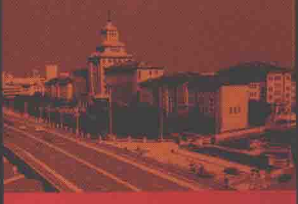

俄罗斯数学精品译丛

"十二五"国家重点图书

Scattered Markov Chain

疏散的马尔可夫链

[苏]罗曼诺夫斯基 著 梁文骐 译

哈爾濱工業大學出版社
HARBIN INSTITUTE OF TECHNOLOGY PRESS

俄罗斯数学精品译丛

"十二五"国家重点图书

疏散的马尔可夫链

Scattered Markov Chain

[苏] 罗曼诺夫斯基 著

梁文骐 译

HITP

哈爾濱工業大學出版社

HARBIN INSTITUTE OF TECHNOLOGY PRESS

内 容 简 介

疏散的马尔可夫链是一般随机过程的一个重要的特殊情形，而其详尽深入的研究则主要是应用矩阵方法. 本书的著者、前苏联已故数学家罗曼诺夫斯基(В. И. Романовский)在这一方面有许多创造性的工作. 本书系其晚年所著，综合了其本人及其他研究者在疏散的马尔可夫链方面的许多研究成果. 除了1938年M. Fréchet的书之外，本书是有系统地、用矩阵方法讲述具有有穷个状态及疏散时间的马尔可夫链的唯一专著.

图书在版编目(CIP)数据

疏散的马尔可夫链/(苏)罗曼诺夫斯基著；梁文骐译. —哈尔滨：哈尔滨工业大学出版社，2014.1

(俄罗斯数学精品译丛)

ISBN 978-7-5603-4047-0

Ⅰ.①疏… Ⅱ.①罗…②梁… Ⅲ.①马尔柯夫链 Ⅳ.①O211.62

中国版本图书馆CIP数据核字(2013)第073531号

策划编辑 张 佳 宋晓翠
责任编辑 张永芹 王 慧
封面设计 孙茵艾
出版发行 哈尔滨工业大学出版社
社 址 哈尔滨市南岗区复华四道街10号 邮编 150006
传 真 0451-86414749
网 址 http://hitpress.hit.edu.cn
印 刷 哈尔滨工业大学印刷厂
开 本 787mm×960mm 1/16 印张 21.25 字数 382千字
版 次 2014年1月第1版 2014年1月第1次印刷
书 号 ISBN 978-7-5603-4047-0
定 价 58.00元

◎ 序

呈献在读者面前的这本书,其目的并不是要对马尔可夫链的理论给出某种程度的完整阐述,马尔可夫链的理论在今日已发展得非常广阔,但是还远没有发展到尽头. 在这本书中所讲述的只是有关具有有穷个状态及疏散时间的链的,以及能够用矩阵研究方法得出的那些基本事实. 至于读者可在 M. Fréchet 的概括性著作[43] 及许多其他马尔可夫链的研究者的论著中找到的材料,本书中包含得并不多,但在本书中考察了许多 M. Fréchet 及其他研究者完全没有涉及的问题:双循环链与复循环链、马尔可夫－布伦斯(Марков－Брунс)链、相关的链、复杂链、马尔可夫链的随机应用以及其他问题. 在本书中,充分注意了链理论的创始人、著名俄罗斯数学家 A·A·马尔可夫的某些论著及其思路,而这些在概率论方面的数学文献中还没有得到足够的估价.

本书的最主要的特色在于推进了链的矩阵研究方法,据我看来,在疏散的马尔可夫链理论的研究中,矩阵乃是基本的和最有力的工具.

如果本书能够对于苏维埃科学以及对于本书范畴内的年青的苏维埃研究工作者有所裨益,则著者将至感荣幸.

В. И. Романовский

◎ 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间疏散的简单的均匀马尔可夫链

我们讲述疏散的马尔可夫(Марков)链的理论,从建立状态数目有限并且时间疏散的简单的均匀马尔可夫链概念入手.这是疏散的马尔可夫链的最简单的情形,同时也是最基本的以及在下文中讲得最多的.

让我们来考察某一体系 S,它具有有限个互不相容的状态

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

每隔有限一段时间之后状态就要变更一次,这种变更遵守如下的随机规律: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,无论我们考察其后的哪一个时刻 $T_k(k=0,1,2,\cdots)$,原来处于状态 $A_\alpha(\alpha=\overline{1,n})$ 的体系 S 在下一个时刻 T_{k+1} 转而呈现状态 $A_\beta(\beta=\overline{1,n})$ 的概率(当体系 S 在时刻 $T_{k+2}, T_{k+3}, \cdots$ 的状态未定时)恒等于一个不依赖于体系 S 在时刻 $T_0, T_1, \cdots, T_{k-1}$ 的状态的非负常数 $p_{\alpha\beta}$. 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

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

这样一来,体系 S 在 $T_1, T_2, \cdots$ 中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无关,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 这就是说得到完全确定的概率.

我们所定义的体系 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间疏散的简单的均匀马尔可夫链，或称狭义的马尔可夫链（A. H. Колмогоров 命名）. 下文中为了简便，我们称之为链 C_n.

我们称矩阵

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

为链 C_n 的规律，而称组成这个矩阵的那些概率 $p_{\alpha\beta}$ 为转移概率. 转移概率须适合以下的显然等式

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

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) 显见概率

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

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

以上定义马尔可夫链，用的是 A. H. Колмогоров 的命名系统. 至于 A·A·马尔可夫本人所研究的就不是那些状态 A_α，而是一系列试验中的那些互不相容的事件 A_α，它们由上面所讲的体系 S 的状态变更所适合的那些条件联结成为链锁；于是 $p_{0\alpha}$ 就是在初始试验中事件 A_α 的概率，而 $p_{\alpha\beta}$ 就是在以后某次试验中，在已知上次试验中出现事件 A_α 的条件下，出现事件 A_β 的概率. A. H. Колмогоров 的命名系统与 A·A·马尔可夫的命名系统相较，具有更富于具体感的优点，并且完全摆脱了隐含在试验这一概念中的某种主观主义色彩. 但是应该注意，在概率论中“试验”这一个词是意味着实现某些条件，在这些条件之下某种事件可能发生，并没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转移而客观地实现的条件的观察. 至于“事件”这个词，比起“体系的状态”这个词来也是较为空泛. 虽然如此，但是如果抽象地、广义地来了解 A·A·马尔可夫以及 A. H. Колмогоров 的命名系统，则可以认为它们是彼此等价的.

为了阐明以上给出的链 C_n 的定义，我们再给出其他可能类型的马尔可夫链的定义.

根据体系 S 的状态的集合，可以把链区分成状态的集合是有穷的或无穷的两种；在状态的集合是无穷的这一情形下，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，时间可能是疏散的或是连续的. 在时间疏散的情形下，又分成简单的链与复杂的链两种，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则仅仅是在给定了这个时刻以前的有穷多个或无穷多个时刻的体系状态之后，才能随机地完全确定. 最后，链还可能是均匀的或是不均匀的，对于前者无论我们考察的是哪一个时刻，其体系状态的转移概率的矩阵保持不变，而后者的转移概率的矩阵则依时刻的改变而改变.

读者对于这些定义，应该看成只不过是为了要阐明"状态数目有限并且时间疏散的简单的均匀马尔可夫链 C_n"在一般的马尔可夫链系统中的位置而预先给出的. 不仅如此，为了要使读者能够更进一步对马尔可夫链在一般随机过程中的位置在心目中有所了解，我们还引入下面的定义；随机过程的一般理论的创立系归功于 A. H. Колмогоров[8]，以下所引入的定义就是依照他的方法来叙述的.

"设 S 是某一体系，它可以处于状态 $x,y,z,\cdots$；而 F 是以 $x,y,z,\cdots$ 为元素所组成的各种集合 $\mathscr{E}$ 的系统. 如果对于任意选择的状态 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 出现 $\mathscr{E}$ 中的一个状态'恒有确定的概率 $P(t_1,x,t_2,\mathscr{E})$① 则我们就称体系 S 的变化过程对于 F 而言是随机地确定了."

由 A. H. Колмогоров 所提出的最早的名词"随机确定的过程"到后来变成了"无后效的随机过程". A. Я. Хинчин 建议称这种过程为马尔可夫过程[42].

从 A. H. Колмогоров 的定义中可以很清楚地看出，链 C_n 是"可能状态 $x,y,z,\cdots$ 的数目有限并且时间疏散的无后效的随机过程". A. H. Колмогоров 称之为狭义马尔可夫链，并且把所有那些时间疏散的无后效的过程的概型称之为广义马尔可夫链.

所有以上定义出的链 C_n 概念的各种扩充，都被包含在无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内容. 但是所有这些，就连无后效的随机过程这一个一般概念在内，都只不过是随机过程一般概念的特殊情形. 关于随机过程的一般概念，A. Я. Хинчин 定义如下[45]：

"随机过程系依赖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，必须对于任何有限的一组值 $t_1,t_2,\cdots,t_n$ 给出相应的变量 $x_{t_1},x_{t_2},\cdots,x_{t_n}$ 的 n 维分布律，而这样定义出来的分布律在其相互关系上应满足概率论中的全部要求."

2. 随机矩阵　在本书所采用的疏散的马尔可夫链的理论的讲述系统之中，非负矩阵中被我们称之为随机矩阵的这一类矩阵，占有重要位置. 因此一开

① 如遵照一般的表示法，即给定 A 之后 B 的条件概率记作 $P(B/A)$，则这一条件概率可用 $P(\mathscr{E},t_2/x,t_1)$ 来表示.

始我们就先来叙述随机矩阵的某一些为我们后来所需要的性质.

正矩阵与非负矩阵是G. I. Frobenius的专门论文[44.1],[44.2]与[44.3]的主题. 关于正矩阵与非负矩阵的很多材料,读者也可以在 Ф. Р. Гантмахер 与 М. Г. Крейн 的书[1] 中找到(参阅译注 1).

如果一个矩阵中所有的元素全都是非负的,则称这个矩阵是非负的;如果一个矩阵中所有的元素全都是正的,则称这个矩阵是正的. 我们要研究的差不多只是正方的矩阵,并对 n 级的非负矩阵采用以下记法

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

如果 n 级的非负矩阵

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

满足以下两个条件:

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

2° 在矩阵的每一纵列中至少有一个元素 $p_{\alpha\beta}$ 不等于 0.

则我们就称之为随机的.

当矩阵 $\boldsymbol{P}$ 是链 C_n 的规律的时候,以上的条件具有简单的随机意义. 根据条件 1°,在矩阵 $\boldsymbol{P}$ 的任何一横行之内不可能所有的概率都等于 0. 这就说明在链 C_n 之中,在某时刻 T_k 体系 S 无论发生任何状态 A_α,到了下一个时刻 T_{k+1},它总至少可以转变成某一个状态. 条件 2° 是说矩阵 $\boldsymbol{P}$ 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发生,因而它只有在初始时刻 T_0 或许可能发生,假如说它的初始概率不是 0 的话. 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,因而这种体系没有什么意思. 可见我们对随机矩阵所加的条件 2° 是十分自然的.

3. 非负矩阵的基本性质 设以 $\boldsymbol{E} = \mathrm{Mt}(e_{\alpha\beta})$ 表示 n 级单位矩阵. 我们称行列式

$$A(\lambda) = |\lambda\boldsymbol{E} - \boldsymbol{A}| \text{ 与 } P(\lambda) = |\lambda\boldsymbol{E} - \boldsymbol{P}|$$

为矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{P}$ 的特征行列式,而称方程

$$A(\lambda) = 0 \text{ 与 } P(\lambda) = 0$$

为矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{P}$ 的特征方程,至于这些方程的根,则称为矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{P}$ 的根或特征数.

3. Ⅰ 如果 $\boldsymbol{A} > 0$,则存在数 $r > 0$,使得数 r 是矩阵 $\boldsymbol{A}$ 的单根,并且数 r 大过矩阵 $\boldsymbol{A}$ 的所有其他的根的绝对值.

这个数 r 称为矩阵 $\boldsymbol{A}$ 的最大根.

3. Ⅱ 如果 $\boldsymbol{A}>0$,则当 $\lambda \geqslant r$ 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.

我们引进这些定理,但不给出证明[①].

根据连续性原理,从定理 3. Ⅰ 可以推知,对于非负的矩阵 $\boldsymbol{A}$ 也存在最大根 r,但这一回可能就不是单根了,并且只是不小于矩阵 $\boldsymbol{A}$ 的所有其他的根的模,而当 $\lambda \geqslant r$ 时,现在就是全部子式 $A_{\alpha\beta}(\lambda)\geqslant 0$.

3. Ⅲ 如果 $\boldsymbol{A}>0$,则最大根 r 介于以下诸和数

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

中的最小者与最大者之间:$\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha \leqslant r \leqslant \max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$;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 $A(\lambda)$ 的各纵列一齐加到具有附标 β 的那一纵列,然后再将 $A(\lambda)$ 按照这一纵列的元素而展开,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 $\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha \leqslant r \leqslant \max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$ 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′ 对于矩阵 $\boldsymbol{A}\geqslant 0$,有 $\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha \leqslant r \leqslant \max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$. 若 $\boldsymbol{A}$ 不可分解则当 $\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha \neq \max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$ 时, 有 $\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha < r < \max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$; 若 $\boldsymbol{A}$ 可分解则即使 $\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha \neq \max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$,以上不等式中等号仍可能成立.

一个矩阵我们称为是不可分解的,假使它不可能通过行与列的同样调动(这种调动不改变矩阵的根)而化成以下形状

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

其中 $\boldsymbol{P}$ 与 $\boldsymbol{R}$ 是两个异于零的正方子阵,$\boldsymbol{Q}$ 一般也是一个异于零的子阵,但有可能是零子阵,而 $\boldsymbol{O}$ 则是零子阵. 反之一个矩阵我们称为是可分解的,假使它可以通过行与列的同样调动而化成以上形状.

这定理的证明完全类似于定理 3. Ⅲ 的证明,并且主要是根据非负不可分

① 读者在[42]90－92 页上可以找到简单的证明.

解矩阵 **A** 的一项性质，这项性质是说 **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. 这一性质在下文中将加以证明. 据此立即可以看出，当矩阵 $\boldsymbol{A} \geqslant 0$ 并且不可分解时，若 $\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha \neq \max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$，则我们仍有 $\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha < r < \max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$. 但若 **A** 可分解，则即使 $\min\limits_{1\leqslant\alpha\leqslant n} a_\alpha \neq \max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$，仍然可能 $r = \max\limits_{1\leqslant\alpha\leqslant n} a_\alpha$ 或 $r = \min\limits_{1\leqslant\alpha\leqslant n} a_\alpha$，为要实现这一点，例如只须令矩阵 **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　为了要使非负矩阵 **A** 是可分解的，必须而且只须，当 $\lambda = r$ 时行列式 $A(\lambda)$ 的诸主子式之中有一个等于零[①].

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

实际上，根据前一定理，非负的不可分解的矩阵 **A** 的所有子式 $A_{\alpha\alpha}(\lambda)(\alpha = \overline{1,n})$ 在 $\lambda = r$ 时皆不等于零，因而当 $\lambda = r$ 时皆是正的. 但是我们有以下的恒等式(参阅译注 3)

$$A(\lambda)A_{\alpha\beta,\alpha\beta}(\lambda) = A_{\alpha\alpha}(\lambda)A_{\beta\beta}(\lambda) - A_{\alpha\beta}(\lambda)A_{\beta\alpha}(\lambda)$$

其中 $A_{\alpha\beta|\alpha\beta}(\lambda)$ 表示行列式 $A(\lambda)$ 的二级子式. 由此立即推知，当 $\lambda = r$ 时，所有子式 $A_{\alpha\beta}(\lambda) > 0$，因而当 $\lambda > r$ 时，亦有 $A_{\alpha\beta}(\lambda) > 0$.

这样一来，定理 3. Ⅲ′ 的证明中我们所引据的非负的不可分解的矩阵 **A** 的那一项性质就得到了证明.

4. 随机矩阵的基本性质　现在我们来证明随机矩阵的若干基本定理.

4. Ⅰ　每一随机矩阵 **P** 皆具有最大根 $\lambda_0 = 1$；如果 $\boldsymbol{P} > 0$ 则这个最大根是单根并且大于所有其他根的模，同时行列式 $P(\lambda)$ 的所有子式 $P_{\alpha\beta}(\lambda)$ 对于 $\lambda \geqslant 1$ 皆是正的；而如果 $\boldsymbol{P} \geqslant 0$，则 $P_{\alpha\beta}(\lambda)$ 对于 $\lambda \geqslant 1$ 皆是非负的.

这个定理乃是定理 3. Ⅰ ～ 3. Ⅲ 的直接推论. 实际上，因为

$$\sum_\beta p_{\alpha\beta} = 1, \alpha = \overline{1,n}$$

所以根据定理 3. Ⅰ，矩阵 **P** 存在有最大根，而根据定理 3. Ⅲ，它等于 1. 定理 4. Ⅰ 的其余断言可以从定理 3. Ⅰ 与 3. Ⅱ 推出.

4. Ⅱ　如果随机矩阵 **P** 是可分解的，并化成了以下形状

$$\boldsymbol{P} = \begin{pmatrix} \boldsymbol{Q} & \boldsymbol{O} \\ \boldsymbol{R} & \boldsymbol{S} \end{pmatrix}$$

其中 $\boldsymbol{R} \neq \boldsymbol{O}$，而 **S** 是不可分解的，则矩阵 **S** 的最大根小于 1.

实际上，在这种情形下，与矩阵 **S** 相对应的，即只取 S 的各横行元素而作成的诸和数

---

① 这个定理的证明可在 G. I. Frobenius 的[44.3] 第 459 页找到.

$$\sum_{\beta} p_{\alpha\beta}$$

之中,有小于1的并且任何一个皆不大于1.因此根据定理3.Ⅲ′推知我们的断言成立.

定理4.Ⅰ 也可以很容易地直接加以证明.以下我们就给出直接证明.

不论矩阵 $\boldsymbol{P}$ 是正的还是非负的,把 $P(\lambda)$ 中除第一纵列以外各纵列都加到第一纵列上之后,我们总可写出

$$P(\lambda)=\begin{vmatrix} \lambda-1 & -p_{12} & \cdots & -p_{1n} \\ \lambda-1 & \lambda-p_{22} & \cdots & -p_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda-1 & -p_{n2} & \cdots & \lambda-p_{nn} \end{vmatrix}$$

由此显见,$\lambda_0=1$ 是 $\boldsymbol{P}$ 的根.

然后我们取线性方程组

$$\lambda x_{\beta}=\sum_{\alpha} x_{\alpha} p_{\alpha\beta} \quad (\beta=\overline{1,n}) \tag{4.1}$$

这个方程组我们称之为 $\boldsymbol{P}$ 的第一共轭线性方程组①.假如 λ' 是 $\boldsymbol{P}$ 的根,则方程组(4.1)有非零解 x'_{β},$\beta=\overline{1,n}$,于是我们就可写出

$$|\lambda'|\sum_{\beta}|x'_{\beta}|\leqslant\sum_{\alpha}|x'_{\alpha}|\sum_{\beta}p_{\alpha\beta}=\sum_{\alpha}|x'_{\alpha}|$$

因此 $|\lambda'|\leqslant 1$,即推得 $\lambda_0=1$ 是矩阵 $\boldsymbol{P}$ 的最大根.

现在设 $\boldsymbol{P}>0$.那么方程组(4.1)就有非零解 x_{β}^{0},$\beta=\overline{1,n}$,其中所有 $x_{\beta}^{0}\neq 0$ 并且具有相同的符号.为了要证明这一断言,我们由方程组(4.1)导出另一个方程组,这个方程组在以后也还是有用的.

假定 $\lambda=|\lambda|(\cos\theta+\mathrm{i}\sin\theta)$ 是矩阵 $\boldsymbol{P}$ 的根,并且

$$x_{\beta}=|x_{\beta}|(\cos\theta_{\beta}+\mathrm{i}\sin\theta_{\beta}) \quad (\beta=\overline{1,n})$$

是相应的方程组(4.1)的非零解.于是从方程组(4.1)即得出以下等式

$$|\lambda||x_{\beta}|\cos(\theta+\theta_{\beta})=\sum_{\alpha}|x_{\alpha}|p_{\alpha\beta}\cos\theta_{\alpha}$$

$$|\lambda||x_{\beta}|\sin(\theta+\theta_{\beta})=\sum_{\alpha}|x_{\alpha}|p_{\alpha\beta}\sin\theta_{\alpha}$$

由此推得一个新的方程组

$$|\lambda||x_{\beta}|=\sum_{\alpha}|x_{\alpha}|p_{\alpha\beta}\cos(\theta_{\alpha}-\theta_{\beta}-\theta),\beta=\overline{1,n} \tag{4.2}$$

这就是我们所要的.

---

① 我们称以下方程组为 $\boldsymbol{P}$ 的第二共轭线性方程组

$$\lambda y_{\alpha}=\sum_{\beta}p_{\alpha\beta}y_{\beta},\alpha=\overline{1,n}$$

因为 $p_{\alpha\beta}$ 皆是实数，所以当 $\lambda=1,\theta=0$ 时，我们可以把 x_β^0 皆取成实数，也就是说可以把 θ_α 皆取成 0 或 π. 于是我们就应该有

$$\cos(\theta_\alpha-\theta_\beta-\theta)=\pm 1$$

但当 $\lambda=1$ 时由(4.2) 可得

$$\sum_\beta |x_\beta^0|=\sum_\alpha |x_\alpha^0| \sum_\beta p_{\alpha\beta}\cos(\theta_\alpha-\theta_\beta)$$

所以所有的 $\cos(\theta_\alpha-\theta_\beta)$ 全都等于 1，亦即所有的 θ_α 或是全等于 0 或是全等于 π；由此推知所有的 x_β^0(不全是零) 具有同样的符号. 最后我们注意，任何一个 x_β 都不可能等于零；因为如若不然，那么对于某个 $x_\beta^0=0$ 我们就有

$$0=x_\beta^0=\sum_\alpha x_\alpha^0 p_{\alpha\beta}$$

因而所有的 x_β^0 就都等于零，但是由于 $x_\beta^0,\beta=\overline{1,n}$ 是非零解，所以这是不可能的.

这样一来，当 $\lambda=1$ 时，(4.1) 的非零解中的 x_β^0 全都是不等于零的实数并且还具有相同的符号. 因为(4.1) 的非零解可以有任意的常系数，所以所有的 x_β^0 可以认为皆是正的.

从适才所得到的结果可以得出两个推论：对于 $\boldsymbol{P}>0,\lambda_0=1$ 是单根，并且当 $\lambda\geqslant 1$ 时，所有的子式 $P_{\alpha\beta}(\lambda)>0$.

实际上，我们有

$$x_1^0 : x_2^0 : \cdots : x_n^0=P_{1\beta}(1) : P_{2\beta}(1) : \cdots : P_{n\beta}(1) \tag{4.3}$$

其中 $\beta=\overline{1,n}$. 另一方面，$\boldsymbol{P}$ 的第二共轭线性方程组

$$\lambda y_\alpha=\sum_\beta p_{\alpha\beta}y_\beta,\alpha=\overline{1,n}$$

当 $\lambda=1$ 时有显而易见的解

$$y_1^0=y_2^0=\cdots=y_n^0=1$$

这就推出

$$P_{\alpha1}(1)=P_{\alpha2}(1)=\cdots=P_{\alpha n}(1) \tag{4.4}$$

其中 $\alpha=\overline{1,n}$. 由于等式(4.3) 与(4.4)，同时由于所有的 $x_\beta^0\neq 0$ 并且具有相同的符号，即可推知所有的 $P_{\alpha\beta}(1)(\alpha,\beta=\overline{1,n})$ 皆不为零并且具有相同的符号.

因为

$$\frac{\mathrm{d}P(\lambda)}{\mathrm{d}\lambda}=\sum_\alpha P_{\alpha\alpha}(\lambda)$$

所以由上可知当 $\lambda=1$ 时有

$$\frac{\mathrm{d}P(\lambda)}{\mathrm{d}\lambda}\neq 0$$

这就推出 $\lambda_0=1$ 是矩阵 $\boldsymbol{P}$ 的单根.

最后,如果 $\boldsymbol{P}>0$,则除 1 以外 $\boldsymbol{P}$ 的所有的根的模全都小于 1,并且当 $\lambda\geqslant 1$ 时所有的子式 $P_{\alpha\beta}(\lambda)>0$.

先来证明第一句话.我们用反证法,假定有一个根 $\lambda_1\neq 1$ 而 $|\lambda_1|=1$. 对于 λ_1, P 的第一共轭线性方程组应具有非零解 x'_β, $\beta=\overline{1,n}$,我们假定其中 $x'_h\neq 0$. 于是方程组(4.2) 只有在条件

$$\sum_\beta p_{h\beta}\cos(\theta_h-\theta_\beta-\theta)=1$$

之下方能成立,因为如若不然,则恒等式

$$\sum_\beta |x'_\beta|=\sum_\alpha |x'_\alpha| \sum_\beta p_{\alpha\beta}\cos(\theta_\alpha-\theta_\beta-\theta)$$

就不可能成立了.由以上的条件立即推知,对于所有的 $x'_h\neq 0$,我们应该有

$$\theta_h-\theta_\beta-\theta=0, \beta=\overline{1,n}$$

因此应该有 $\theta=0$,亦即 $\lambda_1=1$,这就和我们的假定矛盾了.由此推知,当 $\boldsymbol{P}>0$ 时,除去 1 以外矩阵 $\boldsymbol{P}$ 的所有的根的模确实都小于 1.

第二句话的证明如下. 我们来考察一个方程 $P_{\alpha\alpha}(\lambda)=0$, 例如就是 $P_{11}(\lambda)=0$. 假设 λ_1 是这方程的一个根,并且 $\lambda_1\neq 0$. 方程组

$$\lambda_1 x_\beta=\sum_{\alpha=2}^{x} x_\alpha p_{\alpha\beta}\quad(\beta=\overline{2,n})$$

应具有非零解 x'_β, $\beta=\overline{2,n}$. 利用这个解我们就得到了以下的不等式

$$|\lambda_1| \sum_\beta |x'_\beta| \leqslant \sum_\alpha |x'_\alpha| \sum_\beta p_{\alpha\beta}<\sum_\alpha |x'_\alpha|$$

(因为 $\boldsymbol{P}>0$,所有的和数

$$\sum_\beta p_{\alpha\beta}\quad(\alpha,\beta=\overline{2,n})$$

都小于1) 由此显见,方程 $P_{11}(\lambda)=0$ 的所有的根的模皆小于 1. 由此还可推出当 $\lambda\geqslant 1$ 时, $P_{11}(\lambda)>0$. 因为假使对某 $\lambda'\geqslant 1$ 有 $P_{11}(\lambda')<0$,则因

$$\lim_{\lambda\to+\infty} P_{11}(\lambda)=+\infty$$

所以根据 $P_{11}(\lambda)$ 的连续性,在 $(1,+\infty)$ 中 $P_{11}(\lambda)$ 必定有根,这与以上所得结果不符. 此时可知当 $\lambda\geqslant 1$ 时,所有的 $P_{\alpha\beta}(\lambda)>0$,因为前面我们已经看到,所有的 $P_{\alpha\beta}(1)\neq 0$ 且符号相同(参阅译注 2).

为要完成定理 4. Ⅰ 的证明,只须指出,如果矩阵 $\boldsymbol{P}$ 仅仅是非负的,那么根据连续性原理,当 $\lambda\geqslant 1$ 时,子式 $P_{\alpha\beta}(\lambda)$ 就也是非负的.

这样一来,定理 4. Ⅰ 已不依赖于定理 3. Ⅰ ~ 3. Ⅲ 而完全得证. 不过定理 4. Ⅰ 的证明还可以大大简化,假如我们引据下面这简单而重要的 C. A. Гершгорин[2] 定理:

任何一个矩阵

$$\boldsymbol{A}=\mathrm{Mt}(a_{\alpha\beta})\quad(\alpha,\beta=\overline{1,n})$$

的根都在一个闭域 G 上,这个闭域 G 是由以 $a_{\alpha\alpha}$ 为心,以

$$R_\alpha=\sum_{\beta\neq\alpha}|a_{\alpha\beta}|$$

为半径的诸圆所组成的(参阅译注 3).

从这一定理立刻可以看出,矩阵 $\boldsymbol{P}$ 所有的根的模不大于 1,并且如果 $\boldsymbol{P}>0$,则除去 1 以外其所有的根的模皆小于 1.

4. Ⅲ 欲使不可分解的矩阵 $\boldsymbol{P}$ 具有异于 1 而模等于 1 的根,必须其对角线上所有的项 $p_{\alpha\alpha}$ 皆等于零.

4. Ⅳ 如果 $\boldsymbol{P}\geqslant 0$,则行列式 $P(\lambda)$ 的所有的各级主子式当 $\lambda\geqslant 1$ 时皆是非负的.

实际上,我们随便取出这个行列式的某一个若干级的主子式,例如

$$P_{\alpha\beta\cdots\delta|\alpha\beta\cdots\delta}(\lambda)$$

然后对它作出第一共轭线性方程组.借助于这方程组就不难推论出这个子式的根的模不大于1,但因这个子式对于充分大的 λ 是正的,所以显而易见,当 $\lambda\geqslant 1$ 时这个子式不小于零.

4. Ⅴ 欲使 $\lambda_0=1$ 是矩阵 $\boldsymbol{P}$ 的 m 重根,必须而且只须,行列式 $P(\lambda)$ 的所有的 $n-1$ 级,$n-2$ 级,…,$n-m+1$ 级子式当 $\lambda=1$ 时皆等于零,而 $n-m$ 级子式中至少有一个当 $\lambda=1$ 时不等于零.

根据以下的显然等式

$$\frac{\mathrm{d}^r P(\lambda)}{\mathrm{d}\lambda^r}=\sum_{\alpha_1\cdots\alpha_r}P_{\alpha_1\alpha_2\cdots\alpha_\gamma|\alpha_1\alpha_2\cdots\alpha_\gamma}(\lambda),r=1,2,\cdots$$

并且根据定理 4. Ⅳ,即可推知这个定理是正确的.

4. Ⅵ 欲使矩阵 $\boldsymbol{P}$ 是可分解的,必须而且只须,当 $\lambda=1$ 时行列式 $P(\lambda)$ 的主子式中有一个等于零.

这个定理乃是定理 3. Ⅳ(Frobenius) 的简单推论,但是可以不依赖于定理 3. Ⅳ 而加以证明.

假若矩阵 $\boldsymbol{P}$ 是可分解的,那么它就可以表示成以下形状

$$\boldsymbol{P}=\begin{pmatrix}\boldsymbol{Q} & \boldsymbol{O}\\ \boldsymbol{R} & \boldsymbol{S}\end{pmatrix}$$

其中 $\boldsymbol{Q}$ 与 $\boldsymbol{S}$ 是正方矩阵,$\boldsymbol{O}$ 是零矩阵,$\boldsymbol{R}$ 是非零矩阵但也可能是零矩阵.矩阵 $\boldsymbol{P}$ 的这种形状可以借助于行与列的同样调动而得出,如所熟知,行与列的同样调动并不改变矩阵的特征行列式.把 $\boldsymbol{P}$ 表示成以上形状后,$\boldsymbol{P}$ 的特征行列式即具有如下的形状

$$P(\lambda)=Q(\lambda)S(\lambda)$$

由此立刻就可以看出定理 4. Ⅵ 的条件的必要性.