文章编号:1674-7070(2014)02-0097-16

丁锋[1,2,3]　汪菲菲[1]　汪学海[1]

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

摘要

针对多元伪线性滑动平均系统,讨论了多元增广随机梯度算法,为减小算法的计算量,将系统分解为一些子系统,给出了子系统增广随机梯度算法,利用耦合辨识概念和多新息辨识理论,推导了部分耦合(子系统)增广随机梯度算法、部分耦合(子系统)多新息增广随机梯度算法.进一步将提出的方法推广到多元伪线性自回归滑动平均系统,给出了部分耦合(子系统)广义增广随机梯度算法、部分耦合(子系统)多新息广义增广随机梯度算法.文中分析了多元增广随机梯度算法、部分耦合增广随机梯度算法、部分耦合多新息增广随机梯度算法的计算量.

关键词

参数估计;递推辨识;梯度搜索;最小二乘;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;多元系统

中图分类号 TP273

文献标志码 A

收稿日期 2014-04-08

资助项目 国家自然科学基金(61273194);江苏省自然科学基金(BK2012549)

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

系统辨识是研究静态和动态系统的建模理论与方法,建立系统数学模型有机理方法与统计方法[1-2].机理方法可以获得系统的模型结构形式,即表示系统运动规律的方程式.模型的参数(即方程中有关变量的系数)可以通过测量获得,也可以通过试验采集系统的输入输出数据,用某种辨识方法计算得到.后者称为统计建模方法或统计辨识方法[3-5].多少年来,一些辨识方法层出不穷,如最小二乘辨识方法、随机梯度辨识方法、辅助模型辨识方法[6]、迭代辨识方法[7]、多新息辨识方法[8]、递阶辨识方法[9]、耦合辨识方法[10]等.辨识领域存在一系列重要的理论问题,例如存在不可测变量的系统辨识、复杂结构非线性系统的辨识,以及辨识算法的计算效率问题[11-13]、耦合多新息随机梯度算法[14]等.

实际系统是复杂的,而且存在干扰噪声.尽管系统中有许多变量可以用检测仪器测量得到,但有些关键变量还缺乏在线检测手段,或在线检测仪器十分昂贵,需要采集样品在实验室化验分析得到.在这种情形下,可用系统的大量观测信息建立一个辅助模型,利用辅助模型的输出代替系统的不可测变量,从而利用系统的输入输出数据和辅助模型的输出,来研究这类存在不可测变量的辨识问题,这就是辅助模型辨识思想[1,6].

不同的辨识方法具有不同的收敛速度,即使在相同数据长度下,获得的模型参数精度也是不一样的,即不同的辨识方法从输入输出数据中提取信息的能力是不一样的,如最小二乘方法从输入输出数据中提取信息的能力比随机梯度方法强,但随机梯度方法的计算量小.我们将能改进系统参数辨识精度的信息称为新息(innovation).为了提高随机梯度方法的收敛速度,本文作者通过扩展辨识新息长度,来提高参数估计精度,从而提出了多新息随机梯度辨识方法.多新息随机梯度辨识方法从输入输出数据中提取信息的能力比随机梯度方法强,从而改进参数估计精度,这就是多新息辨识理论[1,8].多新息最小二乘方法能提高损失数据系统参数估计精度[15].

实际系统大多是非线性的,辨识模型结构形式十分复杂,使得一些线性参数辨识方法无法应用.对于这样结构复杂的系统,可以根据

模型结构形式,将系统参数分为几个参数集对系统进行分解,使得分解后的子模型(子系统)对每个参数集空间都是线性的,然后分别对每一个子系统进行辨识,并协调各子系统辨识方法间的关联项,从而实现复杂系统的辨识,这就是递阶辨识原理[1,9].递阶辨识原理可以用于大规模系统的辨识和多变量系统的辨识,从而减小辨识方法的计算量[16].

一些典型的工业过程都是复杂多变量系统,且存在大量的输入输出变量.通常一个多变量系统可以按照系统输出数目分解为一些多输入单输出子系统.这些子系统间按照模型参数和信息划分,可以分为子系统间部分参数相同和(或)部分信息向量相同的系统.对于子系统间含有部分相同参数向量的辨识模型,如果每个子系统都估计一次这些参数,会导致大量冗余估计.为了减少冗余参数估计和辨识算法的计算量,应该实现子系统辨识方法间共同参数估计的连接,以至形成耦合辨识,这就是耦合辨识概念[1,10].耦合辨识概念主要用于研究各子系统间存在相同参数的多变量系统或多元系统的辨识问题[1,10,17-18].文献[1,10]研究了类多变量系统的部分耦合最小二乘辨识算法和部分耦合随机梯度辨识算法,文献[17]首次提出了非均匀采样数据系统的部分耦合随机梯度算法,文献[18]研究了多元系统(全)耦合最小二乘辨识方法的等价性和收敛性.

文献[14]将多元线性回归系统分解成 m 个子系统(m 为输出数目),推导出耦合(子系统)随机梯度辨识算法、耦合(子系统)多新息随机梯度辨识算法.本文将这些方法推广用于研究多元伪线性滑动平均系统和多元伪线性自回归滑动平均系统,利用耦合辨识概念和多新息辨识理论,提出了部分耦合(子系统)增广随机梯度算法、部分耦合(子系统)多新息增广随机梯度算法,以及部分耦合(子系统)广义增广随机梯度算法、部分耦合(子系统)多新息广义增广随机梯度算法,并分析了多元增广随机梯度算法、部分耦合增广随机梯度算法、部分耦合多新息增广随机梯度算法的计算量.

1 多元伪线性滑动平均系统

考虑多元伪线性滑动平均系统

$$\boldsymbol{y}(t)=\boldsymbol{\Phi}_{\rm s}(t)\boldsymbol{\theta}+\boldsymbol{D}(z)\boldsymbol{v}(t), \tag{1}$$

其中 $\boldsymbol{y}(t):=[y_1(t),y_2(t),\cdots,y_m(t)]^{\rm T}\in\mathbf{R}^m$ 为观测输出向量,$\boldsymbol{\Phi}_{\rm s}(t)\in\mathbf{R}^{m\times n}$ 是由输入输出数据构成的回归信息矩阵,$\boldsymbol{\theta}\in\mathbf{R}^n$ 为系统参数向量,$\boldsymbol{w}(t):=\boldsymbol{D}(z)\boldsymbol{v}(t)\in\mathbf{R}^m$ 为 m 维干扰噪声向量,$\boldsymbol{v}(t):=[v_1(t),v_2(t),\cdots,v_m(t)]^{\rm T}\in\mathbf{R}^m$ 为 m 维白噪声向量,$\boldsymbol{D}(z)\in\mathbf{R}^{m\times m}$ 是单位移位算子 z^{-1} 的首项为单位阵的多项式矩阵:

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

假设阶次 m,n 和 n_d 已知,且当 $t\leqslant 0$ 时,$\boldsymbol{y}(t)=0,\boldsymbol{\Phi}_{\rm s}(t)=0,\boldsymbol{v}(t)=0$.辨识的目标是基于耦合辨识概念和多新息辨识理论,利用系统的观测数据$\{\boldsymbol{y}(t),\boldsymbol{\Phi}_{\rm s}(t):t=1,2,3,\cdots\}$提出新的算法,估计系统参数向量 $\boldsymbol{\theta}$ 以及噪声模型的参数矩阵 $\boldsymbol{D}_i$.

1.1 多元增广随机梯度辨识算法

定义信息矩阵 $\boldsymbol{\Phi}(t)$ 和参数向量 $\boldsymbol{\vartheta}$ 如下:

$$\boldsymbol{\Phi}(t):=[\boldsymbol{\Phi}_{\rm s}(t),\boldsymbol{\varphi}_{\rm n}^{\rm T}(t)\otimes\boldsymbol{I}_m]\in\mathbf{R}^{m\times n_0},\quad n_0:=n+m^2n_d,$$
$$\boldsymbol{\vartheta}:=\begin{bmatrix}\boldsymbol{\theta}\\ {\rm col}[\boldsymbol{D}]\end{bmatrix}\in\mathbf{R}^{n_0},$$
$$\boldsymbol{\varphi}_{\rm n}(t):=[\boldsymbol{v}^{\rm T}(t-1),\boldsymbol{v}^{\rm T}(t-2),\cdots,\boldsymbol{v}^{\rm T}(t-n_d)]^{\rm T}\in\mathbf{R}^{mn_d},$$
$$\boldsymbol{D}:=[\boldsymbol{D}_1,\boldsymbol{D}_2,\cdots,\boldsymbol{D}_{n_d}]\in\mathbf{R}^{m\times(mn_d)},$$

式中⊗为 Kronecker 积符号,col$[\boldsymbol{X}]$定义为将矩阵 $\boldsymbol{X}$ 的列按次序排成的向量.如 $\boldsymbol{X}=[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n]\in\mathbf{R}^{m\times n}$,$\boldsymbol{x}_i\in\mathbf{R}^m,i=1,2,\cdots,n$,那么

$${\rm col}[\boldsymbol{X}]:=\begin{bmatrix}\boldsymbol{x}_1\\ \boldsymbol{x}_2\\ \vdots\\ \boldsymbol{x}_n\end{bmatrix}\in\mathbf{R}^{mn}.$$

于是,伪线性滑动平均系统(1)可以写成下列多元伪线性回归模型(multivariate pseudo-linear regressive model):

$$\boldsymbol{y}(t)=\boldsymbol{\Phi}(t)\boldsymbol{\vartheta}+\boldsymbol{v}(t). \tag{2}$$

辨识的困难是 $\boldsymbol{\Phi}(t)$ 中包含了未知噪声回归项 $\boldsymbol{v}(t-i)$,解决的办法是用其估计值 $\hat{\boldsymbol{v}}(t-i)$ 代替,代替后的信息矩阵 $\boldsymbol{\Phi}(t)$ 记作 $\boldsymbol{\Psi}(t)$.在辨识算法中使用 $\boldsymbol{\Psi}(t)$,而非 $\boldsymbol{\Phi}(t)$,于是能够获得估计多元伪线性回归辨识模型(2)参数向量 $\boldsymbol{\vartheta}$ 的多元增广随机梯度算法(Multivariate Extended Stochastic Gradient algorithm,M-ESG 算法):

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\frac{\boldsymbol{\Psi}^{\rm T}(t)}{r(t)}[\boldsymbol{y}(t)-\boldsymbol{\Psi}(t)\hat{\boldsymbol{\vartheta}}(t-1)],$$
$$\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_{n_0}/p_0, \tag{3}$$
$$r(t)=r(t-1)+\|\boldsymbol{\Psi}(t)\|^2,\quad r(0)=1, \tag{4}$$
$$\boldsymbol{\Psi}(t)=[\boldsymbol{\Phi}_{\rm s}(t),\hat{\boldsymbol{\varphi}}_{\rm n}^{\rm T}(t)\otimes\boldsymbol{I}_m], \tag{5}$$
$$\hat{\boldsymbol{\varphi}}_{\rm n}(t)=[\hat{\boldsymbol{v}}^{\rm T}(t-1),\hat{\boldsymbol{v}}^{\rm T}(t-2),\cdots,\hat{\boldsymbol{v}}^{\rm T}(t-n_d)]^{\rm T},$$

$$\hat{v}(-i)=\mathbf{1}_m/p_0,\quad i=0,1,\cdots,n_d,\tag{6}$$

$$\hat{v}(t)=\boldsymbol{y}(t)-\boldsymbol{\Psi}(t)\hat{\boldsymbol{\vartheta}}(t).\tag{7}$$

表 1 列出了 M-ESG 算法每一步递推计算中的乘法次数、加法次数和 flop 数[11-13].

表 1 多元增广随机梯度算法的计算量

Table 1 The computational efficiency of the M-ESG algorithm

变量	计算次序	乘法次数	加法次数
$\hat{\boldsymbol{\vartheta}}(t)$	$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}(t-1)+\boldsymbol{\Psi}^{\mathrm{T}}(t)[\boldsymbol{E}(t)/r(t)]\in\mathbf{R}^{n+m^2n_d}$	$m+m(n+m^2n_d)$	$m(n+m^2n_d)$
	$\boldsymbol{E}(t):=\boldsymbol{y}(t)-\Psi(t)\hat{\boldsymbol{\vartheta}}(t-1)\in\mathbf{R}^m$	$m(n+m^2n_d)$	$m(n+m^2n_d)$
$r(t)$	$r(t)=r(t-1)+\Vert\boldsymbol{\Psi}(t)\Vert^2\in\mathbf{R}$	$m(n+m^2n_d)$	$m(n+m^2n_d)$
	总数	$m+3m(n+m^2n_d)$	$3m(n+m^2n_d)$
	总 flop 数	$m+6m(n+m^2n_d)$	

注意到信息矩阵 $\boldsymbol{\Psi}(t)$ 中包含了大量的零元,使得算法的计算量增大.分析式(2),可知其是一个部分参数向量和部分信息向量耦合型辨识模型,即分解后的子系统辨识模型间有部分参数向量相同,也有部分信息向量相同.下面研究这类系统的部分耦合辨识方法.

1.2 部分耦合(子系统)增广随机梯度辨识算法

定义参数矩阵 $\boldsymbol{\alpha}$ 和噪声信息向量 $\boldsymbol{\varphi}_{\mathrm{n}}(t)$ 如下:

$$\boldsymbol{\alpha}^{\mathrm{T}}:=[\boldsymbol{D}_1,\boldsymbol{D}_2,\cdots,\boldsymbol{D}_{n_d}]=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m]^{\mathrm{T}}\in\mathbf{R}^{m\times(mn_d)},$$

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

令 $\boldsymbol{\phi}_i^{\mathrm{T}}(t)$ 为系统信息矩阵 $\boldsymbol{\Phi}_{\mathrm{s}}(t)$ 的第 i 行,即

$$\boldsymbol{\Phi}_{\mathrm{s}}(t)=[\boldsymbol{\phi}_1(t),\boldsymbol{\phi}_2(t),\cdots,\boldsymbol{\phi}_m(t)]^{\mathrm{T}}\in\mathbf{R}^{m\times n}.$$

于是多元伪线性滑动平均系统(1)可以写为

$$\begin{aligned}\boldsymbol{y}(t)&=\boldsymbol{\Phi}_{\mathrm{s}}(t)\boldsymbol{\theta}+\boldsymbol{D}(z)\boldsymbol{v}(t)=\\&\boldsymbol{\Phi}_{\mathrm{s}}(t)\boldsymbol{\theta}+[\boldsymbol{D}_1,\boldsymbol{D}_2,\cdots,\boldsymbol{D}_{n_d}]\begin{bmatrix}\boldsymbol{v}(t-1)\\\boldsymbol{v}(t-2)\\\vdots\\\boldsymbol{v}(t-n_d)\end{bmatrix}+\boldsymbol{v}(t)=\\&\boldsymbol{\Phi}_{\mathrm{s}}(t)\boldsymbol{\theta}+\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\varphi}_{\mathrm{n}}(t)+\boldsymbol{v}(t).\end{aligned}\tag{8}$$

进一步可以表示为

$$\begin{bmatrix}y_1(t)\\y_2(t)\\\vdots\\y_m(t)\end{bmatrix}=\begin{bmatrix}\boldsymbol{\phi}_1^{\mathrm{T}}(t)\\\boldsymbol{\phi}_2^{\mathrm{T}}(t)\\\vdots\\\boldsymbol{\phi}_m^{\mathrm{T}}(t)\end{bmatrix}\boldsymbol{\theta}+\begin{bmatrix}\boldsymbol{\alpha}_1^{\mathrm{T}}\\\boldsymbol{\alpha}_2^{\mathrm{T}}\\\vdots\\\boldsymbol{\alpha}_m^{\mathrm{T}}\end{bmatrix}\boldsymbol{\varphi}_{\mathrm{n}}(t)+\begin{bmatrix}v_1(t)\\v_2(t)\\\vdots\\v_m(t)\end{bmatrix}.\tag{9}$$

定义子系统信息向量为 $\boldsymbol{\psi}_i(t):=\begin{bmatrix}\boldsymbol{\phi}_i(t)\\\boldsymbol{\varphi}_{\mathrm{n}}(t)\end{bmatrix}\in\mathbf{R}^{n+mn_d}$,则式(9)可以分解为 m 个辨识模型(子系统):

$$y_i(t)=\boldsymbol{\phi}_i^{\mathrm{T}}(t)\boldsymbol{\theta}+\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{\varphi}_{\mathrm{n}}(t)+v_i(t)=$$

$$\boldsymbol{\phi}_i^{\mathrm{T}}(t)\boldsymbol{\theta}+\boldsymbol{\varphi}_{\mathrm{n}}^{\mathrm{T}}(t)\boldsymbol{\alpha}_i+v_i(t)=\tag{10}$$

$$[\boldsymbol{\phi}_i^{\mathrm{T}}(t),\boldsymbol{\varphi}_{\mathrm{n}}^{\mathrm{T}}(t)]\begin{bmatrix}\boldsymbol{\theta}\\\boldsymbol{\alpha}_i\end{bmatrix}+v_i(t)=$$

$$\boldsymbol{\psi}_i^{\mathrm{T}}(t)\begin{bmatrix}\boldsymbol{\theta}\\\boldsymbol{\alpha}_i\end{bmatrix}+v_i(t),\quad i=1,2,\cdots,m.\tag{11}$$

在这 m 个子辨识模型(10)中,每个子辨识模型有一个共同的子参数向量 $\boldsymbol{\theta}$ 和一个共同的子信息向量 $\boldsymbol{\varphi}_{\mathrm{n}}(t)$,每个子辨识模型还包含一个不同的子参数向量 $\boldsymbol{\alpha}_i$ 和一个不同的子信息向量 $\boldsymbol{\phi}_i(t)$.为减小计算量,下面研究辨识模型(11)的子系统梯度辨识方法和部分耦合梯度辨识方法.

1) 子系统增广随机梯度辨识算法

使用负梯度搜索,根据辨识模型(11)定义和极小化梯度准则函数

$$J_1(\boldsymbol{\theta},\boldsymbol{\alpha}_i):=\left\{y_i(t)-\boldsymbol{\psi}_i^{\mathrm{T}}(t)\begin{bmatrix}\boldsymbol{\theta}\\\boldsymbol{\alpha}_i\end{bmatrix}\right\}^2,\quad i=1,2,\cdots,m$$

可以得到估计参数向量 $\boldsymbol{\theta}$ 和 $\boldsymbol{\alpha}_i$ 的子系统随机梯度辨识算法:

$$\begin{bmatrix}\hat{\boldsymbol{\theta}}(t)\\\hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\theta}}(t-1)\\\hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\boldsymbol{\psi}_i(t)}{r_i(t)}\left\{y_i(t)-\boldsymbol{\psi}_i^{\mathrm{T}}(t)\begin{bmatrix}\hat{\boldsymbol{\theta}}(t-1)\\\hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}\right\},\tag{12}$$

$$r_i(t)=r_i(t-1)+\Vert\boldsymbol{\psi}_i(t)\Vert^2,\quad r_i(0)=1.\tag{13}$$

由于 $\boldsymbol{\psi}_i(t)$ 中包含了未知噪声项 $\boldsymbol{v}(t-i)$,参数向量 $\boldsymbol{\theta}$ 和 $\boldsymbol{\alpha}_i$ 的估计 $\hat{\boldsymbol{\theta}}(t)$ 和 $\hat{\boldsymbol{\alpha}}_i(t)$ 无法计算得到.借助辅助模型辨识思想[6],将 $\boldsymbol{\psi}_i(t)$ 中未知的 $\boldsymbol{v}(t-i)$ 用其估计 $\hat{\boldsymbol{v}}(t-i)$ 代替,代替后的 $\boldsymbol{\psi}_i(t)$ 记作 $\hat{\boldsymbol{\psi}}_i(t)$,因此有

$$\hat{\boldsymbol{\psi}}_i(t):=\begin{bmatrix}\boldsymbol{\phi}_i(t)\\\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\end{bmatrix}\in\mathbf{R}^{n+mn_d},i=1,2,\cdots,m,$$

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

$$\hat{\boldsymbol{v}}(t) := \begin{bmatrix} \hat{v}_1(t) \\ \hat{v}_2(t) \\ \vdots \\ \hat{v}_m(t) \end{bmatrix} \in \mathbf{R}^m.$$

根据式(10),$v_i(t)$ 的估计 $\hat{v}_i(t)$ 可通过下式计算:

$$\hat{v}_i(t) = y_i(t) - \boldsymbol{\phi}_i^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t) - \hat{\boldsymbol{\varphi}}_{\mathrm{n}}^{\mathrm{T}}(t)\boldsymbol{\alpha}_i(t).$$

作了这种代替后,可实现的子系统增广随机梯度辨识算法表达如下($i=1,2,\cdots,m$):

$$\begin{bmatrix} \hat{\boldsymbol{\theta}}(t) \\ \hat{\boldsymbol{\alpha}}_i(t) \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{\theta}}(t-1) \\ \hat{\boldsymbol{\alpha}}_i(t-1) \end{bmatrix} + \frac{\hat{\boldsymbol{\psi}}_i(t)}{r_i(t)} \left\{ y_i(t) - \hat{\boldsymbol{\psi}}_i^{\mathrm{T}}(t) \begin{bmatrix} \hat{\boldsymbol{\theta}}(t-1) \\ \hat{\boldsymbol{\alpha}}_i(t-1) \end{bmatrix} \right\},$$

$$\hat{\boldsymbol{\theta}}(0) = \mathbf{1}_n/p_0, \quad \hat{\boldsymbol{\alpha}}_i(0) = \mathbf{1}_{mn_d}/p_0, \tag{14}$$

$$r_i(t) = r_i(t-1) + \|\hat{\boldsymbol{\psi}}_i(t)\|^2, \quad r_i(0) = 1, \tag{15}$$

$$\hat{\boldsymbol{\psi}}_i(t) = \begin{bmatrix} \boldsymbol{\phi}_i(t) \\ \hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t) \end{bmatrix}, \tag{16}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t) = [\hat{\boldsymbol{v}}^{\mathrm{T}}(t-1), \hat{\boldsymbol{v}}^{\mathrm{T}}(t-2), \cdots, \hat{\boldsymbol{v}}^{\mathrm{T}}(t-n_d)]^{\mathrm{T}}, \tag{17}$$

$$\hat{\boldsymbol{v}}(t) = [\hat{v}_1(t), \hat{v}_2(t), \cdots, \hat{v}_m(t)]^{\mathrm{T}}, \tag{18}$$

$$\hat{v}_i(t) = y_i(t) - \boldsymbol{\phi}_i^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t) - \hat{\boldsymbol{\varphi}}_{\mathrm{n}}^{\mathrm{T}}(t)\hat{\boldsymbol{\alpha}}_i(t), \tag{19}$$

$$\boldsymbol{\Phi}_{\mathrm{s}}(t) = [\boldsymbol{\phi}_1(t), \boldsymbol{\phi}_2(t), \cdots, \boldsymbol{\phi}_m(t)]^{\mathrm{T}}. \tag{20}$$

在上述算法(14)—(20)中,每个子系统都对 $\boldsymbol{\theta}$ 估计了一次,是一种冗余估计.为区分每个子系统中 $\boldsymbol{\theta}$ 的估计,加下标 i,记作 $\hat{\boldsymbol{\theta}}_i(t)$,便得到下列子系统增广随机梯度辨识算法(Subsystem Extended Stochastic Gradient algorithm, S-ESG 算法):

$$\begin{bmatrix} \hat{\boldsymbol{\theta}}_i(t) \\ \hat{\boldsymbol{\alpha}}_i(t) \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{\theta}}_i(t-1) \\ \hat{\boldsymbol{\alpha}}_i(t-1) \end{bmatrix} + \frac{\hat{\boldsymbol{\psi}}_i(t)}{r_i(t)} \left\{ y_i(t) - \hat{\boldsymbol{\psi}}_i^{\mathrm{T}}(t) \begin{bmatrix} \hat{\boldsymbol{\theta}}_i(t-1) \\ \hat{\boldsymbol{\alpha}}_i(t-1) \end{bmatrix} \right\},$$

$$\hat{\boldsymbol{\theta}}_i(0) = \mathbf{1}_n/p_0, \quad \hat{\boldsymbol{\alpha}}_i(0) = \mathbf{1}_{mn_d}/p_0, \tag{21}$$

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

$$\hat{\boldsymbol{\psi}}_i(t) = \begin{bmatrix} \boldsymbol{\phi}_i(t) \\ \hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t) \end{bmatrix}, \quad i=1,2,\cdots,m, \tag{23}$$

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

$$\hat{\boldsymbol{v}}(t) = [\hat{v}_1(t), \hat{v}_2(t), \cdots, \hat{v}_m(t)]^{\mathrm{T}}, \tag{25}$$

$$\hat{v}_i(t) = y_i(t) - \boldsymbol{\phi}_i^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_i(t) - \hat{\boldsymbol{\varphi}}_{\mathrm{n}}^{\mathrm{T}}(t)\hat{\boldsymbol{\alpha}}_i(t), \tag{26}$$

$$\boldsymbol{\Phi}_{\mathrm{s}}(t) = [\boldsymbol{\phi}_1(t), \boldsymbol{\phi}_2(t), \cdots, \boldsymbol{\phi}_m(t)]^{\mathrm{T}}. \tag{27}$$

2)部分耦合子系统增广随机梯度辨识算法

S-ESG 算法(21)—(27)对参数向量 $\boldsymbol{\theta}$ 估计了 m 次,产生 m 个估计 $\hat{\boldsymbol{\theta}}_i(t)$,$i=1,2,\cdots,m$.实践中只需要一个估计,因此可将每个子系统的估计 $\hat{\boldsymbol{\theta}}_i(t)$ 的平均值作为参数向量 $\boldsymbol{\theta}$ 的估计,即

$$\hat{\boldsymbol{\theta}}(t) = \frac{\hat{\boldsymbol{\theta}}_1(t) + \hat{\boldsymbol{\theta}}_2(t) + \cdots + \hat{\boldsymbol{\theta}}_m(t)}{m} \in \mathbf{R}^n.$$

用平均值 $\hat{\boldsymbol{\theta}}(t-1)$ 代替 S-ESG 算法中的 $\hat{\boldsymbol{\theta}}_i(t-1)$,就得到了一个简单的部分耦合子系统增广随机梯度算法(Partially Coupled Subsystem Extended Stochastic Gradient algorithm, PC-S-ESG 算法):

$$\begin{bmatrix} \hat{\boldsymbol{\theta}}_i(t) \\ \hat{\boldsymbol{\alpha}}_i(t) \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{\theta}}(t-1) \\ \hat{\boldsymbol{\alpha}}_i(t-1) \end{bmatrix} + \frac{\hat{\boldsymbol{\psi}}_i(t)}{r_i(t)} \left\{ y_i(t) - \hat{\boldsymbol{\psi}}_i^{\mathrm{T}}(t) \begin{bmatrix} \hat{\boldsymbol{\theta}}(t-1) \\ \hat{\boldsymbol{\alpha}}_i(t-1) \end{bmatrix} \right\},$$

$$\hat{\boldsymbol{\theta}}(0) = \mathbf{1}_n/p_0, \quad \hat{\boldsymbol{\alpha}}_i(0) = \mathbf{1}_{mn_d}/p_0, \tag{28}$$

$$r_i(t) = r_i(t-1) + \|\hat{\boldsymbol{\psi}}_i(t)\|^2, \quad r_i(0) = 1, \tag{29}$$

$$\hat{\boldsymbol{\psi}}_i(t) = \begin{bmatrix} \boldsymbol{\phi}_i(t) \\ \hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t) \end{bmatrix}, \quad i=1,2,\cdots,m, \tag{30}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t) = [\hat{\boldsymbol{v}}^{\mathrm{T}}(t-1), \hat{\boldsymbol{v}}^{\mathrm{T}}(t-2), \cdots, \hat{\boldsymbol{v}}^{\mathrm{T}}(t-n_d)]^{\mathrm{T}}, \tag{31}$$

$$\hat{\boldsymbol{v}}(t) = \boldsymbol{y}(t) - \boldsymbol{\Phi}_{\mathrm{s}}(t)\hat{\boldsymbol{\theta}}(t) - \hat{\boldsymbol{\alpha}}^{\mathrm{T}}(t)\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t), \tag{32}$$

$$\boldsymbol{\Phi}_{\mathrm{s}}(t) = [\boldsymbol{\phi}_1(t), \boldsymbol{\phi}_2(t), \cdots, \boldsymbol{\phi}_m(t)]^{\mathrm{T}}, \tag{33}$$

$$\hat{\boldsymbol{\theta}}(t) = \frac{\hat{\boldsymbol{\theta}}_1(t) + \hat{\boldsymbol{\theta}}_2(t) + \cdots + \hat{\boldsymbol{\theta}}_m(t)}{m}, \tag{34}$$

$$\hat{\boldsymbol{\alpha}}(t) = [\hat{\boldsymbol{\alpha}}_1(t), \hat{\boldsymbol{\alpha}}_2(t), \cdots, \hat{\boldsymbol{\alpha}}_m(t)]. \tag{35}$$

① 遗忘因子部分耦合子系统增广随机梯度辨识算法

为了提高部分耦合子系统增广随机梯度算法的暂态性能,在式(29)中引入遗忘因子(Forgetting Factor, FF)λ,能够得到估计参数向量 $\boldsymbol{\theta}$ 和 $\boldsymbol{\alpha}$ 的遗忘因子 PC-S-ESG 算法[19-21]:

$$\begin{bmatrix} \hat{\boldsymbol{\theta}}_i(t) \\ \hat{\boldsymbol{\alpha}}_i(t) \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{\theta}}(t-1) \\ \hat{\boldsymbol{\alpha}}_i(t-1) \end{bmatrix} + \frac{\hat{\boldsymbol{\psi}}_i(t)}{r_i(t)} \left\{ y_i(t) - \hat{\boldsymbol{\psi}}_i^{\mathrm{T}}(t) \begin{bmatrix} \hat{\boldsymbol{\theta}}(t-1) \\ \hat{\boldsymbol{\alpha}}_i(t-1) \end{bmatrix} \right\} = \begin{bmatrix} \hat{\boldsymbol{\theta}}(t-1) \\ \hat{\boldsymbol{\alpha}}_i(t-1) \end{bmatrix} + \frac{\hat{\boldsymbol{\psi}}_i(t)}{r_i(t)} [y_i(t) - \boldsymbol{\phi}_i^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1) - \hat{\boldsymbol{\varphi}}_{\mathrm{n}}^{\mathrm{T}}(t)\hat{\boldsymbol{\alpha}}_i(t-1)], \tag{36}$$

$$r_i(t) = \lambda r_i(t-1) + \|\hat{\boldsymbol{\psi}}_i(t)\|^2, \quad 0 \leqslant \lambda \leqslant 1, \tag{37}$$

$$\hat{\boldsymbol{\psi}}_i(t) = \begin{bmatrix} \boldsymbol{\phi}_i(t) \\ \hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t) \end{bmatrix}, \quad i=1,2,\cdots,m, \tag{38}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t) = [\hat{\boldsymbol{v}}^{\mathrm{T}}(t-1), \hat{\boldsymbol{v}}^{\mathrm{T}}(t-2), \cdots, \hat{\boldsymbol{v}}^{\mathrm{T}}(t-n_d)]^{\mathrm{T}}, \tag{39}$$

$$\hat{\boldsymbol{v}}(t) = \boldsymbol{y}(t) - \boldsymbol{\Phi}_{\mathrm{s}}(t)\hat{\boldsymbol{\theta}}(t) - \hat{\boldsymbol{\alpha}}^{\mathrm{T}}(t)\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t), \tag{40}$$

$$\boldsymbol{\Phi}_{\mathrm{s}}(t) = [\boldsymbol{\phi}_1(t), \boldsymbol{\phi}_2(t), \cdots, \boldsymbol{\phi}_m(t)]^{\mathrm{T}}, \tag{41}$$

$$\hat{\boldsymbol{\theta}}(t) = \frac{\hat{\boldsymbol{\theta}}_1(t) + \hat{\boldsymbol{\theta}}_2(t) + \cdots + \hat{\boldsymbol{\theta}}_m(t)}{m}, \tag{42}$$

$$\hat{\boldsymbol{\alpha}}(t) = [\hat{\boldsymbol{\alpha}}_1(t), \hat{\boldsymbol{\alpha}}_2(t), \cdots, \hat{\boldsymbol{\alpha}}_m(t)]. \tag{43}$$

② 修正部分耦合子系统增广随机梯度辨识算法

为了提高暂态和稳态收敛速度，在式(28)中引入收敛指数(convergence index)ε，得到带收敛指数的修正 PC-S-ESG 算法[22-23]：

$$\begin{bmatrix}\hat{\boldsymbol{\theta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\theta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\hat{\boldsymbol{\psi}}_i(t)}{r_i^{\varepsilon}(t)}\left\{y_i(t)-\hat{\boldsymbol{\psi}}_i^{\mathrm{T}}(t)\begin{bmatrix}\hat{\boldsymbol{\theta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}\right\}=$$
$$\begin{bmatrix}\hat{\boldsymbol{\theta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{1}{r_i^{\varepsilon}(t)}\begin{bmatrix}\boldsymbol{\phi}_i(t)\\ \hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\end{bmatrix}[y_i(t)-\boldsymbol{\phi}_i^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)-$$
$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}^{\mathrm{T}}(t)\hat{\boldsymbol{\alpha}}_i(t-1)],\quad \frac{1}{2}<\varepsilon\leqslant 1, \tag{44}$$

$$r_i(t)=r_i(t-1)+\|\boldsymbol{\phi}_i(t)\|^2+\|\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\|^2, \tag{45}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)=[\hat{\boldsymbol{v}}^{\mathrm{T}}(t-1),\hat{\boldsymbol{v}}^{\mathrm{T}}(t-2),\cdots,\hat{\boldsymbol{v}}^{\mathrm{T}}(t-n_d)]^{\mathrm{T}}, \tag{46}$$

$$\hat{\boldsymbol{v}}(t)=\boldsymbol{y}(t)-\boldsymbol{\Phi}_{\mathrm{s}}(t)\hat{\boldsymbol{\theta}}(t)-\hat{\boldsymbol{\alpha}}^{\mathrm{T}}(t)\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t), \tag{47}$$

$$\boldsymbol{\Phi}_{\mathrm{s}}(t)=[\boldsymbol{\phi}_1(t),\boldsymbol{\phi}_2(t),\cdots,\boldsymbol{\phi}_m(t)]^{\mathrm{T}}, \tag{48}$$

$$\hat{\boldsymbol{\theta}}(t)=\frac{\hat{\boldsymbol{\theta}}_1(t)+\hat{\boldsymbol{\theta}}_2(t)+\cdots+\hat{\boldsymbol{\theta}}_m(t)}{m}, \tag{49}$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\boldsymbol{\alpha}}_1(t),\hat{\boldsymbol{\alpha}}_2(t),\cdots,\hat{\boldsymbol{\alpha}}_m(t)]. \tag{50}$$

为避免冗余计算，下面基于部分耦合子系统增广随机梯度算法来推导部分耦合增广随机梯度算法.

3）部分耦合增广随机梯度辨识算法

对于递推参数估计算法，期望参数估计随着数据长度 t 的增大而收敛于真参数，故可以认为第 $i-1$ 个子系统在时刻 t 的参数估计 $\hat{\boldsymbol{\theta}}_{i-1}(t)$ 比第 i 个子系统在时刻 $t-1$ 的参数估计 $\hat{\boldsymbol{\theta}}_i(t-1)$ 更接近真参数 $\boldsymbol{\theta}$. 参考文献[10,14,17]中的部分耦合随机梯度辨识方法，用 $\hat{\boldsymbol{\theta}}_{i-1}(t)$ 代替式(21)右边的 $\hat{\boldsymbol{\theta}}_i(t-1)$(见式(53))，用 $\hat{\boldsymbol{\theta}}_m(t-1)$ 代替式(21)中 $i=1$ 时的 $\hat{\boldsymbol{\theta}}_1(t-1)$(见式(51))，对式(22)的 $r_i(t)$ 也进行类似的耦合(见式(52)和(54))，则得到部分耦合增广随机梯度算法(Partially Coupled Extended Stochastic Gradient algorithm，PC-ESG 算法)：

$$\begin{bmatrix}\hat{\boldsymbol{\theta}}_1(t)\\ \hat{\boldsymbol{\alpha}}_1(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\theta}}_m(t-1)\\ \hat{\boldsymbol{\alpha}}_1(t-1)\end{bmatrix}+\frac{\hat{\boldsymbol{\psi}}_1(t)}{r_1(t)}\left\{y_1(t)-\hat{\boldsymbol{\psi}}_1^{\mathrm{T}}(t)\begin{bmatrix}\hat{\boldsymbol{\theta}}_m(t-1)\\ \hat{\boldsymbol{\alpha}}_1(t-1)\end{bmatrix}\right\},$$
$$\hat{\boldsymbol{\theta}}_m(0)=\mathbf{1}_n/p_0, \tag{51}$$

$$r_1(t)=r_m(t-1)+\|\hat{\boldsymbol{\psi}}_1(t)\|^2,\quad r_1(0)=1,$$
$$\hat{\boldsymbol{\alpha}}_1(0)=\mathbf{1}_{mn_d}/p_0, \tag{52}$$

$$\begin{bmatrix}\hat{\boldsymbol{\theta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\theta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\hat{\boldsymbol{\psi}}_i(t)}{r_i(t)}\left\{y_i(t)-\hat{\boldsymbol{\psi}}_i^{\mathrm{T}}(t)\begin{bmatrix}\hat{\boldsymbol{\theta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}\right\},$$
$$\hat{\boldsymbol{\alpha}}_i(0)=\mathbf{1}_{mn_d}/p_0, \tag{53}$$

$$r_i(t)=r_{i-1}(t)+\|\hat{\boldsymbol{\psi}}_i(t)\|^2,\quad r_i(0)=1,\quad i=2,3,\cdots,m, \tag{54}$$

$$\hat{\boldsymbol{\psi}}_i(t)=\begin{bmatrix}\boldsymbol{\phi}_i(t)\\ \hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\end{bmatrix},\quad i=1,2,\cdots,m, \tag{55}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)=[\hat{\boldsymbol{v}}^{\mathrm{T}}(t-1),\hat{\boldsymbol{v}}^{\mathrm{T}}(t-2),\cdots,\hat{\boldsymbol{v}}^{\mathrm{T}}(t-n_d)]^{\mathrm{T}}, \tag{56}$$

$$\hat{\boldsymbol{v}}(t)=[\hat{v}_1(t),\hat{v}_2(t),\cdots,\hat{v}_m(t)]^{\mathrm{T}}, \tag{57}$$

$$\hat{v}_i(t)=y_i(t)-\boldsymbol{\phi}_i^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_i(t)-\hat{\boldsymbol{\varphi}}_{\mathrm{n}}^{\mathrm{T}}(t)\hat{\boldsymbol{\alpha}}_i(t), \tag{58}$$

$$\boldsymbol{\Phi}_{\mathrm{s}}(t)=[\boldsymbol{\phi}_1(t),\boldsymbol{\phi}_2(t),\cdots,\boldsymbol{\phi}_m(t)]^{\mathrm{T}}. \tag{59}$$

这个算法输出的系统参数估计为 $\hat{\boldsymbol{\theta}}(t):=\hat{\boldsymbol{\theta}}_m(t)$，噪声模型的参数估计为 $\hat{\boldsymbol{\alpha}}_i(t)$. 部分耦合增广随机梯度算法的计算量如表 2 所示.

表 2　部分耦合增广随机梯度算法的计算量

Table 2　The computational efficiency of the PC-ESG algorithm

计算次序	乘法次数	加法次数
$\begin{bmatrix}\hat{\boldsymbol{\theta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\theta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\hat{\boldsymbol{\psi}}_i(t)}{r_i(t)}\left\{y_i(t)-\hat{\boldsymbol{\psi}}_i^{\mathrm{T}}(t)\begin{bmatrix}\hat{\boldsymbol{\theta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}\right\}$	$2m(n+mn_d)+m$	$2m(n+mn_d)$
$r_i(t)=r_i(t-1)+\|\hat{\boldsymbol{\psi}}_i(t)\|^2$	$m(n+mn_d)$	$m(n+mn_d)$
总数	$3m(n+mn_d)+m$	$3m(n+mn_d)$
总 flop 数	$6m(n+mn_d)+m$	

上述算法中，$\hat{\boldsymbol{\theta}}_i(t)\in\mathbf{R}^n$ 和 $\hat{\boldsymbol{\alpha}}_i(t)\in\mathbf{R}^{mn_d}$ 分别为第 i 个子系统在时刻 t 的参数估计向量，并且各个子系统辨识算法间的参数估计都是耦合的. 去掉子系统辨识算法间 $r_i(t)$ 的耦合，便得到一个新的部分耦合增广随机梯度算法：

$$\begin{bmatrix}\hat{\boldsymbol{\theta}}_1(t)\\ \hat{\boldsymbol{\alpha}}_1(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\theta}}_m(t-1)\\ \hat{\boldsymbol{\alpha}}_1(t-1)\end{bmatrix}+\frac{\hat{\boldsymbol{\psi}}_1(t)}{r_1(t)}\left\{y_1(t)-\hat{\boldsymbol{\psi}}_1^{\mathrm{T}}(t)\begin{bmatrix}\hat{\boldsymbol{\theta}}_m(t-1)\\ \hat{\boldsymbol{\alpha}}_1(t-1)\end{bmatrix}\right\},$$
$$\hat{\boldsymbol{\theta}}_m(0)=\mathbf{1}_n/p_0, \tag{60}$$

$$\begin{bmatrix}\hat{\boldsymbol{\theta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\theta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\hat{\boldsymbol{\psi}}_i(t)}{r_i(t)}\left\{y_i(t)-\hat{\boldsymbol{\psi}}_i^{\mathrm{T}}(t)\begin{bmatrix}\hat{\boldsymbol{\theta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}\right\},$$
$$i=2,3,\cdots,m, \tag{61}$$

$$r_i(t)=r_i(t-1)+\|\hat{\boldsymbol{\psi}}_i(t)\|^2,\quad r_i(0)=1,\quad i=1,2,\cdots,m, \tag{62}$$

$$\hat{\boldsymbol{\psi}}_i(t)=\begin{bmatrix}\boldsymbol{\phi}_i(t)\\ \hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\end{bmatrix},\quad \hat{\boldsymbol{\alpha}}_i(0)=\mathbf{1}_{mn_d}/p_0, \tag{63}$$

$$\hat{\boldsymbol{\varphi}}_{\rm n}(t)=[\hat{\boldsymbol{v}}^{\rm T}(t-1),\hat{\boldsymbol{v}}^{\rm T}(t-2),\cdots,\hat{\boldsymbol{v}}^{\rm T}(t-n_d)]^{\rm T}, \tag{64}$$

$$\hat{\boldsymbol{v}}(t)=[\hat{v}_1(t),\hat{v}_2(t),\cdots,\hat{v}_m(t)]^{\rm T}, \tag{65}$$

$$\hat{v}_i(t)=y_i(t)-\boldsymbol{\phi}_i^{\rm T}(t)\hat{\boldsymbol{\theta}}_i(t)-\hat{\boldsymbol{\varphi}}_{\rm n}^{\rm T}(t)\hat{\boldsymbol{\alpha}}_i(t), \tag{66}$$

$$\boldsymbol{\Phi}_{\rm s}(t)=[\boldsymbol{\phi}_1(t),\boldsymbol{\phi}_2(t),\cdots,\boldsymbol{\phi}_m(t)]^{\rm T}. \tag{67}$$

在时刻 t 第 m 个子系统参数估计定义为系统的参数估计:$\hat{\boldsymbol{\theta}}(t):=\hat{\boldsymbol{\theta}}_m(t)$,其示意如图 1 所示[1].

PC-ESG 算法(60)—(67)计算参数估计向量 $\hat{\boldsymbol{\alpha}}_i(t)$ 和 $\hat{\boldsymbol{\theta}}_m(t)$ 的步骤如下.

① 置初值:令 $t=1$, $\hat{\boldsymbol{\theta}}_m(0)=\mathbf{1}_n/p_0$, $r_i(0)=1$, $\hat{\boldsymbol{\alpha}}_i(0)=\mathbf{1}_{mn_d}/p_0$, $\hat{\boldsymbol{v}}(-j)=\mathbf{1}_m/p_0$, $j=0,1,\cdots,n_d$, $i=1,2,\cdots,m$, $p_0=10^6$.

② 收集观测数据 $\boldsymbol{y}(t)$ 和 $\boldsymbol{\Phi}_{\rm s}(t)$,用式(64)构成噪声信息向量 $\hat{\boldsymbol{\varphi}}_{\rm n}(t)$,根据式(67)从 $\boldsymbol{\Phi}_{\rm s}(t)$ 读出 $\boldsymbol{\varphi}_i(t)$,以构成式(63)中的 $\hat{\boldsymbol{\psi}}_i(t)$.

③ 用式(62)计算 $r_i(t)$,用式(60)刷新参数估计 $\hat{\boldsymbol{\alpha}}_1(t)$ 和 $\hat{\boldsymbol{\theta}}_1(t)$.

④ 用式(61)刷新参数估计向量 $\hat{\boldsymbol{\alpha}}_i(t)$ 和 $\hat{\boldsymbol{\theta}}_i(t)$.

⑤ 用式(66)计算 $\hat{v}_i(t)$,用式(65)构成 $\hat{\boldsymbol{v}}(t)$.

⑥ t 增 1,转到第②步.

为了提高参数估计收敛速度,可以在式(60)和(61)中引入收敛指数 $\frac{1}{2}<\varepsilon\leqslant 1$,在式(62)中引入遗忘因子 $0\leqslant\lambda\leqslant 1$,得到

$$\begin{bmatrix}\hat{\boldsymbol{\theta}}_1(t)\\ \hat{\boldsymbol{\alpha}}_1(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\theta}}_m(t-1)\\ \hat{\boldsymbol{\alpha}}_1(t-1)\end{bmatrix}+\frac{\hat{\boldsymbol{\psi}}_1(t)}{r_1^{\varepsilon}(t)}\left\{y_1(t)-\hat{\boldsymbol{\psi}}_1^{\rm T}(t)\begin{bmatrix}\hat{\boldsymbol{\theta}}_m(t-1)\\ \hat{\boldsymbol{\alpha}}_1(t-1)\end{bmatrix}\right\},$$
$$\hat{\boldsymbol{\theta}}_m(0)=\mathbf{1}_n/p_0, \tag{68}$$

$$\begin{bmatrix}\hat{\boldsymbol{\theta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\theta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\hat{\boldsymbol{\psi}}_i(t)}{r_i^{\varepsilon}(t)}\left\{y_i(t)-\hat{\boldsymbol{\psi}}_i^{\rm T}(t)\begin{bmatrix}\hat{\boldsymbol{\theta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}\right\},$$
$$i=2,3,\cdots,m, \tag{69}$$

$$r_i(t)=\lambda r_i(t-1)+\|\hat{\boldsymbol{\psi}}_i(t)\|^2,r_i(0)=1,i=1,2,\cdots,m. \tag{70}$$

式(63)—(70)构成了遗忘因子修正 PC-ESG 辨识算法.

1.3 部分耦合(子系统)多新息增广随机梯度辨识算法

为了改进 PC-S-ESG 算法的收敛速度,下面通过引入新息长度,推导部分耦合多新息增广随机梯度算法.定义子系统的信息矩阵 $\boldsymbol{\Gamma}_i(p,t)$ 和堆积输出向量 $\boldsymbol{Y}_i(p,t)$ 如下:

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

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

令 $p\geqslant 1$ 为新息长度.应用多新息辨识理论,推广 PC-S-ESG 算法(28)—(35)中的标量新息

$$e_i(t):=y_i(t)-\hat{\boldsymbol{\psi}}_i^{\rm T}(t)\begin{bmatrix}\hat{\boldsymbol{\theta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}=$$
$$y_i(t)-\boldsymbol{\phi}_i^{\rm T}(t)\hat{\boldsymbol{\theta}}(t-1)-\hat{\boldsymbol{\varphi}}_{\rm n}^{\rm T}(t)\hat{\boldsymbol{\alpha}}_i(t-1)\in\mathbf{R}$$

为新息向量

$$\boldsymbol{E}_i(p,t):=$$
$$\begin{bmatrix}y_i(t)-\hat{\boldsymbol{\psi}}_i^{\rm T}(t)\begin{bmatrix}\hat{\boldsymbol{\theta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}\\ y_i(t-1)-\hat{\boldsymbol{\psi}}_i^{\rm T}(t-1)\begin{bmatrix}\hat{\boldsymbol{\theta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}\\ \vdots\\ y_i(t-p+1)-\hat{\boldsymbol{\psi}}_i^{\rm T}(t-p+1)\begin{bmatrix}\hat{\boldsymbol{\theta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}\end{bmatrix}=$$
$$\begin{bmatrix}y_i(t)-[\boldsymbol{\phi}_i^{\rm T}(t)\hat{\boldsymbol{\theta}}(t-1)+\hat{\boldsymbol{\varphi}}_{\rm n}^{\rm T}(t)\hat{\boldsymbol{\alpha}}_i(t-1)]\\ y_i(t-1)-[\boldsymbol{\phi}_i^{\rm T}(t-1)\hat{\boldsymbol{\theta}}(t-1)+\hat{\boldsymbol{\varphi}}_{\rm n}^{\rm T}(t-1)\hat{\boldsymbol{\alpha}}_i(t-1)]\\ \vdots\\ y_i(t-p+1)-[\boldsymbol{\phi}_i^{\rm T}(t-p+1)\hat{\boldsymbol{\theta}}(t-1)+\hat{\boldsymbol{\varphi}}_{\rm n}^{\rm T}(t-p+1)\hat{\boldsymbol{\alpha}}_i(t-1)]\end{bmatrix}=$$

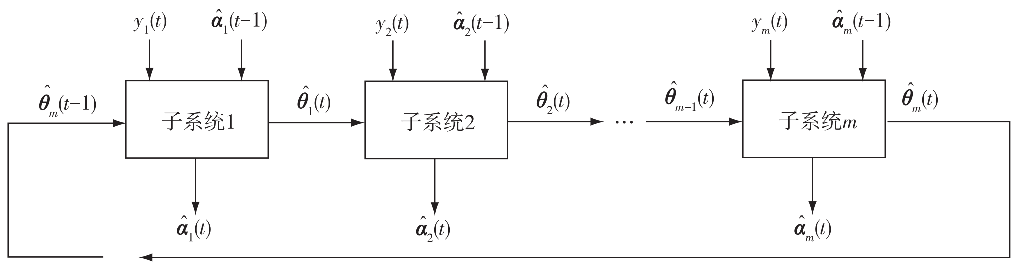

图 1 部分耦合增广随机梯度辨识算法示意

Fig. 1 The schematic diagram of the PC-ESG algorithm

$$\boldsymbol{Y}_i(p,t)-\boldsymbol{\Gamma}_i^{\mathrm{T}}(p,t)\begin{bmatrix}\hat{\boldsymbol{\theta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}\in\mathbf{R}^p.$$

1）部分耦合子系统多新息增广随机梯度算法

参考多新息随机梯度辨识算法与部分耦合随机梯度辨识算法的结构形式[1]，从部分耦合子系统增广随机梯度（PC-S-ESG）算法（28）—（35）可以得到简单的部分耦合子系统多新息增广随机梯度算法（Partially Coupled Subsystem Multi-Innovation Extended Stochastic Gradient algorithm，PC-S-MI-ESG 算法）：

$$\begin{bmatrix}\hat{\boldsymbol{\theta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\theta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\boldsymbol{\Gamma}_i(p,t)}{r_i(t)}\boldsymbol{E}_i(p,t),$$

$$\hat{\boldsymbol{\theta}}(0)=\mathbf{1}_n/p_0,\quad \hat{\boldsymbol{\alpha}}_i(0)=\mathbf{1}_{mn_d}/p_0, \tag{71}$$

$$\boldsymbol{E}_i(p,t)=\boldsymbol{Y}_i(p,t)-\boldsymbol{\Gamma}_i^{\mathrm{T}}(p,t)\begin{bmatrix}\hat{\boldsymbol{\theta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}, \tag{72}$$

$$r_i(t)=r_i(t-1)+\|\hat{\boldsymbol{\psi}}_i(t)\|^2=$$
$$r_i(t-1)+\|\boldsymbol{\phi}_i(t)\|^2+\|\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\|^2,\quad r_i(0)=1, \tag{73}$$

$$\boldsymbol{\Gamma}_i(p,t)=[\hat{\boldsymbol{\psi}}_i(t),\hat{\boldsymbol{\psi}}_i(t-1),\cdots,\hat{\boldsymbol{\psi}}_i(t-p+1)], \tag{74}$$

$$\boldsymbol{Y}_i(p,t)=[y_i(t),y_i(t-1),\cdots,y_i(t-p+1)]^{\mathrm{T}}, \tag{75}$$

$$\hat{\boldsymbol{\psi}}_i(t)=\begin{bmatrix}\boldsymbol{\phi}_i(t)\\ \hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\end{bmatrix},\quad i=1,2,\cdots,m, \tag{76}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)=[\hat{\boldsymbol{v}}^{\mathrm{T}}(t-1),\hat{\boldsymbol{v}}^{\mathrm{T}}(t-2),\cdots,\hat{\boldsymbol{v}}^{\mathrm{T}}(t-n_d)]^{\mathrm{T}}, \tag{77}$$

$$\hat{\boldsymbol{v}}(t)=\boldsymbol{y}(t)-\boldsymbol{\Phi}_{\mathrm{s}}(t)\hat{\boldsymbol{\theta}}(t)-\hat{\boldsymbol{\alpha}}^{\mathrm{T}}(t)\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t), \tag{78}$$

$$\boldsymbol{\Phi}_{\mathrm{s}}(t)=[\boldsymbol{\phi}_1(t),\boldsymbol{\phi}_2(t),\cdots,\boldsymbol{\phi}_m(t)]^{\mathrm{T}}, \tag{79}$$

$$\hat{\boldsymbol{\theta}}(t)=\frac{\hat{\boldsymbol{\theta}}_1(t)+\hat{\boldsymbol{\theta}}_2(t)+\cdots+\hat{\boldsymbol{\theta}}_m(t)}{m}, \tag{80}$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\boldsymbol{\alpha}}_1(t),\hat{\boldsymbol{\alpha}}_2(t),\cdots,\hat{\boldsymbol{\alpha}}_m(t)]. \tag{81}$$

当然，也可以进一步引入遗忘因子或收敛指数，得到相应的遗忘因子 PC-S-MI-ESG 算法和修正 PC-S-MI-ESG 算法.

① 遗忘因子 PC-S-MI-ESG 算法：

$$\begin{bmatrix}\hat{\boldsymbol{\theta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\theta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\boldsymbol{\Gamma}_i(p,t)}{r_i(t)}\boldsymbol{E}_i(p,t), \tag{82}$$

$$\boldsymbol{E}_i(p,t)=\boldsymbol{Y}_i(p,t)-\boldsymbol{\Gamma}_i^{\mathrm{T}}(p,t)\begin{bmatrix}\hat{\boldsymbol{\theta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}, \tag{83}$$

$$r_i(t)=\lambda r_i(t-1)+\|\boldsymbol{\phi}_i(t)\|^2+\|\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\|^2,$$
$$0\leqslant\lambda\leqslant 1, \tag{84}$$

$$\boldsymbol{\Gamma}_i(p,t)=[\hat{\boldsymbol{\psi}}_i(t),\hat{\boldsymbol{\psi}}_i(t-1),\cdots,\hat{\boldsymbol{\psi}}_i(t-p+1)], \tag{85}$$

$$\boldsymbol{Y}_i(p,t)=[y_i(t),y_i(t-1),\cdots,y_i(t-p+1)]^{\mathrm{T}}, \tag{86}$$

$$\hat{\boldsymbol{\psi}}_i(t)=\begin{bmatrix}\boldsymbol{\phi}_i(t)\\ \hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\end{bmatrix},i=1,2,\cdots,m, \tag{87}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)=[\hat{\boldsymbol{v}}^{\mathrm{T}}(t-1),\hat{\boldsymbol{v}}^{\mathrm{T}}(t-2),\cdots,\hat{\boldsymbol{v}}^{\mathrm{T}}(t-n_d)]^{\mathrm{T}}, \tag{88}$$

$$\hat{\boldsymbol{v}}(t)=\boldsymbol{y}(t)-\boldsymbol{\Phi}_{\mathrm{s}}(t)\hat{\boldsymbol{\theta}}(t)-\hat{\boldsymbol{\alpha}}^{\mathrm{T}}(t)\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t), \tag{89}$$

$$\boldsymbol{\Phi}_{\mathrm{s}}(t)=[\boldsymbol{\phi}_1(t),\boldsymbol{\phi}_2(t),\cdots,\boldsymbol{\phi}_m(t)]^{\mathrm{T}}, \tag{90}$$

$$\hat{\boldsymbol{\theta}}(t)=\frac{\hat{\boldsymbol{\theta}}_1(t)+\hat{\boldsymbol{\theta}}_2(t)+\cdots+\hat{\boldsymbol{\theta}}_m(t)}{m}, \tag{91}$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\boldsymbol{\alpha}}_1(t),\hat{\boldsymbol{\alpha}}_2(t),\cdots,\hat{\boldsymbol{\alpha}}_m(t)]. \tag{92}$$

② 修正 PC-S-MI-ESG 算法：

$$\begin{bmatrix}\hat{\boldsymbol{\theta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\theta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\boldsymbol{\Gamma}_i(p,t)}{r_i^{\varepsilon}(t)}\boldsymbol{E}_i(p,t),$$
$$\frac{1}{2}<\varepsilon\leqslant 1, \tag{93}$$

$$\boldsymbol{E}_i(p,t)=\boldsymbol{Y}_i(p,t)-\boldsymbol{\Gamma}_i^{\mathrm{T}}(p,t)\begin{bmatrix}\hat{\boldsymbol{\theta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}, \tag{94}$$

$$r_i(t)=r_i(t-1)+\|\boldsymbol{\phi}_i(t)\|^2+\|\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\|^2, \tag{95}$$

$$\boldsymbol{\Gamma}_i(p,t)=[\hat{\boldsymbol{\psi}}_i(t),\hat{\boldsymbol{\psi}}_i(t-1),\cdots,\hat{\boldsymbol{\psi}}_i(t-p+1)], \tag{96}$$

$$\boldsymbol{Y}_i(p,t)=[y_i(t),y_i(t-1),\cdots,y_i(t-p+1)]^{\mathrm{T}}, \tag{97}$$

$$\hat{\boldsymbol{\psi}}_i(t)=\begin{bmatrix}\boldsymbol{\phi}_i(t)\\ \hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\end{bmatrix},\quad i=1,2,\cdots,m, \tag{98}$$

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

$$\hat{\boldsymbol{v}}(t)=\boldsymbol{y}(t)-\boldsymbol{\Phi}_{\mathrm{s}}(t)\hat{\boldsymbol{\theta}}(t)-\hat{\boldsymbol{\alpha}}^{\mathrm{T}}(t)\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t), \tag{100}$$

$$\boldsymbol{\Phi}_{\mathrm{s}}(t)=[\boldsymbol{\phi}_1(t),\boldsymbol{\phi}_2(t),\cdots,\boldsymbol{\phi}_m(t)]^{\mathrm{T}}, \tag{101}$$

$$\hat{\boldsymbol{\theta}}(t)=\frac{\hat{\boldsymbol{\theta}}_1(t)+\hat{\boldsymbol{\theta}}_2(t)+\cdots+\hat{\boldsymbol{\theta}}_m(t)}{m}, \tag{102}$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\boldsymbol{\alpha}}_1(t),\hat{\boldsymbol{\alpha}}_2(t),\cdots,\hat{\boldsymbol{\alpha}}_m(t)]. \tag{103}$$

2）部分耦合多新息增广随机梯度算法

借助于多新息辨识理论，从部分耦合增广随机梯度算法（60）—（67）可以得到部分耦合多新息增广随机梯度算法（Partially Coupled Multi-Innovation Extended Stochastic Gradient algorithm，PC-MI-ESG 算法）：

$$\begin{bmatrix}\hat{\boldsymbol{\theta}}_1(t)\\ \hat{\boldsymbol{\alpha}}_1(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\theta}}_m(t-1)\\ \hat{\boldsymbol{\alpha}}_1(t-1)\end{bmatrix}+\frac{\boldsymbol{\Gamma}_1(p,t)}{r_1(t)}\boldsymbol{E}_1(p,t), \tag{104}$$

$$\boldsymbol{E}_1(p,t)=\boldsymbol{Y}_1(p,t)-\boldsymbol{\Gamma}_1^{\mathrm{T}}(p,t)\begin{bmatrix}\hat{\boldsymbol{\theta}}_m(t-1)\\ \hat{\boldsymbol{\alpha}}_1(t-1)\end{bmatrix}, \tag{105}$$

$$\begin{bmatrix}\hat{\boldsymbol{\theta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\theta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\boldsymbol{\Gamma}_i(p,t)}{r_i(t)}\boldsymbol{E}_i(p,t),$$
$$i=2,3,\cdots,m, \tag{106}$$

$$\boldsymbol{E}_i(p,t)=\boldsymbol{Y}_i(p,t)-\boldsymbol{\Gamma}_i^{\mathrm{T}}(p,t)\begin{bmatrix}\hat{\boldsymbol{\theta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}, \tag{107}$$

$$r_i(t)=r_i(t-1)+\|\hat{\boldsymbol{\psi}}_i(t)\|^2,i=1,2,\cdots,m, \tag{108}$$

$$\boldsymbol{\Gamma}_i(p,t)=[\hat{\boldsymbol{\psi}}_i(t),\hat{\boldsymbol{\psi}}_i(t-1),\cdots,\hat{\boldsymbol{\psi}}_i(t-p+1)], \tag{109}$$

$$\boldsymbol{Y}_i(p,t)=[y_i(t),y_i(t-1),\cdots,y_i(t-p+1)]^{\mathrm{T}}, \tag{110}$$

$$\hat{\boldsymbol{\psi}}_i(t)=\begin{bmatrix}\boldsymbol{\phi}_i(t)\\ \hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\end{bmatrix}, \tag{111}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)=[\hat{\boldsymbol{v}}^{\mathrm{T}}(t-1),\hat{\boldsymbol{v}}^{\mathrm{T}}(t-2),\cdots,\hat{\boldsymbol{v}}^{\mathrm{T}}(t-n_d)]^{\mathrm{T}}, \tag{112}$$

$$\hat{\boldsymbol{v}}(t)=[\hat{v}_1(t),\hat{v}_2(t),\cdots,\hat{v}_m(t)]^{\mathrm{T}}, \tag{113}$$

$$\hat{v}_i(t)=y_i(t)-\boldsymbol{\phi}_i^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_i(t)-\hat{\boldsymbol{\varphi}}_{\mathrm{n}}^{\mathrm{T}}(t)\hat{\boldsymbol{\alpha}}_i(t), \tag{114}$$

$$\boldsymbol{\Phi}_{\mathrm{s}}(t)=[\boldsymbol{\phi}_1(t),\boldsymbol{\phi}_2(t),\cdots,\boldsymbol{\phi}_m(t)]^{\mathrm{T}}. \tag{115}$$

PC-MI-ESG 算法的计算量如表 3 所示. PC-MI-ESG 算法(104)—(115)计算参数估计向量 $\hat{\boldsymbol{\alpha}}_i(t)$ 和 $\hat{\boldsymbol{\theta}}_m(t)$ 的步骤如下.

表 3　PC-MI-ESG 算法的计算量

Table 3　The computational efficiency of the PC-MI-ESG algorithm

计算次序	乘法次数	加法次数
$\begin{bmatrix}\hat{\boldsymbol{\theta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\theta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\boldsymbol{\Gamma}_i(p,t)}{r_i(t)}\boldsymbol{E}_i(p,t)\in\mathbf{R}^{n+mn_d}$	$m(p+1)(n+mn_d)$	$mp(n+mn_d)$
$\boldsymbol{E}_i(p,t)=\boldsymbol{Y}_i(p,t)-\boldsymbol{\Gamma}_i^{\mathrm{T}}(p,t)\begin{bmatrix}\hat{\boldsymbol{\theta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}\in\mathbf{R}^p$	$mp(n+mn_d)$	$mp(n+mn_d)$
$r_i(t)=r_i(t-1)+\|\hat{\boldsymbol{\psi}}_i(t)\|^2$	$m(n+mn_d)$	$m(n+mn_d)$
总数	$m(2p+2)(n+mn_d)$	$m(2p+1)(n+mn_d)$
总 flop 数	$m(4p+3)(n+mn_d)$	

① 置初值:令 $t=1$, $\hat{\boldsymbol{\theta}}_m(0)=\mathbf{1}_n/p_0$, $r_i(0)=1$, $\hat{\boldsymbol{\alpha}}_i(0)=\mathbf{1}_{mn_d}/p_0$, $\hat{\boldsymbol{v}}(-j)=1/p_0$, $j=0,1,\cdots,n_d$, $i=1,2,\cdots,m$, $p_0=10^6$.

② 用式(112)构成噪声信息向量 $\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)$,从式(115)$\boldsymbol{\Phi}_{\mathrm{s}}(t)$中读出信息向量 $\boldsymbol{\phi}_i(t)$,构成式(111)中的 $\hat{\boldsymbol{\psi}}_i(t)$.

③ 用式(108)计算 $r_i(t)$,用式(109)和(110)构成 $\boldsymbol{\Gamma}_i(p,t)$ 和 $\boldsymbol{Y}_i(p,t)$,用式(105)计算 $\boldsymbol{E}_1(p,t)$.

④ 用式(104)刷新参数估计 $\hat{\boldsymbol{\theta}}_1(t)$ 和 $\hat{\boldsymbol{\alpha}}_1(t)$.

⑤ 依次当 $i=2,3,\cdots,m$ 时,用式(107)计算 $\boldsymbol{E}_i(p,t)$,用式(106)刷新参数估计向量 $\hat{\boldsymbol{\theta}}_i(t)$ 和 $\hat{\boldsymbol{\alpha}}_i(t)$.

⑥ 用式(114)计算 $\hat{v}_i(t)$,用式(113)构成 $\hat{\boldsymbol{v}}(t)$.

⑦ t 增 1,转到第②步.

当然,也可以进一步引入遗忘因子或收敛指数,得到遗忘因子部分耦合多新息增广随机梯度算法(PC-MI-ESG 算法)及修正部分耦合多新息增广随机梯度算法(PC-MI-ESG 算法).

基于上述多元伪线性滑动平均系统的辨识方法,通过类比,下面简单叙述多元伪线性自回归滑动平均系统的部分耦合梯度辨识方法.

2　多元伪线性自回归滑动平均系统(1)

考虑多元伪线性自回归滑动平均系统

$$\boldsymbol{y}(t)=\boldsymbol{\Phi}_{\mathrm{s}}(t)\boldsymbol{\theta}+\frac{\boldsymbol{D}(z)}{c(z)}\boldsymbol{v}(t), \tag{116}$$

其中 $\boldsymbol{y}(t):=[y_1(t),y_2(t),\cdots,y_m(t)]^{\mathrm{T}}\in\mathbf{R}^m$ 为观测输出向量,$\boldsymbol{\Phi}_{\mathrm{s}}(t)\in\mathbf{R}^{m\times n}$ 是由输入输出数据构成的回归信息矩阵,$\boldsymbol{\theta}\in\mathbf{R}^n$ 为系统参数向量,$\boldsymbol{w}(t):=\frac{\boldsymbol{D}(z)}{c(z)}\boldsymbol{v}(t)\in\mathbf{R}^m$ 为 m 维干扰噪声向量,$\boldsymbol{v}(t)\in\mathbf{R}^m$ 为 m 维白噪声向量,$c(z)$ 和 $\boldsymbol{D}(z)$ 为单位后移算子 z^{-1} 的多项式和多项式矩阵:

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

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

假设 m,n,n_c 和 n_d 已知,且当 $t\leqslant 0$ 时,$\boldsymbol{y}(t)=0$, $\boldsymbol{\Phi}_{\mathrm{s}}(t)=0$, $\boldsymbol{v}(t)=0$. 辨识的目标是基于耦合辨识概念和多新息辨识理论,利用系统的观测数据 $\{\boldsymbol{y}(t),\boldsymbol{\Phi}_{\mathrm{s}}(t):t=1,2,3,\cdots\}$ 提出新的算法,估计系统模型参数向量 $\boldsymbol{\theta}$,以及噪声模型的参数 c_i 和参数矩阵 $\boldsymbol{D}_i$.

2.1　部分耦合(子系统)广义增广随机梯度辨识算法

定义参数向量 $\boldsymbol{\vartheta}$,参数矩阵 $\boldsymbol{\alpha}$,信息矩阵 $\boldsymbol{\Phi}(t)$ 和信息向量 $\boldsymbol{\varphi}_{\mathrm{n}}(t)$ 如下:

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

$$\boldsymbol{\alpha}^{\mathrm{T}}:=[\boldsymbol{D}_1,\boldsymbol{D}_2,\cdots,\boldsymbol{D}_{n_d}]=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m]^{\mathrm{T}}\in\mathbf{R}^{m\times(mn_d)},$$

$$\boldsymbol{\Phi}(t):=[\boldsymbol{\Phi}_{\mathrm{s}}(t),-\boldsymbol{w}(t-1),-\boldsymbol{w}(t-2),\cdots,-\boldsymbol{w}(t-n_c)]=[\boldsymbol{\phi}_1(t),\boldsymbol{\phi}_2(t),\cdots,\boldsymbol{\phi}_m(t)]^{\mathrm{T}}\in\mathbf{R}^{m\times(n+n_c)},$$

$\boldsymbol{\varphi}_{\mathrm{n}}(t):=[\boldsymbol{v}^{\mathrm{T}}(t-1),\boldsymbol{v}^{\mathrm{T}}(t-2),\cdots,\boldsymbol{v}^{\mathrm{T}}(t-n_d)]^{\mathrm{T}}\in\mathbf{R}^{mn_d}$,
于是模型(116)可以表示为

$$\boldsymbol{y}(t)=\boldsymbol{\Phi}(t)\boldsymbol{\vartheta}+\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\varphi}_{\mathrm{n}}(t)+\boldsymbol{v}(t). \tag{117}$$

令第 i 个子系统信息向量表示为 $\boldsymbol{\psi}_i(t):=\begin{bmatrix}\boldsymbol{\phi}_i(t)\\ \boldsymbol{\varphi}_{\mathrm{n}}(t)\end{bmatrix}\in\mathbf{R}^{n+n_c+mn_d}$,于是第 i 个子系统可以写为

$$\begin{aligned}y_i(t)&=\boldsymbol{\phi}_i^{\mathrm{T}}(t)\boldsymbol{\vartheta}+\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{\varphi}_{\mathrm{n}}(t)+v_i(t)=\\ &\boldsymbol{\phi}_i^{\mathrm{T}}(t)\boldsymbol{\vartheta}+\boldsymbol{\varphi}_{\mathrm{n}}^{\mathrm{T}}(t)\boldsymbol{\alpha}_i+v_i(t)=\\ &[\boldsymbol{\phi}_i^{\mathrm{T}}(t),\boldsymbol{\varphi}_{\mathrm{n}}^{\mathrm{T}}(t)]\begin{bmatrix}\boldsymbol{\vartheta}\\ \boldsymbol{\alpha}_i\end{bmatrix}+v_i(t)=\\ &\boldsymbol{\psi}_i^{\mathrm{T}}(t)\begin{bmatrix}\boldsymbol{\vartheta}\\ \boldsymbol{\alpha}_i\end{bmatrix}+v_i(t),\quad i=1,2,\cdots,m.\end{aligned} \tag{118}$$

1)子系统广义增广随机梯度辨识算法

参照子系统增广随机梯度算法的推导,可以得到子系统广义增广随机梯度算法(Subsystem Generalized Extended Stochastic Gradient algorithm,S-GESG 算法):

$$\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_i(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\hat{\boldsymbol{\psi}}_i(t)}{r_i(t)}\left\{y_i(t)-\hat{\boldsymbol{\psi}}_i^{\mathrm{T}}(t)\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_i(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}\right\},$$

$$\hat{\boldsymbol{\vartheta}}_i(0)=\mathbf{1}_{n+n_c}/p_0,\quad \hat{\boldsymbol{\alpha}}_i(0)=\mathbf{1}_{mn_d}/p_0, \tag{119}$$

$$r_i(t)=r_i(t-1)+\|\hat{\boldsymbol{\psi}}_i(t)\|^2,\quad r_i(0)=1, \tag{120}$$

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

$$\begin{aligned}\boldsymbol{\Phi}(t)&=[\hat{\boldsymbol{\phi}}_1(t),\hat{\boldsymbol{\phi}}_2(t),\cdots,\hat{\boldsymbol{\phi}}_m(t)]^{\mathrm{T}}=\\ &[\boldsymbol{\Phi}_{\mathrm{s}}(t),-\hat{\boldsymbol{w}}(t-1),-\hat{\boldsymbol{w}}(t-2),\cdots,-\hat{\boldsymbol{w}}(t-n_c)],\end{aligned} \tag{122}$$

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

$$\hat{\boldsymbol{\vartheta}}_i(t)=[\hat{\boldsymbol{\theta}}_i^{\mathrm{T}}(t),\hat{c}_{i1}(t),\hat{c}_{i2}(t),\cdots,\hat{c}_{in_c}(t)]^{\mathrm{T}}, \tag{124}$$

$$\hat{\boldsymbol{w}}(t)=[\hat{w}_1(t),\hat{w}_2(t),\cdots,\hat{w}_m(t)]^{\mathrm{T}}, \tag{125}$$

$$\hat{w}_i(t)=y_i(t)-[\boldsymbol{\Phi}_{\mathrm{s}}(t)\text{的第 }i\text{ 行}]\hat{\boldsymbol{\theta}}_i(t), \tag{126}$$

$$\hat{\boldsymbol{v}}(t)=[\hat{v}_1(t),\hat{v}_2(t),\cdots,\hat{v}_m(t)]^{\mathrm{T}}, \tag{127}$$

$$\hat{v}_i(t)=y_i(t)-\hat{\boldsymbol{\phi}}_i(t)\hat{\boldsymbol{\vartheta}}_i(t)-\hat{\boldsymbol{\alpha}}_i^{\mathrm{T}}(t)\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t). \tag{128}$$

2)部分耦合子系统广义增广随机梯度辨识算法

为了解决上述 S-GESG 算法中的冗余估计 $\hat{\boldsymbol{\vartheta}}_i(t)$,可以用其平均值 $\hat{\boldsymbol{\vartheta}}(t-1)$ 代替上述算法中的 $\hat{\boldsymbol{\vartheta}}_i(t-1)$,根据耦合辨识概念,能够得到部分耦合子系统广义增广随机梯度算法(Partially Coupled Subsystem Generalized Extended Stochastic Gradient algorithm,PC-S-GESG 算法):

$$\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\vartheta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\hat{\boldsymbol{\psi}}_i(t)}{r_i(t)}\left\{y_i(t)-\hat{\boldsymbol{\psi}}_i^{\mathrm{T}}(t)\begin{bmatrix}\hat{\boldsymbol{\vartheta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}\right\},$$

$$\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_{n+n_c}/p_0,\quad \hat{\boldsymbol{\alpha}}_i(0)=\mathbf{1}_{mn_d}/p_0, \tag{129}$$

$$r_i(t)=r_i(t-1)+\|\hat{\boldsymbol{\psi}}_i(t)\|^2,\quad r_i(0)=1, \tag{130}$$

$$\hat{\boldsymbol{\psi}}_i(t)=[\hat{\boldsymbol{\phi}}_i^{\mathrm{T}}(t),\hat{\boldsymbol{\varphi}}_{\mathrm{n}}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{131}$$

$$\begin{aligned}\hat{\boldsymbol{\Phi}}(t)&=[\hat{\boldsymbol{\phi}}_1(t),\hat{\boldsymbol{\phi}}_2(t),\cdots,\hat{\boldsymbol{\phi}}_m(t)]^{\mathrm{T}}=\\ &[\boldsymbol{\Phi}_{\mathrm{s}}(t),-\hat{\boldsymbol{w}}(t-1),-\hat{\boldsymbol{w}}(t-2),\cdots,-\hat{\boldsymbol{w}}(t-n_c)],\end{aligned} \tag{132}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)=[\hat{\boldsymbol{v}}^{\mathrm{T}}(t-1),\hat{\boldsymbol{v}}^{\mathrm{T}}(t-2),\cdots,\hat{\boldsymbol{v}}^{\mathrm{T}}(t-n_d)]^{\mathrm{T}}, \tag{133}$$

$$\hat{\boldsymbol{\vartheta}}_i(t)=[\hat{\boldsymbol{\theta}}_i^{\mathrm{T}}(t),\hat{c}_{i1}(t),\hat{c}_{i2}(t),\cdots,\hat{c}_{in_c}(t)]^{\mathrm{T}}, \tag{134}$$

$$\hat{\boldsymbol{w}}(t)=\boldsymbol{y}(t)-\boldsymbol{\Phi}_{\mathrm{s}}(t)\hat{\boldsymbol{\theta}}(t), \tag{135}$$

$$\hat{\boldsymbol{v}}(t)=\boldsymbol{y}(t)-\hat{\boldsymbol{\Phi}}(t)\hat{\boldsymbol{\vartheta}}(t)-\hat{\boldsymbol{\alpha}}^{\mathrm{T}}(t)\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t), \tag{136}$$

$$\hat{\boldsymbol{\theta}}(t)=\frac{\hat{\boldsymbol{\theta}}_1(t)+\hat{\boldsymbol{\theta}}_2(t)+\cdots+\hat{\boldsymbol{\theta}}_m(t)}{m}, \tag{137}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\frac{\hat{\boldsymbol{\vartheta}}_1(t)+\hat{\boldsymbol{\vartheta}}_2(t)+\cdots+\hat{\boldsymbol{\vartheta}}_m(t)}{m}, \tag{138}$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\boldsymbol{\alpha}}_1(t),\hat{\boldsymbol{\alpha}}_2(t),\cdots,\hat{\boldsymbol{\alpha}}_m(t)]. \tag{139}$$

3)部分耦合广义增广随机梯度辨识算法

利用耦合辨识概念,可以得到部分耦合广义增广随机梯度算法(Partially Coupled Generalized Extended Stochastic Gradient algorithm,PC-GESG 算法):

$$\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_1(t)\\ \hat{\boldsymbol{\alpha}}_1(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_m(t-1)\\ \hat{\boldsymbol{\alpha}}_1(t-1)\end{bmatrix}+\frac{\hat{\boldsymbol{\psi}}_1(t)}{r_1(t)}\left\{y_1(t)-\hat{\boldsymbol{\psi}}_1^{\mathrm{T}}(t)\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_m(t-1)\\ \hat{\boldsymbol{\alpha}}_1(t-1)\end{bmatrix}\right\},$$

$$\hat{\boldsymbol{\vartheta}}_m(0)=\mathbf{1}_{n+n_c}/p_0,\quad \hat{\boldsymbol{\alpha}}_1(0)=\mathbf{1}_{mn_d}/p_0, \tag{140}$$

$$r_1(t)=r_m(t-1)+\|\hat{\boldsymbol{\psi}}_1(t)\|^2,\quad r_1(0)=1, \tag{141}$$

$$\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\hat{\boldsymbol{\psi}}_i(t)}{r_i(t)}\left\{y_i(t)-\hat{\boldsymbol{\psi}}_i^{\mathrm{T}}(t)\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}\right\},$$

$$\hat{\boldsymbol{\alpha}}_i(0)=\mathbf{1}_{mn_d}/p_0,\quad i=2,3,\cdots,m, \tag{142}$$

$$r_i(t)=r_{i-1}(t)+\|\hat{\boldsymbol{\psi}}_i(t)\|^2,\quad r_i(0)=1, \tag{143}$$

$$\hat{\boldsymbol{\psi}}_i(t)=[\hat{\boldsymbol{\phi}}_i^{\mathrm{T}}(t),\hat{\boldsymbol{\varphi}}_{\mathrm{n}}^{\mathrm{T}}(t)]^{\mathrm{T}},\quad i=1,2,\cdots,m, \tag{144}$$

$$\begin{aligned}\hat{\boldsymbol{\Phi}}(t)&=[\hat{\boldsymbol{\phi}}_1(t),\hat{\boldsymbol{\phi}}_2(t),\cdots,\hat{\boldsymbol{\phi}}_m(t)]^{\mathrm{T}}=\\ &[\boldsymbol{\Phi}_{\mathrm{s}}(t),-\hat{\boldsymbol{w}}(t-1),-\hat{\boldsymbol{w}}(t-2),\cdots,-\hat{\boldsymbol{w}}(t-n_c)],\end{aligned} \tag{145}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)=[\hat{\boldsymbol{v}}^{\mathrm{T}}(t-1),\hat{\boldsymbol{v}}^{\mathrm{T}}(t-2),\cdots,\hat{\boldsymbol{v}}^{\mathrm{T}}(t-n_d)]^{\mathrm{T}}, \tag{146}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}_m(t)=[\hat{\boldsymbol{\theta}}_m^{\mathrm{T}}(t),\hat{c}_{m1}(t),\hat{c}_{m2}(t),\cdots,\hat{c}_{mn_c}(t)]^{\mathrm{T}}, \tag{147}$$

$$\hat{\boldsymbol{w}}(t)=\boldsymbol{y}(t)-\boldsymbol{\Phi}_{\mathrm{s}}(t)\hat{\boldsymbol{\theta}}_m(t), \tag{148}$$

$$\hat{\boldsymbol{v}}(t)=\boldsymbol{y}(t)-\hat{\boldsymbol{\Phi}}(t)\hat{\boldsymbol{\vartheta}}(t)-\hat{\boldsymbol{\alpha}}^{\mathrm{T}}(t)\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t), \tag{149}$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\boldsymbol{\alpha}}_1(t),\hat{\boldsymbol{\alpha}}_2(t),\cdots,\hat{\boldsymbol{\alpha}}_m(t)]. \tag{150}$$

去掉 $r_i(t)$ 间的耦合,式(141)和(143)更改为

$$r_i(t)=r_i(t-1)+\|\hat{\boldsymbol{\psi}}_i(t)\|^2,r_i(0)=1,i=1,2,\cdots,m. \tag{151}$$

2.2 部分耦合(子系统)多新息广义增广随机梯度辨识算法

1)部分耦合子系统多新息广义增广随机梯度

算法

定义信息矩阵 $\boldsymbol{\Gamma}_i(p,t)$,堆积输出向量 $\boldsymbol{Y}_i(p,t)$ 和新息向量 $\boldsymbol{E}_i(p,t)$ 如下:

$$\boldsymbol{\Gamma}_i(p,t):=[\hat{\boldsymbol{\psi}}_i(t),\hat{\boldsymbol{\psi}}_i(t-1),\cdots,\hat{\boldsymbol{\psi}}_i(t-p+1)]\in \mathbf{R}^{(n+mn_c+mn_d)\times p},$$

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

$$\boldsymbol{E}_i(p,t):=\begin{bmatrix} y_i(t)-\hat{\boldsymbol{\psi}}_i^{\mathrm{T}}(t)\begin{bmatrix}\hat{\boldsymbol{\vartheta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}\\ y_i(t-1)-\hat{\boldsymbol{\psi}}_i^{\mathrm{T}}(t-1)\begin{bmatrix}\hat{\boldsymbol{\vartheta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}\\ \vdots \\ y_i(t-p+1)-\hat{\boldsymbol{\psi}}_i^{\mathrm{T}}(t-p+1)\begin{bmatrix}\hat{\boldsymbol{\vartheta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}\end{bmatrix}\in \mathbf{R}^p,$$

则可以得到部分耦合子系统多新息广义增广随机梯度算法(Partially Coupled Subsystem Multi-Innovation Generalized Extended Stochastic Gradient algorithm, PC-S-MI-GESG 算法):

$$\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\vartheta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\boldsymbol{\Gamma}_i(p,t)}{r_i(t)}\boldsymbol{E}_i(p,t),$$

$$\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_{n+n_c}/p_0,\quad \hat{\boldsymbol{\alpha}}_i(0)=\mathbf{1}_{mn_d}/p_0, \tag{152}$$

$$\boldsymbol{E}_i(p,t)=\boldsymbol{Y}_i(p,t)-\boldsymbol{\Gamma}_i^{\mathrm{T}}(p,t)\begin{bmatrix}\hat{\boldsymbol{\vartheta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}, \tag{153}$$

$$r_i(t)=r_i(t-1)+\|\hat{\boldsymbol{\phi}}_i(t)\|^2+\|\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\|^2,\quad r_i(0)=1, \tag{154}$$

$$\boldsymbol{\Gamma}_i(p,t)=[\hat{\boldsymbol{\psi}}_i(t),\hat{\boldsymbol{\psi}}_i(t-1),\cdots,\hat{\boldsymbol{\psi}}_i(t-p+1)], \tag{155}$$

$$\boldsymbol{Y}_i(p,t)=[y_i(t),y_i(t-1),\cdots,y_i(t-p+1)]^{\mathrm{T}}, \tag{156}$$

$$\hat{\boldsymbol{\psi}}_i(t)=[\hat{\boldsymbol{\varphi}}_i^{\mathrm{T}}(t),\hat{\boldsymbol{\varphi}}_{\mathrm{n}}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{157}$$

$$\hat{\boldsymbol{\Phi}}(t)=[\hat{\boldsymbol{\phi}}_1(t),\hat{\boldsymbol{\phi}}_2(t),\cdots,\hat{\boldsymbol{\phi}}_m(t)]^{\mathrm{T}}=[\boldsymbol{\Phi}_{\mathrm{s}}(t),-\hat{\boldsymbol{w}}(t-1),-\hat{\boldsymbol{w}}(t-2),\cdots,-\hat{\boldsymbol{w}}(t-n_c)], \tag{158}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)=[\hat{\boldsymbol{v}}^{\mathrm{T}}(t-1),\hat{\boldsymbol{v}}^{\mathrm{T}}(t-2),\cdots,\hat{\boldsymbol{v}}^{\mathrm{T}}(t-n_d)]^{\mathrm{T}}, \tag{159}$$

$$\hat{\boldsymbol{\vartheta}}(t)=[\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t),\hat{c}_1(t),\hat{c}_2(t),\cdots,\hat{c}_{n_c}(t)]^{\mathrm{T}}, \tag{160}$$

$$\hat{\boldsymbol{w}}(t)=\boldsymbol{y}(t)-\boldsymbol{\Phi}_{\mathrm{s}}(t)\hat{\boldsymbol{\theta}}(t), \tag{161}$$

$$\hat{\boldsymbol{v}}(t)=\boldsymbol{y}(t)-\hat{\boldsymbol{\Phi}}(t)\hat{\boldsymbol{\vartheta}}(t)-\hat{\boldsymbol{\alpha}}^{\mathrm{T}}(t)\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t), \tag{162}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\frac{\hat{\boldsymbol{\vartheta}}_1(t)+\hat{\boldsymbol{\vartheta}}_2(t)+\cdots+\hat{\boldsymbol{\vartheta}}_m(t)}{m}, \tag{163}$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\boldsymbol{\alpha}}_1(t),\hat{\boldsymbol{\alpha}}_2(t),\cdots,\hat{\boldsymbol{\alpha}}_m(t)]. \tag{164}$$

2)部分耦合多新息广义增广随机梯度算法

借助于多新息辨识理论,可以得到部分耦合多新息广义增广随机梯度算法(Partially Coupled Multi-Innovation Generalized Extended Stochastic Gradient algorithm, PC-MI-GESG 算法):

$$\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_1(t)\\ \hat{\boldsymbol{\alpha}}_1(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_m(t-1)\\ \hat{\boldsymbol{\alpha}}_1(t-1)\end{bmatrix}+\frac{\boldsymbol{\Gamma}_1(p,t)}{r_1(t)}\boldsymbol{E}_1(p,t),$$

$$\hat{\boldsymbol{\vartheta}}_m(0)=\mathbf{1}_{n+n_c}/p_0,\quad \hat{\boldsymbol{\alpha}}_1(0)=\mathbf{1}_{mn_d}/p_0, \tag{165}$$

$$\boldsymbol{E}_1(p,t)=\boldsymbol{Y}_1(p,t)-\boldsymbol{\Gamma}_1^{\mathrm{T}}(p,t)\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_m(t-1)\\ \hat{\boldsymbol{\alpha}}_1(t-1)\end{bmatrix}, \tag{166}$$

$$\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\boldsymbol{\Gamma}_i(p,t)}{r_i(t)}\boldsymbol{E}_i(p,t),$$

$$\hat{\boldsymbol{\alpha}}_i(0)=\mathbf{1}_{mn_d}/p_0,\quad i=2,3,\cdots,m, \tag{167}$$

$$\boldsymbol{E}_i(p,t)=\boldsymbol{Y}_i(p,t)-\boldsymbol{\Gamma}_i^{\mathrm{T}}(p,t)\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}, \tag{168}$$

$$r_i(t)=r_i(t-1)+\|\hat{\boldsymbol{\phi}}_i(t)\|^2+\|\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\|^2,$$

$$r_i(0)=1,\quad i=1,2,\cdots,m, \tag{169}$$

$$\boldsymbol{\Gamma}_i(p,t)=[\hat{\boldsymbol{\psi}}_i(t),\hat{\boldsymbol{\psi}}_i(t-1),\cdots,\hat{\boldsymbol{\psi}}_i(t-p+1)], \tag{170}$$

$$\boldsymbol{Y}_i(p,t)=[y_i(t),y_i(t-1),\cdots,y_i(t-p+1)]^{\mathrm{T}}, \tag{171}$$

$$\hat{\boldsymbol{\psi}}_i(t)=[\hat{\boldsymbol{\phi}}_i^{\mathrm{T}}(t),\hat{\boldsymbol{\varphi}}_{\mathrm{n}}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{172}$$

$$\hat{\boldsymbol{\Phi}}(t)=[\hat{\boldsymbol{\phi}}_1(t),\hat{\boldsymbol{\phi}}_2(t),\cdots,\hat{\boldsymbol{\phi}}_m(t)]^{\mathrm{T}}=[\boldsymbol{\Phi}_{\mathrm{s}}(t),-\hat{\boldsymbol{w}}(t-1),-\hat{\boldsymbol{w}}(t-2),\cdots,-\hat{\boldsymbol{w}}(t-n_c)], \tag{173}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)=[\hat{\boldsymbol{v}}^{\mathrm{T}}(t-1),\hat{\boldsymbol{v}}^{\mathrm{T}}(t-2),\cdots,\hat{\boldsymbol{v}}^{\mathrm{T}}(t-n_d)]^{\mathrm{T}}, \tag{174}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}_m(t)=[\hat{\boldsymbol{\theta}}_m^{\mathrm{T}}(t),\hat{c}_{m1}(t),\hat{c}_{m2}(t),\cdots,\hat{c}_{mn_c}(t)]^{\mathrm{T}}, \tag{175}$$

$$\hat{\boldsymbol{w}}(t)=\boldsymbol{y}(t)-\boldsymbol{\Phi}_{\mathrm{s}}(t)\hat{\boldsymbol{\theta}}_m(t), \tag{176}$$

$$\hat{\boldsymbol{v}}(t)=\boldsymbol{y}(t)-\hat{\boldsymbol{\Phi}}(t)\hat{\boldsymbol{\vartheta}}(t)-\hat{\boldsymbol{\alpha}}^{\mathrm{T}}(t)\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t), \tag{177}$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\boldsymbol{\alpha}}_1(t),\hat{\boldsymbol{\alpha}}_2(t),\cdots,\hat{\boldsymbol{\alpha}}_m(t)]. \tag{178}$$

3 多元伪线性自回归滑动平均系统(2)

考虑多元伪线性自回归滑动平均系统

$$\boldsymbol{y}(t)=\boldsymbol{\Phi}_{\mathrm{s}}(t)\boldsymbol{\theta}+\boldsymbol{C}^{-1}(z)d(z)\boldsymbol{v}(t), \tag{179}$$

其中 $\boldsymbol{y}(t)\in\mathbf{R}^m$, $\boldsymbol{\Phi}_{\mathrm{s}}(t)\in\mathbf{R}^{m\times n}$, $\boldsymbol{\theta}\in\mathbf{R}^n$ 的定义同上,$\boldsymbol{w}(t):=\boldsymbol{C}^{-1}(z)d(z)\boldsymbol{v}(t)\in\mathbf{R}^m$ 为 m 维干扰噪声向量,$\boldsymbol{v}(t)\in\mathbf{R}^m$ 为 m 维白噪声向量,$\boldsymbol{C}(z)$ 和 $d(z)$ 为单位后移算子 z^{-1} 的多项式矩阵和多项式:

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

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

假设 m,n,n_c 和 n_d 已知,且当 $t\leqslant 0$ 时,$\boldsymbol{y}(t)=\mathbf{0}$, $\boldsymbol{\Phi}_{\mathrm{s}}(t)=\mathbf{0}$, $\boldsymbol{v}(t)=\mathbf{0}$.辨识的目标是基于耦合辨识概念

和多新息辨识理论,利用系统的观测数据$\{\boldsymbol{y}(t),\boldsymbol{\Phi}_s(t):t=1,2,3,\cdots\}$提出新的算法,估计系统模型参数向量$\boldsymbol{\theta}$,以及噪声模型的参数矩阵$\boldsymbol{C}_i$和参数$d_i$.

3.1 部分耦合(子系统)广义增广随机梯度辨识算法

定义参数向量$\boldsymbol{\vartheta}$,参数矩阵$\boldsymbol{\alpha}$,信息矩阵$\boldsymbol{\Phi}(t)$和信息向量$\boldsymbol{\varphi}_n(t)$如下:

$$\boldsymbol{\vartheta}:=[\boldsymbol{\theta}^T,d_1,d_2,\cdots,d_{n_d}]^T\in\mathbf{R}^{n+n_d},$$

$$\boldsymbol{\alpha}^T:=[\boldsymbol{C}_1,\boldsymbol{C}_2,\cdots,\boldsymbol{C}_{n_c}]=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m]^T\in\mathbf{R}^{m\times(mn_c)},$$

$$\boldsymbol{\Phi}(t):=[\boldsymbol{\Phi}_s(t),\boldsymbol{v}(t-1),\boldsymbol{v}(t-2),\cdots,\boldsymbol{v}(t-n_d)]=[\boldsymbol{\phi}_1(t),\boldsymbol{\phi}_2(t),\cdots,\boldsymbol{\phi}_m(t)]^T\in\mathbf{R}^{m\times(n+n_d)},$$

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

于是模型(179)可以表示为

$$\boldsymbol{y}(t)=\boldsymbol{\Phi}(t)\boldsymbol{\vartheta}+\boldsymbol{\alpha}^T\boldsymbol{\varphi}_n(t)+\boldsymbol{v}(t). \tag{180}$$

令第i个子系统信息向量表示为$\boldsymbol{\psi}_i(t):=\begin{bmatrix}\boldsymbol{\phi}_i(t)\\ \boldsymbol{\varphi}_n(t)\end{bmatrix}\in\mathbf{R}^{n+n_d+mn_c}$,那么第$i$个子系统可以写为

$$y_i(t)=\boldsymbol{\phi}_i^T(t)\boldsymbol{\vartheta}+\boldsymbol{\alpha}_i^T\boldsymbol{\varphi}_n(t)+v_i(t)=\boldsymbol{\psi}_i^T(t)\begin{bmatrix}\boldsymbol{\vartheta}\\ \boldsymbol{\alpha}_i\end{bmatrix}+v_i(t),\quad i=1,2,\cdots,m. \tag{181}$$

1) 子系统广义增广随机梯度辨识算法

通过类比,能够获得辨识多元伪线性自回归滑动平均系统(180)的子系统广义增广随机梯度算法(Subsystem Generalized Extended Stochastic Gradient algorithm,S-GESG 算法):

$$\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_i(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\hat{\boldsymbol{\psi}}_i(t)}{r_i(t)}\left\{y_i(t)-\hat{\boldsymbol{\psi}}_i^T(t)\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_i(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}\right\},$$

$$\hat{\boldsymbol{\vartheta}}_i(0)=\mathbf{1}_{n+n_d}/p_0,\quad \hat{\boldsymbol{\alpha}}_i(0)=\mathbf{1}_{mn_c}/p_0, \tag{182}$$

$$r_i(t)=r_i(t-1)+\|\hat{\boldsymbol{\psi}}_i(t)\|^2,\quad r_i(0)=1, \tag{183}$$

$$\hat{\boldsymbol{\psi}}_i(t)=[\hat{\boldsymbol{\phi}}_i^T(t),\hat{\boldsymbol{\varphi}}_n^T(t)]^T, \tag{184}$$

$$\hat{\boldsymbol{\Phi}}(t)=[\hat{\boldsymbol{\phi}}_1(t),\hat{\boldsymbol{\phi}}_2(t),\cdots,\hat{\boldsymbol{\phi}}_m(t)]^T=[\boldsymbol{\Phi}_s(t),\hat{\boldsymbol{v}}(t-1),\hat{\boldsymbol{v}}(t-2),\cdots,\hat{\boldsymbol{v}}(t-n_d)], \tag{185}$$

$$\hat{\boldsymbol{\varphi}}_n(t)=[-\hat{\boldsymbol{w}}^T(t-1),-\hat{\boldsymbol{w}}^T(t-2),\cdots,-\hat{\boldsymbol{w}}^T(t-n_c)]^T, \tag{186}$$

$$\hat{\boldsymbol{\vartheta}}_i(t)=[\hat{\boldsymbol{\theta}}_i^T(t),\hat{d}_{i1}(t),\hat{d}_{i2}(t),\cdots,\hat{d}_{in_d}(t)]^T, \tag{187}$$

$$\hat{\boldsymbol{w}}(t)=[\hat{w}_1(t),\hat{w}_2(t),\cdots,\hat{w}_m(t)]^T, \tag{188}$$

$$\hat{w}_i(t)=y_i(t)-[\boldsymbol{\Phi}_s(t)\text{的第 }i\text{ 行}]\hat{\boldsymbol{\theta}}_i(t), \tag{189}$$

$$\hat{\boldsymbol{v}}(t)=[\hat{v}_1(t),\hat{v}_2(t),\cdots,\hat{v}_m(t)]^T, \tag{190}$$

$$\hat{v}_i(t)=y_i(t)-\hat{\boldsymbol{\phi}}_i^T(t)\hat{\boldsymbol{\vartheta}}_i(t)-\hat{\boldsymbol{\alpha}}_i^T(t)\hat{\boldsymbol{\varphi}}_n(t). \tag{191}$$

2) 部分耦合子系统广义增广随机梯度辨识算法

在上述 S-GESG 算法中,用平均值$\hat{\boldsymbol{\vartheta}}(t-1)$代替$\hat{\boldsymbol{\vartheta}}_i(t-1)$,能够得到部分耦合子系统广义增广随机梯度算法(Partially Coupled Subsystem Generalized Extended Stochastic Gradient algorithm,PC-S-GESG 算法):

$$\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\vartheta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\hat{\boldsymbol{\psi}}_i(t)}{r_i(t)}\left\{y_i(t)-\hat{\boldsymbol{\psi}}_i^T(t)\begin{bmatrix}\hat{\boldsymbol{\vartheta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}\right\},$$

$$\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_{n+n_d}/p_0,\quad \hat{\boldsymbol{\alpha}}_i(0)=\mathbf{1}_{mn_c}/p_0, \tag{192}$$

$$r_i(t)=r_i(t-1)+\|\hat{\boldsymbol{\psi}}_i(t)\|^2,\quad r_i(0)=1, \tag{193}$$

$$\hat{\boldsymbol{\psi}}_i(t)=[\hat{\boldsymbol{\phi}}_i^T(t),\hat{\boldsymbol{\varphi}}_n^T(t)]^T, \tag{194}$$

$$\hat{\boldsymbol{\Phi}}(t)=[\hat{\boldsymbol{\phi}}_1(t),\hat{\boldsymbol{\phi}}_2(t),\cdots,\hat{\boldsymbol{\phi}}_m(t)]^T=[\boldsymbol{\Phi}_s(t),\hat{\boldsymbol{v}}(t-1),\hat{\boldsymbol{v}}(t-2),\cdots,\hat{\boldsymbol{v}}(t-n_d)], \tag{195}$$

$$\hat{\boldsymbol{\varphi}}_n(t)=[-\hat{\boldsymbol{w}}^T(t-1),-\hat{\boldsymbol{w}}^T(t-2),\cdots,-\hat{\boldsymbol{w}}^T(t-n_c)]^T, \tag{196}$$

$$\hat{\boldsymbol{\vartheta}}_i(t)=[\hat{\boldsymbol{\theta}}_i^T(t),\hat{d}_{i1}(t),\hat{d}_{i2}(t),\cdots,\hat{d}_{in_d}(t)]^T, \tag{197}$$

$$\hat{\boldsymbol{w}}(t)=\boldsymbol{y}(t)-\boldsymbol{\Phi}_s(t)\hat{\boldsymbol{\theta}}(t), \tag{198}$$

$$\hat{\boldsymbol{v}}(t)=\boldsymbol{y}(t)-\hat{\boldsymbol{\Phi}}(t)\hat{\boldsymbol{\vartheta}}(t)-\hat{\boldsymbol{\alpha}}^T(t)\hat{\boldsymbol{\varphi}}_n(t), \tag{199}$$

$$\hat{\boldsymbol{\theta}}(t)=\frac{\hat{\boldsymbol{\theta}}_1(t)+\hat{\boldsymbol{\theta}}_2(t)+\cdots+\hat{\boldsymbol{\theta}}_m(t)}{m}, \tag{200}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\frac{\hat{\boldsymbol{\vartheta}}_1(t)+\hat{\boldsymbol{\vartheta}}_2(t)+\cdots+\hat{\boldsymbol{\vartheta}}_m(t)}{m}, \tag{201}$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\boldsymbol{\alpha}}_1(t),\hat{\boldsymbol{\alpha}}_2(t),\cdots,\hat{\boldsymbol{\alpha}}_m(t)]. \tag{202}$$

3) 部分耦合广义增广随机梯度辨识算法

利用耦合辨识概念,能够获得辨识多元伪线性自回归滑动平均系统(179)的部分耦合广义增广随机梯度算法(Partially Coupled Generalized Extended Stochastic Gradient algorithm,PC-GESG 算法):

$$\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_1(t)\\ \hat{\boldsymbol{\alpha}}_1(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_m(t-1)\\ \hat{\boldsymbol{\alpha}}_1(t-1)\end{bmatrix}+\frac{\hat{\boldsymbol{\psi}}_1(t)}{r_1(t)}\left\{y_1(t)-\hat{\boldsymbol{\psi}}_1^T(t)\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_m(t-1)\\ \hat{\boldsymbol{\alpha}}_1(t-1)\end{bmatrix}\right\},$$

$$\hat{\boldsymbol{\vartheta}}_m(0)=\mathbf{1}_{n+n_d}/p_0,\quad \hat{\boldsymbol{\alpha}}_1(0)=\mathbf{1}_{mn_c}/p_0, \tag{203}$$

$$r_1(t)=r_m(t-1)+\|\hat{\boldsymbol{\psi}}_1(t)\|^2,\quad r_1(0)=1, \tag{204}$$

$$\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\hat{\boldsymbol{\psi}}_i(t)}{r_i(t)}\left\{y_i(t)-\hat{\boldsymbol{\psi}}_i^T(t)\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}\right\},$$

$$\hat{\boldsymbol{\alpha}}_i(0)=\mathbf{1}_{mn_c}/p_0,\quad i=2,3,\cdots,m, \tag{205}$$

$$r_i(t)=r_{i-1}(t)+\|\hat{\boldsymbol{\psi}}_i(t)\|^2,\quad r_i(0)=1, \tag{206}$$

$$\hat{\boldsymbol{\psi}}_i(t)=[\hat{\boldsymbol{\phi}}_i^T(t),\hat{\boldsymbol{\varphi}}_n^T(t)]^T,\quad i=1,2,\cdots,m, \tag{207}$$

$$\hat{\boldsymbol{\Phi}}(t)=[\hat{\boldsymbol{\phi}}_1(t),\hat{\boldsymbol{\phi}}_2(t),\cdots,\hat{\boldsymbol{\phi}}_m(t)]^T=[\boldsymbol{\Phi}_s(t),\hat{\boldsymbol{v}}(t-1),\hat{\boldsymbol{v}}(t-2),\cdots,\hat{\boldsymbol{v}}(t-n_d)], \tag{208}$$

$$\hat{\boldsymbol{\varphi}}_n(t)=[-\hat{\boldsymbol{w}}^T(t-1),-\hat{\boldsymbol{w}}^T(t-2),\cdots,-\hat{\boldsymbol{w}}^T(t-n_c)]^T, \tag{209}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}_m(t)=[\hat{\boldsymbol{\theta}}_m^T(t),\hat{d}_{m1}(t),\hat{d}_{m2}(t),\cdots,\hat{d}_{mn_d}(t)]^T, \tag{210}$$

$$\hat{\boldsymbol{w}}(t)=\boldsymbol{y}(t)-\boldsymbol{\Phi}_{\mathrm{s}}(t)\hat{\boldsymbol{\theta}}_m(t), \tag{211}$$

$$\hat{\boldsymbol{v}}(t)=\boldsymbol{y}(t)-\hat{\boldsymbol{\Phi}}(t)\hat{\boldsymbol{\vartheta}}(t)-\hat{\boldsymbol{\alpha}}^{\mathrm{T}}(t)\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t), \tag{212}$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\boldsymbol{\alpha}}_1(t),\hat{\boldsymbol{\alpha}}_2(t),\cdots,\hat{\boldsymbol{\alpha}}_m(t)]. \tag{213}$$

也可去掉式(203)和(206)中 $r_i(t)$ 间的耦合,将其更改为

$$r_i(t)=r_i(t)+\|\hat{\boldsymbol{\psi}}_i(t)\|^2,\quad r_i(0)=1,\quad i=1,2,\cdots,m. \tag{214}$$

3.2 部分耦合(子系统)多新息广义增广随机梯度算法

1)部分耦合子系统多新息广义增广随机梯度算法

对于多元伪线性自回归滑动平均系统(179)对应的辨识模型(181),利用多新息辨识理论,能够得到部分耦合子系统多新息广义增广随机梯度算法(Partially Coupled Subsystem Multi-Innovation Generalized Extended Stochastic Gradient algorithm, PC-S-MI-GESG 算法):

$$\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\vartheta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\boldsymbol{\Gamma}_i(p,t)}{r_i(t)}\boldsymbol{E}_i(p,t),$$

$$\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_{n+n_d}/p_0,\quad \hat{\boldsymbol{\alpha}}_i(0)=\mathbf{1}_{mn_c}/p_0, \tag{215}$$

$$\boldsymbol{E}_i(p,t)=\boldsymbol{Y}_i(p,t)-\boldsymbol{\Gamma}_i^{\mathrm{T}}(p,t)\begin{bmatrix}\hat{\boldsymbol{\vartheta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}, \tag{216}$$

$$r_i(t)=r_i(t-1)+\|\hat{\boldsymbol{\phi}}_i(t)\|^2+\|\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\|^2,\quad r_i(0)=1, \tag{217}$$

$$\boldsymbol{\Gamma}_i(p,t)=[\hat{\boldsymbol{\psi}}_i(t),\hat{\boldsymbol{\psi}}_i(t-1),\cdots,\hat{\boldsymbol{\psi}}_i(t-p+1)], \tag{218}$$

$$\boldsymbol{Y}_i(p,t)=[y_i(t),y_i(t-1),\cdots,y_i(t-p+1)]^{\mathrm{T}}, \tag{219}$$

$$\hat{\boldsymbol{\psi}}_i(t)=[\hat{\boldsymbol{\phi}}_i^{\mathrm{T}}(t),\hat{\boldsymbol{\varphi}}_{\mathrm{n}}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{220}$$

$$\hat{\boldsymbol{\Phi}}(t)=[\hat{\boldsymbol{\phi}}_1(t),\hat{\boldsymbol{\phi}}_2(t),\cdots,\hat{\boldsymbol{\phi}}_m(t)]^{\mathrm{T}}=$$
$$[\boldsymbol{\Phi}_{\mathrm{s}}(t),\hat{\boldsymbol{v}}(t-1),\hat{\boldsymbol{v}}(t-2),\cdots,\hat{\boldsymbol{v}}(t-n_d)], \tag{221}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)=[-\hat{\boldsymbol{w}}^{\mathrm{T}}(t-1),-\hat{\boldsymbol{w}}^{\mathrm{T}}(t-2),\cdots,-\hat{\boldsymbol{w}}^{\mathrm{T}}(t-n_c)]^{\mathrm{T}}, \tag{222}$$

$$\hat{\boldsymbol{\vartheta}}(t)=[\hat{\boldsymbol{\theta}}^{\mathrm{T}}(t),\hat{d}_1(t),\hat{d}_2(t),\cdots,\hat{d}_{n_d}(t)]^{\mathrm{T}}, \tag{223}$$

$$\hat{\boldsymbol{w}}(t)=\boldsymbol{y}(t)-\boldsymbol{\Phi}_{\mathrm{s}}(t)\hat{\boldsymbol{\theta}}(t), \tag{224}$$

$$\hat{\boldsymbol{v}}(t)=\boldsymbol{y}(t)-\hat{\boldsymbol{\Phi}}(t)\hat{\boldsymbol{\vartheta}}(t)-\hat{\boldsymbol{\alpha}}^{\mathrm{T}}(t)\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t), \tag{225}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\frac{\hat{\boldsymbol{\vartheta}}_1(t)+\hat{\boldsymbol{\vartheta}}_2(t)+\cdots+\hat{\boldsymbol{\vartheta}}_m(t)}{m}, \tag{226}$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\boldsymbol{\alpha}}_1(t),\hat{\boldsymbol{\alpha}}_2(t),\cdots,\hat{\boldsymbol{\alpha}}_m(t)]. \tag{227}$$

2)部分耦合多新息广义增广随机梯度算法

进一步可以得到部分耦合多新息广义增广随机梯度算法(Partially Coupled Multi-Innovation Generalized Extended Stochastic Gradient algorithm, PC-MI-GESG 算法):

$$\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_1(t)\\ \hat{\boldsymbol{\alpha}}_1(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_m(t-1)\\ \hat{\boldsymbol{\alpha}}_1(t-1)\end{bmatrix}+\frac{\boldsymbol{\Gamma}_1(p,t)}{r_1(t)}\boldsymbol{E}_1(p,t),$$

$$\hat{\boldsymbol{\vartheta}}_m(0)=\mathbf{1}_{n+n_d}/p_0,\quad \hat{\boldsymbol{\alpha}}_1(0)=\mathbf{1}_{mn_c}/p_0, \tag{228}$$

$$\boldsymbol{E}_1(p,t)=\boldsymbol{Y}_1(p,t)-\boldsymbol{\Gamma}_1^{\mathrm{T}}(p,t)\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_m(t-1)\\ \hat{\boldsymbol{\alpha}}_1(t-1)\end{bmatrix}, \tag{229}$$

$$\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\boldsymbol{\Gamma}_i(p,t)}{r_i(t)}\boldsymbol{E}_i(p,t),$$

$$\hat{\boldsymbol{\alpha}}_i(0)=\mathbf{1}_{mn_c}/p_0,\quad i=2,3,\cdots,m, \tag{230}$$

$$\boldsymbol{E}_i(p,t)=\boldsymbol{Y}_i(p,t)-\boldsymbol{\Gamma}_i^{\mathrm{T}}(p,t)\begin{bmatrix}\hat{\boldsymbol{\vartheta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}, \tag{231}$$

$$r_i(t)=r_i(t-1)+\|\hat{\boldsymbol{\phi}}_i(t)\|^2+\|\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\|^2,$$
$$r_i(0)=1,\quad i=1,2,\cdots,m, \tag{232}$$

$$\boldsymbol{\Gamma}_i(p,t)=[\hat{\boldsymbol{\psi}}_i(t),\hat{\boldsymbol{\psi}}_i(t-1),\cdots,\hat{\boldsymbol{\psi}}_i(t-p+1)], \tag{233}$$

$$\boldsymbol{Y}_i(p,t)=[y_i(t),y_i(t-1),\cdots,y_i(t-p+1)]^{\mathrm{T}}, \tag{234}$$

$$\hat{\boldsymbol{\psi}}_i(t)=[\hat{\boldsymbol{\varphi}}_i^{\mathrm{T}}(t),\hat{\boldsymbol{\varphi}}_{\mathrm{n}}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{235}$$

$$\hat{\boldsymbol{\Phi}}(t)=[\hat{\boldsymbol{\phi}}_1(t),\hat{\boldsymbol{\phi}}_2(t),\cdots,\hat{\boldsymbol{\phi}}_m(t)]^{\mathrm{T}}=$$
$$[\boldsymbol{\Phi}_{\mathrm{s}}(t),\hat{\boldsymbol{v}}(t-1),\hat{\boldsymbol{v}}(t-2),\cdots,\hat{\boldsymbol{v}}(t-n_d)], \tag{236}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)=[-\hat{\boldsymbol{w}}^{\mathrm{T}}(t-1),-\hat{\boldsymbol{w}}^{\mathrm{T}}(t-2),\cdots,-\hat{\boldsymbol{w}}^{\mathrm{T}}(t-n_c)]^{\mathrm{T}}, \tag{237}$$

$$\hat{\boldsymbol{\vartheta}}(t)=\hat{\boldsymbol{\vartheta}}_m(t)=[\hat{\boldsymbol{\theta}}_m^{\mathrm{T}}(t),\hat{d}_{m1}(t),\hat{d}_{m2}(t),\cdots,\hat{d}_{mn_d}(t)]^{\mathrm{T}}, \tag{238}$$

$$\hat{\boldsymbol{w}}(t)=\boldsymbol{y}(t)-\boldsymbol{\Phi}_{\mathrm{s}}(t)\hat{\boldsymbol{\theta}}_m(t), \tag{239}$$

$$\hat{\boldsymbol{v}}(t)=\boldsymbol{y}(t)-\hat{\boldsymbol{\Phi}}(t)\hat{\boldsymbol{\vartheta}}(t)-\hat{\boldsymbol{\alpha}}^{\mathrm{T}}(t)\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t), \tag{240}$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\boldsymbol{\alpha}}_1(t),\hat{\boldsymbol{\alpha}}_2(t),\cdots,\hat{\boldsymbol{\alpha}}_m(t)]. \tag{241}$$

4 多元伪线性自回归滑动平均系统(3)

考虑多元伪线性自回归滑动平均系统[1,2,18]

$$\boldsymbol{y}(t)=\boldsymbol{\Phi}_{\mathrm{s}}(t)\boldsymbol{\theta}+\boldsymbol{C}^{-1}(z)\boldsymbol{D}(z)\boldsymbol{v}(t), \tag{242}$$

其中 $\boldsymbol{y}(t)\in\mathbf{R}^m$, $\boldsymbol{\Phi}_{\mathrm{s}}(t)\in\mathbf{R}^{m\times n}$, $\boldsymbol{\theta}\in\mathbf{R}^n$ 定义同上, $\boldsymbol{w}(t):=\boldsymbol{C}^{-1}(z)\boldsymbol{D}(z)\boldsymbol{v}(t)\in\mathbf{R}^m$ 为 m 维干扰噪声向量, $\boldsymbol{v}(t)\in\mathbf{R}^m$ 为 m 维白噪声向量, $\boldsymbol{C}(z)$ 和 $\boldsymbol{D}(z)$ 均为单位后移算子 z^{-1} 的多项式矩阵:

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

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

假设 m,n,n_c 和 n_d 已知,且当 $t\leqslant 0$ 时, $\boldsymbol{y}(t)=0$, $\boldsymbol{\Phi}_{\mathrm{s}}(t)=\mathbf{0}$, $\boldsymbol{v}(t)=\mathbf{0}$.辨识的目标是基于耦合辨识概念和多新息辨识理论,利用系统的观测数据 $\{\boldsymbol{y}(t),\boldsymbol{\Phi}_{\mathrm{s}}(t):t=1,2,3,\cdots\}$ 提出新的算法,估计系统模型参数向量 $\boldsymbol{\theta}$,以及噪声模型的参数矩阵 $\boldsymbol{C}_i$ 和 $\boldsymbol{D}_i$.

4.1 多元广义增广随机梯度辨识算法

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

$$\boldsymbol{\theta}:=[\theta_1,\theta_2,\cdots,\theta_n]^{\mathrm{T}}\in\mathbf{R}^n,$$

$$\boldsymbol{\alpha}^{\mathrm{T}}:=[\boldsymbol{C}_1,\boldsymbol{C}_2,\cdots,\boldsymbol{C}_{n_c},\boldsymbol{D}_1,\boldsymbol{D}_2,\cdots,\boldsymbol{D}_{n_d}]=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m]^{\mathrm{T}}\in\mathbf{R}^{m\times(mn_c+mn_d)},$$

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

于是系统(242)可以表示为

$$\boldsymbol{y}(t)=\boldsymbol{\Phi}_{\mathrm{s}}(t)\boldsymbol{\theta}+\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\phi}_{\mathrm{n}}(t)+\boldsymbol{v}(t).\tag{243}$$

定义信息矩阵 $\boldsymbol{\Phi}(t)$ 和参数向量 $\boldsymbol{\vartheta}$ 如下：

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

$$n_0:=n+m^2(n_c+n_d),$$

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

则模型(243)可以写为下列多元伪线性回归模型(multivariate pseudo-linear regressive model)

$$\boldsymbol{y}(t)=\boldsymbol{\Phi}(t)\boldsymbol{\vartheta}+\boldsymbol{v}(t).\tag{244}$$

辨识的困难是 $\boldsymbol{\Phi}(t)$ 中包含了未知相关噪声项 $\boldsymbol{w}(t-i)$ 和未知噪声回归项 $\boldsymbol{v}(t-i)$.分别用其估计值 $\hat{\boldsymbol{w}}(t-i)$ 和 $\hat{\boldsymbol{v}}(t-i)$ 代替,代替后的信息矩阵 $\boldsymbol{\Phi}(t)$ 记作 $\boldsymbol{\Psi}(t)$,于是能够获得估计多元伪线性回归辨识模型(244)参数向量 $\boldsymbol{\vartheta}$ 的多元广义增广随机梯度算法(Multivariate Generalized Extended Stochastic Gradient algorithm,M-GESG 算法)：

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

$$\hat{\boldsymbol{\vartheta}}(0)=\mathbf{1}_{n_0}/p_0,\tag{245}$$

$$r(t)=r(t-1)+\|\boldsymbol{\Psi}(t)\|^2,\quad r(0)=1,\tag{246}$$

$$\boldsymbol{\Psi}(t)=[\boldsymbol{\Phi}_{\mathrm{s}}(t),\hat{\boldsymbol{\varphi}}_{\mathrm{n}}^{\mathrm{T}}(t)\otimes\boldsymbol{I}_m],\tag{247}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)=[-\hat{\boldsymbol{w}}^{\mathrm{T}}(t-1),-\hat{\boldsymbol{w}}^{\mathrm{T}}(t-2),\cdots,-\hat{\boldsymbol{w}}^{\mathrm{T}}(t-n_c),\hat{\boldsymbol{v}}^{\mathrm{T}}(t-1),\hat{\boldsymbol{v}}^{\mathrm{T}}(t-2),\cdots,\hat{\boldsymbol{v}}^{\mathrm{T}}(t-n_d)]^{\mathrm{T}},\tag{248}$$

$$\hat{\boldsymbol{w}}(t)=\boldsymbol{y}(t)-\boldsymbol{\Phi}_{\mathrm{s}}(t)\hat{\boldsymbol{\theta}}(t),\quad\hat{\boldsymbol{w}}(-i)=\mathbf{1}_m/p_0,\quad i=0,1,\cdots,n_c,\tag{249}$$

$$\hat{\boldsymbol{v}}(t)=\boldsymbol{y}(t)-\boldsymbol{\Psi}(t)\hat{\boldsymbol{\vartheta}}(t),\hat{\boldsymbol{v}}(-i)=\mathbf{1}_m/p_0,i=0,1,\cdots,n_d.\tag{250}$$

4.2 部分耦合(子系统)广义增广随机梯度辨识算法

尽管算法(245)—(250)能够估计出参数向量 $\hat{\boldsymbol{\vartheta}}(t)$,但是算法中 $\boldsymbol{\Psi}(t)$ 含有大量的零元,导致算法计算量大.为了降低计算量,下面给出多元伪线性自回归滑动平均系统(242)的部分耦合辨识算法.

令 $\boldsymbol{\phi}_i^{\mathrm{T}}(t)\in\mathbf{R}^{1\times n}$ 是 $\boldsymbol{\Phi}_{\mathrm{s}}(t)$ 的第 i 行,$\boldsymbol{\alpha}_i\in\mathbf{R}^{mn_c+mn_d}$ 是参数矩阵 $\boldsymbol{\alpha}$ 的第 i 列,$v_i(t)\in\mathbf{R}$ 是噪声向量 $\boldsymbol{v}(t)$ 的第 i 元,子系统信息向量表示为 $\boldsymbol{\psi}_i(t):=\begin{bmatrix}\boldsymbol{\phi}_i(t)\\ \boldsymbol{\varphi}_{\mathrm{n}}(t)\end{bmatrix}\in\mathbf{R}^{n+mn_c+mn_d}$.参考多元伪线性滑动平均系统的子系统模型(11)的推导,第 i 个子系统可以写为

$$y_i(t)=\boldsymbol{\phi}_i^{\mathrm{T}}(t)\boldsymbol{\theta}+\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{\varphi}_{\mathrm{n}}(t)+v_i(t)=\boldsymbol{\psi}_i^{\mathrm{T}}(t)\begin{bmatrix}\boldsymbol{\theta}\\ \boldsymbol{\alpha}_i\end{bmatrix}+v_i(t),\quad i=1,2,\cdots,m.\tag{251}$$

由此能够得到辨识多元伪线性自回归滑动平均系统(242)的子系统广义增广随机梯度算法(Subsystem Generalized Extended Stochastic Gradient algorithm,S-GESG 算法)：

$$\begin{bmatrix}\hat{\boldsymbol{\theta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\theta}}_i(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\hat{\boldsymbol{\psi}}_i(t)}{r_i(t)}\left\{y_i(t)-\hat{\boldsymbol{\psi}}_i^{\mathrm{T}}(t)\begin{bmatrix}\hat{\boldsymbol{\theta}}_i(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}\right\},$$

$$\hat{\boldsymbol{\theta}}_i(0)=\mathbf{1}_n/p_0,\quad\hat{\boldsymbol{\alpha}}_i(0)=\mathbf{1}_{mn_c+mn_d}/p_0,\tag{252}$$

$$r_i(t)=r_i(t-1)+\|\hat{\boldsymbol{\psi}}_i(t)\|^2,\quad r_i(0)=1,\tag{253}$$

$$\hat{\boldsymbol{\psi}}_i(t)=\begin{bmatrix}\boldsymbol{\phi}_i(t)\\ \hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\end{bmatrix},\tag{254}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)=[-\hat{\boldsymbol{w}}^{\mathrm{T}}(t-1),-\hat{\boldsymbol{w}}^{\mathrm{T}}(t-2),\cdots,-\hat{\boldsymbol{w}}^{\mathrm{T}}(t-n_c),\hat{\boldsymbol{v}}^{\mathrm{T}}(t-1),\hat{\boldsymbol{v}}^{\mathrm{T}}(t-2),\cdots,\hat{\boldsymbol{v}}^{\mathrm{T}}(t-n_d)]^{\mathrm{T}},\tag{255}$$

$$\boldsymbol{\Phi}_{\mathrm{s}}(t)=[\boldsymbol{\phi}_1(t),\boldsymbol{\phi}_2(t),\cdots,\boldsymbol{\phi}_m(t)]^{\mathrm{T}},\tag{256}$$

$$\hat{\boldsymbol{w}}(t)=[\hat{w}_1(t),\hat{w}_2(t),\cdots,\hat{w}_m(t)]^{\mathrm{T}},\tag{257}$$

$$\hat{w}_i(t)=y_i(t)-\boldsymbol{\phi}_i^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_i(t),\tag{258}$$

$$\hat{\boldsymbol{v}}(t)=[\hat{v}_1(t),\hat{v}_2(t),\cdots,\hat{v}_m(t)]^{\mathrm{T}},\tag{259}$$

$$\hat{v}_i(t)=y_i(t)-\boldsymbol{\phi}_i^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_i(t)-\hat{\boldsymbol{\varphi}}_{\mathrm{n}}^{\mathrm{T}}(t)\hat{\boldsymbol{\alpha}}_i(t).\tag{260}$$

式(252)也可以分解为

$$\hat{\boldsymbol{\theta}}_i(t)=\hat{\boldsymbol{\theta}}_i(t-1)+\frac{\boldsymbol{\phi}_i(t)}{r_i(t)}[y_i(t)-\boldsymbol{\phi}_i^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_i(t-1)-\hat{\boldsymbol{\varphi}}_{\mathrm{n}}^{\mathrm{T}}(t)\hat{\boldsymbol{\alpha}}_i(t-1)],\tag{261}$$

$$\hat{\boldsymbol{\alpha}}_i(t)=\hat{\boldsymbol{\alpha}}_i(t-1)+\frac{\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)}{r_i(t)}[y_i(t)-\boldsymbol{\phi}_i^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_i(t-1)-\hat{\boldsymbol{\varphi}}_{\mathrm{n}}^{\mathrm{T}}(t)\hat{\boldsymbol{\alpha}}_i(t-1)].\tag{262}$$

1) 部分耦合子系统广义增广随机梯度辨识算法

根据耦合辨识概念,用平均值 $\hat{\boldsymbol{\theta}}(t-1)$ 代替 $\hat{\boldsymbol{\theta}}_i(t-1)$,可以得到部分耦合子系统广义增广随机梯度算法(Partially Coupled Subsystem Generalized Extended Stochastic Gradient algorithm,PC-S-GESG 算法)：

$$\hat{\boldsymbol{\theta}}_i(t)=\hat{\boldsymbol{\theta}}(t-1)+\frac{\boldsymbol{\phi}_i(t)}{r_i(t)}[y_i(t)-\boldsymbol{\phi}_i^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)-\hat{\boldsymbol{\varphi}}_{\mathrm{n}}^{\mathrm{T}}(t)\hat{\boldsymbol{\alpha}}_i(t-1)],$$

$$\hat{\boldsymbol{\theta}}(0)=\mathbf{1}_n/p_0,\tag{263}$$

$$\hat{\boldsymbol{\alpha}}_i(t)=\hat{\boldsymbol{\alpha}}_i(t-1)+\frac{\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)}{r_i(t)}[y_i(t)-\boldsymbol{\phi}_i^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)-\hat{\boldsymbol{\varphi}}_{\mathrm{n}}^{\mathrm{T}}(t)\hat{\boldsymbol{\alpha}}_i(t-1)],$$

$$\hat{\boldsymbol{\alpha}}_i(0)=\mathbf{1}_{mn_c+mn_d}/p_0,\tag{264}$$

$$r_i(t)=r_i(t-1)+\|\hat{\boldsymbol{\psi}}_i(t)\|^2,\quad r_i(0)=1, \tag{265}$$

$$\hat{\boldsymbol{\psi}}_i(t)=\begin{bmatrix}\boldsymbol{\phi}_i(t)\\ \hat{\boldsymbol{\varphi}}_{\rm n}(t)\end{bmatrix}, \tag{266}$$

$$\hat{\boldsymbol{\varphi}}_{\rm n}(t)=[-\hat{\boldsymbol{w}}^{\rm T}(t-1),-\hat{\boldsymbol{w}}^{\rm T}(t-2),\cdots,-\hat{\boldsymbol{w}}^{\rm T}(t-n_c),\ \hat{\boldsymbol{v}}^{\rm T}(t-1),\hat{\boldsymbol{v}}^{\rm T}(t-2),\cdots,\hat{\boldsymbol{v}}^{\rm T}(t-n_d)]^{\rm T}, \tag{267}$$

$$\boldsymbol{\Phi}_{\rm s}(t)=[\boldsymbol{\phi}_1(t),\boldsymbol{\phi}_2(t),\cdots,\boldsymbol{\phi}_m(t)]^{\rm T}, \tag{268}$$

$$\hat{\boldsymbol{w}}(t)=\boldsymbol{y}(t)-\boldsymbol{\Phi}_{\rm s}(t)\hat{\boldsymbol{\theta}}(t), \tag{269}$$

$$\hat{\boldsymbol{v}}(t)=\boldsymbol{y}(t)-\boldsymbol{\Phi}_{\rm s}(t)\hat{\boldsymbol{\theta}}(t)-\hat{\boldsymbol{\alpha}}^{\rm T}(t)\hat{\boldsymbol{\varphi}}_{\rm n}(t), \tag{270}$$

$$\hat{\boldsymbol{\theta}}(t)=\frac{\hat{\boldsymbol{\theta}}_1(t)+\hat{\boldsymbol{\theta}}_2(t)+\cdots+\hat{\boldsymbol{\theta}}_m(t)}{m}, \tag{271}$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\boldsymbol{\alpha}}_1(t),\hat{\boldsymbol{\alpha}}_2(t),\cdots,\hat{\boldsymbol{\alpha}}_m(t)]. \tag{272}$$

2）部分耦合广义增广随机梯度辨识算法

利用耦合辨识概念,基于 PC-S-GESG 算法,能够得到部分耦合广义增广随机梯度算法(Partially Coupled Generalized Extended Stochastic Gradient algorithm,PC-GESG 算法):

$$\begin{bmatrix}\hat{\boldsymbol{\theta}}_1(t)\\ \hat{\boldsymbol{\alpha}}_1(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\theta}}_m(t-1)\\ \hat{\boldsymbol{\alpha}}_1(t-1)\end{bmatrix}+\frac{\hat{\boldsymbol{\psi}}_1(t)}{r_1(t)}\left\{y_1(t)-\hat{\boldsymbol{\psi}}_1^{\rm T}(t)\begin{bmatrix}\hat{\boldsymbol{\theta}}_m(t-1)\\ \hat{\boldsymbol{\alpha}}_1(t-1)\end{bmatrix}\right\},$$
$$\hat{\boldsymbol{\theta}}_m(0)=\mathbf{1}_n/p_0,\quad \hat{\boldsymbol{\alpha}}_1(0)=\mathbf{1}_{mn_c+mn_d}/p_0, \tag{273}$$

$$\begin{bmatrix}\hat{\boldsymbol{\theta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\theta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\hat{\boldsymbol{\psi}}_i(t)}{r_i(t)}\left\{y_i(t)-\hat{\boldsymbol{\psi}}_i^{\rm T}(t)\begin{bmatrix}\hat{\boldsymbol{\theta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}\right\},$$
$$\hat{\boldsymbol{\alpha}}_i(0)=\mathbf{1}_{mn_c+mn_d}/p_0,\quad i=2,3,\cdots,m, \tag{274}$$

$$r_i(t)=r_i(t)+\|\hat{\boldsymbol{\psi}}_i(t)\|^2,r_i(0)=1,i=1,2,\cdots,m, \tag{275}$$

$$\hat{\boldsymbol{\psi}}_i(t)=\begin{bmatrix}\boldsymbol{\phi}_i(t)\\ \hat{\boldsymbol{\varphi}}_{\rm n}(t)\end{bmatrix}, \tag{276}$$

$$\hat{\boldsymbol{\varphi}}_{\rm n}(t)=[-\hat{\boldsymbol{w}}^{\rm T}(t-1),-\hat{\boldsymbol{w}}^{\rm T}(t-2),\cdots,-\hat{\boldsymbol{w}}^{\rm T}(t-n_c),\ \hat{\boldsymbol{v}}^{\rm T}(t-1),\hat{\boldsymbol{v}}^{\rm T}(t-2),\cdots,\hat{\boldsymbol{v}}^{\rm T}(t-n_d)]^{\rm T}, \tag{277}$$

$$\boldsymbol{\Phi}_{\rm s}(t)=[\boldsymbol{\phi}_1(t),\boldsymbol{\phi}_2(t),\cdots,\boldsymbol{\phi}_m(t)]^{\rm T}, \tag{278}$$

$$\hat{\boldsymbol{w}}(t)=\boldsymbol{y}(t)-\boldsymbol{\Phi}_{\rm s}(t)\hat{\boldsymbol{\theta}}_m(t), \tag{279}$$

$$\hat{\boldsymbol{v}}(t)=\boldsymbol{y}(t)-\boldsymbol{\Phi}_{\rm s}(t)\hat{\boldsymbol{\theta}}_i(t)-\hat{\boldsymbol{\alpha}}^{\rm T}(t)\hat{\boldsymbol{\varphi}}_{\rm n}(t), \tag{280}$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\boldsymbol{\alpha}}_1(t),\hat{\boldsymbol{\alpha}}_2(t),\cdots,\hat{\boldsymbol{\alpha}}_m(t)]. \tag{281}$$

上述算法中去掉了 $r_i(t)$ 间的耦合,输出的参数估计为 $\hat{\boldsymbol{\theta}}(t):=\hat{\boldsymbol{\theta}}_m(t)$ 和 $\hat{\boldsymbol{\alpha}}(t)$.

4.3 部分耦合(子系统)多新息广义增广随机梯度辨识算法

1）部分耦合子系统多新息广义增广随机梯度算法

对于多元伪线性自回归滑动平均系统(242)对应的辨识模型(243),参照部分耦合随机梯度算法(71)—(81),可以得到部分耦合子系统多新息广义增广随机梯度算法(Partially Coupled Subsystem Multi-Innovation Generalized Extended Stochastic Gradient algorithm,PC-S-MI-GESG 算法):

$$\begin{bmatrix}\hat{\boldsymbol{\theta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\theta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\boldsymbol{\Gamma}_i(p,t)}{r_i(t)}\boldsymbol{E}_i(p,t),$$
$$\hat{\boldsymbol{\theta}}(0)=\mathbf{1}_n/p_0,\quad \hat{\boldsymbol{\alpha}}_i(0)=\mathbf{1}_{mn_c+mn_d}/p_0, \tag{282}$$

$$\boldsymbol{E}_i(p,t)=\boldsymbol{Y}_i(p,t)-\boldsymbol{\Gamma}_i^{\rm T}(p,t)\begin{bmatrix}\hat{\boldsymbol{\theta}}(t-1)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}, \tag{283}$$

$$r_i(t)=r_i(t-1)+\|\boldsymbol{\phi}_i(t)\|^2+\|\hat{\boldsymbol{\varphi}}_{\rm n}(t)\|^2,\quad r_i(0)=1, \tag{284}$$

$$\boldsymbol{\Gamma}_i(p,t)=[\hat{\boldsymbol{\psi}}_i(t),\hat{\boldsymbol{\psi}}_i(t-1),\cdots,\hat{\boldsymbol{\psi}}_i(t-p+1)], \tag{285}$$

$$\boldsymbol{Y}_i(p,t)=[y_i(t),y_i(t-1),\cdots,y_i(t-p+1)]^{\rm T}, \tag{286}$$

$$\hat{\boldsymbol{\psi}}_i(t)=\begin{bmatrix}\boldsymbol{\phi}_i(t)\\ \hat{\boldsymbol{\varphi}}_{\rm n}(t)\end{bmatrix}, \tag{287}$$

$$\hat{\boldsymbol{\varphi}}_{\rm n}(t)=[-\hat{\boldsymbol{w}}^{\rm T}(t-1),-\hat{\boldsymbol{w}}^{\rm T}(t-2),\cdots,-\hat{\boldsymbol{w}}^{\rm T}(t-n_c),\ \hat{\boldsymbol{v}}^{\rm T}(t-1),\hat{\boldsymbol{v}}^{\rm T}(t-2),\cdots,\hat{\boldsymbol{v}}^{\rm T}(t-n_d)]^{\rm T}, \tag{288}$$

$$\boldsymbol{\Phi}_{\rm s}(t)=[\boldsymbol{\phi}_1(t),\boldsymbol{\phi}_2(t),\cdots,\boldsymbol{\phi}_m(t)]^{\rm T}, \tag{289}$$

$$\hat{\boldsymbol{w}}(t)=\boldsymbol{y}(t)-\boldsymbol{\Phi}_{\rm s}(t)\hat{\boldsymbol{\theta}}(t), \tag{290}$$

$$\hat{\boldsymbol{v}}(t)=\boldsymbol{y}(t)-\boldsymbol{\Phi}_{\rm s}(t)\hat{\boldsymbol{\theta}}(t)-\hat{\boldsymbol{\alpha}}^{\rm T}(t)\hat{\boldsymbol{\varphi}}_{\rm n}(t), \tag{291}$$

$$\hat{\boldsymbol{\theta}}(t)=\frac{\hat{\boldsymbol{\theta}}_1(t)+\hat{\boldsymbol{\theta}}_2(t)+\cdots+\hat{\boldsymbol{\theta}}_m(t)}{m}, \tag{292}$$

$$\hat{\boldsymbol{\alpha}}(t)=[\hat{\boldsymbol{\alpha}}_1(t),\hat{\boldsymbol{\alpha}}_2(t),\cdots,\hat{\boldsymbol{\alpha}}_m(t)]. \tag{293}$$

2）部分耦合多新息广义增广随机梯度算法

进一步能够得到部分耦合多新息广义增广随机梯度算法(Partially Coupled Multi-Innovation Generalized Extended Stochastic Gradient algorithm,PC-MI-GESG 算法):

$$\begin{bmatrix}\hat{\boldsymbol{\theta}}_1(t)\\ \hat{\boldsymbol{\alpha}}_1(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\theta}}_m(t-1)\\ \hat{\boldsymbol{\alpha}}_1(t-1)\end{bmatrix}+\frac{\boldsymbol{\Gamma}_1(p,t)}{r_1(t)}\boldsymbol{E}_1(p,t),$$
$$\hat{\boldsymbol{\theta}}_m(0)=\mathbf{1}_n/p_0,\quad \hat{\boldsymbol{\alpha}}_1(0)=\mathbf{1}_{mn_c+mn_d}/p_0, \tag{294}$$

$$\boldsymbol{E}_1(p,t)=\boldsymbol{Y}_1(p,t)-\boldsymbol{\Gamma}_1^{\rm T}(p,t)\begin{bmatrix}\hat{\boldsymbol{\theta}}_m(t-1)\\ \hat{\boldsymbol{\alpha}}_1(t-1)\end{bmatrix}, \tag{295}$$

$$\begin{bmatrix}\hat{\boldsymbol{\theta}}_i(t)\\ \hat{\boldsymbol{\alpha}}_i(t)\end{bmatrix}=\begin{bmatrix}\hat{\boldsymbol{\theta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}+\frac{\boldsymbol{\Gamma}_i(p,t)}{r_i(t)}\boldsymbol{E}_i(p,t),\quad i=2,3,\cdots,m, \tag{296}$$

$$\boldsymbol{E}_i(p,t)=\boldsymbol{Y}_i(p,t)-\boldsymbol{\Gamma}_i^{\rm T}(p,t)\begin{bmatrix}\hat{\boldsymbol{\theta}}_{i-1}(t)\\ \hat{\boldsymbol{\alpha}}_i(t-1)\end{bmatrix}, \tag{297}$$

$$r_i(t)=r_i(t-1)+\|\boldsymbol{\phi}_i(t)\|^2+\|\hat{\boldsymbol{\varphi}}_{\rm n}(t)\|^2,\quad r_i(0)=1, \tag{298}$$

$$\boldsymbol{\Gamma}_i(p,t)=[\hat{\boldsymbol{\psi}}_i(t),\hat{\boldsymbol{\psi}}_i(t-1),\cdots,\hat{\boldsymbol{\psi}}_i(t-p+1)], \tag{299}$$

$$\boldsymbol{Y}_i(p,t)=[y_i(t),y_i(t-1),\cdots,y_i(t-p+1)]^{\mathrm{T}}, \quad (300)$$

$$\hat{\boldsymbol{\psi}}_i(t)=\begin{bmatrix}\boldsymbol{\phi}_i(t)\\ \hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)\end{bmatrix}, \quad (301)$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t)=[-\hat{\boldsymbol{w}}^{\mathrm{T}}(t-1),-\hat{\boldsymbol{w}}^{\mathrm{T}}(t-2),\cdots,-\hat{\boldsymbol{w}}^{\mathrm{T}}(t-n_c), \hat{\boldsymbol{v}}^{\mathrm{T}}(t-1),\hat{\boldsymbol{v}}^{\mathrm{T}}(t-2),\cdots,\hat{\boldsymbol{v}}^{\mathrm{T}}(t-n_d)]^{\mathrm{T}}, \quad (302)$$

$$\boldsymbol{\Phi}_{\mathrm{s}}(t)=[\boldsymbol{\phi}_1(t),\boldsymbol{\phi}_2(t),\cdots,\boldsymbol{\phi}_m(t)]^{\mathrm{T}}, \quad (303)$$

$$\hat{\boldsymbol{w}}(t)=\boldsymbol{y}(t)-\boldsymbol{\Phi}_{\mathrm{s}}(t)\hat{\boldsymbol{\theta}}_i(t), \quad (304)$$

$$\hat{\boldsymbol{v}}(t)=\boldsymbol{y}(t)-\boldsymbol{\Phi}_{\mathrm{s}}(t)\hat{\boldsymbol{\theta}}(t)-\hat{\boldsymbol{\alpha}}^{\mathrm{T}}(t)\hat{\boldsymbol{\varphi}}_{\mathrm{n}}(t). \quad (305)$$

本文中有些算法可以引入遗忘因子和修正指数,得到相应的遗忘因子辨识算法和修正辨识算法,这里不一一讨论.

5 结语

本文将耦合辨识概念与多新息辨识理论相结合,研究了多元伪线性滑动平均系统和伪线性自回归滑动平均系统的多元随机梯度类辨识算法、部分耦合子系统随机梯度类算法、部分耦合随机梯度类算法、部分耦合子系统多新息随机梯度类辨识算法、部分耦合多新息随机梯度类辨识算法等.进一步可以研究多元伪线性滑动平均系统的部分耦合随机梯度类辨识算法和部分耦合最小二乘类算法、部分耦合多新息随机梯度类算法、部分耦合多新息最小二乘算法的计算效率及收敛性.

参考文献
References

[1] 丁锋.系统辨识新论[M].北京:科学出版社,2013
DING Feng. System identification: New theory and methods[M].Beijing:Science Press,2013
[2] 丁锋.系统辨识:辨识方法性能分析[M].北京:科学出版社,2014
DING Feng. System identification: Performance analysis for identification methods [M]. Beijing: Science Press,2014
[3] 丁锋.系统辨识(1):辨识导引[J].南京信息工程大学学报:自然科学版,2011,3(1):1-22
DING Feng. System identification. Part A: Introduction to the identification[J].Journal of Nanjing University of Information Science and Technology:Natural Science Edition,2011,3(1):1-22
[4] 丁锋.系统辨识(2):系统描述的基本模型[J].南京信息工程大学学报:自然科学版,2011,3(2):97-117
DING Feng.System identification.Part B:Basic models for system description[J].Journal of Nanjing University of Information Science and Technology:Natural Science Edition,2011,3(2):97-117
[5] 丁锋.系统辨识(3):辨识精度与辨识基本问题[J].南京信息工程大学学报:自然科学版,2011,3(3):193-226
DING Feng. System identification. Part C: Identification accuracy and basic problems[J].Journal of Nanjing University of Information Science and Technology: Natural Science Edition,2011,3(3):193-226
[6] 丁锋.系统辨识(4):辅助模型辨识思想与方法[J].南京信息工程大学学报:自然科学版,2011,3(4):289-318
DING Feng.System identification.Part D:Auxiliary model identification idea and methods[J].Journal of Nanjing University of Information Science and Technology:Natural Science Edition,2011,3(4):289-318
[7] 丁锋.系统辨识(5):迭代搜索原理与辨识方法[J].南京信息工程大学学报:自然科学版,2011,3(6):481-510
DING Feng.System identification.Part E:Iterative search principle and identification methods [J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2011, 3 (6): 481-510
[8] 丁锋.系统辨识(6):多新息辨识理论与方法[J].南京信息工程大学学报:自然科学版,2012,4(1):1-28
DING Feng.System identification.Part F:Multi-innovation identification theory and methods[J].Journal of Nanjing University of Information Science and Technology: Natural Science Edition,2012,4(1):1-28
[9] 丁锋.系统辨识(7):递阶辨识原理与方法[J].南京信息工程大学学报:自然科学版,2012,4(2):97-124
DING Feng. System identification. Part G: Hierarchical identification principle and methods [J]. Journal of Nanjing University of Information Science and Technology:Natural Science Edition,2012,4(2):97-124
[10] 丁锋.系统辨识(8):耦合辨识概念与方法[J].南京信息工程大学学报:自然科学版,2012,4(3):193-212
DING Feng.System identification.Part H:Coupling identification concept and methods[J].Journal of Nanjing University of Information Science and Technology: Natural Science Edition,2012,4(3):193-212
[11] 丁锋.辨识方法的计算效率(1):递推算法[J].南京信息工程大学学报:自然科学版,2012,4(4):289-300
DING Feng.Computational efficiency of the identification methods.Part A:Recursive algorithms[J].Journal of Nanjing University of Information Science and Technology: Natural Science Edition,2012,4(4):289-300
[12] 丁锋.辨识方法的计算效率(2):迭代算法[J].南京信息工程大学学报:自然科学版,2012,4(5):385-401
DING Feng.Computational efficiency of the identification methods. Part B: Iterative algorithms [J]. Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2012, 4 (5): 385-401
[13] 丁锋.辨识方法的计算效率(3):信息向量耦合算法[J].南京信息工程大学学报:自然科学版,2012,4(6):481-495
DING Feng.Computational efficiency of the identification methods. Part C: Coupled information vector algorithms [J].Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 2012, 4(6):

481-495.
[14] 丁锋,汪菲菲.多元系统耦合多新息随机梯度类辨识方法[J].南京信息工程大学学报:自然科学版,2014,6(1):1-16
DING Feng,WANG Feifei.Coupled multi-innovation stochastic gradient type identification methods for multivariate systems[J].Journal of Nanjing University of Information Science and Technology:Natural Science Edition,2014,6(1):1-16
[15] 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
[16] 丁锋,杨家本.大系统的递阶辨识[J].自动化学报,1999,25(5):647-654
DING Feng,YANG Jiaben.Hierarchical identification of large scale systems[J].Acta Automatica Sinica,1999,25(5):647-654
[17] 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
[18] Ding F.Coupled-least-squares identification for multivariable systems[J].IET Control Theory and Applications,2013,7(1):68-79
[19] Ding F,Chen T.Hierarchical gradient-based identification of multivariable discrete-time systems[J].Automatica,2005,41(2):315-325
[20] 丁锋,丁韬,杨家本,等.时变参数遗忘梯度估计算法的收敛性[J].自动化学报,2002,28(6):962-968
DING Feng,DING Tao,YANG Jiaben,et al.Convergence of forgetting gradient estimation algorithm for time-varying parameters[J].Acta Automatica Sinica,2002,28(6):962-968
[21] Ding F,Chen T.Performance analysis of multi-innovation gradient type identification methods[J].Automatica,2007,43(1):1-14
[22] Ding J,Shi Y,Wang H G,et al.A modified stochastic gradient based parameter estimation algorithm for dual-rate sampled-data systems[J].Digital Signal Processing,2010,20(4):1238-1247
[23] Chen J,Ding F.Modified stochastic gradient algorithms with fast convergence rates[J].Journal of Vibration and Control,2011,17(9):1281-1286

Partially coupled multi-innovation stochastic gradient type identification methods for multivariate pseudo-linear regressive systems

DING Feng[1,2,3] WANG Feifei[1] WANG Xuehai[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 For multivariate pseudo-linear regressive moving average systems,a multivariate extended stochastic gradient(ESG) algorithm is discussed.In order to reduce the computational cost of the identification algorithm,we decompose a multivariate system into several subsystems,and derive a partially coupled(subsystem) ESG algorithm and a partially coupled(subsystem) multi-innovation ESG algorithm according to the coupling identification concept and the multi-innovation identification theory.Furthermore,we extend these methods to multivariate pseudo-linear autoregressive moving average systems and present a partially coupled(subsystem) generalized extended stochastic gradient (GESG) algorithm and a partially coupled (subsystem) multi-innovation GESG algorithm. The computational efficiencies of the multivariate ESG algorithm,the partially coupled ESG algorithm and the partially coupled multi-innovation ESG algorithm are analyzed.

Key words parameter estimation;recursive identification;gradient search;least squares;auxiliary model identification idea;multi-innovation identification theory;hierarchical identification principle;coupling identification concept;multivariate system