CS 225: Probability and Computing Homework 3

Zihao Ye

July 25, 2016

PROBLEM 1

(a) *Proof.* 假设 p 为 gambler 赢的概率.

可以估计出 $E[T]$ 的上界:$E[T] \le (a+b)2^{a+b}$(将 $\{X_n\}$ 每 $a+b$ 个分为一段, 用 geometry distribution 的期望计算) 有限.

那么由 *Wald's Equation*, 有

$$pb + (1-p)\cdot -a = E\left[\sum_{i=1}^{T} X_i\right] = E[T]E[X_i] = 0$$

解得 $p = \frac{a+b}{a}$.

如果要写成利用 *O.S.T* 的证明, 令 $Y_t = \sum_{i=1}^{t}(X_i - E[X_i])$, 则 Y_t 是一个 martingale w.r.t $X_0, X_1, \cdots$, 由 *O.S.T* 得 $0 = E[Y_T]$. □

(b) *Proof.* 设 X_i 为第 i 次抛硬币的随机变量, $X_i = 1$ 时为正, $X_i = -1$ 时为负, 特别的, 令 $X_0 = 0$.

定义如下 Z_t, Z_t 由 $X_1, X_2, \cdots, X_t$ 完全决定.

设 $Y_t = \sum_{i=1}^{t} X_i$ 即进行完 t 轮之后的 $\#HEAD - \#TAIL$, 令

$$Z_t = \begin{cases} (Y_t - b)(Y_t + a) - t & \forall i \in [t-1]\, Y_i \in (a, b) \\ Z_{t-1} & \exists i \in [t-1]\, Y_i \ge b \text{ or } Y_i \le -a \end{cases}$$

则可以证明 $E[Z_t \mid X_1, X_2, \cdots, X_{t-1}] = Z_{t-1}$, 因此 $Z_0, Z_1, \cdots$ 是关于 $X_0, X_1, \cdots$ 的 martingale.

由上一问可知 $E[T]$ 有限.

由于 $E[|Z_{t+1}-Z_t \mid X_0,\cdots,X_t] \le 1+(a+b)^2$ 是个常数, 由 *O.S.T* 得:

$$-E[T]=E[Z_T]=E[0]=-ab \Longrightarrow E[T]=ab$$

$\square$

PROBLEM 2

设某次 flip 的 clause 中变量为 σ_i, σ_j, 则 $\sigma_i \ne \pi_i$ or $\sigma_j \ne \pi_j$.
因此设 $D_i^{(1)}$ 为进行第 i 次 flip 时 σ 与 π 的 *Hamming distance*.

$$Pr\left[D_{i+1}^{(1)}-D_i^{(1)}=-1\right] \ge \frac{1}{2}$$

可将此过程设为在一维空间上 $[0,n]$ 区间内进行的随机游走, 且初始位置未知, 每次向 0 方向走一步的概率 $\ge \frac{1}{2}$. 停时定义为 T 时刻 σ 与某个合法的 π 重合, 若 $D_t^{(1)}=0, T\le t$, 但 $D_T^{(2)}$ 不一定 $=0$, 设此过程为 $\mathfrak{M}_1$.
考虑一个类似的问题, 在 $[0,2n]$ 区间内进行随机游走, 初始位置在 $[0,n]$ 之间与 $\mathfrak{M}_1$ 的初始位置相同, 停时为在 T 时刻到达 0 或 $2n$, 设此过程为 $\mathfrak{M}_2$.
我们考虑 coupling, 产生一系列 $[0,1]$ 中独立均匀分布的随机变量 $\{X_i\}$, 则在 $\mathfrak{M}_1$ 中:

$$D_{i+1}^{(1)}-D_i^{(1)}=\begin{cases} -1 & X_i \le Pr\left[D_{i+1}^{(1)}-D_i^{(1)}=-1\right] \\ 1 & \text{otherwise} \end{cases}$$

在 $\mathfrak{M}_2$ 中:

$$D_{i+1}^{(2)}-D_i^{(2)}=\begin{cases} -1 & \left(D_i^{(2)}\le n \text{ and } X_i \le \frac{1}{2}\right) \text{ or } \left(D_i^{(2)} > n \text{ and } X_i \ge \frac{1}{2}\right) \\ 1 & \text{otherwise} \end{cases}$$

则可以证明此时 $Pr\left[D_{i+1}^{(2)}-D_i^{(2)}=-1\right]=\frac{1}{2}$ 且两两独立.
同时有:

$$\min\left\{2n-D_i^{(2)}, D_i^{(2)}\right\} \ge D_i^{(1)}$$

设 T_2 为过程 $\mathfrak{M}_2$ 的停时, 由于 $\min\left\{2n-D_{T_2}^{(2)}, D_{T_2}^{(2)}\right\}=0\ge D_{T_2}^{(1)}$, 因此此时有 $D_{T_2}^{(1)}=0, T_1\le T_2$.
由 *PROBLEM 1* 的结论, $E[T_2]=D_0^{(1)}\left(2n-D_0^{(1)}\right)=n^2-\left(D_0^{(1)}-n\right)^2$.
由此得出 $E[T_1]\le n^2=O(n^2)$, 要使得准确率 $\ge 99\%$, 则设一个此算法的 flip 次数上限 $t=100n^2$:

$$Pr[T_1\ge t]\le \frac{E[T_1]}{t}\le 0.01$$

PROBLEM 3

固定 v, 构造一个 circuit 如下:

$$\begin{aligned} R_{u,v} &= c_{uv} \\ \forall u \in V, C_u^- &= d(u) \\ C_v^+ &= 2m \end{aligned}$$

根据基尔霍夫定律和安培定律可以写出如下等式:

$$\begin{aligned} d(u) &= \sum_{uw\in E} C_{u\to w} \\ &= \sum_{uw\in E} \frac{\varphi_{u,w}}{R_{u,w}} \\ &= \sum_{uw\in E} \frac{1}{R_{u,w}} \varphi_{u,v} - \sum_{uw\in E} \frac{1}{R_{u,w}} \varphi_{w,v} \end{aligned}$$

解得:

$$\varphi_{u,v} = \frac{1}{\sum\limits_{wu\in E} c_{wu}} + \sum_{wu\in E} P_{uw}\varphi_{wv}$$

设 $\tau_{u,v}$ 为在给定的 *Markov Chain* 中从 u 出发, 到达 v 的期望 *cost*.

$$\begin{aligned} \tau_{u,v} &= \sum_{wu\in E} P_{uw}(c_{wu} + \tau_{w,v}) \\ &= \frac{1}{\sum\limits_{wu\in E} c_{wu}} + \sum_{wu\in E} P_{uw}\tau_{wv} \end{aligned}$$

因此可得出:

$$\forall u \in V, \varphi_{u,v} = \tau_{u,v}$$

构造一个类似的电路, 固定 u, 令 $C_u^+ = 2m$, 其余边的入流不变, 类似的, 可以得出:

$$\forall v \in V, \varphi_{v,u} = \tau_{v,u}$$

将两个 flow 叠加, 得到:

$$S_{u,v} = \tau_{uv} + \tau_{vu} = 2mR(u,v)$$

PROBLEM 4

(a) 令

$$P(X,Y)=\begin{cases}\frac{1}{1+\Delta} & Y\neq X \text{ and } X \text{ is adjacent to } Y\\ 1-\frac{\deg(X)}{1+\Delta} & Y=X\end{cases}$$

由于对于每个 state:X, $P(X,X)>0$, 因此这个 random walk 是 lazy 的, 由此可以得出它是 ergodic 的, 当原图连通时, 它是 irreducible 的.

容易验证:

$$\forall X,Y\in V, \frac{1}{n}P(X,Y)=\frac{1}{n}P(Y,X)$$

满足 *detailed balance equation*, 因此是 reversible 的, 它的 *stationary distribution* 就是 *uniform distribution*.

(b) 令

$$\pi'=\min_{v\in V}\{\pi(v)\}>0$$

则可以构造如下的 *Markov Chain*:

$$P(X,Y)=\begin{cases}\frac{\pi'}{(1+\Delta)\cdot\pi(X)} & Y\neq X \text{ and } X \text{ is adjacent to } Y\\ 1-\sum\limits_{V \text{ adjacent to } X}\frac{\pi'}{(1+\Delta)\cdot\pi(V)} & Y=X\end{cases}$$

由于对于每个 state:X, $P(X,X)>0$, 因此这个 random walk 是 lazy 的, 由此可以得出它是 ergodic 的, 当原图连通时, 它是 irreducible 的.

容易验证:

$$\forall X,Y\in V, \pi(X)P(X,Y)=\pi(Y)P(Y,X)$$

满足 *detailed balance equation*, 因此是 reversible 的, 它的 *stationary distribution* 就是 π.

PROBLEM 5

(a) 原来的转移规则等价于, 随机选取一位, 以 p 的概率修改它的值, 这样我们每次转移需要两个随机数 a_i, b_i, 分别用于确定选取的是哪一位以及是否修改.

为了求 *mixing time*, 考虑 coupling, 令 $\{a_i\}, a_i\in[n], \{b_i\}, b_i\in[0,1]$ 为已经生成好的随机变量数列, 且 a_i 之间, b_i 之间, a 与 b 都满足 *mutually independent*.

考虑任意两个 n 位的初始 string:x,y, 对于任一 string:s 每步转移的规则如下:

a) $0\le b_i\le p$: 将 s 的第 b_i 位设为 0.

b) $p<b_i\le 1-p$: 翻转 s 的第 b_i 位.

c) $1-p<b_i\le 1$: 将 s 的第 b_i 位设为 1.

容易证明, 对于任意的当先的 string 状态, 会以 p 的概率停留在原来的状态, 否则以等概率转移到相邻状态.

这样第 k 位, 一旦存在某一步 $i : a_i = k, b_i \le p$ or $b_i > 1 - p$(规则 aa或规则 ac), 则 $x(j) = y(j), \forall j > i$.

$$\begin{array}{rcl} \forall x, y & & \\ \Delta(t) & \le & Pr[X_t \neq Y_t] \\ & \le & \sum_{i=1}^{n} Pr[X_t(i) \neq Y_t(i)] \\ & = & \sum_{i=1}^{n} \left(1 - \frac{2p}{n}\right)^t \\ & = & n\left(1 - \frac{2p}{n}\right)^t \\ & \le & ne^{-\frac{2pt}{n}} \end{array}$$

令其 $\le \frac{1}{2e}$, 求得:

$$\tau_{mix} = \frac{1}{2p}\left(n\log n + n\log 2 + n\right)$$

将第一问的 $p = \frac{1}{n+1}$ 代入, 得:

$$\tau_{mix} = \frac{n+1}{2}\left(n\log n + n\log 2 + n\right)$$

.

(b) 另一种 coupling 方式可以得出一个更低阶的界 (via 游宇榕):

同上文生成两组随机变量 $\{a_i \in [n+1]\}, \{b_i \in [0,1]\}$.

在算法运行到任意一步时, 对于当前的 string:x, y, 设 $next(i) = j$ 当且仅当 $x(i) \neq y(i)$ 且

$$j = \min\left\{k > i \mid x(k) \neq y(k)\right\}$$

(如果 i 是最后一个 $x(k) \neq y(k)$ 则 $next(i) = n+1$, 特别的 $next(n+1) = \min\left\{k \mid x(k) \neq y(k)\right\}$).

定义 $f : [n+1] \to [n+1]$ 如下:

$$f(i) = \begin{cases} i & x(i) = y(i) \\ next(i) & x(i) \neq y(i) \end{cases}$$

可以验证 f 是一个 bijection.

a) 如果 $x(a_i) = y(a_i)$, 则

  i. $0 \le b_i \le p$: 不变.

ii. $p < b_i < 1$: 翻转 $x(a_i)$ 和 $y(a_i)$.

如果 $x(a_i) \neq y(a_i)$, 则

i. $0 \le b_i \le p$: 不变.

ii. $p < b_i < 1$: 翻转 $x(a_i)$ 与 $y(next(a_i))$.

可以验证这样 coupling 不改变 *marginal distribution* 而且可以保证任意两步操作的独立性.

下面开始求 *mixing time*:

$$
\begin{aligned}
\forall x, y & \\
\Delta(t) &\le Pr[X_t \neq Y_t] \\
&\le \sum_{i=1}^{n} Pr[X_t(t) \neq Y_t(i)] \\
&\le \sum_{i=1}^{n} \left(p + (1-p)\frac{n-2}{n}\right)^t \\
&\le n\left(1 - \frac{2(1-p)}{n}\right)^t \\
&\le ne^{-\frac{2(1-p)t}{n}}
\end{aligned}
$$

令其 $\le \frac{1}{2e}$, 得:

$$\tau_{mix} = \frac{1}{2(1-p)}\left(n\log n + n\log 2 + n\right)$$

为了计算第一问, 将 $p = \frac{1}{n+1}$ 代入:

$$\tau_{mix} = \frac{1}{2}\left((n+1)\log n + (n+1)\log 2 + (n+1)\right)$$

Problem 6

(a) 任意两个大小为 k 的集合 S_1, S_2 都可以通过一系列 $S \to S + S_2(i) - S_1(i)$ 的操作转化得到, 因此这个 *random walk* 是 irreducible 的.

当加上 lazy 条件 (即每个 state 有 $\frac{1}{2}$ 的概率不转移), 此时 *random walk* 是 ergodic 的.

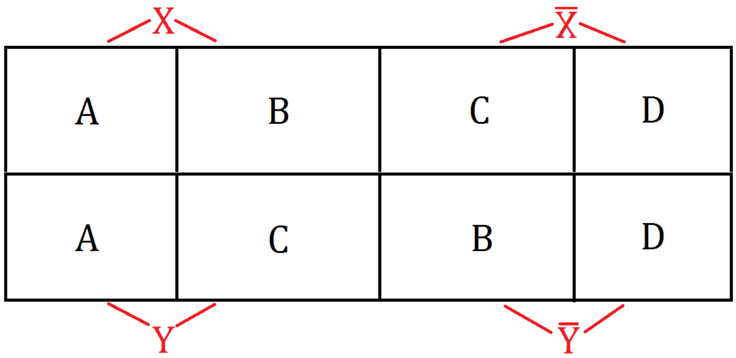
不加 lazy 条件时, 当 $n = 2, k = 1$ 时, *random walk* 不是 ergodic 的. 当 $n \ge 3$ 时, 每个状态的周期都是 $\gcd(2,3) = 1$($S \to S - a + b \to S, S \to S - a + b \to S - a + c \to S$), *random walk* 也是 ergodic 的.

对于任意两个状态 S_1, S_2, 有 $P(S_1, S_2) = P(S_2, S_1)$, 因此转移矩阵是对称的, 有对应于特征值 1 的特征向量 $\mathbf{v}$, 单位化之后得到 *uniform stationary distribution*: π.

(b) 对于任意两个初始状态 X, Y, 在进行第 x 步操作时, Ω 中的元素可以分为如下四类:

- $A_x = X_x \cap Y_x$
- $B_x = X_x \setminus (X_x \cap Y_x)$
- $C_x = Y_x \setminus (X_x \cap Y_x)$
- $D_x = \overline{X_x \cup Y_x}$

其中 $|B_x| = |C_x|$, 设 $|A_x| = i$, 则 $|B_x| = |C_x| = k - i$



考虑如下的 coupling 形式, 我们将 A_x, B_x, C_x, D_x 分别从小到大排序, 用下标表示其中的元素 (下标越界时从头开始), $X_x = A_x \cup B_x$, $Y_x = A_x \cup C_x$.

随机出两个数 $a \in [k], b \in [n-k]$, 则第 x 步对 X_x, Y_x 进行如下操作:

(1) $a \le i, b \le k - i$:

$$X_{x+1} = X_x - A_x(a) + C_x(b)$$

$$Y_{x+1} = Y_x - A_x(a) + B_x(b)$$

(2) $a > i, b \le k - i$:

$$X_{x+1} = X_x - B_x(a) + C_x(b)$$

$$Y_{x+1} = Y_x - C_x(a+1) + B_x(b+1)$$

(3) $a \le i, b > k - i$:

$$X_{x+1} = X_x - A_x(a) + D_x(b)$$

$$Y_{x+1} = Y_x - A_x(a) + D_x(b)$$

(4) $a > i, b > k - i$:

$$X_{x+1} = X_x - B_x(a) + D_x(b)$$

$$X_{x+1} = X_k - C_x(a) + D_x(b)$$

可以证明这样的 coupling 对应的边际分布是与原问题相同的.

除了 (3), 其它几种操作后, $|X \cap Y|$ 都会上升至少 1, 在操作 (3) 后, $|X \cap Y|$ 不变.

这样, 我们可以计算使得 $X_t = Y_t$ 的最小的 t, 设 T_i 为 $|X \cap Y| = i$ 的最早时间, 则有 (第一个不等式由几何分布的期望得出, 不是等号的原因是操作 (2) 使 $|X \cap Y|$ 增长超过 1):

$$
\begin{aligned}
E[T_{i+1}] &\leq E[T_i] + \frac{1}{1-\frac{i}{k}\times\frac{n-2k+i}{n-k}} \\
&\leq E[T_i] + \frac{1}{1-\frac{i}{k}} \\
&\leq E[T_i] + \frac{k}{k-i}
\end{aligned}
$$

解得

$$E[T_n] \leq \sum_{i=0}^{k-1} \frac{k}{k-i} = k\log k + O(1)$$

为了求出 *mixing time,*

$$
\begin{aligned}
&\forall X, Y \\
\Delta(t) &\leq Pr[X_t \neq Y_t] \\
&= Pr[T_n > t] \\
&\leq \frac{k\log k + O(1)}{t}
\end{aligned}
$$

令其 $\leq \frac{1}{2e}$, 求出:

$$\tau_{mix} = 2ek\log k + O(1) = O(n\log k)$$