נושאים מתקדמים באלגוריתמים
הרצאה 1

6 בפברואר 2021

מתן ירושלמי

תוכן העניינים

חלק I

הרצאה 7

1 משפטים על ספקטרום של גרפים

1.1 משפט

יהי $G=(E,V)$ גרף לא מכוון. אזי G רגולרי $\Leftarrow$ מטריצת השכנויות $A=A(G)$ בעלת פירוק ספקטרלי מלא.

הוכחה מהיותו לא מכוון G סימטרי, והערכים במטריצה הם 1 או 0, ובפרט ממשיים. לכן לפי משפט הפירוק הספקטרלי מתקיים הנדרש.

1.2 משפט

$G=(E,V)$ גרף (n,d) רגולרי $\iff$ $\lambda_1(G)=d$, כאשר λ_1 הינו הערך העצמי הגדול ביותר בפירוק הספקטרלי של מטריצת השכנויות $A(G)$.

הוכחה "$\Leftarrow$": יהי $\vec{v}$ וקטור שכל הקורדינטות בו שוות לקבוע כלשהו $c\in\mathbb{R}$. לכל כניסה במטריצה יש בדיוק d יציאות שיש בהן 1. כלומר עבור השורה ה-i-

$$[A\vec{v}]_i=\sum_{j=1}^{n(=|V|)}a_{ij}v_j=\sum_{j=1}^{n(=|V|)}a_{ij}c=dc\,.$$

לפיכך נקבל

$$A\vec{v} = d\vec{v}$$

" $\Rightarrow$ ": טענה 2.10 בספר

1.3 משפט

יהי $G = (E, V)$ גרף (n, d) רגולרי. אזי G קשיר $\lambda_2(G) < d \iff$.

הוכחה " $\Rightarrow$ ": נראה כי אם G לא קשיר, אז $\lambda_2(G) = d$. ראשית לפי המשפט הקודם ידוע כי $\lambda_1 = d$, כלומר d הינו הערך העצמי הגדול ביותר של G. נראה כי $\lambda_2 = \lambda_1$. G לא קשיר, לכן קיימים לפחות שני רכיבי קשירות, נקרא להםS_1, S_2, בגרף. נסדר את הקודקודים במטריצת השכנויות באופן הבא:

$$\begin{aligned} S_1 &= \left\{v_1, \ldots, v_{|S_1|}\right\} \\ S_2 &= \left\{v_{|S_1|+1}, \ldots, v_{|S_1 \cup S_2|}\right\} \\ V \backslash (S_1, S_2) &= \left\{v_{|S_1 \cup S_2|+1}, \ldots, v_n\right\} \end{aligned}$$

כמובן שאם יש שני רכיבי קשירות בלבד אז הקבוצה בשורה האחרונה תהיה ריקה. אז מטריצת השכנויות תהיה מטריצת בלוקים מהצורה:

$$\left[\begin{array}{c|c|c} S_1 & & \\ \ddots & & \\ \ddots & & \\ \hline & S_2 & \\ & \ddots & \\ \hline & & ? \\ & & \ddots \\ & & ? \end{array}\right]$$

כאשר הבלוק השמאלי העליון הוא בגודל $|S_1| \times |S_1|$, הבלוק האמצעי בגודל $|S_2| \times |S_2|$, והבלוק האחרון מגודל $n - (|S_1| + |S_2|)$. ניווכח שהווקטורים u, w המוגדרים על ידי

$$u_i = \begin{cases} 1 & i \leq |S_1| \\ 0 & \text{else} \end{cases}$$
$$w_i = \begin{cases} 1 & |S_1| < i \leq |S_2| \\ 0 & \text{else} \end{cases}$$

הם וקטורים עצמיים עם ערך d. זאת משיקולים קודמים להוכחה הקודמת (ניתן להכפיל ולראות זאת). אז $\lambda_1 = \lambda_2 = d$ ובנוסף $u \perp w$ כי בכל מקום שיש 1 באחד מהם, יש 0 בשני.
"$\Leftarrow$": הוכחה בספר.

1.4 משפט

יהי $G = (E, V)$ גרף (n, d) רגולרי. אזי $\lambda_n(G) = -d \iff G$ דו-צדדי.

הוכחה בספר.

1.5 משפט

יהי $G = (E, V)$ גרף (n, d) רגולרי. אזי G דו-צדדי $\iff$ הספקטרום של G סימטרי (כלומר $\lambda_i = \lambda_{n+1-i}$ לכל $1 \leq i \leq n$).

הוכחה בספר.

2 משפט ההרחבה ומשפט הערבוב

שני המשפטים מתייחסים לגרפים (n, d) רגולריים וכך גם הפרמטרים בהם.

2.1 משפט ההרחבה

יהי $G = (E, V)$ גרף (n, d) רגולרי. אזי לכל תת-קבוצה $S \subseteq V$ מתקיים

$$|E(S, V \backslash S)| \geq (d - \lambda_2)|S| \frac{n - |S|}{n}.$$

2.1.1 ניסוח שקול

$$\left|E\left(S,V\backslash S\right)\right|-d\left|S\right|\frac{n-\left|S\right|}{n}\geq-\lambda_{2}\left|S\right|\frac{n-\left|S\right|}{n}.$$

2.1.2 דוגמאות להמחשה

1. במקרה שהגרף לא קשיר $\lambda_2=d$ ולכן כל מה שהמשפט אומר הוא שמכל תת-קבוצה יוצאות לפחות אפס צלעות.
2. במקרה ש-λ_2 קרוב מאוד ל-λ_1 נקבל יחס הרחבה מאוד נמוך מאוד ובמקרה שהם רחוקים נקבל יחס הרחבה גבוה.

2.2 משפט (=למת) הערבוב

יהי $G=\left(E,V\right)$ גרף $\left(n,d\right)$ רגולרי. נסמן $\lambda=\max\left\{ \left|\lambda_{2}\right|,\ldots,\left|\lambda_{n}\right|\right\}$. אזי לכל $S,T\subseteq V$ מתקיים $\left|E\left(S,T\right)\right|=d\left|S\right|\frac{\left|T\right|}{n}\pm\lambda\sqrt{\left|S\right|\left|T\right|}$.

2.2.1 ניסוח שקול למשפט הערבוב

$$\left|\left|E\left(S,T\right)\right|-d\left|S\right|\frac{\left|T\right|}{n}\right|\leq\lambda\sqrt{\left|S\right|\left|T\right|}$$

הגדרה פער הספקטרלי הינו ההפרש בין $d=\lambda_1$ אל בין $\lambda=\max\left\{ \left|\lambda_{2}\right|,\ldots,\left|\lambda_{n}\right|\right\}$. כלומר $\lambda_1-\lambda$.

הערה נבחין כי $\lambda=\max\left\{ \lambda_{2},-\lambda_{n}\right\}$ משום שהערכים בספקטרום ממויינים, אבל יותר נוח להסתכל על כל האחרים.

הערה שימו לב ש-S,T לאו דווקא זרות.

אינטואיציה המשפט אומר שלכל שתי קבוצות S,T, מספר הקשתות שעוברות מ-S ל-T הוא פחות או יותר מה שאנחנו מצפים ($d\left|S\right|\frac{\left|T\right|}{n}$), עם שגיאה מסויימת $\lambda\sqrt{\left|S\right|\left|T\right|}$. המשפט אומר שאם הפער הספקטרלי מספיק גדול, אז הגרף שלנו "יתערבב" יפה, כלומר לא יהיו שינויים גדולים בדלילות הגרף בחלקים שונים שלו. זאת משום שהשגיאה קטנה יותר ככל ש-λ קטן.

הבחנה נבחין כי משפט ההרחבה הוא מקרה פרטי של משפט הערבוב שבו $S=S$, $T=V\backslash S$. בנוסף נבחין כי מקרה זה מעניין במיוחד משום ש-$\partial\left(S\right)=E\left(S,V\backslash S\right)$ וזה מתקשר לנו ליחס הרחבה.

*נציין שבדרך כלל קוראים לו "למת הערבוב לגרפים מרחיבים" או באנגלית Expander Mixing Lemma.

2.3 תזכורות - אלגברה לינארית

הגדרה וקטור אינדיקטור במקרה שלנו הינו וקטור אשר מקבל 1 לכל קודקוד אשר שייך לקבוצה.

תזכורת

$$\vec{x}^t A\vec{y} = \sum_{1\leq i,j\leq n} x_i A_{i,j} y_j$$

תזכורת נורמה סטנדרטית של וקטור $v \in \mathbb{R}^n$ היא

$$\sqrt{\langle v \mid v\rangle} = \sqrt{\sum_{i=1}^{n} v_i^2}$$

תזכורת נרמול של וקטור - שינוי האורך של הוקטור ל-1 על ידי חילוק כל קורדינטה בנורמה.

תזכורת אורתוגונליות - בעברית ניצבות, היא יחס בין שני וקטורים. באופן גרפי המשמעות היא שהוקטורים ניצבים אחד לשני (90 מעלות). באופן אלגברי המשמעות היא שהמכפלה הפנימית שלהם היא אפס. כלומר, בהנתן מרחב וקטורי V, $u,v \in V$, נסמן $\langle u \mid v\rangle = 0 \iff u \perp v$.

תזכורת בסיס אורתונורמלי - בבסיס אורתונורמלי כל שני וקטורים בבסיס ניצבים אחד לשני, והנורמה של כל וקטור היא אחד. זה מאוד נוח כי עבור בסיס אורתונורמלי $v_1,\dots,v_n$ של תת-מרחב V ממימד n נקבל

$$\langle v_i \mid v_j\rangle = \begin{cases} 1 & i=j \\ 0 & i\neq j \end{cases}.$$

תזכורת בהנתן בסיס $v_1,\dots,v_n$ של תת-מרחב V, לכל $u\in V$ מתקיים $u=\sum_{i=1}^{n}\alpha_i v_i$. אם הבסיס אורתונורמלי, אז $\alpha_i = \langle u \mid v_i\rangle$. בנוסף $|u| = \sqrt{\sum (\alpha_j)^2}$.

משפט אי-שוויון קושי-שוורץ - יהיו u, v וקטורים מעל מרחב מכפלה פנימית, אז

$$|\langle u \mid v\rangle|^2 \leq \langle u \mid u\rangle \cdot \langle v \mid v\rangle$$

כלומר

$$|\langle u \mid v\rangle| \leq \sqrt{\langle u \mid u\rangle}\cdot\sqrt{\langle v \mid v\rangle}$$

2.4 החצי הראשון של ההוכחה של שני המשפטים

ההוכחה של שני שני המשפטים מתחילה אותו דבר ואז מתפצלת לשניים.

הוכחה יהי $G=(E,V)$ גרף (n,d) רגולרי ותהיינה $S,T\subseteq V$. נסמן ב-$\vec{1}_S,\vec{1}_T$ את וקטורי האינדיקטורים של S ושל T. אז

$$\begin{aligned}\vec{1}_S^tA\vec{1}_T &\underset{\text{תזכורת}}{=}\sum_{u,v\in V}\left(\vec{1}_S\right)_uA_{u,v}\left(\vec{1}_T\right)_v\\ &=\sum_{u,v\in V}\begin{cases}1 & u\in S\wedge(u,v)\in E\wedge v\in T\\ 0 & \text{else}\end{cases}\\ &=\left|E\left(S,T\right)\right|\end{aligned}$$

לפי המשפטים מ-1 קיים בסיס אורתונורמלי של וקטורים עצמיים $\vec{v}_1,\ldots,\vec{v}_n$ עם ערכים עצמיים $d=\lambda_1\geq\cdots\geq\lambda_n$. עבור λ_1 נבחר את הוקטור $\vec{1}$, ואחרי נרמול (ראו תזכורת על נורמה של וקטור לעיל) הוא יהפוך להיות $\frac{1}{\sqrt{n}}\cdot\vec{1}$.

על-כן קיימים סקלרים $\alpha_1,\ldots,\alpha_n$ ו-$\beta_1,\ldots,\beta_n$ כך שמתקיים

$$\begin{aligned}\vec{1}_S&=\sum_{i=1}^n\alpha_iv_i\\ \vec{1}_T&=\sum_{i=1}^n\beta_jv_j\end{aligned}$$

אז

$$
\begin{aligned}
|E(S,T)| = \vec{1}_S^t A \vec{1}_T &= \left(\sum_{i=1}^{n} \alpha_i v_i\right) A \left(\sum_{j=1}^{n} \beta_j v_j\right) \\
\text{Linearity} &= \left(\sum_{i=1}^{n} \alpha_i v_i\right) \sum_{j=1}^{n} \beta_j A v_j \\
&= \left(\sum_{i=1}^{n} \alpha_i v_i\right) \sum_{j=1}^{n} \beta_j \lambda_j v_j \\
\text{Linearity} &= \sum_{1 \le i,j \le n} \alpha_i \beta_i \lambda_j \langle v_i \mid v_j \rangle \\
\text{Orthonormal Basis} &= \sum_{j=1}^{n} \alpha_j \beta_j \lambda_j \\
&= \alpha_1 \beta_1 \lambda_1 + \sum_{j=2}^{n} \alpha_j \beta_j \lambda_j \\
&= \left\langle \vec{1}_S \mid \frac{1}{\sqrt{n}} \cdot \vec{1} \right\rangle \cdot \left\langle \vec{1}_T \mid \frac{1}{\sqrt{n}} \cdot \vec{1} \right\rangle \cdot d + \sum_{j=2}^{n} \alpha_j \beta_j \lambda_j \\
\text{indicators} &= \frac{|S|}{\sqrt{n}} \cdot \frac{|T|}{\sqrt{n}} \cdot d + \sum_{j=2}^{n} \alpha_j \beta_j \lambda_j \\
&= d\,|S|\,\frac{|T|}{n} + \sum_{j=2}^{n} \alpha_j \beta_j \lambda_j
\end{aligned}
$$

2.4.1 סיום הוכחת משפט ההרחבה

נותר להוכיח

$$\sum_{j=2}^{n} \alpha_j \beta_j \lambda_j \ge -\lambda_2 \,|S|\, \frac{n-|S|}{n}$$

עבור $T = V \setminus S$.

נבחין כי משום ש-$T = V \setminus S$ אז $\vec{1}_T = \vec{1} - \vec{1}_S$ $(*)$. אז

$$\begin{aligned}
\sum_{j=2}^{n} \alpha_j \beta_j \lambda_j \underset{\text{orthonormality}}{=} & \sum_{j=2}^{n} \alpha_j \left\langle \vec{1}_T \mid v_j \right\rangle \lambda_j \\
(*) = & \sum_{j=2}^{n} \lambda_j \alpha_j \left\langle \vec{1} - \vec{1}_S \mid v_j \right\rangle \\
= & \sum_{j=2}^{n} \lambda_j \alpha_j \left[\left\langle \vec{1} \mid v_j \right\rangle - \left\langle \vec{1}_S \mid v_j \right\rangle \right] \\
(**) = & \sum_{j=2}^{n} \lambda_j \alpha_j \cdot (-\alpha_j) \\
\geq & -\max \{\lambda_2, \ldots, \lambda_n\} \sum_{j=2}^{n} \alpha_j^2 \\
\text{sorted} = & -\lambda_2 \sum_{j=2}^{n} \alpha_j^2 \\
= & -\lambda_2 \left[\left(\sum_{j=1}^{n} \alpha_j^2 \right) - \alpha_1^2 \right] \\
= & -\lambda_2 \left[\left(\sum_{j=1}^{n} \alpha_j^2 \right) - \left(\left\langle \vec{1}_S \mid \vec{v}_1 \right\rangle \right)^2 \right] = -\lambda_2 \left[\left(\sum_{j=1}^{n} \alpha_j^2 \right) - \left(\left\langle \vec{1}_S \mid \frac{1}{\sqrt{n}} \vec{1} \right\rangle \right)^2 \right] \\
\text{indicator} = & -\lambda_2 \left[\left(\sum_{j=1}^{n} \alpha_j^2 \right) - \frac{|S|^2}{n} \right] \\
\text{תזכורת אחרונה} = & -\lambda_2 \left[\left\| \vec{1}_S \right\|^2 - \frac{|S|^2}{n} \right] \\
= & -\lambda_2 \left[\left\langle \vec{1}_S \mid \vec{1}_S \right\rangle - \frac{|S|^2}{n} \right] = -\lambda_2 \left[|S| - \frac{|S|^2}{n} \right] \\
= & -\lambda_2 \left[\frac{|S|}{n} \cdot (n - |S|) \right]
\end{aligned}$$

כנדרש.■

$(**)$ מתקיים כי ידוע ש-$v_1 = \frac{1}{\sqrt{n}} \vec{1}$ הינו וקטור עצמי של λ_1, ובנוסף ידוע שלכל $i \neq 1$ מתקיים $\lambda_i \neq \lambda_1$. באלגברה לינארית יש משפט שאומר שוקטורים עצמיים ששייכים לערכים עצמיים שונים זה מזה אורתוגונליים ולכן $\vec{1} \perp v_i$ לכל $2 \leq i \leq n$.

2.4.2 סיום הוכחת למת הערבוב

נותר להוכיח

$$\left|\sum_{j=2}^{n}\alpha_j\beta_j\lambda_j\right|\leq\lambda\sqrt{|S|\,|T|}.$$

נסמן לשם הנוחות $\vec{\beta}:=\begin{bmatrix}|\beta_2| & \dots & |\beta_n|\end{bmatrix}^t$, $\vec{\alpha}:=\begin{bmatrix}|\alpha_2| & \dots & |\alpha_n|\end{bmatrix}^t$ כלומר וקטורי המקדמים בערך מוחלט.

$$\begin{aligned}
\left|\sum_{j=2}^{n}\alpha_j\beta_j\lambda_j\right| &\underset{\triangle\text{ inequality}}{\leq}\left(\sum_{j=2}^{n}|\alpha_j\beta_j\lambda_j|\right)\\
&=\sum_{j=2}^{n}|\alpha_j|\,|\beta_j|\,|\lambda_j|\\
&\leq\sum_{j=2}^{n}|\alpha_j|\,|\beta_j|\cdot\max\{|\lambda_2|,\dots,|\lambda_n|\}\\
&\overset{\text{def}}{=}\lambda\sum_{j=2}^{n}|\alpha_j|\,|\beta_j|\\
&=\lambda\left\langle\vec{\alpha}\mid\vec{\beta}\right\rangle\underset{(*)}{=}\lambda\left|\left\langle\vec{\alpha}\mid\vec{\beta}\right\rangle\right|\\
\text{schwarz cauchy}\qquad &\leq\lambda\sqrt{\langle\vec{\alpha}\mid\vec{\alpha}\rangle}\sqrt{\left\langle\vec{\beta}\mid\vec{\beta}\right\rangle}\\
\text{תזכורת אחרונה}\qquad &\underset{(**)}{=}\lambda\left\|\vec{1}_S\right\|\left\|\vec{1}_T\right\|\\
&\overset{\text{def}}{=}\lambda\sqrt{\left\langle\vec{1}_S\mid\vec{1}_S\right\rangle}\sqrt{\left\langle\vec{1}_T\mid\vec{1}_T\right\rangle}\\
\text{indicator}\qquad &=\lambda\sqrt{|S|\,|T|}
\end{aligned}$$

כנדרש. ■

$(*)$ ערך מוחלט של ערכים מוחלטים

$(**)$ $\forall x\in\mathbb{R}\quad\sqrt{x\cdot x}=\sqrt{|x|\cdot|x|}=|x|$

חלק II

הגברת נכונות - Error Reduction

3 רקע

3.1 הכרעה עם שגיאה חד צדדית

תהי L שפה (מעל אלפבית Σ). נאמר ש-A הוא אלגוריתם הכרעה עם שגיאה חד-צדדית אם

$$\forall x \in \Sigma^* \backslash L \quad \text{Pr}\left[A\left(x\right) = \text{Reject}\right] = 1.$$

עבור $x \in L$, A לא חייב לקבל את המילה. כלומר,

$$\forall x \in L \quad \text{Pr}\left[A\left(x\right) \neq \text{Accept}\right] \geq 0.$$

תרגיל חשבו על אלגוריתם יעיל לשגיאה חד-צדדית.

פתרון אלגוריתם אשר מחזיר Reject בכל מקרה. כלומר, בהנתן $x \in \Sigma^*$,

$$\text{Pr}\left[A\left(x\right) = \text{Reject}\right] = \text{Pr}\left[A\left(x\right) \neq \text{Accept}\right] = 1.$$

מסקנה כדי לקבל אלגוריתם מעניין, נרצה שההסתברות שהאלגוריתם ידחה מילה בשפה תהיה קטנה. לשם כך, נוסיף פרמטר δ המוגדר להלן

הגדרה δ מסמן את היתרון של A על פני האלגוריתם טריוויאלי (זה שתמיד דוחה את המילה). נסמן:

$$\text{Pr}\left[A\left(x\right) \neq \text{Accept}\right] \leq 1 - \delta.$$

ככל ש-δ גדול, כך ההסתברות שהאלגוריתם יחזיר תשובה שגויה נמוך.

3.1.1 משפט

יהי $0 < \varepsilon \leq 1$, ויהי A אלגוריתם הכרעה יעיל בשגיאה חד-צדדית עם יתרון δ. אז קיים אלגוריתם מוגבר A^* בשגיאה חד-צדדית עם הסתברות שגיאה $\leq \varepsilon$, כש-A^* מריץ את A $m = \theta\left(\frac{1}{\delta}\ln\frac{1}{\varepsilon}\right)$ פעמים.

אינטואיציה המשפט אומר שאפשר להקטין את השגיאה של האלגוריתם החד-צדדי במספר קטן יחסית של הרצות של A. שימו לב ש-A לא דטרמיניסטי ולכן הוא לאו דווקא יחזיר את אותה התוצאה בכל פעם עבור מילים בשפה.

הבחנה כדי לקבל אלגוריתם A^* יעיל (הגברת נכונות יעילה), אנחנו צריכים δ שהיא לכל הפחות $\frac{1}{f(n)}$ עבור פולינום כלשהו f, ו-$n = \text{length}\,(x)$ לכל קלט x. אחרת, למשל עבור $\delta = \frac{1}{2^n}$ נקבל

$$m = \theta\left(2^n \ln\frac{1}{\varepsilon}\right)$$

ואז האלגוריתם A^* לא יעיל.

3.2 הכרעה עם שגיאה דו-צדדית

תהי L שפה (מעל אלפבית Σ). באופן דומה, נשתמש בפרמטר יתרון δ על-פני האלגוריתם הטריוויאלי שבמקרה זה מגריל Reject או Accept ללא תלות בקלט, וצודק בחצי מהזמן. נגדיר A, אלגוריתם הכרעה עם שגיאה דו-צדדית על-ידי:

$$.\forall x \in \Sigma^* \backslash L \quad \Pr\left[A\left(x\right) \neq \text{Reject}\right] \leq \frac{1}{2} - \delta$$

$$.\forall x \in L \quad \Pr\left[A\left(x\right) \neq \text{Accept}\right] \leq \frac{1}{2} - \delta$$

3.2.1 משפט

יהי $1 \geq \varepsilon > 0$, ויהי A אלגוריתם הכרעה יעיל בשגיאה דו-צדדית עם יתרון δ. אז קיים אלגוריתם מוגבר A^* בשגיאה חד-צדדית עם הסתברות שגיאה $\varepsilon \geq$, כש-A^* מריץ את A $m = \Theta\left(\frac{1}{\delta^2}\ln\frac{1}{\varepsilon}\right)$ פעמים.

הבחנה זה אותו משפט כמו עבור אלגוריתם הכרעה בשגיאה חד-צדדית, עם δ^2 במכנה במקום δ.

4 הגברת נכונות באמצעות גרפים מרחיבים

4.1 הגברת נכונות בשגיאה דו-צדדית (BPP)

הגדרה אקספנדר רמנוג'אן - זהו גרף עם פער ספקטרלי מקסימלי. ניזכר בסימון מתחילת ההרצאה

$$.\lambda = \max\left\{\left|\lambda_2\right|, \ldots, \left|\lambda_n\right|\right\} = \max\left\{\lambda_2, -\lambda_n\right\}$$

אז באקספנדר רמנוג'אן

$$.\lambda = 2\sqrt{d-1}$$

קיימת בנייה יעילה במובן החזק לאקספנדר מסוג זה.

הסבר בנייה יעילה במובן החזק היא בנייה של גרף אקספנדר באופן שמאפשר לנו לגשת לחלקים בגרף בלי לבנות את כולו. למשל, אם כל קודקוד בגרף מסמל רצף אפשרי של הטלות מטבע, ואנחנו רוצים להתבונן במרחב הרצפים באורך $\sqrt{n}$ של הטלות, אז נקבל גרף עם $2^{\sqrt{n}}$ קודקודים. ברור שלא ניתן לבנות גרף כזה באופן יעיל. אך עם בנייה יעילה במובן החזק לכל קודקוד נוכל לתמוך בשאילתא על למי הקודקוד הספציפי מחובר.

מה מנסים לעשות? באלגוריתם נאיבי של הגברת נכונות, בהינתן אלגוריתם יעיל A עם שגיאה דו-צדדית ויתרון δ, נפעיל את A שוב ושוב, נספור את מספר הקבלות והדחיות ונחזיר את המקסימום ביניהם. אנחנו רוצים לייעל את התהליך הזה.

דוגמה שלי להרצת אלגוריתם נאיבי של הגברת נכונות תהי $x \in L$ ואלגוריתם יעיל A עם $\delta = 0.1$, אז אחרי 11 הפעלות (בחרתי מספר אי זוגי בשביל שובר שוויון) של A נקבל

$$\begin{aligned}
\Pr\left[A_{11}^*(x) \neq \text{accept}\right] &= \Pr\left[\text{count}\left(A_{rej}\right) > \text{count}\left(A_{acc}\right)\right] \\
&= \sum_{k=1}^{5} \Pr\left[\text{count}\left(A_{acc}\right) = j\right] \\
&= \sum_{k=1}^{5} \binom{11}{k} \cdot 0.6^k \cdot 0.4^{11-k} \\
&\approx 0.246
\end{aligned}$$

אז הקטנו את ההסתברות לשגיאה ל-0.246. לעומת זאת אחרי 101 הפעלות נוכל לראות שההסתברות לשגיאה יורדת באופן משמעותי כצפוי:

$$\begin{aligned}
\Pr\left[A_{101}^*(x) \neq \text{accept}\right] &= \Pr\left[\text{count}\left(A_{rej}\right) > \text{count}\left(A_{acc}\right)\right] \\
&= \sum_{k=1}^{50} \Pr\left[\text{count}\left(A_{acc}\right) = j\right] \\
&= \sum_{k=1}^{50} \binom{101}{k} \cdot 0.6^k \cdot 0.4^{101-k} \\
&\approx 0.02
\end{aligned}$$

4.1.1 תיאור ההגברה

1. ראשית, נייצר גרף רמנוג'ן (**עם בנייה יעילה במובן החזק**). כל קודקוד בגרף שלנו ייצג תוצאה אפשרית של r הטלות מטבע.
2. מגרילים קודקוד יחיד v בהתפלגות אחידה.
3. A^* מריץ את האלגוריתם A על כל השכנים של v, כל פעם עם רצף הטלות המטבע שמופיע בשכן. (A אלגוריתם שמשתמש באקראיות, על-כן יש לספק לו את סדרת הטלות המטבע הנדגמת באופן אחיד).
4. מחזירים את מה שיצא ברוב הפעמים.

תיאור פורמלי של האלגוריתם מגביר הנכונות:

אלגוריתם 1 הגברת נכונות דו-צדדית באמצעות גרף רמנוג'ן
הגרל $v \in V$ בהתפלגות אחידה.
החזר $A^*(x,v) \equiv \text{Maj}\left\{A(x,u_i) \mid u_i \in \Gamma(v)\right\}$

הערה Maj:=Majority, $\equiv$ הוא סימן השקילות.

4.1.2 ניתוח הסתברות השגיאה של האלגוריתם המוגבר

נסמן ב-B את קבוצת הקודקודים בגרף, אשר הפעלת A (האלגוריתם ה**לא** מוגבר) עליהם (עם מילת הקלט x) מובילה לשגיאה.

באופן אנלוגי, נסמן ב-B^* את קבוצת הקודקודים בגרף, אשר הפעלת A^* (האלגוריתם המגביר) עליהם (עם מילת הקלט x) מובילה לשגיאה.

נסמן $n = |V| = 2^r$. שימו לב! n איננו גודל הקלט x.

הבחנה בהנתן מילת קלט x וסדרת הטלות מטבע $\{u_1, \ldots, u_r\}$, A מחזיר פלט **דטרמיניסטי**, כלומר תמיד נקבל את אותה התוצאה. הוא משתמש באקראיות שסיפקנו לו מבחוץ, ואם נספק לו כל פעם את אותו רצף של הטלות (ביטים) אז הוא יחזיר את אותו פלט.

הבחנה אין שום קשר (יחס כמו חיתוך, הכלה וכולי) בין B ל-B^*.

הבחנה $\frac{|B|}{n} = \frac{1}{2} - \delta$, כי בהנתן כל קלט x ההסתברות לשגיאה היא $\frac{1}{2} - \delta$.

הסתברות השגיאה של A^* היא בדיוק ההסתברות שהקודקוד שבחרנו הוא בתוך B^*. דגמנו את v באופן אחיד ועל כן הסתברות זו שווה בדיוק ל-$\frac{|B^*|}{n}$. כדי לחשב את המספר הזה, נספור כמה צלעות יש עוברות מ-B^* אל B, באופן פורמלי $|E(B^*, B)|$.

לכל $v \in B^*$, נמספר את הקשתות שיוצאות מ-v ב-$E_v := \{e_1, \ldots, e_d\}$. אז לפחות חצי מהקשתות ב-E_v מובילות לקודקוד ש-A שוגה עליו (קודקוד ב-B). נסמן את תת-הקבוצה של קשתות אלה ב-E_{vB}. אז

$$|E(B^*, B)| = \sum_{i=1}^{|B^*|} |E_{v_i B}| \geq \sum_{i=1}^{|B^*|} \frac{d}{2} = \frac{d}{2}|B^*| \, .$$

מצד שני, לפי למת הערבוב מתקיים

$$|E(B^*, B)| \leq d|B^*|\frac{|B|}{n} + \lambda\sqrt{|B^*||B|} \, .$$

על-כן:

$$\begin{aligned}
&|B^*|\frac{d}{2} \leq d|B^*|\frac{|B|}{n} + \lambda\sqrt{|B^*||B|} \\
\Rightarrow & d|B^*|\left(\underbrace{\frac{1}{2} - \frac{|B|}{n}}_{=\delta}\right) \leq \lambda\sqrt{|B^*||B|} \\
\Rightarrow & \left|\sqrt{B^*}\right| \leq \frac{\lambda}{d\delta}\sqrt{B} \\
\Rightarrow & |B^*| \leq \left(\frac{\lambda}{d}\right)^2 \frac{1}{\delta^2}|B|
\end{aligned}$$

ניזכר כי $\lambda = 2\sqrt{d-1}$ ובפרט $\lambda \leq 2\sqrt{d}$. לכן $\frac{\lambda}{d} \leq \frac{2}{\sqrt{d}}$. בנוסף אנו מתעניינים בהסתברות לשגיאה ולכן נחלק את שני האגפים של המשוואה לעיל ב-n ונקבל

$$\frac{|B^*|}{n} \leq \frac{4}{d} \cdot \frac{1}{\delta^2}\frac{|B|}{n}$$

אז בעצם קיבלנו

$$\varepsilon = \Pr[\text{error}(A^*)] \leq \frac{4}{d \cdot \delta^2}\Pr[\text{error}(A)] . \blacksquare$$

4.1.3 השוואה של הגברת נכונות בעזרת אקספנדרים להגברת נכונות סטנדרטית

ראשית נחלץ את d מהמשוואה שקיבלנו בסוף ההוכחה הקודמת. ניזכר כי d הוא בדיוק מספר האיטרציות שהאלגוריתם מבצע.

$$d \leq \frac{4}{\varepsilon\delta^2} \cdot \left(\frac{1}{2} - \delta\right) = \Theta\left(\frac{1}{\varepsilon\delta^2}\right)$$

ניזכר בחסם על האלגוריתם הנאיבי, שקיבלנו בטענה $3.2.1$. נסמן ב-d את מספר האיטרציות גם כאן:

$$d = \Theta\left(\frac{1}{\delta^2}\ln\frac{1}{\varepsilon}\right)$$

אז מבחינת משאב הזמן, האלגוריתם הסטנדרטי הוא המנצח הברור.

כעת נתבונן במשאב האקראיות. האלגוריתם הסטנדרטי דורש $dr = \Theta\left(\frac{1}{\delta^2}\ln\frac{1}{\varepsilon}r\right)$ הטלות מטבע, אך לעומת זאת האלגוריתם שמשתמש באקספנדר דורש רק הגרלה של הקודקוד הראשון (מחרוזת בינארית באורך r) כלומר r הטלות מטבע.■