

《信息、控制与系统》系列教材

随 机 控 制

郭尚来 编著

清 华 大 学 出 版 社

(京)新登字 158 号

内 容 简 介

随机控制理论是控制理论的一个重要分支,已渗透到各个领域。本书着重介绍随机控制理论的基础知识,包括随机过程、随机系统模型、最小方差控制、最优状态估计和最优随机控制等内容。选材注意基础性和实用性,基本内容论述严谨,理论表述由浅入深,由简到繁,符合工科学子的认识规律。每章都配有例题和习题。

本书适合作为工科大学本科生和研究生教材,也可供科技工作者学习参考。

版权所有,翻印必究。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

图书在版编目(CIP)数据

随机控制/郭尚来编著. —北京:清华大学出版社, 1999

信息、控制与系统系列教材

ISBN 7-302-03555-5

. 随... . 郭... . 随机控制-教材 . 0231

中国版本图书馆 CIP 数据核字(1999)第 14324 号

出版者: 清华大学出版社(北京清华大学校内, 邮编 100084)

<http://www.tup.tsinghua.edu.cn>

印刷者: 昌平国马印刷厂

发行者: 新华书店总店北京发行所

开 本: 787x 1092 1/16 印张: 16.5 字数: 386 千字

版 次: 1999 年 11 月第 1 版 1999 年 11 月第 1 次印刷

书 号: ISBN 7-302-03555-5/TP · 1954

印 数: 0001 ~ 3000

定 价: 18.00 元

出版说明

《信息、控制与系统》系列教材是一套关于信息、控制和系统学科的基本理论和应用技术的高等学校教材。选题范围包括信号和信息处理、模式识别、知识工程、控制理论、自动化技术、传感技术、自动化仪表、系统理论、系统工程、机器人控制、智能控制、计算机应用和控制等方面。主要读者对象为自动控制、计算机、过程自动化、无线电等系科的高年级大学生和研究生,以及在这些领域和部门工作的科学工作者和工程技术人员。

信息、控制与系统科学是在 20 世纪上半叶形成和发展起来的新兴学科。它们的应用和影响已经遍及众多的部门和领域,贯穿其中的许多思想和方法已用于经济和社会现象的研究,而以这些学科为理论基础的自动化技术的广泛应用更是实现现代化的重要标志之一。这套系列教材,正是在这样的客观要求下,为适应教学和科研工作的需要而编写和出版的。它以清华大学自动化系近年来经过教学实践的新编教材为主,力求反映这些学科的基本理论和最新进展,并且反映清华大学在这些学科中科学研究和教学研究的成果。我们希望这套系列教材,既是在校大学生和研究生的教科书,也能为广大科技人员提供有价值的参考。

组编和出版这套教材是一次尝试。我们热忱欢迎选用本系列教材的老师、学生和科技工作者提出批评、建议。

《信息、控制与系统》系列教材编委会

《信息、控制与系统》系列教材编委会

主	编	常 迥		
编	委	常 迥	童诗白	方崇智
		韩曾晋	李衍达	郑大钟
		夏绍玮	徐培忠	
责任编辑	蔡鸿程	王仁康		

前 言

在自然界以及人类的生产和生活中,“随机”问题到处可见。例如,天气预报就是随机问题的一个例子,一个地区的天气状况是个随机量,受到本地区和邻边地区各种因素,包括各种随机因素的影响,只有全面综合这些因素的作用,才能准确地做出天气预报。又如,股民“炒股”是人们经济生活中的一个随机问题,只有全面分析政治、经济和市场等各种随机因素的影响,掌握股票走势,做出正确预测的股民才有可能获得较好的经济收益。这些都属于随机控制中的预报、预测问题。

随着社会经济活动的更加丰富,生产和科学技术的深入发展,随机控制理论已渗透到各个领域。因此,掌握随机控制理论的基础知识,学会用“随机”的观点分析问题和解决问题,正在逐渐成为众多科技工作者必备的知识素质。

目前,随机控制理论已成为自动控制理论的一个重要分支。数字计算机的出现使得大量的计算变得简捷、快速和可能,同时也使随机控制理论的研究和应用得到了长足的进步。随机控制理论专门研究具有噪声(随机过程)扰动的动力学系统,包括随机过程、系统辨识、状态估计、随机控制和自适应控制等内容。

本书作为本科生或研究生的基础教材,仅着重于介绍随机控制理论的基础知识,包括随机过程、随机系统模型、最小方差控制、最优状态估计和最优随机控制等内容。其他如系统辨识和参数估计、自适应控制等内容,已有很多书籍论述,在本书中不再介绍。

人们往往容易对随机控制理论产生误解,以为用“均值”代替对应的随机变量,就可用确定性控制代替随机控制。实际上,完全用确定性控制系统的方法来研究随机控制系统是不够的,而且,对某些用确定性控制观点难于处理的问题,却能从随机控制的观点加以分析。目前,随机控制理论远没有像确定性控制理论那样成熟,在深度和广度上都有待进一步研究和发展。

在笔者多年教学工作中,感到对工科学生来说,除了授予基础的知识外,还要从工程角度出发,引导学生对该课程中的基本内容进行比较严格的证明,这样才能深入理解本学科内容,以便能在实际工作中灵活运用。本书在编写中注意编入一些较新的内容,例如采用随机微分方程方法处理连续时间最优估计问题,采用连续时间贝尔曼方程直接求解连续时间最优随机控制问题,而不是采用惯用的离散时间极限化间接方法。

随机控制的初学者可能会感到随机控制“定义多、理论多、公式多”,不易掌握。为此,本书在理论表述上注意了由浅入深、由简到繁的原则,如开始只介绍不带控制项的简单模型的解,然后再讨论带有控制项和带有量测方程的情况;先介绍简单的最小方差预测和控制问题,再考虑状态方程的状态估计和随机控制问题。另外,书中还举了一些例题,并有习题供读者选做,有条件的读者可在计算机上安排一些试验,这将有助于掌握和加深理解本书的内容。

一般地说,学习随机控制,读者需具备概率论、线性代数和控制原理的基本知识。本书编写中尽量在不多的概率论知识基础上介绍随机过程的基本知识,并展开其他方面的讨论。阅读本书不需确定性最优控制的基础,没有学习过最优控制的读者,会通过阅读本书,掌握随机控制的基本内容;对于已经掌握最优控制知识的读者,定会对本书内容有更深的理解,实际应用也会更加灵活。

在本书编写和教学过程中,得到同事和同学们的热情支持和帮助,在此表示谢意。由于作者水平有限,不足与疏漏之处,请读者指正。

郭尚来
1997 年 12 月于清华园

目 录

第 1 章	绪论	1
1.1	随机控制的研究对象	1
1.2	随机控制的研究内容	2
1.3	本书内容	3
1.4	随机控制理论的发展概况	3
	习题	5
第 2 章	随机过程	7
2.1	引言	7
2.2	随机过程概念	7
2.2.1	事物分类	7
2.2.2	随机过程定义	8
2.2.3	随机过程分类	8
2.2.4	随机过程描述	9
2.3	工程中常用的随机过程	13
2.3.1	平稳随机过程	13
2.3.2	高斯随机过程	15
2.3.3	马尔可夫过程	18
2.3.4	二阶随机过程	19
2.3.5	独立增量过程	19
2.3.6	维纳过程	21
2.4	协方差函数	22
2.4.1	对称阵及其正定性	22
2.4.2	施瓦茨(Schwartz)不等式	23
2.4.3	协方差函数的几个重要性质	23
2.4.4	平稳随机过程情况	24
2.5	谱密度	26
2.5.1	确定性函数的谱密度	26
2.5.2	随机过程的谱密度	27
2.5.3	互谱密度	32
2.5.4	复谱密度	33
2.6	白噪声	33
2.6.1	白噪声概念	33
2.6.2	带限白噪声和有色噪声	35

2.6.3	连续时间平稳高斯白噪声与维纳过程的关系	36
2.6.4	标准均匀分布伪随机数的生成	37
2.6.5	高斯白噪声的生成	39
2.7	随机分析.....	41
2.7.1	收敛性	41
2.7.2	均方连续性	43
2.7.3	均方可微性	44
2.7.4	均方可积性	45
2.7.5	随机积分	46
2.7.6	伊藤(Itô)微分规则	48
习题	50
第3章	随机系统的数学模型	56
3.1	引言.....	56
3.2	受控自回归平移平均模型.....	56
3.3	离散时间随机状态模型.....	58
3.3.1	随机状态模型和假设	58
3.3.2	模型的解	59
3.4	连续时间随机状态模型.....	62
3.4.1	随机状态模型和假设	62
3.4.2	模型的解	63
3.4.3	关于两种模型的讨论	65
3.5	连续时间与离散时间随机状态模型之间的转换.....	67
3.5.1	连续时间随机状态模型的离散化	67
3.5.2	离散时间随机状态模型的极限化	70
3.6	输入量是随机序列的离散时间系统分析.....	72
3.6.1	系统脉冲响应	72
3.6.2	系统的解	72
3.6.3	平稳过程情况	73
3.6.4	新息的概念	77
3.7	离散时间过程的谱分解.....	78
3.7.1	有理谱密度和谱分解举例	78
3.7.2	谱分解定理	79
3.7.3	谱表示定理	80
3.8	输入量是随机过程的连续时间系统分析.....	82
3.8.1	解的存在性	82
3.8.2	系统的解	82
3.8.3	平稳过程情况	83
3.8.4	高斯白噪声输入情况	84

3.9	连续时间过程的谱分解.....	85
3.9.1	谱分解定理	85
3.9.2	谱表示定理	86
习题	87
第4章	最小方差控制	97
4.1	引言.....	97
4.2	简例.....	98
4.2.1	例	98
4.2.2	结果验证	99
4.2.3	几点说明	99
4.3	ARMA 模型的最优预测	100
4.3.1	简例	100
4.3.2	p 步预测	102
4.3.3	$F(q^{-1})$ 和 $G(q^{-1})$ 的计算	103
4.4	最小方差控制策略.....	106
4.5	次最优控制策略.....	110
4.5.1	稳定性分析	110
4.5.2	次最优控制策略	111
4.6	CARMA 模型与状态空间模型的转换	115
习题	117
第5章	最优状态估计	120
5.1	引言.....	120
5.2	状态估计与条件均值的等价性.....	121
5.2.1	引理	121
5.2.2	定理	123
5.3	多维高斯分布的某些结果及其几何解释.....	124
5.3.1	定理 5.3.1	124
5.3.2	定理 5.3.2	124
5.3.3	定理 5.3.3	126
5.3.4	三个定理的物理意义和几何解释	126
5.3.5	定理 5.3.4(投影定理)	128
5.3.6	定理 5.3.5	129
5.4	离散时间随机系统的预测和滤波.....	132
5.4.1	离散时间随机状态模型	132
5.4.2	一步最优预测	133
5.4.3	最优滤波	136
5.4.4	p 步最优预测	139

5.5	离散时间随机系统的最优平滑.....	140
5.5.1	最优平滑分类	140
5.5.2	一步最优平滑	141
5.5.3	二步最优平滑	143
5.5.4	固定区间最优平滑	144
5.5.5	固定点最优平滑	147
5.5.6	固定滞后最优平滑	149
5.6	有色噪声情况下的最优估计.....	152
5.6.1	成形滤波器	153
5.6.2	模型噪声为有色噪声情况下的最优滤波	154
5.6.3	模型噪声和量测噪声相关的最优滤波	155
5.6.4	量测噪声为有色噪声时的最优滤波	157
5.7	最优估计的几个问题.....	159
5.7.1	稳定性问题	159
5.7.2	模型误差分析	163
5.7.3	发散问题	164
5.8	连续时间随机状态模型的状态估计.....	167
5.8.1	连续时间随机状态模型和新息过程	167
5.8.2	连续时间随机状态模型的最优滤波	169
5.8.3	最优预测	173
5.8.4	最优平滑	173
	习题	175
第 6 章	最优随机控制	185
6.1	引言.....	185
6.2	离散时间随机状态控制系统模型.....	188
6.2.1	模型的描述	188
6.2.2	最优性原理	190
6.2.3	高斯向量二次型的均值	191
6.3	离散时间 LQG 完全状态信息情况	192
6.4	离散时间 LQG 不完全状态信息情况	197
6.5	离散时间 LQG 控制系统小结	204
6.5.1	变分恒等式	204
6.5.2	各种控制情况	205
6.6	离散时间 LQG 最优跟随器	207
6.6.1	问题的描述	207
6.6.2	最优控制策略算法	208
6.7	连续时间确定性最优控制.....	210
6.7.1	一般贝尔曼方程	210

6.7.2	线性贝尔曼方程	211
6.8	连续时间 LQG 完全状态信息情况	215
6.8.1	系统模型	215
6.8.2	贝尔曼方程	216
6.8.3	随机线性调节器的开环控制	219
6.9	连续时间 LQG 不完全状态信息情况	222
6.9.1	问题描述	222
6.9.2	等效模型及其解	224
6.9.3	分离定理	227
6.10	连续时间 LQG 控制系统小结	229
6.10.1	变分恒等式.....	229
6.10.2	各种控制情况.....	230
6.11	对偶性	232
6.11.1	离散时间系统对偶性.....	233
6.11.2	连续时间系统对偶性.....	234
6.12	里卡蒂方程的求解	235
6.12.1	差分方程法.....	235
6.12.2	矩阵微分方程法.....	236
	习题	241
	参考文献	248

第 1 章 绪 论

1.1 随机控制的研究对象

考虑确定性、线性、离散时间控制系统,可用差分方程表示为

$$\begin{aligned} y(k) + a_1 y(k-1) + a_2 y(k-2) + \dots + a_n y(k-n) \\ = b_0 u(k-d) + b_1 u(k-1-d) + \dots + b_m u(k-m-d) \end{aligned} \quad (1.1.1)$$

式中: $y(k)$ 是系统输出; $u(k)$ 是系统输入; k 是时间步数; 输入到输出有 d 步延迟; n 和 m 分别为 $y(k)$ 和 $u(k)$ 的阶数; a_1, a_2, \dots, a_n 和 b_0, b_1, \dots, b_m 是系数。这样的控制系统也可用状态空间方程表示为

$$x(k+1) = \Phi x(k) + \Gamma u(k) \quad (1.1.2)$$

$$y(k) = C x(k) \quad (1.1.3)$$

式(1.1.2)称为对象模型, 式(1.1.3)称为量测(输出)模型。式中: $x(k)$ 是 $n \times 1$ 状态向量; $y(k)$ 是 $m \times 1$ 量测向量; $u(k)$ 是 1×1 控制向量; Φ 是 $n \times n$ 状态转移阵; C 是 $m \times n$ 量测阵; Γ 是 $n \times 1$ 控制阵。它们可以是时不变的, 也可以是时变的。

如果控制系统所处的环境存在噪声(随机扰动), 也就是控制系统受到噪声的作用, 或者说控制系统受到噪声的污染, 那么在式(1.1.1), (1.1.2)和(1.1.3)中应加入相应的项, 分别得到

$$\begin{aligned} y(k) + a_1 y(k-1) + a_2 y(k-2) + \dots + a_n y(k-n) \\ = b_0 u(k-d) + b_1 u(k-1-d) + \dots + b_m u(k-m-d) + e(k) \end{aligned} \quad (1.1.4)$$

$$x(k+1) = \Phi x(k) + \Gamma u(k) + \omega(k) \quad (1.1.5)$$

$$y(k) = C x(k) + v(k) \quad (1.1.6)$$

式中: $e(k)$ 是噪声, 可以是单项式, 也可以是多项式; $\omega(k)$ 是 $n \times 1$ 模型噪声向量; $v(k)$ 是 $m \times 1$ 量测(输出)噪声向量。式(1.1.4), (1.1.5)和(1.1.6)称为随机控制系统。

用上述同样方法可描述连续时间控制系统。考虑确定性、线性、连续时间控制系统, 可用状态空间方程表示为

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t) \quad (1.1.7)$$

$$y(t) = Hx(t) \quad (1.1.8)$$

式(1.1.7)称为对象模型, 式(1.1.8)称为量测(输出)模型。式中: $x(t)$ 是 $n \times 1$ 状态向量; $y(t)$ 是 $m \times 1$ 量测向量; $u(t)$ 是 1×1 控制向量; t 是连续时间; A, B 和 H 分别是 $n \times n, n \times 1$ 和 $m \times n$ 系数阵, 它们可以是时不变的, 也可以是时变的。如果控制系统中存在噪声, 则在式(1.1.7)和(1.1.8)中应加入相应的项, 得到

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t) + \omega(t) \quad (1.1.9)$$

$$y(t) = H x(t) + v(t) \quad (1.1.10)$$

式中: $v(t)$ 是 $n \times 1$ 模型噪声向量; $w(t)$ 是 $m \times 1$ 量测噪声向量。式(1.1.9)和(1.1.10)称为连续时间随机控制系统。

综上所述,随机控制系统就是其扰动可以用随机过程明确表示的受控动力学系统。随机控制(理论)研究的就是这种系统。它分析动力学系统和系统变量的统计特性,研究如何实现参数最优化,并对随机控制系统给定一个性能指标,寻求使性能指标最小的最优控制律。

1.2 随机控制的研究内容

以状态空间模型为例,完整的随机控制系统如图1.2.1所示。图中各部分及有关问题都是随机控制(理论)的研究内容,具体包括下列各部分:

(1) 分析动力学系统和系统变量的统计特性

随机动力学系统中存在噪声,噪声是一种随机过程,这就使状态向量和量测(输出)向量在噪声作用(污染)下,也变成了随机过程。随机过程不同于确定性过程,要用统计特性描述。无疑,为分析和研究随机动力学系统,了解随机动力学系统和系统变量的统计特性是必要的,这是分析和研究随机动力学系统的数学理论基础。

图 1.2.1 随机控制系统

(2) 系统辨识和参数估计

为研究和设计随机控制系统,首先要列写出受控对象的数学表达式,如式(1.1.4)~(1.1.6), (1.1.9), (1.1.10)所示,确定系统结构和参数,这就是本部分内容。在某些场合,系统模型和参数可通过理论计算得到,而在另外一些场合,如对复杂的生产过程,则必须通过试验辨识方法得到。对时变系统,则要求通过在线辨识求得。辨识方法很多,用途很广,这里主要研究能够用于随机控制系统的各种在线、递推辨识方法。

(3) 状态变量的估计(滤波)

像确定性控制系统一样,对随机控制系统,人们感兴趣的是受控对象的状态向量 x , 但不同的是,随机控制系统的状态向量变成了随机过程,得到的量测向量 y 也是随机过程。状态估计(滤波)方法就是由量测向量 y 求出状态向量 x , 计算过程中滤掉噪声信号,得到有用的最优状态向量估计 \hat{x} 。

状态估计与系统辨识相比较,简单地说,两者方程形式相同,不同的是,前者把参数作为已知量求出状态,后者相反,把状态作为已知量求出参数,所以它们的很多算法是相通的。

(4) 最优随机控制

最优随机控制就是利用最优状态估计 \hat{x} , 求出一个状态反馈,实现使给定性能指标为最小的目标。这个状态反馈称为最优控制策略(律)。

完整的随机控制系统包括上述各个部分,但具体的随机控制系统不一定都包括每一

部分,某些部分可能变得简单一些。如有的系统可通过理论计算,而不需用复杂的试验辨识方法来得到系统模型。又如,有的系统的状态向量 x 可直接量测,而不需要量测方程,从而也就不需要状态估计这一部分。

系统辨识和状态估计方法除用于随机控制外,在预报预测、故障诊断、参数监视等方面也得到了广泛的应用,已有大量专著出版。

1.3 本书内容

全书共包括 6 章。

第 1 章绪论,介绍本书的概况。

第 2 章随机过程,介绍随机过程的基本知识,包括随机过程基本概念、工程中常用的随机过程、协方差函数、谱密度和随机分析等内容,这些内容是后续章节要用到的。随机过程作为数学的一个分支,内容很丰富,用途很广泛。本章围绕全书内容介绍了一些随机过程的基本知识,从事其他科技工作的初学者也可阅读。

第 3 章随机系统的数学模型,介绍随机控制常用的三种数学模型,即受控自回归平移平均模型、离散时间随机状态模型和连续时间随机状态模型,着重描述随机状态向量的统计特性。这一章还包括输入量是随机过程的系统分析。

第 4 章最小方差控制,介绍受控自回归平移平均模型的预测、最小方差控制策略和次最优控制策略等内容。

第 5 章最优状态估计,介绍卡尔曼滤波方法,它是一种在线、递推估计方法,适合在随机控制系统中应用。在这一章中,对离散时间随机状态模型,比较严格地证明和推导了卡尔曼滤波方法,并介绍了滤波、预测和平滑等各种不同情况。本章主要讨论模型噪声与量测噪声不相关情况,但也介绍了两者相关的情况。对连续时间随机状态模型,主要介绍用随机微分方程直接推导卡尔曼滤波方程的方法。

第 6 章最优随机控制,介绍用贝尔曼动态规划方法推导出最优随机控制策略(律)的方法。对离散时间随机状态模型,讨论了完全状态信息情况和不完全状态信息情况,并利用一个恒等式讨论了无噪声、存在模型噪声和存在量测噪声等不同情况。对连续时间随机状态模型,主要介绍用随机微分方程直接导出连续时间贝尔曼动态规划方程,进而再求出最优随机控制律的方法。

本书不包括系统辨识和参数估计的内容,也不包括除卡尔曼滤波之外的其他估计方法,需要学习这些内容的读者,请参阅有关书籍和文献。

1.4 随机控制理论的发展概况

随机控制系统的量测量和被控量都是随机过程,因此,随机控制理论的发展是与随机过程理论的发展密切相关的。随机过程理论产生于 20 世纪初期,是为适应物理学、生物学、通信与控制、管理科学等方面的需要而逐步发展起来的。最初在布朗运动、电话信息量和电子管的散粒效应噪声等问题的研究中取得了成果。1931 年,科尔莫戈罗夫(A. N.

Kolmogorov)奠定了随机过程的数学理论基础。1953年,杜布(J. L. Doob)的著作论述了随机过程的数学理论。以后,有关随机过程理论和应用的著作大批出现,并取得丰硕成果。1951年,伊藤(K. Ito)发表了《论随机微分方程》一文。随后,对随机微分方程的研究受到了广泛重视,并渗透到很多领域。随机过程的研究和发展,为随机控制的发展提供了理论基础。

自动控制理论是由于军用和民用工业的需要而提出和发展起来的。1927年已有布莱克(Black)负反馈放大器,1932年奈魁斯特(H. Nyquist)提出的稳定判据是采用频率法的标志,1945年伯德(H. W. Bode)提出的反馈放大器理论奠定了自动控制理论的基础,1948年出现根轨迹法,至此,形成了经典控制理论。在此期间,作为自动控制理论一个分支的随机控制理论也得到了发展。

维纳(N. Wiener)和科尔莫戈罗夫发展起来的滤波和预测理论,使从信号加噪声的观测中抽取有用信号成为可能,这是随机控制理论的一个重要基础,有重大理论价值。但由于维纳-科尔莫戈罗夫理论需要求解一种难于求解的积分方程(维纳-霍甫夫方程),所以未能得到广泛应用。

1956年庞特里亚金(. . .)提出的极大值原理,1957年贝尔曼(R. Bellman)提出的动态规划法,以及1960年卡尔曼(R. E. Kalman)提出的滤波和预测理论使自动控制理论有了重大突破。1960年在美国自动控制第一届联合会上,首次提出了“现代控制理论”这一名称,标志着现代控制理论的正式产生。而这些理论与随机控制直接有关。

数字计算机的广泛应用大大加速了随机控制理论的发展。卡尔曼和布西(R. S. Bucy)1960年提出了求解滤波和预测问题的递推算法,对滤波和预测理论做出了特殊贡献。求解随机控制问题紧密地依赖于动态规划的概念和方法。1961年约瑟夫(P. D. Joseph)和陶(J. T. Tou)提出了分离定理,根据分离定理可把线性随机控制问题分为两部分分别求解,一部分是状态估计器,另一部分是求解最优控制策略。另外,还可以证明,随机最优控制策略与确定性最优控制策略是相同的,这就是确定性等价原理。

自校正自适应控制和模型参考自适应控制的发展,丰富了随机控制内容,使之得到更广泛的应用。

随机控制理论对于具有集中控制器和状态完全能观测的系统已经具有相当完善的理论。几十年来,随机控制理论在很多领域已有广泛和成功的应用。例如,在某些工业批量加工过程中,在阿波罗飞船导航及控制系统中使用了最小方差控制。又如,随机控制理论已用于排队(及流量控制)网络最优控制的模型和分析,水电站的水库管理就是一例。再如,随机控制已用于某些经济问题的模型和分析方法,如非风险性投资和保险有价证券管理的决策方法等。还有,当把先进的随机控制概念与物理仿真模型结合起来,把排队网络理论与扰动分析结合起来,这些用于算法设计的新工具就使得通信及制造网络的优化研究发生了革命。

随机控制理论既有广阔的应用前景,也有很多富有挑战性的研究课题。自动控制理论的传统模型已不能满足要求,应向非线性、随机、分布参数模型等更广的范围扩展。实际系统很少是线性的,高度非线性对象更是如此。因此,这种系统的非线性滤波器和非线性信

号处理器的数字实现,可同时辨识未知参数和系统的自适应随机控制就成为重要而困难的研究课题。在许多实际随机控制问题中,传感器和控制器都受所处位置噪声的干扰,在处理这类多因素或分散式随机控制问题时,还缺少模型、概念体系和分析方法,这也是重要的具有挑战性的研究课题。到目前为止,尚未解决兼顾优化和状态估计的集成控制律问题,从而影响了随机控制的进一步发展,这是又一个重要的研究课题。

习 题

1.1 考虑系统

$$\dot{x}(t) = u(t)$$

它的初始条件为

$$x(0) = 1$$

控制的目的是使性能指标

$$J = \int_0^{\infty} [x^2(t) + u^2(t)] dt$$

达到极小,试证明控制信号

$$u(t) = -e^{-t} \quad (\text{开环程序控制})$$

和控制律

$$u(t) = -x(t) \quad (\text{闭环反馈控制})$$

是最优的。

1.2 考虑实际系统

$$\dot{x}_s(t) = u_s(t)$$

$$x_s(0) = 1$$

假设控制是由模型

$$[\dot{x}_m(t)]^2 = a u_m(t) \quad (a > 1)$$

$$x_m(0) = 1$$

确定的,试相对于开环控制和闭环控制确定性能指标

$$J = \int_0^{\infty} [x^2(t) + u^2(t)] dt$$

的值。

1.3 当实际系统和性能指标由方程

$$\dot{x}(t) = u(t) + v$$

$$x(0) = 1$$

$$J = \int_0^{\infty} [x^2(t) + u^2(t)] dt$$

决定时,试比较开环控制

$$u(t) = -e^{-t}$$

和闭环控制

$$u(t) = -x(t)$$

的性能。其中 v 是一个未知扰动, 特别设 v 是一个未知常数。

1.4 考虑由

$$x(t) = a \cos t$$

描述的系统, 其中 a 是随机变量, 试给出一个准确预测 x 未来值的计算方法。

1.5 考虑由

$$x(k+1) = ax(k) + e(k) \quad (k = k_0, k_0 + 1, \dots)$$

描述的系统, 试证明预测值 $\hat{x}(k+1) = ax(k)$ 在如下意义上是最优的, 即它使由 $E\{x(k+1) - \hat{x}(k+1)\}^2$ 定义的最小二乘预测误差为极小。式中: $x(t)$ 是状态变量, 给定初始条件 $x(k_0) = 1; a \in [-1, 1]; e(k)$ 是一个独立高斯 $N(0, \sigma^2)$ 随机变量序列; $e(k)$ 与 $x(t)$ 是互相独立的。

第 2 章 随 机 过 程

2.1 引 言

随机控制系统是遭受噪声扰动的受控动力学系统,其噪声是一种随机过程,于是整个系统的量测量和被控量(状态向量)都变成了随机过程。因此,在研究随机控制系统之前,学习和了解一些随机过程知识是必不可少的。随机过程理论是研究随机控制系统的数学理论基础。随机过程是数学的一个分支,随机过程理论内容十分丰富,研究日益深入,应用也很广泛。这一章仅介绍一些与随机控制相关的随机过程的基本知识,这些基本知识对其它工程也是有用的。本章内容包括随机过程概念、工程中常用的随机过程、随机过程重要的数字特征协方差函数、用频率表示和分析随机过程的谱密度、作为随机控制系统扰动项的白噪声以及随机分析等。

2.2 随机过程概念

2.2.1 事物分类

自然界中,人们研究的事物按随时间变化的过程,可分为两大类。

一类是确定性(变化)过程。这类事物的变化过程具有必然的变化规律,可用确定性函数描述,即自变量时间与函数值是一一对应的,且保持不变。这是人们研究较多、比较熟悉的一类过程。

另一类是随机过程。这类事物的变化过程没有确定的变化形式,不能用确定性函数描述,即对每一个自变量时间可取不同的函数值。当然,我们所研究的随机过程是具有统计规律的,而不是任意的随机现象。

电子直流放大器输出的零点漂移就是这类随机过程的一个例子,如图 2.2.1 所示。当放大器没有输入,即 $u(t) = 0$ 时,要求输出 $x(t)$ (称为事件或试验结果)也为零。但是,由于放大器内部各种噪声的作用和外部电磁波等各种干扰的影响,使得输出 $x(t)$ 并不为零,造成零点漂移。放大器放大倍数 K 越大,零点漂移越严重。放大器的零点漂移必须限定在一定范围内,零点漂移过大,会使放大器严重失真,甚至不能正常

图 2.2.1 电子直流放大器的零点漂移

工作。

在相同工作条件下,重复量测出 n 条零点漂移曲线 $x_1(t), x_2(t), \dots, x_n(t)$, 如图 2.2.1 所示, 每条曲线都称为样本函数, 各条曲线是各不相同的, 不能用统一的确定性函数表示出来, 所以零点漂移是一个随时间变化的随机过程。由图 2.2.1 进一步看出, 任取固定时刻 t_i 时, $x_1(t_i), x_2(t_i), \dots, x_n(t_i)$ 的取值没有必然的规律, 若把 $x_1(t_i), x_2(t_i), \dots, x_n(t_i)$ 看做是 $x(t)$ 在时刻 t_i 的各种可能的取值, 很明显, $x(t_i)$ 是一个随机变量。

这类的具体例子还很多, 如在每一年的某月份的降雨量, 某日期的湿度, 人踏楼梯脚印的位置等都是随机过程。大量的工业对象以及经济、人口、生态学等领域中, 都存在随机过程问题, 应用随机过程理论加以阐述和研究。

2.2.2 随机过程定义

设 ω 是试验结果, $\Omega = \{\omega\}$ 是样本空间; 时间 $t \in T, T$ 是指标集。下面给出随机过程定义。

定义 2.2.1 随机过程 $x(\omega, t)$ 是一个随机变量族。

随机过程也可定义为: 如果对每一个固定的 $\omega \in \Omega$, 都能确定一个时间 t 的函数 $x(\omega, t)$, 于是对所有 $\omega \in \Omega$, 得到一族时间 t 的函数 $x(\omega, t)$, 则称 $x(\omega, t)$ 为随机过程。随机过程还可定义为: 如果对每一个固定的时间 $t_1 \in T$, 都能得到一个 ω 的随机变量 $x(\omega, t_1)$, 则对所有时间 $t \in T$, 得到一族随机变量 $x(\omega, t)$, 则称 $x(\omega, t)$ 为随机过程。

由定义可知, 一个随机过程 $\{x(\omega, t), \omega \in \Omega, t \in T\}$ 是两个自变量 ω 和 t 的函数。对固定的 $t_1 \in T, x(\omega, t_1) = x(\omega)$ 实际上是一个随机变量, 随机变量是一个特殊的随机过程, 这是概率论研究的内容。对固定的 $\omega \in \Omega, x(\omega, t) = x(t)$ 是一个随机过程, 本书中主要研究这种随机过程。

$x(t)$ 也称为样本函数空间, $x_i(t)$ 表示 $x(t)$ 的某个实现, 也称为样本函数或过程的轨迹。样本函数 $x_i(t)$ 已经实现, 是一个确定性函数。所有样本函数 $x_i(t) (i = 1, 2, \dots, i, \dots)$ 组成样本函数空间

$$x(t) = \{x_1(t), x_2(t), \dots, x_i(t), \dots\} \tag{2.2.1}$$

可把某个样本函数 $x_i(t)$ 看做是样本函数空间 $x(t)$ 的一个元素。

一个随机过程 $x(t)$ 可是标量, 也可以是向量。若 $x(t)$ 和 $y(t)$ 都是随机过程, 而 $i = \overline{1, n}$, 则

$$z(t) = x(t) + iy(t) \tag{2.2.2}$$

称为复随机过程。

2.2.3 随机过程分类

随机过程可按两种方法分类, 下面分别介绍。

1. 按随机变量和指标集类型分类

(1) 连续型随机过程 考虑随机过程 $x(\omega, t)$, 如果随机变量 $x(\omega)$ 是连续变化的, $t \in T$ 也是连续变化的(指标集为 $T = \{t, 0 \leq t < \infty\}$ 或 $T = \{t, -\infty < t < \infty\}$), 则称 $x(\omega, t)$ 为连续型随机过程。如正弦波随机过程就是这类过程。

(2) 离散型随机过程 考虑随机过程 $x(\cdot, t)$, 如果 $x(\cdot)$ 是离散的, 而 t 是连续的, 则称 $x(\cdot, t)$ 为离散型随机过程。如电报信号过程就是这类过程。

(3) 连续型随机序列 考虑随机过程 $x(\cdot, t)$, 如果 $x(\cdot)$ 是连续的, 而 $t \in T$ 是离散变化的($T = \{\dots, -2k, -k, 0, k, 2k, \dots\}$ 或 $T = \{0, k, 2k, \dots\}$, 若 $k=1$, 则 $T = \{\dots, -2, -1, 0, 1, 2, \dots\}$ 或 $T = \{0, 1, 2, \dots\}$), 则称 $x(\cdot, t)$ 为连续型随机序列。对时间连续的随机变量进行采样就是这类过程。

(4) 离散型随机序列 考虑随机过程 $x(\cdot, t)$, 如果 $x(\cdot)$ 是离散的, t 也是离散的, 则称 $x(\cdot, t)$ 为离散型随机序列。伯努利试验、随机游动等都是这类过程。

2. 按随机过程功能分类

可以分为平稳过程、高斯过程、马尔可夫过程(序列)、二阶过程、独立增量过程、维纳过程、白噪声等。这些过程在随机控制中应用较多, 其他过程还很多, 如泊松过程、分枝过程、更新过程、生灭过程等。

2.2.4 随机过程描述

随机过程是一个随机变量族, 简单说, 随机过程就是把随机变量看做是时间的函数。因此, 只要注意到自变量时间, 描述随机变量的方法可完全用来描述随机过程。

像对待随机变量一样, 要全面了解一个随机过程 $x(\cdot, t)$, 最根本的是要知道它的概率分布函数 $F(\cdot, t)$ 或概率密度函数 $f(\cdot, t)$, 它们是 \cdot 和 t 的函数。但为了方便, 在多数场合, 只要知道它的有关数字特征, 如均值函数 $m_x(t)$ (数学期望 $E\{x(t)\}$)、方差函数 $r_x(t)$ 、自协方差函数 $r_x(s, t)$ 等一、二阶矩就够了, 它们都是时间 t 的函数。

考虑随机过程 $x(\cdot, t)$, 它的均值为

$$\begin{aligned} m_x(t) &= E\{x(t)\} \\ &= \int_{-\infty}^{\infty} x dF(\cdot, t) \\ &= \int_{-\infty}^{\infty} x f(\cdot, t) d\cdot \quad (\text{连续情况}) \end{aligned} \tag{2.2.3}$$

$$m_x(k) = \sum_{i=1}^N x_i(k) P(x_i, k) \quad (\text{离散情况}) \tag{2.2.4}$$

式中: $P(x_i, k)$ 是概率函数。它的方差为

$$\begin{aligned} r_x(t) &= \text{var}(x(t)) = \int_{-\infty}^{\infty} [x - m_x(t)]^2 f(\cdot, t) d\cdot \\ &= E\{[x(t) - m_x(t)]^2\} \\ &= E\{x(t)^2\} - m_x(t)^2 \quad (\text{连续情况}) \end{aligned} \tag{2.2.5}$$

$$\begin{aligned} r_x(k) &= \text{var}(x(k)) = \sum_{i=1}^N [x_i(k) - m_x(k)]^2 P(x_i, k) \\ &= E\{[x(k) - m_x(k)]^2\} \\ &= E\{x(k)^2\} - m_x(k)^2 \quad (\text{离散情况}) \end{aligned} \tag{2.2.6}$$

式中: $E\{x(t)^2\}$ 和 $E\{x(k)^2\}$ 称为均方值函数。它的自协方差为

$$r_x(s, t) = \text{cov}[x(s), x(t)]$$

$$\begin{aligned}
&= \int_0^T \int_0^T [x(s) - m_x(s)][x(t) - m_x(t)] d^2 F(\tau_1, \tau_2, s, t) \\
&= \int_0^T \int_0^T [x(s) - m_x(s)][x(t) - m_x(t)] f(\tau_1, \tau_2, s, t) d\tau_1 d\tau_2 \\
&= E\{[x(s) - m_x(s)][x(t) - m_x(t)]\} \\
&= E\{x(s)x(t)\} - m_x(s)m_x(t) \quad (\text{连续情况}) \quad (2.2.7)
\end{aligned}$$

$$\begin{aligned}
r_x(k, l) &= \text{cov}[x(k), x(l)] \\
&= \sum_{n=1}^N \sum_{m=1}^M [x_n(k) - m_x(k)][x_m(l) - m_x(l)] P(x_n, x_m, k, l) \\
&= E\{[x(k) - m_x(k)][x(l) - m_x(l)]\} \\
&= E\{x(k)x(l)\} - m_x(k)m_x(l) \quad (\text{离散情况}) \quad (2.2.8)
\end{aligned}$$

式中: $F(\tau_1, \tau_2, s, t)$, $f(\tau_1, \tau_2, s, t)$ 和 $P(x_n, x_m, k, l)$ 分别为联合分布函数、联合密度函数和联合概率函数。 $E\{x(s)x(t)\}$ 和 $E\{x(k)x(l)\}$ 称为自相关函数, 表示为

$$C_x(s, t) = E\{x(s)x(t)\} \quad (2.2.9)$$

$$C_x(k, l) = E\{x(k)x(l)\} \quad (2.2.10)$$

由式(2.2.7)和(2.2.5), 式(2.2.8)和(2.2.6)可知, 当 $s = t$, $k = l$ 时, 自协方差函数就退化为方差函数。

考虑两个随机过程 $x(s)$ 和 $y(t)$, 它们的互协方差函数 $r_{xy}(s, t)$ 表示为

$$\begin{aligned}
r_{xy}(s, t) &= \text{cov}[x(s), y(t)] \\
&= E\{[x(s) - m_x(s)][y(t) - m_y(t)]\} \\
&= E\{x(s)y(t)\} - m_x(s)m_y(t) \quad (2.2.11)
\end{aligned}$$

式中: $E\{x(s)y(t)\}$ 称为 $x(s)$ 和 $y(t)$ 的互相关函数, 表示为

$$C_{xy}(s, t) = E\{x(s)y(t)\} \quad (2.2.12)$$

若随机过程 $x(s)$ 和 $y(t)$ 分别为 $n \times 1$ 和 $m \times 1$ 列向量, 用 T 表示转置, 则有

$$\begin{aligned}
r_x(s, t) &= E\{[x(s) - m_x(s)][x(t) - m_x(t)]^T\} \\
&= E\{x(s)x(t)^T\} - m_x(s)m_x(t)^T \quad (2.2.13)
\end{aligned}$$

$$\begin{aligned}
r_{xy}(s, t) &= E\{[x(s) - m_x(s)][y(t) - m_y(t)]^T\} \\
&= E\{x(s)y(t)^T\} - m_x(s)m_y(t)^T \\
&= r_{yx}(t, s)^T \quad (2.2.14)
\end{aligned}$$

式中: $r_x(s, t)$ 为 $n \times n$ 矩阵; $r_{xy}(s, t)$ 为 $n \times m$ 矩阵。

若 $x(s)$ 和 $y(t)$ 是两个复随机过程, 则它们均值函数的定义与一般随机过程相同, 但协方差函数定义为

$$r_x(s, t) = E \overline{[x(s) - m_x(s)] [x(t) - m_x(t)]^T} \quad (2.2.15)$$

$$r_{xy}(s, t) = E \overline{[x(s) - m_x(s)] [y(t) - m_y(t)]^T} \quad (2.2.16)$$

考虑两个随机过程 $x(s)$ 和 $y(t)$, 则条件均值为

$$\begin{aligned}
m_{x|y}(s, t) &= E\{x(s)|y(t)\} \\
&= \int_0^T f(\tau, s, t) d\tau \quad (\text{连续情况}) \quad (2.2.17)
\end{aligned}$$

$$m_{x|y}(k, l) = E\{x(k)|y(l)\}$$

$$= \sum_{i=1}^N x_i(k) P(x_i|y, k, l) \quad (\text{离散情况}) \quad (2.2.18)$$

式中: $f(x|y, s, t)$ 和 $P(x_i|y, k, l)$ 分别为条件密度函数和条件概率函数。在 $y(t)$ 条件下的 $x(s)$ 的均值函数 $m_{x|y}(s, t)$ 不仅是 $y(t)$ 的函数, 而且也是 s 的函数。离散情况类似, 不多述。

例 2.2.1 设 $\{e(k), k = \dots, -2, -1, 0, 1, 2, \dots\}$ 是一个独立同分布高斯(正态) $N(0, 1)$ 随机变量序列, 试求 $\{e(k)\}$ 的均值 $E\{e(k)\}$ 和自协方差函数 $r_e(k, l)$ 。

解 按题意, 对不同采样时间 $\dots, -2, -1, 0, 1, 2, \dots, \dots, e(-2), e(-1), e(0), e(1), e(2), \dots$ 相互独立, 但分布相同, 且都是高斯 $N(0, 1)$ 的, 因此, $\{e(k)\}$ 是一个连续型随机序列, 画出它的概率密度函数 $f_e(\cdot, t)$, 如图 2.2.2 所示。 $E\{e(k)\}$ 和 $r_e(k, l)$ 实际上已由 $N(0, 1)$ 给出:

$$E\{e(k)\} = 0 \quad (k = \dots, -2, -1, 0, 1, 2, \dots) \quad (2.2.19)$$

$$r_e(k, l) = \delta_{kl} = \begin{cases} 1 & (k = l) \\ 0 & (k \neq l) \end{cases} \quad (2.2.20)$$

式中: δ_{kl} 称为克罗内克(Kronecker)符号。

图 2.2.2 独立同分布高斯 $N(0, 1)$ 随机变量序列的概率密度函数

对 $r_e(k, l)$, 进行如下计算:

考虑 $\{e(k)\}$ 独立时, 有

$$E\{e(k)e(l)\} = E\{e(k)\}E\{e(l)\} \quad (k \neq l) \quad (2.2.21)$$

则

$$r_e(k, l) = 1 \quad (k = l) \quad (2.2.22)$$

$$r_e(k, l) = E\{e(k)e(l)\} - E\{e(k)\}E\{e(l)\}$$

$$= E\{e(k)\}E\{e(l)\} - E\{e(k)\}E\{e(l)\}$$

$$= 0 \quad (k \neq l) \quad (2.2.23)$$

例 2.2.2 考虑样本空间 $\Omega = \{(x, y)\}$, 各元素 (x, y) 取值为 $(1, 1), (2, 2), (3, 3), (4, 4)$, 其取值概率 $P(x, y)$ 分别为 0.10, 0.35, 0.05, 0.50。求联合分布函数 $F(x, y)$, $F(x), F(y)$, 联合密度函数 $f(x, y)$, 条件概率 $P(x|y = 4)$, 均值 $E\{x\}, E\{xy\}$ 等量。

解 根据题设条件可写出

$$F(x, y) = \sum_{i=1}^4 \sum_{j=1}^4 P_{ij} (x - x_i) (y - y_j)$$

$$\begin{aligned}
 &= 0.10 (x - 1) (y - 1) + 0.35 (x - 2) (y - 2) \\
 &\quad + 0.05 (x - 3) (y - 3) + 0.50 (x - 4) (y - 4)
 \end{aligned}
 \tag{2.2.24}$$

$$\begin{aligned}
 f(x, y) &= \frac{dF(x, y)}{dx dy} \\
 &= 0.10 (x - 1) (y - 1) + 0.35 (x - 2) (y - 2) \\
 &\quad + 0.05 (x - 3) (y - 3) + 0.50 (x - 4) (y - 4)
 \end{aligned}
 \tag{2.2.25}$$

$$\begin{aligned}
 F(x) &= F(x, \quad) \\
 &= 0.10 (x - 1) + 0.35 (x - 2) + 0.05 (x - 3) + 0.50 (x - 4)
 \end{aligned}
 \tag{2.2.26}$$

$$F(y) = 0.10 (y - 1) + 0.35 (y - 2) + 0.05 (y - 3) + 0.50 (y - 4)
 \tag{2.2.27}$$

式中： Φ 和 ϕ 分别表示阶跃函数和 函数。图 2.2.3 和图 2.2.4 画出了 $F(x, y)$ 和 $f(x, y)$ 的图形。

图 2.2.3 例 2.2.2 的联合分布函数

图 2.2.4 例 2.2.2 的联合密度函数

根据 $F(x, y)$, 可导出绝对概率:

$$\begin{aligned}
 P(x_1) &= P(1) = 0.10 & P(x_2) &= P(2) = 0.35 \\
 P(x_3) &= P(3) = 0.05 & P(x_4) &= P(4) = 0.50 \\
 P(y_1) &= P(1) = 0.10 & P(y_2) &= P(2) = 0.35 \\
 P(y_3) &= P(3) = 0.05 & P(y_4) &= P(4) = 0.50
 \end{aligned}
 \tag{2.2.28}$$

并进一步求出条件概率

$$P(x_i | y_k) = \frac{P(x_i, y_k)}{P(y_k)}
 \tag{2.2.29}$$

$$\begin{aligned}
 P(x_i | y_4) &= \frac{P(x_i, 4)}{P(4)} \\
 &= \frac{0.50}{0.50} \\
 &= 1
 \end{aligned}
 \tag{2.2.30}$$

利用上面求出的值, 可得到均值

$$E\{x\} = \sum_{i=1}^4 x_i P_i$$

$$= 2.95 \quad (2.2.31)$$

$$E\{xy\} = \sum_{i=1}^4 \sum_{j=1}^4 x_i y_j P_{ij} = 9.95 \quad (2.2.32)$$

2.3 工程中常用的随机过程

在上节定义了一般的随机过程的基础上, 这节对随机控制和其他工程中常用的随机过程做一简要的介绍。这里我们共介绍六种随机过程, 另一种随机过程白噪声在谱密度之后进行介绍。

2.3.1 平稳随机过程

1. 一阶平稳随机过程

定义 2.3.1 考虑随机过程 $x(t)$, 如果对 $t, t+T$ 的所有 T , 它的一阶密度函数满足

$$f(\tau, t) = f(\tau, t+T) \quad (2.3.1)$$

则称 $x(t)$ 为一阶平稳随机过程。

因为 τ 的选择是任意的, 令 $\tau = -t$, 那么式(2.3.1)变为

$$f(\tau, t) = f(\tau, 0) = f(\tau) \quad (2.3.2)$$

很明显, 一阶平稳过程的密度函数与时间无关。实际上, 一阶平稳随机过程是一个随机变量。

由式(2.3.2)求均值和方差, 得到

$$m_x(t) = \int_{-\infty}^{\infty} \tau f(\tau) d\tau = m_x \quad (2.3.3)$$

$$r_x(t) = \int_{-\infty}^{\infty} (\tau - m_x)^2 f(\tau) d\tau = r_x \quad (2.3.4)$$

一阶平稳随机过程的均值和方差分别为常数 m_x 和 r_x , 而不是时间的函数。

2. 二阶平稳随机过程

定义 2.3.2 考虑随机过程 $x(t)$, 如果对所有 t_1, t_2, t_1+T, t_2+T , 它的二阶联合密度函数满足

$$f(\tau_1, \tau_2, t_1, t_2) = f(\tau_1, \tau_2, t_1+T, t_2+T) \quad (2.3.5)$$

则称 $x(t)$ 为二阶平稳随机过程。

因为 τ_1, τ_2 的选择是任意的, 令 $\tau_1 = -t_1, \tau_2 = t_2 - t_1$, 那么式(2.3.5)变为

$$f(\tau_1, \tau_2, t_1, t_2) = f(\tau_1, \tau_2) \quad (2.3.6)$$

很明显, 二阶平稳随机过程的联合密度函数是两个时间差的函数, 而不是某一个时间的函数。

二阶平稳随机过程的自相关函数用式(2.3.6)计算, 得到

$$C_x(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tau_1 \tau_2 f(\tau_1, \tau_2) d\tau_1 d\tau_2$$

$$= r_x(\tau) \quad (2.3.7)$$

可见, $C_x(t_1, t_2)$ 是两个时间差的函数, 而不是某个时间的函数。自协方差函数也有相同结论。

由于一阶密度函数可由二阶联合密度函数确定, 所以二阶平稳随机过程必然是一阶平稳随机过程。

考虑两个随机过程 $x(s)$ 和 $y(t)$, 如果它们的互相关函数满足

$$C_{xy}(s, t) = C_{xy}(\tau) \quad (2.3.8)$$

式中 $\tau = s - t$, 则称 $x(s)$ 和 $y(t)$ 是联合平稳的。联合平稳过程的互相关函数(包括互协方差函数)也只是两个时间差的函数, 而不是某个时间的函数。

3.k 阶平稳随机过程

定义 2.3.3 考虑随机过程 $x(t)$, 如果对所有 $t_i, t_i + \tau_i$ ($i = 1, 2, \dots, k$) 它的 k 阶联合密度函数满足

$$\begin{aligned} f(\tau_1, \tau_2, \dots, \tau_k, t_1, t_2, \dots, t_k) \\ = f(\tau_1, \tau_2, \dots, \tau_k, t_1 + \tau_1, t_2 + \tau_2, \dots, t_k + \tau_k) \end{aligned} \quad (2.3.9)$$

则称 $x(t)$ 为 k 阶平稳随机过程。

如果 k 为任意阶, 式(2.3.9)都成立, 则该随机过程为严平稳随机过程, 只考虑 $k=2$ 的情况, 则称为宽平稳随机过程。显然, 严平稳必定宽平稳, 而宽平稳却不能保证严平稳。通常, 只考虑宽平稳就够了, 因此, 为简便起见, 若不加说明, 平稳过程就是指宽平稳过程。

例 2.3.1 设 y 是一个随机变量, 试分析随机过程

$$x_1(t) = y \quad (2.3.10)$$

$$x_2(t) = ty \quad (2.3.11)$$

的平稳性。

解 很明显, $x_1(t)$ 是一个随机变量, 其密度函数与时间无关, 是宽平稳过程, 也是严平稳过程。

对 $x_2(t)$ 求均值和相关函数, 得到

$$E\{x_2(t)\} = tE\{y\} \quad (2.3.12)$$

$$E\{x_2(t_1)x_2(t_2)\} = t_1t_2E\{y^2\} \quad (2.3.13)$$

可见, $x_2(t)$ 的均值不是常数, 相关函数也不是时间差的函数, 所以, $x_2(t)$ 不是平稳过程。

4. 平稳随机过程的遍历性

“遍历性”是把统计均值与时间均值联系起来的一个词。用 $A[\cdot]$ 表示一个函数 $[\cdot]$ 的时间均值, 考虑平稳随机过程 $x(t)$, 并用它表示一个样本函数, 则可得到下列两种时间均值:

$$A[x(t)] = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t) dt \quad (2.3.14)$$

$$A[x(t)x(t + \tau)] = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t)x(t + \tau) dt \quad (2.3.15)$$

定义 2.3.4 如果平稳随机过程 $x(t)$ 的统计均值与时间均值相等, 即

$$E\{x(t)\} = A[x(t)] \quad (2.3.16)$$

则称 $x(t)$ 的均值具有遍历性。

定义 2.3.5 如果平稳随机过程 $x(t)$ 的自相关函数与相应时间均值相等, 即

$$C_x(\tau) = A[x(t)x(t+\tau)] \tag{2.3.17}$$

则称 $x(t)$ 的自相关函数具有遍历性。

定义 2.3.6 如果平稳随机过程 $x(t)$ 的均值和自相关函数都具有遍历性, 则称 $x(t)$ 为遍历性过程。

只有平稳随机过程才可能具有遍历性, 但并不是所有平稳随机过程都具有遍历性。下面举一个例子说明这一点。

例 2.3.2 考虑随机过程 $x(t) = y$, y 是高斯随机变量 $N(m_y, \sigma_y^2)$, 试证明 $x(t)$ 是否具有遍历性。

解 例 2.3.1 已说明 $x(t)$ 是平稳随机过程, 但它不具有遍历性, 因为 $x(t)$ 的统计均值和时间均值分别为

$$E\{x(t)\} = m_y \tag{2.3.18}$$

$$A[x(t)] = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T y dt = y \tag{2.3.19}$$

两者是不相等的。

遍历性是平稳随机过程的重要特性, 具有实用价值, 它为人们提供了一种计算平稳随机过程的均值和相关函数的有效方法。测出一条平稳随机过程的样本函数曲线, 若它具有遍历性, 用式(2.3.14)和(2.3.15)求出均值, 就是它的均值和相关函数。

遍历性是平稳随机过程的一种限定形式, 但要证明一个平稳随机过程是否具有遍历性是困难的。实践中, 为了方便, 可凭直觉先假定所研究的平稳随机过程具有遍历性, 求出均值后, 再看这假定是否符合实际情况。

2.3.2 高斯随机过程

1. 定义

定义 2.3.7 考虑随机过程 $x(t)$, 如果它的密度函数的形式为

$$f(x) = (2\pi\sigma_x^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} \frac{(x - m_x)^2}{\sigma_x^2}\right\} \tag{2.3.20}$$

则称 $x(t)$ 为高斯随机过程。式中 $m_x = m_x(t)$, $\sigma_x^2 = \sigma_x^2(t)$, 分别为 $x = x(t)$ 的均值和方差。

高斯随机过程的密度函数或分布函数由其均值和方差唯一地确定。

定义 2.3.8 考虑 $n \times 1$ 向量随机过程 $x(t)$, 如果它的密度函数的形式为

$$f(x) = (2\pi)^{-\frac{n}{2}} (\det R_x)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} [x - m_x]^T R_x^{-1} [x - m_x]\right\} \tag{2.3.21}$$

则称 $x(t)$ 为高斯向量随机过程。式中

$$x(t) = [x_1(t_1) \ x_2(t_2) \ \dots \ x_n(t_n)]^T \tag{2.3.22}$$

m_x 和 R_x 分别为 $x(t)$ 的均值和协方差函数:

$$m_x(t) = [m_1(t_1) \ m_2(t_2) \ \dots \ m_n(t_n)]^T \tag{2.3.23}$$

$$R_x = \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ r_{21} & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \dots & r_{nn} \end{pmatrix} \quad (2.3.24)$$

$$r_{ij} = \text{cov}[x_i(t_i), x_j(t_j)] \quad (i, j = 1, 2, \dots, n) \quad (2.3.25)$$

det 为行列式符号。

定义 2.3.9 考虑 $n \times 1$ 和 $n \times 1$ 向量随机过程 $x(s)$ 和 $y(t)$, 如果它们的联合密度函数的形式为

$$f(x, y) = (2\pi)^{-\frac{n_x+n_y}{2}} (\det R)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \begin{pmatrix} x - m_x \\ y - m_y \end{pmatrix}^T R^{-1} \begin{pmatrix} x - m_x \\ y - m_y \end{pmatrix} \right\} \quad (2.3.26)$$

则称 $x(s)$ 和 $y(t)$ 为联合高斯随机过程。式中 R 为 $x(s)$ 和 $y(t)$ 的联合协方差函数, 表示为

$$R = \begin{pmatrix} R_x & R_{xy} \\ R_{yx} & R_y \end{pmatrix} \quad (2.3.27)$$

考虑多个随机过程, 如果各个随机过程的分布都是高斯的, 可以证明, 这些随机过程和的分布仍是高斯的。

2. 边际密度函数

对二元密度函数 $f(x, y)$, 计算

$$f(x) = \int_{-\infty}^{\infty} f(x, y) dy \quad (2.3.28)$$

$$f(y) = \int_{-\infty}^{\infty} f(x, y) dx \quad (2.3.29)$$

称 $f(x)$ 和 $f(y)$ 为 $f(x, y)$ 的边际密度函数。

对式(2.3.26) 计算边际密度函数, 得到

$$f(x) = (2\pi)^{-\frac{n_x}{2}} (\det R_x)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} [x - m_x]^T R_x^{-1} [x - m_x] \right\} \quad (2.3.30)$$

$$f(y) = (2\pi)^{-\frac{n_y}{2}} (\det R_y)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} [y - m_y]^T R_y^{-1} [y - m_y] \right\} \quad (2.3.31)$$

很明显, 二元高斯密度函数的边际密度函数都是一元高斯密度函数, 这是一个重要结论, 这一结论可推广到多元情况。但要注意, 两个一元高斯边际密度函数不能确定其二元密度函数是高斯的。

3. 条件密度函数

考虑式(2.3.26) 所表示的联合高斯密度函数, 现求条件密度函数 $f(x|y)$, 这是后面求解估计问题的依据。

求解这个问题的步骤是先把式(2.3.26) 展开, 再利用条件密度函数公式

$$f(x|y) = \frac{f(x, y)}{f(y)} \quad (2.3.32)$$

求出结果。

构造矩阵

$$T = \begin{bmatrix} I & -R_{xy}R_y^{-1} \\ 0 & I \end{bmatrix} \quad (2.3.33)$$

$$R_1 = TR = \begin{bmatrix} R_x - R_{xy}R_y^{-1}R_{yx} & 0 \\ R_{yx} & R_y \end{bmatrix} \quad (2.3.34)$$

利用一般矩阵运算公式

$$\begin{bmatrix} A & 0 \\ C & B \end{bmatrix} = \begin{bmatrix} A^{-1} & 0 \\ -B^{-1}CA^{-1} & B^{-1} \end{bmatrix} \quad (2.3.35)$$

$$\begin{bmatrix} A & D \\ 0 & B \end{bmatrix} = \begin{bmatrix} A^{-1} & -A^{-1}DB^{-1} \\ 0 & B^{-1} \end{bmatrix} \quad (2.3.36)$$

求出逆矩阵 T^{-1} 和 R_1^{-1} :

$$T^{-1} = \begin{bmatrix} I & R_{xy}R_y^{-1} \\ 0 & I \end{bmatrix} \quad (2.3.37)$$

$$R_1^{-1} = \begin{bmatrix} (R_x - R_{xy}R_y^{-1}R_{yx})^{-1} & 0 \\ -R_y^{-1}R_{yx}(R_x - R_{xy}R_y^{-1}R_{yx})^{-1} & R_y^{-1} \end{bmatrix} \quad (2.3.38)$$

利用上列各式计算 $\det R$ 和 R^{-1} , 得到

$$\det R = \det R_y \det (R_x - R_{xy}R_y^{-1}R_{yx}) \quad (2.3.39)$$

$$R^{-1} = \begin{bmatrix} (R_x - R_{xy}R_y^{-1}R_{yx})^{-1} & -(R_x - R_{xy}R_y^{-1}R_{yx})^{-1}R_{xy}R_y^{-1} \\ -R_y^{-1}R_{yx}(R_x - R_{xy}R_y^{-1}R_{yx})^{-1} & R_x^{-1}R_{yx}(R_x - R_{xy}R_y^{-1}R_{yx})^{-1}R_xR_y^{-1} + R_y^{-1} \end{bmatrix} \quad (2.3.40)$$

把式(2.3.39)和(2.3.40)代入式(2.3.26), 得到

$$f(x, y) = (2\pi)^{-\frac{n_x}{2}} (\det R_{x@y})^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} [x - m_{x@y}]^T R_{x@y}^{-1} [x - m_{x@y}] \right\} \varphi(y) \quad (2.3.41)$$

其中 $f(y)$ 用式(2.3.31)表示, $m_{x@y}$ 和 $R_{x@y}$ 分别为

$$m_{x@y} = m_x + R_{xy}R_y^{-1}(y - m_y) \quad (2.3.42)$$

$$R_{x@y} = R_x - R_{xy}R_y^{-1}R_{yx} \quad (2.3.43)$$

把式(2.3.41)代入式(2.3.32), 得到结果为

$$f(x@y) = (2\pi)^{-\frac{n_x}{2}} (\det R_{x@y})^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} [x - m_{x@y}]^T R_{x@y}^{-1} [x - m_{x@y}] \right\} \quad (2.3.44)$$

由式(2.3.44)可知, 条件密度函数 $f(x@y)$ 也是高斯的, 并且条件均值 $m_{x@y}$ 和条件协方差 $R_{x@y}$ 分别由式(2.3.42)和(2.3.43)表示。

4. 平稳高斯随机过程

如果 $x(t)$ 是平稳高斯随机过程, 那么是指 $x(t)$ 同时具有平稳过程和高斯过程的所有特性, 其均值是常数, 协方差函数是时间差的函数, 即

$$\text{cov}[x(s), x(t)] = R_x(s, t) \quad (2.3.45)$$

2.3.3 马尔可夫过程

定义 2.3.10 考虑随机过程 $x(t)$, 并设 $t_0 < t_1 < \dots < t_i < \dots < t_k < t_{k+1}$, 如果它的条件概率密度函数满足

$$f[x(t_{k+1}) | x(t_k), x(t_{k-1}), \dots, x(t_0)] = f[x(t_{k+1}) | x(t_k)] \tag{2.3.46}$$

则称 $x(t)$ 为马尔可夫过程。

通常, 当 t_i 取连续值时, $x(t)$ 称为马尔可夫过程; 当 t_i 取离散值时, $x(t)$ 称为马尔可夫序列(链)。

定义式(2.3.46)中, $f[x(t_{k+1}) | x(t_k)]$ 称为转移概率密度, 它是马尔可夫过程(序列)的重要变量。当转移概率密度满足

$$f[x(t_{k+1}) | x(t_k)] = f[x(t_{k+1})] \tag{2.3.47}$$

则称 $x(t)$ 为纯随机过程(序列)。

定义式(2.3.46)表明, $x(t_{k+1})$ 的概率密度函数只取决于对 $x(t_k)$ 的了解, 而与 $x(t_{k-1}), x(t_{k-2}), \dots, x(t_0)$ 无关。所谓对 $x(t_k)$ 的了解, 可以是确定性的(知道 $x(k)$ 的确切值), 也可以是概率意义上的(知道 $f[x(t_k)]$)。马尔可夫性也称为无后效性, 意味着只有“现在” $x(t_k)$ 才对“将来” $x(t_{k+1})$ 有影响, 而“过去” $(x(t_{k-1}), x(t_{k-2}), \dots, x(t_0))$ 对“将来”没有影响, 或更确切地说, “过去”只是通过“现在”来影响“将来”。

考虑马尔可夫过程(序列) $x(t)$, 根据条件密度函数公式和式(2.3.46), 可把联合密度函数写为

$$\begin{aligned} f[x(t_{k+1}), x(t_k), \dots, x(t_0)] &= f[x(t_{k+1}) | x(t_k), x(t_{k-1}), \dots, x(t_0)] f[x(t_k), x(t_{k-1}), \dots, x(t_0)] \\ &= \prod_{i=0}^k f[x(t_{i+1}) | x(t_i)] f(x(t_0)) \end{aligned} \tag{2.3.48}$$

由上式可知, 马尔可夫过程(序列)的联合密度函数等于各个转移概率密度和初始概率密度的乘积。

马尔可夫性看起来似乎有很大局限性, 实际上具有很强的普遍性, 本书中所研究的大多数随机过程都具有马尔可夫性。

例 2.3.3 考虑离散时间随机状态方程

$$x(k+1) = x(k) + v(k) \tag{2.3.49}$$

试问在下列两种情况下 $x(k)$ 是否为马尔可夫序列: (1) 状态 $x(k)$ 与噪声 $v(1) \dots v(k)$ 互相独立; (2) 初始状态 $x(k_0)$ 与初始噪声 $v(k_0)$ 相关。

解 改写给定方程式(2.3.49), $x(k)$ 用 $x(k_0)$ 表示

$$\begin{aligned} x(k) &= x(k-1) + v(k-1) \\ &= [x(k-2) + v(k-2)] + v(k-1) \\ &= \dots = x(k_0) + \sum_{i=0}^{k-k_0-1} v(k-i-1) \end{aligned} \tag{2.3.50}$$

上式右边第二项包括 $v(k-1), v(k-2), \dots, v(k_0)$ 等各个噪声项。下面分别分析两种情况。

(1) 按给定条件, 式(2.3.50)中右边第一项与第二项互相独立, 因此, $x(k) (k = k_0,$

k_0+1, \dots) 都可用 $x(k_0)$ 表示。换句话说, $x(k)$ 包含了 $x(k-1), x(k-2), \dots, x(k_0)$ 的所有信息, 满足马尔可夫性定义式(2.3.46), 所以, $x(k)$ 是马尔可夫序列。

(2) $x(k_0)$ 与 $v(k_0)$ 相关时, $v(k_0)$ 取值影响 $x(k_0)$ 的值, 也影响 $x(k_0+1), \dots, x(k)$ 的值。 $x(k)$ 不再包含 $x(k-1), x(k-2), \dots, x(k_0)$ 的所有信息, 不能使式(2.3.46)成立, 所以, $x(k)$ 不是马尔可夫序列。

随机控制研究的问题多属于第(1)种情况, 它们是马尔可夫序列(过程)。

2.3.4 二阶随机过程

定义 2.3.11 考虑随机过程 $x(t)$, 如果对每个 $t \in T$, 它的均值和协方差函数都存在, 则称 $x(t)$ 为二阶随机过程。

宽平稳随机过程、高斯随机过程等都是二阶随机过程。

随机过程的二阶特性可用于研究随机过程收敛性、连续性、随机微分和随机积分等问题。

2.3.5 独立增量过程

1. 有关定义

定义 2.3.12 考虑随机过程 $x(t)$, 如果对 $t_i \in T (i=1, 2, \dots, k, t_1 < t_2 < \dots < t_{k-1} < t_k)$, 恒有增量 $x(t_k) - x(t_{k-1}), x(t_{k-1}) - x(t_{k-2}), \dots, x(t_2) - x(t_1), x(t_1)$ 互相独立, 则称 $x(t)$ 为独立增量过程。

独立增量过程 $x(t)$ 的增量密度函数满足

$$f[x(t_k) - x(t_{k-1}), x(t_{k-1}) - x(t_{k-2}), \dots, x(t_2) - x(t_1), x(t_1)] = \prod_{i=1}^k f[x(t_i) - x(t_{i-1})] \tag{2.3.51}$$

$$x(t_0) = 0 \tag{2.3.52}$$

即独立增量过程的增量有限维联合密度函数由所有增量密度函数的乘积唯一确定, 其中把 $x(t_1) - x(t_0)$ 也看做是一个增量。

像一般随机过程一样, 若随机过程 $x(t)$ 的增量协方差函数, 对 $t_{j-1} < t_j, t_{i-1} < t_i$ 满足

$$\text{cov}[x(t_i) - x(t_{i-1}), x(t_j) - x(t_{j-1})] = 0 \tag{2.3.53}$$

或增量相关函数满足

$$\begin{aligned} & E\{[x(t_i) - x(t_{i-1})][x(t_j) - x(t_{j-1})]\} \\ &= E\{x(t_i) - x(t_{i-1})\}E\{x(t_j) - x(t_{j-1})\} \end{aligned} \tag{2.3.54}$$

则称 $x(t)$ 为不相关增量过程。

独立增量过程必为不相关增量过程, 反之则不一定, 即不相关增量过程不一定为独立增量过程。

考虑独立增量过程 $x(t)$, 如果它的增量 $x(s) - x(t)$ 的分布是高斯的, 则称 $x(t)$ 为高斯独立增量过程。对高斯过程来说, 高斯不相关增量过程等价于高斯独立增量过程。

定义 2.3.13 考虑随机过程 $x(t)$, 如果对所有 $s, t, s+ \Delta s, t+ \Delta t \in T, \Delta s, \Delta t$ 为任意值, 恒有增量 $x(s) - x(t)$ 的密度函数与增量 $x(s+ \Delta s) - x(t+ \Delta t)$ 的密度函数相等, 则称 $x(t)$ 为平稳

增量过程。

与一般平稳随机过程类似, 平稳增量过程 $x(t)$ 的增量密度函数只是时间差 $s-t$ 的函数, 即

$$f[x(s) - x(t), s, t] = f[x(s) - x(t), s - t] \tag{2.3.55}$$

进一步可导出, 平稳增量过程 $x(t)$ 的方差也只是时间差的函数。

2. 过程协方差函数

考虑独立增量过程 $x(t)$, 为计算方便并不失一般性, 令 $s > t$, 计算过程协方差函数 $r_x(s, t)$:

$$\begin{aligned} r_x(s, t) &= \text{cov}[x(s), x(t)] \\ &= \text{cov}[x(s) - x(t) + x(t), x(t)] \\ &= \text{cov}[x(s) - x(t), x(t)] + \text{cov}[x(t), x(t)] \\ &= r_x(t, t) \end{aligned} \tag{2.3.56}$$

在上式第三个等式右边, 根据独立增量过程定义, 第一项等于零, 就得到第四个等式。

若 $t > s$, 可导出类似结果。总之, 得到

$$\begin{aligned} r_x(s, t) &= \begin{matrix} r_x(t) & (s < t) \\ r_x(s) & (s > t) \end{matrix} \\ &= \text{cov}[x(\min(s, t)), x(\min(s, t))] \end{aligned} \tag{2.3.57}$$

由式(2.3.57)可知, 过程协方差函数并非时间差的函数, 因此, 独立增量过程 $x(t)$ 不是平稳过程。

3. 平稳独立增量方差

考虑随机过程 $x(t)$, 它既具有独立增量, 又具有平稳增量, 计算这种增量的方差, 分两步进行。

首先, 计算独立增量

$$dx = x(s) - x(t) \quad (s > t) \tag{2.3.58}$$

的方差 dV , 得到

$$\begin{aligned} dV &= \text{cov}[dx, dx] \\ &= \text{cov}[x(s) - x(t), x(s) - x(t)] \\ &= \text{cov}[x(s) - x(t), x(s)] - \text{cov}[x(s) - x(t), x(t)] \\ &= \text{cov}[x(s), x(s)] - \text{cov}[x(t), x(s)] \\ &= r_x(s) - r_x(t) \end{aligned} \tag{2.3.59}$$

由上式可知, 独立增量过程的增量方差, 是两个时间的过程方差的差。

其次, 考虑平稳增量的增量方差。按定义式(2.3.55), 平稳增量的增量方差只是时间差的函数, 可写出

$$dV = r(s - t) \tag{2.3.60}$$

由式(2.3.59)等于式(2.3.60), 得到

$$r(s - t) = r_x(s) - r_x(t) \tag{2.3.61}$$

下面进一步推导。把上等式两边不同形式统一起来, 并求解。为此, 令 $t_0 < t < s$, 利用式(2.3.61), 写出下列两个等式:

$$r(s+t) = r_x(s+t+t_0) - r_x(t_0) \quad (2.3.62)$$

$$r(s) = r_x(s+t_0) - r_x(t_0) \quad (2.3.63)$$

由式(2.3.62)减去式(2.3.63), 并且再利用(2.3.61), 得到

$$r(s+t) = r(s) + r(t) \quad (2.3.64)$$

方程式(2.3.64)的解是时间的线性函数, 即

$$r(t) = \sigma^2 t \quad (2.3.65)$$

式中 σ^2 是正实数, 这里称为方差参数。

总之, 平稳独立增量过程的增量方差是时间的线性函数。

2.3.6 维纳过程

维纳过程, 也称为布朗运动, 在随机过程理论的发展中, 曾占有极重要的地位, 很多影响随机控制系统的扰动, 都可由维纳过程来生成。

布朗运动是植物学家罗伯特·布朗(Robert Brown)在 1827 年首先发现的。他发现, 浸入均匀液体中的微小粒子的扩散运动是极不规则的。1905 年, 著名物理学家爱因斯坦(A. Einstein)提出, 布朗运动是由微小粒子与液体分子之间的碰撞引起的。1923 年, 维纳(N. Wiener)对布朗运动给出了严格的数学分析。

直观上分析, 布朗运动是由浸在均匀液体中的微小粒子受到液体分子不断的不规则的碰撞引起的。它在整个空间运动, 这里只考虑一维空间, 设微小粒子在时刻 t 时的位置为 $w(t)$, 于是可自然地导出下列结论: 把微小粒子置入位置作为坐标原点, $w(0) = 0$, 这是一个确定量; 因为 $w(t)$ 是微小粒子与液体分子之间无数次碰撞造成的结果, 所以 $\{w(t), t \geq 0\}$ 是一个随机过程, 通常, 区间 $[0, t]$ 远远大于两次相继碰撞之间的时间间隔, $w(t)$ 是各次碰撞产生的微小位移的总和, 根据中心极限定理, 它是一个高斯随机过程; 假设液体中的分子分布是均匀的, 那么微小粒子由坐标原点 $w(0) = 0$ 向不同方向移动的概率, 以及移动远近的概率都是相同的, 因此可以认定 $E\{w(t)\} = 0$; 有理由假设, 在非交叠的时间间隔中, 微小粒子运动的统计特性是独立的, 即 $w(t)$ 具有独立增量, 也有理由进一步假设, 在不同起点但相同间隔 $[s, s+\Delta t]$ 和 $[t, t+\Delta t]$ 中, $w(t)$ 的统计特性是相同的, 即具有平稳增量, 总之 $w(t)$ 是一个平稳独立增量过程。

把上述过程推广到一般情况, 得到维纳过程定义。

定义 2.3.14 考虑随机过程 $w(t)$, 如果满足

- (1) 初值为零, 即 $w(0) = 0$;
- (2) $w(t)$ 是高斯随机过程;
- (3) $w(t)$ 的均值为零, 即 $E\{w(t)\} = 0$;
- (4) $w(t)$ 具有平稳独立增量;

则称 $w(t)$ 为维纳过程或布朗运动。

由维纳过程定义可给出维纳过程的统计特性。维纳过程 $w(t)$ 的初值和均值直接由定义给出

$$w(0) = 0 \quad (2.3.66)$$

$$E\{w(t)\} = 0 \quad (2.3.67)$$

由于 $w(t)$ 具有平稳独立增量, 它的方差 $r(t)$ 由式(2.3.65)确定, 即

$$r(t) = \sigma^2 t \tag{2.3.68}$$

式中: σ^2 为方差参数。 $w(t)$ 的协方差函数由式(2.3.57)和(2.3.68)确定:

$$r(s, t) = \sigma^2 \min(s, t) \tag{2.3.69}$$

由式(2.3.68)和(2.3.69)可知, $w(t)$ 的方差不是常数, 协方差也不是时间差 $(s - t)$ 的函数, 因此, 维纳过程 $w(t)$ 不是平稳随机过程。

按定义, 维纳过程 $w(t)$ 的分布是高斯的, 已知均值和方差分别由式(2.3.67)和(2.3.68)给定, 容易写出它的密度函数:

$$f(w) = (2\sigma^2 t)^{-\frac{1}{2}} \exp - \frac{1}{2\sigma^2 t} w^2 \tag{2.3.70}$$

由上式可知, 维纳过程 $w(t)$ 是时间 t 的函数, 时间 t 小时, 方差较大, 随时间 t 增长, 方差变小。

维纳过程的定义中, 已设初值为零。实际上, 初值可不为零, 并容易得到相应结果。

例 2.3.4 在应用中, 常利用维纳过程 $w(t)$ 的增量

$$dw(t) = w(t + dt) - w(t) \tag{2.3.71}$$

试求 $dw(t)$ 的方差 $r[dw(t)]$ 。

解 利用式(2.3.59)和(2.3.68), 容易得到

$$\begin{aligned} r[dw(t)] &= r[w(t + dt) - w(t), w(t + dt) - w(t)] \\ &= r[w(t + dt)] - r[w(t)] \\ &= \sigma^2(t + dt) - \sigma^2 t \\ &= \sigma^2 dt \end{aligned} \tag{2.3.72}$$

2.4 协方差函数

协方差函数是随机过程的一个重要数字特征。作为准备知识, 本节先介绍对称阵及其正定性、施瓦茨不等式, 然后再主要讨论协方差函数的几个重要性质。

2.4.1 对称阵及其正定性

1. 定义

定义 2.4.1 考虑一个 n 阶实数方阵 A , 如果满足 $A^T = A$, 即 $a_{ij} = a_{ji} (i, j = 1, 2, \dots, n)$, 则称 A 为 n 阶对称矩阵。

2. 对称阵性质

- (1) 若 A 为对称阵, λ 为任一实数, 则 λA 也是对称阵。
- (2) 若 A 为对称阵, 且其逆 A^{-1} 存在, 则 A^{-1} 也必为对称阵。
- (3) 若 A 和 B 同为对称阵, 则 $A + B$ 也是对称阵, 但 AB 不一定是对称阵。

3. 对称阵的正定性

设 A 为 n 阶对称阵, x 为任意 $n \times 1$ 列向量, 如果由 A 所确定的二次型

$$x^T A x = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \quad (2.4.1)$$

- (1) 不管 x 取何值, 恒满足 $x^T A x \geq 0$, 则称 A 为非负定(半正定)阵, 并记作 $A \geq 0$;
- (2) 不管 x 取何值, 恒满足 $x^T A x \leq 0$, 则称 A 为非正定(半负定)阵, 并记作 $A \leq 0$;
- (3) 对 $x \neq 0$, 恒有 $x^T A x > 0$; 对 $x = 0$, 有 $x^T A x = 0$, 则称 A 为正定阵, 并记作 $A > 0$;
- (4) 对 $x \neq 0$, 恒有 $x^T A x < 0$; 对 $x = 0$, 有 $x^T A x = 0$, 则称 A 为负定阵, 并记作 $A < 0$;
- (5) 既不是半正定, 又不是半负定, 则称 A 为不定。

2.4.2 施瓦茨(Schwarz)不等式

对任意随机过程 $x(s)$ 和 $y(t)$, 必满足不等式

$$[E\{xy\}]^2 \leq E\{x^2\}E\{y^2\} \quad (2.4.2)$$

上式称为施瓦茨不等式。

证明 任选实常数 α , 利用随机过程 $x(s)$ 和 $y(t)$, 构造不等式, 再取均值, 得到

$$[\alpha y - x]^2 \geq 0 \quad (2.4.3)$$

$$E\{[\alpha y - x]^2\} \geq 0 \quad (2.4.4)$$

把上式均值符号内的二次方展开, 对 α 进行配二次方运算, 并考虑到 $\alpha^2 = x^2$, 得到

$$E\{y^2\} + 2E\{\alpha y x\} + \frac{[E\{\alpha y x\}]^2}{E\{y^2\}} - \frac{[E\{\alpha y x\}]^2}{E\{y^2\}} + E\{x^2\} \geq 0 \quad (2.4.5)$$

$$E\{y^2\} + \frac{E\{\alpha y x\}^2}{E\{y^2\}} + E\{x^2\} - \frac{[E\{\alpha y x\}]^2}{E\{y^2\}} \geq 0 \quad (2.4.6)$$

选

$$\alpha = - \frac{E\{xy\}}{E\{y^2\}} \quad (2.4.7)$$

并把它代入式(2.4.6), 就得到式(2.4.2)。

2.4.3 协方差函数的几个重要性质

考虑实二阶随机过程 $\{x(t), t \in T\}$, 它的协方差函数为

$$r_x(s, t) = E\{[x(s) - m_x(s)][x(t) - m_x(t)]\} \quad (2.4.8)$$

则它具有下列四个性质:

- (1) 容易证明, $r_x(s, t)$ 必有

$$r_x(s, t) = r_x(t, s) \quad (2.4.9)$$

当 $x(t)$ 为向量随机过程时, 则有

$$r_x(s, t) = r_x^T(t, s) \quad (2.4.10)$$

- (2) 取时间分点 $t_i \in T (i = 1, 2, \dots, n)$ n 为任意正整数, 构造协方差阵 R_x :

$$R_x = [r_x(t_i, t_j)] \quad (i, j = 1, 2, \dots, n) \quad (2.4.11)$$

$$r_x(t_i, t_j) = E\{[x(t_i) - m_x(t_i)][x(t_j) - m_x(t_j)]\} \quad (2.4.12)$$

则协方差阵 R_x 必是非负定的。

证明 任取 $n \times 1$ 非零列向量 $z = [z_1, z_2, \dots, z_n]^T$, 构造不等式, 并进行运算, 得到

$$E\left\{\left[\sum_{i=1}^n z_i(x(t_i) - m(t_i))\right]^2\right\} = 0 \quad (2.4.13)$$

$$\sum_{i=1}^n z_i z_j r_x(t_i - t_j) = 0 \quad (2.4.14)$$

$$z^T R_x z = 0 \quad (2.4.15)$$

式(2.4.13)左边是二次方运算,必为非负量,它就是 R_x 的二次型,因而式(2.4.15)成立,即得到 $R_x = 0$,协方差阵 R_x 是非负定的。

(3) 协方差 $r_x(s, t)$ 必满足不等式

$$|r_x(s, t)|^2 \leq r_x(s, s)r_x(t, t) \quad (2.4.16)$$

证明 把 $[x(s) - m_x(s)]$ 和 $[x(t) - m_x(t)]$ 分别看做是施瓦茨不等式(2.4.2)中的 x 和 y , 并对应式(2.4.2), 立即得到

$$\begin{aligned} r_x(s, s)r_x(t, t) &= E\{[x(s) - m_x(s)][x(t) - m_x(t)]\}^2 \\ &= |E\{[x(s) - m_x(s)][x(t) - m_x(t)]\}|^2 \\ &= |r_x(s, t)|^2 \end{aligned} \quad (2.4.17)$$

(4) 如果 $r_x(s, t)$ 在 $s = t$ 连续, 则对所有 t 必定连续。

证明 设 s 具有增量 h , t 具有增量 k , 只要证明

$$|r_x(s + h, t + k) - r_x(s, t)| \xrightarrow{h, k \rightarrow 0} 0 \quad (2.4.18)$$

则说明 $r_x(s, t)$ 是连续的。

对式(2.4.18)左边进行运算, 并运用施瓦茨不等式, 得到

$$\begin{aligned} &|r_x(s + h, t + k) - r_x(s, t)| \\ &= |r_x(s + h, t + k) - r_x(s + h, t) + r_x(s + h, t) - r_x(s, t)| \\ &= |r_x(s + h, t + k) - r_x(s + h, t) + r_x(s + h, t) - r_x(s, t)| \\ &= \{r_x(s + h, s + h)r_x(x(t + k) - x(t), x(t + k) - x(t))\}^{\frac{1}{2}} \\ &\quad + \{r_x(x(s + h) - x(s), x(s + h) - x(s))r_x(t, t)\}^{\frac{1}{2}} \\ &= \{r_x(s + h, s + h)[r_x(t + k, t + k) - 2r_x(t + k, t) - r_x(t, t)]\}^{\frac{1}{2}} \\ &\quad + \{[r_x(s + h, s + h) - 2r_x(s + h, s) + r_x(s, s)]r_x(t, t)\}^{\frac{1}{2}} \end{aligned} \quad (2.4.19)$$

按给定条件, $r_x(t, t)$ 是连续的, 即有

$$|r_x(t + k, t) - r_x(t, t)| \xrightarrow{k \rightarrow 0} 0 \quad (2.4.20)$$

$$|r_x(t + k, t + k) - r_x(t, t)| \xrightarrow{k \rightarrow 0} 0 \quad (2.4.21)$$

由式(2.4.20)和(2.4.21)可知, 当 $k \rightarrow 0$ 时, 式(2.4.19)右边第一项趋于零。同理, 当 $h \rightarrow 0$ 时, 式(2.4.19)右边第二项也趋于零。这就证明了式(2.4.18)成立, $r_x(s, t)$ 连续。

$r_x(s, t)$ 是二元函数, $r_x(t, t)$ 是一元函数, 这个性质说明, $r_x(s, t)$ 是否连续, 可由 $r_x(t, t)$ 是否连续决定, 而 $r_x(t, t)$ 是否连续是比较容易证明的。

2.4.4 平稳随机过程情况

工程中遇到的随机过程多为平稳随机过程, 因此把一般随机过程协方差函数的性质

推广到平稳随机过程是必要的。平稳随机过程协方差函数具有下列性质:

(1) 令 $\tau = s - t$, 则协方差函数是 τ 的函数, 并且对 τ 是对称的, 即

$$r_x(\tau) = r_x(-\tau) \tag{2.4.22}$$

(2) 协方差函数阵 R_x 是非负定的, 即 $R_x \geq 0$, R_x 的二次型表示为

$$Z^T R_x Z = \sum_{i=1}^n \sum_{j=1}^n z_i z_j r_x(t_i - t_j) \geq 0 \tag{2.4.23}$$

(3) 平稳随机过程 $r_x(\tau)$ 在 $\tau = 0$ 时取最大值, 即

$$|r_x(\tau)| \leq |r_x(0)| \tag{2.4.24}$$

(4) 如果对 $\tau = 0$, $r_x(\tau)$ 连续, 那么对所有 τ , $r_x(\tau)$ 都连续。

例 2.4.1 讨论下列函数是否为平稳随机过程的协方差函数。

$$r_1(\tau) = \text{常数} \tag{2.4.25}$$

$$r_2(\tau) = \cos \tau \tag{2.4.26}$$

$$r_3(\tau) = \begin{cases} 1 & (|\tau| \leq 1) \\ 0 & (|\tau| > 1) \end{cases} \tag{2.4.27}$$

$$r_4(\tau) = \begin{cases} 1 - |\tau| & (|\tau| \leq 1) \\ 0 & (|\tau| > 1) \end{cases} \tag{2.4.28}$$

解 $r_1(\tau)$, $r_2(\tau)$, $r_4(\tau)$ 满足平稳随机过程各条性质, 因此它们是平稳随机过程的协方差函数。

$r_3(\tau)$ 在 $\tau = 1$ 处不连续, 不满足性质(4), 因此它不是平稳随机过程的协方差函数。

例 2.4.2 已知随机状态方程

$$x(k+1) = 0.6x(k) + 2u(k) + v(k) \tag{2.4.29}$$

$$y(k) = 3x(k) + e(k) \tag{2.4.30}$$

式中 $x(k_0) \sim N(5, 10)$, $v(k) \sim N(0, 6)$, $e(k) \sim N(0, 4)$ 互相独立, $u(k)$ 为控制量。试求协方差函数 $R[y(k_0+1), y(k_0)]$ 。

解 首先列出 $R[y(k_0+1), y(k_0)]$ 的定义式:

$$R[y(k_0+1), y(k_0)] = E\{[y(k_0+1) - m_y(k_0+1)][y(k_0) - m_y(k_0)]\} \tag{2.4.31}$$

再写出各项表达式, 由给定方程导出

$$m_y(k_0) = 3m_x(k_0) \tag{2.4.32}$$

$$y(k_0) - m_y(k_0) = 3[x(k_0) - m_x(k_0)] + e(k_0) \tag{2.4.33}$$

$$y(k_0+1) = 1.8x(k_0) + 6u(k_0) + 3v(k_0) + e(k_0+1) \tag{2.4.34}$$

$$m_y(k_0+1) = 1.8m_x(k_0) + 6u(k_0) \tag{2.4.35}$$

$$y(k_0+1) - m_y(k_0+1) = 1.8[x(k_0) - m_x(k_0)] + 3v(k_0) + e(k_0+1) \tag{2.4.36}$$

最后把式(2.4.36)和(2.4.33)代入式(2.4.31), 并代入给定数据, 得到结果:

$$\begin{aligned} R[y(k_0+1), y(k_0)] &= E\{1.8[x(k_0) - m_x(k_0)] \times 3[x(k_0) - m_x(k_0)]\} \\ &= 1.8 \times 3 \times R_x(k_0) \\ &= 54 \end{aligned} \tag{2.4.37}$$

2.5 谱 密 度

像研究确定性函数一样,既可在时域上,也可在频域上表示和分析随机过程。前几节涉及的概率分布、密度、均值、相关函数和协方差函数等都是在时域上描述随机过程,这一节介绍功率谱密度,它是在频域上表示随机过程。首先给出确定性函数功率谱密度的定义,然后把它推广到随机过程,建立起相关函数与功率谱密度之间的关系,再简单介绍一些功率谱密度的性质和互谱密度等概念。

2.5.1 确定性函数的谱密度

考虑确定性函数 $x(t)$, 设它具有有限个间断点和有限个极值点, 极值点是有界的, 且绝对可积, 则必存在傅里叶变换对

$$F(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t} dt \quad (2.5.1)$$

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)e^{j\omega t} d\omega \quad (2.5.2)$$

式(2.5.1)称为正变换, 式(2.5.2)称为逆变换, 两式关系可简单地表示为

$$x(t) \longleftrightarrow F(\omega) \quad (2.5.3)$$

式中 $F(\omega)$ 称为 $x(t)$ 的频谱, 它是角频率的函数。 $F(\omega)$ 的一般形式为复函数, 当 $x(t)$ 为实函数时, 即 $x(t) = \overline{x(t)}$ 时, 容易由式(2.5.1)得到

$$F(\omega) = \overline{F(-\omega)} \quad \overline{F(\omega)} = F(-\omega) \quad (2.5.4)$$

在 $x(t)$ 和 $F(\omega)$ 之间存在重要关系式, 由如下定理表示:

定理 2.5.1 帕塞瓦尔(Parseval)定理 当 $x(t)$ 是实函数时, 在 $x(t)$ 和 $F(\omega)$ 之间存在关系式:

$$\int_{-\infty}^{\infty} x^2(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |F(\omega)|^2 d\omega \quad (2.5.5)$$

证明 因为 $F(\omega)$ 是复函数, 所以有

$$|F(\omega)|^2 = F(-\omega)F(\omega) \quad (2.5.6)$$

把式(2.5.6)代入式(2.5.5), 并利用式(2.5.1)和式(2.5.2), 则

$$\begin{aligned} \text{右式} &= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(-\omega)F(\omega) d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(-\omega) \int_{-\infty}^{\infty} x(t)e^{-j\omega t} dt d\omega \\ &= \int_{-\infty}^{\infty} x(t) \frac{1}{2\pi} \int_{-\infty}^{\infty} F(-\omega)e^{-j\omega t} d\omega dt \\ &= \text{左式} \end{aligned} \quad (2.5.7)$$

通常, 可把 $x(t)$ 看做是加在 1Ω 电阻上的电压, 那么式(2.5.5)左边就表示总能量(总功), 相应地称 $|F(\omega)|^2$ 为 $x(t)$ 的能谱密度。

在工程技术中, 对很多确定性函数, 其总能量是无限的, 式(2.5.5)无意义, 不能用总

能量表示它。但平均功率(平均能量)总是存在的,可用平均功率表示它。为此,选有限值 T , 对 $x(t)$ 构造截尾函数 $x_T(t)$

$$x_T(t) = \begin{cases} x(t) & (|t| \leq T) \\ 0 & (|t| > T) \end{cases} \tag{2.5.8}$$

由式(2.5.8)和式(2.5.5), 容易写出在区间 $[-T, T]$ 上的平均功率 $P(T)$ 为

$$P(T) = \frac{1}{2T} \int_{-T}^T x^2(t) dt = \frac{1}{2} \int_{-T}^T F_T(\omega) d\omega \tag{2.5.9}$$

式中:

$$F_T(\omega) = \int_{-T}^T x(t) e^{-j\omega t} dt \tag{2.5.10}$$

令 $T \rightarrow \infty$, 对式(2.5.9)两边取极限, 得到总平均功率 P 为

$$\begin{aligned} P &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x^2(t) dt \\ &= \frac{1}{2} \lim_{T \rightarrow \infty} \int_{-T}^T F_T(\omega) d\omega \\ &= \frac{1}{2} \int_{-\infty}^{\infty} S_x(\omega) d\omega \end{aligned} \tag{2.5.11}$$

式中:

$$S_x(\omega) = \lim_{T \rightarrow \infty} \frac{1}{2T} F_T(\omega) F_T^*(\omega) \tag{2.5.12}$$

定义为 $x(t)$ 的功率谱密度函数, 简称谱密度; 而 $S_x(\omega) d\omega$ 称为谱分布函数。

下面把确定性函数的谱密度推广到随机过程情况。

2.5.2 随机过程的谱密度

1. 基本概念

考虑随机过程 $x(t)$, 当然 $x^2(t)$ 也是随机过程。对随机过程 $x(t)$ 直接使用式(2.5.11)是不方便的, 但只要对式(2.5.11)两边取均值, 就可得到适合随机过程的平均功率公式了。于是得到

$$\begin{aligned} \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T E\{x^2(t)\} dt &= \frac{1}{2} \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T E\{F_T(\omega) F_T^*(\omega)\} d\omega \\ &= \frac{1}{2} \int_{-\infty}^{\infty} S_x(\omega) d\omega \end{aligned} \tag{2.5.13}$$

式中:

$$S_x(\omega) = \lim_{T \rightarrow \infty} \frac{1}{2T} E\{F_T(\omega) F_T^*(\omega)\} \tag{2.5.14}$$

定义为随机过程 $x(t)$ 的功率谱密度函数, 简称谱密度; 而 $S_x(\omega) d\omega$ 称为谱分布函数; 另外, $E\{x^2(t)\}$ 是 $x(t)$ 的均方值函数 $C_x(t)$, 即

$$C_x(t) = E\{x^2(t)\} \tag{2.5.15}$$

则由式(2.5.13) ~ (2.5.15) 得到

$$A[C_x(t)] = \frac{1}{2} \int_{-\infty}^{\infty} S_x(\omega) d\omega \tag{2.5.16}$$

如果 $x(t)$ 是平稳随机过程, 均方值函数与时间无关, 有

$$A[C_x(t)] = C_x(0) \quad (2.5.17)$$

则式(2.5.16)变为

$$C_x(0) = \frac{1}{2} \int_{-\infty}^{\infty} x(\tau) d\tau \quad (2.5.18)$$

例 2.5.1 考虑随机过程

$$x(t) = A \cos(\omega_0 t + \theta) \quad (2.5.19)$$

式中: A 和 ω_0 是实常数; θ 是具有一致(均匀)分布

$$f(\theta) = \begin{cases} \frac{1}{2\pi} & (0 \leq \theta < 2\pi) \\ 0 & (\text{其他}) \end{cases} \quad (2.5.20)$$

的随机变量。试验证式(2.5.16)成立。

解 先求式(2.5.16)左边结果, $x(t)$ 的均方值函数为

$$\begin{aligned} C_x(t) &= E\{[A \cos(\omega_0 t + \theta)]^2\} \\ &= E\left\{\frac{A^2}{2} + \frac{A^2}{2} \cos(2\omega_0 t + 2\theta)\right\} \\ &= \frac{A^2}{2} + \frac{A^2}{2} \int_0^{2\pi} \frac{1}{2\pi} \cos(2\omega_0 t + 2\theta) d\theta \\ &= \frac{A^2}{2} - \frac{A^2}{2} \sin 2\omega_0 t \end{aligned} \quad (2.5.21)$$

显然, 由于 $C_x(t)$ 是 t 的函数, 不是常数, $x(t)$ 不是平稳随机过程。均方值函数 $C_x(t)$ 的时间均值为

$$\begin{aligned} A[C_x(t)] &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \left[\frac{A^2}{2} - \frac{A^2}{2} \sin 2\omega_0 t \right] dt \\ &= \frac{A^2}{2} \end{aligned} \quad (2.5.22)$$

其次求式(2.5.16)右边结果, $x_T(t)$ 的频谱 $F_T(\omega)$ 为

$$\begin{aligned} F_T(\omega) &= \int_{-T}^T A \cos(\omega_0 t + \theta) e^{-j\omega t} dt \\ &= \frac{A}{2} e^{j\theta} \int_{-T}^T e^{j(\omega_0 - \omega)t} dt + \frac{A}{2} e^{-j\theta} \int_{-T}^T e^{-j(\omega_0 + \omega)t} dt \\ &= AT e^{j\theta} \frac{\sin[(\omega_0 - \omega)T]}{(\omega_0 - \omega)T} + \frac{A}{2} T e^{-j\theta} \frac{\sin[(\omega_0 + \omega)T]}{(\omega_0 + \omega)T} \end{aligned} \quad (2.5.23)$$

$$\begin{aligned} E\{|F_T(\omega)|^2\} &= E\{F_T(\omega) F_T^*(-\omega)\} \\ &= A^2 T^2 \frac{\sin^2[(\omega_0 - \omega)T]}{(\omega_0 - \omega)^2 T^2} + A^2 T^2 \frac{\sin^2[(\omega_0 + \omega)T]}{(\omega_0 + \omega)^2 T^2} \end{aligned} \quad (2.5.24)$$

利用公式

$$\lim_{T \rightarrow \infty} \frac{T}{2\pi} \frac{\sin^2 \frac{T}{2} x}{\left(\frac{T}{2}\right)^2} = \delta(x) \quad (2.5.25)$$

得到谱密度函数

$$x(\omega) = \frac{A^2}{2} [(\omega - \omega_0) + (\omega + \omega_0)] \quad (2.5.26)$$

计算(2.5.16)右边,得到

$$\frac{1}{2} \int_{-\infty}^{\infty} x(\omega) d\omega = \frac{A^2}{2} \quad (2.5.27)$$

式(2.5.27)与式(2.5.22)结果相同,即验证了式(2.5.16)成立。

2. 相关函数与谱密度的关系

设 $C_x(t_1, t_2)$ 为随机过程 $x(t)$ 的自相关函数:

$$\begin{aligned} C_x(t_1, t_2) &= E\{x(t_1)x(t_2)\} \\ &= E\{x(t)x(t+\tau)\} \end{aligned} \quad (2.5.28)$$

式中已令

$$t = t_1 \quad (t + \tau = t_2) \quad (2.5.29)$$

下面证明一个重要结果,用定理表示。

定理 2.5.2 考虑随机过程 $x(t)$, 则它的自相关函数的时间均值 $A[C_x(t, t+\tau)]$ 与谱密度 $x(\omega)$ 是一傅里叶变换对

$$x(\omega) = \int_{-\infty}^{\infty} A[C_x(t, t+\tau)] e^{-i\omega\tau} d\tau \quad (2.5.30)$$

$$A[C_x(t, t+\tau)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} x(\omega) e^{i\omega\tau} d\omega \quad (2.5.31)$$

证明 对式(2.5.30)和(2.5.31)二式,只要证明其中一个式子成立,另一式必然成立,因此只证明式(2.5.30)成立就够了。

证明 从 $x(\omega)$ 的定义式开始,把式(2.5.10)代入式(2.5.14),得到

$$\begin{aligned} x(\omega) &= \lim_T \frac{1}{2T} E\{F_T(-\omega)F_T(\omega)\} \\ &= \lim_T \frac{1}{2T} E \int_{-T}^T x(t_1) e^{i\omega t_1} dt_1 \int_{-T}^T x(t_2) e^{-i\omega t_2} dt_2 \\ &= \lim_T \frac{1}{2T} \int_{-T}^T \int_{-T}^T E\{x(t_1)x(t_2)\} e^{-i\omega(t_2-t_1)} dt_1 dt_2 \end{aligned} \quad (2.5.32)$$

由式(2.5.29)得到

$$dt = dt_1 \quad d = d(t_2 - t_1) = dt_2 \quad (2.5.33)$$

把式(2.5.28), (2.5.29)和(2.5.33)代入式(2.5.32)进行变量置换,得到

$$\begin{aligned} x(\omega) &= \lim_T \frac{1}{2T} \int_{-(T-\tau)}^{T-\tau} \int_{-T}^T C_x(t, t+\tau) e^{-i\omega\tau} dt d\tau \\ &= \lim_T \frac{1}{2T} \int_{-T}^T C_x(t, t+\tau) dt \int_{-\infty}^{\infty} e^{-i\omega\tau} d\tau \\ &= \int_{-\infty}^{\infty} A[C_x(t, t+\tau)] e^{-i\omega\tau} d\tau \end{aligned} \quad (2.5.34)$$

式中 $A[C_x(t, t+\tau)]$ 是自相关函数的时间均值

$$A[C_x(t, t+\tau)] = \lim_T \frac{1}{2T} \int_{-T}^T C_x(t, t+\tau) dt \quad (2.5.35)$$

式(2.5.34)就是傅里叶正变换式(2.5.30)。

推论 若 $x(t)$ 是平稳随机过程, 则它的自相关函数 $C_x(\tau)$ 和谱密度是一傅里叶变换对

$$x(\omega) = \int_{-\infty}^{\infty} C_x(\tau) e^{-j\omega\tau} d\tau \tag{2.5.36}$$

$$C_x(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} x(\omega) e^{j\omega\tau} d\omega \tag{2.5.37}$$

以上二式称为维纳-辛钦(Wiener-Khintchine)公式。

证明 由于平稳随机过程的自相关函数只是时间差的函数, 有

$$A[C_x(t, t + \tau)] = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T C_x(\tau) d\tau = C_x(\tau) \tag{2.5.38}$$

因此把式(2.5.38)代入式(2.5.30)和(2.5.31), 就得到式(2.5.36)和(2.5.37)。

考虑平稳离散型随机序列 $x(k)$, 它的自相关函数 $C_x(n)$ 为

$$C_x(n) = E\{x(k)x(k + n)\} \tag{2.5.39}$$

同样, 自相关函数 $C_x(n)$ 与谱密度是一离散型傅里叶变换对

$$x(\omega) = \sum_{n=-\infty}^{\infty} C_x(n) e^{-j\omega n} \tag{2.5.40}$$

$$C_x(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} x(\omega) e^{j\omega n} d\omega \tag{2.5.41}$$

例 2.5.2 已知随机过程 $x(t)$ 的谱密度

$$x(\omega) = \frac{\omega^2 + 4}{\omega^4 + 10\omega^2 + 9} \tag{2.5.42}$$

试求 $x(t)$ 的自相关函数和均方值。

解 把式(2.5.42)分解为

$$x(\omega) = \frac{5}{8} \frac{1}{\omega^2 + 9} + \frac{3}{8} \frac{1}{\omega^2 + 1} \tag{2.5.43}$$

利用傅里叶变换公式

$$\frac{2a}{\omega^2 + a^2} \longleftrightarrow e^{-a|\tau|} \tag{2.5.44}$$

则可得到平稳随机过程 $x(t)$ 的自相关函数 $C_x(\tau)$ 和均方值 $C_x(0)$:

$$C_x(\tau) = \frac{5}{48} e^{-3|\tau|} + \frac{3}{16} e^{-|\tau|} \tag{2.5.45}$$

$$C_x(0) = \frac{5}{48} + \frac{3}{16} = \frac{7}{24} \tag{2.5.46}$$

例 2.5.3 已知连续时间平稳随机过程 $x(t)$ 的自相关函数

$$C(\tau) = e^{-a|\tau|} \tag{2.5.47}$$

试确定它的谱密度 $x(\omega)$ 。

解 把式(2.5.47)代入式(2.5.36), 对 τ 的绝对值进行处理, 并进行一般积分运算, 得到

$$x(\omega) = \int_{-\infty}^{\infty} e^{-a|\tau|} e^{-j\omega\tau} d\tau$$

$$\begin{aligned}
 &= \int_{-\infty}^0 e^{-i\omega\tau+a\tau} d\tau + \int_0^{\infty} e^{-i\omega\tau-a\tau} d\tau \\
 &= \frac{2a}{a^2 + \omega^2}
 \end{aligned}
 \tag{2.5.48}$$

例 2.5.4 已知离散时间平稳随机序列 $x(k)$ 的自相关函数

$$C(\tau) = \frac{1}{1-a^2}(-a)^{|\tau|} \tag{2.5.49}$$

试确定它的谱密度 $S_x(\omega)$ 。

解 把式(2.5.49)代入式(2.5.40), 并进行运算, 得到

$$\begin{aligned}
 S_x(\omega) &= \sum_{\tau=-\infty}^{\infty} C(\tau)e^{-i\omega\tau} = C(0) + \sum_{\tau=1}^{\infty} C(\tau)e^{-i\omega\tau} + \sum_{\tau=1}^{\infty} C(-\tau)e^{-i\omega(-\tau)} \\
 &= \frac{1}{1-a^2} \left[\sum_{\tau=-\infty}^{\infty} (-a)^{|\tau|} e^{-i\omega\tau} \right] = \frac{1}{1-a^2} \left[1 + \sum_{\tau=1}^{\infty} (-a)^{\tau} e^{-i\omega\tau} + \sum_{\tau=1}^{\infty} (-a)^{\tau} e^{i\omega\tau} \right] \\
 &= \frac{1}{1-a^2} \left[\frac{1 - (-a)^{\infty} e^{-i\omega\infty}}{1 - (-a)e^{-i\omega}} + \frac{1 - (-a)^{\infty} e^{i\omega\infty}}{1 - (-a)e^{i\omega}} \right] \\
 &= \frac{1}{1 + a(e^{i\omega} + e^{-i\omega}) + a^2} \\
 &= \frac{1}{1 + 2a\cos\omega + a^2}
 \end{aligned}
 \tag{2.5.50}$$

3. 谱密度函数的有关性质

当随机过程 $x(t)$ 为实函数时, 它的谱密度函数 $S_x(\omega)$ 具有下列有关性质:

(1) $S_x(\omega) \geq 0$ (2.5.51)

(2) $S_x(\omega)$ 是实函数

(3) $S_x(\omega) = S_x(-\omega)$ (2.5.52)

由式(2.5.4)和(2.5.14)容易证明上述三条性质成立。

(4) $S_x(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} C(\tau)e^{-i\omega\tau} d\tau$ (2.5.53)

由 $x(t)$ 和 $F(\omega)$ 容易证明

$$\frac{dx(t)}{dt} = (j\omega)F(\omega) \tag{2.5.54}$$

再把上式代入式(2.5.14)就得到式(2.5.53)。

4. 几点说明

(1) 本节介绍谱密度概念, 由确定性情况推广到随机过程情况, 物理概念清楚。另外, 也可由相关函数直接定义谱密度函数, 即直接把平稳过程相关函数的傅里叶变换定义为谱密度函数, 而不涉及一般随机过程情况。

(2) 经推导已得到平稳过程的相关函数与谱密度是一对傅里叶变换的结论。另外, 也可把平稳过程的协方差函数与谱密度看做是一对傅里叶变换。两种做法无本质区别, 只要注意到相关函数 $C(\tau)$ 与协方差函数 $r(\tau)$ 有关系式

$$r(\tau) = C(\tau) - m^2 \tag{2.5.55}$$

(3) 傅里叶变换对公式中, 系数 2 的位置可有不同处理, 使用中要加以注意。

2.5.3 互谱密度

1. 基本概念

互谱密度函数是在频域上描述两个随机过程之间的相关性的。在实际应用中,常常利用测定系统输入输出的互谱密度来确定系统的特性。

考虑两个实随机过程 $x(t_1)$ 和 $y(t_2)$, 它们的互相关函数为

$$\begin{aligned} C_{xy}(t_1, t_2) &= E\{x(t_1)y(t_2)\} \\ &= E\{x(t)y(t+\tau)\} \end{aligned} \tag{2.5.56}$$

式中已令

$$t = t_1 \quad t + \tau = t_2 \tag{2.5.57}$$

在 $x(t_1)$ 和 $y(t_2)$ 之间存在互谱密度, 这里不加证明, 直接给出结果。互谱密度直接定义为

$$S_{xy}(\omega) = \lim_T \frac{1}{2T} E\{F_{xT}(\omega) \overline{F_{yT}(\omega)}\} \tag{2.5.58}$$

互相关函数的时间均值与互谱密度是一傅里叶变换对, 即

$$A[C_{xy}(t, t+\tau)] \quad S_{xy}(\omega) \tag{2.5.59}$$

如果 $x(t_1)$ 和 $y(t_2)$ 是联合平稳的, 则有

$$C_{xy}(t_1, t_2) = C_{xy}(\tau) \tag{2.5.60}$$

$$C_{xy}(\tau) = \overline{xy(\tau)} \tag{2.5.61}$$

2. 互谱密度函数的性质

考虑实随机过程 $x(t_1)$ 和 $y(t_2)$, 互谱密度为 $S_{xy}(\omega)$, 下面只给出 $S_{xy}(\omega)$ 的几个基本性质, 不再加以证明:

$$(1) \quad S_{xy}(\omega) = \overline{S_{yx}(\omega)} = S_{yx}(-\omega) \tag{2.5.62}$$

(2) $\text{Re}[S_{xy}(\omega)]$ 和 $\text{Re}[S_{yx}(\omega)]$ 是 ω 的偶函数, Re 表示实部。

(3) $\text{Im}[S_{xy}(\omega)]$ 和 $\text{Im}[S_{yx}(\omega)]$ 是 ω 的奇函数, Im 表示虚部。

(4) 如果 $x(t_1)$ 和 $y(t_2)$ 正交, 则

$$S_{xy}(\omega) = 0 \quad S_{yx}(\omega) = 0 \tag{2.5.63}$$

(5) 如果 $x(t_1)$ 和 $y(t_2)$ 无关, 且均值为常数 \bar{x} 和 \bar{y} , 则

$$S_{xy}(\omega) = S_{yx}(\omega) = 2\bar{x}\bar{y}\delta(\omega) \tag{2.5.64}$$

例 2.5.5 设联合平稳随机过程 $x(t)$ 和 $y(t)$ 的互谱密度为

$$S_{xy}(\omega) = \begin{cases} a + ib/c & (-c < \omega < c) \\ 0 & (\text{其他}) \end{cases} \tag{2.5.65}$$

式中 a, b, c 是实常数, $c > 0$, 求互相关函数 $C_{xy}(\tau)$ 。

解 按傅里叶逆变换公式求解:

$$\begin{aligned} C_{xy}(\tau) &= \frac{1}{2} \int_{-c}^c (a + i\frac{b}{c} e^{i\omega\tau}) d\omega \\ &= \frac{a}{2} \int_{-c}^c e^{i\omega\tau} d\omega + i\frac{b}{2c} \int_{-c}^c e^{i\omega\tau} d\omega \end{aligned}$$

$$\begin{aligned}
 &= \frac{a}{2} \left. \frac{e^i}{it} \right|_{-c}^c + \frac{ib}{2c} e^i \left. \frac{1}{i} - \frac{1}{(i)^2} \right|_{-c}^c \\
 &= \frac{1}{c^2} [(2c - b) \text{sinc} + bc \cos c]
 \end{aligned} \tag{2.5.66}$$

计算中利用了公式

$$\begin{aligned}
 e^{ax} dx &= \frac{e^{ay}}{a} & x e^{ay} dx &= e^{ax} \left(\frac{x}{a} - \frac{1}{a^2} \right) \\
 \cos &= \frac{1}{2}(e^i + e^{-i}) & \sin &= \frac{1}{2i}(e^i - e^{-i})
 \end{aligned} \tag{2.5.67}$$

2.5.4 复谱密度

考虑平稳复随机过程情况。设

$$\begin{aligned}
 z(t) &= x(t) + iy(t) \\
 \text{是平稳复随机过程, 自相关函数为 } C_z(\cdot), \text{ 复谱密度为 } z(\cdot), \text{ 则有关系式} \\
 C_z(\cdot) &= z(\cdot)
 \end{aligned} \tag{2.5.68}$$

设有联合平稳复随机过程 $z_1(t_1)$ 和 $z_2(t_2)$, 其互相关函数为 $C_{z_1 z_2}(\cdot)$, 互复谱密度为 $z_1 z_2(\cdot)$, 则有关系式

$$C_{z_1 z_2}(\cdot) = z_1 z_2(\cdot) \tag{2.5.69}$$

2.6 白 噪 声

白噪声是一种非常重要的随机过程, 随机系统中的随机扰动常用白噪声表述, 很多随机过程可用白噪声加到一个网络上产生。

2.6.1 白噪声概念

1. 一般定义

定义 2.6.1 考虑随机过程 $v(t)$, 如果它的谱密度等于常数, 即

$$(\cdot) = C \tag{2.6.1}$$

则称 $v(t)$ 为白噪声。

白噪声的“白”字来源于白色光, 白色光由强度相等的各种频率(波长)的光组成。分析白噪声时, 其均值是否为零, 并非本质问题, 但为简化分析, 也符合大多数实际情况, 一般假设均值为零。

离散时间白噪声和连续时间白噪声都由式(2.6.1)定义, 但它们的自相关函数有差别, 并可由自相关函数定义白噪声, 下面加以讨论。

2. 离散时间白噪声

考虑离散时间白噪声 $v(k)$, 它的谱密度由式(2.6.1)决定。设 $v(k)$ 是平稳的, 它的自相关函数为

$$C(n) = E\{v(k)v(l)\} \tag{2.6.2}$$

式中已令 $k-1=n$ 。把式(2.6.1)代入傅里叶逆变换式(2.5.41),可得:

$$\begin{aligned} C(n) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} C e^{j n d} d\omega \\ &= \frac{C \sin n}{n} \\ &= \begin{cases} C & (n=0) \\ 0 & (n=\pm 1, \pm 2, \dots) \end{cases} \\ &= C_n \end{aligned} \tag{2.6.3}$$

由式(2.6.3)可得到离散时间白噪声的等价定义。

定义 2.6.2 考虑随机序列 $v(k)$, 如果它的自相关函数满足式(2.6.3), 则称 $v(k)$ 为离散时间白噪声。

定义式(2.6.3)中, 当 C 为常数时, 则 $v(k)$ 称为平稳白噪声; 当 C 为函数 $C(k)$ 时, 则 $v(k)$ 称为非平稳白噪声。如果 $v(k)$ 的分布是高斯的, 则 $v(k)$ 称为高斯白噪声。

离散时间平稳白噪声的谱密度和自相关函数如图 2.6.1 表示。

图 2.6.1 离散时间白噪声的谱密度 $S(\omega)$ 和自相关函数 $C(n)$

3. 连续时间白噪声

(1) $S(\omega)$ 函数的谱密度

$S(\omega)$ 函数由下列公式定义:

$$S(\omega) = \begin{cases} \frac{1}{2\pi} & (0^- < \omega < 0^+) \\ 0 & (\text{其他}) \end{cases} \tag{2.6.4}$$

$$\int_{-\infty}^{\infty} S(\omega) d\omega = 1 \tag{2.6.5}$$

把式(2.6.4)和(2.6.5)代入式(2.5.36), 容易求出 $S(\omega)$ 函数的谱密度 $S(\omega)$ 为 1, 即

$$S(\omega) = 1 \tag{2.6.6}$$

$$C(\tau) = \delta(\tau) \tag{2.6.7}$$

(2) 连续时间白噪声的自相关函数

考虑连续时间白噪声 $v(t)$, 它的谱密度由式(2.6.1)决定。设 $v(t)$ 是平稳的, 它的自相关函数为

$$C(\tau) = E\{v(s)v(t)\} \tag{2.6.8}$$

式中已令 $s-t=\tau$ 。利用式(2.6.7)直接得到自相关函数:

$$C(\tau) = C(-\tau) \tag{2.6.9}$$

由式(2.6.9)可得到连续时间白噪声的等价定义。

定义 2.6.3 考虑随机过程 $v(t)$, 如果它的自相关函数满足式(2.6.9), 则称 $v(t)$ 为连续时间白噪声。

连续时间白噪声的谱密度和自相关函数如图 2.6.2 所示。式(2.6.9)中, 当 C 为常数时, 则 $v(t)$ 称为平稳白噪声; 当 C 为函数 $C(t)$ 时, 则 $v(t)$ 称为非平稳白噪声。如果 $v(t)$ 的分布是高斯的, 则 $v(t)$ 称为高斯白噪声。

图 2.6.2 连续时间白噪声的谱密度 $S_v(f)$ 和自相关函数 $C(\tau)$

值得注意的是, 由式(2.6.9)可知, 连续时间白噪声的方差为无穷大, 即

$$r(0) = \infty \tag{2.6.10}$$

因此, 连续时间白噪声不是二阶过程, 实际上是不存在的, 它只是一种理想情况, 便于理论分析。

2.6.2 带限白噪声和有色噪声

1. 带限白噪声

带限白噪声能克服连续时间白噪声方差为无穷大的问题。

定义 2.6.4 考虑随机过程 $v(t)$, 如果它的谱密度满足

$$S_v(f) = \begin{cases} C & (|f| \leq B) \\ 0 & (|f| > B) \end{cases} \tag{2.6.11}$$

则称 $v(t)$ 为带限白噪声。

把式(2.6.11)代入傅里叶逆变换式(2.5.37), 得到自相关函数

$$\begin{aligned} C(\tau) &= \frac{1}{2\pi} \int_{-B}^B C e^{j2\pi f\tau} df \\ &= \frac{C}{2} \frac{\sin 2\pi B\tau}{\tau} \end{aligned} \tag{2.6.12}$$

带限白噪声的谱密度和自相关函数如图 2.6.3 所示。

2. 有色噪声

不是白噪声的任何噪声都称为有色噪声。下面仅举一个有色噪声的例子。

例 2.6.1 考虑一种平稳有色噪声 $v(t)$, 它的自相关函数为

$$C(\tau) = e^{-\alpha|\tau|} \cos \omega_0 \tau \tag{2.6.13}$$

式中 α, ω_0 为常数, 把式(2.6.13)代入式(2.5.36), 得到谱密度

$$\begin{aligned}
C(\tau) &= \int_{-\infty}^{\infty} \frac{1}{2} e^{-\omega|\tau|} (e^{i\omega\tau} + e^{-i\omega\tau}) d\omega \\
&= \frac{1}{2} \int_{-\infty}^0 e^{-i\omega\tau + i\omega\tau} d\omega + \frac{1}{2} \int_0^{\infty} e^{-i\omega\tau + i\omega\tau} d\omega \\
&\quad + \frac{1}{2} \int_{-\infty}^0 e^{-i\omega\tau - i\omega\tau} d\omega + \frac{1}{2} \int_0^{\infty} e^{-i\omega\tau - i\omega\tau} d\omega \\
&= \frac{1}{\tau^2 + (\frac{1}{2})^2} + \frac{1}{\tau^2 + (\frac{1}{2})^2}
\end{aligned} \tag{2.6.14}$$

图 2.6.3 带限白噪声的谱密度 $S(\omega)$ 和自相关函数 $C(\tau)$

2.6.3 连续时间平稳高斯白噪声与维纳过程的关系

连续时间平稳高斯白噪声与维纳过程之间的关系是分析连续时间随机状态方程遇到的一个基本问题,两者之间的关系可用定理表示。

定理 2.6.1 维纳过程 $w(t)$ 是连续时间平稳高斯白噪声 $v(t)$ 的积分过程,即

$$w(t) = \int_0^t v(\tau) d\tau \tag{2.6.15}$$

或者说,维纳过程 $w(t)$ 的微分过程是连续时间平稳高斯白噪声 $v(t)$

$$\frac{dw(t)}{dt} = v(t) \tag{2.6.16}$$

证明 设连续时间平稳高斯白噪声给定为 $v(t) \sim N(0, C(\tau))$, C 为常数,即

$$E\{v(t)\} = 0 \tag{2.6.17}$$

$$C(\tau) = E\{v(t)v(t+\tau)\} = C \tag{2.6.18}$$

下面考察式(2.6.15)是否符合维纳过程的四个定义条件:

(1) 由式(2.6.15)容易得到

$$w(0) = \int_0^0 v(\tau) d\tau = 0 \tag{2.6.19}$$

即 $w(t)$ 初值为零。

(2) 已知 $v(t)$ 是高斯的,其积分过程式(2.6.15)可看做是一个求和过程,即

$$w(t) = \lim_{\max \Delta t_i \rightarrow 0} \sum_{i=1}^{n_t} v(t_i) \Delta t_i \tag{2.6.20}$$

所以 $w(t)$ 也是高斯的,因为高斯过程的线性组合仍是高斯过程。其中 n_i ($i=0, 1, \dots, n_t=t$) 是对积分区间 $[0, t]$ 的分点值, $\Delta t_i = n_i - n_{i-1}$ 是相邻分点间的时间增量。

(3) 对式(2.6.15)两边取均值, 得到

$$E\{w(t)\} = E \int_0^t v(\tau) d\tau = \int_0^t E\{v(\tau)\} d\tau = 0 \quad (2.6.21)$$

即 $w(t)$ 的均值为零。

(4) 为证明 $w(t)$ 是否具有平稳独立增量, 设定时间 $t_{j-1} < t_j$, $t_i < t_{i+1}$ 。首先证明 $w(t)$ 具有独立增量, 为此计算式(2.6.15)的增量协方差, 考虑到增量均值为零, 得到

$$\begin{aligned} r(w_i, w_j) &= E\{[w(t_{i+1}) - w(t_i)][w(t_j) - w(t_{j-1})]\} \\ &= E \int_{t_i}^{t_{i+1}} v(\tau_1) d\tau_1 \int_{t_{j-1}}^{t_j} v(\tau_2) d\tau_2 \\ &= \int_{t_i}^{t_{i+1}} d\tau_1 \int_{t_{j-1}}^{t_j} E\{v(\tau_1)v(\tau_2)\} d\tau_2 \\ &= \int_{t_i}^{t_{i+1}} d\tau_1 \int_{t_{j-1}}^{t_j} C(\tau_2 - \tau_1) d\tau_2 \\ &= 0 \end{aligned} \quad (2.6.22)$$

计算中从第三等式到第四等式利用了式(2.6.18), 从第四等式到第五等式利用了

$$C(\tau_2 - \tau_1) = 0 \quad (\tau_2 - \tau_1 > 0) \quad (2.6.23)$$

式(2.6.22)说明 $w(t)$ 具有不相关增量, 考虑到 $w(t)$ 是高斯的, 不相关等价于独立, 因此 $w(t)$ 具有独立增量。

其次, 证明 $w(t)$ 具有平稳增量。为此计算 $w(t)$ 的增量方差:

$$\begin{aligned} r(w_i) &= E\{[w(t_{i+1}) - w(t_i)]^2\} \\ &= C(t_{i+1} - t_i) \end{aligned} \quad (2.6.24)$$

推导过程与式(2.6.22)的相同, 只要注意到

$$C(\tau_2 - \tau_1) = 1 \quad (\tau_2 - \tau_1 = 0) \quad (2.6.25)$$

式(2.6.24)说明, $w(t)$ 的增量方差只是时间差的函数, 因此, $w(t)$ 具有平稳增量。

由式(2.6.22)和(2.6.24)可知, $w(t)$ 具有平稳独立增量。

综上所述, $w(t)$ 符合维纳过程的四个定义条件, 所以, 连续时间平稳高斯白噪声的积分过程是维纳过程。

2.6.4 标准均匀分布伪随机数的生成

本节介绍一种生成白噪声的方法, 即生成标准均匀分布伪随机数的方法, 这是在分析和研究随机系统时, 生成白噪声扰动的基本方法。

1. 基本概念

(1) 标准均匀分布函数 $F(x)$ 表示为

$$F(x) = \begin{cases} 0 & (x < 0) \\ x & (0 \leq x < 1) \\ 1 & (1 \leq x) \end{cases} \quad (2.6.26)$$

其均值 $E\{x\} = \frac{1}{2}$, 方差 $r(x) = \frac{1}{12}$ 。

(2) 若整数 x 与 y 之差 $(x - y)$ 是整数 m 的倍数, 则称 x 与 y 对 m 同余, 记作

$$x = y \pmod{m} \quad (2.6.27)$$

取数 x_{i-1} 乘以常数 a , 然后被 m 除, 取其余数为 x_i , 记作

$$x_i = ax_{i-1} \pmod{m} \quad (2.6.28)$$

例如, $a=8$, $m=64$, $x_{i-1}=19$, 则

$$x_i = 8 \times 19 \pmod{64} = 24 \quad (2.6.29)$$

在一定条件下, 由递推式(2.6.29)得到的序列 $\{x_i\}$ 是均匀分布的, 它是一种伪随机数。这种生成伪随机数的方法称为乘同余法。

2. 生成方法

(1) 给定基数 R , 对十进数, 取 $R=10$ 。

(2) 给定伪随机数的有效字长 N , 选择模 m

$$m = R^N \quad (2.6.30)$$

(3) 确定常数乘子 IX (即式(2.6.28)中的 a)

$$IX = 200IT \pm IR \quad (2.6.31)$$

式中: IR 在 3, 11, 13, 19, 21, 27, 29, 37, 53, 59, 61, 67, 69, 77, 83 和 91 等数中选择; IT 可选定为任意正整数。配合选择 IR 和 IT , 使 IX 接近 $10^{\frac{N}{2}}$ (对 $R=10$)。

(4) 任意给定初值 NR , 其值不能为 R 的因数整除, 对 $k=1, 2, \dots$, 递推计算数列

$$NR_k = IX \times NR_{k-1} \pmod{m} \quad (2.6.32)$$

(5) 把 NR_k ($k=1, 2, \dots$) 置于小数点之后, 就得到一组标准均匀分布的伪随机数。

(6) 伪随机数 NR_k ($k=1, 2, \dots$) 满足

$$NR_k = NR_{k+p} \quad (2.6.33)$$

时, 使上式成立的最小整数 p 称为周期。对 $R=10$, $N=4$, 一般有

$$p = 5 \times 10^{N-2} = \frac{m}{20} \quad (2.6.34)$$

可参考周期 p , 确定 NR_k 的数量。

例 2.6.2 给定 $R=10$, $N=4$, $IR=91$, $IT=0$, $NR_0=1973$, 求标准均匀分布伪随机数序列。

解 利用式(2.6.31)和(2.6.32), 代入给定数据, 得到

$$IX = 91$$

$$NR_1 = 91 \times 1973 \pmod{10^4} = 9543$$

$$NR_2 = 91 \times 9543 \pmod{10^4} = 8413$$

$$NR_3 = 91 \times 8413 \pmod{10^4} = 5583$$

利用式(2.6.34), 求出参考周期 $p=500$ 。

最后得到标准均匀分布伪随机数 0.9543, 0.8413, 0.5583, ...。

把本例编制程序, 得到数据如图 2.6.4 所示。求得均值 $E\{NR_k\} = 0.4989$, 方差 $r(NR_k) = 0.0835$ ($k=1, 2, \dots, 500$)。

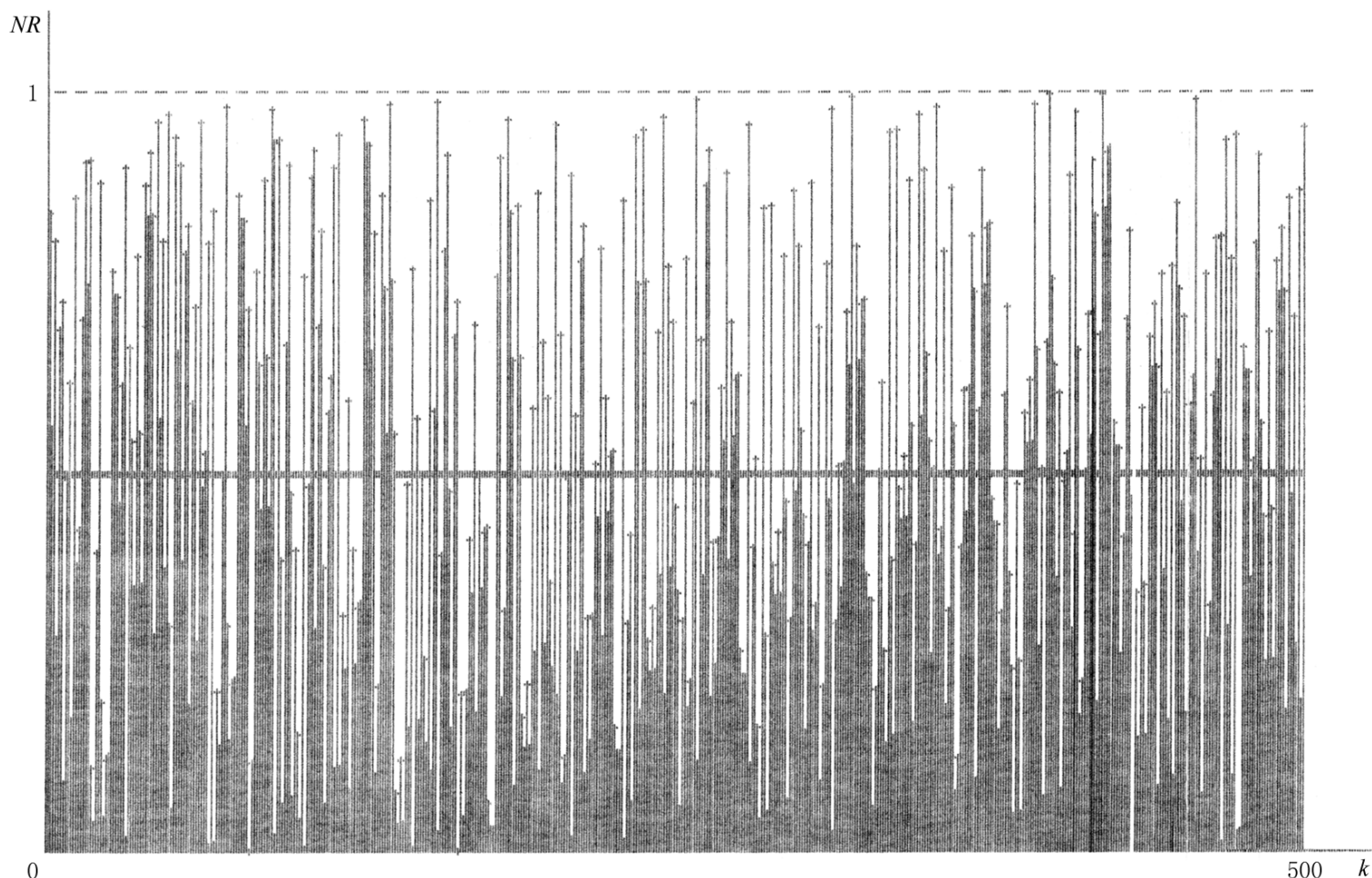


图 2.6.4 标准均匀分布伪随机数序列

2.6.5 高斯白噪声的生成

在分析和研究随机系统时,所加噪声多为高斯白噪声,因此,应当了解如何生成高斯白噪声。本节介绍两种利用均匀分布随机序列生成高斯白噪声的方法。

1. 方法 1

设 x_1 和 x_2 是两个在区间 $[0, 1]$ 上互相独立的均匀分布的随机序列,上节生成的标准均匀分布伪随机数序列就可作为这样的序列。构造

$$y_1 = (-2 \ln n x_1)^{1/2} \cos 2\pi x_2 \quad (2.6.35)$$

$$y_2 = (-2 \ln n x_1)^{1/2} \sin 2\pi x_2 \quad (2.6.36)$$

已知 n 为序列数,则 y_1 和 y_2 是互相独立的高斯分布 $N(0, 1)$ 的随机序列。

2. 方法 2

设 x_1, x_2, \dots, x_N 是互相独立的同为标准均匀分布的随机序列,其均值 $m = \frac{1}{2}$, 方差 $r = \frac{1}{12}$ 。根据中心极限定理,有

$$y = \frac{\sum_{i=1}^N x_i - \frac{m}{N}}{r} \quad (2.6.37)$$

当 N 较大时,其分布近似于高斯分布 $N(0, 1)$ 。一般选取 $N = 12$, 由式(2.6.37)得到

$$y = \frac{\sum_{i=1}^{12} x_i - 6}{6} \quad (2.6.38)$$

考虑 x_i 和 $(1 - x_i)$ 同为标准均匀分布, 可把式(2.6.38)改写为

$$\begin{aligned}
 y &= x_1 + x_2 + \dots + x_6 - x_7 - x_8 - \dots - x_{12} \\
 &= \sum_{i=1}^6 (x_i - x_{i+6})
 \end{aligned}
 \tag{2.6.39}$$

当 x 为标准均匀分布的随机序列(总体)时, 从中抽取 N 个随机子样 x_1, x_2, \dots, x_N , 则它们是互相独立的标准均匀分布的随机序列。根据这种情况, 生成 x , 并抽取 N 个子样, 利用式(2.6.39), 可得到高斯分布 $N(0, 1)$ 的随机序列。

例 2.6.3 利用方法 2 生成高斯白噪声。

解 由 2.6.4 节提供的方法生成标准均匀分布伪随机数序列, 利用方法 2 编制程序, 得到高斯白噪声如图 2.6.5 所示。求得均值 $m = -0.0131$, 方差 $r = 0.9947$ 。

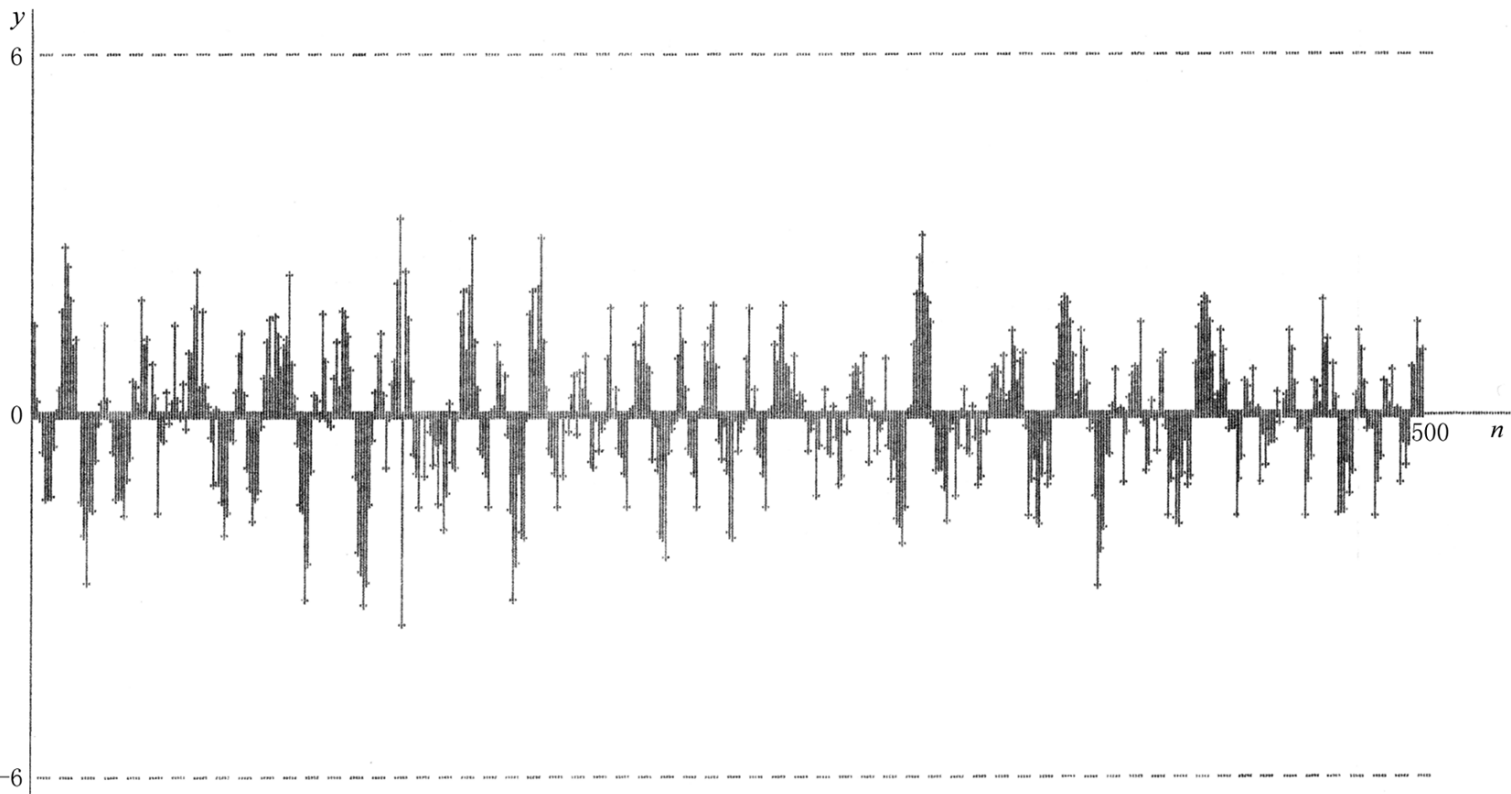


图 2.6.5 由方法 2 生成的高斯白噪声

例 2.6.4 利用方法 1 生成高斯白噪声。

解 由 2.6.4 节提供的方法生成标准均匀分布伪随机数序列, 利用方法 1 编制程序, 得到高斯白噪声的概率密度函数, 如图 2.6.6 所示。

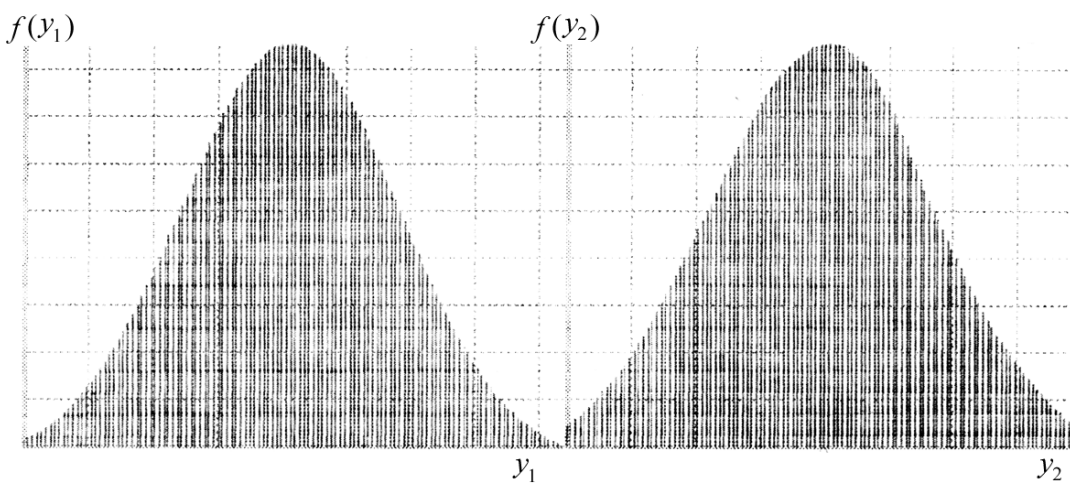


图 2.6.6 由方法 1 生成的高斯白噪声的概率密度函数

2.7 随机分析

为了分析随机动力学系统,有必要建立随机过程的收敛性、连续性、微分和积分等概念,进而研究随机微分方程。对应于人们熟悉的数学分析,这部分内容称为随机分析或随机微积分。

自从伊藤(K. Ito)在 1961 年发表《论随机微分方程》论文以来,随机微分方程受到了广泛的重视,取得飞速发展,渗透到很多领域。例如在随机控制理论、信号分析、分子物理学、原子物理学、化学动力学、固态扩散、结构稳定性等领域都得到了应用。

在这一节中介绍随机分析的基本知识,直接为连续时间的状态估计和随机控制做好理论准备。限于篇幅,只着重介绍基本概念和有关结论,有些证明没有给出。

2.7.1 收敛性

随机过程的收敛性是研究随机分析的基础,由于随机过程的不确定性,收敛性选择也是多种多样的,这一小节中首先给出四种收敛性定义,然后选择均方收敛作为研究问题的出发点,再介绍均方收敛性判据和性质。

1. 内积和范数

定义 2.7.1 设 x 和 y 是两个确定性向量,其夹角为 θ ,则我们知道 x 和 y 的内积 (x, y) 定义为

$$(x, y) = \|x\| \|y\| \cos \theta \tag{2.7.1}$$

式中 $\|x\|$ 称为 x 的范数,表示为

$$\|x\| = (x, x)^{\frac{1}{2}} \tag{2.7.2}$$

内积 (x, y) 是一个标量,它的物理意义可看做是向量力 x 在 y 方向上对物体做的功。若 $x(x_1, x_2, x_3)$ 和 $y(y_1, y_2, y_3)$ 表示直角坐标系上的两个向量,则有

$$(x, y) = x_1 y_1 + x_2 y_2 + x_3 y_3 \tag{2.7.3}$$

$$\|x\|^2 = (x, x) = x_1^2 + x_2^2 + x_3^2 \tag{2.7.4}$$

下面把确定性向量的内积和范数推广到随机过程。

考虑随机过程 x 和 y ,把 x 和 y 看做是两个随机向量,则它们的内积和范数分别定义为

$$\begin{aligned} (x, y) &= E\{x_1 y_1\} + E\{x_2 y_2\} + E\{x_3 y_3\} \\ &= E\{xy\} \end{aligned} \tag{2.7.5}$$

$$\|x\|^2 = E\{x_1^2\} + E\{x_2^2\} + E\{x_3^2\} \tag{2.7.6}$$

由式(2.7.5)可知,随机过程的内积等于它们的相关函数。

2. 收敛性定义

这里介绍四种收敛性定义,并给出各种收敛性之间的关系,最后选择均方收敛作为研究问题的基础。

定义 2.7.2

(1) 依均方收敛 考虑随机变量序列 $\{x_n, n=0, 1, \dots\}$, 如果存在随机变量 x 满足

$$\lim_n E\{X_n - x\}^2 = 0 \quad (2.7.7)$$

即

$$\lim_n E\{X_n - x\}^2 = 0 \quad (2.7.8)$$

则称随机变量序列 X_n 依均方收敛于随机变量 x , 并记作

$$X_n \xrightarrow{m.s} x \quad (2.7.9)$$

或

$$X_n \xrightarrow{m.s} x \quad (2.7.10)$$

$$X_n \xrightarrow{2} x \quad (2.7.11)$$

(2) 依概率 1 收敛 考虑随机变量序列 $\{X_n, n=0, 1, \dots\}$, 如果存在随机变量 x 满足

$$P\{\lim_n X_n = x\} = 1 \quad (2.7.12)$$

则称随机变量序列 X_n 依概率 1 (或极乎必然) 收敛于随机变量 x , 并记作

$$X_n \xrightarrow{a.s} x \quad (2.7.13)$$

(3) 依概率收敛 考虑随机变量序列 $\{X_n, n=0, 1, \dots\}$, 如果存在随机变量 x 满足

$$\lim_n P\{|X_n - x| > \epsilon\} = 0 \quad (2.7.14)$$

式中 ϵ 为任意小正数, 则称随机变量序列 X_n 依概率收敛于随机变量 x , 并记作

$$X_n \xrightarrow{p} x \quad (2.7.15)$$

或

$$X_n \xrightarrow{i.p} x \quad (2.7.16)$$

$$X_n \xrightarrow{p} x \quad (2.7.17)$$

(4) 依分布收敛 考虑随机变量序列 $\{X_n, n=0, 1, \dots\}$, 如果存在随机变量 x 满足

$$\lim_n F_n(x) = F(x) \quad (2.7.18)$$

则称随机变量 X_n 依分布 (或弱) 收敛于随机变量 x , 并记作

$$X_n \xrightarrow{w} x \quad (2.7.19)$$

上述四种收敛性的关系如图 2.7.1 所示。 $m.s$ 收敛和 $a.s$ 收敛是互相独立的, 它们之间没有确定的关系。 w 收敛最弱, 它包含 $i.p$ 收敛, $i.p$ 收敛又包含 $m.s$ 和 $a.s$ 两种收敛。

图 2.7.1 四种收敛性的关系

收敛性选择 按照分析和研究问题的要求, 可选择不同的收敛性。但在一般情况下, 分析和研究随机控制、随机过程或其他工程问题时, 都采用依均方收敛性, 这是因为依均

方收敛性能简化分析并比较实用。今后,本书中分析和研究问题一般都使用均方收敛性。

3. 两个均方收敛性判据

里斯-菲希尔定理(Riesz-Fisher theorem) 对随机变量序列 $\{x_n, n= 0, 1, \dots\}$ 构造柯西序列(Cauchy sequence) $x_n - x_m$, 如果满足

$$\lim_{n, m} E\{(x_n - x_m)^2\} = 0 \tag{2.7.20}$$

则必然存在一个随机变量 x , 使得

$$x_n \overset{m.p.s.}{\rightarrow} x \tag{2.7.21}$$

洛夫准则(Loeve criterion)(也称均方收敛准则) 随机变量序列 $\{x_n, n= 0, 1, \dots\}$ 均方收敛于 x 的充要条件是

$$\lim_{n, m} E\{x_n x_m\} = C \quad (C = \text{常数}) \tag{2.7.22}$$

4. 均方收敛性的性质

(1) 如果随机变量序列 $\{x_n, n= 0, 1, \dots\}$ 依均方收敛于随机变量 x , 则有

$$\lim_n E\{x_n\} = E\{\lim_n x_n\} = E\{x\} \tag{2.7.23}$$

即取极限和求均值的符号可以调换位置。注意第二个式子中括号内是取均方收敛极限, 依概率 1 收敛性不具有这一性质。

(2) 均方收敛性是唯一的。如果 $x_n \overset{m.p.s.}{\rightarrow} x$ 和 $x_n \overset{m.p.s.}{\rightarrow} y$, 则必有 $x = y$ 。

(3) 如果 $x_n \overset{m.p.s.}{\rightarrow} x$ 和 $y_n \overset{m.p.s.}{\rightarrow} y$, 则有

$$\lim_{n, m} E\{x_n y_m\} = E\{xy\} \tag{2.7.24}$$

(4) 如果 $x_n \overset{m.p.s.}{\rightarrow} x$ 和 $y_n \overset{m.p.s.}{\rightarrow} y$, a 和 b 是任意常数, 则有

$$\lim_n E\{ax_n + by_n\} = ax + by \tag{2.7.25}$$

2.7.2 均方连续性

定义 2.7.3 考虑二阶随机过程 $\{x(t), t \in T\}$, 如果对时刻 t 满足

$$\lim_{h \rightarrow 0} E\{x(t+h) - x(t)\}^2 = 0 \tag{2.7.26}$$

即

$$\lim_{h \rightarrow 0} E\{[x(t+h) - x(t)]^2\} = 0 \tag{2.7.27}$$

则称 $x(t)$ 对时刻 t 依均方连续。其中 h 为时间增量。如果 $x(t)$ 对任一 $t \in T$ 都依均方连续, 则称 $x(t)$ 在 T 上依均方连续。

均方连续性判据 二阶随机过程 $\{x(t), t \in T\}$ 对 $t \in T$ 依均方连续的充要条件是 $x(t)$ 的均值函数 $m(t)$ 对 t 连续, 以及自相关函数 $C(s, t)$ 在 (t, t) 连续。

例 2.7.1 试分析维纳过程 $w(t)$ 是否依均方连续。

解 由例 2.3.4 中式(2.3.72) 已知

$$E\{[dw(t)]^2\} = dt \tag{2.7.28}$$

把上式代入均方连续性定义式(2.7.27), 得到

$$\lim_{h \rightarrow 0} E\{[w(t+h) - w(t)]^2\} = \lim_{h \rightarrow 0} E\{[dw(h)]^2\}$$

$$= \lim_{h \rightarrow 0} \frac{1}{h^2} h = 0 \quad (2.7.29)$$

所以维纳过程是依均方连续的。

2.7.3 均方可微性

定义 2.7.4 考虑二阶随机过程 $\{x(t), t \in T\}$, 如果对时刻 t 有

$$\lim_{h \rightarrow 0} E \frac{x(t+h) - x(t)}{h} = x'(t) = 0 \quad (2.7.30)$$

即

$$\lim_{h \rightarrow 0} \frac{x(t+h) - x(t)}{h} = x'(t) \quad (2.7.31)$$

则称 $x(t)$ 对时刻 t 依均方可微, $x'(t)$ 称为均方导数。如果 $x(t)$ 对任一 $t \in T$ 都依均方可微, 则称 $x(t)$ 为可微的随机过程。

均方可微判据 一个二阶随机过程 $x(t)$ 对 $t \in T$ 依均方可微的充要条件是 $x(t)$ 的均值函数 $m(t)$ 对 t 可微, 以及自相关函数 $C(s, t)$ 的广义二阶导数

$$\frac{d^2 C(s, t)}{ds dt} = \lim_{h, k \rightarrow 0} \frac{1}{kh} [C(s+h, t+k) - C(s, t+k) - C(s+h, t) + C(s, t)] \quad (2.7.32)$$

在 $s=t$ 时存在。

均方可微性的性质如下:

(1) 均方导数是唯一的, 即如果 $x'(t) = \mu$ 和 $x'(t) = \nu$, 则必有 $\mu = \nu$ 。

(2) 如果 $x(t)$ 和 $y(t)$ 均方可微, a 和 b 是任意常数, 则

$$[ax(t) + by(t)]' = ax'(t) + by'(t) \quad (2.7.33)$$

(3) 如果 $x(t)$ 在 T 上均方可微, 则必在 T 上连续。反之则不成立。

(4) 如果 $x(t)$ 均方可微, $y(t)$ 是普通的可微函数, 则有

$$[y(t)x(t)]' = y'(t)x(t) + y(t)x'(t) \quad (2.7.34)$$

(5) 如果 $x(t)$ 在 T 上均方可微, 自相关函数 $C(t_1, t_2)$ 是可微函数, 则有

$$\frac{d}{dt} E\{x(t)\} = E\{x'(t)\} \quad (2.7.35)$$

$$\frac{d}{dt_1} C(t_1, t_2) = E\{x'(t_1)x(t_2)\} \quad (2.7.36)$$

$$\frac{d}{dt_1 dt_2} C(t_1, t_2) = E\{x'(t_1)x'(t_2)\} \quad (2.7.37)$$

即均方求导和求均值的符号可调换次序。

例 2.7.2 试分析维纳过程 $w(t)$ 的非均方可微性。

解 按定义式, 但不考虑 $w(t)$ 项, 并注意到式(2.7.28), 得到

$$\lim_{h \rightarrow 0} E \frac{w(t+h) - w(t)}{h} = \lim_{h \rightarrow 0} \frac{h}{h^2} \quad (2.7.38)$$

由上式可见, 不可能存在 $w'(t)$ 项, 所以维纳过程 $w(t)$ 是非均方可微的。

2.7.4 均方可积性

本小节讨论随机过程的均方黎曼积分,下小节讨论随机过程的斯蒂阶积分。

定义 2.7.5 考虑二阶随机过程 $\{x(t), t \in T\}$, 区间 $[a, b] \subset T$ 及其分点 $a = t_0 < t_1 < \dots < t_n = b$, 构造和式

$$I_n = \sum_{k=0}^{n-1} x(t_k)(t_{k+1} - t_k) \quad (2.7.39)$$

式中 t_k 的值可在 $[t_k, t_{k+1}]$ 中任意选择。如果 $\lim_{n \rightarrow \infty} \max_{k=0, \dots, n-1} (t_{k+1} - t_k) = 0$, 和式 I_n 依均方收敛于一个极限 I , 即

$$\lim_n E\{[I_n - I]^2\} = 0 \quad (2.7.40)$$

则称 $x(t)$ 是均方可积的, 式中极限 I 称为 $x(t)$ 在区间 $[a, b]$ 上的均方黎曼积分, 记为

$$I = \int_a^b x(t) dt \quad (2.7.41)$$

均方可积性判据 二阶随机过程 $\{x(t), t \in T\}$ 在区间 $[a, b]$ 上均方可积的充要条件是 $x(t)$ 的自相关函数 $C(s, t)$ 的二重积分 $\int_a^b \int_a^b C(s, t) dt ds$ 存在。

均方可积性的性质如下:

(1) 如果随机过程 $x(t)$ 和 $y(t)$ 在 $[a, b]$, $[a, c]$ 和 $[c, b]$ 上均方可积, α 和 β 为任意常数, 则有

$$\int_a^a x(t) dt = 0 \quad (2.7.42)$$

$$\int_a^b x(t) dt = \int_a^c x(t) dt + \int_c^b x(t) dt \quad (2.7.43)$$

$$\int_a^b [\alpha x(t) + \beta y(t)] dt = \alpha \int_a^b x(t) dt + \beta \int_a^b y(t) dt \quad (2.7.44)$$

(2) 如果随机过程 $x(t)$ 在 $[a, b]$ 上均方可积, 则

$$E \int_a^b x(t) dt = \int_a^b E\{x(t)\} dt \quad (2.7.45)$$

$$E \left| \int_a^b x(t) dt \right|^2 = M(b-a)^2 \quad (2.7.46)$$

式中:

$$M = \max_{a \leq t \leq b} E\{x(t)^2\} \quad (2.7.47)$$

$$E \int_a^b x(t) dt \int_a^b x(t) dt = \int_a^b \int_a^b C(s, t) ds dt \quad (2.7.48)$$

(3) 如果随机过程 $x(t)$ 在 $[a, b]$ 上均方可微, 且其导数 $\dot{x}(t)$ 在 $[a, b]$ 上连续, 则

$$\int_a^b \dot{x}(t) dt = x(b) - x(a) \quad (2.7.49)$$

(4) 如果随机过程 $x(t)$ 在 $[a, b]$ 上均方连续, $f(t, s)$ 是 $[a, b] \times [a, b]$ 上的一般二元连续函数, 则

$$\int_a^b \int_a^r f(r, s) x(s) ds dr = \int_a^b \int_s^b f(r, s) dr x(s) ds \quad (2.7.50)$$

(5) 如果随机过程 $x(t)$ 在 $[a, b]$ 上均方连续, $f(t, s)$ 和 $\frac{f(t, s)}{t}$ 是 $[a, b] \times [a, b]$ 上的一般二元连续函数, 设

$$y(t) = \int_a^t f(t, s)x(s) \, ds \qquad (t \in [a, b]) \tag{2.7.51}$$

则

$$y(t) = f(t, t)x(t) + \int_a^t \frac{f(t, s)}{t}x(s) \, ds \tag{2.7.52}$$

例 2.7.3 试分析维纳过程 $w(t)$ 在 $[0, u]$ 上是否均方可积。

解 已知维纳过程 $w(t)$ 的自相关函数为

$$C(s, t) = \frac{1}{2} \min(s, t) \tag{2.7.53}$$

利用均方可积性判据, 计算 $C(s, t)$ 的二重积分, 得到

$$\begin{aligned} \int_0^u \int_0^u C(s, t) \, dt \, ds &= \int_0^u \int_0^u \frac{1}{2} \min(s, t) \, dt \, ds \\ &= \frac{1}{2} \int_0^u ds \int_0^s t \, dt + \int_0^u s \, ds = \frac{u^3}{3} \end{aligned} \tag{2.7.54}$$

$C(s, t)$ 的二重积分存在, 所以维纳过程 $w(t)$ 是均方可积函数。

2.7.5 随机积分

考虑随机状态方程

$$dx(t) = f(x, t)dt + g(x, t)dw(t) \tag{2.7.55}$$

$$x(t_0) = x_0 \tag{2.7.56}$$

式中: $x(t)$ 为状态; $w(t)$ 为维纳过程; $f(x, t)$ 和 $g(x, t)$ 是非线性函数; x_0 为初值。式 (2.7.55) 的解为

$$x(t) = x_0 + \int_{t_0}^t f(x, s) \, ds + \int_{t_0}^t g(x, s) \, dw(s) \tag{2.7.57}$$

上式右边第二项是随机过程的黎曼积分, 第三项则是随机过程的斯蒂阶积分。下面对 $g(x, t)$ 分两种情况讨论。

(1) $g(x, t)$ 是确定性函数 $g(t)$ 情况

这时, 把积分定义为

$$I = \int_{t_0}^t g(t)dw(t) = \lim_{n \rightarrow \infty} \sum_{k=1}^n g(t_k)[w(t_k) - w(t_{k-1})] \tag{2.7.58}$$

式中: $t_k \in [t_k, t_k]$; $t_n = t_0$ 。可直接求解 I 的均值和相关函数。另外, 还可利用如下公式:

$$\int_{t_0}^t g(s)dw(s) = g(t)w(t) - g(t_0)w(t_0) - \int_{t_0}^t w(s)dg(s) \tag{2.7.59}$$

(2) 当 $g(x, t)$ 是随机过程时的情况

当 $g(x, t)$ 是随机过程, 且为与 $w(t)$ 有关的随机过程时, 积分值将有赖于 t_k 的选择, 见式 (2.7.58)。下面以积分

$$I = \int_a^b w(t)dw(t) \tag{2.7.60}$$

为例来说明这一问题。设 $w(t)$ 是均值为零和增量方差为 1 的维纳过程, 即 $dw(t) \sim N(0,$

dt), 并令 $a = t_0 < t_1 < \dots < t_n = b$, $n = \max_{1 \leq k \leq n} (t_k - t_{k-1})$ 。

首先计算 $k = t_{k-1}$ 时的情况:

$$\begin{aligned} I_0 &= \sum_{k=1}^n w(t_{k-1}) [w(t_k) - w(t_{k-1})] \\ &= - [w^2(t_0) - w(t_0)w(t_1) + w^2(t_1) - w(t_1)w(t_2) + \dots - w(t_{n-1})w(t_n)] \\ &= - \frac{1}{2}w^2(t_0) + \frac{1}{2}w^2(t_0) - w(t_0)w(t_1) + \frac{1}{2}w^2(t_1) + \dots + \frac{1}{2}w(t_n) - \frac{1}{2}w^2(t_n) \\ &= \frac{1}{2}[w^2(b) - w^2(a)] - \frac{1}{2} \sum_{k=1}^n [w(t_k) - w(t_{k-1})]^2 \end{aligned} \tag{2.7.61}$$

考虑在均方意义上

$$\lim_{n \rightarrow 0} \sum_{k=1}^n [w(t_k) - w(t_{k-1})]^2 = b - a \tag{2.7.62}$$

把上式代入式(2.7.61), 当 $n \rightarrow 0$ 时, 得到

$$\begin{aligned} I_0 &= \int_a^b w(t)dw(t) \\ &= \frac{1}{2}[w^2(b) - w^2(a)] - \frac{1}{2}(b - a) \end{aligned} \tag{2.7.63}$$

其次计算 $k = t_k$ 时的情况:

$$I_1 = \sum_{k=1}^n w(t_k) [w(t_k) - w(t_{k-1})] \tag{2.7.64}$$

由上式减去式(2.7.61), 得到

$$I_1 - I_0 = \sum_{k=1}^n [w(t_k) - w(t_{k-1})]^2 \tag{2.7.65}$$

而

$$I_1 - I_0 = \lim_{n \rightarrow 0} (I_1 - I_0) = b - a \tag{2.7.66}$$

把式(2.7.63)代入上式, 得到

$$I_1 = \int_a^b w(t)dw(t) = \frac{1}{2}[w^2(b) - w^2(a)] + \frac{1}{2}[b - a] \tag{2.7.67}$$

由上述计算得到如下结论:

(1) 比较式(2.7.67)和(2.7.63)可知, I_1 和 I_0 结果不同, 随机积分的结果有赖于 k 的选择。

(2) 若 $w(t)$ 是确定性函数, 有

$$\int_a^b w(t)dw(t) = \frac{1}{2} \int_a^b dw^2(t) = \frac{1}{2}[w^2(b) - w^2(a)] \tag{2.7.68}$$

把上式结果与式(2.7.63), (2.7.67)进行比较可知, 随机积分与确定性函数积分是不同的。

通过设定 I 可把上述随机积分统一起来,

$$I = (1 - \alpha)I_0 + \alpha I_1 \tag{2.7.69}$$

式中: $0 \leq \alpha \leq 1$, 选择不同 α 可得到不同的随机积分结果。特别是 $\alpha = 0$ 时, 得到式(2.7.63), 称为伊藤(Ito)积分。 $\alpha = 0.5$ 时, 随机积分结果与确定性函数积分结果相同, 称

为斯特拉托诺维奇(Stratonovich) 积分。

随机积分具有分部积分公式。设 $g(t)$ 和 $y(t)$ 是随机函数, 如果 $\int_0^t g(s) dy(s)$ 存在, 则有

$$\begin{aligned} g(t)y(t) - g(0)y(0) &= I_1(g, dy) + I_0(y, dg) \\ &= I_1(y, dg) + I_0(g, dy) \end{aligned} \tag{2.7.70}$$

式中:

$$I_0(g, dy) = \lim_{n \rightarrow \infty} \sum_{k=1}^n g(t_{k-1})[y(t_k) - y(t_{k-1})] \tag{2.7.71}$$

$$I_1(g, dy) = \lim_{n \rightarrow \infty} \sum_{k=1}^n g(t_k)[y(t_k) - y(t_{k-1})] \tag{2.7.72}$$

以后最常用的是伊藤积分, 下面介绍它的几个性质:

(1) 取均值和求积分符号可调换次序, 其他积分不满足这一性质。

(2) 如果伊藤积分 $\int_a^b x(t)dw(t)$ 和 $\int_a^b y(t)dw(t)$ 存在, 和 α, β 是任意常数, 则

$$\int_a^b [\alpha x(t) + \beta y(t)]dw(t) = \alpha \int_a^b x(t)dw(t) + \beta \int_a^b y(t)dw(t) \tag{2.7.73}$$

(3) 如果 $a < c < b$, 则

$$\int_a^b x(t)dw(t) = \int_a^c x(t)dw(t) + \int_c^b x(t)dw(t) \tag{2.7.74}$$

由上述几个性质可知, 伊藤积分的性质与一般确定性函数积分的性质基本相同。

2.7.6 伊藤(Ito)微分规则

求解随机微分方程基本按伊藤微分规则进行。考虑随机状态方程

$$dx(t) = f(x, t)dt + g(x, t)dw(t) \tag{2.7.75}$$

这就是要研究的随机微分方程, 式中给定增量维纳过程 $dw(t) \sim N(0, Rdt)$ 。这里利用伊藤微分规则(伊藤公式)来求解这类方程。

定理 2.7.1 伊藤(Ito)微分规则 考虑随机过程 $x(t)$, 它满足随机微分方程式(2.7.75), 设随机过程 $y = y(x, t)$ 对 t 连续可微和对 x 二次连续可微, 则有伊藤微分规则:

$$dy = y_t dt + y_x^T f + \frac{1}{2} t_{xx} y_{xx} g R g^T dt + y_x^T g dw \tag{2.7.76}$$

式中: $y_t = \frac{dy}{dt}, y_x = \frac{dy}{dx}, y_{xx} = \frac{d^2 y}{dx^2}$; 并且 $dt^2 = 0, dw(t)dt = 0, dw(t)^2 = Rdt$

证明 把 $y(x, t)$ 展开为泰勒级数

$$y = y_t t + y_x^T x + \frac{1}{2} y_{tt} (t)^2 + \frac{1}{2} (x)^T y_{xx} x + o(t) \tag{2.7.77}$$

把式(2.7.75)代入上式右边第二项, 得到

$$y_x^T x = y_x^T f dt + y_x^T g dw \tag{2.7.78}$$

因为 $dt^2 = 0$, 式(2.7.77)右边第三项趋于零。把式(2.7.75)代入式(2.7.77)右边第四项, 得到

$$\begin{aligned}
(x)^T y_{xx} x &= [f \ t + \ g \ w]^T y_{xx} [f \ t + \ g \ w] \\
&= [f \ t]^T y_{xx} f \ t + [f \ t]^T y_{xx} g \ w \\
&\quad + [g \ w]^T y_{xx} f \ t + [g \ w]^T y_{xx} g \ w \\
&= w^T g^T y_{xx} g \ w + o(t) \\
&= \text{tr} y_{xx} g \ w \ w^T g^T + o(t) \\
&= \text{tr} y_{xx} g R g^T t + o(t)
\end{aligned} \tag{2.7.79}$$

上式第二等式右边前三项都趋于零, 第三等式右边为二次型(为标量), 可加取迹符号 tr , 再把后边各量顺序调换次序就得到第四项。关于“距阵的迹”的概念, 请参阅 6.2.3 小节。

把式(2.7.78)和(2.7.79)代入式(2.7.77), 并取极限, 就得到式(2.7.76)。

下面举两个例子说明伊藤微分规则的应用。

例 2.7.4 设

$$x(t) = w(t) \tag{2.7.80}$$

$$y(t) = [x(t)]^2 \tag{2.7.81}$$

式中: $dw(t) \sim N(0, dt)$ 。试用这个例子说明伊藤微分规则和伊藤积分的一致性。

解 把给定方程式(2.7.80)与式(2.7.75)对照可知, $f(x, t) = 0$, $g(x, t) = 1$ 。对给定方程式(2.7.81)使用伊藤微分规则式(2.7.76), 得到

$$dy = dt + 2w(t)dw(t) \tag{2.7.82}$$

对上式两边取积分, 注意到 $y(0) = w(0)^2 = 0$, 得到

$$y = t + \int_0^t w(t)dw(t) \tag{2.7.83}$$

$$\int_0^t w(t)dw(t) = \frac{1}{2}w(t)^2 - \frac{1}{2}t \tag{2.7.84}$$

把上式与式(2.7.63)进行比较可知, 用伊藤微分规则求得的结果与用伊藤积分求得的结果是相同的, 说明两者是统一的。

例 2.7.5 求解随机微分方程

$$dx(t) = -0.5x(t)dt + x(t)dw(t) \tag{2.7.85}$$

$$x(0) = 1 \tag{2.7.86}$$

式中: $dw(t) \sim N(0, dt)$ 。

解 构造函数

$$y = \ln x(t) \tag{2.7.87}$$

由式(2.7.86)得到初值 $t=0$ 时, $y=0$ 。对式(2.7.87)使用伊藤微分规则, 得到

$$dy = -\frac{1}{2}dt + dw(t) \tag{2.7.88}$$

$$y = -\frac{1}{2}t + w(t) \tag{2.7.89}$$

$$x(t) = \exp\{-\frac{1}{2}t + w(t)\} \tag{2.7.90}$$

若使用一般确定性函数的微分规则, 得到

$$x(t) = \exp\{-0.5t + w(t)\} \tag{2.7.91}$$

比较式(2.7.90)和(2.7.91)可知, 伊藤微分结果不同于一般的微分结果。

综上所述, 伊藤微分规则有其特点, 形式上不同于一般的微分规则。但考察其推导过

程可知, 推导中使用的是一般的微分规则, 要注意的是维纳过程 $w(t)$ 的特点, 即 $w(t)^2$ 的单位是时间的单位, $w(t)$ 的单位是时间单位的 $1/2$ 次方。

习 题

2.1 设 I 是实线上的线段 $[0, 1]$, 且测度(概率) P 为均匀分布; 再设指标集 T 也在区间 $[0, 1]$ 上。考虑由

$$\begin{aligned} x(t, \omega) &= 0 && (\text{对所有 } t \text{ 和 } \omega) \\ y(t, \omega) &= \begin{cases} 1 & (t = \omega) \\ 0 & (\text{其他}) \end{cases} \end{aligned}$$

定义的随机过程 $\{x(t), t \in T\}$ 和 $\{y(t), t \in T\}$, 试证明上述两个随机过程有同样的有限维分布, 并且证明

$$\begin{aligned} P\{\omega, x(t, \omega) < 0.5, \text{ 对所有 } t\} &= 1 \\ P\{\omega, y(t, \omega) < 0.5, \text{ 对所有 } t\} &= 0 \end{aligned}$$

2.2 考虑一阶平移平均过程

$$x(k) = e(k) + 0.5e(k - 1)$$

式中: $\{e(k), k = \dots, -2, -1, 0, 1, 2, \dots\}$ 是独立同分布高斯 $N(0, 1)$ 随机变量序列。试确定 $x(1)$ 和 $x(k)$ 的协方差。

2.3 考虑自回归过程

$$x(k) + ax(k - 1) = ce(k)$$

式中: $|a| < 1$; 而 $\{e(k), k = \dots, -2, -1, 0, 1, 2, \dots\}$ 是独立同分布高斯 $N(0, 1)$ 随机变量序列。试确定 $x(k)$ 和 $x(1)$ 的协方差, 再求 $t_0 \rightarrow -\infty$ 的情况。

2.4 设 $\{x(t), t \in T\}$ 是一个均值为零的高斯过程, 试证明

$$E\{x(t_1)x(t_2)x(t_3)x(t_4)\} = r(t_1, t_2)r(t_3, t_4) + r(t_1, t_3)r(t_2, t_4) + r(t_1, t_4)r(t_2, t_3)$$

式中: $r(t_i, t_j)$ 为 $x(t_i)$ 和 $x(t_j)$ 的协方差函数。

提示: 可利用特征函数进行证明。引入

$$\phi(u) = E\{\exp(-i x^T u)\}$$

式中: $\phi(u)$ 称为 n 维随机过程 $x(t)$ 的特征函数; $i = \sqrt{-1}$; 而 u 是与 $x(t)$ 同维的列向量

$$u^T = [u_1, u_2, \dots, u_n]$$

2.5 有一对讲系统总机, 为六个病房播送音乐, 任一病房开关接通并在任何时刻消耗功率的概率为 0.4。当开关接通时, 一个病房消耗功率为 0.5 W。把总机输送的功率作为随机变量, 求它的密度函数和分布函数。当输送功率大于 2 W 时, 总机放大器过载, 求过载的概率。

2.6 一电压随机变量 x 的分布为高斯 $N(0, 9)$ 的, 电压 x 加到平方律全波检波器上, 如果检波器的传输特性为 $y = 5x^2$, 求输出电压 y 的均值。

2.7 若随机过程 $x(t)$ 和 $y(t)$ 的联合密度函数为

$$f_{xy}(x, y) = u(x)u(y)e^{-x(y+1)}$$

求条件密度函数 $f_x(x|y)$ 。

2.8 设 $x(k)$ 是独立同分布高斯 $N(0, R_x)$ 随机变量序列, 现有 $y_t = \frac{1}{t} \sum_{i=1}^t x(t-i)$, $y_s = \frac{1}{s} \sum_{k=0}^{s-1} x(s-k)$, 令 $t=s$, 试求协方差函数 $R(y_t, y_s)$ 。

2.9 随机过程 $\{x(t), t \geq 0\}$ 具有独立增量, 试证明 $x(t)$ 是或不是平稳过程。

2.10 考虑一阶平移平均过程

$$x(k) = e(k) + ce(k-1)$$

式中: $\{e(k), k = \dots, -2, -1, 0, 1, 2, \dots\}$ 是独立同分布高斯 $N(0, 1)$ 随机变量序列。试问此过程是平稳的、高斯的、马尔可夫的、遍历性的、或是奇异的吗? 此过程有独立增量吗?

2.11 考虑一阶自回归过程

$$x(k+1) + ax(k) = e(k) \quad (k = k_0, k_0 + 1, \dots)$$

式中: $|a| < 1$; $\{e(k)\}$ 是独立同分布高斯 $N(0, 1)$ 随机变量序列; 初始状态 $x(k_0)$ 是高斯 $N(0, \sigma^2)$ 的; 还假设 $\{e(k)\}$ 与 $x(k_0)$ 相互独立。试问此过程是平稳的、高斯的、马尔可夫的、遍历性的、或奇异的吗? 此过程有独立增量吗?

2.12 考虑自回归过程

$$x(k+1) + ax(k) = e(k) \quad (k = k_0, k_0 + 1, \dots)$$

式中: $|a| < 1$; $\{e(k)\}$ 是独立同分布高斯 $N(0, 1)$ 随机变量序列; 初始状态 $x(k_0)$ 是高斯 $N(0, \sigma^2)$ 的; 再假设 $x(k_0)$ 和 $e(k_0)$ 是具有相关系数为 ρ 的联合高斯分布。试问此过程是马尔可夫过程吗?

2.13 考虑由

$$\frac{dx(t)}{dt} = 0$$

定义的奇异随机过程 $\{x(t), 0 \leq t < \infty\}$, 其初始状态 $x(0)$ 是高斯 $N(0, 1)$ 的, 试问此过程是遍历性的吗? 试给出一个基于 $x(t)$ 的量测来预测 $x(t+h)$ 的过程预测器。

2.14 设 $w(t)$ 是维纳过程, 其方差参数 $\sigma^2 = 3$, 试计算 $I = E \int_0^5 0.4dw(t) \int_1^6 0.2dw(t)$ 的值。

2.15 考虑维纳过程 $w(t)$, 其协方差函数 $r(t, t) = t$ 。(1) 试问 $E\{w(t)\}$ 等于什么? (2) 试求 $r(dw(t))$ 。(3) 当 $s=2$ 时, 试画出 $r(s, t)$ 对 t 的关系曲线。

2.16 考虑随机过程

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \quad (1)$$

式中初始状态是高斯的, 其均值为零, 协方差为

$$\text{cov}[x(0), x(0)] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (2)$$

此过程是遍历性的吗? 给出一个基于 $\{x_1(s), t_0 \leq s \leq t\}$ 的观测来预测 $x(t+h)$ 的过程预测器。

2.17 已知 $e(t)$ 是独立同分布高斯 $N(1, 10)$ 随机变量序列, $E\{t\} = \sum_{t=0}^{10} e(t)$, $E\{s\} =$

15

$e(s)$, 试计算协方差函数 $r[E\{t\}, E\{s\}]$ 。

$s=1$

2.18 讨论下列函数是否为(弱)平稳随机过程的协方差函数:

$r_1(\tau) = \text{常数}$

$r_2(\tau) = \cos \tau$

$r_3(\tau) = \begin{cases} 1 & (|\tau| \leq 1) \\ 0 & (|\tau| > 1) \end{cases}$

$r_4(\tau) = \begin{cases} 1 - |\tau| & (|\tau| \leq 1) \\ 0 & (|\tau| > 1) \end{cases}$

$r_5(\tau) = \frac{1}{1 + 2|\tau|}$

$r_6(\tau) = \begin{cases} 2 & (\tau = 0) \\ e^{-|\tau|} & (\tau \neq 0) \end{cases}$

2.19 设 $\{e(k), k = \dots, -2, -1, 0, 1, 2, \dots\}$ 是独立同分布高斯 $N(0, 1)$ 随机变量序列。考虑由

$$x(k) + ax(k-1) = e(k) + ce(k-1)$$

所定义的随机过程, 式中 $|a| < 1$, 试确定此过程的协方差函数。

答案:

$$\text{cov}[x(1), x(k)] = (-a)^{1-k} \frac{1 + c^2 - 2ac}{1 - a^2} - \frac{c}{a} \quad (1 > k)$$

$$\text{cov}[x(1), x(k)] = \frac{1 + c^2 - 2ac}{1 - a^2} \quad (1 = k)$$

2.20 随机电报信号是一种连续时间随机过程 $\{x(t), -\infty < t < \infty\}$ 。它具有下列特性: 若 x 的取值仅为 $+1$ 或 -1 , 则在区间 $(t, t+h)$ 上, x 改变其值的概率为 $h + o(h)$; 在长度为 t 的区间上, x 改变 n 次的概率为 $\frac{(t)^n}{n!} \exp(-t)$ 。试确定此过程的协方差函数。

2.21 设 L 是进行数学期望运算的线性算子, 试证明随机过程 $\{x(t), t \in T\}$ 在线性关系式 $Lx = 0$ 下是奇异(或纯确定性)过程的必要条件是此过程的协方差函数 $r(s, t)$ 有下列特性:

$$Lr(\tau, t) = 0 \quad (t = \text{常数})$$

$$Lr(s, \tau) = 0 \quad (s = \text{常数})$$

2.22 试证明

$$\text{cov}[Ax + a, By + b] = A\text{cov}[x, y]B^T$$

2.23 平稳随机过程有下列相关函数

$C_1(\tau) = e^{-a|\tau|}$

$C_2(\tau) = e^{-a^2 \tau^2}$

$C_3(\tau) = A + B \cos \omega_0 \tau$

$$C_4(\tau) = e^{-a|\tau|} \cos \tau$$

试确定相应的谱密度函数。

2.24 已知下列随机过程

$$x_1(k) = e(k) + ce(k-1)$$

$$x_2(k) + ax_2(k-1) = e(k-1)$$

$$x_3(k) + ax_3(k-1) = e(k) + ce(k-1)$$

的相关函数分别为

$$C_1(\tau) = \begin{cases} 1 + c^2 & (\tau = 0) \\ c & (|\tau| = 1) \\ 0 & (|\tau| > 1) \end{cases}$$

$$C_2(\tau) = \frac{1}{1-a^2} (-a)^{|\tau|}$$

$$C_3(\tau) = \begin{cases} \frac{1+c^2-2ac}{1-a^2} & (\tau = 0) \\ \frac{1+c^2-2ac}{1-a^2} - \frac{c}{a} (-a)^{|\tau|-1} & (|\tau| = 1) \end{cases}$$

式中: $\{e(k), k = \dots, -2, -1, 0, 1, 2, \dots\}$ 是独立同分布高斯 $N(0, 1)$ 随机变量序列; $a \neq 1$ 。试确定相应的谱密度。

2.25 考虑具有下列相关函数的平稳随机过程

$$C_1(\tau) = e^{-a|\tau|}$$

$$C_2(\tau) = \begin{cases} 2 & (\tau = 0) \\ e^{-|\tau|} & (\tau \neq 0) \end{cases}$$

这些过程依均方连续吗?

2.26 考虑具有下列相关函数的平稳随机过程

$$C_1(\tau) = e^{-a|\tau|}$$

$$C_2(\tau) = \frac{a}{a^2 + \tau^2}$$

$$C_3(\tau) = \sin a\tau$$

这些过程依均方可微吗?

2.27 考虑随机电报信号, 样本函数依概率 1 连续吗? 此过程是否依概率 1 可微或是依均方可微?

2.28 考虑相关函数为

$$C(\tau) = (1 + |\tau|)e^{-|\tau|}$$

的高斯平稳过程, 试证明此过程依均方可微, 再证明 $x(t)$ 和 $\frac{dx(t)}{dt}$ 互相独立。

2.29 设 $\{w(t), t \geq 0\}$ 是一个具有单位方差参数的维纳过程, 试证明此过程依均方可积。再确定积分过程

$$z(t) = \int_0^t w(s) ds$$

的协方差函数。

2.30 考虑具有下列性质的随机过程 $\{x(t), -\infty < t < \infty\}$: 假设样本函数的值仅为 $+1$ 或 -1 , $x(t)$ 在区间 $(t, t+h)$ 内的变换概率是 $h+o(h)$ 。试回答下列问题: (1) 样本函数是连续的吗? (2) 此随机过程依均方连续吗? (3) 此随机过程依概率 1 连续吗? (4) 此随机过程依均方可微吗? (5) 样本函数是可微的吗? (6) 此随机过程依均方可积吗? (7) 样本函数是可积的吗?

2.31 考虑维纳过程, 试回答下列问题: (1) 样本函数是连续的吗? (2) 此过程是平稳过程吗? (3) 此过程依均方连续吗? (4) 此过程依概率 1 连续吗? (5) 此过程依均方可微吗? (6) 样本函数是可微的吗? (7) 此过程依均方可积吗? (8) 样本函数是可积的吗?

2.32 用兰杰文方程(Langevin equation)

$$\frac{dv(t)}{dt} = -av(t) + e(t)$$

能相当精确地描述布朗运动, 式中 $v(t)$ 是粒子的速度, $\{e(t), t \in T\}$ 是均值和协方差分别为

$$\begin{aligned} E\{e(t)\} &= 0 \\ r(s-t) &= E\{e(s)e(t)\} \\ &= R_1(s-t) \end{aligned}$$

的连续时间白噪声。试确定速度的协方差函数 $R(t)$ 。

2.33 设 $\{w(t), t \in T\}$ 是一个具有单位方差参数的维纳过程, 取

$$w(t) = w(t+h) - w(t)$$

试证明下式是依概率 1 成立的:

$$\int_0^h (dw)^2 = \lim_{h \rightarrow 0} [w(t_i + 1) - w(t_i)]^2 = h$$

2.34 考虑积分

$$I = \int_0^t f(s) dy(s)$$

式中: $f(s)$ 是确定性函数, $\{y(t), t \in T\}$ 是均值为 $m(t)$ 和增量协方差为 dr 的独立增量过程。设 $y(0) = 0$, 再设 f, m 和 r 连续可微, 并用 \dot{f}, \dot{m} 和 \dot{r} 分别表示它们的导数。若用

$$I = f(t)y(t) - f(0)y(0) - \int_0^t f(s)y(s)ds$$

来定义积分 I , 试证明

$$\begin{aligned} E\{I\} &= \int_0^t f(s)m(s)ds \\ \text{var}I &= \int_0^t f^2(s)r(s)ds \end{aligned}$$

2.35 考虑函数

$$y(t) = e^{w(t)}$$

式中: $\{w(t), t \in T\}$ 是具有单位方差参数的维纳过程。试证明函数 y 满足随机微分方程

$$\begin{aligned} dy &= y \frac{1}{2}dt + dw \\ y(0) &= 1 \end{aligned}$$

2.36 考虑随机微分方程

$$dx = Ax dt + dv$$

式中: $\{v(t), t \in T\}$ 是协方差参数为 $R_1 dt$ 的维纳过程。试求函数

$$y(x, t) = x^T S(t) x$$

的微分 dy , 式中 $S(t)$ 是一对称阵函数。

2.37 设 $\{w(t), t \in T\}$ 是具有单位方差参数的维纳过程, 试证明随机微分方程

$$\begin{aligned} dx &= \overline{x} dw + \frac{1}{4} dt \\ x(0) &= 1 \end{aligned}$$

有解

$$x(t) = 1 + \frac{1}{2} w(t)^2$$

2.38 设 $\{w(t), t \in T\}$ 是具有单位方差参数的维纳过程, 试证明随机微分方程

$$\begin{aligned} dx_1 &= x_2 dw - \frac{1}{2} x_1 dt \\ dx_2 &= -x_1 dw - \frac{1}{2} x_2 dt \\ x_1(0) &= 0 \\ x_2(0) &= 1 \end{aligned}$$

有解

$$\begin{aligned} x_1 &= \sin w(t) \\ x_2 &= \cos w(t) \end{aligned}$$

2.39 求解初始条件为

$$\begin{aligned} x_1(0) &= 1 \\ x_2(0) &= 0 \end{aligned}$$

的随机微分方程

$$\begin{aligned} dx_1 &= x_2 dt \\ dx_2 &= dw \end{aligned}$$

式中: $\{w(t), t \in T\}$ 是具有单位方差参数的维纳过程。

第 3 章 随机系统的数学模型

3.1 引言

在上一章学习了随机过程基本知识的基础上,在这一章介绍随机系统的数学模型,后续章节则围绕这些数学模型讨论它们的状态估计和控制问题。

常用的离散时间随机系统的数学模型有两种,一种是受控自回归平移平均模型(CARMA),另一种是随机状态模型(方程)。常用的连续时间随机系统的数学模型是随机状态模型。这一章主要介绍这三种模型,讨论如何确定模型中有关量(状态、输出量)的值。连续时间随机系统还可以用高阶微分方程描述,在控制中应用较少,本书中不专门介绍。这几种模型是可以互相转换的,本书中也进行一些介绍。

本章还包括了输入量是随机过程的离散时间和连续时间系统的分析。

3.2 受控自回归平移平均模型

1. CARMA 模型

受控自回归平移平均模型(过程)英文缩写为 CARMA (controlled autoregressive moving average process) 或 ARMA. X, 实质上是高阶差分方程,表示式为

$$\begin{aligned} y(k) + a_1 y(k-1) + \dots + a_n y(k-n) \\ = b_0 u(k-d) + b_1 u(k-d-1) + \dots + b_m u(k-d-m) \\ + [e(k) + c_1 e(k-1) + \dots + c_l e(k-l)] \end{aligned} \quad (3.2.1)$$

引入单位后向移位算子

$$y(k-1) = q^{-1} y(k) \quad (3.2.2)$$

用上式运算式(3.2.1), 得到

$$A(q^{-1})y(k) = B(q^{-1})u(k-d) + C(q^{-1})e(k) \quad (3.2.3)$$

式中:

$$A(q^{-1}) = 1 + a_1 q^{-1} + a_2 q^{-2} + \dots + a_n q^{-n} \quad (3.2.4)$$

$$B(q^{-1}) = b_0 + b_1 q^{-1} + b_2 q^{-2} + \dots + b_m q^{-m} \quad (3.2.5)$$

$$C(q^{-1}) = 1 + c_1 q^{-1} + c_2 q^{-2} + \dots + c_l q^{-l} \quad (3.2.6)$$

$y(k)$ 表示系统输出; $u(k)$ 表示系统输入,它是控制量; d 表示延迟步数; $e(k)$ 表示噪声; $A(q^{-1})$ 和 $C(q^{-1})$ 分别为 n 阶和 l 阶首1多项式; $B(q^{-1})$ 为 m 阶多项式; δ 为常数。 $A(q^{-1})$, $B(q^{-1})$, $C(q^{-1})$ 和 δ 都是已知的。

2. 一般假设

对 CARMA 模型式(3.2.3), 一般做两点假设:

(1) 噪声 $e(k)$ 是独立同分布高斯 $N(0, 1)$ 白噪声。由式(3.2.3)可知, 因为存在噪声

$e(k), y(k)$ 也变成了随机过程, $e(k)$ 是高斯的, $y(k)$ 也是高斯的。

(2) 噪声 $e(k)$ 和 $y(1)$ 互相独立, 即满足

$$E\{e(k)y(1)\} = 0 \quad (k > 1) \tag{3.2.7}$$

3. 多种简化形式

CARMA 模型式(3.2.3)可简化为多种形式。当 $u(k) = 0$ 时, 式(3.2.3)简化为

$$A(q^{-1})y(k) = C(q^{-1})e(k) \tag{3.2.8}$$

称为自回归平移平均模型 ARMA($n, 1$)。当 $u(k) = 0$ 和 $l = 0$ 时, 式(3.2.3)简化为

$$A(q^{-1})y(k) = e(k) \tag{3.2.9}$$

称为自回归模型 ARMA($n, 0$) = AR。当 $u(k) = 0$ 和 $n = 0$ 时, 得到

$$y(k) = C(q^{-1})e(k) \tag{3.2.10}$$

称为平移平均模型 ARMA($0, 1$) = MA。当 $e(k) = 0$ 时, 得到

$$A(q^{-1})y(k) = B(q^{-1})u(k - d) \tag{3.2.11}$$

称为脉冲传递函数。

4. 系统分析

这里对 CARMA 系统的稳定性和平稳性进行分析, 下一章讨论系统的预测和控制问题。为便于分析, 举一个例子。

例 3.2.1 考虑自回归过程 ARMA($1, 0$)

$$y(t) + ay(t - 1) = e(t) \tag{3.2.12}$$

已知 $\{e(t), t = \dots, -2, -1, 0, 1, 2, \dots\}$ 是独立同分布高斯 $N(0, 1)$ 随机变量序列, 系数 $|a| < 1$, 给定初值 $y(t_0)$ 。试求自协方差函数 $r(s, t)$, 并考虑初值 t_0 的情况。

解 自协方差函数 $r(s, t)$ 的定义式为

$$r(s, t) = E\{[y(s) - E\{y(s)\}][y(t) - E\{y(t)\}]\} \tag{3.2.13}$$

式中 $y(s)$ 由式(3.2.12)得到:

$$y(s) = (-a)^{s-t_0}y(t_0) + \sum_{l=1}^{s-t_0} (-a)^{s-t_0-l}e(t_0 + l) \tag{3.2.14}$$

因此有

$$E\{y(s)\} = (-a)^{s-t_0}E\{y(t_0)\} = (-a)^{s-t_0}y(t_0) \tag{3.2.15}$$

$$y(s) - E\{y(s)\} = \sum_{l=1}^{s-t_0} (-a)^{s-t_0-l}e(t_0 + l) \tag{3.2.16}$$

同样, 可得到

$$y(t) - E\{y(t)\} = \sum_{k=1}^{t-t_0} (-a)^{t-t_0-k}e(t_0 + k) \tag{3.2.17}$$

把式(3.2.16)和(3.2.17)代入式(3.2.13), 并设 $s \geq t$, 得到

$$r(s, t) = \sum_{l=1}^{s-t_0} \sum_{k=1}^{t-t_0} (-a)^{s+t-2t_0-l-k}E\{e(t_0 + l)e(t_0 + k)\} \tag{3.2.18}$$

注意到

$$E\{e(t_0 + l)e(t_0 + k)\} = \begin{cases} 1 & (l = k) \\ 0 & (l \neq k) \end{cases} \tag{3.2.19}$$

得到

$$\begin{aligned} r(s, t) &= (-a)^{s+t-2t_0} \sum_{k=1}^{t-t_0} (-a)^{-2k} \\ &= \frac{(-a)^{s-t}}{1-a^2} - \frac{(-a)^{s+t-2t_0}}{1-a^2} \end{aligned} \tag{3.2.20}$$

当 $t_0 \rightarrow \infty$ 时, 并有 $|a| < 1$, 得到定态解

$$r(s, t) = \frac{(-a)^{s-t}}{1-a^2} \tag{3.2.21}$$

(1) 系统的稳定性

重写系统方程

$$A(q^{-1})y(k) = B(q^{-1})u(k-d) + C(q^{-1})e(k) \tag{3.2.22}$$

很明显, 只有多项式 $A(q^{-1})$ 的全部零点都在单位圆内时, 式(3.2.22)表述的系统才可能稳定。在这里, $C(q^{-1})e(k)$ 是有色噪声, 一般情况下, 自身是稳定的, 也就是多项式 $C(q^{-1})$ 的全部零点都在单位圆内。若 $C(q^{-1})$ 在单位圆外有零点, 则表示系统噪声随时间增加越来越大, 这是不可能的。对多项式 $B(q^{-1})$, 也要求它的全部零点在单位圆内, 否则系统也是不稳定的, 在下一章将专门分析这个问题。

(2) 系统的平稳性

分析例 3.2.1 结果, 当 $|a| < 1$, 即 $A(q^{-1})$ 零点在单位圆内时, 若 $t_0 \rightarrow \infty$, 由式(3.2.15)得到输出均值, 由式(3.2.21)得到自协方差函数定态解, 分别为

$$E\{y(s)\} = 0 \tag{3.2.23}$$

$$r(s, t) = \frac{(-a)^{|s-t|}}{1-a^2} \tag{3.2.24}$$

由上两式看出, 系统输出均值为零(常数), 自协方差函数是时间差 $|s-t|$ 的函数, 与单个时间 s 和 t 无关, 所以 $y(s)$ 是平稳过程。将这种情况推广到高阶 ARMA($n, 1$) 也是适用的, 即 ARMA($m, 1$) 是平稳过程。

CARMA 模型不仅在随机控制中, 而且在生产过程、经济、生物、医学、天文物理和地震勘探等方面都得到了广泛的应用。

3.3 离散时间随机状态模型

3.3.1 随机状态模型和假设

离散时间随机状态模型(方程)一般表示为

$$x(k+1) = f[x(k), k] + u[x(k), k] + [x(k), k] \tag{3.3.1}$$

$$y(k) = g[x(k), k] + [x(k), k] \tag{3.3.2}$$

对上二式进行线性化, 得到

$$x(k+1) = \Phi(k+1, k)x(k) + u(k) + \omega(k) \tag{3.3.3}$$

$$y(k) = H(k)x(k) + \nu(k) \tag{3.3.4}$$

式(3.3.1)和(3.3.3)称为对象方程, 式(3.3.2)和(3.3.4)称为量测(输出)方程。本节只讨

论式(3.3.3)中没有控制项的简单情况,这时式(3.3.3)变为

$$x(k+1) = (k+1, k)x(k) + (k) \tag{3.3.5}$$

其他情况留待第5章和第6章中介绍。

式(3.3.5)称为离散时间线性随机状态模型(方程),其中 $x(k)$ 为 $n \times 1$ 状态向量, (k) 为 $n \times 1$ 噪声向量, $(k+1, k)$ 为 $n \times n$ 状态转移阵。

和确定性状态模型相同,状态转移阵 (k_2, k_1) 具有下列性质:

$$(1) \quad (k, k) = I \tag{3.3.6}$$

$$(2) \quad (k_3, k_2) (k_2, k_1) = (k_3, k_1) \tag{3.3.7}$$

$$(3) \quad (k_2, k_1)^{-1} = (k_1, k_2) \tag{3.3.8}$$

状态转移阵 (k_2, k_1) 还有其他一些性质,将在下一节介绍。

对研究的模型式(3.3.5),一般做下列假设:

(1) 假设噪声 (k) 为独立同分布高斯 $N(0, R_1(k))$ 向量白噪声,也就是

$$E\{(k)\} = 0 \tag{3.3.9}$$

$$r(l, k) = E\{(l)(k)\} = \begin{matrix} R_1(k) & (l = k) \\ 0 & (l \neq k) \end{matrix} \tag{3.3.10}$$

式中: $R_1(k)$ 一般为常阵,也可以是时变的。

(2) 假设状态初值 $x(k_0)$ 为高斯 $N(m_0, R_0)$ 向量,由于 $x(k_0)$ 和 (k) 都是高斯的,由式(3.3.5)可知, $x(k)$ 可写为 $x(k_0)$ 和 (k) 的线性组合,所以 $x(k)$ 也是高斯的。

(3) 假设 (k) 和 $x(k)$ 互相独立,即满足

$$\text{cov}[(l), x(k)] = 0 \quad (l \neq k) \tag{3.3.11}$$

3.3.2 模型的解

已知状态向量 $x(k)$ 是高斯过程,它完全由均值和协方差函数确定。所谓求模型的解就是求解 $x(k)$ 的均值和协方差函数。

1. 一般情况

(1) 求均值

已知状态方程式(3.3.5),只要对式(3.3.5)两边取均值,并考虑到噪声的均值为零,就能求出状态均值

$$\begin{aligned} m_x(k+1) &= E\{(k+1, k)x(k) + (k)\} \\ &= (k+1, k)m_x(k) \end{aligned} \tag{3.3.12}$$

考虑到初始条件

$$m_x(k_0) = m_0 \tag{3.3.13}$$

则得到

$$m_x(k) = (k, k_0)m_0 \tag{3.3.14}$$

(2) 求协方差函数 $R(l, k)$

设 $1 \leq k$, 反复使用已知状态方程式(3.3.5)和状态转移阵性质(2),并利用式(3.3.14),得到

$$x(l) = (l, l-1)x(l-1) + (l-1)$$

$$= (1, k)x(k) + \sum_{i=k+1}^l (1, i)(i-1) \quad (3.3.15)$$

$$m_x(1) = (1, k)m_x(k) \quad (3.3.16)$$

把式(3.3.15)和(3.3.16)代入协方差函数定义式,并稍加运算,得到

$$\begin{aligned} R(1, k) &= E\{[x(1) - m_x(1)][x(k) - m_x(k)]^T\} \\ &= E\{[(1, k)(x(k) - m_x(k)) + \sum_{i=k+1}^l (1, i)(i-1)][x(k) - m_x(k)]^T\} \\ &= (1, k)P(k) \end{aligned} \quad (3.3.17)$$

式中 $P(k)$ 表示 $x(k)$ 的方差阵

$$P(k) = E\{[x(k) - m_x(k)][x(k) - m_x(k)]^T\} \quad (3.3.18)$$

从式(3.3.17)第二式到第三式利用了 $[(k), (k+1), \dots, (l-1)]$ 与 $[x(k) - m_x(k)]$ 互相独立的性质。

通过简单计算,容易得到 $P(k+1)$ 的递推方程:

$$\begin{aligned} P(k+1) &= E\{[x(k+1) - m_x(k+1)][x(k+1) - m_x(k+1)]^T\} \\ &= E\{[(k+1, k)(x(k) - m_x(k)) + \sum_{i=k+1}^l (k+1, i)(i-1)][x(k) - m_x(k)]^T\} \\ &= (k+1, k)P(k) + R_1(k) \end{aligned} \quad (3.3.19)$$

上式可解释为: 第一项表示方差阵是如何转移的, 而第二项表示由噪声引起的方差增量。

由上述推导可知, 已知方差阵初值 $P(k_0) = R_0$ 和 $R_1(i) (i = k_0, k_1, \dots)$ 可用式(3.3.19)递推得到 $P(k)$, 再用式(3.3.17)就能求出 $R(1, k)$ 。

2. $(k+1, k)$ 和 $R_1(k)$ 为常阵的情况

$(k+1, k)$ 和 $R_1(k)$ 为常阵, 即

$$(k+1, k) = \quad (3.3.20)$$

$$R_1(k) = R_1 \quad (3.3.21)$$

是大量存在的情况。此时, 均值、方差阵和协方差阵分别为

$$\begin{aligned} m_x(k) &= (k, k_0)m_0 \\ &= \Phi^{k-k_0}m_0 \end{aligned} \quad (3.3.22)$$

$$\begin{aligned} P(k) &= P(k-1)^T + R_1 \\ &= [P(k-2)^T + R_1]^T + R_1 \\ &= \Phi^{k-k_0}R_0(\Phi^T)^{k-k_0} + \sum_{i=0}^{k-k_0-1} \Phi^i R_1 (\Phi^T)^i \end{aligned} \quad (3.3.23)$$

$$\begin{aligned} R(1, k) &= (1, k)P(k) \\ &= \Phi^{1-k}P(k) \end{aligned} \quad (3.3.24)$$

如果状态转移阵 Φ 的所有特征值都小于 1, 当 $k \rightarrow \infty$ 时, $P(k)$ 收敛, 有稳态值

$$P = \lim_{k \rightarrow \infty} P(k) = P \quad (3.3.25)$$

此稳态值容易由式(3.3.23)求出:

$$P = P^T + R_1 \quad (3.3.26)$$

例 3.3.1 一动力学系统由随机状态方程

$$x(k+1) = \begin{bmatrix} 1.5 & 1 \\ -0.7 & 0 \end{bmatrix} x(k) + \begin{bmatrix} 1.0 \\ 0.5 \end{bmatrix} w(k) \quad (3.3.27)$$

描述, 式中噪声 $\{w(k), k \in T\}$ 是独立同分布高斯 $N(0, 1)$ 随机变量序列, $w(k)$ 与 $x(k)$ 互相独立。试确定状态变量 $x(k)$ 的定态分布的协方差。

解 先验证给定动力学系统是否稳定, 再求出定态方差 P , 最后确定定态协方差 $R(1, k)$ 。

(1) 验证系统的稳定性

稳定的条件是转移矩阵 A 的特征值 λ 小于 1。对给定系统, 不难求出

$$A = \begin{bmatrix} 1.5 & 1 \\ -0.7 & 0 \end{bmatrix} \quad (3.3.28)$$

$$\det(I - \lambda I) = 0 \quad (3.3.29)$$

$$\lambda = 0.75 \pm 0.37i \quad (3.3.30)$$

$$|\lambda| = 0.84 \quad (3.3.31)$$

系统特征值的模小于 1, 系统稳定, 存在定态方差和定态协方差。

(2) 求定态方差 P

利用公式(3.3.26)

$$P = AP + R \quad (3.3.32)$$

式中:

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \quad (3.3.33)$$

$$\begin{aligned} R &= E\{[w(k)][w(k)]^T\} \\ &= E\{w(k)w(k)^T\} \\ &= \begin{bmatrix} 1.0 & 0.5 \\ 0.5 & 0.25 \end{bmatrix} \end{aligned} \quad (3.3.34)$$

计算中利用了给定条件

$$E\{w(k)w(k)^T\} = I \quad (3.3.35)$$

$$= \begin{bmatrix} 1.0 & 0 \\ 0 & 0.5 \end{bmatrix} \quad (3.3.36)$$

把式(3.3.33), (3.3.28)和(3.3.34)代入式(3.3.32), 经简单计算, 得到 $P_{11} = 18.88, P_{12} = P_{21} = -11.37, P_{22} = 9.5$, 即

$$P = \begin{bmatrix} 18.88 & -11.37 \\ -11.37 & 9.5 \end{bmatrix} \quad (3.3.37)$$

(3) 确定定态协方差 $R(1, k)$

令 $1 \leq k$, 利用公式(3.3.24), 得到最后结果

$$\begin{aligned} R(1, k) &= A^{1-k}P \\ &= \begin{bmatrix} 1.5 & 1 \\ -0.7 & 0 \end{bmatrix}^{1-k} \begin{bmatrix} 18.88 & -11.37 \\ -11.37 & 9.5 \end{bmatrix} \end{aligned} \quad (3.3.38)$$

3.4 连续时间随机状态模型

3.4.1 随机状态模型和假设

连续时间随机状态模型(方程)一般表示为

$$\frac{dx(t)}{dt} = f[x(t), t] + u[x(t), t] + [x(t), t] \quad (3.4.1)$$

$$y(t) = h[x(t), t] + [x(t), t] \quad (3.4.2)$$

对上二式进行线性化, 得到

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t) + (t) \quad (3.4.3)$$

$$y(t) = Hx(t) + (t) \quad (3.4.4)$$

式(3.4.1)和(3.4.3)称为对象方程, 式(3.4.2)和(3.4.4)称为量测(输出)方程。本节只讨论线性化情况。

实际上, 式(3.4.3)是向量线性随机微分方程, 当然可以写为增量的形式, 即

$$\begin{aligned} dx(t) &= Ax(t)dt + Bu(t)dt + (t)dt \\ &= Ax(t)dt + Bu(t)dt + dw(t) \end{aligned} \quad (3.4.5)$$

式中:

$$dw(t) = (t)dt \quad (3.4.6)$$

相应地, 为研究问题方便, 为形式上一致, 把式(3.4.4)改写为式(3.4.5)同样的形式。为此, 引入向量变量

$$z(t) = \int_0^t y(s)ds \quad (3.4.7)$$

于是得到

$$\frac{dz(t)}{dt} = Hx(t) + (t) \quad (3.4.8)$$

$$\begin{aligned} dz(t) &= Hx(t)dt + (t)dt \\ &= Hx(t)dt + dv(t) \end{aligned} \quad (3.4.9)$$

式中:

$$dv(t) = (t)dt \quad (3.4.10)$$

为简化符号, 仍然把式(3.4.9)中的 $z(t)$ 写为 $y(t)$, 最后得到

$$dy(t) = Hx(t)dt + dv(t) \quad (3.4.11)$$

总之, 得到了连续时间线性随机状态模型的两种表示, 一种表示为式(3.4.3)和(3.4.4), 另一种表示为式(3.4.5)和(3.4.11)。

本节只讨论式(3.4.3)和(3.4.5)中没有控制项的简单情况, 即

$$\frac{dx(t)}{dt} = Ax(t) + (t) \quad (3.4.12)$$

$$dx(t) = Ax(t)dt + dw(t) \quad (3.4.13)$$

其他情况留待第5章和第6章中介绍。

式(3.4.12)和(3.4.13)称为连续时间线性随机状态模型(方程)。式中: $x(t)$ 为 $n \times 1$ 状态向量; A 为 $n \times n$ 系数矩阵, 可为时变或时不变的; (t) 和 $dw(t)$ 分别为 $n \times 1$ 噪声向量。

对研究的模型式(3.4.12)和(3.4.13)一般做下列假设:

(1) 假设 $w(t)$ 是连续时间高斯 $N(0, R_1(t))$ 白噪声, 那么由 2.6.3 小节可知, $dw(t)$ 是维纳过程 $N(0, R_1(t)dt)$, 其中 $R_1(t)$ 可以是时不变的, 或时变的。

(2) 假设状态初值 $x(t_0)$ 为高斯 $N(m_0, R_0)$ 随机向量。由于 $x(t_0)$, $w(t)$, $dw(t)$ 都是高斯的, 因此由它们的和得到的状态向量 $x(t)$ 也是高斯的。

(3) 假设 $w(t)$ 与 $x(t)$, $dw(t)$ 与 $x(t)$ 互相独立, 即满足

$$E\{w(s)x(t_0)\} = 0 \quad (s \neq t_0) \quad (3.4.14)$$

$$E\{dw(s)x(t_0)\} = 0 \quad (s \neq t_0) \quad (3.4.15)$$

3.4.2 模型的解

已知状态向量 $x(t)$ 是高斯过程, 它完全由均值和协方差函数确定。所谓求模型的解就是求解 $x(t)$ 的均值和协方差函数。

1. 求均值

根据线性系统理论, 式(3.4.12)或(3.4.13)的解为

$$x(t) = \Phi(t, t_0)x(t_0) + \int_{t_0}^t \Phi(t, s)dw(s) \quad (3.4.16)$$

式中 $\Phi(t, s)$ 就是状态转移阵。

在 3.3.1 小节, 介绍了 $\Phi(t, s)$ 的三个性质, 下面再给出几个比较有用的性质:

$$(1) \quad \frac{\partial}{\partial t} \Phi(t, s) = A(t) \Phi(t, s) \quad (3.4.17)$$

$$\frac{\partial}{\partial t} \Phi^T(t, s) = -\Phi^T(t, s)A^T(t) \quad (3.4.18)$$

$$(2) \quad \frac{\partial}{\partial s} \Phi(t, s) = -\Phi(t, s)A(s) \quad (3.4.19)$$

$$\frac{\partial}{\partial s} \Phi^T(t, s) = -A^T(s) \Phi^T(t, s) \quad (3.4.20)$$

$$(3) \quad \frac{\partial}{\partial t} \det \Phi(t, s) = [\text{tr} A(t)] \det \Phi(t, s) \quad (3.4.21)$$

(4) 若 $[-A^T(\cdot)]$ 的状态转移阵为 $\Phi(t, s)$, 则

$$\Phi(t, s) = [\Phi^T(t, s)]^{-1} \quad (3.4.22)$$

(5) 若 $A(t) = C$, C 为任意实数, 则

$$\Phi(t, s) = \exp[C(t-s)] \quad (t \geq s) \quad (3.4.23)$$

只要对式(3.4.16)两边取均值, 就得到状态向量 $x(t)$ 的均值 $m(t)$

$$m(t) = \Phi(t, t_0)m_0 \quad (3.4.24)$$

上式是 $x(t)$ 均值的一种表示形式。还可以把它表示为微分方程的形式。为此, 对式(3.4.24)两边取导数, 并利用式(3.4.17)和(3.4.24), 得到

$$\begin{aligned} \frac{dm(t)}{dt} &= \frac{\partial}{\partial t} \Phi(t, t_0)m_0 \\ &= A(t) \Phi(t, t_0)m_0 \\ &= A(t)m(t) \end{aligned} \quad (3.4.25)$$

其初值为 $m(t_0) = m_0$ 。

2. 确定协方差函数

令 $s \leq t$, 按协方差函数定义, 并利用式(3.4.16)进行计算, 得到协方差函数

$$\begin{aligned} R(s, t) &= E\{[x(s) - m(s)][x(t) - m(t)]^T\} \\ &= E\left[\left(\int_{t_0}^s (s, \tau)dw(\tau) + (s, t)(x(t) - m(t))\right)[x(t) - m(t)]^T\right] \\ &= (s, t)P(t) \end{aligned} \tag{3.4.26}$$

式中引入了 $x(t)$ 的方差阵 $P(t)$, 即

$$P(t) = E\{[x(t) - m(t)][x(t) - m(t)]^T\} \tag{3.4.27}$$

欲求 $R(s, t)$, 需先计算 $P(t)$ 。为此, 把式(3.4.16)和(3.4.24)代入方差阵定义式(3.4.27), 得

$$\begin{aligned} P(t) &= E\left\{\left[\int_{t_0}^t (t, s)dw(s) + (t, t_0)(x(t_0) - m_0)\right]\left[\int_{t_0}^t (t, s)dw(s) + (t, t_0)(x(t_0) - m_0)\right]^T\right\} \\ &= (t, t_0)E\{[x(t_0) - m_0][x(t_0) - m_0]^T\} + (t, t_0)(t, t_0)^T \\ &\quad + E\left[\int_{t_0}^t (t, s)dw(s) \int_{t_0}^t (t, s)dw(s)^T\right] \\ &= (t, t_0)R_0 + (t, t_0)^T(t, t_0) + \int_{t_0}^t (t, s)R_1(t, s)ds \end{aligned} \tag{3.4.28}$$

上式第二个等式除列出的两项外, 还有两项, 其值为零。为得到第三个等式, 利用了下列两个公式:

$$R_0 = E\{[x(t_0) - m_0][x(t_0) - m_0]^T\} \tag{3.4.29}$$

$$E\{dw(s)dw(s)^T\} = R_1(s)ds \tag{3.4.30}$$

尤其应注意式(3.4.30), 它把二重积分简化为一元函数积分。

式(3.4.28)是 $P(t)$ 的一种表示形式, 还有另一种表示形式, 即微分方程的形式。下面计算中, 除利用状态转移阵 $\Phi(t, t_0)$ 有关公式外, 还利用公式

$$\frac{d}{dt} \int_{t_0}^t F(t, s)ds = F(t, t) + \int_{t_0}^t \frac{d}{dt} F(t, s)ds \tag{3.4.31}$$

对式(3.4.28)两边取导数, 得到

$$\begin{aligned} \frac{dP(t)}{dt} &= \frac{d}{dt} (t, t_0)R_0 + (t, t_0)R_0 \frac{d}{dt} (t, t_0)^T \\ &\quad + (t, t)R_1(t) + \int_{t_0}^t \frac{d}{dt} (t, s)R_1(s) + (t, s)R_1(s) \frac{d}{dt} (t, s)^T ds \\ &= A(t)(t, t_0)R_0 + (t, t_0)R_0 A^T(t) \\ &\quad + R_1(t) + A(t) \int_{t_0}^t (t, s)R_1(s) + (t, s)R_1(s) A^T(t) ds \end{aligned}$$

$$= A(t)P(t) + P(t)A^T(t) + R_1(t) \quad (3.4.32)$$

初值为

$$P(t_0) = R_0 \quad (3.4.33)$$

式(3.4.32)称为里卡蒂方程(Riccati equation)。可通过里卡蒂方程计算状态向量 $x(t)$ 的方差阵 $P(t)$, 再由式(3.4.26)确定 $x(t)$ 的协方差函数 $R(s, t)$ 。

3.4.3 关于两种模型的讨论

上一节中给出了连续时间线性随机状态模型的两种表示式(3.4.12)和(3.4.13)。实际上, 式(3.4.12)是向量微分方程, 而式(3.4.13)是向量增量方程, 这两种表示是人们经常使用的。但由于两种表示中包含的噪声项不同, 式(3.4.12)中包含高斯白噪声 $\dot{w}(t)$, 式(3.4.13)中包含维纳过程 $dw(t)$, 使用效果是不同的。

1. 关于模型式(3.4.12)

随机状态模型式(3.4.12)中含有高斯白噪声项 $\dot{w}(t)$, 它的方差是 $\delta(t)$ 函数, 实际上是不存在的, 也就是说 $\frac{dx(t)}{dt}$ 是不存在的, 因此, 在使用模型式(3.4.12)时可能出现错误, 这一点要引起注意。

这里举一个例子, 说明在计算中使用 $\frac{dx(t)}{dt}$ 可能得到错误结果。利用方差阵 $P(t)$ 的定义式(3.4.27), 直接导出微分方程

$$\begin{aligned} \frac{dP(t)}{dt} &= E \frac{d[x(t) - m(t)]}{dt} [x(t) - m(t)]^T \\ &\quad + E [x(t) - m(t)] \frac{d[x(t) - m(t)]^T}{dt} \\ &= E \{ [A(t)(x(t) - m(t)) + \dot{w}(t)] [x(t) - m(t)]^T \} \\ &\quad + E \{ [x(t) - m(t)] [A(t)(x(t) - m(t)) + \dot{w}(t)]^T \} \\ &= A(t)P(t) + P(t)A^T(t) \end{aligned} \quad (3.4.34)$$

把式(3.4.34)与式(3.4.32)相比较可看到, 式(3.4.34)缺少一项 $R_1(t)$ 。这个错误是怎样造成的? 推导过程是正确的, 这是由于 $\frac{dx(t)}{dt}$ 不存在造成的。

2. 关于模型式(3.4.13)

随机状态模型式(3.4.13)中的噪声项是维纳过程 $dw(t)$, 它与高斯白噪声 $\dot{w}(t)$ 不同, 是存在的, 即增量 $dx(t)$ 是存在的, 因此, 直接利用式(3.4.13)进行计算不会出现错误。

仍以推导 $P(t)$ 的里卡蒂方程为例, 说明计算中使用增量 $dx(t)$ 不会出现任何错误。首先计算增量

$$\begin{aligned} \{[x(t) - m(t)][x(t) - m(t)]^T\} &= [(x(t) - m(t)) + (x(t) - m(t))][x(t) - m(t)]^T - [x(t) - m(t)][x(t) - m(t)]^T \\ &= [x(t) - m(t)][x(t) - m(t)]^T \\ &\quad + [x(t) - m(t)][x(t) - m(t)]^T \\ &\quad + [x(t) - m(t)][x(t) - m(t)]^T \\ &= [x(t) - m(t)][A(t)[x(t) - m(t)]^T t + w(t)]^T \end{aligned}$$

$$\begin{aligned}
& + [A(t)(x(t) - m(t)) \quad t + \quad w(t)][x(t) - m(t)]^T \\
& + [A(t)(x(t) - m(t)) \quad t + \quad w(t)][A(t)(x(t) - m(t)) \quad t + \quad w(t)]^T
\end{aligned} \tag{3.4.35}$$

其次,对上式两边取均值和极限,略掉 t 的高阶无穷小项,得到

$$\begin{aligned}
dP(t) &= \lim_0 E\{ [x(t) - m(t)][\dot{x}]^T \} \\
&= A(t)P(t)dt + P(t)A^T(t)dt + R_1(t)dt
\end{aligned} \tag{3.4.36}$$

把式(3.4.36)与式(3.4.32)进行比较看出,两式是相同的,都是正确的结果。

例 3.4.1 1905 年,爱因斯坦给出了布朗运动的解。可见粒子在均匀液体中运动,他假设可见粒子比液体分子自由路程的均值大,因此可见粒子的运动方程近似为

$$x(t) = v(t) \tag{3.4.37}$$

$$mv(t) = -cv(t) + f(t) \tag{3.4.38}$$

式中: $x(t)$ 表示粒子位置; $v(t)$ 表示粒子速度; m 表示粒子质量; c 表示斯托克斯粘滞力系数(常数); $f(t)$ 表示液体分子间碰撞产生的随机力。由于碰撞的时间均值极短, $f(t)$ 近似为高斯 $N(0, q(\cdot))$ 白噪声, q 为常数。设

$$E\{x(0)\} = E\{v(0)\} = E\{x^2(0)\} = E\{v^2(0)\} = E\{x(0)v(0)\} = 0 \tag{3.4.39}$$

试确定 $E\{v^2(t)\}$, $E\{v(t)x(t)\}$ 和 $E\{x^2(t)\}$ 的表示式。

解 对布朗运动进行分析,可见粒子向各个方向运动的概率、受力大小和运动速度的概率是相同的,因此有

$$E\{x(t)\} = 0 \tag{3.4.40}$$

$$E\{v(t)\} = 0 \tag{3.4.41}$$

所求量就是可见粒子运动速度 $v(t)$ 的方差、位置 $x(t)$ 的方差和 $v(t)$ 与 $x(t)$ 的协方差,分别表示为

$$p_1(t) = E\{v^2(t)\} \tag{3.4.42}$$

$$p_2(t) = E\{x^2(t)\} \tag{3.4.43}$$

$$p_{12}(t) = p_{21}(t) = E\{v(t)x(t)\} \tag{3.4.44}$$

也就是求方差阵

$$P(t) = \begin{pmatrix} p_1(t) & p_{12}(t) \\ p_{21}(t) & p_2(t) \end{pmatrix} \tag{3.4.45}$$

下面利用 $P(t)$ 的里卡蒂方程(3.4.32)求解。

为此,由给定方程式(3.4.37)和(3.4.38)列出系统状态方程

$$dv(t) = -\frac{c}{m}v(t)dt + \frac{1}{m}dw(t) \tag{3.4.46}$$

$$dx(t) = v(t)dt \tag{3.4.47}$$

$$\begin{aligned}
\begin{pmatrix} dv(t) \\ dx(t) \end{pmatrix} &= \begin{pmatrix} -\frac{c}{m} & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} v(t) \\ x(t) \end{pmatrix} dt + \begin{pmatrix} \frac{1}{m} \\ 0 \end{pmatrix} dw(t)
\end{aligned} \tag{3.4.48}$$

式中: $dw(t)$ 为维纳过程

$$dw(t) = f(t)dt \tag{3.4.49}$$

由式(3.4.48)和给定条件写出相应量

$$A = \begin{bmatrix} -\frac{c}{m} & 0 \\ 1 & 0 \end{bmatrix} \tag{3.4.50}$$

$$R_1(t)dt = E \begin{bmatrix} \frac{1}{m} & 0 \\ 0 & 0 \end{bmatrix} dw(t)dw(t) \begin{bmatrix} \frac{1}{m} & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \frac{q}{m^2} & 0 \\ 0 & 0 \end{bmatrix} dt \tag{3.4.51}$$

把式(3.4.45),(3.4.50)和(3.4.51)代入里卡蒂方程式(3.4.32),并通过运算,得到

$$\begin{aligned} p_1 &= -\frac{2c}{m}p_1 + \frac{q}{m^2} \\ p_{12} &= p_{21} = -\frac{c}{m}p_{12} + p_1 \\ p_2 &= 2p_{12} \end{aligned} \tag{3.4.52}$$

解上述方程,得到最后结果

$$\begin{aligned} p_1(t) &= \frac{q}{2cm} \left(1 - e^{-\frac{2c}{m}t} \right) \\ p_{12}(t) &= \frac{q}{c^2} \frac{1}{2} + \frac{1}{2} e^{-\frac{2c}{m}t} - e^{-\frac{c}{m}t} \\ p_2(t) &= \frac{q}{c^2} t + \frac{m}{2c} \left(1 - e^{-\frac{2c}{m}t} \right) - \frac{2m}{c} \left(1 - e^{-\frac{c}{m}t} \right) \end{aligned} \tag{3.4.53}$$

3.5 连续时间与离散时间随机状态模型之间的转换

前两节中,分别讨论了离散时间和连续时间线性随机状态模型,两种模型之间存在内在联系,可以互相转换。连续时间随机状态模型,通过离散化方法,可以转换为离散时间随机状态模型;离散时间随机状态模型,通过对时间极限化方法,可以转换为连续时间随机状态模型。在很多场合,连续时间过程的预测和控制是在数字计算机上实现的,这就需要在采样时刻描述状态向量,描述过程也会大大简化。一般说,描述离散时间过程的理论比描述连续时间过程的理论相对简单,把离散时间过程的理论和公式,经过极限化方法,推广到连续时间过程,是经常采用的方法。

下面分别介绍离散化和极限化两种方法。

3.5.1 连续时间随机状态模型的离散化

考虑连续时间线性随机状态模型

$$dx(t) = A(t)x(t)dt + dw(t) \tag{3.5.1}$$

$$dy(t) = H(t)x(t)dt + dv(t) \tag{3.5.2}$$

关于式(3.5.1),已在3.4节讨论,不重复。式(3.5.2)中: $y(t)$ 为 $n \times 1$ 量测向量; $H(t)$ 为 $n \times n$ 量测矩阵,可以是时变或时不变的; $dv(t)$ 为 $n \times 1$ 量测噪声向量。

假设 $dv(t)$ 为增量维纳过程 $N(0, R_2(t)dt)$,并与 $x(t), dw(t)$ 互相独立,即满足

$$E\{x(t_0)dv(t)\} = 0 \quad (s \neq t_0) \tag{3.5.3}$$

$$E\{dv(t)dw(s)\} = 0 \quad (3.5.4)$$

由于状态向量初值 $x(t_0)$ 和 $dv(t)$ 都是高斯过程, 因此由式(3.5.2)可知, $dy(t)$ 也是高斯过程。

为区分连续时间 t 和离散时间 k , 并把两者联系起来, 令 $t_k = k$ 。

按式(3.4.16)写出式(3.5.1)的解

$$x(t_{k+1}) = (t_{k+1}, t_k)x(t_k) + \int_{t_k}^{t_{k+1}} (t_{k+1}, t)dw(t) \quad (3.5.5)$$

令

$$(t_k) = \int_{t_k}^{t_{k+1}} (t_{k+1}, t)dw(t) \quad (3.5.6)$$

则得到

$$x(k+1) = (k+1, k)x(k) + (k) \quad (3.5.7)$$

对式(3.5.2)积分, 并把式(3.5.5)代入积分式, 得到

$$\begin{aligned} y(t_{k+1}) &= y(t_k) + \int_{t_k}^{t_{k+1}} dy(s) \\ &= y(t_k) + \int_{t_k}^{t_{k+1}} H(s) \left[(s, t_k)x(t_k) + \int_{t_k}^s (s, t)dw(t) \right] ds + \int_{t_k}^{t_{k+1}} dv(s) \\ &= y(t_k) + (t_{k+1}, t_k)x(t_k) + (t_k) \end{aligned} \quad (3.5.8)$$

式中:

$$(t_{k+1}, t_k) = \int_{t_k}^{t_{k+1}} H(s) \int_{t_k}^s (s, t)ds \quad (3.5.9)$$

$$\begin{aligned} (t_k) &= \int_{t_k}^{t_{k+1}} H(s) \int_{t_k}^s (s, t)dw(t)ds + \int_{t_k}^{t_{k+1}} dv(s) \\ &= \int_{t_k}^{t_{k+1}} \int_{t_k}^{t_{k+1}} H(s) (s, t)dsdw(t) + \int_{t_k}^{t_{k+1}} dv(s) \\ &= \int_{t_k}^{t_{k+1}} (t_{k+1}, t)dw(t) + \int_{t_k}^{t_{k+1}} dv(s) \end{aligned} \quad (3.5.10)$$

再令

$$z(t_k) = y(t_{k+1}) - y(t_k) \quad (3.5.11)$$

则由式(3.5.8)得到

$$\begin{aligned} z(t_k) &= y(t_{k+1}) - y(t_k) \\ &= (t_{k+1}, t_k)x(t_k) + (t_k) \end{aligned} \quad (3.5.12)$$

即为

$$\begin{aligned} z(k) &= y(k+1) - y(k) \\ &= (k+1, k)x(k) + (k) \end{aligned} \quad (3.5.13)$$

式(3.5.7)和(3.5.13)就是连续时间线性随机状态模型式(3.5.1)和(3.5.2)的离散化模型, 下面计算模型噪声向量 (k) 和量测噪声向量 (k) 的有关数字特征。由式(3.5.6)计算, (k) 的均值 $E\{(k)\}$ 和协方差函数 $R_1(k)$ 分别为

$$E\{(k)\} = 0 \quad (3.5.14)$$

$$R_1(k) = E\{(t_k)^T(t_k)\}$$

$$\begin{aligned}
&= E \int_{t_k}^{t_{k+1}} \int_{t_k}^{t_{k+1}} (t_{k+1}, t) dw(t) dw^T(s) \int_{t_k}^{t_{k+1}} (t_{k+1}, s) \\
&= \int_{t_k}^{t_{k+1}} (t_{k+1}, t) R_1(t) \int_{t_k}^{t_{k+1}} (t_{k+1}, t) dt
\end{aligned} \tag{3.5.15}$$

由式(3.5.10)计算, (k) 的均值 $E\{ (k) \}$ 和协方差函数 $R_2(k)$ 分别为

$$E\{ (k) \} = 0 \tag{3.5.16}$$

$$\begin{aligned}
R_2(k) &= E\{ (k) (k)^T \} \\
&= \int_{t_k}^{t_{k+1}} \int_{t_k}^{t_{k+1}} (t_{k+1}, t) E\{ dw(t) dw^T(s) \} \int_{t_k}^{t_{k+1}} (t_{k+1}, s) \\
&\quad + \int_{t_{k+1}}^{t_{k+1}} \int_{t_k}^{t_{k+1}} E\{ dv(t) dv^T(s) \} \\
&= \int_{t_k}^{t_{k+1}} (t_{k+1}, t) R_1(t) \int_{t_k}^{t_{k+1}} (t_{k+1}, t) dt + \int_{t_k}^{t_{k+1}} R_2(t) dt
\end{aligned} \tag{3.5.17}$$

由式(3.5.6)和(3.5.10)可知, (k) 和 (k) 中都含有 $dw(t)$, 因此, 两者是相关的, 互协方差函数 $R_{12}(k)$ 为

$$\begin{aligned}
R_{12}(k) &= E\{ (k) (k)^T \} \\
&= \int_{t_k}^{t_{k+1}} (t_{k+1}, t) R_1(t) \int_{t_k}^{t_{k+1}} (t_{k+1}, t) dt
\end{aligned} \tag{3.5.18}$$

现将上述结论用定理 3.5.1 表示。

定理 3.5.1 连续时间线性随机状态模型式(3.5.1)和(3.5.2)在采样时刻 $t_k = k$ 可离散化为离散时间线性随机状态模型式(3.5.7)和(3.5.13), 其中 $(k+1, k)$ 表示状态转移阵, 量测矩阵 $(k+1, k)$ 由式(3.5.9)定义。模型噪声向量 (k) 由式(3.5.6)表示, 其均值和协方差函数分别用式(3.5.14)和(3.5.15)表示。量测噪声向量 (k) 由式(3.5.10)表示, 其均值和协方差函数分别用式(3.5.16)和(3.5.17)表示。尽管 $dw(t)$ 和 $dv(t)$ 是互相独立的, (k) 和 (k) 却是相关的, 其协方差函数由式(3.5.18)表示。

例 3.5.1 考虑随机微分方程

$$dx(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x(t) dt + \begin{bmatrix} 0 \\ 1 \end{bmatrix} dw(t) \tag{3.5.19}$$

$$dy(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) dt + dv(t) \tag{3.5.20}$$

式中: $dw(t)$ 是增量维纳过程 $N(0, dt)$; $dv(t)$ 是增量维纳过程 $N(0, rdt)$; $x(t)$, $dw(t)$ 和 $dv(t)$ 两两互相独立。当采样区间为 h 时, 试确定它的采样型。

解 由给定方程和条件, 得到

$$\begin{aligned}
A &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} & H &= \begin{bmatrix} 1 & 0 \end{bmatrix}, \\
R_1(t) &= \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} & R_2(t) &= r
\end{aligned} \tag{3.5.21}$$

用式(3.5.21)中的数据计算状态转移阵 $(k+1, k)$ 和量测阵 $(k+1, k)$:

$$\begin{aligned}
(k+1, k) &= e^{Ah} \\
&= I + Ah
\end{aligned}$$

$$= \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} \quad (3.5.22)$$

$$\begin{aligned} (k+1, k) &= \int_0^h H(s) \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} ds \\ &= \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} + \frac{h^2}{2} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \end{aligned} \quad (3.5.23)$$

由式(3.5.22)和(3.5.23)列出离散化方程

$$x(k+1) = \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} x(k) + \begin{bmatrix} h \\ 0 \end{bmatrix} u(k) + \begin{bmatrix} w_1(k) \\ w_2(k) \end{bmatrix} \quad (3.5.24)$$

$$y(k+1) = y(k) + h \left[\frac{h}{2} x(k) + u(k) \right] + w_3(k) \quad (3.5.25)$$

其模型噪声 $w(k)$ 的均值为零, 协方差为

$$\begin{aligned} R_1(k) &= \int_0^h \begin{bmatrix} h & s \\ 0 & 1 \end{bmatrix} R_1(t) \begin{bmatrix} h & s \\ 0 & 1 \end{bmatrix}^T ds \\ &= \int_0^h \begin{bmatrix} 1 & h-s & 0 & 0 \\ 0 & 1 & h-s & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} ds \\ &= \begin{bmatrix} \frac{h^3}{3} & \frac{h^2}{2} \\ \frac{h^2}{2} & h \end{bmatrix} \end{aligned} \quad (3.5.26)$$

量测噪声 $w_3(k)$ 的均值为零, 协方差为

$$\begin{aligned} R_2(k) &= \int_0^h \begin{bmatrix} h & s \\ 0 & 1 \end{bmatrix} R_1(t) \begin{bmatrix} h & s \\ 0 & 1 \end{bmatrix}^T ds + \int_0^h R_2(t) ds \\ &= \int_0^h \begin{bmatrix} h-s & \frac{(h-s)^2}{2} & 0 & 0 \\ 0 & 1 & h-s & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} ds + \int_0^h r ds \\ &= \frac{h^5}{20} + rh \end{aligned} \quad (3.5.27)$$

$x(k)$ 和 $y(k)$ 的互协方差为

$$\begin{aligned} R_{12}(k) &= \int_0^h \begin{bmatrix} t_{k+1} & s \\ 0 & 1 \end{bmatrix} R_1(t) \begin{bmatrix} t_{k+1} & s \\ 0 & 1 \end{bmatrix}^T ds \\ &= \int_0^h \begin{bmatrix} 1 & t-s & 0 & 0 \\ 0 & 1 & h-s & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \frac{(h-s)^2}{2} ds \end{aligned} \quad (3.5.28)$$

3.5.2 离散时间随机状态模型的极限化

考虑离散时间线性随机状态模型

$$x(k+1) = \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} x(k) + \begin{bmatrix} h \\ 0 \end{bmatrix} u(k) + \begin{bmatrix} w_1(k) \\ w_2(k) \end{bmatrix} \quad (3.5.29)$$

$$\begin{aligned} z(k) &= y(k+1) - y(k) \\ &= x(k) + u(k) + w_3(k) \end{aligned} \quad (3.5.30)$$

关于式(3.5.29), 已在 3.3 节讨论, 不重复。式(3.5.30)中: $z(k)$ 或 $y(k+1) - y(k)$ 为 $n \times 1$

量测向量; (k) 为 $m \times n$ 量测矩阵, 可以是时变或时不变的; (k) 为 $m \times 1$ 量测噪声。

假设 (k) 为离散时间高斯 $N(0, R_2(k))$ 白噪声, 并与 $x(k)$, (k) 互相独立, 即满足

$$E\{x(k_0) (1)\} = 0 \quad (1 \leq k_0) \tag{3.5.31}$$

$$E\{(k) (1)\} = 0 \tag{3.5.32}$$

由于状态向量初值 $x(k_0)$ 和 (k) 都是高斯过程, 因此由式(3.5.30)可知, $z(k)$ 或 $y(k+1) - y(k)$ 也是高斯过程。

为区分离散时间 k 和连续时间 t , 并把两者联系起来, 令 $k = t_k$ 。

通过极限化的方法把式(3.5.29)和式(3.5.30)转换成连续时间随机状态模型。令采样区间为 h , 由

$$\frac{d}{dt} (t, t_k) = A(t) (t, t_k) \tag{3.5.33}$$

得到

$$(t_k + h, t_k) = I + A(t_k)h \tag{3.5.34}$$

把式(3.5.34)代入式(3.5.29), 得到

$$x(t + h) - x(t) = A(t)x(t)h + (t) \tag{3.5.35}$$

$$x(t) = A(t)x(t)h + w(t) \tag{3.5.36}$$

式中用时间 t 代替了时间 t_k , 结果不变。 $w(t)$ 是一个 $m \times 1$ 的增量维纳过程, 它与 (t) 的关系表示为

$$w(t) = \frac{(t)h}{h} = \frac{w_1(t)}{h} \tag{3.5.37}$$

式中:

$$w_1(t) = (t)h \tag{3.5.38}$$

是增量维纳过程。显然, $w(t)$ 的均值为零, 协方差 $R_1(t_k)$ 为

$$\begin{aligned} R_1(t_k) &= E\{w(t_k) w^T(t_k)\} \\ &= \frac{E\{(k) (k)^T\}h}{h^2} \\ &= \frac{R_1(k)}{h} \end{aligned} \tag{3.5.39}$$

类似地, 把量测方程式(3.5.30)写成

$$\begin{aligned} y(t) &= \frac{(t)}{h} x(t)h + v(t) \\ &= H(t) x(t)h + v(t) \end{aligned} \tag{3.5.40}$$

式中:

$$H(t) = \frac{(t)}{h} \tag{3.5.41}$$

$v(t)$ 是 $m \times n$ 的维纳过程, 其均值为零, 协方差 $R_2(t_k)$ 为

$$R_2(t_k) = \frac{R_2(k)}{h} \tag{3.5.42}$$

当 $h \rightarrow 0$ 时, 式(3.5.36)和(3.5.40)就极限化为连续时间线性随机状态模型式(3.5.1)和(3.5.2)。

3.6 输入量是随机序列的离散时间系统分析

很多控制系统, 尽管被控对象是确定性的, 但输入量是随机信号, 本章后几节就分析这种控制系统。3.6 和 3.7 节分析离散时间情况, 3.8 和 3.9 两节分析连续时间情况。

3.6.1 系统脉冲响应

考虑离散时间控制系统, 如图 3.6.1 所示, $u(k)$ 表示输入, $y(k)$ 表示输出, 被控对象是确定性的, $h(k)$ 表示系统脉冲过渡函数。

图 3.6.1 系统框图

设研究的系统是渐近稳定的, 即脉冲过渡函数满足

$$\lim_{k \rightarrow \infty} h(k) = 0 \quad (3.6.1)$$

已知在输入序列 $\{u(k)\}$ 作用下, 系统输出 $y(k)$ 为卷积形式, 即

$$\begin{aligned} y(k) &= h(k)u(0) + h(k-1)u(1) + \dots + h(0)u(k) \\ &= \sum_{s=0}^k h(s)u(k-s) \end{aligned} \quad (3.6.2)$$

假设 $s > k$ 时, 有

$$u(k-s) = 0 \quad (3.6.3)$$

则 $y(k)$ 可表示为

$$\begin{aligned} y(k) &= \sum_{s=0}^k h(s)u(k-s) \\ &= \sum_{s=-\infty}^k h(k-s)u(s) \end{aligned} \quad (3.6.4)$$

下面要分析的是, $u(k)$ 是随机序列时 $y(k)$ 的变化规律。

3.6.2 系统的解

1. 解的存在性

定理 3.6.1 考虑渐近稳定系统, 脉冲过渡函数为 $h(k)$, 当系统输入 $u(k)$ 为二阶过程时, 则系统输出 $y(k)$ 在均方意义上存在, 且为二阶过程。

证明 用柯西收敛判别准则证明。对确定性序列 $\{x_s\}$, 用柯西收敛判别其一致收敛性时。先构造柯西序列 $\sum_{s=n}^m x_s$, 若满足

$$\lim_{n, m \rightarrow \infty} \sum_{s=n}^m x_s = 0 \quad (3.6.5)$$

则 x_s 收敛, 并存在收敛值 x 。若 $\{x_s\}$ 是随机序列, 只要对式 (3.6.5) 取均值, 就变成随机情况下的柯西收敛判别准则了。

对给定 $h(k)$ 和 $u(k)$, $y(k)$ 由式 (3.6.4) 给出, 构造柯西序列 $\sum_{s=n}^m y(k) = \sum_{s=n}^m h(s)u(k-s)$

- s), 为判别其是否均方收敛, 只要证明

$$\lim_{n, m} E \sum_{s= n}^m h(s)u(k - s)^2 = 0 \tag{3.6.6}$$

就够了。为此, 推导式(3.6.6):

$$\begin{aligned} \lim_{n, m} E \sum_{s= n}^m h(s)u(k - s)^2 &= \lim_{n, m} E \sum_{s= n}^m h(s)u(k - s) \sum_{l= n}^m h(l)u(k - l) \\ &= \lim_{n, m} \sum_{s= n}^m \sum_{l= n}^m h(s)h(l)E\{u(k - s)u(k - l)\} \end{aligned} \tag{3.6.7}$$

已知系统是渐近稳定的, 并且系统输入 $u(k)$ 是二阶过程, 即有

$$\lim_s Ch(s) \odot = 0 \tag{3.6.8}$$

$$E\{u(k - s)u(k - l)\} \leq C(\text{常数}) < \infty \tag{3.6.9}$$

把式(3.6.8)和(3.6.9)代入式(3.6.7)可知, 式(3.6.7)趋于零, 也就是式(3.6.6)成立, 所以 $y(k)$ 均方收敛, 即 $y(k)$ 存在。显然, $y(k)$ 均方收敛就是 $y(k)$ 自相关函数收敛, 因此, $y(k)$ 是二阶过程。

2. 系统的解

考虑脉冲过渡函数为 $h(k)$ 的渐近稳定系统, 给定系统输入 $u(k)$ 为二阶过程, 其均值为 $m_u(k)$ 和自相关函数为 $C_u(s, k)$ 。已经证明, 这时系统输出 $y(k)$ 存在, 且为二阶过程。下面就求出 $y(k)$ 的均值 $m_y(k)$ 、自相关函数 $C_y(s, k)$ 以及 $u(k)$ 和 $y(k)$ 的互相关函数 $C_{uy}(s, k)$ 。

通过 $y(k)$ 与 $u(k)$ 的关系式, 并取均值, 得到

$$y(k) = \sum_{l= 0}^{\infty} h(l)u(k - l) \tag{3.6.10}$$

$$m_y(k) = \sum_{l= 0}^{\infty} h(l)m_u(k - l) \tag{3.6.11}$$

由上式可知, 把系统 $u(k)$ 的均值 $m_u(k)$ 直接加到系统上, 就能简单得到系统输出 $y(k)$ 的均值 $m_y(k)$ 。

利用自相关函数的定义式, 容易求出 $C_y(s, k)$:

$$C_y(s, k) = E\{y(s)y(k)\} = \sum_{t= 0}^{\infty} \sum_{l= 0}^{\infty} h(t)h(l)C_u(s - t, k - l) \tag{3.6.12}$$

同样能求出 $C_{uy}(s, k)$:

$$\begin{aligned} C_{uy}(s, k) &= E\{u(s)y(k)\} \\ &= \sum_{l= 0}^{\infty} h(l)C_u(s, k - l) \end{aligned} \tag{3.6.13}$$

3.6.3 平稳过程情况

通常, 研究的实际问题大多属于平稳过程情况, 因此, 分析平稳过程情况是重要的。平稳过程情况是指系统输入 $u(k)$ 不仅是二阶过程, 而且为平稳过程, 即

$$m_u(k) = m_u(\text{常数}) \tag{3.6.14}$$

$$C_u(s, k) = C_u(s - k) \quad (3.6.15)$$

同样, 要求求出 $m_y(k)$, $C_y(s, k)$ 和 $C_{uy}(s, k)$ 。另外, 已知平稳过程能用谱密度描述, 这里也对此加以分析。

1. 求 $m_y(k)$, $C_y(s, k)$ 和 $C_{uy}(s, k)$

把式(3.6.14)代入式(3.6.11), 得到

$$m_y(k) = \sum_{l=0} m_u h(l) \quad (3.6.16)$$

考虑输入 $u(k)$ 的自相关函数是时间差的函数, 并令 $n = s - k$, 则有

$$\begin{aligned} C_u(s - t, k - l) &= C_u(s - k + l - t) \\ &= C_u(n + l - t) \end{aligned} \quad (3.6.17)$$

$$\begin{aligned} C_u(s, k - l) &= C_u(s - k + l) \\ &= C_u(n + l) \end{aligned} \quad (3.6.18)$$

把式(3.6.17)和(3.6.18)分别代入式(3.6.12)和(3.6.13), 得到

$$\begin{aligned} C_y(s, k) &= C_y(s - k) \\ &= \sum_{t=0} \sum_{l=0} h(t)h(l)C_u(n + l - t) \end{aligned} \quad (3.6.19)$$

$$\begin{aligned} C_{uy}(s, k) &= C_{uy}(s - k) \\ &= \sum_{l=0} h(l)C_u(n + l) \end{aligned} \quad (3.6.20)$$

由式(3.6.16)和(3.6.19)看到, 当 $u(k)$ 是平稳过程时, $y(k)$ 的均值也是常数, $y(k)$ 的自相关函数也是时间差 $(s - k) = n$ 的函数, 因此, $y(k)$ 也是平稳过程。同样, 由式(3.6.20)看到, 当 $u(k)$ 是平稳过程时, $u(k)$ 和 $y(k)$ 也是联合平稳的。

2. 求 m_y , $y(\cdot)$ 和 $u_y(\cdot)$

已知平稳随机过程能用谱密度描述, 谱密度和平稳随机过程的相关函数是一对傅里叶变换。给定系统输入 $u(k)$ 为平稳过程, 已知其均值为 m_u 和自相关函数为 $C_u(n)$, 可计算出系统输入 $u(k)$ 的谱密度 $u(\cdot)$ 。下面就在已知系统输入 $u(k)$ 的谱密度 $u(\cdot)$ 条件下, 求出系统输出 $y(k)$ 的谱密度 $y(\cdot)$, 以及 $u(k)$ 和 $y(k)$ 的互谱密度 $u_y(\cdot)$ 。

首先, 写出系统脉冲传递函数 $H(z)$, 按控制原理, 系统脉冲传递函数 $H(z)$ 是系统过渡函数 $h(n)$ 的 z 变换, 即

$$H(z) = \sum_{n=0} h(n)z^{-n} \quad (3.6.21)$$

$$z = e^{Ts} \quad (3.6.22)$$

式中: T 为采样时间, 本书中都设定为单位时间, 即 $T = 1$; s 可写为 i 。这样, 式(3.6.22)可写为

$$z = e^s = e^i \quad (3.6.23)$$

把上式代入式(3.6.21), 得到关系式

$$H(e^i) = \sum_{n=0} h(n)e^{-in} \quad (3.6.24)$$

$$H(e^{-i\omega}) = \sum_{n=-\infty}^{\infty} h(n)e^{in\omega} \quad (3.6.25)$$

其次, 求出 m_y 。令 $z=1$, 则式(3.6.21)变为

$$H(1) = \sum_{n=-\infty}^{\infty} h(n) \quad (3.6.26)$$

把上式代入式(3.6.16), 得到

$$m_y = H(1)m_u \quad (3.6.27)$$

再次, 求出 $y(\omega)$ 。把式(3.6.19)代入 $y(\omega)$ 的定义式, 并注意到式(3.6.24)和(3.6.25), 得到

$$\begin{aligned} y(\omega) &= \sum_{n=-\infty}^{\infty} C_y(n) e^{-in\omega} \\ &= \sum_{t=-\infty}^{\infty} h(t) e^{-it\omega} \sum_{l=-\infty}^{\infty} h(l) e^{il\omega} \sum_{n=-\infty}^{\infty} C_u(n+l-t) e^{-i(n+l-t)\omega} \\ &= H(e^{i\omega}) H(e^{-i\omega}) u(\omega) \\ &= |H(e^{i\omega})|^2 u(\omega) \end{aligned} \quad (3.6.28)$$

由上式可知, 当系统输入谱密度为 $u(\omega)$ 时, 通过系统脉冲传递函数 $|H(e^{i\omega})|^2$ 就能得到系统输出的谱密度。

最后, 确定 $u_y(\omega)$ 。同样, 把式(3.6.20)代入 $u_y(\omega)$ 的定义式, 并注意到式(3.6.25), 得到

$$\begin{aligned} u_y(\omega) &= \sum_{n=-\infty}^{\infty} C_{uy}(n) e^{-in\omega} \\ &= \sum_{l=-\infty}^{\infty} h(l) e^{il\omega} \sum_{n=-\infty}^{\infty} C_u(n+l) e^{-i(n+l)\omega} \\ &= H(e^{-i\omega}) u(\omega) \end{aligned} \quad (3.6.29)$$

3. 系统输入是白噪声情况

当系统输入 $u(k)$ 是白噪声时, 其谱密度为常数, 为简便起见, 令 $u(\omega) = 1$, 则式(3.6.29)变为

$$u_y(\omega) = |H(e^{-i\omega})| \quad (3.6.30)$$

这表明, 当系统输入 $u(k)$ 是白噪声时, 通过量测互谱密度 $u_y(\omega)$, 就可得到系统的脉冲传递函数 $H(e^{-i\omega})$ 。

在上述系统输入时, 其自相关函数为 $C_u(n) = \delta_n$, 则式(3.6.20)只有在 $l = -n$ 时才有值, 并等于

$$C_{uy}(n) = h(-n) \quad (3.6.31)$$

这表明, 当系统输入 $u(k)$ 是白噪声时, 通过量测互相关函数 $C_{uy}(n)$, 就可确定系统的脉冲过渡函数 $h(-n)$ 。

例 3.6.1 考虑 ARMA(1, 1) 方程

$$y(k) + a y(k-1) = e(k) + c e(k-1) \quad (3.6.32)$$

式中: $y(k)$ 是系统输出; 系统输入 $e(k)$ 是离散时间高斯 $N(0, 1)$ 白噪声; 系数 $|a|, |c| < 1$,

例 1. 试确定 $y(k)$ 的自相关函数 $C_y(n)$, 以及 $e(k)$ 和 $y(k)$ 的互相关函数 $C_{ey}(n)$ 。

解 在 3.2 节已经知道, 对 ARMA 方程, 把噪声项 $e(k)$ 作为系统输入时, 系统输出 $y(k)$ 是平稳过程, 因此, 可以用式(3.6.19)和(3.6.20)求解。

首先, 确定系统的脉冲过渡函数 $h(n)$ 。为此, 写出系统的脉冲传递函数 $H(z)$:

$$\begin{aligned} H(z) &= \frac{z+c}{z+a} \\ &= 1 + (c-a)\frac{z}{z+a}z^{-1} \end{aligned} \tag{3.6.33}$$

对式(3.6.33)进行 z 的反变换, 得到系统的脉冲过渡函数

$$h(n) = \delta_n + (c-a)(-a)^{n-1} \tag{3.6.34}$$

即

$$h(n) = \begin{cases} 1 & (n=0) \\ (c-a)(-a)^{n-1} & (n \geq 1) \end{cases} \tag{3.6.35}$$

其次, 利用式(3.6.19)

$$C_y(n) = \sum_{t=0}^{\infty} \sum_{l=0}^{\infty} h(t)h(l)C_e(n+1-t) \tag{3.6.36}$$

确定 $C_y(n)$ 。因为给定输入 $e(k)$ 是离散时间高斯 $N(0,1)$ 白噪声, 所以有

$$C_e(n+1-t) = \begin{cases} 1 & (t=n+1) \\ 0 & (t \neq n+1) \end{cases} \tag{3.6.37}$$

把式(3.6.37)和(3.6.35)代入式(3.6.36), 得到

$$\begin{aligned} C_y(0) &= \sum_{t=0}^{\infty} \sum_{l=0}^{\infty} h(t)h(l)C_e(1-t) \\ &= \sum_{l=0}^{\infty} h(l)h(l) \\ &= h(0)h(0) + \sum_{l=1}^{\infty} h(l)h(l) \\ &= 1 + (c-a)^2 \sum_{l=1}^{\infty} (-a)^{2(l-1)} \\ &= \frac{1-2ac+c^2}{1-a^2} \quad (n=0) \end{aligned} \tag{3.6.38}$$

$$\begin{aligned} C_y(n) &= \sum_{l=0}^{\infty} h(n+1-l)h(l) \\ &= h(n)h(0) + \sum_{l=1}^{\infty} h(n+1-l)h(l) \\ &= (c-a)(-a)^{n-1} + (c-a)^2 \sum_{l=1}^{\infty} (-a)^{n+1-l-1}(-a)^{l-1} \\ &= -\frac{c}{a} \frac{1-ac}{1-a^2} (-a)^n \quad (n \geq 1) \end{aligned} \tag{3.6.39}$$

总之, 有

$$C_y(n) = \begin{cases} \frac{1 - \frac{2ac + c^2}{1 - a^2}}{1 - \frac{c}{a} \frac{1 - \frac{ac}{1 - a^2}}{(-a)^n}} & (n = 0) \\ 1 - \frac{c}{a} \frac{1 - \frac{ac}{1 - a^2}}{(-a)^n} & (n \neq 0) \end{cases} \quad (3.6.40)$$

最后, 利用式(3.6.20)

$$C_{ey}(n) = \sum_{l=0}^{\infty} h(l) C_e(n + l) \quad (3.6.41)$$

确定 $C_{ey}(n)$ 。同样有

$$C_e(n + l) = \begin{cases} 1 & (n + l = 0) \\ 0 & (n + l \neq 0) \end{cases} \quad (3.6.42)$$

$$h(l) = 0 \quad (l < 0) \quad (3.6.43)$$

把式(3.6.42), (3.6.43)和(3.6.35)代入式(3.6.41), 得到

$$C_{ey}(n) = \begin{cases} 1 & (n = 0) \\ 0 & (n > 0) \\ (c - a)(-a)^{n-1} & (n < 0) \end{cases} \quad (3.6.44)$$

上式结果表明: 对 $e(k + n)$ 和 $y(k)$ 来说, 只有当 $n = 0$ 时, 它们才是相关的; $n > 0$ 时, 它们无关。

3.6.4 新息的概念

考虑脉冲过渡函数为 $h(k)$ 的渐近稳定系统, 系统输入为离散时间高斯 $N(0, 1)$ 白噪声 $e(k)$ 。由式(3.6.4), 系统输出 $y(k)$ 表示为

$$y(k) = \sum_{s=-\infty}^k h(k - s)e(s) \quad (3.6.45)$$

设式(3.6.45)是可逆的, 则存在一个函数 g , 满足

$$e(k) = \sum_{l=-\infty}^k g(k - l)y(l) \quad (3.6.46)$$

式(3.6.45)和(3.6.46)表明, 输出序列 $\{y(k)\}$ 的信息相当于输入序列 $\{e(k)\}$ 的信息, 两个序列包含着相同的信息。

考虑 $y(k + 1)$ 表示式。用式(3.6.45)写出 $y(k + 1)$ 表示式, 并把式(3.6.46)代入, 得到

$$\begin{aligned} y(k + 1) &= \sum_{s=-\infty}^{k+1} h(k + 1 - s)e(s) \\ &= \sum_{s=-\infty}^k h(k + 1 - s) \sum_{l=-\infty}^s g(k - l)y(l) + h(0)e(k + 1) \end{aligned} \quad (3.6.47)$$

由式(3.6.47)看出, 可把 $y(k + 1)$ 写为两项的和。第一项是 $y(k), y(k - 1), \dots$ 的线性组合, 在时刻 k , 可认为这些量是已知的, 因此, 第一项是已知的信息; 第二项是由 $e(k + 1)$ 提供的, 是未知量, 因此, 第二项是未知的信息, 称为新息。所以, $y(k + 1)$ 由已知信息和新息组成, 式(3.6.47)也称为过程的新息表示。

新息表示还有其他形式, 新息表示在研究预测和随机控制问题时是重要的。

3.7 离散时间过程的谱分解

从上节已经知道,对脉冲传递函数为 $H(e^i)$ 的渐近稳定系统,当加入谱密度为 $u(\omega)$ 的二阶平稳过程 $u(k)$ 时,则输出 $y(k)$ 也是平稳过程,且谱密度 $y(\omega)$ 表示为

$$\begin{aligned} y(\omega) &= H(e^{i\omega})H(e^{-i\omega})u(\omega) \\ &= \mathbb{C}H(e^{i\omega})\mathbb{C}_1^2 u(\omega) \end{aligned} \quad (3.7.1)$$

当 $u(k)$ 是白噪声,且 $u(\omega) = 1$ 时,有

$$\begin{aligned} y(\omega) &= H(e^{i\omega})H(e^{-i\omega}) \\ &= \mathbb{C}H(e^{i\omega})\mathbb{C}_1^2 \end{aligned} \quad (3.7.2)$$

这里存在一个相反的问题,如果已知 $y(\omega)$,能否把它分解为式(3.7.2)的形式,这就是谱分解问题。一般说,对任意谱密度进行谱分解是困难的,甚至是不可能的。但是,对有理谱密度进行谱分解是容易的,很多实际问题属于这种情况。

本节讨论有理谱密度的谱分解和谱表示问题。

3.7.1 有理谱密度和谱分解举例

定义 3.7.1 若谱密度 $y(\omega)$ 是 $e^{i\omega}$ (或 $\cos \omega$) 的有理函数,即 $y(\omega)$ 是 ω 的实函数,则称 $y(\omega)$ 为有理谱密度。具有有理谱密度的随机过程称为有理谱密度过程。

下面举一个例子,说明有理谱密度可以进行谱分解。

例 3.7.1 试对有理谱密度

$$y(\omega) = \frac{1.04 + 0.4\cos \omega}{1.25 + \cos \omega} \quad (3.7.3)$$

进行谱分解。

解 把式(3.7.3)的分子和分母做如下分解:

$$\begin{aligned} \text{分子} &= 1 + 0.4j\frac{1}{2}(e^{i\omega} + e^{-i\omega}) + 0.04 \\ &= (e^{i\omega} + 0.2)(e^{-i\omega} + 0.2) \end{aligned} \quad (3.7.4)$$

$$\text{分母} = (e^{i\omega} + 0.5)(e^{-i\omega} + 0.5) \quad (3.7.5)$$

分解中利用了欧拉公式

$$\cos \omega = \frac{1}{2}(e^{i\omega} + e^{-i\omega}) \quad (3.7.6)$$

把式(3.7.4)和(3.7.5)代入式(3.7.3),立即得到

$$y(\omega) = H(e^{i\omega})H(e^{-i\omega}) \quad (3.7.7)$$

式中已令

$$H(e^{i\omega}) = \frac{e^{i\omega} + 0.2}{e^{i\omega} + 0.5} \quad (3.7.8)$$

$$H(e^{-i\omega}) = \frac{e^{-i\omega} + 0.2}{e^{-i\omega} + 0.5} \quad (3.7.9)$$

由式(3.7.7)可知,有理谱密度式(3.7.3)可分解为式(3.7.7),即式(3.7.2)的形式,但分

解不是唯一的, 式(3.7.3) 还容易分解为下列三种形式:

$$\begin{aligned} \Phi(e^{j\omega}) &= \frac{0.4(e^{j\omega} + 5)}{(e^{j\omega} + 2)} \cdot \frac{0.4(e^{-j\omega} + 5)}{(e^{-j\omega} + 2)} \\ &= \frac{0.2(e^{j\omega} + 5)}{(e^{j\omega} + 0.5)} \cdot \frac{0.2(e^{-j\omega} + 5)}{(e^{-j\omega} + 0.5)} \\ &= \frac{2(e^{j\omega} + 0.2)}{(e^{j\omega} + 2)} \cdot \frac{2(e^{-j\omega} + 0.2)}{(e^{-j\omega} + 2)} \end{aligned} \tag{3.7.10}$$

由上面例子可知: 第一, 有理谱密度是可以进行谱分解的; 第二, 谱分解不是唯一的。但为使系统稳定, 要求系统脉冲传递函数在单位圆外没有零点和极点, 这也是可以做到的, 如例 3.7.1 中, 只要采用式(3.7.7)的分解形式, 而不采用式(3.7.10) 所示其他三种分解形式。

下面用定理形式证明上述结论。

3.7.2 谱分解定理

定理 3.7.1 考虑有理谱(密度)过程, 其谱密度 $\Phi(e^{j\omega})$ 的分子为偶数 m 阶, 分母为偶数 n 阶, $m \leq n$, 则必然存在一个有理函数 H , 其全部极点都在单位圆内, 且全部零点都在单位圆内或单位圆上, 并满足

$$\begin{aligned} \Phi(e^{j\omega}) &= H(e^{j\omega})H(e^{-j\omega}) \\ &= \mathcal{C}H(e^{j\omega})\mathcal{C}^H \end{aligned} \tag{3.7.11}$$

证明 证明分公式证明和稳定性证明两步进行。

1. 公式证明

已知 $\Phi(e^{j\omega})$ 是 $e^{j\omega}$ 的有理函数, 即 $\Phi(e^{j\omega})$ 是 ω 的实函数。设 $(e^{j\omega} - z_k)$ 是 $\Phi(e^{j\omega})$ 分子中的一个因子, 其中 z_k 为复数, 表示为

$$z_k = \alpha_k + j\beta_k \tag{3.7.12}$$

为保证 $\Phi(e^{j\omega})$ 是 ω 的实函数, 必存在其共轭因式 $(e^{-j\omega} - \bar{z}_k)$, 这是因为有

$$(e^{j\omega} - z_k)(e^{-j\omega} - \bar{z}_k) = (\cos \omega - \alpha_k)^2 + (\sin \omega - \beta_k)^2 \tag{3.7.13}$$

可想而知, 对分子肯定都能进行上述分解, 这也是要求分子阶数 m 为偶数的原因。对 $\Phi(e^{j\omega})$ 的分母部分, 也能进行同样的分析。于是得到

$$\begin{aligned} \Phi(e^{j\omega}) &= \frac{\prod_{k=1}^{\frac{m}{2}} C_k (e^{j\omega} - z_k)}{\prod_{l=1}^{\frac{n}{2}} D_l (e^{j\omega} - p_l)} \cdot \frac{\prod_{k=1}^{\frac{m}{2}} C_k^* (e^{-j\omega} - \bar{z}_k)}{\prod_{l=1}^{\frac{n}{2}} D_l^* (e^{-j\omega} - \bar{p}_l)} \\ &= H(e^{j\omega})H(e^{-j\omega}) \\ &= \mathcal{C}H(e^{j\omega})\mathcal{C}^H \end{aligned} \tag{3.7.14}$$

式中: C 为常系数; $p_l (l= 1, 2, \dots, \frac{n}{2})$ 是 $\Phi(e^{j\omega})$ 分母的极点, 一般是复数; $H(e^{j\omega})$ 和 $H(e^{-j\omega})$ 分别为

$$H(e^{j\omega}) = \frac{\prod_{k=1}^{\frac{m}{2}} (e^{j\omega} - z_k)}{\prod_{l=1}^{\frac{n}{2}} (e^{j\omega} - p_l)} \quad (3.7.15)$$

$$H(e^{-j\omega}) = \frac{\prod_{k=1}^{\frac{m}{2}} (e^{-j\omega} - \bar{z}_k)}{\prod_{l=1}^{\frac{n}{2}} (e^{-j\omega} - \bar{p}_l)} \quad (3.7.16)$$

当然,可以把 $H(e^{j\omega})$ 写成多项式的形式:

$$\begin{aligned} H(e^{j\omega}) &= \frac{A(e^{j\omega})}{B(e^{j\omega})} \\ &= \frac{\sum_{k=0}^{\frac{m}{2}} a_k e^{j(\frac{m}{2}-k)\omega}}{\sum_{l=0}^{\frac{n}{2}} b_l e^{j(\frac{n}{2}-l)\omega}} \end{aligned} \quad (3.7.17)$$

式中: $A(e^{j\omega})$ 和 $B(e^{j\omega})$ 分别为分子和分母多项式; $a_k, k=1, 2, \dots, \frac{m}{2}$ 和 $b_l, l=1, 2, \dots, \frac{n}{2}$ 分别为分子和分母多项式系数。

式(3.7.14)完成了公式的证明。

2. 稳定性证明

为使系统稳定,要求系统脉冲传递函数 $H(e^{j\omega})$ 的所有零点都在单位圆内或单位圆上,即零点的模小于或等于 1,而所有极点都在单位圆内,即极点的模小于 1。谱密度 $S(\omega)$ 是可积函数,因此,单位圆上不可能有极点。

只要证明,当零点或极点的模大于 1 时,都能变换为小于 1,就能使系统稳定。

设谱密度 $S(\omega)$ 的分子中有模大于 1 的零点 z_k ,则必然存在共轭零点 \bar{z}_k 。做如下运算:

$$\begin{aligned} (e^{j\omega} - z_k)(e^{-j\omega} - \bar{z}_k) &= \bar{z}_k e^{j\omega} - \frac{1}{\bar{z}_k} e^{-j\omega} - \frac{1}{z_k} \\ &= \frac{1}{z_k \bar{z}_k} (e^{j\omega} - z_k)(e^{-j\omega} - \bar{z}_k) \end{aligned} \quad (3.7.18)$$

式中: $\frac{1}{z_k} = \frac{1}{\bar{z}_k}$, $\frac{1}{\bar{z}_k} = \frac{1}{z_k}$ 。当 z_k 和 \bar{z}_k 的模大于 1 时, $\frac{1}{z_k}$ 和 $\frac{1}{\bar{z}_k}$ 的模必然小于 1。把其模大于 1 的其他零点和极点都做式(3.7.18)所示的运算,就满足系统稳定性要求了。

3.7.3 谱表示定理

谱表示定理是谱分解定理的逆定理。

定理 3.7.2 给定有理谱密度 $S(\omega)$,那么必然存在一个渐近稳定系统,只要给该系统

加入离散时间白噪声,就能使系统的输出成为谱密度为 ()的平稳过程。

证明 根据谱分解定理,对给定有理谱密度 (),必然可分解为稳定的有理函数 $H(e^{j\omega})$ 和 $H(e^{-j\omega})$ 的乘积。把有理函数 $H(e^{j\omega})$ 作为系统的脉冲传递函数,并在该系统上加入离散时间白噪声,那么由关系式(3.7.11)可知,系统的输出就是谱密度为 ()的平稳过程。

这个定理说明,通过把离散时间白噪声加到一个渐近稳定的系统上,改变系统的脉冲传递函数,就能得到有理谱密度过程。

例 3.7.2 试证明由两个离散时间线性状态方程

$$x(k+1) = 0.8x(k) - 1.2e(k) \tag{3.7.19}$$

$$y_1(k) = x(k) + e(k) \tag{3.7.20}$$

和

$$x(k+1) = 0.8x(k) + 0.6e(k) \tag{3.7.21}$$

$$y_2(k) = x(k) + 2e(k) \tag{3.7.22}$$

表示的随机过程 $y_1(k)$ 和 $y_2(k)$ 具有相同的谱密度。式中: $e(k)$ 是离散时间高斯 $N(0, 1)$ 白噪声。

解 两个系统的输入量是 $e(k)$, 输出量分别是 $y_1(k)$ 和 $y_2(k)$, 我们顺序求出 $y_1(k)$ 和 $y_2(k)$ 的谱密度。

对式(3.7.19)和(3.7.20)进行 z 变换得:

$$x(z) = \frac{-1.2}{z - 0.8}e(z) \tag{3.7.23}$$

$$y_1(z) = x(z) + e(z) \tag{3.7.24}$$

把式(3.7.23)代入式(3.7.24),得到 $y_1(k)$ 的脉冲传递函数

$$y_1(z) = \frac{z - 2}{z - 0.8}e(z) \tag{3.7.25}$$

考虑到系统输入 $e(k)$ 的谱密度为

$$e(\omega) = 1 \tag{3.7.26}$$

由式(3.7.25)得到 $y_1(k)$ 的谱密度

$$\begin{aligned} y_1(\omega) &= \frac{e^{j\omega} - 2}{e^{j\omega} - 0.8} \frac{e^{-j\omega} - 2}{e^{-j\omega} - 0.8} \\ &= \frac{2(e^{j\omega} - 0.5)}{(e^{j\omega} - 0.8)} j\omega \frac{2(e^{-j\omega} - 0.5)}{(e^{-j\omega} - 0.8)} \end{aligned} \tag{3.7.27}$$

同样地,由式(3.7.21)和(3.7.22)导出 $y_2(k)$ 的脉冲传递函数和谱密度

$$y_2(z) = \frac{2z - 1}{z - 0.8}e(z) \tag{3.7.28}$$

$$\begin{aligned} y_2(\omega) &= \frac{2e^{j\omega} - 1}{e^{j\omega} - 0.8} j\omega \frac{2e^{-j\omega} - 1}{e^{-j\omega} - 0.8} \\ &= \frac{2(e^{j\omega} - 0.5)}{(e^{j\omega} - 0.8)} j\omega \frac{2(e^{-j\omega} - 0.5)}{(e^{-j\omega} - 0.8)} \end{aligned} \tag{3.7.29}$$

由式(3.7.27)和(3.7.29)可知, $y_1(k)$ 和 $y_2(k)$ 具有相同的谱密度,即

$$y_1(\omega) = y_2(\omega) \tag{3.7.30}$$

3.8 输入量是随机过程的连续时间系统分析

3.8.1 解的存在性

考虑渐近稳定系统, 系统的脉冲过渡函数为 $h(t)$ 。根据线性系统理论, 在系统输入 $u(t)$ 作用下, 系统输出为

$$\begin{aligned} y(t) &= \int_0^t h(t-s) u(s) ds \\ &= \int_0^t h(s) u(t-s) ds \end{aligned} \quad (3.8.1)$$

下面用定理形式表示解的存在性。

定理 3.8.1 考虑脉冲过渡函数为 $h(t)$ 的渐近稳定系统, 当系统输入 $u(t)$ 是均值为 $m_u(t)$ 和自相关函数为 $C_u(s, t)$ 的二阶随机过程时, 则系统输出 $y(t)$ 存在, 且为二阶过程。

证明 对式(3.8.1)构造均方意义上的随机柯西函数, 得到

$$\begin{aligned} \lim_{a, b} E\{y(t)^2\} &= \lim_{a, b} E \int_a^b h(k) u(s-k) dk \int_a^b h(l) u(t-l) dl \\ &= \lim_{a, b} \int_a^b \int_a^b h(k) h(l) C_u(s-k, t-l) dk dl \end{aligned} \quad (3.8.2)$$

按给定条件, 系统为渐近稳定系统, 系统输入为二阶过程, 有

$$\lim_k |h(k)| = 0 \quad (3.8.3)$$

$$C_u(s-k, t-l) \leq C(\text{常数}) < \infty \quad (3.8.4)$$

得到

$$\lim_{a, b} E\{y(t)^2\} = 0 \quad (3.8.5)$$

因此, 系统输出 $y(t)$ 在均方意义上收敛, 且 $y(t)$ 的均方值存在, 也就是 $y(t)$ 为二阶过程。

3.8.2 系统的解

考虑脉冲过渡函数为 $h(t)$ 的渐近稳定系统, 设系统输入 $u(t)$ 的均值为 $m_u(t)$ 和自相关函数为 $C_u(s, t)$ 。已经证明, 这时系统输出 $y(t)$ 存在, 且为二阶过程, 下面就顺序求出 $y(t)$ 的均值 $m_y(t)$, 自相关函数 $C_y(s, t)$, 以及 $u(t)$ 和 $y(t)$ 的互相关函数 $C_{uy}(s, t)$ 。

对式(3.8.1)求均值, 得到 $y(t)$ 的均值

$$\begin{aligned} m_y(t) &= E \int_0^t h(s) u(t-s) ds \\ &= \int_0^t h(s) m_u(t-s) ds \end{aligned} \quad (3.8.6)$$

按自相关函数定义, 并考虑式(3.8.1), 得到 $y(t)$ 的自相关函数

$$\begin{aligned} C_y(s, t) &= E\{y(s)y(t)\} \\ &= E \int_0^s h(k) u(s-k) dk \int_0^t h(l) u(t-l) dl \\ &= \int_0^s \int_0^t h(k) h(l) C_u(s-k, t-l) dk dl \end{aligned} \quad (3.8.7)$$

同样, 得到 $u(t)$ 和 $y(t)$ 的互相关函数

$$\begin{aligned} C_{uy}(s, t) &= E\{u(s)y(t)\} \\ &= E\left\{u(s) \int_0^t h(l)u(t-l)dl\right\} \\ &= \int_0^t h(l)C_u(s, t-l)dl \end{aligned} \quad (3.8.8)$$

3.8.3 平稳过程情况

这种情况是指, 系统输入 $u(t)$ 不仅是二阶过程, 而且是平稳过程, 即给定系统输入 $u(t)$ 的均值和自相关函数分别为

$$m_u(t) = m_u = C(\text{常数}) \quad (3.8.9)$$

$$\begin{aligned} C_u(s, t) &= C_u(s - t) \\ &= C_u(\tau) \quad (\tau = s - t) \end{aligned} \quad (3.8.10)$$

要求分别求出有关均值、相关函数和谱密度。

1. 求 $m_y(t)$, $C_y(s, k)$ 和 $C_{uy}(s, t)$

把式(3.8.9)代入式(3.8.6), 得到

$$\begin{aligned} m_y(t) &= m_u \int_0^t h(s)ds \\ &= m_y(\text{常数}) \end{aligned} \quad (3.8.11)$$

把式(3.8.10)代入式(3.8.7)和(3.8.8), 分别得到

$$\begin{aligned} C_y(s, t) &= \int_0^t \int_0^t h(k)h(l)C_u(\tau + l - k)dkdl \\ &= C_y(\tau) \end{aligned} \quad (3.8.12)$$

$$\begin{aligned} C_{uy}(s, t) &= \int_0^t h(l)C_u(\tau + l)dl \\ &= C_{uy}(\tau) \end{aligned} \quad (3.8.13)$$

由式(3.8.11)和(3.8.12)结果可知, 当系统输入 $u(t)$ 是平稳过程时, 系统输出 $y(t)$ 的均值为常数, 自相关函数只是时间差的函数, 因此, $y(t)$ 也是平稳过程。同样, 由式(3.8.13)结果可知, $u(t)$ 和 $y(t)$ 是联合平稳的。

2. 求 m_y , $C_y(\tau)$ 和 $C_{uy}(\tau)$

给定系统输入 $u(t)$ 的均值 m_u 和自相关函数 $C_u(\tau)$, 按平稳过程相关函数与谱密度的关系, 也就给定了系统输入 $u(t)$ 的谱密度

$$C_u(\tau) = \int_{-\infty}^{\infty} C_u(\tau) e^{-i\tau\omega} d\omega \quad (3.8.14)$$

首先, 写出系统的传递函数 $G(s)$, 按控制原理, 当给定系统脉冲过渡函数 $h(t)$ 时, 有

$$G(s) = \int_0^{\infty} h(t)e^{-st}dt \quad (3.8.15)$$

令 $s = i\omega$, 得到

$$G(i\omega) = \int_0^{\infty} h(t)e^{-i\omega t}dt \quad (3.8.16)$$

$$G(-i) = \int_0^\infty h(t)e^{-it}dt \quad (3.8.17)$$

其次, 求出 m_y 。令 $s=0$, 则式(3.8.15)变为

$$G(0) = \int_0^\infty h(t)dt \quad (3.8.18)$$

把上式代入式(3.8.11), 得到

$$m_y = m_u G(0) \quad (3.8.19)$$

再次, 求出 $y(\cdot)$ 。由 $y(\cdot)$ 的定义式及式(3.8.12), (3.8.14), (3.8.16)和(3.8.17), 容易得到

$$\begin{aligned} y(\cdot) &= \int_0^\infty C_y(\cdot) e^{-i\cdot} d \\ &= \int_0^\infty h(k)e^{-ik}dk \int_0^\infty h(l)e^{il}dl \int_0^\infty C_u(\cdot + l - k)e^{-(\cdot + l - k)}d \\ &= G(i)G(-i)u(\cdot) \end{aligned} \quad (3.8.20)$$

最后, 求出 $u_y(\cdot)$ 。同样, 由 $u_y(\cdot)$ 的定义式及式(3.8.13), (3.8.14)和(3.8.17), 得到

$$\begin{aligned} u_y(\cdot) &= \int_0^\infty C_{uy}(\cdot) e^{-i\cdot} d \\ &= \int_0^\infty h(l)e^{il}dl \int_0^\infty C_u(\cdot + l)e^{-i(\cdot + l)}d \\ &= G(-i)u(\cdot) \end{aligned} \quad (3.8.21)$$

3.8.4 高斯白噪声输入情况

仍考虑脉冲过渡函数为 $h(t)$ 的渐近稳定系统, 但系统输入 $u(t)$ 为高斯 $N(0, R_1(t)(t))$ 白噪声。高斯白噪声不是二阶过程, 但在这一特定情况下, 输出仍有意义。

已知高斯白噪声与维纳过程 $w(t)$ 的关系为

$$dw(t) = u(t)dt \quad (3.8.22)$$

式中 $dw(t)$ 是增量维纳过程, 它是高斯 $N(0, R_1(t)dt)$ 过程, 也是二阶过程。把式(3.8.22)代入式(3.8.1), 得到系统输出

$$y(t) = \int_0^t h(t-l)dw(l) \quad (3.8.23)$$

利用维纳过程特性和系统输出表示式, 可求出 $m_y(t)$, $C_y(s, t)$ 和 $y(\cdot)$ 等量。对式(3.8.23)取均值, 容易得到

$$m_y(t) = 0 \quad (3.8.24)$$

$$\begin{aligned} C_y(s, t) &= E\{y(s)y(t)\} \\ &= \int_0^s \int_0^t h(s-k)h(t-l)E\{dw(k)dw(l)\} \\ &= \int_0^t h(s-l)h(t-l)R_1dl \end{aligned} \quad (3.8.25)$$

做变量置换: $t-l=n$, 并令 $s-t=\tau$, 结果为

$$C_y(s, t) = \int_0^\infty h(\tau + n)h(n)R_1dn$$

$$= C_y(\tau) \quad (3.8.26)$$

上述结果表明,当系统输入为高斯白噪声时,系统输出 $y(t)$ 仍为二阶平稳过程。 $y(t)$ 的谱密度为

$$S_y(\omega) = \int_{-\infty}^{\infty} C_y(\tau) e^{-i\omega\tau} d\tau$$

$$= \int_{-\infty}^{\infty} \int_0^\infty h(\tau + n)h(n)R_1dn e^{-i\omega\tau} d\tau$$

$$= \int_0^\infty h(n) e^{-i\omega n} dn \int_{-\infty}^{\infty} h(\tau + n) e^{-i\omega(\tau + n)} d\tau R_1$$

$$= G(-i\omega)G(i\omega)R_1 \quad (3.8.27)$$

从上式第三等式到第四等式利用了

$$h(t) = 0 \quad (t < 0) \quad (3.8.28)$$

例 3.8.1 考虑渐近稳定系统,系统传递函数为

$$G(s) = \frac{\overline{2}(s + \overline{5})}{s^2 + 2s + 5} \quad (3.8.29)$$

系统输入和系统输出的关系式为

$$y(t) = \int_{-\infty}^t h(t - \tau) d\tau \quad (3.8.30)$$

式中: $d\tau(t)$ 为增量维纳过程。试写出系统输出 $y(t)$ 的表达式。

解 对系统传递函数式(3.8.29)进行拉普拉斯反变换,求出系统脉冲过渡函数

$$h(t) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} G(s) e^{st} ds$$

$$= \frac{\overline{2}}{2} e^{-t} \cos 2t + \frac{\overline{2}(\overline{5} - 1)}{2} e^{-t} \sin 2t \quad (3.8.31)$$

把式(3.8.31)代入式(3.8.30),得到最后结果

$$y(t) = \frac{\overline{2}}{2} \int_{-\infty}^t e^{-(t-\tau)} \cos 2(t - \tau) d\tau \quad (1)$$

$$+ \frac{\overline{2}(\overline{5} - 1)}{2} \int_{-\infty}^t e^{-(t-\tau)} \sin 2(t - \tau) d\tau \quad (1) \quad (3.8.32)$$

3.9 连续时间过程的谱分解

和 3.7 节类似,本节讨论连续时间过程的谱分解和谱表示问题。

3.9.1 谱分解定理

定理 3.9.1 考虑有理谱密度,其谱密度 $S_y(\omega)$ 的分子为偶数 m 阶,分母为偶数 n 阶, $m \leq n$, 则必然存在一个有理函数 $G(i\omega)$, 它的全部极点都在左半平面,全部零点都在左半平面或虚轴上,并满足

$$(\quad) = G(i)G(-i) \quad (3.9.1)$$

证明 通过分析 (\quad) 的分子零点、分母极点的特点来证明定理。

零点有三种情况:

(1) k 为实零点

为保证 (\quad) 为实函数, (\quad) 分子中若有一因子 $(i + k)$, 必有另一因子 $(-i + k)$, 这时有

$$(i + k)(-i + k) = k^2 + k^2 \quad (3.9.2)$$

确实得到实函数。

(2) k 是虚轴上的零点

同样, 若有因子 $(i + i k)$, 必有另一因子 $(-i + i k)$, 这时有

$$(i + i k)(-i + i k) = k^2 - k^2 \quad (3.9.3)$$

能保证 (\quad) 为实函数。

(3) k 为复根

同样, 若有因子 $(i + k)$, 必有另外三个因子 $(-i + k)$, $(i + \bar{k})$ 和 $(-i + \bar{k})$ 。令

$$k = k_1 + i k_2 \quad (3.9.4)$$

则有

$$\begin{aligned} & (i + k)(-i + k)(i + \bar{k})(-i + \bar{k}) \\ &= (k^2 + k^2)(\bar{k}^2 + k^2) \\ &= k^4 + 2(k_1^2 - k_2^2)k^2 + (k_1^2 + k_2^2)^2 \end{aligned} \quad (3.9.5)$$

确实得到实函数, 可保证 (\quad) 为实函数。

综上所述, 把 (\quad) 零点因子 $(i + k)$ 和 $(i + \bar{k})$ 部分放在一起, 称为 $B(i)$; 把 $(-i + k)$ 和 $(-i + \bar{k})$ 部分放在一起, 称为 $B(-i)$ 。 (\quad) 分子的阶次必须是偶数才有可能进行如上分解。

对 (\quad) 的分母也做类似处理, 把分母分为 $A(i)$ 和 $A(-i)$ 两部分。但由于 (\quad) 是可积函数, 因此, 在虚轴上不能有极点。

通过上述分析, 可把 (\quad) 写为

$$\begin{aligned} (\quad) &= \frac{B(i)}{A(i)} i^{\alpha} \frac{B(-i)}{A(-i)} \\ &= G(i)G(-i) \end{aligned} \quad (3.9.6)$$

式中:

$$G(i) = \frac{B(i)}{A(i)} \quad (3.9.7)$$

$$G(-i) = \frac{B(-i)}{A(-i)} \quad (3.9.8)$$

显然, $G(i)$ 的零点都在左半平面或虚轴上, 极点都在左半平面, 若把 $G(i)$ 作为系统的传递函数, 它是稳定的。

3.9.2 谱表示定理

定理 3.9.2 给定有理谱密度函数 (\quad) , 则必然存在一个渐近稳定的系统, 它的传递

函数为 $G(s)$, 对应的脉冲过渡函数为 $h(t)$, 使得

$$y(t) = \int_0^t h(t - \tau) d\tau \tag{3.9.9}$$

式中: $y(t)$ 的谱密度为 $S_y(\omega)$; $d\tau$ 为增量维纳过程。

证明 令 $d\tau$ 为 $N(0, d\tau)$, 则有

$$S_y(\omega) = |G(j\omega)|^2 S_d(\omega) \tag{3.9.10}$$

把由 $S_y(\omega)$ 谱分解得到的 $G(j\omega)$ 作为系统传递函数, 它对应 $h(t)$, 必满足式(3.9.9)。

例 3.9.1 某平稳随机过程有自相关函数

$$C(\tau) = e^{-2|\tau|} \cos 2\tau \tag{3.9.11}$$

试求出形式为

$$y(t) = \int_0^t h(t - \tau) d\tau \tag{3.9.12}$$

的过程的表示式, 其中 $d\tau$ 为增量维纳过程, 其均值为零, 增量方差为 $d\tau$ 。

解 先利用给定 $C(\tau)$ 计算出系统输出 $y(t)$ 的谱密度 $S_y(\omega)$, 进而推算系统的传递函数 $G(s)$ 和系统脉冲过渡函数 $h(t)$, 最后写出 $y(t)$ 的表示式。

$S_y(\omega)$ 是 $C(\tau)$ 的傅里叶变换, 由式(3.9.11) 得到

$$\begin{aligned} S_y(\omega) &= \int_{-\infty}^{\infty} C(\tau) e^{-j\omega\tau} d\tau \\ &= \int_{-\infty}^{\infty} e^{-2|\tau|} \cos 2\tau e^{-j\omega\tau} d\tau \\ &= \int_{-\infty}^0 e^{\cos 2\tau} e^{-j\omega\tau} d\tau + \int_0^{\infty} e^{-\cos 2\tau} e^{-j\omega\tau} d\tau \\ &= \frac{-s+1}{(-s+1)^2+2^2} + \frac{s+1}{(s+1)^2+2^2} \\ &= \frac{2(s+5)}{s^2+2s+5} \cdot j\omega \frac{2(-s+5)}{s^2-2s+5} \\ &= G(s) G(-s) \end{aligned} \tag{3.9.13}$$

最后利用了关系式(3.9.10), 式中 $s=j\omega$ 。于是得到系统传递函数

$$G(s) = \frac{2(s+5)}{s^2+2s+5} \tag{3.9.14}$$

上式就是例 3.8.1 给定的系统传递函数, 继而由式(3.8.31) 和(3.8.32) 分别给出 $h(t)$ 和 $y(t)$ 的表达式。

习 题

3.1 某动力学系统由以下差分方程描述:

$$x(k+1) = \begin{bmatrix} \cosh & \sinh \\ -\sinh & \cosh \end{bmatrix} x(k)$$

式中:

$$h = \frac{1}{4n}$$

它的初始状态 $x(0)$ 是高斯的;其均值和协方差分别为

$$E\{x(0)\} = m_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$R_0 = cov[x(0), x(0)] = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

试确定使在 $k = k^*$ 时的分量 x_1 和 x_2 是独立的最小 k^* ,再确定 $x(k^*)$ 的分布。

3.2 一动力学系统由以下随机差分方程给定:

$$x(k+1) = \begin{bmatrix} 1 & 0.5 \\ 0.1 & 0.2 \end{bmatrix} x(k) + \begin{bmatrix} 1.0 \\ 0.5 \end{bmatrix} e(k)$$

式中: $\{e(k), k \in T\}$ 是独立同分布高斯 $N(0, 1)$ 随机变量序列; $e(k)$ 和 $x(k)$ 互相独立。试确定定态分布的协方差。

3.3 某随机过程由以下随机差分方程确定:

$$x(k+1) = ax(k) + e(k) \quad (a \in (-1, 1))$$

式中: $\{e(k), k \in T\}$ 是独立同分布高斯 $N(0, \sigma^2)$ 随机变量序列; 初始状态 $x(k_0)$ 为高斯 $N(0, \sigma_0^2)$; $\{e(k), k \in T\}$ 与 $x(k_0)$ 是独立的。试确定 $x(k)$ 的方差和当 $k \rightarrow \infty$ 或 $k_0 \rightarrow -\infty$ 时方差的极限。当选

$$\sigma_0^2 = \lim_{k \rightarrow \infty} P(k)$$

和取指标集为 $T = \{k_0, k_0+1, \dots\}$ 时, 试证明此过程是平稳的。并确定此平稳过程的协方差函数和谱密度。

3.4 考虑满足以下随机差分方程的平稳随机过程:

$$x(k+1) = x(k) + e(k)$$

式中: $\{e(k), k \in T\}$ 是均值为零和协方差阵为 R_1 的独立同分布向量序列。设 λ 的特征多项式为

$$\det[I - \lambda A] = \lambda^n + a_1 \lambda^{n-1} + \dots + a_n$$

试证明状态变量分量的任意线性组合的协方差函数 $r(\cdot)$ 满足

$$r(k) + a_1 r(k-1) + \dots + a_n r(k-n) = 0 \quad (k \geq n)$$

3.5 火箭发动机推力的变分可近似地用推进航向角 θ 上的白噪声变分来描述, 假设其谱密度为

$$N = 0.0004 \text{ rad}^2/\text{Hz}$$

火箭加速度 a 的大小为 3 m/s^2 。已知横向运动可描述为

$$\ddot{x} = a$$

试确定加速 100 s 后的横向误差和横向误差速度的方差。

3.6 某系统状态为

$$dx = \begin{bmatrix} -a_1 & -a_2 \\ 1 & 0 \end{bmatrix} x dt + \begin{bmatrix} 1 \\ 0 \end{bmatrix} dv \quad (1)$$

式中: $a_1 > 0, a_2 > 0; \{v(t), t \in T\}$ 是具有单位方差参数的维纳过程。试计算其定态协方差阵。

3.7 标量随机过程 $\{x(t), t \in T\}$ 满足随机微分方程

$$dx = \lambda x dt + dv$$

式中: $\{v(t), t \in T\}$ 是一个方差参数为 r_1 的维纳过程。初始状态 x_0 是高斯的, 其均值为 m_0 , 协方差为 r_0 。过程 $v(t)$ 对 $x(t_0)$ 独立。当 $t \rightarrow \infty$ 或 $t_0 \rightarrow -\infty$, 保证 $E\{x^2(t)\}$ 的极限存在时, 试确定 $x(t)$ 的均值函数 $m(t)$ 和协方差函数 $P(t)$ 。如果选择 $m_0 = 0, r_0 = \lim_{t \rightarrow \infty} E\{x^2(t)\}$, 试证明过程 $\{x(t), t_0 \leq t < \infty\}$ 是平稳的, 并计算此平稳过程的谱密度和协方差函数。

3.8 考虑满足如下随机微分方程的平稳随机过程:

$$dx = Ax dt + dv$$

式中: 矩阵 A 的特征多项式为

$$\det[I - \lambda A] = \lambda^n + a_1 \lambda^{n-1} + \dots + a_n$$

它的零点全部在左半平面; 而 $\{v(t), t \in T\}$ 是增量方差为 $R_1 dt$ 的维纳过程。试证明方程式状态变量的任意线性组合的协方差函数满足微分方程

$$\frac{d^n r}{dt^n} + a_1 \frac{d^{n-1} r}{dt^{n-1}} + \dots + a_n r = 0 \quad (t > 0)$$

3.9 考虑随机微分方程

$$dx = Ax dt + b dv$$

式中: A 是 $n \times n$ 常阵; b 是 n 维常向量; $\{v(t), t \in T\}$ 是具有单位方差参数的维纳过程。假设 A 的所有特征值都有负实部, 试证明 x 的定态协方差为

$$r = \int_0^\infty z(t) z^T dt$$

式中: z 是初始条件为

$$z(0) = b$$

的微分方程

$$\frac{dz}{dt} = Az$$

的解。

3.10 考虑平稳随机过程

$$dx = Ax dt + b dv$$

$$y = x_1$$

式中: A 是 $n \times n$ 常阵; b 是 n 维常向量; $\{v(t), t \in T\}$ 是具有单位方差参数的维纳过程。试证明, 可以采用下列步骤, 用模拟计算机计算 y 的协方差函数: 对具有初始条件为

$$z(1) = b$$

的微分方程

$$\frac{dz}{dt} = Az$$

进行积分, 再计算由

$$r_i = \int_0^t z_i(t) z_i(t) dt \quad (i = 1, 2, \dots, n)$$

定义的向量 r 的各个分量, 那么, 协方差函数 $r_y(t)$ 就给定

$$\begin{aligned} \frac{dz^r}{dt} &= Az \\ z^r(0) &= r \\ r_y(t) &= z_i^r(t) \end{aligned}$$

3.11 试证明, 计算随机微分方程

$$dx = Ax dt + dv$$

的状态的协方差阵 R 的问题就是计算初始条件为

$$x(0) = b$$

的动力学系统

$$\frac{dx}{dt} = A^T x$$

的二次指标函数

$$V = \int_0^t x^T(t) R x(t) dt$$

的对偶问题。

3.12 液体中的自由粒子能够用所谓兰杰文方程来描述:

$$m \frac{dv}{dt} + f v = K(t)$$

式中: m 是质量; f 是粘滞摩擦因数; K 是由于与液体其他分子碰撞产生的波动力, 波动力 K 的平均值为零, 并有比 $\frac{m}{f}$ 更迅速趋于零的协方差函数, 因此可以把波动力 K 看做是白噪声过程。由统计力学的均分定律可知, 波动力的强度使得

$$\frac{1}{2} m E\{v^2\} = \frac{1}{2} kT$$

式中: k 是玻耳兹曼常数; T 是周围介质的热力学温度。假设此粒子开始处于静止状态, 试确定作为时间函数的速度的概率分布。

3.13 当随机微分方程

$$dx = Ax dt + dv$$

的解由下式定义:

$$x(t) = \Phi(t, t_0) x(t_0) + \int_{t_0}^t \Phi(t, s) dv(s)$$

式中:

$$\begin{aligned} \frac{d}{dt} \Phi(t, t_0) &= A(t) \Phi(t, t_0) \\ \Phi(t_0, t_0) &= I \end{aligned}$$

而解中的积分又定义为

$$\int_{t_0}^t \Phi(t, s) dv(s) = v(t) - \Phi(t, t_0) v(t_0) - \int_{t_0}^t \frac{d}{ds} \Phi(t, s) v(s) ds$$

试证明微分方程的解是一个均值为 $m_x(t)$ 和协方差为 $R(s, t)$ 的高斯过程, 其中

$$\begin{aligned}\frac{dm_x}{dt} &= Am_x \\ m_x(t_0) &= m_0 \\ R(s, t) &= (s, t)P(t) - (s - t) \\ \frac{dP}{dt} &= AP + PA^T + R_1 \\ P(t_0) &= R_0\end{aligned}$$

3.14 考虑初始条件为零的微分方程

$$\begin{aligned}\frac{dx_1}{dt} &= e \\ \frac{dx_2}{dt} &= x_1e\end{aligned}$$

设 $\{e(t), t \in T\}$ 是均值为零和协方差函数为 $r(\cdot)$ 的平稳随机过程。试计算 x_2 的均值。当 $r(\cdot)$ 趋于 $\delta(\cdot)$ 分布时, 再计算这个均值的极限。

3.15 考虑随机微分方程

$$\begin{aligned}dx_1 &= dw \\ dx_2 &= x_1dw\end{aligned}$$

式中: $\{w(t), t \in T\}$ 是维纳过程。设 x_1 和 x_2 的初值为零, 则微分方程的解为

$$\begin{aligned}x_1(t) &= w(t) \\ x_2(t) &= \int_0^t w(s)dw(s)\end{aligned}$$

定义上述积分

$$\begin{aligned}I &= \int_0^t w(s)dw(s) \\ &= (1 - \frac{1}{2})I_0 + I_1 - \frac{1}{2}(0 - 1)\end{aligned}$$

式中:

$$\begin{aligned}I_0 &= \lim_{N \rightarrow \infty} \sum_{i=1}^N w(t_i) [w(t_{i+1}) - w(t_i)] \\ I_1 &= \lim_{N \rightarrow \infty} \sum_{i=1}^N w(t_{i+1}) [w(t_{i+1}) - w(t_i)]\end{aligned}$$

试计算 x_2 的均值, 并特别考虑 $\rho = 0, 0.5$ 和 1 等几种情况。

3.16 考虑随机差分方程

$$\begin{aligned}x_1(t) &= w(t) \\ x_2(t) &= x_1(t) - w(t)\end{aligned}$$

式中: $w(t)$ 是维纳过程。当

$$\begin{aligned}f(t) &= f(t+h) - f(t) \\ f(t) &= f(t) - f(t-h) \\ f(t) &= \frac{1}{2}[f(t+h) - f(t-h)]\end{aligned}$$

时, 试计算 x_1 和 x_2 的均值。

3.17 设 $\{e(t), t \in T\}$ 是均值为 m_e 和协方差函数为 $r(\cdot)$ 的平稳随机过程。考虑具有初值

$$\begin{aligned}x_1(0) &= 0 \\x_2(0) &= 0\end{aligned}$$

的差方程

$$\begin{aligned}x_1(t+h) &= x_1(t) + he(t) \\x_2(t+h) &= x_2(t) + hx_1(t)e(t) \quad (t = 0, h, 2h, \dots)\end{aligned}$$

试计算 $x_1(nh), x_2(nh)$ 和它们的均值, 再计算当 $h \rightarrow 0$ 使得 nh 为常数时的 $E\{x_2(nh)\}$ 的极限, 最后计算当 $h \rightarrow 0$ 时

$$r(h) = \begin{cases} r(0) & (h = 0) \\ 0 & (h \neq 0) \end{cases}$$

的 $E\{x_2(nh)\}$ 的极限。

3.18 与周围环境处于热平衡状态的电流计的运动可用下面方程描述:

$$J \frac{d^2}{dt^2} + D \frac{d}{dt} + C = M$$

式中: 转矩 M 是由空气分子碰撞反光镜所产生的。试在转矩 M 可用白噪声模拟的条件下, 确定电流计偏转角的协方差和谱密度。注意: 均分定律给出

$$\frac{1}{2} J E \left(\frac{d}{dt} \right)^2 = \frac{1}{2} C E \{ \theta^2 \} = \frac{1}{2} kT$$

3.19 单轴回转稳定台可用以下近似模型描述:

$$\begin{aligned}D \frac{d}{dt} &= m + H \\&= -c\end{aligned}$$

式中: θ 是回转输出信号; ϕ 是台架的角偏差; D 是回转粘滞阻尼系数; H 是回转角动量矩; 而 m 是由分子热运动产生的自由回转的干扰转矩, 假设 m 可用白噪声模拟。试确定由热起伏产生的台架偏转角 ϕ 的方差, 并证明 t 的值极大时, 能得到

$$E\{\phi^2\} = 2kTDH^{-1}t$$

式中: T 是热力学温度; k 是玻耳兹曼常数。当 $D = 0.03 \text{ kg} \cdot \text{m}^2/\text{s}$, $H = 0.01 \text{ kg} \cdot \text{m}^2/\text{s}$ 时, 试计算在 $t = 1 \text{ s}$ 和 1 h 的台架偏移的标准差。注意, 均分定律给出

$$\frac{1}{2} c H E \{ \phi^2 \} = \frac{1}{2} kT$$

3.20 考虑随机微分方程

$$\begin{aligned}dx &= Ax dt + dv \\dy &= Cx dt + de\end{aligned}$$

其中维纳过程 $\{v(t), t \in T\}$ 和 $\{e(t), t \in T\}$ 通过联合增量协方差

$$E \int_{de}^{dv} [dv^T de^T] = \begin{bmatrix} R_1 & R_{12} \\ R_{12}^T & R_2 \end{bmatrix} dt$$

相联系。试证明给定方程的采样型由

$$\begin{aligned}x(t_{i+1}) &= x(t_{i+1}, t_i) + \tilde{v}(t_i) \\ y(t_{i+1}) &= y(t_i) + x(t_{i+1}, t_i) x(t_i) + \tilde{e}(t_i)\end{aligned}$$

给定, 其中 $\tilde{e}(t_i)$ 和 $\tilde{v}(t_i)$ 的均值为零, 协方差为

$$\begin{aligned}E\{\tilde{v}(t_i)\tilde{v}^T(t_i)\} &= \tilde{R}_1(t_i) = \int_{t_i}^{t_{i+1}} (t_{i+1}, s) R_1(s)^T (t_{i+1}, s) ds \\ E\{\tilde{v}(t_i)\tilde{e}^T(t_i)\} &= \tilde{R}_{12}(t_i) = \int_{t_i}^{t_{i+1}} (t_{i+1}, s) [R_1(s)^T (t_{i+1}, s) + R_{12}(s)] ds \\ E\{\tilde{e}(t_i)\tilde{e}^T(t_i)\} &= \tilde{R}_2(t_i) = \int_{t_i}^{t_{i+1}} [(t_{i+1}, s) R_1(s)^T (t_{i+1}, s) \\ &\quad + (t_{i+1}, s) R_{12}(s) + R_{12}^T(s)^T (t_{i+1}, s) + R_2(s)] ds\end{aligned}$$

3.21 考虑具有

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, C = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, R_1 = \begin{bmatrix} q_1 & 0 & 0 \\ 0 & q_2 & 0 \\ 0 & 0 & r \end{bmatrix}, R_{12} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, R_2 = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}$$

的系统, 试确定采样区间为 h 的这个系统的采样型参量 $A, B, \tilde{R}_1, \tilde{R}_{12}$ 和 \tilde{R}_2 。

3.22 考虑随机微分方程

$$\begin{aligned}dx &= Axdt + dv \\ dy &= Cxdt + de\end{aligned}$$

式中: $\{v(t), -\infty < t < \infty\}$ 和 $\{e(t), -\infty < t < \infty\}$ 是增量方差分别为 R_1dt 和 R_2dt 的维纳过程; $v(t)$ 和 $e(t)$ 的互协方差为 $R_{12}dt$ 。定义

$$y(t_{i+1}) = y(t_i) + \int_{t_i}^{t_{i+1}} dy(s) \quad (t_i < t < t_{i+1})$$

试确定微分方程相应的采样型。

3.23 在一般情况下, 平稳动力学系统的平稳单输入 u 和单输出 y 的互协方差函数定义为

$$r_{uy}(\tau) = E\{[x(t+\tau) - E\{x(t+\tau)\}][y(t) - E\{y(t)\}]\}$$

它们的互谱密度为

$$u_y(\omega) = H(e^{-j\omega}) u(\omega)$$

但互协方差函数也可定义为

$$r_{uy}^*(\tau) = E\{[x(t) - E\{x(t)\}][y(t+\tau) - E\{y(t+\tau)\}]\}$$

试证明在这种情况下, 互谱密度变为

$$u_y^*(\omega) = H(e^{j\omega}) u(\omega)$$

3.24 考虑时变系统

$$\begin{aligned}y(t) &= \int_{s=-\infty}^t g(t, s) u(s) ds \\ &= \sum_{k=0}^{\infty} g(t, t-k) u(t-k)\end{aligned}$$

式中: 输入 $\{u(s), s= \dots, -2, -1, 0, 1, 2, \dots\}$ 是协方差函数为 $r_u(s, t)$ 的平稳过程。试证明

在适当条件下, 输出 $y(t)$ 的协方差函数存在, 并且为

$$r_y(s, t) = \sum_{k=-\infty}^s \sum_{l=-\infty}^t g(s, k) r_u(k, l) g(t, l)$$

3.25 设 $\{u(t), t \in T\}$ 和 $\{y(t), t \in T\}$ 是由方程

$$y(t) + a(t-1)y(t-1) = u(t)$$

相联系的两个随机过程, 试证明

$$\begin{aligned} r_{uy}(s, t) + a(t-1)r_{uy}(s, t-1) &= r_u(s, t) \\ r_{yu}(s, t) + a(s-1)r_{yu}(s-1, t) &= r_u(s, t) \\ r_y(s, t) + a(t-1)r_y(s, t-1) &= r_{yu}(s, t) \end{aligned}$$

3.26 把上一题推广到

$$\begin{aligned} y(t) + a_1(t-1)y(t-1) + a_2(t-2)y(t-2) + \dots + a_n(t-n)y(t-n) \\ = u(t) + b_1(t-1)u(t-1) + b_2(t-2)u(t-2) + \dots + b_n(t-n)u(t-n) \end{aligned}$$

时的情况。

3.27 考虑由下式定义的平稳过程 $\{y(t), t \in T\}$:

$$y(t) + ay(t-1) = e(t) + ce(t-1) \quad (|a| < 1)$$

式中: $\{e(t), t \in T\}$ 是协方差函数为

$$r_e(\tau) = e^{-a|\tau|}$$

的平稳过程。试确定 e 的谱密度 $f_e(\omega)$, y 的谱密度 $f_y(\omega)$, 以及 e 和 y 之间的互谱密度 $f_{ey}(\omega)$ 。

3.28 一平稳离散时间随机过程有谱密度

$$f(\omega) = \frac{2 + 2\cos\omega}{5 + 4\cos\omega}$$

试进行谱分解, 并且在输入为白噪声的条件下, 确定使输出的谱密度为 $f(\omega)$ 的稳定系统的脉冲传递函数。

3.29 考虑平移平均过程

$$y(t) = e(t) + 4e(t-1)$$

式中: $e(t)$ 是独立同分布高斯 $N(0, 1)$ 随机变量序列。试证明可由

$$y(t) = [x(t) + c(x(t-1))]$$

产生具有相同谱密度的过程, 已知 x 是独立高斯 $N(0, 1)$ 随机变量序列, 且 $|c| < 1$ 。试确定参数 c 和 σ 。

3.30 考虑由以下式形成的高斯随机过程 $y_1(t)$:

$$y_1(t) = x_1(t) + x_2(t)$$

式中:

$$\begin{aligned} x_1(t+1) &= -ax_1(t) + v_1(t) \\ x_2(t+1) &= -bx_2(t) + v_2(t) \end{aligned}$$

$v_1(t)$ 和 $v_2(t)$ 分别是独立同分布高斯 $N(0, \sigma_1^2)$ 和 $N(0, \sigma_2^2)$ 随机变量序列。试证明具有相同谱密度的随机过程可以表示为

$$y_2(t) = \frac{q + c}{(q + a)(q + b)} e(t)$$

式中: q 是移位算子

$$y(t + 1) = qy(t)$$

而 $e(t)$ 是独立同分布高斯 $N(0, 1)$ 随机变量序列。试确定参数 a 和 c 。

3.31 试证明由

$$x(t + 1) = 0.8x(t) - 1.2e(t)$$

$$y(t) = x(t) + e(t)$$

和

$$x(t + 1) = 0.8x(t) + 0.6e(t)$$

$$z(t) = x(t) + 2e(t)$$

表示的随机过程 $y(t)$ 和 $z(t)$ 具有相同的谱密度, 已知 $e(t)$ 是独立高斯 $N(0, 1)$ 随机变量序列。

3.32 试对连续时间随机过程 $y(t)$ 的谱密度

$$y(\omega) = \frac{\omega^2 + 1}{\omega^4 + 8\omega^2 + 4}$$

进行谱分解。设 $y(\omega)$ 是在输入谱密度

$$u(\omega) = 1$$

的条件下求得的, 试确定系统的传递函数 $G(s)$ 和加权函数 $h(t)$ 。

3.33 某平稳随机过程有协方差函数

$$r(\tau) = e^{-\alpha|\tau|} \cos 2\tau$$

试求出形式为

$$y(t) = \int_{-\infty}^t h(t - s) dv(s)$$

的过程的表示, 已知 $\{v(t), t \in T\}$ 是均值为零和增量方差为 dt 的正交增量过程。

3.34 随机过程 $\{x(t), t \in T\}$ 和 $\{y(t), t \in T\}$ 是平稳的和高斯的, 它们的谱密度为

$$x(\omega) = \frac{1}{\omega^2 + 1}$$

$$y(\omega) = \frac{1}{\omega^2 + 4}$$

$$x_y(\omega) = \frac{1}{\omega^2 + i\omega + 2}$$

试给出向量过程 $\begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$ 的过程表示。并确定把 $x(t)$ 作为输入和把 $y(t)$ 作为输出的动力学系统的传递函数 $G(s)$ 。

3.35 设具有实系数的 n 阶多项式 $A(s)$ 的全部零点都在左半平面上, 试证明多项式

$$\tilde{A}(s) = \frac{1}{2} [A(s) - (-1)^n A(-s)]$$

的全部零点都在虚轴上。

3.36 考虑两个平稳随机过程, 其谱密度为

$$\begin{aligned}x(\omega) &= G_1(i\omega) G_1(-i\omega) \\x_y(\omega) &= G_1(i\omega) G_2(-i\omega) \\y(\omega) &= G_2(i\omega) G_2(-i\omega)\end{aligned}$$

式中：

$$\begin{aligned}G_1(s) &= \frac{1}{s^2 + 2s + 2} \\G_2(s) &= \frac{s}{s^2 + 2s + 2}\end{aligned}$$

试确定 $E\{x^2\}$, $E\{y^2\}$, $E\{xy\}$ 。

第 4 章 最小方差控制

4.1 引言

本章研究受控自回归平移平均模型(CARMA)的预测和控制问题,因为控制的性能指标是使系统输出的方差为最小,故称为最小方差控制。最小方差控制的目标,是对给定 CARMA 模型

$$A(q^{-1})y(k) = B(q^{-1})u(k-d) + C(q^{-1})e(k) \tag{4.1.1}$$

寻找一个控制策略,使系统输出的方差为最小。图 4.1.1 给出了 CARMA 的框图。

CARMA 是离散时间随机系统,它从系统的输入和输出关系解决随机控制问题,具有简单、易于使用计算机等特点,因而较早地得到了实际应用。本章主要讨论单输入和单输出的随机控制问题。

造纸过程的基重控制是最小方差控制最早的一个实例。基重是指单位面积纸张的重量(g/m^2),在造纸的生产过程中,由于各种噪声的影响,基重是随机变量,分布是高斯的,如图 4.1.2 所示。对基重要求,有一个检验界限 T ,基重大于 T 时为合格,小于 T 的数量不得超过一定限度,否则为不合格。不加控制时,为满足要求,设定值不得不设在较大位置 s_0 ,基重方差大,不仅浪费原料,而且纸张质量差。若加入控制,能使基重方差大大缩小,把设定值选在较低位置 s_c ,就能满足要求,既节省了原料,又提高了纸张质量。

图 4.1.1 CARMA 框图

图 4.1.2 基重方差与检验界限的关系

本章先通过一个简例,来说明最小方差控制的基本概念,再介绍一般的预测和控制问题,最后讨论系统的灵敏度。

4.2 简 例

这一节举一个例子,例子虽然简单,但包括了最小方差控制的基本概念,对理解本章内容很有帮助。

4.2.1 例

例 4.2.1 已知 CARMA 模型

$$y(k) + ay(k-1) = u(k-1) + e(k) + ce(k-1) \quad (4.2.1)$$

式中: $y(k)$ 是系统输出,它既是被控量,又是量测量,其均值为零,可设想 $y(k)$ 的设定值为零; $u(k)$ 是控制量,它是 $y(k), y(k-1), \dots, u(k-1), u(k-2), \dots$ 的函数,即

$$u(k) = f[y(k), y(k-1), \dots, u(k-1), u(k-2), \dots] \quad (4.2.2)$$

系数 $a, c \neq 1, \neq 1$ 。控制目标是寻找一个控制策略 $u(k)$, 使性能指标

$$J = E\{y^2(k)\} \quad (4.2.3)$$

为最小。由于已设定系统输出 $y(k)$ 为零, 因此 $E\{y^2(k)\}$ 就是 $y(k)$ 的方差, 这种控制称为最小方差控制。

解 考虑 k 时刻系统输出 $y(k)$ 已被量测, 是已知量, 要求对 $y(k+1)$ 进行控制。为此, 把式(4.2.1)写为 $y(k+1)$ 的形式, 并对 $y(k+1)$ 取方差:

$$y(k+1) = e(k+1) + u(k) - ay(k) + ce(k) \quad (4.2.4)$$

$$\begin{aligned} J &= E\{y^2(k+1)\} \\ &= E\{[e(k+1) + u(k) - ay(k) + ce(k)]^2\} \\ &= E\{e^2(k+1)\} + E\{[u(k) - ay(k) + ce(k)]^2\} \\ &\quad + 2E\{e(k+1)[u(k) - ay(k) + ce(k)]\} \end{aligned} \quad (4.2.5)$$

式(4.2.5)表示的性能指标与式(4.2.3)的意义是相同的。注意到 $e(k+1)$ 是 $y(k+1)$ 的新息, 与 $[u(k) - ay(k) + ce(k)]$ 独立, 由此可得到式(4.2.5)右边第三式第三项为零。在 k 时刻, $y(k)$ 和 $e(k)$ 都是已确定的, $u(k)$ 为待求的确定量, 因此, 式(4.2.5)右边第三式第二项的均值符号可以去掉, 得到

$$J = E\{e^2(k+1)\} + [u(k) - ay(k) + ce(k)]^2 \quad (4.2.6)$$

上式右边两项都是正的, 第一项无法控制, 因此, 当第二项为零时, 式(4.2.6)达到最小, 即得到

$$u(k) - ay(k) + ce(k) = 0 \quad (4.2.7)$$

$$\min J = E\{e^2(k+1)\} \quad (4.2.8)$$

把式(4.2.7)代入式(4.2.4), 得到

$$y(k+1) = e(k+1) \quad (4.2.9)$$

在上式中用 k 代替 $k+1$, 并把它代入式(4.2.7), 得到最小方差控制策略为

$$u(k) = (a - c)y(k) \quad (4.2.10)$$

综上所述, 对式(4.2.4)进行控制, 最小方差控制策略如式(4.2.10)所示, 控制误差如式(4.2.9)所示, 控制方差, 即性能指标如式(4.2.8)所示。

4.2.2 结果验证

用 z 变换方法验证对式(4.2.4)实行式(4.2.10)控制时,是否能使系统输出达到式(4.2.8)所示的最小方差。为此,把式(4.2.10)代入式(4.2.4),得到

$$[y(k+1) - e(k+1)] + c[y(k) - e(k)] = 0 \quad (4.2.11)$$

对上式进行 z 变换:

$$z[y(z) - e(z)] - z[y(k_0) - e(k_0)] + c[y(z) - e(z)] = 0 \quad (4.2.12)$$

$$y(z) - e(z) = [y(k_0) - e(k_0)] \frac{z}{z + c} \quad (4.2.13)$$

式中: $[y(k_0) - e(k_0)]$ 为初值。对式(4.2.13)进行反变换,考虑到采样时间 $T=1$ 时, $\frac{z}{z+c}$ 对应 $(-c)^{k-k_0}$, 得到

$$y(k) - e(k) = [y(k_0) - e(k_0)](-c)^{k-k_0} \quad (4.2.14)$$

定态时, 即 $k \rightarrow \infty$ 时, $(-c)^{k-k_0} \rightarrow 0$, 有

$$y(k) = e(k) \quad (4.2.15)$$

$$E\{y^2(k)\} = E\{e^2(k)\} \quad (4.2.16)$$

得到了与式(4.2.8)相同的结果, 这就验证了式(4.2.8)确实是最小方差控制策略。

4.2.3 几点说明

(1) 分离定理

可把式(4.2.4)看做是一步预测模型

$$y(k+1) = \hat{y}(k+1|k) + e(k+1) \quad (4.2.17)$$

式中: $\hat{y}(k+1|k)$ 称为 $y(k)$ 的一步预测。对照式(4.2.4), 其值为

$$\hat{y}(k+1|k) = u(k) - ay(k) + e(k) \quad (4.2.18)$$

那么, 一步预测误差 $\tilde{y}(k+1|k)$ 为

$$\tilde{y}(k+1|k) = y(k+1) - \hat{y}(k+1|k) = e(k+1) \quad (4.2.19)$$

为实现控制, 当系统输出 $y(k)$ 的初值为零时, 应使控制误差

$$y(k+1) = \hat{y}(k+1|k) \quad (4.2.20)$$

等于预测误差

$$y(k+1) = \tilde{y}(k+1|k) \quad (4.2.21)$$

这种控制就是最小方差控制。

由此可知, 控制问题与预测问题紧密相关, 控制问题由两部分组成, 可分别进行计算。一部分是预测器, 求出预测函数和预测误差。另一部分是最优控制策略, 由预测函数计算出使控制误差等于预测误差的控制策略, 这个控制策略就是最小方差控制策略。把控制问题分为预测和控制两部分, 称为分离定理。

(2) 系统动力学

由式(4.2.14)看出, 最优系统的动力学由参数 c 唯一确定, 定态时系统输出等于白噪声。作为噪声项, 必有 $|c| < 1$, 否则不稳定, 是非最小相问题。

(3) 放松 $e(k)$ 的高斯假设

当要求控制策略限制为量测结果的线性函数时, $e(k)$ 的高斯假设可放松。

4.3 ARMA 模型的最优预测

本节分析 ARMA 模型的一般最优预测问题。用量测值(观测值) $[y(k), y(k-1), \dots, y(k_0)]$ 来预测 $y(k+p)$ 的值, 称为 p 步预测, 表示为 $\hat{y}(k+p|k)$, p 可为 $1, 2, \dots$, p 步预测误差 $\tilde{y}(k+p|k)$ 为 p 步真值与 p 步预测值之差, 即

$$\tilde{y}(k+p|k) = y(k+p) - \hat{y}(k+p|k) \quad (4.3.1)$$

性能指标为

$$J = E\{\tilde{y}^2(k+p|k)\} \quad (4.3.2)$$

预测的目标是寻求一种预测 $\hat{y}(k+p|k)$, 使预测误差的方差式(4.3.2)为最小。

本节先举一个简例, 然后再求解一般问题。

4.3.1 简例

例 4.3.1 已知 ARMA 模型

$$y(k+1) + ay(k) = e(k+1) + ce(k) \quad (4.3.3)$$

式中: $|a| < 1$; $|c| < 1$ 。求使一步预测误差的方差 $J = E\{\tilde{y}^2(k+1|k)\}$ 为最小的一步预测 $\hat{y}(k+1|k)$ 。

解 为了解系统动力学和求解一般问题的思路, 这里用两种方法求解。

方法 1

改写式(4.3.3)为

$$y(k+1) = -ay(k) + ce(k) + e(k+1) \quad (4.3.4)$$

上式右边第一项 $y(k)$ 是量测值, 为已知量; 第二项 $e(k)$ 可由量测值 $[y(k), y(k-1), \dots, y(k_0)]$ 精确表示, 也为已知量; 而第三项 $e(k+1)$ 是新息, 与前两项独立。因此, 容易写出

$$\hat{y}(k+1|k) = -ay(k) + ce(k) \quad (4.3.5)$$

$$\tilde{y}(k+1|k) = e(k+1) \quad (4.3.6)$$

式(4.3.5)中存在 $e(k)$ 项, 需要用量测值 $[y(k), y(k-1), \dots, y(k_0)]$ 表示, 为此, 把式(4.3.3)化为

$$[e(k+1) - y(k+1)] + c[e(k) - y(k)] = (a - c)y(k) \quad (4.3.7)$$

对上式两边进行 z 变换, 得

$$z[e(z) - y(z)] - z[e(k_0) - y(k_0)] + c[e(z) - y(z)] = (a - c)y(z) \quad (4.3.8)$$

$$e(z) = y(z) + (a - c) \frac{z}{z + c} z^{-1} y(z) + \frac{z}{z + c} [e(k_0) - y(k_0)] \quad (4.3.9)$$

对上式进行 z 的反变换, 注意到 $\frac{z}{z + c}$ 对应 $(-c)^{k-k_0}$, $z^{-1} y(z)$ 对应 $y(k-1)$, 以及

$F_1(z)F_2(z)$ 对应 $\sum_{n=k_0}^k f(k-n)f_2(n)$, 得到

$$e(k) = y(k) + (a - c) \sum_{l=k_0}^k (-c)^{k-l} y(l-1) + [e(k_0) - y(k_0)](-c)^{k-k_0} \quad (4.3.10)$$

对上式右边第二项进行变量置换 $n = l-1$, 并考虑到 $y(k_0-1) = 0$, 则得到

$$e(k) = y(k) + (a - c) \sum_{n=k_0}^{k-1} (-c)^{k-n-1} y(n) + [e(k_0) - y(k_0)](-c)^{k-k_0} \quad (4.3.11)$$

定态时, 即 $k = k_0$ 时, $(-c)^{k-k_0} = 0$, 有

$$e(k) = y(k) + (a - c) \sum_{n=k_0}^{k-1} (-c)^{k-n-1} y(n) \quad (4.3.12)$$

把求得的 $e(k)$ 表示式(4.3.12)代入式(4.3.5), 得到结果

$$\hat{y}(k+1|k) = (c - a) \sum_{n=k_0}^k (-c)^{k-n} y(n) \quad (4.3.13)$$

虽然上面导出了一步预测公式(4.3.13), 但使用不方便, 因为为了计算, 不得不在计算机中存储大量量测数据。

为解决这个困难, 稍微改写式(4.3.13), 注意到

$$\hat{y}(k|k-1) = (c - a) \sum_{n=k_0}^{k-1} (-c)^{k-n-1} y(n) \quad (4.3.14)$$

就能得到递推公式

$$\begin{aligned} \hat{y}(k+1|k) &= (c - a)y(k) - c(c - a) \sum_{n=k_0}^{k-1} (-c)^{k-n-1} y(n) \\ &= -c \hat{y}(k|k-1) + (c - a)y(k) \end{aligned} \quad (4.3.15)$$

上述方法中系统动力学特性表示清楚, 但推导复杂, 下面使用简单方法计算。

方法 2

引入后向移位算子 q^{-1} , 使 $y(k-1) = q^{-1}y(k)$, 则式(4.3.3)可改写为

$$(1 + aq^{-1})y(k+1) = (1 + cq^{-1})e(k+1) \quad (4.3.16)$$

即

$$\begin{aligned} y(k+1) &= \frac{1 + cq^{-1}}{1 + aq^{-1}} e(k+1) \\ &= e(k+1) + \frac{c - a}{1 + aq^{-1}} e(k) \end{aligned} \quad (4.3.17)$$

同理, 有

$$(1 + aq^{-1})y(k) = (1 + cq^{-1})e(k) \quad (4.3.18)$$

$$e(k) = \frac{1 + aq^{-1}}{1 + cq^{-1}} y(k) \quad (4.3.19)$$

把上式代入式(4.3.17), 得到

$$y(k+1) = e(k+1) + \frac{c - a}{1 + cq^{-1}} y(k) \quad (4.3.20)$$

下一步计算性能指标, 即计算一步预测误差的方差, 注意到 $e(k+1)$ 与其他项独立, 得到

$$\begin{aligned} J &= E\{\tilde{y}^2(k+1|k)\} \\ &= E\{[y(k+1) - \hat{y}(k+1|k)]^2\} \end{aligned}$$

$$\begin{aligned}
&= E\left\{\left[e(k+1) + \frac{c-a}{1+cq^{-1}}y(k) - \hat{y}(k+1|k)\right]^2\right\} \\
&= E\{e^2(k+1)\} + E\left\{\frac{c-a}{1+cq^{-1}}y(k) - \hat{y}(k+1|k)\right\}^2
\end{aligned} \quad (4.3.21)$$

上式最后一个等式右边第二项中都是确定量, 均值符号可以去掉。不难看出, 第二项为零时, J 取最小值, 即

$$\begin{aligned}
\hat{y}(k+1|k) &= \frac{c-a}{1+cq^{-1}}y(k) \\
&= -c\hat{y}(k|k-1) + (c-a)y(k)
\end{aligned} \quad (4.3.22)$$

$$\min J = E\{e^2(k+1)\} \quad (4.3.23)$$

把式(4.3.22)代入式(4.3.17), 得到一步预测误差

$$\tilde{y}(k+1|k) = e(k+1) \quad (4.3.24)$$

两种方法的结果即式(4.3.22)和式(4.3.15)是相同的, 但从推导过程看, 用后一种方法, 在引入后向移位算子后, 计算简化了。在 4.3.2 节中通过引入后向移位算子, 求解一般的 p 步预测问题。

4.3.2 p 步预测

考虑一般 ARMA 模型

$$A(q^{-1})y(k) = C(q^{-1})e(k) \quad (4.3.25)$$

式中:

$$A(q^{-1}) = 1 + a_1q^{-1} + \dots + a_nq^{-n} \quad (4.3.26)$$

$$C(q^{-1}) = 1 + c_1q^{-1} + \dots + c_nq^{-n} \quad (4.3.27)$$

$y(k)$ 是系统输出, 它既是量测值又是被预测量; $e(k)$ 是独立同分布高斯 $N(0, 1)$ 随机序列, 即离散时间高斯白噪声; a_i, c_i 是常数; $A(q^{-1})$ 和 $C(q^{-1})$ 都是 n 阶首 1 多项式, 其零点都在单位圆内, 两个多项式都取 n 阶, 具有一般性, 对于小于 n 阶的多项式, 只要令有关系数为零即可, a_1, a_2, \dots, a_n 和 c_1, c_2, \dots, c_n 都为已知系数。

现要求对式(4.3.25)求出 p 步预测 $\hat{y}(k+p|k)$, 使性能指标

$$J = E\{\tilde{y}^2(k+p|k)\} \quad (4.3.28)$$

达到最小。

仿照例题 4.3.1 方法 2 求解。把式(4.3.25)改写为

$$e(k) = \frac{A(q^{-1})}{C(q^{-1})}y(k) \quad (4.3.29)$$

$$y(k+p) = \frac{C(q^{-1})}{A(q^{-1})}q^pe(k) \quad (4.3.30)$$

把式(4.3.20)右边分为独立的两部分, 为此, 分解 $C(q^{-1})$:

$$C(q^{-1}) = A(q^{-1})F(q^{-1}) + q^{-p}G(q^{-1}) \quad (4.3.31)$$

式中:

$$F(q^{-1}) = 1 + f_1q^{-1} + f_2q^{-2} + \dots + f_{p-1}q^{-(p-1)} \quad (4.3.32)$$

$$G(q^{-1}) = g_0 + g_1q^{-1} + g_2q^{-2} + \dots + g_{n-1}q^{-(n-1)} \quad (4.3.33)$$

多项式 $F(q^{-1})$ 是 $A(q^{-1})$ 除 $C(q^{-1})$ 的商, 而多项式 $q^{-p}G(q^{-1})$ 是余式。 $F(q^{-1})$ 可取任意阶, 但对 p 步预测, 一定要取为 $p-1$ 阶。当 $F(q^{-1})$ 取 $p-1$ 阶时, $G(q^{-1})$ 必为 $n-1$ 阶。 $F(q^{-1})$ 之所以取 $p-1$ 阶, 是因为只有这样做, 才能把式(4.3.29)右边分为独立的两部分。 $F(q^{-1})$ 和 $G(q^{-1})$ 可用长除法或比较系数法确定, 将在下一小节加以介绍。

把式(4.3.31)代入式(4.3.30), 得到

$$y(k+p) = F(q^{-1})e(k+p) + \frac{G(q^{-1})}{A(q^{-1})}e(k) \quad (4.3.34)$$

再把式(4.3.29)代入式(4.3.34), 消去 $e(k)$ 项, 得到

$$y(k+p) = F(q^{-1})e(k+p) + \frac{G(q^{-1})}{C(q^{-1})}y(k) \quad (4.3.35)$$

注意到

$$F(q^{-1})e(k+p) = e(k+p) + f_1e(k+p-1) + f_2e(k+p-2) + \dots + f_{p-1}e(k+1) \quad (4.3.36)$$

是 $y(k+p)$ 的新息, 与 $y(k)$ 是独立的, 式(4.3.35)右边确实把 $y(k+p)$ 分成了独立的两部分。

把式(4.3.35)代入性能指标式(4.3.28), 并加以计算, 得到

$$\begin{aligned} J &= E\{[y(k+p) - \hat{y}(k+p|k)]^2\} \\ &= E\left\{F(q^{-1})e(k+p) + \frac{G(q^{-1})}{C(q^{-1})}y(k) - \hat{y}(k+p|k)\right\}^2 \\ &= E\{[F(q^{-1})e(k+p)]^2\} + \frac{G(q^{-1})}{C(q^{-1})}y(k) - \hat{y}(k+p|k) \quad (4.3.37) \end{aligned}$$

上式右边两项都是正的, 只有第二项取零时, 性能指标 J 取最小值, 即

$$\begin{aligned} \hat{y}(k+p|k) &= \frac{G(q^{-1})}{C(q^{-1})}y(k) \\ &= -c_1\hat{y}(k+p-1|k-1) - c_2\hat{y}(k+p-2|k-2) - \dots \\ &\quad - c_n\hat{y}(k+p-n|k-n) + g_0y(k) + g_1y(k-1) + \dots \\ &\quad + g_{n-1}y(k-n+1) \end{aligned} \quad (4.3.38)$$

$$\begin{aligned} \min J &= E\{[F(q^{-1})e(k+p)]^2\} \\ &= \sigma^2(1 + f_1^2 + f_2^2 + \dots + f_{p-1}^2) \end{aligned} \quad (4.3.39)$$

把式(4.3.38)代入式(4.3.35), 得到 k 步预测误差

$$\begin{aligned} \tilde{y}(k+p|k) &= F(q^{-1})e(k+p) \\ &= [e(k+p) + f_1e(k+p-1) + f_2e(k+p-2) \\ &\quad + \dots + f_{p-1}e(k+1)] \end{aligned} \quad (4.3.40)$$

式(4.3.38), (4.3.39)和(4.3.40)是预测结果。 p 步预测是一个递推式, p 步预测误差由噪声序列组成, 性能指标由噪声序列方差组成。

4.3.3 $F(q^{-1})$ 和 $G(q^{-1})$ 的计算

1. 长除法

把式(4.3.31)写为

$$\frac{C(q^{-1})}{A(q^{-1})} = F(q^{-1}) + \frac{q^{-p}G(q^{-1})}{A(q^{-1})} \quad (4.3.41)$$

如果 $A(q^{-1})$ 和 $C(q^{-1})$ 均为 n 阶首 1 多项式, 选中 p 值, 用 $A(q^{-1})$ 去除 $C(q^{-1})$, 直到商为 $p-1$ 阶多项式为止, 就可唯一确定: 商为 $F(q^{-1})$, 余式为 $q^{-p}G(q^{-1})$, 分别表示为式 (4.3.32) 和 (4.3.33)。

例 4.3.2 已知 3 阶多项式

$$A(q^{-1}) = 1 + 1.2q^{-1} + 0.11q^{-2} - 0.168q^{-3} \tag{4.3.42}$$

$$C(q^{-1}) = 1 - 0.7q^{-1} - 0.14q^{-2} + 0.12q^{-3} \tag{4.3.43}$$

给定 $p=2$, 试按式 (4.3.41) 确定 $F(q^{-1})$ 和 $G(q^{-1})$ 。

解 由给定条件 $n=3, p=2$, 则有

$$F(q^{-1}) = 1 + f_1q^{-1} \tag{4.3.44}$$

$$q^{-p}G(q^{-1}) = q^{-2}(g_0 + g_1q^{-1} + g_2q^{-2}) \tag{4.3.45}$$

用 $A(q^{-1})$ 去除 $C(q^{-1})$:

$$\begin{array}{r} 1 + 1.2q^{-1} + 0.11q^{-2} - 0.168q^{-3} \quad \overline{) 1 - 1.9q^{-1} - 0.25q^{-2} + 0.288q^{-3}} \\ \underline{1 - 0.7q^{-1} - 0.14q^{-2} + 0.12q^{-3}} \\ 1.9q^{-1} - 0.25q^{-2} + 0.288q^{-3} \\ \underline{1.9q^{-1} - 2.28q^{-2} - 0.209q^{-3} + 0.3192q^{-4}} \\ 2.03q^{-2} + 0.497q^{-3} - 0.3192q^{-4} \end{array}$$

得到 $f_1 = -1.9, g_0 = 2.03, g_1 = 0.497, g_2 = -0.3192$, 最后结果为

$$F(q^{-1}) = 1 - 1.9q^{-1} \tag{4.3.46}$$

$$q^{-p}G(q^{-1}) = q^{-2}(2.03 + 0.497q^{-1} - 0.3192q^{-2}) \tag{4.3.47}$$

2. 比较系数法

把式 (4.3.31)

$$C(q^{-1}) = A(q^{-1})F(q^{-1}) + q^{-p}G(q^{-1}) \tag{4.3.48}$$

展开, 得

$$\begin{aligned} 1 + c_1q^{-1} + \dots + c_nq^{-n} &= (1 + a_1q^{-1} + \dots + a_nq^{-n})(1 + f_1q^{-1} + \dots + f_{p-1}q^{-(p-1)}) \\ &\quad + q^{-p}(g_0 + g_1q^{-1} + \dots + g_{n-1}q^{-(n-1)}) \end{aligned} \tag{4.3.49}$$

比较式 (4.3.49) 两边系数, 使两边 q^{-1} 的各次幂的系数相等, 可列出 $n+p-1$ 个方程式:

$$\begin{aligned} q^{-1} \quad c_1 &= a_1 + f_1 \\ q^{-2} \quad c_2 &= a_2 + a_1f_1 + f_2 \\ &\vdots \\ q^{-(p-1)} \quad c_{p-1} &= a_{p-1} + a_{p-2}f_1 + a_{p-3}f_2 + \dots + a_1f_{p-2} + f_{p-1} \\ q^{-p} \quad c_p &= a_p + a_{p-1}f_1 + a_{p-2}f_2 + \dots + a_1f_{p-1} + g_0 \\ q^{-(p+1)} \quad c_{p+1} &= a_{p+1} + a_pf_1 + a_{p-1}f_2 + \dots + a_2f_{p-1} + g_1 \\ &\vdots \\ q^{-n} \quad c_n &= a_n + a_{n-1}f_1 + a_{n-2}f_2 + \dots + a_{n-p+1}f_{p-1} + g_{n-p} \\ q^{-(n+1)} \quad 0 &= a_nf_1 + a_{n-1}f_2 + \dots + a_{n-p+2}f_{p-1} + g_{n-p+1} \\ &\vdots \\ q^{-(n+p-1)} \quad 0 &= a_nf_{p-1} + g_{n-1} \end{aligned} \tag{4.3.50}$$

可解出 $f_1, f_2, \dots, f_{p-1}, g_0, g_1, \dots, g_{n-1}$ 共 $n+p-1$ 个系数。

求解方程式(4.3.50)比较简单,不用联立求解,逐次求解就够了。如通过式(4.3.50)的第一个方程可求出系数 f_1 , 通过第二个方程可求出系数 f_2 , 等等。

例 4.3.3 利用比较系数法求解例 4.3.2 结果。

解 利用给定条件,列出比较系数方程

$$\begin{aligned} -0.7 &= 1.2 + f_1 \\ -0.14 &= 0.11 + 1.2f_1 + g_0 \\ 0.12 &= 0.168 + 0.11f_1 + g_1 \\ 0 &= -0.168f_1 + g_2 \end{aligned} \tag{4.3.51}$$

由式(4.3.51)容易解出各系数,分别为 $f_1 = -1.9$, $g_0 = 2.03$, $g_1 = 0.497$, $g_2 = -0.3192$ 。结果与长除法相同。

3. $A(q^{-1})$ 和 $C(q^{-1})$ 阶数不等的情况

设 $A(q^{-1})$ 为 n 阶多项式, $C(q^{-1})$ 为 m 阶多项式, 进行 p 步预测, $F(q^{-1})$ 仍为 $p-1$ 阶多项式, 则 $G(q^{-1})$ 的阶数按如下选择:

- (1) 当 $n + p - 1 < m$ 时, $G(q^{-1})$ 为 $m - p$ 阶;
- (2) 当 $n + p - 1 \geq m$ 时, $G(q^{-1})$ 为 $n - 1$ 阶。

例 4.3.4 考虑表示为

$$\begin{aligned} y(k) &= 2.6y(k-1) + 2.85y(k-2) - 1.4y(k-3) + 0.25y(k-4) \\ &= e(k) - 0.7e(k-1) \end{aligned} \tag{4.3.52}$$

的随机过程 $y(k)$, $e(k)$ 为离散时间高斯 $N(0, 1)$ 白噪声。试确定 $y(k)$ 的一步最优预测 $\hat{y}(k+1|k)$, 使一步预测误差的方差为最小。

解 把式(4.3.52)改写为

$$A(q^{-1})y(k) = C(q^{-1})e(k) \tag{4.3.53}$$

式中:

$$A(q^{-1}) = 1 - 2.6q^{-1} + 2.85q^{-2} - 1.4q^{-3} + 0.25q^{-4} \tag{4.3.54}$$

$$C(q^{-1}) = 1 - 0.7q^{-1} \tag{4.3.55}$$

$$n = 4 \qquad m = 1 \qquad p = 1 \tag{4.3.56}$$

$$n + p - 1 > m \tag{4.3.57}$$

因此

$$F(q^{-1}) = 1 \tag{4.3.58}$$

$$G(q^{-1}) = g_0 + g_1q^{-1} + g_2q^{-2} + g_3q^{-3} \tag{4.3.59}$$

把式(4.3.54), (4.3.55), (4.3.58)和(4.3.59)代入恒等式(4.3.48), 得到

$$\begin{aligned} 1 - 0.7q^{-1} &= 1 - 2.6q^{-1} + 2.85q^{-2} - 1.4q^{-3} + 0.25q^{-4} \\ &\quad + q^{-1}(g_0 + g_1q^{-1} + g_2q^{-2} + g_3q^{-3}) \end{aligned} \tag{4.3.60}$$

利用比较系数法列出方程

$$\begin{aligned}
-0.7 &= -2.6 + g_0 \\
0 &= 2.85 + g_1 \\
0 &= -1.4 + g_2 \\
0 &= 0.25 + g_3
\end{aligned} \tag{4.3.61}$$

容易求出系数 $g_0 = 1.9, g_1 = -2.85, g_2 = 1.4, g_3 = -0.25$, 多项式 $G(q^{-1})$ 为

$$G(q^{-1}) = 1.9 - 2.85q^{-1} + 1.4q^{-2} - 0.25q^{-3} \tag{4.3.62}$$

把式(4.3.55), (4.3.58)和(4.3.62)代入式(4.3.38), (4.3.39)和(3.3.40), 得到

$$\begin{aligned}
\hat{y}(k+1|k) &= \frac{1.9 - 2.85q^{-1} + 1.4q^{-2} - 0.25q^{-3}}{1 - 0.7q^{-1}} y(k) \\
&= 0.7 \hat{y}(k|k-1) + 1.9y(k) - 2.85y(k-1) \\
&\quad + 1.4y(k-2) - 0.25y(k-3)
\end{aligned} \tag{4.3.63}$$

$$J = 1 \tag{4.3.64}$$

$$\tilde{y}(k+1|k) = e(k+1) \tag{4.3.65}$$

4.4 最小方差控制策略

考虑 CARMA 系统

$$A(q^{-1})y(k) = B(q^{-1})u(k-d) + C(q^{-1})e(k) \tag{4.4.1}$$

式中:

$$A(q^{-1}) = 1 + a_1q^{-1} + a_2q^{-2} + \dots + a_nq^{-n} \tag{4.4.2}$$

$$B(q^{-1}) = b_0 + b_1q^{-1} + b_2q^{-2} + \dots + b_nq^{-n} \tag{4.4.3}$$

$$C(q^{-1}) = 1 + c_1q^{-1} + c_2q^{-2} + \dots + c_nq^{-n} \tag{4.4.4}$$

$y(k)$ 为系统输出, 它是对设定值的偏差, 既是观测值又是被控量; $u(k)$ 为系统输入, 它是控制量; d 为系统延迟; $e(k)$ 为系统噪声, 假设它是离散时间高斯 $N(0, 1)$ 白噪声; $A(q^{-1})$, $B(q^{-1})$, $C(q^{-1})$ 都为 n 阶多项式, 所有零点都在单位圆内, 且 $A(q^{-1})$ 和 $C(q^{-1})$ 为首 1 多项式, $B(q^{-1})$ 不为首 1 多项式, 各多项式系数都是已知的, 若某多项式的阶数小于 n , 则可令其有关系数为零即可; 为常系数。

假设容许控制 $u(k)$ 是如下函数:

$$u(k) = f[y(k), y(k-1), \dots, y(k_0); u(k-1), u(k-2), \dots, u(k_0)] \tag{4.4.5}$$

性能指标 J 是系统输出 $y(k)$ 的方差:

$$J = E\{y^2(k)\} \tag{4.4.6}$$

问题的提法是: 对 CARMA 系统式(4.4.1), 要求寻找一种使系统输出方差式(4.4.6)为最小的最优控制策略。这种最优控制策略称为最小方差控制策略, 这种控制称为最小方差控制。

按照解决 p 步预测的类似方法求解最小方差控制策略。把式(4.4.1)改写为

$$e(k) = \frac{A(q^{-1})}{C(q^{-1})}y(k) - \frac{B(q^{-1})}{C(q^{-1})}q^{-d}u(k) \tag{4.4.7}$$

由于从控制作用开始到系统输出延迟 d 步, 在时刻 k 加入的系统输入 $u(k)$, 延迟到时刻

$k+d$ 时才在系统输出 $y(k+d)$ 开始显露出来, 因此有

$$y(k+d) = \frac{B(q^{-1})}{A(q^{-1})}u(k) + \frac{C(q^{-1})}{A(q^{-1})}e(k+d) \quad (4.4.8)$$

像求解预测问题一样, 引入恒等式

$$C(q^{-1}) = A(q^{-1})F(q^{-1}) + q^{-d}G(q^{-1}) \quad (4.4.9)$$

以便把式(4.4.8)右边分为独立的两部分, 为此把 $F(q^{-1})$ 选为 $d-1$ 阶多项式, 而 $G(q^{-1})$ 必为 $n-1$ 阶多项式, 即

$$F(q^{-1}) = 1 + f_1q^{-1} + f_2q^{-2} + \dots + f_{d-1}q^{-(d-1)} \quad (4.4.10)$$

$$G(q^{-1}) = g_0 + g_1q^{-1} + \dots + g_{n-1}q^{-(n-1)} \quad (4.4.11)$$

把恒等式(4.4.9)代入式(4.4.8), 得到

$$y(k+d) = F(q^{-1})e(k+d) + \frac{B(q^{-1})}{A(q^{-1})}u(k) + \frac{G(q^{-1})}{A(q^{-1})}e(k) \quad (4.4.12)$$

把式(4.4.7)代入上式, 得到

$$\begin{aligned} y(k+d) = & F(q^{-1})e(k+d) + \frac{G(q^{-1})}{C(q^{-1})}y(k) + \frac{B(q^{-1})}{A(q^{-1})}u(k) \\ & - \frac{G(q^{-1})}{A(q^{-1})} \frac{B(q^{-1})}{C(q^{-1})} q^{-d}u(k) \end{aligned} \quad (4.4.13)$$

由于可把恒等式(4.4.9)改写为

$$\frac{q^{-d}G(q^{-1})}{A(q^{-1})} = \frac{C(q^{-1})}{A(q^{-1})} - F(q^{-1}) \quad (4.4.14)$$

把上式代入式(4.4.13)右边第四项, 容易得到

$$y(k+d) = F(q^{-1})e(k+d) + \frac{G(q^{-1})}{C(q^{-1})}y(k) + \frac{B(q^{-1})F(q^{-1})}{C(q^{-1})}u(k) \quad (4.4.15)$$

经上述推导, 式(4.4.15)右边就把 $y(k+d)$ 分成了独立的两部分, 第一项与第二、第三项是独立的。并且, 在时刻 k , $y(k)$ 已被观测, 是已知的, $u(k)$ 是待确定的, 但也是确定性函数。把式(4.4.15)代入性能指标表示式, 并加计算, 得到

$$\begin{aligned} J = & E\{y^2(k+d)\} \\ = & E\{[F(q^{-1})e(k+d)]^2\} + \left[\frac{G(q^{-1})}{C(q^{-1})}y(k) + \frac{B(q^{-1})F(q^{-1})}{C(q^{-1})}u(k) \right]^2 \end{aligned} \quad (4.4.16)$$

上式右边两项都是正的, 只有第二项为零时, 性能指标取最小值。于是得到

$$u(k) = - \frac{G(q^{-1})}{B(q^{-1})F(q^{-1})}y(k) \quad (4.4.17)$$

$$\begin{aligned} \min J = & E\{[F(q^{-1})e(k+d)]^2\} \\ = & \sigma^2[1 + f_1^2 + f_2^2 + \dots + f_{d-1}^2] \end{aligned} \quad (4.4.18)$$

把式(4.4.17)代入式(4.4.15), 得到

$$\begin{aligned} y(k+d) = & F(q^{-1})e(k+d) \\ = & [e(k+d) + f_1e(k+d-1) + \dots + f_{d-1}e(k+1)] \end{aligned} \quad (4.4.19)$$

总之, 对式(4.4.1)实行最小方差控制时, 最小方差控制策略为式(4.4.17)所示; 控制偏差由一个噪声序列组成, 如式(4.4.19)所示; 最小性能指标为偏差噪声序列的方差, 如式(4.4.18)所示。

例 4.4.1 已知 CARMA 系统式(4.4.1), 其中

$$A(q^{-1}) = 1 - 1.7q^{-1} + 0.7q^{-2} \quad (4.4.20)$$

$$B(q^{-1}) = 1 + 0.5q^{-1} \quad (4.4.21)$$

$$C(q^{-1}) = 1 + 1.5q^{-1} + 0.9q^{-2} \quad (4.4.22)$$

$$= 1 \quad (4.4.23)$$

试求 $d=1$ 和 $d=2$ 时的最小方差控制策略。

解 先求 $d=1$ 的情况。利用恒等式(4.4.9)计算出

$$F(q^{-1}) = 1 \quad (4.4.24)$$

$$G(q^{-1}) = 3.2 + 0.2q^{-1} \quad (4.4.25)$$

由式(4.4.17), (4.4.18)和(4.4.19)分别求出

$$\begin{aligned} u(k) &= - \frac{3.2 + 0.2q^{-1}}{1 + 0.5q^{-1}} y(k) \\ &= - 0.5u(k-1) - 3.2y(k) - 0.2y(k-1) \end{aligned} \quad (4.4.26)$$

$$\min J = 1 \quad (4.4.27)$$

$$y(k+1) = e(k+1) \quad (4.4.28)$$

再求 $d=2$ 的情况。同样求出

$$F(q^{-1}) = 1 + 3.2q^{-1} \quad (4.4.29)$$

$$G(q^{-1}) = 5.64 - 2.24q^{-1} \quad (4.4.30)$$

$$\begin{aligned} u(k) &= - \frac{5.64 - 2.24q^{-1}}{1 + 3.7q^{-1} + 1.6q^{-2}} y(k) \\ &= - 3.7u(k-1) - 1.6u(k-2) - 5.64y(k) + 2.24y(k-1) \end{aligned} \quad (4.4.31)$$

$$\min J = 1 + 3.2^2 = 11.24 \quad (4.4.32)$$

$$y(k+2) = e(k+2) + 3.2e(k+1) \quad (4.4.33)$$

由这个例子可知, 时间延迟越大, 性能指标会越大, 控制误差也大大增加。

例 4.4.2 考虑由

$$y(k) = \frac{1}{1 + 0.7q^{-1}} u(k-1) + \frac{1 + 2q^{-1}}{1 - 0.2q^{-1}} e(k) \quad (4.4.34)$$

描述的系统, 其中 $e(k)$ 为独立同分布高斯 $N(0, 1)$ 随机变量序列。试确定系统的最小方差控制策略。

解 把式(4.4.34)改写为 CARMA 模型的标准形式

$$A(q^{-1})y(k) = B(q^{-1})u(k-1) + C(q^{-1})e(k) \quad (4.4.35)$$

式中

$$A(q^{-1}) = (1 - 0.2q^{-1})(1 + 0.7q^{-1}) = 1 + 0.5q^{-1} - 0.14q^{-2} \quad (4.4.36)$$

$$B(q^{-1}) = 1 - 0.2q^{-1} \quad (4.4.37)$$

$$C(q^{-1}) = (1 + 2q^{-1})(1 + 0.7q^{-1}) \quad (4.4.38)$$

$$= 1 \quad n = 2 \quad d = 1 \quad (4.4.39)$$

由式(4.4.38)可知, $C(q^{-1})$ 在单位圆外有零点, 使系统不稳定, 作为噪声项这是不可能的, 需进行换算。为此, 写出谱密度, 并加以变换

$$\begin{aligned} &= (1 + 2e^{-1})(1 + 2e^{-1}) \\ &= 2(1 + 0.5e^{-1}) \times 2(1 + 0.5e^{-1}) \end{aligned} \quad (4.4.40)$$

由上式可知, $2(1 + 0.5e^{-1})$ 与 $(1 + 2e^{-1})$ 有相同的谱密度, 因此, 可用前者代替后者, 于是式 (4.4.38) 变为

$$\begin{aligned} {}_1C_1(q^{-1}) &= 2(1 + 0.5q^{-1})(1 + 0.7q^{-1}) \\ &= 2(1 + 1.2q^{-1} + 0.35q^{-2}) \end{aligned}$$

系统模型变为

$$A(q^{-1})y(k) = B(q^{-1})u(k-1) + {}_1C_1(q^{-1})e(k) \quad (4.4.41)$$

式中: $n=2$, $p=1$, $r=2$ 。

采用比较系数法, 得到

$$F(q^{-1}) = 1 \quad (4.4.42)$$

$$G(q^{-1}) = 0.7 + 0.49q^{-1} \quad (4.4.43)$$

再由式 (4.4.17), (4.4.18) 和 (4.4.19) 得到结果

$$\begin{aligned} u(k) &= - \frac{0.7 + 0.4q^{-1}}{1 - 0.2q^{-1}} y(k) \\ &= 0.2u(k-1) - 0.7y(k) - 0.49y(k-1) \end{aligned} \quad (4.4.44)$$

$$\min J = 4 \quad (4.4.45)$$

$$y(k+1) = 2e(k+1) \quad (4.4.46)$$

下面给出几点说明。

(1) 分离定理

可把式 (4.4.15) 看做是预测模型, 右边后两项是 d 步预测, 而第一项是预测误差。选择控制策略, 使控制误差等于预测误差, 这种策略就是最小方差控制策略。因此, 可把控制问题分为两部分, 一部分是预测, 另一部分是求控制策略, 这称为分离定理。

(2) 放松 $e(k)$ 的高斯假设

对上述的线性问题, $e(k)$ 的高斯假设可以放松。

(3) 多输入多输出情况

当式 (4.4.1) 中, $y(k)$, $u(k)$ 和 $e(k)$ 为 $n \times 1$ 向量, $A(q^{-1})$, $B(q^{-1})$ 和 $C(q^{-1})$ 为矩阵多项式时, 最小方差控制策略有相应结果。矩阵多项式恒等式为

$$A^{-1}(q^{-1})C(q^{-1}) = F(q^{-1}) + q^{-d}A^{-1}(q^{-1})G(q^{-1}) \quad (4.4.47)$$

式中 $F(q^{-1})$ 和 $G(q^{-1})$ 仍分别由式 (4.4.10) 和 (4.4.11) 表示, 但为矩阵多项式。最小方差控制策略、控制误差和最小性能指标分别为

$$\begin{aligned} u(k) &= -B^{-1}(q^{-1})G(q^{-1})F^{-1}(q^{-1})y(k) \\ y(k+d) &= F(q^{-1})e(k+d) \\ \min J &= \min E\{y^T(k+d)y(k+d)\} \\ &= E\{[F(q^{-1})e(k+d)]^T [q^d]\} \\ &= \text{tr}^2[I + f_1^T f_1 + f_2^T f_2 + \dots + f_{d-1}^T f_{d-1}] \end{aligned}$$

4.5 次最优控制策略

本节针对多项式 $B(q^{-1})$ 讨论系统控制问题。最小方差控制要求多项式 $B(q^{-1})$ 的零点都在单位圆内,即要求 $B(q^{-1})$ 是最小相位的。如果 $B(q^{-1})$ 在单位圆外有零点,即 $B(q^{-1})$ 是非最小相位的,会造成系统不稳定,不能实现最小方差控制策略。这时先要使系统稳定,再采用次最优控制策略。本节先进行稳定性分析,再介绍两种次最优控制方法。

4.5.1 稳定性分析

考虑随机控制系统由已熟悉的 CARMA 模型

$$A(q^{-1})y(k) = B(q^{-1})u(k-d) + C(q^{-1})e(k) \quad (4.5.1)$$

描述,多项式 $A(q^{-1})$, $B(q^{-1})$, $C(q^{-1})$ 和 θ 是真实参量。但它们的值可能难于精确确定,假设提供设计最小方差控制策略用的 CARMA 模型为

$$A_s(q^{-1})y(k) = B_s(q^{-1})u(k-d) + {}_sC_s(q^{-1})e(k) \quad (4.5.2)$$

式中: $A_s(q^{-1})$, $B_s(q^{-1})$, $C_s(q^{-1})$ 和 ${}_s\theta$ 是设计用参量,通过理论计算或辨识等方法得到。应使设计参量值尽可能逼近真实参量值,但一般不能做到完全相同。根据式(4.5.2)设计参量设计的最小方差控制策略,已在 4.4 节讨论,得到下列相关公式:

$$C_s(q^{-1}) = A_s(q^{-1})F_s(q^{-1}) + q^{-d}G_s(q^{-1}) \quad (4.5.3)$$

$$u(k) = - \frac{G_s(q^{-1})}{B_s(q^{-1})F_s(q^{-1})}y(k) \quad (4.5.4)$$

$$y(k) = {}_sF_s(q^{-1})e(k+d) \quad (4.5.5)$$

实际情况是用设计的最小方差控制策略去控制真实系统,其系统传递函数如图 4.5.1 所示,图中引入多项式 $E(k) = \frac{C(q^{-1})}{A(q^{-1})}e(k)$ 。

图 4.5.1 最小方差控制系统传递函数图

为进行稳定性分析,把 $e(k)$ 作为系统输入,把 $y(k)$ 和 $u(k)$ 作为系统输出,写出系统闭环传递函数,即有

$$y(k) = \frac{1}{1 + \frac{G_s(q^{-1})}{B_s(q^{-1})F_s(q^{-1})} i^{\alpha} q^{-d} \frac{B(q^{-1})}{A(q^{-1})}} E(k)$$

$$= \frac{B_s(q^{-1})C(q^{-1})F_s(q^{-1})}{A(q^{-1})B_s(q^{-1})F_s(q^{-1}) - A_s(q^{-1})B(q^{-1})F_s(q^{-1}) + B(q^{-1})C_s(q^{-1})} e(k) \quad (4.5.6)$$

$$u(k) = \frac{-C(q^{-1})G_s(q^{-1})}{A(q^{-1})B_s(q^{-1})F_s(q^{-1}) - A_s(q^{-1})B(q^{-1})F_s(q^{-1}) + B(q^{-1})C_s(q^{-1})} e(k) \quad (4.5.7)$$

为使系统稳定, 要求特征方程

$$A(q^{-1})B_s(q^{-1})F_s(q^{-1}) - A_s(q^{-1})B(q^{-1})F_s(q^{-1}) + B(q^{-1})C_s(q^{-1}) = 0 \quad (4.5.8)$$

的根都在单位圆内。为简化分析, 令

$$A_s(q^{-1}) = A(q^{-1}) \quad B_s(q^{-1}) = B(q^{-1}) \quad C_s(q^{-1}) = C(q^{-1}) \quad (4.5.9)$$

把上式代入式(4.5.8), 特征方程变为

$$B(q^{-1})C(q^{-1}) = 0 \quad (4.5.10)$$

由上式看出, 为使系统稳定, 要求 $B(q^{-1})$ 和 $C(q^{-1})$ 的零点, 即系统极点都在单位圆内。

首先, 分析 $C(q^{-1})$ 的零点对系统稳定性的影响。为使系统稳定, 由式(4.5.10)可知, 要求 $C(q^{-1})$ 的零点都在单位圆内。由 4.2 节的式(4.2.14)

$$y(k) = e(k) + [y(k_0) - e(k_0)](-c)^{k-k_0} \quad (4.5.11)$$

看到, 最小方差控制系统的动力学取决于多项式 $C(q^{-1})$, 如果 $C(q^{-1})$ 在单位圆外有零点, 就会使系统不稳定。幸好, $C(q^{-1})$ 是表示噪声的多项式, 其零点必都在单位圆内。否则, 噪声多项式是发散的, 不能收敛, 这在物理意义上表明噪声随着时间的增加而越来越大, 能量越来越强, 是不可能的。

其次, 分析 $B(q^{-1})$ 的零点对系统稳定性的影响。把式(4.5.9)代入式(4.5.6)和(4.5.7), 得到

$$y(k) = F_s(q^{-1}) e(k) \quad (4.5.12)$$

$$u(k) = -\frac{G_s(q^{-1})}{B(q^{-1})} e(k) \quad (4.5.13)$$

如果 $B(q^{-1})$ 的零点都在单位圆内, 由式(4.5.13)可知, $u(k)$ 可为有限值, 有可能实现最小方差控制策略; 如果 $B(q^{-1})$ 在单位圆外有零点, 式(4.5.13)将是不稳定的, 从理论上说, $u(k)$ 可能达到无穷大值, 这是任何实际设备不能允许的, 因此不能实现最小方差控制策略。甚至, 即使 $B(q^{-1})$ 的零点都在单位圆内, 但有的零点靠近单位圆时, 也会使控制系统对它的参数变化极为敏感, 造成整个控制系统不能稳定运行。由式(4.5.12)可知, 不管 $B(q^{-1})$ 在单位圆外有无零点, 系统输出 $y(k)$ 都能达到最小方差控制策略效果, 只要系统输入 $u(k)$ 为任何值时都能正常工作。

4.5.2 次最优控制策略

当多项式 $B(q^{-1})$ 为非最小相位时, 系统不稳定, 不能实现最小方差控制策略。下面介绍两种方法, 它们能使系统正常工作, 但为次最优控制策略。

1. 极点消去法

设 n 阶多项式 $B(q^{-1})$ 是非最小相位的, 可把它分解为如下两部分:

$$B(q^{-1}) = B_1(q^{-1})B_2(q^{-1}) \quad (4.5.14)$$

$$n = n_1 + n_2 \quad (4.5.15)$$

式中: $B_1(q^{-1})$ 是 n_1 阶多项式, 其零点都在单位圆内; $B_2(q^{-1})$ 是 n_2 阶多项式, 其零点都在单位圆外或单位圆上。解决问题的关键是构造如下形式的恒等式:

$$C(q^{-1}) = A(q^{-1})F_1(q^{-1}) + q^{-d}B_2(q^{-1})G_1(q^{-1}) \quad (4.5.16)$$

$$F_1(q^{-1}) = 1 + f_{11}q^{-1} + f_{12}q^{-2} + \dots + f_{1(n_2+d-1)}q^{-(n_2+d-1)} \quad (4.5.17)$$

$$G_1(q^{-1}) = g_{10} + g_{11}q^{-1} + g_{12}q^{-2} + \dots + g_{1(n-1)}q^{-(n-1)} \quad (4.5.18)$$

$B_2(q^{-1})$ 的零点都在单位圆外, 使系统不稳定, 必须在系统闭环极点中将 $B_2(q^{-1})$ 消去, 对照最小方差控制策略式(4.4.17)

$$u(k) = - \frac{G(q^{-1})}{B(q^{-1})F(q^{-1})} y(k) \quad (4.5.19)$$

如果把 $B_2(q^{-1})$ 和 $G(q^{-1})$ 放在一起, 就可消去 $B_2(q^{-1})$, 所以我们构造了形式为式(4.5.16)的恒等式。有关多项式阶次可做如下解释: 采用最小方差控制策略时, $F(q^{-1})$ 是 $d-1$ 阶的, 则 $q^{-d}G(q^{-1})$ 是 $d+n-1$ 阶的; 次最优控制策略时, $F_1(q^{-1})$ 是 $d+n_2-1$ 阶的, 则 $q^{-d}B_2(q^{-1})G_1(q^{-1})$ 是 $d+n_2+n-1$ 阶的, 也就是 $G_1(q^{-1})$ 是 $n-1$ 阶的。多项式 $F_1(q^{-1})$ 和 $G_1(q^{-1})$ 由比较系数法确定。

按照确定最小方差控制策略式(4.5.19)的同样方法, 可得有关公式:

$$\begin{aligned} u(k) &= - \frac{B_2(q^{-1})G_1(q^{-1})}{B(q^{-1})F_1(q^{-1})} y(k) \\ &= - \frac{G_1(q^{-1})}{B_1(q^{-1})F_1(q^{-1})} y(k) \end{aligned} \quad (4.5.20)$$

$$\begin{aligned} y(k+d) &= F_1(q^{-1})e(k+d) \\ &= [e(k+d) + f_{11}e(k+d-1) + f_{12}e(k+d-2) + \dots + \\ &\quad f_{1(d-1)}e(k+1) + f_{1d}e(k) + \dots + f_{1(n_2+d-1)}e(k-n_2+1)] \end{aligned} \quad (4.5.21)$$

$$\min J = \frac{1}{2}(1 + f_{11}^2 + f_{12}^2 + \dots + f_{1(d-1)}^2 + f_{1d}^2 + \dots + f_{1(n_2+d-1)}^2) \quad (4.5.22)$$

由式(4.5.20)可知, $B_2(q^{-1})$ 已从闭环分母中消去, 分母中不再包含单位圆以外的极点, 系统变成了稳定系统。由式(4.5.22)可知, 系统输出方差大于最小方差控制所达到的方差, 因此, 这种控制策略称为次最优控制策略, 这种控制称为次最优控制。

下面进行稳定性分析。

同样, 用下标 s 表示设计用参量, 进行与 4.5.1 小节同样的分析。令

$$B_{2s}(q^{-1}) = B_2(q^{-1}) \quad (4.5.23)$$

那么闭环传递函数为

$$y(k) = \frac{B_{1s}(q^{-1})C(q^{-1})F_{1s}(q^{-1})}{A(q^{-1})B_{1s}(q^{-1})F_{1s}(q^{-1}) - A_s(q^{-1})B_1(q^{-1})F_{1s}(q^{-1}) + B_1(q^{-1})C_s(q^{-1})} e(k) \quad (4.5.24)$$

$$u(k) = \frac{C(q^{-1})G_{1s}(q^{-1})}{A(q^{-1})B_{1s}(q^{-1})F_{1s}(q^{-1}) - A_s(q^{-1})B_1(q^{-1})F_{1s}(q^{-1}) + B_1(q^{-1})C_s(q^{-1})} e(k) \quad (4.5.25)$$

系统的特征方程为

$$A(q^{-1})B_{1s}(q^{-1})F_{1s}(q^{-1}) - A_s(q^{-1})B_1(q^{-1})F_{1s}(q^{-1}) + B_1(q^{-1})C_s(q^{-1}) = 0 \quad (4.5.26)$$

为简化分析, 令

$$A_s(q^{-1}) = A(q^{-1}) \quad B_{1s}(q^{-1}) = B_1(q^{-1}) \quad C_s(q^{-1}) = C(q^{-1}) \quad (4.5.27)$$

则有

$$B_1(q^{-1})C_s(q^{-1}) = 0 \quad (4.5.28)$$

把上式与式(4.5.10)对照, 在上式中已把 $B_2(q^{-1})$ 消去, 使特征方程的根都在单位圆内, 因此系统是稳定的。

例 4.5.1 考虑系统

$$\begin{aligned} & y(k) + 0.64y(k-1) + 0.22y(k-2) \\ & = 6.4u(k-3) + 19.2u(k-4) + [e(k) - 0.82e(k-1) + 0.21e(k-2)] \end{aligned} \quad (4.5.29)$$

试确定次最优控制策略。

解 把式(4.5.29)写为标准形式

$$A(q^{-1})y(k) = B(q^{-1})u(k-d) + C(q^{-1})e(k) \quad (4.5.30)$$

式中:

$$A(q^{-1}) = 1 + 0.64q^{-1} + 0.22q^{-2} \quad (4.5.31)$$

$$B(q^{-1}) = 6.4(1 + 3q^{-1}) \quad (4.5.32)$$

$$C(q^{-1}) = 1 - 0.82q^{-1} + 0.21q^{-2} \quad (4.5.33)$$

$$n = 2 \quad d = 3 \quad (4.5.34)$$

多项式 $B(q^{-1})$ 在单位圆外有零点, 使系统工作不稳定, 采用极点消去法, 只能求出次最优控制策略。把 $B(q^{-1})$ 分为两部分

$$B(q^{-1}) = B_1(q^{-1})B_2(q^{-1}) \quad (4.5.35)$$

式中:

$$B_1(q^{-1}) = 6.4 \quad (n_1 = 0) \quad (4.5.36)$$

$$B_2(q^{-1}) = 1 + 3q^{-1} \quad (n_2 = 1) \quad (4.5.37)$$

列出恒等式

$$C(q^{-1}) = A(q^{-1})F_1(q^{-1}) + q^{-d}B_2(q^{-1})G_1(q^{-1}) \quad (4.5.38)$$

$F_1(q^{-1})$ 是 $n_2 + d - 1 = 3$ 阶多项式, $G_1(q^{-1})$ 是 $n - 1 = 1$ 阶多项式。通过比较系数法, 得到

$$F_1(q^{-1}) = 1 - 1.46q^{-1} + 0.92q^{-2} - 0.25q^{-3} \quad (4.5.39)$$

$$G_1(q^{-1}) = 0.02 + 0.02q^{-1} \quad (4.5.40)$$

把有关数据代入式(4.5.20), (4.5.21)和(4.5.22), 得到次最优控制策略、控制误差和性能指标, 分别为

$$\begin{aligned} u(k) &= - \frac{-0.02 + 0.02q^{-1}}{6.4(1 - 1.46q^{-1} + 0.92q^{-2} - 0.25q^{-3})} y(k) \\ &= 1.46u(k-1) - 0.92u(k-2) + 0.25u(k-3) \end{aligned}$$

$$+ 0.003y(k) - 0.003y(k-1) \quad (4.5.41)$$

$$y(k) = [e(k) - 1.46e(k-1) + 0.92e(k-2) - 0.25e(k-3)] \quad (4.5.42)$$

$$\begin{aligned} \min J &= \frac{1}{2} [1 + 1.46^2 + 0.92^2 + 0.25^2] \\ &= 4.04 \quad (4.5.43) \end{aligned}$$

为了对比,把最小方差控制策略、控制误差和性能指标列在下面:

$$\begin{aligned} u_0(k) &= \frac{0.27 + 0.2q^{-1}}{6.4(1 + 3q^{-1})(1 - 1.46q^{-1} + 0.92q^{-2})} y(k) \\ &= -1.54u(k-1) + 3.45u(k-2) - 2.77u(k-3) \\ &\quad + 0.04y(k) + 0.03y(k-1) \end{aligned} \quad (4.5.44)$$

$$y_0(k) = [e(k) - 1.46e(k-1) + 0.92e(k-2)] \quad (4.5.45)$$

$$\begin{aligned} \min J_0 &= \frac{1}{2} [1 + 1.46^2 + 0.92^2] \\ &= 3.98 \quad (4.5.46) \end{aligned}$$

比较式(4.5.46)和(4.5.43)可知,次最优控制策略的性能指标比最小方差控制策略的性能指标大。

2. 广义最小方差控制

当多项式 $B(q^{-1})$ 是非最小相位时,最小方差控制策略的困难在于可能出现过大的控制信号 $u(k)$,使实际设备饱和或不允许,而影响调节品质或不能正常工作。针对这个问题,考虑在性能指标函数中加入加权控制信号来解决。为此,把性能指标函数改写为

$$J = E\{y^2(k+d) + Qu^2(k)\} \quad (4.5.47)$$

式中: Q 为加权系数。

现在的问题是,对 CARMA 模型式(4.4.1),寻求一种控制策略 $u(k)$,使性能指标式(4.5.47)为最小。

我们对性能指标进行计算。把式(4.4.15)代入式(4.5.47),并加以推导,得到

$$\begin{aligned} J &= E \left\{ F(q^{-1})e(k+d) + \frac{G(q^{-1})}{C(q^{-1})}y(k) + \frac{B(q^{-1})F(q^{-1})}{C(q^{-1})}u(k) \right\}^2 + Qu^2(k) \\ &= E \{ [F(q^{-1})e(k+d)]^2 \} + \frac{G(q^{-1})}{C(q^{-1})}y(k) + \frac{B(q^{-1})F(q^{-1})}{C(q^{-1})}u(k) \quad (4.5.48) \end{aligned}$$

式中: $y(k)$ 为观测值,是已知量; $u(k)$ 为待确定的确定性函数。因此,上式第二等式右边第二项和第三项可不加均值符号。

为了求 J 的最小值,使式(4.5.48)对 $u(k)$ 的偏导数为零,得到

$$\frac{\partial J}{\partial u(k)} = 2 \frac{G(q^{-1})}{C(q^{-1})}y(k) + \frac{B(q^{-1})F(q^{-1})}{C(q^{-1})}u(k) - b_0 + 2Qu(k) = 0 \quad (4.5.49)$$

式中 b_0 是这样求得的: 因为

$$\frac{B(q^{-1})F(q^{-1})}{C(q^{-1})}u(k) = b_0u(k) + f[u(k-1), u(k-2), \dots] \quad (4.5.50)$$

所以,有

$$\begin{aligned} &\frac{G(q^{-1})}{C(q^{-1})}y(k) + \frac{B(q^{-1})F(q^{-1})}{C(q^{-1})}u(k) \\ &= \frac{1}{u(k)} \{ b_0u(k) + f[u(k-1), u(k-2), \dots] \} \end{aligned}$$

$$= b_0 \quad (4.5.51)$$

由式(4.5.49)求得控制策略

$$u(k) = \frac{G(q^{-1})}{B(q^{-1})F(q^{-1}) + \frac{Q}{b_0}C(q^{-1})} y(k) \quad (4.5.52)$$

把上式代入式(4.5.48), 得到性能指标函数 J 的最小值

$$\begin{aligned} \min J = E\{[F(q^{-1})e(k+d)]^2\} \\ + \frac{G^2(q^{-1})}{C^2(q^{-1})} \left[1 - \frac{B(q^{-1})F(q^{-1})}{B(q^{-1})F(q^{-1}) + \frac{Q}{b_0}C(q^{-1})}\right]^2 \\ + \frac{QC^2(q^{-1})}{B(q^{-1})F(q^{-1}) + \frac{Q}{b_0}C(q^{-1})} y^2(k) \end{aligned} \quad (4.5.53)$$

上式右边第一项就是最小方差控制的最小性能指标, 而第二项是正的, 因此, 广义最小方差控制的最小性能指标肯定大于最小方差控制的最小性能指标, 广义最小方差控制是一种次最优控制。当 Q 等于零时, 广义最小方差控制策略就是最小方差控制策略。

现进行稳定性分析。

令设计参量等于真实参量, 推导出闭环传递函数为

$$y(k) = \frac{B(q^{-1})F(q^{-1}) + \frac{Q}{b_0}C(q^{-1})}{B(q^{-1}) + \frac{Q}{b_0}A(q^{-1})} e(k) \quad (4.5.54)$$

$$u(k) = \frac{G(q^{-1})}{B(q^{-1}) + \frac{Q}{b_0}A(q^{-1})} e(k) \quad (4.5.55)$$

特征方程为

$$B(q^{-1}) + \frac{Q}{b_0}A(q^{-1}) = 0 \quad (4.5.56)$$

由上式看出: 当加权系数 Q 为零时, 其特征方程与最小方差控制情况相同, 闭环极点就是多项式 $B(q^{-1})$ 的零点; 当加权系数 Q 为无穷大时, $B(q^{-1})$ 可以忽略, 闭环极点就是 $A(q^{-1})$ 的零点。一般情况, $A(q^{-1})$ 的零点都在单位圆内。总之, 广义最小方差控制系统的闭环极点, 随 Q 的增加, 由 $B(q^{-1})$ 的零点向 $A(q^{-1})$ 的零点逼近。适当选择 Q 值, 可确定较为理想的闭环极点。

4.6 CARMA 模型与状态空间模型的转换

本节介绍把 CARMA 模型转换为状态空间模型的方法。

考虑 CARMA 模型, 并把它写为如下形式

$$y(k) = \frac{B(q^{-1})}{A(q^{-1})} u(k-d) + \frac{C(q^{-1})}{A(q^{-1})} e(k) \quad (4.6.1)$$

式中:

$$A(q^{-1}) = 1 + a_1 q^{-1} + a_2 q^{-2} + \dots + a_n q^{-n} \quad (4.6.2)$$

$$B(q^{-1}) = b_0 + b_1 q^{-1} + b_2 q^{-2} + \dots + b_n q^{-n} \quad (4.6.3)$$

$$C(q^{-1}) = 1 + c_1 q^{-1} + c_2 q^{-2} + \dots + c_n q^{-n} \quad (4.6.4)$$

计算 $\frac{B(q^{-1})}{A(q^{-1})}$ 和 $\frac{C(q^{-1})}{A(q^{-1})}$, 有

$$\frac{B(q^{-1})}{A(q^{-1})} = b_0 + \frac{B_1(q^{-1})}{A(q^{-1})} \quad (4.6.5)$$

$$\frac{C(q^{-1})}{A(q^{-1})} = 1 + \frac{C_1(q^{-1})}{A(q^{-1})} \quad (4.6.6)$$

式中:

$$B_1(q^{-1}) = (b_1 - b_0 a_1) q^{-1} + (b_2 - b_0 a_2) q^{-2} + \dots + (b_n - b_0 a_n) q^{-n} \quad (4.6.7)$$

$$C_1(q^{-1}) = (c_1 - a_1) q^{-1} + (c_2 - a_2) q^{-2} + \dots + (c_n - a_n) q^{-n} \quad (4.6.8)$$

把式(4.6.5)和(4.6.6)代入式(4.6.1), 得到

$$\begin{aligned} y(k) &= b_0 u(k-d) + \frac{B_1(q^{-1})}{A(q^{-1})} u(k-d) + e(k) + \frac{C_1(q^{-1})}{A(q^{-1})} e(k) \\ &= x_1(k) + b_0 u(k-d) + e(k) \end{aligned} \quad (4.6.9)$$

式中:

$$x_1(k) = \frac{B_1(q^{-1})}{A(q^{-1})} u(k-d) + \frac{C_1(q^{-1})}{A(q^{-1})} e(k) \quad (4.6.10)$$

$$A(q^{-1}) x_1(k) = B_1(q^{-1}) u(k-d) + C_1(q^{-1}) e(k) \quad (4.6.11)$$

对上式令

$$\begin{aligned} x_1(k+1) &= -a_1 x_1(k) + (b_1 - b_0 a_1) u(k-d) + (c_1 - a_1) e(k) + x_2(k) \\ x_2(k+1) &= -a_2 x_1(k) + (b_2 - b_0 a_2) u(k-d) + (c_2 - a_2) e(k) + x_3(k) \end{aligned} \quad (4.6.12)$$

$$\begin{aligned} x_n(k+1) &= -a_n x_1(k) + (b_n - b_0 a_n) u(k-d) + (c_n - a_n) e(k) \\ x(k) &= [x_1(k) \ x_2(k) \ \dots \ x_n(k)]^T \end{aligned} \quad (4.6.13)$$

式(4.6.12)和式(4.6.11)是等价的。把式(4.6.12), (4.6.13)和(4.6.9)写成状态空间方程, 就是要求的结果, 即

$$\begin{aligned} x(k+1) &= \begin{bmatrix} -a_1 & 1 & 0 & \dots & 0 \\ -a_2 & 0 & 1 & w & 0 \\ & & w & w & \\ -a_n & 0 & \dots & & 0 \end{bmatrix} x(k) + \begin{bmatrix} b_1 - b_0 a_1 \\ b_2 - b_0 a_2 \\ \\ b_n - b_0 a_n \end{bmatrix} u(k-d) \\ &\quad + \begin{bmatrix} c_1 - a_1 \\ c_2 - a_2 \\ \\ c_n - a_n \end{bmatrix} e(k) \end{aligned} \quad (4.6.14)$$

$$y(k) = [1 \ 0 \ \dots \ 0] x(k) + b_0 u(k-d) + e(k) \quad (4.6.15)$$

习 题

4.1 考虑由

$$\begin{aligned} y(k) &= 1.5y(k-1) + 0.5y(k-2) \\ &= 2[e(k) - 1.2e(k-1) + 0.6e(k-2)] \end{aligned}$$

定义的随机过程 $y(k)$, 式中 $\{e(t)\}$ 是独立同分布高斯 $N(0, 1)$ 随机变量序列。试确定使均方预测误差为最小的 p 步预测。

4.2 考虑由下式定义的随机过程 $y(k)$:

$$y(k) + ay(k-1) = [e(k) + ce(k-1)]$$

式中: $|a| < 1$; $|c| < 1$; $\{e(k)\}$ 是独立同分布高斯 $N(0, 1)$ 随机变量序列。试确定使均方预测误差为最小的 p 步预测。

4.3 试对于由

$$y(k) = \frac{1}{1 + aq^{-1}}e(k) + v(k)$$

定义的随机过程, 求出其使均方误差为最小的两步预测, 并确定预测误差。式中: $\{e(k), k = T\}$ 和 $\{v(k), k = T\}$ 是两个独立同分布高斯 $N(0, 1)$ 随机变量序列, $|a| < 1$ 。

4.4 随机过程 $\{y(k), k = T\}$ 的表示式为

$$y(k) + 0.7y(k-1) = e(k) + 2e(k-1)$$

式中: $\{e(k), k = T\}$ 是独立同分布高斯 $N(0, 1)$ 随机变量序列。试确定最优一步预测和预测误差的方差, 这个最优预测是均方预测误差为最小的预测。

提示: $C(q^{-1})$ 在单位圆外有零点, 系统结构不稳定, 需在 $C(q^{-1})$ 谱密度保持不变条件下进行换算。

4.5 考虑由下式定义的随机过程 $\{y(k), k = k_0, k_0+1, \dots\}$:

$$y(k) = \sum_{l=k_0}^k g(k, l)e(l)$$

式中: $\{e(k), k = k_0, k_0+1, \dots\}$ 是独立同分布高斯 $N(0, 1)$ 随机变量序列。如果 $g(k, k) = 0$, 则 $y(k)$ 总有逆, 设其逆表示为

$$e(k) = \sum_{l=k_0}^k h(k, l)y(l)$$

试确定使均方预测误差为最小的过程 $y(k)$ 的 p 步预测 $\hat{y}(k+p|k)$ 。

4.6 考虑由下式描述的系统:

$$y(k) = \frac{1}{1 + 0.5q^{-1}}u(k-1) + \frac{1 + 0.7q^{-1}}{1 - 0.2q^{-1}}e(k)$$

式中: $e(k)$ 是独立同分布高斯 $N(0, 1)$ 随机变量序列。试确定最小方差控制策略。

4.7 考虑由下式描述的系统:

$$y(k) + ay(k-1) = bu(k-p) + [e(k) + ce(k-1)]$$

式中: $\{e(k), k = \dots, -2, -1, 0, 1, 2, \dots\}$ 是独立同分布高斯 $N(0, 1)$ 随机变量序列。试分别

在 $p = 1, 2$ 和 3 时, 确定最小方差控制策略和控制误差。

4.8 考虑由

$$A(q^{-1})y(k) = B(q^{-1})u(k - p)$$

描述的确定性系统。试证明控制策略

$$u(k) = - \frac{G(q^{-1})}{B(q^{-1})F(q^{-1})}y(k)$$

具有使系统的输出在 p 步时达到零的特性。上式中多项式 $F(q^{-1})$ 和 $G(q^{-1})$ 由恒等式

$$1 = A(q^{-1})F(q^{-1}) + q^{-p}G(q^{-1})$$

确定。这种策略叫做非周期性策略。

4.9 考虑系统

$$A(q^{-1})y(k) = B(q^{-1})u(k - p) + C(q^{-1})e(k)$$

试在非周期性策略

$$u(k) = - \frac{G(q^{-1})}{B(q^{-1})F(q^{-1})}y(k)$$

$$1 = A(q^{-1})F(q^{-1}) + q^{-p}G(q^{-1})$$

调节时, 确定输出的方差 $\text{var}[y(k)]$ 。

4.10 考虑由

$$y(k) = \frac{B_1(q^{-1})}{A_1(q^{-1})}u(k - p) + \frac{C_1(q^{-1})}{A_2(q^{-1})}e(k)$$

描述的系统。如果多项式 $A_2(z)$ 在 $z = 1$ 有单一零点, 并且 $A_1(1) \neq 0$, 试证明最小方差控制策略将总包含一个积分。

4.11 对于由

$$y(t) + 0.5y(t - 1) = e(t) + 2e(t - 1) + u(t - 1)$$

描述的系统, 试确定最小方差控制策略和控制误差的方差。

4.12 考虑由下式描述的系统:

$$\frac{1}{1 + 2q^{-1}}y(k) = \frac{1}{1 + 2q^{-1}}u(k - 1) + \frac{1 + 0.7q^{-1}}{1 - 0.2q^{-1}}e(k)$$

式中: $e(k)$ 是独立同分布高斯 $N(0, 1)$ 随机变量序列。试确定系统的最小方差控制策略。

4.13 PI 调节器的数字模型可用以下方程来描述:

$$u(k) = u(k - 1) + K \left(1 + \frac{h}{T} \right) y(k) - y(k - 1)$$

式中: K, h 和 T 是正数。试证明上式算法是最小方差控制律的形如

$$A(q^{-1})y(k) = B(q^{-1})u(k - p) + C(q^{-1})e(k)$$

的一般系统。

4.14 考虑多项式恒等式

$$C(z) = A(z)F(z) + B(z)G(z)$$

式中: A, B 和 C 分别是 n_A, n_B 和 n_C 阶的已知多项式。假设 $n_C < n_A + n_B - 1$, 试证明必存在 $n_F = n_B - 1$ 和 $n_G = n_A - 1$ 阶的唯一多项式 $F(z)$ 和 $G(z)$ 满足上等式。并证明存在很多 $n_F > n_B - 1$ 和 $n_G > n_A - 1$ 阶的多项式 $F(z)$ 和 $G(z)$ 满足上等式。

4.15 考虑系统

$$A(q^{-1})y(k) = B(q^{-1})u(k-p) + C(q^{-1})e(k)$$

当 $l > p$ 时, 试证明存在一种控制策略, 使得调节误差等于 l 步预测输出的误差。

4.16 试把次最优控制策略的控制误差的方差 $\text{var}[y(k)]$ 与 $p+n_2$ 步预测的误差的方差 $\text{var}[\tilde{y}(k+p+n_2|k)]$ 进行比较。其中 n_2 是 $B(q^{-1})$ 在单位圆外的零点个数。

4.17 为求系统

$$A(q)y(k) = B_1(q)B_2(q)u(k-p) + C(q)e(k)$$

的次最优策略, 用恒等式

$$H(q)C(q) = A(q)F(q) + B_2(q)G(q)$$

也能确定 $p+n_2-1$ 阶多项式 $F(q)$ 和 $n-1$ 阶多项式 $G(q)$, 式中 $H(q)$ 是 $p+n_2-1$ 阶的任意多项式。试证明控制律

$$u(k) = - \frac{q^p G(q)}{B_1(q)F(q)} y(k)$$

给出的控制误差为

$$H(q)y(k) = F(q)e(k)$$

并证明闭环系统的特征方程为

$$H(z)B_1(z)C(z) = 0$$

4.18 考虑由

$$A(q^{-1})y(k) = B(q^{-1})u(k-p) + C(q^{-1})e(k)$$

给定的系统。试证明控制策略

$$u(k) = - \frac{G_1(q^{-1})}{F_1(q^{-1})} y(k)$$

给出了一个特征方程为

$$z^{n+p-1}C(z) = 0$$

的闭环系统, 式中 $n+p-1$ 阶多项式 $F_1(q^{-1})$ 和 $n-1$ 阶多项式 $G_1(q^{-1})$ 由恒等式

$$C(q^{-1}) = A(q^{-1})F_1(q^{-1}) + q^{-p}B(q^{-1})G_1(q^{-1})$$

来定义。再证明此闭环系统对参数 $B(q^{-1})$ 的变化并不特别敏感。并计算控制误差的方差, 且与最小方差进行比较。

4.19 考虑系统

$$y(k) + ay(k-1) = u(k-1) + 2.5u(k-2) + u(k-3) + e(k) + ce(k-1)$$

试确定最小方差控制策略和最小控制误差, 说明最优系统对参数变化是非常敏感的。再分别利用本章 4.5.2 小节和习题 18 提供的方法, 推导出对参数变化不敏感的两个次最优控制策略和控制误差。

4.20 已知系统

$$x(k+1) = -0.8x(k) + u(k) + 0.3e(k)$$

且 $x(k) \sim N(m_x(k), R_x(k))$ 与 $e(k) \sim N(2, 1)$ 互相独立。试求出使性能指标 $J = E\{x(k+1)^2\}$ 为最小的控制策略 $u(k)$, 并计算 $\min J$ 的值。实行这种控制后, $x(k+1)$ 等于什么?

第 5 章 最优状态估计

5.1 引言

1. 状态估计的概念

我们研究的动力学系统, 不仅受到固有控制输入的作用, 而且也受到扰动的作用, 而这种扰动既不能控制又不能在模型中确定。同时, 传感器也受到噪声的污染, 使量测信号包括了噪声信号, 不能直接把有用信号(状态)量测出来。如果根据这些信号直接对系统进行控制, 不会得到期望的结果。例如, 飞行器驾驶员给定一个方位角控制, 由于风的扰动和执行机构的不精确性, 可能不会产生准确的期望响应。这里存在这样一个问题, 如何把有用信号(状态)估计出来, 以实现准确的控制。

随机系统是一种遭受噪声作用的动力学系统, 简单情况表示为

$$y(t) = s(t) + n(t) \tag{5.1.1}$$

有用信号(状态) $s(t)$ 受到了噪声 $n(t)$ 的污染, 量测信号 $y(t)$ 只能是信噪信号。如何从信噪信号 $y(t)$ 中滤去噪声 $n(t)$ 的影响, 得到有用信号 $s(t)$, 这种问题称为状态估计(state estimation)。

除包括上述简单问题的状态估计外, 一般是指离散时间状态空间方程

$$x(k+1) = \Phi x(k) + \Gamma u(k) \tag{5.1.2}$$

$$y(k) = H x(k) + v(k) \tag{5.1.3}$$

和连续时间状态空间方程

$$\dot{x}(t) = A x(t) + B u(t) + w(t) \tag{5.1.4}$$

$$y(t) = H x(t) + v(t) \tag{5.1.5}$$

的状态估计。问题的提法是, 如何由量测值 $y(t)$ [或 $y(k)$] 把状态值 $x(t)$ [或 $x(k)$] 估计出来, 而使某种性能指标为最小。

线性随机控制问题可分为两部分, 一部分是状态估计, 另一部分是根据状态估计建立状态反馈, 即确定最优控制律, 这称为分离定理。因此, 状态估计是随机控制的重要组成部分。

2. 状态估计的分类

考虑用时刻 k 时的量测值 $y(k)$ 估计时刻 k_1 时的状态值 $x(k_1)$, 表示为 $\hat{x}(k_1|k)$, 其意义为

$$\begin{aligned} \hat{x}(k_1|k) &= \hat{x}(k_1|y(k), y(k-1), \dots, y(k_0)) \\ &= \hat{x}(k_1|Y_k) \end{aligned} \tag{5.1.6}$$

式中:

$$Y_k = [y^T(k) \ y^T(k-1) \ \dots \ y^T(k_0)] \tag{5.1.7}$$

按 k 和 k_1 的关系, 状态估计可分为三类:

- (1) 当 $k_1 > k$ 时, 称为预测(prediction);
- (2) 当 $k_1 = k$ 时, 称为滤波(filter);
- (3) 当 $k_1 < k$ 时, 称为平滑(smoothing) 或内插(intorpolation)。

有时, 状态估计也称为滤波。

3. 对性能指标的规定

状态估计的好与差用性能指标函数来衡量。对性能指标函数的要求是, 能反映估计效果和容易计算。依照不同要求, 可规定不同的性能指标, 从而得到各种状态估计。

对性能指标函数做如下规定: 首先, 规定性能指标函数 $l(s)$ 是一个对称的非递减的正的实可积函数, 即 $l(s) = l(-s)$; 当 $s_2 \geq s_1$ 时, $l(s_2) \geq l(s_1)$; 当 $s = 0$ 时, $l(s) = 0$; $l(s)$ 是一个标量实函数; $l(s)$ 是可积函数。其次, 设状态估计值为 \hat{x} , 把估计的偏差 $x - \hat{x}$ 定为性能指标的自变量, 即 $s = x - \hat{x}$ 是合适的。最后, 由于状态 x 是一个随机过程, 并且其估计是在量测值 y 的条件下得到的, 因此, 用一个条件均值 $E\{l(x - \hat{x})|y\}$ 来表示性能指标是很自然的。综上所述, 将完全的性能指标函数规定为

$$J = E\{l(x - \hat{x})|y\} = \int_{-\infty}^{\infty} l(x - \hat{x}) f(x|y) dx \tag{5.1.8}$$

式中: $f(x|y)$ 表示在 y 条件下的 x 的条件密度函数; 性能指标函数 J 是量测值 y 的函数。使性能指标函数 J 为最小的状态估计 \hat{x} 称为最优估计。

4. 本章主要内容

本章先证明最优状态估计与条件均值的等价性, 再求出在高斯分布下状态估计的基本计算公式, 然后利用这些基本公式求解预测、滤波和平滑问题。目前存在多种状态估计方法, 本书集中介绍卡尔曼滤波方法。卡尔曼滤波方法是一种进行数据处理的最优递推算法。所谓递推是指, 在计算中不需要把所有以前的数据保存在存储器内, 只需处理每一时刻取到的新的量测值。卡尔曼滤波由卡尔曼(Kalman)和布西(Bucy)在 1960 年提出, 为随机控制的发展做出了重大贡献, 并且首先在航天工业, 后在多个领域得到了广泛的应用。在本章前面部分介绍离散时间状态方程的卡尔曼滤波方法, 在后面部分介绍连续时间状态方程的卡尔曼滤波方法。

5.2 状态估计与条件均值的等价性

本节证明在一定条件下, 使性能指标函数式(5.1.8)为最小的最优状态估计就是条件均值。先引入一个引理, 再用引理证明一个定理, 这个定理证明上述结论的正确性。

5.2.1 引理

引理 5.2.1 设 $g(x)$ 为对称非减正实可积函数, 即有 $g(x) = g(-x)$, 对 $x_2 \geq x_1$ 有 $g(x_2) \geq g(x_1)$, 对 $x \geq 0$ 有 $g(x) \geq 0$; 设 $h(x)$ 为对称非增正实可积函数, 即有 $h(x) = h(-x)$, 对 $x_2 \geq x_1$ 有 $h(x_2) \leq h(x_1)$, 对 $x \geq 0$ 有 $h(x) \geq 0$ 。则对任意实数 a , 必有

$$\int_{-\infty}^{\infty} g(x+a)h(x) dx = \int_{-\infty}^{\infty} g(x)h(x) dx \tag{5.2.1}$$

成立。

证明 令 a 为零、负数和正数等三种情况, 分别加以证明。

(1) $a = 0$

式(5.2.1)显然成立。

(2) $a < 0$

改写式(5.2.1), 只要证明

$$[g(x - |a|) - g(x)]h(x)dx \geq 0 \quad (5.2.2)$$

就够了。为此, 对式(5.2.2)左式进行运算

$$\begin{aligned} & [g(x - |a|) - g(x)]h(x)dx \\ &= \int_{-\frac{|a|}{2}}^{\frac{|a|}{2}} [g(x - |a|) - g(x)]h(x)dx + \int_{\frac{|a|}{2}}^{\frac{|a|}{2}} [g(x - |a|) - g(x)]h(x)dx \quad (5.2.3) \end{aligned}$$

对上式右边第一项做 $x = x + |a|$ 变量置换, 并利用给定函数对称性等条件, 得到

$$\begin{aligned} & [g(x - |a|) - g(x)]h(x)dx \\ &= \int_{\frac{|a|}{2}}^{\frac{|a|}{2}} [g(x) - g(x - |a|)]h(x - |a|)dx + \int_{\frac{|a|}{2}}^{\frac{|a|}{2}} [g(x) - g(x - |a|)]h(x)dx \\ &= \int_{\frac{|a|}{2}}^{\frac{|a|}{2}} [g(x) - g(x - |a|)][h(x - |a|) - h(x)]dx \quad (5.2.4) \end{aligned}$$

利用 $g(x)$ 的非减性和 $h(x)$ 的非增性, 有

$$g(x) - g(x - |a|) \geq 0 \quad \frac{|a|}{2} \leq x < \frac{|a|}{2} \quad (5.2.5)$$

$$h(x - |a|) - h(x) \geq 0 \quad \frac{|a|}{2} \leq x < \frac{|a|}{2} \quad (5.2.6)$$

把以上两式代入式(5.2.4), 则式(5.2.2)成立。

(3) $a > 0$

同样, 改写式(5.2.1), 只要证明

$$[g(x + a) - g(x)]h(x)dx \geq 0 \quad (5.2.7)$$

就够了。为此, 对式(5.2.7)左式进行运算:

$$\begin{aligned} & [g(x + a) - g(x)]h(x)dx \\ &= \int_{-\frac{a}{2}}^{\frac{a}{2}} [g(x + a) - g(x)]h(x)dx + \int_{\frac{a}{2}}^{\frac{a}{2}} [g(x + a) - g(x)]h(x)dx \quad (5.2.8) \end{aligned}$$

对上式右边第一项做 $x = x$ 变量置换, 对第二项做 $x = x + a$ 变量置换, 得到

$$\begin{aligned} & [g(x + a) - g(x)]h(x)dx \\ &= \int_{\frac{a}{2}}^{\frac{a}{2}} [g(x - a) - g(x)]h(x)dx + \int_{\frac{a}{2}}^{\frac{a}{2}} [g(x) - g(x - a)]h(x - a)dx \\ &= \int_{\frac{a}{2}}^{\frac{a}{2}} [g(x) - g(x - a)][h(x - a) - h(x)]dx \quad (5.2.9) \end{aligned}$$

对被积函数, 式(5.2.5)和(5.2.6)成立, 于是得到

$$\int [g(x + \mathbb{Q} \mathbb{Y}) - g(x)]h(x)dx = 0 \tag{5.2.10}$$

即式(5.2.7)成立。

综上所述, 式(5.2.1)成立, 引理 5.2.1 得证。

5.2.2 定理

定理 5.2.1 设 $x(t)$ 在 y 条件下的条件密度函数 $f(x|\mathbb{Y})$ 对自身均值

$$E\{x|\mathbb{Y}\} = \int xf(x|\mathbb{Y})dx \tag{5.2.11}$$

是对称的, 且对 $x = E\{x|\mathbb{Y}\}$ 是非增的。规定性能指标函数 $l(t)$ 为对称非减正实可积函数, 即有 $l(t) = l(-t)$, 对 $t_2 > t_1$ 有 $l(t_2) \geq l(t_1)$, 对 $t \geq 0$ 有 $l(t) \geq 0$ 。那么由 y 估计 x 的最优估计 $\hat{x}(y)$ 由条件均值给定, 即

$$\hat{x}(y) = E\{x|\mathbb{Y}\} \tag{5.2.12}$$

证明 最优估计是通过使性能指标函数式(5.1.8)

$$J = \int l[x - \hat{x}(y)]f(x|\mathbb{Y})dx \tag{5.2.13}$$

为最小来获得, 现直接对它进行处理。做变量置换, 令

$$t + E\{x|\mathbb{Y}\} = x \tag{5.2.14}$$

并代入式(5.2.13), 得到

$$J = \int l[t + E\{x|\mathbb{Y}\} - \hat{x}(y)]f(t + E\{x|\mathbb{Y}\}|\mathbb{Y})dt \tag{5.2.15}$$

把上式中 $f(t + E\{x|\mathbb{Y}\})$ 看做是 t 的函数, 其意义是把 x 的坐标原点平移到了 $E\{x|\mathbb{Y}\}$, 把 t 作为自变量, $f(t + E\{x|\mathbb{Y}\})$ 对 $E\{x|\mathbb{Y}\}$ 是对称的, 于是 $f(t + E\{x|\mathbb{Y}\})$ 符合引理 5.2.1 中给定函数 $h(t)$ 的条件。再看 $l[t + E\{x|\mathbb{Y}\} - \hat{x}(y)]$, 令

$$a = E\{x|\mathbb{Y}\} - \hat{x}(y) \tag{5.2.16}$$

则按给定条件, $l(t)$ 满足引理 5.2.1 中给定函数 $g(t)$ 的要求。于是, 由式(5.2.1)得到

$$\int l[t + E\{x|\mathbb{Y}\} - \hat{x}(y)]f(t + E\{x|\mathbb{Y}\}|\mathbb{Y})dt = \int l(t)f(t + E\{x|\mathbb{Y}\}|\mathbb{Y})dt \tag{5.2.17}$$

上式右边是性能指标函数式(5.2.13)的最小值, 很明显, 上式取最小值的条件是式(5.2.16)为零, 即

$$\hat{x}(y) = E\{x|\mathbb{Y}\} \tag{5.2.18}$$

上式即为要求证的式(5.2.12)。

定理 5.2.1 极为重要, 它表明, 只要 $x(t)$ 在 y 条件下的条件密度函数满足一定条件, 如定理 5.2.1 所述, 则使性能指标函数式(5.2.13)为最小的最优估计 $\hat{x}(y)$ 就是条件均值 $E\{x|\mathbb{Y}\}$, 这就提供了一种计算最优估计的算法。高斯分布满足定理 5.2.1 的给定条件, 下

面就对这种情况做进一步的讨论。

5.3 多维高斯分布的某些结果及其几何解释

我们研究的实际过程多数为高斯过程, 高斯过程的条件密度函数 $f(x|y)$ 符合定理 5.2.1 的条件, 因此, 用量测值 y 对 x 的最优估计为

$$\hat{x}(y) = E\{x|y\} \tag{5.3.1}$$

这一节中, 在式(5.3.1)基础上, 导出各种情况的最优估计的计算公式, 并在几何上把 x 在量测 y 下的最优估计解释为 x 在 y 上的投影。

5.3.1 定理 5.3.1

定理 5.3.1 设 x 和 y 分别为 $n \times 1$ 和 $p \times 1$ 的高斯向量, 向量 $\begin{pmatrix} x \\ y \end{pmatrix}$ 也是联合高斯的, 其均值和协方差函数分别为

$$m = \begin{pmatrix} m_x \\ m_y \end{pmatrix} \quad R = \begin{pmatrix} R_x & R_{xy} \\ R_{yx} & R_y \end{pmatrix} \tag{5.3.2}$$

构造向量

$$z = x - m_x - R_{xy}R_y^{-1}(y - m_y) \tag{5.3.3}$$

则向量 z 的均值 m_z 为零, 且对 y 独立, 协方差为

$$R_z = R_x - R_{xy}R_y^{-1}R_{yx} \tag{5.3.4}$$

证明 对式(5.3.3)两边取均值, 立即得到 m_z 为零

$$m_z = E\{x\} - m_x - R_{xy}R_y^{-1}(E\{y\} - m_y) = 0 \tag{5.3.5}$$

利用式(5.3.3)和(5.3.5)计算 z 和 y 的协方差函数

$$\begin{aligned} R_{zy} &= E\{z(y - m_y)^T\} \\ &= E\{(x - m_x)(y - m_y)^T\} - R_{xy}R_y^{-1}E\{(y - m_y)(y - m_y)^T\} \\ &= R_{xy} - R_{xy}R_y^{-1}R_y \\ &= 0 \end{aligned} \tag{5.3.6}$$

上式表明, z 和 y 不相关, 由于对于高斯过程不相关等价于独立, 所以 z 和 y 互相独立。

最后, 求协方差函数 R_z , 计算中注意到 $R_x = R_x^T, R_y = R_y^T, R_{xy} = R_{yx}^T$, 得到

$$\begin{aligned} R_z &= E\{zz^T\} \\ &= E\{(x - m_x)(x - m_x)^T\} - E\{(x - m_x)(y - m_y)^T\}R_y^{-1}R_{yx} \\ &\quad - R_{xy}R_y^{-1}E\{(y - m_y)(x - m_x)^T\} + R_{xy}R_y^{-1}E\{(y - m_y)(y - m_y)^T\}R_y^{-1}R_{yx} \\ &= R_x - R_{xy}R_y^{-1}R_{yx} \end{aligned} \tag{5.3.7}$$

5.3.2 定理 5.3.2

定理 5.3.2 考虑 x 和 y 分别为 $n \times 1$ 和 $p \times 1$ 的高斯向量, 并且是联合高斯的, 那么 x 在 y 条件下的条件分布是高斯的, 其条件均值和条件协方差函数分别为

$$\mathbf{m}_{x|y} = \mathbf{m}_x + \mathbf{R}_{xy}\mathbf{R}_y^{-1}(\mathbf{y} - \mathbf{m}_y) \quad (5.3.8)$$

$$\begin{aligned} \mathbf{R}_{x|y} &= E\{[\mathbf{x} - E\{\mathbf{x}|y\}][\mathbf{x} - E\{\mathbf{x}|y\}]^T\} \\ &= \mathbf{R}_x - \mathbf{R}_{xy}\mathbf{R}_y^{-1}\mathbf{R}_{yx} \end{aligned} \quad (5.3.9)$$

并且 y 和 $\mathbf{x} - \mathbf{m}_{x|y}$ 互相独立。

证明 在已知 $f(x, y)$ 和 $f(y)$ 为

$$f(x, y) = (2\pi)^{-\frac{n+p}{2}} |\mathbf{R}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \begin{bmatrix} \mathbf{x} - \mathbf{m}_x \\ \mathbf{y} - \mathbf{m}_y \end{bmatrix}^T \mathbf{R}^{-1} \begin{bmatrix} \mathbf{x} - \mathbf{m}_x \\ \mathbf{y} - \mathbf{m}_y \end{bmatrix} \right\} \quad (5.3.10)$$

$$f(y) = (2\pi)^{-\frac{p}{2}} |\mathbf{R}_y|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{y} - \mathbf{m}_y)^T \mathbf{R}_y^{-1} (\mathbf{y} - \mathbf{m}_y) \right\} \quad (5.3.11)$$

条件下, 利用条件密度公式

$$f(\mathbf{x}|y) = \frac{f(\mathbf{x}, y)}{f(y)} \quad (5.3.12)$$

求出条件密度函数 $f(\mathbf{x}|y)$, 进而得到结果。为此, 对式(5.3.10)进行分解。首先分解 $|\mathbf{R}|$

$$\begin{aligned} |\mathbf{R}| &= \begin{vmatrix} \mathbf{R}_x & \mathbf{R}_{xy} \\ \mathbf{R}_{yx} & \mathbf{R}_y \end{vmatrix} = \begin{vmatrix} \mathbf{R}_x - \mathbf{R}_{xy}\mathbf{R}_y^{-1}\mathbf{R}_{yx} & \mathbf{R}_{xy} - \mathbf{R}_{xy}\mathbf{R}_y^{-1}\mathbf{R}_y \\ \mathbf{R}_{yx} & \mathbf{R}_y \end{vmatrix} \\ &= |\mathbf{R}_z| |\mathbf{R}_y| \end{aligned} \quad (5.3.13)$$

式中第二等式由第一等式的第一行减去左乘 $\mathbf{R}_{xy}\mathbf{R}_y^{-1}$ 的第二行得到, 第二等式到第三等式利用了式(5.3.4)。

其次分解式(5.3.10)的指数部分, 注意利用式(5.3.3), 得到

$$\begin{aligned} \begin{bmatrix} \mathbf{x} - \mathbf{m}_x \\ \mathbf{y} - \mathbf{m}_y \end{bmatrix} &= \begin{bmatrix} \mathbf{z} + \mathbf{R}_{xy}\mathbf{R}_y^{-1}(\mathbf{y} - \mathbf{m}_y) \\ \mathbf{y} - \mathbf{m}_y \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{I} & \mathbf{R}_{xy}\mathbf{R}_y^{-1} \\ 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{z} \\ \mathbf{y} - \mathbf{m}_y \end{bmatrix} \end{aligned} \quad (5.3.14)$$

$$\begin{bmatrix} \mathbf{x} - \mathbf{m}_x \\ \mathbf{y} - \mathbf{m}_y \end{bmatrix}^T = \begin{bmatrix} \mathbf{z}^T & (\mathbf{y} - \mathbf{m}_y)^T \mathbf{R}_y^{-1} \mathbf{R}_{yx} \end{bmatrix} \begin{bmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{I} \end{bmatrix} \quad (5.3.15)$$

再利用已有矩阵求逆公式

$$\begin{aligned} \mathbf{R}^{-1} &= \begin{bmatrix} \mathbf{R}_x & \mathbf{R}_{xy} \\ \mathbf{R}_{yx} & \mathbf{R}_y \end{bmatrix}^{-1} \\ &= \begin{bmatrix} \mathbf{R}_z^{-1} & -\mathbf{R}_z^{-1}\mathbf{R}_{xy}\mathbf{R}_y^{-1} \\ -\mathbf{R}_y^{-1}\mathbf{R}_{yx}\mathbf{R}_z^{-1} & \mathbf{R}_y^{-1} + \mathbf{R}_y^{-1}\mathbf{R}_{yx}\mathbf{R}_z^{-1}\mathbf{R}_{xy}\mathbf{R}_y^{-1} \end{bmatrix} \end{aligned} \quad (5.3.16)$$

由式(5.3.14), (5.3.15)和(5.3.16)导出式(5.3.10)的指数部分为

$$\begin{aligned} \begin{bmatrix} \mathbf{x} - \mathbf{m}_x \\ \mathbf{y} - \mathbf{m}_y \end{bmatrix}^T \mathbf{R}^{-1} \begin{bmatrix} \mathbf{x} - \mathbf{m}_x \\ \mathbf{y} - \mathbf{m}_y \end{bmatrix} &= \begin{bmatrix} \mathbf{z}^T & (\mathbf{y} - \mathbf{m}_y)^T \mathbf{R}_y^{-1} \mathbf{R}_{yx} \end{bmatrix} \begin{bmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{z} \\ \mathbf{y} - \mathbf{m}_y \end{bmatrix} \\ &= \mathbf{z}^T \mathbf{R}_z^{-1} \mathbf{z} + (\mathbf{y} - \mathbf{m}_y)^T \mathbf{R}_y^{-1} (\mathbf{y} - \mathbf{m}_y) \end{aligned} \quad (5.3.17)$$

把式(5.3.13)和(5.3.17)代入式(5.3.10), 并注意到式(5.3.11), 得到

$$f(\mathbf{x}, y) = f(\mathbf{z})f(y) \quad (5.3.18)$$

把式(5.3.18)代入(5.3.12), 得到

$$f(\mathbf{x}|y) = f(\mathbf{z})$$

$$= (2\pi)^{-\frac{n}{2}} |\mathbf{R}_z|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}[\mathbf{x} - (\mathbf{m}_x + \mathbf{R}_{xy}\mathbf{R}_y^{-1}(\mathbf{y} - \mathbf{m}_y))]^T \mathbf{R}_z^{-1}[\cdot]\right\} \quad (5.3.19)$$

由上式看出, 条件密度函数 $f(\mathbf{x}|\mathbf{y})$ 是高斯的, 其条件均值和条件协方差函数分别由式(5.3.8)和(5.3.9)给出。

由式(5.3.8)和(5.3.3)得到

$$\mathbf{z} = \mathbf{x} - \mathbf{m}_{x|y} \quad (5.3.20)$$

定理 5.3.1 已经证明 y 与 z 互相独立, 也就是 y 与 $\mathbf{x} - \mathbf{m}_{x|y}$ 互相独立。

5.3.3 定理 5.3.3

定理 5.3.3 设 \mathbf{x} , \mathbf{u} 和 \mathbf{v} 是具有联合高斯分布的随机向量, 且 \mathbf{u} 和 \mathbf{v} 是互相独立的, 则

$$E\{\mathbf{x}|\mathbf{u}, \mathbf{v}\} = E\{\mathbf{x}|\mathbf{u}\} + E\{\mathbf{x}|\mathbf{v}\} - \mathbf{m}_x \quad (5.3.21)$$

证明 利用定理 5.3.2 中式(5.3.8)证明, 为此, 令

$$\mathbf{y} = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} \quad (5.3.22)$$

则有

$$\mathbf{m}_y = \begin{bmatrix} \mathbf{m}_u \\ \mathbf{m}_v \end{bmatrix} \quad \mathbf{R}_y = \begin{bmatrix} \mathbf{R}_u & 0 \\ 0 & \mathbf{R}_v \end{bmatrix} \quad \mathbf{R}_y^{-1} = \begin{bmatrix} \mathbf{R}_u^{-1} & 0 \\ 0 & \mathbf{R}_v^{-1} \end{bmatrix} \quad (5.3.23)$$

$$\begin{aligned} \mathbf{R}_{xy} &= E\{(\mathbf{x} - \mathbf{m}_x)(\mathbf{y} - \mathbf{m}_y)^T\} \\ &= E\left\{(\mathbf{x} - \mathbf{m}_x) \begin{bmatrix} \mathbf{u} - \mathbf{m}_u \\ \mathbf{v} - \mathbf{m}_v \end{bmatrix}^T\right\} \\ &= [\mathbf{R}_{xu}, \mathbf{R}_{xv}] \end{aligned} \quad (5.3.24)$$

把式(5.3.22), (5.3.23)和(5.3.24)代入式(5.3.8), 计算 $E\{\mathbf{x}|\mathbf{u}, \mathbf{v}\}$, 得到

$$\begin{aligned} E\{\mathbf{x}|\mathbf{u}, \mathbf{v}\} &= E\{\mathbf{x}|\mathbf{y}\} \\ &= \mathbf{m}_x + \mathbf{R}_{xy}\mathbf{R}_y^{-1}(\mathbf{y} - \mathbf{m}_y) \\ &= \mathbf{m}_x + [\mathbf{R}_{xu} \mathbf{R}_{xv}] \begin{bmatrix} \mathbf{R}_u^{-1} & 0 \\ 0 & \mathbf{R}_v^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{u} - \mathbf{m}_u \\ \mathbf{v} - \mathbf{m}_v \end{bmatrix} \\ &= \mathbf{m}_x + \mathbf{R}_{xu}\mathbf{R}_u^{-1}(\mathbf{u} - \mathbf{m}_u) + \mathbf{R}_{xv}\mathbf{R}_v^{-1}(\mathbf{v} - \mathbf{m}_v) \\ &= E\{\mathbf{x}|\mathbf{u}\} + E\{\mathbf{x}|\mathbf{v}\} - \mathbf{m}_x \end{aligned} \quad (5.3.25)$$

5.3.4 三个定理的物理意义和几何解释

1. 物理意义

由定理 5.2.1 已知, 当 \mathbf{x} 在 y 条件下的条件密度函数对其自身均值 $E\{\mathbf{x}|\mathbf{y}\}$ 对称时, 使给定性能指标函数为最小的最优估计 $\hat{\mathbf{x}}(y)$ 就等于条件均值 $E\{\mathbf{x}|\mathbf{y}\}$

$$\hat{\mathbf{x}}(y) = E\{\mathbf{x}|\mathbf{y}\} \quad (5.3.26)$$

定理 5.3.2 已经证明, 当 \mathbf{x} 和 \mathbf{y} 具有联合高斯分布时, \mathbf{x} 在 y 条件下的条件密度函数是高斯的, 并求出了条件均值 $E\{\mathbf{x}|\mathbf{y}\}$, 因为高斯条件密度函数是符合定理 5.2.1 条件的, 所以 \mathbf{x} 在量测 y 下的最优估计 $\hat{\mathbf{x}}(y)$ 就由式(5.3.8)确定, 即

$$\begin{aligned}\hat{x}(y) &= E\{x|y\} \\ &= m_x + R_{xy}R_y^{-1}(y - m_y)\end{aligned}\quad (5.3.27)$$

估计误差 $\tilde{x}(y)$ 由式(5.3.27)和(5.3.20)得到

$$\begin{aligned}\tilde{x}(y) &= x - \hat{x}(y) = z \\ &= x - [m_x + R_{xy}R_y^{-1}(y - m_y)]\end{aligned}\quad (5.3.28)$$

由上式可知, 定理 5.3.1 中构造的向量 z 就是估计误差 $\tilde{x}(y)$ 。已经证明, z , 即 $\tilde{x}(y)$ 的均值为零, 因此, 这种估计是无偏估计。估计误差的协方差函数由式(5.3.9)给出, 即

$$E\{\tilde{x}(y)\tilde{x}^T(y)\} = R_x - R_{xy}R_y^{-1}R_{yx}\quad (5.3.29)$$

定理 5.3.3 的物理意义也是明显的, 当两个量测值 u 和 v 互相独立时, x 在量测 u 和 v 下的最优估计 $\hat{x}(u, v)$ 等于 x 分别在量测 u 和 v 下的最优估计 $\hat{x}(u)$, $\hat{x}(v)$ 的和, 再减去均值 m_x , 也就是如式(5.3.25)所示。

式(5.3.27), (5.3.28), (5.3.29)和(5.3.25)为我们提供了计算最优估计的具体公式。

2. 几何解释

我们已经证明, 在一定条件下, 最优估计 $\hat{x}(y)$ 就等于条件均值 $E\{x|y\}$, 在几何意义上如何解释呢? 实际上, 这个条件均值 $E\{x|y\}$ 可看做 x 在 y 上的投影, 下面就证明这一点。

设 x 和 y 是直角坐标系的两个随机向量, 其夹角为 θ , 并且为简便起见, 令 x 和 y 的均值为零。

与确定性向量内积定义式(2.7.1)类似, 对 x 和 y 两个随机向量的内积定义为

$$(x, y) = |x| |y| \cos \theta \quad (5.3.30)$$

式中: (x, y) 为随机向量的内积, 是标量; $|x|$ 和 $|y|$ 分别为随机向量 x 和 y 的范数, 由式(2.7.6)定义。

由内积定义式(5.3.30)可得到随机向量 x 在随机向量 y 上的投影值

$$|x| \cos \theta = \frac{(x, y)}{|y|} \quad (5.3.31)$$

把这个投影表示为向量形式, x 在 y 上的投影, 其方向与 y 方向相同, 而 y 方向上的单位向量为 $\frac{y}{|y|}$, 所以 x 在 y 上的投影向量表示为

$$\overline{|x| \cos \theta} = \frac{(x, y)y}{|y|^2} \quad (5.3.32)$$

由 2.7.1 小节已知, 随机向量 x 和 y 的内积等于它们的相关函数, 因为 x 和 y 的均值为零, 所以它们的相关函数就是它们的协方差函数, 也就是

$$(x, y) = E\{xy\} = R_{xy} \quad (5.3.33)$$

$$|y|^2 = (y, y) = R_y \quad (5.3.34)$$

把以上两式代入式(5.3.32), 得到

$$\overline{|x| \cos \theta} = R_{xy}R_y^{-1}y \quad (5.3.35)$$

把上式与式(5.3.27)对照, 并注意到已令 $m_x = 0$, 得到

$$\hat{x}(y) = E\{x|y\} = \overline{|x| \cos \theta} \quad (5.3.36)$$

上式表明, 最优估计 $\hat{x}(y)$ 在几何上就是 x 在 y 上的投影, 如图 5.3.1 所示。把 $\hat{x}(y)$ 看做为 x 的一个分量, 则 $\tilde{x}(y)$ 是另一个分量, 已经证明, $\tilde{x}(y)$ 与 y 独立, 在几何上就表现为两者互相垂直。

最后, 说明定理 5.3.3 的几何解释。当 u 和 v 互相独立和 $m_x = 0$ 时, 式(5.3.25)变为

$$\hat{x}(u, v) = E\{x|u\} + E\{x|v\} = \hat{x}(u) + \hat{x}(v) \tag{5.3.37}$$

这表明, 最优估计 $\hat{x}(u, v)$ 在几何上是 x 在由 u 和 v 张成的线性子空间上的正交投影, 它等于 x 在 u 上的投影 $\hat{x}(u)$ 与 x 在 v 上的投影 $\hat{x}(v)$ 的和, 如图 5.3.2 所示。

图 5.3.1 把最优估计 $\hat{x}(y)$ 表示为 x 在 y 上的投影

图 5.3.2 定理 5.3.3 的几何解释

5.3.5 定理 5.3.4(投影定理)

定理 5.3.4 设 y_1, y_2, \dots, y_l 和 x 是欧几里得空间中的元素, Y 是 y_1, y_2, \dots, y_l 张成的线性子空间, 那么必存在一个唯一的元素 $\hat{x} \in Y$, 使得

$$x - \hat{x} = \inf_{z \in Y} x - z \tag{5.3.38}$$

证明 设 \hat{x} 是 x 在 Y 上的正交投影, 那么有

$$(x - \hat{x}, y_i) = 0 \quad (i = 1, 2, \dots, l) \tag{5.3.39}$$

任取一向量 $z \in Y$,

$$z = \sum_{i=1}^l \alpha_i y_i \tag{5.3.40}$$

式中: $\alpha_i (i = 1, 2, \dots, l)$ 为系数。则有

$$\begin{aligned} (x - \hat{x}, z) &= (x - \hat{x}, \sum_{i=1}^l \alpha_i y_i) \\ &= \sum_{i=1}^l \alpha_i (x - \hat{x}, y_i) \\ &= 0 \end{aligned} \tag{5.3.41}$$

由式(5.3.40)和(5.3.41)表明, 正交投影 $x - \hat{x}$ 与任一 $y_i (i = 1, 2, \dots, l)$ 及它们的任何线性组合都是垂直的。因此有

$$\begin{aligned} x - z &^2 = (x - z, x - z) \\ &= ((x - \hat{x}) - (z - \hat{x}), (x - \hat{x}) - (z - \hat{x})) \\ &= (x - \hat{x}, x - \hat{x}) - 2(x - \hat{x}, z - \hat{x}) + (z - \hat{x}, z - \hat{x}) \\ &= x - \hat{x}^2 + z - \hat{x}^2 \end{aligned} \tag{5.3.42}$$

上式右边两项都是正的,且只有 z 是可变的,于是可把上式进一步写为

$$x - z^2 = x - \hat{x}^2 \tag{5.3.43}$$

上式表明, $x - \hat{x}^2$ 是最小值,只有当 $z = \hat{x}$ 时,式(5.3.43)才取等号,即取最小值。

综上所述,正交投影 \hat{x} 是唯一的,并且使 $x - \hat{x}$ 达到最小值,即满足式(5.3.38)。

为形象说明投影定理,我们画出 Y 为二维线性子空间时的投影图,如图 5.3.3 所示。图中 \hat{x} 表示 x 在 Y 上的正交投影, z 表示 x 在 Y 上的任意非正交投影,则有 $\hat{x}, z, z - \hat{x} \in Y, \tilde{x} = x - \hat{x}$ 为正交投影误差,且

$$x - z = \tilde{x}$$

\tilde{x} 是各种投影误差的最小值,只有 $z = \hat{x}$, 上式取等号,说明正交投影是唯一的。

图 5.3.3 投影定理的证明

5.3.6 定理 5.3.5

定理 5.3.5 考虑定理 5.3.3 中 u 和 v 是相关的,其他条件不变,则有

$$E\{x @l, v\} = E\{x @l, v(u)\} \tag{5.3.44}$$

证明 式(5.3.44)中, $v(u)$ 是由 u 估计 v 的估计误差,有关公式为

$$v(u) = v - \hat{v}(u) \tag{5.3.45}$$

$$\hat{v}(u) = m_v + R_{vu}R_u^{-1}(u - m_u) \tag{5.3.46}$$

$$E\{v(u)\} = 0 \tag{5.3.47}$$

$$R_{v(u)} = R_v - R_{vu}R_u^{-1}R_{uv} \tag{5.3.48}$$

$v(u)$ 与 u 互相独立。式(5.3.44)表明,可把用两个相关量 u 和 v 估计 x 转换为用两个独立量 u 和 $v(u)$ 估计 x ,于是可用定理 5.3.3 进行计算。

我们分别计算式(5.3.44)两边,比较它们是否相等来证明定理。

设 $y = \begin{bmatrix} u \\ v \end{bmatrix}$, 则有

$$R_{xy} = [R_{xu} \ R_{xv}] \tag{5.3.49}$$

$$R_y = \begin{bmatrix} R_u & R_{uv} \\ R_{vu} & R_v \end{bmatrix} \tag{5.3.50}$$

$$R_y^{-1} = \begin{bmatrix} R_u^{-1} + R_u^{-1}R_{uv}R_{v(u)}^{-1}R_{vu}R_u^{-1} & -R_u^{-1}R_{uv}R_{v(u)}^{-1} \\ -R_{v(u)}^{-1}R_{vu}R_u^{-1} & R_{v(u)}^{-1} \end{bmatrix} \tag{5.3.51}$$

$$y - m_y = \begin{bmatrix} u - m_u \\ v - m_v \end{bmatrix} \tag{5.3.52}$$

首先,计算式(5.3.44)左边,利用上述公式,得到

$$\begin{aligned} E\{x @l, v\} &= m_x + R_{xy}R_y^{-1}(y - m_y) \\ &= m_x + [R_{xu} + R_{xu}R_u^{-1}R_{uv}R_{v(u)}^{-1}R_{vu} \\ &\quad - R_{xv}R_{v(u)}^{-1}R_{vu}]R_u^{-1}(u - m_u) \\ &\quad + [R_{xv} - R_{xu}R_u^{-1}R_{uv}]R_{v(u)}^{-1}(v - m_v) \end{aligned} \tag{5.3.53}$$

再计算式(5.3.44)右边, 因为 u 与 $v(u)$ 独立, 有

$$\begin{aligned} E\{x \hat{u}, v(u)\} &= E\{x \hat{u}\} + E\{x \hat{v}(u)\} - m_x \\ &= m_x + R_{xu} R_u^{-1} (u - m_u) + R_{xv(u)} R_{v(u)}^{-1} v(u) \end{aligned} \quad (5.3.54)$$

式中:

$$\begin{aligned} R_{xv(u)} &= E\{(x - m_x) v^T(u)\} \\ &= E\{(x - m_x) [v - m_v - R_{vu} R_u^{-1} (u - m_u)]^T\} \\ &= R_{xv} - R_{xu} R_u^{-1} R_{uv} \end{aligned} \quad (5.3.55)$$

把上式和式(5.3.46)代入式(5.3.54), 得到

$$\begin{aligned} E\{x \hat{u}, v(u)\} &= m_x + R_{xu} R_u^{-1} (u - m_u) - [R_{xu} R_u^{-1} R_{uv} - R_{xv}] R_{v(u)}^{-1} (v - m_v) \\ &\quad - [R_{xv} - R_{xu} R_u^{-1} R_{uv}] R_{v(u)}^{-1} R_{vu} R_u^{-1} (u - m_u) \end{aligned} \quad (5.3.56)$$

对式(5.3.56)和(5.3.53)进行比较, 两式相等, 因此, 式(5.3.44)成立。

将定理 5.3.5 与定理 5.3.3 比较, 不同的只是用 $v(u)$ 置换了 v , 对此可做如下解释: 把 v 分为两个分量, 即 $\hat{v}(u)$ 和 $v(u)$, 于是用 v 估计 x , 分为用 $\hat{v}(u)$ 和 $v(u)$ 两部分分别估计 x 。其中 $\hat{v}(u)$ 为最优估计, 它是 v 在 u 上的投影, 因此 $\hat{v}(u)$ 的信息已由 u 提供, 也就是说, 用 u 估计 x 包含了用 $\hat{v}(u)$ 估计 x 的全部信息。而 $v(u)$ 为估计误差, $v(u)$ 与 u 互相独立, 几何意义上两者互相垂直, 因此 $v(u)$ 给出了 u 所没有的信息, 是新息, 也就是说, 用 $v(u)$ 估计 x 这部分不能用 u 估计 x 代替。所以, 用 u 和 v 估计 x 可用 u 和 $v(u)$ 估计 x 代替。几何意义上 x 在 u 和 v 上的投影等于 x 在 u 和 $\hat{x}(u)$ 上投影。

例 5.3.1 一随机向量 x 具有高斯 $N(m_x, R_x)$ 分布, 用量测仪器可量测其线性组合 y

$$y = cx + e \quad (5.3.57)$$

式中: 量测误差 e 具有高斯 $N(0, R_e)$ 分布; c 为给定常阵。(1) e 与 x 互相独立; (2) e 与 x 相关, 互协方差函数为 R_{xe} 。试求两种情况下最优估计 $\hat{x}(y)$ 和估计误差 $\tilde{x}(y)$ 的协方差 $R_{\tilde{x}(y)}$ 。

解 我们利用公式

$$\hat{x}(y) = m_x + R_{xy} R_y^{-1} (y - m_y) \quad (5.3.58)$$

$$R_{\tilde{x}(y)} = R_x - R_{xy} R_y^{-1} R_{yx} \quad (5.3.59)$$

求解。先导出一般公式, 再分两种情况求算。对式(5.3.57)两边取均值, 得到

$$m_y = cm_x \quad (5.3.60)$$

把式(5.3.57)和(5.3.60)代入有关定义式, 得到

$$\begin{aligned} R_{xy} &= E\{(x - m_x)(y - m_y)\} \\ &= E\{(x - m_x)(cx - cm_y + e)^T\} \\ &= R_x c^T + R_{xe} \end{aligned} \quad (5.3.61)$$

$$\begin{aligned} R_{yx} &= R_{xy}^T \\ &= R_{xe}^T + c R_x \end{aligned} \quad (5.3.62)$$

$$\begin{aligned} R_y &= E\{(y - m_y)(y - m_y)^T\} \\ &= E\{(cx - cm_x + e)(cx - cm_x + e)^T\} \\ &= c R_x c^T + R_e + c R_{xe} + (c R_{xe})^T \end{aligned} \quad (5.3.63)$$

$$R_y^{-1} = [cR_x c^T + R_e + cR_{xe} + (cR_{xe})^T]^{-1} \quad (5.3.64)$$

(1) e 与 x 互相独立情况

把式(5.3.60) ~ (5.3.64) 代入式(5.3.58) 和(5.3.59), 只要注意到 $R_{xe} = 0$, 得到

$$\hat{x}(y) = m_x + R_x c^T (cR_x c^T + R_e)^{-1} (y - cm_x) \quad (5.3.65)$$

$$R_{\tilde{x}(y)} = R_x - R_x c^T (cR_x c^T + R_e)^{-1} cR_x \quad (5.3.66)$$

(2) e 与 x 相关情况

把式(5.3.60) ~ (5.3.64) 代入式(5.3.58) 和(5.3.59), 得到

$$\hat{x}(y) = m_x + (R_x c^T + R_{xe}) [cR_x c^T + R_e + cR_{xe} + (cR_{xe})^T]^{-1} (y - cm_x) \quad (5.3.67)$$

$$R_{\tilde{x}(y)} = R_x - (R_x c^T + R_{xe}) [cR_x c^T + R_e + cR_{xe} + (cR_{xe})^T]^{-1} (cR_x + R_{xe}^T) \quad (5.3.68)$$

例 5.3.2 考虑标量随机过程

$$x(t) = \int_{t_0}^t g(t, s) dv(s) \quad (5.3.69)$$

$$y(t) = \int_{t_0}^t h(t, s) dw(s) \quad (5.3.70)$$

且式(5.3.70) 有逆式

$$w(t) = \int_{t_0}^t k(t, s) dy(s) \quad (5.3.71)$$

式中: $dv(t)$ 和 $dw(t)$ 为增量维纳过程, 具有参数 $N(0, r dt)$ 。试确定利用 $y(s) (t_0 \leq s \leq t)$ 使均方估计误差为最小的 $x(t)$ 的估计 $\hat{x}(t|s)$, 再确定估计误差的方差 $R_{\tilde{x}(t|s)}$ 。

解 把均方估计误差作为性能指标, 使其满足本节定理规定的性能指标条件, 因此可利用公式

$$\hat{x}(t|s) = m_x + R_{xy} R_y^{-1} (y(s) - m_y) \quad (5.3.72)$$

$$R_{\tilde{x}(t|s)} = R_x - R_{xy} R_y^{-1} R_{yx} \quad (5.3.73)$$

求解。对式(5.3.69) 和(5.3.70) 两边取均值, 可得

$$m_x = 0 \quad (5.3.74)$$

$$m_y = 0 \quad (5.3.75)$$

利用式(5.3.69), (5.3.70), (5.3.74) 和(5.3.75), 可得

$$\begin{aligned} R_{xy} &= E\{x(t)y(s)\} \\ &= \int_{t_0}^t \int_{t_0}^s g(t, u) h(s, v) E\{dv(u)dw(v)\} \\ &= r \int_{t_0}^s g(t, u) h(s, u) du \end{aligned} \quad (5.3.76)$$

$$R_{yx} = R_{xy} \quad (5.3.77)$$

$$\begin{aligned} R_y &= E\{y^2(s)\} \\ &= \int_{t_0}^s \int_{t_0}^s h(s, u) h(s, v) E\{dw(u)dw(v)\} \end{aligned}$$

$$= r \int_{t_0}^s h^2(s, u) du \quad (5.3.78)$$

$$R_y^{-1} = r \int_{t_0}^s h^2(s, u) du^{-1} \quad (5.3.79)$$

$$\begin{aligned} R_x &= E\{x^2(t)\} \\ &= \int_{t_0}^t \int_{t_0}^t g(t, u) g(t, s) E\{dv(u) dv(s)\} \\ &= r \int_{t_0}^t g^2(t, u) du \end{aligned} \quad (5.3.80)$$

把式(5.3.74) ~ (5.3.80)代入式(5.3.72)和(5.3.73),得到最后结果

$$\hat{x}(t|s) = \int_{t_0}^s g(t, u) h(s, u) du \int_{t_0}^s h^2(s, u) du^{-1} y(s) \quad (5.3.81)$$

$$R_{\tilde{x}(t|s)} = r \int_{t_0}^t g^2(t, u) du - r \int_{t_0}^s g(t, u) h(s, u) du^2 \int_{t_0}^s h^2(s, u) du \quad (5.3.82)$$

5.4 离散时间随机系统的预测和滤波

在上节,我们介绍了状态估计的基本理论和基本计算方法,导出了在一定条件下,状态估计 $\hat{x}(y)$ 等于条件均值 $E\{x|y\}$,并在几何上解释为 x 在 y 上的正交投影,这种估计是无偏估计,估计误差为最小。后面两节中,我们用这些基本理论和基本计算方法具体讨论离散时间随机状态模型的状态估计问题,这一节中讨论预测和滤波问题,下一节中讨论平滑问题。

5.4.1 离散时间随机状态模型

考虑离散时间随机状态模型

$$x(k+1) = \Phi x(k) + \Gamma w(k) \quad (5.4.1)$$

$$y(k) = H x(k) + v(k) \quad (5.4.2)$$

式中: $x(k)$ 为 $n \times 1$ 随机状态向量; $y(k)$ 为 $p \times 1$ 量测向量; $w(k)$ 为 $n \times 1$ 模型噪声向量; $v(k)$ 为 $p \times 1$ 量测噪声向量; Φ 为 $n \times n$ 状态转移阵; Γ 为 $n \times n$ 量测阵。 Φ 和 Γ 可以是时变的或时不变的。

对式(5.4.1)和(5.4.2)的假设如下:

(1) 模型噪声 $w(k)$ 和量测噪声分别为独立同分布高斯 $N(0, R_1)$ 和 $N(0, R_2)$ 随机向量序列,即为离散时间高斯白噪声, R_1 和 R_2 可以是时变的或时不变的, $w(k)$ 和 $v(k)$ 互相独立,即

$$E\{w(k) w^T(l)\} = 0 \quad (k \neq l) \quad (5.4.3)$$

(2) 状态初值 $x(k_0)$ 为高斯 $N(m_0, R_0)$ 的,且与 $w(k)$ 和 $v(k)$ 互相独立,即

$$E\{[x(k_0) - m_0] w^T(k)\} = 0 \quad (k \geq k_0) \quad (5.4.4)$$

$$E\{[x(k_0) - m_0] v^T(k)\} = 0 \quad (k \geq k_0) \quad (5.4.5)$$

由于 $x(k_0)$, (k) 和 (k) 都是高斯的, 而 $x(k)$ 和 $y(k)$ 可表示为它们的线性组合, 因此也是高斯的。

(3) 容许估计值 $\hat{x}(k_1|k)$ 是指直到 k 时刻量测值 $y(k_0), y(k_1), \dots, y(k)$ 对 k_1 时刻 $x(k_1)$ 的估计, 表示为

$$\begin{aligned}\hat{x}(k_1|k) &= \hat{x}(k_1|y(k_0), y(k_1), \dots, y(k)) \\ &= \hat{x}(k_1|Y_k)\end{aligned}\tag{5.4.6}$$

式中: Y_k 表示量测向量集

$$Y_k = [y^T(k_0) \ y^T(k_1) \ \dots \ y^T(k)]^T\tag{5.4.7}$$

下面推导一步最优预测和最优滤波计算公式, 它们是最优估计的基本计算公式, 再在此基础上, 给出 k 步最优预测公式。

5.4.2 一步最优预测

定理 5.4.1(卡尔曼估计定理) 离散时间随机状态模型式(5.4.1)和(5.4.2)的一步最优预测 $\hat{x}(k+1|k)$ 满足递推方程

$$\hat{x}(k+1|k) = \hat{x}(k|k-1) + K(k|k-1)[y(k) - \hat{x}(k|k-1)]\tag{5.4.8}$$

式中: $K(k|k-1)$ 称为增益矩阵, 表示为

$$K(k|k-1) = P(k|k-1)^T [P(k|k-1)^T + R_2]^{-1}\tag{5.4.9}$$

$P(k+1|k)$ 称为一步预测误差的方差阵, 表示为

$$\begin{aligned}P(k+1|k) &= E\{\tilde{x}(k+1|k) \tilde{x}^T(k+1|k)\} \\ &= [I - K(k|k-1)]P(k|k-1)[I - K(k|k-1)]^T \\ &\quad + R_1 + K(k|k-1)R_2K^T(k|k-1) \\ &= P(k|k-1)^T + R_1 - K(k|k-1)P(k|k-1)^T \\ &= P(k|k-1)^T + R_1 - K(k|k-1)[P(k|k-1)^T + R_2]K^T(k|k-1) \\ &= [I - K(k|k-1)]P(k|k-1)^T + R_1\end{aligned}\tag{5.4.10}$$

初值为

$$\hat{x}(k_0|k_0-1) = m_0\tag{5.4.11}$$

$$P(k_0|k_0-1) = R_0\tag{5.4.12}$$

各式中符号意义和有关给定值由 5.4.1 小节给出。另外假设 R_2 为正定, 以保证式(5.4.9)的存在。

证明 证明分三步进行。

1. 确定 $\hat{x}(k+1|k)$ 的递推公式和 $K(k|k-1)$

由定理 5.3.5, 得到

$$\begin{aligned}\hat{x}(k+1|k) &= E\{x(k+1)|Y_k\} \\ &= E\{x(k+1)|Y_{k-1}, y(k)\} \\ &= E\{x(k+1)|Y_{k-1}, \tilde{y}(k|k-1)\}\end{aligned}\tag{5.4.13}$$

式中: $\tilde{y}(k|k-1)$ 为 $y(k-1), y(k-2), \dots, y(k_0)$ 对 $y(k)$ 的估计误差。由于 Y_{k-1} 和 $\tilde{y}(k|k-1)$ 互相独立, 因此由定理 5.3.3 和式(5.3.8), 并考虑到

$$E\{\tilde{x}(k+1|k)\} = 0 \quad (5.4.14)$$

得到

$$\begin{aligned} \hat{x}(k+1|k) &= E\{x(k+1)|Y_{k-1}\} + E\{x(k+1)|\tilde{y}(k|k-1)\} - m_{x(k+1)} \\ &= E\{x(k+1)|Y_{k-1}\} + R_{x(k+1)\tilde{y}(k|k-1)} R_{\tilde{y}(k|k-1)}^{-1} \tilde{y}(k|k-1) \end{aligned} \quad (5.4.15)$$

下面分别计算式(5.4.15)右边各项:

$$\begin{aligned} E\{x(k+1)|Y_{k-1}\} &= E\{x(k) + (k)|Y_{k-1}\} \\ &= \hat{x}(k|k-1) \end{aligned} \quad (5.4.16)$$

推导中利用了

$$E\{(k)|Y_{k-1}\} = E\{(k)\} = 0 \quad (5.4.17)$$

$$\begin{aligned} \tilde{y}(k|k-1) &= y(k) - \hat{y}(k|k-1) \\ &= y(k) - E\{x(k) + (k)|Y_{k-1}\} \\ &= y(k) - \hat{x}(k|k-1) \\ &= x(k) + (k) - \hat{x}(k|k-1) \\ &= \tilde{x}(k|k-1) + (k) \end{aligned} \quad (5.4.18)$$

上式右边由第二式到第三式利用了

$$E\{(k)|Y_{k-1}\} = E\{(k)\} = 0 \quad (5.4.19)$$

$$\begin{aligned} R_{\tilde{y}(k|k-1)} &= E\{\tilde{y}(k|k-1) \tilde{y}^T(k|k-1)\} \\ &= E\{[\tilde{x}(k|k-1) + (k)][\tilde{x}^T(k|k-1) + (k)^T]\} \\ &= p(k|k-1)^T + R_2 \end{aligned} \quad (5.4.20)$$

上式推导中利用了

$$E\{\tilde{x}(k|k-1) (k)^T\} = E\{(k) \tilde{x}^T(k|k-1)\} = 0 \quad (5.4.21)$$

这是因为 $\tilde{x}(k|k-1)$ 是 Y_{k-1} 和 $y(k)$ 的函数,而 Y_{k-1} 和 $y(k)$ 与 (k) 互相独立。

$$\begin{aligned} R_{x(k+1)\tilde{y}(k|k-1)} &= E\{[x(k+1) - m_{x(k+1)}] \tilde{y}^T(k|k-1)\} \\ &= E\{[x(k) + (k) - m_{x(k)}][\tilde{x}(k|k-1) + (k)]^T\} \\ &= E\{x(k) \tilde{x}^T(k|k-1)\}^T \end{aligned} \quad (5.4.22)$$

其余五项取值为零。注意到 $x(k)$ 可由互相独立的两项组成,即

$$x(k) = \hat{x}(k|k-1) + \tilde{x}(k|k-1) \quad (5.4.23)$$

所以得到

$$R_{x(k+1)\tilde{y}(k|k-1)} = P(k|k-1)^T \quad (5.4.24)$$

令

$$K(k|k-1) = R_{x(k+1)\tilde{y}(k|k-1)} R_{\tilde{y}(k|k-1)}^{-1} \quad (5.4.25)$$

把式(5.4.24)和(5.4.20)代入式(5.4.25),就得到式(5.4.9)。把式(5.4.16),(5.4.25)和(5.4.18)代入式(5.4.15),就得到式(5.4.8)。

2. 确定 $P(k+1|k)$

用 $P(k+1|k)$ 的定义式计算,先求 $\tilde{x}(k+1|k)$:

$$\tilde{x}(k+1|k) = x(k+1) - \hat{x}(k+1|k)$$

$$\begin{aligned}
&= \hat{x}(k) + (k) - \{ \hat{x}(k|k-1) + K(k|k-1)[\tilde{x}(k|k-1) + (k)] \} \\
&= [I - K(k|k-1)]\tilde{x}(k|k-1) + (k) - K(k|k-1)(k) \quad (5.4.26)
\end{aligned}$$

上式右边三项是互相独立的, 因此有

$$\begin{aligned}
P(k+1|k) &= E\{\tilde{x}(k+1|k)\tilde{x}^T(k+1|k)\} \\
&= [I - K(k|k-1)]P(k|k-1)[I - K(k|k-1)]^T + R_1 \\
&\quad + K(k|k-1)R_2K^T(k|k-1) \quad (5.4.27)
\end{aligned}$$

上式就是求证的式(5.4.10)第二式。只要把式(5.4.27)适当展开, 容易得到另外的三个表达式。

3. 确定初值 $\hat{x}(k_0|k_0-1)$ 和 $P(k_0|k_0-1)$

这里, 要证明像式(5.4.11)和(5.4.12)那样设置初值是正确的。证明方法是用式(5.4.8)求得 $\hat{x}(k_0+1|k_0)$, 与用基本估计计算公式求得结果进行比较来加以证明。

首先由式(5.4.8)和(5.4.9), 得到

$$\begin{aligned}
\hat{x}(k_0+1|k_0) &= \hat{x}(k_0|k_0-1) + P(k_0|k_0-1)^T [P(k_0|k_0-1)^T + R_2]^{-1} \\
&\quad \times [y(k_0) - \hat{x}(k_0|k_0-1)] \quad (5.4.28)
\end{aligned}$$

下面再用基本估计计算公式求解 $\hat{x}(k_0+1|k_0)$ 。基本估计计算公式为

$$\hat{x}(k_0+1|k_0) = m_{x(k_0+1)} + R_{x(k_0+1)y(k_0)} R_{y(k_0)}^{-1} [y(k_0) - m_{y(k_0)}] \quad (5.4.29)$$

分别计算上式右边各项:

$$m_{x(k_0+1)} = m_0 \quad (5.4.30)$$

$$x(k_0+1) - m_{x(k_0+1)} = [x(k_0) - m_0] + (k_0) \quad (5.4.31)$$

$$\begin{aligned}
y(k_0) - m_{y(k_0)} &= y(k_0) - m_0 \\
&= [x(k_0) - m_0] + (k_0) \quad (5.4.32)
\end{aligned}$$

$$\begin{aligned}
R_{y(k_0)} &= R_0^T + R_2 \\
R_{x(k_0+1)y(k_0)} &= E\{[x(k_0+1) - m_{x(k_0+1)}][y(k_0) - m_{y(k_0)}]^T\} \\
&= R_0^T \quad (5.4.33)
\end{aligned}$$

把式(5.4.30) ~ (5.4.33)代入式(5.4.29), 得到

$$\hat{x}(k_0+1|k_0) = m_0 + R_0^T [R_0^T + R_2]^{-1} [y(k_0) - m_0] \quad (5.4.34)$$

把式(5.4.28)与上式进行对照, 容易确定式(5.4.11)和(5.4.12)是正确的, 初值选择是合适的。

一步最优预测是一种递推算法, 计算步骤如下:

(1) 利用给定初值 R_0 、式(5.4.9)和(5.4.10)递推求出任意时刻的 $K(k|k-1)$ 和 $P(k+1|k)$ 。由于上述计算与量测数据无关, 可预先计算, 并存入计算机备用, 节省在线计算时间。

(2) 再通过给定初值 m_0 、已得到的增益矩阵 $K(k_0|k_0-1)$, $K(k_0+1|k_0)$, ... 和量测数据 $y(k_0)$, $y(k_0+1)$, ..., 利用式(5.4.8)可得到任意时刻的一步最优预测。一步最优预测计算流程如图 5.4.1 所示。

一步最优预测框图如图 5.4.2 所示。

图 5.4.1 一步最优预测计算流程图

图 5.4.2 一步最优预测框图

5.4.3 最优滤波

定理 5.4.2(卡尔曼滤波定理) 离散时间随机状态模型式(5.4.1)和(5.4.2)的最优滤波 $\hat{x}(k+1|k+1)$ 满足递推方程

$$\hat{x}(k+1|k+1) = \hat{x}(k|k) + K(k+1|k+1)[y(k+1) - \hat{x}(k|k)] \tag{5.4.35}$$

式中: $K(k+1|k+1)$ 称为滤波增益阵, 表示为

$$K(k+1|k+1) = P(k+1|k)^T [P(k+1|k)^T + R_2]^{-1} \tag{5.4.36}$$

式中: 一步预测误差的方差阵 $P(k+1|k)$ 表示为

$$P(k+1|k) = P(k|k)^T + R_1 \tag{5.4.37}$$

$P(k+1|k+1)$ 为滤波误差的方差阵, 表示为

$$\begin{aligned} P(k+1|k+1) &= [I - K(k+1|k+1)]P(k+1|k)[I - K(k+1|k+1)]^T \\ &\quad + K(k+1|k+1)R_2K^T(k+1|k+1) \\ &= [I - K(k+1|k+1)]P(k+1|k) \end{aligned} \tag{5.4.38}$$

初值为

$$\hat{x}(k_0|k_0) = m_0 \tag{5.4.39}$$

$$P(k_0|k_0) = R_0 \tag{5.4.40}$$

各式中符号意义和有关给定值由 5.4.1 小节给出, 另外假设 R_2 为正定, 以保证式(5.4.36)的存在。

证明 分四步进行证明。

(1) 确定 $\hat{x}(k+1|k+1)$ 的递推公式和 $K(k+1|k+1)$

由定理 5.3.5, 得到

$$\begin{aligned}\hat{x}(k+1|k+1) &= E\{x(k+1)|Y_{k+1}\} \\ &= E\{x(k+1)|Y_k, \tilde{y}(k+1|k)\} \\ &= E\{x(k+1)|Y_k\} + E\{x(k+1)|\tilde{y}(k+1|k)\} - m_{x(k+1)} \\ &= \hat{x}(k|k) + R_{x(k+1)\tilde{y}(k+1|k)} R_{\tilde{y}(k+1|k)}^{-1} \tilde{y}(k+1|k)\end{aligned}\quad (5.4.41)$$

下面分别计算上式右边各项:

$$\begin{aligned}\tilde{y}(k+1|k) &= y(k+1) - \hat{y}(k+1|k) \\ &= y(k+1) - \hat{x}(k|k) \\ &= \tilde{x}(k+1|k) + (k+1)\end{aligned}\quad (5.4.42)$$

$$E\{\tilde{y}(k+1|k)\} = 0 \quad (5.4.43)$$

$$\begin{aligned}R_{\tilde{y}(k+1|k)} &= E\{[\tilde{x}(k+1|k) + (k+1)][\tilde{x}(k+1|k) + (k+1)]^T\} \\ &= P(k+1|k)^T + R_2\end{aligned}\quad (5.4.44)$$

$$\begin{aligned}R_{x(k+1)\tilde{y}(k+1|k)} &= E\{[x(k+1) - m_{x(k+1)}][\tilde{x}(k+1|k) + (k+1)]^T\} \\ &= E\{x(k+1)\tilde{x}^T(k+1|k)\}^T \\ &= P(k+1|k)^T\end{aligned}\quad (5.4.45)$$

令

$$K(k+1|k+1) = R_{x(k+1)\tilde{y}(k+1|k)} R_{\tilde{y}(k+1|k)}^{-1} \quad (5.4.46)$$

把式(5.4.44)和(5.4.45)代入式(5.4.46)就得到式(5.4.36)。把式(5.4.46)和(5.4.42)代入式(5.4.41)就得到式(5.4.35)。

(2) 确定 $P(k+1|k)$

把

$$\begin{aligned}\tilde{x}(k+1|k) &= x(k+1) - \hat{x}(k+1|k) \\ &= \tilde{x}(k|k) + (k)\end{aligned}\quad (5.4.47)$$

代入 $P(k+1|k)$ 的定义式, 就得到式(5.4.37)。

(3) 确定 $P(k+1|k+1)$

把式(5.4.35)代入 $\tilde{x}(k+1|k+1)$ 表示式, 得到

$$\begin{aligned}\tilde{x}(k+1|k+1) &= x(k+1) - \hat{x}(k+1|k+1) \\ &= [I - K(k+1|k+1)]\tilde{x}(k+1|k) - K(k+1|k+1)(k+1)\end{aligned}\quad (5.4.48)$$

把上式代入 $P(k+1|k+1)$ 的定义式, 得到结果为

$$\begin{aligned}P(k+1|k+1) &= E\{\tilde{x}(k+1|k+1)\tilde{x}^T(k+1|k+1)\} \\ &= [I - K(k+1|k+1)]P(k+1|k)[I - K(k+1|k+1)]^T \\ &\quad + K(k+1|k+1)R_2K(k+1|k+1)\end{aligned}\quad (5.4.49)$$

上式就是式(5.4.38)第一式, 把它适当展开并化简, 容易得到式(5.4.38)的第二式。

(4) 初值

很自然,把式(5.4.39)和(5.4.40)作为初值是合理的。

最优滤波是一种递推算法,计算步骤如下:

(1) 给定初值 R_0 ,通过式(5.4.37),(5.4.36)和(5.4.38)递推计算所需时刻的 $P(k+1|k)$, $K(k+1|k+1)$ 和 $P(k+1|k+1)$ 。由于上述计算与量测数据无关,可预先计算,并存入计算机备用。

(2) 给定初值 m_0 ,量测数据 $y(k_0+1), y(k_0+2), \dots$, 利用式(5.4.35)递推计算得到状态的最优滤波 $x(k_0+1|k_0+1), x(k_0+2|k_0+2), \dots$ 。最优滤波计算流程如图 5.4.3 所示。

图 5.4.3 最优滤波计算流程图

最优滤波框图如图 5.4.4 所示。

图 5.4.4 最优滤波框图

需注意的是,如果把式(5.4.36),(5.4.37)和(5.4.38)看做为第一套计算公式,则可把

$$K(k|k) = P(k|k)^T R_2^{-1} \tag{5.4.50}$$

$$P^{-1}(k|k) = P^{-1}(k|k-1) + ^TR_2^{-1} \tag{5.4.51}$$

和式(5.4.37)看做为第二套计算公式。

证明 把式(5.4.36)变换为

$$\begin{aligned} K(k|k) &= P(k|k-1)^T [P(k|k-1)^T + R_2]^{-1} \\ &= [(P(k|k-1)^T + R_2)(P(k|k-1)^T)^{-1}]^{-1} \\ &= [^TR_2^{-1} + P^{-1}(k|k-1)]^{-1} ^TR_2^{-1} \end{aligned} \tag{5.4.52}$$

对上式进一步计算,得到

$$\begin{aligned}
 I - K(k) &= [{}^T R_2^{-1} + P^{-1}(k-1)]^{-1} [{}^T R_2^{-1} + P^{-1}(k-1)] \\
 &\quad - [{}^T R_2^{-1} + P^{-1}(k-1)]^{-1} {}^T R_2^{-1} \\
 &= [P^{-1}(k-1) + {}^T R_2^{-1}]^{-1} P^{-1}(k-1)
 \end{aligned} \quad (5.4.53)$$

$$\begin{aligned}
 P(k) &= [I - K(k)]P(k-1)[I - K(k)]^T + K(k)R_2K^T(k) \\
 &= [P^{-1}(k-1) + {}^T R_2^{-1}]^{-1}
 \end{aligned} \quad (5.4.54)$$

上式即为式(5.4.51)。把式(5.4.54)代入式(5.4.52),得到式(5.4.50)。

5.4.4 p 步最优预测

我们把一步最优预测和最优滤波作为基础,导出 p 步最优预测和 p 步最优预测误差的方差。容易导出

$$\hat{x}(k+p) = \hat{x}(k+p-1) \quad (5.4.55)$$

$$P(k+p) = P(k+p-1)^T + R_1 \quad (5.4.56)$$

用以上两式分别递推得到

$$\begin{aligned}
 \hat{x}(k+p) &= {}^{p-1}\hat{x}(k+1) \\
 &= {}^p\hat{x}(k)
 \end{aligned} \quad (5.4.57)$$

$$\begin{aligned}
 P(k+p) &= {}^{p-1}P(k+1)({}^T)^{p-1} + \sum_{n=0}^{p-2} {}^nR_1({}^T)^n \\
 &= {}^pP(k)({}^T)^p + \sum_{n=0}^{p-1} {}^nR_1({}^T)^n
 \end{aligned} \quad (5.4.58)$$

例 5.4.1 已知状态方程

$$x(k+1) = x(k) + u(k) + (k) \quad (5.4.59)$$

$$y(k) = x(k) \quad (5.4.60)$$

式中: $x(k_0) \sim N(m_0, R_0)$ 与噪声 $(k) \sim N(m, R)(k \geq k_0)$ 互相独立; $u(k)$ 为所加控制量; 与 为已知相应常阵。试用卡尔曼估计方法求出 $\hat{x}(k+1)$, $\tilde{x}(k+1)$ 和 $P(k+1)$ 。

解 用基本一步预测公式求解:

$$\hat{x}(k+1) = m_{x(k+1)} + R_{x(k+1)y(k)}R_{y(k)}^{-1}[y(k) - m_{y(k)}] \quad (5.4.61)$$

式中:

$$\begin{aligned}
 m_{x(k+1)} &= E\{x(k) + u(k) + (k)\} \\
 &= m_{x(k)} + u(k) + m
 \end{aligned} \quad (5.4.62)$$

$$\begin{aligned}
 R_{y(k)} &= E\{[y(k) - m_{y(k)}][y(k) - m_{y(k)}]\} \\
 &= E\{[x(k) - m_{x(k)}][x(k) - m_{x(k)}]\} \\
 &= R[x(k)]
 \end{aligned} \quad (5.4.63)$$

$$\begin{aligned}
 R_{x(k+1)y(k)} &= E\{[x(k+1) - m_{x(k+1)}][y(k) - m_{y(k)}]\} \\
 &= E\{[(x(k) - m_{x(k)}) + (k)][x(k) - m_{x(k)}]\} \\
 &= R[x(k)]
 \end{aligned} \quad (5.4.64)$$

把式(5.4.62) ~ (5.4.64)代入式(5.4.61),得到

$$\hat{x}(k+1|k) = y(k) + u(k) + m \quad (5.4.65)$$

$$\begin{aligned} \tilde{x}(k+1|k) &= x(k+1) - \hat{x}(k+1|k) \\ &= (k) - m \end{aligned} \quad (5.4.66)$$

$$\begin{aligned} P(k+1|k) &= E\{\tilde{x}(k+1|k) \tilde{x}(k+1|k)\} \\ &= R \end{aligned} \quad (5.4.67)$$

例 5.4.2 已知给定系统

$$x(k+1) = 0.6x(k) + 2u(k) + (k) \quad (5.4.68)$$

$$y(k) = 3x(k) + (k) \quad (5.4.69)$$

并且 $x(k_0) \sim N(5, 10)$, $(k) \sim N(0, 3)$, $(k) \sim N(1, 2)$ 互相独立, $u(k)$ 为控制量, 已测得 $y(k_0) = 14$ 。试求 $\hat{x}(k_0+1|k_0)$ 的值。

解 给定系统式(5.4.68)和(5.4.69)与系统式(5.4.1)和(5.4.2)的区别在于加入了控制项 $u(k)$ 且量测噪声的均值不为零, 经类似推导可得一步预测公式

$$\hat{x}(k+1|k) = \hat{x}(k|k-1) + u(k) + K(k|k-1)[y(k) - \hat{x}(k|k-1) - m_{(k)}] \quad (5.4.70)$$

考虑初值 $\hat{x}(k_0|k_0-1) = m_0$, $R(k_0|k_0-1) = R_0$, 所求值为

$$\hat{x}(k_0+1|k_0) = m_0 + u(k_0) + K(k_0|k_0-1)[y(k_0) - m_0 - m] \quad (5.4.71)$$

式中 $K(k_0|k_0-1)$ 公式不变, 为

$$K(k_0|k_0-1) = R_0^{-1} [R_0^{-1} + R]^{-1} \quad (5.4.72)$$

代入数据 $= 0.6$, $= 2$, $= 3$, $m_0 = 5$, $R_0 = 10$, $m = 1$, $R = 2$, $y(k_0) = 14$, 得到

$$K(k_0|k_0-1) = \frac{18}{92} = 0.2 \quad (5.4.73)$$

$$\begin{aligned} \hat{x}(k_0+1|k_0) &= 0.6 \times 5 + 2u(k_0) + 0.2[14 - 3 \times 5 - 1] \\ &= 2.6 + 2u(k_0) \end{aligned} \quad (5.4.74)$$

5.5 离散时间随机系统的最优平滑

5.5.1 最优平滑分类

已知当 $k < l$ 时, 最优估计 $\hat{x}(k|l)$ 称为最优平滑, 而按 k 和 l 的关系, 最优平滑可分为三类:

(1) 固定区间平滑

设 N 为正整数, $k = 0, 1, \dots, N-1$, 则 $\hat{x}(k|N)$ 称为固定区间(最优)平滑(fixed-interval smoothing)。

(2) 固定点平滑

令 k 为正整数, $j = k+1, k+2, \dots$, 则 $\hat{x}(k|j)$ 称为固定点(最优)平滑(fixed-point smoothing)或单点平滑(single-point smoothing)。

(3) 固定滞后平滑

令 N 为正整数, $k = 0, 1, 2, \dots$, 则 $\hat{x}(k|k+N)$ 称为固定滞后平滑(fixed-lag

smoothing)。

下面先介绍一步最优平滑和二步最优平滑,然后在此基础上分析各类平滑。各类平滑都是在已知一步最优预测和最优滤波条件下进行计算的。

5.5.2 一步最优平滑

考虑离散时间随机状态模型

$$x(k+1) = x(k) + w(k) \tag{5.5.1}$$

$$y(k) = x(k) + v(k) \tag{5.5.2}$$

式中符号意义和有关假设如以前所述,则有如下定理:

定理 5.5.1 考虑离散时间随机状态模型式(5.5.1)和(5.5.2),一步最优平滑为

$$\begin{aligned} \hat{x}(k+1) &= \hat{x}(k) + A(k)[\hat{x}(k+1) - \hat{x}(k)] \\ &= \hat{x}(k) + A(k)K(k+1)\tilde{y}(k+1) \\ &= \hat{x}(k) + M(k)[y(k+1) - \hat{x}(k)] \end{aligned} \tag{5.5.3}$$

式中:

$$A(k) = P(k)^T P^{-1}(k+1) \tag{5.5.4}$$

$$K(k+1) = P(k+1)^T [P(k+1)^T + R_2]^{-1} \tag{5.5.5}$$

$$M(k+1) = A(k)K(k+1) \tag{5.5.6}$$

一步最优平滑误差 $\tilde{x}(k+1)$ 的方差阵为

$$P(k+1) = P(k) + A(k)[P(k+1) - P(k)]A^T(k) \tag{5.5.7}$$

证明 直接写出一部最优平滑公式:

$$\begin{aligned} \hat{x}(k+1) &= E\{x(k) | Y_{k+1}\} \\ &= E\{x(k) | Y_k, y(k+1)\} \\ &= E\{x(k) | Y_k, \tilde{y}(k+1)\} \\ &= E\{x(k) | Y_k\} + E\{x(k) | \tilde{y}(k+1)\} - m_{x(k)} \\ &= \hat{x}(k) + R_{x(k)\tilde{y}(k+1)} R_{\tilde{y}(k+1)}^{-1} \tilde{y}(k+1) \end{aligned} \tag{5.5.8}$$

下面分别计算各项:

$$\begin{aligned} \tilde{y}(k+1) &= y(k+1) - \hat{y}(k+1) \\ &= y(k+1) - \hat{x}(k) \\ &= \tilde{x}(k+1) + w(k+1) \\ &= \tilde{x}(k) + w(k) + w(k+1) \end{aligned} \tag{5.5.9}$$

$$E\{\tilde{y}(k+1)\} = 0 \tag{5.5.10}$$

$$\begin{aligned} R_{\tilde{y}(k+1)} &= E\{\tilde{y}(k+1) \tilde{y}^T(k+1)\} \\ &= E\{[\tilde{x}(k+1) + w(k+1)][\tilde{x}(k+1) + w(k+1)]^T\} \\ &= P(k+1)^T + R_2 \end{aligned} \tag{5.5.11}$$

$$R_{x(k)\tilde{y}(k+1)} = E\{[x(k) - m_{x(k)}] \tilde{y}^T(k+1)\}$$

$$\begin{aligned}
&= E\{[x(k) - m_{x(k)}][\tilde{x}(k|k) + (k) + (k+1)^T]\} \\
&= E\{x(k) \tilde{x}^T(k|k)\}^T \\
&= P(k|k)^T
\end{aligned} \tag{5.5.12}$$

由式(5.5.11)和(5.5.12), 令

$$\begin{aligned}
M(k|k+1) &= R_{x(k)} \tilde{y}(k+1|k) R_{\tilde{y}(k+1|k)}^{-1} \\
&= P(k|k)^T [P(k+1|k)^T + R_2]^{-1} \\
&= P(k|k)^T P^{-1}(k+1|k) P(k+1|k)^T [P(k+1|k)^T + R_2]^{-1}
\end{aligned} \tag{5.5.13}$$

上式右边恰好可分为式(5.5.4)和(5.5.5)两部分, 即得到

$$M(k|k+1) = A(k)K(k+1|k+1) \tag{5.5.14}$$

式中: $A(k)$ 称为固定区间最优平滑增益阵; $K(k+1|k+1)$ 为卡尔曼滤波增益阵。把式(5.5.14)和(5.5.9)代入式(5.5.8), 得到

$$\begin{aligned}
\hat{x}(k|k+1) &= \hat{x}(k|k) + M(k|k+1)[y(k+1) - \hat{x}(k|k)] \\
&= \hat{x}(k|k) + A(k)K(k+1|k+1) \tilde{y}(k+1|k)
\end{aligned} \tag{5.5.15}$$

通过利用卡尔曼滤波公式

$$\hat{x}(k+1|k+1) = \hat{x}(k|k) + K(k+1|k+1) \tilde{y}(k+1|k) \tag{5.5.16}$$

得到

$$\begin{aligned}
K(k+1|k+1) \tilde{y}(k+1|k) &= \hat{x}(k+1|k+1) - \hat{x}(k|k) \\
&= \hat{x}(k+1|k+1) - \hat{x}(k+1|k)
\end{aligned} \tag{5.5.17}$$

把上式代入式(5.5.15), 得到

$$\hat{x}(k|k+1) = \hat{x}(k|k) + A(k)[\hat{x}(k+1|k+1) - \hat{x}(k+1|k)] \tag{5.5.18}$$

下面求一步最优平滑误差 $\tilde{x}(k|k+1)$ 的方差阵 $P(k|k+1)$ 。把式(5.5.18)改写为

$$\hat{x}(k|k+1) - A(k)\hat{x}(k+1|k+1) = \hat{x}(k|k) - A(k)\hat{x}(k+1|k) \tag{5.5.19}$$

用 $x(k) = \tilde{x}(k) + \hat{x}(k)$ 减上式两边, 得到

$$\tilde{x}(k|k+1) + A(k)\hat{x}(k+1|k+1) = \tilde{x}(k|k) + A(k)\hat{x}(k+1|k) \tag{5.5.20}$$

上式左边两项是互相独立的, 因为第一项与 \mathbf{Y}_{k+1} 在几何意义上是垂直的, 而第二项是在 \mathbf{Y}_{k+1} 上的投影。同样, 上式右边两项也是互相独立的。注意到 $E\{\tilde{x}(k|k+1)\} = 0$, 把上式两边自乘, 并取均值, 得到

$$P(k|k+1) + A(k)P_{\Lambda}(k+1|k+1)A^T(k) = P(k|k) + A(k)P_{\Lambda}(k+1|k)A^T(k) \tag{5.5.21}$$

$$P(k|k+1) = P(k|k) + A(k)[P_{\Lambda}(k+1|k) - P_{\Lambda}(k+1|k+1)]A^T(k) \tag{5.5.22}$$

式中:

$$P_{\Lambda}(k+1|k) = E\{\hat{x}(k+1|k) \hat{x}^T(k+1|k)\} \tag{5.5.23}$$

$$P_{\Lambda}(k+1|k+1) = E\{\hat{x}(k+1|k+1) \hat{x}^T(k+1|k+1)\} \tag{5.5.24}$$

另有

$$\tilde{x}(k+1|k) + \hat{x}(k+1|k) = \tilde{x}(k+1|k+1) + \hat{x}(k+1|k+1) \tag{5.5.25}$$

对上式处理同前, 得到

$$P_{\Lambda}(k+1|k) - P_{\Lambda}(k+1|k+1) = P(k+1|k) - P(k+1|k+1) \tag{5.5.26}$$

把上式代入式(5.5.22), 得到最后结果为

$$P(k\mathbb{K} + 1) = P(k\mathbb{K}) + A(k)[P(k + 1\mathbb{K} + 1) - P(k + 1\mathbb{K})]A^T(k) \tag{5.5.27}$$

5.5.3 二步最优平滑

定理 5.5.2 考虑离散时间随机状态模型式(5.5.1)和(5.5.2), 二步最优平滑为

$$\hat{x}(k\mathbb{K} + 2) = \hat{x}(k\mathbb{K}) + A(k)[\hat{x}(k + 1\mathbb{K} + 2) - \hat{x}(k + 1\mathbb{K})] \tag{5.5.28}$$

式中: $A(k)$ 称为固定区间平滑增益阵, 表示为

$$A(k) = P(k\mathbb{K})^{-1}P^{-1}(k + 1\mathbb{K}) \tag{5.5.29}$$

二步最优平滑误差 $\tilde{x}(k\mathbb{K} + 2)$ 的方差阵为

$$P(k\mathbb{K} + 2) = P(k\mathbb{K}) + A(k)[P(k + 1\mathbb{K} + 2) - P(k + 1\mathbb{K})]A^T(k) \tag{5.5.30}$$

证明 二步最优平滑 $\hat{x}(k\mathbb{K} + 2)$ 表示为

$$\begin{aligned} \hat{x}(k\mathbb{K} + 2) &= E\{x(k) \odot Y_{k+2}\} \\ &= E\{x(k) \odot Y_{k+1}, \tilde{y}(k + 2\mathbb{K} + 1)\} \\ &= E\{x(k) \odot Y_{k+1}\} + E\{x(k) \odot \tilde{y}(k + 2\mathbb{K} + 1)\} - m_{x(k)} \\ &= \hat{x}(k\mathbb{K} + 1) + R_{x(k)\tilde{y}(k+2\mathbb{K}+1)}R_{\tilde{y}(k+2\mathbb{K}+1)}^{-1}\tilde{y}(k + 2\mathbb{K} + 1) \end{aligned} \tag{5.5.31}$$

下面分别计算各项:

$$\begin{aligned} \tilde{y}(k + 2\mathbb{K} + 1) &= y(k + 2) - \hat{y}(k + 2\mathbb{K} + 1) \\ &= y(k + 2) - \hat{x}(k + 1\mathbb{K} + 1) \\ &= \tilde{x}(k + 2\mathbb{K} + 1) + (k + 2) \\ &= \tilde{x}(k + 1\mathbb{K} + 1) + (k + 1) + (k + 2) \\ &= [I - K(k + 1\mathbb{K} + 1)]\tilde{x}(k\mathbb{K}) \\ &\quad + [I - K(k + 1\mathbb{K} + 1)](k) + (k + 1) \\ &\quad - K(k + 1\mathbb{K} + 1)(k + 1) + (k + 2) \end{aligned} \tag{5.5.32}$$

$$E\{\tilde{y}(k + 2\mathbb{K} + 1)\} = 0 \tag{5.5.33}$$

$$R_{\tilde{y}(k+2\mathbb{K}+1)} = P(k + 2\mathbb{K} + 1)^T + R_2 \tag{5.5.34}$$

$$\begin{aligned} R_{x(k)\tilde{y}(k+2\mathbb{K}+1)} &= E\{[x(k) - m_{x(k)}]\tilde{y}^T(k + 2\mathbb{K} + 1)\} \\ &= P(k\mathbb{K})^{-1}[I - K(k + 1\mathbb{K} + 1)]^T P^{-1} \end{aligned} \tag{5.5.35}$$

令

$$M(k\mathbb{K} + 2) = R_{x(k)\tilde{y}(k+2\mathbb{K}+1)}R_{\tilde{y}(k+2\mathbb{K}+1)}^{-1} \tag{5.5.36}$$

把式(5.5.34)和(5.5.35)代入上式, 得到

$$M(k\mathbb{K} + 2) = P(k\mathbb{K})^{-1}[I - K(k + 1\mathbb{K} + 1)]^T P^{-1} [P(k + 2\mathbb{K} + 1)^T + R_2]^{-1} \tag{5.5.37}$$

利用卡尔曼滤波公式(5.4.38)

$$P(k + 1\mathbb{K} + 1) = [I - K(k + 1\mathbb{K} + 1)]P(k + 1\mathbb{K}) \tag{5.5.38}$$

即 $[I - K(k + 1\mathbb{K} + 1)]^T = P^{-1}(k + 1\mathbb{K})P(k + 1\mathbb{K} + 1) \tag{5.5.39}$

和式(5.4.36)

$$K(k+1)K(k+1) = P(k+1)K(k)^T [P(k+1)K(k)^T + R_2]^{-1} \quad (5.5.40)$$

即 $K(k+2)K(k+1) = [P(k+2)K(k+1)^T + R_2]^{-1} = P^{-1}(k+2)K(k+2)K(k+2)K(k+2)$

(5.5.41)

把式(5.5.39)和(5.5.41)代入式(5.5.37), 得到

$$M(k)K(k+2) = A(k)A(k+1)K(k+2)K(k+2) \quad (5.5.42)$$

式中 $A(k)$ 和 $A(k+1)$ 由式(5.5.29) 给出。把式(5.5.32), (5.5.36)和(5.5.42)等式代入式(5.5.31), 得到

$$\begin{aligned} \hat{x}(k)K(k+2) &= \hat{x}(k)K(k+1) + M(k)K(k+2)[y(k+2) - \hat{x}(k+1)K(k+1)] \\ &= \hat{x}(k)K(k+1) + A(k)A(k+1)K(k+2)K(k+2)\tilde{y}(k+2)K(k+2) \end{aligned} \quad (5.5.43)$$

由最优滤波公式(5.4.36), 用 $k+1$ 置换 k , 得到

$$\hat{x}(k+2)K(k+2) = \hat{x}(k+2)K(k+1) + K(k+2)K(k+2)\tilde{y}(k+2)K(k+2) \quad (5.5.44)$$

即 $K(k+2)K(k+2)\tilde{y}(k+2)K(k+2) = \hat{x}(k+2)K(k+2) - \hat{x}(k+2)K(k+1)$

(5.5.45)

把上式代入式(5.5.43), 得到

$$\hat{x}(k)K(k+2) = \hat{x}(k)K(k+1) + A(k)A(k+1)[\hat{x}(k+2)K(k+2) - \hat{x}(k+2)K(k+1)] \quad (5.5.46)$$

再把如下两个一步最优平滑公式

$$\hat{x}(k)K(k+1) = \hat{x}(k)K(k) + A(k)[\hat{x}(k+1)K(k+1) - \hat{x}(k+1)K(k)] \quad (5.5.47)$$

$$\begin{aligned} &A(k+1)[\hat{x}(k+2)K(k+2) - \hat{x}(k+2)K(k+1)] \\ &= \hat{x}(k+1)K(k+2) - \hat{x}(k+1)K(k+1) \end{aligned} \quad (5.5.48)$$

代入式(5.5.46), 得到最后结果为

$$\hat{x}(k)K(k+2) = \hat{x}(k)K(k) + A(k)[\hat{x}(k+1)K(k+2) - \hat{x}(k+1)K(k)] \quad (5.5.49)$$

二步最优平滑误差 $\tilde{x}(k)K(k+2)$ 的方差阵 $P(k)K(k+2)$ 的证明与一步最优平滑类似, 不再赘述。

比较一步最优平滑和二步最优平滑表示式(5.5.3)和(5.5.28)不难发现, 两式形式相同, 只有右边方括弧内第一项不同, 一个是 $\hat{x}(k+1)K(k+1)$, 另一个是 $\hat{x}(k+1)K(k+2)$ 。它启发我们, 固定区间最优平滑 $\tilde{x}(k)K(k)$ 是否也有这样的表示式, 只是方括弧内第一项为 $\hat{x}(k+1)K(k)$, 答案是肯定的, 下面就给出证明。

5.5.4 固定区间最优平滑

定理 5.5.3 考虑离散时间随机状态模型式(5.5.1)和(5.5.2), 固定区间最优平滑为

$$\hat{x}(k)K(k) = \hat{x}(k)K(k) + A(k)[\hat{x}(k+1)K(k) - \hat{x}(k+1)K(k)] \quad (5.5.50)$$

式中: N 为固定区间; $A(k)$ 为

$$A(k) = P(k)K(k)P^{-1}(k+1)K(k) \quad (5.5.51)$$

固定区间最优平滑误差 $\tilde{x}(k|N)$ 的方差阵为

$$P(k|N) = P(k|k) + A(k)[P(k+1|N) - P(k+1|k)]A^T(k) \tag{5.5.52}$$

证明

1. 确定 $\hat{x}(k|N)$ 的递推式

我们用数学归纳法证明。考虑 $\hat{x}(k|N)(1 \leq k)$ ，若初值 $1 = k+1$ 时式(5.5.50)成立，设 $1 = N-1$ 时式(5.5.50)成立，若能证明 $1 = N$ 时式(5.5.50)也成立，则问题得证。

(1) 初值 $1 = k+1$ 时，定理 5.5.1 已给出结果，使式(5.5.50)成立。为方便起见，重写有关结果：

$$\begin{aligned} \hat{x}(k|k+1) &= \hat{x}(k|k) + A(k)[\hat{x}(k+1|k+1) - \hat{x}(k+1|k)] \\ &= \hat{x}(k|k) + A(k)K(k+1|k+1)\tilde{y}(k+1|k) \\ &= \hat{x}(k|k) + M(k|k+1)\tilde{y}(k+1|k) \end{aligned} \tag{5.5.53}$$

式中：

$$A(k) = P(k|k)^T P^{-1}(k+1|k) \tag{5.5.54}$$

$$K(k+1|k+1) = P(k+1|k)^T [P(k+1|k)^T + R_2]^{-1} \tag{5.5.55}$$

$$M(k|k+1) = A(k)K(k+1|k+1) \tag{5.5.56}$$

(2) 设 $1 = N-1$ 时式(5.5.50)成立，即有

$$\hat{x}(k|N-1) = \hat{x}(k|k) + A(k)[\hat{x}(k+1|N-1) - \hat{x}(k+1|k)] \tag{5.5.57}$$

(3) 证明 $1 = N$ 时式(5.5.50)也成立。为此，写出 $\hat{x}(k|N)$ 的基本公式：

$$\begin{aligned} \hat{x}(k|N) &= E\{x(k)|Y_N\} \\ &= \hat{x}(k|N-1) + R_{x(k)\tilde{y}(N|N-1)}R_{\tilde{y}(N|N-1)}^{-1}\tilde{y}(N|N-1) \end{aligned} \tag{5.5.58}$$

下面分别计算各有关项：

$$\begin{aligned} \tilde{y}(N|N-1) &= y(N) - \hat{x}(N-1|N-1) \\ &= \tilde{x}(N|N-1) + (N) \\ &= \tilde{x}(N-1|N-1) + (N-1) + (N) \end{aligned} \tag{5.5.59}$$

$$R_{\tilde{y}(N|N-1)} = P(N|N-1)^T + R_2 \tag{5.5.60}$$

$$\begin{aligned} R_{x(k)\tilde{y}(N|N-1)} &= E\{[x(k) - m_{x(k)}]\tilde{y}^T(N|N-1)\} \\ &= E\{x(k)\tilde{x}^T(N-1|N-1)\}^T \tag{5.5.61} \end{aligned}$$

由式(5.5.60)和(5.5.61)，令

$$\begin{aligned} M(k|N) &= R_{x(k)\tilde{y}(N|N-1)}R_{\tilde{y}(N|N-1)}^{-1} \\ &= E\{x(k)\tilde{x}^T(N-1|N-1)\}^T [P(N|N-1)^T + R_2]^{-1} \end{aligned} \tag{5.5.62}$$

上式中 $\tilde{x}(N-1|N-1)$ 表示为

$$\begin{aligned} \tilde{x}(N-1|N-1) &= x(N-1) - \hat{x}(N-1|N-1) \\ &= x(N-2) + (N-2) - \hat{x}(N-1|N-1) \end{aligned} \tag{5.5.63}$$

式(5.5.63)中 $\hat{x}(N-1|N-1)$ 表示为

$$\hat{x}(N-1|N-1) = \hat{x}(N-2|N-2) + K(N-1|N-1)\tilde{y}(N-1|N-2) \tag{5.5.64}$$

式(5.5.64)中 $\tilde{y}(N-1:N-2)$ 表示为

$$\tilde{y}(N-1:N-2) = \tilde{x}(N-2:N-2) + (N-2) + (N-1) \quad (5.5.65)$$

把式(5.5.65)代入式(5.5.64),再把式(5.5.64)代入式(5.5.63),得到

$$\begin{aligned} \tilde{x}(N-1:N-1) &= [I - K(N-1:N-1)] \tilde{x}(N-2:N-2) \\ &\quad + [I - K(N-1:N-1)] (N-2) \\ &\quad - K(N-1:N-1) (N-1) \end{aligned} \quad (5.5.66)$$

把上式代入式(5.5.62),得到

$$\begin{aligned} M(k:N) &= E\{x(k) \tilde{x}^T(N-2:N-2)\}^T [I - K(N-1:N-1)]^T \\ &\quad \times [P(N:N-1)^T + R_2]^{-1} \end{aligned} \quad (5.5.67)$$

利用卡尔曼滤波 $P(N-1:N-1)$ 和 $K(N:N)$ 公式,得到

$$[I - K(N-1:N-1)]^T = P^{-1}(N-1:N-2)P(N-1:N-1) \quad (5.5.68)$$

$$[P(N:N-1)^T + R_2]^{-1} = P^{-1}(N:N-1)K(N:N) \quad (5.5.69)$$

把以上两式代入式(5.5.67),得到

$$\begin{aligned} M(k:N) &= E\{x(k) \tilde{x}^T(N-2:N-2)\}^T P^{-1}(N-1:N-2) \\ &\quad \times P(N-1:N-1)^T P^{-1}(N:N-1)K(N:N) \\ &= E\{x(k) \tilde{x}^T(N-2:N-2)\}^T P^{-1}(N-1:N-2) \\ &\quad \times A(N-1)K(N:N) \end{aligned} \quad (5.5.70)$$

式中: $A(N-1)$ 称为固定区间最优平滑增益阵,表示为

$$A(N-1) = P(N-1:N-1)^T P^{-1}(N:N-1) \quad (5.5.71)$$

通过由式(5.5.62)至式(5.5.71)的推算,我们把 $M(k:N)$ 中 $E\{x(k) \tilde{x}^T(N-1:N-1)\}$ 变成了 $E\{x(k) \tilde{x}^T(N-2:N-2)\}$,然后再套用这些公式把 $E\{x(k) \tilde{x}^T(N-2:N-2)\}$ 变成 $E\{x(k) \tilde{x}^T(N-3:N-3)\}$,如此反复进行,最后得到

$$M(k:N) = \sum_{i=k}^{N-1} A(i) K(N:N) \quad (5.5.72)$$

把上式代入式(5.5.58),得到

$$\begin{aligned} \hat{x}(k:N) &= \hat{x}(k:N-1) + \sum_{i=k}^{N-1} M(k:N) \tilde{y}(N:N-1) \\ &= \hat{x}(k:N-1) + \sum_{i=k}^{N-1} A(i) K(N:N) \tilde{y}(N:N-1) \\ &= \hat{x}(k:N-1) + A(k) \sum_{i=k+1}^{N-1} A(i) K(N:N) \tilde{y}(N:N-1) \end{aligned} \quad (5.5.73)$$

在上式中用 $k+1$ 置换 k ,得到

$$\hat{x}(k+1:N) = \hat{x}(k+1:N-1) + \sum_{i=k+1}^{N-1} A(i) K(N:N) \tilde{y}(N:N-1) \quad (5.5.74)$$

$$\text{即} \quad \sum_{i=k+1}^{N-1} A(i) K(N:N) \tilde{y}(N:N-1) = \hat{x}(k+1:N) - \hat{x}(k+1:N-1) \quad (5.5.75)$$

把上式代入式(5.5.73),得到

$$\hat{x}(k:N) = \hat{x}(k:N-1) + A(k) [\hat{x}(k+1:N) - \hat{x}(k+1:N-1)] \quad (5.5.76)$$

再把式(5.5.57)代入上式,得到最后结果为

$$\hat{x}(k|N) = \hat{x}(k|k) + A(k)[\hat{x}(k+1|N) - \hat{x}(k+1|k)] \quad (5.5.77)$$

2. 确定 $P(k|N)$ 的递推式

把式(5.5.77)改写为

$$\hat{x}(k|N) - A(k)\hat{x}(k+1|N) = \hat{x}(k|k) - A(k)\hat{x}(k+1|k) \quad (5.5.78)$$

用 $\tilde{x}(k) = x(k) - \hat{x}(k)$ 减去上式两边,得到

$$\tilde{x}(k|N) + A(k)\hat{x}(k+1|N) = \tilde{x}(k|k) + A(k)\hat{x}(k+1|k) \quad (5.5.79)$$

$$[\tilde{x}(k|N) + A(k)\hat{x}(k+1|N)]^T = [\tilde{x}(k|k) + A(k)\hat{x}(k+1|k)]^T \quad (5.5.80)$$

把以上两式相乘并取均值,注意到 $\tilde{x}(k|N)$ 与 $\hat{x}(k+1|N)$ 互相独立, $\tilde{x}(k|k)$ 与 $\hat{x}(k+1|k)$ 互相独立,得到

$$P(k|N) + A(k)P_{\Lambda(k+1|k)}A^T(k) = P(k|k) + A(k)P_{\Lambda(k+1|k)}A^T(k) \quad (5.5.81)$$

$$P(k|N) = P(k|k) + A(k)[P_{\Lambda(k+1|k)} - P_{\Lambda(k+1|N)}]A^T(k) \quad (5.5.82)$$

式中: $P(k|N)$ 为固定区间最优平滑误差的方差阵; $P(k|k)$ 为最优滤波误差的方差阵; $P_{\Lambda(k+1|k)}$ 和 $P_{\Lambda(k+1|N)}$ 分别为

$$P_{\Lambda(k+1|k)} = E\{\hat{x}(k+1|k)\hat{x}^T(k+1|k)\} \quad (5.5.83)$$

$$P_{\Lambda(k+1|N)} = E\{\hat{x}(k+1|N)\hat{x}^T(k+1|N)\} \quad (5.5.84)$$

为置换 $[P_{\Lambda(k+1|k)} - P_{\Lambda(k+1|N)}]$, 考虑等式

$$\tilde{x}(k+1|k) + \hat{x}(k+1|k) = \tilde{x}(k+1|N) + \hat{x}(k+1|N) \quad (5.5.85)$$

$$[\tilde{x}(k+1|k) + \hat{x}(k+1|k)]^T = [\tilde{x}(k+1|N) + \hat{x}(k+1|N)]^T \quad (5.5.86)$$

把以上两式相乘,并取均值,得到

$$P_{\Lambda(k+1|k)} - P_{\Lambda(k+1|N)} = P(k+1|N) - P(k+1|k) \quad (5.5.87)$$

把上式代入式(5.5.82),得到最后结果

$$P(k|N) = P(k|k) + A(k)[P(k+1|N) - P(k+1|k)]A^T(k) \quad (5.5.88)$$

把上式与一步平滑式(5.5.7)和二步平滑式(5.5.30)比较可知,公式的区别只在右边括弧内第一项不同,便于记忆,便于计算。

3. $\hat{x}(k|N)$ 和 $P(k|N)$ 的计算

固定区间最优平滑 $\hat{x}(k|N)$ 的实际计算按逆时间递归进行,把最优滤波 $\hat{x}(N|N)$ 作为初值(边界条件),通过式(5.5.50)先求出 $\hat{x}(N-1|N)$,再求出 $\hat{x}(N-2|N)$,依次类推就可导出 $\hat{x}(k|N)$ 。

同样,固定区间最优平滑误差 $\tilde{x}(k|N)$ 的方差阵 $P(k|N)$ 的实际计算也按逆时递归进行,把 $P(N|N)$ 作为初值,通过式(5.5.52)先求出 $P(N-1|N)$,再求出 $P(N-2|N)$,依次类推就可导出 $P(k|N)$ 。

$\hat{x}(k|N)$ 和 $P(k|N)$ 的计算步骤可如图 5.5.1 表示。

5.5.5 固定点最优平滑

定理 5.5.4 考虑离散时间随机状态模型式(5.5.1)和(5.5.2),对 $j = k+1, k+2, \dots, k$ 为固定点,则固定点最优平滑为

图 5.5.1 固定区间平滑 $\hat{x}(k|N)$ 和 $P(k|N)$ 的计算步骤

$$\begin{aligned}\hat{x}(k|j) &= \hat{x}(k|j-1) + M(k|j) \tilde{y}(j|j-1) \\ &= \hat{x}(k|j-1) + B(j)[\hat{x}(j|j) - \hat{x}(j|j-1)]\end{aligned}\quad (5.5.89)$$

式中:

$$\begin{aligned}M(k|j) &= \sum_{i=k}^{j-1} A(i) K(j|i) \\ &= B(j) K(j|j)\end{aligned}\quad (5.5.90)$$

$$\begin{aligned}B(j) &= \sum_{i=k}^{j-1} A(i) \\ &= B(j-1)A(j-1)\end{aligned}\quad (5.5.91)$$

$$A(i) = P(i|i)^T P^{-1}(i+1|i) \quad (5.5.92)$$

$B(j)$ 称为固定点最优平滑增益矩阵。固定点最优平滑误差 $\tilde{x}(k|j)$ 的方差阵为

$$\begin{aligned}P(k|j) &= P(k|j-1) - B(j)K(j|j)P(j|j-1)B^T(j) \\ &= P(k|j-1) + B(j)[P(j|j) - P(j|j-1)]B^T(j)\end{aligned}\quad (5.5.93)$$

证明

(1) 确定 $\hat{x}(k|j)$ 的递推式

在式(5.5.73)中, 令 k 为固定点, 用 j 置换 N , 就得到式(5.5.89)的第一式, 再把式(5.5.90)代入, 得到

$$\begin{aligned}\hat{x}(k|j) &= \hat{x}(k|j-1) + M(k|j) \tilde{y}(j|j-1) \\ &= \hat{x}(k|j-1) + B(j)K(j|j) \tilde{y}(j|j-1)\end{aligned}\quad (5.5.94)$$

把最优滤波公式

$$K(j|j) \tilde{y}(j|j-1) = \hat{x}(j|j) - \hat{x}(j|j-1) \quad (5.5.95)$$

代入式(5.5.94), 得到式(5.5.89)第二式。

(2) 确定 $P(k|j)$ 的递推式

把式(5.5.89)改写为

$$\hat{x}(k|j) - B(j) \hat{x}(j|j) = \hat{x}(k|j-1) - B(j) \hat{x}(j|j-1) \quad (5.5.96)$$

用 $\tilde{x}(k) = x(k) - \hat{x}(k)$ 减去上式两边得到

$$\tilde{x}(k|j) + B(j) \hat{x}(j|j) = \tilde{x}(k|j-1) + B(j) \hat{x}(j|j-1) \quad (5.5.97)$$

对上式两边取方差, 注意到 \tilde{x} 与 \hat{x} 互相独立, 得到

$$P(k|j) = P(k|j-1) + B(j)[P_{\hat{x}(j|j-1)} - P_{\hat{x}(j|j)}]B^T(j) \quad (5.5.98)$$

采用式(5.5.85) ~ (5.5.87)的同样方法, 容易求出

$$P_{\Lambda}(j|j-1) - P_{\Lambda}(j|j) = P(j|j) - P(j|j-1) \quad (5.5.99)$$

把上式代入式(5.5.98), 得到

$$P(k|j) = P(k|j-1) + B(j)[P(j|j) - P(j|j-1)]B^T(j) \quad (5.5.100)$$

把最优滤波公式式(5.4.38) 改写为

$$P(j|j) - P(j|j-1) = -K(j|j)P(j|j-1) \quad (5.5.101)$$

将上式代入式(5.5.100), 得到

$$P(k|j) = P(k|j-1) - B(j)K(j|j)P(j|j-1)B^T(j) \quad (5.5.102)$$

(3) $\hat{x}(k|j)$ 和 $P(k|j)$ 的计算

$\hat{x}(k|j)$ 的实际计算是把 $\hat{x}(k|k)$ 作为初值, 沿时间递推得到 $\hat{x}(k|k+1)$, $\hat{x}(k|k+2)$, ...。

同样, $P(k|j)$ 的实际计算是把 $P(k|k)$ 作为初值, 沿时间递推得到 $P(k|k+1)$, $P(k|k+2)$, ...。

$\hat{x}(k|j)$ 和 $P(k|j)$ 的计算步骤如图 5.5.2 所示。

图 5.5.2 固定点最优平滑 $\hat{x}(k|j)$ 和 $P(k|j)$ 的计算步骤

5.5.6 固定滞后最优平滑

定理 5.5.5 考虑离散时间随机状态模型式(5.5.1) 和(5.5.2), 对固定滞后时间 N , 固定滞后最优平滑为

$$\begin{aligned} \hat{x}(k+1|k+1+N) = & \hat{x}(k|k+N) + C(k+1+N)K(k+1+N) \\ & \tilde{y}(k+1+N|k+N) + V(k+1)[\hat{x}(k|k+N) - \hat{x}(k|k)] \end{aligned} \quad (5.5.103)$$

式中 $K(k+1+N)$ 表示最优滤波增益阵 $K(k+1+N|k+1+N)$,

$$C(k+1+N) = \sum_{i=k+1}^{k+N} A(i) = B(k+1+N) \quad (5.5.104)$$

$$A(i) = P(i|i)^T P^{-1}(i+1|i) \quad (5.5.105)$$

$$V(k+1) = R_1[{}^T]^{-1} P^{-1}(k|k) \quad (5.5.106)$$

固定滞后最优平滑误差 $\tilde{x}(k+1|k+1+N)$ 的方差阵为

$$\begin{aligned} P(k+1|k+1+N) = & P(k+1|k) - C(k+1+N)K(k+1+N) \\ & \times P(k+1+N|k+N)C^T(k+1+N) - A^{-1}(k) \\ & \times [P(k|k) - P(k|k+N)][A^T(k)]^{-1} \end{aligned} \quad (5.5.107)$$

证明 利用固定区间和固定点最优平滑的结果进行证明。

(1) 确定 $\hat{x}(k+1|k+1+N)$ 的递推式

把

$$\hat{x}(k+1|k) = \hat{x}(k|k) \quad (5.5.108)$$

代入固定区间最优平滑

$$\hat{x}(k|N) = \hat{x}(k|k) + A(k)[\hat{x}(k+1|N) - \hat{x}(k+1|k)] \quad (5.5.109)$$

并进行 $\pm A(k) \hat{x}(k|N)$ 的运算, 得到

$$\begin{aligned} \hat{x}(k+1|N) &= \hat{x}(k|N) + [A^{-1}(k) - 1][\hat{x}(k|N) - \hat{x}(k|k)] \\ &= \hat{x}(k|N) + V(k+1)[\hat{x}(k|N) - \hat{x}(k|k)] \end{aligned} \quad (5.5.110)$$

式中已令

$$V(k+1) = A^{-1}(k) - 1 \quad (5.5.111)$$

为简化上式, 把最优滤波公式

$$P(k+1|k) = P(k|k)^T + R_1 \quad (5.5.112)$$

代入式(5.5.105), 并求逆, 得到

$$A^{-1}(k) = 1 - R_1[1^T]^{-1}P^{-1}(k|k) \quad (5.5.113)$$

把上式代入式(5.5.111), 就得到式(5.5.106)。我们可以把式(5.5.110)看做为固定区间最优平滑的另一种形式。在式(5.5.110)中, 用 $k+N$ 置换 N , 有

$$\hat{x}(k+1|k+N) = \hat{x}(k|k+N) + V(k+1)[\hat{x}(k|k+N) - \hat{x}(k|k)] \quad (5.5.114)$$

把上式代入固定点最优平滑式(5.5.89), 并用 $k+1$ 置换 k , 用 $k+1+N$ 置换 j , 再注意到式(5.5.90), 得到

$$\begin{aligned} \hat{x}(k+1|k+1+N) &= \hat{x}(k+1|k+N) + B(k+1+N)K(k+1+N) \\ &\quad \tilde{y}(k+1+N|k+N) \\ &= \hat{x}(k|k+N) + C(k+1+N)K(k+1+N) \\ &\quad \tilde{y}(k+1+N|k+N) + V(k+1) \\ &\quad [\hat{x}(k|k+N) - \hat{x}(k|k)] \end{aligned} \quad (5.5.115)$$

为区别两种平滑, 上式中采用了新符号, 如式(5.5.104)所示。

(2) 确定 $P(k+1|k+1+N)$ 递推式

对固定区间最优平滑式(5.5.52)进行移项整理, 得到

$$P(k+1|N) = P(k+1|k) - A^{-1}(k)[P(k|k) - P(k|N)][A^T(k)]^{-1} \quad (5.5.116)$$

在上式中用 $k+N$ 置换 N , 得到

$$P(k+1|k+N) = P(k+1|k) - A^{-1}(k)[P(k|k) - P(k|k+N)][A^T(k)]^{-1} \quad (5.5.117)$$

另外, 考虑固定点最优平滑式(5.5.93), 在式中用 $k+1$ 置换 k , 用 $k+1+N$ 置换 j , 有

$$\begin{aligned} P(k+1|k+1+N) &= P(k+1|k+N) - B(k+1+N)K(k+1+N) \\ &\quad P(k+1+N|k+N)B^T(k+1+N) \end{aligned} \quad (5.5.118)$$

把式(5.5.117)和(5.5.104)代入上式, 就得到式(5.5.107)。

(3) $\hat{x}(k|k+N)$ 和 $P(k|k+N)$ 的计算

实际计算 $\hat{x}(k|k+N)$ 时, 把 $\hat{x}(k|k)$ 作为初值, 用式(5.5.103) 沿时间递推得到 $\hat{x}(k+1|k+1)$, $\hat{x}(k+2|k+2)$, ...。

同样, 实际计算 $P(k|k+N)$ 时, 把 $P(k|k)$ 作为初值, 用式(5.5.107) 沿时间递推得到 $P(k+1|k+1)$, $P(k+2|k+2)$, ...。

$\hat{x}(k|k+N)$ 和 $P(k|k+N)$ 的计算步骤如图 5.5.3 所示。

图 5.5.3 固定滞后最优平滑 $\hat{x}(k|k+N)$ 和 $P(k|k+N)$ 的计算步骤

例 5.5.1 已知离散时间随机状态标量方程

$$x(k+1) = x(k) + w(k) \tag{5.5.119}$$

$$y(k) = x(k) + v(k) \tag{5.5.120}$$

式中: $w(k) \sim N(0, 25)$, $v(k) \sim N(0, 15)$, $x(k_0) \sim N(10, 100)$ 互相独立。试求:

- (1) 最优滤波误差的方差阵 $P(k|k)$ ($k=1, 2, 3, 4$); 并求定态值 $P = P(\infty)$;
- (2) 固定区间最优平滑误差的方差阵 $P(k|N)$ ($k=0, 1, 2, 3$), $N=4$ 计算中把(1)中
最优滤波误差的方差阵 $P(4|4)$ 作为初值;
- (3) 固定点最优平滑误差的方差阵 $P(0|j)$ ($j=1, 2, 3, 4$)。

解 (1) 先确定最优滤波 $P(k|k)$ ($k=1, 2, 3, 4$)。把式(5.5.119), (5.5.120) 与式(5.5.1), (5.5.2) 比较, 可知

$$F = 1 \quad H = 1 \quad R_1 = 25 \quad R_2 = 15 \quad R_0 = 100 \tag{5.5.121}$$

把上式数据代入最优滤波公式(5.4.37), (5.4.36)和式(5.4.38), 得到

$$P(k+1|k) = P(k|k) + 25 \tag{5.5.122}$$

$$K(k+1|k+1) = \frac{P(k+1|k)}{P(k+1|k) + 15} \tag{5.5.123}$$

$$P(k+1|k+1) = 15K(k+1|k+1) \tag{5.5.124}$$

把 $R_0 = P(0|0) = 100$ 作为初值, 依次利用以上三式递推求得 $P(k|k-1)$, $K(k|k)$ 和 $P(k|k)$ ($k=1, 2, 3, 4$)。计算结果在表 5.5.1 中列出。

表 5.5.1 最优滤波计算结果

k	$P(k k-1)$	$K(k k)$	$P(k k)$
0			100
1	125	0.89	13.40
2	38.40	0.72	10.80
3	35.80	0.70	10.57
4	35.60	0.70	10.55

再求定态值。由式(5.5.124), (5.5.123)和(5.5.122)导出 P 的定态方程

$$P = \frac{15(P + 25)}{P + 40} \tag{5.5.125}$$

其解为 $P = 10.55, -35.55$, 略去负根, 取 $P = 10.55$ 。把这个解与表 5.5.1 中结果比较可知, $k = 4$ 时, $P(4|4)$ 已达定态。

(2) 把式(5.5.121)中数据代入固定区间最优平滑公式(5.5.7)和(5.5.4), 得到

$$P(k|4) = P(k|k) + A^2(k)[P(k+1|4) - P(k+1|k)] \tag{5.5.126}$$

$$A(k) = \frac{P(k|k)}{P(k+1|k)} \tag{5.5.127}$$

将(1)中 $P(k|k)$ 和 $P(k|k-1)$ 的值代入式(5.5.127), 求出 $A(k)$ ($k = 0, 1, 2, 3$)。把(1)中 $P(k|k)$ 和 $P(k|k-1)$ 代入式(5.5.126), 选择初值为 $P(4|4) = 10.55$, 递推得到 $P(k|N)$ ($k = 0, 1, 2, 3$)。计算结果在表 5.5.2 中列出。

表 5.5.2 固定区间最优平滑计算结果

k	$P(k k)$	$P(k k-1)$	$A(k)$	$P(k 4)$
0	100		0.80	26.23
1	13.40	125	0.35	9.73
2	10.80	38.40	0.30	8.30
3	10.57	35.80	0.30	8.36
4	10.55	35.60		10.55

(3) 把(1)和(2)中有关数据代入式(5.5.93), 得到

$$P(0|j) = P(0|j-1) + B^2(j)[P(j|j) - P(j|j-1)] \tag{5.5.128}$$

$$B(j) = A(0)A(1) \dots A(j-1) \tag{5.5.129}$$

将 $R_0 = P(0|0) = 100$ 作为初值, 利用式(5.5.128), 递推得到 $P(0|j)$ ($j = 1, 2, 3, 4$)。计算结果列于表 5.5.3 中。

表 5.5.3 固定点最优平滑计算结果

k, j	$P(k k)$	$P(k k-1)$	$A(k)$	$B(j)$	$P(0 j)$
0	100		0.80		100
1	13.40	125	0.35	0.80	28.58
2	10.80	38.40	0.30	0.28	26.43
3	10.57	35.80	0.30	0.08	26.25
4	10.55	35.60		0.03	26.23

5.6 有色噪声情况下的最优估计

前面讨论最优估计问题时, 都假设模型噪声 $w(k)$ 和量测噪声 $v(k)$ 都为离散时间高斯

白噪声, 分别表示为 $(k) \sim N(0, R_1)$ 和 $(k) \sim N(0, R_2)$, 并且两者互相独立, 这符合一类实际情况。但在实际应用中还存在另外一类情况, 即 (k) 和 (k) 为有色噪声, 本节中, 我们就讨论这种情况。内容包括: 先介绍成形滤波器, 说明有色噪声如何表示; 之后讨论模型噪声 (k) 为有色噪声情况下的最优估计; 再后介绍模型噪声 (k) 与量测噪声相关的情况, 后面会用到所得的结果; 最后讨论量测噪声 (k) 为有色噪声时的最优估计。

5. 6. 1 成形滤波器

定义 5. 6. 1 考虑 (k) 是均值为零和协方差函数阵为

$$R(k, l) = E\{ (k)^T(l) \} \tag{5. 6. 1}$$

的随机序列, 且 $R(k, l)$ 为正定阵。如果 (k) 可由差分方程(状态方程)

$$(k + 1) = (k + 1, k) (k) + v(k) \tag{5. 6. 2}$$

描述, 式中 $v(k)$ 为白噪声序列, 与 (k) 互相独立, 则称差分方程式(5. 6. 2)为随机序列 (k) 的成形滤波器。

给定 (k) 的均值和协方差阵时, $(k+ 1, k)$ 和 $v(k)$ 的均值和协方差阵由下面定理给出:

定理 5. 6. 1 随机序列 (k) 由成形滤波器式(5. 6. 2)描述, 给定 (k) 的均值为零, 协方差阵为 $R(k, l)$, $v(k)$ 为白噪声序列, 则

(1) $v(k)$ 的均值为零, 即

$$E\{v(k)\} = 0 \tag{5. 6. 3}$$

(2) 状态转移阵 $(k+ 1, k)$ 为

$$(k + 1, k) = R(k + 1, k)R^{-1}(k) \tag{5. 6. 4}$$

(3) $v(k)$ 的方差阵 $R_v(k)$ 为

$$R_v(k) = R(k + 1) - R(k + 1, k)R^{-1}(k, k)R(k, k + 1) \tag{5. 6. 5}$$

证明

(1) 对式(5. 6. 2)两边取均值, 注意到

$$E\{ (k) \} = 0 \tag{5. 6. 6}$$

立即得到式(5. 6. 3)。

(2) 计算 $\{k\}$ 的协方差函数 $R(k+ 1, k)$ 。把式(5. 6. 2)代入 $R(k+ 1, k)$, 得到

$$\begin{aligned} R(k + 1, k) &= E\{ (k + 1)^T(k) \} \\ &= E\{ [(k + 1, k) (k) + v(k)]^T(k) \} \\ &= (k + 1, k)R(k, k) \end{aligned} \tag{5. 6. 7}$$

上式中由第二等式到第三等式利用了 $v(k)$ 与 (k) 互相独立这一性质。由式(5. 6. 7)可立即得到式(5. 6. 4)。

(3) 用式(5. 6. 2)计算 $v(k)$ 的方差阵 $R_v(k)$ 。考虑到式(5. 6. 4), 则有

$$\begin{aligned} R_v(k) &= E\{ v(k)v^T(k) \} \\ &= E\{ [(k + 1) - (k + 1, k) (k)][(k + 1) - (k + 1, k) (k)]^T \} \\ &= E\{ (k + 1)^T(k + 1) \} - E\{ (k + 1)^T(k) \}^T(k + 1, k) \end{aligned}$$

$$\begin{aligned}
& - (k+1, k)E\{ (k)^T(k+1) \} + (k+1, k)E\{ (k)^T(k) \}^T(k+1, k) \\
= & R(k+1) - R(k+1, k)R^{-1}(k)R^T(k+1, k) \\
& - R(k+1, k)R^{-1}(k)R(k, k+1) + R(k+1, k)R^{-1}(k)R(k)R^{-1}(k)R^T(k+1, k) \\
= & R(k+1) - R(k+1, k)R^{-1}(k)R(k, k+1) \tag{5.6.8}
\end{aligned}$$

下面讨论模型噪声 (k) 或量测噪声 (k) 为成形滤波器式(5.6.2)形式描述的有色噪声时的最优估计问题, 讨论中, 我们只推导最优滤波公式, 其他问题可以类推。

5.6.2 模型噪声为有色噪声情况下的最优滤波

重写系统状态模型和模型噪声的成形滤波器方程

$$x(k+1) = x(k) + (k) \tag{5.6.9}$$

$$y(k) = x(k) + (k) \tag{5.6.10}$$

$$(k+1) = (k) + v(k) \tag{5.6.11}$$

式中: $x(k)$ 为 $n \times 1$ 状态向量, 初值为 $x(k_0) \sim N(m_0, R_0)$; $y(k)$ 为 $m \times 1$ 量测向量; Φ 为 $n \times n$ 状态转移阵; H 为 $m \times n$ 量测阵; F 为 $n \times n$ 成形滤波器状态转移阵; (k) 为 $n \times 1$ 模型有色噪声 $N(0, R_1)$; (k) 为 $m \times 1$ 量测白噪声 $N(0, R_2)$; $v(k)$ 为 $n \times 1$ 白噪声。 Φ , H 和 F 可以是时变的或时不变的。

假设 $x(k)$, (k) 和 (k) 两两互相独立, 由于 $x(k_0)$, (k) 和 (k) 是高斯的, 使得 $x(k)$ 和 $y(k)$ 也是高斯的。 R_1 和 R_2 可以是时变的或时不变的。

我们进一步假设 $v(k)$ 与 $x(k)$, (k) 也是互相独立的。

由定理 5.6.1 知道, $(k+1, k)$, $v(k)$ 的均值 m_v 和方差 R_v 分别为

$$(k+1, k) = R_1(k+1, k)R_1^{-1}(k) \tag{5.6.12}$$

$$m_v = 0 \tag{5.6.13}$$

$$R_v(k) = R_1(k+1) - R_1(k+1, k)R_1^{-1}(k)R_1(k, k+1) \tag{5.6.14}$$

对上述问题, 我们利用扩充状态向量法来求解最优滤波问题。

由式(5.6.9)和(5.6.11)组成扩充状态方程

$$\begin{aligned}
\begin{bmatrix} x(k+1) \\ (k+1) \end{bmatrix} &= \begin{bmatrix} (k+1, k) & I \\ 0 & (k+1, k) \end{bmatrix} \begin{bmatrix} x(k) \\ (k) \end{bmatrix} + \begin{bmatrix} 0 \\ v(k) \end{bmatrix} \tag{5.6.15}
\end{aligned}$$

$$x(k+1) = (k+1, k)x(k) + (k) \tag{5.6.16}$$

式中:

$$\begin{aligned}
\begin{bmatrix} x(k) \\ (k) \end{bmatrix} &= \begin{bmatrix} x(k) \\ (k) \end{bmatrix} \quad (k+1, k) = \begin{bmatrix} (k+1, k) & I \\ 0 & (k+1, k) \end{bmatrix} \tag{5.6.17}
\end{aligned}$$

$$\begin{aligned}
\begin{bmatrix} 0 \\ v(k) \end{bmatrix} &= \begin{bmatrix} 0 \\ v(k) \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
E\{ (k) \} &= 0 \quad R_1(k) = \begin{bmatrix} 0 & 0 \\ 0 & R_3(k) \end{bmatrix} \tag{5.6.18}
\end{aligned}$$

相应地, 对量测方程式(5.6.10)也写成扩充状态 $x(k)$ 的量测方程

$$y(k) = [\quad 0] \begin{bmatrix} x(k) \\ (k) \end{bmatrix} + (k) \tag{5.6.19}$$

$$y(k) = x(k) + v(k) \quad (5.6.20)$$

式中:

$$= \begin{bmatrix} 1 & 0 \end{bmatrix} \quad (5.6.21)$$

对照式(5.6.15), (5.6.16)和(5.6.20)可知, $v(k)$ 与 $x(k)$ 互相独立。再令 $v(k)$ 为离散时间高斯白噪声 $N(0, R_2)$, 扩充状态方程式(5.6.16)和(5.6.20)就符合求解最优滤波定理 5.4.2 的要求。下面通过一个定理直接给出最优滤波的结果。

定理 5.6.2 对状态方程式(5.6.9), (5.6.10)和成形滤波器式(5.6.11), 给定条件和假设如本节前面所述, 则最优滤波可通过构造扩充状态方程式(5.6.16)和式(5.6.20)得到

$$\hat{x}(k+1|k+1) = \hat{x}(k|k) + K(k+1|k+1)[y(k+1) - \hat{y}(k|k)] \quad (5.6.22)$$

$$K(k+1|k+1) = P(k+1|k)^T [P(k+1|k)^T + R_2]^{-1} \quad (5.6.23)$$

$$P(k+1|k) = P(k|k)^T + R_1 \quad (5.6.24)$$

$$\begin{aligned} P(k+1|k+1) &= [I - K(k+1|k+1)]P(k+1|k)[I - K(k+1|k+1)]^T \\ &\quad + K(k+1|k+1)R_2K^T(k+1|k+1) \\ &= [I - K(k+1|k+1)]P(k+1|k) \end{aligned} \quad (5.6.25)$$

$$x(k_0|k_0) = E\{x(k_0)\} = \begin{bmatrix} m_0 \\ m_0 \end{bmatrix} = m_0 \quad (5.6.26)$$

$$P(k_0|k_0) = R(x(k_0)) = \begin{bmatrix} R_0 & 0 \\ 0 & R_1 \end{bmatrix} = R_0 \quad (5.6.27)$$

5.6.3 模型噪声和量测噪声相关的最优滤波

考虑系统状态方程式(5.6.9)和(5.6.10), 给定状态初值 $x(k_0) \sim N(m_0, R_0)$; 模型噪声 $w(k)$ 为高斯 $N(0, R_1)$ 白噪声, 与 $x(k_0)$ 互相独立; 量测噪声 $v(k)$ 为高斯 $N(0, R_2)$ 白噪声, 与 $x(k_0)$ 互相独立; 但 $w(k)$ 与 $v(k)$ 相关, 其协方差函数为

$$R \begin{pmatrix} w(k) \\ v(l) \end{pmatrix} = E \begin{bmatrix} w(k) \\ v(l) \end{bmatrix} \begin{bmatrix} w(k) \\ v(l) \end{bmatrix}^T = \begin{bmatrix} R_1(k) & R_{12}(k, l) \\ R_{21}(l, k) & R_2(l) \end{bmatrix} \quad (5.6.28)$$

定理 5.6.3 考虑离散时间随机状态方程式(5.6.9)和(5.6.10), 模型噪声 $w(k)$ 与 $v(k)$ 是相关的, 给定条件和假设如前所述, 则最优滤波 $\hat{x}(k+1|k+1)$ 满足递推方程

$$\hat{x}(k+1|k+1) = \hat{x}(k|k) + K(k+1|k+1)[y(k+1) - \hat{y}(k|k)] \quad (5.6.29)$$

最优滤波增益阵 $K(k+1|k+1)$ 表示为

$$K(k+1|k+1) = [P(k+1|k)^T + R_{12}(k, k+1)][P(k+1|k)^T + R_2]^{-1} \quad (5.6.30)$$

式(5.6.30)中一步预测误差 $\tilde{x}(k+1|k)$ 的方差阵 $P(k+1|k)$ 表示为

$$P(k+1|k) = P(k|k)^T + R_1 \quad (5.6.31)$$

式中最优滤波误差 $\tilde{x}(k|k)$ 的方差阵 $P(k|k)$ 表示为

$$\begin{aligned} P(k+1|k+1) &= [I - K(k+1|k+1)]P(k+1|k)[I - K(k+1|k+1)]^T \\ &\quad + K(k+1|k+1)R_2K^T(k+1|k+1) \end{aligned}$$

$$= [I - K(k+1)C]P(k+1) \quad (5.6.32)$$

初值为

$$\hat{x}(k_0) = m_0 \quad (5.6.33)$$

$$P(k_0) = R_0 \quad (5.6.34)$$

证明 利用状态估计基本计算公式进行证明, 由于读者已经熟悉这些公式, 这里主要给出有关结果。

由定理 5.3.5, 最优滤波递推式为

$$\begin{aligned} \hat{x}(k+1) &= E\{x(k+1) | Y_{k+1}\} \\ &= E\{x(k+1) | Y_k\} + R_{x(k+1)\tilde{y}(k+1)} R_{\tilde{y}(k+1)}^{-1} \tilde{y}(k+1) \end{aligned} \quad (5.6.35)$$

式中:

$$\begin{aligned} E\{x(k+1) | Y_k\} &= \hat{x}(k+1) = \hat{x}(k) \\ \tilde{y}(k+1) &= y(k+1) - \hat{y}(k+1) \end{aligned} \quad (5.6.36)$$

$$\begin{aligned} &= y(k+1) - \hat{x}(k)C \\ &= \tilde{x}(k+1) + (k+1) \end{aligned} \quad (5.6.37)$$

$$\begin{aligned} R_{\tilde{y}(k+1)} &= E\{\tilde{y}(k+1) \tilde{y}^T(k+1)\} \\ &= E\{[\tilde{x}(k+1) + (k+1)][\tilde{x}(k+1) + (k+1)]^T\} \\ &= P(k+1) + R_2 \end{aligned} \quad (5.6.38)$$

$$\begin{aligned} R_{x(k+1)\tilde{y}(k+1)} &= E\{[x(k+1) - m_{x(k+1)}][\tilde{x}(k+1) + (k+1)]^T\} \\ &= E\{[x(k) + (k) - m_{x(k)}][\tilde{x}(k+1) + (k+1)]^T\} \\ &= P(k) + R_{12}(k, k+1) \end{aligned} \quad (5.6.39)$$

令

$$\begin{aligned} K(k+1) &= R_{x(k+1)\tilde{y}(k+1)} R_{\tilde{y}(k+1)}^{-1} \\ &= [P(k) + R_{12}(k, k+1)][P(k) + R_2]^{-1} \end{aligned} \quad (5.6.40)$$

则上式就是式(5.6.30), 把上式和式(5.6.36)和(5.6.37)代入式(5.6.35)就得到式(5.6.29)。

由于

$$\begin{aligned} \tilde{x}(k+1) &= x(k+1) - \hat{x}(k+1) \\ &= \tilde{x}(k) - (k) \end{aligned} \quad (5.6.41)$$

把上式代入 $P(k+1)$ 的定义式就得到式(5.6.31)。

把 $\tilde{x}(k+1)$ 表示式

$$\begin{aligned} \tilde{x}(k+1) &= x(k+1) - \hat{x}(k+1) \\ &= [I - K(k+1)C] \tilde{x}(k) - K(k+1)(k+1) \end{aligned} \quad (5.6.42)$$

代入 $P(k+1)$ 的定义式就得到式(5.6.32)。

很自然, 状态初值由式(5.6.33)和(5.6.34)确定。

下面讨论量测噪声为有色噪声的最优估计时,要利用上述结果。

5.6.4 量测噪声为有色噪声时的最优滤波

考虑离散时间随机状态方程

$$x(k+1) = \Phi(k+1, k)x(k) + \Gamma(k)w(k) \quad (5.6.43)$$

$$y(k) = H(k)x(k) + v(k) \quad (5.6.44)$$

式中: $w(k)$ 为有色噪声, 表示为

$$w(k+1) = \Phi_w(k+1, k)w(k) + e(k) \quad (5.6.45)$$

已知 $w(k) \sim N(0, R_2)$ 时, 由定理 5.6.1 可知, $\Phi_w(k+1, k)$ 和 $e(k) \sim N(0, R_4)$ 表示为

$$\Phi_w(k+1, k) = R_2(k+1, k)R_2^{-1}(k) \quad (5.6.46)$$

$$R_4(k) = R_2(k+1) - R_2(k+1, k)R_2^{-1}(k)R_2(k, k+1) \quad (5.6.47)$$

其他给定条件和假设如前所述。

下面用扩充状态向量法和等效量测方程法来求解最优滤波问题。

1. 扩充状态向量法

把式(5.6.43)和(5.6.45)组合在一起, 得到

$$\begin{bmatrix} x(k+1) \\ w(k+1) \end{bmatrix} = \begin{bmatrix} \Phi(k+1, k) & 0 \\ 0 & \Phi_w(k+1, k) \end{bmatrix} \begin{bmatrix} x(k) \\ w(k) \end{bmatrix} + \begin{bmatrix} \Gamma(k) \\ 0 \end{bmatrix} \begin{bmatrix} w(k) \\ e(k) \end{bmatrix} \quad (5.6.48)$$

$$x^*(k+1) = \Phi^*(k+1, k)x^*(k) + \Gamma^*(k) \begin{bmatrix} w(k) \\ e(k) \end{bmatrix} \quad (5.6.49)$$

式中:

$$\begin{aligned} x^*(k) &= \begin{bmatrix} x(k) \\ w(k) \end{bmatrix} & \Phi^*(k+1, k) &= \begin{bmatrix} \Phi(k+1, k) & 0 \\ 0 & \Phi_w(k+1, k) \end{bmatrix} \\ \Gamma^*(k) &= \begin{bmatrix} \Gamma(k) \\ 0 \end{bmatrix} & & \end{aligned} \quad (5.6.50)$$

相应地, 量测方程式(5.6.44)改写为

$$y(k) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x(k) \\ w(k) \end{bmatrix} \quad (5.6.51)$$

$$y(k) = H^*x^*(k) \quad (5.6.52)$$

式中:

$$H^* = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad (5.6.53)$$

有关噪声参数为

$$E\{\Gamma^*(k)\} = 0 \quad (5.6.54)$$

$$\begin{aligned} R_1^*(k) &= E \begin{bmatrix} \Gamma(k) & 0 \\ 0 & e(k) \end{bmatrix} \begin{bmatrix} \Gamma(k) & 0 \\ 0 & e(k) \end{bmatrix}^T \\ &= \begin{bmatrix} R_1(k) & R_{14}(k, k) \\ R_{41}(k, k) & R_4(k) \end{bmatrix} \end{aligned} \quad (5.6.55)$$

对扩充状态方程式(5.6.49)和(5.6.52), 容易写出最优滤波方程组为

$$\hat{x}^*(k+1|k+1) = \Phi^*(k+1|k)\hat{x}^*(k|k) + K^*(k+1|k+1)[y(k+1) - H^*\hat{x}^*(k|k)] \quad (5.6.56)$$

$$K^*(k+1|k+1) = P^*(k+1|k)^{-T} [- P^*(k+1|k)^{-T}]^{-1} \quad (5.6.57)$$

$$P^*(k+1|k) = - P^*(k|k) + R_1^* \quad (5.6.58)$$

$$P^*(k+1|k+1) = [I - K^*(k+1|k+1)^{-T}] P^*(k+1|k) \\ \times [I - K^*(k+1|k+1)^{-T}]^T \quad (5.6.59)$$

状态初值为

$$\hat{x}^*(k_0|k_0) = m_0^* \quad (5.6.60)$$

$$P^*(k_0|k_0) = R_0^* \quad (5.6.61)$$

$$m_0^* = \begin{matrix} m_0 \\ 0 \end{matrix} \quad (5.6.62)$$

$$R_0^* = \begin{matrix} R_0 & 0 \\ 0 & R_4 \end{matrix} \quad (5.6.63)$$

由量测方程式(5.6.52)看出, 式中没有量测噪声, 这使最优滤波增益阵 $K^*(k+1|k+1)$ 的表达式(5.6.57)右边求逆矩阵中没有 R_2^* 这一项。在求最优滤波时, 已假定 R_2^* 是正定的, 以保证上述逆矩阵存在, 而现在没有 R_2^* 这一项, 无法保证逆矩阵存在, 可能出现病态。为回避这一病态, 我们可采用等效量测方程法。

2. 等效量测方程法

首先, 要构造一个等效的量测方程。为此, 把式(5.6.45)和(5.6.43)代入式(5.6.44), 得到

$$y(k+1) = (k+1)x(k+1) + (k+1) \\ = (k+1)[(k+1, k)x(k) + (k)] + (k+1, k)(k) + e(k) \quad (5.6.64)$$

由上式减去

$$(k+1, k)y(k) = (k+1, k)(k)x(k) + (k+1, k)(k) \quad (5.6.65)$$

得到 $z(k)$ 的表达式为

$$z(k) = y(k+1) - (k+1, k)y(k) \\ = [(k+1)(k+1, k) - (k+1, k)(k)]x(k) + [(k+1)(k) + e(k)] \\ = \underline{\omega}(k)x(k) + \underline{\omega}(k) \quad (5.6.66)$$

式中:

$$\underline{\omega}(k) = (k+1)(k+1, k) - (k+1, k)(k) \quad (5.6.67)$$

$$\underline{\omega}(k) = (k+1)(k) + e(k) \quad (5.6.68)$$

式(5.6.66)是一个等效的量测方程, 等效噪声 $\underline{\omega}(k)$ 是白噪声, 与模型噪声 (k) 是相关的。有关参数由式(5.6.68)得到

$$E\{\underline{\omega}(k)\} = 0 \quad (5.6.69)$$

$$R_2(k) = (k+1)R_1(k)^T(k+1) + R_3(k) \quad (5.6.70)$$

$$R_{12}(k) = E\{(k)\underline{\omega}^T(k)\} \\ = R_1^T(k+1) \quad (5.6.71)$$

对等效随机状态方程式(5.6.43)和(5.6.66), 由定理 5.6.3 得到最优滤波递推方程

组为

$$\hat{x}(k+1|k+1) = \hat{x}(k|k) + K(k+1|k+1)[z(k+1) - \hat{z}(k|k)] \quad (5.6.72)$$

$$K(k+1|k+1) = [P(k+1|k) - ^T + R_{12}(k)][-P(k+1|k) - ^T + R_2]^{-1} \quad (5.6.73)$$

$$P(k+1|k) = P(k|k) - ^T + R_1 \quad (5.6.74)$$

$$P(k+1|k+1) = [I - K(k+1|k+1)]P(k+1|k)[I - K(k+1|k+1) + 1)]^T + K(k+1|k+1)R_2K^T(k+1|k+1) \quad (5.6.75)$$

$$\hat{x}(k_0|k_0) = m_0 \quad (5.6.76)$$

$$P(k_0|k_0) = R_0 \quad (5.6.77)$$

把上述结果总结为如下定理形式:

定理 5.6.4 考虑离散时间随机状态方程式(5.6.43),(5.6.44)和量测噪声 (k) 成形滤波器式(5.6.45), 给定条件和假设如前所述, 则最优滤波可通过等效量测方程法构造状态方程式(5.6.43)和(5.6.66), 由式(5.6.72) ~ (5.6.77) 决定, 有关变量和参数由式(5.6.67) ~ (5.6.71) 给出。

5.7 最优估计的几个问题

在前几节中主要介绍了最优估计的基本理论, 这一节中进一步讨论最优估计的稳定性、模型误差分析和发散等问题。讨论中我们略去繁杂的证明, 主要给出有关基本概念和重要结果。

5.7.1 稳定性问题

稳定性问题是控制系统、最优估计的基本问题, 是系统首要必须保证的性质。这里我们只是针对最优估计的稳定性问题加以讨论。由于论述最优估计的渐近性质需要较多的随机过程的概念, 存在更多的困难, 因此, 一般从估计误差方差阵着手讨论, 因为估计误差方差阵是确定性函数, 相对简单一些。下面以最优滤波为例讨论最优估计的稳定性概念。

1. 滤波的稳定性概念

在进行滤波时, 需要预先给定状态初值 $\hat{x}(k_0)$ 和方差阵 $P(k_0)$, 这些初值与应当取的最优初值 $E\{x(k_0)\}$ 和方差阵 $R(k_0)$ 难免有一定的误差, 这些初值误差对最优滤波有什么影响呢? 为保证滤波值与最优值相差任意小, 是否需要初值误差充分小, 或在时间充分长之后, 初值误差的影响可以忽略? 这就涉及到滤波的稳定性概念。

我们先给出一般系统稳定性定义。考虑系统

$$x(k+1) = (k+1, k)x(k) + v(k) \quad (5.7.1)$$

式中 $(k+1, k)$ 为可逆转移矩阵。

定义 5.7.1 对系统式(5.7.1)

(1) 如果存在常数 $C > 0$, 使得对所有 $k \geq 0$,

$$(k, k_0) \leq C \quad (5.7.2)$$

则称系统式(5.7.1)为稳定的。

(2) 如果当 $k \rightarrow \infty$ 时,

$$\|x(k, k_0)\| \rightarrow 0 \quad (5.7.3)$$

则称系统式(5.7.1)为渐近稳定的。

(3) 如果存在常数 $C_1 > 0$ 和 $C_2 > 0$, 使得对所有 $k \geq 0$ 时,

$$\|x(k, 1)\| \leq C_2 e^{-C_1(k-1)} \quad (5.7.4)$$

则称系统式(5.7.1)为一致渐近稳定的。

由上述定义可以看出, 一致渐近稳定必然渐近稳定, 渐近稳定必然稳定。

根据上述定义, 我们考察初值误差对系统的影响。设从两个初值 $x(k_0)$ 和 $\tilde{x}(k_0)$ 出发, 分别得到系统式(5.7.1)的两个解

$$x(k) = \Phi(k, k_0)x(k_0) + \sum_{l=1}^k \Phi(k, l)v(1-l) \quad (5.7.5)$$

$$\tilde{x}(k) = \Phi(k, k_0)\tilde{x}(k_0) + \sum_{l=1}^k \Phi(k, l)v(1-l) \quad (5.7.6)$$

为分析初值不同对系统的影响, 由式(5.7.6)减去式(5.7.5), 得到

$$x(k) - \tilde{x}(k) = \Phi(k, k_0)[x(k_0) - \tilde{x}(k_0)] \quad (5.7.7)$$

上式两边取范数, 并由范数性质, 有

$$\|x(k) - \tilde{x}(k)\| = \|\Phi(k, k_0)\| \|x(k_0) - \tilde{x}(k_0)\| \quad (5.7.8)$$

把上式对照定义 5.7.1, 可得到如下结论:

(1) 如果系统式(5.7.1)是稳定的, 即 $\|\Phi(k, k_0)\| \leq C$, 那么任意给定 $\epsilon > 0$, 只要初值误差 $\|x(k_0) - \tilde{x}(k_0)\| < \frac{\epsilon}{C}$, 就可保证 $\|x(k) - \tilde{x}(k)\| < \epsilon$, 即可使最后解的误差在给定范围内。

(2) 如果系统式(5.7.1)是渐近稳定的, 即当 $k \rightarrow \infty$ 时, $\|\Phi(k, k_0)\| \rightarrow 0$, 则对任意初值 $x(k_0)$ 和 $\tilde{x}(k_0)$, 都有 $\|x(k) - \tilde{x}(k)\| \rightarrow 0$, 也就是 k 充分大之后初值影响消失。

(3) 如果系统式(5.7.1)一致渐近稳定, 那么设存在正常数 r , 使得对所有 $k \geq 0$ 时, 有 $\|v(k)\| \leq r$ 和 $\|x(k_0)\| \leq r$, 则有

$$\begin{aligned} \|x(k)\| &= \|\Phi(k, k_0)x(k_0) + \sum_{l=1}^k \Phi(k, l)v(1-l)\| \\ &\leq C_2 e^{-C_1(k-k_0)} r + C_2 r \sum_{l=1}^k e^{-C_1(k-l-1)} \\ &= \frac{C_2 r}{1 - e^{-C_1}} \end{aligned} \quad (5.7.9)$$

这表明, 如果系统输入是有界的, 则系统输出也是有界的。

2. 滤波误差方差阵的上下界

考虑离散时间随机状态方程

$$x(k+1) = \Phi(k+1, k)x(k) + \Gamma(k)u(k) \quad (5.7.10)$$

$$y(k) = C(k)x(k) + d(k) \quad (5.7.11)$$

给定条件和假设如前所述。根据线性系统理论, 对系统式(5.7.10)和(5.7.11)完全能控制

和完全能观测的充分和必要条件分别为

$$C(k-N+1,k)=\sum_{i=k-N+1}^k (k,i)R_1(i-1)^T(k,i)>0 \tag{5.7.12}$$

$$O(k-N+1,k)=\sum_{j=k-N+1}^k {}^T(j,k) {}^T(j)R_2^{-1}(j) (j) (j,k)>0 \tag{5.7.13}$$

由式(5.7.12)和(5.7.13)可得如下定义:

定义 5.7.2

(1) 对系统式(5.7.10)和(5.7.11), 如果存在正整数 N 和 $\epsilon > 0, \delta > 0$, 使得对所有 $k \geq N$, 有

$$I - C(k-N+1,k)I \tag{5.7.14}$$

则称系统式(5.7.10)和(5.7.11)为一致完全能控制的。

(2) 对系统式(5.7.10)和(5.7.11), 如果存在正整数 N 和 $\epsilon > 0, \delta > 0$, 使得对所有 $k \geq N$, 有

$$I - O(k-N+1,k)I \tag{5.7.15}$$

则称系统式(5.7.10)和(5.7.11)为一致完全能观测的。

在满足定义 5.7.2 条件下, 即对一致完全能控制和一致完全能观测的线性系统, 可以证明, 它的最优线性滤波误差的方差阵 $P(k|k)$, 在 $k \geq N$ 时满足不等式

$$\frac{1}{1+(n-k)^T}I - P(k|k) \leq \frac{1+(n-k)^T}{1+(n-k)^T}I \tag{5.7.16}$$

且有 $P(k|k) > 0$ 。也就是说, 最优线性滤波误差的方差阵 $P(k|k)$ 对 k 有一致的有限上界和正定的下界。

3. 最优滤波的稳定性定理

按定理 5.4.2, 对系统式(5.7.10)和(5.7.11), 最优滤波递推公式为

$$\begin{aligned} \hat{x}(k+1|k+1) &= (k+1,k) \hat{x}(k|k) + K(k+1|k+1)[y(k+1) \\ &\quad - (k+1,k) \hat{x}(k|k)] \\ &= [I - K(k+1|k+1)(k+1,k)] (k+1,k) \hat{x}(k|k) \\ &\quad + K(k+1|k+1)y(k+1) \\ &= (k+1,k) \hat{x}(k|k) + K(k+1|k+1)y(k+1) \end{aligned} \tag{5.7.17}$$

式中已令

$$(k+1,k) = [I - K(k+1|k+1)(k+1,k)] (k+1,k) \tag{5.7.18}$$

可把式(5.7.17)看做一个线性系统, $K(k+1|k+1)y(k+1)$ 为系统的输入项, 而 $(k+1,k)$ 为系统的状态转移阵。

在上述论述基础上, 可得到如下定理:

定理 5.7.1(滤波稳定性定理) 如果系统式(5.7.10)和(5.7.11)为一致完全能控制和一致完全能观测的, 则它的最优线性滤波系统式(5.7.17)为一致渐近稳定的, 即存在常数 $C_1 > 0, C_2 > 0$, 使得对所有 $k \geq 1, 0$, 有

$$(k,1) \leq C_2 e^{-C_1(k-1)} \tag{5.7.19}$$

由定理 5.7.1 可得到如下结论:

(1) 对于一致完全能控制和一致完全能观测的线性系统, 当时间充分长之后, 它的最优线性滤波值将(渐近地)不依赖于滤波初值的选取, 并且有界的量测输入必导致有界的滤波输出。

(2) 根据定理 5.7.1, 如果从两个不同的初始滤波误差的方差阵 $P^{(1)}(k_0|k_0)$ 和 $P^{(2)}(k_0|k_0)$ 出发, 分别得到第 k 时刻的滤波误差的方差阵 $P^{(1)}(k|k)$ 和 $P^{(2)}(k|k)$, 则存在常数 $C_3 > 0$ 和 $C_4 > 0$, 使得对所有 $k \geq 0$, 有

$$P^{(2)}(k|k) - P^{(1)}(k|k) \leq C_4 e^{-C_3(k-1)} [P^{(2)}(k_0|k_0) - P^{(1)}(k_0|k_0)] \tag{5.7.20}$$

这表明, 当 $k \rightarrow \infty$ 时,

$$P^{(2)}(k|k) - P^{(1)}(k|k) \rightarrow 0 \tag{5.7.21}$$

也就是, 当时间充分长之后, 它的滤波误差方差阵 $P(k|k)$ (从而增益矩阵) 也将(渐近地)不依赖于初始方差阵的选取。

4. 定常系统最优滤波的稳定性

如果定理 5.7.1 所述系统是定常的, 即 A, B, R_1 和 R_2 都为常阵, 并有 $R_1 > 0$ 和 $R_2 > 0$, 则由其完全能控制和完全能观测可导出其为一致完全能控制和一致完全能观测的, 即一致性总是满足的。进一步可得到, 定常系统一致完全能控制和一致完全能观测的充分和必要条件分别为

$$\sum_{i=0}^{n-1} [A^i B]^T > 0 \tag{5.7.22}$$

$$\sum_{i=0}^{n-1} [A^i]^T C^T C > 0 \tag{5.7.23}$$

式中: n 为状态维数。

可以证明, 对于完全能控制和完全能观测的定常线性系统, 必然存在一个唯一的正定阵 P , 从任何初始方差阵 $P(k_0|k_0)$ 出发, 当 $k \rightarrow \infty$ 时, 恒有 $P(k|k) \rightarrow P$, 同时, $P(k+1|k) \rightarrow M$, $K(k+1|k+1) \rightarrow K$, 即达到定态, 其中 M 和 K 都为常阵。

可以通过下述方法计算定态值 M, K 和 P 。重写有关最优滤波公式

$$K(k+1|k+1) = P(k+1|k)^T [P(k+1|k)^T + R_2]^{-1} \tag{5.7.24}$$

$$P(k+1|k) = P(k|k)^T + R_1 \tag{5.7.25}$$

$$P(k|k) = P(k|k-1) - P(k|k-1)^T [P(k|k-1)^T + R_2]^{-1} P(k|k-1) \tag{5.7.26}$$

式(5.7.26)是把式(5.4.36)代入式(5.4.38)得到的。把式(5.7.26)代入式(5.7.25)得到

$$P(k+1|k) = \{P(k|k-1) - P(k|k-1)^T [P(k|k-1)^T + R_2]^{-1} P(k|k-1) + R_2\}^{-1} P(k|k-1)^T + R_1 \tag{5.7.27}$$

上式称为里卡蒂差分方程。

由于当 $k \rightarrow \infty$ 时, $P(k+1|k) \rightarrow M$, 里卡蒂差分方程式(5.7.27)退化为里卡蒂代数方程, 即

$$M = \{M - M^T [M^T + R_2]^{-1} M\}^T + R_1 \tag{5.7.28}$$

求解式(5.7.28)得到 M , 再把 M 代入式(5.7.24)和(5.7.26)定态式, 得到 K 和 P

$$K = M^{-T} [M^{-T} + R_2]^{-1} \quad (5.7.29)$$

$$P = M - M^{-T} [M^{-T} + R_2]^{-1} M \quad (5.7.30)$$

5.7.2 模型误差分析

前面在正确模型条件下讨论了最优滤波稳定性问题, 现在来分析模型不准时对最优滤波的影响。一般说来, 实际工程建模时, 由于对实际问题认识不全面或为了简化计算而做某些近似, 另外, 对噪声的先验统计量很难精确了解, 因此, 在实际方程和建模模型之间存在误差是正常的。

为简化分析, 我们只考虑状态初值、滤波误差方差初值、模型噪声和量测噪声存在误差, 用 $x(k_0|k_0)$, $P(k_0|k_0)$, $R_1(k)$ 和 $R_2(k)$ 表示模型值, 用 $\hat{x}(k_0|k_0)$, $\hat{P}(k_0|k_0)$, $\hat{R}_1(k)$ 和 $\hat{R}_2(k)$ 表示实际值。

建模得到的系统模型和最优滤波公式为

$$x(k+1) = Ax(k) + B(k)u(k) \quad (5.7.31)$$

$$y(k) = C(k)x(k) + d(k) \quad (5.7.32)$$

$$\hat{x}(k+1|k+1) = \hat{x}(k|k) + K(k+1|k+1)[y(k+1) - C(k)\hat{x}(k|k)] \quad (5.7.33)$$

$$K(k+1|k+1) = P(k+1|k)^T [P(k+1|k)^T + R_2]^{-1} \quad (5.7.34)$$

$$P(k+1|k) = P(k|k)^T + R_1(k) \quad (5.7.35)$$

$$P(k+1|k+1) = [I - K(k+1|k+1)]P(k+1|k)[I - K(k+1|k+1)]^T + K(k+1|k+1)R_2(k)K(k+1|k+1) \quad (5.7.36)$$

为考虑模型误差对滤波的影响, 考虑下列实际值与滤波值之差:

$$\tilde{x}(k|k) = x(k) - \hat{x}(k|k)$$

$$\tilde{x}(k+1|k) = x(k+1) - \hat{x}(k+1|k)$$

$$P^a(k|k) = E\{\tilde{x}(k|k)\tilde{x}^T(k|k)\}$$

$$P^a(k+1|k) = E\{\tilde{x}(k+1|k)\tilde{x}^T(k+1|k)\} \quad (5.7.37)$$

经计算得到

$$P^a(k+1|k) = P^a(k|k)^T + R_1(k) \quad (5.7.38)$$

$$P^a(k+1|k+1) = [I - K(k+1|k+1)]P^a(k+1|k)[I - K(k+1|k+1)]^T + K(k+1|k+1)R_2(k)K^T(k+1|k+1) \quad (5.7.39)$$

考虑差值, 得到

$$\begin{aligned} P(k|k) &= P(k|k) - P^a(k|k) \\ &= [I - K(k|k)]P(k|k-1)[I - K(k|k)]^T + K(k|k)[R_2(k) - \hat{R}_2(k)]K(k|k) \end{aligned} \quad (5.7.40)$$

$$\begin{aligned} P(k|k-1) &= P(k|k-1) - P^a(k|k-1) \\ &= P(k-1|k-1)^T + [R_1(k-1) - \hat{R}_1(k-1)] \end{aligned} \quad (5.7.41)$$

由上两式可以看出, 如果选取 $R_1(k) = \hat{R}_1(k)$, $R_2(k) = \hat{R}_2(k)$, 再使 $P(k-1|k-1) = 0$, 即 $P(k-1|k-1) = P^a(k-1|k-1)$, 必得 $P(k|k-1) = 0$, 即 $P(k|k-1) = P^a(k|k-1)$, 进而 $P(k|k) = P^a(k|k)$ 。因此, 由上两式递推, 可得到下列两个结果:

(1) 只在状态初值、滤波误差方差阵初值、模型噪声和量测噪声方差阵有误差的条件下, 如果选取 $P(k_0|k_0) = P^a(k_0|k_0)$, 再对所有 k 选取 $R_1(k) = R_1(k), R_2(k) = R_2(k)$, 则对所有 k , 必有 $P(k|k) = P^a(k|k)$ 。

(2) 在同样条件下, 如果选取 $P(k_0|k_0) = P(k_0|k_0)$, 再对所有 k 选取 $R_1(k) = R_1(k), R_2(k) = R_2(k)$, 则对所有 k , 必有 $P(k|k) = P^a(k|k)$ 。

以上两个结果对解决实际问题有指导作用, 我们虽然不能确定 $\hat{x}(k_0|k_0), P^a(k_0|k_0), R_1(k)$ 和 $R_2(k)$ 的准确值, 但可采用它们的保险值, 限制滤波误差方差阵真值的大小。如果滤波效果不满意, 即 $P(k|k)$ 过大, 可适当往下调整 $P(k_0|k_0), R_1(k)$ 和 $R_2(k)$ 的值, 以使所设计的滤波器在规定误差范围内良好地工作。

在同样条件下, 我们可进一步得到第三个结果。

(3) 在同样条件下, 并设系统是一致完全能控制的和一致完全能观测的, 如果选取 $P(k_0|k_0) = P^a(k_0|k_0), R_1(k) = R_1(k), R_2(k) = R_2(k)$, 则 $P^a(k|k)$ 有一致的上界和下界, 即存在某常数 $r > 0$, 使对所有 k , 有

$$P^a(k|k) \leq P(k|k) \leq rI \tag{5.7.42}$$

由这个结果可以看出, 如果由于模型误差而使 $P(k|k)$ 无上界而发散, 可通过适当选取 $R_1(k)$ 和 $R_2(k)$, 使建模的系统具有一致完全能控制性和一致完全能观测性, 从而使 $P(k|k)$ 具有上界, 能稳定工作。

5.7.3 发散问题

随着处理大量数据, 最优滤波给出状态的最优估计, 这似乎已足够解决工程技术提出的各种滤波问题。但在实际应用中会发现, 由最优滤波得到的状态估计可能与真实状态存在误差, 并且这种误差方差可能大大超过理论计算允许的范围。即使理论计算的误差方差变得很小, 而实际误差可能变成无穷大。这种现象称为滤波的发散现象。最早, 卡尔曼本人就预见到这种发散现象的可能性, 在空间导航和轨道确定的实际应用中, 都发现过这种问题。

出现发散现象的主要原因为:

(1) 由于对实际工程问题了解不精确, 使得对随机动力学系统(包括噪声统计特性)的建模不准确或不符合实际情况。或对实际工程问题了解足够精确, 但所得数学模型较复杂(如非线性、维数过高), 在简化数学模型时(如线性化、降阶)带来不精确性。这样引起的发散称为滤波发散。

(2) 最优滤波的递推计算在有限字长的计算机上实现时, 每步都有舍入误差, 从而可能使计算的滤波误差的方差阵逐渐失去正定性、对称性, 造成发散。这种发散称为计算发散。

下面举一个简单的例子, 说明建模不准确而引起的滤波发散。

例 5.7.1 设实际随机动力学系统为

$$\begin{aligned} x(k+1) &= x(k) + a \\ &= x(k_0) + (k+1)a \quad (k_0 = 0) \end{aligned} \tag{5.7.43}$$

式中: a 为常数。建模时, 错误地忽略了 a , 取数学模型为

$$x(k+1) = x(k) \quad (5.7.44)$$

量测模型为

$$y(k) = x(k) + v(k) \quad (5.7.45)$$

式中: $v(k)$ 为离散高斯 $N(0, \sigma^2)$ 白噪声。

把现系统式(5.7.44), (5.7.45)和滤波方程式(5.7.10), (5.7.11)对照可知,

$$F = 1 \quad G = 1 \quad R_1 = 0 \quad R_2 = \sigma^2 \quad (5.7.46)$$

并设初值为 $\hat{x}(k_0|k_0) = m_0 = 0$, $P(k_0|k_0) = R_0 = \frac{\sigma^2}{2}$ 。按照定理 5.4.2, 得到滤波方程

$$\hat{x}(k+1|k) = \hat{x}(k|k) = \hat{x}(k|k) \quad (5.7.47)$$

$$P(k+1|k) = P(k|k)^T + R_1 = P(k|k) \quad (5.7.48)$$

$$\begin{aligned} P(k|k) &= [P^{-1}(k|k-1) + R_2^{-1}]^{-1} \\ &= [P^{-1}(k-1|k-1) + \frac{1}{\sigma^2}]^{-1} \\ &= [P^{-1}(k_0|k_0) + \frac{k}{\sigma^2}]^{-1} \\ &= \frac{\sigma^2}{k} \end{aligned} \quad (5.7.49)$$

$$\begin{aligned} K(k+1|k+1) &= P(k+1|k)^T [P(k+1|k)^T + R_2]^{-1} \\ &= \frac{\sigma^2}{k} \frac{\sigma^2}{\sigma^2 + \sigma^2} = \frac{1}{2} \\ &= \frac{1}{k+1} \end{aligned} \quad (5.7.50)$$

$$\begin{aligned} \hat{x}(k+1|k+1) &= \hat{x}(k|k) + K(k+1|k+1)[y(k+1) - \hat{x}(k|k)] \\ &= [1 - K(k+1|k+1)] \hat{x}(k|k) + K(k+1|k+1)y(k+1) \\ &= \frac{k}{k+1} \hat{x}(k|k) + \frac{1}{k+1} y(k+1) \\ &= \frac{k-1}{k+1} \hat{x}(k-1|k-1) + \frac{1}{k+1} [y(k) + y(k+1)] \\ &= \frac{1}{k+1} \sum_{i=1}^{k+1} y(i) \end{aligned} \quad (5.7.51)$$

式中: $y(k)$ 是对真实随机动力学系统的量测。而真实量测为

$$\begin{aligned} y(k) &= x(k) + v(k) \\ &= x(k_0) + ka + v(k) \end{aligned} \quad (5.7.52)$$

把式(5.7.52)代入式(5.7.51), 得到滤波为

$$\hat{x}(k+1|k+1) = x(k_0) + \frac{k+2}{2}a + \frac{1}{k+1} \sum_{i=1}^{k+1} v(i) \quad (5.7.53)$$

$$\hat{x}(k|k) = x(k_0) + \frac{k+1}{2}a + \frac{1}{k} \sum_{i=1}^k v(i) \quad (5.7.54)$$

真正状态式(5.7.43)与滤波值 $\hat{x}(k|k)$ 之间的差为

$$\tilde{x}(k|k) = x(k) - \hat{x}(k|k)$$

$$= \frac{k-1}{2}a - \frac{1}{k} \sum_{i=1}^k (i) \quad (5.7.55)$$

滤波误差的方差和均方误差分别为

$$P(k) = \frac{1}{k} \quad (5.7.56)$$

$$E\{\tilde{x}^2(k)\} = \frac{(k-1)^2}{4}a^2 + \frac{1}{k} \quad (5.7.57)$$

虽然当 $k \rightarrow \infty$ 时, $P(k) \rightarrow 0$, 但 $E\{\tilde{x}^2(k)\} \rightarrow \infty$, 这就出现了发散现象。

造成上述发散的原因, 似乎是建模时忽略了常量 a , 错误地把式(5.7.43) 写为式(5.7.44), 好像只要在模型中加入一常量即可。但实际上, 即使在模型中加入一常量 b , 它不等于 a , 仍然会出现发散现象。设模型取为

$$x(k+1) = x(k) + b \quad (b \neq a) \quad (5.7.58)$$

其他条件不变, 则得

$$\tilde{x}(k) = \frac{k-1}{2}(a-b) - \frac{1}{k} \sum_{i=1}^k (i) \quad (i) \quad (5.7.59)$$

$$E\{\tilde{x}^2(k)\} = \frac{(k-1)^2}{4}(a-b)^2 + \frac{1}{k} \quad (5.7.60)$$

当 $k \rightarrow \infty$ 时, 仍有 $E\{\tilde{x}^2(k)\} \rightarrow \infty$ 。

上例发散的直接原因在于, $P(k)$ 随时间 k 的增长趋于零的速度太快, 从而使滤波增益 $K(k)$ 随 k 的增长而迅速变小, 于是新的量测数据 $y(k)$ 对滤波的作用越来越小, 最后失去作用, 因此使滤波与量测相分离, 形成数据饱和, 以致引起发散。

克服上例发散现象的简单方法是, 限定 $K(k)$ 的减小, 取之为

$$K^*(k) = \begin{cases} \frac{1}{k} & (k \leq M) \\ \frac{1}{M} & (k > M) \end{cases} \quad (5.7.61)$$

即在 M 个量测后, $K^*(k)$ 不再减小, 取为固定值 $\frac{1}{M}$ 。这时, 当 $k \rightarrow \infty$ 时, 滤波均方误差为

$$E\{\tilde{x}^2(k)\}^* = (M-1)^2a^2 + \frac{1}{2M-1} \quad (5.7.62)$$

M 的选取可以这样考虑: 为了真正起到滤波作用, 必须使 $E\{\tilde{x}^2(k)\}^*$ 小于量测噪声的方差(均方差) σ^2 , 否则, 不如直接把量测 $y(k)$ 作为 $x(k)$ 的估计, 于是得到

$$1 < M < 1 + \frac{1}{a} \quad (5.7.63)$$

M 取为大于 1, 且小于 $1 + \frac{1}{a}$ 的正整数。

针对滤波发散和计算发散, 可以提出许多种克服发散的方法, 当然, 这些方法都以次优滤波来取得滤波器的稳定工作。其中一类方法的基本思路是限制滤波增益 $K(k)$ 的减小, 以防止滤波脱离量测。这类方法包括: 直接增加滤波增益矩阵, 如例 5.7.1 所述; 限制滤波误差方差阵 $P(k)$ 的增加; 人为地增加模型噪声方差。

另一类方法的基本思路是: 在滤波时, 设法加大新量测数据的作用, 相对地减少早期

数据的作用。“衰减记忆滤波法”就是这种方法,随着时间的增加,过老的量测数据逐渐失去作用。“限定记忆滤波法”也属于这种方法,它在滤波时,只利用离现时刻 k 最近的 N 个量测数据 $y(k), y(k-1), \dots, y(k-N+2), y(k-N+1)$, 而把以前的数据全部甩掉,其中 N 为预先设置的记忆长度。

再一类是针对滤波误差方差阵 $P(k|k)$ 随时间增加可能失去对称性和非负定性引起计算发散而采取的克服方法。“协方差平方根滤波法”就是这种方法,由于对任何矩阵 S , SS^T 总是对称的和非负定的,为防止 $P(k|k), P(k+1|k)$ 失去对称性和非负定性,可把它们分解为 $P(k|k) = S(k|k)S^T(k|k), P(k+1|k) = S(k+1|k)S^T(k+1|k)$, 在计算中以 $S(k|k)$ 的递推式代替 $P(k|k)$ 的递推式,以保证 $P(k|k)$ 的对称性和非负定性。

还有一类是自适应滤波方法,也是克服发散的有效措施。自适应滤波在利用量测数据进行滤波的同时,不断地对系统模型和噪声特性进行估计和修正,以逐步缩小实际的滤波误差,并使滤波器稳定工作。

5.8 连续时间随机状态模型的状态估计

本节讨论连续时间随机状态模型的状态估计,包括滤波、预测和平滑问题。我们首先给出连续时间随机状态模型,然后重点讨论最优滤波,因为滤波是状态估计的基础,最后简单介绍预测和平滑的结果。

目前,一般有三种方法求解连续时间随机状态模型的状态估计:第一种方法是把离散时间随机状态模型的状态估计公式,通过使采样时间 Δt 趋近于零,取极限而得到连续时间随机状态模型的状态估计公式,它是一种间接的方法,推导比较简单;第二种方法是利用连续时间随机状态模型的状态估计与连续时间确定性状态模型的最优控制的对称性(等价性),通过已知确定性最优控制导出随机状态估计,它也是一种间接的方法,其中推导确定性最优控制的结果也同样不容易,甚至也是由离散化取极限得到的;第三种方法是利用随机过程和系统知识直接导出状态估计的结果。

我们采用第三种方法求解状态估计。具体推导中,先定义量测模型的新息过程,并导出一些必要的结果,再以新息过程为出发点,导出状态估计的结果。

5.8.1 连续时间随机状态模型和新息过程

1. 连续时间随机状态模型

3.4.1 小节已列出连续时间随机状态模型,为进行状态估计,所用模型为

$$dx(t) = Ax(t)dt + dw(t) \tag{5.8.1}$$

$$dy(t) = Hx(t)dt + dv(t) \tag{5.8.2}$$

式(5.8.1)称为对象模型,模型中略去了控制项;式(5.8.2)称为量测模型。式中: $x(t)$ 为 $n \times 1$ 状态向量; $y(t)$ 为 $m \times 1$ 量测向量; t 为连续时间; A 和 H 分别为 $n \times n$ 和 $m \times n$ 系数矩阵,它们可以是时不变或时变的; $dw(t)$ 为 $n \times 1$ 模型噪声向量; $dv(t)$ 为 $m \times 1$ 量测噪声向量。

为进行状态估计,模型式(5.8.1)和(5.8.2)的高斯假设可以放松。假设状态初值

$x(t_0)$ 为均值为 m_0 和方差为 R_0 的随机过程, $dw(t)$ 为均值为零和方差为 $R_1 dt$ 的独立增量过程, $dv(t)$ 为均值为零和方差为 $R_2 dt$ 的独立增量过程, 并且它们是两两互相独立的, 即

$$E\{x(t_0) dw(t)\} = 0 \qquad (t \geq t_0) \tag{5.8.3}$$

$$E\{x(t_0) dv(t)\} = 0 \qquad (t \geq t_0) \tag{5.8.4}$$

$$E\{dw(s) dv(t)\} = 0 \tag{5.8.5}$$

2. 新息过程

为表示简便, 引入一些符号。设 H^y_t 表示 $\{y(t)\}$ 所处希尔伯特空间, P_t 表示在 H^y_t 上的正交投影, 则 $\{y(t)\}$ 对 $x(s)$ 的状态估计表示为 $\hat{x}(s|t) = \hat{x}_{s|t} = P_t x(s) \in H^y_t$ 。估计误差 $\tilde{x}(s|t) = x(s) - \hat{x}(s|t)$ 与 H^y_t 互相独立, 几何上相互垂直, 表示为 $\tilde{x}(s|t) \perp H^y_t$ 。

为求最优滤波, 对量测模型式(5.8.2)定义一个新息过程 $d(t) \in H^y_t$, 表示为

$$d(t) = dy(t) - H \hat{x}(t|t) dt \tag{5.8.6}$$

把式(5.8.2)代入上式, 得到 $d(t)$ 的另一种形式:

$$d(t) = H \tilde{x}(t|t) dt + dv(t) \tag{5.8.7}$$

对新息过程 $d(t)$ 有如下定理:

定理 5.8.1 由式(5.8.6)定义的新息过程 $d(t)$ 具有与 $dv(t)$ 相同的统计特性, 即 $d(t)$ 是一个均值为零和方差为 $R_2 dt$ 的独立增量过程。

证明 首先证明 $d(t)$ 是一个独立增量过程, 只要证明 $(t) - (s)$ 与 H^y_s 互相独立, 其中 $t > s$ 。设 $z(s) \in H^y_s$, 则有

$$\begin{aligned} [z(s), (t) - (s)] &= E \{ z(s) \int_s^t H \tilde{x}(u|u) du + \int_s^t dv(u) \} \\ &= \int_s^t H E \{ z(s) \tilde{x}(u|u) \} du + \int_s^t E \{ z(s) dv(u) \} \\ &= 0 \end{aligned} \tag{5.8.8}$$

上式表明 $[(t) - (s)] \perp H^y_s$, $d(t)$ 是一个独立增量过程。

显然, 由式(5.8.7)可知, $d(t)$ 的均值为零。

最后求 $d(t)$ 的方差阵。按方差定义, 并把式(5.8.7)代入, 得到

$$\begin{aligned} R(d(t)) &= E\{d(t)d^T(t)\} \\ &= E\{[H \tilde{x}(t|t) dt + dv(t)][H \tilde{x}(t|t) dt + dv(t)]^T\} \\ &= E\{H P(t|t) H^T dt^2\} + E\{H \tilde{x}(t|t) dt dv(t)\} \\ &\quad + E\{dv(t) dt \tilde{x}^T(t|t) H^T\} + E\{dv(t) dv^T(t)\} \\ &= R_2 dt \end{aligned} \tag{5.8.9}$$

式中第三式右边第一项中 dt^2 是 dt 的高阶无限小, 因此, 第一项为零; 因为 $\tilde{x}(t|t)$ 与 $dv(t)$ 互相独立, 因此, 第二、三两项均为零; 只有第四项有值, 并为第四式结果。

3. 有关滤波的一个结果

定理 5.8.2 对随机过程 $\{x(t)\}$, 以及均值为零和方差为 $\int_0^t R ds$ 的独立增量过程 $\{d(t)\}$, 必有

$$\hat{x}(t|t) = \int_0^t \frac{d}{ds} E\{x(t) - (s)^T\} R^{-1} d(s) \quad (5.8.10)$$

证明 我们已知 $\hat{x}(t|t) = P(t)x(t) - H(t)$, 因此必存在某个函数 $d(t)$, 使得

$$\hat{x}(t|t) = \int_0^t d(s) d(s) \quad (5.8.11)$$

再任选一个函数 $d(t)$, 使满足 $\int_0^t d(s) d(s) = H(t)$, 则滤波误差 $\tilde{x}(t|t) = [x(t) - \hat{x}(t|t)]$
 $\int_0^t d(s) d(s)$ 。于是有

$$\begin{aligned} E\{x(t) - \int_0^t d(s) d(s)\}^T &= E\{\hat{x}(t|t) - \int_0^t d(s) d(s)\}^T \\ &= E\left\{\int_0^t d(s) d(s) - \int_0^t d(s) d(s)\right\}^T \\ &= \int_0^t \int_0^t d(s) E\{d(s) d^T(s)\} d^T(s) \\ &= \int_0^t d(s) R^{-1}(s) ds \end{aligned} \quad (5.8.12)$$

因为 $d(t)$ 是任选的, 现选为 $d(s) = I_{[0,r]}, r = t$, 则式(5.8.12)变为

$$E\{x(t) - (r)\} = \int_0^r d(s) R ds \quad (5.8.13)$$

$$d(s) = \frac{d}{ds} E\{x(t) - (s)\} R^{-1} \quad (5.8.14)$$

把上式代入式(5.8.11), 就得到式(5.8.10)。

5.8.2 连续时间随机状态模型的最优滤波

定理 5.8.3(卡尔曼滤波定理) 对连续时间随机状态模型式(5.8.1)和(5.8.2), 在给定新息过程式(5.8.6)条件下, 最优滤波满足线性随机方程

$$\begin{aligned} d\hat{x}(t|t) &= [A - P(t)H^T R_2^{-1} H] \hat{x}(t|t) dt + P(t)H^T R_2^{-1} dy(t) \\ &= A \hat{x}(t|t) dt + P(t)H^T R_2^{-1} d(t) \end{aligned} \quad (5.8.15)$$

$$\hat{x}(0|0) = E\{x(0)\} = m_0 \quad (5.8.16)$$

式中 $P(t) = P(t|t) = E\{\tilde{x}(t|t) \tilde{x}^T(t|t)\}$ 为最优滤波误差的方差阵, 它满足里卡蒂方程

$$\frac{dP}{dt} = AP + PA^T + R_1 - PH^T R_2^{-1} HP \quad (5.8.17)$$

$$P(0) = \text{var}(x(0)) = P_0 = R_0 \quad (5.8.18)$$

把滤波方程式(5.8.15)第二式与状态方程式(5.8.1)对照可知, 滤波方程具有与状态方程式(5.8.1)相同的动力学系数矩阵 $A(t)$, 也由一个独立增量过程驱动, 不过用 $P(t)H^T R_2^{-1} d(t)$ 替换了原来的 $dw(t)$ 。 $P(t)H^T R_2^{-1}$ 可用 $K(t)$ 表示, $K(t)$ 称为最优滤波增益。

证明 我们先证明 $\hat{x}(0|0) = m(0) = 0$ 的情况, 然后再把结果推广到不为零的情况。我们分几步进行证明。

(1) 列出 $\hat{x}(t|t)$ 表示式

由定理 5.8.2 中式(5.8.10), 有

$$\hat{x}(t|0) = \int_0^t \frac{d}{ds} E\{x(t) - x(s)\} R_2^{-1} ds \quad (5.8.19)$$

我们假设有线性滤波公式

$$d\hat{x}(t|0) = A\hat{x}(t|0)dt + g(t)dw(t) \quad (5.8.20)$$

式中: $g(t)$ 为待定函数。在假设 $\hat{x}(0|0) = m(0) = 0$ 条件下, 式(5.8.20) 有解

$$\hat{x}(t|0) = \int_0^t \Phi(t, s) g(s) dw(s) \quad (5.8.21)$$

式中 $\Phi(t, s)$ 为 A 的状态转移阵, 对照式(5.8.21)和(5.8.19), 有

$$\Phi(t, s) g(s) = \frac{d}{ds} E\{x(t) - x(s)\} R_2^{-1} \quad (5.8.22)$$

当 $t = s$ 时, 得到

$$g(t) = \frac{d}{dt} E\{x(t) - x(t)\} R_2^{-1} \quad (5.8.23)$$

下面我们利用上式确定系数函数 $g(t)$ 。

(2) 确定系数函数 $g(t)$

设 $t > u$, 式(5.8.1)有解

$$x(t) = \Phi(t, s)x(s) + \int_s^t \Phi(t, \tau)dw(\tau) \quad (5.8.24)$$

由新息过程公式(5.8.7), 得到

$$\tilde{x}(s) = \int_0^s H^T \tilde{x}(u|0) du + v(s) \quad (5.8.25)$$

考虑到 $x(s)$ 与 $v(s)$, $dw(\tau)$ 与 $\tilde{x}(u|0)$, 以及 $dw(\tau)$ 与 $v(s)$ 都是互相独立的, 因此得到

$$\begin{aligned} E\{x(t)^T(s)\} &= E\left[\Phi(t, s)x(s) + \int_s^t \Phi(t, \tau)dw(\tau)\right]^T \int_0^s H^T \tilde{x}(u|0) du \\ &= \int_0^s \Phi(t, s) E\{x(s) \tilde{x}^T(u|0)\} H^T du \\ &= \int_0^s \Phi(t, s) P(u) H^T du \end{aligned} \quad (5.8.26)$$

$$\frac{d}{ds} E\{x(t)^T(s)\} = \Phi(t, s) P(s) H^T \quad (5.8.27)$$

$$\frac{d}{dt} E\{x(t)^T(t)\} = P(t) H^T \quad (5.8.28)$$

把上式代入式(5.8.23), 得到

$$g(t) = P(t) H^T R_2^{-1} \quad (5.8.29)$$

再把上式代入式(5.8.20), 得到

$$d\hat{x}(t|0) = A\hat{x}(t|0)dt + P(t) H^T R_2^{-1} dw(t) \quad (5.8.30)$$

上式就是需要证明的式(5.8.15) 第二式, 把式(5.8.6)代入式(5.8.30) 就得到式(5.8.15) 的第一式。

(3) 初值不为零的情况

当 $\hat{x}(0|0) = E\{x(0)\} = m(0) = m_0 \neq 0$ 时, 有

$$m(t) = Am(t) \quad m(0) = m_0 \quad (5.8.31)$$

其中心过程 $\hat{x}^c(t)$ 满足

$$\dot{\hat{x}}^c(t) = A \hat{x}^c(t) - m(t) \tag{5.8.32}$$

由于 $\hat{x}^c(0) = 0$, 因此 $\hat{x}^c(t)$ 满足式(5.8.15)的条件, 所以 $\hat{x}(t) = \hat{x}^c(t) + m(t)$ 满足同样的方程。

(4) 求解滤波误差 $\tilde{x}(t|t)$ 的方差阵 $P(t)$

把式(5.8.7)代入式(5.8.30), 得到

$$d\hat{x}(t|t) = A \hat{x}(t|t)dt + P(t)H^T R_2^{-1} [H \tilde{x}(t|t)dt + dv(t)] \tag{5.8.33}$$

再由状态方程式(5.8.1)减去上式, 得到

$$d\tilde{x}(t|t) = [A - P(t)H^T R_2^{-1} H] \tilde{x}(t|t)dt + dw(t) - P(t)H^T R_2^{-1} dv(t) \tag{5.8.34}$$

上式是一个状态方程, 用 $\Phi(t, s)$ 表示对应系数矩阵 $[A - P(t)H^T R_2^{-1} H]$ 的状态转移阵, 那么这个方程的解为

$$\tilde{x}(t|t) = \Phi(t, 0)x(0) + \int_0^t \Phi(t, s)dw(s) - \int_0^t \Phi(t, s)P H^T R_2^{-1} dv(s) \tag{5.8.35}$$

考虑到 $x(0)$, $dw(s)$ 和 $dv(s)$ 三个量互相独立, 并且

$$E\{\tilde{x}(t|t)\} = 0 \tag{5.8.36}$$

因此有

$$\begin{aligned} P(t) &= E\{\tilde{x}(t|t) \tilde{x}^T(t|t)\} \\ &= \Phi(t, 0)R_0 \Phi^T(t, 0) + \int_0^t \Phi(t, s)R_1 \Phi^T(t, s)ds \\ &\quad + \int_0^t \Phi(t, s)P H^T R_2^{-1} R_2 R_2^{-1} H P^T(t, s)ds \end{aligned} \tag{5.8.37}$$

把 $P(t)$ 写为导数形式, 将上式对 t 求导, 得到

$$\begin{aligned} \frac{dP(t)}{dt} &= \frac{d}{dt} \Phi(t, 0)R_0 \Phi^T(t, 0) + \Phi(t, 0)R_0 \frac{d}{dt} \Phi^T(t, 0) \\ &\quad + \Phi(t, t)R_1 \Phi^T(t, t) + \int_0^t \frac{d}{dt} \Phi(t, s)R_1 \Phi^T(t, s)ds + \int_0^t \Phi(t, s)R_1 \frac{d}{dt} \Phi^T(t, s)ds \\ &\quad + \Phi(t, t)P H^T R_2^{-1} H P^T(t, t) + \int_0^t \frac{d}{dt} \Phi(t, s)P H^T R_2^{-1} H P^T(t, s)ds \\ &\quad + \int_0^t \Phi(t, s)P H^T R_2^{-1} H P^T(t, s) \frac{d}{dt} \Phi^T(t, s)ds \\ &= [A - P H^T R_2^{-1} H] \Phi(t, 0)R_0 \Phi^T(t, 0) + \Phi(t, 0)R_0 \Phi^T(t, 0) [A - P H^T R_2^{-1} H]^T \\ &\quad + R_1 + [A - P H^T R_2^{-1} H] \int_0^t \Phi(t, s)R_1 \Phi^T(t, s)ds + \int_0^t \Phi(t, s)R_1 \Phi^T(t, s)ds \\ &\quad + [A - P H^T R_2^{-1} H]^T + P H^T R_2^{-1} H P + [A - P H^T R_2^{-1} H] \\ &\quad + \int_0^t \Phi(t, s)P H^T R_2^{-1} H P^T(t, s)ds + \int_0^t \Phi(t, s)P H^T R_2^{-1} H P^T(t, s)ds \\ &\quad + [A - P H^T R_2^{-1} H P]^T \end{aligned} \tag{5.8.38}$$

上式右边共 8 项, 其中第一、四、七各项之和为 $[A - P H^T R_2^{-1} H]P$, 第二、五、八各项之和为 $P[A - P H^T R_2^{-1} H]^T$, 于是最后得到

$$\frac{dP(t)}{dt} = AP(t) + P(t)A^T + R_1 - P(t)H^TR_2^{-1}HP(t) \quad (5.8.39)$$

上式就是式(5.8.17)。

定理 5.8.3 在计算上给出了最优滤波的有效解, 滤波误差方差阵 $P(t)$ 方程式(5.8.17)和(5.8.18)不是随机方程, 函数 $P(t)$ 可以提前计算, 然后求解由观测(量测)过程 $\{y(t)\}$ 驱动的线性时变方程式(5.8.15)和(5.8.16), 得到最优(卡尔曼)滤波 $\{\hat{x}(t|t)\}$ 。

然而, 在计算之前, 我们必须确信, 式(5.8.17)没有多个解, 不会出现无界解, 因为方程中有一个二次项, 有可能有多个解或无界解。幸好, 这种情况不会发生, 只要初值条件是非负定的, 而初值 $P(0) = R[x(0)] = R_0$ 总是非负定的。显然, 其解 $P(t) = E\{\tilde{x}(t|t)\tilde{x}^T(t|t)\}$ 也是非负定的, 并且为有界的, 因为 $P(t) = E\{\tilde{x}(t|t)\tilde{x}^T(t|t)\} = E\{x(t)x^T(t)\}$, 后者有界。

例 5.8.1 (Ornstein-Uhlenbeck 过程的噪声观测) 考虑标量方程

$$dx = -xdt + dw \quad (5.8.40)$$

$$dy = xdt + gdv \quad (5.8.41)$$

式中: dw 和 dv 都为均值为零和方差为 dt 的独立增量过程; g 为系数; 给定状态初值 $E\{x(0)\} = 0$; $R(x(0)) = p_0$ 非负定。试列出卡尔曼滤波方程, 列出并求解里卡蒂方程。

解 按定理 5.8.3 容易列出给定方程式(5.8.40)和(5.8.41)的卡尔曼滤波和里卡蒂方程

$$d\hat{x}(t|t) = [-p(t)g^{-2}]\hat{x}(t|t)dt + p(t)g^{-2}dy(t) \quad \hat{x}(t|0) = 0 \quad (5.8.42)$$

$$\frac{dp(t)}{dt} = 1 - p^2(t)g^{-2} - 2p(t) \quad p(0) = p_0 \quad (5.8.43)$$

式中利用了

$$R_2dt = E\{[g dv]^2\} = g^2dt \quad (5.8.44)$$

求解式(5.8.43), 得到

$$p(t) = p_1 + \frac{(p_1 - p_2)(p_0 - p_1)}{(p_0 - p_2)e^{2t} - (p_0 - p_1)} \quad (5.8.45)$$

式中

$$p_1 = (1 + g^{-2})^{-\frac{1}{2}} \quad p_2 = -g^2(1 + g^{-2}) \quad (5.8.46)$$

由 $p(t)$ 的解式(5.8.45)可知, 当 $t \rightarrow \infty$ 时, $p(t) \rightarrow p_1$, 这表明, 当 t 很大时, 式(5.8.42)几乎变成时不变系统($p(t) \rightarrow p_1$), 这种情况甚至发生在 $\{x(t)\}$ 不稳时, 即 g 为负值。这时, 信号 $x(t)dt$ 最终完全淹没了噪声 $dv(t)$, 使得滤波特别容易。

如果选取 $p_0 = p_1$, 那么里卡蒂方程变为 $\dot{p}(t) = 0$, 使得在所有时刻都有 $p(t) = p_1$ 。这表明: 一方面, 在较长时间进行连续观测时, 观测器记忆 $\{x(t)\}$; 另一方面, 当 $x(t)$ 远离初始位置时, $p_0 = p_1$ 确保消除这种趋向, 使偏离程度(误差)的方差保持 $p(t) = p_1$ 。

例 5.8.2 分析下列三个微分方程

$$(1) \dot{p}_1 = -p_1^2, p_{10} = 1 \quad (5.8.47)$$

$$(2) \dot{p}_2 = -p_2^2, p_{20} = 1 \quad (5.8.48)$$

$$(3) \dot{p}_3 = p_3^2, p_{30} = 1 \quad (5.8.49)$$

是否为满足要求的里卡蒂方程。

解 方程式(1),(2)和(3)的解分别为

$$p_1(t) = (t + \tau^{-2})^{-1} \quad (5.8.50)$$

$$p_2(t) = (t - \tau^{-2})^{-1} \quad (5.8.51)$$

$$p_3(t) = (\tau^{-2} - t)^{-1} \quad (5.8.52)$$

由以上三个方程的解可知:第(1)方程是满足要求的里卡蒂方程;第(2)方程的初值为负值,不满足非负定条件,使 $p(t)$ 有无界解;第(3)方程不满足里卡蒂方程条件。

5.8.3 最优预测

为求解最优预测 $\hat{x}(t + \tau | t)$, $\tau > 0$, 写出式(5.8.1)的解

$$x(t + \tau) = \Phi(t + \tau, t)x(t) + \int_t^{t+\tau} \Phi(t + \tau, s)dw(s) \quad (5.8.53)$$

考虑到 $\int_t^{t+\tau} \Phi(t + \tau, s)dw(s) \sim H_t^y$, 最优预测为

$$\begin{aligned} \hat{x}(t + \tau | t) &= P_t^y x(t + \tau) \\ &= \Phi(t + \tau, t) \hat{x}(t | t) \end{aligned} \quad (5.8.54)$$

上式是 $\hat{x}(t + \tau | t)$ 的一种表示形式。还可把 $\hat{x}(t + \tau | t)$ 表示为微分方程的形式。为此,假设系数矩阵 A 是时不变的,则有

$$\Phi(t + \tau, t) = \Phi(\tau, 0) \quad (5.8.55)$$

由式(5.8.54)有

$$\begin{aligned} d\hat{x}(t + \tau | t) &= d[\Phi(\tau, 0) \hat{x}(t | t)] \\ &= \Phi(\tau, 0) d\hat{x}(t | t) \end{aligned} \quad (5.8.56)$$

由于 $\Phi(\tau, 0)$ 是非奇异的,则有

$$\hat{x}(t | t) = \Phi^{-1}(\tau, 0) \hat{x}(t + \tau | t) \quad (5.8.57)$$

把上式和式(5.8.15)代入式(5.8.56),得到

$$d\hat{x}(t + \tau | t) = \Phi(\tau, 0)A^{-1}\Phi(\tau, 0) \hat{x}(t + \tau | t)dt + \Phi(\tau, 0)PH^T R_z^{-1}d(t) \quad (5.8.58)$$

初值为

$$\hat{x}(\tau | 0) = \Phi(\tau, 0)m_0 \quad (5.8.59)$$

5.8.4 最优平滑

平滑也称内插,最优平滑表示为

$$\hat{x}(t | s) = \hat{x}_{t|s} = P_s^y x(t) \sim H_s^y \quad (s > t) \quad (5.8.60)$$

最优平滑包括三种基本形式,即固定点最优平滑 $\hat{x}(0 | b)$ 、固定区间最优平滑 $\hat{x}(t | b)$ 和固滞后最优平滑 $\hat{x}(t - \tau | b)$, 其中 t 是可调的,而 b 和 τ 是固定不变的。我们通过把滤波作为终值条件可计算 $\hat{x}(t | b)$, 它是计算各种平滑的基础。

定理 5.8.4 考虑连续时间随机状态模型式(5.8.1), (5.8.2) 和新息过程式(5.8.6), 在前面假设和符号条件下, 最优平滑 $\hat{x}(t | b, b > t)$ 满足微分方程

$$\frac{d\hat{x}(t|b)}{dt} = A\hat{x}(t|b)dt + R_1P^{-1}(t)[\hat{x}(t|b) - \hat{x}(t|a)]dt \quad (5.8.61)$$

终值条件为

$$\hat{x}(b|b) = \hat{x}_b \quad (5.8.62)$$

证明 式(5.8.15)对所有 t 成立, 可以写为

$$\begin{aligned} \hat{x}(t|b) &= \int_0^b \frac{d}{ds} E\{x(t)^T(s)\} R_2^{-1} d(s) \\ &= \int_0^t \frac{d}{ds} E\{x(t)^T(s)\} R_2^{-1} d(s) + \int_t^b \frac{d}{ds} E\{x(t)^T(s)\} R_2^{-1} d(s) \\ &= \hat{x}(t|a) + \int_t^b \frac{d}{ds} E\{x(t)^T(s)\} R_2^{-1} d(s) \end{aligned} \quad (5.8.63)$$

式中:

$$\frac{d}{ds} E\{x(t)^T(u)\} = E\{x(t) \tilde{x}^T(u|a)\} H^T \quad (5.8.64)$$

当 $u > t$ 时, 上式不正确, 需换算。为此, 由式(5.8.35), 有

$$\tilde{x}(u|a) = \Phi(u, t) \tilde{x}(t|a) + z \quad (5.8.65)$$

式中 z 与 $\hat{x}(t|a)$ 互相独立。把式(5.8.65)代入式(5.8.64), 得到

$$\begin{aligned} \frac{d}{ds} E\{x(t)^T(u)\} &= E\{x(t) \tilde{x}^T(t|a)\}^T \Phi^T(u, t) H^T \\ &= P(t)^T \Phi^T(u, t) H^T \end{aligned} \quad (5.8.66)$$

把上式代入式(5.8.63), 得到

$$\begin{aligned} \hat{x}(t|b) &= \hat{x}(t|a) + P(t) \int_t^b \Phi^T(s, t) H^T R_2^{-1} d(s) \\ &= \hat{x}(t|a) + P(t) \Phi(t, t) \end{aligned} \quad (5.8.67)$$

式中:

$$\Phi(t) = \int_t^b \Phi^T(s, t) H^T R_2^{-1} d(s) \quad (5.8.68)$$

我们已知 $\Phi(s, t)$ 是系数矩阵 $[A - PH^TR_2^{-1}H]$ 的状态转移阵。对 $\Phi(s, t)$ 有

$$-\frac{d}{dt} \Phi^T(s, t) = -[A - PH^TR_2^{-1}H]^T \Phi^T(s, t) \quad (5.8.69)$$

由上式可知, 可把 $\Phi^T(s, t)$ 看做是系数矩阵 $(-[A - PH^TR_2^{-1}H]^T)$ 的状态转移阵。于是式(5.8.68)的 $\{\Phi(t)\}$ 必为随机方程

$$d\Phi(t) = -[A - PH^TR_2^{-1}H]^T \Phi(t)dt - H^TR_2^{-1}d(t) \quad (5.8.70)$$

$$\Phi(b) = 0 \quad (5.8.71)$$

的解。另外, 我们有

$$d[P(t)\Phi(t)] = P(t)d\Phi + P(t)\Phi(t)dt \quad (5.8.72)$$

把式(5.8.67)写成微分形式, 并注意到式(5.8.72), 得到

$$d\hat{x}(t|b) = d\hat{x}(t|a) + P(t)d\Phi(t) + P(t)\Phi(t)d(t) \quad (5.8.73)$$

再把式(5.8.15), (5.8.70)和(5.8.17)等式代入上式, 得到

$$d\hat{x}(t|b) = A\hat{x}(t|a)dt + PH^TR_2^{-1}d(t) + P[-A + PH^TR_2^{-1}H]^T \Phi(t)dt$$

$$-H^T R_2^{-1} d(t)] + [AP + PA^T + R_1 - PH^T R_2^{-1} HP](t) dt \\ = A[\hat{x}(t_0) + P(t)] dt + R_1(t) dt \quad (5.8.74)$$

由式(5.8.67)有

$$(t) = P^{-1}[\hat{x}(t_0) - \hat{x}(t_0)] \quad (5.8.75)$$

把上式和式(5.8.67)代入式(5.8.74), 得到最后结果

$$d\hat{x}(t_0) = A\hat{x}(t_0)dt + R_1 P^{-1}[\hat{x}(t_0) - \hat{x}(t_0)]dt \quad (5.8.76)$$

$$\frac{d\hat{x}(t_0)}{dt} = A\hat{x}(t_0) + R_1 P^{-1}[\hat{x}(t_0) - \hat{x}(t_0)] \quad (5.8.77)$$

习 题

5.1 随机变量 x 呈高斯 $N(a, \sigma^2)$ 分布, 用量测仪器确定 x , 量测误差 e 可用对 x 是独立的高斯 $N(0, \sigma_e^2)$ 随机变量来模拟。试在先验信息和量测使 $E\{\hat{x} - \hat{x}_0\}$ 为极小的意义上, 确定 x 的最优估计。

5.2 随机变量 x 呈高斯 $N(a, \sigma^2)$ 分布, 使用两种量测, 设两种量测误差是独立的, 并分别为高斯 $N(0, \sigma_1^2)$ 和 $N(0, \sigma_2^2)$ 分布。试确定 x 的最优估计和估计误差的方差。当 $\sigma_1, \sigma_2 \rightarrow 0$ 时, 再确定估计极限。

5.3 设 x 和 y 为高斯随机向量。试证明总能找到一个适当维数的矩阵 A 和一个高斯随机向量 e , 使得

$$x = Ay + e$$

式中 e 对 y 是独立的。并给出 Ay 的概率解释。

5.4 考虑标量随机过程

$$x(k) = \sum_{s=k_0}^k g(k, s) v(s) \\ y(k) = \sum_{s=t_0}^t h(k, s) e(s)$$

式中: $v(k)$ 和 $e(k)$ 是两个独立高斯 $N(0, 1)$ 变量的相关序列

$$E\{v(k)e(s)\} = \delta_{k,s}$$

且 $h(k, k) = 0$ 。试确定利用 $y(k), y(k-1), \dots, y(k_0)$ 对 $x(k)$ 的最优均方估计 $\hat{x}(k|k)$ 。

5.5 考虑标量随机过程

$$x(t) = \int_{s=t_0}^t g(t, s) dv(s) \\ y(t) = \int_{s=t_0}^t h(t-s) de(s)$$

式中: $e(t)$ 和 $v(t)$ 是具有单位方差参数的维纳过程

$$E\{dv(u)de(v)\} = \begin{cases} r dv & (u = v) \\ 0 & (u \neq v) \end{cases}$$

并假设 $y(t)$ 有逆

$$e(t) = \int_{s=t_0}^t k(t,s) dy(s)$$

试确定利用 $\{y(s), t_0 \leq s \leq t\}$ 使均方估计误差为极小的 $x(t)$ 的估计 $\hat{x}(t|s)$, 再确定估计误差的方差 $E\{\tilde{x}(t)\tilde{x}(t)|\mathcal{Y}(s)\}$ 。

5.6 考虑动力学系统

$$\begin{aligned} x(k+1) &= x(k) + \quad (k) \\ y(k) &= x(k) + \quad (k) \end{aligned}$$

式中: $\{\quad(k), k \geq T\}$ 和 $\{\quad(k), k \geq T\}$ 分别是两个独立高斯 $N(0, 1)$ 和 $N(0, \quad^2)$ 随机变量序列。初始状态是高斯 $N(1, \quad_0)$ 的。试确定系统的最优一步预测状态估计量和定态估计量。

5.7 考虑系统

$$\begin{aligned} x(k+1) &= x(k) + \quad (k) \\ y(k) &= x(k) + \quad (k) \end{aligned}$$

式中:

$$\begin{aligned} &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = [1 \quad 0] \\ R_1 &= \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix} \quad R_2 = 1 \end{aligned}$$

试确定一步预测定态估计的协方差 $P = \begin{bmatrix} p_1 & p_2 \\ p_2 & p_3 \end{bmatrix}$ 和增益矩阵 K 。

5.8 考虑系统和输出方程

$$\begin{aligned} x(k+1) &= x(k) + u(k) + \quad (k) \\ y(k) &= x(k) + \quad (k) \end{aligned}$$

式中: $\{v(k)\}$ 和 $\{e(k)\}$ 是离散时间高斯白噪声, 其均值为零, 协方差为

$$\begin{aligned} E\{\quad(k)^T(s)\} &= \quad_{k,s}R_1 \\ E\{\quad(k)^T(s)\} &= \quad_{k,s}R_{12} \\ E\{\quad(k)^T(s)\} &= \quad_{k,s}R_2 \end{aligned}$$

试证明最优估计 $\hat{x}(k) = \hat{x}(k|k-1)$ 给定为

$$\hat{x}(k+1) = \hat{x}(k) + u(k) + K[y(k) - \hat{x}(k)]$$

K 的最佳值给定为

$$\begin{aligned} K &= K(k) = K(k|k-1) \\ &= [P(k)^T + R_{12}][P(k)^T + R_2]^{-1} \end{aligned}$$

式中:

$$P(k+1) = P(k+1|k+1) = P(k)^T + R_1 - K(k)[P(k)^T + R_2]K^T(k)$$

5.9 考虑系统

$$\begin{aligned} x(k+1) &= x(k) + e(k) \\ y(k) &= x(k) + e(k) \end{aligned}$$

式中: $\{e(k)\}$ 是均值为零和协方差为 R_2 的独立高斯随机变量序列。初始状态是高斯的, 其均值为 m_0 , 协方差为 R_0 。假设矩阵 \quad , \quad 和 \quad 有常数元, 并且矩阵 \quad 的所有特征值都

在单位圆内。试确定一步预测最优估计量和定态估计误差的协方差。

5.10 设 x 和 y 是标量, 考虑系统

$$\begin{aligned}x(k+1) &= x(k) + be(k) \\ y(k) &= x(k) + e(k)\end{aligned}$$

式中: $\{e(k)\}$ 是独立高斯 $N(0, 1)$ 随机变量序列。初始状态 $x(k_0)$ 是高斯 $N(0, \frac{2}{0})$ 的, 并且所有 k 对 $e(k)$ 是独立的。试确定最优均方估计 $\hat{x}(k+1|k)$, $\hat{y}(k+1|k)$ 和预测误差, 特别要分析 $t_0 -$ 时的情况。

5.11 设 x 和 y 是标量, 考虑系统

$$\begin{aligned}x(k+1) &= x(k) + be(k) \\ y(k) &= x(k) + e(k)\end{aligned}$$

式中: $\{e(k)\}$ 是独立高斯 $N(0, 1)$ 随机变量序列。试利用最小方差预测理论确定一步最优均方预测 $\hat{y}(k+1|k)$ 。

5.12 考虑系统

$$y(k) + a(k-1)y(k-1) = b(k-1)u(k-1) + e(k)$$

式中:

$$\begin{aligned}a(k+1) &= a(k) + v_1(k) \\ b(k+1) &= b(k) + v_2(k)\end{aligned}$$

$\{e(k)\}$, $\{v_1(k)\}$ 和 $\{v_2(k)\}$ 是分别具有方差 r_2 , r_{11} 和 r_{22} 的离散时间高斯白噪声。试证明卡尔曼定理能用来得到参数 a 和 b 的估计。

5.13 已知系统

$$\begin{aligned}x^0(k+1) &= x^0(k) + u^0(k) \\ y^0(k) &= x^0(k) + v^0(k)\end{aligned}$$

式中: $u^0(t)$ 的均值 m 和 $v^0(t)$ 的均值 m 都是常数, 但是未知量。试确定状态量的最优估计, 以及 m 和 m 。

5.14 考虑系统

$$\begin{aligned}x(k+1) &= \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} x(k) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} e(k) \\ y(k) &= [1 \ 0] x(k) + e(k)\end{aligned}$$

式中: $\{e(k)\}$ 是独立高斯 $N(0, 1)$ 随机变量序列。试确定一步预测最优估计和估计误差的定态协方差。

5.15 已知对系统

$$\begin{aligned}x(k+1) &= x(k) + (k) \\ y(k) &= x(k) + (k)\end{aligned}$$

有新息

$$\begin{aligned}\tilde{x}(k+1|k) &= [1 - K(k|k-1)] \tilde{x}(k|k-1) + (k) - K(k|k-1) (k) \\ y(k|k-1) &= \tilde{x}(k|k-1) + (k)\end{aligned}$$

试求解一步预测和滤波问题。

5.16 考虑系统

$$x(k+1) = x(k) + u(k) + (k)$$

$$y(k) = x(k) + (k)$$

假设控制函数 $\{u(k), k \in T\}$ 已知, 试证明在 Y_{k-1} 下的 $x(k)$ 的条件均值为

$$\hat{x}(k+1|k) = \hat{x}(k|k-1) + u(k) + K(k|k-1)[y(k) - \hat{x}(k|k-1)]$$

5.17 考虑系统

$$A(q^{-1})y(t) = C(q^{-1})e(t)$$

式中:

$$A(q^{-1}) = 1 + a_1q^{-1} + \dots + a_nq^{-n}$$

$$C(q^{-1}) = 1 + c_1q^{-1} + \dots + c_nq^{-n}$$

试利用卡尔曼定理证明一步预测为

$$\hat{y}(t+1|t) = \frac{G(q^{-1})}{C(q^{-1})}y(t)$$

式中: $G(q^{-1})$ 是 $n-1$ 次多项式, 由下式决定

$$C(q^{-1}) = A(q^{-1}) + q^{-1}G(q^{-1})$$

5.18 考虑由

$$x(k+1) = (k+1, k)x(k) + (k)$$

$$y(k) = x(k) + (k)$$

描述的随机系统。设 $K(k|k)$ 是状态估计量

$$\begin{aligned} \hat{x}(k|k) = & (k, k-1)\hat{x}(k-1|k-1) \\ & + K(k|k)[y(k) - (k, k-1)\hat{x}(k-1|k-1)] \end{aligned}$$

的最优增益。再设 $K(k+1|k)$ 是 1 步预测

$$\begin{aligned} \hat{x}(k+1|k) = & (k+1, k+1-1)\hat{x}(k+1-1|k-1) \\ & + K(k+1|k)[y(k) - (k, k+1-1)\hat{x}(k+1-1|k-1)] \end{aligned}$$

的最优增益, 试证明

$$\begin{aligned} K(k+1|k) = & (k+1, k)K(k|k) \\ = & (k+1, k+1-1)(k+1-1, k+1-2)\dots(k+1, k)K(k|k) \end{aligned}$$

5.19 给定状态方程

$$x(k+1) = 0.6x(k) + 2u(k) + (k)$$

$$y(k) = 3x(k) + (k)$$

式中: $x(k_0) \sim N(5, 10)$, $(k) \sim N(0, 3)$, $(k) \sim N(1, 2)$ 互相独立; $u(k)$ 是控制量; 已测得 $y(k_0) = 14$ 。试求 $\hat{x}(k_0+1|k_0)$ 的值。

5.20 考虑随机系统

$$dx = xdt + dw$$

$$dy = xdt + dv$$

式中: $\{w(t)\}$ 和 $\{v(t)\}$ 是增量协方差分别为 r_1dt 和 r_2dt 的独立维纳过程。假设初始状态是高斯 $N(m_0, r_0)$ 的, 试证明最优滤波的增益是

$$K(t) = \frac{P(t)}{r_2}$$

式中:

$$P(t)=\frac{\frac{r_1}{r_2}\sinh t + r_0\cosh t + \frac{r_0}{r_2}\sinh t}{\cosh t - \frac{r_1}{r_2}\sinh t + \frac{r_0}{r_2}\sinh t}$$
$$\sigma^2 = \sigma_0^2 + \frac{r_1}{r_2}$$

5. 21 考虑随机微分方程

$$dy + aydt = budt + de$$

式中: {e(t)} 是方差参数为 r 的维纳过程; 参数 a 和 b 由

$$da = -adt + dv$$

$$db = -bdt + dw$$

给定; {v(t)} 和 {w(t)} 是方差参数分别为 r₁₁ 和 r₂₂ 的独立维纳过程。试导出能确定参数 a 和 b 的递推方程。

5. 22 对离散时间随机系统

$$x(k+1) = x(k) + \tilde{w}(k)$$
$$y(k) = x(k) + v(k)$$

已知有卡尔曼一步预测估计定理

$$\hat{x}(k+1) = \hat{x}(k) + K(k)[y(k) - \hat{y}(k)]$$
$$\hat{x}(k_0) = m_0$$
$$K(k) = P(k)^T [P(k)^T + R_2]^{-1}$$
$$P(k+1) = P(k)^T + R_1 - K(k)P(k)^T$$
$$P(k_0) = R_0$$

试通过把离散时间问题取极限的形式来得到连续时间卡尔曼-布西滤波定理。

5. 23 把卡尔曼-布西滤波定理推广到噪声为

$$dw = dv + v_1 dt$$
$$dv = de + e_1 dt$$

的情况, 其中 v 和 e 是维纳过程, v₁ 和 e₁ 是未知常数。

5. 24 考虑具有参数

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad C = [1 \quad 0]$$
$$R_1 = \begin{bmatrix} 1 & 0 \\ 0 & r_2 \end{bmatrix} \quad R_2 = r_2$$

的系统。试确定估计系统状态的定态卡尔曼滤波器。再确定 r₂ → 0 时的最优滤波器的传递函数的极限。

5. 25 考虑系统

$$dx = Axdt + dw$$
$$dy = Cxdt + dv$$

式中:

$$\begin{aligned} A &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} & C &= [1 \quad 0] \\ R_1 &= \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} & R_2 &= 0 \end{aligned}$$

试确定定态卡尔曼滤波器和最优滤波器的传递函数。

5. 26 考虑以下系统的状态估计问题:

$$\begin{aligned} dx &= Axdt + Bdu \\ dy &= Cxdt + de \\ dz &= -A^Tzdt + C^Tdv \\ dy &= B^Tzdt + dn \end{aligned}$$

式中: {u(t), t ∈ T}, {e(t), t ∈ T}, {v(t), t ∈ T} 和 {n(t), t ∈ T} 是增量协方差分别为 R₁dt, R₂dt, R₂⁻¹dt 和 R₁⁻¹dt 的过程。初始状态是高斯的, 其协方差分别为 R₀ 和 R₀⁻¹。试证明这两个状态估计问题是对偶的。

5. 27 考虑系统

$$\begin{aligned} dx &= Axdt + dw \\ dy &= Cxdt + dv \end{aligned}$$

试证明, 当 s = t 时, 在 Y_t 下的 x(s) 的极小均方预测为

$$\begin{aligned} \hat{x}(s|Q_t) &= (s, t) \hat{x}(t) \\ \hat{x}(t) &= \hat{x}(t|Q_t) \end{aligned}$$

预测误差的协方差 P(s|Q_t) 由

$$\begin{aligned} \frac{dP(s|Q_t)}{ds} &= A(s)P(s|Q_t) + P(s|Q_t)A^T(s) + R_1(s) \\ P(t|Q_t) &= P(t) \end{aligned}$$

给定, 并推导预测的递推公式。

5. 28 设 x 和 y 是标量, 考虑系统

$$\begin{aligned} dx &= xdt + bde \\ dy &= xdt + de \end{aligned}$$

式中: {e(t), t ∈ T} 是具有单位方差参数的维纳过程。初始状态是高斯 N(0, r₀) 的。试求出在 Y_t 下的 x(t) 的最优均方滤波, 并确定定态滤波增益和滤波误差的定态方差。

5. 29 陀螺仪的随机漂移用以下简化数学模型来描述:

$$\begin{aligned} dx_1 &= (x_2 + x_3)dt \\ dx_2 &= 0 \\ dx_3 &= -\frac{1}{s}x_3dt + \frac{2}{s}dv \end{aligned}$$

式中: x₁ 是漂移角; x₂ 是日常漂移; x₃ 是“随机漂移”; 而 {v(t), t ∈ T} 是具有单位方差参数的维纳过程。x₁ 是无差量测, 且 x₁(0) = 0; x₂(0) 是高斯 N(0, $\frac{2}{d}$) 的; x₃(0) 是高斯 N(0, $\frac{2}{s}$) 的。试确定分量 x₂ 和 x₃ 的最小方差滤波。

5.30 设信号 $\{s(t), t \in T\}$ 和噪声 $\{n(t), t \in T\}$ 是协方差函数分别为 $r_s(t)$ 和 $r_n(t)$ 的独立平稳随机过程。假设线性组合

$$y = s + n$$

已被观测, 并且希望通过线性运算

$$\hat{s}(t+h)=\int_{-\infty}^tg(t-u)y(u)du\\=\int_0^hg(u)y(t-u)du$$

在长为 h 的区间上预测 s 。试证明均方预测误差为

$$E\{[s(t+h)-\hat{s}(t+h)]^2\}=r_s(0)-2\int_0^hg(u)r_s(h+u)du\\+\int_0^hg(u)du\int_0^hg(v)[r_s(u-v)+r_n(u-v)]dv$$

因此, 预测问题可用公式表示为使均方预测误差为极小。

5.31 考虑由习题(5.30)得到的均方预测误差为预测加权函数 g 的泛函

$$J[g]=r_s(0)-2\int_0^hg(u)r_s(u+h)du+\int_0^h\int_0^hg(u)[r_s(u-v)+r_n(u-v)]g(v)dudv$$

试证明这个泛函的一次和二次变分由

$$J[g+g_1]=J[g]+J_1+J_2$$

给定, 式中:

$$J_1=-2\int_0^hg(u)[r_s(u+h)-\int_0^h[r_s(u-v)+r_n(u-v)]g(v)dv]du\\J_2=\int_0^h\int_0^hg(u)g(v)[r_s(u-v)+r_n(u-v)]dudv$$

然后证明, g 是最优预测加权函数的充要条件为

$$r_s(t+h)-\int_0^h[r_s(t-v)+r_n(t-v)]g(v)dv=0\quad(t\in T)$$

和 r_s+r_n 是正定的。这个积分方程(上式)叫做维纳-霍甫夫(Wiener-Hopf)方程。

5.32 考虑由下式定义的标量离散时间随机过程 $\{y(t), t \in T\}$:

$$y(t)=\sum_{n=t_0}^tg(t,n)e(n)$$

式中: $\{e(t), t \in T\}$ 是独立同分布高斯 $N(0, 1)$ 随机变量序列; 对每个 $t, g(t, t) > 0$ 。试确定过程的最佳 k 步均方预测和预测误差。

5.33 考虑由下式定义的标量连续时间随机过程 $\{y(t), t \in T\}$:

$$y(t)=\int_{t_0}^tg(t,s)dw(s)$$

式中: $\{w(t), t \in T\}$ 是具有单位方差参数的维纳过程。试确定在区间 $(t, t+h)$ 上的最优均方预测 $\hat{y}(t+h|t)$ 和预测误差 $\tilde{y}(t+h|t)$ 。规定 $t_0=-\infty$ 和

$$g(t,s)=(t-s)e^{-(t-s)}$$

5.34 考虑由下式定义的标量连续时间随机过程 $\{y(t), t \in T\}$:

$$y(t) = \int_{t_0}^t g(t,s)dw(s)$$

式中:

$$g(t,s) = (t-s)e^{-(t-s)}$$

{w(t), t ∈ T}是具有单位方差参数的维纳过程

$$E\{dw(t)dw(t)\} = dt$$

并取 t₀= -∞。试确定在区间(t, t+ h)上形式为

$$\hat{y}(t+h|t) = y(t)$$

的最优均方预测。

5.35 考虑动力学系统

$$x(k+1) = x(k) + u(k) + (k)$$

$$y(k) = x(k) + (k)$$

式中: {(k)}和{(k)}是离散时间白噪声,其均值为零,协方差为

$$E\{(s)^T(k)\} = s,kR_1$$

$$E\{(s)^T(k)\} = s,kR_{12}$$

$$E\{(s)^T(k)\} = s,kR_2$$

试证明,本系统的状态变量可用数学模型

$$\hat{x}(k+1) = \hat{x}(k) + u(k) + K(k)[y(k) - \hat{x}(k)]$$

复现,增益向量 K(k)的最优值为

$$K(k) = [P(k)^T + R_{12}](P(k)^T + R_2)^{-1}$$

式中:

$$P(k+1) = P(k)^T + R_1 - K(k)(P(k)^T + R_2)K^T(k)$$

用所示数学模型复现时,其性能指标是使复现误差

$$\tilde{x}(k) = x(k) - \hat{x}(k)$$

的方差

$$P(k) = E\{[\tilde{x}(k) - E\{\tilde{x}(k)\}][\tilde{x}(k) - E\{\tilde{x}(k)\}]^T\}$$

为极小。

5.36 考虑动力学系统

$$x(k+1) = x(k) + u(k) + (k)$$

$$y(k) = x(k) + (k)$$

式中: (k)和 (k)是离散时间白噪声,其均值为零,协方差为

$$E\{(s)^T(k)\} = s,kR_1$$

$$E\{(s)^T(k)\} = s,kR_2$$

$$E\{(s)^T(k)\} = 0$$

(k)和 (k)与 x(k₀)无关。试证明,系统的状态变量用数学模型

$$\hat{x}(k+1) = \hat{x}(k) + u(k) + K(k+1)\{y(k+1) - [\hat{x}(k) + u(k)]\}$$

复现时,K(k+1)的最优选择为

$$\begin{aligned}K(k) &= P(k)^{-T} [P(k)^{-T} + R_2]^{-1} \\P(k) &= S(k-1)^{-T} + R_1 \\S(k) &= P(k) - K(k) P(k) \\S(k_0) &= R_0\end{aligned}$$

式中: $S(k)$ 是复现误差的方差

$$S(k) = E\{\tilde{x}(k) \tilde{x}^T(k)\}$$

5. 37 考虑给定系统

$$\begin{aligned}x(k+1) &= x(k) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(k) \\y(k) &= x(k)\end{aligned}$$

有关参数和初值为

$$\begin{aligned}A &= \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} = [1 \quad 0] \\R_1 &= \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad R_2 = 0 \\S(0) &= \begin{bmatrix} \frac{2}{1} & 0 \\ 0 & \frac{2}{2} \end{bmatrix}\end{aligned}$$

试确定形式为

$$\hat{x}(k+1) = \hat{x}(k) + u(k) + K(k+1)\{y(k+1) - [1 \quad 0] \hat{x}(k) + u(k)\}$$

的最优复现器的增益矩阵 $K(k)$ 和复现误差的协方差阵。

5. 38 考虑动力学系统

$$\begin{aligned}dx &= Axdt + Budy + dv \\dy &= Cxdt + de\end{aligned}$$

式中: y 是观测量。 $v(t)$ 和 $e(t)$ 是均值为零和增量协方差分别为 R_1 和 R_2 的不相关维纳过程。状态初值为 $\hat{x}(t_0) = E\{x(t_0)\}$, $v(t)$ 和 $e(t)$ 对 $x(t_0)$ 独立。试证明形式为

$$\begin{aligned}d\hat{x} &= A\hat{x}dt + Budy + K(dy - C\hat{x}dt) \\K(t) &= P(t)C^TR_2^{-1} \\\frac{dP}{dt} &= AP + PA^T + R_1 - PC^TR_2^{-1}CP \\P(t_0) &= R_0\end{aligned}$$

的复现器在均方意义上是最优的。

5. 39 考虑一个粒子的直线运动, 假设该粒子的加速度可描述为谱密度为 $\frac{1}{2}$ 的白噪声过程, 而观察到的该粒子的位置含有谱密度为 $\frac{r}{2}$ 的白噪声量测误差。试确定该粒子位置和速度的形式为

$$d\hat{x}(t) = A\hat{x}(t)dt + K(t)[dy(t) - C\hat{x}(t)dt]$$

的最小方差复现器。再求出复现误差的协方差阵。

5.40 考虑动力学系统

$$dx = A x dt + B u dt + dv$$

$$dy = C x dt + de$$

其中 v 和 e 是均值为零和增量协方差为

$$E \begin{bmatrix} dv \\ de \end{bmatrix} \begin{bmatrix} dv^T & de^T \end{bmatrix} = \begin{bmatrix} R_1 & R_{12} \\ R_{12}^T & R_2 \end{bmatrix} dt$$

的相关维纳过程。试证明,当初值选为 $\hat{x}(t_0) = E\{x(t_0)\}$ 和增益矩阵给定为

$$K(t) = [P(t)C^T + R_{12}]R_2^{-1}$$

时,就能得到结构为

$$d\hat{x} = A\hat{x} dt + B u dt + K[dy - C\hat{x} dt]$$

的最小方差复现器,上式中复现误差的协方差阵满足里卡蒂方程

$$\begin{aligned} \frac{dP}{dt} = & [A - R_{12}R_2^{-1}C]P + P[A - R_{12}R_2^{-1}C]^T + R_1 \\ & - R_{12}R_2^{-1}R_{12}^T - PC^TR_2^{-1}CP \end{aligned}$$

$$P(t_0) = R_0$$

5.41 图题 5.41 表示一个简单的垂直校准系统,它由平台伺服机构和加速计组成。加速计信号给出垂直指示误差的信息。此信号使平台法线恢复到垂直位置。在偏差较小

图题 5.41 垂直校准系统示意图

时,加速计信号是

$$y = \quad + n$$

式中: y 是垂直指示误差; n 是由于水平加速所引起的扰动。平台用以下方程描述:

$$\frac{d}{dt} = u$$

式中: u 是控制信号。伺服环路可以描述为

$$u = -ky$$

式中: k 是增益。假设噪声可用白噪声模拟,这时系统能由以下随机微分方程描述:

$$d = -k dt - k dw$$

式中: $w(t)$ 是方差参数为 r 的维纳过程。假设初始状态是高斯 $N(0, \quad)$ 的,试求出使垂直校正误差的方差尽量小的时变增益 $k(t)$ 。

第 6 章 最优随机控制

6.1 引言

一般随机控制系统表示为

$$\dot{x}(t) = f[t, x(t), u(t), dw(t)] \quad (6.1.1)$$

它的最优控制问题通过给定性能指标(价值函数) $J(u)$ 来说明。性能指标 $J(u)$ 通常表示为

$$J(u) = E \int_0^T h[t, x(t), u(t)] dt + g(T) \quad (6.1.2)$$

随机最优控制就是通过选择 $u(t) \in U$, 使性能指标 $J(u)$ 达到最小, 这种控制称为随机最优控制。在工程问题中, h 通常选为偏离状态 $\{x(t)\}$ 某个期望值所花费的代价(费用), 或选为控制力或消耗的能量, 而 g 表示在终端时刻 T 达到某个规定目标失败的价值。例如对具体的投资问题, $\{x(t)\}$ 可表示某公司的财富, h 表示消费率, 而 g 则为馈赠函数。

我们知道, 确定性最优控制的性能指标 $J_d(u)$ 一般表示为

$$J_d(u) = \int_0^T h[t, x(t), u(t)] dt + g[x(T)] \quad (6.1.3)$$

对照式(6.1.2)和(6.1.3), 有

$$J(u) = E\{J_d(u)\} \quad (6.1.4)$$

上式似乎表明, 随机控制系统的性能指标, 可通过对确定性控制性能指标各采样函数取平均值来简单地获得。例如, 破产概率为 0.8 的投资策略, 很多人认为是不可取的, 即使他们的收益可能取得极大值。然而, 对随机情况可以通过选择 h 和 g 而加以考虑, 使它成为可以接受的事件。

确定性控制与随机控制之间的基本差别在于, 在确定性控制中开环和闭环控制是等价的, 而随机控制则不同。对确定性控制, 任何开环控制都可看做是退化的闭环控制。下面证明这一点。设确定性动力学系统为

$$\dot{x}(t) = f[t, x(t), u(t)] \quad (6.1.5)$$

考虑闭环控制, 加入反馈控制 $u[t, (x(s), s - t)]$, 使方程式(6.1.5)

$$\dot{x}(t) = f[t, x(t), u(t), (x(s), s - t)] \quad (6.1.6)$$

有唯一解 $x(t)$ 。现考虑开环控制, 把反馈控制作为开环控制的输入量 $u^*(t)$, 即

$$u^*(t) = u[t, (x(s), s - t)] \quad (6.1.7)$$

得到系统解 $x^*(t)$ 为

$$\dot{x}^*(t) = f[t, x^*(t), u^*(t)] \quad (6.1.8)$$

显然 $x^*(t) = x(t)$ 。因此, 只要选开环控制信号 $u^*(t)$ 与闭环控制信号 $u(t)$ 相同, 开环控制与闭环控制就是等价的。

然而, 对随机控制来说, 开环控制与闭环控制却是不同的。下面举一个例子说明这

一点。

例 6.1.1 给定动力学系统和性能指标分别为

$$x_1 = x_0 + u_0 \tag{6.1.9}$$

$$J(u_0) = E\{x_1^2\} \tag{6.1.10}$$

式中：状态初值 x_0 为随机变量, 取值和概率为

$$x_0 = \begin{matrix} 0 & p = \frac{1}{2} \\ 1 & p = \frac{1}{2} \end{matrix} \tag{6.1.11}$$

试分别在开环和闭环情况下, 寻找控制策略 $u_0 = u(x_0)$, 使性能指标 $J(u_0)$ 最小。

解 (1) 开环控制情况

利用式(6.1.9)和(6.1.11), 直接计算式(6.1.10), 得到

$$\begin{aligned} E\{x_1^2\} &= p(0 + u_0)^2 + p(1 + u_0)^2 \\ &= \frac{1}{2}(1 + 2u_0 + 2u_0^2) \end{aligned} \tag{6.1.12}$$

$$E\{x_1^2\}_{u_0} = \frac{1}{2}(2 + 2u_0) = 0 \quad \text{即} \quad u_0 = -\frac{1}{2} \tag{6.1.13}$$

$$\min J(u_0) = \frac{1}{2} \left(1 + 2\left(-\frac{1}{2}\right) + 2\left(-\frac{1}{2}\right)^2 \right) = \frac{1}{4} \tag{6.1.14}$$

求得最优控制策略为 $u_0 = -\frac{1}{2}$, 最小性能指标为 $\min J(u_0) = \frac{1}{4}$ 。

(2) 闭环控制情况

显然, 最优控制策略为 $u_0 = -x_0$, 最小性能指标为 $\min J(u_0) = 0$ 。

计算结果表明：对随机控制来说, 开环控制与闭环控制确实不同, 闭环控制优于开环控制。

一般随机控制系统, 除具有对象模型式(6.1.1)外, 还具有量测(观测)模型。下面也举一个简单例子, 具体说明这种情况。

例 6.1.2 在例 6.1.1 题设条件下, 加入观测噪声 w , 观测方程表示为

$$y = x_0 + w \tag{6.1.15}$$

式中： y 为观测值; w 的分布为

$$w = \begin{matrix} -1 & p = \frac{1}{3} \\ 0 & p = \frac{1}{3} \\ +1 & p = \frac{1}{3} \end{matrix} \tag{6.1.16}$$

试寻找控制策略 $u_0 = u(y)$, 使性能指标 $J(u_0)$ 最小。

解 首先在表 6.1.1 中列出 x_0 , w 和 y 的各种条件分布结果。然后求出各种情况下的最优控制策略 u_0 和最小性能指标 $J(u_0)$ 。

表 6.1.1 x_0, w 和 y 的条件分布表

x_0	0	0	1	0	1	1
w	- 1	- 1	0	0	1	1
y	- 1	0		1		2
p	$\frac{1}{6}$	$\frac{1}{3}$		$\frac{1}{3}$		$\frac{1}{6}$

第一种情况:

$y = - 1 \qquad x_0 = 0 \qquad p = \frac{1}{6}$

$J(u_0)_1 = E\{x_1^2\} = E\{[x_0 + u_0]^2\} = \frac{1}{6}(0 + u_0)^2 = \frac{1}{6}u_0^2 \qquad (6.1.17)$

$[J(u_0)_1]_{u_0} = \frac{1}{6} \cdot 2u_0 = 0 \qquad \text{即} \quad u_0 = 0 \qquad (6.1.18)$

$\min J(u_0)_1 = \frac{1}{6} \cdot 0^2 = 0 \qquad (6.1.19)$

第二种情况:

$y = 0 \qquad x_0 = \begin{matrix} 0 \\ 1 \end{matrix} \qquad \begin{matrix} p = \frac{1}{6} \\ p = \frac{1}{6} \end{matrix}$

$J(u_0)_2 = \frac{1}{6} (0 + u_0)^2 + (1 + u_0)^2 = \frac{1}{6}[1 + 2u_0 + 2u_0^2] \qquad (6.1.20)$

$[J(u_0)_2]_{u_0} = \frac{1}{6}[2 + 2 \cdot 2u_0] = 0 \qquad \text{即} \quad u_0 = - \frac{1}{2} \qquad (6.1.21)$

$\min J(u_0)_2 = \frac{1}{6} [1 + 2 - \frac{1}{2} + 2 - \frac{1}{2}^2] = \frac{1}{12} \qquad (6.1.22)$

第三种情况:

同第二种情况, 得到 $u_0 = - \frac{1}{2}, \min J(u_0)_3 = \frac{1}{12}$ 。

第四种情况:

$y = 2, x_0 = 1, p = \frac{1}{6}$, 经同样计算得到 $u_0 = - 1, \min J(u_0)_4 = 0$,

综上所述, 得到结果

$$\begin{aligned} u_0(- 1) &= 0 \\ u_0(0) &= - \frac{1}{2} \\ u_0(1) &= - \frac{1}{2} \\ u_0(2) &= - 1 \\ \min J(u_0) &= \min_{i=1}^4 J(u_0)_i = \frac{1}{6} \end{aligned}$$

由表 6.1.1 可知, p 为在给定观测 y 条件下的状态 x_0 的条件概率, 如第一种情况表示在 $y = -1$ 条件下, 状态 $x_0 = 0$ 的条件概率。这表示滤波在随机控制中的作用。很明显, 随机控制受状态观测的影响, 而闭环和开环控制可看做是零噪声和“无限”噪声的极限情况。

研究控制理论主要有两种趋势, 一种是基于变分的计算, 后来庞特里亚金创造了极大值原理, 另一种是贝尔曼动态规划法。一般说, 由于技术原因, 前者用在确定性控制更为成功, 后者在随机控制中更为成功。我们采用贝尔曼动态规划法讨论随机控制问题。

我们主要讨论线性调节问题, 线性调节问题有三种情况: 一种情况是选择 $u(t)$ 为量测量 $\{y(s), s \in t\}$ 的函数, 称为不完全状态信息(部分观测)情况; 另一种情况是当状态可在无误差方法量测时, 即不需要量测方程时, 致使 $x(t) = y(t)$, 称为完全状态信息(完全观测)情况; 第三种情况是确定性情况。所有三种情况的解在某种意义上都是相同的。我们主要讨论前两种情况, 它们属于随机控制问题, 但我们也会涉及第三种情况的结果。

本章研究的系统状态模型是线性的 (linear), 选择的性能指标函数是二次型的 (quadratic), 在高斯假设下, 状态向量的分布是高斯的 (Gaussian), 这样的问题称为线性-二次型-高斯控制问题, 简称 LQG 问题。

对 LQG 问题, 我们得到的最优控制 $u(t)$ 为最优滤波 $\hat{x}(t|Q)$ 的线性函数(离散时间系统也一样), 即为 $u(t) = -L(t)\hat{x}(t|Q)$ 。运算分成两部分, 一部分是计算 $\hat{x}(t|Q)$, 另一部分是计算控制增益阵 $L(t)$, 两者基本上独立进行, 这称为分离定理。如果把 $\hat{x}(t|Q)$ 换成 $x(t)$, 也就是把不完全状态信息情况换成完全状态信息情况或确定性情况, 最优控制形式仍然为 $u(t) = -L(t)x(t)$, 只是把 $\hat{x}(t|Q)$ 换成 $x(t)$, 因此, 分离定理也称为必然等价原理。

本章 6.2 ~ 6.6 节讨论离散时间随机状态模型的最优控制问题, 6.7 ~ 6.10 节讨论连续时间随机状态模型的最优控制问题, 6.11 节讨论最优控制与状态估计的对偶性, 最后, 在 6.12 节介绍里卡蒂方程的解法。

6.2 离散时间随机状态控制系统模型

6.2.1 模型的描述

1. 模型与假设

我们主要研究线性、离散时间随机状态控制系统, 为此采用线性、离散时间随机状态模型, 在 3.3 节讨论过这种模型, 表示为

$$x(k+1) = (k+1, k)x(k) + (k)u(k) + (k) \tag{6.2.1}$$

$$y(k) = (k)x(k) + (k) \tag{6.2.2}$$

式(6.2.1)称为对象方程, 式(6.2.2)称为量测方程。式中: $x(k)$ 为 $n \times 1$ 状态向量; $y(k)$ 为 $m \times 1$ 量测向量; $u(k)$ 为 $k \times 1$ 控制向量; $(k+1, k)$ 为 $n \times n$ 状态转移阵; (k) 为 $n \times n$ 量测阵; (k) 为 $n \times 1$ 控制阵; (k) 为 $n \times 1$ 模型噪声向量; (k) 为 $m \times 1$ 量测噪声向量。

$(k+1, k)$, (k) 和 (k) 是已知的, 可以是常阵或是时变的。

对式(6.2.1)和(6.2.2), 我们假设 (k) 和 (k) 分别为给定独立同分布高斯

$N(0, R_1(t))$ 和 $N(0, R_2(t))$ 随机变量序列, 也就是 (k) 和 (k) 都为离散时间高斯白噪声, 表示为

$$E\{ (k) \} = 0 \qquad E\{ (1)^T(k) \} = R_1(k)_{1,k} \qquad (6.2.3)$$

$$E\{ (k) \} = 0 \qquad E\{ (1)^T(k) \} = R_2(t)_{1,k} \qquad (6.2.4)$$

式中: $R_1(t)$ 和 $R_2(t)$ 可以是常阵, 也可以是时变的。在第 5 章已经介绍过, 要求 $R_1(t)$ 是非负定的, $R_2(t)$ 是正定的。一般还要求 (h) 和 (k) 互相独立, 表示为

$$E\{ (1)^T(k) \} = 0 \qquad (6.2.5)$$

我们假设初始状态 $x(k_0) = x(0)$ 为给定高斯分布 $N(m_0, R_0)$ 的, 其中 R_0 为非负定的, 且要求 $x(k_0)$ 与 (k) 和 (k) 互相独立, 即

$$E\{x(k_0)\} = m_0 \qquad R[x(k_0)] = R_0 \qquad (6.2.6)$$

$$\text{cov}[x(k_0), (k)] = 0 \qquad (k \leq k_0) \qquad (6.2.7)$$

$$\text{cov}[x(k_0), (k)] = 0 \qquad (k > k_0) \qquad (6.2.8)$$

在上述假设条件下, 模型式(6.2.1)和(6.2.2)中状态向量 $x(k)$ 和量测向量 $y(k)$ 都成为高斯向量, $y(k)$ 与 (k) 和 (k) 也是独立的, 表示为

$$\text{cov}[y(1), (k)] = 0 \qquad (k \leq 1) \qquad (6.2.9)$$

$$\text{cov}[y(1), (k)] = 0 \qquad (k > 1) \qquad (6.2.10)$$

2. 性能指标

随机控制系统的性能用加权二次型性能指标函数

$$J(u) = E \left\{ x^T(N) Q_0 x(N) + \sum_{k=k_0}^{N-1} x^T(k) Q_1(k) x(k) + u^T(k) Q_2(k) u(k) \right\} \qquad (6.2.11)$$

表征。由于 $\{x(k)\}$ 是随机过程, 把整个二次型取均值是合适的。式中: k_0 和 N 分别表示时间初值和终值; Q_0 , $Q_1(k)$ 和 $Q_2(k)$ 为加权对称阵, 可以是常阵, 也可以是时变的, 为满足后面计算上的要求, 这里假定 Q_0 和 $Q_1(k)$ 是非负定的, $Q_2(k)$ 是正定的。

性能指标函数 $J(u)$ 是 $x(k)$ 和 $u(k)$ 的性能量度, 是 $x(k)$ 和 $u(k)$ 的单调非减函数, $J(u)$ 通过选取 u 而达到最小。关于 $x(k)$ 项, 可看做是对 $x(k)$ 偏离零值大小的量度或对某个值的偏差。若选择 Q_0 和 Q_1 等于零, 则说明该系统对 $x(k)$ 的偏差无要求, 换句话说, 允许 $x(k)$ 的偏差可能达到很大的值。若选择 Q_0 和 Q_1 很大, 则说明该系统对 $x(k)$ 的偏差有严格限制, 偏差不能太大。相对来说, Q_0 和 Q_1 较小时, 对 $x(k)$ 的偏差要求低, 系统调整速度是较快的。相反, Q_0 和 Q_1 较大时, 使 $x(k)$ 的偏差较小, 系统调整速度将是较慢的。 $x(N)$ 是终值条件, 是给定的, 不需要再加以控制。关于 $u(k)$ 项, 可看做是对 $u(k)$ 强度的量度, 例如, 这个控制强度可以是控制能量, 表示该系统消耗能量的多少。一般说, 若选择 Q_2 等于零, 则表示该系统对输入能量无限制, 换句话说, 系统能提供任意需要的能量, 对实际设备来说, 很难满足这一要求, 输入量很大时, 或使设备达到饱和, 或危及设备安全。若选择 Q_2 很大, 则表示该系统对输入有严格限制, 只能消耗较小的能量。相对来说, Q_2 较小时, 系统调整速度较快, Q_2 较大时, 系统调整速度较慢。 Q_0 , Q_1 和 Q_2 就是在这些限制之间通过选择确定。

把 $J(u)$ 与古典控制性能指标相比, 可以看到, 现代控制性能指标用 Q_0, Q_1 和 Q_2 表征, 意义不直观明确, 多采用计算机求解, 计算准确; 而古典控制性能指标用超调量、调节时间和振荡次数等量表示, 意义直观明确, 多采用经验凑试法求解, 不甚准确。

3. 完全和不完全状态信息情况

随机控制系统按照有无量测方程可分为两种情况。

完全状态信息(完全观测)情况: 这种情况下不需量测方程, 状态向量 $x(k)$ 既作被控量又作量测量就够了, 它是随机控制系统的简单情况。

不完全状态信息(部分观测)情况: 这种情况下状态向量 $x(k)$ 不能直接量测, 需另设量测方程, 且存在量测噪声, 它是随机控制系统的一般情况。

4. 容许控制

我们规定容许控制 $u(k)$ 是量测集 U 的函数, 即 $u(k) \in U$ 。对完全状态信息情况:

$$u(k) = u(k, X_k) \tag{6.2.12}$$

式中 X_k 为直到 $x(k)$ 的所有状态向量集合, 即

$$X_k = [x^T(k_0) \ x^T(k_0 + 1) \ \dots \ x^T(k)]^T \tag{6.2.13}$$

对不完全状态信息情况:

$$u(k) = u(k, Y_{k-1}) \quad \text{或} \quad u(k) = u(k + 1, Y_k) \tag{6.2.14}$$

式中 Y_{k-1} 为直到 $y(k-1)$ 的所有量测向量集合, 即

$$Y_{k-1} = [y^T(k_0) \ y^T(k_0 + 1) \ \dots \ y^T(k-1)]^T \tag{6.2.15}$$

5. 问题提法

综上所述, 随机最优控制问题的提法是, 对随机控制系统式(6.2.1)和(6.2.2), 求出使性能指标函数 $J(u)$ 为最小的容许最优控制策略 $u(k)$ 。

6.2.2 最优性原理

我们将利用贝尔曼动态规划法来求解随机控制系统问题, 而动态规划法的核心是最优性原理, 因此, 了解最优性原理是必要的。

定理 6.2.1 (最优性原理) 一个多级决策过程的最优策略具有这种性质, 即不管其初始状态和初始决策如何, 其余的决策必须依初始决策所形成的状态组成一个最优策略。

最优性原理也可简单表述为: 最优策略(状态)的一部分也必为最优策略(状态)。

证明 证明是简单的。考虑随机控制系统式(6.2.1), (6.2.2)和性能指标函数式(6.2.11), 设 $u^*[k_0, N]$ 是使性能指标函数 $J(u)$ 为最小的最优策略, 即满足

$$J(u^*) \leq J(u) \quad (u \in U) \tag{6.2.16}$$

它对应最优状态 $x^*[k_0, N]$ 。

现反设在区间 $[t_1, N]$, $t_1 \in [k_0, N]$ 上存在另一控制策略 $u^0[t_1, N]$, 使性能指标函数更小, 即

$$J(u^*[t_1, N]) < J(u^0[t_1, N]) \tag{6.2.17}$$

于是有

$$J(u^*[k_0, N]) = J(u^*[k_0, t_1]) + J(u^0[t_1, N]) \tag{6.2.18}$$

使 $u^*[k_0, N]$ 不为最优策略, $x^*[k_0, N]$ 也不为最优状态, 因为它们不能使 $J(u^*[k_0, N])$ 最小。这与假设条件不符, 也就是说, $u^*[k_0, N]$ 是最优的, $u^0[t_1, N]$ 是不存在的, 除非 $u^0[t_1, N] = u^*[t, N]$, 这也说明最优策略是唯一的。

6.2.3 高斯向量二次型的均值

由式(6.2.11) 已知, 系统性能指标 $J(u)$ 是 $x(k)$ 和 $u(k) (k \in [k_0, N])$ 的二次型函数。我们还已知, 已假设 $\{x(k)\}$ 是高斯的, 因此, $J(u)$ 是高斯二次型函数。为处理 $J(u)$, 需要计算高斯二次型的均值。本节先介绍矩阵迹的概念, 再用迹表示高斯向量二次型的均值。

1. 矩阵的迹(trace)

定义 6.2.1 $n \times n$ 方阵 A 的主对角线元素 $a_{ii} (i = 1, 2, \dots, n)$ 之和称为矩阵 A 的迹, 用 $\text{tr}A$ 表示, 即

$$\text{tr}A = \sum_{i=1}^n a_{ii} \tag{6.2.19}$$

迹的性质:

(1) 转置矩阵的迹不变

$$\text{tr}A^T = \text{tr}A \tag{6.2.20}$$

(2) 矩阵和的迹等于矩阵迹的和。设 A 和 B 都为方阵, 则有

$$\text{tr}(A + B) = \text{tr}A + \text{tr}B \tag{6.2.21}$$

(3) 常数与矩阵乘积的迹等于常数与矩阵迹相乘。设 c 为常数, 则有

$$\text{tr}(cA) = c \text{tr}A \tag{6.2.22}$$

(4) 标量的迹等于自身。设 a 为标量, 则有

$$\text{tr}a = a \tag{6.2.23}$$

(5) 设 A 为 $n \times m$ 阵, B 为 $m \times n$ 阵, 则有

$$\text{tr}(AB) = \text{tr}(BA) \tag{6.2.24}$$

证明 性质(1) ~ (4) 推导简单, 略去证明。性质(5) 的证明如下:

$$\text{tr}(AB) = \sum_{i=1}^n \sum_{j=1}^m a_{ij} b_{ji} = \sum_{j=1}^m \sum_{i=1}^n b_{ji} a_{ij} = \text{tr}(BA) \tag{6.2.25}$$

2. 高斯向量二次型的均值

设 x 为 $n \times 1$ 高斯 $N(m, R)$ 向量, S 为 $n \times n$ 方阵, 则有

$$E\{x^T S x\} = m^T S m + \text{tr}SR \tag{6.2.26}$$

下面证明式(6.2.26) 成立。

由于

$$E\{(x - m)^T S (x - m)\} = E\{x^T S x\} - m^T S m \tag{6.2.27}$$

另外, 考虑 $(x - m)^T S (x - m)$ 二次型是标量, 它的迹就是它自身。还有, 取均值和取迹符号可以调换位置, 而其值不变。因此有

$$\begin{aligned} E\{(x - m)^T S (x - m)\} &= E\{\text{tr}(x - m)^T S (x - m)\} \\ &= \text{tr}S E\{(x - m)(x - m)^T\} \\ &= \text{tr}SR \end{aligned} \tag{6.2.28}$$

上式与式(6.2.27)相等,就可得到式(6.2.26)。

在 y 条件下, 高斯向量 x 二次型的条件均值为

$$E\{x^T S x | y\} = \hat{x}^T(y) S \hat{x}(y) + \text{tr} S P(\tilde{x}(y)) \tag{6.2.29}$$

式中: $\hat{x}(y)$ 为在 y 条件下 x 的条件均值, 即 y 对 x 的估计; $\tilde{x}(y) = x - \hat{x}(y)$ 为估计误差; $P(\tilde{x}(y)) = E\{[x - \hat{x}(y)][x - \hat{x}(y)]^T\}$, 为估计误差的方差阵。

6.3 离散时间 LQG 完全状态信息情况

我们通过定理来讨论离散时间线性二次型高斯随机状态模型控制系统的完全状态信息情况。

定理 6.3.1 给定系统、性能指标和容许控制分别为

$$x(k+1) = A x(k) + B u(k) + w(k) \tag{6.3.1}$$

$$J(u) = E x^T(N) Q_0 x(N) + \sum_{k=k_0}^{N-1} [x^T(k) Q_1 x(k) + u^T(k) Q_2 u(k)] \tag{6.3.2}$$

$$u(k) = u(X_k) \tag{6.3.3}$$

则使性能指标函数式(6.3.2)为极小的最优控制策略为

$$u(k) = -L(k)x(k) \tag{6.3.4}$$

式中:

$$L(k) = [A^T S(k+1) + Q_2]^{-1} A^T S(k+1) \tag{6.3.5}$$

$$\begin{aligned} S(k) &= S(k+1) + Q_1 - L^T [A^T S(k+1) + Q_2] L(k) \\ &= [A - L(k)]^T S(k+1) [A - L(k)] + Q_1 + L^T(k) Q_2 L(k) \\ &= [A - L(k)]^T S(k+1) + Q_1 \end{aligned} \tag{6.3.6}$$

$$S(N) = Q_0 \tag{6.3.7}$$

$$\min J(u) = m_0^T S(k_0) m_0 + \text{tr} S(k_0) R_0 + \sum_{k=k_0}^{N-1} \text{tr} S(k+1) R_1 \tag{6.3.8}$$

式中符号意义和假设由 6.2 节给出。

证明 我们利用贝尔曼动态规划法分步骤证明。

(1) 列写贝尔曼泛函方程

直接对性能指标函数式(6.3.2)进行处理, 把由初始时间 k_0 开始的整体性能指标函数表示为 $J(k_0)$, 那么把由 $k \in [k_0, N]$ 开始的部分性能指标函数表示为 $J(k)$, 表示式为

$$J(k) = E x^T(N) Q_0 x(N) + \sum_{l=k}^{N-1} [x^T(l) Q_1 x(l) + u^T(l) Q_2 u(l)] \tag{6.3.9}$$

我们先计算 k 的情况, 之后把 k 换成 k_0 就是整体情况。注意到 $J(k)$ 是对 $u(k), u(k+1), \dots, u(N-1)$ 取极小值, 并且是在量测 X_k 条件下取极小值, 另外取极小值符号与取均值符号可以对调, 得到

$$\min J(k) = \min_{u(k), u(k+1), \dots, u(N-1)} E x^T(N) Q_0 x(N) + \sum_{l=k}^{N-1} [x^T(l) Q_1 x(l) + u^T(l) Q_2 u(l)]$$

$$\begin{aligned}
&= E \min_{u(k), u(k+1), \dots, u(N-1)} x^T(N) Q_0 x(N) + \sum_{l=k}^{N-1} [x^T(l) Q_1 x(l) + u^T(l) Q_2 u(l)] \\
&= E \min_{u(k), u(k+1), \dots, u(N-1)} E x^T(N) Q_0 x(N) + \sum_{l=k}^{N-1} [x^T(l) Q_1 x(l) \\
&\quad + u^T(l) Q_2 u(l)] \odot \mathbf{X}_k \\
&= E V(\mathbf{X}_k, k)
\end{aligned} \tag{6.3.10}$$

式中:

$$\begin{aligned}
V(\mathbf{X}_k, k) = \min_{u(k), u(k+1), \dots, u(N-1)} E x^T(N) Q_0 x(N) + \sum_{l=k}^{N-1} [x^T(l) Q_1 x(l) \\
+ u^T(l) Q_2 u(l)] \odot \mathbf{X}_k
\end{aligned} \tag{6.3.11}$$

式(6.3.10)右边由第一式推导到第二式是通过把符号 min 和 E 对调, 第三式取了性能指标对量测量的条件均值, 使得 J(k) 成为 \mathbf{X}_k 的函数, 表示了量测量的作用。V(\mathbf{X}_k, k) 称为风险函数, 考虑到式(6.3.1)是随机差分方程, 在假设条件下具有马尔科夫性, 即未来状态对其之前状态的条件分布只是现在状态的函数, 因此在式(6.3.11)中可以用 $x(k)$ 代替 \mathbf{X}_k , 变为

$$\begin{aligned}
V(\mathbf{X}_k, k) = \min_{u(k), u(k+1), \dots, u(N-1)} E x^T(N) Q_0 x(N) + \sum_{l=k}^{N-1} [x^T(l) Q_1 x(l) \\
+ u^T(l) Q_2 u(l)] \odot \mathbf{X}(k) = V(x(k), k)
\end{aligned} \tag{6.3.12}$$

$$\min J(k) = E\{V[x(k), k]\} \tag{6.3.13}$$

整体性能指标函数为

$$\min J(k_0) = E\{V[x(k_0), k_0]\} \tag{6.3.14}$$

下面进一步把风险函数式(6.3.12)写成贝尔曼动态规划方程的形式, 把式(6.3.12)分解, 得到

$$\begin{aligned}
V[x(k), k] &= \min_{u(k)} E x^T(k) Q_1 x(k) + u^T(k) Q_2 u(k) \\
&\quad + \min_{u(k+1), u(k+2), \dots, u(N-1)} E x^T(N) Q_0 x(N) \\
&\quad + \sum_{l=k+1}^{N-1} [x^T(l) Q_1 x(l) + u^T(l) Q_2 u(l)] \odot \mathbf{x}(k+1) \odot \mathbf{x}(k) \\
&= \min_{u(k)} E x^T(k) Q_1 x(k) + u^T(k) Q_2 u(k) + V(x(k+1), k+1) \odot \mathbf{x}(k) \\
&= \min_{u(k)} x^T(k) Q_1 x(k) + u^T(k) Q_2 u(k) + E V(x(k+1), k+1) \odot \mathbf{x}(k)
\end{aligned} \tag{6.3.15}$$

上式中由第二式推导到第三式, 把 $x(k)$ 和 $u(k)$ 的二次型前的均值符号去掉, 这是因为 $x(k)$ 在 k 时刻已被量测, 是确定量, 而 $u(k)$ 是 $x(k)$ 的函数, 也是确定量。上式称为贝尔曼泛函方程。

(2) 求解贝尔曼泛函方程

贝尔曼泛函方程式(6.3.15)的解是二次型

$$V[x(k), k] = x^T(k) S(k) x(k) + s(k) \tag{6.3.16}$$

式中: $S(k)$ 和 $s(k)$ 是待定函数。

我们利用数学归纳法证明。令 l 表示时间, 若 l 为初值 N 时式(6.3.16)成立, 设 l 为 $k+1$ 时式(6.3.16)正确, 再能证明 l 为 k 时式(6.3.16)成立, 则问题得证。

初值 $l = N$ 时情况

因为 $x(N)$ 是给定的, 不受控, 所以式(6.3.15)变为

$$\begin{aligned} V[x(N), N] &= \min E \quad x^T(N) Q_0 x(N) \odot x(N) \\ &= x^T(N) Q_0 x(N) \end{aligned} \quad (6.3.17)$$

只要令

$$S(N) = Q_0 \quad s(N) = 0 \quad (6.3.18)$$

就满足式(6.3.16)了, 即在初值 $l = N$ 时式(6.3.16)成立。

$l = k+1$ 情况

$l = k+1$ 时, 有

$$V[x(k+1), k+1] = x^T(k+1) S(k+1) x(k+1) + s(k+1) \quad (6.3.19)$$

按式(6.3.15)要求, 对上式两边取条件均值, 得到

$$E \quad V[x(k+1), k+1] \Big| x(k) = E \quad x^T(k+1) S(k+1) x(k+1) \Big| x(k) + s(k+1) \quad (6.3.20)$$

把系统方程式(6.3.1)代入上式, 并注意到用量测量 $x(k)$ 对状态 $x(k)$ 的条件均值(估计)就是它自身, 条件 $x(k)$ 可去掉, 因此得到

$$\begin{aligned} E \{ V[x(k+1), k+1] \odot x(k) \} &= [\quad x(k) + \quad u(k)]^T S(k+1) [\quad x(k) + \quad u(k)] \\ &\quad + \text{tr} S(k+1) R_1 \end{aligned} \quad (6.3.21)$$

计算中利用了式(6.2.26), 其中

$$E \{ x(k+1) \} = \quad x(k) + \quad u(k) \quad (6.3.22)$$

$$\begin{aligned} R[x(k+1)] &= E \{ [x(k+1) - E \{ x(k+1) \}] [x(k+1) - E \{ x(k+1) \}]^T \} \\ &= R_1 \end{aligned} \quad (6.3.23)$$

$l = k$ 情况

把式(6.3.21)展开并代入式(6.3.15), 经整理配方, 并注意到 $S(k)$ 是非负定阵, $[\quad^T S(k+1) + Q_2]$ 是正定阵, 得到

$$\begin{aligned} V[x(k), k] &= \min_{u(k)} x^T(k) [\quad^T S(k+1) + Q_1] x(k) + u^T(k) [\quad^T S(k+1) + Q_2] u(k) \\ &\quad + x^T(k) \quad^T S(k+1) \quad u(k) + u^T(k) \quad^T S(k+1) \quad x(k) \\ &\quad + \text{tr} S(k+1) R_1 + s(k+1) \\ &= \min_{u(k)} x^T(k) S(k) x(k) \\ &\quad + [u(k) + L(k)x(k)]^T [\quad^T S(k+1) + Q_2] [u(k) + L(k)x(k)] \\ &\quad + \text{tr} S(k+1) R_1 + s(k+1) \end{aligned} \quad (6.3.24)$$

式中已令

$$L(k) = [\quad^T S(k+1) + Q_2]^{-1} \quad^T S(k+1) \quad (6.3.25)$$

$$S(k) = \quad^T S(k+1) + Q_1 - L(k) [\quad^T S(k+1) + Q_2] L(k) \quad (6.3.26)$$

$$s(k) = \text{tr}S(k+1)R_1 + s(k+1) \tag{6.3.27}$$

令

$$u(k) = -L(k)x(k) \tag{6.3.28}$$

式(6.3.24)取极小值

$$V[x(k),k] = x(k)S(k)x(k) + s(k) \tag{6.3.29}$$

由上述推导可知, $l=k$ 时式(6.3.16)成立, 同时导出最优控制策略式(6.3.28)。

式(6.3.26)就是要证明的式(6.3.6)第一式, 由第一式经简单变换可得式(6.3.6)的第二、三式。

推导中利用了 $S(k)$ 为非负定阵, $[^TS(k+1)T + Q_2]$ 为正定阵这一性质, 这容易验证。我们已经给定 Q_0 和 Q_1 为对称非负定阵, Q_2 为正定阵, 则由式(6.3.7)知 $S(N)$ 是对称非负定阵, 再由式(6.3.6)得到 $S(N-1), S(N-2), \dots, S(k), \dots$ 也是对称非负定阵。进一步可知, $^TS(k+1)$ 是非负定阵, 而 Q_2 为对称正定性, 因此 $[^TS(k+1) + Q_2]$ 为正定阵。

(3) 确定性能指标极小值

由式(6.3.14)和(6.3.16), 得到

$$\begin{aligned} \min J(k_0) &= E\{V[x(k_0), k_0]\} \\ &= E\{x^T(k_0)S(k_0)x(k_0) + s(k_0)\} \end{aligned} \tag{6.3.30}$$

利用式(6.2.26), (6.3.27)和初值条件式(6.3.18), 反复推导, 得到

$$\begin{aligned} \min J(k_0) &= m_0^TS(k_0)m_0 + \text{tr}S(k_0)R_0 + \text{tr}S(k_0+1)R_1 + s(k_0+1) \\ &= \dots \\ &= m_0^TS(k_0)m_0 + \text{tr}S(k_0)R_0 + \sum_{k=k_0}^{N-1} \text{tr}S(k+1)R_1 \end{aligned} \tag{6.3.31}$$

对上式可做如下物理解释: 在确定性情况, 性能指标极小值只取决于状态初值的二次型, m_0 表示确定性量。状态初值是随机量时, 要附加一项 $\text{tr}S(k_0)R_0$, 使性能指标变大。若系统

存在模型噪声, 还要再附加一项 $\sum_{k=k_0}^{N-1} \text{tr}S(k+1)R_1$, 使性能指标变得更大。

系统框图如图 6.3.1 所示。

图 6.3.1 完全状态信息情况系统框图

定理 6.3.1 的计算是把终值作为初值进行递归计算, 下面举一个例子说明。

例 6.3.1 考虑标量系统方程、指标函数和容许控制分别为

$$x(k+1) = x(k) + u(k) + (k) \quad (6.3.32)$$

$$J(k) = E \sum_{k=1}^N [x^2(k) + u^2(k)] \quad (6.3.33)$$

$$u(k) = u(\mathbf{X}_k) \quad (6.3.34)$$

式中: $\{(k)\}$ 为独立同分布高斯 $N(0, r)$ 随机变量序列。状态初值 $x(k_0)$ 为高斯 $N(m, \sigma^2)$ 的, 并与 (k) 互相独立。试确定达到 $\min J(k)$ 的控制策略和指标函数的极小值。当 N 时, 再确定极限控制策略。

解 由给定方程式(6.3.32)和(6.3.33)可知

$$Q_0 = Q_1 = Q_2 = 1 \quad R_1 = r, m_0 = m, R_0 = \sigma^2 \quad (6.3.35)$$

把给定数据代入定理 6.3.1 各计算公式, 得到

$$u(k) = -L(k)x(k) \quad (6.3.36)$$

$$L(k) = \frac{S(k+1)}{1 + S(k+1)} \quad (6.3.37)$$

$$S(k) = 1 + L(k) = \frac{1 + 2S(k+1)}{1 + S(k+1)} \quad (6.3.38)$$

$$S(N) = Q_0 = 1 \quad (6.3.39)$$

$$\begin{aligned} \min J(k_0) &= m^2 S(k_0) + \text{tr} S(k_0) R_0 + \sum_{k=k_0}^{N-1} \text{tr} S(k+1) R_1 \\ &= S(k_0) m^2 + S(k_0) \sigma^2 + \sum_{k=k_0}^{N-1} S(k+1) r \end{aligned} \quad (6.3.40)$$

下面利用以上公式分别计算 $k = N, N-1, \dots$ 时的各个参数。

$k = N$ 时, 不再进行控制, $u(N) = 0$ 。

$k = N-1$ 时, $L(N-1) = 0.5$, $S(N-1) = 1.5$, $u(N-1) = -0.5x(N-1)$, $\min J(N-1) = 1.5(m^2 + \sigma^2) + r$ 。

$k = N-2$ 时, $L(N-2) = 0.6$, $S(N-2) = 1.6$, $u(N-2) = -0.6x(N-2)$, $\min J(N-2) = 1.6(m^2 + \sigma^2) + 2.5r$ 。

$k = N-3$ 时, $L(N-3) = 0.615$, $S(N-3) = 1.615$, $u(N-3) = -0.615x(N-3)$, $\min J(N-3) = 1.615(m^2 + \sigma^2) + 4.1r$ 。

$k = N-4$ 时, $L(N-4) = 0.618$, $S(N-4) = 1.618$, $u(N-4) = -0.618x(N-4)$, $\min J(N-4) = 1.618(m^2 + \sigma^2) + 5.715r$ 。

$k = N-5$ 时, $L(N-5) = 0.618$, $S(N-5) = 1.618$, $u(N-5) = -0.618x(N-5)$, $\min J(N-5) = 1.618(m^2 + \sigma^2) + 7.333r$ 。

N 时, $S = \lim_{N \rightarrow \infty} S(k) = \lim_{N \rightarrow \infty} S(k+1)$, 由(6.3.38), 得 $S = 1.618$, 其他值为 $L(k) = 0.618$, $u(k) = -0.618x(k)$ 。

6.4 离散时间 LQG 不完全状态信息情况

我们通过定理来讨论离散时间线性二次型高斯随机状态模型控制系统的不完全状态信息情况。

定理 6.4.1 给定系统、性能指标和容许控制分别为

$$x(k+1) = Ax(k) + Bu(k) + w(k) \quad (6.4.1)$$

$$y(k) = Cx(k) + v(k) \quad (6.4.2)$$

$$J(u) = E x^T(N)Q_0x(N) + \sum_{k=k_0}^{N-1} [x^T(k)Q_1x(k) + u^T(k)Q_2u(k)] \quad (6.4.3)$$

$$u(k) = u(Y_{k-1}) \quad (6.4.4)$$

则使性能指标函数式(6.4.3)为极小的最优控制策略为

$$u(k) = -L(k)\hat{x}(k|k-1) \quad (6.4.5)$$

式中:

$$L(k) = [A^T S(k+1) + Q_2]^{-1} A^T S(k+1) \quad (6.4.6)$$

$$\begin{aligned} S(k) &= A^T S(k+1) + Q_1 - L^T(k)[A^T S(k+1) + Q_2]L(k) \\ &= [A - L(k)]^T S(k+1)[A - L(k)] + Q_1 - L^T(k)Q_2L(k) \\ &= [A - L(k)]^T S(k+1) + Q_1 \end{aligned} \quad (6.4.7)$$

$$S(N) = Q_0 \quad (6.4.8)$$

$$\hat{x}(k+1|k) = \hat{x}(k|k-1) + K(k|k-1)[y(k) - C\hat{x}(k|k-1)] \quad (6.4.9)$$

$$\begin{aligned} \min J(u) &= m_0^T S(k_0)m_0 + \text{tr}S(k_0)R_0 + \sum_{k=k_0}^{N-1} \text{tr}S(k+1)R_1 \\ &\quad + \sum_{k=k_0}^{N-1} \text{tr}P(k|k-1)L^T(k)[A^T S(k+1) + Q_2]L(k) \end{aligned} \quad (6.4.10)$$

式中符号意义和假设由 6.2 节给出。

本定理中, 由于增加了量测方程, 使得结果中有两式与定理 6.3.1 不同, 其余各式仍相同。在最优控制策略式中, $\hat{x}(k|k-1)$ 代替了 $x(k)$, $\hat{x}(k|k-1)$ 由式(6.4.9)表示, 其中 $K(k|k-1)$ 等量由定理 5.4.1 给出。在性能指标式中前三项相同, 但增加了第四项。

本定理讨论容许控制 $u(k) = u(Y_{k-1})$ 情况, 下一节我们讨论 $u(k) = u(Y_k)$ 情况。

证明 像定理 6.3.1 一样, 我们利用贝尔曼动态规划法分步骤证明。

(1) 列写贝尔曼泛函方程

直接对性能指标式(6.4.3)进行处理, 把由初始时刻 k_0 开始的整体性能指标表示为 $J(k_0)$, 那么由 $k \in [k_0, N]$ 开始的部分性能指标表示为

$$J(k) = E x^T(N)Q_0x(N) + \sum_{l=k}^{N-1} [x^T(l)Q_1x(l) + u^T(l)Q_2u(l)] \quad (6.4.11)$$

我们先计算 k 的情况, 再把 k 换成 k_0 就得到整体情况。注意符号 \min 和 E 可以对调, 其值不变, $J(k)$ 是对 $u(k), u(k+1), \dots, u(N-1)$ 取极小值, 并且是在量测 Y_{k-1} 条件下取极小

值,得到

$$\begin{aligned} \min J(k) &= \min_{u(k), u(k+1), \dots, u(N-1)} E \left[x^T(N) Q_0 x(N) + \sum_{l=k}^{N-1} [x^T(l) Q_1 x(l) + u^T(l) Q_2 u(l)] \right] \\ &= E \min_{u(k), u(k+1), \dots, u(N-1)} \left[x^T(N) Q_0 x(N) + \sum_{l=k}^{N-1} [x^T(l) Q_1 x(l) + u^T(l) Q_2 u(l)] \right] \\ &= E \min_{u(k), u(k+1), \dots, u(N-1)} E \left[x^T(N) Q_0 x(N) + \sum_{l=k}^{N-1} [x^T(l) Q_1 x(l) \right. \\ &\quad \left. + u^T(l) Q_2 u(l)] \mid Y_{k-1} \right] \\ &= E W(Y_{k-1}, k) \end{aligned} \tag{6.4.12}$$

式中:

$$\begin{aligned} W(Y_{k-1}, k) &= \min_{u(k), u(k+1), \dots, u(N-1)} E \left[x^T(N) Q_0 x(N) + \sum_{l=k}^{N-1} [x^T(l) Q_1 x(l) \right. \\ &\quad \left. + u^T(l) Q_2 u(l)] \mid Y_{k-1} \right] \end{aligned} \tag{6.4.13}$$

整体性能指标函数为

$$\min J(k_0) = E \{ W(Y_{k_0-1}, k_0) \} \tag{6.4.14}$$

$W(Y_{k-1}, k)$ 称为风险函数,下面进一步把式(6.4.12)写成贝尔曼动态规划方程的形式,得到

$$\begin{aligned} W(Y_{k-1}, k) &= \min_{u(k)} E \left[x^T(k) Q_1 x(k) + u^T(k) Q_2 u(k) \right. \\ &\quad \left. + \min_{u(k+1), u(k+2), \dots, u(N-1)} E \left[x^T(N) Q_0 x(N) + \sum_{l=k+1}^{N-1} [x^T(l) Q_1 x(l) \right. \right. \\ &\quad \left. \left. + u^T(l) Q_2 u(l)] \mid Y_k \right] \mid Y_{k-1} \right] \\ &= \min_{u(k)} E \left[x^T(k) Q_1 x(k) + u^T(k) Q_2 u(k) + W(Y_k, k+1) \mid Y_{k-1} \right] \end{aligned} \tag{6.4.15}$$

上式称为贝尔曼泛函方程。

(2) 求解贝尔曼泛函方程

贝尔曼泛函方程式(6.4.15)的解是二次型

$$W(Y_{k-1}, k) = \hat{x}^T(k|k-1) S(k) \hat{x}(k|k-1) + s(k) \tag{6.4.16}$$

式中: $S(k)$ 和 $s(k)$ 是待定函数。

我们利用数学归纳法证明。令 l 表示时间,若式(6.4.16)在初值 $l=N$ 时成立,设其在 $l=k+1$ 时成立,若能证明在 $l=k$ 时也成立,则问题得证。

初值 $l=N$ 时情况

直接由式(6.4.13)写出方程,并注意到符号 \min 无意义,得到

$$\begin{aligned} W(Y_{N-1}, N) &= \min E \left[x^T(N) Q_0 x(N) \mid Y_{N-1} \right] \\ &= \hat{x}^T(N|N-1) Q_0 \hat{x}(N|N-1) + \text{tr} Q_0 P(N|N-1) \end{aligned} \tag{6.4.17}$$

式中: $P(N|N-1)$ 为一步预测误差方差阵,表示为

$$\begin{aligned} P(N \otimes N - 1) &= E [x(N) - \hat{x}(N \otimes N - 1)][x(N) - \hat{x}(N \otimes N - 1)]^T \\ &= E \tilde{x}(N \otimes N - 1) \tilde{x}^T(N \otimes N - 1) \end{aligned} \quad (6.4.18)$$

只要令

$$S(N) = Q_0 \quad s(N) = \text{tr} Q_0 P(N \otimes N - 1) \quad (6.4.19)$$

就满足式(6.4.16)了,即式(6.4.16)在 $l = N$ 时成立。

$l = k + 1$ 情况

$l = k + 1$ 时,有

$$W(\mathbf{Y}_k, k + 1) = \hat{x}^T(k + 1 \otimes k) S(k + 1) \hat{x}(k + 1 \otimes k) + s(k + 1) \quad (6.4.20)$$

其中一步预测估计 $\hat{x}(k + 1 \otimes k)$ 表示为

$$\begin{aligned} \hat{x}(k + 1 \otimes k) &= \hat{x}(k \otimes k - 1) + u(k) + K(k \otimes k - 1)[y(k) - \hat{x}(k \otimes k - 1)] \\ &= \hat{x}(k \otimes k - 1) + u(k) + K(k \otimes k - 1)[\tilde{x}(k \otimes k - 1) + (k)] \end{aligned} \quad (6.4.21)$$

其均值和方差阵分别为

$$\begin{aligned} E\{\hat{x}(k + 1 \otimes k)\} &= \hat{x}(k \otimes k - 1) + u(k) \\ R[\hat{x}(k + 1 \otimes k)] &= E[\hat{x}(k + 1 \otimes k) - E\{\hat{x}(k + 1 \otimes k)\}][\hat{x}(k + 1 \otimes k) \\ &\quad - E\{\hat{x}(k + 1 \otimes k)\}]^T \\ &= E[K(k \otimes k - 1)[\tilde{x}(k \otimes k - 1) + (k)][K(k \otimes k - 1) \\ &\quad + (k)]^T \\ &= K(k \otimes k - 1)[P(k \otimes k - 1)^T + R_2]K^T(k \otimes k - 1) \end{aligned} \quad (6.4.22)$$

按式(6.4.15)要求,对式(6.4.20)两边取条件均值,并注意到 \mathbf{Y}_{k-1} 对 $\hat{x}(k \otimes k - 1)$ 的条件均值就是 $\hat{x}(k \otimes k - 1)$,得到

$$\begin{aligned} E[W(\mathbf{Y}_k, k + 1) | \mathbf{Y}_{k-1}] &= E\{W(\mathbf{Y}_k, k + 1)\} \\ &= E[\hat{x}^T(k + 1 \otimes k) S(k + 1) \hat{x}(k + 1 \otimes k) + s(k + 1)] \\ &= [\hat{x}(k \otimes k - 1) + u(k)]^T S(k + 1) [\hat{x}(k \otimes k - 1) + u(k)] \\ &\quad + \text{tr} S(k + 1) K(k \otimes k - 1) [P(k \otimes k - 1)^T + R_2] \\ &\quad + K^T(k \otimes k - 1) + s(k + 1) \end{aligned} \quad (6.4.23)$$

$l = k$ 情况

计算式(6.4.15)右边第一项

$$E[x^T(k) Q_1 x(k) | \mathbf{Y}_{k-1}] = \hat{x}^T(k \otimes k - 1) Q_1 \hat{x}(k \otimes k - 1) + \text{tr} Q_1 P(k \otimes k - 1) \quad (6.4.24)$$

其中一步预测误差方差阵 $P(k \otimes k - 1)$ 为

$$\begin{aligned} P(k \otimes k - 1) &= E[x(k) - \hat{x}(k \otimes k - 1)][x(k) - \hat{x}(k \otimes k - 1)]^T \\ &= E[\tilde{x}(k \otimes k - 1) \tilde{x}^T(k \otimes k - 1)] \end{aligned} \quad (6.4.25)$$

把式(6.4.23)和(6.4.24)代入式(6.4.15),得到

$$\begin{aligned} W(\mathbf{Y}_{k-1}, k) &= \min_{u(k)} \hat{x}^T(k \otimes k - 1) S(k) \hat{x}(k \otimes k - 1) \\ &\quad + [u(k) + L(k) \hat{x}(k \otimes k - 1)]^T [-^T S(k + 1) + Q_2] \end{aligned}$$

$$j\alpha[u(k) + L(k) \hat{x}(k|k-1)] + s(k) \quad (6.4.26)$$

式中已令

$$L(k) = [{}^T S(k+1) + Q_2]^{-1} {}^T S(k+1) \quad (6.4.27)$$

$$S(k) = {}^T S(k+1) + Q_1 - L^T(k) [{}^T S(k+1) + Q_2] L(k) \quad (6.4.28)$$

$$s(k) = \text{tr} Q_1 P(k|k-1) + \text{tr} S(k+1) K(k|k-1) [P(k|k-1)^T + R_2] \\ j\alpha K^T(k|k-1) + s(k+1) \quad (6.4.29)$$

选择

$$u(k) = -L(k) \hat{x}(k|k-1) \quad (6.4.30)$$

式(6.4.26)取极小值

$$W(\mathbf{Y}_{k-1}, k) = \hat{x}^T(k|k-1) S(k) \hat{x}(k|k-1) + s(k) \quad (6.4.31)$$

由上述推导可知, $l=k$ 时式(6.4.16)成立, 同时导出最优控制策略式(6.4.30)。

式(6.4.28)就是要证明的式(6.4.7)第一式, 由第一式经简单变换可得其余两个表示式。

像定理 6.3.1 一样, $S(k)$ 为非负定阵, $[{}^T S(k+1) + Q_2]$ 为正定阵。

(3) 确定性能指标极小值

由式(6.4.14)和(6.4.16)可得性能指标极小值

$$\min J(k_0) = E \hat{x}^T(k_0|k_0-1) S(k_0) \hat{x}(k_0|k_0-1) + s(k_0) \\ = E \hat{x}^T(k_0|k_0-1) S(k_0) E \{ \hat{x}(k_0|k_0-1) \} \\ + \text{tr} S(k_0) R [\hat{x}(k_0|k_0-1)] + s(k_0) \quad (6.4.32)$$

式中:

$$E \{ \hat{x}(k_0|k_0-1) \} = m_0 \quad (6.4.33)$$

$$R [\hat{x}(k_0|k_0-1)] = E [\hat{x}(k_0|k_0-1) - E \{ \hat{x}(k_0|k_0-1) \}] [\hat{x}(k_0|k_0-1) \\ - E \{ \hat{x}(k_0|k_0-1) \}]^T \\ = 0 \quad (6.4.34)$$

把以上两式和式(6.4.29)代入式(6.4.32), 并经反复推导, 得到

$$\min J(k_0) = m_0^T S(k_0) m_0 + s(k_0) \\ = m_0^T S(k_0) m_0 + \sum_{k=k_0}^{N-1} \text{tr} Q_1 P(k|k-1) \\ + \sum_{k=k_0}^{N-1} \text{tr} S(k+1) K(k|k-1) [P(k|k-1)^T + R_2] \\ j\alpha K^T(k|k-1) + \text{tr} Q_0 P(N|N-1) \quad (6.4.35)$$

上式是性能指标极小值的一种结果。但与式(6.4.10)不同, 前者物理意义不明显, 而后者物理意义明显, 下面证明两者是等价的。

重写式(5.4.10)和(6.4.7), 并列等式

$$P(k+1|k) = P(k|k-1)^T + R_1 - K(k|k-1) [P(k|k-1)^T + R_2] \\ j\alpha K^T(k|k-1) \quad (6.4.36)$$

$$S(k) = {}^T S(k+1) + Q_1 - L^T(k)[{}^T S(k+1) + Q_2]L(k) \quad (6.4.37)$$

$$\begin{aligned} & \text{tr}S(k+1)P(k+1\mathbb{K}) - \text{tr}P(k\mathbb{K}-1)S(k) \\ &= \text{tr}S(k+1)R_1 - \text{tr}S(k+1)K(k\mathbb{K}-1)[P(k\mathbb{K}-1)^T + R_2] \\ & \quad \text{;} K^T(k\mathbb{K}-1) - \text{tr}P(k\mathbb{K}-1)Q_1 + \text{tr}P(k\mathbb{K}-1)L^T(k) \\ & \quad \text{;} [{}^T S(k+1) + Q_2]L(k) \end{aligned} \quad (6.4.38)$$

对上式两边求和“ $\sum_{k=k_0}^{N-1}$ ”，得到

$$\begin{aligned} \text{左式} &= \sum_{k=k_0}^{N-1} [\text{tr}S(k+1)P(k+1\mathbb{K}) - \text{tr}P(k\mathbb{K}-1)S(k)] \\ &= \text{tr}S(N)P(N\mathbb{N}-1) - \text{tr}P(k_0\mathbb{K}_0-1)S(k_0) \\ &= \text{tr}Q_0P(N\mathbb{N}-1) - \text{tr}R_0S(k_0) \end{aligned} \quad (6.4.39)$$

$$\begin{aligned} \text{右式} &= \sum_{k=k_0}^{N-1} \text{tr}S(k+1)R_1 - \sum_{k=k_0}^{N-1} \text{tr}S(k+1)K(k\mathbb{K}-1)[P(k\mathbb{K}-1)^T + R_2] \\ & \quad \text{;} K^T(k\mathbb{K}-1) - \sum_{k=k_0}^{N-1} \text{tr}P(k\mathbb{K}-1)Q_1 - \sum_{k=k_0}^{N-1} \text{tr}P(k\mathbb{K}-1) \\ & \quad \text{;} L^T(k)[{}^T S(k+1) + Q_2]L(k) \end{aligned} \quad (6.4.40)$$

以上两式相等, 并经简单整理, 得到

$$\begin{aligned} & \sum_{k=k_0}^{N-1} \text{tr}Q_1P(k\mathbb{K}-1) + \sum_{k=k_0}^{N-1} \text{tr}S(k+1)K(k\mathbb{K}-1)[P(k\mathbb{K}-1)^T + R_2] \\ & \quad \text{;} K^T(k\mathbb{K}-1) + \text{tr}Q_0P(N\mathbb{N}-1) \\ &= \text{tr}S(k_0)R_0 + \sum_{k=k_0}^{N-1} \text{tr}S(k+1)R_1 + \sum_{k=k_0}^{N-1} \text{tr}P(k\mathbb{K}-1)L^T(k) \\ & \quad \text{;} [{}^T S(k+1) + Q_2]L(k) \end{aligned} \quad (6.4.41)$$

把上式代入式(6.4.35)就得到式(6.4.10)。对式(6.4.10)可做如下物理解释: 在确定性情况, $\min J(k_0)$ 为状态初值的二次型 $m_0^T S(k_0)m$; 初始状态为随机量时则附加一项 $\text{tr}S(k_0)R_0$, R_0 表示随机作用; 若存在模型噪声, 还要附加一项 $\sum_{k=k_0}^{N-1} \text{tr}S(k+1)R_1$, 它表示噪声作用; 若存在量测噪声, 就要附加最后一项 $\sum_{k=k_0}^{N-1} \text{tr}P(k\mathbb{K}-1)L^T(k)[{}^T S(k+1) + Q_2]L(k)$, $P(k\mathbb{K}-1)$ 表示量测噪声的作用。

本定理明显表示, 最优控制分为两部分, 一部分是状态估计器, 另一部分是最优控制策略, 两部分可以分别计算。充分说明分离定理的正确性。

图 6.4.1 给出了系统框图。

定理 6.4.1 的计算是把终值作为初值进行递推计算, 下面举一个例子说明。

例 6.4.1 考虑标量随机控制系统和性能指标函数分别为

图 6.4.1 不完全状态信息情况系统框图

$$x(k+1) = x(k) + 2u(k) + \omega(k) \quad (6.4.42)$$

$$y(k) = x(k) + \omega(k) \quad (6.4.43)$$

$$J(0) = E x^2(3) + \sum_{k=0}^2 u^2(k) \quad (6.4.44)$$

式中： $\omega(k)$ 和 $\omega(k)$ 分别为离散时间高斯 $N(0, 25)$ 和 $N(0, 15)$ 白噪声；给定状态初值 $x(0)$ 为高斯 $N(0, 100)$ 的； $x(0)$ ， $\omega(k)$ 和 $\omega(k)$ 互相独立。试在已知量测数据 y_0, y_1 和 y_2 时，计算有关状态估计、最优控制策略和性能指标极小值。

解 由式(6.4.44)可知，时间初值 $k_0= 0$ ，时间终值 $N= 3$ 。我们先通过定理 5.4.1 预先算出 P 和 K ，通过定理 6.4.1 预先算出 S 和 L ，然后在给定 y 时，按下列顺序计算： $\hat{x}(0|0) \rightarrow u(0) \rightarrow \hat{x}(1|0) \rightarrow u(1) \rightarrow \hat{x}(2|1) \rightarrow u(2) \rightarrow \hat{x}(3|2)$ 。

由给定系统方程和性能指标函数，有

$$\begin{aligned} & a = 1 \qquad b = 2 \qquad m_0 = 0 \qquad R_0 = 100 \qquad R_1 = 25 \\ & R_2 = 15 \qquad Q_0 = 1 \qquad Q_1 = 0 \qquad Q_2 = 1 \end{aligned} \quad (6.4.45)$$

由定理 5.4.1 列出一步预测公式

$$\begin{aligned} K(k|k-1) &= P(k|k-1)^T [P(k|k-1)^T + R_2]^{-1} \\ &= \frac{P(k|k-1)}{P(k|k-1) + 15} \end{aligned} \quad (6.4.46)$$

$$\begin{aligned} P(k+1|k) &= [1 - K(k|k-1)]P(k|k-1)[1 - K(k|k-1)]^T \\ &\quad + R_1 + K(k|k-1)R_2K(k|k-1) \\ &= 15K(k|k-1) + 25 \end{aligned} \quad (6.4.47)$$

由以上二式计算 $P(k|k-1)$ 和 $K(k|k-1)$ ，计算结果由表 6.4.1 给出。由定理 6.4.1 列出最优控制策略公式

$$L(k) = [\begin{matrix} ^T S(k+1) & + & Q_2 \end{matrix}]^{-1} \begin{matrix} ^T S(k+1) \end{matrix}$$

$$= \frac{2S(k+1)}{1+4S(k+1)} \tag{6.4.48}$$

$$S(k) = \begin{matrix} ^T S(k+1) & + & Q_1 \end{matrix} - L^T(k) [\begin{matrix} ^T S(k+1) & + & Q_2 \end{matrix}] L(k)$$

$$= \frac{S(k+1)}{1+4S(k+1)} \tag{6.4.49}$$

$$L(k) = 2S(k) \tag{6.4.50}$$

由以上二式计算 S(k)和 L(k), 计算结果由表 6.4.2 给出。

表 6.4.1 一步预测数据表

k	P (k@k- 1)	K (k@k- 1)
0	100	0. 87
1	38. 04	0. 72
2	35. 80	0. 70
3	35. 50	0. 70

表 6.4.2 最优控制策略数据表

k	S(k)	L(k)
3	1	
2	0. 20	0. 40
1	0. 11	0. 22
0	0. 08	0. 16

利用最优控制策略和一步预测公式

$$u(k) = - \begin{matrix} L(k) \end{matrix} \begin{matrix} \hat{x} (k@k- 1) \end{matrix} \tag{6.4.51}$$

$$\begin{aligned} \hat{x} (k+ 1@k) &= \begin{matrix} \hat{x} (k@k- 1) \end{matrix} + \begin{matrix} u(k) \end{matrix} + \begin{matrix} K (k@k- 1) \end{matrix} [y(k) - \begin{matrix} \hat{x} (k@k- 1) \end{matrix}] \\ &= \begin{matrix} \hat{x} (k@k- 1) \end{matrix} + 2u(k) + \begin{matrix} K (k@k- 1) \end{matrix} [y(k) - \begin{matrix} \hat{x} (k@k- 1) \end{matrix}] \end{aligned} \tag{6.4.52}$$

顺序计算各一步预测和最优控制策略:

$$\hat{x} (0@- 1) = x(0) = 0 \tag{6.4.53}$$

$$u(0) = - \begin{matrix} L(0) \end{matrix} \begin{matrix} \hat{x} (0@- 1) \end{matrix} = 0 \tag{6.4.54}$$

$$\begin{aligned} \hat{x} (1@0) &= \begin{matrix} \hat{x} (0@- 1) \end{matrix} + 2u(0) + \begin{matrix} K (0@- 1) \end{matrix} [y_0 - \begin{matrix} \hat{x} (0@- 1) \end{matrix}] \\ &= 0. 87y_0 \end{aligned} \tag{6.4.55}$$

$$u(1) = - \begin{matrix} 0. 19y_0 \end{matrix} \tag{6.4.56}$$

$$\hat{x} (2@1) = - \begin{matrix} 0. 14y_0 \end{matrix} + \begin{matrix} 0. 72y_1 \end{matrix} \tag{6.4.57}$$

$$u(2) = \begin{matrix} 0. 056y_0 \end{matrix} - \begin{matrix} 0. 29y_1 \end{matrix} \tag{6.4.58}$$

$$\hat{x} (3@2) = \begin{matrix} 0. 07y_0 \end{matrix} - \begin{matrix} 0. 36y_1 \end{matrix} + \begin{matrix} 0. 7y_2 \end{matrix} \tag{6.4.59}$$

把上述有关数据代入性能指标极小值公式得到

$$\begin{aligned} \min J(0) &= m_0^T S(0)m_0 + \begin{matrix} \text{tr}S(0)R_0 \end{matrix} + \sum_{k=0}^2 \begin{matrix} \text{tr}(k+1)R_1 \end{matrix} \\ &\quad + \sum_{k=0}^2 \begin{matrix} \text{tr}P(k@k-1)L^T(k) [\begin{matrix} ^T S(k+1) & + & Q_2 \end{matrix}] L(k) \end{matrix} \\ &= 0 + 8 + 32. 75 + 35. 38 \\ &= 76. 13 \end{aligned}$$

6.5 离散时间 LQG 控制系统小结

本节通过证明一个变分恒等式引理求解确定性、完全状态信息和不完全状态信息情况的最优控制策略问题,把各种情况联系起来。

6.5.1 变分恒等式

引理 6.5.1 给定系统

$$x(k+1) = Ax(k) + Bu(k) + w(k) \quad (6.5.1)$$

设初始条件为

$$S(N) = Q_0 \quad (6.5.2)$$

的差分方程

$$S(k) = A^T S(k+1) + Q_1 - L^T(k)[A^T S(k+1) + Q_2]L(k) \quad (6.5.3)$$

在 $k_0 \leq k \leq N$ 有非负定解,式中

$$L(k) = [A^T S(k+1) + Q_2]^{-1} A^T S(k+1) \quad (6.5.4)$$

则必存在变分恒等式

$$\begin{aligned} J(k_0) &= x^T(N)Q_0x(N) + \sum_{k=k_0}^{N-1} [x^T(k)Q_1x(k) + u^T(k)Q_2u(k)] \\ &= x^T(k_0)S(k_0)x(k_0) + \sum_{k=k_0}^{N-1} [u(k) + L(k)x(k)]^T [A^T S(k+1) + Q_2] \\ &\quad \cdot [u(k) + L(k)x(k)] + \sum_{k=k_0}^{N-1} \{ A^T(k)S(k+1)[x(k) + u(k)] \\ &\quad + [x(k) + u(k)]^T S(k+1)(k) \} + \sum_{k=k_0}^{N-1} A^T(k)S(k+1)(k) \end{aligned} \quad (6.5.5)$$

式中符号意义和假设由 6.2 节给出, $J(k_0)$ 为确定性情况性能指标,随机情况时取为 $E\{J(k_0)\}$ 。

证明 构造恒等式

$$\begin{aligned} &x^T(N)Q_0x(N) - x^T(k_0)S(k_0)x(k_0) \\ &= \sum_{k=k_0}^{N-1} [x^T(k+1)S(k+1)x(k+1) - x^T(k)S(k)x(k)] \end{aligned} \quad (6.5.6)$$

把式(6.5.1)和(6.5.3)代入上式,有

$$\begin{aligned} &x^T(N)Q_0x(N) - x^T(k_0)S(k_0)x(k_0) \\ &= \sum_{k=k_0}^{N-1} [x(k) + u(k)]^T S(k+1)[x(k) + u(k)] \\ &\quad + A^T(k)S(k+1)[x(k) + u(k)] + [x(k) + u(k)]^T S(k+1)(k) \\ &\quad + A^T(k)S(k+1)(k) - x^T(k)A^T S(k+1)x(k) - x^T(k)Q_1x(k) \\ &\quad + x^T(k)L^T(k)[A^T S(k+1) + Q_2]L(k)x(k) \end{aligned}$$

$$\begin{aligned}
&= \sum_{k=k_0}^{N-1} \mathbf{x}^T(k) \mathbf{S}(k+1) [\mathbf{x}(k) + \mathbf{u}(k)] + [\mathbf{x}(k) + \mathbf{u}(k)]^T \mathbf{S}(k+1) \mathbf{x}(k) \\
&\quad + \mathbf{x}^T(k) \mathbf{S}(k+1) \mathbf{x}(k) - \sum_{k=k_0}^{N-1} \mathbf{x}^T(k) \mathbf{Q}_1 \mathbf{x}(k) + \mathbf{u}^T(k) \mathbf{Q}_2 \mathbf{u}(k) \\
&\quad + \sum_{k=k_0}^{N-1} \mathbf{u}^T(k) [\mathbf{x}^T \mathbf{S}(k+1) + \mathbf{Q}_2] \mathbf{u}(k) + \mathbf{u}^T(k) \mathbf{x}^T \mathbf{S}(k+1) \mathbf{x}(k) \\
&\quad + \mathbf{x}^T(k) \mathbf{x}^T \mathbf{S}(k+1) \mathbf{u}(k) + \mathbf{x}^T(k) \mathbf{L}^T(k) [\mathbf{x}^T \mathbf{S}(k+1) + \mathbf{Q}_2] \mathbf{L}(k) \mathbf{x}(k)
\end{aligned} \tag{6.5.7}$$

对上式右边第二式进行 $\pm \mathbf{u}^T(k) \mathbf{Q}_2 \mathbf{u}(k)$ 的运算, 右边第三项恰好是二项式, 把它写成二项式形式。对式(6.5.7)进行整理, 就得到变分恒等式式(6.5.5)。

6.5.2 各种控制情况

1. 确定性情况

考虑确定性系统

$$\mathbf{x}(k+1) = \mathbf{x}(k) + \mathbf{u}(k) \quad (k) \quad 0 \tag{6.5.8}$$

很明显, 只有选择

$$\mathbf{u}(k) = -\mathbf{L}(k) \mathbf{x}(k)$$

时, 性能指标式(6.5.5)取极小值, 并得到

$$\min J(k_0) = \mathbf{m}_0^T \mathbf{S}(k_0) \mathbf{m}_0 \tag{6.5.9}$$

式中: \mathbf{m}_0 是 $\mathbf{x}(k)$ 的初值, 即

$$\mathbf{x}(k_0) = \mathbf{m}_0 \tag{6.5.10}$$

$\mathbf{x}(k)$ 是确定量。

2. 状态初值为随机量情况

考虑系统

$$\mathbf{x}(k+1) = \mathbf{x}(k) + \mathbf{u}(k) \quad (k) \quad 0 \tag{6.5.11}$$

其中状态初值 $\mathbf{x}(k_0)$ 为随机量, 并设其具有高斯 $N(\mathbf{m}_0, \mathbf{R}_0)$ 分布。同前, 只有选择

$$\mathbf{u}(k) = -\mathbf{L}(k) \mathbf{x}(k) \tag{6.5.12}$$

时, 性能指标式(6.5.5)取极小值, 并得到

$$\begin{aligned}
\min E\{J(k_0)\} &= E\{\mathbf{x}^T(k_0) \mathbf{S}(k_0) \mathbf{x}(k_0)\} \\
&= \mathbf{m}_0^T \mathbf{S}(k_0) \mathbf{m}_0 + \text{tr} \mathbf{S}(k_0) \mathbf{R}_0
\end{aligned} \tag{6.5.13}$$

3. 完全状态信息情况

考虑系统

$$\mathbf{x}(k+1) = \mathbf{x}(k) + \mathbf{u}(k) + \mathbf{w}(k) \tag{6.5.14}$$

式中符号意义与假设由 6.2 节给出。我们设 $\mathbf{u}(k) = \mathbf{u}(\mathbf{X}_k)$, 那么性能指标是在 \mathbf{X}_k 给定条件下求得的, 再注意到系统是马尔科夫过程, 则有

$$E\{J(k_0)\} = E\{\mathbf{y}^T \mathbf{X}_k\} = E\{\mathbf{y}^T \mathbf{x}(k)\} \tag{6.5.15}$$

式中用“ \cdot ”简化表示式(6.5.5)。我们还注意到状态 $\mathbf{x}(k)$ 既是被控量又是量测量, $\mathbf{x}(k)$ 成

为已知量, 而性能指标函数式(6.5.5)是状态 $x(k)$ 的函数, 因此在 $x(k)$ 条件下对 $x(k)$ 取均值仍为它自身, 即

$$E\{J(k_0)\} = E\{j\} \tag{6.5.16}$$

只有选择

$$u(k) = -L(k)x(k) \tag{6.5.17}$$

时, 性能指标式(6.5.16)取极小值, 并为

$$\begin{aligned} \min E\{J(k_0)\} &= E\{x^T(k_0)S(k_0)x(k_0) + \sum_{k=k_0}^{N-1} x^T(k)S(k+1)x(k) \\ &\quad + \sum_{k=k_0}^{N-1} u^T(k)R_1u(k) + x^T(N)S(N)x(N)\} \\ &= m_0^T S(k_0)m_0 + \text{tr}S(k_0)R_0 + \sum_{k=k_0}^{N-1} \text{tr}S(k+1)R_1 \end{aligned} \tag{6.5.18}$$

4. 不完全状态信息情况

考虑系统

$$x(k+1) = Ax(k) + Bu(k) + w(k) \tag{6.5.19}$$

$$y(k) = Cx(k) + v(k) \tag{6.5.20}$$

式中符号意义和假设由6.2节给出。按容许控制分两种情况。

- (1) $u(k) = u(\hat{Y}_k)$ 情况
- 这时, 性能指标为

$$E\{J(k_0)\} = E\{j(\hat{Y}_k)\} \tag{6.5.21}$$

为求出最优控制策略, 对上式第二项进行处理:

$$\begin{aligned} E\{[u(k) + L(k)x(k)]^T [{}^T S(k+1) + Q_2][u(k) + L(k)x(k)] \hat{Y}_k\} \\ = E\{[u(k) + L(k)\hat{x}(k|k)]^T [{}^T S(k+1) + Q_2][u(k) + L(k)\hat{x}(k|k)] \\ + \text{tr}P(k|k)L^T(k)[{}^T S(k+1) + Q_2]L(k)\} \end{aligned} \tag{6.5.22}$$

式中利用了

$$\begin{aligned} E\{[u(k) + L(k)x(k)] \hat{Y}_k\} &= u(k) + L(k)\hat{x}(k|k) \\ R[u(k) + L(k)x(k) \hat{Y}_k] \\ &= E\{[u(k) + L(k)x(k) - (u(k) + L(k)\hat{x}(k|k))] \\ &\quad + [u(k) + L(k)\hat{x}(k|k)]\} \\ &= E\{L(k)[x(k) - \hat{x}(k|k)][x(k) - \hat{x}(k|k)]^T L^T(k)\} \\ &= L(k)E\{\tilde{x}(k|k)\tilde{x}^T(k|k)\}L^T(k) \\ &= L(k)P(k|k)L^T(k) \end{aligned} \tag{6.5.24}$$

把式(6.5.22)代入式(6.5.21), 性能指标函数只有选择

$$u(k) = -L(k)\hat{x}(k|k) \tag{6.5.25}$$

条件下才取极小值, 并且极小值为

$$\begin{aligned} \min E\{J(k_0)\} &= m_0^T S(k_0)m_0 + \text{tr}S(k_0)R_0 + \sum_{k=k_0}^{N-1} \text{tr}S(k+1)R_1 \\ &\quad + \sum_{k=k_0}^{N-1} \text{tr}P(k|k)L^T(k)[{}^T S(k+1) + Q_2]L(k) \end{aligned} \tag{6.5.26}$$

(2) $u(k) = u(\hat{Y}_{k-1})$ 情况

处理方法与(1)相同, 得到最优控制策略和性能指标极小值分别为

$$\begin{aligned}
 u(k) &= -L(k)\hat{x}(k|k-1) \quad (6.5.27) \\
 \min E J(k_0) &= m_0^T S(k_0)m_0 + \text{tr}S(k_0)R_0 + \sum_{k=k_0}^{N-1} \text{tr}S(k+1)R_1 \\
 &\quad + \sum_{k=k_0}^{N-1} \text{tr}P(k|k-1)L^T(k)[S(k+1) + Q_2]L(k) \quad (6.5.28)
 \end{aligned}$$

由上述结果可知, 在各种情况下, $L(k)$, $S(k)$ 和 $S(N) = Q_0$ 都相同。最优控制策略 $u(k)$ 都由状态反馈实现, 但所用状态不同, 确定性情况、状态初值为随机量情况和完全状态信息情况下, 所用状态都为 $x(k)$, 只有不完全状态信息情况下所用状态为 $\hat{x}(k|k)$ 或 $\hat{x}(k|k-1)$ 。性能指标函数在各种情况下也不相同, 物理意义明显, 随着噪声项增多, 会使性能指标逐项增加。

6.6 离散时间 LQG 最优跟随器

前两节中, 我们讨论了离散时间 LQG 最优调节器问题, 这一节我们介绍最优跟随器 (随动系统, 跟踪器) 问题。

6.6.1 问题的描述

1. 受控系统

设受控系统和受控向量为

$$x_c(k+1) = a_c x_c(k) + b_c u(k) + c_c(k) \quad (6.6.1)$$

$$y_c(k) = c_c x_c(k) + d_c(k) \quad (6.6.2)$$

$$z_c(k) = C_c x_c(k) \quad (6.6.3)$$

式中符号意义和假设由 6.2 节给出, 不过加了下标“c”。受控向量 $z_c(k)$ 可以是状态向量, 也可以是系统输出, 甚至其他量, C_c 为系数阵。

2. 参考系统

设参考系统和参考向量为

$$x_d(k+1) = a_d x_d(k) + b_d(k) \quad (6.6.4)$$

$$y_d(k) = c_d x_d(k) + d_d(k) \quad (6.6.5)$$

$$z_d(k) = C_d x_d(k) \quad (6.6.6)$$

同样, 式中符号意义和假设由 6.2 节给出, 只是加了下标“d”。参考向量 $z_d(k)$ 可以是状态向量, 也可以是系统输出, 甚至其他量, C_d 为系数阵。

参考向量 z_d 是一个标准, 要求受控向量 z_d 尽量逼近 z_c , 使误差 $e(k)$

$$e(k) = z_d - z_c \quad (6.6.7)$$

尽量小, 使受控系统运动轨迹尽量逼近参考系统轨迹。

我们假设 $x_c(k)$, $x_d(k)$, $u_c(k)$ 和 $u_d(k)$ 互相独立。

3. 性能指标

这里仍采用二次型性能指标

$$J(k_0) = E \{ e^T(N) Q_0 e(N) + \sum_{k=k_0}^{N-1} e^T(k) Q_1 e(k) + u^T(k) Q_2 u(k) \} \tag{6.6.8}$$

上式与式(6.2.11)相比, 只是用 $e(k)$ 和 $e(N)$ 取代了 $x(k)$ 和 $x(N)$, 更为一般化。当 $z_c(k)$ 表示状态时, 则系统称为状态跟随器; 当 $z_c(k)$ 表示输出时, 则系统称为输出跟随器。

6.6.2 最优控制策略算法

首先采用扩充状态向量法把“ c ”和“ d ”系统组成一个系统。做法如下, 令

$$\begin{aligned} x(k) &= \begin{bmatrix} x_c(k) \\ x_d(k) \end{bmatrix} = \begin{bmatrix} x_c(k) & 0 \\ 0 & x_d(k) \end{bmatrix} = \begin{bmatrix} x_c(k) \\ x_d(k) \end{bmatrix} \\ y(k) &= \begin{bmatrix} y_c(k) \\ y_d(k) \end{bmatrix} = \begin{bmatrix} y_c(k) & 0 \\ 0 & y_d(k) \end{bmatrix} = \begin{bmatrix} y_c(k) \\ y_d(k) \end{bmatrix} \end{aligned} \tag{6.6.9}$$

于是新系统为

$$x(k+1) = A x(k) + B u(k) + w(k) \tag{6.6.10}$$

$$y(k) = C x(k) + v(k) \tag{6.6.11}$$

有关参数为

$$\begin{aligned} R_1 &= \begin{bmatrix} R_{1c} & 0 \\ 0 & R_{1d} \end{bmatrix} \quad R_2 = \begin{bmatrix} R_{2c} & 0 \\ 0 & R_{2d} \end{bmatrix} \\ E\{x(k_0)\} &= \begin{bmatrix} E\{x_c(k_0)\} \\ E\{x_d(k_0)\} \end{bmatrix} = \begin{bmatrix} m_{0c} \\ m_{0d} \end{bmatrix} = m_0 \\ R[x(k_0)] &= \begin{bmatrix} R_{0c} & 0 \\ 0 & R_{0d} \end{bmatrix} = R_0 \\ e(k) &= (C_c - C_d) x(k) = C x(k) \end{aligned} \tag{6.6.12}$$

性能指标函数改写为

$$\begin{aligned} J(k_0) &= E \{ x^T(N) C^T Q_0 C x(N) + \sum_{k=k_0}^{N-1} [x^T(k) C^T Q_1 C x(k) + u^T(k) Q_2 u(k)] \} \\ &= E \{ x^T(N) Q_0 x(N) + \sum_{k=k_0}^{N-1} [x^T(k) Q_1 x(k) + u^T(k) Q_2 u(k)] \} \end{aligned} \tag{6.6.14}$$

式中:

$$Q_0 = C^T Q_0 C \quad Q_1 = C^T Q_1 C \quad C = C_c - C_d \tag{6.6.15}$$

经上述扩充状态向量法把跟随器问题转化为由式(6.6.10), (6.6.11) 和(6.6.14) 表述的最优调节器问题。选容许控制 $u(k) = u(Y_k)$, 则得到最优跟随控制策略

$$\begin{aligned} u(k) &= -L(k) \hat{x}(k|k) \\ &= -L_c \hat{x}_c(k|k) - L_d \hat{x}_d(k|k) \end{aligned} \tag{6.6.16}$$

式中 $L(k)$ 及 $S(k)$ 表示式同前, 为

$$L(k) = [-^T S(k+1) + Q_2]^{-1} ^T S(k+1) \tag{6.6.17}$$

$$S(k) = [- - L(k)]^T S(k+1) [- - L(k)] + Q_1 + L^T(k) Q_2 L(k) \tag{6.6.18}$$

$$S(N) = Q_0 \tag{6.6.19}$$

$\hat{x}(k|k)$ 由下列滤波公式(同前)给出:

$$\begin{aligned} \hat{x}(k+1|k+1) = & \hat{x}(k|k) + u(k) + K(k+1|k+1) \\ & ;\hat{y}[y(k+1) - u(k) - \hat{x}(k|k)] \end{aligned} \tag{6.6.20}$$

$$K(k+1|k+1) = P(k+1|k) ^T [P(k+1|k) ^T + R_2]^{-1} \tag{6.6.21}$$

$$P(k+1|k) = P(k|k) ^T + R_1 \tag{6.6.22}$$

$$P(k+1|k+1) = [I - K(k+1|k+1)] P(k+1|k) \tag{6.6.23}$$

$$\hat{x}(k_0|k_0) = m_0 \qquad P(k_0|k_0) = R_0 \tag{6.6.24}$$

极小性能指标为

$$\begin{aligned} \min J(k_0) = & m_0^T S(k_0) m_0 + \sum_{k=k_0}^{N-1} \text{tr} S(k+1) R_1 \\ & + \sum_{k=k_0}^{N-1} \text{tr} P(k|k) L^T(k) [-^T S(k+1) + Q_2] L(k) \end{aligned} \tag{6.6.25}$$

为进一步了解最优控制算法的结构, 把上述结果展开, 写成

$$L_c(k) = [-^T S_{11}(k+1) _c + Q_2]^{-1} _c S_{11}(k+1) _c \tag{6.6.26}$$

$$L_d(k) = [-^T S_{11}(k+1) _c + Q_2]^{-1} ^T_d S_{12}(k+1) _d \tag{6.6.27}$$

$$\begin{aligned} S_{11}(k) = & [_c - _c L_c(k)]^T S_{11}(k+1) [_c - _c L_c(k)] + C_c^T Q_{1c} C_c + L_c^T(k) Q_2 L_c(k) \\ & \tag{6.6.28} \end{aligned}$$

$$\begin{aligned} S_{12}(k) = & - [_c - _c L_c(k)]^T [S_{11}(k+1) _c L_d(k) - S_{12}(k+1) _d] \\ & - C_c^T Q_{1c} C_d + L_c^T(k) Q_2 L_d(k) \end{aligned} \tag{6.6.29}$$

$$\begin{aligned} S_{22}(k) = & L_d^T(k) ^T_c [S_{11}(k+1) _c L_d(k) - S_{12}(k+1) _d] + _d [S_{12}^T(k+1) _c \\ & - S_{22}(k+1) _d] + C_d^T Q_{1c} C_d + L_d^T Q_2 L_d(k) \end{aligned} \tag{6.6.30}$$

$$\begin{aligned} S_{11}(N) = & C_c^T(N) Q_{0c} C_c(N) \qquad S_{12}(N) = - C_c^T(N) Q_{0c} C_d(N) \\ S_{22}(N) = & C_d^T(N) Q_{0c} C_d(N) \end{aligned} \tag{6.6.31}$$

$$\begin{aligned} \hat{x}_c(k+1|k+1) = & _c \hat{x}_c(k|k) + u(k) + K_c(k+1|k+1) \\ & ;\hat{y}[y_c(k+1) - _c u(k) - _c _c \hat{x}_c(k|k)] \end{aligned} \tag{6.6.32}$$

$$K_c(k+1|k+1) = P_c(k+1|k) ^T_c [_c P(k+1|k) ^T_c + R_{2c}]^{-1} \tag{6.6.33}$$

$$P_c(k+1|k) = _c P_c(k|k) ^T_c + R_{1c} \tag{6.6.34}$$

$$P_c(k+1|k+1) = [I - K_c(k+1|k+1) _c] P_c(k+1|k) \tag{6.6.35}$$

$$\hat{x}_c(k_0|k_0) = m_{0c} \qquad P_c(k_0|k_0) = R_{0c} \tag{6.6.36}$$

$$\begin{aligned} \hat{x}_d(k+1|k+1) = & _d \hat{x}(k|k) + K_d(k+1|k+1) [y_d(k+1) - _d _d \hat{x}_d(k|k)] \\ & \tag{6.6.37} \end{aligned}$$

$$K_d(k+1|k+1) = P_d(k+1|k) ^T_d [_d P_d(k+1|k+1) + R_{2d}]^{-1} \tag{6.6.38}$$

$$P_d(k+1) = P_d(k) + R_{1d} \quad (6.6.39)$$

$$P_d(k+1) = [I - K_d(k+1)] P_d(k) \quad (6.6.40)$$

$$\hat{x}(k_0) = m_{0d} \quad P_d(k_0) = R_{0d} \quad (6.6.41)$$

由分解式(6.6.26) ~ (6.6.41)可知:

(1) 原系统控制策略矩阵 $L_c(k)$ 与参考系统无关;但参考系统控制策略矩阵 $L_d(k)$ 与原系统有关。

(2) 在设原系统与参考系统互相独立条件下,两系统的滤波计算可分开进行。

把上述结果总结为定理形式:

定理 6.6.1 由受控系统式(6.6.1) ~ (6.6.3)和参考系统式(6.6.4) ~ (6.6.6)联合组成的跟随器,可通过扩充状态向量法构成由式(6.6.10)和(6.6.11)表示的新系统,则使性能指标函数式(6.6.8)为极小的最优控制策略为式(6.6.16),极小性能指标为式(6.6.25),其中最优滤波由式(6.6.20)给出。最优控制策略和最优滤波的分解形式由式(6.6.16) ~ (6.6.41)给出。

6.7 连续时间确定性最优控制

前几节中讨论了离散时间 LQG 最优控制问题,后几节我们将讨论连续时间 LQG 最优控制问题。我们先介绍确定性最优控制,再求解完全状态和不完全状态信息情况。

讨论连续时间最优控制问题可直接在连续时间下求解,也可通过把离散时间最优控制问题对时间取极限化来求解,我们采用前者。主要方法采用贝尔曼动态规划法,先讨论确定性最优控制,再讨论随机问题。

本节先介绍连续时间确定性问题的—般贝尔曼方法,再求解连续时间线性确定性最优控制问题。

6.7.1 —般贝尔曼方程

连续时间确定性控制系统由微分方程

$$\dot{x}(t) = f(t, x, u) \quad (6.7.1)$$

描述,性能指标函数为

$$J(u, [0, T]) = \int_0^T h(x, u) dt + g[x(T)] \quad (6.7.2)$$

式中: t 表示时间,“0”为初始时间, T 为终端时间; $x = x(t) = x_t$ 表示 $n \times 1$ 维状态向量; $u = u(t, x) \in U$ 表示 $k \times 1$ 控制向量; U 表示控制向量集合; $h(x, u)$ 通常表示偏离 $\{x(t)\}$ 某个轨道所花费的费用; $g[x(T)]$ 表示终端时间失败的费用。

所谓控制问题就是对系统式(6.7.1)寻求一种控制 u , 使性能指标式(6.7.2)达到极小。求出的 u 称为最优控制,一般表示为状态反馈 $u = -kx$, 其中 k 称为控制增益。令性能指标函数极小值为

$$V(t, x) = \min_{u[t, T]} J(u[t, T]) \quad (6.7.3)$$

则最优控制问题可表示为

$$V(t, x) = \min_{u[t, T]} \int_t^T h[x(s), u(s)] ds + g[x(T)] \quad (6.7.4)$$

$$s. t. \quad \dot{x} = f(x, u) \quad x = x(t) = x_0 \text{ (初值)}$$

我们利用贝尔曼最优性原理列出动态规划方程求解, 任取 $[0, T - \Delta t]$, 则有

$$\begin{aligned} V(t, x) &= \min_{u[t, T]} \int_t^{t+\Delta t} h[x(s), u(s)] ds + \int_{t+\Delta t}^T h[x(s), u(s)] ds + g[x(T)] \\ &= \min_{u[t, t+\Delta t]} \int_t^{t+\Delta t} h[x(s), u(s)] ds + \min_{u[t+\Delta t, T]} \int_{t+\Delta t}^T h[x(s), u(s)] ds + g[x(T)] \\ &= \min_{u[t, t+\Delta t]} \int_t^{t+\Delta t} h[x(s), u(s)] ds + V[t+\Delta t, x(t+\Delta t)] \end{aligned} \quad (6.7.5)$$

设 $V(t, x)$ 存在, 且对 (t, x) 可微, 则 $V(t, x)$ 有泰勒级数展开式

$$\begin{aligned} V[t+\Delta t, x(t+\Delta t)] &= V(t, x) + \frac{\partial V(t, x)}{\partial t} \Delta t + \frac{\partial V(t, x)}{\partial x} \Delta x + o(\Delta t, \Delta x) \\ &= V(t, x) + \frac{\partial V(t, x)}{\partial t} \Delta t + \frac{\partial V(t, x)}{\partial x} f(x, u) \Delta t + o(\Delta t) \end{aligned} \quad (6.7.6)$$

式中由第一式到第二式利用了系统方程 $\dot{x} = f(t, u)$ 。另外有 $\frac{\partial V}{\partial x} = V_x$, $\frac{\partial V}{\partial t} = V_t$ 。

把式(6.7.6)代入式(6.7.5), 并且由 $\int_t^{t+\Delta t} h[x(s), u(s)] ds = h(t, x, u) \Delta t$, 得到

$$V(t, x) = \min_u \int_t^{t+\Delta t} h(t, x, u) ds + V(t, x) + \frac{\partial V(t, x)}{\partial t} \Delta t + \frac{\partial V(t, x)}{\partial x} f(x, u) \Delta t + o(\Delta t) \quad (6.7.7)$$

上式两边用 Δt 除, 并令 Δt 趋于 0, 得到

$$-\frac{\partial V(t, x)}{\partial t} + \min_u \frac{\partial V(t, x)}{\partial x} f(x, u) + h(x, u) = 0 \quad (6.7.8)$$

并有初始条件(把终端作为初值)

$$V(T, x) = g[x(T)] \quad (6.7.9)$$

式(6.7.8)和式(6.7.9)称为贝尔曼方程, 或称为哈密顿-雅可比方程, 它是 $V(t, x)$ 的非线性偏微分方程, $V(t, x)$ 称为风险函数。贝尔曼方程的重要性在于, 它把最优控制的极小化问题化为逐点极小化问题, 简化了计算。

6.7.2 线性贝尔曼方程

对线性确定性最优调节问题, 是把一般系统式(6.7.1)线性化为

$$\dot{x} = Ax + Bu \quad x(0) = x_0 \quad (6.7.10)$$

性能指标函数一般采用 x 和 u 的二次型形式

$$h(x, u) = x^T(t) Q_1 x(t) + u^T(t) Q_2 u(t) \quad (6.7.11)$$

$$g[x(T)] = x^T(T) Q_0 x(T) \quad (6.7.12)$$

式中: A 为 $n \times n$ 系数矩阵; B 为 $n \times 1$ 控制矩阵; Q_0 , Q_1 和 Q_2 分别为 $n \times n$, $n \times n$ 和 1×1 加权方阵。各矩阵元素要在优质调节性能与低能量消耗之间进行折衷选择。只要 $\Delta x(t)$ 和 $\Delta u(t)$ 很小, 二次型函数可以逼近任何性能指标函数, 因此把二次型函数形式选为性能指标函数是合适的。于是, 线性调节问题可以写为

$$V(t, x) = \min_{u[t, T]} \int_t^T [x^T(s) Q_1 x(s) + u^T(s) Q_2 u(s)] ds + x^T(T) Q_0 x(T) \quad (6.7.13)$$

$$\text{s. t.} \quad \dot{x} = Ax + Bu \quad x = x_1(\text{初值})$$

把式(6.7.10) ~ (6.7.12)代入式(6.7.8)和(6.7.9),得到线性调节问题式(6.7.13)的贝尔曼方程

$$V_t + \min_u \dot{x}^T (Ax + Bu) + x^T Q_1 x + u^T Q_2 u = 0 \quad (6.7.14)$$

$$V(T, x) = x^T(T) Q_0 x(T) \quad (6.7.15)$$

对式(6.7.14)加以整理得到

$$V_t + V_x^T Ax + x^T Q_1 x + \min_u V_x^T Bu + u^T Q_2 u = 0 \quad (6.7.16)$$

先处理上式取极小值部分,考虑到上式是标量计算,有

$$V_x^T Bu = V_x^T Bu^T = u^T B^T V_x \quad (6.7.17)$$

$$\begin{aligned} \min_u V_x^T Bu + u^T Q_2 u &= \min_u u^T Q_2 u + \frac{1}{2} V_x^T Bu + \frac{1}{2} u^T B^T V_x \\ &= \min_u u + \frac{1}{2} Q_2^{-1} B^T V_x^T Q_2 u + \frac{1}{2} Q_2^{-1} B^T V_x - \frac{1}{4} V_x^T B Q_2^{-1} B^T V_x \\ &= -\frac{1}{4} V_x^T B Q_2^{-1} B^T V_x \end{aligned} \quad (6.7.18)$$

式中取极小值的条件为

$$u = -\frac{1}{2} Q_2^{-1} B^T V_x \quad (6.7.19)$$

把式(6.7.18)代入式(6.7.16),贝尔曼方程变为

$$V_t + V_x^T Ax + x^T Q_1 x - \frac{1}{4} V_x^T B Q_2^{-1} B^T V_x = 0 \quad (6.7.20)$$

$$V(T, x) = x^T Q_0 x$$

这个二次型方程有如下形式的解:

$$V(t, x) = x^T S(t) x + s(t) \quad (6.7.21)$$

式中: $S(t)$ 和 $s(t)$ 为待定函数,且 $S(t)$ 为对称阵。下面对是否存在式(6.7.21)形式的解做一简单证明。由式(6.7.21)求 V_t 和 V_x , 得到

$$V_t(t, x) = \dot{x}^T S(t) x + s(t) \quad (6.7.22)$$

$$V_x(t, x) = 2Sx \quad (6.7.23)$$

把式(6.7.22)和(6.7.23)代入式(6.7.20),并注意到 $x^T A^T S x = x^T S A x$, 得到

$$x^T (S + Q_1 + SA + A^T S - SBQ_2^{-1} B^T S) x + s(t) = 0 \quad (6.7.24)$$

如果令

$$-S = SA + A^T S + Q_1 - SBQ_2^{-1} B^T S \quad (6.7.25)$$

$$s(t) = 0 \quad (6.7.26)$$

$$S(T) = Q_0 \quad (6.7.27)$$

则使式(6.7.21)成立,证明式(6.7.21)是式(6.7.20)的解,其中 S 满足里卡蒂方程式(6.7.25)和(6.7.27)。把式(6.7.23)代入式(6.7.19),得到最优控制为状态的线性函数

$$u(t, x) = - Q_2^{-1} B^T S(t) x(t) \quad (6.7.28)$$

性能指标极小值为

$$V[0, x(0)] = \min_{u \in U} J(u[0, T]) = x^T(0) S(0) x(0) \quad (6.7.29)$$

式中: $x(0) = x_0$ 为给定初值。

在完成连续时间线性确定性最优控制计算后, 让我们简单讨论一个理论问题。假设存在对 (t, x) 可微的函数 $V(t, x)$, 满足贝尔曼方程式(6.7.8)和(6.7.9), 并在 $u^0 \in U$ 时达到极小, 即

$$V_x(t, x) f(x, u^0) + h(x, u^0) \leq V_x(t, x) f(x, u) + h(x, u) \quad (6.7.30)$$

那么 u^0 是最优的, 也就是说贝尔曼方程的解是唯一的。

下面做一个简单证明。设 $u \in U$ 是任意控制, 那么由式(6.7.8)得到

$$V_t(t, x) + V_x(t, x) f(x, u) + h(t, x, u) = 0 \quad (6.7.31)$$

上式前两项是 $V(t, x)$ 对 t 的全微分, 因此, 有

$$\begin{aligned} \frac{d}{dt} V(t, x) &= V_t(t, x) + V_x(t, x) \dot{x} \\ &= V_x(t, x) + V_x(t, x) f(x, u) \\ &\quad - h(x, u) \end{aligned} \quad (6.7.32)$$

选择时间 $t_0 = 0, t_T = T$, 对上式积分, 得到

$$V[T, x(T)] - V[0, x(0)] = - \int_0^T h[x(s), u(s)] ds \quad (6.7.33)$$

由于 $V[T, x(T)] = g[x(T)]$, 得到

$$\int_0^T h[x(s), u(s)] ds + g[x(T)] = V[0, x(0)] \quad (6.7.34)$$

我们已经看到, 当把 u^0 代入上式时, 得到等式

$$\int_0^T h[x(s), u^0(s)] ds + g[x(T)] = V[0, x(0)] \quad (6.7.35)$$

并给出性能指标极小值

$$\min J(u^0) = V[0, x(0)] \quad (6.7.36)$$

由上式和式(6.7.34)可知, u^0 是最优的。并且可以看到, u^0 至多不过满足这些条件, 因此, 贝尔曼方程的解是唯一的。上述结论是一般性的, 当然适用于连续时间线性确定性最优控制问题。

但上述结论只说明了充分条件, 而不是必要条件, 动态规划法的主要缺点就在这一点。在很多问题中, V 是不可微的, 因此贝尔曼方程没有解。在这种情况下, 可供选择的办法就是设法得到更多的信息, 如借助极大值原理导出优化的必要条件。

另外还可以证明, 当线性系统状态方程

$$\dot{x}(t) = Ax(t) + Bu(t, x) \quad x(0) = x_0 \quad (6.7.37)$$

有唯一解, 并且施加控制 $u(t, x)$ 满足利普希茨(Lipschitz)条件, 即对某个常数 K , 有不等式

$$\|u(s, x) - u(s, y)\| \leq K \|x - y\| \quad (6.7.38)$$

则最优控制策略 $u(t, x)$ 为状态向量 $x(t)$ 的线性函数。

最后把上述计算结果总结为定理形式:

定理 6.7.1 由式(6.7.13)描述的连续时间线性确定性控制系统可表征为贝尔曼方程, 如果其中的 V 函数存在且可微, 则使性能指标为极小的最优控制式为(6.7.28), 其中 $S(t)$ 由里卡蒂方程式(6.7.25)给出, 性能指标极小值为式(6.7.29)。

例 6.7.1 给定系统方程和性能指标函数分别为

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \quad x(t_0) = m_0 = \begin{bmatrix} m_{10} \\ m_{20} \end{bmatrix} \quad (6.7.39)$$

$$J(u) = \int_{t_0}^{t_1} [x_1^2 + qu^2] dt \quad (6.7.40)$$

求最优控制。当 $t_0 \rightarrow \infty$ 时, 再求极限情况。

解 此例中,

$$\begin{aligned} A &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} & B &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} & Q_0 &= 0 \\ Q_1 &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} & Q_2 &= q \end{aligned} \quad (6.7.41)$$

令 $S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$, S 是对称阵, 因此有 $S_{12} = S_{21}$ 。把给定数据式(6.7.41)代入里卡蒂方程式(6.7.25), 给出

$$\begin{aligned} - \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} + \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\ - \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & q \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \end{aligned} \quad (6.7.42)$$

$$S(t_1) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (6.7.43)$$

对上式计算和整理, 得到三个方程:

$$S_{11} = \frac{1}{q} S_{12}^2 - 1 \quad S_{11}(t_1) = 0 \quad (6.7.44)$$

$$S_{12} = \frac{1}{q} S_{12} S_{22} - S_{11} \quad S_{12}(t_1) = S_{21}(t_1) = 0 \quad (6.7.45)$$

$$S_{22} = \frac{1}{q} S_{22}^2 - 2S_{12} \quad S_{22}(t_1) = 0 \quad (6.7.46)$$

解出以上三个方程, 得到 S_{11} , S_{12} 和 S_{22} 。由式(6.7.28)计算最优控制 $u(t)$ 为

$$\begin{aligned} u(t) &= - \frac{1}{q} \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= - \frac{1}{q} S_{12} x_1(t) - \frac{1}{q} S_{22} x_2(t) \end{aligned} \quad (6.7.47)$$

式中 $x_1(t)$ 和 $x_2(t)$ 通过求解式(6.7.39)得到。由式(6.7.29)计算性能指标极小值

$$\min J = \begin{bmatrix} m_{10} & m_{20} \end{bmatrix} \begin{bmatrix} S_{11}(t_0) & S_{12}(t_0) \\ S_{21}(t_0) & S_{22}(t_0) \end{bmatrix} \begin{bmatrix} m_{10} \\ m_{20} \end{bmatrix} \quad (6.7.48)$$

下面求解 $t_0 = \infty$ 时的情况。 $t_0 = \infty$ 时, 得到定态, $S = 0$, 里卡蒂方程式(6.7.44) ~ (6.7.46) 退化为代数方程, 其解 S 为

$$S = \frac{\overline{2} \overline{q}}{\overline{q}} + \frac{\overline{q}}{2q \overline{q}} \tag{6.7.49}$$

把上式代入式(6.7.47), 得到最优控制

$$u(t) = - \frac{1}{q} \overline{q} x_1(t) + \frac{\overline{q}}{2q \overline{q}} x_2(t) \tag{6.7.50}$$

把式(6.7.49)代入式(6.7.48), 得到性能指标极小值

$$\min J = \overline{2} \overline{q} m_{10}^2 + 2 \overline{q} m_{10} m_{20} + \overline{2q} \overline{q} m_{20}^2 \tag{6.7.51}$$

6.8 连续时间 LQG 完全状态信息情况

我们先给出系统模型, 再列出贝尔曼方程, 最后给出最优控制解。

6.8.1 系统模型

连续时间随机状态模型 LQG 完全状态信息情况给定为

$$dx(t) = Ax(t)dt + Bu(t) + dw(t) \tag{6.8.1}$$

式中: $x(t)$ 为 $m \times 1$ 随机状态向量, 状态初值 $x(t_0)$ 为高斯 $N(m_0, R_0)$ 向量; $u(t)$ 为 $k \times 1$ 控制向量; $dw(t)$ 为 $n \times 1$ 增量维纳过程 $N(0, R_1 dt)$ 向量; A 为 $n \times n$ 系数矩阵; B 为 $n \times 1$ 控制矩阵。 $x(t_0)$ 与 $dw(t)$ 互相独立。由于状态向量 $x(t)$ 是高斯向量组合, 因此也为高斯过程向量。

性能指标函数为

$$J(u[0, T]) = E \int_0^T x^T(t)Q_1x(t) + u^T(t)Q_2u(t) dt + x^T(T)Q_0x(T) \tag{6.8.2}$$

式中: Q_1 和 Q_0 为对称非负定阵, Q_2 为对称正定阵。

我们规定容许控制 $u(t)$ 是状态向量 $x(t)$ 的线性函数, 即

$$u[t, x(t)] = K(t)x(t) \quad (K \in \mathbf{K}) \tag{6.8.3}$$

式中: \mathbf{K} 为 $[0, T]$ 上 $k \times n$ 矩阵函数集合, 其元素分段连续; K 为 \mathbf{K} 的元素, 称为控制增益。问题就是选择一个 K_0 , 使式(6.8.2)达到极小。这时, 可写为 $J(u) = J(K)$ 。

把式(6.8.3)代入式(6.8.1)和(6.8.2), 系统模型和性能指标可表示为

$$dx(t) = \tilde{A} x(t)dt + dw(t) \tag{6.8.4}$$

$$J(u) = E \int_0^T x^T(t)M(t)x(t) dt + x^T(T)Q_0x(T) \tag{6.8.5}$$

式中:

$$\tilde{A} = A + BK \tag{6.8.6}$$

$$M = Q_1 + K^TQ_2K \tag{6.8.7}$$

6.8.2 贝尔曼方程

1. 二次型积分均值

为把线性确定性调节器问题(LQ)转变为随机系统问题(LQG), 首先计算

$$V(t, x) = E_{t, x} \int_t^T x^T(s) M(s) x(s) ds + x^T(T) Q_0 x(T) \quad (6.8.8)$$

式中: $E_{t, x}$ 表示由初值 $x(t) = x_t = x$ 开始取均值, 并设 x_t 为确定性值。然后再考虑 x_t 为随机量情况。

式(6.8.4)的解为

$$x(s) = \Phi(s, t) x_t + \int_t^s \Phi(s, u) dw(u) \quad (6.8.9)$$

式中: Φ 为对应 \tilde{A} 的转移矩阵。考虑到 $E(\dot{w} dw) = 0$, 我们有二次型均值为

$$E x^T(s) M(s) x(s) = x^T(s) \Phi^T(s, t) M(s) \Phi(s, t) x(s) + \text{tr} M(s) \int_t^s \Phi(s, u) R_1(u) \Phi^T(s, u) du \quad (6.8.10)$$

因此有

$$\begin{aligned} V(t, x) = & x^T(t) \int_t^T \Phi^T(s, t) M(s) \Phi(s, t) ds + \Phi^T(T, t) Q_0 \Phi(T, t) x(t) \\ & + \int_t^T \text{tr} M(s) \int_t^s \Phi(s, u) R_1(u) \Phi^T(s, u) du ds \\ & + \text{tr} Q_0 \int_t^T \Phi(T, s) R_1(s) \Phi^T(T, s) ds \end{aligned} \quad (6.8.11)$$

用 $x^T(t) N(t) x(t)$ 表示上式第一项, 则 N 和 \dot{N} 为

$$N(t) = \int_t^T \Phi^T(s, t) M(s) \Phi(s, t) ds + \Phi^T(T, t) Q_0 \Phi(T, t) \quad (6.8.12)$$

$$\begin{aligned} \dot{N}(t) = & -\dot{M} + \int_t^T \frac{d}{dt} \Phi^T M ds + \frac{d}{dt} \Phi^T(T, t) Q_0 \Phi(T, t) + \dots \\ = & -\dot{M} - \tilde{A}^T \int_t^T \Phi^T M ds + \Phi^T(T, t) Q_0 \Phi(T, t) + \dots \\ = & -\dot{M} - \tilde{A}^T N - N \tilde{A} \end{aligned} \quad (6.8.13)$$

因此, $N(t)$ 满足微分方程

$$\dot{N} + \tilde{A}^T N + N \tilde{A} + \dot{M} = 0 \quad (6.8.14)$$

$$N(T) = Q_0 \quad (6.8.15)$$

式(6.8.11)右边第二项在交换积分次序后, 得到

$$\begin{aligned} \text{第二项} = & \int_t^T \int_u^T \text{tr} M(s) \Phi(s, u) R_1(u) \Phi^T(s, u) ds du \\ = & \int_t^T \int_u^T \text{tr} \Phi^T(s, u) M(s) \Phi(s, u) R_1(u) ds du \end{aligned} \quad (6.8.16)$$

类似地, 式(6.8.11)右边第三项为

$$\text{第三项} = \int_t^T \text{tr} \left[\begin{matrix} T \\ T, u \end{matrix} \right] Q_0 \left(\begin{matrix} T \\ T, u \end{matrix} \right) R_1(u) du \quad (6.8.17)$$

用 $q(t)$ 表示第二项(6.8.16)和第三项(6.8.17)之和:

$$q(t) = \int_t^T \text{tr} N(u) R_1(u) du \quad (6.8.18)$$

把上式和式(6.8.12)代入式(6.8.11), 得到

$$V(t, x) = x^T N(t) x + \int_t^T \text{tr} N(s) R_1(s) ds \quad (6.8.19)$$

式中: N 为式(6.8.14)和(6.8.15)的唯一解。

考虑状态初值 x 为一随机过程, $x(t) \sim N(m, R)$, 那么由式(6.8.14)容易得到

$$\begin{aligned} E\{V(t, x)\} &= E \int_t^T x^T(s) M(s) x(s) ds + x^T(T) Q_0 x(T) \\ &= m^T N(t) m + \text{tr} N(t) R + \int_t^T \text{tr} N(s) R_1(s) ds \end{aligned} \quad (6.8.20)$$

2. 伊藤微分规则算子运算

定义微分算子

$$\mathbf{L} f(t, x) = f_t + f_x^T \tilde{A} x + \frac{1}{2} \text{tr} f_{xx} R_1 \quad (6.8.21)$$

式中: $f_t = \frac{f}{t}$; $f_x = \frac{f}{x}$; $f_{xx} = \frac{f^2}{x^2}$ 。用算子式(6.8.21)对 $V(t, x)$ (6.8.19)进行运算, 考虑到

$$V_t = x^T N x - \text{tr} N R_1 \quad V_x = 2N x \quad V_{xx} = 2N \quad (6.8.22)$$

得到

$$\begin{aligned} \mathbf{L} V(t, x) &= x^T (-\tilde{A}^T N - N \tilde{A} - M) x - \text{tr} N R_1 + 2x^T N \tilde{A} x + \text{tr} N R_1 \\ &= -x^T M x \end{aligned} \quad (6.8.23)$$

把式(6.8.23)代入式(6.8.8), 可得到式(6.8.19)的等价形式

$$V(t, x) = -E_{t, x} \int_t^T \mathbf{L} V(s, x(s)) ds - x^T(T) Q_0 x(T) \quad (6.8.24)$$

我们定义的算子式(6.8.21)就是随机微积分中所谓伊藤微分规则在二次型函数情况下的应用。

3. 控制增益 K 的优化条件

定理 6.8.1 设 $K_0 \in \mathbf{K}$, 对应的 $x_0(t)$ 是式(6.8.4)的解, 对应的 M_0, V_0 等满足式(6.8.8)

$$V_0(t, x) = E_{t, x} \int_t^T x_0^T(s) M_0(s) x_0(s) ds + x^T(T) Q_0 x(T) \quad (6.8.25)$$

如果有特性

$$\mathbf{L}_K V_0(t, x) + x^T M_K x - \mathbf{L}_0 V_0(t, x) + x^T M_0 x = 0 \quad (6.8.26)$$

则 K_0 是 \mathbf{K} 中最优的。

证明 只要证明 V_0 小于或等于任意 V 即可。由式(6.8.24), 有

$$V_0(t, x) = -E_{t, x} \int_t^T \mathbf{L}_K V_0(s, x(s)) ds - x^T(T) Q_0 x(T)$$

$$\begin{aligned}
& - E_{t,x} \int_t^T \mathbf{L}_0 \mathbf{V}_0 + \mathbf{x}^T(s) \mathbf{M}_0 \mathbf{x}(s) - \mathbf{x}^T(s) \mathbf{M}_K \mathbf{x}(s) ds - \mathbf{x}^T(T) \mathbf{Q}_0 \mathbf{x}(T) \\
& = - E_{t,x} \int_t^T - \mathbf{x}^T(s) \mathbf{M}_K \mathbf{x}(s) ds - \mathbf{x}^T(T) \mathbf{Q}_0 \mathbf{x}(T) \\
& = V(t, x)
\end{aligned} \tag{6.8.27}$$

上式中由第一式到第二式利用了给定条件式(6.8.26);第三、四两式即为式(6.8.8)。

把式(6.8.27)推广到一般情况,令 $(t, x(t)) = (0, x(0))$,并且状态初值 $x(0)$ 为高斯 $N(m_0, R_0)$ 过程向量,参照式(6.8.25)和(6.8.20),得到

$$E\{V_0(0, x(0))\} = J(K_0) - J(K) \tag{6.8.28}$$

由上式可知,控制增益 K_0 使性能指标函数极小,因此, K_0 是 \mathbf{K} 中最优的。

4. 随机贝尔曼方程

对 V 进行微分算子式(6.8.21)计算,得到

$$\begin{aligned}
\mathbf{L}_K V &= \frac{1}{2} \text{tr} V_{xx} R_1 + V_x^T \tilde{\mathbf{A}} x + V_t \\
&= \frac{1}{2} \text{tr} V_{xx} R_1 + V_x^T (\mathbf{A}x + \mathbf{B}u) + V_t
\end{aligned} \tag{6.8.29}$$

把上式代入式(6.8.26),得到

$$\min_K V_t + \frac{1}{2} \text{tr} V_{xx} R_1 + V_x (\mathbf{A}x + \mathbf{B}u) + x^T \mathbf{M}_K x = 0 \tag{6.8.30}$$

注意到

$$\begin{aligned}
u &= \mathbf{K}x & \tilde{\mathbf{A}} &= \mathbf{A} + \mathbf{B}\mathbf{K} & \mathbf{M} &= \mathbf{Q}_1 + \mathbf{K}^T \mathbf{Q}_2 \mathbf{K} \\
f(x, u) &= \mathbf{A}x + \mathbf{B}u & h(x, u) &= x^T \mathbf{Q}_1 x + u^T \mathbf{Q}_2 u
\end{aligned} \tag{6.8.31}$$

则式(6.8.30)变为

$$V_t + \frac{1}{2} \text{tr} V_{xx} R_1 + \min_u (V_x f + h) = 0 \tag{6.8.32}$$

这就是随机贝尔曼方程,注意它是在 V 是二次型函数和 u 是 x 的线性函数条件下建立的选择 u 的优化方程。

与确定性贝尔曼方程式(6.7.8)进行比较,随机贝尔曼方程附加了一项 $\frac{1}{2} \text{tr} V_{xx} R_1$ 。

5. 随机贝尔曼方程的解

把随机贝尔曼方程式(6.8.30)写为

$$V_t + \frac{1}{2} \text{tr} V_{xx} R_1 + x^T \mathbf{Q}_1 x + V_x^T \mathbf{A}x + \min_u u^T \mathbf{Q}_2 u + V_x^T \mathbf{B}u = 0 \tag{6.8.33}$$

先计算上式左边最后一项:

$$\begin{aligned}
\min_u u^T \mathbf{Q}_2 u + V_x^T \mathbf{B}u &= \min_u \left[u + \frac{1}{2} \mathbf{Q}_2^{-1} \mathbf{B}^T V_x \right]^T \mathbf{Q}_2 \left[u + \frac{1}{2} \mathbf{Q}_2^{-1} \mathbf{B}^T V_x \right] - \frac{1}{4} V_x^T \mathbf{B} \mathbf{Q}_2^{-1} \mathbf{B}^T V_x \\
&= - \frac{1}{4} V_x^T \mathbf{B} \mathbf{Q}_2^{-1} \mathbf{B}^T V_x
\end{aligned} \tag{6.8.34}$$

式中已取

$$u = - \frac{1}{2} \mathbf{Q}_2^{-1} \mathbf{B}^T V_x \tag{6.8.35}$$

u 满足上式时 V_t 达到极小, 这时式(6.8.33)为

$$V_t + \frac{1}{2} \text{tr} V_{xx} R_1 + x^T Q_1 x + V_x^T A x - \frac{1}{4} V_x^T B Q_2^{-1} B^T V_x = 0 \quad (6.8.36)$$

其初值为 $V_t(T, x(T)) = x^T(T) Q_0 x(T)$ 。类似于确定性贝尔曼方程式(6.7.21), 上式的解为

$$V_t(t, x) = x^T S(t) x + s(t) \quad (6.8.37)$$

$S(t)$ 和 $s(t)$ 为待定函数, 由上式得

$$V_t(t, x) = x^T S(t) x + s(t) \quad (6.8.38)$$

$$V_x(t, x) = 2S(t)x(t) \quad (6.8.39)$$

$$V_{xx}(t, x) = 2S(t) \quad (6.8.40)$$

把以上三式代入式(6.8.36), 经整理得到

$$x^T s + S A + A^T S + Q_1 - S^T B Q_2^{-1} B^T S x + S + \text{tr} S R_1 = 0 \quad (6.8.41)$$

令

$$- S = sA + A^T S + Q_1 - S B Q_2^{-1} B^T S \quad (6.8.42)$$

$$S(T) = Q_0$$

$$s = - \text{tr} S R_1 \quad (6.8.43)$$

$$s(T) = 0$$

则式(6.8.37)成立。把式(6.8.39)代入式(6.8.35), 得到最优控制

$$u(t) = - Q_2^{-1} B^T S(t) x(t) \quad (6.8.44)$$

性能指标函数的极小值由式(6.8.28)给出:

$$\begin{aligned} \min J(u[0, T]) &= E \{ V(0, x(0)) \} \\ &= E \{ x(0)^T S(0) x(0) + s(0) \} \\ &= m_0^T S(0) m_0 + \text{tr} S(0) R_0 + \int_0^T \text{tr} S(t) R_1 dt \end{aligned} \quad (6.8.45)$$

把上述推导结果总结为定理 6.8.2。

定理 6.8.2 考虑由式(6.8.1)和(6.8.2)描述的连续时间 LQG 问题, 设 $S(t)$ 是里卡蒂方程式(6.8.42)的有界对称解, 则控制式(6.8.44)是最优解, 性能指标极小值由式(6.8.45)表示。

把随机 LQG 问题与确定性 LQ 问题对比, 最优控制和里卡蒂方程是相同的, 噪声项 $dw(t)$ 不影响它们。“随机性”只对性能指标函数发生作用, 式(6.8.45)的第一项是确定性 LQ 问题的作用, 第二项是随机状态初值的作用, 而第三项则是噪声项 $dw(t)$ 的贡献。

6.8.3 随机线性调节器的开环控制

在上一小节, 我们讨论了随机调节器的闭环控制, 这一小节, 我们考虑开环控制。我们已经证明, 对确定性控制来说, 开环控制与闭环控制是等价的。但对随机控制来说, 两者是不同的, 并且开环控制一般次于闭环控制, 在例 6.1.1 中已说明了这一点。现在让我们进一步讨论这一问题。

考虑系统

$$dx(t) = Ax(t)dt + Bu(t)dt + dw(t) \quad (6.8.46)$$

式中: 状态初值 $x(0)$ 是高斯 $N(m_0, R_0)$ 的; 噪声 $dw(t)$ 是维纳 $N(0, R_1 dt)$ 过程。 $x(0)$ 与 $dw(t)$ 互相独立。为研究开环控制, 把系统状态 $x(t)$ 分为确定性部分 $x^*(t)$ 和随机部分 $x(t)$ 两部分, 则有

$$x(t) = x^*(t) + x(t) \quad (6.8.47)$$

$$dx^*(t) = Ax^*(t)dt + Bu(t)dt \quad x^*(0) = m_0 \quad (6.8.48)$$

$$dx(t) = Ax(t)dt + dw(t) \quad x(0) = x_0 - m_0 \quad (6.8.49)$$

用确定性控制实现的系统性能指标函数分为两部分:

$$J_0(u[0, T]) = J_1(u[0, T]) + J_2(K) \quad (6.8.50)$$

$$J_1(u[0, T]) = \int_0^T x^{*T}(t)Q_1x^*(t) + u^T(t)Q_2u(t) dt + x^{*T}(T)Q_0x^*(T) \quad (6.8.51)$$

$$J_2(K) = E \int_0^T x^T(t)Q_1x(t)dt + x^T(T)Q_0x(T) \quad (6.8.52)$$

首先讨论式(6.8.48)和(6.8.51)组成的确定性控制问题, 由 6.7 节得到最优控制和性能指标极小值为

$$u(t) = -Q_2^{-1}BSx^*(t) \quad (6.8.53)$$

$$\min J_1(u[0, T]) = m_0^T S(0)m_0 \quad (6.8.54)$$

式中:

$$S + A^T S + SA + Q_1 - SBQ_2^{-1}B^T S = 0 \quad (6.8.55)$$

$$S(T) = Q_0 \quad (6.8.56)$$

再讨论式(6.8.49)和(6.8.52)组成的随机问题, 由式(6.8.4)和(6.8.5)得到

$$\begin{aligned} \min J_2(K) &= E x(0)^T N(0)E\{x(0)\} + \text{tr}N(0)R_0 + \int_0^T \text{tr}N(s)R_1 ds \\ &= \text{tr}N(0)R_0 + \int_0^T \text{tr}N(s)R_1 ds \end{aligned} \quad (6.8.57)$$

式中:

$$E\{x(0)\} = E\{x(0) - m_0\} = 0 \quad (6.8.58)$$

$$N + A^T N + NA + Q_1 = 0 \quad (6.8.59)$$

$$N(T) = Q_0 \quad (6.8.60)$$

把式(6.8.54)和(6.8.57)代入式(6.8.50), 得到开环控制(确定性闭环控制)随机控制系统的性能指标极小值

$$\min J_0(u[0, T]) = m_0^T S(0)m_0 + \text{tr}N(0)R_0 + \int_0^T \text{tr}N(t)R_1 dt \quad (6.8.61)$$

重写随机系统闭环控制最优控制和性能指标函数等有关公式

$$u(t) = -Q_2^{-1}BSx(t) \quad (6.8.62)$$

$$\min J_f(u[0, T]) = m_0^T S(0)m_0 + \text{tr}S(0)R_0 + \int_0^T \text{tr}S(t)R_1 dt \quad (6.8.63)$$

式中:

$$S + A^T S + SA + Q_1 - SBQ_2^{-1}B^T S = 0 \tag{6.8.64}$$

$$S(T) = Q_0 \tag{6.8.65}$$

比较式(6.8.61)中 $\min J_0(u[0, T])$ 和式(6.8.63)中 $\min J_f(u[0, T])$ 可知, 两个公式基本相同, 只有第二项 $N(0)$ 和 $S(0)$, 第三项 $N(t)$ 和 $S(t)$ 的差别。为比较两式大小, 令 $M(t) = N(t) - S(t)$ 和 $Y = SBQ_2^{-1}B^T S$, 用式(6.8.59)减去式(6.8.64), 得到

$$M + A^T M + MA + Y = 0 \tag{6.8.66}$$

$$M(T) = 0 \tag{6.8.67}$$

把以上两式对照式(6.8.14), (6.8.15)和(6.8.12), 可以得到

$$M(t) = \int_t^T (s, t) Y(s) (s, t) ds \tag{6.8.68}$$

由上式可见, $Y(s)$ 是非负定的, 因而 $M(t)$ 也是非负定的, 因此

$$\text{tr} N(0) R_0 \geq \text{tr} S(0) R_0 \tag{6.8.69}$$

$$\int_0^T \text{tr} N(t) R_1 dt \geq \int_0^T \text{tr} S(t) R_1 dt \tag{6.8.70}$$

最后得到

$$\min J_0(u[0, T]) \geq \min J_f(u[0, T]) \tag{6.8.71}$$

该不等式是严格的, 除非在退化情况下。结论是随机闭环控制优于随机开环控制。

例 6.8.1 给定系统方程和性能指标函数分别为

$$dx = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x dt + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u dt + dw(t) \tag{6.8.72}$$

$$J(u) = \int_{t_0}^{t_1} [x_1^2 + qu^2] dt \tag{6.8.73}$$

状态初值 $x(t_0)$ 为高斯 $N(m_0, R_0)$ 向量; $dw(t)$ 为增量维纳 $N(0, R_1 dt)$ 过程; $x(t_0)$ 与 $dw(t)$ 互相独立; 有关参数为

$$m_0 = \begin{bmatrix} m_{10} \\ m_{20} \end{bmatrix} \quad R_0 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \quad R_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

试求最优控制。当 $t_0 \rightarrow \infty$ 时, 再求极限情况。

解 此例中, 除 m_0, R_0 和 R_1 为式(6.8.73)所示外, 其余参数为

$$\begin{aligned} A &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} & B &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} & Q_0 &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\ Q_1 &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} & Q_2 &= q \end{aligned} \tag{6.8.74}$$

在例 6.7.1 中已求出本题的确定性情况, S 的里卡蒂方程和控制矩阵不受随机因素的影响, 因此, 本题中 S_{11}, S_{12} 和 S_{22} 与例 6.7.1 中的相同, 最优控制亦为

$$u(t) = - \frac{1}{q} S_{12} X_1(t) - \frac{1}{q} S_{22} X_2(t) \tag{6.8.75}$$

形式与例 6.7.1 相同, x_1 和 x_2 为量测量, 由仪器测量得到。

随机因素会使系统性能变坏, 即使性能指标增加, 性能指标由式(6.8.45)给出:

$$\min J = m_0^T S(t_0) m_0 + \text{tr} S(t_0) R_0 + \int_{t_0}^{t_1} \text{tr} S(t) R_1 dt \tag{6.8.76}$$

当 $t_0 = 0$ 时, s 和 $u(t)$ 由例 6.7.1 给出。性能指标式(6.8.76)中各项分别为

$$m_0^T S(t_0) m_0 = \frac{1}{2} \overline{m_{10}^2} + \frac{1}{2} \overline{m_{10} m_{20}} + \frac{1}{2q} \overline{m_{20}^2} \tag{6.8.77}$$

$$\text{tr} S(t_0) R_0 = s_{11} + 2s_{22} = \frac{1}{2} \overline{1} + \frac{1}{2} \frac{2}{q} \overline{1} \tag{6.8.78}$$

$$\int_{t_0}^{t_1} \text{tr} S(t) R_1 dt = \frac{1}{2} \overline{1} + \frac{1}{2q} \overline{1} (t_1 - t_0) \tag{6.8.79}$$

把以上各式代入式(6.8.76), 得到

$$\begin{aligned} \min J = & \frac{1}{2} \overline{m_{10}^2} + \frac{1}{2} \overline{m_{10} m_{20}} + \frac{1}{2q} \overline{m_{20}^2} \\ & + \frac{1}{2} \overline{1} + \frac{1}{2} \frac{2}{q} \overline{1} + \frac{1}{2} \overline{1} + \frac{1}{2q} \overline{1} (t_1 - t_0) \end{aligned}$$

6.9 连续时间 LQG 不完全状态信息情况

求解连续时间状态模型 LQG 不完全状态信息情况最优控制问题的思路, 是把它转变为完全状态信息情况, 再用上节结论给出结果。上节给出最优控制是系统状态的线性反馈, 本节的结论类似, 最优控制是状态滤波的线性反馈。本节我们先给出问题描述, 然后引入中间状态向量 z , 把系统模型转变为相当“完全状态信息情况”的等效模型, 进而得到最优控制结果, 最后根据这个结果给出分离定理。

6.9.1 问题描述

1. 系统模型

连续时间随机状态模型 LQG 不完全状态信息情况和性能指标函数分别给定为

$$dx(t) = Ax(t)dt + Bu(t)dt + dw(t) \tag{6.9.1}$$

$$dy(t) = Hx(t)dt + dv(t) \tag{6.9.2}$$

$$J(u) = E \left[x^T(T) Q_0 x(T) + \int_0^T x^T(t) Q_1 x(t) + u^T(t) Q_2 u(t) dt \right] \Big| Y_t \tag{6.9.3}$$

式中: $x(t)$ 为 $n \times 1$ 随机状态向量, 状态初值 $x(t_0)$ 为 $N(m_0, R_0)$ 向量, $t_0 = 0$; $u(t)$ 为 $m \times 1$ 确定性控制向量; $y(t)$ 为 $n \times 1$ 随机量测向量; $dw(t)$ 为 $n \times 1$ 增量维纳过程 $N(0, R_1 dt)$ 模型噪声向量; $dv(t)$ 为 $n \times 1$ 增量维纳过程 $N(0, R_2 dt)$ 量测噪声向量; A, B, H, Q_0, Q_1 和 Q_2 分别为相应矩阵, R_1, Q_0 和 Q_1 是对称非负定阵, R_2 和 Q_2 为正定阵。

$x(t_0)$ 与 $dw(t)$ 和 $dv(t)$ 互相独立, 由于它们都是高斯的, 因此, 系统状态 $x(t)$ 和量测向量 $y(t)$ 也都是高斯的。

2. 容许控制

为反映量测方程式(6.9.2)的作用,由性能指标函数式(6.9.3)也可看出,容许控制 $u(t)$ 为

$$u(t) = u(Y_t) \tag{6.9.4}$$

量测 Y_t 的作用反映为状态滤波 $\hat{x}(t|Q)$, 因此最优控制函数 $u(t)$ 为最优滤波 $\hat{x}(t|Q)$ 的线性函数,即

$$u(t) = K(t) \hat{x}(t|Q) \tag{6.9.5}$$

这一结论是正确的,离散时间系统也是这样表示的,但对连续时间系统在做法上要复杂一些。为此,引入一个中间状态向量 $z(t)$, 并设

$$u(t) = K(t)z(t) \tag{6.9.6}$$

$$dz(t) = \quad z(t)dt + \quad dy(t) \qquad z(t_0) = \tag{6.9.7}$$

式中 $z(t)$ 实际上与 $\hat{x}(t|Q)$ 等价,在后面推导中会看到这一点。 (K, \quad, \quad) 的意义在后面给出,由式(6.9.6)和(6.9.7)两式视出, $u(t)$ 也可视为 (K, \quad, \quad) 的函数。

3. 状态滤波模型

现把 $x(t)$ 和 $y(t)$ 分为随机和确定性两部分:

$$x(t) = x(t) + x^*(t) \tag{6.9.8}$$

$$y(t) = y(t) + y^*(t) \tag{6.9.9}$$

式中: $x(t)$ 和 $y(t)$ 表示随机量部分; $x^*(t)$ 和 $y^*(t)$ 表示确定性量部分。由以上两式和式(6.9.1)和(6.9.2)可得:

$$dx(t) = Ax(t)dt + dw(t) \tag{6.9.10}$$

$$dy(t) = Hx(t)dt + dv(t) \tag{6.9.11}$$

$$dx^*(t) = Ax^*(t)dt + Bu(t)dt \tag{6.9.12}$$

$$dy^*(t) = Hx^*(t)dt \tag{6.9.13}$$

为确定状态滤波模型,看一看式(6.9.11)的新息过程,定义新息过程为

$$d(t) = dy(t) - H\hat{x}(t|Q)dt$$

注意到式(6.9.8)的滤波为

$$\hat{x}(t|Q) = \hat{x}(t|Q) + x^*(t) \tag{6.9.14}$$

因而有

$$\begin{aligned} d(t) &= dy(t) - H\hat{x}(t|Q)dt + [dy^*(t) - Hx^*(t)dt] \\ &= dy(t) - H\hat{x}(t|Q)dt \end{aligned} \tag{6.9.15}$$

这就是式(6.9.2)的新息过程,上式表示式(6.9.11)的新息过程等于式(6.9.2)的新息过程。

由定理 5.8.2, 式(6.9.10)和(6.9.11)的最优滤波(卡尔曼滤波)方程为

$$\begin{aligned} d\hat{x}(t|Q) &= A\hat{x}(t|Q)dt + P(t|Q)H^TR_2^{-1}d(t) \\ &= A\hat{x}(t|Q)dt + P(t|Q)H^TR_2^{-1}[dy(t) - H\hat{x}(t|Q)dt] \end{aligned} \tag{6.9.16}$$

式中: $P(t|Q)$ 为最优滤波误差方差阵。把上式与式(6.9.12)相加,得到

$$d\hat{x}(t|Q) = A\hat{x}(t|Q)dt + Bu(t)dt + P(t|Q)H^TR_2^{-1}d(t)$$

$$= A \hat{x}(t|t)dt + Bu(t)dt + P(t|t)H^T R_2^{-1}(dy(t) - H \hat{x}(t|t)dt) \quad (6.9.17)$$

这就是带有控制项 $Bu(t)$ 的标准最优滤波方程, 其中 $d(t)$ 为增量维纳 $N(0, R_1 dt)$ 过程。把式(6.9.17)与式(6.9.1)进行比较可知, $\hat{x}(t|t)$ 的方程取代了 $x(t)$ 的状态方程, $\hat{x}(t|t)$ 的所有项都是“已知的”。

但是, 性能指标函数式(6.9.3)仍借助 $x(t)$ 给出, 下一步我们将借助 $\hat{x}(t|t)$ 重新阐述它。

4. 性能指标

我们已知, 状态向量 $x(t)$ 由最优滤波 $\hat{x}(t|t)$ 和滤波误差 $\tilde{x}(t|t)$ 两部分组成, $x(t) = \hat{x}(t|t) + \tilde{x}(t|t)$, 且 $\hat{x}(t|t) \perp \tilde{x}(t|t)$, 于是有

$$\begin{aligned} E\{x^T(t)Qx(t)\} &= E\{\hat{x}^T(t|t) + \tilde{x}^T(t|t)\}^T Q \{\hat{x}(t|t) + \tilde{x}(t|t)\} \\ &= E\{\hat{x}^T(t|t)Q\hat{x}(t|t)\} + E\{\tilde{x}^T(t|t)Q\tilde{x}(t|t)\} \end{aligned} \quad (6.9.18)$$

我们还知道, 最优滤波误差 $\tilde{x}(t|t)$ 为高斯 $N(0, P(t|t))$ 过程, 且与 $u(t)$ 无关, 因此有

$$E\{\tilde{x}^T(t|t)Q\tilde{x}(t|t)\} = \text{tr}QP(t|t) \quad (6.9.19)$$

把式(6.9.19)代入式(6.9.18), 得到

$$E\{x^T(t)Qx(t)\} = E\{\hat{x}^T(t|t)Q\hat{x}(t|t)\} + \text{tr}QP(t|t) \quad (6.9.20)$$

考虑到上式结果, 性能指标函数式(6.9.3)可写为

$$\begin{aligned} J(u) &= E\{\hat{x}^T(T|T)Q_0\hat{x}(T|T)\} + \int_0^T \hat{x}^T(t|t)Q_1\hat{x}(t|t) + u^T(t)Q_2u(t) dt \\ &\quad + \int_0^T \text{tr}Q_1P(t|t)dt + \text{tr}Q_0P(T|T) \end{aligned} \quad (6.9.21)$$

式中后两项不依赖于控制, 把上式与式(6.9.3)比较可以看出, 使原性能指标函数式(6.9.3)最小等价于使 $\hat{x}(t|t)$ 的同样函数为最小, 不过要用 $\hat{x}(t|t)$ 代替 $x(t)$ 。

6.9.2 等效模型及其解

对式(6.9.6)和(6.9.7), 我们选择

$$\begin{aligned} K(t) &= K_1(t) \\ &= A - P(t|t)H^T R_2^{-1}H + BK_1(t) \\ &= P(t|t)H^T R_2^{-1} \\ &= m_0 \end{aligned} \quad (6.9.22)$$

则由式(6.9.7)和(6.9.17)得到

$$dz(t) = [A - PH^T R_2^{-1}H + BK_1]z(t)dt + PH^T R_2^{-1}dy \quad z(0) = m_0 \quad (6.9.23)$$

$$d\hat{x}(t|t) = [A - PH^T R_2^{-1}H]\hat{x}(t|t)dt + BK_1 zdt + PH^T R_2^{-1}dy \quad \hat{x}(0|0) = m_0 \quad (6.9.24)$$

由式(6.9.23)减式(6.9.24), $[z - \hat{x}(t|t)]$ 满足

$$d[z - \hat{x}(t|t)] = [A - PH^T R_2^{-1}H][z - \hat{x}(t|t)]dt \quad z(0) - \hat{x}(0|0) = 0 \quad (6.9.25)$$

结果表明, 对所有 t , $z(t) = \hat{x}(t|t)$, 使式(6.9.6)变为 $u(t) = K_1(t)\hat{x}(t|t)$, 因此, 我们选择控制表示式(6.9.6)是允许的。

现把 $u(t)$ 选为更一般的形式, 符号也改变一点, 令

$$u(t) = K_1(t) \hat{x}(t|t) + K_2(t)z(t) = K(t) \begin{pmatrix} \hat{x}(t|t) \\ z(t) \end{pmatrix} \quad (6.9.26)$$

式中: $K(t) = [K_1(t) \quad K_2(t)]$ 。

在以前讨论基础上, 由式(5.9.7), (6.9.15)和(6.9.17)组成等效模型

$$\begin{pmatrix} d\hat{x}(t|t) \\ dz(t) \end{pmatrix} = \begin{pmatrix} A & 0 \\ H & z \end{pmatrix} \begin{pmatrix} \hat{x}(t|t) \\ z \end{pmatrix} dt + \begin{pmatrix} B \\ 0 \end{pmatrix} u(t) dt + \begin{pmatrix} PH^T R_2^{-1} \\ 0 \end{pmatrix} d(t) \quad (6.9.27)$$

即

$$d(t) = \tilde{A}(t)dt + \tilde{B}u(t)dt + \tilde{C}d(t) \quad (6.9.28)$$

式中:

$$\begin{pmatrix} \hat{x}(t|t) \\ z \end{pmatrix} \quad \tilde{A} = \begin{pmatrix} A & 0 \\ H & \end{pmatrix} \quad \tilde{B} = \begin{pmatrix} B \\ 0 \end{pmatrix} \quad \tilde{C} = \begin{pmatrix} PH^T R_2^{-1} \\ 0 \end{pmatrix} \quad (6.9.29)$$

并且控制具有形式 $u(t) = K(t) \begin{pmatrix} \hat{x}(t|t) \\ z \end{pmatrix}$ 。对应等效状态模型, 考虑到式(6.9.21), 性能指标函数给定为

$$J(u) = E \begin{pmatrix} \hat{x}(T|T) \\ z(T) \end{pmatrix}^T \tilde{Q}_0 \begin{pmatrix} \hat{x}(T|T) \\ z(T) \end{pmatrix} + \int_0^T \begin{pmatrix} \hat{x}(t|t) \\ z(t) \end{pmatrix}^T \tilde{Q}_1 \begin{pmatrix} \hat{x}(t|t) \\ z(t) \end{pmatrix} + u(t)Q_2 u(t) dt + \int_0^T \text{tr} Q_1 P(t|t) dt + \text{tr} Q_0 P(T|T) \quad (6.9.30)$$

式中:

$$\tilde{Q}_1 = \begin{pmatrix} Q_1 & 0 \\ 0 & 0 \end{pmatrix} \quad \tilde{Q}_0 = \begin{pmatrix} Q_0 & 0 \\ 0 & 0 \end{pmatrix} \quad (6.9.31)$$

由等效系统模型式(6.9.28)和性能指标函数式(6.9.30)组成了标准的 LQG 完全状态信息情况问题, 由定理 6.8.2 得到, 最优控制为

$$u(t) = -Q_2^{-1} \tilde{B}^T S(t) \quad (6.9.32)$$

式中: S 是里卡蒂方程

$$\dot{S} = -\tilde{A}^T S - S \tilde{A} - \tilde{Q}_1 + S \tilde{B} Q_2^{-1} \tilde{B}^T S \quad (6.9.33)$$

$$S(T) = \tilde{Q}_0 \quad (6.9.34)$$

的解。令

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \quad Y = B Q_2^{-1} B^T \quad (6.9.35)$$

因为 S 是对称阵, 所以 $S_{12} = S_{21}$ 。把上式代入式(6.9.33), 展开得到

$$(S_{11} + A^T S_{11} + S_{11} A + Q_1 - S_{11} Y S_{11}) + S_{12} H + H^T S_{12} = 0 \quad S_{11}(T) = Q_0 \quad (6.9.36)$$

$$S_{12} + S_{12} + A^T S_{12} + S_{11} Y S_{12} + H^T S_{22} = 0 \quad S_{12}(T) = 0 \quad (6.9.37)$$

$$S_{22} + S_{22} + S_{22}^T S_{22} + S_{12} Y S_{12} = 0 \quad S_{22}(T) = 0 \quad (6.9.38)$$

式(6.9.36)左边括号内是完全状态信息情况里卡蒂方程的内容。式(6.9.33)的解不是唯

一的,我们选取其中一种解 $s_1 = 0, s_{12} = s_2 = 0$, 则得到最优控制为

$$u(t) = - Q_2^{-1} B^T S_1 \hat{x}(t|t) \tag{6.9.39}$$

它是与完全状态信息情况形式相同的最优控制,只是用 $\hat{x}(t|t)$ 代替了 $x(t)$ 。

性能指标函数为

$$\begin{aligned} \min J(u) &= [m_0 \quad \quad] S(0) m_0 + \int_0^T \text{tr} S \tilde{C} R_2^{-1} \tilde{C}^T ds + \int_0^T \text{tr} Q_1 P(t|t) dt + \text{tr} Q_0 P(t|t) \\ &= m_0^T S_1(0) m_0 + \int_0^T \text{tr} S_1 P(t|t) H^T R_2^{-1} H P(t|t) dt + \int_0^T \text{tr} Q_1 P(t|t) dt \\ &\quad + \text{tr} Q_0 P(T|T) \end{aligned} \tag{6.9.40}$$

注意式(6.9.17)中不包含 \hat{x} 和 \hat{u} , 因此选取 $\hat{x} = \hat{u} = 0$, 而不影响性能指标函数。另外,我们看到引入的中间状态向量 $z(t)$ 可以复现 $\hat{x}(t|t)$ 的动力学, 因此, $z(t)$ 和 $\hat{x}(t|t)$ 的不同组合都可用于控制, 并得到相应的性能指标函数。

性能指标表示式(6.9.40)的物理意义不明显, 可以把它化为物理意义明显的表示式。为此, 构造函数 $P(t)S(t)$, 其中 $P(t) = P(t|t)$ 为最优滤波误差方差阵, 把 s_1 写为 S , 对 PS 进行微分, 并把 P 和 S 的里卡蒂方程代入, 得到

$$\begin{aligned} \frac{d(P S)}{dt} &= \frac{dP}{dt} S + P \frac{dS}{dt} \\ &= (AP + P A^T + R_1 - P H^T R_2^{-1} H P) S \\ &\quad + P (-SA - A^T S - Q_1 + S B Q_2^{-1} B^T S) \end{aligned} \tag{6.9.41}$$

$$\begin{aligned} d(P S) &= (AP S + P A^T S + R_1 S - P H^T R_2^{-1} H P S \\ &\quad - P S A - P A^T S - P Q_1 + P S B Q_2^{-1} B^T S) dt \end{aligned} \tag{6.9.42}$$

对上式取迹(tr), 消去有关项后进行积分, 得到

$$\begin{aligned} \int_0^T \text{tr} d(P S) &= \int_0^T \text{tr} R_1 S dt - \int_0^T \text{tr} P H^T R_2^{-1} H P S dt \\ &\quad - \int_0^T \text{tr} P Q_1 dt + \int_0^T \text{tr} P S B Q_2^{-1} B^T S dt \end{aligned} \tag{6.9.43}$$

计算上式左边, 注意到 $P(0) = R_0, S(T) = Q_0$, 得到

$$\begin{aligned} \int_0^T \text{tr} d(P S) &= \text{tr} P(T) S(T) - \text{tr} P(0) S(0) \\ &= \text{tr} P(T) Q_0 - \text{tr} R_0 S(0) \end{aligned} \tag{6.9.44}$$

式(6.9.43)右边最后一项还可表示为如下形式:

$$\begin{aligned} \int_0^T \text{tr} P S B Q_2^{-1} B^T S dt &= \int_0^T \text{tr} P S B Q_2^{-1} Q_2 Q_2^{-1} B^T S dt \\ &= \int_0^T \text{tr} P L^T Q_2 L dt \\ &= \int_0^T \text{tr} Q_2 L P L^T dt \end{aligned} \tag{6.9.45}$$

式中已令

$$L(t) = Q_2^{-1} B^T S(t) \tag{6.9.46}$$

把式(6.9.44)和(6.9.45)代入式(6.9.43), 进行移项整理, 得到

$$\begin{aligned} & \int_0^T \text{tr} S P H^T R_2^{-1} H P dt + \int_0^T \text{tr} Q_1 P dt + \text{tr} Q_0 P(T) \\ &= \text{tr} S(0) R_0 + \int_0^T \text{tr} S R_1 dt + \int_0^T \text{tr} Q_2 L P L^T dt \end{aligned} \quad (6.9.47)$$

把上式代入式(6.9.40), 得到性能指标最后表示式

$$\min J(u) = m_0^T S(0) m_0 + \text{tr} S(0) R_0 + \int_0^T \text{tr} S R_1 dt + \int_0^T \text{tr} Q_2 L P L^T dt \quad (6.9.48)$$

性能指标表示式(6.9.48) 物理意义明显, 第一项反映确定性控制情况, 第二项表示状态初值为随机量时的附加量, 第三项表示系统模型噪声项对性能指标所做贡献, 第四项为量测噪声项对性能指标所做贡献。

把上述结果总结为定理形式。

定理 6.9.1 对由式(6.9.1) ~ (6.9.3) 描述的连续时间随机状态模型 LQG 不完全状态信息情况, 使性能指标式(6.9.3) 极小的最优控制为

$$u(t) = - L(t) \hat{x}(t|t) \quad (6.9.49)$$

式中: 控制阵 $L(t)$ 为

$$L(t) = Q_2^{-1} B^T S \quad (6.9.50)$$

S 为里卡蒂方程

$$- \dot{S} = A^T S + S A + Q_1 - S B Q_2^{-1} B^T S \quad (6.9.51)$$

$$S(T) = Q_0 \quad (6.9.52)$$

的解。最优滤波由式(6.9.17) 和定理 5.8.2 给出。性能指标极小值为式(6.9.48)。

6.9.3 分离定理

定理 6.9.1 的结果称为分离定理或必然等价原理。最优控制的计算分为两部分, 一部分是计算最优滤波 $\hat{x}(t|t)$, 另一部分是计算作为 $\hat{x}(t|t)$ 函数的控制值 $u(t) = - L(t) \hat{x}(t|t)$ 。决定性的一点是这两部分计算在某种意义上是独立的, 最优滤波 $\hat{x}(t|t)$ 的计算与性能指标参数 Q_0 , Q_1 和 Q_2 无关, 这三个矩阵定义了控制问题; 控制矩阵 $L(t)$ 也不依赖于 R_0 , R_1 和 R_2 等噪声参数。两部分计算, 基本上分开, 独立进行, 故称为分离定理。如果用实际状态 $x(t)$ 代替最优滤波 $\hat{x}(t|t)$, 最优控制计算同样进行, 所以分离定理也称为必然等价原理。

与完全状态信息情况相同, 在不完全状态信息情况下 $L(t)$ 的性质有赖于如下事实: 一方面, 当性能指标函数改造为式(6.9.21) 时, 作为 $\hat{x}(t|t)$, $u(t)$ 的函数, 它的形式与 $x(t)$, $u(t)$ 的原函数的形式相同, 这是二次型函数的一个性质; 另一方面, 分离结果主要适用于系统动力学的线性结构, 它保证状态的条件分布是高斯的, 由 $\hat{x}(t|t)$ 和 $P(t|t)$ 所规定, 并能让我们把 $x(t)$ 分为“噪声”和“控制”两部分, $x(t) = x(t) + x^*(t)$ 。对于性能指标函数不为二次型的线性系统, 最优控制具有 $u(t) = u(t, \hat{x}(t|t))$ 的形式, 但对完全状态信息情况的最优控制并不是 $u(t, x(t))$, 而是某种其他的函数。因此, 这时分离定理成立, 而等价原理必然不成立。

例 6.9.1 考虑万向平台系统, 设系统状态 $x(t)$ 表示平台转角, 系统方程、量测方程和性能指标函数分别为

$$dx(t) = -\frac{1}{T}x(t)dt + \frac{1}{T}u(t)dt + dw(t) \quad (6.9.53)$$

$$dy(t) = x(t)dt + dv(t) \quad (6.9.54)$$

$$J = E \left[Q_0 x^2(t_1) + \int_{t_0}^{t_1} W_{xx} x^2(t) + W_{uu} u^2(t) dt \right] \quad (6.9.55)$$

式中: 状态初值 $x(t_0)$ 为高斯 $N(m_0, r_0)$ 的; 噪声 $dw(t)$ 和 $dv(t)$ 为增量维纳过程; 参数分别为 $N(0, r_1)$ 和 $N(0, r_2)$ 。 $x(t_0)$, $dw(t)$ 与 $dv(t)$ 两两互相独立。试求最优控制和定态($t_0 \rightarrow \infty$)情况。

解 对照定理 6.9.1, 本例中

$$\begin{aligned} A &= -\frac{1}{T} & B &= \frac{1}{T} & H &= 1 & m_0 &= m_0 & R_0 &= r_0 & R_1 &= r_1 \\ R_2 &= r_2 & Q_0 &= Q_0 & Q_1 &= W_{xx} & Q_2 &= W_{uu} \end{aligned} \quad (6.9.56)$$

把数据代入里卡蒂方程式(6.9.51)和(6.9.52)得到:

$$-\dot{S} = -\frac{2}{T}S + W_{xx} - \frac{1}{T^2 W_{uu}} S^2 \quad (6.9.57)$$

$$S(t_1) = Q_0 \quad (6.9.58)$$

由以上方程解出 S , 把它和式(6.9.56)代入式(6.9.50), 得到最优控制为

$$u(t) = -\frac{S}{T W_{uu}} \hat{x}(t|t) \quad (6.9.59)$$

最优滤波 $\hat{x}(t|t)$ 由式(6.9.17)给出:

$$\begin{aligned} d\hat{x}(t|t) &= Ax(t|t)dt + Bu(t)dt + P(t|t)H^T R_2^{-1}[dy(t) - H\hat{x}(t|t)dt] \\ &= -\frac{1}{T}\hat{x}(t|t)dt + \frac{1}{T}u(t)dt + \frac{P(t|t)}{r_2}[dy(t) - \hat{x}(t|t)dt] \end{aligned} \quad (6.9.60)$$

而 $P(t|t)$ 由里卡蒂方程

$$\dot{P} = -\frac{2}{T}P + r_1 - \frac{P^2}{r_2} \quad (6.9.61)$$

$$P(t_0) = r_0 \quad (6.9.62)$$

确定。

再考虑定态情况。 $t_0 \rightarrow \infty$ 时, $S = 0$, 里卡蒂方程式(6.9.57)退化为代数方程, 解为

$$S = \frac{W_{xx}}{1 + \frac{W_{xx}}{W_{uu}}} = 1 \quad (6.9.63)$$

把上式代入式(6.9.59), 给出最优控制为

$$u(t) = -\frac{1}{1 + \frac{W_{xx}}{W_{uu}}} \hat{x}(t|t) \quad (6.9.64)$$

定态时, $t_1 \rightarrow \infty$, $P = 0$, 里卡蒂方程式(6.9.61)退化为代数方程, 解为

$$P = r_2 \left[\frac{1}{T^2} + \frac{r_1}{r_2} - \frac{1}{T} \right] \quad (6.9.65)$$

把上式代入式(6.9.60), 给出最优滤波为

$$d\hat{x}(t|t) = -\frac{1}{T}\hat{x}(t|t)dt + \frac{1}{T}u(t)dt + \sqrt{\frac{1}{T^2} + \frac{r_1}{r_2} - \frac{1}{T}} dy(t) - \hat{x}(t|t)dt \quad (6.9.66)$$

用式(6.9.48) 计算性能指标函数, 由四项组成, 分别为

$$m_0^T S(t_0)m_0 = \sqrt{1 + \frac{W_{xx}}{W_{uu}}} - 1 \quad m_0^2 \quad (6.9.67)$$

$$\text{tr}S(t_0)R_0 = \sqrt{1 + \frac{W_{xx}}{W_{uu}}} - 1 \quad r_0 \quad (6.9.68)$$

$$\int_{t_0}^{t_1} \text{tr}SR_1 dt = \sqrt{1 + \frac{W_{xx}}{W_{uu}}} - 1 \quad r_1(t_1 - t_0) \quad (6.9.69)$$

$$\int_{t_0}^{t_1} \text{tr}Q_2LPL^T dt = \frac{r_2}{T^2} \sqrt{\frac{1}{T^2} - \frac{r_1}{r_2} - \frac{1}{T}} \sqrt{1 + \frac{W_{xx}}{W_{uu}}} - 1 \quad (t_1 - t_0)^2 \quad (6.9.70)$$

由以上四式给出性能指标函数为

$$\begin{aligned} \min J = & \sqrt{1 + \frac{W_{xx}}{W_{uu}}} - 1 \quad m_0^2 + r_0 + r_1 + \frac{r_2}{T^2} \sqrt{\frac{1}{T^2} - \frac{r_1}{r_2} - \frac{1}{T}} \sqrt{1 + \frac{W_{xx}}{W_{uu}}} - 1 \\ & \int_{t_0}^{t_1} \sqrt{1 + \frac{W_{xx}}{W_{uu}}} - 1 \quad (t_1 - t_0) \end{aligned} \quad (6.9.71)$$

由定态结果可以看出: 对最优控制律的计算只与系统参数和性能指标参数有关, 各种噪声项对它没有影响; 对最优滤波的计算也一样, 只与系统参数和各噪声项有关, 而与性能指标参数无关。两者计算“基本上”独立进行, 它们在计算最优滤波 $\hat{x}(t|t)$ 时通过 $u(t)$ 项建立了联系。上述分析也适合一般非定态结果。从这个例子也验证了分离定理。

6.10 连续时间 LQG 控制系统小结

本节通过证明一个变分恒等式引理来求解确定性、完全状态信息和不完全状态信息情况的最优控制问题, 从而把各种情况联系起来。

6.10.1 变分恒等式

引理 6.10.1 给定系统

$$dx(t) = Ax(t)dt + Bu(t)dt + dw(t) \quad (6.10.1)$$

设初始条件(终端条件)为

$$S(T) = Q_0 \quad (6.10.2)$$

的里卡蒂方程

$$\begin{aligned} -\frac{dS}{dt} &= A^T S + SA + Q_1 - SBQ_2^{-1}B^T S \\ &= A^T S + SA + Q_1 - L^T Q_2 L \end{aligned} \quad (6.10.3)$$

在区间 $[t_0, T]$ 上有非负定解, 一般初值 $t_0 = 0$ 。式中 $L(t)$ 表示为

$$L(t) = Q_2^{-1} B^T S(t) \quad (6.10.4)$$

如果 $x(t)$ 是随机微分方程式(6.10.1)的解, 则必存在变分恒等式

$$\begin{aligned} J(t_0) &= x^T(T) Q_0 x(T) + \int_0^T [x^T(t) Q_1 x(t) + u^T(t) Q_2 u(t)] dt \\ &= x^T(t_0) S(t_0) x(t_0) + \int_{t_0}^T [u + Lx]^T Q_2 [u + Lx] dt \\ &\quad + \int_{t_0}^T \text{tr} S R_1 dt + \int_{t_0}^T dw^T S x + \int_{t_0}^T x^T S dw \end{aligned} \quad (6.10.5)$$

式中符号意义和假设如前所述。

证明 构造恒等式

$$\begin{aligned} x^T(T) Q_0 x(T) &= x^T(T) S(T) x(T) \\ &= x^T(t_0) S(t_0) x(t_0) + \int_{t_0}^T d(x^T S x) \end{aligned} \quad (6.10.6)$$

下面需要计算 $d(x^T S x)$, 由于 x 是随机微分方程式(6.10.1)的解, 不能用一般方法求解, 而 $d(x^T S x)$ 也不遵守一般的微积分规则, 因此需要用伊藤微分规则计算。注意到增量 dx 的方差阵为 $R_1 dt$, 因此有

$$\begin{aligned} d(x^T S x) &= (x^T S x)_t dt + (x^T S x)_x dx + \frac{1}{2} \text{tr}(x^T S x)_{xx} R_1 dt \\ &= x^T \frac{dS}{dt} x dt + dx^T S x + x^T S dx + \text{tr} S R_1 dt \end{aligned} \quad (6.10.7)$$

利用给定里卡蒂方程式(6.10.3)和系统式(6.10.1), 计算上式各项, 得到

$$\begin{aligned} x^T \frac{dS}{dt} x dt &= x^T (-A^T S - S A - Q_1 + L^T Q_2 L) x dt \\ &= -x^T A^T S x - x^T S A x - x^T Q_1 x + L^T Q_2 L x dt \end{aligned} \quad (6.10.8)$$

$$\begin{aligned} dx^T S x &= (Ax dt + Bu dt + dw)^T S x \\ &= (x^T A^T S x + u^T B^T S x) dt + dw^T S x \end{aligned} \quad (6.10.9)$$

$$\begin{aligned} x^T S dx &= x^T S (Ax dt + Bu dt + dw) \\ &= (x^T S A x + x^T S B u) dt + x^T S dw \end{aligned} \quad (6.10.10)$$

把以上三式代入式(6.10.7), 得到

$$\begin{aligned} d(x^T S x) &= (u^T B^T S x - x^T S B u - x^T Q_1 x + x L^T Q_2 L x) dt \\ &\quad + \text{tr} S R_1 dt + dw^T S x + x^T S dw \\ &= (-x^T Q_1 x - u^T Q_2 u) dt + (u + Lx)^T Q_2 (u + Lx) dt \\ &\quad + \text{tr} S R_1 dt + dw^T S x + x^T S dw \end{aligned} \quad (6.10.11)$$

式中加入 $\pm u^T Q_2 u$ 以进行配方运算。把上式代入式(6.10.6), 对各项稍加整理就得到式(6.10.5), 于是引理得证。

6.10.2 各种控制情况

1. 确定性情况

给定确定性系统

$$dx(t) = Ax(t)dt + Bu(t)dt \quad dw(t) = 0 \quad (6.10.12)$$

由于 $dw(t) = 0$, 性能指标式(6.10.5)后三项无值, 于是有

$$\begin{aligned} \min_u J(t_0) &= \min_u x^T(t_0)S(t_0)x(t_0) + \int_{t_0}^T [u + Lx]^T Q_2 [u + Lx] dt \\ &= x^T(t_0)S(t_0)x(t_0) \end{aligned} \quad (6.10.13)$$

性能指标取极小值的条件是

$$u(t) = -Q_2^{-1}B^T S(t)x(t) \quad (6.10.14)$$

式中: $x(t)$ 为确定性状态向量, 状态初值为 $x(t_0)$; 函数 $S(t)$ 由里卡蒂方程式(6.10.2)和(6.10.3)确定。

2. 状态初值为随机量情况

给定系统

$$dx(t) = Ax(t)dt + Bu(t)dt \quad dw(t) = 0 \quad (6.10.15)$$

式中: 状态初值 $x(t_0)$ 为随机量, 并且 $x(t_0) \sim N(m_0, R_0)$ 。这时, 性能指标式(6.10.5)要取均值。只有选择

$$u(t) = -L(t)x(t) \quad (6.10.16)$$

时, 性能指标式(6.10.5)取极小值, 为

$$\begin{aligned} \min E\{J(t_0)\} &= E\{x^T(t_0)S(t_0)x(t_0)\} \\ &= m_0^T S m_0 + \text{tr} S(t_0) R_0 \end{aligned} \quad (6.10.17)$$

3. 完全状态信息情况

给定随机系统

$$dx(t) = Ax(t)dt + Bu(t)dt + dw(t) \quad (6.10.18)$$

式中符号意义和假设如前所述。我们设容许控制 $u(t) = u(X_t)$, 那么性能指标是在给定 X_t 条件下求得的。由于 $x(t)$ 既是被控量又是量测量, 因此条件均值为 $E\{x(t) | X_t\} = E\{x(t)\}$ 。给定 $dw(t)$ 的方差阵为 $R_1 dt$, 性能指标式(6.10.5)中第三项不为零。于是当选择

$$u(t) = -L(t)x(t) \quad (6.10.19)$$

时, 性能指标取极小值, 并为

$$\min E\{J(t_0)\} = m_0^T S(t_0) m_0 + \text{tr} S(t_0) R_0 + \int_{t_0}^T \text{tr} S(t) R_1 dt \quad (6.10.20)$$

4. 不完全状态信息情况

给定随机系统

$$dx(t) = Ax(t)dt + Bu(t)dt + dw(t) \quad (6.10.21)$$

$$dy(t) = Hx(t)dt + dv(t) \quad (6.10.22)$$

式中符号意义和假设如前所述。这种情况下, 性能指标为观测 Y_t 下的条件均值

$$E\{J(t_0)\} = E\{j | Y_t\} \quad (6.10.23)$$

式中“ \cdot ”表示式(6.10.5)右边内容。其中

$$\begin{aligned} E\{[u + Lx]^T Q_2 [u + Lx] | Y_t\} \\ = [u + L\hat{x}(t|Y_t)]^T Q_2 [u + L\hat{x}(t|Y_t)] + \text{tr} Q_2 L(t) P(t|Y_t) L^T(t) \end{aligned} \quad (6.10.24)$$

上式为二次型条件均值, 有关值计算如下:

$$\begin{aligned}
 E\{u(t) + L(t)x(t) \odot Y_t\} &= u(t) + L(t)x(t \odot t) \quad (6.10.25) \\
 R[u(t) + L(t)x(t) \odot Y_t] \\
 &= E[u(t) + L(t)x(t) - u(t) - L(t)\hat{x}(t \odot t)] [i \ddot{x}}^T \\
 &= E\{[L(t)(x(t) - \hat{x}(t \odot t))][i \ddot{x}}^T\} \\
 &= L(t)E[\tilde{x}(t \odot t) \tilde{x}^T(t \odot t)] L^T(t) \\
 &= L(t)P(t \odot t) L^T(t) \quad (6.10.26)
 \end{aligned}$$

式中: $P(t \odot t)$ 为滤波误差的方差阵。

$$\text{tr} Q_2 R[u(t) + L(t)x(t) \odot Y_t] = \text{tr} Q_2 L(t)P(t \odot t)L^T(t) \quad (6.10.27)$$

把式(6.10.24)代入式(6.10.23), 计算性能指标极小值:

$$\begin{aligned}
 \min_u E J(t_0) &= \min_u m_0^T S(t_0) m_0 + \text{tr} S(t_0) R_0 + \int_{t_0}^T \text{tr} S(t) R_1 dt \\
 &\quad + \int_{t_0}^T [u(t) + L(t)\hat{x}(t \odot t)]^T Q_2 [u(t) + L(t)\hat{x}(t \odot t)] dt \\
 &\quad + \int_{t_0}^T \text{tr} Q_2 L(t)P(t \odot t)L^T(t) dt \\
 &= m_0^T S(t_0) m_0 + \text{tr} S(t_0) R_0 + \int_{t_0}^T \text{tr} S(t) R_1 dt \\
 &\quad + \int_{t_0}^T \text{tr} Q_2 L(t)P(t \odot t)L^T(t) dt \quad (6.10.28)
 \end{aligned}$$

最优控制为

$$u(t) = -L(t)\hat{x}(t \odot t) = -Q_2^{-1} B^T S(t)\hat{x}(t \odot t) \quad (6.10.29)$$

由上述计算结果可知, 在各种情况下, 最优控制是相同的, 都由状态反馈实现, $L(t)$, $S(t)$ 和 $S(T)$ 都是相同的。区别只在于所用状态不同, 确定性情况和完全状态信息情况所用状态都为 $x(t)$, 而不完全状态信息情况所用状态为 $\hat{x}(t \odot t)$ 。性能指标在各种情况都不相同, 在确定性情况, 性能指标仅为 $m_0^T S(t_0) m_0$, 状态初值为随机量时, 则增加一项 $\text{tr} S(t_0) R_0$, 完全状态信息情况再增加一项 $\int_{t_0}^T \text{tr} S(t) R_1 dt$, 不完全状态信息情况还要增加一项 $\int_{t_0}^T \text{tr} Q_2 L(t)P(t \odot t)L^T(t) dt$ 。每增加一项噪声, 都要对性能指标做出贡献, 使性能指标变大, 使系统性能变坏。

6.11 对 偶 性

在讨论了随机最优控制和状态估计之后, 我们发现, 不管是离散时间系统, 还是连续时间系统, 最优控制和某些状态估计的主要公式在形式上是相同的, 只是变量和参数不同。但是只要把变量和参数稍加变换, 最优控制和状态估计就可互相转换, 最优控制可以等价于状态估计, 状态估计也可等价于最优控制, 这种在最优控制和状态估计之间存在的性质称为对偶性。可以这样说, 状态估计是最优控制的对偶系统, 最优控制也是状态估计

的对偶系统。

对偶性具有理论意义, 状态估计可以用最优控制来表示, 状态估计问题可通过研究最优控制来解决。反之亦然。对偶性具有实际意义, 可以编制一套程序, 只要改变某些变量和参数, 就可计算最优控制或状态估计问题。

我们分离散时间系统和连续时间系统两种情况讨论。讨论中, 先列出有关公式, 再找出对偶关系。

6.11.1 离散时间系统对偶性

离散时间控制系统的系统模型和性能指标如下(尽管确定性系统和随机系统有所不同, 但由于有关控制结果是相同的, 因此我们只列出确定性系统公式):

$$x(k+1) = Ax(k) + Bu(k) \tag{6.11.1}$$

$$y(k) = Cx(k) \tag{6.11.2}$$

$$J(u) = x^T(N)Q_0x(N) + \sum_{k=k_0}^{N-1} x^T(k)Q_1x(k) + u(k)Q_2u(k) \tag{6.11.3}$$

最优控制、控制增益、里卡蒂方程和边界条件分别为

$$u(k) = -L(k)x(k) \tag{6.11.4}$$

$$L(k) = [A^T S(k+1) + Q_2]^{-1} B^T S(k+1) \tag{6.11.5}$$

$$S(k) = A^T S(k+1) + Q_1 + L^T(k)[A^T S(k+1) + Q_2]L(k) \tag{6.11.6}$$

$$S(N) = Q_0 \tag{6.11.7}$$

以上各式符号意义和假设如前所述(下同)。

离散时间一步预测与上述系统具有对偶关系, 下面分别列出离散时间一步预测的系统模型及给定参数、性能指标、一步预测、一步预测增益、一步预测误差方差阵和边界条件公式:

$$x(k+1) = Ax(k) + Bu(k) + w(k) \tag{6.11.8}$$

$$y(k) = Cx(k) + v(k) \tag{6.11.9}$$

$$x(k_0) \sim N(m_0, R_0) \qquad w(k) \sim N(0, R_1) \qquad v(k) \sim N(0, R_2) \tag{6.11.10}$$

$$\hat{x}(k+1|k) = \hat{x}(k|k-1) + u(k) + K(k|k-1)[y(k) - C\hat{x}(k|k-1)] \tag{6.11.11}$$

$$K(k|k-1) = P(k|k-1)^T [P(k|k-1)^T + R_2]^{-1} \tag{6.11.12}$$

$$P(k+1|k) = P(k|k-1)^T + R_1 - K(k|k-1)[P(k|k-1)^T + R_2] \\ \times K^T(k|k-1) \tag{6.11.13}$$

$$\hat{x}(k_0|k_0-1) = m_0 \qquad P(k_0|k_0-1) = R_0 \tag{6.11.14}$$

对比上述公式, 列出最优控制的有关变量和参数, 再对应列出一部预测的对偶变量和参数, 即只要把一步预测的原有变量和参数进行给定变换, 就把一步预测公式化为最优控制公式, 结果如表 6.11.1 所示。

表 6.11.1 离散时间系统最优控制与一步最优预测的对偶关系

最 优 控 制	一步最优预测
	T
	T
	T
$L(k)$	$K^T(k k-1)$
$S(k)$	$P(k k-1)$
Q_1	R_1
Q_2	R_2
$S(N) = Q_0$	$P(k_0 k_0-1) = R_0$
N	$k_0 = 0$
$k_0 = 0$	N
k	$N - k$

6.11.2 连续时间系统对偶性

类似于离散时间系统, 确定性系统与随机系统最优控制的有关结果是相同的, 因此我们只列出确定性系统最优控制有关公式:

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{6.11.15}$$

$$y(t) = Hx(t) \tag{6.11.16}$$

$$J(u) = \int_0^T [x^T(t)Q_1x(t) + u^T(t)Q_2u(t)]dt + x^T(T)Q_0x(T) \tag{6.11.17}$$

$$u(t) = -L(t)x(t) \tag{6.11.18}$$

$$L(t) = Q_2^{-1}BS^T(t) \tag{6.11.19}$$

$$-\dot{S} = SA + A^TS + Q_1 - SBQ_2^{-1}B^TS \tag{6.11.20}$$

$$S(T) = Q_0 \tag{6.11.21}$$

连续时间系统最优滤波(卡尔曼滤波)与连续时间系统最优控制是一对对偶系统。有关最优滤波公式如下:

$$dx(t) = Ax(t)dt + Bu(t)dt + dw(t) \tag{6.11.22}$$

$$dy(t) = Hx(t)dt + dv(t) \tag{6.11.23}$$

$$x(t_0) \sim N(m_0, R_0) \qquad dw(t) \sim N(0, R_1dt) \qquad dv(t) \sim N(0, R_2dt) \tag{6.11.24}$$

$$\dot{\hat{x}}(t|t) = A\hat{x}(t|t)dt + Bu(t)dt + K(t|t)[dy(t) - H\hat{x}(t|t)dt] \tag{6.11.25}$$

$$K(t|t) = P(t|t)H^TR_2^{-1} \tag{6.11.26}$$

$$\dot{P} = AP + PA^T + R_1 - PH^TR_2^{-1}HP \tag{6.11.27}$$

$$P(k_0|k_0) = R_0 \tag{6.11.28}$$

连续时间系统最优控制与最优滤波的对偶关系如表 6.11.2 所示。

表 6.11.2 连续时间系统最优控制与最优滤波的对偶关系

最 优 控 制	最 优 滤 波
A	A^T
B	H^T
H	B^T
$L(t)$	$K(t t_0)$
$S(t)$	$P(t t_0)$
Q_1	R_1
Q_2	R_2
$S(T) = Q_0$	$P(t_0 t_0) = R_0$
T	$t_0 = 0$
$t_0 = 0$	T
t	$t - T$

6.12 里卡蒂方程的求解

我们在第 3 章讨论随机状态模型系统解时, 在第 5 章讨论随机状态估计时, 以及在本章讨论随机控制时, 不管是在离散时间系统, 或是在连续时间系统, 都遇到了求解里卡蒂方程问题。状态估计误差的方差阵 P 和最优控制增益中的 S 都由里卡蒂方程描述。

如果系统状态向量 $x(t)$ 为 $n \times 1$ 向量, 那么里卡蒂方程由 n^2 个差分方程(离散时间系统)或微分方程(连续时间系统)组成。由于里卡蒂方程具有重要性, 但求解较复杂, 很多学者对它进行研究, 寻找快速、精确的计算方法和提出了多种方法。在本节我们只介绍离散时间里卡蒂方程求解的差分方程方法和连续时间里卡蒂方程求解的矩阵微分方程方法。

6.12.1 差分方程法

我们以计算一步预测误差方差阵 $P(k+1|k)$ 为例说明这种算法, 重列一步预测增益 $K(k|k-1)$, $P(k+1|k)$ 和初值公式如下:

$$K(k|k-1) = P(k|k-1)^T [P(k|k-1)^T + R_2]^{-1} \tag{6.12.1}$$

$$P(k+1|k) = P(k|k-1)^T + R_1 - K(k|k-1) [P(k|k-1)^T + R_2] \\ \times K^T(k|k-1) \tag{6.12.2}$$

$$P(k_0|k_0-1) = m_0 \tag{6.12.3}$$

式中符号意义如前所述。计算是简单和直接的, 给定初值 m_0 后, 即可计算出 $K(k_0|k_0-1)$ 和 $P(k_0+1|k_0)$, 再由 $P(k_0+1|k_0)$ 计算出 $K(k_0+1|k_0)$ 和 $P(k_0+2|k_0+1)$, 依此类推, 可递推求出任意时刻 k 的 $K(k|k-1)$ 和 $P(k+1|k)$ 。如果系统稳定, 一直递推计算下去, 可得到稳态值, 或到相邻 $P(k+1|k)$ 之差足够小为止。

这种方法直接、简单, 适用于各种情况, 但收敛速度慢, 计算量大, 从而使截断和舍入误差较大。

6.12.2 矩阵微分方程法

我们以最优控制中 S 方程

$$-\dot{S}(t) = A^T S(t) + S(t)A + Q_1 - S(t)BQ_2^{-1}B^T S(t) \quad (6.12.4)$$

$$S(T) = Q_0 \quad (6.12.5)$$

为例讨论这种算法, 式中符号意义如前所述。

1. 方程的解

方程的解为

$$S(t) = Y(t)X^{-1}(t) \quad (6.12.6)$$

式中 $Y(t)$ 和 $X(t)$ 为线性矩阵微分方程组

$$\dot{X}(t) = AX(t) - BQ_2^{-1}B^T Y(t) \quad (6.12.7)$$

$$\dot{Y}(t) = -Q_1 X(t) - A^T Y(t) \quad (6.12.8)$$

$$Y(T) = Q_0 \quad X(T) = I \quad (6.12.9)$$

的解。

(1) 首先证明式(6.12.7) ~ (6.12.9) 存在时, 式(6.12.6) 是式(6.12.4) 和(6.12.5) 的解。由式(6.12.6) 得到

$$Y(t) = S(t)X(t) \quad (6.12.10)$$

$$\dot{Y}(t) = \dot{S}(t)X(t) + S(t)\dot{X}(t) \quad (6.12.11)$$

把式(6.12.7), (6.12.8) 和(6.12.10) 代入上式, 得到

$$-Q_1 X(t) - A^T S(t)X(t) = \dot{S}(t)X(t) + S(t)[AX(t) - BQ_2^{-1}B^T S(t)X(t)] \quad (6.12.12)$$

把上式稍加整理, 就得到式(6.12.4)。把式(6.12.9) 代入式(6.12.6), 就得到初值(终端) 条件式(6.12.5)。

(2) 问题的关键是 $X(t)$ 是否存在, 且能保证非奇异。下面就证明, 当式(6.12.4) 和(6.12.5) 的解 $S(t)$ 存在时(对最优控制问题, 能保证 $S(t)$ 的存在和唯一), $X(t)$ 存在, 且为非奇异解。

把式(6.12.10) 代入式(6.12.7), 得到

$$\dot{X}(t) = [A - BQ_2^{-1}B^T S(t)]X(t) \quad (6.12.13)$$

把上式看做是状态方程, 则有解

$$X(t) = \Phi(t, T)X(T) \quad (6.12.14)$$

式中: $\Phi(t, T)$ 为 $[A - BQ_2^{-1}B^T S(t)]$ 的状态转移阵。当 $S(t)$ 存在时, $\Phi(t, T)$ 必存在, 且为非奇异阵。对 $\Phi(t, T)$ 求导数, 有

$$\begin{aligned} \dot{\Phi}(t, T) &= [A - BQ_2^{-1}B^T S(t)]\Phi(t, T) \\ &= A\Phi(t, T) - BQ_2^{-1}B^T S(t)\Phi(t, T) \end{aligned} \quad (6.12.15)$$

再对(6.12.10) 求导数, 并把式(6.12.4) 和(6.12.15) 代入, 得到

$$\begin{aligned} \dot{Y}(t) &= \frac{d}{dt}[S(t)\Phi(t, T)X(T)] \\ &= \dot{S}(t)\Phi(t, T)X(T) + S(t)\dot{\Phi}(t, T)X(T) \end{aligned}$$

$$\begin{aligned}
&= [-A^T S(t) - S(t)A - Q_1 + S(t)BQ_2^{-1}B^T S(t)](t, T)X(T) \\
&\quad + S(t)[A - BQ_2^{-1}B^T S(t)](t, T)X(T) \\
&= -Q_1(t, T)X(T) - A^T S(t)(t, T)X(T)
\end{aligned} \quad (6.12.16)$$

令初值为 $X(T) = I$, 则式(6.12.6)变为

$$\frac{d}{dt}[S(t)(t, T)] = -Q_1(t, T) - A^T S(t)(t, T) \quad (6.12.17)$$

把式(6.12.15)与(6.12.7), 式(6.12.17)与(6.12.8)比较可知, 只要令

$$(t, T) = X(t) \quad (6.12.18)$$

它们就是相等的, 这说明 (t, T) 与 $X(t)$ 等价, 已知 (t, T) 存在且非奇异, 则 $X(t)$ 也存在且非奇异。于是完成了证明。

2. 解题步骤

(1) 把(6.12.7) ~ (6.12.9)写成矩阵方程形式:

$$\begin{aligned}
\begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} &= \begin{bmatrix} A & -BQ_2^{-1}B^T \\ -Q_1 & -A \end{bmatrix} \begin{bmatrix} X(t) \\ Y(t) \end{bmatrix}
\end{aligned} \quad (6.12.19)$$

$$\begin{aligned}
\begin{bmatrix} X(T) \\ Y(T) \end{bmatrix} &= \begin{bmatrix} I \\ Q_0 \end{bmatrix}
\end{aligned} \quad (6.12.20)$$

求出矩阵状态方程式(6.12.19)的转移矩阵

$$(t, T) = \begin{bmatrix} x_x(t, T) & x_y(t, T) \\ y_x(t, T) & y_y(t, T) \end{bmatrix} \quad (6.12.21)$$

(2) 由转移矩阵 (t, T) , 求出矩阵状态方程式(6.12.19)的解

$$\begin{aligned}
\begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} &= \begin{bmatrix} x_x & x_y \\ y_x & y_y \end{bmatrix} \begin{bmatrix} I \\ Q_0 \end{bmatrix}
\end{aligned} \quad (6.12.22)$$

即

$$X(t) = x_x + x_y Q_0 \quad (6.12.23)$$

$$Y(t) = y_x + y_y Q_0 \quad (6.12.24)$$

(3) 把式(6.12.23)和(6.12.24)代入式(6.12.6), 计算里卡蒂方程式(6.12.4)的解

$$S(t) = [y_x + y_y Q_0][x_x + x_y Q_0]^{-1} \quad (6.12.25)$$

例 6.12.1 给定标量系统和性能指标函数为

$$dx(t) = 0.5x(t)dt + u(t)dt + dw(t) \quad (6.12.26)$$

$$J = \int_0^T [e^{-t}x^2(t) + 4e^{-t}u^2(t)]dt \quad (6.12.27)$$

求解里卡蒂方程和最优控制。

解 问题较简单, 可直接求解。按给定方程, 已知

$$A = \frac{1}{2} \quad B = 1 \quad Q_1 = e^{-t} \quad Q_2 = 4e^{-t} \quad (6.12.28)$$

把上式数据代入里卡蒂方程式(6.12.4), 并求解, 得到

$$-S = S + e^{-t} - \frac{e^t}{4}S^2 \quad S(T) = 0 \quad (6.12.29)$$

$$S(t) = 2 \begin{bmatrix} 1 - e^{t-T} & e^t + e^{2t-T} - 1 \end{bmatrix} \quad (6.12.30)$$

由上式和式(6.12.28)得到最优控制

$$\begin{aligned} u(t) &= -Q_2^{-1} B^T S(t) x(t) \\ &= -\frac{e^t}{4} [2 \begin{bmatrix} 1 - e^{t-T} & e^t + e^{2t-T} - 1 \end{bmatrix} x(t)] \\ &= \frac{1}{2} [e^{t-T} - 1 \quad e^{t-T} + 1] x(t) \end{aligned} \quad (6.12.31)$$

例 6.12.2 给定一阶系统和性能指标为

$$dx(t) = -\frac{1}{2}x dt + u(t) dt + dw(t) \quad (6.12.32)$$

$$J = Q_0 x^2(T) + \int_0^T [2x^2(t) + u^2(t)] dt \quad (6.12.33)$$

求解里卡蒂方程和最优控制。

解 我们仍直接求解。按给定方程, 已知

$$A = -\frac{1}{2} \quad B = 1 \quad Q_0 = Q_0 \quad Q_1 = 2 \quad Q_2 = 1 \quad (6.12.34)$$

把给定数据代入里卡蒂方程式(6.12.4), 并求解, 得到

$$S = S + S^2 - 2 \quad S(T) = Q_0 \quad (6.12.35)$$

$$\frac{dS}{S^2 + S - 2} = dt \quad (6.12.36)$$

对上式分母配方后, 两边积分给出

$$-\frac{2}{3} \frac{1}{\tanh} \frac{2}{3} S_1 + \frac{1}{2} = t + c_1 \quad (6.12.37)$$

$$-\frac{2}{3} \frac{1}{\coth} \frac{2}{3} S_2 + \frac{1}{2} = t + c_2 \quad (6.12.38)$$

$$S_1(t) = -\frac{1}{2} + \frac{3}{2} \tanh - \frac{3}{2}t - c_1 \quad (6.12.39)$$

$$S_2(t) = -\frac{1}{2} + \frac{3}{2} \coth - \frac{3}{2}t - c_2 \quad (6.12.40)$$

按 $S(T) = Q_0$ 分两种情况讨论:

(1) $T_1 = 1, Q_{01} = 0$ 。可解得

$$c_1 = -1.845 \quad (6.12.41)$$

由此得到里卡蒂方程解 $S_1(t)$ 和最优控制 $u_1(t)$ 为

$$S_1(t) = -\frac{1}{2} + \frac{3}{2} \tanh - \frac{3}{2}t + 1.845 \quad (6.12.42)$$

$$u_1(t) = \frac{1}{2} - \frac{3}{2} \tanh - \frac{3}{2}t + 1.845 \quad (6.12.43)$$

(2) $T_2 = 10, Q_{02} = 10$, 可解得

$$c_2 = -15.143 \quad (6.12.44)$$

由此得到 $S_2(t)$ 和 $u_2(t)$ 为

$$S_2(t) = -\frac{1}{2} + \frac{3}{2} \coth - \frac{3}{2}t + 15.143 \quad (6.12.45)$$

$$u_2(t) = \frac{1}{2} - \frac{3}{2} \coth - \frac{3}{2}t + 15.143 \quad x(t) \quad (6.12.46)$$

以上两个例子中列举的都是一阶系统,情况简单,里卡蒂方程可以直接求解。但对高阶系统,直接求解就相当困难了,实际上甚至是不可能的,这时,可用矩阵微分方程法求解。即使利用这种方法,求解转移矩阵 $\Phi(t, T)$ (6.12.21) 也是困难的。

3. 当 A, B, Q_1 和 Q_2 为常阵时,里卡蒂方程的解

在这种情况下转移矩阵具有解析表达式,从而里卡蒂方程具有解析表达式。这种方法先求出一个相似矩阵,再求有关转移矩阵,最后得到里卡蒂方程解。

考虑矩阵里卡蒂方程

$$-\dot{S}(t) = A^T S(t) + S(t)A + Q_1 - S(t)BQ_2^{-1}B^T S(t) \quad (6.12.47)$$

$$S(0) = Q_0 \quad t \in (-\infty, 0] \quad (6.12.48)$$

式中: A, B, Q_1 和 Q_2 均为常阵, Q_0 和 Q_1 为对称非负阵, Q_2 为对称正定阵。

对式(6.12.19)系数矩阵

$$M = \begin{bmatrix} A & -BQ_2^{-1}B^T \\ -Q_1 & -A \end{bmatrix} \quad (6.12.49)$$

通过非奇异变换

$$T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \quad (6.12.50)$$

把 M 变换为

$$T^{-1}MT = \begin{bmatrix} - & 0 \\ 0 & \end{bmatrix} \quad (6.12.51)$$

上式为 M 的特征值矩阵。因为在给定条件下, M 没有纯虚数的特征值,并且特征值 λ_i 和 $(-\lambda_i)$ ($i=1, 2, \dots, n$), 成对出现,所以 M 的特征值矩阵可以写为式(6.12.51)。对式(6.12.19)进行如下非奇异变换:

$$\begin{bmatrix} \dot{X} \\ \dot{Y} \end{bmatrix} = T \begin{bmatrix} \dot{X} \\ \dot{Y} \end{bmatrix} \quad (6.12.52)$$

则可得

$$\begin{aligned} \begin{bmatrix} \dot{X} \\ \dot{Y} \end{bmatrix} &= T^{-1} \begin{bmatrix} \dot{X} \\ \dot{Y} \end{bmatrix} = T^{-1}M \begin{bmatrix} X \\ Y \end{bmatrix} = T^{-1}MT \begin{bmatrix} X \\ Y \end{bmatrix} \\ &= \begin{bmatrix} - & 0 \\ 0 & \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} \end{aligned} \quad (6.12.53)$$

由条件式(6.12.20)得

$$\begin{bmatrix} X(0) \\ Y(0) \end{bmatrix} = \begin{bmatrix} I \\ Q_0 \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} X(0) \\ Y(0) \end{bmatrix} \quad (6.12.54)$$

$$Y(0) = RX(0) \quad (6.12.55)$$

式中:

$$R = - [T_{22} - FT_{12}]^{-1} [T_{21} - FT_{11}] \quad (6.12.56)$$

求解状态方程式(6.12.53), 并把式(6.12.55)代入, 得到

$$\begin{aligned} \begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} &= \begin{bmatrix} e^{-t} & 0 \\ 0 & e^{-t} \end{bmatrix} \begin{bmatrix} X(0) \\ Y(0) \end{bmatrix} \\ &= \begin{bmatrix} e^{-t} & 0 \\ 0 & e^{-t} \end{bmatrix} \begin{bmatrix} X(0) \\ Y(0) \end{bmatrix} \end{aligned} \quad (6.12.57)$$

由上式给出

$$Y(t) = e^{-t} Re^{-t} X(t) \quad (6.12.58)$$

再由上式和式(6.12.52)导出

$$\begin{aligned} Y(t) &= T_{21}X(t) + T_{22}Y(t) \\ &= [T_{21} + T_{22}e^{-t}Re^{-t}]X(t) \\ &= [T_{21} + T_{22}e^{-t}Re^{-t}][T_{11} + T_{12}e^{-t}Re^{-t}]^{-1}X(t) \end{aligned} \quad (6.12.59)$$

最后由上式和式(6.12.6)得到里卡蒂方程的解为

$$\begin{aligned} S(t) &= Y(t)X^{-1}(t) \\ &= [T_{21} + T_{22}e^{-t}Re^{-t}][T_{11} + T_{12}e^{-t}Re^{-t}]^{-1} \end{aligned} \quad (6.12.60)$$

由以上解的表达式(6.12.60)可以看出:

(1) 当 $t \rightarrow \infty$ 时, $e^{-t} \rightarrow 0$, 从而有稳态值

$$S = T_{21}T_{11}^{-1} \quad (6.12.61)$$

(2) $S(t)$ 的稳态值 S 与初值 F 无关, F 只影响暂态。

例 6.12.3 给定稳定里卡蒂矩阵微分方程

$$-\dot{S} = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} S + S \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} - S \begin{bmatrix} 1 & 0 \end{bmatrix} S \quad (6.12.62)$$

$$S(0) = 0 \quad (6.12.63)$$

试求 $S(t)$ 表达式。

解 首先构造 M 阵:

$$M = \begin{bmatrix} A & -BQ_2^{-1}B^T \\ -Q_1 & -A \end{bmatrix} = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 1 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & -1 & 0 & \frac{1}{2} \end{bmatrix} \quad (6.12.64)$$

求出 M 阵特征值为 $\lambda_1 = \lambda_2 = 1$, $\lambda_3 = \lambda_4 = -1$, 则有

$$T^{-1}MT = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (6.12.65)$$

求出变换矩阵 T 为

$$T = \begin{bmatrix} \frac{1}{2} & 1 & -\frac{1}{2} & 1 \\ 1 & 0 & -1 & 0 \\ 1 & -1 & 1 & 1 \end{bmatrix} \quad (6.12.66)$$

T 的分块矩阵和 R 阵分别为

$$\begin{aligned} T_{11} &= \begin{bmatrix} \frac{1}{2} & 1 \\ 1 & 0 \end{bmatrix} & T_{12} &= \begin{bmatrix} -\frac{1}{2} & 1 \\ -1 & 0 \end{bmatrix} \\ T_{21} &= \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} & T_{22} &= \begin{bmatrix} 1 & 1 \end{bmatrix} \end{aligned} \quad (6.12.67)$$

$$R = -[T_{22} - FT_{12}]^{-1}[T_{21} - FT_{11}] = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \quad (6.12.68)$$

求式(6.12.53)的转移矩阵, 得到

$$e^{-t} = \begin{bmatrix} e^{-t} & te^{-t} \\ 0 & e^{-t} \end{bmatrix} \quad (6.12.69)$$

$$e^{-t}Re^{-t} = \begin{bmatrix} e^{-2t} - 2te^{-2t} & 2te^{-2t} - 2t^2e^{-2t} \\ -2e^{-2t} & e^{-2t} - 2te^{-2t} \end{bmatrix} \quad (6.12.70)$$

$$T_{21} + T_{22}e^{-t}Re^{-t} = \begin{bmatrix} 1 - e^{-2t} + 2te^{-2t} & -2te^{-2t} + 2t^2e^{-2t} \\ 1 - e^{-2t} - 2te^{-2t} & -1 + e^{-2t} - 2t^2e^{-2t} \end{bmatrix} = M_1 \quad (6.12.71)$$

$$\begin{aligned} &T_{11} + T_{12}e^{-t}Re^{-t} \\ &= \begin{bmatrix} 1 + 3e^{-2t} - 2te^{-2t} & 1 - e^{-2t} + 4te^{-2t} - 2t^2e^{-2t} \\ (-\frac{1}{2} + 1) + (3 - \frac{1}{2})e^{-2t} - 2(-\frac{1}{2} - 1)te^{-2t} & (-\frac{1}{2} + 2) + (-\frac{1}{2} + 2)e^{-2t} + 2te^{-2t} - 2(-\frac{1}{2} - 1)t^2e^{-2t} \end{bmatrix} \\ &= M_2 \end{aligned} \quad (6.12.72)$$

最后得到里卡蒂矩阵微分方程式(6.12.62)和(6.12.63)的解为

$$S(t) = M_1 M_2^{-1}$$

稳态解为

$$S = \begin{bmatrix} 2 & -\frac{3}{2} \\ 3 & -2 \end{bmatrix} \quad (6.12.73)$$

习 题

6.1 设 u 表示控制向量, y 表示输出, x 表示状态, 假设 x 是高斯 $N(m, R)$ 的。向量 x, y 和 u 的关系为

$$y = Au + x$$

考虑指标函数为

$$J = q_0 + q_1^T y + q_2^T u + \frac{1}{2} y^T Q_1 y + y^T Q_{12} u + \frac{1}{2} u^T Q_2 u$$

的最优化问题。当没有状态信息(即控制 u 只是先验信息 y 的函数)时, 或当有完全状态信息(即允许 u 是 x 的函数)时, 试确定使期望指标为极小的控制 u 和极小期望指标。

6.2 考虑系统模型

$$y = Au + x$$

指标函数为

$$l = q_0 + q_1^T y + q_1^T u + \frac{1}{2} y^T Q_1 y + y^T Q_{12} u + \frac{1}{2} u^T Q_2 u$$

当选择参数

$$\begin{aligned} A = & \begin{bmatrix} 0.666 & -0.188 & 0.671 \\ -0.052 & -0.296 & 0.259 \\ 0.285 & 2.358 & -1.427 \end{bmatrix} & m = & \begin{bmatrix} -5.59 \\ -3.704 \\ -0.729 \end{bmatrix} \\ q_0 = q_1 = q_2 = Q_{12} = 0 & & Q_1 = Q_2 = I \end{aligned}$$

时, 试确定不完全状态信息的最优控制。

6.3 如果假设系统中出现的随机向量是不相关向量, 而不是独立向量, 如果容许控制策略 $u(k)$ 限于是 $X(k)$ 的线性函数, 试证明定理 6.3.1 的结论仍然成立。

6.4 考虑指标函数为

$$l = \sum_{k=1}^N x^2(k)$$

的系统

$$x(k+1) = ax(k) + bu(k) + (k)$$

设容许控制策略 $u(k)$ 是 $X(k)$ 的函数, 试确定使期望指标为极小的控制策略。

6.5 考虑指标函数为

$$l = \sum_{k=k_0}^N x^2(k)$$

的标量系统

$$x(k+1) = ax(k) + bu(k) + (k)$$

当容许控制策略 $u(k)$ 是 $X(t-1)$ 的函数时, 试确定使期望指标为极小的控制策略。

6.6 考虑系统

$$x(k+1) = x(k) + u(k) + (t)$$

$$y(k) = x(k) + (k)$$

式中: $\{ (k) \}$ 和 $\{ (k) \}$ 是两个均值为零和协方差分别为 r_1 和 r_2 的独立高斯随机变量序列。假设初始状态是高斯 $N(m, \sigma^2)$ 的, 并设指标函数为

$$l = \sum_{k=1}^N [x^2(k) + qu^2(k)]$$

当

- (1) $r_1 = r_2 = 0$
- (2) $r_1 = 0, r_2 = 0$
- (3) $r_1 = 0, r_2 = 0$

试确定使期望指标为极小的控制策略。在三种情况下, 假设 $u(k)$ 是 $y(k-1), y(k-2), \dots$ 的函数。

6.7 考虑上题。当 $N \rightarrow \infty$ 时, 试确定表示最优控制策略的脉冲传递函数。给出期望指标的极小值的渐近表达式, 并对各项给出物理解释。

6.8 考虑系统

$$x(k+1) = Ax(k) + e(k)$$

式中: $\{e(k)\}$ 是独立高斯 $N(0, R_1(k))$ 随机变量序列, 初始状态 $x(k_0)$ 是高斯 $N(m_0, R_0)$ 的。试证明

$$\begin{aligned} E \{ x^T(k_1) Q_0 x(k_1) + \sum_{k=k_0}^{k_1-1} x^T(k) Q_1 x(k) \} \\ = m^T S(k_0) m + \text{tr} S(k_0) R_0 + \sum_{k=k_0}^{k_1-1} \text{tr} S(k+1) R_1(k) \end{aligned}$$

式中:

$$\begin{aligned} S(k) &= A^T S(k+1) A + Q_1 \\ S(k_1) &= Q_0 \end{aligned}$$

6.9 考虑性能指标为

$$E \{ I \} = E \{ x^T(N) Q_0 x(N) + \sum_{t=k_0}^{N-1} x^T(t) Q_1 x(t) + u^T(t) Q_2 u(t) \}$$

的系统

$$x(k+1) = Ax(k) + Bu(k) + w(k)$$

设容许控制策略 $u(k)$ 是状态 $X(k-1)$ 的函数, 试证明最优策略为

$$u(k) = - [L(k)]^{-1} B^T [X(k-1) - L(k)]$$

式中:

$$L(k) = [Q_2 + B^T S(k+1) B]^{-1} B^T S(k+1)$$

再确定性能指标函数的极小值。

6.10 考虑性能指标为

$$I = x^T(N) Q_0 x(N) + \sum_{t=k_0}^{N-1} [x^T(t) Q_1 x(t) + u^T(t) Q_2 u(t)]$$

的系统

$$x(k+1) = Ax(k) + Bu(k) + w(k)$$

设容许控制 $u(k)$ 只是先验数据的函数(开环控制)。试证明最优控制信号为

$$u(k) = - [L(k)]^{-1} B^T m(k)$$

式中:

$$\begin{aligned} m(k+1) &= A m(k) + B u(k) \\ &= [A - B L(k)] m(k) \\ m(k_0) &= E \{ x(k_0) \} = m_0 \end{aligned}$$

再确定期望指标的极小值。

6.11 考虑性能指标为

$$I = x^T(N)Q_0x(N) + \sum_{t=k_0}^{N-1} [x^T(t)Q_1x(t) + u^T(t)Q_2u(t)]$$

的系统

$$x(k+1) = x(k) + u(k) + (k)$$

在下列不同情况下:

- (1) $u(k)$ 是 $X(k)$ 的函数
- (2) $u(k)$ 是 $m(k)$ 的函数
- (3) $u(k)$ 是 $X(k-1)$ 的函数

试比较最优控制策略的结果。

6.12 考虑系统

$$\begin{aligned} x(k+1) &= ax(k) + u(k-p) + (k) \\ y(k) &= x(k) + (k) \end{aligned}$$

式中: $\{(k)\}$ 和 $\{(k)\}$ 是两个均值为零和方差分别为 r_1 和 r_2 的独立高斯随机变量序列。设初始状态 $x(t_0)$ 是高斯(m_0, σ^2) 的, 指标函数为

$$I = \sum_{t=1}^N [x^2(t) + qu^2(t)]$$

当容许控制策略 $u(k)$ 是 $y(k), y(k-1), \dots$ 的函数时, 试确定使期望指标为极小的控制策略。

6.13 考虑由下式描述的系统:

$$y(k) + ay(k-1) = u(k-1) + bu(k-2) + e(k) + ce(k-1)$$

式中: $\{e(t)\}$ 是独立高斯 $N(0, 1)$ 随机变量序列。试证明系统能表示为

$$\begin{aligned} x(k+1) &= \begin{bmatrix} -a & 1 \\ 0 & 0 \end{bmatrix} x(k) + \begin{bmatrix} 1 \\ b \end{bmatrix} u(k) + \begin{bmatrix} 1 \\ c \end{bmatrix} e(k+1) \\ y(t) &= [1 \quad 0] x(k) \end{aligned}$$

当容许控制策略是使 $u(k)$ 是 $y(k), y(k-1), \dots$ 的函数时, 试确定使期望指标 $E x_1^2(k+1)$ 为极小的控制策略。

6.14 考虑系统

$$\begin{aligned} x(k+1) &= x(k) + u(k) + (k+1) \\ y(k) &= x(k) \end{aligned}$$

式中:

$$\begin{aligned} &= \begin{bmatrix} -a & 1 \\ 0 & 0 \end{bmatrix} & & = \begin{bmatrix} 1 \\ b \end{bmatrix} & & = \begin{bmatrix} 1 \\ c \end{bmatrix} \\ &= [1 \quad 0] & R_1 = 1 & & R_2 = 0 \end{aligned}$$

当容许控制 $u(k)$ 是 Y_k 的函数时, 试确定使期望指标

$$I = E \frac{1}{N} \sum_{t=1}^N y^2(k)$$

为极小的控制策略。

6.15 考虑系统

$$x(k+1) = x(k) + w(k)$$

式中: $\{w(k)\}$ 是独立高斯 $N(0, R_1(k))$ 随机变量序列; 初始状态 $x(k_0)$ 是高斯 $N(m_0, R_0)$ 的, 且与 $w(k)$ 独立。试证明

$$\begin{aligned} I &= E \{ x^T(N) Q_0 x(N) + \sum_{t=k_0}^{N-1} x^T(t) Q_1 x(t) \} \\ &= m^T S(k_0) m + \text{tr} S(k_0) R_0 + \sum_{t=t_0}^{N-1} \text{tr} S(k+1) R_1(k) \\ S(k) &= S(k+1) + Q_1 \\ S(N) &= Q_0 \end{aligned}$$

6.16 考虑系统

$$x(k+1) = x(k) + u(k) + w(k)$$

式中: $\{w(k)\}$ 是独立高斯 $N(0, R_1(t))$ 随机变量序列; 初始状态 $x(t_0)$ 是高斯 $N(m_0, R_0)$ 的, 且与 $w(k)$ 独立。指标函数为

$$I = x^T(N) Q_0 x(N) + \sum_{t=k_0}^{N-1} [x^T(t) Q_1 x(t) + u^T(t) Q_2 u(t)]$$

试求最优开环控制 $[u(k) \text{ 是 } m(k) \text{ 的函数}]$ 策略, 并确定期望指标的极小值。

6.17 考虑系统

$$x(k+1) = x(k) + u(k) + w(k)$$

式中: $\{w(k)\}$ 是独立高斯 $N(0, R_1(k))$ 随机变量序列; 初始状态 $x(k_0)$ 是高斯 $N(m_0, R_0)$ 的, 且与 $w(k)$ 独立。指标函数为

$$I = x^T(N) Q_0 x(N) + \sum_{t=k_0}^{N-1} [x^T(t) Q_1 x(t) + u^T(t) Q_2 u(t)]$$

设 $u(k)$ 是 $x(k-1)$ 的函数, 试求出最优控制策略, 并确定期望指标的极小值。

6.18 考虑系统

$$dx = u dt + dw$$

$$dy = x dt + dv$$

式中: $\{w(t)\}$ 和 $\{v(t)\}$ 是方差参数分别为 r_1 和 r_2 的维纳过程; 初始状态是高斯 $N(m_0, r_0)$ 的。设指标函数为

$$I = \int_0^T [x^2(t) + qu^2(t)] dt$$

控制的目的是使期望指标为极小。试确定开环控制、完全状态信息和不完全状态信息时的最优控制策略。

6.19 考虑上一题的系统, 试确定当 $t \rightarrow \infty$ 时不同情况的定态控制律和期望指标的极小值。

6.20 考虑系统

$$dx = Ax dt + dw$$

式中: $\{w(t)\}$ 是增量协方差为 $R_1 dt$ 的维纳过程; 初始状态 $x(t_0)$ 是高斯 $N(m_0, R_0)$ 的。当

$$-\frac{dS}{dt} = A^T S + SA + Q_1$$

$$S(t_1) = Q_0$$

时, 试证明

$$\begin{aligned} E \quad & x^T(t_1)Q_0x(t_1) + \int_{t_0}^{t_1} x^T(t)Q_1x(t)dt \\ &= m^TS(t_0)m + \int_{t_0}^{t_1} \text{tr}S(t)R_0 + \int_{t_0}^{t_1} \text{tr}S(t)R_1(t)dt \end{aligned}$$

6.21 考虑由

$$dx = Axdt + udt + dw$$

描述的系统, 试证明最优开环系统的指标函数为

$$\begin{aligned} \min J &= E \quad x^T(t_1)Q_0x(t_1) + \int_{t_0}^{t_1} [x^T(t)Q_1x(t) + u^T(t)Q_2u(t)]dt \\ &= m^TS(t_0)m + \int_{t_0}^{t_1} \text{tr}R(t)R_0 + \int_{t_0}^{t_1} \text{tr}R(t)R_1dt \end{aligned}$$

式中:

$$-\frac{dR}{dt} = A^TR + RA + Q_1$$

$$R(t_1) = Q_0$$

6.22 设 P 和 Q 是里卡蒂方程

$$\frac{dP}{dt} = AP + PA^T + R_1 - PC^TR_2^{-1}CP$$

$$P(t_0) = P_0$$

$$\frac{dQ}{dt} = AQ + QA^T + R_1 - QC^TR_2^{-1}CQ$$

$$Q(t_0) = Q_0$$

的解, 试对里卡蒂方程有解的所有 t, 证明 $P_0 > Q_0$ 蕴涵着 $P(t) > Q(t)$ 。

6.23 考虑里卡蒂方程

$$\frac{dP}{dt} = AP + PA^T + R_1 - PC^TR_2^{-1}CP$$

$$P(t_0) = P_0$$

$$\frac{dQ}{dt} = AQ + QA^T + R_3 - QC^TR_2^{-1}CQ$$

$$Q(t_0) = P_0$$

试对使方程有解的所有 t, 证明 $R_1 > R_3$ 蕴涵着 $P(t) > Q(t)$ 。

6.24 考虑里卡蒂方程

$$\frac{dP}{dt} = AP + PA^T + R_1 - PC^TR_2^{-1}CP$$

$$P(t_0) = P_0$$

$$\frac{dQ}{dt} = AQ + QA^T + R_1 - QC^TR_3^{-1}CQ$$

$$Q(t_0) = P_0$$

试对方程有解的所有 t , 证明 $R_2 = R_3$ 蕴涵着 $P(t) = Q(t)$ 。

6.25 考虑里卡蒂方程

$$\frac{dP}{dt} = AP + PA^T + R_1$$

$$P(t_0) = P_0$$

若设

$$P = QQ^T \qquad P_0 = Q_0Q_0^T$$

试证明矩阵满足微分方程

$$\frac{dQ}{dt}Q^T = AQQ^T + \frac{1}{2}R_1$$

$$Q(t_0) = Q_0$$

6.26 假设里卡蒂方程

$$\frac{dP}{dt} = AP + PA^T + R_1 - PC^TR_2^{-1}CP$$

$$P(t_0) = R_0$$

有对所有 t 有正定的解, 试证明逆 P^{-1} 满足方程

$$\frac{dP^{-1}}{dt} = -P^{-1}A - A^TP^{-1} - P^{-1}R_1P^{-1} + C^TR_2^{-1}C$$

$$P^{-1}(t_0) = R_0^{-1}$$

6.27 试给出一种求解里卡蒂方程

$$\frac{dP}{dt} = AP + PA^T - PC^TR_2^{-1}CP$$

$$P(t_0) = R_0$$

的简化算法。

参 考 文 献

1. Astrom K J. Introduction to Stochastic Control Theory. New York: Academic Press, 1970
2. Davis M H A. Linear Estimation and Stochastic Control. London: Chapman and Hall, 1977
3. Maybeck P S. Stochastic Models, Estimation, and Control. Volume 1. New York: Academic Press, 1979
4. Maybeck P S. Stochastic Models, Estimation, and Control. Volume 2. New York: Academic Press, 1982
5. Maybeck P S. Stochastic Models, Estimation, and Control. Volume 3. New York: Academic Press, 1982
6. Mahmoud M S, Singh M G. Discrete Systems Analysis, Control and Optimization. Berlin: Springer-Verlag, 1984
7. Freidlin M I, Wentzell A D. Random Perturbations of Dynamical Systems. Berlin: Springer-Verlag, 1984
8. Stengel R F. Stochastic Optimal Control Theory and Application. New York: John Wiley & Sons, 1986
9. Arkin V I, Shiraev A, Wets R. Stochastic Optimization, Lecture Notes in Control and Information Sciences. Berlin: Springer-Verlag, 1986
10. Bucy R S, Moura J M F. Nonlinear Stochastic Problems. Dordrecht: D Reidel Publishing Company, 1983
11. Skelton R E, Iwasaki T. Liapunov and Covariance Controller. Int J Control, 1993, 57(3): 519 ~ 536
12. Grimmett G, Welsh D. Probability, an Introduction. Oxford: Clarendon Press, 1986
13. Barnett S. Matrix Methods for Engineers and Scientists. London: McGraw-Hill Book Company (UK) Limited, 1979
14. Oksendal B. Stochastic Differential Equations, an Introduction with Applications. Berlin: Springer-Verlag, 1985
15. Kalman R E. A New Approach to Linear Filtering and Prediction Problems. Trans ASME J Basic Eng, 1960, 82: 34 ~ 45
16. Kalman R E, Bucy R S. New Results in Linear Filtering and Prediction Theory. Trans ASME J Basic Eng, 1961, 83: 95 ~ 108
17. Bucy R S. Global Theory of the Riccati Equation. J Comput System Sci, 1967, 1: 349 ~ 361
18. Bucy R S. Optimal Filtering for Correlated Noise. J Math Anal Appl, 1967, 20: 1 ~ 8
19. Bellman R. Dynamic Programming. New Jersey: Princeton University Press, 1957
20. Bellman R. Adaptive Control Processes. A Guided Tour. New Jersey: Princeton University Press, 1961
21. Bellman R, Dreyfus S E. Applied Dynamic Programming. New Jersey: Princeton University Press, 1962
22. 李树英, 许茂增. 随机系统的滤波与控制. 北京: 国防工业出版社, 1991
23. 郑政谋, 朱志祥. 随机控制引论. 西安: 西安电子科技大学出版社, 1991
24. 蔡尚峰. 随机控制理论. 上海: 上海交通大学出版社, 1987

25. 韩崇昭,王月娟,万百五. 随机系统理论. 西安: 西安交通大学出版社, 1987
26. 陈翰馥. 离散时间系统的递推估计与随机控制. 北京: 科学出版社, 1980
27. 麦迪成 J S 著. 赵希人译. 随机最优线性估计与控制. 哈尔滨: 黑龙江人民出版社, 1984
28. 鲍齐克 S M 著. 凌云旦译. 数字滤波和卡尔曼滤波. 北京: 科学出版社, 1984
29. 布赖森 A E, 何毓琦著. 钱洁文, 等译. 应用最优控制—最优化·估计·控制. 北京: 国防工业出版社, 1982
30. 王照林. 现代控制理论基础. 北京: 国防工业出版社, 1981
31. 谢绪恺. 现代控制理论基础. 沈阳: 辽宁人民出版社, 1981
32. 解学书. 最优控制理论与应用. 北京: 清华大学出版社, 1986
33. 帕普力斯 A 著. 谢国瑞译. 概率、随机变量与随机过程. 北京: 高等教育出版社, 1984
34. 布莱克曼 P F 著. 华南工学院自动控制教研组译. 状态变量分析导论. 北京: 科学出版社, 1982
35. 黄琳. 系统与控制理论中的线性代数. 北京: 科学出版社, 1984
36. 须田信英著. 曹长修译. 自动控制中的矩阵理论. 北京: 科学出版社, 1979