

**Version 2.01**

编译日期: 2020-12-27

任何建议及错误信息请发送至邮箱  
ocnzhao@163.com

# 随机信号分析

赵国亮 整理

$$\rho := \frac{1 + \sqrt{-3}}{2}$$



本作品采用知识共享署名-非商业性使用 4.0 国际许可协议  
进行许可. 访问 [查看该许可协议](#).

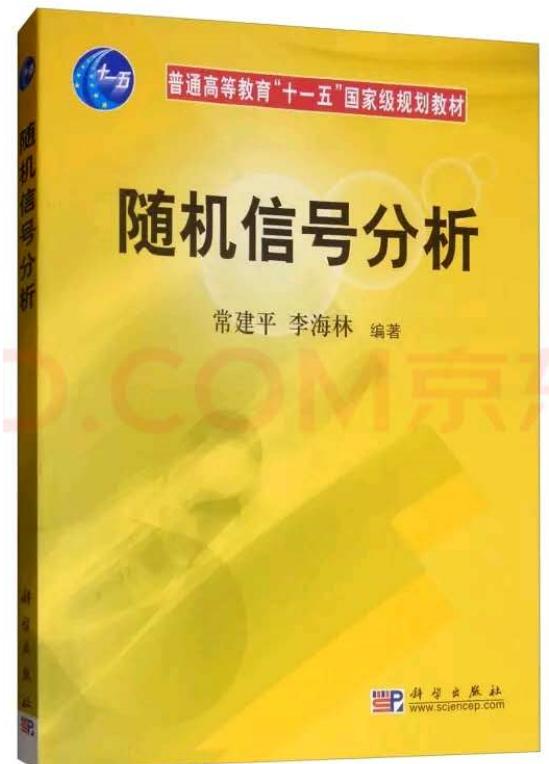
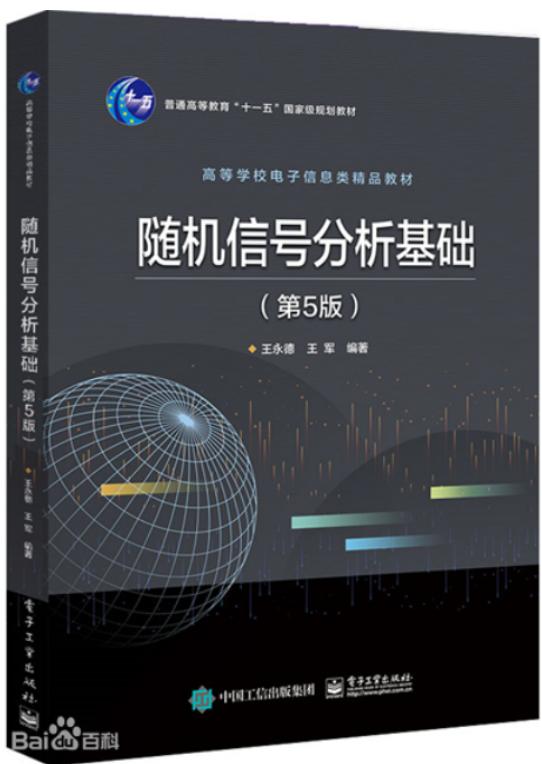
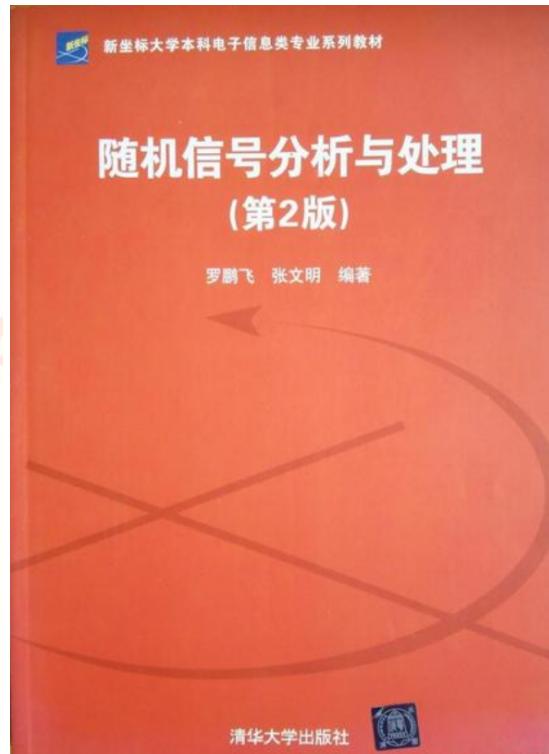
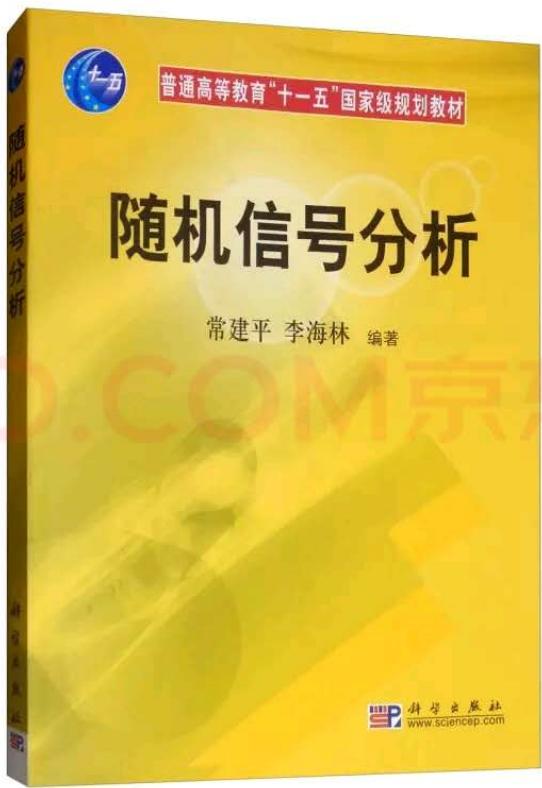


# 前　言

本资料汇编自网上，请勿用于商业用途，作者对此不负任何责任。

参考书籍

- 1) Umberto Spagnolini, 《[工程中的统计信号处理-Statistical Signal Processing in Engineering](#)》 , 2017, [[Spagnolini2017SSPE](#)].
- 2) [Statistical Signal Processing](#).
- 3) Chonavel, T. [Statistical Signal Processing: Modelling and Estimation](#), 2002, [[Chonavel2002-9862](#)].



# 目 录

前 言	i
概 览	ix
<b>第一章 概率论</b>	<b>1</b>
1.1 概率论概要 . . . . .	1
1.1.1 随机事件及其概率 . . . . .	1
1.1.2 随机变量及其分布 . . . . .	5
1.1.3 多维随机变量 . . . . .	19
1.1.4 随机变量函数的分布 . . . . .	23
1.1.5 随机变量的数字特征 . . . . .	41
1.1.6 条件数学期望 . . . . .	49
1.1.7 随机变量的方差和矩 . . . . .	51
1.1.8 相关、正交和独立 . . . . .	59
1.1.9 随机变量的特征函数 . . . . .	65
1.1.10 高斯随机变量 . . . . .	74
1.1.11 复随机变量 . . . . .	82
1.2 线性代数导论 . . . . .	84
1.2.1 矩阵的概念和基本运算 . . . . .	84
1.2.2 特殊矩阵 . . . . .	88
1.2.3 矩阵的逆 . . . . .	92
1.2.4 矩阵分解 . . . . .	93
1.2.5 子空间 . . . . .	98
1.2.6 梯度分析 . . . . .	99
1.3 习题 . . . . .	106

<b>第二章 随机信号的时域分析</b>	<b>111</b>
2.1 随机过程 . . . . .	112
2.1.1 随机过程的基本概念 . . . . .	112
2.1.2 随机过程的分类 . . . . .	114
2.1.3 随机过程的分布 . . . . .	118
2.1.4 随机过程的数字特征 . . . . .	121
2.1.5 随机过程的特征函数 . . . . .	128
2.2 平稳与非平稳随机过程 . . . . .	130
2.2.1 平稳随机过程的分类 . . . . .	130
2.2.2 平稳随机过程自相关函数的性质 . . . . .	135
2.2.3 非平稳随机过程 . . . . .	141
2.3 两个随机过程联合的统计特性 . . . . .	142
2.4 复随机过程 . . . . .	148
2.5 随机过程的积分微分特性 . . . . .	152
2.6 高斯随机过程 . . . . .	154
2.7 各态历经过程 . . . . .	161
2.7.1 各态历经过程 . . . . .	161
2.7.2 广义各态历经过程 . . . . .	163
2.8 小结 . . . . .	168
2.9 习题 . . . . .	169
<b>第三章 随机信号的频域分析</b>	<b>173</b>
3.1 实随机过程的功率谱密度 . . . . .	173
3.1.1 实随机过程的功率谱密度 . . . . .	174
3.1.2 实平稳过程的功率谱密度与自相关函数之间的关系 . . . . .	179
3.2 两个实随机过程的互功率谱密度 . . . . .	186
3.3 理想白噪声 . . . . .	189
3.4 习题 . . . . .	193
<b>第四章 随机信号通过线性系统的分析</b>	<b>197</b>
4.1 信号与系统概述 . . . . .	197
4.1.1 信号及其分类 . . . . .	197
4.1.2 系统及其分类 . . . . .	200
4.1.3 线性时不变系统 . . . . .	202

4.2	随机信号通过线性时不变系统 . . . . .	204
4.2.1	系统输出的均值 . . . . .	205
4.2.2	系统输出的自相关函数和功率谱密度函数 . . . . .	206
4.2.3	系统输入与输出的互相关函数和互功率谱密度函数 . . . . .	207
4.2.4	系统输出的概率密度 . . . . .	214
4.3	随机序列通过线性时不变系统 . . . . .	214
4.3.1	系统输出的均值 . . . . .	215
4.3.2	系统输出的自相关函数 $R_Y(n_1, n_2)$ 和功率谱密度函数 $G_Y(\Omega)$ . . . . .	216
4.3.3	系统输入与输出的互相关函数和互功率谱密度函数 . . . . .	217
4.4	色噪声的产生与白化滤波器 . . . . .	218
4.5	白噪声通过线性时不变系统 . . . . .	221
4.5.1	系统输出的一般特性及等效噪声带宽 . . . . .	221
4.5.2	白噪声通过理想低通系统 . . . . .	223
4.5.3	白噪声通过理想带通系统 . . . . .	224
4.5.4	白噪声通过具有高斯频率特性的带通系统 . . . . .	225
4.6	白噪声序列和平稳随机序列的参数模型 *	227
4.6.1	自回归滑动平均模型 . . . . .	227
4.6.2	自回归模型 . . . . .	228
4.6.3	滑动平均模型 . . . . .	229
4.6.4	三种模型间的联系 . . . . .	230
4.7	习题 . . . . .	231
<b>第五章</b>	<b>随机信号统计特征的实验研究方法</b>	<b>237</b>
5.1	统计特征实验研究的基础 . . . . .	237
5.2	随机信号时域特征的估计 . . . . .	239
5.3	随机信号功率谱密度的估计 . . . . .	246
5.3.1	经典谱估计 . . . . .	247
5.3.2	经典谱估计 . . . . .	247
5.3.3	经典谱估计的改进 . . . . .	250
5.3.4	现代谱估计简介 . . . . .	252
5.4	习题五 . . . . .	256
5.5	知识点总结 . . . . .	257

<b>第六章 窄带随机信号</b>	<b>259</b>
6.1 预备知识 . . . . .	259
6.1.1 信号的解析形式 . . . . .	259
6.1.2 希尔伯特变换的性质 . . . . .	262
6.1.3 高频窄带信号的复指数形式 . . . . .	267
6.1.4 高频窄带信号通过窄带系统 . . . . .	272
6.1.5 随机过程的解析形式及其性质 . . . . .	274
6.2 窄带随机过程 . . . . .	277
6.2.1 窄带随机过程的数学模型及复指数形式 . . . . .	277
6.2.2 窄带随机过程的“垂直”分解 . . . . .	278
6.2.3 窄带随机过程的统计分析 . . . . .	280
6.3 窄带高斯过程包络与相位的分布 . . . . .	286
6.3.1 包络和相位的一维概率分布 . . . . .	287
6.3.2 包络和相位各自的二维概率分布 . . . . .	290
6.3.3 随相余弦信号与窄带高斯噪声之和的包络及相位的概率分布 . . . . .	293
6.4 窄带高斯过程包络平方的概率分布 . . . . .	297
6.4.1 窄带高斯噪声包络平方的分布 . . . . .	298
6.4.2 余弦信号加窄带高斯噪声包络平方的概率分布 . . . . .	298
6.4.3 卡方分布和非中心卡方分布 . . . . .	299
6.5 习题 . . . . .	306
<b>第七章 马尔可夫过程、独立增量过程及独立随机过程</b>	<b>311</b>
7.1 马尔可夫过程 . . . . .	312
7.1.1 马尔可夫序列 . . . . .	312
7.1.2 马尔可夫链 . . . . .	316
7.1.3 马尔可夫过程 . . . . .	328
7.2 独立增量过程 . . . . .	331
7.2.1 概述 . . . . .	332
7.2.2 泊松过程 . . . . .	334
7.2.3 泊松过程的应用实例 . . . . .	340
7.2.4 维纳过程 . . . . .	349
7.3 独立随机过程 . . . . .	353
7.4 习题 . . . . .	355
7.5 参考文献 . . . . .	360





# 概 览

推荐参考文献-来自科学网段法兵教授

**【1】** Signal Detection in Non-Gaussian Noise, Wegman, Edward J. and Kassam, Saleem A., Springer Texts in Electrical Engineering, 1993, [**Wegman1993Signal**]

Kassam 的这本非高斯噪声中的信号检测比较早接触的一本书，书中对于微弱信号检测的理论推导非常严密且易于理解，由此我也对于非线性检测器的性能有了更加深刻的理解，其中有关非高斯噪声的 Fisher 信息对于我的启发最大，背景噪声的 Fisher 信息其实蕴含了将信号能否被更好地检测的一个基本指标，不过这个必须是在弱信号的假设下才能成立，这本书中有关随机信号检测部分还有待进一步理解。

在信号的分析与处理中常常需要用到不少数学基础知识，因此在本章中将对统计信号处理中可能用到的数学知识作简单的介绍。

**【2】** Fundamentals of Statistical Signal Processing, S. M. Kay, Prentice-Hall, 1998, [**SMKay1998FundamentalSSP**]

这部书包含了两个部分，估计和检测，基本都是离散信号处理，与 Trees 的书可以说正好相互契合，一部离散框架，一部连续，两部书都是浩瀚精深，目前只是对于 Kay 的书摸得比较多些，并且还能荣幸地得到手写的习题解答，Trees 的书读还是不多。时常感叹，终究资质有限，即使皓首穷经，也难以望其项背。

**【3】** Elements of Information Theory, T. M. Cover and J. A. Thomas, [**CoverThomas1991EIT**]

这本书出自 1990 年，但是 Cover 作为香农的门徒，是现在信息论的最大权威，老早就买了这本书，这本书和不同学科之间联系以及各种数学的处理无处不挥发着信息理论的强大气场，信息时代说法由来已久，但是信息理论完美的数学理论是该领域最吸引人的地方，不仅仅是数学，其应用是无止境的。





# 1

## 概率论

1.1 节主要是对概率论内容的概述；因此关于数学基础知识的阐述以基本概念和有用的结论为主，而命题的严格证明较少涉及，对此感兴趣的读者可以参阅有关参考文献。

### 1.1 概率论概要

#### 1.1.1 随机事件及其概率

自然界与人类社会的众多现象大致可分为两类，分别称为确定性现象与随机现象。所谓确定性现象，即在一定条件下必然会出现某一结果（或发生某一事件）的现象。例如，纯净水在一个大气压下加热至 100 摄氏度时，必然沸腾；物体以 10 米 / 秒的速度做匀速直线运动 1 分钟，其走过的路程必为 600 米。这类确定性现象由确定的规律所控制，从数量的角度来研究，从而产生了量与量之间确定的函数关系。

所谓随机现象，即在一定条件下可能出现不同结果（或发生不同事件），且不能准确预言究竟出现哪一种结果的现象。例如，相同条件下掷一枚硬币，可能正面向上，也可能反面向上，且在未掷之前无法准确预言究竟哪一面向上；二元数字通信系统发送的信号可能是“1”也可能是“0”，接收机在接收之前无法准确预言接收结果是信号“1”还是信号“0”。这一类现象广泛存在于自然界与社会活动中，而概率论正是探索研究这类随机现象客观规律的一门学科。

本节首先介绍随机事件及其概率，并在此基础上分析随机变量的分布和数字特征，最后对常用的多维随机变量作简单的概述。

### 定义 1.1 基本事件

观察并研究随机现象的手段与过程称为试验。当试验满足下述条件时，称之为随机试验，简称试验，记为  $E$ .



随机试验具有如下特征：

- (1) 试验可在相同条件下重复进行(可重复性);
- (2) 试验可能出现的结果不止一个，并明确知道所有可能的结果;
- (3) 每次试验总是恰好出现这些可能结果中的一个，但是在一次试验之前不能准确预言哪一种结果会出现(结果出现具有随机性)。

如：掷一颗骰子并观察出现的点数，从一批产品中任意抽取若干件，观察其中的次品数等都是随机试验。概率论所要研究的是随机试验中出现的各种情况，为了方便研究，对试验的有关结果给出如下概念。

### 定义 1.2 基本事件

某一随机试验中可能出现的每一结果称为该试验的一个基本事件(样本点)，记为  $e$ . 所有基本事件构成的集合称为该试验的样本空间，记为  $\Omega$ , 由样本空间  $\Omega$  中的若干基本事件构成的子集合称为该试验中的随机事件，简称为事件，记为  $A, B, C$ 。当属于事件  $A$  的某一基本事件发生时，称事件  $A$  发生。



我们在研究随机现象时，不仅需要知道可能会出现哪些事件，更重要且更具实践意义的是了解和研究各种事件发生可能性的大小，并加以度量。我们把刻画事件  $A$  发生可能性大小的数量指标称为事件  $A$  的概率，记为  $P(A)$ .

下面给出计算  $P(A)$  的三种主要方法。

### 定义 1.3 统计概率

在观察某一随机事件  $A$  的随机试验中，随着随机试验次数  $n$  的增大，事件  $A$  发生的频率  $p(A)$  会越来越稳定地在某一常数  $p$  附近摆动，这时就以常数  $p$  作为事件  $A$  的概率，称之为统计概率，即  $P(A) = p$ .



具有以下特征的随机试验称为古典概型：

- (1) 每次试验的样本空间  $\Omega$  只包含有限个基本事件，记为如  $\omega_1, \omega_2, \dots, \omega_n$ ;
- (2) 各个基本事件出现的可能性相同，即基本事件的出现具有等可能性。

**定义 1.4 古典概率**

对古典概型中的任一随机事件  $A$ , 以  $p(A) = \frac{A \text{ 中的基本事件数}}{\Omega \text{ 中的基本事件数}}$  作为事件  $A$  的概率, 称为古典概率。



具有以下特征的随机试验称为几何概型:

(1) 随机试验可归结为在一个可度量的几何图形  $\Omega$  中随机投点(或取点), 以  $m(\Omega)$  表示  $\Omega$  的度量(如长度、面积、体积等), 而事件  $A$  是指所投点(取点)落在(取自)  $\Omega$  中的可度量图形  $A$  中;

(2) 事件  $A$  的概率与  $A$  的度量  $m(\Omega)$  成正比, 而与  $A$  在  $\Omega$  中的位置无关。

**定义 1.5 几何概率**

对几何概型中的任一随机事件  $A$ , 以

$$p(A) = \frac{m(A)}{m(\Omega)} \quad (1.1)$$

作为事件  $A$  的概率, 称为几何概率。



需要指出的是, 随着概率论这门学科研究的深入和发展, 产生了对随机事件概率高度科学概括的公理化定义。

**定义 1.6 古典概率**

设随机试验  $E$  的样本空间为  $\Omega$ , 对于随机试验  $E$  的每一随机事件  $A$ , 都赋予唯一确定的实数  $P(A)$ , 其中满足下列条件的集合函数  $P(\cdot)$  称为事件  $A$  的概率:

- (1) 非负性: 对每一个事件  $A \subset \Omega$ , 都有  $P(A) \geq 0$ ;
- (2) 规范性:  $P(\Omega) = 1$ ;
- (3) 可列可加性: 对任意互不相容的事件  $A_1, A_2, \dots$ , 有  $P\left(\sum_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ .

**定义 1.7 概率空间**

在概率论的公理化结构中, 称三元组  $(\Omega, F, P)$  为概率空间, 其中  $\Omega$  为样本空间,  $F$  为事件域(事件的全体),  $P$  为概率。



在统计信号处理中, 我们还经常用到条件概率。

### 定义 1.8 古典概率

设  $A$  和  $B$  为任意两个随机事件，且  $P(B) > 0$ ，称

$$P(A|B) = \frac{P(AB)}{P(B)}. \quad (1.2)$$

为事件  $B$  发生条件下事件  $A$  发生的条件概率，也称  $A$  对  $B$  的条件概率。

由条件概率定义 (1.2) 式可知

$$P(AB) = P(A|B)P(B).$$

类似地

$$P(AB) = P(A)P(B|A), P(A) > 0.$$

以上两式可以称为概率的乘法公式。乘法公式还可以推广到任意有限事件的情况：

$$\begin{aligned} P\left(\prod_{i=1}^n A_i\right) &= P(A_1 A_2 \cdots A_{n-1}) P(A_n | A_1 A_2 \cdots A_{n-1}) \\ &= P(A_1) P(A_2 | A_1) P(A_3 | A_1 A_2) \cdots P(A_n | A_1 A_2 \cdots A_{n-1}). \end{aligned} \quad (1.3)$$

利用条件概率定义、乘法公式以及概率的可加性，可以推导出两个十分有用的公式：全概率公式和贝叶斯公式。

### 定义 1.9 全概率公式

设某随机试验的样本空间  $\Omega$  中的事件  $A_1, A_2, \dots$  (有限个或可列个)  $A_i$  构成一个完备事件组，且  $P(A_i) > 0 (i = 1, 2, \dots)$ ，则对任一事件  $B$ ，有

$$P(B) = \sum P(A_i) P(B|A_i) \quad (1.4)$$

称为全概率公式。

### 定义 1.10 贝叶斯公式

设某随机试验的样本空间  $\Omega$  中的事件  $A_1, A_2, \dots$  (有限个或可列个)  $A_i$  构成一个完备事件组，且  $P(A_i) > 0 (i = 1, 2, \dots)$ ，则对任一事件  $A \subset \Omega, B \subset \Omega, P(B) > 0$ ，有如下的条件概率：

$$P(A_m|B) = \frac{P(A_m) P(B|A_m)}{\sum P(A_i) P(B|A_i)}, \quad m = 1, 2, \dots \quad (1.5)$$

称为贝叶斯公式。



最后，我们将简要介绍事件和试验的独立性。

### 定义 1.11 事件独立

如果随机事件  $A$  与  $B$  满足如下关系

$$P(A|B) = p(B), \quad (1.6)$$

则称事件  $A$  与  $B$  是相互独立的，简称  $A$  与  $B$  独立。



独立事件具有如下的重要性质：若事件  $A_1, A_2, \dots, A_n$  相互独立，则有

$$P(AB) = P(A)P(B).$$

### 定义 1.12 试验的独立定义

一个试验重复进行  $n$  次，如果在每次试验中，任意事件出现的概率与其他各次试验结果无关，则称这  $n$  次试验是独立的，或称这  $n$  次试验是  $n$  次重复独立试验。



## 1.1.2 随机变量及其分布

在本部分，我们将引入随机变量的概念来表达随机事件。

### 1) 随机变量

#### 定义 1.13 事件独立

某随机试验的样本空间为  $\Omega = \{\omega\}$ ，对于每一个样本点  $\omega \in \Omega$ ，都有唯一的实数  $X(\omega)$  与之对应，这样就得到一个定义在  $\Omega$  上的单值实函数  $X = X(\omega)$ ，如果对任意实数  $x \in \mathbb{R}$ ，“ $X(\omega) \leq x$ ”都是一个随机事件，并有确定的概率，则称  $X = X(\omega)$  为随机变量。



随机变量常用大写拉丁字母  $X, Y, Z$  或希腊字母  $\xi, \eta, \zeta$  等表示，随机变量的取值常用  $x, y, z, a, b, c$  等表示。

由定义可见，随机变量的每一个取值都对应着随机试验样本空间  $\Omega$  中的一个样本点  $e$ 。这样，引入随机变量之后，就将随机试验的结果数量化，从而把对随机试验  $E$  以及其中随机事件  $A$  的研究转化为对随机变量以及其取值的研究。

随机变量可分为两类：离散型随机变量和连续型随机变量。

### 定义 1.14 分布律

如果随机变量  $X$  的全部可能取值为有限个或可列个，则称  $X$  为离散型随机变量。设离散型随机变量  $X$  所有可能的取值为  $x_i (i = 1, 2, \dots)$ ,  $X$  取各个可能值的概率，即事件  $X = x_i$  的概率为

$$P(X = x_i) = p_i, i = 1, 2, \dots \quad (1.7)$$

我们称 (1.7) 式为离散型随机变量  $X$  的概率分布或分布律。

显然分布律也可以用表格的形式表示。

#### 2) 分布函数

### 定义 1.15 分布函数

设  $X$  是一个随机变量，对任意的实数  $x (-\infty < x < +\infty)$ ，令

$$F(x) = P(X \leq x), \quad (1.8)$$

则称  $F(x)$  为随机变量  $X$  的概率分布函数，简称为分布函数。它的定义域是  $(-\infty, +\infty)$ ，值域是  $[0, 1]$ 。如果对于随机变量  $X$  的分布函数  $F(x)$ ，存在一个非负可积函数  $f(x) (-\infty < x < +\infty)$ ，有

$$F(x) = \int_{-\infty}^x f(t) dt, \quad (1.9)$$

则称  $X$  为连续型随机变量，函数  $f(x)$  为其概率密度函数。

### 定义 1.16 0-1 分布

假设随机试验  $E$  只有两个可能结果  $A$  与  $\bar{A}$  时，随机变量

$$X = \begin{cases} 0 & \bar{A} \text{ 出现} \\ 1 & A \text{ 出现} \end{cases} \quad (1.10)$$

$X$  表示在试验中事件  $A$  出现的次数，并设  $P(A) = p (0 < p < 1)$ ，则  $X$  的概率分布为

$$P(X = k) = p^k (1 - p)^{1-k}, k = 0, 1 \quad (1.11)$$

这时称  $X$  服从参数为  $p$  的 0-1 分布，记为  $X \sim B(1, p)$ 。

**定义 1.17 泊松分布**

如果随机变量  $X$  的分布为

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad (1.12)$$

其中  $\lambda > 0$  为常数，则称  $X$  服从参数为  $\lambda$  的泊松分布，记为  $X \sim P(\lambda)$ .

对于连续型随机变量，常见的分布有均匀分布和正态分布等。

**定义 1.18 均匀分布**

如果连续型随机变量  $X$  的概率密度函数为

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{其他} \end{cases} \quad (1.13)$$

其中  $a$  和  $b$  为常数且  $a < b$ ，则称  $X$  在区间  $[a, b]$  上服从均匀分布，记为  $X \sim U[a, b]$ .

**定义 1.19 高斯分布**

如果连续型随机变量  $X$  的概率密度函数为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left[-\frac{(x - m_X)^2}{2\sigma_X^2}\right], \quad -\infty < x < +\infty \quad (1.14)$$

其中  $m_X$  和  $\sigma_X$  为常数，且  $\sigma_X > 0$ ，则称  $X$  服从参数为  $m_X$  和  $\sigma_X$  的高斯分布，记为  $X \sim N(m_X, \sigma_X^2)$ .

除此之外，还有如下几种分布：

**定义 1.20 瑞利分布**

如果连续型随机变量  $X$  的概率密度函数为

$$f(x) = \begin{cases} \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right), & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (1.15)$$

其中  $\sigma$  为常数，且  $\sigma > 0$ ，则称  $X$  服从参数为  $\sigma$  的瑞利分布， $X \sim \text{Rayleigh}(\sigma)$ 。瑞利分布的密度函数如图 1-2.

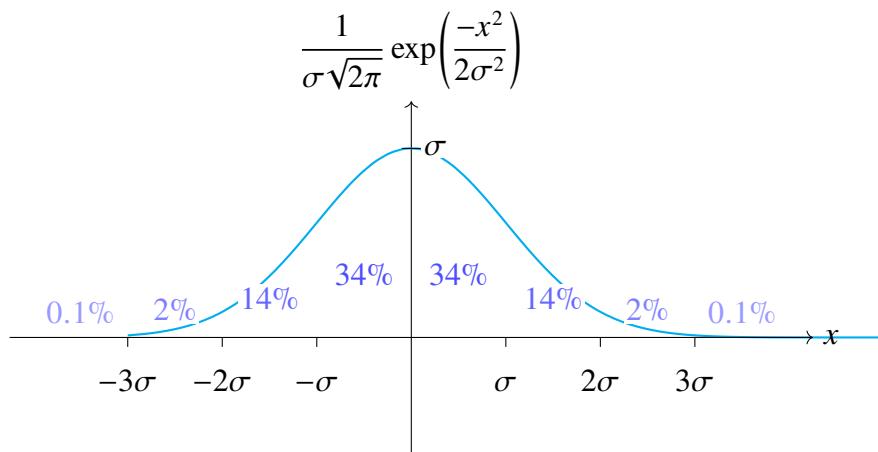


图 1-1 正态分布的分位点

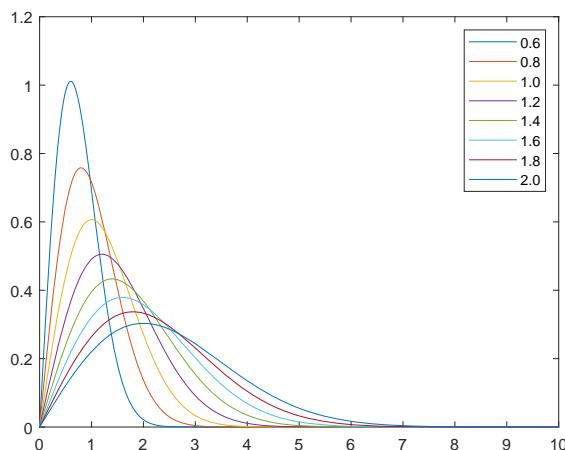


图 1-2 瑞利分布的密度函数

### 定义 1.21 卡方分布

如果连续型随机变量  $X$  的概率密度函数为

$$f(x) = \begin{cases} \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} \exp(-\frac{x}{2}), & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (1.16)$$

其中  $n$  为正整数，则称  $X$  服从自由度为  $n$  的  $\chi^2$  分布， $X \sim \chi^2(n)$ 。



 **注** 若  $X_i \sim e(\lambda)$ ,  $i = 1, 2, \dots, n$ , 且  $X_i$  独立, 则  $\sum_{i=1}^n e(\lambda) \sim \Gamma(n, \lambda)$ , 即  $n$  个独立同分布的指数变量之和为伽玛变量。常用  $\chi^2(n) = \Gamma(\frac{n}{2}, \frac{1}{2})$  来定义卡方分布。

### 定义 1.22 莱斯分布

如果连续型随机变量  $X$  的概率密度函数为

$$f(x) = \begin{cases} \frac{x}{\sigma^2} \exp\left[-\frac{(x^2+v^2)}{2\sigma^2}\right] I_0\left(\frac{xy}{\sigma^2}\right), & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (1.17)$$

其中  $\sigma$  为常数且  $\sigma > 0$ ,  $I_0(z)$  是零阶第一类贝塞尔 (Bessel) 函数, 则称  $X$  服从莱斯分布。

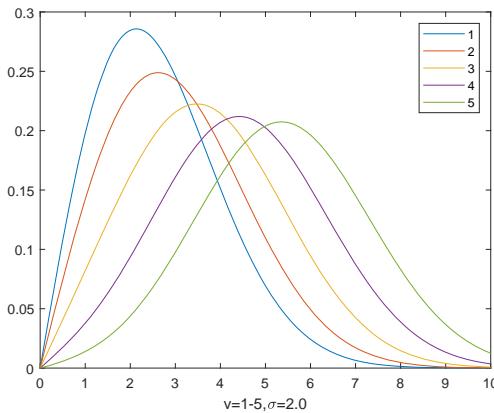


图 1-3 不同  $v$  下的莱斯分布的密度函数  
联合分布函数  $F_{XY}(x, y)$  具有以下基本性质:

- 1°  $F_{XY}(x, y)$  分别对  $x, y$  单调不减。
- 2°  $F_{XY}(x, y)$  对每个变量, 均为右连续。
- 3°  $0 \leq F_{XY}(x, y) \leq 1$  且  $F_{XY}(x, -\infty) = 0, F_{XY}(-\infty, y) = 0$  和  $F_{XY}(+\infty, +\infty) = 1$ .
- 4° 若任意四个实数  $a_1, a_2, b_1, b_2$ , 满足  $a_1 \leq a_2, b_1 \leq b_2$ , 则

$$\begin{aligned} P\{a_1 < X \leq a_2, b_1 < Y \leq b_2\} &= F_{XY}(a_2, b_2) + F_{MY}(a_1, b_1) - F_{XY}(a_1, b_2) \\ &\quad - F_{XY}(a_2, b_1). \end{aligned} \quad (1.18)$$

如图 1-5 所示。

3) 概率密度的定义

### 定义 1.23 联合概率密度

若  $F_{XY}(x, y)$  存在二阶偏导数, 则称

$$f_{XY}(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial x \partial y} \quad (1.19)$$

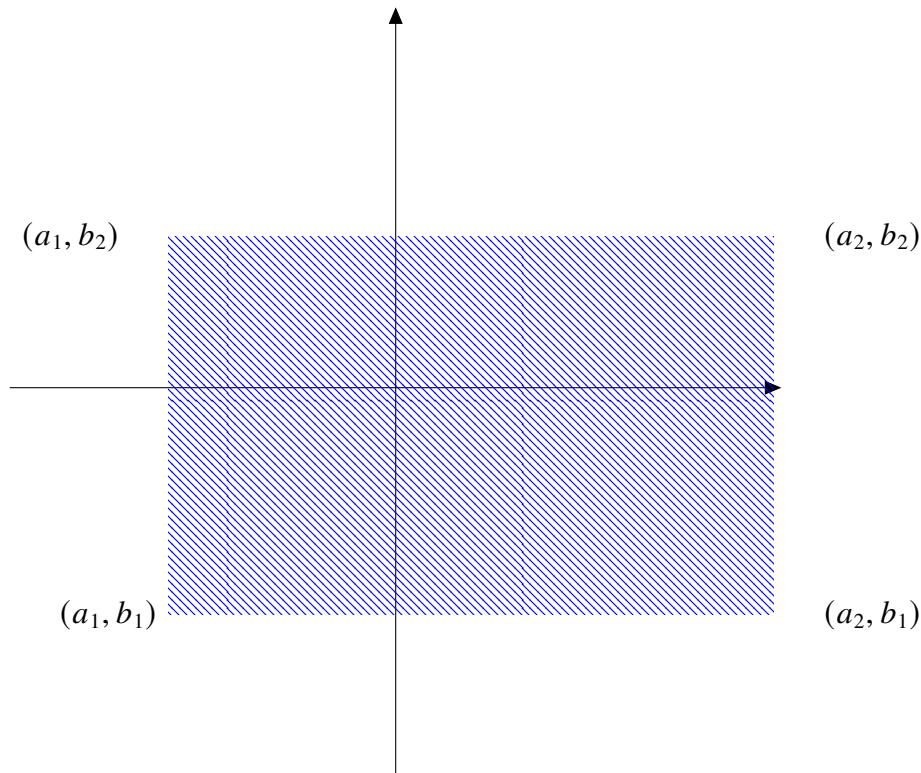


图 1-5 联合分布函数

为二维随机变量  $(X, Y)$  的联合概率密度。



联合概率密度具有以下基本性质:

$$1^\circ \quad f_{XY}(x, y) \geq 0.$$

$$2^\circ \quad \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{XY}(x, y) dx dy = 1.$$

$$3^\circ \quad \int_{-\infty}^y \int_{-\infty}^x f_{XY}(u, v) du dv = F_{XY}(x, y).$$

$$4^\circ \quad P\{(x, y) \in D\} = \iint f_{XY}(u, v) du dv.$$

在几何上,  $P(x, y) \in D$  表示曲面  $f_x(x, y)$  与  $D$  所围的柱体体积, 如图 1-6 所示。

**例 1.24** 设二维随机变量  $(X, Y)$  的概率密度

$$f(x, y) = \begin{cases} e^{-x-y}, & 0 < x < +\infty, 0 < y < +\infty \\ 0, & \text{其他} \end{cases} \quad (1.20)$$

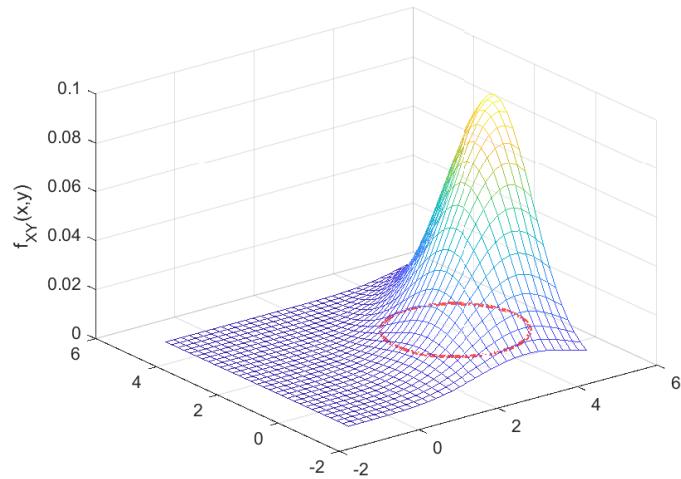
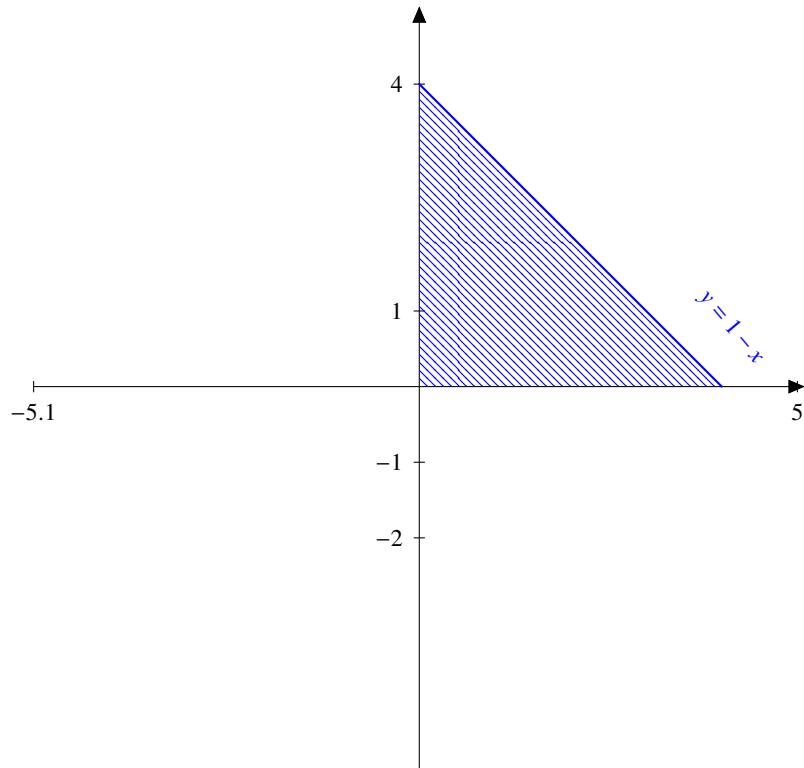


图 1-6 矩形区域上的联合分布函数值

图 1-7 三角积分区域  $G : x > 0, y > 0, x + y < 1$

求: ① 求分布函数  $F_{XY}(x, y)$ 。②  $(X, Y)$  落在如图 1-7 所示的三角形域  $G$  内的概率。

解: ① 分布函数

$$\begin{aligned} F_{XY}(x, y) &= \int_{-\infty}^y \int_{-\infty}^x f(u, v) du dv \\ &= \begin{cases} \int_0^y \int_0^x f(u, v) du dv, & 0 < x < +\infty, 0 < y < +\infty \\ 0, & \text{其他} \end{cases} \\ &= \begin{cases} (1 - e^{-\alpha})(1 - e^{-y}), & 0 < x < +\infty, 0 < y < +\infty \\ 0, & \text{其他} \end{cases} \end{aligned} \quad (1.21)$$

②  $(X, Y)$  落在三角形区域  $G$  内的概率

$$\begin{aligned} P\{(x, y) \in G\} &= \iint_G f(x, y) dx dy = \int_0^1 \int_0^{1-y} e^{-(x+y)} dx dy \\ &= \int_0^1 e^{-y} \left[ \int_0^{1-y} e^{-x} dx \right] dy = \int_0^1 e^{-y} \cdot (1 - e^{-1+y}) dy \\ &= \int_0^1 (e^{-y} - e^{-1}) dy = 1 - 2e^{-1} = 0.2642. \end{aligned} \quad (1.22)$$

#### (4) 离散型二维随机变量

若二维随机变量  $(X, Y)$  的所有可能取值是有限对或可列无限多对, 则称  $(X, Y)$  的所有可能取值为  $(x_i, y_j)$  ( $i, j = 1, 2, \dots$ ), 其上的概率记为

$$P\{X = x_i, Y = y_j\} = p_{ij}, i, j = 1, 2, \dots \quad (1.23)$$

根据概率的性质, 有

$$1^\circ \quad p_{ij} \geq 0,$$

$$2^\circ \quad \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} p_{ij} = 1,$$

则称  $p_{ij}$  为二维离散型随机变量  $(X, Y)$  的分布律, 或称为随机变量  $X$  和  $Y$  的联合分布律。

利用阶跃函数  $U(x)$  与冲激函数  $\delta(x)$ , 离散型二维随机变量的联合分布函数可表示为

$$\begin{aligned} F_{XY}(x, y) &= P\{X \leq x, Y \leq y\} \\ &= \sum_i \sum_j P\{X = x_i, Y = y_j\} U(x - x_i) U(y - y_j) \\ &= \sum_i \sum_j p_{ij} U(x - x_i) U(y - y_j). \end{aligned} \quad (1.24)$$

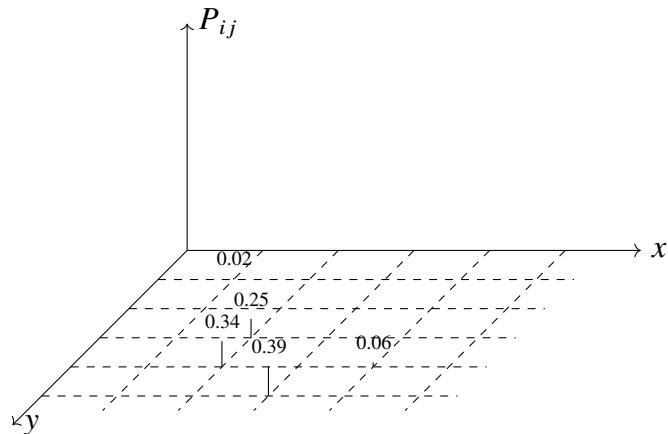


图 1-8 离散型二维随机变量的分布律

离散型二维随机变量的联合概率密度可表示为

$$\begin{aligned} P_{ij} &= \sum_i \sum_j P \{X = x_i, Y = y_j\} \delta(x - x_i) \delta(y - y_j) \\ &= \sum_i \sum_j p_{ij} \delta(x - x_i) \delta(y - y_j). \end{aligned} \quad (1.25)$$

## 2. 二维随机变量的边缘分布和条件分布

### (1) 边缘分布函数和边缘概率密度

二维随机变量  $(X, Y)$  作为一个整体, 它具有联合分布函数  $F_{XY}(x, y)$ ; 而  $X$  和  $Y$  也都是随机变量, 即分布函数为  $F_X(x)$  和  $F_Y(y)$ , 它们与联合分布函数  $F_{XY}(x, y)$  具有如下关系:

$$F_X(x) = F_{XY}(x, \infty), \quad F_Y(y) = F_{XY}(\infty, y), \quad (1.26)$$

则称  $F_X(x)$  和  $F_Y(y)$  分别为  $(X, Y)$  关于  $X$  和  $Y$  的边缘分布函数, 简称  $X$  和  $Y$  的边缘分布函数。

### 例 1.1.1

如 1-8 所示, 试着解释哪张图可以推出联系分布律?

- ① 对于连续型随机变量  $(X, Y)$ , 有  $(X, Y)$  对  $X$  的的分布函数

$$F_X(x) = F_{XY}(x, \infty) = \int_{-\infty}^x \int_{-\infty}^{+\infty} f_{XY}(u, y) dy du. \quad (1.27)$$

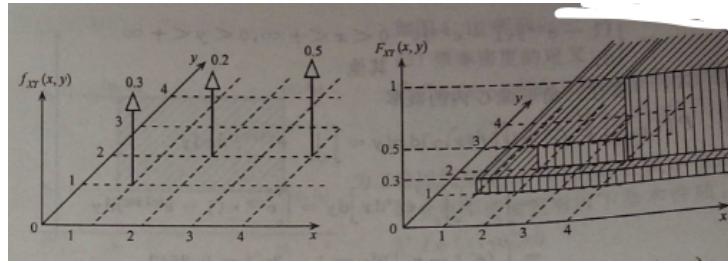


图 1-9 离散型二维随机变量的概率密度  $f_X(x)$  和分布函数  $F_Y(y)$

对  $F_X(x) = F_{XY}(x, \infty) = \int_{-\infty}^x \int_{-\infty}^{+\infty} f_{XY}(u, y) dy du$  求导, 得

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x, y) dy, \quad (1.28)$$

称之为  $(X, Y)$  关于  $X$  的边缘概率密度。

① 同理,  $(X, Y)$  关于  $Y$  的边缘概率密度为

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) dx. \quad (1.29)$$

对于离散型随机变量  $(X, Y)$ , 有  $(X, Y)$  对  $X$  的分布函数

$$F_X(x) = F_{XY}(x, \infty) = \sum_i \sum_{j=1}^{\infty} p_{ij} U(x - x_i) = \sum_{x_i \leq x} \sum_{j=1}^{\infty} p_{ij}. \quad (1.30)$$

可得

$$p_{i \cdot} = P\{X = x_i\} = \sum_{j=1}^{\infty} p_{ij}, i = 1, 2, \dots \quad (1.31)$$

称之为  $(X, Y)$  关于  $X$  的边缘分布律。同理,  $(X, Y)$  关于  $Y$  的边缘分布律为

$$p_{\cdot j} = P\{Y = y_j\} = \sum_{i=1}^{\infty} p_{ij}, j = 1, 2, \dots \quad (1.32)$$

边缘分布函数、边缘概率密度和边缘分布律, 反映了二维随机变量中各随机变量本身的统计特征。

**例 1.25** (续上例) ① 求边缘分布函数  $F_X(x)$  和  $F_Y(y)$ 。② 求边缘概率密度。

解: ① 已知联合分布函数  $F_{XY}(x, y)$  的表达式, 则边缘分布函数

$$\begin{aligned} F_X(x) &= F_X(x, \infty) = \begin{cases} (1 - e^{-x})(1 - e^{-\infty}), & 0 < x < +\infty \\ 0, & \text{其他} \end{cases} \\ &= \begin{cases} 1 - e^{-x}, & 0 < x < +\infty \\ 0, & \text{其他} \end{cases} \end{aligned} \quad (1.33)$$

同理

$$F_Y(y) = F_{XY}(\infty, y) = \begin{cases} 1 - e^{-y}, & 0 < y < +\infty \\ 0, & \text{其他} \end{cases} \quad (1.34)$$

② 边缘概率密度的计算

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{+\infty} f_{XY}(x, y) dy = \begin{cases} \int_0^{+\infty} e^{-(x+y)} dy, & 0 < x < +\infty \\ 0, & \text{其他} \end{cases} \\ &= \begin{cases} e^{-x}, & 0 < x < +\infty \\ 0, & \text{其他} \end{cases} \end{aligned} \quad (1.35)$$

同理

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x, y) dx = \begin{cases} e^{-y}, & 0 < y < +\infty \\ 0, & \text{其他} \end{cases}. \quad (1.36)$$

1) 条件概率的定义

### 定义 1.26 条件分布函数和条件概率密度

1 对于连续型二维随机变量  $(X, Y)$  有

$$\begin{aligned} F_Y(y|X=x) &= P\{Y \leq y | x \leq X \leq x + \Delta\} = P(A|B) = \int_{-\infty}^y \frac{f_{XY}(x, v)}{f_X(x)} dv, \\ f_Y(y|X=x) &= \frac{f_{XY}(x, y)}{f_X(x)}; \end{aligned} \quad (1.37)$$

分别称为给定  $X = x$  的条件下  $Y$  的条件分布函数和条件概率密度，可以简写为  $F_Y(y|x), f_Y(y|x)$ .



2) 推导过程

前面引入了条件概率的概念，即在给定事件  $B$  的条件下，事件  $A$  发生的条件概率。

$$P(A|B) = \frac{P(AB)}{P(B)}. \quad (1.38)$$

把这个概念引入到随机变量的理论中。

1) 先推导  $X \leq x$  条件下的分布函数和分布律：

对于连续型二维随机变量  $(X, Y)$ ，若令  $A = \{Y \leq y\}$ ,  $B = \{X \leq x\}$ ，则称  $P(A|B) = P\{Y \leq y|B\} = F_Y(y|B)$  为给定条件  $B$  下  $Y$  的分布函数。

根据上述讨论，可得条件分布函数  $F_Y(y|B)$ ，联合分布函数  $F_{XY}(x, y)$  及边缘分布函数  $F_X(x)$  三者之间的关系为

$$F_Y(y|X \leq x) = P(A|B) = \frac{P(AB)}{P(B)} = \frac{P\{X \leq x, Y \leq y\}}{P\{X \leq x\}} = \frac{F_{XY}(x, y)}{F_X(x)}. \quad (1.39)$$

若上式对  $y$  的导数存在, 则有

$$f_Y(y|X \leq x) = \frac{\partial F_Y(y|X \leq x)}{\partial y} = \frac{\partial F_{XY}(x, y)/\partial y}{F_X(x)} = \frac{\int_{-\infty}^x f_{XY}(u, y)du}{\int_{-\infty}^x f_X(u)du}. \quad (1.40)$$

2) 接着推导  $X = x$  条件下的分布函数:

若令  $B = \{X = x\}$ , 代入式 (1.39), 得

$$\begin{aligned} F_Y(y|X = x) &= \lim_{\Delta x \rightarrow 0} F_Y(y|x < X \leq x + \Delta x) = \lim_{\Delta x \rightarrow 0} \frac{P\{x < X \leq x + \Delta x, Y \leq y\}}{P\{x < X \leq x + \Delta x\}} \\ &= \lim_{\Delta x \rightarrow 0} \frac{P\{X \leq x + \Delta x, Y \leq y\} - P\{X \leq x, Y \leq y\}}{P\{X \leq x + \Delta x\} - P\{X \leq x\}} \\ &= \lim_{\Delta x \rightarrow 0} \frac{[F_{XY}(x + \Delta x, y) - F_{XY}(x, y)]}{[F_X(x + \Delta x) - F_X(x)]} \\ &= \lim_{\Delta x \rightarrow 0} \frac{[F_{XY}(x + \Delta x, y) - F_{XY}(x, y)]/\Delta x}{[F_X(x + \Delta x) - F_X(x)]/\Delta x} = \frac{\partial F_{XY}(x, y)/\partial x}{\partial F_X(x)/\partial x} \\ &= \frac{\int_{-\infty}^y f_{XY}(x, y)dy}{f_X(x)}. \end{aligned} \quad (1.41)$$

得

$$f_Y(y|X = x) = \frac{\partial}{\partial y} F_Y(y|X = x) = \frac{f_{XY}(x, y)}{f_X(x)}, f_X(x) \neq 0. \quad (1.42)$$

**例 1.27** (续上例) ⑤ 求条件分布函数  $F_X(x|y)$  和  $F_Y(y|x)$ 。⑥ 求条件概率密度  $f_X(x|y)$ .

解: ⑤ 条件分布函数

$$\begin{aligned} F_X(x|y) &= \int_{-\infty}^x \frac{f_{XY}(x, y)dx}{f_Y(y)} dy = \begin{cases} \frac{\int_0^x e^{-(x+y)} dx}{\int_0^y e^{-y} dy}, & 0 < x < +\infty \\ 0, & \text{其他} \end{cases} \\ &= \begin{cases} 1 - e^{-x}, & 0 < x < +\infty \\ 0, & \text{其他} \end{cases}. \end{aligned} \quad (1.43)$$

同理

$$F_Y(y|x) = \begin{cases} 1 - e^{-y}, & 0 < y < +\infty \\ 0, & \text{其他} \end{cases}. \quad (1.44)$$

⑥ 条件概率密度

$$f_X(x|y) = \frac{f_X(x, y)}{f_Y(y)} = \begin{cases} \frac{e^{-(x+y)}}{e^{-y}}, & 0 < x < +\infty \\ 0, & \text{其他} \end{cases} = \begin{cases} e^{-x}, & 0 < x < +\infty \\ 0, & \text{其他} \end{cases} \quad (1.45)$$

3) 条件分布函数  $F_Y(y|B)$  是求在  $B$  发生的条件下, 事件  $\{Y(\zeta) \leq y\}$  发生的概率,  $\zeta \in B$ . 换句话说, 它是求在新的样本空间上事件  $\{Y(\zeta) \leq y\}$  发生的概率. 而无条件的分

布函数  $F_Y(y)$ , 则是在  $\Omega$  上求  $\{Y(\zeta) \leq y\}$  事件发生的概率  $\zeta \in \Omega$ 。因此, 除了样本空间缩小成  $\Omega_B$  以外, 条件分布函数的性质与一般分布函数的性质完全相同。件分布函数的性质如下:

条件分布函数性质如下:

**性质** 1°  $F_Y(\infty|B) = 1, F_Y(-\infty|B) = 0, 0 \leq F_Y(y|B) \leq 1.$

2°  $F_Y(y_2|B) - F_Y(y_1|B) = P\{y_1 < Y \leq y_2|B\}.$

**性质** 条件密度函数的性质如下:

1°  $f_Y(y|B) \geq 0.$

2°  $\int_{-\infty}^{\infty} f_Y(y|B) dy = F_Y(\infty|B) - F_Y(-\infty|B) = 1.$

3°  $F_Y(y|B) = \int_y^\infty f_Y(y|B) dy.$

#### 4) 离散型随机变量的条件分布律

##### 定义 1.28 条件分布律

1 对于离散型随机变量  $X$  和  $Y$ , 其在  $X = x$  的条件下,  $Y = y$  的条件概率可直接定义为

$$P\{Y = y|X = x\} = \frac{P\{X = x, Y = y\}}{P\{X = x\}}, \quad P(X = x) > 0. \quad (1.46)$$

① 二维随机变量  $(X, Y)$ , 对于固定的  $j$ , 若  $P\{Y = y_j\} > 0$ , 则

$$P\{X = x_i|Y = y_j\} = \frac{P\{X = x_i, Y = y_j\}}{P\{Y = y_j\}} = \frac{p_{ij}}{p_{\cdot j}}, \quad i = 1, 2, \dots, \quad (1.47)$$

称为在  $Y = y$  条件下随机变量  $X$  的条件分布律。

② 对于固定的  $i$ , 若  $P\{X = x\} > 0$ , 则

$$P\{Y = y_j|X = x_i\} = \frac{P\{X = x_i, Y = y_j\}}{P\{X = x_i\}} = \frac{p_{ij}}{p_{ii}}, \quad j = 1, 2, \dots, \quad (1.48)$$

称为在  $X = x_i$  条件下随机变量  $Y$  的条件分布律。

#### 3. 随机变量的统计独立

现在把事件独立的概念引入到随机变量中来。

##### 定义 1.29 相互独立

$X, Y$  是两个随机变量, 若对任意实数  $x$  和  $y$ , 有

$$P\{X < x, Y < y\} = P\{(X < x) \cap (Y < y)\} = P\{X < x\} \cdot P\{Y < y\}, \quad (1.49)$$

则称随机变量  $X, Y$  相互独立。

二维随机变量  $(X, Y)$ ,  $X$  与  $Y$  相互独立的条件为

$$F_{XY}(x, y) = F_X(x) \cdot F_Y(y). \quad (1.50)$$

或

$$f_{XY}(x, y) = f_X(x) \cdot f_Y(y). \quad (1.51)$$

把上两式代入条件分布函数和概率密度的定义 (式 (1-73)), 可得

$$\begin{cases} F_Y(y|x) = F_Y(y) \\ f_Y(y|x) = f_Y(y) \end{cases}. \quad (1.52)$$

同理可得

$$\begin{cases} F_X(x|y) = F_X(x) \\ f_X(x|y) = f_X(x) \end{cases}. \quad (1.53)$$

说明: 当  $X$  和  $Y$  相互独立时,  $X$  在  $Y = y$  的条件下的分布与  $X$  的无条件分布相同, 或者  $Y$  在  $X = x$  的条件下的分布与  $Y$  的无条件分布相同。也就是说, 随机变量  $X$  的统计特征与随机变量  $Y$  的统计特征无关。

离散型随机变量  $X$  和  $Y$  独立的条件: 对所有  $i, j$ , 均有  $p_{ij} = p_i \cdot p_j$  ( $i, j = 1, 2, \dots$ ), 即

$$P\{X = x_i, Y = y_j\} = P\{X = x_i\} \cdot P\{Y = y_j\}. \quad (1.54)$$

**例 1.30** (续上例) ⑦  $X$  和  $Y$  是否统计独立?

解: ⑦ 独立。因为存在下面的关系 (只需满足  $a, b, c, d$  任一条件即可推出统计独立)。

a. 由已知条件和求得的结论可知

$$F_{XY}(x, y) = \begin{cases} (1 - e^{-x})(1 - e^{-y}), & 0 < x < +\infty, 0 < y < +\infty \\ 0, & \text{其他} \end{cases}. \quad (1.55)$$

$$F_X(x) = \begin{cases} 1 - e^{-x}, & 0 < x < +\infty, \\ 0, & \text{其他} \end{cases}, \quad F_Y(y) = \begin{cases} 1 - e^{-y}, & 0 < y < +\infty \\ 0, & \text{其他} \end{cases} \quad (1.56)$$

所以  $F_{XY}(x, y) = F_X(x) \cdot F_Y(y)$  成立, 则  $X$  和  $Y$  统计独立。

b. 由已知条件和求得的结论可知

$$F_X(x|y) = \begin{cases} 1 - e^{-x}, & 0 < x < +\infty \\ 0, & \text{其他} \end{cases}, \quad F_X(x) = \begin{cases} 1 - e^{-x}, & 0 < x < +\infty \\ 0, & \text{其他} \end{cases} \quad (1.57)$$

所以  $F_X(x|y) = F_X(x)$  成立, 则  $X$  和  $Y$  统计独立。

c. 由已知条件和求得的结论可知

$$f_{XY}(x, y) = \begin{cases} e^{-(x+y)}, & 0 < x < +\infty, 0 < y < +\infty \\ 0, & \text{其他} \end{cases} \quad (1.58)$$

$$f_X(x) = \begin{cases} e^{-x}, & 0 < x < +\infty \\ 0, & \text{其他} \end{cases}, f_Y(y) = \begin{cases} e^{-y}, & 0 < y < +\infty \\ 0, & \text{其他} \end{cases} \quad (1.59)$$

所以  $f_{xx}(x, y) = f_X(x) \cdot f_Y(y)$  成立, 则  $X$  和  $Y$  统计独立。

d. 由已知条件和求得的结论可知

$$f_X(x|y) = \begin{cases} e^{-x}, & 0 < x < +\infty \\ 0, & \text{其他} \end{cases}, f_X(x) = \begin{cases} e^{-x}, & 0 < x < +\infty \\ 0, & \text{其他} \end{cases} \quad (1.60)$$

所以  $f_X(x|y) = f_X(x)$  成立, 则  $X$  和  $Y$  统计独立。

### 1.1.3 多维随机变量

在实际问题中, 一些随机现象常常需要两个或两个以上的随机变量来描述, 例如高频信号的中心频率、带宽、振幅和初相位等等。为此, 我们需要引入多维随机量(也可称为随机矢量)的概念。

#### 定义 1.31 随机矢量

设某随机试验的样本空间为  $\Omega = \{w\}$ , 对于每一个样本点  $\omega \in \Omega$ ,  $X_i = X_i(\omega), i = 1, 2, \dots$  是定义在同一个样本空间  $\Omega$  上的  $n$  个随机变量, 则  $(X_1, X_2, \dots, X_n)$  称为  $n$  维随机变量, 其矢量形式  $[X_1, X_2, \dots, X_n]^T$  也称为随机矢量。

本节我们简要回忆二维随机变量的情况, 它的很多结果都可以推广到多维随机变量上。

#### 定义 1.32 联合分布

设  $(X, Y)$  是二维随机变量, 对于任意实数  $x$  和  $y$ , 二元函数

$$F(x, y) = P(X \leq x, Y \leq y) \quad (1.61)$$

称为二维随机变量  $(X, Y)$  的分布函数, 或称随机变量  $X$  和  $Y$  的联合分布函数, 简称联合分布。

除了分布函数, 我们还将简单介绍边缘分布函数的概念。二维随机变量  $(X, Y)$  作为一个整体, 具有分布函数  $F(x, y)$ . 而  $X$  和  $Y$  都是随机变量, 我们也可以对其中任何一个

随机变量单独进行研究，即求随机变量  $X$  或  $Y$  的分布，这就是二维随机变量的边缘分布。它与二维变量的分布函数具有如下关系：

$$F_X(x) = P(X \leq x, Y < +\infty) = F(x, +\infty). \quad (1.62)$$

$$F_Y(y) = P(X < +\infty, Y \leq y) = F(+\infty, y). \quad (1.63)$$

同一维随机变量类似，二维随机变量也可以分为离散型和连续型两种形式。

### 定义 1.33 二维离散型随机变量

如果二维随机变量  $(X, Y)$  的所有可能取值都是有限对或可列无限多对，并且以确定的概率取各个不同的数对，则称  $(X, Y)$  为二维离散型随机变量。

### 定义 1.34 联合概率分布

若  $(X, Y)$  是一个二维离散型随机变量，它的一切可能取值为  $(x_i, y_j)$  ( $i, j = 1, 2, \dots$ )，则

$$P(X = x_i, Y = y_j) = p_{ij}, i, j = 1, 2, \dots \quad (1.64)$$

称为  $(X, Y)$  的联合概率分布。

### 定义 1.35 联合概率密度函数

设二维随机变量  $(X, Y)$  的分布函数是  $F(x, y)$ ，如果存在非负函数  $f(x, y)$ ，使得对任意实数  $x, y$ ，有

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv, \quad (1.65)$$

则称  $(X, Y)$  是二维连续型随机变量， $f(x, y)$  为二维连续型随机变量  $(X, Y)$  的联合概率密度函数。

利用二维随机变量的联合分布函数与边缘分布函数，引入随机变量独立的概念。

### 定义 1.36 独立

设  $F(x, y)$  是二维随机变量  $(X, Y)$  的联合分布函数， $F_X(x)$  和  $F_Y(y)$  是其边缘分布函数，若对任意实数  $x$  和  $y$ ，有

$$F(x, y) = F_X(x)F_Y(y), \quad (1.66)$$

则称随机变量  $X$  和  $Y$  是统计独立的。

此外，对于两个随机变量，我们也可以讨论它们的条件分布。对二维离散随机变量，由条件概率的公式可得

$$P(X = x_i | Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)}. \quad (1.67)$$

上式被称作在  $Y = y_i$  条件下  $X = x$  的条件概率，其中  $P(Y = y_i)$  表示边缘分布函数。

由于连续型随机变量取任何单点数值的概率都是零，所以不能像离散型随机变量那样直接利用条件概率公式给出连续型随机变量的条件概率密度。因此，考虑采用极限的办法来解决：设  $y$  是定值，对任  $\Delta y > 0$ ,  $P(y - \Delta y < Y \leq y + \Delta y) > 0$ , 若对任意实数  $X$ , 极限  $\lim_{\Delta y \rightarrow 0} P(X \leq x | y - \Delta y < Y \leq y + \Delta y)$  存在，则称此极限为在  $Y = y$  条件下  $X$  的条件分布函数，记为  $P(X \leq x | Y = y)$ 。

### 定义 1.37 $m$ 维边缘分布函数

$n$  维随机变量中的任意  $m$  ( $m < n$ ) 个分量的联合分布函数，都称为  $n$  维随机变量的  $m$  维边缘分布函数。



由  $n$  维随机变量的联合分布函数  $F_X(x_1, x_2, \dots, x_n)$ , 可以得到它任意  $m$  个分量的边缘分布函数。

### 例 1.1.2

如

$$\begin{aligned} F_X(x_1, x_2, \dots, x_m) &= F_X(x_1, x_2, \dots, x_m, \infty, \dots, \infty), \\ F_X(x_i) &= F_X(\infty, \dots, \infty, x_i, \infty, \dots, \infty). \end{aligned} \quad (1.68)$$



$n$  维随机变量中的任意  $m$  ( $m < n$ ) 个分量的概率密度，都称为  $n$  维随机变量的  $m$  维边缘概率密度。

由  $n$  维随机变量的联合概率密度  $f_X(x_1, x_2, \dots, x_n)$ , 可以得到它任意  $m$  个分量的边缘概率密度。

**例 1.1.3**

如

$$f_X(x_1, x_2, \dots, x_m) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X(x_1, \dots, x_m, x_{m+1}, \dots, x_n) dx_{m+1} dx_{m+2} \cdots dx_n. \quad (1.69)$$

$$f_X(x_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n) dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_n. \quad (1.70)$$

**4. 条件概率密度****定义 1.38 m 维边缘分布函数**

$n$  维随机变量  $(X_1, X_2, \dots, X_n)$  在给定  $X_1 = x_1$  的条件下, 其余  $n - 1$  个分量  $(X_2, X_3, \dots, X_n)$  的条件概率密度为

$$f_X(x_2, \dots, x_n | x_1) = \frac{f_X(x_1, x_2, \dots, x_n)}{f_X(x_1)}. \quad (1.71)$$

**定义 1.39 m 维边缘分布函数**

也可用多个随机变量固定为条件, 如在  $X_1 = x_1$  和  $X_2 = x_2$  条件下, 随机变量  $(X_3, X_4, \dots, X_n)$  的条件概率密度为

$$f_X(x_3, \dots, x_n | x_1, x_2) = \frac{f_X(x_1, x_2, \dots, x_n)}{f_X(x_1, x_2)}. \quad (1.72)$$

在  $X_1, X_2, \dots, X_{n-1}$  固定的条件下, 随机变量  $X_n$  的条件概率密度为

$$f_X(x_n | x_1, x_2, \dots, x_{n-1}) = \frac{f_X(x_1, x_2, \dots, x_n)}{f_X(x_1, x_2, \dots, x_{n-1})}. \quad (1.73)$$

利用上述条件概率密度的定义, 可得  $n$  维随机变量联合概率密度的递推公式:

$$\begin{aligned} f_X(x_1, \dots, x_n) &= f_X(x_n | x_1, \dots, x_{n-1}) f_X(x_{n-1} | x_1, \dots, x_{n-2}) \\ &\quad \cdots f_X(x_2 | x_1) f_X(x_1). \end{aligned} \quad (1.74)$$

**证明.** 归纳证明过程如下

- ① 二维情况  $f_X(x_1, x_2) = f_X(x_2 | x_1) f_X(x_1).$
- ② 三维情况  $f_X(x_1, x_2, x_3) = f_X(x_3 | x_1, x_2) f_X(x_1, x_2).$

可得到  $(X_1, X_2, X_3)$  三维联合概率密度的递推关系为

$$f_X(x_1, x_2, x_3) = f_X(x_3|x_1, x_2) f_X(x_2|x_1) f_X(x_1). \quad (1.75)$$

③ 由数学归纳法可推出  $n$  维随机变量联合概率密度的递推关系 ②  $n-1$  维情况  $f_X(x_1, x_2, x_{n-1}) = f_X(x_n|x_1, x_{n-1}) f_X(x_1, x_2, x_{n-1})$  成立. 当  $X_1, X_2, \dots, X_n$  相互独立时, 可进一步得到

$$\begin{aligned} f_X(x_1, \dots, x_n) &= f_X(x_n|x_1, \dots, x_n) f_X(x_1, \dots, x_n) \\ &= f_X(x_n) f_X(x_{n-1}) \cdots f_X(x_2) f_X(x_1). \end{aligned} \quad (1.76)$$

□

**例 1.40** 四维随机变量  $(X_1, X_2, X_3, X_4)$  中各随机变量相互独立, 且都服从  $(0, 1)$  上的均匀分布。求: ① 四维随机变量的联合概率密度  $f_X(x_1, x_2, x_3, x_4)$ 。② 边缘概率密度  $f_X(x_1, x_2)$ 。③ 条件概率密度  $f_X(x_3|x_1, x_2)$  和  $f_X(x_3, x_4|x_1, x_2)$ 。

解: ①  $X_i$  服从  $(0, 1)$  上的均匀分布, 则  $X_i$  的概率密度为

$$f_X(x_i) = \begin{cases} 1, & 0 < x_i < 1 \\ 0, & \text{其他} \end{cases} \quad (1.77)$$

且随机变量  $X_1, X_2, X_3, X_4$  相互独立, 则四维随机变量的联合概率密度为

$$\begin{aligned} f_X(x_1, x_2, x_3, x_4) &= f_X(x_1) \cdot f_X(x_2) \cdot f_X(x_3) \cdot f_X(x_4) \\ &= \begin{cases} 1, & 0 < x_1, x_2, x_3, x_4 < 1 \\ 0, & \text{其他} \end{cases} \end{aligned} \quad (1.78)$$

② 同理可知  $X_1, X_2$  的联合概率密度为

$$f_X(x_1, x_2) = f_X(x_1) \cdot f_X(x_2) = \begin{cases} 1, & 0 < x_1, x_2 < 1 \\ 0, & \text{其他} \end{cases} \quad (1.79)$$

③ 因为随机变量  $X_1, X_2, X_3, X_4$  相互独立, 所以条件概率密度

$$f_X(x_3|x_1, x_2) = f_X(x_3) = \begin{cases} 1, & 0 < x_3 < 1 \\ 0, & \text{其他} \end{cases} \quad (1.80)$$

$$f_X(x_3, x_4|x_1, x_2) = f_X(x_3, x_4) = f_X(x_3) \cdot f_X(x_4) = \begin{cases} 1, & 0 < x_3, x_4 < 1 \\ 0, & \text{其他} \end{cases} \quad (1.81)$$

## 1.1.4 随机变量函数的分布

### 1. 一维随机变量函数的分布

### 定理 1.41 一维随机变量

$X$  是一个连续型随机变量, 其概率密度函数为  $f(x), -\infty < X < +\infty$ ,  $g(x)$  处处可导且恒有  $g'(x) > 0$  或恒有  $g'(x) < 0$ , 则  $Y = g(X)$  也是一个连续型随机变量, 且其概率密度函数为

$$f(y) = \begin{cases} f(h(y)) |h'(y)|, & a < y < \beta \\ 0, & \text{其他} \end{cases} \quad (1.82)$$

其中,  $\alpha = \min\{g(-\infty), g(+\infty)\}, \beta = \max\{g(-\infty), g(+\infty)\}$ ,  $h(y)$  是  $g(x)$  的反函数. ♥

特别地, 若随机变量  $X \sim N(\mu, \sigma^2)$ , 令随机变量  $Y = \frac{X-\mu}{\sigma}$ , 则  $Y \sim N(0, 1)$ .

证明. 1. 当  $y \leq \alpha, F_Y(y) = P\{Y \leq y\} = P\{\emptyset\} = 0$ .

2. 当  $y \geq \beta, F_Y(y) = P\{Y \leq y\} = P\{\Omega\} = 1$ .

3. 当  $\alpha \leq y < \beta, F_Y(y) = P\{Y \leq y\} = P\{g(X) \leq y\} = P\{X \leq h(y)\} = \int_{-\infty}^{h(y)} f_X(x) dx$ .

概率密度函数

$$f_Y(y) = F'_Y(y) = \begin{cases} f_X(h(y))h'(y), & \alpha \leq y < \beta \\ 0, & \text{其他} \end{cases},$$

对于  $g'(x) < 0$  的情况,

$$f_Y(y) = F'_Y(y) = \begin{cases} f_X(h(y))(-h'(y)), & \alpha \leq y < \beta \\ 0, & \text{其他} \end{cases},$$

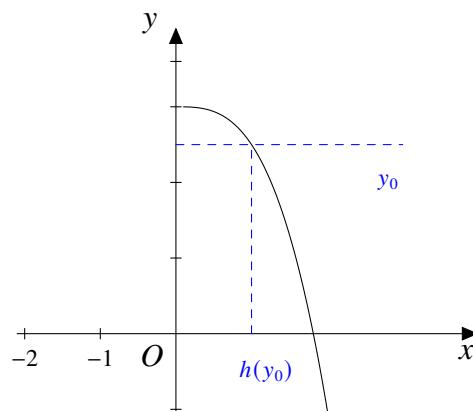


图 1-10 单调递减函数

对于  $g'(x) > 0$  的情况,

$$f_Y(y) = F'_Y(y) \begin{cases} f_X(h(y))(h'(y)), & \alpha \leq y < \beta \\ 0, & \text{其他} \end{cases},$$

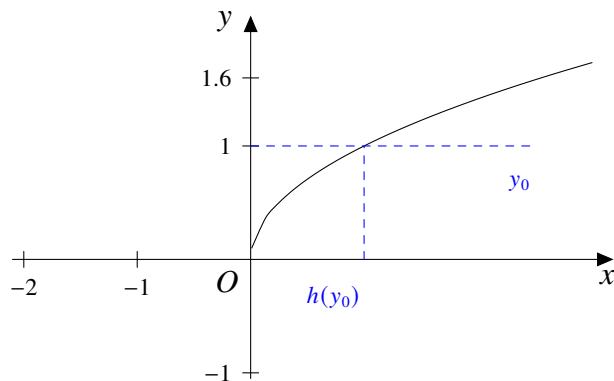


图 1-11 单调递增函数

合并后, 得密度函数

$$f_Y(y) = F'_Y(y) \begin{cases} f_X(h(y))|h'(y)|, & \alpha \leq y < \beta \\ 0, & \text{其他} \end{cases}.$$

□

#### 例 1.1.4

设  $X$  的概率密度为

$$f(x) = \begin{cases} \frac{2x}{\pi^2}, & 0 < x < \pi \\ 0, & \text{其他} \end{cases}$$

求  $Y = \sin X$  的概率密度.



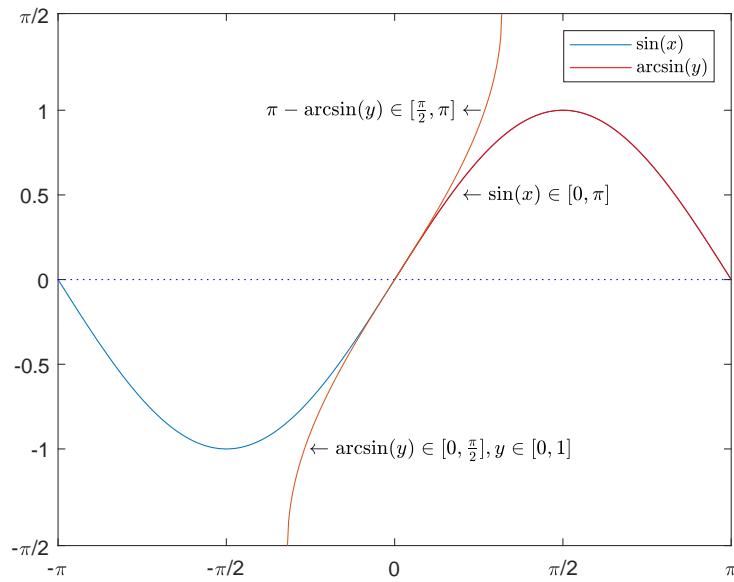


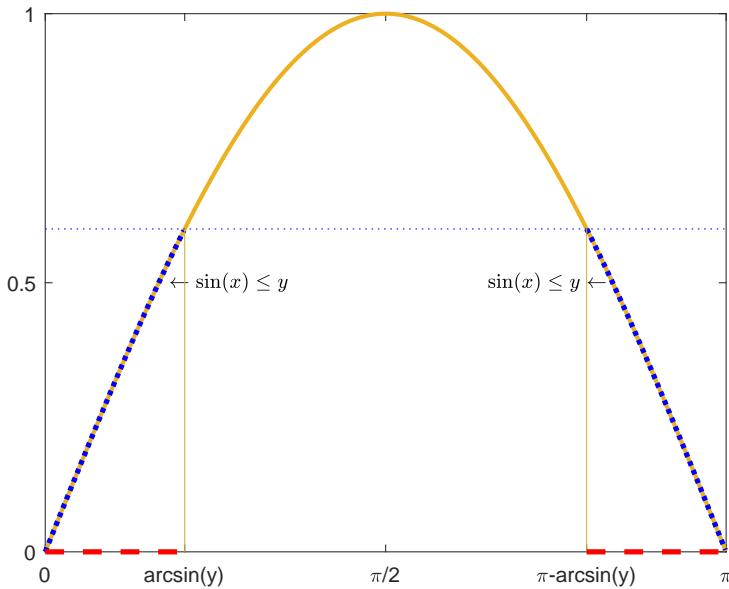
图 1-12  $\sin(x)$  及其反函数的取值区间.

解:  $F_Y(y) = P\{Y \leq y\} = P\{\sin X \leq y\}$ .

- 当  $y < 0$  时,  $F_Y(y) = 0$ .
- 当  $0 \leq y \leq 1$  时:  $F_Y(y) = P\{\sin X \leq y\} = P\{0 \leq X \leq \arcsin y\} + P\{\pi - \arcsin y \leq X \leq \pi\}$ , ( $x = \arcsin y \in [0, \frac{\pi}{2}]$ ,  $0 \leq y \leq 1$ ;  $x = \pi - \arcsin y \in [\frac{\pi}{2}, \pi]$ ,  $0 \leq y \leq 1$ ). 如图 1-12 或者 1-13).

$$F_Y(y) = \int_0^{\arcsin y} \frac{2x}{\pi^2} dx + \int_{\pi - \arcsin y}^{\pi} \frac{2x}{\pi^2} dx.$$

- 当  $1 < y$  时:  $F_Y(y) = 1$ .

图 1-13  $\sin(x) \leq y$  的取值范围.

$Y$  的概率密度  $f(y)$  为:

- $y \leq 0$  时,  $f(y) = [F_Y(y)]' = (0)' = 0$ .
- $0 < y < 1$  时,  $f(y) = [F_Y(y)]' = \left( \int_0^{\arcsin y} \frac{2x}{\pi^2} dx + \int_{\pi - \arcsin y}^{\pi} \frac{2x}{\pi^2} dx \right)'$ ,  
即

$$\begin{aligned} f(y) &= \left( \frac{x^2}{\pi^2} \Big|_0^{\arcsin y} + \frac{x^2}{\pi^2} \Big|_{\pi - \arcsin y}^{\pi} \right)' \\ &= \left( \frac{(\arcsin y)^2}{\pi^2} + 1 - \frac{(\pi - \arcsin y)^2}{\pi^2} \right)' \\ &= \left( \frac{(\arcsin y)^2}{\pi^2} + 1 - \frac{\pi^2 - 2\pi \arcsin y + (\arcsin y)^2}{\pi^2} \right)' \\ &= \left( \frac{2 \arcsin y}{\pi} \right)' = \frac{2}{\pi \sqrt{1 - y^2}}. \end{aligned}$$

- $1 \leq y$  时,  $f(y) = [F_Y(y)]' = (1)' = 0$ . 则  $Y$  的概率密度  $f(y)$  为:

$$f(y) = \begin{cases} 0, & y \geq 1 \text{ 或 } y \leq 0, \\ \frac{2}{\pi \sqrt{1 - y^2}}, & 0 < y < 1. \end{cases}$$

## 2. 多维随机变量函数的分布

### 定理 1.42 多维随机变量函数的分布

$(X_1, X_2, \dots, X_n)$  具有概率密度函数  $f(x_1, x_2, \dots, x_n)$  的连续型  $n$  维随机变量,

(1)  $y_1 = g_1(x_1, x_2, \dots, x_n), \dots, y_n = g_n(x_1, x_2, \dots, x_n)$  是  $n$  维实数空间到自身的一对一的映射, 即存在定义在该变换值域上的逆变换:

$$x_1 = h_1(y_1, y_2, \dots, y_n), \dots, x_n = h_n(y_1, y_2, \dots, y_n).$$

(2) 变换和它的逆变换都是连续的;

(3) 偏导数  $\frac{\partial h_i}{\partial y_j}$  ( $i = 1, 2, \dots, n; j = 1, 2, \dots, n$ ) 存在且连续;

(4) 逆变换的雅可比行列式

$$J(y_1, y_2, \dots, y_n) = \begin{vmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} & \cdots & \frac{\partial h_1}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_n}{\partial y_1} & \frac{\partial h_n}{\partial y_2} & \cdots & \frac{\partial h_n}{\partial y_n} \end{vmatrix} \neq 0,$$

则  $(Y_1, Y_n) = (g_1(X_1, X_2, \dots, X_n), \dots, g_n(X_1, X_2, \dots, X_n))$  的联合概率密度函数函数

$$f(y_1, y_2, \dots, y_n) = f(h_1(y_1, y_2, \dots, y_n), \dots, h_n(y_1, y_2, \dots, y_n)) |J|. \quad (1.83)$$

#### 1.1.4.1 随机变量函数的分布

上节讨论了随机变量的概念及其分布。但实际工作中, 还经常遇到求随机变量函数分布的问题。例如电子系统中, 在  $t$  时刻一个概率密度为  $f(x)$  的随机变量  $X$  通过一个非线性放大器。

$$Y = \begin{cases} X^{1/n}, & X \geq 0 \\ -|X|^{1/n}, & X < 0 \end{cases}, n \in \mathbb{Z}^+ \quad (1.84)$$

如何求出输出随机变量  $Y$  的概率密度呢? 显然, 若能找到求随机变量  $X$  的函数  $Y = g(X)$  的概率密度的方法, 就能解决上述的实际问题。

#### 1.1.4.2 一维随机变量函数的分布

对于一维随机变量函数的分布, 分成两种情况来讨论。

##### 1. 单值变换

随机变量  $X$  和  $Y$  存在单调函数关系  $Y = g(X)$ , 存在唯一反函数  $X = h(Y)$ , 即若有一个  $X$  出现, 则必有一个与其对应的  $Y$  出现。若  $X$  位于  $(x_0, x_0 + dx)$  区间内, 则  $Y$  必位于  $(y_0, y_0 + dy)$  区间内。因此,  $X$  落在区间  $(x_0, x_0 + dx)$  内的概率等于  $Y$  落在区间  $(y_0, y_0 + dy)$

的概率, 有

$$P\{x_0 < X \leq x_0 + dx\} = P\{y_0 < Y \leq y_0 + dy\}. \quad (1.85)$$

可得

$$f_Y(y)dy = f_X(x)dx. \quad (1.86)$$

所以

$$f_Y(y) = f_X(x) \frac{dx}{dy} = [h'(y)] \cdot f_X[h(y)]. \quad (1.87)$$

由于概率密度不可能取负值, 所以  $dx/dy$  应取绝对值, 即

$$f_Y(y) = |h'(y)| \cdot f_X[h(y)]. \quad (1.88)$$

这样, 不论  $h(y)$  是单调增函数 ( $h'(y) > 0$ ), 还是单调减函数 ( $h'(y) < 0$ ), 上式均成立。

**例 1.43** 随机变量  $X$  和  $Y$  之间成线性关系:  $Y = X + 5$ 。已知随机变量  $X$  服从标准高斯分布。求随机变量  $Y$  的概率密度。

解: 随机变量  $X$  和  $Y$  之间存在唯一的反函数, 其表达式为  $X = h(Y) = Y - 5$ , 则  $x = h(y) = y - 5$ , 所以  $|h(y)| = 1$ 。由单值变换公式可得

$$f_Y(y) = |h'(y)| f_X[h(y)] = 1 \cdot f_X(y - 5) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-5)^2}{2}}; \quad (1.89)$$

可见,  $X$  服从高斯分布, 其线性函数  $y = x + 5$  也服从高斯分布。

思考: 若随机变量  $X$  服从高斯分布, 其线性函数  $Y = ax + b$  ( $a \neq 0$ ) 是否服从高斯分布?

## 2. 多值变换

随机变量  $X$  和  $Y$  存在非单调函数关系  $Y = g(X)$ , 反函数  $X = h(Y)$  不唯一。

① 如果  $Y$  值可能对应着两个  $X$  值,  $X_1 = h_1(Y)$  和  $X_2 = h_2(Y)$ 。所以, 当  $X$  位于  $(x_1, x_1 + dx_1)$  内或位于  $(x_2, x_2 + dx_2)$  内时, 两事件中只要有一个发生, 则  $Y$  位于  $(y_0, y_0 + dy)$  内的事件就发生。因此, 根据和事件概率的求法可得

$$f_Y(y)dy = f_X(x_1)dx_1 + f_X(x_2)dx_2. \quad (1.90)$$

将  $x_1$  用  $h_1(y)$  代入,  $x_2$  用  $h_2(y)$  代入, 可得

$$f_Y(y) = |h'_1(y)| f_X[h_1(y)] + |h'_2(y)| f_X[h_2(y)]. \quad (1.91)$$

② 更复杂的是一个  $Y$  值对应多个值的情况。此时, 将上式作进一步推广, 由概率可加性可得

$$f_Y(y)dy = f_X(x_1)dx_1 + f_X(x_2)dx_2 + f_X(x_3)dx_3 + \dots \quad (1.92)$$

则

$$f_Y(y) = |h'_1(y)| f_X[h_1(y)] + |h'_2(y)| \cdot f_X[h_2(y)] + |h'_3(y)| f_X[h_3(y)] + \dots \quad (1.93)$$

**例 1.44** 已知随机变量  $X$  服从标准高斯分布, 求随机变量  $Y = X^2$  的概率。

解: 随机变量  $X$  和  $Y$  之间的反函数关系为

$$X = h(Y) = \pm\sqrt{Y}. \quad (1.94)$$

其反函数导数的绝对值为

$$|h'_1(y)| = |h'_2(y)| = \left| \frac{dx}{dy} \right| = \frac{1}{2\sqrt{y}}. \quad (1.95)$$

① 当  $y < 0$  时,  $\{X^2 \leq y\}$  为不可能事件, 所以  $P\{X^2 \leq y\} = 0$ , 得  $F_Y(y) = 0$ , 因此, 当  $y < 0$  时, 其概率密度  $f(y) = 0$ 。

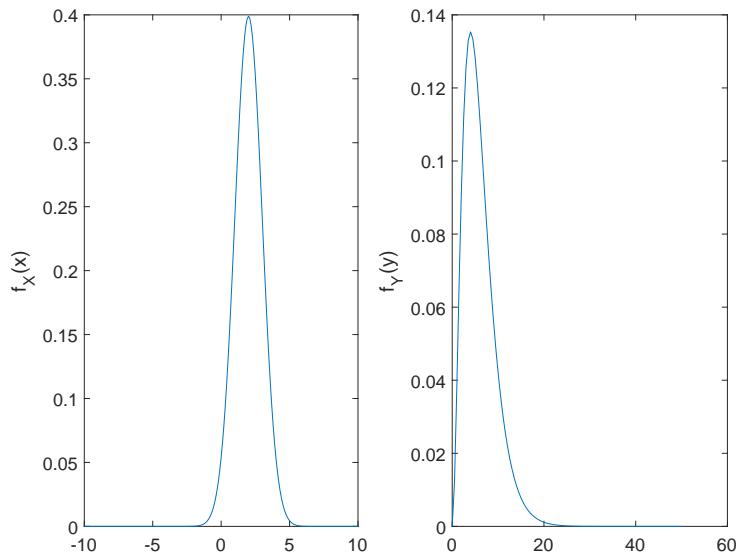
② 当  $y > 0$  时, 反函数为  $X = h(Y) = \pm\sqrt{Y}$  是双值变换。已知变量  $x$  服从标准高斯分布, 则

$$\begin{aligned} f_Y(y) &= |h'_1(y)| f_X[h_1(y)] + |h'_2(y)| f_X[h_2(y)] \\ &= \frac{1}{2\sqrt{y}} \left( \frac{1}{\sqrt{2\pi}} e^{-(\sqrt{y})^2/2} + \frac{1}{\sqrt{2\pi}} e^{-(\sqrt{y})^2/2} \right) \\ &= \frac{1}{2\sqrt{y}} \cdot \frac{2}{\sqrt{2\pi}} e^{-x/2} = \frac{1}{\sqrt{2\pi}} y^{-1/2} e^{-y/2}. \end{aligned} \quad (1.96)$$

综合①②可得

$$f_Y(y) = \begin{cases} \frac{1}{\sqrt{2\pi}} y^{-1/2} e^{-y/2}, & y > 0 \\ 0, & y \leq 0 \end{cases}. \quad (1.97)$$

称之为  $\chi^2$  分布。说明一个高斯变量经过平方变换以后, 其概率密度为  $\chi^2$  分布, 如图 1-14 所示。

图 1-14 输入、输出随机变量的概率密度:  $mu = 2, \sigma = 1, n = 6$ **例 1.1.5**

设随机变量  $X \sim U[-1, 2]$ , 求随机变量函数  $Y = X^2$  的概率密度.

**解:**  $X$  的概率密度

$$f_X(x) = \begin{cases} \frac{1}{3}, & -1 \leq x \leq 2 \\ 0, & \text{其他} \end{cases},$$

由  $Y = X^2 \in [0, 4]$ , 用分布函数求  $Y = X^2$  的概率密度  $f_Y(y)$ .

1. 当  $y < 0$ ,  $F_Y(y) = P\{Y \leq y\} = P\{\emptyset\} = 0$ .
2. 当  $y \geq 4$ ,  $F_Y(y) = P\{Y \leq y\} = P\{\Omega\} = 1$ .
3. 当  $0 \leq y < 4$ ,  $x \in [-1, 2], 0 \leq |x| \leq 2$ ,  $F_Y(y) = P\{Y \leq y\} = P\{X^2 \leq y\} = P\{-\sqrt{y} \leq X \leq \sqrt{y}\}$ .
  - (a) 当  $0 \leq y < 1$  时,  $0 \leq \sqrt{y} < 1$ ,  $-1 < -\sqrt{y} \leq 0$ , 且  $-1 \leq x \leq 2$ .

$$\begin{aligned} F_Y(y) &= P\{-\sqrt{y} \leq X \leq \sqrt{y}\} = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{3} dx \\ &= \frac{2}{3}\sqrt{y}, \quad 0 \leq y < 1. \end{aligned}$$

(b) 当  $1 \leq y < 4$  时,  $1 \leq \sqrt{y} < 2$ ,  $-2 \leq -\sqrt{y} < -1$ , 且  $-1 \leq x \leq 2$ .

$$\begin{aligned} F_Y(y) &= P\{\{-\sqrt{y} \leq X \leq \sqrt{y}\} \cap \{-1 \leq X \leq 2\}\} \\ &= P\{-1 \leq X \leq \sqrt{y}\} = \int_{-1}^{\sqrt{y}} \frac{1}{3} dx = \frac{\sqrt{y} + 1}{3}. \end{aligned}$$

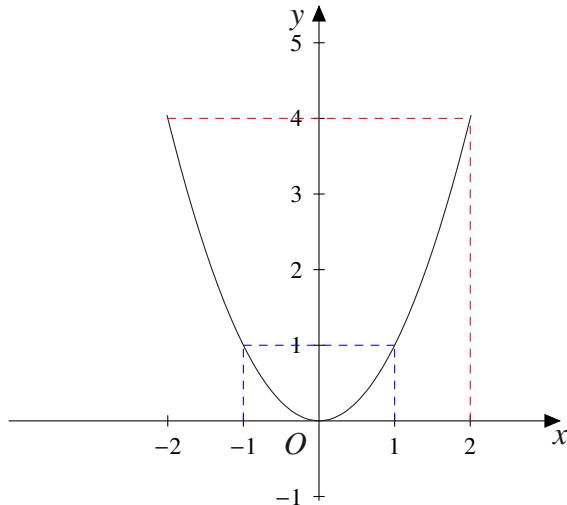


图 1-15 随机变量函数关系  $Y = X^2$

### 分布函数

$$F_Y(y) = \begin{cases} 0, & y < 0 \\ \frac{2\sqrt{y}}{3}, & 0 \leq y < 1 \\ \frac{\sqrt{y}+1}{3}, & 1 \leq y < 4 \\ 1, & y \geq 4 \end{cases}$$

对  $y$  求导, 得  $Y$  的概率密度函数

$$f_Y(y) = \begin{cases} \frac{1}{3\sqrt{y}}, & 0 \leq y < 1 \\ \frac{1}{6\sqrt{y}}, & 1 \leq y < 4 \\ 0, \text{ 其他} \end{cases}$$

### 1.1.4.3 二维随机变量函数的分布

求解二维问题所采用的方法基本上和一维情况相似, 仅仅是稍微复杂一些。已知二维随机变量  $(X_1, X_2)$  的联合概率密度为  $f_{X_1, X_2}(x_1, x_2)$ , 要求新的二维随机变量  $(Y_1, Y_2)$  的联

合概率密度  $(Y_1, Y_2)$ , 其中  $Y_1, Y_2$  分别为  $(X_1, X_2)$  的函数, 则

$$\begin{cases} Y_1 = g_1(X_1, X_2) \\ Y_2 = g_2(X_1, X_2). \end{cases} \quad (1.98)$$

函数  $g_1(\cdot), g_2(\cdot)$  可以是单值变换, 也可以是多值变换。

### 1. 单值变换

若解出的反函数

$$\begin{cases} X_1 = h_1(Y_1, Y_2) \\ X_2 = h_2(Y_1, Y_2) \end{cases} \quad (1.99)$$

是唯一的, 则称二维随机变量  $(X_1, X_2)$  与  $(Y_1, Y_2)$  之间是单值的函数变换, 简称单值变换与一维随机变量类似, 单值变换是一一对应的变换。换句话说, 当随机点落入  $x_1Ox_2$  平面上时, 在  $y_1Oy_2$  平面上有且仅有一个随机点与其对应, 反之亦然。假设  $dS_{x_1x_2}$  是  $x_1Ox_2$  平面上的一个任意闭域,  $dS_{y_1y_2}$  是它在  $y_1Oy_2$  平面上的映射, 如图 1-16 所示那么  $(X_1, X_2)$  点落入  $dS_{x_1x_2}$  区间的概率  $f_X(x_1, x_2) dS_{x_1x_2}$  等于它的映射  $(Y_1, Y_2)$  落入  $dS_{y_1y_2}$  区间的概率  $f_Y(y_1, y_2) dS_{y_1y_2}$ 。

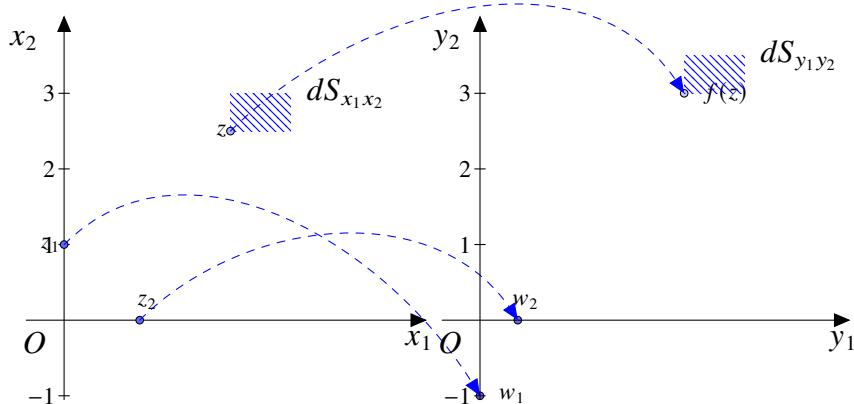


图 1-16 单值映射下, 函数变换对应的区间变换

所以, 新的二维随机变量  $(Y_1, Y_2)$  的概率密度为

$$f_Y(y_1, y_2) = f_X(x_1, x_2) \cdot \frac{dS_{x_1x_2}}{dS_{y_1y_2}}. \quad (1.100)$$

坐标转换中  $dS_{x_1x_2}$  和  $dS_{y_1y_2}$  之间的变换称为雅可比变换。可得雅可比行列式为

$$J = \frac{dS_{x_1x_2}}{dS_{y_1y_2}} = \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{vmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} \\ \frac{\partial h_2}{\partial y_1} & \frac{\partial h_2}{\partial y_2} \end{vmatrix}. \quad (1.101)$$

于是

$$\begin{aligned} f_Y(y_1, y_2) &= f_X(x_1, x_2) \left| \frac{dS_{x_1 x_2}}{dS_{y_1 y_2}} \right| = |J| f_X(x_1, x_2) \\ &= |J| f_X[h_1(y_1, y_2), h_2(y_1, y_2)]. \end{aligned} \quad (1.102)$$

**例 1.45** 已知二维随机变量  $(X_1, X_2)$  的联合概率密度

$$f_X(x_1, x_2) = \begin{cases} e^{-(x_1+x_2)}, & x_1 > 0, x_2 > 0 \\ 0, & \text{其他} \end{cases} \quad (1.103)$$

新的二维随机变量  $(Y_1, Y_2)$  是  $(X_1, X_2)$  的函数, 满足关系

$$Y_1 = \frac{X_1 + X_2}{2}, \quad Y_2 = \frac{X_1 - X_2}{2}. \quad (1.104)$$

求: ① 二维随机变量  $(Y_1, Y_2)$  的联合概率密度  $f_{XY}(y_1, y_2)$ . ② 边缘密度  $f_Y(y_1)$  和  $f_Y(y_2)$ , 说明  $Y_1$  与  $Y_2$  是否相互独立。

解: ① 由函数关系, 可以找出唯一的反函数

$$\begin{cases} x_1 = h_1(y_1, y_2) = y_1 + y_2 \\ x_2 = h_2(y_1, y_2) = y_1 - y_2 \end{cases} \quad (1.105)$$

则其雅可比行列式为

$$J = \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2. \quad (1.106)$$

可得

$$\begin{aligned} f_Y(y_1, y_2) &= |J| f_X[h_1(y_1, y_2), h_2(y_1, y_2)] \\ &= 2f_X[y_1 + y_2, y_1 - y_2] = \begin{cases} 2e^{-2y_1}, & y_1 > |y_2| \geq 0 \\ 0, & \text{其他} \end{cases}. \end{aligned} \quad (1.107)$$

其中, 根据  $(X_1, X_2)$  与  $(Y_1, Y_2)$  的函数关系, 将  $(X_1, X_2)$  的值域映射到  $y_1 O y_z$  平面, 找出  $(Y_1, Y_2)$  的值域, 如图 1-17 所示。 $(Y_1, Y_2)$  的值域满足

$$\begin{cases} x_1 > 0 \\ x_2 > 0 \end{cases} \Rightarrow \begin{cases} y_1 + y_2 > 0 \\ y_1 - y_2 > 0 \end{cases} \Rightarrow y_1 > |y_2| \geq 0. \quad (1.108)$$

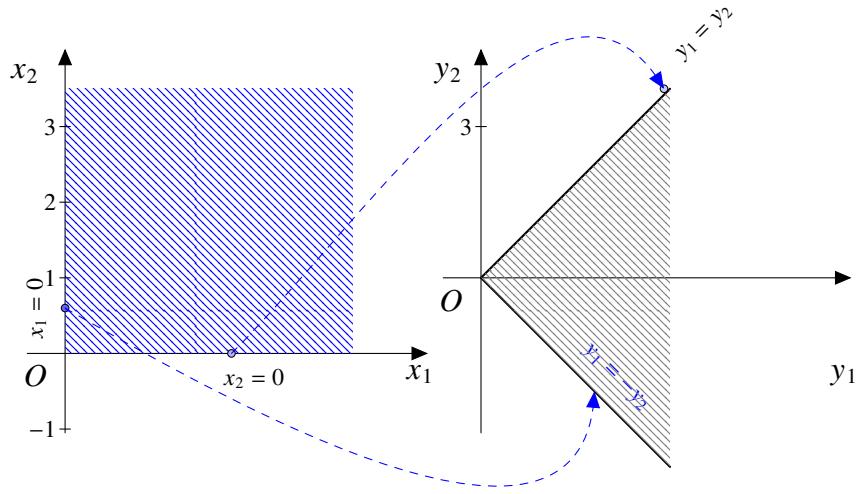


图 1-17 例1.45 的区间变换

② 其边缘分布为

$$\begin{aligned} f_Y(y_1) &= \int_{-y_1}^{y_1} f_Y(y_1, y_2) dy_2 = \int_{-y_1}^{y_1} 2e^{-2y_1} dy_2 = 4y_1 e^{-2y_1}, y_1 > 0 \\ f_Y(y_2) &= \begin{cases} \int_{y_2}^{\infty} 2e^{-2y_1} dy_1 = e^{-2y_2}, y_2 \geq 0 \\ \int_{-y_2}^{\infty} 2e^{-2y_1} dy_1 = e^{2y_2}, y_2 < 0 \end{cases} = e^{-2|y_2|}, -\infty < y_2 < +\infty \end{aligned} \quad (1.109)$$

由于

$$\begin{aligned} f_Y(y_1) \cdot f_Y(y_2) &= 4y_1 e^{-2(y_1+|y_2|)} \\ &\neq f_Y(y_1, y_2) = \begin{cases} 2e^{-2y_1}, & y_1 > |y_2| \geq 0 \\ 0, & \text{其他} \end{cases}. \end{aligned} \quad (1.110)$$

所以  $Y_1$  与  $Y_2$  不是相互独立的。

## 2. 多值变换

若能从  $g_1(\cdot), g_2(\cdot)$  中解出的  $X_1$  和  $X_2$  不是唯一的, 如解出两对反函数

$$\begin{cases} X_{a_1} = h_{a_1}(Y_1, Y_2) \\ X_{a_2} = h_{a_2}(Y_1, Y_2) \end{cases}, \begin{cases} X_{b_1} = h_{b_1}(Y_1, Y_2) \\ X_{b_2} = h_{b_2}(Y_1, Y_2) \end{cases}. \quad (1.111)$$

这种  $(Y_1, Y_2)$  有多个  $(X_1, X_2)$  与其对应的函数变换, 称之为多值变换。与一维的多值变换情况类似, 可用概率的加法定理, 求二维随机变量的多值变换函数的概率密度, 即

$$f_Y(y_1, y_2) = |J_a| f_X[h_{a_1}(y_1, y_2), h_{a_2}(y_1, y_2)] + |J_b| f_X[h_{b_1}(y_1, y_2), h_{b_2}(y_1, y_2)]. \quad (1.112)$$

其中, 雅可比行列式

$$J_a = \frac{\partial(x_{a_1}, x_{a_2})}{\partial(y_1, y_2)} = \begin{vmatrix} \frac{\partial h_{a_1}}{\partial y_1} & \frac{\partial h_{a_1}}{\partial y_2} \\ \frac{\partial h_{a_2}}{\partial y_1} & \frac{\partial h_{a_2}}{\partial y_2} \end{vmatrix}, \quad (1.113)$$

$$J_b = \frac{\partial(x_{b_1}, x_{b_2})}{\partial(y_1, y_2)} = \begin{vmatrix} \frac{\partial h_{b_1}}{\partial y_1} & \frac{\partial h_{b_1}}{\partial y_2} \\ \frac{\partial h_{b_2}}{\partial y_1} & \frac{\partial h_{b_2}}{\partial y_2} \end{vmatrix}. \quad (1.114)$$

#### 1.1.4.4 $n$ 维随机变量函数的分布

由上述的一维和二维随机变量函数变换的结论, 可用归纳法将其扩展到  $n$  维变换的情况。若  $n$  维随机变量  $(X_1, \dots, X_n)$  与其函数  $(Y_1, \dots, Y_n)$  间的变换是单值的, 即有唯一反函数

$$\begin{cases} X_1 = h_1(Y_1, Y_2, \dots, Y_n) \\ X_2 = h_2(Y_1, Y_2, \dots, Y_n) \\ \vdots \\ X_n = h_n(Y_1, Y_2, \dots, Y_n) \end{cases}. \quad (1.115)$$

则

$$\begin{aligned} f_Y(y_1, \dots, y_n) &= |J| \cdot f_X(x_1, \dots, x_n) \\ &= |J| \cdot f_X[h_1(y_1, \dots, y_n), \dots, h_n(y_1, \dots, y_n)]. \end{aligned} \quad (1.116)$$

雅可比行列式为

$$J = \frac{\partial(x_1, \dots, x_n)}{\partial(y_1, \dots, y_n)} = \begin{vmatrix} \frac{\partial h_1(y_1, \dots, y_n)}{\partial y_1} & \dots & \frac{\partial h_1(y_1, \dots, y_n)}{\partial y_n} \\ \vdots & & \vdots \\ \frac{\partial h_n(y_1, \dots, y_n)}{\partial y_1} & \dots & \frac{\partial h_n(y_1, \dots, y_n)}{\partial y_n} \end{vmatrix}. \quad (1.117)$$

**例 1.46** 已知  $n$  维随机变量  $(X_1, \dots, X_n)$  的联合概率密度  $f_X(x_1, \dots, x_n)$ , 求随机变量  $Y$  的概率密度  $f_Y(y)$ , 其中  $Y = \sum_{i=1}^n X_i$ .

解: 要想进行雅可比变换, 必须保证变换的维数相同, 因此必须构造新的  $n$  维随机变量  $(Y_1, \dots, Y_n)$ , 使满足

$$Y_1 = X_1, Y_2 = X_2, \dots, Y_{n-1} = X_{n-1}, Y_n = \sum_{i=1}^n X_i = Y. \quad (1.118)$$

解出反函数

$$X_1 = Y_1, X_2 = Y_2, \dots, X_{n-1} = Y_{n-1}, X_n = Y_n - \sum_{i=1}^{n-1} Y_i. \quad (1.119)$$

则  $n$  维雅可比行列式为

$$J = \begin{vmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ -1 & -1 & \cdots & -1 & 1 \end{vmatrix} = 1. \quad (1.120)$$

所以  $n$  维随机变量  $(Y_1, \dots, Y_n)$  的联合概率密度为

$$f_Y(y_1, \dots, y_n) = f_X\left(y_1, \dots, y_{n-1}, y_n - \sum_{i=1}^{n-1} y_i\right). \quad (1.121)$$

其关于  $Y$  的边缘分布, 即随机变量  $Y$  的概率密度  $f_Y(y)$  为

$$f_Y(y) = f_Y(y_n) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X\left(y_1, \dots, y_{n-1}, y_n - \sum_{i=1}^{n-1} y_i\right) dy_1 \cdots dy_{n-1}. \quad (1.122)$$

**例 1.47** 已知随机变量  $X_i \sim e(\lambda)$  且相互独立, 写出  $n$  维随机变量  $(X_1, \dots, X_n)$  的联合概率密度  $f_X(x_1, \dots, x_n)$ , 并求随机变量  $Y$  的概率密度  $f_Y(y)$ , 其中  $Y = \sum_{i=1}^n X_i$ .

#### 1.1.4.5 随机变量函数分布的广义定积分方法

随机变量的函数分布是概率论与数理统计中的重要内容之一, 关于二维随机变量分布函数与概率密度的求解是一个难点。对于二维随机变量的函数的概率密度的计算, 一般教材和文献仅给出随机变量的和、差、积和商的概率密度公式, 其余均须按定义先求其分布函数, 然后对分布函数求导得到概率密度。尽管求解思路不复杂, 但求分布函数时多需要分区域求二重积分, 而正确划分积分区域给出准确表达积分式往往难度较大, 极易出错, 运算量巨大。为了解决这一难题, 文中应用积分变换给出了二维随机变量的函数的概率密度的新计算公式 [宋明娟 2011]。利用这些公式求概率密度, 仅需完成广义定积分计算即可, 较分布函数法降低了积分重数, 简化了计算难度, 提高了计算效率和准确率。

### 定理 1.1.1

设二维随机变量  $(X, Y)$  的概率密度为  $f(x, y)$ ,  $Z = g(X, Y)$  为关于随机变量  $X$  或  $Y$  的概率密度, 可按如下方式计算

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x, y(x, z)) \left| \frac{\partial y}{\partial z} \right| dx. \quad (1.123)$$

或

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x(y, z), y) \left| \frac{\partial x}{\partial z} \right| dy. \quad (1.124)$$

证明 若当  $y \in (-\infty, +\infty)$  时,  $\frac{\partial z}{\partial y} > 0$ , 则函数  $z = g(x, y)$  关于变量  $y$  严格单调增加, 它的反函数  $y = y(x, z)$  及  $\frac{\partial y}{\partial z}$  必存在, 且  $\frac{\partial y}{\partial z} > 0$ , 于是雅可比行列式

$$J = \begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial y}{\partial z} \\ \frac{\partial x}{\partial \theta} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} 1 & \frac{\partial y}{\partial x} \\ 0 & \frac{\partial y}{\partial z} \end{vmatrix} = \frac{\partial y}{\partial z}.$$

$(X, Z)$  的联合分布为

$$\begin{aligned} F(x, z) &= P\{X \leq x, Z(X, Y) \leq z\} = P\{X \leq x, Y \leq y(x, z)\} \\ &= \int_{-\infty}^x dx \int_{-\infty}^{y(x, z)} f(x, y) dy = \frac{\partial y}{\partial z}. \end{aligned}$$

所以

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x, y(x, z)) \frac{\partial y}{\partial z} dx.$$

若当  $y \in (-\infty, +\infty)$  时,  $\frac{\partial z}{\partial y} < 0$ , 则

$$\begin{aligned} F(x, z) &= P\{X \leq x, Z(X, Y) \leq z\} = P\{X \leq x, Y \geq y(x, z)\} \\ &= \int_{-\infty}^x dx \int_{y(x, z)}^{+\infty} f(x, y) dy. \\ &= \int_{-\infty}^x dx \int_{-\infty}^z f(x, y(x, z)) \left( -\frac{\partial y}{\partial z} \right) dz. \end{aligned}$$

综上可得, 当  $z$  为  $y$  的严格单调函数且  $\frac{\partial y}{\partial z}$  处处存在, 则有

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x, y(x, z)) \left| \frac{\partial y}{\partial z} \right| dx.$$

同理, 当  $z$  为  $x$  的严格单调函数且  $\frac{\partial x}{\partial z}$  处处存在, 则有

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x(y, z), y) \left| \frac{\partial x}{\partial z} \right| dy.$$

推论 1 设二维连续型随机变量  $(X, Y)$  的概率密度为  $f(x, y)$ ,  $a, b$  为非零实数, 则

(1)  $Z = aX + bY$  的概率密度为

$$f_Z(z) = \int_{-\infty}^{+\infty} f\left(x, \frac{z - ax}{b}\right) \left|\frac{1}{b}\right| dx.$$

或

$$f_Z(z) = \int_{-\infty}^{+\infty} f\left(\frac{z - by}{a}, y\right) \left|\frac{1}{a}\right| dy.$$

(2)  $Z = \frac{Y}{aX}$  的概率密度为

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x, axz) |ax| dx.$$

(3)  $Z = aXY$  的概率密度为

$$f_Z(z) = \int_{-\infty}^{+\infty} f\left(x, \frac{z}{ax}\right) \left|\frac{1}{ax}\right| dx.$$

或

$$f_Z(z) = \int_{-\infty}^{+\infty} f\left(\frac{z}{ay}, y\right) \left|\frac{1}{ay}\right| dy.$$

### 例 1.1.6

设  $(X, Y)$  服从  $D = \{(x, y) \mid -1 \leq x \leq 1, 0 \leq y < 1\}$  上的均匀分布。  
试求  $Z = \frac{Y}{3X}$  的概率密度  $f_Z(z)$ .

解  $(X, Y)$  的联合概率密度为

$$f(x, y) = \begin{cases} \frac{1}{2}, & -1 \leq x \leq 1, 0 \leq y < 1 \\ 0, & \text{其他} \end{cases}$$

由  $Z = \frac{Y}{3X}$ , 得  $Y = 3XZ$ ,  $\frac{\partial Y}{\partial Z} = 3X$ ,

$$f(x, 3xz) = \begin{cases} \frac{1}{2}, & -1 \leq x \leq 0, \frac{1}{3x} \leq z < 0 \text{ 或 } 0 \leq x \leq 1, 0 \leq z < \frac{1}{3x} \\ 0, & \text{其他} \end{cases} \quad (1.125)$$

于是

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{+\infty} f(x, 3xz) |3x| dx = \begin{cases} \int_{\frac{1}{3}}^0 \frac{-3x}{2} dx, & -\infty < z < -\frac{1}{3} \\ \int_{-1}^0 \frac{-3x}{2} dx, & -\frac{1}{3} \leq z < 0 \\ \int_0^1 \frac{3x}{2} dx, & 0 \leq z < \frac{1}{3} \\ \int_0^{\frac{1}{3}} \frac{3x}{2} dx, & \frac{1}{3} \leq z < +\infty \end{cases} \\ &= \begin{cases} \frac{1}{12z^2}, & |z| \geq \frac{1}{3} \\ \frac{3}{4}, & |z| < \frac{1}{3} \end{cases} \end{aligned} \quad (1.126)$$

### 定理 1.1.2

设二维连续型随机变量  $(X, Y)$  的概率密度为  $f(x, y)$ ,  $Z = g(X, Y)$  关于随机变量  $X$  或  $Y$  在区间  $I_i (i = 1, 2, \dots, n)$  上为严格单调函数, 且  $\frac{\partial z}{\partial x}$  或  $\frac{\partial z}{\partial y}$  处处存在, 其在区间  $I_i$  上的反函数为  $y = y_i(x, z)$  或  $x = x_i(y, z)$ , 若  $I_i (i = 1, 2, \dots, n)$  互不相交, 且  $\bigcup_{i=1}^n I_i = (-\infty, +\infty)$ , 则随机变量  $Z$  的概率密度为

$$f_Z(z) = \sum_{i=1}^n \int_{-\infty}^{+\infty} f(x, y_i(x, z)) \left| \frac{\partial y_i}{\partial z} \right| dx. \quad (1.127)$$

或

$$f_Z(z) = \sum_{i=1}^n \int_{-\infty}^{+\infty} f(x_i(y, z), y) \left| \frac{\partial x_i}{\partial z} \right| dy. \quad (1.128)$$

证明 若  $Z = g(X, Y)$  关于随机变量  $Y$  在区间  $I_i (i = 1, 2, \dots, n)$  上为严格单调函数, 则随机变量  $(X, Z)$  的联合分布为

$$\begin{aligned} F(x, z) &= P\{X \leq x, Z \leq z\} = P\left\{\bigcup_{i=1}^n (X \leq x, g(X, Y) \leq z, Y \in I_i)\right\} \\ &= \sum_{i=1}^n P\{X \leq x, g(X, Y) \leq z, Y \in I_i\} \\ &= \sum_{i=1}^n \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} f(x, y_i(x, z)) \left| \frac{\partial y_i}{\partial z} \right| dz \\ &= \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} \left[ \sum_{i=1}^n f(x, y_i(x, z)) \left| \frac{\partial y_i}{\partial z} \right| \right] dz. \end{aligned} \quad (1.129)$$

所以  $(X, Z)$  的联合概率密度为

$$\sum_{i=1}^n f(x, y_i(x, z)) \left| \frac{\partial y_i}{\partial z} \right|.$$

$Z$  的概率密度为

$$f_Z(z) = \sum_{i=1}^n \int_{-\infty}^{+\infty} f(x, y_i(x, z)) \left| \frac{\partial y_i}{\partial z} \right| dx. \quad (1.130)$$

同理可证当  $Z = g(X, Y)$  关于随机变量  $X$  在区间  $I_i (i = 1, 2, \dots, n)$  上严格单调函数时,

$$f_Z(z) = \sum_{i=1}^n \int_{-\infty}^{+\infty} f(x_i(y, z), y) \left| \frac{\partial x_i}{\partial z} \right| dy.$$

**例 1.1.7**

设随机变量  $X$  和  $Y$  相互独立, 并且都服从正态分布  $N(0, \sigma^2)$ , 求随机变量  $Z = \sqrt{X^2 + Y^2}$  的概率密度。

解由于  $X$  和  $Y$  相互独立, 因此  $(X, Y)$  的概率密度为

$$f(x, y) = f_X(x)f_Y(y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, x, y \in \mathbb{R}. \quad (1.131)$$

由  $z = \sqrt{x^2 + y^2}$ , 解得

$$\begin{cases} y_1 = \sqrt{z^2 - x^2}, & \frac{\partial y_1}{\partial z} = \frac{z}{\sqrt{z^2 - x^2}} > 0, \quad |x| \leq z \\ y_2 = -\sqrt{z^2 - x^2}, & \frac{\partial y_2}{\partial z} = -\frac{z}{\sqrt{z^2 - x^2}} < 0, \quad |x| \leq z \end{cases}, \quad (1.132)$$

$$f(x, \pm\sqrt{z^2 - x^2}) = \begin{cases} \frac{z}{2\pi\sigma^2} e^{-\frac{j^2}{2\sigma^2}}, & |x| < z \\ 0, & \text{其他} \end{cases}. \quad (1.133)$$

当  $z > 0$  时,

$$\begin{aligned} f_z(z) &= \int_{-z}^z f(x, \sqrt{z^2 - x^2}) \frac{z}{\sqrt{z^2 - x^2}} dx \\ &\quad + \int_{-z}^z f(x, -\sqrt{z^2 - x^2}) \frac{z}{\sqrt{z^2 - x^2}} dx \\ &= 4 \int_0^z \frac{1}{2\pi\sigma^2} e^{-\frac{z^2}{2\sigma^2}} \frac{z}{\sqrt{z^2 - x^2}} dx = \frac{2z}{\pi\sigma^2} e^{-\frac{z^2}{2\sigma^2}} \left[ \arcsin \frac{x}{z} \right]_0^z \\ &= \frac{z}{\sigma^2} e^{-\frac{z^2}{2\sigma^2}}. \end{aligned} \quad (1.134)$$

显然, 当  $z \leq 0$  时,  $f_z(z) = 0$ .

综上可得

$$f_z(z) = \begin{cases} \frac{z}{\sigma^2} e^{-\frac{z^2}{2\sigma^2}}, & z \geq 0 \\ 0, & z < 0 \end{cases}, Z \sim \text{Rayleigh}(\sigma). \quad (1.135)$$

### 1.1.5 随机变量的数字特征

在实用中, 概率分布函数(或概率密度函数)往往很难获得, 有时也仅仅需要随机变量的一些统计特性。下面以连续型随机变量为例, 描述随机变量的期望、方差和相关系数等数字特征。

### 定义 1.48 数学期望

变量  $X$  的数学期望定义为

$$E\{X\} = \int_{-\infty}^{+\infty} xf(x)dx. \quad (1.136)$$

数学期望是在概率统计意义上的一种平均，称为集合均值，简称集平均，因此  $E\{X\}$  又称为均值。



1) 数学期望具有如下一些基本性质：

(1) 常量的数学期望等于常量本身。

(2) 对常数  $a_i$  ( $i = 1, 2, \dots, n$ )，有

$$E\left\{\sum_{i=1}^n a_i X_i\right\} = \sum_{i=1}^n a_i E\{X_i\}. \quad (1.137)$$

(3)  $X_i$  ( $i = 1, 2, \dots, n$ ) 相互独立，则

$$E\{X_1 X_2 \cdots X_n\} = \prod_{i=1}^n E\{X_i\}. \quad (1.138)$$

**例 1.49** 随机变量  $X$  在区间  $(a, b)$  上服从均匀分布，求  $X$  的数学期望。

解：由于  $X$  服从均匀分布，则概率密度为

$$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{其他} \end{cases} \quad (1.139)$$

则数学期望为

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_a^b \frac{x}{b-a} dx = \frac{b+a}{2}. \quad (1.140)$$

2) 一维随机变量函数的数学期望

实际应用中，不仅要求随机变量的期望，还要求随机变量函数的数学期望。例如：飞机机翼受到的压力  $W = KV$  ( $K > 0$ ) 是常数， $V$  表示风速，是个随机变量，若要计算受到的压力  $W$  的统计平均值，即求随机变量  $V$  的函数  $W$  的数学期望。下面讨论已知随机变量  $X$  的分布，求其函数  $Y = g(X)$  的数学期望。

已知随机变量  $X$  的概率密度函数  $f_X(x)$ ，且随机变量  $Y = g(X)$ ，其中  $g(\cdot)$  是连续型实函数。随机变量  $Y$  的数学期望为

$$E[Y] = \int_{-\infty}^{\infty} y f_Y(y) dy. \quad (1.141)$$

由于  $f_Y(y)$  未知, 故下一步利用  $f_X(x)$  表示, 求出  $f_Y(y)$ .

① 若  $g(\cdot)$  是单值变换, 则

$$f_Y(y)dy = f_X(x)dx \Rightarrow f_Y(y) = f_X(x)dx/dy. \quad (1.142)$$

代入期望定义式, 得

$$E[Y] = \int_{-\infty}^{\infty} y f_Y(y) dy = \int_{-\infty}^{\infty} g(x) f_X(x) dx = E[g(X)]. \quad (1.143)$$

② 若  $g(\cdot)$  是多值变换

$$\begin{aligned} f_Y(y)dy &= f_X(x_1)dx_1 + f_X(x_2)dx_2 + \dots = \sum_{i=1}^n f_X(x_i)dx_i \\ f_Y(y) &= [f_X(x_1)dx_1 + f_X(x_2)dx_2 + \dots]/dy = \frac{\sum_{i=1}^n f_X(x_i)dx_i}{dy}. \end{aligned} \quad (1.144)$$

代入期望定义式, 得

$$\begin{aligned} E[Y] &= \int_{D_{x_1}} g(x_1) f_X(x_1) dx_1 + \int_{D_{x_2}} g(x_2) f_X(x_2) dx_2 + \dots \\ &= \int_{-\infty}^{\infty} g(x) f_X(x) dx = E[g(X)]. \end{aligned} \quad (1.145)$$

其中  $D_{x_1}, D_{x_2}, \dots$  为  $f_X(x_1), f_X(x_2), \dots$  的定义域。

综合①②可知, 无论函数  $g(X)$  是单值还是多值变换随机变量。函数的期望定义如下:

### 定义 1.50 随机变量函数的期望

若连续型变量  $X$  的概率密度为  $f_X(x)$ , 且  $\int_{-\infty}^{\infty} |g(x)| f_X(x) dx < \infty$ , 则函数  $g(X)$  的期望为

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx. \quad (1.146)$$

若  $X$  为离散型变量, 且  $\sum_{k=1}^{\infty} |g(x_k)| p_k < \infty$ , 则函数  $g(X)$  的期望为

$$E[g(X)] = \sum_{k=1}^{\infty} g(x_k) p_k. \quad (1.147)$$

**例 1.51** 随机变量  $X$  在区间  $(a, b)$  上服从均匀分布, 求  $g(X) = x^2 + 1$  的数学期望.

解: 由于  $X$  服从均匀分布, 则概率密度为

$$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{其他} \end{cases} \quad (1.148)$$

则函数的数学期望为

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx = \int_a^b \frac{x^2 + 1}{b-a} dx = \frac{1}{3} (a^2 + ab + b^2) + 1. \quad (1.149)$$

### 3. 二维随机变量及其函数的数学期望

#### 定义 1.52 二维连续型随机变量函数的期望

设  $(X, Y)$  是定义在概率空间  $(\Omega, F, P)$  上的二维连续型随机变量, 且联合概率密度  $f_{XY}(x, y)$  已知, 则由联合概率密度与边缘概率密度的关系及期望定义, 可得

$$\begin{cases} E[X] = \int_{-\infty}^{\infty} x f(x) dx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_{XY}(x, y) dx dy; \\ E[Y] = \int_{-\infty}^{\infty} y f(y) dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f_{XY}(x, y) dx dy. \end{cases} \quad (1.150)$$

#### 定义 1.53 二维离散型随机变量函数的期望

若  $(X, Y)$  为离散型随机变量, 且联合概率分布率  $P\{X = x, Y = y\}$  已知, 则

$$\begin{cases} E[X] = \sum_i x_i P\{X = x_i\} = \sum_i \sum_j x_i P\{X = x_i, Y = y_j\}; \\ E[Y] = \sum_i y_j P\{Y = y_j\} = \sum_j \sum_i y_j P\{X = x_i, Y = y_j\}. \end{cases} \quad (1.151)$$

仿照单个随机变量函数求期望的方法, 二维随机变量函数  $g(X, Y)$  的数学期望定义为

#### 定义 1.54 二维离散型随机变量函数的期望

二维随机变量函数  $g(X, Y)$  的数学期望为

$$\begin{aligned} E[g(X, Y)] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{XY}(x, y) dx dy. \\ E[g(X, Y)] &= \sum_i \sum_j g(x_i, y_j) P\{X = x_i, Y = y_j\}. \end{aligned} \quad (1.152)$$

### 4. $n$ 维随机变量的数学期望

#### 定义 1.55 $n$ 维连续型随机变量的期望

设  $(X_1, X_2, \dots, X_n)$  是定义在概率空间  $(\Omega, F, P)$  上的  $n$  维连续型随机变量, 若其联合概率密度分布为  $f_X(x_1, x_2, \dots, x_n)$ , 则与二维情况类似, 有

$$E[X_i] = \underbrace{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty}}_{n-1} x_i f_X(x_1, \dots, x_n) dx_1 \dots dx_n, i = 1, 2, \dots, n; \quad (1.153)$$

### 定义 1.56 $n$ 维连续型随机变量函数的期望

若  $n$  维随机变量的函数为  $g(X_1, X_2, \dots, X_n)$ , 则  $n$  维随机变量函数的数学期望为

$$E[g(X_1, X_2, \dots, X_n)] = \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{n\text{重积分}} g(x_1, \dots, x_n) f_X(x_1, \dots, x_n) dx_1 \cdots dx_n. \quad (1.154)$$

### 定义 1.57 随机矢量的期望

当  $n$  维随机变量  $(X_1, X_2, \dots, X_n)$  用随机矢量  $\mathbf{X} = [X_1 \ \dots \ X_n]^T$  来表示, 且若随机矢量  $\mathbf{X}$  中的每个分量  $X_i$  的数学期望均存在 ( $E[X_i] = m_i$ ), 则随机矢量  $\mathbf{X}$  的数学期望为

$$E(\mathbf{X}) = \begin{bmatrix} E(X_1) \\ \vdots \\ E(X_n) \end{bmatrix} = \begin{bmatrix} m_1 \\ \vdots \\ m_n \end{bmatrix} = \mathbf{M}_X. \quad (1.155)$$

可见, 随机矢量  $\mathbf{X}$  的数学期望是一个常数矢量, 常用  $\mathbf{M}_X$  表示。

**例 1.58** 设  $n$  维随机变量  $(X_1, X_2, \dots, X_n)$  的函数为  $g(X_1, X_2, \dots, X_n) = \sum_{i=1}^n a_i X_i$ , 其中权重  $a_i$  是常数。求数学期望  $E[g(X_1, X_2, \dots, X_n)]$ 。

解: 由  $n$  维随机变量函数的期望定义

$$\begin{aligned} E[g(X_1, X_2, \dots, X_n)] &= \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{n\text{重积分}} g(x_1, \dots, x_n) f_X(x_1, x_2, \dots, x_n) dx_1 \cdots dx_n \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (a_1 x_1 + a_2 x_2 + \cdots + a_n x_n) f_X(x_1, \dots, x_n) dx_1 \cdots dx_n \\ &= \sum_{i=1}^n \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{n\text{重积分}} a_i x_i f_X(x_1, \dots, x_n) dx_1 \cdots dx_n. \end{aligned} \quad (1.156)$$

根据边缘概率密度的公式, 和式 (1.156) 中每一项都可以化成

$$\begin{aligned} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} a_i x_i f_X(x_1, \dots, x_n) dx_1 \cdots dx_n &= \int_{-\infty}^{\infty} a_i x_i f_{X_i}(x_i) dx_i = E[a_i X_i] \\ &= a_i E[X_i]. \end{aligned} \quad (1.157)$$

所以

$$E \left[ \sum_{i=1}^n a_i X_i \right] = \sum_{i=1}^n a_i E [X_i]. \quad (1.158)$$

由此可见, 随机变量加权求和的均值等于各随机变量均值的加权和。

**例 1.59** 已知  $n$  个相互独立的随机变量  $X_1, X_2, \dots, X_n$  分别为  $E [X_1], E [X_2], \dots, E [X_n]$ 。求其函数  $g (X_1, X_2, \dots, X_n) = \prod_{i=1}^n X_i$  的数学期望。

解: 由  $n$  维随机变量函数的期望定义

$$\begin{aligned} E [g (X_1, X_2, \dots, X_n)] &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g (x_1, \dots, x_n) f_X (x_1, x_2, \dots, x_n) dx_1 \cdots dx_n \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1 x_2 \cdots x_n f_X (x_1, \dots, x_n) dx_1 \cdots dx_n \text{ (相互独立条件下)} \\ &= \int_{-\infty}^{\infty} x_1 f_{x_1} (x_1) dx_1 \cdot \int_{-\infty}^{\infty} x_2 f_{x_2} (x_2) dx_2 \cdots \cdots \int_{-\infty}^{\infty} x_n f_{x_n} (x_n) dx_n \\ &= E [X_1] E [X_2] \cdots E [X_n]. \end{aligned} \quad (1.159)$$

即

$$E \left[ \prod_{i=1}^n X_i \right] = \prod_{i=1}^n E [X_i]. \quad (1.160)$$

由上式可见,  $n$  个相互独立的随机变量乘积的期望等于  $n$  个随机变量期望的乘积。

归纳起来, 数学期望的基本性质有

1° 若随机变量满足  $a \leq X \leq b, a, b$  常数, 则其数学期望  $a \leq E[X] \leq b$ 。

2° 常数  $C$  的期望  $E[C] = C$ 。

3° 对任意常数  $b, a_i (i = 1, 2, \dots, n)$ , 有

$$E \left[ \sum_{i=1}^n a_i X_i + b \right] = \sum_{i=1}^n a_i E [X_i] + b. \quad (1.161)$$

4° 若随机变量  $X$  与随机变量  $Y$  互不相关, 则

$$E[XY] = E[X]E[Y]. \quad (1.162)$$

5° 若  $n$  个随机变量  $(X_1, X_2, \dots, X_n)$  相互独立, 则

$$E \left[ \prod_{i=1}^n X_i \right] = \prod_{i=1}^n E [X_i]. \quad (1.163)$$

### 定义 1.1.1

设  $X \sim N(0, 1)$ , 若  $z_\alpha$  满足  $P\{X > z_\alpha\} = \alpha$  ( $0 < \alpha < 1$ ), 则称  $z_\alpha$  为标准正态分布的上  $\alpha$  分位点 (图 1-18).  $\Phi(z_\alpha) = P\{X \leq z_\alpha\} = 1 - \alpha$ .

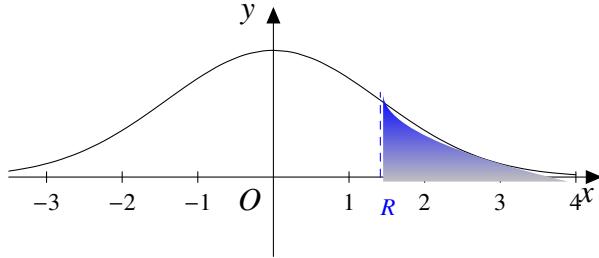


图 1-18 上  $\alpha$  分位点

### 注 1.60.

$$z_{0.05} = 1.645, z_{0.005} = 2.57, z_{0.025} = 1.96.$$

在本小节最后, 我们先列出一些关于多维随机变量函数经常用到的结论:

**性质 (1)** 若  $X_i \sim N(\mu_i, \sigma_i^2)$  ( $i = 1, 2, \dots, n$ ), 则随机变量  $Z = \sum_{i=1}^n c_i X_i \sim N\left(\sum_{i=1}^n c_i \mu_i, \sum_{i=1}^n c_i^2 \sigma_i^2\right)$ , 其中  $c_i$  ( $i = 1, 2, \dots, n$ ) 是任意常数.

**性质 (2)** 若  $X_i \sim N(0, \sigma^2)$  ( $i = 1, 2, \dots, n$ ) 且相互独立, 则  $Z = \sum_{i=1}^n X_i^2$  服从自由度为  $n$  的中心  $\chi^2$  分布:

$$f(z) = \begin{cases} \frac{1}{\sigma^n 2^{\frac{n}{2}} \Gamma(\frac{n}{2})} z^{\frac{n}{2}-1} e^{-\frac{z}{2}}, & z > 0 \\ 0, & z \leq 0 \end{cases} \quad (1.164)$$

**性质 (3)** 若  $X_i \sim N(\mu, \sigma^2)$  ( $i = 1, 2, \dots, n$ ) 且相互独立, 则  $Z = \sum_{i=1}^n X_i^2$  服从自由度为  $n$  的非中心  $\chi^2$  分布

$$f(z) = \frac{\sigma^2}{2} e^{-(x+\lambda)/2} \left(\frac{x}{\lambda}\right)^{n/4-1/2} I_{n/2-1}(\sqrt{\lambda x}), \quad (1.165)$$

式中,  $\lambda = n(\mu/\sigma)^2$  为非中心参数,  $I_{n/2-1}(\cdot)$  为  $n/2 - 1$  阶第一类贝塞尔函数, 具体表达式见 (1.167).

$n$  阶第一类贝塞尔函数,  $c = -n$  时,

$$I_{-n}(x) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m! \Gamma(-n+m+1)} \left(\frac{x}{2}\right)^{-n+2m}, \quad n \neq 1, 2, \dots \quad (1.166)$$

$$I_n(x) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m! \Gamma(n+m+1)} \left(\frac{x}{2}\right)^{n+2m}, \quad (1.167)$$

**性质 (3)** 若随机变量  $X$  和  $Y$  相互独立, 且都服从  $N(0, \sigma^2)$ , 则  $Z = \sqrt{X^2 + Y^2}$  服从参数为  $\sigma$  的瑞利分布

$$f(z) = \begin{cases} \frac{z}{\sigma^2} e^{-\frac{z^2}{2\sigma^2}}, & z \geq 0 \\ 0, & z < 0 \end{cases} \quad (1.168)$$

### 1.1.6 条件数学期望

前面我们已经讨论了条件分布的概念, 现在引入条件数学期望的概念。它在随机过程、时间序列分析和统计判决理论中起着重要作用。

#### 1. 随机变量关于某给定值的条件期望

1) 设  $(X, Y)$  是定义在同一概率空间上的二维连续型随机变量。 $Y$  的数学期望可以通过  $Y$  本身概率密度  $f_Y(y)$  来计算, 也可用  $(X, Y)$  的联合概率密度  $f_{XY}(x, y)$  来计算, 也可以借助条件分布  $f_Y(y|x)$  来计算, 如

$$\begin{aligned} E[Y] &= \int_{-\infty}^{\infty} y f_Y(y) dy = \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f_{XY}(x, y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f_Y(y|x) f_X(x) dx dy \\ &= \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} y f_Y(y|x) dy \right] \cdot f_X(x) dx. \end{aligned} \quad (1.169)$$

注意方括号  $[ \cdot ]$  内的项,  $\int_{-\infty}^{\infty} y f_Y(y|x) dy$  也是  $y$  的一种统计平均, 只不过权重不是一般的非条件概率, 而是条件概率。常称它为  $Y$  在  $\{X = x\}$  下  $Y$  的条件数学期望, 并用符号  $E[Y|X = x]$  表示, 即

$$E[Y|X = x] = \int_{-\infty}^{\infty} y f_Y(y|x) dy. \quad (1.170)$$

2) 上述定义方法也适用于  $(X, Y)$  为离散型随机变量的情况:

$$\begin{aligned} E[Y] &= \sum_j y_j P\{Y = y_j\} = \sum_i \sum_j y_j P_{ij} = \sum_i \sum_j y_j P\{X = x_i, Y = y_j\} \\ &= \sum_i \left[ \sum_j y_j P\{Y = y_j | X = x_i\} \right] P\{X = x_i\}. \end{aligned} \quad (1.171)$$

离散型随机变量  $Y$  在给定条件  $X = x_i$  下的条件数学期望表示为

$$E[Y|X = x_i] = \sum_j y_j P\{Y = y_j | X = x_i\}. \quad (1.172)$$

3) 此外, 若  $X$  是离散型随机变量, 而  $Y$  是连续型随机变量, 而且对所有的  $x_i$  ( $i = 1, 2, \dots$ ) 和  $y$  取值的条件概率密度  $f_Y(y|X = x_i)$  都存在, 则

$$E[Y] = \sum_i \left[ \int_{-\infty}^{\infty} y f_Y(y|X = x_i) dy \right] P\{X = x_i\}. \quad (1.173)$$

连续型随机变量  $Y$  在给定条件  $X = x_i$  下的条件数学期望表示为

$$E[Y|X = x_i] = \int_{-\infty}^{\infty} y f_Y(y|X = x_i) dy. \quad (1.174)$$

由于  $Y$  的条件数学期望是对  $Y$  的所有取值求统计平均, 因此由函数数学期望的求法, 可以推得

$$E[g(Y)|X=x] = \int_{-\infty}^{\infty} g(y)f_Y(y|x)dy. \quad (1.175)$$

$$E[g(X,Y)|X=x] = \int_{-\infty}^{\infty} g(x,y)f_Y(y|x)dy. \quad (1.176)$$

$$E[g_1(X)g_2(Y)|X=x] = g_1(x)E[g_2(Y)|X=x]. \quad (1.177)$$

## 2. 一随机变量关于另一随机变量的条件期望

由条件期望  $E[Y|X=x]$  的定义可知,  $E[Y|X=x]$  是与条件  $X=x$  有关的量。若以随机变量  $X$  替换给定值  $x$ , 则称  $E[Y|X=x]$  为随机变量  $Y$  关于条件  $X$  的条件数学期望。

对于随机变量  $X$  的所有取值而言, 条件期望  $E[Y|X]$  是定义在样本空间  $\Omega_X$  上的函数。 $E[Y|X=x]$  结果取决于  $X$  的值, 而  $E[Y|X]$  则是随机变量  $X$  的函数, 也是个随机变量。所以, 条件期望  $E[Y|X]$  有如下的性质:

$$1^\circ E_X\{E[Y|X]\} = E[Y].$$

推导过程:

$$\begin{aligned} E_X\{E[Y|X]\} &= \int_{-\infty}^{\infty} E[Y|X=x]f_X(x)dx = \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} yf_Y(y|x)dy \right] f_X(x)dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} yf_Y(y|x)f_X(x)dydx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} yf_{XY}(x,y)dydx = E[Y]. \end{aligned} \quad (1.178)$$

即条件期望的期望等于非条件期望。

同理可得其他性质:

$$2^\circ E\{E[g(X,Y)|X]\} = E[g(X,Y)].$$

$$3^\circ E[g(X)Y|X] = g(X)E[Y|X].$$

4° 当随机变量  $X$  和  $Y$  相互独立时,  $E[Y|X] = E[Y]$ , 且  $E[C|X] = E[C] = C$ ,  $C$  为常数(常数与一切随机变量独立)。

**例 1.61** 已知随机变量  $X$  服从  $(0, 1)$  的均匀分布, 随机变量  $Y$  服从  $(X, 1)$  的均匀分布。求① 条件数学期望  $E[Y|X=x]$ 。② 条件数学期望  $E[Y|X]$ 。

解: ① 根据已知条件, 在给定条件  $\{X=x\}$  下, 随机变量  $Y$  的概率密度的表达式为

$$f_Y(y|X=x) = \begin{cases} \frac{1}{1-x}, & x < y < 1 \\ 0, & \text{其他} \end{cases} \quad (1.179)$$

## 条件数学期望

$$E[Y|X=x] = \int_{-\infty}^{\infty} y f_Y(y|x) dy = \int_x^1 \frac{y}{1-x} dy = \frac{1+x}{2}. \quad (1.180)$$

由上可以看出条件数学期望  $E[Y|X=x]$  是关于给定值  $x$  的函数。

② 用随机变量  $X$  替换给定值  $x$ , 则数学期望  $E[Y|X] = \frac{1+X}{2}$  是随机变量  $X$  的函数, 也是个随机变量。因为随机变量  $X$  服从  $(0, 1)$  的均匀分布, 函数  $1+X$  服从  $(1, 2)$  的均匀分布, 则其函数  $\frac{1+X}{2}$  也服从  $(\frac{1}{2}, 1)$  的均匀分布, 即  $E[Y|X]$  服从  $(\frac{1}{2}, 1)$  的均匀分布,  $E[Y|X] \sim U(\frac{1}{2}, 1)$ .

## 1.1.7 随机变量的方差和矩

### 定义 1.62 方差

变量  $X$  的方差定义为

$$\begin{aligned} D(X) &= E\{[X - E\{X\}]^2\} \\ &= \int_{-\infty}^{+\infty} (x - E\{X\})^2 f(x) dx = E\{X^2\} - (E\{X\})^2. \end{aligned} \quad (1.181)$$

方差的大小反映了随机变量相对于其均值的离散程度。随机变量  $X$  的方差是对“ $X$  的取值与其期望  $m_X$  差值的平方”计算平均。通常称  $\sqrt{D(X)}$  为随机变量  $X$  的均方差或标准差, 习惯上用  $\sigma_X^2$  表示  $D(X)$ .



方差具有如下一些基本的性质:

- (1) 对于任意随机变量  $X$ , 有  $D[X] \geq 0$ , 且当  $x = C$  ( $C$  常数) 时,  $D[X] = 0$ .
- (2) 对于任意实数  $C$ , 有

$$D[CX] = C^2 D[X]. \quad (1.182)$$

- (3) 对于常数  $a_i$  ( $i = 1, 2, \dots, n$ ), 随机变量  $X_1, \dots, X_n$  的线性组合  $\sum_{i=1}^n a_i X_i$  的方差为

$$\text{Var}\left\{\sum_{i=1}^n a_i X_i\right\} = \sum_{i=1}^n \sum_{j=1}^n a_i a_j E\{(X_i - E\{X_i\})(X_j - E\{X_j\})\}. \quad (1.183)$$

- (4)  $X_i$  ( $i = 1, 2, \dots, n$ ) 独立, 则

$$\text{Var}\left\{\sum_{i=1}^n a_i X_i\right\} = \sum_{i=1}^n a_i^2 D[X_i]. \quad (1.184)$$

(5) 对于一切实数  $\mu$ , 有

$$E[(X - \mu)^2] \geq E[(X - m_X)^2] = D[X]. \quad (1.185)$$

证:

$$\begin{aligned} E[(X - \mu)^2] &= E\{(X - m_X) + (m_X - \mu)\}^2 \\ &= E[(X - m_X)^2 + 2(X - m_X)(m_X - \mu) + (m_X - \mu)^2] \\ &= E[(X - m_X)^2] + 2E[(X - m_X)(m_X - \mu)] + E[(m_X - \mu)^2] \quad (1.186) \\ &= E[(X - m_X)^2] + 2(m_X - \mu)E[X - m_X] + (m_X - \mu)^2 \\ &= E[(X - m_X)^2] + (m_X - \mu)^2 \geq E[(X - m_X)^2]. \end{aligned}$$

这表明, 随机变量的取值相对于任何其他  $\mu$  值的分散程度不小于其相对于期望的分散程度。可知如下两个结论:

- ① 此统计平均的结果是个大于等于 0 的数。
- ② 方差是用来表征随机变量取值相对数学期望的分散程度。

### 例 1.1.8

高斯变量  $X$  的概率密度为

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x - m_X)^2}{2\sigma^2}\right\}. \quad (1.187)$$

可求出其数学期望为  $m_X$ , 方差为  $\sigma^2$ 。显然, 概率密度  $f_X(x)$  的值与均方差  $\sigma$  的大小成反比。又由概率的性质可知, 无论  $\sigma^2$  是多少,  $f_X(x)$  曲线下的面积都是相同的, 即

$$\int_{-\infty}^{\infty} f(x)dx = 1. \quad (1.188)$$

在  $\sigma$  大小不同的情况下,  $f_X(x)$  的图形如图 1-19 所示。从图中可以看出, 当  $\sigma$  小时,  $X$  的分布相对数学期望  $m$  较集中; 当  $\sigma$  大时,  $X$  的分布相对数学期望  $m$  较分散。由此进一步说明, 方差  $\sigma^2$  的大小正比于随机变量的取值相对于其期望分散程度的大小。

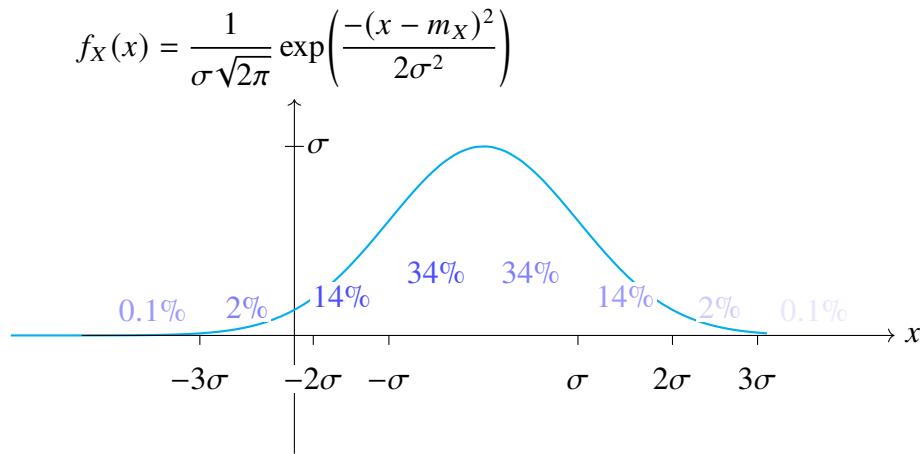


图 1-19 高斯变量的概率密度分布

## 1. 随机变量的矩

### (1) 一维随机变量的矩

随机变量的矩有两类; 原点矩和中心矩

#### 1) $k$ 阶原点矩

##### 定义 1.63 随机变量 $X$ 的 $k$ 阶原点矩

对随机变量  $X$ , 求  $X$  的  $k$  次幂的统计平均

$$E[X^k] = \int_{-\infty}^{\infty} x^k f_X(x) dx. \quad (1.189)$$

称为随机变量  $X$  的  $k$  阶原点矩。



对随机变量  $X$  的  $k$  阶原点矩的讨论如下:

①  $k = 0$  时,  $E[1] = 1$ .

②  $k = 1$  时,  $E[X]$  为随机变量  $X$  的数学期望。

③  $k = 2$  时,  $E[X^2]$  为二阶原点矩, 又称为随机变量  $X$  的均方值。有时还用到  $E[|X|^k]$ , 被称作  $k$  阶绝对原点矩, 即

$$E[|X|^k] = \int_{-\infty}^{\infty} |x|^k f_X(x) dx. \quad (1.190)$$

#### 2) $k$ 阶中心矩

随机变量  $X$  相对于其数学期望  $m_X$  的差  $(X - m_X)$  的  $k$  次幂求统计平均

$$E[(X - m_X)^k] = \int_{-\infty}^{\infty} (x - m_X)^k f_X(x) dx. \quad (1.191)$$

称为随机变量  $X$  的  $k$  阶中心原点矩。讨论如下情况：

$$\textcircled{1} k=0 \text{ 时, } E[(X-m_X)^0]=1;$$

$$\textcircled{2} k=1 \text{ 时, } E[X-m_X]=E[X]-m_X=0;$$

\textcircled{3}  $k=2$  时,  $E[(X-m_X)^2]$  为二阶中心矩, 它是中心矩中最重要也是最常用的一种矩, 通常称为随机变量  $X$  的方差, 用  $D[X]$  或  $\sigma_X^2$  表示。 $D[X]=\sigma_X^2=E[(X-m_X)^2]$ ,  $D[X]$  的正平方根  $\sigma_X$  称为随机变量  $X$  的均方差或标准偏差.

有时还常用到  $E[|X-m_X|^k]$ , 被称作  $k$  阶绝对中心矩, 即

$$E[|X-m_X|^k]=\int_{-\infty}^{\infty}|x-m_X|^k f_X(x)dx. \quad (1.192)$$

### 3) 矩存在的条件

随机变量的各阶矩不是都存在的。

**例 1.64** 若随机变量的各阶绝对矩都存在, 则它相应的各阶矩都存在。但有的随机变量则不满足“各阶绝对矩都存在”的条件, 如柯西分布(图 1-20), 它的概率密度为

$$f_X(x)=\frac{1}{\pi(\sigma^2+\nu_2^2(x-\mu)^2)}, \quad (1.193)$$

其一阶绝对矩  $\int_{-\infty}^{\infty}|x|f_X(x)dx$  是发散的, 所以它的数学期望就不存在。

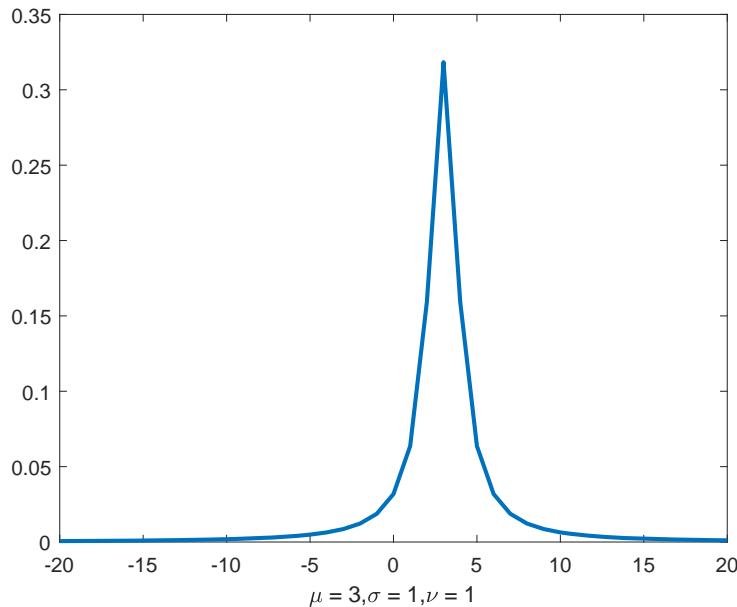


图 1-20 柯西分布

若  $(X_1, X_2, \dots, X_n)$  两两互不相关, 则

$$D[X_1 \pm X_2 \pm \dots \pm X_n] = D[X_1] + D[X_2] + \dots + D[X_n]. \quad (1.194)$$

证: 设随机变量  $Y = X_1 \pm X_2 \pm \dots \pm X_n$ , 由方差的定义  $D[Y] = E\{(Y - E(Y))^2\}$ , 可得

$$\begin{aligned} D[X_1 \pm X_2 \pm \dots \pm X_n] &= E\{(X_1 \pm X_2 \pm \dots \pm X_n) - E(X_1 \pm X_2 \pm \dots \pm X_n)\}^2 \\ &= E\{(X_1 - m_1) \pm (X_2 - m_2) \pm \dots \pm (X_n - m_n)\}^2 \\ &= E\left\{\sum_{i=1}^n (X_i - m_i)^2 \pm 2 \sum_{i < j} (X_i - m_i)(X_j - m_j)\right\} \\ &= \sum_{i=1}^n E[(X_i - m_i)^2] \pm 2 \sum_{i < j} E\{(X_i - m_i)(X_j - m_j)\}. \end{aligned} \quad (1.195)$$

对于所有  $i \neq j$ , 随机变量  $X_i$  和  $X_j$  互不相关。因此根据数学期望的性质 4°, 有

$$\begin{aligned} E[(X_i - m_i)(X_j - m_j)] &= E[X_i - m_i]E[X_j - m_j] \\ &= \{E[X_i] - m_i\}\{E[X_j] - m_j\} = 0. \end{aligned} \quad (1.196)$$

所以

$$D[X_1 \pm X_2 \pm \dots \pm X_n] = \sum_{i=1}^n E[(X_i - m_i)^2] = \sum_{i=1}^n D[X_i]. \quad (1.197)$$

(2)  $n$  维随机变量的矩

1) 联合原点矩

已知二维随机变量  $(X, Y)$  的联合概率密度函数为  $f_{XY}(x, y)$ , 定义二维随机变量  $(X, Y)$  的  $n+k$  阶联合原点矩为

$$E[X^n Y^k] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^n y^k f_{XY}(x, y) dx dy. \quad (1.198)$$

显然:

①  $k = 0$  时,  $E[X^n]$  是随机变量  $X$  的  $n$  阶原点矩。

②  $n = 0$  时,  $E[Y^k]$  是随机变量  $Y$  的  $k$  阶原点矩。

③  $E[XY]$  是联合原点矩中最重要的一个, 它反映了  $X$  与  $Y$  两个随机变量间的关联程度, 称之为随机变量  $X$  和  $Y$  的互相关, 通常用  $R_{XY}$  表示, 如下所示:

$$E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{XY}(x, y) dx dy = R_{XY}. \quad (1.199)$$

已知  $n$  维随机变量  $(X_1, X_2, \dots, X_n)$  的联合概率密度函数为  $f_X(x_1, x_2, \dots, x_n)$ , 定义  $n$  维随机变量的  $(k_1 + k_2 + \dots + k_n) \triangleq k$  阶联合原点矩为

$$E[X_1^{k_1} X_2^{k_2} \cdots X_n^k] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1^{k_1} x_2^{k_2} \cdots x_n^k f_X(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n. \quad (1.200)$$

其中  $k_1, k_2, \dots, k_n$  为正整数。

## 2) 联合中心矩

### 定义 1.65 二维随机变量 $(X, Y)$ 的 $n+k$ 阶联合中心矩

二维随机变量  $(X, Y)$  的  $n+k$  阶联合中心矩为

$$E[(X - m_X)^n (Y - m_Y)^k] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_X)^n (y - m_Y)^k f_{XY}(x, y) dx dy. \quad (1.201)$$

显然:

①  $E[(X - m_X)^2] = \sigma_X^2$  是随机变量  $X$  的二阶中心矩, 即  $X$  的方差。

②  $E[(Y - m_Y)^2] = \sigma_Y^2$  是随机变量  $Y$  的二阶中心矩, 即  $Y$  的方差。

③  $E[(X - m_X)(Y - m_Y)]$  是联合中心矩中最重要的一个, 它也反映了  $X$  与  $Y$  两个随机变量间的关联程度, 称之为  $X$  和  $Y$  的协方差, 通常用  $C_{XY}$  或  $\text{Cov}(X, Y)$  表示如下:

### 定义 1.66 联合中心矩

对于  $n$  维随机变量  $(X_1, X_2, \dots, X_n)$ , 若其联合概率密度函数为  $f_X(x_1, x_2, \dots, x_n)$ 。

定义  $n$  维随机变量的  $(k_1 + k_2 + \dots + k_n)$  阶联合中心矩为

$$\begin{aligned} E[(X - m_X)(Y - m_Y)] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_X)(y - m_Y) f_{XY}(x, y) dx dy \\ &= C_{XY}. \end{aligned} \quad (1.202)$$

其中  $k_1, k_2, \dots, k_n$  为正整数。

**例 1.67** 已知二维随机变量  $(X, Y)$  的联合概率密度为

$$f_{XY}(x, y) = \begin{cases} 1, & 0 < x < 1, 0 < y < 1 \\ 0, & \text{其他} \end{cases} \quad (1.203)$$

求: ① 随机变量  $X$  的一、二阶原点矩。② 随机变量  $X$  的二、三阶中心矩。③ 联合原点矩  $R_{XY}$ 。④ 联合中心矩  $C_{XY}$ 。

解：由二维随机变量的联合概率密度，可得其边缘概率密度

$$\begin{aligned} f_X(x) &= \begin{cases} 1, & 0 < x < 1, \\ 0, & \text{其他} \end{cases} \quad f_Y(y) = \begin{cases} 1, & 0 < y < 1 \\ 0, & \text{其他} \end{cases} \\ E[X] &= \frac{1}{2}, E[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_0^1 x^2 dx = \frac{1}{3}. \\ E[(X - m_X)^2] &= \int_0^1 \left(x - \frac{1}{2}\right)^2 dx = \frac{1}{12}. \\ E[(X - m_X)^3] &= \int_0^1 \left(x - \frac{1}{2}\right)^3 dx = \frac{1}{32}. \\ R_{XY} = E[XY] &= \int_0^1 \int_0^1 xy dx dy = \frac{1}{4}. \\ C_{XY} = E[(X - m_X)(Y - m_Y)] &= \int_0^1 \int_0^1 \left(x - \frac{1}{2}\right) \left(y - \frac{1}{2}\right) dx dy = 0. \end{aligned}$$

## 2. 方差

### (1) 二维随机变量的协方差

#### 定义 1.68 二阶联合中心矩

若二维随机变量  $(X, Y)$  中关于  $X$  和  $Y$  的数学期望和方差均存在，则称

$$E[(X - m_X)(Y - m_Y)] = C_{XY}. \quad (1.204)$$

为随机变量  $X$  与  $Y$  的协方差，又称为“相关矩”或“二阶联合中心矩”。 ♣

二阶联合中心矩通常被用来表征两个随机变量间的关联程度。当  $C_{XY} = 0$  时，称  $X$  与  $Y$  互不相关。协方差也可以通过下式进行计算：

$$C_{XY} = E[(X - m_X)(Y - m_Y)] = E[XY] - m_X m_Y = R_{XY} - m_X m_Y. \quad (1.205)$$

当  $X$  与  $Y$  互不相关时，可知  $C_{XY} = 0$ ，由上式可得

$$E[XY] = E[X]E[Y]. \quad (1.206)$$

## 3. 随机矢量的方差

**定义 1.69 随机矢量  $X$  的方差**

若将  $n$  维随机变量  $(X_1, X_2, \dots, X_n)$  用随机矢量  $\mathbf{X} = [X_1, X_2, \dots, X_n]^T \in \mathbb{R}^{n \times 1}$  表示, 则随机矢量  $\mathbf{X}$  的方差定义为

$$D(\mathbf{X}) = E[(\mathbf{X} - E\mathbf{X})(\mathbf{X} - E\mathbf{X})^T]. \quad (1.207)$$

记作  $D(\mathbf{X}) = \text{Var}\{\mathbf{X}\}$ 。



若将上式展开, 可得随机矢量的方差阵

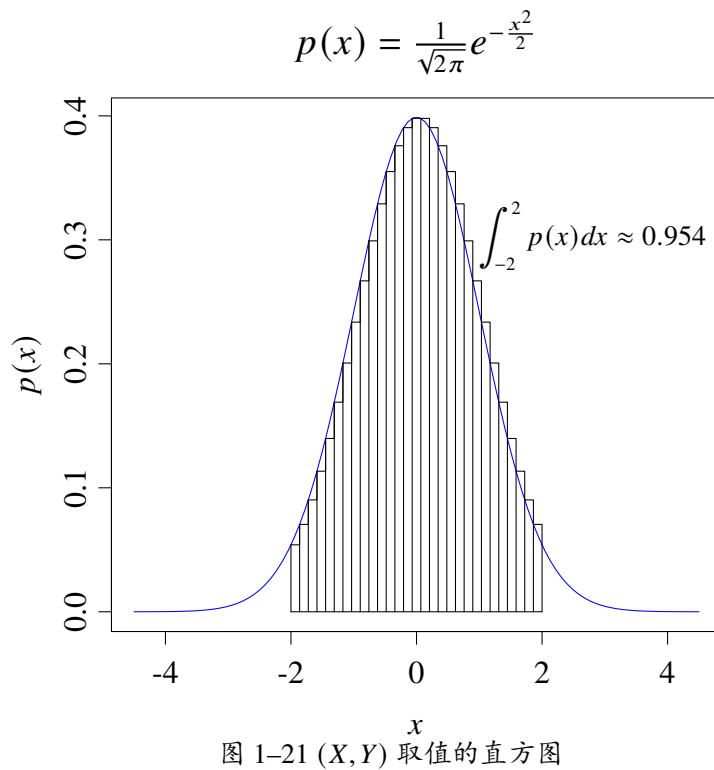
$$\begin{aligned} D(\mathbf{X}) &= E \left\{ \begin{bmatrix} X_1 - m_1 \\ X_2 - m_2 \\ \vdots \\ X_n - m_n \end{bmatrix}_{n \times 1} \times [X_1 - m_1 \quad X_2 - m_2 \quad \cdots \quad X_n - m_n]_{1 \times n} \right\} \\ &= \begin{bmatrix} E[(X_1 - m_1)^2] & E[(X_1 - m_1)(X_2 - m_2)] & \cdots & E[(X_1 - m_1)(X_n - m_n)] \\ E[(X_2 - m_2)(X_1 - m_1)] & E[(X_2 - m_2)^2] & \cdots & E[(X_2 - m_2)(X_n - m_n)] \\ \vdots & \vdots & \ddots & \vdots \\ E[(X_n - m_n)(X_1 - m_1)] & E[(X_n - m_n)(X_2 - m_2)] & \cdots & E[(X_n - m_n)^2] \end{bmatrix}. \end{aligned} \quad (1.208)$$

若  $n$  维变量  $(X_1, X_2, \dots, X_n)$  中每个变量之间的协方差  $C_{ij} = E[(X_i - m_i)(X_j - m_j)]$  (包括自身的方差) 均存在, 则

$$D(\mathbf{X}) = \begin{bmatrix} D[X_1] & C_{12} & \cdots & C_{1n} \\ C_{21} & D[X_2] & \cdots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \cdots & D[X_n] \end{bmatrix} = C_X, \quad (1.209)$$

由式(1.209)可见, 矢量  $\mathbf{X}$  的方差阵中除对角线上均为各分量的方差外, 其余均为分量之间的协方差  $C_{ij}$  ( $i \neq j$ )。若将方差  $D(X_i)$  看成协方差的特例  $C_{ii}$  ( $i = j$ ), 将方差阵中的所有元素均用协方差  $C_{ij}$  表示, 则有

$$D(\mathbf{X}) = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix}_{n \times n} = C_X, \quad (1.210)$$

图 1-21  $(X, Y)$  取值的直方图

又可称为  $n$  维随机变量的协方差矩阵。

由  $C_{ij} = C_{ji}$ , 可得  $C_X = C_X^T$ , 随机矢量  $\mathbf{X}$  的方差阵是对称阵。

### 1.1.8 相关、正交和独立

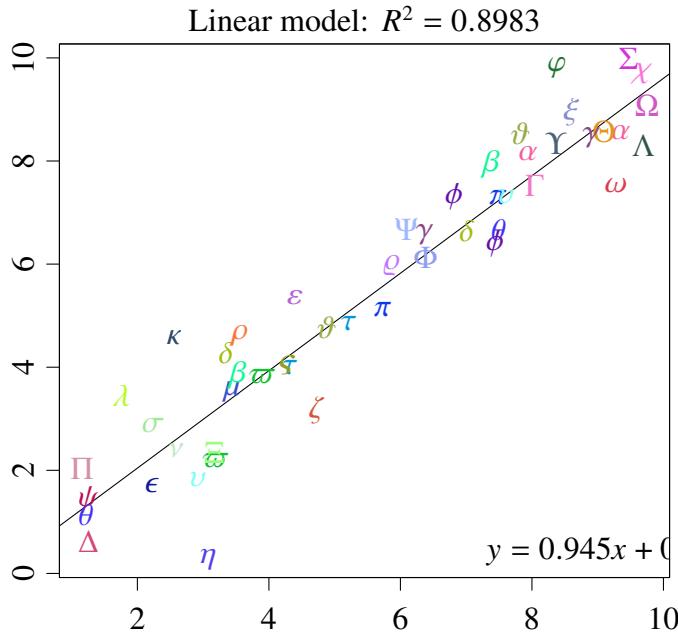
#### 1. 相关系数

##### (1) 两个随机变量相互关系的描述

在实际中, 为了将两个随机变量  $X$  和  $Y$  的相互关系表示出来, 最简单的方法是将两个随机变量的所有取值情况在  $xOy$  的平面上画出来, 称之为随机变量  $X$  和  $Y$  取值的散布图 (图 1-22)。若  $X$  与  $Y$  独立,  $X$  的取值与  $Y$  的取值没有任何关系, 则  $(X, Y)$  在  $xOy$  平面上没有样本点存在。若  $X$  与  $Y$  相关, 对于  $X$  取的每一确定值, 都有  $Y$  的取值与其相对应。因此,  $(X, Y)$  在  $xOy$  平面上有样本点存在。

所以散布图直观反映了  $X$  与  $Y$  之间的相关情况。如果要从数学上表示  $X$  与  $Y$  的函数关系, 则要设法找到逼近其散布点密集分布的一条回归线。如果这条回归线是直线, 我们就说  $X$  与  $Y$  线性相关; 如果这条回归线是曲线, 我们就说  $X$  与  $Y$  非线性相关。

最常用的方法是用一根直线去逼近  $(X, Y)$  取值散布点的密集分布, 即寻找最逼近

图 1-22  $(X, Y)$  取值的散布图

$(X, Y)$  关系的直线(线性回归)。在这条直线上,可由  $X$  的取值预测  $Y$  的取值,即

$$Y_p = a + bX, \quad (1.211)$$

当然  $Y_p$  并不就是  $Y$ ,而是根据  $X$  的取值所得到的线性相关预测值。

如果  $a, b$  构造的是回归直线,那么一定会使得  $Y$  与  $Y_p$  的误差最小。

寻找使均方误差  $\varepsilon = E[(Y - Y_p)^2]$  最小  $a, b$  值的方法,如下

$$\begin{aligned} \varepsilon &= E[(Y - Y_p)^2] = E[(Y - a - bX)^2] \\ &= E[Y^2 + b^2X^2 + a^2 + 2abX - 2aY - 2bXY] \\ &= a^2 - 2aE[Y] + 2abE[X] - 2bE[XY] + b^2E[X^2] + E[Y^2]. \end{aligned} \quad (1.212)$$

令

$$\begin{cases} \frac{\partial \varepsilon}{\partial a} = 2a - 2m_Y + 2bm_X = 0 \\ \frac{\partial \varepsilon}{\partial b} = 2am_X - 2R_{XY} + 2bE[X^2] = 0 \end{cases}. \quad (1.213)$$

$a, b$  的值为

$$\begin{cases} a = m_Y - bm_X = m_Y - \frac{C_{XY}}{\sigma_X^2 m_X} \\ b = \frac{R_{XY} - m_X m_Y}{E[X^2] - m_X^2} = \frac{C_{XY}}{\sigma_X^2} \end{cases}. \quad (1.214)$$

将其代入式(1.212)中, 可验证给出的  $\varepsilon$  是极小值, 从而得到最优预测结果, 直线方程为

$$Y_p = a + bX = m_Y + \frac{C_{XY}}{\sigma_X^2} (X - m_X). \quad (1.215)$$

由式(1.215)可见, 最佳预测线通过  $(m_X, m_Y)$  点。通过这条直线, 可以根据  $X$  的取值, 给出  $Y$  的最佳预测值  $Y$ , 即这条直线最能反映  $X$  与  $Y$  之间的线性关系。

为了简便, 引入  $X$  和  $Y$  的归一化随机变量

$$X = \frac{X - m_X}{\sigma_X}, Y = \frac{Y_p - m_Y}{\sigma_Y}. \quad (1.216)$$

式中斜率  $\frac{C_{XY}}{\sigma_X \sigma_Y}$  的大小反映出随机变量  $X$  与  $Y$  之间线性相关的程度。

### (2) 相关系数的定义

#### 定义 1.70 相关系数

用来表征两个随机变量之间线性相关程度的量

$$\rho_{XY} = \frac{C_{XY}}{\sigma_X \sigma_Y}. \quad (1.217)$$

称之为相关系数。



### (3) 相关系数的性质

1°  $-1 \leq \rho_{XY} \leq 1$ , 即  $|\rho_{XY}| \leq 1$ .

证: 为了使证明简化, 这里研究两个归一化随机变量的相关系数。令

$$U = \frac{X - m_X}{\sigma_X}, V = \frac{Y - m_Y}{\sigma_Y}. \quad (1.218)$$

显然有  $E[U] = E[V] = 0, D[U] = D[V] = 1, E[V^2] = 1$  成立。再由

$$\begin{aligned} E^2[U \pm V] &= \{E[U] \pm E[V]\}^2 = 0. \\ E[(U \pm V)^2] &= E[U^2] \pm 2E[UV] + E[V^2] = 2 \pm 2E[UV]. \\ \rho_{UV} &= \frac{C_{UV}}{\sigma_U \sigma_V} = \frac{R_{UV} - m_U m_V}{\sigma_U \sigma_V} = R_{UV} = E[UV]. \end{aligned} \quad (1.219)$$

把上述三式代入式(1.209), 可得

$$D[U \pm V] = E[(U \pm V)^2] - E^2[U \pm V] \geq 0. \quad (1.220)$$

2° 若  $X$  与  $Y$  间以概率 1 存在线性关系, 即满足  $P\{Y = aX + b\} = 1$  ( $a$  和  $b$  为实常数) 时, 则有  $|\rho_{XY}| = 1$ 。

3° 若  $X$  与  $Y$  线性不相关, 则有  $|\rho_{XY}| = 0$ 。等价于

$$\begin{cases} C_{XY} = 0, \\ R_{XY} = E[X]E[Y], \\ D[X \pm Y] = D[X] + D[Y]. \end{cases} \quad (1.221)$$

注: 通常说的“互不相关”均指的是线性互不相关。

推广到  $n$  维的情况。当  $n$  维随机变量  $X_1, X_2, \dots, X_n$  之间互不相关时, 由于  $C_{ij} = 0 (i \neq j)$ , 则  $n$  维随机变量的协方差矩阵为对角线矩阵, 如下所示

$$C_X = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \sigma_n^2 \end{bmatrix}. \quad (1.222)$$

相关系数矩阵

$$\rho_X = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2n} \\ \vdots & \vdots & & \vdots \\ \rho_{n1} & \rho_{n2} & \cdots & \rho_{nn} \end{bmatrix}. \quad (1.223)$$

当  $X_1, X_2, \dots, X_n$  互不相关时, 由于  $\rho_{ij} = 0 (i \neq j)$  和  $\rho_{ii} = 1 (i = j)$  成立, 则相关系数矩阵为单位矩阵。

## 2. 不相关、独立和正交

### (1) 正交的定义

#### 定义 1.71 $X$ 与 $Y$ 正交

若随机变量  $X$  与  $Y$  满足  $R_{XY} = E[XY] = 0$ , 则称随机变量  $X$  与  $Y$  正交。



考虑到  $C_{XY} = R_{XY} - m_X m_Y$ , 则  $X$  与  $Y$  正交满足下式

$$C_{XY} = -m_X m_Y. \quad (1.224)$$

### (2) 不相关、独立和正交的关系

1° 若随机变量  $X$  与  $Y$  相互独立, 则  $X$  与  $Y$  必定互不相关。

证: 当  $X$  与  $Y$  相互独立时, 满足

$$f_{XY}(x, y) = f_X(x)f_Y(y). \quad (1.225)$$

$$\begin{aligned} C_{XY} &= \int_{-\infty}^{\infty} (x - m_X) f_X(x) dx \cdot \int_{-\infty}^{\infty} (y - m_Y) f_Y(y) dy \\ &= E[(X - m_X)] \cdot E[(Y - m_Y)] = 0. \end{aligned} \quad (1.226)$$

即  $X$  与  $Y$  互不相关。

2° 若随机变量  $X$  与  $Y$  互不相关 ( $\rho_{XY} = 0$ ), 则  $X$  与  $Y$  不一定相互独立。

因为独立表示两个随机变量之间既线性不相关, 又非线性不相关。而相关系数  $\rho_{XY} = 0$  仅仅表征两个随机变量线性不相关, 而不能说明它们之间非线性也不相关。

**例 1.72** 已知随机变量  $X = \cos \varphi$  和  $Y = \sin \varphi$ , 式中  $\varphi$  是在  $(0, 2\pi)$  上均匀分布, 讨论  $X$  和  $Y$  的相关性及独立性。

解:

① 因为  $\varphi$  是在  $(0, 2\pi)$  上均匀分布的随机变量, 其概率密度为

$$f(\varphi) = \begin{cases} \frac{1}{2\pi}, & 0 < \varphi < 2\pi \\ 0, & \text{其他} \end{cases}. \quad (1.227)$$

则

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} \cos \varphi f(\varphi) d\varphi = \frac{1}{2\pi} \int_0^{2\pi} \cos \varphi d\varphi = 0. \\ E[Y] &= \int_{-\infty}^{\infty} \sin \varphi f(\varphi) d\varphi = \frac{1}{2\pi} \int_0^{2\pi} \sin \varphi d\varphi = 0. \\ E[XY] &= E[\cos \varphi \sin \varphi] = \frac{1}{2} E[\sin 2\varphi] = 0. \\ C_{XY} &= E[(X - m_X)(Y - m_Y)] = E[XY] = 0. \end{aligned}$$

即相关系数  $\rho_{XY} = 0$ , 表明  $X$  与  $Y$  是互不相关的。

② 由于随机变量  $X$  与  $Y$  存在  $X^2 + Y^2 = 1$  的关系 (非线性相关),  $Y$  的取值依赖于  $X$  的取值, 所以随机变量  $X$  与  $Y$  之间相互不独立。

3° 若两个随机变量的联合矩对任意  $n \geq 1$  和  $k \geq 1$  均可分解为

$$E[X^n Y^k] = E[X^n] \cdot E[Y^k], \quad (1.228)$$

则  $X$  与  $Y$  统计独立。

4° 当随机变量  $X$  与  $Y$  之间存在线性函数关系  $Y = aX + b$  时, 则有  $\rho_{XY} = \pm 1$ ; 当  $X$  与  $Y$  间存在非线性函数关系时, 则有  $0 < |\rho_{XY}| < 1$ 。

**例 1.73** 已知随机变量  $X$  与  $Y$  有非线性关系  $Y = X^3$ , 且

$$E(X^n) = \begin{cases} 1 \cdot 3 \cdot 5 \cdots (n-1)\sigma_X^n, & n \text{ 为偶数} \\ 0, & n \text{ 为奇数} \end{cases}, n \geq 2. \quad (1.229)$$

求相关系数  $\rho_{XY}$ 。

解: 由于

$$m_X = E[X] = 0, m_Y = E[X^3] = 0. \quad (1.230)$$

则协方差为

$$C_{XY} = R_{XY} - m_X m_Y = R_{XY} = E[XY] = E[XX^3] = E[X^4] = 3\sigma_X^4.$$

相关系数为

$$\rho_{XY} = \frac{3\sigma_X^4}{\sigma_X \sigma_Y} = \frac{3\sigma_X^4}{\sqrt{E[X^2] \cdot E[X^6]}} = \frac{3\sigma_X'}{\sqrt{15\sigma_X^8}} = \frac{3}{\sqrt{15}} = 0.775 < 1. \quad (1.231)$$

5° 针对一般情况而言, 两个随机变量正交不能保证它们不相关; 反之, 两个变量不相关也不能保证正交。由  $C_{XY} = R_{XY} - m_X m_Y$  可以看出, 若  $m_X = 0$  或  $m_Y = 0$  其中一个成立, 则不相关和正交等价。

**例 1.74** 已知随机变量  $X$  的均值  $m_X = 3$ , 方差  $\sigma_X^2 = 2$ , 且另一随机变量  $Y = -6X + 22$ 。讨论  $X$  与  $Y$  的相关性与正交性。

解: 由题可知  $X$  的均值  $m_X = 3$ , 所以  $Y$  的均值为

$$m_Y = E[-6X + 22] = -6E[X] + 22 = 4. \quad (1.232)$$

①  $X$  与  $Y$  的互相关为

$$\begin{aligned} R_{XY} &= E[XY] = E[-6X^2 + 22X] = -6E[X^2] + 22E[X] \\ &= -6(m_X^2 + \sigma_X^2) + 22m_X = 0. \end{aligned} \quad (1.233)$$

可见  $X$  与  $Y$  是正交的。

由于  $\sigma_X^2 = 2$  且  $\sigma_Y^2 = E[(Y - m_Y)^2] = E[(-6X + 18)^2] = 72$ , 所以相关系数为

$$\rho_{XY} = \frac{R_{XY} - m_X m_Y}{\sqrt{\sigma_X^2 \sigma_Y^2}} = \frac{-12}{\sqrt{144}} = -1. \quad (1.234)$$

这说明  $X$  与  $Y$  又是线性相关的。

### 1.1.9 随机变量的特征函数

随机变量的数学期望和方差等数字特征虽可以反映其概率分布的某些特征,但一般无法通过它们来确定其概率分布。而随机变量的特征函数既可以确定其概率分布又具有良好的分析性质,它是概率论中最重要的分析工具之一。它的优越性,首先在于特征函数与概率密度唯一对应。其次,由特征函数计算统计特性比由概率密度函数计算更为方便。

**例 1.75** 由概率密度计算矩需要积分运算,而由特征函数计算矩则只要微分运算;由概率密度求独立随机变量和的分布,需要各个概率密度进行卷积,而由特征函数求独立随机变量和的特征函数,只需要各个特征函数相乘即可。

这些优点将会在下面的介绍中逐一呈现。

#### 1. 一元特征函数及其性质

##### 定义 1.76 随机变量 $X$ 的特征函数

设  $X$  是定义在概率空间  $(\Omega, F, P)$  上的随机变量,其概率密度函数为  $f(x)$ ,则其函数  $e^{jux}$  的数学期望为

$$Q_X(u) = E[e^{jux}] , j = \sqrt{-1}, -\infty < u < \infty. \quad (1.235)$$

称之为随机变量  $X$  的特征函数。



① 当  $X$  为连续型随机变量时,其特征函数为

$$Q_X(u) = \int_{-\infty}^{\infty} e^{jux} f(x) dx. \quad (1.236)$$

② 当  $X$  为离散型随机变量时,其特征函数为

$$Q_X(u) = \sum_i e^{jux_i} P(X = x_i) = \sum_i e^{jux_i} p_i. \quad (1.237)$$

#### (2) 性质

1°  $|Q_X(u)| \leq Q_X(0) = 1$ 。

证: 由特征函数的定义

$$|Q_X(u)| = \left| \int_{-\infty}^{\infty} e^{jux} f(x) dx \right| \leq \int_{-\infty}^{\infty} |e^{jux} f(x)| dx. \quad (1.238)$$

由于  $f(x) \geq 0$ , 且  $|e^{jux}| = 1$ , 可得

$$|Q_X(u)| \leq \left| \int_{-\infty}^{\infty} e^{jux} f(x) dx \right| = \int_{-\infty}^{\infty} |e^{jux} f(x)| dx = Q_X(0) = 1. \quad (1.239)$$

由于  $|Q_X(u)| \leq 1$ , 说明特征函数  $Q_X(u)$  在一切实数  $u$  上都有定义。

2° 特征函数  $Q_X(u)$  是实变量  $u$  在  $(-\infty, +\infty)$  上的连续函数。

3° 特征函数  $Q_X(u)$  是实变量  $u$  的复函数, 有

$$Q_X^*(u) = Q_X(-u). \quad (1.240)$$

4° 随机变量  $X$  的函数  $y = aX + b$  的特征函数

$$Q_Y(u) = e^{iub} Q_X(au). \quad (1.241)$$

证:

$$\begin{aligned} Q_Y(u) &= E[e^{iuY}] = E[e^{iu(aX+b)}] = E[e^{iuaX} \cdot e^{iub}] \\ &= e^{iub} E[e^{iuaX}] = e^{iub} Q_X(au). \end{aligned}$$

5° 相互独立随机变量和的特征函数等于它们特征函数的积。

若随机变量  $X_1, \dots, X_n$  相互独立, 且各自相应的特征函数为  $Q_{X_1}(u), Q_{X_2}(u), \dots, Q_{X_n}(u)$ .

随机变量  $Y = \sum_{k=1}^n X_i$ , 则其特征函数为

$$Q_Y(u) = \prod_{i=1}^n Q_{X_i}(u). \quad (1.242)$$

证

$$Q_Y(u) = E[e^{iuY}] = E[e^{iu(X_1+X_2+\dots+X_n)}] = E\left[\prod_{i=1}^n e^{iux_i}\right]. \quad (1.243)$$

由于  $X_1, \dots, X_n$  相互独立, 根据数学期望的性质 5°, 可得

$$Q_Y(u) = E[e^{iuY}] = E[e^{iu(X_1+X_2+\dots+X_n)}] = \prod_{i=1}^n Q_{X_i}(u). \quad (1.244)$$

6° 若随机变量  $X$  的  $n$  阶绝对矩存在, 则它的特征函数存在  $n$  阶导数。并且当  $1 \leq k \leq n$  时, 有

$$E[X^k] = (-j)^k \left| \frac{d^k Q_X(u)}{du^k} \right|_{u=0}. \quad (1.245)$$

可见, 利用这一性质, 由特征函数  $Q_X(u)$ , 求随机变量  $X$  的  $k$  阶矩, 只要对特征函数  $Q_X(u)$  求  $k$  次导数 (微分) 即可, 显然这比由概率密度  $f(x)$  求  $k$  阶矩的积分运算要简单得多。

7° 若随机变量  $X$  的特征函数  $Q_X(u)$  可以展开成麦克劳林级数, 其特征函数可由该随机变量  $X$  的各阶矩唯一地确定, 即

$$Q_X(u) = \sum_{n=0}^{\infty} E[X^n] \frac{(ju)^n}{n!}. \quad (1.246)$$

这个性质常用在理论推导中。再实际应用中, 可由随机变量的各阶矩  $E(X^n)$  ( $n = 1, 2, \dots$ ) 来求  $Q_X(u)$  不现实, 特别是某些随机变量的各阶矩并不一定都存在, 如柯西分布。

**例 1.77** 随机变量  $X \sim U[a, b]$ ,  $f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0 & \text{其他} \end{cases}$ .

解:

$$\begin{aligned} Q_X(u) &= \int_{-\infty}^{\infty} e^{jux} f(x) dx = \int_a^b e^{jux} \frac{1}{b-a} dx \\ &= \frac{1}{(b-a)ju} e^{jux} \Big|_a^b \\ &= \frac{1}{(b-a)ju} [e^{jub} - e^{jua}] \\ &= \frac{e^{jua}}{(b-a)ju} [e^{ju(b-a)} - 1]. \end{aligned} \quad (1.247)$$

**例 1.78** 随机变量  $X \sim B(1, p)$ , 求特征函数  $Q_X(u)$ .

解:

$$Q_X(u) = e^{ju \times 0} q + e^{ju \times 1} p = q + p(\cos(u) + j \sin(u)). \quad (1.248)$$

**例 1.79** 求二项分布的数学期望、方差和特征函数。

解: 方法一:

二项分布的分布律为

$$P\{Y = k\} = C_n^k p^k q^{n-k}, k = 0, 1, 2, \dots, n. \quad (1.249)$$

由数学期望的定义, 可得

$$E[Y] = \sum_{k=1}^n k P\{Y = k\} = \sum_{k=1}^n k C_n^k p^k (1-p)^{n-k} = \dots = np. \quad (1.250)$$

由方差的定义, 可得

$$\begin{aligned} D[Y] &= \sum_{k=1}^n (k - m_Y)^2 \cdot P\{Y = k\} \\ &= \sum_{k=1}^n (k - np)^2 \cdot C_n^k p^k (1-p)^{n-k} = \dots = npq. \end{aligned} \quad (1.251)$$

直接引用特征函数的定义, 可得

$$Q_Y(u) = \sum_{k=1}^n e^{ju\omega k} P\{Y = k\} = \sum_{k=1}^n e^{ju\omega k} C_n^k p^k q^{n-k} = \dots \quad (1.252)$$

计算较为繁杂。

方法二: 二项分布的随机变量  $Y$  代表的是  $n$  重伯努利试验中随机事件发生的次数。而在每次试验中, 随机事件发生的概率为  $p$ , 随机事件不发生的概率为  $1 - p = q$ , 每次试验中事件发生的次数  $x$  均服从  $(0, 1)$  分布。 $(0, 1)$  分布的特征函数为

$$Q_{X_i}(u) = E[e^{juX_i}] = \sum_{k=0}^1 e^{ju^k} P(X_i = k) = q + pe^{ju}. \quad (1.253)$$

由于  $Y$  代表  $n$  重伯努利试验中随机事件发生的次数, 所以有  $Y = \sum_{i=1}^n X_i$ 。且  $n$  重伯努利试验中各次试验均相互独立, 由性质 5° 可得二项分布的特征函数

$$Q_Y(u) = \prod_{i=1}^n Q_{X_i}(u) = (q + pe^{ju})^n. \quad (1.254)$$

可见, 离散型随机变量的特征函数也是实数的连续函数。利用性质 6° 得

$$\begin{aligned} E[Y] &= -j \frac{d}{du} (pe^{ju} + q)|_{u=0} = np. \\ E[Y^2] &= (-j)^2 \frac{d^2}{du^2} (pe^{ju} + q)|_{u=0} = npq + n^2 p^2. \\ D[Y] &= E[Y^2] - (E[Y])^2 = npq. \end{aligned} \quad (1.255)$$

**例 1.80** 求下列高斯变量的特征函数: ① 标准高斯变量  $Y \sim N(0, 1)$ 。② 高斯变量  $X \sim N(m, \sigma^2)$ 。

解: 先介绍一个从复变函数中得到的积分公式

$$\int_{-\infty}^{\infty} e^{-Ax^2+2Bx-C} dx = \sqrt{\frac{\pi}{A}} e^{-\frac{(AC-B^2)}{A}}. \quad (1.256)$$

① 根据特征函数的定义

$$Q_Y(u) = E[e^{juY}] = \int_{-\infty}^{\infty} e^{juy} \cdot \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2} + 2\frac{ju}{2}} dy. \quad (1.257)$$

令  $A = \frac{1}{2}$ ,  $B = \frac{ju}{2}$ ,  $C = 0$ , 则利用积分公式, 可得标准高斯变量的特征函数如下

$$Q_Y(u) = \frac{1}{\sqrt{2\pi}} \left\{ \sqrt{\frac{\pi}{1/2}} \cdot \exp \left[ \left( \frac{ju}{2} \right)^2 / \frac{1}{2} \right] \right\} = e^{-\frac{u^2}{2}}. \quad (1.258)$$

② 将高斯变量  $X$  归一化处理得

$$X' = \frac{X - m_X}{\sigma_X} = Y, X = \sigma_X Y + m_X. \quad (1.259)$$

利用性质 4 及 ① 的结论 ( $X = \sigma_X Y + m_X$ ), 得一般高斯变量的特征函数如下

$$Q_X(u) = e^{imu} Q_Y(\sigma u) = \exp\left(jum - \frac{1}{2}\sigma^2 u^2\right). \quad (1.260)$$

**例 1.81** 随机变量  $X \sim N(\theta, 1)$ , 求  $X$  的特征函数  $Q_X(u)$ .

因为  $X \sim N(0, 1)$ , 所以  $X$  的概率密度函数为

$$\begin{aligned} p(x, \theta) dx &= (2\pi)^{-1/2} \exp\left\{-\frac{x^2}{2} + x\theta - \frac{\theta^2}{2}\right\} dx \\ &= (2\pi)^{-1/2} e^{-\theta^2/2} e^{\theta x} e^{-x^2/2} dx \doteq C(\theta) e^{\theta x} d\mu(x), \end{aligned} \quad (1.261)$$

其中  $C(\theta) = (2\pi)^{-1/2} e^{-\theta^2/2}$ ,  $d\mu(x) = e^{-x^2/2} dx$ .

易见, 正态分布族  $N(\theta, 1)$  满足指类型分布族  $p_\theta(x) d\mu(x) = C(\theta) e^{\theta T(x)} d\mu(x)$ , 其中  $\theta \in \Theta = \{\theta : \theta = a + bi, a, b \in \mathbb{R}\}$ ,  $T(x)$  为有限的  $B_S$  可测的实函数 [陈艳 2013 指类型分布族的特征函数]. 由定理知,  $T(X) = X$  的特征函数为

$$Q_X(u) = \frac{C(\theta)}{C(\theta + iu)} = \exp\left\{iu\theta - \frac{u^2}{2}\right\}. \quad (1.262)$$

**例 1.82** 随机变量  $X$  服从伽马分布, 即  $X \sim G(\alpha, \beta)$ , 则  $X$  的概率密度函数

$$p(x, \alpha, \beta) dx = \frac{\alpha^\beta}{\Gamma(\beta)} x^{\beta-1} e^{-\alpha x} dx, x > 0, \quad (1.263)$$

$$\Gamma(\beta) = \int_0^\infty x^{\beta-1} e^{-x} dx, \quad (1.264)$$

其中  $\alpha > 0, \beta > 0$  为参数, 则  $X$  的特征函数为  $Q_X(u) = (1 - iu/\alpha)^{-\beta}$ .

① 当  $\beta = 1$  时,  $G(\alpha, 1)$  为指数分布  $E(\alpha)$ , 其特征函数为  $Q_X(u) = (1 - iu/\alpha)^{-1}$ ;

② 当  $\alpha = 1/2, \beta = n/2$  时,  $G(1/2, n/2)$  即为自由度  $n$  的中心化  $\chi^2$  分布  $\chi^2(n)$ , 其特征函数为  $Q_X(u) = (1 - 2iu)^{-n/2}$ . 这里, 可以顺带得到了指数分布和中心化  $\chi^2$  分布的特征函数.

**例 1.83** 随机变量  $X \sim E(\lambda)$ , 求特征函数  $Q_X(u)$ .

解:

$$\begin{aligned} Q_X(u) &= \int_{-\infty}^{\infty} e^{jux} f(x) dx = \int_0^{\infty} e^{jux} \lambda e^{-\lambda x} dx \\ &= \frac{\lambda}{ju - \lambda} e^{ju - \lambda} \Big|_0^\infty \\ &= \frac{\lambda}{ju - \lambda} [e^{(ju - \lambda)\infty} - 1] = \frac{-\lambda}{-\lambda + ju} = \frac{1}{1 - j\frac{u}{\lambda}}. \end{aligned} \quad (1.265)$$

## 2. 特征函数与概率密度的对应关系

由特征函数的定义可知, 已知随机变量  $X$  的概率密度  $f(x)$ , 总可以求得它的特征函数  $Q_X(u)$ , 那么由特征函数能否求得概率密度呢? 在讨论之前, 先回顾一下傅立叶变换对的定义:

$$\begin{cases} S(\omega) = F[s(t)] = \int_{-\infty}^{\infty} s(t) e^{-j\omega t} dt \\ s(t) = F^{-1}[S(\omega)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) e^{j\omega t} d\omega \end{cases} . \quad (1.266)$$

对照特征函数的定义, 可得

$$Q_X(u) = \int_{-\infty}^{\infty} f(x) e^{jux} dx = 2\pi \left[ \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) e^{jux} dx \right] = 2\pi F^{-1}[f(x)]. \quad (1.267)$$

即

$$F^{-1}[f(x)] = \frac{1}{2\pi} Q_X(u). \quad (1.268)$$

则可用傅立叶变换, 求出  $X$  的密度函数

$$f(x) = F \left[ \frac{1}{2\pi} Q_X(u) \right] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} Q_X(u) e^{-jux} du. \quad (1.269)$$

称之为逆转公式。

由傅立叶变换对之间的唯一性可知,  $Q_X(u)$  与  $f(x)$  之间也是唯一对应的。有了特征函数与概率密度的变换关系, 就可以通过特征函数来间接求独立随机变量和的分布, 比直接用概率密度  $f(x)$  来求, 要方便得多。

**例 1.84** 求  $Y = \sum_{i=1}^n X_i$  的分布  $f_Y(y)$  时, 若用概率密度直接求, 则必须对  $f_X(x_1, \dots, x_n)$  做  $n-1$  重积分, 即

$$f_Y(y) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X \left( x_1 \cdots x_{n-1}, y - \sum_{i=1}^{n-1} x_i \right) dx_1 \cdots dx_{n-1}. \quad (1.270)$$

即使当  $X_1, X_2, \dots, X_n$  相互独立时, 也要做  $n-1$  个积分; 但用特征函数间接求解时, 只要先求  $Q_Y(u) = \prod_{i=1}^n Q_{X_i}(u)$ , 再由逆转公式即可求出  $f_Y(y)$ .

**例 1.85** 证明: 若随机变量  $X_1 \sim B(n, p)$  和  $X_2 \sim B(m, p)$  之间相互独立, 则有  $Y = X_1 + X_2 \sim B(n+m, p)$ .

证: 由特征函数的性质 5° 可得

$$\begin{aligned} Q_Y(u) &= Q_{X_1}(u)Q_{X_2}(u) = (pe^{ju} + q)^n \cdot (pe^{ju} + q)^m \\ &= (pe^{ju} + q)^{n+m}. \end{aligned} \quad (1.271)$$

而右端正是二项分布  $B(n+m, p)$  的特征函数, 由特征函数和概率密度对应关系的唯一性, 可知  $Y \sim B(n+m, p)$ , 即二项分布关于参数  $n$  具有再生性。

**例 1.86** 随机变量  $X \sim E(\lambda)$ , 求特征函数  $Q_X(u)$ .

解:

$$\begin{aligned} Q_X(u) &= \int_{-\infty}^{\infty} e^{jux} f(x) dx = \int_0^{\infty} e^{jux} \lambda e^{-\lambda x} dx \\ &= \frac{\lambda}{ju - \lambda} e^{ju - \lambda} \Big|_0^{\infty} \\ &= \frac{\lambda}{ju - \lambda} [e^{(ju - \lambda)\infty} - 1] = \frac{-\lambda}{-\lambda + ju} = \frac{1}{1 - j\frac{u}{\lambda}}. \end{aligned} \quad (1.272)$$

**例 1.87** 已知随机变量  $X$  服从拉普拉斯分布, 其概率密度为

$$f_X(x) = \frac{1}{2} e^{-|x|}, \quad -\infty < x < +\infty, \quad (1.273)$$

求其特征函数。

解:

$$\begin{aligned} Q_X(u) &= \int_{-\infty}^{\infty} e^{jux} f(x) dx = \int_{-\infty}^{\infty} e^{jux} \frac{1}{2} e^{-|x|} dx \\ &= \frac{1}{2} \left[ \int_{-\infty}^0 e^{jux} e^x dx + \int_0^{\infty} e^{jux} e^{-x} dx \right] \\ &= \frac{1}{2} \left[ \frac{1}{1 + ju} + \frac{1}{1 - ju} \right] \\ &= \frac{1}{1 + u^2}. \end{aligned} \quad (1.274)$$

**例 1.88** 如果随机变量  $X$  的对数  $Y = \ln X$  服从正态分布, 称  $X$  服从对数正态分布, 对数正态分布的概率密度函数  $f(x)$  和概率分布函数  $F(x)$  为

$$\begin{aligned} f(x) &= \frac{1}{xs_1 \sqrt{2\pi}} e^{-\frac{(\ln x - \mu_1)^2}{2s_1^2}}, x \geq 0 \\ F(x) &= \int_0^x \frac{1}{xs_1 \sqrt{2\pi}} e^{-\frac{(\ln x - \mu_1)^2}{2s_1^2}} dx \end{aligned} \quad (1.275)$$

式中  $\mu_1$  是对数平均值,  $\mu = \mu_{\ln x} = \frac{1}{n} \sum_{x=1}^n \ln x$ ,  $s_1$  是对数标准离差,  $s_1 = s_{\ln x} = [\frac{1}{n-1} \sum_{x=1}^n (\ln x - \mu_1)^2]^{\frac{1}{2}}$ . 求其特征函数。

**例 1.89** 威布尔分布的概率密度函数  $f(x)$  和概率分布函数  $F(x)$  为

$$f(x) = \frac{\beta}{\eta} \left( \frac{t - \gamma}{\eta} \right)^{\beta-1} e^{-\left( \frac{t-\gamma}{\eta} \right)^\beta}, \beta > 0, \eta > 0, \gamma \leq x \quad (1.276)$$

$$F(x) = 1 - e^{-\left( \frac{x-\gamma}{\eta} \right)^\beta} \text{ 或, } R(x) = e^{-\left( \frac{x-\gamma}{\eta} \right)^\beta} \quad (1.277)$$

$\beta$ ——形状参数。它决定分布密度函数曲线的形状,  $f(X)$  曲线大致可以分为三类:

1)  $\beta < 1$  时,  $f(x)$  曲线以  $x = \gamma$  为渐近线, 失效率  $\lambda(x)$  为递减函数, 可以用来描述产品的造期失效率。

2)  $\beta = 1$  时,  $f(x)$  是指数分布曲线, 失效率  $\lambda(x)$  为常数可以用来描述产品的偶然失效率。

3)  $\beta > 1$  时,  $f(x)$  是单峰曲线,  $2.7 < \beta < 3.7$  为近似正态分布 ( $\beta = 3.313$  时为正态分布), 失效率是递增函数, 可以用来描述产品的耗损失效期。

$\gamma$ ——位置参数。表示产品在  $\gamma$  之前具有 100% 的存活率 (可靠度)。

$\eta$ ——尺度参数 (或特) 征寿命。求其特征函数。

### 3. 多元特征函数

下面将一维随机变量的特征函数, 推广出多维随机变量的特征函数。

#### (1) 随机矢量的定义

若将  $n$  维随机变量  $(X_1, \dots, X_n)$  用随机矢量  $\mathbf{X}$  表示, 用矢量  $\mathbf{U}$  表示取值  $(x_1, \dots, x_n)$ ,  $n$  个参变量  $(u_1, \dots + u_n)$  用矢量  $U$  表示, 即

$$\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}, \mathbf{X}_i = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \mathbf{U} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \quad (1.278)$$

则  $n$  维随机变量  $(X_1, \dots, X_n)$  的联合特征函数 (或随机矢量  $\mathbf{X}$  的特征函数) 为

$$\begin{aligned} Q_X(u_1, \dots, u_n) &= Q_X(\mathbf{U}^T) = E[e^{j\mathbf{U}^T \mathbf{X}}] \\ &= E[e^{j(u_1 X_1 + u_2 X_2 + \dots + u_n X_n)}] = E[e^{j \sum_{k=1}^n u_k X_k}] \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{j(u_1 x_1 + \dots + u_n x_n)} f_X(x_1, \dots, x_n) dx_1 \dots dx_n. \end{aligned} \quad (1.279)$$

逆转公式为

$$\begin{aligned} f_X(x_1, \dots, x_n) &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-j\mathbf{U}^T x} Q_X(\mathbf{U}^T) \frac{du_1}{2\pi} \dots \frac{du_n}{2\pi} \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-j \sum_{k=1}^n u_k x_k} Q_X(u_1, \dots, u_n) \frac{du_1}{2\pi} \dots \frac{du_n}{2\pi}. \end{aligned} \quad (1.280)$$

#### (2) 性质

1°  $|Q_X(u_1, \dots, u_n)| \leq Q_X(0, \dots, 0) = 1$ .

2° 特征函数  $Q_X(u_1, \dots, u_n)$  在  $R^n$  中一致连续, 其中  $R^n$  表示  $u_1, \dots, u_n$  的  $n$  维空间。

3° 特征函数  $Q_X(u_1, \dots, u_n)$  是关于实变量  $u_1, \dots, u_n$  的复函数, 有

$$Q_X^*(u_1, \dots, u_n) = Q_X(-u_1, \dots, -u_n). \quad (1.281)$$

4° 若  $Q_X(u_1, \dots, u_n)$  是随机矢量  $\mathbf{X}$  的特征函数, 矩阵  $\mathbf{A}$  是  $r \times n$  常系数矩阵, 矢量  $\mathbf{B}$  是  $r$  ( $r < n$ ) 维常数矢量, 则随机矢量  $\mathbf{Y} = \mathbf{AX} + \mathbf{B} \in \mathbb{R}^{r \times 1}$  的特征函数为

$$Q_Y(u_1, \dots, u_r, 0, \dots, 0) = e^{jU^T \mathbf{B}'} Q_X(U^T \mathbf{A}'), \quad (1.282)$$

其中  $\mathbf{B}' = \begin{bmatrix} \mathbf{B} \\ \mathbf{0} \end{bmatrix}_{n \times 1}$ ,  $\mathbf{0}_{(n-r) \times 1}$  为补充的零向量。 $\mathbf{A}' = \begin{bmatrix} \mathbf{A}_{r \times n} \\ \mathbf{0}_{(n-r) \times n} \end{bmatrix}_{n \times n}$ ,  $\mathbf{0}_{(n-r) \times n}$  为补充的零矩阵。

下面讨论两种特殊情况:

① 当  $r = n$  时, 矩阵  $\mathbf{A}$  是  $n \times n$  对角矩阵, 矢量  $\mathbf{B}$  是  $n$  维矢量, 即

$$\mathbf{A} = \begin{bmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}. \quad (1.283)$$

此时

$$\mathbf{AX} = \begin{bmatrix} a_1 X_1 \\ \vdots \\ a_n X_n \end{bmatrix}, \quad \mathbf{Y} = \mathbf{AX} + \mathbf{B} = \begin{bmatrix} a_1 X_1 + b_1 \\ \vdots \\ a_n X_n + b_n \end{bmatrix}, \quad (1.284)$$

则随机矢量  $\mathbf{Y}$  的特征函数为

$$Q_Y(u_1, u_2, \dots, u_n) = e^{j \sum_{k=1}^n u_k b_k} Q_X(a_1 u_1, \dots, a_n u_n). \quad (1.285)$$

② 当  $r = 1$  时, 矩阵  $\mathbf{A} = (a_{ij})_{1 \times n}$ , 而  $B$  是一常数  $b$ , 则

$$\mathbf{Y} = \mathbf{AX} + \mathbf{B} = a_1 X_1 + \dots + a_n X_n + b. \quad (1.286)$$

是一维随机变量, 一元函数情形的特征函数为

$$Q_Y(u_1) = e^{ju_1 b} Q_X(a_1 u_1, a_2 u_1, \dots, a_n u_1). \quad (1.287)$$

5°  $X_1, \dots, X_n$  相互独立的充要条件 (联合特征函数刻画)

### 定义 1.90 $n$ 维随机变量相互独立的充要条件

若  $n$  维随机变量  $(X_1, \dots, X_n)$  的联合特征函数为  $Q_X(u_1, \dots, u_n)$ , 其分量  $X$  的特征函数为  $Q_{X_k}(u_k)$ ,  $k = 1, 2, \dots, n$ , 则  $X_1, \dots, X_n$  相互独立的充要条件为

$$Q_X(u_1, \dots, u_n) = \prod_{k=1}^n Q_{X_k}(u_k). \quad (1.288)$$

由逆转公式, 可证

$$f_X(x_1, \dots, x_n) = \prod_{k=1}^n f_{x_k}(x_k). \quad (1.289)$$



思考: 一元特征函数性质 5° 和多元特征函数性质 5° 有什么不同?

6° 若随机矢量  $\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$  的特征函数为  $Q_X(u_1, \dots, u_n)$ , 则其子向量  $\mathbf{Y} = \begin{bmatrix} X_1 \\ \vdots \\ X_k \end{bmatrix}$

的特征函数为

$$Q_Y(u_1, \dots, u_k) = Q_X(u_1, \dots, u_k, 0, \dots, 0). \quad (1.290)$$

称为  $n$  维随机变量的边缘特征函数, 其中  $k < n$ 。

同理, 任取  $k$  个分量组成的子向量  $\mathbf{Y}_i = \begin{bmatrix} X_{i1} \\ \vdots \\ X_{ik} \end{bmatrix}$ , 类似也可得到其边缘特征函数。

7° 若矩  $E[X_1^n X_2^k]$  存在, 则

$$E[X_1^n X_2^k] = (-j)^{n+k} \frac{\partial^{n+k} Q_X(u_1, u_2)}{\partial u_1^n \partial u_2^k} \Big|_{u_1=0, u_2=0}. \quad (1.291)$$

8° 若对所有  $n = 0, 1, 2, \dots$  和所有  $k = 0, 1, 2, \dots$ ,  $E[X_1^n X_2^k]$  均存在, 则

$$Q_X(u_1, u_2) = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} E[X_1^n X_2^k] \cdot \frac{(ju_1)^n}{n!} \cdot \frac{(ju_2)^k}{k!}. \quad (1.292)$$

### 1.1.10 高斯随机变量

在实际应用中, 常常遇到的随机变量是多个随机变量的和, 需要对其进行刻画, 给出统计特征。中心极限定理已证明, 在满足一定条件的情况下, 大量随机变量和的极限分布是高斯分布。因此, 高斯分布占有特殊的地位, 是科技领域中最常遇到的分布, 也是无线电技术理论(包括噪声理论、信号检测理论、信息理论等)中最重要的概率分布。实际

中有许多随机变量，它们是由大量相互独立的随机因素共同作用所形成的，而其中每一个因素的单独作用都很微小。因此高斯随机变量是一类被广泛应用的随机变量。

### 1. 一维高斯随机变量及其特征

一维高斯随机变量  $X$  的概率密度函

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left[-\frac{(x-m_X)^2}{2\sigma_X^2}\right], x \in \mathbb{R}, \quad (1.293)$$

式中， $m_X$  为  $X$  的均值， $\sigma_X$  为  $X$  的方差。

如果  $m_X = 0, \sigma_X = 1$ ，则

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \quad -\infty < x < +\infty. \quad (1.294)$$

对于一给定的  $x$ ，高斯随机变量  $X$  的分布函数为

$$\begin{aligned} F(x) &= P(X \leq x) \\ &= \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left[-\frac{(u-m_X)^2}{2\sigma_X^2}\right] du. \end{aligned} \quad (1.295)$$

令  $\frac{u-m_X}{\sigma_X} = t$ ，则

$$F(x) = \int_{-\infty}^{(x-m_X)/\sigma_X} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt = \Phi\left(\frac{x-m_X}{\sigma_X}\right). \quad (1.296)$$

其中  $\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt$  是概率积分函数。

高斯分布函数还可以表示为

$$\begin{aligned} F(x) &= \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{x-m_X}{\sqrt{2}\sigma_X}\right). \\ F(x) &= 1 - \frac{1}{2} \operatorname{erfc}\left(\frac{x-m_X}{\sqrt{2}\sigma_X}\right). \end{aligned} \quad (1.297)$$

其中  $\operatorname{erf}(x)$  称为误差函数， $\operatorname{erfc}(x)$  称为互补误差函数，它们分别为

$$\begin{aligned} \operatorname{erf}(x) &= \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt; \\ \operatorname{erfc}(x) &= 1 - \operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_x^{+\infty} e^{-t^2} dt. \end{aligned} \quad (1.298)$$

1. 高斯随机变量  $X$  具有下列性质 (式 (1.293) 所表示的随机变量):

1)  $X$  的概率密度函数曲线具有单峰对称形式，并且在均值  $m_X$  处有极大值  $\frac{1}{\sqrt{2\pi}\sigma_X}$ ，当  $x \rightarrow \pm\infty$  时，概率密度函数的值趋于零。均值反映了  $X$  概率密度函数的最大值所在的位

置信息，方差反映了概率密度函数曲线的陡度，随着  $\sigma_X$  的增大曲线变得平坦，随机变量  $X$  落在  $[m_X - 3\sigma_X, m_X + 3\sigma_X]$  区间的概率

$$\begin{aligned} P(m_X - 3\sigma_X \leq x \leq m_X + 3\sigma_X) &= \frac{1}{\sqrt{2\pi}\sigma_X} \int_{m_X-3\sigma_X}^{m_X+3\sigma_X} \exp\left[-\frac{(x-m_X)^2}{2\sigma_X^2}\right] dx \\ &\approx 99.7\%. \end{aligned} \quad (1.299)$$

2) 高斯随机变量  $X$  的特征函数为

$$\varphi(\omega) = \exp\left(j\omega m_X - \frac{1}{2}\omega^2\sigma_X^2\right). \quad (1.300)$$

(3) 高斯随机变量  $X$  的  $n$  阶中心矩

$$E\{(X - m_X)^n\} = \begin{cases} 0, & n = 2k - 1 \\ 1 \times 3 \times 5 \times \cdots \times (n-1)\sigma_X^n, & n = 2k \end{cases}, k = 1, 2, \dots \quad (1.301)$$

## 2. 二维高斯随机变量及特征

设  $X_1$  和  $X_2$  分别是均值为  $m_{X_1}$ 、方差为  $\sigma_{X_1}^2$  和均值为  $m_{X_2}$ 、方差为  $\sigma_{X_2}^2$  的高斯随机变量，它们的相关系数为  $\rho_{X_1 X_2}$ ，那么二维高斯随机变量  $(X_1, X_2)$  的联合概率密度函数为

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_{X_1}\sigma_{X_2}\sqrt{1-\rho_{X_1 X_2}}} \exp\left\{-\frac{1}{2(1-\rho_{X_1 X_2}^2)} \left[ \frac{(x_1 - m_{X_1})^2}{\sigma_{X_1}^2} - \frac{2\rho_{X_1 X_2}(x_1 - m_{X_1})(x_2 - m_{X_2})}{\sigma_{X_1}\sigma_{X_2}} + \frac{(x_2 - m_{X_2})^2}{\sigma_{X_2}^2} \right]\right\}. \quad (1.302)$$

可见，二维高斯随机变量的概率密度函数由  $m_{X_1}, m_{X_2}, \sigma_{X_1}^2, \sigma_{X_2}^2, \rho_{X_1 X_2}$  五个参数完全确定。

高斯随机变量  $X_1$  和  $X_2$  的边缘概率密度函数分别为

$$f_{X_1}(x_1) = \int_{-\infty}^{+\infty} f(x_1, x_2) dx_2 = \frac{1}{\sqrt{2\pi}\sigma_{X_1}} \exp\left[-\frac{(x_1 - m_{X_1})^2}{2\sigma_{X_1}^2}\right]. \quad (1.303)$$

$$f_{X_2}(x_2) = \int_{-\infty}^{+\infty} f(x_1, x_2) dx_1 = \frac{1}{\sqrt{2\pi}\sigma_{X_2}} \exp\left[-\frac{(x_2 - m_{X_2})^2}{2\sigma_{X_2}^2}\right]. \quad (1.304)$$

可见，当相关系数  $\rho_{X_1 X_2} = 0$  时，有  $f(x_1, x_2) = f_{X_1}(x_1)f_{X_2}(x_2)$ ，即两个高斯随机变量  $X_1$  和  $X_2$  互不相关就意味着它们统计独立。如果两个随机变量  $X_1$  和  $X_2$  的联合概率密度函数是高斯的，则  $X_1$  和  $X_2$  各自的概率密度函数也是高斯的。反之，则不一定。

可以证明，二维高斯随机变量的特征函数

$$\varphi(\omega_1, \omega_2) = \exp\left[j(m_{X_1}\omega_1 + m_{X_2}\omega_2) - \frac{1}{2}(\sigma_{X_1}^2\omega_1^2 + 2\rho_{X_1 X_2}\sigma_{X_1}\sigma_{X_2}\omega_1\omega_2 + \sigma_{X_2}^2\omega_2^2)\right]. \quad (1.305)$$

### 3. 多维高斯随机变量的有关性质

设  $X_1, X_2, \dots, X_n$  组成  $n$  维高斯随机变量, 令  $\mathbf{x} = [X_1, X_2, \dots, X_n]^T$ , 均值矢量为

$$\begin{aligned} m_X &= [E\{X_1\}, E\{X_2\}, \dots, E\{X_n\}]^T \\ &= [m_{X_1}, m_{X_2}, \dots, m_{X_n}]^T. \end{aligned} \quad (1.306)$$

它的协方差矩阵为

$$C_X = \begin{bmatrix} E\{(X_1 - m_{X_1})^2\} & \cdots & E\{(X_1 - m_{X_1})(X_n - m_{X_n})\} \\ \vdots & & \vdots \\ E\{(X_n - m_{X_n})(X_1 - m_{X_1})\} & \cdots & E\{(X_n - m_{X_n})^2\} \end{bmatrix}. \quad (1.307)$$

则  $n$  维高斯随机变量的联合概率密度函数  $f(x)$  可以表示为

$$f(x) = \frac{1}{(2\pi)^{\frac{n}{2}} |C_X|^{\frac{1}{2}}} \exp\left[-\frac{(x - m_X)^T C_X^{-1} (x - m_X)}{2}\right]. \quad (1.308)$$

$n$  维高斯随机变量的特征函数为

$$\varphi_X(\omega) = \exp(jm_X^T \omega - \omega^T C_X \omega / 2), \quad (1.309)$$

式中,  $\omega = [\omega_1, \omega_2, \dots, \omega_n]^T$ .

$n$  维高斯随机变量具有以下性质:

- (1)  $n$  维高斯随机变量经过线性变换后仍然是高斯随机变量;
- (2)  $n$  维互不相关的高斯随机变量一定是彼此统计独立的。

由随机变量的数字特征的有关分析, 易知: 一般的  $n$  维随机变量, 若它们彼此统计独立, 则必然互不相关, 反之, 则不一定成立; 对于  $n$  维高斯随机变量, 若它们互不相关, 则必然是彼此统计独立的。这一性质对分析与处理高斯随机信号带来很大的方便。

### 2. 高斯变量的矩

标准高斯变量  $Y$  的  $n$  ( $n \geq 2$ ) 阶矩为

$$E[Y^*] = \begin{cases} 1 \cdot 3 \cdot 5 \cdots (n-1) = (n-1)!! , & n \text{ 为偶数} \\ 0, & n \text{ 为奇数} \end{cases}, \quad n \geq 2. \quad (1.310)$$

一般高斯变量  $X \sim N(m, \sigma^2)$  的  $n$  ( $n \geq 2$ ) 阶中心距为

$$E[(X - m)^n] = \begin{cases} \sigma^n (n-1)!! , & n \text{ 为偶数} \\ 0, & n \text{ 为奇数} \end{cases}.$$

### 3. 高斯变量的和

#### ① 相互独立的高斯变量之和服从高斯分布

设  $n$  个相互独立的高斯变量  $X_k \sim N(m_k, \sigma_k^2), k = 1, \dots, n$ , 那么它的和也服从高斯分布, 即  $Y = \sum_{k=1}^n X_k \sim N(m_Y, \sigma_Y^2)$ , 均值为  $m_Y = \sum_{k=1}^n m_k$ , 方差为  $\sigma_Y^2 = \sum_{k=1}^n \sigma_k^2$ .

#### ② 相关的高斯变量之和服从高斯分布

设  $n$  个相关的高斯变量  $X_k \sim N(m_k, \sigma_k^2), k = 1, \dots, n$ ,  $X_i$  和  $X_j (i, j = 1, 2, \dots, k)$  的相关系数为  $\rho_{ij}$ 。那么其和  $Y = \sum_{k=1}^n X_k$  也服从高斯分布, 其期望为  $m_Y = \sum_{k=1}^n m_k$ , 方差为  $\sigma_Y^2 = \sum_{k=1}^n \sigma_k^2 + 2 \sum_{j < i} \rho_{ij} \sigma_i \sigma_k$ .

#### 1. $n$ 维高斯变量的概率密度函数与特征函数

若  $n$  个随机变量  $X_1, \dots, X_n$  均是高斯变量, 则称  $(X_1, \dots, X_n)$  为  $n$  维高斯变量。若用随机矢量  $X$  来表示  $n$  维随机变量, 用矢量  $x$  表示  $X$  的取值, 用矢量  $M_X$  表示随机矢量  $X$  的均值

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}, x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, M_x = \begin{bmatrix} m_1 \\ \vdots \\ m_n \end{bmatrix}. \quad (1.311)$$

方差为

$$D[X] = C_X = E[(X - M_X)(X - M_X)^T] = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix}_{n \times n}. \quad (1.312)$$

则  $n$  维高斯变量  $(X_1, \dots, X_n)$  的联合概率密度 (或  $n$  维高斯矢量  $X$  的概率密度) 可以写成如下矩阵形式:

$$f_X(x_1, \dots, x_n) = \frac{1}{(2\pi)^{n/2} |C_X|^{1/2}} \exp \left[ -\frac{(x - M_X)^T C_X^{-1} (x - M_X)}{2} \right]. \quad (1.313)$$

式中  $(x - M_X)^T$  表示  $(x - M_X)$  的转置。 $n$  维高斯分布可记为  $X \sim N(M_X, C_X)$ 。

$n$  维高斯变量  $(X_1, \dots, X_n)$  的联合特征函数也可以表示为矩阵形式

$$Q_X(u_1, \dots, u_n) = E[e^{jU^T X}] = \exp \left[ jM_X^T U - \frac{U^T C_X U}{2} \right]. \quad (1.314)$$

其中

$$U = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}, U^T X = \sum_{i=1}^n u_i X_i. \quad (1.315)$$

## 2. $n$ 维高斯变量的性质

1°  $n$  维高斯变量的互不相关与独立是等价的。

对于一般随机变量而言, 独立必定不相关, 不相关不一定独立。但对于高斯变量而言, 互不相关与独立是等价的。

证: 设  $n$  维高斯变量  $X_1, \dots, X_n$  互不相关, 且  $X_k \sim N(m_k, \sigma_k^2)$ ,  $k = 1, \dots, n$ , 即所有协方差  $C = 0$  ( $i \neq 1$ ), 其协方差矩阵为对角阵。

$$C_X = \begin{bmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{bmatrix}_{n \times n}. \quad (1.316)$$

则  $n$  维联合特征函数

$$\begin{aligned} Q_X(u_1, \dots, u_n) &= \exp \left\{ j(m_1, \dots, m_n) \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} - \frac{1}{2} (u_1, \dots, u_n) \begin{bmatrix} \sigma_1^2 & 0 & \\ & \ddots & \\ 0 & & \sigma_n^2 \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \right\} \\ &= \exp \left[ \sum_{k=1}^n jm_k u_k - \sum_{k=1}^n \frac{u_k^2 \sigma_k^2}{2} \right] = \exp \left[ \sum_{k=1}^n \left( jm_k u_k - \frac{u_k^2 \sigma_k^2}{2} \right) \right] \\ &= \prod_{k=1}^n \exp \left[ jm_k u_k - \frac{u_k^2 \sigma_k^2}{2} \right] = \prod_{k=1}^n Q_{x_k}(u_k). \end{aligned} \quad (1.317)$$

由特征函数性质 5° 可知,  $X_1, \dots, X_n$  之间相互独立。

2°  $n$  维高斯变量线性变换后仍服从高斯分布。

证: 先用  $n$  维高斯矢量  $\mathbf{X} = (X_1, \dots, X_n)^\top$  来表示  $n$  维高斯变量  $(X_1, \dots, X_n)$ , 再对  $\mathbf{X}$  作一线性变换  $\mathbf{Y} = \mathbf{AX}$ , 其中  $\mathbf{A} = (a_{ij})_{m \times n}$  为  $m \times n$  阶常系数矩阵。线性变换后

$$\mathbf{Y} = \mathbf{A}_{m \times n} \mathbf{X}_{n \times 1} = \begin{bmatrix} Y_1 = a_{11}X_1 + \dots + a_{1n}X_n \\ Y_2 = a_{21}X_1 + \dots + a_{2n}X_n \\ \vdots \\ Y_m = a_{m1}X_1 + \dots + a_{mn}X_n \end{bmatrix}. \quad (1.318)$$

$Y$  表示新的  $m$  维随机变量  $\{Y_1, \dots, Y_m\}$ . 若  $X \sim N(M_X, C_X)$ , 则线性变换后的数字特征为

$$\begin{aligned} M_Y &= E[Y] = E[AX] = AE[X] = AM_X. \\ C_Y &= D[Y] = E[(Y - M_Y)(Y - M_Y)^T] \\ &= E[(AX - AM_X)(AX - AM_X)^T] \\ &= E[A(X - M_X)(X - M_X)^T A^T] \\ &= AE[(X - M_X)(X - M_X)^T] A^T \\ &= AC_X A^T. \end{aligned} \quad (1.319)$$

$$\begin{aligned} Q_Y(u_1, \dots, u_n) &= E[\exp(jU^T Y)] = E[\exp(jU^T AX)] \\ &= E[\exp(j(A^T U)^T)] \\ &= Q_x(U^T A) = \exp\left[jM_X^T(A^T U) - \frac{(A^T U)^T C_X (A^T U)}{2}\right] \\ &= \exp\left[j(AM_X)^T U - \frac{U^T (AC_X A^T) U}{2}\right] \\ &= \exp\left[jM_Y^T U - \frac{U^T C_X U}{2}\right]. \end{aligned} \quad (1.320)$$

由特征函数的形式可知, 线性变换后的  $m$  维随机变量  $(Y_1, \dots, Y_m)$  仍服从高斯分布。

由于高斯变量经过线性变换后仍服从高斯分布, 因此, 只要常系数矩阵  $A$  选择适当, 使常系数矩阵  $A$  成为对角阵, 则可以使不独立的  $X_1, \dots, X_n$ , 通过线性变换构成相互独立的变量  $Y_1, \dots, Y_m$ 。

$$C_Y = AC_X A^T = \begin{bmatrix} \sigma_{X_1}^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_{X_n}^2 \end{bmatrix}. \quad (1.321)$$

3°  $n$  维高斯变量的边缘分布仍服从高斯分布

证: 设  $n$  维高斯矢量  $\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$ , 其中任一子矢量为  $\bar{\mathbf{X}} = \begin{bmatrix} X_i \\ \vdots \\ X_{ik} \end{bmatrix}$ ,  $k < n$ 。数学期望

$M_X = \begin{bmatrix} m_1 \\ \vdots \\ m_n \end{bmatrix}$ ,  $M_{\bar{X}} = \begin{bmatrix} m_n \\ \vdots \\ m_{ik} \end{bmatrix}$  为  $M$  中相应的子矢量。 $C_{\bar{X}}$  为协方差矩阵  $C_X$  中的子矩阵,

$C_X = \begin{bmatrix} C_X & C_1 \\ C_2 & C_3 \end{bmatrix}_{n \times n}$  其中  $C_{\bar{X}}$  为  $k$  阶矩阵。根据  $n$  维随机变量特征函数的性质可知,  $n$  维

随机矢量  $\mathbf{X}$  的  $k$  维边缘特征函数为

$$Q_{\tilde{\mathbf{X}}}(u_{i1}, \dots, u_{ik}) = Q_{\mathbf{X}}(u_n, \dots, u_{ik}, 0, \dots, 0), \quad (1.322)$$

其中  $\tilde{\mathbf{U}} = (u_{i1}, \dots, u_{ik})^T$  为  $\mathbf{U} = (u_1, \dots, u_n)^T$  的子矢量。

若  $\mathbf{X}$  为  $n$  维高斯矢量, 则由特征函数的矩阵形式与边缘特征函数的求法, 得

$$\begin{aligned} Q_{\tilde{\mathbf{X}}}(u_{i1}, \dots, u_{ik}) &= \exp \left[ j(m_1, \dots, m_n) \begin{pmatrix} u_{i1} \\ \vdots \\ u_{ik} \\ 0 \\ \vdots \\ 0 \end{pmatrix} - \frac{1}{2} (u_{i1}, \dots, u_{ik}, 0, \dots, 0) \mathbf{C}_{\mathbf{X}} \begin{pmatrix} u_{i1} \\ \vdots \\ u_{ik} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right] \\ &= \exp \left[ j \mathbf{M}_{\tilde{\mathbf{X}}}^T \tilde{\mathbf{U}} - \frac{1}{2} \tilde{\mathbf{U}}^T \mathbf{C}_{\tilde{\mathbf{X}}} \tilde{\mathbf{U}} \right]. \end{aligned} \quad (1.323)$$

由逆转公式可知,  $n$  维高斯变量的  $k$  维边缘分布也服从高斯分布。

**例 1.91** 已知三维随机矢量  $\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \sim N(\mathbf{0}, \mathbf{B})$ , 其中  $\mathbf{B}^{-1} = \begin{pmatrix} 3 & 2 & 0 \\ 2 & 4 & -2 \\ 0 & -2 & 5 \end{pmatrix}$ . 令二维随机

矢量  $\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$ , 其中  $\begin{cases} Y_1 = X_1 \\ Y_2 = X_1 + X_2 \end{cases}$ . 求: ① 随机矢量  $\mathbf{X}$  的方差  $D[\mathbf{X}]$ . ② 随机矢量  $\mathbf{Y}$  的概率密度和特征函数。③ 随机变量  $Y_1$  和  $Y_2$  服从什么分布? 它们是否独立? 给出理由。

解:

① 随机矢量  $\mathbf{X}$  的方差

$$D[\mathbf{X}] = \mathbf{B} = [\mathbf{B}^{-1}]^{-1} = \begin{bmatrix} 3 & 2 & 0 \\ 2 & 4 & -2 \\ 0 & -2 & 5 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{4}{7} & -\frac{5}{14} & -\frac{1}{7} \\ -\frac{5}{14} & \frac{15}{28} & \frac{3}{14} \\ -\frac{1}{7} & \frac{3}{14} & \frac{2}{7} \end{bmatrix}. \quad (1.324)$$

① 随机矢量  $\mathbf{Y} = \mathbf{A}\mathbf{X}$ , 其中  $\mathbf{A}$  由性质 2° 可知

$$\mathbf{M}_Y = \mathbf{A}\mathbf{M}_X = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \quad (1.325)$$

$$\begin{aligned}
C_Y &= \mathbf{A} C_X \mathbf{A}^T \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{bmatrix} \frac{4}{7} & -\frac{5}{14} & -\frac{1}{7} \\ -\frac{5}{14} & \frac{15}{28} & \frac{3}{14} \\ -\frac{1}{7} & \frac{3}{14} & \frac{2}{7} \end{bmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} \frac{4}{7} & -\frac{5}{14} & -\frac{1}{7} \\ \frac{3}{14} & \frac{5}{28} & \frac{1}{14} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} \frac{4}{7} & \frac{3}{14} \\ \frac{3}{14} & \frac{11}{28} \end{pmatrix}. \tag{1.326}
\end{aligned}$$

可得  $|C_Y| = \begin{vmatrix} \frac{4}{7} & \frac{3}{14} \\ \frac{3}{14} & \frac{11}{28} \end{vmatrix} = \frac{5}{28}$ , 方差的逆矩阵为  $C_Y^{-1} = \begin{pmatrix} \frac{11}{5} & -\frac{6}{5} \\ -\frac{6}{5} & \frac{16}{5} \end{pmatrix}$ .

则其概率密度

$$\begin{aligned}
f_Y(y_1, j_z) &= \frac{1}{2\pi |C_Y|^{\frac{1}{2}}} \exp \left[ -\frac{(y - M_Y)^T C_Y^{-1} (y - M_Y)}{2} \right] \\
&= \frac{1}{2\pi |C_Y|^{\frac{1}{2}}} \exp \underbrace{[y^T C_Y^{-1} y]}_{\text{}} \\
&= \frac{1}{\pi} \sqrt{\frac{7}{5}} \exp \left[ -\frac{11y_1^2 - 12y_1y_2 + 16y_2^2}{10} \right]. \tag{1.327}
\end{aligned}$$

其特征函数为

$$Q_Y(u_1, u_2) = \exp \left[ jM_Y^T \mathbf{U} - \frac{\mathbf{U}^T C_Y \mathbf{U}}{2} \right] = \exp \left[ -\frac{16u_1^2 + 12u_1u_2 + 11u_2^2}{56} \right]. \tag{1.328}$$

② 由性质 2° 可知, 随机矢量  $Y$  也服从高斯分布; 由性质 3° 可知, 高斯分布的

$$Q_Y(u_1, u_2) = \exp \left[ jM_Y^T \mathbf{U} - \frac{\mathbf{U}^T C_Y \mathbf{U}}{2} \right] = \exp \left[ -\frac{16u_1^2 + 12u_1u_2 + 11u_2^2}{56} \right]. \tag{1.329}$$

边缘分布也服从高斯分布, 即  $Y_1$  和  $Y_2$  服从高斯分布。由于随机矢量  $Y$  的方差  $C$  不是对角阵, 所以  $Y_1$  和  $Y_2$  是相关的, 也表明  $Y_1$  和  $Y_2$  不是相互独立的。

### 1.1.11 复随机变量

上述的讨论是针对实随机变量的, 在实际中我们还经常用到复随机变量中, 这里先定义复随机变量为

$$Z = X_1 + jX_2, \tag{1.330}$$

其中,  $X_1$  和  $X_2$  为两个实随机变量这样一个复随机变量的概率密度函数, 可以定义为如下的联合概率密度函数

$$f(z) = f(x_1, x_2). \quad (1.331)$$

一个复随机变量的均值为

$$\begin{aligned} m_Z &= E\{Z\} = E\{X_1 + jX_2\} \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x_1 + jx_2) f(x_1, x_2) dx_1 dx_2. \end{aligned} \quad (1.332)$$

或者, 定义  $dz = dx_1 dx_2$  作为复  $z$  平面上的积分微元, 是一块无穷小区域, 有

$$m_Z = E\{Z\} = \int_{(z)} z f(z) dz. \quad (1.333)$$

且

$$\begin{aligned} E\{|Z|^2\} &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |x_1 + jx_2|^2 f(x_1, x_2) dx_1 dx_2 \\ &= E\{X_1^2\} + E\{X_2^2\}. \end{aligned} \quad (1.334)$$

则复随机变量的方差为

$$\sigma_Z^2 = E\{|Z - m_Z|^2\} = \sigma_{X_1}^2 + \sigma_{X_2}^2 = E\{|Z|^2\} - m_Z^2. \quad (1.335)$$

对于两个复随机变量,  $W = U_1 + jU_2, Z = X_1 + jX_2$ , 概率密度函数为

$$f(w, z) = f(x_1, x_2, u_1, u_2). \quad (1.336)$$

其协方差定义为

$$C_{WZ} = E\{(W - m_W)(Z - m_Z)\}. \quad (1.337)$$

相关系数为

$$\rho_{WZ} = \frac{C_{WZ}}{\sigma_W \sigma_Z}. \quad (1.338)$$

一般情况下相关系数为复值。

对于复随机变量  $w$  和  $z$ , 我们定义协方差矩阵和互协方差矩阵为

$$\begin{aligned} C_Z &= E\{(Z - m_Z)(Z - m_Z)^H\}. \\ C_{WZ} &= E\{(W - m_W)(Z - m_Z)^H\}. \end{aligned} \quad (1.339)$$

特别地, 对  $n$  维复高斯随机矢量  $\mathbf{z}$ , 若其均值和协方差矩阵用  $\mathbf{m}_Z$  和  $\mathbf{C}_Z$  表示, 则复高斯随机矢量的概率密度函数为

$$f(\mathbf{z}) = \frac{1}{\pi^n |\mathbf{C}_Z|} \exp \left[ -(\mathbf{z} - \mathbf{m}_Z)^H \mathbf{C}_Z^{-1} (\mathbf{z} - \mathbf{m}_Z) \right]. \quad (1.340)$$

高斯随机变量的特征函数由下式给出

$$\varphi(\omega) = \exp \left( j \operatorname{Re} \{ \omega^H \mathbf{m}_Z \} - \frac{\omega^H \mathbf{C}_Z \omega}{4} \right). \quad (1.341)$$

其中  $\operatorname{Re}\{\cdot\}$  表示复数的取实部操作。

## 1.2 线性代数导论

### 1.2.1 矩阵的概念和基本运算

在本小节中, 我们首先引入矩阵的概念并介绍其基本运算方法。

#### 定义 1.92 数域

设它是复数集合  $\mathbb{C}$  的一个非空子集, 若  $\mathbb{F}$  中的任意两个数  $a, b$  的和  $a+b$  从差  $a-b$ , 积  $ab$ , 商  $a/b$  ( $b$  不等于 0) 仍在  $\mathbb{F}$  中, 则称  $\mathbb{F}$  为一个数域, 我们常用  $\mathbb{R}$  表示实数域,  $\mathbb{C}$  表示复数域。

#### 定义 1.93 数域上的矩阵

设  $A$  是复数域  $\mathbb{C}$  中  $m \times n$  个数  $a_{ij}$  ( $i = 1, 2, \dots, m; j = 1, 2, \dots, n$ ) 所组成的数集, 若规定了其中各个数之间的相对位置并排成  $m$  行  $n$  列的数表

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

则称数集  $A$  为复数域  $\mathbb{C}$  上的  $m$  行  $n$  列矩阵, 简称为  $m \times n$  矩阵, 记为  $A = [a_{ij}] = (a_{ij})_{m \times n}$ , 其中数  $a_{ij}$  ( $i = 1, 2, \dots, m; j = 1, 2, \dots, n$ ) 称为矩阵  $A$  的第  $i$  行第  $j$  列元素。数域  $\mathbb{C}$  上的全体  $m \times n$  矩阵的集合记为  $\mathbb{C}^{m \times n}$ (或  $\mathbb{R}^{m \times n}$ )。

**定义 1.94 矩阵相等**

若  $A, B \in C^{m \times n}$ , 并且它们的对应元素相等, 即

$$a_{ij} = b_{ij} \quad (i = 1, 2, \dots, m; j = 1, 2, \dots, n)$$

则称矩阵  $A$  与矩阵  $B$  相等, 记作  $A = B$ .

**定义 1.95 矩阵转置**

设  $A \in C^{m \times n}$ , 矩阵  $A$  的第  $i$  ( $i = 1, 2, \dots, m$ ) 行作为新矩阵  $B$  的第  $i$  列, 则称  $B$  为  $A$  的转置矩阵, 记作  $B = A^T$ , 显然有  $A^T \in C^{n \times m}$ ,  $(A^T)^T = A$ .

**定义 1.96 矩阵共轭**

设  $A \in C^{m \times n}$ , 矩阵  $A$  的第  $i$  个元素换成其共轭, 则称  $A^*$  是  $A$  的共轭矩阵, 记作  $B = A^* = A^H$ , 显然有  $A^H \in C^{n \times m}$ ,  $(A^H)^H = A$ .

**定义 1.97 矩阵和**

设  $A, B \in C^{m \times n}$ ,  $A, B$  和矩阵  $C = [a_{ij} \pm b_{ij}]$ , 记作  $C = A \pm B$ , 显然,  $A \pm B \in C^{m \times n}$ .

**定义 1.98 矩阵积**

设  $A \in C^{m \times s}$ ,  $B \in C^{s \times n}$ ,  $A, B$  中的元素按  $\sum_{k=1}^s a_{ik} b_{kj}$  方式计算, 得到新矩阵  $C = [c_{ij}]$  成为是  $A$  与  $B$  的乘积, 记为  $C = AB$ .



矩阵乘法具有如下运算性质:

设  $A, B, C$  是矩阵,  $K$  是数, 并且运算都是可行的, 则

(1) 结合律

$$\begin{aligned} (AB)C &= A(BC). \\ (kB)C &= k(BC). \end{aligned} \tag{1.342}$$

(2) 分配律

$$\begin{aligned} A(B+C) &= AB+AC. \\ (A+B)C &= AC+BC. \end{aligned} \tag{1.343}$$

(3) 转置律

$$(AB)^T = B^T A^T. \tag{1.344}$$

矩阵的初等变换也是我们经常用到的矩阵操作，它主要包括：位置变换，即互换矩阵的第  $i$  行 (列) 与第  $j$  行 (列)；数乘变换，即用非零数  $k$  乘矩阵的第  $i$  行 (列)。

### 定义 1.99 迹

$A = [a_{ij}] \in C^{n \times n}$  的所有对角元素之和称为  $A$  的迹，记为  $\text{tr}(A) = a_{11} + a_{22} + \dots + a_{nn}$ 。

### 定义 1.100 行列式的阶

$$\det\{A\} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \quad (1.345)$$

表示关于矩阵  $A$  的  $n^2$  个元素的一种特定代数运算，称为矩阵  $A$  的行列式，其中  $n$  称为行列式的阶。

### 定义 1.101 迹

$A \in C^{n \times n}$ ，把  $A$  中元素  $a_{ij}$  所在的第  $i$  行和第  $j$  列划去后，留下来的  $n-1$  阶方阵称为元素  $a_{ij}$  的余子矩阵，记为  $A(i, j)$ ，并且称  $\text{cof}(A_{ij}) = \det\{A(i, j)\}$  为元素  $a_{ij}$  的余子式， $(-1)^{i+j} \det\{A(i, j)\}$  称为元素  $a_{ij}$  的代数余子式，记为  $A_{ij}$ 。

行列式具体的运算法则是

当  $n = 1$  时， $\det\{A\} = a_{11}$ 。

当  $n > 1$  时， $\det\{A\} = a_{11}A_{11} + a_{21}A_{21} + \dots + a_{n1}A_{n1} = \sum_{k=1}^n a_{k1}A_{k1}$ ，其中  $A_{k1}$  是  $a_{ij}$  的代数余子式。

特别地当  $\det(A) = 0$  时， $A$  称为奇异 (退化) 方阵，否则称为非奇异 (非退化) 方阵，也称正则矩阵。

$n$  阶行列式的性质：

(1) 转置变换不改变行列式的值，即  $\det\{A^T\} = \det\{A\}$ ；

(2) 位置变换仅改变行列式的符号，即两行 (列) 互换，行列式的值变号。

### 定理 1.102

数乘变换改变行列式的值，并且仿方阵  $A = [a_1, \dots, a_i, \dots, a_n], f_j|, B =$

$[a_1, \dots, k a_i, \dots, a_n]$ , 其中列矢量  $a_i \in \mathbb{C}^{n \times 1}$  ( $i = 1, 2, \dots, n$ ), 则

$$\det\{\mathbf{B}\} = k \det\{\mathbf{A}\}. \quad (1.346)$$

$$\det\{k \mathbf{A}\} = k^n \det\{\mathbf{A}\}. \quad (1.347)$$

$$\det\{\mathbf{AB}\} = \det\{\mathbf{A}\} \det\{\mathbf{B}\}, \mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}. \quad (1.348)$$

引入矩阵子式的概念, 对于矩阵  $\mathbf{A} \in \mathbb{C}^{m \times n}$ , 其  $k$  行、 $k$  列相交的元素所组成的矩阵的行列式称为  $A$  的  $K$  阶子式。

### 定义 1.103 满秩

设矩阵  $\mathbf{A} \in \mathbb{C}^{m \times n}$ , 有一个  $k$  阶子式不等于零, 而所有  $k+1$  阶子式(若存在)都等于零, 则称正整数  $k$  为矩阵  $A$  的秩, 记为  $\text{rank } A = k$ . 若  $m = n = k$ , 则称  $A$  为满秩矩阵.

### 定义 1.104 迹

矩阵  $\mathbf{A} \in \mathbb{C}^{m \times n}$ , 如果对于任意的非零列矢量  $x \in \mathbb{C}^n$ , 有

- (1) 当  $\text{Re}\{x^H A x\} > 0$  时, 称矩阵  $A$  为正定矩阵;
- (2) 当  $\text{Re}\{x^H A x\} \geq 0$  时, 称矩阵  $A$  为半正定(非负定)矩阵;
- (3) 当  $\text{Re}\{x^H A x\} < 0$  时, 称矩阵  $A$  为负定矩阵.

### 定义 1.105

设  $A$  和  $B$  都是  $n$  阶方阵, 则

- (1) 当方阵  $C = A - B$  为正定矩阵时, 称矩阵  $A > B$ .
- (2) 当方阵  $C = A - B$  为半正定矩阵时, 称矩阵  $A \geq B$ ;
- (3) 当方阵  $C = A - B$  为负定矩阵时, 称矩阵  $A < B$ .

### 定义 1.106 范数

矩阵  $\mathbf{A} \in \mathbb{C}^{m \times n}$  的某个行列式  $\|\mathbf{A}\|$  满足如下性质

- (1)  $A$  为非零矩阵时,  $\|\mathbf{A}\| > 0$ ;  $A$  为零矩阵时,  $\|\mathbf{A}\| = 0$ ;
- (2) 对于任意复数  $c$  有  $\|c \mathbf{A}\| = |c| \|\mathbf{A}\|$ ;
- (3)  $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$ ;
- (4)  $\|\mathbf{AB}\| \leq \|\mathbf{A}\| \|\mathbf{B}\|$ .

则称  $\|\mathbf{A}\|$  是  $\mathbb{C}^{m \times n}$  上的一个范数.

常见的范数有 F 范数、行和范数和列和范数：

$$\begin{aligned}\|A\|_F &= \left( \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right)^{1/2}. \\ \|A\|_{\text{row}} &= \max_{1 \leq i \leq m} \left\{ \sum_{j=1}^n |a_{ij}| \right\}. \\ \|A\|_{\text{col}} &= \max_{1 \leq i \leq n} \left\{ \sum_{i=1}^m |a_{ij}| \right\}.\end{aligned}\quad (1.349)$$

### 定义 1.107 矩阵的内积

对于同型矩阵  $A, B \in \mathbb{C}^{m \times n}$ , 可以定义

$$\{A, B\} = \text{tr} \{AB^H + BA^H\} / 2 \quad (1.350)$$

为矩阵的内积.



### 定理 1.108 Laurent-Moïse Schwartz 不等式

设  $A \in \mathbb{C}^{m \times n}, B \in \mathbb{C}^{n \times l}, AA^T$  可逆, 则

$$B^T B \geq (AB)^T (AA^T)^{-1} (AB) \quad (1.351)$$

等号成立的条件为存在  $C \in \mathbb{C}^{m \times l}, B = A^T C$ .



## 1.2.2 特殊矩阵

### 定义 1.109 矩阵

设矩阵  $A \in \mathbb{C}^{m \times n}$ , 如果  $A$  在主对角线上的元素全为 1 而其他元素全为 0, 则称  $A$  为  $n$  阶单位矩阵, 通常用  $I$  表示; 如果  $A$  在交叉对角线上的元素全为 1 而其他元素全为 0, 则称  $A$  为  $n$  阶反向单位矩阵, 通常用  $J$  表示。



### 定义 1.110 范德蒙矩阵

设矩阵  $A \in \mathbb{C}^{m \times n}$ , 如果  $A = A^T$ , 则称  $A$  为对称矩阵, 通常用  $I$  表示; 如果  $A = -A^T$ , 则称  $A$  为反对称矩阵; 如果  $A = A^H$ , 则称  $A$  为共轭对称矩阵; 如果  $A = -A^H$ , 则称  $A$  为反共轭对称矩阵, 也称反厄米特矩阵。



**定义 1.111 范德蒙矩阵**

设矩阵  $\mathbf{A} \in \mathbb{C}^{m \times n}$ ,

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{m-1} & x_2^{m-1} & \cdots & x_n^{m-1} \end{bmatrix}. \quad (1.352)$$

称为范德蒙矩阵, 记为  $\mathbf{V}$ , 其中第  $j$  列是比值为  $x_j$  的等比数列,  $j = 1, 2, \dots, n$ . ♣

**定理 1.112**

范德蒙矩阵  $\mathbf{V} \in \mathbb{C}^{n \times n}$ , 其行列式

$$\det\{\mathbf{V}\} = \prod_{\Omega^j} (x_j - x_i) \equiv \prod_{i=1}^n \prod_{i=i+1}^n (x_j - x_i) \quad (1.353)$$

**定理 1.113 汉克矩阵**

设矩阵  $\mathbf{A} \in \mathbb{C}^{n+1 \times n+1}$ ,

$$\mathbf{A} = \begin{bmatrix} a_0 & a_1 & a_2 & \cdots & a_n \\ a_1 & a_2 & a_3 & \cdots & a_{n+1} \\ a_2 & a_3 & a_4 & \cdots & a_{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_n & a_{n+1} & a_{n+2} & \cdots & a_{2n} \end{bmatrix}. \quad (1.354)$$

称为汉克矩阵 (Hankel matrix), 其中  $a_{ij} = a_{i+j-2}$  ( $i, j = 1, 2, \dots, n+1$ ). 汉克矩阵是对称矩阵, 完全由其第 1 行和第  $n+l$  行的  $2n+l$  个元素确定。 ♣



### 定理 1.114 汉克矩阵

非奇异矩阵  $H \in \mathbb{C}^{n \times n}$  为汉克矩阵的重要统计是存在  $n \times n$  阶矩阵

$$C = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ p_0 & p_1 & \cdots & p_{n-2} & p_{n-1} \end{bmatrix}. \quad (1.355)$$

使得

$$CH = HC^T \quad (1.356) \heartsuit$$

### 定义 1.115 托普利兹矩阵

$$T = \begin{bmatrix} t_0 & t_1 & t_2 & \cdots & t_{n-1} \\ t_{-1} & t_0 & t_1 & \cdots & t_{n-2} \\ t_{-2} & t_{-1} & t_0 & \cdots & t_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{-n+1} & t_{-n+2} & t_{-n+3} & \cdots & t_0 \end{bmatrix}, \quad (1.357)$$

则称  $T$  为托普利兹矩阵 (Toeplitz matrix), 其中  $T$  完全由其第 1 行和第 1 列的  $2n - 1$  个元素确定。托普利兹矩阵沿平行主对角线的每一对角线上的元素是相等的, 是关于交叉对角线对称的. 显然

$$JT^TJ = T, \quad (1.358)$$

其中  $J$  为反向单位矩阵

简单的托普利兹矩阵有

$$F = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & 0 \end{bmatrix} \in \mathbb{R}^{n \times n}, \quad B = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 1 & 0 \end{bmatrix} \in \mathbb{R}^{n \times n}. \quad (1.359)$$

因它们作用于标准基  $\{e_1, \dots, e_n\}$  (见定义 (1.2.5) 时所产生的直观影响, 故分别称为前向

移位矩阵 (forward shift matrix) 和后向移位矩阵 (backward shift matrix).

### 定义 1.116 托普利兹矩阵

$T \in \mathbb{C}^{n \times n}$  为托普利兹矩阵的充分必要条件是  $T$  可以表示为

$$T = \sum_{k=1}^{n-1} t_{-k} B^k + \sum_{k=0}^{n-1} t_k F^k, \quad (1.360)$$

其中  $F$  是  $B$  由 1.3.22 给定.



### 定理 1.117

非奇异矩阵  $T \in \mathbb{C}^{n \times n}$  为托普利兹矩阵的充分必要条件是存在  $n \times n$  阶矩阵

$$C^{(1)} = \begin{bmatrix} p_{n-1} & p_{n-2} & \cdots & p_1 & p_0 \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \quad (1.361)$$

$$C^{(2)} = \begin{bmatrix} 0 & \cdots & \cdots & 0 & p_0 \\ 1 & \ddots & & \vdots & p_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & p_{n-2} \\ 0 & \cdots & 0 & 1 & p_{n-1} \end{bmatrix}. \quad (1.362)$$

使得

$$C^{(1)}T = TC^{(2)}. \quad (1.363)$$



### 定理 1.118

若矩阵  $H \in \mathbb{C}^{n \times n}$  是汉克矩阵, 则  $JH$  和  $HJ$  是托普利兹矩阵. 如  $T \in \mathbb{C}^{n \times n}$  是托普利兹矩阵, 则  $JT$  和  $TJ$  是汉克矩阵, 其中  $J$  是反向单位矩阵.



### 定义 1.119 托普利兹矩阵

如果满足条件  $A^H A = AA^H$ , 则称  $A$  为正规矩阵; 如  $A^H A = AA^H = I$ , 则称  $A$  为酉矩阵; 如果满足条件  $A^T A = AA^T = I$ , 称  $A$  为正交矩阵.



### 1.2.3 矩阵的逆

对矩阵  $A \in \mathbb{C}^{n \times n}$ , 若  $\det\{A\} \neq 0$ , 则必存在唯一的同阶方阵  $B$  使得

$$AB = BA = I \quad (1.364)$$

称  $B$  是  $A$  的逆矩阵, 记为  $A^{-1}$ . 但当  $A$  不是方阵或者  $A$  是方阵但  $\det A = 0$  时, 则上述的逆矩阵就不存在。

逆矩阵的性质:

- (1)  $(A^{-1})^{-1} = A$ ;
- (2) 若  $A = kI(k \neq 0)$ , 则  $A^{-1} = k^{-1}I$ ;
- (3) 若  $A, B$  都是  $n$  阶可逆方阵, 则  $AB$  也是可逆矩阵, 且  $(AB)^{-1} = B^{-1}A^{-1}$

#### 定义 1.120

(Sherman-Morrison 公式) 设  $A$  是一个  $n \times n$  的可逆矩阵, 并且  $x$  和  $y$  是两个  $n \times l$  矢量。若  $(A + xy^H)$  可逆, 则

$$(A + xy^H)^{-1} = A^{-1} - \frac{A^{-1}xy^H A^{-1}}{1 + y^H A^{-1}x} \quad (1.365)$$

上式可推广为矩阵之和的求逆公式:

$$\begin{aligned} (A + UBV)^{-1} &= A^{-1} - A^{-1}UB(B + BVA^{-1}UB)^{-1}BVA^{-1} \\ &= A^{-1} - A^{-1}U(I + BVA^{-1}U)^{-1}BVA^{-1} \\ &= A^{-1} - A^{-1}UBVA^{-1}(I + UBV A^{-1})^{-1}. \end{aligned} \quad (1.366)$$

或者

$$(A - UV)^{-1} = A^{-1} + A^{-1}U(I - VA^{-1}U)^{-1}VA^{-1} \quad (1.367)$$

分块矩阵的几种求逆公式:

- (1) 矩阵  $A$  可逆时, 为

$$\begin{bmatrix} A & U \\ V & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} + A^{-1}U(D - VA^{-1}U)^{-1}VA^{-1} & -A^{-1}U(D - VA^{-1}U)^{-1} \\ -(D - VA^{-1}U)^{-1}VA^{-1} & (D - VA^{-1}U)^{-1} \end{bmatrix}. \quad (1.368)$$

- 2) 矩阵  $A$  和  $D$  可逆时, 为

$$\begin{bmatrix} A & U \\ V & D \end{bmatrix}^{-1} = \begin{bmatrix} (A - UD^{-1}V)^{-1} & -A^{-1}U(D - VA^{-1}U)^{-1} \\ -D^{-1}V(A - UD^{-1}V)^{-1} & (D - VA^{-1}U)^{-1} \end{bmatrix}. \quad (1.369)$$

(3) 矩阵  $A$  和  $D$  可逆时, 为

$$\begin{bmatrix} A & U \\ V & D \end{bmatrix}^{-1} = \begin{bmatrix} (A - UD^{-1}V)^{-1} & -(A - UD^{-1}V)^{-1}UD^{-1} \\ -(D - VA^{-1}U)^{-1}VA^{-1} & (D - VA^{-1}U)^{-1} \end{bmatrix}. \quad (1.370)$$

或

$$\begin{bmatrix} A & U \\ V & D \end{bmatrix}^{-1} = \begin{bmatrix} (A - UD^{-1}V)^{-1} & -(V - DU^{-1}A)^{-1} \\ (U - AV^{-1}D)^{-1} & (D - VA^{-1}U)^{-1} \end{bmatrix}. \quad (1.371)$$

$A \in \mathbb{C}^{m \times n}, G \in \mathbb{C}^{n \times m}$  满足 Penrose - Moore 方程

$$AGA = A; \quad GAG = G; \quad (AG)^H = AG; \quad (GA)^H = GA \quad (1.372)$$

的全部或一部分, 则称  $G$  为  $A$  的广义逆矩阵。其中满足第一个条件的广义逆矩阵记作  $A^{-1}$  满足全部四个条件的广义逆矩阵记作  $A^\dagger$ .

对于任意  $A \in \mathbb{C}^{m \times n}$ , 满足 (1.372) 式条件的广义逆矩阵  $A^\dagger$  是唯一存在的, 其性质如下:

$$\begin{aligned} (A')' &= A \\ (A^H)' &= (A')^H \\ (A^T)' &= (A')^H \\ A' &= (A^H A)' A^H = A^H (A A^H)'. \end{aligned} \quad (1.373)$$

$A^\dagger$  的计算一般需要用到矩阵的分解。设  $A \in \mathbb{C}^{m \times n}, \text{rank}\{A\} = r, A, A$  的最大秩分解为  $A = BC$ , 其中  $B \in \mathbb{C}^{m \times r}, C \in \mathbb{C}^{r \times n}, \text{rank}\{B\} = \text{rank}\{C\} = \text{rank}\{A\} = r$ , 则有

$$A^\dagger = C^H (C C^H)^{-1} (B B^H)^{-1} B^H. \quad (1.374)$$

特别是

当  $A \in \mathbb{C}^{m \times n}, \text{rank}\{A\} = m$  时, 则  $A^\dagger = A^H (A A^H)^{-1}$ .

当  $A \in \mathbb{C}^{m \times n}, \text{rank}\{A\} = n$  时, 则  $A^\dagger = (A A^H)^{-1} A^H$ .

#### 1.2.4 矩阵分解

所谓矩阵分解, 就是通过线性变换, 将某个给定或已知的矩阵分解为两个或三个矩阵标准型的乘积。常用的矩阵分解方法主要有特征值分解 (EVD) 奇异值分解 (SVD)、Cholesky 分解、LU 分解、QR 分解等。

##### 1. 特征值分解

### 定义 1.121 特征值

设  $A \in C^{n \times n}$ , 如果存在数  $\lambda$  和  $n$  维非零列矢量  $x$ , 使得  $Ax = \lambda x$ , 则称  $\lambda$  是矩阵  $A$  的特征值,  $x$  是  $A$  的属于特征值  $\lambda$  的特征矢量。



特征值与特征矢量的性质:

(1) 若  $A$  是实对称矩阵或厄米特矩阵, 则其所有特征值都是实数。

(2) 对一个  $n \times n$  矩阵  $A$ :

a 若  $\lambda$  是  $A$  的特征值, 则  $\lambda$  也是  $A^T$  的特征值。

b 若  $\lambda$  是  $A$  的特征值, 则  $\lambda^*$  是  $A^H$  的特征值

c. 若入是  $A$  的特征值, 则  $\lambda + \sigma^2$  是  $A + \sigma^2 I$  的特征值。

d. 若入是  $A$  的特征值, 则  $1/\lambda$  是逆矩阵  $A^{-1}$  的特征值。

(3) 若  $A$  是实正交矩阵, 则其所有特征值为 1 或者 -1.

(4) 不同特征值  $\lambda_1, \lambda_2, \dots, \lambda_n$  对应的非零特征矢量  $u_1, u_2, \dots, u_n$  线性无关。

(5) 特征值与秩的关系:

a. 若  $n \times n$  矩阵  $A$  有  $k$  个非零特征值, 则  $\text{rank} A = k$ .

b. 若  $n \times n$  矩阵  $A$  只有一个零特征值, 则  $\text{rank} A = n - l$ .

c. 若  $\text{rank} A - \lambda I \leq n - 1$ , 则入是矩阵  $A$  的特征值。

(6) 对于一个厄米特矩阵  $A$ , 当且仅当它的所有特征值都为正(或者非负)值时, 它是正定(或半正定)的。

(7) 若  $A$  的特征值不相同, 则一定存在一个矩阵  $S$ , 使  $S^{-1}AS = D$ (对角矩阵),  $D$  的对角元素是矩阵  $A$  的特征值。

(8) 一个  $n \times n$  矩阵  $A = [a_{ij}]$  的最大特征值以该矩阵的列元素之和的最大值为界, 即

$$\lambda_{\max} \leq \max_i \left\{ \sum_{j=1}^n a_{ij} \right\} \quad (1.375)$$

(9) 若  $\lambda_1, \lambda_2, \dots, \lambda_n$  是  $n \times n$  矩阵  $A$  的特征值,  $u_1, u_2, \dots, u_n$  分别是与特征值对应的特征矢量, 则

$$A = E \Lambda E^{-1} \quad (1.376)$$

其中  $E = [u_1, u_2, \dots, u_n]$ ,  $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$

**定理 1.122**

设  $\lambda_1, \lambda_2, \dots, \lambda_n$  是  $n \times n$  矩阵  $A$  的  $n$  个特征值，则

$$\begin{aligned}\sum_{i=1}^n \lambda_i &= \sum_{i=1}^n a_{ii} = \text{tr}\{A\} \\ \prod_{i=1}^n \lambda_i &= \lambda_1 \lambda_2 \cdots \lambda_n = \det\{A\}.\end{aligned}\quad (1.377)$$

$n$  阶方阵  $A$  可逆等价于  $A$  的  $n$  个特征值全不为零。 ♡

**定理 1.123**

设  $\lambda$  是  $A$  的特征值， $t$  是一个多项式，则  $f(\lambda) = a_0 + a_1\lambda + \cdots + a_m\lambda^m$  是方阵  $f(A) = a_0I + a_1A + \cdots + a_mA^m$  的一个特征值。 ♡

## 2. 奇异值分解

**定义 1.124 奇异值**

设  $A \in \mathbb{C}^{m \times n}$ ,  $\lambda_i$  为  $A^H A$  的特征值，则称  $\sigma_i = \lambda_i^{1/2}$  ( $i = 1, 2, \dots, n$ ) 为  $A$  的奇异值。 ♣

**定理 1.125**

设  $A \in \mathbb{C}^{m \times n}$ ,  $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$  为  $A$  的奇异值，则存在酉矩阵  $U \in \mathbb{C}^{m \times m}$ ,  $V \in \mathbb{C}^{n \times n}$ , 使得

$$A = U\Sigma V^H \quad (1.378)$$

式中  $\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix}$ ,  $\Sigma_1 = \text{diag}\{\delta_1, \delta_2, \dots, \delta_r\}$ , 其对角元素  $\delta_i$  为复数，且  $|\delta_i| = \sigma_i$ ,  $r = \text{rank}\{A\}$ . ♡

令矩阵  $A$  为  $m \times n$  矩阵，并且  $r = \text{rank}A$ ,  $p = \min\{m, n\}$ . 设矩阵  $A$  的奇异值排列为

$$\sigma_{\max} = \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{p-1} \geq \sigma_p = \sigma_{\min} \geq 0 \quad (1.379)$$

矩阵的各种变形与奇异值的变化有以下关系：

(1) 矩阵  $A$  的共轭转置  $A^H$  的奇异值分解为

$$A^H = V\Sigma^H U^H, \quad (1.380)$$

即矩阵  $A$  和  $A^H$  的具有完全相同的奇异值。

(2)  $\mathbf{P}$  和  $\mathbf{Q}$  分别为  $m \times m$  和  $n \times n$  酉矩阵时,  $\mathbf{PAQ}^H$  的奇异值分解由

$$\mathbf{PAQ}^H = \tilde{\mathbf{U}}\Sigma\tilde{\mathbf{V}}^H \quad (1.381)$$

给出。其中,  $\tilde{\mathbf{U}} = \mathbf{P}\mathbf{U}$ ,  $\tilde{\mathbf{V}} = \mathbf{Q}\mathbf{V}$ 。也就是说, 矩阵  $\mathbf{PAQ}^H$  与  $\mathbf{A}$  具有相同的奇异值, 即奇异值具有酉不变性, 但奇异矢量最不同。

$\mathbf{A}^H\mathbf{A}, \mathbf{AA}^H$  的奇异值分解分别为

$$\mathbf{A}^H\mathbf{A} = \mathbf{V}\Sigma^H\Sigma\mathbf{V}^H, \quad \mathbf{AA}^H = \mathbf{U}\Sigma\Sigma^H\mathbf{U}^H \quad (1.382)$$

其中

$$\begin{aligned} \Sigma^H\Sigma &= \text{diag} \left\{ \sigma_1^2, \sigma_2^2, \dots, \sigma_r^2, \overbrace{0}^{\sim}, \dots, 0 \right\} \\ \Sigma\Sigma^H &= \text{diag} \left\{ \sigma_1^2, \sigma_2^2, \dots, \sigma_r^2, \overbrace{0}^{\sim}, \dots, 0 \right\}. \end{aligned} \quad (1.383)$$

注:  $\mathbf{A}^H\mathbf{A}, \mathbf{AA}^H$  均为厄米特矩阵, 厄米特矩阵的奇异值分解与特征值分解是一致的。

4)  $m \times n$  矩阵  $\mathbf{A}$  的奇异值分解与  $n \times m$  维 Moore - Penrose 广义逆矩阵  $\mathbf{A}^\dagger$

$$\mathbf{A}^\dagger = \mathbf{V}\Sigma^\dagger\mathbf{U}^H \quad (1.384)$$

其中  $\Sigma^+ = \begin{bmatrix} \Sigma^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$ 。

(5) 矩阵  $\mathbf{A}$  的谱范数为  $\mathbf{A}$  的最大奇异值, 即

$$\|\mathbf{A}\|_{\text{spec}} = \sigma_{\max} \quad (1.385)$$

(6) 若  $m \geq n$ , 则对于矩阵  $\mathbf{A}$ , 有

$$\begin{aligned} \sigma_{\min}(\mathbf{A}) &= \min_{x \neq 0} \left\{ \left( \frac{x^H \mathbf{A}^H \mathbf{A} x}{x^H x} \right)^{1/2} \right\} \\ &= \min_{x^H x = 1} \min_{x \in \mathbb{C}^{n \times 1}} \left\{ (x^H \mathbf{A}^H \mathbf{A} x)^{1/2} \right\}. \end{aligned} \quad (1.386)$$

(7)  $m \geq n$ , 则对于矩阵  $\mathbf{A}$ , 有

$$\begin{aligned} \sigma_{\max}(\mathbf{A}) &= \max_{x \neq 0} \left\{ \left( \frac{x^H \mathbf{A}^H \mathbf{A} x}{x^H x} \right)^{1/2} \right\} \\ &= \max_{x^H x = 1} \max_{x \in \mathbb{C}^{n \times 1}} \left\{ (x^H \mathbf{A}^H \mathbf{A} x)^{1/2} \right\}. \end{aligned} \quad (1.387)$$

(8) 若  $\mathbf{A}$  和  $\mathbf{B}$  均为  $m \geq n$  矩阵, 则对千  $1 \leq i, j \leq p, i + j \leq p + 1$ , 有

$$\sigma_{i+j-1}(\mathbf{A} + \mathbf{B}) \leq \sigma_i(\mathbf{A}) + \sigma_j(\mathbf{B}) \quad (1.388)$$

其中  $\sigma_i(\mathbf{A})$  表示矩阵  $\mathbf{A}$  的第  $i$  个大奇异值。特别地, 当  $j = l$  时,  $\sigma_i(\mathbf{A} + \mathbf{B}) \leq \sigma_i(\mathbf{A}) + \sigma_1(\mathbf{B}) (i = 1, 2, \dots, p)$  成立。

(9) 对于  $m \geq n$  矩阵  $\mathbf{A}$  和  $\mathbf{B}$ , 有

$$\sigma_{\max}(\mathbf{A} + \mathbf{B}) \leq \sigma_{\max}(\mathbf{A}) + \sigma_{\max}(\mathbf{B}). \quad (1.389)$$

### 3. Cholesky 分解

设  $\mathbf{A} \in \mathbb{C}^{n \times n}$  是正定矩阵,  $\mathbf{A} = \mathbf{G}\mathbf{A}\mathbf{G}^H$  称为矩阵  $\mathbf{A}$  的 Cholesky 分解, 其中  $\mathbf{G}$

$$\mathbf{G} = \begin{bmatrix} g_{11} & 0 & \cdots & 0 \\ g_{21} & g_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{bmatrix} \quad (1.390)$$

#### 定理 1.126

若  $\mathbf{A} \in \mathbb{C}^{n \times n}$  是正定矩阵, 则 Cholesky 分解  $\mathbf{A} = \mathbf{G}\mathbf{A}\mathbf{G}^{-1}$  是唯一的。



### 4. LU 分解

设  $\mathbf{A} \in \mathbb{C}^{m \times n}$ , 则  $\mathbf{A} = \mathbf{L}\mathbf{U}$  称为矩阵的  $LU$  分解。其中,  $\mathbf{L}$  为  $m \times m$  单位下三角矩阵(对角线为 1 的下三角矩阵),  $\mathbf{U}$  是  $\mathbf{A}$  的为  $m \times n$  上阶梯型矩阵。

#### 定理 1.127

如果  $\mathbf{A} \in \mathbb{C}^{n \times n}$  非奇异, 并且其  $LU$  分解存在, 则  $\mathbf{A}$  的  $LU$  分解是唯一的, 且  
 $\det\{\mathbf{A}\} = u_{11}u_{22} \cdots u_{nn}$  •



### 5. QR 分解

#### 定理 1.128

若  $\mathbf{A} \in \mathbb{C}^{m \times n}$ , 且  $m \geq n$ , 则存在列正交的矩阵  $\mathbf{Q} \in \mathbb{C}^{m \times n}$  和上三角矩阵  $\mathbf{R} \in \mathbb{C}^{n \times n}$  使得  $\mathbf{A} = \mathbf{Q}\mathbf{R}$ . 当  $m = n$  时,  $\mathbf{Q}$  是正交矩阵。若  $\mathbf{A}$  是非奇异的  $n \times n$  矩阵, 则  $\mathbf{R}$  的所有对角线元素均为正, 并且  $\mathbf{Q}$  和  $\mathbf{R}$  二者是唯一的。



### 1.2.5 子空间

所有  $n$  维复矢量的集合构成  $n$  维复矢量空间  $\mathbb{C}^n$ ,  $m(\leq n)$  个  $n$  维复矢量的子集合构成  $\mathbb{C}^n$  内的一个矢量子空间。

若  $S = \{u_1, u_2, \dots, u_m\}$  是矢量空间  $V$  的矢量子集合, 则  $u_1, u_2, \dots, u_m$  的所有线性组合的集合  $W$  称为由  $u_1, u_2, \dots, u_m$  张成的子空间, 为

$$\begin{aligned} W &= \text{Span}\{u_1, u_2, \dots, u_m\} \\ &= \{u; u = a_1u_1 + a_2u_2 + \dots + a_mu_m\} \end{aligned} \quad (1.391)$$

$V$  本身是  $V$  的一个子空间; 只由零矢量构成的矢量集合也是  $V$  的一个子空间, 称为零子空间  $V$  和零子空间称为  $V$  的平凡子空间; 不是平凡的子空间称为  $V$  的真子空间.

#### 定理 1.129

令  $S = \{u_1, u_2, \dots, u_m\}$  是矢量空间  $V$  的一个集合, 且  $W = \text{Span}\{u_1, u_2, \dots, u_m\}$  是由  $S$  的  $m$  个列矢量张成的子空间。

- (1) 在  $S$  内, 若有某个矢量(例如  $U_k$ )是其他矢量的线性组合, 则从  $S$  中删去矢量  $U_k$  后, 其他矢量张成的子空间仍为  $W$ .
- (2) 若  $W$  为非平凡子空间, 则在  $S$  内一定存在某个由线性无关的矢量组成的集合, 它张成子空间  $W$ . ♡

#### 定义 1.130 基

若  $W$  是由矢量集合  $S = \{u_1, u_2, \dots, u_m\}$  张成的子空间, 则  $S$  内的最大线性无关矢量组称为  $W$  的一组基. ♣

#### 定义 1.131 子空间

子空间  $W$  的任何一组基的矢量个数称为  $W$  的维数, 用符号  $\dim(W)$  表示若  $W$  的任何一组基都不是由有限个线性无关的矢量组成时, 则称  $W$  是无限维矢量子空间. ♣

#### 定义 1.132 标准基

$\{e_1, e_2, \dots, e_n\}$  是  $\mathbb{R}^n$  上的一组基矢量, 其中

$$e_1 = [1, 0, \dots, 0]^T, \quad e_2 = [0, 1, \dots, 0]^T, \quad \dots, \quad e_n = [0, 0, \dots, 1]^T \quad (1.392)$$

任一矢量  $x = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n$  可以用矢量  $e_i$  表示为  $x = \sum_{i=1}^n x_i e_i$ , 我们称

$\{e_1, e_2, \dots, e_n\} \mathbb{R}^n$  上的标准基.



### 定理 1.133

$B = \{b_1, b_2, \dots, b_n\}$  是  $n$  维矢量子空间  $\mathbf{W}$  的一组基, 则对于  $\mathbf{W}$  中的任何一个矢量  $x$ , 都存在一组唯一的标量  $c_1, c_2, \dots, c_n$  使得  $x$  可以表示为

$$x = c_1 b_1 + c_2 b_2 + \dots + c_n b_n \quad (1.393)$$

一矢量与子空间  $\mathbf{W}$  的所有矢量都正交, 则称该矢量最正交子空间  $\mathbf{W}$ , 若  $\forall a_i \in W_i, a_j \in W_j (i \neq j)$ , 恒有  $a_i \perp a_j$ , 则子空间  $W_i, W_j$  正交子空间, 记作  $W_i \perp W_j (i \neq j)$ .

特别地, 与子空间  $\mathbf{W}$  正交的所有矢量的集合组成的子空间, 称为  $\mathbf{W}$  的正交补空间, 记作  $\mathbf{W}^\perp$ , 即

$$W^\perp = \{x | x^T y = 0, \forall y \in W\}. \quad (1.394)$$

## 1.2.6 梯度分析

### 1. 实矢量的梯度算子

$n \times l$  实矢量  $\mathbf{x}$  的梯度算子定义为

$$\begin{aligned} \nabla_x &= \left[ \frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n} \right]^T = \frac{\partial}{\partial x} \\ \nabla_{x^T} &= \left[ \frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n} \right] = \frac{\partial}{\partial x^T}. \end{aligned} \quad (1.395)$$

对于  $n \times 1$  实矢量  $\mathbf{x}$ , 有一实标量函数  $f(\mathbf{x})$ , 则  $f(\mathbf{x})$  相对于  $\mathbf{x}$  和  $\mathbf{x}^T$  的梯度分别为

$$\begin{aligned} \nabla_x f(x) &= \frac{\partial f(x)}{\partial x} = \left[ \frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \dots, \frac{\partial f(x)}{\partial x_n} \right]^T \\ \nabla_{x^T} f(x) &= \frac{\partial f(x)}{\partial x^T} = \left[ \frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \dots, \frac{\partial f(x)}{\partial x_n} \right]. \end{aligned} \quad (1.396)$$

从梯度的定义式可以看出:

- (1) 一个以矢量为变元的标量函数的梯度为一矢量;
- (2) 梯度的每个分量给出了标量函数在该分量方向上的变化率。

梯度矢量反映了当变元增大时函数  $t$  的最大增大率。

对于  $n \times l$  实矢量  $\mathbf{x}$ , 有一实函数矢量  $f(\mathbf{x}) = [f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_m(\mathbf{x})]$ , 则  $f(\mathbf{x})$  相

对于  $\mathbf{x}$  的梯度定义为

$$\nabla_{\mathbf{x}} f(\mathbf{x}) = \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \frac{\partial f_2(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial f_m(\mathbf{x})}{\partial x_1} \\ \frac{\partial f_1(\mathbf{x})}{\partial x_2} & \frac{\partial f_2(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_m(\mathbf{x})}{\partial x_2} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_1(\mathbf{x})}{\partial x_n} & \frac{\partial f_2(\mathbf{x})}{\partial x_n} & \cdots & \frac{\partial f_m(\mathbf{x})}{\partial x_n} \end{bmatrix}. \quad (1.397)$$

## 2. 实矩阵的梯度算子

对于  $m \times n$  实矩阵  $\mathbf{A}$ , 有一实值函数  $f(\mathbf{A})$ , 则  $f(\mathbf{A})$  相对于  $\mathbf{A}$  的梯度 (简称梯度矩阵) 为

$$\nabla_{\mathbf{A}} f(\mathbf{A}) = \frac{\partial f(\mathbf{A})}{\partial \mathbf{A}} = \begin{bmatrix} \frac{\partial f(\mathbf{A})}{\partial a_{11}} & \frac{\partial f(\mathbf{A})}{\partial a_{12}} & \cdots & \frac{\partial f(\mathbf{A})}{\partial a_{1n}} \\ \frac{\partial f(\mathbf{A})}{\partial a_{21}} & \frac{\partial f(\mathbf{A})}{\partial a_{22}} & \cdots & \frac{\partial f(\mathbf{A})}{\partial a_{2n}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f(\mathbf{A})}{\partial a_{m1}} & \frac{\partial f(\mathbf{A})}{\partial a_{m2}} & \cdots & \frac{\partial f(\mathbf{A})}{\partial a_{mn}} \end{bmatrix}. \quad (1.398)$$

式中,  $a_{ij}$  是矩阵  $\mathbf{A}$  的第  $i$  行、第  $j$  列元素。

## 3. 迹函数的梯度矩阵

(1) 单个矩阵的迹的梯度矩阵对于  $m \times n$  矩阵  $\mathbf{W}$ , 有

$$\frac{\partial \text{tr}\{\mathbf{W}\}}{\partial \mathbf{W}} = I_m. \quad (1.399)$$

对于  $m \times n$  可逆矩阵  $\mathbf{W}$ , 有

$$\frac{\partial \text{tr}\{\mathbf{W}^{-1}\}}{\partial \mathbf{W}} = -(\mathbf{W}^{-2})^T. \quad (1.400)$$

对于两个列矢量的外积, 有

$$\frac{\partial \text{tr}\{xy^T\}}{\partial x} = \frac{\partial \text{tr}\{yx^T\}}{\partial x} = y. \quad (1.401)$$

(2) 两个矩阵乘积的迹的梯度

对于  $\mathbf{W} \in \mathbb{R}^{m \times n}, A \in \mathbb{R}^{n \times m}$ , 有

$$\frac{\partial \text{tr}\{\mathbf{W}A\}}{\partial \mathbf{W}} = \frac{\partial \text{tr}\{AW\}}{\partial \mathbf{W}} = A^T \quad (1.402)$$

当  $\mathbf{W}$  是对称矩阵时, 有

$$\frac{\partial \text{tr}\{\mathbf{W}A\}}{\partial \mathbf{W}} = \frac{\partial \text{tr}\{AW\}}{\partial \mathbf{W}} = A + A^T - \text{diag}\{A\} \quad (1.403)$$

对于  $\mathbf{W} \in \mathbb{R}^{m \times n}, A \in \mathbb{R}^{n \times m}$ , 有

$$\frac{\partial \operatorname{tr}\{\mathbf{W}^T \mathbf{A}\}}{\partial \mathbf{W}} = \frac{\partial \operatorname{tr}\{\mathbf{A} \mathbf{W}^T\}}{\partial \mathbf{W}} = \mathbf{A} \quad (1.404)$$

对于  $\mathbf{W} \in \mathbb{R}^{m \times n}$ , 有

$$\frac{\partial \operatorname{tr}\{\mathbf{W} \mathbf{W}^T\}}{\partial \mathbf{W}} = \frac{\partial \operatorname{tr}\{\mathbf{W}^T \mathbf{W}\}}{\partial \mathbf{W}} = 2\mathbf{W}. \quad (1.405)$$

对于  $\mathbf{W} \in \mathbb{R}^{m \times m}$ , 有

$$\frac{\partial \operatorname{tr}\{\mathbf{W}^2\}}{\partial \mathbf{W}} = \frac{\partial \operatorname{tr}\{\mathbf{W} \mathbf{W}\}}{\partial \mathbf{W}} = 2\mathbf{W}^T. \quad (1.406)$$

对于  $\mathbf{W} \in \mathbb{R}^{m \times n}, A \in \mathbb{R}^{n \times m}$  且  $\mathbf{W}$  非奇异, 有

$$\frac{\partial \operatorname{tr}\{\mathbf{A} \mathbf{W}^{-1}\}}{\partial \mathbf{W}} = -(\mathbf{W}^{-1} \mathbf{A} \mathbf{W}^{-1})^T. \quad (1.407)$$

(3) 三个矩阵乘积的迹的梯度

对于  $\mathbf{W} \in \mathbb{R}^{m \times n}, A \in \mathbb{R}^{m \times m}$ , 有

$$\frac{\partial \operatorname{tr}\{\mathbf{W}^T \mathbf{A} \mathbf{W}\}}{\partial \mathbf{W}} = (\mathbf{A} + \mathbf{A}^T) \mathbf{W}. \quad (1.408)$$

当  $\mathbf{W}$  是对称矩阵时, 有

$$\frac{\partial \operatorname{tr}\{\mathbf{W}^T \mathbf{A} \mathbf{W}\}}{\partial \mathbf{W}} = 2\mathbf{A} \mathbf{W}. \quad (1.409)$$

对于  $\mathbf{W} \in \mathbb{R}^{m \times n}, A \in \mathbb{R}^{n \times n}$ , 有

$$\frac{\partial \operatorname{tr}\{\mathbf{W} \mathbf{A} \mathbf{W}^T\}}{\partial \mathbf{W}} = \mathbf{W} (\mathbf{A} + \mathbf{A}^T). \quad (1.410)$$

当  $\mathbf{A}$  是对称矩阵时, 有

$$\frac{\partial \operatorname{tr}\{\mathbf{W} \mathbf{A} \mathbf{W}^T\}}{\partial \mathbf{W}} = 2\mathbf{W} \mathbf{A}. \quad (1.411)$$

对于  $\mathbf{W} \in \mathbb{R}^{m \times n}, A \in \mathbb{R}^{n \times m}$  且  $\mathbf{W}$  非奇异, 有

$$\frac{\partial \operatorname{tr}\{\mathbf{A} \mathbf{W}^{-1} \mathbf{B}\}}{\partial \mathbf{W}} = -(\mathbf{W}^{-1} \mathbf{B} \mathbf{A} \mathbf{W}^{-1})^T. \quad (1.412)$$

(4) 四个矩阵乘积的迹的梯度

对于  $\mathbf{W} \in \mathbb{R}^{m \times n}, A \in \mathbb{R}^{p \times m}$ , 有

$$\frac{\partial \operatorname{tr} \{A\mathbf{W}\mathbf{W}^T A^T\}}{\partial \mathbf{W}} = 2A^T A \mathbf{W}. \quad (1.413)$$

对于  $\mathbf{W} \in \mathbb{R}^{m \times n}, A \in \mathbb{R}^{p \times n}$ , 有

$$\frac{\partial \operatorname{tr} \{A\mathbf{W}^T \mathbf{W} A^T\}}{\partial \mathbf{W}} = 2\mathbf{W} A^T A. \quad (1.414)$$

对于  $\mathbf{W} \in \mathbb{R}^{m \times n}, A \in \mathbb{R}^{p \times m}, B \in \mathbb{R}^{m \times p}$ , 有

$$\frac{\partial \operatorname{tr} \{A\mathbf{W}\mathbf{W}^T B\}}{\partial \mathbf{W}} = (\mathbf{B}A + A^T \mathbf{B}^T) \mathbf{W}. \quad (1.415)$$

对于  $\mathbf{W} \in \mathbb{R}^{m \times n}, A \in \mathbb{R}^{p \times n}, B \in \mathbb{R}^{n \times p}$ , 有

$$\frac{\partial \operatorname{tr} \{A\mathbf{W}^T \mathbf{W} B\}}{\partial \mathbf{W}} = \mathbf{W} (BA + A^T B^T). \quad (1.416)$$

对于  $\mathbf{W} \in \mathbb{R}^{m \times n}, A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{m \times m}$ , 有

$$\frac{\partial \operatorname{tr} \{WA^T W^T B\}}{\partial \mathbf{W}} = BWA + B^T WA^T. \quad (1.417)$$

对于  $\mathbf{W} \in \mathbb{R}^{m \times n}, A, B \in \mathbb{R}^{n \times m}$ , 有

$$\frac{\partial \operatorname{tr} \{WAWB\}}{\partial \mathbf{W}} = B^T W^T A^T + A^T W^T B^T. \quad (1.418)$$

#### 4. 行列式的梯度矩阵

若矩阵  $\mathbf{W}$  的元素  $W_{ij}$  相互独立,

$$\frac{\partial \mathbf{W}}{\partial w_{ij}} = E_{ij} = \mathbf{e}_i \mathbf{e}_j^T \quad (1.419)$$

其中,  $E_{ij}$  是一个除  $(i, j)$  元素为 1 外其余元素均为 0 的矩阵;  $\mathbf{e}_i$  为 (1.392) 式给出的标准基, 其第 1 个元素为 1, 其余元素全部为 0.

若矩阵  $\mathbf{W}$  为对称矩阵, 有

$$\frac{\partial \mathbf{W}}{\partial w_{ij}} = E_{ij} + E_{ji} - \delta_{ij} E_{ij} \quad (1.420)$$

$$\text{其中 } \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

记  $n \times n$  矩阵  $\mathbf{W}$  中元素  $W_{ij}$  的余子式为  $\text{cof}(\mathbf{W};;)$ , 则有  $\det(\mathbf{W}) = \sum_{j=1}^n w_{ij} \text{cof}(W_{ij})$  ( $i = 1, 2, \dots, n$ ), 有

$$\frac{\partial \det\{\mathbf{W}\}}{\partial w} = \text{cof}(W_{ij}),$$

$\mathbf{W}$  的元素相互独立.

$$\begin{aligned} \frac{\partial \det\{\mathbf{W}\}}{\partial w_{ij}} &= \text{cof}(W_{ij}) + \text{cof}(W_{ji}) - \delta_{ij} \text{cof}(W_{ij}) \\ &= (1 - \delta_{ij}) \text{cof}(W_{ij}) + \text{cof}(W_{ji}), \end{aligned} \quad (1.421)$$

$\mathbf{W}$  为对称阵.

令  $i, j = 1, 2, \dots, n$  写成矩阵形式, 有

$$\frac{\partial \det\{\mathbf{W}\}}{\partial \mathbf{W}} = (\mathbf{W}^\#)^T = \det\{\mathbf{W}\} (\mathbf{W}^{-1})^T, \quad (1.422)$$

$\mathbf{W}$  的元素相互独立.

$$\begin{aligned} \frac{\partial \det\{\mathbf{W}\}}{\partial \mathbf{W}} &= 2\mathbf{W}^\# - \text{diag}\{\mathbf{W}^\#\} \\ &= \det\{\mathbf{W}\} (2\mathbf{W}^{-1} - \text{diag}\{\mathbf{W}^{-1}\}). \end{aligned} \quad (1.423)$$

$\mathbf{W}$  为对称阵,  $\mathbf{W}^\#$  是  $\mathbf{W}$  伴随矩阵,  $\mathbf{W}^{-1} = \mathbf{W}^\#/\det\{\mathbf{W}\}$ .

若是  $\mathbf{W}$  为非奇异矩阵, 有

$$\frac{\partial}{\partial \mathbf{W}} \log(\det\{\mathbf{W}\}) = \frac{1}{\det\{\mathbf{W}\}} \frac{\partial \det\{\mathbf{W}\}}{\partial \mathbf{W}} \quad (1.424)$$

将 (1.422) 式和 (1.423) 式的结果代入 (1.424) 式, 可得

$$\frac{\partial}{\partial \mathbf{W}} \log(\det\{\mathbf{W}\}) = (\mathbf{W}^{-1})^T \quad (1.425)$$

$\mathbf{W}$  的元素相互独立.

$$\frac{\partial}{\partial \mathbf{W}} \log(\det\{\mathbf{W}\}) = 2\mathbf{W}^{-1} - \text{diag}\{\mathbf{W}^{-1}\} \quad (1.426)$$

$\mathbf{W}$  为对称阵,  $\mathbf{W}^\#$  是  $\mathbf{W}$  伴随矩阵,  $\mathbf{W}^{-1} = \mathbf{W}^\#/\det\{\mathbf{W}\}$ .

行列式梯度其他性质还有:

(1) 满秩矩阵的行列式的梯度

$$\frac{\partial \det\{\mathbf{W}^{-1}\}}{\partial \mathbf{W}} = -\det^{-1}\{\mathbf{W}\} (\mathbf{W}^{-1})^T. \quad (1.427)$$

(2) 两个矩阵乘积的行列式的梯度

$$\begin{aligned}\frac{\partial \det\{W^T\}}{\partial W} &= 2 \det\{WW^T\} (WW^T)^{-1} W, \quad \text{rank}\{W_{m \times n}\} = m \\ \frac{\partial \det\{W^T W\}}{\partial W} &= 2 \det\{W^T W\} W (W^T W)^{-1}, \quad \text{rank}\{W_{m \times n}\} = n \\ \frac{\partial \det\{W^2\}}{\partial W} &= 2 \det^2\{W\} (W^{-1})^T, \quad \text{rank}\{W_{m \times m}\} = m\end{aligned}\quad (1.428)$$

(3) 三个矩阵乘积的行列式的梯度

$$\begin{aligned}\frac{\partial \det\{AWB\}}{\partial W} &= \det\{AWB\} A^T (B^T W^T A^T)^{-1} B^T \\ \frac{\partial \det\{W^T AW\}}{\partial W} &= 2 AW (W^T AW)^{-1}, \quad \det\{W^T AW\} > 0 \\ \frac{\partial \det\{WAW^T\}}{\partial W} &= [(WAW^T)^{-1}]^T W (A^T + A) \\ &= 2 (WAW^T)^{-1} WA, \quad A = A^T.\end{aligned}\quad (1.429)$$

### 5. 复矢量函数的梯度

首先引入复数求导的定义, 若  $f(w)$  是复数  $w$  的函数, 其中  $w = w_r + jw_i$ ,  $w^* = w_r - jw_i$ , 则  $f(w)$  对  $w$  求导定义如下:

$$\begin{aligned}\frac{df(w)}{dw} &= \frac{1}{2} \left[ \frac{\partial f(w)}{\partial w_r} - j \frac{\partial f(w)}{\partial w_i} \right] \\ \frac{df(w)}{dw^*} &= \frac{1}{2} \left[ \frac{\partial f(w)}{\partial w_r} + j \frac{\partial f(w)}{\partial w_i} \right].\end{aligned}\quad (1.430)$$

那么, 目标函数  $f(\mathbf{w})$  相对于复矢量  $\mathbf{w}$  的梯度定义为

$$\frac{\partial f(\mathbf{w})}{\partial \mathbf{w}} = \nabla_{\mathbf{w}} f(\mathbf{w}) = \left[ \frac{\partial f(\mathbf{w})}{\partial w_1}, \frac{\partial f(\mathbf{w})}{\partial w_2}, \dots, \frac{\partial f(\mathbf{w})}{\partial w_n} \right]^T. \quad (1.431)$$

而目标函数  $f(\mathbf{w})$  相对于复数共轭矢量  $\mathbf{w}^*$  的梯度常简称为共轭梯度(矢量), 定义为

$$\frac{\partial f(\mathbf{w})}{\partial \mathbf{w}^*} = \nabla_{\mathbf{w}^*} f(\mathbf{w}) = \left[ \frac{\partial f(\mathbf{w})}{\partial w_i}, \frac{\partial f(\mathbf{w})}{\partial w_2^*}, \dots, \frac{\partial f(\mathbf{w})}{\partial w_n^*} \right]^T. \quad (1.432)$$

若  $\mathbf{f}(\mathbf{w}) = [f_1(\mathbf{w}), f_2(\mathbf{w}), \dots, f_m(\mathbf{w})]$  则  $f(\mathbf{w})$  相对于复列矢量  $\mathbf{w}$  的梯度为一  $n \times m$  矩阵, 并定义为

$$\frac{\partial \mathbf{f}(\mathbf{w})}{\partial \mathbf{w}} = \begin{bmatrix} \frac{\partial f_1(\mathbf{w})}{\partial w_1} & \frac{\partial f_1(\mathbf{w})}{\partial w_2} & \dots & \frac{\partial f_1(\mathbf{w})}{\partial w_n} \\ \frac{\partial f_2(\mathbf{w})}{\partial w_1} & \frac{\partial f_2(\mathbf{w})}{\partial w_2} & \dots & \frac{\partial f_2(\mathbf{w})}{\partial w_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m(\mathbf{w})}{\partial w_1} & \frac{\partial f_m(\mathbf{w})}{\partial w_2} & \dots & \frac{\partial f_m(\mathbf{w})}{\partial w_n} \end{bmatrix}. \quad (1.433)$$

类似地，行矢量函数  $f(\mathbf{w})$  相对千复共轭列矢量  $\mathbf{w}^*$  的梯度称为共轭梯度矩阵，定义为

$$\frac{\partial f(\mathbf{w})}{\partial \mathbf{w}^*} = \begin{bmatrix} \frac{\partial f_1(\mathbf{w})}{\partial w_1^*} & \frac{\partial f_2(\mathbf{w})}{\partial w_1^*} & \cdots & \frac{\partial f_m(\mathbf{w})}{\partial w_1^*} \\ \frac{\partial f_1(\mathbf{w})}{\partial w_i^*} & \frac{\partial f_2(\mathbf{w})}{\partial w_i^*} & \cdots & \frac{\partial f_m(\mathbf{w})}{\partial w_i^*} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_1(\mathbf{w})}{\partial w_n^*} & \frac{\partial f_2(\mathbf{w})}{\partial w_n^*} & \cdots & \frac{\partial f_m(\mathbf{w})}{\partial w_n^*} \end{bmatrix}. \quad (1.434)$$

可得

$$\frac{\partial \mathbf{w}^T}{\partial \mathbf{w}} = \frac{\partial \mathbf{w}^H}{\partial \mathbf{w}^*} = \mathbf{I} \quad (1.435)$$

$$\frac{\partial \mathbf{w}^T}{\partial \mathbf{w}^*} = \frac{\partial \mathbf{w}^H}{\partial \mathbf{w}} = 0. \quad (1.436)$$

其中， $\mathbf{1}$  和  $0$  分别为单位矩阵和零矩阵。

由以上两式可得：任何一个复矢量  $\mathbf{w}$  和它的共轭矢量  $\mathbf{w}^*$  都可以当作两个独立的复变元处理，即在求梯度的过程中，复矢量  $\mathbf{w}$  相对于其共轭矢量  $\mathbf{w}^*$  可视为一常数；反之， $\mathbf{w}^*$  相对于  $\mathbf{w}$  也可视为一常数。

令  $f(\mathbf{w})$  是复矢量  $\mathbf{w}$  的实值函数。通过将  $\mathbf{w}$  和  $\mathbf{w}^*$  视为独立的变元，实目标函数  $f(\mathbf{w})$  的曲率方向由共轭梯度矢量  $\nabla_{\mathbf{w}} \cdot f(\mathbf{w})$  给出（曲率方向就是函数最大变化率方向）。

常用标量函数的共轭梯度公式：

(1) 若  $f(\mathbf{x}) = c$  为常数，则共轭梯度

$$\frac{\partial c}{\partial x^*} = 0 \quad (1.437)$$

(2) 若  $n \times l$  矢量  $a$  为与  $\mathbf{x}$  无关的常数矢量，则

$$\frac{\partial a^H x}{\partial x^*} = \mathbf{0}, \quad \frac{\partial x^H a}{\partial x^*} = a. \quad (1.438)$$

(3) 若  $\mathbf{A}$  是一个与矢量  $x, y$  无关的矩阵，则

$$\begin{aligned} \frac{\partial x^H A x}{\partial x} &= A^T x^*, \quad \frac{\partial x^H A x}{\partial x^*} = A x \\ \frac{\partial x^H A y}{\partial A} &= x^* y^T, \quad \frac{\partial x^H A x}{\partial A} = x^* x^T. \end{aligned} \quad (1.439)$$

## (4) 迹函数的共轭梯度公式

$$\frac{\partial \operatorname{tr}\{x^H y\}}{\partial x^*} = \frac{\partial \operatorname{tr}\{yx^H\}}{\partial x^*} = y \quad (1.440)$$

$$\frac{\partial \operatorname{tr}\{BA^H\}}{\partial A^*} = \frac{\partial \operatorname{tr}\{A^H B\}}{\partial A^*} = B \quad (1.441)$$

$$\frac{\partial \operatorname{tr}\{A^H\}}{\partial A^*} = I, \quad (1.442)$$

$$\frac{\partial \operatorname{tr}\{A\}}{\partial A^*} = 0 \quad (1.443)$$

$$\frac{\partial \operatorname{tr}\{A^H W A\}}{\partial A^*} = WA, \quad (1.444)$$

$$\frac{\partial \operatorname{tr}\{AWA^H\}}{\partial A^*} = AW. \quad (1.445)$$

### 1.3 习题

- 练习 1.1** 写出下列随机试验的样本空间: ① 10 只产品中有 3 只次品, 每次从中取 1 只(不放回), 直到将 3 只次品都取出, 记录抽取的次数。② 甲、乙两人下棋一局, 观察棋赛的结果。③ 口袋中有许多红色、白色、蓝色的乒乓球, 任取 4 只, 观察它们有哪几种颜色。
- 练习 1.2** 已知  $A, B, C$  为三个事件, 用  $A, B, C$  的运算关系表示下列事件: ①  $A$  发生,  $B$  与  $C$  不发生。②  $A, B, C$  都发生。③  $A, B, C$  至少有一个发生。④  $A, B, C$  不多于两个发生。
- 练习 1.3** 已知样本空间  $\Omega = \{1, 2, \dots, 10\}$ , 事件  $A = \{2, 3, 4\}$ ,  $B = \{3, 4, 5\}$ ,  $C = \{5, 6, 7\}$ , 写出下列事件的表达式: ①  $\overline{AB}$ 。②  $\overline{A}B$ 。③  $\overline{ABC}$ 。④  $\overline{A(B \cup C)}$ 。
- 练习 1.4** 我方对敌方雷达设备同时采取  $A, B, C$  三种干扰措施, 它们之间相互独立。根据以往的作战经验, 可估算出  $A, B, C$  三种干扰措施的成功率分别为 0.2、0.3 和 0.4。求: ① 恰好只有一种干扰措施成功的概率。② 敌方雷达被干扰后失效的总概率。
- 练习 1.5** 考察甲、乙两个城市十月份下雨的情况,  $A, B$  分别表示甲、乙两市出现雨天的事件。根据以往气象记录得知  $P(A) = P(B) = 0.4$ ,  $P(AB) = 0.28$ 。求  $P(A|B)$ ,  $P(B|A)$  和  $P(A \cup B)$ 。
- 练习 1.6** 信息  $A$  和信息  $B$  通过发射机发送。接收机接收时,  $A$  被误收作  $B$  的概率为 0.02; 而  $B$  被误收作  $A$  的概率为 0.01。已知发送信息  $A$  和信息  $B$  的次数比为 3:1, 若接收机收到的信息是  $A$  时, 原发送信息是  $A$  的概率为多少。
- 练习 1.7** 已知事件  $A, B$  相互独立, 分别证明:  $A$  和  $B$  相互独立;  $\bar{A}$  和  $B$  相互独立;  $\bar{A}$  和  $\bar{B}$  相互独立。

**练习 1.8** 有朋自远方来, 她乘火车、轮船、汽车或飞机来的概率分别是 0.3、0.20.1 和 0.4。如果她乘火车、轮船或汽车来, 迟到的概率分别是 0.25、0.4 和 0.1, 乘飞机来则不会迟到。结果她迟到了, 问她乘火车来的概率是多少?

**练习 1.9** 已知随机变量  $X$  的分布函数为

$$F_X(x) = \begin{cases} 0, & x < 0 \\ kx^2, & 0 \leq x \leq 1 \\ 1, & x > 1 \end{cases} \quad (1.446)$$

① 求参数  $k$ 。② 求区间概率  $P\{0.3 \leq X \leq 0.7\}$ 。③ 求概率密度  $f(X)$ 。

**练习 1.10** 已知随机变量  $X$  的概率密度为  $f_X(x) = ke^{-|x|}$ ,  $-\infty < x < +\infty$ ,  $X$  服从拉普拉斯分布, 求: ① 系数  $k$ 。②  $X$  落在区间  $(0, 1)$  内的概率。③ 随机变量  $X$  的分布函数。

**练习 1.11** 某繁忙的汽车站, 每天有大量的汽车进出。设每辆汽车在一天内出事故的概率为  $1E-4$ , 若每天有 1000 辆汽车进出汽车站, 问汽车站出事故的次数不小于 2 的概率是多少?

**练习 1.12** 已知随机变量  $(X, Y)$  的概率密度为

$$f(x, y) = \begin{cases} ke^{-(3x+4y)}, & x > 0, y > 0 \\ 0, & \text{其他} \end{cases} \quad (1.447)$$

求: ① 系数  $k$ 。②  $(X, Y)$  的分布函数。③  $P\{0 < X \leq 1, 0 < y \leq 2\}$ 。

**练习 1.13** 已知随机变量  $(X, Y)$  的概率密度为

$$f(x, y) = \begin{cases} 1, & 0 < x < 1, |y| < x \\ 0, & \text{其他} \end{cases} \quad (1.448)$$

① 求条件概率密度  $f_X(x|y)$  和  $f_Y(y|x)$ 。② 判断  $X$  和  $Y$  是否独立, 给出理由。

**练习 1.14** 已知离散型随机变量  $X$  的分布律为

$X$	3	6	7	
$P$	0.2	0.1	0.7	

(1.449)

求: ①  $X$  的分布函数。② 随机变量  $Y = 3X + 1$  的分布律。

**练习 1.15** 已知随机变量  $X$  服从标准高斯分布。求: ① 随机变量  $Y = e^X$  的概率密度。② 随机变量  $Z = |X|$  的概率密度。

**练习 1.16** 已知随机变量  $X_1$  和  $X_2$  相互独立, 概率密度分别为

$$f_{X_1}(x_1) = \begin{cases} \frac{1}{2}e^{-\frac{1}{2}x_1}, & x_1 \geq 0 \\ 0, & x_1 < 0 \end{cases}, \quad f_{X_2}(x_2) = \begin{cases} \frac{1}{3}e^{-\frac{1}{3}x_2}, & x_2 \geq 0 \\ 0, & x_2 < 0 \end{cases} \quad (1.450)$$

求随机变量  $Y = X_1 + X_2$  的概率密度。

**练习 1.17** 已知随机变量  $X, Y$  的联合分布律为

$$P\{X = m, Y = n\} = \frac{3^m 2^n e^{-5}}{m!n!}, \quad m, n = 0, 1, 2, \dots \quad (1.451)$$

求: ① 边缘分布律  $P\{X = m\}, m = 0, 1, 2, \dots$  和  $P\{Y = n\}, n = 0, 1, 2, \dots$ 。② 条件分布律  $P\{X = m, Y = n\}$  和  $P\{Y = n, X = m\}$ 。

**练习 1.18** 已知随机变量  $X_1, X_2, \dots, X_n$  相互独立, 概率密度分别为  $f_1(x_1), \dots, f_n(x_n)$ 。又随机变量

$$\begin{cases} Y_1 = X_1 \\ Y_2 = X_1 + X_2 \\ \dots \\ Y_n = X_1 + X_2 + \dots + X_n \end{cases} \quad (1.452)$$

证明: 随机变量  $Y_1, Y_2, \dots, Y_n$  的联合概率密度为  $f_Y(y_1, y_2, \dots, y_n) = f_1(y_1)f_2(y_2 - y_1) \cdots f_n(y_n - y_{n-1})$ .

**练习 1.19** 已知随机变量  $X$  服从拉普拉斯分布, 其概率密度为

$$f_X(x) = \frac{1}{2}e^{-|x|}, \quad -\infty < x < +\infty \quad (1.453)$$

求其数学期望与方差。

**练习 1.20** 已知随机变量  $X$  的可能取值为  $-4, -1, 2, 3, 4$ , 且每个值出现的概率均为  $1/5$ 。求: ① 随机变量  $X$  的数学期望和方差。② 随机变量  $Y = 3X^2$  的概率密度。③  $Y$  的数学期望和方差。

**练习 1.21** 已知随机变量  $X$  服从  $(0, 1)$  的均匀分布, 随机变量  $Y$  服从高斯分布  $Y \sim N(3X, 1)$ 。求: ① 条件数学期望  $E[Y|X = x]$ 。② 条件数学期望  $E[Y|X]$ 。

**练习 1.22** 已知两个随机变量  $X, Y$  的数学期望为  $m_X = 1, m_Y = 2$ , 方差为  $\sigma_X^2 = 4, \sigma_Y^2 = 1$ , 相关系数  $\rho_{XY} = 0.4$ 。现定义新随机变量  $V, W$  为

$$\begin{cases} V = -X + 2Y \\ W = X + 3Y \end{cases} \quad (1.454)$$

求  $V, W$  的期望, 方差以及它们的相关系数。

- ✉ **练习 1.23** 已知随机变量  $X, Y$  满足  $Y = aX + b, a, b$  皆为常数。证明: ①  $C_{XY} = a\sigma_X^2$ 。②  $\rho_{XY} = \begin{cases} 1, & a > 0 \\ -1, & a < 0 \end{cases}$  ③ 当  $m_X \neq 0$  且  $b = -\frac{aE[X^2]}{E[X]}$  时, 随机变量  $X, Y$  正交。

- ✉ **练习 1.24** 已知二维随机变量  $(X, Y)$  的联合概率密度为

$$f_{XY}(x, y) = \begin{cases} \frac{xy}{9}, & 0 < x < 2, 0 < y < 3 \\ 0, & \text{其他} \end{cases} \quad (1.455)$$

判断随机变量  $X$  和  $Y$  是否正交、不相关和独立, 给出理由。

- ✉ **练习 1.25** 已知随机变量  $X, Y$  相互独立, 分别服从参数为  $\lambda_1$  和  $\lambda_2$  的泊松分布。① 求随机变量  $X$  的期望和方差。② 证明  $Z = X + Y$  服从参数为  $\lambda_1 + \lambda_2$  的泊松分布。

- ✉ **练习 1.26** 已知随机变量  $X, Y$  的联合特征函数为

$$Q_{XY}(u, v) = \frac{6}{6 - 2ju - 3jv - uv}. \quad (1.456)$$

求: ① 随机变量  $X$  的特征函数。② 随机变量  $Y$  的期望和方差。

- ✉ **练习 1.27** 已知随机变量  $X, Y$  对于任意  $n \geq 1, m \geq 1$  都有  $E[XY] = E[X^m]E[Y]$ , 证明  $X, Y$  独立

- ✉ **练习 1.28** 已知两个独立的随机变量  $X, Y$  的特征函数分别是  $Q_X(u)$  和  $Q_Y(u)$ , 求随机变量  $Z = 3(X + 1) + 2(Y - 4)$  的特征函数  $Q_Z(u)$ .

- ✉ **练习 1.29** 已知二维高斯变量  $(X_1, X_2)$  中, 高斯变量  $X_1, X_2$  的期望分别为方差分别为  $\sigma_1^2, \sigma_2^2$ , 相关系数为  $\rho$ 。令

$$Y_1 = \frac{X_1 - m_1}{\sigma_1}, Y_2 = \frac{1}{\sqrt{1 - \rho^2}} \left( \frac{X_2 - m_2}{\sigma_2} - \rho \frac{X_1 - m_1}{\sigma_1} \right). \quad (1.457)$$

① 写出二维高斯变量  $(X_1, X_2)$  的概率密度和特征函数的矩阵形式, 并展开。② 证明  $(Y_1, Y_2)$  相互独立, 皆服从标准高斯分布。

- ✉ **练习 1.30** 已知二维高斯变量  $(X_1, Y_1)$  的两个分量相互独立, 期望皆为 0, 方差皆为  $\sigma^2$ 。令

$$\begin{cases} Y_1 = \alpha X_1 + \beta X_2 \\ Y_2 = \alpha X_1 - \beta X_2 \end{cases} \quad (1.458)$$

其中  $\alpha \neq 0, \beta \neq 0$  为常数。① 证明:  $(Y_1, Y_2)$  服从二维高斯分布。② 求  $(Y_1, Y_2)$  的均值和协方差矩阵。③ 证明:  $Y_1, Y_2$  相互独立的条件为  $\alpha = \pm\beta$ 。

✉ 练习 1.31 已知三维高斯随机矢量  $\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$  的均值为常矢量  $\sigma$ , 其方差阵为

$$\mathbf{B} = \begin{bmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 4 \end{bmatrix}, \quad (1.459)$$

证明:  $X_1, X_2 - X_1, X_1/3 + 2X_2/3 + X_3$  相互独立。

✉ 练习 1.32 已知三维高斯随机变量  $(X_1, X_2, X_3)$  各分量相互独立, 皆服从标准高斯分布。求  $Y_1 = X_1 + X_2$  和  $Y_2 = X_1 + X_3$  的联合特征函数。

✉ 练习 1.33 已知随机变量  $X$  服从拉普拉斯分布, 其概率密度为

$$f_X(x) = \frac{1}{2}e^{-|x|}, \quad -\infty < x < +\infty \quad (1.460)$$

求其特征函数。

# 2

## 随机信号的时域分析

信号是个随时间、空间或其他某个参量变化的、携带某种信息的物理量。例如大气层中的温度信号, 是个温度随高度变化的物理量; 电视中的图像信号, 是个亮度随平面坐标  $(x, y)$  变化的物理量; 而通常遇到最多的时间信号, 则是随时间变化的物理量。在对信号进行分析之前, 首先要给信号一个数学描述, 即信号的数学模型可以用确定的数学关系描述的信号, 称为确定信号。如正弦电压信号, 其信号的幅度随时间作规律性变化; 又如电容器通过电阻放电时, 电容器两端电位差随时间的变化也存在规律; 对于真空中的自由落体运动, 其速度变化也有规律。对于这些实验, 无论试验重复多少次, 其结果都完全相同, 均可用一个确定的时间函数  $g(t)$  来描述。

与确定信号相反, 不能用确定的数学关系描述的信号, 称为随机信号。以时间为参量的随机信号无处不在, 它的特点是随时间作无规律的、随机性的变化。如某条路上每天 24 小时行驶的车辆数目的变化, 某海湾每天 24 小时海浪高度的变化等, 这类信号的变化具有随机性, 即不确定性。这次观察的结果与上次观察的结果可能完全不同。因此找不到一个确定的数学关系 (或确定的时间函数) 来描述它, 然而人们发现, 随机信号的统计规律却是确定的, 所以在数学上, 人们用统计学的方法建立了随机信号的数学模型—“随机过程”。

随机过程理论产生于 20 世纪初, 是因统计物理学、生物学、通信与控制、管理科学等领域的研究需要而逐步发展起来的, 特别是在预测与控制领域中出现的大量的随机过程问题, 这些问题也是随机过程理论发展的重要推动力。同时, 随机过程理论的发展又为研究人员在上述领域中研究随机现象提供了数学模型, 奠定了数学基础。

第 2.1 节主要介绍随机过程的基本内容。在 1.2 节中, 线性代数的知识将是我们阐述的重点, 需要特别说明的是由于本书侧重于数学知识在统计信号处理中的应用,

在概率论概要章节，我们引入了“随机变量”的概念。随机试验的结果或呈离散取值的特性，或呈连续取值的特性，相应的随机变量亦有离散随机变量及连续随机变量。

在自然界中还存在着另一类随机现象，与这类随机现象相联系的随机事件不能用一般的单维或多维随机变量去描述。如：用  $X(t)$  表示  $t$  时刻以前某通信站接到的呼唤次数，显然，对于某个固定的  $t$  而言， $X(t)$  是一个随机量，但  $X(t)$  这类随机量将随着  $t$  的变化而变化。这里，问题所涉及的随机现象必须用一族与  $t$  有关的无限维随机变量才能描述。

把电网电压看作是随时间变化而变化的随机量  $V(t)$ ，它也涉及一种必须用一族与  $t$  有关的无限维的随机变量才能描述的随机现象。

雷达接收机的噪声输出也是一种必须用无限维的随机变量才能描述的随机现象，不妨用  $n(t)$  去描述这种随机现象。对于一台具体的雷达接收机来说，每次测到的接收机输出的噪声电压各不相同，可以用  $n_1(t), n_2(t), n_3(t), \dots$  代表， $n_i(t) (i = 1, 2, \dots)$  称为样本函数，样本函数的全体所构成的集合则称为随机过程  $\{n(t)\}$ ，随机过程中的每一样本函数又称为随机过程的一个现实。

## 2.1 随机过程

### 2.1.1 随机过程的基本概念

所有试验结果  $\zeta$  的集合，称之为样本空间  $\Omega$ 。尽管在测量之前，不能事先确定哪条波形将会出现，但可以确定必为波形集  $x_1(t), x_2(t), \dots, x_m(t)$  中的一个。这是一个典型的随机过程模型。

尽管每次测量的结果可能各不相同，但每次的结果却是一个确定的函数

$$X(t, \zeta_k) = x_k(t),$$

因此，如果能把每个结果用一个确定的函数  $x_k(t)$  来描述，那么，所有这些确定函数的总集  $\{x_1(t), x_2(t), \dots, x_m(t)\}$  就可以描述该随机过程。

仿照随机变量的定义，随机过程的定义如下：

#### 定义 2.1 随机过程

对随机试验  $E$  中的每个结果  $\zeta_k$ ，总有一实时间函数  $X(t, \zeta_k)$  与其对应，而对应于所有不同的试验结果  $\zeta \in \Omega$ ，得到的一簇时间函数  $\{X(t, \zeta_1), X(t, \zeta_2), \dots, X(t, \zeta_m)\}$  的总体  $X(t, \zeta)$ 。称之为随机过程 (stochastic process 或 random process)。簇中的每一个确定函数  $X(t, \zeta_k) = x_k(t)$  为随机过程的样本函数，所有可能结果的集合

$\{x_1(t), x_2(t), \dots, x_m(t)\}$  构成了随机过程的样本函数空间，其中  $t \in T$ ,  $T$  是观测区间或时间域，它可以是实数集或整个时间轴。 $\zeta \in \Omega$ ,  $\Omega$  是随机试验的样本空间。



### 例 2.1.1

随机过程的实例如股票和汇率的波动、语音信号、视频信号、体温的变化，还包括布朗运动和随机徘徊等。

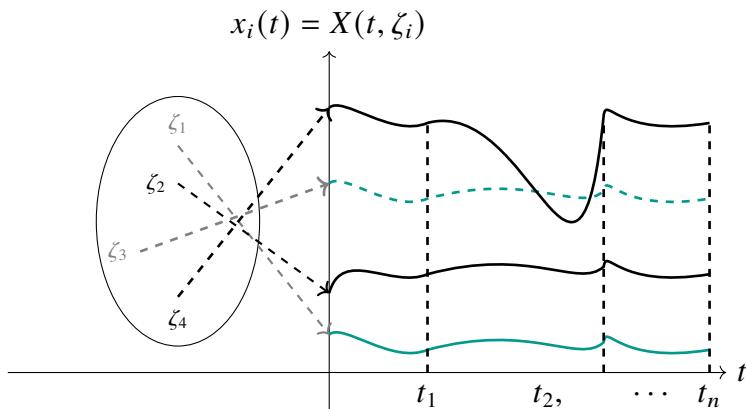


图 2-1 典型的随机过程模型

从上述定义中看到，一个随机过程  $X(t, \zeta)$  ( $t \in T, \zeta \in \Omega$ ) 实际上是时间  $t$  和随机结果  $\zeta$  两个变量的函数。如图 2-1 所示，对于固定时间  $t_i$ ,  $X(t_i, \zeta)$  是定义于概率空间  $\Omega$  上  $\zeta$  的函数，它所有试验结果  $\zeta$  的集合称为样本空间。

尽管在测量之前，不能事先确定哪条波形将会出现，但可以确定必为所有波形  $x_1(t), x_2(t), \dots, x_m(t)$  中的一个，这是一个典型的随机过程模型。

尽管每次测量的结果可能各不相同，但每次的结果却是一个确定的函数  $X(t, \zeta_k) = x_k(t)$ 。因此，如果能把每个结果用一个确定的函数  $x_k(t)$  来描述，那么，所有这些确定函数的总体  $\{x_1(t), x_2(t), \dots, x_m(t)\}$  就可以描述该随机过程。

其每做一次随机试验，随机试验的结果应是某一个随机现实，每一次随机试验之前，其试验结果究竟属于哪一种随机现实，事先不能预测。在数学上，随机过程也可以定义为：

**定义 2.2 随机过程**

设  $T \subset \mathbb{R}^1$ ,  $T$  是实数集合, 有穷或无穷, 可列或不可列。对于每一个  $t \in T$ ,  $x_t(\omega)$  为一随机变量, 其中  $\omega$  代表某概率空间  $(\Omega, F, P)$  的元素.  $\{x(t, \omega), t \in T\}$  称为随机过程。这类问题的随机现象用随机过程才能描述, 通常将  $\{x_t(\omega), t \in T\}$  简记为  $\{x(t)\}, t \in T$ .

**2.1.2 随机过程的分类**

随机过程的分类方法很多, 下面列出几种常见的分类方法:

1. 按随机过程  $X(t)$  的时间和状态的连续性来分类:

**定义 2.3 随机过程**

样本函数在时间上连续的随机过程称为“随机过程”。



样本函数在时间上离散的随机过程称为“随机序列”。

**定义 2.4 随机序列**

设  $Z$  为整数集, 若对于每一整数  $n, n \in Z$ , 均有定义在概率空间  $(\Omega, F, P)$  上的一个随机变量  $X(\zeta, n) (\zeta \in \Omega)$  与之对应, 则称依赖于参数  $n$  的一列随机变量  $X(\xi, n) (n = 1, 2, \dots, m, \dots)$  为离散时间的随机过程或“随机序列”, 记为  $\{X(\zeta, n), \zeta \in \Omega, n \in Z\}$ , 简记为  $\{X(n)\}$ 。



若用序号  $n$  取代随机过程  $X(t)$  的时间  $t$ , 得到一串随  $n$  变化的随机变量序列  $X(1), X(2), \dots, X(n), \dots$ , 即是随机序列  $\{X(n)\}$ 。

- (1) 连续型随机过程—时间连续、状态连续

**例 2.1.2**

接收机输出的噪声为连续型随机过程, 其样本函数  $x(t)$  在时间上是连续的, 状态的取值也是连续的。



如图 2-2 所示, 任意时刻  $t$  的状态  $X$ , 是连续型随变量。

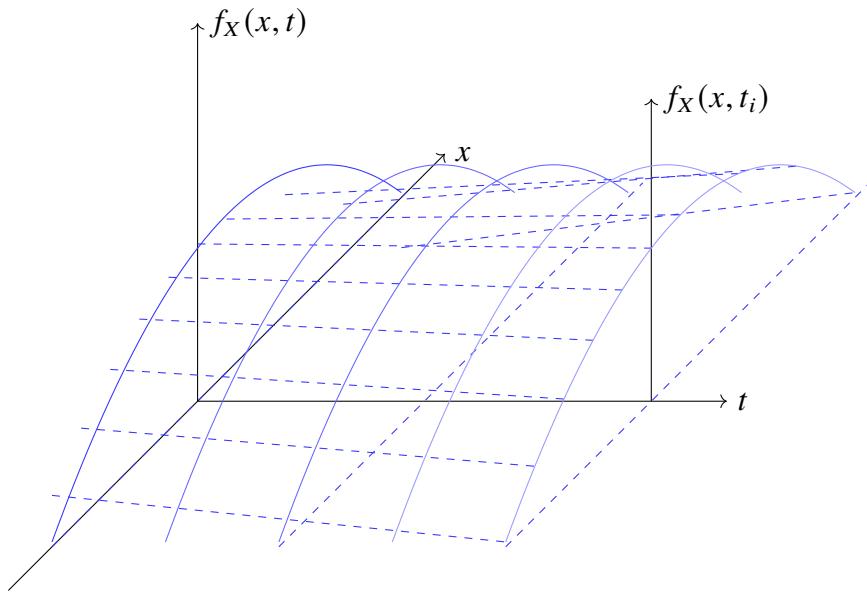


图 2-2 连续型随机过程

## (2) 离散型随机过程—时间连续、状态离散

例如，上述噪声通过一限幅器输出，它的样本函数  $y(t)$  在时间上是连续的，但状态却取离散值，如图 2-3 所示，该过程  $Y(t)$  在任意时刻的状态  $Y$  是离散型随机变量。

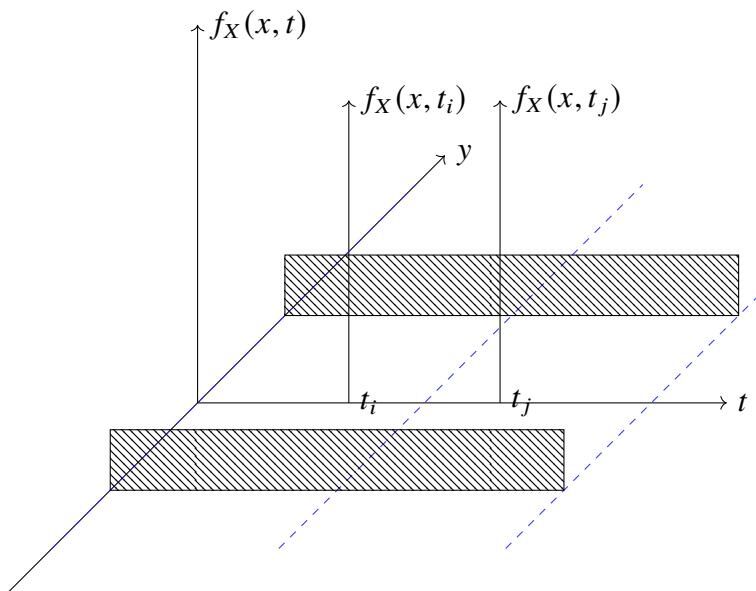


图 2-3 离散型随机过程

### (3) 连续随机序列—时间离散、状态连续

在时间  $\{1, 2, \dots, n, \dots\}$  上测量到的噪声为连续随机序列。这种随机序列  $\{Z(n)\}$  可看成是对连续型随机过程  $X(t)$  等间隔采样(时域离散化)的结果, 所以它的样本函数  $z(m)$  在时间上是离散的, 状态则是连续取值。如图 2-4 所示, 连续型随机序列  $\{Z(n)\}$  在任意时刻  $t$  的状态  $Z$  都是连续型随机变量。

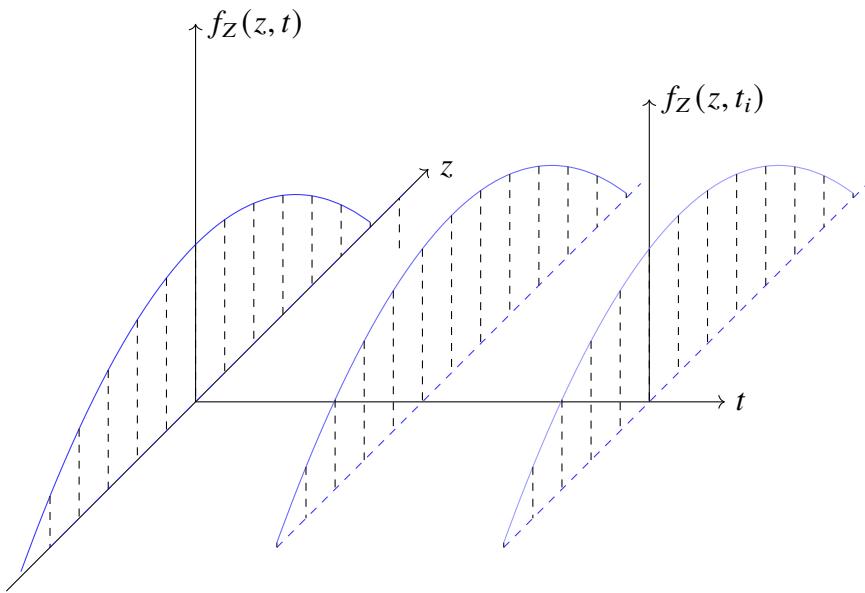


图 2-4 连续随机序列的一个样本函数

### (4) 离散随机序列(随机数字信号) 时间离散、状态离散

常见的随机序列  $\{W(n)\}$  可看成是对连续型随机过程等间隔采样, 并将采样值量化分层的结果。更一般的是, 对连续型随机过程采用其他采样方式, 并将采样值量化分层。如图 2-5 所示, 它的样本函数  $\{k(n)\}$  在时间上是离散的, 状态取值也是离散的。因此, 离散型随机序列  $\{W(n)\}$  在任意时刻  $t$  的状态  $W$ , 都是离散型随机变量。

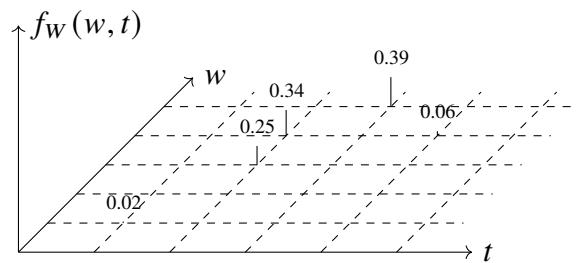


图 2-5 离散随机序列的一个样本函数

可见,最基本的是连续型随机过程,其他三种均可以通过对它进行采样量化、分层采样得到,故本书主要介绍连续型随机过程。

## 2. 按随机过程的概率分布、统计特性进行分类

按随机过程统计特性有无平稳性,分为: 平稳过程和非平稳过程; 也可按随机过程的概率分布,可分为: 高斯(正态)过程、瑞利过程、马尔可夫过程、泊松过程、维纳过程等; 还可按随机过程在频域的带宽分为: 宽带随机过程和窄带随机过程、白噪声随机过程和色噪声随机过程等等。本章将重点介绍平稳随机过程(随机过程统计特性分类)和高斯过程(概率分布分类),其他将在后面几章中进行介绍。

### (1) 确定的随机过程

如果随机过程  $X(t)$  的任意一个样本函数的未来值,都能由过去的观测值确定,即样本函数有确定的形式,则称此类过程为确定的随机过程。

**例 2.5** 正弦随机信号  $X(t) = A \cos(\Omega t + \Phi)$ , 尽管式中振幅  $A$ 、角频率  $\Omega$  和相位  $\Phi$  是随机变量,但对于任一次试验结果  $\zeta_k$  而言,随机变量  $A, \Omega, \Phi$ ,仅取某个具体的值  $a_k, \omega_k, \varphi_k$ ,相应的样本函数  $x_k(t) = a_k \cos(\omega_k t + \varphi_k)$  是一个确定的函数,都能由  $t_i$  以前出现的波形来确定  $t$  以后将会出现的波形,如图 2-6(a) 所示。

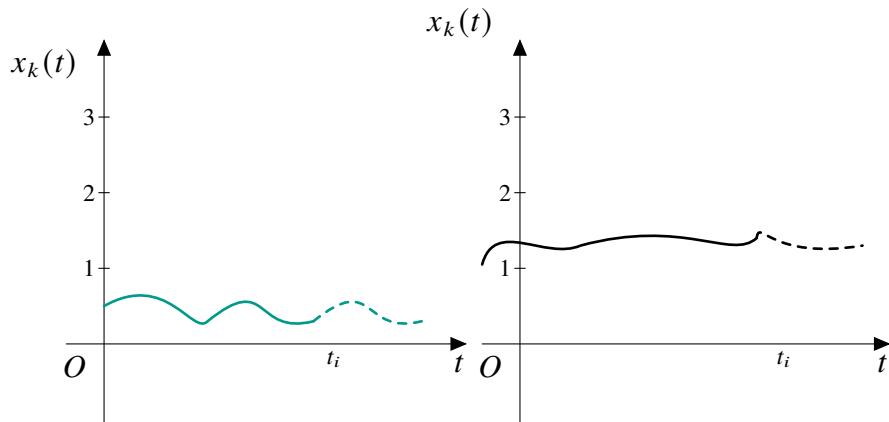


图 2-6 离散随机序列的一个样本函数(左: 确定性过程, 右: 不确定性过程)

### (2) 不确定的随机过程

如果随机过程  $X(t)$  的任意一个样本函数的未来值,都不能由过去的观测值确定,即样本函数无确定形式,如图 2-6(b) 所示。对某次实验而言,虽然样本  $x_k(t)$  在  $t$  以前的一段波形已出现,但仍不能确定在  $t$  以后将会出现什么波形。此类过程称为不确定的随机过程。

### 2.1.3 随机过程的分布

严格地说, 若要通过图形表示一个随机过程, 必须如图 2-1 中那样, 画出它所有的样本函数。但为了便于说明, 暂且将随机过程  $X(t)$  描绘成一条曲线, 如图 2-7 所示。图 2-7 中曲线上的每一点都代表过程的一个状态(随机变量)。

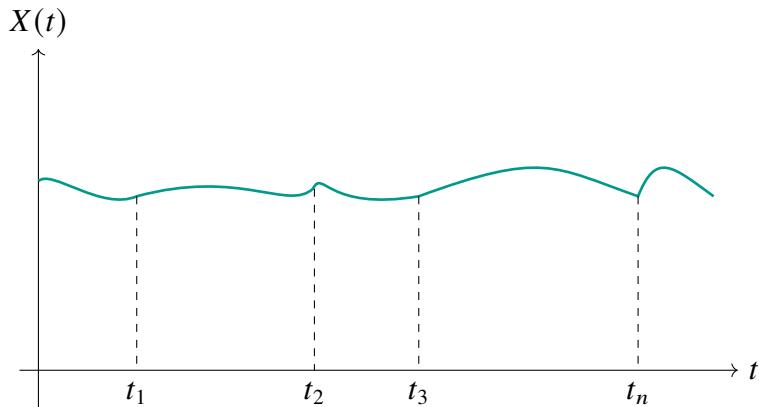


图 2-7 离散随机序列的一个样本函数

用记录器记录一个随机过程  $X(t)$  时, 只能记录  $X(t)$  在确定时刻  $t_1, t_2, \dots, t_n$  下的状态  $X(t_1), X(t_2), \dots, X(t_n)$ , 所以可以用多维随机变量  $[X(t_1), X(t_2), \dots, X(t_n)]$  来近似描述随机过程  $X(t)$ 。时间间隔  $\Delta t = t_i - t_{i-1}$  越小, 多维随机变量的随机过程的描述就越精确。在  $\Delta t \rightarrow 0$  且  $n \rightarrow \infty$  时, 随机过程的概念可以作为多维随机变量的概念在维数无穷多(不可列)的情况下自然推广。

有了上述多维随机变量对随机过程的描述, 可用研究随机变量的方法, 给出描述随机过程统计特性, 比如分布函数和概率密度。

#### 1. 一维分布函数

随机过程  $X(t)$ , 对任一固定时刻  $t_1 \in T$ , 其状态  $X(t_1)$  是一维随机变量

$$F_X(x_1; t_1) = P\{X(t_1) \leq x_1\}. \quad (2.1)$$

它表示过程  $X(t)$  在  $t_1$  时刻的状态  $x(t_1)$  取值小于  $x_1$  的概率, 如图 2-7 所示。

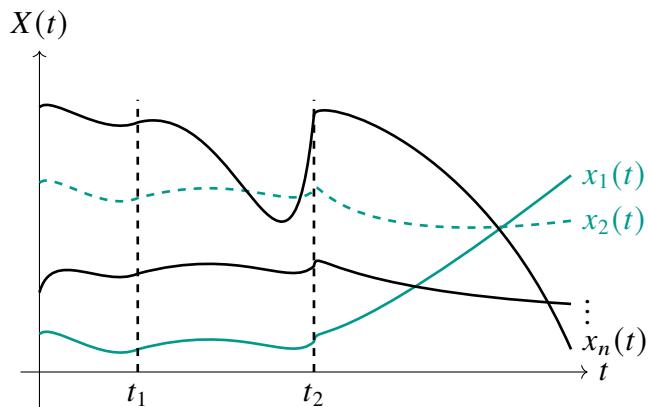


图 2-8 随机过程的分布

**定义 2.6** 随机过程  $X(t)$  的一维分布函数

若将  $x_1, t_1$  作为变量, 那么  $F_X(x_1, t_1)$  是  $x_1$  和  $t_1$  的二元函数, 可写成  $F_X(x, t)$ ; 又  $t \in T$ , 分布函数  $F_X(x; t)$  反映了过程  $X(t)$  在整个时间段  $T$  上的所有一维状态的分布状况。所以定义

$$F_X(x; t) = P\{X(t) \leq x\}, t \in T \quad (2.2)$$

为随机过程  $X(t)$  的一维分布函数。



类似于随机变量的分布函数, 也可以定义随机过程  $X(t)$  的一维概率密度。

**定义 2.7** 随机过程  $X(t)$  的一维概率密度

如果  $F_X(x; t)$  对  $x$  的偏导数存在, 则称

$$f_X(x; t) = \frac{\partial F_X(x; t)}{\partial x} \quad (2.3)$$

为随机过程  $X(t)$  的一维概率密度。 $f_X(x; t)$  也是  $x$  和  $t$  的二元函数。



显然, 随机过程的一维分布函数和一维概率密度具有和一维随机变量的分布函数和概率密度一样的各种性质, 所不同的是它们同时还是时间  $t$  的函数。

**注 2.8.** 一维分布函数和一维概率密度仅给出了随机过程最简单的概率分布特性, 它们只能描述随机过程任一时刻单一状态的统计特性, 而不能反映随机过程各个时刻的多个状态之间的联系。

## 2. 二维分布

随机过程  $X(t)$  在任意的两个固定时刻  $t_1, t_2$  的状态  $X(t_1), X(t_2)$  构成随机变量  $[X(t_1), X(t_2)]$ , 它们的分布函数为

$$F_X(x_1, x_2, t_1, t_2) = P\{X(t_1) \leq x_1, X(t_2) \leq x_2\} \quad (2.4)$$

表示随机过程  $X(t)$  在  $t_1, t_2$  两个不同时刻的两个状态  $X(x_1), X(x_2)$  的取值分别小于  $x_1, x_2$  的概率, 如图 2-8 所示。

#### 定义 2.9 随机过程 $X(t)$ 的二维分布函数

如果  $x_1, x_2, t_1, t_2$  均为变量, 那么分布函数  $F_X(x_1, x_2; t_1, t_2)$  就是  $x_1, x_2, t_1, t_2$  的四元函数, 又因  $t_1 \in T, t_2 \in T$ , 则  $F_X(x_1, x_2; t_1, t_2)$  反映了随机过程  $X(t)$  在整个时间段  $T$  上的任意两个状态间的联合分布状况, 所以定义  $F_X(x_1, x_2; t_1, t_2)$  为随机过程  $X(t)$  的二维分布函数。

#### 定义 2.10 随机过程 $X(t)$ 的二维概率密度

如果  $F_X(x_1, x_2; t_1, t_2)$  对  $x_1, x_2$  的二阶混合偏导数存在, 则称

$$f_X(x_1, x_2; t_1, t_2) = \frac{\partial^2 F_X(x_1, x_2; t_1, t_2)}{\partial x_1 \partial x_2} \quad (2.5)$$

为随机过程  $X(t)$  的二维概率密度。

由于二维分布描述了随机过程在任意两个时刻的状态之间的联系, 并可通过积分求得两个一维概率密度, 即一维边缘概率密度  $f_X(x_1; t_1)$  和  $f_X(x_2; t_2)$ 。因此, 随机过程的二维分布比一维分布含有更多的信息, 对随机过程统计特性的描述也更细致。但是, 二维分布还不能反映随机过程两个以上状态之间的联系, 不能完整地反映出随机过程的全部统计特性。需要给出  $n$  维分布的刻画。

### 3. $n$ 维分布

随机过程  $X(t)$  在任意  $n$  个时刻  $t_1, t_2, \dots, t_n$  的状态  $X(t_1), X(t_2), \dots, X(t_n)$  构成了  $n$  维随机变量  $[X(t_1), X(t_2), \dots, X(t_n)]$ , 即随机矢量  $\mathbf{X}$  表示。用类似上面的方法, 可定义随机过程  $X(t)$  的  $n$  维分布函数和  $n$  维概率密度, 分别为

#### 定义 2.11 随机过程 $X(t)$ 的 $n$ 维分布函数和 $n$ 维概率密度

$$F_X(x_1, x_2, \dots, x_n, t_1, t_2, \dots, t_n) = P\{X(t_1) \leq x_1, X(t_2) \leq x_2, \dots, X(t_n) \leq x_n\}. \quad (2.6)$$

$$f_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \frac{\partial^n F_X(x_1, x_2, \dots, x_n, t_1, t_2, \dots, t_n)}{\partial x_1 \partial x_2 \dots \partial x_n}. \quad (2.7)$$

显然,  $n$  维分布描述了随机过程在任意  $n$  个时刻的  $n$  个状态之间的联系, 比其低维分布含有更多的信息, 对随机过程统计特性的描述也更加细致。

若随机过程的观测点取得越多, 维数  $n$  越大, 则对随机过程的统计特性描述得就越细致。从理论上来说, 只有维数  $n$  为无限多时, 才能完整地描述随机过程  $X(t)$  的统计特性。类似于多维随机变量, 随机过程  $X(t)$  的  $n$  维分布具有下列性质:

$$1^{\circ} F_X(x_1, x_2, \dots, -\infty, \dots, x_n; t_1, t_2, \dots, t_i, \dots, t_n) = 0. \quad (2.8)$$

$$2^{\circ} F_X(\infty, \infty, \dots, \infty, t_1, t_2, \dots, t_n) = 1. \quad (2.9)$$

$$3^{\circ} f_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) \geq 0. \quad (2.10)$$

$$4^{\circ} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) dx_1 dx_2 \cdots dx_n = 1. \quad (2.11)$$

$$\begin{aligned} 5^{\circ} & \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X(x_1, \dots, x_m, x_{m+1}, \dots, x_n; t_1, \dots, t_m, t_{m+1}, \dots, t_n) dx_{m+1} \cdots dx_n \\ &= f_X(x_1, \dots, x_m; t_1, \dots, t_m). \end{aligned} \quad (2.12)$$

6° 如果  $X(t_1), X(t_2), \dots, X(t_n)$  统计独立, 则有  $f_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = f_X(x_1; t_1) f_X(x_2; t_2) \cdots f_X(x_n; t_n)$ 。由于  $n$  越大描述起来越困难, 因此在许多实际应用中, 一般只取二维情况。

#### 2.1.4 随机过程的数字特征

虽然随机过程的分布能较全面地描述整个过程的统计特征, 但要确定一个随机过程的高维分布通常比较困难, 分析处理起来也很复杂。在实际应用中, 往往只需要知道随机过程的几个常用统计平均量(即数字特征), 就能满足要求。随机过程的数字特征既能描述随机过程的重要统计特征, 又便于实际的测量和运算。随机过程常用的基本数字特征有数学期望、方差和相关函数等。

由于随机过程是随时间变化的随机变量, 因此随机过程的数字特征可以由随机变量的数字特征演变而来。

下面介绍随机过程的这些基本数字特征。

##### 1. 一维数字特征

###### (1) 数学期望

对任一固定时刻  $t$ ,  $X(t)$  代表一个随机变量, 它的随机取值  $x(t)$  ( $t$  固定) 简记为 “ $x$ ”, 根据随机变量数学期望的定义, 可得  $X(t)$  的数学期望为

$$E[X(t)] = \int_{-\infty}^{\infty} x f_X(x; t) dx = m_X(t). \quad (2.13)$$

当  $t$  是一个时间变量时,  $X(t)$  代表的是随机过程, 因此定义  $E[X(t)]$  为随机过程的数学期望。由于  $m_X(t)$  是随机过程  $X(t)$  的所有样本函数在  $t$  时刻所取的样本值  $(x_1, \dots, x_m)$  的统计平均(集平均), 随  $t$  而变化, 是时间  $t$  的确定函数。如图 2-9 所示, 粗实线表示随机过程的各个样本函数, 点线表示数学期望。由图可见,  $m_X(t)$  是随机过程  $X(t)$  的所有样本函数在各个时刻摆动的中心, 是  $X(t)$  在各个时刻的状态的概率质量分布的“重心位置”。

**例 2.12** 如果讨论的随机过程是接收机输出端的噪声电压, 这时数学期望  $m_X(t)$  就是次噪声电压的瞬时统计平均值。

## (2) 均方值与方差

对任一固定时刻  $t$  来讲,  $X(t)$  是一个随机变量。

### 定义 2.13 随机过程 $X(t)$ 的均方值

由随机变量二阶原点矩的定义, 有

$$\begin{aligned} E[X^2(t)] &= \int_{-\infty}^{\infty} x^2 f_X(x; t) dx \\ &= \phi_X^2(t). \end{aligned} \quad (2.14)$$

当  $t$  是一个时间变量时,  $X(t)$  代表的是随机过程,  $E[X^2(t)]$  定义为随机过程  $X(t)$  的均方值。

### 定义 2.14 方差和均方差

由随机变量二阶原点矩的定义, 可得

$$\begin{aligned} D[X(t)] &= E\{[X(t) - m_X(t)]^2\} = \int_{-\infty}^{\infty} [x - m_X(t)]^2 f_X(x; t) dx \\ &= \sigma_X^2(t). \end{aligned} \quad (2.15)$$

当  $t$  是一个时间变量时, 称  $D[X(t)]$  为随机过程  $X(t)$  的方差。而方差  $D[X(t)]$  的正平方根

$$\sqrt{D[X(t)]} = \sigma_X(t), \quad (2.16)$$

称为随机过程  $X(t)$  的均方差。

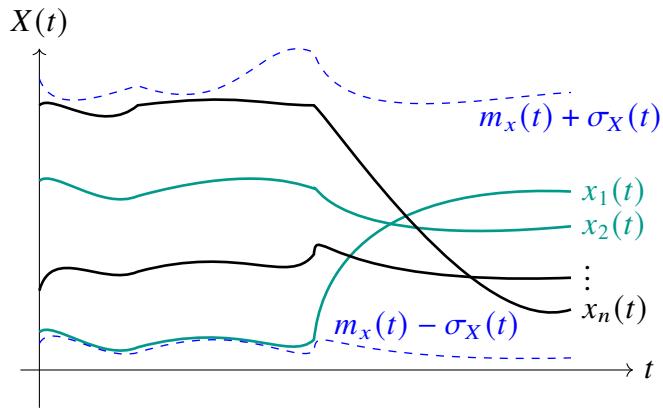


图 2-9 随机过程的数学期望和方差

均方差描绘了随机过程  $X(t)$  各个样本对其数学期望  $m_X(t)$  的偏差(或偏离)程度, 如图 2-9 所示。

**例 2.15** 由定义的积分可知,  $[Y(t), t \in T]$  都是  $t$  的确定函数。如果  $X(t)$  表示噪声电压, 则均方值  $\psi_X^2(t)$  和方差  $\sigma_X^2(t)$  就分别表示消耗在单位电阻上的瞬时功率的统计平均值和瞬时交流功率的统计平均值。

### (3) 离散型随机过程的一维数字特征

若离散型随机过程  $[Y(t), t \in T]$  的所有状态取值来源于离散的样本空间  $\Omega = \{y_1, y_2, \dots, y_n\}$ , 则其一维概率密度可用  $\delta$  函数表示为

$$f_Y(y, t) = \sum_{k=1}^m p_k(t) \delta(y - y_k), \quad k \in I = \{1, \dots, m\}, \quad (2.17)$$

其中  $p_k(t) = P(Y(t) = y_k)$ , 表示  $t$  时刻状态  $Y(t)$  取值为  $y$  的概率。在此基础上, 一维离散型随机过程  $Y(t)$  的期望、均方值和方差分别为

$$m_Y(t) = \int_{-\infty}^{\infty} y \sum_{k=1}^m p_k(t) \delta(y - y_k) dy = \sum_{k=1}^m y_k p_k(t). \quad (2.18)$$

$$\psi_Y^2(t) = E[Y^2(t)] = \sum_{i=1}^m y_i^2 p_k(t). \quad (2.19)$$

$$\sigma_Y^2(t) = D[Y(t)] = \sum_{k=1}^m [y_k - m_Y(t)]^2 p_k(t). \quad (2.20)$$

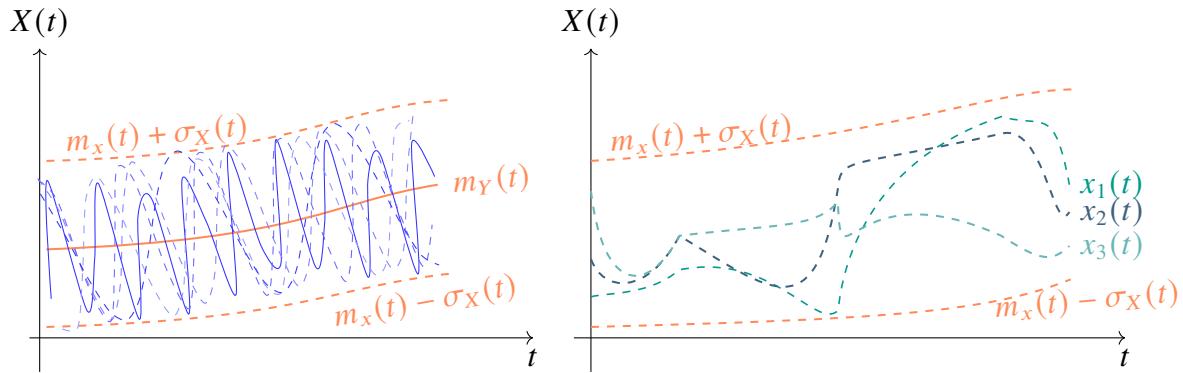


图 2-10 随机过程的自相关函数

如图 2-10 所示, 两个随机过程虽然有相同的均值与方差 (一维数字特征), 但它们有明显不同的内在结构: 一个随时间变化慢, 两个不同时刻状态之间的相互依赖性强 (相关性强); 另一个随时间变化快, 两个不同时刻状态之间的相互依赖性弱 (相关性弱)。可见随机过程的一维数字特征, 不能反映随机过程中两个不同时刻状态之间的相关程度。因此, 要用二维数字特征来描述随机过程任意两个时刻状态间的内在联系。

## 2. 二维数字特征

### (1) 自相关函数

定义随机过程最重要的二维数字特征——自相关函数。

#### 定义 2.16 二维自相关函数

自相关函数为

$$R_X(t_1, t_2) = E[X(t_1)X(t_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_X(x_1, x_2; t_1, t_2) dx_1 dx_2. \quad (2.21)$$

式 (2.21) 是随机过程  $X(t)$  两个不同时刻  $t_1, t_2$  的取值  $X(t_1), X(t_2)$  之间的二阶联合原点矩, 它反映了  $X(t)$  在任意两个时刻状态之间的自相关程度。



 **注 2.17.** 当  $t_1 = t_2 = t$  时,  $X(t)$  的自相关函数就是其均方值, 即

$$R_X(t, t) = E[X(t)X(t)] = E[X^2(t)] = \psi_X^2(t). \quad (2.22)$$

### (2) 自协方差函数

用任意两个不同时刻  $t_1, t_2$  上两个随机变量的协方差来描述自相关程度。

### 定义 2.18 自协方差函数

称

$$\begin{aligned} C_X(t_1, t_2) &= E \{ [X(t_1) - m_X(t_1)] [X(t_2) - m_X(t_2)] \} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x_1 - m_X(t_1)] [x_2 - m_X(t_2)] f_X(x_1, x_2; t_1, t_2) dx_1 dx_2. \end{aligned}$$

为自协方差函数(或中心化自相关函数)。



$C_X(t_1, t_2)$  与  $R_X(t_1, t_2)$  有下列关系:

$$\begin{aligned} C_X(t_1, t_2) &= E \{ [X(t_1) - m_X(t_1)] [X(t_2) - m_X(t_2)] \} \\ &= R_X(t_1, t_2) - m_X(t_1) m_X(t_2). \end{aligned} \quad (2.23)$$

当  $t_1 = t_2 = t$  时,  $X(t)$  的自协方差函数就是方差, 即

$$C_X(t, t) = E \{ [X(t) - m_X(t)]^2 \} = D[X(t)] = \psi_X^2(t) - m_X^2(t) = \sigma_X^2(t). \quad (2.24)$$

### (3) 自相关系数

对  $C_X(t_1, t_2)$  按  $\sigma_X(t)$  归一化, 就得到随机过程  $X(t)$  的自相关系数

$$\rho_X(t_1, t_2) = \frac{C_X(t_1, t_2)}{\sigma_X(t_1) \sigma_X(t_2)}, \quad (2.25)$$

式中  $\sigma_X(t_1) \neq 0, \sigma_X(t_2) \neq 0$ 。

注意:  $\rho_X(t_1, t_2), C_X(t_1, t_2)$  和  $R_X(t_1, t_2)$  只描述  $X(t_1)$  与  $X(t_2)$  间的线性相关的程度。

### (4) 离散型随机过程的二维数字特征

由于离散型随机过程  $Y(t)$  在  $t_1, t_2 \in T$  时的二维概率密度可用二维  $\delta$  函数表示, 即

$$f_Y(y_1, y_2, t_1, t_2) = \sum_{k_1, k_2 \in \varepsilon_Y} \sum P_{k_1 k_2}(t_1, t_2) \delta(y_1 - k_1) \delta(y_2 - k_2), \quad (2.26)$$

其中  $P_{k_1 k_2}(t_1, t_2) = P\{Y(t_1) = k_1, Y(t_2) = k_2\}$ , 代表过程  $Y(t)$  在  $t_1, t_2$  时刻两个状态  $Y(t_1)$  与  $Y(t_2)$  分别取离散值  $k_1, k_2$  的联合概率。离散型随机过程  $Y(t)$  的自相关函数为

$$R_Y(t_1, t_2) = \sum_{k_1, k_2 \in \varepsilon_Y} \sum k_1 k_2 P\{Y(t_1) = k_1, Y(t_2) = k_2\}, \quad (2.27)$$

其中  $\varepsilon_Y$  是随机过程  $Y(t)$  所有状态可能取值的范围。

注意: 随机过程  $X(t)$  的均值、方差和自相关函数等存在的条件是

$$\begin{aligned} 1^\circ \quad &E\{|X(t)|\} < \infty. \\ 2^\circ \quad &E\{|X(t)|^2\} < \infty. \end{aligned} \quad (2.28)$$

**例 2.19** 参数为  $\alpha$  的柯西过程  $X(t)$  的概率密度为

$$f(x, t) = \frac{\alpha t / \pi}{x^2 + (\alpha t)^2}, \quad (2.29)$$

由于

$$E\{|X(t)|\} = \frac{1}{\pi} \int_{-\infty}^{\infty} |x| \frac{\alpha t}{x^2 + (\alpha t)^2} dx = \infty. \quad (2.30)$$

$$E\{|X(t)|^2\} = \frac{1}{\pi} \int_{-\infty}^{\infty} x^2 \frac{\alpha t}{x^2 + (\alpha t)^2} dx = \frac{\alpha t}{\pi} \int_{-\infty}^{\infty} \left[ 1 - \frac{(\alpha t)^2}{x^2 + (\alpha t)^2} \right] dx. \quad (2.31)$$

是发散的，所以柯西过程的均值和方差等均不存在。

**例 2.20** 已知随机过程  $X(t) = V \cos 4t$ ,  $-\infty < t < \infty$ , 式中  $V$  是随机变量, 其数学期望为 5, 方差为 6。求随机过程  $X(t)$  的均值、方差和自相关函数和自协方差函数。

解：由题意可知  $E[V] = 5$ ,  $D[V] = 6$ 。从而得到  $V$  的均方值为

$$E[V^2] = D[V] + E^2[V] = 6 + 5^2 = 31.$$

根据随机过程数字特征的定义和性质, 可求得

$$m_X(t) = E[X(t)] = E[V \cos 4t] = \cos 4t \cdot E[V] = 5 \cos 4t.$$

$$\sigma_X^2(t) = D[X(t)] = D[V \cos 4t] = \cos^2 4t \cdot D[V] = 6 \cos^2 4t.$$

$$\begin{aligned} R_X(t_1, t_2) &= E[X(t_1) X(t_2)] = E[V \cos 4t_1 \cdot V \cos 4t_2] \\ &= \cos 4t_1 \cos 4t_2 \cdot E[V^2] = 31 \cos 4t_1 \cos 4t_2. \end{aligned}$$

$$\begin{aligned} C_X(t_1, t_2) &= E[(X(t_1) - m_X(t_1))(X(t_2) - m_X(t_2))] \\ &= R_X(t_1, t_2) - m_X(t_1)m_X(t_2) \\ &= 31 \cos 4t_1 \cos 4t_2 - 5 \cos 4t_1 \cdot 5 \cos 4t_2 = 6 \cos 4t_1 \cos 4t_2. \end{aligned}$$

**例 2.21** 已知随机过程  $X(t) = Ut$ ,  $-\infty < t < \infty$ , 其中随机变量  $U \sim [0, 1]$  的均匀分布。求随机过程  $X(t)$  的均值、方差、自相关函数和自协方差函数。

解：由题意可知, 随机变量  $U$  的概率密度为

$$f_U(u) = \begin{cases} 1, & 0 \leq u \leq 1 \\ 0, & \text{其他} \end{cases}. \quad (2.32)$$

根据随机过程数学期望的定义, 时间  $t$  与求统计平均无关, 因此可以看成常数, 则

$$\begin{aligned} m_X(t) &= E[X(t)] = E[Ut] = tE[U] = t \int_0^1 u \cdot 1 \cdot du = \frac{t}{2}. \\ R_X(t_1, t_2) &= E[X(t_1)X(t_2)] = E[Ut_1 \cdot Ut_2] = t_1 t_2 E[U^2]. \\ &= t_1 t_2 \int_{-\infty}^{\infty} u^2 f_U(u) du = t_1 t_2 \int_0^1 u^2 \cdot 1 \cdot du = \frac{t_1 t_2}{3}. \\ C_X(t_1, t_2) &= R_X(t_1, t_2) - m_X(t_1)m_X(t_2) = \frac{t_1 t_2}{3} - \frac{t_1}{2} \cdot \frac{t_2}{2} = \frac{t_1 t_2}{12}. \\ \sigma_X^2(t) &= C_X(t, t) = \frac{t^2}{12}. \end{aligned}$$

**例 2.22** 已知一个随机过程由四条样本函数组成, 如图 2-11 所示, 而且每条样本函数出现的概率相等。求自相关函数  $R_X(t_1, t_2)$ 。

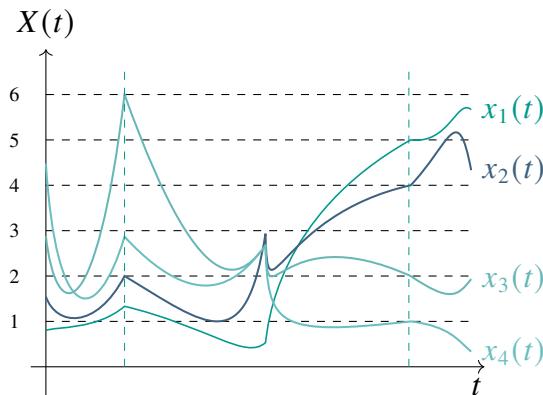


图 2-11 随机过程的样本

解: 由题意可知, 随机过程  $X(t)$  在  $t_1$  和  $t_2$  两个时刻为两个离散随机变量, 根据图 2-11 所示, 可列出联合分布律如下

	$\zeta_1$	$\zeta_2$	$\zeta_3$	$\zeta_4$
$X(t_1)$	1	2	6	3
$X(t_2)$	5	4	2	1
$p_{k_1 k_2}(t_1, t_2)$	0.25	0.25	0.25	0.25

故

$$\begin{aligned}
 R_X(t_1, t_2) &= \sum_{k_1, k_2 \in \epsilon_Y} \sum k_1 k_2 p_{k_1 k_2}(t_1, t_2) \\
 &= \sum_{k_1, k_2 \in \epsilon_Y} \sum k_1 k_2 p\{X(t_1) = k_1, X(t_2) = k_2\} \\
 &= 0.25 \times (1 \times 5 + 2 \times 4 + 6 \times 2 + 3 \times 1) = 7.
 \end{aligned}$$

### 2.1.5 随机过程的特征函数

由第一章可知, 随机变量的特征函数与其概率密度存在着唯一对应性, 从而可以用特征函数来简化随机变量数字特征的运算。随机过程的特征函数与其概率密度之间也存在着唯一对应性, 因此也可利用随机过程的特征函数来简化随机过程数字特征的运算。

#### 1. 一维特征函数

随机过程  $X(t)$  在任一固定时刻  $t_1$  的状态  $X(t_1)$  是一维随机变量,  $X(t_1)$  的特征函数为

$$Q_X(u_1; t_1) = E[e^{ju_1 X(t_1)}] = \int_{-\infty}^{\infty} e^{ju_1 x} f_X(x; t_1) dx, \quad (2.33)$$

式中  $X = x(t_1)$  为随机变量  $X(t_1)$  可能的取值,  $f_X(x; t)$  为过程  $X(t)$  的一维概率密度。若将  $t$  换成变量, 则随机过程  $X(t)$  的一维特征函数为

$$Q_X(u; t) = E[e^{ju X(t)}] = \int_{-\infty}^{\infty} e^{ju x} f_X(x; t) dx, \quad (2.34)$$

它是  $u, t$  的二元函数, 它与  $f_X(x; t)$  构成一对变换, 有

$$f_X(x; t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Q_X(u; t) e^{-jux} du. \quad (2.35)$$

若将特征函数定义式 (2.34) 的两边都对变量  $u$  求  $n$  阶偏导数, 得

$$\frac{\partial^n Q_X(u; t)}{\partial u^n} = j^n \int_{-\infty}^{\infty} x^n e^{ju x} f_X(x; t) dx, \quad (2.36)$$

则随机过程  $X(t)$  的  $n$  阶原点矩函数为

$$E[X^n(t)] = \int_{-\infty}^{\infty} x^n f_X(x; t) dx = (-j)^n \left. \frac{\partial^n Q_X(u, t)}{\partial u^n} \right|_{u=0}. \quad (2.37)$$

可见, 利用特征函数可以将“求积分”转变成“求导”, 简化了运算。

#### 2. 二维特征函数

$X(t)$  在任意两个时刻  $t_1, t_2$  的状态构成二维随机变量  $[X(t_1), X(t_2)]$ , 它们的联合特征函数

$$\begin{aligned} Q_X(u_1, u_2; t_1, t_2) &= E \{ \exp [ju_1 X(t_1) + ju_2 X(t_2)] \} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{j(u_1 x_1 + u_2 x_2)} f_X(x_1, x_2; t_1, t_2) dx_1 dx_2. \end{aligned} \quad (2.38)$$

称为随机过程  $X(t)$  的二维特征函数。它是  $u_1, u_2, t_1, t_2$  的四元函数, 表达式  $X_1 = x(t_1), X_2 = x(t_2)$  分别为随机变量  $X(t_1), X(t_2)$  可能的取值。

若  $f_X(x_1, x_2; t_1, t_2)$  是随机过程  $X(t)$  的二维概率密度, 则它与二维特征函数  $Q_X(u_1, u_2; t_1, t_2)$  构成变换对, 有

$$f_X(x_1, x_2; t_1, t_2) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Q_X(u_1, u_2; t_1, t_2) e^{-j(u_1 x_1 + u_2 x_2)} du_1 du_2. \quad (2.39)$$

若将二维特征函数定义式 2.38 的两边对变量  $u_1, u_2$  各求一次偏导数, 得

$$\frac{\partial^2 Q_X(u_1, u_2; t_1, t_2)}{\partial u_1 \partial u_2} = j^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 e^{j(u_1 x_1 + u_2 x_2)} f_X(x_1, x_2; t_1, t_2) dx_1 dx_2. \quad (2.40)$$

则随机过程  $X(t)$  的相关函数为

$$\begin{aligned} R_X(t_1, t_2) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_X(x_1, x_2; t_1, t_2) dx_1 dx_2 \\ &= - \left. \frac{\partial^2 Q_X(u_1, u_2; t_1, t_2)}{\partial u_1 \partial u_2} \right|_{u_1=u_2=0}. \end{aligned} \quad (2.41)$$

### 3. $n$ 维特征函数

同理, 随机过程  $X(t)$  的  $n$  维特征函数定义为

$$\begin{aligned} Q_X(u_1, \dots, u_n; t_1, \dots, t_n) &= E \{ \exp [ju_1 X(t_1) + \dots + ju_n X(t_n)] \} \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{j(u_1 x_1 + \dots + u_n x_n)} \\ &\quad \cdot f_X(x_1, \dots, x_n; t_1, \dots, t_n) dx_1 \dots dx_n. \end{aligned} \quad (2.42)$$

根据逆转公式, 由过程  $X(t)$  的  $n$  维特征函数, 求得的  $n$  维概率密度为

$$\begin{aligned} f_X(x_1, \dots, x_n; t_1, \dots, t_n) &= \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Q_X(u_1, \dots, u_n; t_1, \dots, t_n) \\ &\quad e^{-j(u_1 x_1 + \dots + u_n x_n)} du_1 \dots du_n. \end{aligned} \quad (2.43)$$

## 4 离散型随机过程的特征函数

设离散型随机过程  $X(t)$  在  $t$  时刻的状态(随机变量)为  $X(t_1)$ 。随机变量  $X(t_1)$  的取值为其分布律为  $p_i(t_1) = P\{X(t_1) = x_i\}$ 。则离散型随机过程的一维特征函数定义如下

$$Q_X(u; t) = \sum_i e^{jux_i} p_i(t), t \in T. \quad (2.44)$$

## 2.2 平稳与非平稳随机过程

在本节，我们首先简要讨论平稳与非平稳随机过程的概念，在此基础上研究随机过程的主要特性，并以高斯随机过程为例，简要讨论它的性质和应用。

如定义 2.1 所述，随机过程在数学上概括为一种必须用一簇与  $t$  有关的随机变量才能描述的随机现象，一般记为  $\{X(t)\}, t \in T$ 。所谓平稳随机过程，是指其统计特性不随时间推移而变化的随机过程。

① 当  $t = t_1$  时，随机过程  $\{X(t)\}$  退化为随机变量  $X_{t_1}$ ，其概率密度函数为  $f(x_{t_1})$ ，记  $f(x_1; t_1)$ 。

② 当  $t = t_1$  及  $t = t_2$  时，随机过程  $\{X(t)\}$  退化为随机变量  $x_{t_1}$  和  $x_{t_2}$ ，其概率密度函数为  $f(x_{t_1}, x_{t_2})$ ，记  $f(x_1, x_2; t_1, t_2)$ 。

③ 当  $t = t_1, t = t_2, \dots, t = t_n$  时，相应的  $n$  维随机变量的概率密度函数  $f(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)$ 。

显然，一般来说，对于一个具体的随机过程  $\{X(t)\}, t_{i+1} - t_i (i = 0, \pm 1, \pm 2, \dots)$  愈小， $n$  愈大，则  $f(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)$ ，对于  $\{X(t)\}$  的概率特性的描绘就愈详尽。

### 2.2.1 平稳随机过程的分类

#### 1. 严平稳随机过程

为便于形象地理解平稳概念，纵坐标表达随机过程  $X(t)$  的所有状态，如图 2-12 所示。

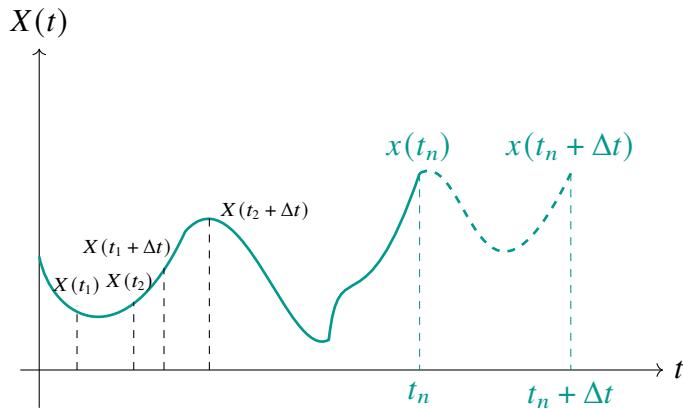


图 2-12 平稳随机过程

如果随机过程  $X(t)$  的任意  $n$  维分布不随时间起点的不同而变化, 即取样点在时间轴上平移了任意  $\Delta t$  后, 其  $n$  维概率密度保持不变, 即

$$f_X(x_1, \dots, x_n; t_1 + \Delta t, \dots, t_n + \Delta t) = f_X(x_1, \dots, x_n; t_1, \dots, t_n). \quad (2.45)$$

则称该过程为严平稳随机过程(或狭义平稳过程)。

### 定义 2.23 严平稳随机过程

随机过程  $\{X(t)\}$  所确定的  $n$  维概率密度函数为  $f(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)$ , 其中  $t_i$  ( $i = 1, 2, \dots, n$ ) 为任意取定的  $n$  个时刻, 如果对任意值  $n$ , 满足

$$f(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = f(x_1, x_2, \dots, x_n; t_1 + \tau, t_2 + \tau, \dots, t_n + \tau). \quad (2.46)$$

那么, 称  $\{X(t)\}$  为严平稳随机过程, 又可称为强平稳随机过程或狭义平稳随机过程。

严平稳随机过程的统计特性与所选取的时间起点无关。

无论从什么时间开始测量  $n$  个状态, 得到的统计特性是一样的, 即  $\{X(t), t \in T\}$  与  $\{X(t + \Delta t), t + \Delta t \in T\}$  具有相同的分布与统计特性。

**例 2.24** 测量优质电阻热噪声的统计特性, 由于它是平稳过程, 因而在任何时间进行测试, 都能得到相同的结果。

因此, 讨论平稳随机过程的实际意义在于平稳过程可使分析大为简化。

### (2) 概率密度及数字特征

严平稳随机过程的  $n$  维概率密度不随时间平移而变化的特性, 反映在它的一维和二维概率密度及数字特征上, 具有下列性质:

1) 若  $X(t)$  是严平稳随机过程, 则它的一维概率密度和数字特征与时间  $t$  无关。

将严平稳随机过程的定义式 (2.46) 用于一维概率密度情况, 令  $\Delta t = -t_1$ , 则有

$$f_X(x_1; t_1) = f_X(x_1; t_1 + \Delta t) = f_X(x_1; 0) = f_X(x_1). \quad (2.47)$$

由此可求得过程  $X(t)$  的均值、均方值和方差, 皆为与时间  $t$  无关的常数, 分别记作  $m_X, \psi_X^2$  和  $\sigma_X^2$ .

$$\begin{aligned} E[X(t)] &= \int_{-\infty}^{\infty} x_1 f_X(x_1) dx_1 = m_X. \\ E[X^2(t)] &= \int_{-\infty}^{\infty} x_1^2 f_X(x_1) dx_1 = \psi_X^2. \end{aligned} \quad (2.48)$$

### 定义 2.25 弱平稳随机过程

当  $E\{X^2(t)\} < +\infty$  时, 由  $\{X(t)\}$  在任意两个时刻  $t_1, t_2$  ( $t_2 > t_1$ ) 所确定的随机变量为  $X(t_1)$  和  $X(t_2)$ . 对于期望

$$E\{X(t)\} = E\{X(t + \tau)\}. \quad (2.49)$$

$$E\{X(t_1) X(t_2)\} = E\{X(t_1 + \tau) X(t_2 + \tau)\}. \quad (2.50)$$

如果将上述时刻移动任意值  $\tau$ , 满足 (2.49) 和 (2.50), 那么, 称  $\{X(t)\}$  为广义平稳随机过程, 又称为弱平稳随机过程。其中 (2.49) 式表示  $\{X(t)\}$  的一阶矩, 具有平移不变性, 式 (2.50) 表示  $\{X(t)\}$  的联合二阶矩, 具有时刻平移不变性。

显然, 当  $E\{X^2(t)\} < +\infty$  时, 严格平稳随机过程必然是广义平稳的, 但是广义平稳随机过程不一定是严格平稳的。

$$D[X(t)] = \int_{-\infty}^{\infty} (x_1 - m_X)^2 f_X(x_1) dx_1 = \sigma_X^2. \quad (2.51)$$

2) 严平稳随机过程  $X$  的二维样本密度和自相关与  $t_1, t_2$  的时间间隔  $t_2 - t_1$  有关, 而与时间起点无关。

将严平稳随机过程的定义式用于二维概率密度情况, 令平移时间  $\Delta t = -t_1$ , 设  $\tau = t_2 - t_1$ , 则

$$\begin{aligned} f_X(x_1, x_2; t_1, t_2) &= f_X(x_1, x_2, t_1 + \Delta t, t_2 + \Delta t) \\ &= f_X(x_1, x_2; 0, t_2 - t_1) = f_X(x_1). \end{aligned} \quad (2.52)$$

概率密度仅依赖于时间差  $t_2 - t_1$ , 而与时刻  $t_1, t_2$  无关。由此可得, 过程  $X(t)$  的自相关函数只是单变量  $\tau$  的函数

$$R_X(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_X(x_1, x_2; \tau) dx_1 dx_2 = R_X(\tau). \quad (2.53)$$

同理, 自协方差函数

$$C_X(t_1, t_2) = C_X(\tau) = R_X(\tau) - m_X^2. \quad (2.54)$$

## 2. 宽平稳随机过程

实际上, 要判定某个具体随机过程严平稳是很困难的。一般在工程应用中, 若产生随机过程的主要物理条件在时间运动过程中不变化, 那么此过程就可认为是平稳的。

**例 2.26** 电子管中的散弹噪声是由器件的颗粒效应引起的, 由于产生它的主要条件与时间无关, 所以此噪声可以认为是平稳随机过程(当然, 在刚接上电源, 接收机还处在“过渡”状态时的输出噪声是非平稳的)。

有些非平稳过程, 在某一时、同一范围内可作为平稳过程来处理。在很多实际问题的研究中, 往往并不需要随机过程在所有时间都平稳, 只要在观测的有限时间内过程平稳就行。因此, 在工程实际的应用中, 通常只在相关理论的范围内考察过程的平稳性问题。所谓相关理论是指仅限于研究与随机过程的一、二阶矩有关的理论, 它主要研究随机过程的数学期望、相关函数及功率谱密度等。随机过程的一、二阶矩函数不能像  $n$  维概率分布那样全面地描述随机过程的统计特性, 但它们在一定程度上相当有效地描述了随机过程的某些重要特性。以电子技术为例, 若平稳过程  $X$  代表某一噪声电压, 则由  $X(t)$  的阶矩函数可以求出该噪声电压的直流平均功率、交流平均功率、总平均功率和功率谱密度等参数。对于很多实际工程技术, 往往获得这些参数也足够。

而对于工程技术中最常用来模拟随机现象的高斯过程来说, 研究其一、二阶矩就能代替对其整个过程性质的研究。因此, 在实际应用中, 经常具体讨论在二阶矩意义上的、较广泛的一类平稳过程—宽平稳随机过程。

### 定义 2.27 广义平稳随机过程

如果随机过程  $X(t)$  的数学期望为常数, 其自相关函数只与时间  $\tau = t_2 - t_1$  有关, 且均方值有限, 即满足三个条件

$$\begin{cases} E[X(t)] = m_X. \\ R_X(t_1, t_2) = E[X(t_1)X(t_2)] = R_X(\tau) . \\ E[X^2(t)] < \infty. \end{cases} \quad (2.55)$$

则称  $X(t)$  为广义平稳随机过程(或宽平稳随机过程)。



由于宽平稳随机过程的定义只涉及与一维和二维概率密度有关的数字特征。

**结论** 一个严平稳随机过程只要均方值有界, 则它必定是宽平稳的, 反之则不一定成立。

 **注** 高斯过程例外, 因为高斯过程的概率密度是由均值和自相关函数完全确定的。如果高斯过程的均值和自相关函数不随时间平移而变化, 那么其概率密度也就不随时间的平移而变化, 即对于高斯过程, 宽平稳和严平稳等价。

注意: 本书以后提到的“平稳随机过程”除特别指明外, 通常是指宽平稳随机过程。

**例 2.28** 设随机过程  $X(t) = a \cos(\omega_0 t + \Phi)$ ,  $a, \omega_0$  皆为常数, 随机变量  $\Phi$  服从  $(0, 2\pi)$  上的均匀分布。判断  $X(t)$  是否为平稳随机过程, 给出理由。

解: 由题意可知, 随机变量  $\Phi$  的概率密度为

$$f_\Phi(\varphi) = \begin{cases} \frac{1}{2\pi}, & 0 < \varphi < 2\pi \\ 0, & \text{其他} \end{cases}. \quad (2.56)$$

根据定义式 (2.55), 求得过程  $X(t)$  的均值、自相关函数和均方值分别为

$$\begin{aligned} m_X(t) &= E[X(t)] = \int_{-\infty}^{\infty} x(t)f_\Phi(\varphi)d\varphi = \int_0^{2\pi} a \cos(\omega_0 t + \varphi) \cdot \frac{1}{2\pi} \cdot d\varphi = 0. \\ R_X(t_1, t_2) &= R_X(t, t + \tau) = E[X(t)X(t + \tau)] \\ &= E[a \cos(\omega_0 t + \Phi) \cdot a \cos(\omega_0(t + \tau) + \Phi)] \\ &= \frac{a^2}{2} E[\cos \omega_0 \tau + \cos(2\omega_0 t + \omega_0 \tau + 2\Phi)] \\ &= \frac{a^2}{2} \left[ \cos \omega_0 \tau + \int_0^{2\pi} \cos(2\omega_0 t + \omega_0 \tau + 2\varphi) \cdot \frac{1}{2\pi} d\varphi \right] \\ &= \frac{a^2}{2} \cos \omega_0 \tau = R_X(\tau). \\ E[X^2(t)] &= R_X(t, t) = R_X(0) = \frac{a^2}{2} < \infty. \end{aligned} \quad (2.57)$$

由上可知, 过程  $X(t)$  的均值为 0 (常数), 自相关函数仅与时间间隔  $\tau$  有关, 均方值为  $a^2/2$  (有限), 故过程  $X(t)$  是(宽) 平稳随机过程。

算法 2-1 符号积分计算.

```

1 clc
2 syms omega0 t tau varphi
3 int(cos(2*omega0*t+omega0*tau+2*varphi)/(2*pi),varphi,0,2*pi)

```

运行结果见图

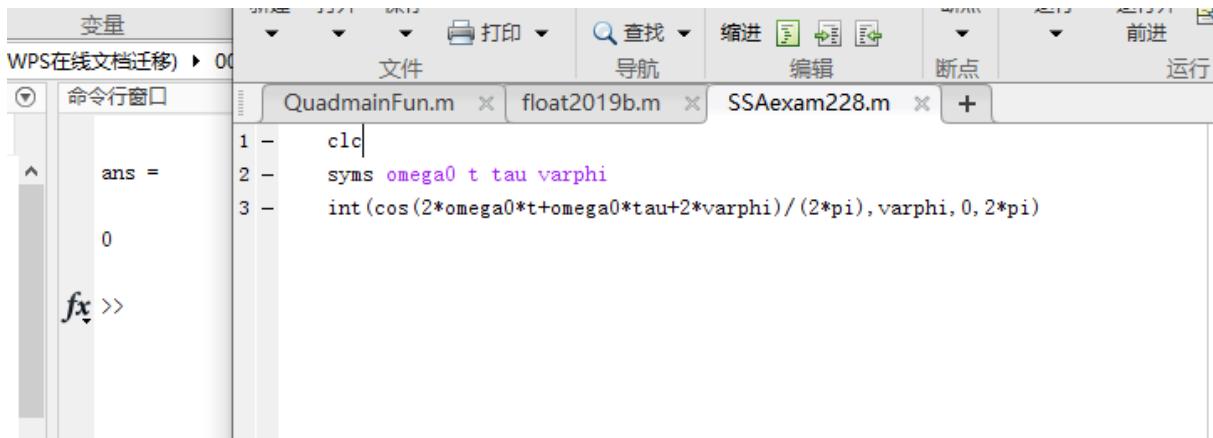


图 2-13 matlab 计算复积分  $\int_0^{2\pi} \cos(2\omega_0 t + \omega_0 t + 2\varphi) \cdot \frac{1}{2\pi} d\varphi$ .

**注 2.29.** 可以证明：仅当随机变量  $\Phi$  服从  $(0, 2\pi)$  或  $(-\pi, \pi)$  上的均匀分布时，过程  $X(t) = a \cos(\omega_0 t + \Phi)$  才是宽平稳随机过程。

**例 2.30** 已知两个随机过程  $X_1(t) = Y$  和  $X_2(t) = tY$ ，式中  $Y$  是随机变量。判断过程  $X_1(t)$  和  $X_2(t)$  的平稳性，给出理由。

解：① 对于过程  $X_1(t)$ ，由于

$$\begin{cases} m_{X_1}(t) = E[X_1(t)] = E[Y] = m_Y = \text{constant} \\ R_{X_1}(t_1, t_2) = E[X_1(t_1) X_1(t_2)] = E[Y \cdot Y] = E[Y^2] = \psi_Y^2 = \text{常数} \\ E[X_1^2(t)] = R_{X_1}(t, t) = \psi_Y^2 < \infty. \end{cases} \quad (2.58)$$

所以  $X_1(t)$  是平稳随机过程。

② 对于过程  $X_2(t)$ ，有

$$\begin{aligned} m_{X_2}(t) &= E[X_2(t)] = E[tY] = tE[Y] = tm_Y. \\ R_{X_2}(t_1, t_2) &= E[X_2(t_1) X_2(t_2)] = E[t_1 Y \cdot t_2 Y] = t_1 t_2 E[Y^2] = t_1 t_2 \psi_Y^2. \end{aligned} \quad (2.59)$$

可见，由于  $X_2(t)$  的均值与时间  $t$  有关，故  $X_2(t)$  不是平稳随机过程；或由于自相关函数与时间  $t_1, t_2$  有关，故  $X_2(t)$  不是平稳随机过程。

## 2.2.2 平稳随机过程自相关函数的性质

随机过程最基本的数字特征是数学期望和自相关函数。因为平稳过程的数学期望是个常数，经中心化后变为零，所以平稳过程主要的数字特征就是自相关函数。自相关函数不仅提供了随机过程各状态间的关联性信息，也是求随机过程功率谱密度必不可少的工具。

平稳过程的自相关函数具有如下性质:

1° 平稳过程的自相关函数在  $\tau = 0$  点的值为过程的均方值, 且非负, 即

$$R_X(0) = E[X^2(t)] = \psi_X^2 \geq 0. \quad (2.60)$$

2° 平稳过程自相关函数和自协方差函数是变量  $X$  的偶函数, 即

$$R_X(-\tau) = R_X(\tau), \quad C_X(-\tau) = C_X(\tau). \quad (2.61)$$

证:  $R_X(\tau) = E[X(t)X(t+\tau)] = E[X(t+\tau)X(t)] = R_X(-\tau)$ .

3° 平稳过程的自相关函数和自协方差函数在  $\tau = 0$  时具有最大值, 即

$$R_X(0) \geq |R_X(\tau)|, \quad C_X(0) = \sigma_X^2 \geq |C_X(\tau)|. \quad (2.62)$$

注意: 这里并不排除在  $\tau \neq 0$  时,  $R_X(\tau)$  和  $C_X(\tau)$  也有可能出现同样的最大值。如随机余弦信号的自相关函数  $R_X(\tau) = \frac{a^2}{2} \cos \omega_0 \tau$ , 在  $\tau = \frac{2n\pi}{\omega_0}, n = 0, \pm 1, \pm 2$  均为最大值  $\frac{a^2}{2}$ 。

4° 若  $X(t)$  为周期平稳过程, 满足  $X(t+T) = X(t)$ , 则其自相关函数必为周期函数, 且它的周期与过程的周期相同, 即

$$R_X(\tau+T) = R_X(\tau). \quad (2.63)$$

5° 若平稳过程  $X(t)$  含有一个周期分量, 则  $R_X(\tau)$  也可能含有一个周期分量, 即

$$X(t) = S(t) + N(t) = a \cos(\omega_0 t + \Phi) + N(t). \quad (2.64)$$

**例 2.31** 某接收机收到的混合信号  $X(t)$  是随机余弦信号  $S(t)$  和噪声  $N(t)$  之和, 式中  $\Phi$  服从  $(0, 2\pi)$  上均匀分布的随机变量,  $N(t)$  为平稳过程, 且  $S(t)$  与  $N(t)$  相互独立。易求  $X(t)$  的自相关函数为

$$R_X(\tau) = R_S(\tau) + R_N(\tau) = \frac{a^2}{2} \cos \omega_0 \tau + R_N(\tau). \quad (2.65)$$

可见,  $R_X(\tau)$  含有的周期分量  $R_S(\tau)$  与  $X(t)$  的周期分量  $S(t)$  的周期相同。

6° 若平稳过程不含有任何周期分量, 则

$$\lim_{|\tau| \rightarrow \infty} R_X(\tau) = R_X(\infty) = m_X^2. \quad (2.66)$$

证: 对此类非周期平稳过程, 当  $|\tau|$  增大时, 随机变量  $X(t)$  与  $X(t+\tau)$  之间的相关性减弱。在  $|x| \rightarrow \infty$  的极限情况下, 两者互不相关。故有

$$\begin{aligned} R_X(\infty) &= \lim_{|\tau| \rightarrow \infty} R_X(\tau) = \lim_{|\tau| \rightarrow \infty} E[X(t)X(t+\tau)] \\ &= \lim_{|\tau| \rightarrow \infty} E[X(t)] \cdot E[X(t+\tau)] = m_X^2. \end{aligned} \quad (2.67)$$

同理可得

$$\lim_{|\tau| \rightarrow \infty} C_X(\tau) = C_X(\infty) = 0. \quad (2.68)$$

7° 若平稳过程含有平均分量(均值)  $m_X$ , 则相关函数也将含有平均分量  $m_X^2$ , 即

$$R_X(\tau) = C_X(\tau) + m_X^2. \quad (2.69)$$

则由性质 6° 和 7° 可推出, 当平稳过程不含有任何周期分量时, 其在  $\tau = 0$  时的方差为

$$\sigma_X^2 = C_X(0) = R_X(0) - R_X(\infty). \quad (2.70)$$

8° 平稳过程的自相关函数不含有阶跃函数  $U(\tau)$  因子。

由平稳过程的频域分析可知(参见维纳-辛钦定理), 平稳过程的功率谱密度  $G_X(\omega) = F[R_X(\tau)]$  是  $\tau$  的实函数。若  $R_X(\omega)$  中含有  $U(\tau)$  因子, 则  $G_X(\omega)$  必含有虚数因子  $\pi\sigma(\omega) + 1/j\omega$ 。

根据以上性质的讨论, 可画出平稳过程自相关函数  $R_X(\tau)$  的典型曲线, 如图 2-14 所示。

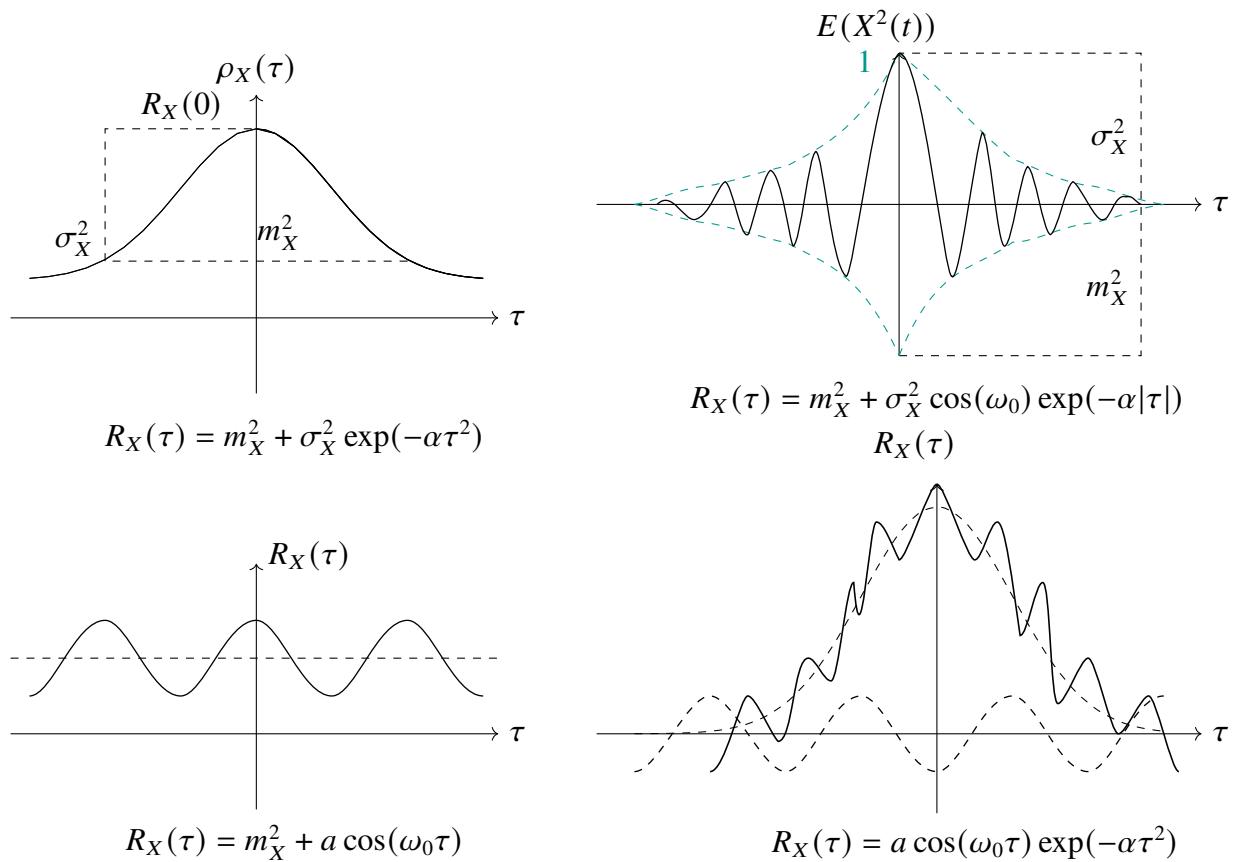


图 2-14 平稳过程的自相关函数

**例 2.32** 已知平稳过程  $X(t)$  的自相关函数为  $R_X(\tau) = \frac{4}{1+5\tau^2} + 36$ , 求  $X(t)$  的均值和方差。

解: 由性质 6° 和性质 7° 可得

$$\begin{aligned} m_X^2 &= R_X(\infty) = 36 \Rightarrow m_X = \pm\sqrt{R_X(\infty)} = \pm 6. \\ \sigma_X^2 &= R_X(0) - R_X(\infty) = 40 - 36 = 4. \end{aligned} \quad (2.71)$$

**例 2.33** 平稳过程  $X(t)$  的自相关函数为  $R_X(\tau) = 100e^{-10|\tau|} + 100 \cos 10\tau + 100$ , 求  $X(t)$  的均值、均方值和方差。

解: 将  $R_X(\tau)$  分解成周期与非周期两部分, 即

$$R_X(\tau) = (100 \cos 10\tau) + \left(100e^{-10|\tau|} + 100\right) = R_{X_1}(\tau) + R_{X_2}(\tau), \quad (2.72)$$

则  $X(t) = X_1(t) + X_2(t)$ , 且  $X_1(t)$  与  $X_2(t)$  相互独立。 $R_{X_1}(\tau) = 100 \cos 10\tau$  是周期分量  $X_1(t)$  的自相关函数, 利用例 2.28 的结论, 此分量的均值  $m_{X_1} = 0$ 。式 (2.72) 中  $m_{X_1} = 0$ 。式中  $R_{X_2}(\tau) = 100e^{-10|\tau|} + 100$  是非周期分量  $X_2(t)$  的自相关函数, 由性质 6°, 可得

$$m_{X_2}^2 = R_{X_2}(\infty) = 100 \Rightarrow m_{X_2} = \pm\sqrt{R_{X_2}(\infty)} = \pm 10. \quad (2.73)$$

则

$$\begin{aligned} m_X &= m_{X_1} + m_{X_2} = \pm 10. \\ E[X^2(t)] &= R_X(0) = 300. \\ \sigma_X^2 &= R_X(0) - m_X^2 = 200. \end{aligned} \quad (2.74)$$

即随机过程  $X(t)$  的均值为  $\pm 10$ , 均方值为 300, 方差为 200。

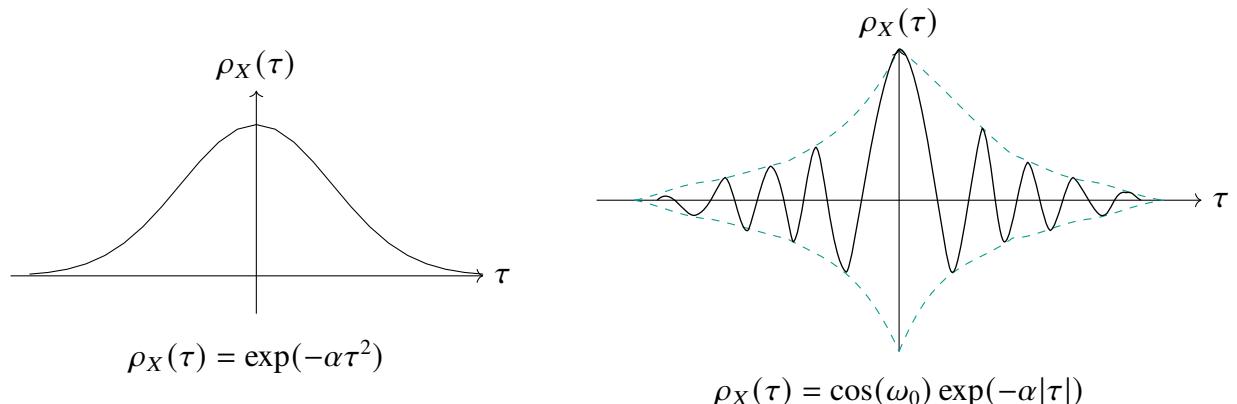


图 2-15 自相关函数  $\rho_X(\tau)$  的两条典型曲线

### 1. 自相关系数

### 定义 2.34 归一化自相关函数

为了表示平稳过程  $X(t)$  在两个不同时刻状态空间的线性相关程度，排除其他因素的影响，要对自相关函数进行归一化处理，从而得到过程  $X(t)$  的自相关系数

$$\rho_X(\tau) = \frac{C_X(\tau)}{\sigma_X^2} = \frac{R_X(\tau) - m_X^2}{\sigma_X^2} \quad (2.75)$$

也称为过程  $X(t)$  的归一化自相关函数。



图 2-15 给出了自相关系数  $\rho_X(\tau)$  的两条典型曲线。

由图可见， $\rho_X(\tau)$  可正、可负。正值表示正相关，即表示随机变量  $X(t)$  与  $X(t+\tau)$  同方向的相关；负值表示负相关，即表示随机变量  $X(t)$  与  $X(t+\tau)$  反方向的相关。 $\rho_X(\tau) = 0$  表示线性不相关， $|\rho_X(\tau)| = 1$  表示最强的线性相关。

## 2. 自相关时间

对于非周期随机过程  $X(t)$ ，随着  $\tau$  的增大， $X(t)$  与  $X(t+\tau)$  的相关程度将减弱。当  $\tau \rightarrow \infty$  时， $\rho_X(\tau) \rightarrow 0$ ，此时的  $X(t)$  与  $X(t+\tau)$  不再相关。实际上，当  $\tau$  大到一定程度时， $\rho_X(\tau)$  就已经很小了，此时， $X(t)$  与  $X(t+\tau)$  可认为已不相关（图 2-16(a))。因此，常常定义一个时间  $\tau_0$ ，当  $\tau > \tau_0$  时，就认为  $X(t)$  与  $X(t+\tau)$  不相关，把这个时间  $\tau_0$  称为相关时间。

### (1) 由相关时间给出自相关系数的一种定义

#### 定义 2.35 相关时间

由最大值  $\rho_X(0) = 1$  下降到  $\rho_X(\tau) = 0.05$  所经历的时间间隔为相关时间  $\tau'_0$ ，即

$$|\rho_X(\tau'_0)| = 0.05. \quad (2.76)$$

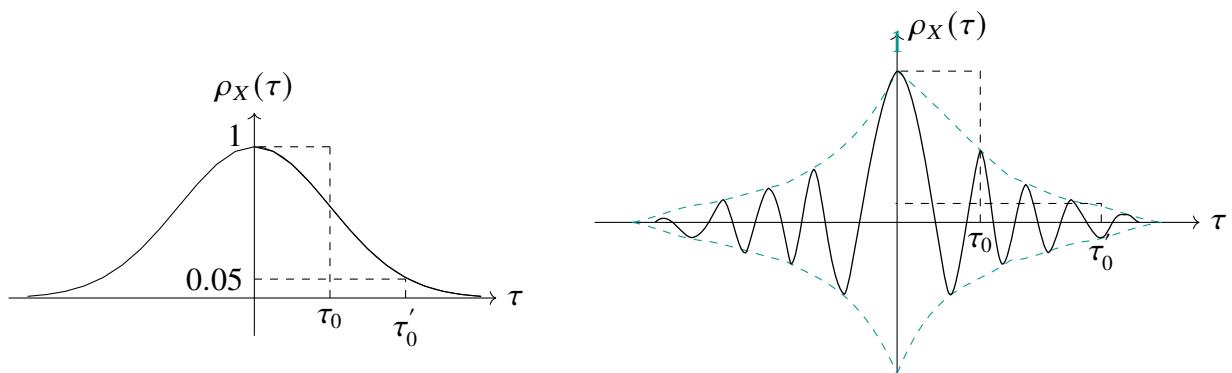


图 2-16 相关时间的定义

## (2) 相关时间的另一种定义

① 对不含高频分量的平稳过程, 用  $\rho_X(\tau)$  积分的一半来定义过程的相关时间, 即

$$\tau_0 = \int_0^\infty \rho_X(\tau) d\tau. \quad (2.77)$$

② 对于含高频分量的平稳过程, 如  $\rho_X(\tau) = a_X(\tau) \cos \omega_0 \tau$ , 则利用其包络  $a_X(\tau)$  积分的一半来定义其相关时间 (图 2-16(b)), 即

$$\tau_0 = \int_0^\infty a_X(\tau) d\tau. \quad (2.78)$$

综上所述, 自相关时间是随机过程的任意两个状态线性互不相关所需时间差的一种量度。信号包络的计算如图 2-17。

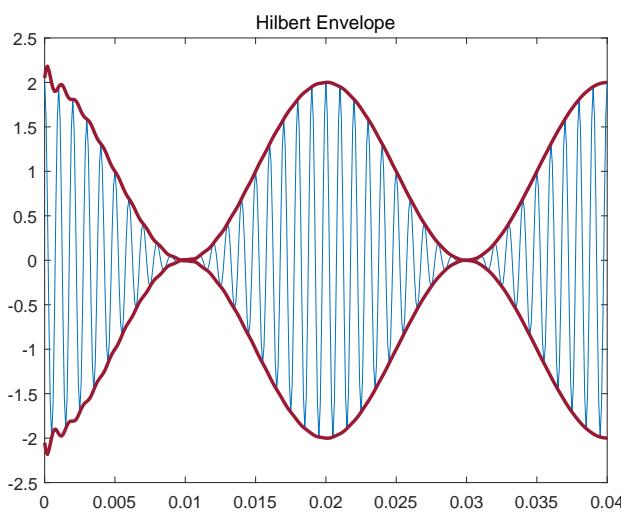


图 2-17 信号包络的计算

如图 2-18 所示,  $X(t)$  曲线越陡, 相关时间  $\tau'_{10}$  越小, 这就意味着随机过程  $X(t)$  的任意两个状态线性互不相关所需的时间差  $d_1$  越短, 过程随时间变化越剧烈, 其样本随时间  $t$  起伏越大; 反之,  $\rho_X(\tau)$  曲线越平缓, 则  $\tau'_{20}$  如越大, 这就意味着过程  $X(t)$  的任意两个状态线性互不相关所需的时间差越长, 随机过程随时间变化越缓慢, 其样本随  $t$  起伏越小。

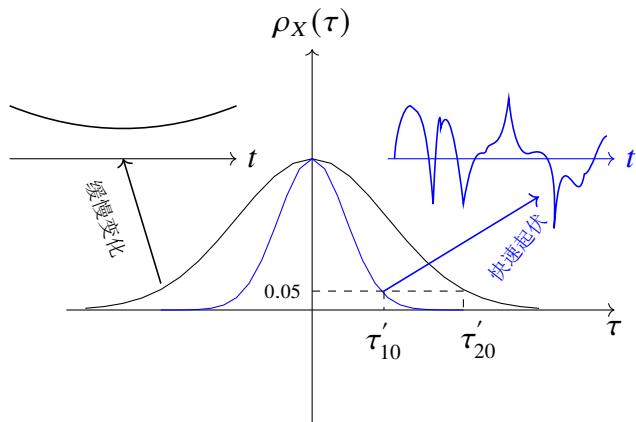


图 2-18 自相关时间

**例 2.36** 已知平稳过程  $X(t)$  的自相关函数  $R_X(\tau) = 3 \exp(-\tau^2)$ , 求自相关系数和自相关时间。

解: 由自相关系数的定义

$$\begin{aligned}\rho_X(\tau) &= \frac{C_X(\tau)}{C_X(0)} = \frac{R_X(\tau) - R_X(\infty)}{R_X(0) - R_X(\infty)} \\ &= \frac{3e^{-\tau^2} - 3e^{-\infty^2}}{3e^{-0^2} - 3e^{-\infty^2}} = \frac{3e^{-\tau^2}}{3} = e^{-\tau^2}.\end{aligned}\quad (2.79)$$

自相关时间定义一

$$|\rho_X(\tau'_0)| = 0.05 \Rightarrow \tau'_0 = \sqrt{-\ln(0.05)} = 1.731. \quad (2.80)$$

自相关时间定义二

$$\tau_0 = \int_0^\infty \rho_X(\tau) d\tau = \int_0^\infty e^{-\tau^2} d\tau = \frac{\sqrt{\pi}}{2} = 0.886. \quad (2.81)$$

### 2.2.3 非平稳随机过程

与平稳随机过程相对的是非平稳随机过程, 凡是不具有平稳随机过程概率特性的随机过程, 即为非平稳随机过程。应该说, 非平稳随机过程是自然界中存在的较之平稳随

机过程更具有广泛意义的一种随机过程，而平稳随机过程往往只是非平稳随机过程的一种近似处理或特殊情况。当非平稳随机过程的概率特性随时间变化比较缓慢时，在一段时间内可近似作为平稳随机过程来处理。

有一类过程，其本质是随机过程，但是其表示形式却类似确定过程，这类过程称为准随机过程。

**例 2.37** 如  $\{X(t) = \sum_{k=0}^{\tau} \xi_k t^k\}$ , 其中  $\xi_k (k = 0, 1, \dots, \tau)$  为随机变量, 又例如  $\{X(t) = A \sin(\omega_0 t + \theta)\}$ , 其中  $A, \theta$  为随机变量.

定义随机过程  $\{X(t)\}$  的  $n$  阶增量为

$$\Delta^n x(t) = x(t) - \binom{n}{1} x(t - \tau) + \binom{n}{2} x(t - 2\tau) + \dots + (-1)^n x(t - n\tau); \quad (2.82)$$

当  $n = 1$  时,  $\Delta x(t) = x(t) - x(t - \tau)$ .

平稳随机过程的增量亦是平稳随机过程, 但反之不一定成立。如:  $\{x(t) = \xi_0 + \xi_1 t\}$ , 其中  $\xi_0$  和  $\xi_1$  为随机变量。它的一阶增量  $\{\Delta x(t) = (\xi_0 + \xi_1 t) - [\xi_0 + \xi_1 (t - \tau)] = \xi_1 \tau\}$  是平稳随机过程, 但  $\{X(t)\}$  本身不是平稳的。

## 2.3 两个随机过程联合的统计特性

前面讨论了单个随机过程的统计特性。在实际应用中, 常需要研究两个或多个过程的统计特性。

如图 2-19, 它收到的通常是混入噪声的信号, 从噪声中检出有用信号, 除了考虑噪声和信号各自的统计特性外, 还要研究它们联合的统计特性。



图 2-19 接收机模块

### 1. 两个随机过程的联合分布

设有两个随机过程

$$\{X(t), t \in T\} \text{ 和 } \{Y(t), t \in T\} \quad (2.83)$$

的概率密度分别为

$$f_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n), f_Y(y_1, y_2, \dots, y_n; t'_1, t'_2, \dots, t'_m). \quad (2.84)$$

(1) 两个过程的  $n+m$  维联合分布函数

$$\begin{aligned} & F_{XY}(x_1, \dots, x_n, y_1, \dots, y_m; t_1, \dots, t_n, t'_1, \dots, t'_m) \\ &= P\{X(t_1) \leq x_1, \dots, X(t_n) \leq x_n, Y(t'_1) \leq y_1, \dots, Y(t'_m) \leq y_m\}. \end{aligned} \quad (2.85)$$

(2) 两个过程的  $n+m$  维联合概率密度

$$\begin{aligned} f_{XY} & (x_1, \dots, x_n, y_1, \dots, y_m; t_1, \dots, t_n, t'_1, \dots, t'_m) \\ &= \frac{\partial^{n+m} F_{XY}(x_1, \dots, x_n, y_1, \dots, y_m; t_1, \dots, t_n, t'_1, \dots, t'_m)}{\partial x_1 \cdots \partial x_n \partial y_1 \cdots \partial y_m}. \end{aligned} \quad (2.86)$$

(3) 随机过程  $X(t)$  和  $Y(t)$  的独立

若对任意的  $m, n$ ,  $X(t)$  和  $Y(t)$  都有

$$\begin{aligned} & F_{XY}(x_1, \dots, x_n, y_1, \dots, y_m; t_1, \dots, t_n, t'_1, \dots, t'_n) \\ &= F_X(x_1, \dots, x_n; t_1, \dots, t_n) F_Y(y_1, \dots, y_m; t'_1, \dots, t'_n). \end{aligned} \quad (2.87)$$

或

$$\begin{aligned} & f_{XY}(x_1, \dots, x_n, y_1, \dots, y_m, t_1, \dots, t_n, t'_1, \dots, t'_m) \\ &= f_X(x_1, \dots, x_n; t_1, \dots, t_n) f_Y(y_1, \dots, y_m, t'_1, \dots, t'_m). \end{aligned} \quad (2.88)$$

成立, 则称随机过程  $X(t)$  和  $Y(t)$  是独立的。

(4) 联合严平稳

#### 定义 2.38 联合严平稳

若两个过程的任意  $n+m$  维联合分布均不随时间平移而变化, 则称这两个过程为联合严平稳。

2. 两个随机过程的相关和正交

(1) 互相关函数

#### 定义 2.39 互相关函数

两个随机过程  $X(t)$  和  $Y(t)$  的互相关函数为

$$R_{XY}(t_1, t_2) = E[X(t_1)Y(t_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{XY}(x, y; t_1, t_2) dx dy, \quad (2.89)$$

式中  $X(t_1), Y(t_2)$  是过程  $X(t), Y(t)$  在  $t_1, t_2$  时刻的状态。

(2) 互协方差函数

**定义 2.40 互协方差函数**

两个过程  $X(t)$  和  $Y(t)$  的互协方差函数为

$$\begin{aligned} C_{XY}(t_1, t_2) &= E\{[X(t_1) - m_X(t_1)][Y(t_2) - m_Y(t_2)]\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x - m_X(t_1)][y - m_Y(t_2)] f_{XY}(x, y; t_1, t_2) dx dy, \end{aligned} \quad (2.90)$$

式中  $m_X(t_1)$  和  $m_Y(t_2)$  分别是随机变量  $X(t_1)$  和  $Y(t_2)$  的数学期望。



上式也可写成

$$C_{XY}(t_1, t_2) = R_{XY}(t_1, t_2) - m_X(t_1)m_Y(t_2). \quad (2.91)$$

**(3) 两个过程正交****定义 2.41 两个过程正交**

若两个过程  $X(t)$  和  $Y(t)$  对任意两个时刻  $t_1, t_2$ , 都有

$$R_{XY}(t_1, t_2) = 0 \text{ 或 } C_{XY}(t_1, t_2) = -m_X(t_1)m_Y(t_2), \quad (2.92)$$

则称  $X(t)$  和  $Y(t)$  两个过程正交。



 **注 2.42.** 随机变量  $X$  和  $Y$  可以被认为是无限维向量空间中的向量, 空间中的向量赋以为内积运算, 也即  $\langle X, Y \rangle = E[XY]$ . 内积可以诱导出范数; 随机变量  $X$  的范数是  $\|X\| \triangleq \sqrt{E[X^2]}$ . 内积运算也满足对称性、线性和非负性。当  $\langle X, Y \rangle = E[XY] = 0$ , 称随机变量  $X$  和  $Y$  正交. 任意两个时刻  $t_1, t_2$  随机变量  $X$  和  $Y$  正交, 则称  $X(t)$  和  $Y(t)$  两个过程正交。

**定义 2.43 同一时刻的状态正交**

若仅在同一时刻  $t$  存在

$$R_{XY}(t, t) = 0, \quad (2.93)$$

则称  $X(t)$  和  $Y(t)$  两个过程在同一时刻的状态正交。

**(4) 两个过程互不相关**

**定义 2.44 两个过程互不相关**

若两个过程  $X(t)$  和  $Y(t)$  对任意两个时刻  $t_1, t_2$ , 都有

$$C_{XY}(t_1, t_2) = 0 \text{ 或 } R_{XY}(t_1, t_2) = m_X(t_1)m_Y(t_2), \quad (2.94)$$

则称  $X(t)$  和  $Y(t)$  两个过程互不相关。

**定义 2.45 同一时刻的状态互不相关**

若仅在同一时刻  $t$  存在

$$C_{XY}(t, t) = 0, \quad (2.95)$$

则称  $X(t)$  和  $Y(t)$  两个过程在同一时刻的状态互不相关。



**例 2.46** 对于平稳随机过程, 规定标准正交随机变量的随机函数形式 [Rodríguez-2003] 为

$$\xi_m = \text{cas}(m\Theta), \quad (m = 1, 2, \dots, M) \quad (2.96)$$

式中:  $\Theta$  为基本随机变量, 是区间  $[-\pi, \pi]$  的均匀分布; 函数  $\text{cas}(mx) = \sqrt{2} \sin(x + \frac{\pi}{4})$  为 Hartley 正交基函数 (标准正交随机变量, 即  $E[\xi_m] = 0$ ,  $E[\xi_m \xi_k] = \delta_{mk}$ )。

容易验证 [刘章军-2014] 式 (2.96) 构造的随机函数形式满足标准正交随机变量的条件, 其标准正交随机变量  $\xi_m (m = 1, 2, \dots, M)$  一般为非高斯分布。利用式(2.96) 中标准正交随机变量的随机函数形式, 即可构造一组高斯标准正交(独立)随机变量 [刘章军-2014], 即

$$\xi_m = \Phi^{-1} \left[ \frac{1}{2} + \frac{1}{\pi} \arcsin \left( \frac{\text{cas}(m\Theta)}{\sqrt{2}} \right) \right], \quad (m = 1, 2, \dots, M) \quad (2.97)$$

式中:  $\Phi^{-1}$  为标准高斯随机变量分布函数  $\Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{t^2}{2}} dt$  的反函数。

### 3. 两个随机过程的联合平稳

#### (1) 两个随机过程的联合宽平稳定义

**定义 2.47 过程的联合宽平稳**

若两个随机过程  $X(t)$  和  $Y(t)$  各自宽平稳, 且它们的互相关函数仅是单变量  $\tau$  的函数, 即

$$R_{XY}(t_1, t_2) = E[X(t_1)Y(t_2)] = R_{XY}(\tau), \quad \tau = t_2 - t_1, \quad (2.98)$$

则称过程  $X(t)$  和  $Y(t)$  联合宽平稳(或联合平稳)。



### 2) 性质

两个联合平稳过程的互相关函数和互协方差函数具有如下性质:

1° 互相关函数和互协方差函数不存在偶对称, 它们满足(注意下标)

$$\begin{cases} R_{XY}(\tau) = R_{YX}(-\tau), \\ C_{XY}(\tau) = C_{YX}(-\tau). \end{cases} \quad (2.99)$$

2° 互相关函数和互协方差函数的取值满足

$$\begin{cases} |R_{XY}(\tau)|^2 \leq R_X(0)R_Y(0), \\ |C_{XY}(\tau)|^2 \leq C_X(0)C_Y(0) = \sigma_X^2\sigma_Y^2. \end{cases} \quad (2.100)$$

$$\begin{cases} |R_{XY}(\tau)| \leq \frac{1}{2}[R_X(0) + R_Y(0)] \\ |C_{XY}(\tau)| \leq \frac{1}{2}[C_X(0) + C_Y(0)] = \frac{1}{2}[\sigma_X^2 + \sigma_Y^2] \end{cases}. \quad (2.101)$$

(3) 两个联合平稳过程的互相关系数

#### 定义 2.48 互相关系数

两个联合平稳过程  $X(t)$  和  $Y(t)$  的互相关系数定义为

$$\rho_{XY}(\tau) = \frac{C_{XY}(\tau)}{\sqrt{C_X(0)C_Y(0)}} = \frac{R_{XY}(\tau) - m_Xm_Y}{\sigma_X\sigma_Y}. \quad (2.102)$$

由性质 2° 易得  $|\rho_{XY}(\tau)| \leq 1$ , 且  $\rho_{XY}(\tau) = 0$  时, 两个平稳过程  $X(t)$  和  $Y(t)$  线性不相关。

**例 2.49** 已知随机过程  $X(t)$  和  $Y(t)$  是平稳随机过程, 当

$$\textcircled{1} \begin{cases} X(t) = U \sin t + V \cos t \\ Y(t) = W \sin t + V \cos t \end{cases}, \textcircled{2} \begin{cases} X(t) = A \cos t + B \sin t \\ Y(t) = A \cos 2t + B \sin 2t \end{cases} \quad (2.103)$$

式中  $U, V, W$  是均值为 0, 方差为 6, 且互不相关的随机变量;  $A, B$  是均值为 0, 方差为 3, 且互不相关的随机变量。判断  $X(t)$  和  $Y(t)$  是否联合平稳, 给出理由。

解:

① 过程  $X(t)$  和  $Y(t)$  的互相关函数为

$$\begin{aligned}
 R_{XY}(t, t + \tau) &= E[X(t)Y(t + \tau)] \\
 &= E\{(U \sin t + V \cos t)[W \sin(t + \tau) + V \cos(t + \tau)]\} \\
 &= E[UW \sin t \sin(t + \tau) + UV \sin t \cos(t + \tau) \\
 &\quad + VW \cos t \sin(t + \tau) + V^2 \cos t \cos(t + \tau)] \\
 &= 0 + 0 + 0 + E[V^2] \cos t \cos(t + \tau) \\
 &= 6 \cdot \frac{1}{2} [\cos(2t + \tau) + \cos t] \\
 &= 3 \cos(2t + \tau) + 3 \cos t.
 \end{aligned} \tag{2.104}$$

可见, 此互相关函数是变量  $t, \tau$  的二元函数, 故  $X(t)$  和  $Y(t)$  不是联合平稳的。

② 过程  $X(t)$  和  $Y(t)$  的互相关函数为

$$\begin{aligned}
 R_{XY}(t, t + \tau) &= E[X(t)Y(t + \tau)] \\
 &= E\{(A \cos t + B \sin t)[A \cos 2(t + \tau) + B \sin 2(t + \tau)]\} \\
 &= E[A^2 \cos t \cos 2(t + \tau) + AB \cos t \sin 2(t + \tau) \\
 &\quad + AB \sin t \cos 2(t + \tau) + B^2 \sin t \sin 2(t + \tau)] \\
 &= E[A^2] [\cos t \cos 2(t + \tau) + \sin t \sin 2(t + \tau)] + 0 + 0 \\
 &= 3 \cos(t + 2\tau).
 \end{aligned} \tag{2.105}$$

可见, 此互相关函数也是变量  $t, \tau$  的二元函数, 故过程  $X(t)$  和  $Y(t)$  也不是联合平稳的。

**例 2.50** 对于实联合平稳的随机过程  $X(t)$  和  $Y(t)$ , 计算

$$E \left[ \left( \frac{X(t)}{\sqrt{R_X(0)}} \pm \frac{Y(t + \tau)}{\sqrt{R_Y(0)}} \right)^2 \right]. \tag{2.106}$$

证明: 式中  $R_X(0)$  和  $R_Y(0)$  分别为随机过程  $X(t)$  和  $Y(t)$  的均方值, 即

$$R_X(0) = E[X^2(t)], \quad R_Y(0) = E[Y^2(t + \tau)]. \tag{2.107}$$

只讨论实随机过程的情形, 就是说  $X(t)$  和  $Y(t)$  皆为时间  $t$  的实函数。而实函数的平方是非负的, 则

$$E \left[ \left( \frac{X(t)}{\sqrt{E[X^2(t)]}} \pm \frac{Y(t + \tau)}{\sqrt{E[Y^2(t + \tau)]}} \right)^2 \right] \geq 0. \tag{2.108}$$

将不等式左边展开, 得

$$\begin{aligned}
 & \frac{E[X^2(t)]}{E[X^2(t)]} \pm 2 \frac{E[X(t)Y(t+\tau)]}{\sqrt{E[X^2(t)]E[Y^2(t+\tau)]}} + \frac{E[Y^2(t+\tau)]}{E[Y^2(t+\tau)]} \geq 0 \\
 & \Rightarrow 2 \pm 2 \frac{R_{XY}(\tau)}{\sqrt{E[X^2(t)]E[Y^2(t+\tau)]}} \geq 0 \\
 & \Rightarrow |R_{XY}(\tau)| \leq \sqrt{E[X^2(t)]E[Y^2(t+\tau)]} \\
 & \Rightarrow |R_{XY}(\tau)|^2 \leq R_X(0)R_Y(0).
 \end{aligned} \tag{2.109}$$

## 2.4 复随机过程

前面讨论的过程都是实随机过程, 即其样本函数是时间的实函数。这种表示方法的优点是直观, 易于接受。但在某些情况下, 如在高频窄带随机信号的处理中, 将信号表示成复函数形式更为方便。将这种用复函数表示的随机过程称为复随机过程。类似于实随机过程, 复随机过程是随时间变化的复随机变量。本节首先引入复随机变量的概念, 然后再介绍复随机过程的概念和有关特性。

### 1. 复随机变量

定义复随机变量  $Z$  为

$$Z = X + jY, \tag{2.110}$$

式中,  $X$  和  $Y$  都为实随机变量。实质上, 复随机变量  $Z$  是实随机变量  $X$  和  $Y$  所组成的二维随机变量, 故  $Z$  的统计特性可以用  $X$  和  $Y$  的联合分布来完整地描述。

将实随机变量的数学期望、方差和相关等概念推广到复随机变量中去时, 必须遵循的原则是: 当  $Y = 0$  时, 复随机变量  $Z$  等于实随机变量  $X$ 。

#### (1) 复随机变量 $Z$ 的数学期望

$$m_Z = E[Z] = E[X] + jE[Y] = m_X + jm_Y. \tag{2.111}$$

#### (2) 复随机变量 $Z$ 的方差

$$D_Z = D[Z] = E[|Z - m_Z|^2] = E[|\tilde{Z}|^2], \tag{2.112}$$

式中  $\tilde{Z} = Z - m_Z$ , 且有

$$D_Z = D_X + D_Y. \tag{2.113}$$

(3) 两个复随机变量  $Z_1, Z_2$  的协方差

$$C_{Z_1 Z_2} = E \left[ (Z_1 - m_{Z_1})^* (Z_2 - m_{Z_2}) \right] = E \left[ \tilde{Z}_1 \tilde{Z}_2 \right], \quad (2.114)$$

式中 “\*” 表示共轭,  $Z_1 = X_1 + jY_1, Z_2 = X_2 + jY_2$ 。当  $Z_1 = Z_2 = Z$  时,  $C_{Z_1 Z_2} = E \left[ |\tilde{Z}|^2 \right] = D_Z$ 。  
还可以把协方差写成下面的形式:

$$C_{Z_1 Z_2} = C_{X_1 X_2} + C_{Y_1 Y_2} + j (C_{X_1 Y_2} - C_{Y_1 X_2}), \quad (2.115)$$

式中  $C_{X_1 X_2}, C_{Y_1 Y_2}, C_{X_1 Y_2}, C_{Y_1 X_2}$  分别是  $(X_1, X_2), (Y_1, Y_2), (X_1, Y_2), (Y_1, X_2)$  的协方差。

(4) 两个复随机变量  $Z_1, Z_2$  的独立

**定义 2.51 相互独立的复随机变量**

若两个复随机变量  $Z_1 = X_1 + jY_1, Z_2 = X_2 + jY_2$  满足

$$f_{Z_1 Z_2}(x_1, y_1, x_2, y_2) = f_{X_1 Y_1, X_2 Y_2}(x_1, y_1, x_2, y_2) = f_{X_1 Y_1}(x_1, y_1) f_{X_2 Y_2}(x_2, y_2), \quad (2.116)$$

则称  $Z_1, Z_2$  相互独立。



(5) 两个复随机变量  $Z_1, Z_2$  的互不相关

若两个复随机变量  $Z_1, Z_2$  满足

$$C_{Z_1 Z_2} = E \left[ (Z_1 - m_{Z_1})^* (Z_2 - m_{Z_2}) \right] = 0, \quad (2.117)$$

则称  $Z_1, Z_2$  互不相关。

(6) 两个复随机变量  $Z_1, Z_2$  的正交

若两个复随机变量  $Z_1, Z_2$  满足

$$R_{z_1 z_2} = E \left[ Z_1^* Z_2 \right] = 0. \quad (2.118)$$

## 2. 复随机过程

如同复随机变量一样, 在信号的统计分析中, 引入复随机过程也是必要的。复随机过程表示为  $\{Z(t) = X(t) + jY(t)\}$ , 其中  $\{X(t)\}$  为复随机过程的实部, 而  $\{Y(t)\}$  则为随机过程的虚部,  $\{X(t)\}, \{Y(t)\}$  都是实随机过程。因此, 在  $t_i, i = 1, 2, \dots, n$  时刻, 由  $\{Z(t)\}$  所确定的  $2n$  维随机变量, 其相应的概率密度函数可用如下的连和概率密度

$$f(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n; y_1, y_2, \dots, y_m; t'_1, t'_2, \dots, t'_m) \quad (2.119)$$

表示。

### 定义 2.52 复随机过程

复随机过程为

$$Z(t) = X(t) + jY(t), \quad (2.120)$$

式中  $X(t)$  和  $Y(t)$  都是实随机过程。



复随机过程  $Z(t)$  的统计特性可由  $X(t)$  和  $Y(t)$  的  $2n$  维联合分布完整描述，其概率密度为

$$f_{XY}(x_1, \dots, x_n, y_1, \dots, y_n; t_1, \dots, t_n, t'_1, \dots, t'_n). \quad (2.121)$$

(1) 复随机过程  $Z(t)$  的数学期望

$$m_Z(t) = E[Z(t)] = E[X(t) + jY(t)] = m_X(t) + jm_Y(t). \quad (2.122)$$

(2) 复随机过程  $Z(t)$  的方差

$$D_Z(t) = E[|Z(t) - m_Z(t)|^2] = E[|\tilde{Z}(t)|^2], \quad (2.123)$$

式中  $\tilde{Z}(t) = Z(t) - m_Z(t)$ , 且有

$$D_Z(t) = D_X(t) + D_Y(t). \quad (2.124)$$

(3) 复随机过程  $Z(t)$  的自相关函数

$$R_Z(t, t + \tau) = E[Z^*(t)Z(t + \tau)]. \quad (2.125)$$

(4) 复随机过程  $Z(t)$  的自协方差函数

$$\begin{aligned} C_Z(t, t + \tau) &= E\{[Z(t) - m_Z(t)] \cdot [Z(t + \tau) - m_Z(t + \tau)]\} \\ &= E[\tilde{Z}^*(t)\tilde{Z}(t + \tau)]. \end{aligned} \quad (2.126)$$

$$C_Z(t, t) = D_Z(t). \quad (2.127)$$

(5) 复随机过程  $Z(t)$  宽平稳

若复随机过程  $Z(t)$  满足下面两个条件：

$$\begin{aligned} m_Z(t) &= m_Z, \\ R_Z(t, t + \tau) &= R_Z(\tau), \end{aligned} \quad (2.128)$$

则称此随机过程  $Z(t)$  宽平稳。

(6) 两个随机过程  $\{X(t)\}, \{Y(t)\}$ 。在不少的实际问题中，可能会涉及同时存在一个以上的随机过程，其中最简单的情况是同时存在两个随机过程  $\{X(t)\}, \{Y(t)\}$ 。可以用以下的  $n+m$  维随机变量的联合概率密度去描述它们

$$\begin{aligned} & f_{XY}(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n; y_1, y_2, \dots, y_m; t'_1, t'_2, \dots, t'_m) \\ &= f_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) \times f_Y(y_1, y_2, \dots, y_m; t'_1, t'_2, \dots, t'_m). \end{aligned} \quad (2.129)$$

显然， $n, m$  愈大，这种描述将愈精细。如果对于  $n$  和  $m$  取自然数中的任意值，有

$$\begin{aligned} & f(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n; y_1, y_2, \dots, y_m; t'_1, t'_2, \dots, t'_m) \\ &= f(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) \times f(y_1, y_2, \dots, y_m; t'_1, t'_2, \dots, t'_m). \end{aligned} \quad (2.130)$$

那么  $\{X(t)\}$  和  $\{Y(t)\}$  彼此统计独立。如果两个过程的联合概率密度函数与时间起点无关，这样的过程称为联合平稳的随机过程。

(7) 两个复随机过程  $Z_1(t)$  和  $Z_2(t)$  的互相关函数和互协方差函数

$$\begin{aligned} R_{Z_1 Z_2}(t, t + \tau) &= E[Z_1^*(t) Z_2(t + \tau)], \\ C_{Z_1 Z_2}(t, t + \tau) &= E\{[Z_1(t) - m_{Z_1}(t)] \cdot [Z_2(t + \tau) - m_{Z_2}(t + \tau)]\}. \end{aligned} \quad (2.131)$$

(8) 两个复随机过程  $Z_1(t)$  和  $Z_2(t)$  的联合平稳

两个复随机过程  $Z_1(t)$  和  $Z_2(t)$  各自平稳，且它们的互相关函数满足

$$R_{Z_1 Z_2}(t, t + \tau) = R_{Z_1 Z_2}(\tau), \quad (2.132)$$

则称这两个复随机过程联合平稳。

(9) 两个复随机过程  $Z_1(t)$  和  $Z_2(t)$  的互不相关

若两个复随机过程  $Z_1(t)$  和  $Z_2(t)$  满足

$$C_{Z_1 Z_2}(t, t + \tau) = 0, \quad (2.133)$$

则称这两个复随机过程互不相关。

(10) 两个复随机过程  $Z_1(t)$  和  $Z_2(t)$  的正交

### 定义 2.53 复随机过程正交

若两个复随机过程  $Z_1(t)$  和  $Z_2(t)$  满足

$$R_{Z_2 Z_2}(t, t + \tau) = 0, \quad (2.134)$$

则称这两个复随机过程正交。



**例 2.54** 已知复随机过程  $V(t)$  由  $N$  个复信号之和组成, 即

$$V(t) = \sum_{n=1}^N A_n \exp [j(\omega_0 t + \Phi_n)], \quad n = 1, 2, \dots, N \quad (2.135)$$

其中  $\omega_0$  为常数, 表示每个复信号的角频率;  $A_n$  是随机变量, 表示第  $n$  个复信号的幅度;  $\Phi_n$  是服从  $(0, 2\pi)$  上均匀分布的随机变量, 表示第  $n$  个复信号的相位。若相位  $\Phi_n$  ( $n = 1, 2, \dots, N$ ) 之间相互独立, 且幅度  $A_n$  和  $\Phi_n$  之间也是相互独立的, 求复过程  $V(t)$  的自相关函数。

解: 由复过程的自相关系数定义, 有

$$\begin{aligned} R_V(t, t + \tau) &= E[V^*(t)V(t + \tau)] \\ &= E\left\{\sum_{n=1}^N A_n \exp[-j(\omega_0 t + \Phi_n)] \sum_{m=1}^N A_m \exp[j(\omega_0(t + \tau) + \Phi_m)]\right\} \\ &= \sum_{n=1}^N \sum_{m=1}^N \{E[A_n A_m \exp[j(\Phi_m - \Phi_n)]] \exp(j\omega_0 \tau)\} \\ &= R_V(\tau). \end{aligned} \quad (2.136)$$

由已知的独立条件可得

$$\begin{aligned} R_V(\tau) &= \exp(j\omega_0 \tau) \sum_{n=1}^N \sum_{m=1}^N E[A_n A_m \exp[j(\Phi_m - \Phi_n)]] \\ &= \exp(j\omega_0 \tau) \sum_{n=1}^N \sum_{m=1}^N \{E(A_n A_m) E[\exp[j(\Phi_m - \Phi_n)]]\}. \end{aligned} \quad (2.137)$$

且

$$E\{\exp[j(\Phi_m - \Phi_n)]\} = \begin{cases} E[e^{j0}] = 1, & m = n \\ E[e^{j\Phi_m}] E[e^{-j\Phi_n}] = 0, & m \neq n \end{cases}. \quad (2.138)$$

所以

$$R_V(\tau) = \exp(j\omega_0 \tau) \sum_{n=1}^N E[A_n^2]. \quad (2.139)$$

## 2.5 随机过程的积分微分特性

在本书中, 我们还可能会用到随机过程积分和微分的特性, 下面给出它们的定义和一些基本性质。随机过程  $\{X(t)\}$  的可微性定义如下: 随机过程  $\{X(t)\}$  在时刻  $t$  处满足下述关系:

随机过程  $\{X(t)\}$  的可微性定义如下：

### 定义 2.55 事件独立

随机过程  $\{X(t)\}$  在时刻  $t$  处满足下述关系：

$$\lim_{T \rightarrow 0} E \left\{ \left[ \frac{x(t+T) - x(t)}{T} - x'(t) \right]^2 \right\} = 0, \quad (2.140)$$

则称在  $t$  时刻  $\{X(t)\}$  可微， $\{X'(t)\}$  为随机过程  $\{X(t)\}$  的导数。



$\{X'(t)\}$  具有如下的数字特征：

$$\begin{aligned} E \{X'(t)\} &= m'_X(t) \\ R_{X'}(t_1, t_2) &= \frac{\partial^2 R_X(t_1, t_2)}{\partial t_1 \partial t_2} E \{X'(t)\} = 0, \\ R_X(t_1, t_2) &= R_{X'}(\tau) = \lim_{T \rightarrow 0} \frac{1}{T^2} R_{x_T}(\tau), \end{aligned} \quad (2.141)$$

其中  $R_{X'}(\tau) = 2R_X(\tau) - R_X(\tau - T) - R_X(\tau + T)$ 。

将上式中的  $R_X(\tau - T)$  及  $R_X(\tau + T)$  按泰勒级数展开，忽略三次以上的微分项，可得  $R_{X'}(\tau) = -R''_X(\tau)$ .

此外，我们有时也会用到随机过程  $\{X(t)\}$  的  $n$  阶导数  $\{X^{(n)}(t)\}$  (在这里，我们假定每一步都必须满足均方可微的要求)，根据均方可微的规则，可以推得

$$R_X^{(n)}(t_1, t_2) = \frac{\partial^{2n} R_X(t_1, t_2)}{\partial t_1^{(n)} \partial t_2^{(n)}}. \quad (2.142)$$

若  $\{X(t)\}$  为平稳随机过程，那么  $R_{X^{(n)}}(\tau) = (-1)^n R_X^{2n}(\tau)$ .

### 定义 2.56 随机过程的均方积分

已知随机过程  $\{X(t)\}$ ，且

$$x_n(t) = \sum_{k=1}^n x(t'_k) (t_k - t_{k-1}), \quad (2.143)$$

其中  $\Delta t_k = t_k - t_{k-1}$  是在  $[a, b]$  内划分的微小区间， $t_{k-1} < t'_k < t_k$ . 如果存在随机过程的样本函数的定积分  $\int_a^b x(t) dt$  满足

$$\lim_{\substack{\Delta t_k \rightarrow 0 \\ n \rightarrow +\infty}} E \left\{ \left[ \sum_{k=1}^n x(t'_k) \Delta t_k - \int_a^b x(t) dt \right]^2 \right\} = 0. \quad (2.144)$$

那么将  $\int_a^b X(t) dt$  记为随机过程  $X(t)$  在  $[a, b]$  区间内的均方积分.



由于在信号统计分析中，经常涉及随机过程通过线性时变系统以后的有关统计性质，因此，有必要研究  $\{X(t)\}$  通过冲击响应为  $h(t, t')$  的线性系统输出的统计特性；这时，将涉及随机积分  $\int_a^b h(t, v)x(v)dv$ .

同样，首先定义

$$y_n(t) = \sum_{k=1}^n h(t, v_k) x(v'_k) (v_k - v_{k-1}), \quad (2.145)$$

式中， $\Delta v_k = v_k - v_{k-1}$ ,  $v_{k-1} < v'_k < v_k$ ,  $\Delta v_k$ ,  $\lim_{\Delta v_k \rightarrow 0} \left| y_n(t) - \int_a^b h(t, v)x(v)dv \right|^2 = 0$ , 则  $\int_a^b h(t, v)x(v)dv$ , 即为均方差下的随机积分。

## 2.6 高斯随机过程

在信号的统计分析与处理中，最经常遇到的随机过程是高斯随机过程。在一定的条件下，存在大量统计独立的随机变量  $X_i$  之和  $Z_n$ ,  $Z_n = \sum_{i=1}^n X_i$ , 当  $n \rightarrow +\infty$  时， $Z_n$  的分布趋于高斯分布。根据这一定理，由无数微观粒子的不规则运动产生的热噪声、散粒噪声等属于高斯随机过程。而许多干扰，如云雨杂波，地物杂波等也往往近似为高斯随机过程。在高斯假设的前提下，往往可以得到随机信号分析与处理的解析解。因此，深入全面地介绍高斯随机过程及其有关特性十分必要。

1. 高斯过程的概念第一章已经较详细地讨论过一维和多维高斯变量，现在将这一概念推广到随机过程中去。

**高斯随机过程的定义如下：**

在任意时刻  $t_1, t_2, \dots, t_n$  ( $n = 1, 2, \dots$ ),  $\{X(t)\}$  所形成的  $n$  维随机变量，其概率密度函数  $f(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)$  为高斯分布

$$f(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \frac{1}{(2\pi)^{\frac{n}{2}} |C_X|^{\frac{1}{2}}} \exp \left[ -\frac{1}{2} (\mathbf{X} - \mathbf{m}_X)^\top C_X^{-1} (\mathbf{X} - \mathbf{m}_X) \right]. \quad (2.146)$$

其中  $\mathbf{X} = [X_1, X_2, \dots, X_n]^\top$ ,  $\mathbf{m}_X = E\{X\} = [E\{X_1\}, E\{X_2\}, \dots, E\{X_n\}]^\top$ ,  $C_X = E\{(X - m_X)(X - m_X)^\top\}$ , 随机过程  $\{X(t)\}$  为高斯过程。

高斯随机过程具有以下性质：

(1) 对于高斯随机过程来说，如果知道了  $E\{X(t)\}$  及  $C_X(t_1, t_2)$ , 其中  $t_1, t_2$  为任意值，那么随机过程的整个概率特性也就完全被决定了。

(2) 若高斯随机过程是广义平稳的，那么，它一定是严格平稳的，其输出仍然是高斯随机过程。

(3) 一个高斯随机过程经过任意线性变换(如线性相加、线性放大、微分和积分), 其输出仍然是高斯随机过程。

中心极限定理已证明, 大量独立的、均匀微小的随机变量之和近似地服从高斯分布。高斯分布是在实际应用中最常遇到的、最重要的分布。同样, 在电子系统中遇到最多的过程也是高斯过程。如电路中最常见的电阻热噪声、电子管(或晶体管)的散粒噪声, 如大气和宇宙噪声, 以及许多积极干扰、消极干扰(包括云雨杂波、地物杂波等)也都可以近似为高斯过程。另一方面, 只有高斯过程的统计特性最简便, 故常用作噪声的理论模型。高斯过程将是以后各章节中的一个主要研讨对象。

### 定义 2.57 高斯随机过程

任意  $n$  维分布都是高斯分布的随机过程, 称为高斯随机过程。



(2) 高斯过程  $X(t)$  的  $n$  维概率密度, 如式(2.146), 式中  $\mathbf{M}_X$  是  $n$  维期望矢量,  $\mathbf{C}$  是协方差矩阵。

$$\mathbf{M}_X = \begin{bmatrix} E[X(t_1)] \\ \vdots \\ E[X(t_n)] \end{bmatrix} = \begin{bmatrix} m_X(t_1) \\ \vdots \\ m_X(t_n) \end{bmatrix}_{n \times 1}, \quad \mathbf{C} = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix}_{n \times n}. \quad (2.147)$$

$$\begin{aligned} C_{ik} &= C_X(t_i, t_k) = E[[X(t_i) - m_X(t_i)][X(t_k) - m_X(t_k)]] \\ &= R_X(t_i, t_k) - m_X(t_i)m_X(t_k). \end{aligned} \quad (2.148)$$

从定义式(2.147)中可看出, 高斯过程的  $n$  维分布完全由均值矢量  $\mathbf{M}_X$  与协方差矩阵  $\mathbf{C}$  所确定, 且有关时间  $(t_1, \dots, t_n)$  的因素, 全部包含在  $\mathbf{M}_X$  和  $\mathbf{C}$  中。

### (3) 平稳高斯过程

若高斯过程  $X(t)$  的数学期望是常数, 自相关函数只与时间差值  $\tau$  有关, 即满足

$$\begin{cases} m_X(t) = m_X, \\ R_X(t_i, t_k) = R_X(\tau_{k-i}), \quad , \quad \tau_{k-i} = t_k - t_i \quad (i, k = 1, 2, \dots, n, \dots) . \\ E[X^2(t)] = R_X(0) < \infty. \end{cases} \quad (2.149)$$

则此高斯过程是宽平稳的。

### (4) 平稳高斯过程的 $n$ 维特征函数

$$Q(u_1, \dots, u_n; \tau_1, \dots, \tau_{n-1}) = \exp \left[ jm_X \sum_{i=1}^n u_i - \frac{1}{2} \sum_{i=1}^n \sum_{k=1}^n C(\tau_{k-i}) u_i u_k \right]. \quad (2.150)$$

## 2. 高斯过程的性质

### (1) 高斯过程的宽平稳与严平稳等价

证明: 在高斯过程的  $n$  维概率密度中, 与时间有关的两个参量是  $\mathbf{M}_X$  和  $\mathbf{C}$ 。因为高斯过程宽平稳, 所以其均值矢量  $\mathbf{M}_X = [m_X, m_X, \dots, m_X]^T$  为常数矢量, 矩阵  $\mathbf{C}$  中的每一个元素  $C_{ik}$  仅取决于时间差  $\tau_1, \tau_2, \dots, \tau_{n-1}$ , 而与时间的起点无关, 即

$$C_{ik} = C(t_i, t_k) = R(\tau_{k-i}) - m_X^2, i, k = 1, 2, \dots, n \quad (2.151)$$

因此, 宽平稳高斯过程的  $n$  维概率密度仅仅是时间差的函数。有

$$f_X(x_1, x_2, \dots, x_n, t_1, t_2, \dots, t_n) = f_X(x_1, x_2, \dots, x_n; \tau_1, \tau_2, \dots, \tau_{n-1}). \quad (2.152)$$

当高斯过程  $X(t)$  的  $n$  维概率密度的取样点  $X_1, X_2, \dots, X_n$  在时间轴上作任意  $\Delta t$  平移后, 由于时间差  $\tau'_{\mu-i} = (t_k + \Delta t) - (t_i + \Delta t) = t_k - t_i = \tau_{k-i}$  不随时间平移  $\Delta t$  变化, 所以过程  $X(t)$  的  $n$  维概率密度也不随时间平移  $\Delta t$  变化。据严平稳定义可知, 满足上述条件的过程是严平稳的。

### (2) 高斯过程在不同时刻状态间的互不相关和独立等价刻画

证明: 设高斯过程  $X(t)$  的  $n$  个不同时刻  $t_1, t_2, \dots, t_n$  的状态为  $X(t_1), X(t_2), \dots, X(t_n)$ 。由高斯过程定义可知, 它们都是高斯变量。

当所有状态互不相关时, 协方差矩阵

$$\mathbf{C} = \begin{bmatrix} \sigma_X^2(t_1) & 0 & \cdots & 0 \\ 0 & \sigma_X^2(t_2) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_X^2(t_n) \end{bmatrix}. \quad (2.153)$$

其中

$$\begin{cases} C_{it} = C_X(t_i, t_k) = E\{[X(t_i) - m_X(t_i)][X(t_k) - m_X(t_k)]\}. \\ C_n = C_X(t_i, t_i) = E\{[X(t_i) - m_X(t_i)]^2\} = \sigma_X^2(t_i). \end{cases} \quad (2.154)$$

代入  $n$  维概率密度表达式, 并展开

$$\begin{aligned} f_X(x_1, x_2, \dots, x_n, t_1, t_2, \dots, t_n) &= \frac{1}{(2\pi)^{n/2}\sigma_X(t_1)\sigma_X(t_2)\cdots\sigma_X(t_n)} \exp\left[-\frac{1}{2}\sum_{i=1}^n \frac{(x_i - m_X(t_i))^2}{\sigma_X^2(t_i)}\right] \\ &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_X(t_i)} \exp\left[-\frac{(x_i - m_X(t_i))^2}{2\sigma_X^2(t_i)}\right] \\ &= f_X(x_1, t_1)f_X(x_2, t_2)\cdots f_X(x_n, t_n). \end{aligned} \quad (2.155)$$

由上式可见, 在  $C_{ik} = 0 (i \neq k)$  的条件下,  $n$  维概率密度等于  $n$  个一维概率密度的乘积, 满足独立的条件。故对高斯过程来说, 不同时刻状态间的互不相关与独立是等价的。

### (3) 平稳高斯过程与确定信号之和仍是高斯过程, 但不一定平稳

在通信和雷达等系统中, 从噪声  $X(t)$  背景中接收、检测有用信号  $s(t)$  时, 往往需要处理的是噪声与信号叠加在一起的随机信号。而噪声  $X(t)$  常常认为是高斯过程。

证明: 设合成的随机信号  $Y(t) = X(t) + s(t)$ 。若已知  $f_X(x; t)$  为噪声  $X(t)$  的一维概率密度; 因  $s(t)$  是确定信号, 故其概率密度可表示为  $\delta[s - s(t)]$ 。利用独立和的卷积公式, 可得到合成信号  $Y(t)$  的一维概率密度为

$$f_Y(y; t) = \int_{-\infty}^{\infty} f_X(x; t) \delta[y - s(t) - x] dx = f_X[y - s(t)]. \quad (2.156)$$

可见, 当  $f_X(x; t)$  为高斯分布时, 合成信号的一维分布  $f(y; t)$  服从高斯分布。

同理可得, 合成信号  $Y(t)$  的二维概率密度

$$f_Y(y_1, y_2; t_1, t_2) = f_X(y_1 - s(t_1), y_2 - s(t_2)). \quad (2.157)$$

也是服从高斯分布的。

依此类推, 只要用  $y_i - s_i(t)$  代替  $x_i$ , 就可得到合成信号  $Y(t)$  的  $n$  维概率密度, 即合成信号的  $n$  维分布也是服从高斯分布。

$$f_Y(y_1, \dots, y_n; t_1, \dots, t_n) = f_X(y_1 - s(t_1), \dots, y_n - s(t_n)). \quad (2.158)$$

还应指出, 虽然平稳高斯过程与确定信号之和的概率分布仍为高斯分布, 但是, 一般情况下的合成信号不再是平稳过程而平稳高斯过程与随相余弦信号的合成, 虽已不再服从高斯分布, 是宽平稳的随机过程。

(4) 若高斯序列矢量  $\{\mathbf{X}(n)\}$  均方收敛于随机矢量  $\mathbf{X}$  (即  $E[|\lim_{n \rightarrow \infty} \{X_n\} - X|^2] = 0$ ), 则这个随机矢量  $\mathbf{X}$  服从高斯分布

证明: 设  $k$  维高斯序列矢量  $\{\mathbf{X}(n)\} = \begin{pmatrix} \{X_1(n)\} \\ \{X_2(n)\} \\ \vdots \\ \{X_k(n)\} \end{pmatrix}_{k \times 1}$  均方收敛于  $k$  维随机矢量

$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{bmatrix}_{k \times 1}$ , 其中高斯序列矢量  $\{\mathbf{X}(n)\}$  的每个分量  $\{\mathbf{X}_i(n)\}$  都是高斯序列, 它均方收

敛于随机矢量  $\mathbf{X}$  的分量  $\mathbf{X}_i$ 。若  $\{\mathbf{X}(n)\}$  和  $\mathbf{X}$  的均值矢量和方差阵分别记为

$$E[\{\mathbf{X}(n)\}] = \begin{bmatrix} m_1(n) \\ m_2(n) \\ \vdots \\ m_k(n) \end{bmatrix} = \mathbf{M}(n), \quad E[\mathbf{X}] = \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_k \end{bmatrix} = \mathbf{M}. \quad (2.159)$$

$$E[(\mathbf{X}(n) - \mathbf{M})(\mathbf{X}(n) - \mathbf{M})^T] = \mathbf{C}(n), \quad E[(\mathbf{X} - \mathbf{M})(\mathbf{X} - \mathbf{M})^T] = \mathbf{C}. \quad (2.160)$$

由于高斯序列矢量  $\{\mathbf{X}(n)\}$  的每个分量  $\{\mathbf{X}_i(n)\}$  均方收敛于随机矢量  $\mathbf{X}$  的分量  $\mathbf{X}_i$ , 即

$$\lim_{n \rightarrow \infty} E[|X_i(n) - X_i|^2] = 0, \quad i = 1, \dots, k. \quad (2.161)$$

根据均方连续的推论 2 可知, 随机序列  $\{\mathbf{X}_i(n)\}$  的数学期望也连续, 即

$$\lim_{n \rightarrow \infty} E[X_i(n)] = E[X_i], \quad i = 1, \dots, k. \quad (2.162)$$

则

$$\begin{aligned} \lim_{n \rightarrow \infty} m_i(n) &= m_i, \quad i, j = 1, \dots, k \\ \lim_{n \rightarrow \infty} C_{ij}(n) &= C_{ij}. \end{aligned} \quad (2.163)$$

从而, 高斯序列矢量  $\{\mathbf{X}(n)\}$  的均值矢量和方差阵, 有

$$\begin{cases} \lim_{n \rightarrow \infty} \mathbf{M}(n) = \mathbf{M} \\ \lim_{n \rightarrow \infty} \mathbf{C}(n) = \mathbf{C} \end{cases}. \quad (2.164)$$

若以  $Q_n(u_1, u_2, \dots, u_k)$  和  $Q_X(u_1, u_2, \dots, u_k)$  分别代表  $\{\mathbf{X}(n)\}$  和  $\mathbf{X}$  的  $k$  维特征函数, 由于  $\{\mathbf{X}(n)\}$  为  $k$  维高斯分布的随机序列矢量, 故

$$Q_n(u_1, u_2, \dots, u_k) = \exp \left[ j \mathbf{U}^T \mathbf{M}(n) - \frac{1}{2} \mathbf{U}^T \mathbf{C}(n) \mathbf{U} \right]. \quad (2.165)$$

对上式两边求极限

$$\begin{aligned} \lim_{n \rightarrow \infty} Q_n(u_1, u_2, \dots, u_k) &= \exp \left\{ j \mathbf{U}^T \left[ \lim_{n \rightarrow \infty} \mathbf{M}(n) \right] - \frac{1}{2} \mathbf{U}^T \left[ \lim_{n \rightarrow \infty} \mathbf{C}(n) \right] \mathbf{U} \right\} \\ &= \exp \left[ j \mathbf{U}^T \mathbf{M} - \frac{1}{2} \mathbf{U}^T \mathbf{C} \mathbf{U} \right]. \end{aligned} \quad (2.166)$$

由于  $\{\mathbf{X}(n)\}$  均方收敛于  $\mathbf{X}$ , 则有

$$\lim_{n \rightarrow \infty} Q_n(u_1, u_2, \dots, u_k) = Q_X(u_1, u_2, \dots, u_k). \quad (2.167)$$

对照上述两式右端, 可得

$$Q_X(u_1, u_2, \dots, u_k) = \exp \left[ jU^T M - \frac{1}{2} U^T C U \right]. \quad (2.168)$$

所以,  $\mathbf{X}$  也是  $k$  维高斯分布的随机矢量。

(5) 均方可微高斯过程的导数是高斯过程

证明: 设高斯过程  $[X(t), t \in T]$  在  $T$  上均方可微。在  $T$  上任取  $t_1$  使  $\{t_1 + \Delta t, \dots, t_t + \Delta t, \dots, t_A + \Delta t\} \in T$ , 构造随机矢量

$$\left[ \frac{X(t_1 + \Delta t) - X(t_1)}{\Delta t}, \dots, \frac{X(t_i + \Delta t) - X(t_i)}{\Delta t}, \dots, \frac{X(t_k + \Delta t) - X(t_k)}{\Delta t} \right]^T,$$

上式是  $k$  维高斯矢量  $[X(t_1), \dots, X(t_i), \dots, X(t_k)]^T$  的线性组合, 所以也是  $k$  维高斯矢量。由于高斯过程  $\mathbf{X}(t)$  均方可微, 故对每个  $t$  而言,  $\frac{X(t_i + \Delta t) - X(t_i)}{\Delta t}$  均方收敛于  $X'(t_i)$  ( $i = 1, \dots, k$ ).

根据性质 (4), 上式构造的高斯矢量的均方极限  $[X'(t_1), X'(t_2), \dots, X'(t_k)]^T$  也是  $k$  维高斯随机矢量, 即  $\mathbf{X}'(t)$  是一个高斯随机过程。

(6) 均方可积高斯过程的积分是高斯过程

高斯过程  $\mathbf{X}(t)$  的均方积分为

$$Y(t) = \int_a^t X(\lambda) d\lambda, a, t \in T. \quad (2.169)$$

证明: 对任意的  $t_1, t_2, \dots, t_k \in T$ , 使  $\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_n$  为  $[a, t]$  区间上的一系列采样点, 即  $a = \lambda_0 < \lambda_1 < \lambda_2 < \dots < \lambda_n < t_3$  ( $j = 1, \dots, k$ )。构造一个线性组合, 令

$$\sum_{i=0}^{m_j} X(\lambda_i) \Delta \lambda_i = Y(n_j), \quad j = 1, \dots, k \quad (2.170)$$

由于  $\mathbf{X}(t)$  是高斯过程、 $\mathbf{X}(\lambda_i)$  是高斯变量, 因此  $\mathbf{X}(\lambda_i)$  的线性组合  $\{Y(n_j)\}$  也是高斯变量, 而  $\{Y(n_j)\}$ ,  $n_j = 0, 1, 2, \dots$  为高斯序列。所以,  $[\{Y(n_1)\}, \{Y(n_2)\}, \dots, \{Y(n_k)\}]^T$  所组成的随机矢量是  $k$  维高斯序列矢量。据随机过程积分的定义, 由于  $\mathbf{X}(t)$  在  $T$  上均方可积, 故对每个  $t_j$  而言, 有

$$\lim_{n_j \rightarrow \infty} Y(n_j) = \lim_{\Delta \lambda \rightarrow 0} \sum_{i=0}^{n_j} X(\lambda_i) \Delta \lambda_i = \int_a^{u_j} X(\lambda) d\lambda = Y(t_j), \quad (2.171)$$

即高斯序列  $\{Y(n_j)\}$  均方收敛于  $Y(t_j)$  ( $j = 1, \dots, k$ ).

由性质 (4) 可得, 高斯序列矢量  $[\{Y(n_1)\}, \{Y(n_2)\}, \dots, \{Y(n_k)\}]^T$  的均方极限  $[Y(t_1), Y(t_2), \dots, Y(t_k)]^T$  也是  $k$  维高斯矢量, 即  $\mathbf{Y}(t)$  为高斯过程。

(7) 同理, 可以证明在确定区间  $[a, b]$  上, 高斯过程  $\mathbf{X}(t)$  与权函数  $h(\lambda, t)$  (它是  $\lambda, t$  的连续函数) 乘积的积分

$$Y(t) = \int_a^b X(\lambda) h(\lambda, t) d\lambda. \quad (2.172)$$

仍为高斯过程。这样, 若高斯过程  $\mathbf{X}(t)$  在  $-\infty < t < \infty$  上均方可积, 则

$$Y(t) = \int_{-\infty}^t X(\lambda) h(\lambda, t) d\lambda \quad (2.173)$$

也为高斯过程。由此可见, 高斯过程经过积分变换后仍为高斯过程。

**例 2.58** 已知平稳高斯过程  $\mathbf{X}(t)$  的均值为零, 它的自相关函数  $R_X(\tau) = \frac{1}{4} \exp(-2|\tau|)$ 。求在  $t_1$  时刻,  $\mathbf{X}(t_1)$  取值在 0.5 与 1 之间的概率。

解: 高斯过程  $\mathbf{X}(t)$  在  $t_1$  时刻的状态为高斯变量  $X(t_1)$ , 其均值  $m_X = 0$ , 均方值  $R_X(0) = 1/4$ , 方差  $\sigma_X^2 = R_X(0) - m_X^2 = 1/4$ 。则

$$\begin{aligned} P\{0.5 \leq X(t_1) \leq 1\} &= P\{\sigma_X \leq X(t_1) \leq 2\sigma_X\} \\ &= \Phi\left(\frac{2\sigma_X - m_X}{\sigma_X}\right) - \Phi\left(\frac{\sigma_X - m_X}{\sigma_X}\right) \\ &= \Phi(2) - \Phi(1) = 0.1359. \end{aligned} \quad (2.174)$$

**例 2.59** 已知随机过程  $\mathbf{X}(t) = A \cos \omega_0 t + B \sin \omega_0 t$ , 其中  $A$  与  $B$  是相互独立的高斯变量, 且  $E[A] = E[B] = 0$ ,  $E[A^2] = E[B^2] = \sigma^2$ ,  $\omega_0$  为常数。求此过程  $\mathbf{X}(t)$  的一、二维概率密度。

解: 在任意时刻  $t$  对过程  $\mathbf{X}(t)$  进行采样, 由于它是高斯变量  $A$  与  $B$  的线性组合, 故  $\mathbf{X}(t)$  也是个高斯变量。从而可知,  $\mathbf{X}(t)$  是一高斯过程。为确定高斯过程  $\mathbf{X}(t)$  的概率密度, 只要求出  $\mathbf{X}(t)$  的均值和协方差函数即可。

可求得  $\mathbf{X}(t)$  的均方值和方差为

$$\begin{aligned} E[X(t)] &= E[A \cos \omega_0 t + B \sin \omega_0 t] \\ &= E[A] \cos \omega_0 t + E[B] \sin \omega_0 t = 0. \end{aligned}$$

$$\begin{aligned} R_X(t, t+\tau) &= E[X(t)X(t+\tau)] \\ &= E[(A \cos \omega_0 t + B \sin \omega_0 t)(A \cos \omega_0(t+\tau) + B \sin \omega_0(t+\tau))] \\ &= E[A^2] \cos \omega_0 t \cos \omega_0(t+\tau) \\ &\quad + E[B^2] \sin \omega_0 t \sin \omega_0(t+\tau) \\ &\quad + E[AB] \cos \omega_0 t \sin \omega_0(t+\tau) \\ &\quad + E[AB] \sin \omega_0 t \cos \omega_0(t+\tau). \end{aligned} \quad (2.175)$$

因  $A$  与  $B$  独立, 有  $E[AB] = E[A]E[B] = 0$ , 则

$$\begin{aligned} R_X(t, t + \tau) &= E[A^2] \cos \omega_0 t \cos \omega_0(t + \tau) + E[B^2] \sin \omega_0 t \sin \omega_0(t + \tau) \\ &= \sigma^2 \cos \omega_0 \tau = R_X(\tau). \end{aligned} \quad (2.176)$$

可求得  $\mathbf{X}(t)$  的均方值和方差为

$$\psi_x^2 = R_X(0) = \sigma^2 < \infty, \quad \sigma_x^2 = R_X(0) - m_x^2 = \sigma^2. \quad (2.177)$$

由上可知, 高斯过程  $\mathbf{X}(t)$  为平稳过程, 其一维概率密度为

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right). \quad (2.178)$$

其二维均值矢量和协方差矩阵为

$$\mathbf{M}_X = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} \sigma^2 & \sigma^2 \cos \omega_0 \tau \\ \sigma^2 \cos \omega_0 \tau & \sigma^2 \end{pmatrix}. \quad (2.179)$$

则其二维概率密度为

$$f_{\mathbf{X}}(x_1, x_2, \tau) = \frac{1}{2\pi\sigma^2\sqrt{1-\cos^2\omega_0\tau}} \exp\left(-\frac{x_1^2 - 2x_1x_2 \cos \omega_0 \tau + x_2^2}{2\sigma^2(1 - \cos^2 \omega_0 \tau)}\right). \quad (2.180)$$

## 2.7 各态历经过程

研究随机过程的统计特性, 从理论上说需要知道过程的  $n$  维概率密度或  $n$  维分布函数, 或者要知道所有的样本函数。这一点在实际问题中往往办不到, 因为这需要对一个过程进行大量重复的实验或观察, 甚至要求实验次数  $N \rightarrow \infty$  才能达到需要。因而, 促使人们提出这样一个问题, 能否用在一段时间范围内观察到的一个样本函数作为提取整个过程数字特征的充分依据? 俄国概率学家辛钦的证明: 当具备一定的补充条件时, 有一种平稳随机过程, 对其任一个样本函数所作的各种时间平均, 从概率意义上趋近于此过程的各种统计平均。对具有这一特性的随机过程, 称之为具有各态历经性的随机过程 (或遍历过程)。从字面上可以理解为: 遍历过程的各个样本函数都同样地经历了整个过程的所有可能状态。

### 2.7.1 各态历经过程

因此, 这类随机过程的任何一个样本函数中含有整个过程的全部统计信息, 即可以用它的任何一个样本函数的时间平均来代替它的统计平均。

**例 2.7.1**

在较长时间  $T$  内，观测工作在稳定状态下的一个噪声二极管的输出电压。欲求时间平均，理论上应该满足下式

$$\overline{x_k(t)} = \frac{1}{T} \int_0^T x_k(t) dt. \quad (2.181)$$

如图 2-20 所示，由于  $x_k(t)$  是噪声电压，所以写不出样本函数  $x_k(t)$  的表达式。

因此，只有对  $x_k(t)$  进行采样。将  $T$  分成  $n$  等份（这个  $n$  应相当大），对在  $T$  时间内采得的  $n$  个电压值  $x_1, \dots, x_n$  进行算术平均才可行。

$$\overline{x'_k(t)} = \frac{\sum_{i=1}^n x_i}{n} \approx \overline{x_k(t)}, \quad (2.182)$$

即以算术平均  $\overline{x'_k(t)}$  来近似时间平均  $\overline{x_k(t)}$ 。

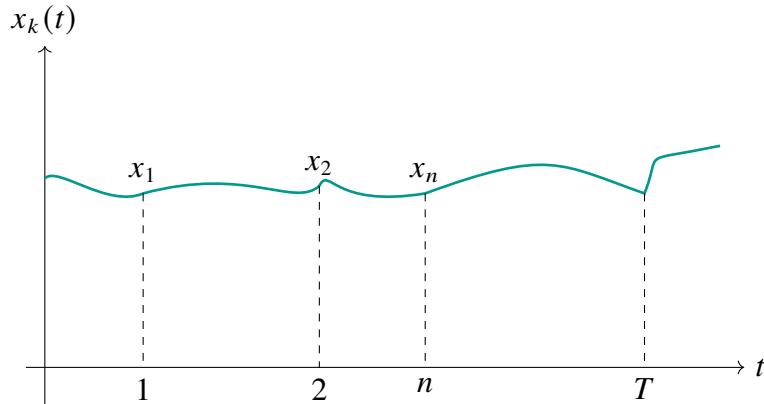


图 2-20 噪声电压

在工作条件不变的情况下，假如对同一个工作在稳定状态下的噪声二极管进行  $n$  次独立重复的试验，取出它的  $n$  条样本函数，如图 2-21 所示，并对任一时刻  $t'$  的状态  $X(t')$  的所有取值进行统计平均

$$E[X(t')] = \sum_{i=1}^n x'_i P\{X(t') = x'_i\} = \frac{1}{n} \sum_{i=1}^n x'_i. \quad (2.183)$$

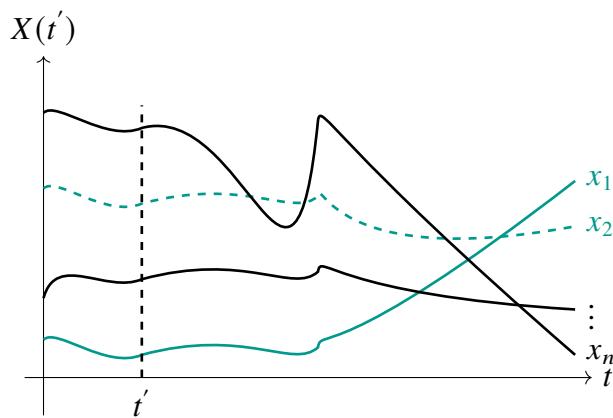


图 2-21 \$n\$ 条样本函数

如果 \$T\$ 取得相当长, 这些样本函数可以看成是从 \$x\_k(t)\$ 上一段段截取下来的, 又由于工作在稳定状态时, 可以把它看成是平稳的, 所以其特性与起点无关。那么只要 \$n\$ 取得相当大, 就找不出任何理由说, 前一种方法得到的时间平均 \$x(t)\$ 与后一方法得到的统计平均值 \$E[X(t')]\$ 有什么差别。也就是说, 当 \$T \rightarrow \infty, n \rightarrow \infty\$ 时, 在概率意义上, 下式成立

$$P\left\{E[X(t)] = \overline{x_k(t)}\right\} = 1. \quad (2.184)$$

噪声电压在时间上的平均值与它的统计平均值相等, 这就是所谓的噪声电压的均值各态历经性。

#### 定义 2.60 严各态历经过程

如果一个平稳随机过程 \$X(t)\$, 它的各种时间平均 (时间足够长) 以概率 1 收敛于相应的统计平均, 即

$$P\left\{\lim_{n \rightarrow \infty} \text{各种时间平均} = \text{相应的统计平均}\right\} = 1. \quad (2.185)$$

则称过程 \$X(t)\$ 具有严格的各态历经性, 或称此过程为严(或狭义)各态历经过程。

#### 2.7.2 广义各态历经过程

但是, 如同在前面讨论随机过程的平稳性时曾经指出的理由一样, 工程上往往只在相关理论的范围内考虑各态历经过程, 称之为宽(或广义)各态历经过程。下面首先引入随机过程的时间平均概念, 然后给出宽各态历经过程的定义。

##### 2. 宽各态历经过程

### (1) 时间均值

一般来说, 若对一个确定的时间函数  $x(t)$  求时间均值, 即

$$\overline{x(t)} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t) dt = m, \quad (2.186)$$

其结果  $m$  是个确定的常数。但由于随机过程  $X(\zeta, t)$  是随时间变化的随机变量, 对不同的试验结果  $\zeta_k \in \Omega$ , 有不同的确定时间函数  $x_k(t)$  与其对应, 所以, 若对数随机过程  $X(\zeta, t)$  求时间均值, 即

$$\overline{X(t)} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T X(t) dt = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T X(\zeta, t) dt = M(\zeta), \quad (2.187)$$

则  $M(\zeta)$  是个随机变量, 对每一个试验结果  $\zeta_k \in \Omega$ , 都有一个确定值  $m_k$  与其对应。其中符号  $(\cdot)$  表示求时间平均。

### (2) 时间自相关

若对确定函数  $x(t)$  求时间自相关, 则

$$\overline{x(t)x(t+\tau)} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t)x(t+\tau) dt = f(\tau), \quad (2.188)$$

其结果  $f(\tau)$  是个确定的时间函数。

若对随机过程  $X(\zeta, t)$  求时间自相关, 则

$$\begin{aligned} \overline{X(t)X(t+\tau)} &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T X(t)X(t+\tau) dt \\ &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T X(\zeta, t)X(\zeta, t+\tau) dt = f(\zeta, \tau). \end{aligned} \quad (2.189)$$

由于  $f(\zeta, \tau)$  对每一个试验结果  $\zeta_i \in \Omega$  而言, 都有一个确定的时间函数  $f_i(\tau)$  与其对应, 所以, 随机过程的时间自相关函数一般是个随机过程。

### (3) 宽各态历经的定义

#### 定义 2.61 均值的各态历经性

设  $X(t)$  是一个平稳随机过程, 如果

$$\overline{X(t)} = E[X(t)] = m_X \quad (2.190)$$

以概率 1 成立, 则称过程  $X(t)$  的均值具有各态历经性。



### 定义 2.62 自相关函数的各态历经性

如果

$$\overline{X(t)X(t+\tau)} = E[X(t)X(t+\tau)] = R_X(\tau) \quad (2.191)$$

以概率 1 成立, 则称过程  $X(t)$  的自相关函数具有各态历经性。

### 定义 2.63 均方值的各态历经性

若在  $\tau = 0$  时, 上式成立, 则称过程  $X(t)$  的均方值具有各态历经性。

### 定义 2.64 宽(或广义)各态历经过程

如果过程  $X(t)$  的均值和自相关函数都具有各态历经性, 则称  $X(t)$  为宽(或广义)各态历经过程。

今后, 凡提到“各态历经”一词时, 除非特别指出, 否则皆指宽各态历经过程。

### 3. 各态历经性的实际意义

#### (1) 随机过程的各态历经性具有重要的实际意义

对一般随机过程而言, 其时间平均是个随机变量。可是, 对各态历经过程来说, 由上述定义时间平均, 得到的结果趋于一个非随机的确定量, 见式(2.190)。由式(2.184)和式(2.190)可得

$$P\left\{\overline{X(t)} = \overline{x_k(t)}\right\} = 1, \quad (2.192)$$

即

$$\overline{X(t, \zeta)} \stackrel{\text{a. e.}}{=} \overline{x_k(t)}. \quad (2.193)$$

上式表明: 各态历经过程诸样本函数的时间平均, 实际上可认为是相同的。因此, 各态历经过程的时间平均就可由它的任一样本函数的时间平均来表示。这样, 对各态历经过程, 可以直接用它的任一个样本函数的时间平均来代替对整个过程统计平均的研究, 故有

$$E[X(t)] = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t) dt. \quad (2.194)$$

$$R_X(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t)x(t+\tau) dt. \quad (2.195)$$

实际上, 这也正是引出各态历经性概念的重要目的, 从而给解决许多工程问题带来极大的方便。

**例 2.65** 测得接收机的噪声, 用一般的方法, 就需要用数量极多的相同的接收机, 在同一条件下同时进行测量和记录, 用统计方法算出所需的数学期望、相关函数等数字特征; 而利

用噪声过程的各态历经性则可以只用一部接收机，在不变的条件下，对其输出噪声作长时间的记录，后用求时间平均的方法，即可求出数学期望、相关函数等数字特征，这就使得工作大大地简化。

当然，在实际工作中，由于对随机过程的观察时间总是有限的，因此，用有限的时间代替无限长的时间，会给结果带来一定的误差，然而，只要所取时间足够长，结果定能满足实际要求。

(2) 电子技术中，噪声电压(或电流)的各态历经过程  $X(t)$  的数字特征的物理意义

① 数学期望代表噪声电压(或电流)的直流分量  $\bar{X}(t) = E[X(t)] = m_X$ 。

② 均方值代表噪声电压(或电流)消耗在 1 欧姆电阻上的总平均功率，令  $\tau = 0$ ，则自相关函数有

$$R_X(0) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t)x(t+\tau) dt \Big|_{\tau=0} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x^2(t) dt. \quad (2.196)$$

可见， $R_X(0)$  代表噪声电压(或电流)消耗在  $1 \Omega$  电阻上的总平均功率。

③ 方差代表噪声电压(或电流)消耗在  $1 \Omega$  电阻上的交流平均功率。

$$\sigma_X^2 = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T [x(t) - m_X]^2 dt, \quad (2.197)$$

$\sigma_X^2$  代表噪声电压(或电流)消耗在  $1 \Omega$  电阻上的交流平均功率。标准差  $\sigma_X$  则代表噪声电压(或电流)的有效值。

**例 2.66** 设随机过程  $X(t) = a \cos(\omega_0 t + \Phi)$ ，式中  $a, \omega_0$  皆为常数， $\Phi$  服从  $(0, 2)$  上均匀分布的随机变量。判断  $X(t)$  是否宽各态历经，给出理由。

解：由于

$$\begin{aligned} \bar{X}(t) &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T a \cos(\omega_0 t + \Phi) dt = \lim_{T \rightarrow \infty} \frac{a \cos \Phi \sin \omega_0 T}{\omega_0 T} = 0. \\ \bar{X}(t)X(t+\tau) &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T a \cos(\omega_0 t + \Phi) a \cos[\omega_0(t+\tau) + \Phi] dt \\ &= \frac{a^2}{2} \cos \omega_0 \tau. \end{aligned} \quad (2.198)$$

由例 2.20 的结果，可得

$$\begin{aligned} \bar{X}(t) &= E[X(t)] = 0. \\ \bar{X}(t)X(t+\tau) &= E[X(t)X(t+\tau)] = \frac{a^2}{2} \cos \omega_0 \tau. \end{aligned} \quad (2.199)$$

所以，过程  $X(t)$  具有宽各态历经性。

**例 2.67** 随机过程  $X(t) = Y$ ，其中  $Y$  是方差不为零的随机变量，讨论过程的各态历经性。

解: 由例 2.21 可知

$$\begin{aligned} E[X(t)] &= E[Y] = \text{常数} \\ E[X(t)X(t+\tau)] &= E[Y^2] = \text{常数}. \end{aligned} \quad (2.200)$$

故过程  $X(t)$  为宽平稳的。然而, 因

$$\overline{X(t)} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T Y dt = Y. \quad (2.201)$$

可见,  $X(t)$  是随机变量  $Y$ , 且时间均值随  $Y$  的取值不同而变化,  $\overline{X(t)} \neq E[X(t)]$ . 所以,  $X(t)$  不是宽各态历经过程。此例表明: 平稳过程不一定具有各态历经性。

#### 4. 随机过程具备各态历经性的条件

##### (1) 各态历经过程一定是平稳过程

各态历经过程一定是平稳随机过程, 但平稳随机过程并不一定具备各态历经性。

简单说明: 由均值各态历经定义可知, 时间均值必定是个与时间无关的常数; 由时间自相关各态历经的定义可知, 时间自相关函数必定只是时间差  $\tau$  的单值函数。这就是说, 因为各态历经过程的数学期望是个常数, 其相关函数仅是  $\tau$  的单值函数, 所以, 它必定是个平稳随机过程。但平稳过程不一定具有各态历经性, 如例 2.67 所示情况。

##### (2) 均值的各态历经性定理

平稳随机过程的均值具有各态历经性的充要条件为

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^{2T} \left(1 - \frac{\tau}{2T}\right) [R_X(\tau) - m_X^2] d\tau = 0. \quad (2.202)$$

##### (3) 自相关函数 $R_X(\tau)$ 具有各态历经性的充要条件

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^{2T} \left(1 - \frac{\tau}{2T}\right) [R_\Phi(\tau) - E^2[\Phi(t)]] dz = 0, \quad (2.203)$$

式子中

$$\begin{aligned} \Phi(t) &= X(t + \tau_1) X(t), \\ R_\Phi(\tau) &= E[\Phi(t + \tau)\Phi(t)], \\ E[\Phi(t)] &= E[X(t + \tau_1) X(t)] = R_X(\tau_1). \end{aligned} \quad (2.204)$$

##### (4) 平稳高斯过程具有各态历经性的充分条件

$$\int_0^\infty |R_X(\tau)| d\tau < \infty, \quad (2.205)$$

在平稳高斯过程的均值为零, 自相关函数  $R_X(\tau)$  连续的条件下, 可以证明式(2.205)成立。

在实际应用中, 要想从理论上确切地证明一个平稳过程是否满足这些条件, 并非易事。事实上, 由于同一随机过程中各样本都出于同一随机因素, 因而各样本函数都具有相同的概率分布特性, 可以认为所遇到的大多数平稳过程都具有各态历经性。因此, 常常凭经验把各态历经性作为一种假设, 然后, 再根据实验来检验此假设是否合理。

**例 2.68** 已知随机电报信号  $X(t)$  的均值和相关函数分别为  $E[X(t)] = 0$  和  $R_X(\tau) = e^{-a|\tau|}$ , 判断  $X(t)$  是否均值各态历经, 给出理由。

解: 由均值的各态历经性定理

$$\begin{aligned} \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^{2T} \left(1 - \frac{\tau}{2T}\right) [e^{-a|\tau|} - 0^2] d\tau &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^{2T} e^{-a\tau} \left(1 - \frac{\tau}{2T}\right) d\tau \\ &= \lim_{T \rightarrow \infty} \left[ \frac{1}{aT} - \frac{1 - e^{-2aT}}{2a^2 T^2} \right] = 0. \end{aligned} \quad (2.206)$$

因此, 随机过程  $X(t)$  的均值具有各态历经性。

<pre> ans1 = 1/2*(2*a*T-1+exp(-2*a*T))/T^2/a^2 </pre>	<pre> 1 -      syms T a tau 2 -      ans1=int((1-tau/(2*T))*exp(-a*tau),0,2*T)/T 3 -      pretty(ans1) </pre> $\frac{2 a T - 1 + \exp(-2 a T)}{T^2 a^2}$
---	--

图 2-22 各态历经中积分的计算

## 2.8 小结

随机信号和统计信号处理是信号处理和机器学习的数学基础, 有了随机信号和统计信号处理课程构建出的思维, 会容易很难理解目前机器学习, 深度学习和神经网络的公共数学基础及其推导过程。在学习随机信号和统计信号处理前, 需要具备两方面的知识: 一是理解随机变量, 二是概率论的基础理论。随机信号和统计信号处理中的估计理论还需要一些数字信号处理中的 Z 变换、傅立叶变换和滤波器等知识。

## 2.9 习题

- 练习 2.1 已知随机过程  $X(t) = A \cos \omega_0 t$ , 其中  $\omega_0$  为常数, 随机变量  $A$  服从标准高斯分布。求  $t = 0, \pi/3\omega_0, \pi/2\omega_0$  三个时刻  $X(t)$  的一维概率密度。
- 练习 2.2 如图 2-23 所示, 已知随机过程  $X(t)$  仅由四条样本函数组成, 出现的概率分别为

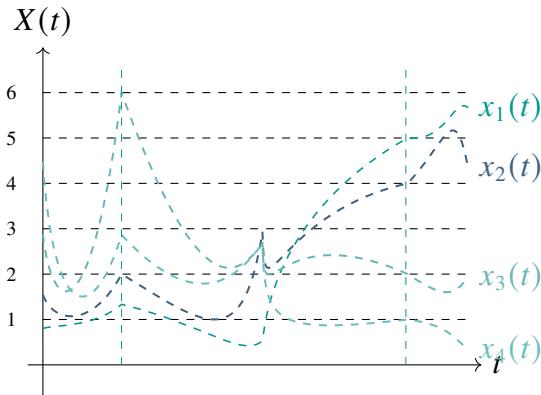


图 2-23 随机过程  $X(t)$  的四条样本函数

在  $t_1$  和  $t_2$  两个时刻的分布律如下:

	$\zeta_1$	$\zeta_2$	$\zeta_3$	$\zeta_4$	
$X(t_1)$	1	2	6	3	
$X(t_2)$	5	4	2	1	
$p_{k_1 k_2}(t_1, t_2)$	1/8	1/4	3/8	1/4	(2.207)

求  $E[X(t_1)], E[X(t_2)], E[X(t_1)X(t_2)]$ 。

- 练习 2.3 已知随机过程  $X(t)$  由  $X(t, \zeta_1) = 1, X(t, \zeta_2) = \sin t$  和  $X(t, \zeta_3) = \cos t$  三条样本函数曲线组成, 三条样本等概率出现。求随机过程的期望  $E[X(t)]$  和自相关函数  $R_X(t_1, t_2)$ 。
- 练习 2.4 已知随机过程  $X(t) = A + Bt$ , 其中  $A, B$  皆为已知的随机变量。
- ① 求随机过程的期望  $E[X(t)]$  和自相关函数  $R_X(t_1, t_2)$ ,
  - ② 若已知随机变量  $A, B$  相互独立, 试用它们的概率密度  $f_A(a)$  和  $f_B(b)$  来表示  $X(t)$  的一维概率密度  $f_X(x, t)$ 。
- 练习 2.5 已知  $X(t)$  为平稳过程随机变量的平稳性。随机变量  $Y = X(t_0), t_0$  为常数。判断随机过程  $Z(t) = X(t) + Y$  是否宽平稳。

**练习 2.6** 已知随机过程  $Y(t) = X(t) \cos(\omega_0 t + \Phi)$ , 其中幅度函数  $X(t)$  是宽平稳随机过程; 角频率  $\omega_0$  为常数; 随机相位  $\Phi$  服从  $(-\pi, \pi)$  的均匀分布, 且与过程  $X(t)$  相互独立。

① 求随机过程  $Y(t)$  的期望和自相关函数。

② 判断随机过程  $Y(t)$  是否宽平稳。

**练习 2.7** 已知随机过程  $X(t)$  由三个样本函数  $X(t, \zeta_1) = 2, X(t, \zeta_2) = 2 \cos t, X(t, \zeta_3) = 3 \sin t$  组成, 每个样本等概率出现。判断该随机过程  $X(t)$  是否宽平稳。

**练习 2.8** 已知平稳过程  $X(t)$  的自相关函数为  $R_X(\tau) = 4e^{-|\tau|} \cos \pi\tau + \cos 3\pi\tau$  求过程  $X(t)$  的均方值和方差。

**练习 2.9** 已知随机过程  $X(t)$  和  $Y(t)$  相互独立且各自平稳, 证明新随机过程  $Z(t) = X(t)Y(t)$  也是平稳的。

**练习 2.10** 已知过程  $X(t) = A \cos t - B \sin t$  和  $Y(t) = B \cos t + A \sin t$ , 其中随机变量  $A, B$  独立, 均值都为 0, 方差都为 5。

① 证明  $X(t)$  和  $Y(t)$  各自平稳且联合平稳。

② 求两个过程的互相关函数。

**练习 2.11** 已知过程  $X(t)$  和  $Y(t)$  联合平稳, 令  $Z(t) = X(t) + Y(t)$ 。试用  $X(t)$  和  $Y(t)$  的自相关函数和互相关函数表示:

①  $Z(t)$  的自相关函数  $R_Z(\tau)$ 。

②  $X(t)$  和  $Y(t)$  独立时的  $R_Z(\tau)$ 。

③  $X(t)$  和  $Y(t)$  独立且均值都为 0 时的  $R_Z(\tau)$ 。

**练习 2.12** 已知随机过程  $X(t)$  和  $Y(t)$  独立且各自平稳, 自相关函数为

$$R_X(\tau) = 2e^{-2|\tau|} \cos \omega_0 \tau, R_Y(\tau) = 9 + \exp(-3\tau^2), \quad (2.208)$$

令随机过程  $Z(t) = AX(t)Y(t)$ , 其中  $A$  是均值为 2, 方差为 9 的随机变量, 且与相互独立。求过程  $Z(t)$  的均值、方差和自相关函数。

**练习 2.13** 已知联合平稳的两个随机过程  $X(t)$  和  $Y(t)$  为

$$\begin{cases} X(t) = a \cos(\omega_0 t + \Phi) \\ Y(t) = b \sin(\omega_0 t + \Phi) \end{cases}, \quad (2.209)$$

其中  $a, b, \omega_0$  皆为常数, 随机相位  $\Phi$  服从  $(0, 2\pi)$  上的均匀分布。求互相关函数  $R_{XY}(\tau), R_{YX}(\tau)$ , 并说明互相关函数在  $\tau = 0$  时的意义。

**练习 2.14** 已知复随机过程

$$Z(t) = \sum_{i=1}^n A_i \exp(j\omega_i t) \quad (2.210)$$

式中  $A_i (i = 1, \dots, n)$  为  $n$  个实随机变量,  $\omega_i (i = 1, \dots, n)$  为  $n$  个实数。求当  $A$  满足什么条件时,  $Z(t)$  复平稳。

**练习 2.15** 已知随机过程  $X(t)$  均方可导, 证明  $X(t)$  和其导数  $Y(t) = X'(t)$  的互相关函数为

$$R_{YX}(t_1, t_2) = \frac{\partial R_X(t_1, t_2)}{\partial t_1}. \quad (2.211)$$

**练习 2.16** 已知平稳过程  $X(t)$  和  $Y(t)$  均方可导, 证明  $X(t), Y(t)$  的互相关函数和  $Y(t)$  的自相关函数分别为

$$R_{XY}(\tau) = \frac{dR_X(\tau)}{d\tau}, R_Y(\tau) = -\frac{d^2R_X(\tau)}{d^2\tau}. \quad (2.212)$$

**练习 2.17** 已知随机过程  $X(t)$  的数学期望  $E[X(t)] = t^2 + 4$ , 求另一个随机过程  $Y(t) = tX'(t) + t^2$  的期望。

**练习 2.18** 已知平稳过程  $X(t)$  的自相关函数  $R_X(\tau) = 2 \exp(-\frac{1}{2}\tau^2)$ , 求:

- ① 其导数  $Y(t) = X'(t)$  的自相关函数和方差。
- ②  $X(t)$  和  $Y(t)$  的方差比。

**练习 2.19** 已知随机过程  $X(t) = V \cos 3t$ , 其中  $V$  是均值和方差皆为 1 的随机变量。令随机过程

$$Y(t) = \frac{1}{t} \int_0^t X(\lambda) d\lambda. \quad (2.213)$$

求  $Y(t)$  的均值、自相关函数、协方差函数和方差。

**练习 2.20** 已知平稳过程  $X(t)$  的自相关函数为 (1)  $R_x(\tau) = 6 \exp\left(-\frac{|\tau|}{2}\right)$ , (2)  $R_x(\tau) = 6 \frac{\sin \pi \tau}{\pi \tau}$  求当  $t$  固定时, 过程  $X(t)$  的四个状态  $X(t), X(t+1), X(t+2), X(t+3)$  的协方差矩阵。

**练习 2.21** 已知平稳高斯过程  $X(t)$  的均值为 0, 令随机过程  $Y(t) = [X(t)]^2$ 。证明

$$R_Y(\tau) = [R_X(0)]^2 + 2[R_X(\tau)]^2. \quad (2.214)$$

**练习 2.22** 已知随机过程  $X(t) = A \cos(\omega_0 t + \Phi)$ , 其中随机相位  $\Phi$  服从  $(0, 2\pi)$  上的均匀分布;  $A$  可能为常数, 也可能为随机变量, 且若  $A$  为随机变量时, 和随机变量  $\Phi$  相互独立。求当  $A$  具备什么条件时, 过程各态历经。



# 3

## 随机信号的频域分析

在信号与系统、信号处理、通信理论以及其他许多领域的理论与实际应用中，广泛应用傅里叶变换这一有效工具，对确定信号在时域上和频域上的状况进行分析。许多情况下，在时域中需要卷积积分运算的问题，放在频域中只需要乘法运算就可以解决，大大减少了运算量。那么是否也能应用傅里叶变换这一工具对随机过程进行频域分析，随机过程是否也有通常意义的“频谱”呢？下面就从理论上详细讨论随机过程的频域特性。

### 3.1 实随机过程的功率谱密度

在研究随机过程的频域特性之前，我们首先对傅里叶变换作一简单回顾。设给定信号  $s(t)$  是时间  $t$  的非周期实函数，其傅里叶变换存在的条件为

1°  $s(t)$  在  $(-\infty, \infty)$  范围内满足狄利克雷条件。

2°  $\int_{-\infty}^{\infty} |s(t)|dt < \infty$  (绝对可积) 的等价条件为  $\int_{-\infty}^{\infty} |s(t)|^2 dt < \infty$  (信号  $s(t)$  的总能量有限)。

若  $s(t)$  满足上述条件 1° 和 2°，则有傅里叶变换对存在：

$$(正变换) 频谱 \quad S(\omega) = \int_{-\infty}^{\infty} s(t)e^{-j\omega t} dt. \quad (3.1)$$

$$(反变换) 信号 \quad s(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega)e^{j\omega t} d\omega. \quad (3.2)$$

或者说信号  $s(t)$  的“频谱”存在。

工程技术中的许多重要的时间函数的总能量是无限的。不能满足傅氏变换的条件，如正弦函数。由于随机过程样本函数的持续时间无限，其总能量也是无限的，不

能满足傅氏变换的条件，随机过程样本函数的“频谱”不存在。所以，随机过程没有通常意义的“频谱”存在。然而人们发现，这类信号尽管总能量是无限的，但它们的平均功率却是有限值。即随机过程样本函数平均功率满足傅里叶变换的条件，随机过程样本函数平均功率的“频谱”——功率谱存在。

### 3.1.1 实随机过程的功率谱密度

#### 1. 实信号的能量谱密度

对于实信号  $s(t)$ ，由定义的傅里叶变换对 (3.1)-(3.2)，可得

$$\begin{aligned} \int_{-\infty}^{\infty} [s(t)]^2 dt &= \int_{-\infty}^{\infty} s(t) \left[ \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) e^{j\omega t} d\omega \right] dt \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) \left[ \int_{-\infty}^{\infty} s(t) e^{j\omega t} dt \right] d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) S^*(\omega) d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} |S(\omega)|^2 d\omega. \end{aligned} \quad (3.3)$$

实信号  $s(t)$  的频谱  $S(\omega)$  一般是  $\omega$  的复函数， $S^*(\omega) = S(-\omega)$  (“\*”表示复共轭)。

对于信号  $s(t)$ ，若  $\int_{-\infty}^{\infty} [s(t)]^2 dt < +\infty$  不一定满足，但  $s(t)$  满足  $\frac{1}{2T} \int_{-T}^T [s(t)]^2 dt < +\infty$ ，这种确定信号称为功率信号。

由上述推导过程得

$$\int_{-\infty}^{\infty} [s(t)]^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |S(\omega)|^2 d\omega. \quad (3.4)$$

式 (3.4) 称为帕塞瓦尔等式。式 (3.4) 右侧等式是  $|S(\omega)|^2$  在整个频域上的积分。因此，则称被积函数  $|S(\omega)|^2$  为  $s(t)$  的能量谱密度。

#### 2. 实随机信号的平均功率

若要将傅里叶变换应用于实随机过程，必须对过程的样本函数做某些限制，最简单的一种方法是应用截取函数，如图 3-1 所示。

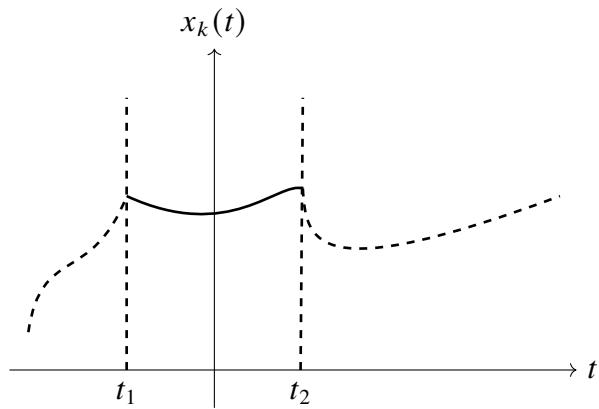


图 3-1 截取函数

在实随机过程  $X(t)$  的样本函数  $X(t, \zeta_k) = x_k(t)$  中, 任意截取  $2T$  长的一段  $x_{kT}(t)$  表示为

$$x_{kT}(t) = \begin{cases} x_k(t), & |t| \leq T \\ 0, & |t| > T \end{cases} \quad (3.5)$$

称  $x_{kT}(t)$  为样本函数  $x_k(t)$  的截取函数。当  $T$  为有限值时, 截取函数  $x_{kT}(t)$  满足绝对可积条件, 其傅里叶变换存在, 即

$$\mathfrak{X}_{kT}(\omega) = \int_{-\infty}^{\infty} x_{kT}(t) e^{-j\omega t} dt = \int_{-T}^{T} x_k(t) e^{-j\omega t} dt. \quad (3.6)$$

$$x_{kT}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X_{kT}(\omega) e^{j\omega t} d\omega. \quad (3.7)$$

可见,  $\mathfrak{X}_{kT}(\omega) = \mathfrak{X}_T(\omega, \zeta_k)$  为  $x_{kT}(t)$  的频谱函数。据帕塞瓦尔等式

$$\begin{aligned} \int_{t_1}^{t_2} f^2(t) dt &= \sum_{n=1}^{\infty} C_n^2 \int_{t_1}^{t_2} g_n^2(t) dt \\ &= \sum_{n=1}^{\infty} \int_{t_1}^{t_2} [C_n g_n(t)]^2 dt \end{aligned} \quad (3.8)$$

 **注 3.1.** 物理意义: 一个信号所含有的能量(功率)恒等于此信号在完备正交函数集中各分量能量(功率)之和。几何意义: 函数可以看作向量, 如果一个函数可以用另一组正交完备函数集表示, 那么这些完备函数集的每一个分量就是一个基矢量。显然, 一个向量的模等于该向量在各个基矢量上投影的平方和。本质上就是说矢量空间信号正交变换的范数不变性。

有

$$\int_{-\infty}^{\infty} x_{kT}^2(t) dt = \int_{-T}^T x_k^2(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\mathfrak{X}_{kT}(\omega)|^2 d\omega. \quad (3.9)$$

若随机过程  $X(t)$  代表一噪声电压(或电流),  $\int_{-T}^T x_k^2(t) dt$  表示噪声的一个样本  $\zeta$  在时间  $(-T, T)$  内消耗在  $1\Omega$  电阻上的总能量。在  $(-T, T)$  上求时间平均的极限为

$$P_k = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x_k^2(t) dt = \lim_{T \rightarrow \infty} \frac{1}{4\pi T} \int_{-\infty}^{\infty} |\mathfrak{X}_{kT}(\omega)|^2 d\omega, \quad (3.10)$$

则  $P$  表示随机过程的样本函数  $x(t)$  消耗在  $1\Omega$  电阻上的平均功率。一般称为随机过程样本函数  $x_k(t)$  的平均(时间平均)功率。

 **注 3.2.** 由于样本  $\zeta$  对应的样本函数  $x_k(t)$  是个确定函数, 所以平均功率  $P_k$  是个值。

$$P_{\Delta}(\zeta) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T X^2(t, \zeta) dt = \lim_{T \rightarrow \infty} \frac{1}{4\pi T} \int_{-\infty}^{\infty} |X_T(\omega, \zeta)|^2 d\omega, \quad (3.11)$$

其中  $X(t, \zeta) = X(t)$  代表一个随机过程,  $X_T(\omega, \zeta) = X_T(\omega)$  代表随机过程截取函数的频谱。若对  $P(\zeta)$  取统计平均,

$$P = E[P_{\Delta}(\zeta)] = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T E[X^2(t)] dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \lim_{T \rightarrow \infty} \frac{1}{2T} E[|X_T(\omega)|^2] d\omega, \quad (3.12)$$

则定义所得的确定值  $P$  为随机过程  $X(t)$  的平均功率。

### 3. 功率谱密度

#### (1) 功率谱密度的定义

由随机过程  $X(t)$  的平均功率的定义式 (3.12), 右端的被积函数记为

$$G_X(\omega) = \lim_{T \rightarrow \infty} \frac{1}{2T} E[|X_T(\omega)|^2], \quad (3.13)$$

则  $G_X(\omega)$  在整个频域上的积分被定义为随机过程的平均功率。则被积函数  $G_X(\omega)$  表示了随机过程  $X(t)$  在不同频率上, 单位频带内消耗在  $1\Omega$  电阻上的平均功率。

同理可得, 样本函数  $x_k(t)$  的功率谱密度为

$$G_k(\omega) = \lim_{T \rightarrow \infty} \frac{1}{2T} |\mathfrak{X}_{kT}(\omega)|^2. \quad (3.14)$$

由于  $\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (\cdot) dt = \overline{(\cdot)}$  表示时间平均, 因此平均功率和功率谱密度的相互关系也可表示为

$$P = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T E[X^2(t)] dt = \overline{E[X^2(t)]} = \frac{1}{2\pi} \int_{-\infty}^{\infty} G_X(\omega) d\omega; \quad (3.15)$$

即随机过程的平均功率，可以通过对过程的均方值  $E[X^2(t)]$  求时间平均来获得。

### (2) 平稳过程的平均功率

若  $X(t)$  为平稳过程，均方值  $E[X^2(t)] = R(0)$  为常数，则平均功率可表示为

$$P = \overline{E[X^2(t)]} = \overline{R(0)} = R(0) = E[X^2(t)]. \quad (3.16)$$

### (3) 各态历经过程的平均功率

由  $\overline{X(t, \zeta)} \stackrel{\text{a.e.}}{=} \overline{x_k(t)}$ ，各态历经过程  $X(t)$  的所有样本函数的时间平均都以概率 1 相同，与  $\zeta \in \Omega$  无关。因此可以由式 (3.10), 式 (3.11) 推出

$$(随机变量) P_\Delta(\zeta) = \overline{X^2(t, \zeta)} \stackrel{\text{a.e.}}{=} \overline{x_k^2(t)} = P_k (\text{常数}) \quad (3.17)$$

而

$$P = E[P_\Delta(\zeta)] \stackrel{\text{a.e.}}{=} E[P_k] = P_k \quad (3.18)$$

即各态历经过程  $X(t)$  的平均功率  $P$  与其样本函数的平均功率  $P_k$  以概率 1 相等。所以，各态历经过程  $X(t)$  的平均功率  $P$  可以由一个样本函数的平均功率  $P_k$  来代替。

### (4) 各态历经过程的功率谱密度

同理，由各态历经性，可以推出

$$G_X(\omega, \zeta) = \lim_{T \rightarrow \infty} \frac{1}{2T} |X_T(\omega, \zeta)|^2 \stackrel{\text{a.e.}}{=} \lim_{T \rightarrow \infty} \frac{1}{2T} |\mathfrak{X}_{kT}(\omega)|^2 = G_k(\omega). \quad (3.19)$$

因此有

$$G_X(\omega) = E[G_X(\omega, \zeta)] \stackrel{\text{a.e.}}{=} E[G_k(\omega)] = G_k(\omega) = \lim_{T \rightarrow \infty} \frac{1}{2T} |\mathfrak{X}_{kT}(\omega)|^2. \quad (3.20)$$

可得结论：

- ① 各态历经过程的平均功率  $P$  与其样本函数的平均功率  $P$  以概率 1 相等。
- ② 各态历经过程的功率谱密度  $G_X(\omega)$  与其样本函数的功率谱密度  $G_k(\omega)$  以概率 1 相等。

综上所述，功率谱密度  $G_X(\omega)$  是从频率角度描述随机过程  $X(t)$  统计规律的最主要的数字特征。

 **注 3.3.** 必须指出， $G_X(\omega)$  仅仅描述了随机过程  $X(t)$  的平均功率按频率分类的情况。

### (5) 实随机过程功率谱密度的性质

功率谱密度是随机过程在频域中主要的统计特征，它具有下列重要性质

1° 功率谱密度非负, 满足

$$G_X(\omega) \geq 0. \quad (3.21)$$

2° 功率谱密度是  $\omega$  的实函数, 满足

$$G_X^*(\omega) = G_X(\omega). \quad (3.22)$$

3° 功率谱密度是  $\omega$  的偶函数, 满足

$$G_X(\omega) = G_X(-\omega). \quad (3.23)$$

证: 根据傅里叶变换的性质, 当截取函数  $x_{iT}(t)$  为  $t$  的实函数时, 其频谱有

$$\mathfrak{X}_{iT}(\omega) = \mathfrak{X}_{iT}^*(-\omega) \Rightarrow X_{iT}^*(\omega) = \mathfrak{X}_{iT}(-\omega). \quad (3.24)$$

所以, 对于过程  $X(t)$  截尾函数的频谱有

$$X_T(\omega) = X_{T^*}(-\omega) \Rightarrow X_{T^*}(\omega) = X_T(-\omega).$$

代入功率谱的定义式, 有

$$\begin{aligned} G_X(\omega) &= \lim_{T \rightarrow \infty} \frac{1}{2T} E [|X_T(\omega)|^2] \\ &= \lim_{T \rightarrow \infty} \frac{1}{2T} E [X_{T^*}(\omega) X_T(\omega)] \\ &= \lim_{T \rightarrow \infty} \frac{1}{2T} E [X_T(-\omega) X_{T^*}^*(-\omega)] = G_X(-\omega). \end{aligned} \quad (3.25)$$

4° 平稳过程的功率谱密度可积, 满足

$$\int_{-\infty}^{\infty} G_X(\omega) d\omega < \infty. \quad (3.26)$$

证: 平稳过程的平均功率

$$P = E [X^2(t)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} G_X(\omega) d\omega. \quad (3.27)$$

由平稳过程的第三个条件可知, 它的均方值有根, 满足  $E [X^2(t)] < \infty$ , 得证。

5° 若平稳过程的功率谱密度可以表示为有理函数形式

$$G_X(\omega) = G_0 \frac{\omega^{2m} + a_{2m-2}\omega^{2m-2} + \cdots + a_0}{\omega^{2n} + b_{2n-2}\omega^{2n-2} + \cdots + b_0}. \quad (3.28)$$

由性质 1 和性质 4。要求上式满足:

(1)  $G_0 > 0$ ;

(2) 有理式的分母无实数根 (即在实轴上无极点), 且  $n > m$ .

**例 3.4** 利用功率谱密度的性质, 判断下列函数, 哪些可能成为平稳过程的功率谱密度。

$$f_1(\omega) = \cos 3\omega, f_2(\omega) = \frac{1}{(\omega - 1)^2 + 2}; f_3(\omega) = \frac{\omega^2 + 1}{\omega^4 + 5\omega^2 + 6}, \quad f_4(\omega) = \frac{\omega^2 + 4}{\omega^4 - 4\omega^2 + 3}.$$

解: 只有  $f_3(\omega)$  可能, 因为  $f_1(\omega) < 0$ ,  $f_2(\omega)$  非偶函数,  $f_4(\omega)$  在实数轴上有极点。

### 3.1.2 实平稳过程的功率谱密度与自相关函数之间的关系

#### 1. 维纳-辛钦定理

众所周知, 确定信号  $s(t)$  与它的频谱  $S(\omega)$  在时域和频域之间构成一对傅里叶变换。可以证明, 平稳随机信号的自相关函数与其功率谱密度之间也构成一对傅里叶变换。下面就来推导这一关系式。

由功率谱密度的推导可知

$$G_X(\omega) = \lim_{T \rightarrow \infty} \frac{1}{2T} E [|X_T(\omega)|^2] = \lim_{T \rightarrow \infty} \frac{1}{2T} E [X_T^*(\omega) X_T(\omega)], \quad (3.29)$$

式中截取函数的频谱

$$X_T(\omega) = \int_{-T}^T X(t) e^{-j\omega t} dt, \quad (3.30)$$

则实过程  $X(t)$  的功率谱密度表示为

$$\begin{aligned} G_X(\omega) &= \lim_{T \rightarrow \infty} \frac{1}{2T} E \left[ \int_{-T}^T X(t_1) e^{j\omega t_1} dt_1 \cdot \int_{-T}^T X(t_2) e^{-j\omega t_2} dt_2 \right] \\ &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \int_{-T}^T E [X(t_1) X(t_2)] e^{-j\omega(t_2-t_1)} dt_1 dt_2 \\ &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \int_{-T}^T R_X(t_1, t_2) e^{-j\omega(t_2-t_1)} dt_1 dt_2. \end{aligned} \quad (3.31)$$

式中  $R_X(t_1, t_2)$  只在  $-T \leq t_1, t_2 \leq T$  存在, 令  $t = t_1, \tau = t_2 - t_1 = t_2 - t$ , 代入上式进行变量置换, 如图 3-2 所示, 可得

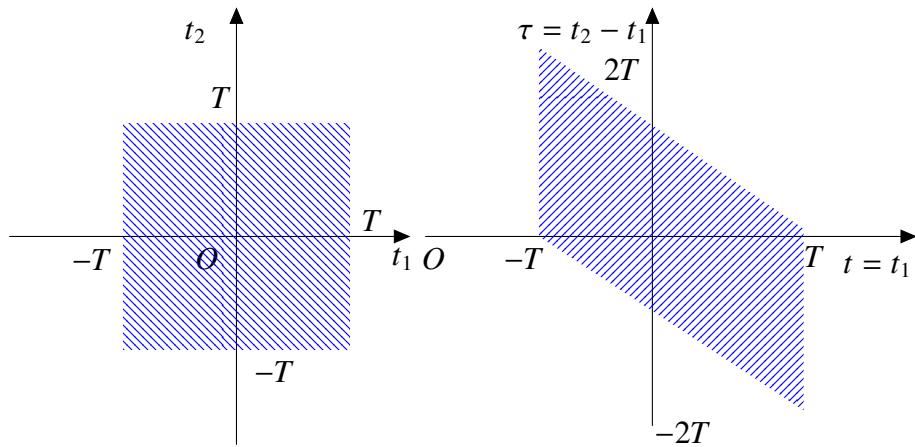


图 3-2 维纳-辛钦定理推导过程中的变量置换

$$\begin{aligned}
 G_X(\omega) &= \lim_{r \rightarrow \infty} \frac{1}{2T} \left\{ \int_{-2T}^0 \left[ \int_{-T-\tau}^T R_X(t, t+\tau) dt \right] e^{-j\omega\tau} d\tau \right. \\
 &\quad \left. + \int_0^{2T} \left[ \int_{-T}^{T-\tau} R_X(t, t+\tau) dt \right] e^{-j\omega\tau} d\tau \right\} \\
 &= \int_{-\infty}^0 \left[ \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T-\tau}^T R_X(t, t+\tau) dt \right] e^{-j\omega\tau} d\tau \\
 &\quad + \int_0^{\infty} \left[ \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{T-\tau} R_X(t, t+\tau) dt \right] e^{-j\omega\tau} d\tau \\
 &= \int_{-\infty}^0 \overline{R_X(t, t+\tau)} e^{-j\omega\tau} d\tau + \int_0^{\infty} \overline{R_X(t, t+\tau)} e^{-j\omega\tau} d\tau \\
 &= \int_{-\infty}^{\infty} \overline{R_X(t, t+\tau)} e^{-j\omega\tau} d\tau \\
 &= F[\overline{R_X(t, t+\tau)}]. \tag{3.32}
 \end{aligned}$$

由傅里叶变换的定义，上式成立的条件是  $\overline{R_X(t, t+\tau)}$  绝对可积。根据傅里叶变换的唯一性，必有

$$\overline{R_X(t, t+\tau)} = F^{-1}[G_X(\omega)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} G_X(\omega) e^{j\omega\tau} d\omega. \tag{3.33}$$

所以，任意随机过程  $X(t)$  自相关函数的时间平均与其功率谱密度互为傅里叶变换，即有

$$\overline{R_X(t, t+\tau)} \xrightleftharpoons[F^{-1}]{F} G_X(\omega). \tag{3.34}$$

若随机过程  $X(t)$  是平稳过程，其自相关函数  $R_X(t_1, t_2) = R_X(t, t + \tau) = R_X(\tau)$  与  $t$  无关。则有

$$\overline{R}_X(t, t + \tau) = \overline{R_X(\tau)} = R_X(\tau). \quad (3.35)$$

只要  $R_X(\tau)$  绝对可积，满足

$$\int_{-\infty}^{\infty} |R_X(\tau)| d\tau < \infty. \quad (3.36)$$

则

$$G_X(\omega) = \int_{-\infty}^{\infty} R_X(\tau) e^{-j\omega\tau} d\tau. \quad (3.37)$$

同理

$$R_X(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G_X(\omega) e^{j\omega\tau} d\omega, \quad (3.38)$$

所以平稳过程的自相关函数与其功率谱密度之间是一对傅里叶变换关系，即

$$R_X(\tau) \xrightleftharpoons[F^{-1}]{F} G_X(\omega). \quad (3.39)$$

这一关系就是著名的维纳-辛钦定理，或称为维纳-辛钦公式。

### 定理 3.5 维纳-辛钦公式

于平稳随机过程  $X(t)$ ，过程的自相关函数记为  $R_X(\tau)$ ，其功率谱密度记为  $G_X(\omega)$ ，下式成立

$$R_X(\tau) \xrightleftharpoons[F^{-1}]{F} G_X(\omega). \quad (3.40)$$

即，平稳过程的自相关函数与其功率谱密度之间是一对傅里叶变换。

它给出了平稳过程的时域特性和频域特性之间的联系。可以说，它是分析随机过程的一个最重要、最基本的公式。

利用平稳过程的自相关函数和功率谱密度皆为偶函数，维纳-辛钦定理可表示为

$$\begin{cases} G_X(\omega) = 2 \int_0^\infty R_X(\tau) \cos \omega\tau d\tau \\ R_X(\tau) = \frac{1}{\pi} \int_0^\infty G_X(\omega) \cos \omega\tau d\omega \end{cases}. \quad (3.41)$$

只要式 (3.41) 右侧的积分有意义，它就是功率谱，否则功率谱不存在。

$$G(\omega) = \int_{-\infty}^{\infty} \left[ \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T R_X(t + \tau, t) dt \right] e^{-j\omega\tau} d\tau. \quad (3.42)$$

或者说，定理成立的条件就是功率谱密度、自相关函数有意义。更进一步，还可以使条件更弱：可以去掉“平稳”。将  $R_X(\tau)$  换成  $R_X(t_1, t_2) = R_X(t + \tau, t)$  之后，继续上述证明过程，最后可以得到。

**证明。(另一种方法)** 维纳-辛钦定理推导中的变形置换的证明：按定义：

$$G(\omega) = E \left\{ \lim_{T \rightarrow \infty} \frac{1}{2T} |X_T(\omega)|^2 \right\}. \quad (3.43)$$

其中

$$X_T(\omega) = \int_{-T}^T X(t) e^{-j\omega t} dt = \int_{-\infty}^{\infty} \text{rect}\left(\frac{t}{2T}\right) X(t) e^{-j\omega t} dt, \quad (3.44)$$

式中

$$\frac{1}{2T} \int_{-T}^T \text{rect}\left(\frac{t+\tau}{2T}\right) dt = \begin{cases} 1 - \frac{|\tau|}{2T}, & |\tau| < 2T \\ 0, & |\tau| \geq 2T \end{cases} \quad (3.45)$$

将 (3.44) 代入 (3.45)，对所有  $\tau \in (-\infty, \infty)$ ，有

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \text{rect}\left(\frac{t+\tau}{2T}\right) dt = 1. \quad (3.46)$$

将 (3.44)-(3.46) 代入 (3.43)：

$$G_X(\omega) = \int_{-\infty}^{\infty} R_X(\tau) e^{-j\omega\tau} d\tau. \quad (3.47)$$

这样一来，维纳-辛钦定理不需要额外的条件。

实际上，上述对功率谱密度的定义适用于任意随机过程（定义本身不要求必须是平稳过程），对上面的式 (3.41) 改变时间平均和傅氏变换运算的次序，就是时变功率谱的时间平均。□

## 2. 维纳-辛钦定理的推广

应该指出，以上我们讨论的维纳-辛钦定理是在随机过程的  $R_X(\tau)$  满足绝对可积的条件下推出的。它要求随机过程的均值为零，且  $R_X(\tau)$  中不能含有周期分量。实际中含有直流分量和周期分量的随机过程很多，绝对可积的条件限制了定理的应用。

通过借助  $\delta$  函数，就可以不受此条件的限制。即将直流分量与周期分量在各个频率点上的无限值用一个  $\delta$  函数来表示，借助  $\delta$  函数的傅里叶变换，则维纳-辛钦公式就可以推广到含有直流或周期成分的平稳过程中来。先介绍一些常见的结论；

1)  $\delta$  函数的时域和频域傅里叶变换对为

$$\begin{cases} \delta(\tau) \Leftrightarrow 1 \\ \frac{1}{2\pi} \Leftrightarrow \delta(\omega) \end{cases} \quad (3.48)$$

2) 周期函数的傅里叶变换对为

$$\begin{cases} \cos(\omega_0\tau) \Leftrightarrow \pi [\delta(\omega - \omega_0) + \delta(\omega + \omega_0)] \\ \sin(\omega_0\tau) \Leftrightarrow j\pi [\delta(\omega + \omega_0) - \delta(\omega - \omega_0)] \end{cases} \quad (3.49)$$

3)  $\delta$  函数与连续函数  $s(t)$  的乘积公式为

$$\begin{cases} s(t)\delta(t-\tau) = s(\tau)\delta(t-\tau) \\ s(t)\delta(t) = s(0)\delta(t) \end{cases} \quad (3.50)$$

**例 3.6** 已知一个电报信号是平稳随机过程, 其自相关函数  $R_X(\tau) = Ae^{-\beta|\tau|}$ ,  $A > 0, \beta > 0$ , 如图 3-3(a) 所示。求该电报信号的功率谱密度。

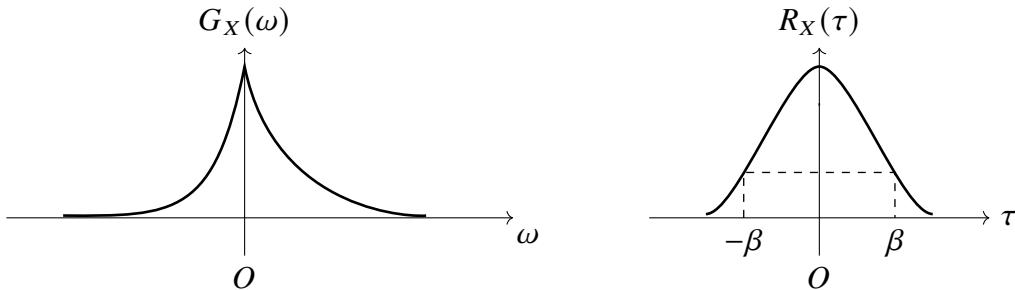


图 3-3 a) 平稳过程的自相关函数 b) 相应的功率谱密度

解: 因为自相关函数  $R_X(\tau)$  的表示式中包含有  $|\tau|$  项, 因此在应用维纳-辛钦公式求积分时, 应将  $|\tau|$  项分成  $+\tau$  和  $-\tau$  两部分进行积分。

$$\begin{aligned} G_X(\omega) &= \int_{-\infty}^0 Ae^{\beta\tau} e^{-j\omega\tau} d\tau + \int_0^\infty Ae^{-\beta\tau} e^{-j\omega\tau} d\tau = A \left[ \frac{e^{(\beta-j\omega)\tau}}{(\beta-j\omega)} \right]_0^\infty + A \left[ \frac{e^{-(\beta+j\omega)\tau}}{-(\beta+j\omega)} \right]_0^\infty \\ &= A \left[ \frac{1}{\beta-j\omega} + \frac{1}{\beta+j\omega} \right] = \frac{2A\beta}{\beta^2 + \omega^2}. \end{aligned} \quad (3.51)$$

计算出来的功率谱密度  $G_X(\omega)$  如图 3-3(b) 所示。

**例 3.7** 已知随机相位过程  $X(t) = A \cos(\omega_0 t + \theta)$ , 其中  $A, \omega_0$  为实常数,  $\theta$  为随机相位, 服从  $(0, 2\pi)$  上的均匀分布。可证其为平稳过程, 且自相关函数为  $R_X(\tau) = \frac{A^2}{2} \cos(\omega_0 \tau) = \frac{A^2}{4} (e^{j\omega_0 \tau} + e^{-j\omega_0 \tau})$ , 求  $X(t)$  的功率谱密度  $G_X(\omega)$ 。

解：自相关函数  $R_X(\tau)$  含有周期分量，引入  $\delta$  函数，可得

$$\begin{aligned} G_X(\omega) &= \int_{-\infty}^{\infty} R_X(\tau) e^{-j\omega\tau} d\tau \\ &= \frac{A^2}{4} \int_{-\infty}^{\infty} (e^{j\omega_0\tau} + e^{-j\omega_0\tau}) e^{-j\omega\tau} d\tau \\ &= \frac{A^2\pi}{2} [\delta(\omega - \omega_0) + \delta(\omega + \omega_0)], \end{aligned} \quad (3.52)$$

表示随机过程  $X(t)$  的功率谱密度为在  $\pm\omega_0$  处的  $\delta$  函数，功率集中在  $\pm\omega_0$  处，如图 3-4 所示。

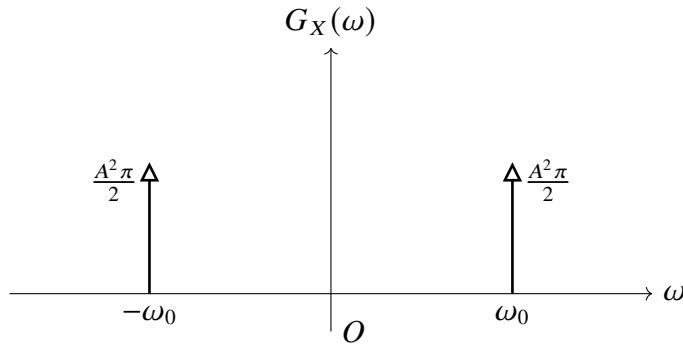


图 3-4

**例 3.8** 已知平稳过程  $X(t)$ ，具有功率谱密度为

$$G_X(\omega) = \frac{16}{\omega^4 + 13\omega^2 + 36},$$

求该平稳过程  $X(t)$  的自相关函数和均方值。

解：由例 3.7 知

$$R_X(\tau) = Ae^{-\beta|\tau|} \Leftrightarrow G_X(\omega) = \frac{2A\beta}{\beta^2 + \omega^2}. \quad (3.53)$$

为了利用这个傅里叶变换关系，可以将  $G_X(\omega)$  用部分分式法展开，有

$$G_X(\omega) = \frac{16}{\omega^4 + 13\omega^2 + 36} = \frac{16}{(\omega^2 + 4)(\omega^2 + 9)} = \frac{16/5}{\omega^2 + 4} - \frac{16/5}{\omega^2 + 9}. \quad (3.54)$$

于是，自相关函数  $R_X(\tau)$  应当具有如下形式：

$$R_X(\tau) = F^{-1}[G_X(\omega)] = F^{-1}\left[\frac{16/5}{\omega^2 + 4}\right] - F^{-1}\left[\frac{16/5}{\omega^2 + 9}\right]. \quad (3.55)$$

由于

$$\frac{16/5}{\omega^2 + 4} = \frac{2 \times 2 \times 4/5}{\omega^2 + 4}, \quad \frac{16/5}{\omega^2 + 9} = \frac{2 \times 3 \times 8/15}{\omega^2 + 9}. \quad (3.56)$$

故

$$A_1 = \frac{4}{5}, \beta_1 = 2; \quad A_2 = \frac{8}{15}, \beta_2 = 3, \quad (3.57)$$

可得

$$R_X(\tau) = \frac{4}{5}e^{-2|\tau|} - \frac{8}{15}e^{-3|\tau|}. \quad (3.58)$$

### 3. 物理功率谱密度

前面定义的随机过程功率谱密度分布为“双边谱”密度。实际应用中  $\omega < 0$ , 负频率外不存在, 公式中的负频率纯粹只有数学上的意义和运算的方便。因此有时也采用另一称“单边谱”密度, 如图 3-5 所示。

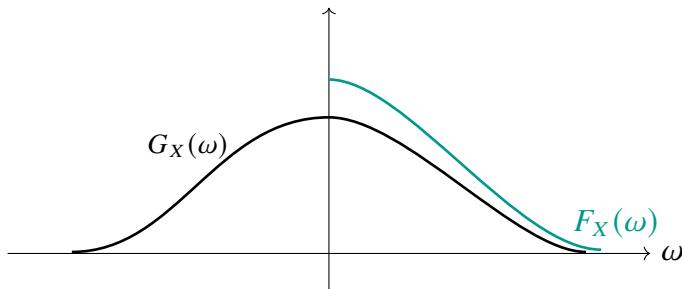


图 3-5 物理功率谱密度

$F_X(\omega)$  与  $G_X(\omega)$  的关系如下:

$$F_X(\omega) = 2G_X(\omega)U(\omega) = \begin{cases} 2G_X(\omega), & \omega \geq 0 \\ 0, & \omega < 0 \end{cases}. \quad (3.59)$$

其中阶跃函数

$$U(\omega) = \begin{cases} 1, & \omega \geq 0 \\ 0, & \omega < 0 \end{cases}. \quad (3.60)$$

若用物理功率谱密度  $F_X(\omega)$  和  $P$  表示平稳随机过程的自相关函数及平均功率, 则为

$$R_X(\tau) = \frac{1}{2\pi} \int_0^\infty F_X(\omega) \cos \omega \tau d\omega. \quad (3.61)$$

$$P = R_X(0) = \frac{1}{2\pi} \int_0^\infty F_X(\omega) d\omega. \quad (3.62)$$

本节中讨论的功率谱密度, 若不加说明, 皆指“双边谱”密度(功率谱密度)。

## 3.2 两个实随机过程的互功率谱密度

在上章中已经建立了两个实随机过程的互相关函数的概念。下面将单个实随机过程的功率谱密度的概念, 以及相应的分析方法推广到两个随机过程中去。

### 1. 互功率谱密度

#### (1) 其平均功率的定义

若  $X(t)$  和  $Y(t)$  的样本函数  $x_k(t)$  和  $y_k(t)$  的两个截取函数  $x_{kT}(t), y_{kT}(t)$  可分别表示为

$$x_{kT}(t) = \begin{cases} x_k(t), & -T < t < T \\ 0, & \text{其他} \end{cases}, \quad y_{kT}(t) = \begin{cases} y_k(t), & -T < t < T \\ 0, & \text{其他} \end{cases} \quad (3.63)$$

因为截取函数  $x_{kT}(t), y_{kT}(t)$  都满足绝对可积的条件, 所以它们的傅氏变换存在, 于是有

$$x_{kT}(t) \xrightarrow[F]{F^{-1}} \mathfrak{X}_{kT}(\omega), \quad y_{kT}(t) \xrightarrow[F]{F^{-1}} \mathcal{Y}_{kT}(\omega). \quad (3.64)$$

由于  $x_{kT}(t), y_{kT}(t)$  的傅氏变换存在, 故帕塞瓦尔定理对它们也适用, 即

$$\int_{-\infty}^{\infty} x_{kT}^*(t) y_{kT}(t) dt = \int_{-T}^T x_k^*(t) y_k(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathfrak{X}_{kT}(\omega) \mathcal{Y}_{kT}(\omega) d\omega. \quad (3.65)$$

由于  $X(t), Y(t)$  为实过程, 所以  $x_{K^*T}(t) = x_{kT}(t), y_{K^*T}(t) = y_{kT}(t)$ . 因此可得两个随机过程的样本函数  $x_k(t)$  和  $y_k(t)$  的互平均功率

$$P_k = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x_k(t) y_k(t) dt = \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{2T} \mathfrak{X}_{kT}(\omega) \mathcal{Y}_{kT}(\omega) d\omega. \quad (3.66)$$

由于  $x_{kT}(t), y_{kT}(t)$  和  $\mathfrak{X}_{kT}(\omega), \mathcal{Y}_{kT}(\omega)$  都是试验结果  $\zeta_k$  的函数, 相对于所有试验结果上的平均功率  $\{P_k\} = P_\Delta(\zeta)$  是一个随机变量. 平均后的确定值  $P_{XY}$  为  $X(t), Y(t)$  两个随机过程的互平均功率, 即

$$\begin{aligned} P_{XY} &= E [P_\Delta(\zeta)] = E \left[ \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T X(t) Y(t) dt \right] \\ &= E \left[ \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{2T} X_{T^*}(\omega) Y_T(\omega) d\omega \right]. \end{aligned} \quad (3.67)$$

交换期望与极限的次序有

$$\begin{aligned} P_{XY} &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T E[X(t)Y(t)] dt \\ &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T R_{XY}(t, t) dt \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \lim_{T \rightarrow \infty} \frac{1}{2T} E[X_T^*(\omega)Y_T(\omega)] d\omega. \end{aligned} \quad (3.68)$$

## (2) 互功率谱密度的定义与关系

$$G_{XY}(\omega) = \lim_{T \rightarrow \infty} \frac{1}{2T} F[X_T^*(\omega)Y_T(\omega)], \quad (3.69)$$

则互平均功率为

$$P_{XY} = \frac{1}{2\pi} \int_{-\infty}^{\infty} G_{XY}(\omega) d\omega. \quad (3.70)$$

同理可得,  $X(t), Y(t)$  的另一个互功率谱密度为

$$G_{YX}(\omega) = \lim_{T \rightarrow \infty} \frac{1}{2T} E[Y_T^*(\omega)X_T(\omega)]. \quad (3.71)$$

$X(t), Y(t)$  的另一个互平均功率为

$$P_{YX} = \frac{1}{2\pi} \int_{-\infty}^{\infty} G_{YX}(\omega) d\omega. \quad (3.72)$$

比较可得两个互平均谱密度关系为

$$G_{XY}(\omega) = G_{YX}^*(\omega). \quad (3.73)$$

## 2. 互谱密度与互相关函数的关系

如同单个实平稳过程自相关函数与其功率谱密度之间的关系一样, 两个实平稳过程互相关系数与互谱密度之间也存在着类似的关系。对于两个实随机过程  $X(t), Y(t)$ , 其互谱密度  $G_{XY}(\omega)$  与其互相关函数  $R_{XY}(t, t + \tau)$  之间的关系为

$$G_{XY}(\omega) = \int_{-\infty}^{\infty} \overline{R_{XY}(t, t + \tau)} e^{-j\omega\tau} d\tau. \quad (3.74)$$

即

$$\overline{R_{XY}(t, t + \tau)} \xrightleftharpoons[F^{-1}]{F} G_{XY}(\omega). \quad (3.75)$$

若  $X(t), Y(t)$  联合平稳, 则有

$$R_{XY}(\tau) \xrightarrow[F]{F^{-1}} G_{XY}(\omega). \quad (3.76)$$

即两个联合平稳的实随机过程, 它们的互谱密度与互相关函数为一傅里叶变换对, 表示如下

$$\begin{cases} G_{XY}(\omega) &= \int_{-\infty}^{\infty} R_{XY}(\tau) e^{-j\omega\tau} d\tau \\ R_{XY}(\tau) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} G_{XY}(\omega) e^{j\omega\tau} d\omega \end{cases}. \quad (3.77)$$

### 3. 互谱密度的性质

两个随机过程的互功率谱密度与单个随机过程的功率谱密度不同, 它不再是频率  $\omega$  的非负、实的、偶函数。下面列举互功率谱密度的一些性质。

1) 互谱密度非偶函数, 满足

$$G_{XY}(\omega) = G_{Y^*X}(\omega) = G_{YX}(-\omega). \quad (3.78)$$

2) 互谱密度的实部为  $\omega$  的偶函数, 即

$$\begin{cases} \operatorname{Re}[G_{XY}(\omega)] = \operatorname{Re}[G_{XY}(-\omega)] \\ \operatorname{Re}[G_{YX}(\omega)] = \operatorname{Re}[G_{YX}(-\omega)] \end{cases}. \quad (3.79)$$

3) 互谱密度的虚部为  $\omega$  的奇函数, 即

$$\begin{cases} \operatorname{Im}[G_{XX}(\omega)] = -\operatorname{Im}[G_{XY}(-\omega)] \\ \operatorname{Im}[G_{YX}(\omega)] = -\operatorname{Im}[G_{YX}(-\omega)] \end{cases}. \quad (3.80)$$

式中  $\operatorname{Im}[\cdot]$  表示虚部。

4) 若  $X(t), Y(t)$  正交, 则有

$$G_{XY}(\omega) = G_{YX}(\omega) = 0. \quad (3.81)$$

5) 若  $X(t), Y(t)$  不相关, 且分别具有常数均值  $m_X$  和  $m_Y$ , 则

$$\begin{cases} R_{XY}(t, t + \tau) = m_X m_Y \\ G_{XY}(\omega) = G_{YX}(\omega) = 2\pi m_X m_Y \delta(\omega) \end{cases}. \quad (3.82)$$

6) 互相关函数和互谱密度满足

$$\begin{cases} \overline{R_{XY}(t, t + \tau)} \Leftrightarrow G_{XY}(\omega), \\ \overline{R_{YX}(t, t + \tau)} \Leftrightarrow G_{YX}(\omega). \end{cases} \quad (3.83)$$

**例 3.9** 设两个随机过程  $X(t), Y(t)$  联合平稳, 其互相关函数

$$R_{XY}(\tau) = \begin{cases} 9e^{-3\tau}, & \tau \geq 0 \\ 0, & \tau < 0 \end{cases} \quad (3.84)$$

求互谱密度  $G_{XY}(\omega)$  和  $G_{YX}(\omega)$ 。

由联合平稳过程互相关函数和互谱密度的傅立叶变换对关系, 可得

$$\begin{aligned} G_{XY}(\omega) &= \int_{-\infty}^{\infty} R_{XY}(\tau) e^{-j\omega\tau} d\tau = \int_0^{\infty} 9e^{-3\tau} e^{-j\omega\tau} d\tau \\ &= 9 \int_0^{\infty} e^{-(3+j\omega)\tau} d\tau = \frac{9}{3+j\omega}. \end{aligned} \quad (3.85)$$

可见,  $G_{XY}(\omega)$  是  $\omega$  的复函数。根据互谱密度的性质 1), 可得

$$G_{YX}(\omega) = G_{XY^*}(\omega) = \frac{9}{3-j\omega}. \quad (3.86)$$

### 3.3 理想白噪声

#### 定义 3.10 白噪声的定义

若平稳过程  $N(t)$  的均值为零, 功率谱密度在整个频率轴  $(-\infty, +\infty)$  上服从均匀分布, 满足

$$G_N(\omega) = \frac{1}{2}N_0, \quad (3.87)$$

其中  $N_0$  为正实常数. 则称此过程为白噪声过程, 简称白噪声。

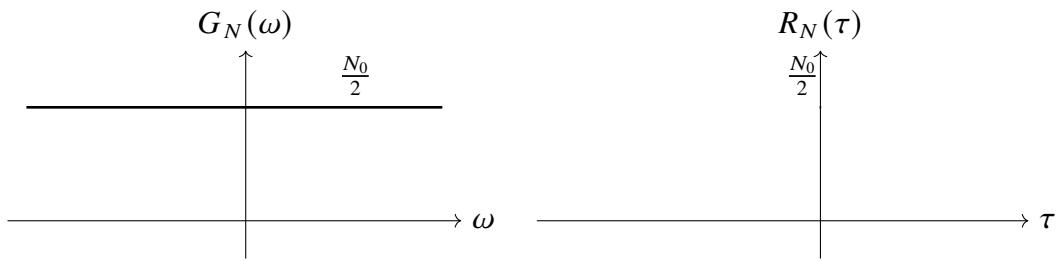
“白”是借用了光学中“白光”这一术语。因为白光的光谱包含了所有可见光的频率分量, 分布在整个频率轴上。任意的非白噪声被定义为色噪声。如图 3-6(b) 所示, 例 3.4 所表示的随机过程就是色噪声的一例。

#### (2) 白噪声的自相关函数

利用维纳-辛钦定理, 不难得到白噪声  $N(t)$  的自相关函数为

$$R_N(\tau) = \frac{1}{2}N_0\delta(\tau). \quad (3.88)$$

上式说明, 白噪声的自相关函数是一个面积等于功率谱密度的  $\delta$  函数。白噪声的功率谱密度和自相关函数的图形如图 3-6 所示。

图 3-6 理想白噪声 (功率谱密度  $G_N(\omega)$  和自相关函数  $R_N(\tau)$ )

### (3) 白噪声的自相关系数

$$\rho_N(\tau) = \frac{R_N(\tau)}{R_N(0)} = \begin{cases} 1, & \tau = 0 \\ 0, & \tau \neq 0 \end{cases}. \quad (3.89)$$

### (4) 白噪声的特点

#### 1) 理想化的数学模型

(1) 由白噪声的自相关系数可见, 白噪声在任何两个相邻时刻的状态 (即使是紧连接的两个时刻), 只要不是同一时刻都是不相关的。因此, 在时域中白噪声的样本函数变化极快。然而任何实际的过程, 无论样本函数变化多快, 紧连卷的两个时刻的状态总存在一定的关联性, 自相关函数不可能是一个  $\delta$  函数。

(2) 由于定义下的白噪声模型的功率谱无限宽, 因此其平均功率就无限大。然而, 物理上存在的任何随机过程, 其平均功率总是有限的。因此, 在这样定义下的白噪声只是一种理想化的数学模型, 在物理上是不存在的。尽管如此, 由于白噪声在数学上具有处理简单, 方便的优点, 所以它在随机过程的理论研究及实际应用中仍占有特别重要的地位。

#### 2) 数学上有很好的运算性质

白噪声的功率谱密度是“常数”, 自相关函数是一个“冲激函数”, 所以将它作为噪声与信号一起分析处理, 运算起来非常方便。

#### 3) 理想白噪声是大多数重要噪声的模型

经过科学家的验证, 大自然中许多重要的噪声过程, 因功率谱近似于常数, 确实可以用白噪声来近似。例如, 对通信系统有很大影响的热噪声, 是由元件中电子原件的热运动产生的。约翰逊和奈奎斯特从实验和理论两个方面研究证明: 在正常室温下, 网络工作频率低于  $10^3\text{Hz}$  时, 阻值为  $R$  的电阻两端噪声电压  $N_V$  的均值为零均方值 (平均功率)

$$E [N_V^2(t)] = 4kT R \Delta f, \quad (3.90)$$

式中  $T$  为绝对零度,  $k = 1.38 \times 10^{23} \text{ J/K}$  为波耳兹曼常数,  $\Delta f$  为噪声带宽。

其功率谱密度为

$$G_{\text{NV}}(\omega) = \frac{E[N_V^2(t)]}{2\Delta f} = 2kTR. \quad (3.91)$$

可见, 热噪声具有平均的功率谱密度, 可以用白噪声来近似。

#### 4) 白噪声可以替代实际应用中的宽带噪声

在实际工作中, 任何一个系统的带宽总是有限的。当噪声通过某一系统时, 只要它在比我们感兴趣的信号频带宽得多的范围内, 都具有近似均匀的功率谱密度, 如图 3-7 所示, 这个噪声就可以被当作白噪声来处理, 而且不会带来很大的误差。

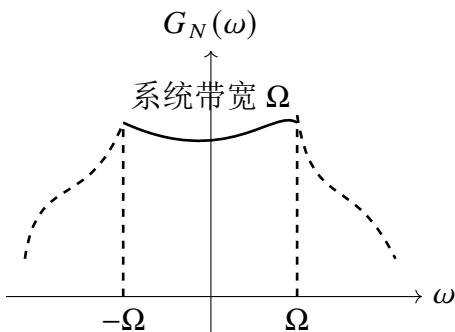


图 3-7 白噪声的特点

因此, 电子设备中出现的各种起伏过程, 大多数都可认为是白噪声。如电阻热噪声、晶体管的散弹噪声等, 在相当宽的频率范围内都具有均匀的功率谱密度, 所以可以把它看成是白噪声。

#### 5) 高斯白噪声

高斯白噪声在任意两相邻时刻的状态有各态历经性。

#### 2. 限带白噪声

##### 定义 3.11 限带白噪声

平稳随机过程  $X(t)$  均值为零, 功率谱密度在有限频率范围内是均匀分布, 在此范围外为零, 则称此过程为限带白噪声。

限带白噪声是另外一个常用的概念。

#### (2) 限带白噪声的分类

**性质** 限带白噪声分为低通型限带白噪声和带通型限带白噪声。

## 1) 低通型限带白噪声

若随机过程  $X(t)$  的功率谱密度满足

$$G_X(\omega) = \begin{cases} G_0, & |\omega| \leq \Omega/2 \\ 0, & |\omega| > \Omega/2 \end{cases} \quad (3.92)$$

则称此过程为低通型限带白噪声。

将白噪声通过一个理想低通滤波器，便产生出低通型限带白噪声。其自相关系数为

$$\begin{aligned} R_X(\tau) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} G_X(\omega) e^{j\omega\tau} d\omega \\ &= \frac{1}{2\pi} \int_{-\Omega/2}^{\Omega/2} G_0 e^{j\omega\tau} d\omega = \frac{\Omega G_0}{2\pi} \cdot \frac{\sin(\Omega\tau/2)}{\Omega\tau/2}. \end{aligned} \quad (3.93)$$

低通型限带白噪声的  $G_X(\omega)$  和  $R_X(\tau)$  的图形如图 3-8 所示。可以看出时间间隔  $\tau$  等于  $2\pi/\Omega$  整数倍的那些随机变量，彼此是不相关的。

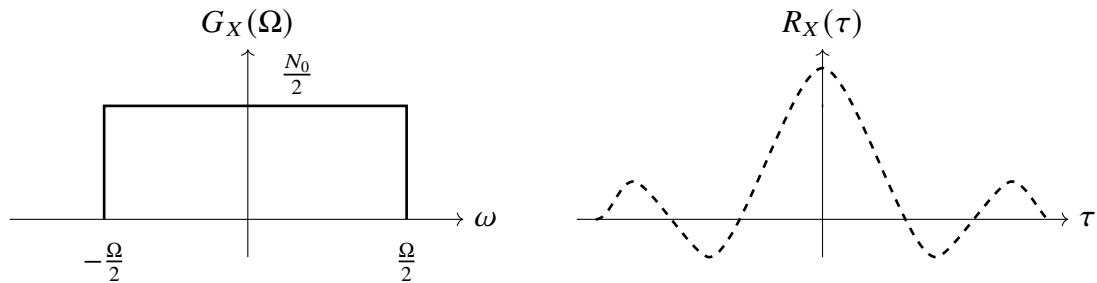


图 3-8 低通型限带白噪声 ( $G_X(\omega)$  和  $R_X(\tau)$ )

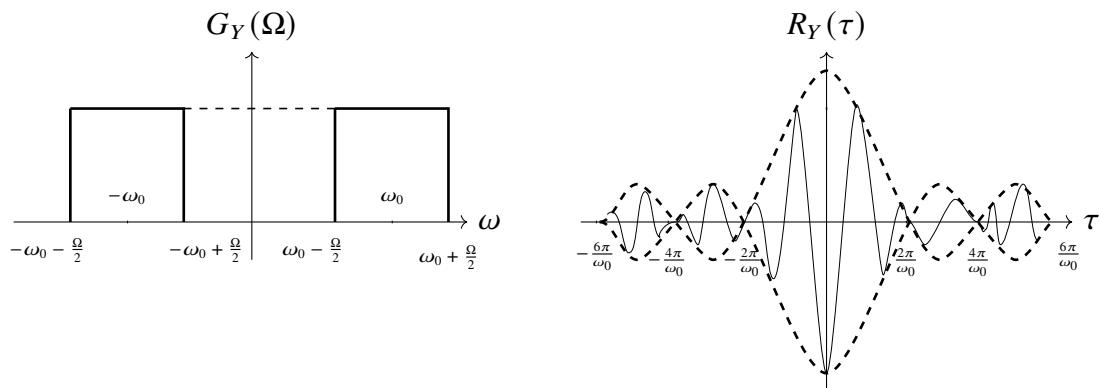


图 3-9 带通型限带白噪声 ( $G_Y(\Omega)$  和  $R_Y(\tau)$ )

## 2) 带通型限带白噪声

类似低通型限带白噪声, 带通型限带白噪声的功率谱密度为

$$G_Y(\omega) = \begin{cases} G_0, & \omega_0 - \Omega/2 < |\omega| < \omega_0 + \Omega/2 \\ 0, & \text{其他} \end{cases} \quad (3.94)$$

应用维纳——辛钦定理, 不难导出它的自相关函数为

$$R_Y(\tau) = \frac{\Omega G_0}{\pi} \cdot \frac{\sin(\Omega\tau/2)}{\Omega\tau/2} \cos \omega_0\tau = 2R_X(\tau) \cos \omega_0\tau. \quad (3.95)$$

带通型限带白噪声  $G_Y(\omega)$  和  $R_Y(\tau)$  的图形由图 3-9 给出。不难看出, 将白噪声通过一个理想带通滤波器便可产生带通型限带白噪声。

 **注** 带通型限带白  
噪声的自相关函  
数计算



## 3.4 习题

✎ **练习 3.1** 已知平稳过程  $X(t)$  的功率谱密度为  $G_X(\omega) = 32/(\omega^2 + 16)$ , 求: ① 该过程的平均功率。②  $\omega$  取值在  $(-4, 4)$  范围内的平均功率。

✎ **练习 3.2**  已知平稳过程的功率谱密度为

$$G_X(\omega) = \frac{\omega^2}{\omega^4 + 3\omega^2 + 2}. \quad (3.96)$$

求此过程的均方值和方差。

✎ **练习 3.3**  已知平稳过程  $X(t)$  的自相关函数如下, 求它们的功率谱密度  $G_X(\omega)$ , 并画出图形。(1)  $R_X(\tau) = e^{-3|\tau|}$ .

$$(2) R_X(\tau) = e^{-|\tau|} \cos \pi\tau.$$

$$(3) R_X(\tau) = 5 \exp\left[-\frac{\tau^2}{8}\right].$$

$$(4) R_X(\tau) = \begin{cases} 1 - |\tau|, & |\tau| \leq 1 \\ 0, & \text{其他} \end{cases}.$$

✎ **练习 3.4** 已知平稳过程  $X(t)$  的自相关函数如下, 求功率谱密度  $G_X(\omega)$ 。

$$\textcircled{1} R_X(\tau) = 4e^{-|\tau|} \cos \pi\tau + \cos 3\pi\tau.$$

$$\textcircled{2} R_X(\tau) = 16e^{-2|\tau|} - 8e^{-4|\tau|}.$$

练习 3.5 已知平稳过程  $X(t)$  在频率  $=0$  时的功率谱密度为零, 证明  $X(t)$  的自相关函数满足

$$\int_{-\infty}^{\infty} R_X(\tau) d\tau = 0. \quad (3.97)$$

练习 3.6 已知平稳过程  $X(t)$  的自相关函数为

$$R_X(\tau) = a \cos^4 \omega_0 \tau. \quad (3.98)$$

其中  $a, \omega_0$  皆为正常数, 求  $X(t)$  的功率谱密度和平均功率。

练习 3.7 如图 3-10 所示, 线性系统的输入  $X(t)$  为平稳过程, 系统的输出为平稳过程  $Y(t) = X(t) - X(t - T)$ 。证明: 系统输出的平稳过程  $Y(t)$  的功率谱密度为  $G_Y(\omega) = 2G_X(\omega)(1 - \cos \omega T)$ 。

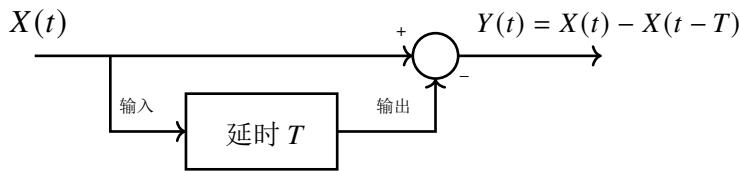


图 3-10 线性系统

练习 3.8 已知平稳过程

$$X(t) = \sum_{i=1}^N a_i Y_i(t), \quad (3.99)$$

式中  $a_i$  是一组常实数, 而随机过程  $Y_i(t)$  皆为平稳过程且相互正交。证明:

$$G_X(\omega) = \sum_{i=1}^N a_i^2 G_{Y_i}(\omega). \quad (3.100)$$

练习 3.9 已知平稳过程  $X(t)$  和  $Y(t)$  相互独立, 功率谱密度分别为

$$G_X(\omega) = \frac{16}{\omega^2 + 16}, \quad G_Y(\omega) = \frac{\omega^2}{\omega^2 + 16}.$$

令新的随机过程

$$\begin{cases} Z(t) = X(t) + Y(t) \\ V(t) = X(t) - Y(t). \end{cases} \quad (3.101)$$

① 证明  $X(t)$  和  $Y(t)$  联合平稳。

- ② 求  $Z(t)$  的功率谱密度  $G_Z(\omega)$ 。
- ③ 求  $X(t)$  和  $Y(t)$  的互谱密度  $G_{XY}(\omega)$ 。
- ④ 求  $X(t)$  和  $Z(t)$  的互相关函数  $R_{XZ}(\tau)$ 。
- ⑤ 求  $V(t)$  和  $Z(t)$  的互相关函数  $R_{VZ}(\tau)$

**练习 3.10** 已知可微平稳过程  $X(t)$  的功率谱密度为

$$G_X(\omega) = \frac{4}{\omega^2 + 1}. \quad (3.102)$$

- ① 证明过程  $X(t)$  和导数  $Y(t) = X'(t)$  联合平稳。
- ② 求互相关函数  $R_{XY}(\tau)$  和互谱密度  $G_X(\omega)$ 。

**练习 3.11** 已知可微平稳过程  $X(t)$  的自相关函数为  $R_X(\tau) = 2 \exp[-\tau^2]$ , 其导数为  $Y(t) = X'(t)$ 。求互谱密度  $G_X(\omega)$  和功率谱密度  $G_X(\omega)$ 。

**练习 3.12** 已知随机过程  $W(t) = X(t)Y(t)$ , 式中随机过程  $X(t), Y(t)$  联合平稳,  $\omega$  为常数。

- ① 讨论  $X(t), Y(t)$  及其均值和自相关函数在什么条件下, 才能使随机过程  $W(t)$  宽平稳。
- ② 利用①的结论, 用功率谱密度  $G_X(\omega), G_Y(\omega), G_{XY}(\omega)$  表示  $W(t)$  的功率谱密度  $G_W(\omega)$ 。
- ③ 若  $X(t), Y(t)$  互不相关, 求  $W(t)$  的功率谱密度  $G_W(\omega)$ 。

**练习 3.13** 已知平稳过程  $X(t), Y(t)$  互不相关, 它们的均值  $m_X, m_Y$  皆不为零。令新的随机过程  $Z(t) = X(t) + Y(t)$ , 求互谱密度  $G_{XY}(\omega)$  和  $G_X(\omega)$ 。

**练习 3.14** 已知复过程  $X(t)$  为宽平稳过程, 证明:

- ①  $X(t)$  的自相关函数为  $R_X(\omega) = R_{X^*}(-\omega)$ 。
- ② 已知复过程  $X$  的自相关函数  $R_X(\omega)$  和其功率谱密度  $G_X(\omega)$  也满足维纳辛钦定理, 为一对傅里叶变换对。证明功率谱密度  $G_X(\omega)$  为实函数。

**练习 3.15** 已知可微平稳过程  $X(t)$  的功率谱密度为

$$G_X(\omega) = \frac{4a^2\beta}{(a^2 + \omega^2)^2}, \quad (3.103)$$

其中  $\alpha, \beta$  皆为正实常数, 求随机过程  $X(t)$  和其导数  $Y(t) = X'(t)$  的互谱密度。

**练习 3.16** 已知随机过程  $X(t), Y(t)$  为

$$\begin{cases} X(t) = a \cos(\omega_0 t + \theta) \\ Y(t) = A(t) \cos(\omega_0 t + \theta) \end{cases} \quad (3.104)$$

式中  $a, \omega_0$  为实正常数,  $A(t)$  是具有恒定均值  $m_A$  的随机过程,  $\theta$  为与  $A(t)$  独立的随机变量。

① 运用互谱密度的定义式

$$G_{XY}(\omega) = \lim_{T \rightarrow \infty} \frac{1}{2T} E [X_T^*(\omega) Y_T(\omega)]. \quad (3.105)$$

证明: 无论随机变量的概率密度形式如何, 总有

$$G_{XY}(\omega) = \frac{\pi a m_A}{2} [\delta(\omega - \omega_0) + \delta(\omega + \omega_0)]. \quad (3.106)$$

② 证明:  $X(t), Y(t)$  的互相关函数为

$$\begin{aligned} R_{XY}(t, t + \tau) = & \frac{am_A}{2} \{ \cos \omega_0 \tau + E[\cos(2\theta)] \cos(2\omega_0 t + \omega_0 \tau) \\ & - E[\sin(2\theta)] \sin(2\omega_0 t + \omega_0 \tau) \}. \end{aligned} \quad (3.107)$$

③ 求互相关函数  $R_{XY}(t, t + \tau)$  的时间平均  $\overline{R_{XY}(t, t + \tau)}$ 。

练习 3.17 已知平稳过程  $X(t)$  的物理功率谱密度为

$$F_X(\omega) = \begin{cases} 4, & \omega \geq 0 \\ 0, & \text{其他} \end{cases}. \quad (3.108)$$

①求  $X(t)$  的功率谱密度  $G_X(\omega)$  和自相关函数  $R_X(\tau)$ 。画出  $F_X(\omega), G_X(\omega), R_X(\tau)$  的图形。

②判断过程  $X(t)$  是白噪声还是色噪声, 给出理由。

# 4

## 随机信号通过线性系统的分析

随机信号通过线性系统的分析是信号进行统计处理的基础。由于随机信号不能像确定信号那样用明确的数学表达式来描述，只能用概率统计的方法进行描述，因此，研究随机信号通过线性系统的输出，也只能从分析系统输出的统计特性入手。直接获取系统输出的概率分布一般比较困难，但在许多实际应用中，如果知道了系统输出的一些统计特性（如均值相关函数、功率谱密度函数等）往往就能解决问题。本章首先介绍信号与系统的基本概念、分类和特点；然后分别从连续系统、离散系统两方面，分析随机信号通过线性时不变系统的统计特性，重点介绍系统输出的一阶和二阶统计特性；在此基础上，介绍平稳随机序列的另一种分析方法及随机序列参数模型的概念。

### 4.1 信号与系统概述

#### 4.1.1 信号及其分类

信号是信息的表现形式，通常反映为随着自变量变化而变化的某种物理量。在数学上，信号一般可以表示成单个或多个自变量的函数。如：电信号可以看作是电流或电压随时间变化的函数，语音信号可以看作是声压随时间变化的函数，而图像信号可以看作是像素灰度随坐标位置变化的函数。自变量可以是时间，也可以是距离、高度、温度等，为了表述方便，这里将这些变量都统称为时间，这样信号就可以理解为随时间变化的函数，称为时间信号，简称信号。

根据信号的函数特点及其频谱分布特性，一般对信号进行如下分类。

- 确定信号与随机信号

根据信号变化规律是否可预知，可以将信号分为确定信号和随机信号。能够用确定的数学表达式来描述变化规律的信号称为确定信号，如我们熟知的正弦信号、指数信号等。对于这类信号，给定一个时刻，就有一个确定的值与之对应。而对于随机信号，其变化规律是不可预知的，不能用明确的数学表达式进行描述，因此也是不可重现的，如接收机内部的热噪声、地震波信号等。对于这类信号，即使在相同的条件下，每次观测到的信号（称为样本函数）也是不同的，只能通过概率统计的方法，分析多次观测得到的样本函数，才能掌握它们的变化规律。还有一些信号，其本质是随机信号，但信号的表现形式却类似确定信号，如随机相位信号  $s(t) = A \cos(\omega t + \phi)$ ，其表征信号的参数  $A$  和  $\omega$  是常量，但参数  $\phi$  是随机量，每次观测到的样本函数均是余弦波，但观测之前具体是什么相位的余弦波不可预测，因此该信号也是随机信号。更进一步，如果由随机信号所确定的任意维概率密度函数都满足时延不变性，该随机信号又称为严格平稳随机信号；如果仅有一阶和二阶统计量满足时延不变性，则称之为广义平稳随机信号；不具有上述特性的随机信号则称为非平稳随机信号。此外，还可以根据分布特性，将随机信号分为高斯信号与非高斯信号等。由于信号在传输过程中一般会受到周围环境的各种干扰及系统内部热噪声的影响，实际系统输出的信号多为随机信号，或者是确定信号与随机信号的叠加。

#### ► 连续信号与离散信号

根据函数自变量的取值是否连续，可以将信号分为连续信号和离散信号。除了若干个不连续点外，在自变量范围内的任何时刻都能连续取值的信号称为连续信号，如正弦信号、单位阶跃信号等。连续信号的函数值既可以是连续的，也可以是离散的，函数值连续取值的信号称为模拟信号，而只能取特定离散值的称为离散信号，它的自变量只能在一些规定的离散时刻才能取值，而在其他时刻则没有定义。离散信号可以由连续信号经过抽样后得到，也有些信号本身就是离散信号，如国家按年度统计的国民生产总值、股票交易数据、公司按月统计的销售额等。如果离散信号的函数值可连续取值，则称为抽样信号，若只能取离散值，则称为数字信号或时间序列。

#### ► 周期信号与非周期信号

根据信号按一定间隔是否周期重复，可以将信号分为周期信号与非周期信号。如果信号  $s(t)$  按一定的时间间隔  $T$  周而复始的变化，即  $s(t) = s(t + kT)$ ,  $k = 0, \pm 1, \pm 2, \dots, \pm n$ . 则称信号  $s(t)$  为周期信号， $T$  为信号的周期，如正弦信号  $s(t) = A \cos(\omega t + \phi)$  就是一个周期为  $2\pi/\omega$  的周期信号。而非周期信号，如指数信号，则不具有这种周而复始变化的特性。

► 能量信号与功率信号

根据信号的能量或功率是否为有限值，可以将信号分为能量信号与功率信号。下式分别定义信号的能量与平均功率：

$$E = \int_{-\infty}^{\infty} s^2(t) dt. \quad (4.1)$$

$$P = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} s^2(t) dt. \quad (4.2)$$

能量  $E < +\infty$  的信号称为能量信号；不满足能量有限，但满足  $P < +\infty$  的信号称为功率信号。根据定义可知，能量信号的平均功率为零，周期信号一般都是功率信号，而非周期信号则可能是能量信号，也可能是功率信号。上面的概念是从信号的时间域对信号进行分类。

► 基带信号与高频带通信号

根据信号频谱的分布特性，也可以将信号分为基带信号与高频带通信号。频谱主要集中在零值附近的信号称为基带信号。如果信号的频谱主要集中在以  $\pm\omega_0$  为中心频率、带宽为  $\Delta\omega$  的范围内，则该信号称为高频带通信号。根据带宽  $\Delta\omega$  相对载波  $\omega_0$  的比例关系，又可进一步将信号进行划分。一般将该比例在 1% 以内的信号称为窄带信号，常规的通信、雷达等系统中涉及的信号大多为窄带信号；该比例在 1%-20% 之间的信号称为宽带信号；超过 20% 的称为超宽带信号。窄带信号通常可表示为如下形式：

$$s(t) = a(t) \cos[\omega_0 t + \theta(t)] \quad (4.3)$$

式中  $a(t)$  和  $\theta(t)$  相对于载波变化缓慢，均为基带慢变信号。为了便于分析和处理问题，通常又将上式表示成复信号的形式：

$$\tilde{s}(t) = a(t) e^{j[\omega_0 t + \theta(t)]} = \tilde{a}(t) e^{j\omega_0 t}, \quad (4.4)$$

其中

$$\tilde{a}(t) = a(t) e^{j\theta(t)} = a(t) \cos \theta(t) + j a(t) \sin \theta(t) = s_I(t) + j s_Q(t), \quad (4.5)$$

称为信号  $s(t)$  的复包络， $s_I(t)$  和  $s_Q(t)$  称为信号  $s(t)$  的两个正交分量。

对于任意实信号  $s(t)$ ,  $-\infty < t < +\infty$ , 其复表示为

$$\tilde{s}(t) = s(t) + j \hat{s}(t), -\infty < t < +\infty, \quad (4.6)$$

其中  $\hat{s}(t) = \frac{1}{T} \int_{-\infty}^{\infty} \frac{s(\tau)}{t - \tau} d\tau$ , 如为  $s(t)$  的希尔伯特变换。 $\tilde{s}(t)$  也称为  $s(t)$  的解析信号或信号预包络。

### 4.1.2 系统及其分类

系统一般定义为由若干个相互联系的事物组成的具有特定功能的整体，它是个应用非常广泛的概念。在信号处理领域，系统可以看作是能够导致任何一种信号变换的过程。对于任何一个系统，不管其内部结构如何，都存在一个输入信号和一个(经过系统变换而与输入信号相关联的)输出信号，或者是多个输入信号和多个与之相关联的输出信号，前者称为单输入单输出系统，后者则称为多输入多输出系统，如图 4-1 所示。根

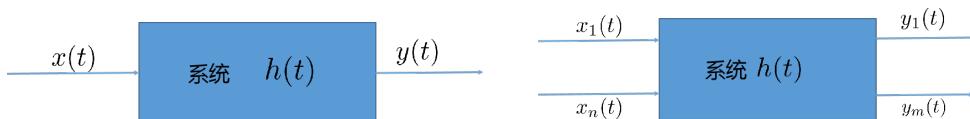


图 4-1 SISO 和 MIMO 系统

据系统处理的信号形式的不同及具有的某些性质，可以将系统分为如下 8 类：

#### 1. 连续系统与离散系统

如果系统的输入和输出均为连续信号，称这样的系统为连续系统。输入和输出均为离散信号的系统称为离散系统。由两者混合组成的系统则称为混合系统。

#### 2. 记忆系统与无记忆系统

若某一时刻系统的输出只与该时刻系统的输入有关，称这样的系统为无记忆系统。反之，若某一时刻系统的输出还与其他时刻的输入有关，则称为记忆系统。常见的无记忆系统有电阻器、检波器、限幅器等，记忆系统有积分器、累积器和平滑器等。

#### 3. 线性系统与非线性系统

能同时满足齐次性和叠加性的系统称为线性系统。其中的齐次性是指，当输入被扩大  $k$  倍后，系统的输出也相应地被扩大  $k$  倍；叠加性是指，当有多个输入同时作用于系统时，系统的输出等于各个输入单独作用后系统的输出之和。也就是说，对于线性系统，如果输入  $X_1(t)$  对应的系统输出为  $Y_1(t)$ ，输入  $X_2(t)$  对应的系统输出为  $Y_2(t)$ ，则输入  $k_1x_1(t) + k_2x_2(t)$  对应的输出为  $k_1y_1(t) + k_2y_2(t)$ 。不满足该性质的系统就称为非线性系统。

#### 4. 时变系统与时不变系统

如果系统的输出与输入的起始作用时间无关，当输入被延时，系统的输出仅作相应的延时而波形不变，这样的系统称为时不变系统。即对于时不变系统，如果输入  $X(t)$  延时到  $x(t - \tau)$ ，输出则从  $y(t)$  延时到  $y(t - \tau)$ 。系统的输出波形与输入的起始作用时间有关的系统就称为时变系统。

#### 5. 稳定系统与不稳定系统

如果输入有界，系统的输出也有界，这样的系统为稳定系统。如果系统对有界输入的输出是无限值，则为不稳定系统。显然，只有稳定系统才具有实际应用价值。

### 6. 因果系统与非因果系统

系统的输出与输入是相关的，如果系统在任一时刻的输出只与该时刻及该时刻以前的输入有关，而与该时刻以后的输入无关，称这样的系统为因果系统。如果输出还跟当前时刻以后的输入有关，则这样的系统称为非因果系统。根据定义，所有的无记忆系统都是因果系统。记忆系统则可能是因果系统，如积分器、累积器等；也可能是非因果系统，如平滑器、一些事后分析系统等。

### 7. 可逆系统与不可逆系统

如果对不同的输入系统能产生不同的输出，即根据系统的输出能唯一确定输入，称这样的系统为可逆系统。根据输出不能唯一确定输入的系统称为不可逆系统。对于任何可逆系统，均存在一个该系统的逆系统，将它与原系统级联后，能恢复出原系统的输入。可逆系统在实际中有着很多非常重要的应用，其中最典型的就是通信系统中的一些处理器件，如发送端的编码器、调制器等都是可逆系统，其对应的逆系统为接收端的解码器、解调器。

以上是根据系统的时域特性对系统进行分类。从频率域来分。

### 8. 最小相移系统与非最小相移系统

如果从系统的复频域上看，一个系统是因果稳定的线性时不变系统，则要求连续系统的所有极点必须位于  $S$  平面的虚轴左侧，或者离散系统的所有极点必须位于  $Z$  平面的单位圆内，而对系统的零点没有限制在此基础上，根据系统零点的分布情况，可以定义出最小相位系统。

#### 定义 4.1 最小相位系统

如果系统的所有零点也都位于  $S$  平面的虚轴左侧或者  $Z$  平面的单位圆内，则称该因果稳定的线性时不变系统为最小相位系统，否则，只要有一个零点位于  $S$  平面的虚轴右侧或者  $Z$  平面的单位圆外，就称为非最小相位系统。可见，最小相位系统一定存在逆系统，而且该逆系统既是因果稳定系统，也是最小相位系统。

此外，根据输入信号、系统和输出信号之间的关系，如果已知一个输入信号，将其通过一个已知的系统来寻求输出信号，一般称之为滤波；如果已知一个输入信号，根据其通过一个系统的输出信号来寻求系统传输函数，称之为系统辨识；如果根据一个已知系统的输出信号来恢复出系统的输入信号，则称之为逆滤波；如果根据一个系统的部分特性和系统输出信号来确定系统传递函数并恢复出输入信号，则称之为盲均衡和盲分离。

### 4.1.3 线性时不变系统

#### (1) 线性时不变冲激响应

根据冲激函数  $\delta(t)$  的性质

$$x(t) = \int_{-\infty}^{\infty} x(\tau) \delta(t - \tau) d\tau, \quad (4.7)$$

带入式  $y(t) = L[x(t)]$ , 并考虑运用线性算子对时间  $t$  进行运算, 故有

$$\begin{aligned} y(t) &= L(x(t)) = L\left[\int_{-\infty}^{\infty} x(\tau) \delta(t - \tau) d\tau\right] \\ &= \int_{-\infty}^{\infty} x(\tau) L[\delta(t - \tau)] d\tau, \end{aligned} \quad (4.8)$$

定义其中的函数  $h(t) = L[\delta(t)]$ , 函数  $h(t)$  是线性系统在输入冲激  $\delta(t)$  情况下的输出, 所以通常称  $h(t)$  为线性系统的冲激响应, 它是反映系统时域特性的函数。于是线性系统的输出为

$$y(t) = \int_{-\infty}^{\infty} x(\tau) h(t, \tau) d\tau \quad (4.9)$$

#### (2) 线性时不变系统

若输入信号  $x(t)$  有时移, 使系统输出  $y(t)$  有一个相同的时间移动, 即

$$y(t - \tau) = L[\delta(t - \tau)], \quad (4.10)$$

则这个系统的算子  $L(\cdot)$  就称为线性时不变系统。因此, 线性时不变系统的冲激响应也应满足下式

$$h(t - \tau) = L[\delta(t - \tau)], \quad (4.11)$$

故可得: 一个线性时不变系统的输出响应为

$$y(t) = \int_{-\infty}^{\infty} x(\tau) h(t - \tau) d\tau, \quad (4.12)$$

变量也可置换, 写成

$$y(t) = \int_{-\infty}^{\infty} x(t - \tau) h(\tau) d\tau; \quad (4.13)$$

 **注 4.2.** 线性时不变系统的输出响应 (4.12) 和 (4.13) 等价。

线性时不变系统的输出响应 (4.12) 和 (4.13) 经常被记作

$$y(t) = x(t) * h(t) = \int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau = \int_{-\infty}^{\infty} x(t-\tau)h(\tau)d\tau, \quad (4.14)$$

上式表明线性时不变系统的响应完全由系统的输入  $x(t)$  与系统的冲激响应  $h(t)$  所确定。

对式 (4.14) 作傅氏变换, 则  $Y(\omega), X(\omega), H(\omega)$  三者之间在频域中的关系为

$$Y(\omega) = X(\omega)H(\omega), \quad (4.15)$$

式中  $H(\omega)$  称系统的传递函数, 传递函数与系统的冲激响应构成一对傅里叶变换, 即

$$h(t) \xleftrightarrow[F]{F^{-1}} H(\omega). \quad (4.16)$$

(3) 系统的稳定性和物理可实现

(1) 稳定系统

### 定义 4.3 系统的稳定

若一个线性时不变系统, 对任意有界输入, 其输出均有界, 则称此系统是稳定的。 

根据定义 4.3, 寻找使得系统稳定的条件。若输入信号有界, 则必有正常数  $A$  存在, 使得对于所有的  $t$ , 有  $|x(t)| \leq A < \infty$  成立。由式 (4.13) 可知, 对所有的  $t$ , 显然有下式成立。

$$|y(t)| = \left| \int_{-\infty}^{\infty} h(\tau)x(t-\tau)d\tau \right| \leq \int_{-\infty}^{\infty} |h(\tau)||x(t-\tau)|d\tau \leq A \int_{-\infty}^{\infty} |h(\tau)|d\tau. \quad (4.17)$$

如果要求系统稳定, 即要求输出有界,  $|y(t)| \leq B < \infty$ , 则稳定系统的冲激应满足下述条件

$$\int_{-\infty}^{\infty} |h(\tau)|d\tau < \infty, \quad (4.18)$$

即冲激响应  $h(t)$  绝对可积。

(2) 物理可实现性(因果性)

工程上为使系统在物理上有可实现性, 必须要求系统在考察 ( $t = 0$ ) 以前不产生响应, 即系统具有因果性。也就是说系统的冲激响应  $h(t)$  函数应满足

$$h(t) = 0, \quad t < 0. \quad (4.19)$$

所有实际运行的物理可实现系统都是因果的, 都满足上式。于是, 对于物理可实现系统来说, 有

$$y(t) = \int_0^{\infty} h(\tau)x(t-\tau)d\tau = \int_{-\infty}^t x(\tau)h(t-\tau)d\tau. \quad (4.20)$$

物理可实现系统的传递函数为

$$H(\omega) = \int_0^{\infty} h(t) e^{-j\omega t} dt, \quad (4.21)$$

上式也可以用复频率 ( $s = \sigma + j\omega$ ) 表示, 若以  $s$  代替  $H(\omega)$  中的  $j\omega$ , 则有

$$H(s) = \int_0^{\infty} h(t) e^{-st} dt. \quad (4.22)$$

### (3) 稳定的物理可实现系统

#### 定义 4.4 稳定的物理可实现系统

若物理可实现系统的传递函数  $H(s)$  的所有极点都位于  $S$  平面的左半平面 (不含虚轴), 则称此系统为“稳定的物理可实现系统”。



#### 例 4.1.1

某系统的传递函数用拉氏变换形式表示, 讨论其稳定性。

$$(1) H_1(s) = \frac{1}{a+s}, \quad a > 0; (2) H_2(s) = \frac{1}{a-s}, \quad a > 0.$$



解: ①  $H_1(s)$  的极点  $s_1 = -a < 0$  在左半平面, 冲激响应为  $h_1(t) = e^{-at} (t > 0)$ , 此系统的冲激响应绝对可积, 所以此系统是稳定的。

②  $H_2(s)$  的极点  $s_2 = a > 0$  在右半平面, 相应的冲激响应  $h_2(t) = e^{at} (t > 0)$ , 此系统的冲激响应非绝对可积, 所以此系统是不稳定的。

## 4.2 随机信号通过线性时不变系统

上节中我们讨论了信号与系统的一些概念、分类和特点, 有关确定信号通过线性系统的分析已经在信号与系统的相关课程中作了介绍。而实际应用中的信号大多是随机信号, 这里我们利用信号与系统、随机过程等课程中的相关知识, 分析随机信号通过线性时不变系统的统计特性, 重点介绍系统输出的一阶和二阶统计特性。

考虑随机信号  $\{X(t)\}$  通过一个连续线性时不变系统, 系统的冲激响应为  $h(t)$ , 输出为随机信号  $\{Y(t)\}$ , 如图 4-2 所示。由于随机信号  $\{X(t)\}$  可以看作是大量样本函数的集合, 每个样本函数  $x(t)$  对应每次试验的结果, 均为确定的时间函数。当样本函数  $X(t)$  通过系统时, 其输出  $y(t)$  与输入  $x(t)$  之间存在如下卷积关系:

$$y(t) = x(t) * h(t) = \int_{-\infty}^{\infty} h(\tau) x(t - \tau) d\tau = \int_{-\infty}^{\infty} h(t - \tau) x(\tau) d\tau. \quad (4.23)$$

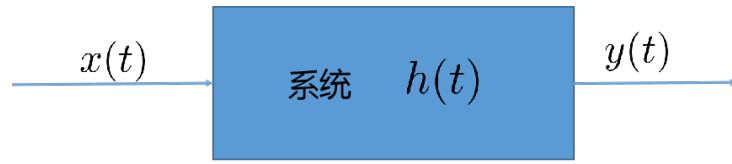


图 4-2 随机信号通过线性时不变系统示意图

如果  $x(t)$  和  $h(t)$  绝对可积，即系统是稳定的，且傅里叶变换分别为  $X(\omega), H(\omega)$ ，则  $y(t)$  的傅里叶变换  $Y(\omega)$  满足如下等式：

$$Y(\omega) = H(\omega)X(\omega), \quad (4.24)$$

其中， $H(\omega)$  又称为时不变线性系统的传递函数，传递与系统的冲激响应  $h(t)$  构成傅里叶变换对，关系如下：

$$H(\omega) = \int_{-\infty}^{\infty} h(t)e^{-j\omega t} dt \Rightarrow H(0) = \int_{-\infty}^{\infty} h(t)dt. \quad (4.25)$$

$$H^*(\omega) = \int_{-\infty}^{\infty} h(t)e^{j\omega t} dt \quad (4.26)$$

$$h(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(\omega)e^{j\omega t} d\omega. \quad (4.27)$$

即

$$h(t) \xrightleftharpoons[F^{-1}]{F} H(\omega). \quad (4.28)$$

这样，对于  $\{X(t)\}$  的一系列样本函数  $x(t)$ ，系统输出端就会得到一系列新的样本函数  $y(t)$ ，这些样本函数就构成随机信号  $\{Y(t)\}$ 。

### 4.2.1 系统输出的均值

根据 (4.23) 式，输出  $\{Y(t)\}$  的均值可表示为

$$\begin{aligned} m_Y(t) &= E\{Y(t)\} = \int_{-\infty}^{+\infty} h(\tau)E\{X(t-\tau)\}d\tau \\ &= \int_{-\infty}^{+\infty} h(\tau)m_X(t-\tau)d\tau. \end{aligned} \quad (4.29)$$

当输入  $\{X(t)\}$  广义平稳时， $m_X(t) = m_X$ ，则由式 (4.25)，可得

$$m_Y(t) = m_X \int_{-\infty}^{+\infty} h(\tau)d\tau = m_X H(0) = m_Y. \quad (4.30)$$

可见，此时输出均值也与时间无关。

## 4.2.2 系统输出的自相关函数和功率谱密度函数

根据 (4.23) 式, 输出  $\{Y(t)\}$  的自相关函数可表示成

$$\begin{aligned} R_Y(t_1, t_2) &= E\{Y(t_1)Y(t_2)\} \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(\tau_1)h(\tau_2) E\{X(t_1 - \tau_1)X(t_2 - \tau_2)\} d\tau_1 d\tau_2 \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(\tau_1)h(\tau_2) R_X(t_1 - \tau_1, t_2 - \tau_2) d\tau_1 d\tau_2. \end{aligned} \quad (4.31)$$

当输入  $\{X(t)\}$  广义平稳,  $R_X(t_1, t_2) = R_X(t_2 - t_1) = R_X(\tau)$ , 则

$$R_Y(t_1, t_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(\tau_1)h(\tau_2) R_X(\tau + \tau_1 - \tau_2) d\tau_1 d\tau_2 = R_Y(\tau). \quad (4.32)$$

此时输出  $\{Y(t)\}$  的均方值, 即平均功率为

$$\begin{aligned} E\{Y^2(t)\} &= R_Y(0) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(\tau_1)h(\tau_2) R_X(\tau_2 - \tau_1) d\tau_1 d\tau_2 \\ &\leq \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |h(\tau_1)h(\tau_2)R_X(\tau_2 - \tau_1)| d\tau_1 d\tau_2 \\ &\leq R_X(0) \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |h(\tau_1)h(\tau_2)| d\tau_1 d\tau_2 \\ &= R_X(0) \left( \int_{-\infty}^{+\infty} |h(\tau)| d\tau \right)^2 < +\infty. \end{aligned} \quad (4.33)$$

(4.30) 式、(4.32) 式以及式 (4.33) 表明: 如果广义平稳随机信号  $\{X(t)\}$  通过线性时不变系统, 则输出随机信号也广义平稳。

进一步, 根据维纳—辛钦定理, 即功率谱密度函数与自相关函数是一组傅里叶变换对, 对 (4.32) 式两边进行傅里叶变换, 可得到输出  $\{Y(t)\}$  的功率谱密度函数

$$\begin{aligned} G_Y(\omega) &= \int_{-\infty}^{+\infty} R_Y(\tau) e^{-j\omega\tau} d\tau \\ &\stackrel{(4.32)}{=} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(\tau_1)h(\tau_2) R_X(\tau - \tau_1 + \tau_2) e^{-j\omega\tau} d\tau_1 d\tau_2 d\tau \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(\tau_2) e^{-j\omega\tau} \int_{-\infty}^{+\infty} h(\tau_1) R_X(\tau - \tau_1 + \tau_2) d\tau_1 d\tau_2 d\tau. \end{aligned} \quad (4.34)$$

定义  $R_1(t) = R_X(t) * h(t) = \int_{-\infty}^{+\infty} R_X(t - \tau_1)h(\tau_1)d\tau_1$ , 则  $R_1(t)$  的傅里叶变换为  $H(\omega)G_X(\omega)$ ,

上式可进一步表示为

$$\begin{aligned}
 G_Y(\omega) &= \int_{-\infty}^{+\infty} h(\tau_2) \int_{-\infty}^{+\infty} R_1(\tau + \tau_2) e^{-j\omega\tau} d\tau d\tau_2 \\
 &= \int_{-\infty}^{+\infty} h(\tau_2) H(\omega) G_X(\omega) e^{j\omega\tau_2} d\tau_2 \\
 &= H(-\omega) H(\omega) G_X(\omega) \\
 &= |H(\omega)|^2 G_X(\omega),
 \end{aligned} \tag{4.35}$$

其中  $G_X(\omega) = \int_{-\infty}^{+\infty} R_X(\tau) e^{-j\omega\tau} d\tau$  为广义平稳输入  $\{X(t)\}$  的功率谱密度函数。式(4.35)表明, 广义平稳随机信号通过线性时不变系统, 输出功率谱密度等于输入功率谱密度乘以系统传输函数模的平方。可见, 输出功率谱密度只与系统的幅频特性有关, 而与系统的相频特性无关。

### 4.2.3 系统输入与输出的互相关函数和互功率谱密度函数

根据(4.23)式, 输入  $\{X(t)\}$  与输出  $\{Y(t)\}$  之间的互相关函数  $R_{XY}(t_1, t_2), R_{YX}(t_1, t_2)$  可表示为

$$\begin{aligned}
 R_{XY}(t_1, t_2) &= E\{X(t_1) Y(t_2)\} \\
 &= \int_{-\infty}^{+\infty} h(\tau_2) E\{X(t_1) X(t_2 - \tau_2)\} d\tau_2 \\
 &= \int_{-\infty}^{+\infty} h(\tau_2) R_X(t_1, t_2 - \tau_2) d\tau_2.
 \end{aligned} \tag{4.36}$$

$$\begin{aligned}
 R_{YX}(t_1, t_2) &= E\{Y(t_1) X(t_2)\} \\
 &= \int_{-\infty}^{+\infty} h(\tau_1) E\{X(t_2) X(t_1 - \tau_1)\} d\tau_1 \\
 &= \int_{-\infty}^{+\infty} h(\tau_1) R_X(t_1 - \tau_1, t_2) d\tau_1.
 \end{aligned} \tag{4.37}$$

当输入  $\{X(t)\}$  广义平稳,  $R_X(t_1, t_2) = R_X(t_2 - t_1) = R_X(t_1 - t_2) = R_X(\tau)$ , 则

$$R_{XY}(t_1, t_2) = \int_{-\infty}^{+\infty} h(\tau_2) R_X(\tau + \tau_2) d\tau_2 = R_{XY}(\tau). \tag{4.38}$$

$$R_{YX}(t_1, t_2) = \int_{-\infty}^{+\infty} h(\tau_1) R_X(\tau - \tau_1) d\tau_1 = R_{YX}(\tau) = R_{XY}(-\tau). \tag{4.39}$$

上式表明, 广义平稳随机信号通过线性时不变系统, 其输入与输出联合广义平稳。同样,

对以上等式两边进行傅里叶变换，可得输入与输出之间的互功率谱密度函数：

$$\begin{aligned} G_{XY}(\omega) &= \int_{-\infty}^{+\infty} R_{XY}(\tau) e^{-j\omega\tau} d\tau = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(\tau_2) R_X(\tau + \tau_2) e^{-j\omega\tau} d\tau_2 d\tau \\ &= H(-\omega) G_X(\omega). \end{aligned} \quad (4.40)$$

$$\begin{aligned} G_{YX}(\omega) &= \int_{-\infty}^{+\infty} R_{YX}(\tau) e^{-j\omega\tau} d\tau = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(\tau_1) R_X(\tau - \tau_1) e^{-j\omega\tau} d\tau_1 d\tau \\ &= H(\omega) G_X(\omega). \end{aligned} \quad (4.41)$$

结合(4.35)式，有如下关系式成立：

$$\begin{aligned} G_Y(\omega) &= H(-\omega) H(\omega) G_X(\omega) = H(\omega) G_{XY}(\omega) \\ &= H(-\omega) G_{YX}(\omega). \end{aligned} \quad (4.42)$$

综上所述，当广义平稳随机信号  $\{X(t)\}$  通过线性时不变系统，系统输出广义平稳，且输出与输入联合广义平稳。输出以及输入、输出之间的统计关系具体可如图 4-3 所示。

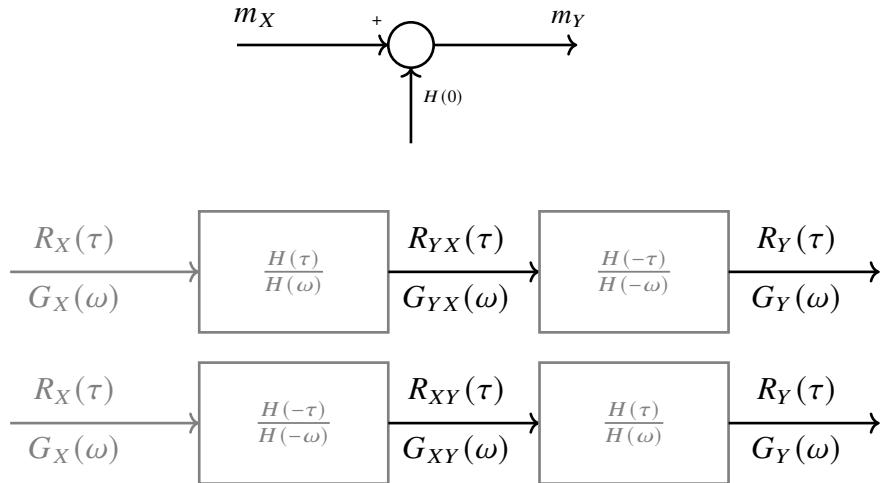


图 4-3 广义平稳随机信号通过线性时不变系统示意图

#### 4.2.3.1 物理可实现的线性时不变系统的平稳性

上述分析虽然基于物理不可实现的线性时不变系统，但是当广义平稳随机信号通过物理可实现的线性时不变系统时，相关结论仍然成立，只是此时  $h(t) = 0, t < 0$ 。而当系统输入  $\{X(t)\}$  只从  $t = 0$  时刻开始加入，即  $x(t) = 0, t < 0$ ，结论就有所不同了。

(1) 输入  $\{X(t)\}$  广义平稳，但输出  $\{Y(t)\}$  的均值不再是常数，这种条件下的系统输

出不再是平稳的。此时输出  $\{Y(t)\}$  的均值和相关函数分别为

$$m_Y(t) = E\{Y(t)\} = \int_{-\infty}^t h(\tau)E\{X(t-\tau)\}d\tau = m_X \int_{-\infty}^t h(\tau)d\tau. \quad (4.43)$$

$$\begin{aligned} R_Y(t_1, t_2) &= E\{Y(t_1)Y(t_2)\} \\ &= \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} h(\tau_1)h(\tau_2)E\{X(t_1-\tau_1)X(t_2-\tau_2)\}d\tau_1d\tau_2 \\ &= \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} h(\tau_1)h(\tau_2)R_X(\tau-\tau_1+\tau_2)d\tau_1d\tau_2. \end{aligned} \quad (4.44)$$

$$\begin{aligned} R_{XY}(t_1, t_2) &= E\{X(t_1)Y(t_2)\} = \int_{-\infty}^{t_2} h(\tau_2)E\{X(t_1)X(t_2-\tau_2)\}d\tau_2 \\ &= \int_{-\infty}^{t_2} h(\tau_2)R_X(\tau+\tau_2)d\tau_2. \end{aligned} \quad (4.45)$$

$$\begin{aligned} R_{YX}(t_1, t_2) &= E\{X(t_2)Y(t_1)\} = \int_{-\infty}^{t_2} h(\tau_1)E\{X(t_2)X(t_1-\tau_1)\}d\tau_1 \\ &= \int_{-\infty}^{t_1} h(\tau_1)R_X(\tau-\tau_1)d\tau_1. \end{aligned} \quad (4.46)$$

根据上述公式，尽管输入  $\{X(t)\}$  广义平稳，但输出  $\{Y(t)\}$  的均值不再是常数，而与时间  $t$  有关，且自相关函数跟时间差值  $T$  以及时间  $t_1$  和  $t_2$  都有关，说明这种条件下的系统输出不再是平稳的。

(2) 若输入  $X(t)$  是宽平稳的，则系统输出  $Y(t)$  也是宽平稳的，且输入与输出联合平稳。

若  $X(t)$  宽平稳，则有

$$\begin{cases} E[X(t)] = m_X (\text{常数}) \\ R_X(t_1, t_2) = R_X(\tau), \quad \tau = t_2 - t_1 \\ R_X(0) = E[X^2(t)] < \infty \end{cases} \quad (4.47)$$

利用时域分析的结果，可得

$$E[Y(t)] = \int_0^\infty h(\tau)E[X(t-\tau)]d\tau = m_X \int_0^\infty h(\tau)d\tau = m_Y. \quad (4.48)$$

$$R_{XY}(t_1, t_2) = \int_0^\infty h(u)R_X(\tau-u)du = R_X(\tau)*h(\tau) = R_{XY}(\tau). \quad (4.49)$$

$$R_{YX}(t_1, t_2) = \int_0^\infty h(u)R_X(\tau+u)du = R_X(\tau)*h(-\tau) = R_{YX}(\tau). \quad (4.50)$$

$$\begin{aligned} R_Y(t_1, t_2) &= \int_0^\infty \int_0^\infty h(u)h(v)R_X(t_2-t_1-v+u)dudv \\ &= \int_0^\infty \int_0^\infty h(u)h(v)R_X(\tau-v+u)dudv = R_Y(\tau). \end{aligned} \quad (4.51)$$

此外, 系统输出的均方值为

$$\begin{aligned}
 E[Y^2(t)] &= \left| E[Y^2(t)] \right| = \left| \int_0^\infty \int_0^\infty h(u)h(v)R_X(u-v)dudv \right| \\
 &\leq \int_0^\infty \int_0^\infty |h(u)||h(v)| |R_X(u-v)| dudv \\
 &\leq R_X(0) \int_0^\infty \int_0^\infty |h(u)||h(v)| dudv \\
 &= R_X(0) \int_0^\infty |h(u)|du \cdot \int_0^\infty |h(v)|dv.
 \end{aligned} \tag{4.52}$$

如果系统是稳定的, 则由  $\int_0^\infty |h(\tau)|d\tau < \infty$ , 可推出

$$|E[Y^2(t)]| < \infty \tag{4.53}$$

① 由 (4.48), (4.51) 和 (4.53), 可证, 输出  $Y(t)$  是宽平稳过程。

② 由 (4.49) 或 (4.50) 可证, 输入与输出之间是联合宽平稳的。

若用卷积形式, 则上述各式可表示成

$$R_{XY}(\tau) = R_X(\tau) * h(\tau) \tag{4.54}$$

$$R_{YX}(\tau) = R_X(\tau) * h(-\tau) \tag{4.55}$$

$$R_Y(\tau) = R_X(\tau) * h(-\tau) = R_{XY}(\tau) * h(-\tau) = R_{YX}(\tau) * h(\tau). \tag{4.56}$$

(3) 若输入  $X(t)$  是严平稳的, 则输出  $Y(t)$  也是严平稳的。

证: 因为是线性时不变系统, 可证对于任何时移  $\tau$ , 都有

$$Y(t + \tau) = L[X(t + \tau)] \tag{4.57}$$

成立。系统算子不变, 输出  $Y(t + \tau)$  和输入  $X(t + \tau)$  之间的关系与  $Y(t), X(t)$  之间的是完一样的。当  $X(t + \tau)$  与  $X(t)$  具有相同的概率密度函数时,  $Y(t + \tau)$  与  $Y(t)$  也具有相同概率密度函数, 所以  $Y(t)$  也是严平稳的。

(4) 若输入  $X(t)$  是宽各态历经的, 则输出  $Y(t)$  也是宽各态历经的。

证: 由随机过程的宽各态历经定义, 输入  $X(t)$  满足

$$\begin{cases} \overline{X(t)} = m_X. \\ \overline{X(t)X(t + \tau)} = R_X(\tau). \end{cases} \tag{4.58}$$

则输出  $Y(t)$  的时间平均  $\overline{Y(t)}$  为

$$\begin{aligned}\overline{Y(t)} &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T Y(t) dt \\ &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-r}^T \left[ \int_0^\infty h(u) X(t-u) du \right] dt \\ &= \int_0^\infty \left[ \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T X(t-u) dt \right] h(u) du \\ &= \int_0^\infty m_X h(u) du = m_Y.\end{aligned}\tag{4.59}$$

$$\begin{aligned}\overline{Y(t)Y(t+\tau)} &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T Y(t)Y(t+\tau) dt \\ &= \int_0^\infty \int_0^\infty h(u)h(v) \left[ \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T X(t-u)X(t+\tau-v) dt \right] du dv \\ &= \int_0^\infty \int_0^\infty h(u)h(v) R_X(\tau+u-v) du dv \\ &= R_Y(\tau),\end{aligned}\tag{4.60}$$

故  $Y(t)$  是宽各态历经的。

**例 4.5** 低通 RC 电路是由电源  $X(t)$ , 电阻  $R$  和电容  $C$  组成的简单电路, 已知输入信号  $X(t)$  是宽平稳的双侧信号, 其均值为  $m_X$ , 求输出均值。

解: 由电路知, 可得此系统的冲激响应为  $h(t) = b e^{-bt} U(t)$ , 其中  $b = 1/RC$ , 则其输出均值为

$$m_Y = m_X \int_0^\infty b e^{-bu} du = -m_X e^{-bu} \Big|_0^\infty = m_X.\tag{4.61}$$

**例 4.6** 若上例中  $X(t)$  是自相关函数为  $\frac{N_0}{2}\delta(\tau)$  的白噪声, ①求输出的自相关函数, ②输出的平均功率, ③输入与输出间的互相关函数  $R_{XY}(\tau)$  和  $R_{YX}(\tau)$ 。

解: 由题意知  $R_X(\tau) = \frac{N_0}{2}\delta(\tau)$ , 则输出自相关函数  $R_Y(\tau)$  为

$$\begin{aligned}R_Y(\tau) &= E[Y(t_1)Y(t_2)] = E \left[ \int_{-\infty}^\infty h(u) \left[ \int_{-\infty}^\infty h(v) X(t_1-u)X(t_2-v) dv \right] du \right] \\ &= \int_0^\infty h(u) \int_0^\infty h(v) E[X(t_1-u)X(t_2-v)] dv du \\ &= \int_0^\infty h(u)h(v) R_X(t_1, t_2) dv du\end{aligned}\tag{4.62}$$

则

$$\begin{aligned} R_Y(\tau) &= \int_0^\infty h(u) \left[ \int_0^\infty \frac{N_0}{2} \delta(\tau + u - v) h(v) dv \right] du \\ &= \frac{N_0}{2} \int_0^\infty h(u) h(\tau + u) du, \end{aligned} \quad (4.63)$$

上式说明：当输入是白噪声时，输出信号的自相关函数正比于系统冲激响应的卷积。

$$R_Y(\tau) = \frac{N_0}{2} \int_0^\infty (be^{-bu}) U(u) (be^{-b(\tau+u)}) U(\tau+u) du. \quad (4.64)$$

将上式分别按  $\tau \geq 0$  与  $\tau < 0$  两种情况求解。

① 当  $\tau \geq 0$  时，有

$$R_Y(\tau) = \frac{N_0 b^2}{2} e^{-b\tau} \int_0^\infty e^{-2bu} du = \frac{N_0 b}{4} e^{-b\tau}. \quad (4.65)$$

② 当  $\tau < 0$  时，由自相关函数的偶对称性，得到输出的自相关函数

$$R_Y(\tau) = R_Y(-\tau) = \frac{N_0 b}{4} e^{b\tau}. \quad (4.66)$$

合并  $\tau \geq 0$  与  $\tau < 0$  时的结果，得到输出的自相关函数

$$R_Y(\tau) = \frac{N_0 b}{4} e^{-b|\tau|}, \quad |\tau| < \infty. \quad (4.67)$$

③ 在上式中令  $\tau = 0$ ，即可得输出的平均功率为

$$E[Y^2(t)] = R_Y(0) = \frac{b N_0}{4}. \quad (4.68)$$

由于  $b$  是时间常数  $RC$  的倒数， $b = 1/RC$ ，因此，也与电路的带宽  $\Delta f$  有关，其中

$$\Delta f = \frac{1}{2\pi RC} = \frac{b}{2\pi}. \quad (4.69)$$

于是输出平均功率又可写成

$$E[Y^2(t)] = \frac{\pi N_0}{2} \Delta f. \quad (4.70)$$

由此可见，该电路的输出平均功率随着电路的带宽变宽而线性地增大。

④ 输入和输出的互相关函数为

$$\begin{aligned} R_{XY}(\tau) &= \int_0^\infty \frac{N_0}{2} \delta(\tau - u) h(u) du = \frac{N_0}{2} h(\tau) U(\tau) \\ &= \begin{cases} \frac{N_0}{2} h(\tau), & \tau \geq 0 \\ 0, & \tau < 0 \end{cases} \end{aligned} \quad (4.71)$$

同理

$$\begin{aligned} R_{YX}(\tau) &= \int_0^\infty \frac{N_0}{2} \delta(\tau+u) h(u) du = \frac{N_0}{2} h(-\tau) U(-\tau) \\ &= \begin{cases} 0 & \tau > 0 \\ \frac{N_0}{2} h(-\tau) & \tau \leq 0 \end{cases} \end{aligned} \quad (4.72)$$

**例 4.7** 在例 4.6 中, 假设  $X(t)$  的自相关函数为  $R_X(\tau) = \frac{\beta N_0}{4} e^{-\beta|\tau|}$ , 式中  $\beta \neq b$ , 求输出的自相关函数。

$$\begin{aligned} R_Y(\tau) &= E[Y(t_1)Y(t_2)] = E \left[ \int_{-\infty}^{\infty} h(u) \left[ \int_{-\infty}^{\infty} h(v) X(t_1-u) X(t_2-v) dv \right] du \right] \\ &= \int_0^{\infty} h(u) \int_0^{\infty} h(v) E[X(t_1-u) X(t_2-v)] dv du \\ &= \int_0^{\infty} h(u) h(v) R_X(t_1, t_2) dv du \end{aligned} \quad (4.73)$$

则对于输入  $X(t)$  的自相关函数为  $R_X(\tau) = \frac{\beta N_0}{4} e^{-\beta|\tau|}$ , 输出的自相关函数可以写成

$$\begin{aligned} R_Y(\tau) &= \int_0^{\infty} \int_0^{\infty} R_X(\tau+u-v) h(u) h(v) du dv \\ &= \int_0^{\infty} \int_0^{\infty} \frac{\beta N_0}{4} e^{-\beta|\tau+u-v|} b e^{-bu} \cdot b e^{-bv} du dv. \end{aligned} \quad (4.74)$$

当  $\tau > 0$  时, 考虑到  $u, v$  均在 0 和  $\infty$  之间变化, 故先对  $v$  积分较方便。

$$\begin{aligned} R_Y(\tau) &= \frac{\beta N_0 b^2}{4} \int_0^{\infty} e^{-bu} \left[ \int_0^{\tau+u} e^{-\beta(\tau+u-v)} e^{-bv} dv + \int_{\tau+u}^{\infty} e^{\beta(\tau+u-v)} e^{-bv} dv \right] du \\ &= \frac{\beta N_0 b^2}{4(b^2 - \beta^2)} \left( e^{-\beta\tau} - \frac{\beta}{b} e^{-b\tau} \right), \quad \tau \geq 0 \end{aligned} \quad (4.75)$$

因自相关函数为  $\tau$  的偶函数, 所以  $\tau < 0$  时的  $R_Y(\tau)$  表达式能直接由  $R_Y(-\tau)$  时的表达写出。综合可得

$$R_Y(\tau) = \frac{b^2 \beta N_0}{4(b^2 - \beta^2)} \left( e^{-\beta|\tau|} - \frac{\beta}{b} e^{-b|\tau|} \right). \quad (4.76)$$

为了做比较, 上式可写为

$$R_Y(\tau) = \left( \frac{b N_0}{4} e^{-b|\tau|} \right) \cdot \left\{ \left[ \frac{1}{1 - b^2/\beta^2} \right] \left[ 1 - \frac{b}{\beta} e^{-(\beta-b)|\tau|} \right] \right\}, \quad (4.77)$$

式中的第一项因子是白噪声输入时系统输出的自相关函数, 第二项因子是非白噪声输入时系统输出的自相关函数附加的相乘因子。显然, 当  $\beta/b \rightarrow \infty$  时, 有

$$\lim_{\beta \rightarrow \infty} R_Y(\tau) = \frac{b N_0}{4} e^{-b|\tau|}, \quad (4.78)$$

关于  $Y$  的自相关函数  $R_Y(\tau)$  趋近于第一项因子。

由此可知, 当  $\beta$  较  $b$  大很多时, 第二项因子接近于 1,  $R_Y(\tau)$  趋近于第一项因子—白噪声输入时, 系统输出的自相关函数。由于  $\beta$  正比于输入信号的带宽  $b$ , 正比于系统的带宽, 因此, 由上面讨论, 可以得到: 在输入信号的带宽远大于系统带宽的情况下, 分析系统输出的统计特性时, 可利用白噪声来近似输入的随机信号。这既可以节省工作量, 又不会使精度降低太多。

例如, 在带宽为 10MHz 的高增益宽带放大器中, 最重要的噪声来源是第一级的热噪声, 这种热噪声的带宽可以到 1000Hz, 因此, 若取  $b/\beta$  为 0.01, 那么用白噪声近似热噪声时, 其误差不超过 1%。

#### 4.2.4 系统输出的概率密度

以上仅从系统输出的均值、相关函数及相应的功率谱密度三方面, 分析了随机信号通过线性时不变系统后的统计特性, 但若希望知道更多的统计特性, 就需要知道系统输出的概率密度函数。根据系统输出与输入的积分关系(4.23)式, 可将其写成求和的形式, 便于分析:

$$y(t) = \int_{-\infty}^{+\infty} h(\tau)x(t-\tau)d\tau = \sum_{i=-\infty}^{+\infty} h(\tau_i)x(t-\tau_i)\Delta\tau_i, \quad (4.79)$$

其中  $\Delta\tau_i = \tau_{i+1} - \tau_i$  足够小。可见, 系统输出是输入的线性组合, 利用随机变量的线性变换的概率密度函数与随机变量的概率密度函数之间的关系(即相差一个雅可比系数行列式, 见(1.82)式), 就可计算出系统输出的概率密度函数。如果系统输入是高斯信号, 由于高斯变量的线性组合仍然是高斯变量, 系统输出也是高斯信号。描述高斯信号的分布特性则比较简单, 此时只需要计算出系统输出的均值和相应的相关函数, 就可得到任意维的概率密度函数。

### 4.3 随机序列通过线性时不变系统

研究随机序列通过线性时不变系统, 分析输出统计特性的方法, 类似于上节中讨论连续随机信号的方法。

随机序列  $\{X(n)\}$  通过一个冲激响应为  $h(n)$  的离散线性时不变系统, 输出为随机序列  $\{Y(n)\}$ 。由于随机序列  $\{X(n)\}$  同样可以看作是大量样本序列的集合, 每个样本序列  $x(n)$  对应每次试验的结果, 均为确定的时间序列。当样本序列  $x(n)$  通过系统时, 其输

出  $Y(n)$  与输入  $X(n)$  之间存在如下卷积关系:

$$y(n) = x(n) * h(n) = \sum_{k=1}^{+\infty} h(k)x(n-k). \quad (4.80)$$

如果  $X(n), h(n)$  绝对可和, 且傅里叶变换分别为  $X(\Omega), H(\Omega)$ , 则  $Y(n)$  的傅里叶变换  $Y(\Omega)$  满足如下等式:

$$Y(\Omega) = H(\Omega)X(\Omega), \quad (4.81)$$

其中, 系统传递函数  $H(\Omega)$  与冲激响应  $h(n)$  构成傅里叶变换对, 满足如下关系:

$$H(\Omega) = \sum_{n=-\infty}^{+\infty} h(n)e^{-jn\Omega}. \quad (4.82)$$

$$h(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H(\Omega)e^{jn\Omega} d\Omega. \quad (4.83)$$

也即

$$h(n) \xrightarrow[F]{F^{-1}} H(\Omega). \quad (4.84)$$

对于  $\{X(t)\}$  的一系列样本序列  $x(n)$ , 系统输出端就会得到一系列新的样本序列  $y(n)$ , 这些样本序列就构成随机序列  $\{Y(t)\}$ . 此外, 在 (4.80) 式中, 当系统是物理可实现的线性时不变系统时,  $h(k) = 0, k < 0$ .

### 4.3.1 系统输出的均值

$$\begin{aligned} m_Y(n) &= E\{Y(n)\} = \sum_{k=-\infty}^{+\infty} h(k)E\{x(n-k)\} \\ &= \sum_{k=-\infty}^{+\infty} h(k)m_X(n-k). \end{aligned} \quad (4.85)$$

当输入  $\{X(n)\}$  广义平稳时,  $m_X(n) = m_X$ , 则

$$m_Y(n) = m_X \sum_{k=-\infty}^{+\infty} h(k) = m_X H(0) = m_Y. \quad (4.86)$$

此时输出均值与时间无关。

### 4.3.2 系统输出的自相关函数 $R_Y(n_1, n_2)$ 和功率谱密度函数 $G_Y(\Omega)$

$$\begin{aligned}
 R_Y(n_1, n_2) &= E\{Y(n_1)Y(n_2)\} \\
 &= \sum_{k_1=-\infty}^{+\infty} \sum_{k_2=-\infty}^{+\infty} h(k_1)h(k_2)E\{X(n_1-k_1)X(n_2-k_2)\} \\
 &= \sum_{k_1=-\infty}^{+\infty} \sum_{k_2=-\infty}^{+\infty} h(k_1)h(k_2)R_X(n_1-k_1, n_2-k_2). \tag{4.87}
 \end{aligned}$$

当输入  $\{X(t)\}$  广义平稳时，则

$$R_Y(n_1, n_2) = \sum_{k_1=-\infty}^{+\infty} \sum_{k_2=-\infty}^{+\infty} h(k_1)h(k_2)R_X(m-k_1+k_2) = R_Y(m), m \in \mathbb{Z}. \tag{4.88}$$

得出的均方值(即平均功率)为

$$\begin{aligned}
 E\{Y^2(n)\} &= R_Y(0) = \sum_{k_1=-\infty}^{+\infty} \sum_{k_2=-\infty}^{+\infty} h(k_1)h(k_2)R_X(-k_1+k_2) \\
 &\leq \sum_{k_1=-\infty}^{+\infty} \sum_{k_2=-\infty}^{+\infty} |h(k_1)h(k_2)R_X(-k_1+k_2)| \\
 &\leq R_X(0) \sum_{k_1=-\infty}^{+\infty} \sum_{k_2=-\infty}^{+\infty} |h(k_1)h(k_2)| \\
 &= R_X(0) \left( \sum_{k_1=-\infty}^{+\infty} |h(k_1)| \right)^2 < +\infty. \tag{4.89}
 \end{aligned}$$

由(4.86)式及上面两式，我们可得如下结论：当广义平稳随机序列  $\{X(t)\}$  仅从  $t_0$  时刻开始就加入到线性时不变系统，其输出随机序列也广义平稳；同样，利用维纳——辛钦公式，即功率谱密度函数与自相关函数是一组傅里叶变换对，对上式两边进行傅里叶变换，可得输出  $\{Y(t)\}$  的功率谱密度函数

$$\begin{aligned}
 G_Y(\Omega) &= \sum_{m=-\infty}^{+\infty} R_Y(m)e^{-j\Omega m} \\
 &= \sum_{m=-\infty}^{+\infty} \sum_{k_1=-\infty}^{+\infty} \sum_{k_2=-\infty}^{+\infty} h(k_1)h(k_2)R_X(m-k_1+k_2)e^{-j\Omega m} \\
 &= \sum_{m=-\infty}^{+\infty} \sum_{k_2=-\infty}^{+\infty} h(k_2)e^{-j\Omega m} \sum_{k_1=-\infty}^{+\infty} h(k_1)R_X(m+k_2-k_1), m \in \mathbb{Z} \tag{4.90}
 \end{aligned}$$

继续化简功率谱密度函数，得

$$\begin{aligned} G_Y(\Omega) &= \sum_{k_2=-\infty}^{+\infty} h(k_2) \sum_{m=-\infty}^{+\infty} e^{-j\Omega m} R_1(m+k_2) \\ &= \sum_{k_2=-\infty}^{+\infty} h(k_2) H(\Omega) G_X(\Omega) e^{j\Omega k_2} \\ &= H(\Omega) G_X(\Omega) H(-\Omega) = |H(\Omega)|^2 G_X(\Omega). \end{aligned}$$

其中  $R_1(n) = R_X(n) * h(n)$ ,  $G_X(\Omega) = \sum_{m=1}^{+\infty} R_X(m) e^{-j\omega m}$  为广义平稳输入,  $\{X(t)\}$  的功率谱密度函数。可见, 输出功率谱密度只与系统的幅频特性有关, 而与系统的相频特性无关。

### 4.3.3 系统输入与输出的互相关函数和互功率谱密度函数

$$\begin{aligned} R_{XY}(n_1, n_2) &= E\{X(n_1)Y(n_2)\} = \sum_{k_2=-\infty}^{+\infty} h(k_2) E\{X(n_1)X(n_2-k_2)\} \\ &= \sum_{k_2=-\infty}^{+\infty} h(k_2) R_X(n_1, n_2 - k_2). \end{aligned} \quad (4.91)$$

$$\begin{aligned} R_{YX}(n_1, n_2) &= E\{Y(n_1)X(n_2)\} = \sum_{k_1=-\infty}^{+\infty} h(k_1) E\{X(n_2)X(n_1-k_1)\} \\ &= \sum_{k_1=-\infty}^{+\infty} h(k_1) R_X(n_1 - k_1, n_2). \end{aligned} \quad (4.92)$$

当输入  $\{X(t)\}$  广义平稳时,  $R_X(n_1, n_2) = R_X(n_1, n_2) = R_X(m)$ , 则

$$R_{XY}(n_1, n_2) = \sum_{k_2=-\infty}^{+\infty} h(k_2) R_X(m + k_2) = R_{XY}(m). \quad (4.93)$$

$$R_{YX}(n_1, n_2) = \sum_{k_1=-\infty}^{+\infty} h(k_1) R_X(m - k_1) = R_{YX}(m) = R_{XY}(-m). \quad (4.94)$$

对以上等式两边进行傅里叶变换，可得输入与输出之间的互功率谱密度函数

$$\begin{aligned}
 G_{XY}(\Omega) &= \sum_{m=-\infty}^{+\infty} R_{XY}(m) e^{-j\Omega m} \\
 &= \sum_{m=-\infty}^{+\infty} \sum_{k_2=-\infty}^{+\infty} h(k_2) R_X(m+k_2) e^{-j\Omega m} \\
 &= H(-\Omega) G_X(\Omega). \\
 G_{YX}(\Omega) &= \sum_{m=-\infty}^{+\infty} R_{YX}(m) e^{-j\Omega m} \\
 &= \sum_{m=-\infty}^{+\infty} \sum_{k_2=-\infty}^{+\infty} h(k_1) R_X(m-k_1) e^{-jm} \\
 &= H(\Omega) G_X(\Omega).
 \end{aligned}$$

结合(4.90)式，有如下关系式：

$$G_Y(\Omega) = H(-\Omega)H(\Omega)G_X(\Omega) = H(-\Omega)G_{YX}(\Omega) = H(\Omega)G_{XY}(\Omega). \quad (4.95)$$

综上讨论，我们可以得到与连续随机信号相类似的结论：当广义平稳随机序列  $x(n)$  从  $-\infty$  时刻开始就加入到线性时不变系统，则系统输出广义平稳，且输出与输入联合广义平稳。系统输出以及系统输入、输出之间的统计关系也可参照图 4-3 所示。但是，如果广义平稳随机序列  $\{X(n)\}$  是从  $n=0$  时刻开始加入系统，即  $X(n)=0, n < 0$ ，此时系统输出不再平稳。

## 4.4 色噪声的产生与白化滤波器

下面将讨论在理论和实际应用中具有重要意义的两个问题。其一，如何设计一个线性系统，使其输入为白噪声时，能输出具有指定功率谱密度的色噪声；其二，如何设计一个线性系统，使其在输入为色噪声时，能输出白噪声。如图 4-4 所示

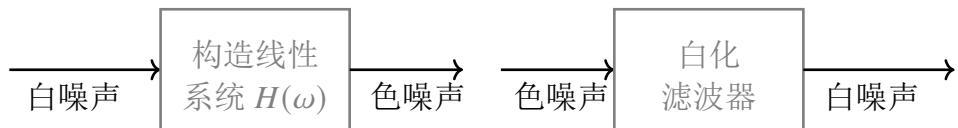


图 4-4 色噪声的产生和白化滤波器

### 1. 色噪声的产生

当具有单位功率谱密度的自噪声通过一个功率传递函数为  $|H(\omega)|$  的线性系统后, 其输出的功率谱密度为

$$G_Y(\omega) = |H(\omega)|^2 \cdot 1, \quad (4.96)$$

上式表明白噪声通过任一线性系统后, 由于输出的功率谱密度取决于系统的功率传递函数  $|H(\omega)|^2$ , 所以输出的功率谱密度不再是常数, 即不再是白噪声。

若我们需要产生(或模拟)一个指定功率谱密度的色噪声, 可以由给定的功率谱密度

$$G_Y(\omega) = H(\omega)H^*(\omega) \quad (4.97)$$

来设计一个线性系统的传递函数  $H(\omega)$ , 然后用白噪声去激励所设计的系统, 由输出端所获得的色噪声  $G_Y(\omega)$ 。注意: 所涉及的系统必须是物理可实现的稳定系统, 即选择的  $H(\omega)$  必满足稳定的物理可实现条件。

为了便于讨论, 将传递函数用复频率来表示, 以  $s$  代替  $H(\omega)$  中的  $j\omega$ , 以  $-s^2$  代替  $G_X(\omega)$  中的  $\omega^2$ , 即 ( $s = j\omega, s^2 = -\omega^2$ ), 则有

$$G_Y(s) = H(s)H(-s) \quad (4.98)$$

因此, 我们可以利用白噪声源和线性系统产生具有指定功率谱密度的色噪声, 方法如下

① 将指定色噪声的功率谱密度  $G_Y(s)$  分解成(示例代码见图 4-5)

$$G_Y(s) = \tilde{G}_Y(s)\tilde{G}_Y(-s). \quad (4.99)$$

② 根据物理可实现稳定的条件, 在  $\tilde{G}_Y(s)$  和  $\tilde{G}_Y(-s)$  中选择满足所有极点都位于  $S$  左半平面(不包含虚轴)的一个为  $H(s)$ 。

#### 例 4.4.1

设计一稳定的线性系统, 使其在单位白噪声激励下的输出功率谱为

$$G_Y(\omega) = \frac{25\omega^2 + 49}{\omega^4 + 10\omega^2 + 9} \quad (4.100)$$

解: 输出的功率谱密度  $G_Y(\omega)$  的复频域表达式( $s = -j\omega, s^2 = -\omega^2$ )为

$$G_Y(s) = \frac{49 - 25s^2}{s^4 - 10s^2 + 9} \quad (4.101)$$

将  $G_Y(\omega)$  进行谱分解

$$G_Y(s) = \frac{(7+5s)(7-5s)}{(1-s^2)(9-s^2)} = \frac{(7+5s)(7-5s)}{(1+s)(1-s)(3+s)(3-s)} \quad (4.102)$$

令

$$\tilde{G}_Y(s) = \frac{(5s+7)}{(s+1)(s+3)}, \quad \tilde{G}_Y(-s) = \frac{-(5s-7)}{(s-1)(s-3)} \quad (4.103)$$

因  $\tilde{G}_Y(s)$  位于  $S$  平面的所有极点  $s_1 = -1, s_2 = -3$  均在左半平面, 所以选  $\tilde{G}_Y(s)$  为  $H(s)$ , 即

$$H(s) = \tilde{G}_Y(s) = \frac{(5s+7)}{(s+1)(s+3)} \quad (4.104)$$

所设计系统的传递函数为

$$H(\omega) = \frac{(5j\omega+7)}{(j\omega+1)(j\omega+3)} \quad (4.105)$$

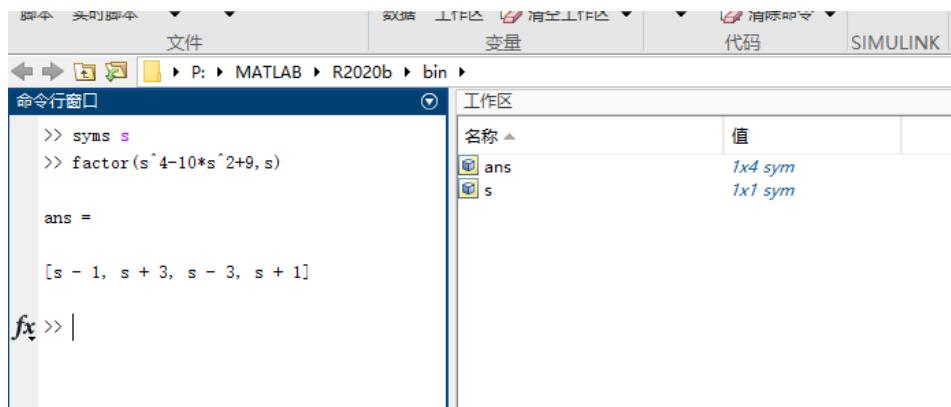


图 4-5  $s^4 - 10s^2 + 9$  的分解

## 2. 白化滤波器

在随机信号处理中, 往往会遇到等待处理的随机信号是非白噪声信号, 这样会给问题的解决带来困难。克服这一困难的措施之一是对色噪声进行白化处理。所谓白化处理就是设计一个稳定的线性滤波器, 使输入色噪声的系统输出是白噪声, 这实质上是求上一节问题的逆。我们称完成这一功能的系统为白化滤波器。色噪声白化的方法为:

① 先将色噪声的有理谱  $G_X(s)$  分解为 (若色噪声的谱密度不是有理函数, 常用有理函数来逼近)

$$G_X(s) = \widetilde{G}_X(s)\widetilde{G}_X(-s) \quad (4.106)$$

② 令系统的传递函数为

$$H(s) = \frac{1}{\widetilde{G}_X(s)} \quad \text{或} \quad H(s) = \frac{1}{\widetilde{G}_X(-s)} \quad (4.107)$$

要求  $H(s)$  是稳定的因果系统的传递函数(即在虚轴与右半平面无极点)。

③ 虽然输入系统的是色噪声, 但系统输出信号的功率谱密度是

$$G_Y(s) = H(s)H(-s)G_X(s) = \frac{1}{\widetilde{G}_X(s)\widetilde{G}_X(-s)} \cdot \widetilde{G}_X(s)\widetilde{G}_X(-s) = 1 \quad (4.108)$$

可见系统输出的是单位谱的白噪声。

由于白噪声具有数学运算简单等特点, 因而白化滤波器在信号检测、估计等方面得到广泛应用。

## 4.5 白噪声通过线性时不变系统

白噪声是实际应用中最常遇见的一种噪声, 它是一种广义平稳的随机过程。其特殊的统计特性使它在线性系统的分析应用中起着非常重要的作用。白噪声  $\{N(t)\}$  的均值为零, 功率谱密度均匀分布在无限宽的频带范围内, 即

$$G_N(\omega) = \frac{N_0}{2}, \quad -\infty < \omega < +\infty, \quad (4.109)$$

其中  $N_0$  为常数。白噪声的“白”字是借用了光学中白光的概念, 白光的光谱覆盖了所有可见光的频率。不具有上述均匀功率谱的噪声称为色噪声。根据维纳—辛钦定理, 我们可以进一步得到白噪声  $N(t)$  的自相关函数  $R_N(\tau)$  以及自相关系数  $\rho_N(\tau)$ :

$$R_N(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} G_N(\omega) e^{j\omega\tau} d\omega = \frac{N_0}{4\pi} \int_{-\infty}^{+\infty} e^{j\omega\tau} d\omega = \frac{N_0}{2} \delta(\tau). \quad (4.110)$$

$$\rho_N(\tau) = \frac{R_N(\tau)}{R_N(0)} = \begin{cases} 1, & \tau = 0 \\ 0, & \tau \neq 0 \end{cases}. \quad (4.111)$$

上式表明, 白噪声在任意相邻时刻的取值都是不相关的, 且平均功率  $R_N(0)$  无限大。然而, 上述定义的白噪声只是一种理想化的噪声模型, 在实际中是不可能存在的, 因为实际系统的带宽总是有限的, 有限带宽系统中白噪声的平均功率不可能是无限大的。通常情况下, 只要平稳随机过程的功率谱密度在比系统带宽大得多的频率范围内近似均匀分布, 就可以将其视为白噪声, 这种白噪声通常又被称为带白噪声。

### 4.5.1 系统输出的一般特性及等效噪声带宽

作为一种特殊的广义平稳随机过程, 白噪声通过线性时不变系统后, 系统的输出  $Y(t)$  和 4.3 节中所列举的一系列统计特性一致。根据 (4.29), ..., (4.41) 式, 我们可得如下

结论：

$$m_Y = m_N \int_{-\infty}^{+\infty} h(\tau) d\tau = m_N H(0) = 0. \quad (4.112)$$

$$\begin{aligned} R_Y(\tau) &= \frac{N_0}{2} \int_{-\infty}^{+\infty} h(\tau_1) h(\tau_2) \delta(\tau - \tau_1 + \tau_2) d\tau_1 d\tau_2 \\ &= \frac{N_0}{2} \int_{-\infty}^{+\infty} h(\tau_2) h(\tau + \tau_2) d\tau_2. \end{aligned} \quad (4.113)$$

$$G_Y(\omega) = \frac{N_0}{2} |H(\omega)|^2. \quad (4.114)$$

$$R_{NY}(\tau) = \frac{N_0}{2} h(-\tau) = R_{YN}(-\tau). \quad (4.115)$$

$$G_{NY}(\omega) = \frac{N_0}{2} H(-\omega) = G_{YN}(-\omega). \quad (4.116)$$

由 (4.114) 式可知，白噪声通过线性时不变系统后，输出一般不再具有均匀分布的功率谱密度，而是由系统的幅频响应来决定，输出一般是色噪声。此时，输出噪声的平均功率为

$$\begin{aligned} E[Y^2(t)] &= R_Y(0) \\ &\stackrel{(4.114)}{=} \frac{N_0}{2} \int_{-\infty}^{+\infty} h^2(\tau) d\tau = \frac{N_0}{4\pi} \int_{-\infty}^{+\infty} |H(\omega)|^2 d\omega \\ &= \frac{N_0}{2\pi} \int_0^{+\infty} |H(\omega)|^2 d\omega, \end{aligned}$$

式 (4.117) 的计算利用了实际系统的冲激响应一般为实函数的性质。

 **注 4.8.** 在实际应用中，为了便于计算系统的输出噪声功率，人们引入了等效噪声带宽的概念，用一个具有矩形幅频响应的理想系统来替代实际系统。当白噪声激励这个理想系统时，输出噪声的功率谱密度在通带范围内均匀分布。

如果白噪声通过频率响应为矩形的理想系统，那么输出的信号就是个“限带白噪声”。而利用白噪声作为背景和数学模型，又将会给随机信号的分析带来极大地方便。在实际应用中，为了分析的方便，通常会将一个频率响应为矩形的理想系统来代替实际系统。而白噪声通过这个系统的输出，就等效成了一个限带白噪声。因此，理想系统的带宽就是等效出来的限带白噪声功率谱的带宽，用  $\Delta\omega_e$  表示等效噪声带宽。若输出噪声的平均功率等于实际系统输出的平均功率，且理想系统  $H_e(\omega)$  在通带范围内的幅度值等于实际系统幅频响应的最大值，则理想系统的带宽  $\Delta\omega_e$  就定义为实际系统的等效噪声带宽，如图 4-6，此时输出噪声的平均功率为

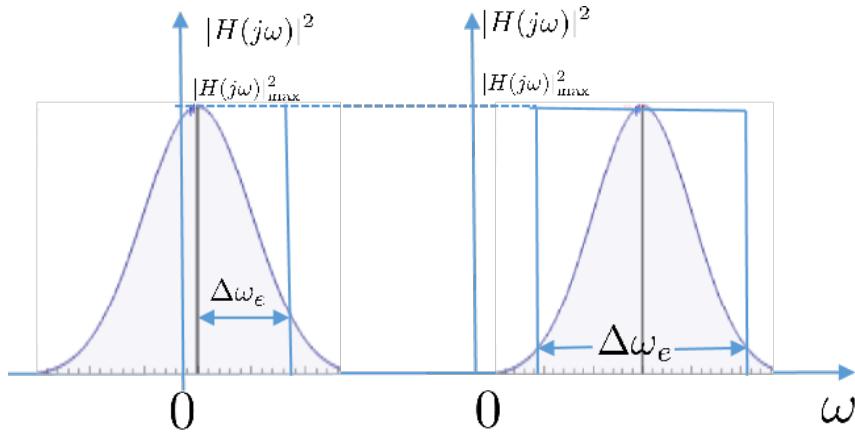


图 4-6 系统的等效噪声带宽 a) 低通系统 b) 带通系统

$$\begin{aligned} R_Y(0) &= \frac{N_0}{2\pi} \int_0^{+\infty} |H_e(\omega)|^2 d\omega = \frac{N_0}{2\pi} \Delta\omega_e |H(\omega)|_{\max}^2 \\ &= \frac{N_0}{2\pi} \int_0^{+\infty} |H(\omega)|^2 d\omega. \end{aligned} \quad (4.117)$$

解得

$$\Delta\omega_e = \frac{\int_0^{+\infty} |H(\omega)|^2 d\omega}{|H(\omega)|_{\max}^2}. \quad (4.118)$$

此外，根据 (4.113) 式和 (4.114) 式，我们可以通过用白噪声激励一个线性时不变系统的方法，来产生具有特定二阶统计特性的平稳随机过程，这是个信号建模问题。但由于输出的功率谱密度只提供了系统的幅度信息，缺少相位信息，我们只能据此得到最小相位系统。反过来，如果知道了平稳随机过程的二阶统计特性，将其激励一个线性时不变系统来产生一个白噪声信号，则称之为对随机过程进行预白化。由自相关函数 (4.115) 式和功率谱函数 (4.116) 式，我们还可以看到：系统输出与输入之间的互相关函数（或互功率谱密度）等于系统的冲激响应（或传递函数）乘上一个常数。也就是说，当白噪声激励一个未知的线性时不变系统时，可以通过估计系统输出与白噪声之间的互相关函数（或互功率谱密度）来估计系统的冲激响应（或传输函数），这是一个系统辨识问题。

## 4.5.2 白噪声通过理想低通系统

图 4-7 所示为一个理想的低通系统，系统的幅频响应为

$$|H(\omega)| = \begin{cases} A, & |\omega| \leq \Delta\omega \\ 0, & |\omega| > \Delta\omega \end{cases} \quad (4.119)$$

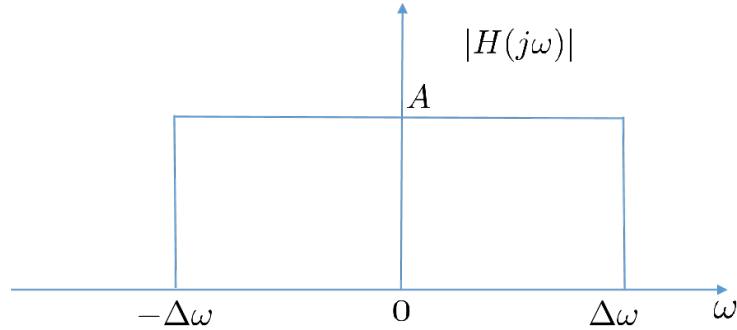


图 4-7 系统的等效噪声带宽

当白噪声通过该系统，输出的功率谱密度函数和自相关函数可以分别表示

$$G_Y(\omega) = \begin{cases} \frac{N_0 A^2}{2}, & |\omega| \leq \Delta\omega \\ 0, & |\omega| > \Delta\omega \end{cases}. \quad (4.120)$$

$$\begin{aligned} R_Y(\tau) &= \frac{N_0 A^2}{2} \cdot \frac{1}{2\pi} \int_{-\Delta\omega}^{\Delta\omega} e^{j\omega\tau} d\omega = \frac{N_0 A^2}{2\pi} \int_0^{\Delta\omega} \cos \omega\tau d\omega \\ &= \frac{N_0 A^2 \Delta\omega}{2\pi} \cdot \frac{\sin \Delta\omega\tau}{\Delta\omega\tau} = R_Y(0) \cdot \frac{\sin \Delta\omega\tau}{\Delta\omega\tau}. \end{aligned} \quad (4.121)$$

此时系统的等效噪声带宽为

$$\Delta\omega_e = \frac{1}{A^2} \int_0^{\Delta\omega} A^2 d\omega = \Delta\omega. \quad (4.122)$$

此外，根据 (4.121) 式可以进一步得到如下结论：当  $\tau = k\pi/\Delta\omega$  ( $k = \pm 1, \pm 2, \dots$ ) 时， $R_Y(\tau) = 0$ ，即白噪声通过理想低通系统后，输出在  $\pi/\Delta\omega$  间隔上抽样，得到的各个样本之间是互不相关的。

### 4.5.3 白噪声通过理想带通系统

图 4-8 所示为一个理想的带通系统，系统的幅频响应为

$$|H(\omega)| = \begin{cases} A, & |\omega \pm \omega_0| \leq \frac{\Delta\omega}{2} \\ 0, & |\omega \pm \omega_0| > \frac{\Delta\omega}{2} \end{cases} \quad (4.123)$$

当白噪声通过该系统，系统输出的功率谱密度函数  $G_Y(\omega)$  和自相关函数  $R_Y(\tau)$  可

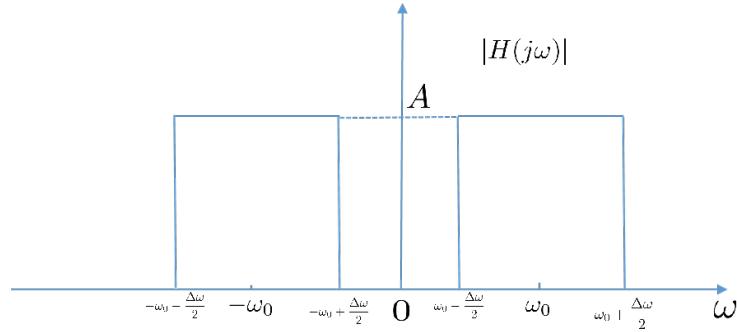


图 4-8 理想的带通系统

以分别表示为：

$$\begin{aligned}
 R_Y(\tau) &= \frac{N_0 A^2}{2\pi} \int_{\omega_0 - \frac{\Delta\omega}{2}}^{\omega_0 + \frac{\Delta\omega}{2}} \cos \omega \tau d\omega \\
 &= \frac{N_0 A^2}{2\pi\tau} \sin \left( \omega_0 + \frac{\Delta\omega}{2} \right) \tau - \frac{N_0 A^2}{2\pi\tau} \sin \left( \omega_0 - \frac{\Delta\omega}{2} \right) \tau \\
 &= \frac{N_0 A^2}{\pi\tau} \sin \frac{\Delta\omega\tau}{2} \cos \omega_0\tau \\
 &= \frac{N_0 A^2 \Delta\omega}{2\pi} \cdot \frac{\sin \frac{\Delta\omega\tau}{2}}{\frac{\Delta\omega\tau}{2}} \cos \omega_0\tau \\
 &= R_Y(0) \cdot \frac{\sin \frac{\Delta\omega\tau}{2}}{\frac{\Delta\omega\tau}{2}} \cos \omega_0\tau. \tag{4.124}
 \end{aligned}$$

$$G_Y(\omega) = \begin{cases} \frac{N_0 A^2}{2}, & |\omega \pm \omega_0| \leq \frac{\Delta\omega}{2} \\ 0, & |\omega \pm \omega_0| > \frac{\Delta\omega}{2} \end{cases}. \tag{4.125}$$

此时，系统的等效噪声带宽为

$$\Delta\omega_e = \frac{1}{A^2} \int_{\omega_0 - \frac{\Delta\omega}{2}}^{\omega_0 + \frac{\Delta\omega}{2}} A^2 d\omega = \Delta\omega. \tag{4.126}$$

#### 4.5.4 白噪声通过具有高斯频率特性的带通系统

由多级单调谐中频放大器级联构成的系统，其幅频特性近似具有高斯特性。图 4-9 所示的是一个具有高斯频率特性的带通系统，系统的幅频响应为

$$|H(\omega)| = \begin{cases} A \exp \left[ -\frac{(\omega - \omega_0)^2}{2\beta^2} \right], & \omega > 0. \\ A \exp \left[ -\frac{(\omega + \omega_0)^2}{2\beta^2} \right], & \omega < 0. \end{cases} \tag{4.127}$$

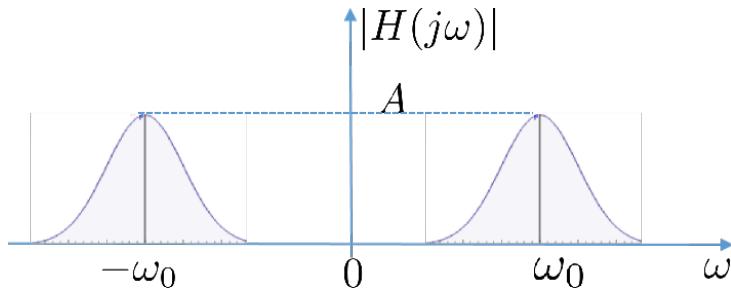


图 4-9 具有高斯频率特性的带通系统

当白噪声通过该系统，输出的功率谱密度函数为

$$G_Y(\omega) = \begin{cases} \frac{N_0 A^2}{2} \exp\left[-\frac{(\omega-\omega_0)^2}{\beta^2}\right], & \omega > 0 \\ \frac{N_0 A^2}{2} \exp\left[-\frac{(\omega+\omega_0)^2}{\beta^2}\right], & \omega < 0 \end{cases}. \quad (4.128)$$

如果将  $G_Y(\omega)$  看作是基带信号  $s_1(\tau)$  经过上变频后的带通信号，即

$$s_1(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} N_0 A^2 \exp\left(-\frac{\omega^2}{\beta^2}\right) e^{j\omega\tau} d\omega. \quad (4.129)$$

此时输出的自相关函数可表示为

$$R_Y(\tau) = s_1(\tau) \cos \omega_0 \tau = \frac{N_0 A^2}{2\pi} \int_{-\infty}^{+\infty} \exp\left(-\frac{\omega^2}{\beta^2}\right) e^{j\omega\tau} \cos \omega_0 \tau d\omega. \quad (4.130)$$

利用  $\exp(-a^2 t^2)$  和  $\frac{\sqrt{\pi}}{a} \exp\left(-\frac{\omega^2}{4a^2}\right)$  为傅里叶变换对，也即

$$\exp(-a^2 t^2) \xrightarrow[F]{F^{-1}} \frac{\sqrt{\pi}}{a} \exp\left(-\frac{\omega^2}{4a^2}\right) \quad (4.131)$$

则

$$\begin{aligned} R_Y(\tau) &= \frac{N_0 A^2}{2\pi} \beta \sqrt{\pi} \exp\left(-\frac{\beta^2 \tau^2}{4}\right) \cos \omega_0 \tau \\ &= \frac{N_0 A^2 \beta}{2\sqrt{\pi}} \exp\left(-\frac{\beta^2 \tau^2}{4}\right) \cos \omega_0 \tau. \end{aligned} \quad (4.132)$$

此时，系统的等效噪声带宽为

$$\Delta\omega_e = \frac{1}{A^2} \int_0^{+\infty} A^2 \exp\left[-\frac{(\omega-\omega_0)^2}{\beta^2}\right] d\omega = \sqrt{\pi} \beta. \quad (4.133)$$

## 4.6 白噪声序列和平稳随机序列的参数模型 \*

与连续的白噪声相对应, 白噪声序列  $\{W(n)\}$  是具有如下统计特性的零均值广义平稳随机序列:

$$\begin{aligned} R_W(m) &= \sigma_W^2 \delta(m). \\ S_W(\Omega) &= \sigma_W^2, \quad -\pi \leq \Omega \leq \pi. \end{aligned} \quad (4.134)$$

当白噪声序列通过离散线性时不变系统时, 根据前面章节中的结论, 输出噪声序列是广义平稳的, 且与输入序列联合平稳, 其具体统计特性可参照式(4.116)。与连续的白噪声相类似, 我们也可以用白噪声序列激励一个离散的最小相位系统进行信号建模, 产生具有特定二阶统计特性的平稳随机序列。

在实际应用中, 为了便于对问题的分析和处理, 线性时不变系统的传递函数通常用一个有限阶数的有理函数来近似表示。相应地, 系统的输出序列也就可以用有限个参数来描述, 该有理函数称为平稳随机序列的参数模型。这种通过建立合适的参数模型来分析平稳随机序列的方法, 因其频谱分辨率高, 已成为现代谱分析中的一种重要方法。这里只介绍三种最常用的参数模型。

考虑具有如下有限阶数的离散线性时不变系统

$$H(\Omega) = \frac{1 + \sum_{k=1}^q b_k e^{-j\Omega k}}{1 + \sum_{k=1}^p a_k e^{-j\Omega k}} = \frac{\sum_{k=0}^q b_k e^{-j\Omega k}}{\sum_{k=0}^p a_k e^{j\Omega k}}. \quad (4.135)$$

其中  $a_0 = b_0 = 1$ , 其他参数由输出序列  $\{Y(n)\}$  的二阶统计特性 (自相关函数或者功率谱函数) 决定. 写成  $Z$  变换的形式

$$H(z) = \frac{1 + \sum_{k=1}^q b_k z^{-k}}{1 + \sum_{k=1}^p a_k z^{-k}} = \frac{\prod_{k=1}^q (1 - \beta_k z^{-1})}{\prod_{k=1}^p (1 - \alpha_k z^{-1})}, \quad (4.136)$$

其中  $\alpha_k, \beta_k$  分别为系统的极点和零点, 为保证系统的因果稳定性, 这里要求  $|\alpha_k| < 1$ .

### 4.6.1 自回归滑动平均模型

如果系数  $a_1, \dots, a_p$  和  $b_1, \dots, b_q$  不全为零, 且  $a_p \neq 0, b_q \neq 0$  时, 此时系统存在零点也存在极点。当白噪声序列  $\{w(n)\}$  激励上述系统后, 系统输出序列  $\{Y(n)\}$  与输入  $\{w(n)\}$  之间的关系可由如下差分方程表示

$$\begin{aligned} y(n) &= -a_1 y(n-1) - a_2 y(n-2) - \cdots - a_p y(n-p) \\ &\quad + w(n) + b_1 w(n-1) + b_2 w(n-2) + \cdots + b_q w(n-q) \\ &= \sum_{k=0}^q b_k w(n-k) - \sum_{k=1}^p a_k y(n-k). \end{aligned} \quad (4.137)$$

称为平稳随机序列  $\{Y(n)\}$  的  $p$  阶自回归  $q$  阶滑动平均模型 (ARMA). 由该模型表征的平稳随机序列  $\{Y(n)\}$ , 也即系统输出的功率谱密度函数和自相关函数分别为

$$G_Y(\Omega) = \sigma_w^2 \frac{\left| \sum_{k=0}^q b_k e^{-j\Omega k} \right|^2}{\left| \sum_{k=0}^p a_k e^{-j\Omega k} \right|^2}. \quad (4.138)$$

$$\begin{aligned} R_Y(m) &= R_Y(n_1 - n_2) = E \left\{ y(n_2) \left[ \sum_{k=0}^q b_k w(n_1 - k) - \sum_{k=1}^p a_k y(n_1 - k) \right] \right\} \\ &= \sum_{k=0}^q b_k R_{wy}(m - k) - \sum_{k=1}^p a_k R_Y(m - k) \\ &= \sum_{k=0}^q b_k \sigma_w^2 h(k - m) - \sum_{k=1}^p a_k R_Y(m - k). \end{aligned} \quad (4.139)$$

上式利用了 4.3 节中的结论

$$R_{wy}(m) = \sum_{k=-\infty}^{+\infty} h(k) R_w(m+k) = \sigma_w^2 h(-m), \quad (4.140)$$

其中, 系统的冲激响应  $h(n)$  可根据 (4.135) 式表示成下式

$$h(n) = \begin{cases} -\sum_{k=1}^p a_k h(n-k) + b_n, & 0 \leq n \leq q \\ -\sum_{k=1}^p a_k h(n-k), & n > q \\ 0, & n < 0 \end{cases}. \quad (4.141)$$

## 4.6.2 自回归模型

当系数  $b_1, \dots, b_q$  均为零且  $a_p \neq 0$  时, 系统为全极点系统, 其传输函数和冲激响应分别为

$$H(\Omega) = \frac{1}{1 + \sum_{k=1}^p \Omega_k e^{-j\Omega k}} = \frac{1}{\sum_{k=0}^p \Omega_k e^{-j\Omega k}}. \quad (4.142)$$

$$h(n) = \begin{cases} -\sum_{k=1}^p a_k h(n-k) + \delta(n), & n \geq 0 \\ 0, & n < 0 \end{cases}. \quad (4.143)$$

系统输出序列  $\{Y(n)\}$  与白噪声序列  $\{W(n)\}$  的当前值和  $\{Y(n)\}$  的  $p$  个过去值有关, 即

$$y(n) = w(n) - \sum_{k=1}^p a_k y(n-k), \quad (4.144)$$

称为平稳随机序列  $\{Y(n)\}$  的  $p$  阶自回归模型 (Auto Regressive model, AR), 记为  $AR(p)$ . 可见  $AR(p)$  模型是由  $ARMA(p, q)$  模型经过  $q = 0$  退化而来。相应地, 由该模型表征的平稳随机序列  $\{Y(n)\}$ , 即系统输出的功率谱密度函数和自相关函数可表示为

$$G_Y(\Omega) = \frac{\sigma_w^2}{\left| \sum_{k=0}^p a_k e^{-j\Omega k} \right|^2}. \quad (4.145)$$

$$\begin{aligned} R_Y(m) &= R_Y(n_1 - n_2) = E \left\{ y(n_2) \left[ w(n_1) - \sum_{k=1}^p a_k y(n_1 - k) \right] \right\} \\ &= R_{my}(m) - \sum_{k=1}^p a_k R_Y(m - k) \\ &= \sigma_m^2 h(-m) - \sum_{k=1}^p a_k R_Y(m - k). \end{aligned} \quad (4.146)$$

### 4.6.3 滑动平均模型

当系数  $a_1, \dots, a_p$  均为零, 且  $b_q \neq 0$  时, 系统为全零点系统, 其传输函数和冲激响应分别为

$$\begin{aligned} H(\Omega) &= 1 + \sum_{k=1}^q b_k e^{-j\Omega k} = \sum_{k=0}^q b_k e^{j\Omega k}. \\ h(n) &= \begin{cases} b_n, & 0 \leq n \leq q \\ 0, & n < 0, n > q \end{cases}. \end{aligned} \quad (4.147)$$

此时,  $\{Y(n)\}$  仅与白噪声序列  $\{W(n)\}$  的当前值和  $q$  个过去值有关, 即

$$y(n) = w(n) + \sum_{k=1}^q b_k w(n - k) = \sum_{k=0}^q b_k w(n - k), \quad (4.148)$$

称为平稳随机序列  $\{Y(n)\}$  的  $q$  阶滑动平均模型 (Moving Average model, MA), 记为  $MA(q)$ . 同样,  $MA(q)$  模型可以看作是由  $ARMA(p, q)$  模型经过  $p = 0$  退化而来, 相应地, 由该模型表征的平稳随机序列  $\{Y(n)\}$ , 即系统输出的功率谱密度函数和自相关函数可分别表示为

$$G_Y(\Omega) = \sigma_W^2 \left| \sum_{k=0}^q b_k e^{-j\Omega k} \right|^2. \quad (4.149)$$

$$\begin{aligned} R_Y(m) &= R_Y(n_1 - n_2) = E \left\{ Y(n_2) \sum_{k=0}^q b_k W(n_1 - k) \right\} \\ &= \sum_{k=0}^q b_k R_{wy}(m-k) = \sum_{k=0}^q b_k \sigma_W^2 h(k-m). \end{aligned} \quad (4.150)$$

#### 4.6.4 三种模型间的联系

在上节，我们提到 AR 模型和 MA 模型是 ARMA 模型分别在不同条件下的特例，除此之外，这三种模型之间还可以相互等价。下面我们就以 MA 模型为例，证明一个有限阶的 MA( $q$ ) 模型可以等价于一个无穷阶的 AR 模型。

对 (4.148) 式两边取 Z 变换

$$Y(z) = H(z)W(z) = \prod_{k=1}^q (1 - \beta_k z^{-1}) W(z), \quad (4.151)$$

其中  $Y(z)$  和  $W(z)$  分别为  $y(n)$  和  $W(n)$  的 Z 变换。若系统  $H(z)$  具有可逆性，即其零点  $|\beta_k| < 1$ ，序列  $\{W(n)\}$  可以看成是序列  $\{Y(n)\}$  激励其逆系统  $H_{inv}$  的输出，即

$$W(z) = H_{inv}Y(z) = \frac{Y(z)}{\prod_{k=1}^q (1 - \beta_k z^{-1})}, \quad (4.152)$$

利用系统的稳定性条件和下面的公式

$$\frac{1}{1 - \beta_k z^{-1}} = \sum_{i=0}^{+\infty} \beta_k^i z^{-i} = 1 + \sum_{i=1}^{+\infty} \beta_k^i z^{-i}, \quad (4.153)$$

将其代入 (4.152) 式得

$$W(z) = \prod_{k=1}^q \left( 1 + \sum_{i=1}^{+\infty} \beta_k^i z^{-i} \right) Y(z) = \left( 1 + \sum_{i=1}^{+\infty} \varepsilon_i z^{-i} \right) Y(z). \quad (4.154)$$

比较上式与 (4.144) 式，两者在形式上完全一致，这就证明，一个可逆的有限阶 MA 模型可以等价于一个无穷阶的 AR 模型。同样，可以用类似的分析方法得到如下结论：一个可逆的有限阶 MA 模型或 ARMA 模型可以等价于一个无穷阶的 AR 模型，一个稳定的有限阶 AR 模型或 ARMA 模型也可以等价于一个无穷阶的 MA 模型。

## 小结

本章首先介绍了信号与系统的基本概念、分类和特点；其次分别从时域和频域两方面分析了平稳随机信号通过线性时不变系统后，系统输出与输入的一阶、二阶统计量之间的关系，得到了一系列的结论；然后将其应用到对白噪声这种特殊的随机过程通过线性系统后的分析，给出了等效噪声带宽的概念；紧接着利用相同的方法，分析了随机序列通过线性时不变系统后的特性分析，得到了类似的一些结论；并在其基础上，引入了通过建立随机序列参数模型来分析平稳随机序列统计特性的方法，介绍了 ARMA、AR、MA 三种模型的概念及其之间的联系；然而，针对非平稳随机信号以及有记忆的非线性等更复杂的系统的处理，限于篇幅没有涉及，感兴趣的读者可查阅相关的文献和书籍。

## 4.7 习题

✎ **练习 4.1** 已知系统的单位冲激响应  $h(t) = 5e^{-3t}U(t)$ ，输入信号为  $X(t) = M + 4\cos(2t + \Phi)$ ，其中  $M$  是随机变量， $\Phi$  是  $(0, 2\pi)$  上均匀分布的随机变量，且  $M$  和  $\Phi$  独立。求输出信号的表达式。

✎ **练习 4.2** 已知线性系统的单位冲激响应

$$h(t) = [5\delta(t) + 3][U(t) - U(t - 1)]. \quad (4.155)$$

输入信号为  $X(t) = 4\sin(2\pi t + \Phi)$ ，其中  $\Phi$  是  $(0, 2\pi)$  上均匀分布的随机变量。求输出信号的表达式均值和方差。

✎ **练习 4.3** 已知系统的单位冲激响应为

- (1)  $h(t) = e^{-2t}U(t);$
- (2)  $h(t) = e^{-2t} \sin tU(t);$
- (3)  $h(t) = te^{-2t}U(t);$
- (4)  $h(t) = \delta(t) + e^{-3t}U(t).$

当输入平稳信号  $X(t)$  的自相关函数为  $R_X(\tau) = 4 + e^{-|\tau|}$  时，求系统输出的均值和方差。

✎ **练习 4.4** 已知有限时间积分系统的单位冲激响应为  $h(t) = (1-t)[U(t) - U(t - 0.5)]$ ，系统输入功率谱密度为  $10V^2/Hz$  的高斯白噪声，求系统输出的（总）平均功率，交流功率和输入输出互相关函数。

✎ **练习 4.5** 已知系统的单位冲激响应  $h(t) = (1-t)[U(t) - U(t - 1)]$ ，信号的自相关函数为  $R_X(\tau) = 2\delta(\tau) + 9$ ，求系统输出的直流功率和相关函数。

练习 4.6 已知系统的输入信号是物理谱密度为  $N_0$  的白噪声, 系统为一带通滤波器, 如图 4-10 所示。求系统输出的(总)噪声平均功率。

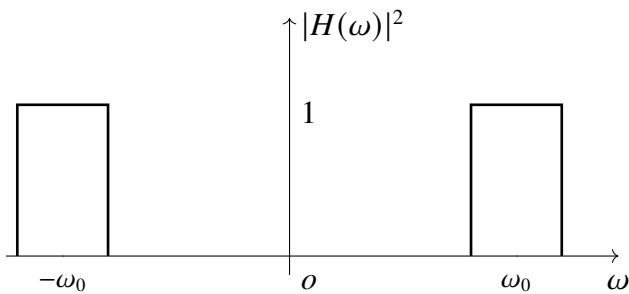


图 4-10

练习 4.7 已知如图 4-11 所示的线性系统, 系统输入信号是物理谱密度为  $N_0$  的白噪声, 求:  
① 系统的传递函数  $H(\omega)$ 。输出  $Z(t)$  的均方值, 其中

$$\int_0^\infty \frac{[\sin(ax)]^2}{x^2} dx = \int_0^\infty a^2 [\text{Sa}(ax)]^2 dx = \frac{\pi}{2}|a|. \quad (4.156)$$

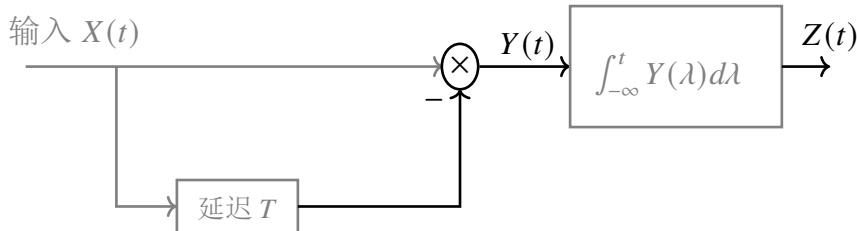


图 4-11

练习 4.8 已知某积分电路输入输出之间满足关系

$$Y(t) = \int_{t-T}^t X(t') dt'. \quad (4.157)$$

式中  $T$  是指积分时间常数。若输入和输出信号皆为平稳过程, 证明输出功率谱密度满足

$$G_Y(\omega) = G_X(\omega) \frac{[\sin(\frac{\omega T}{2})]^2}{(\frac{\omega}{2})^2}. \quad (4.158)$$

练习 4.9 如图 4-12 所示的某单输入双输出的线性系统

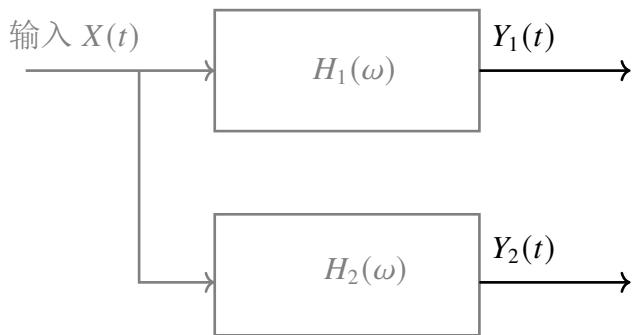


图 4-12

① 若输入信号  $X(t)$  宽平稳, 证明输出  $Y_1(t)$  和  $Y_2(t)$  的互谱密度满足

$$G_{Y_1 Y_2}(\omega) = G_X(\omega) H_1^*(\omega) H_2(\omega) \quad (4.159)$$

② 若输入信号为零均值平稳高斯过程, 为使得  $Y_1(t)$  和  $Y_2(t)$  相互独立, 求传递函数  $H_1(\omega)$  和  $H_2(\omega)$  应满足什么条件, 并画图说明。

✎ 练习 4.10 单位冲激响应  $h(t)$  如图 4-13 所示的线性系统, 若输入为功率谱密度为  $2V^2/Hz$  的白噪声, 求:

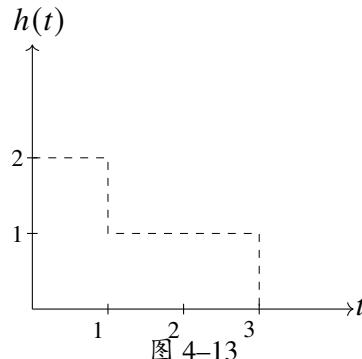


图 4-13

① 系统的等效噪声带宽。② 系统输出的平均功率。

✎ 练习 4.11 已知系统的输入为单位谱密度的白噪声, 输出的功率谱密度为

$$G_Y(\omega) = \frac{\omega^2 + 4}{\omega^4 + 10\omega^2 + 9}. \quad (4.160)$$

求此稳定系统的单位冲激响应  $h(t)$ 。

**练习 4.12** 已知系统输入信号的功率谱密度为设计一稳定的线性系统  $H(\omega)$ , 使得系统的输出为单位谱密度的白噪声。

**练习 4.13** 已知平稳随机信号  $X(t)$  在  $\tau = 0$  时刻加到传递函数  $H(s) = s^{-1}$  的零初始条件的积分器上。

① 若  $X(t)$  的自相关函数为  $R_X(\tau) = \sigma^2 e^{-\beta|\tau|}$ , 求输出  $Y(t)$  的均方值。

② 若  $X(t)$  的自相关函数为  $R_X(\tau) = \delta(\tau)$ , 求输出  $Y(t)$  的自相关函数。

**练习 4.14** 功率谱密度为  $N_0/2$  的白噪声作用于  $|H(0)| = 2$  的低通网络上, 等效噪声带宽为 2MHz。若在  $1\Omega$  电阻上的输出平均功率为 0.1W, 求  $N_0$  的值。

**练习 4.15** 已知线性系统的输入信号是功率谱密度为  $N_0/2$  的白噪声, 现用一等效系统的传递函数  $H_e(\omega)$  来代替原系统的传递函数  $H(\omega)$ , 等效原则为

$$H_e(\omega) = \begin{cases} H(\omega_0), & |\omega| \leq \Omega \\ 0, & \text{其他} \end{cases} \quad (4.161)$$

此时记系统的输出为  $Y(t)$ 。

① 求使得  $E[Y(t)] = E[Y^2(t)]$  时, 等效系统的带宽。

② 若已知  $H(\omega) = \frac{\beta}{j\omega + 2}$ , 求原系统  $H(\omega)$  的等效噪声带宽。

**练习 4.16** 已知一零均值平稳随机信号  $X(t)$  的自相关函数为  $R_X(\tau)$ , 相应的功率谱密度为  $G_X(\omega)$ , 且  $G_X(\omega) \leq G_X(0)$ 。若通过线性低通滤波器  $H(\omega)$ , 输出为  $Y(t)$ , 且系统  $H(\omega)$  的 3dB 带宽小于  $X(t)$  的等效噪声带宽。证明:

(1)  $\tau_X B_X = \tau_Y B_Y$ , (2)  $\frac{\tau_Y}{\tau_X} = \frac{\sigma_X^2}{\sigma_Y^2} |H(0)|^2$ , 其中  $\tau_X, \tau_Y; B_X, B_Y; \sigma_X^2, \sigma_Y^2$  分别为输入信号和输出信号的相关时间, 等效噪声带宽和方差。

**练习 4.17** 已知线性系统的单位冲激响应为  $h(t)$ , 系统输入  $X(t)$  为零均值, 自相关函数  $R_X(\tau) = \delta(\tau)$  的平稳高斯过程。若要使  $X(t_1)$  和  $Y(t_1)$  相互独立,  $h(t)$  应满足什么条件?

**练习 4.18** 如图 4-14 所示的线性系统, 系统输入  $W(t)$  是零均值的白噪声, 且  $h(t) = e^{-t} U(t)$ 。

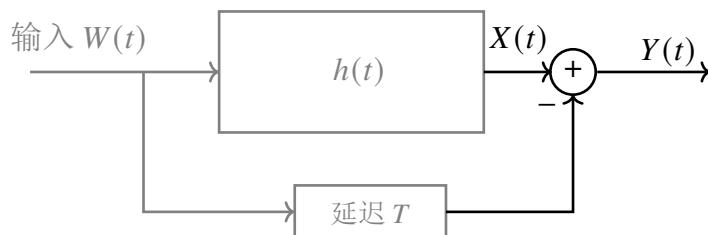


图 4-14

- ① 判断  $X(t)$  和  $Y(t)$  分别服从什么分布，给出理由。
- ② 证明  $Y(t)$  是严平稳过程。
- ③ 求  $W(t)$  和  $X(t)$  的互相关函数， $Y(t)$  的功率谱密度。
- ④ 写出  $Y(t)$  的一维概率密度表达式。
- ⑤ 判断同一时刻  $X(t)$  和  $Y(t)$  是否独立，给出理由。



# 5

## 随机信号统计特征的实验研究方法

### 5.1 统计特征实验研究的基础

#### 1. 随机信号的采样定理

##### 定义 5.1 低通带限信号

如果平稳随机信号  $X(t)$  的功率谱  $G_X(\omega)$  满足:  $G_X(\omega) = 0, |\omega| \geq \Omega_C$ , 则称  $X(t)$  为“低通带限”随机信号, 式中  $\Omega_C$  表示功率谱的最高截止频率。

若以采样间隔  $T$  对“低通带限”的平稳随机信号  $X(t)$  进行采样, 采样后获得随机序列  $\{X(n)\}$ 。只要采样频率  $f_s$  满足

$$\omega_s = 2\pi f_s \geq 2\Omega_c \quad \text{或} \quad T_s \leq \frac{1}{2f_c} = \frac{\pi}{\Omega_c} \quad (5.1)$$

采样插值为

$$\hat{X}(t) = \sum_{n=-\infty}^{\infty} X(nT_s) \frac{\sin \Omega_c (t - nT_s)}{\Omega_c (t - nT_s)}. \quad (5.2)$$

则可以证明

$$E [|X(t) - \hat{X}(t)|^2] = 0. \quad (5.3)$$

表明  $\hat{X}(t)$  在均方意义上等于  $X(t)$ 。

#### 2. 随机信号估计的准则

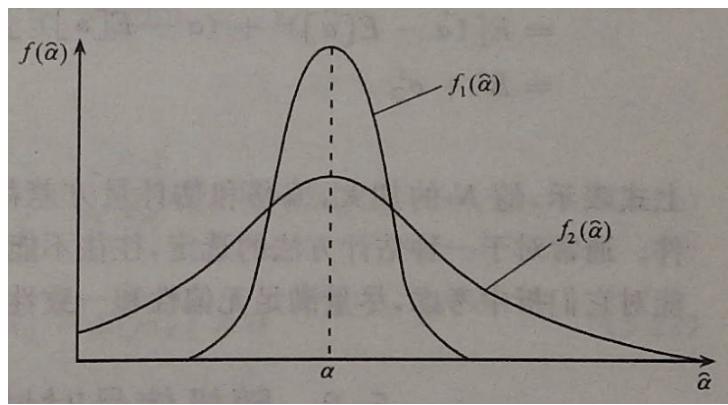


图 5-1 截取函数

统计实验研究的理论是建立在假设随机信号具有各态历经性基础上的。在这个假设的前提下，方可根据一个样本（实现）来“估计”出整个随机信号的有关统计参数。即根据对随机信号  $X(t)$  采样得到的随机序列  $\{X(n)\}$  的一组样本  $x_0, x_1, x_2, \dots, x_{N-1}$ ，来估计  $X(t)$  的统计参数。

一般来说参数的点估计，就是根据观测到的数据  $x_0, x_1, x_2, \dots, x_{N-1}$ ，构造一个用来估计未知参数  $\alpha$  的估计量  $\hat{\alpha}_a$

$$\hat{\alpha} = h(x_0, x_1, x_2, \dots, x_{N-1}). \quad (5.4)$$

由于观测到的数据  $x_0, x_1, x_2, \dots, x_{N-1}$  是随机的，因此，由观测数据构造的估计量  $\hat{\alpha}$  也是随机的，但被估计的参数  $\alpha$  是确定量，因此构造的估计量  $\hat{\alpha}$  与真实参数  $\alpha$  之间存在随机的估计误差，这个误差的大小则是判定估计方法好坏的标准。

估计误差用  $\tilde{\alpha}$  表示

$$\tilde{\alpha} = \hat{\alpha} - \alpha. \quad (5.5)$$

$\tilde{\alpha}$  和  $\hat{\alpha}$  都是随机变量，而随机变量则存在一定的统计分布规律。设  $\hat{\alpha}$  的概率密度曲线如图 5-1 所示，图中  $\alpha$  是要估计的参数。若估计值  $\hat{\alpha}$  接近  $\alpha$  的概率很大，则说明这是一种比较好的估计方法，如图中  $f_1(\hat{\alpha})$  比  $f_2(\hat{\alpha})$  要好。一般来说，一个好的估计值  $\hat{\alpha}$  的概率密度函数分布应比较集中在其估计值的真值  $\alpha$  附近，概率密度曲线必须窄。

通常，评价估计性能好坏的标准有以下三种：

1) 偏移性。令估计量的统计平均值与真值之间的差值为偏移  $B$ ，其公式为

$$B = \alpha - E[\hat{\alpha}]. \quad (5.6)$$

如果  $B = 0$ , 称为无偏估计。无偏估计表示估计量仅在它真值附近摆动, 这是我们希望有的估计特性。如果  $B \neq 0$  则称为有偏估计。如果随着观测次数  $N$  的加大, 能满足

$$\lim_{N \rightarrow \infty} E[\hat{\alpha}] = \alpha \quad (5.7)$$

则称为渐近无偏估计。

2) 估计量的方差。如果两个估计量的观察次数相同, 又都是无偏估计, 哪一个估计量在真值附近摆动更小一些, 即估计量的方差更小一些, 就说这一个估计量的估计更有效。比如  $\hat{\alpha}$  和  $\hat{\alpha}'$  是  $X$  的两个无偏估计量

$$\sigma_{\hat{\alpha}}^2 = E[(\hat{\alpha} - E[\hat{\alpha}])^2] \quad \sigma_{\hat{\alpha}'}^2 = E[(\hat{\alpha}' - E[\hat{\alpha}'])^2]. \quad (5.8)$$

若对任意  $N$ , 它们的方差满足

$$\sigma_{\hat{\alpha}}^2 < \sigma_{\hat{\alpha}'}^2 \quad (5.9)$$

则称  $\hat{\alpha}$  比  $\hat{\alpha}'$  更有效。一般希望当  $N \rightarrow \infty$  时,  $\sigma_{\hat{\alpha}}^2 \rightarrow 0$ 。

3) 一致性——均方误差。

在许多情况下, 比较两个有偏估计值是麻烦的。偏移较小的估计值, 可能有较大的方差, 而方差较小的估计值可能有较大的偏移, 此时使用与估计值有关的均方误差会更方便。估计量的均方误差用下式表示

$$E[\tilde{\alpha}^2] = E[(\hat{\alpha} - \alpha)^2]. \quad (5.10)$$

如果估计量的均方误差随着观察次数的增加趋于 0, 即估计量随着  $N$  的加大, 在均方意义上趋于它的真值, 则称该估计是一致估计。

估计量的均方误差、估计量的方差和偏移的关系, 推导如下

$$\begin{aligned} E[\tilde{\alpha}^2] &= E[(\hat{\alpha} - \alpha)^2] = E\{[(\hat{\alpha} - E[\hat{\alpha}]) - (\alpha - E[\hat{\alpha}])]^2\} \\ &= E[(\hat{\alpha} - E[\hat{\alpha}])^2 + (\alpha - E[\hat{\alpha}])^2] - 2E[(\hat{\alpha} - E[\hat{\alpha}]) \cdot (\alpha - E[\hat{\alpha}])] \\ &= B^2 + \sigma_{\hat{\alpha}}^2. \end{aligned} \quad (5.11)$$

尽量满足无偏性和一致性。

## 5.2 随机信号时域特征的估计

1) 最大似然估计

由于各态历经性的假设, 实验测得的一组样本  $x_0, x_1, x_2, \dots, x_{N-1}$  可看作是随机序列  $\{X(n)\}$  中的一组随机变量。定义在参量  $\alpha$  条件下  $\{X(n)\}$  的多维条件概率密度  $f(x_0, x_1, \dots, x_{N-1}/\alpha)$  为“似然函数”。若估计量  $\hat{\alpha}$  能使似然函数  $f(x_0, x_1, \dots, x_{N-1}/\alpha)$  在  $\alpha = \hat{\alpha}$  时

$$f(x_0, x_1, \dots, x_{N-1}/\hat{\alpha}) = \max \{f(x_0, x_1, \dots, x_{N-1}/\alpha)\}, \quad (5.12)$$

则估计量  $\hat{\alpha}$  被称为最大似然估计量。

## 2) 均值的估计

### (1) 均值的最大似然估计

若对高斯信号  $X(t)$  的采样是独立的, 则  $\{X(n)\}$  就是独立高斯随机序列, 实验测得的一组样本  $x_0, x_1, x_2, \dots, x_{N-1}$  也可看作是相互独立的高斯随机变量。那么可得均值的最大似然估计量为

$$\hat{m}_X = \frac{1}{N} \sum_{i=0}^{N-1} x_i. \quad (5.13)$$

### (2) 估计量的评价

1°  $\hat{m}_X$  是无偏估计量。因为

$$E[\hat{m}_X] = E\left[\frac{1}{N} \sum_{i=0}^{N-1} x_i\right] = \frac{1}{N} \sum_{i=0}^{N-1} E[x_i] = E[x_i] = m_X. \quad (5.14)$$

所以

$$B = m_X - E[\hat{m}_X] = 0, \quad (5.15)$$

这个估计量是无偏估计量。

2°  $\hat{m}_X$  是一致估计量 (假设样本数据  $(x_0, x_1, \dots, x_{N-1})$  之间不存在相关性)。由

$$\begin{aligned} E[\hat{m}_X^2] &= \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} E[x_i x_j] \\ &= \frac{1}{N^2} \left[ \sum_{i=0}^{N-1} \sum_{j=0, j=i}^{N-1} E[x_i x_j] + \sum_{i=0}^{N-1} \sum_{j=0, j \neq i}^{N-1} E[x_i x_j] \right] \\ &= \frac{1}{N^2} \left[ \sum_{i=0}^{N-1} E[x_i^2] + \sum_{i=0}^{N-1} \sum_{j=0, j \neq i}^{N-1} E[x_i] E[x_j] \right] \\ &= \frac{1}{N^2} [N E[x_i^2] + N(N-1) E[x_i] E[x_i]] \\ &= \frac{1}{N} E[x_i^2] + \frac{N-1}{N} m_X^2. \end{aligned} \quad (5.16)$$

可得估计的方差  $\sigma_{m_X}^2$  为

$$\begin{aligned}\sigma_{m_X}^2 &= E[(\hat{m}_X - E[\hat{m}_X])^2] = E[\hat{m}_X^2] - E^2[\hat{m}_X] \\ &= E[\hat{m}_X^2] - m_X^2 = \frac{1}{N}E[x_i^2] - \frac{1}{N}m_X^2 = \frac{1}{N}\sigma_X^2.\end{aligned}\quad (5.17)$$

则其极限

$$\lim_{N \rightarrow \infty} \sigma_{m_X}^2 = \lim_{N \rightarrow \infty} \frac{1}{N} \sigma_X^2 = 0. \quad (5.18)$$

估计量的均方误差

$$E[\tilde{m}_X^2] = E[(\hat{m}_X - m_X)^2] = B^2 + \sigma_{m_X}^2 = 0 + \sigma_{m_X}^2. \quad (5.19)$$

则其极限

$$\lim_{N \rightarrow \infty} E[\tilde{m}_X^2] = \lim_{N \rightarrow \infty} \sigma_{m_X}^2 = 0. \quad (5.20)$$

所以该估计量也是一致估计量。

结论：当样本数据内部不相关时，按照式(5.13)估计均值是一种无偏的一致估计，是一种好的估计方法。但如果数据内部存在关联性，会使一致性的效果下降，估计量的方差比数据内部不存在相关情况的方差要大，达不到信号方差的  $1/N$ 。

### 3) 方差的估计

#### (1) 方差的最大似然估计

假设采样得的  $\{X(n)\}$  为独立的高斯随机序列，用实验测得的一组样本数据  $x_0, x_1, \dots, x_{N-1}$  来估计方差。解出方差的最大似然估计量为

1° 若均值  $m_X$  已知

$$\hat{\sigma}_X^2 = \frac{1}{N} \sum_{i=0}^{N-1} (x_i - m_X)^2. \quad (5.21)$$

2° 若均值  $m_X$  未知

$$\hat{\sigma}_X^2 = \frac{1}{N} \sum_{i=0}^{N-1} (x_i - \hat{m}_X)^2. \quad (5.22)$$

(2) 估计值的评价 (假设样本数据  $(x_0, x_1, \dots, x_{N-1})$  之间不存在相关性)

1°  $\hat{\sigma}_X^2$  是有偏估计量。

$$\begin{aligned} E[\hat{\sigma}_X^2] &= \frac{1}{N} \sum_{i=0}^{N-1} \left\langle E[x_i^2] + E[\hat{m}_X^2] - 2E[x_i \hat{m}_X] \right\rangle \\ &= \frac{1}{N} \sum_{i=0}^{N-1} E[x_i^2] + \frac{1}{N} \sum_{i=0}^{N-1} E[\hat{m}_X^2] - \frac{2}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} E[x_i x_j] \\ &= E[x_i^2] + E[\hat{m}_X^2] - 2E[\hat{m}_X^2]. \end{aligned} \quad (5.23)$$

代入式(5.16)结论可得

$$E[\hat{\sigma}_X^2] = \frac{N-1}{N} [E[x_i^2] - m_X^2] = \frac{N-1}{N} \sigma_X^2. \quad (5.24)$$

即  $\hat{\sigma}_X^2$  为有偏估计量。但当  $N \rightarrow \infty$  时,  $\hat{\sigma}_X^2$  为渐进无偏的。

另外, 为了得到无偏估计, 可以用下式计算:

$$\hat{\sigma}'_X^2 = \frac{1}{N-1} \sum_{i=0}^{N-1} (x_i - \hat{m}_X)^2. \quad (5.25)$$

$\hat{\sigma}_X^2$  和  $\hat{\sigma}'_X^2$  之间的关系为

$$\hat{\sigma}_X^2 = \frac{N}{N-1} \hat{\sigma}'_X^2. \quad (5.26)$$

$$E[\hat{\sigma}'_X^2] = \sigma_X^2. \quad (5.27)$$

所以,  $\sigma'_X^2$  是个无偏估计量。

2°  $\hat{\sigma}_X^2$  是个一致估计量。由  $E[\tilde{\alpha}^2] = E[(\hat{\alpha} - \alpha)^2] = B + \sigma_{\hat{\alpha}}^2$  知, 估计距  $\tilde{\sigma}_X^2$  的均方误差为

$$E[(\bar{\sigma}_X^2)^2] = E[(\tilde{\sigma}_X^2 - \sigma_X^2)^2] = B + D[\hat{\sigma}_X^2]. \quad (5.28)$$

其中

$$D[\tilde{\sigma}_X^2] = E[(\hat{\sigma}_X^2)^2] - (E[\hat{\sigma}_X^2])^2. \quad (5.29)$$

令

$$\hat{\sigma}_X^2 = \frac{1}{N} \sum_{i=0}^{N-1} (x_i - \hat{m}_X)^2 = \frac{1}{N} \sum_{i=0}^{N-1} (y_i)^2. \quad (5.30)$$

有

$$\begin{aligned}
 E[(\tilde{\sigma}_X^2)^2] &= \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (E[y_i^2 y_j^2]) \\
 &= \frac{1}{N^2} \left\{ N E[y_i^4] + N(N-1) (E[y_i^2])^2 \right\} \\
 &= \frac{1}{N} \left\{ E[y_i^4] + (N-1) (E[y_i^2])^2 \right\} \\
 E[\hat{\sigma}_X^2] &= E\left[\frac{1}{N} \sum_{i=0}^{N-1} y_i^2\right] = E[y_i^2]. \tag{5.31}
 \end{aligned}$$

则

$$D[\hat{\sigma}_{\tilde{X}}^2] = E[(\hat{\sigma}_X^2)^2] - (E[\hat{\sigma}_X^2])^2 = \frac{1}{N} \left\{ E[y_i^4] - (E[y_i^2])^2 \right\}. \tag{5.32}$$

估计量  $\hat{\sigma}_X^2$  均方误差为

$$E[(\tilde{\sigma}_X^2)^2] = E[(\hat{\sigma}_X^2 - \sigma_X^2)^2] = B + \frac{1}{N} \left\{ E[y_i^4] - (E[y_i^2])^2 \right\}. \tag{5.33}$$

因为其极限

$$\lim_{N \rightarrow \infty} E[(\tilde{\sigma}_X^2)^2] = \lim_{N \rightarrow \infty} B + \lim_{N \rightarrow \infty} \frac{1}{N} \left\{ E[y_i^4] - (E[y_i^2])^2 \right\} = 0. \tag{5.34}$$

所以  $\hat{\sigma}_X^2$  是个一致估计量。以上由关于  $\{X(n)\}$  为高斯分布的假设, 导出了均值和方差的最大似然估计量。当未知  $\{X(n)\}$  的密度函数形式, 或不是高斯分布时, 也常用式 (5.13) 和式 (5.21) 的估计方法。此时均值估计量仍为无偏一致估计量, 而方差估计量仍为渐近无偏一致估计量。但对有限样本来讲, 它们不再是最大似然估计, 从而不能保证是最佳的了。

#### 4) 自相关函数的估计

设对零均值平稳过程  $X(t)$  采样得随机序列  $\{X(n)\} (n = 0, \pm 1, \pm 2, \dots, \pm \infty)$ , 其自相关函数为

$$R_X(m) = E[X(n)X(n+m)]. \tag{5.35}$$

##### (1) 自相关函数估计量

由实验手段测得  $\{X(n)\}$  的一组样本数据  $x_0, x_1, x_2, \dots, x_{N-1}$  估计的方差  $\hat{\sigma}_X^2$ , 也就是零滞后自相关函数  $R_X(0)$  的估计量

$$\hat{R}_X(0) = \hat{\sigma}_X^2 = \frac{1}{N} \sum_{i=0}^{N-1} x_i^2 = \frac{x_0^2 + x_1^2 + \dots + x_{N-1}^2}{N}. \tag{5.36}$$

$$\hat{R}_X(m) = \begin{cases} \frac{1}{N} \sum_{i=0}^{N-|m|-1} x_i \cdot x_{i+m}, & m = 0, 1, \dots, N-1 \\ 0, & |m| \geq N \end{cases} \quad (5.37)$$

注意：

- (1) 当原过程  $X(t)$  为独立高斯过程时, 该估计量  $\hat{R}_X(0)$  为最大似然估计量。
- (2) 当  $\{X(n)\}$  中各  $X_i$  独立但非高斯时,  $\hat{R}_X(0)$  为渐近无偏一致估计量。
- (3) 当  $\{X(n)\}$  中各  $X_i$  之间不独立时, 往往也采用这样的估计量, 严格来说这时的估计量已不再是最佳的, 甚至可能不再是渐近无偏一致估计量了。

若将  $R_X(0)$  的估计方法推广至非零滞后自相关函数  $R_X(m)$  的估计

$$\begin{aligned} \hat{R}_X(m) &= \frac{x_0x_0 + |m| + x_1x_{1+|m|} + \dots + x_{N-1} - |m|x_{N-1-|m|+|m|}}{N} \\ &= \frac{1}{N} \sum_{i=0}^{N-|m|-1} x_i x_{i+|m|}, \quad m = 0, 1, \dots, N-1 \end{aligned} \quad (5.38)$$

估计量  $\hat{R}_X(m)$  求和的项数一般小于  $N$ . 特别是当  $m$  很大时, 参加求和的乘积项将非常少。当  $m \geq N$  时,  $x_i x_{i+|m|} = 0$ 。因而, 令  $m \geq N$  时的自相关函数为 0, 即

$$\hat{R}_X(m) = 0, \quad |m| \geq N. \quad (5.39)$$

因此, 自相关函数的估计量为

$$\hat{R}_X(m) = \begin{cases} \frac{1}{N} \sum_{i=0}^{N-|m|-1} x_i x_{i+|m|}, & m = 0, 1, \dots, N-1 \\ 0, & |m| \geq N. \end{cases} \quad (5.40)$$

(2) 估计量的评价

1°  $\hat{R}_X(m)$  是渐近无偏估计。由

$$E [\hat{R}_X(m)] = \frac{1}{N} \sum_{i=0}^{N-|m|-1} E [x_i x_{i+|m|}] = \left(1 - \frac{|m|}{N}\right) R_X(m). \quad (5.41)$$

其极限为

$$\lim_{N \rightarrow \infty} E [\hat{R}_X(m)] = \lim_{N \rightarrow \infty} \left|1 - \frac{|m|}{N}\right| R_X(m) = R_X(m). \quad (5.42)$$

偏移

$$B(m) = R_X(m) - E [\hat{R}_X(m)] = \frac{|m|}{N} R_X(m). \quad (5.43)$$

其极限为

$$\lim_{N \rightarrow \infty} B(m) = \lim_{N \rightarrow \infty} \frac{|m|}{N} R_X(m) = 0, \quad (5.44)$$

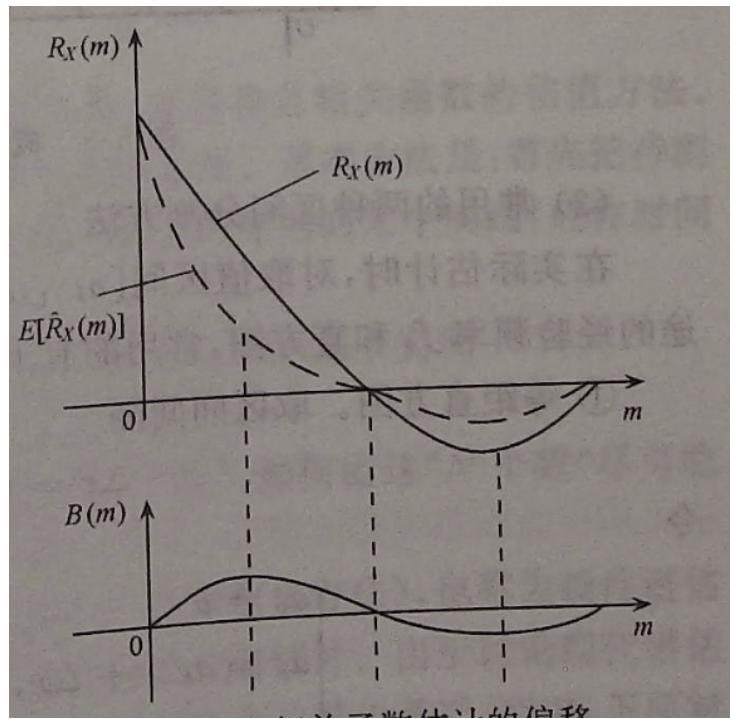


图 5-2 自相关函数估计的偏移

所以  $\hat{R}_X(m)$  是渐近无偏估计。偏移  $B(m) = \frac{|m|}{N} R_X(m)$  是  $m$  的函数, 如图5-2所示。

2°  $\hat{R}_X(m)$  是一致估计量。即需要证明, 当  $N \rightarrow \infty$  时,  $\hat{R}_X(m)$  的方差趋向 0。一般情况下这一结论的证明是比较复杂的。当原过程  $X(t)$  是高斯过程时, 可以导出下列方差的近似公式:

$$D[\hat{R}_X(m)] \approx \frac{2}{N} \sum_{i=0}^{N-|m|-1} \left(1 - \frac{|m|+i}{N}\right) [R_X^2(m) + R_X(i+m)R_X(i-m)]. \quad (5.45)$$

可以证明此式对一般的非高斯过程也是近似成立的。因为尽管  $N \rightarrow \infty$ , 但是当  $i \rightarrow \pm\infty$  时,  $R_X(i) \rightarrow 0$ , 即  $\sum$  号内仅有有限项存在, 则

$$\lim_{N \rightarrow \infty} D[\hat{R}_X(m)] = 0. \quad (5.46)$$

### 5. 随机信号概率密度的估计

多维概率密度的估计十分复杂, 这里仅讨论一种最简单的一维概率密度的估计方法——“直方图法”。这种方法的合理性是建立在“分布函数具有各态历经性(略)”的基础上的。

#### (1) 直方图法

若实验测得随机过程  $X(t)$  的一组样本数据  $x_0, x_1, x_2, \dots, x_{N-1}$ , 由这组样本数据来计算  $X(t)$  的“直方图”。

1) 首先求出其位置参量

极小值为

$$a = \min \{x_0, x_1, \dots, x_{N-1}\}. \quad (5.47)$$

极大值为

$$b = \max \{x_0, x_1, \dots, x_{N-1}\}. \quad (5.48)$$

(2) 分区间

将  $x$  的取值区间  $[a, b)$  分成  $K$  个互不相交的分区间, 取分割点为  $a_0 < a_1 < \dots < a_k < \dots < a_K$ , 且使  $a_0 \leq a, a_K > b$ 。则子区间分别为  $[a_{k-1}, a_k)$  ( $k = 1, 2, \dots, K$ )。

(3) 计算实验数据的经验频率

求出落入每个子区间  $[a_{k-1}, a_k)$  的数据点的个数  $N_k$  ( $k = 1, 2, \dots, K$ ) 即计算  $a_{k-1} \leq x_n < a_k$ ,  $n = 1, 2, \dots, K$  的实验数据的点数。 $N_k$  称经验频数, 则经验频率为

$$f_k = \frac{N_k}{N}, \quad k = 1, 2, \dots, K. \quad (5.49)$$

显然,  $0 \leq f_k \leq 1$ , 且  $\sum_{k=1}^K f_k = 1$ , 当  $N$  充分大时,  $f_k$  可以近似表示随机变量  $X$  在区间  $[a_{k-1}, a_k)$  上取值的概率。则概率密度在  $[a_{k-1}, a_k)$  上的估值  $\hat{f}(x)$  为

$$\hat{f}(x) = \frac{f_k}{a_k - a_{k-1}}, \quad k = 1, 2, \dots, K \quad x \in [a_{k-1}, a_k). \quad (5.50)$$

如图 5-3 所示。

### 5.3 随机信号功率谱密度的估计

前面, 我们已经讨论了有关随机序列的均值、方差和自相关函数的估值方法。本节将讨论概率谱密度的估计方法, 即功率谱的测量法。基本方法是: 首先把待测的随机过程通过取样量化 (A/D 变换) 变换成一个随机序列的某个实现 (或称时间序列) 后, 再对它进行功率谱估值。

谱估值的基本问题是: 已知随机过程  $X(t)$  或  $\{X(n)\}$  的某个实现

$$\dots, x_{-2}, x_{-1}, x_0, x_1, x_2, \dots, x_{N-1}, \dots \quad (5.51)$$

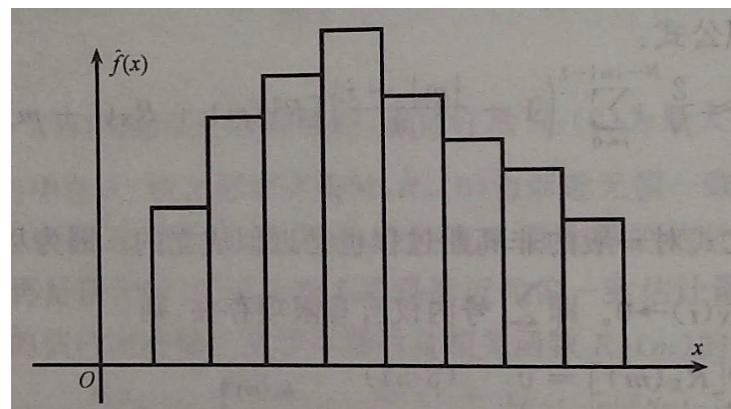


图 5-3 概率密度的直方图

中的有限长序列段  $x_n (0 \leq n \leq N - 1)$ , 或者说“ $N$  个数”, 如何由这“ $N$  个数”尽可能准确地得到  $X(t)$  或  $\{X(n)\}$  的功率谱密度  $G_X(\omega)$ 。

功率谱估计分为两大类:一类是经典谱估计(非参数化方法),也称为线性谱估计;另一类是现代谱估计(参数化方法),也称为非线性谱估计。由于讨论现代谱估计所需要具备的理论基础超出了学生学习本书时所具备的理论基础,所以,下面对常用的几种经典谱估计的方法作原理性介绍,而对现代谱估计仅作一般性介绍。有兴趣的读教可以参看有关班代谱估计的著作。

### 5.3.1 经典谱估计

功率谱估计分为两大类:

- ①一类是经典谱估计(非参数化方法),也称为线性谱估计;
- ②另一类是现代谱估计(参数化方法),也称为非线性谱估计。

下面仅对常用的两种经典谱估计的方法做原理性介绍,而对现代谱估计仅做一般性介绍。

### 5.3.2 经典谱估计

#### 1.BT 法 (Blackman-Tukey)

##### (1) 随机序列的维纳一辛钦定理

由于随机序列  $\{X(n)\}$  的自相关函数  $R_X(m) = E[X(n)X(n+m)]$  定义在离散点  $m$  上,

设取样间隔为  $T_s$ , 则可将随机序列的自相关函数可用连续时间函数表示为

$$R_X^D(\tau) = \sum_{m=-\infty}^{\infty} R_X(\tau) \cdot \delta(\tau - mT_s) = \sum_{m=-\infty}^{\infty} R_X(mT_s) \cdot \delta(\tau - mT_s). \quad (5.52)$$

等式两端取傅立叶变换, 则随机序列的功率谱密度

$$\begin{aligned} G_X(\omega) &= \int_{-\infty}^{\infty} R_X^D(\tau) e^{-j\omega\tau} d\tau \\ &= \sum_{m=-\infty}^{\infty} \int_{-\infty}^{\infty} R_X(mT_s) \delta(\tau - mT_s) e^{-j\omega mT_s} d\tau \\ &= \sum_{m=-\infty}^{\infty} R_X(mT_s) e^{-j\omega mT}. \end{aligned} \quad (5.53)$$

注意:  $G_X(\omega)$  为  $\omega$  的连续函数, 而且在频率轴上以  $2\pi/T_s$  周期重复。我们感兴趣的仅仅是它在所谓的奈奎斯特间隔  $(-\pi/T_s, \pi/T_s)$  上的值。因此, 已知  $G_X(\omega)$  而要求  $R_X(mT_s)$ , 仅需在一个周期  $(-\pi/T_s, \pi/T_s)$  内积分即可

$$R_X(mT_s) = \frac{T_s}{2\pi} \int_{-\frac{\pi}{T_s}}^{\frac{\pi}{T_s}} G_X(\omega) e^{j\omega mT_s} d\omega. \quad (5.54)$$

令  $T_s = 1$ , 则上述两式可写成

$$\begin{aligned} G_X(\omega) &= \sum_{m=-\infty}^{\infty} R_X(m) \cdot e^{-j\omega m} \\ R_X(m) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} G_X(\omega) e^{j\omega m} d\omega. \end{aligned} \quad (5.55)$$

以上两式为随机序列的维纳—辛钦定理。

## (2) 谱估计

*BT* 法是先估计自相关函数  $\widehat{R}_X(m)$  ( $|m| = 0, 1, 2, \dots, N-1$ ), 然后再经过离散傅立叶变换, 求得功率谱密度的估值  $\widehat{G}_X(\omega)$ 。即

$$\widehat{G}_X(\omega) = \sum_{m=-N-1}^{N-1} \widehat{R}_X(m) e^{-j\omega m}. \quad (5.56)$$

用 *BT* 法估计功率谱的流程为

$$\left\{ \frac{1}{N} \sum_{i=0}^{N-|m|-1} x_i \cdot x_{i+|m|} \right\} = \left\{ \widehat{R}_X(m) \right\} \xrightarrow{FFT} \left\{ \widehat{G}_X(\omega) \right\}, \quad (5.57)$$

其中 FFT 表示快速傅立叶变换。

## 2. 周期图法

周期图法是根据各态历经随机过程功率谱的定义来进行谱估计的。在第三章中已知，各态历经的连续随机过程的功率谱密度满足

$$G_X(\omega) = G_i(\omega) = \lim_{T \rightarrow \infty} \frac{1}{2T} |\mathcal{X}_{iT}(\omega)|^2, \quad (5.58)$$

式中  $\mathcal{X}_{iT}(\omega)$  是连续随机过程第  $i$  个样本的截取函数  $x_{iT}(t)$  的频谱。对应在随机序列中则有

$$G_X(\omega) = G_i(\omega) = \lim_{N \rightarrow \infty} \frac{1}{2N+1} |\mathcal{X}_N(\omega)|^2. \quad (5.59)$$

由于随机序列中观测数据  $x_n$  仅在  $0 \leq n \leq N-1$  的点上存在，则  $x_n$  的  $N$  点离散傅立叶变换 (DFT) 为

$$\mathcal{X}_N(\omega) = \sum_{n=0}^{N-1} x_n e^{-j\omega n}. \quad (5.60)$$

因此有随机信号的观测数据  $x_n$  的功率谱估计值称“周期图”如下

$$\widehat{G}_X(\omega) = \frac{1}{N} |\mathcal{X}_N(\omega)|^2 = \frac{1}{N} \left| \sum_{n=0}^{N-1} x_n e^{-j\omega n} \right|^2. \quad (5.61)$$

由于上式中的离散傅立叶变换 (DFT) 可以用快速傅立叶变换 (FFT) 计算，因此，用周期图法估计功率谱的流程如下

$$\{x_n\} \xrightarrow{\text{FFT}} \{x_N(\omega)\} \rightarrow \left\{ \frac{1}{N} |x_N(\omega)|^2 \right\} = \widehat{G}_X(\omega). \quad (5.62)$$

无论采用哪种方法，估计量  $\widehat{G}_X(\omega)$  的性质是一样的。可以证明它们均为有偏估值，但是渐进无偏的。直接采用上面两个公式的估值方法，最大的问题是这个估计量不是一致估计量，即当  $N$  很大时， $\widehat{G}_X(\omega)$  的方差也不小。理论分析可以证明

$$\lim_{N \rightarrow \infty} D \left[ \widehat{G}_X(\omega) \right] \approx G_X^2(\omega). \quad (5.63)$$

即功率谱估值的方差近似等于功率谱真值的平方。真实谱越大的地方（通常是我们于这两点结论的证明超出了本书讨论的范围，有兴趣的读者可以参看有关谱估计的经典著作。因而在实际使用时，必须将周期图法的谱估值加以修正，下面介绍两种卓有成效的修正方法。

### 5.3.3 经典谱估计的改进

从上面的分析可知, 周期图法作为功兹谱的估计不满足一致估计的条件, 必须进行改进。改进的措施主要是将周期图进行平滑(平均属于平滑的一种主要方法), 使估计方差减小, 从而得到一致谱估计。

平滑的方法主要有两种:

一种是常用的窗口处理法“窗函数法”选择适当的窗口函数作为加权函数, 对自相关函数估计值进行加权平均; 另一种是周期图“平均法”先将数据分段, 对各段数据分别求其周期图, 最后将所有段周期图的平均作为功摔谱的估计。由于“平均法”采用快速傅立叶变换(FFT), 运算速度快, 所以是当前用得较多的一种平滑方法。Welch 法是前两种方法的结合, 因此是用得最多的一种方法。

#### 1. 窗函数法

从前面的讨论得知, BT 法实质是对自相关函数的估值求傅里叶变换的一种方法。由于作自相关函数估计时, 滞后量  $m$  增大, 参与求和的项就会变少, 因而平均的效果弱。当  $m = N - 1$  时, 求和项只有一项, 也就是说,  $m$  较大的那些自相关估值  $\hat{R}_X(m)$  可靠性差, 估计的方差会较大。由于 BT 法用了从  $m = 0 \sim (N - 1)$  的全部自相关函数估值, 则方差大的自相关函数估值造成了谱估计较大的方差。如果对计的影响, 就有可能提高谱估计的性能。

为减少谱估计的方差, 经常用窗函数  $\tilde{\omega}(m)$  对自相关函数进行加权, 此时谱估计公式为

$$\hat{G}_x(\omega) = \sum_{m=-M+1}^{M-1} \hat{R}_X(m) \tilde{\omega}(m) e^{-j\omega m}, \quad M \leq N. \quad (5.64)$$

式中窗函数

$$\tilde{\omega}(m) = \begin{cases} \tilde{\omega}(m), & -(M-1) \leq m \leq (M-1) \\ 0, & \text{其他} \end{cases} \quad (5.65)$$

由于功率谱是非负的, 则要求加窗后的功客谱也是非负的, 因此窗函数  $\tilde{\omega}(m)$  的选择必须满足的原则, 即他的傅里叶变换必须是非负的(此时,  $\tilde{\omega}(m)$  也应是一个偶序列)。例如三角(Bartlett)窗和 Parzen 窗是满足此条件的, 而海宁(Hanning)窗和汉明(Hamming)窗就不满足此条件。其中应用较多的 Bartlett 窗为

$$\tilde{\omega}(m) = \begin{cases} 1 - \frac{|m|}{M}, & |m| \leq M, M < N \\ 0, & \text{其他} \end{cases} \quad (5.66)$$

#### 2. 平均法

为了了解周期图经过平均后会使它的方差减小, 达到一致估计的目的, 我们先来看一个定理: 如果  $x_1, x_2, \dots, x_L$  是不相关的随机变量, 且每一个都有均值  $\mu$  及方差  $\sigma^2$ , 则可以证明它们的算术平均  $\bar{x} \triangleq (x_1 + x_2 + \dots + x_L) / L$  的均值等于  $\mu$ , 方差为  $D[\bar{x}] = \sigma^2 / L$ 。

由定理可见: 具有  $L$  个独立同分布随机变量平均的方差, 是单个随机变量方估计量经过算术平均后得到的估计量的方差, 也是原估计量的方差的  $1/L$ 。

平均周期图方法即是将数据  $x_0, x_1, \dots, x_{N-1}$  分段求周期图后再平均。比如, 给定  $N = 1000$  个数据样本 (平均法适用于数据量大的场合), 则可以将它分成 10 个长度为 100 的小段, 分别计算每一小段的周期图

$$G_{100,l}(\omega) = \frac{1}{100} \left| \sum_{n=100(l-1)}^{100l-1} x_n e^{-j\omega n} \right|^2, \quad l = 1, 2, \dots, 10. \quad (5.67)$$

然后将这 10 个周期图加以平均, 得谱估值为

$$\hat{G}_{100}^{10}(\omega) = \frac{1}{10} \sum_{l=1}^{10} G_{100,l}(\omega). \quad (5.68)$$

由于这 10 小段的周期图取决于同一过程, 因而其均值应该相同。若这 10 个小段的周期图是统计独立的, 则这 10 个小段平均之后的方差却是单段方差的  $1/10$ 。

$$D[\hat{G}_{100}^{10}(\omega)] = \frac{1}{10} D[G_{100,l}(\omega)]. \quad (5.69)$$

即: 平均法将  $x_0, x_1, \dots, x_{N-1}$  的  $N$  个数据分成  $L$  段 ( $N = LM$ ), 若各段数据相互独立, 则平均后估计量的方差是原来不分段估计量方差的  $1/L$ 。所以当  $L \rightarrow \infty$  时, 估计的方差趋于零, 达到了一致估计的目的。但是, 随着分段数  $L$  的增加,  $M$  点数减少, 虽然偏差减小, 但方差增大。况且此时各组数据之间因存在一定的相关性而非完全独立, 也会使方差增大。所以, 在实际应用时必须兼顾分辨力与方差的要求, 来适当选择  $L$  的值。

### 3.Welch

Welch 法又称修正周期图法, 它将“窗函数法”与“平均法”结合使用, 其步骤为:

- (1) 先将  $N$  个数据分成  $L$  段, 每段  $M$  个数据,  $N = ML$ .
- (2) 选择适当的窗函数  $\tilde{\omega}(n)$ , 并用该  $\tilde{\omega}(n)$  依次对每段数据作相应的加权, 然后确定每一段的周期图

$$G_{N,l}(\omega) = \frac{1}{MU} \left| \sum_{n=M(l-1)}^{Ml-1} x_n \tilde{\omega}(n) e^{-j\omega n} \right|^2, \quad 1 \leq l \leq L. \quad (5.70)$$

这里  $U$  为归一化因子

$$U = \frac{1}{M} \sum_{n=0}^{M-1} \tilde{\omega}^2(n). \quad (5.71)$$

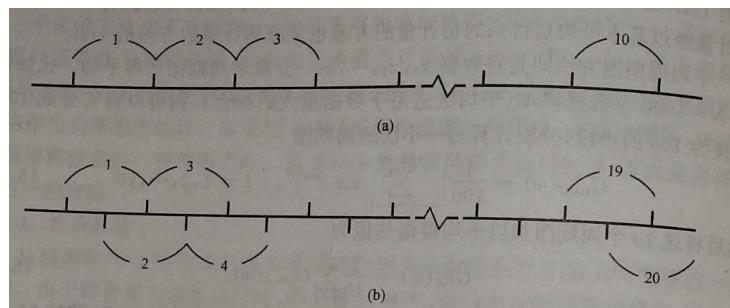


图 5-4 Welch 法

对分段周期图进行平均得到功率谱估计

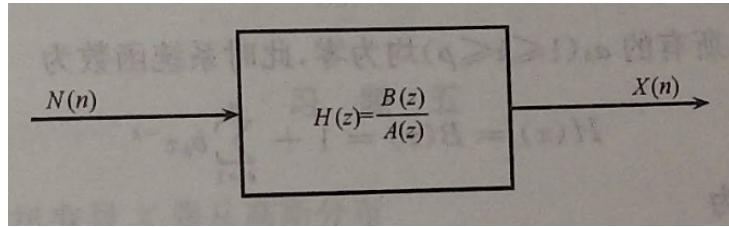
$$\hat{G}_M^L(\omega) = \frac{1}{L} \sum_{l=1}^L G_{M,i}(\omega). \quad (5.72)$$

由前面的分析, 当数据量一定时, 若分段数  $L$  增加,  $M$  点数减少, 则分辨力下降; 若分段数  $L$  增加, 虽  $M$  增加, 但方差增大。数据量一定时, 由于  $N = ML$ , 不可能  $L$  与  $M$  两个都大。解决这一矛盾的方法是, 让数据段之间适当“重造”, 如图 5-4(b) 所示。如果, 数据共有 1000 个点, 每段为 100 个点, 若重复 50%, 则可以分成 20 段。这样的分法, 由于段间相关性变大, 明然不能使方差减少到原来的  $1/20$ , 但肯定小于  $1/10$ , 而分辨力却没变, 这是 Welch 法的一个优点。Welch 法的另一个优点是, 对窗函数  $\tilde{\omega}(n)$  没有要求, 因为在式 (5-83) 中可以看到, 无论什么样的窗函数都可以使谱估计非负。

注意: 段长  $M$  应满足频率分辨力的要求。若要求的频率分辨力是  $\Delta f$ (Hz), 则归一化的频率分辨力要求  $\Delta f/f$ 。因此要求  $1/M \leq \Delta f/f_s$  或  $M \geq f_s/\Delta f$ 。为了增加 FET 后的谱线的密度, 可以在每段数据后适当补零, 一般是对  $M$  长的数据增添  $m$  个零(或者使添零的结果让总数为 2 的方幂), 作  $2M$  个点的 FFT。注意, 补零并不能提高频率分辨力, 只能增加功率谱的谱线密度, 克服栅栏效应。

### 5.3.4 现代谱估计简介

经典谱估计方法可以利用 FFT 计算, 因而有计算效率高的优点, 在谱分辨率要求不是很高的地方常用这种方法。经典谱估计的一个主要缺点是频率分辨率低。这是由于 BT 法仅利用  $N$  个有限的观测数据作自相关函数行计, 也就隐含着建立了在已知数据之外的自相关函数均为零的假设, 或者周期图法在计算中把观测到的有限长的  $N$  个数据以外的数据认为是零。这显然与事实不符。把观测不到的或估计不到的值认为是零, 相当于在时域中乘上了一个矩形窗口函数, 这在频域里相当于引入了一个与之卷积的辛克函

图 5-5  $X(n)$  表示为白噪声通过线性系统的输出

数。由于主瓣宽度反比于数据记录长度  $N$ , 而在实际中一般又不可能获得很长的数据记录, 所以辛克函数主瓣不会很窄。如果原来真实的功率谱是很窄的, 那么与主瓣卷积后会使功率向附近频域扩散, 使得信号模糊, 降低了分辨率。可见主瓣越宽分辨率越差, 严重时, 会使主瓣产生很大的失真, 甚至主瓣中的弱分量会被旁瓣中的强泄漏所掩盖。

为了克服以上缺点, 人们提出了平均、加窗平滑等方法, 在一定程度上改善了经典谱估计的性能。但是, 经典谱估计始终无法解决频率分辨率与谱估计稳定性之间的矛盾。特别是在数据记录很短的情况下, 这一矛盾显得尤为突出。

在第四章的讨论中我们知道, 任何具有有理功率谱密度的随机信号都可以看能克服经典谱估计的缺点。如由这个模型来求功率谱估计, 可望得到比较好的传统的输出, 如图 5-5 所示。其传递函数为

$$H(z) = \frac{B(z)}{A(z)} = \frac{\sum_{k=0}^q b_k z^{-k}}{\sum_{k=0}^p a_k z^{-k}} \quad (5.73)$$

当输入白噪声的功兹谱密度  $G_N(\omega) = \sigma_N^2$  时, 输出的功率谱密度为

$$G_N(\omega) = \sigma_N^2 |H(e^{j\omega})|^2 = \sigma_N^2 \left| \frac{B(e^{j\omega})}{A(e^{j\omega})} \right|^2. \quad (5.74)$$

如果能确定  $\sigma_N^2$  与  $a_k (0 \leq k \leq p)$ ,  $b_k (0 \leq k \leq q)$  值, 通过上式就可得到所需信号的功率谱密度。由于  $|H(e^{j\omega})|$  的增益系数可并入  $\sigma_N^2$  进行考虑, 因此, 为不失一般性可假设  $a_0 = 1, b_0 = 1$ 。我们将这个线性系统分为三种情况。

### 1) AR 模型

除  $b_0 = 1$  外所有的  $b_k (1 \leq k \leq q)$  均为零, 此时系统函数为

$$H(z) = \frac{1}{A(z)} = \frac{1}{1 + \sum_{k=1}^p a_k z^{-k}}. \quad (5.75)$$

其差分方程为

$$X(n) = - \sum_{k=1}^p a_k X(n-k) + N(n). \quad (5.76)$$

其信号流图则如图 5-6 所示。

这种模型称为  $p$  阶自回归模型，简称 AR(Autoregressive) 模型。其系统函数只有极点，没有零点，因此也称为全极点模型。模型输出功率谱为

$$G_X(\omega) = \frac{\sigma_N^2}{|A(e^{j\omega})|^2} = \frac{\sigma_N^2}{1 + \sum_{k=1}^p a_k e^{-j\omega k}}. \quad (5.77)$$

只要我们能求得  $\sigma_N^2$  及所有  $a_k$  值，就可以求得随机信号  $X(n)$  的功率谱。

### 2) MA 模型

除  $a_0 = 1$  外，所有的  $a_k (1 \leq k \leq p)$  均为零，此时系统函数为

$$H(z) = B(z) = 1 + \sum_{k=1}^q b_k z^{-k} \quad (5.78)$$

系统的差分方程为

$$X(n) = \sum_{k=0}^q b_k N(n-k). \quad (5.79)$$

其信号流图如图 5-7 所示。

这种模型称为  $q$  阶滑动平均模型，简称 MA (MovingAverage) 模型。其系统函数只有零点，没有极点，因此也称为全零点模型。模型输出功率谱为

$$G_X(\omega) = \sigma_N^2 |B(e^{j\omega})|^2 = \sigma_N^2 \left| \left( 1 + \sum_{k=1}^q b_k e^{-j\omega k} \right) \right|. \quad (5.80)$$

只要我们能求得  $\sigma_N^2$  及所有  $b_k$  值，就可以求得随机信号  $X(n)$  的功率谱。

### 3) ARMA 模型

设  $a_0 = 1$ ，和  $b_0 = 1$ ，其余所有的  $a_k$  和  $b_k$  不全为零。此时模型的传递函数和输出功率谱分别由式(5.73)和式(5.74)表示。这是一个“极点、零点”模型，称为 ARMA( $p, q$ ) 模型，或直接简称 ARMA 模型。

- (1) 选择一个合适的模型。
- (2) 用已观测到的数据估计模型参数。
- (3) 将模型参数代入功率谱的计算公式就，得到功率谱估值。

由以上讨论可知，用模型法作功率谱估计，实际上要解决的是模型的参数估计问题。所以这类谱估计方法又统称为参数化方法。

基于模型的功率谱估计方法首先要选择一个模型，但在一般情况下，我们设定：有随机信号模型的先验知识。若模型选择不当是否会对谱估计性能产生较大的影响？Word

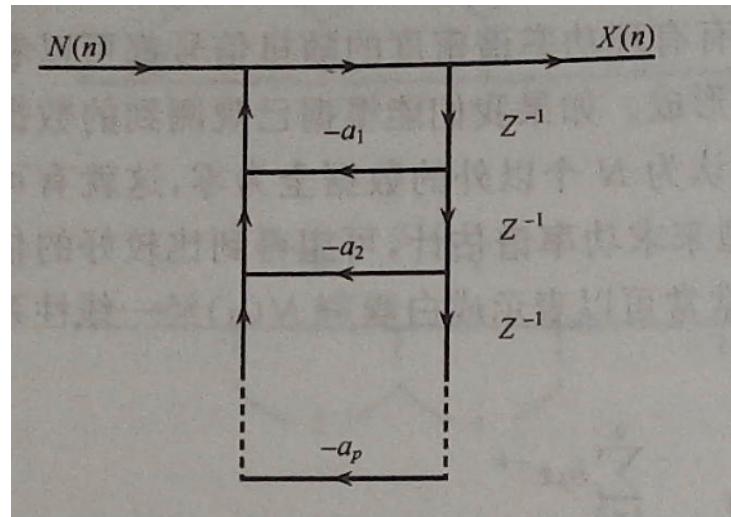


图 5-6 AR 模型

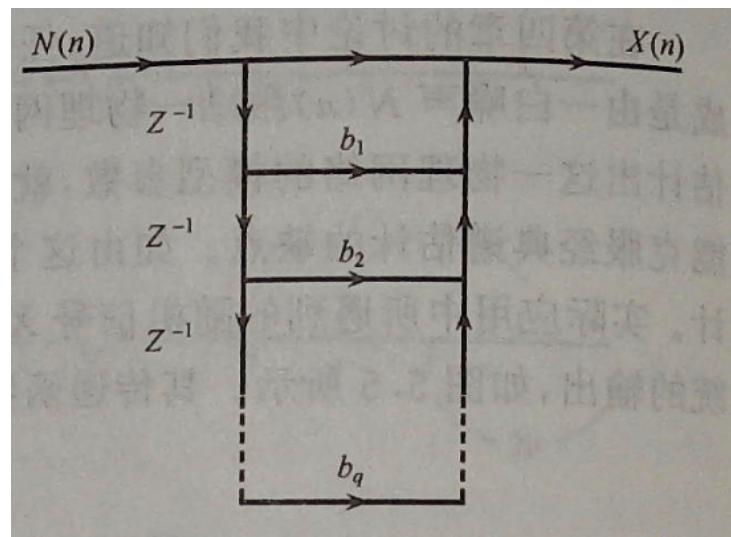


图 5-7 MA 模型

分解定理告诉我们, 如果功率谱是连续的, 则任何 ARMA 过程或 AR 过程也可以用一个无限阶的 MA 过程表示。Kolmogorov 也提出, 任何 ARMA 或 MA 过程也可以用一个无限阶的 AR 过程表示, 所以如果选择了一个不合适的模型, 但只要模型的阶数足的高, 它仍然能够比较好地逼近被建模的随机过程。

在这三种参数模型中, AR 模型得到了普遍应用。其原因为

- (1) 任意 ARMA 或 MA 信号模型都可以用无限阶或阶数足够大的 AR 模型来表示。
- (2) AR 模型的参数计算是线性方程, 比较简便。同时适合表示很窄的频谱。
- (3) AR 模型在作谱估计时, 由于具有递推特性, 所以所需的数据较短, 很适合表示窄的功率谱:

而 MA 模型表示窄谱时, 一般需要数量很多的参数;ARMA 模型虽然所需的参数数量最少, 但参数估计的算法是非线性方程组, 其运算远比 AR 模型复杂。

## 5.4 习题五

✎ **练习 5.1** 已知随机变量  $X$  服从高斯分布

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \quad (5.81)$$

求未知参数  $\mu$  和  $\sigma^2$  的最大似然估计。

✎ **练习 5.2** 已知平稳过程  $X(t)$  的数学期望  $\mu$  和方差  $\sigma^2$  存在,  $x_0, x_1, \dots, x_{N-1}$  是过程的样本数据。证明:

- (1)  $\hat{\mu} = \sum_{i=0}^{N-1} c_i x_i$  是数学期望  $\mu$  的无偏估计量, 其中  $\sum_{i=0}^{N-1} c_i = 1, c_i > 0$ .
- (2)  $\hat{\sigma}_X^2 = \frac{1}{N} \sum_{i=0}^{N-1} (x_i - \hat{\mu})^2$  是  $\sigma^2$  的无偏估计量。
- (3)  $\hat{\sigma}_X^2 = \frac{1}{N} \sum_{i=0}^{N-1} (x_i - \bar{x})^2$  不是  $\sigma^2$  的无偏估计量。

✎ **练习 5.3** 已知  $\hat{\theta}$  是参数  $\theta$  的无偏估计, 且有  $D[\hat{\theta}] > 0$ , 证明:  $(\hat{\theta})^2$  不是  $\theta^2$  的无偏估。

✎ **练习 5.4** 证明: 用式  $\hat{R}_X(m) = \frac{1}{N} \sum_{n=0}^{N-|m|-1} x^*(n)x(n+m)$  进行自相关函数的估计是渐进一致估计。

✎ **练习 5.5** 用周期图法进行谱估计时

$$\hat{G}_X(\omega) = \frac{1}{N} \left| \sum_{n=0}^{N-1} x_n e^{-i\omega n} \right|^2. \quad (5.82)$$

说明为什么可用 FFT 进行计算。周期图法的谱分辨率较低, 且估计的方差也较大, 说明造成这两种缺点的原因, 并说明无论选取什么样的窗函数都难以从根本上解决问题的原因。

- ✉ 练习 5.6 说明白噪声过程的周期图是一个无偏估计器。
- ✉ 练习 5.7 采用下式给出的有偏自相关函数的定义，并加窗  $\tilde{\omega}(k)$  得到 BT 谱估计器

$$\hat{R}_X(k) = \begin{cases} \frac{1}{N} \sum_{n=0}^{N-1} x^*(n)x(n+k), & k = 0, 1, \dots, N-1 \\ \hat{R}_x(-k), & k = -(N-1), -(N-2), \dots, -1 \end{cases}$$

$$\tilde{\omega}(k) = \begin{cases} 1, & |k| \leq N-1 \\ 0, & \text{其他} \end{cases}$$

$$\bar{G}_X(\omega) = \sum_{k=-N-1}^{N-1} \hat{R}_X(k) \tilde{\omega}(k) e^{-j\omega t}$$

说明 BT 谱估计器与周期图相同。

## 5.5 知识点总结

- ▶ 估计的准则
  - 偏移性 (无偏估计、有偏估计、渐进无偏估计)
  - 估计量的方差
  - 均方误差 (一致估计)
- ▶ 时域信号的估计 (最大似然估计)
  - 均值的估计及其评价
  - 方差的估计及其评价
  - 自相关的估计及其评价
  - 概率密度的估计 (直方图法)
  - 频域特征的估计
  - 经典谱估计 (BT 法、周期图法、窗函数法、平均法、welch 法)



# 6

## 窄带随机信号

通信、雷达和广播电视等信息传输系统以及传输的信号均满足窄带条件：中心频率  $\omega_0$  远大于谱宽  $\Delta\omega$ , 即  $\omega_0 \gg \Delta\omega$ 。这类信号和系统分别称为窄带信号和窄带系统。其频率有选择性。工作在高频或中频的无线电系统，多数都是窄带系统。

类似地，一个随机信号的功率谱密度，只要分布在高频载波  $\omega_0$  附近的一个“窄随机过程”。在信息传输系统中，特别是接收机中经常遇到的随机信号，都是窄带随机式及随机过程的复数表示法，因此，本章首先介绍这些方法，然后再将这些方法应用到窄带随机过程的分析中去。

### 6.1 预备知识

#### 6.1.1 信号的解析形式

##### 1. 单边谱信号

我们知道，实信号频谱的数学模型是含有正负频率的双边谱，然而在实际应用中，其负频率 ( $a < 0$ ) 是物理不可实现的。由于实信号的双边谱是关于 0 上的竖轴偶对称，因此，采用其单边谱的信号形式，既可简化对问题的分析，又可恢复原信号。下面对只含正频域部分的信号—单边谱信号进行讨论。

###### (1) 单边谱信号在时域是复信号

设单边谱信号的傅里叶变换为

$$f(t) \xrightleftharpoons[F]{F^{-1}} F(\omega), \omega > 0 \quad (6.1)$$

由于

$$f(t) = \frac{1}{2\pi} \int_0^\infty F(\omega) e^{j\omega t} d\omega, \quad (6.2)$$

则

$$f^*(t) = \frac{1}{2\pi} \int_0^\infty F^*(\omega) e^{-j\omega t} d\omega, \quad (6.3)$$

因为  $f^*(t) \neq f(t)$ , 所以单边谱信号在时域内的表示是个复信号。

(2) 从实信号中分解出单边谱信号

设  $s(t)$  为具有连续频谱的实信号

$$s(t) = \frac{1}{2\pi} \int_{-\infty}^\infty S(\omega) e^{j\omega t} d\omega, \quad (6.4)$$

式中  $S(\omega)$  为信号  $s(t)$  的频谱。由傅里叶变换, 可以证明, 当  $s^*(t) = s(t)$  时, 有

$$S(-\omega) = \int_{-\infty}^\infty s(t) e^{-j\omega t} dt = \int_{-\infty}^\infty s^*(t) (e^{j\omega t})^* dt = S^*(\omega), \quad (6.5)$$

所以实信号  $s(t)$  的频谱  $S(\omega)$  有性质 (6.5), 可知实信号  $s(t)$  的频谱  $S(\omega)$  是  $\omega$  的复函数。

将  $S(\omega)$  用傅里叶逆变换分解成正负两频域部分积分的和, 得到实信号  $s(t)$ ,

$$\begin{aligned} s(t) &= \frac{1}{2\pi} \int_{-\infty}^\infty S(\omega) e^{j\omega t} d\omega \\ &= \frac{1}{2\pi} \int_0^\infty S(\omega) e^{j\omega t} d\omega + \frac{1}{2\pi} \int_{-\infty}^0 S(\omega) e^{j\omega t} d\omega \\ &\stackrel{\omega'=-\omega}{=} \frac{1}{2\pi} \int_0^\infty S(\omega) e^{j\omega t} d\omega + \frac{1}{2\pi} \int_0^\infty S(-\omega') e^{-j\omega' t} d\omega' \\ &= \frac{1}{2\pi} \int_0^\infty S(\omega) e^{j\omega t} d\omega + \left[ \frac{1}{2\pi} \int_0^\infty S(\omega') e^{j\omega' t} d\omega' \right]^* \\ &= \operatorname{Re} \left[ \frac{1}{2\pi} \int_0^\infty 2S(\omega) e^{j\omega t} d\omega \right] = \operatorname{Re}[\tilde{s}(t)], \end{aligned} \quad (6.6)$$

其中单边谱信号

$$\tilde{s}(t) = \frac{1}{2\pi} \int_0^\infty 2S(\omega) e^{j\omega t} d\omega. \quad (6.7)$$

式 (6.7) 和式 (6.4) 只在积分上下限和被积函数有区别; 单边谱信号  $\tilde{s}(t)$  被称为实信号  $s(t)$  的解析信号, 或  $s(t)$  的包络。此外,  $\tilde{s}(t)$  还具有如下的单边频谱

$$\begin{aligned} \tilde{S}(\omega) &= \int_{-\infty}^\infty \tilde{s}(t) e^{-j\omega t} dt = \int_{-\infty}^\infty \frac{1}{2\pi} \int_0^\infty 2S(\omega) e^{j\omega t} d\omega e^{-j\omega t} dt \\ &= 2 \int_0^\infty \frac{1}{2\pi} \int_{-\infty}^\infty S(\omega) e^{j\omega t} d\omega e^{-j\omega t} dt = 2 \int_0^\infty s(t) e^{-j\omega t} dt. \end{aligned} \quad (6.8)$$

$$\tilde{S}(\omega) = 2S(\omega)U(\omega) = \begin{cases} 2S(\omega), & \omega > 0 \\ 0, & \omega < 0 \end{cases}, U(\omega) = \begin{cases} 1, & \omega > 0 \\ 0, & \omega < 0 \end{cases} \quad (6.9)$$

所以, 实信号  $s(t)$  可用一个仅含其正频率成分的解析信号  $\tilde{s}(t)$  的实部来表示, 即  $s(t) = \operatorname{Re}[\tilde{s}(t)]$ 。

## 2. 解析信号的表示方法。

由于单边谱信号  $\tilde{s}(t)$  在时域是个复信号, 所以根据复数的性质,  $s(t)$  的解析信号  $\tilde{s}(t)$  可以表示为

$$\tilde{s}(t) = \operatorname{Re}[\tilde{s}(t)] + j\operatorname{Im}[\tilde{s}(t)] \in \mathbb{C} \times T. \quad (6.10)$$

由  $\operatorname{Re}[\tilde{s}(t)] = s(t)$ , 再令

$$\operatorname{Im}[\tilde{s}(t)] = \hat{s}(t). \quad (6.11)$$

即解析信号  $s(t)$  的虚部用符号  $\hat{s}(t)$  表示, 则

$$\tilde{s}(t) = s(t) + j\hat{s}(t), \quad (6.12)$$

上式是解析信号  $\tilde{s}(t)$  的一般表达式, 然而, 式中的虚部  $\hat{s}(t)$  该如何求呢?

从频域关系出发, 解析信号  $\tilde{s}(t)$  的频谱  $\tilde{S}(\omega)$  满足 (6.9) 关系式  $\tilde{S}(\omega) = 2S(\omega)U(\omega)$ , 对等式两边进行傅里叶逆变换。由于

$$U(\omega) \xrightarrow[F]{F^{-1}} \frac{1}{2} \left[ \delta(t) - \frac{1}{j\pi t} \right], \quad (6.13)$$

则解析信号的时域表达式为

$$\tilde{s}(t) = s(t) * \left[ \delta(t) - \frac{1}{j\pi t} \right] = s(t) + j \left[ \textcolor{blue}{s(t)} * \frac{1}{\pi t} \right]. \quad (6.14)$$

不难得到

$$\hat{s}(t) = s(t) * \frac{1}{\pi t} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s(\tau)}{t - \tau} d\tau = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s(t - \tau)}{\tau} d\tau, \quad (6.15)$$

 **注 6.1.** 式 (6.13) 中复函数的计算基础:

$$\oint_{|z|<0.2} \frac{1}{j\pi z} e^{-j\omega z} dz = \frac{1}{j\pi} \oint_{|z|<0.2} \frac{1}{t} e^{-j\omega z} dz = \frac{2j\pi}{j\pi} = 2. \quad (6.16)$$

### 定义 6.2 希尔伯特变换

给定实信号  $s(t)$ , 解析信号  $\tilde{s}(t)$  的频谱  $\tilde{S}(\omega)$  关系式为  $\tilde{S}(\omega) = 2S(\omega)U(\omega)$ , 时域表达式由 (6.14) 给出, 则实信号  $s(t)$  的希尔伯特 (Hilbert) 变换  $\hat{s}(t)$  可由式 (6.14) 给出, 简记作  $\hat{s}(t) = H[s(t)]$ 。



**结论** 对于任何一个实信号  $s(t)$ , 都可以分解出一个单边谱的解析信号  $\tilde{s}(t)$  与其对应。此解析信号是个复信号, 其实部为原信号  $s(t)$ , 虚部为原信号  $s(t)$  的希尔伯特变换  $\hat{s}(t)$ 。

### 6.1.2 希尔伯特变换的性质

解析信号是最常用的复信号之一, 它在分析窄带随机信号中起着重要作用, 希尔伯特变换是应用解析信号时必不可少的数学工具, 因此有时也将用解析形式表示复信号的方法称为希尔伯特表示法。由定义, 希尔伯特变换 (6.15) 有以下几个重要性质。

#### 1. 希尔伯特变换相当于一个正交滤波器。

希尔伯特变换 (6.15) 是  $s(t)$  和  $1/\pi t$  的卷积, 根据线性系统的输出特征, 可以将  $\hat{s}(t)$  看成是  $s(t)$  通过一个具有冲激响应为  $h_{\wedge} = \frac{1}{\pi t}$  的线性滤波器, 如图 6-1 所示

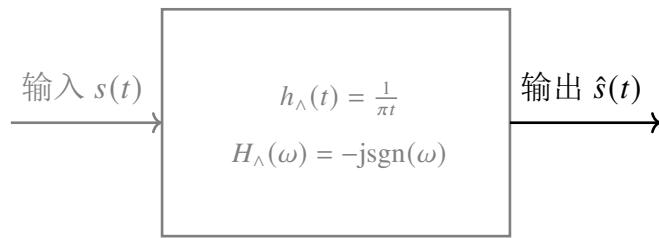


图 6-1 希尔伯特变换

由于如下的傅里叶变换对存在, 即

$$j \frac{1}{\pi t} \xrightarrow[F]{F^{-1}} \text{sgn}(\omega), \quad (6.17)$$

式中  $\text{sgn}(\omega)$  为符号函数, 满足

$$\text{sgn}(\omega) = \begin{cases} 1, & \omega > 0 \\ -1, & \omega < 0 \end{cases} \quad (6.18)$$

等价地, 有如下的傅里叶变换对

$$\frac{1}{\pi t} \xrightarrow[F]{F^{-1}} -j \text{sgn}(\omega). \quad (6.19)$$

则图 6-1 所示的线性滤波器的传递函数 (频率响应) 为

$$H_{\wedge}(\omega) = -j\text{sgn}(\omega). \quad (6.20)$$

其幅频特性和相频特性分别为

$$|H_{\wedge}(\omega)| = 1, \theta_{H_{\wedge}(\omega)} = \begin{cases} -\pi/2, & \omega > 0 \\ +\pi/2, & \omega < 0 \end{cases}. \quad (6.21)$$

与滤波器输出信号  $\hat{s}(t)$  相应的频谱为

$$\hat{S}(\omega) = H_A(\omega)S(\omega) = -j\text{sgn}(\omega)S(\omega) = \begin{cases} -jS(\omega), & \omega > 0 \\ jS(\omega), & \omega < 0 \end{cases}. \quad (6.22)$$

由式(6.22)可见, 通过此滤波器的信号, 其所有频率分量的幅度响应为 1. 而在相位上, 所有正频率分量移相  $-\pi/2$ , 所有负频率分量移相  $+\pi/2$ 。因此说, 希尔伯特变换是一种正交变换, 它的作用相当于一个正交滤波器, 一次变换如图 6-2 中所示。

## 2. 两次希尔伯特变换相当于一个倒相器。

若对  $s(t)$  进行两次希尔伯特变换, 则相当于信号  $s(t)$  通过两个级联的  $h_A(t)$  网络, 即

$$\hat{\hat{s}}(t) = H\{H[s(t)]\} = H[\hat{s}(t)] = \hat{s}(t) * \frac{1}{\pi t} = s(t) * \frac{1}{\pi t} * \frac{1}{\pi t} \quad (6.23)$$

$$\hat{\hat{S}}(\omega) = \hat{S}(\omega)[-j\text{sgn}(\omega)] = S(\omega)[-j\text{sgn}(\omega)][-j\text{sgn}(\omega)] = -S(\omega) \quad (6.24)$$

从而得到时域关系

$$\hat{\hat{s}}(t) = H\{H[s(t)]\} = -s(t). \quad (6.25)$$

如图 6-2 所示, 两次希尔伯特变换将信号  $s(t)$  翻转了  $180^\circ$ 。

## 3. 希尔伯特逆变换等于负的希尔伯特正变换 $H^{-1} = -H$ 。

定义希尔伯特逆变换为

$$s(t) = H^{-1}[\hat{s}(t)]. \quad (6.26)$$

由式(6.25)可得

$$s(t) = -H[\hat{s}(t)]. \quad (6.27)$$

对比式(6.26)和式(6.27)可知

$$H^{-1}[\hat{s}(t)] = -H[\hat{s}(t)]. \quad (6.28)$$

即希尔伯特逆变换等于负的希尔伯特正变换

$$H^{-1}[\cdot] = -H[\cdot]. \quad (6.29)$$

因此, 希尔伯特逆变换的公式为

$$\begin{aligned} s(t) &= H^{-1}[\hat{s}(t)] = -H[\hat{s}(t)] = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\hat{s}(t-\tau)}{\tau} d\tau \\ &\stackrel{x=-\tau}{=} -\frac{1}{\pi} \int_{\infty}^{-\infty} \frac{\hat{s}(t+x)}{-x} dx = x \end{aligned} \quad (6.30)$$

$$\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\hat{s}(t+x)}{x} dx = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\hat{s}(t+\tau)}{\tau} d\tau. \quad (6.31)$$

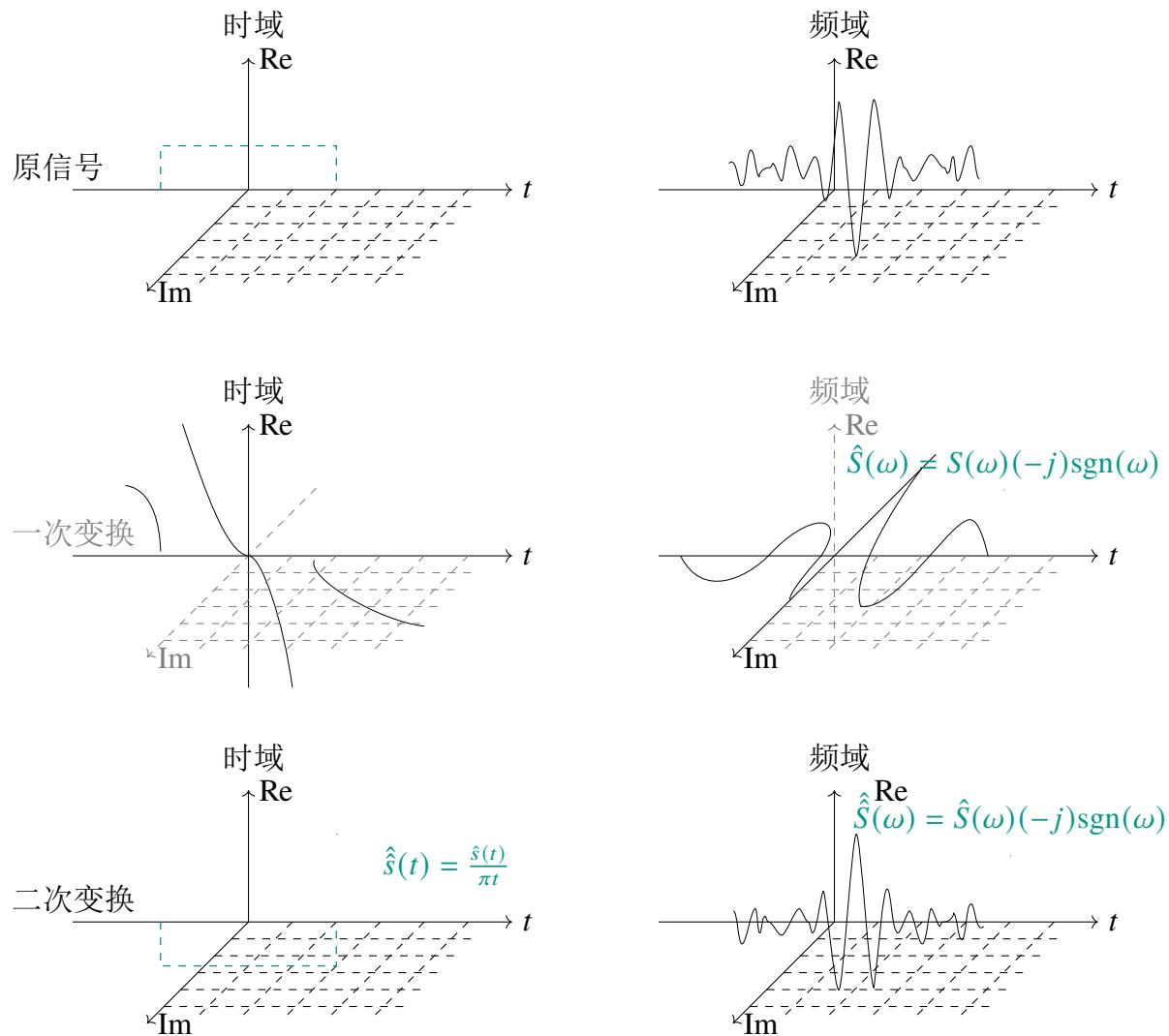


图 6-2 希尔伯特变换

**例 6.3** 求  $\cos \Omega t, \sin \Omega t$  的希尔伯特变换。

解:

①  $H[\cos \Omega t]$  的计算:

$$\begin{aligned}
 H[\cos \Omega t] &= \cos \Omega t * \frac{1}{\pi t} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\cos[\Omega(t-\tau)]}{\tau} d\tau \\
 &= \frac{1}{\pi} \int_{-\infty}^{\infty} \left[ \frac{1}{\tau} \cos \Omega t \cos \Omega \tau + \frac{1}{\tau} \sin \Omega t \sin \Omega \tau \right] d\tau \\
 &= \frac{1}{\pi} \left[ \cos \Omega t \int_{-\infty}^{\infty} \frac{\cos \Omega \tau}{\tau} d\tau + \sin \Omega t \int_{-\infty}^{\infty} \frac{\sin \Omega \tau}{\tau} d\tau \right].
 \end{aligned} \tag{6.32}$$

由于  $\frac{\cos \Omega \tau}{\tau}$  是  $\tau$  的奇函数, 所以第一项积分为 0, 而  $\frac{\sin \Omega \tau}{\tau}$  是  $\tau$  的偶函数, 所以

$$H[\cos \Omega t] = \frac{2}{\pi} \sin \Omega t \int_0^\infty \frac{\sin \Omega \tau}{\tau} d\tau. \quad (6.33)$$

其中

$$\int_0^\infty \frac{\sin \Omega \tau}{\tau} d\tau = \frac{\pi}{2} \operatorname{sgn}(\Omega) = \begin{cases} \frac{\pi}{2}, & \Omega > 0 \\ -\frac{\pi}{2}, & \Omega < 0 \end{cases}. \quad (6.34)$$

所以

$$H[\cos \Omega t] = \operatorname{sgn}(\Omega) \sin \Omega t = \begin{cases} \sin \Omega t, & \Omega > 0 \\ -\sin \Omega t, & \Omega < 0 \end{cases}. \quad (6.35)$$

 **注 6.4.** sine 积分函数定义 [GautschiW1972] 为

$$\operatorname{Si}(x) = \int_0^x \frac{\sin(t)}{t} dt \quad (6.36)$$

对于 (6.34),

$$\int_0^\infty \frac{\sin \Omega \tau}{\tau} d\tau = \int_0^\infty \frac{\sin \Omega \tau}{\Omega \tau} d\Omega \tau = \int_0^\infty \frac{\sin t}{t} dt \quad (6.37)$$

可以使用指令 `sinint(inf)` 求  $\operatorname{Si}(x), x > 0$  的结果。

② 对上式两端再求一次 Hilbert 变换, 有

$$H\{H[\cos \Omega t]\} = H[\operatorname{sgn}(\Omega) \sin \Omega t] = \operatorname{sgn}(\Omega) H[\sin \Omega t]. \quad (6.38)$$

利用性质 2 可知, 两次希尔伯特变换相当于倒相器, 则

$$H\{H[\cos \Omega t]\} = -\cos \Omega t. \quad (6.39)$$

综合上述两式 (6.38) 和 (6.39), 可得

$$H[\operatorname{sgn}(\Omega) \sin \Omega t] = -\operatorname{sgn}(\Omega) H[\sin \Omega t] = H\{H[\cos \Omega t]\} = -\cos \Omega t. \quad (6.40)$$

则有

$$H[\sin \Omega t] = -\operatorname{sgn}(\Omega) \cos \Omega t = \begin{cases} -\cos \Omega t, & \Omega > 0 \\ \cos \Omega t, & \Omega < 0 \end{cases} \quad (6.41)$$

**例 6.5** 设限带信号

$$\begin{cases} s_1(t) = a(t) \cos \omega_0 t \\ s_2(t) = a(t) \sin \omega_0 t \end{cases} \quad (6.42)$$

其中  $a(t)$  为低频限带信号, 其频谱为

$$A(\omega), |\omega| < \frac{\Delta\omega}{2} < \omega_0 \quad (6.43)$$

求  $s_1(t), s_2(t)$  的 Hilbert 变换。

解: ① 利用傅里叶变换的相乘性质, 有

$$\begin{aligned} a(t) \cos \omega_0 t &\xrightarrow[F]{F^{-1}} \frac{1}{2\pi} A(\omega) * \pi [\delta(\omega - \omega_0) + \delta(\omega + \omega_0)] \\ S_1(\omega) &= \frac{1}{2} [A(\omega - \omega_0) + A(\omega + \omega_0)] \end{aligned} \quad (6.44)$$

如图 6-3 所示, 由于  $\frac{\Delta\omega}{2} < \omega_0$ , 可得

$$S_1(\omega) = \begin{cases} \frac{1}{2}A(\omega - \omega_0), & \omega > 0 \\ \frac{1}{2}A(\omega + \omega_0), & \omega < 0 \end{cases} \quad (6.45)$$

所以其希尔伯特变换的频谱为

$$\hat{S}_1(\omega) = -j \operatorname{sgn}(\omega) S_1(\omega) = \begin{cases} -\frac{j}{2}A(\omega - \omega_0), & \omega > 0 \\ \frac{j}{2}A(\omega + \omega_0), & \omega < 0 \end{cases} \quad (6.46)$$

取  $\hat{S}_1(\omega)$  的傅里叶逆变换可得

$$\begin{aligned} \hat{s}_1(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} -j \operatorname{sgn}(\omega) S_1(\omega) e^{j\omega t} d\omega \\ &= -\frac{j}{2} \left[ \frac{1}{2\pi} \int_0^{\infty} A(\omega - \omega_0) e^{j\omega t} d\omega \right] + \frac{j}{2} \left[ \frac{1}{2\pi} \int_{-\infty}^0 A(\omega + \omega_0) e^{j\omega t} d\omega \right] \\ &= -\frac{j}{2} \left[ \frac{1}{2\pi} \int_{-\infty}^{\infty} A(\omega - \omega_0) e^{j\omega t} d\omega \right] + \frac{j}{2} \left[ \frac{1}{2\pi} \int_{-\infty}^{+\infty} A(\omega + \omega_0) e^{j\omega t} d\omega \right]. \end{aligned} \quad (6.47)$$

② 利用傅里叶变换的频移性质

$$\begin{aligned} \hat{s}_1(t) &= H[a(t) \cos \omega_0 t] = -\frac{j}{2}a(t)e^{j\omega_0 t} + \frac{j}{2}a(t)e^{-j\omega_0 t} \\ &= \frac{j}{2}(e^{-j\omega_0 t} - e^{j\omega_0 t})a(t) = a(t) \sin \omega_0 t. \end{aligned} \quad (6.48)$$

③ 利用 Hilbert 性质和二次变换的性质, 可得

$$\hat{s}_2(t) = H[a(t) \sin \omega_0 t] = H\{H[a(t) \cos \omega_0 t]\} = -a(t) \cos \omega_0 t. \quad (6.49)$$

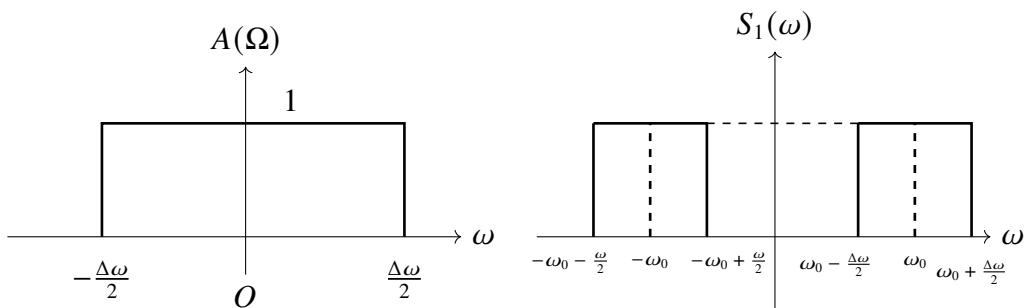


图 6-3 例 6.5 图

### 6.1.3 高频窄带信号的复指数形式

#### 1. 高频窄带信号

##### 定义 6.6 高频窄带信号

高频窄带信号是指信号的中心频率  $\omega_0$  位于高频段的窄带信号。



这类信号的频带  $\Delta\omega$  远远小于其中心频率  $\omega_0$ , 即  $\Delta\omega \ll \omega_0$ 。诸如雷达、通信等电子系统中的一些高频信号, 大多可以近似认为是高频窄带信号。这类信号的典型频谱如图 6-4(a) 所示。如果在示波器上观察这类窄带信号, 它的波形或多或少地有点像正弦波, 如图 6-4(b) 所示。但它的振幅、相位不是常数, 而是随时间  $t$  变化的函数。这类窄带信号  $s(t)$  可以表示为

$$s(t) = a(t) \cos [\omega_0 t + \varphi(t)], \quad (6.50)$$

式中  $a(t)$  是信号  $s(t)$  的振幅调制, 称为包络函数。 $\varphi(t)$  是信号  $s(t)$  的相位调制, 称为相位函数, 相对载波  $\cos \omega_0 t$  来讲,  $a(t)$  与  $\varphi(t)$  都是慢变化的时间函数。

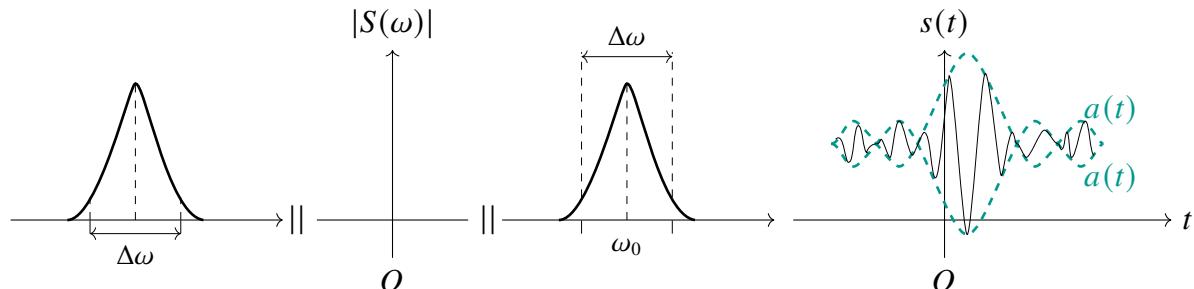


图 6-4 高频窄带信号

如果将上式展开, 则

$$\begin{aligned} s(t) &= a(t) \cos [\omega_0 t + \varphi(t)] \\ &= \underbrace{a(t) \cos \varphi(t)}_{m_c(t)} \cos \omega_0 t - \underbrace{a(t) \sin \varphi(t)}_{m_s(t)} \sin \omega_0 t \\ &= m_c(t) \cos \omega_0 t - m_s(t) \sin \omega_0 t, \end{aligned} \quad (6.51)$$

其中

$$\begin{cases} m_c(t) = a(t) \cos \varphi(t) \\ m_s(t) = a(t) \sin \varphi(t) \end{cases} \quad (6.52)$$

可见, 将正弦和余弦看成是函数基中的基函数,  $m_c(t)$  和  $m_s(t)$  相对  $\cos \omega_0 t$  来讲都是低频信号, 且  $m_c(t)$  和  $m_s(t)$  在几何上彼此正交。

## 2. 高频窄带信号的复指数形式

信号  $s(t)$  的希尔伯特变换为

$$\begin{aligned}\hat{s}(t) &= H[m_c(t) \cos \omega_0 t - m_s(t) \sin \omega_0 t] \\ &= H[m_c(t) \cos \omega_0 t] - H[m_s(t) \sin \omega_0 t].\end{aligned}\quad (6.53)$$

如果  $m_c(t), m_s(t)$  均为低频限带信号, 即满足

$$\begin{aligned}m_c(t) &\xrightleftharpoons[F^{-1}]{F} M_c(\omega), \quad |\omega| < \frac{\Delta\omega_c}{2} < \omega_0 \quad (\text{低频限带}) \\ m_s(t) &\xrightleftharpoons[F^{-1}]{F} M_s(\omega), \quad |\omega| < \frac{\Delta\omega_s}{2} < \omega_0 \quad (\text{低频限带}).\end{aligned}\quad (6.54)$$

由例 6-2 的结论, 可得  $s(t)$  的希尔伯特变换

$$\begin{aligned}\hat{s}(t) &= H[m_c(t) \cos \omega_0 t] - H[m_s(t) \sin \omega_0 t] \\ &= m_c(t) \sin \omega_0 t + m_s(t) \cos \omega_0 t \\ &= a(t) \cos \varphi(t) \sin \omega_0 t + a(t) \sin \varphi(t) \cos \omega_0 t \\ &= a(t) \sin [\omega_0 t + \varphi(t)].\end{aligned}\quad (6.55)$$

此时,  $s(t)$  的解析形式可以用复指数表示为

$$\begin{aligned}\tilde{s}(t) &= s(t) + j\hat{s}(t) = a(t) \cos [\omega_0 t + \varphi(t)] + ja(t) \sin [\omega_0 t + \varphi(t)] \\ &= a(t)e^{j[\omega_0 t + \varphi(t)]} = m(t)e^{j\omega_0 t} = \tilde{s}_{\text{复}}(t),\end{aligned}\quad (6.56)$$

式中

$$m(t) = a(t)e^{j\varphi(t)}, \quad (6.57)$$

通常, 将  $m(t)$  称为信号  $s(t)$  的复包络, 而将  $e^{j\omega_0 t}$  称为复载频。 $a(t)$  称为包络, 解析信号  $\tilde{s}(t)$  称为预包络。复包络  $m(t)$  可展开为

$$m(t) = a(t) \cos \varphi(t) + ja(t) \sin \varphi(t) = m_c(t) + jm_s(t). \quad (6.58)$$

可见, 相对  $\cos \omega_0 t$  来讲, 包络  $a(t)$  和复包络  $m(t)$  都是低频限带信号。

### 误差分析

#### (1) 误差产生的原因

从复指数形式的推导过程可以看出, 尽管解析信号  $\tilde{s}(t)$  是复信号, 但要将它表示成复指数形式, 还必须满足  $m_c(t), m_s(t)$  均为低频限带信号的条件。如果不满足此条件, 就不能由解析信号推出复指数表示。即不满足  $m_c(t)$  和  $m_s(t)$  低频限带信号的条件, 其解析信号表示与复指数表示仍存在一定的误差。

## (2) 误差的计算

下面举例, 在频域上对解析信号  $\tilde{s}(t)$  与复指数信号  $\tilde{s}_{\text{复}}(t)$  进行比较, 计算它们之间的误差。

**例 6.7** 设一实信号  $s(t) = a(t) \cos \omega_0 t$ , 其复包络  $m(t) = a(t)$ , 复包络的频谱  $M(\omega)$  不满足低频限带条件, 如图 6-5 所示

$$m(t) \xrightleftharpoons[F^{-1}]{F} M(\omega), \quad |\omega| < \Delta\omega \nleq \omega_0. \quad (6.59)$$

实信号  $s(t)$  频谱

$$S(\omega) = \frac{1}{2} [M(\omega + \omega_0) + M(\omega - \omega_0)]. \quad (6.60)$$

①  $s(t)$  的解析信号

$$\tilde{s}(t) = s(t) + j\hat{s}(t). \quad (6.61)$$

解析信号的频谱

$$\tilde{S}(\omega) = 2S(\omega)U(\omega) = [M(\omega + \omega_0) + M(\omega - \omega_0)]U(\omega). \quad (6.62)$$

②  $s(t)$  的复指数表示为

$$\tilde{s}_{\text{复}}(t) = m(t)e^{j\omega_0 t}. \quad (6.63)$$

复信号频谱

$$\bar{S}_{\text{复}}(\omega) = M(\omega - \omega_0). \quad (6.64)$$

比较复信号频谱与解析信号频谱间的误差, 如图 6-5 所示, 有

$$\begin{aligned} \varepsilon(\omega) &= \tilde{S}_{\text{复}}(\omega) - \tilde{S}(\omega) \\ &= M(\omega - \omega_0)[1 - U(\omega)] - M(\omega + \omega_0)U(\omega). \end{aligned}$$

这个误差在时域的表达式为

$$\epsilon(x) = \frac{1}{\pi} \operatorname{Im} \left[ \int_{-\infty}^{-\omega_0} M(\omega) e^{i(\omega+\omega_0)t} d\omega \right]. \quad (6.65)$$

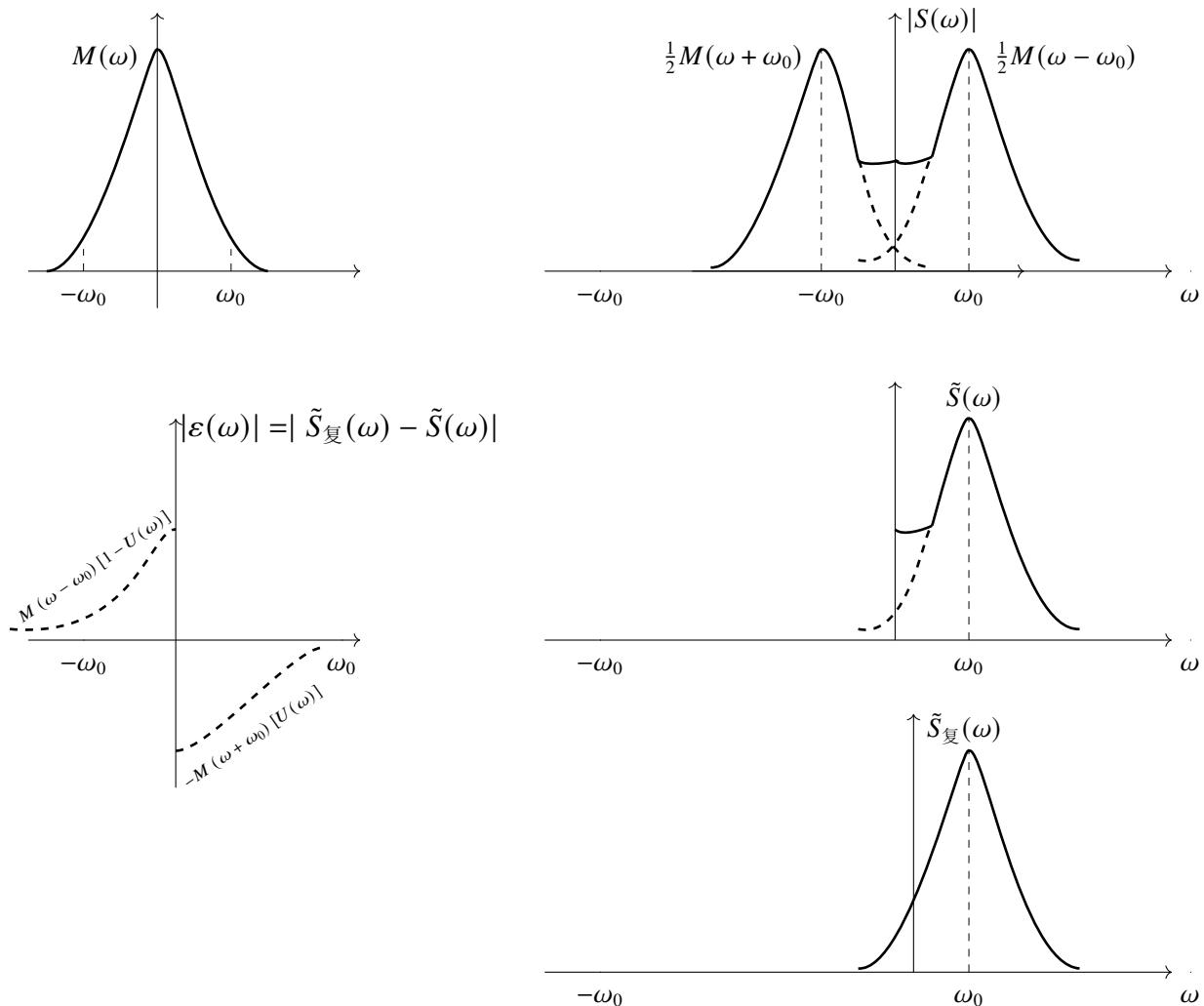


图 6-5 误差分析

### (3) 误差分析

如图 6-5 所示, 误差  $\varepsilon(t)$  的大小取决于  $M(\omega)$  中  $|\omega| > \omega_0$  尾部所包含的能量。若尾部能量为零, 则误差  $\varepsilon(t)$  为零。因此, 若要将  $s(t)$  的解析形式表示成复数形式, 则要求复包络  $m(t)$  必须是个低频限带信号, 即必须满足

$$M(\omega) = 0, \quad |\omega| > \Delta\omega, \quad \Delta\omega < \omega_0 \quad (6.66)$$

也即  $\varepsilon(t) = 0$ 。可见, 此条件与前面从解析形式推导到复指数形式过程中的条件是一致的。

对于一般窄带信号而言, 即使复包络  $m(t)$  不是限带, 但如果能使  $\omega_0 \gg \Delta\omega$ , 使得尾部能量很小, 误差  $\varepsilon(t)$  也很小, 则用  $s(t)$  的复指数信号替代  $s(t)$  的解析信号也是允许的。

当然,对于一个理想的高频窄带信号来讲,由于  $\omega_0 \gg \Delta\omega$ ,且复包络  $m(t)$  又是限带的,可以用指数形式的复信号直接替代其解析信号。

**例 6.8** 若已知高频窄带信号复包络的频谱  $M(\omega)$ ,有

$$\tilde{s}_{\text{复}}(t) = m(t)e^{j\omega_0 t} \xrightarrow[F]{F^{-1}} \tilde{S}_{\text{复}}(\omega) = M(\omega - \omega_0), \quad (6.67)$$

给出高频窄带信号频谱的复指数表示形式。

则

$$\tilde{s}_{\text{复}}^*(t) \xrightarrow[F^{-1}]{F} \tilde{S}_{\text{复}}^*(-\omega) = M^*(-\omega - \omega_0) \quad (6.68)$$

由于可以用复指数形式代替其解析形式  $\tilde{s}(t) = \bar{s}_{\text{复}}(t)$ , 则

$$s(t) = \operatorname{Re}[\tilde{s}(t)] = \operatorname{Re}[\tilde{s}_{\text{复}}(t)] = \frac{1}{2} [\tilde{s}_{\text{复}}(t) + \tilde{s}_{\text{复}}^*(t)] \quad (6.69)$$

则高频窄带信号的频谱可以表示为

$$S(\omega) = \frac{1}{2} [M(\omega - \omega_0) + M^*(-\omega - \omega_0)] \quad (6.70)$$

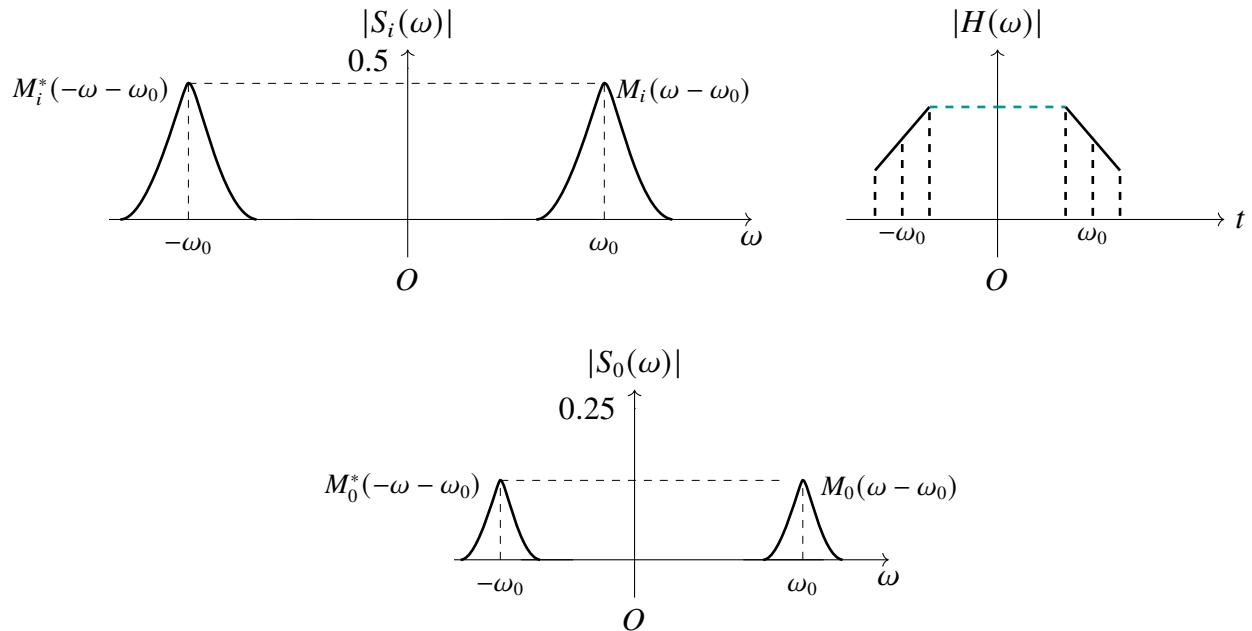


图 6-6 高频窄带信号通过窄带系统

### 6.1.4 高频窄带信号通过窄带系统

窄带信号的复数表示方法同样可以应用到窄带系统上, 以简化对系统的分析。当复包络  $m(t)$  是低频限带信号时, 无论是高频窄带信号还是高频窄带系统, 都可以直接用指数形式替代其解析形式。下面就讨论高频窄带信号通过相同中心频率的高频窄带系统的一种简便分析方法。

设输入的高频窄带信号为

$$s_i(t) = \operatorname{Re} [\tilde{s}_i(t)] = \operatorname{Re} [m_i(t)e^{j\omega_0 t}], \quad (6.71)$$

则其频谱可表示为

$$S_i(\omega) = \frac{1}{2} [M_i(\omega - \omega_0) + M_i^*(-\omega - \omega_0)]. \quad (6.72)$$

高频窄带系统的冲激响应和传递函数为

$$\begin{aligned} h(t) &= \operatorname{Re} [\tilde{h}(t)] = \operatorname{Re} [h_m(t)e^{j\omega_0 t}] \\ H(\omega) &= \frac{1}{2} [H_m(\omega - \omega_0) + H_m^*(-\omega - \omega_0)]. \end{aligned} \quad (6.73)$$

其中  $h_m(t), H_m(a)$  分别为窄带系统的复包络及其频谱。

则系统的输出为

$$\begin{aligned} s_o(t) &= s_i(t) * h(t) \\ S_o(\omega) &= S_i(\omega)H(\omega). \end{aligned} \quad (6.74)$$

将输入窄带信号和窄带系统的频谱代入上式, 得

$$\begin{aligned} S_o(\omega) &= \frac{1}{4} [M_i(\omega - \omega_0) + M_i^*(-\omega - \omega_0)] [H_m(\omega - \omega_0) + H_m^*(-\omega - \omega_0)] \\ &= \frac{1}{4} [M_i(\omega - \omega_0) H_m(\omega - \omega_0) + M_i(\omega - \omega_0) H_m^*(-\omega - \omega_0) \\ &\quad + M_i^*(-\omega - \omega_0) H_m(\omega - \omega_0) + M_i^*(-\omega - \omega_0) H_m^*(-\omega - \omega_0)] \end{aligned} \quad (6.75)$$

由于  $s_i(t)$  的复包络  $m_i(t)$  是低频限带信号, 其频谱  $M_i(\omega)$  如图 6-6 所示

$$\begin{cases} M_i^*(-\omega - \omega_0) H_m(\omega - \omega_0) = 0 \\ M_i(\omega - \omega_0) H_m^*(-\omega - \omega_0) = 0 \end{cases} \quad (6.76)$$

由图可知, 式 (6.76) 显然成立。所以

$$S_o(\omega) = \frac{1}{4} [M_i(\omega - \omega_0) H_m(\omega - \omega_0) + M_i^*(-\omega - \omega_0) H_m^*(-\omega - \omega_0)]. \quad (6.77)$$

相应的输出信号为

$$s_o(t) = \frac{1}{4} \{ [m_i(t) * h_m(t)] e^{j\omega_0 t} + [m_i(t) * h_m(t)]^* e^{-j\omega_0 t} \} \quad (6.78)$$

又因为

$$\tilde{s}_o(t) = \operatorname{Re} [\tilde{s}_o(t)] = \frac{1}{2} [\tilde{s}_o(t) + \tilde{s}_o^*(t)]. \quad (6.79)$$

可得

$$\tilde{s}_o(t) = \frac{1}{2} [m_i(t) * h_m(t)] e^{j\omega_0 t}. \quad (6.80)$$

式 (6.80) 表明, 输出  $s_o(t)$  也是一高频窄带信号, 且其复包络

$$m_o(t) = m_i(t) * \frac{1}{2} h_m(t), \quad (6.81)$$

即输出的复包络  $m_o(t)$  可由输入信号的复包络  $m_i(t)$  与系统冲激响应的复包络  $h_m(t)/2$  求得, 因此, 输出  $s_o(t)$  也可写成

$$s_o(t) = \operatorname{Re} [m_o(t) e^{j\omega_0 t}]. \quad (6.82)$$

结果表明, 一个高频窄带信号通过高频窄带系统, 可以作如图 6-7 所示的等效, 即可以等效为信号的复包络通过一个冲激响应为  $h_m(t)/2$  的低通系统。即输出的复包络  $m_o(t)$  仅由输入信号的复包络  $m_i(t)$  与系统冲激响应的复包络  $h_m(t)/2$  卷积而成。这种处理方法使我们对高频窄带信号通过高频窄带系统这类问题的分析与运算大为简化, 避免了高频信号与高频冲激响应卷积时高频项处理的麻烦。

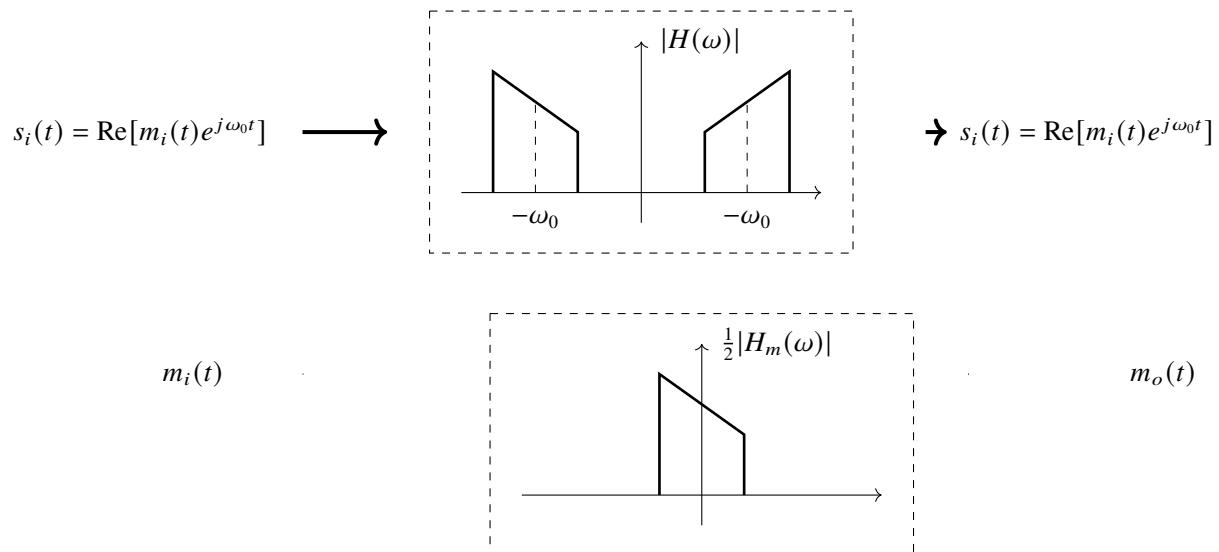


图 6-7 运算的简化

### 6.1.5 随机过程的解析形式及其性质

实随机过程  $X(t)$  的解析形式 (或解析过程) 为

$$\tilde{X}(t) = X(t) + j\hat{X}(t), \quad (6.83)$$

其中

$$\hat{X}(t) = X(t) * \frac{1}{\pi t}. \quad (6.84)$$

称为  $X(t)$  的希尔伯特变换,

由于希尔伯特变换具有线性性质,  $1/\pi t$  可以看成一线性系统的冲激响应。因此,  $\hat{X}(t)$  可以看成是在输入  $X(t)$  下线性系统  $h_{\Lambda}(t)$  的输出, 即

$$Y(t) = X(t) * \frac{1}{\pi t} = \hat{X}(t). \quad (6.85)$$

如图 6-8 所示。正是利用这一等效公式, 使得在下一节中应用解析过程来分析窄带随机信号变得是十分方便。

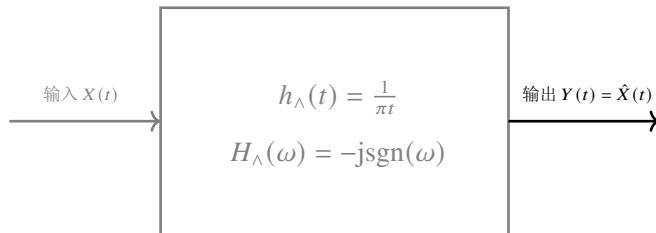


图 6-8 解析过程

既然希尔伯特变换是一种线性变换, 那么随机信号通过线性系统的结论在此解析过程的性质亦可应用。

1° 若  $X(t)$  为宽平稳 (实) 过程, 则  $\hat{X}(t)$  也是宽平 (实) 过程且  $X(t)$  与  $\hat{X}(t)$  联合宽平稳。

2° 实随机过程  $X(t)$  和它的  $\hat{X}(t)$  具有相同的自相关函数和功率谱密度。

证: 从图 6-8 不难看出

$$\begin{aligned} G_{\hat{X}}(\omega) &= G_X(\omega) |H_{\Lambda}(\omega)|^2, \\ |H_{\Lambda}(\omega)| &= |-j\text{sgn}(\omega)| = 1. \end{aligned} \quad (6.86)$$

因此

$$G_{\hat{X}}(\omega) = G_X(\omega) \Rightarrow R_{\hat{X}}(\tau) = R_X(\tau). \quad (6.87)$$

3°  $X(t)$  与  $\hat{X}(t)$  的互相关函数  $R_{X\hat{X}}(\tau)$  等于  $X(t)$  自相关函数的希尔伯特变换。据线性系统输入输出随机信号之间互相关函数的性质, 有

$$\begin{cases} R_{X\hat{X}}(\tau) = E[X(t)\hat{X}(t+\tau)] = R_X(\tau) * h_\Lambda(\tau) = R_X(\tau) * \frac{1}{\pi\tau} = \hat{R}_X(\tau). \\ R_{\hat{X}X}(\tau) = E[\hat{X}(t)X(t+\tau)] = R_X(\tau) * h_\Lambda(-\tau) = -R_X(\tau) * \frac{1}{\pi\tau} = -\hat{R}_X(\tau). \end{cases} \quad (6.88)$$

且有

$$R_{X\hat{X}}(\tau) = -R_{\hat{X}X}(\tau). \quad (6.89)$$

由此可得  $X(t)$  与  $\hat{X}(t)$  的互功率谱密度为

$$\begin{aligned} G_{X\hat{X}}(\omega) &= F[R_{X\hat{X}}(\tau)] = F[\hat{R}_X(\tau)] \\ &= -j\text{sgn}(\omega)G_X(\omega) = \begin{cases} -jG_X(\omega), & \omega > 0 \\ jG_X(\omega), & \omega < 0 \end{cases} \end{aligned} \quad (6.90)$$

4°  $X(t)$  与  $\hat{X}(t)$  的互相关函数是  $\tau$  的奇函数。

证: 由于

$$R_{X\hat{X}}(-\tau) = R_X(-\tau) * h_\Lambda(-\tau). \quad (6.91)$$

且  $R_X(\tau)$  是偶函数, 则

$$R_{X\hat{X}}(-\tau) = R_X(-\tau) * \left(-\frac{1}{\pi\tau}\right) = R_X(\tau) * \left(-\frac{1}{\pi\tau}\right) = -\hat{R}_X(\tau) = -R_{X\hat{X}}(\tau). \quad (6.92)$$

同理可证

$$R_{X\hat{X}}(-\tau) = -R_{\hat{X}X}(\tau). \quad (6.93)$$

5° 随机过程程  $R_{X,\hat{X}}(\tau), R_{X,\hat{X}}(\tau)$  在任何同一时刻的两个状态正交。

证: 因为  $R_{X\hat{X}}(\tau), R_{\hat{X}X}(\tau)$  是  $\tau$  的奇函数, 所以当  $\tau = 0$  时, 有

$$\begin{cases} R_{X\hat{X}}(0) = 0 \\ R_{\hat{X}X}(0) = 0 \end{cases} \quad (6.94)$$

上式说明, 过程  $X(t)$  与  $\hat{X}(t)$  在任何同一时刻 ( $\tau = 0$ ) 的两个状态正交, 即

$$R_{X\hat{X}}(0) = E[X(t)\hat{X}(t)] = 0. \quad (6.95)$$

6° 解析过程的功率谱密度只存在于正频域。

按照复随机过程自相关函数的定义, 解析过程  $\tilde{X}(t)$  的自相关函数为

$$\begin{aligned} R_{\tilde{X}}(\tau) &= E[\tilde{X}^*(t) \cdot \tilde{X}(t+\tau)] \\ &= E\{[X(t) - j\hat{X}(t)][X(t+\tau) + j\hat{X}(t+\tau)]\} \\ &= R_X(\tau) + R_{\hat{X}}(\tau) + j[R_{X\hat{X}}(\tau) - R_{\hat{X}X}(\tau)]. \end{aligned} \quad (6.96)$$

再应用性质 2° 和 3°, 可得

$$R_{\tilde{X}}(\tau) = 2 [R_X(\tau) + j\hat{R}_X(\tau)] = 2\tilde{R}_X(\tau). \quad (6.97)$$

对式 (6.97) 两边求傅里叶变换, 可得解析过程  $\tilde{X}(t)$  的功率谱密度

$$\begin{aligned} G_{\tilde{X}}(\omega) &= 2 \{G_X(\omega) + jF[\hat{R}_X(\tau)]\} \\ &= 2 \{G_X(\omega) + j[-j \operatorname{sgn}(\omega)G_X(\omega)]\} \\ &= 2 [G_X(\omega) + \operatorname{sgn}(\omega)G_X(\omega)] \\ &= \begin{cases} 4G_X(\omega), & \omega > 0 \\ 0, & \omega < 0 \end{cases} \end{aligned} \quad (6.98)$$

上式表明, 解析过程的功率谱密度只存在于正频域, 即它是单边带的功率谱密度其强度等于原实过程功率谱密度强度的 4 倍。 $G_X(\omega)$  与  $G_{\tilde{X}}(\omega)$  的关系如图 6-9 所示。

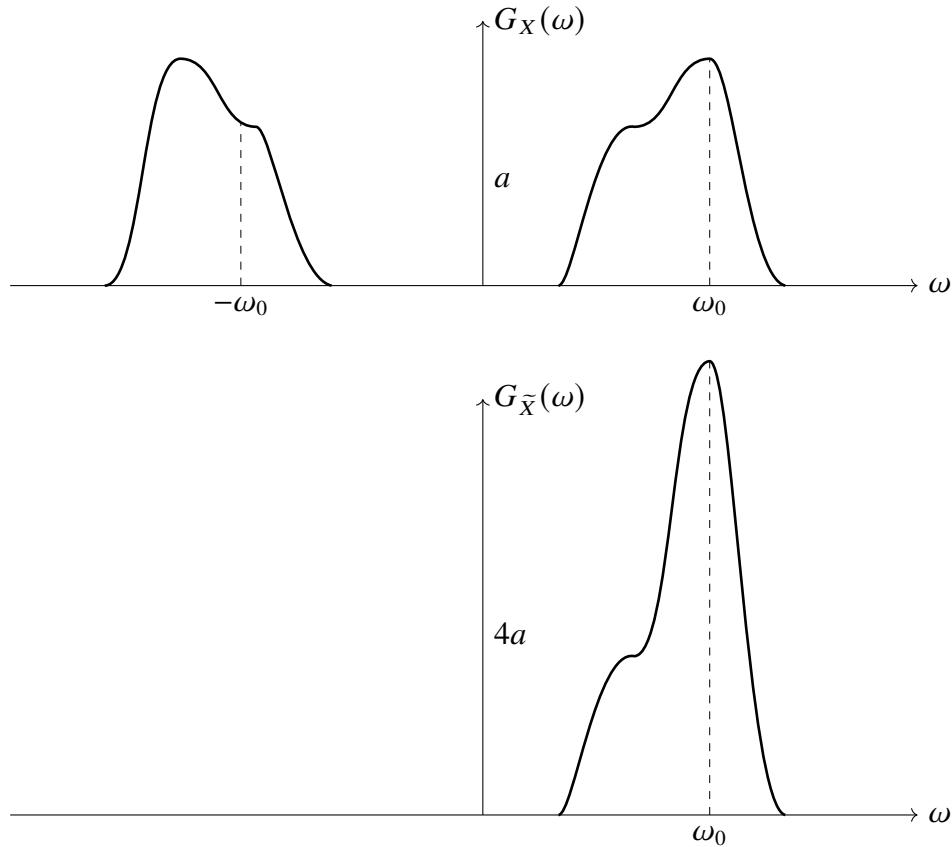


图 6-9 解析过程的功率谱

## 6.2 窄带随机过程

### 6.2.1 窄带随机过程的数学模型及复指数形式

#### 1. 窄带随机过程的数学模型

在雷达、通信等许多电子系统中，通常是用一个宽带随机信号来激励一个窄带滤波器。此时，在滤波器输出端得到的便是一个窄带随机信号，若用一示波器来观测它的某次输出的波形（某个样本），则可以看到，它的样本接近于一个正弦波，但此正弦波的幅度和相位都在作缓慢的随机变化，典型窄带随机过程的功率谱密度及样本函数图形如图 6-10 所示

我们将图 6-10(b) 中的样本函数的图形与图 6-4(b) 所示的波形进行比较，可知窄带随机过程的一个样本函数就是一个高频窄带信号。因此，对于某次观测试验结果，样本函数可写成

$$x_k(t) = a_k(t) \cos [\omega_0 t + \varphi_k(t)], \quad \zeta_k \in \Omega \quad (6.99)$$

而所有样本函数的总体—窄带随机过程，则可写成

$$X(t) = A(t) \cos [\omega_0 t + \Phi(t)]. \quad (6.100)$$

上式就是窄带随机过程常用的数学模型。

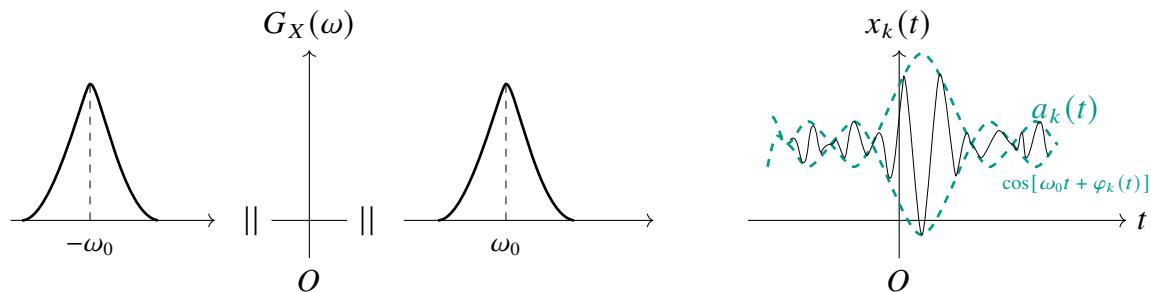


图 6-10 窄带随机过程

由于  $a_k(t)$  和  $\varphi_k(t)$  相对  $\cos \omega_0 t$  来说是慢变化的时间函数，所以  $A(t)$  和  $\Phi(t)$  相对常数来说就是慢变化的随机过程。于是，我们就可以把窄带随机过程看成是一个随机调幅和随机调相的准正弦振荡。

#### 2. 窄带随机过程的复指数形式

若将高频窄带信号的复指数形式应用到窄带随机过程中，则

$$\tilde{X}(t) = A(t)e^{j\Phi(t)}e^{j\omega_0 t} = M(t)e^{j\omega_0 t}, \quad (6.101)$$

式中  $M(t)$  称为  $X(t)$  的复包络,  $A(t)$  称为包络,  $\Phi(t)$  称为相位,  $e^{j\omega_0 t}$  称之为复载频。且  $M(t) = A(t)e^{j\Phi(t)}$ , 如果此窄带随机过  $X(t)$  是平稳过程, 那么用复指数形式表示后, 其统计特性如下:

(1) 自相关函数

$$\begin{aligned} R_{\tilde{X}}(\tau) &= E \left[ \tilde{X}^*(t)\tilde{X}(t+\tau) \right] \\ &= E \left[ M^*(t)e^{-j\omega_0 t} M(t+\tau)e^{j\omega_0(t+\tau)} \right] \\ &= E [M^*(t)M(t+\tau)] e^{j\omega_0 \tau} \\ &= R_M(\tau) e^{j\omega_0 \tau}. \end{aligned} \quad (6.102)$$

(2) 功率谱密度

若傅里叶变换对  $R_M(\tau) \xrightarrow[F]{F^{-1}} G_M(\omega)$ , 则有

$$G_{\tilde{X}}(\omega) = G_M(\omega - \omega_0). \quad (6.103)$$

因为  $R_{\tilde{X}}(\tau) = 2[R_X(\tau) + j\hat{R}_X(\tau)]$ , 可得

$$R_{\tilde{X}}(\tau) + R_{\tilde{X}}^*(\tau) = 2[R_X(\tau) + j\hat{R}_X(\tau)] + 2[R_X(\tau) - j\hat{R}_X(\tau)]. \quad (6.104)$$

由上式解得

$$\begin{aligned} R_X(\tau) &= \frac{1}{4} \left[ R_{\tilde{X}}(\tau) + R_{\tilde{X}}^*(\tau) \right] = \frac{1}{4} [R_M(\tau)e^{j\omega_0 \tau} + R_M^*(-\tau)e^{-j\omega_0 \tau}], \\ G_X(\omega) &= \frac{1}{4} \left[ G_{\tilde{X}}(\omega) + G_{\tilde{X}}^*(-\omega) \right] = \frac{1}{4} [G_M(\omega - \omega_0) + G_M^*(-\omega - \omega_0)]. \end{aligned} \quad (6.105)$$

由此得出  $X(t)$  与  $\tilde{X}(t)$  及  $M(t)$  在频域上的关系, 如图 6-11 所示。

## 6.2.2 窄带随机过程的“垂直”分解

统计分析的对象是随机函数。为了更方便的对窄带随机过程进行统计分析, 先将窄带随机过程中的随机函数与非随机函数以不同的因式表示, 分解开后表达, 故有

$$\begin{aligned} X(t) &= A(t) \cos [\omega_0 t + \Phi(t)] \\ &= A(t) \cos \Phi(t) \cos \omega_0 t - A(t) \sin \Phi(t) \sin \omega_0 t \\ &= A_c(t) \cos \omega_0 t - A_s(t) \sin \omega_0 t = [A_c(t), A_s(t)][\cos \omega_0 t, -\sin \omega_0 t]^T, \end{aligned} \quad (6.106)$$

上式中,  $\cos \omega_0 t$  和  $\sin \omega_0 t$  都是非随机函数, 而随机函数为

$$\begin{cases} A_c(t) = A(t) \cos \Phi(t) \\ A_s(t) = A(t) \sin \Phi(t) \end{cases} \quad (6.107)$$

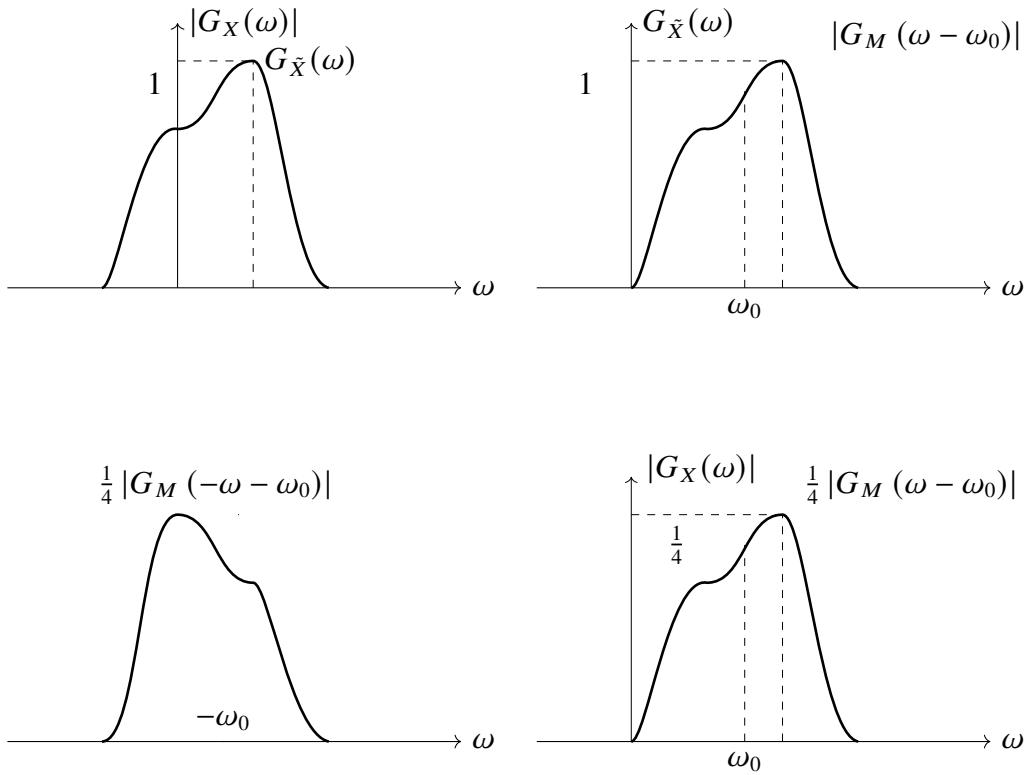


图 6-11 窄带随机过程的功率谱

若将窄带过程的复数形式分解为

$$\tilde{X}(t) = A(t)e^{j\Phi(t)}e^{j\omega_0 t} = M(t)e^{j\omega_0 t}, \quad (6.108)$$

其中  $M(t) = A(t)e^{j\Phi(t)}$ ,  $\tilde{X}(t)$  中的随机分量是  $M(t)$ , 而  $M(t)$  又可分解为

$$M(t) = [A(t) \cos \Phi(t) + jA(t) \sin \Phi(t)] = A_c(t) + jA_s(t), \quad (6.109)$$

其中

$$\begin{cases} A(t) = \sqrt{A_c^2(t) + A_s^2(t)} \\ \Phi(t) = \tan^{-1} \frac{A_s(t)}{A_c(t)} = \arctan \frac{A_s(t)}{A_c(t)} \end{cases} \quad (6.110)$$

可见, 窄带随机过程  $X(t)$  的包络过程的垂直分解  $A(t)$ , 相位  $\Phi(t)$  完全可由  $A_c(t)$  和  $A_s(t)$  确定。且  $A_c(t)$  和  $A_s(t)$  是一对在几何上正交的分量。如图 6-12 所示。

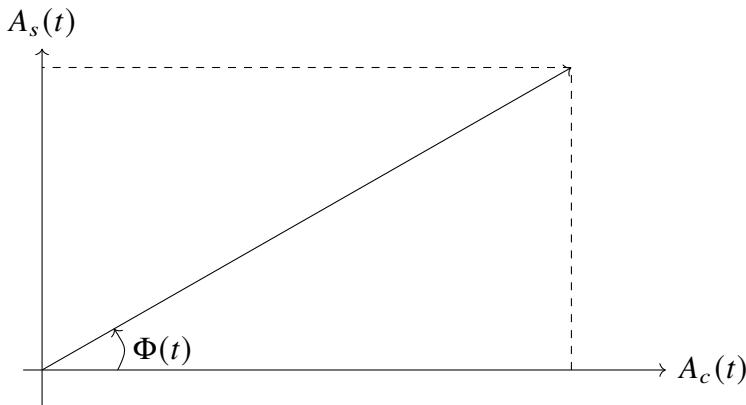


图 6-12 窄带过程的垂直分解

为了与统计意义上的“正交”有所区别,今后我们把几何意义上的“正交”改称为“垂直”,称  $A_c(t)$  和  $A_s(t)$  为一对垂直分量。

经过以上分析可知,  $A_c(t)$  和  $A_s(t)$  包含了窄带随机过程  $X(t)$  的所有随机因素。因此,下面讨论窄带随机过程  $X(t)$  的统计特性, 主要就是讨论这一对垂直分量  $A_c(t)$  和  $A_s(t)$  的统计特性及它们与过程  $X(t)$  之间的统计关系。

在讨论统计特性之前,先导出  $X(t)$ ,  $A_c(t)$  和  $A_s(t)$  之间的函数关系如下

$$\begin{cases} X(t) = A_c(t) \cos \omega_0 t - A_s(t) \sin \omega_0 t \\ \hat{X}(t) = A_c(t) \sin \omega_0 t + A_s(t) \cos \omega_0 t \\ A_c(t) = X(t) \cos \omega_0 t + \hat{X}(t) \sin \omega_0 t \\ A_s(t) = \hat{X}(t) \cos \omega_0 t - X(t) \sin \omega_0 t \end{cases} \quad (6.111)$$

### 6.2.3 窄带随机过程的统计分析

若窄带随机过程  $X(t)$  是零均值平稳的实过程,且功率谱密度如图 6-13 所示,满足

$$G_X(\omega) = \begin{cases} G_X(\omega), & \left( \begin{array}{l} \Omega < \omega - \omega_0 < \Delta\omega - \Omega \\ -\Delta\omega + \Omega < \omega + \omega_0 < \Omega \end{array} \right) \\ 0, & \text{其他} \end{cases} \quad (6.112)$$

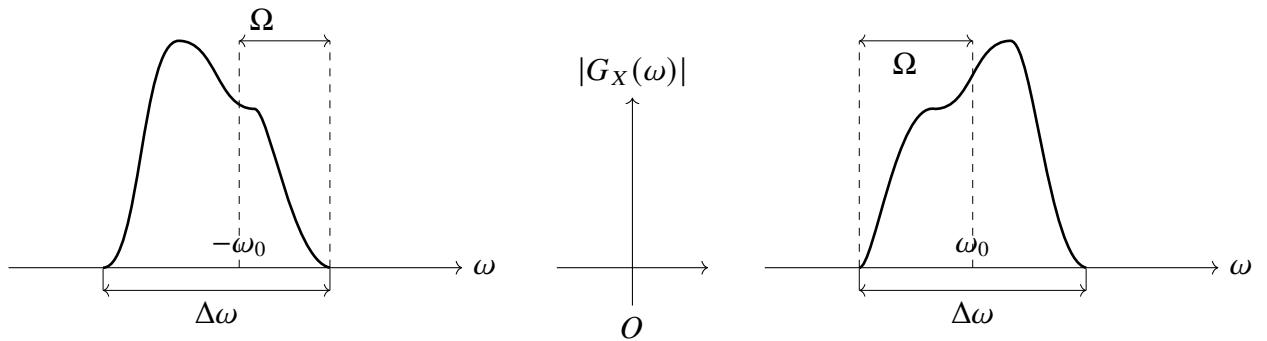


图 6-13 零均值平稳窄带实随机过程的频谱

这里  $\Omega$  和  $\Delta\omega$  皆为正实常数,  $\Delta\omega \ll \omega_0$ 。则  $A_c(t)$  和  $A_s(t)$  这对垂直分量有下面的性质:

- 1)  $A_c(t)$  和  $A_s(t)$  均为实随机过程。
- 2)  $A_c(t)$  和  $A_s(t)$  的期望均为 0, 即满足

$$E [A_c(t)] = E [A_s(t)] = 0. \quad (6.113)$$

- 3)  $A_c(t)$  和  $A_s(t)$  的自相关函数

①  $A_c(t)$  和  $A_s(t)$  各自平稳, 它们的自相关函数为

$$R_{A_c}(\tau) = R_{A_s}(\tau) = R_X(\tau) \cos \omega_0 \tau + \hat{R}_X(\tau) \sin \omega_0 \tau. \quad (6.114)$$

- ② 当  $\tau = 0$  时, 有

$$R_{A_c}(0) = R_{A_s}(0) = R_X(0). \quad (6.115)$$

即

$$E [A_c^2(t)] = E [A_s^2(t)] = E [X^2(t)], \quad (6.116)$$

表示  $X(t)$ ,  $A_c(t)$  和  $A_s(t)$  三者的平均功率皆相等。

由于都是零均值, 因此三者的方差相同, 即

$$\sigma_{A_c}^2 = \sigma_{A_s}^2 = \sigma_X^2. \quad (6.117)$$

- 4)  $A_c(t)$  和  $A_s(t)$  的功率谱密度

$$G_{A_c}(\omega) = G_{A_s}(\omega) = L_p [G_X(\omega - \omega_0) + G_X(\omega + \omega_0)]. \quad (6.118)$$

其中  $L_p[\cdot]$  表示一低通滤波器。

证: 由于

$$\begin{aligned} R_{A_c}(\tau) &= R_X(\tau) \cos \omega_0 \tau + \hat{R}_X(\tau) \sin \omega_0 \tau \\ &= \frac{1}{2} R_X(\tau) [e^{j\omega_0 \tau} + e^{-j\omega_0 \tau}] + \frac{1}{2j} \hat{R}_X(\tau) [e^{j\omega_0 \tau} - e^{-j\omega_0 \tau}]. \end{aligned} \quad (6.119)$$

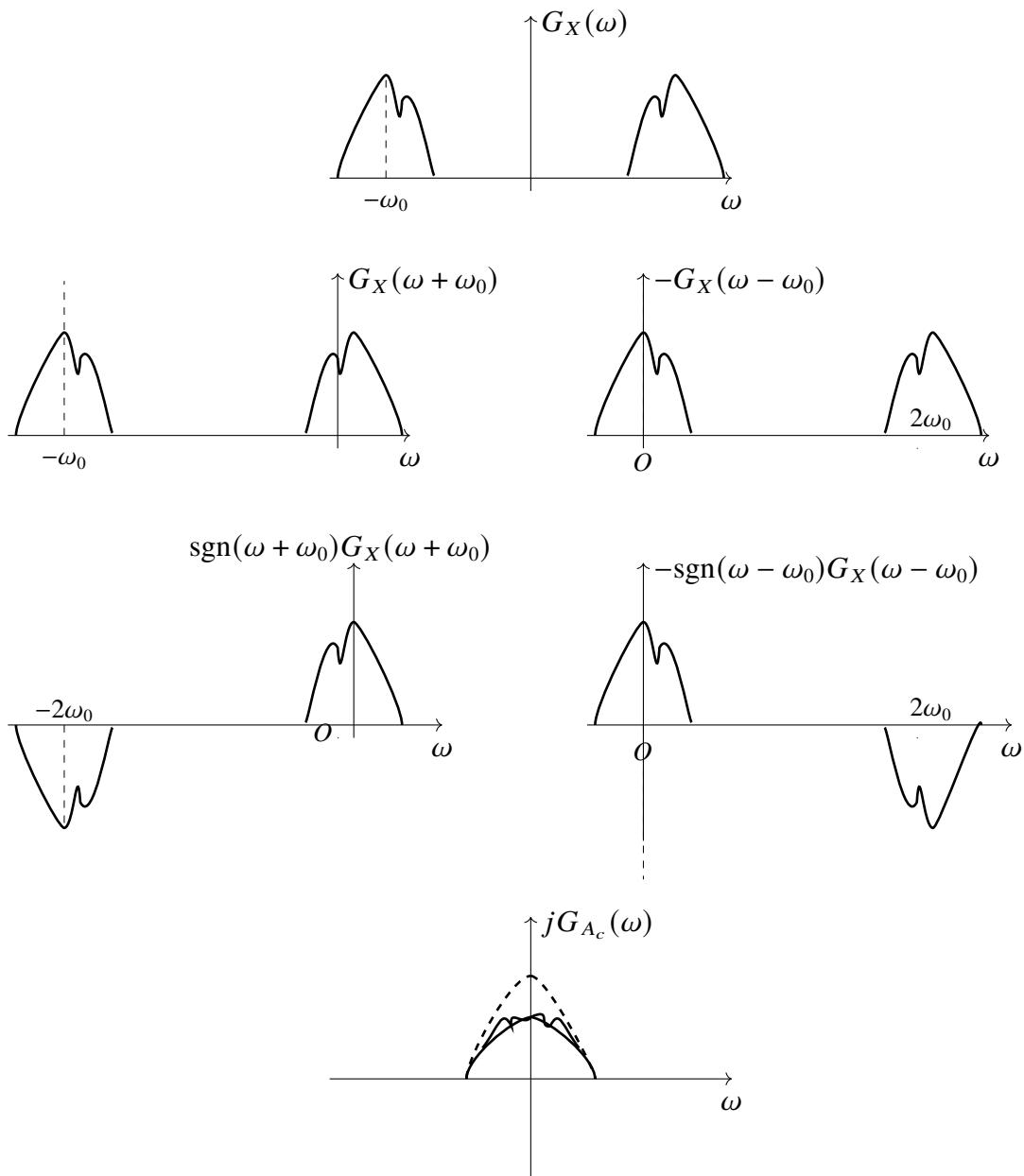


图 6-14 垂直分量的功率谱密度

两边取傅里叶变换，并利用  $F[\hat{R}_X(\tau)] = -j\text{sgn}(\omega)G_X(\omega)$ , 可得

$$\begin{aligned} G_{A_c}(\omega) &= \frac{1}{2} [G_X(\omega - \omega_0) + G_X(\omega + \omega_0)] \\ &\quad + \frac{1}{2} [-\text{sgn}(\omega - \omega_0)G_X(\omega - \omega_0) + \text{sgn}(\omega + \omega_0)G_X(\omega + \omega_0)]. \end{aligned} \quad (6.120)$$

上式各项所对应的功率谱密度如图 6-14 所示。从图形中可以直接得出

$$G_{A_c}(\omega) = L_p[G_X(\omega - \omega_0) + G_X(\omega + \omega_0)]. \quad (6.121)$$

同理可得

$$G_{A_s}(\omega) = L_p[G_X(\omega - \omega_0) + G_X(\omega + \omega_0)]. \quad (6.122)$$

这说明,  $A_c(t)$  和  $A_s(t)$  都是低频限带过程。

5)  $A_c(t)$  和  $A_s(t)$  的互相关函数

①  $A_c(t)$  和  $A_s(t)$  联合平稳, 它们的互相关函数为

$$\begin{cases} R_{A_c A_s}(\tau) = -R_X(\tau) \sin \omega_0 \tau + \hat{R}_X(\tau) \cos \omega_0 \tau \\ R_{A_s A_c}(\tau) = R_X(\tau) \sin \omega_0 \tau - \hat{R}_X(\tau) \cos \omega_0 \tau \\ R_{A_s A_c}(\tau) = -R_{A_c A_s}(\tau) \end{cases} \quad (6.123)$$

② 互相关函数  $R_{A_c A_s}(\tau), R_{A_s A_c}(\tau)$  均是  $\tau$  的奇函数。

$$\begin{cases} R_{A_c A_s}(\tau) = -R_{A_c A_s}(-\tau) \\ R_{A_s A_c}(\tau) = -R_{A_s A_c}(-\tau) \end{cases}. \quad (6.124)$$

③ 当  $\tau = 0$  时, 有

$$R_{A_c A_s}(0) = 0 \quad (6.125)$$

说明随机过程  $A_c(t)$  和  $A_s(t)$  在同一时刻的两个状态 (随机变量) 之间是相互正交的。

因为  $A_c(t)$  和  $A_s(t)$  的均值皆为 0, 所以当  $\tau = 0$  时, 也有

$$C_{A_c A_s}(0) = 0. \quad (6.126)$$

说明随机过程  $A_c(t), A_s(t)$  在同一时刻的两个状态 (随机变量) 之间是不相关的。

6)  $A_c(t)$  和  $A_s(t)$  的互谱密度

① 功率谱密度  $G_X(\omega)$  相对于中心频率  $\omega_0$  非偶对称时, 如图 6-15(a) 所示

$$G_{A_c A_s}(\omega) = -jL_p[G_X(\omega + \omega_0) - G_X(\omega - \omega_0)] = -G_{A_s A_c}(\omega). \quad (6.127)$$

证: 由  $A_c(t)$  和  $A_s(t)$  的互相关函数

$$\begin{aligned} R_{A_c A_s}(\tau) &= -R_X(\tau) \sin \omega_0 \tau + \hat{R}_X(\tau) \cos \omega_0 \tau \\ &= -\frac{1}{2j} R_X(\tau) [e^{j\omega_0 \tau} - e^{-j\omega_0 \tau}] + \frac{1}{2} \hat{R}_X(\tau) [e^{j\omega_0 \tau} + e^{-j\omega_0 \tau}]. \end{aligned} \quad (6.128)$$

两边取傅里叶变换, 并利用  $F[\hat{R}_X(\tau)] = -j \text{sgn}(\omega) G_X(\omega)$ , 可得

$$\begin{aligned} G_{A_c A_s}(\omega) &= -\frac{1}{2j} [G_X(\omega - \omega_0) - G_X(\omega + \omega_0)] \\ &\quad + \frac{1}{2} [-j \text{sgn}(\omega - \omega_0) G_X(\omega - \omega_0) - j \text{sgn}(\omega + \omega_0) G_X(\omega + \omega_0)]. \end{aligned} \quad (6.129)$$

$$\begin{aligned} jG_{A_c A_s}(\omega) &= -\frac{1}{2} [G_X(\omega - \omega_0) - G_X(\omega + \omega_0)] \\ &\quad + \frac{1}{2} [\text{sgn}(\omega - \omega_0) G_X(\omega - \omega_0) + \text{sgn}(\omega + \omega_0) G_X(\omega + \omega_0)] \end{aligned} \quad (6.130)$$

上式各项所对应的功率谱密度图形如图 6-15 所示。

从图上易证

$$jG_{A_c A_s}(\omega) = L_p [G_X(\omega + \omega_0) - G_X(\omega - \omega_0)]. \quad (6.131)$$

即有

$$G_{A_c A_s}(\omega) = -jL_p [G_X(\omega + \omega_0) - G_X(\omega - \omega_0)]. \quad (6.132)$$

同理可得

$$G_{A_s A_c}(\omega) = jL_p [G_X(\omega + \omega_0) - G_X(\omega - \omega_0)]. \quad (6.133)$$

②  $G_X(\omega)$  相对于中心频率  $\omega$  偶对称时, 如图 6-16(a) 所示。

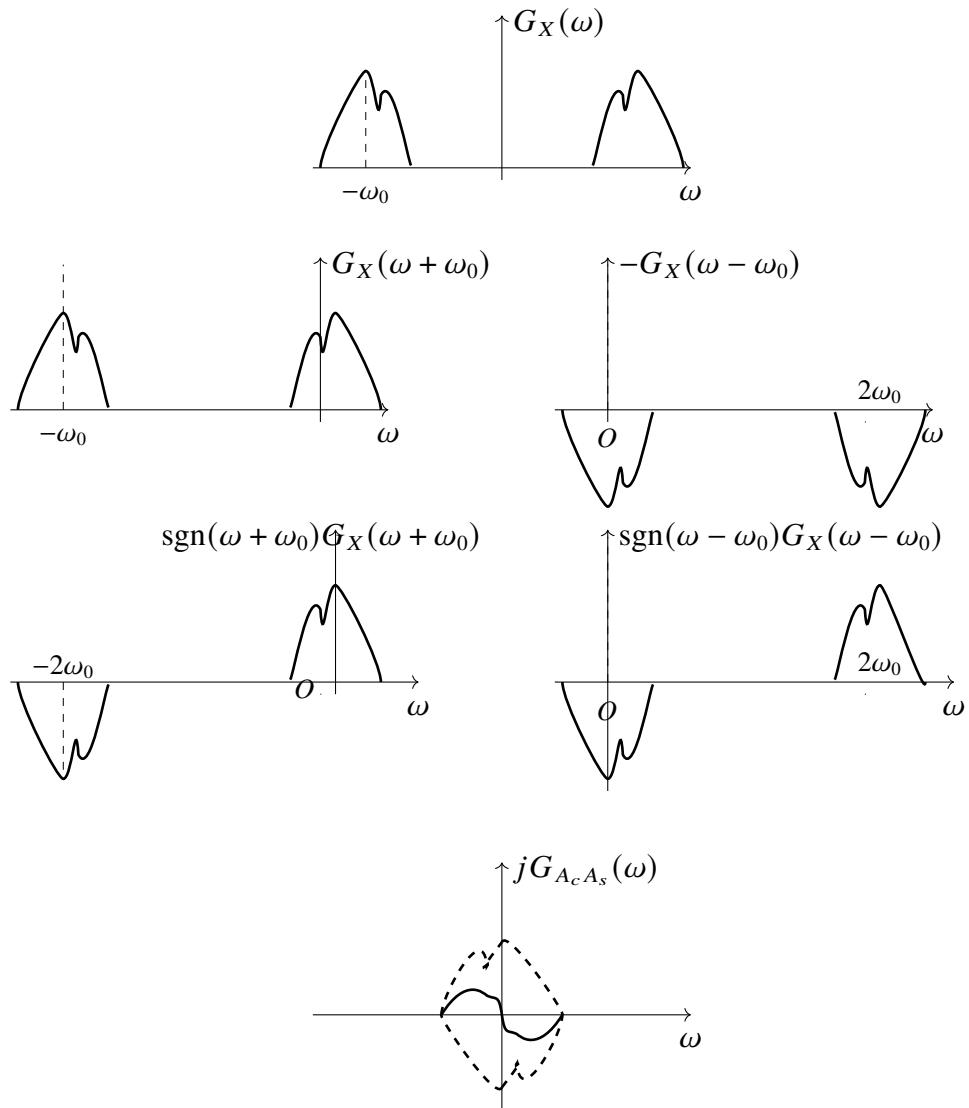


图 6-15 垂直分量的互谱密度

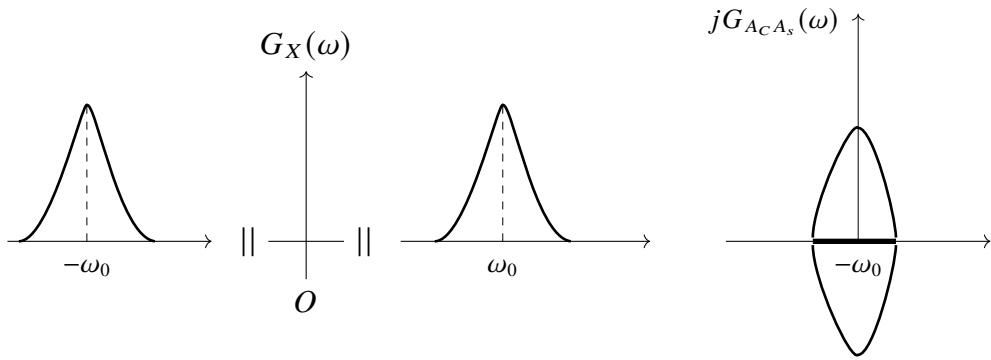
从图中可以看出, 各项叠加后恰好互相抵消, 如图 6-16(b) 所示。所以有

$$G_{A_c A_s}(\omega) = 0. \quad (6.134)$$

由上式可推出, 对任意  $\tau$  值

$$R_{A_c A_s}(\tau) = 0. \quad (6.135)$$

说明当  $X(t)$  具有对称于的功率谱密度时, 随机过程  $A_c(t)$  和  $A_s(t)$  正交, 此处正交是指随机过程正交, 当然可以推出  $A_c(t)$  和  $A_s(t)$  在同一时刻也是正交的。

图 6-16  $G_X(\omega)$  相对于中心频率  $\omega_0$  偶对称

由于  $A_c(t)$  和  $A_s(t)$  的均值皆为 0, 同样可以证明: 当  $X(t)$  具有对称于  $\omega_0$  的功率谱密度时, 两个随机过程  $A_c(t)$  和  $A_s(t)$  互不相关。仿照上述性质的证明方法, 可以证明 (过程略), 窄带随机过程  $X(t)$  的自相关函数为

$$R_X(\tau) = R_{A_c}(\tau) \cos \omega_0 \tau - R_{A_c A_s}(\tau) \sin \omega_0 \tau. \quad (6.136)$$

### 6.3 窄带高斯过程包络与相位的分布

在许多电子系统或电路中, 经常遇到用一个宽带随机过程  $N(t)$  激励一个高频窄带系统(或简称窄带滤波器)的情况, 如图 6-17 所示。由第四章的结论可知, 该情况下系统输出的随机过程都可以认为是一个窄带高斯过程。因此, 窄带高斯过程是通信和电子系统中最常遇到的随机信号模型, 研究窄带高斯过程具有实际意义。

根据上一节 6.2 的分析, 可将任一平稳窄带高斯过程  $X(t)$  表示为准正弦振荡的形式

$$X(t) = A(t) \cos [\omega_0 t + \Phi(t)], \quad (6.137)$$

其中  $A(t)$  和  $\Phi(t)$  分别是  $X(t)$  的包络和相位, 它们都是低频限带过程;  $\omega_0$  是  $X(t)$  的载波频率。

许多实际应用中, 常常需要检测出包络  $A(t)$  或相位  $\Phi(t)$  的信息。若将窄带随机信号  $X(t)$  送入一个包络检波器, 则在检波器输出端可得到包络  $A(t)$ ; 若将  $X(t)$  送入一个相位检波器, 便可检测出  $X(t)$  的相位信息, 如图 6-17 所示。由于  $A(t)$  和  $\Phi(t)$  都是  $X(t)$  的非线性变换, 推导它们的多维概率密度函数十分困难。本书将只限于推导它们的一维和二维概率密度函数。

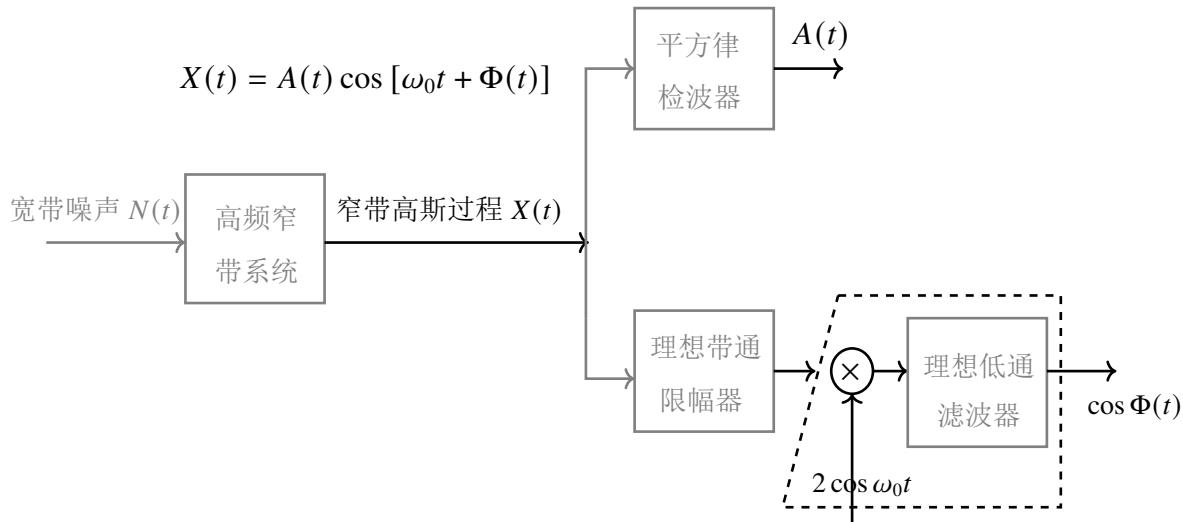


图 6-17 窄带高斯过程的包络与相位

### 6.3.1 包络和相位的一维概率分布

#### 例 6.3.1

设  $X(t)$  是一个窄带平稳高斯实随机过程, 具有零均值和方差  $\sigma^2$ 。现求其包络  $A(t)$  和  $\Phi(t)$  的一维概率密度。

**分析:** 任一给定的时刻, 对  $A(t)$  和  $\Phi(t)$  采样, 便可得到随机变量  $A_t$  和  $\Phi_t$ 。求  $X(t)$  包络  $A(t)$  和  $\Phi(t)$  的一维概率密度就是求  $A_t$  和  $\Phi_t$  的概率密度  $f_A(a_t)$  和  $f_\varphi(\varphi_t)$ 。

由窄带过程的垂直分解,  $X(t)$  可以表示成

$$X(t) = A(t) \cos [\omega_0 t + \Phi(t)] = A_c(t) \cos \omega_0 t - A_s(t) \sin \omega_0 t. \quad (6.138)$$

由上节内容,  $A(t)$ ,  $\Phi(t)$  和  $A_c(t)$  和  $A_s(t)$  具有如下关系

$$\begin{cases} A(t) = \sqrt{A_c^2(t) + A_s^2(t)} \\ \Phi(t) = \tan^{-1} \frac{A_s(t)}{A_c(t)} = \arctan \frac{A_s(t)}{A_c(t)} \end{cases}. \quad (6.139)$$

则  $A_t$  和  $\Phi_t$  以及  $A_{ct}$  和  $A_{st}$  满足

$$\begin{cases} A_t = g_1(A_{ct}, A_{st}) = \sqrt{A_{ct}^2 + A_{st}^2} \\ \Phi_t = g_2(A_{ct}, A_{st}) = \arctan \frac{A_{st}}{A_{ct}} \end{cases}. \quad (6.140)$$

其中  $A_{ct}$  和  $A_{st}$  为垂直分量  $A_c(t)$  和  $A_s(t)$  在固定时刻的采样, 也都是随机变量。则反变换关系为

$$\begin{cases} A_{ct} = h_1(A_t, \Phi_t) = A_t \cos \Phi_t \\ A_{st} = h_2(A_t, \Phi_t) = A_t \sin \Phi_t \end{cases}. \quad (6.141)$$

求解思路:① 首先根据已知的窄带高斯过程垂直分量  $A_c(t), A_s(t)$  的统计特性, 来研究  $A_{ct}$  和  $A_{st}$  的统计特性, 从而得到  $A_{ct}A_{st}$  的联合概率密度  $f_{A_cA_s}(a_{ct}, a_{st})$ 。② 从  $f_{A_cA_s}(a_{ct}, a_{st})$  出发, 利用雅可比变换, 得到  $A_t, \Phi_t$  的联合概率密度  $f_{A\Phi}(a_t, \varphi_t)$ 。③ 最后对的联合概率密度  $f_{A\Phi}(a_t, \varphi_t)$  积分, 求出边缘概率密度  $f_A(a_t)$  和  $f_\Phi(\varphi_t)$ 。

具体的步骤为:

① 求  $f_{A_cA_s}(a_{ct}, a_{st})$

首先根据已知的窄带高斯过程垂直分量  $A_c(t), A_s(t)$  的统计特性, 来研究  $A_{ct}, A_{st}$  以下的统计特性, 从而得到  $A_{ct}, A_{st}$  的联合概率密度  $f_{A_cA_s}(a_{ct}, a_{st})$ 。

再由如下 4 个性质:

**性质 1°**  $A_{ct}$  和  $A_{st}$  都是高斯随机变量。

已知  $X(t)$  是一个平稳高斯过程, 由于  $\hat{X}(t)$  是  $X(t)$  的线性变换, 所以  $\hat{X}(t)$  也是平稳高斯过程。 $A_c(t), A_s(t)$  均为  $X(t), \hat{X}(t)$  的线性组合, 故  $A_c(t), A_s(t)$  也是平稳高斯过程, 所以  $A_{ct}, A_{st}$  均为高斯变量。

**性质 2°**  $A_{ct}$  和  $A_{st}$  的均值皆为零。

根据  $A_c(t), A_s(t)$  的性质, 有  $E[A_c(t)] = E[A_s(t)] = 0$ , 所以

$$E[A_{ct}] = E[A_{st}] = 0. \quad (6.142)$$

**性质 3°**  $A_{ct}$  和  $A_{st}$  具有相同的方差, 并且等于  $X(t)$  的方差  $\sigma^2$ 。

根据  $A_c(t)$  和  $A_s(t)$  的性质, 有  $\sigma_{A_c}^2 = \sigma_{A_s}^2 = \sigma_X^2$ , 所以

$$\sigma_{A_c}^2 = \sigma_{A_s}^2 = \sigma_X^2 = \sigma^2. \quad (6.143)$$

**性质 4°**  $A_{ct}$  和  $A_{st}$  相互独立。

根据  $A_c(t)$  和  $A_s(t)$  的性质可知, 同一时刻的两个状态互不相关, 满足  $C_{A_cA_s}(0) = 0$ 。即  $A_{ct}$  和  $A_{st}$  互不相关。而对于高斯随机变量来说, 互不相关与统计独立等价,  $A_{ct}$  和  $A_{st}$  相互独立。

根据以上的性质可知,  $A_{ct}$  和  $A_{st}$  是均值皆为零, 方差皆为  $\sigma^2$  的高斯变量, 且相互独

立。则其联合概率密度为

$$\begin{aligned} f_{A_c A_s}(a_{ct}, a_{st}) &= f_{A_c}(a_{ct}) f_{A_s}(a_{st}) \\ &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{a_{ct}^2}{2\sigma^2}\right] \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{a_{st}^2}{2\sigma^2}\right] \\ &= \frac{1}{2\pi\sigma^2} \exp\left[-\frac{a_{ct}^2 + a_{st}^2}{2\sigma^2}\right]. \end{aligned} \quad (6.144)$$

② 求  $f_{A\Phi}(a_t, \varphi_t)$

从  $f_{A_c A_s}(a_{ct}, a_{st})$  出发, 利用雅可比变换, 得  $A_t$  和  $\Phi_t$  的联合概率密度  $f_{A\Phi}(a_t, \varphi_t)$ 。根据雅可比变换

$$f_{A\Phi}(a_t, \varphi_t) = |J| f_{A_c A_s}(a_t \cos \varphi_t, a_t \sin \varphi_t), \quad (6.145)$$

其中雅可比行列式为

$$J = \begin{vmatrix} \frac{\partial h_1}{\partial a_t} & \frac{\partial h_1}{\partial \varphi_t} \\ \frac{\partial h_2}{\partial a_t} & \frac{\partial h_2}{\partial \varphi_t} \end{vmatrix} = \begin{vmatrix} \cos \varphi_t & -a_t \sin \varphi_t \\ \sin \varphi_t & a_t \cos \varphi_t \end{vmatrix} = a_t \geq 0. \quad (6.146)$$

则  $A_t$  和  $\Phi_t$  的联合概率密度为

$$f_{A\Phi}(a_t, \varphi_t) = \begin{cases} \frac{a_t}{2\pi\sigma^2} \exp\left\{-\frac{a_t^2}{2\sigma^2}\right\}, & a_t \geq 0, 0 \leq \varphi_t \leq 2\pi \\ 0, & \text{其他} \end{cases}. \quad (6.147)$$

由上式可见,  $A_t$  和  $\Phi_t$  的联合概率密度  $f_{A\Phi}(a_t, \varphi_t)$  与  $\Phi_t$  无关。

③ 求  $f_A(a_t)$  和  $f_\Phi(\varphi_t)$ 。

最后对  $A_t$  和  $\Phi_t$  的联合概率密度  $f_{A\Phi}(a_t, \varphi_t)$  积分, 求边缘概率密度  $f_A(a_t)$  和  $f_\Phi(\varphi_t)$ 。

①  $A_t$  的概率密度为

$$\begin{aligned} f_A(a_t) &= \int_0^{2\pi} f_{A\Phi}(a_t, \varphi_t) d\varphi_t \\ &= f_{A\Phi}(a_t, \varphi_t) \int_0^{2\pi} 1 \cdot d\varphi_t \\ &= \frac{a_t}{\sigma^2} \exp\left\{-\frac{a_t^2}{2\sigma^2}\right\}, \quad a_t \geq 0. \end{aligned} \quad (6.148)$$

上式给出了包络  $A(t)$  的一维概率密度表达式, 通常称它为瑞利概率密度或简称瑞利分布。

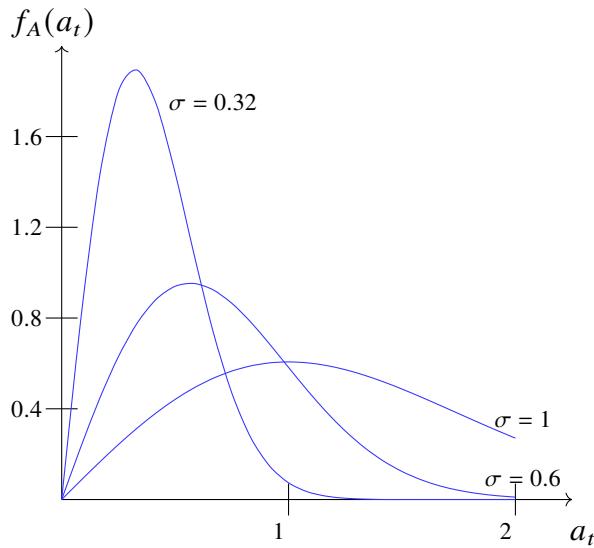


图 6-18 瑞利分布

(1)  $\Phi_t$  的概率密度为

$$f_\phi(\varphi_t) = \int_0^\infty f_{A\Phi}(a_t, \varphi_t) da_t = \frac{1}{2\pi}, 0 \leq \varphi_t \leq 2\pi. \quad (6.149)$$

可见, 相位  $\Phi(t)$  的一维概率密度为  $(0, 2\pi)$  区间上的均匀分布。

(2) 由  $A_t$  和  $\Phi_t$  的概率密度表达式, 易得

$$f_{A\Phi}(a_t, \varphi_t) = f_A(a_t) f_\Phi(\varphi_t). \quad (6.150)$$

上式说明, 随机变量  $A_t \Phi_t$  相互独立, 即随机过程  $A(t)$  与  $\Phi(t)$  在同一时刻相互独立。但应注意, 这并不意味着随机过程  $A(t)$  与  $\Phi(t)$  相互独立。

### 6.3.2 包络和相位各自的二维概率分布

若平稳高斯过程  $X(t)$  在  $t_1$  和  $t_2$  时刻采样为

$$\begin{cases} X(t_1) = A(t_1) \cos [\omega_0 t_1 + \Phi(t_1)] = A_c(t_1) \cos \omega_0 t_1 - A_s(t_1) \sin \omega_0 t_1 \\ X(t_2) = A(t_2) \cos [\omega_0 t_2 + \Phi(t_2)] = A_c(t_2) \cos \omega_0 t_2 - A_s(t_2) \sin \omega_0 t_2 \end{cases}. \quad (6.151)$$

为了便于表示, 把包络  $A(t)$  在  $t_1, t_2$  时刻采样的随机变量  $A(t_1), A(t_2)$  简记为  $A_1, A_2$ ; 把相位  $\Phi(t)$  在  $t_1, t_2$  时刻采样的随机变量  $\Phi(t_1), \Phi(t_2)$  简记为  $\Phi_1, \Phi_2$ ; 把垂直分量  $A_c(t)$  在  $t_1, t_2$  时刻采样的随机变量  $A_c(t_1), A_c(t_2)$  简记为  $A_{c1}, A_{c2}$ ; 垂直分量  $A_s(t)$  在  $t_1, t_2$  时刻采

样的随机变量  $A_s(t_1), A_s(t_2)$  简记为  $A_{s1}, A_{s2}$ 。现在求随机过程  $A(t)$  与  $\Phi(t)$  各自的二维概率密度  $f_A(a_1, a_2)$  和  $f_\Phi(\varphi_1, \varphi_2)$ 。

步骤 1: 求  $f_{A_c A_s}(a_{c1}, a_{s1}, a_{c2}, a_{s2})$ 。

由上已知  $A_c(t), A_s(t)$  都是平稳高斯过程, 具有零均值和相同的方差  $\sigma^2$ 。因此, 可得  $A_{c1}, A_{s1}, A_{c2}, A_{s2}$  为同分布的高斯变量, 具有相同的零均值和相同的方差  $\sigma^2$ 。现在, 我们讨论一种最简单, 也是实际中最常见的情况, 即假定窄带随机过程  $X(t)$  的功率谱密度  $G_X(\omega)$  关于载波频率  $\omega_0$  偶对称, 如图 6-16 所示。根据 6.2.3 节的分析可知, 两个随机过程  $A_c(t), A_s(t)$  正交且互不相关, 满足  $R_{A_c A_s}(\tau) = 0$ , 皆为高斯过程, 则两个过程  $A_c(t), A_s(t)$  是统计独立的。所以四维高斯变量  $(A_{c1}, A_{s1}, A_{c2}, A_{s2})$  的联合概率密度, 可以用二维概率密度来表示, 即二维高斯变量  $(A_{c1}, A_{s2})$  的协方差矩阵为

$$C = \begin{bmatrix} C_{A_{c1} A_{c1}} & C_{A_{c1} A_{c2}} \\ C_{A_{c2} A_{c1}} & C_{A_{c2} A_{c2}} \end{bmatrix} = \begin{bmatrix} C_{A_c}(0) & C_{A_c}(\tau) \\ C_{A_c}(-\tau) & C_{A_c}(0) \end{bmatrix} = \begin{bmatrix} \sigma^2 & R_{A_c}(\tau) \\ R_{A_c}(\tau) & \sigma^2 \end{bmatrix}. \quad (6.152)$$

上式应用了平稳高斯过程  $A_c(t)$  的  $C_{A_c}(\tau) = R_{A_c}(\tau)$  和  $C_{A_c}(0) = \sigma^2$  的性质。因为二维高斯变量  $(X, Y)$  的联合概率密度形式为

$$f_{XY}(x, y) = \frac{1}{2\pi\sqrt{\sigma_X^2\sigma_Y^2 - C_{XY}^2}} \exp\left[-\frac{\sigma_Y^2(x - m_X)^2 - 2C_{XY}(x - m_X)(y - m_Y) + \sigma_X^2(y - m_Y)^2}{2(\sigma_X^2\sigma_Y^2 - C_{XY}^2)}\right]. \quad (6.153)$$

所以二维高斯变量  $(A_{c1}, A_{c2})$  的联合概率密度为

$$f_{A_c}(a_{c1}, a_{c2}) = \frac{1}{2\pi\sqrt{\sigma^4 - R_{A_c}^2(\tau)}} \exp\left[-\frac{\sigma^2 a_{c1}^2 - 2R_{A_c}(\tau)a_{c1}a_{c2} + \sigma^2 a_{c2}^2}{2[\sigma^4 - R_{A_c}^2(\tau)]}\right]. \quad (6.154)$$

同理可得, 二维高斯变量  $(A_{s1}, A_{s2})$  的联合概率密度为

$$f_{A_s}(a_{s1}, a_{s2}) = \frac{1}{2\pi\sqrt{\sigma^4 - R_{A_s}^2(\tau)}} \exp\left[-\frac{\sigma^2 a_{s1}^2 - 2R_{A_s}(\tau)a_{s1}a_{s2} + \sigma^2 a_{s2}^2}{2[\sigma^4 - R_{A_s}^2(\tau)]}\right]. \quad (6.155)$$

应用性质  $R_{A_c}(\tau) = R_A(\tau)$ , 可得四维高斯变量  $(A_{c1}, A_{c2}, A_{s1}, A_{s2})$  的联合概率密度为

$$\begin{aligned} f_{A_c A_s}(a_{c1}, a_{s1}, a_{c2}, a_{s2}) &= f_{A_c}(a_{c1}, a_{c2}) f_{A_s}(a_{s1}, a_{s2}) \\ &= \frac{1}{(2\pi)^2 [\sigma^4 - R_{A_s}^2(\tau)]} \exp\left[-\frac{\sigma^2 (a_{c1}^2 + a_{c2}^2 + a_{s1}^2 + a_{s2}^2) - 2R_{A_c}(\tau)(a_{c1}a_{c2} + a_{s1}a_{s2})}{2[\sigma^4 - R_{A_s}^2(\tau)]}\right]. \end{aligned} \quad (6.156)$$

步骤 2: 求  $f_{A\Phi}(a_1, \varphi_1, a_2, \varphi_2)$

$A_{cl}, A_{s1}, A_{c2}, A_{s2}$  和  $A_1, \Phi_1, A_2, \Phi_2$  的关系为

$$\begin{cases} A_{c1} = h_1(A_1, \Phi_1) = A_1 \cos \Phi_1 \\ A_{n1} = h_2(A_1, \Phi_1) = A_1 \sin \Phi_1 \\ A_{c2} = h_3(A_2, \Phi_2) = A_2 \cos \Phi_2 \\ A_{s2} = h_4(A_2, \Phi_2) = A_2 \sin \Phi_2 \end{cases} . \quad (6.157)$$

雅可比行列式为

$$J = \begin{bmatrix} \cos \varphi_1 & -A_1 \sin \varphi_1 & 0 & 0 \\ \sin \varphi_1 & A_1 \cos \varphi_1 & 0 & 0 \\ 0 & 0 & \cos \varphi_2 & -A_2 \sin \varphi_2 \\ 0 & 0 & \sin \varphi_2 & A_2 \cos \varphi_2 \end{bmatrix} = a_1 a_2 \geq 0. \quad (6.158)$$

四维随机变量  $(A_1, \Phi_1, A_2, \Phi_2)$  的联合概率密度为

$$\begin{aligned} f_{A\Phi}(a_1, \varphi_1, a_2, \varphi_2) &= |J| f_{A_c A_s}(a_{c1}, a_{s1}, a_{c2}, a_{s2}) \\ &= \begin{cases} \frac{a_1 a_2}{(2\pi)^2 [\sigma^4 - R_{A_s}^2(\tau)]} \exp \left[ -\frac{\sigma^2(a_1^2 + a_2^2) - 2R_{A_c}(\tau)a_1 a_2 \cos(\varphi_2 - \varphi_1)}{2[\sigma^4 - R_{A_s}^2(\tau)]} \right] \\ a_1, a_2 \geq 0; 0 \leq \varphi_1, \varphi_2 \leq 2\pi \\ 0, \text{ 其他} \end{cases} . \quad (6.159) \end{aligned}$$

步骤 3: 求  $A(t)$  和  $\Phi(t)$  各自的二维联合概率密度  $f_A(a_1, a_2), f_\Phi(\varphi_1, \varphi_2)$

$$\begin{aligned} f_A(a_1, a_2) &= \int_0^{2\pi} \int_0^{2\pi} f_{A\Phi}(a_1, \varphi_1, a_2, \varphi_2) d\varphi_1 d\varphi_2 \\ &= \begin{cases} \frac{a_1 a_2}{[\sigma^4 - R_{A_s}^2(\tau)]} I_0 \left\{ \frac{a_1 a_2 R_{A_c}(\tau)}{[\sigma^4 - R_{A_s}^2(\tau)]} \right\} \exp \left\{ -\frac{\sigma^2(a_1^2 + a_2^2)}{2[\sigma^4 - R_{A_s}^2(\tau)]} \right\}, a_1, a_2 \geq 0 \\ 0 \end{cases} . \quad (6.160) \end{aligned}$$

式  $I_0(x) = \frac{1}{2\pi} \int_0^{2\pi} \exp[x \cos \varphi] d\varphi$ , 是第一类零阶修正贝塞尔函数,

$$f_\Phi(\varphi_1, \varphi_2) = \begin{cases} \frac{[\sigma^4 - R_{A_s}^2(\tau)]}{4\pi\sigma^2} \left[ \frac{(1-\beta)^{\frac{1}{2}} + \beta(\pi - \arccos \beta)}{(1-\beta^2)^{\frac{3}{2}}} \right], 0 \leq \varphi_1, \varphi_2 \leq 2\pi \\ 0, \text{ 其他} \end{cases} , \quad (6.161)$$

其中

$$\beta = \frac{R_{A_c}(\tau)}{\sigma^2} \cos(\varphi_2 - \varphi_1) \quad (6.162)$$

若令  $q = \varphi_2$ , 不难看出

$$f_{A\Phi}(a_1, \varphi_1, a_2, \varphi_2) \neq f_A(a_1, a_2) f_\Phi(\varphi_1, \varphi_2). \quad (6.163)$$

这一特例说明, 窄带随机过程的包络  $A(t)$  和相位  $\Phi(t)$  彼此不是独立的。

### 6.3.3 随相余弦信号与窄带高斯噪声之和的包络及相位的概率分布

假设

$$X(t) = s(t) + N(t), \quad (6.164)$$

其中信号  $s(t)$  为随相余弦信号

$$s(t) = a \cos(\omega_0 t + \theta), \quad (6.165)$$

式中  $a, \omega_0$  是已知常数, 随机变量  $\theta$  服从  $(0, 2\pi)$  区间上的均匀分布。其中噪声  $N(t)$  为功率谱密度关于中心频率偶对称、零均值、方差为  $\sigma^2$  的平稳窄带高斯实过程。

显然这里  $X(t)$  是一个窄带随机过程。若要求  $X(t)$  的包络及相位的概率密度函数, 则可以仿照前面的方法, 将窄带高斯噪声  $N(t)$  表示为

$$N(t) = A_c(t) \cos \omega_0 t - A_s(t) \sin \omega_0 t. \quad (6.166)$$

将随相余弦信号  $s(t)$  表示为

$$s(t) = a \cos \theta \cos \omega_0 t - a \sin \theta \sin \omega_0 t. \quad (6.167)$$

代入  $X(t)$ , 得

$$X(t) = s(t) + N(t) = [a \cos \theta + A_c(t)] \cos \omega_0 t - [a \sin \theta + A_s(t)] \sin \omega_0 t. \quad (6.168)$$

令

$$\begin{cases} A'_c(t) = a \cos \theta + A_c(t) \\ A'_s(t) = a \sin \theta + A_s(t) \end{cases}. \quad (6.169)$$

则

$$X(t) = A'_c(t) \cos \omega_0 t - A'_s(t) \sin \omega_0 t. \quad (6.170)$$

由 6.2 节的结论可知,  $A'_c(t), A'_s(t)$  都是低频限带过程, 它们随时间的变化  $\cos \omega_0 t$  要缓慢得多。

若将  $X(t)$  表示为准正弦振荡的形式

$$X(t) = A(t) \cos [\omega_0 t + \Phi(t)] \quad (6.171)$$

则  $X(t)$  的包络和相位  $A(t), \Phi(t)$  为

$$\begin{cases} A(t) = \sqrt{[A'_c(t)]^2 + [A'_s(t)]^2} \\ \Phi(t) = \arctan [A'_s(t)/A'_c(t)] \end{cases}. \quad (6.172)$$

易知  $A(t), \Phi(t)$  也都是慢变化随机过程。下面, 求  $A(t), \Phi(t)$  是  $A'_c(t), A'_s(t)$  的一维概率密度  $A'_c(t)$  和  $A'_s(t)$ 。

已知  $A(t), \Phi(t)$  是  $A'_c(t), A'_s(t)$  的函数, 而  $A'_c(t)$  和  $A'_s(t)$  又是随机相位  $\theta$  的函数, 所以  $A(t), \Phi(t)$  也是随机相位  $\theta$  的函数。

1) 求  $A'_c(t), A'_s(t)$  在给定  $\theta$  条件下的二维条件概率密度  $f(a'_{ct}/\theta, a'_{st}/\theta)$

给定  $\theta$  值的条件下,  $t$  时刻对  $A'_c(t), A'_s(t)$  的采样用  $A'_{ct}/\theta, A'_{st}/\theta$  表示。下面讨论  $A'_{ct}/\theta, A'_{st}/\theta$  的统计特性。

1°  $A'_u/\theta$  和  $A'_v/\theta$  都是高斯变量, 且相互独立。

由  $A'_c(t), A'_s(t)$  和  $A_c(t), A_s(t)$  的关系, 可得

$$\begin{cases} A'_{ct} = a \cos \theta + A_{ct} \\ A'_{st} = a \sin \theta + A_{st} \end{cases} . \quad (6.173)$$

因为  $A_a$  和  $A_{st}$  是独立高斯变量, 从上式便可推出  $A'_{ct}/\theta, A'_{st}/\theta$  也是独立高斯变量。

2°  $A'_{ct}/\theta, A'_{st}/\theta$  的均值。

由  $A_{ct}$  和  $A_{st}$  的均值皆为零, 可得

3°  $A'_{ct}/\theta, A'_{st}/\theta$  的方差。由  $A_{ct}$  和  $A_{st}$  的均值皆为  $\sigma^2$ , 可得

$$D[A'_{ct}/\theta] = D[A'_{st}/\theta] = \sigma^2 \quad (6.174)$$

由上述三个结论, 可得  $A'_c(t), A'_s(t)$  在给定条件下的二维条件概率密度为

$$f(a'_{ct}/\theta, a'_{st}/\theta) = \frac{1}{2\pi\sigma^2} \exp \left\{ -\frac{1}{2\sigma^2} \left[ (a'_{ct} - a \cos \theta)^2 + (a'_{st} - a \sin \theta)^2 \right] \right\} . \quad (6.175)$$

2) 求  $A(t)$  和  $\Phi(t)$  在给定  $\theta$  条件下的二维条件概率密度  $f(a_t/\theta, \varphi_t/\theta)$ 。 $A'_a, A'_{st}, A_t$  和  $\Phi_t$  的关系如下

$$\begin{cases} A'_{ct} = h_1(A_t, \Phi_t) = A_t \cos \Phi_t \\ A'_{st} = h_2(A_t, \Phi_t) = A_t \sin \Phi_t \end{cases} . \quad (6.176)$$

利用雅可比变换,  $J = a \geq 0$ , 便可得  $A$ , 在给定条件下的二维条件概率密度。

3) 求包络  $A(t)$  的一维概率密度  $f_A(a_t)$

对二维条件概率积分, 得到边缘条件概率密度为

$$\begin{aligned} f(a_t/\theta) &= \int_0^{2\pi} f(a_t/\theta, \varphi_t/\theta) d\varphi_t \\ &= \frac{a_t}{\sigma^2} I_0 \left( \frac{aa_t}{\sigma^2} \right) \exp \left\{ -\frac{a_t^2 + a^2}{2\sigma^2} \right\}, \quad a_t \geq 0, \end{aligned} \quad (6.177)$$

式中  $I_0(\cdot)$  是第一类零阶修正贝塞尔函数。

由上式可见,  $f(a_t/\theta)$  与  $\theta$  无关, 就是无条件分布  $f_A(a_t)$ 。于是, 可得随机余弦信号加窄带高斯噪声的包络  $A(t)$  的一维概率密度为

$$f_A(a_t) = \frac{a_t}{\sigma^2} I_0\left(\frac{aa_t}{\sigma^2}\right) \exp\left\{-\frac{a_t^2 + a^2}{2\sigma^2}\right\}, \quad a_t \geq 0. \quad (6.178)$$

当随相余弦信号不存在, 即幅度  $a = 0$  时, 式 (6.178) 便退化为式, 即  $A_t$  服从瑞利分布。所以, 称式 (6.178) 为广义瑞利概率密度或莱斯概率密度, 简称莱斯分布。 $a/\sigma$  表示信号幅度与窄带噪声标准差之比, 简称信噪比, 记为  $r$ 。

下面讨论在不同的信噪比  $r$  条件下包络的一维概率密度。为此, 将第一类零阶修正贝塞尔函数  $I_0(x)$  展开成无穷级数为

① 由于  $x \ll 1$  时

$$I_0(x) = \sum_{n=0}^{\infty} \frac{x^{2n}}{2^{2n}(n!)^2}. \quad (6.179)$$

② 当  $r \gg 1$  时

$$f_A(a_t) \approx \frac{a_t}{\sigma^2} \exp\left\{-\frac{a_t^2}{2\sigma^2}\right\}, \quad a_t \geq 0. \quad (6.180)$$

这就是说, 当信噪比很小时, 包络  $A(t)$  的一维概率密度趋近于瑞利分布

③ 由于  $r \ll 1$  时

$$I_0(x) \approx \frac{1}{\sqrt{2\pi x}} e^x. \quad (6.181)$$

在大信噪比条件下 ( $r \gg 1$ ), 包络  $A(t)$  的一维概率密度为

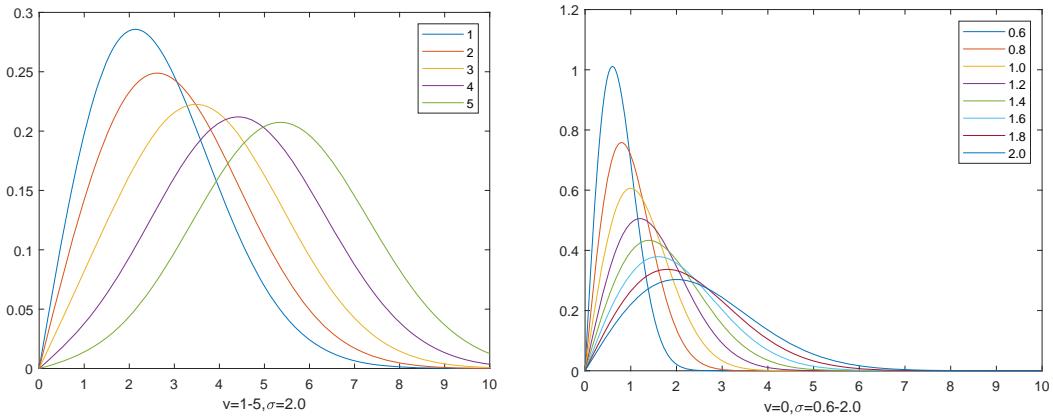
$$f_A(a_t) \approx \sqrt{\frac{a_t}{2\pi a}} \cdot \frac{1}{\sigma} \exp\left\{-\frac{(a_t - a)^2}{2\sigma^2}\right\}, \quad a_t \geq 0. \quad (6.182)$$

从上式中可见, 此概率密度在  $a_t = a$  处取最大值。当  $a_t$  偏离  $a$  时, 它很快下降, 且  $\sqrt{\frac{a_t}{2\pi a}}$  改变的速度比  $\exp\left\{-\frac{(a_t - a)^2}{2\sigma^2}\right\}$  的衰减速度要慢很多, 特别是在  $a$  的附近, 即当  $a_t$  偏离  $a$  很小时, 可以近似地认为  $\sqrt{\frac{a_t}{2\pi a}} \approx \frac{1}{\sqrt{2\pi}}$ 。

于是,  $f_A(a)$  可以近似为

$$f_A(a_t) \approx \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(a_t - a)^2}{2\sigma^2}\right\}, \quad a_t \geq 0. \quad (6.183)$$

上式说明, 在大信噪比的条件下, 在  $a$  附近个包络的一维概率密度近似为高斯分布。以上导出了包络  $A(t)$  的一维概率密度函数, 并得到了在大信噪比和小信噪比条件下它的近似式。图 6-20 所示为不同信噪比条件下莱斯分布的图形。

图 6-19 莱斯分布在不同  $v$  下的密度函数图 6-20 莱斯分布在不同  $\sigma$  下的密度函数

4) 求相位  $\Phi(t)$  的一维概率密度  $f_\Phi(\varphi/\theta)$

对二维条件概率密度积分, 得到边缘条件概率密度为

$$\begin{aligned}
 f_\Phi(\varphi_t/\theta) &= \int_0^\infty f(a_t/\theta, \varphi_t/\theta) da_t \\
 &= \int_0^\infty \frac{a_t}{2\pi\sigma^2} \exp\left\{-\frac{1}{2\sigma^2} [a_t^2 + a^2 - 2aa_t \cos(\theta - \varphi_t)]\right\} da_t \\
 &= \int_0^\infty \frac{a_t}{2\pi\sigma^2} \exp\left\{-\frac{1}{2\sigma^2} [a_t^2 - 2aa_t \cos(\theta - \varphi_t) \right. \\
 &\quad \left. + (a \cos(\theta - \varphi_t))^2 - (a \cos(\theta - \varphi_t))^2 + a^2]\right\} da_t \\
 &= \frac{1}{2\pi} \exp\left\{-\frac{a^2 - [a \cos(\theta - \varphi_t)]^2}{2\sigma^2}\right\} \\
 &\quad \int_0^\infty \frac{a_t}{\sigma^2} \exp\left\{-\frac{[a_t - a \cos(\theta - \varphi_t)]^2}{2\sigma^2}\right\} da_t \\
 &= \frac{1}{2\pi} \exp\left\{-\frac{r^2}{2}\right\} + \frac{r \cos(\theta - \varphi_t)}{\sqrt{2\pi}} \Phi[r \cos(\theta - \varphi_t)] \\
 &\quad \exp\left\{-\frac{1}{2} [r^2 - r^2 \cos^2(\theta - \varphi_t)]\right\}, \tag{6.184}
 \end{aligned}$$

式中  $\Phi(\cdot)$  是概率积分函数。上式是  $f_\Phi(\varphi_t/\theta)$  的一般表示式。

下面讨论在不同信噪比条件下,  $\Phi_t$  的概率密度形式

① 当信噪比  $r = 0$  时, 即不存在信号的条件下

$$f_\Phi(\varphi_t/\theta) = f_\Phi(\varphi_t) = \frac{1}{2\pi}. \tag{6.185}$$

这时, 随机相位为均匀分布。

② 当  $\theta - \varphi \ll 1$  时, 即信噪比很大时

$$\begin{cases} \sin(\theta - \varphi_t) \approx \theta - \varphi_t \\ \cos(\theta - \varphi_t) \approx 1 \end{cases} . \quad (6.186)$$

可以看出,  $f_\Phi(\varphi/\theta)$  的图形关于对称, 并在  $\varphi_t = \theta$  处取得最大值。

显然, 上式为高斯概率密度形式, 其均值为  $\theta$ , 方差为  $1/r^2$ 。

结论: 在信噪比极小时, 相位  $\Phi_t$  接近于均匀分布; 随信噪比的增加,  $f_\Phi(\varphi_t/\theta)$  逐渐接近于高斯分布; 在信噪比很大 ( $r^2 \gg 1$ ) 时,  $f_\Phi(\varphi_t/\theta)$  在  $\theta$  值附近服从高斯分布。由于方差  $1/r^2$  与信噪比  $r$  成反比, 所以当  $r \rightarrow \infty$  时,  $f_\Phi(\varphi_t/\theta)$  趋于  $\delta(\theta - \varphi_t)$ 。图 6-21 所示为不同信噪比条件下相位  $\Phi(t)$  的一维分布图。

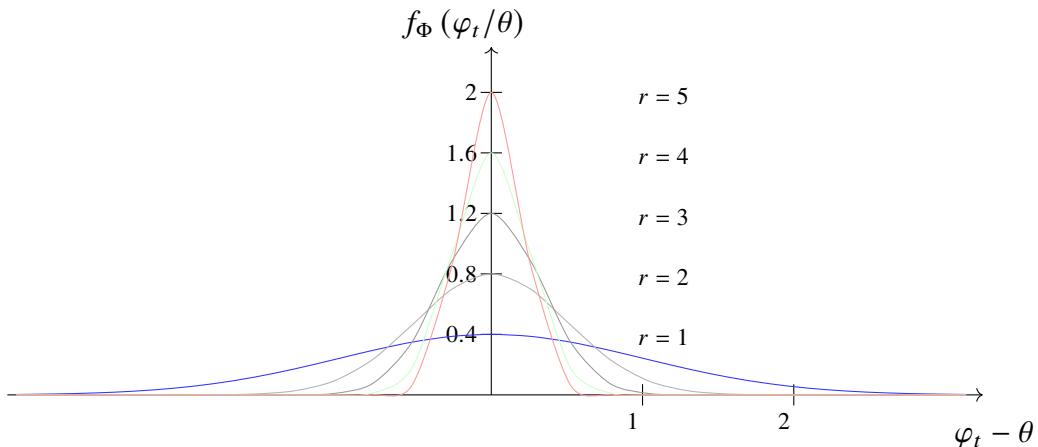


图 6-21  $f_\Phi(\varphi_t/\theta)$  的分布

## 6.4 窄带高斯过程包络平方的概率分布

在许多实际应用中, 也常常在高频窄带滤波器的输出端接入一平方律检波器如图 6-22 所示, 在平方律检波器输出端便得到  $X(t)$  包络的平方  $A^2(t)$ 。

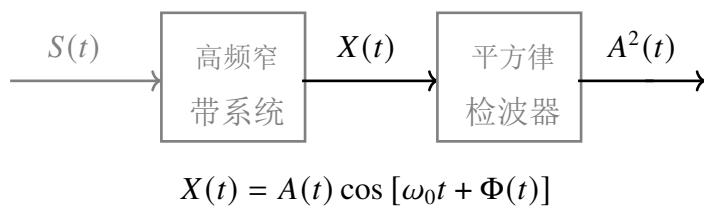


图 6-22 平方律检波器

### 6.4.1 窄带高斯噪声包络平方的分布

前面已经推导出当窄带随机过程为一具有零均值、方差为  $\sigma^2$  的平稳高斯噪声时, 其包络  $A(t)$  的一维概率密度为瑞利分布。

$$f_A(a_t) = \frac{a_t}{\sigma^2} \exp\left\{-\frac{a_t^2}{2\sigma^2}\right\}, a_t \geq 0. \quad (6.187)$$

应用求随机变量函数分布的方法, 很容易求出包络平方的一维概率密度。令

$$U(t) = A^2(t) \quad (6.188)$$

则在时刻  $t$  的采样有

$$\begin{cases} U_t = g(A_t) = A_t^2, & A_t \geq 0 \\ A_t = h(U_t) = \sqrt{U_t}, & U_t \geq 0. \end{cases}. \quad (6.189)$$

其雅可比行列式为

$$J = \frac{1}{2\sqrt{U_t}} \quad (6.190)$$

于是包络平方的一维概率密度为

$$f_U(u_t) = |J| f_A(A_t) = \frac{1}{2\sigma^2} \exp\left\{-\frac{u_t}{2\sigma^2}\right\}, \quad u_t \geq 0. \quad (6.191)$$

上式表明,  $U_t$  服从指数分布。

实际中, 为了分析方便, 常应用归一化随机变量。令归一化随机变量  $V_t = \frac{U_t}{\sigma^2}$ , 则可得到  $V_t$  的概率密度为

$$f_V(v_t) = \frac{1}{2} e^{-\frac{v_t}{2}}, \quad v_t \geq 0. \quad (6.192)$$

### 6.4.2 余弦信号加窄带高斯噪声包络平方的概率分布

当窄带随机过程为余弦信号加窄带高斯噪声时, 即

$$X(t) = a \cos \omega_0 t + N(t) = A(t) \cos [\omega_0 t + \Phi(t)], \quad (6.193)$$

其中  $a, \omega_0$  为已知常数,  $N(t)$  为具有零均值、方差  $\sigma^2$  的窄带高斯噪声。根据 6-3 节分析的结果可知,  $X(t)$  的包络服从广义瑞利分布, 即

$$f_A(a_t) = \frac{a_t}{\sigma^2} I_0\left(\frac{aa_t}{\sigma^2}\right) \exp\left\{-\frac{a_t^2 + a^2}{2\sigma^2}\right\}, \quad a_t \geq 0. \quad (6.194)$$

仿照 6.4.1 节中的方法, 不难导出包络平方  $U^2 = A^2$  的一维概率密度为

$$f_U(u_t) = \frac{1}{2\sigma^2} I_0\left[\frac{a\sqrt{u_t}}{\sigma^2}\right] \exp\left\{-\frac{1}{2\sigma^2}[u_t + a^2]\right\}, \quad u_t \geq 0. \quad (6.195)$$

令  $V_t = \frac{U_t}{\sigma^2}$ , 可得归一化随机变量  $V$  的概率密度函数为

$$f_V(v_t) = \frac{1}{2} I_0 \left[ \frac{\sqrt{v_t} a}{\sigma} \right] \exp \left\{ -\frac{v_t + a^2/\sigma^2}{2} \right\}, \quad v_t \geq 0. \quad (6.196)$$

### 6.4.3 卡方分布和非中心卡方分布

#### 1. $\chi^2$ 分布

在有些应用中, 例如在信号检测中, 为了改进检测性能, 经常采用所谓视频积累技术, 即对包络的平方进行独立采样后再积累, 如图 6-23 所示, 这时输出的随机变量习惯上记为  $\chi^2$ , 它的概率密度习惯上称之为  $\chi^2$  分布。如图 6-23 所示, 让一个具有零均值和方差  $\sigma^2$  的平稳窄带实高斯噪声  $N(t)$  通过一平方律检波器, 而检波器输出的是  $N(t)$  的包络平方  $A^2(t)$ 。然后对随机过程  $A^2(t)$  进行独立采样, 得到  $m$  个独立的随机变量  $A_i^2 = A^2(t_i)$  ( $i = 1, 2, \dots, m$ ), 经归化以后送入累加器。下面讨论累加器输出端随机变量  $\chi^2$  的概率密度。为了避免混淆, 在下面的推导中, 用  $V$  来代替  $\chi^2$ 。

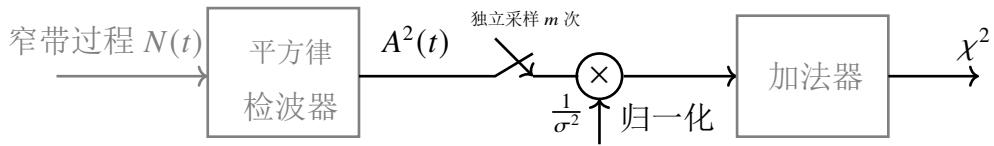


图 6-23 视频积累技术

由于窄带过程  $N(t)$  的包络  $A$  和它的一对垂直分量  $A_c(t), A_s(t)$  有如下关

$$A^2(t) = A_c^2(t) + A_s^2(t), \quad (6.197)$$

式中  $A_c(t), A_s(t)$  是零均值、方差为  $\sigma^2$  的平稳高斯过程。 $A^2(t)$  经采样后, 加法器的输出

$$V = \sum_{i=1}^m (A_{ci}'^2 + A_{si}'^2). \quad (6.198)$$

式中  $A_{ci}', A_{si}'$  是将  $A_{ci}, A_{si}$  归一化以后的随机变量。由于  $A_{ci}', A_{si}'$  都是同分布的高斯变量, 故上式又可表示为

$$V = \sum_{i=1}^{2m} X_i^2. \quad (6.199)$$

式中每一个  $X_i$  都是同分布的标准高斯变量(零均值、方差为 1), 且各  $X_i$  相互独立。为了书写简单, 用  $n$  代替上式中的  $2m$ , 可得

$$V = \sum_{i=1}^n X_i^2. \quad (6.200)$$

于是, 求  $V$  的概率密度便归结为求  $n$  个独立同分布高斯变量平方和的概率密度为此, 首先求每一随机变量  $X_i^2$  的概率密度。已知标准高斯随机变量  $X_i$  的概率密度为

$$f_X(x_i) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x_i^2}{2}\right\}. \quad (6.201)$$

令  $Y_i = X_i^2$ , 不难求出  $Y_i$  的概率密度

$$f_{Y_i}(y_i) = |J| f_X(x_i) = \frac{1}{\sqrt{2\pi_i}} e^{-\frac{y_i}{2}}, \quad y_i \geq 0. \quad (6.202)$$

从而得到  $Y_i$  的特征函数为

$$\begin{aligned} Q_{Y_i}(u) &= \int_{-\infty}^{\infty} f_{Y_i}(y_i) e^{iy_i} dy_i \\ &= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} y_i^{-\frac{1}{2}} \exp\left\{-\left(\frac{1}{2} - ju\right)y_i\right\} dy_i = (1 - 2ju)^{\frac{1}{2}}. \end{aligned} \quad (6.203)$$

由于  $V = \sum_{i=1}^n Y_i$ , 利用特征函数的性质: 独立随机变量之和的特征函数等于各随机变量特征函数的乘积。便可得到  $V$  的特征函数为

$$Q_V(u) = \prod_{i=1}^n Q_{Y_i}(u) = (1 - 2ju)^{-\frac{n}{2}} \quad (6.204)$$

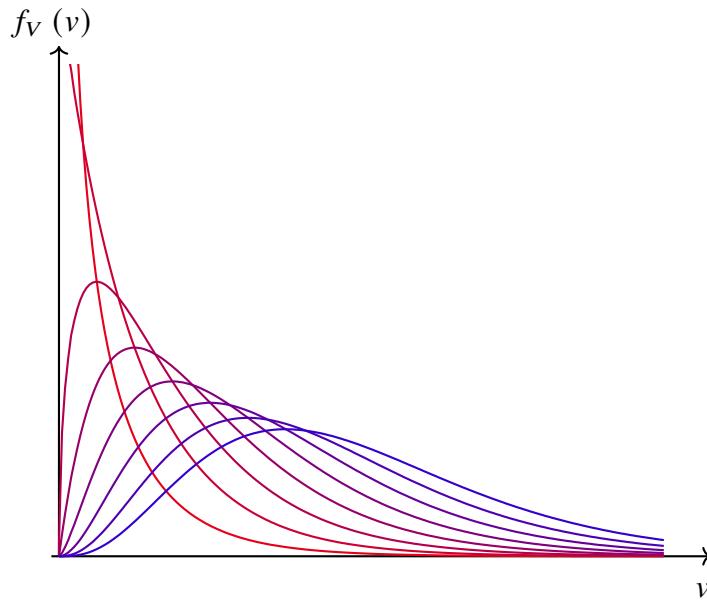
上式进行傅里叶逆变换, 便可求得  $V$  的概率密度为

$$f_V(v) = \frac{1}{\Gamma\left(\frac{n}{2}\right)} 2^{-\frac{n}{2}} v^{\left(\frac{n}{2}-1\right)} e^{-\frac{v}{2}}, \quad v \geq 0. \quad (6.205)$$

式中,  $\Gamma(\cdot)$  为  $\Gamma$  函数, 满足

$$\Gamma(\alpha) = \int_0^{\infty} t^{(\alpha-1)} e^{-t} dt. \quad (6.206)$$

称  $f_v(v)$  为  $n$  个自由度的  $\chi^2$  分布。图 6-24 画出了几个不同自由度下  $f_v(v)$  的图形。

图 6-24 几个不同自由度下的  $\chi^2$  分布

$\chi^2$  分布具有下列性质

1° 两个独立的  $\chi^2$  变量之和仍为  $\chi^2$  变量。若它们各自的自由度分别为  $n_1$  和  $n_2$ , 则它们的和变量为具有  $(n_1+n_2)$  个自由度的  $\chi^2$  分布  $2n$  个自由度的  $\chi^2$  变量的均值  $E[V] = n$ , 方差  $D[V] = 2n$ 。

## 2. 非中心 $\chi^2$ 分布

若窄带过程  $N'(t)$  为余弦信号  $N'(t)$  与窄带高斯噪声  $N(t)$  之和, 则加法器输出的就是非中心  $\chi^2$  分布。

### (1) 信号包络为常数的情况

设信号

$$s(t) = a \cos(\omega_0 t + \pi/4) \quad (6.207)$$

包络  $a$  为常数, 则有

$$\begin{aligned} s(t) &= a \cos \frac{\pi}{4} \cos \omega_0 t + a \sin \frac{\pi}{4} \sin \omega_0 t \\ &= a \frac{\sqrt{2}}{2} \cos \omega_0 t + a \frac{\sqrt{2}}{2} \sin \omega_0 t. \end{aligned} \quad (6.208)$$

若令  $s = a\sqrt{2}/2$ , 则

$$s(t) = s \cos \omega_0 t + s \sin \omega_0 t. \quad (6.209)$$

又由于

$$N(t) = n_e(t) \cos \omega_0 t - n_3(t) \sin \omega_0 t. \quad (6.210)$$

代入  $N(t)$  得

$$\begin{aligned} N'(t) &= s(t) + N(t) \\ &= [s + n_c(t)] \cos \omega_0 t - [s + n_s(t)] \sin \omega_0 t \\ &= A_c(t) \cos \omega_0 t - A_s(t) \sin \omega_0 t. \end{aligned} \quad (6.211)$$

而  $N(t)$  的包络的平方

$$A^2(t) = A_c^2(t) + A_s^2(t) = [s + n_c(t)]^2 + [s + n_s(t)]^2. \quad (6.212)$$

仿照求  $\chi^2$  分布的方法, 加法器输出端的随机变量  $V'$  应为

$$V' = \sum_{i=1}^n (s + X_i)^2 = \sum_{i=1}^n Y_i^2 \quad (6.213)$$

式中  $X_i$  为同分布的独立高斯变量 (均值为零、方差为  $\sigma^2$ ),  $s$  为常数。为了导出  $V$  的概率密度, 首先求  $Y_i = (s + X_i)^2$  的概率密度和特征函数。令

$$R_i = s + X_i. \quad (6.214)$$

显然,  $R_i$  的概率密度为

$$f_{R_i}(r_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(r_i-s)^2}{2\sigma^2}\right\}. \quad (6.215)$$

则  $Y' = R^2$  的概率密度为

$$f_{Y_i}(y'_i) = \frac{1}{\sqrt{8\pi}\sigma^2 y'_i} \left\{ \exp\left[-\frac{(\sqrt{y'_i}-s)^2}{2\sigma^2}\right] + \exp\left[-\frac{(-\sqrt{y'_i}-s)^2}{2\sigma^2}\right] \right\}. \quad (6.216)$$

将上式中指数的平方项展开, 并利用双曲余弦函数  $2 \cosh(b) = e^b + e^{-b}$ , 可得

$$f_{Y_i}(y'_i) = \frac{1}{\sqrt{2\pi}\sigma^2 y'_i} \exp\left\{-\frac{y'_i+s^2}{2\sigma^2}\right\} \cosh\left(\frac{s\sqrt{y'_i}}{\sigma^2}\right). \quad (6.217)$$

由于  $X, Y$  为独立同分布的, 则  $Y_i = (s + X_i)^2$  也是独立同分布的。而  $V' = \sum_{i=1}^n Y'_i$ , 于是  $V$  的特征函数为

$$\mathcal{Q}_{V'}(u) = \prod_{i=1}^n \mathcal{Q}_{Y'_i}(u) = \left(\frac{1}{1-j2\sigma^2 u}\right)^{\frac{n}{2}} \exp\left\{-\frac{ns^2}{2\sigma^2} + \frac{ns^2}{2\sigma^2(1-j2\sigma^2 u)}\right\}. \quad (6.218)$$

对上式作傅里叶逆变换, 可得  $V$  的概率密度为

$$f_{V'}(v') = \frac{1}{2\sigma^2} \left(\frac{v'}{\lambda'^2}\right)^{\frac{n-2}{4}} \exp\left\{-\frac{\lambda' + v'}{2\sigma^2}\right\} I_{\frac{n}{2}-1}\left(\frac{\sqrt{v'\lambda'}}{\sigma^2}\right), \quad v' \geq 0. \quad (6.219)$$

式中,  $\lambda' = s^2 n$  定义为非中心参量,  $I_n(\cdot)$  为第一类  $n$  阶修正贝塞尔函数。

定义归一化变量  $V = V'/\sigma^2$ , 那么

$$V = \sum_{i=1}^n \left( \frac{s}{\sigma} + \frac{X_i}{\sigma} \right)^2 = \sum_{i=1}^n R_i'^2. \quad (6.220)$$

其中变量  $R$  是均值为  $s/\sigma$ 、方差为 1 的相互独立的高斯变量。易证  $V$  的概率密度为

$$f_V(v) = \frac{1}{2} \left( \frac{v}{\lambda} \right)^{\frac{n-2}{4}} \exp \left\{ -\frac{\lambda}{2} - \frac{v}{2} \right\} I_{\frac{n}{2}-1}(\sqrt{v\lambda}), \quad v \geq 0. \quad (6.221)$$

上式是  $n$  个自由度的非中心  $\chi^2$  分布。其中非中心参量  $\lambda = ns^2/\sigma^2$  表示视频积累后的功率信噪比。图 6-25 画出了不同信噪比  $\lambda$  和样本数  $n$  情况下的非中心  $\chi^2$  函数。(2) 信号

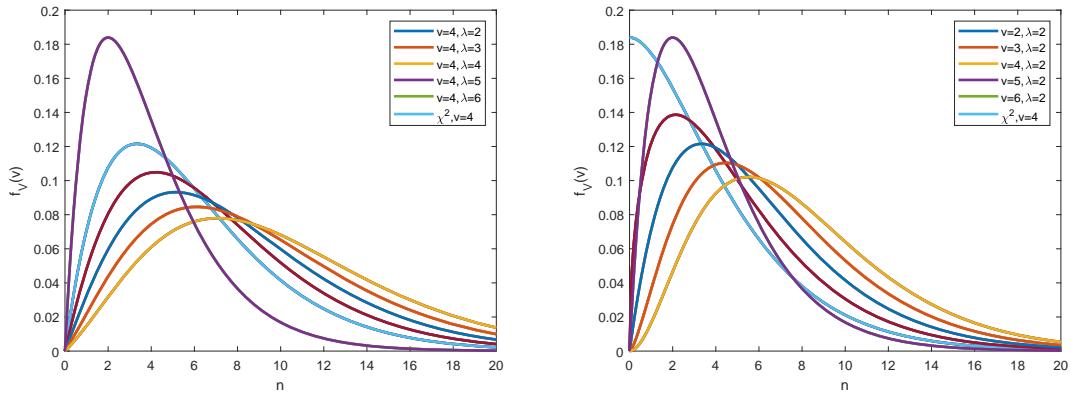


图 6-25 非中心  $\chi^2$  分布: a)  $v = 4, \lambda = 2..6$ , b)  $v = 2..6, \lambda = 2$

包络不为常数的情况

给定信号

$$s(t) = a(t) \cos(\omega_0 t + \pi/4), \quad (6.222)$$

包络  $a(t)$  为确定函数, 则有

$$\begin{aligned} s(t) &= a(t) \cos \frac{\pi}{4} \cos \omega_0 t + a(t) \sin \frac{\pi}{4} \sin \omega_0 t \\ &= a(t) \frac{\sqrt{2}}{2} \cos \omega_0 t + a(t) \frac{\sqrt{2}}{2} \sin \omega_0 t. \end{aligned} \quad (6.223)$$

又由于

$$N(t) = n_c(t) \cos \omega_0 t - n_s(t) \sin \omega_0 t. \quad (6.224)$$

代入  $N'(t)$  得

$$\begin{aligned} N'(t) &= s(t) + N(t) = \left[ a(t) \frac{\sqrt{2}}{2} + n_e(t) \right] \cos \omega_0 t - \left[ a(t) \frac{\sqrt{2}}{2} + n_1(t) \right] \sin \omega_0 t \\ &= A_c(t) \cos \omega_0 t - A_s(t) \sin \omega_0 t. \end{aligned} \quad (6.225)$$

而  $N(t)$  的包络的平方

$$\begin{aligned} A^2(t) &= A_c^2(t) + A_s^2(t) \\ &= \left[ a(t) \frac{\sqrt{2}}{2} + n_c(t) \right]^2 + \left[ a(t) \frac{\sqrt{2}}{2} + n_3(t) \right]^2. \end{aligned} \quad (6.226)$$

在  $t_i$  ( $i = 1, 2, \dots, n$ ) 时刻对  $A^2(t)$  进行独立的采样, 令  $s = a(t)$ , 仿照求  $\chi^2$  分布的方法, 加法器输出端的随机变量  $Q'$  应为

$$Q' = \sum_{i=1}^n (s_i + X_i)^2 = \sum_{i=1}^n Y_i, \quad (6.227)$$

式中  $s$  是对信号包络  $a(t)$  的第  $i$  次采样, 是确定值。由于单个样本  $Y_i = (s_i + X_i)^2$  的特征函数, 可以直接应用上面信号包络为常量的推导结果

$$Q_Y^*(u) = \frac{1}{\sqrt{1 - j2\sigma^2 u}} \exp \left\{ -\frac{s_1^2}{2\sigma^2} + \frac{s_1^2}{2\sigma^2(1 - j2\sigma^2 u)} \right\}. \quad (6.228)$$

又因为  $Y_i^*$  ( $i = 1, 2, \dots, n$ ) 相互独立, 而  $Q' = \sum_{k=1}^n Y_i$ , 于是  $Q'$  的特征函数为

$$Q_\alpha(u) = \prod_{i=1}^n Q_{Y_i^*}(u) = \left( \frac{1}{1 - j2\sigma^2 u} \right)^{\frac{n}{2}} \exp \left\{ -\frac{\sum_{i=1}^n s_i^2}{2\sigma^2} + \frac{\sum_{i=1}^n s_i^2}{2\sigma^2(1 - j2\sigma^2 u)} \right\}. \quad (6.229)$$

对上式作傅里叶逆变换可得  $Q'$  的概率密度

$$f_\alpha(q') = \frac{1}{2\sigma^2} \left( \frac{q'}{\lambda'^2} \right)^{\frac{n-2}{4}} \exp \left\{ -\frac{\lambda' + q'}{2\sigma^2} \right\} \cdot I_{\frac{n}{2}-1} \left( \frac{\sqrt{q'\lambda'}}{\sigma^2} \right), \quad q' \geq 0. \quad (6.230)$$

将式 (6.230) 与式 (6.219) 对照可见,  $Q'$  与  $V'$  具有相同的概率密度, 不同的只是此时的非中心参量  $\lambda' = \sum_{i=1}^n s_i^2$ , 类似地, 定义归一化参量  $Q = Q'/\sigma^2$ , 于是可得

$$Q = \sum_{i=1}^n \left( \frac{s_i}{\sigma} + \frac{X_i}{\sigma} \right)^2 \quad (6.231)$$

令

$$Z_i = \frac{s_i}{\sigma} + \frac{X_i}{\sigma}. \quad (6.232)$$

则  $Z_i (i = 1, 2, \dots, n)$  是具有均值  $s_i/\sigma$  和单位方差的独立高斯变量。于是, 可得具有  $n$  个自由度的非中心  $\chi^2$  分布为

$$f_Q(q) = \frac{1}{2} \left( \frac{q}{\lambda} \right)^{\frac{n-2}{4}} \exp \left\{ -\frac{\lambda}{2} - \frac{q}{2} \right\} I_{\frac{n}{2}-1}(\sqrt{Q\lambda}), \quad (6.233)$$

式中  $\lambda = \sum_{i=1}^n s_i^2 / \sigma^2$  为非中心参量。

不难证明, 两个统计独立的非中心  $\chi^2$  随机变量之和仍为非中心  $\chi^2$  随机变量。

若它们的自由度分别为  $n_1$  与  $n_2$ , 非中心参量分别为  $\lambda_1$  和  $\lambda_2$ , 则和变量的自由度为  $n = n_1 + n_2$ , 非中心参量为  $\lambda = \lambda_1 + \lambda_2$ 。

**例 6.9** 设图 6-2 中加至平方律检波器输入端的窄带随机过程  $X(t)$  为

$$X(t) = a \cos(\omega_0 t + \theta) + N(t), \quad (6.234)$$

其中  $a \cos(\omega_b t + \theta)$  为随相余弦信号,  $a$  和  $\omega_0$  为常数。 $N(t)$  是零均值、方差为  $\sigma^2$  的稳窄带高斯噪声, 其功率谱关于  $\omega_0$  偶对称。 $X(t)$  经检波并作归一化处理以后, 独立采样  $m$  次, 求累加器输出端随机变量的概率密度及其参数。

解: 先将  $N(t)$  表示为

$$N(t) = A_c(t) \cos \omega_0 t - A_s(t) \sin \omega_0 t, \quad (6.235)$$

若用  $A(t)$  表示窄带随机过程  $X(t)$  的包络, 那么在平方律检波器的输出端, 可得到包络平方为

$$A^2(t) = [a \cos \theta + A_c(t)]^2 + [a \sin \theta + A_s(t)]^2. \quad (6.236)$$

于是, 加法器输出端随机变量  $V$  为

$$V = \frac{1}{\sigma^2} \left[ \sum_{i=1}^m (a \cos \theta + A_{ci})^2 + \sum_{i=1}^m (a \sin \theta + A_{si})^2 \right], \quad (6.237)$$

式中  $A_\sigma = A_e(t_i)$  和  $A_{si} = A_s(t_i)$  分别表示  $A_e(t)$  和  $A_s(t)$  在  $t_i$  时刻的状态。根据  $A_c(t), A_s(t)$  的有关性质可知, 各个  $A_\sigma, A_{si}$  是同分布的独立标准高斯变量。对照式 (6.220) 可知, 上式中两个和式分别是自由度为  $m$  的非中心  $\chi^2$  变量, 它们的非中心参量  $\lambda_1$  和  $\lambda_2$  分别为

$$\begin{cases} \lambda_1 = \frac{1}{\sigma^2} \sum_{i=1}^m (a \cos \theta)^2 = \frac{ma^2}{\sigma^2} \cos^2 \theta \\ \lambda_2 = \frac{1}{\sigma^2} \sum_{i=1}^m (a \sin \theta)^2 = \frac{ma^2}{\sigma^2} \sin^2 \theta \end{cases}. \quad (6.238)$$

由于这两个非中心  $\chi^2$  随机变量也彼此独立, 因而它们的和变量  $V$  也是非中心  $\chi^2$  随机变量, 自由度  $n = 2m$ , 非中心参量  $\lambda = \lambda_1 + \lambda_2 = \frac{ma^2}{\sigma^2}$ , 便可得到  $V$  的概率密度函数为

$$f_V(v) = \frac{1}{2} \left( \frac{v}{\lambda} \right)^{\frac{m-1}{2}} \exp \left\{ -\frac{\lambda}{2} - \frac{v}{2} \right\} I_{m-1}(\sqrt{v\lambda}), \quad v \geq 0. \quad (6.239)$$

而非中心参量与自由度之比  $\lambda/n = a^2/2\sigma^2$ , 正好是检波器输入端的功率信噪比。

## 6.5 习题

### 练习 6.1 证明

- ①  $H \left[ \frac{\sin t}{t} \right] = \frac{1-\cos t}{t}$ .
- ②  $H \left[ e^{j\omega_0 t} \right] = -je^{j\omega_0 t}, \omega_0 > 0$ .
- ③  $H^{-1} \left[ \frac{\sin t}{t} \sin 200\pi t \right] = \frac{\sin t}{t} \cos 200\pi t$ .
- ④  $H \left[ e^{j\omega_0 t} \right] = -je^{j\omega_0 t}, \omega_0 > 0$ .
- ⑤  $H^{-1} \left[ \frac{1}{\pi t} \right] = \delta(t)$ .
- ⑥  $H^{-1} \left[ \frac{\sin t}{t} \cos 200\pi t \right] = -\frac{\sin t}{t} \sin 200\pi t$ .

**练习 6.2** 证明: ① 偶函数的希尔伯特变换为奇函数。② 奇函数的希尔伯特变换为偶函数。

**练习 6.3** 当  $\tau, \omega_0$  满足什么条件时, 能使  $y(t) = \frac{\sin(\pi t/\tau)}{\pi t/\tau} e^{j\omega_0 t} = \text{Sa}(\pi t/\tau) e^{j\omega_0 t}$  为解析信号。画出  $\text{Sa}(\pi t/\tau)$  和  $y(t)$  的频谱图。

**练习 6.4** 调频信号  $s(t) = \cos [\omega_0 t + m(t)]$ , 当 ①  $\frac{dm(t)}{dt} \ll \omega_0$  时, 为窄带信号, 求  $s(t)$  的包络和预包络。

**练习 6.5** 已知随机过程

$$X(t) = [X_1(t) \quad X_2(t) \quad X_1(t+\tau) \quad X_2(t+\tau)]^T. \quad (6.240)$$

式中  $X_1(t)$  为平稳标准高斯过程,  $X_2(t)$  为  $X_1(t)$  的希尔伯特变换。证明

$$E \left[ X(t) X^T(t) \right] = \begin{bmatrix} 1 & 0 & R_1(\tau) & \hat{R}_1(\tau) \\ 0 & 1 & -\hat{R}_1(\tau) & R_1(\tau) \\ R_1(\tau) & -\hat{R}_1(\tau) & 1 & 0 \\ \hat{R}_1(\tau) & R_1(\tau) & 0 & 1 \end{bmatrix}, \quad (6.241)$$

其中  $R_1(\tau) = E [X_1(t)X_1(t+\tau)]$ 。

**练习 6.6** 已知平稳过程  $X(t)$  的功率谱密度  $G_X(\omega)$ , 如图 6-26 所示。求随机过程  $W(t)$  的功率谱密度, 并画图表示。

$$W(t) = X(t) \cos \omega_0 t - \hat{X}(t) \sin \omega_0 t. \quad (6.242)$$

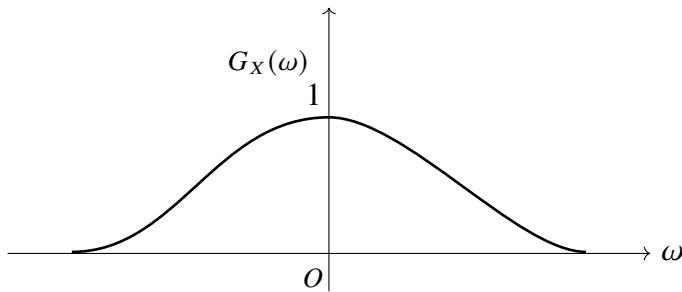


图 6-26

**练习 6.7** 零均值窄带平稳过程  $X(t) = A(t) \cos \omega_0 t - B(t) \sin \omega_0 t$  的功率谱密度  $G_X(\omega)$  在频带内关于中心频率  $\omega$  偶对称, 其中  $A(t), B(t)$  为平稳过程。① 证明:  $X(t)$  的自相关函数  $R_X(\tau) = R_A(\tau) \cos \omega_0 \tau$ 。② 求  $X(t)$  自相关函数的包络和预包络。

**练习 6.8** 已知零均值窄带平稳噪声  $X(t) = A(t) \cos \omega_0 t - B(t) \sin \omega_0 t$ , 其功率谱密度如图 6-27 所示。画出下列情况下随机过程  $A(t), B(t)$  各自的功率谱密度:

- (1)  $\omega_0 = \omega_1$ ;
- (2)  $\omega_0 = \omega_2$ ;
- (3)  $\omega_0 = (\omega_1 + \omega_2) / 2$ .

判断上述每种情况下, 过程  $A(t), B(t)$  是否互不相关, 并给出理由。

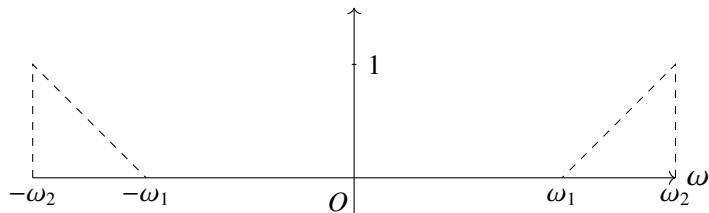


图 6-27

**练习 6.9** 已知平稳噪声  $N(t)$  的功率谱密度, 如图 6-28 所示。求窄带过程

$$X(t) = N(t) \cos(\omega_0 t + \theta) - N(t) \sin(\omega_0 t + \theta) \quad (6.243)$$

的功率谱密度  $G_X(m)$ , 并画图表示。其中  $a_n > a_1$  为常数, 服从  $(0, 2\pi)$  上的均匀分布, 且与噪声  $N(t)$  独立。

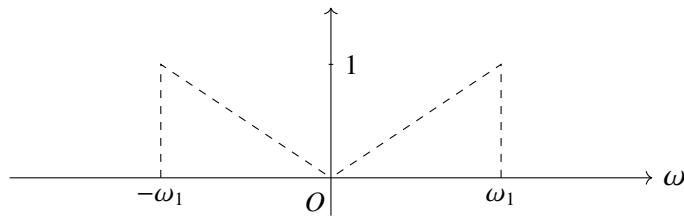


图 6-28

- 练习 6.10 已知零均值, 方差为  $\sigma^2$  的窄带高斯平稳过程  $X(t) = A_c(t) \cos \omega_0 t - A_s(t) \sin \omega_0 t$ , 其中  $A_c(t), A_s(t)$  为过程的一对垂直分解。证明:  $R_X(\tau) = R_{A_c}(\tau) \cos \omega_0 \tau - R_{A_c A_s}(\tau) \sin \omega_0 \tau$ 。

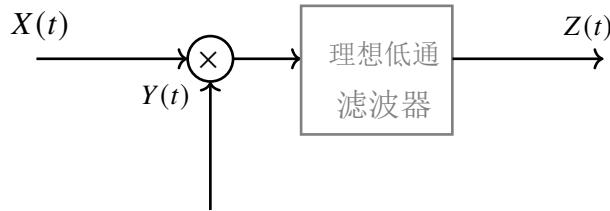


图 6-29

- 练习 6.11 证明: 零均值, 方差为 1 的窄带平稳高斯过程, 其任意时刻的包络平方的数学期望为 2, 方差为 4。
- 练习 6.12 已知窄带高斯平稳过程  $X(t) = A(t) \cos [\omega_0 t + \Phi(t)]$ , 包络  $A(t)$  在任意时刻  $t$  的采样为随机变量  $A_t$ , 求  $A_t$  的均值和方差。
- 练习 6.13 如图 6-29 所示, 同步检波器的输入  $X(t)$  为窄带平稳噪声, 其自相关函数为

$$R_X(\tau) = \sigma_X^2 e^{-\beta|\tau|} \cos \omega_0 \tau, \quad \beta \ll \omega_0. \quad (6.244)$$

若另一输入

$$Y(t) = A \sin (\omega_0 t + \theta), \quad (6.245)$$

其中  $A$  为常数, 服从  $(0, 2\pi)$  上的均匀分布, 且与噪声  $X(t)$  独立。求检波器输出  $Z(t)$  的平均功率。

- 练习 6.14 如图 6-30 所示, 系统 1 是线性系统的传递函数, 关于中心频率  $\omega$  偶对称, 系统 2 是线性系统的传递函数为  $-j\text{sgn}(\omega)$ , 系统 3 为线性微分系统。输入  $N(t)$  为物理谐密度

为  $N_0$  的白噪声, 且系统 1 输出  $X(t)$  的自相关函数的包络为  $\exp\{-\tau^2\}$ 。整个系统进入稳态。

- ① 判断  $X(t)$  和  $Z(t)$  分别服从什么分布, 给出理由。
- ② 证明  $Z(t)$  是严平稳过程。
- ③ 求  $X(t)$  和  $Y(t)$  的互相关函数,  $Y(t)$  的功率谱密度,
- ④ 写出  $Z(t)$  一维概率密度表达式。
- ⑤ 判断同一时刻,  $Y(t)$  和  $Z(t)$  是否独立, 给出理由。



图 6-30

**练习 6.15** 如图 6-31 所示, 系统输入  $N(t)$  为物理谱密度为  $N_0$  的白噪声, 对包络平方检波后的过程进行二次独立采样, 求积累后的输出  $X(t)$  的分布平方律。

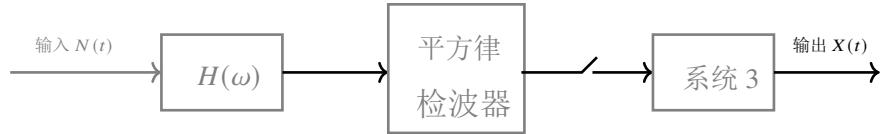


图 6-31



# 7

## 马尔可夫过程、独立增量过程及独立随机过程

第二章详细地介绍了平稳随机过程、高斯随机过程和白噪声随机过程等随机过程，它们是工程领域中最常见、最基本的几种随机过程。本章将讨论另外几种在工程技术中常见且重要的随机过程：马尔可夫过程、独立增量过程、独立随机过程。马尔可夫过程是一种发展很快且应用很广的重要的随机过程，它在信息处理、自动控制、数字计算方法、近代物理、生物(生灭过程)以及公用事业等方面皆有重要的应用。独立增量过程是一种特殊的马尔可夫过程，泊松(Poisson)过程和维纳(Wiene)过程是两个最重要的独立增量过程，也即马尔可夫过程。电子系统中，它们是研究热噪声和散粒噪声的数学基础，具有重要的实用价值。独立随机过程是一种很特殊的随机过程，它的重要应用就是高斯白噪声。连续时间参数的独立随机过程是一种理想化的随机过程，它在数学处理上具有简单且方便的优点。

马尔科夫过程是一类重要的随机过程。原始模型是俄国数学家 A. A. 马尔可夫于 1907 年提出的马尔可夫链。该过程具有已知目前状态(现在)的条件下，它未来的演变(将来)不依赖于它以往的演变(过去)的特性。

例如森林中动物头数的变化构成——马尔可夫过程。在现实世界中，有很多过程都是马尔可夫过程，如液体中微粒所作的布朗运动、传染病受感染的人数、车站的候车人数等，都可视为马尔可夫过程。马尔科夫过程概率分布的演化方程就是主方程。普遍应用于化学、生物学、人口动力学、激光物理学、布朗运动、流体及半导体相关等问题。马尔科夫过程理论主要包括：马尔科夫过程的基本理论、马尔科夫链、马尔科夫方程和布朗运动。

## 7.1 马尔可夫过程

电子系统中, 马尔可夫过程是一种重要的随机过程, 它具有如下特性: 当随机过程在时刻  $t$  所处的状态已知时, 过程在时刻  $t_1$  ( $t_1 > t$ ) 所处的状态仅与过程在  $t$  时刻的状态有关, 而与过程在  $t$  时刻以前所处的状态无关。此特性称为随机过程的无后效性或马尔可夫性。无后效性也可理解为: 随机过程  $X(t)$  在“现在”状态已知的条件下, 过程“将来”的情况与“过去”的情况是无关的。或者说, 随机过程的“将来”只是通过“现在”与“过去”产生关系, 且“现在”已知, 那么“将来”和“过去”就无关了。马尔可夫过程按其状态空间和时间参数集  $T$  是连续还是离散可分成四类。

表 7-1 马尔可夫过程的分类

分类名称	状态空间 I	
	离散	连续
时间参数集 $T$		
离散 ( $n = 0, 1, 2, \dots$ )	马尔可夫链	马尔可夫序列
连续 ( $t \geq 0$ )	可列马尔可夫过程	马尔可夫过程

可见, 如表 7-1 所示, 马尔可夫链是指时间、状态皆离散的马尔可夫过程; 马尔可夫序列是指时间离散、状态连续的马尔可夫过程; 可列马尔可夫过程是指时间连续、状态离散的马尔可夫过程; 至于马尔可夫过程, 有时指时间、状态皆连续的马尔可夫过程, 有时也为此四类过程的总称。

实际上, 我们所观察到的物理过程并不一定是精确的马尔可夫过程。然而, 在很多具体问题中, 有时却能近似地将其看作马尔可夫过程, 这正是我们研讨马尔可夫过程的原因。下面, 我们将对马尔可夫序列、马尔可夫链和一般马尔可夫过程的概念及特性依次加以介绍。

### 7.1.1 马尔可夫序列

随机序列  $\{X(n)\} = \{X_1, X_2, \dots, X_n, \dots\}$  可看作随机过程  $\{X(n)\}$  在  $t$  为整数时的采样值。

#### 定义 7.1 马尔可夫序列

若对于任意的  $n$ , 随机序列  $\{X(n)\}$  的条件分布函数满足

$$F_X(x_n|x_{n-1}, x_{n-2}, \dots, x_1) = F_X(x_n|x_{n-1}), \quad (7.1)$$

则称此随机序列  $\{X(n)\}$  为马尔可夫序列。条件分布函数  $F_x(x_n|x_{n-1})$  常被称为转移分布。



对于连续型随机变量, 由上式可得

$$f_X(x_n|x_{n-1}, x_{n-2}, \dots, x_1) = f_X(x_n|x_{n-1}), \quad (7.2)$$

因此, 利用条件概率的性质

$$f_X(x_1, x_2, \dots, x_n) = f_X(x_n|x_{n-1}, \dots, x_1) \cdots f_X(x_2|x_1) f_X(x_1). \quad (7.3)$$

结合式 (7.2), 可得

$$\begin{aligned} f_X(x_1, x_2, \dots, x_n) &= f_X(x_n|x_{n-1}, \dots, x_1) f_X(x_{n-1}, x_{n-2}, \dots, x_1) \\ &= f_X(x_n|x_{n-1}) f_X(x_{n-1}|x_{n-2}) \cdots f_X(x_2|x_1) f_X(x_1). \end{aligned} \quad (7.4)$$

和初始概率密度  $f_X(x_1)$  所确定。相反地, 若式 (7.4) 对所有的  $n$  皆成立, 则序列是马尔可夫序列。因为

$$f_X(x_n|x_{n-1}, \dots, x_2, x_1) = \frac{f_X(x_1, x_2, \dots, x_{n-1}, x_n)}{f_X(x_1, x_2, \dots, x_{n-1})} = f_X(x_n|x_{n-1}). \quad (7.5)$$

## 2. 马尔可夫序列的性质

1° 马尔可夫序列的子序列仍为马尔可夫序列

给定  $n$  个任意整数  $k_1 < k_2 < \dots < k_n$ , 有

$$f_X(x_{k_n}|x_{k_{n-1}}, \dots, x_{k_1}) = f_X(x_{k_n}|x_{k_{n-1}}). \quad (7.6)$$

马尔可夫序列通常由上式来定义, 但式 (7.2) 的定义更为紧凑。

2° 马尔可夫序列按其相反方向组成的逆序列仍为马尔可夫序列。

对任意的整数  $n$  和  $k$ , 有

$$f_X(x_n|x_{n+1}, x_{n+2}, \dots, x_{n+k}) = f_X(x_n|x_{n+1}). \quad (7.7)$$

证: 由式 (7.4) 知

$$\begin{aligned} f_X(x_n|x_{n+1}, x_{n+2}, \dots, x_{n+k}) &= \frac{f_X(x_n, x_{n+1}, x_{n+2}, \dots, x_{n+k})}{f_X(x_{n+1}, x_{n+2}, \dots, x_{n+k})} \\ &= \frac{f_X(x_{n+k}|x_{n+k-1}) f_X(x_{n+k-1}|x_{n+k-2}) \cdots f_X(x_{n+1}|x_n) f_X(x_n)}{f_X(x_{n+k}|x_{n+k-1}) f_X(x_{n+k-1}|x_{n+k-2}) \cdots f_X(x_{n+2}|x_{n+1}) f_X(x_{n+1})} \\ &= \frac{f_X(x_{n+1}|x_n) f_X(x_n)}{f_X(x_{n+1})} \\ &= \frac{f_X(x_{n+1}, x_n)}{f_X(x_{n+1})} = f_X(x_n|x_{n+1}). \end{aligned} \quad (7.8)$$

3° 马氏序列的条件数学期望满足

$$E [X_n | X_{n-1}, \dots, X_1] = E [X_n | X_{n-1}] . \quad (7.9)$$

### 定义 7.2 鞍

如果马尔可夫序列满足

$$E [X_n | X_{n-1}, \dots, X_1] = X_{n-1} , \quad (7.10)$$

则称此随机序列为“鞍”。



4° 马尔可夫序列中, 若现在已知, 则未来与过去无关。

若  $n > r > s$ , 在假定  $X_r$  已知条件下, 随机变量  $X_n$  与  $X_s$  是独立的。满足

$$f_X (x_n, x_s | x_r) = f_X (x_n | x_r) f_X (x_s | x_r) . \quad (7.11)$$

证: 由式 (7.4) 知

$$\begin{aligned} f_X (x_n, x_s | x_r) &= \frac{f_X (x_n | x_r) f_X (x_r | x_s) f_X (x_s)}{f_X (x_r)} \\ &= \frac{f_X (x_n | x_r) f_X (x_r, x_s)}{f_X (x_r)} \\ &= f_X (x_n | x_r) f_X (x_s | x_r) , \end{aligned} \quad (7.12)$$

可把上述结论推广到具有任意个过去与未来随机变量的情况。

5° 多重马尔可夫序列

多重马尔可夫序列是马尔可夫序列的概念可以推广。

满足式 (7.1) 的随机序列称为 1 重马尔可夫序列。对于任意  $n$ , 满足

$$F_X (x_n | x_{n-1}, x_{n-2}, \dots, x_1) = F_X (x_n | x_{n-1}, x_{n-2}) \quad (7.13)$$

的随机序列称为 2 重马尔可夫序列。依此类推, 可定义多重马尔可夫序列。

6° 齐次马尔可夫序列。

### 定义 7.3 齐次马尔可夫序列

对一般马尔可夫序列而言, 条件概率密度  $f_X (x_n | x_{n-1})$  是  $x$  和  $n$  的函数, 如果条件概率密度  $f_X (x_n | x_{n-1})$  与  $n$  无关, 则称马尔可夫序列是齐次的。用记号  $f_{X_n} (x | X_{n-1} = x_0) = f_X (x | x_0)$  表示  $x_{n-1} = x_0$  条件下  $X_n$  的条件概率密度。



7° 平稳马尔可夫序列。

**定义 7.4 平稳马尔可夫序列**

如果一个马尔可夫序列是齐次的，并且所有的随机变量  $X_n$  具有相同的概率密度，则称马尔可夫序列为平稳的。可以用更精确的记号  $f_{X_n}(x)$  来表示此概率密度，则要求这个函数与  $n$  无关。



**推论** 在一个齐次马尔可夫序列中，若最初的两个随机变量  $X_1$  和  $X_2$  具有相同的概率密度，则此序列是平稳的。

8° 切普曼——柯尔莫哥洛夫 (Chapman-Kolmogorov) 方程。

**定义 7.5 切普曼——柯尔莫哥洛夫方程**

若一个马尔可夫序列的转移概率密度满足

$$f_X(x_n|x_s) = \int_{-\infty}^{\infty} f_X(x_n|x_r) f_X(x_r|x_s) dx_r. \quad (7.14)$$

其中  $n > r > s$  为任意整数，则称该方程为切普曼——柯尔莫哥洛夫方程。



证：对任意三个随机变量  $X_n, X_r, X_s, n > r > s$ ，有

$$\begin{aligned} f_X(x_n|x_s) &= \int_{-\infty}^{\infty} f_X(x_n, x_r|x_s) dx_r = \int_{-\infty}^{\infty} \frac{f_X(x_n, x_r, x_s)}{f_X(x_s)} dx_r \\ &= \int_{-\infty}^{\infty} \frac{f_X(x_n|x_r, x_s) f_X(x_r, x_s)}{f_X(x_s)} dx_r \\ &= \int_{-\infty}^{\infty} f_X(x_n|x_r, x_s) f_X(x_r|x_s) dx_r \\ &= \int_{-\infty}^{\infty} f_X(x_n|x_r) f_X(x_r|x_s) dx_r. \end{aligned} \quad (7.15)$$

最后一步应用了无后效性，即  $f_X(x_n|x_r, x_s) = f_X(x_n|x_r)$ 。反复应用切普曼-柯尔莫哥洛夫方程，我们可根据相邻随机变量的转移概率密度，来求得  $X_s$  条件下  $X_n$  的转移概率密度。

9° 高斯-马尔可夫序列

**定义 7.6 高斯-马尔可夫序列**

如果一个  $n$  维随机序列矢量  $\{X(n)\}$  既是高斯序列，又是马尔可夫序列，则称它为高斯-马尔可夫序列。



高斯-马尔可夫序列的高斯特性决定了它幅度的概率密度分布，而马尔可夫特性则决定了它在时间上的传播方式。这种模型常用在运动目标导弹和飞机的轨迹测量中。

### 7.1.2 马尔可夫链

马尔可夫链就是状态和时间参数皆离散的马尔可夫过程。

#### 定义 7.7 马尔可夫链

随机过程  $X(t)$  在时刻  $t_n (n = 1, 2, \dots)$  的采样为  $X_n = X(t_n)$ , 且  $X_n$  可能取得的状态必为  $a_1, a_2, \dots, a_N$  之一, 其中  $A_I = \{a_1, a_2, \dots, a_N\}$  为有限的状态空间,  $I = \{1, 2, \dots, N\}$ 。随机过程只在  $t_1, t_2, \dots, t_n, \dots$  可列个时刻发生状态转移。若随机过程  $X(t)$  在  $t_{m+k}$  时刻变成任一状态  $a_j$  的概率, 只与过程在  $t_m$  时刻的状态  $a_i$  有关, 而与过程在  $t_m$  时刻以前的状态无关, 则称此随机过程为马尔可夫链, 简称为马氏链。



可用公式表示为

$$P\{X_{m+k} = a_j | X_m = a_i, X_{m-1} = a_p, \dots, X_1 = a_q\} = P\{X_{m+k} = a_j | X_m = a_i\} \quad (7.16)$$

实际上, 过程  $X(t)$  是状态离散的随机序列  $\{X_n\}$ , 所以上式可以看成由式(7.1)演变而来。

#### 2. 马氏链的转移概率及其转移概率矩阵

##### (1) 马氏链的转移概率

马氏链“在  $m$  时刻出现的状态为  $a_i$  的条件下,  $t_{m+k}$  时刻出现的状态为  $a_j$ ”的条件概率可用  $p_{ij}(m, m+k)$  表示, 即

$$p_{ij}(m, m+k) = P\{X_{m+k} = a_j | X_m = a_i\}, \quad (7.17)$$

式中  $i, j = 1, 2, \dots, N$ , 且  $m, k$  皆为正整数, 则称  $p_{ij}(m, m+k)$  为马氏链的转移概率。一般而言,  $p_{ij}(m, m+k)$  不仅依赖于  $i, j, k$ , 而且还依赖于  $m$ 。若  $p_{ij}(m, m+k)$  与  $m$  无关, 则称此马氏链为齐次的。下面只讨论齐次马氏链, 并通常将“齐次”二字省去。

##### (2) 一步转移概率及其转移概率矩阵

当  $k = 1$  时, 马氏链由状态  $a_i$  经过一次转移就到达状态  $a_j$  的转移概率称为一步转移概率, 通常用符号  $p_{ij}$  表示, 即

$$p_{ij} = p_{ij}(m, m+1) = P\{X_{m+1} = a_j | X_m = a_i\}, \quad i, j \in I \quad (7.18)$$

由所有状态  $I = \{1, 2, \dots, N\}$  之间的一步转移概率  $p_{ij}$  构成的矩阵, 称为马氏链的一步转移概率矩阵, 简称为转移概率矩阵,

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ \vdots & \vdots & & \vdots \\ p_{N1} & p_{N2} & \cdots & p_{NN} \end{bmatrix} \quad (7.19)$$

此矩阵决定了状态  $X_1, X_2, \dots, X_N$  转移的概率法则, 具有下列两个性质:

$$1^\circ 0 \leq p_{ij} \leq 1; \quad 2^\circ \sum_{j=1}^N p_{ij} = 1 \quad (7.20)$$

表示转移概率矩阵是一个每行元素之和为 1 的非负元素矩阵。因  $p_{ij}$  为条件概率, 故性质 1 是显然的, 性质 2 可由下式推得:

$$\begin{aligned} \sum_{j=1}^N p_{ij} &= \sum_{j=1}^N P\{X_{m+1} = a_j | X_m = a_i\} \\ &= P\{X_{m+1} = a_1 | X_m = a_i\} + \dots + P\{X_{m+1} = a_N | X_m = a_i\} = 1. \end{aligned} \quad (7.21)$$

任意满足性质 1 和性质 2 的矩阵也称之为随机矩阵。

### (3) $n$ 步转移概率及其转移概率矩阵

与一步转移概率类似, 当  $k = n$  时, 定义马氏链的  $n$  步转移概率  $p_{ij}(n)$  为

$$p_{ij}(n) = p_{ij}(m, m+n) = P(X_{m+n} = a_j | X_m = a_i), \quad n \geq 1. \quad (7.22)$$

表明马氏链在时刻  $t_m$  的状态为  $a_i$  的条件下, 经过  $n$  步转移到达状态  $a_j$  的概率。对应的  $n$  步转移概率矩阵为

$$P(n) = \begin{bmatrix} p_{11}(n) & p_{12}(n) & \cdots & p_{1N}(n) \\ p_{21}(n) & p_{22}(n) & \cdots & p_{2N}(n) \\ \vdots & \vdots & & \vdots \\ p_{N1}(n) & p_{N2}(n) & \cdots & p_{NN}(n) \end{bmatrix}. \quad (7.23)$$

它也是随机矩阵, 显然具有如下性质

$$1^\circ 0 \leq p_{ij}(n) \leq 1; \quad 2^\circ \sum_{j=1}^N p_{ij}(n) = 1 \quad (7.24)$$

当  $n = 1$  时,  $p_{ij}$  就是一步转移概率, 即  $p_{ij}(n) = p_{ij}(1) = p_{ij} = p_{ij}(m, m+1)$ 。通常还规定

$$p_{ij}(0) = p_{ij}(m, m) = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad (7.25)$$

### (4) $n$ 步转移概率 $p_{ij}(n)$ 与一步转移概率 $p_{ij}(1)$ 的关系

对于  $n$  步转移概率, 将问题利用动态规划原理, 划分成 2 个阶段, 利用切普曼-柯尔莫哥洛夫方程的离散形式:

$$p_{ij}(n) = p_{ij}(l+k) = \sum_{r=1}^N p_{ir}(l)p_{rj}(k), \quad n = l+k, \quad (7.26)$$

其中,  $l$  表示第一阶段的转移步数,  $k$  表示第二阶段的转移步数, 且  $n = l + k$ ;  $r$  是马尔科夫链由状态  $a_i$  经过  $l$  步转移到达状态数  $N, r = 1, 2, \dots, N$  指示  $N$  个状态数。

证: 由全概率公式可得

$$\begin{aligned}
p_{ij}(n) &= p_{ij}(n \equiv l + k) = P\{X_{m+l+k} = a_j | X_m = a_i\} = \frac{P\{X_m = a_i, X_{m+l+k} = a_j\}}{P\{X_m = a_i\}} \\
&= \sum_{r=1}^N \frac{P\{X_m = a_i, X_{m+l+k} = a_j, X_{m+l} = a_r\}}{P\{X_m = a_i, X_{m+l} = a_r\}} \cdot \frac{P\{X_m = a_i, X_{m+l} = a_r\}}{P\{X_m = a_i\}} \\
&= \sum_{r=1}^N P\{X_{m+l+k} = a_j | X_{m+l} = a_r, X_m = a_i\} \cdot P\{X_{m+l} = a_r | X_m = a_i\} \\
&= \sum_{r=1}^N p_{ir}(l)p_{rj}(k).
\end{aligned} \tag{7.27}$$

再利用马氏链的无后效性与齐次性, 可得

$$\begin{cases} P\{X_{m+l+k} = a_j | X_{m+l} = a_r, Z_m = a_i\} = p_{rj}(k) \\ P\{X_{m+l} = a_r | X_m = a_i\} = p_{ir}(l) \end{cases} \tag{7.28}$$

式 (7.26) 表明: 由于马氏链的无后效性与齐次性, 该链从状态  $a_i$  经过  $n$  步转移到达状态  $a_j$  这一事件  $(a_i \xrightarrow{n} a_j)$ , 等效于该链先由状态  $a_i$  经过  $l$  步转移到达状态  $a_r$ , ( $r = 1, 2, \dots, N$ ), 再由状态  $a_r$  经过  $k$  步转移到达状态  $a_j$  这一事件  $(a_i \xrightarrow{l} a_r \xrightarrow{k} a_j)$ , 也就是说, 只要  $r \in I = \{1, 2, \dots, N\}$  中有一个事件  $(a_i \xrightarrow{l} a_r \xrightarrow{k} a_j)$  发生, 则事件  $(a_i \xrightarrow{n} a_j)$  就必发生。因此事件的概率是  $r \in I$  中所有事件  $(a_i \xrightarrow{l} a_r \xrightarrow{k} a_j)$  概率的和。

① 当  $l = 1, k = 1$  时

$$p_{ij}(2) = \sum_{r=1}^N p_{ir}(1)p_{rj}(1) = \sum_{r=1}^N p_{ir}p_{rj}. \tag{7.29}$$

② 当  $l = 1, k = 2$  时

$$p_{ij}(3) = \sum_{r=1}^N p_{ir}(1)p_{rj}(2) = \sum_{r=1}^N p_{ir} \sum_{k=1}^N p_{rk}p_{kj}. \tag{7.30}$$

③ 以此类推

$$p_{ij}(n) = \sum_{r=1}^N p_{ir}(1)p_{rj}(n-1) = \sum_{r=1}^N p_{ir}p_{rj}(n-1). \tag{7.31}$$

上述公式给出了随机矩阵具体元素的计算方法。

同理，可得离散切普曼-柯尔莫哥洛夫方程的矩阵形式：

$$\mathbf{P}(n) = \mathbf{P}(l+k) = \mathbf{P}(l)\mathbf{P}(k). \quad (7.32)$$

① 当  $n = 2$  时

$$\mathbf{P}(2) = \mathbf{P}(1)\mathbf{P}(1) = [\mathbf{P}(1)]^2 = \mathbf{P}^2. \quad (7.33)$$

一般将一步转移概率矩阵  $\mathbf{P}(1)$  简写为  $\mathbf{P}$ 。

② 当  $n = 3$  时

$$\mathbf{P}(3) = \mathbf{P}(1)\mathbf{P}(2) = \mathbf{P}(1)[\mathbf{P}(1)]^2 = \mathbf{P}^3. \quad (7.34)$$

③ 当  $n$  为任意正整数时

$$\mathbf{P}(n) = \mathbf{P}(1)\mathbf{P}(n-1) = \cdots = \mathbf{P}^n. \quad (7.35)$$

上式表明： $n$  步转移概率矩阵等于一步转移概率矩阵的  $n$  次方。

由此可见，以一步转移概率  $p_{ij}$  为元素的转移概率矩阵  $\mathbf{P}$  决定了马氏链状态转移过程的概率法则。

(5) 马氏链的有限维分布

1) 初始分布

### 定义 7.8 初始概率

马氏链在  $t = 0$  时所处状态  $a_i$  的概率，通常被称作“初始概率”。



$$p_i(0) = p\{X_0 = a_i\} = p_i, \quad i \in I = \{1, \dots, N\}, \quad (7.36)$$

且有  $0 \leq p_i \leq 1, \sum_{i=1}^N p_i = 1$  成立。

而对于  $N$  个状态而言，所有初始概率的集合  $\{p_{ij}\}$  称为马氏链的“初始分布”

$$\{p_i\} = (p_1, \dots, p_i, \dots, p_N). \quad (7.37)$$

2) 一维分布

### 定义 7.9 马氏链的一维分布

马氏链在第  $n$  步所处状态为  $a_i$  的无条件概率称为马氏链的“一维分布”，也称为“状态概率”。表示为

$$p\{X_n = a_j\} = p_j(n), \quad j \in I = \{1, \dots, N\} \quad (7.38)$$

且有  $0 \leq p(n) \leq 1, p(n) = 1$  成立。



由全概率公式,一维分布可表示为

$$\begin{aligned} P_j(n) &= \sum_{i=1}^N P\{X_n = a_j | X_s = a_i\} P\{X_s = a_i\} \\ &= \sum_{i=1}^N p_i(s) p_{ij}(n-s), i, j \in I. \end{aligned} \quad (7.39)$$

上式给出了不同时刻一维分布  $p_i(s), p_j(n)$  以及  $n-s$  步转移概率  $p_{ij}(n-s)$  之间的关系。

① 当  $s = 0$  时

$$p_j(n) = \sum_{i=1}^N p_i p_{ij}(n), \quad j \in I. \quad (7.40)$$

② 当  $s = n - 1$  时

$$p_j(n) = \sum_{i=1}^N p_i(n-1) p_{ij}, \quad j \in I. \quad (7.41)$$

若将一维分布表示成矢量形式

$$\mathbf{p}(n) = \begin{bmatrix} p_1(n) \\ p_2(n) \\ \vdots \\ p_N(n) \end{bmatrix}_{N \times 1} \quad (7.42)$$

称之为“一维分布矢量”或“状态概率矢量”。且其递推公式 (7.39) 可表示为

$$\mathbf{p}(n) = \mathbf{P}^T(n-s) \mathbf{p}(s). \quad (7.43)$$

3)  $n$  维分布

齐次马氏链在  $t = 0, 1, 2, \dots, n-1$  时刻分别取得状态  $a_{i0}, a_{i1}, a_{i2}, \dots, a_{i(n-1)}$ ,  $(i_0, i_1, \dots, i_{n-1} \in I)$ 。这一事件的概率为  $P\{X_0 = a_{i0}, X_1 = a_{i1}, \dots, X_{n-1} = a_{i(n-1)}\}$  称为马氏链的  $n$  维分布。由全概率公式和无后效性, 可证下式成立

$$\begin{aligned} P\{X_0 = a_{i0}, X_1 = a_{i1}, \dots, X_{n-1} = a_{i(n-1)}\} \\ &= P\{X_0 = a_{i0}\} P\{X_1 = a_{i1} | X_0 = a_{i0}\} \\ &\quad \cdots P\{X_{n-1} = a_{i(n-1)} | X_0 = a_{i0}, \dots, X_{n-2} = a_{i(n-2)}\} \\ &= P\{X_0 = a_{i0}\} P\{X_1 = a_{i1} | X_0 = a_{i0}\} \cdots P\{X_{n-1} = a_{i(n-1)} | X_{n-2} = a_{i(n-2)}\} \\ &= p_{i0} p_{i1, i1} \cdots p_{i(n-2), i(n-1)}. \end{aligned} \quad (7.44)$$

由于  $I = \{1, \dots, N\}$ ,  $a_{i0}, a_{i1}, a_{i2}, \dots, a_{i(n-1)}$  分别可以是  $N$  个状态中的任意一个, 因此, 这种  $n$  维分布可以有许多种。

通过马氏链的一维分布和  $n$  维分布的讨论可知, 马氏链的任意有限维分布完全可以由初始分布和一步转移概率矩阵所确定。因此, 初始分布和一步转移概率矩阵是描述马氏链统计特性的两个重要的分布特征。马氏链在研究质点的随机运动、自动控制、通信技术、气象预报和生物遗传工程等方面皆有广泛的应用。

下面是马氏链应用的一些例子

**例 7.10** 设质点  $M$  在直线段上作随机游动, 如图 7-2 所示。假定质点  $M$  只能停留在  $1, 2, \dots, N$  点上, 且只在  $t_1, t_2, \dots, t_n, \dots$  时刻发生游动。游动的概率法则是:

- ① 若质点  $M$  原来处于  $2, \dots, N-1$  这些点上, 则分别以  $p(0 < p < 1)$  的概率向右移动一步或以  $q(q = 1 - p)$  的概率向左移动一步。
- ② 若质点  $M$  原来处于 1 点, 则以概率 1 移动到 2 点。
- ③ 若质点  $M$  原来处于  $N$  点, 则以概率 1 移动到  $N-1$  点上。求其转移概率矩阵。

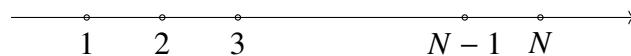


图 7-1 质点  $M$  的随机游动

解: 若以  $X_n = i (i = 1, 2, \dots, N)$  表示质点  $M$  在时刻  $t_n$  位于  $i$  点, 则不难看出  $X_1, X_2, \dots$  是一个齐次马氏链。由于其一步转移概率  $p_{ij}$  为

$$\begin{cases} p_{i,i+1} = p, & 2 \leq i \leq N-1 \\ p_{i,i-1} = q, & 2 \leq i \leq N-1 \\ p_{i,j} = 0, & j \neq i+1, i-1 \\ p_{12} = p_{N,N-1} = 1 \end{cases} \quad (7.45)$$

故质点  $M$  游动的转移概率矩阵为

$$P = \begin{pmatrix} 0 & 1 & & & & & \\ q & 0 & p & & & & 0 \\ & q & & \ddots & & & \\ & \ddots & & & p & & \\ 0 & & & & q & 0 & p \\ & & & & & 1 & 0 \end{pmatrix}. \quad (7.46)$$

$P$  为  $N \times N$  阶方阵。因为质点不能越过 1 和  $N$  两端点, 将 1 和  $N$  这两点称为反射壁(或反射状态), 故称上述游动为一维不可越壁的随机游动, 它仅是一维随机游动的一种。

如果我们改变质点游动的概率法则(即转移概率)就可得到不同类型的随机游动过程。若游动的概率法则改动为:  $n$  当质点  $M$  一旦到达 1 点或  $N$  点, 就永远停留在 1 或  $N$  处, 其他不变。我们将 1 和  $N$  这两点称为吸收壁, 因此得到的是带有两个吸收壁的随机游动。它也是一个齐次马氏链, 由于其一步转移概率为

$$\begin{cases} p_{i,i+1} = p, & 2 \leq i \leq N-1 \\ p_{i,i-1} = q, & 2 \leq i \leq N-1 \\ p_{i,j} = 0, & j \neq i+1, i-1 \\ p_{12} = p_{N,N-1} = 1 \end{cases} \quad (7.47)$$

故质点  $M$  游动的转移概率矩阵为

$$P = \begin{pmatrix} 1 & 0 & & & & 0 \\ q & 0 & p & & & \\ & q & & \ddots & & \\ & & \ddots & & p & \\ 0 & & & q & 0 & p \\ & & & 0 & 0 & 1 \end{pmatrix}. \quad (7.48)$$

它仍是  $N \times N$  阶方阵。

**例 7.11** 在某数字通信系统中传递 0、1 两种信号, 且传递要经过若干级。因为系统中存在噪声, 各级将会造成错误。若某级输入 0、1 数字信号后, 其输出不产生错误的概率为  $p$  (即各级正确传递信息的概率), 产生错误的概率为  $q = 1 - p$ 。该级的输入状态和输出状态构成了一个两状态的马氏链, 它的一步转移概率矩阵为

$$P = \begin{bmatrix} p & q \\ q & p \end{bmatrix}, \quad (7.49)$$

于是, 二步转移概率矩阵为

$$P(2) = (P)^2 = \begin{bmatrix} p & q \\ q & p \end{bmatrix} \begin{bmatrix} p & q \\ q & p \end{bmatrix} = \begin{bmatrix} p^2 + q^2 & 2pq \\ 2pq & p^2 + q^2 \end{bmatrix}. \quad (7.50)$$

**例 7.12** 天气预报问题。若明天是否降雨只与今日的天气(是否有雨)有关, 而与以往的天气无关。并设今日有雨而明日也有雨的概率为 0.6, 今日无雨而明日有雨的概率为 0.3。

另外,假定将“有雨”称作“1”状态天气,而把“无雨”称为“2”状态天气,则本例属于一个两状态的马氏链。

- ① 试求其一步至四步转移概率矩阵。
- ② 今日有雨而后日(第三日)仍有雨的概率为多少?
- ③ 今日有雨而第四日无雨的概率为多少?
- ④ 今日无雨而第五日有雨的概率为多少?

解:由题可知,一步转移概率为

$$\begin{aligned} p_{11} &= 0.6, & p_{12} &= 1 - 0.6 = 0.4 \\ p_{21} &= 0.3, & p_{22} &= 1 - 0.3 = 0.7 \end{aligned} \quad (7.51)$$

① 此马氏链的一步转移概率矩阵为

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{bmatrix}. \quad (7.52)$$

二步转移概率矩阵为

$$\mathbf{P}(2) = (\mathbf{P})^2 = \begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{bmatrix} \cdot \begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{bmatrix} = \begin{bmatrix} 0.48 & 0.52 \\ 0.39 & 0.61 \end{bmatrix}. \quad (7.53)$$

三步转移概率矩阵为

$$\mathbf{P}(3) = (\mathbf{P})^3 = (\mathbf{P})^2 \mathbf{P} = \begin{bmatrix} 0.48 & 0.52 \\ 0.39 & 0.61 \end{bmatrix} \cdot \begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{bmatrix} = \begin{bmatrix} 0.444 & 0.556 \\ 0.417 & 0.583 \end{bmatrix}. \quad (7.54)$$

四步转移概率矩阵为

$$\mathbf{P}(4) = (\mathbf{P})^4 = (\mathbf{P})^2 (\mathbf{P})^2 = \begin{bmatrix} 0.48 & 0.52 \\ 0.39 & 0.61 \end{bmatrix}^2 = \begin{bmatrix} 0.4332 & 0.5668 \\ 0.4251 & 0.5749 \end{bmatrix}. \quad (7.55)$$

- ② 今日有雨而第三日仍有雨的概率为  $p_{11}(2) = 0.48$ .
- ③ 今日有雨而第四日无雨的概率为  $p_{12}(3) = 0.556$ .
- ④ 今日无雨而第五日有雨的概率为  $p_{21}(4) = 0.4251$ .

### 3. 马氏链的平稳分布与遍历性

#### (1) 平稳分布

### 定义 7.13 马氏链的平稳分布

若一个马氏链的概率分布  $P\{X = a_j\} = p_j$  满足

$$p_j = \sum_{i \in I} p_i p_{ij}, \quad j \in I. \quad (7.56)$$

其中  $p_j \geq 0, \sum_{j \in I} p_j = 1$  成立。则称  $\{p_j\} = \{p_1, p_2, \dots, p_N\}$  为该马氏链的“平稳分布”。 ♣

对于平稳分布  $\{p_j\}$ , 有

$$p_j = \sum_{i \in I} \left( \sum_{k \in I} p_k p_{ki} \right) p_{ij} = \sum_{k \in I} p_k \left( \sum_{i \in I} p_{ki} p_{ij} \right) = \sum_{k \in I} p_k p_{kj}(2), \quad j \in I. \quad (7.57)$$

类似可推

$$p_j = \sum_{i \in I} p_i p_{ij}(n), \quad j \in I. \quad (7.58)$$

比较式 (7.56) 与式 (7.58) 可知, 对平稳分布而言, 无论是一步转移到状态  $a_j$ , 还是  $n$  步转移到状态  $a_j$ , 其概率分布不变, 与转移时间  $n$  无关。

**推论** 如果齐次马氏链的初始分布  $\{p_i\}$  是平稳分布, 则对  $n \geq 1$  步后,  $X_n$  的分布  $\{p_j(n)\}$  也是平稳分布, 其中  $p_i = P\{X_0 = a_i\}, p_j(n) = P\{X_n = a_j\}$ 。

证: 由式(7.40) 与式 (7.58), 可得

$$p_j(n) = \sum_{i=1}^N p_i p_i(n) = p_j, \quad j \in I. \quad (7.59)$$

若用概率矢量表示, 则有

$$p(n) = p(0) = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_N \end{bmatrix}_{N \times 1}. \quad (7.60)$$

### 定义 7.14 链的平稳分布

若齐次马氏链的概率分布不随时间  $n$  的变化而改变, 即满足式 (7.59) 或式 (7.60), 则称此链为**平稳马氏链**。称  $\{p_j\} = \{p_1, p_2, \dots, p_N\}$  为该链的平稳分布。 ♣

由于  $p_j(n) = P\{X_n = a_j\}, p_j = P\{X_0 = a_j\}$ , 由式 (7.51) 可得  $p_j(n) = p_i$ , 即  $P\{X_n = a_j\} = P\{X_0 = a_j\}$ , 表示平稳马氏链的一维分布不随时间  $n$  的变化而改变。

若对平稳马氏链的  $m$  维分布在时间上平移  $n$ , 可得

$$\begin{aligned} & P \{ X_{0+n} = a_{i0}, X_{1+n} = a_{i1}, \dots, X_{m-1+n} = a_{i(m-1)} \} \\ &= P \{ X_{0+n} = a_{i0} \} P \{ X_{1+n} = a_{i1} | X_0 = a_{i0} \} \cdots P \{ X_{m-1+n} = a_{i(m-1)} | X_{m-2+n} = a_{i(m-2)} \} \quad (7.61) \\ &= p_{i0}(n) p_{i0,i1} \cdots p_{i(m-2),i(m-1)}. \end{aligned}$$

对于平稳马氏链,  $p_{i0}(n) = p_{i0}(0) = p_i$ , 则

$$\begin{aligned} & P \{ X_{0+n} = a_{i0}, X_{1+n} = a_{i1}, \dots, X_{m-1+n} = a_{i(m-1)} \} \\ &= p_{i0} p_{i0,i1} \cdots p_{i(m-2),i(m-1)} \quad (7.62) \\ &= P \{ X_0 = a_{i0}, X_1 = a_{i1}, \dots, X_{m-1} = a_{i(m-1)} \}. \end{aligned}$$

可见, 该马氏链的  $m$  维分布也不随时间的平移而变化, 说明平稳马氏链是个平稳过程。

## (2) 遍历性

例 7.10 中的马氏链, 求得的各步转移概率矩阵为

$$\mathbf{P} = \begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{bmatrix}, \mathbf{P}(2) = \begin{bmatrix} 0.48 & 0.52 \\ 0.39 & 0.61 \end{bmatrix}, \mathbf{P}(4) = \begin{bmatrix} 0.4332 & 0.5668 \\ 0.4251 & 0.5749 \end{bmatrix}. \quad (7.63)$$

若再求  $\mathbf{P}(8) = \mathbf{P}(4)\mathbf{P}(4)$ , 则

$$\mathbf{P}(8) = \begin{bmatrix} 0.4286 & 0.5714 \\ 0.4285 & 0.5714 \end{bmatrix}. \quad (7.64)$$

可以看出, 随着步长  $n$  增大, 此马氏链的转移概率  $p_1$  与  $p_{21}$ ,  $p_{12}$  与  $p_{22}$  差距越来越小。设想若取  $n \rightarrow \infty$ , 则必有

$$\lim_{n \rightarrow \infty} \mathbf{P}(n) = \begin{bmatrix} p_1 & p_2 \\ p_1 & p_2 \end{bmatrix}. \quad (7.65)$$

即不论从哪个状态  $i$  出发, 只要终点状态  $j$  相同, 则其转移概率相同。

算法 7-1 马氏链的转移概率矩阵的极限.

```

1 p=[.6 0.4;0.3,0.7]
2 p^2
3 p^4
4 p^8
5 p^40
6 p^100
7
8 norm(p^100-p^40,2)

```

### 定义 7.15 马氏链的遍历性

如果一个齐次马氏链对于一切状态  $i$  和  $j$ , 存在不依赖于  $i$  的极限

$$\lim_{n \rightarrow \infty} p_{ij}(n) = p_j, \quad (7.66)$$

则称此马氏链具有遍历性, 这里的  $p(n)$  为此链的  $n$  步转移概率。



由上述定义可知, 当  $n \rightarrow \infty$  时,  $n$  步转移概率  $p_{ij}(n)$  趋近于一个与初始状态  $i$  无关的  $p_j$ 。换言之, 不论过程自哪一状态  $i$  出发, 当转移步数  $n$  充分大时, 转移到达状态  $j$  的概率都趋近于  $p_j$ 。对  $\forall j \in I, p_j$ , 是一种概率分布  $\{p_j\}$ , 满足

$$\sum_{j=1}^N p_j = 1. \quad (7.67)$$

此时  $p_{ij}$  称为极限分布。比较前面介绍的平稳分布可以看出, 马氏链的遍历性可以导致  $n \rightarrow \infty$  的平稳性, 因此平稳分布就是具有遍历性的马氏链的极限分布。

$$\lim_{n \rightarrow \infty} P(n) = \begin{bmatrix} p_1 & p_2 & \cdots & p_N \\ p_1 & p_2 & \cdots & p_N \\ \vdots & \vdots & & \vdots \\ p_1 & p_2 & \cdots & p_N \end{bmatrix}. \quad (7.68)$$

物理上可以理解为: 不管初始状态如何, 系统经过一段时间后 ( $n \rightarrow \infty$ ) 走到稳定状态(平稳状态), 系统的宏观状态不再随时间变化, 即系统处于各个状态的概率不再随时间变化, 是一平稳分布。

以上给出了马氏链具有遍历性的基本定义, 下面的定理给出马氏链具有遍历性的一个简单的充分条件以及求极限分布  $\{p_j\}$  的方法。

### 定理 7.16 有限马氏链具有遍历性的充分条件

对于一有限状态的马氏链, 若存在正整数  $m$ , 使所有的状态满足

$$p_{ij}(m) > 0, \quad i, j \in I \quad (7.69)$$

则马尔科夫链是遍历的。



由于遍历性的马氏链的极限分布  $\{p_j\}$  就是平稳马氏链中的平稳分布, 可推出极限

分布  $\{p_j\}$  是下面方程组的唯一解。

$$\begin{cases} p_j = \sum_{i=1}^N p_i p_{ij} \\ \sum_{j=1}^N p_j = 1 \end{cases} \text{ 或 } \begin{cases} \mathbf{p} = \mathbf{P}^\top \mathbf{p} \\ \sum_{j=1}^N p_j = 1 \end{cases}, \quad \text{其中, } \mathbf{p} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_N \end{bmatrix} \quad (7.70)$$

**例 7.17** 设有三个状态  $\{1, 2, 3\}$  的马氏链, 它的一步转移概率矩阵为

$$\mathbf{P} = \begin{bmatrix} q & p & 0 \\ q & 0 & p \\ 0 & q & p \end{bmatrix}, 0 < p < 1, q = 1 - p \quad (7.71)$$

① 问此链是否具有遍历性? ② 求其极限分布  $\{p_j\}$ 。

解: ① 显然, 当  $m = 1$  时, 有  $\mathbf{P}(1) = \mathbf{P}$ , 因  $\mathbf{P}$  中三个元素  $p_{13}, p_{22}, p_{31}$  为零, 不满足有限马氏链具有遍历性的充分条件。当  $m = 2$  时, 由于

$$\mathbf{P}(2) = (\mathbf{P})^2 = \begin{bmatrix} q^2 + pq & pq & p^2 \\ q^2 & 2pq & p^2 \\ q^2 & pq & pq + p^2 \end{bmatrix} \quad (7.72)$$

的元素皆大于零, 满足遍历的充分条件, 所以此马氏链具有遍历性。

② 据式 (7.70) 有方程组

$$\begin{cases} p_1 = p_1 p_{11} + p_2 p_{21} + p_3 p_{31} \\ p_2 = p_1 p_{12} + p_2 p_{22} + p_3 p_{32} \\ p_3 = p_1 p_{13} + p_2 p_{23} + p_3 p_{33} \end{cases} \quad \text{且} \quad p_1 + p_2 + p_3 = 1 \quad (7.73)$$

将已知条件代入上式可得

$$\begin{cases} P_1 = p_1 q + p_2 q \\ P_2 = p_1 p + p_3 q \\ P_3 = p_2 p + p_3 p \end{cases} \quad \text{且} \quad p_1 + p_2 + p_3 = 1 \quad (7.74)$$

由此解得

$$p_1 = \left[ 1 + \frac{p}{q} + \left( \frac{p}{q} \right)^2 \right]^{-1}, \quad p_2 = \frac{p}{q} p_1, \quad p_3 = \left( \frac{p}{q} \right)^2 p_1 \quad (7.75)$$

可归纳为

$$p_j = \frac{1 - \frac{p}{q}}{1 - \left( \frac{p}{q} \right)^3} \left( \frac{p}{q} \right)^{j-1}, \quad j = 1, 2, 3 \quad (7.76)$$

若  $p = q = 1/2$ , 则得  $p_1 = p_2 = p_3 = 1/3$ .

**例 7.18** 设有两个状态  $\{1, 2\}$  的马氏链, 一步转移概率矩阵为

$$\mathbf{P} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}. \quad (7.77)$$

试问此链是否具有遍历性, 为什么?

解: 因为

$$\left\{ \begin{array}{l} \mathbf{P}(1) = \mathbf{P} = \mathbf{I} \\ \mathbf{P}(2) = (\mathbf{P})^2 = \mathbf{I} \\ \mathbf{P}(3) = (\mathbf{P})^3 = \mathbf{I} \\ \dots \\ \mathbf{P}(n) = (\mathbf{P})^n = \mathbf{I}. \end{array} \right. \quad (7.78)$$

无论  $n$  为多大, 始终有  $p_{12}(n) = p_{21}(n) = 0$ , 不能满足遍历性的充分条件, 故此链不具有遍历性。

### 7.1.3 马尔可夫过程

#### 定义 7.19 马尔可夫过程

设有一随机过程  $X(t), t \in T$ , 若在  $t_1, t_2, \dots, t_{n-1}, t_n (t_1 < t_2 < \dots < t_{n-1} < t_n \in T)$  时刻对  $X(t)$  观测得到相应的观测值  $x_1, x_2, \dots, x_{n-1}, x_n$  满足条件

$$\begin{aligned} P \{X(t_n) \leq x_n | X(t_{n-1}) = x_{n-1}, X(t_{n-2}) = x_{n-2}, \dots, X(t_1) = x_1\} \\ = P \{X(t_n) \leq x_n | X(t_{n-1}) = x_{n-1}\} \end{aligned} \quad (7.79)$$

或

$$\begin{aligned} F_X(x_n; t_n | x_{n-1}, x_{n-2}, \dots, x_2, x_1; t_{n-1}, t_{n-2}, \dots, t_2, t_1) \\ = F_X(x_n; t_n | x_{n-1}; t_{n-1}) \end{aligned} \quad (7.80)$$

则称此类过程为具有马尔可夫性质 (无后效性) 的过程或马尔可夫过程, 简称马氏过程。

其中  $F_X(x_n; t_n | x_{n-1}, \dots, x_2, x_1; t_{n-1}, \dots, t_2, t_1)$  代表  $X(t_{n-1}) = x_{n-1}, \dots, X(t_2) = x_2, X(t_1) = x_1$  的条件下,  $n$  时刻  $X(t_n)$ , 取  $x_n$  值的条件分布函数。

若把  $t_{n-1}$  看作 “现在”, 因为  $t_1 < t_2 < \dots < t_{n-1} < t_n$ , 则  $t_n$  就可看成 “将来”  $t_1, t_2, \dots, t_{n-2}$  就当作 “过去”。因此, 上述定义中的条件可表述为: 在 “现在” 状态  $X(t_{n-1})$  取值为  $x_{n-1}$  的条件下, “将来状态  $X(t_n)$  与 “过去” 状态  $X(t_{n-2}), X(t_{n-3}), \dots, X(t_1)$  是无关的。

## (2) 转移概率分布

**定义 7.20** 马尔可夫过程的条件概率分布

马氏过程的转移概率分布为

$$F_X(x_n; t_n | x_{n-1}; t_{n-1}) = P\{X(t_n) \leq x_n | X(t_{n-1}) = x_{n-1}\} \quad (7.81)$$

或

$$F_X(x; t | x_0; t_0) = P\{X(t) \leq x | X(t_0) = x_0\}, \quad t > t_0 \quad (7.82)$$

转移概率分布是条件概率分布。



对  $X$  而言, 它是一个分布函数, 有以下性质:

- 1°  $F_X(x; t | x_0; t_0) \geq 0$
- 2°  $F_X(\infty; t | x_0, t_0) = 1$
- 3°  $F_X(-\infty; t | x_0; t_0) = 0$
- 4°  $F_X(x; t | x_0; t_0)$  是关于  $x$ , 的单调非降、右连续的函数.
- 5° 满足切普曼-科尔莫哥洛夫方程:

$$\begin{cases} F_X(x; t | x_0, t_0) = \int_{-\infty}^{\infty} F_X(x; t | x_1, t_1) dF_X(x_1, t_1 | x_0, t_0) \\ dF_X(x_1; t_1 | x_0, t_0) = F_X(x_1; t_1 | x_0, t_0) dx_1, \quad t_0 < t_1 < t \end{cases} \quad (7.83)$$

应用全概率公式, 可以证明上式成立。

## (3) 转移概率密度

如果  $F_X(x; t | x_0, t_0)$  关于  $x$  的导数存在, 则

$$F_X(x; t | x_0, t_0) = \frac{\partial}{\partial x} F_X(x; t | x_0, t_0) \quad (7.84)$$

称之为马尔可夫过程的转移概率密度。反之, 可得

$$\int_{-\infty}^x F_X(u; t | x_0, t_0) du = \int_{-\infty}^x dF_X(u; t | x_0, t_0) = F_X(x; t | x_0, t_0). \quad (7.85)$$

并且还有

$$\begin{aligned} \int_{-\infty}^{\infty} F_X(x; t | x_0, t_0) dx &= F_X(\infty; t | x_0, t_0) = 1 \\ F_X(x; t | x_0, t_0) &\xrightarrow{t \rightarrow t_0} \delta(x - x_0). \end{aligned} \quad (7.86)$$

此时, 无后效性可表示为

$$\begin{aligned} F_X(x_n; t_n | x_{n-1}, x_{n-2}, \dots, x_2, x_1; t_{n-1}, t_{n-2}, \dots, t_2, t_1) \\ = F_X(x_n; t_n | x_{n-1}; t_{n-1}). \end{aligned} \quad (7.87)$$

(4) 马氏过程的转移概率密度也满足切普曼-科尔莫哥洛夫方程

$$F_X(x_n; t_n | x_k; t_k) = \int_{-\infty}^{\infty} F_X(x_n; t_n | x_r; t_r) F_X(x_r; t_r | x_k; t_k) dx_r, \quad t_n > t_r > t_k. \quad (7.88)$$

证: 利用概率的乘法定理及马氏过程的无后效性, 可知

$$\begin{aligned} & F_X(x_n, x_r; t_n, t_r | x_k; t_k) \\ &= \frac{F_X(x_n, x_r, x_k; t_s, t_r, t_k)}{F_X(x_k; t_k)} \\ &= \frac{F_X(x_n; t_n | x_r, x_k, t_r, t_k) F_X(x_r, x_k, t_r, t_k)}{F_X(x_k + t_k)} \\ &= F_X(x_n; t_n | x_r; t_r) F_X(x_r; t_r | x_k + t_k). \end{aligned} \quad (7.89)$$

并代入

$$F_X(x_0; t_n | x_k; t_k) = \int_{-\infty}^{\infty} F_X(x_n, x_r; t_n, t_r | x_k; t_k) dx_r. \quad (7.90)$$

可得转移概率密度的切普曼-科尔莫哥洛夫方程。

### 定义 7.21 齐次马尔可夫过程

若马氏过程的转移概率分布  $F_X(x; t|x_0; t_0)$  或转移概率密度  $F_X(x; t|x_0; t_0)$ , 只与转移前后的状态  $x_0, x$  及相应的时间差  $t - t_0 = \tau$  有关, 而与  $t_0, t$  无关。即

$$F_X(x; t|x_0; t_0) = F_X(x|x_0; \tau) \quad (7.91)$$

或

$$F_X(x; t|x_0; t_0) = F_X(x|x_0; \tau). \quad (7.92)$$

具有这种特性的马氏过程称为齐次马尔可夫过程。



## 2. 马尔可夫过程的统计特性及性质

由前面的内容可知, 随机过程的统计特性可由有限维联合概率分布来近似地描述。对于马尔可夫过程  $X(t)$  来说, 其  $n$  维概率密度可以表示为

$$\begin{aligned} & F_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) \\ &= F_X(x_n; t_n | x_1, x_2, \dots, x_{n-1}; t_1, t_2, \dots, t_{n-1}) F_X(x_1, x_2, \dots, x_{n-1}; t_1, t_2, \dots, t_{n-1}) \\ &= \dots \\ &= F_X(x_n; t_n | x_{n-1}; t_{n-1}) F_X(x_{n-1}; t_{n-1} | x_{n-2}; t_{n-2}) \cdots F_X(x_2; t_2 | x_1; t_1) F_X(x_1; t_1) \\ &= F_X(x_1; t_1) \prod_{i=1}^{n-1} F_X(x_{i+1}; t_{i+1} | x_i; t_i), \quad t_1 < t_2 < \dots < t_n. \end{aligned} \quad (7.93)$$

当取  $t_1$  为初始时刻时,  $f_X(x_1; t)$  表示初始概率分布(密度)。上式表明: 马氏过程的统计特性完全由它的初始概率分布(密度)和转移概率分布(密度)所确定。

上面已经介绍了马氏过程的定义及一些特征, 下面给出马氏过程的几个有用性质。

1° 同马尔可夫序列的情况一样, 逆方向的马尔可夫过程仍为马尔可夫过程。

对任意的整数  $n$  和  $k$ , 有

$$\begin{aligned} f_X(x_n; t_n | x_{n+1}, x_{n+2}, \dots, x_{n+k}, t_{n+1}; t_{n+2}, \dots, t_{n+k}) \\ = f_X(x_n; t_n | x_{n+1}; t_{n+1}). \end{aligned} \quad (7.94)$$

证:

$$\begin{aligned} & f_X(x_n; t_n | x_{n+1}, x_{n+2}, \dots, x_{n+k}; t_{n+1}, t_{n+2}, \dots, t_{n+k}) \\ &= \frac{f_X(x_n, x_{n+1}, \dots, x_{n+k}; t_n, t_{n+1}, \dots, t_{n+k})}{f_X(x_{n+1}, \dots, x_{n+k}; t_{n+1}, \dots, t_{n+k})} \\ &= \frac{f_X(x_n; t_n) \prod_{i=n}^{n+k-1} f_X(x_{i+1}; t_{i+1} | x_i; t_i)}{f_X(x_{n+1}; t_{n+1}) \prod_{i=n+1}^{n+k-1} f_X(x_{n+1}; t_{n+1})} \\ &= \frac{f_X(x_{n+1}, x_n; t_{n+1}, t_n)}{f_X(x_{n+1}; t_{n+1})} \\ &= f_X(x_n; t_n | x_{n+1}; t_{n+1}). \end{aligned} \quad (7.95)$$

2° 若马尔可夫过程的现在状态已知, 则将来状态与过去状态无关。

若  $t_n > t_r > t_s$ , 则在已知  $X$ (过程在  $t_n$  时刻的状态)的条件下, 随机变量  $X_n$  和  $X_s$  是独立的, 满足

$$f_X(x_n, x_s; t_n, t_s | x_r, t_r) = f_X(x_n; t_n | x_r; t_r) F_X(x_s; t_s | x_r; t_r). \quad (7.96)$$

3° 若对每个  $t \leq t_1 < t_2$ ,  $X(t_2) - X(t_1)$  与  $X(t)$  皆是独立的, 则过程  $X(t)$  是马氏过程。

4° 由转移概率密度的无后效性, 可推出

$$E[X(t_n) | X(t_{n-1}), \dots, X(t_1)] = E[X(t_n) | X(t_{n-1})]. \quad (7.97)$$

## 7.2 独立增量过程

### 7.2.1 概述

#### 定义 7.22 独立增量过程

设有一个随机过程  $X(t), t \in T$ , 如果对任意的时刻  $0 \leq t_0 < t_1 < t_2 < \dots < t_n < b$  过程的增量  $X(t_1) - X(t_0), X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1})$  是相互独立的随机变量, 则称  $X(t)$  为独立增量过程, 又称为可加过程。



若由独立增量过程  $X(t), t \in T$ , 构造一个新过程  $Y(t) = X(t) - X(t_0), t \in T$ , 则新过程  $Y(t)$  也是一个独立增量过程, 不仅与  $X(t)$  有相同的增量规律, 而且有  $P\{Y(t_0) = 0\} = 1$ 。所以, 对一般的独立增量过程  $X(t)$ , 均假设(规定)其初始概率分布为  $P\{X(t_0) = 0\} = 1$ 。

由定义可见, 独立增量过程有这样的特点: 在任一时间间隔上, 过程状态的改变并不影响将来任一时间间隔上过程状态的改变(称为无后效性)。从而决定了独立增量过程是一种特殊的马尔可夫过程。因此, 同马尔可夫过程一样, 独立增量过程的有限维分布可由它的初始概率分布  $P\{X(t_0) < x_0\}$  及一切增量的概率分布唯一确定, 这里  $t$  为过程的初始时刻。

#### 2. 性质

1° 独立增量过程  $X(t)$  是一种特殊的马尔可夫过程。

证: 设增量为  $Y(t_i) = X(t_i) - X(t_{i-1}), i = 1, 2, \dots, n_0$ 。由于  $X(t)$  为独立增量过程, 故增量  $Y(t_1) = X(t_1) - X(t_0), Y(t_2) = X(t_2) - X(t_1), \dots, Y(t_n) = X(t_n) - X(t_{n-1})$  为相互独立的随机变量。因此有

$$f_Y(y_1, y_2, \dots, y_n; t_1, t_2, \dots, t_n) f_1(y_1; t_1) f_2(y_2; t_2) \cdots f_n(y_n; t_n). \quad (7.98)$$

由  $X(t_0) = C$ , 并利用多维随机变量的函数变换

$$\begin{aligned} F_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) &= f_Y(y_1, y_2, \dots, y_n; t_1, t_2, \dots, t_n) \\ &= f_1(y_1; t_1) f_2(y_2; t_2) \cdots f_n(y_n; t_n) \\ &= f_1(x_1; t_1) f_2(x_2 - x_1; t_2, t_1) \cdots f_n(x_n - x_{n-1}; t_n, t_{n-1}). \end{aligned} \quad (7.99)$$

可得

$$\begin{aligned}
 F_X(x_n; t_n | x_{n-1}, \dots, x_1; t_{n-1}, \dots, t_1) \\
 &= \frac{F_X(x_1, \dots, x_{n-1}, x_n; t_1, \dots, t_{n-1}, t_n)}{F_X(x_1, \dots, x_{n-1}; t_1, \dots, t_{n-1})} \\
 &= \frac{f_1(x_1; t_1) \cdots f_{n-1}(x_{n-1} - x_{n-2}; t_{n-1}, t_{n-2}) f_n(x_n - x_{n-1}; t_n, t_{n-1})}{f_1(x_1; t_1) \cdots f_{n-1}(x_{n-1} - x_{n-2}; t_{n-1}, t_{n-2})} \quad (7.100) \\
 &= f_n(x_n - x_{n-1}; t_n, t_{n-1}) \\
 &= F_X(x_n; t_n | x_{n-1}; t_{n-1}).
 \end{aligned}$$

可见, 在  $x_1$  已知条件下,  $x_n$  与  $x_{n-2}, \dots, x_2, x_1$  无关, 因此  $X(t)$  是马尔可夫过程。

2° 独立增量过程的有限维分布由它的初始概率分布和所有增量的概率分布唯一确定。

证: 设  $Y(t_0) = X(t_0), Y(t_i) = X(t_i) - X(t_{i-1}), i = 1, 2, \dots, n$ , 增量  $Y(t_i)$  的概率分布函数可写成  $F_i(y_i, t_i)$ 。由

$$\left\{
 \begin{array}{l}
 X(t_0) = Y(t_0) \\
 X(t_1) = X(t_1) - X(t_0) + X(t_0) = Y(t_1) + Y(t_0) \\
 X(t_2) = X(t_2) - X(t_1) + X(t_1) - X(t_0) + X(t_0) = Y(t_2) + Y(t_1) + Y(t_0) \\
 \vdots \\
 X(t_n) = Y(t_n) + Y(t_{n-1}) + \cdots + Y(t_1) + Y(t_0) = \sum_{i=0}^n Y(t_i).
 \end{array}
 \right. \quad (7.101)$$

则独立增量过程  $X(t)$  的  $n+1$  维概率分布为

$$\begin{aligned}
 F_X(x_0, x_1, x_2, \dots, x_n; t_0, t_1, t_2, \dots, t_n) \\
 &= P\{X(t_0) \leq x_0, X(t_1) \leq x_1, X(t_2) \leq x_2, \dots, X(t_n) \leq x_n\} \\
 &= P\left\{Y(t_0) \leq x_0, Y(t_1) + Y(t_0) \leq x_1, Y(t_2) + Y(t_1) + Y(t_0) \leq x_2, \dots, \sum_{i=0}^n Y(t_i) \leq x_n\right\}. \quad (7.102)
 \end{aligned}$$

利用条件概率表示  $n$  维分布的方法及马氏过程的无后效性, 有

$$\begin{aligned}
 F_X(x_0, x_1, x_2, \dots, x_n; t_0, t_1, t_2, \dots, t_n) \\
 &= P\{Y(t_0) \leq x_0\} P\{Y(t_1) + Y(t_0) \leq x_1 | Y(t_0) = y_0\} \\
 &\quad P\{Y(t_2) + Y(t_1) + Y(t_0) \leq x_2 | Y(t_0) + Y(t_1) = y_0 + y_1\} \cdots \\
 &\quad P\{\sum_{i=0}^n Y(t_i) \leq x_n | \sum_{i=0}^{n-1} Y(t_i) = \sum_{i=0}^{n-1} y_i\} = P\{Y(t_0) \leq x_0\} P\{Y(t_1) \leq x_1 - y_0\} \\
 &\quad P\{Y(t_2) \leq x_2 - (y_0 + y_1)\} \cdots P\{Y(t_n) \leq x_n - \sum_{i=0}^{n-1} y_i\} \\
 &= F_X(x_0; t_0) F_1(x_1 - y_0; t_1) F_2(x_2 - (y_0 + y_1); t_2) \cdots F_n\left(x_n - \sum_{i=0}^{n-1} y_i; t_n\right). \quad (7.103)
 \end{aligned}$$

因为

$$x_0 = y_0 = 0, \quad y_1 = x_1, \quad y_1 + y_2 = x_2, \quad \dots, \quad \sum_{i=0}^{n-1} y_i = x_{n-1}. \quad (7.104)$$

且当  $X(t_0) = 0$  时,  $F_X(x_0; t_0) = P\{X(t_0) = 0\} = 1$ , 则

$$\begin{aligned} F_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) &= F_X(x_1; t_1) F_2(x_2 - x_1; t_2) \cdots F_n(x_n - x_{n-1}; t_n) \\ &= F_X(x_1; t_1) \prod_{k=2}^n F_k(x_k - x_{k-1}; t_k). \end{aligned} \quad (7.105)$$

上式说明, 用一维增量概率分布  $F_k(x_k - x_{k-1}; t_k)$  ( $k = 2, \dots, n$ ) 与  $X(t)$  的初始分布  $F_X(x_1; t_1)$  就可以充分描述一个独立增量过程的  $n$  维分布。如果独立增量过程  $X(t)$  的增量  $X(t) - X(t-1)$  的分布只与时间差  $t_i, t_{i-1}$  有关, 而与  $t_i, t_{i-1}$  本身无关, 则称  $X(t)$  为齐次独立增量过程或平稳独立增量过程。

## 7.2.2 泊松过程

泊松过程和维纳过程是两个最重要的独立增量过程。在日常生活及工程技术领域中, 常常需要研究这样一类问题, 即研究在一定时间间隔  $[0, t)$  内某随机事件出现次数的统计规律。例如: 公用事业中, 某个固定的时间间隔  $[0, t)$  内, 光顾某商店的顾客数; 通过某交叉路口的电车、汽车数; 某船舶甲板“上浪”的次数; 某电话总机接到的呼唤次数; 电子技术中散粒噪声的冲激脉冲个数; 数字通讯中已编码信号的误码个数等。所有这些问题一般被称为计数过程。

### 1. 计数过程

#### 定义 7.23 计数过程

某事件  $A$  在  $[t_0, t)$  内出现的总次数所组成的过程  $\{X(t), t \geq t_0 \geq 0\}$  称为计数过程。

从定义出发, 任何一个计数过程  $X(t)$  应满足下列条件:

- 1°  $X(t)$  是一个正整数。
- 2° 如果有两个时刻  $t_1, t_2$ , 且  $t_2 > t_1$ , 则  $X(t_2) \geq X(t_1)$ 。
- 3° 当  $t_2 > t_1$  时,  $X(t_2) - X(t_1)$  代表在时间间隔  $t_2 - t_1$  内事件  $A$  出现的次数。

在计数过程中, 如果在不相交叠的时间间隔内事件  $A$  出现的次数是相互独立的, 则该计数过程为独立增量过程。即当  $t_1 < t_2 \leq t_3 < t_4$  时,  $[t_1, t_2]$  和  $[t_3, t_4]$  为两个不相交叠的时间间隔,  $[t_1, t_2]$  内事件  $A$  出现的次数为  $X(t_2) - X(t_1)$ ,  $[t_3, t_4]$  内事件  $A$  出现的次数为  $X(t_4) - X(t_3)$ , 若  $X(t_2) - X(t_1)$  与  $X(t_4) - X(t_3)$  相互独立, 则  $X(t)$  为独立增量过程。

计数过程中, 如果在  $[t_1, t_1 + \tau)$  内事件  $A$  出现的次数仅与时间差  $\tau$  有关, 而与起始时间  $t_1$  无关, 也即  $[X(t_1 + \tau) - X(t_1)]$  仅与  $\tau$  有关而与  $t_1$  无关, 则称该过程为齐次或平稳增量计数过程。

## 2. 泊松过程的一般概念

### 定义 7.24 泊松过程

若有一随机计数过程  $\{X(t), t \geq t_0 \geq 0\}$ , 满足下列假设:

- 1° 从  $t_0$  开始观察事件, 即  $X(t_0) = 0$ 。
- 2° 对任意时刻  $0 \leq t_1 < t_2 < \dots < t_n$ , 出现事件次数  $X(t_{i-1}, t_i) = X(t_i) - X(t_{i-1})$  ( $i = 1, 2, \dots, n$ ) 是相互独立的, 且出现次数  $X(t_{i-1}, t_i)$  仅与时间差  $\tau_i = t_i - t_{i-1}$  有关, 而与起始时间  $t_1$  无关。
- 3° 对于充分小的  $\Delta t$ , 在  $[t, t + \Delta t)$  内出现事件一次的概率为

$$P_1(t, t + \Delta t) = P\{X(t, t + \Delta t) = 1\} = \lambda \Delta t + o(\Delta t). \quad (7.106)$$

其中  $o(\Delta)$  是在  $\Delta t \rightarrow 0$  时关于  $\Delta$  的高阶无穷小量, 常数  $\lambda > 0$  称为过程  $X(t)$  的强度。

- 4° 对于充分小的  $\Delta t$ , 在  $[t, t + \Delta t)$  内出现事件两次及两次以上的概率为

$$\sum_{j=2}^{\infty} P_j(t, t + \Delta t) = \sum_{j=2}^{\infty} P\{X(t, t + \Delta t) = j\} = o(\Delta t). \quad (7.107)$$

此概率与出现一次的概率相比, 可以忽略不计。若将上述两式结合起来, 可得到在  $[t, t + \Delta t)$  内不出现事件(或出现事件零次)的概率为

$$\begin{aligned} P\{X(t, t + \Delta t) = 0\} &= P_0(t, t + \Delta t) = 1 - \left[ P_1(t, t + \Delta t) + \sum_{j=2}^{\infty} P_j(t, t + \Delta t) \right] \\ &= 1 - \lambda \Delta t - o(\Delta t). \end{aligned} \quad (7.108)$$

则称此过程为泊松过程。泊松过程是计数过程, 也是重要的独立增量过程。

泊松过程在任意两时刻  $t_1 < t_2$  所得随机变量的增量  $X(t_1, t_2) = X(t_2) - X(t_1)$  服从期望为  $\lambda(t_2 - t_1)$  的泊松分布, 即对于  $k = 0, 1, 2, \dots$ , 有

$$P_k(t_1, t_2) = P\{X(t_1, t_2) = k\} = \frac{[\lambda(t_2 - t_1)]^k}{k!} e^{-\lambda(t_2 - t_1)}, \quad (7.109)$$

则该过程在  $[t_0, t)$  内出现事件  $k$  次的概率为

$$P_k(t_0, t) = P\{X(t_0, t) = k\} = \frac{[\lambda(t - t_0)]^k}{k!} e^{-\lambda(t-t_0)}, t > t_0, k = 0, 1, 2, \dots \quad (7.110)$$

证: ①首先确定  $P_0(t_0, t)$ , 对于充分小的  $\Delta t > 0$ , 由于

$$\begin{aligned} X(t_0, t + \Delta t) &= X(t + \Delta t) - X(t_0) \\ &= X(t + \Delta t) - X(t) + X(t) - X(t_0) \\ &= X(t, t + \Delta t) + X(t_0, t). \end{aligned} \quad (7.111)$$

故

$$\begin{aligned} P_0(t_0, t + \Delta t) &= P\{X(t_0, t + \Delta t) = 0\} \\ &= P\{[X(t_0, t) + X(t, t + \Delta t)] = 0\} \\ &= P\{X(t_0, t) = 0, X(t, t + \Delta t) = 0\}. \end{aligned} \quad (7.112)$$

由泊松过程定义可知, 满足条件 2, 则

$$\begin{aligned} P_0(t_0, t + \Delta t) &= P\{X(t_0, t) = 0\} P\{X(t, t + \Delta t) = 0\} \\ &= P_0(t_0, t) P_0(t, t + \Delta t) \\ &= P_0(t_0, t) [1 - \lambda \Delta t - O(\Delta t)]. \end{aligned} \quad (7.113)$$

即  $P_0(t_0, t + \Delta t) - P_0(t_0, t) = P_0(t_0, t) [-\lambda \Delta t - O(\Delta t)]$ 。两边除以  $\Delta t$ , 并令  $\Delta t \rightarrow 0$ , 便可得到  $P_0(t_0, t)$  满足的微分方程

$$\frac{dP_0(t_0, t)}{dt} = -\lambda P_0(t_0, t). \quad (7.114)$$

因为  $P_0(t_0, t_0) = P\{X(t_0, t_0) = 0\} = 1$ , 将它看作初始条件, 即可由上式解得

$$P_0(t_0, t) = e^{-\lambda(t-t_0)}, \quad t > t_0. \quad (7.115)$$

②类似地, 可以确定  $P_1(t_0, t)$ 。先考虑

$$\begin{aligned} P_1(t_0, t + \Delta t) &= P\{X(t_0, t + \Delta t) = 1\} \\ &= P\{X(t_0, t) + X(t, t + \Delta t) = 1\} \\ &= P\{X(t_0, t) = 1, X(t, t + \Delta t) = 0\} + P\{X(t_0, t) = 0, X(t, t + \Delta t) = 1\} \\ &= P_1(t_0, t) P_0(t, t + \Delta t) + P_0(t_0, t) P_1(t, t + \Delta t). \end{aligned} \quad (7.116)$$

再将  $P_0(t_0, t)$  代入上式, 经适当整理后, 两边除以  $\Delta t$ , 并令  $\Delta t \rightarrow 0$ , 即可得  $P_1(t_0, t)$  的微分方程

$$\frac{dP_1(t_0, t)}{dt} = -\lambda P_1(t_0, t) + \lambda e^{-\lambda(t-t_0)} \quad (7.117)$$

因为  $P_1(t_0, t_0) = P\{X(t_0, t_0) = 1\} = 0$ , 将它作为初始条件, 可求得上式解为

$$P_1(t_0, t) = \lambda(t - t_0) e^{-\lambda(t-t_0)}, \quad t > t_0 \quad (7.118)$$

③ 重复上述方法, 可求得在  $[t_0, t]$  内事件出现  $k$  次的概率

$$P_k(t_0, t) = P\{X(t_0, t) = k\} = \frac{[\lambda(t - t_0)]^k}{k!} e^{-\lambda(t-t_0)}, \quad t > t_0, k = 0, 1, 2, \dots \quad (7.119)$$

当取  $t_0 = 0$  时, 有

$$P_k(0, t) = P\{X(t) = k\} = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad t > 0, k = 0, 1, 2, \dots \quad (7.120)$$

此式表明: 对于固定的  $t$ , 与泊松过程相应的随机变量  $X(t)$  服从参数为  $\lambda t$  的泊松分布。而  $A$  就是在  $[0, t]$  内出现事件次数的数学期望。换言之, 强度  $\lambda$  就是单位时间内出现事件次数的数学期望。

由于泊松过程是一个计数过程, 泊松过程  $X(t)$  的每一个样本函数  $x(t)$  都呈阶梯形 (图 7-3), 它在每个随机点处产生单位为“1”的阶跃。



图 7-2 质点  $M$  的随机游动

对于给定的  $t$ ,  $X(t)$  等于在时间间隔  $[0, t]$  内的随机点数。所以泊松过程

$$X(t) = \sum_i U(t - t_i), \quad (7.121)$$

其中  $t_i$  是随机变量。

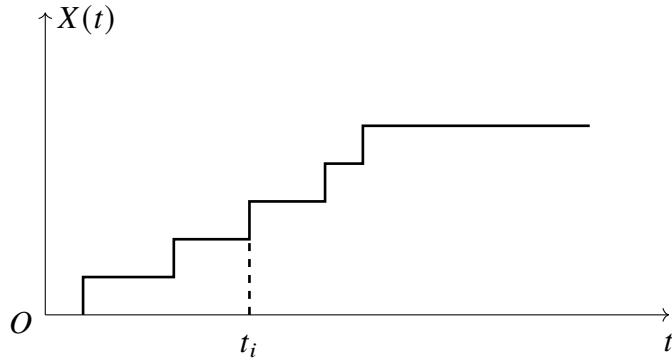


图 7-3 泊松过程的样本函数示意图

### 3. 泊松过程的统计特性

#### (1) 数学期望

若将时间  $t$  固定, 则随机过程  $X(t)$  就是一个泊松分布的随机变量, 因此

$$\begin{aligned} E[X(t)] &= \sum_{k=0}^{\infty} k P\{X(t) = k\} = \sum_{k=0}^{\infty} k \frac{(\lambda t)^k}{k!} e^{-\lambda t} \\ &= \lambda t e^{-\lambda t} \left[ \sum_{k=0}^{\infty} \frac{(\lambda t)^{k-1}}{(k-1)!} \right] = \lambda t e^{-\lambda t} [e^{\lambda t}] = \lambda t. \end{aligned} \quad (7.122)$$

同理引用式 (7.109), 随机过程的增量  $X(t_2) - X(t_1)$  的期望为

$$E[X(t_2) - X(t_1)] = \sum_{k=0}^{\infty} k P\{X(t_1, t_2) = k\} = \lambda(t_2 - t_1). \quad (7.123)$$

#### (2) 均方值与方差

类似于上述方法, 过程  $X(t)$  的均方值

$$\begin{aligned} E[X^2(t)] &= \sum_{k=0}^{\infty} k^2 P\{X^2(t) = k^2\} = \sum_{k=0}^{\infty} k^2 \frac{(\lambda t)^k}{k!} e^{-\lambda t} \\ &= e^{-\lambda t} \left[ \sum_{k=0}^{\infty} k(k-1) \frac{(\lambda t)^k}{k!} + \sum_{k=0}^{\infty} k \frac{(\lambda t)^k}{k!} \right] \\ &= e^{-\lambda t} \left[ (\lambda t)^2 \sum_{k=2}^{\infty} \frac{(\lambda t)^{k-2}}{(k-2)!} + \lambda \sum_{i=1}^{\infty} \frac{(\lambda t)^{i-1}}{(i-1)!} \right] \\ &= e^{-\lambda t} [(\lambda t)^2 e^{\lambda t} + \lambda t e^{\lambda t}] \\ &= \lambda^2 t^2 + \lambda t. \end{aligned} \quad (7.124)$$

同理, 过程  $X(t)$  的方差、过程增量  $X(t_2) - X(t_1)$  的均方值和方差为

$$D[x(t)] = E[X(t)^2] - E^2[X(t)] = \lambda^2 t^2 + \lambda t - (\lambda t)^2 = \lambda t \quad (7.125)$$

$$E\{[X(t_2) - X(t_1)]^2\} = \lambda^2(t_2 - t_1)^2 + \lambda(t_2 - t_1) \quad (7.126)$$

$$D[X(t_2) - X(t_1)] = \lambda(t_2 - t_1). \quad (7.127)$$

### 3) 自相关函数

由定义  $R_X(t_1, t_2) = E[X(t_1)X(t_2)]$ , 可知①和②成立。

① 若  $t_2 > t_1 > 0$ , 如图 7-4 所示, 由于时间间隔  $t_1$  和  $t_2$  相互重叠, 则增量  $X(t_1)$  和  $X(t_2)$  相互不独立, 但时间间隔  $t_2 - t_1$  与  $t_1$  不重叠。因此将增量  $X(t_2)$  变换为两个独立的增量之和。

$$X(t_2) = [X(t_2) - X(t_1)] + X(t_1). \quad (7.128)$$

因此有

$$\begin{aligned} R_X(t_1, t_2) &= E\{X(t_1)[X(t_2) - X(t_1) + X(t_1)]\} \\ &= E[X(t_1)]E[X(t_2) - X(t_1)] + E[X^2(t_1)] \\ &= \lambda_1 \lambda (t_2 - t_1) + \lambda^2 t_1^2 + \lambda_1 \\ &= \lambda^2 t_1 t_2 + \lambda t_1, \quad t_2 > t_1 > 0 \end{aligned} \quad (7.129)$$

② 若  $t_1 > t_2 > 0$ , 同样有

$$R_X(t_1, t_2) = \lambda^2 t_1 t_2 + \lambda t_2, \quad t_1 > t_2 > 0. \quad (7.130)$$

综合上述两式, 则有

$$R_X(t_1, t_2) = \lambda^2 t_1 t_2 + \lambda \min(t_1, t_2). \quad (7.131)$$

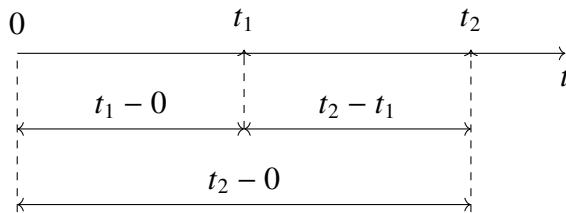


图 7-4 时间关系图

### 7.2.3 泊松过程的应用实例

电报信号在随机点密度  $A$  为常数的均匀情况下, 来研究下述泊松过程的两个应用实例。

#### 例 7.2.1

半随机电报信号  $X(t)$  是只取  $+1$  或  $-1$  的随机过程,  $X(t)$  的一条样本函数曲线如图 7-5 所示。

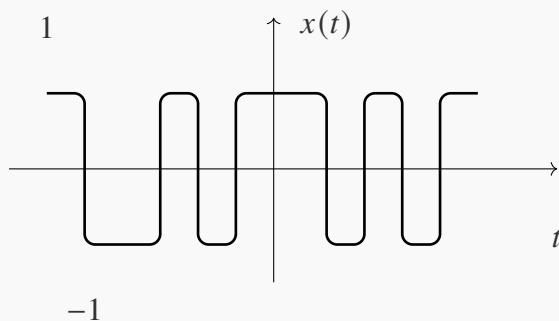


图 7-5 半随机电报信号的样本函数

若在时间间隔  $(0, t)$  内, 信号时刻点的总数为偶数(或 0), 则过程  $X(t) = +1$ ; 若为奇数, 则  $X(t) = -1$ .

#### 1) 半随机电报信号的概率分布

设在时间间隔  $(0, t)$  内有  $k$  个变号点的概率  $P_k(0, t)$  为

$$P_k(0, t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}. \quad (7.132)$$

因事件序列在  $(0, t)$  内出现  $k$  个点 ( $k = 0, 1, 2, \dots$ ) 是互不相容的, 故在  $(0, t)$  内有偶数个变号点的概率为

$$\begin{aligned} P\{X(t) = 1\} &= P_0(0, t) + P_2(0, t) + \dots = e^{-\lambda t} \left[ 1 + \frac{(\lambda t)^2}{2!} + \dots \right] \\ &= e^{-\lambda t} \operatorname{ch}(\lambda t). \end{aligned} \quad (7.133)$$

类似地, 在  $(0, t)$  内有奇数个变号点的概率为

$$\begin{aligned} P\{X(t) = -1\} &= P_1(0, t) + P_3(0, t) + \cdots \\ &= e^{-\lambda} \left[ \lambda t + \frac{(\lambda t)^3}{3!} + \cdots \right] = e^{-\lambda t} \operatorname{sh}(\lambda t). \end{aligned} \quad (7.134)$$

即

$$\begin{cases} P\{X(t) = 1\} &= e^{-\lambda t} \cosh(\lambda t) \\ P\{X(t) = -1\} &= e^{-\lambda t} \sinh(\lambda t) \end{cases}. \quad (7.135)$$

2) 半随机电报信号的均值

$$\begin{aligned} E[X(t)] &= 1 \cdot P\{X(t) = 1\} + (-1) \cdot P\{X(t) = -1\} \\ &= e^{-\lambda} [\operatorname{ch}(\lambda t) - \operatorname{sh}(\lambda t)] = e^{-2\lambda}. \end{aligned} \quad (7.136)$$

其中

$$P\{X(t_1), X(t_2)\} = P\{X(t_2)|X(t_1)\} P\{X(t_1)\}. \quad (7.137)$$

3) 半随机电报信号的自相关函数

$$R_X(t_1, t_2) = \sum_{\substack{X(t_1)=\pm 1 \\ X(t_2)=\pm 1}} X(t_1) X(t_2) P\{X(t_1), X(t_2)\}. \quad (7.138)$$

其中

$$P\{X(t_1), X(t_2)\} = P\{X(t_2)|X(t_1)\} P\{X(t_1)\}. \quad (7.139)$$

① 假设  $t_2 - t_1 = \tau > 0$  且  $X(t_1) = 1$ , 若  $X(t_2) = 1$ , 则在间隔  $(t_2, t_1)$  内有偶数个变号点。故

$$P\{X(t_2) = 1|X(t_1) = 1\} = P\{X(t) = 1\} = e^{-\lambda t} \cosh(\lambda t). \quad (7.140)$$

又由于  $P\{X(t_1) = 1\} = e^{-\lambda t_1} \cosh(\lambda t_1)$ , 就可得到

$$P\{X(t_1) = 1, X(t_2) = 1\} = e^{-\lambda \tau} \cosh(\lambda \tau) e^{-\lambda_1 t} \cosh(\lambda t_1). \quad (7.141)$$

类似地可得

$$\begin{cases} P\{X(t_2) = -1|X(t_1) = -1\} = P\{X(\tau) = 1\} = e^{-\lambda \tau} \cosh(\lambda \tau) \\ P\{X(t_2) = -1|X(t_1) = 1\} = P\{X(\tau) = -1\} = e^{-\lambda \tau} \sinh(\lambda \tau) \\ P\{X(t_2) = 1|X(t_1) = -1\} = P\{X(\tau) = -1\} = e^{-\lambda \tau} \sinh(\lambda \tau) \end{cases} . \quad (7.142)$$

因此有

$$\begin{cases} P\{X(t_1) = -1, X(t_2) = -1\} = e^{-\lambda\tau} \cosh(\lambda\tau) e^{-\lambda t_1} \sinh(\lambda t_1) \\ P\{X(t_1) = 1, X(t_2) = -1\} = e^{-\lambda\tau} \sinh(\lambda\tau) e^{-\lambda t_1} \cosh(\lambda t_1) \\ P\{X(t_1) = -1, X(t_2) = 1\} = e^{-\lambda\tau} \sinh(\lambda\tau) e^{-\lambda t_1} \sinh(\lambda t_1). \end{cases} \quad (7.143)$$

$Y(t)$  的自相关函数为

$$\begin{aligned} R_X(t_1, t_2) &= \sum_{\substack{X(t_1)=\pm 1 \\ X(t_2)=\pm 1}} X(t_1) X(t_2) P\{X(t_1), X(t_2)\} \\ &= 1 \cdot 1 \cdot P\{X(t_1) = 1, X(t_2) = 1\} \\ &\quad + (-1) \cdot (-1) \cdot P\{X(t_1) = -1, X(t_2) = -1\} \\ &\quad + 1 \cdot (-1) \cdot P\{X(t_1) = 1, X(t_2) = -1\} \\ &\quad + (-1) \cdot 1 \cdot P\{X(t_1) = -1, X(t_2) = 1\} = e^{-2\lambda\tau} \end{aligned} \quad (7.144)$$

其中  $t_2 - t_1 = \tau > 0$ .

② 设  $t_1 - t_2 = \tau > 0$ , 同理可推出

$$R_X(t_1, t_2) = e^{-2\lambda\tau}, \quad t_1 - t_2 = \tau > 0. \quad (7.145)$$

所以, 综合①②, 可得  $X(t)$  的自相关函数的最终表达式

$$R_X(t_1, t_2) = e^{-2\lambda|t_1-t_2|} = e^{-2\lambda|\tau|} = R_X(\tau). \quad (7.146)$$

由上式可见, 半随机电报信号的自相关函数仅与时间差  $\tau$  有关, 而与时刻点  $t_1$  本身无关, 如图 7-6 所示。

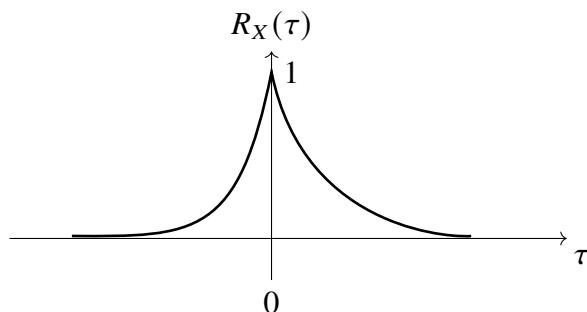


图 7-6 半随机电报信号的自相关函数

#### 4) 半随机电报信号的功率谱密度

对半随机电报信号的自相关函数求傅里叶变换, 即得此电报信号的功率谱密度

$$G_X(\omega) = \int_{-\infty}^{\infty} R_X(\tau) e^{-j\omega\tau} d\tau = \int_{-\infty}^{\infty} e^{-2\lambda|\tau|} e^{-j\omega\tau} d\tau = \frac{4\lambda}{4\lambda^2 + \omega^2}. \quad (7.147)$$

## 2. 随机电报信号

### 例 7.2.2

给定一个随机变量  $A$ , 等概率取  $+1$  或  $-1$ , 即  $P\{A = 1\} = P\{A = -1\} = 0.5$ . 因此,  $E[A] = 0, E[A^2] = 1$ . 假定上述的半随机电报信号  $X(t)$  与随机变量  $A$  统计独立, 即对于每个  $t$ , 随机变量  $X(t)$  与随机变量  $A$  是统计独立的。现在, 定义一个新的随机过程

$$Y(t) = AX(t) \quad (7.148)$$

于是  $Y(t) = X(t)$  或  $Y(t) = -X(t)$ , 为了与  $X(t)$  相区别, 称  $Y(t)$  为随机电报信号。

显然  $Y(t)$  的均值和自相关函数分别为

$$\begin{aligned} E[Y(t)] &= E[AX(t)] = E[A]E[X(t)] = 0 \\ R_Y(t_1, t_2) &= E[Y(t_1)Y(t_2)] = E[A^2X(t_1)X(t_2)] \\ &= E[A^2]E[X(t_1)X(t_2)] = e^{-2\lambda|t_2-t_1|} \\ &= e^{-2\lambda|\tau|} = R_Y(\tau) = R_X(\tau). \end{aligned} \quad (7.149)$$

由上式注意到, 随机过程  $X(t)$  和  $Y(t)$  具有渐进 ( $t \rightarrow \infty$ ) 相等的统计特性.

## 3. 泊松冲激序列

泊松过程  $X(t)$  对时间  $t$  求导, 便可得到与时间轴上的随机点  $t$  相对应的冲激序列  $Z(t)$ , 称为泊松冲激序列。

$$Z(t) = \frac{dX(t)}{dt} = \frac{d}{dt} \sum_i U(t - t_i) = \sum_i \delta(t - t_i), \quad (7.150)$$

其中  $t$  为随机变量。由于泊松过程  $X(t)$  的样本函数是阶梯函数, 则泊松冲激序列的样本函数是一串冲激序列, 如图 7-7 所示。

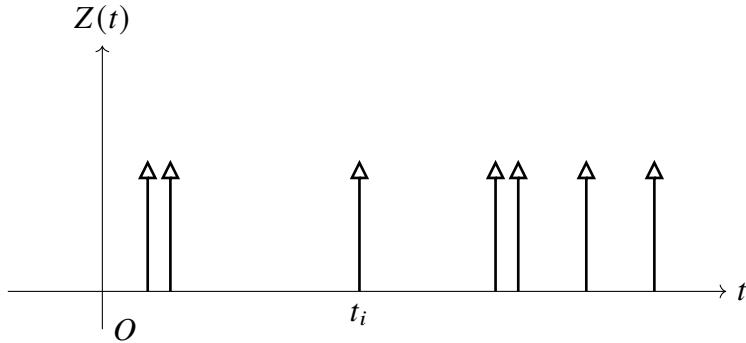


图 7-7 泊松冲激序列的一个样本函数

泊松过程  $X(t)$  及其统计特性均已在前面讨论过, 可得泊松冲激序列  $Z(t)$  的统计特性:

$$\begin{aligned}
 E[Z(t)] &= E\left[\frac{dX(t)}{dt}\right] = \frac{dE[X(t)]}{dt} = \frac{d(\lambda t)}{dt} = \lambda. \\
 R_z(t_1, t_2) &= E[Z(t_1)Z(t_2)] = E\left[\frac{dX(t_1)}{dt_1} \cdot \frac{dX(t_2)}{dt_2}\right] \\
 &= R_{X'X'}(t_1, t_2) = \frac{\partial^2 R_X(t_1, t_2)}{\partial t_1 \partial t_2} = \frac{\partial}{\partial t_1} \left[ \frac{\partial R_X(t_1, t_2)}{\partial t_2} \right] \\
 &= \begin{cases} \frac{\partial}{\partial t_1} (\lambda^2 t_1), & t_1 < t_2 \\ \frac{\partial}{\partial t_1} (\lambda^2 t_1 + \lambda), & t_1 > t_2 \end{cases}. \quad (7.151) \\
 &= \frac{\partial}{\partial t_1} [\lambda^2 t_1 + \lambda U(t_1 - t_2)] \\
 &= \lambda^2 + \lambda \delta(t_1 - t_2) \\
 &= \lambda^2 + \lambda \delta(\tau)
 \end{aligned}$$

由此可见, 泊松冲激序列是平稳过程。

#### 4. 过滤的泊松过程与散粒噪声

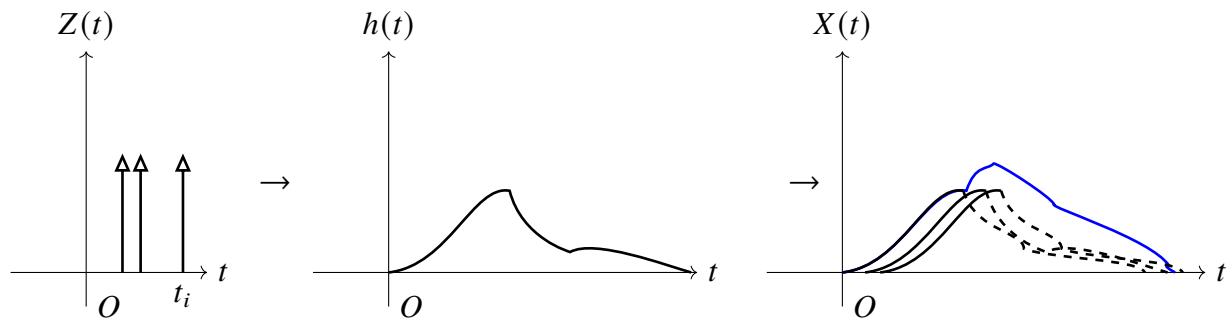


图 7-8 泊松冲激脉冲序列经过线性时不变滤波器

### 定义 7.25 过滤的泊松过程

设有一泊松冲激脉冲序列  $Z(t) = \sum \delta(t - t_i)$  经过线性时不变滤波器, 如图 7-8 所示。则此滤波器输出的随机过程

$$X(t) = Z(t) * h(t) = \sum_{i=1}^{N(T)} h(t - t_i), \quad 0 \leq t < \infty. \quad (7.152)$$

称之为过滤的泊松过程。式中  $h(t)$  为滤波器的冲激响应, 第  $i$  个冲激脉冲出现的时间  $t$  是个随机变量,  $N(T)$  为在  $[0, T]$  内输入到滤波器的冲激脉冲的个数, 它服从泊松分布, 即

$$P\{N(T) = k\} = \frac{(\lambda T)^k}{k!} e^{-\lambda T}, \quad k = 0, 1, 2, \dots \quad (7.153)$$

式中  $\lambda$  为单位时间内的平均脉冲数。

分析可知, 若在  $[0, T]$  内输入到滤波器的冲激脉冲数  $N(T)$  为  $k$ , 则该  $k$  个冲激脉冲出现的时间  $t$  均为独立同分布的随机变量, 且此随机变量均匀分布在  $[0, T]$  内, 即

$$f(t_i | N(T) = k) = \begin{cases} \frac{1}{T}, & 0 \leq t_i < T \\ 0, & \text{其他} \end{cases} \quad (7.154)$$

### 例 7.2.3

温度限制的电子一极管中, 由散粒(或散弹)效应引起的散粒(或散弹)噪声电流是过滤的泊松过程  $n$ 。晶体管中有三种类型的噪声:

- ① 热噪声。
- ② 散粒噪声。
- ③ 闪烁噪声(又称  $1/f$  噪声, 是一种低频噪声)。其中散粒噪声的机理与电子管的相类似, 也是过滤的泊松过程。



换言之, 散粒噪声  $X(t)$  可以表示成

$$X(t) = Z(t) * h(t) = \sum h(t - t_i), \quad (7.155)$$

即把它看成是泊松冲激(脉冲)序列输入到线性时不变系统的输出。

1. 下面来讨论散粒噪声  $X(t)$  的统计特性

① 对于均匀的情况( $\lambda$  为常数), 可以证明  $X(t)$  是平稳的。

已知泊松冲激脉冲序列  $Z(t) = \sum \delta(t - t_i)$  的数学期望和自相关函数为

$$\begin{cases} E[Z(t)] = \lambda \\ R_z(\tau) = \lambda^2 + \lambda\delta(\tau) \end{cases}. \quad (7.156)$$

从而可得泊松冲激序列的功率谱密度

$$G_Z(\omega) = \int_{-\infty}^{\infty} R_z(\tau) e^{-j\omega\tau} d\tau = 2\pi\lambda^2\delta(\omega) + \lambda. \quad (7.157)$$

根据时频域分析, 可得散粒噪声  $X(t)$  的数学期望为

$$\begin{aligned} E[X(t)] &= E[Z(t) * h(t)] = E \left[ \int_{-\infty}^{\infty} Z(t - \eta) h(\eta) d\eta \right] \\ &= \int_{-\infty}^{\infty} E[Z(t - \eta)] h(\eta) d\eta \\ &= \lambda \int_{-\infty}^{\infty} h(\eta) d\eta = \lambda H(0). \end{aligned} \quad (7.158)$$

散粒噪声  $X(t)$  的功率谱密度为

$$\begin{aligned} G_X(\omega) &= |H(\omega)|^2 G_Z(\omega) = |H(\omega)|^2 [2\pi\lambda^2\delta(\omega) + \lambda] \\ &= 2\pi\lambda^2 H^2(0)\delta(\omega) + \lambda|H(\omega)|^2. \end{aligned} \quad (7.159)$$

从而得到其自相关函数为

$$\begin{aligned}
 R_X(\tau) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} G_X(\omega) e^{j\omega\tau} d\omega \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} [2\pi\lambda^2 H^2(0)\delta(\omega) + \lambda|H(\omega)|^2] e^{j\omega\tau} d\omega \\
 &= \lambda^2 H^2(0) + \frac{\lambda}{2\pi} \int_{-\infty}^{\infty} |H(\omega)|^2 e^{j\omega\tau} d\omega \\
 &= \lambda^2 H^2(0) + \lambda \int_{-\infty}^{\infty} h(\tau + \beta)h(\beta) d\beta.
 \end{aligned} \tag{7.160}$$

由上式可见,  $X(t)$  确实是平稳随机过程。均匀的泊松冲激序列  $Z(t)$  和散粒噪声  $N(t)$  的自相关函数及功率谱密度如图 7-9 所示。

则散粒噪声  $X(t)$  的自协方差函数和方差为

$$C_X(\tau) = R_X(\tau) - \{E[X(t)]\}^2 = \lambda \int_{-\infty}^{\infty} h(\tau + \beta)h(\beta) d\beta \tag{7.161}$$

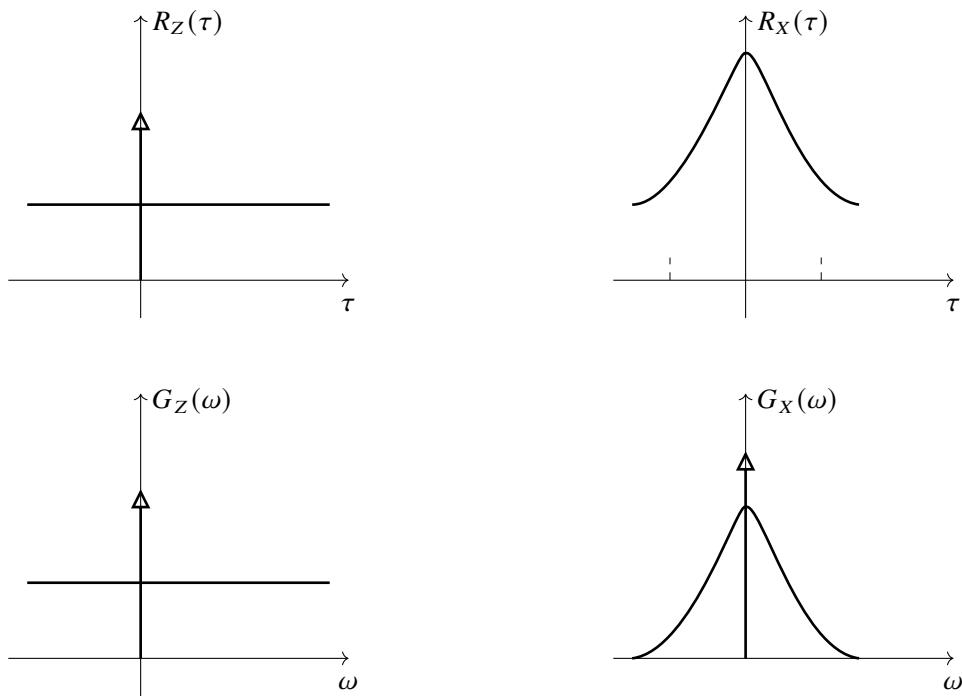
$$\begin{aligned}
 \sigma_X^2 &= C_X(0) = \lambda \int_{-\infty}^{\infty} h^2(\beta) d\beta \\
 &= \lambda \int_{-\infty}^{\infty} h^2(t) dt.
 \end{aligned} \tag{7.162}$$

② 对于非均匀的情况, 即  $\lambda(t)$  不是数  $m_X$  的均值与自协方差函数分别为

$$\begin{aligned}
 E[X(t)] &= E[Z(t) * h(t)] = \int_{-\infty}^{\infty} E[Z(\eta)]h(t - \eta) d\eta \\
 &= \int_{-\infty}^{\infty} \lambda(\eta)h(t - \eta) d\eta.
 \end{aligned} \tag{7.163}$$

其中  $E[Z(t)] = \lambda(t)$ .

$$\begin{aligned}
 C_x(t_1, t_2) &= R_x(t_1, t_2) - E[X(t_1)]E[X(t_2)] \\
 &= \int_{-\infty}^{\infty} \lambda(\eta)h(t_1 - \eta)h(t_2 - \eta) d\eta.
 \end{aligned} \tag{7.164}$$

图 7-9 泊松冲激序列  $Z(t)$  和散粒噪声  $X(t)$  的自相关函数及功率谱密度

③ 如果每个输入冲激脉冲的强度 (面积) 不等于 1, 而是  $q$  (例如电子电荷), 则均匀散粒噪声变为

$$X(t) = \sum_i q h(t - t_i), \quad (7.165)$$

其均值与方差分别为

$$\begin{cases} E[X(t)] = \lambda q \int_{-\infty}^{\infty} h(t) dt = \lambda q H(0) \\ \sigma_X^2 = \lambda q^2 \int_{-\infty}^{\infty} h^2(t) dt \end{cases}. \quad (7.166)$$

反之, 若  $h(t)$  已知, 则由测量  $X(t)$  的均值与方差, 就能求出  $\lambda$  和  $q$ 。

**例 7.26** 图 7.9 所示电路中, 电流源  $i(t)$  由冲激脉冲序列所组成, 即  $i(t) = \sum_i q \delta(t - t_i)$ 。电路两端产生的电压为

$$\begin{aligned} V(t) &= \sum_i q h(t - t_i) \\ \text{且 } h(t) &= \frac{1}{C} e^{-\frac{t}{R}} U(t). \end{aligned} \quad (7.167)$$

式中  $U(t)$  为时域阶跃函数。试求电压  $V(t)$  的均值与方差, 以及  $\lambda$  和  $q$ 。

解：由已知的冲激响应  $h(t)$ , 可方便地求得

$$\begin{aligned} \int_{-\infty}^{\infty} h(t) dt &= \int_0^{\infty} \frac{1}{C} e^{-\frac{t}{RC}} dt = R \\ \int_{-\infty}^{\infty} h^2(t) dt &= \int_0^{\infty} \left( \frac{1}{C} e^{-\frac{t}{RC}} \right)^2 dt = \frac{R}{2C}. \end{aligned} \quad (7.168)$$

可求得  $V(t)$  的均值与方差分别为

$$\begin{cases} m_V = E[V(t)] = \lambda q R \\ \sigma_V^2 = \lambda q^2 \frac{R}{2C} \end{cases}. \quad (7.169)$$

联立后解, 可推导出  $\lambda$  和  $q$  为

$$\lambda = \frac{m_V^2}{2RC\sigma_V^2}, \quad q = \frac{2C\sigma_V^2}{m_V}. \quad (7.170)$$

此例是一个关于限温二极管阳极电压的数学模型, 其中  $i(t)$  是电子电流, 而电子的渡越时间被忽略了。

#### 7.2.4 维纳过程

随机游动的极限形式来研究, 游动过程中的所有轨迹几乎都是连续的。电阻中电子的热运动就具有维纳过程的性质, 可用维纳过程来描述。实际上我们常把白噪声作为热噪声的理想化模型, 而维纳过程可看作是白噪声通过积分器的输出。此外, 维纳过程是一个非平稳的高斯过程。

##### 定义 7.27 维纳过程的定义

若独立增量过程  $X(t)$ , 其增量的概率分布服从高斯分布, 即

$$P\{X(t_2) - X(t_1) < \lambda\} = \frac{1}{\sqrt{2\pi\alpha(t_2-t_1)}} \int_{-\infty}^{\lambda} \exp\left(\frac{-u^2}{2\alpha(t_2-t_1)}\right) du, \quad 0 < t_1 < t_2 \quad (7.171)$$

则称  $X(t)$  为维纳过程。

可以证明, 维纳过程是几乎处处连续的, 但是在任一固定时刻  $t$  上以概率 1 不可微分。

##### 1. 定义

##### 定义 7.28 维纳过程

对于所有样本函数几乎处处连续的齐次独立增量过程 (或齐次独立增量过程  $X(t, \zeta)$ , 几乎对所有在时间轴上连续), 称为维纳过程。

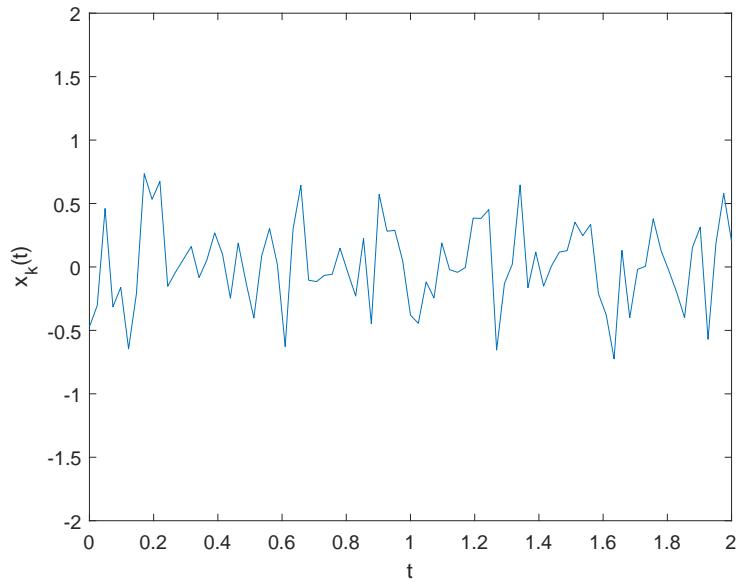


图 7-10 维纳过程的一个样本函数

按照定义给出的条件可以证明: 过程的增量是服从高斯分布的。这里仅对此作简要说明。令  $\Delta = (t_2 - t_1) / n, t_2 > t_1$ 。由于

$$\begin{aligned} X(t_2) - X(t_1) &= [X(t_2) - X(t_2 - \Delta)] + [X(t_2 - \Delta) - X(t_2 - 2\Delta)] \\ &\quad + \cdots + [X(t_2 - (n-1)\Delta) - X(t_1)] \\ &= \sum_{i=1}^n Y_i. \end{aligned} \tag{7.172}$$

由上述条件, 当  $n \rightarrow \infty$  亦即  $\Delta \rightarrow 0$  时,

$$Y_i = X[s - (i-1)\Delta] - X(s - i\Delta) \rightarrow 0 \tag{7.173}$$

故由中心极限定理可得  $X(t_2) - X(t_1)$  趋于高斯分布。

可见, 两种定义是完全一致的。图 7-10 给出了维纳过程一个样本函数的示意图。

## 2. 维纳过程的统计特性

### 1) 维纳过程的数学期望和相关函数

由式 (7.171), 显见

$$E[X(t)] = 0 \tag{7.174}$$

又因为是独立增量过程, 所以有  $X(t_0) = 0$ , 则

$$P\{X(t_1) - X(t_0) < \lambda\} = P\{X(t_1) < \lambda\} = \frac{1}{\sqrt{2\pi\alpha t_1}} \int_{-\infty}^{\lambda} \exp\left(\frac{-u^2}{2\alpha t_1}\right) du. \tag{7.175}$$

故有

$$D[X(t_1)] = E[X^2(t_1)] = \alpha t_1. \quad (7.176)$$

- ① 当  $t_1 = t_2 = t$  时, 有  $R_X(t_1, t_2) = E[X(t_1)X(t_2)] = E[X^2(t)] = \alpha t$ .
- ② 当  $t_1 > t_2$ , 并将  $X(t_1)$  写成  $X(t_1) = X(t_2) + X(t_1) - X(t_2)$ , 则

$$\begin{aligned} R_X(t_1, t_2) &= E[X(t_1)X(t_2)] \\ &= E[(X(t_2) + X(t_1) - X(t_2))X(t_2)] \\ &= E[X^2(t_2)] + E[(X(t_1) - X(t_2))(X(t_2) - X(t_0))] \\ &= E[X^2(t_2)] = \alpha t_2. \end{aligned} \quad (7.177)$$

- ③ 同理, 当  $t_1$  可得

$$R_X(t_1, t_2) = E[X(t_1)X(t_2)] = \alpha t_1. \quad (7.178)$$

综合 ①-③, 可得维纳过程的自相关函数为

$$R_X(t_1, t_2) = E[X(t_1)X(t_2)] = \alpha \min(t_1, t_2). \quad (7.179)$$

### (2) 维纳过程与高斯白噪声

虽然维纳过程在几乎处处 (almost everywhere) 意义下是连续的, 但由上述自相关函数的表达式可知,  $R_X(t_1, t_2)$  在  $t_1 = t_2 = t$  点间断, 所以对于  $t_1 - t_2$ , 该过程的子  $R_x(t_1, t_2)/m_1 a_2$  不存在, 因此维纳过程几乎处处不可微。这样在通常意义下它的导数是不存在的。不过, 若在形式上研究其导数及性质, 则维纳过程  $X(t)$  的导数也是零均值的高斯过程令  $N(t) = X(t)$  为  $X(t)$  形式上的导数, 则  $N(t)$  的自相关函数为

$$R_N(t_1, t_2) = E[X'(t_1)X'(t_2)] = \alpha \delta(t_1 - t_2). \quad (7.180)$$

可见形式导数  $N(t) = X(t)(t \geq 0)$  是高斯白噪声。于是维纳过程  $X(t)$  可以写成白噪声 (具有零均值、均匀谱的平稳高斯过程) 的积分, 即

$$X(t) = \int_0^t N(\tau) d\tau, \quad (7.181)$$

式中的  $N(t)$  有  $E[N(t)] = 0, G_N(\omega) = \alpha$ 。换言之, 维纳过程可看成高斯白噪声通过积分器的输出。

### (3) 维纳过程的概率分布

由上述的讨论结果可以很容易得到维纳过程  $X(t)$  的一维和  $n$  维概率密度为

$$F_X(x; t) = \frac{1}{\sqrt{2\pi\alpha t}} e^{-\frac{x^2}{2\alpha t}}. \quad (7.182)$$

$$\begin{aligned}
F_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) &= f(x_1; t_1) f(x_2 - x_1; t_2, t_1) \cdots f(x_n - x_{n-1}; t_n, t_{n-1}) \\
&= \prod_{i=1}^n \frac{\exp\left[-\frac{1}{2\alpha} \frac{(x_i - x_{i-1})^2}{(t_i - t_{i-1})}\right]}{\sqrt{2\pi\alpha(t_i - t_{i-1})}}.
\end{aligned} \tag{7.183}$$

最后, 将维纳过程的性质归纳为以下几点

- 1°  $X(t_0) = 0$ , 且  $X(t)$  是实过程。
- 2°  $E[X(t)] = 0$ .
- 3° 维纳过程是独立增量过程。
- 4° 维纳过程满足齐次性。换言之,  $X(t_2) - X(t_1)$  的分布只与  $(t_2 - t_1)$  有关, 而与  $t_1$  或  $t_2$  本身无关。
- 5°  $X(t_2) - X(t_1)$  的方差与  $t_2 - t_1$  成正比。

$$\begin{aligned}
D[X(t_2) - X(t_1)] &= E[(X(t_2) - X(t_1))^2] \\
&= E[X^2(t_2)] + E[X^2(t_1)] - 2E[X(t_2)X(t_1)] \\
&= at_2 + \alpha t_1 - 2\alpha t_1 = \alpha(t_2 - t_1), \quad t_2 > t_1.
\end{aligned} \tag{7.184}$$

6° 维纳过程是非平稳高斯过程

### 3. 扩散方程

可以证明, 维纳过程  $X(t)$  满足下列关系式

$$\begin{cases} \frac{\partial f}{\partial t_2} = \frac{\alpha}{2} \frac{\partial f}{\partial x_2^2}; \\ \frac{\partial f}{\partial t_1} + \frac{\alpha}{2} \frac{\partial^2 f}{\partial x_1^2} = 0. \end{cases} \tag{7.185}$$

它们被称作扩散方程。式中  $f = f(x_2, t_2 | x_1, t_1) = f(x_2 | x(t_1) = x_1) (t_2 > t_1)$  是随机变量  $X(t_2)$  在  $X(t_1) = x_1$  条件下的条件概率密度。

证: 因为  $X(t)$  具有零均值并为高斯分布, 有

$$E[X(t_2) | X(t_1) = x_1] = \alpha x_1 = \frac{E[X(t_2)X(t_1)]}{E[X^2(t_1)]} x_1 = x_1, \quad t_2 > t_1. \tag{7.186}$$

即在  $X(t_1) = x_1$  条件下,  $X(t_2)$  的条件方差等于  $\alpha(t_2 - t_1)$ , 故有

$$f(x_2; t_2 | x_1; t_1) = \frac{1}{\sqrt{2\pi\alpha(t_2 - t_1)}} \exp\left[-\frac{(x_2 - x_1)^2}{2\pi(t_2 - t_1)}\right]. \tag{7.187}$$

对上式作求导运算, 即可得扩散方程实际上, 扩散方程是柯尔莫哥洛夫方程 (前进方程和后退方程) 的特例。柯尔莫哥洛夫方程也称为福克尔-普朗克 (Fokker Planck) 方程。这是

因为这组方程在特殊的场合用不十分严格的方法,首先为福克尔和普朗克所获得;而在一般的场合,并用严格的方法,则为柯尔莫哥洛夫所得到。

设扩散过程  $X(t)$  的条件概率密度  $f = F_X(x_2; t_2 | x_1; t_1)$  ( $t_2 > t_1$ ), 则柯尔莫哥洛夫(前进和后退)方程可表示成

$$\begin{cases} \frac{\partial f}{\partial x_1} + a(x_1, t_1) \frac{\partial f}{\partial x_1} + \frac{b(x_1, t_1)}{2} \frac{\partial^2 f}{\partial x_1^2} = 0 \\ \frac{\partial f}{\partial x_2} + \frac{\partial}{\partial x_2} [a(x_2, t_2) f] - \frac{1}{2} \frac{\partial}{\partial x_2^2} [b(x_2, t_2) f] = 0. \end{cases} \quad (7.188)$$

式中,  $a(x, t)$  为  $t$  时刻自  $x$  出发的质点的瞬时平均速度(或者说是过程  $X(t)$  变化的平均速度)。而  $b(x, t)$  则与质点瞬时平均动能成比例,换言之,  $b(x, t)$  是在很小的  $\Delta t$  内,质点位移的平方平均偏差与  $\Delta t$  的比值。

在实际问题中,如果想由实验所得数据资料来直接确定条件分布函数,是极为困难的。但是,可以根据其物理意义,通过比较容易找到的  $a(x, t)$  及  $b(x, t)$  来解柯尔莫哥洛夫微分方程,从而得到条件分布函数。

当随机过程  $X(t)$  为维纳过程,且  $a(x, t) = 0, b(x, t) = a$  时,把这些条件代入柯尔莫哥洛夫方程,则可得到扩散方程,从而解出条件分布函数。也就是说,维纳过程是一种特殊的扩散过程。

#### 例 7.29 研究液体中微粒的随机扩散运动。

设液体的质量是均匀的,由于微粒的运动是由许多分子碰撞所产生的许多小随机位移的和,因而可认为:自时刻  $s$  至的位移  $X_t - X_s$ ,是许多几乎独立的小位移的和,故由中心极限定理可假设  $X_t - X_s$  服从高斯分布。此外由液体的均匀性,还可设  $E[X_i - X_s] = 0$ ,而方差只依赖于时间区间  $t_s$ ,即  $D[X_t - X_s] = a(t - s)$ ,其中  $a$  依赖于液体本身的扩散常数(不同的液体一般有不同的  $a$ )。显然这种情况下的微粒扩散运动—布朗运动,是一个维纳过程,故此维纳过程有时也直接称之为布朗运动。

#### 例 7.30 热噪声是由导线中电子的布朗运动引起的随机现象。

电子通过单位导体截面的瞬时电荷量  $q(t)$  是典型的维纳过程,它服从高斯分布的原因在于它是由许多独立的电子随机运动叠加产生的。每个电子电荷量与通过截面的总电荷量相比为极小量,故过程可看成是连续的。

## 7.3 独立随机过程

这一节扼要介绍独立随机过程。正如本章开始时所指出的,独立随机过程是种很特殊的随机过程。这类随机过程的特点是:过程在任一时刻的状态和任何其他时刻状态之间互不影响。

下面对它给出较严格的规定。

### 定义 7.31 独立随机过程

如果随机过程  $X(t)$  ( $t \in T$ ) 在时间  $t$  的任意  $n$  个时刻  $t_1, t_2, \dots, t_n$  所得到的  $n$  个随机变量  $X(t_1), X(t_2), \dots, X(t_n)$  互为统计独立; 或者说, 过程  $X(t)$  的  $n$  维分布函数可表示为

$$F_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \prod_{i=1}^n F_X(x_i; t_i). \quad (7.189)$$

则称  $X(t)$  为独立随机过程。 ♣

由式 (7-174) 可知, 独立随机过程的一维分布函数包含了该过程的全部统计信息, 按照时间参数  $T$  的连续还是离散, 独立随机过程可分成两类:

① 当参数集  $T$  是一个可列集时, 独立随机过程就是独立离散时间随机过程—独立随机(变量)序列。例如: 在时刻  $F_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \prod_{i=1}^n F_X(x_i; t_i)$  独立地重复抛掷硬币, 正面对应数为 1, 反面对应数为 0,  $X_n$  表示  $t = n$  时抛掷的结果, 这样,  $X_1, X_2, \dots, X_n, \dots$  就构成了一个独立随机序列。

② 当参数集  $T$  是一个不可列集时, 独立随机过程就是独立连续时间随机过程。此过程的样本函数极不规则, 而且可能处处不连续。

**例 7.32** 设独立随机过程  $X(t)$  的一维分布函数为

$$F_X(x; t) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du; \quad (7.190)$$

即过程  $X(t)$  的一维分布函数是与  $t$  无关的标准正态分布。因此,  $X(t_1), X(t_2), \dots, X(t_n)$  的分布函数皆为

$$\int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{4}} du. \quad (7.191)$$

进而可得  $X(t + \Delta t) - X(t)$  的分布函数为

$$\int_{-\infty}^x \frac{1}{\sqrt{4\pi}} e^{-\frac{u^2}{4}} du. \quad (7.192)$$

于是, 对所有的  $\Delta t$ , 当  $\epsilon \rightarrow 0$  时, 则有

$$P\{|X(t + \Delta t) - X(t)| < \epsilon\} = \int_{-\epsilon}^{\epsilon} \frac{1}{\sqrt{4\pi}} e^{-\frac{u^2}{4}} du \rightarrow 0. \quad (7.193)$$

上式表明, 在  $\Delta t \rightarrow 0$  时,  $X(t + \Delta t)$  与  $X(t)$  的差小于任意正数  $\epsilon$  的概率趋于 0。也就是说, 过程  $X(t)$  的样本函数几乎处处不连续。

实际上,此种连续参数的独立随机过程从物理观点来看是不存在的。因为当  $t_1$  与  $t_2$  ( $t_2 > t_1$ ) 充分接近时,我们完全有理由断言,状态  $X(t_2)$  将依赖于  $X(t_1)$  的统计信息。所以,连续参数的独立随机过程被认为是一种理想化的随机过程。由于它在数学处理上较为简便,故常用来分析某些实际问题。

独立随机过程的重要应用就是高斯白噪声,常用它来模拟电子技术中的各种随机噪声,例如,电阻热噪声就是一种正态分布的白噪声。

如果设  $N(t)$  为高斯白噪声,由于此白噪声的相关系数为

$$\rho_N(\tau) = \begin{cases} 1, & \tau = 0 \\ 0, & \tau \neq 0 \end{cases} \quad (7.194)$$

可见,白噪声在任意两个相邻时刻  $t$  和  $t + \tau$  (不管此二时刻多么邻近) 的状态  $N(t)$  和  $N(t + \tau)$  都是不相关的;又因该白噪声服从高斯分布,故不相关与统计独立是等价的,即随机变量  $N(t)$  和  $N(t + \tau)$  总是统计独立的。所以,高斯白噪声可以认为是大量相互独立的无限窄脉冲的随机组合,实际的随机噪声与理想白噪声的波形如图 7-11 所示。

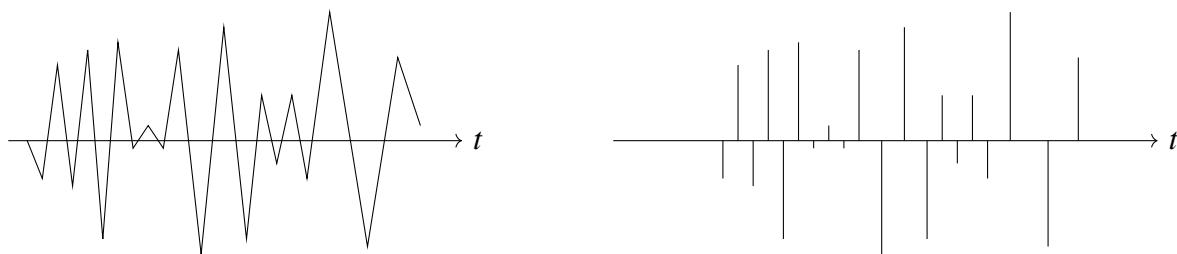


图 7-11 实际的随机噪声与理想白噪声的波形, a) 实际的随机噪声 (b) 理想白噪声

## 7.4 习题

练习 7.1 已知独立随机序列  $\{X(n)\}$  的各个随机变量分别具有概率密度  $F_X(x_1), \dots, F_X(x_n)$ 。

$$\left\{ \begin{array}{l} Y_1 = X_1 \\ Y_2 + CY_1 = X_2 \\ \vdots \\ Y_n + CY_{n-1} = X_n \\ \vdots \end{array} , \quad n \geq 2, C \text{ 为常数} \right. \quad (7.195)$$

令构成一个新序列  $\{Y(n)\}$ 。试证序列  $\{Y(n)\}$  为马尔可夫序列。

练习 7.2 写出下列集合的马尔可夫链的转移概率密度矩阵

①  $I_1 = \{0, 1, 2, \dots, n\}$  ( $n \geq 2$ ) 是有限个正整数集合, 若

$$p_{ij} = \begin{cases} p_{00} = p_m = 1 \\ p, & j = i + 1, \quad q = 1 - p \\ q, & j = i - 1 \\ 0, & \text{其他} \end{cases} \quad (7.196)$$

②  $I_2 = \{\dots, -2, -1, 0, 1, 2, \dots\}$  是全体整数的集合, 若

$$p_{ij} = \begin{cases} p, & j = i + 1 \\ q, & j = i - 1, \quad q = 1 - p \\ 0, & \text{其他} \end{cases} \quad (7.197)$$

**练习 7.3** 带有一个吸收壁的随机游动, 仍考虑质点  $M$  的随机游动, 其基本规律基本上同例 7.10 所述, 但其状态空间  $I = \{0, 1, 2, \dots\}$ , 而且当  $M$  一旦到达  $X_n = 0$  以后,  $X_{n+2}$  也就停留在零这个状态上, 这样的状态称为吸收态。又因“0”为状态空间的个端点(壁), 故称此过程为带有一个吸收壁的随机游动。试确定此过程是否为齐次马尔可夫链, 并写出它的一步转移概率。

**练习 7.4** 带有一个反射壁的随机游动随机游动取状态空间  $I = \{0, 1, 2, \dots\}$ , 这里和上题有所不同, 0 这个状态不再是吸收态了, 一旦质点  $M$  进入 0 状态后, 下步它以概率  $p$  向右移一格, 以概率  $q$  停在 0 状态。这后一种情况可作如下解释: 设想在 “-1/2 处有一反射壁, 每次质点态。正是基于上述解释, 我们把这类过程称为带有一个反射壁的随机游动。试确定此过程是否为齐次马尔可夫链, 并写出它的一步转移概率。

**练习 7.5** 设有三个状态  $\{0, 1, 2\}$  的马尔夫链, 其一步转移概率矩阵为

$$P = \begin{bmatrix} 0 & 1 & 0 \\ q & 0 & p \\ 0 & 1 & 0 \end{bmatrix}, q = 1 - p \quad (7.198)$$

① 试求  $P(2)$ , 并证明  $P(2) = P(4)$ ; ② 求  $P(n), n \geq 1$ 。

**练习 7.6** 天气预报问题。设明日是否有雨仅与今日的天气(是否有雨)有关, 而与过去的天气无关。已知今日有雨而明日也有雨的概率为 0.7, 今日无雨而明日有雨的概率为 0.4。把“有雨”称作“1”状态天气, 而把“无雨”称为“2”状态天气, 则本题属于一个两状态的马尔可夫链。试求: ① 它的一步至四步转移概率矩阵。② 今日有雨而后日(第三日)无雨, 今日有雨而第四日也有雨, 今日无雨而第五日也无雨的概率各是多少?

练习 7.7 设有  $\{0,1\}$  两个状态的马尔可夫链, 其一步转移概率矩阵为

$$P = \begin{bmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{bmatrix}, \quad (7.199)$$

试求以下转移概率:  $f_{00}(1), f_{00}(2), f_{00}(3), f_{01}(1), f_{01}(2), f_{01}(3)$ .

练习 7.8 已知某(齐次)马尔可夫链的一步转移概率矩阵为

$$P = \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{bmatrix}. \quad (7.200)$$

试证:  $n \rightarrow \infty$  时, 有

$$P(n) = P^n \rightarrow \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}. \quad (7.201)$$

练习 7.9 设有  $\{0,1,2\}$  三个状态的马尔可夫链, 其一步转移概率矩阵为

$$P = \begin{bmatrix} p_1 & q_1 & 0 \\ 0 & p_2 & q_2 \\ q_3 & 0 & p_3 \end{bmatrix}, q_i = 1 - p_i, i = 1, 2, 3 \quad (7.202)$$

试求:  $f_{00}(1), f_{00}(2), f_{00}(3), f_{01}(1), f_{01}(2), f_{01}(3)$ .

练习 7.10 已知(齐次)马尔可夫链的一步转移概率矩阵为

$$P = \begin{bmatrix} 1/2 & 1/3 & 1/6 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/2 & 1/6 \end{bmatrix} \quad (7.203)$$

① 问此链共有几个状态? 是否遍历?

② 求其一步转移概率矩阵。

③  $\lim_{n \rightarrow \infty} p_{ij}(n) = p_j$  是否存在? 并求之。

练习 7.11 已知马尔可夫链的一步转移概率矩阵为

$$P = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/3 & 0 & 2/3 \\ 1/2 & 1/2 & 0 \end{bmatrix}. \quad (7.204)$$

① 试问此链共有几个状态?

- ② 是否具有遍历性?  
 ③ 求出极限分布  $\{p_j\}$  的各个概率。

▣ **练习 7.12** 已知有  $\{0, 1, 2\}$  三个状态的马尔可夫链, 它的一步转移概率矩阵为

$$P = \begin{bmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{bmatrix}. \quad (7.205)$$

同此链何时具有遍历性? ② 求出极限分布  $\{p_j\}$  的各个概率。

▣ **练习 7.13** 已知有五个状态  $\{1, 2, 3, 4, 5\}$  的马尔可夫链, 它的一步转移概率矩阵为

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 \\ 0 & 1/3 & 1/3 & 1/3 & 0 \\ 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}. \quad (7.206)$$

同此链何时具有遍历性? ② 求出极限分布  $\{p_j\}$  的各个概率。

▣ **练习 7.14** 对于随机过程  $X(t)$ , 若对每个长  $t \leq t_1 < t_2$ ,  $X(t_2) - X(t_1)$  与  $X(t)$  都是统计独立的, 试证  $X(t)$  为马尔可夫过程。

▣ **练习 7.15** 对于马尔可夫过程  $X(t)$ , 证明  $E[X(t_n)|X(t_{n-1}), X(t_{n-2}), \dots, X(t_1)] = E[X(t_n)|X(t_{n-1})]$ .

▣ **练习 7.16** 证明: 独立增量过程  $X(t)$  是一个特殊的马尔可夫过程。(提示: 只要证明独立增量过程  $X(t)$  具有无后效性即可)。

▣ **练习 7.17** 随机电报信号  $X(t)$ , 其样本函数如图 7-12 所示, 满足下述三个条件

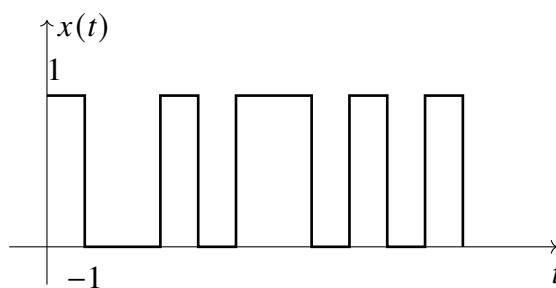


图 7-12 随机电报信号的一个样本函数

① 在任何时刻  $t$ ,  $X(t)$  只能等概率地取 0 或 1 两个状态, 即  $P\{X(t) = 0\} = P\{X(t) = 1\} = 1/2$ 。

② 每个状态的持续时间是随机的, 若在间隔  $(0, t)$  内波形变化的次数  $K$  服从泊松分布, 即

$$P\{K = k\} = p_k(0, t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t} \quad (7.207)$$

式中的  $\lambda$  为单位时间内波形的平均变化次数。

③  $X(t)$  取何值与随机变量  $K$  互为统计独立。试求此随机电报信号  $X(t)$  的均值、自相关函数、自协方差函数和功率谱密度。

✉ **练习 7.18** 已知泊松冲激序列  $Z(t) = \sum_i \delta(t - t_i)$ , 这里  $t_i$  是在时间轴上具有均匀密度  $A$  的随机点。过程  $Z(t)$  通过一个具有冲激响应  $h(t) = e^{-at}U(t)$  的线性系统, 在系统的输出端得到散粒噪声  $X(t)$ 。试求此散粒噪声的均值、方差和功率谱密度。

✉ **练习 7.19** 证明: 维纳过程的一维特征函数为

$$Q_W(u; t) = e^{-atu^2/2}. \quad (7.208)$$

✉ **练习 7.20** 广义维纳过程。已知一个具有参数  $a$  的维纳过程  $W(t)$ , 以及两个连续函数  $m(t), v(t)$ , 其中  $v(t)$  为非减函数, 构成一新的随机过程。

$$Y(t) = m(t) + W[v(t)] \quad (7.209)$$

证明:  $Y(t)$  的一维特征函数为

$$Q_Y(u; t) = \exp \left[ j u m(t) - \frac{1}{2} \alpha v(t) u^2 \right]. \quad (7.210)$$

✉ **练习 7.21** 已知一维纳过程  $W(t)$  的自相关函数为  $R_w(t_1, t_2) = \alpha t_2, t_1 > t_2$ . 现构成新的随机过程  $Y(t) = W^2(t)$ , 求当  $t_1 > t_2$  时,  $Y(t)$  的自相关函数  $R_Y(t_1, t_2)$ 。

## 7.5 参考文献

- 奥本海姆 AV, 谢弗 RW. 1982, 数字信号处理. 董士嘉等译 · 北京: 科学出版社
- 丁玉美, 阎永红, 高新波. 2002. 数字信号处理一时域离散随机信号处理. 西安: 西安电子科技大学出版社
- 林茂庸, 柯有安 1980 雷达信号理论, 北京: 国防工业出版社
- 陆大金. 1986. 随机过程及其应用. 北京: 清华大学出版社
- 陆光华等. 2000 机信号处理, 西安: 西安电子科技大学出版社
- 王永等. 200 随机信号分析基础. 北京: 电子工业出版社
- 吴析耀. 1984. 随机过程北京: 国防工业出版社
- 场福生. 1990 随机信号分析 · 北京: 清华大学出版社
- 伊曼纽尔 · 帕尔逊, 1987 随机过程. 邓水录, 杨振明译. 北京: 高等教育出版社
- 张树京, 张思东 200 统计信号处理 · 北京: 机械工业出版社
- 中山大学数学力学系. 1981. 概率论及数理统计 (上下册) 北京: 人民教育出版
- 朱华等, 1990. 随机信号分析 · 北京: 北京理工大学出版社
- akis DC, Ingle V K, Kogon S M. 2003. 统计与自适应信号处理. 周正等译, 北京: 电子工业出版社
- Papoulis A, 1984. Probability Random Variables and Stochastic Processes. 2nd ed. New York: McGraw-Hill
- Peebles P Z, Jr. 1980. Probability Random Variables and Random Signal Principles. New York: MGraw-Hill

# 8

## 附录——习题

练习 8.1 已知系统的输入为单位谱密度的白噪声, 输出的功率谱密度为

$$G_Y(\omega) = \frac{\omega^2 + 4}{\omega^4 + 10\omega^2 + 9},$$

求此稳定系统的单位冲激响应  $h(t)$ .

解: 分解功率谱密度为

$$\begin{aligned} G_Y(\omega) &= \frac{4}{\omega^4 + 10\omega^2 + 9} = \frac{1}{2} \left[ \frac{1}{(\omega^2 + 1)} - \frac{1}{(\omega^2 + 9)} \right], \\ H(s) &= \frac{4 - s^2}{s^4 - 10s^2 + 9} = \frac{(2-s)(2+s)}{(s^2-1)(s^2-9)} = \frac{(2-s)(2+s)}{(s-1)(s-3)(s+1)(s+3)}. \end{aligned}$$

其中  $s = j\omega$ ,  $H(s) = \frac{s+2}{(s+3)(s+1)} = \frac{1}{2} \left[ \frac{1}{s+3} + \frac{1}{s+1} \right]$ , 所以  $h(t) = \frac{1}{2} (e^{-3t} + e^{-t}) u(t)$ , (其中的一个正确答案)。

练习 8.2 零均值高斯信号  $X(t)$  的自相关函数为  $R_X(\tau) = 0.5e^{-|\tau|}$ , 求  $X(t)$  的一维和二维概率密度。

解: (1) 因为  $m_X(t) = 0$ ,  $D_X(t) = C_X(0) = R_X(0) = 0.5$ , 所以  $X(t)$  的一维概率密度函数为:

$$\begin{aligned} f_X(x; t) &= \frac{1}{\sqrt{2\pi D_X(t)}} \exp \left\{ -\frac{[x - m_X(t)]^2}{2D_X(t)} \right\} \\ &= \frac{1}{\sqrt{\pi}} e^{-x^2}. \end{aligned}$$

(2) 高斯信号  $X(t)$  的二维概率密度函数为:

$$\begin{aligned}\mathbf{X} &= \begin{pmatrix} X(t_1) \\ X(t_2) \end{pmatrix}, \mathbf{t} = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}, \boldsymbol{\mu} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \mathbf{C} &= \begin{pmatrix} C(t_1, t_1) & C(t_1, t_2) \\ C(t_2, t_1) & C(t_2, t_2) \end{pmatrix} = \begin{pmatrix} R_X(t_1, t_1) & R_X(t_1, t_2) \\ R_X(t_2, t_1) & R_X(t_2, t_2) \end{pmatrix} \\ &= \begin{pmatrix} 0.5 & 0.5 \exp(-|\tau|) \\ 0.5 \exp(-|\tau|) & 0.5 \end{pmatrix},\end{aligned}$$

其中  $C(t_i, t_j)$  为协方差, 则

$$f_{\mathbf{X}}(\mathbf{x}, \mathbf{t}) = \frac{1}{2\pi |\mathbf{C}^{1/2}|} \exp\left(-\frac{\mathbf{x}^T \mathbf{C}^{-1} \mathbf{x}}{2}\right).$$

**练习 8.3** 已知随机变量  $X$  和  $Y$ , 有  $E(X) = 1, E(Y) = 1, D(X) = 4, D(Y) = 16, \rho_{XY} = 0.5$ , 又  $U = 3X + Y, V = X - 2Y$ , 试求  $E(U), E(V), D(U), D(V), \rho_{UV}, R_{UV}, C_{UV}$ .

解:  $E(X^2) = D(X) + [E(X)]^2 = 5, E(Y^2) = D(Y) + [E(Y)]^2 = 25$ ; 又因为  $E(XY) = \text{Cov}(X, Y) + E(X) \times E(Y) = \rho_{XY} \sqrt{D(X)D(Y)} + E(X) \times E(Y) = 7$ . 于是有

$$E(U) = E(3X + Y) = 3E(X) + E(Y) = 6, E(V) = E(X - 2Y) = E(X) - 2E(Y) = -5.$$

$$D(U) = E(U^2) - (E(U))^2 = E[9X^2 + 6XY + Y^2] - (E(U))^2 = 76.$$

$$D(V) = E(V^2) - (E(V))^2 = E[X^2 - 4XY + 4Y^2] - (E(V))^2 = 52.$$

$$\begin{aligned}R_{UV} &= E(UV) = E[(3X + Y)(X - 2Y)] = E[3X^2 - 5XY - 2Y^2] \\&= 3E(X^2) - 6E(XY) + E(XY) - 2E(Y^2) = -70.\end{aligned}$$

$$\text{Cov}(U, V) = E(UV) - E(U) \times E(V) = -70 + 30 = -40.$$

$$\rho_{UV} = \frac{\text{Cov}(U, V)}{\sqrt{D(X)D(Y)}} = -\frac{40}{4\sqrt{19 \times 13}}.$$

**练习 8.4** 两个随机信号  $X(t) = \sin(\omega_0 t + \theta)$  与  $Y(t) = \cos(\omega_0 t + \theta)$ ,  $\theta$  为  $[-\pi, \pi]$  上的均匀分布随机变量,  $\omega_0$  是常数。

(1) 试判断  $X(t)$  和  $Y(t)$  在同一时刻的独立性、相关性及正交性。

(2) 试判断  $X(t)$  和  $Y(t)$  是否联合广义平稳。

解: (1) 由于  $X(t)$  和  $Y(t)$  包含同一随机变量  $\theta$ , 因此非独立。根据题意有  $f(\theta) = \frac{1}{2\pi}$ .

$$\begin{aligned} E[X(t)] &= E[\sin(\omega_0 t + \theta)] = \int_{-\pi}^{\pi} \frac{1}{2\pi} \sin(\omega_0 t + \theta) d\theta = 0. \\ E[Y(t)] &= E[\cos(\omega_0 t + \theta)] = \int_{-\pi}^{\pi} \frac{1}{2\pi} \cos(\omega_0 t + \theta) d\theta = 0. \end{aligned}$$

$$\begin{aligned} C_{XY}(t_1, t_2) &= R_{XY}(t_1, t_2) = E[X(t_1)Y(t_2)] \\ &= E[\sin(\omega_0 t_1 + \theta)\cos(\omega_0 t_2 + \theta)] \\ &= \int_{-\pi}^{\pi} \frac{1}{2\pi} \sin(\omega_0 t_1 + \theta)\cos(\omega_0 t_2 + \theta) d\theta \\ &= \int_{-\pi}^{\pi} \frac{1}{4\pi} \{\sin[\omega_0(t_1 + t_2) + 2\theta] + \sin[\omega_0(t_1 - t_2)]\} d\theta \\ &= \frac{1}{2} \sin[\omega_0(t_1 - t_2)]. \end{aligned}$$

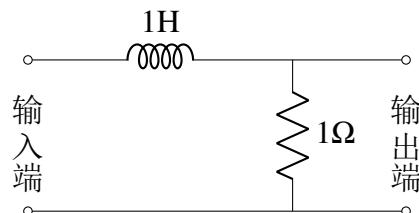
由于  $R_{XY}(t, t) = C_{XY}(t, t) = 0$ ,  $X(t)$  和  $Y(t)$  在同一时刻正交、线性无关。除  $\omega_0(t_1 - t_2) = \pm k\pi$  外的不同时刻, 有  $R_{XY}(t_1, t_2) = C_{XY}(t_1, t_2) \neq 0$ , 所以  $X(t_1)$  和  $Y(t_2)$  非正交且线性相关。

(2) 由于  $E[X(t)] = E[Y(t)] = 0$ ,  $X(t)$  和  $Y(t)$  均值平稳。

$$\begin{aligned} R_X(t_1, t_2) &= E[\sin(\omega_0 t_1 + \theta)\sin(\omega_0 t_2 + \theta)] \\ &= \int_{-\pi}^{\pi} \frac{1}{2\pi} \sin(\omega_0 t_1 + \theta)\sin(\omega_0 t_2 + \theta) d\theta \\ &= \int_{-\pi}^{\pi} \frac{1}{4\pi} \{-\cos[\omega_0(t_1 + t_2) + 2\theta] + \cos[\omega_0(t_1 - t_2)]\} d\theta \\ &= \frac{1}{2} \cos[\omega_0(t_1 - t_2)] = \frac{1}{2} \cos(\omega_0\tau). \end{aligned}$$

同理可得  $R_Y(t_1, t_2) = R_X(t_1, t_2)$ , 因此  $X(t)$  和  $Y(t)$  均广义平稳。由于  $R_{XY}(t_1, t_2) = C_{XY}(t_1, t_2) = \frac{1}{2} \sin[\omega_0(t_1 - t_2)] = \frac{1}{2} \sin(\omega_0\tau)$ , 因此  $X(t)$  和  $Y(t)$  联合广义平稳。

**练习 8.5** 线性时不变系统的框图如图所示, 若输入白噪声的双边功率谱密度  $\frac{N_0}{2} = 2W/Hz$ , 求系统输出噪声的功率谱密度函数和自相关函数, 以及输出噪声总功率。



**解:** 系统的传递函数为  $H(j\omega) = \frac{R}{R+j\omega L} = \frac{1}{j\omega+1}$ , 则系统输出功率谱密度为  $G_Y(\omega) = G_X(\omega) \times |H(j\omega)|^2 = 2 \times \frac{1}{1+\omega^2} = \frac{2}{1+\omega^2}$ 。输出噪声的自相关函数为  $R_Y(\tau) = e^{-|\tau|}$ , 输出噪声总功率为  $P_N = R_Y(0) = 1(W)$ 。

✉ **练习 8.6** 已知平稳过程  $X(t)$  的功率谱密度为  $G_X(\omega)$ , 随机过程  $W(t) = X(t) \cos \omega_0 t - \hat{X}(t) \sin \omega_0 t$ , 求随机过程  $W(t)$  的功率谱密度。

**解:**

$$\begin{aligned} G_W(\omega) &= \frac{1}{2} [G_X(\omega + \omega_0) + G_X(\omega - \omega_0)] \\ &\quad + \frac{1}{2} G_X(\omega) \operatorname{sgn}(\omega - \omega_0) - \frac{1}{4} G_X(\omega + \omega_0) \operatorname{sgn}(\omega + \omega_0). \end{aligned}$$

✉ **练习 8.7** 已知齐次马尔科夫链的一步转移概率矩阵  $\mathbf{P}$  为

$$\mathbf{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$$

- (1) 该马尔科夫链是否具有遍历性?
- (2) 求出其二步转移概率矩阵  $\mathbf{P}(2)$ 。
- (3) 判断  $\lim_{n \rightarrow \infty} P_{ij}(n) = p_j$  是否存在, 并求之。

**解:** (1) 马尔科夫链的遍历性: 存在  $m = 1$ , 一步转移概率矩阵  $\mathbf{P}(1)$  具有如下形式

$$\mathbf{P}(1) = \mathbf{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}.$$

对任意的  $i, j = 1, 2, 3$ , 都有  $p_{ij}(m) > 0$ , 因此, 马尔科夫链具有遍历性。

- (2) 二步转移概率矩阵

$$\mathbf{P}(2) = \mathbf{P}^2 = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}^2 = \begin{pmatrix} \frac{19}{48} & \frac{1}{3} & \frac{13}{48} \\ \frac{13}{36} & \frac{13}{36} & \frac{5}{18} \\ \frac{17}{48} & \frac{17}{48} & \frac{7}{24} \end{pmatrix}.$$

- (3) 若  $\lim_{n \rightarrow \infty} P_{ij}(n) = p_j$  存在, 即过程平稳, 有  $\mathbf{P}^T p = p$ ,

$$p = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \mathbf{P}^T \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}^T \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}$$

$P^T$  增广矩阵为

$$\left( \begin{array}{cccc} -1/2 & 1/3 & 1/4 & 0 \\ 1/4 & -2/3 & 1/2 & 0 \\ 1/4 & 1/3 & -3/4 & 0 \\ 1 & 1 & 1 & 1 \end{array} \right) \sim \left( \begin{array}{cccc} 1 & 0 & 0 & 16/43 \\ 0 & 1 & 0 & 15/43 \\ 0 & 0 & 1 & 12/43 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

由上三角形式，可以求得分布  $\{p_j\}$ ,  $j = 1, 2, 3$ :

$$p_1 = \frac{16}{43}, p_2 = \frac{15}{43}, p_3 = \frac{12}{43}.$$

Matlab 计算代码如 8-1 所示。

算法 8-1 齐次马尔科夫链极限分布的计算代码

```

1 clc, clear
2 format rat
3 A=[1/2,1/4,1/4;1/3,1/3,1/3;1/4,1/2,1/4]
4 A1=A'-eye(3)
5 Aexpand=[A1;1,1,1]
6 AugA=[Aexpand,[0;0;0;1]]
7 rref(AugA)

```

