离散傅里叶变换, 马尔科夫链

参考答案

1. (6 分) 对于一个布尔函数 $f:\{0,1\}^n \to \{0,1\}$, 函数第 i 位的影响被定义为

$$\text{Influence}_i(f) := \Pr_{x\leftarrow\{0,1\}^n}\big[f(x) \neq f(x\oplus e_i)\big],$$

其中 $e_i = (0,\dots,0,1,0,\dots,0)$ 只在第 i 位等于 1.

(1) 布尔函数 f 被称为单调（monotone）, 如果 $x \geq y \implies f(x) \geq f(y)$. （这里 $x \geq y$ 表示 $\forall i, x_i \geq y_i$.）请寻找一个单调的布尔函数 f, 使得 $\sum_i \text{Influence}_i(f)$ 最大, 并证明.

(2) 布尔函数 f 被称为平衡（balanced）, 如果 $\Pr_{x\leftarrow\{0,1\}^n}[f(x)=1] = \frac{1}{2}$. 请寻找一个平衡的布尔函数 f, 使得 $\max_i \text{Influence}_i(f)$ 尽量小, 可以忽略常数系数.

Remark: 证明 $\max_i \text{Influence}_i(f)$ 的下界需要非常有技巧地使用傅里叶变换. 本题不需要证明结果最优.

解

(1) 对于单调布尔函数 f,

$$\begin{aligned}
\sum_i \text{Influence}_i(f) &= \frac{1}{2^n}\sum_i\sum_x \mathbb{1}\big[f(x)\neq f(x\oplus e_i)\big] \\
&= \frac{1}{2^{n-1}}\sum_{i,x \text{ s.t. } x_i=0}\big(f(x\oplus e_i) - f(x)\big) \\
&= \frac{1}{2^{n-1}}\sum_{i,x}\begin{cases} f(x), & \text{if } x_i = 1 \\ -f(x), & \text{if } x_i = 0\end{cases} \\
&= \frac{1}{2^{n-1}}\sum_x\big(x \text{ 中 } 1 \text{ 的个数} - x \text{ 中 } 0 \text{ 的个数}\big).
\end{aligned}$$

因此, 使 $\sum_i \text{Influence}_i(f)$ 最大的 f 即为 majority 函数.

(2) 把 n 个 bit 分成 a 组 $G_1,\cdots,G_a$, 每组大约为 b 个 bit, 忽略常数项, 即 $ab=n$. 取 $f(x)$ 为 $[\exists i\in[a], \forall j \in G_i, x_j = 1]$. 从而有,

$$\Pr_{x\leftarrow\{0,1\}^n}[f(x)=1] = 1-(1-2^{-b})^a = \frac{1}{2}.$$

可推得 $b = \log n - \log(\log n) + O(1)$, 对于 Influence_i 仅需考虑 i 所在的组全是 1 的 x, 因此,

$$\max_i \text{Influence}_i(f) = 2^{1-b} = O(\frac{\log n}{n}).$$

2. (12 分) 对于函数 $f:\{0,1\}^n \to \mathbb{R}$, 用 f 的傅里叶系数表示如下量

(1) $\widehat{g_s}(x)$, 其中 $g(x) := f(x \oplus s)$.

(2) $\widehat{g_y}(x)$, 其中 $g(x) := (-1)^{\langle x,y \rangle} f(x)$.

(3) $\widehat{f_i}(x)$, 其中 $f_i(x) := f(x) - f(x \oplus e_i)$.

(4) $\hat{g}(x)$, 其中 $g(x_1, \ldots, x_k) := f(x_1, \ldots, x_k, 0, \ldots, 0)$.

(5) $\widehat{g_a}(x)$, 其中 $g(x_1, \ldots, x_k) := \mathbb{E}_{y \leftarrow \{0,1\}^{n-k}}[f(x, y)\chi_a(y)]$.

(6) $\mathrm{Var}[f]$.

这里约定 $\hat{f}(a) = \mathbb{E}_x[f(x)\overline{\chi_a(x)}]$, $f(x) = \sum_a \hat{f}(a)\chi_a(x)$.

解

(1) $\widehat{g_s}(x) = \mathbb{E}_y \left[(-1)^{\langle x,y \rangle} f(y \oplus s)\right] = (-1)^{\langle x,s \rangle}\mathbb{E}_y \left[(-1)^{\langle x,y \rangle} f(y)\right] = (-1)^{\langle x,s \rangle}\widehat{f}(x)$.

(2) $\widehat{g_s}(x) = \mathbb{E}_y \left[(-1)^{\langle x \oplus s,y \rangle} f(y)\right] = \widehat{f}(x \oplus s)$.

(3) $\widehat{f_i}(x) = \widehat{f}(x) - \widehat{g_{e_i}}(x) = (1 - (-1)^{x_i})\,\widehat{f}(x)$.

(4) $\widehat{g_k}(x) = \mathbb{E}_y \left[(-1)^{\langle x,y \rangle} f(y, 0)\right] = \mathbb{E}_y \left[\sum_z (-1)^{\langle y,x \oplus z \rangle} \sum_w \widehat{f}(z, w)\right] = \sum_z \sum_w \widehat{f}(z, w)\mathbb{E}_y \left[(-1)^{\langle y,x \oplus z \rangle}\right] = \sum_w \widehat{f}(x, w)$.

(5) $\widehat{g_a}(x) = \mathbb{E}_y \left[(-1)^{\langle x,y \rangle}\mathbb{E}_z \left[(-1)^{\langle a,z \rangle} f(y, z)\right]\right] = \widehat{f}(x, a)$.

(6) $\mathrm{Var}[f] = \mathbb{E}_x \left[f(x)^2\right] - \mathbb{E}_x[f(x)]^2 = \sum_{x \neq 0} \widehat{f}(x)^2$.

3. (4 分) 给定一个函数 $f : \{0,1\}^n \to [-1, 1]$, 你能看到若干独立采样 $(x, f(x))$（其中 x 在 $\{0,1\}^n$ 中均匀分布）.

(1) 需要多少个采样才能以 $1 - \delta$ 的信心和不错过 ε 的绝对误差估计 $\hat{f}(a)$. 换言之, 需要设计一个估计方法使得 $\Pr[|(\text{估计值}) - \hat{f}(a)| > \varepsilon] \leq \delta$.

(2) 问题改为估计 $\sum_{a \in \{0,1\}^n} \hat{f}(a)$, 需要多少个采样.

这里约定 $\hat{f}(a) = \mathbb{E}_x[f(x)\overline{\chi_a(x)}]$, $f(x) = \sum_a \hat{f}(a)\chi_a(x)$.

提示：可以使用上一次作业中 Chernoff bound 的变种.

解

(1) 因为 $\hat{f}(a) = \frac{1}{2^n}\sum_x f(x)(-1)^{\langle x,a \rangle} = \mathbb{E}\left[f(x)(-1)^{\langle x,a \rangle}\right]$, 即估计 2^n 个范围在 $[-1, 1]$ 的数的平均值, 有

$$\Pr\left[\left|\sum_{i=1}^{N} a_i - E\right| \geq \epsilon\right] \leq e^{-O\left(\epsilon^2 N\right)},$$

可得 $N = O\left(\ln \frac{1}{\delta} / \epsilon^2\right)$.

(2) $f(0) = \sum_a \hat{f}(a)$, 那么 N 次不中的概率为 $\left(1 - \frac{1}{2^n}\right)^N$, 取 $N = O\left(2^n \log \frac{1}{\delta}\right)$ 即可.

4. (10 分) 令 $X_1,\dots,X_{2n}$ i.i.d. 服从 Bern(1/2) 分布. 我们要选取一组系数 $c_1,\dots,c_{2n}\in\mathbb{Z}$, 使得 $\sum_i c_iX_i$ 接近均匀分布. 当然, 并不存在 $\mathbb{Z}$ 上的均匀分布, 我们实际的要求是统计距离

$$\Delta\Big(\sum_i c_iX_i,\sum_i c_iX_i+1\Big)\le 2^{-\lambda}. \tag{*}$$

一种显然的做法, 是令 $n=\lambda/2$, 令 $c_i=2^{i-1}$; 这样 $\sum_i c_iX_i$ 服从 $\{0,1,\dots,2^\lambda-1\}$ 上的均匀分布, 而 $\sum_i c_iX_i+1$ 服从 $\{1,2,\dots,2^\lambda\}$ 上的均匀分布, 满足我们对统计距离的要求.

进一步, 假设 $X_1,\dots,X_n$ 中有一半值被泄漏. 要求即使已经看到泄露值, (*) 仍然成立.

为了方便分析, 我们令 $c_1,\dots,c_n$ 是 i.i.d. 从某个分布 P_C 中选取的. 这样不管哪部分值泄露, 分析都相同. 不失一般性, 可以假设前一半值没有泄露. 定义函数

$$\mathrm{Err}(c_1,\dots,c_n)=\Delta\Big(\sum_i c_iX_i,\sum_i c_iX_i+1\Big)$$

我们要求当 $c_1,\dots,c_n$ 是从 $(P_C)^n$ 选取时, 以 $1-2^{-\lambda}$ 的概率（这个随机性只依赖于 $c_1,\dots,c_n$）

$$\mathrm{Err}(c_1,\dots,c_n)\le 2^{-\lambda}.$$

请根据 λ, 选取合适的 n 以及分布 P_C, 使得要求被满足. 请让 n 的取值尽量小, 可以忽略常数系数. 建议选取 P_C 为 $\{1,2,3,\dots,B\}$ 上的均匀分布, 其中 $B=2^{O(\lambda)}$ 根据 λ 选取.

解　在本答案中, 我们定义傅立叶变换和逆傅立叶变换为,

$$\hat f(y)=\sum_{x\in\mathbb{Z}_N} f(x)e^{-2\pi i\frac{xy}{N}},\qquad f(x)=\frac{1}{N}\sum_{y\in\mathbb{Z}_N}\hat f(y)e^{2\pi i\frac{xy}{N}}.$$

为了能使用离散傅立叶分析, 我们选取一个足够大的 $N>nB+1$. 这样, 无论 c_i,X_i 如何取值, 总有 $\sum_i c_iX_i,\sum_i c_iX_i+1<N$. 因此可以把它们看成 $\mathbb{Z}_N$ 上的分布.

用 $\sigma_c:\mathbb{Z}_N\to\mathbb{R}$ 表示 $\{c\}$ 上的退化分布, $\tau_c=\frac12(\sigma_0+\sigma_c)$ 表示 $\{0,c\}$ 的均匀分布, 也就是

$$\sigma_c(x)=\begin{cases}1, & \text{if } x=c\\ 0, & \text{otherwise}\end{cases}\qquad \tau_c(x)=\begin{cases}\frac12, & \text{if } x\in\{0,c\}\\ 0, & \text{otherwise}\end{cases}$$

那么 $\sum_i c_iX_i$ 的分布就是 $\tau_{c_1}*\tau_{c_2}*\cdots*\tau_{c_n}$, 这里 $*$ 表示卷积. 类似地, $\sum_i c_iX_i+1$ 的分布是 $\tau_{c_1}*\tau_{c_2}*\cdots*\tau_{c_n}*\sigma_1$. 这两个分布间的统计距离就是

$$\Delta\Big(\sum_i c_iX_i,\sum_i c_iX_i+1\Big)=\frac12\big\|\tau_{c_1}*\tau_{c_2}*\cdots*\tau_{c_n}-\tau_{c_1}*\tau_{c_2}*\cdots*\tau_{c_n}*\sigma_1\big\|_1=\frac12\big\|\tau_{c_1}*\tau_{c_2}*\cdots*\tau_{c_n}*(\sigma_0-\sigma_1)\big\|_1$$

为了估计 L1 距离的上界, 只需对 L2 距离有足够紧的估计. 而计算 L2 距离可以利用傅立叶系数. 令 $f=\tau_{c_1}*\tau_{c_2}*\cdots*\tau_{c_n}*(\sigma_0-\sigma_1)$, 那么 $\hat f=\widehat{\tau_{c_1}}\cdot\widehat{\tau_{c_2}}\cdot\ldots\widehat{\tau_{c_n}}\cdot\widehat{\sigma_0-\sigma_1}$. 对每个 $a\in\mathbb{Z}_N$,

- $|\hat f(a)|\le|\widehat{\sigma_0-\sigma_1}(a)|\le\frac{2\pi}{N}\cdot|a|$, 其中 $|a|:=\min\{a,N-a\}$. 当 $|a|$ 较小时, 这是一个比较好的估计.

- 当 $|a|$ 较大时, 使用另一个估计. 注意到

  $$\widehat{\tau_c}(a) = \frac{1}{2}(1 + e^{-2\pi i \frac{ac}{N}})$$

  而 c 在 $\{1, \ldots, B\}$ 中均匀取值. 这样当 $|a|$ 较大时, $\widehat{\tau_c}(a)$ 在复平面上, 以 $1/2$ 为中心, 以 $1/2$ 为半径的圆上"均匀"分布. 因此以 $\Omega(1)$ 概率, $\widehat{\tau_c}(a)$ 的绝对值小于 $1 - \Omega(1)$.

  具体来说,

  $$\begin{aligned}
  \mathop{\mathbb{E}}_{c \leftarrow \{1,\ldots,B\}}\left[|\widehat{\tau_c}(a)|^2\right] = \frac{1}{B}\sum_{c=1}^{B} |\widehat{\tau_c}(a)|^2 &= \frac{1}{B}\sum_{c=1}^{B} \frac{1}{4}(1 + e^{-2\pi i \frac{ac}{N}})(1 + e^{2\pi i \frac{ac}{N}}) \\
  &= \frac{1}{B}\sum_{c=1}^{B} \frac{1}{4}(2 + e^{-2\pi i \frac{ac}{N}} + e^{2\pi i \frac{ac}{N}}) \\
  &= \frac{1}{2} + \frac{1}{B}\frac{1}{4}\Big(\frac{1 - e^{-2\pi i \frac{aB}{N}}}{e^{2\pi i \frac{a}{N}} - 1} + \frac{1 - e^{2\pi i \frac{aB}{N}}}{e^{-2\pi i \frac{a}{N}} - 1}\Big) \\
  &\leq \frac{1}{2} + \frac{1}{B}\frac{1}{|e^{2\pi i \frac{a}{N}} - 1|} \\
  &\leq \frac{1}{2} + \frac{1}{B}\frac{1}{2 \cdot |\sin(\pi \frac{a}{N})|}
  \end{aligned}$$

  只要 $|a| \geq N/B$ 且 $B \geq 2$

  $$\mathop{\mathbb{E}}_{c \leftarrow \{1,\ldots,B\}}\left[|\widehat{\tau_c}(a)|^2\right] \leq \frac{1}{2} + \frac{1}{B}\frac{1}{2 \cdot |\sin(\pi \frac{a}{N})|} \leq \frac{1}{2} + \frac{1}{B}\frac{1}{2\sin(\pi/B)} \leq \frac{3}{4}.$$

  根据 Chernoff bound,

  $$\Pr\left[\frac{1}{n}\sum_{i=1}^{n} |\widehat{\tau_{c_i}}(a)|^2 \geq 7/8\right] \leq e^{-n/64}.$$

  利用算术平均和几何平均的关系,

  $$\Pr\left[\prod_{i=1}^{n} |\widehat{\tau_{c_i}}(a)|^2 \geq (7/8)^n\right] \leq e^{-n/64}.$$

  这样以 $1 - N \cdot e^{-n/64}$ 的概率, 对所有 $a \in [N/B, N - N/B]$, 均有

  $$|\hat{f}(a)| = |\widehat{\sigma_0 - \sigma_1}(a)| \cdot \prod_{i=1}^{n} |\widehat{\tau_{c_i}}(a)| \leq \prod_{i=1}^{n} |\widehat{\tau_{c_i}}(a)| \leq (7/8)^{n/2}.$$

综合两种情况, 以 $1 - N \cdot e^{-n/64}$ 的概率

$$\begin{aligned}
\|f\|_2^2 = \sum_x f^2(x) = \frac{1}{N}\sum_a \hat{f}^2(a) &= \frac{1}{N}\sum_{a:|a|<N/B} \hat{f}^2(a) + \frac{1}{N}\sum_{a:|a|\geq N/B} \hat{f}^2(a) \\
&\leq \frac{1}{N}\sum_{a:|a|<N/B} \Big(\frac{2\pi}{N} \cdot |a|\Big)^2 + \frac{1}{N}\sum_{a:|a|\geq N/B} (7/8)^n \\
&\leq \frac{8\pi^2}{B^3} + (7/8)^n.
\end{aligned}$$

根据 Cauchy-Schwarz 不等式, 此时

$$\mathrm{Err}(c_1, \ldots, c_n) = \frac{1}{2}\|f\|_1 \leq \frac{1}{2}\sqrt{N} \cdot \|f\|_2 \leq \frac{1}{2}\sqrt{N\Big(\frac{8\pi^2}{B^3} + (7/8)^n\Big)}$$

为了达到题设的要求, 只需要 n, N, B 满足

$$1 - N \cdot e^{-n/64} \geq 1 - 2^{-\lambda}, \qquad \frac{1}{2}\sqrt{N\left(\frac{8\pi^2}{B^3} + (7/8)^n\right)} \leq 2^{-\lambda}, \qquad N > nB + 1, \qquad B \geq 2.$$

不难验证, 存在 $n = O(\lambda), B = O(2^\lambda \lambda), N = nB + 1$ 使得需要的条件均被满足.

5. (6 分) 考虑一个有限状态空间 Ω 上不可约的（irreducible）马尔可夫核 P. 我们知道 P 存在唯一的稳态分布（stationary distribution）π. 证明对于任意初始分布 μ

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \mu P^j = \pi.$$

解 定义另一个马尔可夫核 $\hat{P} = \frac{1}{2}(P + I)$. 可以把 $\hat{P}$ 理解为, 以 1/2 概率按 P 转移, 以 1/2 概率保持在当前状态. $\pi\hat{P} = \frac{1}{2}(\pi P + \pi) = \pi$, 因此 π 也是 $\hat{P}$ 的稳态分布.

对任意状态 $x \in \Omega$, 都有 $\hat{P}(x|x) \geq 1/2 > 0$, 这样的马尔可夫链被称作 lazy. 一个 lazy 的马尔可夫链一定也无周期. 不可约且无周期的马尔可夫链一定收敛. 因此对任意初始分布 μ

$$\lim_{n\to\infty} \mu\hat{P}^n = \pi,$$

进而有

$$\lim_{n\to\infty} \frac{1}{2n} \sum_{j=0}^{2n-1} \mu\hat{P}^n = \pi.$$

为了证明题设, 我们只需证明

$$\lim_{n\to\infty} \left(\frac{1}{2n}\sum_{j=0}^{2n-1} \mu\hat{P}^j - \frac{1}{n}\sum_{j=0}^{n-1} \mu P^j\right) = 0. \tag{*}$$

(这两者的相似有直观的解释：一边是 $\hat{P}$ 对应的马尔可夫链在前 $2n$ 时刻的平均分布, 一边是 P 对应的马尔可夫链在前 n 时刻的平均分布. 而懒惰的 $\hat{P}$ 大致就是时间放慢两倍的 P.)

$$\frac{1}{2n}\sum_{j=0}^{2n-1} \mu\hat{P}^n = \frac{1}{2n}\sum_{j=0}^{2n-1} \mu\left(\frac{1}{2}P + \frac{1}{2}I\right)^j = \frac{1}{2n}\sum_{i=0}^{2n-1} \mu P^i \sum_{j=i}^{2n-1} \binom{j}{i}\frac{1}{2^j} = \frac{1}{n}\sum_{i=0}^{2n-1} \mu P^i c_i$$

其中 $c_i = \sum_{j=i}^{2n-1} \binom{j}{i}\frac{1}{2^{j+1}}$. 为了证明 (*), 我们希望对大部分 $i < n$, $c_i \approx 1$, 对大部分 $i > n$, $c_i \approx 0$.

我们注意到 c_i 有其它含义. 考虑一列独立的随机变量 $Z_1, Z_2, \cdots \sim \mathrm{Bern}(1/2)$. 注意到

$$\binom{j}{i}\frac{1}{2^{j+1}} = \Pr\Big[\sum_{k\leq j} Z_k = i \wedge Z_{j+1} = 1\Big] = \Pr\Big[j \text{ 是满足 } \sum_{k\leq j+1} Z_k > i \text{ 的最小整数}\Big]$$

$$c_i = \sum_{j=i}^{2n-1}\binom{j}{i}\frac{1}{2^{j+1}} = \Pr\Big[\text{满足 } \sum_{k\leq j+1} Z_k > i \text{ 的最小整数} < 2n\Big] = \Pr\Big[\sum_{j=1}^{2n-1} Z_j > i\Big]$$

(这个关系并不是巧合. Z_i 其实就表示了 $\hat{P}$ 在第 i 时刻是否懒惰. 具体来说, 不妨令 $X_0, X_1, X_2, \ldots$ 表示 P 的马尔可夫链, 不难验证 $\hat{X}_i = X_{\sum_{j\leq i} Z_i}$ 是 $\hat{P}$ 对应的马尔可夫链.)

我们可以用 Chernoff bound 估计 c_i. 令 $\varepsilon = n^{-1/3}$, 那么由 Chernoff bound

$$\begin{aligned} c_i &\geq 1 - O(n^{-1/3}) = 1 - o(1) && \text{if } i \leq n(1-\varepsilon) \\ c_i &\leq O(n^{-1/3}) = o(1) && \text{if } i \geq n(1+\varepsilon) \end{aligned}$$

据此可以证明 (*)

$$\begin{aligned} \frac{1}{2n}\sum_{j=0}^{2n-1} \mu \hat{P}^n &= \frac{1}{n}\sum_{i=0}^{2n-1} \mu P^i c_i \\ &= \frac{1}{n}\sum_{i=0}^{n-\varepsilon n} \mu P^i c_i + \frac{1}{n}\sum_{i=n-\varepsilon n+1}^{n+\varepsilon n-1} \mu P^i c_i + \frac{1}{n}\sum_{i=n+\varepsilon n}^{2n-1} \mu P^i c_i \\ &= \frac{1}{n}\sum_{i=0}^{n-\varepsilon n} \mu P^i (1 - O(n^{-1/3})) + O(n^{-1/3}) + \frac{1}{n}\sum_{i=n+\varepsilon n}^{2n-1} \mu P^i \cdot O(n^{-1/3}) \\ &= \frac{1}{n}\sum_{i=0}^{n-\varepsilon n} \mu P^i + O(n^{-1/3}) \\ &= \frac{1}{n}\sum_{i=0}^{n-1} \mu P^i + O(n^{-1/3}) \end{aligned}$$

另一种解法: 参考下一题中的分解. 我们按 $j \bmod T$ 把求和分成 T 类, 证明每一类都收敛, 从而整个求和收敛（只要收敛就一定收敛到 π）. 每一类中的 Markov 核相当于 pass 到了 P^T, 它在每个 Ω_j 中都是不可约且无周期的 Markov 链, 从而收敛.

6. (4 分) 考虑大小为 n 的有限状态空间 Ω 上的一个不可约（irreducible）马尔可夫核 P, 令 π 是稳态分布.
   (1) 证明 P 只有一个特征值等于 1.
   (2) 假设 P 有周期 $T > 1$. 这里 $T = \gcd\{t : \exists x, P^t(x|x) > 0\}$. 不难证明, 周期性说明状态空间可以划分为 T 个非空子集 $\Omega_0, \Omega_1, \ldots, \Omega_{T-1}$ 满足

$$\forall x, y \in \Omega,\ \forall j \in \mathbb{Z}_T,\ P(y|x) > 0 \wedge x \in \Omega_j \implies y \in \Omega_{j+1}.$$

   证明 1 的所有 T 次根 $e^{2\pi i \frac{k}{T}}$（for $k \in \mathbb{Z}_T$）都是 P 的特征值.

解

(1) 课上已经证明 $\dim(P - I) = n - 1$, 所以特征值为 1 的特征向量只有一维. 我们知道稳态分布 μ 就是这个特征向量.

如果要说明只有一个特征值是 1, 也就是要说明 1 的代数重数等于 1. 反证, 假设 1 的代数重数大于 1. 那么 P 的 Jordan 标准型一定包括一个形如

$$\begin{bmatrix} 1 & 1 & & \\ & 1 & \ddots & \\ & & \ddots & 1 \\ & & & 1 \end{bmatrix}$$

的子块. 因此, Jordan 分解中的基中有一个非零向量 ν 满足

$$\begin{bmatrix} -\nu- \\ -\mu- \end{bmatrix} P = \begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix} \begin{bmatrix} -\nu- \\ -\mu- \end{bmatrix}$$

只要选一个足够小的 ε, 就可以使得 $\mu+\varepsilon\nu$ 每位非负, 于是可以正则化为一个概率分布 $\frac{\mu+\varepsilon\nu}{\|\mu+\varepsilon\nu\|_1}$. 于是

$$\frac{\mu+\varepsilon\nu}{\|\mu+\varepsilon\nu\|_1}P^t = \frac{(1+t\varepsilon)\mu+\varepsilon\nu}{\|\mu+\varepsilon\nu\|_1}$$

也应该是一个分布. 但当 t 充分大时, 它不是分布.

(2) 对任意 $k \in \mathbb{Z}_T$, 构造 π_k 为

$$\pi_k(x) = e^{-2\pi i\cdot jk/T}\cdot \pi(x) \text{ if } x \in \Omega_j.$$

那么对任意 $j \in \mathbb{Z}_T$ 和 $y \in \Omega_j$

$$\begin{aligned}(\pi_k P)(y) = \sum_{x\in\Omega_{j-1}} \pi_k(x)P(y|x) = e^{-2\pi i\cdot(j-1)k/T}\cdot \sum_{x\in\Omega_{j-1}} \pi(x)P(y|x) \\ = e^{-2\pi i\cdot(j-1)k/T}\cdot\pi(y) = e^{2\pi i\cdot k/T}\cdot\pi_k(y).\end{aligned}$$

所以 $e^{2\pi i\cdot k/T}$ 是一个特征值.

另一种解法: 对 $k \in \mathbb{Z}_T$, 我们构造 (列) 特征向量 v, 对 $j \in \mathbb{Z}_T, x \in \Omega_j$, v_x 恒为 $e^{2\pi i\cdot jk/T}$. 即

$$v = \sum_j e^{2\pi i\cdot jk/T}\mathbb{1}_{\Omega_j}$$

由于 $\sum_{y\in\Omega_{j+1}} P(y|x) = 1, \forall x \in \Omega_j$, 我们有 $P\mathbb{1}_{\Omega_{j+1}} = \mathbb{1}_{\Omega_j}$, 从而 $Pv = \sum_j e^{2\pi i\cdot jk/T}P\mathbb{1}_{\Omega_j} = \sum_j e^{2\pi i\cdot jk/T}\mathbb{1}_{\Omega_{j-1}} = \sum_j e^{2\pi i\cdot(j+1)k/T}\mathbb{1}_{\Omega_j} = e^{2\pi i\cdot k/T}v$, 从而 v 是一个特征值为 $e^{2\pi i\cdot k/T}$ 的特征向量.