E. B. 邓肯　A. A. 尤什凯维奇　著

马尔可夫过程

（定理与问题）

科学出版社

马尔可夫过程

（定理与问题）

E. B. 邓肯
A. A. 尤什凯维奇 著

张饴慈 刘吉江 译
王梓坤 校

科学出版社
1988

内 容 简 介

本书用较初等的方法通过典型的例子向数学、物理工作者以及工程技术人员介绍马尔可夫过程的新理论. 第一章介绍最简单的马氏链,即点格上的对称随机徘徊,阐明调和函数等一些古典分析概念与概率论之间的联系. 第二章叙述如何用概率的概念得到分析的结果,证明了 Dirichlet 问题解的存在性. 第三章是最优停止问题. 第四章用概率的观点处理微分方程的边值问题. 每章后附有许多富有启发性的问题.

E. B. Dynkin A. A. Yushkevich
MARKOV PROCESSES,
THEOREMS AND PROBLEMS
Plenum Press 1969

马尔可夫过程

(定理与问题)

E. B. 邓 肯 A. A. 尤什凯维奇 著
张饴慈 刘吉江 译
王梓坤 校

责任编辑 刘嘉善

科学出版社 出版
北京朝阳门内大街 137 号

中国科学院印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1988 年 8 月第 一 版 开本：787×1092 1/32
1988 年 8 月第一次印刷 印张：7 3/8
印数：0001—2,600 字数：166,000

ISBN 7-03-000509-0/O · 135

定价： 4.00 元

前　　言

概率论的概念和方法在自然科学与工程技术中正得到越来越多的应用，而且它们也越来越深入地渗透到数学本身的各个领域．了解这些方法对不同专门化的数学家、物理学家以及工程技术人员同样是有益处的．然而，一些初等教科书对于这个学科的现代发展只能给出有限的概念，而那些致力于这个领域最新方向的专著通常又都是写给一些专家的，并且这些专著又都使用了高深的集合论与分析工具．

为了掌握新的数学概念，必须要了解它们的功能，还要看到它们是怎样起作用的．为此，最好不要从一般性定理开始，而是着手于一些特殊的问题，这些问题应该是实际的并且例子应是典型的，而且还应避免在过于繁杂论证的系统中出现的那些次要的技术上的困难．

本书的目的是用上面指出的方法向读者介绍马尔可夫过程理论中的最新研究成果．

马尔可夫过程是研究得相当彻底并且有多方面应用的一类随机过程．马尔可夫过程理论中已成为经典理论的那些分支可在几本极好的书(例如[1]，[2])中找到．然而，近些年来，人们不仅发现了马尔可夫过程与数学分析之间的新关系，而且还发现了越来越多的新的有意义的方向．在一些专著([3]—[11])中对这些问题作了相当透彻的讨论，不过对刚刚开始熟悉这门学科的人来说很不适合．但是，整个问题的范围主要基于浅显的和图解的想法上，这样，研究这些问题为人们能适应概率论的思想方法提供了大量的材料．

本书共分四章，每一章向读者介绍一个确定的问题的范围．如第一章介绍位势、调和函数与过分函数以及过程轨道的极限状态；第二章介绍微分方程的概率解；第三章介绍某些最优控制问题；第四章介绍分析中边界问题的概率内容．

在第一章中，我们考虑最简单的马尔可夫链，即点格上的对称随机徘徊，说明在分析中一些熟悉的概念，如调和函数、位势、容度等等在离散模型中有它们的类似物，而且可以用于解决纯粹概率论的问题，例如，到达一个给定集合的次数问题．这一章的依据是 Ito 和 McKean 的著作[12]．

在第二章中，说明如何用概率论的思想得到分析中的结果．特别地，用这种方法证明了 Laplace 方程在很广的一类区域上的 Dirichlet 问题解的存在性[1]．

在第三章中，通过对马尔可夫过程最优停止问题的研究，使马尔可夫过程与位势理论之间的关系得到了一个意外的应用，这一章的来源是[13]．

在第四章中，用概率论的观点处理了这样一些问题，即微分方程和其它方程中最广泛的边界条件问题，这些问题近年来得到许多研究工作者的注意．分析最简单的模型（生灭过程），能使我们用完全初等的方法进行严格地处理．最早把概率论的方法应用到边界问题上的人是 Feller，在[14]中，他论述了生灭过程，尽管他是靠概率直观来指引，然而，在他所引进的全部结构中仍然使用纯分析的形式．我们的方法是基于

1) 在马尔可夫过程的一般理论(H.B. Phillips 和 N. Wiener 在 1923年；R. Courant, K. Friedrichs 和 H. Levy 在1928年)出现之前，概率论与 Dirichlet 问题之间的联系就已众所周知，在 A.Ya. Khinchin (1923) 和 I.G. Petrovskii (1934) 的著作中，又详尽地发展了这个思想，用 Wiener 过程的轨道把 Dirichlet 问题的解表示成一个公式是由 J.Doob (1954)得到的．然而，Doob 应用它在方向上与我们正相反，即我们是从分析中的定理推导概率的轨道性质．

考虑轨道的性质并且依赖于特征算子的概念（这一章的概要包含在[15]中）.

在每一章末安排了一系列问题，这些问题超出了简单地提供一些练习的作用，因为有些是正文的补充同时又提出了某些新的内容. 例如，在第三章的习题中包括一个可列马尔可夫链的 Martin 边界的讨论.

为了不中断概率论讨论的主流，一些辅助性的分析问题在附录中叙述.

除了上面提到的主要文献来源外，在全书中将有机会指出其它一些著作(主要在习题和例中).

希望读者能熟悉概率论与分析的基本思想. 当然，某些问题需要较多的知识,在主要内容中,我们有意识地避开了测度论与可测性,因此,对已经有了这方面概念的读者将不难在更严格的集合论的水平上掌握上述内容.

本书基本上来源于它的第一个作者于 1962 至 1963 年在莫斯科大学的讲座的内容（这些讲座由 M.B. Malyutov, S.A. Molchanav 和 M.I. Freidlin 记录）,后来经作者增补并作了根本性的修改,又增加了习题而完成.

作者感谢 I.L. Genis，她在本书出版过程中做了大量工作.

E.B. 邓肯

A.A. 尤什凯维奇

1966.1.22.

目　　录

第一章　常返性判别法

§ 1. 对称随机徘徊

考虑一个质点沿 x 轴在整数值点 $0, \pm 1, \pm 2, \cdots$ 上运动，它每次向左或向右等间隔地跳一个单位．如果它向左和向右跳的概率相等都是 $1/2$，那么我们把这个质点的运动称为**直线上的对称随机徘徊**.

质点能够到达的位置 $0, +1, -1, \cdots$ 称为**状态**.

我们来证明，从任意初始位置出发的质点，迟早总会到达任何一个可能状态的概率为 1．由于所有状态显然都是同等可能的，因此，只需证明，从任意状态出发的质点将在某个时刻到达位置 0．用 $\pi(x)$ 表示从点 x 出发到达 0 的概率，那么 $\pi(0) = 1$，且根据全概率公式，对 $x \neq 0$ 有

$$\pi(x) = \frac{1}{2}\pi(x-1) + \frac{1}{2}\pi(x+1). \tag{1}$$

考虑函数 $\pi(x)$ 的图象，其中 $x = 0, 1, 2, \cdots k, \cdots$. (1) 式意味着这个图象中任意三个相邻点都在一条直线上．因此，当 $x \geqslant 0$ 时，函数 $\pi(x)$ 图象上的所有点都在同一条直线上．由于 $\pi(0) = 1$，所以这条线是自点 $(0, 1)$ 出发的一条射线．如果对某个正数 x，$\pi(x)$ 小于 1，那么这条直线就要和 x 轴相交，从而当 x 很大时，$\pi(x)$ 将是负的，但这是不可能的，因此，对一切 $x \geqslant 0$ 有 $\pi(x) = 1$. 利用随机徘徊的对称性可知，对 $x < 0$，$\pi(x) = 1$ 也成立．这就证明了，从任一初始状态出发到达状态 0 的概率等于 1.

直线上随机徘徊的一个自然推广是 l 维整数值点格 H_l 上的随机徘徊. 这个点格是由形如

$$x = x_1e_1 + x_2e_2 + \cdots + x_le_l$$

的点(向量)组成的,其中 $e_1, e_2, \cdots, e_l$ 是 l 维空间的一组正交基,坐标 $x_1, x_2, \cdots, x_l$ 是任意整数. 当这些坐标中的一个坐标加上或减去 1,而其它坐标保持不变,就得到点 x 的 $2l$ 个邻点(例如,在二维的情形,点格中每点有四个邻点,左、右、上、下各一). 不管质点以前的位置如何,它每经一步运动都以相等的概率 $1/2l$ 到达一个邻点.

对二维情形,和一维情形一样,可以证明,从二维点格中任一点出发的质点到达其它任一点的概率为 1 (见本章后面的习题). 但对三维和更高维点格来说,下面将看到,质点从一状态到达另一状态的概率小于 1,但到达某一集合 B (而不是一个单点) 的概率可能等于 1,也可能小于 1,我们把这个概率记作 $\pi_B(x)$,其中 x 是质点的初始状态. 如果对点格的所有点 x,有 $\pi_B(x) = 1$,称集 B 是**常返的**; 如果点格中至少有一点 x,使 $\pi_B(x) < 1$,那么称集 B 是**非常返的**. 这一章将给出一个判别法,以区分常返集与非常返集.

§2. 转 移 函 数

令 $X(0)$ 表示作随机徘徊的质点的初始位置, $X(n)$ ($n = 1, 2, 3, \cdots$) 表示 n 步后质点的位置.

一个与随机徘徊有关的事件 A 的概率当然依赖于此徘徊的初始状态 x. 我们把这个概率记作 $P_x\{A\}$,用记号 $M_x\xi$ 表示随机变量 ξ 关于这个分布 P_x 的均值.

其次,用 $p(n, x, y)$ 表示质点从位置 x 出发 n 步后到达位置 y 的概率,即

$$p(n, x, y) = P_x\{X(n) = y\}.$$

函数 $p(n, x, y)$ 是随机徘徊的一个重要特征，称它为**转移函数**. 显然，$p(0, x, x) = 1$；　$p(0, x, y) = 0, x \neq y$. 同样显然地有[1] $\sum\limits_y p(n, x, y) = 1$. 称量

$$\sum_{y \in B} p(n, x, y) = P_x\{X(n) \in B\}$$

为从 x 到 B 的 n 步转移概率，其中 B 是 l 维空间中的某一集合.

随机徘徊的一个重要性质(这对随机徘徊的分析很有用)是跳跃 $\xi_k = X(k) - X(k-1)$ $(k = 1, 2, \cdots)$ 相互独立，向量 ξ_k 有相同的分布，并且和质点的初始状态无关. 特别，任一个向量 ξ_k 取它的每一个值 $(\pm e_1, \cdots, \pm e_l)$ 的概率都相等. 利用这个事实，对于转移函数 $p(n, x, y)$ 我们给出一个积分表示.

我们用 $\theta(x)$ 表示在向量 e_k 上取值为 θ_k 的线性形式. 即，如果 $x = x_1e_1 + \cdots + x_le_l$，那么 $\theta(x) = \theta_1x_1 + \cdots \theta_lx_l$. 考虑函数

$$F(\theta) = \sum_y p(n, x, y)e^{i\theta(y)} = M_x e^{i\theta(X(n))}, \tag{2}$$

即随机向量 $X(n)$ 的特征函数.(事实上，在等式(2)中的级数仅有有限项不为零，因为在 n 步后质点可能到达的不同状态不会超过 $(2l)^n$ 个.)利用函数 $F(\theta)$ 较易表示转移函数 $p(n, x, y)$. 例如，令 θ 表示所有这样的线性形式 $\theta(z) = \theta_1z_1 + \cdots + \theta_lz_l$，其中系数 $\theta_1, \theta_2, \cdots, \theta_l$ 的绝对值不超过 π，用 $e^{-i\theta(z)}$ (这里 z 是点格 H^l 中的一个点)乘等式(2)，然后在 θ 上积分，由于 y 和 z 都是整数值坐标的向量，于是有

---

1) 今后符号 $\sum\limits_y$ 表示对点格 H^l 的所有点求和.

$$\int e^{i\theta(y)-i\theta(z)}d\theta=\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}e^{i[\theta_1(y_1-z_1)+\cdots+\theta_l(y_l-z_l)]}d\theta_1\cdots d\theta_l$$

$$=\prod_{k=1}^{l}\int_{-\pi}^{\pi}e^{i\theta_k(y_k-z_k)}d\theta_k$$

$$=\begin{cases}(2\pi)^l & \text{当 } y=z \text{ 时,}\\ 0 & \text{当 } y\neq z \text{ 时.}\end{cases}$$

由此知

$$p(n,x,z)=\frac{1}{(2\pi)^l}\int_Q F(\theta)e^{-i\theta(z)}d\theta. \qquad (3)$$

下面求函数 $F(\theta)$. 由于

$$X(n)=X(0)+\sum_{k=1}^{n}\xi_k,$$

其中 ξ_k 是第 k 步的跳跃,于是有

$$F(\theta)=M_x e^{i\theta(X(n))}=M_x e^{i\theta(X(0))}\prod_{k=1}^{n}e^{i\theta(\xi_k)}.$$

又因为 $X(0)=x$ 的概率为 1, 而且 ξ_k 独立同分布, 故有

$$F(\theta)=e^{i\theta(x)}\Phi^n(\theta), \qquad (4)$$

其中 $\Phi(\theta)=M_x e^{i\theta(\xi_1)}$, 向量 ξ_1 取 $\pm e_1,\cdots\pm e_l$ 的每一个值的概率都是 $\frac{1}{2l}$,因此

$$\Phi(\theta)=\frac{1}{2l}\sum_{m=1}^{l}(e^{i\theta_m}+e^{-i\theta_m})=\frac{1}{l}\sum_{m=1}^{l}\cos\theta_m. \qquad (5)$$

将(4)式代入(3)式并用 y 代替 z, 就得到

$$p(n,x,y)=\frac{1}{(2\pi)^l}\int_Q e^{i\theta(x-y)}\Phi^n(\theta)d\theta. \qquad (6)$$

§ 3. 当 $n\to\infty$ 时质点轨道的状况

在本节假定 $l\geqslant 3$, 我们来证明,当 $n\to\infty$ 时向量 $X(n)$

的长度概率为 1 地趋于无穷. 下面将看到: 由此导致任何有界集的非常返性.

假设我们作了一系列试验，其中第 n 次试验成功的概率为 p_n，于是和 $p_1+p_2+\cdots+p_n+\cdots$ 表示成功次数的均值(事实上，成功次数 η 等于和 $\eta_1+\eta_2+\cdots+\eta_n+\cdots$，其中如果第 n 次试验成功,那么 $\eta_n=1$，否则 $\eta_n=0$).

考虑一个从点 x 出发的随机徘徊，当 $X(n)=y$ 时认为第 n 次试验成功. 于是 $p_n=p(n, x, y)$，而和

$$g(x, y)=\sum_{n=0}^{\infty} p(n, x, y) \tag{7}$$

表示**到达点** y **的次数的均值**,我们来证明

$$g(x, y)<\infty. \tag{8}$$

(可以证明,在一维和二维情形时,对任意 x, y, 有 $g(x, y)=\infty$ (见习题).)

由(5)式所定义的函数 $\Phi(\theta)$ 是连续的，并且 $\Phi(\theta)$ 在点 $(0, \cdots, 0)$ 与形如 $(\pm\pi, \cdots, \pm\pi)$ 的 2^l 个点上满足 $|\Phi(\theta)|=1$，而在 Q 的其它点上满足 $|\Phi(\theta)|<1$. 因此,根据(6)式,有

$$\begin{aligned}(2\pi)^l g(x, y) &\leqslant \sum_{n=0}^{\infty} \int_Q |\Phi^n(\theta)|\, d\theta \\ &= \int_Q \frac{d\theta}{1-|\Phi(\theta)|}.\end{aligned} \tag{9}$$

因为当 $\alpha\to 0$ 时，$\cos\alpha\sim 1-\dfrac{\alpha^2}{2}$,所以存在点 $0=(0, \cdots, 0)$ 的一个邻域 U，使得 U 中的点 θ 满足

$$0<\cos\theta_m\leqslant 1-\frac{\theta_m^2}{4} \quad (m=1, \cdots, l),$$

再由(5)式知

$$|\Phi(\theta)| = \Phi(\theta) < 1 - \frac{1}{4l}(\theta_1^2 + \cdots + \theta_l^2),$$

因此,对 $l \geq 3$ 有

$$\int_U \frac{d\theta}{1 - |\Phi(\theta)|} < \int_U \frac{4l d\theta}{\theta_1^2 + \cdots + \theta_l^2} < \infty.$$

类似地, 在 $\theta = (\pm\pi, \cdots, \pm\pi)$ 的邻域中也可验证积分(9)的收敛性,故

$$\int_Q \frac{d\theta}{1 - |\Phi(\theta)|} < \infty, \tag{10}$$

这就证明了不等式(8).

由这个不等式可知, 质点到达位置 y 只有限次的概率为 1. 因此,质点占据点格中任意给定的一个位置仅仅有限次数的概率为 1. 因为可列个概率为 1 的事件的交仍是一个概率为 1 的事件, 所以质点不可能无穷多次占据一个单点的概率为 1. 因此,对于点格中的任一有界集, 以概率 1 存在这样的时刻,在这时刻以后质点将不再返回该集.

现在不难证明任意有界集 B 是非常返的. 首先我们假定集 B 是常返的. 于是根据全概率公式 (对任意的初始状态 x 和任意的 n) 事件 $A_n = \{$质点在第 n 步后到达 $B\}$ 的概率为

$$\sum_y p(n, x, y)\pi_B(y) = \sum_y p(n, x, y) = 1.$$

因此, 所有事件 A_n 都发生的概率为 1, 即不论时间有多长质点都会到达 B, 这与质点在某个时刻之后将离开 B 的概率为 1 这一事实相矛盾.

由(9)和(10)式可知,级数

$$e^{i\theta(x-y)} \sum_{n=0}^{\infty} \Phi^n(\theta)$$

可以在 Q 上逐项积分,因此,

$$g(x, y) = \sum_{n=0}^{\infty} \frac{1}{(2\pi)^l}\int_Q e^{i\theta(x-y)}\Phi^n(\theta)d\theta$$

$$= \frac{1}{(2\pi)^l}\int_Q \frac{e^{i\theta(x-y)}}{1-\Phi(\theta)}\, d\theta, \qquad (11)$$

最后的式子使我们可以对函数 $g(x, y)$（当 $l \geqslant 3$ 时）给出一个渐近估计，即

$$g(x, y) \sim \frac{c_l}{\|x-y\|^{l-2}}, \quad 当 \|x-y\| \to \infty, \qquad (12)$$

其中 $\|x\|$ 表示向量 x 的长度，c_l 是某个正数（见附录中的 §1）。下面推导常返的判别法时将要用到这个估计.

§4. 调 和 函 数

设 $f(x)$ 是点格 H^l 上的函数，令

$$Pf(x) = M_x f(X(1)) = \sum_y p(1, x, y)f(y). \qquad (13)$$

很自然地，可称 P 为函数 $f(x)$ 的（一步）**推移算子**.

由于 $p(1, x, x+e_k) = \dfrac{1}{2l}$，于是 P 也是**平均算子**

$$Pf(x) = \frac{1}{2l}\sum f(x+e_k)$$

（其中对 k 求和，k 取 $\pm 1, \cdots, \pm l$，且 $e_{-k} = -e_k$）. 很久以前人们就指出线性算子

$$A = P - E$$

是算子 $\dfrac{1}{2}\Delta$ 的离散类似，其中 P 是平均算子，E 是单位算子，而 Δ 是 Laplace 算子，即

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_l^2}.$$

众所周知，对定义在全空间上充分光滑的函数 $f(x)$，有

$$\Delta f(x)=\lim_{h\to 0}\frac{\sum f(x+he_k)-2lf(x)}{h^2},$$

也就是说，把点格无限细分，通过取极限，可以由算子 $P-E$ 得到 Laplace 算子.

算子 $\frac{1}{2}\Delta$ 和 A 之间的类似有着很深远的意义，由于这种类似性，我们可以把随机徘徊中的许多概念用微分方程理论中相应的术语来描述.

称点格 H^l 上的函数 $f(x)$ 是**调和的**，如果 $Af(x)=0$；称函数 $f(x)$ 是**上调和的**，如果对一切 x 有 $Af(x)\leqslant 0$. 换句话说，如果 $Pf=f$，那么函数 f 是调和的；如果 $Pf\leqslant f$，那么函数 f 是上调和的.

显然，任一常数是调和函数. 我们来证明**任一有界调和函数 f 是一常数**.

如果函数 f 在某一点 y_0 达到它的极大值，那么证明是很容易的. 这时，设 $y_1, y_2, \cdots, y_{2l}$ 是点 y_0 的邻点，由 $Pf(y_0)=f(y_0)$ 知，数 $f(y_0)-f(y_k)$ 的算术平均值等于 0，又因为这些数是非负的，所以它们全等于 0. 因此，使得函数 f 取到极大值的点集，除了包含这些点外，还包含它们的邻点，显然，这样的函数是一常数.

对任一有界函数 φ，存在上确界 M. 一般来说，可能没有一个点能达到这个上确界，但是对任意 $\varepsilon>0$，一定存在一点 y，在这点 y 上有 $\varphi(y)\geqslant M-\varepsilon$. 这时，重复上一段讨论，不难证明，如果 φ 是调和的，那么在 y 的任一邻点 y' 上有不等式 $\varphi(y')\geqslant M-2l\varepsilon$. 因此，如果 $M>0$，我们可以取一连串点，$y_0, y_1=y_0+e_1, \cdots, y_n=y_{n-1}+e_1$，使得和

$$s=\varphi(y_0)+\varphi(y_1)+\cdots+\varphi(y_n)$$

大于任一预先指定的数 N.

现在设 f 是任意有界调和函数，于是函数 $\varphi(x)=f(x+e_1)-f(x)$ 也是有界调和函数，对它来说，和

$$s=f(y_n+e_1)-f(y_0)$$

不超过 f 的上界的两倍，因此，函数 φ 的上确界不能是正的. 这意味着对任意 x，有

$$\varphi(x)=f(x+e_1)-f(x)\leqslant 0.$$

在上述讨论中，可以把向量 e_1 换成向量 $-e_1$，于是有

$$f(x+e_1)=f(x),$$

类似地可以证明，对任意 k 有 $f(x+e_k)=f(x)$.

调和函数的一个例子是函数 $\pi_B(x)$，它是从 x 出发无穷次到达集合 B 的概率. 由于，

$$P\bar{\pi}_B(x)=\sum_y p(1, x, y)\bar{\pi}_B(y)$$

是从 x 出发一步之后再无穷次到达 B 的概率，显然，这个概率等于 $\pi_B(x)$.

因为函数 $\pi_B(x)$ 是有界的，所以根据上面的证明它是一个常数. 我们来证明它或者等于 1 或者等于 0，这由集 B 是常返的还是非常返的来决定. 首先设集 B 是非常返的，我们用 $q(n, y)$ 表示从 x 出发在第 n 步首次到达 B 且到达的位置是 y 这一事件的概率，并且同前一样，用 $\pi_B(x)$ 表示从 x 出发到达 B 的概率，显然，有

$$\pi_B(x)=\sum_{n=0}^{\infty}\sum_{y\in B}q(n, y).$$

为了使质点无穷多次到达 B，就必须使它在某一步第一次到达 B，然后再无穷多次到达 B，按照全概率公式计算这个事件的概率，就得到