

从图割看高级算法

随机化、近似与谱松弛

一个问题框架，三条算法主线

本课程由三位教师共同讲授，分别负责随机算法与随机化技术、组合优化的近似算法、谱算法与谱图论三个部分。第一讲选择“图中的割”作为共同入口：全局最小割展示随机收缩、存活概率与概率放大；最大割展示近似保证与去随机化；图的电导率则引出谱松弛、特征向量与扫描舍入。三个问题使用同一套图论语言，却需要截然不同的算法眼光，也由此勾勒出整门课程的三条主线。

高级算法课程教学团队

南京大学

第一讲 · 三部分课程的共同导论

学习目标与阅读路径

本讲并非简单罗列三类算法，而是借助同一个基本对象，考察优化目标、松弛方式与证明语言如何随问题而改变。完成本讲后，读者应当能够：

1. 说明为何同样是寻找一个图割，最小化、最大化和寻求低电导率会导向三种不同的算法范式；
2. 独立复述 Karger 随机收缩算法中“固定最小割存活”的概率分析，并用独立重复把常数或逆多项式成功率放大到高概率；
3. 从均匀随机割推出最大割的 $1/2$ -近似，再用条件期望法将这一算法去随机化；
4. 沿着“Rayleigh 商松弛—第二特征向量—扫描舍入”的链条，解释谱划分如何把连续解转化为具有电导率保证的离散割。

怎样阅读本讲

若目的是先把握课程主线，建议依次阅读第 1 节、第 2 节中 Karger 算法及其基本概率放大、第 3 节直至条件期望法（含），以及第 4 节直至 Cheeger 不等式及其扫描证明（含）。这条路径已经覆盖三部分课程各自最具代表性的基本图景。

希望进一步掌握技术细节的读者，可以继续学习 Karger-Stein 算法的递归分析、基于两两独立性的种子枚举，以及谱坐标的应用。这些内容仍放在各自的算法部分，以保持问题、算法和分析之间的完整联系。附录只整理正文所需的概率论与线性代数工具，既可在遇到技术障碍时查阅，也可作为课前复习材料。

除非另有说明，全文中的图均为有限无向图，且 $n \geq 2$ 。随机收缩过程中保留平行边的重数，并随时删除自环；边默认具有单位权重，带权结论会另行声明。涉及运行时间时，正文会明确指出图的表示方式以及所采用的基本操作成本，避免把概率分析与实现模型混为一谈。

目录

学习目标与阅读路径	i
1 以图割为共同线索	1
一 随机算法与随机化技术	2
2 随机收缩与全局最小割	2
二 组合优化中的近似算法	6
3 最大割：近似算法与去随机化	6
三 谱算法与谱图论	10
4 谱划分与电导率	10
5 习题	16
A 概率与随机化工具	18
B 线性代数工具	19
文献说明	20

1 以图割为共同线索

图割的定义十分简单，却为比较三类算法思想提供了天然的共同语言。同一个边界，既可以成为最小化或最大化的对象，也可以先进入连续松弛，再由谱方法恢复离散解。由此引出本课程三个板块：随机算法、组合优化近似算法、谱算法与谱图论。它们并非彼此割裂：随机割连接了随机化与近似，而“松弛—舍入”又连接了近似算法与谱方法。

设 $G = (V, E)$ 是一个无向无自环多重图，顶点数为 $n = |V|$ ，边副本数为 $m = |E|$ 。对互不相交的顶点集 $S, T \subseteq V$ ，记

$$E(S, T) = \{uv \in E : u \in S, v \in T\}.$$

若 $\emptyset \neq S \subsetneq V$ ，令 $\bar{S} = V \setminus S$ ，并记 $\delta(S) = E(S, \bar{S})$ 。无序二划分 $\{S, \bar{S}\}$ 称为一个割，其大小为 $|\delta(S)|$ 。注意， S 与 $\bar{S}$ 描述的是同一个割。

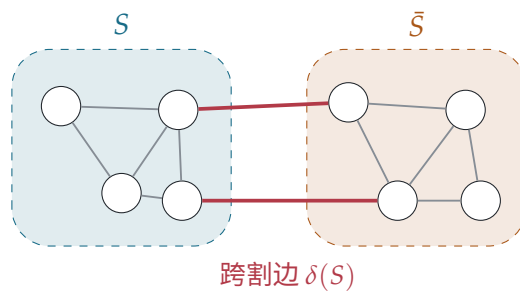


图 1: 一个割就是顶点集的二划分；割的大小由横跨两侧的边数决定。不同优化目标以不同方式衡量同一条边界。

本讲考察以下三个目标：

$$\begin{aligned} \lambda(G) &= \min_{\emptyset \neq S \subsetneq V} |\delta(S)| && \text{(全局最小割)}, \\ \text{OPT}_{\text{MC}}(G) &= \max_{S \subseteq V} |\delta(S)| && \text{(最大割)}, \\ h(G) &= \min_{0 < |S| \leq n/2} \frac{|\delta(S)|}{|S|} && \text{(边扩展率)}. \end{aligned}$$

若 G 是 d -正则图，它的电导率 (conductance) 为

$$\phi(G) = \frac{h(G)}{d} = \min_{\emptyset \neq S \subsetneq V} \frac{|\delta(S)|}{d \min\{|S|, |\bar{S}|\}}.$$

对于非正则图，分母中的顶点数应改为体积 $\text{vol}(S) = \sum_{v \in S} \deg(v)$ ；详见第 4.4 节。

问题	优化目标	对应课程板块与算法视角
全局最小割	$\min_S \delta(S) $	第一部分：精确随机算法；固定最小割的存活概率
最大割	$\max_S \delta(S) $	第二部分：NP-困难组合优化问题的近似与去随机化
电导率	$\min_S \phi(S)$	第三部分：谱松弛与有保证的扫描舍入

表 1: 同一个“割”原语，分别预演由三位教师主讲的三个课程板块。

第一部分

随机算法与随机化技术

2 随机收缩与全局最小割

第一个例子要求得到精确最优解，却采用一个不断改变图结构的随机过程。分析围绕两种反复出现的思路展开：先固定一个最小割，估计它在自适应随机过程中存活至终点的概率；再通过独立重复放大成功概率（amplification）。所需的概率链式法则与重复试验界见附录 A。

若 G 不连通，则 $\lambda(G) = 0$ ，任取一个连通分量便得到最小割；一次图搜索即可先处理这种情形。因此下文假设 G 连通。Karger 算法的出发点看似反常：不断收缩边，在逐步丢失原图信息的同时，设法保住某个固定的最小割。

2.1 边收缩与存活不变量

收缩边（contraction） $e = uv$ ，是把 u, v 合并成一个超顶点，将原来与二者相接的边全部重新接到该超顶点上，保留平行边，并删除由 u, v 之间的边产生的自环。每个超顶点始终代表一组被合并到一起的原图顶点。

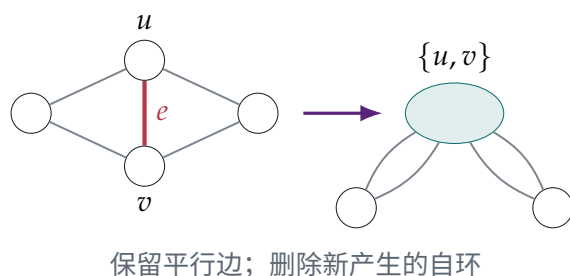


图 2: 收缩 uv 会合并它的两个端点。只要把超顶点记录为原顶点的集合，收缩图中的任意割都能自然提升回原图。

引理 2.1 (收缩下的割)。 设 G/e 是收缩 $e = uv$ 后得到的图。

- (i) G/e 中的每个割都可提升为 G 中一个把 u, v 放在同侧的割，而且跨割边的副本完全相同。
- (ii) 割 $C = \delta_G(S)$ 在 G/e 中仍然存在，当且仅当 $e \notin C$ 。
- (iii) 若 C 是 G 的最小割且 $e \notin C$ ，则它在 G/e 中仍是最小割。

证明。 G/e 的每个顶点都是一块原图顶点。把一个割同侧的各块展开，就得到 G 中的一个割；收缩既不改变不同块之间的边副本，也不改变其中哪些边跨越这个割，故 (i) 成立。收缩后 e 的两个端点同属一块，所以恰在 C 不分开二者、即 $e \notin C$ 时， C 才能由收缩图中的割表示，这证明了 (ii)。最后，若 G/e 存在比 C 更小的割，则由 (i) 将其提升回 G ，便得到 G 中一个比 C 更小的割，与 C 的最优性矛盾，故 (iii) 成立。 $\square$

RandomContraction(G): Karger 随机收缩算法 (1993)

1. 当图中还剩两个以上超顶点时，从当前边的多重集中等概率选择一个边副本并收缩它。
2. 最后两个超顶点之间的所有平行边给出一个割；将其提升回原图并输出。

算法无论如何都会输出原图的一个割；随机性只影响这个割是否最优。下面证明，一个固定的最小割有不可忽略的概率存活到最后。

每一步使用新随机性，并不意味着各步选择相互独立

Karger 算法在每一步都使用新的随机比特，从当前边的多重集中均匀选择一条边。但各步选中的边并非相互独立：此前的选择决定了当前图。因而分析时必须使用条件概率的链式法则，不能把各步的无条件概率直接相乘。

2.2 固定最小割的存活概率

固定一个最小割 C ，并记 $\lambda = |C| = \lambda(G)$ 。假设 C 已经存活到当前图只剩 r 个超顶点的时刻。由引理 2.1，此时 C 仍是当前图的最小割。

引理 2.2 (最小割存活时的边数下界)。 若一个 r 顶点多重图的最小割值为 $\lambda > 0$ ，则它至少有 $r\lambda/2$ 个边副本。

证明。 每个顶点的度都至少为 λ ，因为孤立该顶点的割，其大小恰为该顶点的度。由握手定理， $2m = \sum_v \deg(v) \geq r\lambda$ 。 $\square$

于是，在剩余 r 个超顶点时，当前边副本中属于 C 的比例至多为 $\lambda/(r\lambda/2) = 2/r$ 。换言之，下一次收缩避开 C 的条件概率至少为 $1 - 2/r$ 。

定理 2.3 (Karger 的固定割保证)。 在任意连通的 n 顶点无自环多重图上，RandomContraction 输出任意指定最小割 C 的概率至少为

$$\prod_{r=3}^n \left(1 - \frac{2}{r}\right) = \frac{2}{n(n-1)} = \binom{n}{2}^{-1}.$$

因此，算法输出某个最小割的概率也至少为该值。

证明。 令 A_r 表示“剩余 r 个超顶点时的那次收缩没有击中 C ”这一事件。在此前各步均未击中 C 的条件下，上面的论证给出

$$\mathbb{P}(A_r \mid C \text{ 此前一直存活}) \geq 1 - \frac{2}{r}.$$

由命题 A.1 中的链式法则，

$$\mathbb{P}(C \text{ 存活到最后}) \geq \prod_{r=3}^n \frac{r-2}{r} = \frac{(n-2)!}{n!/2} = \frac{2}{n(n-1)}.$$

如果 C 一直存活到只剩两个超顶点，它就是收缩图中唯一的割，因而必被算法输出。输出 C 当然意味着输出一个最小割，定理得证。 $\square$

注 2.4 (分析是紧的)。 对环图 C_n , 任意两条边都构成一个不同的最小割, 因此它恰有 $\binom{n}{2}$ 个最小割。固定其中一对边 C 。在不收缩 C 中边的条件下, 每个中间图仍是一个环; 当图中有 r 个超顶点时, r 个边副本中恰有两个属于 C 。所以存活概率乘积中的每个不等式都取等号, $\mathbb{P}(\text{输出 } C) = 1/\binom{n}{2}$ 。因此, 固定割的概率下界不能改进。

2.3 成功率放大与一个计数推论

独立运行 RandomContraction 共 R 次, 输出其中最小的割。由命题 A.3 and 定理 2.3,

$$\mathbb{P}(\text{所有运行都没有得到最小割}) \leq \exp\left(-\frac{R}{\binom{n}{2}}\right).$$

取 $R = \lceil c \binom{n}{2} \ln n \rceil$, 失败概率便至多为 n^{-c} 。这是一个蒙特卡洛 (Monte Carlo) 算法: 它总会输出一个合法的割, 但可能不是最优割; 通过增加独立重复次数, 可以把非最优的概率压到任意给定的多项式倒数。

下面给出一种可直接实现上述随机过程的方法。把 m 个边副本存成列表, 先随机打乱, 再借助并查集从前向后扫描。若一条边的两个端点已经属于同一超顶点, 就跳过它; 否则收缩它, 直至只剩两个超顶点。随机序列中第一条仍然有效的边, 在所有当前非自环边副本中均匀分布, 因此该实现与 RandomContraction 完全等价。计入最后一次扫描以读出割, 单次运行的时间为 $O(m\alpha(n) + n)$, 空间为 $O(m + n)$, 其中 $\alpha(n)$ 是增长极慢的反 Ackermann 函数。本节的概率分析则与具体表示无关。

同一个随机试验还会推出一个纯粹的确定性结构结论。

推论 2.5 (最小割的数目)。 一个连通的 n 顶点多重图至多有 $\binom{n}{2}$ 个不同的最小割。这里把割视为无序二划分; 等价地, 也可以按不同的割边集计数。环图 C_n 达到此上界。

证明。 先说明, 在连通图中, 不同的无序二划分必有不同的割边集合。事实上, 若 $\delta(S) = \delta(T)$, 则沿任意边 uv , u, v 关于 $S \triangle T$ 的归属比特必然相同。由连通性, $S \triangle T$ 只能是 $\emptyset$ 或 V , 所以 $T = S$ 或 $T = \bar{S}$ 。

对每个最小割 C , 令 B_C 表示“一次运行输出 C ”。一次运行只有一个输出, 因此各事件 B_C 两两不交; 而定理 2.3 保证 $\mathbb{P}(B_C) \geq 1/\binom{n}{2}$ 。故

$$1 \geq \mathbb{P}\left(\bigcup_C B_C\right) = \sum_C \mathbb{P}(B_C) \geq \frac{\#\{\text{最小割}\}}{\binom{n}{2}}.$$

□

为什么必须假设图连通

推论 2.5 中的连通性不可省略。不连通图的最小割值为零; 只需把各连通分量任意分配到割的两侧, 就可能得到指数多个零割。原始收缩循环也可能在仅剩两个超顶点之前便无边可选。先计算连通分量, 可以同时处理这两个问题。

2.4 Karger-Stein 递归算法

基本收缩算法的大部分成功概率都损失在最后阶段：只有当超顶点已经很少时，因子 $1 - 2/r$ 才会明显偏小。Karger 与 Stein 据此不再一路收缩到底，而是在适当位置提前停止，并从这里独立分支递归。这一改动仍然完全建立在同一个“固定最小割存活”分析之上。

令 $\text{ContractTo}(G, t)$ 表示不断等概率选择当前边并收缩，直到恰剩 t 个超顶点。把定理 2.3 中的乘积截断在 t ，可知对任意固定最小割 C ，

$$\mathbb{P}(C \text{ 从 } n \text{ 个超顶点存活到 } t \text{ 个}) \geq \prod_{r=t+1}^n \left(1 - \frac{2}{r}\right) = \frac{t(t-1)}{n(n-1)}. \quad (1)$$

FastCut(G): Karger-Stein 算法

令 $n = |V(G)|$ 。

1. 若 $n \leq 6$ ，枚举所有非平凡割，返回其中最小者。
2. 令 $t = \lceil n/\sqrt{2} + 1 \rceil$ 。
3. 相互独立地构造 $G_1 = \text{ContractTo}(G, t)$ 与 $G_2 = \text{ContractTo}(G, t)$ 。
4. 分别在两张图上递归运行 FastCut，把两个输出割提升回 G ，并返回较小者。

如此选取并取整 t ，可保证

$$q_n := \frac{t(t-1)}{n(n-1)} \geq \frac{1}{2}.$$

确实， $t-1 \geq n/\sqrt{2}$ 且 $t \geq n/\sqrt{2} + 1$ ，所以 $t(t-1) \geq n^2/2 \geq n(n-1)/2$ 。令 $p(n)$ 为 FastCut 在所有连通 n 顶点多重图上的最坏成功概率，即这些成功概率的下确界。若某个固定最小割在一条分支中存活，则收缩图的最小割值仍等于原图的最小割值；递归返回的任意最小割都可提升为原图的最小割。两条分支使用相互独立的随机性，因此

$$p(n) \geq 1 - (1 - q_n p(t))^2 \geq 1 - (1 - \frac{1}{2} p(t))^2 = p(t) - \frac{p(t)^2}{4}. \quad (2)$$

下面把“递归只损失一个对数因子”精确地写出来。令 $n_0 = n$ ， $n_{j+1} = \lceil n_j/\sqrt{2} + 1 \rceil$ ，再令 L 为首个满足 $n_L \leq 6$ 的下标。当 $n_j \geq 12$ 时，

$$n_{j+1} \leq n_j/\sqrt{2} + 2 \leq 7n_j/8;$$

而在 $7 \leq n_j < 12$ 时，规模每一步都严格下降。因此即使考虑取整，仍有 $L = O(\log n)$ 。

定义 $a_0 = 1$ 以及 $a_{j+1} = a_j - a_j^2/4$ 。由于 $0 < a_j \leq 1$ ，

$$\frac{1}{a_{j+1}} - \frac{1}{a_j} = \frac{1}{4(1 - a_j/4)} \leq \frac{1}{3}.$$

从而 $a_j \geq 3/(j+3)$ 。递归叶子以概率 1 成功。利用函数 $x - x^2/4$ 在 $[0, 1]$ 上单调递增，从叶子开始向上对式 (2) 迭代 L 层，得到

$$p(n) \geq a_L = \Omega(1/\log n).$$

为分析上述时间递推，采用标准的稠密重数矩阵表示，并假设字长为 $\Omega(\log(m+n))$ 比特的 word-RAM 上每次字操作成本为常数。剩余 r 个超顶点时，先按顶点度数成比例抽取第一个

端点，再从对应矩阵行中抽取一个关联边副本；这样得到的无向边副本在当前所有边副本中均匀分布。借助累积度数结构并扫描一行，每次抽样需 $O(r)$ 次字操作；合并相应的两行两列也需 $O(r)$ 次字操作。因此，从 n 个超顶点开始的一段收缩总计 $O(n^2)$ 次字操作，于是

$$T(n) \leq 2T(\lceil n/\sqrt{2} + 1 \rceil) + O(n^2) = O(n^2 \log n).$$

递归树有 $O(\log n)$ 层；除取整带来的常数因子外，每层的收缩总工作量为 $O(n^2)$ 。进一步，若单次运行的成功概率至少为 $c_0/\log n$ ，则独立运行 $R = \lceil (3/c_0) \log^2 n \rceil$ 次，全部失败的概率至多为 $\exp(-c_0 R / \log n) \leq n^{-3}$ 。返回所有运行中最小的割，便可在 $O(n^2 \log^3 n)$ 次字操作内以高概率得到最小割。

定理 2.6 (Karger-Stein)。 在上述 word-RAM 表示下，对任意连通的 n 顶点多重图，FastCut 单次运行使用 $O(n^2 \log n)$ 次字操作，并以 $\Omega(1/\log n)$ 的概率返回最小割。再作独立成功率放大，即可在 $O(n^2 \log^3 n)$ 次字操作内以高概率得到最小割。

注 2.7 (历史视角)。 今天学习收缩算法，主要是为了掌握其证明思想；它已不再代表全局最小割的最佳最坏情形运行时间。Karger 后来给出了精确的随机近线性时间算法，近期又出现了带权图的确定性近线性时间算法。由于这些结果采用不同的计算模型，且隐藏的对数因子也不同，不能脱离模型简单比较快慢。文末列出了相应文献。

第一部分要点

这套概率分析并不试图同时理解算法的所有可能输出。它只固定一个最优割，证明该割以至少 $2/[n(n-1)]$ 的概率存活，再用独立重复放大成功率。Karger-Stein 则提前停止收缩并分支递归，避开存活概率最差的末段，同时完整保留固定割分析。固定最优对象、分析递归随机过程、放大成功概率，是后续随机算法中会反复出现的三条主线。

第二部分

组合优化中的近似算法

3 最大割：近似算法与去随机化

最小割可在多项式时间内精确求解；只把“最小”改为“最大”，计算复杂性便截然不同。最大割的精确求解是 NP-困难的，因此本部分转而寻求多项式时间近似算法，并要求它在每个输入上都具有可证明的最坏情形保证。我们还将看到随机性的另一种典型用法：先用期望证明好解存在，再用条件期望逐步固定随机选择，把存在性证明转化为确定性算法。这个过程称为去随机化（derandomization）。所需的概率论工具汇总于附录 A。

对图 $G = (V, E)$ ，最大割的最优值记为

$$\text{OPT}_{\text{MC}}(G) = \max_{S \subseteq V} |\delta(S)|.$$

判定版问题询问是否存在大小至少为 K 的割；它是 NP-完全的 [6]，因而相应的优化问题是 NP-困难的。若一个算法对每个输入图 G 都能在多项式时间内输出割 S_A ，且满足

$$|\delta(S_A)| \geq \alpha \text{OPT}_{\text{MC}}(G),$$

就称它为最大割的 α -近似算法， α 称为近似比（approximation ratio）。即使最优值为零，上述定义仍然成立。对随机算法，若输出满足 $\mathbb{E}[|\delta(S_A)|] \geq \alpha \text{OPT}_{\text{MC}}(G)$ ，就称它为期望意义下的 α -近似算法。

3.1 一个确定性的 1/2-近似算法

任取顶点次序 $v_1, \dots, v_n$ 。处理 v_i 时，只有 $v_1, \dots, v_{i-1}$ 已经被放置到割的两侧。

GreedyCut(G)

若 $m = 0$ ，返回任意一个两侧均非空的二划分。否则令 $S = T = \emptyset$ 。依次处理 $v_1, \dots, v_n$ ：比较把 v_i 放入 S 或 T 时，它与已处理顶点之间新产生的跨割边数，把 v_i 放到产生更多跨割边的一侧；相等时任意选择。最后返回 (S, T) 。

定理 3.1 (贪心最大割)。 采用邻接表时，GreedyCut 的运行时间为 $O(n + m)$ ，且返回的割至少包含

$$\frac{m}{2} \geq \frac{1}{2} \text{OPT}_{\text{MC}}(G)$$

条边。因此它是一个确定性的 1/2-近似算法。

证明。 在放置 v_i 之前，设它到当前集合 S 和 T 的边数分别为 a_i 和 b_i 。若把 v_i 放入 T ，会新增 a_i 条跨割边；若放入 S ，则会新增 b_i 条。因此贪心选择带来的增量满足

$$\max\{a_i, b_i\} \geq \frac{a_i + b_i}{2}.$$

每条边恰好在两个端点中较晚处理的那个端点出现时，被计入一次 $\sum_i (a_i + b_i)$ ，所以算法最终得到的割至少有 $m/2$ 条边。另一方面，任何割都不可能包含超过全部 m 条边，故 $\text{OPT}_{\text{MC}}(G) \leq m$ 。

实现时，只需在一条边较晚处理的端点处检查该边；孤立点贡献额外的 $O(n)$ 时间。若 $m > 0$ ，则输出自动是非平凡的：要么两侧此前已经都非空，要么第一次遇到一条边的较晚端点时，它会被放到该边另一端点的对侧。□

若边带有非负权重，把“边数”全部替换为“总权重”，上述算法和证明无需其他改动；总权重为零时仍需单独返回一个非平凡二划分。这个分析在允许任意打破平局的意义下是紧的：考虑一条含两条边的路径，先处理两个叶子，并在两次初始平局中把它们放到不同侧，再处理中间顶点。贪心算法只能割开一条边，而最优解能割开两条边。

3.2 贪心规则背后的随机化证明

给每个顶点独立地赋一个均匀随机比特 $X_v \in \{0, 1\}$ ，并据此将顶点放入 S 或 T 。令

$$Z = |\delta(S)| = \sum_{uv \in E} \mathbf{1}_{\{X_u \neq X_v\}}.$$

对每条边都有 $\mathbb{P}(X_u \neq X_v) = 1/2$ 。由期望的线性性（定理 A.2），

$$\mathbb{E}[Z] = \sum_{uv \in E} \mathbb{P}(X_u \neq X_v) = \frac{m}{2} \geq \frac{1}{2} \text{OPT}_{\text{MC}}(G). \quad (3)$$

因此至少有一个割的大小不小于 $m/2$ 。不过，期望值只是对所有随机结果取平均；单独采样一次并不保证达到这个平均值。接下来要把这条存在性结论变成每次都正确的确定性算法。

为使随机试验的样本空间写起来更简洁，我们允许所有顶点落在同一侧，并把这种赋值的割值记为零。这不会影响结论：若 $m > 0$ ，任何割值至少为 $m/2$ 的赋值必然两侧非空；若 $m = 0$ ，任意非平凡二划分都是最优解。

假设已经固定 $X_{v_1}, \dots, X_{v_{i-1}}$ ；这一部分赋值对应一个正概率事件 H 。在给定 H 后，下一个尚未固定的比特仍然均匀，所以当前条件期望（conditional expectation）正好是两个可能分支的平均：

$$\mathbb{E}[Z \mid X_{v_1} = x_1, \dots, X_{v_{i-1}} = x_{i-1}] = \frac{1}{2} \sum_{b \in \{0,1\}} \mathbb{E}[Z \mid X_{v_1} = x_1, \dots, X_{v_i} = b].$$

两个分支中至少有一个不低于当前值。每一步都选择条件期望较大的那个 b ，条件期望便不会下降；全部比特固定以后，条件期望就是所得割的实际大小，所以最终值至少为初始的 $m/2$ 。

这个看似抽象的规则为什么恰好就是 GreedyCut？在第 i 步，除 v_i 与已放置顶点之间的边以外，其余边在两种选择下的条件期望完全相同：从 v_i 连向未放置顶点的边仍以 $1/2$ 的概率跨越割，不与 v_i 相邻的边则根本不受影响。两种选择之间唯一的差别，是到已放置顶点的 a_i 条边或 b_i 条边会成为跨割边。因此，最大化条件期望正是前面的贪心规则。

条件期望法

先设计一个期望目标值足够好的随机过程，再依次公开其中的随机选择。每一步选择一个使条件期望不低于当前值的取值。如果这些条件期望能够高效比较，最终便得到一个确定性对象，其质量至少达到原随机过程的期望保证。

3.3 选读：两两独立已经足够

式 (3) 中对每条边 uv 的计算只涉及一对随机比特 (X_u, X_v) ，因而并不需要全部 n 个比特相互独立。只要每个比特都均匀，并且任意一对彼此独立，即满足两两独立（pairwise independence），同一证明就能成立。

令

$$\ell = \lceil \log_2(n+1) \rceil.$$

记 $\mathbb{F}_2 = \{0,1\}$ 为二元域，其中加法就是异或；对 $a, r \in \mathbb{F}_2^\ell$ 定义 $\langle a, r \rangle = \sum_{j=1}^{\ell} a_j r_j \pmod{2}$ 。若一个线性映射的每个可能输出都有原像，就称它为满射；线性映射的秩是其像空间的维数。由于 $\mathbb{F}_2^\ell$ 中至少有 n 个非零向量，可以给每个顶点分配互不相同的非零标签 $a_v \in \mathbb{F}_2^\ell$ 。现在只随机选取一个均匀种子 $R \in \mathbb{F}_2^\ell$ ，并令

$$Y_v = \langle a_v, R \rangle \pmod{2}.$$

定理 3.2 (小规模的两两独立样本空间)。随机比特族 $(Y_v)_{v \in V}$ 中，每个比特都均匀，且任意两个比特相互独立。

证明。 因为 $a_v \neq 0$ ，线性泛函 $R \mapsto \langle a_v, R \rangle$ 的秩为一，所以恰有一半种子使其取零，另一半使其取一。若 $u \neq v$ ，则 a_u, a_v 是不同的非零向量。在 $\mathbb{F}_2$ 上它们必然线性无关，因为唯一的非零标量是 1。于是映射

$$R \mapsto (\langle a_u, R \rangle, \langle a_v, R \rangle)$$

的秩为二。四种输出各有 $2^{\ell-2}$ 个原像，因此出现概率均为 $1/4$ 。 $\square$

即使使用全部非零向量作为标签，这一随机比特族也不满足三重独立（3-wise independence）。只要 $a, b, a+b$ 都被选作标签，其中 a, b 不同且非零，就有

$$Y_a \oplus Y_b = Y_{a+b},$$

从而 $\mathbb{P}(Y_a = Y_b = Y_{a+b} = 0) = 1/4 \neq 1/8$ 。这个构造的实质，是用一个 ℓ 比特的短随机种子生成许多带相关性的比特，并没有凭空增加随机熵。

由于两两独立保证每条边仍以 $1/2$ 的概率被割开，所得随机割的期望值 依旧是 $m/2$ 。该分布的支撑集至多含 $2^\ell = O(n)$ 个候选割，而非全部 2^n 个赋值。因此，枚举所有种子并返回其中最好的割，即可完成去随机化。

PairwiseSeedSearch(G)

给各顶点选取互不相同的非零标签 $a_v \in \mathbb{F}_2^\ell$ 。对每个种子 $r \in \mathbb{F}_2^\ell$ ，按 $\langle a_v, r \rangle$ 的取值放置顶点 v ，计算所得割的大小；返回所有候选中最大的一个。

种子总数满足 $2^\ell \leq 2(n+1)$ 。在字长为 $\Omega(\log(m+n))$ 比特的 word-RAM 模型下，若逐坐标计算内积，直接枚举使用的字操作数为

$$O(2^\ell(m+n\ell)) = O(nm + n^2 \log n),$$

因为对每个种子既要计算 n 个长度为 ℓ 的内积，也要检查 m 条边来评价候选割；这里既计入内积计算，也计入候选割的求值。它比 GreedyCut 慢，但揭示了伪随机性中的一个基本思想：把所需随机比特数从 n 降到 $O(\log n)$ 。借助 Walsh-Hadamard 变换同时计算所有种子的得分，则只需 $O(m + n \log n)$ 次字操作；这一改进留作进阶练习。

全零种子会把所有顶点放到同一侧。若 $m > 0$ ，所有候选的平均割值为 $m/2 > 0$ ，所以全零种子不可能是最优候选；若 $m = 0$ ，则另行返回任意非平凡二划分。

3.4 越过 $1/2$: Goemans-Williamson 算法

Goemans 和 Williamson 的经典算法使用半正定规划 (semidefinite programming, SDP)：将每个取值为 ± 1 的变量 x_v 松弛为一个单位向量，并最大化 $\frac{1}{2} \sum_{uv \in E} (1 - \langle x_u, x_v \rangle)$ 。随后随机选择一个过原点的超平面，用它把这些向量分到两侧。若一对向量的内积为 ρ ，它们被超平面分开的概率是 $\arccos(\rho)/\pi$ 。逐边比较这一概率与松弛问题中该边的贡献，可证明算法输出割的期望值至少为 $\alpha_{\text{GW}} \text{OPT}_{\text{MC}}(G)$ ，其中

$$\alpha_{\text{GW}} = \min_{-1 \leq \rho < 1} \frac{2 \arccos(\rho)}{\pi(1-\rho)} \approx 0.878567.$$

唯一游戏猜想 (Unique Games Conjecture) 是 Khot 提出的、目前尚未证明的复杂性猜想 [8]。在该猜想成立的前提下，Khot、Kindler、Mossel 和 O' Donnell 证明：对任意 $\varepsilon > 0$ ，达到 $\alpha_{\text{GW}} + \varepsilon$ 近似比是 NP-困难的 [9]。本课程后面还会系统讨论半正定松弛和随机超平面舍入。

第二部分要点

近似算法以可证明的近似比替代精确最优化。对于最大割，均匀随机赋值的期望割值为 $m/2$ ；条件期望法在逐步固定选择的同时维持这一保证，从而得到确定性的 $1/2$ -近似算法。

第三部分

谱算法与谱图论

4 谱划分与电导率

前两部分都直接在顶点集合上进行优化：或精确寻找最小割，或近似求解最大割。本部分改走另一条路线：先把离散划分松弛为实向量上的连续优化问题，再将连续解舍入为一个割。

“连续松弛—离散舍入”是谱算法反复使用的基本范式。Rayleigh–Ritz 变分原理、Cauchy–Schwarz 不等式等通用工具汇总在附录 B；这里集中讨论它们在图上的含义。

从割到谱算法

割目标 $\rightarrow$ Rayleigh 商松弛 $\rightarrow \mu_2$ 的特征向量 $\rightarrow$ 扫描舍入。

第二特征值刻画连续松弛的最优值，扫描引理则控制从连续解回到离散割时的损失。

普通最小割往往偏爱很小的一侧：如果某个顶点度数很小，把它单独切出来就可能最优。若目标是找出一个内部相对紧密、与外部连接较弱的顶点簇，就应当用较小一侧的规模来归一化边界。对一个 $d \geq 1$ 的 d -正则图，定义集合 S 的电导率 (conductance) 为

$$\phi(S) = \frac{|\delta(S)|}{d \min\{|S|, |\bar{S}|\}}, \quad \phi(G) = \min_{\emptyset \neq S \subsetneq V} \phi(S).$$

分母正是较小一侧的体积，因而 $0 \leq \phi(S) \leq 1$ ； $\phi(S)$ 衡量的是这些边端点中有多大比例通向集合外部。

相关术语

在 d -正则图上，

$$h(G) = \min_{|S| \leq n/2} \frac{|\delta(S)|}{|S|}$$

通常称为边扩展率 (edge expansion)，并且 $\phi(G) = h(G)/d$ 。均匀稀疏割 (Uniform Sparsest Cut) 最小化

$$\frac{n|\delta(S)|}{|S||\bar{S}|}.$$

把 S 选作较小一侧后，这一目标与 $|\delta(S)|/|S|$ 相差至多两倍。非均匀稀疏割还给定一张独立的需求图，属于另一个问题。

4.1 邻接矩阵、拉普拉斯矩阵与 Rayleigh 商

设 $A \in \mathbb{R}^{n \times n}$ 是图的邻接矩阵，其中 A_{uv} 等于 u, v 之间平行边的条数。矩阵 A 实对称，因此存在一组标准正交的特征向量，所有特征值均为实数。将它们按降序写作

$$\theta_1 \geq \theta_2 \geq \cdots \geq \theta_n.$$

记 $\mathbf{1} \in \mathbb{R}^V$ 为全一向量。

命题 4.1 (正则图谱的基本性质)。对任意 d -正则多重图：

- (i) 对每个 i , 都有 $|\theta_i| \leq d$;
- (ii) $A\mathbf{1} = d\mathbf{1}$, 所以 $\theta_1 = d$;
- (iii) 特征值 d 的重数恰好等于图的连通分量数。特别地, G 连通当且仅当 $\theta_2 < d$ 。

证明. 若 $Ax = \theta x$, 取满足 $|x_i| = \max_j |x_j| > 0$ 的坐标 i , 则

$$|\theta||x_i| = \left| \sum_j A_{ij}x_j \right| \leq \sum_j A_{ij}|x_j| \leq d|x_i|,$$

故 (i) 成立。每一行的元素和都是 d , 立即得到 (ii)。

下面先设图连通, 并假设 $Ax = dx$ 。必要时把 x 换成 $-x$, 可取 i 使 $x_i = \max_j x_j > 0$ 。于是

$$dx_i = \sum_j A_{ij}x_j \leq \sum_j A_{ij}x_i = dx_i.$$

上式取等意味着 i 的每个邻居都具有同一数值 x_i ; 利用连通性逐层传播, 得到所有坐标均相等。因此连通图的 d -特征空间由 $\mathbf{1}$ 张成。若图不连通, 各连通分量的示性向量是线性无关的 d -特征向量; 在每个连通分量上重复上述论证, 便知它们张成整个 d -特征空间。□

对割问题而言, 使用归一化拉普拉斯矩阵 (*normalized Laplacian*)

$$\mathcal{L} = I - \frac{1}{d}A$$

更为方便。它的特征值满足

$$0 = \mu_1 \leq \mu_2 \leq \cdots \leq \mu_n \leq 2, \quad \mu_i = 1 - \frac{\theta_i}{d}.$$

任取 $f \in \mathbb{R}^V$, 有拉普拉斯能量恒等式

$$f^\top \mathcal{L} f = \frac{1}{d} \sum_{uv \in E} (f_u - f_v)^2. \quad (4)$$

确实, $f^\top A f = 2 \sum_{uv \in E} f_u f_v$, 而 $d \sum_v f_v^2 = \sum_{uv \in E} (f_u^2 + f_v^2)$; 展开右侧平方项即得结论。

若 G 连通, 可把第一个单位特征向量取为 $\mathbf{1}/\sqrt{n}$ 。将任意 $f \perp \mathbf{1}$ 在其余标准正交特征向量上展开, 得到变分刻画

$$\mu_2 = \min_{\substack{f \neq 0 \\ f \perp \mathbf{1}}} \mathcal{R}(f), \quad \mathcal{R}(f) := \frac{\sum_{uv \in E} (f_u - f_v)^2}{d \sum_{v \in V} f_v^2}. \quad (5)$$

这里的 $\mathcal{R}(f)$ 称为 Rayleigh 商 (*Rayleigh quotient*)。在特征向量基底下, 它就是 $\mu_2, \dots, \mu_n$ 的一个加权平均, 因此其最小值为 μ_2 。

到此为止, 原先对顶点集合的离散优化已经变成了对实向量的连续优化。接下来要解决的是舍入问题: 怎样从一个低能量向量中构造一个低电导率的割?

4.2 扫描舍入

取 μ_2 的一个特征向量 f ，等价地，它也是 A 对应于 θ_2 的特征向量，并令 $f \perp \mathbf{1}$ 。按坐标从大到小排列顶点：

$$f(u_1) \geq f(u_2) \geq \cdots \geq f(u_n),$$

再令 $P_i = \{u_1, \dots, u_i\}$ ，其中 $1 \leq i < n$ 。依次检查这些前缀的过程称为扫描舍入 (sweep rounding)。

正则图上的 SpectralSweep(G)

1. 计算一个非零的 μ_2 -特征向量 $f \in \mathbb{R}^n$ ，并使 $f \perp \mathbf{1}$ 。
2. 按 f 的坐标非增次序排列顶点。
3. 对每个 $1 \leq i < n$ ，计算 $\phi(P_i) = |\delta(P_i)| / [d \min\{i, n-i\}]$ 。
4. 返回电导率最小的前缀 P_{i^*} 。

这里不能取 $i = n$ ，因为 $P_n = V$ 的补集为空，电导率没有定义。排序完成后，可以在逐个加入顶点时更新边界，从而在 $O(m+n)$ 时间内求出所有前缀的边界；排序本身需要 $O(n \log n)$ 时间。

为突出数学主线，我们先在精确实数计算模型下描述算法，即假定可以精确求得特征向量。实际计算并不需要这一假设：对任意非零 $f \perp \mathbf{1}$ ，只要 $\mathcal{R}(f) \leq \eta$ ，下面的“中位数加扫描”证明就能找到电导率至多为 $\sqrt{2\eta}$ 的前缀。因此，数值算法应直接在 $\mathbf{1}^\perp$ 上求得满足该 Rayleigh 商界的近似向量；若先得到未正交的向量，则必须保证其投影 $\tilde{f} - \tilde{f}\mathbf{1}$ 仍满足 $\mathcal{R} \leq \eta$ ，再对这一投影应用扫描保证。计算成本取决于精度和谱间隙；坐标相同时可任意决定次序。

4.3 Cheeger 不等式及其舍入证明

连接连续谱量 μ_2 与离散电导率 $\phi(G)$ 的桥梁，是 Cheeger 不等式 (Cheeger inequality)。它的两个方向承担不同任务： $\mu_2 \leq 2\phi(G)$ 说明一个低电导率割可以构造出低能量向量，确立了谱问题作为松弛的合理性； $\phi(G) \leq \sqrt{2\mu_2}$ 则把低能量向量舍入为一个好割，是算法保证的核心。

定理 4.2 (正则图的 Cheeger 不等式)。对任意 d -正则图，

$$\frac{\mu_2}{2} \leq \phi(G) \leq \sqrt{2\mu_2}.$$

而且，SpectralSweep 返回的集合 S 满足

$$\phi(S) \leq \sqrt{2\mu_2} \leq 2\sqrt{\phi(G)}.$$

用邻接矩阵的特征值表示时， $\mu_2 = 1 - \theta_2/d$ 。

先处理不连通的情形。由命题 4.1，此时 $\mu_2 = 0$ ；任取一个连通分量，其边界为空，所以 $\phi(G) = 0$ 。此外，一个与 $\mathbf{1}$ 正交的 μ_2 -特征向量在每个连通分量上为常数，并且至少取两个不同值。取介于两个相邻坐标值之间的阈值，所得前缀是若干连通分量之并，电导率为零。因此定理的数值结论和算法结论均立即成立。以下两段证明只需考虑连通图。

证明：稀疏的割导出低能量向量。固定 S ，令 $s = |S| \leq n/2$ ，并取

$$x = \mathbf{1}_S - \frac{s}{n} \mathbf{1}.$$

显然 $x \perp \mathbf{1}$ 。Rayleigh 商的分子中，只有跨割边有贡献，并且每条这样的边贡献 1；另一方面，

$$\sum_v x_v^2 = \frac{s(n-s)}{n}.$$

由式 (5)，

$$\mu_2 \leq \mathcal{R}(x) = \frac{n|\delta(S)|}{ds(n-s)} \leq \frac{2|\delta(S)|}{ds} = 2\phi(S).$$

对 S 取最小值，即得 $\mu_2/2 \leq \phi(G)$ 。□

反向不等式包含算法的核心：必须把一个低能量向量变回一个割。我们先单独证明所需的舍入事实。

引理 4.3 (扫描引理)。设 $g: V \rightarrow \mathbb{R}_{\geq 0}$ 非零，并且其支撑集至多含 $n/2$ 个顶点。在所有非空阈值集合 $S_t = \{v: g_v^2 \geq t\}$ ($t > 0$) 中，至少有一个满足

$$\phi(S_t) \leq \sqrt{2\mathcal{R}(g)}.$$

证明。分层积分公式（也称离散余面积公式，discrete coarea formula）给出

$$\begin{aligned} \int_0^\infty |\delta(S_t)| dt &= \sum_{uv \in E} |g_u^2 - g_v^2|, \\ \int_0^\infty d|S_t| dt &= d \sum_v g_v^2. \end{aligned}$$

第一式可以逐边理解：边 uv 落在 $\delta(S_t)$ 中，当且仅当 t 位于 g_u^2 与 g_v^2 之间。对每个 $t > 0$ ，集合 S_t 都包含在 g 的支撑集中，故 $|S_t| \leq n/2$ ；单独的阈值 $t = 0$ 是零测集，不影响积分。比较上述两个积分并作平均，可知某个非空 S_t 满足

$$\phi(S_t) \leq \frac{\sum_{uv \in E} |g_u^2 - g_v^2|}{d \sum_v g_v^2}.$$

再利用 Cauchy-Schwarz 不等式和正则性：

$$\begin{aligned} \sum_{uv \in E} |g_u^2 - g_v^2| &\leq \left(\sum_{uv \in E} (g_u - g_v)^2 \right)^{1/2} \left(\sum_{uv \in E} (g_u + g_v)^2 \right)^{1/2} \\ &\leq \left(\sum_{uv \in E} (g_u - g_v)^2 \right)^{1/2} \left(2d \sum_v g_v^2 \right)^{1/2}. \end{aligned}$$

除以 $d \sum_v g_v^2$ ，便得到 $\sqrt{2\mathcal{R}(g)}$ 。□

定理 4.2 中扫描方向的证明。取满足 $f \perp \mathbf{1}$ 的 μ_2 -特征向量。选取其坐标的一个中位数 m ，使得严格大于 m 和严格小于 m 的坐标各至多有 $n/2$ 个。令 $h = f - m\mathbf{1}$ 。平移一个常数不改变任何边上的坐标差，并且由 $f \perp \mathbf{1}$ ，

$$\sum_v h_v^2 = \sum_v f_v^2 + nm^2 \geq \sum_v f_v^2.$$

所以 $\mathcal{R}(h) \leq \mathcal{R}(f) = \mu_2$ 。

把 h 分解为 $h = h_+ - h_-$ ，其中 $h_+ = \max\{h, 0\}$ ， $h_- = \max\{-h, 0\}$ 。两者的支撑集都至多含 $n/2$ 个顶点。对每条边都有

$$(h_u - h_v)^2 \geq (h_{+,u} - h_{+,v})^2 + (h_{-,u} - h_{-,v})^2,$$

同时 $\|h\|_2^2 = \|h_+\|_2^2 + \|h_-\|_2^2$ 。因此，在两个非零向量 $g \in \{h_+, h_-\}$ 中，至少有一个满足 $\mathcal{R}(g) \leq \mathcal{R}(h) \leq \mu_2$ 。

由扫描引理， g 的某个阈值集合的电导率不超过 $\sqrt{2\mu_2}$ 。若 $g = h_+$ ，该阈值集合是按 f 排序后的一个前缀；若 $g = h_-$ ，它是一个后缀，而其补集是具有相同电导率的前缀。因此 SpectralSweep 必然考察到一个电导率不大于该阈值割的前缀。最后，已经证明的 $\mu_2 \leq 2\phi(G)$ 蕴含 $\sqrt{2\mu_2} \leq 2\sqrt{\phi(G)}$ 。□

4.4 非正则图

对非正则图乃至带权图，顶点个数不再能准确反映一侧的“大小”，自然的归一化尺度是顶点体积。设对称权重 $w_{uv} = w_{vu} \geq 0$ ，令 $A_{uv} = w_{uv}$ 、 $d_v = \sum_u w_{uv} > 0$ ，并定义

$$D = \text{diag}(d_v), \quad L = D - A, \quad \mathcal{L} = I - D^{-1/2}AD^{-1/2}.$$

这里仍把 $\mathcal{L}$ 称为归一化拉普拉斯矩阵。进一步定义

$$\begin{aligned} \text{vol}(S) &= \sum_{v \in S} d_v, & w(\delta(S)) &= \sum_{u \in S, v \notin S} w_{uv}, \\ \Phi(S) &= \frac{w(\delta(S))}{\min\{\text{vol}(S), \text{vol}(\bar{S})\}}, & \Phi(G) &= \min_{\emptyset \neq S \subsetneq V} \Phi(S). \end{aligned}$$

以下假定图中没有孤立顶点；若有，可先将其删除。若令 $z = D^{1/2}f$ ，则 $z \perp D^{1/2}\mathbf{1}$ 当且仅当 $\sum_v d_v f_v = 0$ 。Rayleigh-Ritz 原理给出

$$\mu_2 = \min_{\substack{f \neq 0 \\ \sum_v d_v f_v = 0}} \mathcal{R}_D(f), \quad \mathcal{R}_D(f) = \frac{\sum_{uv} w_{uv} (f_u - f_v)^2}{\sum_v d_v f_v^2},$$

其中分子对每条无向带权边只求和一次。因此，若 z 是 $\mathcal{L}$ 的第二特征向量，扫描排序所使用的坐标应当是 $f_v = z_v / \sqrt{d_v}$ ，而不是直接使用 z_v 。

定理 4.4 (带权图的 Cheeger 不等式)。 对上述带权图，

$$\frac{\mu_2}{2} \leq \Phi(G) \leq \sqrt{2\mu_2}.$$

按 $f_v = z_v / \sqrt{d_v}$ 排序并作阈值扫描，可以找到达到上述上界的集合。

证明。 若图不连通，任一具有正体积的连通分量都没有边界，故 $\mu_2 = \Phi(G) = 0$ 。取一个对应特征值 0、并在不同连通分量上取不同常数值特征向量，在两个相邻取值之间选择阈值，就得到相应的零边界扫描割。以下设图连通。

先证容易的方向。取满足 $\text{vol}(S) \leq \text{vol}(V)/2$ 的集合 S ，令

$$x = \mathbf{1}_S - \frac{\text{vol}(S)}{\text{vol}(V)} \mathbf{1}.$$

则 $\sum_v d_v x_v = 0$ ，其能量等于 $w(\delta(S))$ ，而且

$$\sum_v d_v x_v^2 = \frac{\text{vol}(S) \text{vol}(\bar{S})}{\text{vol}(V)}.$$

因此 $\mu_2 \leq \mathcal{R}_D(x) \leq 2\Phi(S)$ 。对 S 取最小值，即得 $\mu_2/2 \leq \Phi(G)$ 。

下面证明扫描方向。取一个达到最小 Rayleigh 商的 f ，再选取它的一个加权中位数 m ，使集合 $\{v : f_v > m\}$ 与 $\{v : f_v < m\}$ 的体积均不超过 $\text{vol}(V)/2$ 。令 $h = f - m\mathbf{1}$ 。平移不改变能量，并且因为 $\sum_v d_v f_v = 0$ ，

$$\sum_v d_v h_v^2 = \sum_v d_v f_v^2 + m^2 \text{vol}(V) \geq \sum_v d_v f_v^2.$$

将 $h = h_+ - h_-$ 分成正部与负部。正则图证明中逐边使用的不等式仍然成立，所以在两个非零向量 $g \in \{h_+, h_-\}$ 中，至少有一个满足 $\mathcal{R}_D(g) \leq \mu_2$ ；同时，它的支撑集体积不超过 $\text{vol}(V)/2$ 。

对阈值集合 $S_t = \{v : g_v^2 \geq t\}$ ，带权余面积公式为

$$\int_0^\infty w(\delta(S_t)) dt = \sum_{uv} w_{uv} |g_u^2 - g_v^2|, \quad \int_0^\infty \text{vol}(S_t) dt = \sum_v d_v g_v^2.$$

所有 S_t 都包含在体积不超过一半的支撑集中。先对阈值平均，再用 Cauchy-Schwarz 不等式，可得某个非空阈值集合满足

$$\begin{aligned} \Phi(S_t) &\leq \frac{\sum_{uv} w_{uv} |g_u^2 - g_v^2|}{\sum_v d_v g_v^2} \\ &\leq \frac{\sqrt{\sum_{uv} w_{uv} (g_u - g_v)^2} \sqrt{\sum_{uv} w_{uv} (g_u + g_v)^2}}{\sum_v d_v g_v^2} \\ &\leq \sqrt{2\mathcal{R}_D(g)} \leq \sqrt{2\mu_2}, \end{aligned}$$

其中最后一步使用了 $\sum_{uv} w_{uv} (g_u + g_v)^2 \leq 2\sum_v d_v g_v^2$ 。 h_+ 的阈值集合是 f 排序后的前缀； h_- 的阈值集合是后缀，其补集是电导率相同的前缀。因此扫描过程一定能找到定理所述的集合。□

若图无权且 d -正则，则 $\text{vol}(S) = d|S|$ 且 $\Phi = \phi$ ，故该定理正好退化为定理 4.2。

4.5 谱坐标的两个应用

图像分割。 把像素或小图像区域看作顶点，在位置相近且视觉特征相似的顶点之间连边，并赋予较大权重。Shi 和 Malik 提出的归一化割 (normalized cut) 目标为 [14]

$$\text{Ncut}(S, \bar{S}) = \frac{w(\delta(S))}{\text{vol}(S)} + \frac{w(\delta(\bar{S}))}{\text{vol}(\bar{S})}.$$

它与电导率并不完全相同。若 $\text{vol}(S) \leq \text{vol}(\bar{S})$ ，则

$$\Phi(S) \leq \text{Ncut}(S, \bar{S}) \leq 2\Phi(S).$$

二者的标准谱松弛都以归一化拉普拉斯矩阵为核心：特征向量为每个像素赋予连续坐标，再以阈值将其划分为离散区域。

图的可视化。 取两个非平凡的拉普拉斯特征向量，把顶点 v 映射到 $(f_2(v), f_3(v))$ 。连接模式相近的顶点往往会在平面中落到相近位置。以顶点编号为 $j = 0, \dots, n-1$ 的环 C_n 为例，

$$f_2(j) = \cos(2\pi j/n), \quad f_3(j) = \sin(2\pi j/n)$$

张成第一个非平凡拉普拉斯特征空间；相应的二维谱嵌入恰好把全部顶点均匀放在一个圆周上。

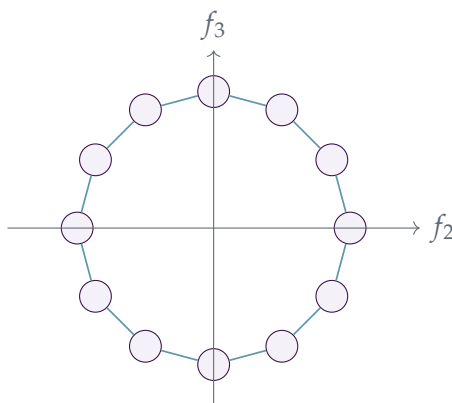


图 3: 环图的精确二维谱嵌入。

第三部分小结

谱算法给出的不是常数近似，而是平方根型保证：它返回的 S 满足 $\phi(S) \leq 2\sqrt{\phi(G)}$ 。它的关键价值在于，连续向量松弛可以高效计算，而舍入过程又具有严格且构造性的性能保证。当图还具有额外的谱间隙或簇结构时，通常能够得到远强于这一通用界的结论。

5 习题

标有 $\star$ 的题目需要在套用正文公式之外再加入一步关键思考；标有 $\star\star$ 的题目可作为小型课程项目。

第一部分：随机算法与随机化技术

1. **成功概率的放大。** 仅从单次成功概率下界 $p \geq 2/[n(n-1)]$ 出发，求出足以保证失败概率不超过任意给定 $\delta \in (0,1)$ 的最小整数重复次数。再将结果具体化到 $\delta = n^{-3}$ 的情形。
2. **收缩操作的不变量。** 用原图顶点集的一个划分来严格描述收缩后的图。直接证明：收缩图中的每个割都能提升为原图中一个大小完全相同的割。
3. **达到分析下界的例子。** 证明环图 C_n 恰有 $\binom{n}{2}$ 个不同的最小割（将互补的两侧视为同一个无序二划分），并证明 RandomContraction 输出其中每一个的概率都恰好为 $1/\binom{n}{2}$ 。
4. $\star$ **带权收缩。** 对正边权图，令 $w(\delta(S)) = \sum_{e \in \delta(S)} w_e$ ，并按与边权成正比的概率选择待收缩边。在这一设定下推广定理 2.3。同时说明运行时间模型需要作哪些调整。
5. $\star$ **Karger-Stein 算法中的常数。** 根据第 2.4 节，给出递归深度的一个显式上界；再给出一个显式常数 $c > 0$ ，使得对所有充分大的 n 都有 $p(n) \geq c/\log_2 n$ 。

第二部分：组合优化中的近似算法

6. **条件期望公式。** 考虑 Max-Cut 的一个部分赋值：已经放置的顶点分别组成 S, T ，尚未放置的顶点组成 U ；用 $E(U)$ 表示两个端点都在 U 中的边集。验证

$$\mathbb{E}[Z \mid S, T] = |E(S, T)| + \frac{1}{2}|E(S \cup T, U)| + \frac{1}{2}|E(U)|.$$

由此重新推出贪心算法在每一步所作的选择。

7. **带权 Max-Cut。** 假设所有边权均非负。证明随机割论证和相应的贪心算法 都能得到权重至少为总边权一半的割，因而给出一个 $1/2$ -近似算法。再举例说明，一旦允许负边权，这套证明为何失效。
8. **两两独立，但不相互独立。** 设 A, B 是两个相互独立的均匀随机比特，且 $C = A \oplus B$ 。证明 A, B, C 中任意两个都相互独立，但三个变量并不相互独立。若将三角形的三个顶点分别赋值为 A, B, C ，所得割值服从怎样的分布？
9. **★★ 快速计算所有种子的得分。** 给每个顶点一个标号 $a_v \in \mathbb{F}_2^\ell$ ，并令 c_z 表示满足 $a_u + a_v = z$ 的边 uv 的条数。证明种子 r 产生的割值为

$$\frac{1}{2} \left(m - \sum_{z \in \mathbb{F}_2^\ell} c_z (-1)^{\langle z, r \rangle} \right).$$

说明如何利用 Walsh-Hadamard 变换，在字长为 $\Omega(\log(m+n))$ 比特的 word-RAM 模型上，用 $O(m+n \log n)$ 次运算同时计算所有种子的割值。

第三部分：谱算法与谱图论

10. **连通分量与特征值重数。** 补全命题 4.1 中按连通分量进行的论证，证明特征值 d 的重数恰好等于图的连通分量数。
11. **检验 Cheeger 不等式中的常数。** 对 $n \geq 3$ 和 $r \geq 2$ ，分别计算 K_n, C_n 与 $K_{r,r}$ 的 μ_2 和 $\phi(G)$ 。写出精确电导率中所有与奇偶性有关的情形，并逐一比较定理 4.2 的两个方向，指出哪些图上、哪一步不等式存在松弛。
12. **★ 线性时间计算扫描割。** 顶点排序完成以后，设计一个 $O(m+n)$ 时间的算法，同时算出所有 $|\delta(P_i)|$ 。提示：当 u_i 加入前缀集合时，根据一条关联边的另一个端点是否已经加入前缀，分别更新割边数。
13. **★ 非正则图上的 Cheeger 扫描。** 令 $L = D - A$ 且 $D = \text{diag}(d_v)$ 。从广义特征值问题 $Lf = \mu Df$ 出发，重新证明容易的一侧 $\mu_2 \leq 2\Phi(G)$ ，并明确指出正则图证明中的顶点数在何处、为何应当改为体积。

A 概率与随机化工具

本讲涉及的随机试验均在有限样本空间 Ω 上进行。概率分布为每个结果 $\omega \in \Omega$ 指定权重 $p(\omega) \geq 0$ ，满足 $\sum_{\omega \in \Omega} p(\omega) = 1$ 。事件就是集合 $A \subseteq \Omega$ ，其概率为

$$\mathbb{P}(A) = \sum_{\omega \in A} p(\omega).$$

由定义立刻得到余事件公式 $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$ ，以及单调性：若 $A \subseteq B$ ，则 $\mathbb{P}(A) \leq \mathbb{P}(B)$ 。

A.1 条件概率与自适应随机过程

若 $\mathbb{P}(B) > 0$ ，则在事件 B 已经发生的条件下，事件 A 的条件概率定义为

$$\mathbb{P}(A | B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

分析自适应随机算法时，每一步的选择概率都必须以前面已经观察到的历史为条件。

命题 A.1 (概率的链式法则)。 设 $A_1, \dots, A_r$ 为一列事件，并记 $H_i = \bigcap_{j=1}^i A_j$ 、 $H_0 = \Omega$ 。若 $\mathbb{P}(H_i) > 0$ 对所有 $0 \leq i < r$ 成立，则

$$\mathbb{P}\left(\bigcap_{i=1}^r A_i\right) = \prod_{i=1}^r \mathbb{P}(A_i | H_{i-1}).$$

证明。 由条件概率的定义， $\mathbb{P}(H_i) = \mathbb{P}(H_{i-1})\mathbb{P}(A_i | H_{i-1})$ 。将这些等式相乘，中间各项逐一消去，即得结论。这里的正概率条件并非技术性的装饰：当条件事件的概率为零时，通常意义下的条件概率没有定义。 $\square$

A.2 期望与条件期望

随机变量是一个函数 $X: \Omega \rightarrow \mathbb{R}$ ，其期望定义为

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} p(\omega)X(\omega).$$

事件 A 的示性变量 $\mathbf{1}_A$ 在 A 上取值 1、在 A 外取值 0，因此 $\mathbb{E}[\mathbf{1}_A] = \mathbb{P}(A)$ 。

若事件 H 满足 $\mathbb{P}(H) > 0$ ，则给定 H 时的条件期望为

$$\mathbb{E}[Z | H] = \sum_{\omega \in H} \mathbb{P}(\omega | H)Z(\omega) = \frac{1}{\mathbb{P}(H)} \sum_{\omega \in H} p(\omega)Z(\omega).$$

定理 A.2 (期望的线性性)。 对任意随机变量 $X_1, \dots, X_r$ 和常数 $a_1, \dots, a_r$ ，都有

$$\mathbb{E}\left[\sum_{i=1}^r a_i X_i\right] = \sum_{i=1}^r a_i \mathbb{E}[X_i].$$

这一结论不要求随机变量之间相互独立。

证明。 只需交换两个有限求和的次序：

$$\sum_{\omega} p(\omega) \sum_i a_i X_i(\omega) = \sum_i a_i \sum_{\omega} p(\omega) X_i(\omega).$$

$\square$

平均值原理

若 Z 是有限概率空间上的实随机变量，则至少存在一个结果 ω 使得 $Z(\omega) \geq \mathbb{E}[Z]$ ；否则所有 $Z(\omega)$ 都小于 $\mathbb{E}[Z]$ ，对它们加权平均后仍小于 $\mathbb{E}[Z]$ ，矛盾。这正是概率方法最基本的逻辑：平均表现达到某个水平，就能保证存在一个达到同样水平的确定性对象。

设事件 H 记录此前已经揭示的随机选择，且尚未揭示的随机比特 B 在给定 H 后仍均匀分布。按 $B = 0$ 与 $B = 1$ 分情形，得到

$$\mathbb{E}[Z | H] = \frac{1}{2}\mathbb{E}[Z | H, B = 0] + \frac{1}{2}\mathbb{E}[Z | H, B = 1].$$

因此， B 的两个取值中至少有一个不会降低条件期望。逐个固定随机比特并反复使用这一观察，就得到条件期望法：只要每一步的两个条件期望都能被高效计算，就可以完成去随机化，同时保留原算法的期望性能保证。

A.3 独立性与成功概率放大

事件 $A_1, \dots, A_r$ 相互独立，是指对每个 $I \subseteq \{1, \dots, r\}$ 都有

$$\mathbb{P}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathbb{P}(A_i).$$

离散随机变量的相互独立性，也用对所有可能取值均成立的相应乘法公式定义。若只要求其中每一对变量独立，便得到较弱的两两独立性。尽管两两独立并不蕴含相互独立，它已经足以支持第 3 节中的随机割分析。

命题 A.3 (独立重复与概率放大)。 若 R 次试验相互独立，且每次成功的概率至少为 p ，则它们全部失败的概率至多为

$$(1 - p)^R \leq e^{-pR}.$$

因此，只要 $R \geq p^{-1} \ln(1/\delta)$ ，失败概率就不超过 δ 。

证明。 独立性使各次试验的失败概率相乘；再应用对所有 $x \geq 0$ 成立的不等式 $1 - x \leq e^{-x}$ 即可。□

本讲义所说的以高概率，统一指概率为 $1 - o(1/n)$ 。每当使用概率放大时，都可以适当选择重复次数中的常数，例如把失败概率进一步压低到 n^{-3} 。

B 线性代数工具

本附录只汇集谱算法部分反复使用的通用线性代数事实。邻接矩阵、图拉普拉斯矩阵、电导率以及扫描舍入等与图有关的概念，均在正文中引入和分析。

对向量 $x, y \in \mathbb{R}^n$ ，记 $\langle x, y \rangle = x^T y$ 以及 $\|x\|_2 = \sqrt{\langle x, x \rangle}$ 。记号 $x \perp y$ 表示 $\langle x, y \rangle = 0$ 。

命题 B.1 (Cauchy-Schwarz 不等式)。 对任意 $x, y \in \mathbb{R}^n$ ，都有

$$|\langle x, y \rangle| \leq \|x\|_2 \|y\|_2.$$

证明。 当 $y = 0$ 时结论显然成立。否则，对每个实数 t 都有 $0 \leq \|x - ty\|_2^2$ 。取 $t = \langle x, y \rangle / \|y\|_2^2$ 并整理，即得 $\langle x, y \rangle^2 \leq \|x\|_2^2 \|y\|_2^2$ 。□

定理 B.2 (实对称矩阵的谱定理)。 每个实对称矩阵 $M \in \mathbb{R}^{n \times n}$ 的全部特征值均为实数，并且存在一组由实特征向量构成的标准正交基。等价地，存在正交矩阵 Q 与实对角矩阵 Λ ，使得 $M = Q\Lambda Q^T$ 。

命题 B.3 (Rayleigh–Ritz 变分原理)。 设 M 是实对称矩阵，其特征值依次为 $\theta_1 \leq \theta_2 \leq \dots \leq \theta_n$ ，相应的标准正交特征向量为 $q_1, \dots, q_n$ 。则

$$\min_{x \neq 0} \frac{x^T M x}{x^T x} = \theta_1, \quad \min_{\substack{x \neq 0 \\ x \perp q_1}} \frac{x^T M x}{x^T x} = \theta_2.$$

证明。 将 x 在特征向量基下展开为 $x = \sum_i a_i q_i$ ，则

$$\frac{x^T M x}{x^T x} = \frac{\sum_i \theta_i a_i^2}{\sum_i a_i^2}.$$

这个商是各特征值的加权平均，因此其最小值为 θ_1 ，并在 q_1 处达到。若再要求 $x \perp q_1$ ，则 $a_1 = 0$ ，此时最小值变为 θ_2 ，并在 q_2 处达到。 $\square$

文献说明

第一部分：随机算法与随机化技术。 Karger 的收缩算法最早发表于 1993 年的 SODA 论文 [1]；递归化的改进由 Karger 和 Stein 提出 [2]。Karger 后来又给出了近线性时间的随机算法 [3]。Stoer 和 Wagner 提出了经典的确定性算法 [4]；Henzinger、Li、Rao 和 Wang 则给出了带权图上的现代确定性近线性时间结果 [5]。关于随机算法的更广泛背景，可参考 Motwani 和 Raghavan 的专著 [15]。

第二部分：组合优化中的近似算法。 Max-Cut 的经典 NP 完全性结果见 [6]。基于半正定规划的近似算法由 Goemans 和 Williamson 提出 [7]。Khot 提出了 Unique Games Conjecture [8]；在该猜想成立的条件下，Goemans–Williamson 近似比的最优性由 Khot、Kindler、Mossel 和 O’Donnell 证明 [9]。Williamson 和 Shmoys 的专著 [16] 系统介绍了近似算法。

第三部分：谱算法与谱图论。 Chung 的专著 [10] 是谱图论的标准参考资料。Cheeger 几何不等式在离散图上的对应结果由 Dodziuk 以及 Alon–Milman 奠定 [11, 12]；带权版本可见 [13]。谱方法与图像分割之间的经典联系由 Shi 和 Malik 系统阐述 [14]。

参考文献

- [1] David R. Karger. Global min-cuts in RNC, and other ramifications of a simple min-cut algorithm. In *Proceedings of SODA*, pages 21-30, 1993. <https://doi.org/10.5555/313559.313605>.
- [2] David R. Karger and Clifford Stein. A new approach to the minimum cut problem. *Journal of the ACM*, 43(4):601-640, 1996. <https://doi.org/10.1145/234533.234534>.
- [3] David R. Karger. Minimum cuts in near-linear time. *Journal of the ACM*, 47(1):46-76, 2000. <https://doi.org/10.1145/331605.331608>.
- [4] Mechthild Stoer and Frank Wagner. A simple min-cut algorithm. *Journal of the ACM*, 44(4):585-591, 1997. <https://doi.org/10.1145/263867.263872>.
- [5] Monika Henzinger, Jason Li, Satish Rao, and Di Wang. Deterministic near-linear time minimum cut in weighted graphs. In *Proceedings of SODA*, pages 3089-3139, 2024. <https://doi.org/10.1137/1.9781611977912.111>.
- [6] Michael R. Garey, David S. Johnson, and Larry Stockmeyer. Some simplified NP-complete graph problems. *Theoretical Computer Science*, 1(3):237-267, 1976. [https://doi.org/10.1016/0304-3975\(76\)90059-1](https://doi.org/10.1016/0304-3975(76)90059-1).
- [7] Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42(6):1115-1145, 1995. <https://doi.org/10.1145/227683.227684>.
- [8] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of STOC*, pages 767-775, 2002. <https://doi.org/10.1145/509907.510017>.
- [9] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O' Donnell. Optimal inapproximability results for Max-Cut and other 2-variable CSPs. *SIAM Journal on Computing*, 37(1):319-357, 2007. <https://doi.org/10.1137/S0097539705447372>.
- [10] Fan R. K. Chung. *Spectral Graph Theory*. CBMS Regional Conference Series in Mathematics 92, AMS, 1997. <https://doi.org/10.1090/cbms/092>.
- [11] Jozef Dodziuk. Difference equations, isoperimetric inequality and transience of certain random walks. *Transactions of the American Mathematical Society*, 284(2):787-794, 1984. <https://doi.org/10.1090/S0002-9947-1984-0743744-X>.
- [12] Noga Alon and Vitali D. Milman. λ_1 , isoperimetric inequalities for graphs, and superconcentrators. *Journal of Combinatorial Theory, Series B*, 38(1):73-88, 1985. [https://doi.org/10.1016/0095-8956\(85\)90092-9](https://doi.org/10.1016/0095-8956(85)90092-9).
- [13] Fan R. K. Chung and Kevin Oden. Weighted graph Laplacians and isoperimetric inequalities. *Pacific Journal of Mathematics*, 192(2):257-273, 2000. <https://doi.org/10.2140/pjm.2000.192.257>.
- [14] Jianbo Shi and Jitendra Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):888-905, 2000. <https://doi.org/10.1109/34.868688>.
- [15] Rajeev Motwani and Prabhakar Raghavan. *Randomized Algorithms*. Cambridge University Press, 1995.
- [16] David P. Williamson and David B. Shmoys. *The Design of Approximation Algorithms*. Cambridge University Press, 2011.