

系统辨识(4): 辅助模型辨识思想与方法

丁锋^{1 2 3}

摘要

辅助模型辨识思想、多新息辨识理论、递阶辨识原理、耦合辨识概念是该文作者提出的研究辨识问题的原创性新方法,已经被用在很多辨识问题的研究中,形成了不同的辨识方法族,可以用于解决许多线性或非线性模型的自适应信号处理、自适应参数估计、自适应滤波和预测、自适应控制等问题.由于客观事物具有双重属性:一些特征变量是可观测的;一些是不可测的.如果表征系统特征的观测变量都是可测的,就容易建立描述其运动规律的数学模型.客观事物的不可测属性给建立系统数学模型带来特别的困难.在这种情况下,如何利用系统的可测信息,实现对系统未知变量的估算,来建立系统的数学模型,是辨识领域极具挑战性的研究课题.辅助模型辨识思想就是在这样的背景下发展起来的.该文介绍辅助模型辨识思想和一些基于辅助模型的辨识方法.

关键词

辅助模型; 递推辨识; 参数估计; FIR 模型; CAR 模型; CARMA 模型; CARAR 模型; CARARMA 模型; 输出误差模型; OEMA 模型; OEAR 模型; 辅助模型辨识; 多新息辨识; 递阶辨识; 耦合辨识

中图分类号 TP273

文献标志码 A

收稿日期 2011-07-18

资助项目 国家自然科学基金(60973043)

作者简介

丁锋,男,博士,教授,博士生导师,主要从事系统辨识、过程建模、自适应控制方面的研究. fding@jiangnan.edu.cn

0 引言

自动化控制科学的高度发展,以及对科学最伟大贡献的标志在于计算机的诞生和一些自动化电子产品的普及^[1-7].在我们欣赏和享受这些打上时代烙印和刻画了控制科学家痕迹的自动化电子产品时,在我们感谢自动化科学家给人类铸造了高度的科学文明和人类的辉煌成就时,我们不禁要问:科学是什么?

科学就是解释自然现象,探索未知世界,揭示事物的运动规律.探索未知世界,其结果具有不可预见性.探索事物的运动规律具有不可预见性,探索事物的本质特征具有不可预见性.事物的一些表层特征可以通过观察得到,事物的本质特征可以通过对观测信息(数据)的处理、加工、抽象、推断获得.然而,客观事物具有双重属性:一些特征变量是可观测的;一些是不可测的.对于控制科学而言,就利用系统观测输入输出数据了解系统本质特征——建立系统数学模型的系统辨识而言,情况更是如此.对于可测性,如果表征系统特征的输入输出数据都是可测的,其内部过程变量(或状态变量)也是可测的,这样的系统对我们来说就是“透明”的,很多辨识方法可以用来建立描述其运动规律的数学模型.由于客观事物的不可测属性给我们认识世界,给我们建立描述事物运动规律的数学模型带来特别的困难,在这种情况下,如何利用系统的可测信息,实现对系统未知变量的估算,来建立系统的数学模型是摆在控制科学家面前的首要任务.这也是辅助模型辨识思想产生的根源.

辅助模型辨识思想(auxiliary model identification idea)是借助于系统的可测信息,建立一个辅助模型,用辅助模型的输出代替系统的不可测变量的一种辨识方法.辅助模型辨识思想起源于本文作者的硕士学位论文《多变量系统辨识》(1990年)^[8].其第1篇基于辅助模型辨识思想的研究论文“传递函数阵子模型参数递推估计:辅助模型方法”发表在1991年《控制与决策》第6期上^[9].起初,辅助模型思想被用来研究多变量系统传递函数阵(transfer function matrix)子模型(Sub-SubModel SSM)的辨识问题^[8-9];1992年被用于研究输出误差模型描述的多变量系统辨识问题,提出了输出误差多变量系统的辅助模型递推最小二乘辨识算法^[10];1997年利用鞅收敛定理详细地分析了文献[10]中的辅助模型递推最小二乘辨识方法的性能^[11].后经发展,2004年辅助模型辨识思想被首次介绍到国际控制领域,一些辅

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

助模型辨识研究成果发表在国际著名期刊《Automatica》^[12-13]和《IEEE Transactions on Automatic Control》^[14]上. 如今辅助模型辨识思想已经推广用于解决双率(多率)采样数据系统^[12, 14]、损失数据系统或稀少量测数据系统^[13, 15], 以及输入非线性输出误差系统^[16]的辨识问题.

众所周知, 在 CARMA 系统的最小二乘辨识模型 $y(t) = \varphi^T(t) \theta + v(t)$ 中, 信息向量 $\varphi(t)$ 除包含系统可测输入输出数据 $u(t-i)$ 和 $y(t-i)$ 外, 还包含了不可测噪声项 $v(t-i)$, 递推增广最小二乘辨识算法中是用其估计残差 $\hat{v}(t-i)$ 代替信息向量中的不可测项 $v(t-i)$. 将估计值代替不可测未知变量的思想加以发展, 也可解释我们的辅助模型辨识思想.

辅助模型辨识是用系统的可测信息建立一个辅助模型, 用辅助模型的输出代替系统的不可测变量(如系统的真实输出或系统的状态)的一种辨识方法, 并且通过选择辅助模型的参变量(参数), 使辅助模型(auxiliary model)的输出逼近这些不可测变量, 从而获得系统参数的一致估计^[12-17]. 辅助模型辨识方法是估计一些存在未知变量系统的一种有效辨识方法, 如输出误差类系统, 本文即将讨论的输入非线性方程误差类系统, 输入非线性输出误差类系统等.

本文将以最小二乘方法为主线研究辅助模型类辨识方法, 即研究输出误差类(OET)系统的辅助模型递推最小二乘方法, 输入非线性方程误差类(IN-EET)系统的辅助模型递推最小二乘辨识方法, 输入非线性输出误差类(IN-OET)系统的辅助模型递推最小二乘辨识方法等. 当然, 也可以用随机梯度方法、多新息随机梯度方法、多新息最小二乘方法来研究本文的辅助模型类辨识方法.

本文较长, 为便于阅读, 特将本文框架结构列示如下.

- 0 引言
- 1 辅助模型辨识思想
- 2 线性输出误差类系统(OET)
 - 2.1 线性输出误差系统(OE)
 - 2.2 线性输出误差滑动平均系统(OEMA)
 - 2.3 线性输出误差自回归系统(OEAR)
 - 2.4 线性 Box-Jenkins 系统(BJ)
- 3 输入非线性方程误差类系统(IN-EET)
 - 3.1 输入非线性有限脉冲响应系统(IN-FIR)

- 3.2 输入非线性受控自回归系统(IN-CAR)
- 3.3 输入非线性受控自回归滑动平均系统(IN-CARMA)
- 3.4 输入非线性受控自回归自回归系统(IN-CARAR)
- 3.5 输入非线性受控自回归自回归滑动平均系统(IN-CARARMA)
- 4 输入非线性输出误差类系统(IN-OET)
 - 4.1 输入非线性输出误差系统(IN-OE)
 - 4.2 输入非线性输出误差滑动平均系统(IN-OEMA)
 - 4.3 输入非线性输出误差自回归系统(IN-OEAR)
 - 4.4 输入非线性 Box-Jenkins 系统(IN-BJ)
 - 4.5 其他输入非线性系统
- 5 输出非线性输出误差类系统(ON-OET)
- 6 结语

1 辅助模型辨识思想

为方便起见, 设 $\{u(t)\}$ 为系统输入序列, $\{y(t)\}$ 为系统观测输出序列, $\{v(t)\}$ 是零均值方差为 σ^2 的白噪声序列, z^{-1} 为单位后移算子: $z^{-1}y(t) = y(t-1)$ 或 $zy(t) = y(t+1)$, $A(z)$, $B(z)$, $C(z)$ 和 $D(z)$ 是算子 z^{-1} 的常系数时不变多项式, 定义如下:

$$A(z) : = 1 + a_1z^{-1} + a_2z^{-2} + \cdots + a_{n_a}z^{-n_a} \in \mathbf{R};$$

$$B(z) : = b_1z^{-1} + b_2z^{-2} + \cdots + b_{n_b}z^{-n_b} \in \mathbf{R};$$

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

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

多项式系数 a_i , b_i , c_i 和 d_i 为模型参数. 根据移位算子的性质, 有

$$A(z)y(t) = (1 + a_1z^{-1} + a_2z^{-2} + \cdots + a_{n_a}z^{-n_a})y(t) = y(t) + a_1y(t-1) + a_2y(t-2) + \cdots + a_{n_a}y(t-n_a),$$

$$B(z)u(t) = (b_1z^{-1} + b_2z^{-2} + \cdots + b_{n_b}z^{-n_b})u(t) = b_1u(t-1) + b_2u(t-2) + \cdots + b_{n_b}u(t-n_b),$$

$$D(z)v(t) = (1 + d_1z^{-1} + d_2z^{-2} + \cdots + d_{n_d}z^{-n_d})v(t) = v(t) + d_1v(t-1) + d_2v(t-2) + \cdots + d_{n_d}v(t-n_d) \text{ 等.}$$

实际工业系统中, 经常有一些变量的值不可能得到(如系统的状态, 内部变量, 或系统的某些输出值), 对于这类存在一些不可测变量的系统, 如何估计其参数, 就是辅助模型辨识方法要回答的问题.

递推最小二乘算法处理的 CAR 系统辨识模型

$$y(t) = \phi^T(t) \theta + v(t),$$

假设由系统量测数据构成的信息向量 $\phi(t)$ 的各元都是可以得到的. 递推增广最小二乘算法处理的 CARMA 系统辨识模型

$$y(t) = \varphi^T(t) \theta + v(t),$$

假设信息向量 $\varphi(t)$ 除包含系统的可测输入输出数据外,还包含了不可测噪声项 $v(t-i)$. 解决的办法是这些不可测噪声项 $v(t-i)$ 用其估计残差 $\hat{v}(t-i)$ 代替.

类似地,如果信息向量 $\varphi(t)$ 中还包括系统的不可测内部变量,如系统的真实输出(即无噪输出) $x(t-i)$,我们可以用辅助模型的输出 $x_a(t-i)$ 或其估计 $\hat{x}(t-i)$ 代替,基于这种代替思想的辨识方法,我们称为辅助模型辨识方法. 这种“代替”的思想用于系统辨识,就是辅助模型辨识思想. 下面简单介绍基于辅助模型辨识思想的辨识原理.

假设线性参数系统的输入输出关系可以用下列函数关系表示:

$$y(t) = g(\vartheta, x(t), x(t-1), \dots, x(t-n), \mu(t), u(t-1), \dots, \mu(t-n)) + v(t); \quad (1)$$

$$x(t) = f(\theta, x(t-1), x(t-2), \dots, x(t-n), \mu(t), \mu(t-1), \dots, \mu(t-n)). \quad (2)$$

其中: $y(t)$ 和 $u(t)$ 是可测的输出和输入; $x(t)$ 是中间变量(未知内部变量); ϑ 和 θ 是系统的参数向量; $v(t)$ 是零均值随机白噪声. 对于线性参数系统,函数 $g(\vartheta, x(t), x(t-1), \dots, x(t-n), \mu(t), \mu(t-1), \dots, \mu(t-n))$ 是参数向量 ϑ 的线性函数,是 $x(t), x(t-1), \dots, x(t-n), \mu(t), \mu(t-1), \dots, \mu(t-n)$ 的非线性函数;函数 $f(\theta, x(t-1), x(t-2), \dots, x(t-n), \mu(t), \mu(t-1), \dots, \mu(t-n))$ 是参数向量 θ 的线性函数,是 $x(t-1), x(t-2), \dots, x(t-n), \mu(t), \mu(t-1), \dots, \mu(t-n)$ 的非线性函数. 为简化表达,引入移位算子 $z: zx(t) = x(t+1)$ 或 $z^{-1}x(t) = x(t-1)$,它们可以简单记作:

$$g(\vartheta, x(t), \mu(t), z) := g(\vartheta, x(t), x(t-1), \dots, x(t-n), \mu(t), \mu(t-1), \dots, \mu(t-n));$$

$$f(\theta, x(t-1), \mu(t), z) := f(\theta, x(t-1), x(t-2), \dots, x(t-n), \mu(t), \mu(t-1), \dots, \mu(t-n)).$$

式(1) — (2) 可以等价写为

$$y(t) = g(\vartheta, x(t), \mu(t), z) + v(t), \quad (3)$$

$$x(t) = f(\theta, x(t-1), \mu(t), z). \quad (4)$$

假设函数 $g(\ast)$ 可以写为或分解为下列形式,

$$g(\vartheta, x(t), \mu(t), z) =$$

$$x(t) + g'(\vartheta, x(t-1), \mu(t), z), \quad (5)$$

那么式(3)可以写为

$$y(t) = x(t) + g'(\vartheta, x(t-1), \mu(t), z) + v(t) = f(\theta, x(t-1), \mu(t), z) + g'(\vartheta, x(t-1), \mu(t), z) + v(t).$$

因为假设是线性参数系统 $f(\ast)$ 和 $g'(\ast)$ 分别是参数向量 θ 和 ϑ 的线性函数,所以上式可以写为下列辨识模型

$$y(t) = \varphi^T(x(t-1), \mu(t), z) \begin{bmatrix} \theta \\ \vartheta \end{bmatrix} + v(t) = \varphi^T(x(t-1), \mu(t), z) \Theta + v(t), \quad (6)$$

其中 $\Theta = \begin{bmatrix} \theta \\ \vartheta \end{bmatrix}$, 信息向量 $\varphi(x(t-1), \mu(t), z) \in \mathbf{R}^m$ 的元 $\varphi_i(x(t-1), \mu(t), z)$ 也是 $x(t-1), x(t-2), \dots, x(t-n), \mu(t), \mu(t-1), \dots, \mu(t-n)$ 的非线性函数,即

$$\varphi(x(t-1), \mu(t), z) = \begin{bmatrix} \varphi_1(x(t-1), \mu(t), z) \\ \varphi_2(x(t-1), \mu(t), z) \\ \vdots \\ \varphi_m(x(t-1), \mu(t), z) \end{bmatrix} \in \mathbf{R}^m,$$

$$\varphi(x(t-1), \mu(t), z) = \varphi(x(t-1), x(t-2), \dots, x(t-n), \mu(t), \mu(t-1), \dots, \mu(t-n)) \in \mathbf{R}^m,$$

$$\varphi_i(x(t-1), \mu(t), z) = \varphi_i(x(t-1), x(t-2), \dots, x(t-n), \mu(t), \mu(t-1), \dots, \mu(t-n)) \in \mathbf{R}.$$

$$\text{设 } \hat{\Theta}(t) := \begin{bmatrix} \hat{\theta}(t) \\ \hat{\vartheta}(t) \end{bmatrix} \text{ 是时刻 } t \text{ 参数向量 } \Theta =$$

$\begin{bmatrix} \theta \\ \vartheta \end{bmatrix}$ 的估计. 对于辨识模型(6),极小化准则函数(criterion function)

$$J_1(\Theta) := \sum_{j=1}^t [y(j) - \varphi^T(x(j-1), \mu(j), z) \Theta]^2,$$

可以得到估计参数向量 Θ 的递推最小二乘算法^[18]:

$$\hat{\Theta}(t) = \hat{\Theta}(t-1) + L(t) [y(t) - \varphi^T(x(t-1), \mu(t), z) \hat{\Theta}(t-1)], \quad (7)$$

$$L(t) = P(t-1) \varphi(x(t-1), \mu(t), z) [1 + \varphi^T(x(t-1), \mu(t), z) P(t-1) \varphi(x(t-1), \mu(t), z)]^{-1}, \quad (8)$$

$$P(t) = [I - L(t) \varphi^T(x(t-1), \mu(t), z)] P(t-1). \quad (9)$$

由于中间变量 $x(t-i)$ 是未知的,上述计算参数估计 $\hat{\Theta}(t)$ 的算法无法实现. 借助于辅助模型辨识思想: 设 $x_a(t)$ 是辅助模型的输出,用 $x_a(t-1)$ 代替 $\varphi(x(t-1), \mu(t), z)$ 中未知 $x(t-1)$,代替后的信息向量记作

$$\hat{\varphi}(x_a(t-1), \mu(t), z) := \hat{\varphi}(x_a(t-1), x_a(t-2), \dots, x_a(t-n), \mu(t), \mu(t-1), \dots, \mu(t-n)) \in \mathbf{R}^m.$$

根据式(4),辅助模型可以取为

$$x_a(t) = f(\hat{\theta}(t) \ x_a(t-1) \ \mu(t) \ z) = f(\hat{\theta}(t) \ x_a(t-1) \ x_a(t-2) \ ; \dots \ x_a(t-n) \ , u(t) \ \mu(t-1) \ ; \dots \ \mu(t-n) \) .$$

用 $\hat{\varphi}(x_a(t-1) \ \mu(t) \ z)$ 代替式(7) — (8) 中 $\varphi(x(t-1) \ \mu(t) \ z)$, 我们可以得到估计参数向量 θ 的辅助模型递推最小二乘算法(AM-RLS):

$$\hat{\theta}(t) = \hat{\theta}(t-1) + P(t) [y(t) - \hat{\varphi}^T(x_a(t-1) \ , u(t) \ z) \hat{\theta}(t-1)] , \quad (10)$$

$$L(t) = P(t-1) \hat{\varphi}(x_a(t-1) \ \mu(t) \ z) [1 + \hat{\varphi}^T(x_a(t-1) \ \mu(t) \ z) P(t-1) \hat{\varphi}(x_a(t-1) \ \mu(t) \ z)]^{-1} , \quad (11)$$

$$P(t) = [I - L(t) \hat{\varphi}^T(x_a(t-1) \ \mu(t) \ z)] P(t-1) , \quad P(0) = p_0 I > 0 , \quad (12)$$

$$\hat{\varphi}(x_a(t-1) \ \mu(t) \ z) = \hat{\varphi}(x_a(t-1) \ x_a(t-2) \ , \dots \ x_a(t-n) \ \mu(t) \ \mu(t-1) \ ; \dots \ \mu(t-n) \) , \quad (13)$$

$$x_a(t) = f(\hat{\theta}(t) \ x_a(t-1) \ \mu(t) \ z) = f(\hat{\theta}(t) \ x_a(t-1) \ x_a(t-2) \ ; \dots \ x_a(t-n) \ , u(t) \ \mu(t-1) \ ; \dots \ \mu(t-n) \) , \quad (14)$$

$$\hat{\theta}(t) = \begin{bmatrix} \hat{\theta}(t) \\ \hat{\varphi}(t) \end{bmatrix} . \quad (15)$$

这就是基于辅助模型辨识思想推导出的递推最小二乘辨识算法. 系统的输出预报为

$$\hat{y}(t) = g(\hat{\vartheta}(t) \ x_a(t) \ \mu(t) \ z) = g(\hat{\vartheta}(t) \ x_a(t) \ x_a(t-1) \ ; \dots \ x_a(t-n) \ , u(t) \ \mu(t-1) \ ; \dots \ \mu(t-n) \) . \quad (16)$$

为简化, 如果用 $\hat{x}(t)$ 表示 $x_a(t)$, 即令 $\hat{x}(t) := x_a(t)$, 那么算法(10) — (15) 可以等价写为

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t) [y(t) - \hat{\varphi}^T(\hat{x}(t-1) \ , u(t) \ z) \hat{\theta}(t-1)] , \quad (17)$$

$$L(t) = P(t-1) \hat{\varphi}(\hat{x}(t-1) \ \mu(t) \ z) [1 + \hat{\varphi}^T(\hat{x}(t-1) \ \mu(t) \ z) P(t-1) \hat{\varphi}(\hat{x}(t-1) \ \mu(t) \ z)]^{-1} , \quad (18)$$

$$P(t) = [I - L(t) \hat{\varphi}^T(\hat{x}(t-1) \ \mu(t) \ z)] P(t-1) , \quad P(0) = p_0 I , \quad (19)$$

$$\hat{\varphi}(\hat{x}(t-1) \ \mu(t) \ z) = \hat{\varphi}(\hat{x}(t-1) \ \hat{x}(t-2) \ , \dots \ \hat{x}(t-n) \ \mu(t) \ \mu(t-1) \ ; \dots \ \mu(t-n) \) , \quad (20)$$

$$\hat{x}(t) = f(\hat{\theta}(t) \ \hat{x}(t-1) \ \mu(t) \ z) = f(\hat{\theta}(t) \ \hat{x}(t-1) \ \hat{x}(t-2) \ ; \dots \ \hat{x}(t-n) \ , u(t) \ \mu(t-1) \ ; \dots \ \mu(t-n) \) , \quad (21)$$

$$\hat{\theta}(t) = \begin{bmatrix} \hat{\theta}(t) \\ \hat{\vartheta}(t) \end{bmatrix} . \quad (22)$$

系统输出预报可以等价写为

$$\hat{y}(t) = g(\hat{\vartheta}(t) \ \hat{x}(t) \ \mu(t) \ z) = g(\hat{\vartheta}(t) \ \hat{x}(t) \ \hat{x}(t-1) \ ; \dots \ \hat{x}(t-n) \ \mu(t) \ ,$$

$$u(t-1) \ ; \dots \ \mu(t-n) \) . \quad (23)$$

考虑下列伪线性回归模型 II (一类白噪声干扰的输出误差模型)

$$y(t) = \frac{f(\vartheta \ \mu_1(t) \ z)}{A(z)} + B_2(z) u_2(t) + v(t) ,$$

其中 $u_1(t)$ 和 $u_2(t)$ 是系统的 2 个输入 $f(\vartheta \ \mu_1(t) \ z) := f(\vartheta \ \mu_1(t) \ \mu_1(t-1) \ ; \dots \ \mu_1(t-n) \)$ 是参数向量 ϑ 的线性函数, 是 $u_1(t) \ \mu_1(t-1) \ ; \dots \ \mu_1(t-n)$ 的非线性函数. 因此, 这个伪线性回归模型 II 是非线性系统, 可以用辅助模型方法来解决其辨识问题, 只需定义下列辅助模型

$$\hat{x}(t) = \frac{f(\hat{\vartheta}(t) \ \mu_1(t) \ z)}{\hat{A}(t, z)} ,$$

其中 $\hat{A}(t, z)$ 是 t 时刻 $A(z)$ 的估计.

辅助模型辨识思想可以解决一些存在不可测变量的系统辨识问题. 基于辅助模型辨识思想的辨识理论和方法已成为一类辨识方法族, 可以用于研究(包括有色噪声模型干扰下)线性和非线性模型的自适应信号处理、自适应滤波和预测、自适应辨识和参数估计等问题. 辅助模型辨识的相关成果和值得研究的课题如下:

- 1) 稀少量测数据系统的辅助模型多新息随机梯度辨识方法^[13];
- 2) 损失数据输出误差系统的辅助模型(遗忘因子)递推最小二乘辨识方法^[15];
- 3) 输出误差系统(输出误差滑动平均系统)的辅助模型(增广)随机梯度辨识算法和辅助模型递推(增广)最小二乘辨识算法^[19];
- 4) 输出误差系统辅助模型多新息随机梯度算法的性能分析^[20];
- 5) 多输入单输出系统的辅助模型多新息随机梯度算法^[21];
- 6) Hammerstein 输出误差滑动平均系统的辅助模型递推增广最小二乘辨识算法和辅助模型多新息增广最小二乘辨识算法^[22];
- 7) 输入非线性输出误差自回归系统的辅助模型递推广义最小二乘参数估计算法^[23];
- 8) 输出误差滑动平均模型(OEMA)的辅助模型多新息增广随机梯度辨识方法^[24];
- 9) 多率多输入系统的辅助模型递推最小二乘算法^[25];
- 10) 输出误差滑动平均模型的辅助模型递推增广最小二乘辨识方法^[26];
- 11) 非均匀周期采样系统的辅助模型递推最小

二乘算法及其收敛性^[27-28];

12.) Box-Jenkins 系统的辅助模型多新息广义增广随机梯度辨识方法^[29];

13) 非线性 Hammerstein 模型的辅助模型最小二乘辨识方法、辅助模型随机梯度辨识方法^[16];

14) 基于辅助模型的多输入单输出系统多新息最小二乘辨识方法^[30];

15) 基于辅助模型的双率采样数据系统的联合参数与损失输出估计^[12];

16) 基于有限脉冲响应辅助模型的双率采样数据系统辨识^[17];

17) 传递函数阵子模型参数的辅助模型最小二乘辨识方法^[8-9];

18) 多变量输出误差系统的辅助模型辨识方法及其收敛性分析^[10-11];

19) 随机干扰系统的辅助模型递推广义增广最小二乘辨识方法、辅助模型广义增广随机梯度辨识方法^[31];

20) 辅助模型多新息辨识方法族、辅助模型递阶辨识方法族。

2 线性输出误差类系统(OET)

2.1 线性输出误差系统(OE)

辅助模型辨识方法是研究一些存在不可测内部变量系统的一种有效辨识方法,下面以简单线性系统的输出误差模型为基础,介绍辅助模型辨识方法。这种方法可以推广到双线性参数输出误差系统和非线性输出误差系统^[16,18]。本节主要参考文献[10-12, 14,18-19]。

2.1.1 辅助模型递推最小二乘法(AM-RLS)

考虑输出误差模型(Output Error model,OE)描述的系统(参见图1):

$$y(t) = \frac{B(z)}{A(z)}u(t) + v(t), \quad (24)$$

其中 $\{u(t)\}$ 和 $\{y(t)\}$ 是可测的系统输入和输出序列, $G(z) = \frac{B(z)}{A(z)}$ 为系统的传递函数。

输出误差模型示意如图1所示,其中 $x(t)$ 为系统的真实输出(true output)或无噪输出(noise-free output)(不可得到) $y(t)$ 是 $x(t)$ 的含噪量测。

定义参数向量 θ 和信息向量 $\varphi(t)$ 如下:

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

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

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

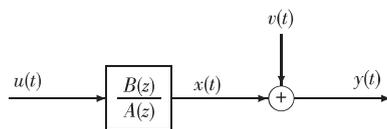


图1 输出误差系统(OE)

Fig.1 The output error system

定义中间变量(inner variable):

$$x(t) := \frac{B(z)}{A(z)}u(t) \quad (25)$$

或

$$A(z)x(t) = B(z)u(t).$$

式(25)写为差分方程

$$x(t) + a_1x(t-1) + a_2x(t-2) + \dots + a_{n_a}x(t-n_a) = b_1u(t-1) + b_2u(t-2) + \dots + b_{n_b}u(t-n_b)$$

或

$$x(t) = \varphi^T(t)\theta. \quad (26)$$

从而可得输出误差系统的辨识模型(identification model):

$$y(t) = x(t) + v(t) = \varphi^T(t)\theta + v(t). \quad (27)$$

这个模型属于伪线性回归模型II,因为信息向量 $\varphi(t)$ 中只包含了不可测真实输出 $x(t-i)$ 。仿照文献中RLS算法的推导,极小化准则函数

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

可得下列递推最小二乘算法:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + P(t)\varphi(t)[y(t) - \varphi^T(t)\hat{\theta}(t-1)], \quad (28)$$

$$P(t) = P(t-1) - \frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)}{1 + \varphi^T(t)P(t-1)\varphi(t)}. \quad (29)$$

因为信息向量 $\varphi(t)$ 中包含了未知的中间变量 $x(t-i)$,所以上述算法不可实现,解决方法是借助于辅助模型辨识思想。参见图2,建立一个与 $\frac{B(z)}{A(z)}$ 结

构相同的辅助模型 $\frac{B_a(z)}{A_a(z)}$,辅助模型的输出 $x_a(t)$ 可以表示为

$$x_a(t) = \frac{B_a(z)}{A_a(z)}u(t).$$

参照式(25)写为式(26)的写法,上式可以表达为

$$x_a(t) = \varphi_a^T(t)\theta_a(t). \quad (30)$$

其中 $\varphi_a(t)$ 和 $\theta_a(t)$ 是 t 时刻辅助模型的信息向量和参数向量。

如果用图2中辅助模型的输出 $x_a(t)$ 代替系统

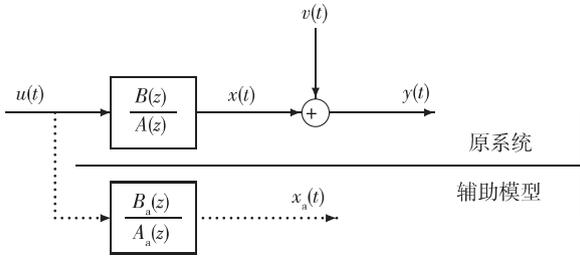


图2 带辅助模型的输出误差系统

Fig.2 The output error system with the auxiliary model

的真实输出 $x(t)$,那么参数向量 θ 的辨识问题可以利用 $u(t)$ 和 $x_a(t)$ 得到解决. 这就是辅助模型辨识思想. 基于这种思想的辨识方法称为辅助模型辨识方法 (auxiliary model based identification method) ,或参考模型辨识方法 (reference model identification method) ,或输出误差辨识方法 (output error identification method) . 如果 $x_a(t)$ 逼近于 $x(t)$,就可以把辅助模型的输出 $x_a(t)$ 作为 $x(t)$ 的一个估计 $\hat{x}(t)$.

问题的关键是 ,如何构造辅助模型(30) ,以便利用数据 $\{u(t) \ y(t)\}$ 获得 $x(t)$ 的估计 $x_a(t)$,并保证 $x_a(t)$ 收敛于 $x(t)$: $x_a(t) \rightarrow x(t)$,那么 $\frac{B(z)}{A(z)}$ 的辨识问题通过使用 $x_a(t)$ 而不是 $x(t)$ 而得到解决. 辅助模型的选择方式将决定辨识算法的具体形式和收敛性能. 当然 ,有许多方法选择辅助模型^[9, 12, 14, 17, 24, 26, 29] ,这里介绍把 $\frac{B(z)}{A(z)}$ 的估计作为辅助模型^[12, 14, 19] 的辨识方法 ,文献[17]介绍了把 $\frac{B(z)}{A(z)}$ 的脉冲响应模型估计作为辅助模型.

设 $\hat{\theta}(t)$ 代表参数向量 θ 的估计. 我们把 $\hat{\theta}(t)$ 作为辅助模型的参数向量 $\theta_a(t)$,用 $x_a(t-i)$ 代替 $x(t-i)$ 后的信息向量 $\varphi(t)$ 记作 $\hat{\varphi}(t)$,同时也把 $\hat{\varphi}(t)$ 作为辅助模型的信息向量 $\varphi_a(t)$,因此有^[12-14, 19]

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

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

在 t 时刻 $\hat{\varphi}(t)$ 是已知的. 因此极小化准则函数

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

可以得到估计参数向量 θ 的基于辅助模型的递推最小二乘算法 ,简称为辅助模型递推最小二乘算法 (Auxiliary Model based Recursive Least Squares algorithm ,AM-RLS) :

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t) [y(t) - \hat{\varphi}^T(t) \hat{\theta}(t-1)] , \quad (31)$$

$$L(t) = P(t) \hat{\varphi}(t) = \frac{P(t-1) \hat{\varphi}(t)}{1 + \hat{\varphi}^T(t) P(t-1) \hat{\varphi}(t)} , \quad (32)$$

$$P(t) = P(t-1) - \frac{P(t-1) \hat{\varphi}(t) \hat{\varphi}^T(t) P(t-1)}{1 + \hat{\varphi}^T(t) P(t-1) \hat{\varphi}(t)} = [I - L(t) \hat{\varphi}^T(t)] P(t-1) \quad P(0) = p_0 I , \quad (33)$$

$$\hat{\varphi}(t) = [-x_a(t-1) \ , -x_a(t-2) \ , \dots \ , -x_a(t-n_a) \ , \mu(t-1) \ \mu(t-2) \ , \dots \ \mu(t-n_b)]^T , \quad (34)$$

$$x_a(t) = \hat{\varphi}^T(t) \hat{\theta}(t) . \quad (35)$$

这里 辅助模型为

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

辅助模型辨识算法是联合系统参数和系统不可测真实输出同时估计的一种联合辨识方法. 它是一个随 t 增加的递阶交互估计计算过程 ,因为计算参数估计需要用到辅助模型的输出 $x_a(t)$,而计算辅助模型输出 $x_a(t)$ 又要用到参数估计 $\hat{\theta}(t)$. $x_a(t)$ 可以作为 $x(t)$ 的估计 $\hat{x}(t)$. AM-RLS 算法也可以把算法(28) — (29) 中的 $\varphi(t)$ 换为 $\hat{\varphi}(t)$ 得到.

本文作者已证明了辅助模型递推最小二乘辨识算法的一致收敛性^[11-12, 15] . 文献[15]给出了损失数据输出误差系统的遗忘因子辅助模型递推最小二乘辨识算法 ,但只分析了遗忘因子 $\lambda = 1$ 时的情形 ,即不带遗忘因子的辅助模型递推最小二乘算法的收敛性. 带遗忘因子辅助模型递推最小二乘算法的收敛条件以及参数误差界分析 ,也是有待辨识爱好者解决的科学难题^[15] .

AM-RLS 算法的计算步骤如下.

- 1) 初始化. 令 $t=1$. 置初值 $P(0) = p_0 I \ \hat{\theta}(0) = 1/p_0 \ p_0 = 10^6 \ x_a(-i) = 1/p_0 \ i=0, 1, 2, \dots, n_a-1$.
- 2) 收集数据 $u(t)$ 和 $y(t)$,用式(34)构造 $\hat{\varphi}(t)$.
- 3) 用式(32)计算 $L(t)$,用式(33)计算 $P(t)$.
- 4) 用式(31)刷新估计 $\hat{\theta}(t)$.
- 5) 用式(35)计算 $x_a(t)$.
- 6) t 增 1 转步骤 2 ,继续递推计算.

AM-RLS 算法计算参数估计 $\hat{\theta}(t)$ 的流程如图 3 所示.

仿真试验如下.

例 1 考虑仿真对象(OE 模型) :

$$\begin{cases} y(t) = \frac{B(z)}{A(z)} u(t) + v(t) , \\ A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} = 1 + 0.412z^{-1} + 0.309z^{-2} , \\ B(z) = b_1 z^{-1} + b_2 z^{-2} = 0.6804z^{-1} + 0.6303z^{-2} , \\ \theta = [a_1 \ a_2 \ b_1 \ b_2]^T = [0.412 \ 0.309 \ 0.6804 \ 0.6303]^T . \end{cases}$$

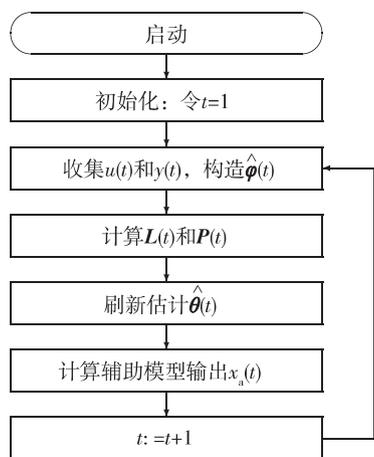


图3 计算 AM-RLS 参数估计 $\hat{\theta}(t)$ 的流程

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

仿真时 输入 $\{u(t)\}$ 采用零均值单位方差不相关可测随机信号序列 $\{v(t)\}$ 采用零均值方差为 σ^2 白噪声序列 改变 σ^2 可以控制噪信比 δ_{ns} (噪信比的定义见文献 [32]). 对于输出误差模型 噪信比定义为

$$\delta_{ns} = \sqrt{\frac{\text{var}[v(t)]}{\text{var}[x(t)]}} \times 100\%.$$

基于量测数据 $\{u(t) \ y(t)\}$,用 AM-RLS 算法估计这个系统的参数 不同噪声方差或噪信比下仿真结果示于表 1. 其中参数估计误差 $\delta: = \|\hat{\theta}(t) - \theta\| / \|\theta\|$ 随 t 变化曲线如图 4 所示 θ 为系统模型参数, $\hat{\theta}(t)$ 为 θ 的估计. 当噪声方差为 $\sigma^2 = 0.10^2$ 和 $\sigma^2 = 0.50^2$ 时,系统的噪信比分别为 $\delta_{ns} = 11.75\%$ 和 $\delta_{ns} =$

58.75%. δ 随 t 的变化曲线可以用来衡量参数估计收敛于真值的速率和精度.

为了说明 AM-RLS 算法的辨识效果,在相同的条件下 ($\sigma^2 = 0.50^2 \ \delta_{ns} = 58.75\%$) ,使用递推最小二乘法 (RLS)

$$\hat{\theta}(t) = \hat{\theta}(t-1) + P(t) \psi(t) [y(t) - \psi^T(t) \hat{\theta}(t-1)], \quad (36)$$

$$P(t) = P(t-1) - \frac{P(t-1) \psi(t) \psi^T(t) P(t-1)}{1 + \psi^T(t) P(t-1) \psi(t)},$$

$$P(0) = p_0 I,$$

$$\psi(t) = [-y(t-1), -y(t-2), \dots, -y(t-n_a), u(t-1), \mu(t-2), \dots, \mu(t-n_b)]^T$$

估计这个输出误差系统的参数,仿真结果示于表 2, 参数估计误差 δ 随 t 变化曲线如图 4 所示.

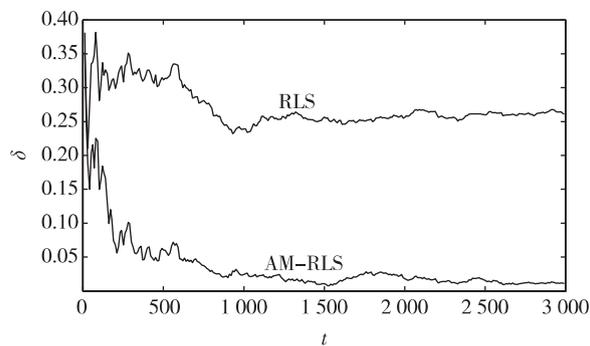


图4 例1的 δ 随 t 的变化曲线 ($\sigma^2 = 0.50^2 \ \delta_{ns} = 58.75\%$)

Fig.4 The estimation errors δ versus t ($\sigma^2 = 0.50^2 \ \delta_{ns} = 58.75\%$)

表 1 例 1 模型参数的 AM-RLS 估计

Table 1 The AM-RLS estimates and errors

δ_{ns}	t	a_1	a_2	b_1	b_2	$\delta / \%$
11.75%	100	0.395 65	0.301 77	0.689 90	0.617 58	2.254 19
	200	0.409 62	0.303 11	0.689 98	0.624 71	1.204 65
	500	0.413 89	0.311 05	0.685 79	0.619 07	1.202 89
	1 000	0.413 35	0.312 94	0.682 83	0.628 25	0.493 65
	2 000	0.410 78	0.310 86	0.682 87	0.627 98	0.382 47
	3 000	0.411 96	0.309 71	0.682 09	0.629 01	0.211 27
	58.75%	100	0.307 86	0.272 83	0.730 35	0.543 58
200		0.404 90	0.274 19	0.728 98	0.606 78	6.091 03
500		0.416 99	0.317 00	0.707 65	0.571 79	6.148 87
1 000		0.417 26	0.327 67	0.692 66	0.619 43	2.392 95
2 000		0.404 67	0.317 71	0.692 87	0.617 90	1.974 82
3 000		0.411 38	0.312 18	0.688 89	0.623 67	1.060 48
真值			0.412 00	0.309 00	0.680 40	0.630 30

表2 例1 模型参数的 RLS 估计
Table 2 The RLS estimates and errors

δ_{ns}	t	a_1	a_2	b_1	b_2	$\delta/\%$
11.75%	100	0.386 07	0.303 12	0.693 43	0.607 52	3.521 07
	200	0.390 71	0.306 34	0.689 13	0.610 19	2.891 77
	500	0.394 45	0.314 65	0.687 90	0.604 68	3.058 16
	1 000	0.398 57	0.310 33	0.683 32	0.616 86	1.816 32
	2 000	0.395 75	0.307 69	0.682 71	0.617 07	1.991 02
	3 000	0.396 10	0.306 74	0.682 39	0.618 07	1.911 61
58.75%	100	0.194 63	0.237 76	0.742 72	0.449 00	28.133 25
	200	0.178 14	0.266 00	0.727 86	0.431 37	29.563 50
	500	0.182 17	0.288 28	0.716 97	0.397 99	31.057 29
	1 000	0.217 93	0.271 32	0.694 23	0.473 14	23.841 41
	2 000	0.198 53	0.257 58	0.693 57	0.471 68	25.563 65
	3 000	0.193 36	0.248 03	0.689 76	0.472 47	26.075 25
真值		0.412 00	0.309 00	0.680 40	0.630 30	

从表1—2和图4可知:递推最小二乘算法给出的参数估计误差大,收敛效果差,参数估计是有偏差的;而辅助模型最小二乘算法参数估计误差随 t 增加而减少,并不断趋于零,说明该方法是有效的。

把下列程序写到 AM_RLS.m 文件中,当 $\sigma = 0.10$ 或 $\sigma = 0.50$ 时,依次运行 $SW = 1$ (AM-RLS 算法)和 $SW = 2$ (RLS 算法),可得到上述例子的仿真结果(参数估计表和误差曲线图)。

```

1 % -----*
2 % Filename: AM_RLS.m for the output error models: *
3 %  $y(t) = B(z)/A(z)u(t) + v(t)$  *
4 % The noise variance  $\sigma^2 = 0.10^2$  and  $\sigma^2 = 0.50^2$  *
5 % The auxiliary model based RLS algorithm *
6 % -----*
7 clear; format short g
8 M = The AM - RLS algorithm ( SW = 1 ) or RLS algorithm
   ( SW = 2 ) '
9 FF = 1; % The Forgetting Factor
10 sigma = 0.10; % The noise variance  $\sigma^2 = 0.10^2$  and
    $\sigma^2 = 0.50^2$ 
11 SW = 2; % SW = 1 for the AM - RLS algorithm
12     % SW = 2 for the RLS algorithm
13 PlotLength = 3000; length1 = PlotLength + 100;
14 na = 2; nb = 2; n = na + nb;
15 a = [1 0.412 0.309]; b = [0 0.6804 0.6303]; d = [1];
16 par0 = [a(2:na+1) b(2:nb+1)]'; % The parameter vector
17
18 p0 = 1e6; P = eye(n) * p0; r = 1;
19 par1 = ones(n,1)/p0; % The parameter estimation
   vector theta

```

```

20 % -----Compute the noise - to - signal ratio
21 sy = f_integral( a ,b ); sv = 1;
22 delta_ns = sqrt( sv/sy ) * 100* sigma;
23 [sy sv delta_ns ]
24 % -----Generate the input - output data
25 rand( 'state' ,rho ); randn( 'state' ,rho );
26 u = ( rand( length1 ,1) - 0.5 ) * sqrt( 12 ); v = randn
   ( length1 ,1 ) * sigma;
27 y = ones( n ,1 ) /p0; x = ones( n ,1 ) /p0; xa = ones( n ,
   1 ) /p0;
28 % -----AM-RLS algorithm ( SW = 1 ) or RLS algorithm
   ( SW = 2 )
29 jj = 0; j1 = 0;
30 for t = n: length1
31     jj = jj + 1;
32     x( t ) = par0 * [ - x( t - 1: - 1: t - na ); u( t - 1: - 1: t
   - nb ) ];
33     y( t ) = x( t ) + v( t );
34     if SW = = 1
35         varphi = [ - xa( t - 1: - 1: t - na ); u( t - 1: - 1: t
   - nb ) ];
36     else
37         varphi = [ - y( t - 1: - 1: t - na ); u( t - 1: - 1: t
   - nb ) ];
38     end
39     L = P* varphi / ( FF + varphi* P* varphi );
40     P = ( P - L* ( varphi* P ) ) / FF;
41     par1 = par1 + L* ( y( t ) - varphi* par1 );
42     delta = norm( par1 - par0 ) / norm( par0 );
43     ls( jj , : ) = [jj par1' delta ];
44     xa( t ) = varphi* par1;

```

```

45 if (jj == 100) | (jj == 200) | jj == 500 | ( mod( jj ,
    1000) == 0)
46     j1 = j1 + 1;
47     ls_100( j1 , : ) = [jj , par1 , delta * 100 ];
48 end
49 end
50 ls_100( j1 + 1 , : ) = [0 , par0 , rho ];
51 fprintf( '\n %s\n' , ...
52     ' $ t $ $ a_1 $ $ a_2 $ $ b_1 $ $ b_2 $
53     $ \delta \ ( \% ) \ $ \ \ ' );
54 fprintf( ' %5d %10.5f %10.5f %10.5f %10.5f %10.5f \
55     \ \ \n ' , ls_100 );
56 figure( 1 );
57 jk = ( 11 : 10 : PlotLength - 1 ) ;
58 plot( ls( jk , 1 ) , ls( jk , n + 2 ) , 'm' )
59
60 if SW == 1 % AM - RLS
61     data1 = [ls( : , 1 ) , ls( : , n + 2 ) ];
62     save data1 data1
63 else % RLS
64     load data1 data1
65     z0 = [data1 , ls( : , n + 2 ) ];
66     figure( 2 );
67     plot( z0( jk , 1 ) , z0( jk , 2 ) , 'k' , z0( jk , 1 ) , z0( jk ,
68         3 ) , 'b' )
69 end
70 xlabel( ' \hat{y}(t) ' ); ylabel( ' \hat{y}(t) - \hat{\theta}(t) ' )

```

2.1.2 辅助模型随机梯度辨识算法(AM-SG)

输出误差类系统的辅助模型随机梯度算法是方程误差类系统随机梯度辨识方法的延伸和推广。

对于输出误差系统(24)对应的辨识模型(27),应用辅助模型辨识思想,极小化准则函数

$$J_4(\theta) = E\{ [y(t) - \hat{\varphi}^T(t)\theta]^2 \}$$

可以得到辨识参数向量 θ 的辅助模型随机梯度辨识算法(Auxiliary Model based Stochastic Gradient algorithm, AM-SG):

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\varphi(t)}{r(t)} [y(t) - \varphi^T(t)\hat{\theta}(t-1)], \quad (37)$$

$$r(t) = r(t-1) + \|\varphi(t)\|^2, \quad r(0) = 1, \quad (38)$$

$$\varphi(t) = [-x_a(t-1), -x_a(t-2), \dots, -x_a(t-n_a), u(t-1), \mu(t-2), \dots, \mu(t-n_b)]^T, \quad (39)$$

$$x_a(t) = \varphi^T(t)\hat{\theta}(t), \quad (40)$$

$$\hat{\theta}(t) = [\hat{a}_1(t), \hat{a}_2(t), \dots, \hat{a}_{n_a}(t), \hat{b}_1(t), \hat{b}_2(t), \dots, \hat{b}_{n_b}(t)]^T. \quad (41)$$

文献[14]研究了双率采样数据系统辅助模型随机梯度算法的收敛性。为了提高AM-SG算法的收敛

速度,在AM-SG算法中引入遗忘因子 λ ,得到辅助模型遗忘因子随机梯度辨识算法(Auxiliary Model based Forgetting Factor Stochastic Gradient algorithm, AM-FFSG),简称辅助模型遗忘梯度辨识算法(Auxiliary Model based Forgetting Gradient algorithm, AM-FG):

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\varphi(t)}{r(t)} [y(t) - \varphi^T(t)\hat{\theta}(t-1)], \quad (42)$$

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

$$\varphi(t) = [-x_a(t-1), -x_a(t-2), \dots, -x_a(t-n_a), u(t-1), \mu(t-2), \dots, \mu(t-n_b)]^T, \quad (44)$$

$$x_a(t) = \varphi^T(t)\hat{\theta}(t), \quad (45)$$

$$\hat{\theta}(t) = [\hat{a}_1(t), \hat{a}_2(t), \dots, \hat{a}_{n_a}(t), \hat{b}_1(t), \hat{b}_2(t), \dots, \hat{b}_{n_b}(t)]^T. \quad (46)$$

当 $\lambda = 1$ 时,AM-FG算法退化为AM-SG算法;当 $\lambda = 0$ 时,AM-FG算法退化为辅助模型投影算法。对于输出误差模型,本文作者等提出了梯度迭代辨识算法和基于最小二乘的迭代辨识方法^[19]。

本文作者2008年发表在《中国科学》上的论文证明了辅助模型随机梯度算法的收敛性,证明中推导出2个精妙的等式,收敛结论也很漂亮^[14, 33-34];发表在控制领域国际著名期刊《Automatica》2011年上的论文证明了稀少数据输出误差系统的辅助模型多新息随机梯度算法的收敛性,包括了辅助模型随机梯度算法作为特殊情形^[13]。辅助模型遗忘梯度算法和辅助模型投影算法的收敛条件以及参数误差界分析,至今仍是辨识领域的重要难题。

仿真试验如下。

例2 考虑仿真对象(OE模型):

$$\begin{cases} y(t) = \frac{B(z)}{A(z)}u(t) + v(t), \\ A(z) = 1 + a_1z^{-1} + a_2z^{-2} = 1 + 0.412z^{-1} + 0.309z^{-2}, \\ B(z) = b_1z^{-1} + b_2z^{-2} = 0.6804z^{-1} + 0.6303z^{-2}, \\ \theta = [a_1 \ a_2 \ b_1 \ b_2]^T = [0.412 \ 0.309 \ 0.6804 \ 0.6303]^T. \end{cases}$$

仿真条件同例1。考虑3种情形:噪声方差与信噪比分别为 $\sigma^2 = 0.00^2$, $\delta_{ns} = 0.00\%$; $\sigma^2 = 0.10^2$, $\delta_{ns} = 11.75\%$; $\sigma^2 = 0.50^2$, $\delta_{ns} = 58.75\%$ 。用AM-SG算法和AM-FG算法估计这个系统的参数,不同噪声方差或信噪比下AM-SG参数估计及其误差如表3所示,AM-SG参数估计误差 $\delta: = \|\hat{\theta}(t) - \theta\| / \|\theta\|$ 随 t 变化曲线如图5所示。

不同遗忘因子下,AM-SG($\lambda = 1$)算法和AM-FG算法($\lambda = 0.95$, $\lambda = 0.99$)参数估计及其误差如表4—6所示;参数估计误差 $\delta: = \|\hat{\theta}(t) - \theta\| /$

$\|\theta\|$ 随 t 变化曲线如图 6 所示; 表 6 是变遗忘因子时的 AM-FG 参数估计.

由表 3—6, 图 5—6 可知: AM-SG 算法收敛速度慢, 参数估计变化平稳, 但参数估计精度低; AM-FG 算法具有较快的收敛速度, 遗忘因子 λ 越小, 算法收敛速度越快, 但参数估计波动越大, λ 越大, 算法的跟踪速度越慢, 但参数估计越平稳. 一种折衷方案是在算法起动的开始阶段选择较小 λ 值, 随后选择较大 λ 值, 并使 λ 逐渐趋于 1, 这样既可以保证算法有一定的收敛速度, 又可减小参数估计波动的方差, 从而得到较高精度的参数估计. 例如, 表 6 中, 当 $t <$

1 330 时, 取 $\lambda = 0.95$; 当 $t \geq 1 330$ 时, 取 $\lambda = 0.99$. 而 AM-SG 算法的收敛速度很慢, 几乎没有跟踪(时变)参数的能力, 可见引入遗忘因子的重要性. 从表 4 和图 6 可知, AM-FG 算法通过选择适当的遗忘因子, 可以获得满意的辨识效果.

把下列程序写到 AM_FG1.m 文件中, 当 $\sigma = 0.10$ 或 $\sigma = 0.50$ 时, 当 $SW = 1, FF = 1$ 时, 依次运行当 $\sigma = 0, 0.1, 0.50$; 当 $SW = 2, \sigma = 0.50$ 时, 依次运行 $FF = 1, 0.99, 0.95$, 可得到上述例子的仿真结果(参数估计表和误差曲线图). 适当修改一下, 可得到变遗忘因子的仿真结果.

表 3 不同噪信比下例 2 模型参数的 AM-SG 估计
Table 3 The AM-SG estimates and errors with different δ_{ns}

δ_{ns}	t	a_1	a_2	b_1	b_2	$\delta/\%$
00.00%	100	0.071 46	0.228 29	0.604 33	0.341 06	43.393 78
	200	0.083 95	0.249 89	0.625 54	0.351 62	41.280 58
	500	0.097 14	0.272 56	0.639 36	0.364 03	39.211 87
	1 000	0.106 59	0.287 47	0.648 93	0.376 29	37.616 76
	2 000	0.116 94	0.299 54	0.656 77	0.385 62	36.211 28
	3 000	0.121 99	0.305 21	0.659 84	0.391 24	35.482 35
11.75%	100	0.048 15	0.228 00	0.626 61	0.319 80	46.010 49
	200	0.060 50	0.250 05	0.644 40	0.329 79	44.074 46
	500	0.075 79	0.274 35	0.653 34	0.338 77	42.150 59
	1 000	0.083 07	0.290 09	0.659 78	0.354 25	40.563 15
	2 000	0.092 52	0.302 56	0.665 84	0.364 93	39.176 99
	3 000	0.097 45	0.308 27	0.667 81	0.371 01	38.444 00
58.75%	100	-0.032 14	0.208 22	0.722 10	0.23215	57.157 36
	200	-0.021 03	0.233 39	0.725 72	0.240 45	55.548 25
	500	0.001 46	0.265 96	0.714 97	0.237 26	53.826 39
	1 000	-0.001 51	0.286 36	0.708 13	0.265 26	52.103 29
	2 000	0.003 04	0.302 06	0.706 21	0.281 68	50.717 38
	3 000	0.006 86	0.308 82	0.703 36	0.289 71	49.938 27
真值		0.412 00	0.309 00	0.680 40	0.630 30	

表 4 例 2 模型参数的 AM-SG 估计 ($\sigma^2 = 0.50^2, \delta_{ns} = 58.75\%$)
Table 4 The AM-SG estimates and errors ($\sigma^2 = 0.50^2, \delta_{ns} = 58.75\%$)

t	a_1	a_2	b_1	b_2	$\delta/\%$
100	-0.032 14	0.208 22	0.722 10	0.232 15	57.157 36
200	-0.021 03	0.233 39	0.725 72	0.240 45	55.548 25
500	0.001 46	0.265 96	0.714 97	0.237 26	53.826 39
1 000	-0.001 51	0.286 36	0.708 13	0.265 26	52.103 29
1 500	0.001 46	0.294 52	0.704 23	0.274 16	51.297 11
2 000	0.003 04	0.302 06	0.706 21	0.281 68	50.717 38
2 500	0.005 86	0.306 20	0.704 02	0.285 70	50.257 45
3 000	0.006 86	0.308 82	0.703 36	0.289 71	49.938 27
真值	0.412 00	0.309 00	0.680 40	0.630 30	

表5 不同遗忘因子下例2模型参数的AM-FG估计($\sigma^2 = 0.50^2$ $\delta_{ns} = 58.75\%$)

Table 5 The AM-FG estimates and errors ($\sigma^2 = 0.50^2$ $\delta_{ns} = 58.75\%$)

λ	t	a_1	a_2	b_1	b_2	$\delta/\%$
0.99	100	-0.032 02	0.226 20	0.722 37	0.242 94	56.228 03
	200	-0.005 81	0.266 47	0.731 51	0.254 51	53.339 81
	500	0.052 00	0.333 56	0.713 32	0.265 18	48.487 45
	1 000	0.078 48	0.363 28	0.688 66	0.388 67	39.165 51
	1 500	0.158 90	0.338 64	0.673 69	0.413 36	31.552 82
	2 000	0.206 43	0.340 02	0.711 62	0.452 84	25.932 33
	2 500	0.260 16	0.341 18	0.682 90	0.479 98	20.368 30
	3 000	0.281 20	0.331 92	0.696 08	0.527 66	15.889 63
0.95	100	-0.021 27	0.294 12	0.714 26	0.287 93	52.169 38
	200	0.088 77	0.339 94	0.744 79	0.306 42	43.654 21
	500	0.234 40	0.351 27	0.758 60	0.404 62	28.337 33
	1 000	0.338 02	0.367 59	0.680 35	0.623 44	8.918 91
	1 500	0.450 95	0.256 03	0.636 71	0.566 75	9.552 72
	2 000	0.406 43	0.281 74	0.723 37	0.610 95	5.158 26
	2 500	0.458 89	0.347 69	0.672 61	0.581 75	7.370 06
	3 000	0.385 70	0.348 31	0.700 10	0.666 77	5.927 78
真值		0.412 00	0.309 00	0.680 40	0.630 30	

表6 变遗忘因子下例2模型参数的AM-FG估计($\sigma^2 = 0.50^2$ $\delta_{ns} = 58.75\%$) (当 $t < 1 330$ 时 $\lambda = 0.95$; 当 $t \geq 1 330$ 时 $\lambda = 0.99$)

Table 6 The AM-FG estimates and errors ($\sigma^2 = 0.50^2$ $\delta_{ns} = 58.75\%$) ($\lambda = 0.95$ for $t < 1 330$; $\lambda = 0.99$ for $t \geq 1 330$)

t	a_1	a_2	b_1	b_2	$\delta/\%$
100	-0.021 27	0.294 12	0.714 26	0.287 93	52.169 38
200	0.088 77	0.339 94	0.744 79	0.306 42	43.654 21
500	0.234 40	0.351 27	0.758 60	0.404 62	28.337 33
1 000	0.338 02	0.367 59	0.680 35	0.623 44	8.918 91
1 500	0.401 32	0.280 74	0.676 13	0.620 05	3.033 93
2 000	0.407 09	0.293 13	0.708 17	0.612 58	3.478 01
2 500	0.419 47	0.309 56	0.681 57	0.607 77	2.240 70
3 000	0.410 93	0.307 71	0.692 10	0.628 55	1.125 93
真值	0.412 00	0.309 00	0.680 40	0.630 30	

```

1 %-----*
2 % Filename: AM_FG1.m for the output error models: *
3 % y(t) = B(z) / A(z) u(t) + v(t) *
4 % The AM - FG algorithm *
5 % The forgetting factor FF = \lambda *
6 % The noise variance sigma^2 *
7 % When FF = 1 ,sigma = 0 ., 10 and 0.50 *
8 % When sigma = 0.50 ,FF = 1 ., 99 ., 95 *
9 %-----*
10 clear; format short g
11 M = The AM - FG algorithm for the output error model'
12 SW = 2; % SW = 1: Plot for fixed FF ,sigma = 0 ., 1
    and 0.5
13 % SW = 2: Plot for fixed sigma ,FF = 1 ., 99 ., 95
14 FF = 0.95; % The forgetting factor
15 sigma = 0.5; % sigma = 0.5 FF = 0.9 ,0.95 ,0.99 for
    AM-FG other case
16 PlotLength = 3000; length1 = PlotLength + 100;
17 na = 2; nb = 2; n = na + nb;
18 a = [1 0.412 0.309]; b = [0 0.6804 0.6303];
19 par0 = [a(2:na+1) b(2:nb+1)]; % The
    parameter vector
20 p0 = 1e6; r = 1;
21 par1 = ones(n,1) / p0; % The parameter estimation vector
22 % Compute the noise - to - signal ratio
23 sy = f_integral(a, b); sv = 1;
24 delta_ns = sqrt(sv/sy) * 100 * sigma;
25 [sy delta_ns]
    
```

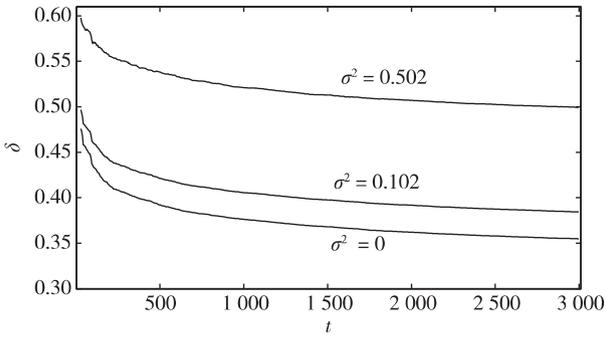


图5 不同噪声比下例2的AM-SG
估计误差 δ 随 t 的变化曲线

fig. 5 The AM-SG estimation errors δ versus t with different δ_{ns}

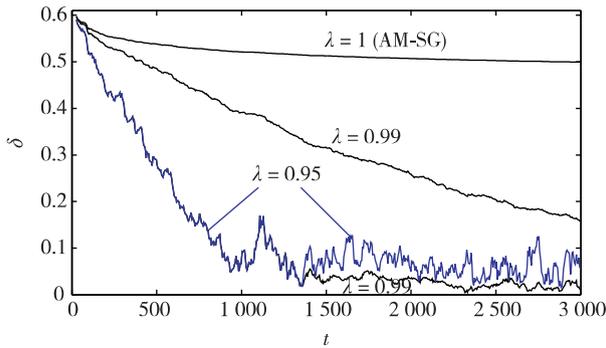


图6 不同遗忘因子下例2的 δ 随
 t 的变化曲线($\sigma^2 = 0.50^2$)

Fig. 6 The AM-FG estimation errors δ versus
 t with different λ ($\sigma^2 = 0.50^2$)

```

26 %——Generate the input - output data
27 rand( 'state' , rho ); randn( 'state' , rho );
28 u = ( rand( length1 , 1) - 0.5) * sqrt( 12); v = randn
( length1 , 1) * sigma;
29
30 y = ones( length1 , 1) / p0;
31 x = ones( length1 , 1) / p0; x1 = x;
32 %——Compute AM-FG estimates
33 jj = 0; j1 = 0;
34 for t = n: length1
35     jj = jj + 1;
36     if jj == 1330
37         FF = 0.99; % Appropriate modification: add a %
38     end
39     x( t) = - par0( 1: na) * x( t - 1: - 1: t - na) +
par0( na + 1: n) * u( t - 1: - 1: t - nb);
40     y( t) = x( t) + v( t);
41     varphi = [ - x1( t - 1: - 1: t - na); u( t - 1: - 1: t
- nb) ];

```

```

42     r = FF* r + varphi* varphi;
43     par1 = par1 + varphi/r* ( y( t) - varphi* par1) ;
44     x1( t) = varphi* par1;
45     delta = norm( par1 - par0) / norm( par0) ;
46
47     ls( jj ; ) = [jj par1' delta];
48     if ( jj == 100) | ( jj == 200) | ( jj == 500) | mod
( jj 500) == 0
49         j1 = j1 + 1;
50         ls_100( j1 ; ) = [jj par1' delta* 100];
51     end
52     if jj == PlotLength
53         break
54     end
55 end
56 ls_100( j1 + 1 ; ) = [0 par0' rho];
57 fprintf( ' FF = %4. 2f \ $ \ \sigma^2 = %4. 2f $ \ $ \ del-
ta_{ \ \ns} = %6. 2f % s' , ...
58         FF , sigma , delta_ns , '\%$ ) )
59 fprintf( ' \n %s \n' , ...
60         '$ t $ $ a_1 $ $ a_2 $ $ b_1 $ $ b_2
$ $ \delta \ ( \% ) \ $ \ \hline' );
61 fprintf( '%5d %10. 5f %10. 5f %10. 5f %10. 5f %10. 5f
\ \ \n' , ls_100' );
62
63 figure( 1) ; jk = ( 28: 10: PlotLength - 1) ' ;
64 plot( ls( jk , 1) , ls( jk , n + 2) )
65 figure( 2) ;
66 if SW == 1
67     if sigma == 0
68         data1 = [ls( : , 1) , ls( : , n + 2) ];
69         save data1 data1
70     elseif sigma == 0. 10
71         load data1
72         data2 = [data1 , ls( : , n + 2) ];
73         save data2 data2
74         z0 = data2;
75         jk = ( 28: 5: PlotLength - 1) ' ;
76         plot( z0( jk , 1) , z0( jk , 2) , 'k' , z0( jk , 1) , z0( jk ,
3) , 'b' )
77         axis( [0 , PlotLength , 0 , 0. 61] )
78     else
79         load data2
80         z0 = [data2 , ls( : , n + 2) ];
81         jk = ( 28: 5: PlotLength - 1) ;
82         plot( z0( jk , 1) , z0( jk , 2) , 'k' , z0( jk , 1) , z0( jk ,
3) , 'b' , ...
83         z0( jk , 1) , z0( jk , 4) , 'b' )

```

```

84     text( 1500 0.525 , '\sigma^2 = 0.50^2)
85     text( 1500 0.41 , '\sigma^2 = 0.10^2)
86     text( 1500 0.353 , '\sigma^2 = 0)
87     end
88 else % SW = = 2
89     if FF = = 1
90         data1 = [ls( : ,1) ls( : ,n+2) ];
91         save data1 data1
92     elseif FF = = 0.99001 % Appropriate modification
93         load data1
94         data2 = [data1 ls( : ,n+2) ];
95         save data2 data2
96     elseif FF = = 0.95001 % Appropriate modification
97         load data2
98         data3 = [data2 ls( : ,n+2) ];
99         save data3 data3
100    else
101        load data3
102        z0 = [data3 ls( : ,n+2) ];
103        jk = ( 28:5:PlotLength - 1) ;
104        figure ( 2)
105        plot( z0( jk ,1) z0( jk ,2) , 'k' z0( jk ,1) z0( jk ,
106            3) , 'k' , ...
107            z0( jk ,1) z0( jk ,5) , 'k' z0( jk ,1) z0
108            ( jk ,4) , 'b' )
109        axis( [0 ,PlotLength 0 0.61] )
110        text( 1500 ,0.535 , '\lambda = 1 ( AM -
111            SG ) ' )
112        text( 1500 0.34 , '\lambda = 0.99' )
113        text( 1100 0.25 , '\lambda = 0.95' )
114        line( [803 ,1100] , [0.138 0.23] )
115        line( [1648 ,1350] , [0.1218 0.23] )
116        text( 1550 0.015 , '\lambda = 0.99' )
117    end
118 end
119 xlabel( '\it t' ) ; ylabel( '\it \delta' )

```

2.2 线性输出误差滑动平均系统(OEMA)

输出误差类系统包括基本的输出误差(OE)系统、输出误差滑动平均(OEMA)系统、输出误差自回归(OEAR)系统和输出误差自回归滑动平均(OE-ARMA = Box-Jenkins)系统. 这些模型的辅助模型辨识方法在文献 [18] 中有详细研究, 这里进行简单讨论.

随机梯度算法、最小二乘算法、多新息辨识方法等都可以用于输出误差类模型的辨识. 例如, 对于 OEMA 系统, 我们可以得到辅助模型增广随机梯度

算法、辅助模型递推增广最小二乘算法、辅助模型多新息增广随机梯度算法、辅助模型多新息增广最小二乘算法等.

输出误差类的输出误差滑动平均模型(Output Error Moving Average model ,OEMA) 如图 7 所示, 其输入输出表达如下:

$$y(t) = \frac{B(z)}{A(z)}u(t) + D(z)v(t). \quad (47)$$

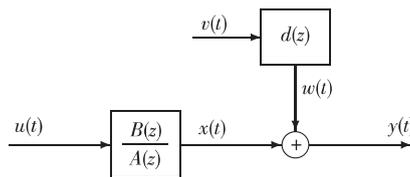


图 7 输出误差滑动平均系统(OEMA)

Fig.7 The output error moving average system

定义系统真实输出 $x(t)$ 和噪声模型输出 $w(t)$ 分别为

$$x(t) = \frac{B(z)}{A(z)}u(t), \quad (48)$$

$$w(t) = D(z)v(t). \quad (49)$$

置系统参数向量 θ 和信息向量 $\varphi(t)$ 如下:

$$\theta = [a_1 \ a_2 \ \dots \ a_{n_a} \ b_1 \ b_2 \ \dots \ b_{n_b} \ d_1 \ d_2 \ \dots \ d_{n_d}]^T \in \mathbf{R}^{n_a+n_b+n_d},$$

$$\varphi(t) = [-x(t-1) \ -x(t-2) \ \dots \ -x(t-n_a) \ u(t-1) \ \mu(t-2) \ \dots \ \mu(t-n_b) \ v(t-1) \ v(t-2) \ \dots \ v(t-n_d)]^T \in \mathbf{R}^{n_a+n_b+n_d}.$$

借助于上述定义, 式(48) — (49) 可以写为

$$\begin{aligned}
 x(t) &= [1 - A(z)]x(t) + B(z)u(t) = \\
 &= -\sum_{i=1}^{n_a} a_i x(t-i) + \sum_{i=1}^{n_b} b_i u(t-i), \quad (50) \\
 w(t) &= \sum_{i=1}^{n_d} d_i v(t-i) + v(t).
 \end{aligned}$$

则 OEMA 系统(47) 可以写为下列辨识模型:

$$\begin{aligned}
 y(t) &= x(t) + w(t) = \\
 &= -\sum_{i=1}^{n_a} a_i x(t-i) + \sum_{i=1}^{n_b} b_i u(t-i) + \sum_{i=1}^{n_d} d_i v(t-i) + v(t) = \\
 &= \varphi^T(t) \theta + v(t). \quad (51)
 \end{aligned}$$

因为式(51) 辨识模型信息向量 $\varphi(t)$ 中不仅包含了不可测真实输出 $x(t-i)$, 而且包含了噪声项 $v(t-i)$, 故这是一个伪线性回归模型 III.

辨识 OEMA 系统的辅助模型的递推增广最小二乘算法(Auxiliary Model based Recursive Extended

Least Squares algorithm ,AM-RELS) 如下:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t) [y(t) - \hat{\varphi}^T(t) \hat{\theta}(t-1)], \quad (52)$$

$$L(t) = P(t-1) \hat{\varphi} \hat{\varphi}^T(t) P(t-1) \hat{\varphi} \hat{\varphi}^T(t)]^{-1}, \quad (53)$$

$$P(t) = [I - L(t) \hat{\varphi} \hat{\varphi}^T(t)] P(t-1) \quad P(0) = p_0 I, \quad (54)$$

$$\hat{\varphi} \hat{\varphi}^T(t) = [\hat{\varphi}_s^T(t) \quad \hat{\rho}(t-1) \quad \hat{\rho}(t-2) \quad ; \cdots \quad \hat{\rho}(t-n_d)]^T, \quad (55)$$

$$\hat{\varphi}_s(t) = [-\hat{x}(t-1) \quad , -\hat{x}(t-2) \quad ; \cdots \quad , -\hat{x}(t-n_a) \quad , \quad u(t-1) \quad \mu(t-2) \quad ; \cdots \quad \mu(t-n_b)]^T, \quad (56)$$

$$\hat{x}(t) = \hat{\varphi}_s^T(t) \hat{\theta}_s(t), \quad (57)$$

$$\hat{v}(t) = y(t) - \hat{\varphi}^T(t) \hat{\theta}(t), \quad (58)$$

$$\hat{\theta}(t) = [\hat{\theta}_s^T(t) \quad \hat{a}_1(t) \quad \hat{a}_2(t) \quad ; \cdots \quad \hat{a}_{n_d}(t)]^T, \quad (59)$$

$$\hat{\theta}_s(t) = [\hat{a}_1(t) \quad \hat{a}_2(t) \quad ; \cdots \quad \hat{a}_{n_a}(t) \quad \hat{b}_1(t) \quad , \quad \hat{b}_2(t) \quad ; \cdots \quad \hat{b}_{n_b}(t)]^T. \quad (60)$$

对于 OEMA 系统,王冬青^[26]提出了基于辅助模型的递推增广最小二乘辨识方法(AM-RELS),还给出了通过模型变换的递推增广最小二乘算法,AM-RELS 方法原理简单、计算量小,可以给出高精度参数估计,且能够用于在线辨识;本文作者等提出了基于辅助模型的多新息增广随机梯度辨识算法(AM-MI-ESG)^[24]和基于辅助模型的多新息增广最小二乘辨识算法(AM-MI-ELS)和变间隔 AM-MI-ELS 算法^[35]还提出了输出误差滑动平均模型的基于梯度和基于最小二乘的迭代辨识方法^[19].

2.3 线性输出误差自回归系统(OEAR)

输出误差自回归模型(Output Error Autoregressive model ,OEAR) 如图 8 所示,其输入输出表达如下:

$$y(t) = \frac{B(z)}{A(z)} u(t) + \frac{1}{C(z)} v(t). \quad (61)$$

对于 OEAR 系统,许多算法能够估计其参数,如辅助模型广义随机梯度算法、辅助模型递推广义最小二乘算法、辅助模型多新息广义随机梯度算法、辅助模型多新息广义最小二乘算法等.

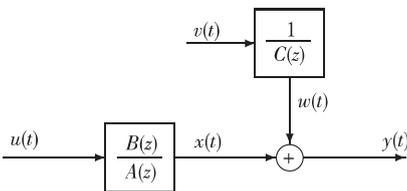


图 8 输出误差自回归系统(OEAR)

Fig. 8 The output error autoregressive system

定义系统真实输出 $x(t)$ 和噪声模型输出 $w(t)$ 分别为

$$x(t) := \frac{B(z)}{A(z)} u(t), \quad (62)$$

$$w(t) := \frac{1}{C(z)} v(t). \quad (63)$$

置系统参数向量 θ 和信息向量 $\varphi(t)$ 如下:

$$\theta := [a_1 \quad a_2 \quad ; \cdots \quad a_{n_a} \quad b_1 \quad b_2 \quad ; \cdots \quad b_{n_b} \quad c_1 \quad c_2 \quad ; \cdots \quad c_{n_c}]^T \in \mathbf{R}^{n_a+n_b+n_c},$$

$$\varphi(t) := [-x(t-1) \quad , -x(t-2) \quad ; \cdots \quad , -x(t-n_a) \quad , \quad u(t-1) \quad \mu(t-2) \quad ; \cdots \quad \mu(t-n_b) \quad , -w(t-1) \quad , \quad -w(t-2) \quad ; \cdots \quad , -w(t-n_c)]^T \in \mathbf{R}^{n_a+n_b+n_c}.$$

借助于上述定义,式(62) — (63) 可以写为

$$\begin{aligned} x(t) &= [1 - A(z)]x(t) + B(z)u(t) = \\ &= -\sum_{i=1}^{n_a} a_i x(t-i) + \sum_{i=1}^{n_b} b_i u(t-i), \\ w(t) &= [1 - C(z)]w(t) + v(t) = \\ &= -\sum_{i=1}^{n_c} c_i w(t-i) + v(t). \end{aligned}$$

则 OEAR 系统(61) 可以写为下列辨识模型:

$$\begin{aligned} y(t) &= x(t) + w(t) = \\ &= -\sum_{i=1}^{n_a} a_i x(t-i) + \sum_{i=1}^{n_b} b_i u(t-i) - \sum_{i=1}^{n_d} c_i w(t-i) + v(t) = \\ &= \varphi^T(t) \theta + v(t). \end{aligned} \quad (64)$$

因为式(64) 辨识模型信息向量 $\varphi(t)$ 中不仅包含了不可测真实输出 $x(t-i)$,而且包含了噪声项 $v(t-i)$,故这是一个伪线性回归模型 III.

辨识 OEAR 系统的辅助模型的递推广义最小二乘算法(Auxiliary Model based Recursive Generalized Least Squares algorithm ,AM-RGLS) 如下:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t) [y(t) - \hat{\varphi}^T(t) \hat{\theta}(t-1)], \quad (65)$$

$$L(t) = P(t-1) \hat{\varphi} \hat{\varphi}^T(t) P(t-1) \hat{\varphi} \hat{\varphi}^T(t)]^{-1}, \quad (66)$$

$$P(t) = [I - L(t) \hat{\varphi} \hat{\varphi}^T(t)] P(t-1) \quad P(0) = p_0 I, \quad (67)$$

$$\hat{\varphi} \hat{\varphi}^T(t) = [\hat{\varphi}_s^T(t) \quad , -\hat{\omega}(t-1) \quad , -\hat{\omega}(t-2) \quad ; \cdots \quad , \quad -\hat{\omega}(t-n_d)]^T, \quad (68)$$

$$\hat{\varphi}_s(t) = [-\hat{x}(t-1) \quad , -\hat{x}(t-2) \quad ; \cdots \quad , -\hat{x}(t-n_a) \quad , \quad u(t-1) \quad \mu(t-2) \quad ; \cdots \quad \mu(t-n_b)]^T, \quad (69)$$

$$\hat{x}(t) = \hat{\varphi}_s^T(t) \hat{\theta}_s(t), \quad (70)$$

$$\hat{\omega}(t) = y(t) - \hat{x}(t) = y(t) - \hat{\varphi}_s^T(t) \hat{\theta}_s(t), \quad (71)$$

$$\hat{\theta}(t) = [\hat{\theta}_s^T(t) \quad \hat{c}_1(t) \quad \hat{c}_2(t) \quad ; \cdots \quad \hat{c}_{n_d}(t)]^T, \quad (72)$$

$$\hat{\theta}(t) = [\hat{a}_1(t) \quad \hat{a}_2 \quad ; \cdots \quad \hat{a}_{n_a}(t) \quad \hat{b}_1(t) \quad \hat{b}_2 \quad ; \cdots \quad \hat{b}_{n_b}(t)]^T. \quad (73)$$

2.4 线性 Box-Jenkins 系统(BJ)

Box-Jenkins 模型(BJ model) 如图 9 所示,其输入输出表达如下:

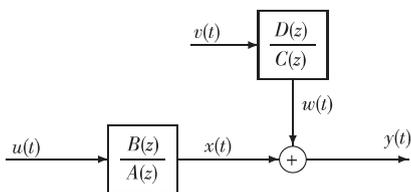


图9 Box-Jenkins 系统(BJ)

Fig.9 The Box-Jenkins system

$$y(t) = \frac{B(z)}{A(z)}u(t) + \frac{D(z)}{C(z)}v(t) \quad (74)$$

辨识这个 Box-Jenkins 系统的方法有辅助模型广义增广随机梯度算法、辅助模型递推广义增广最小二乘算法、辅助模型多新息广义增广随机梯度算法、辅助模型多新息广义增广最小二乘算法等。

定义系统真实输出 $x(t)$ 和噪声模型输出 $w(t)$ 分别为

$$x(t) := \frac{B(z)}{A(z)}u(t) \quad (75)$$

$$w(t) := \frac{D(z)}{C(z)}v(t) \quad (76)$$

置系统参数向量 θ 和信息向量 $\varphi(t)$ 如下:

$$\theta := [a_1 \ a_2 \ ; \dots \ a_{n_a} \ b_1 \ b_2 \ ; \dots \ b_{n_b} \ c_1 \ c_2 \ ; \dots \ c_{n_c}]^T \in \mathbf{R}^{n_a+n_b+n_c}$$

$$\varphi(t) := [-x(t-1) \ , -x(t-2) \ ; \dots \ , -x(t-n_a) \ , \\ u(t-1) \ \mu(t-2) \ ; \dots \ \mu(t-n_b) \ , \\ -w(t-1) \ , -w(t-2) \ ; \dots \ , -w(t-n_c) \ , \\ v(t-1) \ \nu(t-2) \ ; \dots \ \nu(t-n_d)]^T \in \mathbf{R}^{n_a+n_b+n_c+n_d}$$

借助于上述定义,式(75)-(76)可以写为

$$x(t) = [1 - A(z)]x(t) + B(z)u(t) = \\ - \sum_{i=1}^{n_a} a_i x(t-i) + \sum_{i=1}^{n_b} b_i u(t-i) \ , \\ w(t) = [1 - C(z)]w(t) + D(z)v(t) = \\ - \sum_{i=1}^{n_c} c_i w(t-i) + \sum_{i=1}^{n_d} d_i v(t-i) + v(t) \ .$$

则 Box-Jenkins 系统(74)可以写为下列辨识模型:

$$y(t) = x(t) + w(t) = \\ \sum_{i=1}^{n_a} a_i x(t-i) + \sum_{i=1}^{n_b} b_i u(t-i) - \sum_{i=1}^{n_c} c_i w(t-i) + \\ \sum_{i=1}^{n_d} d_i v(t-i) + v(t) = \varphi^T(t) \theta + v(t) \ . \quad (77)$$

因为式(77)辨识模型信息向量 $\varphi(t)$ 中不仅包含了不可测真实输出 $x(t-i)$,而且包含了相关噪声项 $w(t-i)$,故这是一个伪线性回归模型Ⅲ。

辨识 Box-Jenkins 系统的辅助模型的递推广义最小二乘算法(Auxiliary Model based Recursive Generalized Least Squares algorithm, AM-RGLS) ,

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t) [y(t) - \hat{\varphi}^T(t) \hat{\theta}(t-1)] \ , \quad (78)$$

$$L(t) = \frac{P(t-1) \hat{\varphi}(t)}{1 + \hat{\varphi}^T(t) P(t-1) \hat{\varphi}(t)} \ , \quad (79)$$

$$P(t) = [I - L(t) \hat{\varphi}^T(t)] P(t-1) \quad P(0) = p_0 I \ , \quad (80)$$

$$\hat{\varphi}(t) = [\hat{\varphi}_s^T(t) \ , -\hat{w}(t-1) \ , -\hat{w}(t-2) \ ; \dots \ , -\hat{w}(t-n_c) \ , \\ \hat{v}(t-1) \ \hat{p}(t-2) \ ; \dots \ \hat{p}(t-n_d)]^T \ , \quad (81)$$

$$\hat{\varphi}_s(t) = [-\hat{x}(t-1) \ , -\hat{x}(t-2) \ ; \dots \ , -\hat{x}(t-n_a) \ , \\ u(t-1) \ \mu(t-2) \ ; \dots \ \mu(t-n_b)]^T \ , \quad (82)$$

$$\hat{x}(t) = \hat{\varphi}_s^T(t) \hat{\theta}_s(t) \ , \quad (83)$$

$$\hat{w}(t) = y(t) - \hat{x}(t) \ , \quad (84)$$

$$\hat{v}(t) = y(t) - \hat{\varphi}^T(t) \hat{\theta}(t) \ , \quad (85)$$

$$\hat{\theta}(t) = [\hat{\theta}_s^T(t) \ \hat{c}_1(t) \ \hat{c}_2(t) \ ; \dots \ \hat{c}_{n_c}(t) \ \hat{d}_1(t) \ , \\ \hat{d}_2(t) \ ; \dots \ \hat{d}_{n_d}(t)]^T \ , \quad (86)$$

$$\hat{\theta}_s(t) = [\hat{a}_1(t) \ \hat{a}_2(t) \ ; \dots \ \hat{a}_{n_a}(t) \ \hat{b}_1(t) \ , \\ \hat{b}_2(t) \ ; \dots \ \hat{b}_{n_b}(t)]^T \ . \quad (87)$$

对于 Box-Jenkins 系统,王冬青和本文作者提出了基于辅助模型多新息广义增广随机梯度辨识方法(AM-MI-GESG)^[29];刘艳君^[36]等提出了基于最小二乘的迭代辨识方法;王冬青^[37]等提出了基于梯度的迭代参数估计方法;谢莉^[38]等提出了非均匀采样数据系统的辅助模型多新息广义增广随机梯度辨识方法。

CARAR 系统的 RGLS 算法、CARARMA 系统的 RGELS 算法、本节讨论的 OEMA 系统的 AM-RELS 算法、OEAR 系统的 AM-RGLS 算法、Box-Jenkins 系统的 AM-RGELS 算法的收敛性分析(参数估计的一致收敛性、估计误差的有界收敛性)都是控制领域中国际著名的研究难题。

3 输入非线性方程误差类系统(IN-EET)

输入非线性方程误差类(IN-EET)系统包括输入非线性有限性脉冲响应(IN-FIR)系统、输入非线性受控自回归(IN-CAR)系统、输入非线性受控自回归滑动平均(IN-CARMA)系统、输入非线性受控自回归自回归(IN-CARAR)系统、输入非线性受控自回归自回归滑动平均(IN-CARARMA)系统。

3.1 输入非线性有限脉冲响应系统(IN-FIR)

输入非线性有限脉冲响应(Input Nonlinear Finite Impulse Response, IN-FIR)系统如图10所示。它

是由一个无记忆非线性环节 $f(\cdot)$ 串连一个线性动态 FIR 模型 $G(z)$ 构成的. 非线性部分是系数为 (c_1, c_2, \dots, c_m) 的已知非线性基 $f = (f_1, f_2, \dots, f_m)$ 的线性函数:

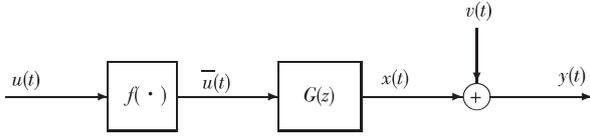


图 10 输入非线性有限脉冲响应系统 (IN-FIR)

Fig. 10 The input nonlinear finite impulse response (IN-FIR) system

$$\begin{aligned} \bar{u}(t) &= f(u(t)) = \\ c_1 f_1(u(t)) &+ c_2 f_2(u(t)) + \dots + c_m f_m(u(t)) = \\ \sum_{j=1}^m c_j f_j(u(t)) &= f(u(t)) \mathbf{c}, \end{aligned} \quad (88)$$

其中 $f(u(t)) := [f_1(u(t)), f_2(u(t)), \dots, f_m(u(t))] \in \mathbf{R}^{1 \times m}$ 是基函数构成的行向量 $\mathbf{c} := [c_1, c_2, \dots, c_m]^T \in \mathbf{R}^m$ 是非线性部分的参数向量.

线性部分是一个参数为 $(b_0, b_1, b_2, \dots, b_n)$ 的 FIR 模型, 其传递函数为

$$G(z) = b_0 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_n z^{-n} = \sum_{i=0}^n b_i z^{-i}. \quad (89)$$

因此, 线性部分可以表示为

$$\begin{aligned} y(t) &= x(t) + v(t) = G(z) \bar{u}(t) + v(t) = \\ (b_0 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_n z^{-n}) \bar{u}(t) &+ v(t) = \\ b_0 \bar{u}(t) + b_1 \bar{u}(t-1) + b_2 \bar{u}(t-2) + \dots + b_n \bar{u}(t-n) &+ v(t) = \\ b_0 f(u(t)) \mathbf{c} + b_1 f(u(t-1)) \mathbf{c} + b_2 f(u(t-2)) \mathbf{c} + \dots + \\ b_n f(u(t-n)) \mathbf{c} + v(t) &= \mathbf{b}^T \mathbf{F}(t) \mathbf{c} + v(t) \end{aligned} \quad (90)$$

其中

$$\mathbf{b} := [b_0, b_1, b_2, \dots, b_n]^T \in \mathbf{R}^{n+1}, \quad (91)$$

$$\mathbf{F}(t) := \begin{bmatrix} f(u(t)) \\ f(u(t-1)) \\ \vdots \\ f(u(t-n)) \end{bmatrix} =$$

$$\begin{bmatrix} f_1(u(t)) & f_2(u(t)) & \dots & f_m(u(t)) \\ f_1(u(t-1)) & f_2(u(t-1)) & \dots & f_m(u(t-1)) \\ f_1(u(t-2)) & f_2(u(t-2)) & \dots & f_m(u(t-2)) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(u(t-n)) & f_2(u(t-n)) & \dots & f_m(u(t-n)) \end{bmatrix} \in \mathbf{R}^{(n+1) \times m}, \quad (92)$$

式(90)是一个双线性参数模型 (bilinear-parameter model), 式中出现了 2 个参数集的乘积项, 故该系统不是参数可辨识的.

为了辨识这个系统, 需要规范化系统参数. 基本的规范化方法有: 1) 固定 b_i 中的一个, 或者固定 c_j 中的一个; 2) 设 $(b_0, b_1, b_2, \dots, b_n)$ 或 (c_1, c_2, \dots, c_m) 的模为 1, 即 $\|\mathbf{b}\|^2 = b_0^2 + b_1^2 + b_2^2 + \dots + b_n^2 = 1$, 或 $\|\mathbf{c}\|^2 = c_1^2 + c_2^2 + \dots + c_m^2 = 1$; 3) 设线性子系统的增益为 1, 即 $G(1) = b_0 + b_1 + b_2 + \dots + b_n = 1$ 或非线性函数的系数和为 1, 即 $c_1 + c_2 + \dots + c_m = 1$.

为方便起见, 本文假设 $G(z)$ 的第 1 个系数 $b_0 = 1$. 这样式(90)可以写为

$$y(t) = \bar{u}(t) + b_1 \bar{u}(t-1) + b_2 \bar{u}(t-2) + \dots + b_n \bar{u}(t-n) + v(t).$$

将式(88)代入上式右边的第 1 项可得

$$y(t) = c_1 f_1(u(t)) + c_2 f_2(u(t)) + \dots + c_m f_m(u(t)) + b_1 \bar{u}(t-1) + b_2 \bar{u}(t-2) + \dots + b_n \bar{u}(t-n) + v(t) = \boldsymbol{\varphi}^T(t) \boldsymbol{\vartheta} + v(t), \quad (93)$$

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

$$\boldsymbol{\vartheta} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \\ c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix} \in \mathbf{R}^{n+m}; \quad \boldsymbol{\varphi}(t) = \begin{bmatrix} \bar{u}(t-1) \\ \bar{u}(t-2) \\ \vdots \\ \bar{u}(t-n) \\ f_1(u(t)) \\ f_2(u(t)) \\ \vdots \\ f_m(u(t)) \end{bmatrix} \in \mathbf{R}^{n+m}. \quad (94)$$

对于辨识模型(93) 极小化准则函数

$$J_5(\boldsymbol{\vartheta}) := \sum_{j=1}^t [y(j) - \boldsymbol{\varphi}^T(j) \boldsymbol{\vartheta}]^2$$

可得下列递推最小二乘算法:

$$\hat{\boldsymbol{\vartheta}}(t) = \hat{\boldsymbol{\vartheta}}(t-1) + \mathbf{L}(t) [y(t) - \boldsymbol{\varphi}^T(t) \hat{\boldsymbol{\vartheta}}(t-1)], \quad (95)$$

$$\mathbf{L}(t) = \mathbf{P}(t) \boldsymbol{\varphi}(t) = \frac{\mathbf{P}(t-1) \boldsymbol{\varphi}(t)}{1 + \boldsymbol{\varphi}^T(t) \mathbf{P}(t-1) \boldsymbol{\varphi}(t)}, \quad (96)$$

$$\mathbf{P}(t) = [\mathbf{I} - \mathbf{L}(t) \boldsymbol{\varphi}^T(t)] \mathbf{P}(t-1), \quad \mathbf{P}(0) = p_0 \mathbf{I}. \quad (97)$$

假设系统的输入输出数据为 $\{u(t), y(t)\}$, 非线性基函数 $f_j(u(t))$ 是已知的, 式(94)中 $\boldsymbol{\varphi}(t)$ 包含了未知中间变量 (非线性部分的输出) $\bar{u}(t-i)$, 使得上述算法无法实现. 解决的办法是借助于辅助模型辨识思想: 未知变量用辅助模型的输出代替或用其估计代替. 具体方法如下. 设 $\hat{\bar{u}}(t)$ 是辅助模型的输出 (即 $\bar{u}(t)$ 的估计), 用 $\hat{\bar{u}}(t-i)$ 代替式(94) $\boldsymbol{\varphi}(t)$ 中的未知项 $\bar{u}(t-i)$, 代替后的信息向量记作

$$\hat{\boldsymbol{\varphi}}(t) := [\hat{\bar{u}}(t-1), \hat{\bar{u}}(t-2), \dots, \hat{\bar{u}}(t-n)],$$

$$f_1(u(t)) \ f_2(u(t)) \] \cdots \ f_m(u(t)) \]^T =$$

$$[\hat{u}(t-1) \ \hat{u}(t-2) \ ; \cdots \ \hat{u}(t-n) \ \mathbf{f}(u(t)) \]^T \in$$

$$\mathbf{R}^{n+m}.$$

设 $\hat{\vartheta}$ 在时刻 t 的估计为

$$\hat{\vartheta}(t) := [\hat{b}_1(t) \ \hat{b}_2(t) \ ; \cdots \ \hat{b}_n(t) \ \hat{c}^T(t) \]^T \in \mathbf{R}^{n+m},$$

$$\hat{c}(t) := [\hat{c}_1(t) \ \hat{c}_2(t) \ ; \cdots \ \hat{c}_m(t) \]^T \in \mathbf{R}^m.$$

用 $\hat{c}_i(t)$ 代替式 (88) 中参数 c_i , 故辅助模型输出 $\hat{u}(t)$ 可通过下式计算,

$$\hat{u}(t) = \hat{c}_1(t)f_1(u(t)) + \hat{c}_2(t)f_2(u(t)) + \cdots +$$

$$\hat{c}_m(t)f_m(u(t)) =$$

$$\sum_{j=1}^m \hat{c}_j(t)f_j(u(t)) = \mathbf{f}(u(t)) \hat{c}(t).$$

上式即为计算未知变量 $\bar{u}(t)$ 的辅助模型. 用 $\hat{\varphi}(t)$ 代替式 (95) — (97) 中未知 $\varphi(t)$, 可以总结出估计输入非线性 FIR 系统参数向量 ϑ 的辅助模型递推最小二乘辨识算法 (AM-RLS):

$$\hat{\vartheta}(t) = \hat{\vartheta}(t-1) + L(t) [y(t) - \hat{\varphi}^T(t) \hat{\vartheta}(t-1)], \quad (98)$$

$$L(t) = P(t) \hat{\varphi}(t) = \frac{P(t-1) \hat{\varphi}(t)}{1 + \hat{\varphi}^T(t) P(t-1) \hat{\varphi}(t)}, \quad (99)$$

$$L(t) = [I - L(t) \hat{\varphi}^T(t)] P(t-1) P(0) = p_0 I, \quad (100)$$

$$\hat{\varphi}(t) = [\hat{u}(t-1) \ \hat{u}(t-2) \ ; \cdots \ \hat{u}(t-n) \ \mathbf{f}(u(t)) \]^T, \quad (101)$$

$$\hat{u}(t) = \hat{c}(t)f_1(u(t)) + \hat{c}(t)f_2(u(t)) + \cdots +$$

$$\hat{c}_m(t)f_m(u(t)) = \mathbf{f}(u(t)) \hat{c}(t), \quad (102)$$

$$\mathbf{f}(u(t)) = [f_1(u(t)) \ f_2(u(t)) \ ; \cdots \ f_m(u(t)) \], \quad (103)$$

$$\hat{\vartheta}(t) = [\hat{b}_1(t) \ \hat{b}_2(t) \ ; \cdots \ \hat{b}_n(t) \ \hat{c}^T(t) \]^T. \quad (104)$$

输入非线性 FIR 系统 AM-RLS 算法 (98) — (104) 的计算步骤如下.

1) 初始化: 令 $t=1$. 置初值 $P(0) = p_0 I$, $\hat{\vartheta}(0) = 1/p_0$, $p_0 = 10^6$, $\hat{u}(-i) = 1/p_0$, $i = 0, 1, 2, \dots, n-1$. 给定基函数 $f_j(\cdot)$.

2) 收集数据 $u(t)$ 和 $y(t)$, 用式 (103) 构造基函数行向量 $\mathbf{f}(u(t))$, 用式 (101) 构造信息向量 $\hat{\varphi}(t)$.

3) 用式 (99) 计算增益向量 $L(t)$, 用式 (100) 计算协方差阵 $P(t)$.

4) 用式 (98) 刷新参数估计向量 $\hat{\vartheta}(t)$.

5) 根据式 (104), $\hat{\vartheta}(t)$ 中读出 $\hat{c}(t)$, 用式 (102) 计算辅助模型输出 $\hat{u}(t)$.

6) t 增 1, 转步骤 2, 继续递推计算.

输入非线性 FIR 系统 AM-RLS 算法计算参数估计 $\hat{\vartheta}(t)$ 的流程如图 11 所示.



图 11 计算输入非线性 FIR 系统 AM-RLS

参数估计 $\hat{\vartheta}(t)$ 的流程

Fig. 11 The flowchart of computing the AM-RLS estimate $\hat{\vartheta}(t)$ for the IN-FIR system

3.2 输入非线性受控自回归系统 (IN-CAR)

输入非线性方程误差类系统 (Input Nonlinear Equation Error Type system, IN-EET) 包括基本的输入非线性方程误差模型, 即输入非线性受控自回归模型 (IN-CAR)、输入非线性受控自回归滑动平均模型 (IN-CARMA)、输入非线性受控自回归自回归模型 (IN-CARAR)、输入非线性受控自回归自回归滑动平均模型 (IN-CARARMA)。

借助于辅助模型辨识思想, 随机梯度算法、最小二乘算法、多新息辨识方法等都可以发展用于这类输入非线性方程误差类模型的辨识. 例如, 对于 IN-CAR 系统, 我们可以得到辅助模型随机梯度算法、辅助模型递推最小二乘算法、辅助模型多新息随机梯度算法、辅助模型多新息最小二乘算法等. 在下面的讨论中, 我们只给出最小二乘辨识方法.

最近, 本文作者详细综述了输入非线性方程误差系统 (即 IN-CAR 系统) 的各种辨识方法, 给出了过参数化 (over-parameterization) 随机梯度辨识方法, 提出了投影辨识方法、随机梯度算法、遗忘梯度算法 (遗忘因子随机梯度算法)、牛顿递推辨识算法、牛顿迭代辨识算法等^[39].

输入非线性方程误差类的输入非线性受控自回归系统 (Input Nonlinear CAR system, IN-CAR) 结构如图 12 所示. 其非线性部分方程为

$$\bar{u}(t) = \mathbf{f}(u(t)) = c_1 f_1(u(t)) + c_2 f_2(u(t)) + \cdots +$$

$$c_m f_m(u(t)) = \mathbf{f}(u(t)) \mathbf{c}, \quad (105)$$

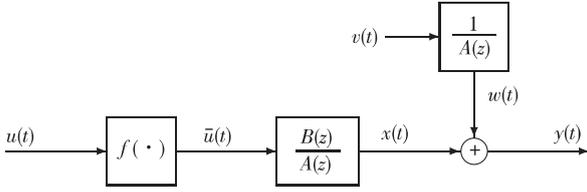


图 12 输入非线性受控自回归系统 (IN-CAR)
Fig. 12 The input nonlinear CAR system

其中 $f(u(t)) := [f_1(u(t)) \ f_2(u(t)) \ \dots \ f_m(u(t))] \in \mathbf{R}^{1 \times m}$ 是基函数构成的行向量, $\mathbf{c} := [c_1 \ c_2 \ \dots \ c_m]^T \in \mathbf{R}^m$ 是非线性部分的参数向量. 线性部分的方程为

$$A(z)y(t) = B(z)\bar{u}(t) + v(t), \quad (106)$$

其中 $A(z)$ 和 $B(z)$ 是后移算子 z^{-1} 的多项式:

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

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

为了获得唯一的参数估计, 我们直接假定了多项式 $B(z)$ 的首项 $b_0 = 1$. 参数向量 $\boldsymbol{\vartheta}$ 和信息向量 $\boldsymbol{\varphi}(t)$ 定义如下:

$$\boldsymbol{\vartheta} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{c} \end{bmatrix} \in \mathbf{R}^{2n+m}; \quad \mathbf{a} := \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathbf{R}^n; \quad \mathbf{b} := \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \in \mathbf{R}^n,$$

$$\boldsymbol{\varphi}(t) := [-y(t-1) \ -y(t-2) \ \dots \ -y(t-n) \ \bar{u}(t-1) \ \bar{u}(t-2) \ \dots \ \bar{u}(t-n) \ f(u(t))]^T \in \mathbf{R}^{2n+m}.$$

将式 (105) 代入式 (106) 得到辨识模型.

$$y(t) = [1 - A(z)]y(t) + [B(z) - 1]\bar{u}(t) + \bar{u}(t) + v(t) = -\sum_{i=1}^n a_i y(t-i) + \sum_{i=1}^n b_i \bar{u}(t-i) + f(u(t))\mathbf{c} + v(t) = \boldsymbol{\varphi}^T(t)\boldsymbol{\vartheta} + v(t). \quad (107)$$

信息向量 $\boldsymbol{\varphi}(t)$ 中未知中间变量 $\bar{u}(t-i)$ 用辅助模型的输出 $\hat{\bar{u}}(t-i)$ 代替, 可以得到估计输入非线性 CAR 系统参数向量 $\boldsymbol{\vartheta}$ 的辅助模型递推最小二乘辨识算法 (AM-RLS):

$$\hat{\boldsymbol{\vartheta}}(t) = \hat{\boldsymbol{\vartheta}}(t-1) + L(t)[y(t) - \hat{\boldsymbol{\varphi}}(t)\hat{\boldsymbol{\vartheta}}(t-1)], \quad (108)$$

$$L(t) = P(t-1)\hat{\boldsymbol{\varphi}}(t)[1 + \hat{\boldsymbol{\varphi}}(t)P(t-1)\hat{\boldsymbol{\varphi}}(t)]^{-1}, \quad (109)$$

$$P(t) = [I - L(t)\hat{\boldsymbol{\varphi}}(t)]P(t-1) \quad P(0) = p_0I, \quad (110)$$

$$\hat{\boldsymbol{\varphi}}(t) = [-y(t-1) \ -y(t-2) \ \dots \ -y(t-n) \ \hat{\bar{u}}(t-1) \ \hat{\bar{u}}(t-2) \ \dots \ \hat{\bar{u}}(t-n) \ f(u(t))]^T, \quad (111)$$

$$\hat{\bar{u}}(t) = f(u(t))\hat{\mathbf{c}}(t) \quad (112)$$

$$f(u(t)) = [f_1(u(t)) \ f_2(u(t)) \ \dots \ f_m(u(t))] \quad (113)$$

$$\hat{\boldsymbol{\vartheta}}(t) = [\hat{\mathbf{a}}^T(t) \ \hat{\mathbf{b}}^T(t) \ \hat{\mathbf{c}}^T(t)]^T. \quad (114)$$

值得指出的是: 3.1 节 IN-FIR 系统的 AM-RLS 算法, 3.2 节 IN-CAR 系统的 AM-RLS 算法的收敛性 (参数估计的一致收敛性、估计误差的有界收敛性) 是有待解决的辨识课题. 本文作者正在研究这 2 个算法的性能.

3.3 输入非线性受控自回归滑动平均系统 (IN-CARMA)

输入非线性受控自回归滑动平均系统 (Input Nonlinear CARMA system, IN-CARMA) 结构如图 13 所示. 其非线性部分方程为

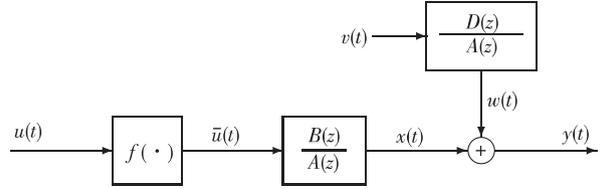


图 13 输入非线性受控自回归滑动平均系统 (IN-CARMA)
Fig. 13 The input nonlinear CARMA system

$$\bar{u}(t) = f(u(t)) = c_1 f_1(u(t)) + c_2 f_2(u(t)) + \dots + c_m f_m(u(t)) = f(u(t))\mathbf{c}, \quad (115)$$

其中 $f(u(t)) := [f_1(u(t)) \ f_2(u(t)) \ \dots \ f_m(u(t))] \in \mathbf{R}^{1 \times m}$ 是基函数构成的行向量, $\mathbf{c} := [c_1 \ c_2 \ \dots \ c_m]^T \in \mathbf{R}^m$ 是非线性部分的参数向量. 线性部分的方程为

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

或

$$A(z)y(t) = B(z)\bar{u}(t) + D(z)v(t), \quad (116)$$

其中 $A(z)$ 、 $B(z)$ 和 $D(z)$ 是后移算子 z^{-1} 的多项式:

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

$$B(z) := 1 + b_1z^{-1} + b_2z^{-2} + \dots + b_nz^{-n},$$

$$D(z) := 1 + d_1z^{-1} + d_2z^{-2} + \dots + d_{n_d}z^{-n_d}.$$

同样, 我们直接假定了多项式 $B(z)$ 的首项 $b_0 = 1$. 参数向量 $\boldsymbol{\vartheta}$ 和信息向量 $\boldsymbol{\varphi}(t)$ 定义如下:

$$\boldsymbol{\vartheta} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{c} \\ \mathbf{d} \end{bmatrix} \in \mathbf{R}^{2n+m+n_d}; \quad \mathbf{a} := \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathbf{R}^n;$$

$$\mathbf{b} := \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \in \mathbf{R}^n; \quad \mathbf{d} := \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_{n_d} \end{bmatrix} \in \mathbf{R}^{n_d},$$

$$\varphi(t) := [-y(t-1), -y(t-2), \dots, -y(t-n), \bar{u}(t-1), \bar{\mu}(t-2), \dots, \bar{\mu}(t-n), \mathbf{f}(u(t)), v(t-1), v(t-2), \dots, v(t-n_d)]^T \in \mathbf{R}^{2n+m+n_d}$$

将式(115)代入式(116)得到辨识模型

$$y(t) = [1 - A(z)]y(t) + [B(z) - 1]\bar{u}(t) + \bar{u}(t) + [D(z) - 1]v(t) + v(t) = -\sum_{i=1}^n a_i y(t-i) + \sum_{i=1}^n b_i \bar{u}(t-i) + \mathbf{f}(u(t)) \mathbf{c} + \sum_{i=1}^n d_i v(t-i) + v(t) = \varphi^T(t) \boldsymbol{\vartheta} + v(t). \quad (117)$$

信息向量 $\varphi(t)$ 中未知中间变量 $\bar{u}(t-i)$ 用辅助模型的输出 $\hat{\bar{u}}(t-i)$ 代替,未知噪声项 $v(t-i)$ 用其估计 $\hat{v}(t-i)$ 代替,可以得到估计 IN-CARMA 系统参数向量 $\boldsymbol{\vartheta}$ 的辅助模型递推增广最小二乘辨识算法 (AM-RELS):

$$\hat{\boldsymbol{\vartheta}}(t) = \hat{\boldsymbol{\vartheta}}(t-1) + \mathbf{L}(t) [y(t) - \hat{\varphi}^T(t) \hat{\boldsymbol{\vartheta}}(t-1)], \quad (118)$$

$$\mathbf{L}(t) = \mathbf{P}(t-1) \hat{\varphi}(t) [1 + \hat{\varphi}^T(t) \mathbf{P}(t-1) \hat{\varphi}(t)]^{-1}, \quad (119)$$

$$\mathbf{P}(t) = [\mathbf{I} - \mathbf{L}(t) \hat{\varphi}^T(t)] \mathbf{P}(t-1) \quad \mathbf{P}(0) = p_0 \mathbf{I}, \quad (120)$$

$$\hat{\varphi}(t) = [-y(t-1), -y(t-2), \dots, -y(t-n), \hat{\bar{u}}(t-1), \hat{\bar{\mu}}(t-2), \dots, \hat{\bar{\mu}}(t-n), \mathbf{f}(u(t))] \quad (121)$$

$$\hat{\bar{u}}(t) = \mathbf{f}(u(t)) \hat{\mathbf{c}}(t), \quad (122)$$

$$\mathbf{f}(u(t)) = [f_1(u(t)), f_2(u(t)), \dots, f_m(u(t))] \quad (123)$$

$$\hat{\mathbf{v}}(t) = y(t) - \hat{\varphi}^T(t) \hat{\boldsymbol{\vartheta}}(t), \quad (124)$$

$$\hat{\boldsymbol{\vartheta}}(t) = [\hat{\mathbf{a}}^T(t) \quad \hat{\mathbf{b}}^T(t) \quad \hat{\mathbf{c}}^T(t) \quad \hat{\boldsymbol{\gamma}}^T(t)]^T. \quad (125)$$

本文作者等利用过参数化方法 (over-parameterization method) 提出了 IN-CARMA 系统的最小二乘迭代辨识方法和递推增广最小二乘辨识方法^[40],以及梯度迭代辨识方法和增广随机梯度辨识方法^[41].

3.4 输入非线性受控自回归自回归系统 (IN-CARAR)

输入非线性受控自回归自回归系统 (Input Non-linear CARAR system, IN-CARAR) 结构如图 14 所示. 其非线性部分方程为

$$\bar{u}(t) = \mathbf{f}(u(t)) = c_1 f_1(u(t)) + c_2 f_2(u(t)) + \dots + c_m f_m(u(t)) =$$

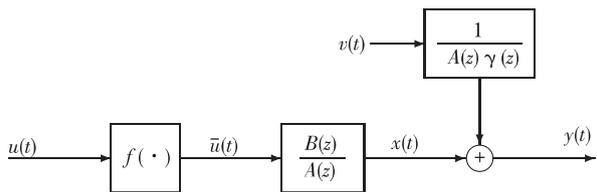


图 14 输入非线性受控自回归自回归系统 (IN-CARAR)

Fig. 14 The input nonlinear CARAR system

$$\mathbf{f}(u(t)) \mathbf{c}, \quad (126)$$

其中 $\mathbf{f}(u(t)) := [f_1(u(t)), f_2(u(t)), \dots, f_m(u(t))] \in \mathbf{R}^{1 \times m}$ 是基函数构成的行向量, $\mathbf{c} := [c_1, c_2, \dots, c_m]^T \in \mathbf{R}^m$ 是非线性部分的参数向量. 线性部分的方程为

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

或

$$A(z) y(t) = B(z) \bar{u}(t) + \frac{1}{\gamma(z)} v(t), \quad (127)$$

其中 $A(z)$, $B(z)$ 和 $\gamma(z)$ 是后移算子 z^{-1} 的多项式:

$$A(z) := 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n},$$

$$B(z) := 1 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_n z^{-n},$$

$$\gamma(z) := 1 + \gamma_1 z^{-1} + \gamma_2 z^{-2} + \dots + \gamma_{n_\gamma} z^{-n_\gamma}.$$

同样假定了多项式 $B(z)$ 的首项 $b_0 = 1$. 定义中间相关噪声变量:

$$w(t) := \frac{1}{\gamma(z)} v(t), \quad (128)$$

或

$$w(t) = [1 - \gamma(z)] w(t) + v(t) = -\sum_{i=1}^{n_\gamma} \gamma_i w(t-i) + v(t). \quad (129)$$

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

$$\boldsymbol{\vartheta} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{c} \\ \boldsymbol{\gamma} \end{bmatrix} \in \mathbf{R}^{2n+m+n_\gamma}; \quad \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathbf{R}^n;$$

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \in \mathbf{R}^n; \quad \boldsymbol{\gamma} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_{n_\gamma} \end{bmatrix} \in \mathbf{R}^{n_\gamma},$$

$$\varphi(t) := [-y(t-1), -y(t-2), \dots, -y(t-n), \bar{u}(t-1), \bar{\mu}(t-2), \dots, \bar{\mu}(t-n), \mathbf{f}(u(t)), -w(t-1), -w(t-2), \dots, -w(t-n_\gamma)]^T \in \mathbf{R}^{2n+m+n_\gamma}. \quad (130)$$

将式(126)和(128)代入式(127),利用式(129)可得到辨识模型

$$y(t) = [1 - A(z)]y(t) + [B(z) - 1]\bar{u}(t) + \bar{u}(t) + w(t) = -\sum_{i=1}^n a_i y(t-i) + \sum_{i=1}^n b_i \bar{u}(t-i) + \mathbf{f}(u(t)) \mathbf{c} - \sum_{i=1}^{n_\gamma} \gamma_i w(t-i) + v(t) = \varphi^T(t) \boldsymbol{\vartheta} + v(t). \quad (131)$$

信息向量 $\varphi(t)$ 中未知中间变量 $\bar{u}(t-i)$ 用辅助模型的输出 $\hat{\bar{u}}(t-i)$ 代替, 未知相关噪声项 $w(t-i)$ 用其估计 $\hat{w}(t-i)$ 代替, 可以得到估计 IN-CARAR 系统参数向量 ϑ 的辅助模型递推广义最小二乘辨识算法 (AM-RGLS):

$$\hat{\vartheta}(t) = \hat{\vartheta}(t-1) + L(t) [y(t) - \hat{\varphi}^T(t) \hat{\vartheta}(t-1)] \quad (132)$$

$$L(t) = P(t-1) \hat{\varphi}(t) [1 + \hat{\varphi}^T(t) P(t-1) \hat{\varphi}(t)]^{-1} \quad (133)$$

$$P(t) = [I - L(t) \hat{\varphi}^T(t)] P(t-1) \quad P(0) = p_0 I \quad (134)$$

$$\hat{\varphi}(t) = [\hat{\varphi}_s^T(t), -\hat{w}(t-1), -\hat{w}(t-2), \dots, -\hat{w}(t-n_\gamma)]^T \quad (135)$$

$$\hat{\varphi}_s(t) = [-y(t-1), -y(t-2), \dots, -y(t-n)] \quad (136)$$

$$\hat{\bar{u}}(t-1) \hat{\bar{u}}(t-2) \dots \hat{\bar{u}}(t-n) \mathbf{f}(u(t))]^T \quad (136)$$

$$\hat{\bar{u}}(t) = \mathbf{f}(u(t)) \hat{\mathbf{c}}(t) \quad (137)$$

$$\mathbf{f}(u(t)) = [f_1(u(t)) \ f_2(u(t)) \ \dots \ f_m(u(t))] \quad (138)$$

$$\hat{w}(t) = y(t) - \hat{\varphi}_s^T(t) [\hat{\mathbf{a}}^T \ \hat{\mathbf{b}}^T \ \hat{\mathbf{c}}^T(t)]^T \quad (139)$$

$$\hat{\vartheta}(t) = [\hat{\mathbf{a}}^T(t) \ \hat{\mathbf{b}}^T(t) \ \hat{\mathbf{c}}^T(t) \ \hat{\boldsymbol{\gamma}}^T(t)]^T \quad (140)$$

3.5 输入非线性受控自回归自回归滑动平均系统 (IN-CARARMA)

输入非线性受控自回归自回归滑动平均系统 (Input Nonlinear CARARMA system, IN-CARARMA) 结构如图 15 所示. 其非线性部分方程为

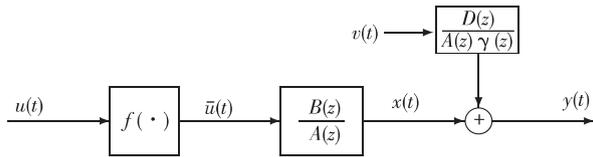


图 15 输入非线性受控自回归自回归滑动平均系统 (IN-CARARMA)

Fig. 15 The input nonlinear CARARMA system

$$\bar{u}(t) = \mathbf{f}(u(t)) = c_1 f_1(u(t)) + c_2 f_2(u(t)) + \dots + c_m f_m(u(t)) = \mathbf{f}(u(t)) \mathbf{c} \quad (141)$$

其中 $\mathbf{f}(u(t)) := [f_1(u(t)) \ f_2(u(t)) \ \dots \ f_m(u(t))] \in \mathbf{R}^{1 \times m}$ 是基函数构成的行向量, $\mathbf{c} := [c_1 \ c_2 \ \dots \ c_m]^T \in \mathbf{R}^m$ 是非线性部分的参数向量. 线性部分的方程为

$$y(t) = \frac{B(z)}{A(z)} \bar{u}(t) + \frac{D(z)}{A(z) \gamma(z)} v(t) \quad (142)$$

或

$$A(z) y(t) = B(z) \bar{u}(t) + \frac{D(z)}{\gamma(z)} v(t) \quad (142)$$

其中 $A(z)$ $B(z)$ $D(z)$ 和 $\gamma(z)$ 是后移算子 z^{-1} 的多

项式:

$$A(z) := 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n} \quad (143)$$

$$B(z) := 1 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_n z^{-n} \quad (143)$$

$$D(z) := 1 + d_1 z^{-1} + d_2 z^{-2} + \dots + d_{n_d} z^{-n_d} \quad (143)$$

$$\gamma(z) := 1 + \gamma_1 z^{-1} + \gamma_2 z^{-2} + \dots + \gamma_{n_\gamma} z^{-n_\gamma} \quad (143)$$

同样假定了多项式 $B(z)$ 的首项 $b_0 = 1$. 定义中间相关噪声变量:

$$w(t) := \frac{D(z)}{\gamma(z)} v(t) \quad (143)$$

或

$$w(t) = [1 - \gamma(z)] w(t) + [D(z) - 1] v(t) + v(t) = - \sum_{i=1}^{n_\gamma} \gamma_i w(t-i) + \sum_{i=1}^{n_d} d_i v(t-i) + v(t) \quad (144)$$

参数向量 ϑ 和信息向量 $\varphi(t)$ 定义如下:

$$\vartheta(t) := \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{c} \\ \boldsymbol{\gamma} \\ \mathbf{d} \end{bmatrix} \in \mathbf{R}^{2n+m+n_\gamma+n_d}; \quad \mathbf{a} := \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathbf{R}^n;$$

$$\mathbf{b} := \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \in \mathbf{R}^n; \quad \boldsymbol{\gamma} := \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_{n_\gamma} \end{bmatrix} \in \mathbf{R}^{n_\gamma}; \quad \mathbf{d} := \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_{n_d} \end{bmatrix} \in \mathbf{R}^{n_d}$$

$$\varphi(t) := [-y(t-1), -y(t-2), \dots, -y(t-n), \bar{u}(t-1) \ \bar{\mu}(t-2) \ \dots \ \bar{\mu}(t-n) \ \mathbf{f}(u(t)), -w(t-1) \ w(t-2) \ \dots \ w(t-n_\gamma) \ v(t-1) \ v(t-2) \ \dots \ v(t-n_d)]^T \in \mathbf{R}^{2n+m+n_\gamma+n_d}$$

将式(141)和(143)代入式(142), 利用式(144)可得辨识模型,

$$y(t) = [1 - A(z)] y(t) + [B(z) - 1] \bar{u}(t) + \bar{u}(t) + w(t) = - \sum_{i=1}^n a_i y(t-i) + \sum_{i=1}^n b_i \bar{u}(t-i) + \mathbf{f}(u(t)) \mathbf{c} - \sum_{i=1}^{n_\gamma} \gamma_i w(t-i) + \sum_{i=1}^{n_d} d_i v(t-i) + v(t) = \varphi^T(t) \vartheta + v(t) \quad (145)$$

信息向量 $\varphi(t)$ 中未知中间变量 $\bar{u}(t-i)$ 用辅助模型的输出 $\hat{\bar{u}}(t-i)$ 代替, 未知噪声项 $w(t-i)$ 和白噪声项 $v(t-i)$ 分别用其估计 $\hat{w}(t-i)$ 和 $\hat{v}(t-i)$ 代替, 可以得到估计 IN-CARARMA 系统参数向量 ϑ 的辅助模型递推广义增广最小二乘辨识算法 (AM-RGELS):

$$\hat{\vartheta}(t) = \hat{\vartheta}(t-1) + L(t) [y(t) - \hat{\varphi}^T(t) \hat{\vartheta}(t-1)] \quad (146)$$

$$L(t) = P(t-1) \hat{\varphi}(t) [1 + \hat{\varphi}^T(t) P(t-1) \hat{\varphi}(t)]^{-1} \quad (147)$$

$$P(t) = [I - L(t)] \hat{\varphi}^T(t)] P(t-1) \quad P(0) = p_0 I, \quad (148)$$

$$\hat{\varphi}^T(t) = [\hat{\varphi}_s^T(t), -\hat{w}(t-1), -\hat{w}(t-2), \dots, -\hat{w}(t-n_y), \hat{\nu}(t-1), \hat{\nu}(t-2), \dots, \hat{\nu}(t-n_d)]^T, \quad (149)$$

$$\hat{\varphi}_s^T(t) = [-y(t-1), -y(t-2), \dots, -y(t-n), \hat{u}(t-1), \hat{u}(t-2), \dots, \hat{u}(t-n), f(u(t))]^T, \quad (150)$$

$$\hat{u}(t) = f(u(t)) \hat{c}(t), \quad (151)$$

$$f(u(t)) = [f_1(u(t)), f_2(u(t)), \dots, f_m(u(t))] , \quad (152)$$

$$\hat{w}(t) = y(t) - \hat{\varphi}_s^T(t) [\hat{a}^T(t) \hat{b}^T(t) \hat{c}^T(t)]^T, \quad (153)$$

$$\hat{v}(t) = y(t) - \hat{\varphi}^T(t) \hat{\theta}(t), \quad (154)$$

$$\hat{\theta}(t) = [\hat{a}^T(t) \hat{b}^T(t) \hat{c}^T(t) \hat{d}^T(t)]^T. \quad (155)$$

本节讨论的一些输入非线性系统 (IN-CARMA, IN-CARAR, IN-CARARMA) 的辅助模型递推最小二乘算法的收敛性分析 (参数估计的一致收敛性、估计误差的有界收敛性) 仍然是控制科学家有待解决的辨识难题。

4 输入非线性输出误差类系统 (IN-OET)

输入非线性输出误差类系统 (Input Nonlinear Output Error Type system, IN-OET) 包括基本的输入非线性输出误差系统 (IN-OE)、输入非线性输出误差滑动平均系统 (IN-OEMA)、输入非线性输出误差自回归系统 (IN-OEAR)、输入非线性输出误差自回归滑动平均系统 (IN-OEARMA) 即输入非线性 Box-Jenkins 系统 (IN-Box-Jenkins)。

4.1 输入非线性输出误差系统 (IN-OE)

输入非线性输出误差系统 (Input Nonlinear Output Error system, IN-OE) 如图 16 所示, 其中线性动态部分的输出 (也称系统的真实输出或无噪输出) $x(t)$ 和内部变量 $\bar{u}(t)$ (非线性环节的输出) 是不可测的 $u(t)$ 是系统输入 $y(t)$ 是 $x(t)$ 的含噪量测 $v(t)$ 是量测白噪声。(原始) Hammerstein 模型的非线性部分是一个已知阶次的输入的多项式:

$\bar{u}(t) = f(u(t)) = c_1 u(t) + c_2 u^2(t) + \dots + c_m u^m(t)$, c_i 为多项式的系数, m 为多项式的阶次。后来, Hammerstein 模型非线性部分扩展为一个已知基 (known basis) (f_1, f_2, \dots, f_m) 的非线性函数:

$$\bar{u}(t) = f(u(t)) = c_1 f_1(u(t)) + c_2 f_2(u(t)) + \dots + c_m f_m(u(t)) = f(u(t)) \mathbf{c}, \quad (156)$$

其中 $f(u(t)) := [f_1(u(t)), f_2(u(t)), \dots, f_m(u(t))] \in \mathbf{R}^{1 \times m}$ 是基函数构成的行向量, $\mathbf{c} := [c_1, c_2, \dots, c_m]^T \in \mathbf{R}^m$ 是非线性部分的参数向量。

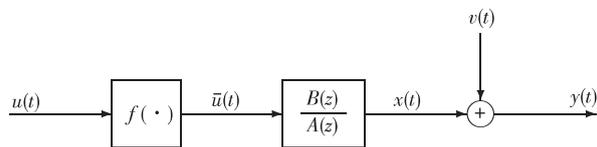


图 16 输入非线性输出误差系统 (IN-OE)

Fig. 16 The input nonlinear output error system

这个 Hammerstein 模型线性部分是一个输出误差子系统, 其传递函数为

$$G(z) := \frac{B(z)}{A(z)}, \quad (157)$$

其中 $A(z)$ 和 $B(z)$ 是后移算子 z^{-1} 的多项式:

$$A(z) := 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n},$$

$$B(z) := b_0 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_n z^{-n}.$$

因此, 线性部分可以表示为

$$y(t) = x(t) + v(t),$$

$$x(t) = \frac{B(z)}{A(z)} \bar{u}(t). \quad (158)$$

为了获得唯一的参数估计, 固定 $b_0 = 1$. 上式可以写为

$$x(t) = [1 - A(z)]x(t) + [B(z) - 1]\bar{u}(t) + \bar{u}(t) = -\sum_{i=1}^n a_i x(t-i) + \sum_{i=1}^n b_i \bar{u}(t-i) + f(u(t)) \mathbf{c}, \quad (159)$$

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

$$\boldsymbol{\theta} := \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{c} \end{bmatrix} \in \mathbf{R}^{2n+m}; \quad \mathbf{a} := \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathbf{R}^n;$$

$$\mathbf{b} := \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \in \mathbf{R}^n; \quad \mathbf{c} := \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix} \in \mathbf{R}^m;$$

$$\boldsymbol{\varphi}(t) := [-x(t-1), -x(t-2), \dots, -x(t-n), \bar{u}(t-1), \bar{u}(t-2), \dots, \bar{u}(t-n), f(u(t))]^T \in \mathbf{R}^{2n+m}. \quad (160)$$

将式 (159) 代入式 (160) 得到辨识模型

$$y(t) = [1 - A(z)]x(t) + [B(z) - 1]\bar{u}(t) + \bar{u}(t) + v(t) = -\sum_{i=1}^n a_i x(t-i) + \sum_{i=1}^n b_i \bar{u}(t-i) + f(u(t)) \mathbf{c} + v(t) = \boldsymbol{\varphi}^T(t) \boldsymbol{\theta} + v(t). \quad (161)$$

对于辨识模型 (161), 极小化准则函数

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

可得下列递推最小二乘算法:

$$\boldsymbol{\vartheta}(t) = \boldsymbol{\vartheta}(t-1) + \mathbf{L}(t) [y(t) - \boldsymbol{\varphi}^T(t) \boldsymbol{\vartheta}(t-1)], \quad (162)$$

$$\mathbf{L}(t) = \frac{\mathbf{P}(t-1) \boldsymbol{\varphi}(t)}{1 + \boldsymbol{\varphi}^T(t) \mathbf{P}(t-1) \boldsymbol{\varphi}(t)}, \quad (163)$$

$$\mathbf{P}(t) = [\mathbf{I} - \mathbf{L}(t) \boldsymbol{\varphi}^T(t)] \mathbf{P}(t-1) \mathbf{P}(0) = p_0 \mathbf{I}. \quad (164)$$

类似地, 这个算法不能实现, 因为信息向量 $\boldsymbol{\varphi}(t)$ 中包含了系统未知真实输出 $x(t-i)$ 和未知中间变量 $\bar{u}(t-i)$, 解决的办法是它们分别用辅助模型的输出代替 $\hat{x}(t-i)$ 和 $\hat{\bar{u}}(t)$ 代替. 分别用 2 个辅助模型的输出 $\hat{x}(t-i)$ 和 $\hat{\bar{u}}(t-i)$ 代替式 (161) 信息向量 $\boldsymbol{\varphi}(t)$ 中的未知项 $x(t-i)$ 和 $\bar{u}(t-i)$, 代替后的信息向量记作为

$$\hat{\boldsymbol{\varphi}}(t) := [-\hat{x}(t-1), -\hat{x}(t-2), \dots, -\hat{x}(t-n), \hat{\bar{\mu}}(t-1), \hat{\bar{\mu}}(t-2), \dots, \hat{\bar{\mu}}(t-n), \mathbf{f}(u(t))]^T \in \mathbf{R}^{2n+m}.$$

设 $\boldsymbol{\vartheta}$ 在时刻 t 的估计为

$$\hat{\boldsymbol{\vartheta}}(t) := \begin{bmatrix} \hat{\mathbf{a}}(t) \\ \hat{\mathbf{b}}(t) \\ \hat{\mathbf{c}}(t) \end{bmatrix} \in \mathbf{R}^{2n+m}; \quad \hat{\mathbf{a}}(t) := \begin{bmatrix} \hat{a}_1(t) \\ \hat{a}_2(t) \\ \vdots \\ \hat{a}_n(t) \end{bmatrix} \in \mathbf{R}^n;$$

$$\hat{\mathbf{b}}(t) := \begin{bmatrix} \hat{b}_1(t) \\ \hat{b}_2(t) \\ \vdots \\ \hat{b}_n(t) \end{bmatrix} \in \mathbf{R}^n; \quad \hat{\mathbf{c}}(t) := \begin{bmatrix} \hat{c}_1(t) \\ \hat{c}_2(t) \\ \vdots \\ \hat{c}_m(t) \end{bmatrix} \in \mathbf{R}^m. \quad (165)$$

用 $\hat{\boldsymbol{\varphi}}(t)$ 和 $\hat{\boldsymbol{\vartheta}}(t)$ 代替式 (159) 中的 $\boldsymbol{\varphi}(t)$ 和 $\boldsymbol{\vartheta}$, 那么计算真实输出的辅助模型可以表示为

$$\hat{x}(t) = \hat{\boldsymbol{\varphi}}^T(t) \hat{\boldsymbol{\vartheta}}(t).$$

用 $\hat{\mathbf{c}}(t)$ 代替式 (153) 中参数 c_i , 那么计算非线性环节输出 $\bar{u}(t)$ 的辅助模型为

$$\hat{\bar{u}}(t) = \hat{c}_1(t) f_1(u(t)) + \hat{c}_2(t) f_2(u(t)) + \dots + \hat{c}_m(t) f_m(u(t)) = \sum_{j=1}^m \hat{c}_j(t) f_j(u(t)) = \mathbf{f}(u(t)) \hat{\mathbf{c}}(t).$$

用 $\hat{\boldsymbol{\varphi}}(t)$ 代替式 (162) — (164) 中未知 $\boldsymbol{\varphi}(t)$, 可以总结出估计输入非线性输出误差系统参数向量 $\boldsymbol{\vartheta}$ 的辅助模型递推最小二乘辨识算法 (AM-RLS):

$$\hat{\boldsymbol{\vartheta}}(t) = \hat{\boldsymbol{\vartheta}}(t-1) + \mathbf{L}(t) [y(t) - \hat{\boldsymbol{\varphi}}^T(t) \hat{\boldsymbol{\vartheta}}(t-1)], \quad (166)$$

$$\mathbf{L}(t) = \mathbf{P}(t-1) \hat{\boldsymbol{\varphi}}(t) [1 + \hat{\boldsymbol{\varphi}}^T(t) \mathbf{P}(t-1) \hat{\boldsymbol{\varphi}}(t)]^{-1}, \quad (167)$$

$$\mathbf{P}(t) = [\mathbf{I} - \mathbf{L}(t) \hat{\boldsymbol{\varphi}}^T(t)] \mathbf{P}(t-1) \mathbf{P}(0) = p_0 \mathbf{I}, \quad (168)$$

$$\hat{\boldsymbol{\varphi}}(t) = [-\hat{x}(t-1), -\hat{x}(t-2), \dots, -\hat{x}(t-n),$$

$$\hat{\bar{u}}(t-1), \hat{\bar{u}}(t-2), \dots, \hat{\bar{u}}(t-n), \mathbf{f}(u(t))]^T, \quad (169)$$

$$\hat{x}(t) = \hat{\boldsymbol{\varphi}}^T(t) \hat{\boldsymbol{\vartheta}}(t), \quad (170)$$

$$\hat{\bar{u}}(t) = \mathbf{f}(u(t)) \hat{\mathbf{c}}(t), \quad (171)$$

$$\mathbf{f}(u(t)) = [f_1(u(t)), f_2(u(t)), \dots, f_m(u(t))], \quad (172)$$

$$\hat{\boldsymbol{\vartheta}}(t) = [\hat{\mathbf{a}}^T(t), \hat{\mathbf{b}}^T(t), \hat{\mathbf{c}}^T(t)]^T. \quad (173)$$

这个 IN-OE 系统辅助模型最小二乘算法的收敛性有待研究.

输入非线性输出误差系统 AM-RLS 算法 (166) — (173) 的计算步骤如下.

1) 初始化: 令 $t=1$. 置初值 $\mathbf{P}(0) = p_0 \mathbf{I}$, $\hat{\boldsymbol{\vartheta}}(0) = \mathbf{1}/p_0$, $p_0 = 10^6$, $\hat{x}(-i) = 1/p_0$, $\hat{\bar{u}}(-i) = 1/p_0$, $i=0, 1, 2, \dots, n-1$. 给定基函数 $f_j(\cdot)$.

2) 收集数据 $u(t)$ 和 $y(t)$, 用式 (172) 构造基函数行向量 $\mathbf{f}(u(t))$, 用式 (169) 构造信息向量 $\hat{\boldsymbol{\varphi}}(t)$.

3) 用式 (167) 计算增益向量 $\mathbf{L}(t)$, 用式 (165) 计算协方差阵 $\mathbf{P}(t)$.

4) 用式 (166) 刷新参数估计向量 $\hat{\boldsymbol{\vartheta}}(t)$.

5) 根据式 (172), 从 $\hat{\boldsymbol{\vartheta}}(t)$ 中读出 $\hat{\mathbf{c}}(t)$, 用式 (171) 计算辅助模型输出 $\hat{\bar{u}}(t)$.

6) t 增 1, 转步骤 2, 继续递推计算.

输入非线性输出误差系统 AM-RLS 算法计算参数估计 $\hat{\boldsymbol{\vartheta}}(t)$ 的流程如图 17 所示.

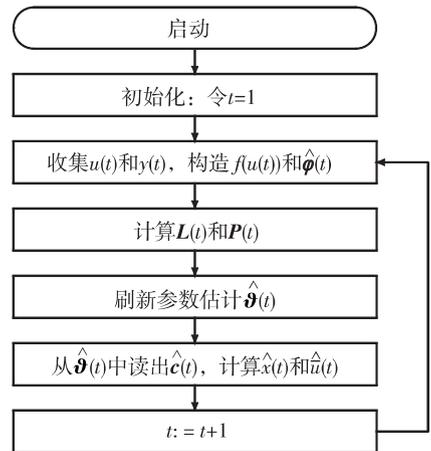


图 17 计算输入非线性输出误差系统 AM-RLS 参数估计 $\hat{\boldsymbol{\vartheta}}(t)$ 的流程

Fig. 17 The flowchart of computing the AM-RLS estimate $\hat{\boldsymbol{\vartheta}}(t)$ for the IN-OE system

本文作者等^[16]利用过参数化方法,提出了 IN-OE 系统的辅助模型递推最小二乘算法,并研究了提出算法的收敛性.

4.2 输入非线性输出误差滑动平均系统(IN-OEMA)

输入非线性输出误差滑动平均系统 (Input Nonlinear OEMA system, IN-OEMA) 结构如图 18 所示,其非线性部分方程为

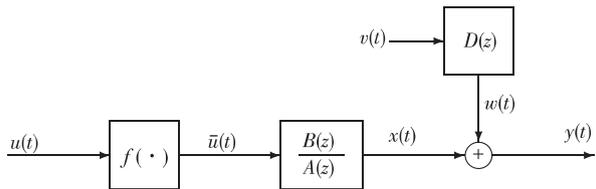


图 18 输入非线性输出误差滑动平均系统(IN-OEMA)

Fig. 18 The input nonlinear OEMA system

$$\bar{u}(t) = f(u(t)) = c_1 f_1(u(t)) + c_2 f_2(u(t)) + \dots + c_m f_m(u(t)) = f(u(t)) \mathbf{c}, \quad (174)$$

其中 $f(u(t)) := [f_1(u(t)) \ f_2(u(t)) \ \dots \ f_m(u(t))]$ $\in \mathbf{R}^{1 \times m}$ 是基函数构成的行向量, $\mathbf{c} := [c_1 \ c_2 \ \dots \ c_m]^T \in \mathbf{R}^m$ 是非线性部分的参数向量. 线性部分的方程为

$$y(t) = \frac{B(z)}{A(z)} \bar{u}(t) + D(z) v(t), \quad (175)$$

其中 $A(z)$ 、 $B(z)$ 和 $D(z)$ 是后移算子 z^{-1} 的多项式:

$$A(z) := 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n},$$

$$B(z) := 1 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_n z^{-n},$$

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

同样,我们直接假定了多项式 $B(z)$ 的首项 $b_0 = 1$.

1. 定义未知真实输出

$$x(t) := \frac{B(z)}{A(z)} \bar{u}(t). \quad (176)$$

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

$$\boldsymbol{\vartheta} := \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{c} \\ \mathbf{d} \end{bmatrix} \in \mathbf{R}^{2n+m+n_d}, \quad \mathbf{a} := \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathbf{R}^n;$$

$$\mathbf{b} := \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \in \mathbf{R}^n; \quad \mathbf{d} := \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_{n_d} \end{bmatrix} \in \mathbf{R}^{n_d};$$

$$\boldsymbol{\varphi}(t) := [-x(t-1) \ -x(t-2) \ \dots \ -x(t-n) \ \bar{u}(t-1) \ \bar{u}(t-2) \ \dots \ \bar{u}(t-n)],$$

$$[f(u(t)) \ v(t-1) \ v(t-2) \ \dots \ v(t-n_d)]^T \in \mathbf{R}^{2n+m+n_d}.$$

式(176)可以写为

$$x(t) = [1 - A(z)]x(t) + [B(z) - 1]\bar{u}(t) + \bar{u}(t) = -\sum_{i=1}^n a_i x(t-i) + \sum_{i=1}^n b_i \bar{u}(t-i) + f(u(t)) \mathbf{c}. \quad (177)$$

将式(176)代入式(175),利用式(177)得到辨识模型

$$y(t) = x(t) + [D(z) - 1]v(t) + v(t) = -\sum_{i=1}^n a_i x(t-i) + \sum_{i=1}^n b_i \bar{u}(t-i) + f(u(t)) \mathbf{c} + \sum_{i=1}^n d_i v(t-i) + v(t) = \boldsymbol{\varphi}^T(t) \boldsymbol{\vartheta} + v(t). \quad (178)$$

信息向量 $\boldsymbol{\varphi}(t)$ 中未知真实输出 $x(t-i)$ 用辅助模型(184)输出 $\hat{x}(t-i)$ 代替,未知中间变量 $\bar{u}(t-i)$

用辅助模型(185)的输出 $\hat{\bar{u}}(t-i)$ 代替,未知噪声项 $v(t-i)$ 用式(186)中估计残差 $\hat{v}(t-i)$ 代替,可以得到估计 IN-OEMA 系统参数向量 $\boldsymbol{\vartheta}$ 的辅助模型递推增广最小二乘辨识算法 (AM-RELS):

$$\hat{\boldsymbol{\vartheta}}(t) = \hat{\boldsymbol{\vartheta}}(t-1) + \mathbf{L}(t) [y(t) - \hat{\boldsymbol{\varphi}}^T(t) \hat{\boldsymbol{\vartheta}}(t-1)], \quad (179)$$

$$\mathbf{L}(t) = \mathbf{P}(t-1) \hat{\boldsymbol{\varphi}}(t) [1 + \hat{\boldsymbol{\varphi}}^T(t) \mathbf{P}(t-1) \hat{\boldsymbol{\varphi}}(t)]^{-1}, \quad (180)$$

$$\mathbf{P}(t) = [\mathbf{I} - \mathbf{L}(t) \hat{\boldsymbol{\varphi}}^T(t)] \mathbf{P}(t-1) \quad \mathbf{P}(0) = p_0 \mathbf{I}, \quad (181)$$

$$\hat{\boldsymbol{\varphi}}(t) = [\hat{\boldsymbol{\varphi}}_s^T(t) \ \hat{\mathbf{v}}(t-1) \ \hat{\mathbf{v}}(t-2) \ \dots \ \hat{\mathbf{v}}(t-n_d)]^T, \quad (182)$$

$$\hat{\boldsymbol{\varphi}}_s(t) = [-\hat{x}(t-1) \ -\hat{x}(t-2) \ \dots \ -\hat{x}(t-n) \ \hat{\bar{\mu}}(t-1) \ \hat{\bar{\mu}}(t-2) \ \dots \ \hat{\bar{\mu}}(t-n) \ f(u(t))]^T, \quad (183)$$

$$\hat{x}(t) = \hat{\boldsymbol{\varphi}}_s^T(t) [\hat{\mathbf{a}}^T(t) \ \hat{\mathbf{b}}^T(t) \ \hat{\mathbf{c}}^T(t)]^T, \quad (184)$$

$$\hat{\bar{u}}(t) = f(u(t)) \hat{\mathbf{c}}(t), \quad (185)$$

$$\hat{v}(t) = y(t) - \hat{\boldsymbol{\varphi}}^T(t) \hat{\boldsymbol{\vartheta}}(t), \quad (186)$$

$$f(u(t)) = [f_1(u(t)) \ f_2(u(t)) \ \dots \ f_m(u(t))], \quad (187)$$

$$\hat{\boldsymbol{\vartheta}}(t) = [\hat{\mathbf{a}}^T(t) \ \hat{\mathbf{b}}^T(t) \ \hat{\mathbf{c}}^T(t) \ \hat{\mathbf{d}}^T(t)]^T. \quad (188)$$

王冬青等^[22]提出了 IN-OEMA 系统的辅助模型增广最小二乘辨识算法和多新息增广最小二乘算法. 陈晶等提出了多变量 IN-OEMA 系统的辅助模型多新息增广随机梯度辨识算法^[42].

4.3 输入非线性输出误差自回归系统(IN-OEAR)

输入非线性输出误差自回归系统 (Input Nonlinear OEAR system, IN-OEAR) 结构如图 19 所示,其非线性部分方程为

$$\bar{u}(t) = f(u(t)) =$$

$$c_1 f_1(u(t)) + c_2 f_2(u(t)) + \cdots + c_m f_m(u(t)) = f(u(t)) \mathbf{c}, \quad (189)$$

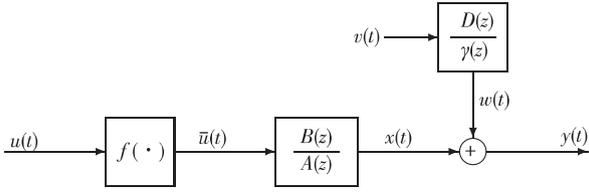


图 19 输入非线性输出误差自回归系统 (IN-OEAR)

Fig. 19 The input nonlinear OEAR system

其中 $f(u(t)) := [f_1(u(t)) \ f_2(u(t)) \ \cdots \ f_m(u(t))] \in \mathbf{R}^{1 \times m}$ 是基函数构成的行向量, $\mathbf{c} := [c_1 \ c_2 \ \cdots \ c_m]^T \in \mathbf{R}^m$ 是非线性部分的参数向量. 线性部分的方程为

$$y(t) = \frac{B(z)}{A(z)} \bar{u}(t) + \frac{1}{\gamma(z)} v(t), \quad (190)$$

其中 $A(z)$ 、 $B(z)$ 和 $\gamma(z)$ 是后移算子 z^{-1} 的多项式:

$$A(z) := 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_n z^{-n},$$

$$B(z) := 1 + b_1 z^{-1} + b_2 z^{-2} + \cdots + b_n z^{-n},$$

$$\gamma(z) := 1 + \gamma_1 z^{-1} + \gamma_2 z^{-2} + \cdots + \gamma_{n_\gamma} z^{-n_\gamma}.$$

同样, 我们直接假定了多项式 $B(z)$ 的首项 $b_0 = 1$. 定义未知真实输出 $x(t)$ 和中间相关噪声变量 $w(t)$ 如下:

$$x(t) := \frac{B(z)}{A(z)} \bar{u}(t); \quad (191)$$

$$w(t) := \frac{1}{\gamma(z)} v(t). \quad (192)$$

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

$$\boldsymbol{\vartheta} := \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{c} \\ \boldsymbol{\gamma} \end{bmatrix} \in \mathbf{R}^{2n+m+n_\gamma}; \quad \mathbf{a} := \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathbf{R}^n;$$

$$\mathbf{b} := \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \in \mathbf{R}^n; \quad \boldsymbol{\gamma} := \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_{n_\gamma} \end{bmatrix} \in \mathbf{R}^{n_\gamma};$$

$$\boldsymbol{\varphi}(t) := [-x(t-1) \ -x(t-2) \ \cdots \ -x(t-n) \ \bar{u}(t-1) \ \bar{u}(t-2) \ \cdots \ \bar{u}(t-n) \ f(u(t)) \ -w(t-1) \ -w(t-2) \ \cdots \ -w(t-n_\gamma)]^T \in \mathbf{R}^{2n+m+n_\gamma}.$$

式(191) — (192) 可以写为

$$x(t) = [1 - A(z)]x(t) + [B(z) - 1]\bar{u}(t) + \bar{u}(t) = -\sum_{i=1}^n a_i x(t-i) + \sum_{i=1}^n b_i \bar{u}(t-i) + f(u(t)) \mathbf{c}, \quad (193)$$

$$w(t) = [1 - \gamma(z)]w(t) + v(t) = -\sum_{i=1}^n \gamma_i w(t-i) + v(t). \quad (194)$$

将式(191) — (192) 代入式(190), 利用式(193) — (194) 得到辨识模型,

$$y(t) = x(t) + w(t) = -\sum_{i=1}^n a_i x(t-i) + \sum_{i=1}^n b_i \bar{u}(t-i) + f(u(t)) \mathbf{c} - \sum_{i=1}^{n_\gamma} \gamma_i w(t-i) + v(t) = \boldsymbol{\varphi}^T(t) \boldsymbol{\vartheta} + v(t). \quad (195)$$

信息向量 $\boldsymbol{\varphi}(t)$ 中未知真实输出 $x(t-i)$ 用辅助模型(201) 输出 $\hat{x}(t-i)$ 代替, 未知中间变量 $\bar{u}(t-i)$ 用辅助模型(202) 的输出 $\hat{\bar{u}}(t-i)$ 代替, 未知噪声项 $w(t-i)$ 用式(203) 中估计残差 $\hat{w}(t-i)$ 代替, 可以得到估计 IN-OEAR 系统参数向量 $\boldsymbol{\vartheta}$ 的辅助模型递推推广最小二乘辨识算法 (AM-RGLS):

$$\hat{\boldsymbol{\vartheta}}(t) = \hat{\boldsymbol{\vartheta}}(t-1) + L(t) [y(t) - \hat{\boldsymbol{\varphi}}^T(t) \hat{\boldsymbol{\vartheta}}(t-1)], \quad (196)$$

$$L(t) = P(t-1) \hat{\boldsymbol{\varphi}}(t) [1 + \hat{\boldsymbol{\varphi}}^T P(t-1) \hat{\boldsymbol{\varphi}}(t)]^{-1}, \quad (197)$$

$$P(t) = [I - L(t) \hat{\boldsymbol{\varphi}}^T(t)] P(t-1) \quad P(0) = p_0 I, \quad (198)$$

$$\hat{\boldsymbol{\varphi}}(t) = [\hat{\boldsymbol{\varphi}}_s^T(t) \ -\hat{w}(t-1) \ -\hat{w}(t-2) \ \cdots \ -\hat{w}(t-n_\gamma)]^T, \quad (199)$$

$$\hat{\boldsymbol{\varphi}}_s(t) = [-\hat{x}(t-1) \ -\hat{x}(t-2) \ \cdots \ -\hat{x}(t-n) \ \hat{\bar{u}}(t-1) \ \hat{\bar{u}}(t-2) \ \cdots \ \hat{\bar{u}}(t-n) \ f(u(t))]^T \quad (200)$$

$$\hat{x}(t) = \hat{\boldsymbol{\varphi}}_s^T(t) [\hat{\mathbf{a}}^T(t) \ \hat{\mathbf{b}}^T(t) \ \hat{\mathbf{c}}^T(t)]^T, \quad (201)$$

$$\hat{\bar{u}}(t) = f(u(t)) \hat{\mathbf{c}}^T(t)]^T, \quad (202)$$

$$\hat{w}(t) = y(t) - \hat{x}(t), \quad (203)$$

$$f(u(t)) = [f_1(u(t)) \ f_2(u(t)) \ \cdots \ f_m(u(t))] , \quad (204)$$

$$\hat{\boldsymbol{\vartheta}}(t) = [\hat{\mathbf{a}}^T(t) \ \hat{\mathbf{b}}^T(t) \ \hat{\mathbf{c}}^T(t) \ \hat{\boldsymbol{\gamma}}^T(t)]^T. \quad (205)$$

4.4 输入非线性 Box-Jenkins 系统 (IN-BJ)

输入非线性 Box-Jenkins 系统 (Input Nonlinear Box-Jenkins system IN-BJ) 结构如图 20 所示, 其非线性部分方程为

$$\bar{u}(t) = f(u(t)) = c_1 f_1(u(t)) + c_2 f_2(u(t)) + \cdots + c_m f_m(u(t)) = f(u(t)) \mathbf{c}, \quad (206)$$

其中 $f(u(t)) := [f_1(u(t)) \ f_2(u(t)) \ \cdots \ f_m(u(t))] \in \mathbf{R}^{1 \times m}$ 是基函数构成的行向量 $\mathbf{c} := [c_1 \ c_2 \ \cdots \ c_m]^T \in \mathbf{R}^m$ 是非线性部分的参数向量. 线性部分的方程为

$$y(t) = \frac{B(z)}{A(z)} \bar{u}(t) + \frac{D(z)}{\gamma(z)} v(t), \quad (207)$$

其中 $A(z)$ 、 $B(z)$ 、 $D(z)$ 和 $\gamma(z)$ 是后移算子 z^{-1} 的多项式:

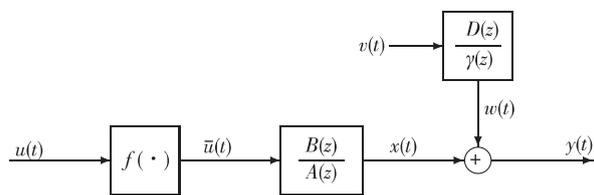


图 20 输入非线性 Box-Jenkins 系统 (IN-BJ)

Fig. 20 The input nonlinear Box-Jenkins system

$$A(z) := 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n},$$

$$B(z) := 1 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_n z^{-n},$$

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

$$\gamma(z) := 1 + \gamma_1 z^{-1} + \gamma_2 z^{-2} + \dots + \gamma_{n_\gamma} z^{-n_\gamma},$$

定义未知真实输出 $x(t)$ 和中间相关噪声变量

$w(t)$ 如下:

$$x(t) := \frac{B(z)}{A(z)} \bar{u}(t); \quad (208)$$

$$w(t) := \frac{D(z)}{\gamma(z)} v(t). \quad (209)$$

参数向量 ϑ 和信息向量 $\varphi(t)$ 定义如下:

$$\vartheta = \begin{bmatrix} a \\ b \\ c \\ \gamma \\ d \end{bmatrix} \in \mathbf{R}^{2n+m+n_\gamma+n_d}; \quad a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathbf{R}^n;$$

$$b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \in \mathbf{R}^n; \quad \gamma = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{bmatrix} \in \mathbf{R}^{n_\gamma};$$

$$d = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_{n_d} \end{bmatrix} \in \mathbf{R}^{n_d};$$

$$\varphi(t) := [-x(t-1), -x(t-2), \dots, -x(t-n), \bar{u}(t-1), \bar{u}(t-2), \dots, \bar{u}(t-n), f(u(t)), -w(t-1), -w(t-2), \dots, -w(t-n_\gamma), v(t-1), v(t-2), \dots, v(t-n_d)]^T \in \mathbf{R}^{2n+m+n_\gamma+n_d}.$$

式 (208) — (209) 可以写为

$$\begin{aligned} x(t) &= [1 - A(z)]x(t) + [B(z) - 1]\bar{u}(t) + \bar{u}(t) = \\ &= -\sum_{i=1}^n a_i x(t-i) + \sum_{i=1}^n b_i \bar{u}(t-i) + f(u(t))c, \quad (210) \\ w(t) &= [1 - \gamma(z)]w(t) + [D(z) - 1]v(t) + v(t) = \end{aligned}$$

$$-\sum_{i=1}^{n_\gamma} \gamma_i w(t-i) + \sum_{i=1}^{n_d} d_i v(t-i) + v(t). \quad (211)$$

将式 (208) — (209) 代入式 (204), 利用式 (210) — (211) 可得辨识模型

$$y(t) = x(t) + w(t) = -\sum_{i=1}^n a_i x(t-i) + \sum_{i=1}^n b_i \bar{u}(t-i) +$$

$$f(u(t))c - \sum_{i=1}^{n_\gamma} \gamma_i w(t-i) +$$

$$\sum_{i=1}^{n_d} d_i v(t-i) + v(t) = \varphi^T(t)\vartheta + v(t). \quad (212)$$

信息向量 $\varphi(t)$ 中未知真实输出项 $x(t-i)$ 用辅助模型 (218) 输出 $\hat{x}(t-i)$ 代替, 未知中间变量 $\bar{u}(t-i)$ 用辅助模型 (219) 的输出 $\hat{\bar{u}}(t-i)$ 代替, 未知相关噪声项 $w(t-i)$ 用式 (220) 的估计 $\hat{w}(t-i)$ 代替, 未知噪声项 $v(t-i)$ 用是 (221) 中估计残差 $\hat{v}(t-i)$ 代替, 可以得到估计 IN-BJ 系统参数向量 ϑ 的辅助模型递推广义增广最小二乘辨识算法 (AM-RGELS):

$$\hat{\vartheta}(t) = \hat{\vartheta}(t-1) + L(t) [y(t) - \hat{\varphi}^T(t)\hat{\vartheta}(t-1)], \quad (213)$$

$$L(t) = P(t-1)\hat{\varphi}(t) [I + \hat{\varphi}^T(t)P(t-1)\hat{\varphi}(t)]^{-1}, \quad (214)$$

$$P(t) = [I - L(t)\hat{\varphi}^T(t)]P(t-1) \quad P(0) = p_0 I, \quad (215)$$

$$\hat{\varphi}(t) = [\hat{\varphi}_s^T(t), -\hat{w}(t-1), -\hat{w}(t-2), \dots, -\hat{w}(t-n_\gamma), \hat{\mu}(t-1), \hat{\mu}(t-2), \dots, \hat{\mu}(t-n_d)]^T, \quad (216)$$

$$\hat{\varphi}_s(t) = [-\hat{x}(t-1), -\hat{x}(t-2), \dots, -\hat{x}(t-n), \hat{\bar{u}}(t-1), \hat{\bar{u}}(t-2), \dots, \hat{\bar{u}}(t-n), f(u(t))]^T, \quad (217)$$

$$\hat{x}(t) = \hat{\varphi}_s^T(t) [\hat{a}^T(t) \hat{b}^T(t) \hat{c}^T(t)]^T, \quad (218)$$

$$\hat{\bar{u}}(t) = f(u(t))\hat{c}(t), \quad (219)$$

$$\hat{w}(t) = y(t) - \hat{x}(t), \quad (220)$$

$$\hat{v}(t) = y(t) - \hat{\varphi}^T(t)\hat{\vartheta}(t), \quad (221)$$

$$f(u(t)) = [f_1(u(t)) \ f_2(u(t)) \ \dots \ f_m(u(t))], \quad (222)$$

$$\hat{\vartheta}(t) = [\hat{a}^T(t) \ \hat{b}^T(t) \ \hat{c}^T(t) \ \hat{\gamma}^T(t) \ \hat{d}^T(t)]^T. \quad (223)$$

本节讨论的一些输入非线性输出误差类系统 (IN-OEMA, IN-OEAR, IN-BJ) 的辅助模型递推最小二乘算法的收敛性分析 (参数估计的一致收敛性、估计误差的有界收敛性) 仍然是控制科学家有待解决的辨识难题.

4.5 其他输入非线性系统

上述讨论的输入非线性系统假设输入端非线性特性 $\bar{u} = f(u)$ 是已知非线性基 $f = (f_1, f_2, \dots, f_m)$ 的线性参数函数, 即非线性特性可以写为参数 (c_1, c_2, \dots, c_m) 的线性组合形式

$$\begin{aligned} \bar{u}(t) &= f(u(t)) = \\ &= c_1 f_1(u(t)) + c_2 f_2(u(t)) + \dots + c_m f_m(u(t)). \end{aligned}$$

还有一些非线性特性,如预负载死区非线性、饱和非线性、滞后继电器非线性、滞后非线性等(参见图 21),其输出只能表示为输入的分段函数,难以写成像上式那样的单一表达式形式.例如,对于图 21b 中的饱和非线性特性,其输入输出关系如下

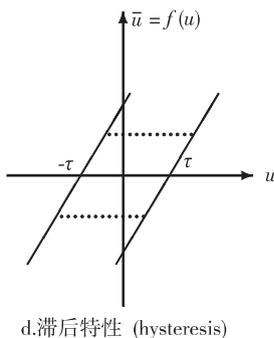
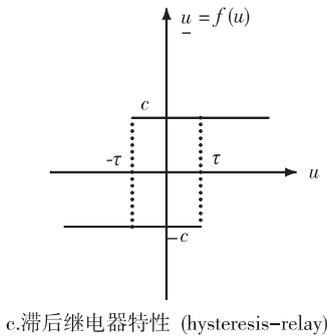
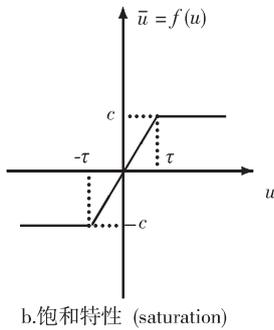
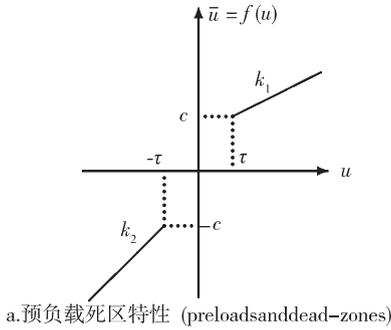


图 21 其他输入非线性特性
Fig. 21 Other input nonlinearities

$$\bar{u}(t) = \begin{cases} -c & \mu < -\tau, \\ \frac{c}{\tau}u, & -\tau \leq u < \tau, \\ c & \mu \geq \tau. \end{cases}$$

具有这样非线性特性的系统称为硬非线性系统 (systems with hard nonlinearities). 在这方面,陈晶等^[43]针对饱和和死区输入非线性 CAR 系统,引入开关函数和利用辅助模型辨识思想,提出了有限数据长度下的梯度迭代辨识算法;王冬青等^[23]借助于辅助模型辨识思想和关键项分离原理,提出了分段非线性 IN-OEAR 系统的辅助模型递推广义最小二乘辨识算法.

由于非线性特性形状各异,不可能存在通用的辨识方法,必须针对不同的非线性特性,采用不同技巧,研究相应的递推辨识方法或迭代辨识方法.这也给辨识科学家提出了许多研究课题.

5 输出非线性输出误差类系统(ON-OET)

输出非线性系统类型要复杂得多,难以分为输出非线性方程误差类系统(Output Nonlinear Equation Error Type system, ON-EET)和输出非线性输出误差类系统(Output Nonlinear Output Error Type system, ON-OET) 2 类.基本的输出非线性输出误差系统(Output Nonlinear Output Error system, ON-OE)如图 22 所示.其中 $u(t)$ 为系统输入, $x(t)$ 为线性部分的输出(不可测中间变量), $\bar{y}(t)$ 为非线性部分的输出(未知的), $p(t)$ 为量测噪声(可以假定是零均值和有限方差的), $y(t)$ 是系统输出,即 $\bar{y}(t)$ 的含噪量测, $f(\cdot)$ 是输出端的非线性函数, $G(z) = \frac{B(z)}{A(z)}$ 是线性部分的传递函数,其中 $A(z)$ 和 $B(z)$ 是后移算子 z^{-1} 的多项式:

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

$$B(z) = b_0 + b_1z^{-1} + b_2z^{-2} + \dots + b_nz^{-n}.$$

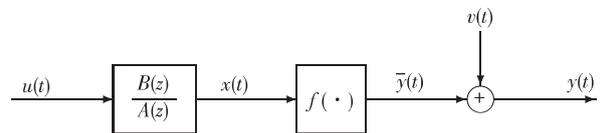


图 22 输出非线性输出误差系统(ON-OE)

Fig. 22 The output nonlinear output error system

这个 ON-OE 系统可以表示为

$$x(t) = \frac{B(z)}{A(z)}u(t), \quad (224)$$

$$\bar{y}(t) = f(x(t)), \quad (225)$$

$$y(t) = \bar{y}(t) + v(t) = f(x(t)) + v(t). \quad (226)$$

定义参数向量 θ 和信息向量 $\varphi(t)$ 如下:

$$\begin{aligned} \theta &= [a_1 \ a_2 \ \dots \ a_n \ b_0 \ b_1 \ b_2 \ \dots \ b_n]^T \in \mathbf{R}^{2n+1}, \\ \varphi(t) &:= [-x(t-1) \ -x(t-2) \ \dots \ -x(t-n) \ \mu(t) \ \mu(t-1) \ \mu(t-2) \ \dots \ \mu(t-n)]^T \in \mathbf{R}^{2n+1}. \end{aligned}$$

式(229)可以写为

$$A(z)x(t) = B(z)u(t)$$

或

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

因此,ON-OE 模型可以等价表示为

$$y(t) = f(x(t)) + v(t), \quad (227)$$

$$x(t) = \varphi^T(t)\theta. \quad (228)$$

对于这样一个简单的输出非线性输出误差系统,即使假定非线性函数是一个二次函数

$$\bar{y} = f(x) = x^2,$$

或

$$\bar{y}(t) = f(x(t)) = x^2(t),$$

其辨识问题也是不简单的。

当然,也可以假设这个输出非线性特性 $\bar{y} = f(x)$ 是已知非线性基 $f = (f_1 \ f_2 \ \dots \ f_m)$ 的线性函数,即非线性特性可以写为参数 $(c_1 \ c_2 \ \dots \ c_m)$ 的线性组合形式,

$$\begin{aligned} \bar{y}(t) = f(x(t)) &= \\ c_1 f_1(x(t)) + c_2 f_2(x(t)) + \dots + c_m f_m(x(t)) &= \\ f(x(t))c, \end{aligned}$$

其中 $f(x(t)) := [f_1(x(t)) \ f_2(x(t)) \ \dots \ f_m(x(t))] \in \mathbf{R}^{1 \times m}$ 是基函数构成的行向量 $c := [c_1 \ c_2 \ \dots \ c_m]^T \in \mathbf{R}^m$ 是非线性部分的参数向量。

对于输出非线性系统,也需要规范化模型参数,规范化方法有:1) 固定 b_i 中的一个,或者固定 c_j 中的一个;2) 设 $(b_0 \ b_1 \ b_2 \ \dots \ b_n)$ 或 $(c_1 \ c_2 \ \dots \ c_m)$ 的模为 1,即 $b_0^2 + b_1^2 + b_2^2 + \dots + b_n^2 = 1$ 或 $c_1^2 + c_2^2 + \dots + c_m^2 = 1$;3) 设线性子系统的增益为 1,即 $G(1) = b_0 + b_1 + b_2 + \dots + b_n = 1$,或非线性函数的系数和为 1,即 $c_1 + c_2 + \dots + c_m = 1$ 。

具有已知非线性基的输入非线性 CAR 系统可以表示为

$$y(t) = [1 - A(z)]y(t) + B(z)f(u(t)) + v(t).$$

我们也可以仿照输入非线性系统的上述表达式,推广出一类输出非线性系统

$$y(t) = [1 - A(z)]f(y(t)) + B(z)u(t) + v(t) =$$

$$\begin{aligned} &- a_1 f(y(t-1)) - a_2 f(y(t-2)) - \dots - \\ &a_n f(y(t-n)) + b_0 u(t) + b_1 u(t-1) + \dots + \\ &b_n u(t-n) + v(t). \end{aligned}$$

或更一般形式,

$$\begin{aligned} y(t) &= -a_1 f_1(y(t-1)) - a_2 f_2(y(t-2)) - \dots - \\ &a_n f_n(y(t-n)) + b_0 u(t) + b_1 u(t-1) + \dots + \\ &b_n u(t-n) + v(t). \end{aligned}$$

以上 2 式是输出非线性 CAR 系统. 输入输出非线性系统(2 个静态非线性夹着一个线性子系统),即 Hammerstein-Wiener 非线性系统的一种形式

$$\begin{aligned} y(t) &= -a_1 f_1(y(t-1)) - a_2 f_2(y(t-2)) - \dots - \\ &a_n f_n(y(t-n)) + b_0 h_0(u(t)) + \\ &b_1 h_1(u(t-1)) + \dots + b_n h_n(u(t-n)) + v(t), \end{aligned}$$

其中 $h = (h_0 \ h_1 \ \dots \ h_n)$ 是已知非线性基数函数. 至于其他的输出非线性系统,我们将在非线性系统辨识的系列论文中介绍。

王冬青等^[44]针对一类 Hammerstein-Wiener AR-MAX 非线性系统,提出了增广随机梯度辨识算法。

6 结语

系统辨识是研究建立系统数学模型的理论与方法. 尽管建立系统数学模型有很多方法,但是对于存在不可测变量的系统,其辨识就更加困难,必须探索新的辨识方法. 本文主要讨论了一些辅助模型最小二乘辨识方法. 需要指出的是,随机梯度辨识方法、多新息随机梯度辨识方法^[45-48]、多新息最小二乘辨识方法^[35]、有限数据窗最小二乘算法^[49-50]等都可以用于本文系统的辨识,得出相应的辨识方法. 这些方法的性能分析也是有待研究的辨识难题。

参考文献

References

- [1] 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 & Technology: Natural Science Edition 2011 3(1):1-22
- [2] 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 & Technology: Natural Science Edition 2011 3(2):97-117
- [3] 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 & Technology: Natural Science Edition 2011, 3(3): 193-226
- [4] 丁锋. 基于输出估计的多输入系统随机梯度估计算法 [J]. 南京信息工程大学学报: 自然科学版, 2010, 2(6): 481-488
DING Feng. Stochastic gradient estimation algorithm for multiple-input systems based on the output estimation [J]. Journal of Nanjing University of Information Science & Technology: Natural Science Edition, 2010, 2(6): 481-488
- [5] 万百五. 控制论创立六十年 [J]. 控制理论与应用, 2008, 25(4): 597-602
WAN Baiwu. Sixty years of cybernetics since founding [J]. Control Theory and Applications 2008, 25(4): 597-602
- [6] 廖晓昕. 漫谈 Lyapunov 稳定性的理论、方法和应用 [J]. 南京信息工程大学学报: 自然科学版, 2009, 1(1): 1-15
LIAO Xiaoxin. Talking on the theory, methods and applications of Lyapunov stability [J]. Journal of Nanjing University of Information Science & Technology: Natural Science Edition 2009, 1(1): 1-15
- [7] 廖晓昕. 综合国力非线性扩散模型稳定性分析 [J]. 南京信息工程大学学报: 自然科学版, 2009, 1(3): 247-251
LIAO Xiaoxin. Stability analysis of nonlinear diffusion model describing comprehensive national power [J]. Journal of Nanjing University of Information Science & Technology: Natural Science Edition, 2009, 1(3): 247-251
- [8] 丁锋. 多变量系统辨识 [D]. 北京: 清华大学自动化系, 1990
DING Feng. Identification of multivariable systems [D]. Beijing: Department of Automation, Tsinghua University, 1990
- [9] 丁锋, 谢新民. 传递函数阵子模型参数递推估计: 辅助模型方法 [J]. 控制与决策, 1991, 6(6): 447-452
DING Feng, XIE Xinmin. Recursive estimation of parameters of transfer function matrix sub-submodels: Instrumental model method [J]. Control and Decision, 1991, 6(6): 447-452
- [10] 丁锋, 谢新民. 多变量系统的辅助模型辨识算法 [J]. 清华大学学报: 自然科学版, 1992, 32(4): 100-106
DING Feng, XIE Xinmin. Instrumental model identification algorithm for multivariable systems [J]. Journal of Tsinghua University: Science and Technology, 1992, 32(4): 100-106
- [11] 丁锋. 多变量系统的辅助模型辨识方法的收敛性分析 [J]. 控制理论与应用, 1997, 14(2): 192-200
DING Feng. Convergence analysis of the auxiliary model identification algorithm for multivariable systems [J]. Control Theory and Applications, 1997, 14(2): 192-200
- [12] Ding F, Chen T. Combined parameter and output estimation of dual-rate systems using an auxiliary model [J]. Automatica 2004, 40(10): 1739-1748
- [13] Ding F, Liu G J, Liu X P. Parameter estimation with scarce measurements [J]. Automatica, 2011, 47(8): 1645-1655
- [14] Ding F, Chen T. Parameter estimation of dual-rate stochastic systems by using an output error method [J]. IEEE Transactions on Automatic Control, 2005, 50(9): 1436-1441
- [15] Ding F, Ding J. Least squares parameter estimation for systems with irregularly missing data [J]. International Journal of Adaptive Control and Signal Processing, 2010, 24(7): 540-553
- [16] Ding F, Shi Y, Chen T. Auxiliary model based least-squares identification methods for Hammerstein output-error systems [J]. Systems & Control Letters, 2007, 56(5): 373-380
- [17] Ding F, Chen T. Identification of dual-rate systems based on finite impulse response models [J]. International Journal of Adaptive Control and Signal Processing, 2004, 18(7): 589-598
- [18] 丁锋. 系统辨识理方法论 [M]. 北京: 中国电力出版社, 2012
DING Feng. System identification theory and methods [M]. Beijing: China Electric Power Press, 2012
- [19] Ding F, Liu X P, Liu G J. Gradient based and least-squares based iterative identification methods for OE and OEMA systems [J]. Digital Signal Processing, 2010, 20(3): 664-677
- [20] Wang D Q, Ding F. Performance analysis of the auxiliary models based multi-innovation stochastic gradient estimation algorithm for output error systems [J]. Digital Signal Processing, 2010, 20(3): 750-762
- [21] Liu Y J, Xiao Y S, Zhao X L. Multi-innovation stochastic gradient algorithm for multiple-input single-output systems using the auxiliary model [J]. Applied Mathematics and Computation, 2009, 215(4): 1477-1483
- [22] Wang D Q, Chu Y Y, Ding F. Auxiliary model-based RELS and MI-ELS algorithms for Hammerstein OEMA systems [J]. Computers & Mathematics with Applications, 2010, 59(9): 3092-3098
- [23] Wang D Q, Chu Y Y, Yang G W, et al. Auxiliary model-based recursive generalized least squares parameter estimation for Hammerstein OEAR systems [J]. Mathematical and Computer Modelling, 2010, 52(1-2): 309-317
- [24] Ding F, Liu X P, Liu G. Auxiliary model based multi-innovation extended stochastic gradient parameter estimation with colored measurement noises [J]. Signal Processing, 2009, 89(10): 1883-1890
- [25] Han L L, Sheng J, Ding F, et al. Auxiliary model identification method for multirate multi-input systems based on least squares [J]. Mathematical and Computer Modelling, 2009, 50(7/8): 1100-1106
- [26] 王冬青. 基于辅助模型的递推增广最小二乘辨识方法 [J]. 控制理论与应用, 2009, 26(1): 51-56
WANG Dongqing. Recursive extended least squares identification method based on auxiliary models [J]. Control

- Theory and Applications 2009 26(1):51-56
- [27] 刘艳君, 谢莉, 丁锋. 非均匀采样数据系统的 AM-RLS 辨识方法及仿真研究 [J]. 系统仿真学报, 2009, 21(19):6186-6189
LIU Yanjun, XIE Li, DING Feng. AM-RLS identification and simulation studies for non-uniformly sampled-data systems [J]. Journal of System Simulation, 2009, 21(19):6186-6189
- [28] Liu Y J, Xie L, Ding F. An auxiliary model based on a recursive least-squares parameter estimation algorithm for non-uniformly sampled multirate systems [J]. Proceedings of the Institution of Mechanical Engineers. Part I: Journal of Systems and Control Engineering, 2009, 223(4):445-454
- [29] 王冬青, 丁锋. 基于辅助模型的多新息广义增广随机梯度算法 [J]. 控制与决策, 2008, 23(9):999-1003, 1010
WANG Dongqing, DING Feng. Auxiliary models based multi-innovation generalized extended stochastic gradient algorithms [J]. Control and Decision 2008 23(9):999-1003, 1010
- [30] Ding F, Chen H B, Li M. Multi-innovation least squares identification methods based on the auxiliary model for MISO systems [J]. Applied Mathematics and Computation 2007 187(2):658-668
- [31] 谢莉, 王冬青, 丁锋. 随机干扰系统的辅助模型递推广义增广最小二乘辨识方法 [J]. 科学技术与工程, 2008 8(14):3944-3945, 3965
XIE Li, WANG Dongqing, DING Feng. AM-RGELS algorithms for general stochastic systems [J]. Science Technology and Engineering 2008 8(14):3944-3945, 3965
- [32] 丁锋, 杨家本. 输入输出系统噪信比的计算 [J]. 清华大学学报:自然科学版, 1998 38(9):107-110
DING Feng, YANG Jiaben. Computation formulas of the noise-to-signal ratio of input-output systems [J]. Journal of Tsinghua University: Science and Technology, 1998 38(9):107-110
- [33] 丁锋, 杨慧中, 刘飞. 弱条件下随机梯度算法性能分析 [J]. 中国科学 E 辑: 信息科学, 2008, 38(12):2173-2184
- [34] Ding F, Yang H Z, Liu F. Performance analysis of stochastic gradient algorithms under weak conditions [J]. Science in China Series F-Information Sciences 2008 51(9):1269-1280
- [35] 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
- [36] Liu Y J, Wang D Q, Ding F. Least-squares based iterative algorithms for identifying Box-Jenkins models with finite measurement data [J]. Digital Signal Processing, 2010, 20(5):1458-1467
- [37] Wang D Q, Yang G W, Ding R F. Gradient-based iterative parameter estimation for Box-Jenkins systems [J]. Computers and Mathematics with Applications, 2010, 60(5):1200-1208
- [38] Xie L, Liu Y J, Yang H Z, et al. Modeling and identification for non-uniformly periodically sampled-data systems [J]. IET Control Theory & Applications, 2010, 4(5):784-794
- [39] Ding F, Liu P X, Liu G J. Identification methods for Hammerstein nonlinear systems [J]. Digital Signal Processing 2011 21(2):215-238
- [40] Ding F, Chen T. Identification of Hammerstein nonlinear ARMAX systems [J]. Automatica, 2005, 41(9):1479-1489
- [41] Ding F, Shi Y, Chen T. Gradient-based identification methods for Hammerstein nonlinear ARMAX models [J]. Nonlinear Dynamics 2006 45(1-2):31-43
- [42] Chen J, Zhang Y, Ding R F. Auxiliary model based multi-innovation algorithms for multivariable nonlinear systems [J]. Mathematical and Computer Modelling 2010 52(9-10):1428-1434
- [43] Chen J, Wang X P, Ding R F. Gradient based estimation algorithm for Hammerstein systems with saturation and dead-zone nonlinearities [J]. Applied Mathematical Modelling 2011(in press)
- [44] Wang D Q, Ding F. Extended stochastic gradient identification algorithms for Hammerstein-Wiener ARMAX systems [J]. Computers & Mathematics with Applications, 2008 56(12):3157-3164
- [45] 丁锋, 谢新民, 方崇智. 时变系统辨识的多新息方法 [J]. 自动化学报, 1996 22(1):85-91
DING Feng, XIE Xinmin, FANG Chongzhi. Multi-innovation identification methods for time-varying systems [J]. Acta Automatica Sinica, 1996 22(1):85-91
- [46] 丁锋, 萧德云, 丁韬. 多新息随机梯度辨识方法 [J]. 控制理论与应用 2003 20(6):870-874
DING Feng, XIAO Deyun, DING Tao. Multi-innovation stochastic gradient identification methods [J]. Control Theory and Applications 2003 20(6):870-874
- [47] Ding F, Chen T. Performance analysis of multi-innovation gradient type identification methods [J]. Automatica, 2007 43(1):1-14
- [48] Ding F. Several multi-innovation identification methods [J]. Digital Signal Processing 2010 20(4):1027-1039
- [49] 丁锋, 丁韬, 萧德云, 等. 时变系统有限数据窗最小二乘辨识的有界收敛性 [J]. 自动化学报, 2002 28(5):754-761
DING Feng, DING Tao, XIAO Deyun, et al. Bounded convergence of finite data window least squares identification for time-varying systems [J]. Acta Automatica Sinica, 2002 28(5):754-761
- [50] 丁锋, 杨慧中, 纪志成. 时变系统辨识方法及其收敛定理 [J]. 江南大学学报:自然科学版, 2006, 5(1):115-126
DING Feng, YANG Huizhong, JI Zhicheng. Time-varying system identification methods and convergence theorems [J]. Journal of Southern Yangtze University: Natural Science Edition 2006 5(1):115-126

System identification. Part D: Auxiliary model identification idea and methods

DING Feng^{1, 2, 3}

1 School of Internet of Things Engineering, Jiangnan University, Wuxi 214122

2 Control Science and Engineering Research Center, Jiangnan University, Wuxi 214122

3 Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi 214122

Abstract The auxiliary model identification idea, the multi-innovation identification theory, the hierarchical identification principle and the coupled identification concept are new methods for studying identification problems proposed by the author of this article. They have been applied to many identification researches and resulted in different identification method families and can be used to solve adaptive signal processing, adaptive parameter estimation, adaptive filtering and prediction, adaptive control and other issues for many linear or nonlinear models. Any objective thing has a dual property: some of the characteristics variables are observable and some are unmeasurable. For the observed system, it is easy to set up the mathematical model for describing its law of motion. The unmeasurable property of thing brings us particular difficulties for setting up the mathematical model. In this case, how to use the measured information of the system to estimate the unknown variables of a system and then to establish the mathematical model of the system is a challenging research topic in the area of system identification, which is the root of the auxiliary model identification idea. This article introduces the auxiliary model identification idea and some auxiliary model based identification methods.

Key words auxiliary model; recursive identification; parameter estimation; FIR model; CAR model; CARMA model; CARAR model; CARARMA model; output error model; OEMA model; OEAR model; multi-innovation identification; hierarchical identification; coupled identification