



概率论与随机过程

蔡远利 教授
西安交通大学自动化学院

0. Outline

- 1 概率空间与贝叶斯定理 / 3
- 2 随机向量及其分布 / 9
- 3 随机向量的数字特征 / 16
- 4 两个常用的随机分布 / 21
- 5 向量随机过程的含义 / 31
- 6 随机过程的数字特征 / 32
- 7 随机过程的平稳性 / 35

8 高斯-马尔可夫过程 / 39

9 白噪声与维纳过程 / 47

10 有色噪声仿真方法 / 53

1. 概率空间与贝叶斯定理

Definition 1.1 (σ -代数) 设 \mathcal{F} 为样本空间 Ω 的子集构成的集合, 满足

(1) $\Omega \in \mathcal{F}$;

(2) 若 $A \in \mathcal{F}$, 则 $\bar{A} \in \mathcal{F}$;

(3) 若 $A_n \in \mathcal{F}, n = 1, 2, \dots$, 则 $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ 。

那么称 \mathcal{F} 为 σ -代数, 也称为 σ -域。

Definition 1.2 (事件域) 设 Ω 是样本空间, \mathcal{F} 是由样本空间 Ω 的一些子集构成的一个 σ -代数, 则称 \mathcal{F} 为事件域。 \mathcal{F} 中的元素称为事件, Ω 称为必然事件, \emptyset 称为不可能事件。

Definition 1.3 (概率测度) 设 $P(A)$ 是定义在事件域 \mathcal{F} 上的实值集合函数, 如果满足

(1) 非负性: 对任一 $A \in \mathcal{F}$, $P(A) \geq 0$;

(2) 规范性: 对必然事件 Ω , $P(\Omega) = 1$;

(3) 可列可加性: 设 A_1, A_2, \dots 是两两互不相容事件序列, 即对于 $i \neq j$, $A_i A_j = \emptyset$ ($i, j = 1, 2, \dots$), 则有 $P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k)$ 。

那么称 $P(A)$ 为事件域 \mathcal{F} 上事件 A 的概率测度, 简称概率。

一般称三元组 (Ω, \mathcal{F}, P) 为概率空间, 其中 Ω 是样本空间, \mathcal{F} 是事件域, P 是概率。可以容易验证概率的如下性质:

(1) $P(\emptyset) = 0, P(\Omega) = 1$ 。

(2) 对 n 个两两互不相容事件 A_1, A_2, \dots, A_n , 有 $P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k)$ 。

(3) 对于任一事件 A , 有 $P(\bar{A}) = 1 - P(A)$ 。

(4) 对于任意两个事件 A, B , 有 $P(A - B) = P(A) - P(AB)$ 。特别地, 当 $B \subset A$ 时, 有 $P(A - B) = P(A) - P(B)$, 且 $P(A) \geq P(B)$ 。

(5) 对于任意两个事件 A, B , 有 $P(A \cup B) = P(A) + P(B) - P(AB)$ 。

Definition 1.4 (条件概率) 设 A 、 B 是两个事件, 且 $P(A) > 0$, 称

$$P(B|A) = \frac{P(AB)}{P(A)} \quad (1)$$

为在事件 A 已经发生的条件下事件 B 发生的条件概率。

Theorem 1.1 设 A 、 B 是两个事件, 且 $P(A) > 0$, 有

$$P(AB) = P(B|A)P(A) \quad (2)$$

若 $P(B) > 0$, 则有 $P(AB) = P(A|B)P(B)$ 。

Theorem 1.2 (全概率公式) 设 A_1, A_2, \dots, A_n 是样本空间 Ω 的一个完备事件组, 且 $P(A_i) > 0 (i = 1, 2, \dots, n)$ 。对于任意事件 B , 有

$$P(B) = \sum_{i=1}^n P(B|A_i) P(A_i) \quad (3)$$

Theorem 1.3 (贝叶斯公式) 设 A_1, A_2, \dots, A_n 是样本空间 Ω 的一个完备事件组, 且 $P(A_i) > 0 (i = 1, 2, \dots, n)$ 。对于任意事件 B , $P(B) > 0$, 则有

$$P(A_i|B) = \frac{P(B|A_i) P(A_i)}{\sum_{i=1}^n P(B|A_i) P(A_i)} \quad (4)$$

贝叶斯公式也称为**贝叶斯定理**，是概率论中最重要的定理之一。其简化版为：

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (5)$$

2. 随机向量及其分布

Definition 2.1 (随机向量) 设 $X_1(\omega), X_2(\omega), \dots, X_n(\omega)$ 是定义在概率空间 (Ω, \mathcal{F}, P) 上的 n 个随机变量, 称

$$\mathbf{X}(\omega) = [X_1(\omega), X_2(\omega), \dots, X_n(\omega)]^T \quad (6)$$

为 n 维随机向量。

n 维随机向量取值于 n 维欧几里得空间 \mathbb{R}^n 。对 n 个实数 x_1, x_2, \dots, x_n , $\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n\} = \bigcap_{i=1}^n \{X_i \leq x_i\}$ 有定义, 并属于 \mathcal{F} 。

Definition 2.2 (随机向量分布函数) 如下 n 元函数:

$$F(x_1, x_2, \dots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n) \quad (7)$$

称为 n 维随机变量 $\mathbf{X}(\omega) = [X_1(\omega), X_2(\omega), \dots, X_n(\omega)]^T$ 的 (联合) 分布函数。

Definition 2.3 (随机向量概率密度) 若对任意的 n 个实数 x_1, x_2, \dots, x_n , 存在非负实函数 $f(x_1, x_2, \dots, x_n)$, 使随机向量 \mathbf{X} 的分布函数

$$\begin{aligned} F(\mathbf{x}) &= F(x_1, x_2, \dots, x_n) \\ &= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} f(u_1, u_2, \dots, u_n) du_1 du_2 \cdots du_n \\ &= \int_{-\infty}^{\mathbf{x}} f(\mathbf{u}) d\mathbf{u} \end{aligned} \quad (8)$$

则称函数 $f(\mathbf{x}) = f(x_1, x_2, \dots, x_n)$ 为随机向量 \mathbf{X} 的概率分布密度或概率密度函数。

注意, 这里我们用 $d\mathbf{x}$ 表示 $dx_1 dx_2 \cdots dx_n$ 。

Property 2.1 随机向量概率密度 $f(\mathbf{x})$ 具有如下性质:

(1) 对 $\forall \mathbf{x} \in \mathbb{R}^n$, $f(\mathbf{x}) \geq 0$;

(2) $\int_{-\infty}^{+\infty} f(\mathbf{x}) d\mathbf{x} = \int \cdots \int_{-\infty}^{+\infty} f(x_1, x_2, \cdots, x_n) dx_1 dx_2 \cdots dx_n = 1$;

(3) 若 $F(\mathbf{x})$ 在 \mathbf{x} 处连续, 则有

$$\frac{\partial^n F(\mathbf{x})}{\partial x_1 \cdots \partial x_n} = f(x_1, \cdots, x_n) = f(\mathbf{x}) \quad (9)$$

(4) 设 V 是 \mathbb{R}^n 中任一区域, 随机点 \mathbf{x} 落入区域 V 内的概率为

$$P(\mathbf{x} \in V) = \int_V f(\mathbf{x}) d\mathbf{x} \quad (10)$$

Definition 2.4 (边缘分布) 设 n 维随机向量 \mathbf{X} 和 m 维随机向量 \mathbf{Y} 的联合分布函数为 $F_{XY}(\mathbf{x}, \mathbf{y})$, 其中 $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^m$, 那么称

$$\begin{aligned} F_X(\mathbf{x}) &= F_{XY}(\mathbf{x}, +\infty) = P(\mathbf{X} \leq \mathbf{x}, \mathbf{Y} \leq +\infty) \\ &= P(X_1 \leq x_1, \dots, X_n \leq x_n; Y_1 \leq +\infty, \dots, Y_m \leq +\infty) \end{aligned} \quad (11)$$

为随机向量 \mathbf{X} 的边缘分布函数。称

$$\begin{aligned} F_Y(\mathbf{y}) &= F_{XY}(+\infty, \mathbf{y}) = P(\mathbf{X} \leq +\infty, \mathbf{Y} \leq \mathbf{y}) \\ &= P(X_1 \leq +\infty, \dots, X_n \leq +\infty; Y_1 \leq y_1, \dots, Y_m \leq y_m) \end{aligned} \quad (12)$$

为随机向量 \mathbf{Y} 的边缘分布函数。

Definition 2.5 (边缘分布密度函数) 设随机向量 X 和随机向量 Y 的联合分布函数为 $F_{XY}(x, y)$, 对应的联合概率密度函数为 $f_{XY}(x, y)$, 即

$$F_{XY}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{XY}(u, v) du dv \quad (13)$$

称

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x, y) dy \quad (14)$$

为随机向量 X 的边缘分布密度函数。称

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x, y) dx \quad (15)$$

为随机向量 Y 的边缘分布密度函数。

Theorem 2.1 设随机变量 X, Y 的联合概率密度为 $f_{XY}(x, y)$, $(x, y) \in D$ 。令 $U = g_1(X, Y), V = g_2(X, Y)$ 。假设 $x = h_1(u, v), y = h_2(u, v)$ 对 u, v 有连续的偏导数, 并且雅可比矩阵 (Jacobian matrix)

$$J(u, v) \triangleq \frac{\partial(h_1, h_2)}{\partial(u, v)} = \begin{bmatrix} \frac{\partial h_1}{\partial u} & \frac{\partial h_1}{\partial v} \\ \frac{\partial h_2}{\partial u} & \frac{\partial h_2}{\partial v} \end{bmatrix}$$

非奇异。记

$$G = \{(u, v) | u = g_1(x, y), v = g_2(x, y), (x, y) \in D\}$$

那么

$$f_{UV}(u, v) = \begin{cases} f_{XY}(h_1(u, v), h_2(u, v)) |J(u, v)|, & (u, v) \in G; \\ 0, & \text{其他.} \end{cases}$$

其中 $|J(u, v)|$ 是雅可比矩阵的行列式。

3. 随机向量的数字特征

Definition 3.1 (数学期望) 设随机向量 \mathbf{X} 的概率密度函数为 $f(\mathbf{x})$, 如果

$$E(\mathbf{X}) = \int_{-\infty}^{+\infty} \mathbf{x} f(\mathbf{x}) d\mathbf{x} \quad (16)$$

存在, 则称 $E(\mathbf{X})$ 为随机向量 \mathbf{X} 的数学期望, 有时又称为均值。

上述定义表明, 数学期望 $E(\mathbf{X})$ 是与随机向量 \mathbf{X} 同维的常向量。经常也记为 $E\mathbf{X}, \bar{\mathbf{X}}, \mu(\mathbf{X})$ 等。数学期望表示随机量的平均值, 我们称扣除均值后的随机变量为中心随机变量, 例如 $\dot{\mathbf{X}} = \mathbf{X} - E(\mathbf{X})$ 。

Definition 3.2 (协方差矩阵) 设 \mathbf{X} 是随机向量, 若

$$E(\mathring{\mathbf{X}}\mathring{\mathbf{X}}^T) = \int_{-\infty}^{+\infty} \mathring{\mathbf{X}}\mathring{\mathbf{X}}^T f(\mathbf{x})d\mathbf{x} \quad (17)$$

存在, 则称其为随机向量 \mathbf{X} 的协方差矩阵, 记为 $\text{cov}(\mathbf{X}, \mathbf{X})$ 或 $D(\mathbf{X})$ 。

对于随机变量, 即一维随机向量, 上述协方差的概念退化为方差。随机变量 X 的方差记为 $D(X)$ 或 $\sigma^2(X)$ 。显然随机变量的方差是非负的, 其平方根称为均方差、标准差, 记为 $\sigma(X)$ 。即 $\sigma(X) = \sqrt{D(X)} = \sqrt{\sigma^2(X)}$ 。

Theorem 3.1 随机向量的协方差矩阵是对称的非负定方阵。

Definition 3.3 (互协方差矩阵) 设 \mathbf{X}, \mathbf{Y} 是两个随机向量, 若

$$E(\mathring{\mathbf{X}}\mathring{\mathbf{Y}}^T) = \int_{-\infty}^{+\infty} \mathring{\mathbf{X}}\mathring{\mathbf{Y}}^T f(\mathbf{x}, \mathbf{y}) d\mathbf{x}d\mathbf{y} \quad (18)$$

存在, 则称其为随机向量 \mathbf{X} 与 \mathbf{Y} 的互协方差矩阵, 记为 $\text{cov}(\mathbf{X}, \mathbf{Y})$ 。

两个随机变量的互协方差是一个标量, 简称为协方差, 反映了两者之间的统计关联程度。

Definition 3.4 (独立性) 设 X, Y 是两个随机向量, 如果

$$F_{XY}(\mathbf{x}, \mathbf{y}) = F_X(\mathbf{x})F_Y(\mathbf{y}) \quad (19)$$

等价地 (假设相应概率密度函数存在)

$$f_{XY}(\mathbf{x}, \mathbf{y}) = f_X(\mathbf{x})f_Y(\mathbf{y}) \quad (20)$$

那么称 X 与 Y 相互独立。

Definition 3.5 (相关性) 设 X, Y 是两个随机向量, 如果

$$E(\overset{\circ}{X}\overset{\circ}{Y}^T) = \mathbf{0} \quad (21)$$

即 $E\mathbf{X}\mathbf{Y}^T = E(\mathbf{X})E(\mathbf{Y}^T)$, 那么称 X 与 Y 不相关。

Theorem 3.2 (贝叶斯法则) 设 $f_Y(\mathbf{y}) > 0$, 有

$$f_{X|Y}(\mathbf{x}|\mathbf{y}) = \frac{f_{Y|X}(\mathbf{y}|\mathbf{x})f_X(\mathbf{x})}{f_Y(\mathbf{y})} \quad (22)$$

这是我们后面要经常用到的一个结论, 以后会称 $f_{X|Y}(\mathbf{x}|\mathbf{y})$ 为后验概率密度, 称 $f_X(\mathbf{x})$ 为先验概率密度。

4. 两个常用的随机分布

Definition 4.1 (均匀分布) 设随机变量 X 具有概率密度

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{其他} \end{cases} \quad (23)$$

则称 X 在区间 (a, b) 上服从均匀分布, 记为 $X \sim U(a, b)$ 。

Theorem 4.1 设 $X \sim U(a, b)$, 那么

$$E(X) = \frac{1}{2}(a + b) \quad (24)$$

$$\sigma^2(X) = \frac{1}{12}(b - a)^2 \quad (25)$$

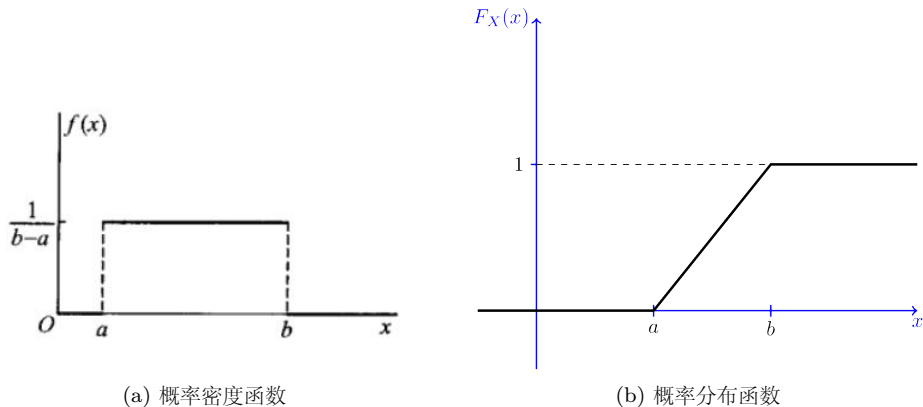


Figure 1: 均匀分布随机变量

Definition 4.2 (正态分布) 设随机变量 X 具有概率密度

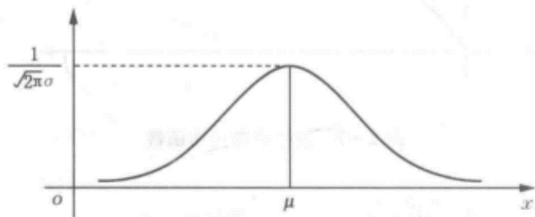
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < +\infty \quad (26)$$

其中 $\mu, \sigma (\sigma > 0)$ 为常数, 则称 X 服从参数为 μ, σ 的正态分布 (*normal distribution*) 或高斯 (*Gaussian distribution*) 分布, 记为 $X \sim \mathcal{N}(\mu, \sigma^2)$ 。

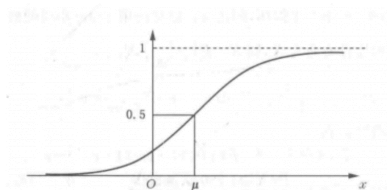
Property 4.1 设 $X \sim \mathcal{N}(\mu, \sigma^2)$, 那么 $E(X) = \mu, D(X) = \sigma^2$ 。

当 $\mu = 0, \sigma = 1$ 时, 称 X 服从标准正态分布。任何一个一般的正态分布都可以通过线性变换转化为标准正态分布。

Property 4.2 设 $X \sim \mathcal{N}(\mu, \sigma^2)$, 则 $Y = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1)$ 。



(a) 概率密度函数



(b) 概率分布函数

Figure 2: 正态分布随机变量

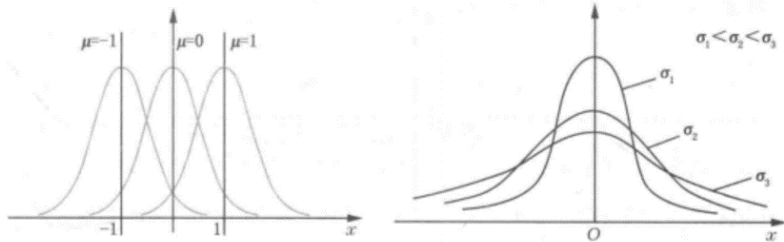


Figure 3: 正态分布均值和方差参数的物理意义

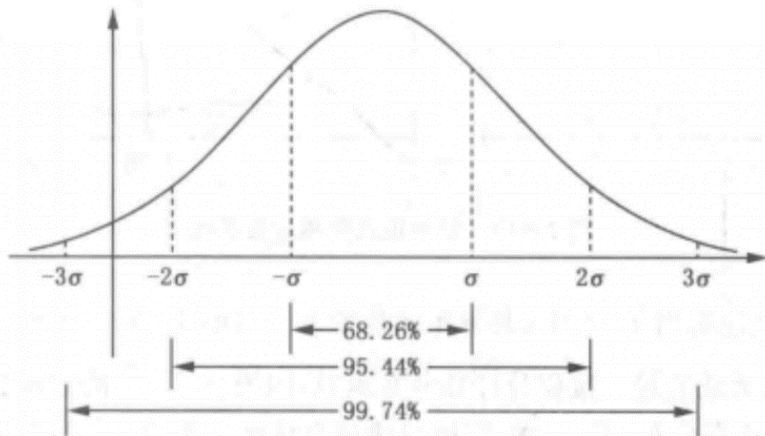


Figure 4: 正态分布方差参数与概率的关系

Definition 4.3 (高维正态分布) 设随机向量 X 的均值向量为 μ , 记

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix}.$$

如果 X 的概率密度为

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n |P|}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T P^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\} \quad (27)$$

则 X 服从正态分布, 记为 $X \sim \mathcal{N}(\boldsymbol{\mu}, P)$ 。其中 $P = \text{cov}(X, X)$, 而且 P 对称正定。

Theorem 4.2 设 X 和 Y 分别服从正态分布, 两者相互独立的充要条件是两者不相关, 即

$$f_{XY}(\mathbf{x}, \mathbf{y}) = f_X(\mathbf{x})f_Y(\mathbf{y}) \iff \text{cov}(\mathbf{X}, \mathbf{Y}) = 0.$$

对一维正态分布随机变量来说, 上述定理表明 $f_{XY}(x, y) = f_X(x)f_Y(y)$ 与 $E(XY) = E(X)E(Y)$ 相互等价。

Theorem 4.3 (中心极限定理) 设 $X^i (i = 1, 2, \dots, r)$ 是一组相互独立、同分布的 n 维随机向量, 具有有限均值 $E(X^i)$ 和协方差矩阵 P^i , 令

$$Y^r = \sum_{i=1}^r X^i, \quad Z^r = (P^r)^{-\frac{1}{2}}(Y^r - \bar{Y}^r).$$

其中, $\bar{Y}^r = \sum_{i=1}^r X^i$, $P^r = \sum_{i=1}^r P^i$ 。那么

$$\lim_{r \rightarrow +\infty} f(z^r) = \frac{1}{\sqrt{(2\pi)^n}} \exp\left\{-\frac{z^T z}{2}\right\} \quad (28)$$

上述定理说明, 当 $r \rightarrow +\infty$ 时, Z^r 趋于标准正态分布的随机向量。因此, 大量的微观上独立随机因素之和在宏观上可以用正态分布来描述。

Remark 4.1 1 到目前为止，我们分别用 X, \mathbf{X} 表示随机变量和随机向量。在不会产生歧义情况下，后面我们将分别用 x, \mathbf{x} 表示随机变量和随机向量。例如， $x \sim \mathcal{N}(\mu, \sigma^2)$ 表示服从均值为 μ 、方差为 σ^2 的正态分布随机变量 x ， $\mathbf{x} \sim \mathcal{N}(\bar{\mathbf{x}}, P_x)$ 表示服从均值为 $\bar{\mathbf{x}}$ 、协方差矩阵为 P_x 的正态分布随机向量 \mathbf{x} 。

Remark 4.2 2 如果一个随机向量 \mathbf{x} 的均值和协方差矩阵分布为 $\bar{\mathbf{x}}, P_X$ ，我们经常简记为 $\mathbf{x} \sim (\bar{\mathbf{x}}, P_X)$ 。当然，随机变量也类似地简记为 $x \sim (\bar{x}, \sigma_X^2)$ ，注意未必是正态分布。

5. 向量随机过程的含义

Definition 5.1 (随机过程) 给定概率空间 (Ω, \mathcal{F}, P) 和参数集 T , 若对于每个 $\omega \in \Omega$ 和 $t \in T$ 都有一个定义在概率空间上的随机向量 $x(\Omega, t)$ 与它对应, 则称依赖于参数 t 的随机变向量 $\{x(\Omega, t), t \in T\}$ 为 (向量) 随机过程。简单记为 $\{x(t), t \in T\}$ 或 $x(t)$ 。

6. 随机过程的数字特征

Definition 6.1 (随机过程的数学期望) 设 $\{x(t), t \in T\}$ 是一个随机过程, 如果对每一个 $t \in T$, 随机向量 $x(t)$ 的均值都存在, 则称

$$E[x(t)] = \bar{x}(t) = \int_{-\infty}^{+\infty} x f(x, t) dx \quad (29)$$

为随机过程 $\{x(t), t \in T\}$ 的数学期望或均值。也记为 $E x(t)$ 或 $m_X(t)$ 。

Definition 6.2 (随机过程的协方差) 设 $\{x(t), t \in T\}$ 是一个随机过程, 如果对每一个 $t \in T$, 随机向量 $x(t)$ 的协方差矩阵存在, 则称

$$P_X(t) = E[x(t) - \bar{x}(t)][x(t) - \bar{x}(t)]^T \quad (30)$$

为随机过程 $\{x(t), t \in T\}$ 的协方差矩阵。有时也记为 $C_X(t)$ 。

对于一维随机过程 $\{x(t), t \in T\}$, 上述协方差称为随机过程 $\{x(t), t \in T\}$ 的方差函数, 经常记为 $\sigma_X^2(t)$ 或 $D_X(t)$ 。而 $\sigma_X(t) = \sqrt{D_X(t)}$ 称为随机过程 $\{x(t), t \in T\}$ 的均方差函数。

Definition 6.3 (随机过程之间的相关性) 对任意的 $t, \tau \in T$, 若

$$E[\mathbf{x}(t)\mathbf{y}^T(\tau)] = E\mathbf{x}(t)E\mathbf{y}^T(\tau) \quad (31)$$

或等价地 $C_{XY}(t, \tau) = \mathbf{0}$, 则称随机过程 $\{\mathbf{x}(t), t \in T\}$ 和 $\{\mathbf{y}(t), t \in T\}$ 不相关。若 $R_{XY}(t, \tau) = \mathbf{0}$, 则称 $\{\mathbf{x}(t), t \in T\}$ 与 $\{\mathbf{y}(t), t \in T\}$ 正交。

Definition 6.4 (随机过程之间的独立性) 对任意的 $t, \tau \in T$, 若

$$F_{XY}[\mathbf{x}(t), \mathbf{y}(\tau)] = F_X[\mathbf{x}(t)]F_Y[\mathbf{y}(\tau)] \quad (32)$$

或者 $f_{XY}[\mathbf{x}(t), \mathbf{y}(\tau)] = f_X[\mathbf{x}(t)]f_Y[\mathbf{y}(\tau)]$, 则称随机过程 $\{\mathbf{x}(t), t \in T\}$ 与 $\{\mathbf{y}(t), t \in T\}$ 相互独立。其中 $F_{XY}(\cdot, \cdot)$ 和 $f_{XY}(\cdot, \cdot)$ 分别为联合分布函数及联合概率密度函数, $F_X(\cdot), F_Y(\cdot)$ 和 $f_X(\cdot), f_Y(\cdot)$ 分别为对应的边缘分布函数及边缘概率密度。

7. 随机过程的平稳性

Definition 7.1 (严平稳随机过程) 设 $\{\mathbf{x}(t), t \in T\}$ 是一个随机过程, 若对任意正整数 m , 任意 $t_1, t_2, \dots, t_m \in T$ 及使 $t_1 + \tau, t_2 + \tau, \dots, t_m + \tau \in T$ 的 τ , 随机向量族 $\mathbf{x}(t_1), \mathbf{x}(t_2), \dots, \mathbf{x}(t_m)$ 的联合分布函数与 $\mathbf{x}(t_1 + \tau), \mathbf{x}(t_2 + \tau), \dots, \mathbf{x}(t_m + \tau)$ 的联合分布函数满足

$$\begin{aligned} & F(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m; t_1, t_2, \dots, t_m) \\ &= F(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m; t_1 + \tau, t_2 + \tau, \dots, t_m + \tau) \end{aligned} \quad (33)$$

则称 $\{\mathbf{x}(t), t \in T\}$ 是严格平稳随机过程, 简称严平稳随机过程。

Definition 7.2 (宽平稳随机过程) 设 $\{\mathbf{x}(t), t \in T\}$ 是一个二阶矩随机过程，若

(1) $E\mathbf{x}(t) = \bar{\mathbf{x}}(t) = \mathbf{m}_X$ (不随时间变化);

(2) $E\mathbf{x}(t)\mathbf{x}^T(t+\tau) = R_X(\tau), \forall t \in T, \tau \geq 0$ 。

则称 $\{\mathbf{x}(t), t \in T\}$ 是宽平稳随机过程，简称平稳随机过程。

对宽平稳随机过程，显然有 $R_X(0) = E\mathbf{x}(t)\mathbf{x}^T(t)$, $R_X(-\tau) = R_X(\tau)$ 。
对应标量随机过程 $\{x(t), t \in T\}$, 则有 $|R_X(\tau)| \leq R_X(0)$ 。

Definition 7.3 (谱密度) 设 $R_X(\tau)$ 是平稳随机过程 $x(t)$ 的相关函数, 那么

$$\phi_X(\omega) = \int_{-\infty}^{+\infty} R_X(\tau) e^{-j\omega\tau} d\tau \quad (34)$$

称为随机过程 $x(t)$ 的谱密度。

上述谱密度在文献中也称为功率谱、功率密度谱、功率谱密度或功率密度。由傅里叶反变换公式, 可知

$$R_X(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \phi_X(\omega) e^{j\omega\tau} d\omega \quad (35)$$

式 (34) 与式 (35) 合称为维纳-辛钦 (Wiener - Khintchine) 公式。

Definition 7.4 (互谱密度) 设平稳随机过程 $x(t)$ 与 $y(t)$ 的互相关函数为 $R_{XY}(\tau)$, 称

$$\phi_{XY}(\omega) = \int_{-\infty}^{+\infty} R_{XY}(\tau) e^{-j\omega\tau} d\tau \quad (36)$$

为 $x(t)$ 与 $y(t)$ 的互谱密度。

8. 高斯-马尔可夫过程

Definition 8.1 (高斯过程) 设 $\{x(t), t \in T\}$ 是定义在概率空间 (Ω, \mathcal{F}, P) 上的 n 维向量随机过程, 对任意的 $t_1, t_2, \dots, t_m \in T$, 如果随机向量组 $\{x(t_1), x(t_2), \dots, x(t_m)\}$ 服从 $n \times m$ 维的正态分布, 则称 $\{x(t), t \in T\}$ 是高斯过程。

Definition 8.2 (马尔可夫过程) 设 $\{\mathbf{x}(t), t \in T\}$ 是概率空间 (Ω, \mathcal{F}, P) 上定义的向量随机过程, 对任意整数 n 及任意的 $t_1 < t_2 < \cdots < t_n, \forall t_i \in T$, 如果随机向量 $\mathbf{x}(t_n)$ 的条件概率密度

$$f[\mathbf{x}(t_n) | \mathbf{x}(t_{n-1}), \mathbf{x}(t_{n-2}), \cdots, \mathbf{x}(t_1)] = f[\mathbf{x}(t_n) | \mathbf{x}(t_{n-1})] \quad (37)$$

则称随机过程 $\mathbf{x}(t)$ 是马尔可夫过程。如果参数集 T 是离散时间集, 此时也称 $\mathbf{x}(t)$ 为马尔可夫链。

不难验证，马尔可夫过程有如下重要的性质：

$$\begin{aligned} & f[\mathbf{x}(t_n), \mathbf{x}(t_{n-1}), \mathbf{x}(t_{n-2}), \cdots, \mathbf{x}(t_1)] \\ &= f[\mathbf{x}(t_n) | \mathbf{x}(t_{n-1})] f[\mathbf{x}(t_{n-1}) | \mathbf{x}(t_{n-2})] \cdots f[\mathbf{x}(t_2) | \mathbf{x}(t_1)] f[\mathbf{x}(t_1)] \end{aligned} \quad (38)$$

其中的条件概率密度 $f[\mathbf{x}(t_{k+1}) | \mathbf{x}(t_k)]$ 称为马尔可夫过程 $\{\mathbf{x}(t), t \in T\}$ 的转移概率密度。

关于转移概率，由贝叶斯公式可知

$$f[\mathbf{x}(t_{k+1}) | \mathbf{x}(t_k)] = \frac{f[\mathbf{x}(t_{k+1}), \mathbf{x}(t_k)]}{f[\mathbf{x}(t_k)]} = \frac{f[\mathbf{x}(t_{k+1}), \mathbf{x}(t_k)]}{\int_{-\infty}^{+\infty} f[\mathbf{x}(t_{k+1}), \mathbf{x}(t_k)] d\mathbf{x}(t_{k+1})}$$

说明马尔可夫过程的任意有限维分布密度函数可以由二维的分布密度函数确定。

Definition 8.3 (独立随机过程) 设 $\{x(t), t \in T\}$ 是概率空间 (Ω, \mathcal{F}, P) 上定义的向量随机过程, 对任意整数 n 及任意的 $t_1 < t_2 < \cdots < t_n, \forall t_i \in T$, 如果随机向量 $x(t_1), x(t_2), \cdots, x(t_n)$ 是相互独立的, 那么称随机过程 $x(t)$ 是独立随机过程, 也称为纯随机过程。

显然独立随机过程是马尔可夫过程。

Definition 8.4 (高斯-马尔可夫过程) 如果随机过程 $\{x(t), t \in T\}$ 既是高斯过程, 又是马尔可夫过程, 则称 $\{x(t), t \in T\}$ 为高斯-马尔可夫过程。

Example 8.1 考虑离散时间随机系统

$$\mathbf{x}_{k+1} = F_k \mathbf{x}_k + \mathbf{w}_k$$

其中 $\{\mathbf{w}_k, k \in (0, 1, 2, \dots)\}$ 是零均值的高斯独立随机过程，且与高斯随机向量 \mathbf{x}_0 独立。判断 $\{\mathbf{x}_k, k \in (0, 1, 2, \dots)\}$ 是什么属性的随机过程。

Solution 8.1 因为

$$\mathbf{x}_1 = F_0 \mathbf{x}_0 + \mathbf{w}_0$$

$$\mathbf{x}_2 = F_1 \mathbf{x}_1 + \mathbf{w}_1 = F_1 F_0 \mathbf{x}_0 + F_1 \mathbf{w}_0 + \mathbf{w}_1$$

$$\mathbf{x}_3 = F_2 \mathbf{x}_2 + \mathbf{w}_2 = F_2 F_1 F_0 \mathbf{x}_0 + F_2 F_1 \mathbf{w}_0 + F_2 \mathbf{w}_1 + \mathbf{w}_2$$

$$\vdots$$

$$\mathbf{x}_{k+1} = F_k \mathbf{x}_k + \mathbf{w}_k = \prod_{i=0}^k F_i \mathbf{x}_0 + \mathbf{w}_k + \sum_{i=0}^{k-1} B_{i,k-1} \mathbf{w}_i$$

其中, $B_{i,k} = \prod_{j=i}^k F_j$ 。可见, 在时刻 k , \mathbf{x}_k 由高斯分布的 $\mathbf{x}_0, \mathbf{w}_{k-1}, \dots, \mathbf{w}_0$ 的线性组合构成, 因此随机过程 $\{\mathbf{x}(k), k = 0, 1, \dots\}$ 是高斯过程。另外, 给定 \mathbf{x}_k 下, \mathbf{x}_{k+1} 仅依赖于 \mathbf{w}_k , 而 \mathbf{w}_k 与 $\mathbf{x}_0, \mathbf{w}_{k-1}, \dots, \mathbf{w}_0$ 独立, 即 \mathbf{w}_k 与 $\mathbf{x}_{k-1}, \dots, \mathbf{x}_0$ 独立。因此, 随机过程 $\{\mathbf{x}(k), k = 0, 1, \dots\}$ 是高斯-马尔可夫

过程。

Example 8.2 考虑连续时间随机系统

$$\dot{\boldsymbol{x}}(t) = A(t)\boldsymbol{x}(t) + \boldsymbol{w}(t)$$

其中, $\{\boldsymbol{w}(t), t \geq t_0\}$ 是均值为 0 的高斯独立随机过程, 且与高斯随机向量 $\boldsymbol{x}(t_0)$ 独立。判断 $\{\boldsymbol{x}(t), t > t_0\}$ 是什么属性的随机过程。

Solution 8.2 对应任意时刻 $t > t_0$, 有

$$\mathbf{x}(t) = \Phi(t, t_0)\mathbf{x}(t_0) + \int_{t_0}^t \Phi(t, \tau)\mathbf{w}(\tau)d\tau$$

可见, $\{\mathbf{x}(t), t > t_0\}$ 是高斯分布的。同时, 对于 $t > t_1 \geq t_0$ 有

$$\mathbf{x}(t) = \Phi(t, t_1)\mathbf{x}(t_1) + \int_{t_1}^t \Phi(t, \tau)\mathbf{w}(\tau)d\tau$$

可见给定 $\mathbf{x}(t_1)$ 条件下, $\{\mathbf{x}(t), t > t_1\}$ 只依赖于 $\mathbf{w}(\tau)(t_1 \leq \tau \leq t)$, 而 $\mathbf{w}(\tau)(t_1 \leq \tau \leq t)$ 与 $\mathbf{x}(t)(t < t_1)$ 独立, 也就是说 $\{\mathbf{x}(t), t > t_1\}$ 与 $\mathbf{x}(t)(t < t_1)$ 独立, 所以 $\{\mathbf{x}(t), t > t_0\}$ 也是一个马尔可夫过程。因此, $\{\mathbf{x}(t), t > t_0\}$ 是高斯-马尔可夫过程。

9. 白噪声与维纳过程

连续时间白噪声

对于连续时间随机过程 $\mathbf{w}(t)$, 如果其谱密度 (矩阵) 为

$$\Phi(\omega) = Q \quad (39)$$

其中 Q 是非负定常值矩阵, 则称 $\mathbf{w}(t)$ 为白噪声过程。

根据维纳-辛钦公式, 可知

$$E\mathbf{w}(t)\mathbf{w}^T(t+\tau) = R_W(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} Q e^{j\omega\tau} d\omega = Q\delta(\tau) \quad (40)$$

即 $E\mathbf{w}(t)\mathbf{w}^T(t+\tau) = Q\delta(\tau)$ 。

狄拉克 δ 函数

$\delta(\cdot)$ 为连续时间单位脉冲函数, 又称为狄拉克 (Dirac) δ 函数, 可以认为是单位阶跃信号的 (广义) 导数。

对于任意函数 $f(x)$, 如果 x_0 是其连续点, 那么

$$\int_{-\infty}^{+\infty} f(x)\delta(x-x_0)dx = f(x_0) \quad (41)$$

这是狄拉克 δ 函数的一个重要性质。

离散时间白噪声

对于离散时间随机过程 $\{\mathbf{w}(k), k = 0, 1, 2, \dots\}$, 若

$$E\mathbf{w}(i)\mathbf{w}^T(j) = Q\delta_{ij}, \quad \forall i, j \geq 0 \quad (42)$$

则称 $\{\mathbf{w}(k), k = 0, 1, 2, \dots\}$ 为白噪声序列。

其中, δ_{ij} 是离散时间单位脉冲函数, 即克罗内克 (Kronecker) δ 函数, 定义为

$$\delta_{ij} = \begin{cases} 1, & i = j; \\ 0, & i \neq j. \end{cases} \quad (43)$$

Definition 9.1 (白噪声) 如果随机过程 $\{w(t), t \in T\}$ 时间集上任意的两个时间点 $t_1 \neq t_2$, 对应的随机向量 $w(t_1)$ 与 $w(t_2)$ 相互独立, 那么称 $\{w(t), t \in T\}$ 为白噪声; 否则称为有色噪声。

对于连续时间白噪声

$$Ew(t)w^T(t+\tau) = Q(t)\delta(\tau) \quad (44)$$

其中, $Q(t)$ 是随时间变化的非负定矩阵, 称为均值为零白噪声 $\{w(t), t \in T\}$ 的功率谱密度矩阵。

对于离散时间白噪声

$$Ew(i)w^T(j) = Q(i)\delta_{ij} \quad (45)$$

其中, $Q(i)$ 是随时间变化的非负定矩阵, 称为均值为零白噪声序列 $\{w(k), k = 0, 1, 2, \dots\}$ 的协方差矩阵。

Definition 9.2 (维纳过程) 设 $w(t)$ 是连续时间平稳的高斯白噪声，称

$$N(t) = \int_0^t w(\tau) d\tau \quad (46)$$

为维纳过程。

维纳过程 $N(t)$ 的基本性质：

- (1) 初值为 0，即 $N(0) = 0$;
- (2) $N(t)$ 是高斯过程；
- (3) $N(t)$ 的均值为零，即 $EN(t) = 0$;
- (4) $N(t)$ 具有平稳独立增量。

可以验证

$$R_N(t, s) = EN(t)N(s) = \sigma^2 \min\{t, s\} \quad (47)$$

$$P_N(t) = \sigma_N^2(t) = \sigma^2 t \quad (48)$$

其中 σ^2 是维纳过程 $N(t)$ 的方差参数 (常数)。

维纳过程 $N(t)$ 任意时刻的概率密度函数为

$$f(x, t) = \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{1}{2} \frac{x^2}{\sigma^2 t}} \quad (49)$$

由于维纳过程的方差随时间增加, 不是常数, 所以维纳过程是非平稳随机过程。

10. 有色噪声仿真方法

设 $x_i \sim U(0, 1)$ ，根据中心极限定理，当 N 充分大时，近似地有

$$y = \sum_{i=1}^N \frac{x_i - 0.5}{\sqrt{N/12}} \sim \mathcal{N}(0, 1) \quad (50)$$

如果取 $N = 12$ ，则

$$y = \sum_{i=1}^{12} x_i - 6 \quad \text{or} \quad y = \sum_{i=1}^6 (x_i - x_{i+5}) \quad (51)$$

假设 $r_i \sim \mathcal{N}(0, 1), i = 1, 2, \dots$ 。那么

$$\mathbf{w} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix} \sim \mathcal{N}(\mathbf{0}, I_n) \quad (52)$$

其中, I_n 表示 $n \times n$ 维单位矩阵。

一般地, 如果 $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, Q)$, 其中

$$Q = \begin{bmatrix} \sigma_1^2 & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma_n^2 \end{bmatrix} \quad (53)$$

要产生此时的随机向量 $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, Q)$, 需要进行一些分析和处理。

基本思路

- (1) 设协方差矩阵 Q 的特征值为 $\lambda_i(Q) = \mu_i^2$, $i = 1, 2, \dots, n$, 对应的特征向量为 $\mathbf{t}_i(Q)$;
- (2) 用特征向量 $\mathbf{t}_i(Q)$ 构造正交变换矩阵 T , $T^T = T^{-1}$;
- (3) 正交变换后, $\bar{Q} = T^T Q T = \text{diag}\{\mu_1^2, \mu_2^2, \dots, \mu_n^2\}$;
- (4) 定义新的随机向量 $\mathbf{v} = T^T \mathbf{w}$, 那么 \mathbf{v} 的协方差矩阵为 $\bar{Q} = T^T Q T$ 。

由于我们已经有了生成协方差矩阵是对角线矩阵随机向量的方法, 因此可以生成足够多的 \mathbf{v} 样本, 从而得到足够多的随机向量 \mathbf{w} 。

一般随机向量的生成方法

- (1) 生成独立同分布的随机数 $r_i \sim \mathcal{N}(0, 1), i = 1, 2, \dots$;
- (2) 令 $v_i = \mu_i r_i$, 这是 $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \bar{Q})$ 的一个实现;
- (3) 利用反变化 $\mathbf{w} = T\mathbf{v}$, 得到随机向量 $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, Q)$ 的一个实现;
- (4) 根据需要生成需要数量的 \mathbf{w} 样本。

Questions?