כלים מתמטיים למדעי המחשב*

סוכם ע"י אביב יעיש

תוכן עניינים

*כפי שהועבר בתשע"ז על ידי המרצה נתי ליניאל והמתרגלים מיכאל סימקין ויובל פלד

חלק I
הסתברות

1 שיעור 1

1.1 בסיס

הגדרה 1.1 מרחב הסתברות בדיד (סופי)

קבוצה סופית ("מרחב המדגם"\"מרחב ההסתברות") Ω ביחד עם העתקה $Pr:\Omega\Rightarrow\mathbb{R}_{\geq 0}$ עם התכונה ש- $\sum_{x\in\Omega}\Pr(x)=1$.

הגדרה 1.2 מאורע

לאיברי Ω אנחנו קוראים מאורעות אלמנטריים. ובאופן כללי: מאורע הוא תת קבוצה של Ω. אם $A\subseteq\Omega$ מאורע אז $Pr(A)=\sum_{x\in A}Pr(x)$. לכל מאורע מוגדרת ההסתברות שלו $Pr(A)$ עם הזהויות הברורות

$$\begin{aligned}&Pr(\emptyset)=0\\&Pr(\Omega)=1\\&A\cap B=\emptyset\Rightarrow Pr(A\cup B)=Pr(A)+Pr(B)\end{aligned}$$

אחרי ההרחבה הזו $Pr:2^{\Omega}\Rightarrow\mathbb{R}_{\geq 0}$.

ללא מאמץ רב אפשר להרחיב את ההגדרה גם למקרה שΩ אינסופית ובת מניה (ז"א $|\Omega|=\aleph_0$ $\Leftrightarrow$ יש התאמה חח"ע $\Omega\leftrightarrow\mathbb{N}$).

הערה 1.3 מתי לא די במערכת המושגים הזאת? איך מתגברים על חולשותיה?

דוגמה: מטילים קוביה מאוזנת עד שיוצא המס 6. עוצרים כאשר לראשונה מתקבל המס 6. רוצים להבין באופן אופייני לכמה הטלות נזדקק, מה ההסתברות שנסטה מזה באופן משמעותי. כדי להבין איך לענות על שאלה זו, יש לדבר על המרחב:

$$\Omega=\{a_1,a_2,....|a_i\in\{1,..,6\}\}$$

נשים לב $|\Omega|=\aleph=[0,1]$.

ע"מ להניח את היסודות לטיפול במרחבים כאלה נדרשת עבודת הכנה בתורת המידה (measure theory). זו תורה שמאפשרת לייחס "גודל" (למשל שטח\נפח) למשפחה של תת קבוצות של קבוצה.

הגדרה 1.4 התניה

$$Pr(B/A)=\frac{Pr(A\cap B)}{Pr(A)}$$

הגדרה 1.5 אי תלות

אומרים ששני מאורעות $A,B\subseteq\Omega$ הם בלתי תלויים אם:

$$Pr\left(A\cap B\right)=Pr\left(A\right)\cdot Pr\left(B\right)$$

וזה שקול ל$Pr\left(B/A\right)=Pr\left(B\right)$ (זה שאמרנו שהמאורע A התקיים לא משפיע על ההסתברות שהמאורע B התקיים).

הערה 1.6 דוגמא: מטילים קוביה פעמיים. A הוא המאורע שיצא 3 בהטלה הראשונה, B הוא המאורע שיצא 5 בהטלה השניה. אז

$$Pr\left(B\right)=\frac{1}{6}=Pr\left(A\right)$$
$$Pr\left(A\cap B\right)=\frac{1}{36}$$

הגדרה 1.7 משתנה מקרי זו פונקציה ... $f:\left(\Omega,Pr\right)\rightarrow$.

הגדרה 1.8 מ"מ ממשי: הטווח הוא $\mathbb{R}$.

הגדרה 1.9 מ"מ מציין ($indicator$): הטווח הוא $\{0,1\}$. יש התאמה חח"ע בין מ"מ מציינים לבין מאורעות. למ"מ מציין f מתאים המאורע $\{\omega\in\Omega|f\left(\omega\right)=1\}$.

1.1.1 מהם הגדלים שאנחנו רוצים לייחס למ"מ?

הגדרה 1.10 תוחלת ("ממוצע"): אם $f:\Omega\rightarrow\mathbb{R}$, אז $\mathbb{E}\left[f\right]=\sum_{x\in\Omega}Pr\left(x\right)\cdot f\left(x\right)$.

טענה 1.11 תוחלת היא לינארית: ז"א אם $X_1,...,X_k$ מ"מ על אותו מרחב הסתברות, ואם $a_1,...,a_k\in\mathbb{R}$, אז: $\mathbb{E}\left[\sum a_iX_i\right]=\sum a_i\mathbb{E}\left[X_i\right]$

חשוב להבין עד כמה המ"מ X "מרוכז", כלומר - נוטה בהסתברות גבוהה לקבל ערכים קרובים לתוחלתו $E\left[X\right]$.

הגדרה 1.12 השונות של X:

$$\begin{aligned}Var\left[X\right]:&=E\left[\left(X-E\left[X\right]\right)^2\right]=E\left[X^2-2XE\left[X\right]+E\left[X\right]^2\right]=\\&=E\left[X^2\right]-2E\left[X\right]^2+E\left[X\right]^2=E\left[X^2\right]-E\left[X\right]^2\end{aligned}$$

1.2 אי שוויונות בסיסיים בתורת ההסתברות

טענה 1.13 חסם האיחוד: $Pr\left(\cup_{i=1}^{n}A_i\right)\leq\sum Pr\left(A_i\right)$

טענה 1.14 אי שוויון מרקוב:
אם X מ"מ אי שלילי, אז עבור קבוע $c>1$ נקבל: $Pr\left(X\geq cE\left[X\right]\right)\leq\frac{1}{c}$.

טענה 1.15 החסם של אי שוויון מרקוב הדוק.

הוכחה: עבור

$$X:\Omega\to\{0,1\}$$
$$Pr\left(X=1\right)=\frac{1}{c}$$
$$Pr\left(X=0\right)=1-\frac{1}{c}$$
$$\Downarrow$$
$$\mathbb{E}\left[X\right]=\frac{1}{c}$$
$$\Downarrow$$
$$Pr\left(X\geq cE\left[X\right]\right)=Pr\left(X\geq1\right)=\frac{1}{c}$$

■

טענה 1.16 אי שוויון צ'בישב:
יהי X מ"מ עם תוחלת $E\left[X\right]=\mu$ ושונות $Var\left[X\right]=\sigma^{2}$ (σ סטיית תקן standard deviation). יהי $c>0$, אז:

$$Pr\left(\left|X-\mu\right|\geq c\sigma\right)\leq\frac{1}{c^{2}}$$

הוכחה: נגדיר מ"מ חדש $Y=\left(X-\mu\right)^{2}$, נציב במרקוב.

$$Pr\left(Y\geq c^{2}\sigma^{2}\right)=Pr\left(\left|X-\mu\right|\geq c\sigma\right)\leq\frac{1}{c^{2}}$$

■

הערה 1.17 ל $E\left[X^{k}\right]$ קוראים המומנט הk של X. תוחלת = מומנט ראשון, שונות (זה בערך) מומנט שני.

1.3 תמורות

הגדרה 1.18 תמורה - העתקה חח"ע $\pi:\{1,...,n\}\to\{1,..,n\}$

הגדרה 1.19 נקודת שבת: $i\in[n]$ היא נקודת שבת (fixed point) של התמורה π אם $\pi\left(i\right)=i$.

דוגמה: מרחב ההסתברות שלנו יהיה

$$\Omega=\{\text{All permutations of }\{1,...,n\}\text{ in uniform probability}\}$$

נגדיר

$$X:\Omega\to\mathbb{N}_{\geq0}$$
$$\{S_{n}=\Omega,unif\}$$

כש $X\left(\pi\right)=\text{number of fixed points of }\pi$, והשאלה היא $E\left[X\right]=?$.

פתרון:

נגדיר מ"מ מציינים $X_1,...,X_n$ כש $X_i(\pi)=\begin{cases}1 & \pi(i)=i\\ 0 & otherwise\end{cases}$, נשים לב:

$$X=\sum X_i$$

מלינאריות התוחלת: $E[X]=\sum_{i=1}^{n}[X_i]$. לכל i, היות שX_i מ"מ מציין:

$$E[X_i]=Pr_{\pi\in S_n}(\pi(i)=i)=\frac{\text{number of permutations where }\pi(i)=i}{n!}=\frac{(n-1)!}{n!}=\frac{1}{n}$$
$$\Downarrow$$
$$E[X]=n\frac{1}{n}=1$$

הגדרה 1.20 התפלגות פואסונית $Poi(\lambda)$ עם פרמטר λ זו התפלגות על המספרים השלמים האי שליליים:

$$Pr(k)=e^{-\lambda}\frac{\lambda^k}{k!}$$

(נזכר ש $\sum_{k=0}^{\infty}e^{-\lambda}\frac{\lambda^k}{k!}=1$).

מס' נקודות השבת הוא פואוסני בגבול עם $\lambda=1$, כלומר לכל k טבעי:

$$lim_{n\to\infty}Pr(X=k)=\frac{1}{e}\cdot\frac{1}{k!}$$

1.3.1 עוד משהו על ההתנהגות האופיינית של תמורות

הצגנו מקודם תמורה כך:

12345678
56217438

דרך אחרת:

$$1\to5\to7\to3\to2\to6\to4\to1$$

נשים לב שזהו מעגל.

מתברר שהמבנה האופייני של תמורה מקרית בהצגה זו הוא כמו כמה מעגלים, כשאחד בערך באורך $\frac{n}{2}$, אחד בערך באורך $\frac{n}{4}$, ..., והאחרון מאורך 1 (והוא נקודת השבת).

דוגמה:

שוב עובדים ב$(S_n, uniform)$, קובעים k טבעי, ומגדירים מ"מ

$$X(\pi) = \text{number of circles of length k in } \pi$$

אם $k=1$ זו בדיוק השאלה הקודמת. רוצים להציג את X כסכום של משתנים מקריים מציינים.

לכל סידרה $\alpha_1,...,\alpha_k$ של מס' שונים זה מזה ב$\{1,..,n\}$, נרצה להתאים מ"מ מציין. עלינו לחשוב על $\alpha_1,..,\alpha_k$ ועל $a_2,...,\alpha_k,\alpha_1$, וכו', כעל אותו אובייקט. נסמן את מחלקת השקילות הזו ב$\tau = [\alpha_1,...,\alpha_k]$.

$$X = \sum_{\tau\,\in \text{equiv. groups}} X_\tau$$

נגדיר

$$X_\tau(\pi) = \begin{cases} 1 & \tau\in\pi \\ 0 & otherwise \end{cases}$$

$$E[X] = \sum_{|\tau|=k} E[X_\tau] = |\text{equivalence groups}|\, Pr_{\pi\in_{uniform} S_n}(\tau\in\pi) =$$
$$= \left(\binom{n}{k}(k-1)!\frac{(n-k)!}{n!}\right) = \frac{n!}{k!(n-k)!}(k-1)!\frac{(n-k)!}{n!} = \frac{1}{k}$$

1.4 גרפים

משפט 1.21 משפט Ramsey בגרסה כמותית (Erdos-Szekeres):
יהיו $k,l\in\mathbb{N}$, ויהיה $N=\binom{k+l-2}{k-1}$ (נשים לב ש$N=\binom{k+l-2}{l-1}$). אם צובעים את הצלעות של K_N (הגרף השלם עם N קודקודים) בכחול ובאדום, אז יש בהכרח kיה כחולה או lיה אדומה (או שניהם).

הערה 1.22 נשים לב: אם $k=l$ אז $\binom{2k-2}{k-1}\approx\frac{4^k}{\sqrt{k}}$. אז:

$$N=4^k \Leftrightarrow k=\frac{1}{2}log_2 N$$

משפט 1.23 מקרה פרטי של המשפט של רמזי: בכל גרף על N קודקודים יש קליקה מגודל $<\frac{1}{2}log_2N$ או אנטי קליקה (indipendent set) מגודל $<\frac{1}{2}log_2N$.

משפט 1.24 יש גרפים על N קודקודים ללא קליקה וללא אנטיקליקה מגודל $<2log_2N$.

הוכחה: רעיון ההוכחה: נעבוד במרחב ההסתברות $\Omega = G\left(N,\frac{1}{2}\right)$, זה מרחב ההסתברות הכולל את כל הגרפים על קבוצת הקודקודים $\{1,..,N\}$ בהתפלגות אחידה. יש $2^{\binom{N}{2}}$ גרפים כאלה, ולכן ההסתברות של כל גרף היא

$2^{-\binom{N}{2}} = \frac{1}{2^{\binom{N}{2}}}$. דרך אחרת לתאר את המרחב הזה: לכל זוג קודקודים נטיל מטבע מאוזן ונשים צלע ביניהם אס"ם המטבע יוצא ראש. נקבע מס' טבעי t, ונגריל 2 מ"מ:

$$X(G) = \text{how many t-cliques are in G}$$
$$Y(G) = \text{" t-anticliques "}$$

מראים שאם $t > 2log_2 N$ אז $E[X+Y] = E[X] + E[Y] = o(1)$, ומסיקים שיש G שבשבילו $X(G) = Y(G) = 0$, בG כזה אין t קליקה, ואין t־אנטיקליקה.

$$X = \sum_{T\subseteq[N],|T|=t} X_T;$$
$$E[X] = \sum E[X_T] = \binom{N}{t} Pr(\text{t-ary is a clique}) = \binom{N}{t}\left(\frac{1}{2}\right)^{\binom{t}{2}}$$

■

2 תרגול 1

הגדרה 2.1 התפלגות ברנולי Bernoulli

ממ X מתפלג ברנולי עם פרמטר $p \in [0,1]$ אם

$$Pr[X=1] = p$$
$$Pr[X=0] = 1-p$$

ה־דוגמא: ממ מציין ־ בהנתן מאורע A, הממ המציין של A מקבל 1 אסם A קרה, ו־0 אחרת.

נשים לב:

$$E[X] = 1\cdot p + 0\cdot(1-p) = p$$
$$Var[X] = E\left[|X - E[X]|^2\right] = E\left[X^2\right] - E[X]^2 =$$
$$= E[X] - E[X]^2 = p - p^2 = p(1-p)$$

הגדרה 2.2 התפלגות בינומית

נאמר $X \sim Bin(n \in N, p \in [0,1])$ אם לכל $0 \le k \le n$:

$$Pr[X=k] = \binom{n}{k} p^k (1-p)^{n-k}$$

אפשר לחשוב על X כך: $X=\sum_{i=1}^{n}X_i$, כאשר: $\forall i: X_i\sim Ber\left(p\right)$,וכולם ב"ת. נשים לב:

$$\begin{aligned}E\left[X\right]&=E\left[\sum_{i=1}^{n}X_i\right]=^{linear}\sum_{i=1}^{n}E\left[X_i\right]=\sum_{i=1}^{n}p=np\\E\left[X^2\right]&=E\left[\left(\sum_{i=1}^{n}X_i\right)^2\right]=\sum_{i=1}^{n}\sum_{j=1}^{n}E\left[X_iX_j\right]=\sum_{i=1}^{n}E\left[X_i^2\right]+\sum_{i=1}^{n}\sum_{j\neq i}^{n}E\left[X_iX_j\right]=\\&=\sum_{i=1}^{n}E\left[X_i\right]+\sum_{i=1}^{n}\sum_{j\neq i}^{n}E\left[X_iX_j\right]=np+\sum_{i=1}^{n}\sum_{j\neq i}^{n}E\left[X_iX_j\right]=\\&=np+n\left(n-1\right)p^2\end{aligned}$$

המעבר האחרון נובע מכך שההסתברות שX_i,X_j שניהם יחד יהיו שונים מ0 היא p^2.

$$Var\left[X\right]=E\left[X^2\right]-E\left[X\right]^2=np+n\left(n-1\right)p^2-n^2p^2=np-np^2=np\left(1-p\right)=nVar\left[X_1\right]$$

2.1 דוגמא - גרפים מקריים

ראינו את מודל $G\left(n,\frac{1}{2}\right)$ ליצירת גרפים. גם מודל שנקרא $G\left(n,p\right)$ הוא מעניין, כאשר $p\in\left[0,1\right]$ וכל קשת נכנסת לגרף בהסתברות p במקום בהסתברות $\frac{1}{2}$. היום נראה מודל שנקרא $G\left(n,n,p\right)$, המכיל שתי קבוצות במקום קבוצה אחת של n קודקודים. נקרא לקבוצה אחת X, ולשניה Y. במודל זה, כל הקשתות בין X לY. מודל זה יוצר גרפים דו צדדיים מקריים.

מ"מ מציין שמקבל 1 אםם $\left(x,y\right)\in E$ מתפלג ברנולי עם פרמטר p. עבור $x\in X$ נסמן בR_x את הממ המציין של המאורע בו x מבודד, כלומר - לא יושב באף צלע.

$$R_x\sim Ber\left(\left(1-p\right)^n\right)$$

נשים לב שההסתברות היא כיוון שההסתברות שלא תיהיה צלע לקודקוד אחד אחר היא $\left(1-p\right)$. יש n קודקודים מהקבוצה האחרת, ולכן נקבל הסתברות זו. נסמן את הקודקודים בX ע"י $\left[n\right]=\left\{1,..,n\right\}$. נסמן:

$$R=\sum_{i=1}^{n}R_i$$

נשים לב שR סופר את הקודקודים הבודדים בגרף. כיוון שהR_i מתפלגים ברנולי עם אותו פרמטר, וכיוון שהם ב"ת, ניתן לאמר שR מתפלג בינומי. מדוע הם ב"ת? ניתן לראות שהם ב"ת בזוגות כיוון שההסתברות ששני קודקודים יהיו בודדים שווה למכפלת ההסתברות שכל אחד בנפרד בודדים. צריך להוכיח שכל תת קבוצה שלהם ב"ת, וניתן לעשות זאת באינדוקציה על אותו טיעון. קיבלנו: $R\sim Bin\left(n,\left(1-p\right)^n\right)$

כיצד המאורע $R>0$ תלוי בp?

התכונה "אין קודקודים מבודדים" היא מונוטונית (אם אני מוסיף צלעות לגרף בלי קודקודים מבודדים, הוא עדיין ישאר בלי קודקודים מבודדים). מהו $Pr\left[R=0\right]$ כפונק' של p? אם $p=0$ אין צלעות בגרף, ולכן יש קודקודים מבודדים, ולכן $Pr\left[R=0\right]=0$. אם $p=1$ נקבל $Pr\left[R=0\right]=1$.

משפט 2.3 השאפה של הסתברות לקודקודים מבודדים

א. יהי $c<1$, $p\le c\cdot\frac{ln(n)}{n}$ אז

$$\lim_{n\to\infty}Pr\left[G\left(n,n,p\right)\text{ has a solitary vertex}\right]=1$$

ב. יהיו $c>1$ ו $p\ge c\frac{ln(n)}{n}$ אז

$$\lim_{n\to\infty}Pr\left[G\left(n,n,p\right)\text{ has no solitary vertices}\right]=1$$

הוכחה: הוכחת ב:
נסמן בI_X את המאורע בו יש קודקוד מבודד בX. באופן דומה נסמן I_Y ו$I=I_X\cup I_Y$ המאורע בו יש קודקוד מבודד.

$$\begin{aligned}
Pr\left[I\right]&\le Pr\left[I_X\right]+Pr\left[I_Y\right]\le^{symmetry}2Pr\left[I_X\right]=2Pr\left[I_Y\right]\\
Pr\left[I_X\right]&=Pr\left[R>0\right]=^{\text{R takes only integer values}}Pr\left[R\ge 1=\frac{1}{E\left[R\right]}E\left[R\right]\right]\le E\left[R\right]\\
E\left[R\right]&=n\left(1-p\right)^n\le_{\forall x\in R:1+x\le exp(x)}n\cdot exp\left(-pn\right)\le
\end{aligned}$$

כיוון שעכשיו $p\ge c\frac{ln(n)}{n}$ נקבל $-pn\le -c\cdot ln\left(n\right)$ ולכן:

$$\le n\cdot exp\left(-c\cdot ln\left(n\right)\right)=n^{1-c}\to_{\text{because 1-c<0}}0$$

■

3 תרגול אקסטרה 1

פרטים

יובל פלד, נמצא ב$A422$ בימי ב' 14-15. ליצור קשר לפני במייל!

$$yuvalp@cs.huji.ac.il$$

דוגמה

נתונה רשת מעגלית של רכבות עם n תחנות וn מסילות. נתון שכל תחנה פעילה בהסתברות $\frac{1}{2}$ ובאופן ב"ת משאר התחנות. מסילה פעילה אם התחנות בשני קצותיה פעילות. נסמן בX את מס' המסילות הפעילות.

1. מה התוחלת של X?

נציג את X כסכום של ממ מציינים:

$$X_i=\begin{cases}1 & \text{active rail}\\ 0 & otherwise\end{cases}\Rightarrow X=\sum_{i=1}^{n}X_i$$

נשתמש בלינאריות של התוחלת:

$$E\left[X_i\right] = 1 \cdot Pr\left[X_i = 1\right] + 0 \cdot Pr\left[X_i = 1\right] = Pr\left[X_i = 1\right] = \frac{1}{4}$$
$$\Downarrow$$
$$E\left[X\right] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E\left[X_i\right] = \frac{n}{4}$$

הערה 3.1 אם Y מ"מ מציין, אז $E\left[Y\right] = Pr\left[Y = 1\right]$.

2. $Pr\left(X = 0\right) = ?$

נעשה זאת בעזרת מרקוב!

הערה 3.2 תזכורת לא"ש מרקוב: X הוא מ"מ אי שלילי ו$a > 0$, אז $Pr\left(X \geq a\right) \leq \frac{E[X]}{a}$

נשים לב:

$$Pr\left(X = 0\right) = Pr\left(X \leq 0\right)$$

א"ש מרקוב עוסק בסטיות מהתוחלת כלפי מעלה, הפתרון הוא להשתמש במשלים, למשל במקרה שלנו נגדיר: $Y = n - X$. כיוון ש$X \leq n$ נקבל $0 \leq Y$, ואז:

$$E\left[Y\right] =^{linear} n - E\left[X\right] = \frac{3n}{4}$$
$$Pr\left(X = 0\right) = Pr\left(X \leq 0\right) = Pr\left(Y \geq n\right) \leq \frac{E\left[Y\right]}{n} = \frac{3}{4}$$

3. $Var\left[X\right] = ?$

הגדרה 3.3 שונות משותפת:

$$Cov\left(X_i, X_j\right) = \mathbb{E}\left[X_i X_j\right] - \mathbb{E}\left[X_i\right]\mathbb{E}\left[X_j\right]$$

טענה 3.4 אם $X = \sum X_i$, אז

$$Var\left[X\right] = \sum Var\left[X_i\right] + \sum_{i \neq j} Cov\left(X_{i,} X_j\right)$$

הוכחה: $Var\left[X\right]=\mathbb{E}\left[X^{2}\right]-\mathbb{E}^{2}\left[X\right]$ + לינאריות של תוחלת. ■

נשתמש בטענה זו כדי להוכיח את הנדרש.

$$Var\left[X_{i}\right]=\mathbb{E}\left[X_{i}^{2}\right]-\mathbb{E}^{2}\left[X_{i}\right]=\mathbb{E}\left[X_{i}\right]-\mathbb{E}^{2}\left[X_{i}\right]=\frac{1}{4}-\frac{1}{16}=\frac{3}{16}$$

הערה 3.5 אם Y הוא מ"מ מציין ו$\mathbb{E}\left[Y\right]=p$ אז $Var\left[Y\right]=p-p^{2}=p\left(1-p\right)$.

אם המסילות הi והj לא נחתכות (אין תחנה משותפת) אז המאורעות שהמסילה הi פעילה והמסילה הj פעילה ב"ת, ולכן:

$$\begin{aligned}\mathbb{E}\left[X_{i}X_{j}\right]&=Pr\left(\text{the i and j rails are active}\right)=\\&=Pr\left(\text{the i-th rail is active}\right)Pr\left(\text{the j-th rail is active}\right)=\\&=\frac{1}{16}\end{aligned}$$

נשים לב:

$$\mathbb{E}\left[X_{i}X_{j}\right]=\begin{cases}1 & \text{both i and j are active}\\0\end{cases}$$

$$Cov\left(X_{i},X_{j}\right)=\mathbb{E}\left[X_{i}X_{j}\right]-\mathbb{E}\left[X_{i}\right]\mathbb{E}\left[X_{j}\right]=\mathbb{E}\left[X_{i}X_{j}\right]-\frac{1}{16}$$

לכן עבור מסילות i,j שלא נחתכות נקבל $Cov\left(X_{i},X_{j}\right)=0$. במקרה האחר:

$$\mathbb{E}\left[X_{i}X_{j}\right]=Pr\left(\text{i \& j rails are active}\right)=\frac{1}{8}\Rightarrow Cov\left(X_{i},X_{j}\right)=\frac{1}{8}-\frac{1}{16}=\frac{1}{16}$$

סה"כ:

$$\begin{aligned}Var\left[X\right]&=\sum Var\left[X_{i}\right]+\sum_{\text{i,j intersecting}}Cov\left(X_{i},X_{j}\right)+\sum_{\text{i,j not intersecting}}Cov\left(X_{i},X_{j}\right)=\\&=\frac{3}{16}n+2n\frac{1}{16}+0=\frac{5n}{16}\end{aligned}$$

הערה: $\sum_{\text{i,j intersecting}}$ הוא סכום על הזוגות הסדורים, לפי הנוסחה של חישוב Var לפי Cov.

4. $Pr\left(X=0\right)=?$

נעשה זאת בעזרת צב'ישב!

הערה 3.6 תזכורת לא"ש צ'בישב: $Pr\left(\left|X-\mathbb{E}\left[X\right]\right|\geq t\right)\leq\frac{Var\left[X\right]}{t^{2}}$

במקרה שלנו:

$$Pr\left(X=0\right)\leq^{1}Pr\left(\left|X-\mathbb{E}\left[X\right]\right|\geq\mathbb{E}\left[X\right]\right)\leq_{chebyshev}\frac{Var\left[X\right]}{\mathbb{E}^{2}\left[X\right]}=\frac{\frac{5n}{16}}{\frac{n^{2}}{16}}=\frac{5}{n}$$

הערה 3.7 $\leq^{1}$ נובע מכך ש:

$$X=0\Rightarrow\left|X-\mathbb{E}\left[X\right]\right|\geq\mathbb{E}\left[X\right]$$

כלומר, המאורע $X=0$ מוכל במאורע $\left|X-\mathbb{E}\left[X\right]\right|\geq\mathbb{E}\left[X\right]$ ולכן $Pr\left(X=0\right)\leq^{1}Pr\left(\left|X-\mathbb{E}\left[X\right]\right|\geq\mathbb{E}\left[X\right]\right)$.

4 תרגיל 1

טענה 4.1 אם X,Y מ"מ ב"ת אז $\mathbb{E}\left[XY\right]=\mathbb{E}\left[X\right]\mathbb{E}\left[Y\right]$.

טענה 4.2 ההפך אינו נכון!

טענה 4.3 אם $\left\{X_{i}\right\}_{i=1}^{n}$ ב"ת בזוגות אז $Var\left[\sum_{i=1}^{n}X_{i}\right]=\sum_{i=1}^{n}Var\left[X_{i}\right]$

5 שיעור 2

5.1 חזרה על שיעור קודם

משפט 5.1 משפט $Ramsey$ כמותי \ משפט Erdos Szekeres

אם צובעים את צלעות K_{n} בכחול ואדם, אז בהכרח יש K_{l} כחול או K_{l} אדום (או שניהם) $\Leftrightarrow$ $=\binom{k+l-2}{k-1}\geq\left(k,l\right)$ $\binom{k+l-2}{l-1}$.

מסקנה 5.2 מהצבה $k=l$ מקבלים: בכל גרף על n קודקודים יש קליקה או אנטיקליקה מגודל $\left(\frac{1}{2}-o\left(1\right)\right)log_{2}n\leq$.

מאידך גיסא: יש גרפים (בעצם רוב הגרפים) מסדר n (= בעלי n קודקודים) בהם אין קליקה \ אנטיקליקה $\left(2+o\left(1\right)\right)log_{2}n>$.

הוכחה: נביט במרחב ההסתברות $G\left(n,\frac{1}{2}\right)$ $\Leftrightarrow$התפלגות אחידה על כל הגרפים מסדר n. נגדיר פרמטר k שיקבע בהמשך, ונגדיר שני מ"מ

$$X\left(G\right)=\text{number of k-cliques in G}$$
$$Y\left(G\right)=\text{number of k-anticliques in G}$$

אם נראה ש $\mathbb{E}\left[X+Y\right]=2\mathbb{E}\left[X\right]<1$ זה היה גורר שיש גרפים ללא k קליקות וללא k אנטיקליקות. ז"א $\exists G\ s.t.\ X\left(G\right)=Y\left(G\right)=0$. בעצם יתברר שע"י בחירה מתאימה של k יצא:

$$\mathbb{E}\left[X+Y\right]=o\left(1\right)$$

ומכך ינבע שבהסתברות $1-\epsilon_n >$ עבור $n\to\infty,\ \epsilon_n\to 0$ אין בגרף מקרי k קליקה ולא k אנטיקליקה.
נכתוב:

$$X=\sum_{T\subseteq[n],|T|=k}X_T$$
$$X_T\left(G\right)=\begin{cases}1 & \text{T-clique in G}\\ 0 & otherwise\end{cases}$$
$$\mathbb{E}\left[X\right]=\binom{n}{k}\left(\frac{1}{2}\right)^{\binom{k}{2}}<^{k=?}\frac{1}{2}$$

נשים לב, מסטרלינג נקבל עבור $n>>k>>1$:

$$\binom{n}{k}=\frac{n\left(n-1\right)...\left(n-k+1\right)}{k!}=\left(1+o\left(1\right)\right)\left(\frac{ne}{k}\right)^{k}$$

אז:

$$\mathbb{E}\left[X\right]=\binom{n}{k}\left(\frac{1}{2}\right)^{\binom{k}{2}}=\left(\left(\frac{ne}{k}\right)\left(\frac{1}{2}\right)^{\frac{k-1}{2}}\right)^{k}<^{k=?}\frac{1}{2}$$

הוכחה: מהו k שבשבילו $\left(\frac{ne}{k}\right)\left(\frac{1}{2}\right)^{\frac{k-1}{2}}<1-\epsilon$? אם נמצא כזה k, נקבל ש$\mathbb{E}\left[X\right]\to 0$. נוציא log:

$$\frac{ne}{k}<\left(1-\epsilon\right)2^{\frac{k-1}{2}}$$
$$log_2n+log_2k+c<\frac{k-1}{2}-\epsilon$$
$$\Downarrow$$
$$\left(2+o\left(n\right)\right)log_2n<k$$

■

■

5.2 פונקציית הסף להופעה של K_4 ב$G\left(n,p\right),\ \ p=n^{-\frac{2}{3}}$

5.2.1 תזכורות

משפט 5.3 אי שוויון צ'בישב
אם X מ"מ ממשי עם תוחלת $\mu=\mathbb{E}\left[X\right]$ וסטיית תקן σ, ז"א $\sigma^2=Var\left[X\right]$, אז:

$$\Pr\left(\left|X-\mu\right|\geq c\sigma\right)\leq\frac{1}{c^2}$$

מסקנה 5.4 מקרה פרטי: אם X מ"מ אישלילי עם תוחלת μ ושונות σ^2 אז

$$c = \frac{\mu}{\sigma} \Rightarrow^{chebyshev} \Pr\left(|X-\mu| \geq \mu\right) = \Pr\left(X=0\right) \leq \frac{Var\left[X\right]}{\mu^2}$$

אם $\frac{Var[X]}{\mu^2} = o_n\left(1\right)$ אז אסימפטוטית כמעט בוודאות $X>0$.

5.2.2 פונקציית הסף להופעה של K_4 ב $G\left(n,p\right),\ p = n^{-\frac{2}{3}}$

משפט 5.5 נביט בגרף מקרי $G\left(n,p\right)$. אם $p = o\left(n^{-\frac{2}{3}}\right)$ אז אכב"ו (אסימפטוטית כמעט בוודאות) אין K_4 בG .
מצד שני, אם $p >> n^{\frac{-2}{3}}$ יש K_4 בG.

הוכחה: הוכחה לא'. נגדיר מ"מ X המונה את מס' הK_4 שב G.

$$X_S\left(G\right) = \begin{cases} 1 & \text{S is a clique} \\ 0 & otherwise \end{cases}$$

$$X = \sum_{|S|=4} X_S$$

$$\mathbb{E}\left[X\right] = \binom{n}{4} p^6 << n^4 \cdot n^{-4} = 1$$

כי $n^{-4} << p^6 \Leftarrow n^{-\frac{2}{3}} >> p \leftarrow \left(n^{-\frac{2}{3}}\right) p = o$.

הוכחה לב'. כאשר $p >> n^{-\frac{2}{3}}$ אז

$$\mathbb{E}\left[X\right] = \binom{n}{4} p^6 \to \infty$$

ע"פ המקרה הפרטי של צ'בישב נוכל להוכיח את ב' אם נראה ש

$$Var\left[X\right] = o\left(\mu^2\right)$$

נתחיל:

$$X = \sum_{|S|=4} X_S$$

$$Var\left[X\right] = \mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right]^2 = \mathbb{E}\left[\left(\sum_{|S|=4} X_S\right)^2\right] - \mathbb{E}\left[\sum_{|S|=4} X_S\right]^2 =^1$$

נשים לב ש

$$\left(\sum a_i\right)^2 = \sum a_i^2 + 2\sum_{i<j} a_i a_j$$

אז:

$$=^{1} \sum_{|S|=4} \left(\mathbb{E}\left[X_S^2\right] - \left(\mathbb{E}\left[X_S\right]\right)^2\right) + 2\sum_{S<T} \left(\underbrace{\mathbb{E}\left[X_S X_T\right] - \mathbb{E}\left[X_S\right]\mathbb{E}\left[X_T\right]}_{Cov(X_S,X_T)}\right) =$$

$$=^{X_S\ is \text{only either 1 or 0}} \underbrace{\sum_{|S|=4} \left(\mathbb{E}\left[X_S\right] - \left(\mathbb{E}\left[X_S\right]\right)^2\right)}_{\leq \mu = \sum_{|S|=4}\mathbb{E}[X_S]} + 2\sum_{S<T} \left(\mathbb{E}\left[X_s X_T\right] - p^{12}\right) =$$

נשים לב:

$$\mathbb{E}\left[X_S X_T\right] = \Pr\left(S \text{ is a clique } \wedge T\ is\ a\ clique\right) = p^{11}$$

נשים לב שS, T נחתכים, לכן סה"כ יש בהם 6 קודקודים, לכן:

$$\binom{n}{6}\binom{6}{2,2,2}$$

לכן:

$$\theta\left(n^6\left(p^{11} - p^{12}\right)\right) = \theta\left(n^6 p^{11}\right) = \theta\left(n^6 - 11 \cdot \frac{2}{3}\right) = n^{-\frac{4}{3}}$$

■

5.3 שתי עובדות חשובות urn models

משליכים באקראי כדורים לכדים.

1. מתי לראשונה תחול התנגשות? ז"א, מתי צפוי שיהיה כד עם יותר מכדור אחד.

   תשובה: בערך אחרי $\sqrt{n}$ זריקות. פרדוקס ימי ההולדת.

2. אספן הקופונים - באיזה רגע לא יוותרו כדים ריקים?

   תשובה: בערך $nln\left(n\right)$.

נבנה מרחב הסתברות: כל המילים $\omega = \omega_1\omega_2....$ בא"ב $1, .., n$ ($\forall i:\ \omega_i \in [n]$).
נגדיר ממ X:

$$X\left(\omega\right) = min\left\{t \middle| \left\{w_1, ..., w_t\right\} = [n]\right\}$$

X שואל מתי המילה מכילה את כל האותיות, כלומר - מתי לא יוותרו כדים ריקים.

משפט 5.6 סטרלינג: $n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$

מסקנה 5.7 $\frac{n^n}{n!} \approx e^n (2\pi n)^{-\frac{1}{2}}$

טענה 5.8 אי שוויון הממוצעים (AMGM):

$$(a_1 \cdot .... \cdot a_k)^{\frac{1}{k}} \leq \frac{a_1 + ... + a_k}{k}$$

הוכחה: נשים לב ש$logx$ קמורה וקעורה עבור $x > 0$:

$$(logx)' = \frac{1}{x} > 0$$
$$(logx)'' = -\frac{1}{x^2} < 0$$

מהקעירות נובע שלכל $a_1, ..., a_k$ הלוג של הממוצע $\leq$ הממוצע של הלוגים (הדגמה לפי ציור). בקמירות, אי השוויון מתהפך. נכתוב זאת:

$$\frac{1}{k} (loga_1 + ... + loga_k) \leq log\frac{a_1 + ... + a_k}{k}$$
$$\Updownarrow$$
$$\frac{1}{k} log (a_1 \cdot ... \cdot a_k) \leq log\frac{a_1 + ... + a_k}{k}$$
$$\Updownarrow$$
$$log\left((a_1 \cdot ... \cdot a_k)^{\frac{1}{k}}\right) \leq log\frac{a_1 + ... + a_k}{k}$$

■

הערה 5.9

$$ln (n!) = ln1 + ln2 + ... + ln (n) \approx \int^n ln (x) dx = xlnx - x|_1^n = nlnn - (n-1) = n (lnn - lne) = nln\frac{n}{e}$$

מסקנה 5.10 אם $a_i = i$, לפי סטרלינג:

$$\frac{n}{e} \approx (a_1 \cdot .... \cdot a_k)^{\frac{1}{k}} \leq \frac{a_1 + ... + a_k}{k} = \frac{n+1}{2}$$

5.3.1 פרדוקס יום ההולדת

זורקים k כדורים לn כדים ורוצים להבין מה ההסתברות שכל כדור נופל בכד אחר. ההסתברות היא:

$$1 \cdot \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdot ... \cdot \left(1 - \frac{k-1}{n}\right) =^1$$

הערה 5.11 טריק שימושי:

$$e^x = 1 + x + \frac{x^2}{2!} + ...$$

בפרט, כש$|x|$ קטן:

$$e^x = 1 + x + O\left(x^2\right)$$

$$\approx^1 e^0 e^{-\frac{1}{n}} e^{-\frac{2}{n}} ... e^{-\frac{k-1}{n}} = e^{-\frac{1}{n}(1+2+...+(k-1))} =$$
$$= e^{-\frac{k(k-1)}{2n}}$$

לכן, אם $k^2 << n$, ז"א $k << \sqrt{n}$, אז ההסתברות שלא חלה התנגשות (שני כדורים באותו כד) היא $1 - o(1)$.
מצד שני, אם $k^2 >> n$, ז"א $k >> \sqrt{n}$, אז ההסתברות הנ"ל שלא חלה התנגשות היא $o(1)$.

5.3.2 חיתוך קבוצות מקריות

נאמר ש$A, B \subseteq [n]$ תת קבוצות מקריות מעוצמות k, l בהתאמה. רוצים לדעת מה ההסתברות שהקבוצות זרות.

$$Pr_{A,B}(A \cap B = \emptyset) = \frac{\binom{n-k}{l}}{\binom{n}{l}} = \frac{(n-k)(n-k-1)...(n-k-l+1)}{n(n-1)....(n-l+1)} = \frac{n-k}{n} \cdot .... \cdot \frac{n-k-l+1}{n-l+1} =$$
$$= \left(1 - \frac{k}{n}\right)\left(1 - \frac{k}{n-1}\right)....\left(1 - \frac{k}{n-l+1}\right) \approx e^{-\frac{k}{n}\cdots-\frac{k}{n-l+1}} \approx e^{-\frac{kl}{n}}$$

מסקנה 5.12 אם $kl >> n$ אכב"ו A, B תחתכנה ואם $kl << n$ אכב"ו תהיינה זרות ואם $kl \sim n$ אז ההסתברות שתחתכנה חסום מ0 ומ1.

5.3.3 אספן הקופונים

מתי לא יוותר כד ריק?

דרך א n כדים, משליכים k כדורים. מהי התוחלת של מס' הכדים הריקים? אם $o(1) \to \mathbb{E}[X]$ אכב"ו אין כדים ריקים. אם $\mathbb{E}[X] \to \infty$ אז יש לקוות שעל ידי שיקול מומנט שני כמעט בוודאות יש כד ריק.
X מ"מ מקרי שמונה את מס' הקדים הריקים:

$$X_i(\omega) = \begin{cases} 1 & \text{i is empty} \\ 0 & otherwise \end{cases}$$
$$X = \sum X_i$$
$$\mathbb{E}[X] = \sum \mathbb{E}[X_i] = n \Pr(\text{the first urn is empty}) = n\left(1 - \frac{1}{n}\right)^k$$

נרצה $n\left(1-\frac{1}{n}\right)^k \approx ne^{-\frac{k}{n}}$. ואכן, אם $k << nln\left(n\right)$ אז $n\left(1-\frac{1}{n}\right)^k \to \infty$ ואילו אם $k >> nln\left(n\right)$ אז $n\left(1-\frac{1}{n}\right)^k \to 0$. הוכחנו: אם $k >> nln\left(n\right)$, אז אכב"ו אין כדים ריקים.
נגדיר מ"מ X_t שסופר כמה צעדים חלפו מהרגע שבו היו $t-1$ כדים שמכילים כדור עד לרגע שבו יש t כדים שמכילים כדור.

$$X = \sum X_i = \text{time until there are no empty urns}$$
$$\mathbb{E}\left[X\right] = \sum \mathbb{E}\left[X_i\right]$$

X_i הוא מ"מ גיאומטרי עם סיכוי הצלחה $p = \frac{n-i+1}{n}$, לכן התוחלת:

$$\mathbb{E}\left[X_i\right] = \frac{1}{p} = \frac{n}{n-i+1}$$
$$\mathbb{E}\left[X\right] = \sum_{i=1}^{n} \frac{n}{n-i+1} = n\sum_{j=1}^{n}\frac{1}{j} = n\left(logn + O\left(1\right)\right)$$

דרך ב סיימנו כאן.

6 תרגול 2

6.1 קירובים שימושיים

הערה 6.1 בשבוע שעבר ראינו:

$$\forall x \in \mathbb{R}:\ 1 + x \le e^x$$

למה 6.2 קיים $c > 0$ כך שלכל $x \in \left[-c, c\right]$:

$$1 - x \ge exp\left(-\left(x + x^2\right)\right)$$

הוכחה: ניתן להוכיח עם טיילור, לא נעשה זאת. נגדיר את פונקציית ההפרש:

$$f\left(x\right) = 1 - x - e^{-\left(x+x^2\right)}$$

המטרה: f אי שלילית בסביבה של אפס. מספיק להוכיח: $f\left(0\right) \ge 0$ וגם f מקבלת מינימום מקומי באפס. נתחיל:

$$f\left(0\right) = 1 - 0 - e^0 = 0 \ge 0$$

כעת נוכיח את המינימום המקומי:

$$f'(x) = -1 - -(1+2x)\,e^{-\left(x+x^2\right)} \Rightarrow f'(0) = -1 + e^0 = 0$$

עכשיו, כדי לסיים את ההוכחה שזהו מינימום מקומי, יש להראות שהנגזרת השניה חיובית בנקודה זו:

$$f''(x) = 2e^{-\left(x+x^2\right)} - (1+2x)^2\, e^{-\left(x+x^2\right)}$$
$$f''(0) = 2e^0 - e^0 = 1 > 0$$

לכן לf יש מינימום מקומי באפס, ו$0 \le f(0)$, כנדרש. ∎

משפט 6.3 קירוב סטירלינג

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$
$$\lim_{n\to\infty}\frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n}{n!} = 1$$

הוכחנו בשיעור: $n! = \left(1+o(1)\,\frac{n}{e}\right)^n$, כלומר: אם נגדיר את $f:\mathbb{N}\to\mathbb{R}$ כך ש: $n! = \left(1+f(1)\,\frac{n}{e}\right)^n$, אז: $f = o(1)$.

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n = \left(\underbrace{(2\pi n)^{\frac{1}{2n}}}_{1+o(1)}\cdot\frac{n}{e}\right)^n = \left((1+o(1))\,\frac{n}{e}\right)^n$$

6.2 הגבלות של שיטת המומנט הראשון - קודקודים מבודדים והתאמות מושלמות

6.2.1 תזכורת $G(n,n,p)$

יש שתי קבוצות של קודקודים X,Y כך ש$|X| = |Y| = n$, וההסתברות שיש צלע בין קודקוד מקבוצה X לקודקוד מקבוצה Y היא p.

קודקוד מבודד הוא קודקוד עם דרגה 0.

טענה 6.4 יהי I המאורע בו יש קודקודים מבודדים. אזי:

1. אם $c > 1$ ו $p \ge c\frac{ln(n)}{n}$, אז $\lim_{n\to\infty}\Pr(I) = 0$
2. אם $0 \le c < 1$ ו $p < c\frac{ln(n)}{n}$, אז $\lim_{n\to\infty}\Pr(I) = 1$

הוכחה: נוכיח את חלק ב' ונשתמש בשיטת המומנט השני. לכל $i \in [n]$ נגדיר את המ"מ R_i בתור האינדיקטור ש$i \in X$ מבודד. נשים לב: $R_i \sim Ber\left((1-p)^n\right)$ (יש n אפשרויות לקודקודים מהם תצא צלע לi , ההסתברות לכך שלא תיהיה צלע היא $1-p$). נגדיר:

$$R = \sum_{i=1}^{n} R_i = \text{number of solitary vertices in X}$$
$$\Downarrow$$
$$R \sim Bin\left(n,(1-p)^n\right)$$

נסתכל על:

$$\mathbb{E}\left[R\right]=n\left(1-p\right)^{n}$$

במצב א': $c>1$ ו $p\geq c\frac{ln(n)}{n}$, לכן:

$$\mathbb{E}\left[R\right]=n\left(1-p\right)^{n}\leq n\cdot exp\left(-pn\right)\leq n\cdot exp\left(-cln\left(n\right)\right)=n^{1-c}\rightarrow0$$

במצב ב': $c<1$ ו $p\leq c\frac{ln(n)}{n}$, לכן:

$$\mathbb{E}\left[R\right]=n\left(1-p\right)^{n}\geq n\cdot\left(1-c\frac{ln\left(n\right)}{n}\right)^{n}\overset{\text{for big enough n}}{\geq}n\cdot exp\left(-n\left(c\frac{ln\left(n\right)}{n}+c^{2}\frac{ln^{2}\left(n\right)}{n^{2}}\right)\right)\geq$$

$$\geq n\cdot\underbrace{exp\left(-cln\left(n\right)\right)}_{=n^{-c}}\cdot\underbrace{exp\left(\underbrace{-c^{2}\frac{ln^{2}\left(n\right)}{n}}_{\rightarrow0}\right)}_{\rightarrow1}=n^{1-c}\left(1+o\left(1\right)\right)\rightarrow\infty$$

נסתכל על:

$$Var\left[R\right]=\underbrace{n\left(1-p\right)^{n}}_{=\mathbb{E}\left[R\right]}\underbrace{\left(1-\left(1-p\right)^{n}\right)}_{\leq1}\leq\mathbb{E}\left[R\right]$$

$$\underbrace{\Pr\left(I^{c}\right)}_{\text{no solitary vertex}}\leq\Pr\left(R=0\right)\leq\Pr\left(\left|R-\mathbb{E}\left[R\right]\geq\mathbb{E}\left[R\right]\right|\right)\overset{chebyshev}{\leq}\frac{Var\left[R\right]}{\mathbb{E}\left[R\right]^{2}}\leq\frac{\mathbb{E}\left[R\right]}{\mathbb{E}\left[R\right]^{2}}=\frac{1}{\mathbb{E}\left[R\right]}\rightarrow0$$

$$\Downarrow$$

$$\Pr\left(I\right)\rightarrow1$$

■

7 תרגול אקסטרה 2

לא התקיים.

8 תרגיל 2

טענה 8.1 חוק ההסתברות השלמה: אם $\{A_i\}_{i=1}^{n}$ חלוקה של Ω ו$B\subseteq\Omega$ אז

$$\Pr\left(B\right)=\sum_{i=1}^{n}\left(\Pr\left(B|A_{i}\right)\Pr\left(A_{i}\right)\right)$$

טענה 8.2 חוק התוחלת השלמה: אם $\{A_i\}_{i=1}^n$ חלוקה של Ω ו$X : \Omega \to \mathbb{R}$ אז

$$\mathbb{E}[X] = \sum_{i=1}^{n} \left(\mathbb{E}[X|A_i] \Pr(A_i)\right)$$

טענה 8.3 אינפי שימושי: $\sum_{n=0}^{\infty} \frac{\lambda^n}{n} = e^{\lambda}$

טענה 8.4 אם $X \sim Pois(\lambda)$ עם $\lambda > 0$ אז:

1. $\forall \lambda \in \mathbb{N} \cup \{0\} : \Pr(X = n) = e^{-\lambda} \frac{\lambda^n}{n!}$
2. $\mathbb{E}[X] = \lambda$
3. $Var[X] = \lambda$
4. אם $Y \sim Pois(\lambda')$ וX, Y ב"ת אז

$$X + Y \sim Pois(\lambda + \lambda')$$

טענה 8.5 קושי שוורץ עבור מ"מ: אם X, Y מ"מ מוגדרים מעל אותו מרחב הסתברות אז:

$$\mathbb{E}[XY]^2 \leq \mathbb{E}\left[X^2\right] \mathbb{E}\left[Y^2\right]$$

טענה 8.6 יהי X מ"מ אי שלילי, אז $\Pr(X > 0) \geq \frac{\mathbb{E}[X]^2}{\mathbb{E}[X^2]}$

9 שיעור 3

9.0.1 ספרים על הסתברות

1. Alon & Spencer
2. Mitzenmacher Upfal

9.1 כמה טענות

טענה 9.1 $\forall x : e^x \geq 1 + x$

הוכחה: כיוון א: מוכיחים בעזרת ציור או לפי טור טיילור של e^x.
כיוון ב: $e^x \geq 1 + x + x^2$ לכל $x < 1.9$. הוכחה עם וולפרם אלפא או שוב לפי טור טיילור. ■

9.1.1 פרדוקס יום ההולדת

לא מדויק: כאשר משליכים כדורים באקראי לn כדים, צפוי שההתנגשות הראשונה תחול בערך בצעד ה$\sqrt{n}$.

באופן קונקרטי: אם בוחרים באקראי קבוצה מגודל a וקבוצה מגודל b מתוך עולם מגודל n, אז ההסתברות שהן זרות היא בקירוב $e^{-\frac{ab}{n}}$.

9.1.2 אספן הקופונים

לא מדויק: משליכים באקראי כדורים לn כדים, צפוי שבערך בצד ה$nln\left(n\right)$ כבר בכל כד יהיה לפחות כדור אחד.

באופן קונקרטי: הגדרנו T_i מ"מ שהוא משך הזמן בין הרגע שיש $i-1$ כדים לא ריקים לרגע שבו יש i כדים לא ריקים. שמנו לב שT_i הוא מ"מ גיאומטרי עם פרמטר $\frac{n-i+1}{n}$. ולכן, התוחלת היא: $\mathbb{E}\left[T_i\right]=\frac{n}{n-i+1}$. ואם $T=\sum_{i=1}^{n}T_i$ אז:

$$T=\sum_{i=1}^{n}T_i$$

$$\Downarrow$$

$$\mathbb{E}\left[T\right]=\sum_{i=1}^{n}\mathbb{E}\left[T_i\right]=n\sum_{i=1}^{n}\frac{1}{n-i+1}=n\overbrace{\underbrace{H_n}_{=\sum_{i=1}^{n}\frac{1}{i}}}^{\text{the harmonic sum}}=n\left(ln\left(n\right)+\gamma+o\left(1\right)\right)$$

במגוון רחב של יישומי הסתברות ל"חיים" אנחנו לא מסתפקים בידיעה מהי התוחלת μ של מ"מ מעניין. חשוב לנו לדעת דברים מעניינים, כמו:

$$\Pr\left(X>\mu+\Delta\right)<?$$
$$\Pr\left(X<\mu-\Delta\right)<?$$
$$\Pr\left(\left|X-\mu\right|>\Delta\right)<?$$

9.2 תזכורת

$X\geq0$ מ"מ עם תוחלת μ ושונות σ^2.

9.2.1 אי שוויון מרקוב

$$\Pr\left(X\geq c\mathbb{E}\left[X\right]\right)\leq\frac{1}{c}$$

דוגמא עם אספן הקופונים:

$$\Pr\left(T>10nln\left(n\right)\right)<\frac{1}{10}$$

9.2.2 אי שוויון צ'בישב

$$\Pr\left(\left|X-\mu\right|\geq c\sigma\right)\leq\frac{1}{c^2}$$

דוגמא עם אספן הקופונים: המטרה: לקבל משפט ריכוז מידה על T (הזמן עד שאין כדים ריקים) באמצעות אי שוויון צ'בישב:

$$\{T_i\} \text{ are indenpendent}$$
$$\Downarrow$$
$$Var[T] = \sum Var[T]$$

אבל כל T_i הוא מ"מ גיאומטרי. נחשב את השונות של מ"מ גיאומטרי: נניח X מ"מ גיאומטרי עם פרמטר p.

$$\Pr(X = k) = (1-p)^{k-1} p$$
$$\mathbb{E}[X] = \sum_{k=1}^{\infty} k \Pr(X = k) = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = p\sum_{k=1}^{\infty} k(1-p)^{k-1} =$$
$$= p\sum_{k=1}^{\infty} k(1-p)^{k-1} =^1 p\frac{1}{p^2} = \frac{1}{p}$$

We used the following calculations in $=^1$

$$\sum_{t=0}^{\infty} x^t = \frac{1}{1-x}$$
$$\sum_{t=0}^{\infty} tx^t = \frac{1}{\{1-x\}^2}$$

נמשיך:

$$\mathbb{E}\left[X^2\right] = \sum_k k^2 \Pr(X = k) = p\sum_k k^2 (1-p)^{k-1}$$
$$\Downarrow k = t-1$$
$$\sum_{t=2}^{\infty} t(t-1)x^{t-2} = \frac{2}{(1-x)^3}$$
$$\sum \left(k^2 + k\right)x^{k-1} = \sum (k+1)kx^{k-1} = \frac{2}{(1-x)^2}$$
$$\Downarrow$$
$$Var[X] = \frac{1-p}{p^2}$$
$$\Downarrow$$
$$Var[T] = \sum Var[T_i] < \sum_{i=1}^{n}\left(\frac{n}{n-i+1}\right)^2 \le n^2 \sum \frac{1}{k^2} \overset{\sum \frac{1}{k^2} = \frac{\pi^2}{6}}{\le} n^2\frac{\pi^2}{6} =: \sigma^2$$
$$\Downarrow$$
$$\sigma = \frac{\pi}{\sqrt{6}}n = o(\mathbb{E}[T])$$

כשמציבים באי שוויון צ'בישב מתקבלים:

$$\Pr\left(\left|T-nln\left(n\right)\right|>c\frac{\pi}{\sqrt{6}}n\right)\leq\frac{1}{c^{2}}$$

c מעניין אם:

1. $c\underset{n\rightarrow\infty}{\rightarrow}\infty$ ולכן החסם הוא $o\left(1\right)$.
2. וכן $c\frac{\pi}{\sqrt{6}}n>>nln\left(n\right)$, למשל $c=\sqrt{log\left(n\right)}$.

9.3 תהליכים סטוכסטיים

9.3.1 מהלכים מקריים (random walks)

נסתכל על הגרף:

$$-----(-2)-(-1)-0-1-2--------$$

יש לנו מהלך מקרי שמגדיר לו מ"מ X שהוא מקומו לאחר n צעדים (כאשר הקודקודים ממוספרים).

$$X_{i}\sim Bernoulli$$
$$\Pr\left(X_{i}=1\right)=\frac{1}{?}$$
$$X=\sum_{i=1}^{n}X_{i}$$
$$\Pr\left(X=0\right)=\frac{\binom{n}{\frac{n}{2}}}{2^{n}}=\theta\left(\frac{1}{\sqrt{n}}\right)$$

$$2^{n}=\binom{n}{0}+\binom{n}{1}+...+\binom{n}{\frac{n}{2}}+...+\binom{n}{n}$$
$$\binom{n}{\frac{n}{2}}=\frac{n!}{\left(\frac{n}{2}!\right)^{2}}\overset{stirling}{=}\frac{\left(\frac{n}{e}\right)^{n}\sqrt{2\pi n}}{\left(\left(\frac{n}{2e}\right)^{\frac{n}{2}}\sqrt{\pi n}\right)^{2}}=\frac{2^{n}\sqrt{2\pi n}}{\pi n}=\frac{2^{n}}{\sqrt{n}}\cdot\sqrt{\frac{2}{\pi}}$$

הערה 9.2 נניח ש

$$Y_{i}=\begin{cases}0 & \frac{1}{2}\\1 & \frac{1}{2}\end{cases}$$
$$Y=\sum Y$$
$$\mathbb{E}\left[Y\right]=\frac{n}{2}$$

ובסיכוי גבוה לא נסטה מהתוחלת יותר מאשר $O\left(\sqrt{n}\right)$.

נחזור למהלכים: מה בדבר $\Pr\left(X=k\right)$?
נניח שהיו $a=\frac{n+k}{2}$ צעדים ימינה ולכן $n-a$ צעדים שמאלה $\Leftarrow$ נקודת הסיום היא $a-\left(n-a\right)=2a-n=k$.
לכן:

$$\Pr\left(X=k\right)=\frac{\binom{n}{\frac{n+k}{2}}}{2^{n}}$$

אמרנו ש

$$\Pr\left(X=0\right)=\theta\left(\frac{1}{\sqrt{n}}\right)$$

וכשנשווה בין שכנים עבור $k\geq0$:

$$\frac{\Pr\left(X=k+1\right)}{\Pr\left(X=k\right)}=\frac{\binom{n}{\frac{n+k}{2}+1}}{\binom{n}{\frac{n+k}{2}}}=\frac{n!\left(\frac{n+k}{2}\right)!\left(\frac{n-k}{2}\right)!}{\left(\frac{n+k}{2}+1\right)!\left(\frac{n-k}{2}-1\right)!n!}=\frac{\frac{n-k}{2}}{\frac{n+k}{2}+1}=\frac{n-k}{n+k+2}=$$
$$=1-\frac{2k+2}{n+k+2}\overset{k<n^{\frac{1}{2}+\epsilon}}{\approx}e^{-\frac{2k}{n}}$$

$$\Pr\left(X=l\right)=\frac{1}{2^{n}}\binom{n}{\frac{n+l}{2}}=\underbrace{\Pr\left(X=0\right)}_{\frac{1}{\sqrt{n}}}\cdot\underbrace{\frac{\Pr\left(X=2\right)}{\Pr\left(X=0\right)}}_{e^{-\frac{1}{n}\cdot2}}\frac{\Pr\left(X=4\right)}{\Pr\left(X=2\right)}....\frac{\Pr\left(X=l\right)}{\Pr\left(X=l-2\right)}=$$
$$=\theta\left[\frac{1}{\sqrt{n}}e^{-\frac{1}{n}\left(\overbrace{2+4+...+l}^{\frac{l^{2}}{2}}\right)}\right]=\theta\left[\frac{1}{\sqrt{n}}e^{-\frac{l^{2}}{2n}}\right]$$

בפרט אם $l>>\sqrt{nln\left(n\right)}$ אז זה $o\left(1\right)$.

9.3.2 מקדמים בינומיים

טענה 9.3 אם $\alpha\in\left(0,1\right)$ אז:

$$\binom{n}{\alpha n}=2^{n\left(H\left(\alpha\right)+o\left(1\right)\right)}$$

כש$H\left(x\right)=-xlog_{2}x-\left(1-x\right)log_{2}\left(1-x\right)$ היא פונקציית האנטרופיה. היא סימטרית סביב חצי, ערכה ב$\frac{1}{2}$ הוא 1, וערכה הוא 0 ב$1,0$.

נסתכל על:

$$\sum \binom{n}{j}\alpha^{j}\left(1-\alpha\right)^{n-j}=\left(\alpha+\left(1-\alpha\right)\right)^{n}=1$$

המחובר הגדול ביותר הוא בין 1 ל$\frac{1}{n+1}$. איך נמצא אותו? נשווה שני איברים עוקבים על ידי חישוב המנה ביניהם ונמצא מתי המנה היא $1<$:

$$\frac{\binom{n}{j}\alpha^{j}\left(1-\alpha\right)^{n-j}}{\binom{n}{j+1}\alpha^{j+1}\left(1-\alpha\right)^{n-j-1}}\approx 1$$

כשj קטן המנה $1<$, כשj גדול המנה $1>$.

$$\frac{\binom{n}{j}\alpha^{j}\left(1-\alpha\right)^{n-j}}{\binom{n}{j+1}\alpha^{j+1}\left(1-\alpha\right)^{n-j-1}}=\frac{1-\alpha}{\alpha}\frac{\left(j+1\right)!\left(n-j-1\right)!}{j!\left(n-j\right)!}=\frac{1-\alpha}{\alpha}\cdot\frac{j+1}{n-j}\overset{\text{for which j?}}{=}1$$

$$\Downarrow$$

$$\left(1-\alpha\right)j=\alpha\left(n-j\right)$$

$$\Downarrow$$

$$j-\alpha j=\alpha n-\alpha j$$

$$\Downarrow$$

$$j=\alpha n$$

נסתכל על האיבר הj, שהוא האיבר הגדול ביותר:

$$\frac{1}{n+1}<\binom{n}{\alpha n}\alpha^{\alpha n}\left(1-\alpha\right)^{\left(1-\alpha\right)n}<1$$

$$\Downarrow$$

$$\frac{1}{n+1}\left(\frac{1}{\alpha^{\alpha}\left(1-\alpha\right)^{1-\alpha}}\right)^{n}<\binom{n}{\alpha n}<\left(\frac{1}{\alpha^{\alpha}\left(1-\alpha\right)^{1-\alpha}}\right)^{n}$$

נותר רק לשים לב ש:

$$\frac{1}{\alpha^{\alpha}\left(1-\alpha\right)^{1-\alpha}}=2^{H(\alpha)}$$

ע"י מעבר לlog_2 בשני האגפים:

$$log_2\frac{1}{\alpha^{\alpha}\left(1-\alpha\right)^{1-\alpha}}=H\left(\alpha\right)$$

$$=log\frac{1}{\alpha^{\alpha}}+log\frac{1}{\left(1-\alpha\right)^{1-\alpha}}=-\alpha log\alpha-\left(1-\alpha\right)log\left(1-\alpha\right)$$

9.4 משפטים

משפט 9.4 הגבול המרכזי CLT - Central Limit Theorem

הגדרה 9.5 יהיו $X_1, X_2, ...$ מ"מ שווי התפלגות וב"ת (או בקיצור - iid). נניח ש

$$\begin{aligned}\mathbb{E}\left[X_1\right] &= \mu \\ Var\left[X_1\right] &= \sigma^2 \\ Y &= \sum_{i=1}^{n} X_i \\ Z &= \frac{Y - n\mu}{\sigma\sqrt{n}}\end{aligned}$$

משפט 9.6 תהיה $\varphi\left(x\right) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ פונקצית הצפיפות של מ"מ $N\left(0,1\right)$ ו-$\Phi\left(x\right) = \int_{-\infty}^{x}\varphi\left(x\right)dx$ אז לכל $a \in \mathbb{R}$:

$$\lim_{n\to\infty}\Pr\left(Z < a\right) = \Phi\left(a\right)$$

הערה 9.7 ההתפלגות של Z דומה להתפלגות הנורמלית $N\left(0,1\right)$.

משפט 9.8 אי שוויון צ'רנוף (Chernoff)

יהיו $X_1, ..., X_n$ מ"מ מציינים ב"ת, $X = \sum_{i=1}^{n} X_i$. ונסמן: $\mu = \mathbb{E}\left[X\right]$. אזי לכל $\delta > 0$:

$$\begin{aligned}\Pr\left(X \geq a := \left(1+\delta\right)\mu\right) &\leq \left(\frac{e^{\delta}}{\left(1+\delta\right)^{1+\delta}}\right)^{\mu} \\ \Pr\left(X < \left(1-\delta\right)\mu\right) &\leq \left(\frac{e^{-\delta}}{\left(1-\delta\right)^{1-\delta}}\right)^{\mu}\end{aligned}$$

הוכחה: נוכיח רק את אי השוויון הראשון. יהי $t > 0$ שאת ערכו נקבע בהמשך, אז:

$$\Pr\left(X \geq a\right) = \Pr\left(e^{tx} \geq e^{ta}\right) \overset{markov}{\leq} \frac{\mathbb{E}\left[e^{tx}\right]}{e^{ta}}$$

נחשוב על $\mathbb{E}\left[e^{tx}\right]$ כעל פונקציה במשתנה הממשי t, נכתוב את טור הטיילור שלו:

$$\mathbb{E}\left[e^{tx}\right] = \mathbb{E}\left[1 + tx + \frac{t^2x^2}{2} + \cdots\right] = 1 + t\underbrace{\mathbb{E}\left[X\right]}_{\text{first moment}} + \frac{t^2}{2}\underbrace{\mathbb{E}\left[X^2\right]}_{\text{second moment}} + ... := \text{moment generating function}$$

■

10 תרגול 3

10.1 סכומים נבדלים\שונים

הגדרה 10.1 $\{x_1,...,x_k\}=S\subseteq[n]$ נקראת בעלת סכומים נבדלים (או שונים) אם כל המספרים $\sum_{i\in I}x_i,\ I\subseteq[k]$ שונים.

שאלה מה הגודל המרבי של תת קבוצה של $[n]$ בעלת סכומים נבדלים?

תשובה נגדיר פונקציה $f:\mathbb{N}\to\mathbb{N}$ כש

$$\begin{aligned}f(n)=k\text{ s.t. there is a subset of }[n]\text{ of size k, and it has different sums,}\\ \text{and there is no subset of }[n]\text{ with different sums of size }k+1\end{aligned}$$

חסם קל: $\forall n:\ f(n)\le n$

חסם תחתון נסתכל על $S_n:=\left\{1,2,4,...,2^{\lfloor log_2 n\rfloor}\right\}$. תרגיל: S_n בעלת סכומים נבדלים:

$$|S_n|=\lfloor log_2 n\rfloor+1$$
$$f(n)\ge\lfloor log_2 n\rfloor+1$$

חסם עליון נניח $f(n)=k$ ונניח $S=\{x_1,...,x_k\}\subseteq[n]$ בעלת סכומים נבדלים. יש 2^k סכומים שונים וכולם קטנים (ממש) מ-nk. כיוון שכל הסכומים שונים וחיוביים, קיבלנו שיש לכל היותר nk סכומים שונים. כיוון ש-S קבוצה עם סכומים נבדלים, אפשר להפיק ממנה $2^{|S|}=2^k$ סכומים, לכן:

$$2^{f(n)}=2^k\le nk$$
$$\Downarrow log$$
$$f(n)\le log_2 n+log_2 k=log_2 n+log_2 f(n)$$

$$log_2(f(n))\le log_2 log_2\underbrace{nk}_{\le n^2}\le log_2(2log_2(n))=log_2log_2 n+log2=log_2log_2 n+O(1)$$
$$\Downarrow$$
$$f(n)\le log_2 n+log_2log_2 n+O(1)$$

חסם עליון טוב יותר נתחיל שוב עם S כנ"ל. נגדיר $\epsilon_1,...,\epsilon_k\overset{i.i.d}{\sim}Ber\left(\frac{1}{2}\right)$. נגדיר מ"מ $X=\sum_{i=1}^k\epsilon_i x_i$ (ה-x_iים הם איברי S, כמו בהגדרה של S מקודם).

$$\mu=\mathbb{E}[X]=\sum_{i=1}^k\mathbb{E}[\epsilon_i x_i]=\sum_{i=1}^k x_i\mathbb{E}[\epsilon_i]=\frac{x_1+...+x_k}{2}$$
$$\sigma^2=Var[X]=\sum_{i=1}^k x_i^2 Var[\epsilon_i]=\frac{1}{4}\sum_{i=1}^k x_i^2\le\frac{kn^2}{4}\Rightarrow\sigma\le\frac{\sqrt{k}}{2}n$$

לכל $\lambda > 1$ (המשפט נכון ל$\lambda > 0$, אבל נרצה דווקא $\lambda > 1$):

$$\Pr\left(|X-\mu| \geq \lambda\sigma\right) \leq \frac{1}{\lambda^2}$$
$$\Downarrow$$
$$\Pr\left(|X-\mu| < \lambda\sigma\right) \geq 1-\frac{1}{\lambda^2}$$

מהו הגודל של $(\mu-\lambda\sigma,\mu+\lambda\sigma)$? כלומר, כמה מספרים טבעיים נכנסים לכל היותר לקטע הזה? לכל היותר $2\lambda\sigma+1$. לכל $m\in\mathbb{N}$: $\Pr\left(X=m\right)\leq 2^{-k}$ (כי אם ההסתברות היא לא 0, בגלל שזו קבוצה בעלת סכומים נבדלים יש לכל היותר קבוצה אחת בעלת הסכום הזה), לכן:

$$\begin{aligned}
1-\frac{1}{\lambda^2} &\leq \Pr\left(|X-\mu<\lambda\sigma|\right) \leq \frac{2\lambda\sigma+1}{2^k}\\
&\Downarrow\\
1-\frac{1}{\lambda^2} &\leq \frac{2\lambda\sigma+1}{2^k}\\
&\Downarrow\\
2^{f(n)}=2^k &\leq \frac{2\lambda\sigma+1}{1-\frac{1}{\lambda^2}} \leq \frac{2\lambda\frac{\sqrt{k}}{2}n+1}{1-\frac{1}{\lambda^2}}\\
&\Downarrow log_2\\
f\left(n\right) &\leq log_2\left(2\lambda\frac{\sqrt{k}}{2}n+1\right)-log_2\left(1-\frac{1}{\lambda^2}\right)=\\
&=-log_2\left(1-\frac{1}{\lambda^2}\right)+log_2 n+\frac{1}{2}log_2 k+O\left(1\right)\leq\\
&\leq log_2 n+\frac{1}{2}log_2 k+O\left(1\right)\\
&\Downarrow log_2\\
log_2 f\left(n\right) &\leq log_2 log_2\left(\underbrace{n\sqrt{k}}_{\leq n^{\frac{3}{2}}}\right)+O\left(1\right)\leq log_2\left(\frac{3}{2}log_2 n\right)+O\left(1\right)=\\
&=log_2 log_2 n+log_2\frac{3}{2}+O\left(1\right)=log_2 log_2 n+O\left(1\right)
\end{aligned}$$

נציב:

$$f\left(n\right)\leq log_2 n+\frac{1}{2}log_2 log_2 n+O\left(1\right)$$

10.2 תנאי החתונה ב$G\left(n,n,p\right)$

הגדרה 10.2 תנאי החתונה

גרף דו"צ $G=(U,V,E)$ מקיים את תנאי החתונה אם לכל $U \subseteq W$ אז

$$\left|\underbrace{N_G(W)}_{\text{neighbors of W}}\right| \geq |W|$$

11 תרגול אקסטרה 3

11.1 ריכוז מידה

11.1.1 כמעט כל הגרפים הם כמעט רגולרים

הגדרה 11.1 גרף הוא d רגולרי אם לכל קודקודיו יש דרגה d.

הגדרה 11.2 גרף $G=(V,E)$ הוא (d,ϵ)־כמעט־רגולרי אם

$$\forall v \in V: \ d(1-\epsilon) \leq deg(v) \leq d(1+\epsilon)$$

טענה 11.3 אם $G \sim G\left(n,\frac{1}{2}\right)$ אז:

$$\forall \epsilon > 0: Pr_G\left(G \text{ is } \left(\frac{n-1}{2},\epsilon\right)\text{-almost-regular}\right) \overset{n\to\infty}{\to} 1$$

הוכחה: למה $\frac{n-1}{2}$? נסמן בX את הדרגה של איזשהו $v \in V$:

$$X = \sum_{u \in V\setminus\{v\}} X_u$$

$$X_u = \begin{cases} 1 & (v,u) \in E \\ 0 & otherwise \end{cases}$$

$$\Downarrow$$

$$X \sim Bin\left(n-1,\frac{1}{2}\right)$$

$$\Downarrow$$

$$\mathbb{E}[X] = \frac{n-1}{2}$$

נביט במאורע המשלים ונראה מה אנו רוצים להוכיח:

$$\Pr\left(\exists v\in V:\ s.t.\ \underbrace{\left|deg\left(v\right)-\frac{n-1}{2}\right|>\epsilon\frac{n-1}{2}}_{Bv}\right)\overset{n\to\infty}{\to}0$$

$$0\overset{n\to\infty}{\leftarrow}\Pr\left(\cup_{v\in V}Bv\right)\overset{\text{union bound}}{\leq}\sum_{v\in V}\Pr\left(Bv\right)\overset{\text{symmetry between vertices}}{=}n\Pr\left(B_1\right)$$

צ"ל $n\Pr\left(B_1\right)\to 0$**,** או במילים אחרות

$$\Pr\left(B_1\right)=o\left(\frac{1}{n}\right)$$

ננסה את אי שוויון צ'בישב

הדרגה של הקודקוד 1:

$$X\sim Bin\left(n-1,\frac{1}{2}\right)$$

וגם

$$\mathbb{E}\left[X\right]=\frac{n-1}{2}$$
$$Var\left[X\right]=\frac{n-1}{4}$$
$$B_1=\left\{\left|X-\mathbb{E}\left[X\right]\right|>\epsilon\frac{n-1}{2}\right\}$$

לכן:

$$\Pr\left(B_1\right)\overset{chebyshev}{\leq}\frac{Var\left[X\right]}{\left(\epsilon\frac{n-1}{2}\right)^2}=\frac{\frac{n-1}{4}}{\epsilon^2\frac{(n-1)^2}{4}}=\frac{1}{\epsilon^2\left(n-1\right)}\overset{n\to\infty}{\to}0$$

זה נחמד, אבל לא מספיק הדוק בשביל הטענה שלנו.

נשתמש באי שוויון צ'רנוף

מ"מ X הוא סכום של מ"מ מציינים ב"ת ולכן $\mu=\mathbb{E}\left[X\right]$, ולכל $\epsilon>0$:

$$\Pr\left(\left|X-\mu\right|>\epsilon\mu\right)\leq 2e^{-\frac{\epsilon^2\mu}{3}}$$

נפעיל על הדרגה של קדוקוד 1:

$$\Pr\left(B_1\right) \leq 2e^{-\frac{\epsilon^2\frac{n-1}{2}}{3}} = 2e^{-\frac{\epsilon^2(n-1)}{6}}$$

לכן:

$$\Pr\left(G \text{ isn't } \left(\frac{n-1}{2}, \epsilon\right)\text{-almost-regular}\right) \leq n \cdot 2 \cdot e^{-\frac{\epsilon^2(n-1)}{6}} \overset{n\to\infty}{\to} 0$$

■

11.2 מתי גרף מקרי מפסיק להיות "זיווג"

הגדרה 11.4 "זיווג"
גרף שכל קודקודיו בדרגה ≤ 1.

הערה 11.5 צלע כפולה בין שני קודקודים מונעת מגרף להיות זיווג.

נביט בתהליך המקרי הבא:

- מתחילים עם n קודקודים מבודדים
- בכל צעד נבחר זוג קודקודים באקראי ובאופן אחיד, ונחבר ביניהם.

טענה 11.6 נסמן בm את מספר הצעדים שעשינו, אז:

$$\Pr\left(\underbrace{G\left(n,m\right)}_{\text{the graph after m steps}} \text{ is a "matching"}\right) \overset{n\to\infty}{\to} \begin{cases} 1 & m << \sqrt{n} \\ 0 & m >> \sqrt{n} \end{cases}$$

הערה 11.7 נזכר בבעיית הכדורים והסלים:
יש לנו n סלים וm כדורים, ושמים את הכדורים באקראי בסלים. ראינו שאם $m << \sqrt{n}$ אז

$$\Pr\left(\text{there is no bucket with } \geq 2 \text{ balls}\right) \to \begin{cases} 1 & m >> \sqrt{n} \\ 0 & m << \sqrt{n} \end{cases}$$

נשים לב שניתן להקביל את הבעיה שלנו לבעיית הכדורים והסלים: יש לנו n קודקודים וm צלעות, רק שהפעם במקום שכל כדור יכנס לסל, כל צלע "נכנסת" לשני קודקודים (שונים). **הוכחה:** הוכחה לטענה:

$$\begin{aligned}
&\Pr\left(\text{not a "matching"}\right) = \\
&= \Pi_{j=1}^{m} \Pr\left(\overset{\text{the j-th edge didn't hit}}{\text{a non isolated vertex}} | G\left(n, j-1\right) \text{ is a "matching"}\right) = \\
&= 1 \cdot \frac{\binom{n-2}{2}}{\binom{n}{2}} \frac{\binom{n-4}{2}}{\binom{n}{2}} \cdot \ldots \cdot \frac{\binom{n-2(m-1)}{2}}{\binom{n}{2}}
\end{aligned}$$

■

12 תרגיל 3

טענה 12.1 אם

$$X = \#\text{ of fixed points of a permutation chosen uniformly at random}$$

אז:

$$\mathbb{E}[X] = 1 = Var[X]$$

טענה 12.2 אינפי שימושי: טור טיילור של e:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

טענה 12.3 תוחלת של פונקציה יוצרת מומנטים של פואסון:

$$\mathbb{E}\left[e^{tX}\right] = e^{\lambda e^t - \lambda}$$

הגדרה 12.4 זיווג מושלם
גרף $H = (U, V, E)$ עם $|V| = |U|$ מכיל זיווג מושלם אם קיים $M \subseteq E$ כך ש$|M| = |U|$.

משפט 12.5 החתונה של $Hall$
H מכיל זיווג מושלם אם"ם הוא מספק את תנאי החתונה.

הגדרה 12.6 פונקציית האנטרופיה

$$f(x) = \begin{cases} -x \log_2 x - (1-x) \log_2 (1-x) & x \in (0,1) \\ 0 & x = 0, 1 \end{cases}$$

טענה 12.7 פונקציית האנטרופיה רציפה ב$[0,1]$ ומקבלת מקסימום ב$\frac{1}{2}$.

13 שיעור 4

13.1 עוד הסתברות

13.1.1 אי שוויון $Chernoff$

טענה 13.1 אי שוויון $Chernoff$

יהיו $X_1, ..., X_n$ מ"מ מציינים ב"ת ויהיה $X = \sum_{i=1}^{n} X_i$, נסמן $\mathbb{E}[X] = \mu$ ונבחר $\delta > 0$, אז:

$$\Pr(X \geq (1+\delta)\mu) \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu}$$
$$\Pr(X \leq (1-\delta)\mu) \leq \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu}$$

הערה 13.2 בדיקה מהירה: לכל $\delta > 0$:

$$\begin{aligned} & e^{\delta} < (1+\delta)^{1+\delta} \\ & \Updownarrow \\ & \delta < (1+\delta)\, ln(1+\delta) \ \text{(for } \delta = 0 \text{ both sides=0)} \\ & \Downarrow deriviation \\ & 1 < ln(1+\delta) + 1 \\ & \Downarrow \\ & 0 < (1+\delta)^{1+\delta} - e^{\delta} \text{ is rising} \\ & \Downarrow \\ \delta > 0 \Rightarrow & \frac{e^{\delta}}{(1+\delta)^{1+\delta}} < 1 \end{aligned}$$

הערה 13.3 נזכיר שאם X מ"מ אז e^{tX} היא פונקציית יוצרת המומנטים של X, ז"א:

$$e^{tX} = 1 + tX + \frac{t^2}{2}X + \frac{t^3}{3!}X^3 + ...$$
$$\Downarrow$$
$$\mathbb{E}\left[e^{tx}\right] = 1 + t\mathbb{E}[X] + \frac{t^2}{2}\mathbb{E}\left[X^2\right] + ...$$

הוכחה: אי שוויון $Chernoff$
נסמן: $a := (1+\delta)\mu$, אז:

$$\Pr(X \geq a) = \Pr\left(e^{tx} \geq e^{ta}\right)$$

נגדיר:

$$c := \frac{e^{ta}}{\mathbb{E}[e^{tx}]}$$
$$\Downarrow$$
$$\Pr\left(e^{tx} \geq e^{ta}\right) = \Pr\left(e^{tx} \geq c\mathbb{E}\left[e^{tx}\right]\right) \overset{markov}{\leq} \frac{\mathbb{E}[e^{tx}]}{e^{ta}}$$

ובסוף נבחר ערך של t שיקטין את החסם ככל האפשר. נסמן:

$$X=\sum X_i$$
$$\mathbb{E}\left[X_i\right]=p_i \Leftrightarrow \Pr\left(X_i=1\right)=p_i$$
$$\Downarrow$$
$$\mathbb{E}\left[e^{tx}\right]=\mathbb{E}\left[e^{t\sum X_i}\right]=\mathbb{E}\left[\Pi e^{tX_i}\right] \overset{X_i \text{ are i.i.d so } e^{tX_i} \text{ are i.i.d}}{=} \Pi_{i=1}^{n}\mathbb{E}\left[e^{tX_i}\right]$$

נשים לב:

$$\mathbb{E}\left[e^{tX_i}\right]=\underbrace{p_i\cdot e^t}_{X_i=1}+\underbrace{\left(1-p_i\right)}_{X_i=0}=1+p_i\left(e^t-1\right)\le e^{p_i\left(e^t-1\right)}$$
$$\Downarrow$$
$$\mathbb{E}\left[e^{tX}\right]\le \Pi e^{p_i\left(e^y-1\right)}=exp\left(\left(e^t-1\right)\underbrace{\sum p_i}_{=\mu}\right)$$
$$\Downarrow$$
$$\Pr\left(X\ge a\right)\le\frac{exp\left(\left(e^t-1\right)\mu\right)}{e^{ta}}=exp\left(\mu\left(\left(e^t-1\right)-t\left(1+\delta\right)\right)\right)$$

נחפש t שימזער את צד ימין. נגזור ונשווה לאפס:

$$e^t=1+\delta$$
$$\Downarrow$$
$$t=ln\left(1+\delta\right)$$
$$\Downarrow$$
$$\Pr\left(X\ge a\right)\le exp\left(\mu\left(\delta-\left(1+\delta\right)ln\left(1+\delta\right)\right)\right)=\left(\frac{e^\delta}{\left(1+\delta\right)^{1+\delta}}\right)^\mu$$

■

13.1.2 אי תלות של מ"מ

Ω מרחב הסתברות, $X;Y$ מ"מ מציינים עליו:

$$\Omega=\begin{array}{cc} X=0,\ Y=0 & X=0,\ Y=1\\ X=1,\ Y=0 & X=1,\ Y=1\end{array}$$

נזכר ש

$$X=1\Leftrightarrow\left\{\omega\in\Omega|X\left(\omega\right)=1\right\}$$

באופן כללי, אם X,Y מ"מ על מרחב הסתברות Ω ואם לכל S (מדידה) חלקית לטווח של X וT (מדידה) חלקית לטווח של Y, אומרים שX,Y ב"ת אם מתקיים:

$$\Pr\left((X\in S)\wedge(Y\in T)\right)=\Pr\left(X\in S\right)\Pr\left(Y\in T\right)$$

13.2 מרטינגלים (המרטינגל של Doob)

החומר הנ"ל נמצא בProbability & Computing של Upfal & Mitzenmacher.

הגדרה 13.4 עידון

$\sigma > \Pi$. נגיד שΠ מעדנת את σ אם שתיהן חלוקות ולכל חלק S בσ יש מס' חלקים $P_1,...,P_1$ בחלוקה Π כך ש$S=\cup_{i=1}^{r}P_i$.

הגדרה 13.5 מרטינגל (ע"פ $Doob$)

יהי Ω מרחב הסתברות, יהי $X:\Omega\to\mathbb{R}$ מ"מ ממשי. נניח שמוגדרת על Ω סדרה של חלוקות הולכות ומתעדנות. Π_0 היא החלוקה הטריוויאלית, שמעודנת על ידי Π_1, וכו'. בזה מוגדרת סידרה של מ"מ

$$X_0=\mathbb{E}\left[X\right]$$
$$X_k=\mathbb{E}\left[X|\Pi_k\right]$$

כשX_k זהו מ"מ שקבוע על כל חלק בחלוקה Π_k וערכו הוא התוחלת המותנית של X בחלק הזה.

הערה 13.6 מביטים במהמר שמתחיל עם הון התחלתי של X_0. נניח שהוא מהמר כנגד גוף הוגן, ז"א תוחלת הרווח = 0. יש לו הימור Z_1, ואחריו הונו משתנה לX_1. התהליך נמשך והוא נוקט צעד Z_k (אקראי) בזמן j. ההגינות של הקזינו מתבטאת בכך ש:

$$\forall i:\ \mathbb{E}\left[X_i|Z_1,...,Z_{i-1}\right]=X_{i-1}$$

13.2.1 אי שוויון $Azuma$

משפט 13.7 גרסה לא ברורה

אם יש לנו מרטינגל $X_0,X_1,...$ כך שההפרשים X_i-X_{i-1} "קטנים", אז X מקיים משפט ריכוז מידה מעריכי.

משפט 13.8 אי שוויון $Azuma$

יהיה $X_0,X_1,...,X_n$ מרטינגל ונניח שלכל i מתקיים: $c_i\geq|X_i-X_{i-1}|$. אז לכל $\lambda>0$:

$$\Pr\left(|X_n-X_0|\geq\lambda\right)\leq exp\left(-\frac{\lambda^2}{\sum_{i=1}^{n}c_i^2}\right)$$

הוכחה: בה"כ נניח ש$X_0 = 0$ (נחליף כל X_j ב$X_j - X_0$ ונמשיך). נסמן:

$$\begin{aligned} Y_i &= X_i - X_{i-1} \\ &\Downarrow \\ \forall i : |Y_i| &\le c_i \\ &\Downarrow \text{by definition} \\ \mathbb{E}\left[Y_i | X_0, ..., X_{i-1}\right] &= 0 \end{aligned}$$

נביט ב:$\mathbb{E}\left[e^{\alpha Y_i} | X_0, ..., X_{i-1}\right]$. חשבון:

$$e^{\alpha Y_i} \le \frac{e^{\alpha c_i} + e^{-\alpha c_i}}{2} + \frac{Y_i}{2c_i}\left(e^{\alpha c_i} - e^{-\alpha c_i}\right)$$

טענה 13.9 תרגיל באינפי:

$$\forall t : \underbrace{\frac{e^t + e^{-t}}{2}}_{LHS} < \underbrace{e^{\frac{t^2}{2}}}_{RHS}$$

הוכחה: נפתח את שני האגפים לטורי טיילור:

$$LHS = \frac{1}{2}\left(\left(1 + t + \frac{t^2}{2!} + ...\right) + \left(1 - t + \frac{t^2}{2!} - \frac{t^3}{3!} + ...\right)\right) = 1 + \frac{t^2}{2!} + \frac{t^4}{4!} + ... + \frac{t^{2k}}{(2k)!}$$
$$RHS = 1 + \frac{t^2}{2} + \frac{t^4}{2! \cdot 4} + \frac{t^6}{3! \cdot 8} + ... + \frac{t^{2k}}{k! \cdot 2^k}$$

נשים לב ש

$$\frac{1}{(2k)!} \le \frac{1}{k!2^k}$$

כי

$$1 < \frac{(k+1) \cdot ... \cdot 2k}{2 \cdot ... \cdot 2}$$

■

בחזרה להוכחה הקודמת:

$$Y_i = \frac{1 - \frac{Y_i}{c_i}}{2} \cdot (-c_i) + \frac{1 + \frac{Y_i}{c_i}}{2} c_i$$

הצגנו את Y_i כצירוף קמור של $-c_i$ ו-c_i. הפונקציה $e^{\alpha Y_i}$ (כפונקציה של α) היא קמורה.

$$e^{\alpha Y_i} \leq \frac{1-\frac{Y_i}{c_i}}{2} \cdot e^{-\alpha c_i} + \frac{1+\frac{Y_i}{c_i}}{2} \cdot e^{\alpha c_i}$$

נחשב תוחלת מותנית על $X_0, ..., X_{i-1}$ ונקבל:

$$\mathbb{E}\left[e^{\alpha Y_i} | X_0, .., X_{i-1}\right] \leq \frac{e^{\alpha c_i} + e^{-\alpha c_i}}{2} \overset{\text{infi exercise}}{\leq} e^{\frac{\alpha^2 c_i^2}{2}}$$

$$\Downarrow$$

$$\mathbb{E}\left[e^{\alpha(X_n - X_0)}\right] \overset{\text{telescopic sum}}{=} \mathbb{E}\left[\Pi_{i=1}^{n} e^{\alpha Y_i}\right] =^{*} \mathbb{E}\left[\Pi_{i=1}^{n-1} e^{\alpha Y_i}\right] \mathbb{E}\left[e^{\alpha Y_n} | X_0, ..., X_{n-1}\right] \leq$$

$$\leq \mathbb{E}\left[\Pi_{i=1}^{n-1} e^{\alpha Y_i}\right] e^{\frac{\alpha^2 c_n^2}{2}} \overset{\text{induction}}{\leq} e^{\alpha^2 \sum_{j=1}^{n} \frac{c_j^2}{2}}$$

$$\Downarrow$$

$$\Pr\left(X_n - X_0 \geq \lambda\right) = \Pr\left(e^{\alpha(X_n - X_0)} \geq e^{\alpha \lambda}\right) \overset{\text{markov}}{\leq} \frac{\mathbb{E}\left[e^{\alpha(X_n - X_0)}\right]}{e^{\alpha \lambda}} \leq e^{\alpha^2 \frac{\sum c_j^2}{2} - \alpha \lambda}$$

אגף ימין מתמזער כאשר $\alpha = \frac{\lambda}{\sum c_j^2}$ וזה נותן בדיוק את מה שרצינו. ■

13.2.2 בעיה עם מרטינגיילים

הגדרה 13.10 מספר צביעה
המס' המינימלי של צבעים כך ששני קודקודים סמוכים צבועים בצבעים שונים.

מהו מס' הצביעה האופייני של גרף ב-$G\left(n, \frac{1}{2}\right)$?
ראינו שהגודל המירבי של אנטיקליקה ב-$G\left(n, \frac{1}{2}\right)$ הוא $\left(2 + o\left(1\right)\right) log_2\left(n\right)$. מכאן נובע שנדרשים לפחות $\frac{n}{2 log_2 n}$ צבעים.
בעזרת $Azuma$, אחד B. Bollobas הראה משפט חזק יותר: שכמעט בכל גרף ב-$G\left(n, \frac{1}{2}\right)$ גם בכל קבוצה של לפחות $\frac{n}{log^2 n}$ קודקודים יש אנטיקליקה ומגודל $\left(2 + o\left(1\right)\right) log_2 n$.

14 תרגול 4

14.1 חסמי צ'רנוף

משפט 14.1 יהיו $X_1,...,X_n$ מ"מ המקבלים ערכים ב$\{0,1\}$ (אינדיקטורים), נניח שכולם ב"ת. אזי, לכל $\epsilon\in(0,1)$:

$$X=\sum_{i=1}^{n}X_i$$
$$\Pr\left(X\le(1-\epsilon)\mathbb{E}\left[X\right]\right)\le\exp\left(-\frac{\epsilon^2}{2}\mathbb{E}\left[X\right]\right)$$
$$\Pr\left(X\ge(1+\epsilon)\mathbb{E}\left[X\right]\right)\le\exp\left(-\frac{\epsilon^2}{3}\mathbb{E}\left[X\right]\right)$$
$$\Downarrow$$
$$\Pr\left(X\notin(1\pm\epsilon)\mathbb{E}\left[X\right]\right)\le 2\exp\left(-\frac{\epsilon^2}{3}\mathbb{E}\left[X\right]\right)$$

14.2 תכונות אופייניות של $G\left(n,p\right)$

טענה 14.2 יהי $G\sim G\left(n,p\right)$ ויהי $e\left(G\right)$ מס' הצלעות בG, אזי:

$$\Pr\left(e\left(G\right)\notin(1\pm\epsilon)\binom{n}{2}p\right)\le 2\exp\left(-\frac{\epsilon^2}{3}\mathbb{E}\left[X\right]\right)$$

הוכחה: אפשר לכתוב $e\left(G\right)=\sum_{1\le i<j\le n}X_{ij}$ כאשר X_{ij} הוא האינדיקטור של המאורע $\{i,j\}\in E$. נציב בצ'רנוף, ונסיים. ■

מסקנה 14.3 אם $p=\omega\left(\frac{1}{n^2}\right)$ (נזכר: $\frac{f(n)}{g(n)}\to\infty\Leftrightarrow f\left(n\right)=\omega\left(g\left(n\right)\right)$) אז לכל $\epsilon>0$ אכ"ב (אסימפטוטית כמעט בוודאות) ל$G\left(n,p\right)$ יש $\frac{n^2}{2}p$ $(1\pm\epsilon)$ קשתות.

הוכחה: נשים לב: עבור n מספיק גדול

$$(1\pm\epsilon)\frac{n^2}{2}p\subseteq\left(1\pm\frac{\epsilon}{2}\right)\binom{n}{2}p$$

לכן:

$$\Pr\left(e\left(G\right)\notin(1\pm\epsilon)\frac{n^2}{2}p\right)\le\Pr\left(e\left(G\right)\notin\left(1\pm\frac{\epsilon}{2}\right)\binom{n}{2}p\right)\le$$
$$\le 2\exp\left(-\frac{\epsilon^2}{3\cdot 4}\underbrace{\frac{n\left(n-1\right)p}{2}}_{\substack{\to\infty\\ \text{because p}=\omega\left(\frac{1}{n^2}\right)}}\right)\overset{n\to\infty}{\to}0$$

■

טענה 14.4 אם $G \sim G(n,p)$ ו$\epsilon \in (0,1)$, ואם לכל $i \in [n]$ נסמן ב$d(i)$ את הדרגה של קודקוד i, אזי:

$$\Pr\left(\exists i \in [n]:\ d(i) \notin (1 \pm \epsilon)(n-1)p\right) \le 2n \exp\left(-\frac{\epsilon^2}{3}(n-1)p\right)$$

הוכחה: נשים לב

$$\begin{aligned}
&d(i) \sim Bin(n-1,p)\\
&\qquad \Downarrow chernoff\\
&\forall i: \Pr\left(d(i) \notin (1 \pm \epsilon)(n-1)p\right) \le 2\exp\left(-\frac{\epsilon^2}{3}(n-1)p\right)\\
&\qquad \Downarrow \text{union bound}\\
&\qquad \Pr\left(\exists i \in [n]:\ d(i) \notin (1 \pm \epsilon)(n-1)p\right) \le \sum_{i=1}^{n} \Pr\left(d(i) \notin (1 \pm \epsilon)(n-1)p\right) \le\\
&\qquad \le 2n \exp\left(-\frac{\epsilon^2}{3}(n-1)p\right)
\end{aligned}$$

■

מסקנה 14.5 אם $p = \omega\left(\frac{log(n)}{n}\right)$ אז לכל $\epsilon > 0$ אכ"ב לכל הקודקודים ב$G(n,p)$ יש דרגה $(1 \pm \epsilon)np$.

הוכחה: נשים לב:

$$\begin{aligned}
\Pr\left(\exists i:\ d(i) \notin (1 \pm \epsilon)np\right) &\underset{\text{for large enough n}}{\le} \Pr\left(\exists i:\ d(i) \notin \left(1 \pm \frac{\epsilon}{2}\right)(n-1)p\right) \le\\
&\le 2n \exp\left(-\frac{\epsilon^2}{12}(n-1)p\right) = 2\exp\left(-\frac{\epsilon^2}{12}(n-1)p + \log(n)\right) =\\
&= 2\exp\left(-O(1)\underbrace{\frac{n}{\log n}}_{\to\infty} p \left(\underbrace{\log n}_{\to\infty} - \underbrace{\frac{\log n}{n}}_{\to 0} - \underbrace{\frac{\log^2 n}{pn}}_{=\log n \underbrace{\frac{\log n}{pn}}_{\to 0}}\right)\right) =\\
&= 2\exp\left(-O(1)\frac{np}{\log n}\left(\log n(1 - o(1)) - o(1)\right)\right) \to 0
\end{aligned}$$

■

15 תרגול אקסטרה 4

15.1 א"ש אזומה ושימושיו

הערה 15.1 תזכורת לc־ליפשיציות
אם f היא c־ליפשיצית אז:

$$\forall x_1,..,x_n,\hat{x_i}\neq x_i:\ \left|f\left(x_1,..,x_n\right)-f\left(x_1,...,\hat{x_i},..,x_n\right)\right|\leq c$$

משפט 15.2 גרסה ידידותית של אזומה

אם $X=\left(X_1,..,X_m\right)$ הם מ"מ ב"ת, וf היא פונקציה c־ליפשיצית ו$t>0$ אז:

$$\Pr\left(f\left(x\right)\geq\mathbb{E}\left[f\left(x\right)\right]+t\right)\leq e^{-\frac{2t^2}{mc^2}}$$

טענה 15.3 מקרה פרטי
אם $f=\sum_{i=1}^{m}X_i$, אזי f היא 1־ליפשיצית ומקבלים את חסם צ'רנוף עבור משתנה בינומי.

15.1.1 משולשים ב$G\left(n,\frac{1}{2}\right)$

נסמן בT את מס' המשולשים ב$G\sim G\left(n,\frac{1}{2}\right)$. נזכור:

$$T_{i,j,k}=\begin{cases}1 & i,j,k\text{ form a triangle}\Leftrightarrow\left\{i,j\right\},\left\{i,k\right\},\left\{j,k\right\}\in\text{E}\\ 0 & otherwise\end{cases}$$
$$T=\sum_{i,j,k}T_{i,j,k}$$

נשים לב ש$T_{i,j,k}$ אינם ב"ת (למשל נסתכל על 4 קודקודים, בהינתן שיודעים ש3 חולקים משולש ההסתברות שקיים משולש נוסף גבוהה יותר) ולכן לא ניתן להשתמש בצ'רנוף. נרצה:

$$\Pr\left(T\geq\left(1+\epsilon\right)\mathbb{E}\left[T\right]\right)\leq ?$$

היום נראה שבמקרים מסויימים, ניתן להציג סכום של מ"מ שאינם ב"ת כפונקציה ליפשיצית של מ"מ מקריים אחרים ב"ת.
נבחר:

$$X=\left(X_{ij}\right)_{1\leq i<j\leq n}$$

כאשר X_{ij} מ"מ מציין האם $\left\{i,j\right\}\in E$. $m=\binom{n}{2}$. נגדיר:

$$f\left(x\right)=T=\text{number of triangles in the graph}$$

בצורה מפורשת:

$$f\left(x\right)=\sum_{i,j,k}X_{ij}X_{ik}X_{jk}$$

טענה 15.4 f היא $\left(n-2\right)$־ליפשיצית.

הוכחה: כל צלע משתתפת בלכל היותר $n-2$ משולשים פוטנציאליים. ■

נפעיל את אזומה:

$$\Pr\left(T\geq\left(1+\epsilon\right)\mathbb{E}\left[T\right]\right)=\Pr\left(T\geq\left(1+\epsilon\right)\frac{1}{8}\binom{n}{3}\right)\leq e^{\frac{-2\epsilon^{2}\binom{n}{3}^{2}}{64\binom{n}{2}\left(n-2\right)^{2}}}=e^{-\theta\left(n^{2}\right)}$$

נפתח את המעריך ונתעלם מקבועים:

$$\frac{-2\epsilon^{2}\binom{n}{3}^{2}}{64\binom{n}{2}\left(n-2\right)^{2}}=-\theta\left(\frac{n^{6}}{n^{4}}\right)=-\theta\left(n^{2}\right)$$

טענה 15.5 המעריך הדוק:

$$\Pr\left(T\geq\left(1+\epsilon\right)\mathbb{E}\left[T\right]\right)\geq e^{-\theta\left(n^{2}\right)}$$

הוכחה: נשים לב: G גרף שלם $\Leftarrow$ $T=\binom{n}{3}$, לכן:

$$\Pr\left(T\geq\left(1+\epsilon\right)\mathbb{E}\left[T\right]\right)\geq\Pr\left(\text{G is the complete graph}\right)=\left(\frac{1}{2}\right)^{\binom{n}{2}}=e^{-\theta\left(n^{2}\right)}$$

■

15.1.2 תבניות במחרוזות מקריות

נתון אלפבית Σ בגודל s ומילה $y\in\Sigma^k$ באורך קבוע k. כמה עותקים (=תת מילה רציפה) של y נמצא במחרוזת מקרית X באורך n? נסמן בM את מס' העותקים של y. נציג את M כסכום של מ"מ מציינים:

$$M_i=\begin{cases}1 & (x_i,...,x_{i+k-1})=y\\ 0 & otherwise\end{cases}$$

$$M=\sum_{i=1}^{n-k+1}M_i$$

$$\Downarrow$$

$$\Pr(M_i)=\left(\frac{1}{s}\right)^k=\mathbb{E}[M_i]$$

$$\Downarrow$$

$$\mathbb{E}[M]=(n-k+1)\left(\frac{1}{s}\right)^k$$

נרצה לחשב:

$$\Pr(M\geq(1+\epsilon)\mathbb{E}[M])\leq ?$$

נשים לב: $\{M_i\}_i$ אינם ב"ת! נגדיר:

$$f(X)=\#\text{ of copies of y in X}$$

נשים לב:

$$X=(X_1,...,X_n)$$
$$X_i=\text{i-th letter in X}$$

טענה 15.6 f היא k-ליפשיצית.

נפעיל את א"ש אזומה:

$$\Pr(M\geq(1+\epsilon)\mathbb{E}[M])\leq e^{-\theta\left(\frac{2\epsilon^2n^2}{nk^2}\right)}=e^{-\theta(n)}$$

16 תרגיל 4

טענה 16.1 יהיו $X_1,...,X_n$ מ"מ $i.i.d$ המקבלים מס' סופי של ערכים, יהי $X=\sum_{i=1}^n X_i$. קיים $C>0$ שלא תלוי ב n כך שלכל $\epsilon\in(0,1)$:

$$\Pr(X\notin(1\pm\epsilon)\mathbb{E}[X])\leq C\exp\left(-\frac{1}{C}\epsilon^2n\right)$$

16.1 מרטינגיילים

הגדרה 16.2 מרטינגיילים
סדרה סופית או ספירה $X_1, X_2, ...$ מוגדרת על אותו מרחב הסתברות נקראת מרטינגייל אם לכל n:

$$\mathbb{E}\left[X_{n+1}|X_1, X_2, ..., X_n\right] = X_n$$

משפט 16.3 אי שוויון Azuma-Hoeffding
אם $X_1, X_2, ...$ הוא מרטינגייל וקיימים קבועים חיוביים $c_1..., c_{n-1}$ כך שלכל k:

$$|X_{k+1} - X_k| \leq c_k$$

אז אי שוויון $Azuma - Hoeffding$ אומר לנו שלכל $\lambda > 0$ ולכל n:

$$\Pr\left(|X_n - X_1| \geq \lambda\right) \leq 2\exp\left(-\frac{\lambda^2}{2\sum_{k=1}^{n-1} c_k^2}\right)$$

17 בוחן 1

טענה 17.1 עבור X מ"מ המקבל ערכים מבין $\{0, 1, 2....\}$ ו$\mathbb{E}\left[X\right] < \infty$ מתקיים:

$$\Pr\left(X > 0\right) \leq \mathbb{E}\left[X\right]$$

חלק II
אלגברה לינארית

18 שיעור 5

18.1 הגדרות בסיס

הגדרה 18.1 מרחב וקטורי
V מרחב וקטורי מעל שדה $\mathbb{F}$. איברי V נקראים וקטורים ואיברי $\mathbb{F}$ נקראים סקלרים.
יש איבר $0 \in V$ כך ש $\forall v \in V: \ v + 0 = v$ תחת פעולת החיבור שבקרוב נגדיר.
V הוא מבנה אלגברי עם הפעולות הבאות:

1. $+$ הוא חיבור וקטורי. קומוטטיבי, אסוציאטיבי.
2. ויש פעולה של כפל וקטור בסקלר: $v \in V, \ \alpha \in \mathbb{F}: \ \alpha v$

הגדרה 18.2 $T: \ V_{\mathbb{F}} \to W_{\mathbb{F}}$ נקראת **העתקה לינארית** אם:

$$T\left(u +_V v\right) = Tu +_W Tv$$
$$T\left(\alpha v\right) = \alpha Tv$$

הגדרה 18.3 תלות לינארית
אומרים שהוקטורים $v_1, .., v_n \in V$ **תלויים לינארית** אם יש סקלארים $\alpha_1, .., \alpha_n \in \mathbb{F}$ (לא כולם אפס) כך ש$\sum_{i=1}^{n} \alpha_i v_i = 0$

הגדרה 18.4 **המימד** של V זהו הn הגדול ביותר כך שיש $v_1, ..., v_n \in V$ שהם בת"ל, ומסמנים $\dim V = n$. אם אין n כזה, אומרים שV הוא אינסוף מימדי: $\dim V = \infty$.

הגדרה 18.5 אם $\dim V = n$ אז קבוצה של n וקטורים בת"ל נקראת **בסיס**.

משפט 18.6 לבסיס $v_1, ..., v_n$ של מרחב וקטורי V יש התכונה שלכל וקטור $u \in V$ יש ייצוג אחד ויחיד מהצורה $u = \beta_1 v_1 + ... + \beta_n v_n$.

הערה 18.7 יוצא שאם $\mathbb{F}$ שדה כלשהו אז יש מרחב וקטורי n מימדי קנוני מעל $\mathbb{F}$ שנסמן ב: $V = \mathbb{F}^n$:

$$V = \left\{\left(\gamma_1, ..., \gamma_n\right) | \forall i: \ \gamma_i \in \mathbb{F}\right\}$$

$$\text{Operations:}$$
$$\left(\gamma_1, ..., \gamma_n\right) + \left(\gamma_1', ..., \gamma_n'\right) = \left(\gamma_1 + \gamma_1', ..., \gamma_n + \gamma_n'\right)$$
$$\alpha\left(\gamma_1, ..., \gamma_n\right) = \left(\alpha\gamma_1, ..., \alpha\gamma_n\right)$$

הערה 18.8 כאשר עובדים עם מרחבים וקטורים המוגדרים כך יש גם תיאור קונקרטי של העתקות לינאריות $V \to W$ באמצעות מטריצות:

$$\dim V = n,\ \dim W = m$$
$$(a_{ij} \in \mathbb{F})_{m \times n}$$
$$\begin{pmatrix} a_{11} & ... & a_{1n} \\ & ... & \\ a_{m1} & ... & a_{mn} \end{pmatrix} \underbrace{\begin{pmatrix} x_1 \\ ... \\ x_n \end{pmatrix}}_{\in V} = \underbrace{\begin{pmatrix} y_1 \\ ... \\ y_n \end{pmatrix}}_{\in W}$$
$$y_i = \sum_{j=1}^{n} a_{ij} x_j$$

18.2 חבורות

הגדרה 18.9 חבורה סופית
זו קבוצה G עם פעולה $\cdot$. יש בG איבר מיוחד שנקרא איבר היחידה ונסמנו ב$id = id_G$ עם התכונה:

$$\forall g \in G : id \cdot g = g$$
$$\forall g \in G : g \cdot id = g$$

הכפל הוא אסוציאטיבי ולכל איבר $g \in G$ יש הופכי g^{-1} המקיים:

$$g \cdot g^{-1} = id$$

אומרים שהחבורה G קומוטטיבית ("אבלית") אם:

$$\forall g, h \in G :\ \ g \cdot h = h \cdot g$$

הערה 18.10 דוגמה לחבורה קומוטטיבית: חיבור מודולו n.
דוגמה לחבורה לא קומוטטיבית: S_n, "החבורה הסימטרית". זו החבורה של כל התמורות (פרמוטציות): $\pi :\ [n] \to [n]$ עם פעולת ההרכבה.

הגדרה 18.11 תת חבורה
$H \leq G$ היא חבורה בפני עצמה עם אותה הפעולה:

$$h \in H \Rightarrow h^{-1} \in H$$
$$h_1, h_2 \in H \Rightarrow h_1 \cdot h_2 \in H$$
$$id_G = id_H$$

הגדרה 18.12 חבורות נורמליות
יש תת חבורות $H \leq G$ שמקיימות תנאי נוסף שנקרא נורמליות: "H היא תת חבורה נורמלית של G". כשזה המצב אפשר לראות את G כמכפלה של H בחבורת המנה: G/H.

הגדרה 18.13 חבורה פשוטה
אם לG אין תת חבורות נורמליות, אומרים שG **חבורה פשוטה**. יש משפט ("משפט המיון של חבורות פשוטות סופיות") המתאר באופן מלא מיהן כל החבורות הפשוטות הסופיות.

18.3 חזרה ללינארית

משפט 18.14 משפט הSVD
"כל העתקה לינארית היא הרכבה של מתיחות לאורך צירים ושל סיבובים".

הגדרה 18.15 דרגת שורות של מטריצה
נניח A היא מטריצה, אז דרגת השורות של A היא הגודל המירבי של קבוצת שורות שהיא בת"ל.

הגדרה 18.16 דרגת עמודות של מטריצה
נניח A היא מטריצה, אז דרגת העמודות של A היא הגודל המירבי של קבוצת עמודות שהיא בת"ל.

משפט 18.17 דרגה של מטריצה
דרגת השורות שווה לדרגת העמודות.

הגדרה 18.18 נאמר שמטריצה $A_{m\times n}$ היא מדרגה 1 אם יש שני וקטורים u_m, v_n כך ש

$$A = u \otimes v$$
$$a_{ij} = u_i v_i$$

הערה 18.19 קל לראות שכל מטריצה אפשר להציג כסכום של מטריצות מדרגה 1.

הגדרה 18.20 דרגה של מטריצה
מגדירים דרגה של מטריצה A כמספר המזערי של מטריצות מדרגה 1 שסכומן הוא A.

משפט 18.21 הדרגה של מטריצה = מס' המירבי של שורות\עמודות במטריצה שהן בת"ל.

הגדרה 18.22 מתיחה
אם $x \in V$ ו$T : V \to V$ טרנספורמציה לינארית, נאמר שT מותחת פי λ בכיוון x אם $Tx = \lambda x$.

הערה 18.23 בטרמינולוגיה המקובלת: x הוא וקטור עצמי (eigenvector) של T עם ערך עצמי (eigenvalue) λ.

הערה 18.24 נביט במשוואה $Tx = \lambda x$:

$$\underbrace{(Ax)_i}_{vector} = \sum_j a_{ij}x_j$$
$$\underbrace{(yA)_j}_{vector} = \sum_i y_i a_{ij}$$

הגדרה 18.25 גרעין
אם $T_{m\times n}:\ V_n \to W_m$ הגרעין (הימני) של T הוא:

$$rker\left(T\right)=\left\{x\in V \mid Tx=0_W\right\}$$

אם A מטריצה, אז בדומה:

$$lker\left(A\right)=\left\{z\in V \mid zA=0\right\}$$

הערה 18.26 נשים לב ש $Tx=\lambda x \Leftrightarrow \left(\lambda I-T\right)x=0$.
"T מותחת פי λ בכיוון x" שקול ל$x\in rker\left(\lambda I-T\right)$.

הערה 18.27 אם A מטריצה אז $r\ker A$ הוא מרחב לינארי ($r\ker A$ הוא תת מרחב של A) וגם $l\ker A$ הוא מרחב לינארי.

הערה 18.28 אם A מטריצה ריבועית , לשניהם יש אותו המימד:

$$\dim\left(l\ker A\right)=\dim\left(r\ker A\right)=\dim\left(\ker A\right)=n-rank\left(A\right)$$

טענה 18.29

$$\det\left(A\right)=0\Leftrightarrow rank\left(A\right)<n$$

טענה 18.30

$$\det\left(\lambda I-T\right)=0\Leftrightarrow \lambda \text{ is an eigenvalue of T}$$

הגדרה 18.31 פולינום אופייני
אם T_{nm} אז הפולינום האופייני שלה הוא פולינום ממעלה n במשתנה λ שמוגדר כך:

$$\det\left(\lambda I-T\right)=\det\begin{pmatrix}\lambda-t_{11} & -t_{12} & ... & -t_{1n}\\ ... & ... & ... & ...\\ -t_{n1} & ... & ... & \lambda-t_{nm}\end{pmatrix}$$

הערה 18.32 T מותחת בכיוון x פי λ אס"ם x הוא ו"ע של T עם ע"ע λ אס"ם $\det\left(\lambda I-T\right)=0$ ו$\left(\lambda I-T\right)x=0$.

הגדרה 18.33 סיבוב
העתקה לינארית שמשמרת אורכים, זוויות $\Leftrightarrow$ מכפלות פנימיות.

הגדרה 18.34 מכפלה פנימית

$$x,y\in\mathbb{F}^{n}:\ \langle x,y\rangle:=\sum x_{i}y_{i}$$
$$x,y\in\mathbb{C}^{n}:\ \langle x,y\rangle:=\sum x_{i}\overline{y_{i}}$$
$$\langle x,y\rangle=\|x\|\cdot\|y\|\cos\theta$$

כשθ היא הזוויות שבין הוקטור x לוקטור y.

הערה 18.35 כיוון ש$\|x\|_{2}^{2}=\langle x,x\rangle$ אז שמירה על מכפלה פנימית מבטיחה שמירה על אורכי הוקטורים. מההגדרה עם הזווית של מכפלה פנימית יוצא ששמירה על מכפלות פנימיות גורר שמירה גם על הזויות.

מסקנה 18.36 סיבוב זו טרנספורמציה לינארית ששומרת על המכפלות הפנימיות. כלומר:

$$\forall x,y\in V:\ \langle Ax,Ay\rangle=\langle x,y\rangle$$

$$\begin{aligned} yA^{T}Ax&=\langle x,y\rangle\\ &\Updownarrow\\ \forall y\in V:yA^{T}A&=y\\ &\Updownarrow\\ A^{T}A&=I\end{aligned}$$

הערה 18.37 מה שנאמר עכשיו תקף גם למטריצות מלבניות: אם $A_{m\times n}$ אז $AA^{T}_{m\times m}$ זו מטריצה שהאיבר הi,j שלה הוא המכפלה הפנימית של השורה הi בA עם השורה הj בA.

$A^{T}A_{n\times n}$ זו מטריצה שהאיבר הi,j שלה הוא המכפלה הפנימית של העמודה הi בA עם העמודה הj בA.

לכן, $A^{T}A=I$ פירושו שכל עמודה בA היא וקטור יחידה (כשמסתכלים על האלכסון $A^{T}A$), וכל שתי עמודות ניצבות זו לזו.

מסקנה 18.38 A היא מטריצה אורתוגונלית.

הגדרה 18.39 מטריצה אורתוגונלית

$A\in M\left(\mathbb{R}\right)$ תקרא אורתוגונלית אס"ם כל שורה בA היא וקטור מאורך יחידה וכל שתי שורות ניצבות זו לזו ($AA^{T}=I\Leftrightarrow$).

הגדרה 18.40 מטריצה אוניטרית

$A\in M\left(\mathbb{C}\right)$ תקרא אוניטרית אס"ם $I=\underbrace{A^{*}}_{transpose+adjunct}A$.

18.3.1 וקטורים עצמיים

נתונה העתקה לינארית $T : V \to V$. היינו רוצים לבחור בסיס לV שכל איבריו הם וקטורים עצמיים $\{(v_i, \lambda_i)\}_{i=1}^n$ כך ש:

$$\forall y \in V \quad \underbrace{\exists!}_{\text{exists and unique}} \quad \alpha_1, ..., \alpha_n : \ y = \sum_{i=1}^{n} \alpha_i v_i$$

ולכן:

$$Ty = T \sum \alpha_i v_i = \sum \alpha_i T v_i = \sum \alpha_i \lambda_i v_i$$

צרות לא לכל העתקה לינארית יש בסיס של ו"ע.

הצלה (חלקית) "כמעט כל ההעתקות הלינאריות הן טובות" - אם תוסיפו למטריצה רעש מקרי קטן ככל שתוסיפו, אז בהסתברות גבוהה מקבלים מטריצה שיש לה בסיס של וקטורים עצמיים.

טענה 18.41 תהיה A מטריצה ממשית, ויהיו $\{u_i\}_{i=1}^m$ ו"ע של A עם ע"ע שונים זה מזה. אז $\{u_i\}_{i=1}^m$ בת"ל.

הוכחה: נסמן בλ_i את הע"ע המתאים לu_i. נניח שk הוא האורך של התלות הלינארית הקצרה ביותר בין ה$\{u_i\}_{i=1}^m$. בה"כ $\sum_{i=1}^k \alpha_i u_i = 0 = *$ אבל אין תלות קצרה יותר (ז"א תלות לינארית שמעורבים בה $\geq k-1$ מהu_i). נשים לב:

$$A \sum_{i=1}^{k} \alpha_i u_i = A0 = 0$$

$$\Downarrow$$

$$A \sum_{i=1}^{k} \alpha_i u_i = \sum_{i=1}^{k} \alpha_i A u_i =^{**} \sum_{i=1}^{k} \alpha_i \lambda_i u_i = 0$$

אבל אם נכפיל את $*$ בλ_k ונחסיר מ$**$ נקבל ת"ל על$\geq k-1$ מהu_i:

$$0 - 0 = \sum_{i=1}^{k} \alpha_i \lambda_i u_i - \lambda_k \sum_{i=1}^{k} \alpha_i u_i = \sum_{i=1}^{k-1} \alpha_i \left(\lambda_i - \lambda_k\right) u_i = 0$$

■

משפט 18.42 המשפט הספקטרלי למטריצות ממשיות סימטריות (ללא הוכחה)
לכל מטריצה ממשית סימטרית A יש מטריצה אורתוגונלית U ומטריצה אלכסונית Λ כך ש:

$$A = U \Lambda U^T$$

איברי Λ (שבאלכסון) הם הע"ע של A. בפרט למטריצה ממשית סימטרי כל הע"ע הם ממשיים.

מסקנה 18.43 השורות של U הן הו"ע של A.

הוכחה: תהי φ העמודה ה־j שב־U. אז:

$$A\varphi = U\Lambda U^T\varphi = U\Lambda e_j = U\underbrace{\lambda_j}_{vector} = \underbrace{\lambda_j}_{scalar}\varphi$$

■

הערה 18.44 נשים לב:

$$\forall x:\ \ Ax = U\Lambda U^T x$$

כש־U^T מבצע סיבוב, Λ מבצע מתיחה אחת לאורך כל ציר, ו־U מבצע סיבוב בחזרה.

19 תרגול 5

19.1 תזכורת

הגדרה 19.1 ע"ע ו־ו"ע

אם V מ"ו מעל $\mathbb{C}$, $T: V \to V$ ה"ל אז $\lambda \in \mathbb{C}$ נקרא **ערך עצמי** של T אם קיים $V \ni v \neq 0$ כך ש־$Tv = \lambda v$.
במקרה זה v נקרא **וקטור עצמי** של λ.

הגדרה 19.2 פולינום אופייני

אם V סוף מימדי ו־A מטריצה שמייצגת את T לפי בסיס קבוע, אז הגדרנו

$$p_A(x) = det(xI - A)$$

זהו הפולינום האופייני של T.

הערה 19.3 אם λ ע"ע של T $\Leftrightarrow$ $p_A(\lambda) = 0$.

הערה 19.4 $p_A(x)$ לא תלוי בבחירת הבסיס, ולכן ניתן לדבר על **ה**פולינום האופייני של T.

$$\begin{aligned} det\left(xI - SAS^{-1}\right) &= det\left(S(xIS - A)S^{-1}\right) = det(S)\,det(xIS - A)\,det\left(S^{-1}\right) = \\ &= \underbrace{det(S)\,det\left(S^{-1}\right)}_{=1} det(xIS - A) = det(xIS - A) \end{aligned}$$

19.2 נוסחאות נסיגה

הגדרה 19.5 $a_{n+k+1}=\sum_{i=0}^{k}\alpha_i a_{n+i}$ עבור $\alpha_0,...,\alpha_k\in\mathbb{C}$.

הגדרה 19.6 פתרון למשוואת הרקורסיה הנ"ל היא סדרה $\{a_n\}_{n=0}^{\infty}\subseteq\mathbb{C}$ כך שכל

$$a_{n+k+1},a_{n+k},...,a_n$$

מקיים את המשוואה.

הערה 19.7 דוגמא: פיבונאצ'י: $a_{n+2}=a_{n+1}+a_n$
הפתרון עבור פיבונאצ'י:

$$a_0=0,a_1=1,a_2=1,a_3=2,....$$

טענה 19.8 לכל $a_0,...,a_k\in\mathbb{C}$ קיימת ויחידה השלמה לסדרה $\{a_n\}_{n=0}^{\infty}$ שהיא פתרון.

הוכחה: לבנות בצורה אינדוקטיבית איבר־איבר, ומראים את היחידות באינדוקציה על ידי שתי סדרות שמתחילות בצורה זהה ומקיימות את התנאי ולהראות שההפרש ביניהן הוא 0. ■

טענה 19.9 נסמן את אוסף הפתרונות בV, אזי V הוא תת מרחב של הסדרות המרוכבות ממימד $k+1$. כלומר: נסמן $W=\{\{c_n\}_{n=0}^{\infty}\subseteq\mathbb{C}\}$ אזי W הוא מ"ו עם הפעולות:

$$\{c_n\}_{n=0}^{\infty}+\{d_n\}_{n=0}^{\infty}=\{c_n+d_n\}_{n=0}^{\infty}$$
$$\alpha\in\mathbb{C},\alpha\{c_n\}_{n=0}^{\infty}=\{\alpha c_n\}_{n=0}^{\infty}$$

הוכחה: $0\in V$ כי הסדרה הקבועה 0 היא תמיד פתרון.
סגירות לחיבור: אם $\{a_n\}_{n=0}^{\infty},\{b_n\}_{n=0}^{\infty}\in V$ אזי:

$$\forall n,\ a_{n+k+1}+b_{n+k+1}=\sum_{i=0}^{k}\alpha_i a_{n+i}+\sum_{i=0}^{k}\alpha_i b_{n+i}=\sum_{i=0}^{k}\alpha_i\left(a_{n+i}+b_{n+i}\right)$$
$$\Downarrow$$
$$\{a_n\}_{n=0}^{\infty}+\{b_n\}_{n=0}^{\infty}\in V$$

סגירות לכפל: יהי $\beta\in\mathbb{C},\{a_n\}\in V$, אזי לכל n:

$$\beta a_{n+k+1}=\beta\sum_{i=0}^{k}\alpha_i a_{n+i}=\sum_{i=0}^{k}\beta a_{n+i}\alpha_i$$
$$\Downarrow$$
$$\beta\{a_n\}\in V$$

קיבלנו שV ת"מ (תת מרחב) של W. ■

הגדרה 19.10 איזומורפיזם
העתקה לינארית חח"ע ועל.

הגדרה 19.11 נגדיר: $T:\ \mathbb{C}^{k+1}\to V$ כך שT איזומורפיזם. נגדיר:

$$T\begin{pmatrix} a_0 \\ ... \\ a_k \end{pmatrix} = \text{the only series that completes } a_0,..,a_k \text{ and is a solution}$$

T מוגדר היטב לפי הטענה.

הוכחה: T ה"ל: צריך להסתכל על נוסחת הנסיגה ולוודא.

T חח"ע: אם $T\begin{pmatrix} a_0 \\ ... \\ a_k \end{pmatrix}=0$ אז בפרט $a_0=...=a_k=0$. (נזכיר שT חח"ע $\Leftrightarrow$ הגרעין של T הוא 0).

T על: אם $\{a_n\}\in V$ אז $\{a_n\}=T\begin{pmatrix} a_0 \\ ... \\ a_k \end{pmatrix}$. ■

מסקנה 19.12 $dimV=k+1$

נשים לב שאם $\{a_n\}\in V$, אז לכל n:

$$\begin{pmatrix} a_{n+k+1} \\ a_{n+k} \\ ... \\ a_{n+1} \end{pmatrix} = \underbrace{\begin{pmatrix} \alpha_k & ... & ... & \alpha_0 \\ 1 & 0 & ... & 0 \\ 0 & 1 & 0 & \\ & & ... & \\ 0 & ... & 0 & 0 \end{pmatrix}}_{A} \begin{pmatrix} a_{n+k} \\ a_{n+k-1} \\ ... \\ a_n \end{pmatrix}$$

נשים לב שאם λ ע"ע של A, אז נניח $v=\begin{pmatrix} v_k \\ \dots \\ \dots \\ v_0 \end{pmatrix}$ ו"ע מתאים:

$$A\begin{pmatrix} v_k \\ \dots \\ \dots \\ v_0 \end{pmatrix}=\lambda\begin{pmatrix} v_k \\ \dots \\ \dots \\ v_0 \end{pmatrix}=\begin{pmatrix} \sim \\ v_k \\ \dots \\ v_1 \end{pmatrix}$$
$$\Downarrow$$
$$v_1=\lambda v_0$$
$$\dots$$
$$v_k=\lambda^k v_0$$
$$\Downarrow$$
$$\begin{pmatrix} \lambda^k \\ \dots \\ \dots \\ 1 \end{pmatrix}\text{ is an eigenvector}$$
$$\Downarrow$$
$$A\begin{pmatrix} \lambda^k \\ \dots \\ \dots \\ 1 \end{pmatrix}=\begin{pmatrix} \lambda^{k+1} \\ \dots \\ \dots \\ \lambda \end{pmatrix}$$
$$\Downarrow$$
$$\lambda^{k+1}=\sum_{i=0}^{k}\alpha_i\lambda^i$$

אם נגדיר סדרה $a_n=\lambda^n$, אזי

$$a_{n+k+1}=\lambda^{n+k+1}=\lambda^n\lambda^{k+1}=\lambda^n\sum_{i=0}^{k}\alpha_i\lambda^i=\sum_{i=0}^{k}\alpha_i\underbrace{\lambda^{n+i}}_{a_{n+i}}$$

20 תרגול אקסטרה 5

20.1 ע"ע ו־ו"ע

20.1.1 הגדרות

A מטריצה $m\times m$ מעל $\mathbb{R}$, סימטרית ($A^T=A$).

הגדרה 20.1 $\lambda\in\mathbb{R}$ הוא ע"ע של A אם קיים וקטור $v\neq 0$ כך ש: $Av=\lambda v$.

הגדרה 20.2 **המרחב העצמי** של λ:

$$V_\lambda = \{v \in \mathbb{R}^m | Av = \lambda v\}$$

הגדרה 20.3 **הריבוי** של λ:

$$r(\lambda) = \dim V_\lambda$$

משפט 20.4 $A \in M_m(\mathbb{R})$ סימטרית, ו$\lambda_1, ..., \lambda_k$ הם ע"ע של A, אז:

$$r(\lambda_1) + ... + r(\lambda_k) = m$$

20.1.2 דוגמאות

1. $A = I_m$ מטריצת היחידה. נשים לב ש$\lambda = 1$ ע"ע וגם: $V_\lambda = \mathbb{R}^m,\ r(1) = m$
2. $A = J = \begin{bmatrix} 1 & ... & 1 \\ ... & ... & ... \\ 1 & ... & 1 \end{bmatrix}$, אם $v \in \mathbb{R}^m$:

$$Jv = \begin{bmatrix} \sum_{i=1}^m v_i \\ ... \\ \sum_{i=1}^m v_i \end{bmatrix} = \underbrace{\sum_{i=1}^m v_i}_{scalar} \begin{bmatrix} 1 \\ ... \\ 1 \end{bmatrix}$$

(א) אם $\sum_{i=1}^m v_i = 0$ אז $Jv = 0$, משמע $\lambda = 0$ הוא ע"ע של J ו:

$$U = V_0 = \left\{ v \in \mathbb{R}^m | \sum_{i=1}^m v_i = 0 \right\}$$
$$r(0) =^1 m - 1$$

הוכחה: כאשר את $=^1$ ניתן להוכיח בכמה אופנים:

i. נמצא $m-1$ וקטורים בV_0, שהם בת"ל, ופורשים כל וקטור בV_0:

$$u_i = \begin{bmatrix} \\ 1 \\ \\ \\ -1 \\ \\ \end{bmatrix} \begin{array}{l} \\ \leftarrow \text{i-th coordinate} \\ \\ \\ \leftarrow \text{m-th coordinate} \\ \\ \end{array}$$

אם $\alpha_1, ..., \alpha_{m-1} \in \mathbb{R}$ נדרוש:

$$\begin{bmatrix} \alpha_1 \\ ... \\ \alpha_{m-1} \\ -\alpha_1 - .... - \alpha_{m-1} \end{bmatrix} = \sum_{i=1}^{m-1} \alpha_i u_i = 0 \Rightarrow \forall i: \ \alpha_i = 0$$

.ii $v \in V_0$ מחפשים $\alpha_1, ..., \alpha_{m-1} \in \mathbb{R}$ כך ש $v = \sum \alpha_i u_i$. נבחר $\alpha_i = v_i$ ואז צ"ל

$$-\sum_{i=1}^{m-1} \alpha_i = v_m \overset{v \in V_0}{=} -\sum_{i=1}^{m-1} v_i$$

■

(ב) נשאר עוד ע"ע אחד λ בריבוי $r(\lambda) = 1$: $\lambda = m$ עם הו"ע $v = \begin{bmatrix} 1 \\ ... \\ 1 \end{bmatrix}$:

$$Jv = mv$$

.3 $A = \begin{bmatrix} a & b & b & b \\ b & a & b & b \\ b & b & ... & b \\ b & b & b & a \end{bmatrix}$, $0 < b < a \in \mathbb{R}$. נשתמש בעובדה הבאה:

טענה 20.5 אם $v \in \mathbb{R}^m$ הוא ו"ע של A_1 עם ע"ע λ_1, וגם ו"ע של A_2 עם ע"ע λ_2. אז לכל $\alpha, \beta \in \mathbb{R}$ נקבל:

$$\Bigg(\underbrace{\alpha A_1 + \beta A_2}_{A}\Bigg) v = (\alpha \lambda_1 + \beta \lambda_2) v$$

משמע, v הוא ו"ע של A עם ע"ע $\alpha \lambda_1 + \beta \lambda_2$.

נשים לב ש:

$$A = bJ + (a-b) I$$

לכן, כל וקטור בU (שהגדרנו קודם) הוא ו"ע של A עם ע"ע $b \cdot m + (a-b) 1$, והוקטור $v = \begin{bmatrix} 1 \\ ... \\ 1 \end{bmatrix}$ הוא ו"ע של A עם ע"ע $b \cdot 0 + (a-b) \cdot 1$, וסה"כ $\lambda_1 = a + b(m-1)$ עם ריבוי של $m-1$ ו$\lambda_2 = a - b$ בריבוי של 1.

מסקנה 20.6 $rank(A) = m$ אם $rank(A) < m$ אז $\dim \ker(A) > 0$ ואז $\lambda = 0$ הוא ע"ע של A, וזה לא נכון. מכיוון שסכום הריבויים של λ_1, λ_2 הוא m, אין לA עוד ערכים עצמיים.

טענה 20.7 הדטרמיננטה של A היא מכפלת הע"ע שלה עם ריבוי.

מהטענה נקבל: $\det(A) = (a-b)(a + b(m-1))^{m-1}$

20.2 נשתמש באלגברה כדי להבין קומבינטוריקה

משפט 20.8 א"ש פישר
אם $S_1,...,S_m \subseteq \{1,...,n\}$ כך ש:

1. $\forall i:\ |S_i| = a$
2. $\forall i \neq j:\ |S_i \cap S_j| = b$

ו$a > b$ אז: $m \leq n$

הוכחה: נגדיר מטריצה $B \in M_{n\times m}(\mathbb{R})$:

$$B_{k,i} = \begin{cases} 1 & k \in S_i \\ 0 & k \notin S_i \end{cases}$$

נסתכל על $A = B^T B$.

טענה 20.9 לכל i,j: $A_{i,j} = |S_i \cap S_j|$

הוכחה: נשים לב:

$$A_{i,j} \overset{\text{matrx multiplication}}{=} \sum_{k=1}^{n} B_{i,k}^T B_{kj} = \sum_{k=1}^{n} \underbrace{B_{k,i}B_{kj}}_{\begin{cases} 1 & k \in S_i \cap S_j \\ 0 & otherwise \end{cases}} = |S_i \cap S_j| = \begin{cases} a & i=j \\ b & otherwise \end{cases}$$

■

טענה 20.10

$$\forall C,D:\ rank(CD) \leq rank(D)$$

משתי הטענות נקבל:

$$A = \begin{bmatrix} a & b & b \\ b & ... & b \\ b & b & a \end{bmatrix}$$

$$m = rank(A) \leq rank(B) \underset{\text{B has only n rows}}{\leq} n$$

■

21 תרגיל 5

טענה 21.1 $p\left(x\right)=x^{k+1}-\sum_{i=0}^{k}\alpha_{i}x^{i}$ הוא הפולינום האופייני של

$$A=\begin{pmatrix}\alpha_{k} & \alpha_{k-1} & ... & \alpha_{0}\\ 1 & 0\\ & ... & 0\\ & & 1 & 0 \end{pmatrix}$$

משפט 21.2 יהיו $\alpha_{0},...,\alpha_{k}\in\mathbb{C}$ כש$\alpha_{0}\neq0$, ונסמן ב

$$a_{n+k+1}=\sum_{i=0}^{k}$$

נוסחאת נסיגה. אז לפי המשפט היסודי של האלגברה קיימים $\lambda_{1},...,\lambda_{l}\in\mathbb{C}$ ו$p_{1},..,p_{l}\in\mathbb{N}$ כך ש

$$p\left(x\right)=\Pi_{i=1}^{l}\left(x-\lambda_{i}\right)^{p_{i}}$$

$$\sum_{i=1}^{l}p_{i}=k+1$$

במקרה זה, לכל $0\leq j<p_{i}$ נקבל ש $a_{n}=n^{j}\lambda_{i}^{n}$ הוא גם פתרון.

מסקנה 21.3 כל פתרון ניתן לכתיבה באופן ייחודי כ:

$$a_{n}=\left(A_{1,1}+A_{1,2}n+...+A_{1,p_{1}}n^{p_{1}-1}\right)\lambda_{1}^{n}+\left(A_{l,1}+A_{l,2}n+...+A_{l,p_{l}}n^{p_{l}-1}\right)\lambda_{l}^{n}$$

טענה 21.4 אם $T:V\to W$ ה"ל חח"ע ועל ו$\{v_{i}\}$ בסיס בV, אז $\{Tv_{i}\}$ בסיס בW.

טענה 21.5 חישוב דטרמיננטה לפי מינורים:

$$\det\left(A\right)=\sum_{i=1}^{n}\left(-1\right)^{i+j}A_{i,j}\det\left(\text{matrix A without row i and column j}\right)$$

טענה 21.6 אם A מטריצה משולשית, אז $det\left(A\right)=\Pi_{i}A_{i,i}$.

טענה 21.7 דטרמיננטה של מטריצת בלוקים
אם $A_{n\times n},B_{n\times m},C_{m\times n},D_{m\times m}$, אז:

$$\det\begin{pmatrix}A & 0\\ C & D \end{pmatrix}=\det\left(A\right)\det\left(D\right)=\det\begin{pmatrix}A & B\\ 0 & D \end{pmatrix}$$

טענה 21.8 $det\left(A\right)\neq0\Leftrightarrow A\text{ is invertible}$

טענה 21.9 סכום סדרה גיאומטרית, אם $r\neq1$:

$$a+ar+\cdots+ar^{n-1}=a\frac{r^{n}-1}{r-1}$$

22 שיעור 6

22.1 תזכורות

הגדרה 22.1 העתקה $T:\ V\to W$
עבור V,W מרחבים לינאריים מעל אותו השדה $\mathbb{F}$, T נקרעת העתקה לינארית אם:

$$\forall x,y\in V:T\left(x+y\right)=Tx+_W Ty$$
$$\forall x\in V,\alpha\in\mathbb{F}:T\left(\alpha x\right)=\alpha T\left(x\right)$$

נקרא גם טרנספורמציה לינארית (ט"ל).

משפט 22.2 כל העתקה לינארית היא הרכבה של סיבובים ומתיחות בכיוון הצירים.

הגדרה 22.3 טרנספורמציה לינארית $T:V\to V$ נקראת **סיבוב** אם היא שומרת על אורכים ועל זוויות.

הערה 22.4 ראינו שמעל $\mathbb{R}$ סיבוב מתואר על ידי מטריצה אורתוגונלית A כך ש$AA^T=I$, ומעל $\mathbb{C}$: $BB^*=B\overline{B}^T=I$ (העתקה אוניטרית).

הערה 22.5 תת מרחב לינארי עובר דרך הראשית, תת מרחב אפיני הוא הזזה של תת מרחב לינארי.

הערה 22.6 כאשר $L\subseteq V$ תת מרחב חד מימדי ורוצים להבין באילו תנאים

$$\underbrace{AL}_{=\{Ax\ \mid\ x\in L\}}=L$$

במקרה הזה A מכפילה כל וקטור שבL בקבוע λ, ואז אומרים שL הוא מרחב עצמי חד מימדי עם ע"ע λ. כלומר, לכל $x\in L$ מתקיים $Ax=\lambda x$.

הערה 22.7 השורשים של $|\lambda I-A|=0$ הם הע"ע העצמיים.

הערה 22.8 מקרה פשוט במיוחד של מתיחה הוא זה שבו הוקטורים העצמיים הם וקטורי היחידה הסטנדרטיים:

$$Ax=\left(\lambda_1x_1,...,\lambda_nx_n\right)$$
$$A=\begin{bmatrix}\lambda_1 & & 0\\ & ... & \\ 0 & & \lambda_n\end{bmatrix}$$

משפט 22.9 המשפט הספקטרלי למטריצות ממשיות סימטריות
כל מטריצה ממשית סימטרית A ניתן להציג בצורה $A=U\Lambda U^T$ כשU מטריצה אורתוגונלית וΛ אלכסונית, כשאיברי האלכסון שלה הם הע"ע של A והשורות של U הן הו"ע המתאימים (השורה הi היא הו"ע של הע"ע הi).

22.2 איך מייחסים אורך\גודל לוקטור?

22.2.1 הגדרות

הגדרה 22.10 אורך של וקטור
האורך של וקטור = **מרחקו** מהראשית.
במימד אחד, על הישר: המרחק בין $x,y\in\mathbb{R}$ מוגדר כ-$|x-y|$.

הגדרה 22.11 מרחב מטרי
מרחב מטרי (X,d) בנוי מקבוצה X ומפונקציה $d:\ X\times X\to\mathbb{R}$ המקיימת את התכונות הבאות:

$$\forall x,y\in X: d(x,y)\geq 0$$
$$d(x,y)=0\Leftrightarrow x=y$$

$$\forall x,y\in X: d(x,y)=d(y,x)$$

$$\forall x,y,z\in X: d(x,y)+d(y,z)\geq d(x,z)\ \text{(triange inequality)}$$

22.2.2 מטריקות

רוצים מטריקה "טובה" על $\mathbb{R}^k$. $\mathbb{R}^k$ הוא כמובן גם מרחב לינארי.

הגדרה 22.12 נורמה $\|x\|\leftrightarrow$ "האורך של x"
זו העתקה $\mathbb{R}^k\to\mathbb{R}_{\geq 0}$ עם התכונות הבאות:

$$\forall x\in\mathbb{R}^k: 0\leq\|x\|$$
$$0=\|x\|\Leftrightarrow x=0$$

$$\forall\lambda\in\mathbb{R},\ x\in\mathbb{R}^k:\ \|\lambda x\|=|\lambda|\,\|x\|$$

$$\forall x,y\in\mathbb{R}^k:\ \|x+y\|\leq\|x\|+\|y\|\ \text{(triange inequality)}$$

כל מה שנאמר כאן עובר ללא שינוי אם מחליפים את $\mathbb{R}^k$ במרחב לינארי $V_{\mathbb{R}}$ (או $V_{\mathbb{C}}$).

22.2.3 דוגמה חשובה

אם V מצוייד במכפלה פנימית: $x,y\to\langle x,y\rangle$ (מעל $\mathbb{R}$: $\sum x_jy_j$, מעל $\mathbb{C}$: $\sum x_j\overline{y_j}$) מקבלים את הנורמה הקרויה נורמת l_2,L_2:

$$\|x\|_2=\sqrt{\langle x,x\rangle}$$

הערה 22.13 אי שוויון המשולש הוא בדיוק קושי שוורץ

$$\begin{aligned}
\sqrt{\langle x+y,x+y\rangle} &\overset{?}{\leq} \sqrt{\langle x,x\rangle}+\sqrt{\langle y,y\rangle} \\
&\Updownarrow \\
\langle x+y,x+y\rangle &\overset{?}{\leq} \langle x,x\rangle+\langle y,y\rangle+2\sqrt{\langle x,x\rangle}\sqrt{\langle y,y\rangle}= \\
&= \langle x,x\rangle+\langle y,y\rangle+2\left\|x\right\|_2\left\|y\right\|_2 \\
&\Updownarrow \\
\langle x,y\rangle &\leq \left\|x\right\|_2\left\|y\right\|_2
\end{aligned}$$

דוגמה

נביט במרחב הפונקציות הרציפות $f:[0,1]\to\mathbb{R}$:

$$\left\|f\right\|_2 := \sqrt{\int\limits_0^1 \left(f\left(x\right)\right)^2 dx}$$

אפשר לחשוב על וקטור ב$\mathbb{R}^n$ בתור פונקציה $\{1,...,n\}\to\mathbb{R}$:

$$\left(\begin{matrix} x_1, & ..., & x_n \\ \underset{1}{\uparrow} & & \underset{n}{\uparrow}\end{matrix}\right)$$

כשרוצים לדעת עד כמה שתי פונקציות רציפות $f,g:[0,1]\to\mathbb{R}$ שונות זו מזו ("מה מרחקן"). דרך סבירה לעשות זאת היא להביט ב:

$$\left\|f-g\right\|_2 = \sqrt{\int_0^1 \left(f\left(x\right)-g\left(x\right)\right)^2 dx}$$

הערה 22.14 כותבים l_2 כשמתכוונים לנורמה על וקטורים וL_2 כשמתכוונים לנורמה על פונקציות רציפות.

22.2.4 משפחה חשובה של נורמות: נורמות l_p (ובצידן L_p)

נגדיר:

$$\begin{aligned}
x &= [x_1,...,x_n]\in\mathbb{R}^n \\
\left\|x\right\|_p &:= \left(\sum\left|x_i\right|^p\right)^{\frac{1}{p}}
\end{aligned}$$

זו נורמה לכל $1\leq p\leq\infty$ כשהחשובות ביותר מתקבלות עבור $p=1,2,\infty$.

נורמת l_1 **:** נקראת גם נורמת מנהטן:

$$l_1 : \|x_1\| = \sum |x_i|$$

שימוש חשוב: נניח שנתונות לנו שתי התפלגויות π, σ. דרך חשובה למדוד את המרחק ביניהן:

$$d_{Total-Variation} = d_{TV} = \frac{1}{2} \|\pi - \sigma\|_1$$

למשל, אם $\Omega = \{1, ..., n\}$ ויש לנו

$$P : \ p_1, ..., p_n \geq 0, \ \sum p_i = 1$$
$$Q : \ q_1, ..., q_n \geq 0, \ \sum q_i = 1$$

אז:

$$d_{TV}(P,Q) = \frac{1}{2} \sum |p_i - q_i| = \max_{A \subseteq \Omega} (\pi(A) - \sigma(A))$$

l_0 **:** לא נורמה (לא מקיים את המכפלה בסקלר). נקרא התומך של וקטור ושווה למס' הקורדינטות בוקטור ששונות מ0.

נורמת l_∞ **:**

$$\|x\|_\infty = \lim_{p \to \infty} \|x\|_p = \lim_{p \to \infty} \left(\sum |x_i|^p \right)^{\frac{1}{p}} = \max_i |x_i|$$

22.2.5 נקודת מבט גיאומטרית

נניח ש$\|\|$ נורמה על $\mathbb{R}^n$ מותאמת לה קבוצה חלקית ל$\mathbb{R}^n$ הקרויה כדור היחידה של הנורמה:

$$B_{\|\|} = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$$

B_1 הוא ריבוע.

B_2 זוהי הספירה האוקלידית.

במימד כללי זהו אוקטהדר cross-polytope - פאון עם $2n$ קודקודים $\pm e_i, \ i = 1, ..., n$ ויש לו 2^n דפנות $\sum |x_i| \leq 1$.

עבור ∞ ב$\mathbb{R}^n$ הקוביה הn מימדית, ויש לה 2^n קודקודים כאשר כל קודקוד מהצורה $(\epsilon_1, .., \epsilon_n)$ כשכל $\epsilon_i = \pm 1$. לקוביה יש $2n$ דפנות: $x_i = \pm 1, \ i = 1, .., n$.

מה ניתן לומר באופן כללי על $B_{\|\|}$ לנורמה כללית?

1. $B_{\|\|}$ הוא גוף קמור.
2. $B_{\|\|}$ סימטרי יחסית לראשית. ז"א $x \in B_{\|\|}$ אז גם $-x \in B_{\|\|}$.
3. $B_{\|\|}$ הוא ממימד מלא (ז"א לא מוכל בתת מרחב ממש).

בהינתן $B \subseteq \mathbb{R}^n$ עם התכונות: מימד מלא, סימטריה יחסית לראשית וקמירות, מוגדרת נורמה כדלקמן:

$$x \in \mathbb{R}^n$$
$$\|x\| = \inf\left\{\lambda \mid \frac{x}{\lambda} \in B\right\}$$

22.2.6 גודל של מטריצה

יהיו שני מרחבים נורמיים:

$$(V, \|\|_V), \ (W, \|\|_W)$$

ותיהיה $T : V \to W$. נרצה להגדיר "גודל" לT ("נורמה אופרטורית").
נשאל: עד כמה יכולה T להאריך אורכים?

הגדרה 22.15 נורמה אופרטורית:

$$\|T\|_{\substack{op \\ (V,\|\|_V)\to(W,\|\|_W)}} = \sup_{x\in V} \frac{\|Tx\|_W}{\|x\|_V}$$

באופן כללי: נשתמש בנורמה הזו ובעוד מדדים ל"גודל" של מטריצות.

הערה 22.16 סוג חשוב ומעניין במיוחד של נורמות אופרטוריות לוקח בחשבון גם את המבנה הכפלי של אופרטורים (הרכבה):

$$U \xrightarrow{S} V \xrightarrow{T} W$$

היינו רוצים נורמות שמקיימות:

$$\|T \circ S\| \leq \|T\| \, \|S\|$$

22.2.7 נורמה דואלית

הגדרה 22.17 הנורמה הדואלית של $\|\|$

הגדרה 22.18 נניח $(\mathbb{R}^n, \|\|)$ מרחב נורמי. נשים לב שכל $x \in \mathbb{R}^n$ מקודד העתקה $\mathbb{R}^n \to \mathbb{R}$: $y \to \langle x, y \rangle$ ("פונקציונל לינארי"). נגדיר:

$$\|x\|^* := \max_y \frac{|\langle x.y \rangle|}{\|y\|} = $$
$$= \max_{\|y\|=1} |\langle x, y \rangle|$$

מתברר שזו נורמה שנקראת הנורמה הדואלית של $\|\|$.

טענה 22.19 $\|\|^*$ הוא נורמה.

דוגמאות

טענה 22.20 $\|\|_1^* = \|\|_\infty$

הוכחה: בהינתן $x \in \mathbb{R}^n$ נרצה לחשב את $\|x\|^*$:

$$x = (x_1, ..., x_n)$$
$$\|x\|^* = \max_{\|y\|_1=1} |\langle x, y \rangle|$$

מחפשים $y_1, ..., y_n$ כך ש$\sum |y_i| = 1$ וכך ש$|\sum x_i y_i|$ גדול ביותר. בה"כ $sgn(x_i) = sgn(y_i)$ ולכן בה"כ $\forall i \ x_i \geq 0, \ y_i \geq 0$, ונקבל:

$$\|x\|^* = \max_{\substack{y_i \geq 0 \\ \sum y_i = 1}} \sum x_i y_i$$

כמובן הבחירה המיטבית היא $y_j = 1$ כשx_j הוא הגדול מבין הxים, ואז:

$$\|x\|^* = \|x\|_\infty$$

■

23 תרגול 6

23.1 מרחבים מטרים

(X, d) כש d מקיימת סימטריה, חיוביות, א"ש המשולש.

הגדרה 23.1 מרחב נורמי
מרחב נורמי זה מ"ו V מעל $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ ונורמה $\|\cdot\| : V \to \mathbb{R}$ שמקיימת חיוביות, הומוגניות וא"ש המשולש.

הגדרה 23.2 הומוגניות

$$\forall \alpha \in \mathbb{F}, \ v \in V : \ \|\alpha v\| = |\alpha| \|v\|$$

23.1.1 דוגמאות

- $\mathbb{R}^n: \ l_p, l_\infty, l$
- $\mathbb{C}^n: \ "...."$
- $\mathbb{C}: \ |\cdot|$
- בהינתן מ"נ $(V, \|\cdot\|)$ אז $d(u,v) = \|u-v\|$ היא מטריקה על V.
- $V = Hom(\mathbb{R}^n, \mathbb{R}^m)$, $\|\cdot\|_{\mathbb{F}}$.

23.1.2 מרחב ההומומורפיזמים

V נקרא מרחב ההומומורפיזמים בין $\mathbb{R}^n$ ל$\mathbb{R}^m$ והוא מרחב הה"ל (העתקות הלינאריות) מ$\mathbb{R}^n$ ל$\mathbb{R}^m$. נגדיר עבור $S, T: \ \mathbb{R}^n \to \mathbb{R}^m$:

$$\begin{aligned}(S+T) : \mathbb{R}^n &\to \mathbb{R}^m \\ v_n &\to S(v) + T(v)\end{aligned}$$

$$\alpha \in \mathbb{R}: \ (\alpha T) : \mathbb{R}^n \to \mathbb{R}^m$$
$$v \to \alpha T(v)$$

אנחנו יודעים

$$\mathbb{R}^{nm} \cong M_{m,n}(\mathbb{R}) \leftrightarrow Hom(\mathbb{R}^n, \mathbb{R}^m)$$

הגדרה 23.3 נורמות אופרטורית

בהניתן נורמה $|\cdot|_1$ על $\mathbb{R}^n$, ובהינתן נורמה $|\cdot|_2$ על $\mathbb{R}^m$, אז לכל $T: \ \mathbb{R}^n \to \mathbb{R}^m$ אז:

$$\|T\|_{op} = \max_{|x|_1=1} |Tx|_2 = \max_{0 \neq x \in \mathbb{R}^n} = \frac{|Tx|_2}{|x|_1}$$

הערה 23.4 עם ציורים: ב$\mathbb{R}^2$:

$$\frac{\|\|_1}{\sqrt{n}} \leq \|\|_2 \leq \|\|_1 \leq \sqrt{2}\,\|\|_2$$

משפט 23.5 כל הנורמות על $\mathbb{R}^n$ שקולות

תהי $|\cdot|$ נורמה על $\mathbb{R}^n$ אזי קיימים $C, D > 0$ כך שלכל $v \in \mathbb{R}^n$:

$$C\|v\|_1 \leq |v| \leq D\|v\|_1$$

המשפט נכון גם לכל נורמה אחרת במקום $\|\|_1$.

הוכחה: תחילה נזכר במשפט בולצאנו־ויירשטראס:

משפט 23.6 בולצאנו־ויירשטראס:
אם $\{x_n\} \subseteq [a,b]$ אז ל־$\{x_n\}$ יש תת סדרה מתכנסת.

משפט 23.7 בולצאנו־ויירשטראס++:
אם $\{v_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}^d$ והסדרה חסומה ב־l_1 (כלומר קיים $M > 0$ כך ש: $\forall n: \ \|v_n\|_1 < M$) אז יש לה ת"ס מתכנסת ב־$l_1$. כלומר: יש $v \in \mathbb{R}^d$ ותת סדרה $\{v_{n_k}\}_{n\in\mathbb{N}} \subseteq \{v_n\}_{n\in\mathbb{N}}$ כך ש

$$\|v_{n_k} - v\|_1 \overset{k\to\infty}{\to} 0$$

נחזור להוכחת משפט השקילות:
יהי $v = [v_1, ..., v_n] \in \mathbb{R}^n$ אזי:

$$|v| = \left|\sum_{i=1}^{n} v_i e_i\right| \overset{\text{triangle inequality}}{\leq} \sum_{i=1}^{n} |v_i e_i| = \sum_{i=1}^{n} \underbrace{|v_i|}_{\text{abs. value}} \underbrace{|e_i|}_{\text{norm}} \leq$$

$$\leq \overbrace{\max_{i\in[n]} |e_i|}^{D} \cdot \underbrace{\sum_{i=1}^{n} |v_i|}_{=\|v\|_1}$$

$$\Downarrow$$

$$|v| \leq D\,\|v\|_1$$

נניח בשלילה שלכל $C > 0$ יש $v \in V$ כך ש: $|v| > C\,\|v\|_1$. נקח $\{v_n\}$ כך ש:

$$\forall n: \ \frac{1}{n}\|v_n\|_1 > |v_n|$$

נגדיר: $w_n = \frac{v_n}{\|v_n\|_1}$, אזי:

$$\forall n: \ \frac{1}{n}\|w_n\|_1 = \frac{1}{n}\frac{\|v_n\|_1}{\|v_n\|_1} = \frac{1}{n} > \frac{|v_n|}{\|v_n\|_1} = \left|\frac{v_n}{\|v_n\|_1}\right| = |w_n|$$

$$\Downarrow$$

$$\frac{1}{n} > |w_n|$$

$$\|w_n\|_1 = 1$$

נפעיל את $BW++$ ונקבל ת"ס מתכנסת

$$w_{n_k} \overset{l_1}{\to} w \in \mathbb{R}^n$$

כלומר:

$$\begin{aligned}
\left\|w_{n_k}-w\right\|_1 &\overset{k\to\infty}{\to} 0\\
&\Downarrow\\
\left|w_{n_k}-w\right| &\le D\left\|w_{n_k}-w\right\|_1 \to 0\\
|w| &\le \underbrace{\left|w-w_{n_k}\right|}_{\to 0}+\underbrace{\left|w_{n_k}\right|}_{\to 0}\\
&\Downarrow\\
|w| &= 0\\
&\Downarrow\\
w &= 0\\
&\Downarrow\\
\left\|w_{n_k}-w\right\|_1 &= \left\|w_{n_k}\right\|_1 = 1 \to 0
\end{aligned}$$

קיבלנו סתירה! ולכן קיים C שמקיים את א"ש.

■

24 תרגול אקסטרה 6

24.1 תזכורות

הגדרה 24.1 גרעין של מטריצה

$$\begin{aligned}
&\ker A \le \mathbb{R}^n\\
&\ker A = \{v\in\mathbb{R}^n:\ Av=0\}
\end{aligned}$$

הגדרה 24.2 תמונה של מטריצה:

$$Im\left(A\right)=\{Av:\ v\in\mathbb{R}^n\}$$

הגדרה 24.3 דרגה של מטריצה

$$rank\left(A\right)=rk\left(A\right)=\dim Im\left(A\right)$$

משפט 24.4 סכום דרגה ומימד הגרעין

$$rank\left(A\right)+\dim\ker A=n$$

הגדרה 24.5 עקבה

$$trace\left(A\right)=\sum_{i=1}^{m}A_{i,i}$$

משפט 24.6 משפט מלינארית 1

$$trace\left(AB\right)=trace\left(BA\right)$$

הגדרה 24.7 אורתונורמליות

$v_1,..,v_n$ יקראו אורתונורמליים אם

$$\begin{aligned}\left\langle v_i,v_i\right\rangle &= 1\\ \forall i\neq j:\ \left\langle v_i,v_j\right\rangle &= 0\end{aligned}$$

24.2 ערכים עצמיים

משפט 24.8 ספקטרלי

$A\in M_n\left(\mathbb{R}\right)$ סימטרית אז:

$$\begin{aligned}A&=V\Lambda V^T\\ \Lambda&=\begin{pmatrix}\lambda_1 & & 0\\ & ... & \\ 0 & & \lambda_n\end{pmatrix}\\ V&=\begin{pmatrix}| & & |\\ v_1 & ... & v_n\\ | & & |\end{pmatrix}\\ V^TV&=I\end{aligned}$$

לנוחות, נניח: $|\lambda_1|\geq...\geq|\lambda_n|$.

הערה 24.9 $V^TV=I$ שקול לכך ש$v_1,..,v_n$ הם א"נ ולכן הם בסיס.

24.2.1 לקרוא פרמטרים של A סימטרית בעזרת הערכים העצמיים

טענה 24.10

$$\dim\ker\left(A\right)=\left|\left\{i:\ \ \lambda_i=0\right\}\right|$$

הוכחה:

$$\ker A = \{v : \ Av = 0 \cdot v\}$$

■

מסקנה 24.11

$$rank\,(A) = |\{i : \ \lambda_i \neq 0\}|$$

טענה 24.12

$$trace\,(A) = \sum_{i=1}^{n} \lambda_i$$

הוכחה:

$$trace\,(A) = trace\left((V\Lambda)\,V^T\right) = trace\left(\underbrace{V^T \cdot V}_{I}\Lambda\right) = trace\,(\Lambda) = \sum_{i=1}^{n} \lambda_i$$

■

הגדרה 24.13 נורמת פרוביניוס

$$\|A\|_F := \sqrt{\sum_{i,j=1}^{n} A_{i,j}^2}$$

טענה 24.14

$$\|A\|_F = \sqrt{\sum_{i=1}^{n} \lambda_i^2}$$

הוכחה: נשים לב:

$$\|A\|_F^2 = trace\left(A \cdot A^T\right)$$

כיוון שA סימטרית:

$$\|A\|_F^2 = trace\left(A^2\right) = trace\left(V\Lambda\underbrace{V^T \cdot V}_{I}\Lambda V^T\right) = trace\left(V\Lambda^2 V^T\right) = \sum \lambda_i^2$$

■

הגדרה 24.15 נורמה אופרטורית

$$\|A\|_{2\to 2} = \|A\|_{op} = \max_{0\neq v\in\mathbb{R}^n} \frac{\|Av\|_2}{\|v\|_2}$$

טענה 24.16 אם A סימטרית

$$\|A\|_{2\to 2} = \max|\lambda_i| = |\lambda_1|$$

הוכחה: נוכיח שני דברים:

1. יש $v\in\mathbb{R}^n$ כך ש$\frac{\|Av\|}{\|v\|} = |\lambda_1|$ וזהו הוקטור העצמי של λ_1.
2. לכל $v\in\mathbb{R}^n$ מתקיים $\frac{\|Av\|}{\|v\|} \leq |\lambda_1|$.

יהי $v\in\mathbb{R}^n$, ניתן להציג אותו כצ"ל של $v_1,...,v_n$:

$$v = \sum_{i=1}^{n}\alpha_i v_i$$

נחשב את $\|Av\|, \|v\|$:

$$\|v\|^2 = \langle v,v\rangle = \left\langle \sum_{i=1}^{n}\alpha_i v_i, \sum_{j=1}^{n}\alpha_j v_j \right\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j \underbrace{\langle v_i,v_j\rangle}_{i\neq j\Rightarrow\langle v_i,v_j\rangle=0} = \sum_{i=1}^{n}\alpha_i^2$$

מהו Av?

$$Av = \sum_{i=1}^{n}\alpha_i Av_i = \sum_{i=1}^{n}\alpha_i\lambda_i v_i$$
$$\Downarrow$$
$$\|Av\|^2 = \sum_{i=1}^{n}(\alpha_i\lambda_i)^2 \underbrace{\leq}_{|\lambda_1|\geq|\lambda_2|\geq...} \sum_{i=1}^{n}\alpha_i^2\lambda_1^2 = \lambda_1^2\|v\|^2$$

■

משפט 24.17 אם A היא מטריצה סימטרית אז

$$rank\left(A\right) \geq \frac{trace\left(A\right)^2}{\|A\|_F^2} = \frac{trace\left(A\right)^2}{trace\left(A^2\right)}$$

הוכחה: נוכיח תחילה טענה **מאוד שימושית**:

טענה 24.18 אם $a_1, .., a_r \in \mathbb{R}$ אז:

$$\left(\sum_{i=1}^{r} a_i\right)^2 \leq r\sum_{i=1}^{r} a_i^2$$

הוכחה: נגדיר:

$$a = [a_1, .., a_r] \in \mathbb{R}^r$$
$$v = [1, ..., 1] \in \mathbb{R}^r$$

לפי קושי שוורץ:

$$|\langle v, a\rangle| \leq \|v\| \, \|a\|$$
$$\Updownarrow$$
$$\left|\sum_{i=1}^{r} 1 \cdot a_i\right| = \left|\sum_{i=1}^{r} a_i\right| \leq \sqrt{\sum_{i=1}^{r} 1^2}\sqrt{\sum_{i=1}^{r} a_i^2} = \sqrt{r}\sqrt{\sum_{i=1}^{r} a_i^2}$$
$$\Updownarrow$$
$$\left(\sum_{i=1}^{r} a_i\right)^2 \leq r\sum_{i=1}^{r} a_i^2$$

■

נחזור להוכחת המשפט:

$$trace\left(A\right)^2 = \left(\sum_{i=1}^{n} \lambda_i\right)^2$$

נסמן $r = rank\left(A\right)$, אז לכל $i > r$ נקבל $\lambda_i = 0$, ולכן:

$$trace\left(A\right)^2 = \left(\sum_{i=1}^{r} \lambda_i\right)^2 \overset{\text{claim}}{\leq} r\sum_{i=1}^{r} \lambda_i^2 = rank\left(A\right)\|A\|_F^2$$

■

25 תרגיל 6

טענה 25.1 אם $D=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}$ אז:

$$\left\|D\right\|_{op}=\max_{i=1}^{n}\left|\lambda_i\right|$$

טענה 25.2 אם $U\in M_n\left(\mathbb{R}\right)$ אורתוגונלית (כלומר $UU^T=I$) אז:

$$\forall A\in M_n\left(\mathbb{R}\right):\ \left\|UA\right\|_{op}=\left\|A\right\|_{op}$$

טענה 25.3 אם $\left|\right|_1,\left|\right|_2$ הן נורמות על $\mathbb{R}^n$ אז קיימים $C,D>0$ כך שלכל $v\in\mathbb{R}^n$:

$$C\left|v\right|_1\leq\left|v\right|_2\leq D\left|v\right|_1$$

הגדרה 25.4 יהי V מ"ו וd מטריקה על V. תהי $\left\|\cdot\right\|$ נורמה על V. אומרים שd נובעת מ$\left\|\cdot\right\|$ אם:

$$\forall u,v\in V:\ d\left(u,v\right)=\left\|u-v\right\|$$

טענה 25.5 קיימת מטריקה d על $\mathbb{R}^n$ שלא נובעת מנורמה.

הגדרה 25.6 נניח V מ"ו, אז $\left\{v_i\right\}_{i=1}^{n}\subseteq V$ נקראים אורתונורמליים אם:

$$\forall i,j:\ \left\langle v_i,v_j\right\rangle=\begin{cases}1 & i=j\\ 0 & otherwise\end{cases}$$

הגדרה 25.7 נניח V מ"ו, אז עבור $S\subseteq V$:

$$S^{\perp}=\left\{v\in V:\ \forall s\in S,\ \left\langle v,s\right\rangle=0\right\}$$

טענה 25.8 תהליך גרהאם־שמידט:
יהי V מ"ו, יהיו $\left\{v_i\right\}_{i=1}^{k}\subseteq V$ ב"ת. אז קיימים $\left\{w_i\right\}_{i=1}^{k}\subseteq V$ אורתונורמליים כך ש:

$$\forall 1\leq i\leq k:\ span\left\{w_1,...,w_i\right\}=span\left\{v_1,...,v_i\right\}$$

טענה 25.9 אם $S \subseteq V$ אז $S^{\perp}$ הוא תמ"ו לינארי של V.

טענה 25.10 אם $W \subseteq V$ הוא תמ"ו אז לכל $v \in V$ קיים ייצוג ייחודי

$$v = w + u,\ w \in W, u \in W^{\perp}$$

טענה 25.11 $S \subseteq V$, אז $(span\,(S))^{\perp} = S^{\perp}$.

טענה 25.12 $S \subseteq V$, אז $span\,(S) = \left(S^{\perp}\right)^{\perp}$.

26 שיעור 7

26.1 תזכורת

הגדרה 26.1 נורמה n־מימדית על V מרחב לינארי מעל $\mathbb{R}$
זו העתקה $\|\| : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ עם התכונות:

$$\forall x \in \mathbb{R}^n :\ \|x\| \geq 0$$
$$x = 0 \Leftrightarrow \|x\| = 0$$

$$\forall x \in \mathbb{R}^n,\ \forall \lambda \in \mathbb{R} :\ \|\lambda x\| \overset{\text{linearity}}{=} |\lambda| \, \|x\|$$

$$\forall x, y \in \mathbb{R}^n :\ \|x + y\| \overset{\text{triangle inequality}}{\leq} \|x\| + \|y\|$$

הגדרה 26.2 ליפשיץ
אומרים ש$f : (V, \|\|) \to \mathbb{R}$ היא k־ליפשיץ אם

$$\forall x, y \in V :\ |f(x) - f(y)| \leq k \, \|x - y\|$$

הערה 26.3 דוגמאות חשובות
נורמות ℓ, L_p:

$$\|x\|_p := \left(\sum |x_i|^p\right)^{\frac{1}{p}}$$

המקרים החשובים במיוחד:

1. ℓ_2 ־ נורמה אוקלידית.
2. ℓ_∞ ־ נורמת המקסימום: $\|x\|_\infty = \max_i |x_i|$

הגדרה 26.4 כדור היחידה
בהינתן נורמה $\|\|$ על מרחב ליניארי V מוגדר כדור היחידה:

$$B = \{x \in V|\ \|x\| \leq 1\}$$

הערה 26.5 לא קשה לראות: $B = B_{\|\|}$ זו קבוצה קמורה סימטרית יחסית ל0 ("סימטרית מרכזית") ובעלת מימד מלא. ראינו גם איך קבוצה כנ"ל מגדירה באופן יחיד נורמה.

דוגמאות

1. הכדור של ℓ_2 הוא $\left\{x \in \mathbb{R}^n | \sum x_i^2 \leq 1\right\}$.
2. כדור היחידה של ℓ_∞ זו הקובייה: $\{x \in \mathbb{R}^n | \forall i:\ -1 \leq x_i \leq 1\}$.
3. כדור היחידה של ℓ_1 הוא האוקטהדר או ה$cross-polytope$: $\{x \in \mathbb{R}^n | \sum |x_i| \leq 1\}$

הגדרה 26.6 נורמה אופרטורית

בהינתן נורמות, נוח להגדיר גם "גודל" העתקות. למשל אם

$$(V, \|\|_V)\,,\ (W, \|\|_W)$$

ונניח $T: V \to V$ העתקה לינארית, אז הנורמה האופרטורית של T:

$$\|T\|_{op} = \|T\|_{V \to W} := \sup \frac{\|Tx\|_W}{\|x\|_V} = \sup_{x \in V, \|x\|_V = 1} \|Tx\|_W$$

הגדרה 26.7 נורמה דואלית

בהינתן $(V, \|\|)$ מוגדרת הנורמה הדואלית: $\|\|^*$, כדלקמן: נייחס לוקטור $x \in V$ את ההעתקה:

$$\begin{aligned} V &\to \mathbb{R} \\ v &\to \langle x, v \rangle \end{aligned}$$

וזה מוליך להגדרה:

$$\|x\|^* = \sup_{v \in V} \frac{|\langle x, v \rangle|}{\|v\|} = \sup_{\|v\|=1} |\langle x, v \rangle|$$

הערה 26.8 הנורמה הדואלית לנורמת ℓ_p היא נורמת ℓ_q כש

$$\frac{1}{p} + \frac{1}{q} = 1$$

טענה 26.9 הנורמה הדואלית לℓ_2 היא ℓ_2

הוכחה: נסמן $\|\|_2 =: \|\|$. נשים לב:

$$\begin{aligned}\|\|^* &= \sup_{\|v\|_2=1} |\langle x,v\rangle| =^1 \\ &\text{the optimal choice is } v=\frac{x}{\|x\|_2} \\ &=^1 \frac{\langle x,x\rangle}{\|x\|_2}=\frac{\|x\|_2^2}{\|x\|_2}=\|x\|_2\end{aligned}$$

■

טענה 26.10 $\ell_\infty^* = \ell_1$

הוכחה: נשים לב:

$$\|x\|_\infty^* = \max_{\|v\|_\infty=1} |\langle v,x\rangle|$$

בה"כ $v_i x_i \geq 0\ \ \forall i:$, אחרת נחליף ל$-v_i$. המקסימום מושג כאשר

$$\begin{gathered}\forall i:\ \ v_i = sgn\left(x_i\right) \\ \Downarrow \\ \sum x_i \cdot sgn\left(x_i\right) = \sum |x_i| = \|x\|_1\end{gathered}$$

■

טענה 26.11 נורמת $\|\|^*$ היא נורמה.

הוכחה: אי שליליות: היא $0 \leq$ (כי היא סופרמום על אי שליליים), וגם $\|x\|^*=0 \Leftrightarrow x=0$.
לינאריות: מהלינאריות של המכפלה הפנימית:

$$\|\lambda x\|^* = \sup_{\|v\|=1} |\lambda\left\langle x,v\right\rangle| = \lambda \sup_{\|v\|=1} |\langle x,v\rangle|$$

אי שוויון המשולש:

$$\|x+y\|^* \overset{\text{we want}}{\leq} \|x\|^* + \|y\|^*$$

נביט בוקטור v בנורמת יחידה המשיג את $\|x+y\|^* = \langle v,x+y\rangle$. נשים לב:

$$\begin{aligned}|\langle v,x\rangle| &\leq \|x\|^* \\ |\langle v,y\rangle| &\leq \|y\|^*\end{aligned}$$

ומקבלים את הנדרש. ■

משפט 26.12 קושי־שוורץ מוכלל
בהינתן מרחב נורמי $(V, \|\|)$:

$$\forall x, y \in V: \ |\langle x, y\rangle| \leq \|x\| \cdot \|y\|^*$$

הוכחה: יהיו $x, y \in V$:

$$\begin{gathered}\|y\|^* = \sup_v \frac{|\langle v, y\rangle|}{\|v\|} \geq \frac{|\langle x, y\rangle|}{\|x\|} \\ \Downarrow \\ |\langle x, y\rangle| \leq \|x\| \cdot \|y\|^*\end{gathered}$$

■

הגדרה 26.13 כדור היחידה הקוטבי
אם $(V, \|\|)$ מרחב נורמי וB הוא כדור היחידה שלו, אז כדור היחידה של הנורמה הדואלית ("הקוטבי" - polar).

$$B^* = \{x \in V | \forall v \in B: \ \langle x, v\rangle \leq 1\}$$

הגדרה 26.14 קונוס
$C \subseteq \mathbb{R}^n$ הוא קונוס אם הוא סגור לחיבור ולכפל בקבוע אי שלילי:

$$\begin{gathered}\forall x, y \in C: x + y \in C \\ \forall x \in C, \lambda \geq 0: \lambda x \in C\end{gathered}$$

דוגמאות

1. כל $\mathbb{R}^n$
2. האורתנטה החיובית:

$$\forall x \in \mathbb{R}^n: \ \forall i: \ x_i \geq 0$$

3. בוחרים איזושהי רשימה של וקטורים $a_1, ..., a_k \in \mathbb{R}^n$:

$$C = \{x | \ \forall i: \ \langle x, a_i\rangle \geq 0\}$$

הגדרה 26.15 מטריצה מוגדרת חיובית\אי שלילית
מטריצה ממשית סימטרית נקראת מוגדרת חיובית (positive definite=PD)\מוגדרת אי שלילית (positive semidefinite=PSD) אם כל הע"ע שלה חיוביים\אי שליליים.

הערה 26.16 את אוסף כל המטריצות הPSD בגודל $n \times n$ נסמן בPSD_n.

משפט 26.17 התנאים הבאים על מטריצה A סימטרית ממשית שקולים זל"ז (זה לזה):

1. $A \in PSD_n$
2. אפשר להציג את $A = MM^T$
3. התבנית הריבועית אי שלילית: $\forall x : \;\; xAx^T \geq 0$

הוכחה: גרירות:

- $2 \Rightarrow 3$:

$$xAx^T = xMM^Tx^T = \langle xM, xM \rangle = \|xM\|_2^2 \geq 0$$

- $3 \Rightarrow 1$:
נניח שx ו"ע עם ע"ע λ ונרצה לטעון ש$\lambda \geq 0$:

$$Ax^T = \lambda x^T$$

נכפיל משמאל בx:

$$\underbrace{xAx^T}_{\geq 0} = \lambda \underbrace{\|x\|^2}_{\geq 0}$$

- $1 \Rightarrow 2$:
בהסתמך על המשפט הספקטרלי, ניתן להציג $A = V\Lambda V^T$ כשV אורתוגונלי.

$$\Lambda^{\frac{1}{2}} = \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{bmatrix}$$

$$M := V\Lambda^{\frac{1}{2}}$$

$$\Downarrow$$

$$MMT = V\Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}}V^T = A$$

■

משפט 26.18 התנאים הבאים על מטריצה A סימטרית ממשית שקולים זל"ז (זה לזה):

1. $A \in PD$
2. אפשר להציג את A כך $A = MM^T$, כשM לא סינגולרית.
3. התבנית הריבועית חיובית: $\forall x: \ xAx^T > 0$

הערה 26.19 PSD_n הוא קונוס.

הוכחה: נשים לב:
כפל בסקלר חיובי $\mu > 0$ מכפיל את הע"ע בμ.
חיבור: מתנאי 3 במשפט על הPSD.
■

משפט 26.20 הSVD (Singular Value Decomposition)
כל מטריצה ממשית $A_{m\times n}$ ניתן להציג בצורה:

$$A = U_{m\times m} D_{m\times n} V_{n\times n}^T$$
$$U, V \text{ are orthogonal}$$
$$D = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_n \\ 0 & 0 & 0 \end{pmatrix}, \ \sigma_1 \geq ... \geq \sigma_n \geq 0$$

המספרים $\sigma_1, ..., \sigma_n$ נקראים הערכים הסינגולריים של A והם נקבעים ביחידות.

הערה 26.21

$$\dim A = arg \max_m (\text{"}\sigma_m > 0\text{"})$$

הוכחה: היחידות של הσ_i די קלה:

הערה 26.22 תהיה M מטריצה כלשהי, אז MM^T זו מטריצה שבמקום ה(i,j) שלה מופיעה המכפלה הפנימית של השורה הi והשורה הj שבM.
$M^T M$ כנ"ל עם עמודות.

נביט בAA^T - זו מטריצה PSD.

$$AA^T = UD\underbrace{V^T V}_{=I} D^T U^T = UDD^T U^T$$

נשים לב:

$$\underbrace{U^T A \underbrace{A^T U}_{=(U^T A)^T}}_{PSD} = D_{m\times n} D^T_{n\times m} = \begin{pmatrix} \sigma_1^2 & & & \\ & \ddots & & 0 \\ & & \sigma_n^2 & \\ & 0 & & \end{pmatrix}_{m\times m}$$

טענה 26.23 למטריצה ממשית סימטרית M ולUMU^T כשU אורתוגונלית יש אותם הע"ע. **הוכחה:** אם $Mx = \lambda x$, אז:

$$y := Ux$$
$$UMU^T Ux = UMx = \lambda Ux = \lambda y$$

■

$\sigma_1^2, ..., \sigma_n^2$ הם הע"ע של המטריצה הPSD AA^T.

■

הערה 26.24 מכפלה בבלוקים:

$$\begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} = \begin{pmatrix} P_{11}Q_{11} + P_{12}Q_{21} & ... \\ ... & ... \end{pmatrix}$$

הוכחה: באינדוקציה על הגודל של A. ספציפית, נניח שהמשפט ידוע למטריצות $(n-1) \times (m-1)$.
תוכנית הפעולה היא שנרצה למצוא וקטור u_1 ומטריצה U_2 כך ש $\begin{pmatrix} u_1 & U_2 \end{pmatrix}_{m\times m}$ אורתוגונלית, וכן וקטור v_1 ומטריצה V_2 כך ש $\begin{pmatrix} v_1 & V_2 \end{pmatrix}_{n\times n}$ אורתוגונלית:

$$\begin{pmatrix} u_1 & U_2 \end{pmatrix} A \begin{pmatrix} v_1 & V_2 \end{pmatrix}^T \overset{\text{we want}}{=} \begin{pmatrix} \sigma_1 & 0 \\ 0 & B \end{pmatrix}$$

הערה 26.25 מתברר שσ_1 אינו אלא $\|A\|_{op}$.

נגדיר $\sigma_1 := \|A\|_{op}$. ז"א יש וקטור יחידה v_1 כך ש

$$\sigma_1 u_1 = Av_1$$
$$\|v_1\| = 1$$
$$\|u_1\| = 1$$

את וקטורי היחידה u_1, v_1 ניתן להשלים לבסיס אורתונורמלי, ואלו תהיינה המטריצות U_2, V_2 בהתאמה. נסתכל על:

$$\begin{pmatrix} u_1^T \\ U_2^T \end{pmatrix} A \begin{pmatrix} v_1 & V_2 \end{pmatrix} = \begin{pmatrix} u_1^T A v_1 & \overbrace{u_1^T A V_2}^{:=w} \\ U_2^T A v_1 & \underbrace{U_2^T A V_2}_{:=B} \end{pmatrix} = \begin{pmatrix} \sigma_1 \overbrace{\|u_1\|^2}^{=1} & w \\ \underbrace{\sigma_1 U_2^T u_1}_{\text{because of orthogonality}=0} & B \end{pmatrix} = \begin{pmatrix} \sigma_1 & w \\ 0 & B \end{pmatrix}$$

נטען ש$w = 0$ ונראה שאם זה אינו המצב, אז יש וקטור שתמונתו נמתחת בשיעור $> \sigma_1$.

הערה 26.26 למטריצה $\begin{pmatrix} u_1^T \\ U_2^T \end{pmatrix} A \begin{pmatrix} v_1 & V_2 \end{pmatrix}$ יש אותה נורמה אופ' ($= \sigma_1$) כמו לA.

הבה נפעיל את הטרנספורמציה הלינארית הזו על הוקטור $\begin{pmatrix} \sigma_1 & w \end{pmatrix}$, ונקבל:

$$\begin{pmatrix} u_1^T \\ U_2^T \end{pmatrix} A \begin{pmatrix} v_1 & V_2 \end{pmatrix} \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} = \begin{pmatrix} \sigma_1^2 + \|w\|^2 \\ Bw \\ \vdots \end{pmatrix}$$

מהי המתיחה המתקבלת?

לקחנו וקטור שאורכו (נורמת ℓ_2 שלו): $\sqrt{\sigma_1^2 + \|w\|^2}$, וקיבלנו בתמונה וקטור שאורכו $\leq \sigma_1^2 + \|w\|^2$. המתיחה היא לפחות

$$0 \leq \sqrt{\sigma_1^2 + \|w\|^2} = \frac{\sigma_1^2 + \|w\|^2}{\sqrt{\sigma_1^2 + \|w\|^2}}$$

שוויון אס"ם $w = 0$, ומכיוון שσ_1 הוגדר כ$\|A\|_{op}$ יוצא ש$w = 0$.
קיבלנו שאם u_1, v_1 הוגדרו כנ"ל, וU_2, V_2 השלמות שלהם לבסיסים אורתונורמליים אז:

$$\begin{pmatrix} u_1^T \\ U_2^T \end{pmatrix} A \begin{pmatrix} v_1 & V_2 \end{pmatrix} = \begin{pmatrix} \sigma_1 & 0 \\ 0 & B \end{pmatrix}$$

ע"פ הנחת האינדוקציה יש הצגה

$$B = U\tilde{_{m-1}} D\tilde{_{diag}} \tilde{V}_{n-1}^T$$

ומה שראינו מקודם הוא

$$A = \begin{pmatrix} u_1 & U_2 \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} v_1^T \\ V_2^T \end{pmatrix}$$

ולכן:

$$A = \begin{pmatrix} u_1 & U_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \tilde{U} \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \tilde{D} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \tilde{V}^T \end{pmatrix} \begin{pmatrix} v_1^T \\ V_2^T \end{pmatrix}$$

■

27 תרגול 7

27.1 ערכים סינגולריים וSVD

משפט 27.1 משפט הSVD
אם $A \in M_{m,n}(\mathbb{R})$, אז קיימות מטריצות

$$U \in M_m(\mathbb{R}), V \in M_n(\mathbb{R}), D \in M_{m,n}(\mathbb{R})$$

כך ש $A = UDV^T$, ומתקיים:

1. U, V אורתוגונליות (אוניטריות)
2. D "אלכסונית" ($\forall i \neq j: \ D_{i,j} = 0$)
3. וגם $\sigma_1 := D_{11} \geq ... \geq \sigma_{\min(n,m)} := D_{\min(n,m),\min(n,m)} \geq 0$

תרגיל

מצאו SVD למטריצה:

$$A = \begin{pmatrix} 1 & & & \\ & 2 & & \\ & & 3 & \\ & & & 4 \end{pmatrix} \in M_4(\mathbb{R})$$

פתרון

$$A = (?) \begin{pmatrix} 4 & & & \\ & 3 & & \\ & & 2 & \\ & & & 1 \end{pmatrix} (?)$$

נשים לב שלערכים העצמיים $1, 2, 3, 4$ מתאימים הו"ע: $e_1, ..., e_4$ נגדיר: $v_1 = e_4, ..., v_4 = e_1$. ברור ש$\{v_i\}$ הוא בסיס אורתונורמלי של ו"ע. בבסיס זה:

$$e_1 = \begin{pmatrix} \\ \\ \\ 1 \end{pmatrix}, ..., e_4 = \begin{pmatrix} 1 \\ \\ \\ \end{pmatrix}$$

לכן:

$$A = \begin{pmatrix} & & & 1 \\ & & 1 & \\ & 1 & & \\ 1 & & & \end{pmatrix} \begin{pmatrix} 4 & & & \\ & 3 & & \\ & & 2 & \\ & & & 1 \end{pmatrix} \begin{pmatrix} & & & 1 \\ & & 1 & \\ & 1 & & \\ 1 & & & \end{pmatrix}$$

תרגיל

מצאו SVD ל

$$A=\begin{pmatrix}0 & 2\\ 2 & 0\end{pmatrix}$$

פתרון

A סימטרית. נוציא ע"ע:

$$\det\left(xI-A\right)=\det\begin{pmatrix}x & -2\\ -2 & x\end{pmatrix}=x^{2}-4=\left(x-2\right)\left(x+2\right)$$

ולכן הע"ע הם ±2. הו"ע המתאימים:

$$A\begin{pmatrix}1\\ 1\end{pmatrix}=\begin{pmatrix}2\\ 2\end{pmatrix}=2\cdot\begin{pmatrix}1\\ 1\end{pmatrix}$$

נקח $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ 1\end{pmatrix}$ (הו"ע שמצאנו, מנורמל) עבור 2.

נשים לב ש$\begin{pmatrix}1\\ -1\end{pmatrix}$ מאונך לו"ע הקודם. אם כך, נקח אותו מנורמל עבור -2: $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -1\end{pmatrix}$, ולכן:

$$A=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}\begin{pmatrix}2 & 0\\ 0 & -2\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}$$

זה לא פירוק SVD (יש ע"ע שלילי). נשים לב:

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}=\underbrace{\overbrace{\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}}^{S}}_{45\,\text{deg rotation}}\underbrace{\overbrace{\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}}^{R}}_{reflection}$$

$$\Downarrow$$

$$A=SR\begin{pmatrix}2 & 0\\ 0 & -2\end{pmatrix}SR=SR\begin{pmatrix}2 & 0\\ 0 & 2\end{pmatrix}RSR$$

סדרה של ציורים לתאר את הפעולה של A.
נשים לב שRSR היא אורתוגונלית (כמכפלה של אורתוגונליות), SR גם אורתוגונלית. נסמן:

$$U=SR,\ V^{T}=RSR$$

וקיבלנו פירוק SVD.

טענה 27.2 הערכים הסינגולריים הם השורשים של הע"ע $AA^{T}\in M_{m}\left(\mathbb{R}\right)$.

הוכחה: AA^T סימטרית:

$$\left(AA^T\right)^T = AA^T$$

ולכן יש לה לכסון א"ג. נקח SVD:

$$A = UDV^T$$

ונקבל:

$$AA^T = UDV^TVD^TU^T = U \underbrace{DD^T}_{\begin{pmatrix}\sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2\end{pmatrix}} U^T$$

■

תרגיל

מצאו SVD:

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \sqrt{2} & -\sqrt{2} \\ & \end{pmatrix}$$

פתרון

נרצה למצוא את σ_1. נזכר שבהוכחה של SVD:

$$\sigma_1 = \|A\|_{op} = \max_{x\in\mathbb{R}^2, \|x\|^2=1} \|Ax\|_2$$

זה שקול למקסום הפונקציה:

$$f(x,y) = \left\|A\begin{pmatrix}x\\y\end{pmatrix}\right\|^2 = \left(\frac{x}{\sqrt{2}} + \frac{y}{\sqrt{2}}\right)^2 + \left(\sqrt{2}x - \sqrt{2}y\right)^2$$

תחת האילוץ $x^2+y^2=1$. אז:

$$f(x,y) = \frac{x^2+2xy+y^2}{2} + 2x^2 - 4xy + 2y^2 = \frac{1}{2} + xy + 2 - 4xy = \frac{5}{2} - 3xy$$

נכתוב:

$$xy = \begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix}$$

ואז נמקסם:

$$g\left(\theta\right)=\frac{5}{2}-3\cos\theta\sin\theta$$

בתחום $\theta\in\left[0,2\pi\right]$. נגזור:

$$\begin{aligned}
g'\left(\theta\right) & =3\sin^{2}\theta-3\cos^{2}\theta\\
g'\left(\theta\right) & =0\\
 & \Updownarrow\\
\sin^{2}\theta & =3\cos^{2}\theta\\
 & \Updownarrow\\
\theta & \in\left\{ \frac{\pi}{4},\frac{3\pi}{4},\frac{5\pi}{4},\frac{7\pi}{4}\right\} \\
 & \Downarrow\\
\left\Vert A\begin{pmatrix}\frac{1}{\sqrt{2}}\\
-\frac{1}{\sqrt{2}}
\end{pmatrix}\right\Vert _{2} & =2\\
 & \Downarrow\text{we know}\\
A & =\begin{pmatrix}0\\
1\\
0
\end{pmatrix}_{3}\begin{pmatrix}2 & 0\\
0 & ?\\
0 & 0
\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{pmatrix}_{2}
\end{aligned}$$

אפשר לקחת וקטור שמאונך ל$\frac{1}{\sqrt{2}}\quad-\frac{1}{\sqrt{2}}$ על מנת להשלים את המטריצה הימנית.

28 תרגול אקסטרה 7

28.1 פירוק ספקטרלי $\Leftarrow$ SVD

משפט 28.1 משפט הSVD

$A\in M_{m\times n}\left(\mathbb{R}\right)$, $m\geq n$:

$$A_{m\times n}=U_{m\times m}\Sigma_{m\times n}V_{n\times n}^{T}$$

כשU,V אורתוגונליות, ו:

$$\begin{aligned}
\Sigma & =\begin{pmatrix}\sigma_{1}\\
 & \ddots\\
 & & \sigma_{n}\\
\\
\end{pmatrix}\\
\sigma_{1} & \geq...\geq\sigma_{n}\geq0
\end{aligned}$$

28.1.1 הספקטרום של A^TA וAA^T

נשים לב ש$A^TA \in M_n$, והיא סימטרית, לכן יש הצגה:

$$A^TA \overset{\text{spectral decomposition}}{=} V\Lambda V^T$$

ובאופן דומה:

$$AA^T = U\tilde{\Lambda}U^T$$

מטרה: מה הקשר בין Λ ל$\tilde{\Lambda}$ ובין U לV?
נסמן:

$$\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},\ V = \begin{pmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{pmatrix}$$

טענה 28.2 $\lambda_i = \|Av_i\|^2$

הוכחה: $x, y \in \mathbb{R}^n$:

$$x^Ty = \sum_{i=1}^{n} x_iy_i = \langle x, y\rangle$$
$$\Downarrow$$
$$\|Av_i\|^2 = \langle Av_i, Av_i\rangle = (Av_i)^T Av_i = v_i^TA^TAv_i = \lambda_iv_i^Tv_i = \lambda_i$$

■

למה 28.3 $\lambda_i \geq 0$

טענה 28.4 $\lambda > 0$ ע"ע של A^TA אם"ם $\lambda > 0$ ע"ע של AA^T.

הוכחה: נוכיח את כיוון $\Leftarrow$:
נניח $A^TAv = \lambda v$, נכפיל משמאל בA:

$$AA^TAv = \lambda Av$$

אם $Av \neq 0$ אז הוא ו"ע של AA^T עם ע"ע λ. נשים לב ש$Av \neq 0$ כי $\lambda > 0$.
את כיוון $\Rightarrow$ נוכיח באופן דומה. ■

נסמן:

$$\lambda_1 \geq \dots \geq \lambda_r > 0 = \lambda_{r+1} = \dots = \lambda_n$$

לכן:

$$\Lambda_n = \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

$$\tilde{\Lambda}_m = \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

טענה 28.5 נסמן לכל $1 \le i \le r$: $\sigma_i = \sqrt{\lambda_i}$, וגם $u_i := \frac{Av_i}{\sigma_i} \in \mathbb{R}^m$, אז:

1. $u_1, .., u_r$ הם א"נ.
2. u_i הוא ו"ע של AA^T עם ע"ע λ_i.

הוכחה: :

1.

$$\|u_i\| = \frac{\|Av_i\|}{\sigma_i} = 1$$

$$\langle u_i, u_j \rangle = u_i^T u_j = \left(\frac{Av_i}{\sigma_i}\right)^T \left(\frac{Av_j}{\sigma_j}\right) = \frac{1}{\sigma_i \sigma_j} v_i^T A^T A v_j = \frac{1}{\sigma_i \sigma_j} v_i^T \lambda_j v_j = \frac{\sigma_j}{\sigma_i} \langle v_i, v_j \rangle = 0$$

2. כבר הוכחנו. הראנו שאם v הוא ו"ע של A^TA עם ע"ע $\lambda \neq 0$ אז $Av = \sigma_i u$ הוא ו"ע של AA^T.

■

מסקנה 28.6 אם $V = \begin{pmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{pmatrix}$ אז r העמודות הראשונות של U הן:

$$U = \begin{pmatrix} | & & | \\ u_1 & \dots & u_r \\ | & & | \end{pmatrix}$$

הוכחה: נוכיח את משפט הSVD. נראה ש$U^TAV=\Sigma$.
רעיון ההוכחה: **לא לפחד.**

$$U^T=\begin{pmatrix}- & u_1^T & -\\ & \vdots & \\ - & u_r^T & -\\ & \vdots & \\ - & u_m^T & -\end{pmatrix}\in M_{m\times m}$$

$$AV=\begin{pmatrix}| & & | & \\ Av_1 & \dots & Av_r & 0\\ | & & | & \end{pmatrix}=\begin{pmatrix}| & & | & \\ \sigma_1u_1 & \dots & \sigma_ru_r & 0\\ | & & | & \end{pmatrix}\in M_{m\times n}$$

$$U^TAV=\begin{pmatrix}u_1^T\sigma_1u_1 & \dots\\ \vdots & \ddots\end{pmatrix}\overset{\text{orthonormality}}{=}\begin{pmatrix}\sigma_1 & & & & \\ & \sigma_2 & & & \\ & 0 & \ddots & & \\ & & & \sigma_r & \\ & & & & 0\end{pmatrix}=\Sigma=M_{m\times n}$$

■

מסקנה 28.7 ראינו:

1. הערכים הסינגולריים = שורש(ע"ע A^TA, AA^T).
2. העמודות של U = ו"ע של AA^T.
3. העמודות של V=ו"ע של A^TA.
4.

$$\forall i\le r:\ Av_i=\sigma_iu_i$$

ולכן כדי למצוא SVD של A, מספיק למצוא פירוק ספקטרלי של A^TA או של AA^T.

29 תרגיל 7

טענה 29.1 טענות שימושיות לדטרמיננטות:

$$\begin{aligned}\det\left(A^T\right)&=\det(A)\\ \forall B,C\in M_n:\ \det(BC)&=\det(B)\det(C)\\ A\in M_n:\ \det(cA)&=c^n\det(A)\\ \det\left(A^{-1}\right)&=\frac{1}{\det(A)}\end{aligned}$$

טענה 29.2 אם A' מהווה פרמוטציה של שורה\עמודה אחת ביחס לA אז:

$$\det\left(A'\right)=-\det\left(A\right)$$

טענה 29.3 $u,v\in\mathbb{R}^{n}$ אז $uv^{T}\in M_{n}\left(\mathbb{R}\right)$ היא מדרגה 1 או 0.

טענה 29.4 $A\in M_{n}\left(\mathbb{R}\right)$ היא סכום של k מטריצות מדרגה 1.

טענה 29.5 קיימים וקטורים א"נ $v_{1},\cdots,v_{k}\in\mathbb{R}^{n}$ ווקטורים א"נ $v_{1},\cdots,v_{k}\in\mathbb{R}^{m}$ ו$\alpha_{1},\cdots,\alpha_{k}>0$ כך ש$\sum_{i=1}^{k}\alpha_{i}u_{i}v_{i}^{T}=A\in M_{n}\left(\mathbb{R}\right)$

טענה 29.6 נניח $\sigma_{1}\geq\cdots\geq\sigma_{n}\geq0$ הם הערכים הסינגולריים של $A\in M_{m,n}\left(\mathbb{R}\right)$, אז לכל $x\in\mathbb{R}^{m}$:

$$\left\Vert Ax\right\Vert _{2}\geq\sigma_{n}\left\Vert x\right\Vert$$

טענה 29.7 תהי $N_{m}:\mathbb{R}^{m}\to\mathbb{R}$ נורמה על $\mathbb{R}^{m}$, תהי $N_{n}:\mathbb{R}^{n}\to\mathbb{R}$ נורמה על $\mathbb{R}^{n}$. תהי $A\in M_{m,n}\left(\mathbb{R}\right)$, אז:

$$\left\Vert A\right\Vert _{N_{n}\to N_{m}}=\max_{x\in\mathbb{R}^{n}:N_{n}(x)=1}N_{m}\left(Ax\right)$$

טענה 29.8 אי שוויון המשולש ההפוך
אם $\left\Vert \cdot\right\Vert$ היא נורמה אז:

$$\forall x,y:\ \left|\left\Vert x\right\Vert -\left\Vert y\right\Vert \right|\leq\left\Vert x-y\right\Vert$$

30 שיעור 8

30.1 חזרה

משפט 30.1 הSVD
לכל מטריצה ממשית $A_{m\times n}$ יש הצגה:

$$\begin{aligned}A & =UDV^{T}\\
D & =\begin{bmatrix}\sigma_{1} & & & 0\\
 & \ddots\\
 & & \sigma_{n}\\
 & 0
\end{bmatrix}
\end{aligned}$$

כש$U_{m\times m}$ אורתוגונלית, $V_{n\times n}$ אורתוגונלית, D "אלכסונית" ו$\sigma_{1}\geq\cdots\geq\sigma_{m}\geq0$ נקבעים ביחידות.

הערה 30.2 ראינו $\sigma_1 = \|A\|_{op}$.

הערה 30.3 בנוגע ליחידות של הσ_i: המספרים σ_i^2 הם הע"ע של המטריצה הPSD AA^T.

הערה 30.4 אם X מטריצה ממשית כלשהי אז הספקטרום של XX^T ושל X^TX הם זהים למעט הוספה של אפסים במספר הדרוש.

30.2 מסקנות מSVD

הגדרה 30.5 נורמת ℓ_2 של מטריצות קרויה גם נורמת Frobenius או גם Hilbert-Schmidt:

$$\|A\|_F := \sqrt{\sum a_{i,j}^2}$$

מסקנה 30.6 מסקנה נוספת ממשפט הSVD:

$$\|A\|_F^2 \underbrace{=}_{\text{from the definition}} trA^TA \underbrace{=}_{SVD} \sum \sigma_i^2$$

הוכחה: נשים לב:

$$A = UDV^T$$
$$A^TA = VD^TU^TUDV^T = VD^TDV^T$$

טענה 30.7

$$tr\left(PQ\right) = tr\left(QP\right)$$
$$tr\left(P_1P_2\cdots P_k\right) = tr\left(P_kP_1\cdots P_{k-1}\right)$$

הוכחה:

$$tr\left(PQ\right) = \sum_i p_{i,j}q_{j,i}$$

■

לפי הטענה:

$$tr\left(A^TA\right) = tr\left(VD^TDV^T\right) = tr\left(V^TVD^TD\right) = tr\left(D^TD\right) = \sum \sigma_i^2$$

כעת נשים לב:

$$tr\left(MM^T\right) = \sum m_{i,j}m_{j,i}^T = \sum m_{i,j}^2 = \|M\|_F^2$$

וסה"כ קיבלנו את המבוקש. ■

הערה 30.8 נשים לב: $rank\left(A_{m\times n}\right) = d$ה המינימלי כך שניתן להציג $A = X_{m\times d}Y_{d\times n}$.

30.2.1 בעיה

נתונה מטריצה $A_{m\times n}$, רוצים למצוא לה קירוב B "פשוט", למשל $rk\left(B\right)\leq k$.

משפט 30.9 (מסקנה מSVD)
תהיה $A=UDV^T$ (בהצגת SVD), אז הקירוב המיטבי במונחי נורמת פרובניוס והן במונחי הנורמה האופרטורית לA ע"י מטריצה B מדרגה $\geq k$ הוא $B=UD^{(k)}V^T$. $D^{(k)}$ זו המטריצה המתקבלת מD ע"י החלפת $\sigma_{k+1},\cdots$ באפסים.

הוכחה: לקירוב בנורמה אופרטורית.
נשים לב $\left\|A-B\right\|_{op}=\sigma_{k+1}$ מפני שזהו הערך הסינגולרי הגדול ביותר של $A-B$. יש להוכיח, אם כן, שלכל מטריצה C מדרגה $\geq k$ מתקיים $\left\|A-C\right\|_{op}\geq\sigma_{k+1}$. כלומר, שיש z וקטור יחידה כך ש

$$\left\|\left(A-C\right)z\right\|_2\geq\sigma_{k+1}$$

נבנה z כזה כך ש$Cz=0$. מכיוון ש$\text{rank}\left(C\right)\leq k$, לגרעין הימני של C יש מימד $\leq n-k$. כלומר, מטרתנו למצוא וקטור יחידה $z\in r\ker\left(C\right)$ כך ש$\left\|Az\right\|\geq\sigma_{k+1}$.
נסמן את העמודות של V ב: $v_1,v_2,\cdots$. בגלל חשבון המימדים הנ"ל אנחנו יודעים ש

$$r\ker\left(C\right)\cap span\left\{v_1,\cdots,v_{k+1}\right\}\neq\emptyset$$

במילים אחרות, יש וקטור יחידה $z=\sum_{i-1}^{k+1}\alpha_iv_i$ כך ש$Cz=0$ ולכן:

$$\left(A-C\right)z=Az$$

הערה 30.10 ניתן לכתוב את $A=UDV^T$ כך: $A=\sum_{i=1}^{\min(m,n)}\sigma_iu_i\otimes v_i$
מההערה נקבל:

$$Az=\sum_{i}^{\min(m,n)}\sigma_iu_i\otimes v_iz=\sum_{i}^{k+1}\sigma_iu_i\otimes\underbrace{v_iz}_{=v_i\sum_{i-1}^{k+1}\alpha_iv_i\overset{\text{orthonormality}}{=}\alpha}=\sum_{i}^{k+1}\alpha_i\sigma_iu_i$$

$$\Downarrow$$

$$\left\|\left(A-C\right)z\right\|^2=\sum\alpha_i^2\sigma_i^2\geq\sigma_{k+1}^2\underbrace{\sum\alpha_i^2}_{=1\text{ (z is a unit vector, }\{v_i\}\text{ is an orthonormal basis)}}=\sigma_{k+1}^2$$

■

הערה 30.11 נשים לב:

$$xMy^T=\sum x_iy_jM_{i,j}$$

אם $A=UDV^T$ (פירוק SVD) אז נחשוב על U בתור מטריצת בלוקים שכל בלוק הוא עמודה u_i, ונחשוב על V^T בתור מטריצת בלוקים באותו האופן:

$$A=\begin{bmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{bmatrix}\begin{bmatrix} \sigma_1 & \\ & \ddots \\ & \end{bmatrix}\begin{bmatrix} - & v_1 & - \\ & \vdots & \\ - & v_m & - \end{bmatrix}$$

$$\Downarrow$$

$$A=\sum_{i}^{\min(m,n)} \sigma_i u_i \otimes v_i$$

משפט 30.12 ריילי־ריטז Rayleigh-Ritz

תהיה $A_{n\times n}$ מטריצה ממשית סימטרית, ויהיו $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_n$ הע"ע של A, אז:

$$\lambda_1=\max_{x\neq 0}\underbrace{\frac{xAx^T}{\|x\|_2^2}}_{\text{Rayleigh's fraction}}=\max_{\|x\|_2=1} xAx^t$$

יהיו $v_1,\cdots,v_n$ ו"ע המתאימים ל$\lambda_1,\cdots,\lambda_n$, אז

$$\lambda_{k+1}=\max_{x\perp v_1,\cdots,v_k}\frac{xAx^T}{\|x\|_2^2}$$

ולדקדקנים:

$$\lambda_{k+1}=\max_{x\perp\{\text{any eigenvector of }\lambda_1,\cdots,\lambda_k\}}\frac{xAx^T}{\|x\|_2^2}$$

הערה 30.13 אם A ממשית סימטרית, ו

$$Ax=\lambda x,\ Ay=\mu y,\ \lambda\neq\mu$$

אז $x\perp y$.

הוכחה: נביט:

$$\begin{aligned}\lambda\langle x,y\rangle&=xAy=\mu\langle x,y\rangle\\ &\Downarrow \lambda\neq\mu\\ \langle x,y\rangle&=0\\ &\Downarrow\\ x&\perp y\end{aligned}$$

■

30.3 כופלי לגרנז'

30.3.1 מה רוצים

רוצים למצוא את ערכי max/min של $f(x_1,\cdots,x_n)$ בכפוף לכך ש:

$$\begin{aligned} g_1(x_1,\cdots,x_n)&=0\\ &\vdots\\ g_k(x_1,\cdots,x_n)&=0 \end{aligned}$$

30.3.2 השיטה

מגדירים:

$$F(x_1,\cdots,x_n):=f(x_1,\cdots,x_n)+\sum_{i=1}^{k}\lambda_i g_i(x_1,\cdots,x_n)$$

λ_i הם משתנים פורמליים שנקראים כופלי לגראנז'. פותרים:

$$\begin{aligned} &\forall i\in[n]:\frac{\partial F}{\partial x_i}=0\\ &\forall j\in[k]:g_j(x_1,\cdots,x_n)=0 \end{aligned}$$

משפט 30.14 כל הנקודות הקריטיות של f על המשטח $\forall i\in[k]:\ g_i=0$ מתקבלות כך.

הערה 30.15 התנאי במשפט של כופלי לגראנז' אומר ש$grad(F)$ הוא ניצב למשטח המוגדר ע"י המשוואות:

$$g_1=\cdots=g_k=0$$

30.3.3 נשתמש

הוכחה: נוכיח את משפט $Rayleigh-Ritz$. נשתמש בכופלי לגראנז': נשים לב ש$\max_{\|x\|_2=1}xAx^t$ הוא בעצם $\max\sum a_{ij}x_ix_j$ בכפוף לתנאי $\sum x_i^2=1$.

$$\begin{aligned} F(x_1,\cdots,x_n)&=\sum_{1\le i,j\le n}a_{ij}x_ix_j-\lambda\sum x_i^2\\ \forall t:\frac{\partial F}{\partial x_t}&\overset{\text{we want}}{=}0\\ &\Updownarrow\text{we'll show}\\ 2(Ax)_t&=2\lambda x_t \end{aligned}$$

נראה זאת:

$$\frac{\partial}{\partial x_t}\left(\sum_{1\le i,j\le n} a_{ij}x_ix_j\right) = \frac{\partial}{\partial x_t}\left(a_{tt}x_t^2 + \sum_{i\neq t} a_{it}x_ix_t + \sum_{j\neq t} a_{tj}x_tx_j\right) =$$

$$= 2a_{tt}x_t + \sum_{i\neq t} a_{it}x_i + \sum_{j\neq t} a_{tj}x_j =$$

$$= \underbrace{\sum_i a_{it}x_i}_{(xA)_t} + \underbrace{\sum_j a_{tj}x_j}_{(Ax)_t} \overset{symmetry}{=} 2\left(Ax\right)_t$$

$$\Downarrow \frac{\partial F}{\partial x_t} = 0$$

$$\frac{\partial}{\partial x_t}\left(\sum_{1\le i,j\le n} a_{ij}x_ix_j - \lambda\sum x_i^2\right) = 0$$

$$\Updownarrow$$

$$2\left(Ax\right)_t = 2\lambda x_t$$

ריילי־ריץ עבור $\lambda_1, k=0$: יהי x ווקטור יחידה. כעת נשתמש במשפט הספקטרלי:

$$xAx^T = \underbrace{xV}_{vec}\Lambda\underbrace{V^Tx^T}_{vec^T} =^* \sum \lambda_i\left(xV\right)_i^2 \le^1$$

$=^*$ נובע מ:

$$\varphi M\psi = [\varphi_i\psi_j m_{ij}]$$
$$\Downarrow M \text{ is diagonal}$$
$$\varphi M\psi = [\varphi_i\psi_i m_{ii}]$$

נמשיך:

$$\le^1 \max\lambda_i \underbrace{\sum\left(xV\right)_i^2}_{\text{x is a unit vector, V is orthogonal, so }=1} = \lambda_1\cdot 1 = \lambda_1$$

מצד שני, אם נקח $x=v_1$ אז

$$v_1Av_1^T = \lambda_1 v_1v_1^T = \lambda_1$$

כעת המקרה הכללי: נחזור על אותו החשבון. עכשיו, כיוון ש$x\perp v_1,\cdots,v_k$ יוצא ש$\left(Vx\right)_i=0$ כש $i=1,\cdots,k$.

■

30.4 משפטים

משפט 30.16 בדומה ל$Rayleigh-Ritz$
תהיה $A_{n\times n}$ מטריצה ממשית סימטרית, ויהיו $\lambda_n \geq \cdots \geq \lambda_2 \geq \lambda_1$ הע"ע של A, יהיו $v_n, \cdots, v_1$ ו"ע המתאימים ל$\lambda_n, \cdots, \lambda_1$, אז בדומה ל$Rayleigh-Ritz$:

$$\lambda_n = \min \frac{xAx^T}{\|x\|_2^2}$$
$$\lambda_{k-1} = \min_{x\perp v_k,\cdots,v_n} \frac{xAx^T}{\|x\|_2^2}$$

משפט 30.17 קורנט־פישר (Courant-Fisher)
תהיה $A_{n\times n}$ ממשית סימטרית, ויהיו $\lambda_n \geq \cdots \geq \lambda_1$ הע"ע שלה, אז:

$$\forall i : \lambda_{i+1} = \min_{F:\dim F=i} \max_{x\perp F} \frac{xAx^T}{\|x\|_2^2}$$
$$\lambda_i = \min_{G:\dim G=n-i} \min_{x\perp G} \frac{xAx^T}{\|x\|_2^2}$$

מסקנה 30.18 משפט השירוג (interlacing)
תהיה $A_{n\times n}$ מטריצה ממשית סימטרית ויהיו $\lambda_n \geq \cdots \geq \lambda_1$ הע"ע שלה. נמחק לאיזשהו i את השורה הi ואת העמודה הi בA ונקבל את המטריצה $B_{(n-1)\times(n-1)}$ עם ע"ע $\mu_{n-1} \geq \cdots \geq \mu_1$ אז:

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \cdots \geq \lambda_{n-1} \geq \mu_{n-1} \geq \lambda_n$$

הוכחה: איך מסיקים זאת מקורנט־פישר? נמשיך לעבוד עם A, אבל במקום להרשות כל וקטור x, נרשה רק כזה בשבילו הקורדינטה הi היא 0. ∎

משפט 30.19 פרון־פרובניוס (Perron-Frobenios)
תהיה $M_{n\times n}$ מטריצה ממשית כך שכל איבריה $m_{ij} \geq 0$, ויהיה

$$\rho(M) = \max\{|\lambda| \, | \lambda \text{ is an eigenvalue of M}\}$$

ל$\rho(M)$ קוראים הרדיוס הספקטרלי של M. אזי $\rho(M)$ הוא ע"ע פשוט (בלי ריבוי) של M, ולכל שאר הע"ע יש מודולוס $< \rho(M)$. הו"ע המתאים לע"ע $\rho(M)$ הוא חיובי. לכן, אם x_l, x_r הו"ע הימני והשמאלי בהתאמה של הע"ע $\rho(M)$, ז"א

$$x_l M = \rho(M)\, x_l$$
$$M x_r = \rho(M)\, x_r$$

אזי:

$$\lim_{t\to\infty}\left(\frac{M}{\rho\left(M\right)}\right)^{t}=\frac{x_{r}x_{l}^{T}}{\left\langle x_{r},x_{l}\right\rangle }$$

יהיה $G=\left(V,E\right)$ הגרף המכוון שבו $m_{ij}<0\Leftrightarrow\left(i\to j\right)\in E$, נניח ש$G$ קשיר חזק ולא בעל מבנה ציקלי.

הערה 30.20

$$\begin{aligned}z & =x+iy\\ \left|z\right| & =\sqrt{x^{2}+y^{2}}\end{aligned}$$

31 תרגול 8

31.1 גישה וריאציונית לערכים סינגולריים

הגדרה 31.1 נקודה קריטית
יהיו $U\subseteq\mathbb{R}^{n}$ פתוחה ו$f:U\to\mathbb{R}$ גזירה. $x\in U$ נקודה קריטית של f אם

$$\Delta f\left(x\right)=0$$

כלומר:

$$\forall i\in\left[n\right]:\ \frac{\partial f}{\partial x_{i}}\left(x\right)=0$$

משפט 31.2 תהי $A\in M_{n}\left(\mathbb{R}\right)$ מטריצה סימטרית, ו$\lambda\in\mathbb{R}$ ו$v\in\mathbb{R}^{n}$, התב"ש:

- $Av=\lambda v$
- v נקודה קריטית של הפונקציה f:

$$\begin{aligned}f & :\mathbb{R}^{n}\setminus\left\{ 0\right\} \to\mathbb{R}\\ f\left(x\right) & =\frac{x^{T}Ax}{\left\Vert x\right\Vert _{2}^{2}}\\ f\left(v\right) & =\lambda\end{aligned}$$

הערה 31.3 תזכורת: σ ערך סינגולרי של A אם"ם σ^{2} הוא ע"ע של $A^{T}A$.

הוכחה:

$$A^TA=\left(UDV^T\right)^T\left(UDV^T\right)=VD^TU^TUDV^T=V\left(D^TD\right)V^T=V\begin{pmatrix}\sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2\end{pmatrix}V^T$$

■

הגדרה 31.4 עבור $A\in M_{m\times n}\left(\mathbb{R}\right)$ ווקטורי יחידה $u\in\mathbb{R}^m,\ v\in\mathbb{R}^n$ ו-$\sigma\geq 0$, אם

$$\begin{aligned}Av&=\sigma u\\ A^Tu&=\sigma v\end{aligned}$$

אז v נקרא וקטור סינגולרי ימני, u נקרא וקטור סינגולרי שמאלי, σ נקרא ערך סינגולרי.

הוכחה: נראה שההגדרה השניה שקולה להערה הקודמת.
כיוון 1: נניח ש-u,v,σ סינגולריים, אז:

$$A^TAv=A^T\sigma u=\sigma^2v$$

כיוון 2: נניח $A^TAv=\sigma^2v$. אם נקח $u=\frac{Av}{\sigma}$ אז הכל יעבוד. ■

משפט 31.5 תהי $A\in M_{m\times n}\left(\mathbb{R}\right)$, יהיו $u\in\mathbb{R}^m,v\in\mathbb{R}^n$ וקטורי יחידה, $\sigma\geq 0$. התב"ש:

- u,v,σ סינגולריים
- (v,u) היא נקודה קריטית של הפונקציה:

$$\begin{aligned}f:\mathbb{R}^n\setminus\{0\}\times\mathbb{R}^m\setminus\{0\}&\to\mathbb{R}\\ (x,y)&\to\frac{y^TAx}{\|y\|_2\|x\|_2}\\ f(v,u)&=\sigma\end{aligned}$$

הוכחה: נחשב:

$$\begin{gathered}\left(\frac{u}{v}\right)'=\frac{u'v-v'u}{v^2}=0\\ \Updownarrow\\ u'v=v'u\end{gathered}$$

ולכן:

$$\frac{\partial f\left(x,y\right)}{\partial x_i}=0$$
$$\Updownarrow$$
$$\frac{\partial\left(y^TAx\right)}{\partial x_i}\left\|y\right\|\left\|x\right\|=y^TAx\frac{\partial\left(\left\|y\right\|\left\|x\right\|\right)}{\partial x_i}$$

נשים לב ש:

$$y^TAx=\sum_{k=1}^{n}\sum_{l=1}^{n}y_lA_{lk}x_k$$
$$\Downarrow$$
$$\frac{\partial\left(y^TAx\right)}{\partial x_i}=\sum_{l=1}^{n}y_lA_{li}$$
$$\Downarrow$$
$$\frac{\partial\left(y^TAx\right)}{\partial x_i}\left\|y\right\|\left\|x\right\|=\left(\sum_{l=1}^{m}y_lA_{li}\right)\left\|y\right\|\left\|x\right\|$$

ו:

$$y^TAx\frac{\partial\left(\left\|y\right\|\left\|x\right\|\right)}{\partial x_i}=y^TAx\left\|y\right\|\frac{\partial\left(\left\|x\right\|\right)}{\partial x_i}=y^TAx\left\|y\right\|\frac{\partial\sqrt{\sum_{l=1}^{n}x_l^2}}{\partial x_i}=$$
$$=y^TAx\left\|y\right\|\frac{1}{2\left\|x\right\|}\cdot 2x_i=y^TAx\left\|y\right\|\frac{x_i}{\left\|x\right\|}$$

כלומר:

$$\frac{\partial f\left(x,y\right)}{\partial x_i}=0$$
$$\Updownarrow$$
$$\left(\sum_{l=1}^{m}y_lA_{li}\right)\left\|y\right\|\left\|x\right\|=y^TAx\frac{\left\|y\right\|}{\left\|x\right\|}x_i$$

לפי חישוב דומה:

$$\frac{\partial f\left(x,y\right)}{\partial y_i}=0$$
$$\Updownarrow$$
$$\left(\sum_{l=1}^{n}A_{il}x_l\right)\left\|y\right\|\left\|x\right\|=y^TAx\frac{\left\|x\right\|}{\left\|y\right\|}y_i$$

כל המשוואות האלה שקולות ל:

$$\begin{pmatrix} 0 & A^T \\ A & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}\|x\|\|y\| = y^T Ax \begin{pmatrix} \frac{\|y\|}{\|x\|}x \\ \frac{\|x\|}{\|y\|}y \end{pmatrix}$$

נניח שu, v, σ סינגולריים. אזי צ"ל: $\Delta f(v,u) = 0$ ו$f(v,u) = \sigma$. אז כיוון שv, u וקטורי יחידה:

$$\begin{pmatrix} 0 & A^T \\ A & 0 \end{pmatrix}\begin{pmatrix} v \\ u \end{pmatrix} = \begin{pmatrix} A^T u \\ Av \end{pmatrix} = \begin{pmatrix} \sigma v \\ \sigma u \end{pmatrix} = \begin{pmatrix} v \\ u \end{pmatrix}\sigma =^* \begin{pmatrix} v \\ u \end{pmatrix} u^T Av$$

$$=^* \text{ is because:} u^T(Av) = u^T\sigma u \overset{u \text{ is a unit vector}}{=} \sigma$$

ועכשיו הכיוון השני: נניח ש$u, v, \sigma = f(v,u)$ כש u, v וקטורי יחידה, ו(v,u) נקודה קריטית של f. צ"ל $Av = \sigma u$ ו$A^T u = \sigma v$. כיוון שזו נקודה קריטית נציב:

$$\begin{pmatrix} 0 & A^T \\ A & 0 \end{pmatrix}\begin{pmatrix} v \\ u \end{pmatrix} = \underbrace{v^T Au}_{\sigma}\begin{pmatrix} v \\ u \end{pmatrix}$$

$$\Downarrow$$

$$\begin{pmatrix} A^T u \\ Av \end{pmatrix} = \begin{pmatrix} \sigma v \\ \sigma u \end{pmatrix}$$

■

32 תרגול אקסטרה 8

32.1 מטריצות שכנויות של גרפים

הגדרה 32.1 מטריצת שכנויות של גרף
יהי $G = (V = \{1, \cdots, n\}, E)$ גרף לא מכוון. A היא מטריצת השכנויות של G:

$$A_{ij} = \begin{cases} 1 & \{i,j\} \in E \\ 0 & otherwise \end{cases}$$

מסקנה 32.2 אבחנה: A סימטרית ולכן יש לה פירוק ספקטרלי:

$$A = V\Lambda V^T$$

טענה 32.3

$$\left(A^2\right)_{i,j} = \# \text{ of paths from i to j of length 2}$$

הוכחה:

$$\left(A^2\right)_{i,j}=\sum_{k=1}^{n}\underbrace{A_{i,k}A_{k.j}}_{=\begin{cases}1 & \{i,k\},\{k,j\}\in E\\ 0 & otherwise\end{cases}}$$

■

טענה 32.4 נניח ש$\lambda_n \geq \cdots \geq \lambda_1$ הם הע"ע של A, אז דרגה מקסימלית ב$G \geq^2 \lambda_1 \geq^1$ דרגה ממוצעת ב$G>0$.

הוכחה: נראה את $\geq^1$: לכל $0\neq x\in\mathbb{R}^n$ אז:

$$\lambda_1 \overset{Rayleigh-Ritz}{\geq} \frac{x^TAx}{x^Tx}$$

$$x^TAx=\sum_{j,i=1}^{n}A_{ij}x_ix_j=\sum_{i\sim j}1\cdot1\cdot1=2\left|E\right|=\sum_{v\in V}\deg\left(V\right)$$

נסתכל על $x=\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}$ אז $x^Tx=n$, ולכן:

$$\lambda_1\geq\frac{1}{n}\sum_{v\in V}\deg\left(v\right)$$

נראה את $\geq^2$: צ"ל כל ע"ע λ של A מקיים:

$$\left|\lambda\right|\leq\triangle=\text{max degree in G}$$

נניח ש$Ax=\lambda x$, יהי i כך ש$|x_i|$ מקסימלי. אז:

$$\left|\lambda\right|\left|x_i\right|=\left|\lambda x_i\right|=\left|\left(Ax\right)_i\right|=\left|\sum_{j:\ j\sim i}x_j\right|\overset{T.I}{\leq}\sum_{j:\ j\sim i}\left|x_j\right|\leq d_i\left|x_i\right|\leq\triangle\left|x_i\right|$$

■

מסקנה 32.5 אם G הוא d־רגולרי אז $\lambda_1=d$ וגם $v_1=\frac{1}{\sqrt{n}}\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}$.

מסקנה 32.6

$$\lambda_n<0$$

הוכחה:

$$0=\sum A_{i,i}=tr\left(A\right)=\sum_{i=1}^{n}\lambda_i$$

■

הגדרה 32.7 אנטי־קליקה
קבוצה $S\subseteq V$ נקראת אנטי־קליקה אם:

$$\forall u,v\in S:\ u\nsim v$$

הערה 32.8 זו בעיה חישובית חשובה ו־NP־קשה לחסום את גודלה של האנטי־קליקה הגדולה ביותר בגרף נתון G.

משפט 32.9 אם S היא אנטי־קליקה בגרף d־רגולרי עם n קודקודים, אז

$$\frac{|S|}{n}\leq\frac{-\lambda n}{d-\lambda n}=\frac{-\lambda n}{\lambda_1-\lambda n}$$

הוכחה: נסתכל על הוקטור המציין של S:

$$x_i=\begin{cases}1 & i\in S\\ 0 & i\notin S\end{cases}$$

נחשב את x^TAx בשני אופנים:

1.

$$x^TAx=\sum_{i\sim j}x_ix_j\overset{\text{S is an anticlique}}{=}0$$

2. נציג את x כצ"ל של ו"ע:

$$\begin{aligned}x&=\sum_{i=1}^{n}\alpha_iv_i\\ &\Downarrow\\ x^TAx&=\left\langle\sum_{i=1}^{n}\alpha_iv_i,\sum_{i=1}^{n}\alpha_i\lambda_iv_i\right\rangle\overset{\{v_i\}\text{ are orthonormal}}{=}\sum_{i=1}^{n}\alpha_i^2\lambda_i\geq\\ &\geq\underbrace{\alpha_1^2d}_{i=1}+\lambda_n\sum_{i=2}^{n}\alpha_i^2=^*\end{aligned}$$

נשים לב:

$$v_1 = \frac{1}{\sqrt{n}}\begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix}$$

$$\Downarrow$$

$$1: \alpha_1 = \langle x, v_1\rangle = \frac{|S|}{\sqrt{n}}$$

$$2: \ |S| = \|x\|^2 = \left\langle \sum_{i=1}^n \alpha_i v_i, \sum_{i=1}^n \alpha_i v_i \right\rangle = \sum_{i=1}^n \alpha_i^2$$

$$\Downarrow$$

$$=^* \alpha_1^2 d + \underbrace{\lambda_n\left(|S| - \alpha_1^2\right)}_{\text{from 2}} = \lambda_n |S| + \underbrace{\frac{|S|^2}{n}}_{\text{from 1}}(d-\lambda_n)$$

נאחד את שתי צורות ההצגה של x^TAx ונקבל:

$$\underbrace{0}_{\text{first}} \geq \underbrace{\lambda_n |S| + \frac{|S|^2}{n}(d-\lambda_n)}_{\text{second}}$$

$$\Downarrow$$

$$\frac{|S|}{n} \leq \frac{-\lambda_n}{d-\lambda_n}$$

■

33 תרגיל 8

טענה 33.1 יהיו $(V, \langle\cdot,\cdot\rangle_V), (W, \langle\cdot,\cdot\rangle_W)$ מרחבי מכפלה פנימית סופיים מעל $\mathbb{R}$, אז קיים מיפוי לינארי יחודי: $T^* = T^T$ המקיים:

$$\forall v \in V, \ \forall w \in W: \ \langle Tv, w\rangle_W = \langle v, T^*w\rangle_V$$

טענה 33.2 יהיו $(V, \langle\cdot,\cdot\rangle_V), (W, \langle\cdot,\cdot\rangle_W)$ מרחבי מכפלה פנימית סופיים מעל $\mathbb{R}$, אז קיימים בסיסים א"נ $v_1, \cdots, v_n$ לV ו$w_1, \cdots, w_n$ לW ו$\sigma_1 \geq \cdots \geq \sigma_n \geq 0$ כך ש:

$$\forall v \in V : Tv = \sum_{i=1}^n \langle v, v_i\rangle_V \sigma_i w_i$$

$$\forall w \in W : T^*w = \sum_{i=1}^n \langle w, w_i\rangle_W \sigma_i v_i$$

הגדרה 33.3 מטריצה $positive-definite$

$A\in M_n(\mathbb{R})$ תקרא $positive-definite$ אם לכל $x\in\mathbb{R}^n$ מתקיים $x^TAx=\langle Ax,x\rangle\geq 0$ ושוויון אס"ם $x=0$.

הגדרה 33.4 מטריצה $positive-semidefinite$

$A\in M_n(\mathbb{R})$ תקרא $positive-semidefinite$ אם לכל $x\in\mathbb{R}^n$ מתקיים $x^TAx=\langle Ax,x\rangle\geq 0$.

טענה 33.5 יהי V מ"ו ו$\{v_i\}_{i=1}^n$ בסיס א"נ לו תחת המכפלה הפנימית $\langle\cdot,\cdot\rangle_V$. תהי $\langle\cdot,\cdot\rangle_1$ מכפלה פנימית על V, אז קיימת מטריצה ייחודית $A\in M_n(\mathbb{R})$ כך שA סימטרית ו$positive-definite$, וכך שלכל $u,v\in V$ מתקיים:

$$\langle u,v\rangle_1=\sum_{i=1}^n\sum_{j=1}^n\langle u,v_i\rangle_V\langle v,v_j\rangle_V A_{i,j}$$

טענה 33.6 יהי V מ"ו ו$\{v_i\}_{i=1}^n$ בסיס א"נ לו תחת המכפלה הפנימית $\langle\cdot,\cdot\rangle_V$. תהי B סימטרית ו$positive-definite$, נגדיר:

$$T_B=\sum_{j=1}^n B_{i,j}v_j$$
$$\langle x,y\rangle_B=\langle T_Bx,y\rangle_V$$

אז $\langle\cdot,\cdot\rangle_B$ היא מכפלה פנימית על V.

הגדרה 33.7 קונוס

$\emptyset=C\subseteq\mathbb{R}^m$ הוא קונוס אם לכל $x,y\in C$ ו$\alpha,\beta\geq 0$ מתקיים: $\alpha x+\beta y\in C$

טענה 33.8 יהי $P_n=\{A\in M_n(\mathbb{R}):\ \text{A is symmetric and positive-semidefinite}\}$, אז:

1. P_n הוא קונוס.
2. $\forall v\in\mathbb{R}^n:\ vv^T\in P_n$
3. לכל $A\in P_n$ קיימים $\{v_i\}_{i=}^n\in\mathbb{R}^n$ אורתוגונליים ו$\lambda_1\geq\cdots\geq\lambda_n\geq 0$ כך ש $A=\sum_{i=}^n\lambda_iv_iv_i^T$.

34 שיעור 9

34.1 שיטת Monte-Carlo להערכת הסתברות של מאורע

ישנה התפלגות P ומאורע A. נרצה להעריך את $p(A)$. נגריל באופן ב"ת סדרה של נקודות $\omega_1,\cdots,\omega_N\in\Omega$ לפי התפלגות P. נגדיר מ"מ:

$$X_i=\begin{cases}1 & \omega_i\in A\\ 0 & otherwise\end{cases}=\begin{cases}1 & p\\ 0 & 1-p\end{cases}$$

נגדיר $\overline{X}=\frac{1}{N}\sum_{i=1}^{N}X_i$. נשים לב:

$$\mathbb{E}\left[\overline{X}\right]=\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[X_i\right]=p$$

$$Var\left[\overline{X}\right]=\frac{1}{N^2}Var\left[\sum_{i=1}^{N}X_i\right]\overset{\{X_i\}\text{are i.i.d}}{=}\frac{1}{N^2}\sum_{i=1}^{N}Var\left[X_i\right]=\frac{1}{N^2}Np\left(1-p\right)=\frac{p\left(1-p\right)}{N}\leq\frac{1}{4N}$$

לפי א"ש צ'בישב:

$$\Pr\left(\left|\overline{X}-p\right|\geq\epsilon p\right)=\Pr\left(\left|\overline{X}-\mathbb{E}\left[\overline{X}\right]\right|\geq\epsilon p\right)\leq\frac{Var\left[\overline{X}\right]}{\epsilon^2p^2}\leq\frac{1}{4N\epsilon^2p^2}$$

עבור $N>>\frac{1}{\epsilon^2p^2}$ נקבל כי בסיכוי גבוה $\left|\overline{X}-p\right|\leq\epsilon p$.

34.1.1 שיטת MCMC - Markov Chain Monte Carlo (Broder, 86)

בעיה

נתון גרף דו צדדי $G=(V,E)$, כש $|V|=n$. כמה זיווגים מושלמים יש בגרף הזה? הבעיה הזו היא NP קשה (Valiant, 79).

הצעה

נגדיר (Ω,P) כש Ω היא אוסף כל הזיווגים המושלמים, וP היא התפלגות אחידה. נניח שנוכל להגריל זיווג $M\in\Omega$ לפי התפלגות אחידה. נגדיר $\Omega\supseteq A=\{M_0\}$ כש $M_0\in\Omega$ הוא זיווג שמצאנו. יהי Z מספר הזיווגים. נקבל:

$$\Pr\left(A\right)=\frac{1}{Z}$$

הצעה אחרת לA

נקבע צלע $e\in E$ ונגדיר $A=\{M\in\Omega:\ e\in M\}$ (מספר הזיווגים שעוברים דרך צלע מסוימת). $\Pr\left(A\right)$ "אמורה להיות" גדולה $\sim\frac{1}{n}$, ולכן ב$\theta\left(n^2\right)$ הגרלות של זיווג נוכל להעריך $\Pr\left(A\right)$ עם שגיאה כפלית ϵ קטנה. נשים לב:

$$\Pr\left(A\right)=\frac{\#\left\{M:\ e\in M\right\}}{\#\left\{M\right\}}\cdot\frac{\#\left\{M:\ e,e'\in M\right\}}{\#\left\{M:\ e\in M\right\}}\cdots=\frac{1}{\#\left\{M\right\}}$$

(הרעיון: עובדים בשיטת ה"בצל", קודם בוחרים צלע אחת, מורידים אותה ואת הקודקודים שלה, ואז מסתכלים על הגרף שנשאר. בוחרים שוב צלע, וכו').

איך נוכל להגריל בצורה יעילה $M\in\Omega$ בהתפלגות אחידה? נבצע מהלך מקרי על גרף של הזיווגים.

שרשראות מרקוב

נתון גרף $G=(V,E)$ כש$V=\{1,\cdots,n\}$ ולכל שני קודקודים הסתברות המעבר ביניהם. תהי λ_0 התפלגות התחלתית של מיקום המהלך, ונגדיר λ_t לכל $t\geq 0$ להיות התפלגות אחרי t צעדים. רוצים להבין את ההתפלגות של λ_t כאשר $t\to\infty$. נגדיר X_t להיות המשתנה המקרי המתאר את מצב המהלך אחרי t צעדים ואז $X_t\sim\lambda_t$.

הגדרה 34.1 מטריצת המעבר של המהלך המקרי
מטריצה P עם n שורות וn עמודות:

$$P_{ij}=\Pr\left(X_{t+1}=j|X_t=i\right)$$

נניח שP_{ij} לא תלוי בt (זה נקרא שרשרת מרקוב הומוגנית).

דוגמה מהלך מקרי על מעגל עם שלושה קודקודים:

$$P=\begin{array}{cccc} & 1 & 2 & 3\\ 1 & 0 & \frac{1}{2} & \frac{1}{2}\\ 2 & \frac{1}{2} & 0 & \frac{1}{2}\\ 3 & \frac{1}{2} & \frac{1}{2} & 0\end{array}=\begin{pmatrix}\frac{1}{2} & \frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2}\end{pmatrix}-\begin{pmatrix}\frac{1}{2} & & \\ & \frac{1}{2} & \\ & & \frac{1}{2}\end{pmatrix}=\frac{1}{2}\underbrace{J}_{\text{all ones}}-\frac{1}{2}I=\frac{1}{2}\left(J-I\right)$$

$$\lambda_0=\begin{bmatrix}1 & 0 & 0\end{bmatrix}$$

$$\lambda_1=?$$

טענה 34.2 $\lambda_1=\lambda_0P$.

הוכחה: נשים לב:

$$\lambda_1\left(j\right)=\sum_{i=1}^{3}\Pr\left(X_1=y|X_0=i\right)\Pr\left(X_0=i\right)=\sum_{i=1}^{3}P_{ij}\lambda_0\left(i\right)=\left(\lambda P\right)\left(j\right)$$

■

אחרי t צעדים נקבל

$$\lambda_t=\lambda_{t-1}P=\cdots=\lambda_0P^t$$

נשים לב שהע"ע של J הם: $3,0,0$ והע"ע של I הם: $1,1,1$. אם כך, כיוון ש

$$P=\frac{1}{2}\left(P-I\right)$$

אז כיוון שP וI מתחלפות הע"ע של P הם $1,-\frac{1}{2},-\frac{1}{2}$. נסמן את הו"ע v_1,v_2,v_3. זהו בסיס כי P מטריצה סימטרית ואי שלילית, אז נסמן:

$$\begin{aligned}&\lambda_0=a_1v_1+a_2v_2+a_3v_3\\ &\Downarrow\\ &\lambda_0P^t=\left(a_1v_1+a_2v_2+a_3v_3\right)P^t=a_1v_1P^t+a_2v_2P^t+a_3v_3P^t=\\ &=a_1\left(1\right)^tv_1+a_2\left(-\frac{1}{2}\right)^tv_2+a_3\left(-\frac{1}{2}\right)^tv_3\underset{t\to\infty}{\rightarrow}a_1\begin{pmatrix}1 & 1 & 1\end{pmatrix}=\begin{pmatrix}\frac{1}{3} & \frac{1}{3} & \frac{1}{3}\end{pmatrix}\end{aligned}$$

נשים לב ש$a_1 = \frac{1}{3}$ כי אנו רוצים לקבל התפלגות.

הערה 34.3 נשים לב שהטכניקה הזו לא תעבוד ב4 כי תמיד נדע מאיפה התחלנו (אחרי מס' זוגי של צעדים ניהיה בקודקוד אי זוגי ואחרי מס' אי זוגי של צעדים נהיה בקודקוד זוגי). הבעיה היא שב4 הגרף דו צדדי.

הגדרה 34.4 שרשרת מרקוב
סדרה של משתנים מקריים $X_0, X_1, \cdots$ המקבלים ערכים ב$\{1, \cdots, n\}$ ועם התכונה הבאה: לכל $1 \le t$ מתקיים:

$$\Pr\left(X_t = a_t | X_0 = a_0, \cdots, X_{t-1} = a_{t-1}\right) = \Pr\left(X_t = a_t | X_{t-1} = a_{t-1}\right)$$

מטריצת המעבר של השרשרת היא

$$P_{ij}^t = \Pr\left(X_t = j | X_{t-1} = i\right)$$

ומניחים כי היא לא תלויה בt.

מסקנה 34.5 בתנאים האלה, לפי נוסחת ההסתברות השלמה, אם נסמן בλ_t את ההתפלגות של X_t מתקיים:

$$\begin{gathered} \lambda_{t+1} = \lambda_t P \\ \Downarrow \\ \lambda_t = \lambda_0 P^t \end{gathered}$$

משפט 34.6 פרון־פרובניוס Perron-Frobenius
תהי P מטריצה $n \times n$ עם מקדמים אי־שליליים. ונניח כי מתקיימים שני תנאים:

1. הגרף המכוון $G = (V, E)$ המוגדר על ידי P הוא קשיר.
2. לא ניתן לחלק את קבוצת הקודקודים V לקבוצות $S_0, \cdots, S_{d-1}$ כך ש:

$$E \subseteq (S_0 \times S_1) \cup (S_1 \times S_2) \cup \cdots \cup (S_{d-1} \times S_0)$$

יהי $\rho(P)$ הרדיוס הספקטרלי של P:

$$\rho(P) = \max\left\{|\lambda| : \ \lambda \text{ is an eigenvalue of P}\right\}$$

אזי $\rho(P)$ הוא ע"ע של P. יהיה x הו"ע מימין שמתאים לע"ע $\rho(P)$ וy הו"ע משמאל:

$$\begin{aligned} y^T P &= \rho(P)\, y^T \\ Px &= \rho(P)\, x \end{aligned}$$

אז x, y וקטורים חיוביים ו:

$$\left(\frac{P}{\rho(P)}\right)^t \underset{t\to\infty}{\to} \frac{xy^T}{\langle x, y\rangle}$$

טענה 34.7 תהי P מטריצת מעבר של שרשרת מרקוב, ונניח כי P מקיימת את התכונות של משפט PF. אז $\rho\left(P\right)=1$ והוקטור העצמי מימין של P המתאים לע"ע 1 הוא $\vec{1}=\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}=x$.

הוכחה: אמנם, כל שורה של P מסתכמת ל1 (ז"א P מטריצה סטוכסטית), ולכן $P\cdot\vec{1}$, כלומר $\vec{1}$ הוא ו"ע של P מימין עם ע"ע 1. נניח בשלילה כי $\rho=\rho\left(P\right)>1$ אזי קיים לפי PF וקטור עצמי חיובי y של P משמאל כך ש$yP=\rho y$. נגיע לסתירה על ידי כך שנראה ש $\left\langle y,\vec{1}\right\rangle=0$ (סתירה לכך שy חיובי). אמנם נחשב את $\left\langle yP,\vec{1}\right\rangle$ בשתי צורות.
מצד אחד:

$$\left\langle yP,\vec{1}\right\rangle=\left\langle P^{T}y,1\right\rangle=\left\langle \rho y,1\right\rangle=\rho\left\langle y,1\right\rangle$$

מצד שני:

$$\left\langle P^{T}y,1\right\rangle=\left\langle y,P1\right\rangle=\left\langle y,1\right\rangle$$

ולכן קיבלנו:

$$\left\langle y,1\right\rangle=\rho\left\langle y,1\right\rangle$$

ולכן, כיוון ש$\rho\neq1$ אז:

$$\left\langle y,1\right\rangle=0$$

והגענו לסתירה. ■

מסקנה 34.8 עבור מטריצה סטוכסטית P המקיימת את הדרישות של PF מתקיים

$$P^{t}\to\frac{\vec{1}\cdot y^{T}}{\left\langle \vec{1},y\right\rangle}=\frac{1}{\sum_{i=1}^{n}y_{i}}\begin{pmatrix}- & y & -\\ & \vdots\\ - & y & -\end{pmatrix}=\begin{pmatrix}- & \frac{y}{\sum y_{i}} & -\\ & \vdots\\ - & \frac{y}{\sum y_{i}} & -\end{pmatrix}$$

$\frac{y}{\sum y_{i}}$ מסומן בΠ והוא נקרא ההתפלגות הסטציונרית של השרשרת כי $\Pi P=\Pi$. ולכן, לכל התפלגות התחלתית λ_{0} של השרשרת (או של המהלך המקרי) מתקיים:

$$\lambda_{0}P^{t}\underset{t\to\infty}{\to}\lambda_{0}\begin{pmatrix}- & \Pi & -\\ & \vdots\\ - & \Pi & -\end{pmatrix}\underset{\sum\lambda_{0}(i)=1}{=}\Pi$$

משפט 34.9 בתנאים הנ"ל, לכל התפלגות התחלתית λ_{0}, ההתפלגות אחרי t צעדים $t\to\infty$, שואפת להתפלגות הסטציונרית.

נתבונן במהלך מקרי על גרף רגולרי. ז"א יהי G גרף רגולרי עם n קודקודים עם דרגה d, אז מטריצת המעבר P של ההילוך המקרי:

$$P=\begin{pmatrix}0 & & \frac{1}{d}\\ & \ddots & \\ \frac{1}{d} & & 0\end{pmatrix}$$

P היא מטריצה סימטרית ולכן ההתפלגות הסטציונרית שלה היא $\Pi=\begin{pmatrix}\frac{1}{n} & \cdots & \frac{1}{n}\end{pmatrix}$.

הערה 34.10 כדי שהתנאים של PF יתקיימו נדרוש כי:

1. הגרף G הוא קשיר.
2. ושהגרף אינו דו צדדי.

34.2 גרפים מרחיבים

הגדרה 34.11 גרף מרחיב (Expander)
יהי $G=(V,E)$ גרף על n קודקודים. נאמר כי G הוא מרחיב־δ עבור $\delta>0$ אם לכל תת קבוצה $S\subseteq V$ מתקיים:

$$\frac{\#\text{ of outgoing edges from S}}{|S|}=\frac{E\left(S,\overline{S}\right)}{|S|}\geq\delta$$

משפט 34.12 משפטים על הקשר בין ההרחבה של גרף d־רגולרי להתנהגות הערכים העצמיים שלו
יהיו $d=\lambda_1\geq\cdots\geq\lambda_n$ הע"ע של G (כש $G_{ij}=\begin{cases}1 & (i,j)\in E\\ 0 & otherwise\end{cases}$). יהי $h(G)$:

$$h\left(G\right)=\min_{S\subseteq V,\ |S|\leq\frac{|V|}{2}}\frac{E\left(S,\overline{S}\right)}{|S|}$$

אזי:

$$\frac{d-\lambda_2}{2}\leq h\left(G\right)\underbrace{\leq}_{Cheeger}\sqrt{2d\left(d-\lambda_2\right)}$$

אלה המשפטים המקשרים בין ההרחבה של הגרף לפער הספקטרלי שלו (ולכן לתכונת העירבוב המהיר שלו).

35 תרגול 9

35.1 הילוך מקרי פשוט על גרף

משפט 35.1 יהי $G=(V,E)$ גרף רגולרי (לכל הקודקודים אותה הדרגה), קשיר, **לא** דו צדדי. יהי $X_0,X_1,\cdots$ הילוך מקרי פשוט על G, כאשר X_0 מפולגת בהתפלגות כלשהי על V. אזי:

$$\forall v\in V:\ \lim_{t\to\infty}\Pr\left(X_t=v\right)=\frac{1}{|V|}$$

הבהרה: X_t הוא מ"מ שמקבל ערכים בV. בהנתן X_t, X_{t+1} מתפלג לפי:

$$\forall u,v\in V:\ \Pr\left(X_{t+1}=v|X_t=u\right)=\begin{cases}0 & \{u,v\}\notin E\\ \frac{1}{\deg(u)} & \{u,v\}\in E\end{cases}$$

35.1.1 דוגמאות

הגרף K_n

הגדרה 35.2 K_n
הקודקודים הם $[n]$, ויש את כל הצלעות האפשריות.

K_2 - דו"צ ולא מעניין.
$K_{n\geq 3}$ - הגרף לא דו צדדי (יש מעגל באורך 3). הגרף קשיר (מאוד!). נקבע $X_0=1$. מה ההתפלגות של X_1?

$$\begin{aligned}\forall j>1:\ \Pr\left(X_1=j\right)&=\frac{1}{n-1}\\ \Pr\left(X_1=1\right)&=0\\ \Pr\left(X_2=1\right)&=\frac{1}{n-1}\\ \forall j>1:\ \Pr\left(X_2=j\right)&=\left(1-\frac{1}{n-1}\right)\frac{1}{n-1}\end{aligned}$$

לכל t, לכל j:

$$\Pr\left(X_{t+1}=j\right)=\Pr\left(X_t\neq j\right)\frac{1}{n-1}=\left(1-\Pr\left(X_t=j\right)\right)\frac{1}{n-1}$$
$$\Downarrow$$
$$\forall t:\ \Pr\left(X_t=1\right)=\frac{1}{n}+\left(1-\frac{1}{n}\right)\left(\frac{-1}{n-1}\right)^t$$
$$\forall t\ \forall j>1:\ \Pr\left(X_t=j\right)=\frac{1}{n}-\frac{1}{n}\left(\frac{-1}{n-1}\right)^t$$

עתה נניח ש$\forall j:\ \Pr\left(X_0=j\right)=p_j$. נשים לב:

$$\forall t\ \forall j:\ \Pr\left(X_t=j\right)=\frac{1}{n}+\left(p_j\left(1-\frac{1}{n}\right)-\left(1-p_j\right)\right)\left(\frac{-1}{n-1}\right)^t$$

הגרף C_n

הגדרה 35.3 C_n
מעגל פשוט על n קודקודים שנסמן אותם ב: $V=\{0,1,\cdots,n-1\}$

נניח שn אי זוגי, אחרת C_n דו צדדי. אם n אי זוגי אז C_n הוא מעגל באורך אי זוגי. נניח ש$X_0 = 0$.

$$\begin{aligned}
&\Pr\left(X_1 = 0\right) = 0 \\
&\Pr\left(X_1 = 1\right) = \frac{1}{2} = \Pr\left(X_1 = n-1\right) = \frac{1}{2} \\
&\Pr\left(X_2 = 0\right) = \frac{1}{2} \\
&\Pr\left(X_2 = 1\right) = 0 \\
&\Pr\left(X_2 = 3\right) = \Pr\left(X_2 = n-2\right) = \frac{1}{4}
\end{aligned}$$

מטריצת השכנויות של C_n היא:

$$\begin{array}{cccccc}
 & 0 & 1 & 2 & \cdots & n-1 \\
0 & 0 & 1 & 0 & 0 & 1 \\
1 & 1 & 0 & 1 & \ddots & 0 \\
2 & 0 & 1 & \ddots & \ddots & 0 \\
\vdots & 0 & \ddots & \ddots & \ddots & 1 \\
n-1 & 1 & 0 & 0 & 1 & 0
\end{array}$$

$$\Downarrow$$

$$A = \frac{1}{2}\begin{matrix}
0 & 1 & 0 & 0 & 1 \\
1 & 0 & 1 & \ddots & 0 \\
0 & 1 & \ddots & \ddots & 0 \\
0 & \ddots & \ddots & \ddots & 1 \\
1 & 0 & 0 & 1 & 0
\end{matrix}$$

אם $p = (p_0, \cdots, p_{n-1})$ ההתפלגות בזמן t, אז pA ההתפלגות בזמן $t+1$.

טענה 35.4 תהי $X = (X_0, \cdots, X_{n-1})$ ההתפלגות של X_0, אזי:

$$\lim_{t\to\infty} \left\| xA^t - \begin{pmatrix} \frac{1}{n} & \cdots & \frac{1}{n} \end{pmatrix} \right\|_1 = 0$$

הוכחה: נראה:

1. לA יש ע"ע 1, עם ריבוי 1 (כלומר המימד של המ"ע המתאים הוא 1).
2. אם $x \neq 1$ ע"ע של A אז $|\lambda| < 1$.

$$u_0 = \begin{pmatrix} \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} \end{pmatrix}$$
$$\lambda_0 = 1$$
$$\Downarrow$$
$$\langle x, u_1 \rangle = \sum_i x_i \frac{1}{\sqrt{n}} = \frac{1}{\sqrt{n}} \sum_i x_i = \frac{1}{\sqrt{n}}$$

טענה 35.5 השורשים של הפולינום האופייני של A הם בדיוק (כולל ריבוי):

$$\cos\left(\frac{2\pi}{n}k\right),\ k = 0, \cdots, n-1$$

נשים לב ש1 כתוב פעם אחת, ו-1 לא כתוב, ולכן אם נסמן $\lambda_k = \cos\left(\frac{2\pi}{n}k\right)$ אז לכל $k \geq 1$: $|\lambda_k| < 1$

הוכחה: הוקטור $\begin{pmatrix} e^{\frac{2\pi k i}{n}0} \\ \vdots \\ e^{\frac{2\pi k i}{n}(n-1)} \end{pmatrix}$ הוא ו"ע של λ_k . תרגיל: מכאן λ_k הוא שורש של $p_A(x)$ כאשר $A \in M_n(\mathbb{C})$, אבל זה מספיק. ■

הע"ע הכי גדול בערך מוחלט מלבד 1 הוא $1 - \frac{4\pi^2}{2n^2} \overset{Taylor}{\approx} \cos\left(\frac{2\pi}{n}\right)$. עכשיו נרשום:

$$u_0, \cdots, u_{n-1} \text{ is an orthonormal basis of eigenvectors}$$
$$x = \sum_i \langle x, u_i \rangle u_i$$
$$xA^t = \frac{1}{\sqrt{n}} u_0 + \sum_{i \geq 1} \langle x, u_i \rangle x_i^t u_i$$
$$\left\| xA^t - \begin{pmatrix} \frac{1}{n} & \cdots & \frac{1}{n} \end{pmatrix} \right\|_2^2 = \sum_{i \geq 1} \langle x, u_i \rangle^2 x_i^{2t} \approx\leq \left(1 - \frac{2\pi}{n^2}\right)^{2t} \sum_{i \geq 1} \langle x, u_i \rangle^2 \leq \|x\|_2^2$$

■

36 תרגול אקסטרה 9

36.1 הילוך מקרי פשוט על גרף רגולרי

גרף d-רגולרי $G = (V = \{1, \cdots, n\}, E)$.

הגדרה 36.1 הילוך מקרי על G

סדרה $X_0, X_1, \cdots$ של קודקודים בG, כשX_0 הוא קודקוד התחלתי, ולכל t, X_{t+1} הוא שכן מקרי של X_t.

נסמן $p^{(t)} \in \mathbb{R}^n$:

$$\forall i:\ p_i^{(t)} = \Pr(X_t = i)$$

אבחנה:

$$A_{ij}=\begin{cases}1 & \{i,j\}\in E\\ 0 & otherwise\end{cases}$$

$$p^{(t+1)}=\left(\frac{1}{d}A\right)p^{(t)}$$

תזכורת: הע"ע של A הם: $d=\lambda_1\geq\cdots\geq\lambda_n\geq -d$, לכל הע"ע של $\frac{A}{d}$: $1=\frac{\lambda_1}{d}\geq\cdots\geq\frac{\lambda_n}{d}\geq -1$.
הוקטור העצמי הראשון (זה נכון לכל גרף d־רגולרי):

$$v_1=\frac{1}{\sqrt{n}}\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}$$

טענה 36.2 $\lambda_n=-d\Leftrightarrow G$ הוא דו"צ.

הערה 36.3 אם $\lambda_n=-d$:

$$v_n=\frac{1}{\sqrt{n}}\begin{array}{c}\begin{pmatrix}1\\ \vdots\\ 1\\ -1\\ \vdots\\ -1\end{pmatrix}\end{array}\begin{array}{c}A\\ A\\ A\\ B\\ B\\ B\end{array}$$

כש A,B הם החלוקה לצדדים של G.
נסמן ב$\Pi\in\mathbb{R}^n$ את ההתפלגות האחידה:

$$\Pi=\begin{pmatrix}\frac{1}{n}\\ \vdots\\ \frac{1}{n}\end{pmatrix}$$

הגדרה 36.4 פער ספקטרלי של G

$$\delta\left(G\right)=\max_{2\leq i\leq n}\left\{\frac{|\lambda_i|}{d}\right\}=\max\left\{\frac{|\lambda_2|}{d},\frac{|\lambda_n|}{d}\right\}$$

נראה שאם $\delta\left(G\right)<1$ אז $p^{(t)}\overset{t\to\infty}{\to}\pi$, ונראה שככל ש$\delta\left(G\right)$ יותר קטן, קצב ההתכנסות יותר מהיר.

משפט 36.5 $\left\|p^{(t)}-\Pi\right\|_2\leq\left(\delta\left(G\right)\right)^t$

הוכחה: נציג את

$$p^{(0)} = \sum_{i=1}^{n} \alpha_i v_i$$

טענה 36.6 $\alpha_1 v_1 = \Pi$

הוכחה:

$$\alpha_1 = \left\langle p^{(0)}, v_1 \right\rangle = \frac{1}{\sqrt{n}} \underbrace{\sum_{i=1}^{n} p_i^{(0)}}_{=1} = \frac{1}{\sqrt{n}}$$

$$\Downarrow$$

$$\alpha_1 v_1 = \frac{1}{n} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \Pi$$

■

טענה 36.7 $\sum_{i=2}^{n} \alpha_i^2 \leq 1$

הוכחה:

$$\sum_{i=2}^{n} \alpha_i^2 \leq \sum_{i=1}^{n} \alpha_i^2 = \left\| p^{(0)} \right\|^2 = \sum_{i=1}^{n} \left(p_i^{(0)} \right)^2 \leq \sum_{i=1}^{n} p_i^{(0)} = 1$$

■

סה"כ קיבלנו:

$$p^{(0)} = \Pi + \sum_{i=2}^{n} \alpha_i v_i$$

$$\Downarrow$$

$$p^{(t)} = \left(\frac{1}{d} A \right)^t p^{(0)} = \left(\frac{1}{d} A \right)^t \Pi + \left(\frac{1}{d} A \right)^t \sum_{i=2}^{n} \alpha_i v_i =$$

$$\overset{\Pi \text{ is an eigenvector with eigenvalue } 1}{=}$$

$$= \Pi + \sum_{i=2}^{n} \left(\frac{\lambda_i}{d} \right)^t \alpha_i v_i$$

אם $\frac{\lambda_i}{d} < 1$ אז הביטוי הימני ידעך ונשאר עם Π:

$$\left\|p^{(t)} - \Pi\right\|_2^2 = \sum_{i=2}^{n} \left(\frac{\lambda_i}{d}\right)^{2t} \alpha_i^2 \le \delta(G)^{2t} \underbrace{\sum_{i=2}^{n} \alpha_i^2}_{\le 1} \le \delta(G)^{2t}$$

$$\Downarrow$$

$$\left\|p^{(t)} - \Pi\right\|_2 \le \delta(G)^t$$

■

מסקנה 36.8 יהי $\epsilon > 0$. יהי $t \ge \frac{1}{1-\delta(G)} \log\left(\frac{\sqrt{n}}{\epsilon}\right)$. תהי $A \subseteq V$.

$$\left|\Pr(X_t \in A) - \frac{|A|}{n}\right| < \epsilon$$

הוכחה:

$$\begin{aligned}
\left|\Pr(X_t \in A) - \frac{|A|}{n}\right| &= \left|\sum_{i\in A} p_i^{(t)} - \frac{1}{n}\right| \overset{\triangle}{\le} \sum_{i\in A}\left|p_i^{(t)} - \frac{1}{n}\right| \le \\
&\le \sum_{i=1}^{n}\left|p_i^{(t)} - \frac{1}{n}\right| = \left\|p^{(t)} - \Pi\right\|_1 \le \sqrt{n}\left\|p^{(t)} - \Pi\right\|_2 \le \\
&\le \sqrt{n}\,\delta(G)^t \\
&\Downarrow t = \frac{1}{1-\delta(G)}\log\left(\frac{\sqrt{n}}{\epsilon}\right) \\
\left|\Pr(X_t \in A) - \frac{|A|}{n}\right| &\le \sqrt{n}\left(\underbrace{\delta(G)^{\frac{1}{1-\delta(G)}}}_{\le \frac{1}{e}}\right)^{\log\left(\frac{\sqrt{n}}{\epsilon}\right)} \le \sqrt{n}\frac{\epsilon}{\sqrt{n}} = \epsilon
\end{aligned}$$

נשים לב ש

$$\delta^{\frac{1}{1-\delta}} \le \frac{1}{e}$$

$$\Updownarrow$$

$$\delta \le \frac{1}{e^{1-\delta}} = e^{\delta-1}$$

$$\Updownarrow x := \delta - 1$$

$$x + 1 \le e^x$$

■

37 תרגיל 9

טענה 37.1 יהי $G=([n],E)$ גרף d רגולרי, ותהי A מטריצת השכנויות של G. יהיו $\lambda_1 \geq \cdots \geq \lambda_n$ הע"ע של A. אז:

1. A לכסינה אורתוגונלית.
2. $\lambda_1 = d$
3. $\forall i \in [n] : \ |\lambda_i| \leq d$
4. אם יש k רכיבי קשירות בG אז $\lambda_1 = \cdots = \lambda_k = d$ וגם $\lambda_{k+1} < d$.
5. אם G דו צדדית, אז $\forall i \in [n] : \ \lambda_i = -\lambda_{n+1-i}$
6. לG יש רכיב קשירות דו צדדי אס"ם $\lambda_n = -d$.

הגדרה 37.2 Total Variation Distance

יהיו $p=(p_1,\cdots,p_n), q=(q_1,\cdots,q_n)$ התפלגויות על $[n]$. אז הTotal Variation Distance בין p וq מוגדר כך:

$$\delta(p,q) = \frac{1}{2}\|p-q\|_1 = \frac{1}{2}\sum_{i=1}^{n}|p_i - q_i|$$

טענה 37.3 יהיו p,q התפלגויות על $[n]$, עבור $A \subseteq [n]$ אם:

$$P(A) = \sum_{i \in A} p_i, \ Q(A) = \sum_{i \in A} q_i$$

אז:

$$\delta(p,q) = \max_{A \subseteq [n]} |P(A) - Q(A)|$$

טענה 37.4 יהי $G=(V,E)$ גרף קשיר לא דו צדדי עם $|V| \geq 2$, אז ההתפלגות הסטציונרית של מהלך מקרי על G היא:

$$P(v) = \frac{deg(v)}{2|E|}$$

הגדרה 37.5 מהלך מקרי עצל

יהי $G=(V,E)$ כך שבG אין קודקודים מבודדים. מהלך מקרי עצל על G הוא שרשרת מרקוב $X_0, X_1, \cdots$ כך ש:

$$\forall t \in \mathbb{N} : X_t \in V$$

$$\forall t \in \mathbb{N} : \ \forall u,v \in V : \ \Pr(X_t = v \mid X_{t-1} = u) = \begin{cases} \frac{1}{2} & u=v \\ \frac{1}{2deg(u)} & \{u,v\} \in E \\ 0 & otherwise \end{cases}$$

טענה 37.6 יהי $G = ([n], E)$ גרף קשיר d־רגולרי (עם $d \geq 1$), תהי x^0 הסתברות על $[n]$, יהיו $X_0, X_1, \cdots$ הילוך מקרי עצל על G, כש$X_0 \sim x^0$. אז:

$$\forall k \in [n]: \lim_{t \to \infty} \Pr(X_t = k) = \frac{1}{n}$$

הגדרה 37.7 הקוביה הדיסקרטית
$G = (\{0,1\}^n, E)$ כש $\{u,v\} \in E$ אס"ם u, v נבדלים בקורדינטה אחת.

טענה 37.8 מהלך מקרי עצל על הקוביה הדיסקרטית כש$X_0 = (0, \cdots, 0)$ שקול ל: בכל צעד, קורדינטה נבחרת באקראי (באופן אחיד) ומוגדרת להיות 0 או 1 (בהסתברות חצי לכל אפשרות).

טענה 37.9 תהי G הקוביה הדיסקרטית. אז:

1. G דו צדדי.
2. יהי A_t המאורע בו עד זמן t כל הקורדינטות נבחרו במהלך המקרי העצל. אז לכל n $t \geq$ ולכל $x \in \{0,1\}^n$ מתקיים:

$$\Pr(X_t = x | A_t) = \frac{1}{2^n}$$

3. מתקיים: $\forall t: \Pr(A_t) \geq 1 - n\left(1 - \frac{1}{n}\right)^t$
4. נניח $X_t \sim p_t$, אז קיים $C > 0$ שלא תלוי בn כך ש:

$$t \geq Cn^2 \Rightarrow \delta\left(p_t, \left(2^{-n}, \cdots, 2^{-n}\right)\right) \leq \frac{1}{4}$$

38 בוחן 2

טענה 38.1 אם v, v' ו"ע של A סימטרית אז $\langle v, v' \rangle = 0$.

חלק III
אופטימיזציה

39 שיעור 10

39.1 אינטואיציה

הגדרה 39.1 בעיית אופטימיזציה
קבוצה (תחת הגדרה) Ω, פונקציה $f:\Omega\to\Re$, ומחפשים $x\in\Omega$ שממקסם את f:

$$\forall y\in\Omega:\ \ f\left(y\right)\leq f\left(x\right)$$

הגדרה 39.2 תכנון לינארי
סוג של בעיית אופטימיזציה שבה $\Omega\subseteq\mathbb{R}^n$ קבוצה המוגדרת ע"י משוואות ואי־שווינים לינאריים, ובה f אף היא לינארית.

הגדרה 39.3 פאון (polytope)
קבוצה חסומה ב$\mathbb{R}^n$ המוגדרת ע"י רשימה סופית של משוואות לינאריות ואי שווינים לינאריים.

הגדרה 39.4 תכנון לינארי
דרך אחרת להגדרה: רוצים למקסם פונקציה לינארית $f\left(x\right)$ כשx מוגבל לאיזשהו פאון P.

מסקנה 39.5 האופטימון מושג תמיד (לא בהכרח אך ורק ב) בקודקוד של הפאון. זה מאפשר לפחות לתת פתרון במס' צעדים סופי לבעיית הLP.

39.2 בעיות

הגדרה 39.6 בעיית התזונה של פרות
יש מטריצה $A_{m\times n}$, וקטורים $b\in\mathbb{R}^n, c\in\mathbb{R}^m$ כך ש:

$$\begin{aligned}a_{ij}&=\text{how many units of the j-th nutrient is in one unit of food type i}\\ b_i&=\text{the minimal amount of nutrient i needed so the cow will be healthy}\\ c_j&=\text{the price of one unit of food type j}\end{aligned}$$

והמטרה היא להזין במחיר מזערי את הפרות בתזונה הבריאה להן:

$$\begin{aligned}&\min\left\langle c,x\right\rangle\\ &s.t.\\ &\quad x\geq 0\\ &\quad xA\geq b\end{aligned}$$

כשהאילוץ $xA\geq b$ אומר "לכל אב מזון: תנו לפחות את המנה המינימלית הנדרשת."

הערה 39.7 האילוצים $x \geq 0,\ xA \geq b$ מגדירים פוליהדר.

הגדרה 39.8 בעיית על מישור מפריד עם שולי ביטחון
ב$\mathbb{R}^n$ נתונות נקודות $\{x_i\}_{i \in I}$ "חיוביות" ונקודות $\{y_j\}_{j \in J}$ "שליליות". מחפשים על מישור שמפריד ביניהן עם שולי ביטחון ϵ גדולים ככל האפשר. המשתנים הם α, β, ϵ, והבעיה היא:

$$\begin{aligned} &\max \epsilon \\ &s.t. \\ &\epsilon \geq 0 \\ &\forall i \in I: \ \langle \vec{\alpha}, \vec{x_i} \rangle \geq \beta + \epsilon \\ &\forall j \in J: \ \langle \vec{\alpha}, \vec{y_j} \rangle \leq \beta - \epsilon \end{aligned}$$

הגדרה 39.9 בעית הזרימה ברשת
נתון גרף מכוון $G = (V, E)$ ("רשת") עם מקור $s \in V$ ובור $t \in V$, וכן נתונים קיבולים לצלעות $c: E \to \mathbb{R}_{\geq 0}$.
זרימה בGזו פונקציה $f: E \to \mathbb{R}_{\geq 0}$ המקיימת את הדרישות הבאות:

$$\begin{aligned} &\forall e \in E: \ f(e) \in [0, c(e)] \\ &\forall v \neq s, t \in V: \ \sum_{e \to v} f(e) = \sum_{v \to e} f(e) \ \text{(preservation of material)} \end{aligned}$$

בכפוף לתנאים האלה רוצים למקסם את השטף שיוצא מs (שנכנס לt):

$$\sum_{s \to e} f(e) \left(= \sum_{e \to t} f(e) \right)$$

הגדרה 39.10 ניסוח אחר של בעיית הזרימה ברשת
נתונים שוב G, c, s, t. תהיה A המשפחה של כל המסילות המכוונות מs לt בG. זרימה בנוסח זה מתוארת ע"י רשימה של מקדמים אי שליליים $\{x_P\}_{P \in A}$ והכוונה היא שאנו מזרימים x_P לאורך המסילה P. בניסוח זה ברור ששימור החומר מתקיים, ונותר רק לדאוג לכך שלא נעבור את הקיבולים המותרים.

$$\forall e \in E, P \in A: \ M_{e,P} = \begin{cases} 1 & e \in P, \ s \to^P t \\ 0 & otherwise \end{cases}$$

בעיית הזרימה: רוצים $x \geq 0$ שקבוצת האינדקסים שלו היא A, ובנוסף התנאי $Mx \leq c$. נשים לב שאילוצים אלו הם הפאון שלנו. פונקציית המטרה היא: $\max \sum_{P \in A} x_P$.

39.3 אלגוריתמים לפתרון LP

1. אלגוריתמים מטיפוס סימפלקס. הרעיון: בכל צעד עוברים מקודקוד של הפאון לקודקוד שכן באופן שמשפר את פונקציית המטרה.
2. אלגוריתמים של נקודה פנימית.
3. אלגוריתמים של נקודה קיצונית (לא נדבר כלל).

הערה 39.11 $Klee - Minty$ מצאו ב1972 פאון שמכשיל גירסה מסויימת של אלגוריתם הסימפלקס. כמעט כל גרסה של אלגוריתם הסימפלקס נכשלה ברבות השנים בדרך דומה.

שאלה קרובה: מהו הקוטר המירבי של הגרף של פאון d-מימדי, עם n דפנות.

39.3.1 הסברים לכך שאלגוריתמים מטיפוס סימפלקס נוטים לעבוד היטב בבעיות מציאותיות

1. מראים שאם מגרילים את התוכנית הלינארית אז בהסתברות קרובה לאחד גרסה מסויימת של הסימפלקס עובדת מהר.
2. Smoothed Analysis (Spielman Tang)

39.3.2 אלגוריתם האליפסואידים (L.Khachyan)

זהו אלגוריתם פולינומי בגודל הקלט לבעיית LP. לא ידוע האם יש אלגוריתם פוליומי חזק לבעיית LP.

הגדרה 39.12 אלגוריתם פולינומי חזק
נניח שאיברי $A_{m\times n}, b, c$ נתונים בk ביטים של דיוק. אלגוריתם פולינומי חזק הוא אלגוריתם שזמן הריצה שלו $poly\left(m,n\right)$ כשכל הפעולות מתבצעות בk ביטים של דיוק.

כפי שציינו, בעיית הספיקות קשה לפחות כמו LP. ז"א, נניח יש לנו תוכנית שמכריעה אם מערכת נתונה של משוואות ואי שווינים לינאריים היא ספיקה, אז אפשר לקרוא לה מס' קטן של פעמים ע"מ לפתור LP. נניח הבעיה:

$$\begin{aligned} \max\left\langle c,x\right\rangle \\ x\geq 0 \\ Ax\leq b \end{aligned}$$

אז נקבע T ונשאל האם

$$\begin{aligned} \left\langle c,x\right\rangle \geq T \\ x\geq 0 \\ Ax\leq b \end{aligned}$$

ספיקה.

הגדרה 39.13 בעיית הספיקות
נתון גוף קמור $P\subseteq\mathbb{R}^n$. ידוע:

1. $P\subseteq B\left(0,R\right)$
2. אם $P\neq\emptyset$ אז הוא מכיל כדור מרדיוס $\leq r$.

הבעיה: להכריע אם $P=\emptyset$.

הגדרה 39.14 אליפסואיד
התמונה של כדור היחידה כשמפעילים טרנספורמציה אפינית (= לינארית + הזהה).

40 תרגול 10

40.1 פוליהדרונים קמורים ואופטימיזציה לינארית

מעוניינים לפתור

מקסם את $c^T x$ כאשר $Ax \leq b,\ x \in \mathbb{R}^n$. קבוצת הנקודות $\{x:\ Ax \leq b\}$ נקראת קבוצת הנקודות הפיזביליות. היא פוליהדר.

40.1.1 סימונים

הגדרה 40.1 חצי מרחב
עבור $y \in \mathbb{R}^n,\ a \in \mathbb{R}$ נגדיר את חצי המרחב:

$$\Lambda(y,a) := \{x \in \mathbb{R}^n:\ \langle y,x\rangle \leq a\}$$

הגדרה 40.2 על מישור

$$H(y,a) = \{x \in \mathbb{R}^n:\ \langle y,x\rangle = a\} = \Lambda(y,a) \cap \Lambda(-y,-a)$$

הגדרה 40.3 פוליהדר קמור (convex polyhedron)
חיתוך של מס' סופי של חצאי מרחבים. אם הפוליהדר **חסום**, אז הוא נקרא פאון קמור (convex polytope).

40.1.2 דוגמאות

הגדרה 40.4 סימפלקס n־מימדי:

$$\begin{aligned}
\triangle_n &:= \left\{x \in \mathbb{R}^{n+1}:\ x = \sum_{i=1}^{n+1} \lambda_i e_i,\ \sum_{i=1}^{n+1} \lambda_i = 1,\ \forall i:\ \lambda_i \geq 0\right\} = \\
&= conv\{e_1, \cdots, e_{n+1}\} = \\
&= \left\{x \in \mathbb{R}^{n+1}:\ \forall i:\ x_i \geq 0,\ \sum_{i=1}^{n+1} x_i = 1\right\} = \\
&= \cap_{i=1}^{n+1} \Lambda(-e_i, 0) \cap \underbrace{H\left(\begin{pmatrix}1 & \cdots & 1\end{pmatrix}, 1\right)}_{\text{same as } \sum_{i=1}^{n+1} x_i = 1}
\end{aligned}$$

נסתכל על דוגמאות:

$$\begin{aligned}
\triangle_0 &= \{(1)\} \\
\triangle_1 &= diagram \\
\triangle_2 &= diagram
\end{aligned}$$

הגדרה 40.5 פאה

יהי P פוליהדר. פאה של P היא קבוצה מהצורה $P\cap H\left(y,a\right)$ כך ש: $P\subseteq\Lambda\left(y,a\right)$. קודקוד של P זה פאה מגודל 1.

פאות של $\triangle_{1}$:

$$\emptyset,\left\{ \begin{pmatrix}1\\0\end{pmatrix}\right\} ,\left\{ \begin{pmatrix}0\\1\end{pmatrix}\right\} ,\triangle_{1}$$

הגדרה 40.6 קובייה n־מימדית

$$\begin{aligned}C_{n} & :=\left[-1,1\right]^{n}\subseteq\mathbb{R}^{n}=\\
 & =\left\{ x\in\mathbb{R}^{n}:\ \forall i:\ x_{i}\in\left[-1,1\right]\right\} =\\
 & =\cap_{i=1}^{n}\Lambda\left(e_{i},1\right)\cap\Lambda\left(-e_{i},1\right)=\\
 & =conv\left\{ \sum_{i=1}^{n}\epsilon_{i}e_{i}:\ \epsilon_{1},\cdots,\epsilon_{n}\in\left\{ \pm1\right\} \right\}
\end{aligned}$$

הוכחה: נראה שאכן $\epsilon_{1}e_{1}+\cdots+\epsilon_{n}e_{n}$ הוא קודקוד. נקח:

$$H\left(\epsilon_{1}e_{1}+\cdots+\epsilon_{n}e_{n},n\right)$$

נשים לב:

$$\left\langle \epsilon_{1}e_{1}+\cdots+\epsilon_{n}e_{n},\epsilon_{1}e_{1}+\cdots+\epsilon_{n}e_{n}\right\rangle =n$$

אם $x\in C_{n}$ אז:

$$\left\langle x,\epsilon_{1}e_{1}+\cdots+\epsilon_{n}e_{n}\right\rangle =\sum_{i=1}^{n}\epsilon_{i}x_{i}\leq^{*}n$$

$\leq^{*}$ הוא שוויון אם"ם $\sum\epsilon_{i}e_{i}=x$.

$\blacksquare$

הגדרה 40.7 אורתנט n־מימדי

$$\begin{aligned}O_{n} & =\left\{ x\in\mathbb{R}^{n}:\ \forall i:\ x_{i}\geq0\right\} =\\
 & =\cap_{i=1}^{n}\Lambda\left(-e_{i},0\right)
\end{aligned}$$

40.1.3 פאונים שנוצרים מתוכנית לינארית

נגדיר:

$$A = \{(0,1), (1,2)\}$$
$$B = \{(-1,-1), (3,1)\}$$

נרצה ישר שעובר בין A לB כך שהמרחקים על ציר y בין הנקודות לבין הישר יהיה מקסימלי. המשתנים הם a, b שייצגו את השיר:

$$y = ax + b$$

אילוצים:

$$b \leq 1$$
$$a + b \leq 2$$
$$-a + b \geq -1$$
$$3a + b \geq 1$$

כשאי השוויון הראשון מתאים ל$(0,1)$, אי השוויון השני מתאים ל$(1,2)$. שניהם אמורים להיות מעל לישר. אי השוויונות האחרונים מתאימים לנקודות בB שמתחת לישר.

ציור שמראה את קבוצת הנקודות הפיסביליות.

פונקציית מטרה: למקסם:

$$1 - b + 2 - (a + b) + -a + b + 1 + 3a + b - 1 = 3a + 3$$

נשים לב שזה שקול למיקסום $3a$, ששקול למקסם את a.

41 תרגול אקסטרה 10

41.1 תכנון לינארי ואלגוריתמי קירוב

41.1.1 בעיית SET-COVER

קלט

קבוצה של איברים U, כשn$= |U|$
אוסף של תתי קבוצות: $S_1, \cdots, S_m \subseteq U$

מטרה

למצוא תת אוסף מגודל מינימלי שמכסה את כל איברי U.

דוגמה

$$
\begin{aligned}
U &= \{1,2,3\} \\
S_1 &= \{1,2\} \\
S_2 &= \{1,3\} \\
S_3 &= \{2,3\}
\end{aligned}
$$

הפתרון האופטימלי הוא $OPT = 2$ (נדרשים במינימום 2 סטים כדי לכסות את U).

ננסח את SET-COVER כבעיית תכנון בשלמים

נגדיר

$$
\begin{aligned}
b_1, \cdots, b_m &\in \{0,1\} \\
b_i = 1 &\Leftrightarrow S_i \text{ is a part of the collection}
\end{aligned}
$$

אז הבעיה היא:

$$
\begin{aligned}
\min \sum_{i=1}^{m} b_i \\
b_1, \cdots, b_m \in \{0,1\} \\
\forall u \in U : \sum_{i:\ u \in S_i} b_i \geq 1
\end{aligned}
$$

רלקסציית LP

נחליף את הb_iים בx_iים:

$$
\begin{aligned}
\min \sum_{i=1}^{m} x_i \\
x_1, \cdots, x_m \in [0,1] \\
\forall u \in U : \sum_{i:\ u \in S_i} x_i \geq 1
\end{aligned}
$$

ואפשר לחשוב על x_i ככמה אני באמת רוצה את S_i באוסף.

הדוגמא הקודמת לאחר רלקסציית LP

$$\begin{aligned} \min x_1 + x_2 + x_3 \\ 0 \le x_1, x_2, x_3 \le 1 \\ x_1 + x_2 \ge 1 \\ x_1 + x_3 \ge 1 \\ x_2 + x_3 \ge 1 \end{aligned}$$

אז הפתרון האופטמלי מתקבל כש$x_1 = x_2 = x_3 = \frac{1}{2}$ והוא שווה ל

$$OPT_{LP} = \frac{3}{2}$$

הערה 41.1 נשים לב ש

$$OPT_{LP} \le OPT$$

כי התחומים המותרים מוכלים. אבל יש אתגר - כיצד ניתן לקחת את הפתרון האופטימלי של הLP (אותו ניתן לחשב בזמן פולינומיאלי) ולחלץ ממנו לבעיה המקורית? (עיגול או $rounding$ של הפתרון של הLP).

אלגוריתמי עיגול

אלגוריתם 1 אלגוריתם ראשון

קלט: פתרון אופטימלי $x_1, \cdots, x_m$ לLP.
פלט: פתרון c לבעיה המקורית כך ש $|c| \le f \cdot OPT_{LP}$ כאשר f היא השכיחות המקסימלית (כל איבר $u \in U$ מופיע בלכל היותר בf קבוצות מבין $S_1, \cdots, S_m$.

1. נחזיר $c \leftarrow \left\{ i : \ x_i \ge \frac{1}{f} \right\}$

טענה 41.2 שתי טענות:

1. $c \le f \cdot OPT_{LP}$
2. c הוא כיסוי של U.

הוכחה: כל טענה בנפרד:

1. נסמן $b_i = \begin{cases} 1 & x_i \ge \frac{1}{f} \\ 0 & otherwise \end{cases}$ ואז $b_i \le f \cdot x_i$. נקבל:

$$|c| = \sum_{i=1}^{m} b_i \le f \cdot \sum_{i=1}^{m} x_i = f \cdot OPT_{LP}$$

2. יהי $u\in U$. צ"ל: יש i כך ש$u\in S_i$ וגם $x_i\geq\frac{1}{f}$. ואכן, $\sum_{i:\ u\in S_i}x_i\geq1$ ויש לכל היותר f מחוברים בסכום זה, ולכן לפחות אחד מהם הוא לפחות $\frac{1}{f}$.

■

אלגוריתם 2 אלגוריתם שני

קלט: פתרון אופטימלי $x_1,\cdots,x_m$ לLP.

פלט: בהסתברות לפחות $\frac{1}{2}$ - פתרון c לבעיה המקורית כך ש $|C|\leq O\left(\log n\right)\cdot OPT_{LP}\leq O\left(\log n\right)\cdot OPT$ כאשר $n=|U|$.

1. נקבע $k\leftarrow\ln\left(4n\right)$
2. עבור $j\leftarrow1,\cdots,k$:
   (א) נבחר $C_j\subseteq[m]$ באקראי באופן ב"ת בבחירות הקודמות כך ש $\Pr\left(i\in C_j\right)=x_i$
3. נחזיר $C\leftarrow C_1\cup\cdots\cup C_k$

טענה 41.3 לכל $u\in U$ ולכל $j\in[k]$: $\Pr\left(u\text{ is not covered by }C_j\right)\leq\frac{1}{e}$

הוכחה: איך בוחרים את C_1 ? לכל אחת מתתי הקבוצות האפשריות $S_1,\cdots,S_m$ נכניס את S_i לC_1 בהסתברות x_i, אז:

$$\Pr\left(u\text{ is not covered by }C_1\right)=\Pi_{i:\ u\in S_i}\left(1-x_i\right)\leq$$

נסמן בf_u את מס' הקבוצות מבין $S_1,\cdots,s_m$ שu שייך אליהן, אז:

$$\overset{\text{means inequality}}{\leq}\left(\frac{1}{f_u}\sum_{i:\ u\in S_i}\left(1-x_i\right)\right)^{f_u}\leq\left(1-\frac{\sum_{i:\ u\in S_i}x_i}{f_u}\right)^{f_u}\leq\left(1-\frac{1}{f_u}\right)^{f_u}\leq\frac{1}{e}$$

אז באופן זהה

$$\Pr\left(u\text{ is not covered by }C_j\right)\leq\frac{1}{e}$$

■

מסקנה 41.4 נשים לב:

$$\Pr\left(u\text{ is not covered by any of the }C_j\right)\leq\left(\frac{1}{e}\right)^k=\left(\frac{1}{e}\right)^{\ln(4n)}=\frac{1}{4n}$$
$$\Downarrow\text{Union bound}$$
$$\Pr\left(C\text{ is not a cover}\right)\leq\frac{1}{4}$$

טענה 41.5

$$\mathbb{E}\left[|C|\right] \leq k\mathbb{E}\left[|C_1|\right] \leq k\sum_{i=1}^{m} x_i = kOPT_{LP}$$

ומא"ש מרקוב:

$$\Pr\left(|C| \geq 4kOPT\right) \leq \frac{1}{4}$$

ואז נקבל את החסם על האופטימליות של האלגוריתם.

42 תרגיל 10

הגדרה 42.1 $Cross - Polytope$

$$P_n = \left\{x \in \mathbb{R}^n : \ \|x\|_1 \leq 1\right\}$$

טענה 42.2 עבור P_n מתקיים:

1. P_n הוא חיתוך של מס' סופי של חצאי מרחבים ולכן הוא פוליהדרון. כיוון שהוא חסום הוא פוליטופ.
2. הקודקודים של P_n הם $V = \left\{\pm e_i\right\}_{i=1}^{n}$.
3. מתקיים: $P_n = conv\left(V\right)$

43 שיעור 11

43.1 LP

הגדרה 43.1 בעיית LP
למצוא מקסימום של פונקציה לינארית ע"פ תחום ב$\mathbb{R}^n$ המוגדר ע"י משוואות ואי־שוויונים לינארים.

הגדרה 43.2 על מישור ב$\mathbb{R}^n$

$$H = \left\{x \in \mathbb{R}^n \middle| \left\langle c, x\right\rangle = \alpha\right\}, \ c \in \mathbb{R}^n, \alpha \in \mathbb{R}$$

הערה 43.3 על מישורים מקבילים שומרים על c ומשנים את $\alpha \to \beta$

הגדרה 43.4 חצאי מרחב
לכל על מישור מוגדרים שני חצאי מרחב:

$$H^{+}=\left\{x|\left\langle c,x\right\rangle \underbrace{\geq}_{\substack{\geq\text{ closed half space}\\ >\text{ open half space}}}\alpha\right\}$$

$$H^{-}=\left\{x|\left\langle c,x\right\rangle \underbrace{\leq}_{\substack{\leq\text{ closed half space}\\ <\text{ open half space}}}\alpha\right\}$$

הגדרה 43.5 פוליהדר
חיתוך סופי של חצאי מרחב ב$\mathbb{R}^n$.

הגדרה 43.6 פאון
פוליהדר חסום.

הערה 43.7 LP זו הבעיה של מציאת המקסימום של פונקציה לינארית ע"פ פאון.

הערה 43.8 יש אלגוריתם לLP בעל זמן ריצה פולינומי באורך הקלט - אלגוריתם האליפסואידים. לא מוכר אלגוריתם פולינומי חזק.

הגדרה 43.9 אלגוריתם פולינומי חזק עבור LP
אם בעיית הLP מוגדרת על ידי $A_{m\times n}$ והקלט נתון בk ספרות דיוק, אז האלגוריתם רץ בזמן $poly\left(m,n\right)$ ומחשב בדיוק של k ביטים.

הגדרה 43.10 תוכנית לינארית
תוכנית לינארית נראית כך:

$$\begin{aligned}\max\left\langle c,x\right\rangle \\ Ax\leq b\\ x\geq0\end{aligned}$$

הגדרה 43.11 תוכנית לינארית בשלמים ILP
כמו קודם, רק

$$x\in\left\{0,1\right\}^{n}$$

הערה 43.12 תכנון לינארי בשלמים זו בעיה NP-שלמה. אבל, אפשר לנסות את הגישה הבאה: את התנאי $x\in\left\{0,1\right\}^{n}$ נחליף ב$\vec{0}\leq x\leq\vec{1}$. נקבל כרגיל וקטור x שכל הקורדינטות שלו הן בין 0 ל1. עכשיו אפשר לעגל את x בכל מיני שיטות, למשל עיגול מקרי. לא צפוי שיתקבל פתרון אופטימלי, אבל יש בעיות חשובות רבות שבהן עיגול כנ"ל מוביל לפתרון מקורב טוב.

הגדרה 43.13 בעיית הSet cover

יש קבוצת בסיס X ומשפחה f של תתי קבוצות של X. הבעיה הדיסקרטית: מצא תת משפחה קטנה ככל האפשר של f שאיחודה הוא כל X. נניח $|X|=m$ ו$|f|=n$ אז נגדיר:

$$\forall i\in[m],\forall j\in[n]:\ A_{i,j}=\begin{cases}1 & X_i\in f_j\\ 0 & otherwise\end{cases}$$

בעיית הset cover מחפשת וקטור $x\in\{0,1\}^n$ תחת האילוצים

$$\begin{aligned}&\min\left\langle\vec{1},x\right\rangle\\ &Ax\geq\vec{1}\end{aligned}$$

הגדרה 43.14 שיכוך רציונלי של set cover

$$\begin{aligned}&\min\langle 1,x\rangle\\ &0\leq x\leq 1\\ &Ax\geq 1\end{aligned}$$

הגדרה 43.15 צורה 1 של LP

$$\begin{aligned}&\max\langle c,x\rangle\\ &Ax\leq b\end{aligned}$$

הגדרה 43.16 צורה 2 של LP

$$\begin{aligned}&\max\langle c,x\rangle\\ &Ax=b\\ &x\geq 0\end{aligned}$$

משפט 43.17 כל LP אפשר להביא לכל אחת משתי הצורות הנ"ל תוך הגדלת גודל הבעיה בגורם קבוע.

הוכחה: כדי להביא את צורה 2 לצורה 1 כל מה שנדרש הוא להפטר מהמשוואות. נחליף כל שוויון בזוג אישוויונים משלימים: $\leq,\geq$.

כדי להביא את צורה 1 לצורה 2 צריך להפטר מאי שווינים. אי שוויון כללי נראה כך: $\langle a,x\rangle\leq\delta$. ע"מ להפטר ממנו נוסיף משתנה עזר $z\leq 0$ ו:

$$\langle a,x\rangle+z=\delta$$

אם יש משתנה x_i שלא מוגבל בסימנו אז:

$$x_i = u - v$$
$$0 \le u, v$$

■

הגדרה 43.18 פתרון בסיסי מותר basic feasible solution (BFS)
זה וקטור x שמקיים: הוא פתרון ($Ax = b$), הוא מותר ($x \ge 0$), והוא בסיסי: כלומר העמודות המאונדקסות בA ע"י $supp\,(x)$ הן בת"ל. ניסוח שקול ומקובל לבסיסיות: יש בסיס עמודות B של A כך ש$supp\,(x) \subseteq B$.

הערה 43.19 זה שקול לקודקוד של הפאון.

הגדרה 43.20 תומך של וקטור support
יהי x וקטור, התומך שלו הוא קבוצת כל האינדקסים כך ש$x_i \neq 0$. מסומן ב$supp\,(x)$.

משפט 43.21 נביט בLP:

$$\max \langle c, x\rangle$$
$$Ax = b$$
$$x \ge 0$$

כש$A_{m\times n}$ ו$m < n$ $\operatorname{rank}(A) =$. נניח שהמערכת ספיקה (ז"א יש $x \ge 0$ המקיים $Ax = b$), וכן שפונקציית המטרה חסומה. אזי המקסימום של $\langle c, x\rangle$ יתקבל (לאו דווקא כן) בפתרון בסיסי מותר.

הוכחה: רוצים להראות שלכל $x \ge 0$ המקיים $Ax = b$ יש z שהוא bfs כך ש$\langle c, x\rangle \ge \langle c, z\rangle$. אם x עצמו הוא bfs אז נקח $z = x$. אחרת, x אינו בסיסי. ז"א, קבוצת העמודות של A המאונדקסות ע"י $supp\,(x)$ הן כן תלויות לינארית. כלומר, יש וקטור y המקיים $Ay = 0$ וגם $supp\,(y) \subseteq supp\,(x)$. ז"א y מערב רק עמודות המצויינות ע"י התומך של x. נביט בווקטורים מהצורה $x + \epsilon y$:

$$A\,(x + \epsilon y) = \underbrace{Ax}_{=b} + \underbrace{\epsilon Ay}_{=0} = b$$

האם מתקיים $x + \epsilon y \ge 0$? כן, לפחות אם $|\epsilon|$ קטן דיו. רוצים להבין

$$\langle c, x + \epsilon y\rangle = \langle c, x\rangle + \epsilon\,\langle c, y\rangle$$

אם $\langle c, y\rangle \neq 0$ אז ע"י בחירה מתאימה של הסימן של ϵ ניתן להשיג

$$\langle c, x + \epsilon y\rangle > \langle c, x\rangle$$

נגדיל את $|\epsilon|$ עד שאיזושהי קורדינטה תתאפס, ז"א

$$supp\,(x + \epsilon y) \subsetneq supp\,(x)$$

ונמשיך כך עד שנצמצם את התומך של הוקטור שבידינו לקבוצות עמודות בת"ל.

אם $\langle c,y\rangle=0$ אז פונקציית המטרה אינה משתנה, אבל אנחנו מצליחים כנ"ל לצמצם את התומך עד שהוא נהיה בת"ל. ■

43.2 גיאומטריה

הערה 43.22 נזכור: פאון P ב$\mathbb{R}^n$ זו קבוצה חסומה המוגדרת כחיתוך של מס' סופי של חצאי מרחב.

הגדרה 43.23 על מישור תומך (supporting hyperplane)

על מישור $H\subseteq\mathbb{R}^n$ נקרא על מישור תומך של P אם:

$$P\cap H\neq\emptyset$$
$$P\subseteq H^+$$

ל$H\cap P$ קוראים פאה (face) של P, והמימד שלה הוא המימד האפיני של הקבוצה הזו. (לפאה 0 מימדית נקרא קודקוד, לפאה 1 מימדית נקרא צלע, וכו)

משפט 43.24 יהיה P הפאון המוגדר ע"י $Ax=b,\ x\geq0$, ותיהיה $v\in P$. אז התנאים הבאים שקולים:

1. v הוא קודקוד של P
2. v הוא bfs כפתרון מותר של $Ax=b,\ x\geq0$

הוכחה: $2\Leftarrow1$:

ע"פ ההגדרה הגיאומטרית של קודקוד, יש על מישור $H=\{x|\,\langle c,x\rangle=\alpha\}$ כך ש$H\cap P=\{v\}$ ו$P\subseteq H^+$ ז"א $\langle c,v\rangle=\alpha$ ולכל $z\in P\setminus\{v\}$ מתקיים $\langle c,z\rangle<\alpha$. $v\in P$ אומר שv הוא פתרון מותר. חסר: בסיסי? נזכיר את המשפט הקודם:

משפט 43.25 "האופטימום של LP מתקבל תמיד בBFS."

נביט בתוכנית הלינארית:

$$\max\langle c,x\rangle$$
$$x\geq0$$
$$Ax=b$$

לפי ההנחות המקסימום הזה $=\alpha$, הוא מתקבל בv ורק בv.

$1\Leftarrow2$:

נתון שv הוא bfs, רוצים למצוא על מישור שתומך בP בv ובו בלבד. ננסה למצוא c כזה ש

$$\langle c,v\rangle=0$$
$$\forall x\in P\setminus\{v\}:\ \langle c,x\rangle<0$$

בלי מאמץ ברור ש$\langle c,x\rangle\leq0$ לכל $x\in P$ כי $c\leq0\leq x$. מטרתנו: להראות שלכל $x\in P\setminus\{v\}$ בעצם $\langle c,x\rangle<0$, ז"א ש$supp\,(x)$ מכיל גם קורדינטות מחוץ ל$supp\,(v)$. נניח בשלילה: $supp\,(x)\subseteq supp\,(v)$. מכך ש$x\in P\setminus\{v\}$ מתקיים ש$Ax=b$, ולכן $Ax=b=Av$. אבל, מכך שv הוא bfs אז העמודות המאונדקסות על ידי $supp\,(v)$ הן בת"ל. מכך, ומכך ש$v\neq x$, קיבלנו סתירה לייחודיות ההצגה באמצעות בסיס פורש. ■

44 תרגול 11

44.1 שיטת הסימפלקס

44.1.1 אינטואיציה

מיקסום $c^T x$ תחת האילוצים:

$$\begin{aligned} Ax &\le b \\ x &\ge 0 \end{aligned}$$

שמגדירים פאון.

44.1.2 תרגיל

מקסמו את $5x_1 + 4x_2 + 3x_3$ תחת האילוצים:

$$\begin{aligned} 2x_1 + 3x_2 + x_3 &\le 5 \\ 4x_1 + x_2 + 2x_3 &\le 11 \\ 3x_1 + 4x_2 + 2x_3 &\le 8 \\ x_1, x_2, x_3 &\ge 0 \end{aligned}$$

יותר נוח לעבוד עם שוויונות, לכן נוסיף משתנים עודפים:

$$\begin{aligned} 0 &\le 5 - 2x_1 + 3x_2 + x_3 \\ 0 &\le 11 - 4x_1 + x_2 + 2x_3 \\ 0 &\le 8 - 3x_1 + 4x_2 + 2x_3 \\ x_1, x_2, x_3 \ge 0 \\ &\Downarrow \\ s_1 &= 5 - 2x_1 - 3x_2 - x_3 \\ s_2 &= 11 - 4x_1 - x_2 - 2x_3 \\ s_3 &= 8 - 3x_1 - 4x_2 - 2x_3 \\ 0 &\le x_1, x_2, x_3, s_1, s_2, s_3 \end{aligned}$$

רוצים למקסם:

$$z = 5x_1 + 4x_2 + 3x_3$$

הפתרון הפיזבילי הראשון הוא:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ s_1 \\ s_2 \\ s_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 5 \\ 11 \\ 8 \end{pmatrix}, z = 0$$

ננסה לשפר את הפתרון הראשון. נשאיר את x_2, x_3 כמו שהם וננסה לשפר את x_1:

$$0 \le 5 - 2x_1 \Rightarrow x_1 \le \frac{5}{2}$$
$$0 \le 11 - 4x_1 \Rightarrow x_1 \le \frac{11}{4}$$
$$0 \le 8 - 3x_1 \Rightarrow x_1 \le \frac{8}{3}$$

נקבל מהמשוואה הראשונה:

$$\begin{aligned}
&2x_1 = 5 - 3x_2 - x_3 - s_1 \Rightarrow x_1 = \frac{5}{2} - \frac{3}{2}x_2 - \frac{1}{2}x_3 - \frac{1}{2}s_1\\
&\quad\Downarrow\\
&s_2 = 1 + 5x_2 + 2s_1\\
&s_3 = \frac{1}{2} + \frac{1}{2}x_2 - \frac{1}{2}x_3 + \frac{3}{2}s_1
\end{aligned}$$

מקסמו:

$$z = \frac{25}{2} - \frac{7}{2}x_2 + \frac{1}{2}x_3 - \frac{5}{2}s_1$$

קיבלנו את הפתרון:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ s_1 \\ s_2 \\ s_3 \end{pmatrix} = \begin{pmatrix} \frac{5}{2} \\ 0 \\ 0 \\ 0 \\ 1 \\ \frac{1}{2} \end{pmatrix}, z = \frac{25}{2}$$

נגדיל את x_3, נקבל את האילוצים:

$$\begin{aligned}
\text{first equation} &\Rightarrow x_3 \le 5\\
\text{first equation} &\Rightarrow nothing\\
\text{third equation} &\Rightarrow x_3 \le 1 \Leftarrow tight
\end{aligned}$$

נעביר את x_3 אגף במשוואה השלישית:

$$\begin{aligned}
x_3 &= 1 + x_2 + 3s_1 - 2s_3\\
x_1 &= 2 - 2x_2 - 2s_1 + s_3\\
s_2 &= 1 + 5x_2 + 2s_1\\
z &= 13 - 3x_2 - s_1 - s_3
\end{aligned}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ s_1 \\ s_2 \\ s_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, z = 13$$

נשים לב שזהו הפתרון האופטימלי - x_2, s_1, s_3 כולם אי שליליים, אך מופיעים בz עם $-$, ולכן לא ניתן לשפר את z.

44.1.3 איך מוצאים פתרון פיזבילי בסיסי ראשון למערכת $c^T x,\ \ Ax \leq b, x \geq 0$?

מערכת המשוואות שלנו היא:

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &\leq b_1 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &\leq b_m \end{aligned}$$

נוסיף לכל המשוואות:

$$\begin{aligned} s + a_{11}x_1 + \cdots + a_{1n}x_n &\leq b_1 \\ &\vdots \\ s + a_{m1}x_1 + \cdots + a_{mn}x_n &\leq b_m \end{aligned}$$

ונרצה למקסם את s.

45 תרגול אקסטרה 11

45.1 דואליות בתכנון לינארי

45.1.1 דוגמא

$$\begin{aligned} \max & x_1 + x_2 + x_3 \\ 0 \leq & x_1, x_2, x_3 \\ 1 \geq & x_1 + x_2 \\ 1 \geq & x_1 + x_3 \\ 1 \geq & x_2 + x_3 \end{aligned}$$

נשים לב שפתרון אפשרי הוא:

$$x_1 = x_2 = x_3 = \frac{1}{2}$$

לכן: $OPT \geq \frac{3}{2}$. כיצד ניתן להוכיח\להשתכנע בקלות ש: $OPT \leq \frac{3}{2}$? נחבר את כל אי השוויונים ונחלק ב2, ואז לכל פתרון פיזיבילי x מתקיים:

$$x_1 + x_2 + x_3 \leq \frac{1}{2}(x_1 + x_2) + \frac{1}{2}(x_1 + x_3) + \frac{1}{2}(x_3 + x_2) \leq \frac{3}{2}$$

45.1.2 דוגמא

$$\begin{aligned} &\max 2x_1 + 3x_2 \\ 0 &\leq x_1, x_2 \\ 3 &\geq_1 2x_1 + x_2 \\ 4 &\geq_2 3x_1 + 2x_2 \\ 12 &\geq_3 4x_1 + 8x_2 \end{aligned}$$

ננסה למצוא חסמים עליונים לOPT:
נחלק את אי שוויון 3:

$$2x_1 + 3x_2 \leq 2x_1 + 4x_2 \leq 6$$

נקח את אי שוויון 1 ועוד אי שוויון 3 ונחלק ב3:

$$2x_1 + 3x_2 \leq 5$$

שאלה: מה החסם הטוב ביותר ששיטה כזאת יכולה לתת?

תשובה: ניקח $y_1, y_2, y_3 \geq 0$ ונבצע:

$$2x_1 + 3x_2 \underset{\text{if y satisfies constaints}}{\overset{\text{we want}}{\leq}} y_1(2x_1 + x_2) + y_2(3x_1 + 2x_2) + y_3(4x_1 + 8x_2) \leq 3y_1 + 4y_2 + 12y_3$$

כיוון שאנחנו כופלים את y בx, ע"מ לשמר את הסימן נרצה $y \geq 0$. נחפש

$$\begin{aligned} &\min 3y_1 + 4y_2 + 12y_3 \\ &s.t. \\ &y \geq 0 \\ 2y_1 + 3y_2 + 4y_3 &\geq 2 \text{ (we're using the coefficients of } x_1\text{)} \\ y_1 + 2y_2 + 8y_3 &\geq 3 \text{ (we're using the coefficients of } x_2\text{)} \end{aligned}$$

תוכנית זו נקראת התוכנית הדואלית לתוכנית המקורית.

טענה 45.1 לכל y_1, y_2, y_3 שהוא פתרון פיזיבילי לתוכנית הדואלית, ולכל פתרון פיזיבילי x_1, x_2 לתוכנית המקורית, מתקיים:

$$2x_1 + 3x_2 \leq 3y_1 + 4y_2 + 12y_3$$

משפט 45.2 הדואליות
למעשה, הפתרונות האופטימיליים מקיימים שוויון:

$$2x_1 + 3x_2 = 3y_1 + 4y_2 + 12y_3$$

באופן כללי, נניח $c \in \mathbb{R}^n, b \in \mathbb{R}^m, A \in M_{m\times n}$. נניח התוכנית הפרימלית היא:

$$\begin{aligned} &\max c^T x,\ \ x \in \mathbb{R}^n \\ &Ax \leq b \\ &x \geq b \end{aligned}$$

התוכנית הדואלית היא:

$$\begin{aligned} &\min y^T b,\ \ y \in \mathbb{R}^m \\ &A^T y \geq c \\ &y \geq 0 \end{aligned}$$

טענה 45.3 לכל x, y פיזיביליים מתקיים ש: $c^T x \leq y^T b$ ולכן אם מצאנו x, y כך ש$c^T x = y^T b$, אזי זהו הפתרון האופטימלי.

הוכחה: נסתכל על:

$$\begin{aligned} y^T Ax &= \sum_{i=1}^{n} y^T A_i x_i \geq \sum_{i=1}^{n} c_i x_i = c^T x \\ y^T Ax &= \sum_{j=1}^{m} y_j \left(Ax\right)_j \leq \sum_{j=1}^{m} y_j b_j = y^T b \\ &\Downarrow \\ c^T x &\leq y^T b \end{aligned}$$

■

משפט 45.4 אם x, y הם הפתרונות האופטימליים מתקיים $c^T x = y^T b$.

45.1.3 $Min-Cut-Max-Flow$

הגדרה 45.5 $Max-Flow$

יש t s, נסמן בP את אוסף המסלולים מs לt.

$$
\begin{aligned}
&max-flow = \\
&\quad \max \sum_{p\in P} x_p \\
&\quad \forall e: \sum_{p:\ e\in p} x_p \le c_e \\
&\quad x \ge 0
\end{aligned}
$$

נניח הצלעות ממוספרות והמסלולים $s \to t$ ממוספרים, אז ניתן להציג:

$$
A_{ij} = \begin{cases} 1 & e_i \in p_j \\ 0 & otherwise \end{cases}
$$

נסתכל על הדואלי של הבעיה:

$$
\begin{aligned}
&\min \sum_{e\in E} c_e y_e \\
&y \ge 0 \\
&\forall p \in P: \sum_{e\in P} y_e \ge 1
\end{aligned}
$$

טענה 45.6 $max-flow \le min-cut$

נקח חתך מינימלי, נגדיר:

$$
y_e = \begin{cases} 1 & \text{e crosses the cut} \\ 0 & otherwise \end{cases}
$$

46 תרגיל 11

הגדרה 46.1 פתרון בסיסי מותר

יהיו $A \in M_{m,n}(\mathbb{R})$ כש $m \le n$ ו$rank(A) = m$, $b \in \mathbb{R}^m$. אז $y \in \mathbb{R}^n$ יקרא פתרון בסיסי מותר לתוכנית הלינארית $Ax = b,\ x \ge 0$ אם קיים $B \in \binom{[n]}{m}$ כך ש:

1. A_B לא סינגולרית (כש$A_B \in M_{m\times m}(\mathbb{R})$ והעמודות שלה הן העמודות של A המאונדקסות לפי B).
2. $supp(y) \subseteq B$ (כלומר $y_i \neq 0 \Rightarrow i \in B$)

3. y פתרון מותר ($Ay=b, y\geq 0$).

הגדרה 46.2 נקודה קיצונית בסט קמור
נקודה $z\in C$ נקראת קיצונית אם: $\nexists x\neq y\in C \ s.t. \ z=\lambda x+(1-\lambda)y, \ \lambda\in(0,1)$

טענה 46.3 יהיו $A\in M_{m,n}(\mathbb{R})$ כש $m\leq n$ ו$rank(A)=m$, $b\in\mathbb{R}^m$, יהיה $B\in\binom{[n]}{m}$ כך שA_B לא סינגולרית, אז קיים $x\in\mathbb{R}^n$ ייחודי כך ש $Ax=b$ ו$supp(x)\subseteq B$.

טענה 46.4 יהיו $A\in M_{m,n}(\mathbb{R})$ כש $m\leq n$ ו$rank(A)=m$, $b\in\mathbb{R}^m$, x הוא נקודה קיצונית של הפוליהדרון $P=\{y\in\mathbb{R}^n: \ y\geq 0, \ Ay=b\}$ אם"ם x הוא פתרון בסיסי מותר.

טענה 46.5 אם לתוכנית הלינארית $\max c^Tx, \ x\in P$ יש פתרון אופטימלי, אז יש לה פתרון אופטימלי שהוא פתרון בסיסי מותר.

הגדרה 46.6 מטריצה totally unimodular
מטריצה A תקרא totally unimodular אם לכל תת מטריצה C מתקיים $\det(C)\in\{-1,0,1\}$

טענה 46.7 תהי $A\in M_{m\times n}(\mathbb{Z})$, $m\leq n$, $\text{rank}(A)=m$ וA היא totally unimodular. יהיו $b\in\mathbb{Z}^m$ ו$c\in\mathbb{R}^n$.
אז אם קיים פתרון אופטימלי לתוכנית:

$$\begin{aligned} &\max c^Tx\\ &s.t.\\ &Ax=b\\ &x\geq 0\end{aligned}$$

אז יש פתרון אופטימלי עם קורדינטות בשלמים.

משפט 46.8 קריימר (שימושי לטענה הקודמת)
אם $A_{n\times n}x=b$ כש$\det(A)\neq 0$ אז למערכת משוואות הנ"ל יש פתרון ייחודי והוא:

$$\forall i\in[n]: \ x_i=\frac{\det(A_i')}{\det(A)}$$

כשA_i' זהה למטריצה A רק שהחליפו את העמודה הi בוקטור b.

טענה 46.9 תנאים מספיקים למטריצה totally unimodular
תנאים מספיקים כדי שמטריצה A תהיה TU:

1. $A_{ij}\in\{-1,0,1\}$
2. אחד מהתנאים הבאים צריך להתקיים:

(א) קיימת שורה או עמודה עם ערך אחד שונה מאפס.
(ב) קיימת שורה או עמודה שכולה אפסים.
(ג) לכל עמודה יש בדיוק 1 אחד ו-1 אחד.

47 שיעור 12

47.1 פתרון LP

47.1.1 מתכון

מביטים בLP:

$$\begin{aligned} &\max \langle c,x\rangle \\ &s.t. \\ &Ax \le b \\ &x \ge 0 \end{aligned}$$

האם המערכת

$$\begin{aligned} &\langle c,x\rangle \ge T \\ &s.t. \\ &Ax \le b \\ &x \ge 0 \end{aligned}$$

ספיקה?

ע"י ביצוע חיפוש בינארי על ערכי T, ניתן להגיע קרוב מאוד לפתרון. מכאן, אנו יודעים שיש בסיס של עמודות A, נסמנו מB, כך שאיתו מוצאים את האופטימום:

$$Bx| = b$$

47.1.2 אישור (certificate) לאי ספיקות

עקרון הדואליות אומר שאם מערכת של משוואות ואי שווינים לינארים אינה ספיקה, אז יש לכך הוכחה מהאופי הבא: יש צירוף מותר (כלומרת כופלים משוואות בגדלים כרצוננות אי שווינים בגדלים אי שליליים) הנותן $0 > 1$.

47.1.3 ההיבט הגיאומטרי

משפטי הפרדה בין קבוצות קמורות.

47.2 דואליות

משפט 47.1 המשפט החזק על דואליות LP

נביט בזוג התוכניות הלינאריות הבאות:

DLP	LP
$\min \langle b,y\rangle$	$\max \langle c,x\rangle$
$s.t.$	$s.t.$
$yA \ge c$	$Ax \le b$
$y \ge 0$	$x \ge 0$

נניח שהתוכנית LP ספיקה וחסומה. אז גם DLP היא כזו, ומתקיים השוויון:

$$\max \langle c, x\rangle = \min \langle b, y\rangle$$

משפט 47.2 המשפט החלש על דואליות LP
בהנחה שLP, DLP ספיקים, אם x מספק את LP, y מספק את DLP, אז:

$$\langle c, x\rangle \leq \langle b, y\rangle$$

הוכחה: נסתכל על:

$$\langle b, y\rangle \geq \langle y, Ax\rangle = yAx = \langle yA, x\rangle \geq \langle c, x\rangle$$

■

למה 47.3 פרקש
תהיה $A_{m\times n}$ ממשית, ו$b \in \mathbb{R}^m$, אז:

$$\exists y \in \mathbb{R}^m \ .s.t \ \langle y, b\rangle < 0, \ 0 \leq yA \Leftrightarrow \nexists x \geq 0 \ .s.t \ Ax = b$$

הוכחה: כיוון $\Rightarrow$:
אם יש y כזה: נניח שיש $x \geq 0$ כך ש$Ax = b$, ונביט ב:

$$0 \leq \left\langle \underbrace{yA,}_{\geq 0} \underbrace{x}_{\geq 0} \right\rangle = yAx = \langle y, Ax\rangle = \langle y, b\rangle < 0$$

וקיבלנו סתירה. ■

הערה 47.4 אינטרפרטציה גיאומטרית של למת פרקש
הקונוס הנוצר ע"י הוקטורים $a_1, \cdots, a_n \in \mathbb{R}^m$:

$$C = cone\left(a_1, \cdots, a_n\right) = \left\{\sum \lambda_i a_i \mid \forall i: \ \lambda_i \geq 0\right\}$$

נשים לב ש:

$$\{Ax \mid x \geq 0\} = C = cone\left(a_1, \cdots, a_n\right)$$

כשa_i הן עמודות של A. הנחת המשפט היא ש$b \notin C$.

למה 47.5 ניסוח אחר של פרקש
אם $C \subseteq \mathbb{R}^n$ קונוס, ו$b \notin C$, אז יש על מישור דרך הראשית שמפריד את b מC:

$$H = \{z \mid \langle y, z\rangle = 0\}$$
$$b \in H^-$$
$$C \subseteq H^+$$

מסקנה 47.6 איך מסיקים את הדואליות החזקה מלמת פרקש?
נאמר שרוצים לברר אם יש $x \geq 0$ המקיים $Ax \leq b$ שבשבילו $\langle c, x\rangle > T$, ונניח שהתשובה שלילית. אם המערכת הזו אינה ספיקה, יש לכך הוכחה "לינארית", ז"א יש צירוף במקדמים ≥ 0 של מערכת האי שווינים $Ax \leq b$ וכן $yA \geq c$, יש $y \geq 0$ שכאשר נכפיל אותו סקלרית בשני אגפי האי שוויון:

$$\langle c, x\rangle \leq yAx \leq \langle b, y\rangle < T$$

הגדרה 47.7 הפרדה לינארית חזקה
אומרים שיש הפרדה לינארית חזקה בין הקבוצות $S, T \subseteq \mathbb{R}^n$ אם יש על מישור H כך ש:

$$S \subseteq H^+_{open},\ T \subseteq H^-_{open}$$

למה 47.8 הטלה
תהיה $K \subseteq \mathbb{R}^n$ קבוצה קמורה וסגורה, ו$b \notin K,\ b \in \mathbb{R}^n$, אז:

1. יש וקטור אחד ויחיד $z \in K$ שממזער את המרחק $\|b - x\|_2$ ע"פ כל $x \in K$. לz הזו קוראים ההיטל של b על K.
2. ההיטל מאופיין על ידי כך שלכל $x \in K$ מתקיים: $\langle x - z, b - z\rangle \leq 0$

47.2.1 בעיות דואליות

הגדרה 47.9 הבעיה הדואלית של בעיית התזונה של הפרה
בעיית התזונה של הפרה מוגדרת כך :נתון וקטור $b \geq 0$: דרישות הימנימום של הפרה בכ"א מאבות המזון, וכן נתון וקטור $c \geq 0$ של מחיריהם של סוגי האוכל השונים. ממשפט הדואליות נקבל:

LP	DLP
$\min \langle c, x\rangle$	$\max \langle b, y\rangle$
$s.t.$	$s.t.$
$Ax \geq b$	$yA \leq c$
$x \geq 0$	$y \geq 0$

ניתן להסתכל על הDLP כך: יש תעשיין שיכול לייצר "גלולות" של אבות המזון. y_i הוא המחיר שהוא יגבה על יחידה אחת של גלולת אב המזון הi. האילוץ $yA \leq c$ מבטא את האילוץ של התעשיין להיות תחרותי עם סוגי המזון המסורתיים. ז"א $(yA)_j \leq c_j$ אומר שצירוף הגלולות השקול ליח' אחת של סוג אוכל j אומר לעלות לכל היותר כמחיר יחידה אחת של סוג זה, דהיינו c_j.

הגדרה 47.10 בעיית הזרימה
נתונה רשת זרימה: גרף מכוון $G=(V,E)$ עם מקור $s\in V$ ובור $t\in V$, וקיבולים לצלעות $c:E\to\mathbb{R}_{\geq 0}$.
זרימה זו פונקציה על צלעות מכוונות שמקיימת את שימור החומר בכל קודקוד השונה מs,t וחסומה מלמעלה על ידי c. מחפשים זרימה עם שטף מירבי, ז"א $\max\sum_{s\to_e} f(e)$. או בניסוח אחר, אם נגדיר:

$$P=\{\text{All directed rails from s to t}\}$$

$$A_{ij}=\begin{cases}1 & e_i\in P_j\\ 0 & otherwise\end{cases}$$

אז הבעיה היא

$$\begin{aligned}&\max\langle x,1\rangle\\ &s.t.\\ &Ax\leq c\\ &x\geq 0\end{aligned}$$

משפט 47.11 משפט השטוף והחתך ($maxflow=min-cut$)
משפט הדואליות של LP אומר שבעיית הזרימה שקולה ל:

$$\begin{aligned}&\min\langle y,c\rangle\\ &s.t.\\ &yA\geq 1\\ &y\geq 0\end{aligned}$$

פירוש: אם $y\in\{0,1\}^{|E|}$, האילוץ $yA\geq 1$ אומר שזו קבוצת צלעות הפוגשת כל מסילה $s\to t$.

47.3 אלגוריתם האליפסואידים

47.3.1 הגדרת הבעיה

עוסקים רק בבעיות ספיקות.
יש גוף קמור $P\subseteq\mathbb{R}^n$, ועלינו לברר אם $P\overset{?}{=}\emptyset$.
הנחות:

1. ידוע לנו $R>0$ כך ש$P\subseteq B(0,R)$
2. אם $P\neq 0$, אז הוא מכיל כדור ברדיוס $r>0$.

איך מתואר לנו הגוף P?
ע"י שני אובות:

1. אוב שייכות (membership oracle) - יודע לענות על השאלה: בהינתן $x\in\mathbb{R}^n$, האם $x\in P$?
2. אוב הפרדה (separation oracle). בהנחה ש$x\notin P$ (כפי שענה לנו אוב השייכות), אוב ההפרדה ייתן לנו על מישור שx נמצא מצידו האחד וP מצידו האחר.

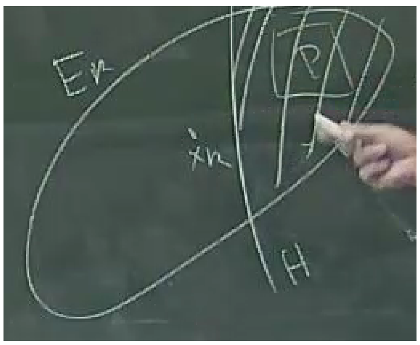
47.3.2 האלגוריתם

הגדרה 47.12 אליפסואיד
תמונה של כדור היחידה תחת העתקה אפינית.

בכל צעד יהיה לנו אליפסואיד E_k שמרכזו x_k ו-$P \subseteq E_k$. הערכים ההתחלתיים הם:

$$x_0 = 0$$
$$E_0 = B\left(0, R\right)$$

בצעד ה-k שואלים את אוב השייכות אם $x_k \in P$. אם כן - $P \neq \emptyset$ וסיימנו. אם לא - פונים לאוב ההפרדה ומקבלים על מישור H.

למה 47.13 יש אליפסואיד E_{k+1} שמכיל את השטח המסומן 

ונפחו מקיים

$$vol\left(E_{k+1}\right) < \left(1 - \frac{1}{n^2}\right) vol\left(E_k\right)$$

הגדרה 47.14 חצי אליפסואיד
החיתוך של אליפסואיד עם חצי מרחב המוגדר ע"י על מישור שעובר דרך מרכז האליפסואיד.

למה 47.15 יהיה $E \subseteq \mathbb{R}^n$ אליפסואיד שמרכזו ב-x, ויהיה E^+ חצי מ-E כלשהו, אז יש אליפסואיד E' המכיל את E^+ ומקיים:

$$vol\left(E'\right) < \left(1 - \frac{1}{n^2}\right) vol\left(E\right)$$

הוכחה: רעיון ההוכחה: יוצרים העתקה בין E לבין כדור היחידה, ומוכיחים את הלמה השקולה עליו:

למה 47.16 יש אליפסואיד ב-$\mathbb{R}^n$ בנפח $< \left(1 - \frac{1}{n^2}\right) vol\left(\text{unit ball in } \mathbb{R}\text{ˆn}\right)$, ואליפסואיד זה מכיל את חצי הכדור:

$$\|x\| \leq 1,\ x_n > 0$$

■

48 תרגול 12

48.1 מבנה המבחן

יהיו 4 שאלות, לבחור 3. כל שאלה תקבל ציון בין 0 ל10. אם הציונים הם $x_1 \geq x_2 \geq x_3$ אז הציון יהיה:

$$5x_1 + 3x_2 + 2x_3$$

ובנוסף יש 2 נקודות בונוס על לכתוב במחברת על אילו שאלות ענינו.

48.2 משחקים סכום אפס ומשפט המינימקס

48.2.1 אבן, נייר ומספריים

$$A := \text{rows (R)} \begin{array}{c|ccc} & \multicolumn{3}{c}{\text{columns player (C)}} \\ & rock & paper & scissors \\ rock & 0 & -1 & 1 \\ paper & 1 & 0 & -1 \\ scissors & -1 & 1 & 0 \end{array}$$

מה כדאי לR לשחק? אם האסטרטגיה של C דולפת לR לפני המשחק, אז R יכול לדאוג לתשלום 1.
הקבוצה $\{$אבן, נייר, מספריים$\}$ נקראת קבוצת האסטרטגיות **הטהורות**.
אסטרטגיה **מעורבת** זו התפלגות על האסטרטגיות הטהורות.
נניח C משחק לפי האסטרטגיה המעורבת $q^T = (q_1, q_2, q_3)$ וR משחק לפי $p^T = (p_1, p_2, p_3)$, ואת המטריצה מקודם נסמן בA, אז התוחלת של התשלום לR היא:

$$p^T Aq$$

48.2.2 באופן כללי

הגדרה 48.1 משחק סכום אפס של שני שחקנים
זוהי מטריצה $A \in M_{m\times n}(\mathbb{R})$, האסטרטגיות הטהורות של שחקן השורות R הן $[m]$, של שחקן העמודות C הן $[n]$. אסטרטגיה מעורבת היא התפלגות על הטהורות. בהינתן אסטרטגיה p של R וq של C, תוחלת התשלום לR היא:

$$\varphi(p,q) = p^T Aq$$

R ממקסם, C ממזער.

מה כדאי לעשות? נניח שR יודע שהאסטרטגיה p^* תגיע לC. בהנחה זו C יבחר אסטרטגיה q שתמזער

$$p^{*T} Aq$$

ואז התשלום לR יהיה

$$\min_{q\in\Delta_{n-1}}\left(p^{*T}Aq\right)=$$
$$\min_{i\in[n]}\left(p^{*T}Ae_i\right)$$

כל e_i מתאים לאסטרטגיה טהורה אחרת. אמנם $q\in\mathbb{R}^n$, אבל כיוון שזו הסתברות אז יש רק $n-1$ דרגות חופש. R יחפש אסטרטגיה p שממקסמת את $\min_{q\in\Delta_{n-1}}\left(p^{*T}Aq\right)=\min_{i\in[n]}\left(p^{*T}Ae_i\right)$. R יפתור את התוכנית הבאה:

$$\begin{aligned}\max v&=v^+-v^-\\ \forall i:v&\leq p^TAe_i\\ \begin{pmatrix}1&\cdots&1\end{pmatrix}p&=1\\ p&\geq 0\end{aligned}$$

התוכנית תתן ערך v ואסטרטגיה p^* כך שלכל אסטרטגיה q של C:

$$p^{*T}Aq\geq v=\max_{p\in\Delta_{m-1}}\min_{q\in\Delta_{n-1}}p^TAq$$

שחקן העמודות מעוניין לפתור:

$$\min_{q\in\Delta_{n-1}}\max_{p\in\Delta_{m-1}}p^TAq=\min_{q\in\Delta_{n-1}}\max_{i\in[m]}e_i^TAq$$

הוא יפתור את התוכנית הלינארית:

$$\begin{aligned}\min u&=u^+-u^-\\ \forall i:u&\geq e_i^TAq\\ \begin{pmatrix}1&\cdots&1\end{pmatrix}q&=1\\ q&\geq 0\end{aligned}$$

נשים לב:

$$\text{"}\forall i:\ u\geq e_i^TAq\text{"}=\text{"}A^Tq-u\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}\leq 0\text{"}=\text{"}\begin{pmatrix}A&-\vec{1}\\ \vec{1}&0\\ -\vec{1}&\end{pmatrix}\begin{pmatrix}q\\ u\end{pmatrix}\leq\begin{pmatrix}0\\ 1\\ -1\end{pmatrix}\text{"}$$
$$\text{"}\forall i:\ v\leq p^TAe_i\text{"}=\text{"}A^Tp-\begin{pmatrix}v\\ \vdots\\ v\end{pmatrix}\geq 0\text{"}=\text{"}\begin{pmatrix}A^T&-\vec{1}\\ \vec{1}&0\\ -\vec{1}&\end{pmatrix}\begin{pmatrix}p\\ u\end{pmatrix}\geq\begin{pmatrix}0\\ 1\\ -1\end{pmatrix}\text{"}$$

49 תרגול אקסטרה 12

49.1 דואליות

49.1.1 בעיה

נתון גרף מכוון $G=(V,E)$, $V=\{1,\cdots,n\}$, $w:E\to[0,\infty]$. מטרה: מסלול הקצר מ1 לn.
נביט בLP הבאה:

$$\begin{aligned} & \max x_n \\ & s.t. \\ & x\in\mathbb{R}^n \\ & x\geq 0 \\ \forall e=(u\to v): & \ x_v-x_u\leq w(e) \\ & x_1=0 \end{aligned}$$

נשים לב שזה שקול ל

$$\begin{aligned} & \max x_n \\ & s.t. \\ & x\in\mathbb{R}^n \\ & x\geq 0 \\ \forall e=(u\to v): & \ x_v-x_u\leq w(e) \\ & x_1\leq 0 \end{aligned}$$

תוכנית

1. נמצא השמה $x\in\mathbb{R}^n$ פיזיבילית כך ש$x_n=d(1,n)$ (ולכן $OPT\geq d(1,n)$).
2. נחשב ונבין את התוכנית הדואלית.
3. נמצא השמה לתוכנית הדואלית שערכה הוא גם $d(1,n)$ (ולכן $OPT\leq d(1,n)$).

השמה פיזיבילית

נגדיר לכל i: $x_i=d(1,i)$ (המשקל של המסלול המינימלי מ1 לi). צ"ל:

$$\forall e=u\to v:\ \ d(1,v)\leq d(1,u)+w(u\to v)$$

נשים לב שזה נכון. נניח בשלילה שלא, אז $d(1,u)+w(u\to v)<d(1,v)$, כלומר מצאנו מסלול קצר יותר מ1 לu מהמסלול המינימלי. סתירה.
בנוסף, מתקיים מההגדרה $d(1,1)$, ולכן ההשמה פיזיבילית. נקבל שפונקציית המטרה היא:

$$x_n=d(1,n)$$

תוכנית דואלית

נסתכל על היצוג המטריציוני של הבעיה הפרימלית:

$$\begin{array}{cc} & \text{vertices (variables)} \\ \text{edges} & \begin{pmatrix} 0 & -1 & 0 & 1 & 0 \\ \\ 1 & 0 & \cdots & \cdots & 0 \end{pmatrix} \end{array} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \leq \begin{pmatrix} w(e) \\ \\ 0 \end{pmatrix}$$

$$A_{e=(u\to v),u} = -1$$
$$A_{e=(u\to v),v} = 1$$
$$A_e x \leq w(e)$$

נסתכל על הבעיה הדואלית: יש משתנה לכל צלע: $y \in \mathbb{R}^{m+1}$, ומשתנה אחרון y_{m+1}. הייצוג המטריציוני נראה כך:

$$\begin{array}{cc} & \text{edges} \\ \text{vertices (variables)} & \begin{pmatrix} & & & 1 \\ 1\,(\text{u}\to\text{v}) & -1\,(\text{v}\to\text{w}) & -1 & 0 \\ & & & \vdots \\ & & & \vdots \\ & & & 0 \end{pmatrix} \end{array} \begin{pmatrix} y \end{pmatrix} \geq \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

פונקציית המטרה היא:

$$\min \left\langle \begin{pmatrix} w(e) \\ \\ 0 \end{pmatrix}, y \right\rangle = \min \left(\sum_e w(e)\, y_e \right)$$

והבעיה הלינארית היא:

$$\begin{aligned} &\min \left(\sum_e w(e)\, y_e \right) \\ &s.t. \\ \text{first vertex:}\, &y_{m+1} + \sum_{e:\to^e 1} y_e - \sum_{e:1\to^e} y_e \geq 0 \\ \text{last vertex;}\, &\sum_{e:\to^e n} y_e - \sum_{e:n\to^e} y_e \geq 1 \\ \forall v \neq 1, n :\, &\sum_{e:\to^e v} y_e - \sum_{e:nv\to} y_e \geq 0 \\ &0 \leq y \in \mathbb{R}^{m+1} \end{aligned}$$

השמה לתוכנית הדואלית

המטרה: למצוא השמה חוקית $y\in\mathbb{R}^{m+1}$ כך ש:

$$\sum w\left(e\right)y_{e}=d\left(1,n\right)$$

הצעה: יהי P המסלול הקצר ביותר בין 1 לn. נגדיר:

$$y_{e}=\begin{cases}1 & e\in P\\0 & otherwise\end{cases}$$
$$y_{m+1}=1$$

נשים לב שההשמה הנ"ל עונה על כל האילוצים, כשהאילוץ הראשון הוא אילוץ ריק כי ניתן להגדיל את y_{m+1} עד אינסוף והאילוץ יתקיים.

מסקנה

ניחשנו פתרון x לתוכנית הפרימלית, ניחשנו פתרון y לתוכנית הדואלית. בגלל שהערכים של הפתרונות זהים, זהו הפתרון האופטימלי.

50 בוחן 3

הגדרה 50.1 בעיית יצירת גרף חסר משולשים
יהי $G=\left(V,E\right)$ גרף עם n קודקודים. רוצים להסיר כמה שפחות צלעות מG כך שניוותר עם גרף חסר משולשים.

טענה 50.2 נניח המשולשים ממוספרים $t_{1},\cdots$ והצלעות ממוספרות $e_{1},\cdots$. נגדיר $A_{t,e}=\begin{cases}1 & e\in t\\0 & otherwise\end{cases}$. תוכנית לינארית בשלמים לבעיה הנ"ל:

$$\begin{aligned}&\min\left\langle \vec{1},x\right\rangle \\&s.t.\\&Ax\geq1\\&x\in\left\{0,1\right\}^{\left|E\right|}\end{aligned}$$

נשים לב שהבעיה הנ"ל שקולה ל$set-cover$ (המשולשים הם הסטים, הצלעות הן האיברים בכל סט).

חלק IV
חזרה

51 תרגול מיכאל

51.1 דואליות

נניח $A\in M_{m\times n}\left(\mathbb{R}\right)$, $g,b\in\mathbb{R}^{m},\ f,c\in\mathbb{R}^{n}$. נסתכל על הבעיה:

$$\begin{aligned} & \max\left(c^{T}x\right)\\ & s.t.\\ & Ax\leq b\\ & x\geq0 \end{aligned}$$

הבעיה הדואלית היא

$$\begin{aligned} & \max\left(b^{T}y\right)\\ & s.t.\\ & A^{T}y\geq c\\ & y\geq0 \end{aligned}$$

טענה 51.1 f,g פתרונות אופטימליים לבעיות (f לפרימלי, g לדואלי), אם"ם:

$$(*):\ \forall i:\ \left(Af\right)_{i}=b_{i}\vee g_{i}=0\text{ and }\forall j:\ \left(A^{T}g\right)_{j}=c_{j}\vee f_{j}=0$$

תנאים אלו נקראים complementary slackness conditions.

הערה 51.2 דרך אחרת לכתוב את התנאי:

$$g^{T}\left(Af-b\right)=0\text{ and }f^{T}\left(A^{T}g-c\right)=0$$

נשים לב שזה נובע מכך ש$f,A^{T}g-c$ הם אי שליליים.

הוכחה: כיוון 1: נניח שf,g אופטימליים. ממשפט הדואליות החזק:

$$c^Tf = b^Tg$$

$$\begin{aligned}
&Af \leq b \\
&\Downarrow \\
&0 \leq b - Af \\
&\Downarrow \\
&0 \leq g^T(b - Af) = g^Tb - g^TAf \\
&\Downarrow \\
&0 \geq^1 g^TAf - g^Tb
\end{aligned}$$

$$\begin{aligned}
&A^Tg \geq c \\
&\Downarrow \\
&0 \leq A^Tg - c \\
&\Downarrow \\
&0 \leq f^T\left(A^Tg - c\right) = f^TA^Tg - f^Tc \overset{\text{strong duality}}{=} \\
&= f^TA^Tg - b^Tg \leq^1 0 \\
&\Downarrow \text{sandwich} \\
&0 = g^T(b - Af)
\end{aligned}$$

כיוון 2: עתה, נניח ש$(*)$ מתקיים, ונוכיח שf,g אופטימליים. צ"ל (לפי דואליות חלשה): $c^Tf = b^Tg$. כלומר, אם y פיזיבילי דואלית, אז הפתרון של הפרימאלי מקיים $b^Ty \geq$. מ$(*)$:

$$\begin{aligned}
&g^TAf = \underbrace{g^Tb}_{\text{a number}} = b^Tg \\
&g^TAf = f^TA^Tg = f^Tc = c^Tf \\
&\Downarrow \\
&b^Tg = c^Tf
\end{aligned}$$

■

51.2 השיטה ההסתברותית

טענה 51.3 יהי $G \sim G(n,p)$. אז:

1. אם $p = (1-\epsilon)\frac{\ln(n)}{n}$ אזי אכב"ו G לא קשיר.
2. אם $p = (1+\epsilon)\frac{\ln(n)}{n}$ אזי אכב"ו G קשיר.

הוכחה: כל סעיף בנפרד:

1. אם יש קודקוד מבודד $\Leftrightarrow G$ לא קשיר. נעריך את ההסתברות שזה יקרה. נסמן בX_i את המ"מ המציין שקודקוד i מבודד.

$$X = \sum_{i\in[n]} X_i$$
$$\Downarrow$$
$$0 < X \Leftrightarrow \text{there is an isolated vertex}$$

ההסתברות לכל צלע ב"ת כי אנחנו ב$G\left(n,p\right)$, לכן:

$$\Pr\left(X_i > 0\right) \overset{\text{indicator}}{=} \mathbb{E}\left[X_i\right] = \left(1-p\right)^{n-1}$$
$$\Downarrow$$
$$\mathbb{E}\left[X\right] = n\left(1-p\right)^{n-1}$$

נחשב את השונות של X:

$$\mathbb{E}\left[X^2\right] = \mathbb{E}\left[\left(\sum_i X_i\right)^2\right] = \sum_i \mathbb{E}\left[X_i^2\right] + \sum_{j\neq i}\mathbb{E}\left[X_iX_j\right] = n\mathbb{E}\left[X_1\right] + \sum_{j\neq i}\mathbb{E}\left[X_iX_j\right] =$$
$$= \mathbb{E}\left[X\right] + \sum_{j\neq i}\mathbb{E}\left[X_iX_j\right] = \mathbb{E}\left[X\right] + n\left(n-1\right)\Pr\left(X_1, X_2 \text{ are both isolated}\right) =^1$$

נשים לב שאם אחד מבודד, אז ידוע לנו שהצלע בינו לבין האחר לא קיימת, ולכן:

$$=^1 \mathbb{E}\left[X\right] + n\left(n-1\right)\left(1-p\right)^{2n-3} = \mathbb{E}\left[X\right] + \frac{n^2\left(1-p\right)^{2n-2}}{1-p} - n\left(1-p\right)^{2n-3} =$$
$$= \mathbb{E}\left[X\right] + \frac{\mathbb{E}\left[X\right]^2}{1-p} - \mathbb{E}\left[X\right]\left(1-p\right)^{n-2}$$

מהנתון על p:

$$(1-p)^{n-2}=\left(1-(1-\epsilon)\frac{\ln(n)}{n}\right)^{n-2}=\exp\left(-(1-\epsilon)^{\ln(n)}\right)(1+o(1))=\frac{1+o(1)}{n^{1-\epsilon}}$$
$$\Downarrow$$
$$\mathbb{E}\left[X^2\right]=\mathbb{E}[X]+\frac{\mathbb{E}[X]^2}{1-p}-\mathbb{E}[X](1-p)^{n-2}=\mathbb{E}[X]+\frac{\mathbb{E}[X]^2}{1+o(1)}-\mathbb{E}[X]\,o(1)$$
$$\Downarrow$$
$$Var[X]=\mathbb{E}[X](1+o(1))+o(1)\,\mathbb{E}[X]^2$$
$$\Downarrow$$
$$\Pr(X=0)\le\Pr(|X-\mathbb{E}[X]|\ge\mathbb{E}[X])\overset{\text{chebyshev}}{\le}\frac{Var[X]}{\mathbb{E}[X]^2}=\frac{o(1)\left(\mathbb{E}[X]+\mathbb{E}[X]^2\right)}{\mathbb{E}[X]^2}=$$
$$=\frac{o(1)}{\mathbb{E}[X]^2}+o(1)$$

חישוב מהיר:

$$\mathbb{E}[X]=n(1-p)^{n-1}=n\left(1-(1-\epsilon)\frac{\ln(n)}{n}\right)^{n-1}\approx\frac{n}{n^{1-\epsilon}}=n^{\epsilon}\to\infty$$

2. אבחנה: G לא קשיר גורר שקיימת קבוצה $W\subseteq V$ כך ש$0<|W|\le\frac{n}{2}$ שאין לה צלעות יוצאות, וגם היא קשירה. עבור $|W|=k$, נחסום מלעיל את ההסתברות שW מקיים את התנאים:

$$\Pr(W\text{ has no outgoing edges})=(1-p)^{k(n-k)}$$
$$\Pr(W\text{ is connected})=\Pr(\text{W has a spanning tree})\le\mathbb{E}[\text{the number of spanning trees in W}]=$$
$$=p^{k-1}\underbrace{k^{k-2}}_{\text{Cayley's formula}}$$
$$\Downarrow$$
$$\Pr\left(\substack{\text{W has no outgoing}\\ \text{edges and is connected}}\right)\le(1-p)^{k(n-k)}p^{k-1}k^{k-2}$$

נשים לב שהמאורעות הנ"ל ב"ת. כעת נחשב:

$$\binom{n}{k}(1-p)^{k(n-k)}p^{k-1}k^{k-2}\le\left(\frac{en}{k}\right)^k\exp\left(-pk\left(n-k\right)\right)p^{k-1}k^{k-2}\le$$
$$\le\exp\left(k\ln\left(e\right)+k\ln\left(n\right)-k\ln\left(k\right)-\left(1+\epsilon\right)\frac{\ln\left(n\right)}{n}k\left(n-k\right)+\left(k-1\right)\ln\left(p\right)+\left(k-2\right)\ln\left(k\right)\right)\le$$
$$\le\exp\left(-k\left(\epsilon\ln\left(n\right)-O\left(\ln\left(\ln\left(n\right)\right)\right)\right)\right)\le q\left(n\right)^k,\ q\left(n\right)\to0$$
$$\Downarrow$$
$$\sum_{k=1}^{\frac{n}{2}}\binom{n}{k}(1-p)^{k(n-k)}p^{k-1}k^{k-2}\le\sum_{k=1}^{\infty}q\left(n\right)^k=\underbrace{q\left(n\right)}_{\to0}\underbrace{\frac{1}{1-q\left(n\right)}}_{\to1}\to0$$

■

51.3 מרטינגיילים וריכוז מידה

הגדרה 51.4 מרטינגיילים
(Ω,Pr) מרחב הסתברות, X מ"מ, נגדיר עידונים:

$$\{\emptyset,\Omega\}=\Omega_1\supseteq\Omega_1\supseteq\cdots\supseteq2^{\Omega}$$

נגדיר:

$$X_t=\mathbb{E}\left[X|\Omega_0,\cdots,\Omega_t\right]$$

אז $X_0=\mathbb{E}\left[X\right],X_1,\cdots$ נקרא מרטינגייל.

משפט 51.5 א"ש אזומה
אם $X_0,\cdots,X_k$ מרטינגייל, ולכל i: $|X_i-X_{i-1}|\le a_i$, אז:

$$\Pr\left(\left|X_k-\mathbb{E}\left[X_0\right]\right|\ge\lambda\right)\le2\exp\left(-\frac{\lambda^2}{\sum_{i=1}^{k}a_i^2}\Theta\left(1\right)\right)$$

בעיה

יהי $n\ge3$ כך ש$\binom{n}{2}$ זוגי, יהי $m=\frac{\binom{n}{2}}{2}$. נגדיר סדרה של גרפים: $G_0,G_1,\cdots,G_m$ ע"י:

1. $G_0=K_n$.
2. לכל t, G_{t+1} מתקבל מG_t על ידי מחיקת צלע מקרית.

עבור גרף G נסמן ב$T\left(G\right)$ את מס' המשולשים בG. נרצה לדעת כמה משולשים יש בG_m.

שאלות

1. מהו $\mu := \mathbb{E}\left[T\left(G_m\right)\right]$?
2. עבור $0 \le t < m$ כתבו את $\mathbb{E}\left[T\left(G_m\right)|G_0,\cdots,G_t\right]$ כביטוי התלוי ב$t,n,m,T\left(G_t\right)$.
3. הוכיחו שלכל $0 \le t < m$ מתקיים: $\left|T\left(G_{t+1}\right)-T\left(G_t\right)\right| \le n$
4. הוכיחו שלכל $0 \le t < m$ מתקיים:

$$\left|\mathbb{E}\left[T\left(G_m\right)|G_0,\cdots,G_{t+1}\right]-\mathbb{E}\left[T\left(G_m\right)|G_0,\cdots,G_t\right]\right| = O\left(n\right)$$

5. הוכיחו:

$$\Pr\left(\left|T\left(G_m\right)-\mu\right| \ge n^2\log n\right) = O\left(\frac{1}{n}\right)$$

תשובות

1.

$$\mu = \mathbb{E}\left[T\left(G_m\right)\right] = \binom{n}{3}\frac{m}{\binom{n}{2}}\frac{m-1}{\binom{n}{2}-1}\frac{m-2}{\binom{n}{2}-2}$$

2.

$$\mathbb{E}\left[T\left(G_m\right)|G_0,\cdots,G_t\right] = T\left(G_t\right)\frac{m}{\binom{n}{2}-t}\frac{m-1}{\binom{n}{2}-t-1}\frac{m-2}{\binom{n}{2}-t-2}$$

מתקיים:

$$\mathbb{E}\left[\mathbb{E}\left[T\left(G_m\right)|G_0,\cdots,G_t\right]|\mathbb{E}\left[T\left(G_m\right)|G_0\right],\cdots,\mathbb{E}\left[T\left(G_m\right)|G_0,\cdots,G_{t-1}\right]\right] = \mathbb{E}\left[T\left(G_m\right)|G_0,\cdots,G_{t-1}\right]$$

ולכן זהו מרטינגייל.

3. ברור: בגרף עם n קודקודים, כל צלע משתתפת בפחות מn משולשים. מכאן:

$$\left|T\left(G_t\right)-T\left(G_{t+1}\right)\right| \le n$$

.4

$$\left|\mathbb{E}\left[T\left(G_m\right)|G_0,\cdots,G_t\right]-\mathbb{E}\left[T\left(G_m\right)|G_0,\cdots,G_{t+1}\right]\right| \overset{\text{from } 2}{=}$$
$$=\left|T\left(G_m\right)\frac{m}{\binom{n}{2}-t}\frac{m-1}{\binom{n}{2}-t-1}\frac{m-2}{\binom{n}{2}-t-2}-T\left(G_{t+1}\right)\frac{m}{\binom{n}{2}-t-1}\frac{m-1}{\binom{n}{2}-t-2}\frac{m-2}{\binom{n}{2}-t-3}\right|=$$
$$=\frac{m\left(m-1\right)\left(m-2\right)}{\left(\binom{n}{2}-t-1\right)\left(\binom{n}{2}-t-2\right)}\left|\frac{T\left(G_m\right)}{\binom{n}{2}-t}-\frac{T\left(G_{t+1}\right)}{\binom{n}{2}-t-3}\right|\leq$$
$$\leq m\left|\frac{\left(T\left(G_t\right)-T\left(G_{t+1}\right)\right)\left(\binom{n}{2}-t\right)-3T\left(G_t\right)}{\Omega\left(n^4\right)}\right|\leq\frac{O\left(n^2\right)}{\Omega\left(n^4\right)}\left(n\cdot n^2+3n^3\right)=O\left(n^{5-4}\right)=O\left(n\right)$$

5. נשים לב ש: $\left\{\mathbb{E}\left[T\left(G_m\right)|G_0,\cdots,G_t\right]\right\}_{t=0}^{m}$ מרטינגייל (כמו ב2), כמו כן גם:

$$\mu=\mathbb{E}\left[T\left(G_m\right)\right]=\mathbb{E}\left[T\left(G_m\right)|G_0\right]$$

לכן לפי א"ש אזומה:

$$\Pr\left(\left|T\left(G_m\right)-\mu\right|\geq n^2\log\left(n\right)\right)\leq 2\exp\left(-\frac{n^4\log^2\left(n\right)}{\sum_{i=0}^{\frac{n^2}{4}}O\left(n^2\right)}\right)=2\exp\left(-\frac{n^4\log^2\left(n\right)}{O\left(n^4\right)}\right)=$$
$$=2\exp\left(-\Omega\left(\log^2\left(n\right)\right)\right)\leq\frac{1}{n^{\Omega\left(\log\left(n\right)\right)}}=O\left(\frac{1}{n}\right)$$

51.4 שרשראות מרקוב (הילוך מקרי פשוט)

הגדרה 51.6 שרשרת מרקוב אי פריקה
שרשרת היא א"פ אם לכל שני מצבים i,j קיים t כך שההסתברות להגיע מi לj בt מהלכים חיובי.

הגדרה 51.7 שרשרת מרקוב לא מחזורית
שרשרת היא לא מחזורית אם קיים זמן t כך שלכל i: $\Pr\left(X_t=i\right)>0$.

טענה 51.8 הילוך מקרי על גרף הוא מחזורי אס"ם הגרף דו"צ.

טענה 51.9 גרם הוא דו"צ אס"ם אין בו מעגל באורך א"ז.

הגדרה 51.10 התפלגות סטציונארית
אם A מטריצת מעברים של שרשרת מרקוב, אז π (וקטור הסתברות) נקרא התפלגות הסטציונרית אם $\pi A=\pi$.

תרגיל

נתון לוח שח $n\times n$. דמויות לא יכולות לצאת מהלוח. על הלוח יש מלך שיכול לנוע לכל המשבצות מסביבו. מלך מרקובי מתחיל במשבצת X_0, ובכל שלב $t+1$ בוחר צעד חוקי בצורה ב"ת ואחידה וזז למשבצת X_{t+1}.

1. האם השרשרת אי־פריקה? כן.
2. האם היא מחזורית? יש פה מעגל אי זוגי, ולכן לא.

מסקנה 51.11 ההתפלגות של X_t שואפת להתפלגות הסטציונרי (תוצאה של פרון פרובניוס).

תרגיל

שאלה זהה רק עם פרש, שיכול לזוז שני צעדים בכיוון אחד ואז צעד אחד בכיוון המאונך אליו.

1. האם השרשרת אי־פריקה? כן. נוכיח שהפרש יכול להגיע לכל מקום. מספיק להוכיח שאם הוא בפינה הקיצונית, הוא יכול לזוז משבצת למעלה ומשבצת באלכסון.
2. האם היא מחזורית? כן ־ נשים לב שהגרף הוא דו צדדי (אם נצבע את הלוח בשחור ולבן, אז בכל צעד הפרש מחליף צבע).

תרגיל

הילוך מקרי על טורוס $\mathbb{Z}_m\times\mathbb{Z}_n$ עם הצעדים $\{(-1,0),(1,0),(0,-1),(0,1)\}$.

1. מצאו אפיון תורת מספרי על m,n כך שההילוך יהיה אי פריק ולא מחזורי. הנ"ל יתקבל אם"ם m או n א"ז. הוכחה: נניח m א"ז, אז הנה מעגל א"ז: $(0,0),(1,0),\cdots,(m-1,0),(0,0)$. אם הגרף מחזורי, אז שני המעגלים הבאים באורך זוגי:
$$\begin{gathered}(0,0),\cdots,(m-1,0)\\(0,0),\cdots,(0,n-1)\end{gathered}$$
אם הגרף לא דו"צ, אז שני המעגלים באורך זוגי $\Leftarrow m,n$ זוגיים.
2. נקבע $m=n$ א"ז. מתחילים הילוך מקרי ב$X_0=(0,0)$. נסמן בπ_t את ההתפלגות של X_t. הוכיחו:
$$\lim_{n\to\infty}\frac{1}{2}\left\Vert \pi_n-\underbrace{\pi}_{\text{stationary dist.}}\right\Vert =1$$
נסמן ב$X_1,\cdots,X_4$ את מס' המהלכים מכל סוג. אז:
$$\begin{aligned}X_i&\sim Bin\left(n,\frac{1}{4}\right)\\\mathbb{E}\left[X_i\right]&=\frac{n}{4}\\Var\left[X_i\right]&=\Theta\left(n\right)\end{aligned}$$
נשים לב:
$$\forall i:\ \Pr\left(\left|X_i-\frac{n}{4}\right|>\sqrt{n}\log\left(n\right)\left(=\frac{n}{4}\frac{4\log\left(n\right)}{\sqrt{n}}\right)\right)\le2\exp\left(-\frac{64\log^2\left(n\right)}{n}\frac{n}{4}\Theta\left(1\right)\right)=O\left(\frac{1}{n}\log\left(n\right)\right)$$

נגדיר את ה"אזור הכחול": $[-\sqrt{n}\log(n), -\sqrt{n}\log(n)]^2$ ונסמן: $A = \mathbb{Z}_n^2 \setminus [-\sqrt{n}\log(n), -\sqrt{n}\log(n)]^2$,
נשים לב:

$$P_\pi(A) = \frac{|A|}{n^2} = \frac{n^2 - 2n\log^2(n)}{n^2} = 1 - \frac{2\log^2(n)}{n} = 1 - o(1)$$

$$P_{\pi_t}(A) \leq O\left(\frac{1}{n^{\log(n)}}\right) = o(1)$$

$$\Downarrow$$

$$\lim_{n\to\infty} \frac{1}{2}\|\pi_n - \pi\|_1 \geq |P_{\pi_t}(A) - P_\pi(A)|$$

52 תרגול יובל