

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 2.8 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 ζ 的集合,称之为样本空间 Ω 。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 ζ_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

$\{x_1(t), x_2(t), \cdots, x_m(t)\}$ 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。 $\zeta \in \Omega$, Ω 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。

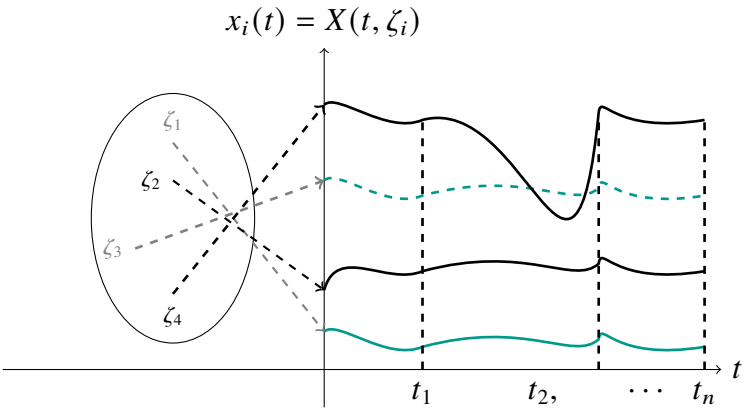


图 2-1 典型的随机过程模型

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 ζ 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 Ω 上 ζ 的函数, 它所有试验结果 ζ 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 $x_1(t), x_2(t), \cdots, x_m(t)$ 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 $\{x_1(t), x_2(t), \cdots, x_m(t)\}$ 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 ω 代表某概率空间 (Ω, 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 (Ω, 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 $X(\zeta, n)$ ($n = 1, 2, \dots, m, \dots$) 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。

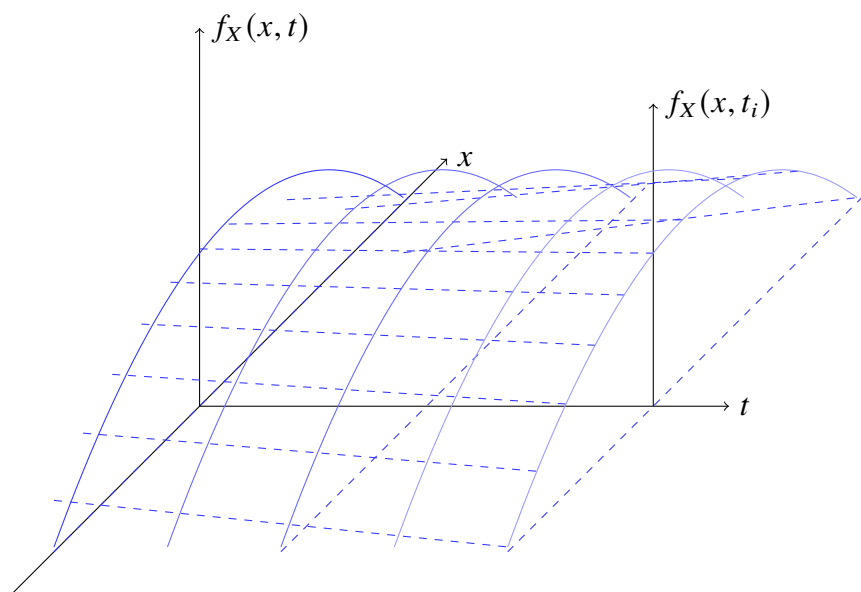


图 2-2 连续型随机过程

(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。

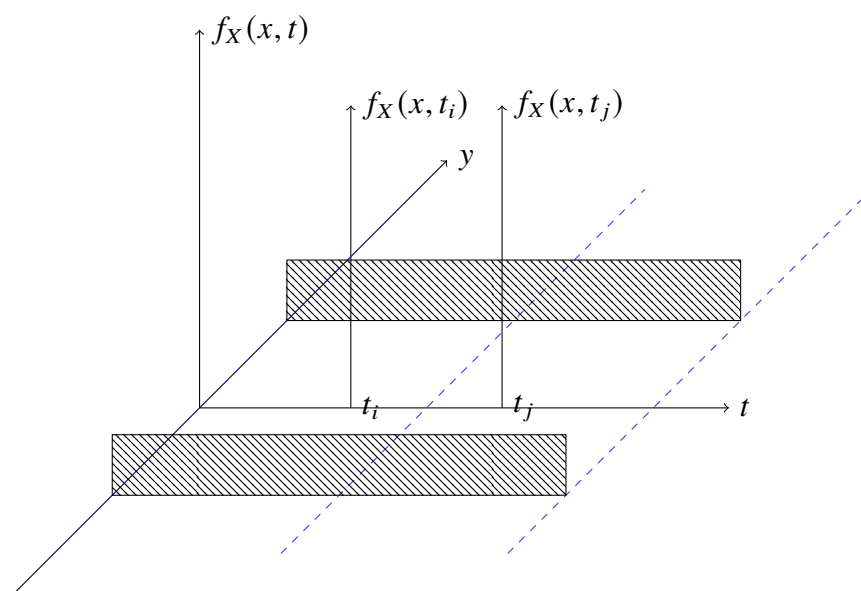


图 2-3 离散型随机过程

(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。

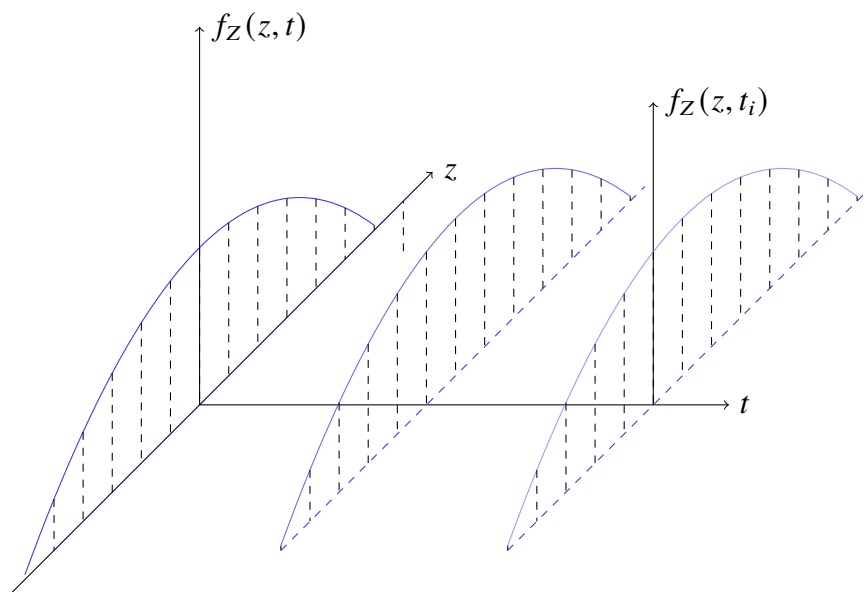


图 2-4 连续随机序列的一个样本函数

(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。

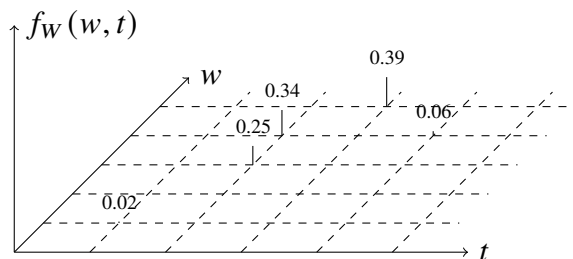


图 2-5 离散随机序列的一个样本函数

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 Ω 和相位 Φ 是随机变量,但对于任一次试验结果 ζ_k 而言,随机变量 A, Ω, Φ , 仅取某个具体的值 a_k, ω_k, φ_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。

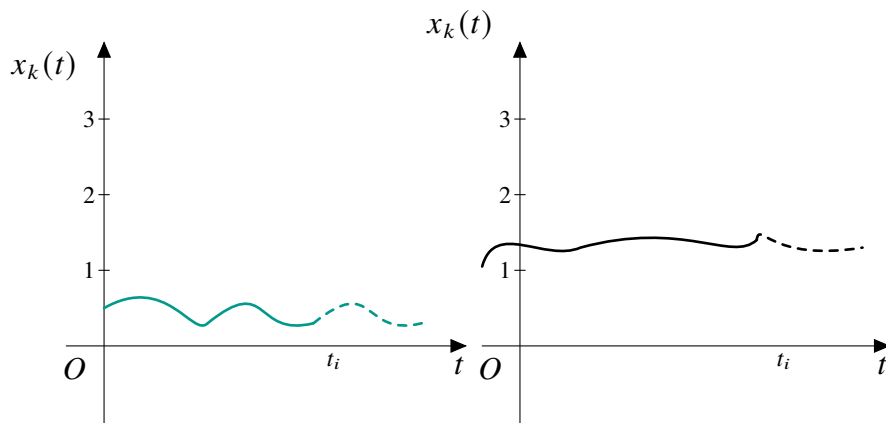


图 2-6 离散随机序列的一个样本函数 (左: 确定性过程, 右: 不确定性过程)

(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)。

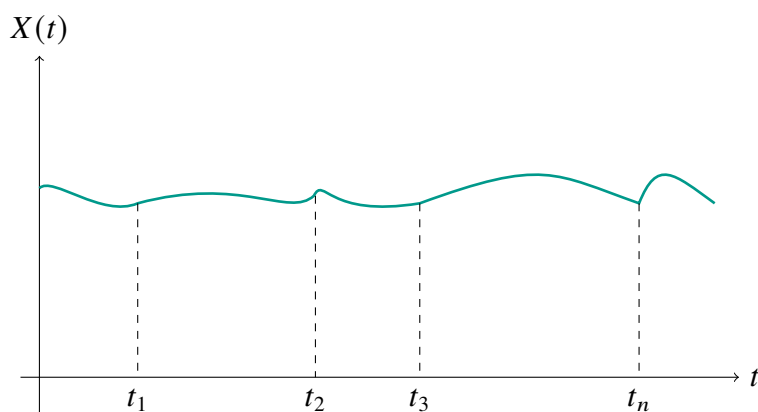


图 2-7 离散随机序列的一个样本函数

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。

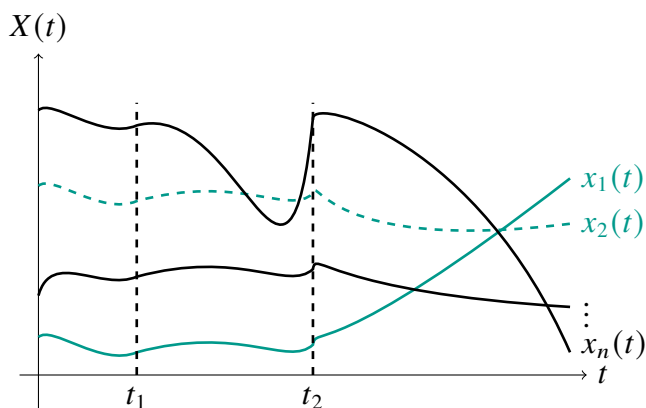


图 2-8 随机过程的分布

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:

$$1^\circ F_X(x_1, x_2, \dots, -\infty, \dots, x_n; t_1, t_2, \dots, t_i, \dots, t_n) = 0. \quad (2.8)$$

$$2^\circ F_X(\infty, \infty, \dots, \infty, t_1, t_2, \dots, t_n) = 1. \quad (2.9)$$

$$3^\circ f_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) \geq 0. \quad (2.10)$$

$$4^\circ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) dx_1 dx_2 \cdots dx_n = 1. \quad (2.11)$$

$$5^\circ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_X(x_1, \dots, x_m, x_{m+1}, \dots, x_n; t_1, \dots, t_m, t_{m+1}, \dots, t_n) dx_{m+1} \cdots dx_n \\ = f_X(x_1, \dots, x_m; t_1, \dots, t_m). \quad (2.12)$$

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。

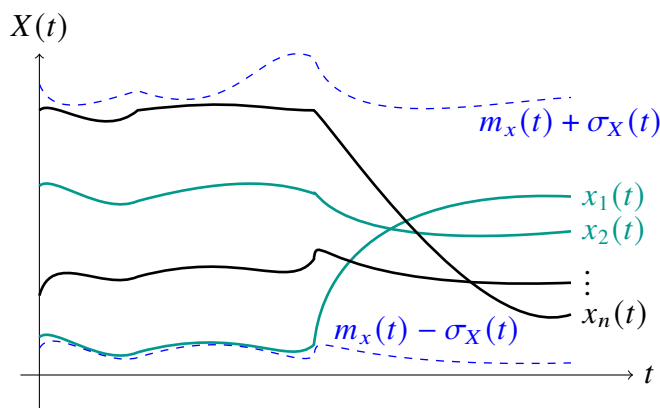


图 2-9 随机过程的数学期望和方差

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 δ 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

$$m_Y(t) = \int_{-\infty}^{\infty} y \sum_{k=1}^m p_k(t) \delta(y - y_k) dy = \sum_{k=1}^m y_k p_k(t). \quad (2.18)$$

$$\psi_Y^2(t) = E[Y^2(t)] = \sum_{i=1}^m y_i^2 p_i(t). \quad (2.19)$$

$$\sigma_Y^2(t) = D[Y(t)] = \sum_{k=1}^m [y_k - m_Y(t)]^2 p_k(t). \quad (2.20)$$

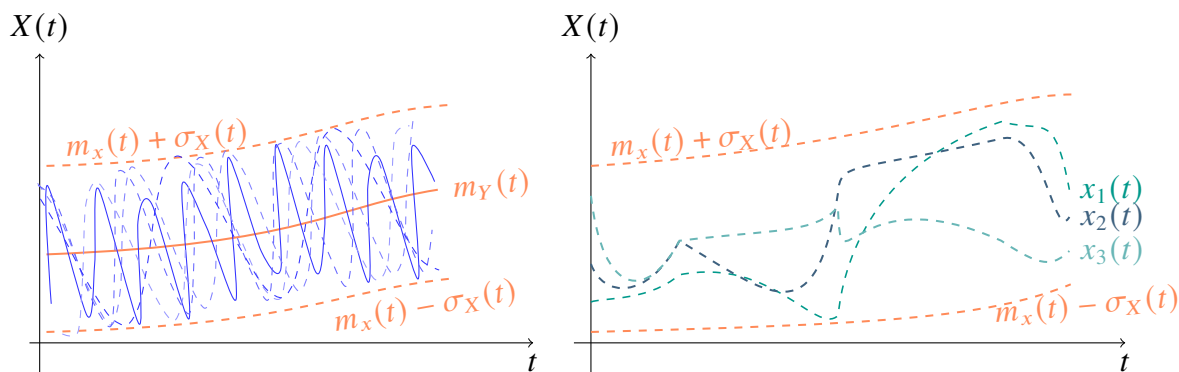


图 2-10 随机过程的自相关函数

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); 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 δ 函数表示, 即

$$f_Y(y_1, y_2, t_1, t_2) = \sum_{k_1, k_2 \in \mathcal{E}_Y} P_{k_1 k_2}(t_1, t_2) \delta(y_1 - k_1) \delta(y_2 - k_2), \quad (2.26)$$

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

$$R_Y(t_1, t_2) = \sum_{k_1, k_2 \in \mathcal{E}_Y} k_1 k_2 P\{Y(t_1) = k_1, Y(t_2) = k_2\}, \quad (2.27)$$

其中 \mathcal{E}_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

$$\begin{aligned} 1^\circ E\{|X(t)|\} &< \infty. \\ 2^\circ E\{|X(t)|^2\} &< \infty. \end{aligned} \quad (2.28)$$

例 2.19 参数为 α 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

$$m_X(t) = E[X(t)] = E[V \cos 4t] = \cos 4t \cdot E[V] = 5 \cos 4t.$$

$$\sigma_X^2(t) = D[X(t)] = D[V \cos 4t] = \cos^2 4t \cdot D[V] = 6 \cos^2 4t.$$

$$R_X(t_1, t_2) = E[X(t_1) X(t_2)] = E[V \cos 4t_1 \cdot V \cos 4t_2].$$

$$= \cos 4t_1 \cos 4t_2 \cdot E[V^2] = 31 \cos 4t_1 \cos 4t_2.$$

$$C_X(t_1, t_2) = E[(X(t_1) - m_X(t_1))(X(t_2) - m_X(t_2))]$$

$$= R_X(t_1, t_2) - m_X(t_1) m_X(t_2)$$

$$= 31 \cos 4t_1 \cos 4t_2 - 5 \cos 4t_1 \cdot 5 \cos 4t_2 = 6 \cos 4t_1 \cos 4t_2.$$

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

$$\begin{aligned}
 m_X(t) &= E[X(t)] = E[Ut] = tE[U] = t \int_0^1 u \cdot 1 \cdot du = \frac{t}{2}. \\
 R_X(t_1, t_2) &= E[X(t_1)X(t_2)] = E[U t_1 \cdot U t_2] = t_1 t_2 E[U^2]. \\
 &= t_1 t_2 \int_{-\infty}^{\infty} u^2 f_U(u) du = t_1 t_2 \int_0^1 u^2 \cdot 1 \cdot du = \frac{t_1 t_2}{3}. \\
 C_X(t_1, t_2) &= R_X(t_1, t_2) - m_X(t_1) m_X(t_2) = \frac{t_1 t_2}{3} - \frac{t_1}{2} \cdot \frac{t_2}{2} = \frac{t_1 t_2}{12}. \\
 \sigma_X^2(t) &= C_X(t, t) = \frac{t^2}{12}.
 \end{aligned}$$

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 $R_X(t_1, t_2)$ 。

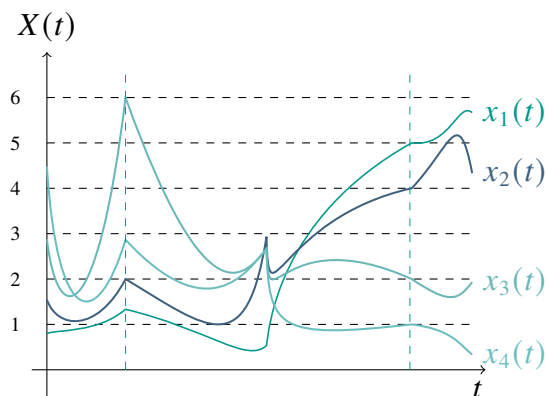


图 2-11 随机过程的样本

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

| | ζ_1 | ζ_2 | ζ_3 | ζ_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

$$\begin{aligned}
 R_X(t_1, t_2) &= \sum_{k_1, k_2 \in \mathcal{E}_Y} \sum k_1 k_2 p_{k_1 k_2}(t_1, t_2) \\
 &= \sum_{k_1, k_2 \in \mathcal{E}_Y} \sum k_1 k_2 p\{X(t_1) = k_1, X(t_2) = k_2\} \\
 &= 0.25 \times (1 \times 5 + 2 \times 4 + 6 \times 2 + 3 \times 1) = 7.
 \end{aligned}$$

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

$$Q_X(u; t) = E[e^{juX(t)}] = \int_{-\infty}^{\infty} e^{jux} f_X(x; t) dx, \quad (2.34)$$

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

$$\frac{\partial^n Q_X(u; t)}{\partial u^n} = j^n \int_{-\infty}^{\infty} x^n e^{jux} f_X(x; t) dx, \quad (2.36)$$

则随机过程 $X(t)$ 的 n 阶原点矩函数为

$$E[X^n(t)] = \int_{-\infty}^{\infty} x^n f_X(x; t) dx = (-j)^n \frac{\partial^n Q_X(u, v)}{\partial u^n} \Big|_{u=0}. \quad (2.37)$$

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

$$f_X(x_1, \dots, x_n; t_1, \dots, t_n) = \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Q_X(u_1, \dots, u_n; t_1, \dots, t_n) e^{-j(u_1 x_1 + \dots + u_n x_n)} du_1 \dots du_n. \quad (2.43)$$

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族与 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。

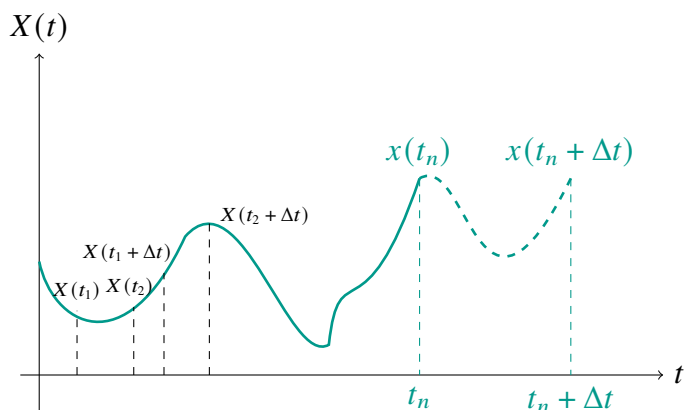


图 2-12 平稳随机过程

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 Δt 后, 其 n 维概率密度保持不变, 即

$$f_X(x_1, \cdots, x_n; t_1 + \Delta t, \cdots, t_n + \Delta t) = f_X(x_1, \cdots, x_n; t_1, \cdots, t_n). \quad (2.45)$$

则称该过程为严平稳随机过程(或狭义平稳过程)。

定义 2.23 严格平稳随机过程

随机过程 $\{X(t)\}$ 所确定的 n 维概率密度函数为 $f(x_1, x_2, \cdots, x_n; t_1, t_2, \cdots, t_n)$, 其中 t_i ($i = 1, 2, \cdots, n$) 为任意取定的 n 个时刻, 如果对任意值 n , 满足

$$f(x_1, x_2, \cdots, x_n; t_1, t_2, \cdots, t_n) = f(x_1, x_2, \cdots, x_n; t_1 + \tau, t_2 + \tau, \cdots, t_n + \tau). \quad (2.46)$$

那么, 称 $\{X(t)\}$ 为严格平稳随机过程, 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, ψ_X^2 和 σ_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 τ , 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

$$\begin{aligned} f_X(x_1, x_2; t_1, t_2) &= f_X(x_1, x_2, t_1 + \Delta t, t_2 + \Delta t) \\ &= f_X(x_1, x_2; 0, t_2 - t_1) = f_X(x_1, x_2; \tau). \end{aligned} \quad (2.52)$$

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 τ 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, 在某一时刻、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, ω_0 皆为常数, 随机变量 Φ 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 Φ 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

$$\begin{aligned} m_X(t) &= E[X(t)] = \int_{-\infty}^{\infty} x(t) f_{\Phi}(\varphi) d\varphi = \int_0^{2\pi} a \cos(\omega_0 t + \varphi) \cdot \frac{1}{2\pi} \cdot d\varphi = 0. \\ R_X(t_1, t_2) &= R_X(t, t + \tau) = E[X(t)X(t + \tau)] \\ &= E[a \cos(\omega_0 t + \Phi) \cdot a \cos(\omega_0(t + \tau) + \Phi)] \\ &= \frac{a^2}{2} E[\cos \omega_0 \tau + \cos(2\omega_0 t + \omega_0 \tau + 2\Phi)] \\ &= \frac{a^2}{2} \left[\cos \omega_0 \tau + \int_0^{2\pi} \cos(2\omega_0 t + \omega_0 \tau + 2\varphi) \cdot \frac{1}{2\pi} d\varphi \right] \\ &= \frac{a^2}{2} \cos \omega_0 \tau = R_X(\tau). \end{aligned} \quad (2.57)$$

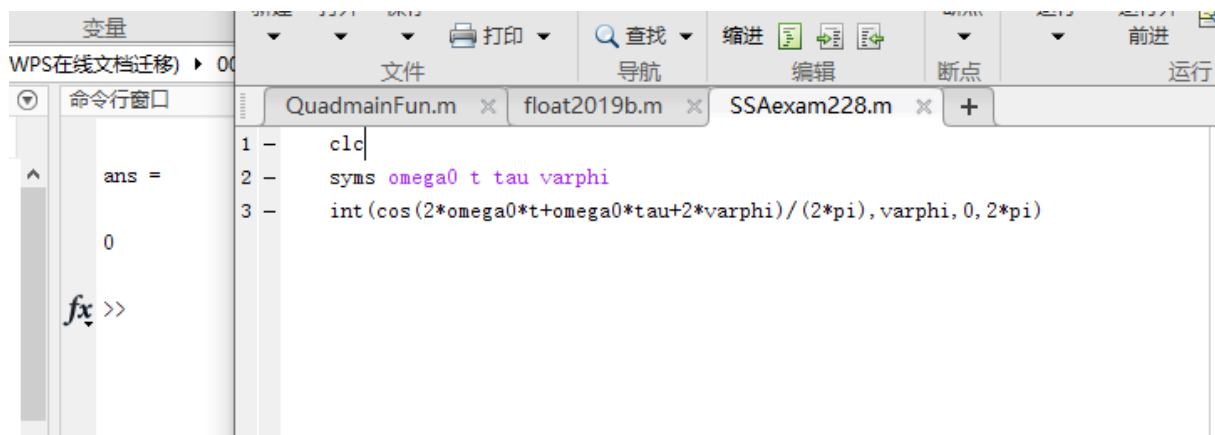
$$E[X^2(t)] = R_X(t, t) = R_X(0) = \frac{a^2}{2} < \infty.$$


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 τ 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 Φ 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

$$\begin{aligned} m_{X_2}(t) &= E[X_2(t)] = E[tY] = tE[Y] = tm_Y. \\ R_{X_2}(t_1, t_2) &= E[X_2(t_1)X_2(t_2)] = E[t_1Y \cdot t_2Y] = t_1t_2E[Y^2] = t_1t_2\psi_Y^2. \end{aligned} \quad (2.59)$$

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 Φ 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 , 则相关函数也将会含有平均分量 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 τ 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。

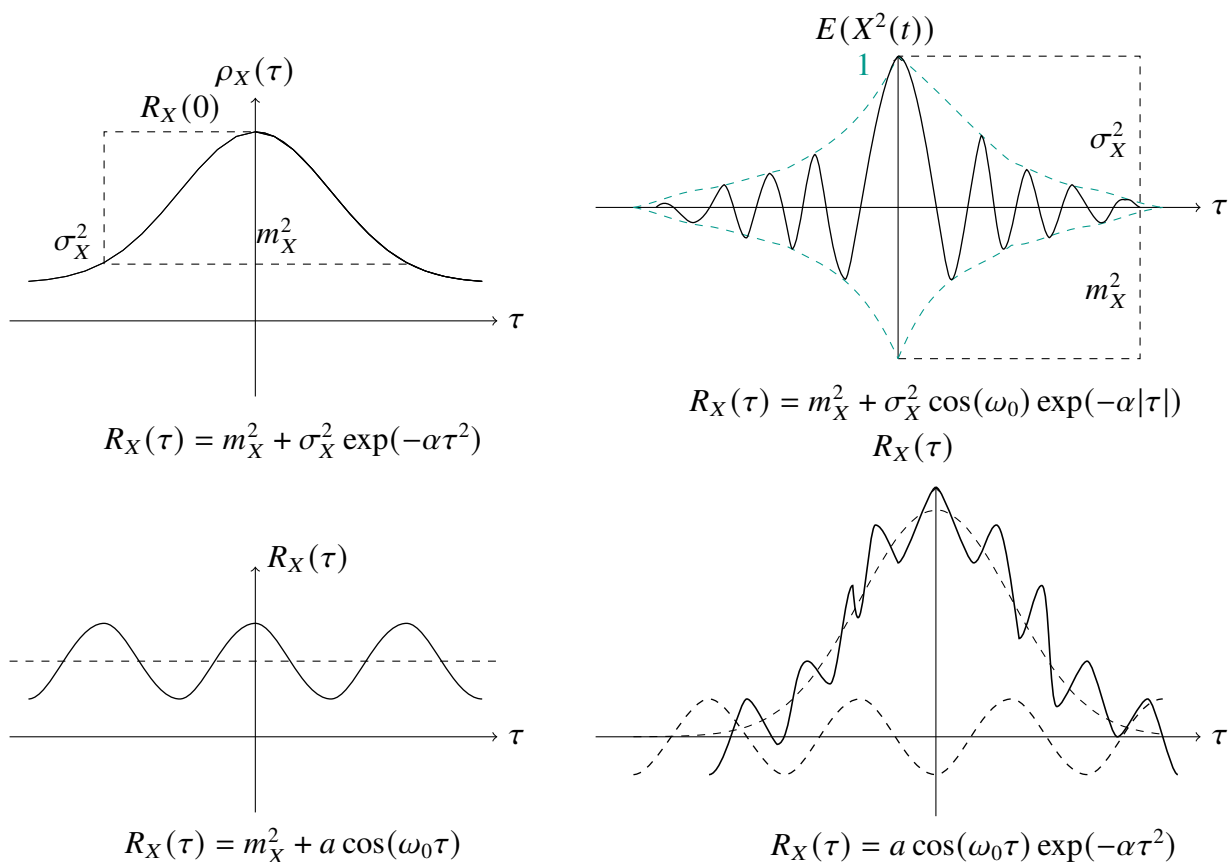


图 2-14 平稳过程的自相关函数

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

$$R_X(\tau) = (100\cos 10\tau) + (100e^{-10|\tau|} + 100) = R_{X_1}(\tau) + R_{X_2}(\tau), \quad (2.72)$$

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 ± 10 , 均方值为 300, 方差为 200。

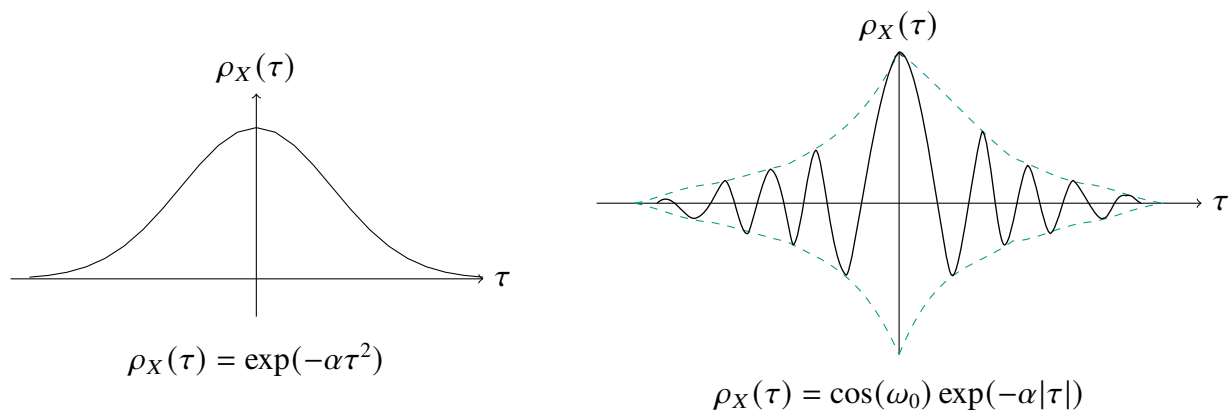


图 2-15 自相关函数 $\rho_X(\tau)$ 的两条典型曲线

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 τ 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 τ 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 τ_0 , 当 $\tau > \tau_0$ 时, 就认为 $X(t)$ 与 $X(t+\tau)$ 不相关, 把这个时间 τ_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 τ'_0 , 即

$$|\rho_X(\tau'_0)| = 0.05. \quad (2.76)$$



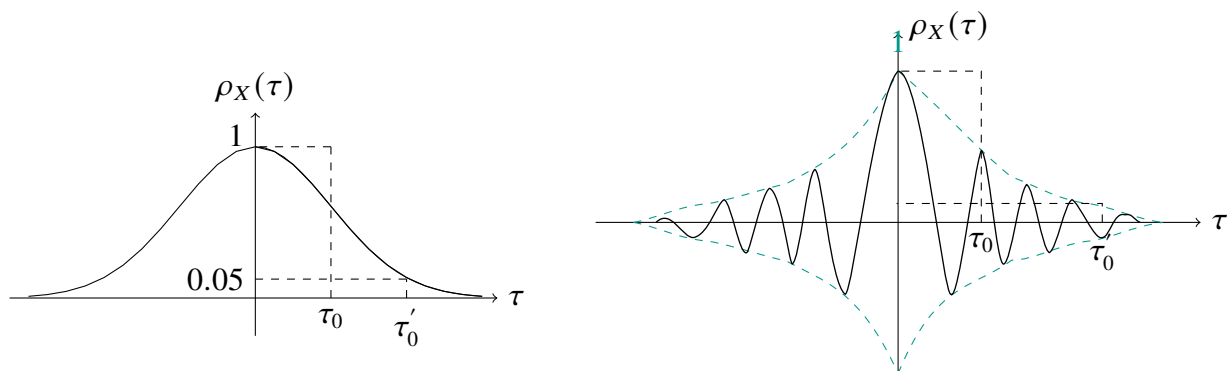


图 2-16 相关时间的定义

(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。

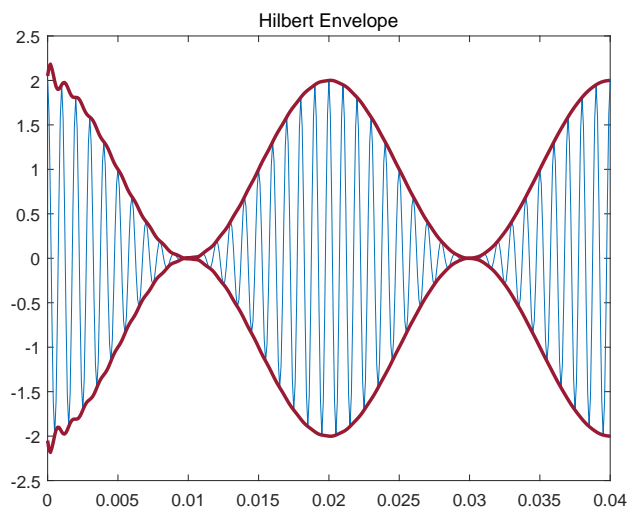


图 2-17 信号包络的计算

如图 2-18 所示, $X(t)$ 曲线越陡, 相关时间 τ'_{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 τ'_{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。

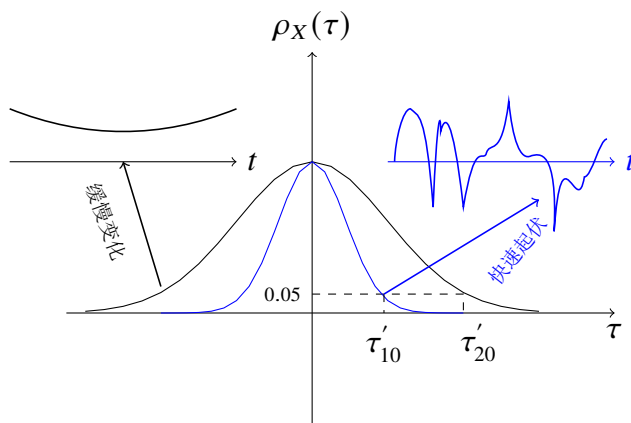


图 2-18 自相关时间

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 ξ_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, θ 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 ξ_0 和 ξ_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

$$f_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n), f_Y(y_1, y_2, \dots, y_n; t'_1, t'_2, \dots, t'_n). \quad (2.84)$$

(1) 两个过程的 $n+m$ 维联合分布函数

$$\begin{aligned} F_{XY}(x_1, \cdots, x_n, y_1, \cdots, y_m; t_1, \cdots, t_n, t'_1, \cdots, t'_m) \\ = P\{X(t_1) \leq x_1, \cdots, X(t_n) \leq x_n, Y(t'_1) \leq y_1, \cdots, Y(t'_m) \leq y_m\}. \end{aligned} \quad (2.85)$$

(2) 两个过程的 $n+m$ 维联合概率密度

$$\begin{aligned} f_{XY}(x_1, \cdots, x_n, y_1, \cdots, y_m; t_1, \cdots, t_n, t'_1, \cdots, t'_m) \\ = \frac{\partial^{n+m} F_{XY}(x_1, \cdots, x_n, y_1, \cdots, y_m; t_1, \cdots, t_n, t'_1, \cdots, t'_m)}{\partial x_1 \cdots \partial x_n \partial y_1 \cdots \partial y_m}. \end{aligned} \quad (2.86)$$

(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

$$\begin{aligned} F_{XY}(x_1, \cdots, x_n, y_1, \cdots, y_m; t_1, \cdots, t_n, t'_1, \cdots, t'_m) \\ = F_X(x_1, \cdots, x_n; t_1, \cdots, t_n) F_Y(y_1, \cdots, y_m; t'_1, \cdots, t'_m). \end{aligned} \quad (2.87)$$

或

$$\begin{aligned} f_{XY}(x_1, \cdots, x_n, y_1, \cdots, y_m, t_1, \cdots, t_n, t'_1, \cdots, t'_m) \\ = f_X(x_1, \cdots, x_n; t_1, \cdots, t_n) f_Y(y_1, \cdots, y_m, t'_1, \cdots, t'_m). \end{aligned} \quad (2.88)$$

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

$$\begin{aligned} C_{XY}(t_1, t_2) &= E\{[X(t_1) - m_X(t_1)] \cdot [Y(t_2) - m_Y(t_2)]\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x - m_X(t_1)] [y - m_Y(t_2)] f_{XY}(x, y; t_1, t_2) dx dy, \end{aligned} \quad (2.90)$$

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: Θ 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: Φ^{-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 τ 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

$$\rho_{XY}(\tau) = \frac{C_{XY}(\tau)}{\sqrt{C_X(0)C_Y(0)}} = \frac{R_{XY}(\tau) - m_X m_Y}{\sigma_X \sigma_Y}. \quad (2.102)$$

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, τ 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, τ 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

$$f(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n; y_1, y_2, \dots, y_n; t'_1, t'_2, \dots, t'_n) \quad (2.119)$$

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

$$R_{Z_2 Z_1}(t, t+\tau) = 0, \quad (2.134)$$

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 ω_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; Φ_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 Φ_n ($n = 1, 2, \dots, N$) 之间相互独立, 且幅度 A_n 和 Φ_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, $\Delta v_k = v_k - v_{k-1}, v_{k-1} < v'_k < v_k \cdot \Delta v_k, \lim_{\Delta v_k \rightarrow 0} \left| y_n(t) - \int_a^b h(t, v)x(v)dv \right|^2 = 0$, 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,深入地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

$$f(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \frac{1}{(2\pi)^{\frac{n}{2}} |C_X|^{\frac{1}{2}}} \exp \left[-\frac{1}{2} (X - m_X)^T C_X^{-1} (X - m_X) \right]. \quad (2.146)$$

其中 $X = [X_1, X_2, \dots, X_n]^T, m_X = E\{X\} = [E\{X_1\}, E\{X_2\}, \dots, E\{X_n\}]^T, C_X = E\{(X - m_X)(X - m_X)^T\}$, 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为遇到的、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 M_X 是 n 维期望矢量, C 是协方差矩阵。

$$M_X = \begin{bmatrix} E[X(t_1)] \\ \vdots \\ E[X(t_n)] \end{bmatrix} = \begin{bmatrix} m_X(t_1) \\ \vdots \\ m_X(t_n) \end{bmatrix}_{n \times 1}, \quad C = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix}_{n \times n}. \quad (2.147)$$

$$\begin{aligned} C_{ik} &= C_X(t_i, t_k) = E[(X(t_i) - m_X(t_i))(X(t_k) - m_X(t_k))] \\ &= R_X(t_i, t_k) - m_X(t_i)m_X(t_k). \end{aligned} \quad (2.148)$$

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 M_X 与协方差矩阵 C 所确定, 且有关时间 (t_1, \dots, t_n) 的因素, 全部包含在 M_X 和 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 τ 有关, 即满足

$$\begin{cases} m_X(t) = m_X. \\ R_X(t_i, t_k) = R_X(\tau_{k-i}), \quad \tau_{k-i} = t_k - t_i \quad (i, k = 1, 2, \dots, n, \dots). \\ E[X^2(t)] = R_X(0) < \infty. \end{cases} \quad (2.149)$$

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

$$C_u = C(t_i, t_k) = R(\tau_{k-i}) - m_X^2, i, k = 1, 2, \dots, n \quad (2.151)$$

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 Δt 平移后, 由于时间差 $\tau'_{k-i} = (t_k + \Delta t) - (t_i + \Delta t) = t_k - t_i = \tau_{k-i}$ 不随时间平移 Δt 变化, 所以过程 $X(t)$ 的 n 维概率密度也不随时间平移 Δt 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

$$\mathbf{C} = \begin{bmatrix} \sigma_X^2(t_1) & 0 & \dots & 0 \\ 0 & \sigma_X^2(t_2) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \sigma_X^2(t_n) \end{bmatrix}. \quad (2.153)$$

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

$$\begin{aligned} f_X(x_1, x_2, \dots, x_n, t_1, t_2, \dots, t_n) &= \frac{1}{(2\pi)^{n/2} \sigma_X(t_1) \sigma_X(t_2) \dots \sigma_X(t_n)} \exp \left[-\frac{1}{2} \sum_{i=1}^n \frac{(x_i - m_X(t_i))^2}{\sigma_X^2(t_i)} \right] \\ &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi} \sigma_X(t_i)} \exp \left[-\frac{(x_i - m_X(t_i))^2}{2\sigma_X^2(t_i)} \right] \\ &= f_X(x_1, t_1) f_X(x_2, t_2) \dots f_X(x_n, t_n). \end{aligned} \quad (2.155)$$

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

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{bmatrix}_{k \times 1}, \text{ 其中高斯序列矢量 } \{\mathbf{X}(n)\} \text{ 的每个分量 } \{X_i(n)\} \text{ 都是高斯序列, 它均方收}$$

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

$$Q_X(u_1, u_2, \dots, u_k) = \exp \left[j U^T M - \frac{1}{2} U^T C U \right]. \quad (2.168)$$

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 $\{t_1 + \Delta t, \dots, t_l + \Delta t, \dots, t_A + \Delta\} \in T$, 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

$$\lim_{n_j \rightarrow \infty} Y(n_j) = \lim_{\substack{n_j \rightarrow \infty \\ \Delta \lambda \rightarrow 0}} \sum_{i=0}^{n_j} X(\lambda_i) \Delta \lambda_i = \int_a^{u_j} X(\lambda) d\lambda = Y(t_j), \quad (2.171)$$

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 λ, 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

$$Y(t) = \int_{-\infty}^{\infty} X(\lambda)h(\lambda, t)d\lambda \quad (2.173)$$

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 $\mathbf{X}(t_1)$, 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 $\mathbf{X}(t) = A \cos \omega t + B \sin \omega t$, 其中 A 与 B 是相互独立的高斯变量, 且 $E[A] = E[B] = 0, E[A^2] = E[B^2] = 2, \omega$ 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

$$\begin{aligned} E[X(t)] &= E[A \cos \omega_0 t + B \sin \omega_0 t] \\ &= E[A] \cos \omega_0 t + E[B] \sin \omega_0 t = 0. \\ R_X(t, t + \tau) &= E[X(t)X(t + \tau)] \\ &= E[(A \cos \omega_0 t + B \sin \omega_0 t)(A \cos \omega_0(t + \tau) + B \sin \omega_0(t + \tau))] \\ &= E[A^2] \cos \omega_0 t \cos \omega_0(t + \tau) \\ &\quad + E[B^2] \sin \omega_0 t \sin \omega_0(t + \tau) \\ &\quad + E[AB] \cos \omega_0 t \sin \omega_0(t + \tau) \\ &\quad + E[AB] \sin \omega_0 t \cos \omega_0(t + \tau). \end{aligned} \quad (2.175)$$

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

$$f_X(x_1, x_2, \tau) = \frac{1}{2\pi\sigma^2\sqrt{1 - \cos^2 \omega_0 \tau}} \exp\left(-\frac{x_1^2 - 2x_1x_2 \cos \omega_0 \tau + x_2^2}{2\sigma^2(1 - \cos^2 \omega_0 \tau)}\right). \quad (2.180)$$

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。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。

$$\overline{x'_k(t)} = \frac{\sum_{i=1}^n x_i}{n}, \quad (2.182)$$

即以算术平均 $\overline{x'_k(t)}$ 来近似时间平均 $\overline{x_k(t)}$ 。

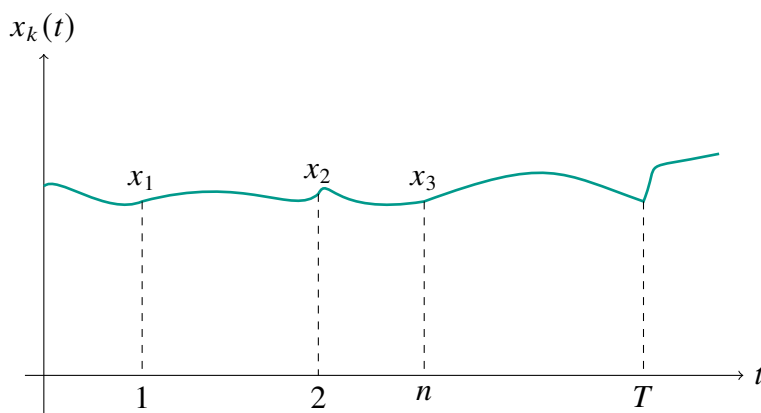
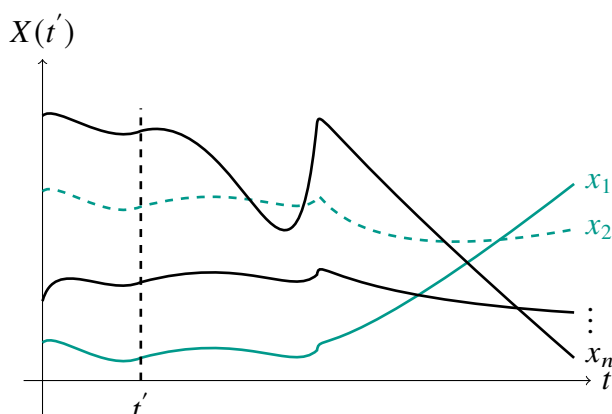


图 2-20 噪声电压

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 $\mathbf{X}(t)$ 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 $\bar{x}(t)$ 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, 在概率意义个义上

$$P\{E[X(t)] = \overline{x_k(t)}\} = 1. \quad (2.184)$$

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

$$P\{\lim_{n \rightarrow \infty} \text{各种时间平均} = \text{相应的统计平均}\} = 1. \quad (2.185)$$

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。♣

2.7.1 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

$$\overline{X(t)} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T X(t) dt = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T X(\zeta, t) dt = M(\zeta). \quad (2.187)$$

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

$$\overline{X(t, \zeta)} \stackrel{\text{a. c.}}{=} \overline{x_k(t)}. \quad (2.193)$$

上式表明: 各态历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, 代表噪声电压 (或电流) 的各态历过程 $X(t)$, 其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 $\overline{X(t)} = E[X(t)] = m_X$.

② 均方值代表噪声电压 (或电流) 消耗在 1Ω 电阻上的总平均功率, 令 $\tau = 0$, 则自相关函数有

$$R_X(0) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t)x(t+\tau)dt = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x^2(t)dt. \quad (2.196)$$

可见, $R_X(0)$ 代表噪声电压 (或电流) 消耗在 1Ω 电阻上的总平均功率。

③ 方差代表噪声电压 (或电流) 消耗在 1Ω 电阻上的交流平均功率。

$$\sigma_X^2 = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T [x(t) - m_X]^2 dt, \quad (2.197)$$

σ_X^2 代表噪声电压 (或电流) 消耗在 1Ω 电阻上的交流平均功率。标准差 σ_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, ω_0 皆为常数, 是服从 $(0, 2)$ 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

$$\begin{aligned} \overline{X(t)} &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T a \cos(\omega_0 t + \Phi) dt = \lim_{T \rightarrow \infty} \frac{a \cos \Phi \sin \omega_0 T}{\omega_0 T} = 0. \\ \overline{X(t)X(t+\tau)} &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T a \cos(\omega_0 t + \Phi) a \cos[\omega_0(t+\tau) + \Phi] dt \\ &= \frac{a^2}{2} \cos \omega_0 \tau. \end{aligned} \quad (2.198)$$

由例 2.20 的结果, 可得

$$\begin{aligned} \overline{X(t)} &= E[X(t)] = 0. \\ \overline{X(t)X(t+\tau)} &= E[X(t)X(t+\tau)] = \frac{a^2}{2} \cos \omega_0 \tau. \end{aligned} \quad (2.199)$$

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。现简单说明如下: 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 t 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 t 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, 可以证明上式 (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

$$\begin{aligned} \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^{2T} \left(1 - \frac{\tau}{2T}\right) [e^{-a|\tau|} - 0^2] d\tau &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T e^{-a\tau} \left(1 - \frac{\tau}{2T}\right) d\tau \\ &= \lim_{T \rightarrow \infty} \left[\frac{1}{aT} - \frac{1 - e^{-2aT}}{2a^2T^2} \right] = 0. \end{aligned} \quad (2.206)$$

因此, $X(t)$ 的均值具有各态历经性。

2.8 线性代数导论

2.8.1 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 2.69 数域

设它是复数集合 \mathbb{C} 的一个非空子集, 若 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 F 中, 则称 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 2.70 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 2.71 矩阵相等

若 $A, B \in \mathbb{C}^{m \times n}$, 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 2.72 矩阵转置

设 $A \in \mathbb{C}^{m \times n}$, 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 $A^T \in \mathbb{C}^{n \times m}$, $(A^T)^T = A$ 。

定义 2.73 矩阵共轭

设 $A \in \mathbb{C}^{m \times n}$, 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 $A^H \in \mathbb{C}^{n \times m}$, $(A^H)^H = A$ 。

定义 2.74 矩阵和

设 $A, B \in \mathbb{C}^{m \times n}$, A 和矩阵 $C = [a_{ij} \pm b_{ij}]$, 记作 $C = A \pm B$, 显然, $A \pm B \in \mathbb{C}^{m \times n}$ 。

定义 2.75 矩阵积

设 $A \in \mathbb{C}^{m \times s}$, $B \in \mathbb{C}^{s \times n}$, A, B 中的元素按 $\sum_{k=1}^s a_k b_{kj}$ 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

$$\begin{aligned} (AB)C &= A(BC). \\ (kB)C &= k(BC). \end{aligned} \tag{2.207}$$

(2) 分配律

$$\begin{aligned} A(B+C) &= AB+AC. \\ (A+B)C &= AC+BC. \end{aligned} \quad (2.208)$$

(3) 转置律

$$(AB)^T = B^T A^T. \quad (2.209)$$

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 2.76 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 $\text{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}$ 。

定义 2.77 行列式的阶

$$\det\{A\} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \quad (2.210)$$

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 2.78 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, $\det\{A\} = a_{11}A_{11} + a_{21}A_{21} + \cdots + a_{n1}A_{n1} = \sum_{k=1}^n a_{k1}A_{k1}$, 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 2.79

数乘变换改变行列式的值, 并且仿方阵 $A = [a_1, \dots, a_i, \dots, a_n], f_j|, B = [a_1, \dots, ka_i, \dots, a_n]$, 其中列矢量 $a_i \in \mathbb{C}^{n \times 1} (i = 1, 2, \dots, n)$, 则

$$\det\{B\} = k \det\{A\}. \quad (2.211)$$

$$\det\{kA\} = k^n \det\{A\}. \quad (2.212)$$

$$\det\{AB\} = \det\{A\} \det\{B\}, A, B \in \mathbb{R}^{n \times n}. \quad (2.213)$$

引入矩阵子式的概念, 对于矩阵 $A \in \mathbb{C}^{m \times n}$, 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 2.80 满秩

设矩阵 $A \in \mathbb{C}^{m \times n}$, 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 2.81 迹

矩阵 $A \in \mathbb{C}^{m \times n}$, 如果对于任意的非零列矢量 $x \in \mathbb{C}^n$, 有

- (1) 当 $\text{Re}\{x^H Ax\} > 0$ 时, 称矩阵 A 为正定矩阵;
- (2) 当 $\text{Re}\{x^H Ax\} \geq 0$ 时, 称矩阵 A 为半正定 (非负定) 矩阵;
- (3) 当 $\text{Re}\{x^H Ax\} < 0$ 时, 称矩阵 A 为负定矩阵。

定义 2.82

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 2.83 范数

矩阵 $A \in \mathbb{C}^{m \times n}$ 的某个行列式 $\|A\|$ 满足如下性质

- (1) A 为非零矩阵时, $\|A\| > 0$; A 为零矩阵时, $\|A\| = 0$;

(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:

$$\begin{aligned}\|\mathbf{A}\|_F &= \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right)^{1/2}. \\ \|\mathbf{A}\|_{\text{row}} &= \max_{1 \leq i \leq m} \left\{ \sum_{j=1}^n |a_{ij}| \right\}. \\ \|\mathbf{A}\|_{\text{col}} &= \max_{1 \leq i \leq n} \left\{ \sum_{i=1}^m |a_{ij}| \right\}.\end{aligned}\tag{2.214}$$

定义 2.84 矩阵的内积

对于同型矩阵 $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m \times n}$, 可以定义

$$\{\mathbf{A}, \mathbf{B}\} = \text{tr} \{ \mathbf{AB}^H + \mathbf{BA}^H \} / 2 \tag{2.215}$$

为矩阵的内积.



定理 2.85 Laurent-Moïse Schwartz 不等式

设 $\mathbf{A} \in \mathbb{C}^{m \times n}$, $\mathbf{B} \in \mathbb{C}^{n \times l}$, \mathbf{AA}^T 可逆, 则

$$\mathbf{B}^T \mathbf{B} \geq (\mathbf{AB})^T (\mathbf{AA}^T)^{-1} (\mathbf{AB}) \tag{2.216}$$

等号成立的条件为存在 $\mathbf{C} \in \mathbb{C}^{m \times l}$, $\mathbf{B} = \mathbf{A}^T \mathbf{C}$.



2.8.2 特殊矩阵

定义 2.86 矩阵

设矩阵 $\mathbf{A} \in \mathbb{C}^{m \times n}$, 如果 \mathbf{A} 在主对角线上的元素全为 1 而其他元素全为 0, 则称 \mathbf{A} 为 n 阶单位矩阵, 通常用 \mathbf{I} 表示; 如果 \mathbf{A} 在交叉对角线上的元素全为 1 而其他元素全为 0, 则称 \mathbf{A} 为 n 阶反向单位矩阵, 通常用 \mathbf{J} 表示.



定义 2.87 范德蒙矩阵

设矩阵 $\mathbf{A} \in \mathbb{C}^{m \times n}$, 如果 $\mathbf{A} = \mathbf{A}^T$, 则称 \mathbf{A} 为对称矩阵, 通常用 \mathbf{I} 表示; 如果 $\mathbf{A} = -\mathbf{A}^T$, 则称 \mathbf{A} 为反对称矩阵; 如果 $\mathbf{A} = \mathbf{A}^H$, 则称 \mathbf{A} 为共轭对称矩阵; 如果 $\mathbf{A} = -\mathbf{A}^H$, 则

称 \mathbf{A} 为反共轭对称矩阵, 也称反厄米特矩阵。



定义 2.88 范德蒙矩阵

设矩阵 $\mathbf{A} \in \mathbb{C}^{m \times n}$,

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{m-1} & x_2^{m-1} & \cdots & x_n^{m-1} \end{bmatrix}. \quad (2.217)$$

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 2.89

范德蒙矩阵 $\mathbf{V} \in \mathbb{C}^{n \times n}$, 其行列式

$$\det\{\mathbf{V}\} = \prod_{\Omega^i} (x_j - x_i) \equiv \prod_{i=1}^n \prod_{i=i+1}^n (x_j - x_i) \quad (2.218)$$



定理 2.90 汉克矩阵

设矩阵 $\mathbf{A} \in \mathbb{C}^{(n+1) \times (n+1)}$,

$$\mathbf{A} = \begin{bmatrix} a_0 & a_1 & a_2 & \cdots & a_n \\ a_1 & a_2 & a_3 & \cdots & a_{n+1} \\ a_2 & a_3 & a_4 & \cdots & a_{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_n & a_{n+1} & a_{n+2} & \cdots & a_{2n} \end{bmatrix}. \quad (2.219)$$

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 $n+1$ 行的 $2n+1$ 个元素确定。



定理 2.91 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

$$C = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ p_0 & p_1 & \cdots & p_{n-2} & p_{n-1} \end{bmatrix}. \quad (2.220)$$

使得

$$CH = HC^T \quad (2.221)$$

定义 2.92 托普利兹矩阵

$$T = \begin{bmatrix} t_0 & t_1 & t_2 & \cdots & t_{n-1} \\ t_{-1} & t_0 & t_1 & \cdots & t_{n-2} \\ t_{-2} & t_{-1} & t_0 & \cdots & t_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{-n+1} & t_{-n+2} & t_{-n+3} & \cdots & t_0 \end{bmatrix}, \quad (2.222)$$

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

$$JT^T J = T, \quad (2.223)$$

其中 J 为反向单位矩阵

简单的托普利兹矩阵有

$$F = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & 0 \end{bmatrix} \in \mathbb{R}^{n \times n}, \quad B = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 1 & 0 \end{bmatrix} \in \mathbb{R}^{n \times n}. \quad (2.224)$$

因它们作用于标准基 $\{e_1, \dots, e_n\}$ (见定义 (2.8.5)) 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 2.93 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

$$T = \sum_{k=1}^{n-1} t_{-k} B^k + \sum_{k=0}^{n-1} t_k F^k, \quad (2.225)$$

其中 F 是 B 由 1.3.22 给定.



定理 2.94

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

$$C^{(1)} = \begin{bmatrix} p_{n-1} & p_{n-2} & \cdots & p_1 & p_0 \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \quad (2.226)$$

$$C^{(2)} = \begin{bmatrix} 0 & \cdots & \cdots & 0 & p_0 \\ 1 & \ddots & & \vdots & p_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & p_{n-2} \\ 0 & \cdots & 0 & 1 & p_{n-1} \end{bmatrix}. \quad (2.227)$$

使得

$$C^{(1)}T = TC^{(2)}. \quad (2.228)$$



定理 2.95

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 2.96 托普利兹矩阵

如果满足条件 $A^H A = A A^H$, 则称 A 为正规矩阵; 如 $A^H A = A A^H = I$, 则称 A 为酉矩阵; 如果满足条件 $A^T A = A A^T = I$, 称 A 为正交矩阵.



2.8.3 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

$$AB = BA = I \quad (2.229)$$

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:

$$(1) (A^{-1})^{-1} = A;$$

$$(2) \text{ 若 } A = kI (k \neq 0), \text{ 则 } A^{-1} = k^{-1}I;$$

$$(3) \text{ 若 } A, B \text{ 都是 } n \text{ 阶可逆方阵, 则 } AB \text{ 也是可逆矩阵, 且 } (AB)^{-1} = B^{-1}A^{-1} \text{ 主}$$

定义 2.97

(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 $n \times 1$ 矢量。若 $(A + xy^H)$ 可逆, 则

$$(A + xy^H)^{-1} = A^{-1} - \frac{A^{-1}xy^HA^{-1}}{1 + y^HA^{-1}x} \quad (2.230)$$

上式可推广为矩阵之和的求逆公式:

$$\begin{aligned} (A + UBV)^{-1} &= A^{-1} - A^{-1}UB(B + BVA^{-1}UB)^{-1}BVA^{-1} \\ &= A^{-1} - A^{-1}U(I + BVA^{-1}U)^{-1}BVA^{-1} \\ &= A^{-1} - A^{-1}UBVA^{-1}(I + UBV A^{-1})^{-1}. \end{aligned} \quad (2.231)$$

或者

$$(A - UV)^{-1} = A^{-1} + A^{-1}U(I - VA^{-1}U)^{-1}VA^{-1} \quad (2.232)$$

分块矩阵的几种求逆公式:

(1) 矩阵 A 可逆时, 为

$$\begin{bmatrix} A & U \\ V & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} + A^{-1}U(D - VA^{-1}U)^{-1}VA^{-1} & -A^{-1}U(D - VA^{-1}U)^{-1} \\ -(D - VA^{-1}U)^{-1}VA^{-1} & (D - VA^{-1}U)^{-1} \end{bmatrix}. \quad (2.233)$$

2) 矩阵 A 和 D 可逆时, 为

$$\begin{bmatrix} A & U \\ V & D \end{bmatrix}^{-1} = \begin{bmatrix} (A - UD^{-1}V)^{-1} & -A^{-1}U(D - VA^{-1}U)^{-1} \\ -D^{-1}V(A - UD^{-1}V)^{-1} & (D - VA^{-1}U)^{-1} \end{bmatrix}. \quad (2.234)$$

(3) 矩阵 A 和 D 可逆时, 为

$$\begin{bmatrix} A & U \\ V & D \end{bmatrix}^{-1} = \begin{bmatrix} (A - UD^{-1}V)^{-1} & -(A - UD^{-1}V)^{-1}UD^{-1} \\ -(D - VA^{-1}U)^{-1}VA^{-1} & (D - VA^{-1}U)^{-1} \end{bmatrix}. \quad (2.235)$$

或

$$\begin{bmatrix} A & U \\ V & D \end{bmatrix}^{-1} = \begin{bmatrix} (A - UD^{-1}V)^{-1} & -(V - DU^{-1}A)^{-1} \\ (U - AV^{-1}D)^{-1} & (D - VA^{-1}U)^{-1} \end{bmatrix}. \quad (2.236)$$

$A \in \mathbb{C}^{m \times n}, G \in \mathbb{C}^{n \times m}$ 满足 Penrose - Moore 方程

$$AGA = A; \quad GAG = G; \quad (AG)^H = AG; \quad (GA)^H = GA \quad (2.237)$$

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 (2.237) 式条件的广义逆矩阵 A^\dagger 是唯一存在的, 其性质如下:

$$\begin{aligned} (A')' &= A \\ (A^H)' &= (A')^H \\ (A^T)' &= (A')^H \\ A' &= (A^H A)' A^H = A^H (A A^H)'. \end{aligned} \quad (2.238)$$

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

$$A^\dagger = C^H (CC^H)^{-1} (BB^H)^{-1} B^H. \quad (2.239)$$

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 $A^\dagger = (A A^H)^{-1} A^H$.

2.8.4 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 2.98 特征值

设 $A \in C^{n \times n}$, 如果存在数 λ 和 n 维非零列矢量 x , 使得 $Ax = \lambda x$, 则称 λ 是矩阵 A 的特征值, x 是 A 的属于特征值 λ 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 λ 是 A 的特征值, 则 λ 也是 A^T 的特征值。

b 若 λ 是 A 的特征值, 则 λ^* 是 A^H 的特征值

c. 若 λ 是 A 的特征值, 则 $\lambda + \sigma^2$ 是 $A + \sigma^2 I$ 的特征值。

d. 若 λ 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 $\text{rank} A = n - 1$.

c. 若 $\text{rank} A - \lambda I \leq n - 1$, 则 λ 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

$$\lambda_{\max} \leq \max_i \left\{ \sum_{j=1}^n a_{ij} \right\} \quad (2.240)$$

(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

$$A = E \Lambda E^{-1} \quad (2.241)$$

其中 $E = [u_1, u_2, \dots, u_n]$, $\Lambda = \text{diag} \{\lambda_1, \lambda_2, \dots, \lambda_n\}$

定理 2.99

设 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 $n \times n$ 矩阵 A 的 n 个特征值, 则

$$\begin{aligned}\sum_{i=1}^n \lambda_i &= \sum_{i=1}^n a_{ii} = \text{tr}\{A\} \\ \prod_{i=1}^n \lambda_i &= \lambda_1 \lambda_2 \cdots \lambda_n = \det\{A\}.\end{aligned}\quad (2.242)$$

n 阶方阵 A 可逆等价于 A 的 n 个特征值全不为零。

**定理 2.100**

设 λ 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 2.101 奇异值

设 $A \in \mathbb{C}^{m \times n}$, λ_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 2.102

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

$$A = U \Sigma V^H \quad (2.243)$$

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 δ_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

$$\sigma_{\max} = \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{p-1} \geq \sigma_p = \sigma_{\min} \geq 0 \quad (2.244)$$

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

$$A^H = V \Sigma^H U^H, \quad (2.245)$$

即矩阵 A 和 A^H 的具有完全相同的奇异值。

(2) P 和 Q 分别为 $m \times m$ 和 $n \times n$ 酉矩阵时, PAQ^H 的奇异值分解由

$$PAQ^H = \tilde{U}\Sigma\tilde{V}^H \quad (2.246)$$

给出。其中, $\tilde{U} = PU, \tilde{V} = QV$. 也就是说, 矩阵 PAQ^H 与 A 具有相同的奇异值, 即奇异值具有酉不变性, 但奇异矢最不同.

$A^H A, AA^H$ 的奇异值分解分别为

$$A^H A = V\Sigma^H\Sigma V^H, \quad AA^H = U\Sigma\Sigma^H U^H \quad (2.247)$$

其中

$$\begin{aligned} \Sigma^H\Sigma &= \text{diag} \left\{ \sigma_1^2, \sigma_2^2, \dots, \sigma_r^2, \overbrace{0}^{\quad}, \dots, 0 \right\} \\ \Sigma\Sigma^H &= \text{diag} \left\{ \sigma_1^2, \sigma_2^2, \dots, \sigma_r^2, \overbrace{0}^{\quad}, \dots, 0 \right\}. \end{aligned} \quad (2.248)$$

注: $A^H A, AA^H$ 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 A 的奇异值分解与 $n \times m$ 维 Moore - Penrose 广义逆矩阵 A^\dagger

$$A^\dagger = V\Sigma^\dagger U^H \quad (2.249)$$

其中 $\Sigma^\dagger = \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix}$ 。

(5) 矩阵 A 的谱范数为 A 的最大奇异值, 即

$$\|A\|_{\text{spec}} = \sigma_{\max} \quad (2.250)$$

(6) 若 $m \geq n$, 则对于矩阵 A , 有

$$\begin{aligned} \sigma_{\min}(A) &= \min_{x \neq 0} \left\{ \left(\frac{x^H A^H A x}{x^H x} \right)^{1/2} \right\} \\ &= \min_{x^H x=1} \min_{x \in \mathbb{C}^{n \times 1}} \left\{ (x^H A^H A x)^{1/2} \right\}. \end{aligned} \quad (2.251)$$

(7) $m \geq n$, 则对于矩阵 A , 有

$$\begin{aligned} \sigma_{\max}(A) &= \max_{x \neq 0} \left\{ \left(\frac{x^H A^H A x}{x^H x} \right)^{1/2} \right\} \\ &= \max_{x^H x=1} \max_{x \in \mathbb{C}^{n \times 1}} \left\{ (x^H A^H A x)^{1/2} \right\}. \end{aligned} \quad (2.252)$$

(8) 若 A 和 B 均为 $m \geq n$ 矩阵, 则对千 $1 \leq i, j \leq p, i+j \leq p+1$, 有

$$\sigma_{i+j-1}(A+B) \leq \sigma_i(A) + \sigma_j(B) \quad (2.253)$$

其中 $\sigma_i(A)$ 表示矩阵 A 的第 i 个大奇异值. 特别地, 当 $j=1$ 时, $\sigma_i(A+B) \leq \sigma_i(A) + \sigma_1(B) (i=1, 2, \dots, p)$ 成立.

(9) 对于 $m \geq n$ 矩阵 A 和 B , 有

$$\sigma_{\max}(A+B) \leq \sigma_{\max}(A) + \sigma_{\max}(B). \quad (2.254)$$

3. Cholesky 分解

设 $A \in \mathbb{C}^{n \times n}$ 是正定矩阵, $A = GAG^H$ 称为矩阵 A 的 Cholesky 分解, 其中 G

$$G = \begin{bmatrix} g_{11} & 0 & \cdots & 0 \\ g_{21} & g_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{bmatrix} \quad (2.255)$$

定理 2.103

若 $A \in \mathbb{C}^{n \times n}$ 是正定矩阵, 则 Cholesky 分解 $A = GAG^{-1}$ 是唯一的.



4. LU 分解

设 $A \in \mathbb{C}^{m \times n}$, 则 $A = LU$ 称为矩阵的 LU 分解. 其中, L 为 $m \times m$ 单位下三角矩阵 (对角线为 1 的下三角矩阵), U 是 A 的为 $m \times n$ 上阶梯型矩阵.

定理 2.104

如果 $A \in \mathbb{C}^{n \times n}$ 非奇异, 并且其 LU 分解存在, 则 A 的 LU 分解是唯一的, 且 $\det\{A\} = u_{11}u_{22} \cdots u_{nn}$.



5. QR 分解

定理 2.105

若 $A \in \mathbb{C}^{m \times n}$, 且 $m \geq n$, 则存在列正交的矩阵 $Q \in \mathbb{C}^{m \times n}$ 和上三角矩阵 $R \in \mathbb{C}^{n \times n}$ 使得 $A = QR$. 当 $m = n$ 时, Q 是正交矩阵. 若 A 是非奇异的 $n \times n$ 矩阵, 则 R 的所有对角线元素均为正, 并且 Q 和 R 二者是唯一的.



2.8.5 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

$$\begin{aligned} W &= \text{Span} \{u_1, u_2, \dots, u_m\} \\ &= \{u; u = a_1 u_1 + a_2 u_2 + \dots + a_m u_m\} \end{aligned} \quad (2.256)$$

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; 不是平凡子空间称为 V 的真子空间。

定理 2.106

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 2.107 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最组称为 W 的一组基。



定义 2.108 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 2.109 标准基

$\{e_1, e_2, \dots, e_n\}$ 是 \mathbb{R}^n 上的一组基矢量, 其中

$$e_1 = [1, 0, \dots, 0]^T, \quad e_2 = [0, 1, \dots, 0]^T, \quad \dots, \quad e_n = [0, 0, \dots, 1]^T \quad (2.257)$$

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

$\{e_1, e_2, \dots, e_n\}$ 上的标准基.

定理 2.110

$B = \{b_1, b_2, \dots, b_n\}$ 是 n 维矢量子空间 W 的一组基, 则对于 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

$$x = c_1 b_1 + c_2 b_2 + \dots + c_n b_n \quad (2.258)$$

一矢量与子空间 W 的所有矢量都正交, 则称该矢最正交子空间 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 W 正交的所有矢量的集合组成的子空间, 称为 W 的正交补空间, 记作 W^\perp , 即

$$W^\perp = \{x | x^T y = 0, \forall y \in W\}. \quad (2.259)$$

2.8.6 梯度分析

1. 实矢量的梯度算子

$n \times l$ 实矢量 x 的梯度算子定义为

$$\begin{aligned} \nabla_x &= \left[\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n} \right]^T = \frac{\partial}{\partial x} \\ \nabla_{x^T} &= \left[\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n} \right] = \frac{\partial}{\partial x^T}. \end{aligned} \quad (2.260)$$

对于 $n \times 1$ 实矢量 x , 有一实标量函数 $f(x)$, 则 $f(x)$ 相对于 x 和 x^T 的梯度分别为

$$\begin{aligned} \nabla_x f(x) &= \frac{\partial f(x)}{\partial x} = \left[\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \dots, \frac{\partial f(x)}{\partial x_n} \right]^T \\ \nabla_{x^T} f(x) &= \frac{\partial f(x)}{\partial x^T} = \left[\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \dots, \frac{\partial f(x)}{\partial x_n} \right]. \end{aligned} \quad (2.261)$$

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 x , 有一实函数矢量 $f(x) = [f_1(x), f_2(x), \dots, f_m(x)]$, 则 $f(x)$ 相

对于 \mathbf{x} 的梯度定义为

$$\nabla_{\mathbf{x}} f(\mathbf{x}) = \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \frac{\partial f_2(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial f_m(\mathbf{x})}{\partial x_1} \\ \frac{\partial f_1(\mathbf{x})}{\partial x_2} & \frac{\partial f_2(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_m(\mathbf{x})}{\partial x_2} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_1(\mathbf{x})}{\partial x_n} & \frac{\partial f_2(\mathbf{x})}{\partial x_n} & \cdots & \frac{\partial f_m(\mathbf{x})}{\partial x_n} \end{bmatrix}. \quad (2.262)$$

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

$$\nabla_{\mathbf{A}} f(\mathbf{A}) = \frac{\partial f(\mathbf{A})}{\partial \mathbf{A}} = \begin{bmatrix} \frac{\partial f(\mathbf{A})}{\partial a_{11}} & \frac{\partial f(\mathbf{A})}{\partial a_{12}} & \cdots & \frac{\partial f(\mathbf{A})}{\partial a_{1n}} \\ \frac{\partial f(\mathbf{A})}{\partial a_{21}} & \frac{\partial f(\mathbf{A})}{\partial a_{22}} & \cdots & \frac{\partial f(\mathbf{A})}{\partial a_{2n}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f(\mathbf{A})}{\partial a_{m1}} & \frac{\partial f(\mathbf{A})}{\partial a_{m2}} & \cdots & \frac{\partial f(\mathbf{A})}{\partial a_{mn}} \end{bmatrix}. \quad (2.263)$$

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

$$\frac{\partial \text{tr}\{\mathbf{W}\}}{\partial \mathbf{W}} = \mathbf{I}_m. \quad (2.264)$$

对于 $m \times n$ 可逆矩阵 \mathbf{W} , 有

$$\frac{\partial \text{tr}\{\mathbf{W}^{-1}\}}{\partial \mathbf{W}} = -(\mathbf{W}^{-2})^T. \quad (2.265)$$

对于两个列矢量的外积, 有

$$\frac{\partial \text{tr}\{\mathbf{xy}^T\}}{\partial \mathbf{x}} = \frac{\partial \text{tr}\{\mathbf{yx}^T\}}{\partial \mathbf{x}} = \mathbf{y}. \quad (2.266)$$

(2) 两个矩阵乘积的迹的梯度

对于 $\mathbf{W} \in \mathbb{R}^{m \times n}$, $\mathbf{A} \in \mathbb{R}^{n \times m}$, 有

$$\frac{\partial \text{tr}\{\mathbf{WA}\}}{\partial \mathbf{W}} = \frac{\partial \text{tr}\{\mathbf{AW}\}}{\partial \mathbf{W}} = \mathbf{A}^T \quad (2.267)$$

当 \mathbf{W} 是对称矩阵时, 有

$$\frac{\partial \text{tr}\{\mathbf{WA}\}}{\partial \mathbf{W}} = \frac{\partial \text{tr}\{\mathbf{AW}\}}{\partial \mathbf{W}} = \mathbf{A} + \mathbf{A}^T - \text{diag}\{\mathbf{A}\} \quad (2.268)$$

对于 $\mathbf{W} \in \mathbb{R}^{m \times n}$, $\mathbf{A} \in \mathbb{R}^{n \times m}$, 有

$$\frac{\partial \operatorname{tr}\{\mathbf{W}^T \mathbf{A}\}}{\partial \mathbf{W}} = \frac{\partial \operatorname{tr}\{\mathbf{A} \mathbf{W}^T\}}{\partial \mathbf{W}} = \mathbf{A} \quad (2.269)$$

对于 $\mathbf{W} \in \mathbb{R}^{m \times n}$, 有

$$\frac{\partial \operatorname{tr}\{\mathbf{W} \mathbf{W}^T\}}{\partial \mathbf{W}} = \frac{\partial \operatorname{tr}\{\mathbf{W}^T \mathbf{W}\}}{\partial \mathbf{W}} = 2\mathbf{W}. \quad (2.270)$$

对于 $\mathbf{W} \in \mathbb{R}^{m \times m}$, 有

$$\frac{\partial \operatorname{tr}\{\mathbf{W}^2\}}{\partial \mathbf{W}} = \frac{\partial \operatorname{tr}\{\mathbf{W} \mathbf{W}\}}{\partial \mathbf{W}} = 2\mathbf{W}^T. \quad (2.271)$$

对于 $\mathbf{W} \in \mathbb{R}^{m \times n}$, $\mathbf{A} \in \mathbb{R}^{n \times m}$ 且 \mathbf{W} 非奇异, 有

$$\frac{\partial \operatorname{tr}\{\mathbf{A} \mathbf{W}^{-1}\}}{\partial \mathbf{W}} = -(\mathbf{W}^{-1} \mathbf{A} \mathbf{W}^{-1})^T. \quad (2.272)$$

(3) 三个矩阵乘积的迹的梯度

对于 $\mathbf{W} \in \mathbb{R}^{m \times n}$, $\mathbf{A} \in \mathbb{R}^{m \times m}$, 有

$$\frac{\partial \operatorname{tr}\{\mathbf{W}^T \mathbf{A} \mathbf{W}\}}{\partial \mathbf{W}} = (\mathbf{A} + \mathbf{A}^T) \mathbf{W}. \quad (2.273)$$

当 \mathbf{W} 是对称矩阵时, 有

$$\frac{\partial \operatorname{tr}\{\mathbf{W}^T \mathbf{A} \mathbf{W}\}}{\partial \mathbf{W}} = 2\mathbf{A} \mathbf{W}. \quad (2.274)$$

对于 $\mathbf{W} \in \mathbb{R}^{m \times n}$, $\mathbf{A} \in \mathbb{R}^{n \times n}$, 有

$$\frac{\partial \operatorname{tr}\{\mathbf{W} \mathbf{A} \mathbf{W}^T\}}{\partial \mathbf{W}} = \mathbf{W} (\mathbf{A} + \mathbf{A}^T). \quad (2.275)$$

当 \mathbf{A} 是对称矩阵时, 有

$$\frac{\partial \operatorname{tr}\{\mathbf{W} \mathbf{A} \mathbf{W}^T\}}{\partial \mathbf{W}} = 2\mathbf{W} \mathbf{A}. \quad (2.276)$$

对于 $\mathbf{W} \in \mathbb{R}^{m \times n}$, $\mathbf{A} \in \mathbb{R}^{n \times m}$ 且 \mathbf{W} 非奇异, 有

$$\frac{\partial \operatorname{tr}\{\mathbf{A} \mathbf{W}^{-1} \mathbf{B}\}}{\partial \mathbf{W}} = -(\mathbf{W}^{-1} \mathbf{B} \mathbf{A} \mathbf{W}^{-1})^T. \quad (2.277)$$

(4) 四个矩阵乘积的迹的梯度

对于 $\mathbf{W} \in \mathbb{R}^{m \times n}, \mathbf{A} \in \mathbb{R}^{p \times m}$, 有

$$\frac{\partial \operatorname{tr} \{ \mathbf{A} \mathbf{W} \mathbf{W}^T \mathbf{A}^T \}}{\partial \mathbf{W}} = 2 \mathbf{A}^T \mathbf{A} \mathbf{W}. \quad (2.278)$$

对于 $\mathbf{W} \in \mathbb{R}^{m \times n}, \mathbf{A} \in \mathbb{R}^{p \times n}$, 有

$$\frac{\partial \operatorname{tr} \{ \mathbf{A} \mathbf{W}^T \mathbf{W} \mathbf{A}^T \}}{\partial \mathbf{W}} = 2 \mathbf{W} \mathbf{A}^T \mathbf{A}. \quad (2.279)$$

对于 $\mathbf{W} \in \mathbb{R}^{m \times n}, \mathbf{A} \in \mathbb{R}^{p \times m}, \mathbf{B} \in \mathbb{R}^{m \times p}$, 有

$$\frac{\partial \operatorname{tr} \{ \mathbf{A} \mathbf{W} \mathbf{W}^T \mathbf{B} \}}{\partial \mathbf{W}} = (\mathbf{B} \mathbf{A} + \mathbf{A}^T \mathbf{B}^T) \mathbf{W}. \quad (2.280)$$

对于 $\mathbf{W} \in \mathbb{R}^{m \times n}, \mathbf{A} \in \mathbb{R}^{p \times n}, \mathbf{B} \in \mathbb{R}^{n \times p}$, 有

$$\frac{\partial \operatorname{tr} \{ \mathbf{A} \mathbf{W}^T \mathbf{W} \mathbf{B} \}}{\partial \mathbf{W}} = \mathbf{W} (\mathbf{B} \mathbf{A} + \mathbf{A}^T \mathbf{B}^T). \quad (2.281)$$

对于 $\mathbf{W} \in \mathbb{R}^{m \times n}, \mathbf{A} \in \mathbb{R}^{n \times n}, \mathbf{B} \in \mathbb{R}^{m \times m}$, 有

$$\frac{\partial \operatorname{tr} \{ \mathbf{W} \mathbf{A}^T \mathbf{W}^T \mathbf{B} \}}{\partial \mathbf{W}} = \mathbf{B} \mathbf{W} \mathbf{A} + \mathbf{B}^T \mathbf{W} \mathbf{A}^T. \quad (2.282)$$

对于 $\mathbf{W} \in \mathbb{R}^{m \times n}, \mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times m}$, 有

$$\frac{\partial \operatorname{tr} \{ \mathbf{W} \mathbf{A} \mathbf{W} \mathbf{B} \}}{\partial \mathbf{W}} = \mathbf{B}^T \mathbf{W}^T \mathbf{A}^T + \mathbf{A}^T \mathbf{W}^T \mathbf{B}^T. \quad (2.283)$$

4. 行列式的梯度矩阵

若矩阵 \mathbf{W} 的元素 w_{ij} 相互独立,

$$\frac{\partial \mathbf{W}}{\partial w_{ij}} = \mathbf{E}_{ij} = \mathbf{e}_i \mathbf{e}_j^T \quad (2.284)$$

其中, \mathbf{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 (2.257) 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

$$\frac{\partial \mathbf{W}}{\partial w_{ij}} = \mathbf{E}_{ij} + \mathbf{E}_{ji} - \delta_{ij} \mathbf{E}_{ij} \quad (2.285)$$

$$\text{其中 } \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

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.

$$\begin{aligned} \frac{\partial \det\{\mathbf{W}\}}{\partial w_{ij}} &= \text{cof}(W_{ij}) + \text{cof}(W_{ji}) - \delta_{ij} \text{cof}(W_{ij}) \\ &= (1 - \delta_{ij}) \text{cof}(W_{ij}) + \text{cof}(W_{ji}), \end{aligned} \quad (2.286)$$

\mathbf{W} 为对称阵.

令 $i, j = 1, 2, \dots, n$ 写成矩阵形式, 有

$$\frac{\partial \det\{\mathbf{W}\}}{\partial \mathbf{W}} = (\mathbf{W}^\#)^T = \det\{\mathbf{W}\} (\mathbf{W}^{-1})^T, \quad (2.287)$$

\mathbf{W} 的元素相互独立.

$$\begin{aligned} \frac{\partial \det\{\mathbf{W}\}}{\partial \mathbf{W}} &= 2\mathbf{W}^\# - \text{diag}\{\mathbf{W}^\#\} \\ &= \det\{\mathbf{W}\} (2\mathbf{W}^{-1} - \text{diag}\{\mathbf{W}^{-1}\}). \end{aligned} \quad (2.288)$$

\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

$$\frac{\partial}{\partial \mathbf{W}} \log(\det\{\mathbf{W}\}) = \frac{1}{\det\{\mathbf{W}\}} \frac{\partial \det\{\mathbf{W}\}}{\partial \mathbf{W}} \quad (2.289)$$

将 (2.287) 式和 (2.288) 式的结果代入 (2.289) 式, 可得

$$\frac{\partial}{\partial \mathbf{W}} \log(\det\{\mathbf{W}\}) = (\mathbf{W}^{-1})^T \quad (2.290)$$

\mathbf{W} 的元素相互独立.

$$\frac{\partial}{\partial \mathbf{W}} \log(\det\{\mathbf{W}\}) = 2\mathbf{W}^{-1} - \text{diag}\{\mathbf{W}^{-1}\} \quad (2.291)$$

\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

$$\frac{\partial \det\{\mathbf{W}^{-1}\}}{\partial \mathbf{W}} = -\det^{-1}\{\mathbf{W}\} (\mathbf{W}^{-1})^T. \quad (2.292)$$

(2) 两个矩阵乘积的行列式的梯度

$$\begin{aligned}\frac{\partial \det\{W^T\}}{\partial W} &= 2 \det\{WW^T\} (WW^T)^{-1} W, \quad \text{rank}\{W_{\max n}\} = m \\ \frac{\partial \det\{W^T W\}}{\partial W} &= 2 \det\{W^T W\} W (W^T W)^{-1}, \quad \text{rank}\{W_{m \times n}\} = n. \\ \frac{\partial \det\{W^2\}}{\partial W} &= 2 \det^2\{W\} (W^{-1})^T, \quad \text{rank}\{W_{m \times m}\} = m\end{aligned}\quad (2.293)$$

(3) 三个矩阵乘积的行列式的梯度

$$\begin{aligned}\frac{\partial \det\{AWB\}}{\partial W} &= \det\{AWB\} A^T (B^T W^T A^T)^{-1} B^T \\ \frac{\partial \det\{W^T AW\}}{\partial W} &= 2AW (W^T AW)^{-1}, \quad \det\{W^T AW\} > 0 \\ \frac{\partial \det\{WAW^T\}}{\partial W} &= \left[(WAW^T)^{-1}\right]^T W (A^T + A) \\ &= 2(WAW^T)^{-1} WA, \quad A = A^T.\end{aligned}\quad (2.294)$$

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:

$$\begin{aligned}\frac{df(w)}{dw} &= \frac{1}{2} \left[\frac{\partial f(w)}{\partial w_r} - j \frac{\partial f(w)}{\partial w_i} \right] \\ \frac{df(w)}{dw^*} &= \frac{1}{2} \left[\frac{\partial f(w)}{\partial w_r} + j \frac{\partial f(w)}{\partial w_i} \right].\end{aligned}\quad (2.295)$$

那么, 目标函数 $f(w)$ 相对于复矢量 w 的梯度定义为

$$\frac{\partial f(w)}{\partial w} = \nabla_w f(w) = \left[\frac{\partial f(w)}{\partial w_1}, \frac{\partial f(w)}{\partial w_2}, \dots, \frac{\partial f(w)}{\partial w_n} \right]^T. \quad (2.296)$$

而目标函数 $f(w)$ 相对于复数共扼矢量 w^* 的梯度常简称为共扼梯度 (矢量), 定义为

$$\frac{\partial f(w)}{\partial w^*} = \nabla_{w^*} f(w) = \left[\frac{\partial f(w)}{\partial w_i}, \frac{\partial f(w)}{\partial w_2^*}, \dots, \frac{\partial f(w)}{\partial w_n^*} \right]^T. \quad (2.297)$$

若 $f(w) = [f_1(w), f_2(w), \dots, f_m(w)]$ 则 $f(w)$ 相对于复列矢量 w 的梯度为一 $n \times m$ 矩阵, 并定义为

$$\frac{\partial f(w)}{\partial w} = \begin{bmatrix} \frac{\partial f_1(w)}{\partial w_1} & \frac{\partial f_2(w)}{\partial w_1} & \dots & \frac{\partial f_m(w)}{\partial w_1} \\ \frac{\partial f_1(w)}{\partial w_2} & \frac{\partial f_2(w)}{\partial w_2} & \dots & \frac{\partial f_m(w)}{\partial w_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_1(w)}{\partial w_n} & \frac{\partial f_2(w)}{\partial w_n} & \dots & \frac{\partial f_m(w)}{\partial w_n} \end{bmatrix}. \quad (2.298)$$

类似地，行矢量函数 $f(\mathbf{w})$ 相对于复共轭列矢量 \mathbf{w}^* 的梯度称为共轭梯度矩阵，定义为

$$\frac{\partial f(\mathbf{w})}{\partial \mathbf{w}^*} = \begin{bmatrix} \frac{\partial f_1(\mathbf{w})}{\partial w_1^*} & \frac{\partial f_2(\mathbf{w})}{\partial w_2^*} & \cdots & \frac{\partial f_m(\mathbf{w})}{\partial w_m^*} \\ \frac{\partial f_1(\mathbf{w})}{\partial w_i^*} & \frac{\partial f_2(\mathbf{w})}{\partial w_i^*} & \cdots & \frac{\partial f_m(\mathbf{w})}{\partial w_i^*} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_1(\mathbf{w})}{\partial w_n^*} & \frac{\partial f_2(\mathbf{w})}{\partial w_n^*} & \cdots & \frac{\partial f_m(\mathbf{w})}{\partial w_n^*} \end{bmatrix}. \quad (2.299)$$

可得

$$\frac{\partial \mathbf{w}^T}{\partial \mathbf{w}} = \frac{\partial \mathbf{w}^H}{\partial \mathbf{w}^*} = \mathbf{I} \quad (2.300)$$

$$\frac{\partial \mathbf{w}^T}{\partial \mathbf{w}^*} = \frac{\partial \mathbf{w}^H}{\partial \mathbf{w}} = \mathbf{0}. \quad (2.301)$$

其中， \mathbf{I} 和 $\mathbf{0}$ 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 $\nabla_{\mathbf{w}^*} f(\mathbf{w})$ 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

$$\frac{\partial c}{\partial \mathbf{x}^*} = \mathbf{0} \quad (2.302)$$

(2) 若 $n \times l$ 矢量 \mathbf{a} 为与 \mathbf{x} 无关的常数矢量，则

$$\frac{\partial \mathbf{a}^H \mathbf{x}}{\partial \mathbf{x}^*} = \mathbf{0}, \quad \frac{\partial \mathbf{x}^H \mathbf{a}}{\partial \mathbf{x}^*} = \mathbf{a}. \quad (2.303)$$

(3) 若 \mathbf{A} 是一个与矢量 \mathbf{x}, \mathbf{y} 无关的矩阵，则

$$\begin{aligned} \frac{\partial \mathbf{x}^H \mathbf{A} \mathbf{x}}{\partial \mathbf{x}} &= \mathbf{A}^T \mathbf{x}^*, \quad \frac{\partial \mathbf{x}^H \mathbf{A} \mathbf{x}}{\partial \mathbf{x}^*} = \mathbf{A} \mathbf{x} \\ \frac{\partial \mathbf{x}^H \mathbf{A} \mathbf{y}}{\partial \mathbf{A}} &= \mathbf{x}^* \mathbf{y}^T, \quad \frac{\partial \mathbf{x}^H \mathbf{A} \mathbf{x}}{\partial \mathbf{A}} = \mathbf{x}^* \mathbf{x}^T. \end{aligned} \quad (2.304)$$

(4) 迹函数的共扼梯度公式

$$\frac{\partial \operatorname{tr}\{x^H y\}}{\partial x^*} = \frac{\partial \operatorname{tr}\{y x^H\}}{\partial x^*} = y \quad (2.305)$$

$$\frac{\partial \operatorname{tr}\{B A^H\}}{\partial A^*} = \frac{\partial \operatorname{tr}\{A^H B\}}{\partial A^*} = B \quad (2.306)$$

$$\frac{\partial \operatorname{tr}\{A^H\}}{\partial A^*} = I, \quad (2.307)$$

$$\frac{\partial \operatorname{tr}\{A\}}{\partial A^*} = 0 \quad (2.308)$$



$$\frac{\partial \operatorname{tr}\{A^H W A\}}{\partial A^*} = W A, \quad (2.309)$$



$$\frac{\partial \operatorname{tr}\{A W A^H\}}{\partial A^*} = A W. \quad (2.310)$$

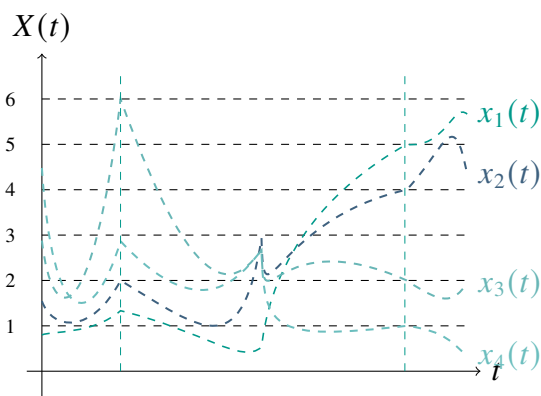
2.9 小结

随机信号和统计信号处理是机器学习和信号处理的数学基础，有了随机信号和统计信号处理构建出的思维，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。

2.10 习题

 **练习 2.1**  已知随机过程 $X(t) = A \cos \omega_0 t$ ，其中 ω_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 2-22 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 2-22 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:

| | ζ_1 | ζ_2 | ζ_3 | ζ_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 |

(2.311)

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 ω_0 为常数; 随机相位 Φ 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

$$R_X(\tau) = 2e^{-2|\tau|} \cos \omega_0\tau, R_Y(\tau) = 9 + \exp(-3\tau^2). \quad (2.312)$$

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

$$\begin{cases} X(t) = a \cos(\omega_0 t + \Phi) \\ Y(t) = b \sin(\omega_0 t + \Phi) \end{cases} \quad (2.313)$$

其中 a, b, ω_0 皆为常数, 随机相位 Φ 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

$$Z(t) = \sum_{i=1}^n A_i \exp(j\omega_i t) \quad (2.314)$$

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

$$R_{YX}(t_1, t_2) = \frac{\partial R_X(t_1, t_2)}{\partial t_1}. \quad (2.315)$$

- 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

$$R_{XY}(\tau) = \frac{dR_X(\tau)}{d\tau}, R_Y(\tau) = -\frac{d^2 R_X(\tau)}{d^2 \tau}. \quad (2.316)$$

- 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

$$Y(t) = \frac{1}{t} \int_0^t X(\lambda) d\lambda. \quad (2.317)$$

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) $R_X(\tau) = 6 \exp(-\frac{|\tau|}{2})$, (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

$$R_Y(\tau) = [R_X(0)]^2 + 2[R_X(\tau)]^2. \quad (2.318)$$

- 练习 2.22 已知随机过程 $X(t) = A \cos(\omega_0 t + \Phi)$, 其中随机相位 Φ 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 Φ 相互独立。求当 A 具备什么条件时, 过程各态历经。