

随机信号处理mooc

2021年9月27日 10:55

参考：《随机信号处理-西安电子科技大学-赵国庆》

<https://www.bilibili.com/video/BV16s411p7iX>

动态的随机过程-具有明确的时间概念

随机试验

三个特征

- 可重复性：可以在同一条件下多次进行
- 结果的可限定性：确定的结果范围
- 结果的随机性

样本空间

定义：随机试验中所有可能结果的集合称为样本空间 Ω

定义波雷尔实践域 F

Ω 中若干元素的集合称为 F ，满足下列性质

- 若 $A \in F$ ，则 $\neg A \in F$ ， $\neg A = \Omega - A$
- $\Omega \in F$
- 若 $A_i \in F$ ， $i=1,2,3,\dots$ ，则 $\cup A_i \in F$ （可加性）

F 的含义：在 Ω 中讨论是有意义的

概率空间

定义：设 $X(\omega)$ 为 Ω 中的一个函数， $\omega \in F$ ， $X(\omega) = x \in R^1$ （ ω 可以看成是一个事件， x 是随机变量），若满足：

- 非负性 $P(X) \geq 0$
 - 完备性 $P(\Omega) = 1$
 - 可列可加性 $P(\cup X_i) = \sum P(X_i)$ ，对于任意 i, j ， $X_i \cap X_j = \emptyset$
- P 为概率（测度）

(Ω, F, P) 称为概率空间

要从古典模型脱离出来，在数学上进行讨论

随机变量的概念

从随机事件到随机数值

定义：函数 $X(\omega)$ ， $\omega \in \Omega$ ， $X(\omega) \in R^1$ ， $X(\omega) = x$ ，之后的讨论就可以按照数值去讨论 x ，随机变量

- $X(\omega) = x \in R^1$ 中离散数值的集合，称之为离散的随机变量
- $x \in R^1$ 中某个区间或全部，称之为连续的随机变量，例如噪声

离散随机变量

表格法

$$\begin{pmatrix} x_i \\ P(x_i) \end{pmatrix}_{i=1}^{\infty}$$

概率密度函数/概率分布

$$\sum P(x_i) = 1, P(x_i) \geq 0$$

连续随机变量

定义: x 的分布函数 $F(x)$, 概率密度函数 $f(x)$

$F(x) = P(X \leq x)$, X — 随机变量, x — R^1 中任意实数

$F(x_1) \leq F(x_2)$, $x_2 \geq x_1$, 单调不减

$F(x) = \int_{-\infty}^x f(s) ds$ $f(x)$ 为 X 的概率密度函数

$$f(x) = \frac{\partial F(x)}{\partial x} \geq 0$$

$-\infty$ 为 0, ∞ 为 1

离散随机变量也有 $F(x) = \sum P(x_i)$, $x_i \leq x$

引入 δ 函数

$$f(x) = \sum P(x_i) \delta(x - x_i)$$

多维随机变量

$x_1, x_2 \dots x_n \in R^1$

多维概率分布

$$\Omega = \begin{pmatrix} x_1, x_2 \dots x_n \\ P(x_1 = x_1 \dots x_n = x_n) \dots \dots \end{pmatrix}$$

$F_x(x_1, x_2 \dots x_n)$ 多维

$$\frac{\partial^n F_x(x_1, x_2 \dots x_n)}{\partial x_1 \partial x_2 \dots \partial x_n} = f_x(x_1, x_2 \dots x_n) \quad \text{多维概率密度函数}$$

若 $x_1, x_2 \dots x_n$ 相互独立 $F_x(x_1, x_2 \dots x_n) = \prod_{i=1}^n F_x(x_i)$

随机变量函数的分布

以随机变量为自变量的函数 $g(x)$ 称为随机变量函数

条件: 已知随机变量 x 的分布, $y = g(x)$, y 的分布一般是随机变量

$$F(y) = P(Y \leq y) = \int_{-\infty}^y f_Y(p) dp$$

$f_Y(y)$ 概率密度函数

$$\int_{-\infty}^y f_Y(p) dp = \int_{-\infty}^x f_x(q) dq \quad \text{其中 } x = g^{-1}(y)$$

两侧对 y 求偏导

$$f_Y(y) = f_x(g^{-1}(y)) \left| \frac{\partial g^{-1}(y)}{\partial y} \right| \quad \text{雅可比变换}$$

补充: 雅可比变换 <https://blog.csdn.net/haoshu1231/article/details/116978706>

当 $z' = f(z)$ 且 q 表示分布函数时:

$$q(z') = q(z) \left| \det \frac{\partial f^{-1}}{\partial z'} \right| = q(z) \left| \det \frac{\partial f}{\partial z} \right|^{-1}, \quad (5)$$

这里补充记录下雅可比变换的数学知识。当我们知道 x 的概率分布时, 雅可比变换是一种确定变量 y 的概率分布的代数方法, 其中 y 是关于 x 的函数。首先定义:

- 变量 x 的概率密度函数为 $f(x)$, 累积分布函数为 $F(x)$;
- 变量 y 的概率密度函数为 $f(y)$, 累积分布函数为 $F(y)$;
- y 与 x 具有函数关系, 且呈单调递增

那么我们认为累积分布函数的变化是一致的:

$$dF(y) = dF(x)$$

从而有:

$$|f(y)dy| = |f(x)dx|$$

重构之后, 可以得到:

$$f(y) = \left| \frac{dx}{dy} \right| f(x) \text{ 其中, } \left| \frac{dx}{dy} \right| \text{ 就是神奇的Jacobian(雅可比行列式)}$$

使用雅可比变换求 y 的分布 (正态分布可以利用线性变换性质)

例:

已知 x 有概率密度函数 $f_x(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, $-\infty < x < \infty$, 即标准正态分布 $N(0,1)$, 若有

$y = \sigma x + a$, 求 y 的概率分布密度函数

解:

$$x = (y-a)/\sigma = g^{-1}(y)$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-a)^2}{2\sigma^2}} \quad -\infty < y < \infty, \text{ 即为 } N(a, \sigma^2)$$

Tips: 求 $g^{-1}(y)$ 时能否写成 x 的显函数

随机变量的数字特征

数学期望 (均值) $E[x]$

离散随机变量 $m_x = \sum_{i=-\infty}^{\infty} x_i P_i \quad P_i = P(x = x_i)$

连续随机变量 $m_x = \int_{-\infty}^{\infty} x f_x(x) dx$

$E[g(x)]$ 求均值的算子

性质:

a. $E[c] = c$

b. $E[kX] = k m_x$

c. $E[X+Y] = m_x + m_y$

(2, 3满足线性系统, E 算子是一个线性算子)

- a. 若X与Y独立, $E[XY]=E[X]E[Y]$

方差和标准差

方差 $D[X]=E[(x - m_x)^2]=\int_{-\infty}^{\infty}(x - m_x)^2 f_x(x)dx$ 离散情况 $\sum_{-\infty}^{\infty}(x - m_x)^2 P_i$

标准差 $\sqrt{D[X]}$

性质:

- $D[c]=0$
- $D[x]=E[X^2]-(m_x)^2$ $E[X^2]$ 总功率 $D[x]$ 交流功率(起伏功率) $(m_x)^2$ 直流功率
- $D[x] \geq 0$
- 若X与Y不相关, $D[X \pm Y]=D[X] \pm D[Y]$

协方差和相关函数

$C(X,Y)=E[(x-m_x)(y-m_y)]=\iint_{-\infty}^{\infty}(x - m_x)(y - m_y) f_{X,Y}(x,y)dxdy$

- $C(X,Y)=0$, X, Y不相关
- 若X, Y独立, 则 $C(X,Y)=0$
独立肯定不相关, 不相关不一定独立
- $C(aX,bY)=abC(X,Y)$
- 若 $C(X,X)$, 和方差一样
- $C(X,Y)=C(Y,X)$ 复随机变量不能随意交换
- 若 $C(X,c)$, 其中c为常数, 则 $C(X,c)=0$
- $C(X,Y \pm Z)=C(X,Y) \pm C(X,Z)$

相关函数 $R(X,Y)=E[XY]=\iint_{-\infty}^{\infty} xyf(x,y)dxdy$

$C(X,Y)=R(X,Y)-m_x m_y - m_y m_x + m_x m_y = R(X,Y)-m_x m_y$

所以 $R(X,Y)=C(X,Y)+m_x m_y$

相关函数比协方差函数多了各自直流量的乘积

若 $R(X,Y)=0$, 称之为正交

相关系数

对协方差函数进行归一化

$$\rho(X,Y) = \frac{C(X,Y)}{\sqrt{D(X)D(Y)}}$$

由于 $|C(X,Y)| \leq \sqrt{D(X)D(Y)}$

所以 $|\rho(X,Y)| \leq 1$

不相关 $\rho(X,Y) = 0$

若 $\rho(X,Y) = \pm 1$, 全相关

C会收到能量大小的影响, ρ 值表现两者的关系

随机变量的特征函数

特征函数是一个数学概念, 并没有对应的物理意义

复随机变量

定义：若 x, y 为实随机变量，则 $z=x+jy$ 为复随机变量

$$E[z]=E[x+jy]=\iint_{\Omega}(x+jy)f(x,y)dxdy=m_x+jm_y=m_z$$

$$D[z]=E[(z-m_z)^*(z-m_z)] \text{ 为求模的平方取共轭} = D[x]+D[y]$$

$$C[z_1, z_2] = E[(z_1 - m_{z_1})^* (z_2 - m_{z_2})]$$

$$R[z_1, z_2] = E[z_1^* z_2]$$

$$\text{不相关 } C[z_1, z_2] = 0$$

$$\text{正交 } R[z_1, z_2] = 0$$

独立 若 $f(x_1, y_1, x_2, y_2) = f(x_1, y_1)f(x_2, y_2)$ 则 z_1, z_2 相互独立

随机变量的特征函数

$$\text{定义: } \varphi_x(\lambda) = E[e^{j\lambda x}] = \int_{-\infty}^{\infty} e^{j\lambda x} f_x(x) dx$$

少了一个负号，不用 ω ，为了和傅里叶变换区分，只是为了利用其数学工具

特征函数的性质

1. $\varphi_x(\lambda)$ ，若 $\lambda = 0$ ，则 $\varphi_x(\lambda) = 1$
2. 若 $\varphi_x(\lambda)$ 已知，若有 $y = ax + b$ ， $\varphi_y(\lambda) = E[e^{j\lambda y}] = E[e^{j\lambda(ax+b)}] = e^{j\lambda b} E[e^{j\lambda ax}] = e^{j\lambda b} \varphi_x(a\lambda)$
3. 若 x, y 是独立的随机变量， $z=x+y$ ，则 $\varphi_z(\lambda) = \varphi_x(\lambda)\varphi_y(\lambda)$
 $E[e^{j\lambda z}] = E[e^{j\lambda(x+y)}] = E[e^{j\lambda x} \cdot e^{j\lambda y}] = \varphi_x(\lambda)\varphi_y(\lambda)$
4. $f_x(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi_x(\lambda) e^{-j\lambda x} d\lambda$

特征函数与矩函数的关系

矩：对随机变量本身的幂或随机变量减均值以后的幂，求概率平均，称为矩

原点矩：对随机变量本身的幂

中心矩：对随机变量减均值以后的幂

总的幂次称之为矩的阶数

均值：一阶原点矩

方差：二阶中心矩

协方差：两个变量的二阶中心矩

相关系数：二阶原点矩

$$\text{求 } x \text{ 的 } n \text{ 阶原点矩 } E[x^n] = \frac{\partial^n \varphi_x(\lambda)}{j^n \partial \lambda^n}$$

$$\text{令 } \lambda=0 \quad \frac{\partial^n}{j^n \partial \lambda^n} E[e^{j\lambda x}] = \frac{1}{j^n} E\left[\frac{\partial^n e^{j\lambda x}}{\partial \lambda^n}\right] = \frac{1}{j^n} E[(jx)^n e^{j\lambda x}] = E[x^n e^{j\lambda x}] \quad \text{带入 } \lambda=0 = E[x^n]$$

利用求导运算替代了 n 次积分运算

多维随机变量的特征函数

$$\varphi_{x_1 \cdots x_n}(\lambda_1 \cdots \lambda_n) = E[e^{j(\sum_{i=1}^n \lambda_i x_i)}] = n \text{重积分} \iint_{-\infty}^{\infty} f(x_1 \cdots x_n) e^{j(\sum_{i=1}^n \lambda_i x_i)} dx_1 \cdots dx_n$$

$$\text{若 } x_1 \cdots x_n \text{ 相互独立 } \varphi_{x_1 \cdots x_n}(\lambda_1 \cdots \lambda_n) = \prod_{i=1}^n \varphi_{x_i}(\lambda_i)$$

推广，求原点矩

记 $\sum_{i=1}^n i_j$ 为 K

$$E[x_1^{i_1} \cdots x_n^{i_n}] = \frac{\partial^K [\varphi_{x_1 \cdots x_n}(\lambda_1 \cdots \lambda_n)]}{j^K} \text{ 令所有的 } \lambda_n = 0$$

特征函数的目的：便于计算，利用傅里叶中的数学工具，解决之后的问题

切比雪夫不等式与极限定理

切比雪夫不等式

定理： $P(|x - m_x| \geq \varepsilon) \leq \frac{D[x]}{\varepsilon^2}$, $\varepsilon \in \forall R^1$

证明：

若 x 为离散随机变量，左边 = $\sum_{|x_i - m_x| \geq \varepsilon} P_i = \sum_{(x_i - m_x)^2 \geq \varepsilon^2} P_i \leq$

$$\sum_{(x_i - m_x)^2 \geq \varepsilon^2} P_i \frac{(x_i - m_x)^2}{\varepsilon^2} \leq \sum P_i \frac{(x_i - m_x)^2}{\varepsilon^2} = \frac{1}{\varepsilon^2} \sum P_i (x_i - m_x)^2 = \frac{D[x]}{\varepsilon^2}$$

告诉了某个随机变量的取值范围和方差的关系

(如果是连续变量，同理)

中心极限定理

1. 独立同分布 $x_1 \cdots x_n$, $y = \sum_{i=1}^n x_i$

的分布，当 n 足够大的时候，其概率分布趋于正态分布 $N(nm, n\sigma^2)$

2. 独立不同分布但有相同的均值方差， n 足够大， $y = \sum_{i=1}^n x_i$ 趋于正态分布 $N(nm, n\sigma^2)$

3. 二项分布 $P, (1-P)$ ，一次实验记为 x_i

可以看成定理1的一个具体应用

$$m = P, E[X^2] = P, D[x] = P(1-P)$$

正态分布为 $N(np, np(1-p))$

Chapter One 随机过程

随机过程的基本概念

定义

1. 设随机试验 E 具有样本空间 S ，对 S 的任意元素 e 都按照某种规则，确定了一个样本函数 $x(e, t)$ ，则 S 中全体元素构成的样本函数族成为随机过程（每个样本都是时间的函数）

2. 对于任意给定的时间 t ，都有一个随机变量 $x(t)$ 与之对应，则 $x(t)$ 称为随机过程

两个特点：

1. 随着时间随机变化

2. 给定时间为随机变量

(以前没有讨论随机变量和时间的关系，只是单纯数值的概念)

分类

按时域和值域中的类型来划分

离散/连续

- a. 离散随机序列 离散时域+离散值域 数字信号处理非常多
- b. 连续随机序列 连续+离散
- c. 离散随机过程 离散+连续
- d. 连续随机过程 连续+连续 之后主要讨论这部分

概率分布

一维概率分布：给定一个时间 t $x(t)$ 为随机变量, $F_x(x, t) = P(x(t) \leq x)$

n 维概率分布：在同一个时间过程中抽样了 n 次 $F_x(x_1 \cdots x_n, t_1 \cdots t_n) = P(x(t_1) \leq x_1 \cdots x(t_n) \leq x_n)$

概率密度 一维离散随机过程/序列 $\left(\frac{x_i(t)}{P_i(t)} \right)_i$ (不是分式是上下的两个值)

$$f_x(x_1 \cdots x_n, t_1 \cdots t_n) = \frac{\partial^n F(x_1 \cdots x_n, t_1 \cdots t_n)}{\partial x_1 \cdots \partial x_n}$$

独立随机过程：白噪声

矩函数

均值: $E[x(t)] = \int_{-\infty}^{\infty} x(t) f_x(x(t), t) dx(t)$

时间函数 $m_x(t) = \int_{-\infty}^{\infty} x f_x(x, t) dx$ 例如：一天内的平均温度，和所取时间有关

方差: $D[x(t)] = E[(x(t) - m_x(t))^2] = E[x(t)^2] - m_x(t)^2 = \sigma_t^2(t)$

协方差: $C_x(t_1, t_2) = E[(x(t_1) - m_x(t_1))(x(t_2) - m_x(t_2))] = E[x(t_1)x(t_2)] - m_x(t_1)m_x(t_2)$

相关函数: $R_x(t_1, t_2) = E[x(t_1)x(t_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_x(x_1, x_2, t_1, t_2) dx_1 dx_2$

将复杂的问题，简化为两个时间之间的关系

特征函数

$$\varphi_x(\lambda, t) = \int_{-\infty}^{\infty} e^{j\lambda x} f_x(x, t) dx = E[e^{j\lambda x(t)}]$$

$$\varphi_x(\lambda_1 \cdots \lambda_n, t_1 \cdots t_n) = E[e^{j(\sum_{i=1}^n \lambda_i x(t_i))}]$$

依然没有意义，用于数学计算

特征函数和矩函数的关系同样适用

平稳随机过程

定义和分类

若 $F_x(x_1 \cdots x_n, t_1 \cdots t_n) = F_x(x_1 \cdots x_n, t_1 + \varepsilon \cdots t_n + \varepsilon)$ ，对于任意 $\varepsilon \in \mathbb{R}$ ，任意 $n < \infty$ ，均满足，称 $x(t)$ 为严格平稳的随机过程/强平稳的随机过程

此时一阶矩函数和时间无关，二阶矩函数和时间的间隔长度有关

广义平稳的随机过程/弱平稳

若，对于 $\forall t$, $m_x(t) = m_x$ ，且 $R_x(t_1, t_2) \equiv R_x(t_2, t_1)$ ，为广义平稳的随机过程

推论: $C_x(t_1, t_2) = R_x(t_1, t_2) - m_x(t_1)m_x(t_2) = R_x(t_1, t_2) - m_x^2 = C_x(t_2, t_1)$

例:

噪声调幅信号 $X(t) = N(t) \cos(\omega t + \varphi)$

其中 $N(t)$ 为广义平稳的随机过程, ω 为常数, $\varphi \in [0, 2\pi)$

上均匀分布的随机变量, 证明 $X(t)$ 为平稳随机过程

若 φ 为常数, 不再是广义平稳过程

各态历经过程

定义时间平均算子 $\langle g(x(t)) \rangle = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T g(x(t)) dt$

在一个足够长的时间内, 计算均值, 即使含有随机过程

定义各态历过程

若 $X(t)$ 任意阶的矩都与其时间平均以概率1相等 (不相等的概率为0), 则称其为各态历过程

$$P(|E[x^n(t)] - \langle x^n(t) \rangle| \leq \varepsilon) = 1, \forall \varepsilon > 0$$

广义各态历经

只要求均值, 相关函数满足

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

$$E[x(t_1)x(t_1 + \tau)] = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t)x(t + \tau) dt$$

随机过程范围很大, 各态历经说明, 只需要取其中一个时间足够长的样本, 就可以代表全部

充要条件

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-\infty}^{\infty} \left(1 - \frac{|\tau|}{T}\right) C_x(\tau) d\tau = 0$$

$\lim_{\tau \rightarrow \infty} C_x(\tau) = 0$ 时间取的足够长, 就认为相距 τ 的两个时间不相关

相关函数性质

1. $R_x(\tau) = R_x(-\tau)$ 相同的间隔时间, 结果是相等的

2. $R_x(0) \geq 0$ 总平均功率

3. $|R_x(\tau)| \leq R_x(0)$

$$\text{证明: } E[(X(t) - X(t + \tau))^2] \geq 0$$

$$E[X(t)^2 + X(t + \tau)^2 - 2X(t)X(t + \tau)] \geq 0$$

$$= R_x(0) + R_x(0) - 2R_x(x) \geq 0$$

$$= 2R_x(0) - 2R_x(x) \geq 0$$

4. $\lim_{\tau \rightarrow \infty} R_x(\tau) = m_x^2$ 直流功率

5. 周期平稳的随机过程 $X(t)$

$$R_x(\tau) = R_x(\tau + nT), \text{ 其中 } n = 1, 2, 3 \dots, T \text{ 为周期}$$

$$X(t) = X(t + nT)$$

6. 非负定性

$$\sum_{i=1}^n \sum_{j=1}^n R_x(t_j - t_i) f(t_i) f(t_j) \geq 0, \text{ 对于 } \forall f, \text{ 任意函数, } \forall t_i, \forall t_j \in R^1$$

相关系数与相关时间

相关系数: $r_x(\tau) = \frac{c_x(\tau)}{c_x(0)} \in [-1, 1]$

归一化, 去除了绝对能量的影响

相关时间: $\tau_0 = \int_0^\infty r_x(\tau) d\tau$

反映变化的快慢