16.6.09

סיבוכיות - שיעור 14

השיעור התקיים בהופעת אורח מיוחדת של גליה קמפה במקום עודד!

חסם צ'רנוף

בשיעור שעבר השתמשנו את חסם צ'רנוף על מנת להוכיח כי ניתן לבנות לכל ~~[illegible]~~ בעיה ב- BPP מכונה אשר נותנת תשובה בדיוק גדול כפי שנרצה, אך לא הוכחנו אותו (או אפילו הראינו מהו).

נניח שיש בחירות בין המועמדים לנשיאות א' ו-ב', ואנחנו רוצים לערוך סקר בחירות. אנחנו רוצים לדעת את אחוז התמיכה במועמד א' בדיוק של 1%. אך בדרך כלל, לא נוכל להציג ~~[illegible]~~ תשובה בדיוק של 1% בוודאות (נצטרך לשאול 99% מהאוכלוסיה, וחלק מהם בכלל לא עונים לסקרים). לכן, נרצה להחזיר תשובה בדיוק של 1% ובוודאות של 90%, כלומר 90% שאנחנו נותנים תשובה מדויקת עד כדי 1%. באופן כללי, נרצה:

$$\text{Accuracy} = \delta$$

$$\text{Certainty} = 1-\varepsilon \quad (\varepsilon \text{ הסיכוי ל"טעות"})$$

נגדיר משתנים מקריים $X_1, X_2, \dots, X_k$ כך:

$$X_i = \begin{cases} 1 & \text{אם הנשאל ה-}i\text{ הצביע ל-א'} \\ 0 & \text{אם הנשאל ה-}i\text{ הצביע ל-ב'} \end{cases}$$

נניח לשם החישוב כי 60% תומכים ב-א', ו-40% תומכים ב-ב'.
אז $E[X_i]=0.6$. נגדיר את הממוצע להיות $X=\frac{1}{k}\sum_{i=1}^{k}X_i$.
אז גם $E[X]=0.6$.

אז חסם צ'רנוף אומר לנו שעבור $X_1, X_2, \dots, X_k$ משתנים מקריים בלתי תלויים בינאריים (אפס או אחד) מתקיים:

$$\Pr\left[\left|\frac{1}{k}\sum_{i=1}^{k}X_i - E[X]\right| > \delta\right] \le 2\cdot e^{-2\delta^2 k}$$

כלומר שהסיכוי שאנחנו רחוקים מהתוחלת ביותר מ-δ הוא $2\cdot e^{-2\delta k}$.
נרצה לבדוק מהו ה-k המינימלי (כמה אנשים נצטרך לשאול) על מנת שהסיכוי לטעות יהיה קטן מ-ε.

$$2e^{-2\delta^2 k} \le \varepsilon$$

$$\ln 2 - 2\delta^2 k \le \ln \varepsilon$$

$$2\delta^2 k - \ln 2 \ge \ln \frac{1}{\varepsilon}$$

$$k = O\left(\ln\frac{1}{\varepsilon} \cdot \frac{1}{\delta^2}\right)$$

נשים לב ש- $\ln\frac{1}{\varepsilon}$ הוא יחסית זניח כאן, והערך המשמעותי הוא $\frac{1}{\delta^2}$.
לכן אם נרצה דיוק של 1% נצטרך לשאול 100,000 אנשים (אולי עם פקטור 2). נשים לב שהחישוב אינו תלוי בגודל האוכלוסיה, לכן גם בישראל עם 7 מיליון אנשים וגם בארה"ב עם כמה מאות מ' בארה"ב, עדיין צריך 100,000 אנשים כדי לקבל אותו דיוק.

שאלה מסבירה שהעניין הזה לא רק חשוב לחיים, אלא גם חשוב לסיבוכיות!

אחרי שכיסינו את החלק הזה שלא באמת היה קשור לשיעור (מה היה אמור לעשות אותו), נתחיל בנושא השיעור - אלגוריתמים רנדומיים.

ניזכר בבעיית SetCover. הקלט הוא $s_i \subseteq \{1,2,\dots,n\}$, $C = \{s_1, s_2, \dots, s_k\}$.
הפלט הוא $S \subseteq C,\ S = \{s_{i_1}, s_{i_2}, \dots, s_{i_e}\}$

$$\text{s.t. } \forall j \in \{1,\dots,n\}\ \exists s_i \in S .\ j \in s$$

נרצה למצוא כיסוי <u>מינימלי</u>. בעיה זו כידוע היא NP-שלמה.

נראה את הבעיה כבעיית Integer Programming (כפי שעשינו גם ל-VertexCover).
נגדיר משתנים $x_1, x_2, \dots, x_k$ כך שלכל s_i מתאים x_i כך ש- $x_i = 1 \leftrightarrow s_i \in S$, $x_i = 0 \leftrightarrow s_i \notin S$. התוכנית שנרצה להביא לאופטימיזציה יהיה $\min \sum_{i=1}^{k} x_i$ (כלומר שכמה שפחות קבוצות יהיו שייכות ל-S). האילוצים שלנו יהיו אלה:

$\forall x: \quad 0 \le x_i \le 1$ (כל המשתנים הם אפס או אחד)

$\forall j \in \{1,2,\dots,n\} \quad \sum_{\forall i:\, j \in S_i} x_i \ge 1$ (כל איבר בעולם שלנו שייך לקבוצה כלשהי)

אבל כידוע, IP הוא בעיה NP-שלמה, ולכן אנחנו לא יודעים לפתור אותה.
נעשה את אותו תרגיל שעשינו עבור VC, נתרגם את הבעיה לבעיית Linear Programming. אך כידוע, כעת הפתרון של LP אינו בדיוק הפתרון שאנחנו מחפשים, כי איננו יכולים לבחור $\frac{2}{3}$ קבוצה וכדומה.

עבור VC בפעם שעברה עיגלנו (Rounding), אם המשתנה קטן מחצי נתנו לו אפס, ואם הוא היה גדול שווה חצי נתנו לו אחד. הפעם זה לא יעבוד, וניאלץ להיות יותר מתוחכמים. נשתמש בשיטה שנקראת Randomized Rounding (RR).

נניח שפתרנו את בעיית ה-LP וקיבלנו פתרון $(x_1^*, x_2^*, \dots, x_k^*)$, $0 \le x_i^* \le 1$. אז $OPT_{LP} = \sum_{i=1}^{k} x_i^* \le minSetCover$

(אי השוויון מתקיים בגלל העובדה ש-LP הוא רלקסציה של IP)

נעבור לשלב העיגול הרנדומי:

ניקח את S_1 בהסתברות x_1^*
ניקח את S_2 בהסתברות x_2^*
$\vdots$
ניקח את S_k בהסתברות x_k^*

נסמן את כל הקבוצות שבחרנו ב- A_1.

מה התוחלת של גודל A_1?

$$E[|A_1|] = Pr[take\ S_1] + \dots + Pr[take\ S_k] = \sum_{i=1}^{k} x_i^* = OPT_{LP}$$

כלומר בתוחלת הקבוצה גדלו קטנה מהאופטימום של SetCover, אך לא בטוח שבאמת כל האיברים באמת שייכים לקבוצה. נחשב את הסיכוי שאיבר j לא יהיה מכוסה:

$j \in S_1, S_2, \dots, S_m$

$$Pr[j \text{ not covered}] = (1-x_1^*)(1-x_2^*)\cdots(1-x_m^*) \le e^{-x_1^*} e^{-x_2^*} e^{-x_3^*} \cdots e^{-x_m^*}$$ (כי $1-x \le e^{-x}$)

$$= e^{-\sum_{i=1}^{m} x_i^*} \le e^{-1} = \frac{1}{e}$$

כי לפי תנאי LP $\sum_{i=1}^{m} x_i^* \ge 1$

קיבלנו שהסיכוי ש-j לא שייך ל-A_1 הוא $\frac{1}{e}$, עד לא אק. Pr
יש סיכוי קבוע של איבר לא נמצא, אם אין הרבה איברים נקבל

סיכויים מאוד קטנים לה-B-ה (בכל לא מעל $\frac{2}{3}$).

נריץ את האלגוריתם הרנדומי $d\cdot\log n$ פעמים ונקבל $A_1, A_2, \ldots, A_{d\log n}$.

נגדיר $A = A_1 \cup A_2 \cup \ldots \cup A_{d\log n}$. כעת:

$$E[|A|] \le d\cdot\log n\cdot OPT_{LP} \le d\cdot\log n\cdot \min SetCover$$

מכיוון שאיחדנו $d\cdot\log n$ קבוצות שתוחלת גודלן הוא OPT_{LP}, תוחלת גודל האיחוד ביניהן הוא לכל היותר סכום תוחלות גודליהן.

נחשב כעת מה הסיכוי שלא כיסינו את האיבר j (ונחשב נגיד מהו ה-d שאנחנו צריכים):

$$Pr[j \text{ not covered by } A] = Pr[j \text{ not covered by } A_1 \text{ and } j \text{ not covered by } A_2 \text{ and} \ldots]$$
$$\le \left(\frac{1}{e}\right)^{d\cdot\log n}$$

לכן נבחר d כך ש- $\left(\frac{1}{e}\right)^{d\log n} \le \frac{1}{4n}$ *

* כשנשאלה למה דווקא $4n$ היא אמרה שהיא נראה אוהבת את המספר 4. אפשר גם לבחור $3n$ עם גם יעבוד (גילן תראו למה) אבל לא נתווכח.

כעת נחשב את הסיכוי שיש איבר j שלא מכוסה ע"י A:

$$Pr[\exists j \in \{1,\ldots,n\} \text{ not covered by } A]$$
לפי Union Bound
$$\le Pr[1 \text{ not covered by } A] + Pr[2 \text{ not covered by } A] + Pr\ldots$$
$$\le \frac{1}{4}$$

לכן הסיכוי שהכיסוי שמצאנו אינו באמת כיסוי קטן מרבע (אז אנו בוחרים n הינו מקבלים שהסיכוי לטעות קטן מ-$\frac{1}{3}$, שזה גם יסתדר לפי הגדרת BPP). קיבלנו אלגוריתם הסתברותי שהוא $\log n$-קירוב ל-SetCover.

נעבור לבעיית <u>SemiDefinite Programming - SDP</u>

הבעיה דומה לבעיית LP, רק קצת יותר קשה. ניזכר במבנה של LP:

ערך לאופטימיזציה
$$\max \sum_{i=1}^{n} A_i x_i$$
אילוצים
$$\text{s.t.} \quad \sum_{i=1}^{n} B_i^1 x_i \ge b^1$$
$$\sum B_i^2 x_i \ge b^2$$
$$\vdots$$
אי שליליות
$$x_1 \ge 0, x_2 \ge 0 \ldots$$
משתנים
$$x_1, x_2, \ldots, x_n$$

בעיית SDP מוגדרת כך:

בכתיב מטריצי

$$\max \sum_{i,j=1}^{n} A_{i,j} v_i \cdot v_j$$

$$\text{s.t.} \quad \sum_{i,j=1}^{n} B^1_{i,j} v_i v_j \geq b^1$$

$$\sum_{i,j=1}^{n} B^2_{i,j} v_i \cdot v_j \geq b^2$$

$$\vdots$$

משתנים $v_1, v_2, \dots, v_n \in \mathbb{R}^n$ *

תזכורת

$$v \cdot u = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \sum_{i=1}^{n} v_i \cdot u_i$$

*נשים לב שמספר הוקטורים ומימד המרחב זהה. זה לא במקרה, כך דרוש של SDP

כאשר כל המשתנים שלנו הפכו לוקטורים, ואנחנו מרשים לעצמנו לא רק להשתמש בצירופים לינאריים שלהם אלא במכפלות פנימיות בין זוגות שלהם.

ישנם אלגוריתמים יעילים אשר פותרים בעיה זו (כמו למשל האלגוריתם האליפסואידי) אבל הם מסובכים. נניח שהם קיימים ונפתור בעיה אחרת בעזרת בעיה זו.

ניזכר כי $\|v\| = \sqrt{v \cdot v}$. לכן $\|v_1\| = 1 \Leftrightarrow v_1 \cdot v_1 = 1$.

לכן ניתן באמצעות האילוצים לדרוש $\|v_i\| = 1$ לכל v_i.

נפתור את בעיית MAX-CUT.

בעיית MAX-CUT מקבלת גרף לא מכוון, והמטרה למצוא $S \subseteq V$ כך ש- $|E(S, \bar{S})|$ הוא מקסימלי. בעיה זו היא כזכור, NP-שלמה.

נמצא אלגוריתם שהוא 0.878-קירוב לבעיה*.

*0.878 נראה כמו מספר מוזר ולא קשור, ולא ברור כמה איך נגיע אליו. אבל בעצם, החוכמה הנ יפהפייה ועד סוף הסיכום יהיה מאוד ברור מה המשמעות מאחורי 0.878. אתם מוזמנים לנסות לחשוב עכשיו מה ...0.878 יכול לבוא

נגדיר $x_1, x_2, \dots, x_n \in \{-1, 1\}$, $x_i = 1 \Leftrightarrow i \in S$, $x_i = -1 \Leftrightarrow i \notin S$ ($i \in \bar{S}$).

נשים לב כי $x_i x_j \in \{-1, 1\}$. כאשר $x_i x_j = 1 \Leftrightarrow x_i, x_j \in S$ או $x_i, x_j \in \bar{S}$

$x_i x_j = -1 \Leftrightarrow x_i \in S, x_j \in \bar{S}$ או $x_i \in \bar{S}, x_j \in S$

נרצה ~~...~~ למצוא ביטוי שעבור קשת בחתך יהיה שווה 1 ועבור קשת לא בחתך יהיה שווה 0. נשתמש ב- $\frac{1 - x_i x_j}{2}$

לכן אנחנו בעצם מחפשים

$$\max \sum_{(i,j) \in E} \frac{1 - x_i \cdot x_j}{2}$$

$$= \frac{1}{2}|E| - \min \sum_{(i,j) \in E} \frac{x_i \cdot x_j}{2}$$

קיבלנו ביטוי ואילוצים עבור בעיית IP. אך כידוע IP היא

NP-קשה (ובכל, הבעיה עדין שקולה ל-MAXCUT ולכן עדין NP-קשה).

נתרגם הבעיה במקום ל-LP, ל-SDP.

נשים לב ש~~[illegible]~~ שכשדרשנו $x_i \in \{-1,1\}$ למעשה דרשנו $|x_i|=1$.

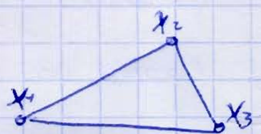
נתרגם:

$$\max \sum_{(i,j)\in E} \frac{1-v_i \cdot v_j}{2}$$

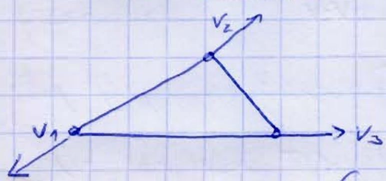
$$\text{s.t.} \quad \|v_i\| = 1 \quad \forall i \in V$$

כעת הפתרון של מערך (של וקטורים) לוקטור ממימד n. את הבעיה הזו אנחנו יכולים לפתור, אבל המטרה שלנו לפתור את MAXCUT, אז צריך להשתמש בבעיה החדשה שביכולתנו לפתור על מנת לפתור את הבעיה המקורית.

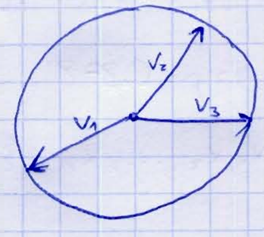
נניח היה לנו במקור את הגרף הבא:



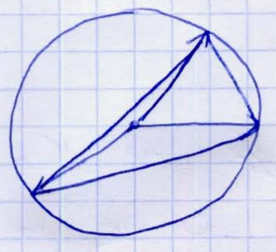
התאמנו לכל קודקוד וקטור כך:



כעת, נשים לב שדרשנו שכל הוקטורים יהיו על ספירת היחידה ($\|v_i\|=1$), ולכן הם נראים כך (לאחר הצבה לדו-מימד)

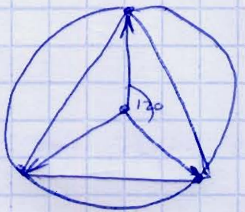


נצייר את המשולש מחדש, בין הקודקודים החדשים הנוצרים על ידי הוקטורים:



אנחנו רוצים להביא למקסימום את $\sum_{(i,j)\in E}\frac{1-v_i\cdot v_j}{2}$ כלומר להביא למינימום את $v_i\cdot v_j$. כלומר נרצה שכל שני וקטורים יהיו רחוקים זה מזה ככל האפשר (אם הם היו הפוכים היה הכי טוב).
אך אם יש יותר משני וקטורים, אי אפשר שכולם יהיו הפוכים

מכולו. לא נוכיח כאן אך זה די אינטואיטיבי שעבור שלושה קודקודים הסידור הכי טוב הוא משולש שווה צלעות:



ואז $v_1 \cdot v_2 = \cos 120 = -\frac{1}{2}$. נקבל ש-

$$\max\left(\frac{1-v_1v_2+1-v_2v_3+1-v_1v_3}{2}\right)=\frac{3-\left(-\frac{3}{2}\right)}{2}$$
$$=\frac{9}{4}>2$$

כלומר קיבלנו פתרון שהוא טוב יותר מהפתרון האופטימלי של MAX-CUT ונבין היגיון כי הבעיה אותה אנו פותרים הוא רלקסציה של הבעיה המקורית. במקום לבחור "וקטורים ממימד אחד" (דרישה של הרכיבים להיות מהראשון הם ±1) בחרנו וקטור כללי ב-$\mathbb{R}^n$.

כדי לעבור מהווקטורים שקיבלנו מ-SDP, נחתוך את הספירה באמצעות היפר מישור (במקרה התלת מימדי זה פשוט מישור, אחרת זה כמו מישור רק ממימד יותר גבוה) אשר עובר במרכז הספירה. כל מה שנמצא בצד אחד של ההיפר מישור יהיה S, והצד השני $\bar{S}$.

פורמלית:

1. solve SDP and get $v_1, \ldots, v_n \in \mathbb{R}^n$
2. randomly choose $\|\vec{n}\| = 1$ — וקטור אקראי מגדיר לנו מישור אקראי (נורמל)
3. define $S = \{ i \mid v_i \cdot n \geq 0 \}$ — שקול לכך ש-v_i נמצא ב"חיובי" של המישור

כעת נוכיח ש- $E[|E(S,\bar{S})|] \geq 0.878^{*}\ldots$ MAX-CUT, = התוחלת של גודל החתך.

* מה שאנחנו מקרבים לצורך של מהו 0.878... וכן הוא בא לידי ביטוי בהוכחה שלנו. תזכרו את הרעיון הגה"ה שלכם מספר כי למה לא מתכלא ש-0.878... כפי שזה מספר אקראי שנה...

נניח כי הזווית בין v_i ל-v_j היא θ.

אז
$$\Pr[(i,j) \text{ is cut}] = \frac{\theta}{\pi}$$

לכן
$$E[|E(S,\bar{S})|] = \sum_{(i,j)\in E} \Pr[(i,j) \text{ is cut}]$$

<u>טענה:</u>
$$\Pr[(i,j) \text{ is cut}] \geq 0.878 \cdot \frac{1-v_1v_2}{2}$$

אם נוכיח את הטענה נקבל:

$$E[|E(S,\bar{S})|] = \sum_{(i,j)\in E} \Pr[(i,j) \text{ is cut}] \geq 0.878 \sum_{(i,j)\in E} \frac{1-v_iv_j}{2} = 0.878 \cdot OPT_{SDP}$$
$$\geq 0.878 \cdot \text{MAX-CUT}$$

לכן נוכיח את הטענה: נרצה להראות

$$\Pr[\cancel{e}\ (i,j) \text{ is cut}] = \frac{\theta}{\pi} \geq 0.878\,\frac{1-v_1v_2}{2} = \frac{1-\cos\theta}{2}\cdot 0.878$$

$$\frac{\theta}{\pi} \geq 0.878\,\frac{1-\cos\theta}{2}$$

המשוואה מתקיימת עבור $\theta=0$, עבור המקרה בו $\theta\neq 0$:

$$\frac{2}{\pi}\frac{\theta}{1-\cos\theta} \geq 0.878...$$

אז למעשה נרצה להוכיח שהמשוואה הזו מתקיימת לכל θ. ~~x~~

בעצם, לא הצלחתי למצוא לזה פתרון אנליטי, פשוט הרצתי את הנוסחה $\frac{2}{\pi}\frac{\theta}{1-\cos\theta}$ במחשב על כמה ה-θ-ות האפשריות וקיבלתי ש- 0.878... זה המינימום.*

פסקת הסברה: הבנו שמגרילים וקטור אקראי עם נורמה 1 כאילו זה דבר טריוויאלי, אבל זה לא באמת. אם נשתמש בשיטה הנאיבית (התפלגות אחידה ~~x~~ בין -1 ל-1 של כל הרכיבים ואז נירמול), לא באמת נקבל התפלגות שתהיה של וקטורי היחידה. למעשה נקבל התפלגות אחידה על ריבוע היחידה, וזה יגרום לוקטורים שרחוקים מהצירים להיות נפוצים יותר מוקטורים קרובים לצירים. על מנת לקבל התפלגות אחידה של וקטורי מעגל היחידה צריך (סתם לידע כללי) לבחור התפלגות נורמלית בין -1 ל-1 של כל רכיבי הוקטור (באופן ב"ת כמובן).

* מזל טוב! כרגע עברתם את המכשלה הגדולה ביותר בקורס שלכם (אולי חוץ מפרסום הציונים בבדידה...). אם עד עכשיו חשבתם שהוכחות מתמטיות במחשב זה דבר מגניב, עכשיו אתם יודעים שזה דבר מבאס. ואם חשבתם שאני מסכם (ודאי שמבאס אתכם יותר מדי) אתם הולכים להחרים מעכשיו את כל הסיכומים שלי, אז גם טוב, כי זה סוף הסיכום האחרון (בשבילכם השנה, בשבילי בתואר!) אתם מוזמנים אם אתם רוצים להחרים את הפסקה האחרונה היא של כבר לא למבחן.

מקווה שנהניתם מהסיכומים, ותזכרו - סיכום אחד ביד עדיף משניים של האינטרנט שצריך להתמקד עם pdf-ים וסריקות וכל זה, במיוחד אם זה סיכומים במדמ"ח, איכסה.

בהצלחה במבחן!