Homework5 solution

授课时间: 2019 年 7 月 12 日 授课教师: John E. Hopcroft 批改：陈铖，孙周易

所有结论正确，过程合理的解答均可得到全部分数，结论错误，过程正确也可以得到相应的步骤分。

前两题的一些前提请参考课本 4.4.1。

Exercise 4.22

What is the mixing time for the undirected graphs

1. *Two cliques connected by a single edge? To simplify the problem assume vertices have self loops except for the two vertices with the edge connecting the two cliques. Thus all vertices have the same degree.*

参考解答： （20 分）

设每个团包含 n 个节点，整个图一共有 $2n$ 个节点。

由前提易知对于图中任意点 x，$\pi(x)=1/2n$，且对于相邻的点对 x,y，有 $p_{xy}=1/n$。

则对于图中的一个大小不超过 n 的点集 S，$\Phi(S)=\frac{\sum_{(x,y)\in(S,\bar{S})}\pi(x)p_{xy}}{\pi(S)}=\frac{E(S)/n}{|S|}$，其中 $E(S)$ 为 S 的出边集合。

故当 S 的出边尽可能少，$|S|$ 又尽可能大时，$\Phi(S)$ 会尽可能的小。

容易发现这两个条件可以同时取到，即令 S 为任意一个团。此时 $E(S)=1, |S|=n, \Phi=\Phi(S)=1/n^2$。

从而 ϵ-mixing time $=O\left(\frac{\ln(1/\pi_{\min})}{\Phi^2\epsilon^3}\right)=O\left(\frac{n^4\ln n}{\epsilon^3}\right)$

评分标准：

写出 π 与 p 得 5 分。

写出 $\Phi(S)$ 的表达式并求得其最小值 Φ 得 10 分。

求得正确的 ϵ-mixing time 得 5 分。

两个团顶点节点个数不同的情况几乎没有区别，提不提到都视作正确。

Exercise 4.24

Find the ϵ-mixing time for a 2-dimensional lattice with n vertices in each coordinate direction with a uniform probability distribution. To do this solve the following problems.

1. *The minimum number of edges leaving a set S of size greater than or equal to $n^2/4$ is n.*
2. *The minimum number of edges leaving a set S of size less than of equal to $n^2/4$ is $\left\lceil\sqrt{2|S|}\right\rceil$.*
3. *Compute $\Phi(S)$*
4. *Compute Φ*

5. *Compute the ϵ-mixing time*

参考解答： （45 分）

首先依次考虑以下事实：S 一定是连通的；S 是“紧凑”的；S 一定与紧贴边缘；由外到内长条长度不断递减，并最终 S 一定可以通过把角上的方块补到之前的列成为一个矩形加上一个可能的长条，使得往外连边的数目不变。

1. 假设 S 往外的连边小于 n，那么这个对应矩形的长也一定小于 n。而此时考虑它的长宽（忽略长条，因为对出边没有影响，还会让面积变小），不妨记作 a, b，有 $a+b<n$，则 $ab<n^2/4$，矛盾。
2. 同理，依旧考虑这个角落中的矩形（加一个长条）。若矩形的长为 n，则向外连边为 n。若小于 n，则同上可知，当其近似为一个正方形时最优，此时出边数量至少为 $\left\lceil\sqrt{2|S|}\right\rceil$。又由 $|S|<n^2/4$，命题得证。
3. $$\Phi(S)=\frac{\sum_{(x,y)\in(S,\bar{S})}\pi(x)p_{xy}}{\pi(S)}=\frac{1/4\times 1/n^2\times|E(S)|}{|S|/n^2}$$

   其中 $E(S)$ 为 S 的出边集合。

   当 $|S|\geq n^2/4$ 时，$\Phi(S)\geq\Omega\left(\frac{cn/n^2}{|S|/n^2}\right)=\Omega(1/n)$。

   当 $|S|\leq n^2/4$ 时，$\Phi(S)\geq\frac{c\sqrt{|S|}/n^2}{|S|/n^2}=\Omega\left(\frac{1}{\sqrt{|S|}}\right)=\Omega(1/n)$
4. 由上可知 $\Phi=\Omega(1/n)$。
5. 从而 ϵ-mixing time $=O\left(\frac{\ln(1/\pi_{\min})}{\Phi^2\epsilon^3}\right)=O\left(\frac{n^2\ln n}{\epsilon^3}\right)$

评分标准： 具体描述 S 的分布状况得 10 分。

1、2 证明各 7 分（不是特别关心具体能不能取到的问题，所以只要证明说得通而且能得到最后的结果都会给分）。

求得 $\Phi(S)$ 的表达式得 5 分，之后正确讨论两种情形各得 5 分。

4、5 答案各 3 分。

Exercise 5.1

Each of the following data sets consists of a subset of the d-dimensional $0/1$ *vectors labeled* $+1$. *The remaining* $0/1$ *vectors are labeled* -1. *Which sets are linearly separable?*

1. *{010, 011, 100, 111}*
2. *{011, 100, 110, 111}*
3. *{0100, 0101, 0110, 1000, 1100, 1101, 1110, 1111}*

参考解答： （25 分）

记给出的集合为 S，则线性可分即存在向量 $\mathbf{w}$ 与常数 t，使得对于任意 $\mathbf{x} \in S$ 都有 $\mathbf{w} \cdot \mathbf{x} > t$，对于任意 $\mathbf{x} \in \bar{S}$ 都有 $\mathbf{w} \cdot \mathbf{x} < t$。故可以列出线性规划解决。

另一方面，也可以考虑线性可分与两个点集对应凸包的关系。点集线性可分当且仅当两个凸包线性可分，而两个凸包线性可分当且仅当它们没有交集。用不同方式判断两个点集所对应凸包是否有交即可。

结果为 1、2 不可分，3 可分。

评分标准：

正确描述判断线性可分这一问题的转化得 10 分。之后每一个正确判断得 5 分。

Exercise 5.7

Find the mapping $\phi(x)$ that gives rise to the kernel

$$K(\mathbf{x}, \mathbf{y}) = (x_1 y_1 + x_2 y_2)^2 .$$

参考解答： （10 分）

$$\begin{aligned} K(\mathbf{x}, \mathbf{y}) &= (x_1 y_1 + x_2 y_2)^2 \\ &= x_1^2 y_1^2 + x_2^2 y_2^2 + 2 x_1 x_2 y_1 y_2 \\ &= \phi_1(\mathbf{x})\phi_1(\mathbf{y}) + \phi_2(\mathbf{x})\phi_2(\mathbf{y}) + \phi_3(\mathbf{x})\phi_3(\mathbf{y}) \end{aligned}$$

故令 $\phi(\mathbf{x}) = (x_1^2, x_2^2, \sqrt{2} x_1 x_2)$，有 $\phi(\mathbf{x})^T \phi(\mathbf{y}) = K(\mathbf{x}, \mathbf{y})$。

评分标准：

答案正确得 10 分。