{\psi}}_{\mathrm{n}}(t-p+1)\end{bmatrix},\qquad(353)$$

$$\boldsymbol{\varphi}(t)=[\boldsymbol{u}^{\mathrm{T}}(t-1),\boldsymbol{u}^{\mathrm{T}}(t-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(t-n)]^{\mathrm{T}},\qquad(354)$$

$$\boldsymbol{\psi}_{\mathrm{s}}(t)=[\boldsymbol{y}(t-1),\boldsymbol{y}(t-2),\cdots,\boldsymbol{y}(t-n)],\qquad(355)$$

$$\hat{\boldsymbol{\psi}}_{\mathrm{n}}(t)=[\hat{\boldsymbol{w}}(t-1),\hat{\boldsymbol{w}}(t-2),\cdots,\hat{\boldsymbol{w}}(t-n_c),-\hat{\boldsymbol{v}}(t-1),-\hat{\boldsymbol{v}}(t-2),\cdots,-\hat{\boldsymbol{v}}(t-n_d)],\qquad(356)$$

$$\hat{\boldsymbol{w}}(t)=\boldsymbol{y}(t)+\boldsymbol{\psi}_{\mathrm{s}}(t)\hat{\boldsymbol{\alpha}}(t-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t-1)\boldsymbol{\varphi}(t),\qquad(357)$$

$$\hat{\boldsymbol{v}}(t)=\hat{\boldsymbol{y}}_{\mathrm{f}}(t)+\hat{\boldsymbol{\psi}}_{\mathrm{f}}(t)\hat{\boldsymbol{\alpha}}(t)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t)\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t),\qquad(358)$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\boldsymbol{\alpha}}_1(t),\hat{\boldsymbol{\alpha}}_2(t),\cdots,\hat{\boldsymbol{\alpha}}_n(t)]^{\mathrm{T}},\qquad(359)$$

$$\hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(t)=[\hat{c}_1(t),\hat{c}_2(t),\cdots,\hat{c}_{n_c}(t),\hat{d}_1(t),\hat{d}_2(t),\cdots,\hat{d}_{n_d}(t)]^{\mathrm{T}}.\qquad(360)$$

F-MIHLS 算法的计算步骤如下：

1）令 $t=1$，$\hat{\boldsymbol{\alpha}}(0)=\mathbf{1}_n/p_0$，$\hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(0)=\mathbf{1}_{n_c+n_d}/p_0$，$\hat{\boldsymbol{\theta}}(0)=\mathbf{1}_{nr\times m}/p_0$，$\boldsymbol{P}_1(0)=p_0\boldsymbol{I}_n$，$\boldsymbol{P}_2(0)=p_0\boldsymbol{I}_{nr}$，$\boldsymbol{P}_3(0)=p_0\boldsymbol{I}_{n_c+n_d}$，$\hat{\boldsymbol{u}}_{\mathrm{f}}(i)=\mathbf{1}_m/p_0$，$\hat{\boldsymbol{y}}_{\mathrm{f}}(i)=\mathbf{1}_m/p_0$，$\hat{\boldsymbol{w}}(i)=\mathbf{1}_m/p_0$，$\hat{\boldsymbol{v}}(i)=\mathbf{1}_m/p_0$，$\boldsymbol{u}(i)=0$，$\boldsymbol{y}(i)=0$，$i\leqslant 0$，$p_0=10^6$.

2）采集输入输出数据 $\boldsymbol{u}(t)$ 和 $\boldsymbol{y}(t)$，根据式(352)—(356)可以构造 $\boldsymbol{\varphi}(t)$，$\boldsymbol{\psi}_{\mathrm{s}}(t)$，$\hat{\boldsymbol{\psi}}_{\mathrm{n}}(t)$ 和 $\hat{\boldsymbol{\Psi}}_{\mathrm{n}}(p,t)$，$\hat{\boldsymbol{W}}(p,t)$，用式(339)—(346)构造 $\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)$，$\hat{\boldsymbol{\psi}}_{\mathrm{f}}(t)$，$\hat{\boldsymbol{Y}}_{\mathrm{f1}}(p,t)$，$\hat{\boldsymbol{\Psi}}_{\mathrm{f1}}(p,t)$，$\hat{\boldsymbol{\Phi}}_{\mathrm{f1}}(p,t)$ 和 $\hat{\boldsymbol{Y}}_{\mathrm{f2}}(p,t)$，$\hat{\boldsymbol{\Psi}}_{\mathrm{f2}}(p,t)$，$\hat{\boldsymbol{\Phi}}_{\mathrm{f2}}(p,t)$.

3）根据式(357)计算 $\hat{\boldsymbol{w}}(t)$，用式(350)和式(351)计算增益 $\boldsymbol{L}_3(t)$ 和协方差阵 $\boldsymbol{P}_3(t)$.

4）用式(349)刷新 $\hat{\boldsymbol{\vartheta}}_{\mathrm{n}}(t)$.

5）从式(360)中读出 $\hat{c}_i(t)$ 和 $\hat{d}_i(t)$，并用式(347)和式(348)计算 $\hat{\boldsymbol{u}}_{\mathrm{f}}(t)$ 和 $\hat{\boldsymbol{y}}_{\mathrm{f}}(t)$.

6）用式(334)和式(335)计算增益 $\boldsymbol{L}_1(t)$ 和协方差阵 $\boldsymbol{P}_1(t)$，用式(337)和式(338)计算 $\boldsymbol{L}_2(t)$ 和 $\boldsymbol{P}_2(t)$.

7）分别用式(333)和式(336)刷新 $\hat{\boldsymbol{\alpha}}(t)$ 和

$\hat{\boldsymbol{\theta}}(t)$,由式(358)计算 $\hat{\boldsymbol{v}}(t)$.

8) t 增 1,转到第 2)步.

9 结语

本文针对类多变量方程误差类系统,利用递阶辨识原理、多新息辨识理论、数据滤波技术,研究和提出了递阶辨识方法、递阶迭代辨识方法、递阶多新息辨识方法,以及基于数据滤波的递阶辨识方法和基于滤波的递阶多新息辨识方法.这些方法可以推广到下列类多变量方程误差类系统

$$\alpha(z)\boldsymbol{y}(t)=\boldsymbol{Q}(z)\boldsymbol{u}(t)+\frac{\boldsymbol{D}(z)}{C(z)}\boldsymbol{v}(t), \quad (361)$$

$$\alpha(z)\boldsymbol{y}(t)=\boldsymbol{Q}(z)\boldsymbol{u}(t)+\boldsymbol{C}^{-1}(z)D(z)\boldsymbol{v}(t), \quad (362)$$

$$\alpha(z)\boldsymbol{y}(t)=\boldsymbol{Q}(z)\boldsymbol{u}(t)+\boldsymbol{C}^{-1}(z)\boldsymbol{D}(z)\boldsymbol{v}(t) \quad (363)$$

和类多变量输出误差类系统

$$\boldsymbol{y}(t)=\frac{\boldsymbol{Q}(z)}{\alpha(z)}\boldsymbol{u}(t)+\boldsymbol{v}(t), \quad (364)$$

$$\boldsymbol{y}(t)=\frac{\boldsymbol{Q}(z)}{\alpha(z)}\boldsymbol{u}(t)+\frac{D(z)}{C(z)}\boldsymbol{v}(t), \quad (365)$$

$$\boldsymbol{y}(t)=\frac{\boldsymbol{Q}(z)}{\alpha(z)}\boldsymbol{u}(t)+\frac{\boldsymbol{D}(z)}{C(z)}\boldsymbol{v}(t), \quad (366)$$

$$\boldsymbol{y}(t)=\frac{\boldsymbol{Q}(z)}{\alpha(z)}\boldsymbol{u}(t)+\boldsymbol{C}^{-1}(z)D(z)\boldsymbol{v}(t), \quad (367)$$

$$\boldsymbol{y}(t)=\frac{\boldsymbol{Q}(z)}{\alpha(z)}\boldsymbol{u}(t)+\boldsymbol{C}^{-1}(z)\boldsymbol{D}(z)\boldsymbol{v}(t) \quad (368)$$

或

$$f(z)\boldsymbol{y}(t)=\frac{\boldsymbol{Q}(z)}{\alpha(z)}\boldsymbol{u}(t)+\boldsymbol{v}(t), \quad (369)$$

$$f(z)\boldsymbol{y}(t)=\frac{\boldsymbol{Q}(z)}{\alpha(z)}\boldsymbol{u}(t)+\frac{D(z)}{C(z)}\boldsymbol{v}(t), \quad (370)$$

$$f(z)\boldsymbol{y}(t)=\frac{\boldsymbol{Q}(z)}{\alpha(z)}\boldsymbol{u}(t)+\frac{\boldsymbol{D}(z)}{C(z)}\boldsymbol{v}(t), \quad (371)$$

$$f(z)\boldsymbol{y}(t)=\frac{\boldsymbol{Q}(z)}{\alpha(z)}\boldsymbol{u}(t)+\boldsymbol{C}^{-1}(z)D(z)\boldsymbol{v}(t), \quad (372)$$

$$f(z)\boldsymbol{y}(t)=\frac{\boldsymbol{Q}(z)}{\alpha(z)}\boldsymbol{u}(t)+\boldsymbol{C}^{-1}(z)\boldsymbol{D}(z)\boldsymbol{v}(t), \quad (373)$$

其中 $\alpha(z)$,$C(z)$,$D(z)$ 和 $\boldsymbol{Q}(z)$ 定义同上,$f(z)$ 为首 1 多项式,$\boldsymbol{C}(z)$ 和 $\boldsymbol{D}(z)$ 为首单位阵的多项式矩阵.

参考文献

References

[1] 丁锋.系统辨识新论[M].北京:科学出版社,2013
DING Feng. System identification: New theory and methods[M].Beijing:Science Press,2013

[2] 丁锋.系统辨识:辨识方法性能分析[M].北京:科学出版社,2014
DING Feng. System identification: Performance analysis for identification methods [M]. Beijing: Science Press,2014

[3] 丁锋.系统辨识(1):辨识导引[J].南京信息工程大学学报:自然科学版,2011,3(1):1-22
DING Feng.System identification.Part A:Introduction to the identification[J].Journal of Nanjing University of Information Science and Technology:Natural Science Edition,2011,3(1):1-22

[4] 丁锋.系统辨识(2):系统描述的基本模型[J].南京信息工程大学学报:自然科学版,2011,3(2):97-117
DING Feng.System identification.Part B:Basic models for system description[J].Journal of Nanjing University of Information Science and Technology: Natural Science Edition,2011,3(2):97-117

[5] 丁锋.系统辨识(3):辨识精度与辨识基本问题[J].南京信息工程大学学报:自然科学版,2011,3(3):193-226
DING Feng. System identification. Part C: Identification accuracy and basic problems[J].Journal of Nanjing University of Information Science and Technology: Natural Science Edition,2011,3(3):193-226

[6] 丁锋.系统辨识(4):辅助模型辨识思想与方法[J].南京信息工程大学学报:自然科学版,2011,3(4):289-318
DING Feng.System identification.Part D:Auxiliary model identification idea and methods[J].Journal of Nanjing University of Information Science and Technology: Natural Science Edition,2011,3(4):289-318

[7] 丁锋.系统辨识(5):迭代搜索原理与辨识方法[J].南京信息工程大学学报:自然科学版,2011,3(6):481-510
DING Feng.System identification.Part E:Iterative search principle and identification methods [J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2011, 3(6): 481-510

[8] 丁锋.系统辨识(6):多新息辨识理论与方法[J].南京信息工程大学学报:自然科学版,2012,4(1):1-28
DING Feng.System identification.Part F:Multi-innovation identification theory and methods[J].Journal of Nanjing University of Information Science and Technology: Natural Science Edition,2012,4(1):1-28

[9] 丁锋.系统辨识(7):递阶辨识原理与方法[J].南京信息工程大学学报:自然科学版,2012,4(2):97-124
DING Feng. System identification. Part G: Hierarchical identification principle and methods [J]. Journal of Nanjing University of Information Science and Technology:Natural Science Edition,2012,4(2):97-124

[10] 丁锋.系统辨识(8):耦合辨识概念与方法[J].南京信息工程大学学报:自然科学版,2012,4(3):193-212
DING Feng.System identification.Part H:Coupling identification concept and methods[J].Journal of Nanjing University of Information Science and Technology: Natural Science Edition,2012,4(3):193-212

[11] 丁锋.辨识方法的计算效率(1):递推算法[J].南京信息工程大学学报:自然科学版,2012,4(4):289-300
DING Feng.Computational efficiency of the identification methods.Part A:Recursive algorithms[J].Journal of Nanjing University of Information Science and Technology: Natural Science Edition,2012,4(4):289-300

[12] 丁锋.辨识方法的计算效率(2):迭代算法[J].南京信息工程大学学报:自然科学版,2012,4(5):385-401
DING Feng.Computational efficiency of the identification methods. Part B: Iterative algorithms [J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2012, 4 (5): 385-401
[13] 丁锋.辨识方法的计算效率(3):信息向量耦合算法[J].南京信息工程大学学报:自然科学版,2012,4(6):481-495
DING Feng.Computational efficiency of the identification methods. Part C: Coupled information vector algorithms [J].Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2012, 4 (6): 481-495
[14] 丁锋,汪菲菲.多元系统耦合多新息随机梯度类辨识方法[J].南京信息工程大学学报:自然科学版,2014,6(1):1-16
DING Feng, WANG Feifei. Coupled multi-innovation stochastic gradient type identification methods for multivariate systems[J].Journal of Nanjing University of Information Science and Technology: Natural Science Edition,2014,6(1):1-16
[15] 丁锋,汪菲菲,汪学海.多元伪线性回归系统部分耦合多新息随机梯度类辨识方法[J].南京信息工程大学学报:自然科学版,2014,6(2):97-112
DING Feng,WANG Feifei,WANG Xuehai.Partially Coupled multi-innovation stochastic gradient type identification methods for multivariate pseudo-linear regressive systems[J].Journal of Nanjing University of Information Science and Technology:Natural Science Edition,2014,6(2):97-112
[16] 丁锋,汪菲菲,汪学海.类多变量输出误差系统的耦合多新息辨识方法[J].南京信息工程大学学报:自然科学版,2014,6(3):193-210
DING Feng, WANG Feifei, WANG Xuehai. Coupled multi-innovation identification methods for multivariable output-error-like systems [J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition,2014,6(3):193-210
[17] 丁锋,汪菲菲,汪学海.多变量方程误差类系统的部分耦合迭代辨识方法[J].南京信息工程大学学报:自然科学版,2014,6(4):289-305
DING Feng,WANG Feifei,WANG Xuehai.Partially coupled iterative identification methods for multivariable equation error type systems[J].Journal of Nanjing University of Information Science and Technology: Natural Science Edition,2014,6(4):289-305
[18] 丁锋,杨家本.大系统的递阶辨识[J].自动化学报,1999,25(5):647-654
DING Feng, YANG Jiaben. Hierarchical identification of large scale systems[J].Acta Automatica Sinica,1999,25(5):647-654
[19] Ding F,Chen T.Hierarchical gradient-based identification of multivariable discrete-time systems [J]. Automatica, 2005,41(2):315-325
[20] Ding F, Chen T. Hierarchical least squares identification methods for multivariable systems[J].IEEE Transactions on Automatic Control,2005,50(3):397-402
[21] Ding F, Chen T. Gradient based iterative algorithms for solving a class of matrix equations[J].IEEE Transactions on Automatic Control,2005,50(8):1216-1221
[22] Ding F, Chen T. Iterative least squares solutions of coupled Sylvester matrix equations [J]. Systems and Control Letters,2005,54(2):95-107
[23] Ding F,Chen T.On iterative solutions of general coupled matrix equations[J].SIAM Journal on Control and Optimization,2006,44(6):2269-2284
[24] Ding F,Liu X P,Ding J.Iterative solutions of the generalized Sylvester matrix equations by using the hierarchical identification principle [J]. Applied Mathematics and Computation,2008,197(1):41-50
[25] Ding F,Zhang H M.Gradient-based iterative algorithm for a class of the coupled matrix equations related to control systems[J].IET Control Theory and Applications 2014, 8.doi:10. 1049/iet-cta.2013. 1044
[26] Han H Q,Xie L,Ding F,et al.Hierarchical least squares based iterative identification for multivariable systems with moving average noises[J].Mathematical and Computer Modelling,2010,51(9/10):1213-1220
[27] ZNZhang 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
[28] Ding F,Chen T.Hierarchical identification of lifted state-space models for general dual-rate systems [J]. IEEE Transactions on Circuits and Systems.I:Regular Papers, 2005,52(6):1179-1187
[29] 丁锋,萧德云.多变量系统状态空间模型的递阶辨识[J].控制与决策,2005,20(8):848-853,859
DING Feng, XIAO Deyun. Hierarchical identification of state space models for multivariable systems[J].Control and Decision,2005,20(8):848-853,859
[30] Wang L Y, Ding F, Liu X P. Consistency of HLS estimation algorithms for MIMO ARX-like systems[J]. Applied Mathematics and Computation, 2007, 190(2): 1081-1093
[31] Ding F,Qiu L,Chen T.Reconstruction of continuous-time systems from their non-uniformly sampled discrete-time systems[J].Automatica,2009,45(2):324-332
[32] Ding J,Ding F,Liu X P,et al.Hierarchical least squares identification for linear SISO systems with dual-rate sampled-data[J].IEEE Transactions on Automatic Control, 2011,56(11):2677-2683
[33] Ding F.Hierarchical multi-innovation stochastic gradient algorithm for Hammerstein nonlinear system modeling [J]. Applied Mathematical Modelling, 2013, 37 (4): 1694-1704
[34] Liu Y J,Ding F,Shi Y.An efficient hierarchical identification method for general dual-rate sampled-data systems [J].Automatica,2014,50(3):962-970.
[35] 丁锋,谢新民,方崇智.时变系统辨识的多新息方法[J].自动化学报,1996,22(1):85-91
DING Feng, XIE Xinmin, FANG Chongzhi. Multi-innovation identification methods for time-varying systems [J].Acta Automatica Sinica,1996,22(1):85-91
[36] Ding F,Chen T.Performance analysis of multi-innovation gradient type identification methods [J]. Automatica, 2007,43(1):1-14
[37] 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
[38] Ding F,Liu G,Liu X P.Parameter estimation with scarce measurements[J].Automatica,2011,47(8):1646-1655
[39] 丁锋,萧德云,丁韬.多新息随机梯度辨识方法[J].控制理论与应用,2003,20(6):870-874
DING Feng, XIAO Deyun, DING Tao. Multi-innovation stochastic gradient identification methods [J]. Control Theory and Application,2003,20(6):870-874
[40] 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
[41] Liu Y J, Yu L, Ding F. Multi-innovation extended stochastic gradient algorithm and its performance analysis [J]. Circuits, Systems and Signal Processing, 2010, 29 (4):649-667
[42] Ding F. Several multi-innovation identification methods [J].Digital Signal Processing,2010,20(4):1027-1039
[43] 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

Hierarchical multi-innovation identification methods for multivariable equation-error-like type systems

DING Feng[1,2,3] WANG Yanjiao[1]
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 According to the hierarchical identification principle,this paper presents the hierarchical stochastic gradient algorithms and the hierarchical gradient based iterative algorithms, the hierarchical least squares algorithms and the hierarchical least squares based iterative algorithms for multivariable equation-error-like systems and multivariable equation-error ARMA-like systems, and further derives the hierarchical multi-innovation gradient algorithms and the hierarchical multi-innovation least squares algorithms. In order to reduce computational burdens, this paper derives the filtering based hierarchical identification algorithms and the filtering based hierarchical multi-innovation identification algorithms for multivariable equation-error ARMA-like systems using the filtering technique. Finally, the computational efficiency and the computational steps of some typical identification algorithms are discussed.

Key words parameter estimation; recursive identification; gradient search; least squares search; multi-innovation identification theory; hierarchical identification principle; multivariable-like system; data filtering technique