

# 系统辨识(2):系统描述的基本模型

丁锋<sup>1,2</sup>

## 摘要

控制是一切科学问题的核心.数学模型是一切控制问题的基础.事物的运动规律用方程描述就是数学模型.不同学科的发展就是建立其数学模型的过程.本文首次把线性动态系统数学模型分为三类:时间序列模型,方程误差类模型,输出误差类模型;详细介绍了线性系统的一些基本数学模型,包括连续系统离散化和模型等价变换,单输入单输出随机系统模型,多变量系统模型,类多变量系统模型,多输入和多输出系统模型(传递函数阵主模型、子模型、子子模型,多输入单输出系统模型,单输入多输出系统模型等).

## 关键词

数学模型;系统辨识;参数估计;CAR模型;ARX模型;CARMA模型;ARMAX模型;CARARMA模型;OEMA模型;OEAR模型

中图分类号 TP273

文献标志码 A

## 0 引言

现代科学文明的标志之一是控制理论与自动化技术的飞速发展.控制论与自动化的辉煌成就——高集成度计算机芯片的生产和广泛应用,就是控制理论与自动化技术发展的极致和光辉典范<sup>[1-2]</sup>.控制论对社会和经济发展起到了巨大的推动作用<sup>[3-5]</sup>.控制是许多科学问题的核心,一架飞机(飞行器)升天执行任务,就是一个控制问题.尽管材料很重要,但对于一座化工厂或一座炼油厂,一旦设备安装好,剩余的生产都是一个自动化控制问题,而且生产这些设备的过程也是一个自动化控制问题.

当今物联网技术(传感网技术)是控制理论与自动化技术发展的最高体现,而自动化就是对客观事物的感知、检测、传输、信息采集、加工处理,实现对客观事物的控制,达到认识自然,改造自然,为人类造福的目的.

万物离不开控制.科学技术的发展离不开计算,离不开控制.计算离不开方程,离不开模型,计算机更离不开模型.数学模型是控制理论与自动化的基础.事物的运动规律用方程描述就是数学模型.

数学模型有静态与动态之分,随时间演变的模型称为动态模型,否则称为静态模型.数学模型有线性与非线性之分,输入输出满足叠加原理的模型称为线性模型,否则称为非线性模型.数学模型有确定性与随机性之分,模型中变量受到随机干扰的系统称为随机系统(随机模型),否则称为确定性系统(确定性模型).本文首次把随机系统分为“时间序列模型”、“方程误差类模型”和“输出误差类模型”3大类,使系统模型的定义和辨识算法的类别变得十分清晰.在此基础上,详细介绍线性系统的基本数学模型,包括连续系统离散化方法和模型等价变换(阶跃不变变换、双线性变换和欧拉变换、脉冲不变  $z$ - $s$  变换、离散状态空间模型化为差分方程模型),单输入单输出随机系统、多变量系统、类多变量系统、多输入和多输出系统(时间序列模型、方程误差模型和输出误差类模型),以及传递函数阵主模型、子模型、子子模型,多输入单输出系统模型,单输入多输出系统模型等.

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 线性系统模型变换

#### 1.1 阶跃不变变换

收稿日期 2011-04-06

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](mailto:fding@jiangnan.edu.cn)

1 江南大学 物联网工程学院,无锡,214122

2 江南大学 控制科学与工程研究中心,无锡,214122

- 1.2 双线性变换和欧拉变换
- 1.3 脉冲不变  $z$ - $s$  变换
- 1.4 离散模型化为差分方程模型
- 2 随机系统模型
  - 2.1 时间序列模型
  - 2.2 方程误差类模型
  - 2.3 输出误差类模型
  - 2.4 特殊方程误差模型
  - 2.5 特殊输出误差模型
  - 2.6 一般随机系统模型
- 3 多变量系统
  - 3.1 多变量时间序列模型
  - 3.2 多变量方程误差类模型
  - 3.3 多变量输出误差类模型
  - 3.4 特殊多变量方程误差模型
  - 3.5 特殊多变量输出误差模型
  - 3.6 一般多变量随机系统模型
- 4 类多变量系统
  - 4.1 状态空间描述到输入输出表达
  - 4.2 类多变量方程误差类模型
  - 4.3 类多变量输出误差类模型
  - 4.4 类特殊多变量方程误差模型
  - 4.5 类特殊多变量输出误差模型
  - 4.6 一般类多变量随机系统模型
- 5 多输入和多输出系统
  - 5.1 传递函数阵主模型
  - 5.2 传递函数阵子模型
  - 5.3 传递函数阵主子模型
  - 5.4 多输入单输出系统模型
  - 5.5 单输入多输出系统模型
  - 5.6 马可夫参数或脉冲响应模型
- 6 结语

## 1 线性系统模型变换

特作申明:本文中有的变量上下文使用了不同的定义,只要一个变量重新定义了,就采用新的定义。“ $A = :X$ ”或“ $X := A$ ”表示“ $A$  记作(定义为)  $X$ ”之意。

(线性)系统可以用多种模型描述,如微分方程模型,状态空间模型,传递函数模型等。尽管连续系统辨识方法也有不少研究,但是由于数字计算机的使用,使得离散系统辨识方法在理论与应用上都占主流。

本节介绍连续系统离散化、阶跃不变变换、双线性变换、欧拉变换,以及本文作者提出的广义双线性变换和  $z$ - $s$  变换以及离散模型化为差分方程模型。

假设单输入单输出连续时间系统的状态空间模型(state space model)描述为

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t), & \mathbf{x}(t_0) = \mathbf{x}_0, \\ y(t) = \mathbf{c}\mathbf{x}(t) + \mathbf{d}u(t), \end{cases} \quad (1)$$

其中  $\mathbf{x}(t) \in \mathbf{R}^n$  是系统的( $n$  维)状态变量(state variable),  $u(t)$  和  $y(t)$  分别为系统的输入和输出,  $\mathbf{A} \in \mathbf{R}^{n \times n}$ ,  $\mathbf{b} \in \mathbf{R}^n$ ,  $\mathbf{c} \in \mathbf{R}^{1 \times n}$  和  $\mathbf{d} \in \mathbf{R}$  均为常数系统矩阵(system matrices)。状态空间模型(1)常记作  $[\mathbf{A}, \mathbf{b}, \mathbf{c}, \mathbf{d}]$ 。方程(1)的第1式称为状态方程(state equation),第2式称为输出方程(output equation)。

### 1.1 阶跃不变变换

下面先推导状态方程的解,然后给出阶跃不变变换。方程(1)的第1式两边左乘以  $e^{-At}$  移项可得

$$e^{-At}\dot{\mathbf{x}}(t) - e^{-At}\mathbf{A}\mathbf{x}(t) = e^{-At}\mathbf{b}u(t)$$

或

$$d[e^{-At}\mathbf{x}(t)] = e^{-At}\mathbf{b}u(t)dt.$$

上式两边从  $t_0$  到  $t$  积分可得

$$e^{-At}\mathbf{x}(t) - e^{-At_0}\mathbf{x}(t_0) = \int_{t_0}^t e^{-A\tau}\mathbf{b}u(\tau)d\tau.$$

两边左乘  $e^{At}$ , 移项得到状态解和输出解为

$$\mathbf{x}(t) = e^{A(t-t_0)}\mathbf{x}(t_0) + \int_{t_0}^t e^{A(t-\tau)}\mathbf{b}u(\tau)d\tau, \quad (2)$$

$$y(t) = \mathbf{c}\mathbf{x}(t) + \mathbf{d}u(t) =$$

$$\mathbf{c}e^{A(t-t_0)}\mathbf{x}(t_0) + \mathbf{c} \int_{t_0}^t e^{A(t-\tau)}\mathbf{b}u(\tau)d\tau + \mathbf{d}u(t), \quad t \geq t_0. \quad (3)$$

由于同一系统可以用不同模型描述,为了揭示模型之间的联系,或是人们出于某种需要,经常要进行模型等价变换(model equivalence transform)。在 Matlab 中,有些函数直接实现模型之间相互转换,如 c2d 把连续系统转换为离散系统, d2c 把离散系统转换为连续系统, tf2ss 把传递函数转换为状态空间模型, ss2tf 把状态空间模型转化为传递函数模型等。

连续系统离散化,需要一定的假设。在不同假设下,得到的离散模型是不同的。如果输入采用方波信号,即输入端采用零阶保持器,得到离散系统的方法称为阶跃响应不变变换,此外还有脉冲响应不变变换、斜坡响应不变变换、双线性变换等。这些方法可借助于 Matlab 的帮助命令“help c2d”查询其功能。常用的离散化方法是阶跃响应不变变换,它保证离散系统的输出等于连续系统在采样点的输出值,因而在计算机控制系统中广泛使用。

设采样周期(sampling period)为  $T$ , 采用零阶保持器(zero-order hold), 即  $u(t) = u(kT), kT \leq t < (k+1)T$ . 记  $\mathbf{x}(kT) = : \mathbf{x}(k), u(kT) = : u(k), y(kT) = : y(k)$ , 则连续时间状态空间模型(1)对应的离散时间状态空间模型为

$$\begin{cases} \mathbf{x}(k+1) = \mathbf{G}\mathbf{x}(k) + \mathbf{f}u(k), \\ y(k) = \mathbf{c}\mathbf{x}(k) + \mathbf{d}u(k). \end{cases} \quad (4)$$

其中:

$$\mathbf{G} := e^{A^T}, \quad \mathbf{f} := \int_0^T e^{A^t} dt \mathbf{b} = \mathbf{A}^{-1}[\mathbf{G} - \mathbf{I}]\mathbf{b}. \quad (5)$$

这里  $\mathbf{I}$  为适当维数的单位阵, 最后一个等式在  $\mathbf{A}$  可逆时成立.

**证明** 将  $t_0 = kT$  和  $t = (k+1)T$  代入式(2), 可得

$$\begin{aligned} \mathbf{x}(k+1) &= e^{A^T} \mathbf{x}(k) + \int_{kT}^{(k+1)T} e^{A(kT+T-\tau)} \mathbf{b}u(\tau) d\tau = \\ &e^{A^T} \mathbf{x}(k) + \int_{kT}^{(k+1)T} e^{A(kT+T-\tau)} d\tau \mathbf{b}u(k). \end{aligned}$$

进行变量置换:  $t = kT + T - \tau$ , 可得

$$\begin{aligned} \mathbf{x}(k+1) &= e^{A^T} \mathbf{x}(k) - \int_T^0 e^{A^t} dt \mathbf{b}u(k) = \\ &e^{A^T} \mathbf{x}(k) + \int_0^T e^{A^t} dt \mathbf{b}u(k) = \mathbf{G}\mathbf{x}(k) + \mathbf{f}u(k), \\ y(k) &= \mathbf{c}\mathbf{x}(k) + \mathbf{d}u(k). \end{aligned}$$

由此直接得到关系式(5), 证毕.

从式(1)到式(4)的变换称为阶跃响应不变变换, 简称为阶跃不变变换(step-invariance transform). 当然, 还有脉冲响应不变变换(impulse-invariance transform)、斜坡响应不变变换(slope-invariance transform)等.

**例 1** 试将下列状态空间模型离散化(discretization):

$$\begin{cases} \dot{\mathbf{x}}(t) = \begin{bmatrix} -1 & 0 \\ 1 & -2 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 1 \\ 2 \end{bmatrix} u(t), \\ y(t) = [2, 1] \mathbf{x}(t). \end{cases}$$

**解** 先求转移矩阵  $\Phi(t) = e^{At}$ ,

$$(s\mathbf{I} - \mathbf{A})^{-1} = \begin{bmatrix} s+1 & 0 \\ -1 & s+2 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{s+1} & 0 \\ \frac{1}{s+1} - \frac{1}{s+2} & \frac{1}{s+2} \end{bmatrix}.$$

进行逆 Laplace 变换得到

$$\Phi(t) = e^{At} = \mathcal{L}^{-1}[(s\mathbf{I} - \mathbf{A})^{-1}] = \begin{bmatrix} e^{-t} & 0 \\ e^{-t} - e^{-2t} & e^{-2t} \end{bmatrix}.$$

于是有

$$\mathbf{G} = e^{AT} = \begin{bmatrix} e^{-T} & 0 \\ e^{-T} - e^{-2T} & e^{-2T} \end{bmatrix},$$

$$\begin{aligned} \mathbf{f} &= \int_0^T e^{At} dt \mathbf{b} = \begin{bmatrix} e^{-t} & 0 \\ e^{-t} + \frac{1}{2}e^{-2t} & -\frac{1}{2}e^{-2t} \end{bmatrix}_{t=0}^{t=T} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \\ &\begin{bmatrix} 1 - e^{-T} & 0 \\ \frac{1}{2} - e^{-T} + \frac{1}{2}e^{-2T} & \frac{1}{2} - \frac{1}{2}e^{-2T} \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \\ &\begin{bmatrix} 1 - e^{-T} \\ \frac{3}{2} - e^{-T} - \frac{1}{2}e^{-2T} \end{bmatrix}. \end{aligned}$$

## 1.2 双线性变换和欧拉变换

### 1) 双线性变换

$Z$  变换算子  $z$  与拉普拉斯变换(Laplace transform)算子  $s$  间有如下关系

$$z = e^{Ts}.$$

其中  $T$  为采样周期. 用级数展开, 取其一次项作下列近似

$$z = e^{Ts} = \frac{e^{\frac{T}{2}s}}{e^{-\frac{T}{2}s}} \approx \frac{1 + \frac{T}{2}s}{1 - \frac{T}{2}s} = \frac{2 + Ts}{2 - Ts}, \quad (6)$$

或

$$s = \frac{2}{T} \frac{z-1}{z+1}. \quad (7)$$

式(6)和(7)即为双线性变换(bilinear transform). 因此,  $G(s)$  与  $G(z)$  之间的近似变换关系为

$$\begin{cases} G(z) = G(s) \Big|_{s = \frac{2}{T} \frac{z-1}{z+1}}, \\ G(s) = G(z) \Big|_{z = \frac{2+Ts}{2-Ts}}. \end{cases}$$

### 2) 欧拉变换

我们也可以作如下一次近似

$$z = e^{Ts} \approx 1 + Ts, \quad (8)$$

或

$$s = \frac{z-1}{T}. \quad (9)$$

式(8)和(9)即为欧拉变换(Euler transform).

### 3) 广义双线性变换

下面导出一种具有更普遍意义的近似变换<sup>[6-7]</sup>. 用级数展开取其一次项近似, 可得

$$z = e^{Ts} = \frac{e^{asT}}{e^{-(1-a)Ts}} \approx \frac{1 + aTs}{1 - (1-a)Ts}, \quad 0 \leq a \leq 1, \quad (10)$$

或

$$s = \frac{1}{T} \frac{z-1}{a + (1-a)z} = \frac{1}{(1-a)T} \frac{z-1}{z + \frac{a}{1-a}}. \quad (11)$$

式(10)和(11)是本文作者提出的广义双线性变换(generalized bilinear transform). 当  $a = 0.5$  时, 广义双线性变换退化为双线性变换.

当  $a = 0, a = \frac{1}{3}, a = \frac{1}{2}, a = \frac{2}{3}, a = 1$  和  $a \in (0, 1)$  时, 广义双线性变换如表 1 所示.

表 1 广义双线性变换表

Table 1 The generalized bilinear transform

序号	$a$ 值	变换式	名称
1	0	$z = \frac{1}{1 - Ts}$	$s = \frac{1}{T} \frac{z-1}{z}$
2	$\frac{1}{3}$	$z = \frac{3 + Ts}{3 - 2Ts}$	$s = \frac{3}{2T} \frac{z-1}{z + \frac{1}{2}}$
3	$\frac{1}{2}$	$z = \frac{2 + Ts}{2 - Ts}$	$s = \frac{2}{T} \frac{z-1}{z+1}$ 双线性变换
4	$\frac{2}{3}$	$z = \frac{3 + 2Ts}{3 - Ts}$	$s = \frac{3}{T} \frac{z-1}{z+2}$
5	1	$z = 1 + Ts$	$s = \frac{1}{T}(z-1)$ Euler 变换
6	$(0, 1)$	$z = \frac{1 + aTs}{1 - (1-a)Ts}$	$s = \frac{1}{(1-a)T} \frac{z-1}{z + \frac{1}{1-a}}$

双线性变换和欧拉变换的精度低、误差大(参见文献[8]), 下面讨论保持变换精度的  $z$ - $s$  变换.

### 1.3 脉冲不变 $z$ - $s$ 变换

双线性变换和欧拉变换的方法简单, 但精度低、误差大, 往往难以满足工程要求. 本文作者在其博士论文中, 根据香农采样定理(Shannon sampling theorem)和恢复定理推导了  $z$ - $s$  变换, 它可以保证连续时间模型与离散时间模型转换的精度<sup>[6-7]</sup>.

设  $G(s)$  为连续系统传递函数,  $G(z)$  为离散系统传递函数. 一般从  $G(s)$  获得  $G(z)$  的步骤如下:

$$G(s) \xrightarrow{\mathcal{L}^{-1}} g(t) \xrightarrow{t=kT} g(kT) \xrightarrow{Z} G(z). \quad (12)$$

其中  $\mathcal{L}^{-1}$  为拉普拉斯逆变换,  $T$  为采样周期,  $Z$  为  $Z$  变换. 实现过程: 先把  $G(s)$  进行拉普拉斯逆变换得到脉冲响应  $g(t)$ , 对  $g(t)$  进行采样, 得到  $g(kT)$  ( $k=0, 1, 2, \dots$ ), 再对  $g(kT)$  进行  $Z$  变换得到  $G(z)$ .

从  $G(z)$  获得  $G(s)$  的步骤如下:

$$G(z) \xrightarrow{Z^{-1}} g(kT) \xrightarrow{f(*)} g(t) \xrightarrow{\mathcal{L}} G(s). \quad (13)$$

其中  $Z^{-1}$  为逆  $Z$  变换(inverse  $Z$  transform),  $f(*)$  为内插函数(interpolation function),  $\mathcal{L}$  为拉普拉斯变换. 实现过程如下: 先把  $G(z)$  进行逆  $Z$  变换得到  $g(kT)$ , 再根据香农恢复定理, 进行插值得到  $g(t)$ , 再对  $g(t)$  进行拉普拉斯变换得到  $G(s)$ .

在一般情况下, 若给定  $G(s)$  和采样周期  $T$ , 由

$G(s)$  得到的  $G(z)$  是唯一的, 由  $G(z)$  得到的  $G(s)$  不是唯一的, 但我们希望得到一个通解的方法. 以上计算过程特别繁琐, 这里给出由  $G(s)$  直接求  $G(z)$ , 或由  $G(z)$  直接求  $G(s)$  的一种方法, 称为  $z$ - $s$  变换.

**定理 1**<sup>[6-7]</sup> 有理传递函数  $G(s)$  与离散传递函数  $G(z)$  构成下列  $z$ - $s$  变换对:

$$G(z) = \frac{1}{2\pi j} \oint_c G(s) \frac{z}{z - e^{Ts}} ds; \quad (14)$$

$$G(z) = \frac{1}{2\pi j} \oint_c \frac{G(s)}{z \left( s - \frac{1}{T} \ln z \right)} dz. \quad (15)$$

其中  $c$  是包围  $G(s)$  或  $G(z)$  所有极点(pole)的曲线. 用留数(residue)形式给出的  $z$ - $s$  变换对为

$$G(z) = \sum_{i=1}^m \frac{1}{(n_i - 1)!} \frac{d^{n_i-1}}{ds^{n_i-1}} \left[ (s - s_i)^{n_i} G(s) \frac{z}{z - e^{Ts}} \right]_{s=s_i} = \sum \left[ G(s) \frac{z}{z - e^{Ts}} \text{ 在 } G(s) \text{ 极点处的留数} \right], \quad (16)$$

$$G(z) = \sum_{i=1}^m \frac{1}{(n_i - 1)!} \frac{d^{n_i-1}}{dz^{n_i-1}} \left[ (z - z_i)^{n_i} \frac{G(z)}{z \left( s - \frac{1}{T} \ln z \right)} \right]_{z=z_i} = \sum \left[ \frac{G(z)}{z \left( s - \frac{1}{T} \ln z \right)} \text{ 在 } G(z) \text{ 极点处的留数} \right]. \quad (17)$$

其中  $s_i$  是  $G(s)$  的极点,  $z_i$  是  $G(z)$  的极点,  $n_i$  是极点的阶数.

以上是脉冲响应不变变换的  $z$ - $s$  变换对(简称脉冲不变  $z$ - $s$  变换). 阶跃响应不变变换的  $z$ - $s$  变换对(简称阶跃不变  $z$ - $s$  变换), 这里暂不作介绍. 只有当输入为脉冲信号时, 脉冲不变  $z$ - $s$  变换离散系统的输出等于连续系统在采样点的值.

**例 2** 设系统的传递函数为

$$G(s) = \frac{cs + d}{(s + a)(s + b)},$$

采样周期为  $T$ , 试求  $G(z)$ .

**解** 根据  $z$ - $s$  变换, 有

$$G(z) = \frac{1}{2\pi j} \oint_c G(s) \frac{z}{z - e^{Ts}} ds = \frac{1}{2\pi j} \oint_c \frac{cs + d}{(s + a)(s + b)} \frac{1}{1 - e^{Ts} z^{-1}} ds = \frac{-ac + d}{b - a} \frac{1}{1 - e^{-aT} z^{-1}} + \frac{-bc + d}{a - b} \frac{1}{1 - e^{-bT} z^{-1}} = \frac{1}{b - a} \frac{(b - a)c + [(ac - d)e^{-bT} + (d - bc)e^{-aT}] z^{-1}}{1 - (e^{-aT} + e^{-bT}) z^{-1} + e^{-(a+b)T} z^{-2}}.$$

当  $c = 0$  时, 有变换对:

$$G(s) = \frac{d}{(s+a)(s+b)};$$

$$G(z) = \frac{1}{b-a} \frac{d(e^{-aT} - e^{-bT})z^{-1}}{1 - (e^{-aT} + e^{-bT})z^{-1} + e^{-(a+b)T}z^{-2}}.$$

$$y(t) = \frac{\beta(z)}{\alpha(z)}u(t). \quad (23)$$

式(23)又可表示为算子多项式方程

$$\alpha(z)y(t) = \beta(z)u(t). \quad (24)$$

方程(24)被称为确定性自回归滑动平均模型 (Deterministic AutoRegressive Moving Average model, DARMA). 将多项式  $\alpha(z)$  和  $\beta(z)$  的定义式代入式(24)可得

$$(1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \dots + \alpha_n z^{-n})y(t) = (\beta_0 + \beta_1 z^{-1} + \beta_2 z^{-2} + \dots + \beta_n z^{-n})u(t).$$

展开得到差分方程 (difference equation),

$$y(t) + \alpha_1 y(t-1) + \alpha_2 y(t-2) + \dots + \alpha_n y(t-n) = \beta_0 u(t) + \beta_1 u(t-1) + \beta_2 u(t-2) + \beta_3 u(t-3) \dots + \beta_n u(t-n). \quad (25)$$

式(25)中  $\alpha_i$  和  $\beta_i$  称为系统参数 (system parameter),  $n$  称为系统阶次 (system order).

由于一般状态空间模型(18)中有  $n^2 + 2n + 1 = (n+1)^2$  个参数, 而差分方程模型(25)或传递函数表达(23)只有  $2n+1$  个参数, 规范状态空间规范型中也只有  $2n+1$  个参数, 因此, 一般把状态空间模型化成差分方程模型或状态空间规范型进行辨识.

### 例3 试将状态空间模型

$$\begin{cases} \mathbf{x}(t+1) = \begin{bmatrix} -1 & 1 \\ 1 & -2 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 1 \\ 2 \end{bmatrix} u(t) \\ y(t) = [2, 1] \mathbf{x}(t) \end{cases}$$

化为差分方程模型.

解 对于本例, 有

$$\mathbf{A} = \begin{bmatrix} -1 & 1 \\ 1 & -2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \mathbf{c} = [2, 1], \mathbf{d} = 0.$$

由于

$$(\mathbf{zI} - \mathbf{A})^{-1} = \begin{bmatrix} z+1 & -1 \\ -1 & z+2 \end{bmatrix}^{-1} =$$

$$\frac{1}{(z+1)(z+2) - 1} \begin{bmatrix} z+1 & 1 \\ 1 & z+1 \end{bmatrix} = \frac{\begin{bmatrix} z+2 & 1 \\ 1 & z+1 \end{bmatrix}}{z^2 + 3z + 1},$$

根据  $y(t) = [\mathbf{c}(\mathbf{zI} - \mathbf{A})^{-1} \mathbf{b} + \mathbf{d}]u(t)$ , 有

$$y(t) = [2, 1] \frac{\begin{bmatrix} z+2 & 1 \\ 1 & z+1 \end{bmatrix}}{z^2 + 3z + 1} \begin{bmatrix} 1 \\ 2 \end{bmatrix} u(t) = \frac{[2z+5, z+3]}{z^2 + 3z + 1} \begin{bmatrix} 1 \\ 2 \end{bmatrix} u(t) = \frac{4z+11}{z^2 + 3z + 1} u(t),$$

或

$$(z^2 + 3z + 1)y(t) = (4z + 11)u(t).$$

两边乘以  $z^{-2}$  得到

## 1.4 离散模型化为差分方程模型

在连续时间系统中, 用  $t$  表示连续时间变量; 在离散时间系统中, 用  $k$  表示离散时间变量, 表示变量在时刻  $t = kT$  取值 ( $T$  为采样周期). 在这里作一个约定: 都用  $t$  表示时间变量, 连续系统中  $t$  表示连续时间变量, 离散系统中  $t$  表示离散时间变量.

下面讨论化一般状态空间模型为差分方程的方法. 考虑单输入单输出时不变离散时间系统状态空间模型

$$\begin{cases} \mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t), \\ y(t) = \mathbf{c}\mathbf{x}(t) + \mathbf{d}u(t). \end{cases} \quad (18)$$

其中:  $\mathbf{x}(t) \in \mathbf{R}^n$  是系统的状态向量,  $u(t)$  和  $y(t)$  分别为系统的输入和输出,  $\mathbf{A} \in \mathbf{R}^{n \times n}$ ,  $\mathbf{b} \in \mathbf{R}^n$ ,  $\mathbf{c} \in \mathbf{R}^{1 \times n}$  和  $\mathbf{d} \in \mathbf{R}$  均为常数矩阵.

引入单位前移算子  $z$  或后移算子  $z^{-1}$ :  $zy(t) = y(t+1)$  和  $z^{-1}y(t) = y(t-1)$ . 式(18)用移位算子表示为

$$\begin{cases} \mathbf{z}\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t), \\ y(t) = \mathbf{c}\mathbf{x}(t) + \mathbf{d}u(t). \end{cases}$$

解得

$$\mathbf{x}(t) = (\mathbf{zI} - \mathbf{A})^{-1} \mathbf{b}u(t),$$

$$y(t) = [\mathbf{c}(\mathbf{zI} - \mathbf{A})^{-1} \mathbf{b} + \mathbf{d}]u(t) =: G(z)u(t). \quad (19)$$

从输入  $u(t)$  到输出  $y(t)$  的传递函数为

$$G(z) = \mathbf{c}(\mathbf{zI} - \mathbf{A})^{-1} \mathbf{b} + \mathbf{d} = \frac{\mathbf{c} \text{adj}[\mathbf{zI} - \mathbf{A}] \mathbf{b}}{\det[\mathbf{zI} - \mathbf{A}]} + \mathbf{d} = \frac{z^{-n} \mathbf{c} \text{adj}[\mathbf{zI} - \mathbf{A}] \mathbf{b} + \mathbf{d} z^n \det[\mathbf{zI} - \mathbf{A}]}{z^{-n} \det[\mathbf{zI} - \mathbf{A}]} \quad (20)$$

这揭示了系统传递函数  $G(z)$  与状态空间模型间的关系. 这里  $\text{adj}[\mathbf{zI} - \mathbf{A}]$  是  $(\mathbf{zI} - \mathbf{A})^{-1}$  的伴随矩阵 (adjoint matrix), 有下列关系

$$(\mathbf{zI} - \mathbf{A})^{-1} = \frac{\text{adj}[\mathbf{zI} - \mathbf{A}]}{\det[\mathbf{zI} - \mathbf{A}]}.$$

定义  $n$  次多项式:

$$\alpha(z) = z^{-n} \det[\mathbf{zI} - \mathbf{A}] = z^{-n} (z^n + \alpha_1 z^{n-1} + \alpha_2 z^{n-2} + \dots + \alpha_n) = 1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \dots + \alpha_n z^{-n}, \quad (21)$$

$$\beta(z) = z^{-n} \mathbf{c} \text{adj}[\mathbf{zI} - \mathbf{A}] \mathbf{b} + \mathbf{d} \mathbf{A}(z) = \beta_0 + \beta_1 z^{-1} + \beta_2 z^{-2} + \dots + \beta_n z^{-n}. \quad (22)$$

这里  $\alpha(z)$  为系统 (矩阵  $\mathbf{A}$ ) 的特征多项式. 使用上述定义和式(20), 由式(19)有

$$(1 + 3z^{-1} + z^{-2})y(t) = (4z^{-1} + 11z^{-2})u(t),$$

对应的差分方程为

$$y(t) + 3y(t-1) + y(t-2) = 4u(t-1) + 11u(t-2).$$

实际系统经常存在这样或那样的干扰作用,且这些干扰往往具有随机性,通常把这些干扰统称为随机噪声.上面讨论了 DARMA 模型是一类确定性系统模型(deterministic system model),当把干扰噪声项加到确定性模型上,就构成了系统的随机模型,用随机模型描述的系统称为随机系统模型(stochastic system model).由于存在干扰项,对随机系统而言,经常使用的术语是系统模型(确定性部分)和噪声模型(noise model)(干扰部分).

## 2 随机系统模型

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,  $\{v(t)\}$  是零均值随机白噪声序列,  $z^{-1}$  为单位后移算子:  $z^{-1}y(t) = y(t-1)$  或  $zy(t) = y(t+1)$ ,  $A(z)$ ,  $B(z)$ ,  $C(z)$ ,  $D(z)$ ,  $F(z)$  是算子  $z^{-1}$  的常数项不变多项式,定义如下:

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

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

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

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

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

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

$$A(z)y(t) =$$

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

$$y(t) + a_1y(t-1) + a_2y(t-2) + \cdots + a_{n_a}y(t-n_a),$$

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

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

$$b_1u(t-1) + b_2u(t-2) + \cdots + b_{n_b}u(t-n_b),$$

$$D(z)v(t) =$$

$$(1 + d_1z^{-1} + d_2z^{-2} + \cdots + d_{n_d}z^{-n_d})v(t) =$$

$$v(t) + d_1v(t-1) + d_2v(t-2) + \cdots + d_{n_d}v(t-n_d), \text{等}.$$

系统模型可以分为时间序列模型、方程误差类模型和输出误差类模型 3 类.

### 2.1 时间序列模型

时间序列模型(time series models)有 3 个基本模型:自回归(AR)模型、滑动平均(MA)模型、自回归滑动平均(ARMA)模型.此外,还有一种确定性 ARMA 模型(简称为 DARMA 模型)和带积分 ARMA 模型(简称为 ARIMA 模型),也可归于时间序列模

型.时间序列模型中有 2 个变量,一个是观测  $y(t)$ ,一个是随机白噪声  $v(t)$ .

1) 自回归模型(AutoRegressive model, AR),简称为 AR 模型:

$$A(z)y(t) = v(t),$$

其特征是方程的一边为白噪声,写成差分方程为

$$y(t) + a_1y(t-1) + a_2y(t-2) + \cdots + a_{n_a}y(t-n_a) = v(t).$$

2) 滑动平均模型(Moving Average model, MA),简称为 MA 模型:

$$y(t) = D(z)v(t),$$

其特征是方程的一边为观测项  $y(t)$ ,写成差分方程为

$$y(t) = v(t) + d_1v(t-1) + d_2v(t-2) + \cdots + d_{n_d}v(t-n_d).$$

3) 自回归滑动平均模型(AutoRegressive Moving Average model, ARMA),简称为 ARMA 模型:

$$A(z)y(t) = D(z)v(t),$$

写成差分方程为

$$y(t) + a_1y(t-1) + a_2y(t-2) + \cdots + a_{n_a}y(t-n_a) =$$

$$v(t) + d_1v(t-1) + d_2v(t-2) + \cdots + d_{n_d}v(t-n_d).$$

文献[9]提出了依等价 AR 模型阶次递增的自回归滑动平均模型辨识方法.

4) 确定性 ARMA 模型(Deterministic ARMA model, DARMA),简称为 DARMA 模型:

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

对应的差分方程为

$$y(t) + a_1y(t-1) + a_2y(t-2) + \cdots + a_{n_a}y(t-n_a) =$$

$$b_1u(t-1) + b_2u(t-2) + \cdots + b_{n_b}u(t-n_b).$$

因为在这个模型中,除观测  $\{y(t)\}$  是已知的,系统输入  $\{u(t)\}$  是确定性信号,也是已知的,这就是 DARMA 名称的来历.

5) 带积分 ARMA 模型(AutoRegressive Integrated Moving Average model, ARIMA),简称为 ARIMA 模型:

$$A(z)(1 - z^{-1})^d y(t) = D(z)v(t),$$

或

$$A(z)y(t) = \frac{D(z)v(t)}{(1 - z^{-1})^d}.$$

其中整数  $d \geq 0$  为积分次数.当  $d = 0$  时,ARIMA = ARMA;当  $d = 1$  时,为单重积分 ARIMA 模型.

系统除了 DARMA 模型,时间序列模型(AR 模型、MA 模型和 ARMA 模型)外,还有 2 类重要的随机系统模型:方程误差类模型和输出误差类模型.

### 2.2 方程误差类模型

方程误差类模型(equation error type models)具

有下列形式

$$A(z)y(t) = B(z)u(t) + w(t), \quad (26)$$

其中  $w(t)$  为白噪声或有色噪声,包括:白噪声过程  $w(t) := v(t)$ ; MA 过程  $w(t) := D(z)v(t)$ ; AR 过程

$$w(t) := \frac{1}{C(z)}v(t); \text{ARMA 过程 } w(t) := \frac{D(z)}{C(z)}v(t).$$

因此,方程误差类模型包括下列几种模型.

1) 方程误差模型 (Equation Error model, EE), 也称受控自回归模型 (即 CAR 模型) 或带外加输入的回归模型 (即 ARX 模型):

$$A(z)y(t) = B(z)u(t) + v(t).$$

这里  $v(t)$  为白噪声.

2) 受控自回归滑动平均模型,即受控 ARMA 模型,简称为 CARMA 模型或称为 ARMAX 模型 (即带外加输入的 ARMA 模型) (又称方程误差滑动平均模型):

$$A(z)y(t) = B(z)u(t) + D(z)v(t).$$

这里  $w(t) := D(z)v(t)$  为 MA 模型.

带积分 CARMA 模型 (简称为 CARIMA 模型或 ARIMAX 模型) 为

$$A(z)y(t) = B(z)u(t) + \frac{D(z)}{(1-z^{-1})^d}v(t).$$

当  $d=1$  时,有

$$A(z)(1-z^{-1})y(t) = B(z)(1-z^{-1})u(t) + d(z)v(t),$$

即

$$A(z)\Delta y(t) = B(z)\Delta u(t) + D(z)v(t).$$

其中

$$\Delta y(t) := (1-z^{-1})y(t) = y(t) - y(t-1),$$

$$\Delta u(t) := (1-z^{-1})u(t) = u(t) - u(t-1).$$

因此,带单重积分模型 ( $d=1$  时) 实际上是利用系统的输入输出增量进行辨识,相当于剔除了输入和输出数据中的直流项 (均值). 类推,带双重积分 ( $d=2$  时) 模型是利用系统的输入输出的二阶增量进行辨识,相当于剔除了输入和输出数据中的斜坡项 (均值).

当然,带积分环节也可与方程误差类的其他模型和输出误差类模型结合,构成相应的带积分模型,如带积分 CARMA 模型 (即 CARIMA 模型),带积分 Box-Jenkins 模型等.

3) CARAR 模型 (即受控 ARAR 模型) 或 ARARX 模型 (即带外加输入的 ARAR 模型),也称动态调节模型 (又称方程误差自回归模型):

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

这里  $w(t) := \frac{1}{C(z)}v(t)$  为 AR 模型.

4) CARARMA 模型 (即受控 ARARMA 模型) 或 ARARMAX 模型 (即带外加输入 ARARMA 模型) (又称方程误差自回归滑动平均模型)<sup>[10]</sup>:

$$A(z)y(t) = B(z)u(t) + \frac{D(z)}{C(z)}v(t). \quad (27)$$

这里  $w(t) := \frac{D(z)}{C(z)}v(t)$  为 ARMA 模型,  $w(t)$  是有色噪声.

CARARMA 模型是方程误差类的一般模型结构形式. 当  $C(z)=1$  和 (或)  $D(z)=1$  时,就得到其他方程误差模型.

置有关参数向量  $\theta$  和信息向量  $\phi(t)$ , 则 CARARMA 系统 (27) 的辨识模型为<sup>[8,11]</sup>

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

$$\theta := \begin{bmatrix} \theta_s \\ \theta_n \end{bmatrix} \in \mathbf{R}^n, n := n_a + n_b + n_c + n_d,$$

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

$$\theta_n := [c_1, c_2, \dots, c_{n_c}, d_1, d_2, \dots, d_{n_d}]^T \in \mathbf{R}^{n_c + n_d},$$

$$\varphi(t) := \begin{bmatrix} \varphi_s(t) \\ \varphi_n(t) \end{bmatrix} \in \mathbf{R}^n,$$

$$\varphi_s(t) := [-y(t-1), -y(t-2), \dots, -y(t-n_a), u(t-1), u(t-2), \dots, u(t-n_b)]^T \in \mathbf{R}^{n_a + n_b},$$

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

这里的下标  $s$  和  $n$  分别表示系统模型和噪声模型之意,取自于英文“system”和“noise”的首字母.

文献[10]提出了 CARARMA 系统递推广义增广最小二乘辨识方法,文献[11]提出了基于数据滤波的 CARARMA 系统递推最小二乘辨识方法.

### 2.3 输出误差类模型

输出误差类模型 (output error type models) 具有下列形式

$$y(t) = \frac{B(z)}{A(z)}u(t) + w(t). \quad (28)$$

模型特征是包含了有理分式项  $\frac{B(z)}{A(z)}u(t) =: x(t)$ , 因为  $x(t)$  是系统真实输出 (无噪输出), 那么输出量测  $y(t)$  中包含了噪声  $w(t)$ , 这是这类模型名称的来历. 这里  $w(t)$  为白噪声或有色噪声. 输出误差类模型包括下列几种模型.

1) 输出误差模型 (Output Error model, OE):

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

文献[12]提出了双率采样数据输出误差系统的辅助模型递推最小二乘算法,文献[13]研究了不规则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算法,文献[14]提出了输出误差模型的梯度迭代辨识方法和最小二乘迭代辨识方法,文献[15]研究了基于辅助模型的输出误差系统多新息随机梯度辨识方法.

2) 输出误差滑动平均模型(Output Error Moving Average model, OEMA)<sup>[16-17]</sup>:

$$y(t) = \frac{B(z)}{A(z)}u(t) + D(z)v(t). \quad (29)$$

文献[14]提出了 OEMA 输出误差模型的梯度迭代辨识方法和最小二乘迭代辨识方法,王冬青<sup>[16]</sup>提出了 OEMA 系统的辅助模型递推增广最小二乘法,文献[17]提出了 OEMA 系统的辅助模型增广随机梯度算法、辅助模型多新息增广随机梯度算法,文献[18]提出了 OEMA 系统的辅助模型多新息增广最小二乘辨识方法.

3) 输出误差自回归模型(Output Error AutoRegressive model, OEAR):

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

4) Box-Jenkins 模型(B-J model),又称为输出误差自回归滑动平均模型(OE-ARMA 模型):

$$y(t) = \frac{B(z)}{A(z)}u(t) + \frac{D(z)}{C(z)}v(t). \quad (30)$$

Box-Jenkins 模型(30)是输出误差类的一般模型结构形式.当  $C(z) = 1$  和(或)  $D(z) = 1$  时,就得到其他输出误差模型.

置

$$x(t) := \frac{B(z)}{A(z)}u(t), \quad w(t) := \frac{D(z)}{C(z)}v(t),$$

系统参数向量  $\theta$ , 系统模型参数向量  $\theta_s$  和噪声模型参数向量  $\theta_n$ , 信息向量  $\varphi(t)$ , 系统模型信息向量  $\varphi_s(t)$  和噪声模型信息向量  $\varphi_n(t)$ , 则 Box-Jenkins 模型(30)的辨识模型为

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

$$\theta := \begin{bmatrix} \theta_s \\ \theta_n \end{bmatrix} \in \mathbf{R}^{n_a+n_b+n_c+n_d},$$

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

$$\theta_n := [c_1, c_2, \dots, c_{n_c}, d_1, d_2, \dots, d_{n_d}]^T \in \mathbf{R}^{n_c+n_d},$$

$$\varphi(t) := \begin{bmatrix} \varphi_s(t) \\ \varphi_n(t) \end{bmatrix} \in \mathbf{R}^{n_a+n_b+n_c+n_d},$$

$$\varphi_s(t) := [-x(t-1), -x(t-2), \dots, -x(t-n_a), u(t-1), u(t-2), \dots, u(t-n_b)]^T \in \mathbf{R}^{n_a+n_b},$$

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

文献[19-20]提出了 Box-Jenkins 系统的最小二乘迭代辨识算法和梯度辨识算法,王冬青等<sup>[21]</sup>提出了 Box-Jenkins 系统的辅助模型多新息广义增广随机梯度算法.

5) 一般输出误差模型或称为自回归 Box-Jenkins 模型(AR-BJ model): 上述所有模型的一般结构形式就是下列一般输出误差模型

$$F(z)y(t) = \frac{B(z)}{A(z)}u(t) + \frac{D(z)}{C(z)}v(t). \quad (31)$$

如果系统没有迟延,多项式  $B(z)$  就定义为  $B(z) = b_0 + b_1z^{-1} + b_2z^{-2} + \dots + b_{n_b}z^{-n_b} \in \mathbf{R}$ .

如果系统有迟延,多项式  $B(z)$  的前几个系数为零.如果迟延为  $d$ ,则多项式  $B(z)$  定义为

$$B(z) = b_dz^{-d} + b_{d+1}z^{-d-1} + \dots + b_{n_b}z^{-n_b} \in \mathbf{R}.$$

问题:模型(31)中的 5 个多项式,为什么有 4 个多项式可归一化为首一多项式,解释其缘故.

## 2.4 特殊方程误差模型

当模型(31)中多项式取不同值时,就得到一些特殊的方程误差模型.

1) 有限脉冲响应模型(FIR 模型):

$$y(t) = B(z)u(t) + v(t).$$

2) 受控滑动平均模型(C-MA 模型):

$$y(t) = B(z)u(t) + D(z)v(t).$$

文献[22]提出了 C-MA 系统的交互随机梯度辨识方法.

3) 受控自回归滑动平均模型(C-ARMA 模型):

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

文献[23]提出了 C-ARMA 系统的交互最小二乘辨识方法.

## 2.5 特殊输出误差模型

当模型(31)中多项式取不同值时,就得到一些特殊的输出误差模型.

1) 自回归输出误差模型(AR-OE model):

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

2) 自回归输出误差滑动平均模型(AR-OEMA



model):

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

3) 自回归输出误差自回归模型 (AR-OEAR model):

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

尽管一般随机系统模型派生出如此多的特殊模型,但这些模型名称十分容易记忆. 记忆方法如下: 模型中的项  $F(z)y(t)$  或  $A(z)y(t)$  称为自回归项;  $B(z)u(t)$  称为受控项;  $\frac{B(z)}{A(z)}u(t)$  称为输出误差项;  $D(z)v(t)$  称为滑动平均项 (因为  $w(t) := D(z)v(t)$  是滑动平均模型);  $\frac{1}{C(z)}v(t)$  称为自回归项 (因为  $w(t) := \frac{1}{C(z)}v(t)$  是自回归模型). 记住了这些,上述模型的名称就一目了然.

## 2.6 一般随机系统模型

一般随机系统模型结构如下:

$$y(t) = G(z)u(t) + H(z)v(t). \quad (32)$$

有时把  $G(z) \in \mathbf{R}$  称为系统模型 (system model) 的 (有理分式) 传递函数,  $H(z) \in \mathbf{R}$  称为噪声模型 (noise model) 的 (有理分式) 传递函数,其结构如图 1 所示. 图 1 中不可测内部变量  $x(t) := G(z)u(t)$  为系统的无噪输出 (noise-free output) 或真实输出 (true output),  $w(t) := H(z)v(t)$  为噪声模型输出 (不可测). 只有可测的输入输出数据  $\{u(t), y(t)\}$  可用来辨识模型参数.

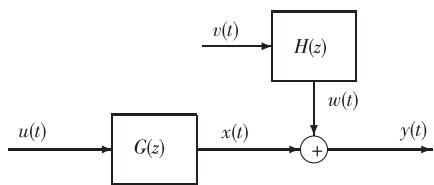


图 1 一般随机系统的结构

Fig. 1 The general stochastic systems

上述所有随机系统模型都可以化为式(32)的结构形式. 例如,

1) 对于 Box-Jenkins 模型(30), 有

$$G(z) := \frac{B(z)}{A(z)}, \quad H(z) := \frac{D(z)}{C(z)}.$$

2) 对于 CARARMA 模型(27), 两边除以  $A(z)$ , 就有

$$G(z) := \frac{B(z)}{A(z)}, \quad H(z) := \frac{D(z)}{A(z)C(z)}.$$

实际中,有时也把 CARARMA 模型(27)中的  $\frac{D(z)}{C(z)}$  称为噪声模型.

3) 对于一般输出误差系统(31), 两边除以  $F(z)$ , 就有

$$G(z) := \frac{B(z)}{F(z)A(z)}, \quad H(z) := \frac{D(z)}{F(z)C(z)}.$$

注 1: 尽管这些模型可以在两边同乘以或同除以一多项式,变为另一种形式的模型,但是从辨识的角度讲,这样做一般会增加辨识(模型)参数的数目. 例如, OEMA 模型(29)包含 3 个多项式  $A(z)$ ,  $B(z)$  和  $D(z)$ , 它们的系数是待辨识的参数, 3 个多项式包含  $n_a + n_b + n_d$  个参数. 如果把模型(29)两边同乘以  $A(z)$  得到

$$A(z)y(t) = B(z)u(t) + D'(z)v(t), \\ D'(z) := A(z)D(z).$$

这是个 CARMA 模型, 还是包含 3 个多项式  $A(z)$ ,  $B(z)$  和  $D'(z)$ , 它们的系数是待辨识的参数, 需要辨识的参数数目从  $n_a + n_b + n_d$  增加到  $2n_a + n_b + n_d$ .

## 3 多变量系统

多变量系统辨识方法可以看作是标量系统辨识方法的直接推广和延伸, 但是多变量系统辨识有比标量系统辨识更丰富的内容. 多变量系统辨识算法的复杂性、计算量及参数估计性质 (无偏性、有效性和收敛性等) 完全取决于所采用的模型及其待辨识参数数目的多少. 因为一种辨识算法在一种模型结构下可获参数的一致无偏估计, 而在另一种模型结构下的参数估计可能是有偏的. 因此, 针对具体对象选择合适的模型, 对于提高辨识精度和较小辨识算法计算量有着重要意义.

单输入单输出系统 (即标量系统) 推广就是多输入多输出系统 (即多变量系统), 也就是将输入扩展为输入向量  $u(t) \in \mathbf{R}^r$ , 将输出扩展为输出向量  $y(t) \in \mathbf{R}^m$ , 移位算子多项式也扩展为多项式矩阵:

$$A(z) := I + A_1 z^{-1} + A_2 z^{-2} + \dots + A_{n_a} z^{-n_a} \in \mathbf{R}^{m \times m}; \\ B(z) := B_1 z^{-1} + B_2 z^{-2} + \dots + B_{n_b} z^{-n_b} \in \mathbf{R}^{m \times r}; \\ C(z) := I + C_1 z^{-1} + C_2 z^{-2} + \dots + C_{n_c} z^{-n_c} \in \mathbf{R}^{m \times m}; \\ D(z) := I + D_1 z^{-1} + D_2 z^{-2} + \dots + D_{n_d} z^{-n_d} \in \mathbf{R}^{m \times m}; \\ F(z) := I + F_1 z^{-1} + F_2 z^{-2} + \dots + F_{n_f} z^{-n_f} \in \mathbf{R}^{m \times m}.$$

假设阶次  $n_a + n_b + n_c + n_d$  和  $n_f$  已知, 移位算子

$z^{-1}$ 多项式矩阵  $A(z), B(z), C(z), D(z)$  和  $F(z)$  的系数矩阵  $A_i \in \mathbf{R}^{m \times m}, B_i \in \mathbf{R}^{m \times r}, C_i \in \mathbf{R}^{m \times m}, D_i \in \mathbf{R}^{m \times m}$  和  $F_i \in \mathbf{R}^{m \times m}$  是待辨识的系统参数矩阵,  $v(t) \in \mathbf{R}^m$  为零均值白噪声向量,  $z^{-1}y(t) = y(t-1)$ , 目标就是利用系统的输入输出数据  $\{u(t), y(t)\}$  估计系统参数矩阵  $(A_i, B_i, C_i, D_i, F_i)$ .

多变量系统模型也可以分为多变量时间序列模型、多变量方程误差类模型和多变量输出误差类模型 3 类.

### 3.1 多变量时间序列模型

多变量时间序列模型 (multivariable time series models) 有 3 个基本模型: 多变量 AR 模型, 多变量 MA 模型, 多变量 ARMA 模型. 现简介如下.

1) 多变量 AR 模型 (multivariable AR model):

$$A(z)y(t) = v(t),$$

或

$$y(t) + A_1 y(t-1) + A_2 y(t-2) + \cdots + A_{n_a} y(t-n_a) = v(t).$$

2) 多变量 MA 模型 (multivariable MA model):

$$y(t) = D(z)v(t),$$

或

$$y(t) = v(t) + D_1 v(t-1) + D_2 v(t-2) + \cdots + D_{n_d} v(t-n_d).$$

3) 多变量 ARMA 模型 (multivariable ARMA model):

$$A(z)y(t) = D(z)v(t),$$

或

$$y(t) + A_1 y(t-1) + A_2 y(t-2) + \cdots + A_{n_a} y(t-n_a) = v(t) + D_1 v(t-1) + D_2 v(t-2) + \cdots + D_{n_d} v(t-n_d).$$

4) 多变量 DARMA 模型 (multivariable DARMA model):

$$A(z)y(t) = B(z)u(t). \quad (33)$$

5) 带积分多变量 ARMA 模型 (multivariable ARIMA model, 多变量 ARIMA 模型):

$$A(z)(1-z^{-1})^d y(t) = D(z)v(t).$$

### 3.2 多变量方程误差类模型

多变量方程误差类模型具有下列形式

$$A(z)y(t) = B(z)u(t) + w(t). \quad (34)$$

其中  $w(t) \in \mathbf{R}^m$  是零均值白噪声或有色噪声向量. 包括: 白噪声过程  $w(t) = v(t)$ ; 多变量 MA 过程  $w(t) = D(z)v(t)$ ; 多变量 AR 过程  $w(t) = C^{-1}(z)v(t)$ ; 多变量 ARMA 过程  $w(t) = C^{-1}D(z)v(t)$ .

1) 多变量方程误差模型 (多变量 CAR 模型或多变量 ARX 模型):

$$A(z)y(t) = B(z)u(t) + v(t).$$

丁锋等<sup>[24-25]</sup>研究了多变量方程误差模型随机梯度辨识方法的收敛性.

2) 多变量 CARMA 模型 (多变量 ARMAX 模型):

$$A(z)y(t) = B(z)u(t) + D(z)v(t).$$

带积分多变量 CARMA 模型 (多变量 CARIMA 模型或多变量 ARIMAX 模型) 为

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

或

$$A(z)\Delta y(t) = B(z)\Delta u(t) + D(z)v(t), \Delta = 1 - z^{-1}.$$

对应的辨识模型为

$$y(t) = \theta^T \varphi(t) + v(t), \quad n = mn_a + rn_b + mn_d,$$

$$\theta^T = [A_1, A_2, \cdots, A_{n_a}, B_1, B_2, \cdots, B_{n_b}, D_1, D_2, \cdots, D_{n_d}] \in \mathbf{R}^{m \times n},$$

$$\varphi(t) = [-\Delta y^T(t-1), -\Delta y^T(t-2), \cdots, -\Delta y^T(t-n_a), \Delta u^T(t-1), \Delta u^T(t-2), \cdots, \Delta u^T(t-n_b), v^T(t-1), v^T(t-2), \cdots, v^T(t-n_d)]^T \in \mathbf{R}^n.$$

3) 多变量 CARAR 模型 (多变量 ARARX 模型, 也称多变量动态调节模型):

$$A(z)y(t) = B(z)u(t) + C^{-1}(z)v(t).$$

4) 多变量 CARARMA 模型 (多变量 ARARMAX 模型):

$$A(z)y(t) = B(z)u(t) + C^{-1}(z)D(z)v(t). \quad (35)$$

令

$$w(t) = C^{-1}(z)D(z)v(t),$$

则多变量 CARARMA 系统(35)的辨识模型为

$$y(t) = \theta^T \varphi(t) + v(t), \quad n = mn_a + rn_b + mn_c + mn_d,$$

$$\theta^T = [A_1, A_2, \cdots, A_{n_a}, B_1, B_2, \cdots, B_{n_b}, C_1, C_2, \cdots, C_{n_c}, D_1, D_2, \cdots, D_{n_d}] \in \mathbf{R}^{m \times n},$$

$$\varphi(t) = [-y^T(t-1), -y^T(t-2), \cdots, -y^T(t-n_a), u^T(t-1), u^T(t-2), \cdots, u^T(t-n_b), -w^T(t-1), -w^T(t-2), \cdots, -w^T(t-n_c), v^T(t-1), v^T(t-2), \cdots, v^T(t-n_d)]^T \in \mathbf{R}^n.$$

当  $C(z) = I$  和 (或)  $D(z) = I$  时, 由多变量 CARARMA 系统(35)可得到其他方程误差模型.

文献[26]提出了多变量 CARARMA 模型的递推广义增广最小二乘算法.

### 3.3 多变量输出误差类模型

多变量输出误差类模型具有下列形式:

$$y(t) = A^{-1}(z)B(z)u(t) + w(t).$$

1) 多变量输出误差模型 (multivariable OE model):

$$y(t) = A^{-1}(z)B(z)u(t) + v(t). \quad (36)$$

2) 多变量输出误差滑动平均模型 (multivariable OEMA model):

$$\mathbf{y}(t) = \mathbf{A}^{-1}(z)\mathbf{B}(z)\mathbf{u}(t) + \mathbf{D}(z)\mathbf{v}(t).$$

3) 多变量输出误差自回归模型 (multivariable OEAR model):

$$\mathbf{y}(t) = \mathbf{A}^{-1}(z)\mathbf{B}(z)\mathbf{u}(t) + \mathbf{C}^{-1}(z)\mathbf{v}(t).$$

4) 多变量 Box-Jenkins 模型 (multivariable Box-Jenkins model):

$$\mathbf{y}(t) = \mathbf{A}^{-1}(z)\mathbf{B}(z)\mathbf{u}(t) + \mathbf{C}^{-1}(z)\mathbf{D}(z)\mathbf{v}(t). \quad (37)$$

$$\mathbf{x}(t) := \mathbf{A}^{-1}(z)\mathbf{B}(z)\mathbf{u}(t),$$

$$\mathbf{w}(t) := \mathbf{C}^{-1}(z)\mathbf{D}(z)\mathbf{v}(t).$$

则多变量 Box-Jenkins 模型 (37) 的辨识模型为

$$\mathbf{y}(t) = \boldsymbol{\theta}^T \boldsymbol{\varphi}(t) + \mathbf{v}(t),$$

$$n := mn_a + rn_b + mn_c + mn_d,$$

$$\boldsymbol{\theta}^T := [\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_{n_a}, \mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_{n_b}, \mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_{n_c}, \mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_{n_d}] \in \mathbf{R}^{m \times n},$$

$$\boldsymbol{\varphi}(t) := [-\mathbf{x}^T(t-1), -\mathbf{x}^T(t-2), \dots, -\mathbf{x}^T(t-n_a),$$

$$\mathbf{u}^T(t-1), \mathbf{u}^T(t-2), \dots, \mathbf{u}^T(t-n_b),$$

$$-\mathbf{w}^T(t-1), -\mathbf{w}^T(t-2), \dots, -\mathbf{w}^T(t-n_c),$$

$$\mathbf{v}^T(t-1), \mathbf{v}^T(t-2), \dots, \mathbf{v}^T(t-n_d)]^T \in \mathbf{R}^n.$$

5) 一般多变量输出误差模型 (多变量自回归 Box-Jenkins 模型, multivariable AR-BJ model):

$$\mathbf{F}(z)\mathbf{y}(t) = \mathbf{A}^{-1}(z)\mathbf{B}(z)\mathbf{u}(t) + \mathbf{C}^{-1}(z)\mathbf{D}(z)\mathbf{v}(t). \quad (38)$$

### 3.4 特殊多变量方程误差模型

当模型 (38) 中多项式矩阵取不同值时, 就得到一些特殊的多变量方程误差模型.

1) 多变量 FIR 模型 (multivariable FIR model):

$$\mathbf{y}(t) = \mathbf{B}(z)\mathbf{u}(t) + \mathbf{v}(t). \quad (39)$$

2) 多变量受控滑动平均模型 (多变量 C-MA 模型, multivariable C-MA model):

$$\mathbf{y}(t) = \mathbf{B}(z)\mathbf{u}(t) + \mathbf{D}(z)\mathbf{v}(t).$$

3) 多变量受控自回归滑动平均模型 (多变量 C-ARMA 模型, multivariable C-ARMA model):

$$\mathbf{y}(t) = \mathbf{B}(z)\mathbf{u}(t) + \mathbf{C}^{-1}(z)\mathbf{D}(z)\mathbf{v}(t).$$

### 3.5 特殊多变量输出误差模型

当模型 (38) 中多项式矩阵取不同值时, 就得到一些特殊的多变量输出误差模型.

1) 多变量自回归输出误差模型 (multivariable AR-OE model):

$$\mathbf{F}(z)\mathbf{y}(t) = \mathbf{A}^{-1}(z)\mathbf{B}(z)\mathbf{u}(t) + \mathbf{v}(t).$$

2) 多变量自回归输出误差滑动平均模型 (multivariable AR-OEMA model):

$$\mathbf{F}(z)\mathbf{y}(t) = \mathbf{A}^{-1}(z)\mathbf{B}(z)\mathbf{u}(t) + \mathbf{D}(z)\mathbf{v}(t).$$

3) 多变量自回归输出误差自回归模型 (multivariable AR-OEAR model):

$$\mathbf{F}(z)\mathbf{y}(t) = \mathbf{A}^{-1}(z)\mathbf{B}(z)\mathbf{u}(t) + \mathbf{C}^{-1}(z)\mathbf{v}(t).$$

当多项式矩阵  $\mathbf{C}(z)$ ,  $\mathbf{D}(z)$  或  $\mathbf{F}(z)$  退化为标量多项式  $C(z)$ ,  $D(z)$  或  $F(z)$ , 那么上述多变量系统变为一类新的多变量系统, 其特点是噪声模型具有相同的传递函数.

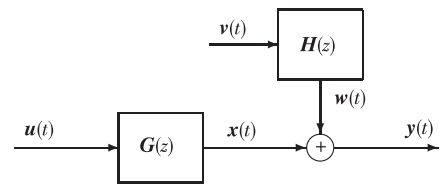


图2 一般多变量随机系统的结构

Fig. 2 The general multivariable stochastic systems

### 3.6 一般多变量随机系统模型

一般多变量随机系统模型结构如下:

$$\mathbf{y}(t) = \mathbf{G}(z)\mathbf{u}(t) + \mathbf{H}(z)\mathbf{v}(t). \quad (40)$$

其中  $\mathbf{G}(z) \in \mathbf{R}^{m \times r}$  为系统模型的 (有理分式) 传递矩阵,  $\mathbf{H}(z) \in \mathbf{R}^{m \times m}$  为噪声模型的 (有理分式) 传递矩阵, 其结构如图 2 所示. 图 2 中  $\mathbf{x}(t) := \mathbf{G}(z)\mathbf{u}(t)$  为系统的无噪声输出向量或真实输出向量 (不可测),  $\mathbf{w}(t) := \mathbf{H}(z)\mathbf{v}(t)$  为噪声模型输出 (不可测).

上述所有多变量随机系统模型都可以化为式 (40) 的结构形式. 例如,

1) 对于多变量 Box-Jenkins 模型 (37), 有

$$\mathbf{G}(z) := \mathbf{A}^{-1}(z)\mathbf{B}(z),$$

$$\mathbf{H}(z) := \mathbf{C}^{-1}(z)\mathbf{D}(z).$$

2) 对于多变量 CARARMA 模型 (35), 两边乘以  $\mathbf{A}^{-1}(z)$ , 就有

$$\mathbf{G}(z) := \mathbf{A}^{-1}(z)\mathbf{B}(z),$$

$$\mathbf{H}(z) := \mathbf{A}^{-1}(z)\mathbf{C}^{-1}(z)\mathbf{D}(z).$$

3) 对于一般输出误差系统 (38), 两边乘以  $\mathbf{F}^{-1}(z)$ , 就有

$$\mathbf{G}(z) := \mathbf{F}^{-1}(z)\mathbf{A}^{-1}(z)\mathbf{B}(z),$$

$$\mathbf{H}(z) := \mathbf{F}^{-1}(z)\mathbf{C}^{-1}(z)\mathbf{D}(z).$$

**注 2:** 上述多变量系统模型可以在两边同乘以某一多项式矩阵, 变为另一种形式的模型, 但是变化后模型参数数目将会增多.

## 4 类多变量系统

### 4.1 状态空间描述到输入输出表达

考虑  $r$  个输入  $m$  个输出的线性多变量系统, 其状态空间模型为

$$\begin{cases} \mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t). \end{cases} \quad (41)$$

其中  $\mathbf{x}(t) \in \mathbf{R}^n$  为系统状态向量,  $\mathbf{u}(t) = [u_1(t), u_2(t), \dots, u_r(t)]^T \in \mathbf{R}^r$  为系统输入向量,  $\mathbf{y}(t) = [y_1(t), y_2(t), \dots, y_m(t)]^T \in \mathbf{R}^m$  为系统输出向量,  $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$  为适当维数的系统矩阵.

使用移位算子的性质, 可以推导出状态空间模型(41)与其传递函数矩阵  $\mathbf{G}(z) \in \mathbf{R}^{m \times r}$  间的关系. 方程(41)可以写作

$$\mathbf{y}(t) = [\mathbf{C}(z\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}]\mathbf{u}(t),$$

或

$$\mathbf{y}(t) = \mathbf{G}(z)\mathbf{u}(t). \quad (42)$$

其中  $\mathbf{G}(z) := \mathbf{C}(z\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D} \in \mathbf{R}^{m \times r}$  为系统的传递矩阵. 进一步展开得到

$$\begin{aligned} \mathbf{G}(z) &= \mathbf{C}(z\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D} = \frac{\mathbf{C}\text{adj}[z\mathbf{I} - \mathbf{A}]\mathbf{B}}{\det[z\mathbf{I} - \mathbf{A}]} + \mathbf{D} = \\ &= \frac{z^{-n}\mathbf{C}\text{adj}[z\mathbf{I} - \mathbf{A}]\mathbf{B}}{z^{-n}\det[z\mathbf{I} - \mathbf{A}]} + \mathbf{D} =: \frac{\mathbf{Q}(z)}{\alpha(z)}. \end{aligned}$$

其中  $\alpha(z)$  是系统的  $n$  次特征多项式, 它定义为传递矩阵的各元最小公分母首一多项式, 即  $\mathbf{G}(z)$  的最小公分母,  $\mathbf{Q}(z)$  是  $z^{-1}$  的矩阵多项式, 它们可以表达为

$$\begin{aligned} \alpha(z) &:= z^{-n}\det[z\mathbf{I} - \mathbf{A}] = \\ &= 1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \dots + \alpha_n z^{-n}, \quad \alpha_i \in \mathbf{R}, \quad (43) \\ \mathbf{Q}(z) &:= z^{-n}\mathbf{C}\text{adj}[z\mathbf{I} - \mathbf{A}]\mathbf{B} + \mathbf{D}\alpha(z) = \\ &= \mathbf{Q}_0 + \mathbf{Q}_1 z^{-1} + \mathbf{Q}_2 z^{-2} + \dots + \mathbf{Q}_n z^{-n}, \quad \mathbf{Q}_i \in \mathbf{R}^{m \times r}. \end{aligned} \quad (44)$$

特征多项式  $\alpha(z)$  的辨识, 对于控制系统极点配置、观测器设计以及稳定性分析十分重要.

方程(42)可以写作

$$\mathbf{y}(t) = \frac{\mathbf{Q}(z)}{\alpha(z)}\mathbf{u}(t).$$

我们把这个模型称为多变量系统传递函数阵主模型(main model), 其特征是传递函数阵  $\mathbf{G}(z) = \frac{\mathbf{Q}(z)}{\alpha(z)}$  有一个系统特征多项式作为公分母, 分子为一个矩阵多项式. 主模型是相对于分子系统的子模型(submodel)而言. 上式又可写为

$$\alpha(z)\mathbf{y}(t) = \mathbf{Q}(z)\mathbf{u}(t). \quad (45)$$

这里  $\alpha(z)$  是一个标量多项式, 这是与多变量 DAR-

MA 模型的差别, 所以称式(45)为多变量 DARMA-like 模型.

实际系统经常存在这样或那样的干扰作用, 且这些干扰往往具有随机性, 可能是白噪声, 也可能是相关噪声. 在确定性系统(45)中引入不同的干扰项, 就得到一些不同多变量系统模型. 其中多变量 AR-like 模型如下:

$$\alpha(z)\mathbf{y}(t) = \mathbf{v}(t),$$

或

$$\begin{cases} \alpha(z)y_1(t) = v_1(t), \\ \alpha(z)y_2(t) = v_2(t), \\ \vdots \\ \alpha(z)y_m(t) = v_m(t). \end{cases}$$

这  $m$  个 AR 模型具有相同的结构形式, 具有相同的标量多项式  $\alpha(z)$ , 右边  $v_i(t)$  都是零均值白噪声过程,  $y_i(t)$  为观测. 从理论上讲, 似乎研究这个多变量 AR 模型没有什么意义. 然而, 任何理论, 只要赋予它实际的含义, 就有价值. 对于预测, 人们总是假设预测对象符合一定的规律, 比如说符合 AR 规律. 最典型的是根据过去的观测预测天气预报. 当然, 可以根据某个地点的天气预报观测量  $y_i(t)$ , 建立 AR 预测模型. 要使建立的预测模型也适用于另一地区, 最好联合  $m$  个地点的观测量  $\{y_i(t), i=1, 2, \dots, m\}$ , 建立一个适用于  $m$  个地区的 AR 预测模型, 这就是多变量 AR-like 模型研究的意义所在.

### 4.2 类多变量方程误差类模型

类多变量方程误差类模型具有下列形式:

$$\alpha(z)\mathbf{y}(t) = \mathbf{Q}(z)\mathbf{u}(t) + \mathbf{w}(t). \quad (46)$$

其中  $\mathbf{w}(t) \in \mathbf{R}^m$  是零均值白噪声或有色噪声向量.

1) 多变量 CAR-like 模型(多变量 ARX-like 模型)<sup>[27-29]</sup>:

$$\alpha(z)\mathbf{y}(t) = \mathbf{Q}(z)\mathbf{u}(t) + \mathbf{v}(t). \quad (47)$$

下面推导多变量 ARX-like 系统的辨识模型. 将式(43)的  $\alpha(z)$  和式(44)的  $\mathbf{Q}(z)$  代入式(47), 利用移位算子  $z^{-1}$  的性质得到

$$(1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \dots + \alpha_n z^{-n})\mathbf{y}(t) = (\mathbf{Q}_0 + \mathbf{Q}_1 z^{-1} + \mathbf{Q}_2 z^{-2} + \dots + \mathbf{Q}_n z^{-n})\mathbf{u}(t) + \mathbf{v}(t),$$

或

$$\begin{aligned} \mathbf{y}(t) + \alpha_1 \mathbf{y}(t-1) + \alpha_2 \mathbf{y}(t-2) + \dots + \alpha_n \mathbf{y}(t-n) = \\ \mathbf{Q}_0 \mathbf{u}(t) + \mathbf{Q}_1 \mathbf{u}(t-1) + \mathbf{Q}_2 \mathbf{u}(t-2) + \dots + \\ \mathbf{Q}_n \mathbf{u}(t-n) + \mathbf{v}(t). \end{aligned}$$

进一步可以写为

$$y(t) + [y(t-1), y(t-2), \dots, y(t-n)] \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = [\mathbf{Q}_0, \mathbf{Q}_1, \dots, \mathbf{Q}_n] \begin{bmatrix} u(t) \\ u(t-1) \\ \vdots \\ u(t-n) \end{bmatrix} + v(t). \quad (48)$$

定义参数矩阵  $\theta$ , 参数向量  $\alpha$ , 输入信息向量  $\varphi(t)$  和输出信息矩阵  $\psi(t)$  如下:

$$\theta^T := [\mathbf{Q}_0, \mathbf{Q}_1, \dots, \mathbf{Q}_n] \in \mathbf{R}^{m \times (nr+r)},$$

$$\alpha := \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} \in \mathbf{R}^n, \quad \varphi(t) := \begin{bmatrix} u(t) \\ u(t-1) \\ \vdots \\ u(t-n) \end{bmatrix} \in \mathbf{R}^{nr+r},$$

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

于是,从式(48)可以得到多变量 CAR-like 系统的递阶辨识模型(hierarchical identification model):

$$y(t) + \psi(t)\alpha = \theta^T \varphi(t) + v(t). \quad (49)$$

这个模型待辨识的参数包括 2 部分:由系统特征多项式系数构成的参数向量  $\alpha \in \mathbf{R}^n$  和由传递矩阵分子多项式矩阵系数构成的参数矩阵  $\theta \in \mathbf{R}^{(nr+r) \times m}$ . 由于这个辨识模型(49)既包含一个参数向量,又包含一个参数矩阵,所以标准最小二乘算法不能直接使用,但是,可以把参数矩阵  $\theta$  化为一个堆积向量  $\text{col}[\theta]$ ,使用 Kronecker 积把模型(49)表示为

$$y(t) = [-\psi(t), \varphi^T(t) \otimes \mathbf{I}_m] \begin{bmatrix} \alpha \\ \text{col}[\theta^T] \end{bmatrix} + v(t) = [-\psi(t), \mathbf{I}_m \otimes \varphi^T(t)] \begin{bmatrix} \alpha \\ \text{col}[\theta] \end{bmatrix} + v(t), \quad (50)$$

其中  $\text{col}[X]$  表示将矩阵  $X$  的列按次序排成的向量,如

$$X := [x_1, x_2, \dots, x_n] \in \mathbf{R}^{m \times n}, x_i \in \mathbf{R}^m, i = 1, 2, \dots, n, \text{ 那么}$$

$$\text{col}[X] = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbf{R}^{mn}.$$

虽然这个辨识模型(50)可以使用递推最小二乘算法进行辨识,但计算量很大,需要计算的协方差矩阵的维数从  $(nr+r)$  增加到  $(mnr+mr+n)$ . 本文作者提出的递阶梯度辨识方法和递阶最小二乘辨识方

法具有小的计算量<sup>[8,27-28]</sup>.

2) 多变量 CARMA-like 模型(多变量 ARMAX-like 模型):

$$\alpha(z)y(t) = \mathbf{Q}(z)u(t) + \mathbf{D}(z)v(t). \quad (51)$$

带积分多变量 CARMA 模型(多变量 CARIMA 模型或多变量 ARIMA 模型)为

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

或

$$\alpha(z)\Delta y(t) = \mathbf{Q}(z)\Delta u(t) + \mathbf{D}(z)v(t).$$

当多项式矩阵  $\mathbf{D}(z)$  退化为标量多项式  $D(z)$  时,得到

$$\alpha(z)y(t) = \mathbf{Q}(z)u(t) + D(z)v(t). \quad (52)$$

文献[30]提出了多变量 CARMA-like 系统(52)的递阶最小二乘迭代辨识方法.

3) 多变量 CARAR-like 模型(多变量 ARARX-like 模型):

$$\alpha(z)y(t) = \mathbf{Q}(z)u(t) + \mathbf{C}^{-1}(z)v(t).$$

4) 多变量 CARARMA-like 模型(多变量 ARARMAX-like 模型):

$$\alpha(z)y(t) = \mathbf{Q}(z)u(t) + \mathbf{C}^{-1}(z)\mathbf{D}(z)v(t). \quad (53)$$

当  $\mathbf{C}(z) = \mathbf{I}$  和(或)  $\mathbf{D}(z) = \mathbf{I}$  时,就得到其他方程误差模型.

### 4.3 类多变量输出误差类模型

类多变量输出误差类模型具有下列形式:

$$y(t) = \frac{\mathbf{Q}(z)}{\alpha(z)}u(t) + w(t). \quad (54)$$

1) 类多变量输出误差模型(多变量 OE-like 模型, multivariable OE-like model):

$$y(t) = \frac{\mathbf{Q}(z)}{\alpha(z)}u(t) + v(t).$$

2) 类多变量输出误差滑动平均模型(多变量 OEMA-like 模型, multivariable OEMA-like model):

$$y(t) = \frac{\mathbf{Q}(z)}{\alpha(z)}u(t) + \mathbf{D}(z)v(t). \quad (55)$$

文献[31]提出了类多变量输出误差滑动平均系统的递阶梯度迭代参数估计算法.

3) 类多变量输出误差自回归模型(多变量 OE-AR-like 模型, multivariable OEAR-like model):

$$y(t) = \frac{\mathbf{Q}(z)}{\alpha(z)}u(t) + \mathbf{C}^{-1}(z)v(t).$$

4) 多变量 Box-Jenkins-like 模型( multivariable BJ-like model):

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

5) 一般类多变量输出误差模型(多变量自回归 Box-Jenkins-like 模型, multivariable AR-BJ-like model):

$$\zeta(z)y(t) = \frac{Q(z)}{\alpha(z)}u(t) + C^{-1}(z)D(z)v(t), \quad (56)$$

其中  $\zeta(z) := 1 + \zeta_1 z^{-1} + \zeta_2 z^{-2} + \dots + \zeta_n z^{-n} \in \mathbf{R}$ .

6) 辨识模型举例: 当多项式矩阵  $D(z)$  退化为标量多项式  $D(z)$  时, 由式(55)得到一特殊多变量 OEMA-like 系统

$$y(t) = \frac{Q(z)}{\alpha(z)}u(t) + D(z)v(t). \quad (57)$$

下面推导这一特殊多变量 OEMA-like 系统(57)的辨识模型. 令

$$x(t) := \frac{Q(z)}{\alpha(z)}u(t) \in \mathbf{R}^m. \quad (58)$$

或

$$\alpha(z)x(t) = Q(z)u(t).$$

将式(43)的  $\alpha(z)$  和式(44)的  $Q(z)$  代入上式, 利用移位算子  $z^{-1}$  的性质得到

$$(1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \dots + \alpha_n z^{-n})x(t) = (Q_0 + Q_1 z^{-1} + Q_2 z^{-2} + \dots + Q_n z^{-n})u(t),$$

或

$$x(t) + \alpha_1 x(t-1) + \alpha_2 x(t-2) + \dots + \alpha_n x(t-n) = Q_0 u(t) + Q_1 u(t-1) + Q_2 u(t-2) + \dots + Q_n u(t-n).$$

$$x(t) = -\alpha_1 x(t-1) - \alpha_2 x(t-2) - \dots - \alpha_n x(t-n) + Q_0 u(t) + Q_1 u(t-1) + Q_2 u(t-2) + \dots + Q_n u(t-n). \quad (59)$$

将式(59)以及多项式  $D(z)$  的表达式代入式(57)得到

$$y(t) = x(t) + (1 + d_1 z^{-1} + d_2 z^{-2} + \dots + d_{n_d} z^{-n_d})v(t).$$

将式(59)代入上式可得

$$y(t) = -\alpha_1 x(t-1) - \alpha_2 x(t-2) - \dots - \alpha_n x(t-n) + Q_0 u(t) + Q_1 u(t-1) + Q_2 u(t-2) + \dots + Q_n u(t-n) + v(t) + d_1 v(t-1) + d_2 v(t-2) + \dots + d_{n_d} v(t-n_d). \quad (60)$$

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

$$\mathbf{v} := \begin{bmatrix} \boldsymbol{\alpha} \\ \mathbf{d} \end{bmatrix} \in \mathbf{R}^{n+n_d};$$

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

$$\mathbf{d} := [d_1, d_2, \dots, d_{n_d}]^T \in \mathbf{R}^{n_d};$$

$$\boldsymbol{\theta}^T := [Q_0, Q_1, Q_2, \dots, Q_n] \in \mathbf{R}^{m \times (nr+r)};$$

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

$$\boldsymbol{\psi}(t) := [\boldsymbol{\psi}_s(t), \boldsymbol{\psi}_n(t)] \in \mathbf{R}^{m \times (n+n_d)};$$

$$\boldsymbol{\psi}_s(t) := [x(t-1), x(t-2), \dots, x(t-n)] \in \mathbf{R}^{m \times n};$$

$$\boldsymbol{\psi}_n(t) := [-v(t-1), -v(t-2), \dots, -v(t-n_d)] \in \mathbf{R}^{m \times n_d}.$$

于是, 由式(59)可得

$$x(t) = -\boldsymbol{\psi}_s(t)\boldsymbol{\alpha} + \boldsymbol{\theta}^T \boldsymbol{\varphi}(t). \quad (61)$$

从式(60)可以得到辨识模型

$$y(t) + \boldsymbol{\psi}(t)\mathbf{v} = \boldsymbol{\theta}^T \boldsymbol{\varphi}(t) + v(t). \quad (62)$$

这就是多变量 OEMA-like 系统(57)的辨识模型. 文献[31]提出了多变量 OEMA-like 系统(57)的递阶梯度迭代参数估计算法.

#### 4.4 类特殊多变量方程误差模型

当模型(56)中多项式矩阵取不同值时, 就得到一些特殊的多变量方程误差模型.

1) 多变量 C-MA-like 模型(multivariable C-MA-like model):

$$y(t) = Q(z)u(t) + D(z)v(t).$$

2) 多变量 C-ARMA-like 模型(multivariable C-ARMA-like model):

$$y(t) = Q(z)u(t) + \frac{D(z)}{C(z)}v(t).$$

#### 4.5 类特殊多变量输出误差模型

当模型(56)中多项式矩阵取不同值时, 就得到一些特殊的多变量输出误差模型.

1) 多变量 AR-OE-like 模型(multivariable AR-OE-like model):

$$\zeta(z)y(t) = \frac{Q(z)}{\alpha(z)}u(t) + v(t).$$

2) 多变量 AR-OEMA-like 模型(multivariable AR-OEMA-like model):

$$\zeta(z)y(t) = \frac{Q(z)}{\alpha(z)}u(t) + D(z)v(t).$$

3) 多变量 AR-OEAR-like 模型(multivariable AR-OEAR-like model):

$$\zeta(z)y(t) = \frac{Q(z)}{\alpha(z)}u(t) + C^{-1}(z)v(t).$$

当多项式矩阵  $C(z)$  或  $D(z)$  退化为下列标量多项式  $C(z)$  或  $D(z)$ , 那么上述类多变量系统派生出一些新的多变量系统模型.

#### 4.6 一般类多变量随机系统模型

一般类多变量随机系统模型结构如下(图3):

$$y(t) = \frac{Q(z)}{\alpha(z)}u(t) + \frac{N(z)}{\gamma(z)}v(t), \quad (63)$$

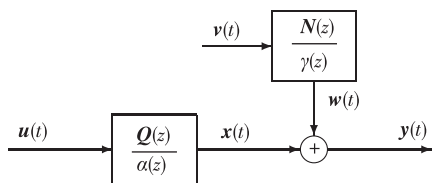


图3 一般多变量随机系统的结构

Fig. 3 The general multivariable stochastic systems

其中  $G(z) := \frac{Q(z)}{\alpha(z)} \in \mathbf{R}^{m \times r}$  为系统模型的传递矩阵,

$H(z) := \frac{N(z)}{\gamma(z)} \in \mathbf{R}^{m \times m}$  为噪声模型的传递矩阵,  $Q(z) \in \mathbf{R}^{m \times r}$  和  $N(z) \in \mathbf{R}^{m \times m}$  均为多项式矩阵或矩阵多项式,  $\alpha(z) \in \mathbf{R}$  和  $\gamma(z) \in \mathbf{R}$  均为标量多项式, 定义如下:

$$\begin{aligned} \alpha(z) &:= 1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \dots + \alpha_n z^{-n} \in \mathbf{R}; \\ Q(z) &:= Q_0 + Q_1 z^{-1} + Q_2 z^{-2} + \dots + Q_n z^{-n} \in \mathbf{R}^{m \times r}; \\ \gamma(z) &:= 1 + \gamma_1 z^{-1} + \gamma_2 z^{-2} + \dots + \gamma_n z^{-n} \in \mathbf{R}; \\ N(z) &:= I + N_1 z^{-1} + N_2 z^{-2} + \dots + N_n z^{-n} \in \mathbf{R}^{m \times m}. \end{aligned}$$

这里  $\alpha_i \in \mathbf{R}, \gamma_i \in \mathbf{R}, Q_i \in \mathbf{R}^{m \times r}$  和  $N_i \in \mathbf{R}^{m \times m}$  是待辨识的系统参数和参数矩阵.

这个多变量随机系统模型(63)的特征是系统模型和噪声模型都有一个标量特征值多项式作分母, 分子是一个多项式矩阵或矩阵多项式.

### 5 多输入和多输出系统

前2节介绍的多变量系统输入输出表达, 在模型中更多的是把系统的多个输入作为一个整体(即输入向量  $u(t) \in \mathbf{R}^r$ ) 对待, 把系统的多个输出看作一个整体输出向量  $y(t) \in \mathbf{R}^m$  对待. 本节的“多输入和多输出系统”中间有一个“和”字, 本来也是多变量系统, 但这里把输入和输出分别对待, 而演变出多输入单输出系统、单输入多输出系统、多输入多输出系统, 它们都是多变量系统. 从系统辨识的角度, 这种分别对待很有意义.

考虑传递函数阵描述的  $r$  个输入  $m$  个输出线性多变量系统:

$$y(t) = G(z)u(t) + w(t),$$

其中  $u(t) = [u_1(t), u_2(t), \dots, u_r(t)]^T \in \mathbf{R}^r$  为系统输入向量,  $y(t) = [y_1(t), y_2(t), \dots, y_m(t)]^T \in \mathbf{R}^m$  为系统输出向量,  $w(t) = [w_1(t), w_2(t), \dots, w_m(t)]^T \in \mathbf{R}^m$  为零均值随机干扰噪声向量,  $G(z) \in \mathbf{R}^{m \times r}$  是系统模型的传递函数阵,  $z$  是单位前移算子:  $zy(t) = y(t +$

1) 和  $z^{-1}y(t) = y(t - 1)$ . 传递矩阵  $G(z)$  一般具有下列形式:

$$G(z) = \begin{bmatrix} g_{11}(z) & g_{12}(z) & \dots & g_{1r}(z) \\ g_{21}(z) & g_{22}(z) & \dots & g_{2r}(z) \\ \vdots & \vdots & & \vdots \\ g_{m1}(z) & g_{m2}(z) & \dots & g_{mr}(z) \end{bmatrix} \in \mathbf{R}^{m \times r}, \quad (65)$$

或

$$G(z) = \begin{bmatrix} \frac{b'_{11}(z)}{a'_{11}(z)} & \frac{b'_{12}(z)}{a'_{12}(z)} & \dots & \frac{b'_{1r}(z)}{a'_{1r}(z)} \\ \frac{b'_{21}(z)}{a'_{21}(z)} & \frac{b'_{22}(z)}{a'_{22}(z)} & \dots & \frac{b'_{2r}(z)}{a'_{2r}(z)} \\ \vdots & \vdots & & \vdots \\ \frac{b'_{m1}(z)}{a'_{m1}(z)} & \frac{b'_{m2}(z)}{a'_{m2}(z)} & \dots & \frac{b'_{mr}(z)}{a'_{mr}(z)} \end{bmatrix} \in \mathbf{R}^{m \times r},$$

其中  $g_{ij}(z)$  是第  $j$  个输入到第  $i$  个输出间的传递函数, 通常是一个阶次  $n_{ij}$  的有理分式:

$$g_{ij}(z) := \frac{b'_{ij}(z)}{a'_{ij}(z)} =$$

$$\frac{b_{ij}(0)z^{n_{ij}} + b_{ij}(1)z^{n_{ij}-1} + b_{ij}(2)z^{n_{ij}-2} + \dots + b_{ij}(n_{ij})}{z^{n_{ij}} + a_{ij}(1)z^{n_{ij}-1} + a_{ij}(2)z^{n_{ij}-2} + \dots + a_{ij}(n_{ij})}, \quad (66)$$

$a'_{ij}(z)$  与  $b'_{ij}(z)$  是  $z$  的  $n_{ij}$  次互质多项式(无公因子):

$$a'_{ij}(z) = z^{n_{ij}} + a_{ij}(1)z^{n_{ij}-1} + a_{ij}(2)z^{n_{ij}-2} + \dots + a_{ij}(n_{ij}),$$

$$b'_{ij}(z) = b_{ij}(0)z^{n_{ij}} + b_{ij}(1)z^{n_{ij}-1} + b_{ij}(2)z^{n_{ij}-2} + \dots + b_{ij}(n_{ij}).$$

对于辨识而言, 通常在式(66)右边分子分母同乘以  $z^{-n_{ij}}$ , 得到

$$g_{ij}(z) = \frac{z^{-n_{ij}}b'_{ij}(z)}{z^{-n_{ij}}a'_{ij}(z)} = \frac{b_{ij}(z)}{a_{ij}(z)}, \quad (67)$$

其中

$$a_{ij}(z) := z^{-n_{ij}}a'_{ij}(z) = 1 + a_{ij}(1)z^{-1} + a_{ij}(2)z^{-2} + \dots + a_{ij}(n_{ij})z^{-n_{ij}}, \quad (68)$$

$$b_{ij}(z) := z^{-n_{ij}}b'_{ij}(z) = b_{ij}(0) + b_{ij}(1)z^{-1} + b_{ij}(2)z^{-2} + \dots + b_{ij}(n_{ij})z^{-n_{ij}}. \quad (69)$$

于是, 传递矩阵  $G(z)$  可以写为

$$G(z) = \begin{bmatrix} \frac{b_{11}(z)}{a_{11}(z)} & \frac{b_{12}(z)}{a_{12}(z)} & \dots & \frac{b_{1r}(z)}{a_{1r}(z)} \\ \frac{b_{21}(z)}{a_{21}(z)} & \frac{b_{22}(z)}{a_{22}(z)} & \dots & \frac{b_{2r}(z)}{a_{2r}(z)} \\ \vdots & \vdots & & \vdots \\ \frac{b_{m1}(z)}{a_{m1}(z)} & \frac{b_{m2}(z)}{a_{m2}(z)} & \dots & \frac{b_{mr}(z)}{a_{mr}(z)} \end{bmatrix} \in \mathbf{R}^{m \times r}. \quad (70)$$

如果作用于系统(64)的噪声  $w(t) \in \mathbf{R}^m$  是通过零均值白噪声向量  $v(t) = [v_1(t), v_2(t), \dots, v_m(t)]^T \in \mathbf{R}^m$

驱动噪声模型传递函数阵  $\mathbf{H}(z) \in \mathbf{R}^{m \times m}$  的输出, 即  $\mathbf{w}(t) := \mathbf{H}(z)\mathbf{v}(t)$ , 那么系统(64)可以写为

$$\mathbf{y}(t) = \mathbf{G}(z)\mathbf{u}(t) + \mathbf{H}(z)\mathbf{v}(t). \quad (71)$$

类似地, 噪声模型传递矩阵有形式

$$\mathbf{H}(z) = \begin{bmatrix} \frac{d_{11}(z)}{c_{11}(z)} & \frac{d_{12}(z)}{c_{12}(z)} & \cdots & \frac{d_{1m}(z)}{c_{1m}(z)} \\ \frac{d_{21}(z)}{c_{21}(z)} & \frac{d_{22}(z)}{c_{22}(z)} & \cdots & \frac{d_{2m}(z)}{c_{2m}(z)} \\ \vdots & \vdots & & \vdots \\ \frac{d_{m1}(z)}{c_{m1}(z)} & \frac{d_{m2}(z)}{c_{m2}(z)} & \cdots & \frac{d_{mm}(z)}{c_{mm}(z)} \end{bmatrix} \in \mathbf{R}^{m \times m}, \quad (72)$$

其中

$$c_{ij}(z) := 1 + c_{ij}(1)z^{-1} + c_{ij}(2)z^{-2} + \cdots + c_{ij}(\mu_{ij})z^{-\mu_{ij}}, \quad (73)$$

$$d_{ij}(z) := 1 + d_{ij}(1)z^{-1} + d_{ij}(2)z^{-2} + \cdots + d_{ij}(\mu_{ij})z^{-\mu_{ij}}. \quad (74)$$

如果对式(70)的  $\mathbf{G}(z)$  和式(72)的  $\mathbf{H}(z)$  的分母分别进行通分得到最小公分母多项式, 分子就是矩阵多项式, 系统(70)就转化式(63)的类多变量随机系统; 如果对式(70)的  $\mathbf{G}(z)$  和式(72)的  $\mathbf{H}(z)$  使用右矩阵分式描述(右 MFD, Matrix-Fraction Description), 系统(70)就转化为式(37)的多变量 Box-Jenkins 随机系统.

## 5.1 传递函数阵主模型

假设式(70)传递矩阵  $\mathbf{G}(z)$  的最小公分母特征值多项式, 即分母的最小公倍数(LCM)多项式为  $\alpha(z) := \text{LCM}[a_{ij}(z), i=1, 2, \dots, m, j=1, 2, \dots, r] = 1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \cdots + \alpha_n z^{-n} \in \mathbf{R}$ .

令

$$Q_{ij}(z) := \frac{b_{ij}(z)}{a_{ij}(z)}\alpha(z) = \beta_{ij}(0) + \beta_{ij}(1)z^{-1} + \beta_{ij}(2)z^{-2} + \cdots + \beta_{ij}(n)z^{-n}.$$

那么式(70)传递矩阵  $\mathbf{G}(z)$  可以等价写为

$$\mathbf{G}(z) = \frac{1}{\alpha(z)} \begin{bmatrix} Q_{11}(z) & Q_{12}(z) & \cdots & Q_{1r}(z) \\ Q_{21}(z) & Q_{22}(z) & \cdots & Q_{2r}(z) \\ \vdots & \vdots & & \vdots \\ Q_{m1}(z) & Q_{m2}(z) & \cdots & Q_{mr}(z) \end{bmatrix} =: \frac{\mathbf{Q}(z)}{\alpha(z)} \in \mathbf{R}^{m \times r}.$$

将上式代入式(64)得到

$$\mathbf{y}(t) = \frac{\mathbf{Q}(z)}{\alpha(z)}\mathbf{u}(t) + \mathbf{w}(t), \quad (75)$$

这就是式(54)的类多变量输出误差类系统, 其中

$$\mathbf{Q}(z) := [\mathbf{Q}_{ij}(z)] = \mathbf{Q}_0 + \mathbf{Q}_1 z^{-1} + \mathbf{Q}_2 z^{-2} + \cdots + \mathbf{Q}_n z^{-n},$$

$$\mathbf{Q}_i := \begin{bmatrix} \beta_{11}(i) & \beta_{12}(i) & \cdots & \beta_{1r}(i) \\ \beta_{21}(i) & \beta_{22}(i) & \cdots & \beta_{2r}(i) \\ \vdots & \vdots & & \vdots \\ \beta_{m1}(i) & \beta_{m2}(i) & \cdots & \beta_{mr}(i) \end{bmatrix}, \quad i=0, 1, 2, \dots, n.$$

相对于分子系统辨识而言, 系统(75)一般采用直接辨识整个系统的参数, 故把系统(75)称为传递函数阵主模型(Main Model, MM).

## 5.2 传递函数阵子模型

将式(70)的传递函数阵  $\mathbf{G}(z)$  代入式(64)得到

$$\begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_m(t) \end{bmatrix} =$$

$$\begin{bmatrix} \frac{b_{11}(z)}{a_{11}(z)} & \frac{b_{12}(z)}{a_{12}(z)} & \cdots & \frac{b_{1r}(z)}{a_{1r}(z)} \\ \frac{b_{21}(z)}{a_{21}(z)} & \frac{b_{22}(z)}{a_{22}(z)} & \cdots & \frac{b_{2r}(z)}{a_{2r}(z)} \\ \vdots & \vdots & & \vdots \\ \frac{b_{m1}(z)}{a_{m1}(z)} & \frac{b_{m2}(z)}{a_{m2}(z)} & \cdots & \frac{b_{mr}(z)}{a_{mr}(z)} \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_r(t) \end{bmatrix} + \begin{bmatrix} w_1(t) \\ w_2(t) \\ \vdots \\ w_m(t) \end{bmatrix}. \quad (76)$$

这个多变量系统可以分解为  $m$  个多输入单输出系统, 式(76)的第  $i$  行就是第  $i$  个子系统, 即传递函数阵子模型(SM; SubModel),

$$y_i(t) = \frac{b_{i1}(z)}{a_{i1}(z)}u_1(t) + \frac{b_{i2}(z)}{a_{i2}(z)}u_2(t) + \cdots + \frac{b_{ir}(z)}{a_{ir}(z)}u_r(t) + w_i(t) = \sum_{j=1}^r \frac{b_{ij}(z)}{a_{ij}(z)}u_j(t) + w_i(t), \quad i=1, 2, \dots, m. \quad (77)$$

对于这  $m$  个子系统, 每个子系统的输出都无关联, 所以一个多变量系统可以看作多个多输入单输出子系统, 并且每个子系统具有相同的结构, 研究其中一个子系统的辨识问题也具有普遍意义. 后面我们将介绍多输入单输出系统.

**注3:** 对于第3节和第4节的所有多变量系统, 都不应该分解为  $m$  个子系统进行辨识, 否则会大幅度增加辨识算法的计算量.

## 5.3 传递函数阵子子模型

如果把子模型(77)右边各项  $\frac{b_{ij}(z)}{a_{ij}(z)}u_j(t)$  看作一个子系统, 即子子模型(Sub-SubModel, SSM), 也就



是把传递函数阵  $\mathbf{G}(z)$  各元  $\frac{b_{ij}(z)}{a_{ij}(z)}$  看作子子模型, 设其输出为  $y_{ij}(t)$ , 则有

$$y_{ij}(t) = \frac{b_{ij}(z)}{a_{ij}(z)} u_j(t), \quad i=1, 2, \dots, m, \quad j=1, 2, \dots, r.$$

这个子子模型的输入  $u_j(t)$  是可得到的, 但输出  $y_{ij}(t)$  是不可测的, 它与可测输出  $y_i(t)$  的关系如下:

$$y_i(t) = \sum_{j=1}^r y_{ij}(t) = y_{i1}(t) + y_{i2}(t) + \dots + y_{ir}(t) + w_i(t).$$

因为只有输入数据  $\{u_j(t)\}$  和输出数据  $\{y_i(t)\}$  可得到, 子子模型输出  $\{y_{ij}(t)\}$  是未知, 造成辨识的困难, 解决的途径是用辅助模型辨识思想或辅助变量方法. 与子系统辨识方法相对应, 把直接辨识传递函数阵  $\mathbf{G}(z)$  各元参数的方法称为子子系统辨识方法(子子模型参数辨识方法). 因此, 一个  $r$  输入  $m$  输出多变量系统共有  $mr$  个单输入单输出子子系统(子子模型).

辨识的目标是利用系统输入输出数据  $\{u_j(t), y_i(t)\}$  和适当的辨识方法估计系统的未知参数  $a_{ij}(l)$  和  $b_{ij}(l), l=1, 2, \dots, n_{ij}$ .

## 5.4 多输入单输出系统模型

为了变量符号简便, 把式(77)的一个多输入单输出子系统表示为

$$y(t) = \frac{B_1(z)}{A_1(z)} u_1(t) + \frac{B_2(z)}{A_2(z)} u_2(t) + \dots + \frac{B_r(z)}{A_r(z)} u_r(t) + w(t) = \sum_{j=1}^r \frac{B_j(z)}{A_j(z)} u_j(t) + w(t), \quad (78)$$

其中  $u_j(t) \in \mathbf{R}, j=1, 2, \dots, r$ , 是系统的  $r$  个输入,  $y(t) \in \mathbf{R}$  是系统输出,  $w(t) \in \mathbf{R}$  为零均值白噪声或有色噪声, 多项式  $A_j(z)$  和  $B_j(z)$  定义为

$$A_j(z) = 1 + a_j(1)z^{-1} + a_j(2)z^{-2} + \dots + a_j(n_j)z_j^{-n_j}, \\ B_j(z) = b_j(0) + b_j(1)z^{-1} + b_j(2)z^{-2} + \dots + b_j(n_j)z_j^{-n_j}.$$

根据式(78)中分母多项式  $A_j(z)$  是否相同, 以及噪声  $w(t)$  的不同性质, 又可演化出下列多输入方程误差类模型和多输入输出误差类模型.

### 5.4.1 多输入方程误差类模型

设式(78)中分母多项式  $A_j(z)$  都相同, 且为

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

有多输入方程误差类模型:

$$A(z)y(t) = B_1(z)u_1(t) + B_2(z)u_2(t) + \dots + B_r(z)u_r(t) + w(t), \quad (79)$$

其中  $w(t) \in \mathbf{R}$  是零均值白噪声或有色噪声.

1) 多输入方程误差模型(多输入 CAR 模型或多输入 ARX 模型):

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

2) 多输入 CARMA 模型(多输入 ARMAX 模型):

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

带积分多输入 CARMA 模型(多输入 CARIMA 模型或多输入 ARIMAX 模型)为

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

3) 多输入 CARAR 模型(多输入 ARARX 模型, 也称多输入动态调节模型):

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

4) 多输入 CARARMA 模型(多输入 ARARMAX 模型):

$$A(z)y(t) = \sum_{j=1}^r B_j(z)u_j(t) + \frac{D(z)}{C(z)}v(t). \quad (80)$$

令

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

则多输入 CARARMA 系统(80)的辨识模型为

$$y(t) = \boldsymbol{\varphi}^T(t)\boldsymbol{\theta} + v(t), \\ n_0 := n + n_1 + n_2 + \dots + n_r + r + n_c + n_d, \\ \boldsymbol{\theta} := [a_1, a_2, \dots, a_n, b_1(0), b_1(1), \dots, b_1(n_1), \\ b_2(0), b_2(1), \dots, b_2(n_2), \dots, b_r(0), b_r(1), \dots, \\ b_r(n_r), c_1, c_2, \dots, c_{n_c}, d_1, d_2, \dots, d_{n_d}]^T \in \mathbf{R}^{n_0}, \\ \boldsymbol{\varphi}(t) := [-y(t-1), -y(t-2), \dots, -y(t-n), \\ u_1(t), u_1(t-1), \dots, u_1(t-n_1), u_2(t), u_2(t-1), \\ \dots, u_2(t-n_2), \dots, u_r(t), u_r(t-1), \\ \dots, u_r(t-n_r), -w(t-1), -w(t-2), \dots, \\ -w(t-n_c), v(t-1), v(t-2), \dots, \\ v(t-n_d)]^T \in \mathbf{R}^{n_0}.$$

### 5.4.2 多输入输出误差类模型

根据式(78), 多输入输出误差类模型具有下列形式:

$$y(t) = \sum_{j=1}^r \frac{B_j(z)}{A_j(z)} u_j(t) + w(t), \quad (81)$$

其中  $w(t) \in \mathbf{R}$  是零均值白噪声和有色噪声向量.

1) 多输入输出误差模型(multi-input OE model):

$$y(t) = \sum_{j=1}^r \frac{B_j(z)}{A_j(z)} u_j(t) + v(t).$$

文献[32]提出了多输入单输出输出误差系统的

辅助模型多新息随机梯度算法。

2) 多输入输出误差滑动平均模型 (multi-input OEMA model):

$$y(t) = \sum_{j=1}^r \frac{B_j(z)}{A_j(z)} u_j(t) + D(z)v(t).$$

3) 多输入输出误差自回归模型 (multi-input OEAR model):

$$y(t) = \sum_{j=1}^r \frac{B_j(z)}{A_j(z)} u_j(t) + \frac{1}{C(z)} v(t).$$

4) 多输入 Box-Jenkins 模型 (multi-input Box-Jenkins model):

$$y(t) = \sum_{j=1}^r \frac{B_j(z)}{A_j(z)} u_j(t) + \frac{D(z)}{C(z)} v(t). \quad (82)$$

5) 一般多输入输出误差模型 (多输入自回归 Box-Jenkins 模型, multi-input AR-BJ model):

$$F(z)y(t) = \sum_{j=1}^r \frac{B_j(z)}{A_j(z)} u_j(t) + \frac{D(z)}{C(z)} v(t). \quad (83)$$

写出以上 5 系统的辨识模型, 需定义  $r$  个中间

变量:  $x_j(t) = \frac{B_j(z)}{A_j(z)} u_j(t), j=1, 2, \dots, r.$

当  $A_j(z)$  都相同, 且为  $A(z)$  时, 上述多输入输出误差类模型退化为

1) 多输入输出误差模型 (multi-input OE model):

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

文献[33-34]分别提出了这个多输入单输出 (MISO) 输出误差系统的偏差补偿递推最小二乘方法和辅助模型多新息最小二乘方法。

2) 多输入输出误差滑动平均模型 (multi-input OEMA model):

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

3) 多输入输出误差自回归模型 (multi-input OEAR model):

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

4) 多输入 Box-Jenkins 模型 (multi-input Box-Jenkins model):

$$y(t) = \frac{1}{A(z)} \sum_{j=1}^r B_j(z) u_j(t) + \frac{D(z)}{C(z)} v(t). \quad (84)$$

5) 一般多输入输出误差模型 (多输入自回归 Box-Jenkins 模型, multi-input AR-BJ model):

$$F(z)y(t) = \frac{1}{A(z)} \sum_{j=1}^r B_j(z) u_j(t) + \frac{D(z)}{C(z)} v(t). \quad (85)$$

写出以上 5 个系统的辨识模型, 需定义一个中

间变量:  $x(t) = \frac{1}{A(z)} \sum_{j=1}^r B_j(z) u_j(t).$

5.4.3 特殊多输入方程误差模型

当模型(83)中多项式取不同值时, 就得到一些特殊的方程误差模型。

1) 多输入 FIR 模型 (multi-input FIR model):

$$y(t) = \sum_{j=1}^r B_j(z) u_j(t) + v(t).$$

2) 多输入受控滑动平均模型 (多输入 C-MA 模型, multi-input C-MA model):

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

3) 多输入受控自回归滑动平均模型 (多输入 C-ARMA 模型, multi-input C-ARMA model):

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

5.4.4 特殊多输入输出误差模型

当模型(83)中多项式取不同值时, 就得到一些特殊的输出误差模型。

1) 多输入自回归输出误差模型 (multi-input AR-OE model):

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

或

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

2) 多输入自回归输出误差滑动平均模型 (multi-input AR-OEMA model):

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

或

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

3) 多输入自回归输出误差自回归模型 (multi-input AR-OEAR model):

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

或

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

5.5 单输入多输出系统模型

尽管在控制系统中, 单输入多输出系统不太常见, 研究意义不是很大, 但在通讯领域研究得很多。因为一个信号源发出的信号, 可以在不同的地方接收到, 相当于多输出。

在前面的多输入多输出系统模型中,将输入向量  $\mathbf{u}(t) \in \mathbf{R}^r$  的维数  $r$  置为 1,就得到单输入多输出系统.由多变量 FIR 模型得到单输入多输出 FIR 模型:

$$\mathbf{y}(t) = \mathbf{B}(z)\mathbf{u}(t) + \mathbf{v}(t), \quad (86)$$

或

$$\begin{cases} y_1(t) = B_1(z)u(t) + v_1(t), \\ y_2(t) = B_2(z)u(t) + v_2(t), \\ \vdots \\ y_m(t) = B_m(z)u(t) + v_m(t), \end{cases}$$

相应地,有单输入多输出输出误差模型:

$$\begin{cases} y_1(t) = \frac{B_1(z)}{A_1(z)}u(t) + v_1(t), \\ y_2(t) = \frac{B_2(z)}{A_2(z)}u(t) + v_2(t), \\ \vdots \\ y_m(t) = \frac{B_m(z)}{A_m(z)}u(t) + v_m(t), \end{cases}$$

当  $A_i(z) = A(z)$  相同时(或进行通分),可得

$$\begin{cases} y_1(t) = \frac{B_1(z)}{A(z)}u(t) + v_1(t), \\ y_2(t) = \frac{B_2(z)}{A(z)}u(t) + v_2(t), \\ \vdots \\ y_m(t) = \frac{B_m(z)}{A(z)}u(t) + v_m(t). \end{cases}$$

在盲辨识方法研究中,也经常假设系统是单输入多输出系统.

尽管多输入多输出系统和多输入单输出系统有很多辨识方法,但是在多变量系统辨识建模实验中,也经常采用分别改变每个输入进行实验,就得到单输入多输出系统.

多变量系统的参数辨识可以采用下列 2 种途径进行.设系统工作处在平衡点(或称稳定工作点),这时的输入输出可认为是零(零均值化),一种途径是改变所有输入信号,用数据采集仪记录所有输入和输出的值,用多输入多输出系统辨识方法估计系统参数,这种方法的不足之处是要准备多个( $r$ 个)信号发生器(设备),同时产生  $r$  个不相关持续激励信号作为输入,操作复杂,增加了设备投资;另一种途径操作比较简单,只需要一个输入信号发生器,每次实验只改变某一通道输入信号,固定其他通道的输入(认为输入为零),用数据采集仪记录这个通道

的输入和所有通道的输出值,分别对每一个输入通道进行实验,这时的系统是一个单输入多输出(SIMO)系统,共需  $r$  次实验,共得到  $r$  个单输入多输出子系统,这种方法可以辨识出多变量系统(76)每个子模型的传递函数.所以从实用地观点看,研究单输入多输出多变量系统辨识方法是必要的.

只改变第  $j$  个输入  $u_j(t)$ ,其他不改变的输入认为是零  $u_i(t) = 0, i \neq j$ ,由多输入系统多输出(76)得到一个单输入多输出系统:

$$\begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_m(t) \end{bmatrix} = \begin{bmatrix} \frac{b_{1j}(z)}{a_{1j}(z)} \\ \frac{b_{2j}(z)}{a_{2j}(z)} \\ \vdots \\ \frac{b_{mj}(z)}{a_{mj}(z)} \end{bmatrix} u_j(t) + \begin{bmatrix} w_1(t) \\ w_2(t) \\ \vdots \\ w_m(t) \end{bmatrix}, j = 1, 2, \dots, r. \quad (87)$$

### 5.6 马可夫参数或脉冲响应模型

设系统(64)稳定,采用长除法可将式(70)传递矩阵  $\mathbf{G}(z)$  的各元  $\frac{b_{ij}(z)}{a_{ij}(z)}$  展开成幂级数,则  $\mathbf{G}(z)$  可以表示为

$\mathbf{G}(z) = \mathbf{J}_0 + \mathbf{J}_1 z^{-1} + \mathbf{J}_2 z^{-2} + \dots + \mathbf{J}_n z^{-n} + \dots + \mathbf{J}_i z^{-i} + \dots$ , 其中  $\mathbf{J}_i$  是待估计的系统参数矩阵,称为马可夫参数矩阵(Markov parameter matrix)或脉冲响应参数矩阵(impulse response parameter matrix).上式可近似表示为

$$\mathbf{G}(z) = \mathbf{J}_0 + \mathbf{J}_1 z^{-1} + \mathbf{J}_2 z^{-2} + \dots + \mathbf{J}_N z^{-N}.$$

于是,多变量系统(64)可以近似表示为有限脉冲相应参数模型,

$$\mathbf{y}(t) = [\mathbf{J}_0 + \mathbf{J}_1 z^{-1} + \mathbf{J}_2 z^{-2} + \dots + \mathbf{J}_N z^{-N}] \mathbf{u}(t) + \mathbf{w} = \mathbf{J}_0 \mathbf{u}(t) + \mathbf{J}_1 \mathbf{u}(t-1) + \mathbf{J}_2 \mathbf{u}(t-2) + \dots + \mathbf{J}_N \mathbf{u}(t-N) + \mathbf{w}(t). \quad (88)$$

由于假设系统稳定,故当  $i \rightarrow \infty$  时,  $\mathbf{J}_i \rightarrow 0$ . 因此,只要  $N$  充分大,式(88)可以满足任何精度要求,即以任意精度逼近系统(64).

该模型待辨识参数数目为  $S = mrN$ ,它取决于所要求的参数估计精度和系统的快速性( $N$ 的大小).马可夫参数模型是一种非参数模型,其精度随  $N$  不同而异,通常不便直接衡量,一般把估计的马可夫参数转化成等价的参数模型(如状态空间模型),来衡量参数估计精度.

视干扰噪声为 MA 过程  $\mathbf{w}(t) := \mathbf{D}(z)\mathbf{v}(t)$ , AR 过程  $\mathbf{w}(t) := \mathbf{C}^{-1}(z)\mathbf{v}(t)$  和 ARMA 过程  $\mathbf{w}(t) :=$

$C^{-1}(z)D(z)v(t)$ , 由式(89)可以得到不同有色噪声干扰的有限脉冲相应参数模型.

## 6 结语

控制学科已经渗透到社会和经济各个发展各个领域. 在不同领域, 为了方便处理问题, 人们经常使用不同的数学模型, 但是有些数学模型在某种条件下是等价的, 如状态空间模型与确定性 ARMA 模型, 而且同一个模型在不同领域有不同的名称, 甚至同一名称表示不同的模型. 本文把随机系统模型分为 3 大类: 时间序列模型、方程误差类模型和输出误差类模型, 使系统模型的定义和辨识算法的类别变得十分清晰. 根据单输入或多输入、单输出或多输出, 白噪声干扰或有色噪声干扰等, 详细总结了标量系统和多变量系统的一些基本描述模型. 尽管这些模型有成百上千种, 但是每一个模型的构成元素十分明了, 记住这些模型的名称、模型的结构和含义就十分容易, 一目了然.

## 参考文献

### References

- [ 1 ] 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
- [ 2 ] 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
- [ 3 ] 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
- [ 4 ] 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
- [ 5 ] 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
- [ 6 ] 丁锋. 时变参数系统辨识及其应用 [D]. 北京: 清华大学自动化系, 1994  
DING Feng. Time-varying parameter system identification and its applications [D]. Beijing: Department of Automation, Tsinghua University, 1994
- [ 7 ] 谢新民, 丁锋. 自适应控制系统 [M]. 北京: 清华大学出版社, 2002  
XIE Xinmin, DING Feng. Adaptive control systems [M]. Beijing: Tsinghua University Press, 2002
- [ 8 ] 丁锋. 系统辨识理论与方法 [M]. 北京: 电力出版社, 2011  
DING Feng. System identification theory and methods [M]. Beijing: China Electric Power Press, 2011
- [ 9 ] 周毅, 丁锋. 依等价 AR 模型阶次递增的自回归滑动平均模型辨识 [J]. 华东理工大学学报: 自然科学版, 2008, 34(3): 425-431  
ZHOU Yi, DING Feng. Identification of ARMA models with the equivalent AR model order increasing [J]. Journal of East China University of Science and Technology: Natural Science Edition, 2008, 34(3): 425-431
- [ 10 ] 丁锋. 辨识 Box-Jenkins 模型参数的递推广义增广最小二乘法 [J]. 控制与决策, 1990, 5(6): 53-56  
DING Feng. Recursive generalized extended least squares (RGELS) method of identifying the Box-Jenkin model [J]. Control and Decision, 1990, 5(6): 53-56
- [ 11 ] Wang D Q, Ding F. Input-output data filtering based recursive least squares parameter estimation for CARARMA systems [J]. Digital Signal Processing, 2010, 20(4): 991-999
- [ 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, Ding J. Least squares parameter estimation with irregularly missing data [J]. International Journal of Adaptive Control and Signal Processing, 2010, 24(7): 540-553
- [ 14 ] Ding F, Liu X P, Liu G. Gradient based and least-squares based iterative identification methods for OE and OEMA systems [J]. Digital Signal Processing, 2010, 20(3): 664-677
- [ 15 ] 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
- [ 16 ] 王冬青. 基于辅助模型的递推广义增广最小二乘辨识方法 [J]. 控制理论与应用, 2009, 26(1): 51-56  
WANG Dongqing. Recursive extended least squares identification method based on auxiliary models [J]. Control Theory and Applications, 2009, 26(1): 51-56
- [ 17 ] Ding F, Liu X P, Liu G. Auxiliary model based multi-innovation extended stochastic gradient parameter estimation with colored measurement noises [J]. Signal Processing, 2009, 89(10): 1883-1890
- [ 18 ] Ding F, Liu X P, Liu G. Multi-innovation least squares i-

- dentification for system modeling [J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 2010, 40(3): 767-778
- [19] 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
- [20] Wang D Q, Yang G W, Ding F. Gradient-based iterative parameter estimation for Box-Jenkins systems with finite measurement data [J]. Computers & Mathematics with Applications, 2010, 60(5): 1200-1208
- [21] 王冬青, 丁锋. Box-Jenkins 模型的基于辅助模型的多新息广义增广随机梯度算法[J]. 控制与决策, 2008, 23(9): 999-1003, 1010  
WANG Dongqing, DING Feng. Auxiliary model based multi-innovation generalized extended stochastic gradient (AM-MI-GESG) algorithms for Box-Jenkins models [J]. Control and Decision, 2008, 23(9): 999-1003, 1010
- [22] Wang L Y, Xie L, Wang X F. The residual based interactive stochastic gradient algorithms for controlled moving average models [J]. Applied Mathematics and Computation, 2009, 211(2): 442-449
- [23] Xiao Y S, Zhang Y, Ding J, et al. The residual based interactive least squares algorithms and simulation studies [J]. Computers & Mathematics with Applications, 2009, 58(6): 1190-1197
- [24] 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
- [25] 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
- [26] 丁锋, 谢新民. 线性多变量系统的联合辨识算法[J]. 控制理论与应用, 1992, 9(5): 545-550  
DING Feng, XIE Xinmin. Combined identification algorithms for linear multivariable systems [J]. Control Theory and Applications, 1992, 9(5): 545-550
- [27] Ding F, Chen T. Hierarchical gradient-based identification of multivariable discrete-time systems [J]. Automatica, 2005, 41(2): 315-325
- [28] Ding F, Chen T. Hierarchical least squares identification methods for multivariable systems [J]. IEEE Transactions on Automatic Control, 2005, 50(3): 397-402
- [29] Liu Y J, Sheng J, Ding R F. Convergence of stochastic gradient algorithm for multivariable ARX-like systems [J]. Computers & Mathematics with Applications, 2010, 59(8): 2615-2627
- [30] Han H Q, Xie L, Ding F, et al. Hierarchical least squares based iterative identification for multivariable systems with moving average noises [J]. Mathematical and Computer Modelling, 2010, 51(9/10): 1213-1220
- [31] Zhang Z N, Ding F, Liu X G. Hierarchical gradient based iterative parameter estimation algorithm for multivariable output error moving average systems [J]. Computers & Mathematics with Applications, 2011, 61(3): 672-682
- [32] 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
- [33] 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
- [34] Ding F, Chen T, Qiu L. Bias compensation based recursive least squares identification algorithm for MISO systems [J]. IEEE Transactions on Circuits and Systems-II: Express Briefs, 2006, 53(5): 349-353

## System identification. Part B: Basic models for system description

DING Feng<sup>1,2</sup>

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

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

**Abstract** Control is the core of all scientific issues. Mathematical models are the basis of all control problems. The movement law of things described by equations is the mathematical model. The development of different disciplines is to establish the process of their mathematical models. This paper divides the mathematical models of the linear dynamic systems into three categories: the time-series models, the equation error type models and the output error type models, and introduces basic mathematical models of linear systems in detail, including the discretization of continuous-time system models and model equivalence transform, single-input single-output stochastic system, multivariable systems, multivariable-like systems, multiple-input and multiple-output systems such as the main model, submodel and sub-submodel of the transfer function matrix, multiple-input single-output systems, and single-input multiple-output systems.

**Key words** mathematical model; system identification; parameter estimation; CAR model; ARX model; CARMA model; ARMAX model; CARARMA model; OEMA model; OEAR model