疏散的馬尔可夫鏈

B. И. 罗曼諾夫斯基著

科 学 出 版 社

[illegible][illegible]的马尔可夫纳

疏散的馬尔可夫鏈

В. И. 罗曼諾夫斯基 著

梁 文 騏 譯

科 学 出 版 社

В. И. РОМАНОВСКИЙ

ДИСКРЕТНЫЕ ЦЕПИ МАРКОВА

Гостехиздат

Москва Ленинград

1949

內 容 提 要

疏散的馬尔可夫鏈是一般随机过程的一个重要的特殊情形，而其詳尽深入的研究則主要是应用矩陣方法。本書的著者、苏联已故数学家罗曼諾夫斯基(В. И. Романовский)在这一方面有許多創造性的工作。本書系其晚年所著，綜合了其本人及其他研究者在疏散的馬尔可夫鏈方面的許多研究成果。除了1938年 M. Fréchet 的書之外，本書是有系統地、用矩陣方法講述具有有穷个状态及疏散时間的馬尔可夫鏈的唯一专著。

疏散的馬尔可夫鏈

В. И. 罗曼諾夫斯基 著

梁 文 騏 譯

*

科学出版社出版(北京朝陽門大街117号)

北京市書刊出版業營業許可証出字第061号

中国科学院印刷厂印刷 新华書店总經售

*

1958年4月第一版 書号：1091 印張：13 5/8

1958年4月第一次印刷 开本：850×1168 1/32

(京) 0001—1,265 字数：332,000

定价：(10) 2.50 元

序

呈献在讀者面前的这本書，其目的并不是要对馬尔可夫鏈的理論給出某种程度的完整闡述，馬尔可夫鏈的理論在今日已發展得非常广闊，但是还远沒有發展到尽头。在这本書中所講述的仅只是有关具有有穷个状态及疏散时間的鏈的，以及能够用矩陣研究方法得出的那些基本事实。至于讀者可在 M. Fréchet 的概括性著作[43]及許多其他馬尔可夫鏈的研究者的論著中找到的材料，本書中包含得并不多，但在本書中考察了許多 M. Fréchet及其他研究者完全沒有涉及到的問題：双循环鏈与复循环鏈、馬尔可夫–布倫斯(Марков–Брунс)鏈、相关的鏈、复杂鏈、馬尔可夫鏈的随机应用以及其他問題。在本書中，充分注意了鏈理論的創始人、著名俄罗斯数学家 A. A. 馬尔可夫的某些論著及其思路，而这些在概率論方面的数学文献中則还沒有得到足够的估价。

本書的最主要的特色在于推进了鏈的矩陣研究方法，据我看来，在疏散的馬尔可夫鏈理論的研究中，矩陣乃是基本的和最有力的工具。

如果本書能够对于苏維埃科学以及对于本書范疇內的年青的苏維埃研究工作者有所裨益，則著者将至感荣幸。

В. И. Романовский

以 此 紀 念

偉 大 的 导 师

安德烈·安德烈維奇·馬尔可夫

目　录

第一章

一些基本概念与基本定理

1. 状态数目有限并且时間疏散的簡单的均勻馬尔可夫鏈. 我們講述疏散的馬尔可夫(Марков)鏈的理論，是从建立状态数目有限并且时間疏散的簡单的均勻馬尔可夫鏈概念入手。这是疏散的馬尔可夫鏈的最簡单的情形，同时也是最基本的以及在下文中講得頂多的。

讓我們来考察某一体系 S，它具有有限个互不相容的状态：

$$A_1, A_2, \cdots, A_n,$$

每隔有限一段时間之后状态就要变更一次，这种变更遵守如下的随机規律：在某一初始时刻 T_0，这些状态的概率分别等于

$$p_{01}, p_{02}, \cdots, p_{0n},$$

其次，無論我們考察其后的那一个时刻 T_k $(k=0,1,2,\cdots)$，原来处于状态 A_α $(\alpha=\overline{1,n})$ 的体系 S 在下一个时刻 T_{k+1} 轉而呈現状态 A_β $(\beta=\overline{1,n})$ 的概率（当体系 S 在时刻 $T_{k+2}, T_{k+3}, \cdots$ 的状态未定时）恒等于一个不依賴于体系 S 在时刻 $T_0, T_1, \cdots, T_{k-1}$ 的状态的非負常数 $p_{\alpha\beta}$。概率 $p_{\alpha\beta}$ 的这一定义利用通常的条件概率写法就可以写成以下的形状：

$$p_{\alpha\beta}=P(A_\beta, T_{k+1}/A_\alpha, T_k), \qquad k=0,1,2,3\cdots.$$

这样一来，体系 S 在 $T_1, T_2, \cdots$ 等任何一个时刻的状态，与上一个时刻以前的那些时刻的状态無关，而在以后各时刻的状态未定的条件下，被上一个时刻的状态随机地完全确定下来了[1)]。

1) 这就是說得到完全确定的概率。

我們所定义的体系 S 的状态的变更系一般的随机过程的一个特殊情形，称为状态数目有限并且时間疏散的簡单的均匀馬尔可夫鏈，或称狭义的馬尔可夫鏈 (A. H. Колмогоров 命名)。下文中为了簡便我們将称之为鏈 C_n。

我們称矩陣

$$P\equiv \mathrm{Mt}(p_{\alpha\beta})\equiv\begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n}\\ p_{21} & p_{22} & \cdots & p_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ p_{n1} & p_{n2} & \cdots & p_{nn}\end{pmatrix}$$

为鏈 C_n 的規律而称組成这个矩陣的那些概率 $p_{\alpha\beta}$ 为轉移概率；轉移概率須适合以下的显然等式

$$\sum_{\beta} p_{\alpha\beta}=1,\qquad \alpha=\overline{1,n}. \tag{1.1}$$

概率 $p_{0\alpha}\,(\alpha=\overline{1,n})$ 称为初始概率；对于初始概率也有

$$\sum_{\alpha} p_{0\alpha}=1. \tag{1.2}$$

由等式(1.1)与(1.2)显見概率

$$p_{\alpha 1}, p_{\alpha 2}, \cdots, p_{\alpha n}$$

不全是零（無論 $\alpha=\overline{1,n}$ 是什么），而且初始概率也不全是零。

以上定义馬尔可夫鏈，用的是 A. H. Колмогоров 的命名系統。至于 A. A. 馬尔可夫本人所研究的就不是那些状态 A_α，而是一系列試驗中的那些互不相容的事件 A_α，它們由上面所講的体系 S 的状态变更所适合的那些条件联結成为鏈鎖；于是 $p_{0\alpha}$ 就是在初始試驗中事件 A_α 的概率，而 $p_{\alpha\beta}$ 就是在以后某次試驗中，在已知上次試驗中出現事件 A_α 的条件下，出現事件 A_β 的概率。A. H. Колмогоров 的命名系統与 A. A. 馬尔可夫的命名系統相較，具有更富于具体感的优点，并且完全摆脫了隐含在試驗这一概念中的某种主观主义色彩。但是应該注意，在概率論中“試驗”这一个

詞是意味着实現某些条件，在这些条件之下某种事件可能發生，幷沒有包含什么必然的主观因素，它可以只表示不随人的意志为轉移而客观地实現的条件的观察。至于"事件"这个詞，比起"体系的状态"这个詞来也是較为空泛。虽然如此，但是如果抽象地、广义地来了解 A. A. 馬尔可夫以及 A. H. Колмогоров 的命名系統，則可以認为它們是彼此等价的。

为了闡明以上給出的鏈 C_n 的定义，我們再給出其他可能类型的馬尔可夫鏈的定义。

根据体系 S 的状态的集合，可以把鏈区分成状态的集合是有穷的或無穷的两种；在状态的集合是無穷的这一情形下，这集合又可能是可数的或不可数的。然后在这所有的三种情形之下，时間可能是疏散的或是連續的。在时間疏散的情形下，又分成簡单的鏈与复杂的鏈两种，前者的体系在任何一个时刻的状态是由其紧上一个时刻的状态所随机地完全确定，而后者的体系在任何一个时刻的状态則仅仅是在給定了这个时刻以前的有穷多个或無穷多个时刻的体系状态之后，才能随机地完全确定。最后，鏈还可能是均匀的或是不均匀的，对于前者無論我們考察的是那一个时刻，其体系状态的轉移概率的矩陣保持不变，而后者的轉移概率的矩陣則依时刻的改变而改变。

讀者对于这些定义，应該看成只不过是为了要闡明"状态数目有限幷且时間疏散的簡单的均匀馬尔可夫鏈 C_n"在一般的馬尔可夫鏈系統中的位置而預先給出的。不仅如此，为了要使讀者能够更进一步对馬尔可夫鏈在一般随机过程中的位置在心目中有所了解，我們还引入下面的定义；按随机过程的一般理論的首先創立系归功于 A. H. Колмогоров[8]，以下所引入的定义就是依照他的方法来叙述的。

"設 S 是某一体系，它可以处于状态 $x, y, z, \cdots$；而 F 是以 $x, y, z, \cdots$ 为元素所組成的各种集合 $\mathscr{E}$ 的系統。如果对于任意选擇的

状态 x，集合 $\mathscr{E}$，与时刻 $t_1, t_2(t_1<t_2)$，事件'在于时刻 t_1 具有状态 x 的前提之下，于时刻 t_2 出現 $\mathscr{E}$ 中的一个状态'恒有确定的概率 $P(t_1, x, t_2, \mathscr{E})$ 1)，則我們就称体系 S 的变化过程对于 F 而言是随机地确定了。"

由 A. H. Колмогоров 所提出的最早的名詞"随机确定的过程"到后来变成了"無后效的随机过程"。А. Я. Хинчин 建議称这种过程为馬尔可夫过程[42]。

从 A. H. Колмогоров 的定义中可以很清楚地看出，鏈 C_n 是"可能状态 $x, y, z, \cdots$ 的数目有限并且时間疏散的無后效的随机过程"。A. H. Колмогоров 称之为狹义馬尔可夫鏈，并且把所有那些时間疏散的無后效的过程的概型称之为广义馬尔可夫鏈。

所有以上定义出的鏈 C_n 概念的各种扩充，都被包含在無后效的随机过程这一概念之中，同时也就构成了这一概念的全部內容。但是所有这些，就連無后效的随机过程这一个一般概念在內，都只不过是随机过程一般概念的特殊情形。关于随机过程一般概念，А. Я. Хинчин 定义如下[45]：

"随机过程系依賴于一个参数的随机变量 $x_t(-\infty<t<+\infty)$ 的总体；为了要确定这个过程，必須对于任何有限的一組值 $t_1, t_2, \cdots, t_n$ 給出相应的变量 $x_{t_1}, x_{t_2}, \cdots, x_{t_n}$ 的 n 維分布律，而这样定义出来的分布律在其相互关系上应滿足概率論中的全部要求。"

2. 随机矩陣. 在本書所采用的疏散的馬尔可夫鏈的理論的講述系統之中，非負矩陣中被我們称之为随机矩陣的这一类矩陣，占有重要位置。因此一开始我們就先来叙述随机矩陣的某一些为我們后来所需要的性質。

正矩陣与非負矩陣是 G. I. Frobenius 的专門論文[44.1]，[44.2]与[44.3]的主題。关于正矩陣与非負矩陣的很多材料，讀者

1) 如遵照一般的表示法，即給定 A 之后 B 的条件概率記作 $P(B/A)$，則这一条件概率可用 $P(\mathscr{E}, t_2/x, t_1)$ 来表示。

也可以在 Ф. Р. Гантмахер 与 М. Г. Крейн 的書[1]中找到（参閱譯注 1）。

如果一个矩陣中所有的元素全都是非負的，則称这个矩陣是非負的；如果一个矩陣中所有的元素全都是正的，則称这个矩陣是正的。我們要研究的差不多仅只是正方的矩陣，并对 n 級的非負矩陣采用以下記法：

$$A=\mathrm{Mt}(a_{\alpha\beta})=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

如果 n 級的非負矩陣

$$P=\mathrm{Mt}(p_{\alpha\beta})$$

滿足以下两个条件：

1° $\sum\limits_{\beta} p_{\alpha\beta}=1,\ \alpha=\overline{1,n}$；

2° 在矩陣的每一縱列中至少有一个元素 $p_{\alpha\beta}$ 不等于 0；

則我們就称之为随机的。

当矩陣 P 是鏈 C_n 的規律的时候，以上的条件具有簡单的随机意义。根据条件 1°，在矩陣 P 的任何一横行之內不可能所有的概率都等于 0；这就說明在鏈 C_n 之中，在某时刻 T_k 体系 S 無論發生任何状态 A_α，到了下一个时刻 T_{k+1}，它总至少可以轉变成某一个状态。条件 2° 是說矩陣 P 不能有一个 0 列；这就意味着对于体系 S 不存在如下的这种状态：于时刻 $T_1, T_2, \cdots$，它在任何状态之后都不可能發生，因而它只有在初始时刻 T_0 或許可能發生，假如說它的初始概率不是 0 的話。具有这种状态的体系在过了初始时刻之后也就变成不具有这种状态的体系了，因而这种体系沒有什么意思。可見我們对随机矩陣所加的条件 2° 是十分自然的。

3. 非負矩陣的基本性質. 設以 $E=\mathrm{Mt}(e_{\alpha\beta})$ 表示 n 級单位矩

陣。我們称行列式

$$A(\lambda)=|\lambda E-A| \qquad 与 \qquad P(\lambda)=|\lambda E-P|$$

为矩陣 A 与 P 的特征行列式，而称方程

$$A(\lambda)=0 \qquad 与 \qquad P(\lambda)=0$$

为矩陣 A 与 P 的特征方程，至于这些方程的根，則称为矩陣 A 与 P 的根或特征数。

3.I. 如果 $A>0$，則存在数 $r>0$，使得数 r 是矩陣 A 的单根，并且数 r 大过矩陣 A 的所有其他的根的絕对值。

这个数 r 称为矩陣 A 的最大根。

3.II. 如果 $A>0$，則当 $\lambda\geqslant r$ 时，全部 $A_{\alpha\beta}(\lambda)>0$。

在这里以及下文中各处，$A_{\alpha\beta}(\lambda)$ 总是表示行列式 $A(\lambda)$ 的对应于元素 $\lambda e_{\alpha\beta}-a_{\alpha\beta}$ 的(代数余)子式。

我們引进这些定理，但不給出証明1)。

根据連續性原理，从定理 3.I 可以推知，对于非負的矩陣 A 也存在有最大根 r，但这一回可能就不是单根了，并且仅只是不小于矩陣 A 的所有其他的根的模，而当 $\lambda\geqslant r$ 时，現在就是全部子式 $A_{\alpha\beta}(\lambda)\geqslant 0$。

3.III. 如果 $A>0$，則最大根 r 介于以下諸和数

$$a_\alpha=\sum_\beta a_{\alpha\beta} \qquad (\alpha=\overline{1,n})$$

中的最小者与最大者之間：$\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha\leqslant r\leqslant\max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$；特别当 $\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha\neq\max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$ 时，以上不等式中的等号可以取消：$\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha<r<\max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$。

实际上，把行列式 $A(\lambda)$ 的各縱列一齐加到具有附标 β 的那一縱列，然后再将 $A(\lambda)$ 按照这一縱列的元素而展开，我們就得到

$$0=A(r)=\sum_\alpha (r-a_\alpha)A_{\alpha\beta}(r).$$

因为全部 $A_{\alpha\beta}(r)>0$，所以以上等式只有在 $\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha\leqslant r\leqslant\max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$

1) 讀者在[42] 90—92 頁上可以找到簡单的証明。

的情况下才可能成立；而当 $\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha \neq \max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$ 时，以上等式只有在 $\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha < r < \max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$ 的情况下才可能成立。

这个定理可推广如下，这对以后甚为重要。

3. III′. 对于矩陣 $A\geqslant 0$，我們有 $\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha \leqslant r \leqslant \max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$。若 A 不可分解則当 $\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha \neq \max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$ 时，我們有 $\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha < r < \max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$；若 A 可分解則即使 $\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha \neq \max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$，以上不等式中等号仍可能成立。

一个矩陣我們称为是不可分解的，假使它不可能通过行与列的同样調动（这种調动不改变矩陣的根）而化成以下形状：

$$A=\begin{pmatrix} P & O \\ Q & R \end{pmatrix},$$

其中 P 与 R 是两个异于零的正方子陣，Q 一般也是一个异于零的子陣，但有可能是零子陣，而 O 則是零子陣。反之一个矩陣我們称为是可分解的，假使它可以通过行与列的同样調动而化成以上形状。

这定理的証明完全类似于定理 3. III 的証明，并且主要是根据非負不可分解矩陣 A 的一項性質，这項性質是說 A 的特征行列式的所有子式 $A_{\alpha\beta}(\lambda)$ 当 $\lambda\geqslant r$ 时都是正的。这一性質在下文中将加以証明。据此立即可以看出，当矩陣 $A\geqslant 0$ 并且不可分解时，若 $\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha \neq \max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$，則我們仍有 $\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha < r < \max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$。但若 A 可分解，則即使 $\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha \neq \max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$，仍然可能 $r=\max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$ 或 $r=\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha$，为要实現这一点，例如只須令矩陣 P 的每一横行元素的和数 a_α 皆等于 $\max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$ 就行了。

3. IV. 为了要使非負矩陣 A 是可分解的，必須而且只須，当 $\lambda=r$ 时行列式 $A(\lambda)$ 的諸主子式之中有一个等于零[1]。

3. V. 如果 $A\geqslant 0$ 是不可分解的，那么当 $\lambda\geqslant r$ 时，全部子式 $A_{\alpha\beta}(\lambda)>0$。

1）这个定理的証明可在 G. I. Frobenius 的[44.3]第 459 頁找到。