

系统辨识

柯益华 编著

四川大学出版社



系统辨识

柯益华 编著

四川大学出版社

责任编辑:王泽彬
封面设计:罗光
版式设计:王泽彬

图书在版编目(CIP)数据

系统辨识 / 柯益华编著. —成都: 四川大学出版社,
1994.7 (2003.6 重印)

ISBN 7-5614-2634-8

I. 系... II. 柯... III. 系统辨识-高等学校-教材 IV. N945.14

中国版本图书馆 CIP 数据核字 (2003) 第 051557 号

书名 系统辨识

作者 柯益华
出版 四川大学出版社
地址 成都市一环路南一段 24 号 (610065)
印刷 新都华兴印务有限公司
发行 四川大学出版社
开本 780 mm×1 092 mm 1/16
印张 14.625
字数 338 千字
版次 1994 年 7 月第 1 版
印次 2003 年 6 月第 2 次印刷
印数 801~1 300 册
定价 18.00 元

◆读者邮购本书,请与本社发行科
联系。电话:85408408/85401670/
85408023 邮政编码:610065

◆本社图书如有印装质量问题,请
寄回印刷厂调换。

◆网址:www.scupress.com.cn

版权所有◆侵权必究

前 言

系统辨识作为一种独立的建模方法,自 60 年代诞生以来,其发展极为迅速,已在空间技术、生产过程、生物工程、经济与社会系统等广阔领域中取得大量应用成果,成为现代控制理论及应用中充满活力的重要分支。因此,《系统辨识》是高等学校自动化类专业本科生和研究生的一门重要选修课程。

本书是在成都科技大学自动控制专业讲授《系统辨识》课程的讲义和总结十余年教学经验的基础上编写而成的。

考虑到工程专业的特点,本书内容以适用于控制的线性系统数学模型辨识为主,同时兼顾其它方面的发展概况,对系统辨识的全过程进行了比较系统的归纳。

全书共十章和一个附录。第一章概述了系统辨识的目的、定义和基本过程;第二章介绍系统辨识中适宜采用的线性系统模型类。第三至六章是全书的重点,比较详细的讲解了线性系统参数估计的确定性方法、概率性方法和辨识非参数模型的相关分析法。为了进一步了解系统辨识近年来的一些新的发展方向,安排在第七章概括介绍非线性系统辨识的若干基本方法;在第八章讨论闭环系统辨识出现的问题和解决的途径。最后,在第九章中专门分析了模型结构辨识问题;在第十章中讨论了作为系统辨识的基础和归宿的两个重要内容——辨识的试验设计与模型验证。附录中列出了随机过程的有关结论和矩阵运算与分析的有关公式。

本书适于作为自动控制类专业研究生和高年级大学生选修课的教材。考虑到本科选修课学时的限制,对本书目录中注有*号的章节可以不作要求。

由于编者学识水平的限制,书中若有不妥之处,敬请读者批评指正。

柯益华

1994 年 3 月

目 录

第一章 概论

1.1 建模与辨识的一般概念	(1)
1.1.1 模型的种类及用途	(1)
1.1.2 两类建模方法	(2)
1.1.3 系统辨识的定义	(3)
1.1.4 模型的可辨识性	(5)
1.2 常用数学模型分类	(5)
1.2.1 静态模型与动态模型	(5)
1.2.2 线性模型与非线性模型	(6)
1.2.3 参数模型与非参数模型	(7)
1.2.4 连续时间模型与离散时间模型	(7)
1.3 辨识过程的基本环节	(7)
1.3.1 辨识目的和验前知识	(8)
1.3.2 试验设计	(9)
1.3.3 模型结构的确定	(9)
1.3.4 参数估计	(10)
1.3.5 模型验证	(10)

第二章 线性动态系统数学模型的典范表达式

2.1 单输入-单输出线性系统的典范表达式	(12)
2.1.1 连续时间典范状态模型	(12)
2.1.2 离散时间典范状态模型	(13)
2.1.3 典范输入-输出差分模型	(15)
2.2 线性多变量系统确定性典范状态模型	(17)
2.2.1 状态模型的等价关系	(17)
2.2.2 两种能观典范结构	(18)
2.2.3 能观典范型结构的待辨识参数量	(21)
2.3 随机性典范状态模型	(23)
2.3.1 具有等值噪声的随机状态模型	(24)
2.3.2 随机状态模型的典范结构	(25)
2.4 确定性典范差分模型	(25)
2.4.1 扩展状态模型	(26)
2.4.2 扩展状态模型的等价性	(28)
2.4.3 典范型差分模型	(28)
2.5 随机性典范差分模型	(29)
2.5.1 随机性典范差分模型的一般表达式	(29)
2.5.2 随机差分模型的特殊类型	(30)

2.6	预测误差模型	(32)
2.6.1	随机典范差分模型的转换	(32)
2.6.2	随机状态模型的转换	(33)

第三章 线性系统参数估计的最小二乘法

3.1	参数的最小二乘估计量	(34)
3.1.1	最小二乘结构	(34)
3.1.2	最小二乘解	(35)
3.1.3	加权最小二乘估计	(38)
3.2	最小二乘估计的性质	(38)
3.2.1	点估计理论的有关结论	(38)
3.2.2	最小二乘估计的统计特性	(41)
3.2.3	最小二乘估计的一致性条件	(45)
3.3	静态模型参数最小二乘估计算法	(48)
3.3.1	一维线性回归分析	(48)
3.3.2	多维线性回归分析	(52)
3.3.3	平方根估计算法	(55)
3.4	动态模型参数最小二乘估计算法	(59)
3.4.1	离线辨识的一次性算法	(59)
3.4.2	过程噪声模型及其相关性	(61)
3.4.3	广义最小二乘法	(63)
3.4.4	辅助变量法	(65)
3.4.5	多阶段最小二乘法	(65)

第四章 递推最小二乘估计算法

4.1	最小二乘估计的递推形式	(68)
4.1.1	递推估计的基本格式	(68)
4.1.2	递推计算的初值选择	(70)
4.1.3	递推算法举例及程序框图	(71)
4.2	时变参数的适应性算法	(73)
4.2.1	增长记忆算法存在的问题	(73)
4.2.2	渐消记忆的递推算法	(74)
4.2.3	限定记忆的递推算法	(75)
4.3	相关噪声时的递推算法	(77)
4.3.1	递推广义最小二乘法	(77)
4.3.2	递推辅助变量法	(78)
4.3.3	扩展最小二乘法	(79)
4.3.4	随机逼近法	(80)
* 4.4	数值稳定的递推算法	(83)
4.4.1	递推算法中的数值稳定问题	(83)
4.4.2	递推平方根估计算法	(84)
4.4.3	误差矩阵的平方根分解算法	(86)

4.4.4	误差矩阵的 U-D 分解算法	(91)
* 4.5	递推算法的收敛性分析	(95)
4.5.1	递推算法的一般格式	(95)
4.5.2	收敛性分析的工具	(96)
4.5.3	收敛性分析的结果	(97)

第五章 参数估计的概率性方法

5.1	极大似然估计法原理	(100)
5.1.1	极大似然估计法的目标函数	(100)
5.1.2	不同观测条件下的似然函数	(101)
5.2	预测误差估计法原理	(102)
5.2.1	预测误差模型	(102)
5.2.2	预测误差估计的目标函数	(103)
5.3	正态分布下的极大似然估计	(103)
5.3.1	似然函数的变量替换	(104)
5.3.2	正态分布下的似然函数	(105)
5.4	线性系统参数极大似然估计离线迭代算法	(107)
5.4.1	参数的极大似然估计量	(107)
5.4.2	参数估计的离线迭代算法	(109)
5.5	近似极大似然估计递推算法	(110)
5.5.1	模型及目标函数	(110)
5.5.2	目标函数的近似公式	(111)
5.5.3	参数估计量的递推算式	(112)
* 5.6	预测误差估计量的一致性和渐近正态性	(114)
5.6.1	预测误差估计量的一致性	(114)
5.6.2	预测误差估计量的渐近正态性	(116)

第六章 线性系统非参数模型相关分析辨识法

6.1	相关分析辨识法原理	(118)
6.1.1	模型与目标函数	(118)
6.1.2	相关函数与极小均方误差估计量	(119)
6.2	二位式伪随机信号的产生及性质	(121)
6.2.1	二位式随机序列的基本特性	(121)
6.2.2	二位式伪随机 M 序列的产生	(122)
6.2.3	二位式 M 序列的自相关函数	(124)
6.2.4	其它伪随机序列信号	(126)
6.3	脉冲响应函数的辨识算法	(129)
6.3.1	逐点估计算法	(130)
6.3.2	多点估计的一次性算法	(131)
6.3.3	多点估计的递推算法	(132)
6.4	脉冲响应估计量的统计性质	(134)
6.4.1	估计量的无偏性	(134)

6.4.2	估计量的一致性	(135)
6.5	相关分析辨识法在参数估计中的应用	(136)
6.5.1	白噪声输入时的估计算法	(136)
6.5.2	其它输入信号时的估计算法	(138)

* 第七章 非线性系统辨识概论

7.1	非线性系统模型结构	(140)
7.1.1	Volterra 泛函级数模型	(141)
7.1.2	分块组合模型	(141)
7.1.3	非线性差分模型	(143)
7.2	非线性静态模型参数估计	(144)
7.2.1	迭代算法原理	(145)
7.2.2	常用的优化迭代算法	(146)
7.2.3	直接搜索法	(149)
7.3	非线性系统的非参数模型辨识法	(152)
7.3.1	Volterra 核的辨识方法	(152)
7.3.2	Wiener 核的辨识方法	(154)
7.4	非线性系统参数模型的辨识	(155)
7.4.1	Hammerstein 模型的辨识	(155)
7.4.2	非线性差分模型的辨识	(158)
7.5	非线性解析模型的参数估计	(159)
7.5.1	参数估计的最优化方法	(160)
7.5.2	状态和参数的联合估计方法	(161)

* 第八章 闭环系统辨识概论

8.1	闭环辨识的基本概念	(163)
8.1.1	闭环辨识的必要性	(163)
8.1.2	闭环辨识的特殊问题	(164)
8.1.3	闭环辨识方法分类	(165)
8.2	线性单变量系统的间接辨识	(165)
8.2.1	闭环系统的 ARMA 模型	(165)
8.2.2	过程参数的可辨识条件	(167)
8.2.3	切换调节器的辨识方法	(168)
8.3	线性单变量系统的直接辨识	(169)
8.3.1	无外部输入信号的参数可辨识条件	(169)
8.3.2	有外部输入信号的直接辨识法	(170)
8.4	线性多变量系统的闭环辨识	(171)
8.4.1	多变量系统闭环可辨识性定义	(171)
8.4.2	多变量系统闭环可辨识的结论	(173)
8.4.3	闭环辨识的精度问题	(175)
8.5	闭环系统联合过程谱分解辨识法	(177)
8.5.1	联合过程的传递矩阵分解	(177)

8.5.2 联合过程的谱分解辨识法	(178)
-------------------------	-------

第九章 模型结构的辨识

9.1 模型结构辨识的概念	(179)
9.2 线性单变量系统差分模型结构直观检验法	(180)
9.2.1 纯时延时间的确定方法	(180)
9.2.2 差分阶次的直观检验法	(180)
9.2.3 模型参数估计值按阶次的递推算法	(181)
9.3 线性单变量系统差分阶次的显著性检验法	(183)
9.3.1 模型阶的假设检验及统计量	(183)
9.3.2 统计量的分布特性及自由度	(184)
9.3.3 模型阶的 F 检验法	(185)
* 9.4 线性单变量系统差分阶次检验的信息准则	(186)
9.4.1 AIC 法定阶的原理	(186)
9.4.2 AIC 定阶法的应用	(187)
9.4.3 AIC 法与 F 检验法的关系	(188)
* 9.5 线性静态模型输入量逐步回归选择法	(189)
9.5.1 输入量在回归模型中的作用	(189)
9.5.2 偏残差平方和的计算公式	(190)
9.5.3 输入量重要性的 F 检验	(191)
9.5.4 逐步回归的计算步骤	(192)
* 9.6 线性多变量系统状态模型结构的辨识	(194)
9.6.1 线性多变量系统的输入输出关系	(194)
9.6.2 结构不变量的辨识	(199)

第十章 辨识的试验设计与模型验证

10.1 辨识试验设计概论	(201)
10.1.1 试验设计的内容	(201)
10.1.2 试验设计的验前知识	(202)
10.2 最优输入信号的设计	(203)
10.2.1 最优输入信号设计准则	(203)
10.2.2 基于特殊 CAR 模型的最优输入信号设计	(204)
10.2.3 基于一般 CAR 模型的最优输入信号设计	(207)
10.2.4 递推输入信号的设计	(208)
10.3 试验数据的滤波处理	(209)
10.4 采样周期及试验长度设计	(210)
10.4.1 采样周期 T_s 的选择	(210)
10.4.2 试验长度 N 的选择	(211)
10.5 模型的验证方法	(212)
10.5.1 仿真检验法	(212)
10.5.2 实际检验法	(213)

附录

附录 A	随机过程的有关结论.....	(214)
附录 B	向量与矩阵的导数公式.....	(218)
附录 C	矩阵公式.....	(221)
附录 D	矩阵分析	(222)

参考文献

第一章 概论

本章概要介绍系统辨识的任务和辨识过程的基本环节,为掌握系统辨识方法建立初步的概念.对于初学者来说,树立总体概念是非常重要的.

1.1 建模与辨识的一般概念

1.1.1 模型的种类及用途

在自然科学、工程技术和社会科学中,都广泛使用模型来描述被研究的对象.

所谓模型(model),就是用来表征被研究对象或系统所具有的特性或运动规律的工具.模型着重于描述对象或系统的行为,而不着重于刻画其实际结构或构造.

随着人类社会的发展,人们研究的对象在广度和深度上不断扩展,研究方法不断进步,模型的种类也越来越多.采用何种形式的模型来表示一个对象或系统,与所研究的学科领域、研究的目的和对被研究对象认识的深度有关.从广义的角度来看,模型的种类和用途可以归纳如下.

(1) 概念模型

概念模型是用语言和文字来描述被研究的事物,表现为某种观念或范畴的规定性.概念模型在任何一门学科和人们日常生活中都普遍使用.特别是在哲学和某些社会科学中,它几乎是唯一能够利用的模型.在自然科学中,概念模型是一种定性的模型.

(2) 物理模型

物理模型是一种形象的和可观察的实体模型,包括实物模型和模拟模型两类.

实物模型是依据模型简化规则制作的一种被描述对象或系统的缩小复制品,并使其结构和行为精确地等同于原系统.例如用于风洞试验的飞机模型,用于水力学实验的水利工程模型等.

模拟模型是一种在结构上不同于原系统,但其行为与原系统相似、易于制作的简单装置,用来模拟不易制作的复杂系统.例如,用电阻和电容组成的电路,其输出电压对输入电压的响应,是和一个节流器与气容元件组成的气路系统中的气压响应以相同方式进行的,而前者比后者易于实现.

在物理模型上进行实验与观察,可以获得在制作模型时还不了解的对象特性.因此,物理模型具有比较丰富的内涵.

物理模型的缺点主要有两个方面:一是在数量的精度上受模型制作技术水平的限制,不能满足现代科学技术的要求;二是模型的通用性较差,外延较窄,一个模型只能描述一个或一类对象,导致模型成本昂贵.此外,某些特殊系统至今尚无法用物理模型来描述,如社会经济系统、环境生态系统等等.这些缺点限制了物理模型的应用范围.

(3) 数学模型

数学模型是一种用来描述被研究系统的内部及其有关的外部因素之间所服从的定量关系的数学结构,表现为不同形式的方程、图形或数表.

数学模型的优点,除了可能满足数量关系上精确性的要求外,还具有宽广的外延.对于内容和结构上完全不同的系统,只要被研究的各种量之间的关系服从相同的数学结构,就可以用同一个数学模型来描述.例如,一个二阶线性常系数微分方程,既可以表示RLC电路的振荡特性,又可以描述由质量、阻尼器和弹簧组成的力学系统的动态行为.因此,虽然数学模型在形式上更抽象,它却更本质、更普遍地概括出实际系统的性质或行为.

数学模型的出现和应用,是利用模型进行研究的科学方法的一大进步.马克思认为,数学是一种精确的科学语言.任何一门自然科学和技术科学,当它还不能使用数学语言来描述时,就不能称为一门科学,至少不能算作一门精确的科学.这个思想在恩格斯的《自然辩证法》一书也有记载.恩格斯在当时的情况下就看到,数学“在刚体力学中是绝对的,在气体力学中是近似的,在液体力学中就已经比较困难了;在物理学中是试验性的和相对的;在化学中是最简单的一次方程式,在生物学中等于零”.经过近百年来科学技术的飞跃发展,数学已经广泛地渗透到各门学科之中.电子计算机的出现和广泛应用,使数学的应用领域和价值发生了革命性的变化.除了自然科学以外,像经济学、心理学、社会学和军事科学等,都已成为使用数学模型和计算机的热门领域.

(4) 结构模型

结构模型主要描述系统内部的结构特点和因果关系,往往表现为网络的形式.如运筹学中的活动网络模型,大系统理论中的布尔网络模型,经济系统分析中的解释性结构模型和生物系统中的房室模型等.

(5) 仿真模型

仿真模型是一种计算机程序.数学模型和结构模型都可以按适当的仿真语言演化成仿真模型.在计算机辅助制造技术中,大量使用仿真模型.

在控制理论中,数学模型具有重要意义.没有数学模型就没有系统分析与综合.建立数学模型的方法也是控制理论的重要分支.

1.1.2 两类建模方法

寻求一个用来描述被研究对象或系统的行为所服从的数学结构或数学模型的过程,我们称为建模(Model Building)或模型化(Modelling).建模方法可以分为理论分析方法和试验分析方法两大类.

理论分析方法建模,在控制理论诞生之前,已在许多学科广泛应用,并获得很大的成就.它是通过具体地分析各个实际系统内在过程的运动机理,运用各种专门学科的基本原理和规律,特别是有关物质、能量、动量的守恒性与连续性原理来建立描述该实际系统的数学模型.这种方法所建立的数学模型称为系统的机理模型.

应当指出,用理论分析方法如果能够获得系统的机理模型,当然是合乎理想的.但其可能性与精确性却受到人们对系统结构及其运动规律的瞭解程度的限制.尽管近百年来各门学科都有了长足的进步,但对客观世界各种复杂系统的运动规律,人们的认识仍然是非常有限的.例如,在化学反应过程中,催化作用机理和催化剂的活性问题,在微生物的生

长与代谢过程中,酶的作用机理和酶的活性问题,迄今为止人们都还没有真正理解,无法用理论分析方法获得这类过程的机理模型.在控制系统中,当被研究的系统规模越来越大,系统构成越来越复杂时,企图用理论分析方法建立机理模型是非常困难的.还应当看到,建立机理模型所依据的系统内部运动规律,都是在一定的假设条件下获得的,必然包含了对环境条件影响等次要因素的忽略与简化,从而带来误差.因此,机理模型不一定是十分精确的.一个全面考虑一切影响因素的精确的机理模型,即使能够建立,它必然是非常复杂的高阶模型,在控制理论中称之为“维数的灾难”.因为一个维数过高的数学模型,不但在理论分析中难于处理,而且在实践上更难于实现.

上述情况说明,用理论分析方法建立机理模型,主要应是各专门学科自身的研究任务,而且并非总是有效的.

试验分析方法是一种具有普遍意义的跨学科的建模方法.早期的试验分析方法仅限于利用试验数据确定机理模型中的待定参数,还没有形成独立的建模方法.本世纪四十年代,当经典控制理论形成的时候,为了确定被控制对象的响应特性或频率特性,也采用了试验分析方法,提出了著名的“黑箱”概念,即人们无需知道系统的内部结构和运动机理,凭借试验取得的输出数据就可以拟合出该系统的外部特性所服从的规律,获得描述系统行为的传递函数或频率特性函数.但是,这种数学模型能否反映系统内部的运动状态,当时还未解决.

1960年以来,在以状态空间描述法、动态规划和极大值原理为基础的现代控制理论的推动下,用试验分析方法建模逐渐发展成为控制理论的一个重要分支——辨识(Identification).特别是 R. E. Kalman 的状态空间理论提出了系统状态的能控性、能观测性概念,证明了完全能控、完全能观系统的状态模型与其输入、输出模型的等价性.这就从理论上证明,根据系统的输入、输出数据可以判明系统中完全能控、完全能观部分的结构和状态行为,从而奠定了用辨识方法建立适合现代控制理论需要的数学模型的基础.近三十年来,辨识方法得到迅速发展,乃至成为现代控制理论的一个基本问题.

实际上,建立数学模型的方法常常是理论建模与系统辨识相结合.辨识方法可以在对系统没有任何先验知识的情况下,凭借观测数据构造数学模型,因而比理论建模具有更大的可能性和现实性.但是,当人们对实际系统中的某些子系统的运动规律已经掌握时,就可以先建立这些子系统的机理模型,然后再用辨识方法建立其余部分的数学模型,或者可以利用对实际系统运动规律的某些初步认识,确定模型类和模型结构,辨识的任务便简化为估计模型中的参数.总之,理论分析可以使辨识过程大为简化.特殊地说,当我们无法通过试验取得观测数据的时候,或当被建模的系统尚不存在、不能进行试验的时候,理论分析方法就成为获得数学模型的唯一途径.

1.1.3 系统辨识的定义

关于系统辨识的定义,已有各种不同的提法.下面介绍两种主要的定义.

定义 1.1 系统辨识是在输入和输出观测的基础上,从指定的一类模型中确定一个与被辨识系统等价的模型.

定义 1.1 是美国加利福尼亚大学 Zadeh 在 1962 年提出的.它表明对一个系统用辨识

方法建立数学模型,必须具备三个要素:规定模型类,取得观测信息和提出目标函数.

(1) 模型类 \mathcal{M}

挑选哪一类模型来描述被辨识的系统,属于规定模型类的问题.被挑选的模型类应具有有效性和实用性.

模型类的有效性是指模型应能反映被辨识系统的全部实质性信息,模型应有必要的精确度,模型应有唯一性.

实用性的要求包括:模型应尽可能采用典型型结构,使待辨识参数最少,以求简化辨识算法;模型实现的费用低;模型在应用时具有广泛性和一定的灵活性.

实际上,由于系统本身的复杂性,模型在结构上不可能做到与被辨识系统同构,只能要求模型与系统的主要特性同态.因此,由系统辨识方法得到的拟合模型,在一定程度上都存在未建模的动力学特性.

(2) 观测信息 φ

观测信息主要包括输入信息 \mathcal{U} 和输出信息 \mathcal{Y} . 由于观测信息 φ 是用辨识方法建模的最重要的依据,在 \mathcal{Y} 中应含有系统主要振型的特性. 然而,对于大多数实际生产过程而言,都是一种因果系统. 在 \mathcal{Y} 中是否包含被辨识系统的特性,主要取决于 \mathcal{U} 的性质. 一个恒定的或低频的输入信号,能使输出信息中载有系统的稳态增益或低频特性等信息,但无法反映系统的高频特性. 这就要求输入信号的频带宽度必须覆盖系统的主要频段,一般要求 \mathcal{U} 应为持续激励信号,且为系统的容许输入信号.

(3) 目标函数 \mathcal{J}

目标函数是由某种意义上的误差函数 e 构成的非负函数 $J[e(\cdot)]$,用以作为拟合模型优劣程度的直接量度. 通常采用的误差函数是图 1.1 所示的输出误差形式.

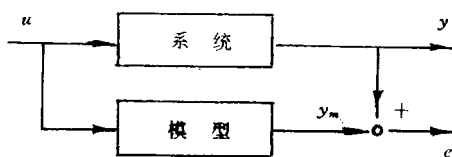


图 1.1 输出误差函数形式

设 $m \in \mathcal{M}$ 是系统 S 的一个模型, $u \in \mathcal{U}$ 是 S 的一个容许输入. 设

$$\left. \begin{aligned} u^i &= \{u(i), i \leq k\} \\ y(k) &= f(u^i, k), \forall k \\ y_m(k) &= f_m(u^i, k), \forall k \\ e(k) &= y(k) - y_m(k) \end{aligned} \right\} \quad (1.1)$$

目标函数 $J[e(\cdot)]$ 的含义是当容许控制 u 施加于系统 S 和模型 m 后,输出误差 e 引起的损失. 故 \mathcal{J} 又称为损失函数.

\mathcal{J} 的选择应满足三个条件:

- ① $J[e(\cdot)] \geq 0$;
- ② $J[0] = 0$;
- ③ 如果 $\|e_1(\cdot)\| < \|e_2(\cdot)\|$, 则有

$$J[e_1(\cdot)] \leq J[e_2(\cdot)], \forall m_1, m_2 \in \mathcal{M}$$

式中

$$e_1(k) = y(k) - y_{m_1}(k), \quad e_2(k) = y(k) - y_{m_2}(k)$$

显然,二次型函数能满足上述要求.

定义 1.2 系统辨识是从载有随机噪声的输入-输出观测数据构成系统实验模型的全过程.

定义 1.2 的特点在于强调系统辨识所依据的观测信息 φ 中不可避免的含有随机噪声,因而辨识过程中必然要涉及对随机噪声的描述和处理;系统辨识所获得的是实验模型,其内容包括从试验设计到模型验证的多个环节,要反复多次迭代才能完成. 这个定义的意义在于划清了系统辨识与单纯的确定性试验数据处理之间的界线.

1.1.4 模型的可辨识性

可辨识性问题的实质,是判断用辨识方法获得的拟合模型,与被辨识系统是否具有某种意义下的等价性. 为此,需要给出模型与系统等价的定义.

定义 1.3 对于模型 $m^* \in \mathcal{M}$,如果在某个容许输入 $u \in \mathcal{U}$ 下,有

$$J[e^*(\cdot)] \leq J[e(\cdot)], \forall m \in \mathcal{M}$$

式中

$$e^*(k) = y(k) - y_{m^*}(k)$$

则称模型 m^* 与被辨识系统 S 等价.

现在,可以把定义 1.1 规定的系统辨识问题转化为下述最优化问题.

系统辨识是基于被辨识系统 S 的观测信息 φ ,按目标函数 J 达到最小的原则,从模型类 \mathcal{M} 中拟合出一个模型 $m^* \in \mathcal{M}$,作为被辨识系统 S 的数学模型.

对于这个优化问题,需要回答的问题是:

(1) 是否存在 $m^* \in \mathcal{M}$,使得

$$J[e^*(\cdot)] = J_{\inf}[e(\cdot)], \forall m \in \mathcal{M}$$

(2) 如果存在最小解,它是否唯一?

(3) 唯一最小解是否与容许输入 $u \in \mathcal{U}$ 的选择有关?

如果对于任意容许输入 $u \in \mathcal{U}$,上述最优化问题都有解且唯一,则称模型类 \mathcal{M} 是可辨识的;如果模型结构已知,模型类定义为一组模型参数的集合,则模型可辨识问题就简化为参数可辨识问题.

1.2 常用数学模型的分类

数学模型的结构形式有很多种类型. 本节介绍在控制系统中通常采用的模型结构的分类方法,以及在本课程中将要涉及的模型种类.

1.2.1 静态模型与动态模型

按照数学模型所描述的系统信号与时间的关系,可以分为静态模型与动态模型两类.

如果在数学模型中不考虑时间的因素,只描述系统的输入量、输出量和状态变量这些

信号之间所服从的数学关系,称为静态模型.静态模型通常可以用代数方程表示.

一般情况下,系统在受到扰动后的暂态过程中,其输入量、输出和状态变量这些信号都是时间的函数,它们之间的关系也与时间有关系,必须用微分方程或差分方程才能描述这些信号的运动状态,称为动态模型.

例如,在状态空间描述法中,状态方程和输出方程

$$\left. \begin{aligned} \dot{x}(t) &= f[x(t), u(t), t] \\ y(t) &= g[x(t), u(t), t] \end{aligned} \right\} \quad (1.2)$$

就是系统的动态模型.当系统进入稳定状态以后,状态变量 x , 输入量 u 和输出量 y 都不随时间而变化, $\dot{x}(t)=0$, (1.2)式变为代数方程

$$\left. \begin{aligned} f[x, u] &= 0 \\ y &= g[x, u] \end{aligned} \right\} \quad (1.3)$$

即为系统的静态模型,它是动态模型的特例.

在控制系统中主要使用动态模型,但静态模型也有很大用途.有的系统本身就是静态的;有的系统暂态过程变化极其缓慢,以致可以看作是静态的;有的系统暂态过程变化极快,在控制时可以只考虑其静态的关系;有些特别复杂的系统,如果无法建立其动态模型,便只能利用静态模型研究稳定状态下各信号之间的关系.例如,在大系统的多级递阶控制中,其最高一级的控制所采用的数学模型往往是静态的.

1.2.2 线性模型与非线性模型

按照数学模型所描述的系统中信号或参数等变量之间的组合关系,可以分为线性模型与非线性模型两类.

如果数学模型中只含有变量及其各阶导数的一次项,模型中的某一变量是由其它变量线性组合而成的,称之为线性模型;否则,即称之为非线性模型.例如,当(1.2)和(1.3)式中的 f 和 g 为任意非线性函数时,即为非线性状态模型.当 f 和 g 为下述线性函数时:

$$\left. \begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned} \right\} \quad (1.4)$$

即称为线性状态模型.信号 $x(t)$, $u(t)$, $y(t)$ 之间的线性组合关系由参数矩阵 A, B, C, D 决定.

一般而论,实际系统都是非线性系统,应采用非线性模型来描述.所谓线性模型只不过是在一定假设条件下采用某种线性化方法处理的结果.但是,由于非线性系统问题本身的复杂性,无论在系统的分析、设计或辨识等领域中,都还没有一般的解决方法.因此,设法将非线性系统线性化,使之能够采用比较成熟的线性系统的方法来解决有关问题,仍然是十分可取的.经典的线性化方法有变量代换法、经 Taylor 级数展开的切线法和割线法等.近年来基于现代微分几何理论发展起来的坐标变换和反馈线性化,是一种全局精确线性化方法,具有十分重要的理论意义和应用价值.

1.2.3 参数模型与非参数模型

按照数学模型所描述的系统信号之间的关系是否具有明确的数学结构,可以分为参数模型与非参数模型两类.

当系统的信号之间的数学关系可由理论分析或验前知识确定为某种结构关系式,且其中包含或多或少的待定参数,即称为参数模型.由式(1.4)表示的状态模型就是一种参数模型.其结构为一阶线性微分方程组和代数方程,待定参数为 A, B, C, D 矩阵中的元素.

当系统的信号之间的数学关系不显含一定的结构关系式和参数,而是以图线或数表的形式来表示时,称为非参数模型.例如,频率响应和脉冲响应曲线都是非参数模型.当然,非参数模型可以通过一定的处理方法转化为参数模型.

1.2.4 连续时间模型与离散时间模型

按照数学模型所描述的系统信号取值的时间类型,可以分为连续时间模型与离散时间模型两类.

在实际系统中,其信号取值的时间类型,客观上存在连续时间和离散时间两种情况.例如,在生产系统中的化工过程、热工过程、电力传输过程、电力拖动过程等都是连续时间过程;而在经济系统中,价格的涨落过程、商品的批发或零售过程等就是离散时间过程.连续时间的动态系统可用传递函数、微分方程来描述,称为连续时间模型;离散时间的动态系统可用脉冲传递函数、差分方程来描述,称为离散时间模型.

在系统辨识中,为了辨识模型结构和估计模型参数,涉及对大量输入输出观测数据的处理,主要使用计算机作为计算工具,因而只能使用离散模型和离散型数据.为此,对于采用连续时间模型的系统,应经离散化处理,转化为离散时间模型.

对于连续时间系统的不同数学描述形式,有不同的离散化方法.常用的离散化方法主要有 Z 变换法、采样保持法和差分法等.例如,连续时间系统用传递函数描述时,可经零阶保持采样后,对传递函数进行 Z 变换,即可得离散时间系统的脉冲传递函数;连续时间系统用状态空间模型描述时,在零阶保持采样条件下解状态方程,可获得离散时间状态空间模型;连续时间系统的输入-输出关系用微分方程进行时域描述时,在一定的采样周期下用差商代替微商,可获得离散时间系统的差分方程.

在本课程中,主要讨论离散时间的线性动态系统参数模型的辨识问题,同时对静态模型、非参数模型和非线性系统数学模型的辨识,也将作简要的介绍.

此外,按照被描述系统本身的某些特性,还有确定性模型与随机性模型、定常参数模型与时变参数模型、集中参数模型与分布参数模型的区别.

1.3 辨识过程的基本环节

系统辨识作为一种建模方法,是一个完整的过程,其内容可划分为若干基本环节.各环节之间的相互关系可用图1.2来表示.