第 2 章　可数马尔可夫链

2.1　引言

在本章中，我们考虑状态空间为可数无限集的（时齐）马尔可夫链．如果一个集合与非负整数集 $\{0,1,2,\cdots\}$ 一一对应，就称它为可数无限集．例如，这样的集合有：$\mathbf{Z}$，所有整数构成的集合；$2\mathbf{Z}$，偶数集；$\mathbf{Z}^2$，平面上的格点集合，即

$$\mathbf{Z}^2=\{(i,j):i,j\text{ 均为整数}\}.$$

（希望读者思考 $\mathbf{Z}^2$ 与 $\{0,1,2,\cdots\}$ 是如何一一对应的．）并不是所有的无限集合都是可数无限的，例如，实数集就不能和正整数集一一对应．

我们同样令 X_n 表示一个马尔可夫链．对有限状态空间马尔可夫链描述的一些结论同样适用于无限的情形，不过，有些时候情况会变得相对复杂些．我们同样可以考虑转移矩阵，但是这种情形下，它为无限的矩阵．所以在这里，我们并不使用矩阵符号表示它，只是简单地写出转移概率，即

$$p(x,y)=\mathbf{P}\{X_1=y\mid X_0=x\},\quad x,y\in S.$$

其中转移概率均为非负，并且每“行”的概率之和为 1，即对于每个 $x\in S$，有

$$\sum_{y\in S}p(x,y)=1.$$

我们用 x，y，z 表示状态空间 S 中的元素，同样定义 n 步转移概率为

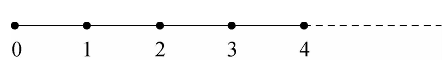
$$p_n(x,y)=\mathbf{P}\{X_n=y\mid X_0=x\}.$$ 43

若 $0<m$，$n<\infty$，那么有

$$\begin{aligned}p_{m+n}(x,y)&=\mathbf{P}\{X_{m+n}=y\mid X_0=x\}\\&=\sum_{z\in S}\mathbf{P}\{X_{m+n}=y,X_m=z\mid X_0=x\}\\&=\sum_{z\in S}p_m(x,z)p_n(z,y).\end{aligned}$$

这个方程有时也称为查普曼-柯尔莫戈洛夫方程．它可以被看做无限矩阵的矩阵乘法定义．

例 1　**在 0 点处带部分反射壁的随机游动．** 令 $0<p<1$，$S=\{0,1,2,\cdots\}$．



转移概率由如下形式给出：

$$p(x,x-1)=1-p,\quad p(x,x+1)=p,\quad x>0,$$

以及

$$p(0,0)=1-p,\quad p(0,1)=p.$$

例 2　**整格点上的简单随机游动．** 令 $\mathbf{Z}^d$ 为 d 维整数格点集，即

$$\mathbf{Z}^d=\{(z_1,\cdots,z_d):z_i\in\mathbf{Z}\}.$$

注意到 $\mathbf{Z}^d$ 中的每个元素 x 在 $\mathbf{Z}^d$ 中都有 $2d$ 个与之距离为 1 的“最近邻”．$\mathbf{Z}^d$ 上的简单随机游

动是指过程 X_n，它在 $\mathbf{Z}^d$ 中取值，并且它每次移动都尽可能地到其当前位置的 $2d$ 个最近邻之一的位置. 更精确地，即它表示状态空间为 $S=\mathbf{Z}^d$ 的马尔可夫链，且

$$p(x,y)=\begin{cases}\dfrac{1}{2d}, & 若\ |x-y|=1,\\ 0, & 其他.\end{cases}$$

例 3 **排队模型.** 令 X_n 表示在某服务系统中等待的顾客数. 我们认为队列中的第一个顾客得到服务而其他顾客则排队等待，在每个时间间隔内一位新顾客到达的概率为 p，第一个顾客的
44 服务结束并且顾客离开队列的可能性为 q. 我们并不限制排队等候的顾客数. 这就是一个马尔可夫链，其状态空间为 $\{0,1,2,\cdots\}$，转移概率（见 1.1 节中的例 2）为

$$p(x,x-1)=q(1-p),\quad p(x,x)=qp+(1-q)(1-p),$$
$$p(x,x+1)=p(1-q),\quad x>0;$$
$$p(0,0)=1-p,\quad p(0,1)=p.$$

在有限马尔可夫链中，我们的目标是理解极限行为. 在有限马尔可夫链中的一些概念同样适合于无限情形. 例如，互通类的概念. 同样，若一个马尔可夫链的所有状态是互通的，则称它为不可约的. 本章讨论的所有例子除了少数情形外都是不可约的，这里的少数情形指除了吸收状态 x 外的所有状态都是互通的，即 $p(x, x)=1$. 我们同样也可以讨论不可约马尔可夫链的周期. 上面的例 1 和例 3 都是非周期的，而例 2 中马尔可夫链的周期为 2. 然而无限状态空间下的不可约、非周期马尔可夫链并不是总收敛到平稳的概率分布.

2.2 常返和非常返

假定 X_n 为一个不可约的马尔可夫链，其状态空间 S 为可数无限的，转移概率为 $p(x, y)$，若对于每一个状态 x 都有

$$\mathbf{P}\{有无限多个\ n,使得\ X_n=x\}=1$$

成立，则称 X_n 为常返（recurrent）链，即此马尔可夫链可以返回到 x 无限多次. 若一个不可约马尔可夫链可以返回到固定状态 x 无限多次，那么它也可以访问其他状态无限多次.（根本原因在于，若 y 为另一状态，则存在一个正的概率从 x 可以到达 y. 如果该马尔可夫链可以访问 x 无限多次，如此反复，那么它也将到达 y 无限多次. 即若某特定事件发生的概率为正，那么当我们进行无限次试验时，此事件将会发生无限次.）如果该马尔可夫链不是常返的，则返回到每个状态的次数是有限的. 在这种情形下，称该马尔可夫链为非常返的. 对于一个给定的马尔可夫链，并不容易判断其是常返的还是非常返的. 在本节中，我们将给出两条准则来判定
45 这一问题.

固定某一状态 x，并且假定 $X_0=x$，考虑随机变量 R 为访问 x 的总的次数，包括在 0 时刻的访问. R 可以表示为

$$R=\sum_{n=0}^{\infty}I\{X_n=x\},$$

这里同样用 I 表示示性函数，即当事件发生时，$I=1$，其他情况为 0. 若考虑的马尔可夫链为常返的，则 R 恒等于无穷大；若所考虑的马尔可夫链为非常返的，则 R 以概率 1 小于∞. 假

定 $X_0=x$，我们可以计算 R 的期望

$$\mathbf{E}(R)=\mathbf{E}\sum_{n=0}^{\infty}I\{X_n=x\}=\sum_{n=0}^{\infty}\mathbf{P}\{X_n=x\}=\sum_{n=0}^{\infty}p_n(x,x).$$

现在用另一种方式计算 $\mathbf{E}(R)$. 设 T 为首次返回到 x 的时间，

$$T=\min\{n>0:X_n=x\}.$$

若所考虑的马尔可夫链永远不返回到 x，则 $T=\infty$. 假定 $\mathbf{P}\{T<\infty\}=1$，那么该马尔可夫链以概率 1 经常返回到状态 x，如此继续，我们可以知道该马尔可夫链无限次返回到 x 的概率为 1，则该马尔可夫链为常返的. 假定 $\mathbf{P}\{T<\infty\}=q<1$，我们根据 q 来计算 R 的分布. 首先 $R=1$ 当且仅当所考虑的马尔可夫链从来不返回到 x，因此 $\mathbf{P}\{R=1\}=1-q$. 若 $m>1$，则 $R=m$ 当且仅当该马尔可夫链返回到状态 x 为 $m-1$ 次且不会返回第 m 次. 因此 $\mathbf{P}\{R=m\}=q^{m-1}(1-q)$. 所以，在非常返情形中，$q<1$，

$$\mathbf{E}(R)=\sum_{m=1}^{\infty}m\mathbf{P}\{R=m\}=\sum_{m=1}^{\infty}mq^{m-1}(1-q)=\frac{1}{1-q}<\infty.$$

我们得到以下结论：

结论 一个不可约马尔可夫链是非常返的当且仅当返回到某一状态的期望值为有限值，即

$$\sum_{n=0}^{\infty}p_n(x,x)<\infty.$$

例 $\mathbf{Z}^d$ **上的简单随机游动**. 首先取定 $d=1$，并考虑取值为整数的马尔可夫链，转移概率为

$$p(x,x+1)=p(x,x-1)=\frac{1}{2}.$$ 46

我们集中考虑 $x=0$ 的状态，并且假定 $X_0=0$. 由于此马尔可夫链的周期为 2，那么对于奇数 n，有 $p_n(0,0)=0$. 我们接下来要写下 $p_{2n}(0,0)$ 的准确表达式. 假定质点在 $2n$ 步后处于位置 0，则质点必须是向右走了 n 步，向左走了 n 步. 任意长度为 $2n$ 且 n 步向右和 n 步向左的“路径”都具有相同结果.

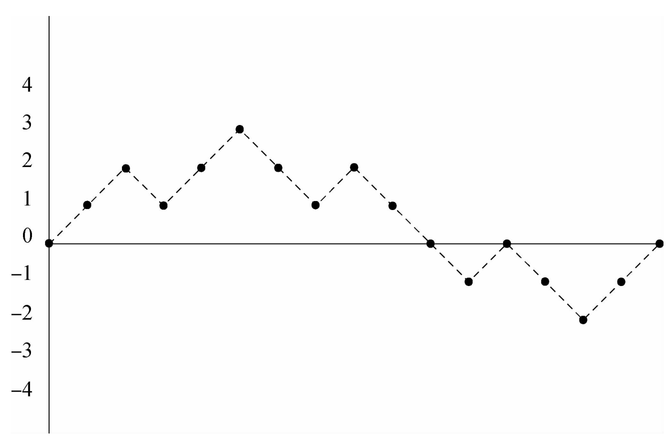


图 2.1　16 步后到达原点的随机游动路径图

既然它包含了 $2n$ 个发生概率为$\frac{1}{2}$的事件，则每个这样的路径的发生概率为$\left(\frac{1}{2}\right)^{2n}$. 从 $2n$ 步中选择 n 步向右，剩下的 n 步向左，共有$\binom{2n}{n}$种不同的方式，因此，

$$p_{2n}(0,0)=\binom{2n}{n}\left(\frac{1}{2}\right)^{2n}=\frac{(2n)!}{n!n!}\left(\frac{1}{2}\right)^{2n}.$$

当 n 很大时，并不容易看出 $p_{2n}(0, 0)$ 的表现形式. 然而，我们可以利用斯特林（Stirling）公式估计阶乘. 斯特林公式（见习题 2.18）为

$$n!\sim\sqrt{2\pi n}n^n\mathrm{e}^{-n},$$

47 其中～意味着当 n 趋于∞时，它两边的比率接近 1. 若将此式代入上面的表达式，则得

$$p_{2n}(0,0)\sim\frac{1}{\sqrt{\pi n}}. \tag{2.1}$$

特别地，因为$\sum n^{-\frac{1}{2}}=\infty$，所以

$$\sum_{n=0}^{\infty}p_{2n}(0,0)=\infty,$$

一维上的简单随机游动是常返的.

现在讨论 $d>1$，此时该马尔可夫链的状态空间为 d 维整数格点 $\mathbf{Z}^d$ 且转移概率为

$$p(x,y)=\frac{1}{2d},\quad |x-y|=1.$$

同样令游动起始位置为 $0=(0,\cdots,0)$. 我们尝试得到 $p_{2n}(0, 0)$ 的渐近表达式（同样，对奇数 n，有 $p_n(0, 0)=0$）. 在此种情形下，组合数有些复杂，所以我们将仅仅给出简要的推导过程. 假定质点移动了 $2n$ 步. 依据大数定律，对于非常大的 n，我们期望在每一维上移动的步数为$\frac{2n}{d}$. 若有机会在移动 n 步后到达 0，则需要在每一维上的移动步数都为偶数. 对大的 n，此事件发生的概率大约为$\left(\frac{1}{2}\right)^{d-1}$（在前 $d-1$ 维的每一维上是否移动了偶数步几乎是独立事件，然
48 而，我们知道若前 $d-1$ 维移动了偶数步，因为总的移动步数为偶数，所以最后一维上的移动也是偶数步.）在每一维上，若移动了大约$\frac{2n}{d}$步，则由（2.1）我们预期每一维到达 0 点的概率大约为$\left(\frac{n\pi}{d}\right)^{-\frac{1}{2}}$. 结合此结果，我们可得渐近表达式

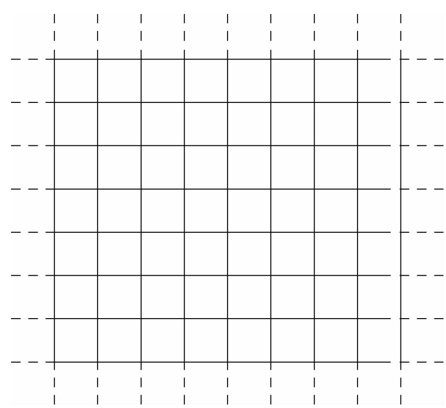


图 2.2　$\mathbf{Z}^2$ 格点

$$p_{2n}(0,0)\sim\left(\frac{1}{2}\right)^{d-1}\left(\frac{d}{n\pi}\right)^{\frac{d}{2}}.$$

回忆一下，当且仅当 $a>1$ 时$\sum n^{-a}<\infty$. 因此

$$\sum_{n=0}^{\infty} p_{2n}(0,0) = \begin{cases} = \infty & d = 1,2, \\ < \infty & d \geqslant 3. \end{cases}$$

我们已推导出如下结论.

结论 当 $d=1$ 或 2 时，$\mathbf{Z}^d$ 上的简单随机游动为常返的，当 $d\geqslant 3$ 时为非常返的.

现在考虑另一种判定常返或非常返的方法. 假设 X_n 是一个不可约的马尔可夫链，考虑一个固定的状态，我们标记为 z. 对于每个状态 x，令

$$\alpha(x) = \mathbf{P}\{X_n = z, \text{对某些 } n \geqslant 0 \mid X_0 = x\}.$$

显然 $\alpha(z)=1$. 若该马尔可夫链为常返的，则对所有的 x 都有 $\alpha(x)=1$. 然而，若该马尔可夫链是非常返的，则必存在 x，使得 $\alpha(x)<1$. 尽管没有那么明显，但事实上若该马尔可夫链为非常返的，那么无论怎样减小 $\alpha(x)$，总存在距离 z "越来越远"的点 x.

若 $x\neq z$，则

$$\begin{aligned} \alpha(x) &= \mathbf{P}\{X_n = z, \text{对某些 } n \geqslant 0 \mid X_0 = x\} \\ &= \mathbf{P}\{X_n = z, \text{对某些 } n \geqslant 1 \mid X_0 = x\} \\ &= \sum_{y\in S} \mathbf{P}\{X_1 = y \mid X_0 = x\}\mathbf{P}\{X_n = z, \text{对某些 } n \geqslant 1 \mid X_1 = y\} \\ &= \sum_{y\in S} p(x,y)\alpha(y). \end{aligned}$$

总之，$\alpha(x)$ 满足如下条件：

$$0 \leqslant \alpha(x) \leqslant 1, \tag{2.2}$$

$$\alpha(z) = 1, \quad \inf\{\alpha(x): x \in S\} = 0, \tag{2.3}$$

以及

$$\alpha(x) = \sum_{y\in S} p(x,y)\alpha(y), \quad x \neq z. \tag{2.4}$$ 49

这表明若 X_n 是非常返的，则（2.2）～（2.4）存在唯一解，且对应于适当的概率. 另外，可证明若（2.2）～（2.4）无解，则 X_n 是常返的（我们将在 5.5 节例 5 中证明此结论）. 这给出了判定常返或非常返的另一种方法.

结论 一个不可约的马尔可夫链是非常返的当且仅当对任意的 z 都能找到满足（2.2）～（2.4）的函数 $\alpha(x)$.

例 考虑前一节中的例 1，即带有部分反射壁的随机游动. 令 $z=0$，尝试求出（2.2）～（2.4）的一个解. 第三个方程表述为

$$\alpha(x) = (1-p)\alpha(x-1) + p\alpha(x+1), \quad x > 0.$$

由（0.5）和（0.6）得出上述方程的解的唯一形式为

$$\alpha(x) = c_1 + c_2\left(\frac{1-p}{p}\right)^x, \quad p \neq \frac{1}{2},$$

$$\alpha(x) = c_1 + c_2 x, \quad p = \frac{1}{2}.$$

从（2.3）的第一个条件得 $\alpha(0)=1$；将此代入上式，

$$\alpha(x) = (1 - c_2) + c_2\left(\frac{1-p}{p}\right)^x, \quad p \neq \frac{1}{2}, \tag{2.5}$$

$$\alpha(x) = 1 + c_2 x, \quad p = \frac{1}{2}. \tag{2.6}$$

若选择 $c_2=0$，则对所有的 x 都有 $\alpha(x)=1$，显然不满足（2.3）. 若 $p=\frac{1}{2}$ 且 $c_2\neq0$，则解无界，因此不满足（2.2）. 类似地，若 $p<\frac{1}{2}$，则当 $c_2\neq0$ 时，（2.5）的解无界. 在这种情形下，我们得出当 $p\leqslant\frac{1}{2}$ 时，该马尔可夫链为常返的. 当 $p>\frac{1}{2}$ 时，我们可以求出一个解. （2.3）中的第二个条件本质上归结为当 $x\to\infty$ 时，$\alpha(x)\to0$，我们得到

$$\alpha(x) = \left(\frac{1-p}{p}\right)^x.$$

因此，当 $p>\frac{1}{2}$ 时，该马尔可夫链为非常返的.

2.3 正常返和零常返

50 假定 X_n 是无限状态空间 S 上的一个不可约、非周期的马尔可夫链. 在这一节我们将研究什么情况下极限概率分布存在. 所谓极限概率 $\pi(x)$（$x\in S$），是指 S 上的一个概率分布，使得对每一个 x，$y\in S$，

$$\lim_{n\to\infty} p_n(y,x) = \pi(x).$$

若 X_n 为非常返的，则对所有的 x，y，

$$\lim_{n\to\infty} p_n(y,x) = 0, \tag{2.7}$$

所以 X_n 不存在极限概率分布. 然而对于一些常返链，（2.7）也可能成立. 例如，考虑上节描述的 $\mathbf{Z}$ 上的简单随机游动（事实上，它是一个周期链，但如果做些小的修改，它可以变成一个非周期的例子）. 它是常返的，但是当 $n\to\infty$ 时，$p_{2n}(0,0)\to0$. 若马尔可夫链为常返的，且

$$\lim_{n\to\infty} p_n(x,y) = 0,$$

则称该马尔可夫链为*零常返的*（null recurrent）. 否则，一个常返链称为*正常返的*（positive recurrent）.

正常返马尔可夫链有着同有限马尔可夫链相似的性质. 若 X_n 为一个不可约、非周期的正常返马尔可夫链，则对每一个 x，y，极限

$$\lim_{n\to\infty} p_n(y,x) = \pi(x) > 0$$

存在，且独立于初始状态 y. 此 $\pi(x)$ 称为 S 上的不变概率分布，即

$$\sum_{y\in S}\pi(y)p(y,x) = \pi(x). \tag{2.8}$$

再者，若返回到一个状态 x 的时间为

$$T = \min\{n>0 \mid X_n = x\},$$

则对于一个正常返马尔可夫链，

$$\mathbf{E}(T \mid X_n = x) = \frac{1}{\pi(x)}.$$

若 X_n 是零常返的，则以概率 1 有 $T<\infty$，但 $\mathbf{E}(T)=\infty$；若 X_n 是非常返的，则以正概率有 $T=\infty$.

判定一个马尔可夫链是否为正常返的一种方法是尝试找到一个不变概率分布．可以证明若一个不可约马尔可夫链是正常返的，则存在满足（2.8）的唯一的概率分布．再者，若一个马尔可夫链不是正常返的，则不存在满足（2.8）的概率分布．这里给出了一个好的准则：尝试找到不变概率分布．若其存在，则该马尔可夫链为正常返的；若不存在，则该马尔可夫链为零常返的或非常返的． 51

例 再次考虑带有部分反射壁的随机游动的例子．我们将尝试找到满足（2.8）的不变概率分布，即非负函数 $\pi(x)$ 满足（2.8）且

$$\sum_{x \in S} \pi(x) = 1. \tag{2.9}$$

在这个例子中，由（2.8）给出

$$\pi(x+1)(1-p) + \pi(x-1)p = \pi(x), \quad x > 0, \tag{2.10}$$

$$\pi(1)(1-p) + \pi(0)(1-p) = \pi(0). \tag{2.11}$$

由（0.5）和（0.6），（2.10）的通解为

$$\pi(x) = c_1 + c_2\left(\frac{p}{1-p}\right)^x, \quad p \neq \frac{1}{2},$$

$$\pi(x) = c_1 + c_2 x, \quad p = \frac{1}{2}.$$

从等式（2.11）得 $\pi(0)=[(1-p)/p]\pi(1)$，将此条件加入到上式，得

$$\pi(x) = c_2\left(\frac{p}{1-p}\right)^x, \quad p \neq \frac{1}{2},$$

$$\pi(x) = c_1, \quad p = \frac{1}{2}.$$

现在加上条件（2.9）：我们能否选择 c_1 或 c_2，使得 $\sum\pi(x)=1$？当 $p=\frac{1}{2}$ 时，显然不行．假定 $p\neq\frac{1}{2}$，显然我们需要 $c_2\neq 0$．若 $p>\frac{1}{2}$，$\sum\left(\frac{p}{1-p}\right)^x=\infty$，则不可能找到满足条件的 c_2（我们已经知道在这样的情形下该马尔可夫链为非常返，所以它不可能为正常返的）．然而若 $p<\frac{1}{2}$，和为有限的，并且我们可以选择

$$\pi(x) = \left(\frac{p}{1-p}\right)^x \left[\sum_{y=0}^{\infty}\left(\frac{p}{1-p}\right)^y\right]^{-1} = \left(\frac{1-2p}{1-p}\right)\left(\frac{p}{1-p}\right)^x.$$

在这种情形下，该马尔可夫链为正常返，且这里给出了不变概率．总结上两节的讨论，关于带部分反射壁的随机游动，我们有如下结论：

$$\text{当 } p < \frac{1}{2} \text{ 时，} \quad \text{正常返，}$$

$$\text{当 } p = \frac{1}{2} \text{ 时，}\quad \text{零常返，}$$

52

$$\text{当 } p > \frac{1}{2} \text{ 时，}\quad \text{非常返.}$$

2.4 分支过程

本节将研究群体增长的随机模型．考虑一些个体组成的群体．令 X_n 为时刻 n 时此群体中的个体数．在每个时间间隔内，群体依下面规则变化：每个个体产生后代的个数都是随机的，个体产生后代后，它自身的生命结束，离开群体．我们对产生后代的过程做两个假定：

1. 每个个体产生后代的个数服从相同的概率分布：一个个体恰好产生 k 个后代的概率为 p_k，已知概率 p_0，p_1，p_2，…非负，且和为 1.
2. 各个个体产生后代的个数是相互独立的.

群体第 n 代的个体数 X_n 是一个马尔可夫链，其状态空间为 $\{0,1,2,\cdots\}$．记 0 为吸收状态；一旦群体灭绝，不会再有个体产生．该马尔可夫链的确切的转移概率不容易写出．假定 $X_n=k$．则 k 个个体产生的后代形成第 $n+1$ 代个体．若 $Y_1,\cdots,Y_k$ 为独立的随机变量，且每个变量的分布为 $\mathbf{P}\{Y_i=j\}=p_j$，则

$$p(k,j) = \mathbf{P}\{X_{n+1} = j \mid X_n = k\} = \mathbf{P}\{Y_1 + \cdots + Y_k = j\}.$$

$Y_1+\cdots+Y_k$ 的实际分布可以根据卷积公式求出，但是这里我们不需要知道它的确切分布形式．令 μ 表示每个个体产生后代的均值，则

$$\mu = \sum_{i=0}^{\infty} i p_i.$$

那么

$$\mathbf{E}(X_{n+1} \mid X_n = k) = \mathbf{E}(Y_1 + \cdots + Y_k) = k\mu.$$

可以相对简单地计算出个体数的均值 $\mathbf{E}(X_n)$：

$$\begin{aligned}\mathbf{E}(X_n) &= \sum_{k=0}^{\infty} \mathbf{P}\{X_{n-1} = k\}\mathbf{E}(X_n \mid X_{n-1} = k) \\ &= \sum_{k=0}^{\infty} k\mu \mathbf{P}\{X_{n-1} = k\} = \mu \mathbf{E}(X_{n-1}).\end{aligned}$$

或者，若重复做 n 次，则得

53

$$\mathbf{E}(X_n) = \mu^n \mathbf{E}(X_0).$$

从这个表达式可得到一些有趣的结论．若 $\mu<1$，则当 n 很大时，后代个数的均值接近于 0．简单估计

$$\mathbf{E}(X_n) = \sum_{k=0}^{\infty} k\mathbf{P}\{X_n = k\} \geqslant \sum_{k=1}^{\infty} \mathbf{P}\{X_n = k\} = \mathbf{P}\{X_n \geqslant 1\}$$

可用来推导出群体最终会灭绝，即

$$\lim_{n\to\infty} \mathbf{P}\{X_n = 0\} = 1.$$

若 $\mu=1$，则期望群体规模保持不变，而当 $\mu>1$ 时，期望群体规模会逐渐变大．在这些情形中

并不能很清楚地看出群体是否以概率 1 灭绝.（尽管 $\mathbf{E}(X_n)$ 不太小，但 X_n 有可能以非常接近 1 的概率达到 0）下面我们将研究如何确定群体灭绝的概率. 为了避免平凡情况，假定

$$p_0 > 0; \quad p_0 + p_1 < 1. \tag{2.12}$$

令

$$a_n(k) = \mathbf{P}\{X_n = 0 \mid X_0 = k\}$$

并且令 $a(k)$ 为假定最初有 k 个个体时群体最终灭绝的概率，即

$$a(k) = \lim_{n\to\infty} a_n(k).$$

若群体在某个时刻有 k 个个体，那么群体灭绝的唯一方法是这 k 个分支都灭绝. 既然各个分支的行为是相互独立的，

$$a(k) = [a(1)]^k.$$

所以只需考虑 $a(1)$，将 $a(1)$ 简记为 a，称为消亡概率（extinction probability）. 现假定 $X_0 = 1$，若仅考虑一代，则得

$$\begin{aligned} a &= \mathbf{P}\{\text{群体消亡} \mid X_0 = 1\} \\ &= \sum_{k=0}^{\infty} \mathbf{P}\{X_1 = k \mid X_0 = 1\}\mathbf{P}\{\text{群体消亡} \mid X_1 = k\} \\ &= \sum_{k=0}^{\infty} p_k a(k) = \sum_{k=0}^{\infty} p_k a^k. \end{aligned}$$

等式右边会引起足够的兴趣，并给予命名. 若 X 是一个在 $\{0,1,2,\cdots\}$ 中取值的随机变量，则 X 的母函数（generating function）为函数

$$\phi(s) = \phi_X(s) = \mathbf{E}(s^X) = \sum_{k=0}^{\infty} s^k \mathbf{P}\{X = k\}.$$ 54

注意到当 $s \geqslant 0$ 时，$\phi(s)$ 是关于 s 的增函数，且有 $\phi(0) = \mathbf{P}\{X=0\}$，$\phi(1)=1$. 求导，我们可得

$$\phi'(s) = \sum_{k=1}^{\infty} k s^{k-1} \mathbf{P}\{X = k\},$$

$$\phi''(s) = \sum_{k=2}^{\infty} k(k-1) s^{k-2} \mathbf{P}\{X = k\}.$$

因此，

$$\phi'(1) = \sum_{k=1}^{\infty} k \mathbf{P}\{X = k\} = \mathbf{E}(X), \tag{2.13}$$

且对于 $s>0$，若 $\mathbf{P}\{X \geqslant 2\} > 0$，

$$\phi''(s) > 0. \tag{2.14}$$

若 $X_1, \cdots, X_m$ 是取值为非负整数的独立随机变量，则

$$\phi_{X_1+\cdots+X_m}(s) = \phi_{X_1}(s) \cdots \phi_{X_m}(s).$$

证明上述等式的最简单的方法是利用表达式 $\phi_X(s) = \mathbf{E}(s^X)$ 和独立随机变量的乘法法则.

回到分支过程，我们已知消亡概率 a 满足方程

$$a = \phi(a).$$

显然，$a=1$ 满足方程，但可能还有其他解. 我们再次假定 $X_0 = 1$，随机变量 X_0 的母函数为 a，

X_1 的母函数为 $\phi(a)$. 令 $\phi^n(a)$ 为 X_n 的母函数，我们将证明

$$\phi^n(a) = \phi(\phi^{n-1}(a)).$$

为此，首先注意到

$$\begin{aligned}\phi^n(a) &= \sum_{k=0}^{\infty} \mathbf{P}\{X_n = k\}a^k \\ &= \sum_{k=0}^{\infty} \left[\sum_{j=0}^{\infty} \mathbf{P}\{X_1 = j\}\mathbf{P}\{X_n = k \mid X_1 = j\}\right]a^k \\ &= \sum_{j=0}^{\infty} p_j \sum_{k=0}^{\infty} \mathbf{P}\{X_{n-1} = k \mid X_0 = j\}a^k.\end{aligned}$$

55

现在，若 $X_0=j$，那么 X_{n-1} 可以表示为 j 个独立的随机变量之和，且每一个随机变量的分布等同于给定 $X_0=1$ 时 X_{n-1} 的分布. 因此，对 k 求和得到的是 j 个独立随机变量之和的母函数，其中每个变量的母函数为 $\phi^{n-1}(a)$，因此

$$\sum_{k=0}^{\infty} \mathbf{P}\{X_{n-1} = k \mid X_0 = j\}a^k = [\phi^{n-1}(a)]^j,$$

$$\phi^n(a) = \sum_{j=0}^{\infty} p_j[\phi^{n-1}(a)]^j = \phi(\phi^{n-1}(a)).$$

我们现在利用递归的方法得到 $\phi^n(a)$，因此可得

$$a_n(1) = \mathbf{P}\{X_n = 0 \mid X_0 = 1\} = \phi^n(0).$$

我们已准备好证明如下结论：消亡概率 a 是方程 $a=\phi(a)$ 的最小正数解. 我们已知 a 必须满足此方程. 令 $\hat{a}$ 为最小正数解. 我们将利用归纳法证明对每一个 n，都有 $a_n=\mathbf{P}\{X_n=0\}\leqslant\hat{a}$（暗含着 $a=\lim a_n\leqslant\hat{a}$）. 因为 $a_0=0$，显然当 $n=0$ 时结论成立. 假设 $a_{n-1}\leqslant\hat{a}$，则

$$\mathbf{P}\{X_n = 0\} = \phi^n(0) = \phi(\phi^{n-1}(0)) = \phi(a_{n-1}) \leqslant \phi(\hat{a}) = \hat{a}.$$

上式中的不等式是依据 ϕ 为增函数的事实得出的.

例 1 令 $p_0=\frac{1}{4}$，$p_1=\frac{1}{4}$，$p_2=\frac{1}{2}$，则 $\mu=\frac{5}{4}$，且

$$\phi(a) = \frac{1}{4} + \frac{1}{4}a + \frac{1}{2}a^2.$$

求解 $a=\phi(a)$，得 $a=1$，$\frac{1}{2}$. 消亡概率为 $\frac{1}{2}$.

例 2 令 $p_0=\frac{1}{2}$，$p_1=\frac{1}{4}$，$p_2=\frac{1}{4}$，则 $\mu=\frac{3}{4}$，且

$$\phi(a) = \frac{1}{2} + \frac{1}{4}a + \frac{1}{4}a^2.$$

求解 $a=\phi(a)$，得 $a=1$，2. 消亡概率为 1（我们已证明了当 $\mu<1$ 时消亡概率为 1 的结论）.

例 3 令 $p_0=\frac{1}{4}$，$p_1=\frac{1}{2}$，$p_2=\frac{1}{4}$，则 $\mu=1$，且

56

$$\phi(a) = \frac{1}{4} + \frac{1}{4}a + \frac{1}{4}a^2.$$

求解 $a=\phi(a)$，得 $a=1$，1. 消亡概率为 1.

本章最后我们建立一个判定 a 是否小于 1 的准则. 我们已经知道当 $\mu<1$ 时，$a=1$. 假定 $\mu=1$，由（2.13），$\phi'(1)=1$，所以由（2.14），当 $s<1$ 时，$\phi'(s)<1$. 因此对于任意 $s<1$，

$$1-\phi(s)=\int_s^1\phi'(s)\mathrm{d}s<1-s,$$

即 $\phi(s)>s$. 因此，当 $\mu=1$ 时，消亡概率为 1. 这是一个有趣的结论：尽管期望群体规模保持为 1，但是群体消亡的概率增长为 1. 此结论的一个推论是在群体没有灭绝的条件下，群体的规模将随时间增大. 这表明，假如某人被告知一个群体在一个很大的时刻还没有灭绝，那么此人将预测该群体的规模很大.

现假定 $\mu>1$，则 $\phi'(1)>1$. 因此，必存在某个 $s<1$，使得 $\phi(s)<s$. 但 $\phi(0)>0$. 由一般的连续函数的结论可知，必存在 $a\in(0, s)$，使得 $\phi(a)=a$. 又因为对 $s\in(0, 1)$有 $\phi''(s)>0$，所以曲线为凸的，并且至多存在一个 $s\in(0, 1)$，使得 $\phi(s)=s$. 这种情形下，群体以一个正的概率永远生存下去. 总结上述结论，给出如下定理.

定理 若 $\mu\leqslant1$ 且 $p_0>0$，则消亡概率 $a=1$，即群体最终灭绝. 若 $\mu>1$，则消亡概率 $a<1$，a 为方程

$$t=\phi(t),$$

的唯一解，其中 $t\in(0, 1)$.

2.5 习题

2.1 考虑排队模型（2.1 节例 3）. 当 p，q 取何值时，马尔可夫链为零常返的、正常返的、非常返的？
在正常返情形下，给出极限概率分布 π. 在平稳状态下，队列的平均长度是多少？
在非常返情形下，给出开始状态为 x 且马尔可夫链在某个时刻到达 0 的概率.

2.2 考虑如下的马尔可夫链，其状态空间为 $S=\{0,1,\cdots\}$. 给定一个正数序列 p_1，p_2，…且 $\sum_{i=1}^{\infty}p_i=1$. 每当该马尔可夫链到达状态 0 时，它便依据概率 p_i 选择一个新状态. 每当该马尔可夫链不在状态 0 时，它便确定地每次向 0 的方向移动一步. 换句话说，该马尔可夫链有转移概率 57

$$p(x,x-1)=1,\quad x>0,$$
$$p(0,x)=p_x,\quad x>0.$$

既然该马尔可夫链能持续地返回 0，则该马尔可夫链是个常返链. 当 p_x 在何种条件下，该马尔可夫链为正常返？在这种情形下，极限概率分布 π 是什么？[提示：或许直接计算 $\mathbf{E}(T)$ 更为简单些，其中 T 是起始为 0 并第一次返回到 0 的时间.]

2.3 考虑状态空间为 $S=\{0,1,\cdots\}$ 的一个马尔可夫链，且转移概率为

$$p(x,x+1)=\frac{2}{3};\qquad p(x,0)=\frac{1}{3}.$$

证明：该马尔可夫链为正常返链，并给出极限概率 π.

2.4 考虑状态空间为 $S=\{0,1,2,\cdots\}$ 的一个马尔可夫链，且转移概率为

$$p(x,x+2)=p,\quad p(x,x-1)=1-p,\quad x>0.$$

$$p(0,2)=p,\quad p(0,0)=1-p.$$

当 p 取何值时，该链是一个非常返链？

2.5 令 X_n 为状态空间 $\mathbf{Z}$ 上的一个马尔可夫链，转移概率为

$$p(n,n+1)=p,\quad p(n,n-1)=1-p,$$

其中 $p>\frac{1}{2}$. 假定 $X_0=0$.

(a) 令 $Y=\min\{X_0,X_1,\cdots\}$，Y 的分布是什么？

(b) 对正整数 k，令 $T_k=\min\{n: X_n=k\}$，且 $e(k)=\mathbf{E}(T_k)$. 解释 $e(k)=ke(1)$.

(c) 求出 $e(1)$（提示：(b) 或许有用）.

(d) 利用 (c) 得出另一结论：当 $p=\frac{1}{2}$时，$e(1)=\infty$.

2.6 假定 J_1，J_2，…为独立的随机变量且 $\mathbf{P}\{J_j=1\}=1-\mathbf{P}\{J_j=0\}=p$. 令 k 为一正整数，且 T_k 为首次连续 k 次出现 1 的时间. 换句话说，若 $J_n=J_{n-1}=\cdots=J_{n-(k-1)}=1$，则 $T_k=n$，且不存在 $m<n$，使得 $J_m=J_{m-1}=\cdots=J_{m-(k-1)}=1$. 令 $X_0=0$，对于 $n>0$，令 X_n 为最近的 1 连续出现的次数，即若 $J_{n-k}=0$ 且对于 $n-k<i\leqslant n$ 有 $J_i=1$，则 $X_n=k$.

(a) 解释 X_n 为什么是一个马尔可夫链，且状态空间为 $\{0, 1, 2, \cdots\}$，并给出转移概率.

58

(b) 证明：该马尔可夫链为不可约的正常返链，并给出不变概率 π.

(c) 用 $\mathbf{E}[T_{k-1}]$ 来表示出 $\mathbf{E}[T_k]$，求解此递归方程，从而解出 $\mathbf{E}[T_k]$.

(d) 以不同的方式求 $\mathbf{E}[T_k]$. 假定该马尔可夫链的初始状态为 k，并令$\hat{T}_k$为直到返回状态 k 的时间，$\hat{T}_0$为该马尔可夫链直到返回状态 0 的时间. 解释为什么

$$\mathbf{E}[\hat{T}_k]=\mathbf{E}[\hat{T}_0]+\mathbf{E}[T_k],$$

求 $\mathbf{E}[\hat{T}_0]$，并利用 (b) 去求得 $\mathbf{E}[\hat{T}_k]$.

2.7 令 X_n 为状态空间 $S=\{0,1,2,\cdots\}$ 上的一个马尔可夫链. 对下面的每一个转移概率，表述该马尔可夫链是否为正常返、零常返或非常返. 若为正常返，给出平稳概率分布：

(a) $p(x, 0)=\dfrac{1}{(x+2)}$，$p(x, x+1)=\dfrac{(x+1)}{(x+2)}$.

(b) $p(x, 0)=\dfrac{(x+1)}{(x+2)}$，$p(x, x+1)=\dfrac{1}{(x+2)}$.

(c) $p(x, 0)=\dfrac{1}{(x^2+2)}$，$p(x, x+1)=\dfrac{(x^2+1)}{(x^2+2)}$.

2.8 给定分支过程的后代分布如下，求消亡概率 a.

(a) $p_0=0.25$，$p_1=0.4$，$p_2=0.35$.

(b) $p_0=0.5$，$p_1=0.1$，$p_3=0.4$.

(c) $p_0=0.91$，$p_1=0.05$，$p_2=0.01$，$p_3=0.01$，$p_6=0.01$，$p_{13}=0.01$.

(d) $p_i=(1-q)q^i$，$0<q<1$.

2.9 考虑分支过程，其后代分布如习题 2.8 (b) 所示，令 $X_0=1$.

(a) 已知群体在第一代中不会灭绝（$X_1>0$），群体在第二代中灭绝的概率为多少？

(b) 已知群体不会在第二代中灭绝，群体在第三代灭绝的概率为多少？

2.10 考虑分支过程，其后代的分布由 $\{p_n\}$ 决定．我们在该过程中加入下面的条件，使得该过程变成不可约马尔可夫链：群体一旦灭绝，下一代将会有一个新的个体．[换句话说，$p(0,1)=1$.] 那么当 $\{p_n\}$ 取何值时，该马尔可夫链为正常返、零常返和非常返？

2.11 考虑下述分支过程的变化．在每个时刻 n，每个个体依后代分布 $\{p_n\}$ 独立地产生后代子孙，且个体死亡概率为 q，$q\in(0,1)$．这样，每个个体可以产生 j 次后代，其中 j 为个体的寿命．当 q，$\{p_n\}$ 取何值时，群体最终的消亡概率为 1． 59

2.12 考虑分支过程，其中 $p_0=\dfrac{1}{3}$，$p_1=\dfrac{1}{3}$，$p_2=\dfrac{1}{3}$．在电脑的辅助下，求群体在 n 步后灭绝的概率．其中 $n=20$，100，200，1 000，1 500，2 000，5 000．分别考虑 $p_0=0.35$，$p_1=0.33$，$p_2=0.32$ 和 $p_0=0.32$，$p_1=0.33$，$p_2=0.35$ 时的情况．

2.13 考虑一个生物群体依下述规则进行无性繁殖：每个个体出生后能存活足够长时间以繁殖后代的概率为 q．若个体确实繁殖后代，则每次它以相等的概率繁殖 1 个或 2 个后代．当这个个体不再繁殖后，则最终消亡．假定群体最初有 4 个个体．

(a) 当 q 取何值时，能保证群体会最终灭绝．

(b) 若 $q=0.9$，则该群体永远存活下去的概率是多少？

2.14 令 X_n 是一个 $\mu>1$ 的分支过程中时刻 n 的个体数．假定 $X_0=1$．令 ϕ 为后代分布的母函数，并令 $a<1$ 为消亡概率．

(a) 解释为什么 $\phi'(a)<1$．

(b) 令 $a_n=\mathbf{P}\{X_n=0\}$，利用 (a) 证明：存在 $\rho<1$，使得对所有充分大的 n，

$$a-a_{n+1}\leqslant\rho(a-a_n).$$

(c) 证明：存在 $b>0$，$c<\infty$，使得对所有的 n，

$$\mathbf{P}\{\text{灭绝}\mid X_n\neq 0\}\leqslant ce^{-bn}.$$

换句话说，若群体将要灭绝，那么很有可能在前几代就会灭绝．

2.15 令 X_1，X_2，…为独立同分布的随机变量，取值为整数且均值为 0．令 $S_0=0$ 且

$$S_n=X_1+\cdots+X_n.$$

(a) 令

$$G_n(x)=\mathbf{E}\left[\sum_{j=0}^{n}I\{S_j=x\}\right]$$ 60

为在最初的 n 步中返回到 x 的期望次数．证明：对所有的 n 和 x，$G_n(0)\geqslant G_n(x)$．[提示：考虑使得 $S_j=x$ 的 j 的最小值．]

(b) 回忆一下，由大数定律知，对每个 $\varepsilon>0$，有

$$\lim_{n\to\infty}\mathbf{P}\{|S_n|\leqslant n\varepsilon\}=1.$$

据此证明对每个 $\varepsilon>0$，

$$\lim_{n\to\infty}\frac{1}{n}\sum_{|x|\leqslant\varepsilon n}G_n(x)=1.$$

(c) 利用 (a) 和 (b)，证明：对每个 $M<\infty$ 都存在一个 n，使得 $G_n(0)\geqslant M$．

(d) 推导 S_n 为一个常返马尔可夫链.

2.16 令 $p_1, p_0, p_{-1}, \cdots$ 为 $\{\cdots, -2, -1, 0, 1\}$ 上的一个概率分布，且均值为负，即

$$\sum_n n p_n = \mu < 0.$$

定义非负整数上的一个马尔可夫链 X_n，转移概率为

$$p(n,m) = p_{m-n}, \quad m > 0,$$

$$p(n,0) = \sum_{m \leqslant 0} p_{m-n}.$$

换句话说，X_n 像一个增长依赖于 p_i 的随机游动，只是此游动不允许跃到 0 以下. 这个练习的主要目的是证明该马尔可夫链为正常返的.

(a) 令 $\pi(n)$ 为该马尔可夫链的一个不变概率. 证明：对每个 $n>0$，有

$$\pi(n) = \sum_{m=n-1}^{\infty} \pi(m) p_{n-m}.$$

(b) 令 $q_n = p_{1-n}$. 证明：存在一个 $\alpha \in (0, 1)$，满足

$$\alpha = q_0 + q_1 \alpha + q_2 \alpha^2 + \cdots.$$

[提示：q_n 为一随机变量的概率分布，且均值大于 1，右边为 q_n 的母函数.]

61 (c) 利用 (b) 中的 α 求该马尔可夫链的不变概率分布.

2.17 令 $p(x, y)$ 为一个马尔可夫链的转移概率，其状态空间为 S. 如果

$$\sum_{y \in S} p(x,y) f(y) \leqslant f(x)$$

对一个固定状态 $z \in S$ 成立，则称 f 为 p 在 x 的上调和函数.

(a) 令$\mathcal{A}$为满足 $f(z)=1$ 的所有函数 f 的集合，且对所有的 $y \in S$，满足 $0 \leqslant f(y) \leqslant 1$，在所有 $y \neq z$ 上 f 为上调和函数. 令 g 定义为

$$g(x) = \inf_{f \in \mathcal{A}} f(x).$$

证明：$g \in \mathcal{A}$.

(b) 证明：对所有 $x \neq z$，有

$$\sum_{y \in S} p(x,y) g(y) = g(x).$$

[提示：假定对某个 x，有 $\sum_y p(x, y) g(y) < g(x)$，证明如何在 x 处递减 g，使得函数保持上调和.]

(c) g 的定义如 (a). 证明：如果对某个 x，$g(x) < 1$，则

$$\inf_{x \in S} g(x) = 0.$$

[提示：令 $\varepsilon = \inf_x g(x)$，考虑 $h(x) = \dfrac{(g(x) - \varepsilon)}{(1-\varepsilon)}$.]

(d) 推导如下结论：假定一个不可约马尔可夫链，其转移概率为 $p(x, y)$，存在函数 f，使得对某些 $x \in S$ 有 $f(x) < 1$，且对所有的 $y \neq z$ 上调和，其中 $f(z) = 1$，$0 \leqslant f(y) \leqslant 1$，$y \in S$. 则该马尔可夫链为非常返的.

2.18 在此习题中，我们将证明斯特林公式

$$n! \sim \sqrt{2\pi} n^{n+\left(\frac{1}{2}\right)} e^{-n}.$$

设 $X_1, X_2, \cdots$ 为独立的泊松（Poisson）随机变量，均值为 1，而 $Y_n = X_1 + \cdots + X_n$ 为一个均值为 n 的泊松随机变量．设

$$p(n,k) = \mathbf{P}\{Y_n = k\} = e^{-n}\frac{n^k}{k!}.$$

（a）利用中心极限定理证明：若 $a>0$，则

$$\lim_{n\to\infty}\sum_{n\leqslant k<n+a\sqrt{n}} p(n,k) = \int_0^a \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx.$$

（b）证明：若 $a>0$，n 为一个正整数，且 $n\leqslant k<n+a\sqrt{n}$，则

$$e^{-a^2}p(n,n)\leqslant p(n,k)\leqslant p(n,n).$$ 62

（c）利用（a）和（b）得到结论

$$p(n,n)\sim\frac{1}{\sqrt{2\pi n}}.$$

斯特林公式（2.15）即可得． 63