פתרון תרגיל מספר 8 ־ כלים מתמטיים

שם: מיכאל גרינבאום, **ת.ז:** 211747639

21 בדצמבר 2020

1. **פתרון:**

(א) **צ"ל:** $d_{TV}(p,q)=\max_{S\subseteq[n]}\left[p(S)-q(S)\right]$

הוכחה:

נסמן $A=\{i:p_i\geq q_i\}$,

תהי $S\subseteq[n]$, נשים לב כי

$$\begin{aligned}p(S)-q(S)&=\sum_{i\in S}\left[p_i-q_i\right]=\sum_{i\in S\wedge i\in A}\left[p_i-q_i\right]+\sum_{i\in S\wedge i\notin A}\left[p_i-q_i\right]\\&\overset{\text{by } A\text{'s definition}}{=}\sum_{i\in S\wedge i\in A}|p_i-q_i|-\sum_{i\in S\wedge i\in[n]\setminus A}|p_i-q_i|\\&\leq\sum_{i\in S\wedge i\in A}|p_i-q_i|\leq\sum_{i\in A}|p_i-q_i|=p(A)-q(A)\end{aligned}$$

כלומר הראנו כי $\max_{S\subseteq[n]}\left[p(S)-q(S)\right]=p(A)-q(A)$,

עתה נשים לב כי

$$\sum_{i\in A}\left[p_i-q_i\right]+\sum_{i\in[n]\setminus A}\left[p_i-q_i\right]=\sum_{i\in[n]}\left[p_i-q_i\right]=\sum_{i\in[n]}p_i-\sum_{i\in[n]}q_i=1-1=0$$

כלומר קיבלנו כי $\sum_{i\in A}\left[p_i-q_i\right]=-\sum_{i\in[n]\setminus A}\left[p_i-q_i\right]$ ולכן

$$\sum_{i\in A}|p_i-q_i|=\sum_{i\in A}\left[p_i-q_i\right]=-\sum_{i\in[n]\setminus A}\left[p_i-q_i\right]=-1\cdot-1\sum_{i\in[n]\setminus A}|p_i-q_i|=\sum_{i\in[n]\setminus A}|p_i-q_i|$$

ולכן

$$\begin{aligned}2\cdot\sum_{i\in A}|p_i-q_i|&=\sum_{i\in A}|p_i-q_i|+\sum_{i\in A}|p_i-q_i|=\sum_{i\in A}|p_i-q_i|+\sum_{i\in[n]\setminus A}|p_i-q_i|\\&=\sum_{i\in[n]}|p_i-q_i|=\|p-q\|_1=2\cdot d_{TV}(p,q)\end{aligned}$$

ולכן קיבלנו כי

$$d_{TV}(p,q)=\sum_{i\in A}|p_i-q_i|=\sum_{i\in A}\left[p_i-q_i\right]=p(A)-q(A)=\max_{S\subseteq[n]}\left[p(S)-q(S)\right]$$

כלומר קיבלנו $\boxed{d_{TV}(p,q)=\max_{S\subseteq[n]}\left[p(S)-q(S)\right]}$, כנדרש

מ.ש.ל.א.☺

(ב) **צ"ל:** $\varepsilon_p(S)+\varepsilon_q(S)\geq 1-d_{TV}(p,q)$

הוכחה:

נשים לב ש־ $\varepsilon_p(S)$ היא בדיוק השגיאה שניחשנו q_i עבור בחירה של $i\in S$, שזה בדיוק $q(S)$,
נשים לב ש־ $\varepsilon_q(S)$ היא בדיוק השגיאה של p_i עבור בחירה של $i\notin S$, שזה בדיוק $p\left(\overline{S}\right)$,
וגם מתקיים כי $\varepsilon_q(S)+\varepsilon_q\left(\overline{S}\right)=p(S)+p\left(\overline{S}\right)=1$,
לאחר אבחנות אלו, נקבל כי

$$\begin{aligned}\varepsilon_p(S)+\varepsilon_q(S)&=\varepsilon_p(S)+1-\varepsilon_q\left(\overline{S}\right)=1+\left(\varepsilon_p(S)-\varepsilon_q\left(\overline{S}\right)\right)\geq 1-\max_{S\subseteq[n]}\left[\varepsilon_p(S)-\varepsilon_q\left(\overline{S}\right)\right]\\&=1-\max_{S\subseteq[n]}\left[q(S)-p(S)\right]=1-d_{TV}(q,p)=1-d_{TV}(p,q)\end{aligned}$$

מ.ש.ל.ב.☺

2. **פתרון:**

(א) **צ"ל:** $\pi=\left(\frac{1}{n},\ldots,\frac{1}{n}\right)$

הוכחה:

נשים לב כי P^T היא גם מטריצה סטוכסטית וגם מתקיים $\pi P^T=\pi P=\pi$,
לכן מתקיים $\pi^T=\left(\pi P^T\right)^T=P\cdot\pi^T$,
ראינו לפי משפט פרון פורבליוס שיש וקטור עצמי חיובי יחיד עד כדי כפולה בסקלר למטריצה ארגודית, ובהרצאה ראינו שהוקטור הימני הזה הוא 1_n למטריצה סטוכסטית P,
לכן קיים סקלר c כך שמתקיים $\pi^T=c\cdot 1_n$ וגם מתקיים $\sum_{i=1}^n\pi_i=1$ בגלל ש־ π וקטור התפלגות ולכן $\sum_{i=1}^n c\cdot 1=$
$nc=1\implies c=\frac{1}{n}$,
כלומר $\boxed{\pi^T=1_n\cdot\frac{1}{n}=\left(\frac{1}{n},\ldots,\frac{1}{n}\right)^T\implies\pi=\left(\frac{1}{n},\ldots,\frac{1}{n}\right)}$, כנדרש

מ.ש.ל.א.☺

(ב) **צ"ל:** $d_{TV}\left(x^{(t+1)},\pi\right)\leq d_{TV}\left(x^t,\pi\right)$

הוכחה:

נשים לב כי

$$\begin{aligned}d_{TV}\left(x^{(t+1)},\pi\right)&=\frac{1}{2}\cdot\left\|x^{(t+1)}-\pi\right\|_1=\frac{1}{2}\cdot\left\|x^{(t)}\cdot P-\pi\right\|_1=\frac{1}{2}\cdot\left\|x^{(t)}\cdot P-\pi\cdot P\right\|_1\\&=\frac{1}{2}\cdot\left\|\left(x^{(t)}-\pi\right)\cdot P\right\|_1=\frac{1}{2}\cdot\sum_{i=1}^n\left|\left[\left(x^{(t)}-\pi\right)\cdot P\right]_i\right|\\&=\frac{1}{2}\cdot\sum_{i=1}^n\left|\sum_{j=1}^n\left(x^{(t)}-\pi\right)_j\cdot P_{j,i}\right|\leq\frac{1}{2}\cdot\sum_{i=1}^n\sum_{j=1}^n\left|\left(x^{(t)}-\pi\right)_j\right|\cdot\left|P_{j,i}\right|\\&=\frac{1}{2}\cdot\sum_{j=1}^n\sum_{i=1}^n\left|\left(x^{(t)}-\pi\right)_j\right|\cdot\left|P_{j,i}\right|=\frac{1}{2}\cdot\sum_{j=1}^n\left[\left|\left(x^{(t)}-\pi\right)_j\right|\cdot\sum_{i=1}^n\left|P_{j,i}\right|\right]\\&=\frac{1}{2}\cdot\sum_{j=1}^n\left[\left|\left(x^{(t)}-\pi\right)_j\right|\cdot\sum_{i=1}^n P_{j,i}\right]=\frac{1}{2}\cdot\sum_{j=1}^n\left[\left|\left(x^{(t)}-\pi\right)_j\right|\cdot 1\right]\\&=\frac{1}{2}\cdot\sum_{j=1}^n\left|\left(x^{(t)}-\pi\right)_j\right|=\frac{1}{2}\cdot\left\|x^{(t)}-\pi\right\|_1=d_{TV}\left(x^{(t)},\pi\right)\end{aligned}$$

כלומר קיבלנו כי $\boxed{d_{TV}\left(x^{(t+1)},\pi\right)\leq d_{TV}\left(x^t,\pi\right)}$, כנדרש.

מ.ש.ל.ב.☺

(ג) **צ"ל:** $\tau(\varepsilon)=\max_{1\le i\le n}\tau(\varepsilon\mid e_i)$

הוכחה:

תהי x_0 התפלגות התחלתית ונסמן $x_0=\sum_{i=1}^{n}a_i\cdot e_i$, ו־ $t\in\mathbb{N}$,

$$\begin{aligned}
d_{TV}\left(x_0^{(t)},\pi\right)&=\frac{1}{2}\cdot\left\|x_0^{(t)}-\pi\right\|_1=\frac{1}{2}\cdot\left\|x_0\cdot P^t-\pi\right\|_1=\frac{1}{2}\cdot\left\|\sum_{i=1}^{n}a_i\cdot e_i\cdot P^t-\pi\right\|_1\\
&\overset{\sum_{i=1}^{n}a_i=1}{=}\frac{1}{2}\cdot\left\|\sum_{i=1}^{n}\left[a_i\cdot e_i\cdot P^t-a_i\cdot\pi\right]\right\|_1\le\frac{1}{2}\cdot\sum_{i=1}^{n}\left\|a_i\cdot e_i\cdot P^t-a_i\cdot\pi\right\|_1\\
&=\frac{1}{2}\cdot\sum_{i=1}^{n}|a_i|\cdot\left\|e_i\cdot P^t-\pi\right\|_1\overset{a_i\in[0,1]}{=}\frac{1}{2}\cdot\sum_{i=1}^{n}a_i\cdot\left\|e_i\cdot P^t-\pi\right\|_1\\
&=\sum_{i=1}^{n}a_i\cdot\frac{1}{2}\cdot\left\|e_i\cdot P^t-\pi\right\|_1=\sum_{i=1}^{n}a_i\cdot d_{TV}\left(e_i^{(t)},\pi\right)
\end{aligned}$$

כלומר הראנו כי

$$d_{TV}\left(x_0^{(t)},\pi\right)\le\sum_{i=1}^{n}a_i\cdot d_{TV}\left(e_i^{(t)},\pi\right)$$

נסמן ב־ $\alpha=\max_{1\le i\le n}d_{TV}\left(e_i^{(t)},\pi\right)$ ונקבל כי

$$\sum_{i=1}^{n}a_i\cdot d_{TV}\left(e_i^{(t)},\pi\right)\le\sum_{i=1}^{n}a_i\cdot\alpha=\alpha\cdot\sum_{i=1}^{n}a_i=\alpha\cdot 1=\alpha$$

כלומר הראנו כי $d_{TV}\left(x_0^{(t)},\pi\right)\le\max_{1\le i\le n}d_{TV}\left(e_i^{(t)},\pi\right)$ לכל התפלגות התחלתית x_0, ולכן

$$\tau(\varepsilon\mid x_0)=\min_{t\in\mathbb{N}}\left\{t:d_{TV}\left(x_0^{(t)},\pi\right)\le\varepsilon\right\}\le\min_{t\in\mathbb{N}}\left\{t:\max_{1\le i\le n}d_{TV}\left(e_i^{(t)},\pi\right)\le\varepsilon\right\}=\max_{1\le i\le n}\tau(\varepsilon\mid e_i)$$

ולכן נקבל כי

$$\tau(\varepsilon)=\max_{x_0}\tau(\varepsilon\mid x_0)\le\max_{1\le i\le n}\tau(\varepsilon\mid e_i)$$

וגם נשים לב כי $\tau(\varepsilon)=\max_{x_0}\tau(\varepsilon\mid x_0)\ge\max_{1\le i\le n}\tau(\varepsilon\mid e_i)$ כי e_i בעצמם חלק מהאפשרויות ל־ x_0 ולכן קיבלנו כי $\boxed{\tau(\varepsilon)=\max_{1\le i\le n}\tau(\varepsilon\mid e_i)}$

מ.ש.ל.ג.☺

3. **צ"ל:** $\tau\left(n^{-4}\right)=\Omega\left(\frac{\log(n)}{\gamma}\right)$

הוכחה:

נסמן $P=\frac{1}{d}A_G$,

תחילה נסמן ב־ $u_1,\dots,u_n$ בסיס של וקטורים עצמיים אורתונורמלי של $\frac{1}{d}A_G$ שמתאימים לערכים העצמיים $\lambda_1\ge\lambda_2\ge\cdots\ge\lambda_n$ בהתאמה. נשים לב כי ראינו בתרגול ובהרצאה כי $u_1=n^{-1}\cdot 1=\pi^T$ כאשר π זאת ההתפלגות הסטציונרית.

הערה: הבסיס אורתונורמלי אבל u_1 הוא מנורמל ביחס לנורמה 1 ולא 2 כמו השאר

למשך ההוכחה נניח בלי הגבלת הכלליות כי מתקיים $\lambda_2\ge|\lambda_n|$,

נסתכל על הוקטור $x_0=u_1=n^{-1}\cdot 1=\pi^T+n^{-1}\cdot u_2$,

תחילה נשים לב כי $x_0\ge 0$, וגם כי $0=\langle u_1,u_2\rangle=\sum_{i=1}^{n}[u_1]_i\cdot[u_2]_i=\sum_{i=1}^{n}[u_2]_i$ ולכן

$$\sum_{i=1}^{n}[x_0]_i=\sum_{i=1}^{n}[u_1]_i+[u_2]_i=\sum_{i=1}^{n}[u_1]_i+\sum_{i=1}^{n}[u_2]_i=1+0=1$$

וגם נשים לב שכל קורדיאנטה ב־ u_2 היא לכל היותר $\frac{1}{2}$ (נסכמים ל1 והסכום הוא 0) ולכן $[x_0]_i = [u_1]_i + [u_2]_i \leq \frac{1}{n} + \frac{1}{2} \leq 1$,
כלומר הראנו ש־ $0 \leq x_0 \leq 1$ וגם $\sum_{i=1}^{n} [x_0]_i = 1$ ולכן הוקטור שהוגדר הוא אכן וקטור התפלגות.
יהי $t \in \mathbb{N}$, תחילה נשים לב כי

$$\begin{aligned} x_0^T \cdot P^t &= \left(u_1 + n^{-1} \cdot u_2\right)^T \cdot P^t = \left(\pi + n^{-1} \cdot u_2^T\right) \cdot P^t = \pi \cdot P^t + n^{-1} \cdot u_2^T \cdot P^t \\ &= \pi + n^{-1} \cdot u_2^T \cdot P^t = \pi + n^{-1} \cdot u_2^T \cdot \lambda_2^t \end{aligned}$$

ולכן נקבל כי

$$\begin{aligned} \left\|x_0^T \cdot P^t - \pi\right\|_1 &= \frac{1}{2} \left\|\pi + n^{-1} \cdot u_2^T \cdot \lambda_2^t - \pi\right\|_1 = \frac{1}{2} \left\|n^{-1} \cdot u_2^T \cdot \lambda_2^t\right\|_1 \\ &\overset{\lambda_2 = 1-(1-\max\{\lambda_2, |\lambda_n|\}) = \gamma}{=} \frac{1}{2} \left\|n^{-1} \cdot u_2^T \cdot (1-\gamma)^t\right\|_1 \geq \frac{1}{2} \left\|n^{-1} \cdot u_2^T \cdot (1-\gamma)^t\right\|_2 = \frac{1}{2} \frac{(1-\gamma)^t}{n} \end{aligned}$$

ונשים לב כי

$$\frac{(1-\gamma)^t}{2n} \geq n^{-4} \iff (1-\gamma)^t \geq 2n^{-3} \iff t \cdot \ln(1-\gamma) \geq \ln\left(2n^{-3}\right) \iff t \geq \frac{-3\ln(2n)}{\ln(1-\gamma)} = \Theta\left(\frac{\ln(n)}{\gamma}\right)$$

כלומר הראנו כי עבור $t = \Omega\left(\frac{\ln(n)}{\gamma}\right)$ מתקיים $d_{TV}\left(x_0^T, \pi\right) = \frac{1}{2}\left\|x_0^T \cdot P^t - \pi\right\|_1 \geq \frac{(1-\gamma)^t}{2n} \geq n^{-4}$,
כלומר מצאנו וקטור התפלגות x_0 כך ש־ $d_{TV}\left(x_0^T, \pi\right) \geq n^{-4}$ כאשר $t = \Omega\left(\frac{\ln(n)}{\gamma}\right)$ ולכן $\tau(\varepsilon) \geq \tau\left(\varepsilon \mid x_0^T\right) = \Omega\left(\frac{\ln(n)}{\gamma}\right)$,
ולכן $\boxed{\tau(\varepsilon) = \Omega\left(\frac{\ln(n)}{\gamma}\right)}$, כנדרש.

מ.ש.ל.☺

4. **פתרון:**

(א) **צ"ל:** מטריצת מעבר
הוכחה:
נסמן ב־ P את מטריצת המעבר של ה־ SRW,
נשים לב כי $\mathbb{P}(X_{n+1} = i \mid X_n = i) = \frac{1}{2}$ לפי ההגדרה,
וגם $\mathbb{P}(X_{n+1} = j \mid X_n = i) = \mathbb{P}(\text{we didn't stay in i}) \cdot P_{i,j} = \frac{1}{2}P_{i,j}$ עבור $i \neq j$.
נשים לב כי מטריצת המעבר המקורית תקרה חצי מהזמן וחצי מהזמן נשאר באותו המצב ולכן ניתן לכתוב $\boxed{P' = \frac{1}{2}P + \frac{1}{2}I_n}$
וזאת תהיה מטריצת המעבר של LRW

מ.ש.ל.א.☺

(ב) **צ"ל:** אותה התפלגות סטציונרית
הוכחה:
תהי π התפלגות סטציונרית של ה־ SRW ולכן $\pi \cdot P = \pi$, ונשים לב כי

$$\pi \cdot P' = \pi \cdot \left(\frac{1}{2}P + \frac{1}{2}I_n\right) = \frac{1}{2} \cdot \pi \cdot P + \frac{1}{2} \cdot \pi \cdot I_n = \frac{1}{2} \cdot \pi + \frac{1}{2} \cdot \pi = \pi$$

ולכן ההתפלגות הסטציונית היא גם של ה־ LRW

מ.ש.ל.ב.☺

(ג) **צ"ל:** ה־ LRW היא ארגודית
הוכחה:
נסמן $Q = P'$,
נוכיח כי $Q^n > 0$, כלומר לכל i, j יש מסלול באורך n בינהם.

יהיו v_i, v_j קודקודים ב־ G.
תחילה מהיות ו־ G קשיר, קיים מסלול בין i ל־ j באורך $k \leq n$ ונסמן $v_i, v_{l_1}, \ldots, v_{l_{k-1}}, v_j$ מסלול זה.
נשים לב כי $\underbrace{v_i, \ldots, v_i}_{n-k \text{ times}}, v_{l_1}, \ldots, v_{l_{k-1}}, v_j$ הוא מסלול באורך n בגרף ה־ LRW בין v_i ל־ v_j ולכן $Q^n_{i,j} > 0$.
כלומר מצאנו חזקה n כך ש־ $Q^n > 0$ ולכן המטריצה אכן ארגודית.

מ.ש.ל.ג.☺

(ד) **צ"ל:** $\tau_{LRW}(\varepsilon) = \mathbf{polylog}(n)$
הוכחה:
נלך לפי טענת העזר, נבחר $t = \left(\tau_{SRW}\left(\frac{\varepsilon}{2}\right)\right)^2$, נשים לב ש־

$$\begin{aligned}
d_{TV}(x_t, \pi) &= \frac{1}{2} \cdot \|x_t - \pi\|_1 = \frac{1}{2} \cdot \left\| x_0 \cdot \left(\frac{1}{2}P + \frac{1}{2} \cdot I_n\right)^t - \pi \right\|_1 \\
&= \frac{1}{2} \cdot \left\| \frac{1}{2^t} \cdot x_0 \cdot \sum_{i=0}^{t} \binom{t}{i} \cdot P^i - \pi \right\|_1 = \frac{1}{2} \cdot \frac{1}{2^t} \cdot \left\| \sum_{i=0}^{t} \binom{t}{i} \cdot x_0 \cdot P^i - \pi \right\|_1 \\
&= \frac{1}{2} \cdot \frac{1}{2^t} \cdot \left\| \sum_{i=0}^{t} \binom{t}{i} \cdot x_0 \cdot P^i - \pi \cdot P^i \right\|_1 \leq \frac{1}{2} \cdot \frac{1}{2^t} \sum_{i=0}^{t} \binom{t}{i} \cdot \left\| x_0 \cdot P^i - \pi \cdot P^i \right\|_1 \\
&= \frac{1}{2^t} \sum_{i=0}^{t} \binom{t}{i} \cdot \left(\frac{1}{2} \cdot \left\| x_0 \cdot P^i - \pi \right\|_1\right) = \frac{1}{2^t} \sum_{i=0}^{t} \binom{t}{i} \cdot d_{TV}\left(x_0 \cdot P^i, \pi\right) \\
&= \frac{1}{2^t} \left[\sum_{i=0}^{\sqrt{t}} \binom{t}{i} \cdot d_{TV}\left(x_0 \cdot P^i, \pi\right) + \sum_{i=\sqrt{t}+1}^{t} \binom{t}{i} \cdot d_{TV}\left(x_0 \cdot P^i, \pi\right) \right] \\
&\leq \frac{1}{2^t} \left[\sum_{i=0}^{\sqrt{t}} \binom{t}{i} \cdot 1 + \sum_{i=\sqrt{t}+1}^{t} \binom{t}{i} \cdot \frac{\varepsilon}{2} \right] \leq \frac{1}{2^t} \left[\sum_{i=0}^{\sqrt{t}} \binom{t}{i} + \frac{\varepsilon}{2} \cdot \sum_{i=0}^{t} \binom{t}{i} \right] \\
&= \frac{1}{2^t} \left[\sum_{i=0}^{\sqrt{t}} \binom{t}{i} + \frac{\varepsilon}{2} \cdot 2^l \right] = \frac{1}{2^t} \cdot \sum_{i=0}^{\sqrt{t}} \binom{t}{i} + \frac{\varepsilon}{2} = \sum_{i=0}^{\sqrt{t}} \frac{1}{2^t} \cdot \binom{t}{i} + \frac{\varepsilon}{2}
\end{aligned}$$

נסמן $X \sim Bin\left(t, \frac{1}{2}\right)$, לכן נקבל לפי הופדינג כי

$$\begin{aligned}
\sum_{i=0}^{\sqrt{t}} \frac{1}{2^t} \cdot \binom{t}{i} &= \mathbb{P}\left(X \leq \sqrt{t}\right) = \mathbb{P}\left(X - \frac{1}{2}t \leq \sqrt{t} - \frac{1}{2}t\right) \leq \mathbb{P}\left(\left|X - \frac{1}{2}t\right| \leq \frac{1}{2}t - \sqrt{t}\right) \leq 2e^{-\frac{\left(\frac{1}{2}t - \sqrt{t}\right)^2}{2t}} \\
&\overset{\star}{\leq} 2\, e^{-t \cdot c}
\end{aligned}$$

נשים לב ש־ $\star$ נכון מהיות $\frac{\frac{\left(\frac{1}{2}t - \sqrt{t}\right)^2}{2t}}{t} \to \frac{1}{2}$ אז מתקיים ש־ $c_1 \cdot t \leq \frac{\left(\frac{1}{2}t - \sqrt{t}\right)^2}{t} \leq c_2 \cdot t$ כמעט תמיד (כש־ c_1, c_2 קבועים),
ולכן ניתן להחליף את הביטוי ב־ $t \cdot c$ כש־ c קבוע כלשהו (קרוב ל־ $\frac{1}{2}$)
כלומר נציב ונקבל כי $d_{TV}(x_t, \pi) \leq 2e^{-t \cdot c} + \frac{\varepsilon}{2}$ ונשים לב שעבור $t_1 = \max\left\{t, -\frac{\log\left(\frac{\varepsilon}{4}\right)}{c}\right\}$ ונקבל כי

$$d_{TV}\left(x_{t_1}, \pi\right) \leq 2e^{-t \cdot c} + \frac{\varepsilon}{2} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

כלומר הראנו כי $\tau(\varepsilon) \leq t_1$ ונשים לב כי $t_1 = \max\left\{t, -\frac{\log\left(\frac{\varepsilon}{4}\right)}{c}\right\} = \text{polylog}(n)$ ולכן $\tau(\varepsilon) = \text{polylog}(n)$, כנדרש.

מ.ש.ל.ד.☺