随机过程笔记

上官凝

2020 年 12 月 13 日

目录

第 1 章　引论

第 1.1 节　引言

Def 1.1.1 (严格平稳). 如果随机过程 $X(t)$ 对任意的 $t_1, \cdots, t_n \in T$ 和任何 h 有

$$(X(t_1+h), \cdots, X(t_n+h)) \stackrel{d}{=} (X(t_1), \cdots, X(t_n)),$$

则称为严格平稳的

Def 1.1.2 (宽平稳). 如果随机过程的所有二阶矩存在并有 $EX(t) = m$ 及协方差函数 $R_X(t,s)$ 只与时间差 $t-s$ 有关，则称为宽平稳的或二阶平稳的

Def 1.1.3 (独立增量过程). 如果对任意的 $t_1 < t_2 \cdots < t_n, t_1, \cdots, t_n \in T$，随机变量 $X(t_2) - X(t_1), X(t_3) - X(t_2), \cdots, X(t_n) - X(t_{n-1})$ 是相互独立的，则称 $X(t)$ 为**独立增量过程**。如果进一步有对任意的 $t_1, t_2, X(t_1+h) - X(t_1) \stackrel{d}{=} X(t_2+h) - X(t_2)$，则过程称为有平稳独立增量的过程

第 1.2 节　条件期望与矩母函数

对于连续型随机变量，定义其条件概率为：

$$P(X \in A \mid Y = y) = \lim_{\Delta y \downarrow 0} P(X \in A \mid Y \in \Delta y)$$

条件期望的表达式为：

$$E(X \mid Y = y) = \sum_x x P\{X = x \mid Y = y\} \qquad \text{(离散型随机变量 Y)}$$

$$E(X \mid Y = y) = \int x f(x \mid y)\, dx \qquad \text{连续型随机变量 Y}$$

$$E(X \mid Y = y) = \int x\, dF(x \mid y) \qquad \text{统一记作这个}$$

条件期望的性质有：

1. 若 X 和 Y 独立，则 $E(X \mid Y = y) = EX$
2. 条件期望有所谓的平滑性：

$$EX = \int E(X \mid Y = y)\, dF_Y(y) = E[E(X \mid Y)]$$

3. 对随机变量 X, Y 的函数 $\phi(X, Y)$ 恒有

$$E(\phi(X,Y) \mid Y=y) = E(\phi(X,y) \mid Y=y)$$

Def 1.2.1 (矩母函数)**.** 随机变量 X 的矩母函数定义为随机变量 $\exp tX$ 的期望，记作 $g(t)$，即

$$g(t) = E(\exp\{tx\}) = \int \exp\{tx\}\, dF(x)$$

通过矩母函数可以求出 X 的各阶矩，即有

$$E[X^n] = g^{(n)}(0), \quad n \geq 1$$

而对相互独立的随机变量 X 和 Y，他们和的矩母函数就等于其矩母函数的积：

$$g_{X+Y}(t) = g_X(t) g_Y(t)$$

Def 1.2.2 (生成函数)**.** 若 X 为离散随机变量，则期望 $E(s^X)$ 为其概率生成函数，记作 $\phi_X(s)$。特别地，若 $P(X=k) = p_k, k = 0, 1, 2, \cdots$，则 $\phi_X(s) = \sum_{k=0}^{\infty} p_k s^k$

随机和的矩母函数 (例 1.12) 记 $X_1, X_2, \cdots$ 为一串独立同分布的随机变量，N 为非负整数值随机变量且与 X 序列相互独立。Y 为随机和 $\sum_{i=1}^{N} X_i$。求 Y 的矩母函数 $g_Y(t)$

例 1.12 随机和的矩母函数. 记 $X_1, X_2, \cdots$ 为一串独立同分布的随机变量, N 为非负整数值随机变量且与 X 序列相独立. Y 为随机和 $\sum_{i=1}^{N} X_i$. 求 Y 的矩母函数 $g_Y(t)$.

解 为求 g_Y, 先算条件期望

$$\begin{aligned} E[e^{tY} \mid N=n] &= E\left[\exp\left\{t\sum_{i=1}^{N} X_i\right\} \mid N=n\right] \\ &= E\left[\exp\left\{t\sum_{i=1}^{n} X_i\right\} \mid N=n\right] \\ &= E\left[\exp\left\{t\sum_{i=1}^{n} X_i\right\}\right] = [g_X(t)]^n. \end{aligned}$$

于是有 $g_Y(t) = E[\exp\{tY\}] = E\{E[\exp\{tY\}|N]\} = E[(g_X(t))^N]$. 对 $g_Y(t)$ 关于 t 求导即有

$$\begin{aligned} g_Y'(t) &= E[N(g_X(t))^{N-1} g_X'(t)], \\ g_Y''(t) &= E[N(N-1)(g_X(t))^{N-2}(g_X'(t))^2 + N(g_X(t))^{N-1} g_X''(t)]. \end{aligned}$$

将 $t=0$ 代入上面两式得

$$\begin{aligned} EY &= E[NE(X)] = EN \cdot EX, \\ EY^2 &= EN \cdot \mathrm{Var}X + EN^2 \cdot E^2X, \qquad (1.20) \\ \mathrm{Var}Y &= EN \cdot \mathrm{Var}X + E^2X \cdot \mathrm{Var}N. \end{aligned}$$

第 2 章　Poisson 过程

本章的符号:

1. X_n：第 $n-1$ 次与第 n 次事件间的间隔时间
2. $W_n = \sum_{i=1}^{n} X_i$ 为第 n 次事件的到达或等待时间

一定要记住，Poisson 的 Poisson 是对于**区间上的事件发生**而言的。

第 2.1 节　Poisson 过程

Poisson 过程有如下两个性质：一是在时间或空间上的均匀性，二是未来的变化与过去的变化没有联系

Def 2.1.1 (Poisson 过程)**.** 一个整数值随机变量 $\{N(t), t \geq 0\}$ 满足下述三个条件就称作强度为 $\lambda > 0$ 的 Poisson 过程:

1. $N(0) = 0$
2. $N(t)$ 是独立增量过程（其定义如1.1.3所示）
3. 对任何 $t > 0, s \geq 0$，增量 $N(s+t) - N(s)$ 服从参数为 λt 对 Poisson 分布，即

$$P\{N(s+t) - N(s) = k\} = \frac{(\lambda t)^k \exp\{-\lambda t\}}{k!}, k = 0, 1, \cdots$$

Def 2.1.2 (Poisson 过程的假定)**.** 对于 Poisson 过程，我们有如下四条**充分必要的**假定（注意和后面其变种的区别）:

1. 在不相交区间中事件发生的数目相互独立，也即对任何整数 $n = 1, 2, \cdots$，设时刻 $t_0 = 0 < t_1 < t_2 < \cdots < t_n$，增量 $N(t_1) - N(t_0), N(t_2) - N(t_1), \cdots, N(t_n) - N(t_{n-1})$ 相互独立。为前后的独立性，说明试验是独立的
2. 对任何时刻 t 和正数 h，随机变量（增量）$N(t+h) - N(t)$ 对分布只依赖于区间长度 h 而不依赖于时刻 t。为时间上的均匀性或齐次性，说明在每个长度相同的小区间上事件有着相同的概率 p

3. 存在正常数 λ，当 $h \downarrow 0$ 时，在长度为 h 的小区间中事件至少发生一次的概率为

$$P\{N(t+h)-N(t)\geq 1\}=\lambda h+o(h)$$

事件是稀有的，其发生概率为 $p=\lambda h$，而且 p 很小

4. 在长度为 h 的小区间上发生两个或两个以上事件的概率为 $o(h)$，即

$$\lim_{h\downarrow 0} P\{N(t+h)-N(t)\geq 2\}=o(h)$$

为相继性，指事件是一件一件地发生的，在同一时间同时发生多个事件的可能性很小很小，而不发生的概率为 $1-\lambda h=1-p$

第 2.2 节　与 Poisson 过程相联系的若干分布

下面这个样本累计图对理解和掌握 Poisson 过程十分有用：

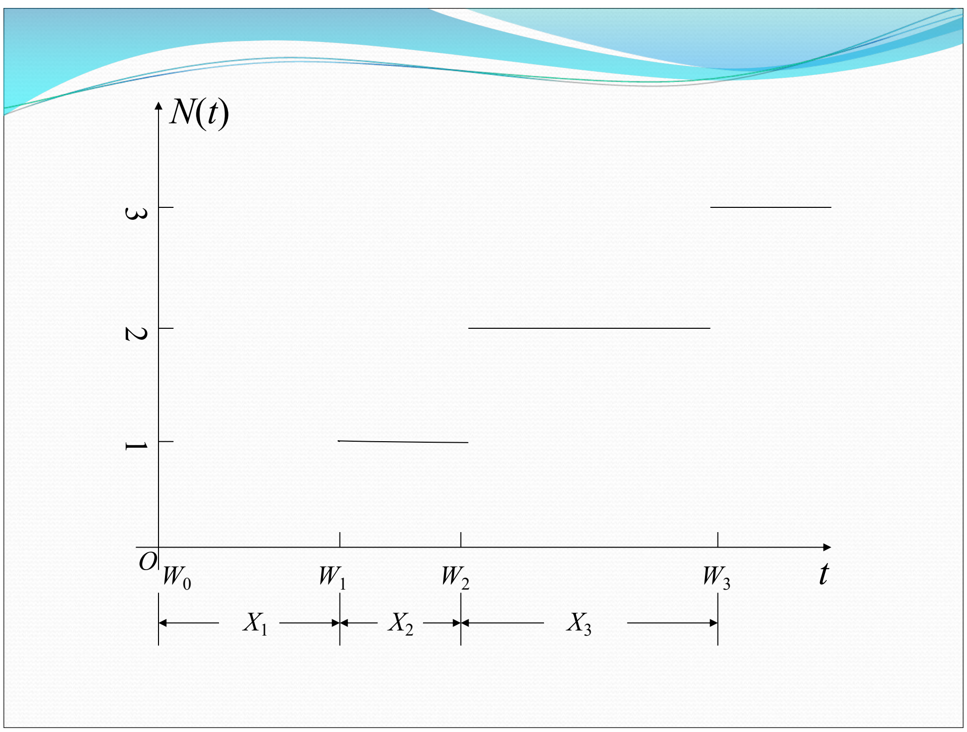

图 2.1: Poisson 过程的样本路径

Thm 2.2.1 (一条很重要的性质).

$$\boxed{P\{W_n\leq t\}=P\{N(t)\geq n\}=\sum_{j=n}^{\infty}e^{-\lambda t}\frac{(\lambda t)^j}{j!}}$$

即事件 $\{N(t)\geq n\}$ 是与 $W_n\leq t$ 等价的，它们都表明第 n 次事件发生在时刻 t 之前，或者换言之，到时刻 t 已经发生了 n 件事

$$X_1, X_2, \cdots, X_n \; i.i.d \sim exp(\frac{1}{\lambda})$$
$$W_n \sim \Gamma(n, \lambda)$$
$$f_{W_n}(t) = \lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}$$

Thm 2.2.2. 若 $N(t), t \geq 0$ 为 Poisson 过程，则给定 $N(t) = n$ 下等待时间 $W_1, \cdots, W_n$ 的联合密度为

$$f_{W_1,\cdots,W_n|N(t)=n}(\omega_1, \cdots, \omega_n \mid n) = \frac{n!}{t^n}, \; 0 < W_1 < \cdots < W_n \leq t$$

第 2.3 节　Poisson 过程的推广

2.3.1　非齐次 Poisson 过程

当允许强度 λ（即事件在某一小区间上的概率与区间长度的比例因子）依赖于时刻 t，这提供了过程增量不平稳的例子，此时，

$$P\{N(t+h) - N(t) = k\} = \frac{(\int_t^{t+h} \lambda(u)\, du)^k \exp(\int_t^{t+h} \lambda(u)\, du)}{k!}, k = 0, 1, \cdots$$

非齐次 Poisson 过程的假定只需要将2.1.2中第二条去掉，并将 (3) 换为

$$P\{N(t+h) - N(t) \geq 1\} = \lambda(t)h + o(h)$$

2.3.2　复合 Poisson 过程

累计值过程 $X(t) = \sum_{i=1}^{N(t)} Y_i$，其中 Y_i 为独立同分布的随机变量，有分布函数 $G(y)$，期望 $EY = \mu$，方差 $VarY = \tau^2$。$N(t)$ 是参数为 λ 的 Poisson 过程。则 $X(t)$ 就是复合 Poisson 过程，他是随机和并且 $E[X(t)] = \lambda\mu t, Var[X(t)] = \lambda(\tau^2 + \mu^2)t$

2.3.3　更新过程

将时间间隔 X_i 服从的指数分布改为一般的分布函数 $F(x)$，就得到所谓的更新过程

第 3 章　Markov 过程

独立随机实验模型最直接的推广就是 Markov 链模型。粗略而言，一个随机过程如果给定了当前时刻 t 的值 X_t，未来 $X_s(s>t)$ 的值不受过去的值 $X_u(u<t)$ 的影响就称为有 Markov 性。

记号：

1. 当 $X_n = i$，就称过程在时间 n 处于状态 i
2. 转移概率矩阵 $P=(P_{ij})$，矩阵的第 $i+1$ 行就是给定 $X_n = i$ 时，X_{n+1} 的概率分布
3. 一步转移概率 $P_{ij}^{n,n+1}$ 和一步转移概率矩阵 $P^{n,n+1}$（当此矩阵与 n 无关时成为上面的 P）
4. n 步转移概率和 n 步转移概率矩阵 $P^{(n)}$

第 3.1 节　Markov 链的定义和例子

Def 3.1.1 (Markov 性质)**.** 如果对任何一列状态 $i_0, i_1, \cdots, i_{n-1}, i_n, j$，以及对任何 $N \geq 0$，随机过程 $\{X_n, n \geq 0\}$ 满足 Markov 性质

$$P\{X_{n+1}=j \mid X_0=i_0, X_1=i_1, \cdots, X_{n-1}=i_{n-1}, X_n=i\} = P\{X_{n+1}=j \mid X_n=i\}$$

则称 X_n 为离散时间 Markov 链

Def 3.1.2 (平稳转移概率)**.** 平稳转移概率 $P_{ij}^{n,n+1}$ 为从（X_n，状态 i）到（X_{n+1}，状态 j）的转移概率，当此概率与 n 无关时记为 P_{ij}，并称为平稳转移概率

Def 3.1.3 (n 步转移概率矩阵)**.** n 步转移概率矩阵 $\mathbf{P}^{(n)}$

$$\mathbf{P}^{(n)} = \mathbf{P}^n = \sum_{k=0}^{\infty} P_{ik} P_{kj}^{(n-1)}$$

$$\mathbf{P}^{(n+m)} = \sum_{k=0}^{\infty} P_{ik}^{(n)} P_{kj}^{(m)}$$

Thm 3.1.1 (求 n 步转移概率矩阵)**.** n 步转移概率矩阵 $P^{(n)} = P^n$

$$P\{X_0=i_0, X_1=i_1, \cdots, X_n=i_n\} = p_i P_{i_0,i_1} \cdots P_{i_{n-1},i_n}$$

这个结论在直接计算的时候很平凡，但是在证明的时候不一定会想起来

3.1.1 求矩阵的 n 次方

由线性代数的知识可以知道，实对称矩阵一定可以对角化[1]。下面对一个 n 阶的实对称矩阵 A 进行对角化，即求正交矩阵 T，使得 $T^{-1}AT$ 为对角矩阵：[2]

1. 求出矩阵 A 的特征多项式 $\det(\lambda I - A) = f_n(\lambda)$ 为一个关于 λ 的 n 次多项式
2. 对于 n 个特征值 λ_i，分别求解 $(\lambda_i I - A)x_1 = 0$
3. 将 x_i 分别单位化，得到 e_i（均为列向量）
4. 将 e_i 自左向右合为一个矩阵，即为 T

第 3.2 节 Markov 链的状态分类

Def 3.2.1 (可达与互达)**.** 如果对某一 $n \geq 0$，有 $P^{(}_{ij}n) > 0$，则称状态 j 是从状态 i 可达的，记作 $i \to j$。他表示从状态 i 经过有限步的转移可以到达状态 j。两个互相可达的状态 i 和 j 称为互达的，记作 $i \leftrightarrow j$

互达性是等价关系。

如果在互达性这一等价关系下 Markov 链的所有状态都居于同一类，那么称这个 Markov 链是不可约的

Def 3.2.2 (Markov 链的周期)**.** 设 i 为 Markov 链的一个状态，使 $P_{ii}^{(n)} > 0$ 的所有正整数 $n(n \geq 1)$ 的最大公约数称作是状态 i 的周期，记作 $d(i)$。如果对所有 $n \geq 1$，都有 $P_{ii}^{(n)} = 0$ 则约定周期为 ∞；周期为 1 的状态非周期的

如果 $i \leftrightarrow j$，那么 $d(i) = d(j)$

Thm 3.2.1. 如果状态 i 有周期 $d(i)$，那么存在整数 N，使得对所有 $n > N$ 恒有 $P_{ii}^{(nd(i))} > 0$

$P_{ji}^{(m+nd(i))} \geq P_{ji}^{(m)} P_{ii}^{(nd(i))}$

如果 $P_{ji}^{(m)} > 0$，则存在正整数 N，使得对所有 $n > N$ 恒有 $P_{ji}^{(m+nd(i))} > 0$

Thm 3.2.2. 令 P 为不可约、非周期、有限状态 Markov 链的转移矩阵，则必存在 N，使得当 $n \geq N$ 时，n 步转移概率阵 $P^{(n)}$ 的所有元素都非零

Def 3.2.3 (常反和瞬过)**.** $f_{ij}^{(n)}$：从 i 出发在 n 步转移时首次到达 j 的概率

f_{ij}：从 i 出发最终转入状态 j 的概率

如果 $j_{ii} = 1$（也等价于 $\sum_{n=1}^{\infty} P_{ii}^{(n)} = \infty$），称状态 i 是常反的

Def 3.2.4 (常反时、正常反和零常反)**.** 对常反状态 i，定义 T_i 为首次返回状态 i 的时刻，称作常反时。记 $\mu_i = ET_i$，则有

$$\mu_i = \sum_{n=1}^{\infty} n f_{ii}^{(n)}$$

一个常反状态 i 当且仅当 $\mu_i = \infty$ 时称为零常反的，当且仅当 $\mu_i < \infty$ 时称为正常反的

[1]《线性代数与解析几何第二版》陈发来等著，P176 §6.4

[2]《线性代数与解析几何第二版》陈发来等著，P208 §7.3 例 7.3.1

3.2.1 可达、周期、常反的判断

可达性、等价类个数可以由 Markov 链状态转换图直接看出来。周期利用同一个等价类中所有状态周期一样的性质，找一个最好算的。常反可以用 $\sum f_{ii}^{(n)}$，找规律来计算，然后利用同一等价类中是否瞬过相同。正常反、零常反、瞬过也可以用下一节第一个定理来判断

第 3.3 节 Markov 链的极限定理和平稳分布

第 3.4 节 分支过程

第 3.5 节 连续时间 Markov 链

Def 3.5.1. 若对所有 $s,t \geq 0$ 和任何非负整数 $i,j,x(u),0 \leq u \leq s$，随机过程 $X(t),t \geq 0$，满足

$$P\{X(t+s)=j \mid X(s)=i, X(x)=x(u), 0 \leq u < s\} = P\{X(t+s)=j \mid X(s)=i\}$$

连续时间 Markov 链的转移概率 $P_{ij}(i)$ 和 p_i

第 4 章　平稳过程

平稳过程 $X = \{X(t), t \in T\}$ 是其概率性质在时间平移下不变的随机过程

第 4.1 节　定义和例子

回顾一下第一章对于严平稳（1.1.1）和宽平稳（1.1.2）的定义。

设 T 是具有以下性质的下标集合（对加法封闭）：若 $t_1, t_2 \in T$，则 $t_1 + t_2 \in T$。通常取以下几种集合之一：

(i) $T = \{0, 1, 2, \cdots\}$;

(ii) $T = \{0, \pm 1, \pm 2, \cdots\}$;

(iii) $T = \{t : t \geq 0\}$;

(iv) $T = \{t : -\infty < t < \infty\}$.

Def 4.1.1 (严平稳过程)**.** 设 $X = \{X(t), t \in T\}$ 为一随机过程，若对任意正整数 k 及 T 中任意 k 个时刻 $t_1 < \cdots < t_k \in T$ 和任何 $h \in T$ 有

$$(X(t_1 + h), \cdots, X(t_k + h)) \stackrel{d}{=} (X(t_1), \cdots, X(t_k)),$$

则随机过程 X 称为严格平稳过程。

如果 $T = \{0, \pm 1, \pm 2, \cdots\}$，我们一般把 X 称为随机序列。如果 X 还是严平稳的，则称为严平稳序列。

Thm 4.1.1. 设 $X = \{X(t), t \in T\}$ 为严平稳过程，则：

(a) $m(t) = EX(t) = m$（期望，是常数若存在）

(b) $Var(X(t)) = E(X(t) - m)^2 = \sigma^2$（方差，是常数若存在）

(c) $E(X(t) - m)(X(s) - m) = E(X(t - s) - m)(X(0) - m) \stackrel{h=t-s}{=} R(h)$（协方差函数，仅与时间差有关）

(d) $Var(X(t)) = R(0)$

(e) $r(\tau) = EX(t)X(t + \tau)$（自相关函数，与起点 t 无关）

(f) $\rho(v) = R(v)/\sigma^2 = R(v)/R(0)$（标准自相关函数，$\rho(0) = 1, | \rho(v) | \leq 1$）

严平稳过程要求所有有限维分布都与起点无关

Def 4.1.2 (宽平稳过程)**.** 设 $X = \{X(t), t \in T\}$ 为一实值随机过程，如果对 $\forall t \in T, EX^2(t) < \infty, EX(t) = m$ 以及协方差函数 $E(X(t) - m)(X(s) - m)$ 仅与 $t - s$ 有关，则称 X 为宽平稳随机过程

Def 4.1.3 (Gauss 过程)**.** 设 $G = \{G(t), -\infty < t < \infty\}$ 为一随机过程，如果对任一正整数 k 以及 k 个时刻 $t_i \leq t_2 \leq \cdots \leq t_k$，有 $(G(t_1), G(t_2), \cdots, G(t_k))$ 的联合分布为 k 维正态分布，则称 G 为 Gauss 过程