19.04.2023 | ***ניתוח מידע אדפטיבי***

הרצאה 4: אורך ייצוג חסום + יציבות

Source: Lecture notes by Aaron Roth and Adam Smith

מרצה: אורי שטמר

איך נוכל לקבל מכניזם שממש עונה על הרבה שאילתות בלי שהוא יהיה צריך לקבל "ניחושים"?

לצורך כך נעזר בטענה הבאה:

<u>טענה 1:</u> לכל $\alpha > 0$ ולכל קביעה של k שאילתות סטטיסטיות $q_1, \dots, q_k: X \to [0,1]$ ולכל דטהבייס $S \in X^n$, קיים דטהבייס $S' \in X^{n'}$ בגודל $n' = \frac{\ln(4k)}{2\alpha^2}$ כך ש:

$$\max_i |q_i(S) - q_i(S')| \leq \alpha$$

<u>הוכחה:</u> יהי S' דטהבייס המתקבל על ידי דגימה של n' נקודות מ- S באופן בלתי תלוי (עם חזרות). כלומר S' הוא דטהבייס בגודל n' שנדגם iid מההתפלגות האחידה על פני האיברים ב- S, שנסמנה U_S. כלומר $S' \sim (U_S)^{n'}$.
אזי, לפי חסם צ'רנוף, בהסתברות לפחות ½ מתקיים ש-

$$\max_i |q_i(S') - q_i(U_S)| \leq \sqrt{\frac{\ln(4k)}{2n'}} = \alpha$$

נשים לב שעבור ההתפלגות הזאת, לכל שאילתא סטטיסטית q מתקיים ש-

$$q(U_S) = \mathbb{E}_{x \sim U_S}[q(x)] = \frac{1}{|S|}\sum_{x \in S} q(x) = q(S)$$

כלומר, התוחלת של השאילתא (תחת U_S) שווה לממוצא האמפירי של q על S. לכן, בהסתברות לפחות ½ מתקיים

$$\max_i |q_i(S') - q_i(S)| \leq \alpha \qquad ((1))$$

מה הראינו? אם מגרילים את S' בצורה הזאת, אזי בהסתברות לפחות ½ מקבלים S' שמקיים את $((1))$. בפרט, ז"א <u>שקיים</u> דטהבייס S' בגודל n' שמקיים את $((1))$.
(אחרת לא היינו יכולים לקבל כזה בהסתברות ½ כשאנחנו מגרילים את S'...)

מ.ש.ל.

עכשיו, בעזרת שילוב של טענה 1 עם אלגוריתם GuessAndCheck, אנחנו יכולים להציג מכניזם שעונה על k שאילתות סטטיסטיות (ולא מקבל "ניחושים"...).

MedianMechanism$(S, q_1, q_2, \dots)$

Input:

- Sample S containing n elements from the domain X

Parameters:

- Accuracy parameter $\eta = \left(\frac{\ln(4k)}{2n}\right)^{1/4}$
- Subsample size $n' = \frac{8 \cdot \ln(4k)}{\eta^2} = 8 \cdot \sqrt{\ln(4k) \cdot 2n}$
- Number of guesses $m = n' \cdot \log|X| = 8 \cdot \sqrt{\ln(4k) \cdot 2n} \cdot \log|X|$

2. Initialize an instance of GuessAndCheck(η, m) on S
3. Initialize a set of all "possible" datasets $\mathbb{S}_0 = X^{n'}$ (that is, $\mathbb{S}_0$ contains all possible datasets containing n' elements from X)
4. For $i = 1,2, \dots, k$ do
    a. Accept the next query q_i
    b. Construct a guess $g_i = \text{median}\{q_i(S') \ : \ S' \in \mathbb{S}_{i-1}\}$
    c. Feed the query (q_i, g_i) to GuessAndCheck and receive an answer a_i
    d. If $a_i = g_i$ then set $\mathbb{S}_i \leftarrow \mathbb{S}_{i-1}$
    e. Else set $\mathbb{S}_i = \mathbb{S}_{i-1} \setminus \left\{ S' \in \mathbb{S}_{i-1} \ : \ |q_i(S') - a_i| > \frac{\eta}{2} \right\}$
    f. Return the answer a_i

<u>משפט 2:</u> לכל $\beta > 0$, מתקיים ש MedianMechanism הוא (α, β)-מדוייק סטטיסטית עבור k שאילתות סטטיסטיות, כאשר

$$\alpha = \tilde{O}\left(\frac{(\log k)^{3/4} \cdot \sqrt{\log\left(\frac{|X|}{\beta}\right)}}{n^{1/4}}\right)$$

<u>הוכחה:</u> ראשית נשים לב ש- MedianMechanism מריץ את אלגוריתם GuessAndCheck ותמיד עונה כמוהו. לכן, לפי המסקנה מסוף השיעור הקודם, אנחנו יודעים שאלגוריתם MedianMechanism הוא (α, β)-מדוייק-סטטיסטית עבור כל השאילתות שהוא מקבל <u>לפני שאלגוריתם GuessAndCheck עוצר</u>, עבור

$$\alpha = O\left(\sqrt{\frac{m \cdot \log\left(\frac{kn}{m}\right) + \log\left(\frac{1}{\beta}\right)}{n}}\right) = \tilde{O}\left(\frac{(\log k)^{3/4} \cdot \sqrt{\log\left(\frac{|X|}{\beta}\right)}}{n^{1/4}}\right)$$

כאשר השיוויון האחרון נובע מהצבת m ופישוט הביטויי (יכולנו לקבל ביטויי הדוק יותר).

לכן, כל מה שנותר לנו להראות זה שאלגוריתם GuessAndCheck <u>לא</u> עוצר לפני ששאלנו את כל k השאילתות.

לפי הגדרת GuessAndCheck, זה שקול ללהראות שישנם לכל היותר m סיבובים בהם $|q_i(S) - g_i| > \eta$, כלומר צ"ל כי ישנם לכל היותר m סיבובים בהם הניחוש שלנו שגויי.
נראה זאת על ידי מעקב אחרי $|\mathbb{S}_i|$.

תחילה נשים לב כי לפי הגדרה $|\mathbb{S}_0| = |X|^{n'}$.

<u>טענת עזר (נוכיח בהמשך):</u> בכל סיבוב i שבו הניחוש שלנו שגוי מתקיים $|\mathbb{S}_i| \leq |\mathbb{S}_{i-1}|/2$.

בנוסף, לפי טענה 1 ולפי הבחירה של n' באלגוריתם, אנחנו יודעים שלכל אוסף של k שאילתות סטטיסטיות $q_1, \dots, q_k$ <u>קיים</u> דטהבייס $S^* \in \mathbb{S}_0$ כך שלכל i מתקיים $|q_i(S^*) - q_i(S)| \leq \frac{\eta}{4}$. הדטהבייס S^* הזה אף פעם לא נמחק מ- $\mathbb{S}_i$, כלומר לכל i מתקיים ש- $S^* \in \mathbb{S}_i$ ולכן $|\mathbb{S}_i| \geq 1$.

מדוע S^* אף פעם לא נמחק? סיבובים בהם יש עדכון אלו סיבובים בהם הניחוש שלנו היה שגוי. לפי התכונות של אלגוריתם GuessAndCheck, בסיבובים אלו התשובה שהוא מחזיר לנו a_i היא $\eta/4$-מדוייקית אמפירית. כלומר

$$|a_i - q_i(S)| \leq \frac{\eta}{4}$$

לכן לפי א"ש המשולש מתקיים

$$|q_i(S^*) - a_i| \leq \frac{\eta}{2}$$

אבל אנחנו מוחקים רק דטהבייסים S' עבורם

$$|q_i(S') - a_i| > \frac{\eta}{2}$$

ולכן S^* לא נמחק.

מסקנה: נסמן ב- t את מספר הסיבובים שבהם הניחוש שלנו שגוי. אזי $|\mathbb{S}_0| \cdot \left(\frac{1}{2}\right)^t \geq 1$, ולכן

$$t \leq \log|\mathbb{S}_0| = n' \cdot \log|X| = m$$

כעת נוכיח את טענת העזר:
צ"ל כי בכל סיבוב i שבו הניחוש שלנו שגוי מתקיים $|\mathbb{S}_i| \leq |\mathbb{S}_{i-1}|/2$.

זכרו כי, לפי הגדרה, סיבוב בו הניחוש שלנו שגוי הוא סיבוב בו
$$|g_i - q_i(S)| > \eta$$

בסיבוב כזה אנחנו מוחקים מ- $\mathbb{S}$ את כל הדטהבייסים S' כך ש-
$$|q_i(S') - a_i| > \frac{\eta}{2}$$

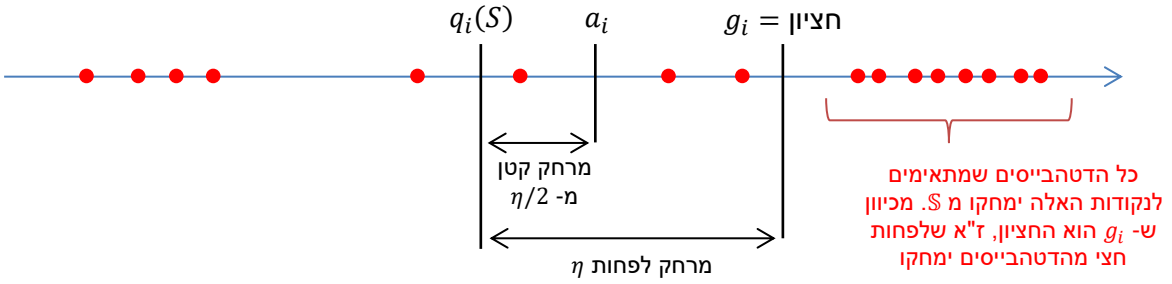
עוד זכרו כי בסיבובים כאלה, אלגוריתם GuessAndCheck מחזיר תשובה a_i כך ש-
$$|a_i - q_i(S)| \leq \frac{\eta}{2}$$

לבסוף, זכרו כי $g_i = \text{median}\{q_i(S') \ : \ S' \in \mathbb{S}_{i-1}\}$ ולכן לפחות חצי מהדטהבייסים ב- $\mathbb{S}$ נמחקים כתוצאה מכך.

מ.ש.ל.

<u>ציור עבור הוכחת טענת העזר:</u>

הנקודות האדומות מייצגות ערכים שונים של $q_i(S')$ עבור $S' \in \mathbb{S}_{i-1}$.
אנחנו קובעים את g_i להיות חציון של הנקודות האדומות האלה.



<u>דיון לגבי משפט 2:</u>
מצד אחד, קיבלנו מכניזם עם שגיאה שגדלה רק פוליגוריתמית עם מספר השאילתות k, שזה קרוב לדבר הכי טוב שיכולנו להשיג גם במקרה הלא-אדפטיבי! מעולה!
מצד שני, לתוצאה הזאת יש כמה חסרונות: (1) האלגוריתם שהצגנו הוא לא יעיל חישובית. (2) השגיאה דועכת רק כמו $1/n^{1/4}$ ולא כמו $1/\sqrt{n}$ כמו שהיינו רוצים. (3) השגיאה גדלה עם $\log|X|$, שזה משהו שאנחנו יכולים לחשוב עליו כמו "המיימד" של הדטה. במקרה הלא-אדפטיבי לא היה לנו דבר כזה...

Stability

עכשיו נתחיל לדבר על תכונה אחרת של מכניזמים אשר גם בעזרתה נוכל להבטיח דיוק-סטטיסטי. נראה שנוכל לקבל תוצאות טובות יותר מאשר בעזרת דחיסת-טרנסקריפט.

<u>הגדרה 3:</u>

א. שני דטהבייסים $S = (x_1, \dots, x_n) \in X^n$ ו- $S' = (x'_1, \dots, x'_n) \in X^n$ יקראו <u>שכנים</u> אם קיים $1 \le i \le n$ כך שלכל $j \neq i$ מתקיים $x_j = x'_j$.

ב. יהי $M: X^* \times P \to F$ מכניזם אשר מקבל כקלט דטהבייס $S \in X^*$ ופרמטר $p \in P$ ומחזיר פלט מקבוצה F. נאמר כי M הוא (ε, δ)-DP-יציב אם לכל פרמטר $p \in P$, לכל זוג דטהבייסים שכנים S, S' ולכל תת קבוצה $H \subseteq F$ מתקיים

$$\Pr[M(S,p) \in H] \le e^{\varepsilon} \cdot \Pr[M(S',p) \in H] + \delta$$

כאשר ההסתברות היא מעל האקראיות של M.

ג. יהי M מכניזם אשר מקבל כקלט דטהבייס S (ואולי גם פרמטר p) ולאחר מכן עונה על שאילתות. עבור יריב A שמציג שאילתות, נסמן ב $A \circ M$ את המכניזם שלוקח דטהבייס S ופרמטר p, מסמלץ את האינטראקציה בין A לבין $M(S,p)$, ובסיום השיחה פולט את הטרנסקריפט (ואת האקראיות של A אם הוא אקראי). נאמר כי M הוא (ε, δ)-DP-יציב אם לכל יריב A מתקיים ש- $A \circ M$ הוא (ε, δ)-DP-יציב.

<u>הערה:</u> עבור המקרה בו $\delta = 0$ ניתן לפשט קצת את ההגדרה הנ"ל (אם F בת מנייה) ולקבל את ההגדרה השקולה הבאה: מכניזם $M: X^n \to F$ מקיים $(\varepsilon, 0)$-*DP*-יציבות אם לכל $S, S' \in X^n$ שכנים ולכל **<u>איבר</u>** $h \in F$ מתקיים

$$\Pr[M(X) = h] \le e^{\varepsilon} \cdot \Pr[M(X') = h]$$

כלומר כאן h הוא <u>איבר</u> ב- F ולא תת קבוצה...

Post-processing

משפט 4: יהי $M: X^n \to R$ המקיים (ε, δ)-*DP*-יציבות ותהי $A: R \to R'$.
אזי גם המכניזם $A\big(M(\cdot)\big)$ מקיים (ε, δ)-*DP*-יציבות.

הוכחה (עבור A דטרמניסטי):
יהיו S, S' שכנים (שונים בבדיוק כניסה אחת), ותהי $H \subseteq R'$.
עלינו להראות שמתקיים

$$\Pr\big[A\big(M(S)\big) \in H\big] \leq e^{\varepsilon} \cdot \Pr\big[A\big(M(S')\big) \in H\big] + \delta$$

נסמן

$$B = \{r \in R : A(r) \in H\}$$

ונקבל

$$\Pr\big[A\big(M(S)\big) \in H\big] = \Pr[M(S) \in B] \leq e^{\varepsilon} \cdot \Pr[M(S') \in B] + \delta = e^{\varepsilon} \cdot \Pr\big[A\big(M(S')\big) \in H\big] + \delta$$

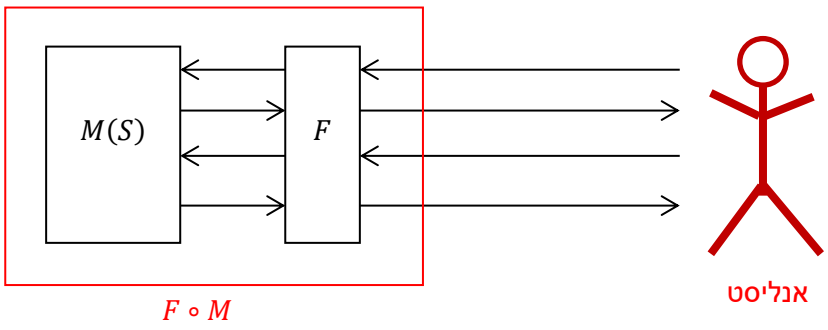
מ.ש.ל.

מה שהטענה הזאת אומרת לנו זה שאם לקחנו דטהבייס ועשינו עליו חישוב יציב, אז עכשיו אנחנו יכולים לקחת את תוצאת החישוב ולעשות איתה כל מה שאנחנו רוצים. לא משנה מה נעשה איתה – זה לא יפר יציבות.

תרגיל כיתה: מדוע תכונת ה post-processing נכונה גם עבור אלגוריתמים אינטראקטיביים (שעונים על שאילתות)?

כלומר, נניח שיש לנו מכניזם M שעונה על שאילתות שהוא (ε, δ)-DP-יציב. יהי F מכניזם כלשהו אשר מאפשר "לעבד" תשובות ש- M פולט (לפני שהתשובות האלה ניתנות לאנליסט) ומאפשר "לעבד" את שאילתות שהאנליסט בוחר (לפני שהשאילתות האלה ניתנות ל- M). מה נוכל לומר על השילוב של M ו- F, שנסמנו $F \circ M$?

בציור:



פתרון:

נקבע אנליסט A. עלינו להראות שהטרנסקריפט בין A לבין F הוא יציב. אנחנו יכולים להשוב על F כעל חלק מהאנליסט ולכן אנחנו יודעים שהטרנסקריפט בין M לבין F הוא יציב (ביחד עם כל האקראיות של A, F). לכן גם כל post-processing שלו הוא יציב, כולל זה שמשחזר את הטרנסקריפט בין A ל- F.