

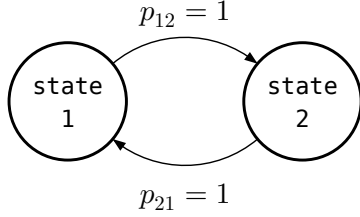
Mathematics for Computer Science:

Homework 9

Yiming Liu
2023010747

Problem 4.1

Answer: (i) We assume that $p(0) = (x_1, x_2)$.

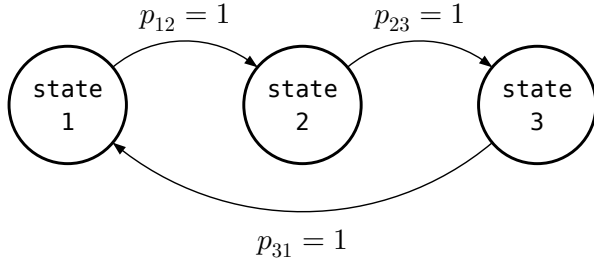


So, the transition matrix P is $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Since $p(t) = p(t-1)P$, we have $p(t) = p(0)P^t$.

$$P^t = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^t = \begin{cases} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & \text{if } 2|t \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \text{if } 2 \nmid t. \end{cases}$$

Therefore $p(t) = \begin{cases} (x_1, x_2) & \text{if } 2 \nmid t \\ (x_2, x_1) & \text{if } 2|t \end{cases}$. It converge iff $x_1 = x_2$.

(ii) We assume that $p(0) = (x_1, x_2, x_3)$.



So the transition matrix P is $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$. Since $p(t) = p(t-1)P$, we have $p(t) = p(0)P^t$.

$$P^t = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}^t = \begin{cases} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} & \text{if } t \equiv 1 \pmod{3} \\ \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} & \text{if } t \equiv 2 \pmod{3} \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \text{if } t \equiv 0 \pmod{3}. \end{cases}$$

Therefore $p(t) = \begin{cases} (x_1, x_2, x_3) & \text{if } t \equiv 0 \pmod{3} \\ (x_3, x_1, x_2) & \text{if } t \equiv 1 \pmod{3} \\ (x_2, x_3, x_1) & \text{if } t \equiv 2 \pmod{3} \end{cases}$. It converge iff $x_1 = x_2 = x_3$.

We can generalize these examples to n states with $p_{12} = p_{23} = p_{34} = \dots = p_{(n-1)n} = p_{n1} = 1$ and the other $p_{ij} = 0$. Similarly, we have P as a cyclic permutation matrix and $p(t)$ converge iff $x_1 = x_2 = \dots = x_n$. So $p(t)$ does not converge for every connected Markov Chain.

Problem 4.2

Answer: Let $a(t) = \log(t)$. We have $\lim_{t \rightarrow \infty} a(t) - a(t+1) = \lim_{t \rightarrow \infty} \log\left(1 - \frac{1}{t+1}\right) = 0$. But $a(t)$ does not converge.

Problem 4.3

Answer: (i) We can tell from the graph that the transition matrix P is

$$\begin{pmatrix} 0.4 & 0.6 & 0 & 0 & 0 \\ 0.4 & 0 & 0.6 & 0 & 0 \\ 0 & 0.4 & 0 & 0.6 & 0 \\ 0 & 0 & 0.4 & 0 & 0.6 \\ 0 & 0 & 0 & 0.4 & 0.6 \end{pmatrix}$$

We assume $\pi = (x_1, x_2, x_3, x_4, x_5)$. From $\pi = \pi P$ we know that

$$\begin{aligned} (x_1, x_2, x_3, x_4, x_5) &= (0.4x_1 + 0.4x_2, 0.6x_1 + 0.4x_3, 0.6x_2 + 0.4x_4, 0.6x_3 + 0.4x_5, 0.6x_4 + 0.6x_5) \\ x_1 + x_2 + x_3 + x_4 + x_5 &= 1 \end{aligned}$$

Solving the equations, we have $x_1 = \frac{16}{211}, x_2 = \frac{24}{211}, x_3 = \frac{36}{211}, x_4 = \frac{54}{211}, x_5 = \frac{81}{211}$. So the stationary distribution is $(\frac{16}{211}, \frac{24}{211}, \frac{36}{211}, \frac{54}{211}, \frac{81}{211})$.

(ii) We can tell from the graph that the transition matrix P is

$$\begin{pmatrix} 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

We assume $\pi = (x_1, x_2, x_3, x_4, x_5)$. From $\pi = \pi P$ we know that

$$\begin{aligned} (x_1, x_2, x_3, x_4, x_5) &= (0.5x_1, 0.5x_1 + 0.5x_2, 0.5x_2 + 0.5x_3, 0.5x_3 + 0.5x_4, 0.5x_4 + x_5) \\ x_1 + x_2 + x_3 + x_4 + x_5 &= 1 \end{aligned}$$

Solving the equations, we have $x_1 = x_2 = x_3 = x_4 = 0, x_5 = \frac{1}{5}$. So the stationary distribution is $(0, 0, 0, 0, 1)$.

Problem 4.10

Answer: The current state stay put with probability $\frac{1}{2}$ and move as in the old chain with probability $\frac{1}{2}$. Therefore,

$$p'_{ij} = \begin{cases} \frac{1}{2} + \frac{1}{2}p_{ii} & \text{if } i = j \\ \frac{1}{2}p_{ij} & \text{if } i \neq j. \end{cases}$$

We just need to prove that

$$\begin{aligned}
\pi_x &= \sum_y \pi_y p'_{yx} \\
\Leftrightarrow \pi_x &= \pi_x \frac{1}{2} + \pi_x \frac{1}{2} p_{ii} + \sum_{y \neq x} \pi_y \frac{1}{2} p_{yx} \\
\Leftrightarrow \pi_x &= \frac{\pi_x}{2} + \frac{\pi_x p_{ii}}{2} + \frac{1}{2} \left(\sum_y \pi_y p_{yx} - \pi_x p_{xx} \right) \\
\Leftrightarrow \pi_x &= \pi_x.
\end{aligned}$$

Therefore, the new chain has the same stationary distribution as the old chain.

And since the new chain tends to stay put instead of getting out, the conductance is much lower than the old chain. (We can also observe this from $\Phi = \frac{\sum_{(x,y) \in (S, \bar{S})} \pi_x p'_{xy}}{\min(\pi(S), \pi(\bar{S}))}$ since $p'_{xy} = \frac{p_{xy}}{2}$ for $x \neq y$) Thus the convergence time will be longer.

Problem 4.19

Answer: (1) For two n vertices cliques connected by a single edge, we have $p_{xy} = \frac{1}{n-1}$ if x, y are not the vertices on that connected single edge and $p_{x,y} = \frac{1}{n}$ if x, y at least one is a vertex of that connected single edge. Thus, $\pi_x = \frac{n}{2(n^2-n+1)}$ if vertex x is on that connected single edge and $\pi_x = \frac{n-1}{2(n^2-n+1)}$ if vertex x is not on that connected single edge.

With $\pi(S) \leq \frac{1}{2}$, we want S to have the smallest probability mass exiting it, which is that connected single edge.

Thus, we have $\Phi = \frac{\frac{n}{2(n^2-n+1)} \frac{1}{n}}{\frac{1}{2}} = \frac{1}{n^2-n+1}$.

Thus the ε -mixing time is $O\left(\frac{\ln\left(\frac{1}{\frac{n-1}{2(n^2-n+1)}}\right)}{\frac{\Phi^2 \varepsilon^3}{1}}\right) = O\left(\frac{(n^2-n+1)^2 \ln\left(\frac{2(n^2-n+1)}{n-1}\right)}{\varepsilon^3}\right) = O\left(\frac{n^4 \ln n}{\varepsilon^3}\right)$.

(2) For a n vertices cliques connected by a single edge, we have $p_{xy} = \frac{1}{n-1}$ if x, y are not the vertices on that connected single edge and $p_{x,y} = \frac{1}{n}$ if one x, y is a vertex of that connected single edge and $p_{x,y} = \frac{1}{n}$ and $p_{x,y} = 1$ if edge xy is that additional edge and x is on the clique. Thus, $\pi_x = \frac{n-1}{n^2-n+2}$ if vertex x is not on that connected single edge and $\pi_x = \frac{n}{n^2-n+2}$, $\pi_y = \frac{1}{n^2-n+2}$ if edge xy is that additional edge and x is on the clique.

Next, we prove a lemma.

We define canonical path $\gamma_{ij} = \{i = x_0, x_1, \dots, x_k = j\}$ and the congestion of the Markov chain as

$$\rho = \max_{(u,v)} \left\{ \frac{1}{\pi(u)P_{uv}} \sum_{\substack{x,y \in \chi \\ \gamma_{xy} \text{ uses } (u,v)}} \pi(x)\pi(y) \right\}.$$

Then we can show that $\Phi \geq \frac{1}{2\rho}$.

Proof. It suffices to show that for any S such that $\pi(S) \leq \frac{1}{2}$, we have $\frac{\text{flow}(S, \bar{S})}{\pi(S)} \geq \frac{1}{2\rho}$.

$$\begin{aligned}
2\rho\text{flow}(S, \bar{S}) &= 2\rho \sum_{\substack{u \in S \\ v \in \bar{S}}} \pi(u)P(u, v) \\
&\geq 2 \sum_{\substack{u \in S \\ v \in \bar{S}}} \rho_{(u,v)} \pi(u)P(u, v) \geq 2 \sum_{\substack{u \in S \\ v \in \bar{S}}} \sum_{\substack{x, y \in \chi \\ \gamma_{xy} \text{ uses } (u,v)}} \pi(u)P(u, v) \\
&\geq 2 \sum_{\substack{x \in S \\ y \in \bar{S}}} \pi(x)\pi(y) = 2\pi(S)\pi(\bar{S}) \geq \pi(S).
\end{aligned}$$

Therefor $\Phi \geq \frac{1}{2\rho}$.

so $t_{\text{mix}} \leq \frac{2}{\Phi^2} \log\left(\frac{1}{\pi_{\min}\varepsilon}\right) \leq 8\rho^2 \log\left(\frac{1}{\pi_{\min}\varepsilon}\right)$.

Now we go back to the original problem, we design the path to be straight, so every two vertices in the clique has path with one edge. For the addition vertex, the path is first go to the vertex on the clique and then go straight to the destination.

From this design of path we can know that

$$\begin{aligned}
\rho &= \max \left(\frac{1}{\frac{1}{n^2-n+2}} (n-1) \frac{1}{n^2-n+2} \frac{n}{n^2-n+2}, \right. \\
&\quad \frac{1}{\frac{1}{n^2-n+2}} 1 \frac{n-1}{n^2-n+2} \frac{n}{n^2-n+2}, \\
&\quad \left. \frac{1}{\frac{1}{n^2-n+2}} 1 \frac{n-1}{n^2-n+2} \frac{n-1}{n^2-n+2} \right) \\
&= \frac{n^2-n}{n^2-n+2}.
\end{aligned}$$

Therefore we have $t_{\text{mix}} = O\left(\left(\frac{n^2-n}{n^2-n+2}\right)^2 \log\left(\frac{n^2-n+2}{\varepsilon}\right)\right)$.

Problem 4.21

Answer: (1) We know that π_x and p_{xy} is the same for all vertices. Thus, if we want to minimize the conductance, we only need to minimize the number of edges between S and \bar{S} . (If two vertices are adjacent, we think they have an edge between them.) We use set A to denote the edges between S and \bar{S} .

Notice two fact:

1. If a vertex is not connected with other vertices, we can always adjust it to connected with other vertices and it will not increase the conductance.
2. If we have vertices connected to each other, we can always adjust it to form many square next to each other and it will not increase the conductance.

Because we know that:

1. We add a isolated vertex at corner and $|A|$ will increase by 2.
2. We add a isolated vertex in the border and $|A|$ will increase by 3.
3. We add a isolated vertex in the center and $|A|$ will increase by 4.
4. We connected a vertex and another vertex and $|A|$ will decrease by 2.

Thus it's always better to connect two isolated vertices to each other. And to form square is another effective way to reduce $|A|$ since it connected more vertices. So we have two cases: one

is that all the vertices are around the corner and the other is that the vertices take at least two corners.

The first case, we only take one corner.

For a given number of small square with 4 vertex, if the total size is given, then $|A|$ is minimized when these squares build a big square since square has the smallest perimeter.

Therefore, the minimum number of edges leaving S is at least $2\lfloor\sqrt{|S|}\rfloor$.

The second case, we take more than one corner.

If we take two corners, then the minimum number of edges leaving S is at least n . If we take more than two corners, then the \bar{S} will take less than two corners and we can use the above result.

Thus, the minimum number of edges leaving S is $\min(n, 2\sqrt{|S|})$.

(1) If $|S| \geq \frac{n^2}{4}$, the minimum number of edges leaving S is $\min(n, 2\sqrt{\frac{n^2}{4}}) = n$.

(2) If $|S| \leq \frac{n^2}{4}$, the minimum number of edges leaving S is $\min(n, 2\sqrt{|S|}) = 2\sqrt{|S|}$.

(3) If $|S| \geq \frac{n^2}{4}$, $\Phi = \Omega\left(\frac{n^{\frac{1}{2}\frac{1}{4}}}{\frac{n^2}{4}\frac{1}{n^2}}\right) = \Omega\left(\frac{1}{n}\right)$.

If $|S| \leq \frac{n^2}{4}$, $\Phi(S) = \Omega\left(\frac{\frac{\sqrt{|S|}}{n^2}\frac{1}{4}}{\frac{|S|}{n^2}}\right) = \Omega\left(\frac{1}{\sqrt{|S|}}\right) = \Omega\left(\frac{1}{n}\right)$

(4) $\Phi = \min_S(\Phi(S)) = \Omega\left(\frac{1}{n}\right)$.

(5) The ε -mixing time is $O\left(\frac{\ln(n^2)}{\Phi^2\varepsilon^3}\right) = O\left(\frac{n^2\ln n}{\varepsilon^3}\right)$.