马尔可夫链

参考答案

1. (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.)

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

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 $Z_1,Z_2,\dots\sim\mathrm{Bern}(1/2)$. 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 $X_0,X_1,X_2,\dots$ 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

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

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.

2. (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.

3. (6 分) 考虑 $\mathbb{Z}$ 上的随机游走. 马尔可夫核是

$$P(x+1|x) = p, \quad P(x-1|x) = 1-p$$

其中 $p \in (0,1)$ 是参数. 请计算这个马尔可夫链返回初始点的概率.

$$\Pr[\exists i > 0 \text{ such that } X_i = X_0].$$

解 不失一般性, 可以假定 $X_0 = 0$, 然后考虑返回 0 点的概率

$$\Pr_0[\exists i > 0 \text{ such that } X_i = 0].$$

定义 $v(j)$ 为如果马尔可夫链从状态 j 出发, 经过 0 点的概率

$$v(j) := \Pr_j[\exists i \geq 0 \text{ such that } X_i = 0].$$

这里在 Pr 的下标规定 X_0 的（退化）分布.

我们关心 v 的值. 显然 $v(0)=1$. 对 $j\neq 0$,

$$v(j)=\Pr_j[\exists i\geq 0, X_i=0]=p\Pr_j[\exists i\geq 0, X_i=0|X_1=j+1]$$
$$+(1-p)\Pr_j[\exists i\geq 0, X_i=0|X_1=j-1]=p\,v(j+1)+(1-p)v(j-1).$$

对于 $j\geq 0$, 数列 $v(0),v(1),v(2),\ldots$ 有简单的线性递推公式, 它们一定满足

$$v(j)=\begin{cases}\alpha j+\beta, & \text{if } p=\frac{1}{2}\\ \alpha(\frac{1-p}{p})^j+\beta, & \text{if } p\neq\frac{1}{2} \text{ 其中 } \frac{1-p}{p} \text{ 是特征方程 } 1=px+(1-p)/x \text{ 除了 1 之外的另一个根}\end{cases}$$

其中 α,β 是待定常数. 这时分三种情况考虑,

- 如果 $p=\frac{1}{2}$: 因为 $v(0)=1$ 且 $\forall j, v(j)\in[0,1]$, 只能是 $v(j)=1$.
- 如果 $p<\frac{1}{2}$: 这时 $\frac{1-p}{p}>1$. 因为 $v(0)=1$ 且 $\forall j, v(j)\in[0,1]$, 只能是 $v(j)=1$.
- 如果 $p>\frac{1}{2}$: 这时 $\frac{1-p}{p}<1$. 根据 $v(0)=1$ 且 $\forall j, v(j)\in[0,1]$ 只能推出 $v(j)=\alpha(\frac{1-p}{p})^j+(1-\alpha)$ 其中 $\alpha\in[0,1]$. 因此我们要额外证明 $\lim_{j\to+\infty}v(j)=0$, 这样可以说明 $v(j)=(\frac{1-p}{p})^j$.

  $$v(j)=\Pr_j[\exists n\geq 0, X_n=0]\leq\sum_{n\geq 0}\Pr_j[X_n=0]\leq\sum_{n\geq j}\Pr_j[X_n\leq j]$$
  $$\leq\sum_{n\geq j}\exp\Big(-n\cdot D\big(\frac{1}{2}\parallel p\big)\Big)=e^{-\Theta(j)}.$$

  其中一步使用了 Chernoff bound.

综合三种情况, $p\,v(1)=p\min(1,\frac{1-p}{p})=\min(p,1-p)$. 对称地, $(1-p)\,v(-1)=\min(p,1-p)$. 所以

$$\Pr[\exists i>0, X_i=X_0]=p\,v(1)+(1-p)v(-1)=2\min(p,1-p).$$

另一种解法: 由第 4 题中的引理 1, 我们只需要计算 $\mathbb{E}[N]=\sum_{n\geq 1}\Pr[X_{2n}=X_0]$（$N$ 和 E 的定义也请参照那里）. $\Pr[X_{2n}=X_0]=\binom{2n}{n}(p(1-p))^n$. 由生成函数 $\sum_{n\geq 0}\binom{2n}{n}x^n=\frac{1}{\sqrt{1-4x}}$ 代入 $x=p(1-p)$, 可以得知 $\mathbb{E}[N]=\frac{1}{\sqrt{1-4p(1-p)}}-1=\frac{1}{|1-2p|}-1$, 从而 $\Pr[E]=1-\frac{1}{1+\mathbb{E}[N]}=1-|1-2p|$.

4. (6 分) 考虑 $\mathbb{Z}^d$ 上的随机游走. 马尔可夫核是

$$P(y_1,\ldots,y_d|x_1,\ldots,x_d)=\begin{cases}1/3^d, & \text{if } \forall i, |y_i-x_i|\leq 1\\ 0, & \text{otherwise}\end{cases}$$

这个马尔可夫核在各个维度上独立, 便于分析. 证明

$$\Pr[\exists i>0 \text{ such that } X_i=X_0]=\begin{cases}1, & \text{if } d=2\\ 1-\Omega(1), & \text{if } d>2\end{cases}$$

提示：考虑

$$\mathbb{E}[\text{number of } i > 0 \text{ such that } X_i = X_0].$$

解 定义随机变量 N 和事件 E 为

$$N := \text{number of } i > 0 \text{ such that } X_i = X_0,$$
$$E := \{\exists i > 0 \text{ such that } X_i = X_0\}.$$

引理 1. $\Pr[E] = 1$ 当且仅当 $\mathbb{E}[N] < \infty$. 当 $\Pr[E] < 1$ 时, 有 $\mathbb{E}[N] = \frac{\Pr[E]}{1-\Pr[E]}$, 或言 $\Pr[E] = \frac{\mathbb{E}[N]}{1+\mathbb{E}[N]}$.

证明. 由于 $N \geq 0$, $\mathbb{E}[N]$ 一定存在（可能为 ∞）. 如果 E 未发生, 那么 $N = 0$, 因此 $\mathbb{E}[N] = \mathbb{E}[N|E]\Pr[E]$. 现在 condition on E 发生, 我们记 i_0 是最小的 $i > 0$ such that $X_i = X_0$, 此时 $N = 1 + (\text{number of } i > i_0 \text{ such that } X_i = X_0)$. 由 Markov 链的无记忆性, $\mathbb{E}[\text{number of } i > i_0 \text{ such that } X_i = X_0|E] = \mathbb{E}[N]$. 由于 $N \geq 0$ 以及期望线性性, 我们有 $\mathbb{E}[N] = \Pr[E](1 + \mathbb{E}[N])$. 如果 $\Pr[E] = 1$, $\mathbb{E}[N]$ 只能为 ∞, 否则 $\mathbb{E}[N] = \frac{\Pr[E]}{1-\Pr[E]}$ 是有限数. $\square$

由期望线性性, 我们有 $\mathbb{E}[N] = \sum_{n\geq 1}\Pr[X_n = X_0]$. 设 p_n 表示 $d = 1$ 时的随机游走满足 $X_n = X_0$ 的概率, 由于每一维独立, 有 $\Pr[X_n = X_0] = p_n^d$. 在剩下的答案中, 我们要证明 $p_n = \Theta(n^{-1/2})$. 这说明, 当 $d \leq 2$ 时, $\mathbb{E}[N] = \infty$; 当 $d \geq 3$ 时, $\mathbb{E}[N] < \infty$. 结合引理 1, 便得到题目要求的结论.

估计 p_n 有多种办法, 下面我们展示一种完全初等的办法. 不失一般性, 可以假定 $X_0 = 0$. 注意到给定的马尔可夫核可以直观地理解为如下两步过程: 首先抛一个概率为 2/3 的硬币, 根据硬币结果决定是否留在当前状态. 如果不留在当前状态, 那么以 50-50 的概率 +1 或 −1. 具体来说, 定义随机变量 $Z_i = \mathbb{1}[X_i \neq X_{i-1}]$. 那么 $Z_i \sim \mathrm{Bern}(2/3)$. condition on $Z_1 + \cdots + Z_n = m$, X_n 等于 m 个独立随机的 ± 1 的和, 因此 $\Pr[X_n = 0|Z_1 + \cdots + Z_n = m] = \binom{m}{m/2}/2^m$（如果 m 是偶数）.

$$\begin{aligned}\Pr[X_n = 0] &= \sum_m \Pr[X_n = 0|Z_1 + \cdots + Z_n = m]\Pr[Z_1 + \cdots + Z_n = m] \\ &= \sum_{\text{even } m} \frac{\binom{m}{m/2}}{2^m}\Pr[Z_1 + \cdots + Z_n = m] = \sum_{\text{even } m > n/2} \frac{\binom{m}{m/2}}{2^m}\Pr[Z_1 + \cdots + Z_n = m] + 2^{-\Omega(n)}. \quad (*)\end{aligned}$$

最后一步利用 Chernoff bound, $m \leq n/2$ 的可能性可以忽略.

引理 2. 对 $k \geq 1$, 我们有

$$\binom{2k}{k} \Big/ 2^{2k} = \Theta\left(\frac{1}{\sqrt{k}}\right).$$

证明. 记 $a_k = \binom{2k}{k}2^{-2k}$, 展开得 $a_k = (1 - \frac{1}{2k})a_{k-1}$.

$$\ln a_k = \sum_{i=1}^{k}\log\Big(1 - \frac{1}{2k}\Big) \leq -\sum_{i=1}^{k}\frac{1}{2k} \leq -\frac{1}{2}\log k + \text{常数}.$$

因而 $a_k = O(k^{-1/2})$. 下界类似, 只需使用 $\ln(1-x) \geq -x - x^2$ 对 $x \in [0, 1/2]$ 成立.

下界也可以用二阶矩估计. 注意到 $a_k = \mathrm{Binom}(2k, \frac{1}{2})(k)$. 二项分布 $\mathrm{Binom}(2k, \frac{1}{2})$ 的期望是 k, 方差是 $k/2$. 根据 Chebyshev 不等式, 以至少 $1/2$ 的概率落在 $(k-\sqrt{k}, k+\sqrt{k})$ 内. 我们又知道二项分布中, 期望的概率最高. 所以 $a_k = \mathrm{Binom}(2k, \frac{1}{2})(k) \geq \frac{1/2}{2\sqrt{k}} = \Omega(k^{-1/2})$. □

引理 2 实际在说, 存在常数 $C > c > 0$, 使得对任意 $k > 1$,

$$\binom{2k}{k} \Big/ 2^{2k} \in \left[\frac{c}{\sqrt{k}}, \frac{C}{\sqrt{k}}\right].$$

回到 (*), 我们可以得到

$$\Pr[X_n = 0] \leq \max_{\text{even } m \in (n/2, n]} \frac{\binom{m}{m/2}}{2^m} + 2^{-\Omega(n)} \leq \frac{C}{\sqrt{n/2}} + 2^{-\Omega(n)} = O\Big(\frac{1}{\sqrt{n}}\Big),$$
$$\Pr[X_n = 0] \geq \max_{\text{even } m \in [0, n]} \frac{\binom{m}{m/2}}{2^m} \cdot \Pr[Z_1 + \cdots + Z_n \text{ is even}] \geq \frac{c}{\sqrt{n}} \cdot \frac{1}{3} = \Omega\Big(\frac{1}{\sqrt{n}}\Big).$$

5. (5 分) 对一个马尔可夫链 P, 用 π 表示它的一个稳态分布, 用 $\tau(\varepsilon)$ 表示它的混合时间.

$$\tau(\varepsilon) = \text{smallest } t \text{ s.t. } d(t) \leq \varepsilon$$
$$d(t) = \max_x \Delta_{\mathrm{TV}}(P^t(x, \cdot), \pi)$$

证明, 对任意 $\varepsilon > 0$, $\tau(2\varepsilon^2) \leq 2\tau(\varepsilon)$.

解 固定 ε, 记 $\tau = \tau(\varepsilon)$.

用 coupling 方法证明题目. 对任意 x_0, 构造 $X_0, X_\tau, X_{2\tau}, Y_0, Y_\tau, Y_{2\tau}$, 使得 $X_0 = x_0$, $Y_0 \sim \pi$, 两边的边缘分布都是马尔可夫链 (即 $\Pr[Y_\tau = y_\tau, Y_{2\tau} = y_{2\tau} | Y_0 = y_0] = P^\tau(y_0, y_\tau) P^\tau(y_\tau, y_{2\tau})$, 对 X_t 类似), 并且 $\Pr[X_{2\tau} \neq Y_{2\tau}] \leq 2\varepsilon^2$.

因为 $\Delta_{\mathrm{TV}}(X_\tau, Y_\tau) = \Delta_{\mathrm{TV}}(P^t(x_0, \cdot), \pi) \leq \varepsilon$, 所以通过 coupling 可以令 $\Pr[X_\tau = Y_\tau] \geq 1 - \varepsilon$.

如果 $X_\tau = Y_\tau$, 可以让两个马尔可夫链保持相同, $\Pr[X_{2\tau} = Y_{2\tau} \mid X_\tau = Y_\tau] = 1$.

如果 $X_\tau = x \neq y = Y_\tau$, 这时 $X_{2\tau}, Y_{2\tau}$ 的条件分布分别是 $P^t(x, \cdot), P^t(x, \cdot)$. 因为

$$\Delta_{\mathrm{TV}}(P^t(x, \cdot), P^t(y, \cdot)) \leq \Delta_{\mathrm{TV}}(P^t(x, \cdot), \pi) + \Delta_{\mathrm{TV}}(P^t(y, \cdot), \pi) \leq 2\varepsilon,$$

可以通过 coupling 令 $\Pr[X_{2\tau} = Y_{2\tau} \mid X_\tau = x, Y_\tau = y] \geq 1 - 2\varepsilon$.

于是

$$\Pr[X_{2\tau} \neq Y_{2\tau}] = \Pr[X_{2\tau} \neq Y_{2\tau} \mid X_\tau \neq Y_\tau] \Pr[X_\tau \neq Y_\tau] \leq 2\varepsilon^2.$$

6. (10 分) 简单图 G 中有 n 个点, 最大度数记为 Δ. 用 $C > 5\Delta$ 种颜色对 G 随机点染色, 要求任意一对相邻点的染色不同. 为了均匀采样一个随机染色, 我们使用 MCMC 方法. 马尔可夫核是:

- 假设当前染色为 $f : V \to C$.
- 随机选取一个点 $v \in C$, 随机选取一个颜色 $c \in C$.

- 如果 v 的邻居的颜色都不是 c, 就将 v 的染色修改为 c; 否则保持染色不变.

请估算混合时间 $\tau(\varepsilon)$, 给出一个尽量好的上界.

$$\tau(\varepsilon) = \text{smallest } t \text{ s.t. } d(t) \le \varepsilon$$
$$d(t) = \max_{\mu} \Delta_{\mathrm{TV}}(\mu P^t, \pi)$$

注. *如果想用 coupling 分析 $C > 2\Delta$ 的情形, 建议用 $S_t \subseteq V$ 表示 t 时刻 coupling 中两个染色一致的点集. 考虑被 S_t 切的边 (一个端点在 S_t 中, 另一个端点在 S_t 外) 有怎样的影响.*

注. *马尔可夫链可以优化为"随机选取一个 v 的邻居中未出现的颜色 $c \in C$".*

解 使用 coupling 分析, 定义马尔可夫链 $\{(X_t, Y_t)\}_{t\ge 0}$. 马尔可夫核是:

- 当前染色为 X_{t-1}, Y_{t-1}.
- 随机选取一个点 $v \in V$. 在 v 以外的部分, 令 X_t 与 X_{t-1} 相同, 令 Y_t 与 Y_{t-1} 相同.

  用 $\mathcal{X} = \{X_{t-1}(u) \mid \{u, v\} \in E\}$ 和 $\mathcal{Y} = \{Y_{t-1}(u) \mid \{u, v\} \in E\}$ 分别表示 v 的邻居的颜色集合.
- 如果 $X_{t-1}(v) \ne Y_{t-1}(v)$, 随机选取一个颜色 $c \in C$. 如果 $c \notin \mathcal{X}$, 就令 $X_t(v) = c$; 否则保持染色不变. 如果 $c \notin \mathcal{Y}$, 就令 $Y_t(v) = c$; 否则保持染色不变.

  这样的话, $X_t(v) = Y_t(v)$ 的概率至少是 $\dfrac{C - |\mathcal{X} \cup \mathcal{Y}|}{C}$.
- 如果 $X_{t-1}(v) = Y_{t-1}(v)$. 从一个特定的分布采样 (c, c'). 边缘分布是均匀分布. 如果 $c \notin \mathcal{X}$, 就令 $X_t(v) = c$; 否则保持染色不变. 如果 $c' \notin \mathcal{X}$, 就令 $Y_t(v) = c'$; 否则保持染色不变.

  为了让 $X_t(v) = Y_t(v)$ 的概率尽量大, 需要适当地设置 (c, c') 的分布. 不难做到,

  $$\Pr[c = c' \land c \notin \mathcal{X} \cup \mathcal{Y}] = \frac{C - |\mathcal{X} \cup \mathcal{Y}|}{C}, \quad \Pr[c \in \mathcal{X} \land c' \in \mathcal{Y}] = \frac{\min(|\mathcal{X}|, |\mathcal{Y}|)}{C}.$$

  这样的话, $X_t(v) = Y_t(v)$ 的概率是 $\dfrac{C - |\mathcal{X} \cup \mathcal{Y}| + \min(|\mathcal{X}|, |\mathcal{Y}|)}{C} = \dfrac{C - \min(|\mathcal{X} \setminus \mathcal{Y}|, |\mathcal{Y} \setminus \mathcal{X}|)}{C}$.

根据提示, 定义 $S_t = \{v | X_t(v) = Y_t(v)\}$. 定义 $N_t = |S_t|$.

考虑 X_{t-1}, Y_{t-1} 到 X_t, Y_t 的马尔可夫核的采样过程.

- 如果 $X_{t-1}(v) \ne Y_{t-1}(v)$ (即 $v \notin S_{t-1}$), 只要采样得到的颜色 c 不在 $\mathcal{X} \cup \mathcal{Y}$ 中, 那么 v 在两边的颜色都会被更新到 c. 用 $E(v, S)$ 表示 v 和 S 之间的边数. 那么

  $$|\mathcal{X} \cup \mathcal{Y}| \le E(v, S_{t-1}) + 2E(v, V \setminus S_{t-1}) \le 2\Delta - E(v, S_{t-1}).$$

  于是 $X_t(v) = Y_t(v)$ 的概率至少是 $\frac{C - 2\Delta + E(v, S_{t-1})}{C}$.
- 如果 $X_{t-1}(v) = Y_{t-1}(v)$ (即 $v \in S_{t-1}$), $X_t(v) \ne Y_t(v)$ 的概率至多是

  $$\frac{\min(|\mathcal{X} \setminus \mathcal{Y}|, |\mathcal{Y} \setminus \mathcal{X}|)}{C} \le \frac{E(v, V \setminus S_{t-1})}{C}.$$

观察上面的概率. 当 $v \notin S_{t-1}$ 时, v 与 S_{t-1} 的连边 "有益"; 当 $v \in S_{t-1}$ 时, v 与 $V \setminus S_{t-1}$ 的连边 "有害". 因为

$$\sum_{v \notin S_{t-1}} E(v, S_{t-1}) = E(V \setminus S_{t-1}, S_{t-1}) = \sum_{v \in S_{t-1}} E(v, V \setminus S_{t-1}),$$

这两种作用恰好可以相互抵消.

具体来说, 给定 X_{t-1}, Y_{t-1} 时,

$$\begin{aligned}
\mathbb{E}\left[N_t - N_{t-1} \,\middle|\, X_{t-1}, Y_{t-1}\right] &\geq \sum_{v \notin S_{t-1}} \frac{1}{n} \frac{C - 2\Delta + E(v, S_{t-1})}{C} - \sum_{v \in S_t} \frac{1}{n} \frac{E(v, V \setminus S_{t-1})}{C} \\
&= \sum_{v \notin S_{t-1}} \frac{1}{n} \frac{C - 2\Delta}{C} \\
&= \frac{n - N_{t-1}}{n} \frac{C - 2\Delta}{C}.
\end{aligned}$$

对两边同时求期望, 可以得到

$$\mathbb{E}[N_t] - \mathbb{E}[N_{t-1}] = \frac{n - \mathbb{E}[N_{t-1}]}{n} \frac{C - 2\Delta}{C}.$$

于是解得

$$\mathbb{E}[N_t] = n - \left(1 - \frac{C - 2\Delta}{nC}\right)^t (n - \mathbb{E}[N_0]).$$

当 $t \geq \frac{\log(\varepsilon/n)}{\log(1 - \frac{C-2\Delta}{nC})} \geq \frac{nC}{C-2\Delta} \log(n/\varepsilon)$ 时, $\mathbb{E}[N_t] \geq n - \varepsilon$, 根据 Markov bound,

$$\Pr[X_t \neq Y_t] = \Pr[N_t \neq n] = \Pr[N_t \leq n - 1] \leq \varepsilon.$$

上述分析不依赖于 X_0, Y_0 的分布. 只要令 X_0 是任意染色, Y_0 服从稳态分布 (均匀分布), 就得到

$$\tau(\varepsilon) \leq \frac{nC}{C - 2\Delta} \log(n/\varepsilon).$$