文章编号:1674-7070(2015)01-0001-23

丁锋[1,2,3]　毛亚文[1]

输入非线性方程误差自回归系统的多新息辨识方法

摘要

典型块结构非线性系统包括基本的输入非线性系统、输出非线性系统、输入输出非线性系统、反馈非线性系统等.输入非线性系统包括输入非线性方程误差类系统和输入非线性输出误差类系统.以输入非线性方程误差自回归系统,即输入非线性受控自回归自回归(IN-CARAR)系统为例,分别基于过参数化模型,基于关键项分离原理,基于数据滤波技术以及基于辨识模型分解技术,研究和提出了IN-CARAR系统的随机梯度辨识方法、多新息随机梯度辨识方法、递推最小二乘辨识方法、多新息最小二乘辨识方法.这些方法可以推广到其他输入非线性方程误差系统、输入非线性输出误差类系统、输出非线性方程误差类系统、输出非线性输出类系统、反馈非线性系统等.同时,给出了几个典型辨识算法的计算步骤、流程图和计算量.

关键词

参数估计;递推辨识;梯度搜索;最小二乘;过参数化模型;关键项分离原理;数据滤波技术;模型分解;辅助模型辨识思想;多新息辨识理论;递阶辨识原理;耦合辨识概念;输入非线性系统;输出非线性系统

中图分类号 TP273
文献标志码 A

收稿日期 2015-01-25
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

系统辨识作为现代控制理论的基础,自形成以来得到了空前的发展,并已在自然科学和社会科学领域如物理学、化学、心理学等诸多方面得到了广泛应用.在此背景下,2011—2012年、2014年笔者等已经在《南京信息工程大学学报》连载了系统辨识论文17篇,并出版了著作《系统辨识新论》[1]与《系统辨识——辨识方法性能分析》[2].2014年的连载论文[3-8]研究了多元系统、类多变量系统、多变量系统、状态空间系统的耦合多新息辨识方法、递阶多新息辨识方法等,但它们都是针对线性系统展开的.近年来,出于工程应用和控制发展的需要,非线性系统的建模与辨识已成为国际控制领域的研究难点和热点,本文研究输入非线性系统的多新息辨识方法.

实际工业系统大部分都具有一定的非线性特性,这类系统往往需要用非线性模型建模方能充分表达其工作特性.因此对非线性特性进行建模和精确辨识,已成为高精度控制系统设计的必由之路.由于非线性系统的机理复杂性,以往的多新息系统辨识方法大都是针对线性系统的,而非线性系统多新息辨识的研究在很大程度上被忽略了.

最小二乘算法收敛速度快,但由于需要计算协方差阵,因而计算量较大.随机梯度算法的计算量小,但收敛速度慢.出于改进随机梯度算法收敛速度的目的,本文第1作者引入了新息长度的概念,建立了多新息辨识理论,其主要思想是将标量新息扩展为新息向量,向量新息扩展为新息矩阵,充分使用系统的辨识新息,以提高量测数据的利用效率和参数估计精度[1,9-11].与随机梯度算法相比,多新息随机梯度算法不仅能提高收敛速度,而且能提高算法的参数辨识精度,还避免计算协方差阵,其实质就是最小二乘算法和随机梯度算法在计算量与收敛速度间的一种折中,是"性价比"比较高的一类辨识方法.自从其诞生以来,便受到了人们的青睐,一些多新息辨识成果相继发表在国际著名期刊上.如今,线性系统的多新息辨识得到了较多研究[9-23],但非线性系统多新息辨识[24]、多率采样数据系统多新息辨识[25-26],以及多新息滤波、多新息状态估计、多新息信号处理、多新息故障检测与诊断、多新息自适应控制与自校正控制[27]、多新息预报与多新息预测控制等方面的研究成果还不多见,这给辨识科学家提供了广阔的研究空间.

数据滤波可以借助数字信号处理等滤波技术减少干扰信号在有用信号中的比重,即减小信号的噪信比(noise-to-signal ratio)[1].由于工业控制对象的环境一般比较恶劣,干扰一般为有色噪声(相关噪声).典型的相关噪声包括自回归(AR)过程、滑动平均(MA)过程、自回归滑动平均(ARMA)过程.为了提高参数辨识精度,创新性工作是将数据滤波技术应用于系统辨识中,提出基于输入输出数据滤波(观测数据滤波)的辨识方法.系统辨识中滤波的主要思想是用一个滤波器(通常取作噪声模型的传递函数)对系统输入输出数据进行滤波,这种滤波不会改变系统输入输出信号的噪信比,但可以改变系统辨识模型的结构,从而提高辨识精度.由于噪声模型是未知的,导致滤波无法进行,解决方法是采用交互估计理论,同时估计噪声模型的参数,用噪声模型的估计值进行滤波,从而实现递推参数估计或迭代参数估计.

在基于数据滤波的辨识方面,本文第1作者在总结前人有关滤波技术与辨识技术的研究成果基础上,开辟和提出了基于数据滤波的辨识理论与方法,首先将这一方法写入尚未出版的系统辨识中文著作中,随后授权于作者的访问学者王冬青博士,将其中“基于数据滤波的 CARARMA 系统递推最小二乘辨识方法”整理翻译写成英文论文[28],及时公布到国际控制领域.最近一些基于滤波的辨识方法陆续出现在国际期刊上,如输出误差滑动平均系统的基于数据滤波的辨识方法[29]、Hammerstein 非线性系统的基于数据滤波的递推最小二乘算法[30]、基于数据滤波的一类线性参数系统的梯度辨识方法[31]、状态空间系统的滤波辨识算法[8,32]等,这使得基于数据滤波的辨识方法[28,33]与梯度辨识、最小二乘辨识、牛顿辨识[34]、极大似然辨识、Bayes 辨识、辅助模型辨识[35-36]、多新息辨识[9-10]、递阶辨识[37-42]、耦合辨识[43-44]等方法齐名,是线性系统、非线性系统辨识方法的重要补充.这些方法也可以加以组合来研究线性系统与非线性系统的辨识问题,如输出误差滑动平均系统的辅助模型多新息随机梯度辨识方法[17]、稀少量测数据系统的辅助模型多新息随机梯度辨识方法[23]、Hammerstein 非线性受控自回归系统的递阶多新息辨识方法[24]、基于滤波与辅助模型的输出误差自回归系统最小二乘参数辨识方法[33]、类多变量输出误差系统的耦合多新息辨识方法[5]、类多变量方程误差类系统的递阶多新息辨识方法[7]等.最近,作者指导的硕士生毛亚文研究了基于数据滤波的 Hammerstein 受控自回归自回归系统多新息随机梯度辨识方法[45].

本文提出了基于过参数化模型、基于关键项分离原理、基于数据滤波技术、基于辨识模型分解的输入非线性方程误差自回归系统(即输入非线性受控自回归自回归系统)的多新息辨识方法.这些方法可以推广到其他输入非线性方程误差系统、输入非线性输出误差类系统、输出非线性方程误差类系统、输出非线性输出误差类系统.

1 基于过参数化模型的多新息辨识方法

1.1 系统描述与辨识模型

输入非线性系统是指一类静态非线性环节串联一个动态线性子系统构成的块结构非线性系统,其结构简单,且广泛存在于实际工程中.

1) 当动态子系统是一个有限脉冲响应(FIR)模型时,就得到一个输入非线性有限脉冲响应(IN-FIR)系统;

2) 当动态子系统是一个受控自回归(CAR)模型时,就得到一个输入非线性受控自回归(IN-CAR)系统(或称输入非线性方程误差系统);

3) 当动态子系统是一个受控自回归滑动平均(CARMA)模型时,就得到一个输入非线性受控自回归滑动平均(IN-CARMA)系统(或称输入非线性方程误差滑动平均系统);

4) 当动态子系统是一个受控自回归自回归(CARAR)模型时,就得到一个输入非线性受控自回归自回归(IN-CARAR)系统(或称输入非线性方程误差自回归系统);

5) 当动态子系统是一个受控自回归自回归滑动平均(CARARMA)模型时,就得到一个输入非线性受控自回归自回归滑动平均(IN-CARARMA)系统(或称输入非线性方程误差自回归滑动平均系统).

本文以 IN-CARAR 系统为例,研究输入非线性方程误差类系统的多新息辨识方法,这些方法可以推广到其他输入非线性系统和输出非线性系统中.

考虑输入非线性受控自回归自回归模型(Input Nonlinear CARAR model,IN-CARAR 模型)描述的系统,其结构如图1所示,输入输出关系表达如下:

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

其中 $u(t)$ 和 $y(t)$ 分别为系统的输入和输出,$v(t)$ 是均值为零的白噪声,非线性块输出 $\bar{u}(t)$ 是系数为

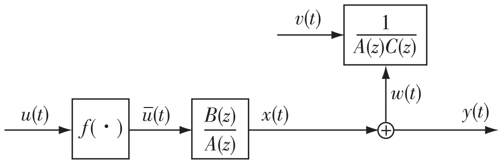

图 1　输入非线性受控自回归自回归系统

Fig. 1　An input nonlinear controlled autoregressive autoregressive system

$(\gamma_1,\gamma_2,\cdots,\gamma_{n_\gamma})$ 的已知非线性基函数 $\boldsymbol{f}:=(f_1,f_2,\cdots,f_{n_\gamma})$ 的线性组合:

$$\bar{u}(t)=f(u(t))=\gamma_1f_1(u(t))+\gamma_2f_2(u(t))+\cdots+\gamma_{n_\gamma}f_{n_\gamma}(u(t))=\boldsymbol{f}(u(t))\boldsymbol{\gamma},\tag{2}$$

$\boldsymbol{f}(u(t)):=[f_1(u(t)),f_2(u(t)),\cdots,f_{n_\gamma}(u(t))]\in\mathbf{R}^{1\times n_\gamma}$ 是基函数构成的行向量,$\boldsymbol{\gamma}:=[\gamma_1,\gamma_2,\cdots,\gamma_{n_\gamma}]^{\mathrm{T}}\in\mathbf{R}^{n_\gamma}$ 是非线性部分的参数向量,$A(z)$,$B(z)$ 和 $C(z)$ 是单位后移算子 $z^{-1}[z^{-1}y(t)=y(t-1)]$ 的常系数时不变多项式:

$$A(z):=1+a_1z^{-1}+a_2z^{-2}+\cdots+a_{n_a}z^{-n_a},$$
$$B(z):=b_0+b_1z^{-1}+b_2z^{-2}+\cdots+b_{n_b}z^{-n_b},$$
$$C(z):=1+c_1z^{-1}+c_2z^{-2}+\cdots+c_{n_c}z^{-n_c}.$$

设阶次 n_a,n_b 和 n_c 已知,且当 $t\leqslant0$ 时,$u(t)=0$,$y(t)=0$,$v(t)=0$.

定义中间噪声变量:

$$w(t):=\frac{1}{C(z)}v(t)=[1-C(z)]w(t)+v(t)=-\sum_{i=1}^{n_c}c_iw(t-i)+v(t).$$

定义参数向量

$$\boldsymbol{\theta}:=[\boldsymbol{a}^{\mathrm{T}},b_0\boldsymbol{\gamma}^{\mathrm{T}},b_1\boldsymbol{\gamma}^{\mathrm{T}},\cdots,b_{n_b}\boldsymbol{\gamma}^{\mathrm{T}},\boldsymbol{c}^{\mathrm{T}}]^{\mathrm{T}}\in\mathbf{R}^{n},$$
$$n:=n_a+(n_b+1)n_\gamma+n_c,$$
$$\boldsymbol{a}:=[a_1,a_2,\cdots,a_{n_a}]^{\mathrm{T}}\in\mathbf{R}^{n_a},$$
$$\boldsymbol{b}:=[b_1,b_2,\cdots,b_{n_b}]^{\mathrm{T}}\in\mathbf{R}^{n_b},$$
$$\boldsymbol{c}:=[c_1,c_2,\cdots,c_{n_c}]^{\mathrm{T}}\in\mathbf{R}^{n_c},$$
$$\boldsymbol{\gamma}:=[\gamma_1,\gamma_2,\cdots,\gamma_{n_\gamma}]^{\mathrm{T}}\in\mathbf{R}^{n_\gamma}$$

和信息向量

$$\boldsymbol{\varphi}(t):=\begin{bmatrix}\boldsymbol{\varphi}_{\mathrm{s}}(t)\\ \boldsymbol{\varphi}_{\mathrm{n}}(t)\end{bmatrix}\in\mathbf{R}^{n},$$
$$\boldsymbol{\varphi}_{\mathrm{s}}(t):=[-y(t-1),-y(t-2),\cdots,-y(t-n_a),\boldsymbol{\psi}_0^{\mathrm{T}}(t),\boldsymbol{\psi}_1^{\mathrm{T}}(t),\cdots,\boldsymbol{\psi}_{n_b}^{\mathrm{T}}(t)]^{\mathrm{T}}\in\mathbf{R}^{n_a+(n_b+1)n_\gamma},$$
$$\boldsymbol{\psi}_j(t):=[f_1(u(t-j)),f_2(u(t-j)),\cdots,f_{n_\gamma}(u(t-j))]^{\mathrm{T}}\in\mathbf{R}^{n_\gamma},$$
$$\boldsymbol{\varphi}_{\mathrm{n}}(t):=[-w(t-1),-w(t-2),\cdots,-w(t-n_c)]^{\mathrm{T}}\in\mathbf{R}^{n_c},$$

系统中的相关干扰噪声 $w(t)$ 是一个自回归(AR)过程.

将式(2)和(3)代入(1)得到辨识模型

$$\begin{aligned}y(t)&=[1-A(z)]y(t)+B(z)\bar{u}(t)+w(t)=\\&-\sum_{i=1}^{n_a}a_iy(t-i)+\sum_{i=1}^{n_\gamma}[b_0\gamma_if_i(u(t))+b_1\gamma_if_i(u(t-1))+\cdots+b_{n_b}\gamma_if_i(u(t-n_b))]+w(t)=\qquad(4)\\&-\sum_{i=1}^{n_a}a_iy(t-i)+\sum_{i=1}^{n_\gamma}[b_0\gamma_if_i(u(t))+b_1\gamma_if_i(u(t-1))+\cdots+b_{n_b}\gamma_if_i(u(t-n_b))]-\sum_{i=1}^{n_c}c_iw(t-i)+v(t)=\\&\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta}+v(t).\qquad(5)\end{aligned}$$

注 1　这就是输入非线性 CARAR 系统的过参数化辨识模型.该模型的参数向量 $\boldsymbol{\theta}$ 包含 $n=n_a+(n_b+1)n_\gamma+n_c$ 个参数,它大于系统的实际参数数目 $n_a+n_b+1+n_c+n_r$(当 $n_b,n_\gamma\geqslant2$ 时).这也是过参数化模型名称的来历.

注 2　过参数化辨识模型(5)是一个双线性参数模型,模型中出现了 2 个参数集$\{b_i\}$与$\{\gamma_j\}$的乘积项$\{b_i\gamma_j\}$,故参数是不可辨识的.为了得到唯一的参数估计,需要规范化系统参数.基本的规范化方法有 3 种:

1) 固定 b_i 中的一个,或者固定 γ_j 中的一个,如 $b_0=1$ 或 $\gamma_1=1$;

2) 固定$(b_0,b_1,b_2,\cdots,b_{n_b})$的模为 1,或$(\gamma_1,\gamma_2,\cdots,\gamma_{n_\gamma})$的模为 1,如 $b_0^2+b_1^2+b_2^2+\cdots+b_{n_b}^2=1$,或 $\|\boldsymbol{\gamma}\|^2:=\gamma_1^2+\gamma_2^2+\cdots+\gamma_{n_\gamma}^2=1$;

3) 固定线性动态子系统的增益,如 $G(1):=\dfrac{B(1)}{A(1)}=b_0+b_1+b_2+\cdots+b_{n_b}=1$.

注 3　因为过参数化辨识模型(5)的参数向量 $\boldsymbol{\theta}$ 包含了原系统参数的乘积项$\{b_i\gamma_j\}$,故参数数目比系统的实际参数数目多,而在过参数化模型的辨识算法中,把这些乘积参数作为独立的参数进行辨识,因此辨识算法的计算量变大,特别当阶次 n_b,n_γ 很大时,最小二乘辨识算法的计算量大更为明显,且当辨识出乘积参数后,还需附加计算从参数估计中分离出原系统的参数估计.

注 4　基于过参数化模型的辨识算法需要从获得的过参数化估计中分离出原系统的参数估计.分离参数的方法有 SVD 方法、平均值方法等.

设 $\boldsymbol{\theta},\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{\gamma}$ 在时刻 t 的估计分别为

$$\hat{\boldsymbol{\theta}}(t):=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\ \widehat{b_0\boldsymbol{\gamma}}^{\mathrm{T}}(t),\ \widehat{b_1\boldsymbol{\gamma}}^{\mathrm{T}}(t),\cdots,\ \widehat{b_{n_b}\boldsymbol{\gamma}}^{\mathrm{T}}(t),\hat{\boldsymbol{c}}^{\mathrm{T}}(t)]^{\mathrm{T}}\in\mathbf{R}^{n},$$

$$\hat{\boldsymbol{a}}(t):=[\hat{a}_1(t),\hat{a}_2(t),\cdots,\hat{a}_{n_a}(t)]^{\mathrm{T}}\in\mathbf{R}^{n_a},$$

$$\hat{\boldsymbol{b}}(t):=[\hat{b}_1(t),\hat{b}_2(t),\cdots,\hat{b}_{n_b}(t)]^{\mathrm{T}}\in\mathbf{R}^{n_b},$$

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

$$\hat{\boldsymbol{\gamma}}(t):=[\hat{\gamma}_1(t),\hat{\gamma}_2(t),\cdots,\hat{\gamma}_{n_\gamma}(t)]^{\mathrm{T}}\in\mathbf{R}^{n_\gamma}.$$

有时为表达简洁,把估计中的“(t)”省略,比如用 $\hat{\boldsymbol{\theta}}$ 表示 $\hat{\boldsymbol{\theta}}(t)$.

这里假设 $B(z)$ 的第 1 个系数是 1,即 $b_0=1$.在此条件下,$\boldsymbol{a},\boldsymbol{c}$ 和 $\boldsymbol{\gamma}$ 的估计可以直接从 $\hat{\boldsymbol{\theta}}$ 中读出来.令 $\hat{\theta}_i$ 是 $\hat{\boldsymbol{\theta}}$ 的第 i 个元素,根据 $\boldsymbol{\theta}$ 的定义式,可知 $b_j(j=1,2,\cdots,n_b)$ 的估计可以用下式计算[46-48]:

$$\hat{b}_j=\frac{\hat{\theta}_{n_a+jn_\gamma+i}}{\hat{\gamma}_i},\quad j=1,2,3,\cdots,n_b,\quad i=1,2,\cdots,n_\gamma.$$

从上式可以看出,计算每一个系数 $\hat{b}_j$ 都出现许多冗余,因为对每一个 $\hat{b}_j(j=1,2,3,\cdots,n_b)$,有 n_γ 个估计.这里采取的措施是取这些估计的平均值作为 b_j 的估计,即

$$\hat{b}_j=\frac{1}{n_\gamma}\sum_{i=1}^{n_\gamma}\frac{\hat{\theta}_{n_a+jn_\gamma+i}}{\hat{\gamma}_i},\quad j=1,2,3,\cdots,n_b.$$

下面简单给出过参数化辨识模型(5)的广义随机梯度辨识方法、多新息广义随机梯度辨识方法、递推广义最小二乘辨识方法、多新息广义最小二乘辨识方法.

1.2 广义随机梯度辨识方法

对于辨识模型(5),使用负梯度搜索,极小化准则函数

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

可得下列递推关系[1]:

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\boldsymbol{\varphi}(t)}{r(t)}e(t),\tag{6}$$

$$e(t)=y(t)-\boldsymbol{\varphi}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1),\tag{7}$$

$$r(t)=r(t-1)+\|\boldsymbol{\varphi}(t)\|^2,\quad r(0)=1.\tag{8}$$

从信息向量 $\boldsymbol{\varphi}(t)$ 的定义可以看出,$\boldsymbol{\varphi}(t)$ 中包含了未知中间变量 $w(t-i)$,因此上述算法无法实现.解决办法是借助于辅助模型辨识思想[1-2,35],使用估计 $\hat{w}(t-i)$ 代替 $\boldsymbol{\varphi}(t)$ 中的未知中间变量 $w(t-i)$,代替后的信息向量记作

$$\hat{\boldsymbol{\varphi}}(t):=\begin{bmatrix}\boldsymbol{\varphi}_{\mathrm{s}}(t)\\ \hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\end{bmatrix}\in\mathbf{R}^{n},$$

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

用 $\hat{\boldsymbol{\varphi}}(t)$ 代替式(6)—(8)中的未知 $\boldsymbol{\varphi}(t)$,可以得到估计参数向量 $\boldsymbol{\theta}$ 的广义随机梯度算法(Generalized Stochastic Gradient algorithm,GSG 算法):

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\hat{\boldsymbol{\varphi}}(t)}{r(t)}e(t),\tag{9}$$

$$e(t)=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1),\tag{10}$$

$$r(t)=r(t-1)+\|\hat{\boldsymbol{\varphi}}(t)\|^2,\quad r(0)=1,\tag{11}$$

$$\hat{\boldsymbol{\varphi}}(t)=\begin{bmatrix}\boldsymbol{\varphi}_{\mathrm{s}}(t)\\ \hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\end{bmatrix},\tag{12}$$

$$\boldsymbol{\varphi}_{\mathrm{s}}(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a),\boldsymbol{\psi}_0^{\mathrm{T}}(t),\boldsymbol{\psi}_1^{\mathrm{T}}(t),\cdots,\boldsymbol{\psi}_{n_b}^{\mathrm{T}}(t)]^{\mathrm{T}},\tag{13}$$

$$\boldsymbol{\psi}_j(t)=[f_1(u(t-j)),f_2(u(t-j)),\cdots,f_{n_\gamma}(u(t-j))]^{\mathrm{T}},\tag{14}$$

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

$$\hat{w}(t)=y(t)-\boldsymbol{\varphi}_{\mathrm{s}}^{\mathrm{T}}(t)[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\ \widehat{b_0\boldsymbol{\gamma}}^{\mathrm{T}}(t),\ \widehat{b_1\boldsymbol{\gamma}}^{\mathrm{T}}(t),\cdots,\ \widehat{b_{n_b}\boldsymbol{\gamma}}^{\mathrm{T}}(t)]^{\mathrm{T}},\tag{16}$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\ \widehat{b_0\boldsymbol{\gamma}}^{\mathrm{T}}(t),\ \widehat{b_1\boldsymbol{\gamma}}^{\mathrm{T}}(t),\cdots,\ \widehat{b_{n_b}\boldsymbol{\gamma}}^{\mathrm{T}}(t),\hat{\boldsymbol{c}}^{\mathrm{T}}(t)]^{\mathrm{T}}.\tag{17}$$

注 5 随机梯度辨识算法计算量小,但存在参数估计收敛速度慢的问题.为提高参数估计精度,可引入遗忘因子(forgetting factor)λ,将式(11)修改为

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

也可引入新息长度,导出下列多新息广义随机梯度辨识算法.

1.3 多新息广义随机梯度辨识方法

参考文献[1,9]中多新息辨识理论,为了提高随机梯度算法的参数估计收敛速度,将式(10)中标量新息 $e(t)\in\mathbf{R}$ 扩展为新息向量

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

其中正整数 p 表征新息长度,且

$$e(t-i)=y(t-i)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t-i)\hat{\boldsymbol{\theta}}(t-i-1).\tag{18}$$

假设时刻$(t-1)$的参数估计 $\hat{\boldsymbol{\theta}}(t-1)$ 比之前时刻 $(t-i)(i=2,3,4,\cdots,p-1)$ 的估计 $\hat{\boldsymbol{\theta}}(t-i)$ 更接近真参数向量 $\boldsymbol{\theta}$.因此,新息向量可以取为

$$\boldsymbol{E}(p,t)=\begin{bmatrix}y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)\\ y(t-1)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t-1)\hat{\boldsymbol{\theta}}(t-1)\\ \vdots\\ y(t-p+1)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t-p+1)\hat{\boldsymbol{\theta}}(t-1)\end{bmatrix}\in\mathbf{R}^{p}.\tag{19}$$

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

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

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

则式(19)可以表述为

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

参考文献[1,9]中方法，可以得到辨识 IN-CARAR 系统向量 $\boldsymbol{\theta}$ 的多新息广义随机梯度算法(Multi-Innovation Generalized Stochastic Gradient algorithm，MI-GSG 算法)：

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\hat{\boldsymbol{\Phi}}(p,t)}{r(t)}\boldsymbol{E}(p,t), \tag{20}$$

$$\boldsymbol{E}(p,t)=\boldsymbol{Y}(p,t)-\hat{\boldsymbol{\Phi}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\theta}}(t-1), \tag{21}$$

$$r(t)=r(t-1)+\|\hat{\boldsymbol{\varphi}}(t)\|^2,\quad r(0)=1, \tag{22}$$

$$\boldsymbol{Y}(p,t)=[y(t),y(t-1),\cdots,y(t-p+1)]^{\mathrm{T}}, \tag{23}$$

$$\hat{\boldsymbol{\Phi}}(p,t)=[\hat{\boldsymbol{\varphi}}(t),\hat{\boldsymbol{\varphi}}(t-1),\cdots,\hat{\boldsymbol{\varphi}}(t-p+1)], \tag{24}$$

$$\hat{\boldsymbol{\varphi}}(t)=\begin{bmatrix}\boldsymbol{\varphi}_{\mathrm{s}}(t)\\ \hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\end{bmatrix}, \tag{25}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)=[-\hat{w}(t-1),-\hat{w}(t-2),\cdots,-\hat{w}(t-n_c)]^{\mathrm{T}}, \tag{26}$$

$$\hat{w}(t)=y(t)-\boldsymbol{\varphi}_{\mathrm{s}}^{\mathrm{T}}(t)[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\ \widehat{b_0\boldsymbol{\gamma}}^{\mathrm{T}}(t),\ \widehat{b_1\boldsymbol{\gamma}}^{\mathrm{T}}(t),\ \cdots,\ \widehat{b_{n_b}\boldsymbol{\gamma}}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{27}$$

$$\boldsymbol{\varphi}_{\mathrm{s}}(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a),\ \boldsymbol{\psi}_0^{\mathrm{T}}(t),\boldsymbol{\psi}_1^{\mathrm{T}}(t),\cdots,\boldsymbol{\psi}_{n_b}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{28}$$

$$\boldsymbol{\psi}_j(t)=[f_1(u(t-j)),f_2(u(t-j)),\cdots,f_{n_\gamma}(u(t-j))]^{\mathrm{T}}, \tag{29}$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\ \widehat{b_0\boldsymbol{\gamma}}^{\mathrm{T}}(t),\ \widehat{b_1\boldsymbol{\gamma}}^{\mathrm{T}}(t),\cdots,\ \widehat{b_{n_b}\boldsymbol{\gamma}}^{\mathrm{T}}(t),\ \hat{\boldsymbol{c}}^{\mathrm{T}}(t)]^{\mathrm{T}}. \tag{30}$$

当新息长度 $p=1$ 时，MISG 算法退化为 SG 算法(9)—(17).式(22)中 $r(t)$ 也可修改为

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

注 6 为了使算法获得更快的收敛速度，也可在式(22)中引入遗忘因子 λ：

$$r(t)=\lambda r(t-1)+\|\hat{\boldsymbol{\varphi}}(t)\|^2,\quad 0\leqslant\lambda\leqslant 1,\quad r(0)=1, \tag{31}$$

则式(20)—(21)，(23)—(31)构成多新息遗忘因子广义随机梯度算法，简称多新息遗忘广义随机梯度算法.当新息长度 $p=1$ 时，多新息遗忘因子广义随机梯度算法退化为遗忘因子广义随机梯度算法.在相同的数据长度下，增大 p 能提高参数估计精度.

1.4 递推广义最小二乘辨识方法

对于辨识模型(5)，利用最小二乘原理，极小化准则函数(criterion function)

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

并借助辅助模型辨识思想，即用估计 $\hat{\boldsymbol{\varphi}}(t)$ 代替 $\boldsymbol{\varphi}(t)$，可以得到辨识 IN-CARAR 系统参数向量 $\boldsymbol{\theta}$ 的递推广义最小二乘算法(Recursive Generalized Least Squares algorithm，RGLS 算法)：

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{L}(t)[y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)], \tag{32}$$

$$\boldsymbol{L}(t)=\boldsymbol{P}(t)\hat{\boldsymbol{\varphi}}(t)=\frac{\boldsymbol{P}(t-1)\hat{\boldsymbol{\varphi}}(t)}{1+\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\boldsymbol{P}(t-1)\hat{\boldsymbol{\varphi}}(t)}, \tag{33}$$

$$\boldsymbol{P}(t)=[\boldsymbol{I}_n-\boldsymbol{L}(t)\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)]\boldsymbol{P}(t-1),\quad \boldsymbol{P}(0)=p_0\boldsymbol{I}_n, \tag{34}$$

$$\hat{\boldsymbol{\varphi}}(t)=\begin{bmatrix}\boldsymbol{\varphi}_{\mathrm{s}}(t)\\ \hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\end{bmatrix}, \tag{35}$$

$$\boldsymbol{\varphi}_{\mathrm{s}}(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a),\ \boldsymbol{\psi}_0^{\mathrm{T}}(t),\boldsymbol{\psi}_1^{\mathrm{T}}(t),\cdots,\boldsymbol{\psi}_{n_b}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{36}$$

$$\boldsymbol{\psi}_j(t)=[f_1(u(t-j)),f_2(u(t-j)),\cdots,f_{n_\gamma}(u(t-j))]^{\mathrm{T}}, \tag{37}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)=[-\hat{w}(t-1),-\hat{w}(t-2),\cdots,-\hat{w}(t-n_c)]^{\mathrm{T}}, \tag{38}$$

$$\hat{w}(t)=y(t)-\boldsymbol{\varphi}_{\mathrm{s}}^{\mathrm{T}}(t)[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\ \widehat{b_0\boldsymbol{\gamma}}^{\mathrm{T}}(t),\ \widehat{b_1\boldsymbol{\gamma}}^{\mathrm{T}}(t),\ \cdots,\ \widehat{b_{n_b}\boldsymbol{\gamma}}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{39}$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\ \widehat{b_0\boldsymbol{\gamma}}^{\mathrm{T}}(t),\ \widehat{b_1\boldsymbol{\gamma}}^{\mathrm{T}}(t),\cdots,\ \widehat{b_{n_b}\boldsymbol{\gamma}}^{\mathrm{T}}(t),\ \hat{\boldsymbol{c}}^{\mathrm{T}}(t)]^{\mathrm{T}}. \tag{40}$$

注 7 辨识算法的计算量是评价计算效率的一个重要指标.辨识算法的计算量可用其乘法运算次数和加法运算次数表示.一次加法运算为一次浮点运算(floating point operation)，称为一个 flop，一次乘法运算也为一个 flop.除法作为乘法对待，减法作为加法对待.这样就可以用 flop 数，即浮点运算数来表示计算量的大小.表 1 列出了基于过参数化模型的递推广义最小二乘算法(32)—(40)的计算量，其中 $n=n_a+(n_b+1)n_\gamma+n_c$，$\hat{\boldsymbol{\theta}}_{\mathrm{s}}(t):=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\ \widehat{b_0\boldsymbol{\gamma}}^{\mathrm{T}}(t),\ \widehat{b_1\boldsymbol{\gamma}}^{\mathrm{T}}(t),\cdots,\ \widehat{b_{n_b}\boldsymbol{\gamma}}^{\mathrm{T}}(t)]^{\mathrm{T}}\in\mathbf{R}^{n_a+(n_b+1)n_\gamma}$.

1.5 多新息广义最小二乘辨识方法

同样，将式(32)中标量新息 $e(t):=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\cdot\hat{\boldsymbol{\theta}}(t-1)\in\mathbf{R}$ 扩展为新息向量

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

可以得到辨识 IN-CARAR 系统参数向量 $\boldsymbol{\theta}$ 的多新息广义最小二乘算法(Multi-Innovation Generalized Least Squares algorithm，MI-GLS 算法)：

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{L}(t)[\boldsymbol{Y}(p,t)-\hat{\boldsymbol{\Phi}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\theta}}(t-1)], \tag{41}$$

$$\boldsymbol{L}(t)=\boldsymbol{P}(t-1)\hat{\boldsymbol{\Phi}}(p,t)[\boldsymbol{I}_p+\hat{\boldsymbol{\Phi}}^{\mathrm{T}}(p,t)\boldsymbol{P}(t-1)\hat{\boldsymbol{\Phi}}(p,t)]^{-1}, \tag{42}$$

$$\boldsymbol{P}(t)=\boldsymbol{P}(t-1)-\boldsymbol{L}(t)\hat{\boldsymbol{\Phi}}^{\mathrm{T}}(p,t)\boldsymbol{P}(t-1),\quad \boldsymbol{P}(0)=p_0\boldsymbol{I}_n, \tag{43}$$

$$\boldsymbol{Y}(p,t)=[y(t),y(t-1),\cdots,y(t-p+1)]^{\mathrm{T}}, \tag{44}$$

$$\hat{\boldsymbol{\Phi}}(p,t)=[\hat{\boldsymbol{\varphi}}(t),\hat{\boldsymbol{\varphi}}(t-1),\cdots,\hat{\boldsymbol{\varphi}}(t-p+1)], \tag{45}$$

表 1　RGLS 算法的计算量

Table 1　The computational efficiency of the RGLS algorithm

变量	表达式	乘法次数	加法次数
$\hat{\boldsymbol{\theta}}(t)$	$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{L}(t)e(t)\in\mathbf{R}^n$	n	n
	$e(t):=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)\in\mathbf{R}$	n	n
$\boldsymbol{L}(t)$	$\boldsymbol{L}(t)=\boldsymbol{\zeta}(t)/[1+\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\boldsymbol{\zeta}(t)]\in\mathbf{R}^n$	$2n$	n
	$\boldsymbol{\zeta}(t):=\boldsymbol{P}(t-1)\hat{\boldsymbol{\varphi}}(t)\in\mathbf{R}^n$	n^2	n^2-n
$\boldsymbol{P}(t)$	$\boldsymbol{P}(t)=\boldsymbol{P}(t-1)-\boldsymbol{L}(t)\boldsymbol{\zeta}^{\mathrm{T}}(t)\in\mathbf{R}^{n\times n}$	n^2	n^2
$\hat{w}(t)$	$\hat{w}(t)=y(t)-\boldsymbol{\varphi}_{\mathrm{s}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_{\mathrm{s}}(t)\in\mathbf{R}$	$n_a+(n_b+1)n_\gamma$	$n_a+(n_b+1)n_\gamma$
总数		$2n^2+4n+n_a+n_bn_\gamma+n_\gamma$	$2n^2+2n+n_a+n_bn_\gamma+n_\gamma$
总 flop 数		$4n^2+6n+2n_a+2n_bn_\gamma+2n_\gamma$	

$$\hat{\boldsymbol{\varphi}}(t)=\begin{bmatrix}\boldsymbol{\varphi}_{\mathrm{s}}(t)\\ \hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\end{bmatrix},\tag{46}$$

$$\begin{aligned}\boldsymbol{\varphi}_{\mathrm{s}}(t)=[&-y(t-1),-y(t-2),\cdots,-y(t-n_a),\\&\boldsymbol{\psi}_0^{\mathrm{T}}(t),\boldsymbol{\psi}_1^{\mathrm{T}}(t),\cdots,\boldsymbol{\psi}_{n_b}^{\mathrm{T}}(t)]^{\mathrm{T}},\end{aligned}\tag{47}$$

$$\boldsymbol{\psi}_j(t)=[f_1(u(t-j)),f_2(u(t-j)),\cdots,f_{n_\gamma}(u(t-j))]^{\mathrm{T}},\tag{48}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)=[-\hat{w}(t-1),-\hat{w}(t-2),\cdots,-\hat{w}(t-n_c)]^{\mathrm{T}},\tag{49}$$

$$\begin{aligned}\hat{w}(t)=y(t)-\boldsymbol{\varphi}_{\mathrm{s}}^{\mathrm{T}}(t)[&\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\ \widehat{b_0\boldsymbol{\gamma}}^{\mathrm{T}}(t),\ \widehat{b_1\boldsymbol{\gamma}}^{\mathrm{T}}(t),\\&\cdots,\ \widehat{b_{n_b}\boldsymbol{\gamma}}^{\mathrm{T}}(t)]^{\mathrm{T}},\end{aligned}\tag{50}$$

$$\begin{aligned}\hat{\boldsymbol{\theta}}(t)=[&\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\ \widehat{b_0\boldsymbol{\gamma}}^{\mathrm{T}}(t),\ \widehat{b_1\boldsymbol{\gamma}}^{\mathrm{T}}(t),\cdots,\ \widehat{b_{n_b}\boldsymbol{\gamma}}^{\mathrm{T}}(t),\\&\hat{\boldsymbol{c}}^{\mathrm{T}}(t)]^{\mathrm{T}}.\end{aligned}\tag{51}$$

当新息长度 $p=1$ 时，MI-GLS 算法退化为 RGLS 算法(32)—(40).

注 8　值得注意的是，在相同数据长度下，随机梯度算法比最小二乘算法计算出的参数估计精度低，这是因为 2 个算法虽然使用了同样的一批数据，但随机梯度算法的数据利用率低，即从测量数据中提取的信息量少于递推最小二乘算法. 多新息随机梯度算法重复使用了可测的输入输出数据，提高了数据的使用效率，因而能提高参数估计精度. 相对而言，由于最小二乘算法的收敛速度本身就很高，所以多新息最小二乘算法对参数估计精度的改进是有限的，往往在数据缺失情况下，采用变递推间隔时，才显示出多新息最小二乘算法的优越性[1].

注 9　过参数化模型待辨识的参数比系统实际参数数目多，因而基于过参数化模型辨识算法的计算量大(指同类算法间的比较). 为减小计算量，下面研究基于关键项分离的辨识方法.

2　基于关键项分离的多新息辨识方法

关键项分离原理是 Vörös 提出的，类似于笔者提出的辅助模型辨识思想. 基于关键项分离的辨识方法相当于辅助模型辨识思想在非线性系统中的应用. 关键项就是系统中的未知变量，而关键项的估算就是利用辅助模型计算的. 在辨识算法中，信息向量中的未知关键项是用辅助模型的输出代替的，也可用其前一时刻的估计代替.

对输入非线性系统而言，关键项分离是将非线性环节当前时刻的输出作为关键项，将系统输出表示为关键项与余项之和，再利用非线性环节方程，从而使系统输出可以表示为所有待估参数的线性组合形式，因此基于辅助模型的辨识方法可以应用. 与过参数化辨识模型相比，基于关键项分离的辨识模型可以避免产生冗余参数，从而减小辨识算法的计算量.

下面基于关键项分离原理，利用辅助模型辨识思想来估算关键项，研究输入非线性 CARAR 系统的广义随机梯度辨识方法、多新息广义随机梯度辨识方法、递推广义最小二乘辨识方法、多新息广义最小二乘辨识方法.

2.1　基于关键项分离的广义随机梯度辨识方法

考虑输入非线性受控自回归自回归系统(1)，重写如下：

$$A(z)y(t)=B(z)\bar{u}(t)+\frac{1}{C(z)}v(t),\tag{52}$$

$$\begin{aligned}\bar{u}(t)=&f(u(t))=\\&\gamma_1f_1(u(t))+\gamma_2f_2(u(t))+\cdots+\gamma_{n_\gamma}f_{n_\gamma}(u(t))=\\&\boldsymbol{f}(u(t))\boldsymbol{\gamma},\end{aligned}\tag{53}$$

$$w(t)=\frac{1}{C(z)}v(t).\tag{54}$$

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

$$\boldsymbol{\theta}:=\begin{bmatrix}\boldsymbol{a}\\ \boldsymbol{b}\\ \boldsymbol{\gamma}\\ \boldsymbol{c}\end{bmatrix}\in\mathbf{R}^{n_0},\quad n_0:=n_a+n_b+n_\gamma+n_c,$$

$$\boldsymbol{\varphi}(t):=\begin{bmatrix}\boldsymbol{\varphi}_{\rm s}(t)\\ \boldsymbol{\varphi}_{\rm n}(t)\end{bmatrix}\in\mathbf{R}^{n_0},$$

$$\boldsymbol{\varphi}_{\rm s}(t):=[-y(t-1),-y(t-2),\cdots,-y(t-n_a),\bar{u}(t-1),\bar{u}(t-2),\cdots,\bar{u}(t-n_b),\boldsymbol{f}(u(t))]^{\rm T}\in\mathbf{R}^{n_a+n_b+n_\gamma},$$

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

其中参数向量 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 和 $\boldsymbol{\gamma}$ 的定义同上.

利用式(54)，从式(52)可得

$$y(t)=[1-A(z)]y(t)+(b_0+b_1z^{-1}+b_2z^{-2}+\cdots+b_{n_b}z^{-n_b})\bar{u}(t)+w(t)=[1-A(z)]y(t)+b_0\bar{u}(t)+b_1\bar{u}(t-1)+\cdots+b_{n_b}\bar{u}(t-n_b)+w(t). \quad (55)$$

同样假设 $B(z)$ 的第 1 个系数 $b_0=1$，则式(55)可以写为

$$y(t)=[1-A(z)]y(t)+\bar{u}(t)+b_1\bar{u}(t-1)+\cdots+b_{n_b}\bar{u}(t-n_b)+w(t)=-\sum_{i=1}^{n_a}a_iy(t-i)+\sum_{i=1}^{n_b}b_i\bar{u}(t-i)+\bar{u}(t)-\sum_{i=1}^{n_c}c_iw(t-i)+v(t).$$

根据关键项分离原理，将式(53)中的 $\bar{u}(t)$ 代入上式，得到辨识模型：

$$y(t)=-\sum_{i=1}^{n_a}a_iy(t-i)+\sum_{i=1}^{n_b}b_i\bar{u}(t-i)+\boldsymbol{f}(u(t))\boldsymbol{\gamma}-\sum_{i=1}^{n_c}c_iw(t-i)+v(t)=\boldsymbol{\varphi}^{\rm T}(t)\boldsymbol{\theta}+v(t). \quad (56)$$

借助辅助模型思想，信息向量 $\boldsymbol{\varphi}(t)$ 中未知中间变量 $\bar{u}(t-i)$ 和 $w(t-i)$ 用辅助模型(62)—(63)的输出 $\hat{\bar{u}}(t-i)$ 和 $\hat{w}(t-i)$ 代替，可以得到辨识 IN-CARAR 系统参数向量 $\boldsymbol{\theta}$ 的基于关键项分离的广义随机梯度算法(Key Term separation based Generalized Stochastic Gradient algorithm，KT-GSG 算法)：

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\hat{\boldsymbol{\varphi}}(t)}{r(t)}e(t), \quad (57)$$

$$e(t)=y(t)-\hat{\boldsymbol{\varphi}}^{\rm T}(t)\hat{\boldsymbol{\theta}}(t-1), \quad (58)$$

$$r(t)=r(t-1)+\|\hat{\boldsymbol{\varphi}}(t)\|^2,\quad r(0)=1, \quad (59)$$

$$\hat{\boldsymbol{\varphi}}(t)=[\hat{\boldsymbol{\varphi}}_{\rm s}^{\rm T}(t),-\hat{w}(t-1),-\hat{w}(t-2),\cdots,-\hat{w}(t-n_c)]^{\rm T}, \quad (60)$$

$$\hat{\boldsymbol{\varphi}}_{\rm s}(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a),\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_b),\boldsymbol{f}(u(t))]^{\rm T}, \quad (61)$$

$$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t), \quad (62)$$

$$\hat{w}(t)=y(t)-\hat{\boldsymbol{\varphi}}_{\rm s}^{\rm T}(t)[\hat{\boldsymbol{a}}^{\rm T}(t),\hat{\boldsymbol{b}}^{\rm T}(t),\hat{\boldsymbol{\gamma}}^{\rm T}(t)]^{\rm T}, \quad (63)$$

$$\boldsymbol{f}(u(t))=[f_1(u(t)),f_2(u(t)),\cdots,f_{n_\gamma}(u(t))], \quad (64)$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{a}}^{\rm T}(t),\hat{\boldsymbol{b}}^{\rm T}(t),\hat{\boldsymbol{\gamma}}^{\rm T}(t),\hat{\boldsymbol{c}}^{\rm T}(t)]^{\rm T}. \quad (65)$$

注 10 式(62)可以看作是计算关键项估计 $\hat{\bar{u}}(t)$ 的辅助模型，式(63)可以看作是计算噪声估计 $\hat{w}(t)$ 的辅助模型.同样，式(59)中可引入遗忘因子来改进参数估计精度.其他算法也可引入遗忘因子来改进参数估计精度，故不再一一说明.

2.2 基于关键项分离的多新息广义随机梯度辨识方法

令新息长度为 p，定义堆积信息矩阵 $\hat{\boldsymbol{\Phi}}(p,t)$ 和堆积输出向量 $\boldsymbol{Y}(p,t)$ 为

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

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

参考多新息广义随机梯度算法(20)—(30)的推导，将式(58)中标量新息 $e(t)\in\mathbf{R}$ 扩展为新息向量 $\boldsymbol{E}(p,t):=\boldsymbol{Y}(p,t)-\hat{\boldsymbol{\Phi}}^{\rm T}(p,t)\hat{\boldsymbol{\theta}}(t-1)\in\mathbf{R}^p$，可以得到辨识 IN-CARAR 系统参数向量 $\boldsymbol{\theta}$ 的基于关键项分离的多新息广义随机梯度算法(Key Term separation based Multi-Innovation Generalized Stochastic Gradient algorithm，KT-MI-GSG 算法)：

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\hat{\boldsymbol{\Phi}}(p,t)}{r(t)}\boldsymbol{E}(p,t), \quad (66)$$

$$\boldsymbol{E}(p,t)=\boldsymbol{Y}(p,t)-\hat{\boldsymbol{\Phi}}^{\rm T}(p,t)\hat{\boldsymbol{\theta}}(t-1), \quad (67)$$

$$r(t)=r(t-1)+\|\hat{\boldsymbol{\varphi}}(t)\|^2,\quad r(0)=1, \quad (68)$$

$$\boldsymbol{Y}(p,t)=[y(t),y(t-1),\cdots,y(t-p+1)]^{\rm T}, \quad (69)$$

$$\hat{\boldsymbol{\Phi}}(p,t)=[\hat{\boldsymbol{\varphi}}(t),\hat{\boldsymbol{\varphi}}(t-1),\cdots,\hat{\boldsymbol{\varphi}}(t-p+1)], \quad (70)$$

$$\hat{\boldsymbol{\varphi}}(t)=[\hat{\boldsymbol{\varphi}}_{\rm s}^{\rm T}(t),-\hat{w}(t-1),-\hat{w}(t-2),\cdots,-\hat{w}(t-n_c)]^{\rm T}, \quad (71)$$

$$\hat{\boldsymbol{\varphi}}_{\rm s}(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a),\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_b),\boldsymbol{f}(u(t))]^{\rm T}, \quad (72)$$

$$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t), \quad (73)$$

$$\hat{w}(t)=y(t)-\hat{\boldsymbol{\varphi}}_{\rm s}^{\rm T}(t)[\hat{\boldsymbol{a}}^{\rm T}(t),\hat{\boldsymbol{b}}^{\rm T}(t),\hat{\boldsymbol{\gamma}}^{\rm T}(t)]^{\rm T}, \quad (74)$$

$$\boldsymbol{f}(u(t))=[f_1(u(t)),f_2(u(t)),\cdots,f_{n_\gamma}(u(t))], \quad (75)$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{a}}^{\rm T}(t),\hat{\boldsymbol{b}}^{\rm T}(t),\hat{\boldsymbol{\gamma}}^{\rm T}(t),\hat{\boldsymbol{c}}^{\rm T}(t)]^{\rm T}. \quad (76)$$

当新息长度 $p=1$ 时，KT-MI-GSG 算法退化为 KT-GSG 算法(57)—(65).

2.3 基于关键项分离的递推广义最小二乘辨识方法

对辨识模型(56)，利用最小二乘原理，极小化准则函数

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

并借助辅助模型思想，信息向量 $\boldsymbol{\varphi}(t)$ 中未知中间变量 $\bar{u}(t-i)$ 和 $w(t-i)$ 用其估计 $\hat{\bar{u}}(t-i)$ 和 $\hat{w}(t-i)$ 代替，可以得到辨识 IN-CARAR 系统参数向量 $\boldsymbol{\theta}$ 的基于关

键项分离的递推广义最小二乘算法（Key Term separation based Recursive Generalized Least Squares algorithm, KT-RGLS 算法）：

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{L}(t)[y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)], \tag{77}$$

$$\boldsymbol{L}(t)=\frac{\boldsymbol{P}(t-1)\hat{\boldsymbol{\varphi}}(t)}{1+\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\boldsymbol{P}(t-1)\hat{\boldsymbol{\varphi}}(t)}, \tag{78}$$

$$\boldsymbol{P}(t)=[\boldsymbol{I}_{n_0}-\boldsymbol{L}(t)\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)]\boldsymbol{P}(t-1),\quad \boldsymbol{P}(0)=p_0\boldsymbol{I}_{n_0}, \tag{79}$$

$$\hat{\boldsymbol{\varphi}}(t)=[\hat{\boldsymbol{\varphi}}_{\mathrm{s}}^{\mathrm{T}}(t),-\hat{w}(t-1),-\hat{w}(t-2),\cdots,-\hat{w}(t-n_c)]^{\mathrm{T}}, \tag{80}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{s}}(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a),\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_b),\boldsymbol{f}(u(t))]^{\mathrm{T}}, \tag{81}$$

$$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t), \tag{82}$$

$$\hat{w}(t)=y(t)-\hat{\boldsymbol{\varphi}}_{\mathrm{s}}^{\mathrm{T}}(t)[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{b}}^{\mathrm{T}}(t),\hat{\boldsymbol{\gamma}}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{83}$$

$$\boldsymbol{f}(u(t))=[f_1(u(t)),f_2(u(t)),\cdots,f_{n_\gamma}(u(t))], \tag{84}$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{b}}^{\mathrm{T}}(t),\hat{\boldsymbol{\gamma}}^{\mathrm{T}}(t),\hat{\boldsymbol{c}}^{\mathrm{T}}(t)]^{\mathrm{T}}. \tag{85}$$

表 2 列出了基于关键项分离的递推广义最小二乘算法(77)—(85)的计算量，$n_0=n_a+n_b+n_\gamma+n_c$.

2.4 基于关键项分离的多新息广义最小二乘辨识方法

参考多新息广义随机梯度算法(41)—(51)的推导，从 KT-RGLS 算法(77)—(85)，可以得到辨识 IN-CARAR 系统参数向量 $\boldsymbol{\theta}$ 的基于关键项分离的多新息广义最小二乘算法（Key Term separation based Multi-Innovation Generalized Least Squares algorithm, KT-MI-GLS 算法）：

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{L}(t)[\boldsymbol{Y}(p,t)-\hat{\boldsymbol{\Phi}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\theta}}(t-1)], \tag{86}$$

$$\boldsymbol{L}(t)=\boldsymbol{P}(t-1)\hat{\boldsymbol{\Phi}}(p,t)[\boldsymbol{I}_p+\hat{\boldsymbol{\Phi}}^{\mathrm{T}}(p,t)\boldsymbol{P}(t-1)\hat{\boldsymbol{\Phi}}(p,t)]^{-1}, \tag{87}$$

$$\boldsymbol{P}(t)=\boldsymbol{P}(t-1)-\boldsymbol{L}(t)\hat{\boldsymbol{\Phi}}^{\mathrm{T}}(p,t)\boldsymbol{P}(t-1),\quad \boldsymbol{P}(0)=p_0\boldsymbol{I}_{n_0}, \tag{88}$$

$$\boldsymbol{Y}(p,t)=[y(t),y(t-1),\cdots,y(t-p+1)]^{\mathrm{T}}, \tag{89}$$

$$\hat{\boldsymbol{\Phi}}(p,t)=[\hat{\boldsymbol{\varphi}}(t),\hat{\boldsymbol{\varphi}}(t-1),\cdots,\hat{\boldsymbol{\varphi}}(t-p+1)], \tag{90}$$

$$\hat{\boldsymbol{\varphi}}(t)=[\hat{\boldsymbol{\varphi}}_{\mathrm{s}}^{\mathrm{T}}(t),-\hat{w}(t-1),-\hat{w}(t-2),\cdots,-\hat{w}(t-n_c)]^{\mathrm{T}}, \tag{91}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{s}}(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a),\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_b),\boldsymbol{f}(u(t))]^{\mathrm{T}}, \tag{92}$$

$$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t), \tag{93}$$

$$\hat{w}(t)=y(t)-\hat{\boldsymbol{\varphi}}_{\mathrm{s}}^{\mathrm{T}}(t)[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{b}}^{\mathrm{T}}(t),\hat{\boldsymbol{\gamma}}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{94}$$

$$\boldsymbol{f}(u(t))=[f_1(u(t)),f_2(u(t)),\cdots,f_{n_\gamma}(u(t))], \tag{95}$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{b}}^{\mathrm{T}}(t),\hat{\boldsymbol{\gamma}}^{\mathrm{T}}(t),\hat{\boldsymbol{c}}^{\mathrm{T}}(t)]^{\mathrm{T}}. \tag{96}$$

当新息长度 $p=1$ 时，KT-MI-GLS 算法退化为 KT-RGLS 算法(77)—(85).

注 11 上述方法是将输入非线性 CARAR 系统中噪声模型的参数 c_i 加到参数向量中，导出广义辨识方法，如广义随机梯度辨识方法、多新息广义随机梯度辨识方法、递推广义最小二乘辨识方法等.下面讨论几种基于数据滤波的参数辨识方法.

3 基于数据滤波的多新息辨识方法(1)

下面利用噪声模型传递函数对输入输出数据进行滤波，继而推导几种基于数据滤波的辨识方法.由于噪声模型是未知的，导致滤波无法进行，解决的方法是采用交互估计理论，同时估计噪声模型的参数，用噪声模型的估计值进行滤波，从而实现递推参数估计.

3.1 基于滤波的随机梯度辨识方法

考虑输入非线性受控自回归自回归系统(1)，重写如下：

$$A(z)y(t)=B(z)\bar{u}(t)+\frac{1}{C(z)}v(t), \tag{97}$$

$$\bar{u}(t)=f(u(t))=\gamma_1f_1(u(t))+\gamma_2f_2(u(t))+\cdots+\gamma_{n_\gamma}f_{n_\gamma}(u(t))=\boldsymbol{f}(u(t))\boldsymbol{\gamma}, \tag{98}$$

表 2 KT-RGLS 算法的计算量

Table 2 The computational efficiency of the KT-RGLS algorithm

变量	表达式	乘法次数	加法次数
$\hat{\boldsymbol{\theta}}(t)$	$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{L}(t)e(t)\in\mathbf{R}^{n_0}$	n_0	n_0
	$e(t):=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)\in\mathbf{R}$	n_0	n_0
$\boldsymbol{L}(t)$	$\boldsymbol{L}(t)=\boldsymbol{\zeta}(t)[1+\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\boldsymbol{\zeta}(t)]\in\mathbf{R}^{n_0}$	$2n_0$	n_0
	$\boldsymbol{\zeta}(t):=\boldsymbol{P}(t-1)\hat{\boldsymbol{\varphi}}(t)\in\mathbf{R}^{n_0}$	n_0^2	$n_0^2-n_0$
$\boldsymbol{P}(t)$	$\boldsymbol{P}(t)=\boldsymbol{P}(t-1)-\boldsymbol{L}(t)\boldsymbol{\zeta}^{\mathrm{T}}(t)\in\mathbf{R}^{n_0\times n_0}$	n_0^2	n_0^2
$\hat{w}(t)$	$\hat{w}(t)=y(t)-\boldsymbol{\varphi}_{\mathrm{s}}^{\mathrm{T}}(t)[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{b}}^{\mathrm{T}}(t),\hat{\boldsymbol{\gamma}}^{\mathrm{T}}(t)]^{\mathrm{T}}$	$n_a+n_b+n_\gamma$	$n_a+n_b+n_\gamma$
$\hat{\bar{u}}(t)$	$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t)$	n_r	n_r-1
总数		$2n_0^2+4n_0+n_a+n_b+2n_\gamma$	$2n_0^2+2n_0+n_a+n_b+2n_\gamma-1$
总 flop 数		$4n_0^2+6n_0+2n_a+2n_b+4n_\gamma-1$	

$$w(t)=\frac{1}{C(z)}v(t). \tag{99}$$

定义线性动态子系统的滤波输入变量 $\bar{u}_{\mathrm{f}}(t)$ 和系统的滤波输出变量 $y_{\mathrm{f}}(t)$ 分别为

$$\begin{aligned}\bar{u}_{\mathrm{f}}(t):=&C(z)\bar{u}(t)=\\&C(z)[\gamma_1 f_1(u(t))+\gamma_2 f_2(u(t))+\cdots+\gamma_{n_\gamma}f_{n_\gamma}(u(t))]=\\&\gamma_1 U_1(t)+\gamma_2 U_2(t)+\cdots+\gamma_{n_\gamma}U_{n_\gamma}(t),\end{aligned} \tag{100}$$

$$\begin{aligned}y_{\mathrm{f}}(t):=&C(z)y(t)=\\&y(t)+c_1y(t-1)+c_2y(t-2)+\cdots+c_{n_c}y(t-n_c),\end{aligned} \tag{101}$$

其中

$$U_i(t):=C(z)f_i(u(t)),\quad i=1,2,\cdots,n_\gamma. \tag{102}$$

定义参数向量 $\boldsymbol{\theta}_{\mathrm{s}}$ 和信息向量 $\boldsymbol{\varphi}_{\mathrm{s}}(t)$, $\boldsymbol{\varphi}_{\mathrm{f}}(t)$ 和 $\boldsymbol{\varphi}_{\mathrm{n}}(t)$ 如下:

$$\boldsymbol{\theta}_{\mathrm{s}}:=\begin{bmatrix}\boldsymbol{a}\\\boldsymbol{b}\\\boldsymbol{\gamma}\end{bmatrix}\in\mathbf{R}^{n_a+n_b+n_\gamma},$$

$$\begin{aligned}\boldsymbol{\varphi}_{\mathrm{s}}(t):=&[-y(t-1),-y(t-2),\cdots,-y(t-n_a),\\&\bar{u}(t-1),\bar{u}(t-2),\cdots,\bar{u}(t-n_b),\boldsymbol{f}(u(t))]^{\mathrm{T}}\in\\&\mathbf{R}^{n_a+n_b+n_\gamma},\end{aligned}$$

$$\begin{aligned}\boldsymbol{\varphi}_{\mathrm{f}}(t):=&[-y_{\mathrm{f}}(t-1),-y_{\mathrm{f}}(t-2),\cdots,-y_{\mathrm{f}}(t-n\_a),\\&\bar{u}_{\mathrm{f}}(t-1),\bar{u}_{\mathrm{f}}(t-2),\cdots,\bar{u}_{\mathrm{f}}(t-n\_b),U_1(t),\\&U_2(t),\cdots,U_{n_\gamma}(t)]^{\mathrm{T}}\in\mathbf{R}^{n_a+n_b+n_\gamma},\end{aligned}$$

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

其中参数向量 $\boldsymbol{a}$,$\boldsymbol{b}$ 和 $\boldsymbol{\gamma}$,以及后面的 $\boldsymbol{c}$ 的定义同上.

由式(99)可以得到噪声模型的辨识表达式:

$$\begin{aligned}w(t)=&[1-C(z)]w(t)+v(t)=\\&-\sum_{i=1}^{n_c}c_iw(t-i)+v(t)=\boldsymbol{\varphi}_{\mathrm{n}}^{\mathrm{T}}(t)\boldsymbol{c}+v(t).\end{aligned} \tag{103}$$

由式(97)可得

$$\begin{aligned}y(t)=&-\sum_{i=1}^{n_a}a_iy(t-i)+\sum_{i=1}^{n_b}b_i\bar{u}(t-i)+\\&\boldsymbol{f}(u(t))\boldsymbol{\gamma}+w(t)=\boldsymbol{\varphi}_{\mathrm{s}}^{\mathrm{T}}(t)\boldsymbol{\theta}_{\mathrm{s}}+w(t).\end{aligned} \tag{104}$$

式(97)两边同时乘以 $C(z)$ 得到

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

即

$$A(z)y_{\mathrm{f}}(t)=B(z)\bar{u}_{\mathrm{f}}(t)+v(t). \tag{105}$$

由式(105)展开得

$$\begin{aligned}y_{\mathrm{f}}(t)=&[1-A(z)]y_{\mathrm{f}}(t)+B(z)\bar{u}_{\mathrm{f}}(t)+v(t)=\\&-\sum_{i=1}^{n_a}a_iy_{\mathrm{f}}(t-i)+\sum_{i=1}^{n_b}b_i\bar{u}_{\mathrm{f}}(t-i)+\\&\bar{u}_{\mathrm{f}}(t)+v(t).\end{aligned} \tag{106}$$

将式(100)中 $\bar{u}_{\mathrm{f}}(t)$ 代入式(106)右边倒数第 2 项,可以得到滤波辨识模型:

$$\begin{aligned}y_{\mathrm{f}}(t)=&-\sum_{i=1}^{n_a}a_iy_{\mathrm{f}}(t-i)+\sum_{i=1}^{n_b}b_i\bar{u}_{\mathrm{f}}(t-i)+\\&\sum_{i=1}^{n_\gamma}\gamma_iU_i(t)+v(t)=\boldsymbol{\varphi}_{\mathrm{f}}^{\mathrm{T}}(t)\boldsymbol{\theta}_{\mathrm{s}}+v(t).\end{aligned} \tag{107}$$

由滤波辨识模型(107)和噪声模型(103)可以看出,滤波辨识模型(107)中信息向量 $\boldsymbol{\varphi}_{\mathrm{f}}(t)$ 包含了未知变量 $y_{\mathrm{f}}(t-i)$,$\bar{u}_{\mathrm{f}}(t-i)$ 和 $U_i(t)$,计算这些变量需要用到噪声模型中的未知参数向量 $\boldsymbol{c}$ 或未知多项式 $C(z)$,所以无法对参数向量 $\boldsymbol{\theta}_{\mathrm{s}}$ 进行估计;同样,噪声模型(103)中的 $w(t)$ 和 $\boldsymbol{\varphi}_{\mathrm{n}}(t)$ 是不可测的,使得参数向量 $\boldsymbol{c}$ 也无法辨识.针对这种困难,本文利用递阶辨识原理(交互估计理论)和辅助模型辨识思想,在滤波中使用的未知多项式 $C(z)$ 用其在时刻 t 的估计 $\hat{C}(t,z)$ 代替,未知滤波变量 $y_{\mathrm{f}}(t-i)$,$\bar{u}_{\mathrm{f}}(t-i)$ 和 $U_i(t)$ 也用其对应的估计 $\hat{y}_{\mathrm{f}}(t-i)$,$\hat{\bar{u}}_{\mathrm{f}}(t-i)$ 和 $\hat{U}_i(t)$ 代替,这样执行一个递阶参数估计计算过程.具体做法如下.

借助于辅助模型思想,利用 $\boldsymbol{c}$ 的估计 $\hat{\boldsymbol{c}}(t):=[\hat{c}_1(t),\hat{c}_2(t),\cdots,\hat{c}_{n_c}(t)]^{\mathrm{T}}\in\mathbf{R}^{n_c}$ 构造多项式 $C(z)$ 的估计 $\hat{C}(t,z):=1+\hat{c}_1(t)z^{-1}+\hat{c}_2(t)z^{-2}+\cdots+\hat{c}_{n_c}(t)z^{-n_c}$.由式(101)和(102)可以得到 $y_{\mathrm{f}}(t)$ 和 $U_i(t)$ 的估计:

$$\begin{aligned}\hat{y}_{\mathrm{f}}(t)=&\hat{C}(t,z)y(t)=\\&\hat{c}_1(t)y(t-1)+\hat{c}_2(t)y(t-2)+\cdots+\hat{c}_{n_c}(t)y(t-n_c)+y(t)=\\&[y(t-1),y(t-2),\cdots,y(t-n_c)]\hat{\boldsymbol{c}}(t)+y(t),\end{aligned}$$

$$\begin{aligned}\hat{U}_i(t)=&\hat{C}(t,z)f_i(u(t))=\\&\hat{c}_1(t)f_i(u(t-1))+\hat{c}_2(t)f_i(u(t-2))+\cdots+\\&\hat{c}_{n_c}(t)f_i(u(t-n_c))+f_i(u(t))=\\&[f_i(u(t-1)),f_i(u(t-2)),\cdots,f_i(u(t-n_c))]\hat{\boldsymbol{c}}(t)+f_i(u(t)).\end{aligned}$$

令 $\hat{\boldsymbol{\theta}}_{\mathrm{s}}(t):=\begin{bmatrix}\hat{\boldsymbol{a}}(t)\\\hat{\boldsymbol{b}}(t)\\\hat{\boldsymbol{\gamma}}(t)\end{bmatrix}$ 为参数向量 $\boldsymbol{\theta}_{\mathrm{s}}=\begin{bmatrix}\boldsymbol{a}\\\boldsymbol{b}\\\boldsymbol{\gamma}\end{bmatrix}$ 在时刻 t 的估计.根据式(98)和式(100),可知未知变量 $\bar{u}(t)$ 和 $\bar{u}_{\mathrm{f}}(t)$ 的估计分别为

$$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t),$$

$$\begin{aligned}\hat{\bar{u}}_{\mathrm{f}}(t)=&\hat{\gamma}_1(t)\hat{U}_1(t)+\hat{\gamma}_2(t)\hat{U}_2(t)+\cdots+\hat{\gamma}_{n_\gamma}(t)\hat{U}_{n_\gamma}(t)=\\&[\hat{U}_1(t),\hat{U}_2(t),\cdots,\hat{U}_{n_\gamma}(t)]\hat{\boldsymbol{\gamma}}(t).\end{aligned} \tag{108}$$

用估计 $\hat{y}_{\mathrm{f}}(t-i)$,$\hat{\bar{u}}_{\mathrm{f}}(t-i)$,$\hat{U}_i(t)$ 和 $\hat{\bar{u}}(t-i)$ 分别构造信息向量 $\boldsymbol{\varphi}_{\mathrm{f}}(t)$ 和 $\boldsymbol{\varphi}_{\mathrm{s}}(t)$ 的估计:

$$\begin{aligned}\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t):=&[-\hat{y}_{\mathrm{f}}(t-1),-\hat{y}_{\mathrm{f}}(t-2),\cdots,-\hat{y}_{\mathrm{f}}(t-n_a),\\&\hat{\bar{u}}_{\mathrm{f}}(t-1),\hat{\bar{u}}_{\mathrm{f}}(t-2),\cdots,\hat{\bar{u}}_{\mathrm{f}}(t-n_b),\hat{U}_1(t),\\&\hat{U}_2(t),\cdots,\hat{U}_{n_\gamma}(t)]^{\mathrm{T}}\in\mathbf{R}^{n_a+n_b+n_\gamma},\end{aligned}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{s}}(t):=[-y(t-1),-y(t-2),\cdots,-y(t-n_a),\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_b),\boldsymbol{f}(u(t))]^{\mathrm{T}}\in\mathbf{R}^{n_a+n_b+n_\gamma}.$$

由式(104)可知未知变量 $w(t)$ 可用下式进行估计,

$$\hat{w}(t)=y(t)-\hat{\boldsymbol{\varphi}}_{\mathrm{s}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_{\mathrm{s}}(t-1).$$

上式中之所以使用 $\hat{\boldsymbol{\theta}}_{\mathrm{s}}(t-1)$,而没有使用 $\hat{\boldsymbol{\theta}}_{\mathrm{s}}(t)$,是为了使算法能进行递推计算.当然,上式如果使用 $\hat{\boldsymbol{\theta}}_{\mathrm{s}}(t)$,就必须用 $\hat{C}(t-1,z)$ 进行滤波,算法才能实现递推计算.

用估计 $\hat{w}(t-i)$ 构造信息向量 $\boldsymbol{\varphi}_{\mathrm{n}}(t)$ 的估计:

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

对于滤波辨识模型(107)和噪声模型(103),利用梯度搜索,极小化准则函数

$$J_4(\boldsymbol{\theta}_{\mathrm{s}}):=[\hat{y}_{\mathrm{f}}(t)-\hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\mathrm{T}}(t)\boldsymbol{\theta}_{\mathrm{s}}]^2,$$

$$J_5(\boldsymbol{c}):=[\hat{w}(t)-\hat{\boldsymbol{\varphi}}_{\mathrm{n}}^{\mathrm{T}}(t)\boldsymbol{c}]^2,$$

可以得到第 1 种辨识 IN-CARAR 系统参数向量 $\boldsymbol{\theta}$ 的基于数据滤波的随机梯度算法(data Filtering based Stochastic Gradient algorithm, F-SG 算法):

$$\hat{\boldsymbol{\theta}}_{\mathrm{s}}(t)=\hat{\boldsymbol{\theta}}_{\mathrm{s}}(t-1)+\frac{\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)}{r_1(t)}e_{\mathrm{f}}(t),\tag{109}$$

$$e_{\mathrm{f}}(t)=\hat{y}_{\mathrm{f}}(t)-\hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_{\mathrm{s}}(t-1),\tag{110}$$

$$r_1(t)=r_1(t-1)+\|\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)\|^2,\quad r_1(0)=1,\tag{111}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)=[-\hat{y}_{\mathrm{f}}(t-1),-\hat{y}_{\mathrm{f}}(t-2),\cdots,-\hat{y}_{\mathrm{f}}(t-n_a),\hat{\bar{u}}_{\mathrm{f}}(t-1),\hat{\bar{u}}_{\mathrm{f}}(t-2),\cdots,\hat{\bar{u}}_{\mathrm{f}}(t-n_b),\hat{U}_1(t),\hat{U}_2(t),\cdots,\hat{U}_{n_\gamma}(t)]^{\mathrm{T}},\tag{112}$$

$$\hat{y}_{\mathrm{f}}(t)=[y(t-1),y(t-2),\cdots,y(t-n_c)]\hat{\boldsymbol{c}}(t)+y(t),\tag{113}$$

$$\hat{\bar{u}}_{\mathrm{f}}(t)=[\hat{U}_1(t),\hat{U}_2(t),\cdots,\hat{U}_{n_\gamma}(t)]\hat{\boldsymbol{\gamma}}(t),\tag{114}$$

$$\hat{U}_i(t)=[f_i(u(t-1)),f_i(u(t-2)),\cdots,f_i(u(t-n_c))]\hat{\boldsymbol{c}}(t)+f_i(u(t)),\tag{115}$$

$$\hat{\boldsymbol{c}}(t)=\hat{\boldsymbol{c}}(t-1)+\frac{\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)}{r_2(t)}e_{\mathrm{n}}(t),\tag{116}$$

$$e_{\mathrm{n}}(t)=\hat{w}(t)-\hat{\boldsymbol{\varphi}}_{\mathrm{n}}^{\mathrm{T}}(t)\hat{\boldsymbol{c}}(t-1),\tag{117}$$

$$r_2(t)=r_2(t-1)+\|\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\|^2,\quad r_2(0)=1,\tag{118}$$

$$\hat{w}(t)=y(t)-\hat{\boldsymbol{\varphi}}_{\mathrm{s}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_{\mathrm{s}}(t-1),\tag{119}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{s}}(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a),\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_b),\boldsymbol{f}(u(t))]^{\mathrm{T}},\tag{120}$$

$$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t),\tag{121}$$

$$\boldsymbol{f}(u(t))=[f_1(u(t)),f_2(u(t)),\cdots f_{n_\gamma}(u(t))],\tag{122}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)=[-\hat{w}(t-1),-\hat{w}(t-2),\cdots,-\hat{w}(t-n_c)]^{\mathrm{T}},\tag{123}$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{\theta}}_{\mathrm{s}}^{\mathrm{T}}(t),\hat{\boldsymbol{c}}^{\mathrm{T}}(t)]^{\mathrm{T}},\tag{124}$$

$$\hat{\boldsymbol{\theta}}_{\mathrm{s}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{b}}^{\mathrm{T}}(t),\hat{\boldsymbol{\gamma}}^{\mathrm{T}}(t)]^{\mathrm{T}}.\tag{125}$$

注 12 F-SG 算法的性能与 KT-GSG 算法(57)—(65)性能类似,只是这里交互估计参数向量 $\boldsymbol{\theta}_{\mathrm{s}}$ 和 $\boldsymbol{c}$,是因为通过滤波把系统变为 2 个辨识模型所致.当然,也可以引入遗忘因子来改善 F-SG 算法的参数估计精度;另一种改善参数精度的方法是引入新息长度,导出基于滤波的多新息随机梯度辨识方法.

3.2 基于滤波的多新息随机梯度辨识方法

同样,为进一步提高基于数据滤波的随机梯度辨识算法的收敛速度和参数辨识精度,可根据多新息辨识理论,推导出基于数据滤波的多新息随机梯度辨识算法.

定义堆积滤波信息矩阵 $\hat{\boldsymbol{\Phi}}_{\mathrm{f}}(p,t)$ 和 $\hat{\boldsymbol{\Phi}}_{\mathrm{n}}(p,t)$,堆积滤波输出向量 $\hat{\boldsymbol{Y}}_{\mathrm{f}}(p,t)$ 和噪声向量 $\hat{\boldsymbol{W}}(p,t)$ 为

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

$$\hat{\boldsymbol{\Phi}}_{\mathrm{f}}(p,t):=[\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t),\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-1),\cdots,\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-p+1)]\in\mathbf{R}^{(n_a+n_b+n_\gamma)\times p},$$

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

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

扩展式(110)和式(117)中标量新息 $e_{\mathrm{f}}(t)\in\mathbf{R}$ 和 $e_{\mathrm{n}}(t)\in\mathbf{R}$ 为新息向量:

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

$$\boldsymbol{E}_{\mathrm{n}}(p,t):=\hat{\boldsymbol{W}}(p,t)-\hat{\boldsymbol{\Phi}}_{\mathrm{n}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{c}}(t-1)\in\mathbf{R}^{p},$$

可以得到第 1 种辨识 IN-CARAR 系统参数向量 $\boldsymbol{\theta}$ 的基于数据滤波的多新息随机梯度算法(data Filtering based Multi-Innovation Stochastic Gradient algorithm, F-MISG 算法):

$$\hat{\boldsymbol{\theta}}_{\mathrm{s}}(t)=\hat{\boldsymbol{\theta}}_{\mathrm{s}}(t-1)+\frac{\hat{\boldsymbol{\Phi}}_{\mathrm{f}}(p,t)}{r_1(t)}\boldsymbol{E}_{\mathrm{f}}(p,t),\tag{126}$$

$$\boldsymbol{E}_{\mathrm{f}}(p,t)=\boldsymbol{Y}_{\mathrm{f}}(p,t)-\hat{\boldsymbol{\Phi}}_{\mathrm{f}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_{\mathrm{s}}(t-1),\tag{127}$$

$$r_1(t)=r_1(t-1)+\|\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)\|^2,\quad r_1(0)=1,\tag{128}$$

$$\hat{\boldsymbol{Y}}_{\mathrm{f}}(p,t)=[\hat{y}_{\mathrm{f}}(t),\hat{y}_{\mathrm{f}}(t-1),\cdots,\hat{y}_{\mathrm{f}}(t-p+1)]^{\mathrm{T}},\tag{129}$$

$$\hat{\boldsymbol{\Phi}}_{\mathrm{f}}(p,t)=[\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t),\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-1),\cdots,\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-p+1)],\tag{130}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)=[-\hat{y}_{\mathrm{f}}(t-1),-\hat{y}_{\mathrm{f}}(t-2),\cdots,-\hat{y}_{\mathrm{f}}(t-n_a),\hat{\bar{u}}_{\mathrm{f}}(t-1),\hat{\bar{u}}_{\mathrm{f}}(t-2),\cdots,\hat{\bar{u}}_{\mathrm{f}}(t-n_b),\hat{U}_1(t),\hat{U}_2(t),\cdots,\hat{U}_{n_\gamma}(t)]^{\mathrm{T}},\tag{131}$$

$$\hat{y}_{\mathrm{f}}(t)=[y(t-1),y(t-2),\cdots,y(t-n_c)]\hat{\boldsymbol{c}}(t)+y(t),\tag{132}$$

$$\hat{\bar{u}}_{\mathrm{f}}(t)=[\hat{U}_1(t),\hat{U}_2(t),\cdots,\hat{U}_{n_\gamma}(t)]\hat{\boldsymbol{\gamma}}(t),\tag{133}$$

$$\hat{U}_i(t)=[f_i(u(t-1)),f_i(u(t-2)),\cdots,$$

$$f_i(u(t-n_c))]\hat{\boldsymbol{c}}(t)+f_i(u(t)),\tag{134}$$

$$\hat{\boldsymbol{c}}(t)=\hat{\boldsymbol{c}}(t-1)+\frac{\hat{\boldsymbol{\Phi}}_{\mathrm{n}}(p,t)}{r_2(t)}\boldsymbol{E}_{\mathrm{n}}(p,t),\tag{135}$$

$$\boldsymbol{E}_{\mathrm{n}}(p,t)=\hat{\boldsymbol{W}}(p,t)-\hat{\boldsymbol{\Phi}}_{\mathrm{n}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{c}}(t-1),\tag{136}$$

$$r_2(t)=r_2(t-1)+\|\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\|^2,\quad r_2(0)=1,\tag{137}$$

$$\hat{\boldsymbol{W}}(p,t)=[\hat{w}(t),\hat{w}(t-1),\cdots,\hat{w}(t-p+1)]^{\mathrm{T}},\tag{138}$$

$$\hat{\boldsymbol{\Phi}}_{\mathrm{n}}(p,t)=[\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t),\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t-1),\cdots,\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t-p+1)],\tag{139}$$

$$\hat{w}(t)=y(t)-\hat{\boldsymbol{\varphi}}_{\mathrm{s}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_{\mathrm{s}}(t-1),\tag{140}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{s}}(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a),\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_b),\boldsymbol{f}(u(t))]^{\mathrm{T}},\tag{141}$$

$$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t),\tag{142}$$

$$\boldsymbol{f}(u(t))=[f_1(u(t)),f_2(u(t)),\cdots,f_{n_\gamma}(u(t))],\tag{143}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)=[-\hat{w}(t-1),-\hat{w}(t-2),\cdots,-\hat{w}(t-n_c)]^{\mathrm{T}},\tag{144}$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{\theta}}_{\mathrm{s}}^{\mathrm{T}}(t),\hat{\boldsymbol{c}}^{\mathrm{T}}(t)]^{\mathrm{T}},\tag{145}$$

$$\hat{\boldsymbol{\theta}}_{\mathrm{s}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{b}}^{\mathrm{T}}(t),\hat{\boldsymbol{\gamma}}^{\mathrm{T}}(t)]^{\mathrm{T}}.\tag{146}$$

当新息长度 $p=1$ 时,这个 F-MISG 算法退化为 F-SG 算法(109)—(125).

F-MISG 算法(126)—(146)的计算步骤如下:

1) 初始化:令 $t=1$. 置初值 $\hat{\boldsymbol{\theta}}(0)=\mathbf{1}_{n_0}/p_0$, $r_1(0)=1$, $r_2(0)=1$, $\hat{w}(-i)=1/p_0$, $\hat{y}_{\mathrm{f}}(-i)=1/p_0$, $\hat{\bar{u}}_{\mathrm{f}}(-i)=1/p_0$ 和 $\hat{\bar{u}}(-i)=1/p_0$, $i=0,1,\cdots,n_0$, $p_0=10^6$, 给定基函数 $f_i(\cdot)$ 和新息长度 p.

2) 收集数据 $u(t)$ 和 $y(t)$,用式(143)构造基函数行向量 $\boldsymbol{f}(u(t))$.

3) 用式(144)构造信息向量 $\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)$,用式(139)构造信息矩阵 $\hat{\boldsymbol{\Phi}}_{\mathrm{n}}(p,t)$,用式(141)构造信息向量 $\hat{\boldsymbol{\varphi}}_{\mathrm{s}}(t)$.

4) 用式(140)计算 $\hat{w}(t)$,用式(138)构造噪声向量 $\hat{\boldsymbol{W}}(p,t)$.

5) 用式(136)计算新息向量 $\boldsymbol{E}_{\mathrm{n}}(p,t)$,用式(137)计算 $r_2(t)$.

6) 根据式(135)刷新参数估计向量 $\hat{\boldsymbol{c}}(t)$.

7) 用式(132)计算 $\hat{y}_{\mathrm{f}}(t)$,用式(134)计算 $\hat{U}_i(t)$.

8) 用式(129)构造堆积输出向量 $\boldsymbol{Y}_{\mathrm{f}}(p,t)$,用式(131)构造信息向量 $\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)$,用式(130)构造信息矩阵 $\hat{\boldsymbol{\Phi}}_{\mathrm{f}}(p,t)$.

9) 用式(127)计算新息向量 $\boldsymbol{E}_{\mathrm{f}}(p,t)$,用式(128)计算 $r_1(t)$.

10) 根据式(126)刷新参数估计向量 $\hat{\boldsymbol{\theta}}_{\mathrm{s}}(t)$.

11) 从式(146)的 $\hat{\boldsymbol{\theta}}_{\mathrm{s}}(t)$ 中读出 $\hat{\boldsymbol{\gamma}}(t)$,用式(142)计算 $\hat{\bar{u}}(t)$,用式(133)计算滤波输入 $\hat{\bar{u}}_{\mathrm{f}}(t)$.

12) t 增 1,转到第 2)步.

F-MISG 算法(126)—(146)计算系统参数估计 $\hat{\boldsymbol{\theta}}(t)$ 的流程如图 2 所示.

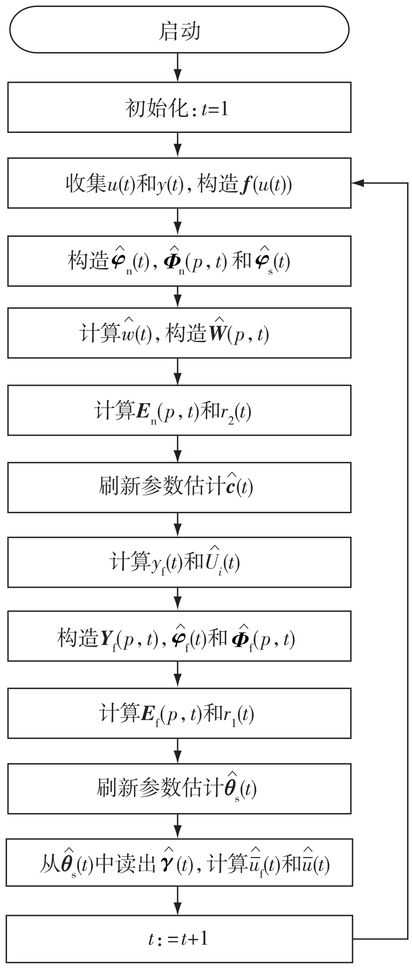

图 2 计算 F-MISG 参数估计 $\hat{\boldsymbol{\theta}}(t)$ 的流程

Fig. 2 The flowchart of computing the F-MISG parameter estimate $\hat{\boldsymbol{\theta}}(t)$

3.3 基于滤波的递推最小二乘辨识方法

对于滤波辨识模型(107)和噪声模型(103),利用最小二乘原理,极小化准则函数

$$J_6(\boldsymbol{\theta}_{\mathrm{s}}):=\sum_{j=1}^{t}[\hat{y}_{\mathrm{f}}(j)-\hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\mathrm{T}}(j)\boldsymbol{\theta}_{\mathrm{s}}]^2,$$

$$J_7(\boldsymbol{c}):=\sum_{j=1}^{t}[\hat{w}(j)-\hat{\boldsymbol{\varphi}}_{\mathrm{n}}^{\mathrm{T}}(j)\boldsymbol{c}]^2,$$

可以得到第 1 种辨识 IN-CARAR 系统参数向量 $\boldsymbol{\theta}$ 的基于数据滤波的递推最小二乘算法(data Filtering based Recursive Least Squares algorithm, F-RLS 算法):

$$\hat{\boldsymbol{\theta}}_{\rm s}(t)=\hat{\boldsymbol{\theta}}_{\rm s}(t-1)+\boldsymbol{L}(t)[\hat{y}_{\rm f}(t)-\hat{\boldsymbol{\varphi}}_{\rm f}^{\rm T}(t)\hat{\boldsymbol{\theta}}_{\rm s}(t-1)], \quad (147)$$

$$\boldsymbol{L}(t)=\frac{\boldsymbol{P}(t-1)\hat{\boldsymbol{\varphi}}_{\rm f}(t)}{1+\hat{\boldsymbol{\varphi}}_{\rm f}^{\rm T}(t)\boldsymbol{P}(t-1)\hat{\boldsymbol{\varphi}}_{\rm f}(t)}, \quad (148)$$

$$\boldsymbol{P}(t)=[\boldsymbol{I}_{n_a+n_b+n_\gamma}-\boldsymbol{L}(t)\hat{\boldsymbol{\varphi}}_{\rm f}^{\rm T}(t)]\boldsymbol{P}(t-1),\ \boldsymbol{P}(0)=p_0\boldsymbol{I}_{n_a+n_b+n_\gamma}, \quad (149)$$

$$\hat{\boldsymbol{\varphi}}_{\rm f}(t)=[-\hat{y}_{\rm f}(t-1),-\hat{y}_{\rm f}(t-2),\cdots,-\hat{y}_{\rm f}(t-n_a),\hat{\bar{u}}_{\rm f}(t-1),\hat{\bar{u}}_{\rm f}(t-2),\cdots,\hat{\bar{u}}_{\rm f}(t-n_b),\hat{U}_1(t),\hat{U}_2(t),\cdots,\hat{U}_{n_\gamma}(t)]^{\rm T}, \quad (150)$$

$$\hat{y}_{\rm f}(t)=[y(t-1),y(t-2),\cdots,y(t-n_c)]\hat{\boldsymbol{c}}(t)+y(t), \quad (151)$$

$$\hat{\bar{u}}_{\rm f}(t)=[\hat{U}_1(t),\hat{U}_2(t),\cdots,\hat{U}_{n_\gamma}(t)]\hat{\boldsymbol{\gamma}}(t), \quad (152)$$

$$\hat{U}_i(t)=[f_i(u(t-1)),f_i(u(t-2)),\cdots,f_i(u(t-n_c))]\hat{\boldsymbol{c}}(t)+f_i(u(t)), \quad (153)$$

$$\hat{\boldsymbol{c}}(t)=\hat{\boldsymbol{c}}(t-1)+\boldsymbol{L}_{\rm n}(t)[\hat{w}(t)-\hat{\boldsymbol{\varphi}}_{\rm n}^{\rm T}(t)\hat{\boldsymbol{c}}(t-1)], \quad (154)$$

$$\boldsymbol{L}_{\rm n}(t)=\frac{\boldsymbol{P}_{\rm n}(t-1)\hat{\boldsymbol{\varphi}}_{\rm n}(t)}{1+\hat{\boldsymbol{\varphi}}_{\rm n}^{\rm T}(t)\boldsymbol{P}_{\rm n}(t-1)\hat{\boldsymbol{\varphi}}_{\rm n}(t)}, \quad (155)$$

$$\boldsymbol{P}_{\rm n}(t)=[\boldsymbol{I}_{n_c}-\boldsymbol{L}_{\rm n}(t)\hat{\boldsymbol{\varphi}}_{\rm n}^{\rm T}(t)]\boldsymbol{P}_{\rm n}(t-1),\quad \boldsymbol{P}_{\rm n}(0)=p_0\boldsymbol{I}_{n_c}, \quad (156)$$

$$\hat{w}(t)=y(t)-\hat{\boldsymbol{\varphi}}_{\rm s}^{\rm T}(t)\hat{\boldsymbol{\theta}}_{\rm s}(t-1), \quad (157)$$

$$\hat{\boldsymbol{\varphi}}_{\rm s}(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a),\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_b),\boldsymbol{f}(u(t))]^{\rm T}, \quad (158)$$

$$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t), \quad (159)$$

$$\boldsymbol{f}(u(t))=[f_1(u(t)),f_2(u(t)),\cdots,f_{n_\gamma}(u(t))], \quad (160)$$

$$\hat{\boldsymbol{\varphi}}_{\rm n}(t)=[-\hat{w}(t-1),-\hat{w}(t-2),\cdots,-\hat{w}(t-n_c)]^{\rm T}, \quad (161)$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{\theta}}_{\rm s}^{\rm T}(t),\hat{\boldsymbol{c}}^{\rm T}(t)]^{\rm T}, \quad (162)$$

$$\hat{\boldsymbol{\theta}}_{\rm s}(t)=[\hat{\boldsymbol{a}}^{\rm T}(t),\hat{\boldsymbol{b}}^{\rm T}(t),\hat{\boldsymbol{\gamma}}^{\rm T}(t)]^{\rm T}. \quad (163)$$

注 13 F-RLS 算法比 F-SG 算法(109)—(125)有更快的收敛速度,在相同数据长度下(即相同递推步数下),前者能给出更高精度的参数估计.

3.4 基于滤波的多新息最小二乘辨识方法

扩展式(147)和式(154)中标量新息 $e_{\rm f}(t):=\hat{y}_{\rm f}(t)-\hat{\boldsymbol{\varphi}}_{\rm f}^{\rm T}(t)\hat{\boldsymbol{\theta}}_{\rm s}(t-1)\in\mathbf{R}$ 和 $e_{\rm n}(t):=\hat{w}(t)-\hat{\boldsymbol{\varphi}}_{\rm n}^{\rm T}(t)\cdot\hat{\boldsymbol{c}}(t-1)\in\mathbf{R}$ 为新息向量:

$$\boldsymbol{E}_{\rm f}(p,t)=\boldsymbol{Y}_{\rm f}(p,t)-\hat{\boldsymbol{\Phi}}_{\rm f}^{\rm T}(p,t)\hat{\boldsymbol{\theta}}_{\rm s}(t-1)\in\mathbf{R}^p,$$

$$\boldsymbol{E}_{\rm n}(p,t)=\hat{\boldsymbol{W}}(p,t)-\hat{\boldsymbol{\Phi}}_{\rm n}^{\rm T}(p,t)\hat{\boldsymbol{c}}(t-1)\in\mathbf{R}^p,$$

可以得到第 1 种辨识 IN-CARAR 系统参数向量 $\boldsymbol{\theta}$ 的基于数据滤波的多新息最小二乘算法(data Filtering based Multi - Innovation Least Squares algorithm, F-MILS算法):

$$\hat{\boldsymbol{\theta}}_{\rm s}(t)=\hat{\boldsymbol{\theta}}_{\rm s}(t-1)+\boldsymbol{L}(t)[\boldsymbol{Y}_{\rm f}(p,t)-\hat{\boldsymbol{\Phi}}_{\rm f}^{\rm T}(p,t)\hat{\boldsymbol{\theta}}_{\rm s}(t-1)], \quad (164)$$

$$\boldsymbol{L}(t)=\frac{\boldsymbol{P}(t-1)\hat{\boldsymbol{\Phi}}_{\rm f}(p,t)}{\boldsymbol{I}_p+\hat{\boldsymbol{\Phi}}_{\rm f}^{\rm T}(p,t)\boldsymbol{P}(t-1)\hat{\boldsymbol{\Phi}}_{\rm f}(p,t)}, \quad (165)$$

$$\boldsymbol{P}(t)=\boldsymbol{P}(t-1)-\boldsymbol{L}(t)\hat{\boldsymbol{\Phi}}_{\rm f}^{\rm T}(p,t)\boldsymbol{P}(t-1),\ \boldsymbol{P}(0)=p_0\boldsymbol{I}_{n_a+n_b+n_\gamma}, \quad (166)$$

$$\hat{\boldsymbol{Y}}_{\rm f}(p,t)=[\hat{y}_{\rm f}(t),\hat{y}_{\rm f}(t-1),\cdots,\hat{y}_{\rm f}(t-p+1)]^{\rm T}, \quad (167)$$

$$\hat{\boldsymbol{\Phi}}_{\rm f}(p,t)=[\hat{\boldsymbol{\varphi}}_{\rm f}(t),\hat{\boldsymbol{\varphi}}_{\rm f}(t-1),\cdots,\hat{\boldsymbol{\varphi}}_{\rm f}(t-p+1)], \quad (168)$$

$$\hat{\boldsymbol{\varphi}}_{\rm f}(t)=[-\hat{y}_{\rm f}(t-1),-\hat{y}_{\rm f}(t-2),\cdots,-\hat{y}_{\rm f}(t-n_a),\hat{\bar{u}}_{\rm f}(t-1),\hat{\bar{u}}_{\rm f}(t-2),\cdots,\hat{\bar{u}}_{\rm f}(t-n_b),\hat{U}_1(t),\hat{U}_2(t),\cdots,\hat{U}_{n_\gamma}(t)]^{\rm T}, \quad (169)$$

$$\hat{y}_{\rm f}(t)=[y(t-1),y(t-2),\cdots,y(t-n_c)]\hat{\boldsymbol{c}}(t)+y(t), \quad (170)$$

$$\hat{\bar{u}}_{\rm f}(t)=[\hat{U}_1(t),\hat{U}_2(t),\cdots,\hat{U}_{n_\gamma}(t)]\hat{\boldsymbol{\gamma}}(t), \quad (171)$$

$$\hat{U}_i(t)=[f_i(u(t-1)),f_i(u(t-2)),\cdots,f_i(u(t-n_c))]\hat{\boldsymbol{c}}(t)+f_i(u(t)), \quad (172)$$

$$\hat{\boldsymbol{c}}(t)=\hat{\boldsymbol{c}}(t-1)+\boldsymbol{L}_{\rm n}(t)[\hat{\boldsymbol{W}}(t)-\hat{\boldsymbol{\Phi}}_{\rm n}^{\rm T}(p,t)\hat{\boldsymbol{c}}(t-1)], \quad (173)$$

$$\boldsymbol{L}_{\rm n}(t)=\frac{\boldsymbol{P}_{\rm n}(t-1)\hat{\boldsymbol{\Phi}}_{\rm n}(p,t)}{\boldsymbol{I}_p+\hat{\boldsymbol{\Phi}}_{\rm n}^{\rm T}(p,t)\boldsymbol{P}_{\rm n}(t-1)\hat{\boldsymbol{\Phi}}_{\rm n}(p,t)}, \quad (174)$$

$$\boldsymbol{P}_{\rm n}(t)=\boldsymbol{P}_{\rm n}(t-1)-\boldsymbol{L}_{\rm n}(t)\hat{\boldsymbol{\Phi}}_{\rm n}^{\rm T}(p,t)\boldsymbol{P}_{\rm n}(t-1),\ \boldsymbol{P}_{\rm n}(0)=p_0\boldsymbol{I}_{n_c}, \quad (175)$$

$$\hat{\boldsymbol{W}}(p,t)=[\hat{w}(t),\hat{w}(t-1),\cdots,\hat{w}(t-p+1)]^{\rm T}, \quad (176)$$

$$\hat{\boldsymbol{\Phi}}_{\rm n}(p,t)=[\hat{\boldsymbol{\varphi}}_{\rm n}(t),\hat{\boldsymbol{\varphi}}_{\rm n}(t-1),\cdots,\hat{\boldsymbol{\varphi}}_{\rm n}(t-p+1)], \quad (177)$$

$$\hat{w}(t)=y(t)-\hat{\boldsymbol{\varphi}}_{\rm s}^{\rm T}(t)\hat{\boldsymbol{\theta}}_{\rm s}(t-1), \quad (178)$$

$$\hat{\boldsymbol{\varphi}}_{\rm s}(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a),\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_b),\boldsymbol{f}(u(t))]^{\rm T}, \quad (179)$$

$$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t), \quad (180)$$

$$\boldsymbol{f}(u(t))=[f_1(u(t)),f_2(u(t)),\cdots,f_{n_\gamma}(u(t))], \quad (181)$$

$$\hat{\boldsymbol{\varphi}}_{\rm n}(t)=[-\hat{w}(t-1),-\hat{w}(t-2),\cdots,-\hat{w}(t-n_c)]^{\rm T}, \quad (182)$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{\theta}}_{\rm s}^{\rm T}(t),\hat{\boldsymbol{c}}^{\rm T}(t)]^{\rm T}, \quad (183)$$

$$\hat{\boldsymbol{\theta}}_{\rm s}(t)=[\hat{\boldsymbol{a}}^{\rm T}(t),\hat{\boldsymbol{b}}^{\rm T}(t),\hat{\boldsymbol{\gamma}}^{\rm T}(t)]^{\rm T}. \quad (184)$$

当新息长度 $p=1$ 时,F-MILS 算法退化为 F-RLS 算法(147)—(163).

注 14 本节讨论的基于数据滤波的辨识算法是将系统的参数分成 2 个参数集,一个是线性动态子系统参数加上静态非线性环节的参数,另一个是噪声模型的参数.算法包括耦合的 2 个子算法,它们交互递推计算每一时刻 2 个参数集的估计.下面讨论另外 2 种基于数据滤波的辨识方法.

4 基于数据滤波的多新息辨识方法(2)

本节讨论第 2 种基于数据滤波的辨识算法.该

算法采用的是一个辨识模型，即一个参数向量，不需要交互计算，就能递推计算系统的参数估计，简化了算法的步骤.

4.1 基于滤波的随机梯度辨识方法

定义滤波输入 $\bar{u}_{\rm f}(t)$ 和滤波输出 $y_{\rm f}(t)$ 为

$$\bar{u}_{\rm f}(t):=C(z)\bar{u}(t)=\bar{u}(t)+c_1\bar{u}(t-1)+c_2\bar{u}(t-2)+\cdots+c_{n_c}\bar{u}(t-n_c)=\sum_{i=1}^{n_\gamma}\gamma_i f_i(u(t))+\sum_{i=1}^{n_c}c_i\bar{u}(t-i),\tag{185}$$

$$y_{\rm f}(t):=C(z)y(t)=y(t)+c_1y(t-1)+c_2y(t-2)+\cdots+c_{n_c}y(t-n_c).\tag{186}$$

将滤波后的模型(105)展开为

$$y_{\rm f}(t)=[1-A(z)]y_{\rm f}(t)+[B(z)-1]\bar{u}_{\rm f}(t)+\bar{u}_{\rm f}(t)+v(t)=-\sum_{i=1}^{n_a}a_iy_{\rm f}(t-i)+\sum_{i=1}^{n_b}b_i\bar{u}_{\rm f}(t-i)+\bar{u}_{\rm f}(t)+v(t).\tag{187}$$

与前节不同的是这里将滤波后变量反代入滤波后的模型中，也就是将式(186)中的滤波输入 $\bar{u}_{\rm f}(t)$ 代入式(187)右边倒数第2项，得到滤波辨识模型：

$$y_{\rm f}(t)=-\sum_{i=1}^{n_a}a_iy_{\rm f}(t-i)+\sum_{i=1}^{n_b}b_i\bar{u}_{\rm f}(t-i)+\sum_{i=1}^{n_\gamma}\gamma_i f_i(u(t))+\sum_{i=1}^{n_c}c_i\bar{u}(t-i)+v(t)\tag{188}$$

$$=:\boldsymbol{\varphi}_{\rm f}^{\rm T}(t)\boldsymbol{\theta}+v(t),\tag{189}$$

$$\boldsymbol{\varphi}_{\rm f}(t):=[-y_{\rm f}(t-1),-y_{\rm f}(t-2),\cdots,-y_{\rm f}(t-n\_a),\bar{u}_{\rm f}(t-1),\bar{u}_{\rm f}(t-2),\cdots,\bar{u}_{\rm f}(t-n_b),\boldsymbol{f}(u(t)),\bar{u}(t-1),\bar{u}(t-2),\cdots,\bar{u}(t-n_c)]^{\rm T}\in\mathbf{R}^{n_0},$$

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

其中参数向量 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 和 $\boldsymbol{\gamma}$ 的定义同上.

同样，由于多项式 $C(z)$ 未知，因此信息向量 $\boldsymbol{\varphi}_{\rm f}(t)$ 中的 $y_{\rm f}(t-i)$ 和 $\bar{u}_{\rm f}(t-i)$ 是未知的，中间变量 $\bar{u}(t-i)$ 也是未知的，这是辨识的困难所在.这里借助于辅助模型的思想，用它们的估计来代替.

令 $\hat{\boldsymbol{\theta}}(t):=[\hat{\boldsymbol{a}}^{\rm T}(t),\hat{\boldsymbol{b}}^{\rm T}(t),\hat{\boldsymbol{\gamma}}^{\rm T}(t),\hat{\boldsymbol{c}}^{\rm T}(t)]^{\rm T}$ 是参数向量 $\boldsymbol{\theta}=[\boldsymbol{a}^{\rm T},\boldsymbol{b}^{\rm T},\boldsymbol{\gamma}^{\rm T},\boldsymbol{c}^{\rm T}]^{\rm T}$ 在时刻 t 的估计.用 $\boldsymbol{\gamma}$ 的估计 $\hat{\boldsymbol{\gamma}}(t)$ 计算未知中间变量 $\bar{u}(t)$ 的估计：

$$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t).$$

用参数向量 $\boldsymbol{c}$ 的估计 $\hat{\boldsymbol{c}}(t):=[\hat{c}_1(t),\hat{c}_2(t),\cdots,\hat{c}_{n_c}(t)]^{\rm T}\in\mathbf{R}^{n_c}$ 来构造多项式 $C(z)$ 的估计：

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

根据式(185)和(187)可知 $\bar{u}_{\rm f}(t)$ 和 $y_{\rm f}(t)$ 的估计 $\hat{\bar{u}}_{\rm f}(t)$ 和 $\hat{y}_{\rm f}(t)$ 可用下式计算：

$$\hat{\bar{u}}_{\rm f}(t)=\hat{C}(t,z)\hat{\bar{u}}(t)=\hat{c}_1(t)\hat{\bar{u}}(t-1)+\hat{c}_2(t)\hat{\bar{u}}(t-2)+\cdots+\hat{c}_{n_c}(t)\hat{\bar{u}}(t-n_c)+\hat{\bar{u}}(t)=[\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_c)]\hat{\boldsymbol{c}}(t)+\hat{\bar{u}}(t),$$

$$\hat{y}_{\rm f}(t)=\hat{C}(t-1,z)y(t)=\hat{c}_1(t-1)y(t-1)+\hat{c}_2(t-1)y(t-2)+\cdots+\hat{c}_{n_c}(t-1)y(t-n_c)+y(t)=[y(t-1),y(t-2),\cdots,y(t-n_c)]\hat{\boldsymbol{c}}(t-1)+y(t).$$

这两式的多项式滤波使用了不同时刻的估计，是为了使递推计算可以实现辨识算法.因为人们总是希望随着数据长度的增加，参数估计收敛于真参数，对于大 t，相邻时刻的参数估计一般很接近，这在仿真中也得到了验证.

作了以上准备，根据式(189)，定义准则函数

$$J_8(\boldsymbol{\theta}):=[y_{\rm f}(t)-\boldsymbol{\varphi}_{\rm f}^{\rm T}(t)\boldsymbol{\theta}]^2,$$

使用梯度搜索，极小化 $J_8(\boldsymbol{\theta})$，相关未知变量用其估计代替，可以得到第2种辨识IN-CARAR系统参数向量 $\boldsymbol{\theta}$ 的基于数据滤波的随机梯度算法(data Filtering based Stochastic Gradient algorithm，F-SG算法)：

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\hat{\boldsymbol{\varphi}}_{\rm f}(t)}{r(t)}e_{\rm f}(t),\tag{190}$$

$$e_{\rm f}(t)=\hat{y}_{\rm f}(t)-\hat{\boldsymbol{\varphi}}_{\rm f}^{\rm T}(t)\hat{\boldsymbol{\theta}}(t-1),\tag{191}$$

$$r(t)=r(t-1)+\|\hat{\boldsymbol{\varphi}}_{\rm f}(t)\|^2,\quad r(0)=1,\tag{192}$$

$$\hat{\boldsymbol{\varphi}}_{\rm f}(t)=[-\hat{y}_{\rm f}(t-1),-\hat{y}_{\rm f}(t-2),\cdots,-\hat{y}_{\rm f}(t-n_a),\hat{\bar{u}}_{\rm f}(t-1),\hat{\bar{u}}_{\rm f}(t-2),\cdots,\hat{\bar{u}}_{\rm f}(t-n_b),\boldsymbol{f}(u(t)),\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_c)]^{\rm T},\tag{193}$$

$$\hat{y}_{\rm f}(t)=[y(t-1),y(t-2),\cdots,y(t-n_c)]\hat{\boldsymbol{c}}(t-1)+y(t),\tag{194}$$

$$\hat{\bar{u}}_{\rm f}(t)=[\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_c)]\hat{\boldsymbol{c}}(t)+\hat{\bar{u}}(t),\tag{195}$$

$$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t),\tag{196}$$

$$\boldsymbol{f}(u(t))=[f_1(u(t)),f_2(u(t)),\cdots,f_{n_\gamma}(u(t))],\tag{197}$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{a}}^{\rm T}(t),\hat{\boldsymbol{b}}^{\rm T}(t),\hat{\boldsymbol{\gamma}}^{\rm T}(t),\hat{\boldsymbol{c}}^{\rm T}(t)]^{\rm T}.\tag{198}$$

注15 这个F-SG算法(190)—(198)比上节的F-SG算法(109)—(125)要简单，不同的是这个算法采用2个不同时刻多项式 $C(z)$ 的估计进行滤波，系统的所有参数都集中在一个参数向量中，而前面的算法是分为2个子参数向量，进行交互估计.

4.2 基于滤波的多新息随机梯度辨识方法

定义堆积滤波信息矩阵 $\hat{\boldsymbol{\Phi}}_{\rm f}(p,t)$ 和堆积滤波输出向量 $\boldsymbol{Y}_{\rm f}(p,t)$ 分别为

$$\hat{\boldsymbol{\Phi}}_{\rm f}(p,t):=[\hat{\boldsymbol{\varphi}}_{\rm f}(t),\hat{\boldsymbol{\varphi}}_{\rm f}(t-1),\cdots,\hat{\boldsymbol{\varphi}}_{\rm f}(t-p+1)]\in\mathbf{R}^{n_0\times p},$$

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

扩展式(191)中标量新息 $e_{\rm f}(t)\in\mathbf{R}$ 为新息向量

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

可以得到第 2 种辨识 IN-CARAR 系统参数向量 $\boldsymbol{\theta}$ 的基于数据滤波的多新息随机梯度算法(data Filtering based Multi-Innovation Stochastic Gradient algorithm, F-MISG 算法):

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\hat{\boldsymbol{\Phi}}_{\rm f}(p,t)}{r(t)}\boldsymbol{E}_{\rm f}(p,t),\tag{199}$$

$$\boldsymbol{E}_{\rm f}(p,t)=\boldsymbol{Y}_{\rm f}(p,t)-\hat{\boldsymbol{\Phi}}_{\rm f}^{\rm T}(p,t)\hat{\boldsymbol{\theta}}(t-1),\tag{200}$$

$$r(t)=r(t-1)+\|\hat{\boldsymbol{\varphi}}_{\rm f}(t)\|^{2},\quad r(0)=1,\tag{201}$$

$$\boldsymbol{Y}_{\rm f}(p,t)=[\hat{y}_{\rm f}(t),\hat{y}_{\rm f}(t-1),\cdots,\hat{y}_{\rm f}(t-p+1)]^{\rm T},\tag{202}$$

$$\hat{\boldsymbol{\Phi}}_{\rm f}(p,t)=[\hat{\boldsymbol{\varphi}}_{\rm f}(t),\hat{\boldsymbol{\varphi}}_{\rm f}(t-1),\cdots,\hat{\boldsymbol{\varphi}}_{\rm f}(t-p+1)],\tag{203}$$

$$\begin{aligned}\hat{\boldsymbol{\varphi}}_{\rm f}(t)=[&-\hat{y}_{\rm f}(t-1),-\hat{y}_{\rm f}(t-2),\cdots,-\hat{y}_{\rm f}(t-n_a),\\&\hat{\bar{u}}_{\rm f}(t-1),\hat{\bar{u}}_{\rm f}(t-2),\cdots,\hat{\bar{u}}_{\rm f}(t-n_b),\boldsymbol{f}(u(t)),\\&\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_c)]^{\rm T},\end{aligned}\tag{204}$$

$$\hat{y}_{\rm f}(t)=[y(t-1),y(t-2),\cdots,y(t-n_c)]\hat{\boldsymbol{c}}(t-1)+y(t),\tag{205}$$

$$\hat{\bar{u}}_{\rm f}(t)=[\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_c)]\hat{\boldsymbol{c}}(t)+\hat{\bar{u}}(t),\tag{206}$$

$$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t),\tag{207}$$

$$\boldsymbol{f}(u(t))=[f_1(u(t)),f_2(u(t)),\cdots,f_{n_\gamma}(u(t))],\tag{208}$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{a}}^{\rm T}(t),\hat{\boldsymbol{b}}^{\rm T}(t),\hat{\boldsymbol{\gamma}}^{\rm T}(t),\hat{\boldsymbol{c}}^{\rm T}(t)]^{\rm T}.\tag{209}$$

当新息长度 $p=1$ 时,F-MISG 算法退化为 F-SG 算法(190)—(198).F-MISG 算法(199)—(209)的计算步骤如下:

1)初始化:令 $t=1$.置初值 $\hat{\boldsymbol{\theta}}(0)=\mathbf{1}_{n_0}/p_0$,$r(0)=1$,$\hat{y}_{\rm f}(-i)=1/p_0$,$\hat{\bar{u}}_{\rm f}(-i)=1/p_0$,$\hat{\bar{u}}(-i)=1/p_0$,$i=0,1,\cdots,n_0$,$p_0=10^6$,给定基函数 $f_i(\cdot)$ 和新息长度 p.

2)收集数据 $u(t)$ 和 $y(t)$,用式(208)构造基函数行向量 $\boldsymbol{f}(u(t))$.

3)用式(205)计算滤波输出 $\hat{y}_{\rm f}(t)$,用式(202)构造堆积滤波输出向量 $\boldsymbol{Y}_{\rm f}(p,t)$.

4)用式(204)构造信息向量 $\hat{\boldsymbol{\varphi}}_{\rm f}(t)$,用式(203)构造信息矩阵 $\hat{\boldsymbol{\Phi}}_{\rm f}(p,t)$.

5)用式(200)计算新息向量 $\boldsymbol{E}_{\rm f}(p,t)$,用式(201)计算 $r(t)$.

6)根据式(199)刷新参数估计向量 $\hat{\boldsymbol{\theta}}(t)$.

7)从式(209)的 $\hat{\boldsymbol{\theta}}(t)$ 中读出 $\hat{\boldsymbol{\gamma}}(t)$ 和 $\hat{\boldsymbol{c}}(t)$.

8)用式(207)计算 $\hat{\bar{u}}(t)$,用式(206)计算滤波输入 $\hat{\bar{u}}_{\rm f}(t)$.

9)t 增 1,转到第 2)步.

F-MISG 算法(199)—(209)计算系统参数估计 $\hat{\boldsymbol{\theta}}(t)$ 的流程如图 3 所示.

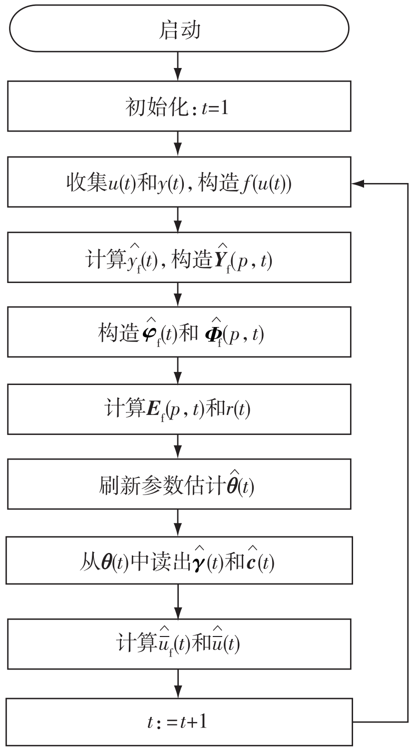

图 3 计算 F-MISG 参数估计 $\hat{\boldsymbol{\theta}}(t)$ 的流程

Fig. 3 The flowchart of computing the F-MISG parameter estimate $\hat{\boldsymbol{\theta}}(t)$

4.3 基于滤波的递推最小二乘辨识方法

对于辨识模型(189),借助于辅助模型思想,可以得到第 2 种辨识 IN-CARAR 系统参数向量 $\boldsymbol{\theta}$ 的基于数据滤波的递推最小二乘算法(data Filtering based Recursive Least Squares algorithm, F-RLS 算法):

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

$$\boldsymbol{L}(t)=\frac{\boldsymbol{P}(t-1)\hat{\boldsymbol{\varphi}}_{\rm f}(t)}{1+\hat{\boldsymbol{\varphi}}_{\rm f}^{\rm T}(t)\boldsymbol{P}(t-1)\hat{\boldsymbol{\varphi}}_{\rm f}(t)},\tag{211}$$

$$\boldsymbol{P}(t)=[\boldsymbol{I}_{n_0}-\boldsymbol{L}(t)\hat{\boldsymbol{\varphi}}_{\rm f}^{\rm T}(t)]\boldsymbol{P}(t-1),\quad \boldsymbol{P}(0)=p_0\boldsymbol{I}_{n_0},\tag{212}$$

$$\begin{aligned}\hat{\boldsymbol{\varphi}}_{\rm f}(t)=[&-\hat{y}_{\rm f}(t-1),-\hat{y}_{\rm f}(t-2),\cdots,-\hat{y}_{\rm f}(t-n_a),\\&\hat{\bar{u}}_{\rm f}(t-1),\hat{\bar{u}}_{\rm f}(t-2),\cdots,\hat{\bar{u}}_{\rm f}(t-n_b),\boldsymbol{f}(u(t)),\\&\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_c)]^{\rm T},\end{aligned}\tag{213}$$

$$\hat{y}_{\rm f}(t)=[y(t-1),y(t-2),\cdots,y(t-n_c)]\hat{\boldsymbol{c}}(t-1)+y(t),\tag{214}$$

$$\hat{\bar{u}}_{\rm f}(t)=[\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_c)]\hat{\boldsymbol{c}}(t)+\hat{\bar{u}}(t),\tag{215}$$

$$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t),\tag{216}$$

$$\boldsymbol{f}(u(t))=[f_1(u(t)),f_2(u(t)),\cdots,f_{n_\gamma}(u(t))],\tag{217}$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{a}}^{\rm T}(t),\hat{\boldsymbol{b}}^{\rm T}(t),\hat{\boldsymbol{\gamma}}^{\rm T}(t),\hat{\boldsymbol{c}}^{\rm T}(t)]^{\rm T}.\tag{218}$$

注 18 这个 F-RLS 算法比 F-SG 算法(190)—(198)有更快的收敛速度，在相同数据长度下(即相同递推步数下)，前者能给出更高精度的参数估计.

4.4 基于滤波的多新息最小二乘辨识方法

扩展式(210)中标量新息 $e_{\mathrm{f}}(t):=\hat{y}_{\mathrm{f}}(t)-\hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\mathrm{T}}(t)\cdot\hat{\boldsymbol{\theta}}(t-1)\in\mathbf{R}$ 为新息向量

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

可以得到第 2 种辨识 IN-CARAR 系统参数向量 $\boldsymbol{\theta}$ 的基于数据滤波的多新息最小二乘算法(data Filtering based Multi-Innovation Least Squares algorithm，F-MILS 算法)：

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{L}(t)[\boldsymbol{Y}_{\mathrm{f}}(p,t)-\hat{\boldsymbol{\Phi}}_{\mathrm{f}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\theta}}(t-1)],\quad(219)$$

$$\boldsymbol{L}(t)=\frac{\boldsymbol{P}(t-1)\hat{\boldsymbol{\Phi}}_{\mathrm{f}}(p,t)}{\boldsymbol{I}_p+\hat{\boldsymbol{\Phi}}_{\mathrm{f}}^{\mathrm{T}}(p,t)\boldsymbol{P}(t-1)\hat{\boldsymbol{\Phi}}_{\mathrm{f}}(p,t)},\quad(220)$$

$$\boldsymbol{P}(t)=\boldsymbol{P}(t-1)-\boldsymbol{L}(t)\hat{\boldsymbol{\Phi}}_{\mathrm{f}}^{\mathrm{T}}(p,t)\boldsymbol{P}(t-1),\quad \boldsymbol{P}(0)=p_0\boldsymbol{I}_{n_0},\quad(221)$$

$$\hat{\boldsymbol{Y}}_{\mathrm{f}}(p,t)=[\hat{y}_{\mathrm{f}}(t),\hat{y}_{\mathrm{f}}(t-1),\cdots,\hat{y}_{\mathrm{f}}(t-p+1)]^{\mathrm{T}},\quad(222)$$

$$\hat{\boldsymbol{\Phi}}_{\mathrm{f}}(p,t)=[\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t),\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-1),\cdots,\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-p+1)],\quad(223)$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)=[-\hat{y}_{\mathrm{f}}(t-1),-\hat{y}_{\mathrm{f}}(t-2),\cdots,-\hat{y}_{\mathrm{f}}(t-n_a),\hat{\bar{u}}_{\mathrm{f}}(t-1),\hat{\bar{u}}_{\mathrm{f}}(t-2),\cdots,\hat{\bar{u}}_{\mathrm{f}}(t-n_b),\boldsymbol{f}(u(t)),\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_c)]^{\mathrm{T}},\quad(224)$$

$$\hat{y}_{\mathrm{f}}(t)=[y(t-1),y(t-2),\cdots,y(t-n_c)]\hat{\boldsymbol{c}}(t-1)+y(t),\quad(225)$$

$$\hat{\bar{u}}_{\mathrm{f}}(t)=[\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_c)]\hat{\boldsymbol{c}}(t)+\hat{\bar{u}}(t),\quad(226)$$

$$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t),\quad(227)$$

$$\boldsymbol{f}(u(t))=[f_1(u(t)),f_2(u(t)),\cdots,f_{n_\gamma}(u(t))],\quad(228)$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{b}}^{\mathrm{T}}(t),\hat{\boldsymbol{\gamma}}^{\mathrm{T}}(t),\hat{\boldsymbol{c}}^{\mathrm{T}}(t)]^{\mathrm{T}}.\quad(229)$$

当新息长度 $p=1$ 时，F-MILS 算法退化为 F-RLS 算法(210)—(218).

5 基于数据滤波的多新息辨识方法(3)

本节讨论第 3 种基于数据滤波的辨识算法.该算法与第 2 种类似，也是采用一个辨识模型，即一个参数向量，不需要交互计算，就能递推计算系统的参数估计，简化了算法的步骤，但对输入和输出数据滤波的多项式都是当前时刻的估计.

5.1 基于滤波的随机梯度辨识方法

将式(187)代入式(188)左边的 $y_{\mathrm{f}}(t)$ 可得

$$y(t)+\sum_{i=1}^{n_c}c_iy(t-i)=-\sum_{i=1}^{n_a}a_iy_{\mathrm{f}}(t-i)+\sum_{i=1}^{n_b}b_i\bar{u}_{\mathrm{f}}(t-i)+\sum_{i=1}^{n_\gamma}\gamma_if_i(u(t))+\sum_{i=1}^{n_c}c_i\bar{u}(t-i)+v(t).$$

移项得到滤波辨识模型：

$$y(t)=-\sum_{i=1}^{n_a}a_iy_{\mathrm{f}}(t-i)+\sum_{i=1}^{n_b}b_i\bar{u}_{\mathrm{f}}(t-i)+\sum_{i=1}^{n_\gamma}\boldsymbol{\gamma}_if_i(u(t))+\sum_{i=1}^{n_c}c_i[\bar{u}(t-i)-y(t-i)]+v(t)=:\boldsymbol{\varphi}_{\mathrm{f}}^{\mathrm{T}}(t)\boldsymbol{\theta}+v(t),\quad(230)$$

$$\boldsymbol{\varphi}_{\mathrm{f}}(t):=[-y_{\mathrm{f}}(t-1),-y_{\mathrm{f}}(t-2),\cdots,-y_{\mathrm{f}}(t-n\_a),\bar{u}_{\mathrm{f}}(t-1),\bar{u}_{\mathrm{f}}(t-2),\cdots,\bar{u}_{\mathrm{f}}(t-n\_b),\boldsymbol{f}(u(t)),\bar{u}(t-1)-y(t-1),\bar{u}(t-2)-y(t-2),\cdots,\bar{u}(t-n_c)-y(t-n_c)]^{\mathrm{T}}\in\mathbf{R}^{n_0},$$

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

其中参数向量 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 和 $\boldsymbol{\gamma}$ 的定义同上.

对于辨识模型(230)，参考第 2 种基于数据滤波的随机梯度算法(190)—(198)的推导过程，可以得到第 3 种辨识 IN-CARAR 系统参数向量 $\boldsymbol{\theta}$ 的基于数据滤波的随机梯度算法(data Filtering based Stochastic Gradient algorithm，F-SG 算法)：

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)}{r(t)}e_{\mathrm{f}}(t),\quad(231)$$

$$e_{\mathrm{f}}(t)=y(t)-\hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1),\quad(232)$$

$$r(t)=r(t-1)+\|\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)\|^2,\quad r(0)=1,\quad(233)$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)=[-\hat{y}_{\mathrm{f}}(t-1),-\hat{y}_{\mathrm{f}}(t-2),\cdots,-\hat{y}_{\mathrm{f}}(t-n_a),\hat{\bar{u}}_{\mathrm{f}}(t-1),\hat{\bar{u}}_{\mathrm{f}}(t-2),\cdots,\hat{\bar{u}}_{\mathrm{f}}(t-n_b),\boldsymbol{f}(u(t)),\hat{\bar{u}}(t-1)-y(t-1),\hat{\bar{u}}(t-2)-y(t-2),\cdots,\hat{\bar{u}}(t-n_c)-y(t-n_c)]^{\mathrm{T}},\quad(234)$$

$$\hat{y}_{\mathrm{f}}(t)=[y(t-1),y(t-2),\cdots,y(t-n_c)]\hat{\boldsymbol{c}}(t)+y(t),\quad(235)$$

$$\hat{\bar{u}}_{\mathrm{f}}(t)=[\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_c)]\hat{\boldsymbol{c}}(t)+\hat{\bar{u}}(t),\quad(236)$$

$$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t),\quad(237)$$

$$\boldsymbol{f}(u(t))=[f_1(u(t)),f_2(u(t)),\cdots,f_{n_\gamma}(u(t))],\quad(238)$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{b}}^{\mathrm{T}}(t),\hat{\boldsymbol{\gamma}}^{\mathrm{T}}(t),\hat{\boldsymbol{c}}^{\mathrm{T}}(t)]^{\mathrm{T}}.\quad(239)$$

5.2 基于滤波的多新息随机梯度辨识方法

定义堆积滤波信息矩阵 $\hat{\boldsymbol{\Phi}}_{\mathrm{f}}(p,t)$ 和堆积输出向量 $\boldsymbol{Y}(p,t)$ 分别为

$$\hat{\boldsymbol{\Phi}}_{\mathrm{f}}(p,t):=[\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t),\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-1),\cdots,\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-p+1)]\in\mathbf{R}^{n_0\times p},$$

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

扩展式(232)中标量新息 $e_{\mathrm{f}}(t)\in\mathbf{R}$ 为新息向量

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

可以得到第 3 种辨识 IN-CARAR 系统参数向量 $\boldsymbol{\theta}$ 的基于数据滤波的多新息随机梯度算法(data Filtering based Multi-Innovation Stochastic Gradient algorithm, F-MISG 算法):

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\hat{\boldsymbol{\Phi}}_{\mathrm{f}}(p,t)}{r(t)}\boldsymbol{E}_{\mathrm{f}}(p,t), \tag{240}$$

$$\boldsymbol{E}_{\mathrm{f}}(p,t)=\boldsymbol{Y}(p,t)-\hat{\boldsymbol{\Phi}}_{\mathrm{f}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\theta}}(t-1), \tag{241}$$

$$r(t)=r(t-1)+\|\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)\|^2,\quad r(0)=1, \tag{242}$$

$$\boldsymbol{Y}(p,t)=[y(t),y(t-1),\cdots,y(t-p+1)], \tag{243}$$

$$\hat{\boldsymbol{\Phi}}_{\mathrm{f}}(p,t)=[\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t),\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-1),\cdots,\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-p+1)], \tag{244}$$

$$\begin{aligned}\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)=[&-\hat{y}_{\mathrm{f}}(t-1),-\hat{y}_{\mathrm{f}}(t-2),\cdots,-\hat{y}_{\mathrm{f}}(t-n_a),\\&\hat{\bar{u}}_{\mathrm{f}}(t-1),\hat{\bar{u}}_{\mathrm{f}}(t-2),\cdots,\hat{\bar{u}}_{\mathrm{f}}(t-n_b),\boldsymbol{f}(u(t)),\\&\hat{\bar{u}}(t-1)-y(t-1),\hat{\bar{u}}(t-2)-y(t-2),\cdots,\\&\hat{\bar{u}}(t-n_c)-y(t-n_c)]^{\mathrm{T}},\end{aligned} \tag{245}$$

$$\hat{y}_{\mathrm{f}}(t)=[y(t-1),y(t-2),\cdots,y(t-n_c)]\hat{\boldsymbol{c}}(t)+y(t), \tag{246}$$

$$\hat{\bar{u}}_{\mathrm{f}}(t)=[\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_c)]\hat{\boldsymbol{c}}(t)+\hat{\bar{u}}(t), \tag{247}$$

$$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t), \tag{248}$$

$$\boldsymbol{f}(u(t))=[f_1(u(t)),f_2(u(t)),\cdots,f_{n_\gamma}(u(t))], \tag{249}$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{b}}^{\mathrm{T}}(t),\hat{\boldsymbol{\gamma}}^{\mathrm{T}}(t),\hat{\boldsymbol{c}}^{\mathrm{T}}(t)]^{\mathrm{T}}. \tag{250}$$

当新息长度 $p=1$ 时,这个 F-MISG 算法退化为 F-SG 算法(231)—(239).

5.3 基于滤波的递推最小二乘辨识方法

对于辨识模型(230),借助于辅助模型思想,可以得到第 3 种辨识 IN-CARAR 系统参数向量 $\boldsymbol{\theta}$ 的基于数据滤波的递推最小二乘算法(data Filtering based Recursive Least Squares algorithm, F-RLS 算法):

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{L}(t)[y(t)-\hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)], \tag{251}$$

$$\boldsymbol{L}(t)=\frac{\boldsymbol{P}(t-1)\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)}{1+\hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\mathrm{T}}(t)\boldsymbol{P}(t-1)\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)}, \tag{252}$$

$$\boldsymbol{P}(t)=[\boldsymbol{I}_{n_0}-\boldsymbol{L}(t)\hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\mathrm{T}}(t)]\boldsymbol{P}(t-1),\quad \boldsymbol{P}(0)=p_0\boldsymbol{I}_{n_0}, \tag{253}$$

$$\begin{aligned}\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)=[&-\hat{y}_{\mathrm{f}}(t-1),-\hat{y}_{\mathrm{f}}(t-2),\cdots,-\hat{y}_{\mathrm{f}}(t-n_a),\\&\hat{\bar{u}}_{\mathrm{f}}(t-1),\hat{\bar{u}}_{\mathrm{f}}(t-2),\cdots,\hat{\bar{u}}_{\mathrm{f}}(t-n_b),\boldsymbol{f}(u(t)),\\&\hat{\bar{u}}(t-1)-y(t-1),\hat{\bar{u}}(t-2)-y(t-2),\cdots,\\&\hat{\bar{u}}(t-n_c)-y(t-n_c)]^{\mathrm{T}},\end{aligned} \tag{254}$$

$$\hat{y}_{\mathrm{f}}(t)=[y(t-1),y(t-2),\cdots,y(t-n_c)]\hat{\boldsymbol{c}}(t)+y(t), \tag{255}$$

$$\hat{\bar{u}}_{\mathrm{f}}(t)=[\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_c)]\hat{\boldsymbol{c}}(t)+\hat{\bar{u}}(t), \tag{256}$$

$$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t), \tag{257}$$

$$\boldsymbol{f}(u(t))=[f_1(u(t)),f_2(u(t)),\cdots,f_{n_\gamma}(u(t))], \tag{258}$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{b}}^{\mathrm{T}}(t),\hat{\boldsymbol{\gamma}}^{\mathrm{T}}(t),\hat{\boldsymbol{c}}^{\mathrm{T}}(t)]^{\mathrm{T}}. \tag{259}$$

表 3 列出了第 3 种基于数据滤波的递推最小二乘算法(251)—(259)的计算量,其中 $n_0=n_a+n_b+n_\gamma+n_c$.

5.4 基于滤波的多新息最小二乘辨识方法

扩展式(251)中标量新息 $e_{\mathrm{f}}(t):=y(t)-\hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\mathrm{T}}(t)\cdot\hat{\boldsymbol{\theta}}(t-1)\in\mathbf{R}$ 为新息向量

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

可以得到第 3 种辨识 IN-CARAR 系统参数向量 $\boldsymbol{\theta}$ 的基于数据滤波的多新息最小二乘算法(data Filtering based Multi-Innovation Least Squares algorithm, F-MILS算法):

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{L}(t)[\boldsymbol{Y}(p,t)-\hat{\boldsymbol{\Phi}}_{\mathrm{f}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\theta}}(t-1)], \tag{260}$$

$$\boldsymbol{L}(t)=\frac{\boldsymbol{P}(t-1)\hat{\boldsymbol{\Phi}}_{\mathrm{f}}(p,t)}{\boldsymbol{I}_p+\hat{\boldsymbol{\Phi}}_{\mathrm{f}}^{\mathrm{T}}(p,t)\boldsymbol{P}(t-1)\hat{\boldsymbol{\Phi}}_{\mathrm{f}}(p,t)}, \tag{261}$$

$$\begin{aligned}\boldsymbol{P}(t)&=\boldsymbol{P}(t-1)-\boldsymbol{L}(t)\hat{\boldsymbol{\Phi}}_{\mathrm{f}}^{\mathrm{T}}(p,t)\boldsymbol{P}(t-1),\\&\boldsymbol{P}(0)=p_0\boldsymbol{I}_{n_0},\end{aligned} \tag{262}$$

$$\boldsymbol{Y}(p,t)=[y(t),y(t-1),\cdots,y(t-p+1)]^{\mathrm{T}}, \tag{263}$$

$$\hat{\boldsymbol{\Phi}}_{\mathrm{f}}(p,t)=[\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t),\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-1),\cdots,\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t-p+1)], \tag{264}$$

表 1 F-RLS 算法的计算量

Table 1 The computational efficiency of the F-RLS algorithm

变量	表达式	乘法次数	加法次数
$\hat{\boldsymbol{\theta}}(t)$	$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{L}(t)e_{\mathrm{f}}(t)\in\mathbf{R}^{n_0}$	n_0	n_0
	$e_{\mathrm{f}}(t):=y(t)-\hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)\in\mathbf{R}$	n_0	n_0
$\boldsymbol{L}(t)$	$\boldsymbol{L}(t)=\boldsymbol{\zeta}(t)/[1+\hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\mathrm{T}}(t)\boldsymbol{\zeta}(t)]\in\mathbf{R}^{n_0}$	$2n_0$	n_0
	$\boldsymbol{\zeta}(t):=\boldsymbol{P}(t-1)\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)\in\mathbf{R}^{n_0}$	n_0^2	$n_0^2-n_0$
$\boldsymbol{P}(t)$	$\boldsymbol{P}(t)=\boldsymbol{P}(t-1)-\boldsymbol{L}(t)\boldsymbol{\zeta}^{\mathrm{T}}(t)\in\mathbf{R}^{n_0\times n_0}$	n_0^2	n_0^2
$\hat{y}_{\mathrm{f}}(t)$	$\hat{y}_{\mathrm{f}}(t)=[y(t-1),y(t-2),\cdots,y(t-n_c)]\hat{\boldsymbol{c}}(t)+y(t)\in\mathbf{R}$	n_c	n_c
$\hat{\bar{u}}_{\mathrm{f}}(t)$	$\hat{\bar{u}}_{\mathrm{f}}(t)=[\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_c)]\hat{\boldsymbol{c}}(t)+\hat{\bar{u}}(t)\in\mathbf{R}$	n_c	n_c
$\hat{\bar{u}}(t)$	$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t)\in\mathbf{R}$	n_r	n_r-1
	总数	$2n_0^2+4n_0+2n_c+n_\gamma$	$2n_0^2+2n_0+2n_c+n_\gamma-1$
	总 flop 数	$4n_0^2+6n_0+4n_c+2n_\gamma-1$	

$$\hat{\boldsymbol{\varphi}}_{\rm f}(t)=[-\hat{y}_{\rm f}(t-1),-\hat{y}_{\rm f}(t-2),\cdots,-\hat{y}_{\rm f}(t-n_a),\hat{\bar{u}}_{\rm f}(t-1),\hat{\bar{u}}_{\rm f}(t-2),\cdots,\hat{\bar{u}}_{\rm f}(t-n_b),\boldsymbol{f}(u(t)),\hat{\bar{u}}(t-1)-y(t-1),\hat{\bar{u}}(t-2)-y(t-2),\cdots,\hat{\bar{u}}(t-n_c)-y(t-n_c)]^{\rm T},\tag{265}$$

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

$$\hat{\bar{u}}_{\rm f}(t)=[\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_c)]\hat{\boldsymbol{c}}(t)+\hat{\bar{u}}(t),\tag{267}$$

$$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t),\tag{268}$$

$$\boldsymbol{f}(u(t))=[f_1(u(t)),f_2(u(t)),\cdots,f_{n_\gamma}(u(t))],\tag{269}$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{a}}^{\rm T}(t),\hat{\boldsymbol{b}}^{\rm T}(t),\hat{\boldsymbol{\gamma}}^{\rm T}(t),\hat{\boldsymbol{c}}^{\rm T}(t)]^{\rm T}.\tag{270}$$

当新息长度 $p=1$ 时,这个 F-MILS 算法退化为 F-RLS 算法(251)—(259). F-MILS 算法(260)—(270)的计算步骤如下:

1) 初始化:令 $t=1$. 置初值 $\hat{\boldsymbol{\theta}}(0)=\mathbf{1}_{n_0}/p_0$, $\boldsymbol{P}(0)=p_0\boldsymbol{I}_{n_0}$, $\hat{y}_{\rm f}(-i)=1/p_0$, $\hat{\bar{u}}_{\rm f}(-i)=1/p_0$, $\hat{\bar{u}}(-i)=1/p_0$, $i=0,1,\cdots,n_0$, $p_0=10^6$,给定基函数 $f_i(\cdot)$ 和新息长度 p.

2) 收集数据 $u(t)$ 和 $y(t)$,用式(269)构造基函数行向量 $\boldsymbol{f}(u(t))$.

3) 用式(263)构造堆积输出向量 $\boldsymbol{Y}(p,t)$. 用式(265)构造信息向量 $\hat{\boldsymbol{\varphi}}_{\rm f}(t)$.

4) 用式(264)构造堆积滤波信息矩阵 $\hat{\boldsymbol{\Phi}}_{\rm f}(p,t)$,用式(261)计算 $\boldsymbol{L}(t)$,用式(262)计算 $\boldsymbol{P}(t)$.

5) 根据式(260)刷新参数估计向量 $\hat{\boldsymbol{\theta}}(t)$.

6) 从式(270)的 $\hat{\boldsymbol{\theta}}(t)$ 中读出 $\hat{\boldsymbol{\gamma}}(t)$ 和 $\hat{\boldsymbol{c}}(t)$.

7) 用式(266)计算滤波输出 $\hat{y}_{\rm f}(t)$,用式(267)计算滤波输入 $\hat{\bar{u}}_{\rm f}(t)$,用式(268)计算 $\hat{\bar{u}}(t)$.

8) t 增 1,转到第 2)步.

F-MILS 算法(260)—(270)计算系统参数估计 $\hat{\boldsymbol{\theta}}(t)$ 的流程如图 4 所示.

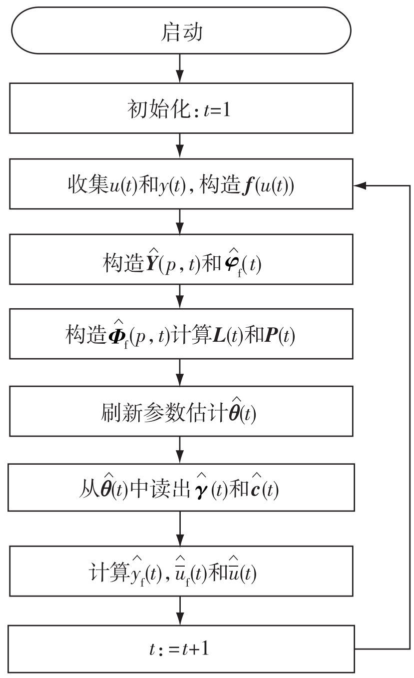

图 4 计算 F-MILS 参数估计 $\hat{\boldsymbol{\theta}}(t)$ 的流程
Fig. 4 The flowchart of computing the F-MILS parameter estimate $\hat{\boldsymbol{\theta}}(t)$

6 基于辨识模型分解的多新息辨识方法

基于辨识模型分解的辨识方法可以大幅度减小最小二乘类辨识算法的计算量. 基于辨识模型分解的辨识方法也称为递阶辨识方法、两阶段辨识方法、三阶段辨识方法或多阶段辨识方法. 基于辨识模型分解的辨识方法已经被用于线性系统[49-52]和非线性系统[24,53-55].

考虑输入非线性受控自回归自回归系统(1),重写如下:

$$A(z)y(t)=B(z)\bar{u}(t)+\frac{1}{C(z)}v(t),\tag{271}$$

$$\bar{u}(t)=f(u(t))=\gamma_1f_1(u(t))+\gamma_2f_2(u(t))+\cdots+\gamma_{n_\gamma}f_{n_\gamma}(u(t))=\boldsymbol{f}(u(t))\boldsymbol{\gamma},\tag{272}$$

$$w(t)=\frac{1}{C(z)}v(t).\tag{273}$$

定义参数向量

$$\boldsymbol{\theta}:=[\boldsymbol{a}^{\rm T},\boldsymbol{b}^{\rm T}]^{\rm T}\in\mathbf{R}^{n_a+n_b},$$
$$\boldsymbol{\vartheta}:=[\boldsymbol{\gamma}^{\rm T},\boldsymbol{c}^{\rm T}]^{\rm T}\in\mathbf{R}^{n_\gamma+n_c}$$

和信息向量

$$\boldsymbol{\varphi}(t):=[-y(t-1),-y(t-2),\cdots,-y(t-n_a),\bar{u}(t-1),\bar{u}(t-2),\cdots,\bar{u}(t-n_b)]^{\rm T}\in\mathbf{R}^{n_a+n_b},$$
$$\boldsymbol{\psi}(t):=[\boldsymbol{f}(u(t)),-w(t-1),-w(t-2),\cdots,-w(t-n_c)]^{\rm T}\in\mathbf{R}^{n_\gamma+n_c}.$$

假设多项式 $B(z)$ 的第 1 个系数 $b_0=1$,利用关键项分离原理,可以得到 IN-CARAR 系统(271)—(273)的辨识模型:

$$y(t)=[1-A(z)]y(t)+[B(z)-1]\bar{u}(t)+\bar{u}(t)+w(t)=-\sum_{i=1}^{n_a}a_iy(t-i)+\sum_{i=1}^{n_b}b_i\bar{u}(t-i)+\boldsymbol{f}(u(t))\boldsymbol{\gamma}-\sum_{i=1}^{n_c}c_iw(t-i)+v(t)=\boldsymbol{\varphi}^{\rm T}(t)\boldsymbol{\theta}+\boldsymbol{\psi}^{\rm T}(t)\boldsymbol{\vartheta}+v(t).\tag{274}$$

6.1 基于分解的随机梯度辨识方法

定义中间变量:

$y_1(t):=y(t)-\boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\vartheta}$,

$y_2(t):=y(t)-\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta}$,

则模型(274)分解为如下 2 个子模型:

$$y_1(t):=\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta}+v(t), \tag{275}$$

$$y_2(t):=\boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\vartheta}+v(t). \tag{276}$$

设 $\hat{\boldsymbol{\theta}}(t)$ 和 $\hat{\boldsymbol{\vartheta}}(t)$ 分别是参数向量 $\boldsymbol{\theta}$ 和 $\boldsymbol{\vartheta}$ 在时刻 t 的估计.对于辨识模型(275)和(276),使用负梯度搜索,极小化准则函数

$J_9(\boldsymbol{\theta}):=[y_1(t)-\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta}]^2$,

$J_{10}(\boldsymbol{\vartheta}):=[y_2(t)-\boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\vartheta}]^2$,

可以得到下列递推关系:

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\boldsymbol{\varphi}(t)}{r_1(t)}[y_1(t)-\boldsymbol{\varphi}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)]=$$
$$\hat{\boldsymbol{\theta}}(t-1)+\frac{\boldsymbol{\varphi}(t)}{r_1(t)}[y(t)-\boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\vartheta}-\boldsymbol{\varphi}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)], \tag{277}$$

$$r_1(t)=r_1(t-1)+\|\boldsymbol{\varphi}(t)\|^2,\quad r_1(0)=1, \tag{278}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\frac{\boldsymbol{\psi}(t)}{r_2(t)}[y_2(t)-\boldsymbol{\psi}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t-1)]=$$
$$\hat{\boldsymbol{\vartheta}}(t-1)+\frac{\boldsymbol{\psi}(t)}{r_2(t)}[y(t)-\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta}-\boldsymbol{\psi}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t-1)], \tag{279}$$

$$r_2(t)=r_2(t-1)+\|\boldsymbol{\psi}(t)\|^2,\quad r_2(0)=1. \tag{280}$$

借助于辅助模型思想,式(277)—(280)中右边未知量 $\boldsymbol{\varphi}(t)$, $\boldsymbol{\vartheta}$, $\boldsymbol{\psi}(t)$ 和 $\boldsymbol{\theta}$ 用其估计代替,可以得到辨识 IN-CARAR 系统参数向量的基于分解的随机梯度算法(Decomposition based Stochastic Gradient algorithm, D-SG 算法):

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\hat{\boldsymbol{\varphi}}(t)}{r_1(t)}e(t), \tag{281}$$

$$e(t)=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)-\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t-1), \tag{282}$$

$$r_1(t)=r_1(t-1)+\|\hat{\boldsymbol{\varphi}}(t)\|^2,\quad r_1(0)=1, \tag{283}$$

$$\hat{\boldsymbol{\varphi}}(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a),\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_b)]^{\mathrm{T}}, \tag{284}$$

$$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t), \tag{285}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\frac{\hat{\boldsymbol{\psi}}(t)}{r_2(t)}e(t), \tag{286}$$

$$r_2(t)=r_2(t-1)+\|\hat{\boldsymbol{\psi}}(t)\|^2,\quad r_2(0)=1, \tag{287}$$

$$\hat{\boldsymbol{\psi}}(t)=[\boldsymbol{f}(u(t)),-\hat{w}(t-1),-\hat{w}(t-2),\cdots,-\hat{w}(t-n_c)]^{\mathrm{T}}, \tag{288}$$

$$\hat{w}(t)=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t)-\hat{\bar{u}}(t), \tag{289}$$

$$\boldsymbol{f}(u(t))=[f_1(u(t)),f_2(u(t)),\cdots,f_{n_\gamma}(u(t))], \tag{290}$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{b}}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{291}$$

$$\hat{\boldsymbol{\vartheta}}(t)=[\hat{\boldsymbol{\gamma}}^{\mathrm{T}}(t),\hat{\boldsymbol{c}}^{\mathrm{T}}(t)]^{\mathrm{T}}. \tag{292}$$

6.2 基于分解的多新息随机梯度辨识方法

定义堆积信息矩阵 $\boldsymbol{\Phi}(p,t)$ 和 $\hat{\boldsymbol{\Psi}}(p,t)$ 和堆积输出向量 $\boldsymbol{Y}(p,t)$ 分别为

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

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

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

扩展式(282)中标量新息 $e(t)\in\mathbf{R}$ 为新息向量

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

可以得到辨识 IN-CARAR 系统参数向量的基于分解的多新息随机梯度算法(Decomposition based Multi-innovation Stochastic Gradient algorithm, D-MISG 算法):

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\hat{\boldsymbol{\Phi}}(p,t)}{r_1(t)}\boldsymbol{E}(p,t), \tag{293}$$

$$\boldsymbol{E}(p,t)=\boldsymbol{Y}(p,t)-\hat{\boldsymbol{\Phi}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\theta}}(t-1)-\hat{\boldsymbol{\Psi}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\vartheta}}(t-1), \tag{294}$$

$$r_1(t)=r_1(t-1)+\|\hat{\boldsymbol{\varphi}}(t)\|^2,\quad r_1(0)=1, \tag{295}$$

$$\hat{\boldsymbol{\Phi}}(p,t)=[\hat{\boldsymbol{\varphi}}(t),\hat{\boldsymbol{\varphi}}(t-1),\cdots,\hat{\boldsymbol{\varphi}}(t-p+1)], \tag{296}$$

$$\boldsymbol{Y}(p,t)=[y(t),y(t-1),\cdots,y(t-p+1)]^{\mathrm{T}}, \tag{297}$$

$$\hat{\boldsymbol{\varphi}}(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a),\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_b)]^{\mathrm{T}}, \tag{298}$$

$$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t), \tag{299}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\frac{\hat{\boldsymbol{\Psi}}(p,t)}{r_2(t)}\boldsymbol{E}(p,t), \tag{300}$$

$$r_2(t)=r_2(t-1)+\|\hat{\boldsymbol{\psi}}(t)\|^2,\quad r_2(0)=1, \tag{301}$$

$$\hat{\boldsymbol{\Psi}}(p,t)=[\hat{\boldsymbol{\psi}}(t),\hat{\boldsymbol{\psi}}(t-1),\cdots,\hat{\boldsymbol{\psi}}(t-p+1)], \tag{302}$$

$$\hat{\boldsymbol{\psi}}(t)=[\boldsymbol{f}(u(t)),-\hat{w}(t-1),-\hat{w}(t-2),\cdots,-\hat{w}(t-n_c)]^{\mathrm{T}}, \tag{303}$$

$$\hat{w}(t)=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t)-\hat{\bar{u}}(t), \tag{304}$$

$$\boldsymbol{f}(u(t))=[f_1(u(t)),f_2(u(t)),\cdots,f_{n_\gamma}(u(t))], \tag{305}$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{b}}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{306}$$

$$\hat{\boldsymbol{\vartheta}}(t)=[\hat{\boldsymbol{\gamma}}^{\mathrm{T}}(t),\hat{\boldsymbol{c}}^{\mathrm{T}}(t)]^{\mathrm{T}}. \tag{307}$$

当新息长度 $p=1$ 时, D-MISG 算法退化为 D-SG 算法(281)—(292).

6.3 基于分解的递推最小二乘辨识方法

对于辨识模型(275)和(276),使用最小二乘原理,极小化准则函数

$$J_{11}(\boldsymbol{\theta}):=\sum_{j=1}^{t}[y_1(j)-\boldsymbol{\varphi}^{\mathrm{T}}(j)\boldsymbol{\theta}]^2=\sum_{j=1}^{t}[y(j)-\boldsymbol{\psi}^{\mathrm{T}}(j)\boldsymbol{\vartheta}-\boldsymbol{\varphi}^{\mathrm{T}}(j)\boldsymbol{\theta}]^2,$$

$$J_{12}(\boldsymbol{\vartheta}):=\sum_{j=1}^{t}[y_2(j)-\boldsymbol{\psi}^{\mathrm{T}}(j)\boldsymbol{\vartheta}]^2=\sum_{j=1}^{t}[y(j)-\boldsymbol{\varphi}^{\mathrm{T}}(j)\boldsymbol{\theta}-\boldsymbol{\psi}^{\mathrm{T}}(j)\boldsymbol{\vartheta}]^2,$$

并借助于辅助模型思想，未知量 $\boldsymbol{\varphi}(t)$，$\boldsymbol{\vartheta}$，$\boldsymbol{\psi}(t)$ 和 $\boldsymbol{\theta}$ 用其估计代替，可以得到辨识 IN-CARAR 系统参数向量的基于分解的递推最小二乘算法（Decomposition based Recursive Least Squares algorithm，D-RLS 算法）：

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{L}_1(t)[y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)-\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t-1)], \tag{308}$$

$$\boldsymbol{L}_1(t)=\frac{\boldsymbol{P}_1(t-1)\hat{\boldsymbol{\varphi}}(t)}{1+\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\boldsymbol{P}_1(t-1)\hat{\boldsymbol{\varphi}}(t)}, \tag{309}$$

$$\boldsymbol{P}_1(t)=[\boldsymbol{I}_{n_a+n_b}-\boldsymbol{L}_1(t)\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)]\boldsymbol{P}_1(t-1),\quad \boldsymbol{P}_1(0)=p_0\boldsymbol{I}_{n_a+n_b}, \tag{310}$$

$$\hat{\boldsymbol{\varphi}}(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a),\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_b)]^{\mathrm{T}}, \tag{311}$$

$$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t), \tag{312}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\boldsymbol{L}_2(t)[y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)-\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t-1)], \tag{313}$$

$$\boldsymbol{L}_2(t)=\frac{\boldsymbol{P}_2(t-1)\hat{\boldsymbol{\psi}}(t)}{1+\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)\boldsymbol{P}_2(t-1)\hat{\boldsymbol{\psi}}(t)}, \tag{314}$$

$$\boldsymbol{P}_2(t)=[\boldsymbol{I}_{n_\gamma+n_c}-\boldsymbol{L}_2(t)\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)]\boldsymbol{P}_2(t-1),\quad \boldsymbol{P}_2(0)=p_0\boldsymbol{I}_{n_\gamma+n_c}, \tag{315}$$

$$\hat{\boldsymbol{\psi}}(t)=[\boldsymbol{f}(u(t)),-\hat{w}(t-1),-\hat{w}(t-2),\cdots,-\hat{w}(t-n_c)]^{\mathrm{T}}, \tag{316}$$

$$\hat{w}(t)=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t)-\hat{\bar{u}}(t), \tag{317}$$

$$\boldsymbol{f}(u(t))=[f_1(u(t)),f_2(u(t)),\cdots,f_{n_\gamma}(u(t))], \tag{318}$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{b}}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{319}$$

$$\hat{\boldsymbol{\vartheta}}(t)=[\hat{\boldsymbol{\gamma}}^{\mathrm{T}}(t),\hat{\boldsymbol{c}}^{\mathrm{T}}(t)]^{\mathrm{T}}. \tag{320}$$

表 4 列出了基于分解的递推最小二乘辨识方法(308)—(320)的计算量，其中 $n_0=n_1+n_2$，$n_1:=n_a+n_b$，$n_2:=n_\gamma+n_c$.

6.4 基于分解的多新息最小二乘辨识方法

扩展式(308)和式(313)中标量新息 $e(t):=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)-\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t-1)\in\mathbf{R}$ 为新息向量 $\boldsymbol{E}(p,t):=\boldsymbol{Y}(p,t)-\hat{\boldsymbol{\Phi}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\theta}}(t-1)-\hat{\boldsymbol{\Psi}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\vartheta}}(t-1)\in\mathbf{R}^p$，可以得到辨识 IN-CARAR 系统参数向量的基于分解的多新息最小二乘算法（Decomposition based Multi-innovation Least Squares algorithm，D-MILS 算法）：

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{L}_1(t)[\boldsymbol{Y}(p,t)-\hat{\boldsymbol{\Phi}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\theta}}(t-1)-\hat{\boldsymbol{\Psi}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\vartheta}}(t-1)], \tag{321}$$

$$\boldsymbol{L}_1(t)=\frac{\boldsymbol{P}_1(t-1)\hat{\boldsymbol{\Phi}}(p,t)}{\boldsymbol{I}_p+\hat{\boldsymbol{\Phi}}^{\mathrm{T}}(p,t)\boldsymbol{P}_1(t-1)\hat{\boldsymbol{\Phi}}(p,t)}, \tag{322}$$

$$\boldsymbol{P}_1(t)=\boldsymbol{P}_1(t-1)-\boldsymbol{L}_1(t)\hat{\boldsymbol{\Phi}}^{\mathrm{T}}(p,t)\boldsymbol{P}_1(t-1),\quad \boldsymbol{P}_1(0)=p_0\boldsymbol{I}_{n_a+n_b}, \tag{323}$$

$$\hat{\boldsymbol{\Phi}}(p,t)=[\hat{\boldsymbol{\varphi}}(t),\hat{\boldsymbol{\varphi}}(t-1),\cdots,\hat{\boldsymbol{\varphi}}(t-p+1)], \tag{324}$$

$$\boldsymbol{Y}(p,t)=[y(t),y(t-1),\cdots,y(t-p+1)]^{\mathrm{T}}, \tag{325}$$

$$\hat{\boldsymbol{\varphi}}(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a),\hat{\bar{u}}(t-1),\hat{\bar{u}}(t-2),\cdots,\hat{\bar{u}}(t-n_b)]^{\mathrm{T}}, \tag{326}$$

表 4 D-RLS 算法的计算量

Table 4 The computational efficiency of the D-RLS algorithm

变量	表达式	乘法次数	加法次数
$\hat{\boldsymbol{\theta}}(t)$	$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(t-1)+\boldsymbol{L}_1(t)e(t)\in\mathbf{R}^{n_1}$	n_1	n_1
	$e(t):=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)-\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}(t-1)\in\mathbf{R}$	n_1	n_1
$\boldsymbol{L}_1(t)$	$\boldsymbol{L}_1(t)=\boldsymbol{\zeta}_1(t)/[1+\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\boldsymbol{\zeta}_1(t)]\in\mathbf{R}^{n_1}$	$2n_1$	n_1
	$\boldsymbol{\zeta}_1(t):=\boldsymbol{P}_1(t-1)\hat{\boldsymbol{\varphi}}(t)\in\mathbf{R}^{n_1}$	n_1^2	$n_1^2-n_1$
$\boldsymbol{P}_1(t)$	$\boldsymbol{P}_1(t)=\boldsymbol{P}_1(t-1)-\boldsymbol{L}_1(t)\boldsymbol{\zeta}_1^{\mathrm{T}}(t)\in\mathbf{R}^{n_1\times n_1}$	n_1^2	n_1^2
$\hat{\bar{u}}(t)$	$\hat{\bar{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t)\in\mathbf{R}$	n_r	n_r-1
$\hat{\boldsymbol{\vartheta}}(t)$	$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\boldsymbol{L}_2(t)e(t)\in\mathbf{R}^{n_2}$	n_2	n_2
$\boldsymbol{L}_2(t)$	$\boldsymbol{L}_2(t)=\boldsymbol{\zeta}_2(t)/[1+\hat{\boldsymbol{\psi}}^{\mathrm{T}}(t)\boldsymbol{\zeta}_2(t)]\in\mathbf{R}^{n_2}$	$2n_2$	n_2
	$\boldsymbol{\zeta}_2(t):=\boldsymbol{P}_2(t-1)\hat{\boldsymbol{\psi}}(t)\in\mathbf{R}^{n_2}$	n_2^2	$n_2^2-n_2$
$\boldsymbol{P}_2(t)$	$\boldsymbol{P}_2(t)=\boldsymbol{P}_2(t-1)-\boldsymbol{L}_2(t)\boldsymbol{\zeta}_2^{\mathrm{T}}(t)\in\mathbf{R}^{n_2\times n_2}$	n_2^2	n_2^2
$\hat{w}(t)$	$\hat{w}(t)=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t)-\hat{\bar{u}}(t)\in\mathbf{R}$	n_1	n_1+1
总数		$4n_0+2n_1^2+2n_2^2+n_1-n_c$	$2n_0+2n_1^2+2n_2^2+n_1-n_c$
总 flop 数		$6n_0+4n_1^2+4n_2^2+2n_1-2n_c$	

$$\hat{\hat{u}}(t)=\boldsymbol{f}(u(t))\hat{\boldsymbol{\gamma}}(t), \tag{327}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\boldsymbol{L}_2(t)[\boldsymbol{Y}(p,t)-\hat{\boldsymbol{\Phi}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\theta}}(t-1)-\hat{\boldsymbol{\Psi}}^{\mathrm{T}}(p,t)\hat{\boldsymbol{\vartheta}}(t-1)], \tag{328}$$

$$\boldsymbol{L}_2(t)=\frac{\boldsymbol{P}_2(t-1)\hat{\boldsymbol{\Psi}}(p,t)}{\boldsymbol{I}_p+\hat{\boldsymbol{\Psi}}^{\mathrm{T}}(p,t)\boldsymbol{P}_2(t-1)\hat{\boldsymbol{\Psi}}(p,t)}, \tag{329}$$

$$\boldsymbol{P}_2(t)=\boldsymbol{P}_2(t-1)-\boldsymbol{L}_2(t)\hat{\boldsymbol{\Psi}}^{\mathrm{T}}(p,t)\boldsymbol{P}_2(t-1),\quad \boldsymbol{P}_2(0)=p_0\boldsymbol{I}_{n_\gamma+n_c}, \tag{330}$$

$$\hat{\boldsymbol{\Psi}}(p,t)=[\hat{\boldsymbol{\psi}}(t),\hat{\boldsymbol{\psi}}(t-1),\cdots,\hat{\boldsymbol{\psi}}(t-p+1)], \tag{331}$$

$$\hat{\boldsymbol{\psi}}(t)=[\boldsymbol{f}(u(t)),-\hat{w}(t-1),-\hat{w}(t-2),\cdots,-\hat{w}(t-n_c)]^{\mathrm{T}}, \tag{332}$$

$$\hat{w}(t)=y(t)-\hat{\boldsymbol{\varphi}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t)-\hat{\hat{u}}(t), \tag{333}$$

$$\boldsymbol{f}(u(t))=[f_1(u(t)),f_2(u(t)),\cdots,f_{n_\gamma}(u(t))], \tag{334}$$

$$\hat{\boldsymbol{\theta}}(t)=[\hat{\boldsymbol{a}}^{\mathrm{T}}(t),\hat{\boldsymbol{b}}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{335}$$

$$\hat{\boldsymbol{\vartheta}}(t)=[\hat{\boldsymbol{\gamma}}^{\mathrm{T}}(t),\hat{\boldsymbol{c}}^{\mathrm{T}}(t)]^{\mathrm{T}}. \tag{336}$$

当新息长度 $p=1$ 时,D-MILS 算法退化为 D-RLS 算法(308)—(320).D-MISG 算法(321)—(336)的计算步骤如下:

1) 初始化:令 $t=1$.置初值 $\hat{\boldsymbol{\theta}}(0)=\mathbf{1}_{n_a+n_b}/p_0$, $\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_{n_\gamma+n_c}/p_0$, $\boldsymbol{P}_1(0)=p_0\boldsymbol{I}_{n_a+n_b}$, $\boldsymbol{P}_2(0)=p_0\boldsymbol{I}_{n_\gamma+n_c}$, $\hat{w}(-i)=1/p_0$, $\hat{\hat{u}}(-i)=1/p_0$, $i=0,1,\cdots,n_0$, $p_0=10^6$,给定基函数 $f_i(\cdot)$ 和新息长度 p.

2) 收集数据 $u(t)$ 和 $y(t)$,用式(334)构造基函数向量 $\boldsymbol{f}(u(t))$.

3) 用式(326)构造信息向量 $\hat{\boldsymbol{\varphi}}(t)$,用式(331)构造信息矩阵 $\hat{\boldsymbol{\Phi}}(p,t)$.用式(325)构造堆积输出向量 $\boldsymbol{Y}(p,t)$.

4) 用式(332)构造信息向量 $\hat{\boldsymbol{\psi}}(t)$,用式(331)构造信息矩阵 $\hat{\boldsymbol{\Psi}}(p,t)$.

5) 用式(322)计算 $\boldsymbol{L}_1(t)$,用式(323)计算 $\boldsymbol{P}_1(t)$.

6) 根据式(321)刷新参数估计向量 $\hat{\boldsymbol{\theta}}(t)$.

7) 用式(329)计算 $\boldsymbol{L}_2(t)$,用式(330)计算 $\boldsymbol{P}_2(t)$.

8) 根据式(328)刷新参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$.

9) 从式(336)的 $\hat{\boldsymbol{\vartheta}}(t)$ 中读出 $\hat{\boldsymbol{\gamma}}(t)$,用式(327)计算 $\hat{\hat{u}}(t)$,用式(333)计算 $\hat{w}(t)$.

10) t 增 1,转到第 2)步.

D-MILS 算法(321)—(336)计算系统参数估计 $\hat{\boldsymbol{\theta}}(t)$ 和 $\hat{\boldsymbol{\vartheta}}(t)$ 的流程如图 5 所示.

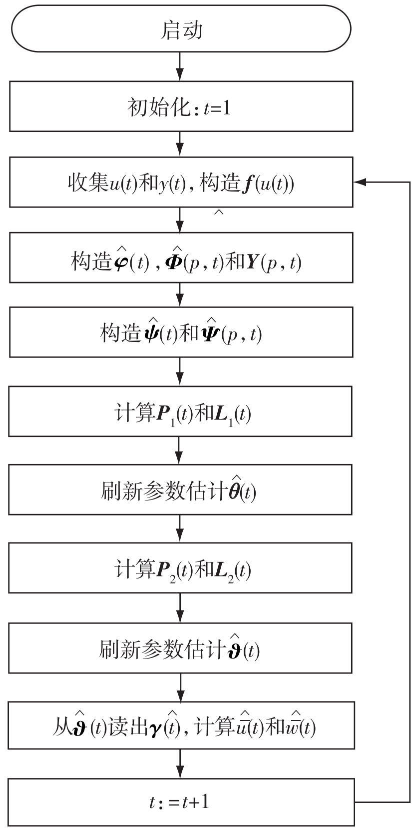

图 5 计算 D-MILS 参数估计 $\hat{\boldsymbol{\theta}}(t)$ 和 $\hat{\boldsymbol{\vartheta}}(t)$ 的流程

Fig. 5 The flowchart of computing the D-MILS parameter estimates $\hat{\boldsymbol{\theta}}(t)$ and $\hat{\boldsymbol{\vartheta}}(t)$

7 结语

针对输入非线性受控自回归自回归(IN-CARAR)系统,研究了基于过参数化模型的辨识方法(包括随机梯度算法、多新息随机梯度算法、最小二乘算法、多新息最小二乘算法,下同).由于过参数化模型带来很多冗余参数,需要辨识的参数比系统实际参数数目多,因而算法的计算量大(是指同类算法之间的比较),特点是不涉及估算未知中间变量(非线性块的输出).为了减小过参数化模型辨识算法计算量大的问题,分别使用关键项分离原理,利用数据滤波技术,利用辨识模型分解技术,推导了 IN-CARAR 系统的基于关键项分离的辨识方法,基于数据滤波的辨识方法,基于辨识模型分解的辨识方法.

值得指出的是,这些算法的计算量可以通过浮点运算次数进行粗略比较,但是一般很难从理论上分析和判断算法的参数估计精度(同类算法在相同数据长度、相同噪声水平等相同条件下),只能从计算机仿真中得出结论.这些算法的收敛性分析都是

需要研究的辨识课题.尽管本文的方法是针对输入非线性系统提出的,但是可以用于线性系统,也可以推广到输出非线性系统、反馈非线性系统等.

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,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

[4] 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

[5] 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

[6] 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.

[7] 丁锋,王艳娇.类多变量方程误差类系统的递阶多新息辨识方法[J].南京信息工程大学学报:自然科学版,2014,6(5):385-404
DING Feng, WANG Yanjiao. Hierarchical multi-innovation identification methods for multivariable equation-error-like type systems[J].Journal of Nanjing University of Information Science and Technology: Natural Science Edition,2014,6(5):385-404

[8] 丁锋,马兴云.规范状态空间系统辨识方法[J].南京信息工程大学学报:自然科学版,2014,6(6):481-504
DING Feng, MA Xingyun. Identification methods for canonical state space systems[J].Journal of Nanjing University of Information Science and Technology: Natural Science Edition,2014,6(6):481-504

[9] Ding F,Chen T.Performance analysis of multi-innovation gradient type identification methods [J]. Automatica,2007,43(1):1-14

[10] 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

[11] Ding F. Several multi-innovation identification methods [J].Digital Signal Processing,2010,20(4):1027-1039

[12] 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

[13] 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

[14] Ding F, Chen H B, Li M. Multi-innovation least squares identification methods based on the auxiliary model for MISO systems [J]. Applied Mathematics and Computation,2007,187(2):658-668

[15] 王冬青,丁锋.Box-Jenkins 模型的基于辅助模型的多新息广义增广随机梯度算法[J].控制与决策,2008,23(9):999-1003,1010
WANG Dongqing, DING Feng. Auxiliary model based multi-innovation generalized extended stochastic gradient algorithms for Box-Jenkins models[J].Control and Decision,2008,23(9):999-1003,1010

[16] 王冬青.基于辅助模型的递推增广最小二乘辨识方法[J].控制理论与应用,2009,26(1):51-56
WANG Dongqing.Recursive extended least squares identification method based on auxiliary models[J].Control Theory and Applications,2009,26(1):51-56

[17] 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

[18] Han L L,Ding F.Multi-innovation stochastic gradient algorithms for multi-input multi-output systems[J].Digital Signal Processing,2009,19(4):545-554

[19] 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

[20] Liu Y J,Xiao Y S,Zhao X L.Multi-innovation stochastic gradient algorithm for multiple-input single-output systems using the auxiliary model [J]. Applied Mathematics and Computation,2009,215(4):1477-1483

[21] 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

[22] Wang D Q, Chu Y Y, Ding F. Auxiliary model-based RELS and MI-ELS algorithms for Hammerstein OEMA systems [J]. Computers and Mathematics with Applications,2010,59(9):3092-3098

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

[24] Ding F. Hierarchical multi-innovation stochastic gradient algorithm for Hammerstein nonlinear system modeling [J]. Applied Mathematical Modelling, 2013, 37 (4): 1694-1704

[25] Xie L,Yang H Z,Ding F.Modeling and identification for non-uniformly periodically sampled-data systems[J].IET Control Theory and Applications,2010,4(5):784-794

[26] Han L L,Ding F.Identification for multirate multi-input systems using the multi-innovation identification theory [J]. Computers and Mathematics with Applications, 2009,57(9):1438-1449

[27] Zhang J B,Ding F,Shi Y.Self-tuning control based on multi-innovation stochastic gradient parameter estimation [J].Systems Control Letters,2009,58(1):69-75

[28] Wang D Q,Ding F.Input-output data filtering based recursive least squares parameter estimation for CARARMA systems[J].Digital Signal Processing,2010, 20(4):991-999

[29] Wang D Q.Least squares-based recursive and iterative estimation for output error moving average systems using data filtering[J].IET Control theory and Applications, 2011,5(14):1648-1657

[30] Wang D Q, Ding F, Chu Y Y. Data filtering based recursive least squares algorithm for Hammerstein systems using the key-term separation principle[J].Information Sciences,2013,222:203-212

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

[32] Ding F.State filtering and parameter estimation for state space systems with scarce measurements[J].Signal Processing,2014,104:369-380

[33] Ding F,Wang Y J,Ding J.Recursive least squares parameter identification for systems with colored noise using the filtering technique and the auxiliary model[J].Digital Signal Processing,2015,37:100-108

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

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

[36] Ding F,Chen T.Parameter estimation of dual-rate stochastic systems by using an output error method[J]. IEEE Transactions on Automatic Control,2005,50(9): 1436-1441

[37] Ding F,Chen T.Hierarchical gradient-based identification of multivariable discrete-time systems[J].Automatica, 2005,41(2):315-325

[38] Ding F,Chen T.Hierarchical least squares identification methods for multivariable systems[J].IEEE Transactions on Automatic Control,2005,50(3):397-402

[39] 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

[40] 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

[41] 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

[42] 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.

[43] Ding F,Liu G,Liu X P.Partially coupled stochastic gradient identification methods for non-uniformly sampled systems[J].IEEE Transactions on Automatic Control, 2010,55(8):1976-1981

[44] Ding F.Coupled-least-squares identification for multivariable systems[J].IET Control Theory and Applications, 2013,7(1):68-79

[45] Mao Y W,Ding F.Multi-innovation stochastic gradient identification for Hammerstein controlled autoregressive autoregressive systems based on the filtering technique [J].Nonlinear Dynamics,2015.doi 10.1007/s11071-014-1771-9

[46] Ding F, Shi Y, Chen T. Gradient-based identification methods for Hammerstein nonlinear ARMAX models[J]. Nonlinear Dynamics,2006,45(1-2):31-43

[47] Ding F,Chen T.Identification of Hammerstein nonlinear ARMAX systems [J]. Automatica, 2005, 41 (9): 1479-1489

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

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

[50] Ding F.Two-stage least squares based iterative estimation algorithm for CARARMA system modeling[J].Applied Mathematical Modelling,2013,37(7):4798-4808

[51] Ding F,Liu X M,Chen H B,et al.Hierarchical gradient based and hierarchical least squares based iterative parameter identification for CARARMA systems[J].Signal Processing,2014,97:31-39

[52] Chen H B,Xiao Y S,Ding F.Hierarchical gradient parameter estimation algorithm for Hammerstein nonlinear systems using the key term separation principle[J].Applied Mathematics and Computation, 2014, 247: 1202-1210

[53] Ding F,Liu X G,Chu J.Gradient based and least squares based iterative algorithms for Hammerstein systems using the hierarchical identification principle[J].IET Control Theory and Applications,2013,7(2):176-184.

[54] Ding F，Ma J X，Xiao Y S.Newton iterative identification for a class of output nonlinear systems with moving average noises[J].Nonlinear Dynamics，2013，74(1/2)：21-30

[55] Wang D Q，Ding F.Least squares based and gradient based iterative identification for Wiener nonlinear systems[J].Signal Processing，2011，91(5)：1182-1189

Multi-innovation identification methods for input nonlinear equation-error autoregressive systems

DING Feng[1,2,3] MAO Yawen[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 Typical block-oriented structure nonlinear systems include the basic input nonlinear systems，the output nonlinear systems，the input-output nonlinear systems and the feedback nonlinear systems.The input nonlinear systems include the input nonlinear equation-error type systems and the input nonlinear output-error type systems.Taking the input nonlinear equation-error autoregressive systems (namely the input nonlinear controlled autoregressive autoregressive (IN-CARAR) systems as an example，this paper studies and presents stochastic gradient (SG) identification methods，multi-innovation SG methods，recursive least squares (LS) identification methods and multi-innovation LS identification methods for IN-CARAR systems based on the over-parameterization model，the key term separation principle and the data filtering technique，the model decomposition technique.These methods can be extended to other input nonlinear equation-error systems，input nonlinear output-error type systems，output nonlinear equation-error type systems and output nonlinear output-error systems，and feedback nonlinear systems.Finally，the computational efficiency，the computational steps and the flowcharts of several typical identification algorithms are discussed.

Key words parameter estimation；recursive identification；gradient search；least squares；over-parameterization model；key term separation principle；data filtering technique；model decomposition etchnique；auxiliary model identification ideal；multi-innovation identification theory；hierarchical identification principle；coupling identification concept；input nonlinear system；output nonlinear system