arXiv:2301.08181v1 [math.CO] 19 Jan 2023

非负矩阵的组合性质与代数性质

Jenish C.
Mehta 的学位论文

部分满足计算机科学博士学位要求



加州理工学院 帕萨迪纳，加利福尼亚州

答辩日期：2021年12月8日

© 2022年
Jenish C. Mehta 保
留所有权利

摘要

我们研究非负矩阵的组合性质和代数性质。我们的结果分为三个不同的类别。

1. 我们给出了对已有 100 年历史的佩龙-弗罗贝尼乌斯定理的首个定量推广。该定理是一项基础定理，已应用于数学的多个领域。佩龙-弗罗贝尼乌斯定理指出，每个不可约非负矩阵 R 都有一个最大的正特征值 r，且 R 的其他每个特征值 λ 都满足 $\mathrm{Re}\lambda < r$，并且 $|\lambda| \le r$。我们通过广为研究的边扩张概念 ϕ 来刻画不可约性；直观而言，该概念衡量 R 的底层有向图 R 的连通程度。我们还证明谱隙 $\Delta = 1 - \mathrm{Re}\lambda/r$（其中 $\lambda \neq r$ 是 R 中实部最大的特征值）与边扩张 ϕ 之间存在如下定量关系。

$$\frac{1}{15} \cdot \frac{\Delta(R)}{n} \le \phi(R) \le \sqrt{2 \cdot \Delta(R)}.$$

此外，由于该结果适用于任意非负矩阵，因此比切格-布泽不等式更具一般性。

2. 我们研究具有小边扩张但大谱隙的特定非对称矩阵（或非可逆马尔可夫链）的构造。这使我们的研究方向比研究边扩张为常数且已得到广泛研究的对称矩阵更新颖、也更少有人探索。我们首先分析一些已知但研究较少的马尔可夫链，然后提出一种新颖的非可逆链构造，对于这种链，

$$\phi(R) \le \frac{\Delta(R)}{\sqrt{n}},$$

得到的界比已知界好得多，且改进幅度呈指数级。我们还提出了一种候选矩阵构造，使得

$$\phi(R) \le 2\frac{\Delta(R)}{n},$$

这是本学位论文中最精彩的成果。我们认为，这些矩阵具有足够突出的性质，值得单独研究。

3. 我们将边扩张和谱隙与非对称矩阵的其他组合性质联系起来。研究最为广泛的性质是混合时间，我们对已知结果以及将混合

时间、边扩张、谱隙和加性对称化的谱隙联系起来的结果给出初等证明，并提出新结果。此外，我们统一阐释容量和归一化容量的概念，并证明非可逆链的容量单调性，以及关于非对称矩阵的其他结果。最后，我们讨论并证明关于不同扩张概念的有趣引理，并展示关于张量游走或非负张量的首批结果。

致谢

决定攻读博士学位花了我相当长的时间。作出这个决定后，生命中许多人以各种方式、从不同方面给予了我支持，对此我深怀感激。

从印度申请不同的学校，却没有机会与这些学校的人实际交流，这个过程就像一个结果难以预料的黑箱。我感谢所有给我攻读博士学位机会的人——Lance Fortnow（佐治亚理工学院）、Laszlo Babai（芝加哥大学）、David Zuckerman（德克萨斯大学奥斯汀分校）、Anup Rao（华盛顿大学，2011），以及 Thomas Vidick 和 Leonard Schulman（加州理工学院）。

我想感谢 Thomas Vidick，他几乎是一位完美的导师——精力充沛、勤奋刻苦、真正的量子专家，头脑极其敏锐，能剖析并洞察各种想法，聪明绝顶，而且拥有超人的能力，能同时处理多项任务。他为加州理工学院的理论组注入了活力，并总是尽力帮助学生。他确保我始终能获得所需的各种资源，让任何事情，无论多么微不足道，都不会妨碍我的研究。我在加州理工学院的头几年里，我们围绕对易哈密顿量进行了许多有趣的讨论；即使后来我开始研究电路复杂性中的其他问题，他也总会定期过问，确认一切顺利。他也从未阻止我钻研极其困难的问题，尽管这样做几乎注定收获甚微。

我想感谢 Chris Umans。在不同的时期，我们曾愉快地讨论复杂性理论中的各种问题。从某种意义上说，在最初几周里，我思考问题的方式与他的最为接近。他思路极其清晰、精准，总是立即深入问题的细枝末节，并以研究生般的热忱，数年如一日地攻克难题，这始终激励着我。

感谢 Shuki Bruck 担任我的学位论文委员会成员，慷慨地拨冗相助，并总是在报告期间提出极有见地的问题。

最后，由于过去几年一直与 Leonard Schulman 共事，我可以写得更详细，因此我想由衷地感谢他，感谢他的一切。他热衷于有趣的问题，不论问题源自何处；他敏锐地洞察问题的本质，也凭借在数学各个领域的广博知识解决问题。事实上，我第一次访问加州理工学院时，

我们只聊了我高中时做的一个项目，研究牛顿摆中权重各异的球，完全没谈主流理论。他待人极其和善（虽然从他简洁的邮件中看不出来），品味高雅而精致，还有一种独特、略显冷峻的幽默感，为每次交谈增添了魅力；他谈论任何话题时流露出的哲思，也展现了他的深邃智慧。他对几乎所有类型的随机对象都有许多有力而新颖的直觉，也有一种难得的能力，能够质疑显而易见的事物，并且（几乎总是）极其清晰地把握复杂问题。有一次，他提出了一个关于有向图切格型不等式的问题，我觉得非常有趣；过去几年与他一起研究这个问题，最终促成了这篇学位论文。他总是让我感到自在，也让每次讨论都妙趣横生、令人愉快，以至于我总盼着与他见面；他还总能确保我们每次都能对这个问题有新的认识。我们的会面几乎总是灵活安排，而且可以持续两三个、四个甚至更多小时，这一点对我帮助极大。我将永远感激他，让我与他一起做研究的经历如此美好。

最重要的是，我要感谢 Thomas Vidick 和 Leonard Schulman，他们给予我自由，让我发展自己对理论的感受、直觉和认知，并让我能够根据自己的兴趣和性情选择问题。这份支持弥足珍贵，我将永远铭记于心。

来到加州理工学院之前，我大多是自学复杂性理论，因此在知识储备和对不同领域概念的理解上都有许多欠缺。在加州理工学院的头几年，对我而言就像是延长版的本科阶段。我要感谢许多人开设了极为出色的课程——Joel Tropp（线性代数）、Chris Umans（复杂性理论，或许是全世界最好的复杂性课程）、Alexei Kitaev（量子计算）、Leonard Schulman（随机性与算法）、Thomas Vidick（哈密顿量复杂性）、Gil Cohen（随机性提取器）、Adam Sheffer（加法组合学；多项式方法与关联），以及 Erik Winfree（生物分子计算）等。我还要感谢 Zeyu Guo，让我对加法组合学产生了兴趣。

最初几年，我从事电路复杂性研究（尝试证明计算内积函数的常数深度 $AC^0 \cdot [\oplus]$ 电路规模的指数下界）。尽管我没能取得多少有意义的成果，但这段经历让我对电路的行为以及各种工具的局限性有了深刻的认识和理解，我希望这些收获能对今后研究这个问题有所帮助。

这段时间里，我想感谢 Johan Hastad、Ben Rossman 和 Srikanth Srinivasan 与我进行了启发性和／或富有成果的交流。我还要感谢 Jaikumar Radhakrishnan 和 Prahlad Harsha 在两个暑假期间两次邀请我到塔塔基础研究院访问，感谢 Meena Mahajan 和 Samir Datta 在我本科期间的两个暑假接待我，也感谢 Samir Datta 为我提出了第一个理论问题。

加州理工学院是一个美丽的地方，我曾在龟池和其他各处度过许多惬意的时光，沉思各种各样的事情。我感谢它给予我这样的机会。我也感谢所有工作人员——特别是 Maria 和 Linda——让各种事务安排都不必让我操心；感谢 Linda 与我进行了许多温暖而愉快的交谈；也感谢加州理工学院国际学生项目，让所有政府相关事宜都变得简单。我还要感谢整个加州理工学院西海岸摇摆舞团体，教会了像我这样四肢僵硬的人如何舞动。

我想感谢许多新老朋友。虽然无法一一列举，但我会提及几位亲近的朋友和／或容易想起的人：研究生时期的朋友——Andrea（最亲密的研究生朋友）、Jing、Zofii、Tal、Palma、Swagato、Florian、Fabien、Sudhi、Rohit、Ben Lee、Anand、Saeed 和 Pranav；本科时期的朋友——Anant（总是带我去冒险）、Praneeth，以及所有至今仍保持联系的人；初高中时期的朋友——Miren（如今就像我的兄弟）、Dhruman、Saagar 和 Preet；还有在我博士学位即将结束时认识的 Corinne，感谢她成为如此美好的人。

最后，我想感谢家人给予我无微不至的支持。我想感谢我的父亲。他天赋卓绝、纯粹无瑕、浑然一体、真实不虚，且近乎超凡的才智，总让我觉得自己不过是受过训练的头脑，只是在符号（或可用符号表达的思想）间移动，以生成所需的字符串。因此，在我于加州理工学院求学期间，他始终是我的动力，尤其是在问题（电路下界）迟迟得不到解决的时候。他对我所做的一切给予的支持，其数量之多、意义之深，难以言表；若没有他坚如磐石的陪伴与支持，我便不可能开始攻读博士学位。遗憾的是，他于2021年因新冠肺炎离世，就在我答辩前几个月。我们之间有着深厚而难以言喻的联结。他每天都会给我打电话、与我交谈（尽管他有千头万绪要忙），在不知不觉、无意之间，以间接的方式给予我继续前行的动力。我想感谢母亲，感谢她深切的关怀、无条件且无私奉献的爱、超凡脱俗而令人安心的陪伴，也感谢她一直以来的微笑与支持。她以富有同理心的方式理解这个世界，与各种现实接触时总是细腻而真诚；

因此，她深深影响了我自身的审美观。我想感谢姐姐 Rutvi，她开朗活泼、纯真无邪；我们俩共有的那股疯狂劲儿，也帮我保留了自己的一点天真。我感谢表亲 Rishabh，感谢我们之间的情谊，也感谢他一直以来事事陪伴着我；还要感谢 Dhwani、Chintan、Dhruvi，以及整个大家族。

最后，我想感谢所有可能被我遗漏的人，以及我目前无法回忆起的所有直接或间接的关联。

谨献给我的父母,

Chetan 和 Bina Mehta

（以及他们可能将其献给的任何人,

等等)

目录

摘要.iii 致谢.v 目录.x 第一章：引言.1 1.1 随机性及其演变.1 1.2 马尔可夫链与佩龙-弗罗贝尼乌斯定理.2 1.3 切格-布泽不等式.3 1.4 佩龙-弗罗贝尼乌斯定理的定量推广.5 1.5 非可逆矩阵的构造.8 1.6 与其他组合性质的关系.10 第二章：预备知识——谱隙、边扩张与佩龙-弗罗贝尼乌斯定理.16 第三章：佩龙-弗罗贝尼乌斯定理的推广与切格-布泽不等式.21 3.1 Buser 不等式—— ϕ (R)的下界.24 3.2 Cheeger不等式—— ϕ(R)的上界.25 3.3 Cheeger-Buser 不等式的紧性.29 3.4 Fiedler的证明——根据 Δ(R)给出 ϕ(R)的上界 . . . . . 29 3.5 奇异值的作用.31 3.6 微扰理论的局限.32 3.7 根据 Δ(R)给出 ϕ(R)的下界.35 3.8 定理 3.2 向所有非负矩阵的推广.46

第四章：非可逆链的构造.48
4.1 非扩张 .49 4.2 更进一步1/2 .50 4.3 仿射线性构造 .51 4.4 de Bruijn 构造 .53 4.5 搜索空间的约束 .59 4.6 满足约束的特定矩阵的特殊情形 . . . . . 63 4.7 一般情形及相关推理 .65 4.8 Rootn 矩阵——固定谱隙并最小化边扩张 .69 4.9 Rootn 矩阵构造中的观察 . . . . . 74 4.10 Chet 矩阵——固定边扩张并最大化谱隙 .79 4.11 Chet 矩阵的观察与性质 .82 4.12 证明 Chet 矩阵非负性的猜想 . . . . . 90 第五章：联系与扩展 .94 5.1 混合时间——定义与预备引理 .95

5.2 混合时间与奇异值.98 5.3 混合时间与边扩张.99 5.4 混合时间与可逆矩阵的谱隙.101 5.5 混合时间与非可逆矩阵的谱隙.103 5.6 非负矩阵及其加性对称化的混合时间 105 5.7 连续算子的混合时间.107 5.8 与典范路径法的比较.109 5.9 舒尔补.110 5.10 舒尔补的性质.112 5.11 容量与狄利克雷引理.114 5.12 归一化容量与有效电导.128 5.13 与对数 Sobolev 常数的关系.138 5.14 一种不同的扩张概念.138 5.15 张量及更广领域.142 结果与问题总结.148 参考文献.151 附录 A：Chet 矩阵——精确计算与数值计算 . . . . . 157 A.1 精确计算得到的 $n = 16$ 的 Chet 矩阵.157 A.2 计算输入整数 n 对应的 Chet 矩阵的伪代码. . . . . 161 A.3 数值计算得到的 $n = 500$ 的 Chet 矩阵 162

第 *1* 章

引言

> 太阳明天不会升起，这一命题并非难以理解，也不比断言太阳明天会升起包含更多矛盾。
>
> ~ 大卫·休谟，《人类理解研究》

1.1 随机性及其演变

随机性是人类思想史上的一个基本观念，而其可测量性所赋予的功能性，更使它不可或缺。随机性的可测量性不同于其含义——后者指对随机性的不同诠释，其中较为重要的两种是贝叶斯视角——纯粹用于理性决策——以及频率学派视角——这是对事件序列本质更强烈、依据也更薄弱的断言。除此之外，还有多种其他诠释，但没有一种足够令人满意或全面，无法为随机性提供严谨且合理的普遍定义。从足够批判的视角来看，与随机性相比，我们发现某种形式的条件非决定性更为根本和普遍，更适合用来理解观测到的事件。然而，尽管这一问题深刻，而且可以说比我们下文将要论述的任何内容都更根本、更重要，本学位论文关注的并非随机性的含义（甚至也不是其可测量性），而主要是其演变。

接下来，我们假设存在一组可能的 n 状态，其数量 n 可以是有限的，也可以是无限的（$\aleph_0, \aleph_1$, 或更多），并且这组状态上存在相应的随机性 p 。请注意，不同状态通常可以对应于任何系统或宇宙的状态。设随机性 p 在不同的时间点演变为 p'

（我们隐含地认为存在不同的事件，这些事件为我们带来了隐含的时间或步数概念）。为了用更数学化的方式理解随机性的演变，我们假设 $p' = Ap$，而将把 p 映射为 p' 的算子 A 作为主要研究对象。

1.2 马尔可夫链与佩龙-弗罗贝尼乌斯定理

我们可以考察 A （以及 p）的许多性质，但为了以具有数学意义的方式理解算子 A，我们对 A 和 p 作出各种简化假设。我们的第一个假设是，在本学位论文中，我们只考虑有限状态空间，并始终用 n 指代它。我们假设 A 独立于 p，且有限维。因此， A 可以表示为有限个常数。在整篇学位论文中（张量部分除外），我们进一步假设 A 是线性的，因此可以表示为一个 $n \times n$ 矩阵。至此，研究范围缩小到一般线性算子；最后，我们假设 A 是逐项非负的，以便对 A作出比一般线性算子更强的论述。

这样的 A 称为马尔可夫链，最早由安德烈·马尔可夫于1906年进行研究。其研究动机源于他与帕维尔·涅克拉索夫的一场争论；涅克拉索夫曾声称，独立性是大数弱定律成立的必要条件 [Mar06；Sen96]。马尔可夫链历史悠久，在过去100多年里得到了深入研究，并已应用于数学的诸多领域。

可以说，关于马尔可夫链最基本的定理是佩龙-弗罗贝尼乌斯定理。为简便起见，假设 A 逐项为正。那么佩龙定理[Per07] 指出，存在一个正数 r，它是 A的特征值； A 的所有其他特征值 $\lambda \neq r$ 均满足 $|\lambda| < r$，且与 r 对应的左右特征向量均逐项为正。弗罗贝尼乌斯 [Fro12], 将这一结果推广到了所有不可约的 A。若带有非负边权$A_{i,j}$ 的底层图是强连通的，也就是说，任意一对顶点之间都存在一条权重非零的有向路径，则 A 是不可约的。在不可约的情形下，唯一的具体区别是：对于 A的所有其他特征值 $\lambda \neq r$，有 $|\lambda| \leq r$ 且 $\mathrm{Re}\lambda < r$。由此得到的定理称为佩龙-弗罗贝尼乌斯定理。

佩龙-弗罗贝尼乌斯定理是数学诸多领域的基础，并广泛应用于大量实际问题（例如参见 [Mac00], [BP94]），例如马尔可夫链 [LPW09]（传统上，马尔可夫链一直被表示为

非负行随机矩阵，以保持概率单纯形）；动力系统理论 [KH97]；经济学中的大量结果，例如置盐定理 [Bow81], 霍金斯-西蒙条件 [HS49], 列昂惕夫投入/产出经济模型 [Leo86], 以及竞争市场的瓦尔拉斯稳定性[TA85]；人口学中的莱斯利人口年龄分布模型 [Les45; Les48]；社交网络中的德格鲁特学习过程 [DeG74; Fre56; Har59]；PageRank 和互联网搜索引擎 [Pag+99; LM11]；低维拓扑中的瑟斯顿曲面微分同胚分类 [Thu88]；流行病学中的克马克–麦肯德里克阈值 [KM91; KM32; KM33]；统计力学（具体而言，配分函数） [Tol79]；矩阵迭代分析中的斯坦–罗森伯格定理以及求解线性方程的塞德尔迭代法与雅可比迭代法之比较[Var62]；更多内容请参见综合性资料 [BP94] 。

然而，早在此时，就出现了一个有趣的问题，这将成为本学位论文第一部分的主要研究重点。佩龙-弗罗贝尼乌斯定理给出了关于矩阵的不可约性和特征值的定性结论，即它表明，如果 A 是不可约的，那么 A 的每个特征值 λ（r 除外）都满足 $\mathrm{Re}\lambda < r$。我们的主要问题是：这一结果的定量版本是什么？假设 λ是除 r 之外实部最大的某个特征值。那么具体而言，我们可以提出这样的问题：

问题 1。 A 的某种不可约性度量与间隔 r –Reλ之间是否存在定量关系？

这将是我们首先重点考虑的问题。事实上，在 A 为对称矩阵的情形下，这一问题已经得到了完整解答。对于可逆的 A，该结果仍然成立，其中 A 是可逆的，当且仅当 $D_u A D_v = D_v A^T D_u$，其中 u和 v 分别是 R 对应于特征值 r的左、右特征向量；但从语法（以及直观）上看，可逆性与对称性实际上完全相同（见引理 3.3）。为了理解这一问题的对称版本，我们还需要一些定义。

1.3 切格-布泽不等式

我们首先定义对称矩阵 A 的谱隙和边扩张。为便于说明，假设 A 是不可约的，且 A 的最大特征值为 1，相应的正特征向量为 w。第一个概念是谱隙，对于对称矩阵而言，它同样易于定义。由于 A 是对称的，设 A 的实特征值为 $1 \geq \lambda_2(A) \geq \ldots\ldots \geq \lambda_n(A)$。

将 A 的谱隙定义为

$$\Delta(A) = 1 - \lambda_2(A).$$

第二个概念是 A 的边扩张（或扩张），它有助于量化 A 的不可约性，定义如下：

$$\phi(A) = \min_{S\subset[n]:\sum_{i\in S} w_i^2 \le \frac{1}{2}\sum_i w_i^2} \frac{\sum_{i\in S, j\in\overline{S}} A_{i,j} w_i w_j}{\sum_{i\in S,j} A_{i,j} w_i w_j}.$$

从符号上看，这个定义可能显得繁琐，但只要看一下对称双随机矩阵 A 的等价定义，直观上就很容易理解；在这种情况下， w 就是全一向量。此时，我们有

$$\phi(A) = \min_{S:|S|\le n/2} \frac{\sum_{i\in S, j\in\overline{S}} A_{i,j}}{|S|}.$$

因此，给定一个集合 S，该集合的边扩张恰好等于离开集合 S 的边的总平均权重，也就是说，这些边从 S 通向 S。对于一般的 A，同样的直觉也成立，只是边会按照特征向量 w 重新加权（其原因稍后会变得清楚）。边扩张这一概念是对 A 的不可约性进行量化的一种简单且足够一般的方式。注意，如果 $\phi(A) = 0$，那么存在某个集合 S，使得离开该集合的边的总质量为 0；这意味着无法从该集合到达 S 中的顶点，因此矩阵是不连通的，其不可约性为 0，或者说它是可约的。同样，如果 $\phi(A) \approx 0$，这意味着 A 中有两个几乎不连通的集合，因此 A 连通性较差，或者说可约性较高。同样，如果 $\phi(A) \ge c$（其中 c 是一个与 A 的大小无关的任意常数），则意味着对于每个集合 S，都有一定比例的边质量离开 S。这意味着，从任意集合 S 中均匀随机选取一个顶点并从此开始随机游走，很可能会离开 S，不会被限制在 S 内；这说明 A 连通性很好，或者说不可约性很高。因此，边扩张 ϕ 清晰地体现了不可约性的概念。此外，我们指出，给定 A，判断 A 的边扩张是否大于给定的输入数 [GJS74] 是 NP困难的。

有了这两个定义，我们就可以通过将 $\phi(A)$ 与 $\Delta(A)$ 联系起来，回答关于对称矩阵的问题 1。Cheeger [Che70]，得到了用 $\Delta(A)$ 表示 $\phi(A)$ 的上界 [，尽管其研究对象是黎曼流形；随后， [AM85; SJ89; Dod84; Nil91] 将其推广到对称矩阵。Buser [Bus82] 为黎曼流形得到了用 $\Delta(A)$ 表示 $\phi(A)$ 的下界，

而将其推广到对称矩阵，则是对称矩阵特征值变分刻画的直接推论。合起来，以下界称为切格-布泽不等式：

$$\frac{1}{2}\Delta(A) \leq \phi(A) \leq \sqrt{2 \cdot \Delta(A)}. \tag{1.1}$$

这些界是紧的：下界对超立方体恰好取等，而无向环的上界则在常数因子范围内取紧。请注意，这有助于我们从谱隙的角度理解 A 的连通性或不可约性，正如我们所希望的那样，并为对称（或可逆）矩阵的佩龙-弗罗贝尼乌斯定理提供了一个定量版本。切格-布泽不等式已被用于许多不同的结果，例如聚类和扩展图构造 [TM06；ARV09；HLW06；ST96], 而高阶切格不等式也是近来才得到证明的 [LOT14]。

任意非负矩阵（不一定是对称或可逆矩阵）的一般问题此前一直悬而未决，而这确实将引出我们的主要结果。

1.4 佩龙-弗罗贝尼乌斯定理的定量推广

我们首先重新定义一般不可约非负矩阵 R 的边扩张和谱隙，该矩阵不一定是对称的。对于最大特征值对应的左右特征向量不同的矩阵，我们记作 R；否则记作 A。同样，为了简化表述，假设 R 的最大特征值为 1，并将 R 的特征值（方程 $\det(\lambda I - R) = 0$ 的根）排列如下： $1 \geq \mathrm{Re}\lambda_2(R) \geq \ldots\ldots \geq \mathrm{Re}\lambda_n(R)$。于是定义 R 的谱隙为

$$\Delta(R) = 1 - \mathrm{Re}\lambda_2(R).$$

与基无关的特征值不同，边扩张取决于矩阵的具体矩阵元素（只有这样才能从概率的角度理解其意义）。因此，我们假设 R 是非负且不可约的（出于纯粹的技术原因），并且根据佩龙-弗罗贝尼乌斯定理， R 的特征值 1 对应的正的左、右特征向量分别为 u 和 v。接着，将 R 的边扩张定义为

$$\phi(R) = \min_{S \subset [n]: \sum_{i \in S} u_i v_i \leq \frac{1}{2}\sum_i u_i v_i} \frac{\sum_{i \in S, j \in \overline{S}} R_{i,j} u_i v_j}{\sum_{i \in S, j \in [n]} R_{i,j} u_i v_j}.$$

我们将这些定义扩展到可约 R：通过在 R周围的不可约矩阵球内取下极限，并且 $\phi(R)$ 的正式定义见2.3 和

2.4，而 $\Delta(R)$ 的定义见2.2 和 2.6。有了这些定义，我们便可在此陈述本文的主定理。

定理 1.2。 设 R 是一个 $n \times n$ 非负矩阵，其边扩张 $\phi(R)$定义为 *2.4* 和 *2.6*，谱隙 $\Delta(R)$ 定义为 *2.2* 和 *2.6*。那么

$$\frac{1}{15} \cdot \frac{\Delta(R)}{n} \leq \phi(R) \leq \sqrt{2 \cdot \Delta(R)}.$$

我们指出，这恰好有助于回答问题 1，并推广了佩龙-弗罗贝尼乌斯定理。它比切格-布泽不等式更一般，因为它对所有非负矩阵都成立；但它并非对该不等式的推广，因为对于对称矩阵情形，下界会变得更弱。

我们对定理 1.2作几点评论。首先，上界与对称矩阵情形中的上界完全相同。Fiedler 已证明双随机矩阵 A 满足该上界，将其推广到所有 R是直接的。事实上，给定切格不等式，很容易利用该不等式证明定理 $\phi(R)$ 中的 $\phi(R)$ 上界。定理 1.2。 然而，与对称情形不同，要得到 $\phi(R)$ 的下界，或得到与 Buser 不等式等价的结果，要困难得多。困难的主要来源在于，非对称矩阵不存在 Courant-Fisher 特征值变分刻画。事实上，对称矩阵的这一刻画是谱理论中大多数关于对称矩阵的定理的核心，而对称矩阵通过这一刻画已得到全面理解。然而，处理非对称情形需要采用不同的思路。

定理 1.2 中关于 $\phi(R)$ 的下界是通过一系列引理证明的，每条引理都可以独立地用于不同的情境。主要思路是证明 $\phi(R^k) \leq k \cdot \phi(R)$（对所有整数 k 均成立），并利用 R 的舒尔分解将 $\phi(R^k)$ 与 R 的谱隙联系起来。定理证明之前，第 3.7 节对此作了详细阐述。定理 1.2 与切格-布泽不等式的主要区别在于因子 n，它使得该下界极弱。不过，该定理中关于 $\phi(R)$ 的下界之所以有力，在于其中包含了那些并未出现的因子。

首先，显而易见的一点是，对 n 的依赖是线性的，而非指数级的。这一点至关重要，因为我们最初尝试证明这一结果时，采用了微扰理论中的工具——因为 $\phi(R)$ 可以看作是对一个由非连通分量构成的矩阵的微扰。我们可以根据矩阵的变化来理解特征多项式系数的变化，并

进一步将其与谱隙的变化联系起来。然而，这会引入对 n 呈指数依赖的因子。另一种思路是使用类似鲍尔-菲克[BF60]（见引理 3.13）的结果，但其适用范围非常有限，因为它仅适用于可对角化R矩阵，且取决于特征向量条件数。特征向量条件数为 $\mathrm{K}(R) = \|V\|_2\|V^{-1}\|_2$，其中 V 是一个可逆矩阵，可将 R 变换为其若尔当分解 $R = VJV^{-1}$。由于即使对于小矩阵，特征向量条件数也可能大得惊人，因此对 $\mathrm{K}(R)$ 的依赖使鲍尔-菲克结果远弱于我们所寻求的结果。Saad [Saa11], 将鲍尔-菲克结果推广到了不可对角化矩阵，但该推广对 n 有指数级依赖，并且依赖于 $\mathrm{K}(R)$。不过请注意，我们的定理 1.2 中的结果不依赖于 $\mathrm{K}(R)$。我们还利用舒尔分解重新证明了鲍尔-菲克结果，消除了对 $\mathrm{K}(R)$ 的依赖，但转而引入了对 $\exp(n)$ 和 $\sigma(R) = \|R\|_2$ 的依赖，其中 $\sigma(R)$ 是 R 的最大奇异值（见引理 3.15）。我们的定理 1.2 的证明中还可能出现另外两个因子；事实上，它们确实出现在我们证明的初始版本中，但在后续迭代中被逐一消除——R 的最大奇异值 $\sigma(R)$，以及特征值条件数$\kappa(R) = \min_i u_i \cdot v_i$，其中 u 和 v 是对应于 R 的最大特征值 1 的正特征向量，并经过归一化，使得 $\sum_i u_i v_i = 1$。通过一个巧妙的技巧（引理 3.3），可以消除对 σ 的依赖，但消除对 $\kappa(R)$ 的依赖则并非易事。事实上，在我们结果的已发表版本 [MS19], 中，我们给出的界较弱，为

$$\frac{1}{15} \cdot \frac{\Delta(R)}{n + \ln\left(\frac{1}{\kappa(R)}\right)} \leq \phi(R).$$

在尝试构造能够说明 $\kappa(R)$ 对于下界而言不可或缺的矩阵时，我们发现，对于矩阵所有可能的“极端情形”，$\kappa(R)$ 确实并非必要。这促使我们重新审视证明中的所有引理；事实上，我们可以改写这些引理，完全消除对 $\kappa(R)$ 的依赖，并得出定理 1.2。

我们还想讨论一下定理 1.2 中 $\phi(R)$ 的下界在哪些情形下有用。由于存在因子 n，当 $\phi(R) \ll 1/n$ 时，这个界最有意义。因此，这个界或许可以作为一项扰动结果来使用：微小扰动——例如 $O(1/n^2)$、$O(1/n \log n)$ 或 $O(1/n\log^2(n))$——会显著影响谱隙；而受扰矩阵中的特征值与正特征值之间的间隔，可以用以 ϕ 表示的相应扰动来界定（见第 3.6 节）。

Th对我们而言，不依赖于 $\exp(n)$（因为我们不使用微扰理论 ），

K(R)（因为我们使用舒尔分解而非若尔当分解），$\sigma(R)$（因为我们将 R 变换为一个 A，使其保留特征值和边扩张，且最大奇异值与特征值相等），以及 $\kappa(R)$（因为我们确保所有中间引理都是紧的，而且关键在于，特征值 1 对应的右特征向量 v 属于 $I - R$的核），正是 $\phi(R)$下界的重要性所在，因为它有助于回答第一个问题，并精确量化佩龙-弗罗贝尼乌斯定理。这意味着 $\Delta(R)$与 $\phi(R)$成比例，比例因子为 n。然而，由于在对称情形下并不依赖于 n （不等式 1.1），这就引出了第二个主要问题——定理 1.2 中对n 的依赖是否必要，还是说它更多是我们证明方法所造成的假象？更具体地说，

问题 2。 是否存在矩阵 R，使得 $\phi(R) \approx \overset{\Delta}{}(R)$? $\frac{}{n}$

1.5 非可逆矩阵的构造

探讨问题 2 会引领我们进入一些新颖且鲜少有人涉足的领域。为了以最严格的方式回答这个问题，我们现在将寻找双随机矩阵，使其达到问题 2 所要求的界。主要原因在于，这些矩阵以均匀向量（其方差为零）作为特征值 1 对应的左、右特征向量；乍看之下，至少对于双随机矩阵而言，定理 ϕ 中的下界 1.2似乎可能要弱得多。为了弄清这一点，我们将先仔细研究特定的不可逆或非对称矩阵的边扩张。

对称矩阵的边扩张已得到充分研究，具有常数边扩张的矩阵称为（组合）扩展图。这些矩阵具有显著的性质，是组合学中的基本构件；而具有常数边扩张且非零元素（边）较少的矩阵，其构造或证明颇为复杂，并且历史悠久，因为它们具有“神奇”的性质，有助于构造看似不可能的对象或算法（见 [HLW06]）。然而，问题 2 与对称情形无关，因为边扩张与谱隙密切相关，而我们寻求的是这两个量之间相差 n 倍的谱隙。

Maria Klawe 的研究已经迈出了这一方向的第一步[Kla81；Kla84] 他证明，某些矩阵的仿射线性构造具有逆多对数(n) 边扩张（见构造 4.5）。Umesh Vazirani [Vaz17], 指出，可以对 Klawe 构造中的边进行定向

从而得到谱隙为常数、但扩张为 1/polylog(n) 的双随机矩阵！这相当令人惊讶，因为它表明存在这样的双随机矩阵：

$$\phi(A_{KV}) \approx \frac{\Delta(A_{KV})}{(\log n)^c}$$

对于某个常数 c。事实上，还有许多其他仿射线性构造，都能达到类似的下界。值得特别一提的是德布鲁因矩阵的构造（第 4.4 节将对此进行详细描述和分析）。对于这些矩阵，结果表明

$$\phi(A_{dB}) \approx \frac{\Delta(A_{KV})}{\log n}.$$

这些矩阵十分优美，其性质列于引理 4.11。这些构造表明，无法避免对 n 的某个函数存在依赖。不过，问题 2 可以改述如下。

问题 3。 是否存在双随机矩阵 A 使得

$$\frac{\phi(A)}{\Delta(A)} \in o\left(\frac{1}{\log n}\right)$$

还是说，对于所有双随机矩阵，都存在某个常数 c，使得

$$c \cdot \frac{\Delta(A)}{\log(n)} \leq \phi(A)?$$

这是我们要考虑的主要问题，因为如果边扩张至少为谱隙乘以 $\log(n)$，就意味着定理 1.2 中关于 $\phi(R)$ 的下界与真实值相差呈指数级（至少对于双随机矩阵而言），因此需要采用新方法来获得更强的下界。此外，这也意味着谱隙能够很好地估算边扩张，而边扩张本身很难计算，因为在大多数应用中，$\log(n)$ 因子基本可以忽略。不过，我们将展示以下令人瞩目的构造。

定理 1.3. 存在一族双随机矩阵 A_n，称为 *Rootn* 矩阵，使得对于每个 n，

$$\phi(A_n) \leq \frac{\Delta(A_n)}{\sqrt{n}}.$$

本节介绍 Rootn 矩阵的构造，见第 4.8 节；该构造在以下各节中得出、介绍并进行了详细讨论：4.5、4.6、4.7、4.8。这一构造比已知构造的效率呈指数级提升，并表明 n 中的因子

定理 1.2 确实更接近真相。Rootn 矩阵的构造方法是：仔细调整舒尔分解中三角矩阵的特征值，以产生一个常数谱隙；再选择一个酉矩阵，使所得矩阵为双随机矩阵，且具有最小边扩张。

因此，最后剩下的问题是 $\sqrt{n}$ 与 n之间的差距。为了解决这个问题，我们采取了不同于构造 Rootn 矩阵时所用的方法。通过观察 Rootn 矩阵，我们发现了矩阵可能具有的一种结构，它有助于固定边扩张；随后，我们非常谨慎地设定矩阵元素，以最大化谱隙，详见第 4.9节。这引出了本学位论文中最精彩的成果——Chet 矩阵，它有助于证明以下结论。

定理 1.4。 存在一族矩阵 C_n ，称为 *Chet* 矩阵，使得$\sum_i C_n(i,j) = 1$且 $\sum_j C_n(i,j) = 1$对所有 i, j均成立；若 $\phi(C_n)$像双随机矩阵那样定义，

$$\phi(C_n) = \min_{S,|S|\le n/2} \frac{\sum_{i\in S, j\in \overline{S}} C_n(i,j)}{|S|},$$

则对所有 n,

$$\phi(C_n) \le 2 \cdot \frac{\Delta(C_n)}{n}.$$

该构造见第 4.10节。然而，我们无法证明对所有 n，Chet 矩阵都是非负的。对于 n，我们进行的数值测试表明，当其不超过 500 时似乎确实如此（见附录 A.3），但我们无法证明这一结论对所有 n都成立。尽管如此，这些矩阵仍具有第 4.11节列出的许多显著性质，而证明它们的非负性是本学位论文的主要开放问题之一。为证明它们的非负性，我们在第4.12节提出了一系列迹猜想；如果这些猜想成立，便几乎可以推出 Chet 矩阵的非负性。我们认为这些猜想本身也很有趣。

目前为止的结果完整回答了我们最初提出的第一个问题，因为我们得到了紧界和（几乎）匹配的构造。现在我们将转向其他组合性质，并探讨它们与边扩张和谱隙之间的关系。

1.6 与其他组合性质的关系

混合时间

马尔可夫链最受广泛研究、且对算法不可或缺的性质是混合时间。为简化说明，考虑一个不可约双随机

矩阵 A，它是$\frac{1}{2}$-惰性的；也就是说，对于每个 i，都有 $A_{i,i} \geq \frac{1}{2}$。这一条件纯粹是出于技术原因（遍历性）而要求的，任何其他常数也足够。该条件保证，除特征值 1 外，A 的所有其他特征值 λ 都满足 $|\lambda| < 1$。因此，若将 $A = VJV^{-1}$ 写成若尔当分解，显然，当 t 足够大时，$A^t \approx \frac{1}{n}\mathbf{1} \cdot \mathbf{1}^T$（其中 $\mathbf{1}$ 是全一向量），因为 A 中所有其他特征值都会趋近于零，只剩下对应于特征值 1 的空间。注意，这也意味着，若我们从某个概率分布 p 开始，当 t 足够大时，$A^t p \approx_\epsilon \frac{1}{n}\mathbf{1}$（例如，按 l_1 范数衡量）。对于任意初始 p 都成立的最小 t，称为 A 或 τ_ϵ 的混合时间。类似地，我们也可以正式定义任意不可约且$\frac{1}{2}$-惰性的非负矩阵 R 的混合时间，该矩阵的最大特征值为 1，对应的左右特征向量分别为 u 和 v，并满足 $\langle u, v \rangle = 1$。这将是满足 $R^t \approx_\epsilon v \cdot u^T$ 的最小 t。

关于混合时间、对其进行界定的不同方法，以及可用于特定类型链的不同工具，已有大量文献。我们的目标主要是获得混合时间的一般界，并将其与其他组合和代数量联系起来。

对称（或可逆）A 的混合时间已得到全面研究，且我们已充分了解它（参见 [AF02； LPW09； MT06]）。根据 A的谱分解，可逆链的混合时间 τ_ϵ 满足以下界：

$$c_1 \cdot \frac{1}{\Delta(A)} \leq \tau_\epsilon(A) \leq c_2 \cdot \frac{\ln(\frac{n}{\epsilon \cdot \kappa(A)})}{\Delta(A)}$$

其中常数为 c_1 和 c_2。因此，混合时间约为谱隙的倒数，差一个 $\log(n)$ 因子和特征值条件数$\kappa(A)$。因子 $\kappa(A)$ 实际上会出现在所有与混合时间相关的界中，而且很容易看出，即使取一个非常接近可约的 3×3 矩阵，也无法消去该因子。

对于不一定对称的一般链，针对特定链也已有许多不同的已知结果，其中最一般的是 Mihail 的结果[Mih89]，，它直接将混合时间与边扩张联系起来（Mihail 对行随机 R证明了这一结果，但只需进行语法上的修改，就能直接推广到所有 R）。

$$\tau_\epsilon(R) \leq c \cdot \frac{\ln\left(\frac{n}{\epsilon \cdot \kappa(R)}\right)}{\phi^2(R)}.$$

这表明，混合时间与 R 的边扩张成反比。我们的第一个结果本质上是对上述结果的一行证明，可由

我们关于矩阵混合时间的主要引理之一推导得出 A （由R导出），该矩阵在特征值 1 对应的左右特征向量相同，且其 w 最大特征值奇异值为 A。我们还给出了 τ_ϵ（R）关于 ϕ（R）的一个下界，该结果也广为人知（见 [LPW09]），不过我们将这个界写成适用于一般 ϵ的形式。

我们还将不可逆情形下的混合时间与谱隙联系起来，并得到以下定理。

定理 1.5. 设 $\tau_\epsilon(R)$ 为不可约1 2-惰性非负矩阵 R的混合时间。那么

$$\tau_\epsilon(R) \leq c \cdot \frac{n + \ln\left(\frac{1}{\epsilon \cdot \kappa(R)}\right)}{\Delta(R)},$$

我们在第 *4.8* 节和*4.10* 节中对 *Rootn* 和 *Chet* 矩阵的构造 n 再次表明这一因子是必要的。

结果1.5中有趣之处在于，分子中的 n 是相加关系，而非相乘关系。接下来，我们研究 A（不可约、非负、1 2-惰性，且具有对应于特征值 1 的左右特征向量 w ）的混合时间与 $\tilde{A} = \frac{1}{2}(A + A^T)$ 的混合时间之间的关系，并给出它们之间双向且紧的界。我们还将 A 的混合时间与 $\exp(t \cdot (I - A))$ 的混合时间联系起来。

我们主要给出许多关于混合时间的已知界的简明／初等证明，并在第 5.1 节介绍一些新结果。我们的主要贡献是展示如何仅用少数几个关键工具和引理就能达到所有（最优）界，这也帮助我们以此前未曾见过的形式获得了许多界。

容量与归一化容量

接下来要探讨的量，在不同领域和不同历史时期，对于对称矩阵已有大量研究，名称也各不相同；我们称之为容量。这一术语及相关问题源于对调和函数和狄利克雷问题的研究。为了再次给出一个简化的定义，假设我们有一个不可约非负矩阵 A ，其最大特征值为 1，对应的左右特征向量为 w。设 A 的拉普拉斯矩阵为 $L = I - A$，设 U 是顶点的某个子集，且 $a \in \mathbb{R}^{|U|}$ 是 U 上的某个实向量。我们希望找到一个向量 $q \in \mathbb{R}^n$，使得对于 $i \in U$，有 $q_i = a_i$；也就是说， q 在 U上与 a 相同，并且对于 $i \in U$，有 $(Lq)_i = 0$。由于 A 不可约，不难看出，满足这些方程的向量 q 唯一。给定这样的 q，容量定义如下。

$$\mathrm{cap}_A(U, \overline{a}) = \langle q, Lq \rangle.$$

向量 a 满足 $a_i = a_i/w_i$。这给出了矩阵 A 下向量 a 在集合 U 上的容量（由于一个技术条件，这并非 a 的容量）。我们的定义完全一般；对称矩阵的容量特例已有研究（参见 [DS84]），其方法是将图解释为电网络，并将边权重 $A_{i,j} = A_{j,i}$ 视为边 $\{i, j\}$ 的导通率。此外，如果 $a \in \{0, 1\}^{|U|}$，那么 U 可写为 $U = S \cup T$，其中 a 在 S 上取值为 1，在 T 上取值为 0；由此得到的容量可记为 $\mathrm{cap}(S, T)$。对于对称矩阵，向量 q 正是各顶点处的电压：将 S 中的顶点设为电压 1，将 T 中的顶点设为电压 0。此外，如果对整个图进行修改，为 S 设一个顶点 s，为 T 设一个顶点 t，并在二者之间连一条边，那么修改后图中 s 与 t 之间的有效电阻，或原图中集合 S 与 T 之间的有效电阻，恰好为

$$\frac{1}{\mathrm{cap}_{S,T}(A)}.$$

事实上，通过将容量用于对称矩阵，可以得到许多不同的结果并设计出相应的算法，在这方面已有大量成果，例如参见 [DS84；SS11；Lyo83；Cha+96；Tet91]。

我们的目标是以最一般的形式研究非对称矩阵的容量。我们给出了许多基本结果，其中有些对于对称矩阵而言是众所周知的，但我们的主要结果如下，将在第5.11节讨论并证明。

定理1.6。设 $A_\alpha = \alpha A + (1-\alpha) A^T$，那么对于每个 U 和 a，若 $0 \le \beta \le \alpha \le \frac{1}{2}$，则

$$cap_{U,a}(A_\alpha) \le cap_{U,a}(A_\beta).$$

关于这一结果，我们再补充几点说明。首先，如果将非对称矩阵与对称矩阵相比较，这个结果很容易得到；也就是说，很容易证明，对于 $\tilde{A} = \frac{1}{2}(A + A^T)$，

$$\mathrm{cap}_{U,a}(\tilde{A}) \le \mathrm{cap}_{U,a}(A).$$

这实际上是狄利克雷定理的直接推论，而狄利克雷定理本身又是柯西-施瓦茨不等式的推论。定理 1.6 的非凡之处在于，它有助于比较两个非对称矩阵的容量，因此无法使用针对对称矩阵的一般工具——具体来说，在此情形下就是柯西-施瓦茨不等式。证明相对冗长且复杂，但我们认为，容量会随着 α 减小而单调增加，并从1 2 变为 0，这一点十分引人注目。其他相关结果也可见第5.11节。

另一种扩张概念与张量游走

在最后一节中，我们初步探讨了一种新的扩张定义，其灵感来自近期对高维扩张器中扩张现象的研究 [Baf+20]，并在第 5.14 节给出了一个简洁的引理，类似于该定义下标准的边扩张概念。我们认为，这一引理可能有助于研究高维对象。我们研究的第二个也是最后一个主题是张量游走。随着人们对机器学习和量子信息的兴趣迅速升温，张量已成为这两个领域的基础工具。然而，与张量相关的研究结果都极为依赖具体应用，而证明一般性结果的工作才刚刚起步。这主要是因为张量的行为几乎在所有方面都不同于矩阵。为说明这一点，考虑一个三阶张量 T，其索引为$T_{i,j,k}$，其中 $i, j, k \in [n]$。给定概率向量 $p_t \in \mathbb{R}^n_{\geq 0}$，我们可以将张量游走的一步定义为

$$p_t(i) = \sum_{j,k} T_{i,j,k} p_{t-1}(j) p_{t-2}(k).$$

为了确保 p_t 也是概率分布，我们可以要求对任意 j， k，$\sum_i T_{i,\ j,\ k} = 1$均成立。需要注意的是，张量游走的一步不再是线性运算，但它是一种非常有意义的游走，而且这类游走自然出现在许多领域中。同一定义可以推广到任意k阶张量。令人惊讶的是，即使是三阶张量，也可能有数量呈指数级增长（关于n）的特征值（以适当方式定义）。非负张量甚至有多个佩龙-弗罗贝尼乌斯定理变体（例如参见 [CPZ08；FGH13] ），然而，许多基本问题仍未得到解答。首先要问的是，正张量是否有唯一不动点？事实上，令人惊讶的是，答案是否定的， [CZ13]中给出了一个例子。我们的主要结果是阐明一个充分条件，使得该游走具有唯一不动点。若一个 k阶张量在指标 i上是随机的，是指对其他指标的每一种固定方式 l_1， l_2，……， l_{k-1}，

$$\sum_{i=1}^{n} T_{l_1,l_2,i,l_3,\ldots,l_{k-1}} = 1.$$

如果 T 在输出指标 i 和任意输入指标上都是随机的，我们就称其为 2 行随机。我们的主定理如下。

定理 1.7。 设 T 是一个具有正矩阵元素的 n 维 k 阶张量，且 T 是2-行随机的。那么 T 有唯一的正不动点。

这一结果可见第 5.15 节。它引出了一个极其重要的问题：游走收敛到不动点的速度有多快？这个问题留待今后探讨。

本文的结构安排如下。我们首先在第2节给出预备定义，接着在第1.2节证明定理 3，在第 1.3节证明定理4.8和 1.4 节中的定理 4.10，并分别在第 1.5、1.6 和1.7 节中证明定理5.1、 5.11 和 5.15。最后，我们将在第 5.15节总结全文。

第 2 章

预备知识——谱隙、边扩张与佩龙-弗罗贝尼乌斯定理

> 至于其余读者，他们可以接受适用于自己的部分。我相信，没有人会为了穿上这件外衣而把接缝撑开，因为它对适合它的人会大有裨益。
>
> ~亨利·戴维·梭罗，《瓦尔登湖》

在本学位论文中，我们将始终通过 R 或 A来研究非负矩阵，它们也称为链或图。对于任意非负矩阵 $R \in \mathbb{R}_{\geq 0}^{n \times n}$，都隐含着一个具有 n 个顶点的底层图，其中边 (i, j) 的权重为 $R_{j,i}$；我们假设当 $R_{j,i} = 0$ 时不存在边。请注意，由于我们假设向量是右乘的，下标的顺序是反过来的；不过，除了为保持引理和证明内部一致性之外，这一点在本学位论文中不会产生任何影响。

给定矩阵 R，如果底层有向图中，对每一对顶点 (s, t)，在权重为正的边构成的图上都存在一条从 s 到 t 的路径，我们就称 R 是强连通的或不可约的；也就是说，对于每一对 (s, t)，都存在 $k > 0$，使得 $R^k(s, t) > 0$。如果存在一对顶点 (s, t)，使得在底层有向图（由权重为正的边构成）中存在一条从 s 到 t 的路径，但不存在从 t 到 s 的路径，我们就称 R 是弱连通的。

我们首先陈述佩龙-弗罗贝尼乌斯定理。该定理关于正矩阵的结论由佩龙提出 R，发表于 [Per07]，；关于不可约矩阵的结论则由弗罗贝尼乌斯提出 R，发表于 [Fro12]。在过去一百年中，正如引言中所述，该定理已广泛应用于数学的许多不同领域，并已成为谱理论和动力系统中的基本工具。

定理 2.1。（佩龙-弗罗贝尼乌斯定理 [Per07；Fro12]）设 $R \in \mathbb{R}^{n\times n}$ 为一个非负矩阵。则对于 R，以下结论成立。

1. R 有一个非负特征值 r，使得所有其他特征值的模至多为 r，并且 R 有对应于 r的非负左、右特征向量 u 和v 。*2.* 如果 R 有对应于某个特征值 λ的正左、右特征向量 u' 和 v' ，那么 $\lambda = r$。*3.* 如果 R 不可约，那么特征值 r 为正且是单特征值（唯一），相应的左、右特征向量 u 和 v 均为正且唯一，并且所有其他特征值 λ 都满足 $|\lambda| \leq r$，且 $Re\lambda < r$。

在线上不同的讲义中可以找到许多关于该定理的优美证明，直接证明它也并不困难。我们将对不可约矩阵陈述所有结果，并可利用下极限将它们推广到一般矩阵。根据佩龙-弗罗贝尼乌斯定理（第2.1节第 3 部分），不可约非负矩阵 R 有一个单特征值，且该特征值为正 r，并且所有特征值的模都严格小于 r；这个特征值称为平凡或随机或 R的佩龙–弗罗贝尼乌斯特征值，而 R 的所有其他特征值称为非平凡特征值。对应于 r 的左、右特征向量称为平凡或佩龙–弗罗贝尼乌斯左特征向量（通常记作 u）和平凡或佩龙–弗罗贝尼乌斯右特征向量（通常记作 v）。由此引出以下定义。

定义 2.2.（不可约非负矩阵的谱隙）设 R 是一个 $n \times n$不可约非负矩阵。设特征值 λ_1 至 λ_n 属于 R，并按如下方式排列 $\lambda_1 > \mathrm{Re}\lambda_2 \geq \mathrm{Re}\lambda_3 \geq \ldots \geq \mathrm{Re}\lambda_n$，其中根据佩龙-弗罗贝尼乌斯定理2.1， λ_1 是实数且为正，并且 $\mathrm{Re}\lambda_2 < \lambda_1$。定义 R 的谱隙为

$$\triangle(R) = 1 - \frac{\mathrm{Re}\lambda_2(R)}{\lambda_1(R)}.$$

◊

注意， $0 \leq \Delta(R) \leq 2$根据佩龙-弗罗贝尼乌斯定理可得，因为 $|\lambda_i(R)| < \lambda_1(R)$。提醒读者， R的特征值就是特征多项式的根，即方程 $\det(\lambda I - R) = 0$的根。我们还将考虑非负矩阵 A 的奇异值，其中正左、右特征向量相同，且对应佩龙–弗罗贝尼乌斯特征值 1，记为 $1 = \sigma_1(A) \geq \sigma_2(A) \geq \cdots \geq \sigma_n(A) \geq 0$

（证明见引理 3.3： $\sigma_1(A) = 1$）。我们将 $M \in \mathbb{C}^{n\times n}$的第 (i, j) 个元素记为 $M_{i,j}$；如果 M已经带有下标，则记为 $M_t(i, j)$，具体含义由上下文确定。我们将 M 的共轭转置记为 M^*，将 M 的转置记为 M^T。任何$M \in \mathbb{C}^{n\times n}$ 都有舒尔分解（例如参见 [Lax07]） $M = UTU^*$，其中 T 是上三角矩阵，其对角线元素为 M的特征值，而 U 是酉矩阵（$UU^* = U^*U = I$）。默认情况下，“向量”指列向量。对于向量 v，我们也用或 v_i 表示其第 i 个元素；如果已有下标，则记为 $v_t(i)$。对于任意两个向量 $x, y \in \mathbb{C}^n$，我们采用标准内积 $\langle x, y\rangle = \sum_{i=1}^n x_i^* \cdot y_i$，并据此定义范数 $\|x\|_2 = \sqrt{\langle x, x\rangle}$。我们用 $u \perp v$ 表示 $\langle u, v\rangle = 0$。注意， $\langle x, My\rangle = \langle M^*x, y\rangle$。我们将 M的算子范数记为 $\|M\|_2 = \max_{u:\|u\|_2=1} \|Mu\|_2$，并回顾一下，算子范数不超过弗罗贝尼乌斯范数，即 $\|M\|_2 \le \|M\|_F := \sqrt{\sum_{i,j} |M_{i,j}|^2}$。我们用$D_u$ 表示对角线由向量 u构成的对角矩阵。我们用“正”表示逐项正，用“扩张”表示边扩张。我们还注意到伯克霍夫-冯诺依曼定理：每个双随机矩阵 A 都可写成 $A = \sum \alpha_i P_i$，其中 P_i 为置换矩阵，且 $\alpha_i \ge 0$ 和 $\sum \alpha_i = 1$。

回顾实对称矩阵特征值的柯朗–费希尔变分刻画，并将其应用于第二大特征值：

$$\max_{u\perp v_1} \frac{\langle u, Mu\rangle}{\langle u, u\rangle} = \lambda_2(M),$$

其中， v_1 是 M的最大特征值对应的特征向量。我们用符号J表示全 1 矩阵除以 n，即 $J = \frac{1}{n}\mathbf{1}\cdot \mathbf{1}^T$。我们将标准基向量记为 e_1，……， e_n。我们用$\mathbf{1}$表示全 1 向量，并称任意子集 $S \subseteq [n]$ 为一个割，将其补集记为 S，并将割的示性向量记为 $\mathbf{1}_S$，其中 $\mathbf{1}_S(i) = 1$ 当 $i \in S$ 时取值为 1，否则取值为 0。我们首先介绍双随机矩阵的边扩张定义。

定义 2.3。 （双随机矩阵的边扩张）对于双随机矩阵 A，割 S 的边扩张定义为

$$\phi_S(A) := \frac{\langle \mathbf{1}_S, A\mathbf{1}_{\overline{S}}\rangle}{\langle \mathbf{1}_S, A\mathbf{1}\rangle}$$

并且 的边扩张定义为

$$\phi(A) = \min_{S,|S|\le n/2} \phi_S(A) = \min_{S,|S|\le n/2} \frac{\sum_{i\in S, j\in \overline{S}} A_{i,j}}{|S|}.$$

◊

我们希望将这些概念推广到一般非负矩阵 R。由于实矩阵的特征值和奇异值不受我们考虑 R 还是R^T的影响，因此有意义的边扩张定义也应具有同样的性质。不过请注意，定义2.3 只有在矩阵是欧拉矩阵时才具有这种独立性，即$R\mathbf{1} = R^T\mathbf{1}$。因此，为一般矩阵定义边扩张时，我们利用其左右特征向量 R u 和 v 对 D_uRD_v进行变换，得到，它确实是欧拉矩阵，因为

$$D_uRD_v\mathbf{1} = D_uRv = rD_uv = rD_uD_v\mathbf{1} = rD_vD_u\mathbf{1} = rD_vu = D_vR^Tu = D_vR^TD_u\mathbf{1}.$$

由于 D_uRD_v 是欧拉的，我们可以仿照双随机矩阵定义 R 的边扩张：

定义 2.4。（不可约非负矩阵的边扩张）设 $R \in \mathbb{R}^{n\times n}$是一个不可约非负矩阵，其对应于佩龙–弗罗贝尼乌斯特征值 r的正左、右特征向量分别为 u 和v。割的边扩张 S 定义为

$$\phi_S(R) := \frac{\langle \mathbf{1}_S, D_uRD_v\mathbf{1}_{\overline{S}}\rangle}{\langle \mathbf{1}_S, D_uRD_v\mathbf{1}\rangle} \tag{2.5}$$

而 的边扩张定义为

$$\phi(R) = \min_{S:\sum_{i\in S} u_i\cdot v_i \le \frac{1}{2}\sum_i u_iv_i} \phi_S(R) = \min_{S:\sum_{i\in S} u_i\cdot v_i \le \frac{1}{2}\sum_i u_iv_i} \frac{\sum_{i\in S, j\in\overline{S}} R_{i,j}\cdot u_i\cdot v_j}{\sum_{i\in S,j} R_{i,j}\cdot u_i\cdot v_i}.$$

◊

边扩张 ϕ（R）可以直观地理解为衡量 不可约程度的指标（尽管严格来说，两者并非完全对应）。如果 ϕ（R）较高，比如为一个常数，甚至达到 1/polylog(n)，那么 R 就是良好连通的（强连通）。随着 ϕ（R）变小，这意味着 R中存在两个（强）非连通分量，或者 R 具有一个类似汇点的分量。

我们的下一个目标是为非不可约的非负矩阵 R 定义 Δ（R）和 ϕ（R）。为此，我们考虑 R周围一个小球内的不可约矩阵，并取下确界的极限，从而得到一个定义良好的量。

定义 2.6。*(*任意非负矩阵的谱隙与边扩张*)* 令 E 为所有不可约非负矩阵组成的集合。对于任意非负矩阵 R，令

$$R_\epsilon = \{H : H \in E,\ \|R - H\|_F \le \epsilon\}.$$

那么，我们将 R 的谱隙定义为

$$\Delta(R) = \lim_{\epsilon \to 0} \inf_{H \in R_\epsilon} \Delta(H),$$

以及 R 的边扩张为

$$\phi(R) = \lim_{\epsilon \to 0} \inf_{H \in R_\epsilon} \phi(H).$$

◊

我们首先可以得出一个简单结论。

引理 2.7. 设 R 是一个非负矩阵。若 R 不可约，则 $\phi(R) > 0$。

证明. 由于 R 不可约，左、右佩龙–弗罗贝尼乌斯特征向量 u 和 v 均为正。反设 $\phi(R) = 0$，则存在某个 S 使得 $\phi_S(R) = 0$，这意味着 $\langle \mathbf{1}_S, D_u R D_v \mathbf{1}_{\bar{S}} \rangle = 0$；又由于 u和 v均为正，因此对于 $i \in S, j \in \bar{S}$，有 $R_{i,j} = 0$；同理，由于 $D_u R D_v$是欧拉的，对于 $\bar{i} \in S, j \in \bar{S}$，有 $R_{\bar{i},j} = 0$。这意味着不存在从 $i \in S$ 到 $j \in \bar{S}$的路径，因而 R不是不可约的，矛盾。

□

根据佩龙-弗罗贝尼乌斯定理（定理2.1第3部分），如果 R 是不可约的，那么 $\mathrm{Re}\lambda_2(R) < \lambda_1(R)$。然而，由于不可约性意味着边扩张为正（见引理2.7），这意味着定理1.2 的极限情形表明$\phi(R) > 0 \Leftrightarrow \Delta(R) > 0$，由此得到的结论比佩龙-弗罗贝尼乌斯定理强得多、也紧确得多。

第 3 章

佩龙-弗罗贝尼乌斯定理和切格-布泽不等式的推广

> 人类理性有一种独特的命运：在其认知中，它背负着一些无法摆脱的问题，因为这些问题是理性本性本身赋予它的；但它也无法回答这些问题，因为它们超出了人类理性的全部能力。
>
> ~ 伊曼努尔·康德，《纯粹理性批判》

本节旨在证明我们的主定理 1.2。正如引言中所述，谱理论中有两个著名定理。第一个是佩龙-弗罗贝尼乌斯定理，它除了给出其他结果外，还提供了一个将矩阵的边扩张与其第二特征值联系起来的定性表述。更具体地说，它指出，对于任意不可约非负矩阵 R ——该矩阵总是满足 $\phi(R) > 0$ ——谱隙 $\Delta(R) > 0$。第二个定理是如下的切格 - 布泽不等式。

定理 3.1。 (切格-布泽不等式 [Che70； Bus82； AM85； SJ89； Dod84； Nil91]) 设 R 是一个 可逆 非负矩阵。那么

$$\frac{1}{2} \cdot \Delta(R) \leq \phi(R) \leq \sqrt{2 \cdot \Delta(R)}.$$

我们对这一问题的理解仍存在一个空白：非可逆情形只有定性结果，而可逆情形则有定量结果。本节将主要关注这一问题。定理 1.2 的主要目标是证明一个形式类似于上述不等式的结论，但适用于 任意非负矩阵（不一定可逆）。

我们将首先针对不可约非负矩阵证明定理 1.2，并在第 3.8 节将其推广到所有矩阵（不一定不可约）。我们重述本章要证明的主定理。

定理 3.2。 设 R 为不可约非负矩阵。那么

$$\frac{1}{15} \cdot \frac{\Delta(R)}{n} \leq \phi(R) \leq \sqrt{2 \cdot \Delta(R)}.$$

关于 $\phi(R)$ 的上界类似于切格不等式，可通过将 Fiedler 对双随机矩阵的证明直接推广 [Fie95], 而下界则需要大量工作。

不过，在详细证明定理 1.2之前，我们将先证明切格-布泽不等式，既为求完整，也为求具体，因为我们对边扩张的定义2.4 与某些文献中的定义不同。我们将首先证明一个关键引理，它将在后续所有章节中用到。

引理 3.3。 设 R 为不可约非负矩阵，其对应于佩龙–弗罗贝尼乌斯特征值 1 的正（左、右）特征向量分别为 u 和 v，并已归一化使得 $\langle u, v \rangle = 1$。定义 $\bar{A} = \bar{D}_u^2 \bar{D}_v^2 R D_u^2 D_v^2$。那么，对于 A，以下结论成立：

1. $\phi(A) = \phi(R)$。2. 对于每个 i, $\lambda_i(A) = \lambda_i(R)$。

3. $\|A\|_2 = 1$。

4. 若 R 可逆，即 $D_u R D_v = D_v R^T D_u$，则 A 对称。

证明。 设矩阵 A 如定义所示，并设 $w = D_u^2 D_v^2 \mathbf{1}$。容易验证，$Aw = w$ 且 $A^T w = w$。此外，

$$\langle w, w \rangle = \langle D_u^{\frac{1}{2}} D_v^{\frac{1}{2}} \mathbf{1}, D_u^{\frac{1}{2}} D_v^{\frac{1}{2}} \mathbf{1} \rangle = \langle D_u \mathbf{1}, D_v \mathbf{1} \rangle = \langle u, v \rangle = 1$$

这里用到了以下事实：矩阵 D_u^2 和 D_v^2 是对角线矩阵，因此它们可交换，并且转置后保持不变。设 S 为任意集合。条件 $\sum_{i \in S} u_i \cdot v_i \leq \frac{1}{2}$ 转化为 $\sum_{i \in S} w_i^2 \leq \frac{1}{2}$，因为 $u_i \cdot v_i = w_i^2$。因此，对于任意集合 S，

其中 $\sum_{i\in S} u_i \cdot v_i = \sum_{i\in S} w_i^2 \leq \frac{1}{2}$，

$$\begin{aligned}
\phi_S(R) &= \frac{\langle \mathbf{1}_S, D_u R D_v \mathbf{1}_{\overline{S}} \rangle}{\langle \mathbf{1}_S, D_u D_v \mathbf{1}_S \rangle} \\
&= \frac{\langle \mathbf{1}_S, D_u D_u^{-\frac{1}{2}} D_v^{\frac{1}{2}} A D_u^{\frac{1}{2}} D_v^{-\frac{1}{2}} D_v \mathbf{1}_{\overline{S}} \rangle}{\langle \mathbf{1}_S, D_u^{\frac{1}{2}} D_v^{\frac{1}{2}} D_u^{\frac{1}{2}} D_v^{\frac{1}{2}} \mathbf{1}_S \rangle} \\
&= \frac{\langle \mathbf{1}_S, D_u^{\frac{1}{2}} D_v^{\frac{1}{2}} A D_u^{\frac{1}{2}} D_v^{\frac{1}{2}} \mathbf{1}_{\overline{S}} \rangle}{\langle \mathbf{1}_S, D_u^{\frac{1}{2}} D_v^{\frac{1}{2}} D_u^{\frac{1}{2}} D_v^{\frac{1}{2}} \mathbf{1}_S \rangle} \\
&= \frac{\langle \mathbf{1}_S, D_w A D_w \mathbf{1}_{\overline{S}} \rangle}{\langle \mathbf{1}_S, D_w^2 \mathbf{1}_S \rangle} \\
&= \phi_S(A)
\end{aligned}$$

并且 (1) 成立。此外，由于 A 是 R 的相似变换，所有特征值均保持不变，因此 (2) 成立。对于 (3)，考虑矩阵 $H = A^T A$。由于 w 是 A 的正左、右特征向量，即 $Aw = w$ 和 $A^T w = w$，我们有 $Hw = w$。但是，由于 A 是非负的，H 也非负；又由于它有一个对应特征值 1 的正特征向量 w，根据佩龙-弗罗贝尼乌斯（定理 2.1，第 2 部分），H 的佩龙–弗罗贝尼乌斯特征值为 1。但是 $\lambda_i(H) = \sigma_i^2(A)$，其中 σ_i（A）是 A 的第 i 大奇异值。因此，我们得到 σ_1^2（A）$= \lambda_1$（H）$= 1$，从而有 $\|A\|_2 = 1$。对于 (4)，

$$\begin{gathered}
D_u R D_v = D_v R^T D_u \Leftrightarrow \\
D_u D_u^{-\frac{1}{2}} D_v^{\frac{1}{2}} A D_u^{\frac{1}{2}} D_v^{-\frac{1}{2}} D_v = D_v D_u^{\frac{1}{2}} D_v^{-\frac{1}{2}} A^T D_u^{-\frac{1}{2}} D_v^{\frac{1}{2}} \Leftrightarrow \\
D_u^{\frac{1}{2}} D_v^{\frac{1}{2}} A D_u^{\frac{1}{2}} D_v^{\frac{1}{2}} = D_u^{\frac{1}{2}} D_v^{\frac{1}{2}} A^T D_u^{\frac{1}{2}} D_v^{\frac{1}{2}} \Leftrightarrow \\
A = A^T
\end{gathered}$$

如所需。 □

由引理 3.3 可知，我们可以将证明限制为展示 A，而非 R，因为谱隙和边扩张均不受影响。

现在我们将证明切格-布泽不等式。文献中有许多广为人知的证明，但针对一般矩阵，$\phi(R)$ 的定义往往不尽如人意（尽管在技术上可行）。我们将完整证明定义 2.4 所给出的 $\phi(R)$，该定义与 [Mih89]中的定义相同。本文的论述也有助于将其与非可逆情形下类似的引理进行对比。

3.1 Buser 不等式—— $\phi(R)$的下界

我们首先证明可逆 R的 $\phi(R)$下界。为此，注意在可逆 R的情形下，可约 R 的 $\phi(R)$定义与不可约情形完全相同（依据预备知识中所述的第二特征值的变分刻画），因此我们只需证明不可约可逆 R的定理。对于任意这样的 R，由 Perron-Frobenius 定理 2.1 可知 $\lambda_1(R) > 0$；将矩阵乘以一个常数进行重新缩放——这不会改变 $\Delta(R)$或 $\phi(R)$（引理 3.3）——之后，我们可假设 $\lambda_1(R) = 1$。此外，我们将考察 $A = D_u^2 D_v{}^2 R D_u{}^2 D_v^2$；它是对称的，因为根据引理 3.3，$R$可逆；它也是不可约的，并且与$R$具有相同的谱隙和边扩张。因此，我们将证明 $\phi(A)$的下界，其中假设 $Aw = w$ 和$A^T w = w$ 对某个正 w 成立（定理 2.1）。利用 A 的谱分解即可轻松得到该下界。注意，以下引理不要求 A 可逆。

引理 3.4. 设 A 是任意一个最大特征值为 *1*、对应正左、右特征向量为 w 的不可约非负矩阵。若 $\langle v, Av\rangle \le \gamma\langle v, v\rangle$ 满足：对于每个向量 v，若 $v \perp w$, 则有 $0 \le \gamma \le 1$，那么

$$\frac{1-\gamma}{2} \le \phi(A).$$

证明. 设 $A = ww^* + B$ 满足 $Bw = 0$，则 $\langle v, Av\rangle = \langle v, Bv\rangle$；对于任意集合 S, 令 $D_w 1_S = c_0 w + v$，其中 $v \perp w$。那么

$$\langle w, D_w 1_S\rangle = c_0$$

and

$$\langle D_w 1_S, D_w 1_S\rangle = |c_0|^2 + \langle v, v\rangle$$

并注意到

$$\langle w, D_w 1_S\rangle = \langle D_w 1_S, D_w 1_S\rangle = \langle D_w 1, D_w 1_S\rangle.$$

那么，对于任意满足 $\langle D_w 1_S, D_w 1_S\rangle \le 1/2$ 的集合 S，我们有

$$\begin{aligned}
\phi_S(A) &= \frac{\langle D_w 1_{\bar{S}}, A D_w 1_S\rangle}{\langle D_w 1, D_w 1_S\rangle} \\
&= 1 - \frac{\langle D_w 1_S, A D_w 1_S\rangle}{\langle D_w 1, D_w 1_S\rangle} \\
&= 1 - \frac{\langle c_0 w + v, A(c_0 w + v)\rangle}{\langle D_w 1, D_w 1_S\rangle} \\
&= 1 - \frac{\langle c_0 w + v, c_0 w + Av\rangle}{\langle D_w 1, D_w 1_S\rangle} \\
&= 1 - \frac{|c_0|^2 + \langle v, Av\rangle}{\langle D_w 1, D_w 1_S\rangle} \\
&\ge 1 - \frac{|c_0|^2 + \gamma\langle v, v\rangle}{\langle D_w 1, D_w 1_S\rangle} \\
&= 1 - \frac{\langle D_w 1_S, D_w 1_S\rangle^2 + \gamma(\langle D_w 1_S, D_w 1_S\rangle - \langle D_w 1_S, D_w 1_S\rangle^2)}{\langle D_w 1, D_w 1_S\rangle} \\
&= 1 - (\gamma + (1-\gamma)\langle D_w 1_S, D_w 1_S\rangle) \\
&\quad [\text{since } \langle D_w 1_S, D_w 1_S\rangle \le 1/2] \\
&\ge \frac{1-\gamma}{2}
\end{aligned}$$

如所需。 □

Buser 不等式可由下述内容轻易推出。

引理 3.5。对于任意 $v \perp w$，$\langle v, Av\rangle \le \lambda_2(A)\langle v, v\rangle$

证明。设 $u_1 = w$，u_2，u_3，……，u_n 是 A 的一组正交特征向量（参见预备知识 2），并设 $v = \sum_i c_i u_i$ 满足 $c_1 = 0$ 因为 $v \perp w$，且

$$\langle v, Av\rangle = \langle \sum_i c_i u_i, A\sum_i c_i u_i\rangle = \langle \sum_i c_i u_i, \sum_i c_i \lambda_i u_i\rangle = \sum_{i\ge 2} c_i^2 \lambda_i \le \lambda_2 \sum_i c_i^2 = \lambda_2\langle v, v\rangle$$

如所需。 □

结合引理 3.4 和引理 3.5，可得定理 $\phi(A)$ 中 3.1关于 下界。

3.2 Cheeger不等式——ϕ（R）的上界

现在我们继续讨论 ϕ（A）的上界，这相对较为复杂。我们的证明方法与 [Mih89], [Chu07] 中的方法类似，并可追溯至 [AM85]。

引理 3.6. $\phi(A) \le \sqrt{2 \cdot \Delta(A)}$。

证明。令 v为对应特征值 $\lambda_2(A)$的特征向量，其中 $v \perp w$，并假设 $\langle w, w\rangle = 1$。于是我们有

$$1 - \lambda_2 = \frac{\langle v, (I - A)v\rangle}{\langle v, v\rangle}.$$

令 $u = D_w^{-1}v - c \cdot \mathbf{1}$，并注意到 $\langle D_w u, D_w u\rangle = \langle v, v\rangle + c^2$，因为 $v \perp w$。我们的目标是选取 c，以便将 u 的正、负矩阵元素分别分成 S和 $\bar{S}$两组，使得 $\sum_{i\in\bar{S}} w_i^2 \le 1/2$ 且 $\sum_{i\in S} w_i^2 \le 1/2$。为此，不失一般性，假设 $D_w^{-1}v$ 的矩阵元素按递减顺序排列（指标从 1 到 n递增），并令 r 为满足 $\sum_{i=1}^{r} w_i^2 > 1/2$ 的最小指标。因此，$\sum_{i=1}^{r-1} w_i^2 \le 1/2$ 且 $\sum_{i=r+1}^{n} w_i^2 \le 1/2$，因为 $\sum_i w_i^2 = 1$。接着我们选取 $c = v_r/w_r$，由此得到 $i < r$对应的 $u_i \ge 0$、$u_r = 0$以及对应于 $i > r$ 的 $u_i \le 0$。于是，令 x 为一个向量，其在 $i < r$ 处的值为 $x_i = u_i$，其他位置为 0；令 y 为一个向量，其在$i > r$ 处的值为 $y_i = u_i$，其他位置为 0；我们得到 $u = x - y$，并且当 i遍历 x 或 y的非零项时，$\sum_i w_i^2 \le 1/2$，且 $\langle x, y\rangle = 0$。因此，我们有

$$\begin{aligned}
1 - \lambda_2 &= \frac{\langle v, (I - A)v\rangle}{\langle v, v\rangle} \\
&= \frac{\langle D_w u, (I - A)D_w u\rangle}{\langle D_w u, D_w u\rangle + c^2} \\
&\quad [\text{since } (I - A)w = 0 \text{ and } (I - A)^T w = 0] \\
&\ge \frac{\langle D_w x, (I - A)D_w x\rangle + \langle D_w y, (I - A)D_w y\rangle - 2\langle D_w x, (I - A)D_w y\rangle}{\langle D_w x, D_w x\rangle + \langle D_w y, D_w y\rangle} \\
&\ge \frac{\langle D_w x, (I - A)D_w x\rangle + \langle D_w y, (I - A)D_w y\rangle}{\langle D_w x, D_w x\rangle + \langle D_w y, D_w y\rangle} \\
&\quad [\text{since } \langle D_w x, D_w y\rangle - \langle D_w x, A \cdot D_w y\rangle = 0 - \langle D_w x, A \cdot D_w y\rangle \le 0, \\
&\quad \text{as all entries of } x, y, w, A \text{ are nonnegative}] \\
&\ge \frac{\langle D_w x, (I - A)D_w x\rangle}{\langle D_w x, D_w x\rangle} \qquad (3.7)
\end{aligned}$$

其中我们在最后一行假设

$$\frac{\langle D_w x, (I - A)D_w x\rangle}{\langle D_w x, D_w x\rangle} \le \frac{\langle D_w y, (I - A)D_w y\rangle}{\langle D_w y, D_w y\rangle}.$$

注意到 $Aw = w$和 $A^T w = w$，因此对任意 i, $w_i^2 = \sum_j a_{i,j} w_i w_j$，对任意 j,

$w_j^2 = \sum_i a_{i,\ j} w_i w_j$。令 $a'_{i,\ j} = a_{i,\ j} w_i w_j$。因此

$$\begin{aligned}
\langle D_w x, D_w x \rangle &= \sum_i w_i^2 x_i^2 = \frac{1}{2} \cdot \left(\sum_i w_i^2 x_i^2 + \sum_j w_j^2 x_j^2 \right) \\
&= \frac{1}{2} \cdot \left(\sum_{i,j} a_{i,j} w_j w_i x_i^2 + \sum_{j,i} a_{i,j} w_i w_j x_j^2 \right) \\
&= \frac{1}{2} \cdot \sum_{i,j} a_{i,j} w_i w_j (x_i^2 + x_j^2) \\
&\geq \frac{1}{4} \cdot \sum_{i,j} a'_{i,j} (x_i + x_j)^2
\end{aligned}$$

and

$$\langle D_w x, (I - A) D_w x \rangle = \frac{1}{2} \sum_{i,j} a'_{i,j} (x_i - x_j)^2$$

并将两个方程的左侧相乘，得到

$$\begin{aligned}
\langle D_w x, D_w x \rangle \langle D_w x, (I - A) D_w x \rangle &\geq \frac{1}{8} \sum_{i,j} a'_{i,j} (x_i + x_j)^2 \sum_{i,j} a'_{i,j} (x_i - x_j)^2 \\
&\geq \frac{1}{8} \left(\sum_{i,j} \sqrt{a'_{i,j}} \left| x_i - x_j \right| \cdot \sqrt{a'_{i,j}} \left| x_i + x_j \right| \right)^2 \\
&\quad \text{[Cauchy-Shwarz]} \\
&= \frac{1}{8} \left(\sum_{i,j} a'_{i,j} \left| x_i^2 - x_j^2 \right| \right)^2
\end{aligned}$$

由于 $x_i \geq x_j$对于 $i < j$，因此对于$m'_{i,j} = a'_{i,j} + a'_{j,i}$

$$\sum_{i,j} a'_{i,j} \left| x_i^2 - x_j^2 \right| = \sum_{i<j} m'_{i,j}(x_i^2 - x_j^2)$$

$$= \sum_{i=1}^{n} \sum_{j=i+1}^{n} \sum_{k=i}^{j-1} m'_{i,j}(x_k^2 - x_{k+1}^2)$$

$$= \sum_{i=1}^{n} \sum_{k=i}^{n-1} \sum_{j=k+1}^{n} m'_{i,j}(x_k^2 - x_{k+1}^2)$$

$$= \sum_{k=1}^{n-1} \sum_{i=1}^{k} \sum_{j=k+1}^{n} m'_{i,j}(x_k^2 - x_{k+1}^2)$$

$$= \sum_{k=1}^{r-1} (x_k^2 - x_{k+1}^2) \sum_{i=1}^{k} \sum_{j=k+1}^{n} m'_{i,j}$$

[since for $k \geq r$, $x_k = 0$]

$$= 2 \sum_{k=1}^{r-1} (x_k^2 - x_{k+1}^2) \cdot \phi_k \cdot \mu_k$$

[2 的因子，因为 $m_{i,\ j}$ 已经是两个矩阵元素之和，

and $\mu_k = \sum_{i=1}^{k} w_i^2 \leq 1/2$ for any $k \leq r$]

$$\geq 2\alpha \cdot \sum_{k=1}^{r-1} (x_k^2 - x_{k+1}^2) \cdot \mu_k$$

[letting α=$\min_k \phi_k$]

$$= 2\alpha \cdot \sum_{k=1}^{r-1} w_k^2 x_k^2$$

$$= 2\alpha \cdot \langle D_w x, D_w x \rangle$$

最终将所有内容合并后得到,

$$\begin{aligned}
1-\lambda_2 &= \frac{\langle v,(I-A)v\rangle}{\langle v,v\rangle} \\
&\geq \frac{\langle D_w x,(I-A)D_w x\rangle}{\langle D_w x,D_w x\rangle} \\
&\quad \text{[from equation 3.7]} \\
&\geq \frac{1}{8}\cdot\frac{(2\alpha\cdot\langle D_w x,D_w x\rangle)^2}{\langle D_w x,D_w x\rangle^2} \\
&= \frac{1}{2}\cdot\alpha^2
\end{aligned}$$

如所要求。 □

3.3 Cheeger-Buser不等式的紧性

通过直接计算可得，对于具有 n 个顶点的超立方体 H_n，其中 n 是 2 的幂,有

$$\phi(H_n)=\frac{1}{2}\cdot\Delta(H_n)$$

这表明 Buser 不等式恰好是紧的；对于具有 n 个顶点的无向环 C_n，我们得到

$$\phi(C_n)\geq\sqrt{\Delta(C_n)}$$

这表明 Cheeger不等式在常数范围内是紧的。接下来，我们证明定理 $\phi(R)$ 中的 3.2 上界。

3.4 Fiedler的证明– 关于 $\Delta(R)$ 的 $\phi(R)$ 上界

根据切格不等式（引理 3.6），ϕ（R）的上界相对容易得到。我们证明引理 A 3.3 中定义的上界。

引理 3.8。（*Fiedler [Fie95]*的推广）设 R 是一个不可约非负矩阵，其对应于特征值 *1* 的正（左、右）特征向量为 u 和 v，并设 $A=D_u^{\frac{1}{2}}D_v^{-\frac{1}{2}}RD_u^{-\frac{1}{2}}D_v^{\frac{1}{2}}$。则

$$\phi(A)\leq\sqrt{2\cdot\Delta(A)}.$$

证明。 给定如上所述的 R，并令

$$A=D_u^{\frac{1}{2}}D_v^{-\frac{1}{2}}RD_u^{-\frac{1}{2}}D_v^{\frac{1}{2}},$$

注意,

$$w=D_u^{\frac{1}{2}}D_v^{\frac{1}{2}}\mathbf{1},$$

其中对角矩阵中的矩阵元素取正平方根。由于 A的左右特征向量均为 w，且对应特征值 1，因此 $M = \frac{A+A^T}{2}$ 也具有相同的性质。 ——

如前文关于一般非负矩阵的边扩张的定义 2.4 中所述， D_wAD_w是欧拉的，因为 $D_wAD_w\mathbf{1} = D_wAw = D_ww = D_w^2\mathbf{1} = D_ww = D_wA^Tw = D_wA^TD_w\mathbf{1}$。因此，对于任意 S,

$$\langle \mathbf{1}_S, D_wAD_w\mathbf{1}_{\overline{S}}\rangle = \langle \mathbf{1}_{\overline{S}}, D_wAD_w\mathbf{1}_S\rangle = \langle \mathbf{1}_S, D_wA^TD_w\mathbf{1}_{\overline{S}}\rangle,$$

因此，对于任意满足 $\sum_{i\in S} w_i^2 \le \frac{1}{2}\sum_i w_i^2$ 的集合 S，

$$\phi_S(A) = \frac{\langle \mathbf{1}_S, D_wAD_w\mathbf{1}_{\overline{S}}\rangle}{\langle \mathbf{1}_S, D_wAD_w\mathbf{1}\rangle} = \frac{1}{2}\cdot\frac{\langle \mathbf{1}_S, D_wAD_w\mathbf{1}_{\overline{S}}\rangle + \langle \mathbf{1}_S, D_wA^TD_w\mathbf{1}_{\overline{S}}\rangle}{\langle \mathbf{1}_S, D_wAD_w\mathbf{1}\rangle} = \phi_S(M)$$

因此

$$\phi(A) = \phi(M). \tag{3.9}$$

对于任意矩阵 H，令

$$R_H(x) = \frac{\langle x, Hx\rangle}{\langle x, x\rangle}.$$

对于每个 $x \in \mathbb{C}^n$,

$$\mathrm{Re}R_A(x) = R_M(x), \tag{3.10}$$

因为 A 和 $\langle x, x\rangle$ 均为非负的，并且我们可以写成

$$2\cdot R_M(x) = \frac{\langle x, Ax\rangle}{\langle x, x\rangle} + \frac{\langle x, A^*x\rangle}{\langle x, x\rangle} = \frac{\langle x, Ax\rangle}{\langle x, x\rangle} + \frac{\langle Ax, x\rangle}{\langle x, x\rangle} = \frac{\langle x, Ax\rangle}{\langle x, x\rangle} + \frac{\langle x, Ax\rangle^*}{\langle x, x\rangle} = 2\cdot\mathrm{Re}R_A(x).$$

此外,

$$\mathrm{Re}\lambda_2(A) \le \lambda_2(M). \tag{3.11}$$

为说明这一点，首先注意到 $\lambda_2(A) < 1$，因为 A 是不可约的。因此，由于 $\lambda_2(A) \neq 1$，令 v 为对应于 $\lambda_2(A)$ 的特征向量。又由于 $Aw = w$，对于 $\lambda_2 = \lambda_2(A)$，我们有,

$$Av = \lambda_2 v \Rightarrow \langle w, Av\rangle = \langle w, \lambda_2 v\rangle \Leftrightarrow \langle A^Tw, v\rangle = \lambda_2\langle w, v\rangle \Leftrightarrow (1-\lambda_2)\langle w, v\rangle = 0$$

这意味着 $v \perp w$。因此，我们有

$$\mathrm{Re}\lambda_2(A) = \mathrm{Re}\frac{\langle v, Av\rangle}{\langle v, v\rangle} = \frac{\langle v, Mv\rangle}{\langle v, v\rangle} \le \max_{u\perp w}\frac{\langle u, Mu\rangle}{\langle u, u\rangle} = \lambda_2(M)$$

其中，第二个等号用到了方程3.10，最后一个等号则由预备知识 2中陈述的特征值的变分刻画得出。因此，利用方程 3.11、方程 3.9 以及关于 M 的切格不等式（定理 3.1），得到

$$\phi(A) = \phi(M) \le \sqrt{2\cdot(1-\lambda_2(M))} \le \sqrt{2\cdot(1-\mathrm{Re}\lambda_2(A))} = \sqrt{2\cdot\Delta(A)}$$

如所要求。 □

3.5 奇异值的作用

在本节中，我们尝试通过模仿定理 ϕ 中 Buser 不等式的证明，得到 3.1 的一个下界。下面的引理可立即得出。

引理 3.12. 设 A 是最大特征值为 *1* 的不可约非负矩阵， w 为其对应的正左、右特征向量。那么

$$\frac{1-\sigma_2(A)}{2} \le \phi(A),$$

其中， $\sigma_2(A)$ 是 A 的第二大奇异值。

证明. 证明可直接得出，因为对于 $v \perp w, \langle v, Av\rangle \le \sigma_2(A)\langle v, v\rangle$ 由基本线性代数可知，再利用引理 3.4 即可得到结果。 □

我们注意到，引理3.12 在以下意义上并无太大意义——设 A 为一个有向环，则 $\sigma_2(A) = 1$ 但 $\phi(A) \approx 1/n$，这表明该界并不连续。使用德布鲁因矩阵可以得到更大的差距——第 4.4 节将对此进行讨论。对于这些矩阵， $\sigma_2(A) = 1$ 但 $\phi(A) \approx 1/\log n$。

另一种思路或许是先确保 A 是惰性的，或者写作 $A' = \frac{1}{2}A + \frac{1}{2}I$；这样一来，$\sigma_2(A')$ 对于 $\phi(A') = \frac{1}{2}\phi(A)$ 可能会更有意义。不过，注意到 $A = w \cdot w^T + B$ 和 $A^T = w \cdot w^T + B^T$，我们有： $\tilde{A} = \frac{1}{2}(A + A^T)$ 的第二大特征值

$$\frac{1}{2}(A + A^T) = w \cdot w^T + \frac{1}{2}(B + B^T)$$

恰好是 $\tilde{B} = \frac{1}{2}(B + B^T)$ 的最大特征值。此外，由于 A 满足引理 3.3 中的条件，即 $\|A\|_2 = 1$ ，且它的特征值 1 对应的左右特征向量相同，均为 w ，因此可知 $\sigma_2(A) = \|B\|_2$。于是，设 x，其中 $\|x\|_2 = 1$ 是 $\frac{1}{2}(B + B^T)$ 最大特征值对应的特征向量，则

$$\lambda_2\left(\frac{1}{2}(A + A^T)\right) = \langle x, \frac{1}{2}(B+B^T)x\rangle \le \frac{1}{2}\|x\|_2\|Bx\|_2 + \frac{1}{2}\|x\|_2\|B^T x\|_2 = \|B\|_2 = \sigma_2(A).$$

因此，使用简易切格尔，我们可以得到

$$\frac{1}{2} - \frac{1}{2}\sigma_2(A) \le \frac{1}{2} - \frac{1}{4}\lambda_2(A + A^T) \le \phi(A),$$

而奇异值也不再有意义，因为它们给出的下界较弱。

3.6 微扰理论的局限

本节将讨论一个微扰界，以说明如何利用微扰理论中的工具，根据谱隙证明 $\phi(R)$ 的下界有多困难，同时也阐明它与定理 1.2 之间的差异。鲍尔 - 菲克界表述如下。

引理 3.13. *(鲍尔-菲克 [BF60])* 设 A 是任意一个可对角化矩阵，满足 $A = VGV^{-1}$ 是 A 的若尔当标准形，且 $\mathcal{K}_p(A) = \|V\|_p\|V^{-1}\|_p$。设 E 是 A的一个微扰，满足 $A' = A + E$。那么，对于每个特征值 μ， A'都有一个特征值 λ A 使得

$$|\mu - \lambda| \leq \mathcal{K}_p(A) \cdot \|E\|_p.$$

证明很简单，这里不再重述。为了与定理 1.2形成对比，设S 是使 ϕ 达到最大值的集合， A (假设其最大特征值为 1)，再设 A' 是两个非连通分量，其中流出集合 S和流入集合的质量已被移除，并重新放回各集合内部，以确保 A' 的最大特征值为 1。注意， $\lambda_2(A') = \lambda_2(A') = 1 = \lambda_1(A)$，因此，由上述引理可得

$$\Delta(A) \leq \mathcal{K}_p(A) \cdot \|E\|_p.$$

请注意，即使我们能够得到这样的界 $\|E\|_p \approx \phi(A)$，$\mathcal{K}(A)$对它的依赖也无法避免，而且它仍然仅适用于可对角化矩阵。Saad [Saa11] 对一般矩阵的推广。

引理 3.14. *(Saad [Saa11]* 对鲍尔-菲克的推广）设 A 为任意矩阵（不一定可对角化），并设 $\mathcal{K}_2(A) = \|V\|_2\|V^{-1}\|_2$为其若尔当分解 $A = VGV^{-1}$。设 E 是 A的一个微扰，且 $A' = A + E$。那么，对于每个特征值 μ， A'都存在 λ 的一个特征值 A 使得

$$1 \leq \frac{1}{|\lambda - \mu|} \sum_{i=0}^{l-1} \left(\frac{1}{|\lambda - \mu|} \right)^i \mathcal{K}_2(A)\|E\|_2$$

其中 l 是最大若尔当块的大小。

我们可以用舒尔分解而不是若尔当分解重新证明该引理，并得到以下结论。

引理 3.15。 设 A 为任意矩阵（不一定可对角化），其最大奇异值为 $\sigma = \|A\|_2$。设 E是 A的微扰，且满足 $A' = A + E$。那么

对于 A'的每个特征值 μ， A 都有一个特征值 λ，使得

$$1 \leq \frac{1}{|\lambda - \mu|} \left(1 + \frac{\sigma}{|\lambda - \mu|}\right)^{n-1} \|E\|_2.$$

证明。令 $A = UTU^*$ 为 A的舒尔分解。令 T 为一个上三角矩阵，满足 $|T_{i,i}| \geq \alpha > 0$ 这意味着 T 可逆，且 $\|T\| \leq \sigma$。令 T^{-1} 中与对角线距离为 k 的元素的最大值为 γ_k。注意，对于 T^{-1} 的对角线元素，我们有 $|T^{-1}(k, k)| \leq \frac{1}{\alpha}$。因此 $\gamma_0 \leq \frac{1}{\alpha}$。假设对于 $k \geq 1$，

$$\gamma_k \leq \frac{\sigma}{\alpha^2} \left(1 + \frac{\sigma}{\alpha}\right)^{k-1}.$$

对于 $k = 1$，容易验证，对于任意行 r，

$$T^{-1}(r, r)T(r, r+1) + T^{-1}(r, r+1)T(r+1, r+1) = 0$$

因此

$$|T^{-1}(r, r+1)| \leq \frac{|T^{-1}(r, r)| \cdot |T(r, r+1)|}{|T(r+1, r+1)|} \leq \frac{\frac{1}{\alpha} \cdot \sigma}{\alpha} = \frac{\sigma}{\alpha^2}$$

正如所要求的。假设上述等式对所有 γ_l 和 $l \leq k$ 都成立，我们将证明 γ_{k+1}。取 T^{-1} 中任意一个与对角线距离为 $x = k + 1$ 的元素 x。由于 $T^{-1}T = I$，我们可得，对于某个固定的行 r 和列 $r + x$，

$$\sum_j T^{-1}(r, j)T(j, r+x) = 0$$

$$\sum_{j=r}^{r+x} T^{-1}(r, j)T(j, r+x) = 0$$

$$\sum_{j=r}^{r+x-1} T^{-1}(r, j)T(j, r+x) + T^{-1}(r, r+x)T(r+x, r+x) = 0$$

因此

$$
\begin{aligned}
|T^{-1}(r,r+x)| &\le \frac{1}{T(r+x,r+x)} \sum_{j=r}^{r+x-1} |T^{-1}(r,j)| \cdot |T(j,r+x)| \\
&\le \frac{1}{\alpha}\left(\frac{\sigma}{\alpha} + \sum_{j=r+1}^{r+x-1} \gamma_{j-r}\cdot\sigma\right) \\
&= \frac{\sigma}{\alpha}\left(\frac{1}{\alpha} + \sum_{j-r=1}^{j-r=x-1} \gamma_{j-r}\right) \\
&= \frac{\sigma}{\alpha}\left(\frac{1}{\alpha} + \sum_{l=1}^{l=k} \gamma_{l}\right) \\
&= \frac{\sigma}{\alpha}\left(\frac{1}{\alpha} + \sum_{l=1}^{l=k} \frac{\sigma}{\alpha^2}\left(1+\frac{\sigma}{\alpha}\right)^{l-1}\right) \\
&= \frac{\sigma}{\alpha}\left(\frac{1}{\alpha} + \frac{\sigma}{\alpha^2}\frac{(1+\frac{\sigma}{\alpha})^{k}-1}{(1+\frac{\sigma}{\alpha})-1}\right) \\
&= \frac{\sigma}{\alpha^2}\left(1+\frac{\sigma}{\alpha}\right)^{k} \\
&= \gamma_{k+1}.
\end{aligned}
$$

因此我们有

$$
\|T^{-1}\|_2 \le n\sum_{i=0}^{n-1}\gamma_i \le \frac{1}{\alpha}\left(1+\frac{\sigma}{\alpha}\right)^{n-1} \tag{3.16}
$$

并利用舒尔分解重新证明鲍尔-菲克定理，可得对于某个特征值μ（属于$A+E$）及其对应的特征向量 v，有

$$
(A+E)v = \mu v,
$$

$$
1 \le \|(A-\mu I)^{-1}\|_2 \|E\|_2,
$$

因此，记 $A = UTU^*$，由方程 3.16 可得，λ 的特征值 A与 μ 的距离满足

$$
1 \le \frac{1}{|\lambda-\mu|}\left(1+\frac{\|A\|_2}{|\lambda-\mu|}\right)^{n-1}\|E\|_2.
$$

同样，将 $A = VGV^{-1}$ 作若尔当分解，并对每个块使用以下展开式进行展开，可得

$$
(|\lambda-\mu|I+E)^{-1} = \frac{1}{|\lambda-\mu|}\sum_{i=0}^{l-1}\left(\frac{1}{|\lambda-\mu|}\right)^{i},
$$

that

$$1 \leq \frac{1}{|\lambda-\mu|}\sum_{i=0}^{l-1}\left(\frac{1}{|\lambda-\mu|}\right)^{i}\mathcal{K}_2(A)\|E\|_2.$$

□

因此，对于一般矩阵，会损失一个$\exp(n)$因子；若使用引理 3.15，还会进一步损失一个$\mathcal{K}(A)$因子；若使用引理 3.15，则会进一步损失一个 $\sigma(A)$因子。即使适当选择 E，这些结果对 ϕ 给出的下界仍极其弱。因此，我们需要不同的思路。

3.7 关于 $\phi(R)$ 的下界，以 $\Delta(R)$为依据

作为前两节讨论的结果，我们希望得到一个比奇异值所给出的界更有意义的界，并且该界不依赖于条件数、指数因子或奇异值，这样与引理 3.8结合后，便能得到对佩龙-弗罗贝尼乌斯定理的完整推广。这个下界将由接下来陈述的一系列引理得出。在此之前，我们先概述一下证明。为便于理解，我们将针对双随机矩阵的情形进行讨论——这类矩阵以均匀分布为不动点；事实证明，它们是一个足够丰富的特殊情形，包含了一般情形的大多数性质。

考虑一个不可约双随机矩阵 A。请注意，迄今为止，我们所有的谱推理都至关重要地依赖于引理3.6、3.1、3.12中某个矩阵的谱分解。然而，对于证明定理1.2 中这类关于 $\phi(A)$ 下界的结果，我们无法使用谱分解，因为 A 不能通过酉矩阵对角化；而尝试模仿 Buser 不等式的证明（第3.1节）也只能得到类似于引理3.12的结果。如上一节所述，文献中的其他定理只能在有限的情形下使用——例如，当 A 可对角化时可以使用鲍尔-菲克定理；然而，将这些定理通过舒尔分解或若尔当分解推广至一般的不可对角化矩阵，或使用微扰理论中的其他工具，得到的下界会带有关于 n的指数因子，和/或依赖于 A 的特征值/特征向量条件数，和/或依赖于 A的最大奇异值，如使用舒尔分解而非若尔当分解的引理3.15 所示。若 A 是双随机矩阵，其奇异值和特征值条件数的意义较小；但在考虑定理1.2所要求的一般矩阵时，它们就会成为主导因素。

解开整个证明的关键思路，是尝试理解 $\phi(A^2)$ 和 $\phi(A^k)$。结果表明，$\phi(A^k) \leq k \cdot \phi(A)$。请注意，通过理解 A^k 的边扩张，这有助于我们给出 A 的边扩张下界。设 $A = J + UTU^*$，其中 $J = \frac{1}{n}\mathbf{1}\cdot\mathbf{1}^T$，且 U是酉矩阵，T是 $A-J$ 的舒尔分解中的上三角矩阵。此外，容易看出 $A^k = J + UT^kU^*$。我们需要找到一个足够小的 k，使得 A^k 的扩张易于界定。为此，我们将尝试确保对于某个 k，有 $\|T^k\| \approx 0$，从而 $\phi(A^k) \approx \phi(J) =$常数。

注意，由于对应于最大特征值 1 的特征子空间已在 J 中被移除，A 的对角线中所有特征值 1 都被替换为 0。设模最大的特征值为 λ_m，并假设 $|\lambda_m| < 1$。若对 $T^{1/1-|\lambda_m|}$ 取幂，则显然，其对角线上的每个元素都约为e^{-1}，且继续取更高次幂时，对角线会呈指数级衰减。因此，$T^{\log n/(1-|\lambda_m|)}$ 的对角线元素都将是多项式倒数级小，此时我们假设它们近似为 0。因此，该矩阵现在的行为将类似于幂零矩阵，由此可得 $\|T^{n\log n/(1-|\lambda_m|)}\| \approx 0$。因此，对于 $k \approx n\cdot\log\frac{n}{1-|\lambda_m|}$，我们有

$$\phi(A) \geq \frac{1}{k}\phi(A^k) \approx \frac{1}{k}\phi(J) \approx \frac{1-|\lambda_m|}{n\cdot\log n},$$

定性地给出了我们所需类型的界。许多技术上的细节会使结果依赖于 A 的最大奇异值以及 A 最大特征值的特征值条件数，但通过仔细分析（并运用足够的技巧），我们可以避免这两个因子，也可以去掉 $\log n$ 这一因子，并将依赖关系从 $|\lambda_m|$ 转换为 $\mathrm{Re}\lambda_2$。

现在我们继续证明定理 3.2 中的下界。

第一个引理指出，ϕ 在以下意义下是次乘法的。

引理 3.17。（ϕ_S 的次乘法性）设 R 和 B 是非负矩阵，它们对于特征值 *1* 具有相同的左右特征向量 u 和 v。那么对于任意割 S，有

$$\phi_S(RB) \leq \phi_S(R) + \phi_S(B).$$

证明。令

$$\gamma_S(R) = \langle \mathbf{1}_{\bar{S}}, D_u R D_v \mathbf{1}_S\rangle + \langle \mathbf{1}_S, D_u R D_v \mathbf{1}_{\bar{S}}\rangle$$

类似地，$\gamma_S(B)$。不失一般性，我们假设 R 和 B 的最大特征值为 1，因为根据引理 3.3，边扩张在缩放下保持不变。

那么，对于每个割 S，我们将证明

$$\gamma_S(RB) \le \gamma_S(R) + \gamma_S(B).$$

固定任意一个割 S。假设 $R = \begin{bmatrix} P & Q \\ H & V \end{bmatrix}$ 和 $B = \begin{bmatrix} X & Y \\ Z & W \end{bmatrix}$ 根据割 S 自然地划分。对于任意向量 u，令 D_u 为以 u 为对角线元素的对角矩阵。由于 $Rv = v, R^T u = u, Bv = v, B^T u = u$，我们有

$$P^T D_{u_S}\mathbf{1} + H^T D_{u_{\bar{S}}}\mathbf{1} = D_{u_S}\mathbf{1}, \tag{3.18}$$

$$ZD_{v_S}\mathbf{1} + WD_{v_{\bar{S}}}\mathbf{1} = D_{v_{\bar{S}}}\mathbf{1}, \tag{3.19}$$

$$XD_{v_S}\mathbf{1} + YD_{v_{\bar{S}}}\mathbf{1} = D_{v_S}\mathbf{1}, \tag{3.20}$$

$$Q^T D_{u_S}\mathbf{1} + V^T D_{u_{\bar{S}}}\mathbf{1} = D_{u_{\bar{S}}}\mathbf{1}, \tag{3.21}$$

其中 u 自然地划分为 u_S 和 $u_{\bar{S}}$，而 v 根据割 S 自然地划分为 v_S 和 $v_{\bar{S}}$。此外，在上述方程及下文中，向量$\mathbf{1}$ 表示维数为 $|S|$ 或 $|\bar{S}|$ 的全 1 向量，具体维数由方程的上下文确定；为简化记号，我们不使用不同的向量。于是，根据 γ_S 的定义，

$$\begin{aligned}\gamma_S(R) &= \langle \mathbf{1}_{\bar{S}}, D_u R D_v \mathbf{1}_S\rangle + \langle \mathbf{1}_S, D_u R D_v \mathbf{1}_{\bar{S}}\rangle \\ &= \langle \mathbf{1}, D_{u_S} Q D_{v_{\bar{S}}}\mathbf{1}\rangle + \langle \mathbf{1}, D_{u_{\bar{S}}} H D_{v_S}\mathbf{1}\rangle\end{aligned}$$

类似地，

$$\gamma_S(B) = \langle \mathbf{1}, D_{u_S} Y D_{v_{\bar{S}}}\mathbf{1}\rangle + \langle \mathbf{1}, D_{u_{\bar{S}}} Z D_{v_S}\mathbf{1}\rangle.$$

矩阵 RB 还具有 u 和 v，它们分别是特征值 1 对应的左、右特征向量，因此，

$$\begin{aligned}\gamma_S(RB) &= \langle \mathbf{1}, D_{u_S} PY D_{v_{\bar{S}}}\mathbf{1}\rangle + \langle \mathbf{1}, D_{u_S} QW D_{v_{\bar{S}}}\mathbf{1}\rangle + \langle \mathbf{1}, D_{u_{\bar{S}}} HX D_{v_S}\mathbf{1}\rangle + \langle \mathbf{1}, D_{u_{\bar{S}}} VZ D_{v_S}\mathbf{1}\rangle \\ &= \langle P^T D_{u_S}\mathbf{1}, YD_{v_{\bar{S}}}\mathbf{1}\rangle + \langle \mathbf{1}, D_{u_S} QW D_{v_{\bar{S}}}\mathbf{1}\rangle + \langle \mathbf{1}, D_{u_{\bar{S}}} HX D_{v_S}\mathbf{1}\rangle + \langle V^T D_{u_{\bar{S}}}\mathbf{1}, ZD_{v_S}\mathbf{1}\rangle \\ &= \langle D_{u_S}\mathbf{1} - H^T D_{u_{\bar{S}}}\mathbf{1}, YD_{v_{\bar{S}}}\mathbf{1}\rangle + \langle \mathbf{1}, D_{u_S} Q(D_{v_{\bar{S}}}\mathbf{1} - ZD_{v_S}\mathbf{1})\rangle \\ &\quad + \langle \mathbf{1}, D_{u_{\bar{S}}} H(D_{v_S}\mathbf{1} - YD_{v_{\bar{S}}}\mathbf{1}\rangle + \langle D_{u_{\bar{S}}}\mathbf{1} - Q^T D_{u_S}\mathbf{1}, ZD_{v_S}\mathbf{1}\rangle \\ &\qquad \text{[from equations 3.18, 3.19, 3.20, 3.21 above]} \\ &= \langle D_{u_S}\mathbf{1}, YD_{v_{\bar{S}}}\mathbf{1}\rangle + \langle \mathbf{1}, D_{u_S} QD_{v_{\bar{S}}}\mathbf{1}\rangle - \langle H^T D_{u_{\bar{S}}}\mathbf{1}, YD_{v_{\bar{S}}}\mathbf{1}\rangle - \langle \mathbf{1}, D_{u_S} QZD_{v_S}\mathbf{1}\rangle \\ &\quad + \langle \mathbf{1}, D_{u_{\bar{S}}} HD_{v_S}\mathbf{1}\rangle + \langle D_{u_{\bar{S}}}\mathbf{1}, ZD_{v_S}\mathbf{1}\rangle - \langle \mathbf{1}, D_{u_{\bar{S}}} HYD_{v_{\bar{S}}}\mathbf{1}\rangle - \langle Q^T D_{u_S}\mathbf{1}, ZD_{v_S}\mathbf{1}\rangle \\ &\le \langle D_{u_S}\mathbf{1}, YD_{v_{\bar{S}}}\mathbf{1}\rangle + \langle \mathbf{1}, D_{u_S} QD_{v_{\bar{S}}}\mathbf{1}\rangle + \langle \mathbf{1}, D_{u_{\bar{S}}} HD_{v_S}\mathbf{1}\rangle + \langle D_{u_{\bar{S}}}\mathbf{1}, ZD_{v_S}\mathbf{1}\rangle \\ &\qquad \text{[since every entry of the matrices is nonnegative]} \\ &= \gamma_S(R) + \gamma_S(B)\end{aligned}$$

如所需。注意，由于 R 和 B 具有相同的左、右特征向量， RB 也如此；又由于 $\langle \mathbf{1}_S, D_u R D_v \mathbf{1}_S \rangle = \langle \mathbf{1}_S, D_u R D_v \mathbf{1}_S \rangle$，我们恰好有

$$\phi_S(R) = \frac{1}{2} \frac{\gamma_S(R)}{\sum_{i \in S} u_i v_i}$$

由此得到该引理。 □

下面的引理可直接由引理 3.17 推论得出，本质上也是整个证明的核心。

引理 3.22。 （ϕ的次乘性）设 $R \in \mathbb{R}^{n \times n}$ 是一个不可约非负矩阵，其对应于佩龙–弗罗贝尼乌斯特征值 1 的左右特征向量为 u 和 v 。则

$$\phi(R^k) \leq k \cdot \phi(R).$$

证明。 注意到 R^k 具有 u 和 v 作为任意 k对应的左右特征向量，利用引理3.17 ，我们归纳得到

$$\phi_S(R^k) \leq \phi_S(R) + \phi_S(R^{k-1}) \leq \phi_S(R) + (k-1) \cdot \phi_S(R) = k \cdot \phi_S(R),$$

令 S 为使 $\phi(R)$达到最小值的集合，便可得到

$$\phi(R^k) \leq k \cdot \phi(R).$$

□

对于 对称双随机矩阵 R，引理3.22 可由 Blakley and Roy 的一个定理推出 [BR65] （该情形不属于将该结果扩展到非对称情形的框架 [Pat12]）。引理3.22可通过取 ϕ（R）的幂来给出 R的下界。这很有用，因为我们可以取足够高的幂，使矩阵变得足够简单，从而容易计算其边扩张。

接下来的两个引理可通过技术性计算得出。

引理 3.23。 （幂的有界范数）设 $T \in \mathbb{C}^{n \times n}$ 是一个上三角矩阵，具有 $\|T\|_2 = \sigma$ ，且对每个 i， $|T_{i,\ i}| \leq \alpha < 1$。那么

$$\|T^k\|_2 \leq n \cdot \sigma^n \cdot \binom{k+n}{n} \cdot \alpha^{k-n}.$$

证明。令 $g_r(k)$ 表示 T^k 中与对角线距离为 r 的矩阵元素的绝对值的最大值，其中对角线上的元素与对角线的距离为 0，非对角线上的元素与对角线的距离为 1，依此类推。更正式地说，

$$g_r(k) = \max_i |T^k(i, i+r)|.$$

我们将通过归纳法证明，对于 $\alpha \le 1$，且 $r \ge 1$，

$$g_r(k) \le \binom{k+r}{r} \cdot \alpha^{k-r} \cdot \sigma^r, \tag{3.24}$$

其中 $\sigma = \|T\|_2$。首先注意，对于 $r = 0$，由于 T 是上三角的，T^k 的对角线为 α^k，因此该假设对 $r = 0$ 以及所有 $k \ge 1$ 均成立。此外，对于 $k = 1$，若 $r = 0$，则 $g_0(1) \le \alpha$；若 $r \ge 1$，则 $g_r(1) \le \|T\|_2 \le \sigma$，此时归纳假设也成立，因为 $r \ge k$ 且 $\alpha^{k-r} \ge 1$。对于归纳步骤，假设对所有 $r \ge 1$ 及所有 $j \le k-1$，都有 $g_r(j) \le \binom{j+r}{r} \cdot \alpha^{j-r} \cdot \sigma^r$。下面给出 $g_r(k)$ 的计算。

由于 $|a+b| \le |a| + |b|$，

$$\begin{aligned} g_r(k) &\le \sum_{i=0}^{r} g_{r-i}(1) \cdot g_i(k-1) \\ &= g_0(1) \cdot g_r(k-1) + \sum_{i=0}^{r-1} g_{r-i}(1) \cdot g_i(k-1). \end{aligned}$$

第一项可写为：

$$\begin{aligned} g_0(1) \cdot g_r(k-1) &= \alpha \cdot \binom{k-1+r}{r} \cdot \alpha^{k-1-r} \cdot \sigma^r \\ &\quad \text{[利用 } g_0(1) \le \alpha \text{ 以及针对第二项的归纳假设]} \\ &\le \alpha^{k-r} \cdot \sigma^r \cdot \binom{k+r}{r} \cdot \frac{\binom{k-1+r}{r}}{\binom{k+r}{r}} \\ &\le \frac{k}{k+r} \cdot \binom{k+r}{r} \cdot \alpha^{k-r} \cdot \sigma^r \end{aligned} \tag{3.25}$$

而第二项为

$$
\begin{aligned}
g_r(k) &\le \sum_{i=0}^{r-1} g_{r-i}(1)\cdot g_i(k-1) \\
&\le \sigma\cdot\sum_{i=0}^{r-1}\binom{k-1+i}{i}\cdot\alpha^{k-1-i}\cdot\sigma^i \\
&\quad \text{[使用 } g_{r-i}(1)\le\sigma \text{ 以及关于第二项的归纳假设]} \\
&\le \sigma^r\cdot\alpha^{k-r}\cdot\binom{k+r}{r}\cdot\sum_{i=0}^{r-1}\cdot\frac{\binom{k-1+i}{i}}{\binom{k+r}{r}}\cdot\alpha^{r-1-i} \\
&\quad \text{[using } \sigma^i\le\sigma^{r-1} \text{ since } \sigma\ge 1 \text{ and } i\le r-1] \\
&\le \sigma^r\cdot\alpha^{k-r}\cdot\binom{k+r}{r}\cdot\sum_{i=0}^{r-1}\cdot\frac{\binom{k-1+i}{i}}{\binom{k+r}{r}} \\
&\quad \text{[using } \alpha\le 1 \text{ and } i\le r-1]
\end{aligned}
$$

现在我们将证明求和中的量至多为 $\frac{r}{k+r}$。根据归纳假设，对于 ~~$r=1$~~，该陈述成立，并假设对于任意其他 r，

$$\sum_{i=0}^{r-1}\frac{\binom{k-1+i}{i}}{\binom{k+r}{r}}\le\frac{r}{k+r}.$$

于是我们有

$$
\begin{aligned}
\sum_{i=0}^{r}\frac{\binom{k-1+i}{i}}{\binom{k+r+1}{r+1}} &= \frac{\binom{k+r}{r}}{\binom{k+r+1}{r+1}}\cdot\sum_{i=0}^{r-1}\frac{\binom{k-1+i}{i}}{\binom{k+r}{r}}+\frac{\binom{k-1+r}{k-1}}{\binom{k+r+1}{k}} \\
&\le \frac{r+1}{k+r+1}\cdot\frac{r}{k+r}+\frac{(r+1)\cdot k}{(k+r+1)\cdot(k+r)} \\
&= \frac{r+1}{k+r+1}
\end{aligned}
$$

因此我们得到第二项至多为 $\sigma^r\cdot\alpha^{k-r}\cdot\binom{k+r}{r}\cdot\frac{r}{k+r}$，将其与第一项（方程 3.25）结合，即完成了归纳假设。

注意到算子范数至多为弗罗贝尼乌斯范数，并且由于 $g_r(k)$ 关于 r 是递增的，且 r 的最大值为 n，利用方程 3.24，我们得到

$$
\begin{aligned}
\|T^k\|_2 &\le \sqrt{\sum_{i,j}|T^k(i,j)|^2} \\
&\le n\cdot\sigma^n\cdot\alpha^{k-n}\cdot\binom{k+n}{n}
\end{aligned}
$$

如所要求。 □

利用引理 3.23，我们可以证明以下适用于算子范数至多为 1 的上三角矩阵这一特殊情形的引理。

引理 3.26. 设 $T \in \mathbb{C}^{n\times n}$ 是一个上三角矩阵，满足 $\|T\|_2 \leq 1$ 和$|T_{i,i}| \leq \alpha < 1$对每个 i都成立。假设对于每个实数 k， T^k都有定义。那么 $\|T^k\| \leq \epsilon$对于

$$k \geq \frac{4n + 2\ln(\frac{n}{\epsilon})}{\ln\left(\frac{1}{\alpha}\right)}.$$

证明。 设 $X = T^{k_1}$， 其中 $k_1 = \frac{c_1}{\ln(\frac{1}{\alpha})}$对 $\alpha < 1$成立。则 $\|X\|_2 \leq \|T\|_2^{k_1} \leq 1$, 且对每个 i,

$$|X_{i,i}| \leq |T_{i,i}|^{k_1} \leq |\lambda_m|^{c_1/\ln(\frac{1}{\alpha})} = e^{-c_1}.$$

利用引理 3.23 对 X 取 $\sigma = 1$ 和 $\beta = e^{-c_1}$， 可得对 $k_2 = c_2 \cdot n$ ，

$$\begin{aligned}
\|X^{k_2}\| &\leq n \cdot \binom{k_2 + n}{n} \cdot e^{-c_1(k_2-n)} \\
&\leq n \cdot e^n \cdot (c_2+1)^n \cdot e^{-c_1(c_2-1)n} \\
&\quad [\text{using } \binom{a}{b} \leq \left(\frac{ea}{b}\right)^b] \\
&= \exp\left(n \cdot \left(\frac{\ln n}{n} + \ln(c_2+1) + 1 + c_1 - c_1c_2\right)\right)
\end{aligned}$$

并且要使该量小于 ϵ，我们要求

$$\begin{aligned}
&\exp\left(n \cdot \left(\frac{\ln n}{n} + \ln(c_2+1) + 1 + c_1 - c_1c_2\right)\right) \leq \epsilon \\
\Leftrightarrow &\left(\frac{1}{\epsilon}\right)^{\frac{1}{n}} \leq \exp\left(-1 \cdot \left(\frac{\ln n}{n} + \ln(c_2+1) + 1 + c_1 - c_1c_2\right)\right) \\
\Leftrightarrow &\frac{1}{n}\ln\frac{n}{\epsilon} + 1 + c_1 + \ln(c_2+1) \leq c_1c_2 \qquad (3.27)
\end{aligned}$$

并令

$$c_1 = 1 + \frac{1}{2.51} \cdot \frac{1}{n}\ln\left(\frac{n}{\epsilon}\right)$$

且 $c_2 = 3$。51 它始终满足不等式 3.27。因此，对于

$$\begin{aligned}
k &= k_1 \cdot k_2 \\
&= \frac{c_1 \cdot c_2 \cdot n}{\ln(\frac{1}{\alpha})} \\
&= \frac{3.51 \cdot n + 1.385 \cdot \ln\left(\frac{n}{\epsilon}\right)}{\ln(\frac{1}{\alpha})}
\end{aligned}$$

我们得到,

$$\|T^k\|_2 \le \|X^{k_2}\| \le \epsilon$$

如要求。 □

根据引理 3.22 和 3.26，我们可以用 $\ln(\frac{1}{|\lambda_m|})$ 给 $\phi(R)$ 一个下界（其中 λ_m 是模最大的非平凡特征值）。我们的目标是用 $\phi(R^k)$ 给 $\phi(R)$ 一个下界，但由于 R^k 的范数会随着幂运算增大，我们不能直接使用这些引理，因为我们不希望最终的下界依赖于 $\sigma(R)$。为了解决这一问题，我们将 R 变换为 A，使得 $\phi(R) = \phi(A)$， R 和 A 的特征值相同，但无论 R 的范数如何， $\sigma(A) = \|A\|_2 = 1$，这使用了引理 3.3。

根据引理 3.3，我们使用 $\phi(A^k)$关于 $1 - |\lambda_m(A)|$的结果，给出 $\phi(A)$的下界，以获得 R 的相应界。

引理 3.28。 设 R 是一个不可约非负矩阵，其佩龙–弗罗贝尼乌斯特征值为 *1*，对应的正左、右特征向量分别为 u 和 v，并已归一化，使得 $\langle u, v\rangle = 1$。设 λ_m 是 R 中模最大的非平凡特征值，并且对每个实数 k， R^k 都定义良好。那么

$$\frac{1}{15} \cdot \frac{\ln\left(\frac{1}{|\lambda_m|}\right)}{n} \le \phi(R).$$

证明。 由给定的 R 及其正左、右特征向量 u 和 v 对应于特征值 1（如上所述），令 $A = D_u^{\frac{1}{2}} D_v^{-2} R D_u^{-2} D_v^{\frac{1}{2}}$以及 $w = D_u^{\frac{1}{2}} D_v^{\frac{1}{2}} \mathbf{1}$如引理 3.3中所示。注意， w 为正，并且

$$\langle w, w\rangle = \langle D_u^{\frac{1}{2}} D_v^{\frac{1}{2}} \mathbf{1}, D_u^{\frac{1}{2}} D_v^{\frac{1}{2}} \mathbf{1}\rangle = \langle u, v\rangle = 1.$$

此外， $Aw = w$ 和 $A^T w = w$，并且由于 $A^k = D_u^{\frac{1}{2}} D_v^{-2} R^k D_u^{-2} D_v^{\frac{1}{2}}$，它对每个实数 k 都定义良好。

设 $A = w\cdot w^T + B$。由于 $(w\cdot w^T)^2 = w\cdot w^T$，且 $Bw\cdot w^T = w\cdot w^T B = 0$，因此得到

$$A^k = w \cdot w^T + B^k.$$

令 $B = UTU^*$ 为 B的舒尔分解，其中 T 的对角线包含 A除随机特征值以外的所有特征值，该随机特征值被替换为 0，因为对于 A的特征值 1， w 同时是左特征向量和右特征向量，而在 $w\cdot w^T$中移除了该空间。此外， T 的最大对角线元素至多为 $|\lambda_m|$，其中 λ_m 是非平凡

特征值 A （或 R）中模最大的一个。注意， $|\lambda_m| < 1$ 因为 A（以及 R）是不可约的。由于 $w \cdot w^T B = Bw \cdot w^T = 0$ 且 $\|A\|_2 \leq 1$ 根据引理3.3，我们有 $\|B\|_2 \leq 1$。

因此，利用引理3.26（实际上，利用引理 3.26证明中的最后几行），对于

$$k \geq \frac{3.51 \cdot n + 1.385 \cdot \ln\left(\frac{n}{\epsilon}\right)}{\ln\left(\frac{1}{|\lambda_m|}\right)},$$

我们得到

$$\|B^k\|_2 = \|T^k\|_2 \leq \epsilon,$$

并且，对于 e_i 为在位置 i 处取值为 1、其余位置均为 0 的向量时，利用柯西-施瓦茨不等式可得

$$|B^k(i,j)| = \left|\langle e_i, B^k e_j \rangle\right| \leq \|e_i\|_2 \|B\|_2 \|e_j\|_2 \leq \epsilon.$$

此外，注意到 $B^k w = 0$ 和 $w^T B^k = 0$。这意味着，对于任意 i 和 $S \subseteq [n]$,

$$\sum_{j \in S} B^k(i,j) w_j = -\sum_{j \in \overline{S}} B^k(i,j) w_j$$

因此，对于任意满足 $\sum_{i \in S} w_i^2 \leq \frac{1}{2}$ 的集合 S，我们得到

$$\begin{aligned}
\phi_S(A^k) &= \frac{\langle \mathbf{1}_S, D_w A^k D_w \mathbf{1}_{\overline{S}} \rangle}{\langle \mathbf{1}_S, D_w D_w \mathbf{1} \rangle} \\
&= \frac{\sum_{i \in S, j \in \overline{S}} A^k(i,j) \cdot w_i \cdot w_j}{\sum_{i \in S} w_i^2} \\
&= \frac{\sum_{i \in S, j \in \overline{S}} \left(w \cdot w^T(i,j) + B^k(i,j)\right) \cdot w_i \cdot w_j}{\sum_{i \in S} w_i^2} \\
&= \frac{\sum_{i \in S} w_i^2 \sum_{j \in \overline{S}} w_j^2 - \sum_{i \in S, j \in S} B^k(i,j) \cdot w_i \cdot w_j}{\sum_{i \in S} w_i^2} \\
&\geq \sum_{j \in \overline{S}} w_j^2 - \epsilon \frac{(\sum w_i^2)^2}{\sum_{i \in S} w_i^2} \\
&\geq \frac{1}{2}(1 - \epsilon)
\end{aligned}$$

因为 $\sum_{i \in S} w_i^2 \leq \frac{1}{2}$。 注意，这对每个集合 S 都成立。因此，我们得到，对于

$$k \geq \frac{3.51n + 1.385 \cdot \ln\left(\frac{n}{\epsilon}\right)}{\ln\left(\frac{1}{|\lambda_m|}\right)}$$

边扩张

$$\phi(A^k) \geq \frac{1}{2}(1-\epsilon),$$

因此，利用引理3.22，我们得到

$$\phi(A) \geq \frac{1}{k} \cdot \phi(A^k) \geq \frac{1-\epsilon}{2} \cdot \frac{\ln\left(\frac{1}{|\lambda_m|}\right)}{3.51 \cdot n + 1.385 \cdot \ln(n/\epsilon)}$$

并令 $\epsilon = \frac{1}{2e}$，再利用 $\ln(e \cdot x) \leq x$，我们得到

$$\phi(A) \geq \frac{1}{15} \cdot \frac{\ln\left(\frac{1}{|\lambda_m|}\right)}{n},$$

根据引理 R，它会传递到 3.3。 □

我们几乎已经得到了所需的证明，但还需要处理两个问题。我们需要得到关于 $\mathrm{Re}\lambda_2$ 而非 $|\lambda_m|$ 的依赖关系，并且需要确保可以对应用这些引理所涉及的底层矩阵取实数次幂。为了解决这两个问题，接下来我们将重点考察 R 的指数版本，即 $E_R = \exp(R - I)$。我们注意到这个矩阵具有以下性质。

引理 3.29。 令 E_R 为 R 的指数版本，即 $E_R = \exp(R - I)$。则对于 E_R，以下结论成立。

1. $E_A = D_u^2 D_v^2 E_R D_u^2 D_v^2$ 对于 $A = D_u^2 D_v^2 R D_u^2 D_v^2$*2.* E_R 的最大特征值为 *1*，E_R 和 R 的左右特征向量相同，且 E_R 不可约。*3.* $(E_R)^k$ 对所有实数 k 都定义良好。*4.* $\phi(E_R^t) \leq t \cdot \phi(R)$

证明。 前三个性质显然成立，且由定义可得。对于最后一个性质，我们证明 A，而 R 的情形可由引理3.3推出。对于

任意集合 S，我们可以写成

$$\begin{aligned}
\phi_S(E_A^t) &= \phi_S(\exp(t(A-I))) \\
&= e^{-t}\phi_S(\sum_{i=0}\frac{1}{i!}t^iA^i) \\
&= e^{-t}\left(\phi_S(I)+\sum_{i=1}^{\infty}\frac{t^i\cdot\phi_S(A^i)}{i!}\right) \\
&\le e^{-t}\left(0+\sum_{i=1}^{\infty}i\frac{t^i\cdot\phi_S(A)}{i!}\right) \\
&\quad \text{[using Lemma 3.22]} \\
&= e^{-1}\left(t\cdot\phi_S(A)\sum_{i=1}^{\infty}\frac{t^{i-1}}{(i-1)!}\right) \\
&= t\cdot\phi_S(A)
\end{aligned}$$

因此

$$\phi(E_A) \le t\cdot\phi(A)$$

如所需，且由引理 R 3.3可将其推广至。 □

根据引理 3.28，我们使用 R 的指数版本，得到定理 3.2中边扩张的下界，现将其重述如下。

引理 3.30。 设 R 是一个不可约非负矩阵，其佩龙–弗罗贝尼乌斯特征值为 1，具有正的（左、右）特征向量 u 和 v，且已归一化，使得 $\langle u,v\rangle = 1$。则

$$\frac{1}{15}\cdot\frac{1-Re\lambda_2(R)}{n} \le \phi(R).$$

证明。 我们先证明 A 的情形，根据引理 R 可推广至3.3。注意，对于 $\lambda = a + i\cdot b$ 的每个特征值 A，其在 E_A 中的对应特征值为

$$\begin{aligned}
\exp(\lambda-1) &= \exp(a-1)\exp(i\cdot b) \\
|\exp(\lambda-1)| &= \exp(a-1).
\end{aligned}$$

因此，要使 E_A 的特征值的模 $|\exp(\lambda-1)|$ 最大，我们需要使 $\exp(a-1)$ 最大。由于 $a<1$ 在 A 下不可约，当 a 最接近 1 时，值 $|\exp(\lambda-1)|$ 最大。换言之，令 $\lambda_m(E_A)$ 为 E_A 中模最大的特征值，则有

$$|\lambda_m(E_A)| = \exp(\mathrm{Re}\lambda_2(A)-1).$$

应用引理 3.28 于 E_A，可得

$$\phi(E_A) \geq \frac{1}{15} \cdot \frac{\ln\left(\frac{1}{|\lambda_m(E_A)|}\right)}{n} = \frac{1}{15} \cdot \frac{1 - \mathrm{Re}\lambda_2(A)}{n},$$

并由引理3.29，最终得到

$$\phi(A) \geq \frac{1}{15} \cdot \frac{1 - \mathrm{Re}\lambda_2(A)}{n}$$

如所需，并通过引理 R 将其推广至 3.3。 □

这就完成了我们的第一个主定理的证明，即关于不可约非负矩阵的 ϕ 下界。将谱隙和边扩张的定义推广后，便可简单地将其推广到所有矩阵。

3.8 定理 3.2 向所有非负矩阵的推广

下面这个简单的引理可由下极限的定义推出。

引理 3.31。设 $p: E \to \mathbb{R}$ 和 $q: E \to \mathbb{R}$ 是两个函数，使得对于所有 $H \in E$, $p(H) \leq q(H)$ 均成立。那么对于所有 R,

$$\lim_{\epsilon \to 0} \inf_{H \in R_\epsilon} p(H) \leq \lim_{\epsilon \to 0} \inf_{H \in R_\epsilon} q(H).$$

证明。注意，若 $R \in E$，则显然成立。假设 $R \notin E$。固定 ϵ。令$a_\epsilon = \inf_{H \in R_\epsilon} p(H)$ 和 $b_\epsilon = \inf_{H \in R_\epsilon} q(H)$。首先，我们断言 $a_\epsilon \leq b_\epsilon$。为导出矛盾，假设

$$a_\epsilon > b_\epsilon. \tag{3.32}$$

令 $\delta_\epsilon = (a_\epsilon - b_\epsilon)/2 > 0$，则存在 $T \in E$，使得 $q(T) \leq b_\epsilon + \delta_\epsilon$(否则 $\inf_{H \in R_\epsilon} q(H) \geq b_\epsilon + \delta_\epsilon > b_\epsilon$)。此外，

$$a_\epsilon \leq p(T) \leq q(T) \leq b_\epsilon + \delta_\epsilon$$

或者 $a_\epsilon \leq b_\epsilon$，这与 3.32 矛盾。因此，$a_\epsilon \leq b_\epsilon$。

现在令 $c_\epsilon = a_\epsilon - b_\epsilon$，并且我们知道，对所有 $\epsilon > 0$，$c_\epsilon \leq 0$。由此可得

$$\lim_{\epsilon \to 0} c_\epsilon \leq 0 \tag{3.33}$$

再次反证，假设相反的情况成立，即 $\lim_{\epsilon \to 0} c_\epsilon > 0$ 且$\delta = \frac{1}{2} \cdot \lim_{\epsilon \to 0} c_\epsilon > 0$。根据极限的定义，存在 $\epsilon > 0$ 使得 $c_\epsilon > \delta$，或者 $c_\epsilon > 0$，这与 3.33矛盾。因此引理得证。 □

断言 3.34. 主定理 *3.2*对所有非负矩阵 R均成立。

证明。 由引理 3.31在适当设定 p 和 q 后可得。 □

至此，我们完成了主定理的证明 1.2。

第 4 章

非可逆链的构造

> 我们现在已经认识到，人们通常对自然所作的一切解释，都不过是想象的方式，并未揭示任何事物的真正本性，只表明想象力的构造。我不把美或丑、秩序或混乱归于自然。只有相对于我们的想象，事物才能被称为美或丑、有序或混乱。
>
> ~ 巴鲁赫·斯宾诺莎，《伦理学》

本节的出发点是理解定理 1.2 的最优性。定理1.2 与 Cheeger-Buser 不等式（定理3.1）之间的主要区别在于，在 n 非可逆情形下会损失一个因子；我们想弄清楚这种损失是否确实不可避免，还是说它只是我们证明技巧局限性的遗留问题。我们还将寻求具有这些性质的双随机 矩阵的构造，以检验我们的定理是否最优，即使在非负矩阵的主特征向量均匀的情形下也是如此；在这种情况下，人们或许会预期损失是 $\mathrm{polylog}(n)$，而不是定理 n 中的损失 1.2。我们说明，将在第4.5、 4.6、 4.7 和 4.9 节详细解释我们如何得到第4.8 和 4.10 节中的构造，使其看起来不至于神秘。简洁固然是智慧的标志，但我们认为全面才是理解的标志；因此，我们大体上会遵循后者，并详细阐述得出这些构造的思路，这或许能启发人们寻找其他类似的构造。

在着手构造之前，我们需要理解图的非扩张。

4.1 非扩张

边扩张 ϕ 如2.4中所定义，在过去70年里，组合学和谱理论领域对其进行了广泛研究，尽管研究对象大多是可逆矩阵，以及边数较少（或矩阵中非零项较少）且扩张率为常数的无向图（或对称矩阵）——也称为（组合）扩展图。这些对象具有许多非凡、近乎神奇的性质（见 [HLW06]），是重要的组合对象。以它们为基础，可以构造许多其他达到最优（忽略低阶项）的伪随机对象，例如纠错码和伪随机生成器等；此外，它们还是数学中大量构造的基石。

我们现在要探讨的是与图的扩张方向相反的性质——图的非扩张。对于无向图而言，这并无特别之处，因为对称矩阵的非扩张性仅意味着谱隙较小，这是由 Cheeger-Buser 不等式3.1得出的，而且构造这类图也很简单。然而，对于有向图（或不可逆矩阵），非扩张性则并非平凡性质，因为原则上，不可逆矩阵可能具有常数谱隙，但其边扩张会随着矩阵规模增大而减小。

本节介绍的所有构造都是非扩张不可逆矩阵的构造，尽管它们与引人注目且备受追求的常数扩张性质相去甚远，但它们以一种不同且独特的方式展现出真正的美，并且值得作为独立对象加以研究。本节的目标有两个方面——分析并阐述一些已知的有趣构造，并介绍一些新构造；就定理 *1.2*中的下界而言，这些构造比已知构造具有指数级的改进，同时也改进了下文定义的图的非扩张性。事实上，即使是次优的构造（就定理1.2中的下界而言）也将具有许多有趣且美观的性质。为使论述系统化，我们定义如下数量。

定义 4.1。 *(矩阵的非扩张)* 令 E_n 为所有 $n \times n$ 双随机矩阵的集合。对于任意 $A \in E_n$，定义 A 的非扩张为

$$\Gamma(A) = \frac{\phi(A)}{1 - \mathrm{Re}\lambda_2(A)}$$

并覆盖记号，定义

$$\Gamma(n) = \inf_{A \in E_n} \Gamma(A).$$

◊

引理 4.2. 对于非可逆矩阵，$\frac{1}{15n} \leq \Gamma(n) \leq \frac{1}{2}$；对于可逆矩阵，$\Gamma^{rev}(n) = \frac{1}{2}$

证明。 下界由定理 1.2和 Cheeger-Buser 不等式3.1得出，而上界则通过取一个具有 n 个顶点的超立方体 E_n得出，详见第3.3节。

□

4.2 超越1/2

首先自然会想到，寻找一个双随机 A，使得 $\Gamma(A) < \frac{1}{2}$，因为我们知道对于任意对称 A，由4.2可得 $\Gamma(A) \geq \frac{1}{2}$。事实上，用数值方法找到这样的矩阵并不难。一种直接的做法是从 J开始，选取一些随机的(i, j)和(k, l)以及随机的 δ，令$A_{i,j} \pm \delta, A_{k,l} \pm \delta, A_{i,l} \mp \delta$，以及$A_{k,j} \mp \delta$（其中 δ 的取值范围经过选择，以确保 A 的矩阵元素不会变成负数）。通过这个简单的算法，可以找到许多超过 1/2 的例子；观察这些矩阵后，我们构造出了下面这个例子。

引理 4.3.（超越 1/2）设矩阵 A 如下（其中· 表示零）：

$$A = \begin{bmatrix} \cdot & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \cdot & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \cdot & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 1 & \cdot & \cdot & \cdot \end{bmatrix}.$$

那么 $\Gamma(A) = \frac{1}{3} < \frac{1}{2}$，优于引理 4.2。

证明。 注意到 A 是双随机的，并且对于 $S = \{2, 3\}$，$\phi(A) = \phi_S(A) = \frac{1}{3}$，且 A 的所有非平凡特征值均为 0。因此，引理得证。 □

注记 4.4。（关于副对角线的对称性）引理 A 中的矩阵4.3表示一条有趣的链，其邻接矩阵关于 副对角线对称。尽管这一性质看起来与边扩张或特征值无关，但它将在我们的最后一个例子中再次出现，颇为奇妙。◊

4.3 仿射线性构造

文献中已有许多仿射线性构造，其中一种构造是 Maria Klawe 的一项成果 [Kla84]。事实上，Klawe 撰写论文的目的正是展示某些 d-正则图 [Vaz17] 在 20 世纪 80 年代，Umesh Vazirani 观察到，可以用一种自然的方式为边定向，从而构造出具有相似扩展性质、入度和出度均为 $d/2$ 的有向图；我们也是从他那里了解到这一构造的。此类构造有一整个系列，参数各不相同，但从最小化 $\Gamma(n)$的角度来看，它们对我们而言都等价。下面给出一种具体而简洁的构造。

构造4.5. *(Klawe-Vazirani矩阵 [Kla84; Vaz17])* 设 $n > 2$ 为素数，并构造一个具有 n 个顶点、入度和出度均为 2 的图，方法是将每个顶点 $v \in \mathbb{Z}/n$ 连接到两个顶点 $1 + v$ 和 $2v$，每条边的权重均为 1/2。 ◊

例 4.6。下方展示了 $n = 7$ 的一个示例，其中对向量进行右乘（左乘时取矩阵的转置）：

$$A_{KV} = \begin{bmatrix} \frac{1}{2} & \cdot & \cdot & \cdot & \cdot & \cdot & \frac{1}{2} \\ \frac{1}{2} & \cdot & \cdot & \cdot & \frac{1}{2} & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \frac{1}{2} & \cdot & \cdot & \frac{1}{2} & \cdot \\ \cdot & \cdot & \frac{1}{2} & \frac{1}{2} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \frac{1}{2} & \cdot & \frac{1}{2} \\ \cdot & \cdot & \cdot & \frac{1}{2} & \cdot & \frac{1}{2} & \cdot \end{bmatrix}.$$

◊

我们注意到这些图具有以下性质。

引理 4.7. 构造 A_{KV} *4.5* 中的矩阵具有以下性质：

1. 这些矩阵有一个特征值 *1*、一个特征值 *0*，以及 $n - 2$ 个λ 特征值 $\lambda = \frac{1}{2}\exp\left(\frac{2\pi I}{n-1} \cdot k\right)$，其中 $k = 1、\ldots、n - 2$。
2. *[Kla84]* $\phi(A) \leq c \cdot \left(\frac{\log\log n}{n}\right)^{1/5}\log$ 其中 c 是一个常数。
3. $\Gamma(n) \in O\left(\frac{\log_{\text{对数}} n}{n}\right)^{1/5}$ 对数 改进了引理 *4.3*。

证明。（1）设矩阵如下： $A = \frac{1}{2}(A' + A'')$，其中 A' 是由 n 个顶点构成的有向环，而 A'' 表示循环（因为 $n > 2$ 是素数），该循环从顶点 i 出发，到达 $(2 \cdot i) \bmod n$。考虑与对顶点为 $\{0, ..., n-1\}$ 的向量进行右乘相对应的矩阵，我们有 $A'_{i,i-1} = 1$，其余情况下为 $A'_{i,j} = 0$；若 $i = 2 \cdot j$，则有 $A''_{i,j} = 1$，否则为 $A''_{i,j} = 0$。令 U 表示在域 $\mathbb{Z}_n$ 上的傅里叶变换，其中 $U_{i,j} = \frac{1}{\sqrt{n}}\omega^{i \cdot j}$，$\omega = \exp(2\pi I/n)$ 是 n 次单位根，且 $I = \sqrt{-1}$。令 $B = U^* A U$, $B' = U^* A' U$，以及 $B'' = U^* A'' U$。注意，B' 是一个对角矩阵，其 $B'_{i,i} = w^i$，其余位置为 0。类似地，

$$B''_{i,j} = \langle U_i, A'' U_j \rangle = \frac{1}{n} \sum_{k=0}^{n-1} \omega^{-i \cdot k} \omega^{j \cdot (k/2)} = \frac{1}{n} \sum_{k=0}^{n-1} \omega^{(j/2 - i) \cdot k}$$

因此 $B''_{i,\,2 \cdot i} = 1$ 其他情况下为 0。注意 $B_{0,\,0} = 1$ 以及 $B_{0,\,j} = 0$ 和 $B_{i,\,0} = 0$ 对所有 i, j都成立；这就是对应于特征值 1 的平凡块，而 B 的其余特征值都在 $n - 1 \times n - 1$ 块中。对于任意这样的特征值 λ 及其对应的 B的特征向量 v，当 $1 \le i \le n - 1$ 时，我们有

$$\frac{1}{2}(\omega^i v_i + v_{2 \cdot i}) = \lambda v_i$$

or

$$(2\lambda - \omega^i) v_i = v_{2 \cdot i}.$$

注意 $(2\lambda - \omega^i) \neq 0$，否则 v 将是全零向量。因此，我们得到方程

$$\prod_{i=1}^{n-1} (2\lambda - \omega^i) = 1, \tag{4.8}$$

其根将是 B 和 A的非平凡特征值。注意，$\lambda = 0$ 是一个根，因为 $\prod_{i=1}^{n-1} \omega^i = 1$，并且 $(-1)^{n-1} = 1$，因为 n 是素数。我们声称其余的$n - 2$ 个根都满足 $|\lambda| = \frac{1}{2}$，其中 $\lambda \neq \frac{1}{2}$，即 $\lambda = \frac{1}{2}\exp(2\pi k \frown I/(n-1))$, −其中$k \in [\pi - 1]$。为此，假设 $2\lambda \neq 1$，并将方程 4.8 乘以并除以 $(2\lambda - 1)$，我们得到

$$1 = \frac{\prod_{i=0}^{n-1}(2\lambda - \omega^i)}{2\lambda - 1} = \frac{(2\lambda)^n - 1}{2\lambda - 1}$$

or

$$2\lambda = (2\lambda)^n$$

证明该断言。

(2) 令 $s(A) = \frac{A + A^T}{2}$。文献 [Kla84] [定理 2。1] 证明，顶点扩张 $\mu(s(A))$ 满足以下界：

$$\mu(s(A)) \leq c_1 \cdot \left(\frac{\log\log n}{\log n}\right)^{1/5}.$$

注意到，由于 $s(A)$ 的度为 4，我们得到

$$\phi(A) = \phi(s(A)) \leq 4 \cdot c_1 \left(\frac{\log\log n}{\log n}\right)^{1/5}$$

这就证明了该结论。

(3) 由于 $\mathrm{Re}\lambda_2 \leq \frac{1}{2}$，由 (1) 和 (2) 可得 $\Gamma(n)$ 的所述上界。 □

事实上，已知有许多不同的仿射线性构造。回顾一下，Alon 和 Boppana [Alo86; Nil91] 证明，对于任意一个无限族的 d-正则无向图，其邻接矩阵经归一化后成为双随机矩阵，并且特征值为 $1 = \lambda_1 \geq \lambda_2 \geq \ldots \geq -1$，具有 $\lambda_2 \geq \frac{2\sqrt{d-1}}{d} - o(1)$。Feng 和 Li [Li92] 证明，无向性对于这一上界至关重要：他们给出了一种循环有向 r-部 ($r \geq 2$) d-正则有向图的构造（对于 $k > d$，其顶点数为 $n = kr$，且 $\gcd(k, d) = 1$），其归一化邻接矩阵除 r “平凡”特征值外，所有特征值的范数均为 $\leq 1/d$。这一构造具有仿射线性性质，与 Klawe-Vazirani 中的矩阵颇为相似；据我们所知，它并未给出比那些结果更强的 Γ 上界。

4.4 de Bruijn 构造

最优美的构造之一是 de Bruijn 图，或者（隐含地）de Bruijn 序列。考虑以下问题：设 s 为一个比特串，使得每个 $x \in \{0, 1\}^3$ 都恰好作为连续子串出现一次。此类比特串最早见于 Pingala 的梵语韵律学著作 [Bro69], , 其中一个这样的字符串 s 是 s =yam¯at¯ar¯ajabh¯anasalag¯am；其中每种三音节模式都从其名称对应的位置开始出现：’yam¯at¯a’具有短–长–长的模式，’m¯at¯ar¯a’具有长–长–长的模式，以此类推，直到’salag¯am’，其模式为短–短–长。一般而言，de Bruijn 序列 dB(n, k) 是一个字符串，其中包含大小为 k 的字母表中的每个 n 字母序列，且每个序列恰好作为连续子序列出现一次；而韵律学中包含一个 dB(3, 2) 字符串。de Bruijn [Bru46]和 I J Good [Goo46]分别独立描述了这些序列；此前 Camille Flye Sainte-Marie 也针对大小为 2 [Bru75]的字母表描述过它们。这些序列也被 Karl Popper [Pop05] 称为“最短类随机序列”。

构造 4.9。 *(德布鲁因矩阵 [Bru46])* 在顶点$\{0,1\}^k$上定义 de Bruijn 图，并令每个顶点 $|a\rangle = |a_1 a_2 \cdots a_k\rangle$, 都有两条出边，分别通向顶点$|a_2 \cdots a_k 0\rangle$和 $|a_2 \cdots a_k 1\rangle$，每条边的权重均为 1/2 换句话说，从一个 k比特顶点 v出发，均匀地转移到以下两个顶点之一：将比特向左移动 v 一位，并均匀地追加 0 或 1。 ◊

例 4.10。 对于 $k = 3$ ($n = 8$)，邻接矩阵（用于向量右乘）如下所示：

$$A_{dB} = \begin{bmatrix} \frac{1}{2} & \cdot & \cdot & \cdot & \frac{1}{2} & \cdot & \cdot & \cdot \\ \frac{1}{2} & \cdot & \cdot & \cdot & \frac{1}{2} & \cdot & \cdot & \cdot \\ \cdot & \frac{1}{2} & \cdot & \cdot & \cdot & \frac{1}{2} & \cdot & \cdot \\ \cdot & \frac{1}{2} & \cdot & \cdot & \cdot & \frac{1}{2} & \cdot & \cdot \\ \cdot & \cdot & \frac{1}{2} & \cdot & \cdot & \cdot & \frac{1}{2} & \cdot \\ \cdot & \cdot & \frac{1}{2} & \cdot & \cdot & \cdot & \frac{1}{2} & \cdot \\ \cdot & \cdot & \cdot & \frac{1}{2} & \cdot & \cdot & \cdot & \frac{1}{2} \\ \cdot & \cdot & \cdot & \frac{1}{2} & \cdot & \cdot & \cdot & \frac{1}{2} \end{bmatrix}.$$

◊

事实上，可以构造许多具有不同移位和不同基域的矩阵来代替 $\mathbb{F}_2$，但所有这些构造都具有相似的性质。

构造 4.9 中的矩阵具有许多显著而优美的性质。首先，这种游走恰好在 k 步后混合，即 $A^k = J$，因为经过 k 步后，到达的顶点完全均匀，且与起始顶点无关。事实上，这些矩阵最显著的性质是，所有（非平凡）特征值都为 0。这意味着每个若尔当块的大小至多为 k，且为幂零矩阵。此外，一半的奇异值为 1，另一半全为 0。值得注意的是，这些矩阵的边扩张为 $O(1/k) = O(1/\log n)$，这表明 $\Gamma(n) \in O(1/\log n)$。这些矩阵的边扩张曾在 [DT98]，中得到研究，但该研究并未给出实际的非扩展集。在下面的引理中，我们将证明 de Bruijn 矩阵的所有相关性质。

引理 4.11。 设 A 为定义在 $n = 2^k$ 个顶点上的 *de Bruijn* 矩阵。则 A 满足以下性质：

1. $\dfrac{1}{2\log n} \le \phi(A) \le \dfrac{8}{\log n}.$

2. A 是双随机的，且 $\Gamma(n) \leq \frac{8}{\text{对数 } n} \in O\left(\frac{1}{\text{对数 } n}\right)$，改进了引理 4.7。的所有非平凡特征值均为零。

4. 若尔当形式与舒尔形式相同，A 有一个大小为 1、对应特征值 1 的平凡块，并且对于特征值 0，它恰有 2^{k-1-r} 个大小为 $r \in [k-1]$的若尔当块和一个大小为 k 的块。

5. A 有 $n/2$个奇异值为 1，$n/2$个奇异值为 0。

证明。 (1) 令 S 为所有包含 $r = \lceil k/2 \rceil$ 个连续的 1 的 k位字符串构成的集合。令 $S = \bigcup_{i=r}^{k} T_i$，其中 T_i 是所有在以 i结尾的 r 个位置上均为 1 的字符串组成的集合，并且不存在具有相同性质的 $j > i$，也就是说，不存在使子串 $[j-r+1, j]$ 全为 1 的 $j > i$。注意，T_i彼此不相交。此外，T_i 的定义意味着，以 i 结尾的 r 个位置全为 1，而 $i+1$的第 $i+1$ 个位置（T_k除外）为 0。事实上，这两个约束足以定义 T_i，并确保它们互不重叠，因为若要重叠，就必须存在一个字符串，其中以 i 结尾的 r 个连续的 1 和以 j结尾的 r 个连续的 1 同时出现；但如果字符串中的 $i+1$ 位置为 0，那么留给 r 个 1 的位置只有 $k-r-1 < k/2 \leq r$ 个，这是不可能的。

因此，对于 T_i，所有字符串都有 $r+1$ 个位置固定——指标 $i-r+1$ 至 i 处包含 1，而位置 $i+1$ 处包含 0。注意，位置 $\leq i-r$ 可以是任意比特，所有位置 $> i+1$也都如此，因为后缀的长度对于所有 $i \geq r$ 都是 $k-(i+1) < r$。因此，我们得到 $|T_i| = 2^{k-r-1}$（对于 $r \leq i \leq k-1$），并且 $|T_k| = 2^{k-r}$。由于$\frac{k}{2} \leq r \leq \frac{k}{2}+1$，我们有

$$\begin{aligned} |S| &= (k-r) \cdot 2^{k-r-1} + 2^{k-r} \\ &\leq \frac{k}{2} 2^{k/2-1} + 2^{k/2} \\ &\leq \frac{5}{4} \cdot k \cdot 2^{k/2} \\ &\ll 2^{k-1} \\ &= \frac{n}{2} \end{aligned}$$

因此， S 是一个有效集合，可用于确定 ϕ。此外，

$$\begin{aligned}|S| &= (k-r)\cdot 2^{k-r-1} + 2^{k-r} \\ &\geq \left(\frac{k}{2} - 1\right) 2^{k/2-2} + 2^{k/2-1} \\ &\geq \frac{k}{8} 2^{k/2}\end{aligned}$$

所有这些集合的关键性质如下：对于任意 $i > r$，集合T_i 中的顶点，其所有出边都位于集合 S内；也就是说，集合 T_i 中的顶点到 S没有边，因为左移一位后，这些字符串总会包含一个由 r 个 1 构成的连续子串。唯一可能有边连向 S 的集合是 T_r，且它最多有 $2\cdot|T_r|$ 条出边。因此，我们得到

$$\phi_S(A) \leq \frac{2\cdot|T_r|}{|S|} \leq \frac{2\cdot 2^{k/2-1}}{\frac{1}{8}\cdot k\cdot 2^{k/2}} \leq \frac{8}{k}$$

如所需。

Fur因此，由于 $A^k = J$根据 (2)，由引理 3.22可知 $\phi(A^k) \leq k\cdot\phi(A)$o r

$$\phi(A) \geq \frac{1}{2k}.$$

(2) 从定义可直接看出， $A\mathbf{1} = \mathbf{1}$ 和 $A^T\mathbf{1} = \mathbf{1}$ ，因此 A 是双随机的。设 $a = a_1a_2\cdots a_k$ 为一个 k位字符串，令 $|a\rangle$ 表示相应的 n维空间中的标准基向量，并且对于任意 $S \subseteq [k]$，令 $a_S = \sum_{i\in S} a_i$ 和

$$v_S = \sum_a (-1)^{a_S}|a\rangle.$$

令 $V = \{v_S\colon\ S \subseteq [k]\}$ 和 $W = \{v_S\colon\ 1 \in S,\ S \subseteq [k]\}$。然后我们证明，对于任意$v_S \in W$

$$Av_S = 0.$$

注意，根据定义，

$$A|a_1\ldots a_k\rangle = \frac{1}{2}|a_2\ldots a_k0\rangle + \frac{1}{2}|a_2\ldots a_k1\rangle$$

通过直接计算，令 $T = S\backslash\{1\}$，可得

$$\begin{aligned} Av_S &= \sum_a (-1)^{a_S} A|a_1 \cdots a_k\rangle \\ &= \sum_a (-1)^{a_1}(-1)^{a_T} A|a_1 \cdots a_k\rangle \\ &= \frac{1}{2}\sum_a (-1)^{a_T}(-1)^{a_1}\left(|a_2 \cdots a_k 0\rangle + |a_2 \cdots a_k 1\rangle\right) \\ &= \frac{1}{2}\sum_{a_2,\cdots a_k=0}^{1} (-1)^{a_T}\left(|a_2 \cdots a_k 0\rangle + |a_2 \cdots a_k 1\rangle\right)\sum_{a_1=0}^{1}(-1)^{a_1} \\ &= 0 \end{aligned}$$

如前所述。此外，对于 $S \neq T$，令 $U = (S\backslash(S \cap T)) \cup (T\backslash(S \cap T)) \neq \emptyset$，则不难看出

$$\langle v_S, v_T\rangle = \sum_a \sum_b (-1)^{a_S}(-1)^{b_T}\langle a|b\rangle = \sum_a (-1)^{a_S}(-1)^{a_T} = \sum_a (-1)^{a_U} = 0.$$

因此， v_S彼此正交，且 $|W| = 2^{n-1}$；又因为对于 $v_S \in W$ $Av_S = 0$，所以核 W of A 的维数为 $n-1$。由于 A 对 $v_\emptyset$ 具有平凡特征值 1，这意味着 A 的所有非平凡特征值均为 0。

(3) 结合 (1) 和 (2)，可得 gamma 的界，因为 $\lambda_2(A) = 0$。

(4) 我们将使用（2）中的记号。按（2）中定义的集合 V 包含正交向量；如果我们用 V中的向量作为基来表示 A，就会得到一个由若尔当块组成的矩阵，因此 A 的若尔当形式和舒尔形式完全相同（因为 V 中的向量构成一个酉矩阵）。现在我们旨在理解 A 对不属于W的向量所产生的影响。这将有助于我们理解广义特征向量的链。考虑 $S \subseteq [k]$，其中 $S = \{i_1, \ldots\ldots, i_r\}$，且 $i_1 < i_2 < \ldots \ldots < i_r$，并令 $T = \{i_1 - 1, \ i_2 - 1, \ \ldots\ldots, \ i_r - 1\}$。令 $a = a_1 \ldots\ldots a_k$、 $b = b_1 \ldots \ldots b_k$，其中 $b_i = a_i + 1$ 且$b_k = 0$，并令 $c = c_1 \ldots\ldots c_k$，其中 $c_i = a_i + 1$ 且 $c_k = 1$。那么注意，如果 $i_1 = 1$，则

由（2）中的结论， $Av_S = 0$，且如果 $i_1 > 1$，则

$$\begin{aligned}
Av_S &= A\sum_a (-1)^{a_S}|a_1...a_r\rangle \\
&= \frac{1}{2}\sum_a (-1)^{a_S}|a_2...a_r 0\rangle + \frac{1}{2}\sum_a (-1)^{a_S}|a_2...a_r 1\rangle \\
&= \sum_{a_2,...,a_k=0}^{1} (-1)^{a_S}|a_2...a_r 0\rangle + \sum_{a_2,...,a_k=0}^{1} (-1)^{a_S}|a_2...a_r 1\rangle \\
&\quad [\text{since } 1 \notin S] \\
&= \sum_{b_1,...,b_{k-1}=0}^{1} (-1)^{b_T}|b\rangle + \sum_{c_1,...,c_{k-1}=0}^{1} (-1)^{c_T}|c\rangle \\
&= \sum_a (-1)^{a_T}|a\rangle \\
&= v_T.
\end{aligned}$$

因此，对于任意 $S \subseteq [k]$，记 $S + j = \{i \in [k] : i - j \in S\}$，并令 $1 \in S$ 和 $\max S = r$，则对于 $1 \le j \le k - r$，有

$$Av_{S+j} = v_{S+j-1}$$

从而得到一条由 $k + 1 - r$ 个向量组成的链，其末尾是核中的一个向量。因此，这构成了一个特定的若尔当块。一般而言，任意向量 v_S 都属于一个大小为 $k -$（$\max S - \min S$）的若尔当块。要计算不同若尔当块的数量，考虑满足 $1 \in S$ 的 S。那么，核中的向量 v_S 将是以 $v_{S+k-\max S}$ 开始的链的最后一个向量，该链位于一个大小为 $k + 1 - \max S$ 的块中。因此，对于每个 $r \in [k-1]$，大小为 r 的若尔当块的数量，等于满足 $1 \in S$ 和 $k + 1 - r \in S$ 的向量v_S的数量，这恰好是 2^{k-1-r}。此外，唯一的大小为 k 的块对应于以 $v_{\{1\}}$ 结束的链，而大小为 1 的平凡块对应于 $v_\emptyset$。作为合理性检查，如果将所有块的大小相加，我们得到

$$\sum_{r=1}^{k-1} r \cdot 2^{k-1-r} + k + 1 = 2^k\left(1 - 2^{-(k-1)} - (k-1)2^{-k}\right) + k + 1 = 2^k = n$$

正如预期。

(5) 注意，利用由 $V =$ 中的向量构成的酉矩阵 $U\{v_S : S \subseteq [k]\}$（经过适当归一化），我们得到 $A = UTU^*$，其中 T 是如上文 (4) 所示的一组若尔当块。由于大小为 r 的若尔当块乘以一个置换矩阵后，可转换为大小为 r 的对角矩阵，其中有 $r - 1$ 个 1 和一个 0 元素，

因此我们得到 $T = DP$，其中 P 是一个置换矩阵，可将每个若尔当块转换为对角矩阵。因此，由于 Q 是酉矩阵，我们得到 $A = UDPU^* = UDQ$ 的奇异值分解。要计算 D 的对角线上 0 的个数，请注意，每个非平凡若尔当块恰好贡献一个 0，因此 0 的总数等于非平凡若尔当块的个数，这恰好是

$$\sum_{r=1}^{k-1} 2^{k-1-r} + 1 = 2^{k-1} = \frac{n}{2}$$

这意味着对角线上有 $n - n/2$ 个 1。因此，A 有 $n/2$ 个奇异值为 1，另有 $n/2$ 个奇异值为 0。 □

这种构造的妙处在于，它极其简单，却仍具有上述引理中的显著性质。此外，这种构造是其他构造的基准；尽管它未能使 $\Gamma(n)$ 降至足以满足定理 1.2的水平，但它将成为第 4.5节中构造的起点。

现在，我们的目标是突破引理 4.11中对 Γ 给出的上界，并弄清定理 1.2 中对 ϕ 给出的下界是否紧，还是比真实值差了指数级。注意，如果以下结论成立：

$$\Gamma(n) \in \Omega\left(\frac{1}{\text{polylog}(n)}\right),$$

这意味着我们用于证明定理 1.2 的技术极其薄弱，要得到更紧的界就需要采用不同的技术。从实用角度看，这意味着即使在非可逆情形下，谱隙也是 ϕ 的良好估计（相差至多多对数级的(n) 项；在输入长度为 $O(\log n)$的简洁定义的链中，这些项基本可以忽略）。为此，我们将尝试构造矩阵，以突破引理 Γ 4.11中给出的上界。

4.5 搜索空间的约束

此时，我们正在寻找一个双随机矩阵 A，以帮助改进引理4.11中的界。由于从非扩张的角度来看，搜索空间（所有双随机矩阵）难以理解，我们将系统地尝试对其施加有意义的约束，以找出我们所需矩阵的类型和形式

。为此，我们首先要回答的问题——而事实证明，这个问题也足以作为最后一个问题——是理解以下内容：

主要问题： 当所有非平凡特征值都为 0 时，双随机矩阵的边扩张可以有多小？

如果我们能证明 $\Gamma(n) \in \Omega(1/\log n)$ 对所有 $n \times n$ 非平凡特征值全为 0 的矩阵都成立，那么就能说明德布鲁因矩阵在定义 4.1 和第 4.1 节所述的非扩张意义下是最优的，也就是说，在所有（非平凡）特征值全为 0 的矩阵中，它们的边扩张最小。这将是我们的第一个约束；此外，正如接下来所述，这一约束也足够一般。

约束 1. 将所有非平凡特征值限制为 0。

选择所有特征值为 0 的理由如下：正如引理 3.26 证明的前几步所示，对于双随机矩阵 A，如果每个非平凡特征值的模至多为某个常数 $1 - c$ c，那么只需将矩阵取 $O(\log n)$ 次幂，对角线元素的模就会变成多项式倒数级小，因此该矩阵的行为似乎应当与所有特征值均为 0 的矩阵相似。因此，如果 A 的所有特征值的模都小于 $1 - c$，我们只需将矩阵 A^k 取 $k \in O(\log n)$ 次幂，并将其作为起始矩阵。我们知道，其扩张至多为 A 扩张的 $O(\log n)$ 倍（由引理 3.22 可知）；因此，如果 A 的扩张较小——约为 $O(1/\sqrt{n})$——那么 A^k 的扩张也会相近（相差一个 $\log n$ 因子）。对于任意具有 $\mathrm{Re}\lambda_2 < 1$ 的双随机矩阵 A，通过惰化很容易确保 $|\lambda_i| \leq 1 - c$ 满足某个特定的 c；也就是说，可以考虑 $\exp(t \cdot (A - I))$ 或 $1\ 2\ (\ A + I)$，这实质上就是引理 3.30 证明的第一步，或定理 1.2 下界证明的第一步。

–

在将讨论范围限制为所有（非平凡）特征值均为 0 的矩阵后，我们始终主要关注矩阵 $B = A - J$。如果 A 的所有非平凡特征值均为 0，这意味着 B 的所有特征值均为 0；我们还将把 B 的舒尔形式写为 $B = UTU^*$，其中 U 是酉矩阵，T 是上三角矩阵，且其所有对角线元素均为 0。

引理 4.12. 如果 $A = J + B$ 的所有（非平凡）特征值均为 0，那么 $B^{n-1} = 0$。

证明。 由于 T 是幂零矩阵，证明显而易见。 □

注意，矩阵的各行和各列中的矩阵元素 A 之和始终为 1，因为我们在 J 中移除了全 1 特征向量，并确保 U 的第一列是向量 $1/\sqrt{n}\cdot\mathbf{1}$，这意味着 U 的所有其他列都与向量$\mathbf{1}$ 正交（因为 U 是酉矩阵），进而可得 $B\mathbf{1}=0$和 $B^T\mathbf{1}=0$。此外，约束1 还带来另一个有用的推论：由于 A 的所有特征值都是实数，我们可以将范围限制为实酉矩阵 U（考察获得舒尔分解的过程即可看出，因为 A 是实矩阵且具有实特征值）。这实际上给出了第二个隐含条件：

约束 2. 酉矩阵 U 是实矩阵，且其第一列为 $\frac{1}{\sqrt{n}}\mathbf{1}$。

基于这一观察，下一步是考察引理 3.26 的证明。在引理 3.26 中，主要目标是为 k 找到一个上界，使得 $T^k\approx 0$，并且我们发现，取约为 $n/1-\mathrm{Re}\lambda_2$ 的 k 就足够了。现在，我们的主要目标是找到一种矩阵构造，使该引理中的界能够取到。换言之，我们现在要寻找一种 T（以及 B）的构造，使得当 k 足够小时，范数 $\|T^k\|_2$ 充分远离零。考虑德布鲁因矩阵以及矩阵 $B_{dB}=A_{dB}-J$。我们知道，当 $k=\log n$ 时，$B_{dB}^k=0$。此外，由引理 4.11 可知，由于 $\phi(A_{dB})\le 8/k$，这也意味着对于任意 $j\ll k$，$\|B^j\|\gg 0$；否则，根据引理 4.11 的证明，我们会得到 $\phi(A)>1/j$，这将产生矛盾。

构造一个 T 使得 $T^k\gg 0$ 对于 $k\in o(n)$ 并不困难。例如，如果令舒尔形式为一个大型若尔当块，即 T 只是一个对角线上方全为 1 的矩阵，那么就很容易得到这一结果。

$$T=\begin{bmatrix}\mathbf{0}&0&0&0&0&0\\0&\mathbf{0}&1&0&0&0\\0&0&\mathbf{0}&1&0&0\\0&0&0&\ddots&\ddots&0\\0&0&0&0&\mathbf{0}&1\\0&0&0&0&0&\mathbf{0}\end{bmatrix},\ T^2=\begin{bmatrix}\mathbf{0}&0&0&0&0&0\\0&\mathbf{0}&0&1&0&0\\0&0&\mathbf{0}&0&\ddots&0\\0&0&0&\ddots&\ddots&1\\0&0&0&0&\mathbf{0}&0\\0&0&0&0&0&\mathbf{0}\end{bmatrix},\ T^3=\begin{bmatrix}\mathbf{0}&0&0&0&0&0\\0&\mathbf{0}&0&0&1&0\\0&0&\mathbf{0}&0&\ddots&1\\0&0&0&\ddots&\ddots&0\\0&0&0&0&\mathbf{0}&0\\0&0&0&0&0&\mathbf{0}\end{bmatrix}$$

对于这个矩阵，显然，对于任意 $k<n-1$，$\|T^k\|=1$，因为由 1 构成的对角线会随着 T 的幂次增大而不断远离主对角线。然而，选择如上所述的矩阵 T 时，眼下的问题是，可能不存在任何酉矩阵 U 使得 $J+UTU^*=A$ 是一个非负矩阵。事实上，构造这些矩阵时的主要问题之一就是——

确保它们为正。注意，由于条件2，矩阵各行和各列的元素之和始终为 1。此时，很难找到一个酉矩阵 U，将上述矩阵 T 转换为双随机矩阵，因此我们需要以某种方式修改 T。注意，本质上，对 T 的唯一要求是它为对角线元素为 0 的上三角矩阵，但要处理一个完全一般的 T 极其困难。因此，我们接下来对 T 作如下进一步放宽/约束：

约束3。对于 $A = J + UTU^*$，令 T 在对角线上方有某个数值 r，而不是 1，并令 T 中的其余矩阵元素为 0。

由上述约束可得

$$T = \begin{bmatrix} \mathbf{0} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{0} & r & 0 & 0 & 0 \\ 0 & 0 & \mathbf{0} & r & 0 & 0 \\ 0 & 0 & 0 & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \mathbf{0} & r \\ 0 & 0 & 0 & 0 & 0 & \mathbf{0} \end{bmatrix}, \quad T^2 = \begin{bmatrix} \mathbf{0} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{0} & 0 & r^2 & 0 & 0 \\ 0 & 0 & \mathbf{0} & 0 & \ddots & 0 \\ 0 & 0 & 0 & \ddots & \ddots & r^2 \\ 0 & 0 & 0 & 0 & \mathbf{0} & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{0} \end{bmatrix}, \quad T^3 = \begin{bmatrix} \mathbf{0} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{0} & 0 & 0 & r^3 & 0 \\ 0 & 0 & \mathbf{0} & 0 & \ddots & r^3 \\ 0 & 0 & 0 & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \mathbf{0} & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{0} \end{bmatrix}$$

注意，对于 $r \approx 1 - 1/n$，由于 $\|T^k\| = r^k$，对于任意 $k < n/2$，我们都有 $\|T^k\| = r^k \gtrsim e^{-1/2}$，这远大于 0。因此，如果我们能通过任意满足 $r \approx 1 - 1/n$ 的有效酉矩阵 U 将 T变换为双随机矩阵，就能得到我们所要寻找的那类矩阵。此时，我们的目标不再是让 r 约为 $O(1/n)$ 且远离 1，而是找到一个矩阵 T，使得存在某个酉矩阵可将其变换为双随机矩阵，并使 r 尽可能大。

此外，T 仅在非对角线上有非零矩阵元素的理由如下：对于任何 T 对角线元素全为零的上三角矩阵，非对角线矩阵元素对其幂中的矩阵元素影响最大，因为距离对角线较远的矩阵元素经过几次 T幂运算后就会失去作用。因此，选择具有若尔当块非零元素分布模式的 T，其结果不会离最优解太远。

注意，我们的结构 T 给出了以下引理，有助于说明试图构造出能够超越 Γ 在引理 4.11中的上界时所面临的主要问题。

引理 4.13。设 $A = J + UTU^*$ 是一个双随机矩阵，其中 T 是全零矩阵，但从第 2 行到 $n-1$ 的非对角线矩阵元素中有一个元素为 r。那么

$$r^2 = \frac{\sum A_{i,j}^2 - 1}{n-2}$$

证明。考察 AA^T 的迹，我们得到

$$\sum_{i,j} A_{i,j}^2 = \mathrm{Tr}(AA^T) = 1 + r^2(n-2)$$

由此可得该引理。 □

请注意，根据该引理，最大化 r 似乎是一项简单的任务——最大化 A的矩阵元素平方和；而且很容易构造这样的A，例如，对于适当的 p，$A = (1-p)I + pJ$。这说明了两个问题。该引理没有用到我们对 A 的约束——非负性，以及所有非平凡特征值均为 0 这一事实。这是所有试图构造目标矩阵的方法面临的首要问题。

注 4.14。（模拟的问题）我们还想重申引言中提到的一点：无法通过模拟找到不同的 A，使其为双随机矩阵、扩张较小但谱隙较大；而且，要模拟 T 和 U 并使其具备我们所关心的性质，则更加困难。主要问题在于这些矩阵对微小扰动极其敏感；事实上，我们的目标正是找到这样的矩阵：它们会因微小扰动（对应于扩张的微小变化）而受到严重影响（就特征值的变化而言）。由于有限位数的模拟本身就会产生扰动，因此数值上极不可能找到这样的矩阵。◊

4.6 满足约束的特定矩阵的特殊情形

为了着手实现我们的目标——找到 T，使其 r 的值较大，并找到某个能将其变换为双随机矩阵的酉矩阵，我们将从一个非常简单的情形开始。起初，我们不愿考察具体的小例子，因为看起来，对于任何有限的 n（例如，足够小的 $n \le 12$），边扩张和谱隙的数值都会相差很大的常数倍；而且由于它们是标量值，本质上并不能提供有用的信息。然而，结果表明，这些矩阵揭示了一些我们起初没有预料到的东西。

考虑 $n = 3$的情形。设我们的 T和 U如下：

$$T = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & r \\ 0 & 0 & 0 \end{bmatrix}, \quad U = \begin{bmatrix} \frac{1}{\sqrt{3}} & a_1 & b_1 \\ \frac{1}{\sqrt{3}} & a_2 & b_2 \\ \frac{1}{\sqrt{3}} & a_3 & b_3 \end{bmatrix}$$

其中 a_i和 b_i都是实数，如约束2所述。现在，我们可以将它们视为 x-y 平面上的 3 个点，并利用 U是酉矩阵这一事实，得到以下方程：

$$\begin{aligned} a_1 + a_2 + a_3 &= 0 \\ b_1 + b_2 + b_3 &= 0 \\ a_1^2 + b_1^2 &= \frac{2}{3} \\ a_2^2 + b_2^2 &= \frac{2}{3} \\ a_3^2 + b_3^2 &= \frac{2}{3} \end{aligned}$$

最后三个方程告诉我们，这三个点都位于半径为 $t = \sqrt{2/3}$ 的圆上；前两个方程则表明，$\overline{\text{向}}$量 $(a_3, \ b_3) = -(a_1, \ b_1) - (a_2, \ b_2)$。由于这三个点都在圆上，因此第三个点位于从原点指向前两个点的向量所成角的角平分线上。然而，由于这些方程完全对称，这一条件对三个点都成立，因此可知这三个点位于等边三角形的三个顶点上。由此可得，这些点彼此相差 $2\pi/3$ 的角度，从而得到

$$\begin{aligned} (a_1, b_1) &= (a, b) \\ (a_2, b_2) &= (-\frac{1}{2}a - \frac{\sqrt{3}}{2}b, -\frac{1}{2}b + \frac{\sqrt{3}}{2}a) \\ (a_3, b_3) &= (-\frac{1}{2}a + \frac{\sqrt{3}}{2}b, -\frac{1}{2}b - \frac{\sqrt{3}}{2}a) \end{aligned}$$

因此，我们得到了酉矩阵 U，接下来需要选择（$a, \ b$）（其中 $a^2 + b^2 = \frac{2}{3}$），以便在保持矩阵非负的同时，使 r 尽可能大。注意，由于 T和 U的结构，我们得到矩阵元素 $B_{i, \ j} = r \cdot a_i \cdot b_j$，并且需要确保 $B_{i, \ j} + \frac{1}{3} \geq 0$。因此，我们的问题转化为求解以下问题：

最大化 r

s.t. $B_{i,j} + \frac{1}{3} \geq 0$ for all i, j, and $a^2 + b^2 = \frac{2}{3}$.

上述问题有一个简单的解，可通过以下方式得到

$$a = b = \frac{1}{\sqrt{3}}$$

给出

$$r = \frac{\sqrt{3}+1}{\sqrt{3}+2} = 1 - \frac{1}{\sqrt{3}+2}.$$

请注意，此时我们得到一个双随机矩阵 A，其所有特征值均为 0，并且具有上文给出的特定值 r；但所有这些数值都可能存在很大的误差，因此实际上毫无意义。

真正有意义的是考察酉矩阵 U:

$$U = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{2}(1+\frac{1}{\sqrt{3}}) & \frac{1}{2}(1-\frac{1}{\sqrt{3}}) \\ \frac{1}{\sqrt{3}} & -\frac{1}{2}(1-\frac{1}{\sqrt{3}}) & -\frac{1}{2}(1+\frac{1}{\sqrt{3}}) \end{bmatrix}$$

令人着迷的是，除了第一列是**1**的倍数之外，甚至 的第一行也是全一向量的倍数。并且，由于 U 表示在 T 的空间中的旋转，而且它是一个向量，不同于作为标量的数值 ϕ、 r、 λ_2，因此这表明 r 的最大值是在全 1 向量的方向上取得的。这给出了下一个关键约束:

约束 4。 r 的最大值在 U 的第一行包含全一向量时取得。

4.7 一般情形及相关推理

现在，我们希望了解，对于第一行和列向量均为全一向量的倍数的任意酉矩阵， r 的值可以有多大。事实上，我们可以证明如下结论:

引理 4.15。 设 $A = J + UTU^*$ 为一个双随机矩阵，其中 T 的对角线元素为零，而 r 的非对角线元素（从第 2 行到 n）为 U

是一个实酉矩阵，其第一行和第一列均为全一向量的倍数。那么

$$r \leq 1 - \frac{\sqrt{n}+2}{n-4} \in 1 - \Omega\left(\frac{1}{\sqrt{n}}\right).$$

证明。设 U 的第 k列为 u_k，其中 $u_{i,\ k}$ 表示第 i列的第 k 个元素，或者表示 U的第（$i,\ k$）个元素，那么根据我们对 T 的选择以及关于 U的条件，可得

$$A = J + r \cdot \sum_{k=2}^{n-1} u_k u_{k+1}^*$$

and

$$A_{i,j} = \frac{1}{n} + r \sum_{k=2}^{n-1} u_{i,k} u_{j,k+1}$$

由于酉矩阵是实数矩阵（如约束2所示）。注意到 U 的第一行和第一列中的每个元素都是 $\frac{1}{\sqrt{n}}$，因此对于所有 i, $u_{i,\ 1} = \frac{1}{\sqrt{n}}$；对于所有 k，$u_{1,\ k} = \frac{1}{\sqrt{n}}$。又由于 U 的第一行和第一列都是全一向量的倍数，U 的其他任意行或列的元素之和均为 0。因此可得

$$A_{1,1} = \frac{1}{n} + r \cdot \frac{n-2}{n}$$
$$1 - A_{1,1} = 1 - \frac{1}{n} - r + \frac{2r}{n}$$

对于 $j \geq 2$,

$$\begin{aligned} A_{1,j} &= \frac{1}{n} + r \sum_{k=2}^{n-1} u_{1,k} u_{j,k+1} \\ &= \frac{1}{n} + \frac{r}{\sqrt{n}} \sum_{k=2}^{n-1} u_{j,k+1} \\ &= \frac{1}{n} + \frac{r}{\sqrt{n}}(-u_{j,1} - u_{j,2}) \\ &\quad [\text{since} \sum_{k=1}^{n} u_{j,k} = 0] \\ &= \frac{1-r}{n} - \frac{r}{\sqrt{n}} \cdot u_{j,2} \end{aligned}$$

这给出 $j \geq 2$

$$u_{j,2} = \frac{\sqrt{n}}{r} \cdot \left(\frac{1-r}{n} - A_{1,j}\right).$$

现在考虑 u_2（或 U 的第二列），则有

$$\begin{aligned}
1 &= \sum_{j=1}^{n} u_{j,2}^2 \\
&= \frac{1}{n} + \frac{n}{r^2} \sum_{j=2}^{n} \left(\frac{1-r}{n} - A_{1,j} \right)^2 \\
\frac{n-1}{n^2} \cdot r^2 &= \frac{n-1}{n^2} \cdot (1-r)^2 + \sum_{j=2}^{n} A_{1,j}^2 - 2 \cdot \frac{1-r}{n} \cdot \sum_{j=2}^{n} A_{1,j}
\end{aligned} \tag{4.16}$$

注意到由于该矩阵是双随机的，$\sum_{j=2}^{n} A_{1,\ j} = 1 - A_{1,\ 1}$ 以及每个$A_{1,\ j} \geq 0$，从而得到

$$\sum_{j=2}^{n} A_{1,j}^2 \leq \left(\sum_{j=2}^{n} A_{1,j} \right)^2 = (1 - A_{1,1})^2.$$

将这些代入方程 4.16，得到

$$\begin{aligned}
\frac{n-1}{n^2} \cdot r^2 &\leq \frac{n-1}{n^2} \cdot (1-r)^2 + (1 - A_{1,1})^2 - 2 \cdot \frac{1-r}{n} \cdot (1 - A_{1,1}) \\
\frac{n-1}{n^2} \cdot r^2 &\leq \frac{n-1}{n^2} \cdot (1-r)^2 + (1 - A_{1,1}) \left(1 - A_{1,1} - 2 \cdot \frac{1-r}{n} \right) \\
(n-1)(2r-1) &\leq (n(1-r) + 2r - 1)(n(1-r) + 4r - 3) \\
0 &\leq (n^2 - 6n + 8)r^2 + (-2n^2 + 8n - 8)r + n^2 - 3n + 2
\end{aligned}$$

并且由于 $r \leq 1$（否则该矩阵的范数将大于 1，因而不是双随机矩阵，参见引理 3.3），解上述二次方程所得的 r 有效范围为

$$r \leq 1 - \frac{\sqrt{n} + 2}{n - 4}$$

这意味着

$$r \in 1 - \Omega\left(\frac{1}{\sqrt{n}} \right)$$

证明完毕。 □

换言之，引理 4.15 指出，将酉矩阵的第一行固定为全一向量后，在给定我们已固定的 T 形式的情况下，我们所能期望达到的 r 最佳值约为 $1 - \frac{1}{\sqrt{n}}$，这也将成为我们接下来的目标。请注意，如果~~可~~行，这将带来相较于指数级的提升

引理 $\Gamma(n)$ 中的4.11所述的 $r$deBruin 构造。于是，为了使 r取得最大值，我们的优化问题如下：

优化问题 1。

最大化 r

使得：

$\frac{1}{n} + r\sum_{k=2}^{n-1} u_{i,k}u_{j,k+1} \geq 0$ for all $2 \leq i, j \leq n$,

$\frac{1}{n} + \frac{r}{\sqrt{n}}\sum_{k=2}^{n-1} u_{j,k+1} \geq 0$ for all $2 \leq j \leq n$

$\frac{1}{n} + \frac{r}{\sqrt{n}}\sum_{k=2}^{n-1} u_{i,k} \geq 0$ for all $2 \leq i \leq n$

$u_{i,k} = \langle e_i, Ue_k \rangle$

$UU^T = U^TU = I.$

$u_{1,j} = \frac{1}{\sqrt{n}}, u_{i,1} = \frac{1}{\sqrt{n}}$, for all $1 \leq i, j \leq n.$

尽管我们可以尝试求解上述问题，但它十分棘手，而且关于（$n-1$）^{2} 变量在 U 中的条件使其几乎难以处理。为解决这一问题，并了解该问题的解会是什么样子，我们将通过固定酉矩阵 U的其他矩阵元素来简化问题。这种固定方式会稍有创造性，并基于我们在 $n = 3$ 中的观察以及其他小 n情形下的情况进行经验性选择。请注意，直观而言，根据酉矩阵 U的约束，它是在两个不同的“舒尔”块所张成的空间中进行旋转；舒尔块与若尔当块相同，只是其非对角线矩阵元素为 r，而不是 1。第一个是仅包含特征值 1 的平凡块，第二个大小为 $n-1$ 的块包含其余的零特征值，并且在对角线上方有 r，其他位置均为零。由于对舒尔块只执行两种操作，我们将把酉矩阵 U 任意一列中（位置大于 1 处）的矩阵元素设为仅由两个不同的值组成。在这一限制下，由于 U 在第一行和第一列中有 $1/\sqrt{n}$，且 U 必须是酉矩阵，因此每列仅含两个不同值的酉矩阵恰好只有一个；结果表明，它是一个对称酉矩阵，其中 $U_{i,i} = \alpha$，且 $U_{i,j} = \beta$，对于 $2 \leq i, j \leq n$。

固定酉矩阵 U和上三角矩阵 T后，我们的目标是选择一个 r 尽可能大的值，使得所得矩阵 A 为非负。因此，优化问题1 变为：

优化问题 2。

最大化 r

满足：

$\frac{1}{n} + r\sum_{k=2}^{n-1} u_{i,k}u_{j,k+1} \geq 0$ for all $2 \leq i, j \leq n$

$\frac{1}{n} + \frac{r}{\sqrt{n}}\sum_{k=2}^{n-1} u_{j,k+1} \geq 0$ for all $2 \leq j \leq n$

$\frac{1}{n} + \frac{r}{\sqrt{n}}\sum_{k=2}^{n-1} u_{i,k} \geq 0$ for all $2 \leq i \leq n$

在这些不等式所给出的约束条件下，我们的目标是最小化 ϕ 或最大化 r。由于我们对 U的限制，所得矩阵 A 中的割 $S = \{1\}$具有特殊性质，因此我们通过最小化该割的边扩张$1 - A_{1,\ 1}$ 来求解上述问题。对于 r 可取的值，我们注意到，若将所有矩阵元素$A_{1,\ i}$ （ $3 \leq i \leq n$ ）的值强制设为 0，就能得到最小化$1 - A_{1,\ 1}$ 或最大化$A_{1,\ 1}$ 这一优化问题的一组极点。随后，我们对所得矩阵的 r 进行最大化（此时实际上恰有两个可能的双随机矩阵），结果便是以下构造。

4.8 Rootn 矩阵 – 固定谱隙并最小化边扩张

构造 4.17。令 $m = \sqrt{n},$

$$a_n = \frac{m^2 + m - 1}{m \cdot (m+2)}, \qquad b_n = \frac{m+1}{m \cdot (m+2)}, \qquad c_n = \frac{1}{m \cdot (m+1)},$$

$$d_n = \frac{m^3 + 2m^2 + m + 1}{m \cdot (m+1) \cdot (m+2)}, \quad e_n = \frac{1}{m \cdot (m+1) \cdot (m+2)}, \quad f_n = \frac{2m+3}{m \cdot (m+1) \cdot (m+2)},$$

并定义 $n \times n$ 矩阵

$$A_n = \begin{bmatrix} a_n & b_n & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & c_n & d_n & e_n & e_n & e_n & e_n & \cdots & e_n \\ 0 & c_n & e_n & d_n & e_n & e_n & e_n & \cdots & e_n \\ 0 & c_n & e_n & e_n & d_n & e_n & e_n & \cdots & e_n \\ 0 & c_n & e_n & e_n & e_n & d_n & e_n & \cdots & e_n \\ 0 & c_n & e_n & e_n & e_n & e_n & d_n & \cdots & e_n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & c_n & e_n & e_n & e_n & e_n & e_n & \ldots & d_n \\ b_n & f_n & c_n & c_n & c_n & c_n & c_n & \cdots & c_n \end{bmatrix}$$

◊

例 4.18。构造 4.17 的 $n = 9$ 如下所示：

$$A_9 = \frac{1}{60}\begin{bmatrix} 44 & 16 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 49 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 5 & 1 & 49 & 1 & 1 & 1 & 1 & 1 \\ 0 & 5 & 1 & 1 & 49 & 1 & 1 & 1 & 1 \\ 0 & 5 & 1 & 1 & 1 & 49 & 1 & 1 & 1 \\ 0 & 5 & 1 & 1 & 1 & 1 & 49 & 1 & 1 \\ 0 & 5 & 1 & 1 & 1 & 1 & 1 & 49 & 1 \\ 0 & 5 & 1 & 1 & 1 & 1 & 1 & 1 & 49 \\ 16 & 9 & 5 & 5 & 5 & 5 & 5 & 5 & 5 \end{bmatrix}.$$

注 $\lambda_2(A_9) = 0$, 和 $\phi(A_9) = \frac{16}{\cancel{60}} < \frac{1}{\sqrt{\cancel{9}}}$。 ◊

定理 4.19。 以下结论适用于矩阵 A_n 在构造*4.17*中：

1. A_n 是双随机的。
2. A_n 的每个非平凡特征值均为 *0*。
3. 边扩张有如下界：

$$\frac{1}{6\sqrt{n}} \leq \phi(A_n) \leq \frac{1}{\sqrt{n}}.$$

4. 由 *1*、*2*、*3* 可得，

$$\phi(A_n) \leq \frac{1 - Re\lambda_2(A_n)}{\sqrt{n}}$$

因此

$$\Gamma(n) \leq \frac{1}{\sqrt{n}},$$

相比引理 *4.11* 中的界，呈指数级改进。

证明。 以下计算很容易验证，可知 A_n 是一个双随机矩阵：

1. $a_n \geq 0, b_n \geq 0, c_n \geq 0, d_n \geq 0, e_n \geq 0, f_n \geq 0.$
2. $a_n + b_n = 1.$

3. $c_n + d_n + (n-3)e_n = 1$。

4. $b_n + f_n + (n-2)c_n = 1$。

这就完成了 (1) 的证明。

A_n 由酉矩阵 U_n三角化为 T_n，即

$$A_n = U_n T_n U_n^*,$$

其中 T_n 和 U_n 定义如下。回忆一下，$m = \sqrt{n}$。令 $^-$

$$r_n = 1 - \frac{1}{m+2},$$

$$\alpha_n = \frac{-n^2 + 2n - \sqrt{n}}{n \cdot (n-1)} = -1 + \frac{1}{m \cdot (m+1)},$$

$$\beta_n = \frac{n - \sqrt{n}}{n \cdot (n-1)} = \frac{1}{m \cdot (m+1)},$$

$$T_n = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & r_n & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & r_n & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & r_n & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & r_n & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & r_n & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & r_n \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

and

$$U_n = \begin{bmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} \\ \frac{1}{\sqrt{n}} & \alpha_n & \beta_n & \beta_n & \beta_n & \beta_n & \cdots & \beta_n & \beta_n \\ \frac{1}{\sqrt{n}} & \beta_n & \alpha_n & \beta_n & \beta_n & \beta_n & \cdots & \beta_n & \beta_n \\ \frac{1}{\sqrt{n}} & \beta_n & \beta_n & \alpha_n & \beta_n & \beta_n & \cdots & \beta_n & \beta_n \\ \frac{1}{\sqrt{n}} & \beta_n & \beta_n & \beta_n & \alpha_n & \beta_n & \cdots & \beta_n & \beta_n \\ \frac{1}{\sqrt{n}} & \beta_n & \beta_n & \beta_n & \beta_n & \alpha_n & \cdots & \beta_n & \beta_n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{1}{\sqrt{n}} & \beta_n & \beta_n & \beta_n & \beta_n & \beta_n & \cdots & \alpha_n & \beta_n \\ \frac{1}{\sqrt{n}} & \beta_n & \beta_n & \beta_n & \beta_n & \beta_n & \cdots & \beta_n & \alpha_n \end{bmatrix}.$$

要证明 U_n 是一个酉矩阵，可以很容易地验证以下计算 :

1. $\frac{1}{n} + \alpha_n^2 + (n-2)\cdot\ \beta_n^2 = 1.$ 2.
$\frac{1}{\sqrt{n}} + \alpha_n + (n-2)\cdot\ \beta_n = 0.$ 3.
$\frac{1}{n} + 2\cdot\ \alpha_n \cdot\ \beta_n + (n-3)\cdot\ \beta_n^2 = 0.$

Als因此，要看出 $A_n = U_n T_n U_n^*$, 以下计算同样很容易验证：

1.
$$A_n(1,1) = a_n = \langle u_1, Tu_1\rangle = \frac{1}{n} + \frac{1}{n}\cdot(n-2)\cdot r_n.$$

2.
$$A_n(1,2) = b_n = \langle u_1, Tu_2\rangle = \frac{1}{n} + \frac{1}{\sqrt{n}}\cdot(n-2)\cdot r_n\cdot\beta_n.$$

3.
$$A_n(n,1) = b_n = \langle u_n, Tu_1\rangle = \frac{1}{n} + \frac{1}{\sqrt{n}}\cdot(n-2)\cdot r_n\cdot\beta_n.$$

4. 对于 $3 \le j \le n$,
$$A_n(1,j) = 0 = \langle u_1, Tu_j\rangle = \frac{1}{n} + \frac{1}{\sqrt{n}}\cdot\alpha_n\cdot r_n + (n-3)\cdot\frac{1}{\sqrt{n}}\cdot\beta_n\cdot r_n.$$

5. 对于 $2 \le i \le n-1$,
$$A_n(i,1) = 0 = \langle u_i, Tu_1\rangle = \frac{1}{n} + \frac{1}{\sqrt{n}}\cdot\alpha_n\cdot r_n + \frac{1}{\sqrt{n}}\cdot(n-3)\cdot\beta_n\cdot r_n.$$

6. 对于 $2 \le i \le n-1$,
$$A_n(i,2) = c_n = \langle u_i, Tu_2\rangle = \frac{1}{n} + \alpha_n\cdot\beta_n\cdot r_n + (n-3)\cdot\beta_n^2\cdot r_n.$$

7. 对于 $3 \le j \le n$,
$$A_n(n,j) = c_n = \langle u_n, Tu_j\rangle = \frac{1}{n} + \alpha_n\cdot\beta_n\cdot r_n + (n-3)\cdot\beta_n^2\cdot r_n.$$

8. 对于 $2 \le i \le n-1$,
$$A_n(i,i+1) = d_n = \langle u_i, Tu_{i+1}\rangle = \frac{1}{n} + \alpha_n^2\cdot r_n + (n-3)\cdot\beta_n^2\cdot r_n.$$

9. 对于 $2 \le i \le n-2$、 $3 \le j \le n$、 $i+1 \ne j$,
$$A_n(i,j) = e_n = \langle u_i, Tu_j\rangle = \frac{1}{n} + 2\cdot\alpha_n\cdot\beta_n\cdot r_n + (n-4)\cdot\beta_n^2\cdot r_n.$$

10.

$$A_n(n,2) = f_n = \langle u_n, Tu_2 \rangle = \frac{1}{n} + (n-2) \cdot r_n \cdot \beta_n^2.$$

由此，我们得到 A_n 的舒尔分解。由于 T_n 的对角线除平凡特征值 1 外全为零，因此 A_n 的所有非平凡特征值均为零。至此，(2) 的证明完成。

若令集合为 $S = \{1\}$，则可得

$$\phi(A_n) \leq \phi_S(A_n) = b_n < \frac{1}{\sqrt{n}}.$$

此外，由于 T_n可写为 $\Pi_n D_n$，其中 $D_n(1,1) = 1$，$D_n(i,i) = r_n$对 $i = 2$到 $n-1$成立，且 $D_n(n,n) = 0$对某个置换 Π_n成立，因此可得 $A_n = (U_n \Pi_n) D_n U_n^*$，这给出了 A_n 的奇异值分解，因为 $U_n \Pi_n$ 和 U_n^* 是酉矩阵。因此，A_n 恰有一个奇异值为 1、$n-2$ 个奇异值为 r_n，以及一个奇异值为 0。因此，由引理 3.12 可得

$$\phi(A) \geq \frac{1 - r_n}{2} = \frac{1}{2 \cdot (\sqrt{n} + 2)} \geq \frac{1}{6\sqrt{n}}$$

由此完成了 (3) 的证明。 □

注 4.20。我们注意到，对于 A_n 在定理 4.19中构造的矩阵，有

$$\phi(A_n) \leq \frac{1 - |\lambda_i(A)|}{\sqrt{n}}$$

对任意 $i \neq 1$均成立，这比定理 1.2所要求的保证更强 .

◊

我们再次强调，通过算法模拟得出这样的构造并不太可能，因为矩阵 A_n 的特征值极其敏感。尽管$\lambda_2(A_n) = 0$，但如果我们只移动矩阵 A_n 中的 $O(1/\sqrt{n})$ 质量，通过将 a_n 替换为 $a_n' = a_n + b_n$、将 b_n 替换为 $b_n' = 0$、将 f_n 替换为 $f_n' = f_n + b_n$，并保持 c_n、d_n、e_n 不变，从而构造出矩阵A_n'，那么 $\lambda_2(A_n') = 1$。因此，由于对 $O(1/\sqrt{n})$ 的扰动（当 n 很大时，这个量非常小）会使第二特征值从 0 跃变到 1（谱隙则从 1 跃变到 0），所以不可能通过对随机矩阵进行微小改动，得到满足定理 4.19 所需性质的构造。

4.9 Rootn 矩阵构造的观察

此时，我们知道定理 ϕ 1.2 中的 下界已接近最优，而对于 n^{α} 而言，损失 $\frac{1}{2} \leq \alpha \leq 1$是不可避免的。现在我们的目标是找到一个能达到界 $\Gamma(n) \in O(1/n)$ 的构造。不过，这将把我们引向与此前不同的方向。要踏上这条新方向，我们需要先从 Rootn 矩阵的构造中得到一些观察。我们的第一个观察来自考察其背后的马尔可夫链。注意，构造4.17中的 Rootn 矩阵看起来大致如下，其中 $\epsilon \in \Theta(1/\sqrt{n})$。

$$A_n = \begin{bmatrix} 1-\epsilon & \epsilon & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \epsilon^2 & 1-\epsilon & \epsilon^3 & \epsilon^3 & \epsilon^3 & \epsilon^3 & \cdots & \epsilon^3 \\ 0 & \epsilon^2 & \epsilon^3 & 1-\epsilon & \epsilon^3 & \epsilon^3 & \epsilon^3 & \cdots & \epsilon^3 \\ 0 & \epsilon^2 & \epsilon^3 & \epsilon^3 & 1-\epsilon & \epsilon^3 & \epsilon^3 & \cdots & \epsilon^3 \\ 0 & \epsilon^2 & \epsilon^3 & \epsilon^3 & \epsilon^3 & 1-\epsilon & \epsilon^3 & \cdots & \epsilon^3 \\ 0 & \epsilon^2 & \epsilon^3 & \epsilon^3 & \epsilon^3 & \epsilon^3 & 1-\epsilon & \cdots & \epsilon^3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \epsilon^2 & \epsilon^3 & \epsilon^3 & \epsilon^3 & \epsilon^3 & \epsilon^3 & \dots & 1-\epsilon \\ \epsilon & \epsilon^2 & \epsilon^2 & \epsilon^2 & \epsilon^2 & \epsilon^2 & \epsilon^2 & \cdots & \epsilon^2 \end{bmatrix}.$$

注意，对于任意顶点 $k \in [3, n]$，由于 $A_{i,\ j}$ 是（向量右乘时）从顶点j 到顶点 i 的概率，因此从顶点 k 出发的链几乎总是以概率 $\approx 1-\epsilon$从 k 向后移动到 $k-1$，否则就近似均匀地移动到除顶点 1 以外的任意其他顶点。因此，我们有以下结论。

注 4.21. 构造 4.17 中的 Rootn 矩阵几乎 是一个循环。 ◊

这一观察实际上进一步引出了以下具体引理。

引理 4.22. 对于 *Rootn* 矩阵 A_n，对于任意集合 $S = \{i, i+1, \dots, j\}$，

$$\phi_S(A_n) \in O\left(\frac{1}{\sqrt{n}}\right).$$

证明. 直接对构造 A 中的矩阵元素求和，即可立即看出这一点4.17。

□

引理 4.22的主要观察结果是，尽管我们选取了一个大小为 $n/2$ 的集合，比如 $\{2, ..., n/2+1\}$，但其中恰好有一个元素 d_n 贡献了约 1，它位于

补集内，其余 $O(n^2)$ 个数值贡献了约 $O(\sqrt{n})$ 的质量，流出该集合，因此流出该集合的平均质量约为$n^2\epsilon^3/(n/2) \in O(1/\sqrt{n})$。因此，如果对角线以上的元素 e_n 被零取代，那么对于大小为 $n/2$ 的集合，恰好会有一个元素d_n 贡献约 1，矩阵的边扩张便会变成 $O(1/n)$，这正是我们所希望的。注意，就扩张而言，元素 e_n 的意义较小，因为从矩阵中去掉因子 $n\cdot e_n \cdot J$，只会使非对角线块中的元素 d_n 保留下来。

这使我们朝着一个全新的方向前进，因为通过观察构造 4.17 并注意到4.21 和引理 4.22，我们可以看到，我们能够 固定矩阵的边扩张，并以某种方式设置矩阵元素，使其具有所需的特征值，这与我们在构造 4.17 Rootn 矩阵时所采取的方向相反。实现这一点的一种简单方法是使用如下矩阵（此前曾针对另一项要求讨论过）–

$$A_n = \begin{bmatrix} \mathbf{0} & 1 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{0} & 1 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{0} & 1 & 0 & 0 \\ 0 & 0 & 0 & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \mathbf{0} & 1 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{0} \end{bmatrix}$$

其边扩张为 $O(1/n)$，且其所有特征值均为 0。然而，该矩阵不是双随机的；通过置入 $A_{n,\,1} = 1$ 使其成为一个循环，这会使谱隙接近 ϕ^2，从而使其严格处于 Cheeger 区间内（见第 3.3 节）。不过，通过观察 4.17 中 Rootn 矩阵里的最重置换，即观察 4.21，我们可以首先将矩阵的 结构固定为如下形式：

$$A_n = \begin{bmatrix} \mathbf{1-r} & r & 0 & 0 & 0 & 0 \\ \cdot & \cdot & r & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & r & 0 & 0 \\ \cdot & \cdot & \cdot & \ddots & \ddots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & r \\ \cdot & \cdot & \cdot & \cdot & \cdot & \mathbf{1-r} \end{bmatrix} \tag{4.23}$$

注意，该矩阵与构造4.17中所述4.21的 Rootn 矩阵几乎完全相似，其中由 · 表示的矩阵元素尚未填入

。由于该矩阵必须是双随机的， $r \leq 1$，且无论 r的值如何，

$$\phi(A_n) \in O(1/n),$$

因为考虑集合 $S = \{1, ..., n/2\}$，

$$\phi_S(A) = \frac{r}{n/2} \leq \frac{2}{n} \in O(1/n).$$

需要注意的关键是，这与 Rootn 矩阵舒尔形式中的三角矩阵结构几乎完全相同；本质上，它与 Rootn 矩阵的结构相同，只是经过酉矩阵旋转。我们只需填充 A_n 的矩阵元素，使其保持双随机，并使第二特征值远离 1，这将得到 $\Gamma(n) \in O(1/n)$

然而，在固定这一结构后，即使我们选择一个 5×5 矩阵，要通过设置变量来获得某些期望的特征值似乎也很困难，因为特征多项式的度为 ≥ 5，而且它没有解析解。尽管解析解并非必需，但缺少解析解会使矩阵分析几乎不可能。为了解决这个问题，主要思路如下：我们不直接尝试控制矩阵的特征值，而是通过控制特征多项式的系数来间接控制特征值 。对于任意给定的矩阵 A_n，设特征多项式中 λ^{n-k} 的系数为 a_k，则

$$a_k = \frac{(-1)^k}{k!} \begin{vmatrix} \mathrm{Tr}A & 1 & 0 & 0 & 0 & 0 \\ \mathrm{Tr}A^2 & \mathrm{Tr}A & 2 & 0 & 0 & 0 \\ \mathrm{Tr}A^3 & \mathrm{Tr}A^2 & \mathrm{Tr}A & \ddots & 0 & 0 \\ \ddots & \mathrm{Tr}A^3 & \mathrm{Tr}A^2 & \mathrm{Tr}A & k-2 & 0 \\ \mathrm{Tr}A^{k-1} & \ddots & \mathrm{Tr}A^3 & \mathrm{Tr}A^2 & \mathrm{Tr}A & k-1 \\ \mathrm{Tr}A^k & \mathrm{Tr}A^{k-1} & \ddots & \mathrm{Tr}A^3 & \mathrm{Tr}A^2 & \mathrm{Tr}A \end{vmatrix}$$

其中 $|\cdot|$ 是行列式。因此，我们发现，控制矩阵 A_n 的幂的迹，有助于控制特征多项式的系数，进而控制特征值。事后看来，这并不令人意外，因为 $\mathrm{Tr}A^k = \sum \lambda_i^k$，但除非 $\sum \lambda_i^k$的取值简单，否则仅凭这一事实并没有多大用处。不过，对我们而言，在约束 1下，若要使所有特征值均为 0，只需令所有 $A^k = 1$ 均满足 $k \geq 1$，这意味着该矩阵有一个（平凡的）特征值 1，而其余特征值均为 0。因此，给定结构 4.23，我们将设定 A 的其余取值，以确保 $\mathrm{Tr}A^k = 1$。

为了具体设定这些取值，注意我们矩阵的底层图如下，其中 $A_{i,\ j}$ 是从 $j \to i$ 出发的边的权重（表示右乘向量）：图中唯一的“后向”边是从顶点 i 指向 $i-1$ 的边（对应 $i \geq 2$），其权重为 r；其他所有边都是从 i 指向 j 的“前向边”（$j \geq i$）。考虑 $\mathrm{Tr}A^k$的组合意义。它将所有从任一顶点出发并回到该顶点的长度为 k的路径（游走）的权重相加（路径的权重是边的乘积的权重）。固定任意一个顶点 i，考虑一条从 i 出发并回到该顶点、长度为 k 的路径。这条路径最远能向前走多远？注意，如果路径一步从 i 走到 $i+k$，由于一步最多只能后退一个顶点，因此这条路径不可能在 k 步内回到顶点 i。因此，从任意顶点出发，长度为 k 的路径在一步中最多能向前走 $k-1$。所以，在 $\mathrm{Tr}A^k$中，恰好 有 一条长度为k 的路径，从顶点 i 出发并回到该顶点，且向后退了 $k-1$ 步，其权重为 $r^{k-1}a_{i,\ i-k+1}$。类似地，任何向后退了 l 步的长度为 k 的路径（其中 $l \leq k-1$）的权重都是形如 $r^l \cdot c$ 的项，其中 c 是图中沿路径经过的其他前向边的乘积。因此，我们得到，对于任意 $k \geq 1$,

$$\mathrm{Tr}A^k = \sum_{i=0}^{k-1} r^i w_i(k), \tag{4.24}$$

或者，$\mathrm{Tr}A^k$ 是关于 r 的 $k-1$ 度多项式，其中 w_i（k）是其余所有非零项 A 的函数。如上所述， r^{k-1} 在$\mathrm{Tr}A^k$ 中的系数很容易求出，且恰好为

$$[r^{k-1}]\mathrm{Tr}A^k = w_{k-1}(k) = \sum_{j=k-1}^{n} a_{j,j-k+1}$$

并且尤其只包含与对角线距离为 $k-1$ 的矩阵元素。这一关键事实极其重要，我们将其明确陈述为一个将要加以利用的事实：

事实 4.25。 对于 A 的结构，如前所述（参见 *4.23*），r^{k-1} 在 TrA^k 中的系数是矩阵元素的线性函数。

观察 4.25 提示我们可以采用一种 归纳 方式来赋予前向边权重。注意，如果我们只使用方程 $\mathrm{Tr}\ k-2$，为 A 中距离对角线$A^l = 1$ 为 $1 \leq l \leq k-1$的所有边设置权重，那么方程4.24 就变为

$$\mathrm{Tr}A^k = W + r^{k-1} \cdot \sum_{j=k-1}^{n} a_{j,j-k+1}$$

并将其设为 1，得到

$$\sum_{j=k-1}^{n} a_{j,j-k+1} = \frac{1-W}{r^{k-1}}, \tag{4.26}$$

其中，W 是所有至多回溯 $k-2$ 次的路径的权重之和。

由于 $A^n = J$，我们有 $n-1$ 个独立方程，$\mathrm{Tr}A^k = 1$ 对于 $1 \le k \le n-1$。给定这些 $n-1$ 方程，如果我们的矩阵仅由 $n-1$ 个不同的变量组成，我们就可以用一个方程来确定每个变量；并且，由于我们可以利用方程4.26 （该方程来自 $\mathrm{Tr}A^k = 1$）来确定与对角线距离为 $k-1$的变量，因此它为我们提供了 $n-2$ 个不同的变量。因此，我们将距离原点 k 的所有变量设为 c_k。这是最重要也是最终的结论。

我们还需要确保矩阵保持双随机；第一列和最后一行将具有特殊值，其余数值则按上述方式设定。由此得到以下矩阵：

$$A_n = \begin{bmatrix} b_0 & r & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ b_1 & c_0 & r & 0 & 0 & 0 & 0 & \cdots & 0 \\ b_2 & c_1 & c_0 & r & 0 & 0 & 0 & \cdots & 0 \\ b_3 & c_2 & c_1 & c_0 & r & 0 & 0 & \cdots & 0 \\ b_4 & c_3 & c_2 & c_1 & c_0 & r & 0 & \cdots & 0 \\ \vdots & \vdots & c_3 & c_2 & c_1 & c_0 & r & \cdots & 0 \\ b_{n-3} & c_{n-4} & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ b_{n-2} & c_{n-3} & c_{n-4} & \cdots & c_3 & c_2 & c_1 & c_0 & r \\ b_{n-1} & b_{n-2} & b_{n-3} & \cdots & b_4 & b_3 & b_2 & b_1 & b_0 \end{bmatrix}$$

其中，对于 $0 \le i \le n-2$

$$b_i = 1 - r - \sum_{j=0}^{i-1} c_i$$

and

$$b_{n-1} = 1 - \sum_{i=0}^{n-2} b_i.$$

利用这一结构，我们可以设定 c_0 使用 $\mathrm{Tr}A = 1$，以及 c_1 使用 $\mathrm{Tr}A^2 = 1$，类似地c_{k-1} 使用 $\mathrm{Tr}A^k = 1$，并注意到 $\mathrm{Tr}A^k$ 包含 b_{k-1}，但 b_{k-1} 不依赖于 c_{k-1}。因此，利用方程4.26，我们可以将 c_k 的每个值表示为 r的函数（多项式，尽管带有负指数）。最后，神奇值 r

我们用它来计算每个 c_k，而它也是使整个构造得以成立的关键：令 r 为 $1/n$ 的 $(n-1)$ 次方根：

$$r = \left(\frac{1}{n}\right)^{\frac{1}{n-1}}$$

r 的值可以通过多种方式求得，例如使用方程 $\mathrm{Tr}A^{n-1} = 1$，该方程只含有 r 这一变量，因为所有 a_i都已确定，然后选取方程中 r 的正实根。或者，也可以取 $A^{n-1} = J$ 并考察$a_{1,\ n}$。由此得到以下构造。

4.10 Chet 矩阵——固定边扩张并最大化谱隙

构造 4.27。 对于任意 n，令 C_n 具有以下结构。

$$C_n = \begin{bmatrix} b_0 & r & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ b_1 & c_0 & r & 0 & 0 & 0 & 0 & \cdots & 0 \\ b_2 & c_1 & c_0 & r & 0 & 0 & 0 & \cdots & 0 \\ b_3 & c_2 & c_1 & c_0 & r & 0 & 0 & \cdots & 0 \\ b_4 & c_3 & c_2 & c_1 & c_0 & r & 0 & \cdots & 0 \\ \vdots & \vdots & c_3 & c_2 & c_1 & c_0 & r & \cdots & 0 \\ b_{n-3} & c_{n-4} & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ b_{n-2} & c_{n-3} & c_{n-4} & \cdots & c_3 & c_2 & c_1 & c_0 & r \\ b_{n-1} & b_{n-2} & b_{n-3} & \cdots & b_4 & b_3 & b_2 & b_1 & b_0 \end{bmatrix}$$

其中对于 $0 \le i \le n-2$,

$$b_i = 1 - r - \sum_{j=0}^{i-1} c_i$$

and

$$b_{n-1} = 1 - \sum_{i=0}^{n-2} b_i.$$

对 $k = 0$ 到 $n-3$进行归纳，利用方程 $\mathrm{Tr}(C_n)^{k+1} = 1$，将 c_k 设为关于 r 的多项式。更具体地，令 $C_n|_k$ 为具有 C_n结构的矩阵，其中所有满足 $l \ge k$ 的 c_l 均设为 0，所有满足 $l \ge k+1$ 的 b_l 均设为 0。对 $k = 0$ 到 $n-3$进行归纳，将多项式设为

$$c_k = \frac{1 - \mathrm{Tr}(C_n|_k)^{k+1}}{(n-k-2)\cdot(k+1)\cdot r^k}.$$

将 r 设为 $1/n$的第 $(n-1)$ 个根，即，

$$r = \left(\frac{1}{n}\right)^{\frac{1}{n-1}}.$$

◊

现在我们来看一些具体的情形，以便实际计算并观察矩阵 A_n 的样子。

例 4.28。对于 $n = 5$，我们得到

$$\begin{aligned} c_0 &= \frac{2}{3}r - \frac{1}{3} \\ c_1 &= \frac{5}{6}r - \frac{1}{3}\cdot\frac{1}{r} \\ c_2 &= \frac{55}{27}r + \frac{2}{9}\cdot\frac{1}{r} - \frac{20}{27}\cdot\frac{1}{r^2} \\ r &= 5^{-1/4} \end{aligned}$$

将矩阵近似到小数点后 5 位，结果如下：

$$C_5 \approx \begin{bmatrix} 0.33126 & 0.66874 & & & \\ 0.21870 & 0.11249 & 0.66874 & & \\ 0.15993 & 0.05886 & 0.11249 & 0.66874 & \\ 0.12170 & 0.03820 & 0.05886 & 0.11249 & 0.66874 \\ 0.16810 & 0.12170 & 0.15993 & 0.21870 & 0.33126 \end{bmatrix}$$

◊

同样，我们还可以写出另一个例子。

例4.29。对于 $n = 8$，我们得到

$$
\begin{aligned}
c_0 &= \frac{1}{6} \cdot (2r - 1) \\
c_1 &= \frac{1}{5r} \cdot \left(\frac{4}{3} r^2 - \frac{7}{12} \right) \\
c_2 &= \frac{1}{4r^2} \cdot \left(\frac{152}{135} r^3 + \frac{7}{45} r - \frac{14}{27} \right) \\
c_3 &= \frac{1}{3r^3} \cdot \left(\frac{778}{675} r^4 + \frac{49}{900} r^2 + \frac{7}{27} r - \frac{1981}{3600} \right) \\
c_4 &= \frac{1}{2r^4} \cdot \left(\frac{112}{75} r^5 + \frac{133}{4050} r^3 + \frac{7}{135} r^2 + \frac{1981}{5400} r - \frac{1043}{1620} \right) \\
c_5 &= \frac{1}{r^5} \cdot \left(\frac{27056}{10125} r^6 + \frac{2863}{121500} r^4 + \frac{28}{1215} r^3 + \frac{1981}{81000} r^2 + \frac{1043}{2430} r - \frac{193417}{243000} \right) \\
r &= 8^{-1/7}
\end{aligned}
$$

而矩阵取近似值并保留 5 位小数后如下：

$$
C_8 \approx \begin{bmatrix}
0.25700 & 0.74300 & 0.17601 & 0 & \dots & 0.08100 & 0.74300 & 0.13489 & 0 & 0.04112 & 0 & \dots & 0.08100 & 0.74300 & 0.10823 & 0 & 0.02666 & 0 & 0.04112 \\
0 & \dots & 0.08100 & 0 & 0.74300 & 0.08901 & 0 & 0.01922 & 0 & 0.02666 & 0 & 0.04112 & 0 & \dots & 0.08100 & 0 & 0.74300 & 0.07431 & 0 & 0.01470 & 0 & 0.01922 & 0 & 0.02666 & 0 & 0.04112 & 0 & \dots & 0.08100 & 0 & 0.74300 & 0.06261 & 0 & 0.01170 & 0 & 0.01470 & 0 & 0.01922 & 0 & 0.02666 & 0 & 0.04112 & 0 & \dots & 0.08100 & 0 & 0.74300 & 0.09800 & 0 & 0.06261 & 0 & 0.07431 & 0 & 0.08901 & 0 & 0.10823 & 0 & 0.13489 & 0 & 0.17601 & 0 & \dots & 0.25700
\end{bmatrix}
$$

◊

我们给出 $n = 16$在附录A.1中的方程

给定构造4.27，我们有如下定理。

定理4.30。令 C_n 为由算法4.27生成的矩阵。那么，对于 C_n，以下结论成立。

1.（猜想） C_n 的每个元素都非负。*2.* C_n 的所有非平凡特征值均为 *0*。*3.* $\phi(C_n) \leq 2/n \in O(1/n)$

4. 由 *2* 和 *3* 可得

$$\Gamma(n) \in O\left(\frac{1}{n}\right)$$

改进引理中的界*4.19*并结合定理*1.2*，可得

$$\Gamma(n) \in \Theta\left(\frac{1}{n}\right)$$

并且定理 $\phi(R)$ 中的下界 *1.2* 在常数因子范围内是紧的。

证明。引理的第 (1) 部分是一个猜想，我们在其余部分中假设该猜想成立。对于 (3)，由于 C_n 由 (1) 可知是非负的，$\phi(C_n)$ 有定义，并且考察$S = \{1, ..., n/2\}$，可得 $\phi_S(C_n) \leq r/2 \leq n \in O(1/n)$。对于 (2)，由于 A_n 的每一行和每一列之和均为 1，它有一个对应于全 1 特征向量的特征值 1；进一步，由于对于所有 $k \geq 1$，$\mathrm{Tr}C_n^k = \sum_{i=1}^n \lambda_i^k(C_n) = 1 + \sum_{i=2}^n \lambda_i^k(C_n) = 1$，可知 C_n 除 1 以外的所有特征值均为 0。 □

对于这一构造，唯一尚不确定的是如何证明 c_i和 b_i 在构造4.27 中均为非负。我们目前还不知道如何证明这一点，这将是本学位论文的主要猜想，并将在第4.12节陈述和讨论。不过，我们先尝试列出 Chet 矩阵的一些性质。

4.11 Chet 矩阵的观察与性质

我们认为 Chet 矩阵足够优美，值得单独研究，并将探讨它们的许多性质。

1. ***(Hessenberg-Toeplitz*** 结构*)* Chet 矩阵具有 Hessenberg-Toeplitz 结构。海森堡矩阵 $H_{k,\ l}$ 是指所有距对角线超过 k 的下方矩阵元素以及距对角线超过 l 的上方矩阵元素均为零的矩阵。此外，该矩阵具有类 Toeplitz 结构，除第一列和最后一行外，每条对角线上的矩阵元素都相同。

2. （积和式）设 C'_n 为一个矩阵，它包含 C_n 的所有矩阵元素，但将对角线上方的 $n-1$ 个 r 矩阵元素替换为 $-r$。那么

$$\mathrm{permanent}(C_n) = \mathrm{determinant}(C'_n)$$

证明可由直接归纳法得出。

3. （近似值）通过数值和解析观察， C_n 的矩阵元素近似值如下：

$$r \approx 1 - \frac{\log n}{n-1}.$$

对于 C_n的矩阵元素，

$$c_i \in O\left(\frac{1}{i \cdot n}\right).$$

参见附录 A.3 ，其中列出了上述内容关于 $n = 500$的表格。因此，请注意

$$\sum_{i=0}^{O(\log n)} c_i \in O\left(\frac{\log\log n}{n}\right),$$

因此，对于 $i \in O(\log\ n)$,

$$b_i = 1 - r - \sum_{j=0}^{i} c_i \approx \frac{\log n - \log\log n}{n} > 0.$$

可以直接证明一个较弱的结论。

引理 4.31。 对于任意 *Chet* 矩阵 C_n，若 $c_i \geq 0$ 且对于 $i \in O(n/\ln\ n)$有 $b_i \geq 0$ ，则对于 $i \in O(n/\ln\ n)$,

$$c_i \in O\left(\frac{1}{i \cdot (n-i)}\right).$$

证明。 注意到，由于 c_i 是通过使用 $\mathrm{Tr}(C_n^{i+1}) = 1$来定义的，因此可得

$$1 = \mathrm{Tr}C_n^{i+1} = S + r^i \cdot c_i \cdot i \cdot (n-i)$$

其中， S 是所有不经过 c_i的其他长度为 $i+1$ 的路径的加权和。由于$S \geq 0$ ，根据 b_i， $c_i \geq 0$的假设，并且

$$r^i = n^{-\frac{i}{n-1}} \geq \exp\left(-\frac{\ln n}{n-1} \cdot i\right) \in \Omega(1)$$

由于 $i \in O$（n/ln n），结论立即成立。 □

4. （解析值）对于任意 n，前三个矩阵元素均可通过解析方法计算如下。令 $r = n^{-1/(n-1)}$。

$$c_0(n) = \frac{2 \cdot r - 1}{n-2}$$

$$c_1(n) = \frac{(2r^2 - 1) \cdot n + 1}{2(n-3)(n-2) \cdot r}$$

$$c_2(n) = \frac{(2r^3 - 1) \cdot n^3 + (-8r^3 + 3r + 4) \cdot n^2 + (12r^3 - 15r - 3) \cdot n + 12r}{3(n-4)(n-3)(n-2)^2 r^2}$$

可以证明，这些值始终为非负。

5. （不同的矩阵元素）这种构造有个有趣之处：矩阵中有Θ（n）个不同的矩阵元素，而构造4.17 中的 Rootn 矩阵有 6 个不同的矩阵元素，de Bruijn 矩阵则只有 2 个。注意，从构造4.17中矩阵的结构出发，并再次使用方程 Tr（A^k） = 1，令 $k = 1$ 取值从到 6（因为有 6 个不同的变量），我们得到的变量值与通过矩阵的舒尔变换得到的相同，因此对 Rootn 矩阵完全避免了舒尔分解。

6. （构造的历程及无限多个类似构造）回顾抵达 Chet 矩阵构造4.27 的历程，着实不同寻常。我们最初想要理解所有特征值均为 0 的矩阵的边扩张，因为它已相当接近最优（见第4.5节）；最终则通过在双随机矩阵的舒尔分解中仔细选择 T 和 U，得到了 Rootn 矩阵（构造4.17）。该构造让我们了解到一种值得考虑的矩阵结构；基于这一结构及后续推理，我们将 Tr$A^k = 1$ 作为主要条件，并通过对角线设置矩阵中的数值，从而得到 Chet 矩阵。此时，我们可以从任意一个边扩张较小的矩阵出发，使用相同的算法4.27 ，利用 Tr $A^k = 1$来设置距离对角线 k 处的数值。此外，只需选取某个特征值集合（例如$1\ 2$、$1\ 4$、……、 $\frac{1}{2^{n-1}}$），并使用相应的迹不等式，我们就能构造出无限多个具有 Chet 矩阵结构、但 r 以及 b_i和 c_i取值不同的矩阵。

7. （由近似方程得到的矩阵）注意到，由于很难理解矩阵元素的多项式系数如何变化， c_i 人们很容易想尝试对构造4.27中的 Chet 矩阵元素进行近似，看看近似是否有效。然而，结果表明，正如预期的那样，矩阵的特征值极其敏感，谱隙会随着近似迅速缩小至零。另一种想法是，不去近似矩阵元素，而是直接近似方程

本身，即令 $\mathrm{Tr}A^k \approx 1$。然而，很容易证明这样行不通。

引理 4.32. 存在双随机矩阵 A，满足 $TrA^k \approx 1$，但 $Re\lambda_2(A) \approx 1$。

证明。 考虑图的邻接矩阵 C。该图由 1 个孤立顶点和一个长度为 $n-1$ 的有向环组成。注意， C 是双随机且不连通的。 C 的特征值为 $1, \omega_i$，其中 $i=0$ 到 $n-2$，且 $\omega_i = e^{\frac{2\pi \cdot i}{n-1}}$。令 $A = \alpha J + (1-\alpha)C$。注意， A 的特征值为 $1, (1-\alpha)\omega_i$，其中 $0 \le i \le n-2$，因为对应于第一个特征值 1 的特征向量是全一向量，而 J 是其上的投影。另一个等于 1 的特征值对应于 ω_0，它位于与全一向量正交的空间中，并且仅缩小 $(1-\alpha)$ 倍。令 $\alpha = {}^{2\log}{}_{n-1}{}^{n}$，则 A 的第二特征值为 $\lambda_2 = 1-\alpha = 1 - {}^{2}{}_{\log_{n-1}}{}^{n} \approx 1$。

——　　——

我们有，对于任意 $k \le n-2$，$\mathrm{Tr}A^k = 1 + (1-\alpha)^k \sum_{i=0}^{n-2} \omega_i^k = 1$，因为 $\sum_{i=0}^{n-2} \omega_i^k = 0$ 对于 $k \le n-2$ 成立。对于 $k = n-1$，我们有

$$\begin{aligned}
\mathrm{Tr}A^{n-1} &= 1 + (1-\alpha)^{n-1} \sum_{i=0}^{n-2} \omega_i^{n-1} \\
&= 1 + (n-1)(1-\alpha)^{n-1} \\
&\le 1 + (n-1)\exp(-2\log n) \\
&\le 1 + \frac{1}{n} \\
&\approx 1
\end{aligned}$$

由此完成证明。注意，这类方程或许有助于构造具有 $\Gamma(n) \approx {}^{1}\log\ n$ 的矩阵，但无助于得到 $\Gamma(n) \in o\left({}^{1}\log\ n\right)$。　　——　　□

8. （若尔当标准形）现在我们尝试理解 Chet 矩阵的若尔当标准形。海森堡矩阵的若尔当标准形已广为人知（例如参见 [Zem06]），并且可以利用预解式 $(t \cdot I - A)^{-1}$轻松推导出来，因此这里只给出最终形式。假设 $A = VQV^{-1}$ 是A的若尔当分解，其中 Q 包含两个块：一个长度为 1、对应特征值 1 的块，以及另一个长度为 $n-1$、对应特征值 0 的块。具体而言， Q 除$Q_{1,\ 1} = 1$和$Q_{i,\ i+1} = 1$ （$i = 2$ 至 $n-1$）之外，其余元素全为零。 V 的列是最有意思的部分，可以描述如下。设 $A_{[1,j]}$ 表示 A 的子矩阵，它由指标为 {1，2，..... j}的所有行和列组成，其中 $A_{[1,0]} = 1$。令 V的列

为 V_i（$1 \le i \le n$）。注意，由于 A 是双随机矩阵，第一列 V_1 只是全一向量的一个倍数。来看第二列。第二列中，行号从 $i = 1$ 到 n的矩阵元素为

$$V_2(i) = \frac{\det(t \cdot I - A_{[1,i-1]})}{r^{i-1}}\bigg|_{t=0}$$

（注意，I 与 $A_{[1,i-1]}$的维度相同，因此对于 $i = 1$，分子就是 -1）。同理，第三列 V_3可以描述如下，

$$V_3(i) = \frac{1}{r^{i-1}} \cdot \frac{d}{dt}(\det(t \cdot I - A_{[1,i-1]}))\bigg|_{t=0}$$

一般而言，对于 $2 \le j \le n$,

$$V_j(i) = \frac{1}{(j-2)! \cdot r^{i-1}} \cdot \frac{d^{(j-2)}}{dt^{(j-2)}}(\det(t \cdot I - A_{[1,i-1]}))\bigg|_{t=0}.$$

因此请注意，甚至若尔当标准形也具有Hessenberg结构！这显然是因为 $\det(t \cdot I - A_{[1,i-1]})$ 关于 t的度为 $i - 1$，而当对其求导超过 $i - 1$ 次时，该项就变为零。我们认为这一性质很优美，但目前还不知道如何利用它来证明变量的非负性。

9.（生成函数）现在，我们用生成函数表示矩阵的所有矩阵元素，以便将其与底层链的组合学联系起来，并得到某些量的闭式表达式。由于该矩阵除平凡特征值外的所有特征值均为 0，令 $\chi_t(C_n)$为 C_n的特征多项式，则 $\chi_t(C_n) = t^{n-1} \cdot (t - 1)$，因此，除 t^n $(= 1)$和 t^{n-1} $(= -1)$的系数外，其余系数均为 0。然而，我们知道$\chi_t(C_n) = \det(t \cdot I - C_n)$，并且由于 $t \cdot I - C_n$ 也具有Hessenberg结构（见第（1）点），令 C_n' 表示将 C_n 中矩阵元素 r 替换为 $-r$ 后得到的矩阵，由第（2）点可得 $\operatorname{perm}(t \cdot I - C_n') = \det(t \cdot I - C_n)$。因此，我们有

$$\operatorname{perm}(t \cdot I - C_n') = t^{n-1} \cdot (t - 1). \tag{4.33}$$

积和式的重要之处在于，相较于行列式，其背后的组合学更适合用生成函数处理。积和式是矩阵所有环覆盖的权重之和，比起行列式所涉及的长度为 k 的加权圈数量（对所有 k而言），处理起来简单得多。

我们令 $Z_n = t \cdot I - C'_n$，并计算 Z中的环覆盖。令 $Z_{1,1} = Z_{n,n} = q_0 = t - b_0$, $Z_{i,i} = a_0 = t - c_0$ 对应于 $2 \le i \le n-1$，并令 $q_i = -b_i$ 和 $a_i = -c_i$分别对应于 $i \ge 2$ 在 Z中的相应矩阵元素。注意，Z 的对角线上方有 r，因为 C'_n 有 $-r$。要计算 Z_n中的环覆盖，注意到由于矩阵的结构，从顶点 i出发，如果一个循环沿着反方向走 k 步到达顶点 $i-k$，那么其中一个循环必定是长度为 $k+1$的循环 $\{i, i-k, i-k+1, i-k+2, \ldots, i-1, i\}$，因为这是该循环能够成为环覆盖一部分的唯一方式。注意，该循环的权重为 $c \cdot r^k$，其中 c 表示从顶点 i 到顶点 $i-k$的边的权重。因此，任何特定的循环都会将图分成两部分，于是我们可以定义以下两个多项式：

$$R(x) = \sum_{i \ge 0} a_i r^i x^{i+1}$$
$$G(x) = \sum_{i \ge 0} q_i r^i x^{i+1}$$

其中假设 $a_l = 0$ 对于 $l \ge n-1$ 成立，且 $q_l = 0$ 对于 $l \ge n$成立。注意，R 和 G 没有常数项，并且由于选择 b_i是为了确保 C_n 的第一列和最后一行的和为 1，因此我们有

$$\begin{aligned} c_i &= b_i - b_{i+1} \\ a_i &= q_i - q_{i+1} \\ R &= G - \frac{G - q_0 x}{rx} \\ G - R &= \frac{G'}{rx} \end{aligned} \tag{4.34}$$

其中 $G' = G - q_0 x$。这个等式稍后会用到。

$k \ge 1$的两个多项式中 x^{k+1} 的系数为

$$[x^k]R(x) = a_k r^k$$
$$[x^k]G(x) = q_k r^k.$$

第一个系数恰好是一个长度为 $k+1$ 的循环的长度，该循环从除第一个和最后一个顶点之外的任意顶点开始；第二个系数则是一个长度为 $k+1$ 的循环的长度，该循环从第一个或最后一个顶点开始。因此，由于 Z 的积和式是我们图中所有长度为 n 的环覆盖的权重之和，我们得到

$$\begin{aligned} H(x) &= G(x)^2 \sum_{i \ge 0} R(x)^i + G(x) \\ &= \frac{G(x)^2}{1 - R(x)} + G(x) \end{aligned}$$

其中，第一项通过固定环覆盖，并考虑包含第一个顶点的循环、包含最后一个顶点的循环，以及将剩余顶点划分为 i 个循环的不同方式的数量，来统计所有环覆盖的权重；第二项则统计所有顶点都出现在同一个循环中的那个环覆盖。注意， $H(x)$ 中 x^k 的系数是关于 t 的多项式。因此，我们有

$$\begin{aligned} t^{n-1} \cdot (t-1) = \operatorname{perm}(Z) &= [x^n] H(x) \\ &= [x^n] \frac{G(x)^2}{1-R(x)} - b_n r^{n-1} \end{aligned}$$

因此，为了比较上述方程两边的系数，我们得到以下 $n+1$ 方程：

$$\begin{aligned} 1 &= \frac{1}{n!} \cdot \frac{d^n}{dt^n} [x^n] \frac{G(x)^2}{1-R(x)} \Big|_{t=0} \\ -1 &= \frac{1}{(n-1)!} \cdot \frac{d^{n-1}}{dt^{n-1}} [x^n] \frac{G(x)^2}{1-R(x)} \Big|_{t=0} \\ 0 &= \frac{1}{k!} \cdot \frac{d^k}{dt^k} [x^n] \frac{G(x)^2}{1-R(x)} \Big|_{t=0} \\ 0 &= [x^n] \frac{G(x)^2}{1-R(x)} \Big|_{t=0} - b_n r^{n-1} \end{aligned} \tag{4.35}$$

其中，倒数第二个方程对于 $1 \le k \le n-2$成立。最后一个方程是冗余的，因为它表示 $0 = \det(t \cdot I - C_n)|_{t=0}$ ，根据我们在 C_n中设置 c_i和 b_i的值的方式，我们知道它为真。要理解其他方程，请注意， $G(x)$中只有一个项/系数依赖于 t，即 $q_0 x = (t - b_0)x$ ；同样，对于 $R(x)$，项 $a_0 x = (t - c_0)x$依赖于相应变量，因此

$$\begin{aligned} \frac{d}{dt} G(x)^2 &= 2G(x) \cdot x \\ \frac{d^2}{dt^2} G(x)^2 &= 2x^2 \\ \frac{d^i}{dt^i} G(x)^2 &= 0 \end{aligned}$$

对于 $i \ge 3$，类似地，我们可以写成：

$$\begin{aligned} \frac{d}{dt} (1-R(x))^{-1} &= (1-R(x))^{-2} x \\ \frac{d^i}{dt^i} (1-R(x))^{-1} &= i! \cdot x^i \cdot (1-R(x))^{-(i+1)} \end{aligned}$$

对所有 i均成立。因此我们得
到

$$\begin{aligned}\frac{d^k}{dt^k}\frac{G(x)^2}{1-R(x)} &= \sum_{i=0}^{k}\binom{k}{i}\frac{d^i}{dt^i}(1-R)^{-1}\frac{d^{k-i}}{dt^{k-i}}G^2 \\ &= k!x^k(1-R)^{-k+1}G^2 + \\ &\quad k(k-1)!x^{k-1}(1-R)^{-k}\cdot 2\cdot G\cdot x + \\ &\quad \binom{k}{2}(k-2)!x^{k-2}(1-R)^{-(k-1)}2x^2 \\ &= k!\cdot x^k\cdot\left(G^2(1-R)^{-(k+1)} + 2G(1-R)^{-k} + (1-R)^{-(k-1)}\right) \\ &= k!\cdot x^k\cdot(1-R)^{-(k+1)}\cdot(G+1-R)^2.\end{aligned}$$

因此，由方程4.35可得

$$\begin{aligned}\frac{1}{k!}\cdot\frac{d^k}{dt^k}[x^n]\frac{G(x)^2}{1-R(x)} &= \frac{1}{k!}\cdot[x^n]\frac{d^k}{dt^k}\frac{G(x)^2}{1-R(x)} \\ &= [x^{n-k}](1-R)^{-(k+1)}\cdot(G+1-R)^2 \\ &= \sum_{j=0}^{n-k}[x^j](1-R)^{-(k+1)}\cdot[x^{n-k-j}]\,(G+1-R)^2\,. \qquad (4.36)\end{aligned}$$

为了计算上述等式中的各项，我们可以写成

$$\begin{aligned}[x^j](1-R)^{-(k+1)} &= [x^j]\left(\sum_{i=0}\binom{k+i}{k}R^i\right) \\ &= \sum_{i=0}\binom{k+i}{k}[x^j]R^i \\ &= \sum_{i=0}^{j}\binom{k+i}{i}[x^j]R^i \\ &\quad [\text{since deg}_x(\text{R})\geq 1] \\ &= [x^j]\sum_{i=0}^{j}\binom{k+i}{i}R^i\end{aligned}$$

注意，前几项如下所示：

$$
\begin{aligned}
[x^0](1-R)^{-(k+1)} &= 1 \\
[x^1](1-R)^{-(k+1)} &= (k+1)\cdot a_0 \\
[x^2](1-R)^{-(k+1)} &= \binom{k+1}{1}\cdot ra_1 + \binom{k+2}{2}\cdot a_0^2 \\
[x^3](1-R)^{-(k+1)} &= \binom{k+1}{1}\cdot r^2a_2 + \binom{k+2}{2}\cdot r\cdot 2a_0a_1 + \binom{k+3}{3}\cdot a_0^3 \\
[x^4](1-R)^{-(k+1)} &= \binom{k+1}{1}\cdot r^3a_3 + \binom{k+2}{2}\cdot r^2\cdot\left(2a_0a_2+a_1^2\right) \\
&\quad + \binom{k+3}{3}\cdot r\cdot\left(3a_1a_0^2\right) + \binom{k+4}{4}\cdot a_0^4
\end{aligned}
$$

同样，对于4.36中的第二项，系数要容易得多。

$$
\begin{aligned}
[x^0]\,(G+1-R)^2 &= 1 \\
[x^{n-k-j}]\,(G+1-R)^2 &= [x^{n-k-j}]\left(\frac{G'}{rx}+1\right)^2 \\
&\qquad [\text{from } 4.34] \\
&= r^{n-k-j-2}\left(2\cdot q_{n-k-j}\cdot r + \sum_{i=1}^{n-k-j-1} q_iq_{n-k-j-i}\right)
\end{aligned}
$$

A因此，方程4.35 的最终表达式变为如下形式 g

$$
\begin{aligned}
0 &= \frac{1}{k!}\cdot\frac{d^k}{dt^k}[x^n]\frac{G(x)^2}{1-R(x)}\bigg|_{t=0} \\
&= \sum_{j=0}^{n-k}[x^j](1-R)^{-(k+1)}\cdot[x^{n-k-j}]\,(G+1-R)^2 \\
&= \sum_{j=0}^{n-k} r^{n-k-j-2}\left(2\cdot q_{n-k-j}\cdot r + \sum_{i_1=1}^{n-k-j-1} q_{i_1}q_{n-k-j-i_1}\right)\cdot[x^j]\sum_{i_2=0}^{j}\binom{k+i_2}{i_2}R^{i_2}
\end{aligned}
$$

这为特征多项式4.33的系数给出了一个闭式表达式，并可用于推断构造 c_i 4.27中 Chet 矩阵的各项取值；但目前我们还不知道如何利用这些方程证明变量的非负性。现在，我们正式陈述一些可能适用于这些方程的猜想。

4.12 证明 Chet 矩阵非负性的猜想

我们在本节陈述本学位论文的主要猜想，以及一系列相关猜想；这些猜想各自都很有趣。

猜想 4.37。 *(Chet 猜想)* 设 C_n 表示 $n \times n$ *Chet* 矩阵，并令$C = \{n : \ C_n$ 逐项非负$\}$。则以下结论成立：

$$|C| = \infty.$$

更强地说，

$$C = \mathbb{N}.$$

这个更强的猜想可能更容易证明或证伪。

数值模拟。 我们对 Chet 矩阵进行了精确模拟 C_n，精度一直达到 $n = 21$ 位，并在 Maple 中进行计算，所有矩阵均为非负；我们还使用附录 $n = 500$ 中的伪代码，在 Matlab 中以 100 位精度对矩阵进行了数值模拟，范围达到 A.2。结果不仅表明这些矩阵均为非负，而且矩阵元素 c_i 与 0 相距甚远，并且随着 i 增加而平缓衰减。对于 $n = 500$，最小矩阵元素中的第一个非零数字出现在小数点后 6 位；相关图表见附录A.3。图表显示，数值 c_i 和 b_i 变化非常平滑，随着 i 增加，其数值没有突然跳变或不连续的情况，这有力地支持了人们凭直觉认为 Chet 猜想成立的观点。

我们提出一些有助于证明 Chet 猜想的思路、进展和其他猜想。

迹猜想。 我们提出一些有趣的猜想，若这些猜想成立，将有助于证明 Chet 矩阵 C_n的非负性（虽然并非通过直接推论，而是通过加深一般性理解）。这些猜想如下。

猜想 4.38。 （迹猜想）设 A 是一个次随机非负矩阵，也就是说，$\sum_i A_{i,\ j} \leq 1$且 $\sum_j A_{i,\ j} \leq 1$对所有 j和 i均成立。假设如下：在对角线上方， A 只有与对角线距离为 1 的位置上才有非零项；在对角线下方， A 只有与对角线距离至多为 k 的位置上才有非零项，其中对角线到自身的距离为零。假设对于 $l \leq k+1$，$TrA^l \leq 1$，则对于所有 l，$TrA^l \leq 1$。

显然，迹猜想4.38 对称矩阵时成立，条件是$k \geq 1$，因为对称矩阵的所有特征值均为实数，且

$$\mathrm{Tr}A^2 = \sum \lambda_i^2 \leq 1$$

意味着 $|\lambda_i| \le 1$，且对于任意 l，

$$\mathrm{Tr}A^l = \sum \lambda_i^l \le \sum_i |\lambda_i|^l \le \sum_i |\lambda_i|^2 \le 1.$$

注意，如果猜想成立，就意味着总能在距离对角线 $k+1$ 处的矩阵元素中填入非负值，并确保 $\mathrm{Tr}A^{k+2} = 1$；进而，可以逐步用非负元素填满矩阵。由于 Chet 矩阵中存在首行和末列，它们是确保矩阵 C_n 保持双随机所必需的，因此迹猜想并不能直接推出 Chet 猜想；对于任何固定的 k，矩阵元素 $C_n[k+1,1]$ 和 $C_n[n,n-(k+1)]$始终为 ≥ 0，而在迹猜想中它们则为零。尽管如此，我们认为，尝试证明迹猜想将有助于证明 Chet 猜想。

迹猜想4.38 本身就很有趣，而且由于我们尚不知其证明，下面的猜想应该相对容易证明。

猜想 4.39. *(托普利茨迹猜想)* 设 A 为一个次随机非负矩阵，即 $\sum_i A_{i,j} \le 1$ 且 $\sum_j A_{i,j} \le 1$ 对所有 j 和 i 均成立。假设如下：在对角线上方，A 只有距离对角线为 1 的矩阵元素非零；在对角线下方，A 只有距离对角线至多为 k 的矩阵元素非零，其中距离对角线下方 l 处的元素为 c_l，而对角线到自身的距离为零。假设对 $1 \le l \le k+1$，$TrA^l \le 1$，则对所有 l，$TrA^l \le 1$。

托普利茨迹猜想的无限版本也很有意思，截至本文撰写之时，它仍未解决。

猜想 4.40. *(无限托普利茨迹猜想)* 考虑定义在 $\mathbb{N}$ 上的以下马尔可夫链。从每个顶点 i 出发，都有一条权重为 r、通往 $i-1$ 的边，以及权重为 c_l、通往顶点 $i+l$（其中 $0 \le l \le k$）的边，并假设 $r + \sum_{i=0}^{k} c_i \le 1$。设 p_l 为从 i 出发并在恰好 l 步后返回 i 的概率，且对于某个常数 α，当 $1 \le l \le k+1$ 时，$p_l \le \alpha/k$。那么对所有 l，$p_l \le \alpha/k$。

理解海森堡矩阵或海森堡–托普利茨矩阵的特征值的行为也很有意思，这将进一步帮助我们理解迹猜想4.38、4.39 和4.40，并由此引出以下猜想。

猜想 4.41. （托普利茨-黑森堡矩阵的特征值）设 A 是一个次随机非负矩阵，也就是说，$\sum_i A_{i,\ j} \leq 1$ 且 $\sum_j A_{i,\ j} \leq 1$ 对所有 j 和 i 都成立。假设如下：在对角线以上，A 只有距离对角线为 1 的位置上才有非零项；在对角线以下，A 只有距离对角线至多为 k 的位置上才有非零项，其中对角线到其自身的距离为零。那么，A 的特征值都包含在复平面上 $k+1$ 个单位根的凸包内。</*output*>

还可以构造许多类似的猜想，对其中任何一个的深入研究都可能有助于证明我们最初的 Chet 猜想 4.37。

第 5 章

联系与扩展

> 然而，随着更多技艺被发明出来，其中一些用于满足生活所需，另一些则用于消遣，人们自然总是认为后者的发明者比前者更有智慧，因为他们所掌握的知识并非以实用为目的。因此，等到这类发明都已确立之后，那些既不以带来愉悦、也不以满足生活所需为目的的科学才得以发现；而且最先出现在人们最早开始拥有闲暇的地方。这就是数学技艺在埃及创立的原因；在那里，祭司阶层获准享有闲暇。
>
> ~ 亚里士多德，《形而上学》

本章将研究一些其他组合量，并探讨它们与边扩张和谱隙之间的联系。对于非可逆链，这些联系中的许多也尚未得到充分研究（但也有一些例外，尤其是 Mihail 关于混合时间与边扩张之间联系的结果）；我们的目标是推导新的结果，或为已知结果提供初等证明。此外，本节的另一个目标是连贯地梳理一些分散在不同领域、不同文献中的主题，并补上缺失的引理。我们将从研究最受关注的组合量——混合时间——开始（这意味着我们能做的贡献不多）。

5.1 混合时间——定义与预备引理

为了说明一般非负矩阵的混合时间定义，我们首先考虑双随机矩阵的混合时间。双随机矩阵 A（即其对应的马尔可夫链）的混合时间，是从任意顶点出发的随机游走达到在各顶点上近似均匀的分布所需的最坏情况下的步数。为避免周期链带来的复杂性，我们假设 A 是1/2-惰性的，也就是说，对每个 i，$A_{i,i} \geq \frac{1}{2}$。给定任意双随机矩阵 A，很容易将其转换为惰性随机游走 $\frac{1}{2}I + \frac{1}{2}A$。转换后仍为双随机矩阵，且 ϕ（A）和谱隙均减半。只要该链是连通的，混合时间就是有限的。考虑任意顶点 i对应的指示向量 $\mathbf{1}_{\{i\}}$。我们要找出满足 $A^\tau \mathbf{1}_{\{i\}} \approx \frac{1}{n}\mathbf{1}$ 或 $A^\tau \mathbf{1}_{\{i\}} - \frac{1}{n}\mathbf{1} \approx 0$ 的最小 τ，这还可以进一步写成 $\left(A^\tau - \frac{1}{n}\mathbf{1}\cdot\mathbf{1}^T\right)\mathbf{1}_{\{i\}} \approx 0$。具体而言，对任意 ϵ，我们要找出 $\tau = \tau_\epsilon$（A），使得对任意 i，

$$\left\| \left(A^\tau - \frac{1}{n}\mathbf{1}\cdot\mathbf{1}^T\right)\mathbf{1}_{\{i\}} \right\|_1 \leq \epsilon.$$

给定这样的 τ 值，对于任意满足 $\|x\|_1 = 1$ 的向量 x，我们得到

$$\begin{aligned}
\left\| \left(A^\tau - \frac{1}{n}\mathbf{1}\cdot\mathbf{1}^T\right)x \right\|_1 &= \left\| \sum_i \left(A^\tau - \frac{1}{n}\mathbf{1}\cdot\mathbf{1}^T\right)x_i\mathbf{1}_{\{i\}} \right\|_1 \\
&\leq \sum_i |x_i| \left\| \left(A^\tau - \frac{1}{n}\mathbf{1}\cdot\mathbf{1}^T\right)\mathbf{1}_{\{i\}} \right\|_1 \\
&\leq \sum_i |x_i| \cdot \epsilon \\
&= \epsilon.
\end{aligned}$$

因此，混合时间 $\tau_\epsilon(A)$ 是满足以下条件的数 τ：对于任意满足 $\|x\|_1 = 1$ 的 x，均有 $\|(A^\tau - J)\cdot x\|_1 \leq \epsilon$。

我们希望将此定义推广到任意非负矩阵 R，其佩龙–弗罗贝尼乌斯特征值为 1，且对应的正左、右特征向量为 u 和 v。在研究边扩张和谱隙时，将最大特征值固定为 1 并不重要，因为这两个量在标量乘法下均保持不变；但在研究混合时间时，这样做就很重要，可以避免在每个方程中都出现最大特征值。注意，若 R 可约，则混合时间为无穷大。此外，如果R 是周期的，那么混合时间也仍未定义。因此，我们再次假设R 不可约且1/2-惰性，即对于每个 i，$R_{i,i} \geq \frac{1}{2}$。为便于说明，设 x 为任意非负向量，尽管我们最终给出的定义并不要求

非负性，并且适用于任意 x。我们希望找到 τ，使得 $R^\tau x$ 与 x 沿着 v方向的分量大致相同。此外，由于我们进行的是右乘，并且希望收敛到右特征向量 v，因此将使用左特征向量 u来定义l_1-范数。因此，对于初始向量 x，我们不再像双随机情形那样要求 $\|x\|_1 = 1$，而是要求 $\|D_u x\|_1 = 1$。由于 x非负，$\|D_u x\|_1 = \langle u, x\rangle = 1$。因此，我们希望找到 τ，使得$R^\tau x \approx v$，或 $\left(R^\tau - v\cdot\ u^T\right) x \approx 0$。由于我们是相对于 u测量初始向量x 的范数，因此也将相对于 u测量最终向量 $\left(R^\tau - v\cdot\ u^T\right) x$的范数。由此得到以下定义。

定义 5.1。（一般非负矩阵的混合时间 R）设 R 是一个 $_1$ $_2$-惰性、不可约_非负矩阵，其佩龙–弗罗贝尼乌斯特征值为 $\mathbf{1}$，u 和 v 是相应的正左、右特征向量，其中 u 和 v 经过归一化，使得 $\langle u, v\rangle = \|D_u v\|_1 = 1$。那么，混合时间 $\tau_\epsilon(R)$ 是满足以下条件的最小数 τ：对于每个满足 $\|D_u x\|_1 = 1$ 的向量 x，都有 $\left\|D_u\left(R^\tau - v\cdot u^T\right) x\right\|_1 \le \epsilon$。◊

我们指出，与双随机情形类似，利用三角不等式，只需确定标准基向量 $\mathbf{1}_{\{i\}}$ 的混合时间即可。我们将其表述为一个初等引理。

引理 5.2。设 R 如定义 *5.1*所述。那么，混合时间 $\tau_\epsilon(R)$ 是满足以下条件的最小数 τ：$\left\|D_u\left(R^\tau - v\cdot u^T\right) y\right\|_1 \le \epsilon$对每个向量 $y = \frac{\mathbf{1}_{\{i\}}}{\|D_u \mathbf{1}_{\{i\}}\|_1}$，其中 $\mathbf{1}_{\{i\}}$ ~~表~~示标准基向量。

证明。设 $y_i = \frac{\mathbf{1}_{\{i\}}}{\|D_u \mathbf{1}_{\{i\}}\|_1}$，则 y_i是非负的，且 $\|D_u y_i\|_1 = \langle u, y_i\rangle = 1$。于是，对任意满足 $\|D_u x\|_1 = 1$ 的 x，我们可以写成

$$x = \sum_i c_i \mathbf{1}_{\{i\}} = \sum_i c_i \|D_u \mathbf{1}_{\{i\}}\|_1 y_i$$

with

$$\|D_u x\|_1 = \left\|D_u \sum_i c_i \mathbf{1}_{\{i\}}\right\|_1 = \sum_i |c_i| \|D_u \mathbf{1}_{\{i\}}\|_1 = 1.$$

因此，如果对于每个 $i, \left\|D_u\left(R^\tau - v\cdot u^T\right) y_i\right\|_1 \le \epsilon$，那么

$$\begin{aligned}\left\|D_u\left(R^\tau - v\cdot u^T\right) x\right\|_1 &= \left\|D_u\left(R^\tau - v\cdot u^T\right) \sum_i c_i \|D_u \mathbf{1}_{\{i\}}\|_1 y_i\right\|_1 \\ &\le \sum_i |c_i| \|D_u \mathbf{1}_{\{i\}}\|_1 \left\|D_u\left(R^\tau - v\cdot u^T\right) y_i\right\|_1 \\ &\le \epsilon.\end{aligned}$$

因此，只需找出每个满足 $\|D_u x\|_1 = \langle u, x\rangle = 1$ 的非负 x 的混合时间，该结论便对所有 x 都成立。

□

下一个引理与 3 3, R $A =$ 相似，但它将 $D_u^2 D_v^2 R D_u^2 D_v^2$ 的混合时间联系起来，表明二者相同。

引理 5.3。 R 和 $A = D_u^2 D_v^2 R D_u^2 D_v^2$ 的混合时间相同，其中 R 是一个不可约的 12-惰性非负矩阵，其最大特征值为 1，对应的正特征向量为 u 和 v，并已归一化使得 $\langle u, v\rangle = 1$。

证明。 我们首先证明 R 和 $A = D_u^2 D_v^2 R D_u^2 D_v^2$ 的混合时间相同。注意，如果 R 是 12-惰性的，那么 A 也 12-惰性，因为如果 $R = \frac{1}{2}I + \frac{1}{2}C$，其中 C 是非负的，那么

$$A = \frac{1}{2} D_u^{\frac{1}{2}} D_v^{-\frac{1}{2}} I D_u^{-\frac{1}{2}} D_v^{\frac{1}{2}} + D_u^{\frac{1}{2}} D_v^{-\frac{1}{2}} C D_u^{-\frac{1}{2}} D_v^{\frac{1}{2}} = \frac{1}{2} I + D_u^{\frac{1}{2}} D_v^{-\frac{1}{2}} C D_u^{-\frac{1}{2}} D_v^{\frac{1}{2}}.$$

令 $w = D_u^2 D_v^2 \mathbf{1}$ 为 A 对应特征值 1 的左右特征向量。我们将证明，对于满足 $\|D_u x\|_1 = 1$ 的每个 x，都存在某个 y，使得 $\|D_w y\|_1 = 1$。

$$\|D_u (R^\tau - v \cdot u^T) x\|_1 = \|D_w (A^\tau - w \cdot w^T) y\|_1$$

根据定义，这将意味着 $\tau_\epsilon(R) = \tau_\epsilon(A)$。

令 x 为满足 $\|D_u x\|_1 = 1$ 的向量，并令 $y = D_u^2 D_v^2 x$。那么，由于 $D_w = D_u^2 D_v^2$，我们得到 $\|D_w y\|_1 = \|D_u x\|_1 = 1$。令 $R = v \cdot u^T + B_R$ 和 $A = w \cdot w^T + B_A$，其中 $B_A = D_u^2 D_v^2 B_R D_u^2 D_v^2$ 且 $B_A^\tau = D_u^2 D_v^2 B_R^\tau D_u^2 D_v^2$。那么

$$\begin{aligned} D_u (R^\tau - v \cdot u^T) x &= D_u B_R^\tau x \\ &= D_u (D_u^{-\frac{1}{2}} D_v^{\frac{1}{2}} B_A^\tau D_u^{\frac{1}{2}} D_v^{-\frac{1}{2}}) x \\ &= D_w B_A^\tau y \\ &= D_w (A^\tau - w \cdot w^T) y \end{aligned}$$

如所要求。 □

这里给出最后一个简单引理，它说明当 AA^T 和 $\tilde{A}$ 满足 A 为惰性矩阵时，其特征值之间的不等式。Fill [Fil91] 证明了这一点（尽管针对的是行随机矩阵），也可以通过多种方式得到这一结论。

引理 5.4。*(Fill [Fil91])* 设 A 是一个$1\,2$-惰性不可约非负矩阵，其最大特征值为 *1*，对应的左右特征向量为 w。则

$$\lambda_2(AA^T) \leq \lambda_2\left(\frac{A+A^T}{2}\right).$$

证明。 由于 A 是$1\,2$-惰性的，我们知道 $2A - I$ 是非负的，佩龙–弗罗贝尼乌斯特征值为 1，且对应于特征值 1 的左右特征向量相同 w；由佩龙-弗罗贝尼乌斯定理（定理 2.1，第 2 部分），可知佩龙–弗罗贝尼乌斯特征值为 1；根据引理 3.3，其最大奇异值也为 1。此外，$1\,2(A+A^T)$ 也具有相同的性质。因此，对于任意 x，

$$\begin{aligned}
AA^T &= \frac{A+A^T}{2} + \frac{(2A-I)(2A^T-I)}{4} - \frac{I}{4} \\
\Rightarrow \langle x, AA^T x\rangle &\leq \langle x, \frac{A+A^T}{2}x\rangle + \|x\|_2 \frac{\|(2A-I)\|_2\|(2A^T-I)\|_2}{4}\|x\|_2 - \frac{\|x\|_2^2}{4} \\
&\leq \langle x, \frac{A+A^T}{2}x\rangle \\
\Rightarrow \max_{x\perp w}\langle x, AA^T x\rangle &\leq \max_{x\perp w}\langle x, \frac{A+A^T}{2}x\rangle.
\end{aligned}$$

最后一个推论是由特征值的变分刻画得出的，因为 AA^T 和 $1\,2(A+A^T)$ 都是对称的，因此 -

$$\lambda_2(AA^T) \leq \lambda_2\left(\frac{A+A^T}{2}\right)$$

□

我们先给出非负矩阵的一般界，并将根据这些方程推导出可逆情形下的界。

5.2 混合时间与奇异值

我们首先证明一个简单的引理，建立非负矩阵的混合时间与第二大奇异值之间的关系。这个引理足以用初等方法恢复 Fill [Fil91] 和 Mihail [Mih89]所得的界，并给出所有依赖于第二特征值或边扩张的可逆链混合时间界。由于任意一般非负矩阵 R 的佩龙–弗罗贝尼乌斯特征值为 1 时，其最大奇异值可能远大于 1，因此混合时间与第二大奇异值之间的关系只对左右特征向量相同且对应特征值 1 的非负矩阵才有意义；根据引理 3.3，这类矩阵的最大奇异值为 1。

引理 5.5。（混合时间与第二大奇异值）设 A 是一个佩龙–弗罗贝尼乌斯特征值为 *1* 的非负矩阵（不一定是惰性的），满足 $Aw = w$ 和 $A^T w = w$，其中某个 w 满足 $\langle w, w\rangle = 1$，并设 $\kappa = \text{最小值}_i\, w_i^2$。那么对于每个 $c > 0$，

$$\tau_\epsilon(A) \;\le\; \frac{\ln\left(\sqrt{\frac{n}{\kappa}} \cdot \frac{1}{\epsilon}\right)}{\ln\left(\frac{1}{\sigma_2(A)}\right)} \;\le\; \frac{c \cdot \ln\left(\frac{n}{\kappa \cdot \epsilon}\right)}{1 - \sigma_2^c(A)}.$$

证明。将 τ 简写为 $\tau_\epsilon(A)$ 并且由于 $A = w \cdot w^T + B$ 满足 $Bw = 0$ 且 $B^T w = 0$，我们有 $A^\tau = w \cdot w^T + B^\tau$。令 x 为一个非负向量，满足 $\langle w, x\rangle = \|D_w x\|_1 = 1$。正如定义 5.1之后所讨论的，这足以界定所有 x的混合时间。于是我们得到

$$\|D_w(A^\tau - w \cdot w^T)x\|_1 = \|D_w B^\tau x\|_1 = \|D_w B^\tau D_w^{-1} y\|_1 \le \|D_w\|_1 \|B^\tau\|_1 \|D_w^{-1}\|_1 \|y\|_1$$

其中 $y = D_w x$ 且 $\|y\|_1 = 1$。此外，由于 $\|w\|_2 = 1$ 且 $\kappa = \min_i w_i^2$，我们有 $\|D_w\|_1 \le 1$ 和 $\|D_w^{-1}\|_1 \le \frac{1}{\sqrt{\kappa}}$，利用这些界继续推导上述不等式，可得

$$\|D_w(A^\tau - w \cdot w^T)x\|_1 \le \frac{1}{\sqrt{\kappa}}\|B^\tau\|_1 \le \frac{\sqrt{n}}{\sqrt{\kappa}}\|B^\tau\|_2 \le \frac{\sqrt{n}}{\sqrt{\kappa}}\|B\|_2^\tau \le \frac{\sqrt{n}}{\sqrt{\kappa}}(\sigma_2(A))^\tau \le \epsilon$$

其中，第二个不等式是柯西-施瓦茨不等式；第四个不等式使用了 $\|B\|_2 \le \sigma_2(A)$，因为 A 具有相同的左、右佩龙–弗罗贝尼乌斯特征向量；最后一个不等式则通过令 $\tau = \tau_\epsilon(A) = \ln\left(\frac{\sqrt{n}}{\sqrt{\kappa}\epsilon}\right) \Big/ \ln\left(\frac{1}{\sigma_2(A)}\right)$ 得到，后续不等式由 $1 - x \le e^{-x}$ 推出。

□

对于 $c = 2$ 的情形，引理 5.5 由 Fill 提出 [Fil91] 时采用了不同的方法，但我们发现，我们这个更一般的证明要简单得多。

5.3 混合时间与边扩张

现在，我们将一般非负矩阵 R 的混合时间与其边扩张 $\phi(R)$联系起来。行随机矩阵 R 关于 $\phi(R)$的上界由 Mihail 提出 [Mih89]，随后 Fill 使用引理 [Fil91] 5.5 对$c = 2$进行了简化。因此，下面的引理并非新结果，但我们在此给出证明，因为我们的证明同样初等，且适用于任意非负矩阵 R。

引理 5.6. *(*混合时间与边扩张的上界，由 *Mihail* 提出 *[Mih89])* 设$\tau_\epsilon(R)$ 为一个1 2-惰性非负矩阵 R 的混合时间，其佩龙–弗罗贝尼乌斯特征值为 *1*，且

相应的正左、右特征向量为 u 和 v，并设 $\kappa = \min_i u_i \cdot v_i$。那么

$$\frac{\frac{1}{2} - \epsilon}{\phi(R)} \leq \tau_\epsilon(R) \leq \frac{4 \cdot \ln\left(\frac{n}{\kappa \cdot \epsilon}\right)}{\phi^2(R)}.$$

证明。由引理 5.3可知，R和 $A = D_u^2 D_v^{-2} R D_u^{-2} D_v^2$的混合时间相同。此外，由引理 3.3可知 $\phi(R) = \phi(A)$。因此，我们证明 $\tau_\epsilon(A)$和 $\phi(A)$的界，R的界则随之可得。

我们首先用边扩张 $\tau_\epsilon(A)$给出 ϕ的上界，从而以基本上一行的证明得到 Mihail 的结果。由引理 5.5，对于 $c = 2$，可得

$$\tau_\epsilon(A) \leq \frac{2 \cdot \ln\left(\frac{n}{\kappa \cdot \epsilon}\right)}{1 - \sigma_2^2(A)} = \frac{2 \cdot \ln\left(\frac{n}{\kappa \cdot \epsilon}\right)}{1 - \lambda_2(AA^T)}, \tag{5.7}$$

并在等式 5.4中替换引理5.7，得到

$$\tau_\epsilon(A) \leq \frac{2 \cdot \ln\left(\frac{n}{\kappa \cdot \epsilon}\right)}{1 - \sigma_2^2(A)} = \frac{2 \cdot \ln\left(\frac{n}{\kappa \cdot \epsilon}\right)}{1 - \lambda_2(AA^T)} \leq \frac{2 \cdot \ln\left(\frac{n}{\kappa \cdot \epsilon}\right)}{1 - \lambda_2\left(\frac{A + A^T}{2}\right)} \leq \frac{4 \cdot \ln\left(\frac{n}{\kappa \cdot \epsilon}\right)}{\phi^2\left(\frac{A + A^T}{2}\right)} = \frac{4 \cdot \ln\left(\frac{n}{\kappa \cdot \epsilon}\right)}{\phi^2(A)}$$

其中倒数第二个不等式由对称矩阵 3.1 的切格不等式 $(A + A^T)/2$得到。

现在，我们用 $\tau_\epsilon(A)$给出 $\phi(A)$的下界。我们将证明非负向量 x的情形；根据定义 5.1 及其后的讨论，该结论对所有 x都成立。根据混合时间的定义，对于任意满足 x 的非负$\|D_w x\|_1 = \langle w, x \rangle = 1$，由于 $A^\tau = w \cdot w^T + B^\tau$，

$$\|D_w (A^\tau - w \cdot w^T) x\|_1 = \|D_w B^\tau x\|_1 \leq \epsilon$$

并令 $y = D_w x$，我们得到，对于任意满足 $\|y\|_1 = 1$ 的非负 y，我们有ve

$$\|D_w B^\tau D_w^{-1} y\|_1 \leq \epsilon.$$

将标准基向量代入 i，我们得到，对于每个 i，

$$\sum_j \left| \frac{1}{w(i)} \cdot B^\tau(j,i) \cdot w(j) \right| = \frac{1}{w(i)} \cdot \sum_j |B^\tau(j,i)| \cdot w(j) \leq \epsilon.$$

因此，对于任意集合 S，

$$\sum_{i \in S} w(i)^2 \cdot \frac{1}{w(i)} \cdot \sum_j |B^\tau(j,i)| \cdot w(j) = \sum_{i \in S} \sum_j w(i) \cdot |B^\tau(j,i)| \cdot w(j) \leq \sum_{i \in S} w(i)^2 \cdot \epsilon. \tag{5.8}$$

此外，对于满足 $\sum_{i\in S} w_i^2 \le \frac{1}{2}$ 的任意集合 S，

$$\begin{aligned}
\phi_S(A^\tau) &= \frac{\langle \mathbf{1}_S, D_w A^\tau D_w \mathbf{1}_{\overline{S}}\rangle}{\langle \mathbf{1}_S, D_w^2 \mathbf{1}\rangle} = \frac{\langle \mathbf{1}_{\overline{S}}, D_w A^\tau D_w \mathbf{1}_S\rangle}{\langle \mathbf{1}_S, D_w^2 \mathbf{1}\rangle} \\
&\quad \text{[since } D_w A^\tau D_w \text{ is Eulerian, i.e. } D_w A^\tau D_w \mathbf{1} = D_w (A^\tau)^T D_w \mathbf{1}\text{]} \\
&= \frac{\sum_{i\in S}\sum_{j\in\overline{S}} A^\tau(j,i)\cdot w(i)\cdot w(j)}{\sum_{i\in S} w_i^2} \\
&= \frac{\sum_{i\in S}\sum_{j\in\overline{S}} w(j)\cdot w(i)\cdot w(i)\cdot w(j) + \sum_{i\in S}\sum_{j\in\overline{S}} B^\tau(j,i)\cdot w(i)\cdot w(j)}{\sum_{i\in S} w_i^2} \\
&\ge \sum_{j\in\overline{S}} w_j^2 - \frac{\sum_{i\in S}\sum_{j\in\overline{S}} |B^\tau(j,i)|\cdot w(i)\cdot w(j)}{\sum_{i\in S} w_i^2} \\
&\ge \sum_{j\in\overline{S}} w_j^2 - \frac{\sum_{i\in S}\sum_{j} |B^\tau(j,i)|\cdot w(i)\cdot w(j)}{\sum_{i\in S} w_i^2} \\
&\ge \frac{1}{2} - \epsilon \\
&\quad \text{[since } \sum_{i\in S} w_i^2 \le \frac{1}{2} \text{ and } \sum_i w_i^2 = 1\text{, and the second term follows from equation 5.8]}
\end{aligned}$$

因此

$$\phi(A^\tau) \ge \frac{1}{2} - \epsilon,$$

并利用引理3.22，我们得到

$$\phi(A^\tau) \le \tau \cdot \phi(A),$$

or

$$\frac{\frac{1}{2} - \epsilon}{\phi(A)} \le \tau_\epsilon(A).$$

□

5.4 可逆矩阵的混合时间与谱隙

我们推导可逆非负矩阵的混合时间与谱隙之间的界。这些界众所周知，我们在此列出，以便与一般矩阵的界进行对比。此外，利用引理 5.5，这些界很容易推导。

引理 5.9。 （可逆矩阵的混合时间与谱隙）设 $\tau_\epsilon(R)$ 为具有佩龙–弗罗贝尼乌斯特征值 1 的1 2-惰性非负可逆矩阵 R 的混合时间

以及相应的正左、右特征向量 u 和 v，归一化使得$\langle u, v\rangle = 1$，并令 $\kappa =$ 最小值$_i\, u_i v_i$。那么对于 $\epsilon < 1$,

$$\frac{1-\epsilon}{2\cdot\Delta(R)} \;\le\; \frac{\ln\left(\frac{1}{\epsilon}\right)}{\ln\left(\frac{1}{\lambda_2(R)}\right)} \;\le\; \tau_\epsilon(R) \;\le\; \frac{\ln\left(\frac{n}{\kappa\cdot\epsilon}\right)}{\Delta(R)}.$$

证明。由于 R 是可逆的，所有特征值均为实数；又由于它是惰性的，对于所有 i，$\lambda_i(R) \ge 0$，因此 $\sigma_2(R) = \lambda_2(R)$。因此，上界由引理 1 1 1 1 ¯ ¯5.5 得出。我们证明 $A = D_u^2 D_v^{-2} R D_u^{-2} D_v^2$的下界；由 5.3 可得 R的下界。设 x是对应特征值 $\lambda_2(A)$ 的特征向量，则 $\langle x, w\rangle = 0$；令 x归一化，使得 $\|D_w x\|_1 = 1$，如定义 5.1 所要求，其中 $w = D_u^2 D_v^2 \mathbf{1}$。因此，

$$\begin{aligned}
\|D_w(A^\tau - w\cdot w^T)x\|_1 &= \|D_w(A^\tau x - w\cdot w^T x)\|_1 \\
&= \|D_w(\lambda_2^\tau(A)x - w\cdot \mathbf{0})\|_1 \\
&= \lambda_2^\tau(A)\|D_w x\|_1 \\
&\quad [\text{since } A \text{ is } \tfrac{1}{2}\text{-lazy and thus } \lambda_2(A)\ge 0] \\
&= \lambda_2^\tau(A) \\
&= (1-\Delta(A))^\tau
\end{aligned}$$

其大于 ϵ，除非

$$\tau \ge \frac{\ln\left(\frac{1}{\epsilon}\right)}{\ln\left(\frac{1}{1-\Delta(A)}\right)}.$$

如果 $\Delta(A) \le \frac{1}{2}$，那么

$$\begin{aligned}
1-\Delta(A) &\ge \exp(-2\cdot\Delta(A)), \\
2\cdot\Delta(A) &\ge \ln\left(\frac{1}{1-\Delta(A)}\right)
\end{aligned}$$

又由于 $\ln(\frac{1}{\epsilon}) \ge 1-\epsilon$，这将得到

$$\tau \ge \frac{1-\epsilon}{2\cdot\Delta(A)}.$$

现在假设 $\Delta(A) > \frac{1}{2}$或 $\lambda_2(A) < \frac{1}{2}$，那么我们知道，当 $\lambda_2(A) = \frac{1}{2}$ 时，混合时间将达到最大，因此至少为

$$\frac{\ln\left(\frac{1}{\epsilon}\right)}{\ln 2} \geq \frac{1-\epsilon}{\ln(2)} \geq \frac{1-\epsilon}{2} \geq 1$$

这始终成立，因为混合时间至少为 1，因此我们再次得到

$$\tau \geq \frac{1-\epsilon}{2 \cdot \Delta(A)}$$

如所要求的。 □

5.5 非可逆矩阵的混合时间与谱隙

我们使用与推导 ϕ 的上界（见定理 1.2）时相似的方法，得到非负矩阵混合时间关于谱隙的界。

引理 5.10. （混合时间与谱隙）设 $\tau_\epsilon(R)$ 为1 2惰性非负矩阵 R 的混合时间，其佩龙–弗罗贝尼乌斯特征值为 *1*，对应的正左、右特征向量分别为 u 和 v，并设 $\kappa = \min_i u_i \cdot v_i$。则

$$(1-\Delta)\frac{1-\epsilon}{\Delta} \leq \tau_\epsilon(R) \leq 20 \cdot \frac{n + \ln\left(\frac{1}{\kappa \cdot \epsilon}\right)}{\Delta(R)}.$$

证明。 由于混合时间（如引理5.3的证明所示）、特征值、1 1 1 1 ¯ ¯边扩张以及 κ 对 R 和 $A = D_u^2 D_v{}^2 R D_u{}^2 D_v^2$ 的取值均相同，我们给出 A 的界，而 R 的界可类似得到。

为得到下界，令 y 为A的第二特征值 λ_2 对应的特征向量，并将其归一化，使得 $\|D_w y\| = 1$。注意，由于 w是 A对应特征值 1 的左右特征向量，因此有 $\langle w, y\rangle = 0$。考虑 y的混合时间时，注意

$$\|D_w(A^\tau - w \cdot w^T)y\|_1 = \|D_w(A^\tau - w \cdot w^T)y\|_1 = \|D_w \lambda_2^\tau y\|_1 = |\lambda_2|^\tau$$

对于.....，该值小于 ϵ

$$\tau_\epsilon(A) \geq \frac{\ln(\frac{1}{\epsilon})}{\ln(\frac{1}{|\lambda_2|})} \geq \frac{1-\epsilon}{1-|\lambda_2|}|\lambda_2| \geq (1-\Delta)\frac{1-\epsilon}{\Delta}$$

wh因为 $1 - x \leq \ln(1/x) \leq 1/x - 1$ 且 $\mathrm{Re}\lambda_2 \leq |\lambda_2$，所以最后一个不等式成立 |.

对于 $\tau_\epsilon(A)$的上界，我们也只考虑非负向量；一般向量的上界可根据引理5.2中讨论的三角不等式得出。与引理 5.5的证明类似，对于任意满足 x 的非负向量 $\langle w, x\rangle = 1$，对于$A = w \cdot w^T + B$,

$$\|D_w(A^\tau - w \cdot w^T)x\|_1 \leq \sqrt{\frac{n}{\kappa}} \cdot \|B^\tau\|_2$$

且具有

$$\|B^\tau\|_2 \leq \frac{\epsilon\sqrt{\kappa}}{\sqrt{n}}$$

是充分的。令 T 为 B的舒尔形式中的三角矩阵。由引理3.26的证明可知，对于

$$k \geq \frac{3.51n + 1.385 \ln\left(\frac{n}{\delta}\right)}{1 - |\lambda_m(A)|},$$

范数

$$\|B^k\|_2 \leq \delta,$$

因此令 $\delta = \frac{\epsilon\sqrt{\kappa}}{\sqrt{n}}$，可得

$$\tau_\epsilon(A) \leq \frac{3.51n + 1.385 \ln\left(n \cdot \frac{\sqrt{n}}{\sqrt{\kappa}\cdot\epsilon}\right)}{1 - |\lambda_m(A)|}.$$

此外，由于 A 是┤2-惰性的， $2A - I$ 也非负，且具有相同的正左右特征向量 w ，其对应佩龙 – 弗罗贝尼乌斯特征值 1，因此最大奇异值为 1，故 $2A - I$ 的每个特征值的模都至多为 1。类似于引理 3.30的证明，若 $\lambda_r = a + i \cdot b$ 是 A的任一特征值，则其在$2A - I$ 中的对应特征值为 $2a - 1 + i \cdot 2b$，其模至多为 1，从而得到 $(2a - 1)^2 + 4b^2 \leq 1$或 $a^2 + b^2 \leq a$。进一步可得

$$1 - |\lambda_r| = 1 - \sqrt{a^2 + b^2} \geq 1 - \sqrt{1 - (1 - a)} \geq 1 - e^{\frac{1}{2}(1-a)} \geq \frac{1}{4}(1 - a)$$

or

$$1 - |\lambda_m| \geq \frac{1}{4}(1 - \mathrm{Re}\lambda_m) \geq \frac{1}{4}(1 - \mathrm{Re}\lambda_2),$$

由此得到

$$\begin{aligned}\tau_\epsilon(A) &\leq 4 \cdot \frac{3.51n + 1.385 \ln\left(n \cdot \frac{\sqrt{n}}{\sqrt{\kappa}\cdot\epsilon}\right)}{1 - \mathrm{Re}\lambda_2(A)} \\ &\leq 20 \cdot \frac{n + \ln\left(\frac{1}{\kappa \cdot \epsilon}\right)}{1 - \mathrm{Re}\lambda_2(A)}\end{aligned}$$

完成证明。 □

我们注意到，这里仅存在对 $\ln\left(\frac{n}{\kappa}\cdot\epsilon\right)$ 的加性依赖，而不存在乘性依赖。此外，我们的构造 4.17 和 4.27 也表明，引理 5.10 中利用 $\mathrm{Re}\lambda_2$ 得到的 τ 上界也是（几乎）紧的。对于定理 4.19 中 A_n 的构造，令 U_n 的列为 $u_1,\ldots,u_n$，其中 $x=u_2$ 为 $(A_n^k-J)u_2=(1-(2+\sqrt{n})^{-1})^k u_3$，因此对于 $k=O(\sqrt{n})$，$A^{O(\sqrt{n})}$ 的三角块范数约为 $1/e$，经过约 $\ln\left(\frac{n}{\epsilon}\right)$ 次幂后，该范数进一步降至小于 ϵ。因此，对于矩阵 $A_n,\tau_\epsilon(A_n)\in O\left(\sqrt{n}\cdot\ln\left(\frac{n}{\epsilon}\right)\right)$。这表明引理 5.10 也是（几乎）紧的，因为 $\lambda_2(A_n)=0$。

5.6 非负矩阵及其加性对称化的混合时间

对于具有相同左右特征向量 w、对应佩龙–弗罗贝尼乌斯特征值 1 的非负矩阵 A，我们还可以用其加性对称化 $M=\frac{1}{2}(A+A^T)$ 的混合时间来界定其混合时间。由于 $\phi(A)=\phi(M)$，我们可以利用引理 5.6 中边扩张与混合时间之间的双边界来界定 $\tau_\epsilon(A)$ 和 $\tau_\epsilon(M)$。对于下界，我们得到 $\gamma_1\cdot\sqrt{\tau_\epsilon(M)}\le\tau_\epsilon(A)$，而对于上界，我们得到

$$\tau_\epsilon(A)\le\gamma_2\cdot\tau_\epsilon^2(M),$$

其中 γ_1 和 γ_2 是关于 n、κ、$\frac{1}{\epsilon}$ 的多对数函数。然而，通过界定相应的算子，我们可以证明 $\tau_\epsilon(A)$ 有一个更紧的上界，其对 $\tau_\epsilon(M)$ 的依赖仅为线性而非二次。

引理 5.11。设 A 是一个具有正左、右特征向量 w、对应佩龙–弗罗贝尼乌斯特征值 1 的 $\frac{1}{2}$-惰性非负矩阵，令 $M=\frac{1}{2}(A+A^T)$，且 $\kappa=\min_i w_i^2$。那么

$$\frac{1-2\epsilon}{4\cdot\ln^{\frac{1}{2}}\left(\frac{n}{\kappa\cdot\epsilon}\right)}\cdot\tau_\epsilon^{\frac{1}{2}}(M)\le\tau_\epsilon(A)\le\frac{2\cdot\ln\left(\frac{n}{\kappa\cdot\epsilon}\right)}{\ln\left(\frac{1}{\epsilon}\right)}\cdot\tau_\epsilon(M).$$

证明。由于 $\phi(A)=\phi(M)$ 且二者具有相同的左右特征向量 w，对应佩龙–弗罗贝尼乌斯特征值 1，因此，$\tau_\epsilon(A)$ 关于 $\tau_\epsilon(M)$ 的下界由引理 5.6 立即得出，因为

$$\sqrt{\tau_\epsilon(M)}\le\ln^{\frac{1}{2}}\left(\frac{n}{\kappa\cdot\epsilon}\right)\cdot\frac{1}{\phi(M)}=\ln^{\frac{1}{2}}\left(\frac{n}{\kappa\cdot\epsilon}\right)\cdot\frac{1}{\phi(A)}\le\ln^{\frac{1}{2}}\left(\frac{n}{\kappa\cdot\epsilon}\right)\cdot\frac{2\tau_\epsilon(A)}{1-2\epsilon}.$$

对于上界，我们首先为具有佩龙–弗罗贝尼乌斯特征值 1 的半正定非负矩阵 M 定义一个新量，其中 w 是对应的左右特征向量。设 $T_\epsilon(M)$ 为满足下式的最小数 k

$$\|M^k-w\cdot w^T\|_2=\epsilon.$$

由于 M 是对称矩阵，我们可以写成 $M = w\cdot\ w^T + UDU^*$，其中酉矩阵 U 的第一列为 w，对角矩阵 D 包含 M 的所有特征值，但将特征值 1 替换为 0。此外，由于 M 是半正定矩阵，$\lambda_2(M)$ 是 M 的第二大特征值，因此对于每个 $i > 2$，$0 \le \lambda_i(M) \le \lambda_2(M)$。由此可得

$$\|M^k - w \cdot w^T\|_2 = \|UD^kU^*\|_2 = \|D^k\|_2 = \lambda_2^k(M)$$

and

$$T_\epsilon(M) = \frac{\ln\left(\frac{1}{\epsilon}\right)}{\ln\left(\frac{1}{\lambda_2(M)}\right)}.$$

此外，对于对应于 λ_2 的 M 的特征向量 y，我们有，对于 $k = T_\epsilon$（M），

$$(M^k - w \cdot w^T)y = \lambda_2^k \cdot y = \epsilon \cdot y.$$

令 $x = \frac{y}{\|D_w y\|_1}$，我们有 $\|D_w x\|_1 = 1$，于是得到

$$\|D_w(M^k - w \cdot w^T)x\|_1 = \left\|D_w \frac{\epsilon \cdot y}{\|D_w y\|_1}\right\|_1 = \epsilon,$$

这意味着

$$\tau_\epsilon(M) \ge k = T_\epsilon(M), \tag{5.12}$$

因为存在某个向量 x，满足 $\|D_w x\| = 1$ 且具有 $\|D_w$（$M^k - w\cdot\ w^T$）$x\|_1 = \epsilon$；并且对于每个 $t < k$，也有 $\|D_w$（$M^t - w\cdot\ w^T$）$x\|_1 > \epsilon$。

现在我们注意到，由于 A 是 1 2-惰性的，且佩龙–弗罗贝尼乌斯特征值为 1，并且对应于特征值 1 的左右特征向量相同，为 w，因此 $M = \frac{1}{2}(A + A^T)$ 是半正定的。由引理 5.5，可得

$$\tau_\epsilon(A) \le \frac{2 \cdot \ln\left(\frac{\sqrt{n}}{\sqrt{\kappa}\cdot\epsilon}\right)}{\ln\left(\frac{1}{\sigma_2^2(A)}\right)}$$

且由于 A 是 1-2-惰性的，如引理 5.6 的证明所示，

$$\sigma_2^2(A) = \lambda_2(AA^T) \le \lambda_2(M),$$

给出

$$\tau_\epsilon(A) \le \frac{2 \cdot \ln\left(\frac{\sqrt{n}}{\sqrt{\kappa}\cdot\epsilon}\right)}{\ln\left(\frac{1}{\lambda_2(M)}\right)} = \frac{2 \cdot \ln\left(\frac{\sqrt{n}}{\sqrt{\kappa}\cdot\epsilon}\right)}{\ln\left(\frac{1}{\epsilon}\right)} \cdot \frac{\ln\left(\frac{1}{\epsilon}\right)}{\ln\left(\frac{1}{\lambda_2(M)}\right)} = \frac{2 \cdot \ln\left(\frac{\sqrt{n}}{\sqrt{\kappa}\cdot\epsilon}\right)}{\ln\left(\frac{1}{\epsilon}\right)} \cdot T_\epsilon(M) \le \frac{2 \cdot \ln\left(\frac{\sqrt{n}}{\sqrt{\kappa}\cdot\epsilon}\right)}{\ln\left(\frac{1}{\epsilon}\right)} \cdot \tau_\epsilon(M)$$

其中最后一个不等式是由方程 5.12 得出的。 □

引理 5.11的一个应用如下：给定任意无向图 G，其中每个顶点的度为 d，以任何 方式对 G的边进行定向，使所得图中每个顶点的入度和出度均为 $d/2$，都不会增加随机游走的混合时间（至多相差 $\ln\left(\frac{n}{\kappa\cdot\epsilon}\right)$ 倍）。

—

5.7 连续时间算子的混合时间

设 R 是佩龙–弗罗贝尼乌斯特征值为 1 的非负矩阵，其对应的正左、右特征向量分别为 u 和 v。与 R 相关联的连续时间算子定义为 $\exp\left(t\cdot(R-I)\right)$，其中，对于任意矩阵 M，我们形式上定义 $\exp(M)=\sum_{i=0}\frac{1}{i!}M^i$。之所以将此算子视为连续时间算子，是因为从任意向量 x_0 出发，时间 $t\in\mathbb{R}_{\geq 0}$ 时的向量 x_t 定义为 $x_t=\exp\left(t\cdot(R-I)\right)x_0$。由于

$$\exp\left(t\cdot(R-I)\right)=\exp(t\cdot R)\cdot\exp(-t\cdot I)=e^{-t}\sum_{i=0}^{\infty}\frac{1}{i!}t^iR^i$$

由于 R 和 I 可交换，我们将算子拆分为两项，因此 $\exp\left(t\cdot(R-I)\right)$ 是非负的；若 R 的任意特征值为 λ，对应的特征向量为 y，则 $e^{t(\lambda-1)}$ 是 $\exp\left(t\cdot(R-I)\right)$ 的特征值，对应的特征向量仍为 y。由此进一步可知，u 和 v 是 $\exp\left(t\cdot(R-I)\right)$ 的左、右特征向量，其佩龙–弗罗贝尼乌斯特征值为 1。$\exp\left(t\cdot(R-I)\right)$ 的混合时间是满足以下条件的 t 值：

$$\left\|D_u\left(\exp\left(t\cdot(R-I)\right)-v\cdot u^T\right)v_0\right\|_1\leq\epsilon$$

对于任意 v_0使得 $\|D_uv_0\|_1=1$，因此，这与按照定义 5.1 的意义考察$\exp(R-I)$ 的混合时间完全相同。

引理 5.13. 设 R 是一个非负矩阵（不一定是惰性的），其对应于佩龙–弗罗贝尼乌斯特征值 *1* 的正左、右特征向量分别为 u 和 v，且归一化为 $\langle u,v\rangle=1$；令 $\kappa=\min_i u_i\cdot v_i$。则 $\exp(t\cdot(R-I))$，或 $\tau_\epsilon(\exp(R-I))$ 的混合时间满足如下界

$$\frac{\frac{1}{2}-\epsilon}{\phi(R)}\ \leq\ \tau_\epsilon\left(\exp(R-I)\right)\ \leq\ \frac{100\cdot\ln\left(\frac{n}{\kappa\cdot\epsilon}\right)}{\phi^2(R)}.$$

证明。 我们首先求算子 $\exp\left(\frac{R-I}{2}\right)$ 的混合时间，而算子 $\exp(R-I)$ 的混合时间恰好是这个数的两倍。根据 exp 函数的级数展开可知，$\exp\left(\frac{R-I}{2}\right)$ 的佩龙–弗罗贝尼乌斯特征值为 1，且 u和 v 分别是对应的左、右特征向量。此外，

$$\exp\left(\frac{R-I}{2}\right)=e^{-\frac{1}{2}}\left(I+\sum_{i\geq 1}\frac{1}{i!}\frac{R^i}{2^i}\right)$$

由于首项以及 $e^{-\frac{1}{2}} \geq \frac{1}{2}$ 和所有其他项均为非负数，因此该算子是 $\frac{1}{2}$-惰性的。此外，对于任意满足 $\sum_{i \in S} u_i v_i \leq \frac{1}{2}$ 的集合 S，令 $\delta = \frac{1}{2}$，则

$$\begin{aligned}
\phi_S\left(\exp\left(\frac{R-I}{2}\right)\right) &= e^{-\delta} \cdot \frac{\langle \mathbf{1}_S, D_u \exp\left(\frac{R}{2}\right) D_v \mathbf{1}_{\overline{S}} \rangle}{\langle \mathbf{1}_S, D_u D_v \mathbf{1} \rangle} \\
&= e^{-\delta} \cdot \sum_{i \geq 1} \frac{\delta^i}{i!} \cdot \frac{\langle \mathbf{1}_S, D_u R^i D_v \mathbf{1}_{\overline{S}} \rangle}{\langle \mathbf{1}_S, D_u D_v \mathbf{1} \rangle} \\
&\quad \text{[since } \langle \mathbf{1}_S, D_u I D_v \mathbf{1}_{\overline{S}} \rangle = 0\text{]} \\
&= e^{-\delta} \cdot \sum_{i \geq 1} \frac{\delta^i}{i!} \cdot \phi_S(R^i) \\
&\quad \text{[since } R^i \text{ also has } u \text{ and } v \text{ as the left and right eigenvectors for eigenvalue 1]} \\
&\leq e^{-\delta} \cdot \sum_{i \geq 1} \frac{\delta^i}{i!} \cdot i \cdot \phi_S(R) \\
&\quad \text{[using Lemma 3.22]} \\
&= e^{-\delta} \cdot \delta \cdot \phi_S(R) \sum_{i \geq 1} \frac{\delta^{i-1}}{(i-1)!} \\
&= e^{-\delta} \cdot \delta \cdot \phi_S(R) \cdot e^{\delta} \\
&= \delta \cdot \phi_S(R)
\end{aligned}$$

因此，

$$\phi\left(\exp\left(\frac{R-I}{2}\right)\right) \leq \frac{1}{2}\phi(R). \tag{5.14}$$

此外，考虑级数展开中的第一项，可得

$$\phi_S\left(\exp\left(\frac{R-I}{2}\right)\right) = e^{-\delta} \sum_{i \geq 1} \frac{\delta^i}{i!} \cdot \phi_S(R^i) \geq e^{-\delta} \cdot \delta \cdot \phi_S(R)$$

or

$$\phi\left(\exp\left(\frac{R-I}{2}\right)\right) \geq \frac{3}{10} \cdot \phi(R). \tag{5.15}$$

由于 $\exp\left(\frac{R-I}{2}\right)$ 具有左右特征向量 u 和 v，且是 $\frac{1}{2}$-惰性的，由引理5.6 以及方程 5.14 和 5.15 可得

$$\frac{1}{2} \cdot \frac{\frac{1}{2} - \epsilon}{\frac{1}{2}\phi(R)} \leq \tau_\epsilon\left(\exp(R-I)\right) \leq 2 \cdot \frac{4 \cdot \ln\left(\frac{n}{\kappa \cdot \epsilon}\right)}{\left(\frac{3}{10}\right)^2 \phi^2(R)} \leq \frac{100 \cdot \ln\left(\frac{n}{\kappa \cdot \epsilon}\right)}{\phi^2(R)}$$

从而得到该结果。 □

5.8 与典范路径方法的比较

对于特征值为 1 的对称非负矩阵 M，如引理 5.9 所示，由于 τ 与 $1-\lambda_2$ 成反比（相差一个 $\ln(\kappa^n \cdot \epsilon)$ 因子），因此任何谱隙下界都可用于给出 $\tau_\epsilon(M)$ 的上界。此外，由于 $1-\lambda_2$ 可以写成对称矩阵的最小化问题（见第 2 节），任何对该优化问题的松弛都可用于得到 $1-\lambda_2$ 的下界，而由此得到的不等式称为庞加莱不等式。一种方法是在底层加权图中使用典范路径 [Sin92]，这有助于在某些无法计算 λ_2 或 ϕ 的情况下界定混合时间。然而，由于典范路径可以有许多不同的定义方式，因此会产生多种用于界定 $1-\lambda_2$ 的松弛方法，各自在不同情形下有用。这里我们给出一个特定的定义和引理，因为它与我们在定理 4.19 中的构造相关；为此，我们对其作适当修改，使之适用于双随机情形。

引理 5.16. *[Sin92]* 设 M 表示一个对称双随机矩阵。设 W 是 M 中的一个路径集，每对顶点之间各有一条路径。对于顶点 (u,v) 之间的任意路径 $\gamma_{u,v} \in S$，其中 $\gamma_{u,v}$ 只是 u 与 v 之间的一组边，令边的数量（或路径的非加权长度）为 $|\gamma_{u,v}|$。令

$$\rho_W(M) = \max_{e=(x,y)} \frac{\sum_{(u,v):e\in\gamma_{u,v}} |\gamma_{u,v}|}{n \cdot M_{x,y}}.$$

那么对于任意 W

$$1-\lambda_2(M) \geq \frac{1}{\rho_W(M)}$$

因此，

$$\tau_\epsilon(M) \leq \rho_W(M) \cdot \ln\left(\frac{n}{\epsilon}\right).$$

推论 5.17. 结合引理 *3.12* 和引理 *5.16*，可知对于任意双随机矩阵 A 以及 W 在 AA^T 的底层图中的路径集，

$$\tau_\epsilon(A) \leq \frac{2\cdot\ln\left(\frac{n}{\epsilon}\right)}{1-\sigma_2^2(A)} = \frac{2\cdot\ln\left(\frac{n}{\epsilon}\right)}{1-\lambda_2(AA^T)} \leq 2\cdot\rho_W(AA^T)\cdot\ln\left(\frac{n}{\epsilon}\right).$$

考虑定理 A_n 中的例子 4.19。不难看出

$$\tau_\epsilon(A_n) \in O\left(\sqrt{n}\cdot\ln\left(\frac{n}{\epsilon}\right)\right). \tag{5.18}$$

这是因为 $\sqrt{n}$ 这一因子确保了 $\bar{A}^{\lceil\sqrt{n}\rceil}$ 在舒尔形式下的三角矩阵 T_n 中，只有 e^{-1} 附近的元素非零，而因子

$\ln\left(\frac{n}{\epsilon}\right)$ 进一步将这些元素的模限制为在 A^{τ} 中至多为 $\frac{\epsilon}{n}$。因此，算子范数约为 $\frac{\epsilon}{n}$，而 l_1 范数的上界为 ϵ。然而，由定理 4.19，由于 $\phi(A_n) \geq \frac{1}{6\sqrt{n}}$，根据引理 5.6 可得 $\tau_\epsilon(A_n) \in O\left(n \cdot \ln\left(\frac{n}{\epsilon}\right)\right)$，与式 5.18 中的实际上界相比，差了一个二次因子。此外，根据定理 4.19，A_n 的第二特征值为 0，即使应用引理 5.10，也会比实际上界多出一个二次因子。不过，引理 5.5 和推论 5.17 确实给出了正确的上界。由于根据定理 4.19，$\sigma_2(A_n) = 1 - \frac{1}{\sqrt{n+2}}$，由引理 5.5 可得，对于 $c = 1$，$\tau_\epsilon(A_n) \in O\left(\sqrt{n} \cdot \ln\left(\frac{n}{\epsilon}\right)\right)$，与式 5.18 中的上界一致。现在来看典范路径和推论 5.17 给出的上界，考虑矩阵 $M = A_n A_n^T$。事实证明，M 的每个元素都是正的，因此选取集合 W，使得任意一对顶点之间的路径就是连接这两个顶点的边。此外，对于证明定理 4.19 时定义的 r_n、α_n、β_n，有 $M = J + r_n^2 B$，其中

$$B_{1,1} = \frac{n-2}{n},\ B_{n,\,n} = (n-2)\cdot\beta_n^2,\ B_{i,i} = \alpha_n^2 + (n-3)\cdot\beta_n^2,\ B_{1,n} = B_{n,1} = \frac{n-2}{\sqrt{n}}\cdot\beta_n,$$

$$B_{n,j} = B_{j,n} = \alpha_n\cdot\beta_n + (n-3)\cdot\beta_n^2,\quad B_{1,j} = B_{j,1} = \frac{1}{\sqrt{n}}\cdot(\alpha_n + (n-3)\cdot\beta_n),\quad B_{i,\,j} = 2\cdot\alpha_n\cdot\beta_n + (n-4)\cdot\beta_n^2,$$

并且 $2 \leq i,\ j \leq n-1$。由此可知，矩阵 M 的任意一个元素都至少为 $c \cdot n^{-\frac{3}{2}}$（其中某个常数为 c），根据推论5.17，可得 $\tau_\epsilon(A_n) \in O\left(\sqrt{n} \cdot \ln\left(\frac{n}{\epsilon}\right)\right)$，与公式5.18中的界一致。

5.9 舒尔补

这是本学位论文中第一次讨论拉普拉斯矩阵。此前一直避免讨论它，是因为虽然它隐含地存在于所有论证中，但我们似乎并不需要它来得到非负矩阵所需的任何界。我们的假设基本上与此前相同。

设 R 是一个不可约非负矩阵。我们需要不可约性的概念，因为我们将处理 R 的子矩阵的逆；若不具备这一性质，这些逆将不存在。此外，R 具有正的 u，并且 v 是对应于特征值 1 的左、右特征向量，且归一化为 $\sum_i u_i v_i = 1$。

定义 5.19。（非负矩阵的拉普拉斯矩阵）设 R 是一个佩龙–弗罗贝尼乌斯特征值为 1 的不可约非负矩阵。定义 R 的拉普拉斯矩阵为 $L = I - R$。◊

注 5.20。文献中对非负矩阵的拉普拉斯矩阵有许多不同的定义 R。为简单起见，假设我们想要了解

不可约非负矩阵的 A 拉普拉斯矩阵 w，其最大特征值为 1，且左右正特征向量为A，那么 $I - \frac{1}{2}(A + A^T)$和 $I - A$是其拉普拉斯矩阵最重要的两种定义。第一种定义被广泛用于推导有关不可约矩阵的结果，这主要是因为对称矩阵更易处理，所得结果也更简洁。然而，第一种定义丢失了关于 A特征值的全部信息，因此我们主要采用第二种定义。由于缺乏对称情形下那种特征值的变分刻画，这一定义更难处理，需要全新的思路和工具。该定义也用于许多其他领域，例如统计物理（见 [Tol79]）。◊

记号。 令 $L_{S,T}$ 表示由 S 中的行和T中的列所确定的 L 的子矩阵，为简洁起见，记$L_S = L_{S,S}$。类似地，对于某个向量 $x \in \mathbb{R}^n$，我们将 x_S 记为 $R^{|S|}$ 中的一个向量，其矩阵元素对应于S中的指标。需要说明的是，对于单元素集合，我们总是写作$x_{\{i\}}$。

第一个简单的引理如下。

引理 5.21。 设 $L = I - R$ 是不可约非负 R 的拉普拉斯矩阵，其最大特征值为 *1*，对应的正左、右特征向量分别为 u 和 v，则 $L_{U,U}$ 可逆，除非 $U = [n]$。

证明。 反设结论不成立，则存在特征值 0 对应于 $L_{U,U}$ 的特征向量 $x_U \neq 0$，将其扩展为向量 x，使其在 U之外取值为 0，于是$Lx = 0$，或者 $Rx = x$，这与佩龙-弗罗贝尼乌斯定理 2.1矛盾，因为 v 是 R对应于特征值 1 的唯一正向量，而 x 在 U之外取值为 0，因此不是正向量。

□

定义 5.22.（拉普拉斯矩阵的逆）线性代数中的一个基本问题，是在给定 b（其中 $L = I - R$，且 R 是最大特征值为 1 的不可约非负矩阵）的情况下，求解方程 $Lx = b$ 中的 x。由于 L 的核维数为 1，这意味着对于 $b \neq 0$，存在唯一的 x 解，使得 $\langle x, v\rangle = 0$，其中 $x = L^+ b$。◊

为了更清楚地理解解 $x = L^+ b$，对于某个固定集合 $U \subset [n]$，令矩阵 $R = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$给出 $L = \begin{bmatrix} I - A & -B \\ -C & I - D \end{bmatrix} = \begin{bmatrix} A' & B' \\ C' & D' \end{bmatrix}$，并假设 A 所在的顶点集合为 U，为简便起见，记作 $x_U = y, x_U = z$，类似地 $\bar{}$

对于 b，那么根据 $Lx = b$，我们得到

$$A'y + B'z = b_1 \tag{5.23}$$

and

$$C'y + D'z = b_2 \tag{5.24}$$

and 若 D' 可逆，这在 L 是引理中的拉普拉斯矩阵时成立5.21，则

$$z = b_2 - D'^{-1}C'y \tag{5.25}$$

and

$$(A' - B'D'^{-1}C')y = b_1 - B'b_2. \tag{5.26}$$

由此引出舒尔补的定义。

定义 5.27。（舒尔补）设拉普拉斯矩阵 $L = I - R$，其中

$$R = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

是最大特征值为 1 的不可约非负矩阵。相对于某个集合 $U \subset [n]$，L 的舒尔补记作 $L|_U$，定义为

$$L|_U = A' - B'D'^{-1}C' = (I - A) - B(I - D)^{-1}C = L_U - L_{U,\overline{U}}L_{\overline{U}}^{-1}L_{\overline{U},U}.$$

我们也允许算子 $|$ 作用于非负矩阵 R，并将其定义为

$$R|_U = I_U - (I - R)|_U = I_U - L|_U = A + B(I - D)^{-1}C.$$

◊

5.10 舒尔补的性质

下面我们列出舒尔补的一些性质，或者更具体地说，列出 $|$ 算子的性质，该算子定义于 5.27。

1. **定义良好**。 舒尔补定义良好。这可由引理5.21得出，因为 $I - D$ 可逆。

2. **传递性**。 设 U 和 V 是指标的两个子集，且 $U \subseteq V$。那么

   $$(L|_V)|_U = L|_U$$

   这可以很容易地从方程5.23、5.24、5.25、5.26中看出，因为求解方程的一个子集不会改变最终解。也可以根据定义直接计算得出，但过程冗长且没有什么启发性。

3.**与逆运算的交换性**。 $(L^{-1})_U = (L|_U)^{-1}$. 再次由方程 5.23、5.24、5.21、5.26 可得，也可以根据定义直接计算得出。

4. **与转置的交换性**。 $L^T|_U = (L|_U)^T$，再次由转置定义立即可得。

5. **拉普拉斯矩阵的封闭性**。 $L|_U$ 是不可约非负矩阵 $R|_U$ 的拉普拉斯矩阵，其左右特征向量为 u_U 和 v_U，对应的特征值为 1。为说明这一点，首先注意到

$$R|_U = A + B(I - D)^{-1}C = A + B\sum_{i=0}^{\infty} D^i C$$

是非负的，其中级数展开定义良好，因为 $I - D$ 是可逆的（引理 5.21），且 $\|D\| < 1$，因为 R 是不可约的，并且 A、B、C、D 均为非负，因为 R 是非负的。此外，注意到

$$R|_U v_U = v_U, \quad R^T|_U u_U = u_U.$$

我们证明 $R|_U$，另一等式可直接由上面的 (4) 得出。为证明该等式，注意到 $Lv = 0$，由此得到 $A'v_1 + B'v_2 = 0$ 和 $C'v_1 + D'v_2 = 0$，或者 $v_2 = -D'^{-1}C'v_1$（因为 D' 可逆），代入第一个等式可得 $L|_U v_U = 0$，从而得到 $(I - R|_U)v_U = 0$，正如所需。

6. **半正定拉普拉斯矩阵的闭包**。 如果 L 是半正定的，那么 $L|_U$ 也是。为此，注意到由于 L 是半正定的，对于所有 x，y，都有 $\langle x_1, A'x_1\rangle + \langle x, B'y\rangle + \langle y, B^*x\rangle + \langle y, Dy\rangle \geq 0$。现在令 $y = -D^{-1}B^*x$，由于 L 是厄米的，D 也是厄米的，而且 $BD^{-1}B^*$（事实上也是半正定的；考虑仅支撑在 D 上的向量可知，D 是半正定的，因此 D^{-1} 也是，$BD^{-1}B^*$ 亦然）。因此，对 y 作上述变换可得 $\langle x, Ax\rangle - \langle x, BD^{-1}B^*x\rangle \geq 0$，这表明舒尔补是半正定的。

7.**特征值-特征向量对**。 设 $Rx = \lambda x$，则 $(\lambda, x|_U)$ 是 $(A + B(\lambda I - D)^{-1}C)$ 的一个特征对。这可以再次通过使用方程 5.23、5.24、5.25、5.26 以及上面的第（5）部分轻易得出。

8. **组合解释**。 拉普拉斯矩阵 L 在某个顶点集合 U 上的舒尔补，恰好就是这样一种操作：将 R 中每条边 (i, j) 的权重替换为所有长度的路径的权重之和，这些路径从 i 出发，经过 U，到达 j；也就是说，从 i 到 j 的路径中的所有中间顶点都来自 U（包括空集）。直接考察 $R|_U$ 即可得出这一点:

$$I - L|_U = R|_U = A + B(I - D)^{-1}C = A + B\sum_{i=0}^{\infty} D^i C.$$

这让我们能够从组合意义上深入理解拉普拉斯矩阵的舒尔补。实质上，它是通过从图中移除顶点，并将这些顶点对图中其余部分产生的影响（就不同路径而言）添加回来而得到的。

有了这些性质，我们继续讨论本节的主要问题，并在推导更多引理的过程中，进一步探究舒尔补的若干性质。

5.11 容量与狄利克雷引理

本节将开始探讨另一种组合扩展概念。不同群体曾在不同时期以不同名称研究过这一概念。我们的首要目标是阐述一个统一的框架，全面涵盖所有相关概念和定义，并尽可能从多个角度加以呈现。接下来，我们将为这些量的非可逆版本推导许多引理；据我们所知，这些引理此前尚未以我们所述的形式出现。

我们从另一个与基本问题5.22非常相似的问题开始，称为狄利克雷问题。我们假设 R 是一个不可约非负矩阵，具有正的 u，且 v 是对应于特征值 1 的左右特征向量，并已归一化，使得$\sum_i u_i v_i = 1$。此外，由于 R 不可约，根据引理5.21 可知， $L = I - R$的右核仅包含由 v 张成的空间，其左核仅包含由 u 张成的空间。

请注意，在问题5.22中，我们处理的是已知 b，并希望求出 x 使得 $Lx = b$的情形。现在，我们将问题修改为：已知集合中的一部分取值 b 以及 x，并求出其余取值。这称为狄利克雷问题。

定义 5.28。（狄利克雷问题）设 U 是顶点的一个子集，则给定边界 U、向量 $a \in R^{|U|}$ 和 $b \in R^{|U|}$ 的狄利克雷问题，是求一个向量 $q \in R^n$（称为狄利克雷向量），使得 $q_U = a$且 $(LD_v q)_U = b$。通常， $b = 0$。◊

请注意，在考察 L 对向量 q 的影响之前，我们先将它乘以D_v 进行“归一化”（或“重新加权”）。这与我们处理边扩张定义时的做法类似（参见定义2.4）。狄利克雷问题的直接解法

可直接通过使用类似于5.23、5.24、5.25、5.26的方程得到。这引出了本节最重要的定义。

定义 5.29。（容量）考虑上述狄利克雷问题的情形，其中$b = 0$，也就是说，设 q是满足 $q_U = a$且 $(LD_v q)_U = 0$的某个向量。则我们将边界上取值为 a 的集合 U 的容量定义为

$$\mathrm{cap}_{U,a}(R) = \langle D_u q, L_R D_v q \rangle.$$

当 U 和 a 可从上下文中明确时，我们将记 $\mathrm{cap}_{U,\ a}(R) = \mathrm{cap}(R)$。 ◊

根据定义可直接得到以下结论，它与引理 3.3和 5.3类似。

Le引理 5.30。 设 $A = D^{1/2} R D^{-1/2}$，其中 $D = D_u D_v^{-1}$，则对于任意 U和任意 a,

$$cap_{U,a}(A) = cap_R(U, a)$$

证明。 由定义可立即得证，注意到 u， v， w都是正向量。

□

根据引理 5.30，我们将只讨论 $A = D_u^{1/2} D_v^{-1/2} R D_v^{1/2} D_u^{-1/2}$及其后续情形，因为它具有相同的主左、右特征向量，因而计算更简单明了。

注 5.31. (容量的另一种定义) 容量还可以按如下方式定义。考虑上文狄利克雷问题中带有 $b = 0$ 的情形，即设 q 为某个向量，满足 $q_U = a$ 且 $(LD_v q)_U = 0$。那么，我们已将边界上取值为 a 的集合 U 的容量定义为

$$\mathrm{cap}_{U,a}(R) = \langle D_u q, L_R D_v q \rangle.$$

我们指出两种不会改变容量本质含义的修改。

首先，将 $y = D_v q$ 视为一个向量并施加条件 $y_S = D_v^{-1} a$，即可得到等价的定义。不过，由于字符串 a 是任意的，我们甚至可以写成 $y_S = a$，其中定义 5.29 中的容量 $\mathrm{cap}_{U,a}(R)$ 在新定义中将变为容量 $\mathrm{cap}_{U,D_v^{-1}a}(R)$。因此我们有

$$\mathrm{cap}_{U,a}(R) = \langle D_u D_v^{-1} y, L_R y \rangle.$$

其次，对角矩阵 $D_u D_v^{-1}$只是重新定义了内积，因为我们可以将容量写为

$$\mathrm{cap}_{U,a}(R) = \langle D_u^{1/2} D_v^{-1/2} y, D_u^{1/2} D_v^{-1/2} L_R y \rangle$$

这只是对容量进行重新缩放，因此甚至不需要对角线乘法。因此，我们也可以直接将容量定义为

$$\mathrm{cap}_{U,a}(R) = \langle q, L_R q \rangle$$

其中 $q_U = a$ 且 $(L_R q)_{\overline{U}} = 0$。不过，我们更倾向于使用定义 5.29 而非上述定义，因为它将等价于我们稍后定义的一种新型扩张。

◊

基于上述评论，我们确实可以如下定义具有相同左右特征向量 w 的非负矩阵 A 的容量。

引理 5.32. （容量的较简定义）设 $L = I - A$，其中 A 是一个最大特征值为 1 的不可约非负矩阵，且 w 是其对应的左右特征向量。设 $U \subset [n]$ 和 $a \in \mathbb{R}^{|U|}$，并设 q 满足 $q_U = (D_w)_U a$，以及 $(Lq)_U = 0$。那么

$$cap_{U,a}(A) = cap(A) = \langle q, Lq \rangle.$$

或者，设 q 是一个满足 $q_U = a$ 和 $(Lq)_U = 0$ 的向量，那么对于 $\overline{a} = (D_w^{-1})_U a$，

$$cap_{U,\overline{a}}(A) = cap(A) = \langle q, Lq \rangle$$

证明。 由定义可立即得出这些结论。 □

这就是本节接下来要用到的两个定义。

引理 5.33。 （容量的等价定义）以下都是容量的等价定义。设 $r = L_A D_w q$ 或 $r = L_R D_v q$（具体取哪个由上下文可知）为满足 $r_U = 0$的向量，且 $q_U = a$。那么，

$$\begin{aligned} cap_{U,a}(R) &= \langle D_u q, L_R D_v q \rangle \\ &= \langle (D_u)_U a, L_R|_U (D_v)_U a \rangle \\ &= \langle D_u D_v^{-1} L_R^+ r, r \rangle \\ &= \langle r_U, (L_R^+)_U r_U \rangle \end{aligned}$$

and

$$\begin{aligned} cap_{U,a}(A) &= \langle D_w q, L_A D_w q\rangle \\ &= \langle D_w a, L_A|_U D_w a\rangle \\ &= \langle r, L_A^+ r\rangle \\ &= \langle r_U, (L_A^+)_U r_U\rangle. \end{aligned}$$

证明。 第一个等式由定义得出，第二个等式通过注意到以下事实得到：$(L_A q)_U = 0$，并使用与下列方程类似的方程：5.23、5.24、5.25、5.26；将 $q = D_v^{-1} L_R^+ r$ 或 $q = D_w^{-1} L_A^+ r$ 代入定义即可得到第三个等式。注意到 $r_U = 0$，由舒尔补的性质 3 可得最后一个等式。 □

现在我们来到了第一个重要的引理，不过它针对的是可逆的 R 或对称的 A（见引理 3.3）。

引理 5.34。 （对称 A 的狄利克雷引理）设 $L = I - A$ 为半正定的，并且对于某个固定的 a，令 q 满足 $q_U = a$ 以及 $(Lq)_U = 0$。令 x 为满足 $x_U = a$ 的任意向量。则对于 $a = (D_w^{-1})_U a$，

$$cap_{U,\bar{a}}(A) = \langle q, Lq\rangle \leq \langle x, Lx\rangle.$$

证明。 由于 L 是半正定的，由柯西-施瓦茨不等式立即得到

$$\langle y, Lz\rangle^2 \leq \langle y, Ly\rangle\langle z, Lz\rangle$$

由于 L 是半正定的，且其平方根总是存在，因此

$$\langle y, Lz\rangle^2 = \langle L^{1/2}y, L^{1/2}z\rangle^2 \leq \langle L^{1/2}y, L^{1/2}y\rangle\langle L^{1/2}z, L^{1/2}z\rangle^2.$$

令 $y = D_w x$ 和 $z = D_w q$，则由于 $(Lz)_U = 0$，我们得到 $\langle y, Lz\rangle = \langle z, Lz\rangle$，因为 y 和 z 在属于 U 的指标上相等。将二者结合即得引理，
$\langle z, Lz\rangle^2 = \langle y, Lz\rangle^2 \leq \langle y, Ly\rangle\langle z, Lz\rangle$。 □

这个引理很重要，因为它是许多其他引理的基本步骤。它表明，对于所有满足 $x_U = a$ 的 x，q 实际上是 $\langle x, Lx\rangle$ 的极小化者。这给出了我们关于非可逆矩阵容量的第一个引理。

引理 5.35. （非可逆矩阵容量的下界）对于任意 U 和 a，

$$cap_{U,a}(\tilde{A}) \leq cap_{U,a}(A),$$

其中 $\tilde{A} = \frac{1}{2}(A + A^T)$。

证明。 注意 L和 $\tilde{L} = \frac{1}{2}(L + L^T)$具有相同的左右特征向量 w，对应于˜佩龙–弗罗贝尼乌斯特征值 1。设 q 为 L 的狄利克雷向量，并设 g 为 L 的狄利克雷向量，其中 $q_U = a$ 且˜$g_U = a$。由于 $\langle x, Lx\rangle = \langle x, Lx\rangle$ 对任意实数 x成立，在引理 5.34 中令 $x = q$，可得

$$\langle g, \tilde{L}g\rangle \leq \langle q, \tilde{L}q\rangle = \langle q, Lq\rangle$$

由此得到

$$\mathrm{cap}_{U,\bar{a}}(\tilde{A}) \leq \mathrm{cap}_{U,\bar{a}}(A)$$

这正是所需的；由于 w 为正，因此对所有 a 都成立，而且任意向量都可以表示为 a。 □

事实上，此时我们可以得出比引理 5.35强得多的结论。在继续之前，我们需要理解顶点聚类这一概念。它本质上是对线性代数进行修改，以确保方程的解仍然有效。这一思想的特例曾在许多地方使用，却缺乏充分论证；在此，我们将对其多种用途作出完整说明。

顶点聚类。 基本思路如下。设 $Q \in \mathbb{C}^{n\times n}$ 为任意矩阵，并且不失一般性地假设 $Q = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$，其中 A 和 D 的大小均不为 0。设 A 的支集所包含的顶点为 $U = \{1, ..., r\}$。对于某个向量 $x \in \mathbb{C}^n$，若 $x \in \mathrm{Ker}(Q)$，则 $\langle x, Qx\rangle = 0$。假设 $x \notin \mathrm{Ker}(Q)$。顶点聚类的主要目的是将其写成

$$\langle x, Qx\rangle = \frac{1}{\langle y, Ty\rangle}$$

对于某个独立于 x 和 Q的 y，以及一个可表示为 Q 和 x的函数的 T。这里的主要结果是以下引理。

引理 5.36。 （顶点聚类） 设 $Q = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \mathbb{C}^{n\times n}$，其中 A 支撑于顶点 $U \subset [n]$；给定字符串 $a \in \mathbb{C}^{|U|}$，使得对于任意 $w \in Ker（Q）$ 都有 $a \neq w_U$，以及狄利克雷向量 x，使得 $x_U = a$ 且 $（Qx）_U = \underline{0}$，那么令$x_U = \underline{z}$，并令 $y \in \mathbb{C}^{n-r+1}$ 和 $T \in \mathbb{C}^{(n-r+1)\times(n-r+1)}$，其中

$$y = \begin{bmatrix} 1 \\ z \end{bmatrix} \text{ and } T = \begin{bmatrix} a^*Aa & a^*B \\ Ca & D \end{bmatrix}.$$

Then

$$\langle x, Qx\rangle = \langle y, Ty\rangle = \frac{1}{\langle e_1, T^+e_1\rangle}$$

其中， e_1 是标准基向量。

证明。 注意

$$\begin{aligned}\langle x, Qx\rangle &= \begin{bmatrix} a^* & z^* \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} a \\ z \end{bmatrix} \\ &= a^*Aa + a^*Bz + z^*Ca + z^*Dz \\ &= \begin{bmatrix} 1 & z^* \end{bmatrix} \begin{bmatrix} a^*Aa & a^*B \\ Ca & D \end{bmatrix} \begin{bmatrix} 1 \\ z \end{bmatrix} \\ &= \langle y, Ty\rangle.\end{aligned}$$

因此，如果 $b = Qx$和 $c = Ty$或 $y = T^+c$，令 $c_{\{1\}} = r \neq 0$（因为 x不在 Q 的核中），并且由于 $c_{\{2,\ldots,n-r+1\}} = 0$，我们得到

$$\begin{aligned}\langle a, Ba\rangle + \langle a, Bz\rangle &= c_{\{1\}} = r \\ b_{\{r+1,\ldots,n\}} &= c_{\{2,\ldots,n-r+1\}} = 0 \\ \langle x, Qx\rangle = \langle y, Ty\rangle = \langle Q^+b, b\rangle &= \langle T^+c, c\rangle = r\end{aligned}$$

这里实质上所做的，是针对 x对 Q 中的顶点 U 进行聚类，方法是对与 U对应的簇中的所有边取加权和。由于$Qx|_U = 0$，因此
$\langle x, Qx\rangle = \langle y, Ty\rangle = \langle y, c\rangle = r$，并且还有 $c = r\cdot\ e_1$（其中 e_1是标准基向量），由此得到

$$\langle T^+c, c\rangle = r^2\langle T^+e_1, e_1\rangle$$

由此得到

$$\langle T^+e_1, e_1\rangle = \frac{1}{r} = \frac{1}{\langle x, Qx\rangle}.$$

□

作为顶点聚类的结果，我们得到以下关键引理。这种形式的引理过去虽已在民间流传，但它们存在以下约束。

1. 它们始终针对对称拉普拉斯矩阵

2. 它们的字符串 a 始终是一个位字符串，对应于势为 0 和 1 的顶点

3. 图中有两个顶点簇，分别对应电流流入和流出图的位置；他们没有采用 e_1 引理5.36，而是使用了形如 $e_i - e_j$ 的向量来表示电流。

相比之下，引理 5.36 适用于任意拉普拉斯矩阵（不一定是对称的），只需要一个顶点簇，并且完全摒弃了电学观点。

引理 5.37. 设 A 为任意非负矩阵，其最大特征值为 1，左右特征向量为 w，设 $L = I - A$，则对于任意集合 $U \subset [n]$ 和 $a = (D_w^{-1})_U a$，有

$$L = \begin{bmatrix} A & B \\ C & D \end{bmatrix}, \text{ and } T = \begin{bmatrix} a^* A a & a^* B \\ Ca & D \end{bmatrix},$$

可得

$$cap_{U,\vec{a}}(A) = \frac{1}{\langle e_1, T^+ e_1 \rangle}$$

其中 e_1 是标准基向量。

证明。 证明可直接由引理 5.36中的顶点聚类得出。 □

由此，我们现在可以陈述一个主要引理。

引理 5.38. 设 $R \in \mathbb{C}^{n \times n}$ 为任意矩阵，使得 $\tilde{R} = \frac{1}{2}(R + R^*)$ 为正定矩阵。则有以下结论：

1. R 可逆
2. $\frac{1}{2}(R^{-1} + (R^*)^{-1})$ 是正定的。
3. $$\left(\frac{R + R^*}{2}\right)^{-1} \succcurlyeq \frac{R^{-1} + (R^*)^{-1}}{2}$$
4. $$\frac{R + R^*}{2} \succcurlyeq \frac{R^* R^{-1} R^* + R(R^*)^{-1} R}{2}$$
5. 设 $\tilde{R} = \frac{1}{2}(R + R^*)$ 和 $\bar{R} = \frac{1}{2}(R - R^*)$。设 $W_\alpha = \tilde{R} + \alpha \bar{R}$ 对 $-1 \le \alpha \le 1$ 成立。那么对于任意满足 $|\alpha| \le |\beta|$ 的 α、β，

$$\frac{W_\alpha^{-1} + (W_\alpha^*)^{-1}}{2} \succcurlyeq \frac{W_\beta^{-1} + (W_\beta^*)^{-1}}{2}.$$

证明。 对于 (1)，设 (λ, v) 是 R 的某个特征值—特征向量对，其中 v 非零，则

$$\mathrm{Re}\lambda = \mathrm{Re}\langle v, Rv \rangle = \frac{1}{2}(\langle v, Rv \rangle + \overline{\langle v, Rv \rangle}) = \frac{1}{2}(\langle v, Rv \rangle + \langle Rv, v \rangle) = \langle v, \frac{R + R^*}{2} v \rangle > 0$$

因此 R 没有 0 特征值，且可逆。

对于 (2)，注意，对于任意满足 u 非零的 $u \in \mathbb{C}^n$，

$$\begin{aligned}\left\langle u, \frac{R^{-1} + (R^*)^{-1}}{2} u \right\rangle &= \left\langle u, R^{-1} \cdot \frac{R + R^*}{2} \cdot (R^*)^{-1} u \right\rangle \\ &= \left\langle (R^*)^{-1} u, \frac{R + R^*}{2} \cdot (R^*)^{-1} u \right\rangle \\ &= \left\langle v, \frac{R + R^*}{2} \cdot v \right\rangle \\ &> 0\end{aligned}$$

其中，最后一个不等式成立是因为 $u \neq 0$，并且 $(R^*)^{-1} u$ 也非零，因为 R^*（由 (1) 可知），因此 $(R^*)^{-1}$ 可逆，故其秩满，说明其核是平凡的。

对于 (3)，令 $L = \tilde{R} = \frac{1}{2}(R + R^*)$ 和 $Q \equiv R = \frac{1}{2}(R - R^*)$（为使证明更易读，我们定义了这些新符号）。注意，$L = L^*$，且由于 L 是正定的，L^{-1} 存在且定义良好。此外，Q 是斜厄米的，即$Q^* = -Q$。我们还可以将 $R = L + Q$ 和 $R^* = L - Q$改写为。于是我们有

$$\left\langle u, \left(\frac{R+R^*}{2}\right)^{-1} u \right\rangle - \left\langle u, \frac{R^{-1}+(R^*)^{-1}}{2} u \right\rangle = \left\langle u, L^{-1}u \right\rangle - \left\langle u, R^{-1} \cdot L \cdot (R^*)^{-1}u \right\rangle$$

$$= \left\langle u, L^{-1}u \right\rangle - \left\langle (R^*)^{-1}u, L \cdot (R^*)^{-1}u \right\rangle$$

$$= \left\langle R^*v, L^{-1}R^*v \right\rangle - \langle v, Lv \rangle$$

[setting $u = R^*v$]

$$= \left\langle (L-Q)^*v, L^{-1}(L-Q)^*v \right\rangle - \langle v, Lv \rangle$$

$$= \left\langle v, (L-Q)L^{-1}(L-Q)^*v \right\rangle - \langle v, Lv \rangle$$

$$= \left\langle v, (L-Q)L^{-1}(L+Q)v \right\rangle - \langle v, Lv \rangle$$

[since $(L-Q)^* = L^* - Q^* = L+Q$]

$$= \left\langle v, (L-Q+Q-QL^{-1}Q)v \right\rangle - \langle v, Lv \rangle$$

$$= \left\langle v, -QL^{-1}Qv \right\rangle$$

$$= \left\langle v, Q^*L^{-1}Qv \right\rangle$$

[since $Q^* = -Q$]

$$= \langle Qv, L^{-1}Qv \rangle$$

$$\geq 0$$

[since L^{-1} is positive definite]

对于 (4)，从 (3) 出发，我们有

$$\begin{aligned}
&\frac{1}{2}\left(\left(\frac{R+R^*}{2}\right)^{-1}+\left(\frac{R+R^*}{2}\right)^{-1}\right)\geqslant\frac{R^{-1}+(R^*)^{-1}}{2}\\
\Leftrightarrow&\left(\frac{R+R^*}{2}\right)^{-1}L\left(\frac{R+R^*}{2}\right)^{-1}\geqslant R^{-1}L(R^*)^{-1}\\
\Leftrightarrow&L\geqslant LR^{-1}L(R^*)^{-1}L\\
\Leftrightarrow&L\geqslant\left(\frac{R+R^*}{2}\right)R^{-1}\left(\frac{R+R^*}{2}\right)(R^*)^{-1}\left(\frac{R+R^*}{2}\right)\\
\Leftrightarrow&L\geqslant\frac{1}{8}\left(I+R^*R^{-1}\right)(R+R^*)\left((R^*)^{-1}R+I\right)\\
\Leftrightarrow&L\geqslant\frac{1}{8}\left(R+R^*+R^*+R^*R^{-1}R^*\right)\left((R^*)^{-1}R+I\right)\\
\Leftrightarrow&L\geqslant\frac{1}{8}\left(R+R^*+R^*+R^*R^{-1}R^*+R(R^*)^{-1}R+R+R+R^*\right)\\
\Leftrightarrow&L\geqslant\frac{3}{4}L+\frac{1}{4}\cdot\frac{R^*R^{-1}R^*+R(R^*)^{-1}R}{2}\\
\Leftrightarrow&\frac{R+R^*}{2}\geqslant\frac{R^*R^{-1}R^*+R(R^*)^{-1}R}{2}
\end{aligned}$$

如所需。

对于 (5)，注意 $W_\alpha^*=W_{-\alpha}$。在上面的 (3) 中，我们证明了

$$W_0^{-1}\geqslant\frac{W_1^{-1}+W_{-1}^{-1}}{2}.$$

现在我们希望推广这个不等式。首先注意，对于每个 α，$\frac{1}{2}(W_\alpha+W_\alpha^*)=L$ 因此由 (1) 可知，W_α 对每个 α 都是可逆的。为了证明该不等式，令 $K=W_\beta$，则

$$W_\alpha=\frac{K+K^*}{2}+\frac{\alpha}{\beta}\frac{K-K^*}{2}=\frac{1}{2}\left(1+\frac{\alpha}{\beta}\right)K+\frac{1}{2}\left(1-\frac{\alpha}{\beta}\right)K^*=cK+(1-c)K^*$$

for

$$0\leq c=\frac{1}{2}\left(1+\frac{\alpha}{\beta}\right)\leq 1.$$

于是我们有

$$\begin{aligned}
&\frac{W_\alpha^{-1}+(W_\alpha^*)^{-1}}{2} \succcurlyeq \frac{W_\beta^{-1}+(W_\beta^*)^{-1}}{2}\\
\Leftrightarrow & W_\alpha^{-1}\frac{W_\alpha+W_\alpha^*}{2}(W_\alpha^*)^{-1} \succcurlyeq K^{-1}\frac{K+K^*}{2}(K^*)^{-1}\\
\Leftrightarrow & L \succcurlyeq W_\alpha K^{-1}L(K^*)^{-1}W_\alpha^*\\
\Leftrightarrow & L \succcurlyeq (cK+(1-c)K^*)\,K^{-1}L(K^*)^{-1}(cK^*+(1-c)K)\\
\Leftrightarrow & L \succcurlyeq \left(cI+(1-c)K^*K^{-1}\right)L\left(cI+(1-c)\,(K^*)^{-1}K\right)\\
\Leftrightarrow & L \succcurlyeq \left(cI+(1-c)K^*K^{-1}\right)L\left(cI+(1-c)\,(K^*)^{-1}K\right)\\
\Leftrightarrow & L \succcurlyeq \left(c^2L+(1-c)^2K^*K^{-1}L(K^*)^{-1}K+c(1-c)\,L(K^*)^{-1}K+c(1-c)K^*K^{-1}L\right)\\
\Leftrightarrow & L \succcurlyeq \left(c^2+(1-c)^2\right)L+c(1-c)\,\left(L(K^*)^{-1}K+K^*K^{-1}L\right)\\
& \quad [\text{因为 } K^*K^{-1}L(K^*)^{-1}K=L]\\
\Leftrightarrow & L \succcurlyeq \left(c^2+(1-c)^2\right)L+c(1-c)L+c(1-c)\,\left(K(K^*)^{-1}K+K^*K^{-1}K^*\right)\\
& \quad [\text{展开 } L=\frac{1}{2}(K+K^*)]\\
\Leftrightarrow & (c+(1-c))^2\,L \succcurlyeq \left(c^2+(1-c)^2\right)L+c(1-c)L+c(1-c)\,\left(K(K^*)^{-1}K+K^*K^{-1}K^*\right)\\
\Leftrightarrow & L \succcurlyeq K(K^*)^{-1}K+K^*K^{-1}K^*\\
& \quad [\text{since } 0\le c\le 1]\\
\Leftrightarrow & \frac{K+K^*}{2} \succcurlyeq K(K^*)^{-1}K+K^*K^{-1}K^*
\end{aligned}$$

根据 (4)，这是真的，因为 $\frac{1}{2}(K+K^*)=L$ 是正定的。 □

作为这些引理的推论，我们得到了这篇学位论文的第二个主要定理。

定理 5.39. *(容量单调性)* 设 A 是任意一个最大特征值为 *1* 的非负矩阵，其左右特征向量为 w，设 $L=I-A$，并设 $L_\alpha=L+\alpha L$对于某个 $-1\le\alpha\le\overline{1}$，其中 $\tilde{L}=\frac{1}{2}(L+L^T)$且 $L=\frac{1}{2}(L-L^T)$，并类似地定义A_α、 A 和 A。那么对于 $|\alpha|\le|\overline{\beta}|$，我们有

$$cap_{U,a}(A_\alpha)\le cap_{U,a}(A_\beta).$$

证明。 设 α 和 β 如上所述。设

$$L=\begin{bmatrix}A & B\\ C & D\end{bmatrix},\ \text{and } H=\begin{bmatrix}a^*Aa & a^*B\\ Ca & D\end{bmatrix},$$

其中 $\tilde{H} = \frac{1}{2}(H + H^T)$ 和 $\overline{H} = \frac{1}{2}(H - H^T)$ 的定义类似于 $\tilde{L}$和 $L_\circ$ 注意

$$L_\alpha = \begin{bmatrix} \frac{1+\alpha}{2}A + \frac{1-\alpha}{2}A^T & \frac{1+\alpha}{2}B + \frac{1-\alpha}{2}C^T \\ \frac{1+\alpha}{2}C + \frac{1-\alpha}{2}B^T & \frac{1+\alpha}{2}D + \frac{1-\alpha}{2}D^T \end{bmatrix},$$

并令

$$\begin{aligned} H_\alpha &= \begin{bmatrix} a^*(\frac{1+\alpha}{2}A + \frac{1-\alpha}{2}A^T)a & a^*(\frac{1+\alpha}{2}B + \frac{1-\alpha}{2}C^T) \\ (\frac{1+\alpha}{2}C + \frac{1-\alpha}{2}B^T)a & \frac{1+\alpha}{2}D + \frac{1-\alpha}{2}D^T \end{bmatrix} \\ &= \frac{1+\alpha}{2}\begin{bmatrix} a^*Aa & a^*B \\ Ca & D \end{bmatrix} + \frac{1-\alpha}{2}\begin{bmatrix} a^*A^Ta & a^*C^T \\ B^Ta & D^T \end{bmatrix} \\ &= \tilde{H} + \alpha\overline{H}. \end{aligned}$$

由于 $a \neq w_U$ 其中 w 是 L_A的核中的唯一向量，对于任意不属于 w的张成空间的向量 x， H_α 对于任意 α都是可逆的。因此，根据引理5.38的第（5）部分，我们得到

$$\frac{H_\alpha^{-1} + (H_\alpha^*)^{-1}}{2} \geqslant \frac{H_\beta^{-1} + (H_\beta^*)^{-1}}{2},$$

并且由于 H_α 是实数，我们有

$$\langle e_1, H_\alpha^{-1} e_1\rangle = \langle e_1, (H_\alpha^*)^{-1} e_1\rangle$$

从而

$$\langle e_1, H_\alpha^{-1} e_1\rangle \geq \langle e_1, H_\beta^{-1} e_1\rangle$$

并且由于

$$\mathrm{cap}_{U,\overline{a}}(A_\alpha) = \frac{1}{\langle e_1, H_\alpha^{-1} e_1\rangle}$$

根据引理5.37 （注意，我们是在核之外的空间中进行运算，因此$H_\alpha^{-1} = H_\alpha^+$），由此得到该定理，因为字符串 a 是任意的，因为 w 是正的。

□

这为我们充分细致地理解非对称非负矩阵的容量提供了依据，因为它告诉我们，随着矩阵偏离对称性，容量会严格增加；即使比较两个非对称矩阵，偏离对称性更远的那个（按照定理 5.39的意义）也严格具有更高的容量。

我们还想指出，由于定理 5.39比较的是非对称矩阵的两个内积，因此这一结果极为深刻，因为对称矩阵的所有工具都不再适用。

我们的下一个目标是证明一个类似于 5.34 的狄利克雷引理，不过针对的是非对称非负矩阵。Slowik [Slo12]相对近期证明了这一结果的一部分；然而，那里的结果将容量表示为不同表达式的下确界和上确界，且没有给出显式解。我们将在此给出这些引理的显式表达式。[1]

引理 5.40。（非对称矩阵的狄利克雷引理）设 $L = I - A$ 为某个不可约非负矩阵，其最大特征值为 1，对应的左右特征向量为 w。设 q 满足 $q_U = a \neq w_U$，且 $(Lq)_U = 0$。设 x 为满足 $x_U = a$的向量。那么

$$cap_{U,\bar{a}}(A) = \langle q, Lq\rangle \leq \langle x, L\tilde{L}^+L^T x\rangle,$$

并且当 $U = \{s,t\}$ 满足 $a_s = 1$ 和 $a_t = 0$时，不等式取等号。

证明。注意到 $\langle x, Lq\rangle = \langle q, Lq\rangle$，因为 $(Lq)_U = 0$；又由于所有量均为实数，$\langle q, Lq\rangle = \langle q, \tilde{L}q\rangle$。根据柯西-施瓦茨不等式，对于 $z = \tilde{L}^+L^T x$，有

$$\langle q, \tilde{L}z\rangle^2 \leq \langle q, \tilde{L}q\rangle\langle z, \tilde{L}z\rangle$$

or

$$\langle q, L^T x\rangle^2 \leq \langle q, \tilde{L}q\rangle\langle x, L\tilde{L}^+L^T x\rangle$$

这就得到了引理。

令 $x = (L^T)^+\tilde{L}q$，则注意到

$$\langle (L^T)^+\tilde{L}q, L\tilde{L}^+L^T(L^T)^+\tilde{L}q\rangle = \langle q, \tilde{L}L^+L\tilde{L}^+L^T(L^T)^+\tilde{L}q\rangle = \langle q, \tilde{L}q\rangle = \langle q, Lq\rangle = \mathrm{cap}_{U,\bar{a}}(A).$$

请注意，唯一的难点在于验证 $x_U = a$。为此，我们首先假设集合 $U = \{s,t\}$ 和 $a_s = 1$ 以及 $a_t = 0$。接着注意，若 q 是这个特定的 U和 a对应的狄利克雷向量，那么由于 $(Lq)_U = 0$且 $L^T w = 0$，令 $Lq = b$，可得 $w_s b_s + w_t b_t = 0$。类似地，令 p为 L^T对应的狄利克雷向量，并且由于 $(L^T p)_U = 0$ 且 $Lw = 0$，令 $L^T p = c$，同样可得$w_s c_s + w_t c_t = 0$。但是，由于 $\langle q, Lq\rangle = \langle p, L^T p\rangle$，且 L和 L^T的容量相同（使用 5.33 或性质 (4)），因此有 $b_s = c_s$，从而得到 $b_t = c_t$，于是

[1] Slowik 中关于显式表达式有一条评论 [Slo12] 但这是不正确的。

$b = c$。由此可见，我们的 x 变为如下形式。

$$\begin{aligned} x &= (L^T)^+ \tilde{L} q \\ &= (L^T)^+ \cdot \frac{1}{2}(L + L^T)q \\ &= \frac{1}{2}q + \frac{1}{2}(L^T)^+ Lq \\ &= \frac{1}{2}q + \frac{1}{2}(L^T)^+ b \\ &= \frac{1}{2}q + \frac{1}{2}(L^T)^+ c \\ &= \frac{1}{2}q + \frac{1}{2}p \end{aligned}$$

并且由于 p 和 q 在 U上相同， x也如此，因此该引理是紧的。 □

我们指出，引理 5.40 是引理5.34的一般化；若 L 是对称矩阵，则后者是前者的一个推论。引理 5.40 的一个显著之处在于，我们可以利用矩阵 $H = L\tilde{L}^+ L^T$来表示任意矩阵 L的容量。我们认为这个矩阵本身就很值得研究。下面给出一个引理，以展示它的用处。

引理 *5.41*。（引理 *5.35* 的加强）设 $L = I - A$ 是某个最大特征值为 *1*、相应的左右特征向量为 w的不可约非负 A 。设 $U = \{s, t\}$，其中 $a_s = 1$ 且 $a_t = 0$；再设 q 满足$q_U = a \neq w_U$且 $(Lq)_U = 0$，并类似地设 p为 L^T的狄利克雷向量。设$a = (D_w^{-1})_U a$且 $H = L\tilde{L}^+ L^T$。那么

$$\langle p, Hp \rangle = \frac{\left(cap_{U,\bar{a}}(A)\right)^2}{cap_{U,\bar{a}}(\tilde{A})}.$$

˜证明。我们首先注意以下事实。首先，令 $\text{cap}_{U,a}(A) = c$ 以及$\text{cap}_{U,a}(A) = \tilde{c}$。类似于引理 5.40 中关于紧性的证明，令 p 为L^T的狄利克雷向量，则我们知道 $Lq = L^T p = b$，并且知道 $b_s = \text{cap}(A) = c$ 以及$b_t = -c \cdot w_s / w_t$。因此，类似于顶点聚集引理，令 μ 为满足 $\mu_s = 1$、$\mu_t = -w_s/w_t$，且在其他情况下为 $\mu_i = 0$ 的向量。于是我们有

$$\begin{aligned} c &= \langle q, Lq \rangle = \langle b, L^+ b \rangle = c^2 \langle \mu, L^+ \mu \rangle \\ \frac{1}{c} &= \langle \mu, L^+ \mu \rangle \end{aligned}$$

类似地，我们有

$$\langle \mu, \tilde{L}^{+}\mu \rangle = \frac{1}{\tilde{c}}.$$

现在考虑

$$\begin{aligned}\langle p, L\tilde{L}^{+}L^{T}p \rangle &= \langle b, \tilde{L}^{+}b \rangle \\ &= c^{2}\langle \mu, \tilde{L}^{+}\mu \rangle \\ &= \frac{c^{2}}{\tilde{c}}\end{aligned}$$

正如所声称的。 □

5.12 归一化容量与有效电导

现在考虑这样的特殊情形： $U = S \cup T$ 具有互不相交的 S 和 T，并且 $a_S = 1$ 和 $a_T = 0$。在这种情况下，我们将 $\mathrm{cap}_{U,a}(A) = \mathrm{cap}_{S,T}(A)$ 称为集合 S 与 T 之间的有效电导，而 $1/\mathrm{cap}(S,T)$ 通常称为 S 与 T 之间的有效电阻。事实上，我们可以对容量进行归一化，从而得到归一化容量的定义。

定义 5.42.（归一化容量） 设 A 是一个最大特征值为 1、对应的左右特征向量为 w 的不可约非负矩阵，且 $L = I - A$ 是相应的拉普拉斯矩阵。以下量称为归一化容量。

$$\sigma_A = \min_{\langle 1_S, D_w^2 1_S \rangle \le \langle 1_T, D_w^2 1_T \rangle} \frac{\mathrm{cap}_{S,T}(A)}{\langle 1_S, D_w^2 1_S \rangle}.$$

为便于说明，注意对于双随机 A,

$$\sigma_A = \min_{|S| \le |T|} \frac{\mathrm{cap}_{S,T}(A)}{|S|}$$

◊

从组合角度理解这一概念的另一种截然不同的方式，是借助舒尔补的性质（8）中讨论的组合视角。这将为理解归一化容量提供一个优美的视角。

引理 5.43。 设 A 是一个不可约非负矩阵，其最大特征值为 1，对应的左右特征向量为 w， $L = I - A$ 为相应的拉普拉斯矩阵。那么

$$\sigma_A = \min_{\langle 1_S, D_w^2 1_S \rangle \le \langle 1_T, D_w^2 1_T \rangle} \frac{cap_{S,T}(A)}{\langle 1_S, D_w^2 1_S \rangle} = \min_{U} \phi(A|_U)$$

证明。 注意引理5.33,

$$\begin{aligned}
\mathrm{cap}_{S,T}(A) &= \langle a, (D_w)_U L|_U (D_w)_U a\rangle \\
&= \langle 1_S, (D_w)_U (I - A|_U)(D_w)_U 1_S\rangle \\
&= \langle 1_S, (D_w^2)_U 1_S\rangle - \langle 1_S, (D_w)_U A|_U (D_w)_U 1_S\rangle \\
&= \langle 1, (D_w^2)_U 1_S\rangle - \langle 1_S, (D_w)_U A|_U (D_w)_U 1_S\rangle \\
&= \langle 1, (D_w)_U A|_U (D_w)_U 1_S\rangle - \langle 1_S, (D_w)_U A|_U (D_w)_U 1_S\rangle \\
&\quad \text{[since } A_u^T w_U = w_U \text{ from properties (5) and (7) of Schur Complements]} \\
&= \langle 1_{\overline{S}}, (D_w)_U A|_U (D_w)_U 1_S\rangle
\end{aligned}$$

因此

$$\frac{\mathrm{cap}_{S,T}(A)}{\langle 1_S, D_w^2 1_S\rangle} = \frac{\langle 1_{\overline{S}}, (D_w)_U A|_U (D_w)_U 1_S\rangle}{\langle 1_S, D_w^2 1_S\rangle} = \phi_S(A|_U)$$

先对 S 取最小值，再对 U 取最小值，即得所需结果。 □

引理 5.43 现在帮助我们对（归一化）容量的概念获得组合性理解，这与舒尔补的性质（8）类似，因为它恰好 是移除图中某些顶点后的边扩张。我们先证明一个简单的引理，然后给出两种不同的解释。

引理 5.44。 （最大值原理）设 $L = I - A$，其中 A 是最大特征值为 1、对应左右特征向量为 w 的不可约非负矩阵。设 $a \in \{0、1\}^{|U|}$，其中 $U = S \cup T$ 且 $a_S = 1$ 和 $a_T = 0$。设 q 为向量，满足 $q_U = (D_w)_U a$ 且 $(Lq)_U = 0$。则对于 q 的每个元素，

$$0 \le q_i \le w_i.$$

此外，对于所有 $i \in S$ 和 $j \in T$,

$$(Lq)_i \ge 0,\ (Lq)_j \le 0.$$

证明。 该结论对矩阵元素 $i \in S \cup T$成立，我们需要证明它对矩阵元素$i \in U$也成立。令 $A = \begin{bmatrix} P & Q \\ R & W \end{bmatrix}$为不可约非负矩阵，且令 $L = I - A$，则由

$(Lq)_{\overline{U}} = 0$，可得

$$\begin{aligned} -Rq_{\overline{U}} + (I - W)q_U &= 0 \\ q_{\overline{U}} &= (I - W)^{-1}Rq_U \\ &= \sum_{i \geq 0} W^i R q_U \\ &= \sum_{i \geq 0} W^i R (D_w)_U a \end{aligned}$$

其中，由于引理 5.21，$I - S$ 是可逆的。因此，由上面的等式可知，由于 W、R、w 和 a 均为非负，我们得到对所有 $i \in U$，$q_i \geq \overline{0}$。为了说明这个上界，首先注意到，随着 a 中 1 的数量增加，q_U 的矩阵元素不会减小，这同样是因为涉及的所有矩阵都是非负的。因此，当 a 为全一向量时，q_U 中的最大矩阵元素便能取到。但在这种情况下，注意 $q_U = w_U$，而满足 $(Lq)_U = 0$ 的向量 q 恰好是 w，这将意味着 $q_U = w_U$，从而表明对于任意 $i \in U$，都有 $q_i \leq w_i$。

为了证明对于 $i \in S$，$(Lq)_i \geq 0$ 成立，首先令 $b = (Lq)_U$，则

$$b = ((I - P) - Q(I - W)^{-1}R)(D_w)_U a.$$

然而，根据舒尔补的性质 (5)，我们知道，取舒尔补会得到另一个非负矩阵的拉普拉斯矩阵，具体而言，$A' = P + Q(I - W)^{-1}R$，其中 w_U 是对应于特征值 1 的左、右佩龙–弗罗贝尼乌斯特征向量，因此我们有

$$b = (I - A')(D_w)_U a.$$

现在我们知道，a 是全一向量 $b = 0$。因此，假设 a 在前 $|S|$ 个矩阵元素中取值为 1，那么对于 $i \in S$，

$$b_i = (1 - A'_{i,i})w_i - \sum_{j \in S, j \neq i} A'_{i,j} w_j,$$

由于 A' 非负，当且仅当对 $A'_{i,\,j}$ 的整行求和时，$b_i = 0$ 恰好成立，因此我们得到 $b_i \geq 0$。类似地，对于 $i \in T$，我们可以写成

$$b_i = -\sum_{j \in T} A_{i,j} w_j \leq 0$$

如所需。 □

与引理 5.43 中舒尔补的组合解释和性质（8）类似，现在我们也可以对容量 $\mathrm{cap}(S,\ T)$ 给出类似的解释。然而，为了

从直观上真正理解容量，我们将为双随机矩阵 A 写出这种解释。由于它具有均匀的佩龙-弗罗贝尼乌斯特征向量，且对角矩阵变为单位矩阵，因此无需通过 D_w 进行重新缩放。我们可以为完全一般的矩阵写出这种解释，但那样并无多大意义。对于双随机矩阵，这种解释一目了然。

引理 5.45。 （从概率角度理解容量）设 A 是一个拉普拉斯矩阵为 $I - A$ 的双随机矩阵。设 S 和 T 是两个集合，满足 $U = S \cup T$，并设 q 是满足以下条件的（狄利克雷）向量： $q_i = 1$、 $q_j = 0$ 对于 $i \in S$、 $j \in T$，且$(Lq)_{S \cup T} = 0$。设 概率$_{S,\,T}(i)$ 表示从 i 出发的随机游走在到达集合 T 之前到达集合 S 的概率。设 v 是满足 $v_i =$ 概率$_{S,\,T}(i)$ 的向量。那么 $v = q$。

证明。 证明可由以下内容直接得到。对于 $i \in S$， $v_i = 1$ ；对于$j \in T$， $v_j = 0$，因此我们需要计算 $i \in U$ 的 v_i。设 $A = \begin{bmatrix} P & Q \\ R & W \end{bmatrix}$，其中我们按照向量右乘来考察矩阵元素，也就是说，$A_{j,\,i}$ 表示从顶点 i 到 j 的概率。接下来注意，从顶点 $i \in U$ 出发，在恰好 t 步内到达集合 S，且先于集合 T 的概率恰好为

$$(W^t R q_U)_i.$$

因此，在 S 之前到达 T 的总概率为

$$\sum_{t \geq 0} (W^t R q_U)_i$$

这恰好等于q_U。 □

在给出第二种解释之前，我们注意到，也可以再次将 S 和 T 中的顶点合并为单个顶点 s' 和 t'，并添加所有相应的边，方法与引理 5.36类似。不过请注意，引理5.36 更简洁且更具一般性，因为它将所有顶点合并为一个超级顶点，而且不需要取值 1 和 0。

引理 5.46。 考虑引理 *5.36*中的情形。令 L' 为一个矩阵，其中所有对应于 S 的行和列都已求和合并为单独的一行/列；对于 T也同样处理，从而在底层图中创建新的顶点 s' 和 t' （此时该图不再是双随机的）。那么

$$cap_{S,T}(A) = cap_{s',t'}(A)$$

证明。 可以将各行相加，因为这只是对线性方程求和；也可以将各列相加，因为 v 在 S的每个顶点处都取相同的值 1，并且在 T的每个顶点处都同样取值 0。由于线性方程没有变化，解也不会改变。

□

由此，对于特定情形 $|S| = |T|$，我们现在可以用期望来解释。

引理 5.47。 （用期望解释容量）设 $L = I - A$ 是不可约双随机 A，且满足 $|S| = |T| = 1$，并将顶点 s 和 t 称为 .....。设 q 是满足以下条件的唯一向量：$(Lq)_s = 1$，$(Lq)_t = -1$，并且 Lq 在其他所有位置均为零，且 $q_t = 0$。那么

$$q_s = \frac{1}{cap_{s,t}(A)},$$

and

$$q_i = E_{s,t}(i),$$

其中，$E_{s,t}(i)$ 表示从 s 出发的随机游走在到达 t之前访问顶点 i 的期望访问次数。

证明。为完整起见，我们给出详细证明。设 s and t 固定。设 X_u是在从 s 出发并在到达 t时停止的随机游走中， u 被访问的次数；类似地，对于任意边 $i \to j$，$Y_{i,j}$ 是在从 s 出发并在 t结束的随机游走中，该边被访问的次数。于是 $X_u = \sum_i Y_{i,u}$。现在固定某个顶点 k，并考虑 $Y_{k,u}$。设 X_k的分布为 r_j，对任意 $j \in \mathbb{N}$均成立。因此 $\sum_j r_j = 1$。现在固定某个 j，并设 X_k恰好被访问 j次。设 Z_1，.....， Z_j为随机变量，使得当且仅当游走在第 t步从顶点 k 走到顶点 u时， $Z_t = 1$，因此 $\mathbb{E}(Z_t) = p_{k,u}$。那么，在已知 $X_k = j$的条件下，我们

有 $Y_{k,\ u} = \sum_{t=1}^{t-j} Z_{t_0}$。因此，$\mathbb{E}(Y_{k,\ u}) = j \cdot p_{k,\ u_0}$。因此我们可以写成

$$\begin{aligned}
\mathbb{E}(Y_{k,u}) &= \sum_j \mathbb{E}(Y_{k,u} | X_k = j) \cdot \Pr[X_k = j] \\
&= \sum_{j=1}^{\infty} \mathbb{E}\left(\sum_{t=1}^{j} Z_t^{(j)} \right) \cdot r_j \\
&= \sum_{j=1}^{\infty} \sum_{t=1}^{j} \mathbb{E} Z_t^{(j)} \cdot r_j \\
&= \sum_{j=1}^{\infty} j \cdot p_{k,u} \cdot r_j \\
&= p_{k,u} \cdot \mathbb{E}(X_k)
\end{aligned}$$

因此

$$\begin{aligned}
\mathbb{E}(X_u) &= \sum_i \mathbb{E}(Y_{i,u}) \\
&= \sum_i p_{i,u} \mathbb{E}(X_i)
\end{aligned}$$

对每个 u，除 s 和 t 外。由于我们从不访问 t，而且总会至少访问一次 s，因此有

$$\begin{aligned}
\mathbb{E}(X_t) &= 0 \\
\mathbb{E}(X_s) &= 1 + \sum_i \mathbb{E}(X_i) \cdot p_{i,s}
\end{aligned}$$

注意，除 t 外所有顶点对应的方程完全确定了 t对应的方程，其结果为

$$\mathbb{E}(X_t) = -1 + \sum_i \mathbb{E}(X_i) \cdot p_{i,s}.$$

注意，这些恰好是针对$i \neq s, t$和 $q_t = 0$的方程 $(Lq)_{\ s} = 1$、$(Lq)_{\ t} = -1$、$(Lq)_{\ i} = 0$。由于 A 不可约且核的维数为 1，令$q_t = 0$，便得到满足该方程的唯一向量。因此，

$$q_u = \mathbb{E}(X_u)$$

按要求。现在我们继续求解方程。由 $L = I - A$，由$\chi_{s,t} = 1_s - 1_t$，可得

$$Lq = \chi_{s,t},$$

其解为

$$q = L^+\chi_{s,t} + c \cdot \mathbf{1}$$

并且在约束 $v_t = 0$ 下，得到

$$q = L^+\chi_{s,t} - \langle 1_t, L^+\chi_{s,t}\rangle \cdot \mathbf{1}.$$

注意，

$$q_s = q_s - q_t = \langle \chi_{s,t}, q\rangle = \langle \chi_{s,t}, L^+\chi_{s,t}\rangle = \frac{1}{\mathrm{cap}_{s,t}(A)}.$$

其中，最后一个等式成立是因为取向量 $v = q/q_s$ 时，我们有 $v_s = 1, v_t = 0$，并且对于 $i \neq s$， t，有 $(Lv)_i = 0$，因此

$$\mathrm{cap}_{s,t}(A) = \langle v, Lv\rangle = \frac{1}{q_s^2}\langle q, Lq\rangle = \frac{1}{q_s^2}\langle \chi_{s,t}, L\chi_{s,t}\rangle = \frac{1}{q_s^2}q_s = \frac{1}{q_s}.$$

□

事实上，对于对称矩阵 A，引理 5.45 中的概率解释可以用电压来表述，而引理5.47中的期望解释可以用电流来表述，但我们更倾向于采用数学解释，以避免对现实作出假设。

现在，我们的目标是进一步探究归一化容量的概念，即 σ_A。为此，注意到我们可以立即将定理5.39作为推论，得到以下引理。

引理 5.48。 *(归一化容量的单调性)* 设 A 是任意一个最大特征值为 *1*、左右特征向量为 w 的非负矩阵，设 $L = I - A$，并设对于某个 $-1 \leq \alpha \leq 1$，有 $A_\alpha = \tilde{A} + \alpha\bar{A}$，其中 $\tilde{A} = \frac{1}{2}(A + A^T)$ 且 $\bar{A} = \frac{1}{2}(A - A^T)$。那么对于 $|\alpha| \leq |\beta|$，有

$$\sigma_{A_\alpha} \leq \sigma_{A_\beta}.$$

证明。 与引理5.39的证明相同。 □

在理解了 σ_A之后，一个有趣的问题是将 σ_A 与 A的谱隙联系起来。事实证明，归一化容量与谱隙在相差常数因子的意义下等价，如 Schild [Sch18]和 Miller 等人 [MWW18]所示。为完整起见，我们给出证明，主要遵循 [MWW18], 的论证，并做了一些修改，使其更简短、更流畅，随后讨论非对称情形下可能的版本。

继续之前，我们先给出一个引理序列。

引理 5.49. $\forall x \in \mathbb{R}^n$, $(x_i - x_j)^2 \le \sum_{k=i}^{j-1} \frac{1}{p_k} (x_k - x_{k+1})^2$ 对于 $\sum p_k \le 1$。

证明。 利用柯西-施瓦茨不等式,

$$(x_i - x_j)^2 = \left(\sum_k \sqrt{p_k} \cdot \frac{x_k - x_{k+1}}{\sqrt{p_k}} \right)^2 \le \sum_{k=i}^{j-1} \frac{1}{p_k} (x_k - x_{k+1})^2.$$

□

引理 5.50。 设 $w_1, ..., w_{n-1}$ 为任意非负权重，且 $f_w(x) = \sum_{k=1}^{n-1} w_k (x_k - x_{k+1})^2$。那么，对于任意满足 $i < j$ 的(i, j)，存在一个满足 $x_k = 1$ 的 x，使得对于 $k \le i, x_k = 0$，有 $k \ge j$，且 $f_w(x) = \left(\sum_{k=i}^{j-1} \frac{1}{w_k} \right)^{-1}$。

证明。 令 x在矩阵元素 $\le i$上取 1，在矩阵元素 $\ge j$上取 0；对于 $i \le k \le j$，令 $x_k = \frac{\sum_{r=k}^{j-1} \frac{1}{w_r}}{\sum_{r=i}^{j-1} \frac{1}{w_r}}$. Then

$$f_w(x) = \sum w_k (x_k - x_{k+1})^2 = \sum_{k=i}^{j-1} w_k (x_k - x_{k+1})^2 = \sum_{k=i}^{j-1} w_k \frac{\frac{1}{w_k^2}}{\left(\sum_{r=i}^{j-1} \frac{1}{w_r} \right)^2} = \left(\sum_{k=i}^{j-1} \frac{1}{w_k} \right)^{-1}$$

如所需。 □

现在我们可以给出 [MWW18 中的主要引理；Sch18]。为简便起见，我们对双随机矩阵证明该结论。

引理 5.51。 *(归一化容量与谱隙 [MWW18；Sch18])* 设 A是一个对称不可约双随机矩阵，其最大特征值为 *1*，对应的左右特征向量为 **1**。那么， A 的归一化容量与 A的谱隙在相差常数因子的意义下等价。定量地说，

$$\frac{1}{2} \cdot \Delta(A) \le \sigma_A \le 4 \cdot \Delta(A).$$

证明。 我们先证明下界，这很直接。只需证明 $\lambda_2(L) \le \lambda_2(L|_U)$对任意 U 均成立，而该不等式可由 Cheeger-Buser 不等式及等价关系 5.43推出。

设 $z = [x, y]^T$，其中 $(Lz)(j) = 0$对 $j \in M$成立。令 $x \perp 1$为满足 $\lambda_2(L|_M) = \dfrac{\langle x, L|_M x\rangle}{\langle x, x\rangle}$的向量。因此，存在一个 y，使得对于 $z = [x, y]^T$，$\langle z, Lz\rangle = \langle x, L|_M x\rangle$其中 Lz在 M的任意顶点处均为 0。令 $c_0 = \langle z, 1\rangle/n$，且 $z' = z - c_0 1$，则$z' \perp 1$ 和 $\langle z, Lz\rangle = \langle z', Lz'\rangle$，并且

$$\begin{aligned}\langle z', z'\rangle &= \langle z, z\rangle - nc_0^2 \\ &= \langle x, x\rangle + \langle y, y\rangle - \frac{\langle y, 1\rangle^2}{n} \\ &\quad [\text{since } x \perp 1] \\ &\geq \langle x, x\rangle\end{aligned}$$

由于令 $m = |M| \leq n$，并将 $y = d_0 1 + \mu$ 写为 $\mu \perp 1$ 和 $d_0 = \langle y, 1\rangle/m$，我们得到

$$\langle y, y\rangle - \frac{\langle y, 1\rangle^2}{n} = md_0^2 + \langle \mu, \mu\rangle - m \cdot \frac{m}{n} d_0^2 \geq 0.$$

现在我们继续证明上界。设 x 是一个非负向量，满足$x_i \geq x_{i+1}$ 并设 $r \leq n/2$ 满足 $x_k = 0$ 对于 $k > r$，由引理3.6可知这样的对象存在。

定义权重 $w^i, {}_1{}^j, \ldots, w^i, {}_r{}^j$，其中 $w^i, {}_k{}^j = \dfrac{x_i - x_j}{x_k - x_{k+1}}$，并约定$w^i, {}_k{}^j$仅对 $i \leq k < j$ 成立，而 $\sum_{k=i}^{j-i} \dfrac{1}{w^i, {}_k{}^j} = 1$。然后注意到 $(x_i - x_j)^2 = \sum_{k=i}^{j-1} w^i, {}_k{}^j (x_k - x_{k+1})^2$。定义 $w_k = \sum_i, {}_j w^i, {}_k{}^j$。于是得到 $\langle x, Lx\rangle = f_w(x)$（在引理 5.50 中定义），并且对所有 y，应用引理 5.49 于每条边可得 $\langle y, Ly\rangle \leq f_w(y)$。

令 $R_i = \sum_{j=i}^{r} \dfrac{1}{w_j}$，并令向量 v_i 满足：它们的前 i 个矩阵元素为 1，最后 $n/2$ 个矩阵元素为 0，其余矩阵元素满足 $f_w(v_i) = R_i^{-1}$（根据引理 5.50，这些向量确实存在），结合 σ 5.42 的定义、狄利克雷引理 5.34 以及上述观察，我们得到

$$\sigma \leq \frac{\langle v_i, Lv_i\rangle}{i} \leq \frac{f_w(v_i)}{i} \leq \frac{1}{i \cdot R_i}.$$

注意到 $\dfrac{1}{w_j} = R_j - R_{j+1}$，并利用 $\dfrac{a^2 - b^2}{a} \leq 2(a - b)$，

$$\begin{aligned}
\sum_i x_i^2 &= \sum_i^r (\sum_{j=i}^r x_j - x_{j+1})^2 = \sum_i^r (\sum_{j=i}^r (x_j - x_{j+1}) w_j^{1/2} R_j^{1/4} \frac{1}{w_j^{1/2} R_j^{1/4}})^2 \\
&\le \sum_i^r \sum_{j=i}^r (x_j - x_{j+1})^2 w_j R_j^{1/2} \sum_{j=i}^r \frac{1}{w_j R_j^{1/2}} \\
&= \sum_i^r \sum_{j=i}^r (x_j - x_{j+1})^2 w_j R_j^{1/2} \sum_{j=i}^r \frac{R_j - R_{j+1}}{R_j^{1/2}} \\
&\le 2 \sum_i^r \sum_{j=i}^r (x_j - x_{j+1})^2 w_j R_j^{1/2} R_i^{1/2} \\
&\le \frac{2}{\sigma} \sum_{i=1}^r \sum_{j \ge i}^r (x_j - x_{j+1})^2 w_j \frac{1}{\sqrt{j}} \frac{1}{\sqrt{i}} \\
&= \frac{2}{\sigma} \sum_{j=1}^r (x_j - x_{j+1})^2 w_j \frac{1}{\sqrt{j}} \sum_{i=1}^j \frac{1}{\sqrt{i}} \\
&\le \frac{4}{\sigma} \sum_{j=1}^r (x_j - x_{j+1})^2 w_j \\
&= \frac{4}{\sigma} f_w(x) \\
&= \frac{4}{\sigma} \langle x, Lx \rangle \\
\frac{\sigma}{4} &\le \frac{\langle x, Lx \rangle}{\langle x, x \rangle}
\end{aligned}$$

如所需。 □

˜这引出了第一个有趣的问题。既然我们知道 $\rho_{\tilde{L}} \le \rho_L$和 $\tau(A) \le \tau(A)$，那么对于非可逆链，是否也存在类似于引理 5.51 的关系？也就是说，是否可能有 $\sigma(A) \le c \cdot \Delta(A)$，其中某个常数为 c？这样的等价关系很有吸引力，而且由于混合时间与 σ 呈线性关系，而不像定理 ϕ 1.2中的情况那样呈二次关系，因此还能给出更好的界。不过遗憾的是，我们将证明这种关系并不成立。

引理 5.52。 事实并非如此： $\sigma_A \in O(\Delta(A))$。具体而言，令 A 为有向环。那么

$$\sigma_A \approx \sqrt{\Delta(A)}.$$

因此， σ_A 与 $\sigma_{\tilde{A}}$ 之间的谱隙可能呈二次关系，因为如果 A 是有向环，˜那么 A 就是无向环，并且

$$\sigma_{\tilde{A}} \le \Delta(\tilde{A}) \le \Delta(A).$$

证明。 设 A 是有向环。注意，对于任意集合 S 和 T ，其中$|S| = |T| = n/4$，且顶点位于正方形的相对两侧，我们得到狄利克雷向量在一半顶点上的值为 1，在另一半顶点上的值为 0，而 $A|_{S\cup T}$ 仍然是一个循环，因此谱保持不变；不过，谱隙现在增大了，因为第二特征值将变为实部最接近 1 的 $n/2$ 次单位根，而不是实部最接近 1 的 n次单位根。因此$\sigma(A) = \phi(A)$，引理得证。

□

因此， σ_A 与 $\sigma_{\tilde{A}}$ 之间的差距可能是二次的。尽管如此，利用引理 H 中得到的矩阵5.40来为非可逆链证明一个类似于5.51 的引理，仍然很有意思。

5.13 与对数Sobolev常数的关系

为了与其他量进行比较，我们陈述对数Sobolev常数与谱隙之间的关系，内容摘自 [DS96],。相关定义见 [DS96] 。令 A 为一个最大特征值为 1、对应左特征向量为 w的对称矩阵。令 w 归一化，使得 $\sum w_i^2 = 1$，并令 $w_0 =$最小值 $_i\, w_i$。如下定义 A 的对数Sobolev常数。令 α 为（对称 A， $\lambda = 1 - \lambda_2(A)$的）对数Sobolev常数，则

$$\frac{\lambda}{1 + \frac{1}{2}\log(1/\pi_0)} \leq \alpha \leq \lambda$$

团对应下界，超立方体对应上界。因此，

$$\frac{c_1.\rho}{2 + \log(1/\pi_0)} \leq \alpha \leq c_2 \cdot \rho.$$

5.14 一种不同的扩张概念

到目前为止，我们已经看到扩张的两个主要定义。第一个定义是标准的定义 2.4 ，表示为边扩张 ϕ；第二个定义源于归一化容量 5.42 ，表示为 σ。在本小节中，我们提出一个新的定义，将其表示为 μ。我们通过尝试理解 [Baf+20] 中高维扩展图的扩张而得出这个定义。和前文一样，为了从概念上理解该定义，我们将范围限制在双随机矩阵上。

定义 5.53。设 A 是一个不可约双随机矩阵。设 1_S 是这样一个向量：在集合 S 上取值为 1，在其他位置取值为 0，并设 $1'_S = 1_S/|S|$。定义

$$\mu_S(A) = \frac{1}{2}\|A1'_S - A1'_{\overline{S}}\|_1$$

并令

$$\mu(A) = \max_{S:|S|\leq n/2} \mu_S(A).$$

◊

为了理解这个定义，请注意，μ 的行为应当近似于 $1-\phi$。考虑 $A = J = \frac{1}{n}\mathbf{1}\cdot\mathbf{1}^T$。显然，$\mu(J) = 0$。考虑一个图 A，它有两个几乎断开的连通分量，分别对应于 S 和 $\overline{S}$。那么，$\mu(A) \approx 1$。然而，在某些图中，$\overline{\mu}$ 的行为也不同于 $1-\phi$。考虑一个有两个连通分量的二分图 A。对于支撑在其中一个连通分量上的集合 S，有 $\mu(A) = 1$；同时还有 $\phi(A) \approx \frac{1}{2}$，这表明 μ和 ϕ 都是常数。

引理 5.54. 设 A 是一个不可约双随机矩阵。那么

$$\mu(A) \geq 1 - 2\cdot\phi(A).$$

证明。对于任意集合 $|S| = k \leq n/2$，令 $k' = n-k$，并设 v 为 A 在 S中的列和向量。根据 $\phi_S(\overline{A})$的定义，有

$$\phi_S(A) = \frac{1}{k}\sum_{i=1}^{k} v_i.$$

此外，由于 A 是1_2-惰性的，我们有 $v_i \leq 1/2$ 对于 $i \leq k$。因此，我们得到

$$\begin{aligned}
\mu_S(A) &= \frac{1}{2}\|A1'_S - A1'_{\bar{S}}\|_1 \\
&= \frac{1}{2}\sum_i \left|\frac{1-v_i}{k} - \frac{v_i}{k'}\right| \\
&= \frac{1}{2}\sum_i \left|\frac{1}{k} - \frac{n\cdot v_i}{k\cdot k'}\right| \\
&= \frac{1}{2}\frac{n}{k\cdot k'}\sum_i \left|\frac{k'}{n} - v_i\right| \\
&\geq \frac{1}{2}\frac{n}{k\cdot k'}\sum_{i=1}^{k}\left(\frac{k'}{n} - v_i\right) \\
&\quad \text{[since } A \text{ is } 1/2\text{-lazy]} \\
&= \frac{1}{2}\left(1 - \frac{n}{k'}\phi_S(A)\right) \\
&\geq \frac{1}{2} - \phi_S(A)
\end{aligned}$$

如所需。对于一般的 A，注意 $\mathrm{t}\sum_i v_i = k'$，因此，

$$\sum_{i>k} v_i = k' - k\cdot\phi_S(A).$$

因此，我们得到：

$$\begin{aligned}
\mu_S(A) &= \frac{1}{2}\|A1'_S - A1'_{\bar{S}}\|_1 \\
&= \frac{1}{2}\frac{n}{k\cdot k'}\sum_i \left|\frac{k'}{n} - v_i\right| \\
&= \frac{1}{2}\frac{n}{k\cdot k'}\sum_i \left(\frac{k'}{n} - v_i\right) + 2\cdot\frac{1}{2}\frac{n}{k\cdot k'}\sum_{i:v_i\geq k'/n}\left(v_i - \frac{k'}{n}\right) \\
&= 0 + \frac{n}{k\cdot k'}\sum_{i:v_i\geq k'/n}\left(v_i - \frac{k'}{n}\right) \\
&\geq \frac{n}{k\cdot k'}\sum_{i>k:v_i\geq k'/n}\left(v_i - \frac{k'}{n}\right)
\end{aligned}$$

为了解求和可能取得的最小值，设有 r 指标 $i > k$满足 $v_i = k'/n$，其余指标均大于 k'/n。因此我们有

$$\sum_{i>k+r} v_i = k'\left(1 - \frac{r}{n}\right) - k\cdot\phi_S(A),$$

继续上述求和，可得

$$\begin{aligned}
\mu_S(A) &\geq \frac{n}{k \cdot k'} \sum_{i>k : v_i \geq k'/n} \left(v_i - \frac{k'}{n} \right) \\
&= \frac{n}{k \cdot k'} \sum_{i>k+r} \left(v_i - \frac{k'}{n} \right) \\
&= \frac{n}{k \cdot k'} \left(k' \left(1 - \frac{r}{n} \right) - k \cdot \phi_S(A) - \frac{k'}{n}(k' - r) \right) \\
&= \frac{n}{k \cdot k'} \left(k' \left(1 - \frac{k'}{n} \right) - k \cdot \phi_S(A) \right) \\
&= \frac{n}{k} \left(1 - \frac{k'}{n} \right) - \frac{n}{k'} \cdot \phi_S(A) \\
&= 1 - \frac{n}{k'} \cdot \phi_S(A) \\
&\geq 1 - 2 \cdot \phi_S(A)
\end{aligned}$$

如所要求。 □

我们要证明的第二个引理更简洁，因为它与引理 3.17类似，只是二元函数由加法改为取最小值。由引理 3.17可知，对于双随机 A 和 B,

$$\phi_S(AB) \leq \phi_S(A) + \phi_S(B).$$

μ 的对应引理如下。

引理 5.55。 设 μ 如5.54中所定义，并设 A 和 B 是两个双随机矩阵。那么

$$\mu_S(AB) \leq \min\{\mu_S(A), \mu_S(B)\}$$

因此

$$\mu(AB) \leq \min\{\mu(A), \mu(B)\}.$$

证明。 任取一个集合 S。首先，注意到

$$\mu_S(AB) = \frac{1}{2}\|AB1'_S - AB1'_{\overline{S}}\|_1 \leq \|A\|_1 \frac{1}{2}\|B1'_S - B1'_{\overline{S}}\|_1 \leq \mu_S(B).$$

我们的目标是现在证明 $\mu_S(AB) \leq \mu_S(A)$。设 B的列为 $B = (b_1, \ldots, b_n)$，设 S包含前 k个满足 $k' = n - k$的顶点，设 $u = b_1 + \ldots + b_k$，则$B1'_S = \frac{1}{k}u$，且 $B1'_{\overline{S}} = \frac{1}{k'}(1 - u)$，并且

$$B1'_S - B1'_{\overline{S}} = \frac{1}{k}u - \frac{1}{k'}(1-u) = \frac{n}{k\cdot k'}u - \frac{1}{k'}1 = \frac{n}{k'}\left(\frac{u}{k} - \frac{1}{n}\mathbf{1}\right)$$

其中 $0 \le u_i \le 1$ 且 $\sum_i u_i = k$，因为 B 是双随机的。

令 v 为一个向量，其前 k 个矩阵元素为 1，其余为 0，则 $u \le v$（或 u 被 v所支配），因此 $u = Rv$ 其中 R 是双随机矩阵。根据伯克霍夫-冯诺依曼定理（见第2章），R 可写成 $R = \sum \alpha_i P_i$，其中 P_i 是置换，且 $\alpha_i \ge 0$，$\sum \alpha_i = 1$。因此

$$u = \sum_i \alpha_i W_i$$

其中 W_i 是一个向量，有 k 个矩阵元素为 1，其余均为 0。现在我们知道，对于任意集合 S,

$$\begin{aligned}
\mu_S(A) &\ge \frac{1}{2}\|A1'_S - A1'_{\overline{S}}\|_1 \\
&= \frac{1}{2}\|\frac{1}{k}A1_S - \frac{1}{k'}A1_{\overline{S}}\|_1 \\
&= \frac{1}{2}\|\frac{1}{k}A1_S - \frac{1}{k'}(1 - A1_S)\|_1 \\
&= \frac{1}{2}\frac{n}{k'}\|\frac{1}{k}A1_S - \frac{1}{n}1\|_1
\end{aligned}$$

特别是对于所有 $|S| = k$。因此，对于任意满足 $|S| = k$ 的集合 S，我们有

$$\begin{aligned}
\mu_S(AB) &= \frac{1}{2}\|AB1'_S - AB1'_{\overline{S}}\|_1 \\
&= \frac{1}{2}\frac{n}{k'}\|\frac{1}{k}Au - \frac{1}{n}1\|_1 \\
&= \frac{1}{2}\frac{n}{k'}\|\sum \alpha_i \left(\frac{1}{k}AW_i - \frac{1}{n}1\right)\|_1 \\
&\le \frac{1}{2}\frac{n}{k'}\sum \alpha_i \|\frac{1}{k}AW_i - \frac{1}{n}1\|_1 \\
&\le \mu_S(A)
\end{aligned}$$

正如所要求的。 □

5.15 张量及其他

最后，我们尝试理解不再是线性的系统的扩张和混合。现在，我们的算子将是一个定义在 n 维空间上的 k 阶张量，记作 T，其中$T_{i_1,\ i_2,\ \ldots\ldots i_k}$ 是一些常数矩阵元素，且每个指标的维数均为 n。全文中我们将以三阶张量为例进行阐述。近年来，相关研究成果不断涌现

这得益于张量在量子信息和机器学习中的应用，但关于随机游走及其收敛的研究成果相对较少，而且非常具体。令人惊讶的是，佩龙-弗罗贝尼乌斯定理有许多不同版本，可以针对张量进行实例化并作多种解释（见[CPZ08; FGH13]）。事实上，对于最自然的张量特征值定义（见 [CQZ13]），一个 k 阶张量可能有 $n(k-1)^{n-1}$ 个特征值，当 $k > 2$时，这一数量呈指数级增长。本节将聚焦于一些非常具体的问题；我们认为这些问题十分基础且重要，但尚未得到解决。

第一个问题是基于 T 定义游走。这相对简单，有两种等价且在数学上颇为优美的方式来定义这种游走。假设我们有一个分布 $p \in \mathbb{R}^n$，为简单起见，假设 T 是一个非负三阶张量。那么我们可以如下定义游走的一步：

$$p_t(i) = \sum_{j=1,k=1}^{n} T_{i,j,k} p_{t-1}(j) \cdot p_{t-1}(k).$$

因此，我们的输出分布是输入分布的二次函数。注意，我们将第一个指标用作输出，不过也可以使用任何一个指标，这类似于矩阵情形中的转置。另一种定义游走的方式是考虑最近两个状态，并根据这两个状态演化到新状态。因此我们可以写成

$$p_t(i) = \sum_{j=1,k=1}^{n} T_{i,j,k} p_{t-1}(j) \cdot p_{t-2}(k).$$

这两种定义本身都很有趣，而对于我们所关注的问题，采用哪种定义并不重要。第一步，为确保我们始终处于概率分布的单纯形内，我们要求 T 是 1-线随机的，即对于每个固定的 j， k，

$$\sum_{i=1}^{n} T_{i,j,k} = 1.$$

因此请注意，如果 p_{t-1} 和 p_{t-2} 是概率分布，那么 p_t也是概率分布，因为 $p_t(i) \geq 0$，又因为 T 是非负的，并且

$$\sum_{i=1}^{n} p_t(i) = \sum_{j=1,k=1}^{n} p_{t-1}(j) \cdot p_{t-2}(k) \sum_{i=1}^{n} T_{i,j,k} = \sum_{j=1}^{n} p_{t-1}(j) \cdot \sum_{k=1}^{n} p_{t-2}(k) = 1.$$

因此，我们可以正式定义张量游走如下。

定义 5.56。 设 T 为一个维数为 n的非负 k阶张量。设 T 为 1-线随机的，即对每个固定的 j_1， j_2，……， j_{k-1}，

$$\sum_{i=1}^{n} T_{i,j_1,j_2,\ldots,j_{k-1}} = 1.$$

那么，由如下定义的张量游走可得到概率分布序列 p_t。

$$p_t(i) = \sum_{j_1,j_2,\ldots,j_{k-1}=1}^{n} T_{i,j_1,j_2,\ldots,j_{k-1}} p_{t-1}(j_1) p_{t-2}(j_2) \ldots p_{t-k+1}(j_{k-1})$$

简洁地说，我们写作

$$p_t = T(p_{t-1}, p_{t-2}, \ldots, p_{t-k}).$$

◊

我们指出，可以考虑任意数量的输入和输出，而在定义 $k-1$ 5.56 中，我们只考虑了 1 个输出和 个输入。对于概率分布的演化而言，这确实是最有意义的方式。于是，我们将演化的不动点 T 定义为在 T 下保持不变的概率分布。

定义 5.57。 (张量演化的不动点) 设 T 为一个 k 阶张量，具有 n 个维度且为 1 行随机，其演化由 T 按 5.56 中的定义给出。若分布 p 是 T 的不动点，则称之为

$$p = T(p, p, \ldots, p)$$
$$p(i) = \sum_{j_1,j_2,\ldots,j_{k-1}=1}^{n} T_{i,j_1,j_2,\ldots,j_{k-1}} p(j_1) p(j_2) \ldots p(j_{k-1}).$$

◊

根据这个定义，第一个问题如下。

If T is entry-wise positive, does T have a unique fixed point?

要理解的首要问题就是这一点，因为根据佩龙-弗罗贝尼乌斯定理，在矩阵情形下，正性意味着存在唯一不动点。直观上看似乎确实如此，但事实证明 T 可能有多个不动点。下面是一个构造于 [CZ13]中的例子。

引理 5.58。 *[CZ13]* 设 T 为一个二维正4阶张量，定义如下：

$$T_{1,1,1,1} = 0.872,\ T_{1,1,1,2} = 2.416/3,\ T_{1,1,2,1} = 2.416/3,\ T_{1,1,2,2} = 0.616/3,$$

$$T_{1,2,1,1} = 2.416/3,\ T_{1,2,1,2} = 0.616/3,\ T_{1,2,2,1} = 0.616/3,\ T_{1,2,2,2} = 0.072,$$

$$T_{2,1,1,1} = 0.128,\ T_{2,1,1,2} = 0.584/3,\ T_{2,1,2,1} = 0.584/3,\ T_{2,1,2,2} = 2.384/3,$$

$$T_{2,2,1,1} = 0.584/3,\ T_{2,2,1,2} = 2.384/3,\ T_{2,2,2,1} = 2.384/3,\ T_{2,2,2,2} = 0.928,$$

注意到 T 是1行随机的。因此，以下是 T 的两个不同的正不动点：

$$p_1 = 0.2,\ p_2 = 0.8$$

and

$$q_1 = 0.6,\ q_2 = 0.4.$$

证明可以通过直接计算进行验证。鉴于此，似乎连最基本的要求张量都无法满足。对于矩阵而言，几乎每个简单的定义，在应用于张量时都有多种阐释方式。事实上，即使采用最直观的特征值定义，对于三阶张量，也会得到 2^n 个特征值。尽管如此，我们实际上可以证明，在 2 行随机张量的情形下，唯一不动点是存在的。

定义 5.59。 设 T 是 n 维空间中的一个非负 k阶张量。如果除输出指标之外，还存在一个输入指标，使得张量沿该指标是行随机的，则称 T 是 2 行随机的。形式上，如果存在某个固定的输入指标 j 以及输入指标 i，使得对于任意固定的 l_1，..... l_{k-2}，都有

$$\sum_{i=1}^{n} T_{i,j,l_1,\ldots,l_{k-2}} = 1$$

and

$$\sum_{j=1}^{n} T_{i,j,l_1,\ldots,l_{k-2}} = 1.$$

◊

给定一个 2 行随机张量，我们确实可以证明，均匀分布是唯一不动点！这就是本节的主定理。

定理 5.60. 设 T 是一个正的 2-行随机张量。那么，对于 *5.56*中定义的张量游走，均匀分布是唯一不动点。

证明. 在深入证明之前，我们先指出两个关键要点。

1. 由于 T 在输出指标 i 上是行随机的，因此游走将始终处于概率单纯形内。

2. 由于 T 在输入指标 j 上是行随机的，因此将存在唯一不动点。

首先注意到，均匀分布是 T的不动点，因为

$$\begin{aligned} p_t(i) &= \sum_{j,l_1,\ldots,l_{k-2}=1}^{n} T_{i,j,l_1,\ldots,l_{k-2}} \left(\frac{1}{n}\right)^{k-1} \\ &= \left(\frac{1}{n}\right)^{k-1} \sum_{l_1,\ldots,l_{k-2}=1}^{n} \sum_{j=1}^{n} T_{i,j,l_1,\ldots,l_{k-2}} \\ &= \left(\frac{1}{n}\right)^{k-1} n^{k-2} \\ &= \frac{1}{n}. \end{aligned}$$

接下来我们证明，对于 $t \geq k$ ，每个 p_t 都是正向量。为此，设 p_1 至 p_{k-1} 为提供给我们的 $k-1$ 个非零初始分布。注意， p_i的 $1 \leq i \leq k-1$ 可能包含零矩阵元素，但并非所有矩阵元素都为零。令$x_1, \ldots, x_{k-1} \in [n]$ 分别为 p_i 中非零矩阵元素对应的指标。因此我们有

$$\begin{aligned} p_k(i) &= \sum_{j_1,j_2,\ldots,j_{k-1}=1}^{n} T_{i,j_1,j_2,\ldots,j_{k-1}} p_{k-1}(j_1) p_{k-2}(j_2) \ldots p_1(j_{k-1}) \\ &\geq T_{i,x_1,x_2,\ldots,x_{k-1}} p_{k-1}(x_1) p_{k-2}(x_2) \ldots p_1(x_{k-1}) \\ &> 0 \end{aligned}$$

由于 T 是正的，而且根据归纳法，该结果对所有 p_t 都成立。因此，不可能存在某些起始概率向量 $p_1 = p_2 = \ldots\ldots = p_{k-1} = q$ 不是正的，而不动点却是正的，因为 $p = T(q, \ldots, q)$将是正的，因此 $p \neq q$。这证明了不动点必定是正的。现在我们证明它必定唯一。

下一步，我们定义一个双随机矩阵 A_t ，其使用 T 和 p_t 如下。

$$A_t(i,j) = \sum_{l_1,\ldots,l_{k-2}=1}^{n} T_{i,j,l_1,\ldots,l_{k-2}} p_{t-2}(l_1) \ldots p_{t-k+1}(l_{k-2}).$$

我们注意到， A_t 是双随机的，这是因为二行随机条件可以通过直接计算进行验证。此外，请注意

$$p_{t+1} = A_t p_t.$$

现在假设 q 是 T 的一个非均匀正不动点。令

$$A_q(i,j) = \sum_{l_1,\ldots,l_{k-2}=1}^{n} T_{i,j,l_1,\ldots,l_{k-2}} q_{l_1} \ldots q_{l_{k-2}}.$$

注意， A_q 是双随机的。此外，由于 q 是正的， A_q 也为正，因此是不可约的。由于 q 是 T的不动点，我们有

$$q = A_q q.$$

然而，这与事实矛盾，因为 A_q 是不可约且双随机的，并且其平稳分布只有均匀分布。因此，该张量有唯一不动点。

□

接下来一个有趣的问题是，了解从任意固定分布出发时，张量收敛到不动点的速度有多快。实现这一目标的一种方法是采用 Mihail 的证明，该证明表明，非负矩阵的收敛速度与扩张的平方成反比 5.6。Mihail 的证明有一个优点：与我们在引理 5.6 中的较简单证明不同，它是组合性的，不依赖于底层对象的谱。然而，这种证明无法直接推广，因为引理 A_t 中的每个矩阵 5.60虽然是双随机矩阵，却是一个不同的 双随机矩阵。因此，我们将收敛到不动点的速度这一问题留作后续研究中定理5.60 最重要的后续工作。

结果与问题总结

我们在此汇总本学位论文的主要结果和一些重要问题。首先重述我们的部分主要定理。

定理 5.61. （谱隙与边扩张）设 R 是一个 $n \times n$ 非负矩阵，其边扩张 $\phi(R)$ 定义为2.4 和2.6，谱隙 $\Delta(R)$ 定义为2.2 和2.6。则

$$\frac{1}{15} \cdot \frac{\Delta(R)}{n} \leq \phi(R) \leq \sqrt{2 \cdot \Delta(R)}.$$

定理 5.62. （混合时间、奇异值、边扩张与谱隙）。设 R 和 A 是不可约且 1 2-惰性的非负矩阵，最大特征值为 1，其中 A 具有相同的对应左右特征向量。设混合时间如5.1中所定义，则

$$\tau_\epsilon(A) \leq \frac{\ln\left(\sqrt{\frac{n}{\kappa}} \cdot \frac{1}{\epsilon}\right)}{\ln\left(\frac{1}{\sigma_2(A)}\right)},$$

$$\frac{\frac{1}{2} - \epsilon}{\phi(R)} \leq \tau_\epsilon(R) \leq \frac{4 \cdot \ln\left(\frac{n}{\kappa \cdot \epsilon}\right)}{\phi^2(R)},$$

$$(1 - \Delta)\frac{1 - \epsilon}{\Delta} \leq \tau_\epsilon(R) \leq 20 \cdot \frac{n + \ln\left(\frac{1}{\kappa \cdot \epsilon}\right)}{\Delta(R)}.$$

定理 5.63. （*Rootn* 矩阵 4.17）对于每个 A_n，都存在一个双随机矩阵 n，使得

$$\phi(A_n) \leq \frac{\Delta(A_n)}{\sqrt{n}}.$$

定理 5.64. （*Chet* 矩阵 4.27）对于每个矩阵 C_n，都存在 n，使得$\sum_i C_{i,}$ $_j = 1$且 $\sum_j C_{i,\ j} = 1$对于所有 i，j，并且令 $\phi(C_n)$ 按与2.3中双随机矩阵类似的方式定义，则有

$$\phi(C_n) \leq 2 \cdot \frac{\Delta(C_n)}{n}.$$

定理 5.65. *(*容量单调性*)* 设 A 是一个不可约非负矩阵，其最大特征值为 1，对应的左右特征向量为 w。设

$A_\alpha = \tilde{A} + \alpha\overline{A}$其中 $\tilde{A} = \frac{1}{2}(A + A^T)$，$\overline{A} = \frac{1}{2}(A - A^T)$，且 $-1 \leq \alpha \leq 1$。那么对于$|\alpha| \leq |\beta|$以及每个 $U \subseteq [n]$ 和 $a \in \mathbb{R}^{|U|}$，有

$$cap_{U,a}(A_\alpha) \leq cap_{U,a}(A_\beta).$$

定理 5.66。 *(*张量的唯一不动点*)* 设 T 是一个定义在 n 维空间中的 k阶张量，且逐项为正。设 T 是2-行随机的，使得对于输出指标i和某个输入指标 j，对于任意 i、 j、 l_1、...、 l_{k-2}，均有

$$\sum_i T_{i,j,l_1,\ldots,l_{k-2}} = 1$$

and

$$\sum_j T_{i,j,l_1,\ldots,l_{k-2}} = 1.$$

令 T 如5.56所述作用于每个概率分布。则均匀分布是 T的唯一不动点。

我们在此汇总一些主要猜想和开放问题。

猜想 5.67。 *(Chet* 猜想*)* 设 C_n 表示在构造 *4.27* 中定义的 *Chet* 矩阵，并设 $C = \{n : C_n$ 是逐项非负的$\}$。则以下结论成立：

$$|C| = \infty.$$

更强地，

$$C = \mathbb{N}.$$

猜想 5.68。 *(*迹猜想*)* 设 A 是一个次随机的非负矩阵，即$\sum_i A_{i,\ j} \leq 1$且 $\sum_j A_{i,\ j} \leq 1$对所有 j和 i均成立。假设如下：在对角线上方， A 仅在距离对角线为 *1* 的矩阵元素处有非零项；在对角线下方， A 仅在距离对角线至多为 k 的矩阵元素处有非零项，其中对角线到其自身的距离为零。假设对于 $l \leq k + 1$，$TrA^l \leq 1$ 。则对于所有 l，

$$TrA^l \leq 1.$$

第一个放宽是证明该结论适用于 *Toeplitz* 情形，即每条对角线上的元素都相同；第二个放宽是证明该结论适用于无限情形，此时条件 $TrA^l \leq 1$ 被替换为

$$p_l \leq \alpha \cdot \frac{1}{k}$$

其中某个常数为 α， p_l 表示恰好经过 l 步后返回起始顶点的概率。

事实上，每一章都可以提出大量开放问题，例如：能否利用引理 H 中的矩阵5.40，得到比定理5.39更精确的容量界，类似于引理5.41 是引理 5.40的精确版本？是否存在一种经过修改的归一化容量概念，使其与一般非负 R 的谱隙仅相差常数因子，类似于引理 5.51？对于 [LOT14], 中定义的高阶扩张概念，它与一般非负R的谱隙有何关系？诸如此类。不过，在我们看来，最有趣且最重要的开放问题是下面这个；我们相信，这将只是一个广阔而未知领域中的第二步。

开放问题 5.69。 对于 2 行随机张量，是否存在一种扩张概念 $\phi(T)$，T使得张量游走经过大约 $1/\phi^2(T)$步后便收敛到均匀分布，类似于引理5.6？

我们希望这些定理、猜想和问题在未来得到应用、解决或深入探讨。

参考文献

[AF02] David Aldous 和 Jim Fill。《图上的可逆马尔可夫链与随机游走》。2002年（引文见第11页）。[Alo86] N. Alon。《特征值与扩展图》。载于：《组合数学》6.2（1986），第83–96页（引文见第53页）。[AM85] N. Alon 和 V. Milman。《图的λ_1、等周不等式与超集中器》。载于：《组合理论杂志》B辑 38（1985），第73–88页（引文见第4、21、25页）。[ARV09] Sanjeev Arora、Satish Rao 和 Umesh Vazirani。《扩展流、几何嵌入与图划分》。载于：《美国计算机学会杂志》（JACM）56.2（2009），第1–37页（引文见第5页）。[Baf+20] Mitali Bafna 等。《高维扩展图：特征值剥离、伪随机性与唯一游戏》。载于：arXiv 预印本（2020），arXiv–2011（引文见第14、138页）。[BF60] F. L. Bauer 和 C. T. Fike。《范数与排斥定理》。载于：《数值数学》2.1（1960），第137–141页（引文见第7、32页）。[Bow81] Samuel Bowles。《技术变革与利润率：置盐定理的一个简单证明》。载于：《剑桥经济学杂志》5.2（1981），第183–186页（引文见第3页）。[BP94] Abraham Berman 和 Robert J Plemmons。《数学科学中的非负矩阵》。SIAM，1994年（引文见第2、3页）。[BR65] G. R. Blakley 和 P. Roy。《关于非负元素对称矩阵的 Hölder型不等式》。载于：美国数学会会刊 6.16（1965），第1244–1245页（引文见第38页）。[Bro69] Charles Philip Brown。《梵语韵律学与数字符号释义》。Trübner & Company，1869年（引文见第53页）。[Bru46] N. G. de Bruijn。《一个组合问题》。载于：荷兰皇家艺术与科学院科学部论文集 49.7（1946），第758–764页（引文见第53、54页）。[Bru75] Nicolaas Govert de Bruijn。《承认 C. Flye Sainte-Marie 在计数圆形排列方面的优先权：其中每个由 2^n 个零和一构成的 n字母词恰好出现一次》。载于：（1975）（引文见第53页）。[Bus82] P. Buser。《关于等周常数的一则注记》。载于：《高等师范学校科学年刊》第4辑 15.2（1982），第213–230页（引文见第4、21页）。

[Cha+96] Ashok K Chandra 等：《图的电阻刻画了其往返时间和覆盖时间》。载于：computational complexity 6.4（1996），第 312–340 页（引文见第 13 页）。[Che70] J. Cheeger：《拉普拉斯矩阵最小特征值的一个下界》。载于：为纪念 S. Bochner 举办的《分析问题研讨会》。R. C. Gunning 编。普林斯顿大学出版社，1970 年，第 195–199 页（引文见第 4、21 页）。[Chu07] Fan Chung：《Cheeger不等式和图划分算法的四种证明》。载于：ICCM 会议论文集。第 2 卷。Citeseer。2007 年，第 378 页（引文见第 25 页）。[CPZ08] Kung-Ching Chang、Kelly Pearson 和 Tan Zhang：《非负张量的佩龙-弗罗贝尼乌斯定理》。载于：《数学科学通讯》6.2（2008），第 507–520 页（引文见第 14、143 页）。[CQZ13] Kungching Chang、Liqun Qi 和 Tan Zhang：《非负张量谱理论综述》。载于：《数值线性代数及其应用》20.6（2013），第 891–912 页（引文见第 143 页）。[CZ13] Kung Chang 和 T. Zhang：《关于转移概率张量的 Z-特征向量的唯一性与非唯一性》。载于：Journal of Mathematical Analysis and Applications 408（2013 年 12 月），第 525–540 页。doi：10.1016/j.jmaa.2013.04.019（引文见第 14、144、145 页）。[DeG74] Morris H DeGroot：《达成共识》。载于：Journal of theAmerican Statistical Association 69.345（1974），第 118–121 页（引文见第 3 页）。[Dod84] J. Dodziuk：《差分方程、等周不等式与某些随机游走的暂态性》。载于：Trans. Amer. Math. Soc. 284.2（1984），第 787–794 页（引文见第 4 、 21 页）。[DS84] Peter G Doyle 和 J Laurie Snell：《随机游走与电网络》。第 22 卷。美国数学会，1984 年（引文见第 13 页）。[DS96] Persi Diaconis 和 Laurent Saloff-Coste：《有限马尔可夫链的对数 Sobolev 不等式》。载于：《应用概率年刊》6.3（1996），第 695–750 页（引文见第 138 页）。[DT98] C. Delorme 和 J.-P. Tillich：《de Bruijn 图和 Kautz图的谱》。载于：《欧洲组合学杂志》19.3（1998），第 307–319 页（引文见第 54 页）。[FGH13] Shmuel Friedland 、 StéphaneGaubert 和 Lixing Han：《非负多重线性形式及其扩展的佩龙-弗罗贝尼乌斯定理》。载于：Linear Algebra and its Applications 438.2（2013），第 738–749 页（引文见第 14、143 页）。[Fie95] M. Fiedler：《双随机矩阵非随机特征值的一个估计》。载于：Linear algebra and its applications 214（1995），第 133–143 页（引文见第 22、29 页）。

[Fil91] J. A. Fill：《非可逆马尔可夫链收敛到平稳性的特征值界及其在排斥过程中的应用》。载于：《应用概率年刊》1.1（1991），第 62–87 页（引文见第 97–99 页）。[Fre56] John RP French Jr.：《社会权力的形式理论》。载于：《心理学评论》63.3（1956），第 181 页（引文见第 3 页）。[Fro12] G. Frobenius：《关于由非负元素构成的矩阵》。载于：（1912）（引文见第 2、16、17 页）。[GJS74] M. R. Garey、D. S. Johnson 和 L. Stockmeyer：《一些简化的 NP 完全问题》。载于：第六届 ACM 计算理论年度研讨会论文集。ACM。1974 年，第 47–63 页（引文见第 4 页）。[Goo46] Irving John Good：《正规循环小数》。载于：《伦敦数学学会杂志》1.3（1946），第 167–169 页（引文见第 53 页）。[Har59] Frank Harary：《法国社会权力理论中的一致性判据》。载于：（1959）（引文见第 3 页）。[HLW06] Shlomo Hoory、Nathan Linial 和 Avi Wigderson：《扩展图及其应用》。载于：《美国数学会通报》43.4（2006），第 439–561 页（引文见第 5、8、49 页）。[HS49] David Hawkins 和 Herbert A Simon：《注：宏观经济稳定性的一些条件》。载于：《计量经济学》，Journal of the Econometric Society（1949），第 245–248 页（引文见第 3 页）。[KH97] Anatole Katok 和 Boris Hasselblatt：《动力系统现代理论导论》。第 54 卷。剑桥大学出版社，1997 年（引文见第 3 页）。[Kla81] M. Klawe：《一维扩张图不存在》。载于：第 22 届计算机科学基础年度研讨会论文集。IEEE。1981 年，第 109–114 页（引文见第 8 页）。[Kla84] M. Klawe：《扩张图显式构造的局限性》。载于：SIAM计算杂志 13.1（1984），第 156–166 页（引文见第 8、51、53 页）。[KM32] William Ogilvy Kermack 和 Anderson G McKendrick：《流行病数学理论的贡献。II——地方性流行问题》。载于：《伦敦皇家学会会刊》A辑，刊载数学与物理类论文 138.834（1932），第 55–83 页（引文见第 3 页）。

[KM33] William Ogilvy Kermack 和 Anderson G McKendrick。“对流行病数学理论的贡献。III——关于地方性流行问题的进一步研究”。载于：《伦敦皇家学会会刊》A辑，数学与物理特征论文集 141.843（1933），第 94–122 页（引于第 3 页）。[KM91] William O Kermack 和 Anderson G McKendrick。“对流行病数学理论的贡献——I. 1927”。载于：《数学生物学通报》53.1-2（1991），第 33–55 页（引于第 3 页）。[Lax07] P. D. Lax。《线性代数及其应用》。第二版。新泽西州霍博肯：Wiley-Interscience，2007。ISBN：9780471751564 0471751561（引于第 18 页）。[Leo86] 瓦西里·列昂惕夫。《投入产出经济学》。牛津大学出版社，1986（引于第 3 页）。[Les45] Patrick H Leslie。“矩阵在某些人口数学问题中的应用”。载于：《生物计量学》33.3（1945），第 183–212 页（引于第 3 页）。[Les48] Patrick H Leslie。“关于矩阵在人口数学中应用的进一步札记”。载于：《生物计量学》35.3/4（1948），第 213–245 页（引于第 3 页）。[Li92] W.-C. W. Li。“特征和与阿贝尔拉曼努金图”。载于：《数论杂志》41（1992）。附录作者：K. Feng 和 W.-C. W. Li，第 199–217 页（引于第 53 页）。[LM11] Amy N Langville 和 Carl D Meyer。《Google 的 PageRank 及其扩展》。普林斯顿大学出版社，2011（引于第 3 页）。[LOT14] J. R. Lee 、 S. Oveis Gharan 和 L. Trevisan。“ 多路谱划分与高阶切格不等式”。载于：《美国计算机学会会刊》（JACM）61.6（2014），第 37 页（引于第 5、150 页）。[LPW09] D. A. Levin、Y. Peres 和 E. L. Wilmer。《马尔可夫链与混合时间》。美国数学会，2009（引于第 2、11、12 页）。[Lyo83] Terry Lyons。“可逆马尔可夫链暂态性的一条简单判据”。载于：《概率年刊》（1983），第 393–402 页（引于第 13 页）。[Mac00] Charles R MacCluer。“ 佩龙定理的多种证明及其应用 ”。载于：《SIAM评论》42.3（2000），第 487–498 页（引于第 2 页）。[Mar06] AA Markov。“将大数定律推广到相互依赖的量，喀山大学物理数学学会会刊，第2辑，第15卷，第135–156页”。载于：俄语。英文译文：《大数定律向相依量的推广》（1906）（引于第 2 页）。

[Mih89] M. Mihail。“马尔可夫链的导通率与收敛：扩展图的组合式处理”。载于：第*30*届计算机科学基础年会。IEEE。1989，第 526–531 页（引于第 11、23、25、98、99 页）。[MS19] Jenish C Mehta 和 Leonard J Schulman。“非负矩阵的边扩张与谱隙”。载于：*arXiv* 预印本 *arXiv:1909.12497*（2019）（引于第7页）。[MT06] R. Montenegro 和 P. Tetali。“马尔可夫链混合时间的数学方面”。载于：理论计算机科学基础与趋势 1.3（2006），第 237–354 页。DOI：10.1561/0400000003（引于第 11 页）。[MWW18] Gary L Miller、Noel J Walkington 和 AlexL Wang。“ Hardy-Muckenhoupt

t

拉普拉斯特征值的界”。载于：arXiv预印本 arXiv:1812.02841（2018）（参见第134、135页）。[Nil91] A. Nilli。“关于图的第二特征值”。载于：《离散数学》91（2 1991），第207–210页（参见第4、21、53页）。[Pag+99] Lawrence Page 等。《PageRank引文排名：为网络带来秩序》。技术报告。斯坦福InfoLab，1999（参见第3页）。[Pat12] T. H. Pate。“将 Blakley and Roy 的 Hölder 型不等式推广到非对称非方阵 ”。载于：美国数学会汇刊 364.8（2012），第4267–4281页（参见第38页）。[Per07] O. Perron。“关于矩阵理论”。载于：《数学年刊》64.2（1907），第248–263页（参见第2、16、17页）。[Pop05] Karl Popper。《科学发现的逻辑》。Routledge，2005（参见第53页）。[Saa11] Y. Saad。《大型特征值问题的数值方法》。第二版。SIAM，2011（参见第7、32页）。[Sch18] Aaron Schild。“舒尔补切格不等式”。载于：arXiv预印本 arXiv:1811.10834（2018）（参见第134、135页）。[Sen96] Eugene Seneta。“马尔可夫与链依赖理论的诞生”。载于：《国际统计评论》/《国际统计评论》（1996），第255–263页（参见第2页）。[Sin92] A. Sinclair。“马尔可夫链混合率及多商品流的改进界”。载于：《组合学、概率与计算》1.4（1992），第351–370页（参见第109页）。

[SJ89] A. Sinclair 和 M. Jerrum。“近似计数、均匀生成与快速混合的马尔可夫链”。载于：Information and Computation 82.1（1989），第93–133页。ISSN：0890-5401。DOI：http://dx.doi.org/10.1016/0890-5401(89)90067-9。网址：http://www.sciencedirect.com/science/article/pii/0890540189900679（参见第4、21页）。[Slo12] Martin Slowik。“可逆与非可逆马尔可夫链容量的变分表示札记”。载于：未发表，柏林工业大学（2012）（参见第126页）。[SS11] Daniel A Spielman 和 Nikhil Srivastava。“利用有效电阻进行图稀疏化”。载于：SIAM计算杂志 40.6（2011），第1913–1926页（参见第13页）。[ST96] Daniel A Spielman 和 Shang-Hua Teng。“谱划分有效：平面图与有限元网格”。载于：第37届计算机科学基础会议论文集。IEEE。1996，第96–105页（参见第5页）。[TA85] Akira Takayama 和 Takayama Akira。《数理经济学》。剑桥大学出版社，1985（参见第3页）。[Tet91] Prasad Tetali。“网络中的随机游走与有效电阻”。载于：Journal of Theoretical Probability 4.1（1991），第101–109页（参见第13页）。[Thu88] William P Thurston。“关于曲面微分同胚的几何与动力学”。载于：《美国数学会通报》（新系列）19.2（1988），第417–431页（参见第3页）。[TM06] David A Tolliver 和 Gary L Miller。“通过谱舍入进行图划分：在图像分割与聚类中的应用”。载于：2006年IEEE计算机学会计算机视觉与模式识别会议（CVPR’06）。第1卷。IEEE。2006，第1053–1060页（参见第5页）。[Tol79] Richard Chace Tolman。《统计力学原理》。Courier Corporation，1979（参见第3、111页）。[Var62] Richard S Varga。《迭代分析》。施普林格，1962（参见第3页）。[Vaz17] U. Vazirani。个人通信。2017（参见第8、51页）。[Zem06] Jens-Peter M Zemke。“Hessenberg特征值—特征矩阵关系”。载于：Linear Algebra and its Applications 414.2-3（2006），第589–606页（参见第85页）。

附录 A

CHET 矩阵——精确计算与数值计算

我们给出一些通过精确计算和数值计算得到的 Chet 矩阵。精确计算是在 Maple 中完成的，数值计算则是在 Matlab 中完成的。

A.1 精确计算的 $n = 16$ Chet 矩阵

关于 C_{16} （其中矩阵定义于构造 4.27 中）的方程如下。

$$r = \left(\frac{1}{16}\right)^{\frac{1}{15}}$$

$$14 \cdot c_0 = 2 \cdot r - 1$$

$$13 \cdot r \cdot c_1 = \frac{8}{7} \cdot r^2 - \frac{15}{28}$$

$$12 \cdot r^2 \cdot c_2 = \frac{528}{637} \cdot r^3 + \frac{45}{637} \cdot r - \frac{20}{49}$$

$$11 \cdot r^3 \cdot c_3 = \frac{38980}{57967} \cdot r^4 + \frac{7095}{231868} \cdot r^2 + \frac{110}{1029} \cdot r - \frac{334605}{927472}$$

$$10 \cdot r^4 \cdot c_4 = \frac{2622960}{4463459} \cdot r^5 + \frac{15675}{811538} \cdot r^3 + \frac{3650}{93639} \cdot r^2 + \frac{5019075}{35707672} \cdot r - \frac{21653}{62426}$$

$$9 \cdot r^5 \cdot c_5 = \frac{221814104}{406174769} \cdot r^6 + \frac{23279535}{1624699076} \cdot r^4 + \frac{4980}{218491} \cdot r^3 + \frac{135515025}{3249398152} \cdot r^2 + \frac{194877}{1092455} \cdot r - \frac{4578373505}{12997592608}$$

$$8 \cdot r^6 \cdot c_6 = \frac{96435386896}{179123073129} \cdot r^7 + \frac{33490290}{2843223383} \cdot r^5 + \frac{10590760}{656128473} \cdot r^4 + \frac{63909555}{2843223383} \cdot r^3 + \frac{4070764}{99413405} \cdot r^2 + \frac{22891867525}{102356041788} \cdot r - \frac{569210253}{1530966437}$$

$$7 \cdot r^7 \cdot c_7 = \frac{14468197830722}{25614599457447} \cdot r^8 + \frac{4659561625}{443542847748} \cdot r^6 + \frac{8450450}{656128473} \cdot r^5$$
$$+ \frac{98852689755}{6505295100304} \cdot r^4 + \frac{2013729}{99413405} \cdot r^3 + \frac{389161747925}{10645028345952} \cdot r^2$$
$$+ \frac{1707630759}{6123865748} \cdot r - \frac{42029290110657}{104084721604864}$$

$$6 \cdot r^8 \cdot c_8 = + \frac{12646467661976}{19922466244681} \cdot r^9 + \frac{36526797305}{3622266589942} \cdot r^7 + \frac{2018448010}{179123073129} \cdot r^6$$
$$+ \frac{268437195855}{22768532851064} \cdot r^5 + \frac{74507973}{5686446766} \cdot r^4 + \frac{68675602575}{4139733245648} \cdot r^3$$
$$+ \frac{15368676831}{557271783068} \cdot r^2 + \frac{126087870331971}{364296525617024} \cdot r - \frac{53944771607543}{120370705142688}$$

$$5 \cdot r^9 \cdot c_9 = \frac{4192421772289592}{5438833284797913} \cdot r^{10} + \frac{225785202013625}{21755333139191652} \cdot r^8 + \frac{280853043700}{26331091749963} \cdot r^7$$
$$+ \frac{10494908689675}{1035968244723412} \cdot r^6 + \frac{197518666}{19902563681} \cdot r^5 + \frac{42922251609375}{4143872978893648} \cdot r^4$$
$$+ \frac{11384205060}{975225620369} \cdot r^3 + \frac{210146450553285}{16575491915574592} \cdot r^2 + \frac{269723858037715}{631946201999112} \cdot r$$
$$- \frac{293344384450753313}{580142217045110720}$$

$$4 \cdot r^{10} \cdot c_{10} = + \frac{583988144046138064}{571077494903780865} \cdot r^{11} + \frac{433842562719595}{38071832993585391} \cdot r^9$$
$$+ \frac{284474332622540}{26357422841712963} \cdot r^8 + \frac{239093054413515}{25381221995723594} \cdot r^7$$
$$+ \frac{136555337396}{16300199654739} \cdot r^6 + \frac{504083501274005}{65265999417574956} \cdot r^5$$
$$+ \frac{7074714234537}{976200845989369} \cdot r^4 + \frac{42029290110657}{7251777713063884} \cdot r^3$$
$$- \frac{269723858037715}{28753552190959596} \cdot r^2 + \frac{2640099460056779817}{5076244399144718800} \cdot r$$

$$-\frac{487051361453894455}{843437530934814816}$$

$$3 \cdot r^{11} \cdot c_{11} = \frac{2582373218800886646164}{1714945717196053937595} \cdot r^{12} + \frac{60910874936948457}{4619382403221694108} \cdot r^{10}$$
$$+\frac{704066732737790}{61500653297330247} \cdot r^{9} + \frac{1873806347700790155}{203252825741754540752} \cdot r^{8}$$
$$+\frac{10063658054264}{1331182971803685} \cdot r^{7} + \frac{454789092312238525}{71270471363991851952} \cdot r^{6}$$
$$+\frac{74092391002251}{13666811843851166} \cdot r^{5} + \frac{948306872766753891}{232288943704862332288} \cdot r^{4}$$
$$-\frac{53944771607543}{134183243557811448} \cdot r^{3} - \frac{71282685421533055059}{1847752961288677643200} \cdot r^{2}$$
$$+\frac{2435256807269472275}{3936041811029135808} \cdot r - \frac{250148541000967840970441}{376308088801876978306560}$$

$$2 \cdot r^{12} \cdot c_{12} = \frac{8292429130967703269984}{3273987278283375699045} \cdot r^{13} + \frac{251191067356462921}{16167838411275929378} \cdot r^{11}$$
$$+\frac{204277892726612218}{16789678350171157431} \cdot r^{10} + \frac{52904039025561315}{5879213967736701592} \cdot r^{9}$$
$$+\frac{54680493901861873}{7995084928652932110} \cdot r^{8} + \frac{3520738115297033525}{654797455656675139809} \cdot r^{7}$$
$$+\frac{5335710693495901}{1243679877790456106} \cdot r^{6} + \frac{237927811316429277}{73910118451547105728} \cdot r^{5}$$
$$+\frac{1059097700970891719}{805904560808215556688} \cdot r^{4} - \frac{36081359287442657499}{6467135364510371751200} \cdot r^{3}$$
$$-\frac{75492961025353640525}{1074539414410954075584} \cdot r^{2} + \frac{250148541000967840970441}{3592031756745189338380} \cdot r$$
$$-\frac{611028831268072088101}{795955121785891907840}$$

$$1 \cdot r^{13} \cdot c_{13} = \frac{2167687960275208704611884}{4171059792533020640583333} \cdot r^{14} + \frac{14252777333017827436531}{8739363374831090865984120} \cdot r^{12}$$

$$+\frac{4397770296526056}{395716324751508761}\cdot r^{11}+\frac{43699381559221754727}{5885093181704438293592}\cdot r^{10}$$

$$+\frac{145628182696554137}{27982797250285262385}\cdot r^{9}+\frac{802269369938308587191 75}{20974472099594618078361888}\cdot r^{8}$$

$$+\frac{16063263041956539}{5540028546521122654}\cdot r^{7}+\frac{43750487609031995683}{20177462337272359863744}\cdot r^{6}$$

$$+\frac{1255133000992702981}{940221987609584816136}\cdot r^{5}-\frac{984757098601178871741}{25894409999499528491804 80}\cdot r^{4}$$

$$-\frac{2435256807269472275}{341898904585303569504}\cdot r^{3}-\frac{250148541000967840970441}{29715899078528384526604 80}\cdot r^{2}$$

$$+\frac{1833086493804216264303}{2785842926250621677440}\cdot r-\frac{41502574124308301104426128989}{46982817503091944495530629120}$$

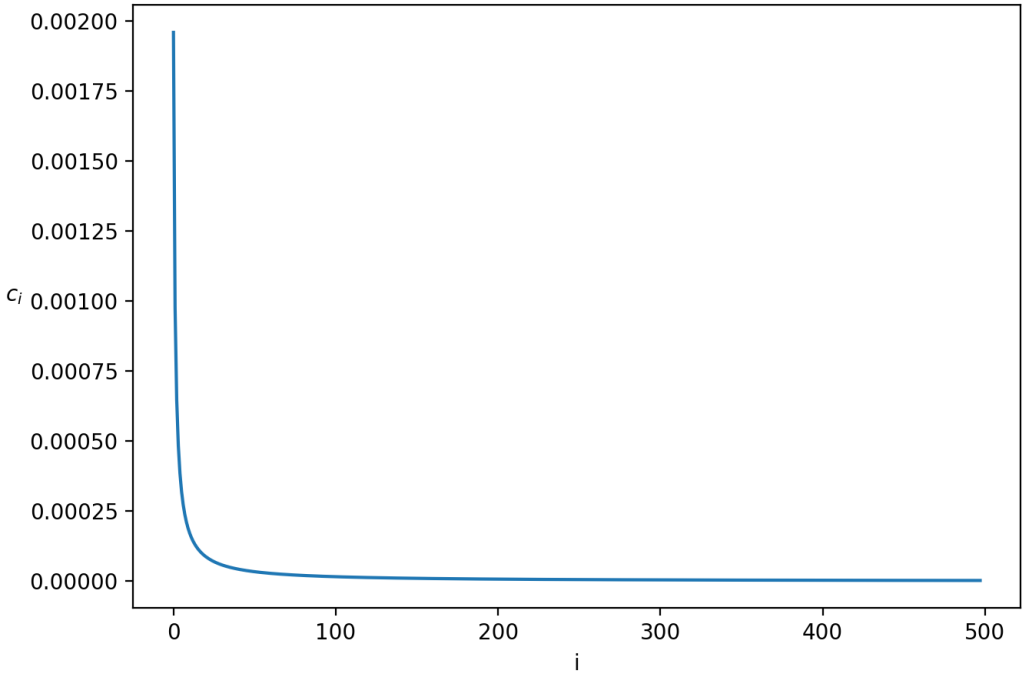
A.2 针对输入整数 n 计算 Chet 矩阵的伪代码

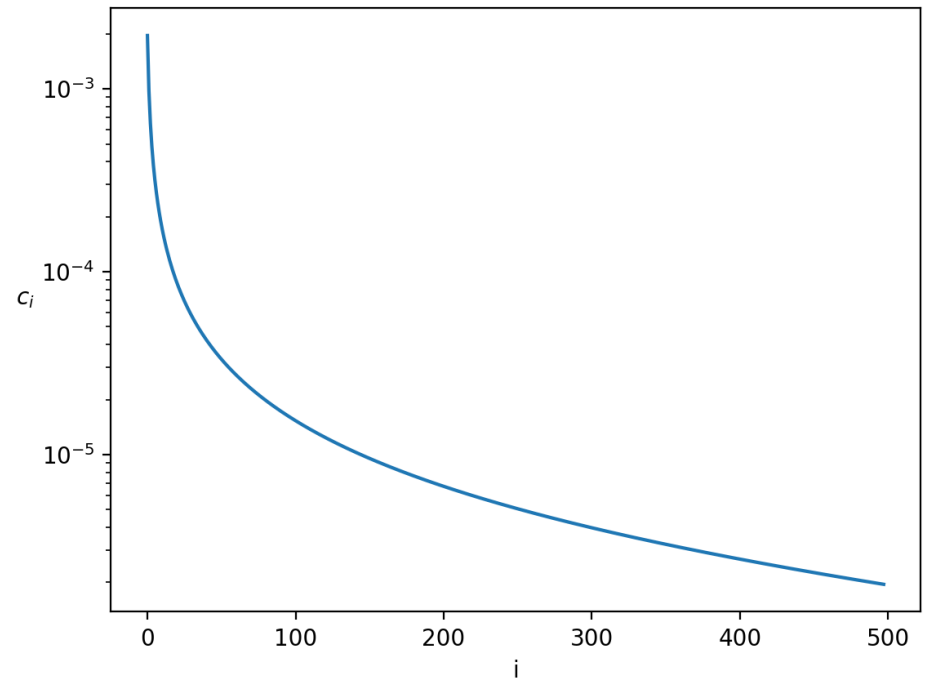
```
chet(n): C =零矩阵, 索引从1到n r =(1/n)^(1/(n-1))fori
    = 1ton-1: C[i+1,i] =rbval = 1-rC[1,1] =bvalC[n,
  n] =bvalfork = 1ton-1: cval =(1 -迹(C^k)))/(k*(
  n-k-1)*r^(k-1))bval =bval -cvalfori=2ton-k: C[i,
  i+k-1] =cvalC[1,1+k] =bvalC[n-k,n] =bval
```

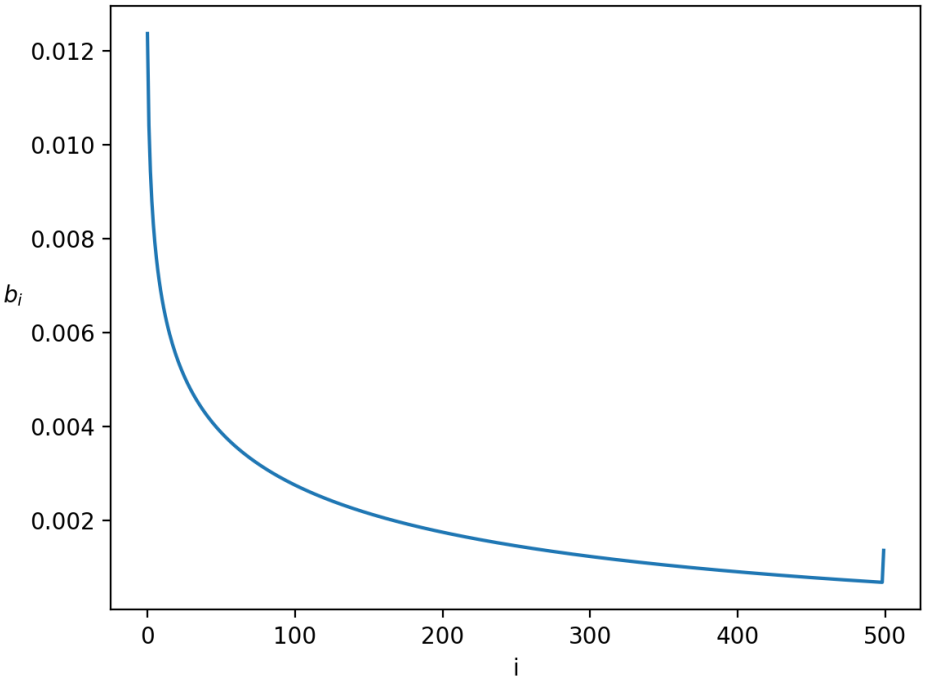
```
  C[1,n] = 1fork = 1ton-1:
  C[1,n] =C[1,n] -C[1,k]
```

A.3 $n = 500$ 的 Chet 矩阵（数值计算）

我们绘制了以下数值： c_i 和 b_i 对于 $n = 500$ 个 Chet 矩阵。第一张图包含以下数值： c_i，第二张图包含以下数值： c_i ，其中 y 轴采用对数刻度；第三张图包含以下数值： b_i。请注意，在第三张图中，最后一个元素的值为 b_{499}，它不一定会随着其他b_i而递减，因为它只是其他 b_i 的线性函数，这些元素的范围从 $i = 0$ 到 498 (详见构造4.27 ）。对于不同大小的所有矩阵，也观察到了相同现象。




10^{-3}
10^{-4}
c_i
10^{-5}
0
100
200
300
400
500
i


0.012
0.010
0.008
b_i
0.006
0.004
0.002
0
100
200
300
400
500
i