סיכום כלים מתמטיים למדעי המחשב

מההרצאות של פרופסור נתן ליניאל, על ידי ליאת שפירא liat.shapira@mail.huji.ac.il

18 בינואר 2021

תוכן עניינים

1 הסתברות

שבוע 2

1.1 שיטת מומנטים

בהינתן משתנה מקרי, אנחנו יודעים שהתוחלת היא סוג של ממוצע משוכלל של ערכיו של המשתנה, תוך התיחסות להסתברות של כל ערך להופיע. התוחלת כשל עצמה עדיין לא אומרת לנו מספיק - כדי להפיק את מירב התועלת ממנה נרצה לשאול מה היא מנבאת עבור המשתנה המקרי. באופן אידיאלי, היינו רוצים שהתוחלת תגיד לנו "בערך" מה אפשר לצפות בתור ערכו של המשתנה המקרי, כלומר היינו רוצים שבהסתברות גבוהה המשתנה המקרי יהיה קרוב מאוד לתוחלת שלו.

כמובן שזה לא תמיד קורה, אבל יש מקרים מסויימים שבהם אכן נוכל לתאר מצב שבו בהסתברות גבוהה מאוד X יהיה קרוב לתוחלת שלו. עולם הבעיה שלנו יכיל בעיקר בעיות שהן תלויות בn כלשהו, ונשאל מה קורה כאשר n שואף לאינסוף, כלומר מה המגמה של התהליך.

הבעיות יראו דומה למה שראינו בהסתברות לגבי התכנסות סדרת הסתברויות: נשתמש באי שיוויון מרקוב ואי שיוויון צ'בישב כדי לחסום הסתברויות. לשיטות שנציג קוראים **שיטת מומנטים.**

שיטת המומנט הראשון היא תוך שימוש באי שיוויון מרקוב, שנותן לנו:

$$\mathbb{P}(X_n \geq c_n \mathbb{E}(X_n)) \leq \frac{1}{c_n}$$

אם $\frac{1}{c_n} \rightarrow_{n\rightarrow\infty} 0$, נקבל שסדרת ההסתברויות הזו חסומה על ידי משהו שהולך ל0, ולכן כאשר $n \rightarrow \infty$, היא הולכת ל0 גם כן.

השימוש העיקרי עבורנו יהיה במקרה פרטי חשוב, כאשר X_n מקבל ערכים טבעיים והתוחלת שלו שואפת לאפס. במקרה כזה אי שיוויון מרקוב יוכל לומר לנו שמתקיים:

$$\mathbb{P}(X_n > 0) \leq \mathbb{E}[X_n] \longrightarrow_{n\rightarrow\infty} 0$$

כלומר, כשn נהיה מאוד גדול, ההסתברות שX_n יהיה חיובי נהיית מאוד מאוד קטנה, ושואפת ל0.

ממה שכתוב למעלה ניתן לראות שכשהתוחלת מאוד קטנה, היא מחייבת את המשתנה המקרי האי שלילי להיות קטן גם כן (בהסתברות ששואפת ל1). נשאלת השאלה - האם ההיפך נכון? כלומר, האם תוחלת גדולה יכולה להבטיח לנו שהמשתנה המקרי מקבל ערכים גבוהים בהסתברות ששואפת ל1? והתשובה המפתיעה אולי היא שלא. קונספטואלית אפשר לחשוב על מדינה שבה רוב מכריע של האנשים הוא עני מאוד ומרוויח משכורת שקרובה לאפס, אבל יש כמה אנשים עשירים באופן קיצוני, עד כדי כך שהם משפיעים על התוחלת של המשכורת הממוצעת להיות ערך גבוהה, למרות שההסתברות לקבל משכורת נמוכה היא כמעט 1.

נראה דוגמא ממשית לעניין: יהי X משתנה מקרי שתלוי בערך $\epsilon > 0$. X מקבל $\frac{1}{\epsilon^2}$ בהסתברות ϵ, ו0 בהסתברות $1-\epsilon$. התוחלת של X תהיה $\mathbb{E}[X] = 0 + \epsilon \cdot \frac{1}{\epsilon^2} = \frac{1}{\epsilon}$, וכאשר ϵ שואף ל0 ברור ש$\mathbb{E}[X]$ שואפת לאינסוף. אבל מצד שני, כאשר ϵ שואף לאפס, $1-\epsilon$ שואף ל1, ולכן X מקבל את הערך 0 בהסתברות ששואפת ל1.

כלומר, תוחלת גדולה כשלעצמה לא מאפשרת להבטיח שהמשתנה המקרי עצמו יהיה גדול. אך מסתבר שבשילוב עם שונות קטנה, כן אפשר יהיה להבטיח את זה. רעיונית, כאשר השונות קטנה אז X בהסתברות גבוהה קרוב לתוחלת שלו, והתוחלת שלו מאוד גדולה. למשל בדוגמא שלמעלה, השונות תהיה $Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{1}{\epsilon^3} - \frac{1}{\epsilon^2}$, וכאשר ϵ הולך ל0, $\frac{1}{\epsilon^3} \gg \frac{1}{\epsilon^2}$ והשונות תהיה ענקית.

שוב נתעניין בעיקר במקרה פרטי חשוב, בו X הוא משתנה מקרי שמקבל ערכים טבעיים ו0.

כשהתוחלת שואפת לאינסוף, נשתמש **בשיטת המומנט השני**. שיטת המומנט השני מסתמכת על המסקנה הבאה ממשפט צ'בישב:

מסקנה אי שיוויון צ'בישב: יהי X משתנה מקרי ממשי עם תוחלת שונה מאפס. אז $\mathbb{P}(X=0) \leq \frac{Var X}{\mathbb{E}[X]^2}$.

ההוכחה היא כמעט מיידית: $\mathbb{P}(X=0) \leq \mathbb{P}(X - \mathbb{E}[X] = \mathbb{E}[X]) \leq \mathbb{P}(|X - \mathbb{E}[X]| \geq \mathbb{E}[X]) \leq \frac{Var(X)}{\mathbb{E}[X]^2}$ ■

המשמעות היא שאם השונות קטנה יחסית לתוחלת, כלומר $Var(X_n) = o(\mathbb{E}[X_n]^2)$, וגם התוחלת של X_nשואפת לאינסוף, נקבל $\frac{Var(X_n)}{\mathbb{E}[X_n]^2} \longrightarrow 0$, ולכן ההסתברות ש$X_n = 0$ חסומה על ידי משהו שהולך לאפס, ולכן היא מאוד נמוכה.

1.1.1 דוגמא - הסף להופעת K_4:

בהינתן n,p, נציג את מרחב ההסתברות $G(n,p)$: מרחב זה מכיל גרפים עם n קודקודים. הדגימה של גרף נעשית באופן הבא: לכל זוג קודקודים, בהסתברות p שמים צלע ביניהם, ובהסתברות $1-p$ לא מחברים אותם בצלע. הצלעות מוגרלות באופן בלתי תלוי אחת מהשנייה. ההסתברות להגריל גרף מסויים $G=([n],E)$ היא:

$$\mathbb{P}(G)=p^{|E|}\cdot(1-p)^{\binom{n}{2}-|E|}$$

לכל $\{x,y\}\in E$ אנחנו צריכים להגריל צלע. לכן יש לנו $p^{|E|}$ צלעות להגריל, ובכל שאר הפעמים נצטרך להגריל שלא תהיה צלע. מספר הצלעות האפשרי סך הכל הוא $\binom{n}{2}$, ולכן נקבל את ההסתברות שמופיעה למעלה (נשים לב שזה הסתברות "עם חשיבות לסדר", כיוון שמאוד אכפת לנו בין איזה קודקודים יש צלע ובין איזה אין).

אינטואיטיבית, מובן שאם p מאוד קטן הגרף יראה יחסית ריק, בעוד שאם p מאוד קרוב ל1 נקבל הרבה מאוד צלעות.

נניח שp נקבע כפונקציה של n, כלומר $p=p(n)$, ונשאל עבור איזה פונקציה $p(n)$ תהיה הסתברות מאוד גבוהה להופע של K_4 כתת גרף מגרף הנדגם מהמרחב $G(n,p)$, כאשר n שואף לאינסוף. בתור ידע כללי, הצורה של ההסתברות להופעתו של K_4 כתלות בפונקציה $p(n)$ יהיה מאוד חד: תהיה "הסתברות קריטית" שעד אליה נקבל הסתברות מאוד נמוכה להתרחשות המאורע, והחל ממנה נקבל הסתברות מאוד גבוהה. אך זה לא הכרחי על מנת לשאול את מה שאנחנו רוצים לשאול, ולמעשה אנחנו יודעים שהp הקריטי הזה הולך להיות $n^{\frac{-2}{3}}$. נסמן את המאורע "הגרף מכיל את K_4" בB_n.

משפט 1: אם $p\ll n^{-(\frac{2}{3})}$, אז $lim_{n\to\infty}\mathbb{P}(B_n)=0$. כלומר, אם p קטן בסדר גודל מ$n^{-(\frac{2}{3})}$, אז ההסתברות להופעה של K_4 תשאף ל0 כאשר n שואף לאינסוף.

הוכחה: נסמן בX את המשתנה המקרי שמקבל את **מספר** הגרפים K_4 שמופיעים בגרף שנבחר.אנחנו מנסים להוכיח שההסתברות שהגרף מכיל את K_4 היא קטנה מאוד, כלומר נרצה להראות שבהסתברות מאוד גבוהה, $X=0 \Leftarrow$ נרצה להשתמש בשיטת המומנט הראשון, ונצפה לכך שהתוחלת של X תשאף לאפס.

כרגיל נרצה לפרק את X לסכום של משתנים מקריים שקל יותר לתאר: עבור $T\subseteq[n]$, $|T|=4$, נגדיר את המשתנה המקרי X_T להיות המשתנה המקרי האינדיקטור לכך שקודקודי T מהווים תת גרף של K_4 בגרף הנבחר. לכן $X=\Sigma_T X_T$, כאשר T עובר על כל תת הקבוצות הללו.

מליניאריות התוחלת:

$$\mathbb{E}[X]=\sum_T\mathbb{E}[X_T]$$

וכמשתנה אינדיקטור, התוחלת של X_T היא ההסתברות של המאורע. בגרף K_4 יש $\binom{4}{2}=6$ צלעות, ולכן $\mathbb{P}(X_T=1)=p^6$. כמות תת הקבוצות בגודל 4 של $[n]$ הוא כמובן $\binom{n}{4}$, ולכן נקבל סך הכל:

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

אם $p=o(n^{-\frac{2}{3}})$, אז $p^6=o(n^{-\frac{2\cdot 6}{3}})=o(n^{-\frac{12}{3}})=o(n^{-4})$. כמו כן, $\binom{n}{4}$ חסום על ידי n^4, לכן סך הכל נקבל:

$$\mathbb{E}[X]\leq n^4p^6=o(n^4\cdot n^{-4})=o(1)$$

כלומר, $\mathbb{E}[X]$ שואפת ל0, ולכן על ידי שימוש בשיטת המומנט הראשון נוכל להסיק שמתקיים:

$$lim\mathbb{P}(X>0)=0$$

■

משפט 2: אם $p \gg n^{-\frac{2}{3}}$, אז $lim_{n\to\infty}\mathbb{P}(B_n) = 1$. כלומר, אם p גדול בסדר גודל מהערך המיוחד של n, ההסתברות להופעה של K_4 תשאף ל1 כאשר n שואף לאינסוף.

הוכחה: נשתמש שוב במשתנה המקרי X, ובחלוקה שלו לסכום של X_T. כדי להראות שבהסתברות מאוד גבוהה יש K_4 בגרף, נרצה להראות שההסתברות ש$X = 0$ היא מאוד קטנה ⇐ נשתמש בשיטת המומנטים השנייה.

אכן הפעם בהנחות המשפט, $p = \omega(n^{-\frac{2}{3}})$, וכיוון ש$\binom{n}{4} = \Theta(n^4)$, נקבל

$$\mathbb{E}[X] \approx n^4p^6 = \omega(n^4 \cdot n^{-4}) = \omega(1)$$

כלומר, התוחלת שואפת לאינסוף. נרצה להשתמש באי השיוויון שהתקבל כמסקנה ממשפט צ'בישב:

$$\mathbb{E}[X] \approx n^4p^6 \Rightarrow \mathbb{E}[X]^2 \approx n^8p^{12}$$

נצטרך לחשב את $Var(X)$. נשתמש במשפט עבור שונות של סכום משתנים מקריים, אך נשים לב שהמשתנים X_T חלקם תלויים, ולכן לא נוכל לבטל מיידית את כל השונויות המשותפות:

$$Var(X) = \sum_T Var(X_T) + 2\sum_{S\neq T} Cov(X_s, X_T)$$

חשוב מאוד להבין שאנחנו לא מחפשים את הערך המדוייק של השונות המשותפת, כל המטרה שלנו היא להראות שהיא זניחה יחסית לתוחלת. לכן לא יהיה לנו אכפת "להיפתר" מכל מני גורמים על מנת לפשט את המשוואה - אם נראה שערך גדול יותר מהשונות הוא זניח יחסית לתוחלת, בוודאי נראה עבור השונות עצמה גם כן.

תחילה נשים לב שאם S וT זרים או חולקים קודקוד יחיד, אין להם צלעות משותפות ולכן הם אכן בלתי תלויים והשונות המשותפת שלהם מתאפסת. לכן:

$$Var(X) = \sum_T Var(X_T) + 2\sum_{|S\cap T|=2} Cov(X_s, X_T) + 2\sum_{|S\cap T|=3} Cov(X_s, X_T)$$

התוחלת של X_T היא כאמור p^6, והתוחלת של משתנה אינדיקטור בריבוע היא עדיין אותה התוחלת, ולכן נקבל

$$Var(X_T) = \mathbb{E}[X_T^2] - \mathbb{E}[X]^2 = p^6 - p^{12} \leq p^6$$

אמור יש $\binom{n}{4}$ משתני X_T, ולכן הסיגמא הראשונה בסכום חסומה על ידי $\binom{n}{4}p^6 = \mathbb{E}[X]$.

בהינתן S, T, לפי ההגדרה $Cov(X_s, X_T) = \mathbb{E}[X_sX_T] - \mathbb{E}[X_s]\cdot\mathbb{E}[X_T]$, לכן אם ניפתר מ$\mathbb{E}[X_s]\mathbb{E}[X_T]$ רק הגדלנו את הביטוי. לכן נותר לנו לחסום את $\mathbb{E}[X_sX_T]$ בשני המקרים.

המכפלה של משתנים מקריים אינדיקטורים היא המשתנה המקרי שהוא אינדיקטור לחיתוך שלהם. לכן בהינתן שתי קבוצות S, T שהחיתוך ביניהם הוא 2, כלומר יש 6 קודקודים סך הכל, נבדוק מה ההסתברות שיחד 6 הקודקודים האלה יוצרים שני גרפים של K_4. יש $\binom{6}{4}$ דרכים לבחור את הקודקודים המשותפים. בכל גרף K_4 בנפרד יש 6 צלעות, ויש להם שני קודקודים משותפים לכן יש צלע אחת משותפת ⇐ סך הכל, בהינתן שני הקודקודים במשותפים יש צורך ב$6 + 6 - 1 = 11$ צלעות. מכאן שמתקיים:

$$\mathbb{E}[X_sX_T] = \binom{6}{4}p^{11}$$

מספר הזוגות של S, T כאלו הוא כמספר הדרכים לבחור שישייה של קודקודים, ולכן הסיגמא השנייה חסומה על ידי $\binom{n}{6}\binom{6}{4}p^{11} = \Theta(n^6p^{11})$ (כמובן ש$\binom{6}{4}$ הוא קבוע ולא משפיע על הסכימה.

הניתוח זהה עבור הסיגמא השלישית, אלא שבמקרה הזה יש לנו 3 קודקודים משותפים, לכן סך הכל יש 5 קודקודים. בין 3 הקודקודים יש 3 צלעות משותפות, לכן מספר הצלעות שנצטרך יהיה $6 + 6 - 3 = 9$, ולכן התוחלת של המכפלה תהיה:

$$\mathbb{E}[X_sX_T] = \binom{5}{3}p^9$$

ויש לנו $\binom{n}{5}$ דרכים לבחור תת קבוצה של 5 קודקודים כזו, ולכן הסיגמא השלישית חסומה על ידי $\binom{n}{5}\binom{5}{3}p^9 = \Theta(n^5p^9)$.

עד כה, הצלחנו לחסום את השונות על ידי:

$$Var(X) \leq \mathbb{E}[X] + \Theta(n^6 p^{11}) + \Theta(n^5 p^9)$$

נשים לב שהתעלמנו לחלוטין מה2 שמופיע לפני שני הסיגמאות השניות, כיוון שאנחנו מעוניינים רק בגודל האסימפטוטי: מבחינתינו היה אפשר להכפיל את כל הביטוי בקבוע וזה עדיין לא היה אכפת לנו.

כעת נצטרך להראות שיתקיים $\frac{Var(X)}{\mathbb{E}[X]^2} \to 0$. $\mathbb{E}[X]$ בוודאי זניחה יחסית ל$\mathbb{E}[X]^2 = \Theta(n^8 p^{12})$. נחסום גם את שני הביטויים הבאים:
$p = \omega(n^{-\frac{2}{3}})$, ולכן בפרט $p = \omega(n^{-2})$. לכן: $n^8 p^{12} = n^8 \cdot p^{11} \cdot p \gg n^8 p^{11} n^{-2} = n^6 p^{11}$, וכך חסמנו את הביטוי השני.
באותו אופן, $p = \omega(n^{-1})$, לכן $p^3 = \omega(n^{-3})$ ולכן: $n^8 p^{12} = n^8 p^9 \cdot p^3 \gg n^8 p^9 n^{-3} = n^5 p^9$, וכך חסמנו גם את הביטוי השלישי.

סך הכל, קיבלנו שהתוחלת $Var(X)$ חסומה על ידי סכום של שלושה ביטויים שכל אחד מהם זניח יחסית ל$\mathbb{E}[X]^2$, ולכן נוכל להציב במסקנה מצ'בישב ולקבל שההסתברות ש$X = 0$ חסומה על ידי משהו ששואף ל0. כלומר, בהסתברות מאוד גבוהה $X > 0$, כלומר קיים גרף K_4 כלשהו ■

1.2 כדורים וכדים

עולם הבעיה שלנו היא כזה: נתונים n כדים, ומשליכים באקראי כדורים לתוך הכדים. כל השלכת כדור נוחתת בתוך כד כלשהו, באופן בלתי תלוי במה שהיה בכד לפני כן ובכל הזריקות האחרות. יהיו כל מני שאלות שיעניינו אותנו בהקשר של מודל הבעיה הזו: למשל, לאחר כמה זריקות נקבל שיש יותר מכדור אחד בכד כלשהו? מתי כל הכדים יכילו כדור? וכך הלאה. מסתבר שהמודל הזה של "כדורים וכדים" מתאים להרבה עולמות בעיה אמיתיים ובעלי משמעות.

הערה: באופן כללי חשוב לדעת שלכל $x \in \mathbb{R}$, $e^x \geq 1 + x$, והם מאוד קרובים כשx קרוב ל0.

1.2.1 דוגמא: פרדוקס יום ההולדת:

נניח שיש לנו חדר עם אנשים. נרצה לשאול - כמה אנשים צריכים להיות בחדר על מנת שנוכל להבטיח בהסתברות טובה שלשניים מהם יש יום הולדת באותו היום? התשובה המפתיע היא שבכלל לא הרבה: כיוון שיש 365 ימים בשנה, ברור שאם יש 366 אנשים בחדר אז מעקרון שובך היונים יהיו שניים עם אותו התאריך. אך מסתבר שכבר החל מ23 אנשים, סביר יותר שיש שניים עם אותו תאריך משלא.

בגרסא לכדורים וכדים, נניח שמשליכים r כדורים לתוך n כדים. מה ההסתברות שאין התנגשות, כלומר שכל כדור נמצא בכד אחר?
כל כדור נזרק באופן בלתי תלוי באחרים כאמור. בזריקת הכדור הראשון זה בוודאות קורה. עבור הכדור השני, יש כד אחד תפוס, ויש לו סיכוי של $\frac{1}{n}$ לנחות בו, ולכן סיכוי של $1 - \frac{1}{n}$ לנחות בכד אחר. כך עבור הכדור הk, נזרקו לפניו $(k-1)$ כדורים, והוא צריך לנחות בכד אחר מכל אחד מהם, ולכן ההסתברות שלו לעשות כך היא $1 - \frac{k-1}{n}$. סך הכל נקבל שההסתברות של המאורע שלא היתה התנגשות בכלל היא:

$$1 \cdot (1 - \frac{1}{n}) \cdot ... \cdot (1 - \frac{r-1}{n})$$

כעת, $(1 - \frac{x}{n}) \approx e^{-\frac{x}{n}}$, ולכן:

$$\approx e^{-(\frac{1}{n} + ... + \frac{r-1}{n})} = exp(-\frac{1 + ... + r - 1}{n})$$

נפרק לפי סכום סדרה חשבונית, ונקבל:

$$= exp(-\frac{r(r-1)}{2n})$$

כלומר, ההסתברות היא בערך $exp(-\frac{r(r-1)}{2n})$.

ננתח את המשמעויות עבורנו: ההסתברות היא מהצורה $e^{f(r,n)}$. כעת, אם $f(r,n)$ מאוד קרוב ל0, הביטוי יהיה מאוד קרוב ל1. השבר יהיה קרוב ל0 כאשר $r(r+1) \approx r^2 = o(n)$. לכן כאשר r מאוד קטן ביחס ל$\sqrt{n}$, נקבל שההסתברות שלא תהיה התנגשות היא גבוהה. באותו אופן בדיוק, אם $f(r,n)$ הולך ל$-\infty$, אז הביטוי יהיה מאוד קרוב ל0. זה יקרה כאשר $r(r+1) = \omega(n)$, כלומר המונה גדל יותר מהר מהמכנה, כלומר כאשר r מאוד גדול ביחס ל$\sqrt{n}$.

שאלה אחרת באותו כיוון: יש לנו עולם שבו יש n אנשים, ונסמן באקראי אנשים לקבוצה A וB באקראי, כאשר $|A| = a$ ו$|B| = b$ (מישהו יכול להיות שייך לשתי הקבוצות). נשאל מה ההסתברות ש$A \cap B = \phi$.
נוכל להתיחס לA כנתון, וברגע שיש לנו את A אנחנו צריכים לחשב את הסיכוי שלא בחרנו עבור B אף אחד שבטעות נמצא גם בA. ההסתברות תהיה בערך:

$$(1-\frac{a}{n})\cdot(1-\frac{a+1}{n})\cdot..\cdot(1-\frac{a+b-1}{n}) \approx exp(-\frac{a+a+1+...+a+b-1}{n}) = exp(-\Theta(\frac{ab}{n}))$$

ולכן כמו קודם, אם $ab \gg n$, אז ההסתברות תהיה מאוד קטנה, ואם $ab \ll n$, ההסתברות תהיה מאוד גדולה.

1.2.2 אספן הקופונים

נניח שיש לנו מישהו שאוסף קופונים, שהוא מקבל למשל בקופסאות של דגנים. נניח שיש n סוגים שונים של קופונים שהאספן יכול לקבל ־ וכמובן שיתכן והוא יקבל קופון מאותו הסוג מספר פעמים. האספן ירצה לדעת כמה חבילות של דגנים הוא יצטרך לקנות לפני שבהסתברות טובה יהיו לו את כל הקופונים.

בגרסא לכדורים וכדים, אנחנו זורקים כדורים לתוך n כדים שוב ושוב, ואנחנו שואלים מתי בהסתברות טובה כל הכדים יהיו מלאים.
המודל שנבחר כדי למדל את הבעיה באופן הפשוט ביותר יהיה לשאול כמה כדורים היינו צריכים להשליך עד שמילאנו עוד כד: נגדיר משתנה מקרי T שיקבל את מספר הזריקות שהיה צריך כדי למלא את כל הכדים. נפשט את המשתנה המקרי: לכל $0 \le k \le n-1$, נגדיר משתנה מקרי T_k, שמקבל את מספר השלכות הכודרים שנצרכו מהרגע שהתמלאו k כדים עד שמילאנו $k+1$ כדים. כלומר, בהנחה שברגע כלשהו בדיוק התמלאו k כדים, ואנחנו ממשיכים לזרוק כדורים, כמה כדורים נצטרך לזרוק עד שנצליח למלא עוד כד.
נמצא את התוחלת של T_k. נשים לב שT_k הוא משתנה מקרי גאומטרי: נזרוק כדורים שוב ושוב עד שנצליח למלא כד נוסף. בהינתן k כדים מלאים, ההסתברות לקלוע לכד חדש ובכך למלא עוד כד היא $\frac{n-k}{n}$, כלומר יש לנו משתנה מקרי גאומטרי עם סיכויי הצלחה $\frac{n-k}{n}$. לכן לפי המשפט לגבי תוחלת של משתנה מקרי גאומטרי,

$$\mathbb{E}[T_k] = \frac{n}{n-k}$$

כעת, $T = \Sigma_{k=0}^{n-1} T_k$, ולכן מליניאריות התוחלת,

$$\mathbb{E}[T] = \Sigma_{k=0}^{n-1}\frac{n}{n-k} = n\Sigma_{k=0}^{n-1}\frac{1}{n-k} = n\cdot\Sigma_{j=1}^{n}\frac{1}{j}$$

כעת, לפי הנוסחה של טור הרמוני:

$$\le n(ln(n)+\Theta(1))$$

כלומר, אנחנו יודעים שהתוחלת של מספר הזריקות שנצטרך היא בערך $n(lnn+\Theta(1))$.

כמובן שזה עדיין לא ממצא את מה שהיינו רוצים לדעת ־ היינו רוצים לחסום את הסיכוי שהמשתנה המקרי T יפול רחוק מהתוחלת. למשל, אם נשאל מה הסיכוי שיקח לנו יותר מפי 10 זריקות מהתוחלת, נקבל לפי מרקוב:

$$\mathbb{P}(T > 10nlnn) \ge \frac{1}{10}$$

(ויתרנו על $\Theta(1)$ בתוחלת, כיוון ששוב זה קבוע ולא מעניין אותנו).

זה נותן לנו חסם אבל לא מאוד הדוק. צ'בישב מסוגל לתת לנו חסם הדוק יותר:

$$\mathbb{P}(|X-\mu| \geq c\sigma) \leq \frac{1}{c^2}$$

נחשב את השונות של T, ובכך נקבל את סטיית התקן שלו. חישוב השונות יהיה קל יחסית, כיוון שT הוא סכום של משתנים מקריים בלתי תלוים שמתפלגים גאומטרית, ואנחנו יודעים מה השונות של משתנה מקרי גאומטרי:

$$Var(T) = Var(\Sigma T_k) = \Sigma_{k=0}^{n-1} Var(T_k) \leq \Sigma_{k=0}^{n-1} (\frac{n}{n-k})^2 - n^2 \Sigma_{j=1}^{n} \frac{1}{j^2} \leq n^2 \cdot \Sigma_{j=1}^{\infty} \frac{1}{j^2} = n^2 \frac{\pi^2}{6}$$

כיוון שהסכום של הטור שאיברו הכללי הוא $\frac{1}{n^2}$ הוא $\frac{\pi^2}{6}$.
כעת נוכל לחשב מה ההסתברות שT יפול רחוק מהתוחלת שלו, באופן יותר מדוייק מה שמרקוב איפשר לנו.

שבוע 3

משתנה מקרי הוא כאמור סוג של פונקציה, $X:\Omega\to\mathbb{R}$, פונקציה ממרחב הסתברות Ω לתוך הממשיים (או תת קבוצה שלהם). כפי שראינו, נתעניין בשאלות כגון "מה ההסתברות שהמשתנה המקרי יקבל ערך מסויים", "איזה סוג של ערכים הוא מקבל 'בדרך כלל'" וכן הלאה.

בהינתן $c \in \mathbb{R}$, אפשר להסתכל על $\mathbb{P}(X \leq c)$. היא למעשה נותנת לנו פונקציה מהממשיים ל$[0,1]$. לפונקציה הזו קוראים **פונקציית ההתפלגות המצטברת** של המשתנה המקרי. נשים לב שהיא פונקציה מונוטונית עולה, שכן אם $c_1 \leq c_2$, המאורע $X \leq c_1$ מוכל במאורע $X \leq c_2$, ולכן $\mathbb{P}(X \leq c_1) \leq \mathbb{P}(X \leq c_2)$. פונקציית ההתפלגות המצטברת מכילה את כל המידע על המשתנה המקרי - במובן מסויים ברגע שיש לנו אותה, יש לנו את כל מה שיעניין אותנו בקשר למשתנה המקרי.

אך ברוב המקרים לא תהיה לנו אפשרות לתאר את פונקציית ההתפלגות המצטברת של המשתנה המקרי, ויהיה לנו מידע חלקי יותר לגביו. למשל, התוחלת והשונות נותנים לנו תחושה מסויימת ביחס להתנהגות של המשתנה המקרי. יש גם מושגים נוספים שמתארים את ההתנהגות של המשתנה המקרי שפחות נתעסק איתם: למשל **חציון** - החציון של המשתנה המקרי הוא ערך m כך שעבורו מתקיים:

$$\mathbb{P}(X \leq m) \geq \frac{1}{2}, \mathbb{P}(X \geq m) \geq \frac{1}{2}$$

כלומר ההסתברות ליפול מעליו וליפול מתחתיו הן שוות. לא ניכנס לפרטי ההגדרה של המושג הזה (למשל למה צריך גדול שווה חצי, ולא דורשים שיוויון לחצי), רק נציין שהוא קיים והוא עוד דרך לנתח תכונה של המשתנה המקרי.

כאשר המידע היחיד שיש לנו הוא התוחלת והשונות, היכולת שלנו לנתח את ההתנהגות של המשתנה המקרי היא מוגבלת בהתאם. במקרה כזה יהיה לנו את אי שיוויון מרקוב, ואי שיוויון צ'בישב, אך לא מעבר. לא נוכל לומר דברים מעניינים מעבר לכך. אך ברגע שנדע יותר על המשתנה המקרי, נוכל בהתאם גם לתת חסמים מדוייקים יותר על הערכים שהוא עשוי לקבל.

1.3 מהלך מקרי (חד מימדי)

נניח שיש לנו איש שמהלך על ציר המספרים. הוא מתחיל בראשית, ובכל צעד מטיל מטבע הוגן, לפיו הוא מחליט האם ללכת ימינה או שמאלה.
נרצה לשאול כל מני שאלות בקשר למיקום שלו:

- היכן הוא עשוי להימצא אחרי n צעדים?
- כמה הוא עשוי להתרחק מהראשית אחרי n צעדים?
- כמה פעמים צפוי שהוא יחזור לראשית בn הצעדים הראשונים?

וכן הלאה. מסתבר שהמודל הזה והשאלות לגבי המיקומים של האיש המהלך הוא חשוב כדי לתאר כל מני בעיות.

נמדל את הבעיה על ידי משתנים מקריים. השאלה היא מה קורה אחרי n צעדים, ובכל צעד האיש מחליט באופן בלתי תלוי בצעדים הקודמים לאיזה כיוון ללכת. הוא מתחיל בראשית, הנקודה 0. לכל צעד i, נגדיר משתנה מקרי X_i, שיקבל 1 אם הוא בחר ללכת ימינה, ו-1 אם הוא בחר ללכת שמאלה. באופן הזה נקבל אוסף של משתנים מקריים $X_1,..,X_n$ שהם בלתי תלוים ושווי התפלגות (נסמן זאת גם כמשתנים מקריים **IID - independent identically distributed**). נתאר גם משתנה מקרי $X=\sum_{i=1}^{n}X_i$, שהוא יקבל את המיקום אחרי n צעדים.

מסתבר שעבור אוסף של משתנים מקריים בלתי תלויים ושווי התפלגות כבר אפשר לומר דברים חזקים בהרבה ממה שאפשר לומר על משתנה מקרי שכל מה שידוע לגביו זה התוחלת והשונות. נתחיל לפתח כלים לכך.

תחילה נראה טענה טכנית:

טענה: לכל $t\in\mathbb{R}$,

$$\frac{e^t+e^{-t}}{2}\le e^{\frac{t^2}{2}}$$

הוכחה: נוכיח דרך טורי טיילור. נזכיר שלכל $x\in\mathbb{R}$, מתקיים:

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

ולכן:

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

$$e^{-t}=\sum_{n=0}^{\infty}\frac{(-t)^n}{n!}=1-t+\frac{t^2}{2!}-\frac{t^3}{3!}+\frac{t^4}{4!}-....$$

כעת, כשנסכום את הטורים, האיברים עם החזקה האי זוגית מתבטלים, ונישאר עם טור של האיברים בעלי החזקה הזוגית בלבד:

$$e^t+e^{-t}=1+\frac{t^2}{2!}+\frac{t^4}{4!}+...=\sum_{k=0}^{\infty}\frac{t^{2k}}{(2k)!}\Rightarrow\frac{e^t+e^{-t}}{2}=\sum_{k=0}^{\infty}\frac{t^{2k}}{2\cdot(2k)!}$$

באותו אופן, נסתכל על טור הטיילור של אגף ימין:

$$e^{\frac{t^2}{2}}=\sum_{k=0}^{\infty}\frac{(\frac{t^2}{2})^k}{k!}=\sum_{k=0}^{\infty}\frac{t^{2k}}{2^k\cdot k!}$$

כעת, נשווה בין האיבר הכללי של שני הטורים האלה. עבור $(2k)!$, מדובר ב$k!$ כפול עוד k איברים, שכל אחד מהם גדול שווה לk. בפרט אם נחליף כל אחד מהם ב2, נקבל ערך קטן יותר, ולכן:

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

כלומר כל איבר בטור השמאלי קטן מכל איבר בטור הימני, ולכן אי השיוויון עובר גם בין הטורים עצמם ■

1.3.1 מקדמים בינומיים

לפני שנמשיך להתעסק בשאלות של מהלך מקרי, ננתח מקדמים בינומיים. נרצה לדעת איך מתנהג הביטוי $\binom{n}{k}$. נציג כמה דרכים להעריך אותו.

הנוסחה הכללית היא:

$$\binom{n}{k} \sim (\frac{ne}{k})^k$$

כאשר $k \ll n$. כלומר, אם k קטן בסדר גודל מn, זו הנוסחה עליה נרצה להסתכל.

נרצה לשאול מה קורה כאשר k בסדר הגודל של n, כלומר מה קורה כאשר $k = \alpha n$.

כלי מאוד חשוב שנשתמש בו בהקשר הזה הוא **נוסחת סטירלינג**, שנותנת הערכה בקשר ל$n!$:

$$n! = (1 + O(n))(\frac{n}{e})^n \sqrt{2\pi n}$$

לא נוכיח את נוסחת סטירלינג, אבל כן ניתן לה אינטואיציה מסויימת. המשמעות הפרקטית כשמעניינת אותנו כמובן גדילה אסימפטוטית היא שמתקיים:

$$n! \sim (\frac{n}{e})^n \sqrt{2\pi n}$$

אינטואיציה לנוסחת סטירלינג: ננסה לתת קצת אינטואיציה כדי להבין מאיפה הגורמים בנוסחת סטירלינג הגיעו. בעצם לומר את נוסחת סטירלינג זה כמו לומר שמתקיים:

$$ln(n!) - [nln(n) - n + \frac{1}{2}ln(n) + \frac{1}{2}ln2\pi] \to 0$$

כאשר $n \to \infty$. נשים לב שמתקיים:

$$ln(n!) = ln(\Pi_{i=1}^{n} i) = \sum_{i=1}^{n} ln(i)$$

כפי שרואים בתרגיל, זה בערך כמו האינטגרל של lnx בין 1 לn. נחשב את האינטגרל:

$$\int_1^n lnxdx = xlnx - x|_1^n = nlnn - n$$

כלומר קיבלנו ש$ln(n!)$ זה בערך כמו $nlnn - n$, ואז נשאר הגורם $\frac{1}{2}lnn + \frac{1}{2}ln2\pi = \Theta(lnn)$.
זה מראה לנו מאיפה מגיע המחובר $nlnn - n$. כדי לראות מאיפה מגיע הביטוי שאותו סימנו ב$\Theta(lnn)$ נצטרך לעבור יותר קשה, ולא נראה זאת כאן.

בהינתן נוסחת סטירלינג, ניתן הערכה תחילה ל$k = \frac{1}{2}n$

$$\binom{n}{\frac{1}{2}n} = \frac{n!}{(\frac{n}{2})! \cdot (\frac{n}{2})!} \sim \frac{(\frac{n}{e})^n \sqrt{2\pi n}}{((\frac{n}{2e})^{\frac{n}{2}} \sqrt{2\pi \cdot \frac{n}{2}})^2} = \frac{(\frac{n}{e})^n \sqrt{2\pi n}}{(\frac{n}{2e})^n \cdot \pi n} = \frac{n^n \sqrt{2\pi} \cdot \sqrt{n} \cdot 2^n e^n}{e^n \cdot n^n \cdot \pi n} = \frac{\sqrt{2\pi} \cdot 2^n}{\pi \cdot \sqrt{n}} \sim \frac{2^n}{\sqrt{n}}$$

כיוון שלא אכפת לנו מקבועים, אנחנו רוצים לדעת מה המגמה, כלומר לאן הביטוי הולך כשn מאוד גדול.

דרך נוספת להעריך את $\binom{n}{\frac{n}{2}}$ היא לפי נוסחת הבינום:

$$2^n = (1+1)^n = \sum_{k=0}^{n} \binom{n}{k}$$

המקדם הבינומי הגדול ביותר שיתקבל כאן הוא האמצעי, כלומר $\binom{n}{\frac{n}{2}}$. כיוון שהסכום מכיל אוסף של מחוברים אי שליליים, ו$\binom{n}{\frac{n}{2}}$ הוא רק אחד מהם, נקבל $\binom{n}{\frac{n}{2}} < 2^n$. מצד שני, בתור המחובר הגדול ביותר הוא חייב להיות גדול יותר מהממוצע של הסכום, שזה $\frac{2^n}{n+1}$ - יש $n+1$ איברים שסכומם הוא 2^n. סך הכל:

$$\frac{2^n}{n+1} < \binom{n}{\frac{n}{2}} < 2^n$$

וזה נותן לנו הערכה נוספת לגבי ערכו של $\binom{n}{\frac{n}{2}}$.

נחזור לנסיון להבין את $k = \alpha n$. כיוון ש$k \leq n$, נוכל להסתכל רק על $0 < \alpha < 1$. כמו כן, כיוון ש$\binom{n}{k} = \binom{n}{n-k}$, מתקיים: $\binom{n}{\alpha n} = \binom{n}{(1-\alpha)n}$ ולכן מספיק להסתכל על $0 < \alpha < \frac{1}{2}$.
נציב בסטירלינג:

$$\binom{n}{\alpha n} = \frac{n!}{(\alpha n)! \cdot [(1-\alpha)n]!} \sim \frac{(\frac{n}{e})^n \sqrt{2\pi n}}{(\frac{\alpha n}{e})^{\alpha n}(\frac{(1-\alpha)n}{e})^{(1-\alpha)n} \cdot \sqrt{2\pi\alpha n} \cdot \sqrt{2\pi(1-\alpha)n}}$$

$$= \frac{n^n e^{\alpha n} \cdot e^{(1-\alpha)n} \cdot \sqrt{2\pi n}}{\alpha^{\alpha n} \cdot n^{\alpha n} \cdot e^n \cdot (1-\alpha)^{(1-\alpha)n} \cdot n^{(1-\alpha)n} \cdot \sqrt{2\pi n} \cdot \sqrt{\alpha} \cdot \sqrt{2\pi n} \cdot \sqrt{(1-\alpha)}}$$

$$= \frac{n^n \cdot e^n}{(\alpha^\alpha \cdot (1-\alpha)^{1-\alpha})^n \sqrt{\alpha(1-\alpha)} n^{\alpha n + n - \alpha n} \sqrt{2\pi n} \cdot e^n} = \frac{1}{(\alpha^\alpha \cdot (1-\alpha)^{1-\alpha})^n \sqrt{2\pi n} \cdot \sqrt{\alpha(1-\alpha)}}$$

והגורם הדומיננטי ביותר הוא הגורם האקספוננציאלי במכנה, $(\alpha^\alpha \cdot (1-\alpha)^{1-\alpha})^n$ - ככל שn גדל, הוא גדל הכי מהר ומשפיע הכי הרבה על הביטוי.

נפתח את הביטוי ונקבל:

$$2^{-log_2[(\alpha^\alpha \cdot (1-\alpha)^{1-\alpha})^n \cdot \sqrt{2\pi n} \cdot \sqrt{\alpha(1-\alpha)}]} \sim 2^{-nlog_2(\alpha^\alpha) - nlog_2(1-\alpha)^{1-\alpha}} = 2^{n[-\alpha log_2(\alpha) - (1-\alpha)log_2(1-\alpha)]}$$

מסתבר שלפונקציה שמופיעה בתוך הסוגריים המרובעים יש משמעות מיוחדת, היא פונקציה שמופיעה בפעמים רבות ונקראת **פונקציית האנטרופיה**, המוגדרת כ:

$$H(x) = -xlog_2 x - (1-x)log_2(1-x)$$

כלומר, סך הכל קיבלנו:כ

$$\binom{n}{\alpha n} = 2^{nH(\alpha) + O(logn)}$$

נחזור לנושא של מהלכים מקריים:

ניתן לדבר גם על מהלכים מקריים עם יותר מימדים. למשל, היה אפשר לתאר מהלך בשני מימדים: במהלך מקרי בשני מימדים האיש זורק שני מטבעות, אחד מהם קובע אם ללכת צפון או דרום, והשני קובע האם ללכת מזרח או מערב. באותו אופן אפשר לתאר גם מהלך בשלושה מימדים: כשמתווסף מטבע ל"למעלה למטה".

נשאל כמה רחוק אפשר לצפות שהאיש יגיע אחרי n צעדים. ברור שהוא לא יכול להתרחק יותר מn מהראשית בכל כיוון, אך נשאלת השאלה מה הטווח ה"סביר" שהוא ימצא בו. מסתבר שהטווח הסביר שהוא ימצא בו הוא בטווח $(-\sqrt{N}, \sqrt{N})$ מהראשית, וההסתברות שהוא יסיים את המהלך בראשית היא בערך $\frac{1}{\sqrt{N}}$.

אפשר להגיע מכך לניתוח שבמהלך מקרי אינסופי על הישר הממשי, האיש יעבור אינסוף פעמים בראשית, כיוון שהטור שמציין את תוחלת מספר הביקורים בראשית הוא:

$$\sum_{N=1}^{\infty} \frac{1}{\sqrt{N}} = \infty$$

באופן דומה, גם מהלך מקרי במישור האיש יעבור אינסוף פעמים בראשית, כיוון שהטור שמציין את תוחלת מספר הביקורים שלו יהיה:

$$\sum_{N=1}^{\infty} \frac{1}{\sqrt{N} \cdot \sqrt{N}} = \sum_{N=1}^{\infty} \frac{1}{N} = \infty$$

(כיוון שהוא זורק שני מטבעות, בלתי תלויים אחד בשני, וההסתברות שלו לחזור לראשית שקולה לכך שהוא חוזר לראשית בשני המימדים).

אבל מסתבר שכשמסתכלים על מהלך מקרי ב3 מימדים, כבר לא צפוי שהוא יעבור אינסוף פעמים בראשית, כיוון שהטור המתאים במקרה הזה הוא:

$$\sum_{N=1}^{\infty} \frac{1}{N\sqrt{N}}$$

וזה טור מתכנס. נציין שמהלך שבו חוזרים בתוחלת אינסוף פעמים לראשית נקרא **מהלך שיכור.**

כעת נפתח כלים שיאפשרו לנו לומר דברים על המהלך המקרי.
טענה: יהיו $X_1, .., X_n$ אוסף משתנים מקריים שמקבלים 1 או -1 בהסתברות חצי, בלתי תלויים, ו$X = \sum_{i=1}^{n} X_i$. אז לכל $a > 0$ מתקיים:

$$\mathbb{P}(X \geq a) \leq e^{-\frac{a^2}{2n}}$$

הטענה הזו תאפשר לנו לחסום את ההסתברות שמספר הצעדים שהוא התרחק מהראשית גדול מa.

בהנחה שהטענה נכונה, נקבל שההסתברות קטנה כאשר החזקה של e בערך מוחלט גדולה (ולכן הולכת ל$-\infty$), כלומר כאשר $\frac{a^2}{2n}$ גדול - כלומר כאשר a גדול יחסית ל$\sqrt{n}$. מכאן נסיק שהתחום הסביר שימצא ההולך הוא במרחק של $\sqrt{n}$.

הוכחה: יהי $t > 0$ כלשהו, נשתמש בו בהמשך כדי למצוא את החזקה האופטימלית עבור e. כיוון שהפונקציה e^x היא מונוטונית, נקבל ש$X \geq a$ אם ורק אם $e^X \geq e^a$, ולכן:

$$\mathbb{P}(X \geq a) = \mathbb{P}(tX \geq ta) = \mathbb{P}(e^{tX} \geq e^{ta}) \leq \frac{\mathbb{E}[e^{tX}]}{e^{ta}}$$

לפי אי שיוויון מרקוב. נחשב את התוחלת:

$$\mathbb{E}[e^{tX}] = \mathbb{E}[e^{t\sum_{i=1}^{n} X_i}] = \mathbb{E}[\Pi_{i=1}^{n} e^{tX_i}]$$

כיוון שכל המשתנים המקריים בלתי תלויים, תוחלת המכפלה שווה למכפלת התוחלות, ולכן:

$$= \Pi_{i=1}^{n} \mathbb{E}[e^{tX_i}]$$

לכל i, מתקיים:

$$\mathbb{E}[e^{tX_i}] = \frac{e^{t\cdot 1} + e^{t\cdot(-1)}}{2} = \frac{e^t + e^{-t}}{2}$$

לפי הגדרת התוחלת של פונקציה של משתנה מקרי. מכאן שמתקיים:

$$\mathbb{E}[e^{tX}] = (\frac{e^t + e^{-t}}{2})^n$$

נציב את אי השיוויון הטכני שראינו בהתחלה, ונקבל:

$$\mathbb{E}[e^{tX}] = (\frac{e^t + e^{-t}}{2})^n \leq e^{\frac{t^2 n}{2}}$$

עד כה קיבלנו את החסם הבא:

$$\mathbb{P}(X \geq a) \leq \frac{e^{\frac{t^2 n}{2}}}{e^{ta}} = e^{\frac{nt^2}{2} - ta}$$

כאשר t היה פרמטר כלשהו שנקבע כרצוננו. נרצה למצוא מהו הt שיגרום לביטוי להיות מינימלי - ובכך לתת לנו את החסם ההדוק ביותר על ההסתברות. כדי לעשות זאת, נתיחס לפונקציה $f(t) = \frac{t^2}{2}n - ta$, ונשאל מתי יש לה נקודת מינימום. $f'(t) = 0 \iff tn - a = 0 \Leftarrow t = \frac{a}{n}$. נציב באי השיוויון ונקבל:

$$\mathbb{P}(X \geq a) \leq e^{\frac{a^2}{2n} - \frac{a^2}{n}} = e^{\frac{-a^2}{2n}}$$

כנדרש ■

מסקנה: אם $a \gg \sqrt{n}$, אז ההסתברות שהמהלך המקרי באורך n יסתיים במקום גדול מa או קטן מ$-a$ היא אפסית, כיוון שהביטוי $\frac{-a^2}{2n}$ שואף ל$-\infty$ ולכן ההסתברות חסומה על ידי משהו שהולך לאפס. במילים אחרות, כמעט בוודאות המהלך המקרי מסתיים בטווח $(-\sqrt{n}, \sqrt{n})$ של הראשית.

טענה: ההסתברות לסיים בדיוק הראשית היא בערך $\frac{1}{\sqrt{N}}$, במהלך מקרי של N צעדים.

הוכחה: באופן כללי, בהינתן t כלשהו, על מנת שהמהלך יסיים בדיוק במקום הt הוא צריך לבצע $\frac{N+t}{2}$ צעדים ימינה, ו$\frac{N-t}{2}$ צעדים שמאלה (יש כאן פשטנות מסויימת - עקרונת צריך להתיחס גם לעניין הזוגיות של הצעדים, המספרים האלה חייבים להיות שלמים וכן הלאה, אבל נניח שהכל מסתדר כדי לא להסתבך עם הפרטים). לכן:

$$\mathbb{P}(X = t) = \frac{1}{2^N}\binom{N}{\frac{N-t}{2}}$$

כיוון שההסתברות של כל משתנה מקרי X_i לקבל -1 היא $\frac{1}{2}$, ויש לנו $\binom{N}{\frac{N-t}{2}}$ משתנים מקריים, והם בלתי תלוים לפי ההגדרה.

בפרט אם נרצה לסיים בראשית, ההסתברות תהיה:

$$\mathbb{P}(X = 0) = \frac{1}{2^N} \cdot \binom{N}{\frac{1}{2}N}$$

מקודם הערכנו:

$$\binom{N}{\frac{1}{2}N} \sim \frac{2^N}{\sqrt{n}}$$

ולכן סך הכל:

$$\mathbb{P}(X = 0) = \Theta(\frac{1}{\sqrt{N}})$$

■

מסתבר שהתוצאה שקיבלנו, בה המשתנה המקרי נמצא בטווח $(-\sqrt{N},\sqrt{N})$, היא מקרה פרטי של תוצאה של משפט חשוב הנקרא משפט הגבול המרכזי.

תחילה נזכיר מהסתברות מהי **התפלגות נורמלית**. משתנה מקרי המתפלג נורמלית $X\sim N(\mu,\sigma^2)$ הוא משתנה מקרי רציף, כך שלכל $a\in\mathbb{R}$ מתקיים:

$$\mathbb{P}(X\leq a)=\int_{-\infty}^{a}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}dx$$

כאשר המשמעות היא שהתוחלת של X היא μ, והשונות היא σ^2. ההתפלגות של משתנה מקרי נורמלית היא ההתפלגות בצורת הפעמון המוכרת.

משפט הגבול המרכזי: יהיו $X_1,..,X_n$ אוסף משתנים מקריים בלתי תלויים ושווי התפלגות, עם תוחלת μ ושונות σ^2. נסמן: $Z_n=\frac{\sum_{i=1}^{n}(X_i-\mu)}{\sqrt{n}\cdot\sigma}$, ואז מתקיים:

$$lim_{n\to\infty}\mathbb{P}(Z_n\leq a)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a}e^{-\frac{t^2}{2}dt}$$

כלומר, כאשר $n\to\infty$, ההתפלגות של המשתנה המקרי Z_n נראית יותר ויותר דומה להתפלגות הנורמלית $N(0,1)$.

הערה: הגרסא הפשוטה יותר של משפט הגבול המרכזי מתיחסת לסכומים של משתנים מקריים עם תוחלת 0 ושונות 1. המעבר מX_iים לZ_n הוא הסכום אחרי "תיקון" של התוחלת להיות 0 והשונות להיות 1.

לא נוכיח את משפט הגבול המרכזי, אך נשים לב של שמה שמעניין בו הוא שבסופו של דבר יש כאן אוסף די כללי של משתנים מקריים: היה ידוע לנו שהם בלתי תלויים, שווי התפלגות, וידענו גם מה התוחלת והשונות שלהם - ובכל זאת ידענו להגיד משהו מאוד ספציפי לגבי ההתנהגות שלהם.

משפט אי שיוויון צ'רנוף: יהיו $X_1,X_2,..,X_n$ אוסף משתנים מקריים מציינים (מקבלים 0 או 1 בלבד) בלתי תלויים (אך לאו דווקא שווי התפלגות!). נסמן: $X=\sum_{i=1}^{n}X_i$, $\mu=\mathbb{E}[X]$. אז לכל $\delta>0$:

$$\mathbb{P}(X\geq(1+\delta)\mu)\leq(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}})^{\mu}$$

יש כל מני גרסאות למשפט, למשל אפשר גם לומר:

$$\mathbb{P}(X\leq(1-\delta)\mu)\leq(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}})^{\mu}$$

הוכחה: על ידי שימוש במרקוב ואותו הטריק שראינו למעלה: נבחר $t>0$ כלשהו, ונוכל לקבל:

$$\mathbb{P}(X\geq a)\leq\mathbb{P}(e^{tX}\geq e^{ta})\leq\frac{\mathbb{E}[e^{tX}]}{e^{ta}}$$

נסמן לכל i, $\mathbb{P}(X_i=1)=p_i$. כעת:

$$\mathbb{E}[e^{tX}]=\mathbb{E}[e^{t\sum_{i=1}^{n}X_i}]=\mathbb{E}[\Pi e^{tX_i}]=\Pi_{i=1}^{n}\mathbb{E}[e^{tX_i}]$$

במהלך מאוד דומה למה שהיה למעלה - תוחלת של מכפלת משתנים מקריים בלתי תלויים שווה למכפלת התוחלות. כעת, בהסתברות p_i $X_i = 1$, ובהסתברות $1 - p_i$, $X_i = 0$, ולכן:

$$\Pi_{i=1}^{n}(p_i \cdot e^t + (1 - p_i) \cdot e^0) = \Pi_{i=1}^{n}(1 + p_i(e^t - 1))$$

כעת, כיוון שלכל $x \in \mathbb{R}$ מתקיים: $1 + x \leq e^x$, נקבל:

$$\leq \Pi_{i=1}^{n} exp(p_i(e^t - 1)) = e^{\sum_{i=1}^{n} p_i(e^t-1)} = e^{\mu(e^t-1)}$$

כיוון שהתוחלת של X היא $\mathbb{E}[X] = \sum_{i=1}^{n} p_i \cdot 1$, כתוחלת של סכום משתנים מקריים מציינים.
כלומר:

$$\mathbb{P}(X \geq (1+\delta)\mu) \leq \frac{e^{\mu(e^t-1)}}{e^{ta}}$$

כמו קודם, יש לנו בצד ימין $e^{f(t)}$, כאשר $f(t) = \mu(e^t - 1) - ta$. נוכל לגזור ולהשוות לאפס, ובכך למצוא מהו t עבורו הביטוי הוא המינימלי - ולכן נותן את החסם ההדוק ביותר. מסתבר ש$t = ln(1+\delta)$, מציבים ומקבלים את הדרוש ■

שבוע 4

2 אלגברה ליניארית

ראינו בליניארית את המושג של **מרחב וקטורי**, או **מרחב ליניארי** מעל שדה: הוא מכיל אובייקטים הנקראים וקטורים, שמוגדרות עליהם פעולת חיבור וכפל בסקלרים של השדה.
אנחנו נתעניין במרחבים וקטוריים מעל הממשיים.

בנוסף למושג של מרחבים וקטוריים, נרצה לפתח מושג של "גודל", של "מרחק" בין וקטורים, ולכן נעשיר את המרחב הוקטורי עם פונקציות שמאפשרות לעשות זאת.

באופן כללי, בהינתן אובייקט מתמטי עם מבנה, אחד מהדברים הראשונים שנרצה הוא העתקה שמשמרת את המבנה בין האובייקטים המתמטיים. במקרה של מרחבים וקטוריים ההעתקה שמשמרת את המבנה שלו היא ההעתקה הליניארית:

הגדרה: יהיו U, V מרחבים ליניאריים, **העתקה ליניארית** $T : U \to V$ זו העתקה שמקיימת:

1. הומוגניות: לכל $\alpha \in \mathbb{R}$ ו$x \in U$, $T(\alpha x) = \alpha T(x)$
2. אדיטיביות: לכל $x, y \in U$, $T(x + y) = T(x) + T(y)$

נזכיר שמרחב ליניארי יכול להיות **ממימד סופי** או **ממימד אינסופי** (בהינתן מרחב וקטורי, נאמר שאוסף של וקטורים $v_1, .., v_n$ הם **בלתי תלויים** (נכתוב בת"ל) אם הצירוף הליניארי היחיד שלהם שמתאפס הוא הצירוף הטריוויאלי, כלומר אם לכל סקלרים $\alpha_1, .., \alpha_n \in \mathbb{R}$ כך שמתקיים $\sum_{i=1}^{n} \alpha_i v_i = 0$, בהכרח לכל i, $\alpha_i = 0$. כמו כן, נאמר שאוסף של וקטורים $v_1, .., v_n$ **פורשים** את המרחב אם לכל וקטור $v \in V$ קיימים סקלרים $\alpha_1, .., \alpha_n$ כך ש$v = \sum_{i=1}^{n} \alpha_i v_n$, כלומר כל וקטור במרחב הוא צירוף ליניארי של אוסף הוקטורים הפורש. נאמר שאוסף של וקטורים הוא **בסיס** למרחב אם הוא בת"ל וגם פורש. במקרה כזה התצוגה של כל וקטור היא יחידה. כמו כן, ראינו בליאינרית שמספר הוקטורים שיש בבסיס הוא תכונה של המרחב - כל הבסיסים עבור המרחב הם בעלי אותה כמות של וקטורים. מספר הוקטורים שיש בבסיס של המרחב נקרא **המימד** של המרחב. יש מרחבים שיש להם בסיס סופי, ויש מרחבים שנדרש אוסף אינסופי של וקטורים בלתי תלויים כדי לפרוס את הבסיס [עבור אוסף אינסופי, נאמר שהוא בלתי תלוי אם כל צירוף סופי שמתאפס הוא בהכרח הצירוף הליניארי הטריוויאלי).

עבור מרחב וקטורי ממימד סופי, ראינו בליניארית שהעתקה ליניארית היא בעצם מיוצגת על ידי מטריצה, כלומר אם יש לנו העתקה T, קיימת מטריצה A כך שלכל x, $T(x) = Ax$ - הפעלה של העתקה ליניארית זה בעצם מכפלה במטריצה. בהמשך נראה שישנם שני סוגים בסיסיים של העתקות ליניאריות: מתיחה וסיבוב, ושלמעשה כל ההעתקות הליניאריות מורכבות ממתיחות וסיבובים (זהו אחד מהמשפטים היסודיים של האלגברה).

נציין שיש לנו עניין גם במרחבים אינסוף מימדיים: זה לא רק עניין אבסטרקטי, יש הרבה בעיות פרקטיות אותן נרצה לפתור באמצעות מרחבים אינסוף מימדיים. למשל, מרחב הפולינומים, או מרחב הפונקציות הרציפות מעל קטע, הם שניהם מרחבים וקטוריים אינסוף מימדיים.

2.1 מטריקות ונורמות

כפי שאמרנו, בהרבה סיטואציות נרצה לתת גודל לוקטורים. דוגמא מאוד חשובה היא הבעיה של קירוב: נניח שיש לנו מערכת משוואות כך שיש יותר משוואות מנעלמים. במצב כזה יתכן ויש למערכת פתרון יחיד, אך יתכן גם שאין פתרון בכלל – למשל במצב שחלק מהמשוואות סותרות אחת את השנייה. במצב כזה לא נוכל לתת השמות לנעלמים כך שיספקו את כל המשוואות. במקרים כאלה נרצה למצוא "קירוב" לבעיה – אך כאן כבר נכנסת השאלה, מה הכוונה בקירוב? האם זה פתרון שנותן תשובה קרובה כמה שיותר לתשובה המקורית? שמספק כמה שיותר משוואות? וכך הלאה. נרצה לפתח כלים כדי להיות מסוגלים לומר למשל כמה קירוב הוא טוב.

באופן כללי, על מנת לתאר אבסטרקציה של מרחק משתמשים במושג של מטריקה:
הגדרה: **מרחב מטרי** הוא זוג (X,d), כאשר X הוא קבוצה כלשהי ו־$d:X\times X\to\mathbb{R}$ היא העתקה הנקראת **מטריקה**, כך שהיא מקיימת:

1. סימטריה: $d(x,y)=d(y,x)$
2. אי שליליות: $d(x,y)\geq 0$, ו־$d(x,y)=0$ אם ורק אם $x=y$
3. אי שיוויון המשולש: לכל $x,y,z\in X$, $d(x,z)\leq d(x,y)+d(y,z)$

המטריקה היא ההכללה שהיינו מצפים למושג של מרחק בין שתי נקודות: המרחק הוא אי שלילי, הוא זהה אם הולכים מ־x ל־y או להיפך, וללכת מ־x ישירות ל־z זה תמיד קצר לפחות כמו ללכת דרך נקודה אמצעית.

הערה: יש גם מושג הנקרא **פסאודו־מטריקה**, בה התנאי 2 מרוכך יותר, ומופיע כך ש־$d(x,y)\geq 0$ לכל x,y, ו־$d(x,x)=0$. כלומר, יתכן מצב שבו $d(x,y)=0$ אבל $x\neq y$.

המטריקה היא אבסטרקטית, ויכולה לתאר גם עוד הרבה דברים בעולם, לא בהכרח מרחק. למשל, היא יכולה לאפיין את כמות הזמן שלוקח לעבור בין שתי נקודות, או את "מרחק העריכה" בין שני טקסטים (מרחק עריכה הוא כמה הטקסטים רחוקים אחד מהשני, במובן הזה של כמה "קשה" להפוך את אחד מהם לשני).

במיוחד נרצה סוג מיוחד של מטריקה שהיא "מכבדת" את המבנה של המרחב הוקטורי. יחד עם הדיון על מטריקה אפשר להתיחס גם לסוג מיוחד של מטריקות, שהן מטריקות שמתקבלות מתוך נורמה. נורמה היא פונקציה שמתארת לנו את האורך של וקטור במרחב וקטורי.

הגדרה: יהי U מרחב ליניארי ממשי. **נורמה** היא העתקה $\|\cdot\|:U\to\mathbb{R}$ המקיימת:

1. אי שליליות: לכל $u\in U$, $\|u\|\geq 0$, ו־$\|u\|=0$ אם ורק אם $u=\bar{0}$, כלומר אם ורק אם u הוא וקטור האפס.
2. הומוגניות: לכל $u\in U$ ו־$\alpha\in\mathbb{R}$, $\|\alpha u\|=|\alpha|\cdot\|u\|$
3. אי שיוויון המשולש: לכל $u,v\in U$, $\|u+v\|\leq\|u\|+\|v\|$

בהינתן מרחב וקטורי עם נורמה, שיקרה לפעמים **מרחב נורמי**, $(U,\|\cdot\|)$, אפשר להסתכל על **כדור היחידה**, שהוא אוסף כל הוקטורים ב־U שהנורמה שלהם לכל היותר 1:

$$B_{\|\cdot\|}=\{x\in U:\|x\|\leq 1\}$$

B שכזו היא קבוצה **קמורה** (לכל שתי נקודות ב־B גם הקו המחבר ביניהם נמצא בתוך הקבוצה), **וסימטרית יחסית לראשית** ($x\in B \Leftarrow -x\in B$).
הרעיון של כדור היחידה יכול לתת לנו דרך חלופית להגדיר נורמה: נוכל להשתמש בכדור היחידה כדי להגדיר נורמה. נתחיל מלתאר את אוסף הוקטורים שירכיבו את כדור היחידה מבחינתנו – פשוט נבחר אוסף כלשהו של וקטורים ונחליט שיש להם נורמה 1. נוכל לבחור כל אוסף שנרצה, כל עוד הוא יתן לנו קבוצה קמורה וסימטרית יחסית לראשית. כעת, בהינתן וקטור כללי $x\in U$, נוכל לשרטט ישר ממנו לוקטור האפס. יש z כלשהו שהוא החיתוך של הישר עם השפה של כדור היחידה B. נוכל לשאול – עבור איזה $\alpha\in\mathbb{R}$ יתקיים $x=\alpha z$, כלומר בכמה נצטרך למתוח (או לכווץ) את z כדי להגיע ל־x. כעת נגדיר: $\alpha=\|x\|$.
כמובן שצריך להראות שההגדרה הזו שבנינו נותנת לנו נורמה חוקית, לא הראנו זאת אבל חלק מהתכונות נובעות מכך שדרשנו שהקבוצה תהיה קמורה.

נרצה לחשוב על נורמה בשני הקשרים: בהינתן וקטור $x\in V$, הנורמה $\|x\|$ נותנת לנו מושג של **גודל**, או של **אורך**, של הוקטור, כמה הוא רחוק מהראשית, כמה הוא "גדול". ההקשר השני בו נחשוב על נורמה יהיה כדי למדוד מרחק – בהינתן שני וקטורים $x,y\in V$, הנורמה תספק לנו דרך למדוד את המרחק ביניהם (למעשה היא משרה מטריקה), על ידי הביטוי $\|x-y\|$, שנותן לנו כימות לכמה הם רחוקים אחד מהשני, כמה המרחק ביניהם גדול.

2.1.1 דוגמאות לנורמות

יש מספר עצום של נורמות (כפי שאפשר לראות לפי ההגדרה השקולה שלהן: יש המון קבוצות קמורות וסימטריות ביחס לראשית שאפשר לבחור בתור כדור היחידה). נציג כמה נורמות שהן במיוחד שימושיות. תחילה נציג כמה נורמות מעל $\mathbb{R}^n$. נניח שיש לנו $x=(x_1,..,x_n)$ ו$y=(y_1,...,y_n)$.

נורמת 2, שתסומן l_2, מוגדרת להיות:

$$\|x\|_2=\sqrt{\sum_{i=1}^n x_i^2}$$

זו הנורמה האוקלידית המוכרת.
נראה שהיא מקיימת את התכונות להיות נורמה: מובן שהיא אי שלילית, ושהנורמה של x היא אפס אם ורק אם הוא בעצמו אפס.
נראה הומוגניות:

$$\|\alpha x\|=\sqrt{\sum_{i=1}^n(\alpha^2x_i^2)}=\sqrt{\alpha^2\sum_{i=1}^n x_i^2}=|\alpha|\cdot\sqrt{\sum_{i=1}^n x_i^2}=|\alpha|\cdot\|x\|$$

אי שיוויון המשולש נובע מ**אי שיוויון קושי שוורץ**. נזכיר שבהינתן $x,y\in\mathbb{R}^n$, ניתן להגדיר את המכפלה הפנימית שלהם על ידי:

$$<x,y>=\sum_{i=1}^n x_iy_i$$

כעת, אי שיוויון קושי שוורץ קובע שמתקיים:

$$<x,y>\leq\|x\|\cdot\|y\|$$

(האינטואיציה לאי שיוויון קושי שוורץ היא ש$<x,y>=\|x\|\cdot\|y\|\cdot cos\theta$, כאשר θ היא הזוית בין הוקטורים. כיוון ש$|cos\theta|\leq 1$, אי השיוויון מובן).
נשתמש באי השיוויון ונקבל:

$$\sqrt{\sum_{i=1}^n(x_i+y_i)^2}\leq\sqrt{\sum_{i=1}^n x_i^2}+\sqrt{\sum_{i=1}^n y_i^2}\iff\sum_{i=1}^n(x_i+y_i)^2\leq\sum_{i=1}^n x_i^2+\sum_{i=1}^n y_i^2+2\sqrt{\sum_{i=1}^n x_i^2}\cdot\sqrt{\sum_{i=1}^n y_i^2}$$

$$\iff\sum_{i=1}^n x_i^2+\sum_{i=1}^n y_i^2+2\sum_{i=1}^n x_iy_i\leq\sum_{i=1}^n x_i^2+\sum_{i=1}^n y_i^2+2\sqrt{\sum_{i=1}^n x_i^2}\cdot\sqrt{\sum_{i=1}^n y_i^2}$$

$$\iff\sum_{i=1}^n x_iy_i\leq\sqrt{\sum_{i=1}^n x_i^2}\cdot\sqrt{\sum_{i=1}^n y_i^2}$$

וזה בדיוק אי שיוויון קושי שוורץ.
הערה: ראינו הוכחה כללית יותר לאי שיוויון קושי שוורץ בליניארית, שם ראינו שהוא עובד לעוד מכפלות פנימיות ונורמות **כל עוד הנורמה היא זו שמושרית מהמכפלה הפנימית**. לא לכל מכפלה פנימית ונורמה אי השיוויון עובד, אלא רק לנורמה שמורשית מהמכפלה הפנימית.

הערה: אי שיוויון קושי שוורץ הוא מקרה פרטי של אי שיוויון הלדר, אותו נלמד בהמשך

נורמת 1, שתסומן l_1, מוגדרת להיות:

$$\|x\|_1=\sum_{i=1}^n|x_i|$$

התכונות של הנורמה עבורה נובעות מכך שערך מוחלט הוא נורמה על $\mathbb{R}$, ולכן מתקיים עבורו אי שיוויון המשולש $|a+b| \leq |a| + |b|$ ו$|\alpha a| = |\alpha| \cdot |a|$.

נורמת אינסוף, שתסומן l_∞, מוגדרת להיות:

$$\|x\|_\infty = max_{1\leq i\leq n}|x_i|$$

כלומר, היא מקבלת את הערך של הקואורדינטה המקסימלית (בערך מוחלט). מובן שהיא מקיימת הומוגניות ואי שליליות. אי שיוויון המשולש נובע מכך ש:

$$max_{1\leq i\leq n}|x_i + y_i| \leq max_{1\leq i\leq n}|x_i| + max_{1\leq i\leq n}|y_i|$$

(באיבר שבצד ימין אנחנו יכולים לבחור לכל מחובר את הקואורדינטה המקסימלית, בעוד שבמחובר של צד שמאל אנחנו חייבים לבחור את אותה הקואורדינטה בשניהם).

נרצה גם נורמות עבור מרחבים אינסוף מימדיים. בעיה אמיתית יכולה להיות למשל בהינתן פונקציה רציפה $f : [a,b] \to \mathbb{R}$, למצוא לה קירוב "טוב" על ידי פולינום, כלומר למצוא פולינום שהוא "קרוב מאוד" למה שf היתה נותנת. נשים לב שf יכולה להיות פונקציה כזו שהיא בכלל לא דומה לפולינום (למשל אם יש לה שפיץ). זו שאלה שיש לה שימושים בעולם האמיתי.
השאלה הגדולה היא כמובן, בהינתן פולינום P, כיצד אנחנו מחליטים האם הוא מקרב את f מספיק? נרצה נורמה, שתתאר לנו את ה"מרחק" בין f לP, ולשם כך נרצה נורמות על פולינומים.

$$\|f-P\|_2 = \int_a^b (f(x)-P(x))^2 dx$$

(נשים לב שאינטגרל זה ההכללה הנכונה לסכום)

$$\|f-P\|_1 = \int_a^b |f(x)-P(x)dx$$

$$\|f-P\|_\infty = max_{x\in[a,b]}|f(x)-P(x)|$$

(אנחנו מניחים שf רציפה על קטע סגור ולכן מקבלת מקסימום, אפשר להגדיר גם עם סופרמום).

בשביל קירוב טוב לf על ידי P, נוכל לדרוש למשל שהמרחק ביניהם יהיה קטן מϵ, עבור $\epsilon > 0$ כלשהו.

נוכל לתת גם גודל למטריצה, כלומר להעתקה ליניארית. עבור העתקה ליניארית, המושג של גודל שנרצה יהיה עד כמה היא מותחת את הוקטורים שעוברים דרכה:
בהינתן מרחבים וקטוריים $(U, \|\cdot\|_U)$ ו$(V, \|\cdot\|_V)$, והעתקה ליניארית $T : U \to V$, נגדיר את **הנורמה האופרטורית** של T להיות:

$$\|T\|_{op} = max_{u\in U, u\neq 0}\frac{\|T(u)\|_V}{\|u\|_U}$$

<u>דוגמא לשימוש בנורמה האופרטורית:</u> נבחר $U = V = \mathbb{R}^2$, $T = id_{\mathbb{R}^2}$, $\|\cdot\|_U = l_1$, $\|\cdot\|_V = l_\infty$. נטען ש$\|T\|_{op} = 1$:
נשים לב שלכל וקטור $x \in \mathbb{R}^2$, $\|x\|_\infty \leq \|x\|_{l_1}$, כיוון ש$\|\cdot\|_{l_1}$ היא מכילה את הקואורדינטה הגדולה ביותר של x בערך מוחלט, בתוספת עוד ערכים, וששיוויון אפשרי (למשל אם כל שאר הקואורדינטות חוץ מהמקסימום מתאפסות). לכן המנה $\frac{\|x\|_\infty}{\|x\|_{l_1}}$ תמיד לכל היותר 1 , ולפעמים יש שיוויון.

2.2 לכסון

הגדרה: בהינתן העתקה לינארית T, ווקטור x **שונה מאפס**, אם קיים $\lambda \in \mathbb{R}$ כך ש$T(x) = \lambda x$, נאמר שx הוא **וקטור עצמי** של T, וλ הוא **ערך עצמי** של T.

בהינתן העתקה ליניארית T נרצה למצוא את כל הערכים העצמיים שלה. נשים לב שעבור וקטור $x \in U$, מתקיים:

$$T(x) = \lambda x \iff Ax = \lambda x = \lambda I_n x \iff (\lambda I_n - A)x = 0$$

וזה קורה אם ורק אם x נמצא ב**גרעין הימני** של המטריצה $\lambda I_n - A$ (נאמר שx נמצא בגרעין הימני של מטריצה B אם $Bx = 0$. נאמר שהוא נמצא בגרעין השמאלי של B אם $xB = 0$).

כלומר, וקטור עצמי הוא וקטור הנמצא בגרעין של $\lambda I_n - A$, וקיים וקטור כזה אם ורק אם המטריצה איננה הפיכה, כלומר היא **סינגולרית**, וזה קורה אם ורק אם הדטרמיננטה שלה מתאפסת.

כלומר, נחפש ערכים של λ עבורם $det(\lambda I_n - A) = 0$.

בהינתן מטריצה A, נגדיר את ה**פולינום האופייני** של A להיות בדיוק $P_A(\lambda) = det(\lambda I_n - A)$, כאשר A מטריצה $n \times n$.
הפולינום האופייני הוא פולינום בλ עם n משתנים, והשורשים שלו הם בדיוק הערכים העצמיים של המטריצה.

כל המטרה בדיון הזה היא להגיע למושג של **לכסון** - בהינתן מטריצה, נרצה "להעביר" אותה (אם אפשר) למודל נוח יותר. המודל הנוח יותר הוא חשוב, כיוון שהוא יכול לאפשר לנו למשל לעשות חישובים יעילים שאחרת היו עולים בהרבה זמן.

הגדרה: נאמר שמטריצות A, B הן **דומות** אם קיימת מטריצה הפיכה P כך ש$B = PAP^{-1}$, או באופן סימטרי לחלוטין $A = P^{-1}AP$ (שתי ההגדרות זהות, ולא יהיה לנו אכפת באיזה אחת נשתמש).

לכסון זו הפעולה בה מוצאים מטריצה אלכסונית (מטריצה שיש לה איברים על האלכסון ואפס בכל מקום אחר) שהיא דומה למטריצה A. כפי שראינו בליניארית, לא כל מטריצה ניתנת ללכסון.

שבוע 5

שיעור שעבר דיברנו על נורמה, שהיא מאפשרת לנו למדוד את הגודל של וקטורים, וכך גם את המרחק ביניהם (המרחק בין שני וקטורים הוא הגודל של ההפרש שלהם).

ראינו שבהינתן מרחב וקטורי V מעל $\mathbb{R}$, נורמה היא העתקה $\|\cdot\| : V \to \mathbb{R}$ כך שלכל $v \in V$, $\|v\| \geq 0$ ושיוויון אם ורק אם $v = 0$ (אי שליליות), לכל $\alpha \in \mathbb{R}$ $\|\alpha \cdot v\| = |\alpha| \cdot \|v\|$ (הומוגניות), ולכל $u, v \in V$, $\|u + v\| \leq \|u\| + \|v\|$ (אי שיוויון המשולש).

בהינתן נורמה $\|\cdot\|$ על מרחב וקטורי V, מגדירים את כדור היחידה להיות:

$$B_{\|\cdot\|} = \{v \in V : \|v\| \leq 1\}$$

טענה: כדור היחידה $B_{\|\cdot\|}$ הוא קבוצה קמורה וסימטרית מרכזית (כלומר סימטרית יחסית לאפס).
כאשר נאמר שקבוצה X היא **קמורה** אם לכל $x, y \in X$, הקטע $L_{x,y} = \{\alpha x + (1-\alpha)y : \alpha \in \mathbb{R}, 0 \leq \alpha \leq 1\} \subseteq X$. (באופן כללי, שתי נקודות במרחב וקטורי מגדירות את הישר שעובר ביניהן, שהוא יהיה **תת מרחב אפיני**, כלומר תת מרחב מוזז. נזכיר שבליניארית הגדרנו **תת מרחב וקטורי** להיות תת קבוצה לא ריקה שהיא סגורה לחיבור ולכפל בסקלר. בפרט, תת מרחב חייב להכיל את אפס. בהינתן תת מרחב W של V ווקטור $v \in V$, נוכל להסתכל על הקבוצה $W + v$, והיא תיקרא תת מרחב אפיני. הישר העובר בין שתי נקודות x, y הוא הקבוצה $\{\alpha x + (1-\alpha)y : \alpha \in \mathbb{R}\}$, ואז הקטע המחבר בין x לy הוא הקטע בישר שהוא בין שתי הנקודות בלבד).

הוכחה: סימטריות זה קל להראות: בהינתן $v \in B_{\|\cdot\|}$, אז $\|v\| \leq 1$. כעת, $\|-v\| = |-1| \cdot \|v\| = 1 \cdot \|v\| = \|v\| \leq 1$ כלומר $-v \in B_{\|\cdot\|}$.
כעת, יהיו $x, y \in B_{\|\cdot\|}$. אז $\|x\|, \|y\| \leq 1$. יהי $0 \leq \alpha \leq 1$ ממשי. אז:

$$\|\alpha x + (1-\alpha)y\| \leq \|\alpha x\| + \|(1-\alpha)y\| = |\alpha|\|x\| + |1-\alpha| \cdot \|y\| = \alpha \cdot \|x\| + (1-\alpha)\|y\| \leq \alpha + (1-\alpha) = 1$$

בגלל אי שיוויון המשולש, הומוגניות הנורמה, העובדה שα בין 0 ל1, וכיוון שהנורמות של x ושל y קטנות שוות ל1. לכן $\alpha x + (1-\alpha)y \in B_{\|\cdot\|}$ כנדרש ■

בהינתן קבוצה קמורה סימטרית מרכזית, אפשר להגדיר נורמה כפי שאמרנו בשיעור קודם: בהינתן וקטור u כלשהו, אפשר לכווץ או למתוח אותו עד שהוא יחתוך את מעגל היחידה (המעגל שנוצר על ידי אוסף הוקטורים שהנורמה שלהם בדיוק 1). למשל, אם עבור וקטור u, היה צריך להכפיל אותו ב$\frac{1}{2}$ כדי שיהיה על מעגל היחידה, נגדיר את הנורמה שלו להיות 2. באופן פורמלי:

$$\|u\|_B = \frac{1}{max\{\lambda > 0 : \lambda u \in B\}}$$

היינו צריכים לקחת את הλ המקסימלי כיוון שהיה אפשר להקטין גם ביותר ועדיין לקבל נקודה בתוך כדור היחידה, וההסבר הפורמלי לכך שהוקטור יהיה "בדיוק" על המעגל הוא מסובך. הפונקציה $\|\cdot\|_B$ היא אכן נורמה, אך לא הראנו זאת בכיתה.

2.3 ריבוי אלגברי וריבוי גאומטרי

הערה: כפי שראינו קודם, כל העתקה ליניארית היא בעצם מטריצה, וכל מטריצה היא העתקה ליניארית.

תזכורת: בהינתן מטריצה A, נאמר שהיא מותחת וקטור $x \in V$ פי λ אם $Ax = \lambda x$. אם בנוסף $x \neq 0$, אז נאמר שx הוא וקטור עצמי של A, וλ יקרא ערך עצמי של A.
הערכים העצמיים של A הם השורשים של הפולינום האופייני של A, כלומר הם השורשים של הפולינום $det(\lambda I - A) = 0$, וזה פולינום ממעלה n עם משתנה λ.

המשפט היסודי של האלגברה: לפולינום ממשי ממעלה n יש n שורשים מרוכבים.

כלומר, מובטח לנו שלפולינום האופייני יהיו n שורשים, אך יתכן והם יהיו מרוכבים (ולכן לא רלוונטיים בתור ערכים עצמיים למטריצה ממשית).

דבר נוסף שחשוב לשים לב אליו הוא שn השורשים המובטחים לנו הם לא בהכרח שונים. למשל, לפולינום $x^4 = 0$ יש מבחינת המשפט היסודי 4 שורשים, כיוון ש0 מופיע 4 פעמים, אך הם לא שונים.
כלומר, בפולינום האופייני ערך עצמי יכול להופיע יותר מפעם אחת. אולי היינו מצפים שאם ערך עצמי מופיע יותר מפעם אחת אז בהתאם יהיו לו גם מספר וקטורים עצמיים - אך מסתבר שזה לא המצב.

הגדרה: מספר הפעמים שערך עצמי מופיע בפולינום האופייני נקרא **הריבוי האלגברי** של הערך העצמי.

הגודל באמת מנבא כמה וקטורים עצמיים "שונים" יש הוא המימד של המרחב של הוקטורים העצמיים: אוסף הוקטוריים העצמיים של ערך עצמי מהווים מרחב וקטורי, והמימד שלו הוא הגודל הנכון כדי לתפוס את כמות הוקטורים העצמיים השונים של הערך העצמי:
הגדרה: מימד מרחב הוקטורים העצמיים השייכים לערך עצמי מסויים נקרא **הריבוי הגאומטרי** של הערך העצמי.

טענה: הריבוי האלגברי הוא תמיד גדול שווה לריבוי הגאומטרי.

יש גם מקרים בהם הריבוי האלגברי גדול ממש מהריבוי הגאומטרי, למשל:
$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \Leftarrow \lambda I - A = \begin{bmatrix} \lambda & -1 \\ 0 & \lambda \end{bmatrix} \Leftarrow det(\lambda I - A) = \lambda^2$ ולכן יש לנו את 0 עם ריבוי אלגברי של 2.

נחפש וקטור עצמי: כלומר, נחפש $x=\begin{bmatrix}x_1\\x_2\end{bmatrix}$ כך ש$Ax=\lambda=0 \Leftarrow 0=\begin{bmatrix}x_2\\0\end{bmatrix}=\begin{bmatrix}0&1\\0&0\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}$. כלומר, כל וקטור מהצורה $\begin{bmatrix}x_2\\0\end{bmatrix}$ יהיה וקטור עצמי עבור A, וזה מרחב שנפרש על ידי וקטור יחיד - למשל $\begin{bmatrix}1\\0\end{bmatrix}$. כלומר, הריבוי הגאומטרי של 0 הוא 1.

הערה: מטריצה תיקרא **מטריצה רגילה** אם כל הערכים העצמיים שלה הם עם ריבוי של 1 ("פשוטים"). המטריצות הרגילות הן צפופות בין כל שאר המטריצות, כלומר לכל A ממשית ולכל $\epsilon>0$ קיימת מטריצה רגילה D כך שלכל i,j, $|a_{ij}-d_{ij}|<\epsilon$.
לא נוכיח זאת.
באופן כללי, מעניין לחשוב על המצב שבו העולם הכללי של האובייקטים איתם אנחנו עוסקים הוא מאוד גדול, אך בפועל מעניין אותנו רק תת אוסף מאוד מסויים שלהם, שיכול להיות מאוד קטן יחסית למרחב הגדול.

הערה: נשים לב שבהינתן מטריצה A וסקלר λ, יכול להיות ערך עצמי ימני $Ax=\lambda x$ וערך עצמי שמאלי $yA=\lambda y$, אך המושג של קיום של ערך עצמי הוגדר דרך הדרמיננטה של המטריצה $\lambda I-A$, והוא לא מתיחס לסדר במכפלה. למעשה, מובטח לנו קיום של וקטור עצמי ימני וגם וקטור עצמי שמאלי, אך אין סיבה לחשוב שיש איזשהו קשר ביניהם.

כפי שאמרנו קודם, המטריצות האלכסוניות יעניינו אותנו מאוד. נזכיר שנאמר שמטריצות A,B הן דומות אם יש מטריצות הפיכה (כלומר לא סיגנולרית) P כך ש$B=PAP^{-1}$, ובאופן סימטרי $A=P^{-1}BP$.
נשים לב שיחס הדמיון של מטריצות הוא יחס שקילות.

טענה: למטריצות דומות יש את אותו הפולינום האופייני ולכן גם את אותם הערכים העצמיים.

הוכחה: נניח ש$B=PAP^{-1}$. אז:

$$det(\lambda I-B)=det(\lambda I-PAP^{-1})$$

כיוון ש$I=PP^{-1}$, וכפל בסקלר מתחלף עם כפל מטריצות, נוכל להמשיך לפתח ולקבל:

$$=det(P\lambda IP^{-1}-PAP^{-1})=det(P(\lambda IP^{-1}-AP^{-1}))=det(P(\lambda I-A)P^{-1})$$

דטרמיננטה היא כפלית, ולכן:

$$=det(P)\cdot det(P^{-1})\cdot det(\lambda I-A)=det(\lambda I-A)$$

כיוון ש$det(P)\cdot det(P^{-1})=1$ ■

נרצה למצוא מטריצה אלכסנית שדומה למטריצה שלנו, ולתהליך זה כאמור קוראים לכסון.

במטריצה אלכסונית, הערכים העצמיים הם הערכים שעל האלכסון. יתר על כן, אם $A=QDQ^{-1}$, כאשר D היא מטריצה אלכסונית, אז נוכל להכפיל בQ ולקבל $AQ=QD$. העמודות של Q הם הוקטורים העצמיים המתאימים לערכים העצמיים של D, בסדר שבו הם מופיעים, (ובפרט יש לD ולA את אותם הערכים העצמיים).

למה: יהיו $v_1,..,v_k$ וקטורים עצמיים של המטריצה A המתאימים לערכים העצמיים **שונים** $\lambda_1,..,\lambda_k$. אז הם בלתי תלויים ליניארית.

הוכחה: נניח בשלילה שלא, ונסתכל על הקבוצה הקטנה ביותר שהיא תלויה ליניארית. בלי הגבלת הכלליות, נניח שאלה הוקטורים $v_1,..,v_l$. אז יש לנו סלקרים לא טריוויאליים כך שמתקיים:

$$\sum_{i=1}^{l} \alpha_i v_i = 0$$

נפעיל את A, ונקבל:

$$A \cdot (\sum_{i=1}^{l} \alpha_i v_i) = \sum_{i=1}^{l} \alpha_i A v_i = \sum_{i=1}^{l} \alpha_i \lambda_i v_i$$

נכפיל את השיוויון הראשון בλ_l, ונקבל $\sum_{i=1}^{l} \alpha_i \lambda_l v_i = 0$. נחסר בין המשוואות ונקבל:

$$\sum_{i=1}^{l} \alpha_i (\lambda_i - \lambda_l) v_i = 0$$

בפרט, המחובר הl מתבטל כיוון ש$\alpha_i(\lambda_l - \lambda_l)v_l = \alpha_i \cdot 0 \cdot v_l = 0$, ולכן נסיק:

$$\sum_{i=1}^{l-1} \alpha_i (\lambda_i - \lambda_l) v_i = 0$$

לפי הנתון, לכל $i \neq l$, $\lambda_i \neq \lambda_l$, ולפי הנחת השלילה $\alpha_i \neq 0$, ולכן יש כאן צירוף ליניארי לא טריוויאלי של $v_1,..,v_{l-1}$ שמתאפס - סתירה למינימליות l ■

2.4 סיבוב, מטריצות אורתוגונליות והמשפט הספקטרלי למטריצות ממשיות

הגדרה: **סיבוב** הוא העתקה לינאירית שמשמרת את נורמת l_2. כלומר, זו העתקה הנתונה על ידי מטריצה A כך שלכל $x \in V$, $\|x\|_2 = \|Ax\|_2$. העתקה כזו תיקרא גם **איזומטריה של** l_2.

נזכיר שאם $x = (x_1,..,x_n)$, אז

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

ואם בנוסף $y = (y_1,..,y_n)$, אז

$$< x, y >= \sum_{i=1}^{n} x_i y_i$$

(מכפלה פנימית). לכן לפי ההגדרה,

$$\|x\|_2^2 =< x, x >$$

הגדרה: נאמר שמטריצה A היא **אורתוגונלית** אם $A^T A = I$
זה שקול לכך ש$A^T = A^{-1}$, כאשר A^T היא המטריצה שמתקבלת על ידי החלפת האיברים של A לפי האלכסון הראשי. כלומר, אם $A = (a_{ij})$, אז $A^T = (a_{ji})$.

הערה: אם M מטריצה, אז MM^T זו המטריצה שבמקום הij שלה נמצאת המכפלה הפנימית של השורה הi והשורה הj של M. זה ממש מההגדרה של מכפלה פנימית: אם מכפילים מטריצות A וB, אז $(AB)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$, כלומר מכפלה פנימית של השורה הi של A עם העמודה הj אצל B. כיוון שהעמודה הj של M^Tזה בדיוק השורה הj של M, הטענה מובנת.
בפרט, $AA^T = I$ אם ורק אם לכל שתי שורות שונות המכפלה הפנימית מתאפסת, כלומר כל שתי שורות שונות הן וקטורים אורתוגונליים אחד לשני, והמכפלה של כל שורה עם עצמה נותן את 1, כלומר כל שורה היא וקטור יחידה.

טענה: כפל בA זו העתקת סיבוב אם ורק אם A היא מטריצה אורתוגונלית

הוכחה: $\Rightarrow$: אם A היא מטריצה אורטוגונילית, אז לפי ההגדרה $A^TA = AA^T = I$. אז לכל וקטור x,

$$\|x\|_2^2 = \|Ix\|_2^2 = \|A^TAx\|_2^2 =< A^TAx, A^TAx >= (A^TAx)^T(A^TAx) = x^TA^T(AA^T)Ax = x^TA^TAx =< Ax, Ax >= \|Ax\|_2^2$$

הנורמה היא אי שלילית, ולכן זה גורר $\|Ax\|_2 = \|x\|_2$ כלומר A סיבוב.
$\Leftarrow$: נניח שA היא העתקת סיבוב. נשים לב שלפי ליניאריות המכפלה הפנימית, לכל x, y:

$$\|x+y\|_2^2 =< x+y, x+y >=< x, x > + < y, y > +2 < x, y >= \|x\|_2^2 + \|y\|_2^2 + 2 < x, y >$$

זה נכון על פי ההגדרה:

$$\|x+y\|_2^2 = \sum_{i=1}^{n}(x_i+y_i)^2 = \sum_{i=1}^{n}(x_i^2+y_i^2+2x_iy_i) = \sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} y_i^2 + 2\sum_{i=1}^{n} x_iy_i = \|x\|_2^2 + \|y\|_2^2 + 2 < x, y >$$

כעת, נציב את $A(x+y)$ בשיוויון ונקבל:

$$\|A(x+y)\|_2^2 = \|Ax\|_2^2 + \|Ay\|_2^2 + 2 < Ax, Ay >= \|x\|_2^2 + \|y\|_2^2 + 2 < A^TAx, y >$$

מצד שני:

$$\|A(x+y)\|_2^2 = \|x+y\|_2^2 = \|x\|_2^2 + \|y\|_2^2 + 2 < x, y >$$

כיוון שA היא סיבוב. לכן מכל השיוויונות נסיק:

$$< x, y >=< A^TAx, y >$$

מכפלה פנימית היא ליניארית ב$\mathbb{R}$, ולכן:

$$< (A^TA - I)x, y >= 0$$

וזה נכון **לכל** $x, y \in V$. כלומר, כל y הוא אורתוגונלי ל$(A^TA - I)x$, וזה יכול לקרות אם ורק אם $(A^TA - I)x = 0$. x היה שרירותי גם כן, לכן נסיק $A^TA - I = 0$ כלומר $A^TA = I \Leftarrow A$ אורתוגונלית ■

נאמר שA ניתנת **ללכסון אורתוגנלי** אם היא ניתנת ללכסון ובנוסף המטריצה המלכסנת O היא מטריצה אורתוגונלית.

המשפט הספקטרלי למטריצות ממשיות: כל מטריצה ממשית סימטרית ($A^T = A$) ניתנת ללכסון אורתוגונלי, כלומר לכל A ממשית סימטרית יש O אורתוגונלית כך ש$A = ODO^T$, כאשר D היא מטריצה אלכסונית.

נשים לב שכפי שאמרנו קודם, האיברים שיושבים על האלכסון של D הם הערכים העצמיים של A, והעמודות של O הם וקטוריים עצמיים המתאימים לערכים העצמיים שמופיעים על האלכסון של D לפי הסדר.

2.4.1 נורמה אופרטורית ונורמה דואלית

יש כל מני דרכים למדוד את ה"גודל" של מטריצה. אחד מהדרכים החשובות היא לפי "כמה היא מותחת", כפי שמודדים העתקות ליניאריות - כפי שהראנו קודם:

הגדרה: בהינתן $(V, \|\cdot\|_V)$, $(U, \|\cdot\|_U)$ מרחבים נורמיים, העתקה ליניארית $T : U \to V$, נגדיר את **הנורמה האופרטורית** של T להיות:

$$\|T\|_{op} = \max_{x \in U} \frac{\|T(x)\|_V}{\|x\|_U} = \max_{x \in U, \|x\|_U = 1} \|T(x)\|_V$$

וכך גם עבור מטריצה A.

נקבע וקטור $u \in U$ כלשהו, ונסתכל על ההעתקה $x \to < u, x >$, כלומר ההעתקה שלוקחת כל וקטור $x \in U$ למכפלה הפנימית של x עם u. זו העתקה ליניארית שמוגדרת דרך u. נגדיר:

$$\|u\|^* = \max_{x \in U} \frac{| < u, x > |}{\|x\|_U} = \max_{\|x\|_U = 1} | < u, x > |$$

מסתבר שהבנייה הזו אכן נותנת נורמה, והיא נקראת **הנורמה הדואלית** של הנורמה $\|\cdot\|_U$.

בהינתן נורמה, מעניין להסתכל ולראות מהי הנורמה הדואלית שלה:
טענה:

1. $\|\cdot\|_1^* = \|\cdot\|_\infty$
2. $\|\cdot\|_\infty^* = \|\cdot\|_1$
3. $\|\cdot\|_2^* = \|\cdot\|_2$

הוכחה (חלקית): לפי ההגדרה, בהינתן וקטור $u = (u_1, .., u_n)$:

$$\|u\|_1^* = \max_{\|x\|_1 = 1} |\sum_{I=1}^{n} u_i x_i|$$

כלומר, הסכום של כל הקואורדינטות של x בערך מוחלט צריך להיות 1, והמכפל ההפנימית שלו עם $\underline{u}$ צריכה להיות מקסימלית. מיידי שנרצה $sign(x_i) = sign(u_i)$ כדי שלא יתבטלו איברים. כיוון שהסכום של כל הקואורדינטות של x הן קטנות שוות 1, למעשה הוא נותן משקל לקואורדניטות של u, כאשר כל קואורדינטה מקבלת משקל בין 0 ל1. לכן הגיוני שנרצה לקחת את הקואורדינטה של u עם המשקל הגבוהה ביותר ולתת לה משקל 1, ולבטל את כל האחרות. לכן נגדיר את x לקבל 1 במקום הi (אם הקואורדינטה המקסימלית נמצאת במקום הi), ובכל מקום אחר 0. לכן:

$$\|u\|_1^* = \max_{1 \le i \le n} |u_i| = \|u\|_{infty}$$

באותו אופן מראים גם עבור 2.
עבור 3, שוב נרצה:

$$\|u\|_2^* = \max_{\sum x_i^2 = 1} |\sum_{i=1}^{n} u_i x_i|$$

נרצה שוב $sign(x_i) = sign(u_i)$ בכל i. הכי נוח לחשוב על מה שקורה כאן מבחינה גרפית: המכפלה הפנימית $< u,x >$ היא $\|u\| \cdot \|x\| \cdot cos\theta$, כאשר θ היא הזוית שביניהם. בהכרח $\|x\| = 1$. נרצה למקסם את $cos\theta$ - ה־cos המקסימלי מתקבל כאשר $\theta = 0$, ואז $cos\theta = 1 \Leftarrow$ נרצה לבחור וקטור x כזה שהוא וקטור יחידה ובדיוק באותו כיוון כמו u, ולכן נצטרך לבחור:

$$x = \frac{u}{\|u\|_2}$$

ובכך נסיים ■

הערה: במרחבים סוף מימדיים, $\|u\|^{**} = \|u\|$

הגדרה: למעשה, הנורמות l_1, l_2, l_∞ הן מקרים פרטיים של משפחה שלמה של נורמות שנקראות **נורמות־**p. בהינתן $1 \leq p \leq \infty$, ווקטור $x = (x_1, .., x_n)$, נגדיר:

$$\|x\|_p = (\sum_{i=1}^{n} |x_i|^p)^{\frac{1}{p}}$$

זו אכן נורמה, אך לא נראה זאת. בייחוד קשה להראות את אי שיוויון המשולש.

משפט: בהינתן $1 \leq p \leq \infty$, נבחר את q להיות המספר שיקיים את המשוואה:

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

אז הנורמה הדואלית ל־L_p היא L_q

לא הוכחנו משפט זה.

טענה: לכל וקטורים x, y, $| < x,y > | \leq \|x\| \cdot \|y\|^*$

הוכחה: זה מיידי מההגדרה: לפי ההגדרה,

$$\|y\|^* = \max_{x \in V} \frac{| < x,y > |}{\|x\|}$$

ולכן בפרט עבור x כלשהו, $\|y\|^* \leq \frac{|<x,y>|}{\|x\|}$. נכפיל ב־$\|x\|$ ונקבל את הדרוש ■

אי שיוויון הלדר: יהיו p, q כך שמתקיים $\frac{1}{p} + \frac{1}{q} = 1$. אז:

$$\|x\|_p \|y\|_q \geq | < x,y > |$$

אי שיוויון הלדר נובע ישירות משתי הטענות הקודמות, אבל כיוון שלא הראנו הוכחה לכך שאם $\frac{1}{p} + \frac{1}{q} = 1$ אז נורמת p ונורמת q הן דואליות, נראה הוכחה ישירה לאי שיוויון הלדר בפעם הבאה.

שבוע 6

תזכורת: בהינתן מרחב עם נורמה $(U, \|\cdot\|)$ ווקטור $x \in U$, אפשר להסתכל על ההעתקה שלוקחת כל וקטור $y \in U$ למכפלה $< x, y >$. זו העתקה ליניארית, ולכן ניתן לדבר על הנורמה האופרטורית שלה. הנורמה האופרטורית של ההעתקה הזו נקראת **הנורמה הדואלית** לנורמה $\|\cdot\|$, והיא מוגדרת להיות:

$$\|x\|^* = \max_{y \in U, \|y\|=1} | < x, y > |$$

כיוון שההעתקה הליניארית $< x, \cdot >$ לוקחת וקטור בU ומחזירה מספר $< x, y >$ ב$\mathbb{R}$ - שניתן לחשוב עליו בתור מרחב נורמי עם הערך המוחלט בתור הנורמה. לכן זה בדיוק ההגדרה של נורמה אופרטורית.

בשיעור שעבר הזכרנו משפט חשוב, מבלי להוכיח אותו:
משפט: יהיו $1 \leq p, q \leq \infty$ כך ש$\frac{1}{p} + \frac{1}{q} = 1$. אז הנורמות L_p, L_q הן נורמות דואליות.

הערה: לא הראנו שנורמת־p היא נורמה, כיוון שזה קצת ארוך ומסורבל (בפרט להראות את אי שיוויון המשולש).

משפט אי שיוויון הלדר: יהיו $x, y \in \mathbb{R}^n$, $1 \leq p, q \leq \infty$ כך ש$\frac{1}{p} + \frac{1}{q} = 1$. אז:

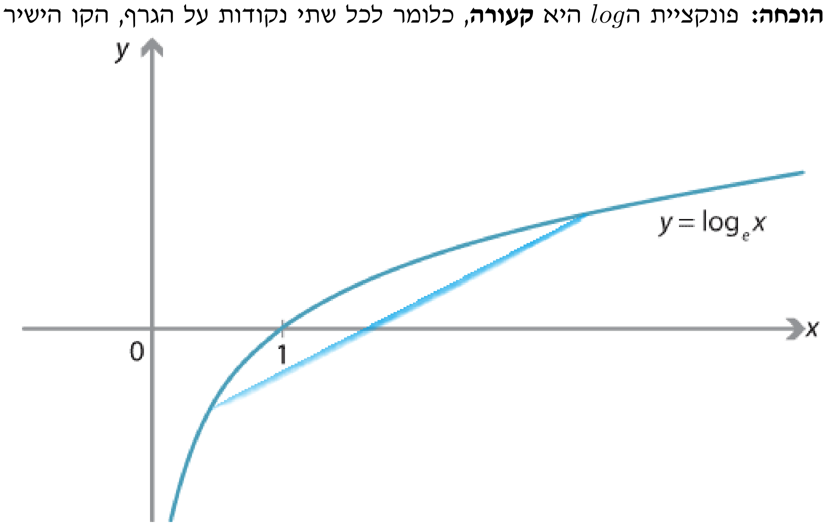
$$| < x, y > | \leq \|x\|_p \|y\|_q$$

$smallskip$
בהינתן המשפט הקודם, כלומר אם אנחנו "מאמינים" שL_p וL_q הן נורמות דואליות אחת לשנייה, המשפט הזה מתקבל כחלק מטענה שהראנו שיעור קודם, שנובעת מהגדרת הנורמה הדואלית. נוכיח אותו ישירות, מבלי להסתמך על כך.

תחילה נוכיח טענת עזר:
משפט אי שיוויון יאנג: יהיו $1 \leq p, q \leq \infty$ כך שמתקיים: $\frac{1}{p} + \frac{1}{q} = 1$, ויהיו $a, b > 0$. אז:

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

הוכחה: פונקציית הlog היא **קעורה**, כלומר לכל שתי נקודות על הגרף, הקו הישיר המחבר ביניהן נמצא מתחת לגרף.



המשמעות היא שלכל $0 \leq \alpha \leq 1$ ולכל $x, y > 0$ מתקיים:

$$log(\alpha x + (1-\alpha)y) \geq \alpha log x + (1-\alpha) log y$$

כעת, נציב $\alpha = \frac{1}{p}$, $1-\alpha = \frac{1}{q}$, ולכן:

$$log(ab) = log(a) + log(b) = log((a^p)^{\frac{1}{p}}) + log((b^q)^{\frac{1}{q}}) = \frac{1}{p}log(a^p) + \frac{1}{q}log(b^q)$$

$$= \alpha log(a^P) - (1-\alpha)log(b^q) \le log(\alpha a^p + (1-\alpha)b^q) = log(\frac{a^p}{p} + \frac{b^q}{q})$$

כלומר, קיבלנו:

$$log(ab) \le log(\frac{a^p}{p} + \frac{b^q}{q})$$

אבל log היא פונקציה מונוטונית עולה, ולכן נובע מכך:

$$ab \le \frac{a^p}{p} + \frac{b^q}{q}$$

כנדרש ■

הוכחת אי שיוויון הלדר: יהיו $x, y \in \mathbb{R}^n$, ונרצה להראות $|< x, y >| \le \|x\|_p \|y\|_q$. זה שקול לכך ש: $\frac{|<x,y>|}{\|x\|_p\|y\|_q} \le 1$. נעת:

$$\frac{|< x, y >|}{\|x\|_p\|y\|_q} = \frac{|\sum_{i=1}^n x_i y_i|}{\|x\|_p\|y\|_q} \le \sum_{i=1}^n \frac{|x_i| \cdot |y_i|}{\|x\|_p\|y\|_q}$$

לפי אי שיוויון המשולש. כעת נשתמש באי שיוויון יאנג, עבור

$$a_i = \frac{|x_i|}{\|x\|_p}, b_i = \frac{|y_i|}{\|y\|_q}$$

נשים לב שאי שיוויון יאנג נכון גם אם a או b הם אפס, לכן נוכל להציב את a_i וb_i באי השיוויון.
נקבל:

$$= \sum_{i=1}^n a_i b_i \le \sum_{i=1}^n (\frac{a_i^p}{p} + \frac{b_i^q}{q}) = \sum_{i=1}^n (\frac{|x_i|^p}{p\|x\|_p^p} + \frac{|y_i|^q}{q\|y\|_q^q}) = \frac{1}{p} \cdot \frac{\sum_{i=1}^n |x_i|^p}{\|x\|_p^p} + \frac{1}{q} \cdot \frac{\sum_{i=1}^n |y_i|^q}{\|y\|_q^q}$$

כעת, לפי ההגדרה:

$$\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}} \Rightarrow \|x\|_p^p = \sum_{i=1}^n |x_i|^p$$

וכנ"ל גבי $\|y\|_q$, לכן נסיק:

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

לפי הנתון. כלומר, הראנו

$$\frac{|< x, y >|}{\|x\|_p\|y\|_q} \le 1 \Rightarrow |< x, y >| \le \|x\|_p\|y\|_q$$

ולכן סיימנו ■

2.5 מטריצות מוגדרות חיובית (PD) ומטריצות מוגדרות אי שליליות (PSD)

הגדרה: נאמר שמטריצה ממשית סימטרית היא **מוגדרת חיובית** **(positive definite - PD)** אם כל הערכים העצמיים שלה הם חיוביים.
באופן דומה אומרים שמטריצה ממשית סימטרית היא **מוגדרת אי שלילית** **(positive semi definite - PSD)** אם כל הערכים העצמיים שלה אי שליליים.

נזכיר שלפי המשפט הספקטרלי, כל מטריצה ממשית סימטרית ניתנת ללכסון אורתוגונלי, כלומר קיימת מטריצה אורתוגונלית U ומטריצה אלכסונית Λ כך שמתקיים: $A = U\Lambda U^T$.

משפט: תהי A מטריצה ממשית סימטרית. אז התנאים הבאים שקולים:

1. A היא מוגדרת אי שלילית
2. קיימת מטריצה ממשית M עבורה $A = MM^T$
3. התבנית הריבועית xAx^T היא אי שלילית, כאשר **תבנית ריבועית** מוגדרת באופן הבא: לכל $x, y \in \mathbb{R}^n$, התבנית $G_A(x,y) = \sum_{i,j} a_{ij}x_iy_j$.
התבנית היא אי שלילית אם לכל $x \in \mathbb{R}^n$ מתקיים $G(x,x) \geq 0$, כלומר $\sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j \geq 0$

הוכחה:
(1)⇐(2): A מטריצה ממשית סימטרית ולכן ניתן לכתוב $A = U\Lambda U^T$ כאשר U אורתוגונלית וΛ אלכסונית. כיוון ש$A \in PSD$, נובע שהערכים העצמיים שלה הם אי שליליים, כלומר $\Lambda = diag(\lambda_1, .., \lambda_n)$, ולכל $1 \leq i \leq n$, $\lambda_i \geq 0$ ⇐ מכאן שלכל $1 \leq i \leq n$, $\lambda_i^{\frac{1}{2}}$ מוגדר. נגדיר את המטריצה $\Lambda^{\frac{1}{2}} = diag(\lambda_1^{\frac{1}{2}}, .., \lambda_n^{\frac{1}{2}})$, ונגדיר $M = U\Lambda^{\frac{1}{2}}$. כעת:

$$MM^T = (U\Lambda^{\frac{1}{2}})(\Lambda^{\frac{1}{2}}U^T)$$

שכן שחלוף של מטריצה אלכסונית לא משנה אותה. מכפלה בין מטריצות אלכסוניות זו המטריצה האלכסונית שאיברי האלכסון שלה הם מכפלות האיברים של המטריצות המתאימות, לכן $\Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}} = diag(\lambda_1, ...\lambda_n) = \Lambda$, ומכאן:

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

(2)⇐(3): נניח ש$A = MM^T$. בהינתן וקטור שורה x כלשהו:

$$xAx^T = xMM^Tx^T =< xM, xM >= \|xM\|_2^2 \geq 0$$

כאשר השתמשנו בכך שאם x, y וקטורי שורה אז $xy^T =< x, y >$, ואז מהגדרת הנורמה ואי שליליות הנורמה.

(3)⇐(1): נניח שהתבנית הריבועית xAx^T היא אי שלילית, ונראה שA מוגדרת אי שלילית. כלומר, נצטרך להראות שכל הערכים העצמיים של A הם אי שליליים - ואכן, יהי x וקטור עצמי לערך עצמי λ של A. אז:

$$xA = \lambda x \Rightarrow xAx^T = \lambda xx^T = \lambda\|x\|_2^2$$

וכיוון ש$xAx^T \geq 0$, ו$\|x\|_2^2 \geq 0$, נובע שגם $\lambda \geq 0$ ■

הערה: נשים לב שבהוכחה של המשפט השתמשנו בוקטור עצמי מהכיוון ההפוך למה שבדרך כלל אנחנו משתמשים. אמרנו מקודם שאם λ הוא ערך עצמי, אז מובטחים לנו וקטורים עצמיים מכל כיוון (כאשר הם לא בהכרח אותו הוקטור, כלומר לא בהכרח שוקטור אחד יהיה וקטור עצמי לשני הכיוונים). אך כאשר המטריצה היא סימטרית, אז אכן אותו הוקטור יעבוד לשני הכיוונים.

מסקנה: האוסף PSD_n (אוסף המטריצות הממשיות הסימטריות שהן מוגדרות אי שליליות) מהווה **קונוס**, כלומר זו קבוצה סגורה לכפל בסקלר חיובי ולחיבור.
כיוון שאם $A,B\in PSD_n$ ו$c>0$, אז:

$$x(A+B)x^T=xAx^T+xBx^T\geq 0, xcAx^T=cxAx^T\geq 0$$

2.6 משפט הSVD (singular value decomposition)

לפני שננסח את המשפט, נשים לב לכך שאוסף המטריצות האורתוגונליות הוא סגור לכפל ולהופכי. כלומר, אם U,V הן מטריצות אורתוגונליות, אז גם UV היא מטריצה אורתוגונלית, וגם $U^{-1}=U^T$ היא מטריצה אורתוגונלית (סגירות להופכי מובנת דרך התיאור של אופן הפעולה של מטריצה אורתוגונלית: אם הפעולה של U היתה לסובב בזוית כלשהי θ, אז הפעולה של U^{-1} תצטרך להיות לסובב חזרה, כלומר לסובב בזוית $-\theta$, ולכן לפי טענה שראינו קודם מיוצגת על ידי מטריצה אורתוגונלית גם כן). (מבחינה מתמטית המשמעות היא שאוסף המטריצות האורתוגונליות מהווה חבורה).

מסקנה: הכפלה מימין ומשמאל של מטריצה A במטריצות אורתוגונליות והופכיות שלהם לא משנה את הנורמה האופרטורית.
זאת מכיוון שהכפלה במטריצה אורתוגונלית מימין ומשמאל לא משנה את הנורמה האופרטורית, ולפי מה שכתוב למעלה גם מכפלה של מטריצות אורתוגונליות וההופכיות שלהן היא מטריצה אורתוגונלית, ולכן כפל בה לא משנה את הנורמה האופרטורית של A.

משפט הSVD: לכל מטריצה ממשית $A\in M_{m\times n}(\mathbb{R})$ קיימות מטריצות $U_{m\times m},V_{n\times n}$ אורתוגונליות ו$D_{m\times n}$ אלכסונית, כך ש$A=UDV^T$.
יתר על כן, המטריצה $D=diag(\sigma_1,\sigma_2,..)$ נקבעת ביחידות, ו$\sigma_1=\|A\|_{op}$

הערכים $\sigma_1,\sigma_2...$ נקראים **הערכים הסינגולריים** של המטריצה A

הערה: הסימון $A=UDV^T$ הוא פשוט סימון מקובל - V^T היא גם כן מטריצה אורתוגונלית ולכן היינו יכולים עקרונית לסמן את V^T בתור V, אלא שהסימון הזה הוא הסימון שמקובל להשתמש בו.

הערה: נשים לב שA **איננה** בהכרח ריבועית, המשפט נכון גם למטריצות מלבניות כלליות. במקרה הזה, כשאנחנו אומרים שD היא מטריצה "אלכסונית" הכוונה היא שD נראית באחת משתי הצורות הבאות:

$$\begin{bmatrix}D'\\ \overline{0}\end{bmatrix},\begin{bmatrix}D'| & 0\end{bmatrix}$$

כאשר D' היא מטריצה אלכסונית ריבועית רגילה. כלומר, לD יש איברים על האלכסון שלה כמה שאפשר, ואז משלימים את מה שנשאר עם אפסים.

לפני הוכחת המשפט, ננסח מסקנות ממנו:

2.6.1 מסקנות ממשפט הSVD

קודם כל, משפט הSVD נותן לנו את מה שדיברנו עליו עוד מההתחלה של הפרק על אלגברה ליניארית: כל העתקה ליניארית היא הרכבה של סיבוב, מתיחה וסיבוב. אם T העתקה ליניארית כלשהי, היא בעצם מכפלה במטריצה A כלשהי. פירוק הSVD נותן לנו $A=UDV^T$. כעת, בהינתן וקטור x, A פועלת עליו בכך שקודם מכפילים בV^T (סיבוב), ואז בD (מתיחה) ואז בU (סיבוב).

מסקנה: אם $\sigma_1,...$ הם הערכים הסינגולריים של A, אז $\sigma_1^2,\sigma_2^2...$ הם הערכים העצמיים של המטריצה (הסימטרית) AA^T.

הוכחה: לפי משפט הSVD, $A=UDV^T$, $D=diag(\sigma_1,\sigma_2,..)$. נחשב:

$$AA^T=UDV^TVD^TU^T=UDD^TU^T$$

כיוון שV היא מטריצה אורתוגונלית. לכן:

$$= U \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & ..... \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & ..... \end{bmatrix} U^T = U \begin{bmatrix} \sigma_1^2 & & \\ & \sigma_2^2 & \\ & & ..... \end{bmatrix} U^T$$

AA^T היא מטריצה PSD לפי משפט קודם, ולכן נסיק ש...σ_1^2, σ_2^2 הם הערכים העצמיים של AA^T ■

המסקנה האחרונה מסבירה גם מדוע הערכים הסינגולריים של A נקבעים ביחידות.

העובדה ש$\sigma_1 = \|A\|_{op}$ תתקבל לפי הבנייה של ההוכחה.

2.6.2 הוכחת משפט הSVD

הוכחת המשפט: המשפט שקול לכך שלכל מטריצה $A_{m\times n}$ ישנן מטריצות $U_{m\times m}$ ו$V_{n\times n}$ אורתוגונליות כך שU^TAV אלכסונית:

$$U^TAV = D \iff AV = UD \iff A = UDV^T$$

לכן נראה שלכל מטריצה A נוכל למצוא מטריצות אורתוגונליות כנ"ל.

נראה את המשפט באינדוקציה על גודל המטריצה: נניח שידוע לפתור עבור מטריצות בגודל $(m-1)\times(n-1)$, ונראה עבור מטריצה בגודל $m\times n$. נעבוד עם מטריצות בלוקים, כאשר נזכיר שמכפלה בין מטריצות בלוקים היא מטריצה המתקבלת מהכפלת הבלוקים.

בשלב הראשון, נמצא מטריצות אורתוגונליות U', V' כך שיתקיים:

$$U'^TAV' = \begin{bmatrix} \sigma_1 & | & 0 \\ 0 & | & B \end{bmatrix} = \begin{array}{c|c} \sigma_1 & 0 \\ \hline 0 & B \\ & \end{array}$$

כלומר, מטריצת בלוקים שהבלוק העליון שלה הוא σ_1, והבלוק התחתון הוא מטריצה כלשהי B.

בהנחה ויש לנו מטריצות U', V' שכאלה, נראה שאפשר להגיע למטריצה אלכסונית. לפי ההנחה יש לנו מטריצות U'', V'' כך שהמטריצה $(U")^TBV"$ היא אלכסונית. כעת, נכפיל את המטריצה שהתקבלה במטריצות $U''' = \begin{array}{c|c} 1 & \\ \hline & (U")^T \end{array}$, $V'''\begin{array}{c|c} 1 & \\ \hline & V" \end{array}$. אלה מטריצות אורתוגונליות, ואז יתקיים:

$$\begin{array}{c|c} 1 & \\ \hline & (U")^T \end{array} \cdot \begin{bmatrix} \sigma_1 & | & 0 \\ 0 & | & B \end{bmatrix} \cdot \begin{array}{c|c} 1 & \\ \hline & V" \end{array} = \begin{array}{c|c} \sigma_1 & \\ \hline & (U")^TBV" \end{array} = \begin{array}{c|c} \sigma_1 & \\ & .... \end{array}$$

לכן הכפלה של A ב$U'''^TU'^T$ וב$V'''V'$ משני הצדדים יתנו לנו מטריצה אלכסונית. כיוון שמכפלה של מטריצות אורתוגונליות היא אורתוגונלית גם כן, נקבל שבהנחה ומצאנו את U', V', סיימנו את ההוכחה.

את הפיתוח של המטריצות U', V' נבצע בשני שלבים:
בשלב הראשון, נמצא מטריצות אורתוגונליות U', V' שהמכפלה בהן תיתן לנו מטריצה מהצורה:

$$\begin{array}{c|c} \sigma_1 & W \\ \hline 0 & B \end{array}$$

כאשר B מטריצה וW וקטור שורה כלשהם ו$\sigma_1 = \|A\|_{op}$.

כיוון ש$\sigma_1 = \|A\|_{op}$, יש וקטורי יחידה u_1, v_1 כך ש$Av_1 = \sigma_1 u_1$, וσ_1 הוא המספר החיובי הגדול ביותר שיקיים זאת (נציין שכשנכתוב נורמה אופרטורית בלי לציין באיזה נורמה, נתכוון תמיד בנורמת l_2, כלומר במקרה הזה $\|A\|_{op} = \|A\|_{2,2}$).
את u_1 נוכל להשלים לבסיס אורתונורמלי למרחב, וכך גם את v_1. נגדיר:

$$U' = \left(u_1 \mid U_2 \right)$$

כלומר, U' היא המטריצה שעמודותיה הן הבסיס האורתונורמלי שקיבלנו מלהרחיב את u_1, וכך גם נגדיר:

$$V' = \left(v_1 \mid V_2 \right)$$

נטען שU', V' נותנים לנו את החלק הראשון. נחשב:

$$(U')^T AV' = \left(\frac{u_1^T}{U_2^T}\right) \cdot A \cdot \left(v_1 \mid V_2 \right) = \left(\frac{u_1^T}{U_2^T}\right) \cdot \left(Av_1 \mid AV_2 \right) = \left(\begin{array}{c|c} u_1^T A v_1 & u_1^T AV_2 \\ \hline U_2^T A v_1 & U_2^T AV_2 \end{array}\right)$$

לפי הבנייה, $\sigma_1 = \sigma_1 \|u_1\| = u_1^T \sigma_1 u_1 = u_1^T A v_1$. נסמן: $U_2^T AV_2 = B$, $u_1^T AV_2 = W$.
כמו כן, $Av_1 = \sigma_1 u_1$, ואז כיוון שהעמודות של U_2 הם וקטורים של בסיס אורתונורמלי, וu_1 וקטור אחר מהבסיס, נקבל $U_2^T A v_1 = \sigma_1 U_2^T u_1 = \begin{pmatrix} \sigma_1 < u_2, u_1 > \\ \dots \\ \sigma_1 < u_m, u_1 > \end{pmatrix} = 0$.

כלומר, אכן הגענו למטריצה מהצורה: $\left(\begin{array}{c|c} \sigma_1 & W \\ \hline 0 & B \end{array}\right)$

נותר לנו להגיע למצב של $\begin{bmatrix} \sigma_1 & | & 0 \\ 0 & | & B \end{bmatrix}$. נטען שאנחנו כבר שם: נניח בשלילה $W \neq 0$. עד כה לא השתמשנו בכך שσ_1 הוא הנורמה האופרטורית של A, אלא רק בכך שהוא ערך עצמי. נשתמש בכך עכשיו. נשים לב שהכפלה מימין ומשמאל במטריצה אורתונוגלית לא משנה את הנורמה האופרטורית, לכן $\|U'^T AV'\|_{op} = \|A\|_{op}$. נראה שקיים וקטור שהמטריצה הזו מותחת אותו יותר מפי σ_1, ובכך נקבל סתירה. נסתכל על הוקטור: $\begin{pmatrix} \sigma_1 \\ W^T \end{pmatrix}$. נחשב:

$$(U')^T AV' \cdot \begin{pmatrix} \sigma_1 \\ W^T \end{pmatrix} = \left(\begin{array}{c|c} \sigma_1 & W \\ \hline 0 & B \end{array}\right) \cdot \begin{pmatrix} \sigma_1 \\ W^T \end{pmatrix} = \begin{pmatrix} \sigma_1^2 + W \cdot W^T \\ \dots \end{pmatrix} = \begin{pmatrix} \sigma_1^2 + \|W\|_2^2 \\ \dots \end{pmatrix}$$

כעת, כיוון שהנורמה של וקטור היא לפחות כמו הקואורדינטה הראשונה שלו, נקבל שהוקטור $\begin{pmatrix} \sigma_1 \\ W^T \end{pmatrix}$ מועתק לוקטור שאורכו פי $\sqrt{\sigma_1^2 + \|W\|_2^2}$.
לפי הנחת השלילה $W \neq 0$ ולכן זהו גודל שהוא גדול ממש מσ_1 - סתירה לכך ש$\sigma_1 = \|A\|_{op}$ ■

2.7 קירובים של מטריצות

עד כה ראינו שאת הגודל של מטריצה ניתן למדוד על ידי שימוש בנורמה האופרטורית שלו. לא תמיד הנורמה האופרטורית תתפוס את מושג המרחק או האורך שהיינו רוצים, לכן יש גם נורמות נוספות על מטריצות.
אחד מהדברים שאפשר לעשות הוא לחשוב על מטריצה כעל וקטור: במקום לחשוב על האיברים של המטריצה כמסודרים במטריצה, נוכל "לשכוח" שהם מסודרים בשורות ועמודות ולחשוב עליהם כעל וקטור. הרעיון הזה פותח בפנינו את האפשרות להשתמש בכל נורמות l_p שהגדרנו עבור וקטורים.

במיוחד מעניין וחשוב הנורמה שמתקבלת מנורמת l_2 באופן הזה. הנורמה הזו נקראת **נורמת פרוביניוס**, והיא מוגדרת כך:

$$\|A\|_F = \sqrt{\sum_{i,j} a_{ij}^2}$$

לא קשה לראות שמתקיים:

$$\|A\|_F^2 = tr(AA^T)$$

טענה: נורמת פרובניוס לא מושפעת מהכפלה מימין ומשמאל במטריצות אורתונורמליות.

הוכחה: יהיו X, Y מטריצות ריבועיות. תכונה חשובה של העקבה היא שמתקיים:

$$tr(XY) = tr(YX)$$

(כמובן שלרוב $XY \neq YX$, אבל העקבות שלהן בכל זאת שוות).
תהי A מטריצה כלשהי, U מטריצה אורתונורמלית כך שAU מוגדר. אזי:

$$\|AU\|_F^2 = tr(AU(AU)^T) = tr(AUU^TA^T) = tr(AA^T) = \|A\|_F^2$$

כמו כן, אם V מטריצה אורתונורמלית כך שVA מוגדר, אזי:

$$\|VA\|_F^2 = tr(VA(VA)^T) = tr(VAA^TV^T) = tr(AA^TV^TV) = tr(AA^T) = \|A\|_F^2$$

■

הערה: חשוב לציין ש$tr(XY) = tr(YX)$ לא אומר שאם יש מכפלה של מספר מטריצות אפשר לסדר אותן כרצוננו - זה לא נכון לכל פרמוטציה של המטריצות אלא רק לשינויים מהסוג הזה.

השאלה הבאה שנעסוק בה היא כיצד ניתן לקרב מטריצה על ידי מטריצה אחרת. כלומר, יש לנו מטריצה A ממימד מסויים, ונרצה למצוא מטריצה אחרת, שהמימד שלה קטן ממש מהמימד של A אבל שהנורמה שלה קרובה לזו של A.
בהינתן מטריצה A ומספר k, נרצה להסתכל על כל המטריצות שדרגתן לכל היותר k, ולמצוא את המטריצה שמרחקה מA מינימלי, כלומר נחפש את המטריצה שנותנת את הערך:

$$\min_{rank(B)\leq k} \|A - B\|_{op}$$

ונרצה גם מטריצה שתיתן את הערך:

$$\min_{rank(B)\leq k} \|A - B\|_F$$

בפעם הבאה נראה שלמעשה זו אותה המטריצה B שנותנת קירוב מירבי לA בנורמה האופרטורית וגם בנורמת פרובניוס, ונראה כיצד היא מתקבלת.

שבוע 7

משפט: תהי $A_{m\times n}$ מטריצה ממשית, ויהי $k < n, m$ מספר טבעי. אז קיימת מטריצה B שנותנת את $\min_{rankB\leq k}\|A - B\|_{op}$ ואת $\min_{rankB\leq k}\|A - B\|_F$ ומקיימת $rankB \leq k$.
המטריצה B מתקבלת באופן הבא: אם UDV^T הוא הפירוק של A המתקבל ממשפט SVD, אז $B = UD^{(k)}V^T$, כאשר $D^{(k)}$ היא המטריצה המתקבלת מD על ידי איפוס המקומות החל מ$k+1$.

כלומר, החלק הראשון של המשפט מבטיח לנו שאם נסתכל על קירובים של A על ידי מטריצות מדרגה לכל היותר k, קיימת מטריצה שמקבלת את הערכים המינימליים הללו. מה שמעניין הוא שהמטריצה המקרבת שמתקבלת היא אותה המטריצה בשתי הנורמות. יתר על כן, אנחנו יודעים בדיוק איך היא מתקבלת: אם $D = diag(\sigma_1, \sigma_2, .., \sigma_n)$, (כאשר $\sigma_1 \geq \sigma_2 \geq .. \geq \sigma_n$ כרגיל) אז $D^{(k)} = diag(\sigma_1, ..., \sigma_k, 0, .., 0)$, ואז $B = UD^{(k)}V^T$.

מה המוטיבציה לנסות לתת קירובים למטריצה על ידי מטריצות מדרגה לכל היותר k? בליניארית הגדרנו את הדרגה של מטריצה בתור המימד של תת המרחב הנפרש מהעמודות של המטריצה (או השורות שלה, ראינו ששני המרחבים יהיו בעלי אותו המימד). מסתבר שהגדרה שקולה למושג הדרגה הוא:

$$rankM = \min_d \{d : \exists X_{m\times d}, Y_{d\times n},\ XY = M\}$$

כלומר, הדרגה של המטריצה היא המספר המינימלי שאפשר למצוא שתי מטריצות שמכפלתן תיתן את המטריצה. אפשר לחשוב על הרעיון כמו "לקפוץ דרך טבעת": יש לנו את המטריצה M, ואת כל המידע שהיא מכילה. כשאנחנו מסתכלים על מטריצות $X_{m\times d}$ ו$Y_{d\times n}$ המקיימות $XY = M$, אנחנו במובן מסויים דוחסים את המידע של M דרך ה"טבעת" d- אנחנו מחייבים את המידע לעבור דרך מספר קטן. הרעיון הוא שאנחנו יכולים לעבור דרך מימד נמוך ולא לאבד מידע. זה מקושר גם לרעיון של התלות הליניארית בין העמודות: הבסיס עבור העמודות בוודאי מכיל את כל המידע לגביהן, ולכן ככל שיש יותר עמודות תלויות ליניארית יש פחות מידע לספק.

דרך נוספת שקולה לחשוב על דרגה היא בתור ה"סיבוכיות" של המטריצה, עד כמה היא מורכבת. נוכל להתחיל מלהגדיר אובייקטים פשוטים שיהיו אבני הבניין של האובייקטים שלנו, ואז הסיבוכיות תהיה מדד של כמה אבני בניין נצרכות על מנת לקבל את האובייקט שלנו. במקרה של מטריצות, אבני הבניין יהיו המטריצות הפשוטות ביותר: המטריצות מדרגה 1. מטריצה מדרגה 1 יכולה להתקבל כ**מכפלה חיצונית** של שני וקטורים (בהינתן שני וקטורים, המכפלה החיצונית שלהם היא מטריצה המתקבלת מהמכפלה של כל איבר בוקטור הראשון בכל איבר בוקטור השני. מבחינה של כתיב, $(u\otimes v)_{ij} = u_i v_j$). לכן אפשר לחשוב על מטריצה מדרגה d כך שהיא מתקבלת כסכום של d מטריצות מדרגה 1.

כלומר, הסיבה שנרצה מלכתחילה להסתכל על קירובים של מטריצה A על ידי מטריצות מדרגות נמוכות יותר היא שקירוב כזה מאפשר לנו להתקרב לאובייקט המורכב (A) עם אובייקט פשוט יותר (B).

נוכיח את המשפט:

הוכחה: נסתכל על הפירוק של A המתקבל ממשפט הSVD, $A = UDV^T$ כאשר $U\in O(m)$, $V\in O(n)$ (אורתוגונליות בגודל $m\times m$ ו$n\times n$), ו$D = diag(\sigma_1, .., \sigma_n)$, כאשר אלה הערכים הסינגולריים של A.

נגדיר את $D^{(k)}$ כמו במשפט: $D^{(k)} = diag(\sigma_1, .., \sigma_k, 0, .., 0)$, כלומר היא המטריצה האלכסונית שמתקבלת מD על ידי החלפה של כל σ_i עבורו $i\geq k+1$ ב0.

כפי שאמרנו בשיעור שעבר:

$$A = \sum_{i=1}^{min(m,n)} \sigma_i u_i \otimes v_i$$

(כאשר u_i וv_i הם וקטורים כמובן).

נגדיר כפי שרצינו $B = UD^{(k)}V^T$. אז B הוא כמעט A, אלא שכל האיברים שמוכפלים בσ_i כאשר $i\geq k+1$ מתאפסים. לכן:

$$B = \sum_{i=1}^{k} \sigma_i u_i \otimes v_i$$

נצטרך להראות ש$rankB\leq k$ ושB נותנת את האופטימזציה הדרושה. כיוון שממש קיבלנו את B כסכום של k מטריצות מדרגה 1, אנחנו רואים שאכן מתקיים $rank(B)\leq k$.

לגבי האופטימיזציה, נראה זאת עבור הנורמה האופרטורית, ולנורמת פרובניוס ההוכחה דומה.

באופן כללי, הנורמה האופרטורית של מטריצה הוא הערך הסינגולרי הגדול ביותר שלה. כיוון שעבור המטריצה $A-B$, כל הערכים הסינגולריים עם אינדקס קטן מk מתאפסים, נסיק שהערך הסינגולרי הגדול ביותר של $A-B$ הוא σ_{k+1}, ולכן:

$$\|A-B\|_{op} = \sigma_{k+1}$$

לכן נצטרך להראות שאם C מטריצה עבורה $rank(C)\leq k$, $\|A-C\|_{op}\geq \sigma_{k+1}$, כלומר שלא ניתן לקבל ערך טוב יותר ממה שB נותנת לנו.

עבור מטריצה C כנ"ל, $\|A-C\|_{op}\geq \sigma_{k+1}$, משמעו לפי ההגדרה שיש וקטור יחידה z כך ש$\|(A-C)z\|_2\geq \sigma_{k+1}$ (הערה: באופן כללי אם לא נציין עם איזה נורמה אנחנו עובדים זו תהיה נורמת l_2). לכן על מנת להראות את הדרוש, נצטרך למצוא וקטור z שכזה.

על C אין לנו הרבה מידע, אבל את A אנחנו כן "מכירים". לכן נרצה להסתכל על $z\in ker(C)$ - אם נבחר z שכזה, $(A-C)z = Az - Cz = Az - 0 = Az$, כלומר נצטרך להסתכל רק על $\|Az\|$ ובכך קיבלנו בעיה פשוטה יותר.

ידוע ש$rank(C) \leq k$, לכן נרצה להשתמש בכך: נסתכל על $span(v_1, .., v_{k+1})$ ($k+1$ הוקטורים הראשונים של המטריצה V, שמהווים בסיס אורתונורמלי). זהו מרחב $k+1$ מימדי. כמו כן, כיוון ש$rank(C) \leq k$, אנחנו יודעים $dimker(C) \geq n-k$. נשתמש באבחנה מאלגברה ליניארית: אם שני תת מרחבים הם כך שסכום המימדים שלהם גדול מהמימד של המרחב כולו, יש למרחבים חיתוך שמכיל וקטור שונה מאפס. במקרה הזה, קיים $z \neq 0$ כך ש$z \in ker(C)$ וגם $z \in span(v_1, .., v_{k+1})$.

נוכל לנרמל את הוקטור z שהתקבל: נניח $\|z\| = 1$, ואז $z = \sum_{i=1}^{k+1} \alpha_i v_i$, כאשר זה הייצוג של z לפי הבסיס האורתונורמלי $v_1, .., v_n$, וכן $Cz = 0$.

נצטרך להראות אם כן $\|Az\| \geq \sigma_{k+1}$. לפי ההגדרה:

$$Az = \sum_{i=1}^{min(m,n)} \sigma_i < u_i \otimes v_i, z >$$

(נשים לב שאנחנו מכפילים כאן מטריצה מדרגה 1 בוקטור במקום מכפלה פנימית רגילה, אבל נתיחס לזה באותו האופן - נרשה לעצמינו לכתוב את המכפלה כך בכל זאת. אפשר לפתוח את ההגדרה ולהשתכנע שמה שקיבלנו אכן הגיוני).
נוכל להוציא את u_i מחוץ למכפלה ולקבל:

$$Az = \sum_{i=1}^{min(m,n)} \sigma_i u_i < v_i, z >$$

כעת, $z \in span(v_1, .., v_{k+1})$ ומדובר במשפחה אורתוגונלית, לכן המכפלה הפנימית של z עם כל וקטור v_i עבור $i > k+1$ יתאפס:

$$= \sum_{i=1}^{k+1} \sigma_i u_i < v_i, z > = \sum_{i=1}^{k+1} \sigma_i < v_i, z > u_i$$

כלומר, כרגע אנחנו יודעים איך נראה הוקטור Az. מה שבאמת מעניין אותנו זה הנורמה שלו, $\|Az\|$. יהיה לנו יותר קל להסתכל על $\|Az\|^2$.
כיוון שיש לנו כאן בעצם וקטור שמיוצג בבסיס $u_1, .., u_m$, כלומר יש לנו וקטור שמיוצג בבסיס אורתונורמלי, לפי משפט פיתגורס (נורמה בריבוע שווה לסכום של ריבועי המקדמים) נוכל לקבל:

$$\|Az\|^2 = \sum_{i=1}^{k+1} \sigma_i^2 < v_i, z >^2$$

לכל $1 \leq i \leq k+1$, $\sigma_i^2 \geq \sigma_{k+1}^2$, וכל המחוברים אי שליליים לכן:

$$\geq \sigma_{k+1}^2 \sum_{i=1}^{k+1} < v_i, z >^2$$

ולפי משפט פיתגורס שוב, כיוון ש$< v_i, z >$ הם המקדמים של z לפי הבסיס האורתונורמלי $v_1, .., v_n$, נקבל:

$$= \sigma_{k+1}^2 \|z\|^2 = \sigma_{k+1}^2 \cdot 1 = \sigma_{k+1}^2$$

לסיכום, קיבלנו שמצד אחד B מקיימת $rank(B) \leq k$ ו$\|A - B\|_{op} = \sigma_{k+1}$, ומצד שני לכל מטריצה C המקיימת $rank(C) \leq k$, $\|A - C\|_{op} \geq \sigma_{k+1}$, ולכן נסיק שאכן B נותנת את המרחק המינימלי בנורמה האופרטורית מבין הקירובים על ידי מטריצות מדרגה k כנדרש ■

2.8 התיאור הווריאציוני של ערכים עצמיים

עד כה דיברנו על ערכים עצמיים מהזוית של משפט הSVD. הוא נתן לנו דרך מצויינת להפוך העתקות ליניאריות כלליות (שהן אובייקטים מופשטים) לאובייקט הרבה יותר קונקרטי: משפט הSVD איפשר לנו לתאר כל העתקה ליניארית על ידי סיבובים ומתיחות, שאלה העתקות ליניאריות קונקרטיות מאוד.
כיוון שהעתקות ליניאריות הן אובייקט כל כך חשוב, יש עוד הרבה פרספקטיביות מעניינות מהן אפשר לתאר אותו, ובחלק הזה נפתח אחת מהן.

2.8.1 משפט $Rayleigh-Ritz$

משפט Rayleigh-Ritz: תהי A מטריצה ממשית סימטרית, ויהיו $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n$ הערכים העצמיים של A. אז מתקיים:

.1

$$\lambda_1 = \max \frac{xAx^T}{\|x\|^2} = \max_{\|x\|=1} xAx^T$$

(ניתן לראות שכפל בסקלר לא ישנה למנה ולכן מספיק להסתכל רק על וקטורי יחידה)

.2 באופן כללי,

$$\lambda_{k+1} = \max_{x \perp x_1,..,x_k} \frac{xAx^T}{\|x\|^2}$$

כלומר, בהינתן מטריצה ממשית סימטרית, אנחנו יודעים בדיוק מי הם הערכים העצמיים שלה (נזכיר שמטריצה ממשית סימטרית היא לכסינה ולכן בפרט כל הערכים העצמיים שלה ממשיים). המשפט אומר לנו שλ_1 הוא הערך המקסימלי שיכול להתקבל לתבנית הבילינארית xAx^T. עבור λ_{k+1}, בהינתן שבחרו כבר $x_1,..,x_k$ כשלהם שנותנים לנו את הערך המקסימלי עבור $\lambda_1,..,\lambda_k$, λ_{k+1} מתקבל על ידי לקיחת המקסימום על גבי כל הוקטורים שמאונכים להם. מתבצע כאן תהליך איטרטיבי: נוכל לאפיין את λ_{k+1} בהינתן שסיימנו לאפיין את $\lambda_1,..,\lambda_k$. הוקטורים $x_1,.,x_k$ שנבחר לא יהיו יחידים, אבל בכל זאת הערך יהיה מוגדר היטב, מסיבות שלא נעמיק בהן.

הערה: באופן דומה אפשר לאפיין מהסוף:

$$\lambda_n = \min \frac{xAx^T}{\|x\|^2}$$

וכן בהנחה שמצאנו כבר את $x_n,..,x_{n-k}$,

$$\lambda_{n-k-1} = \min_{x \perp x_n,..,x_{n-k}} \frac{xAx^T}{\|x\|^2}$$

הוכחה ראשונית: בשלב ראשון נראה את המשפט רק עבור λ_1. אפשר להראות בצורה דומה למקרה הכללי גם כן, אבל זה דורש שימוש בטכניקת כופלי לגרנג' שלא נלמד כאן.

נסתכל על הפונקציה שמתקימה לכל וקטור x את $\frac{xAx^T}{\|x\|^2}$, על ספירת היחידה. ספירת היחידה היא **קבוצה קומפקטית** (קבוצה קומפקטית הוא מושג שהתעסקנו בו באינפי 3 וטופולוגיה. הדבר שבעיקר חשוב לנו לדעת עבור קבוצות קומפקטיות הוא שפונקציה רציפה המוגדרת על קבוצה קומפקטית מקבלת עליה מקסימום. חשוב גם לדעת שב$\mathbb{R}^n$ קבוצה קומפקטית זו קבוצה סגורה וחסומה, כאשר קבוצה תיקרא סגורה אם לכל סדרה שמתכנסת בקבוצה, גבול הסדרה גם כן נמצא בסדרה - ב$\mathbb{R}^n$ לרוב קבוצה סגורה תהיה מוגדרת על ידי תנאי כמו $... \leq x$, או $x \geq ...$). הפונקציה שלנו היא רציפה, ולכן מקבלת מקסימום על ספירת היחידה.

הטענה שלנו היא שהמקסימום הזה הוא בדיוק λ_1. כלומר, השאלה שלנו היא בסך הכל למצוא נקודת מקסימום לפונקציה רציפה $\Leftarrow$ ואנחנו יודעים כיצד לעשות זאת בכלים של אינפי: נגזור, נשווה לאפס, ונמצא כך מי הן נקודות המקסימום של הפונקציה. נשים לב שכאן יש לנו פונקציה בn משתנים, לכן נצטרך להראות בעצם שכל הנגזרות החלקיות מתאפסות.

עבור מנה של פונקציות $\frac{f}{g}$, הנגזרת מתאפסת אם ורק אם $f'g - fg' = 0$, כלומר אם ורק אם $f'g = fg'$. לכן זה התנאי שנצטרך לבדוק.
עבור וקטור y כלשהו,

$$\|y\|^2 = \sum_{j=1}^{n} y_j^2$$

לפי ההגדרה. לכן לכל $1 \leq i \leq n$, כשנגזור לפי הקואורדינטה הi, בסך הכל אנחנו גוזרים פולינום לפי אחד מהמשתנים שלו, ואז נקבל:

$$\frac{\partial}{\partial y_i}\|y\|^2 = 2y_i$$

כיוון שרק המחובר y_i^2 נשאר, וכל השאר הם קבועים כשנחשוב על $\|y\|^2$ כפונקציה בy_i, ולכן מתבטלים.

באותו אופן,

$$yAy^T = \sum_{j,k} a_{jk} y_j y_k$$

כשנגזור את הפולינום הזה לפי y_i, רק המחוברים עם y_i "ישרדו", ונישאר רק עם המקדמים:

$$\frac{\partial}{\partial y_i} yAy^T = 2\sum_k a_{ik} y_k = 2(Ay)_i$$

כאשר ה2 הגיע כיוון שA מטריצה סימטרית, לכן $a_{jk} = a_{kj}$ ומספיק לסרוק רק את k.

סך הכל, לכל $1 \le i \le n$,

$$2(Ay)_i \|y\|^2 = 2y_i(yAy^T) \iff \forall i,\ Ay_i = \frac{yAy^T}{\|y\|^2} y_i \iff Ay = \frac{yAy^T}{\|y\|^2} y$$

כיוון שהשיוויון נכון בכל קואורדינטה, אפשר לומר שהוא מתקיים לוקטורים עצמם.

כלומר, y הוא וקטור עצמי של A, השייך לערך עצמי $\frac{yAy^T}{\|y\|^2}$. כלומר, כל הנקודות הקריטיות של הפונקציה הם וקטורים עצמיים של A עם הערך העצמי הזה, ולכן מצאנו מאפיין לכל נקודות הקיצון של הפונקציה.
נצטרך להראות עדיין שהמקסימום הוא אכן מי שנותן לנו את λ_1, אבל נראה זאת כשנראה את ההוכחה של המשפט בשלמותה, שנעשה עוד מעט.

כעת נוכיח את המשפט בשלמותו:
הוכחה: כיוון שA מטריצה ממשית סימטרית, לפי המשפט הספקטרלי היא לכסינה ולכן:

$$A = V\Lambda V^T$$

כאשר V אורתוגונלית, Λ אלכסונית, העמודות של V הם הוקטורים העצמיים של A, וΛ היא כך שהערכים העצמיים נמצאים על האלכסון. כרגיל נניח שהערכים העצמיים מסודרים בסדר יורד.
בהינתן וקטור x כלשהו:

$$xAx^T = xV\Lambda V^T x^T$$

נחשוב על זה כמכפלה של xV עם Λ, ואז נקבל:

$$= \sum_i \lambda_i (xV)_i^2$$

(כיוון שעבור מטריצה אלכסונית B ווקטור y כלשהו:

$$yBy^T = \sum_{i,j} b_{ij} y_i y_j = \sum_i b_{ii} y_i y_i = \sum_i b_{ii} y_i^2$$

שכן מדובר במטריצה אלכסונית וכל האיברים שאינם על האלכסון מתבטלים).

כיוון שהערכים העצמיים מסודרים לפי הסדר,

$$xAx^T = \sum_i \lambda_i (xV)_i^2 \le \lambda_1 \sum_i (xV)_i^2 = \lambda_1 \|xV\|^2 = \lambda_1 \|x\|^2$$

כאשר השיוויון הלפני אחרון נכון לפי ההגדרה של הנורמה, והשיוויון האחרון נובע מכך שV מטריצה אורתוגונלית. כלומר, λ_1 הוא אכן חסם על הפונקציה הזו. נותר להראות שאפשר למצוא וקטור כלשהו שישי את הערך הזה, ולפי מה ההוכחה הראשונית שעשינו למעלה נדע לחפש אותו בין וקטורים עצמיים של λ_1.

באותו אופן בדיוק אפשר להראות במקרה הכללי: אם $x\perp v_1,..,v_k$, (נראה בהמשך שנרצה לבחור את $x_1,..,x_k$ להיות בדיוק הוקטורים העצמיים המתאימים, כלומר את $v_1,..,v_k$)

$$xAx^T=\sum_{i=1}^{n}\lambda_i(xV_i)_i^2=\sum_{i=k+1}^{n}\lambda_i(xV_i)^2\leq\lambda_{k+1}\sum_i(xV)_i^2=\lambda_{k+1}\|x\|^2$$

כלומר, קיבלנו שלכל וקטור x, $xAx^T\leq\lambda_1\|x\|^2 \Leftarrow \frac{xAx^T}{\|x\|^2}\leq\lambda_1$.

נראה את השיוויון: אם $x_1=v_1$, אז הוא וקטור עצמי המתאים לערך עצמי λ_1, ולכן:

$$xAx^T=\lambda_1\|x\|^2$$

באותו אופן, אם $x\perp x_1,..,x_k$, אז נבחר את x להיות הוקטור העצמי של λ_{k+1}, ואז

$$xAx^T=\lambda_{k+1}\|x\|^2$$

■

מבחינה רעיונית, מה שקורה כאן הוא שהתחלנו מלהסתכל על הערכים של xAx^T על ספירת היחידה. לאחר מכן, בחרנו רק את הוקטורים שהם מאונכים לוקטור הראשון שבחרנו, מה שיתן לנו את החיתוך של ספירת היחידה עם מישור כלשהו, למעשה את ספירת היחידה במימד אחד פחות. כלומר, בתהליך שעשינו הסתכלנו על xים שנבחרו מתוך תת מרחבים ספציפיים מאוד. נשאלת השאלה, מה היה קורה אם היינו מסתכלים על תת מרחבים אחרים במימדים יורדים, מה היה מתקבל אז? התשובה המפתיעה אולי היא שהאפיון שקיבלנו למעלה על ידי תת מרחבים של ספירת היחידה הוא האופטימלי:

משפט Courant-Fischer: תהי $A_{n\times n}$ מטריצה ממשית סימטרית, ויהיו $\lambda_1\geq...\geq\lambda_n$ הערכים העצמיים שלה. אז לכל i:

$$\lambda_{i+1}=\min_{F\text{ is a subspace, dimF=i}}\max_{x\perp F}\frac{xAx^T}{\|x\|^2}$$

וכן:

$$\lambda_{i+1}=\max_{F\text{ is a subspace,}dimF=n-i-1}\min_{x\perp F}\frac{xAx^T}{\|x\|^2}$$

כלומר, λ_{i+1} מתקבל כמינימום (בגרסא הראשונה) על גבי הערכים שמתקבלים מכל הבחירות של תת מרחבים F - ואילו כשהסתכלנו על ספירת היחידה אכן קיבלנו אותו בדיוק, כלומר הסדרה שקיבלנו של תתי מרחבים של ספירת היחידה היתה הסדרה שאכן נותנת את הערך המינימלי מבין כל הסדרות האפשריות של תתי מרחבים.

מסקנה - משפט הinterlacing: תהי $A_{n\times n}$ מטריצה ממשית, סימטרית, עם ערכים עצמיים $\lambda_1\geq..\geq\lambda_n$, ונניח ש$B$ מתקבלת ממנה על ידי מחיקת השורה הi והעמודה הi, לאיזשהו אינדקס i. B היא גם כן ממשית וסימטרית, $B_{(n-1)\times(n-1)}$. נסמן את הערכים העצמיים של B ב$\mu_1\geq..\geq\mu_{n-1}$, ואז:

$$\lambda_1\geq\mu_1\geq\lambda_2\geq...\geq\lambda_{n-1}\geq\mu_{n-1}\geq\lambda_n$$

כלומר, הערכים העצמיים של B "נכנסים בין" הערכים העצמיים של A.

הרעיון בהוכחת המסקנה היא שבמקום לחשוב על מחיקה של שורה ועמודה, נסתכל רק על אוסף הוקטורים שיש להם 0 בקואורדינטה הi. זה יתן לנו אוסף מאוד מסויים של וקטורים, כלומר קיבלנו תת מרחב מאוד מסויים איתו אנחנו עובדים.

2.8.2 משפט פרון־פרובניוס

הגדרה: נאמר שמטריצה A היא **מטריצה אי שלילית** אם כל האיברים שלה אי שליליים, כלומר לכל i,j, $A_{ij} \geq 0$.
באותו אופן, נאמר שA היא **מטריצה חיובית** אם כל האיברים שלה חיוביים, כלומר $A_{ij} > 0$ לכל i,j

חשוב שלא להתבלבל עם מטריצות מוגדרות חיובית או מוגדרות אי שלילית, מה שסימנו כמטריצות PD או PSD.

המשפט יאמר לנו שאפשר להגיד דברים מעניינים על מטריצה אי שלילית או חיובית. במובן מסויים זה מפתיע, כיוון שעל פניו לא נראה שעצם זה שהאיברים של המטריצה הם אי שליליים או חיוביים מחייב איזשהו מבנה על המטריצה.

לא נוכיח את המשפט בשלמותו, אלא רק נראה חלקים ממנו.

הגדרה: **שרשרת מרקוב** היא מערכת שתתואר על ידי n מצבים. לכל זוג מצבים $1 \leq i,j \leq n$, אם המערכת נמצאת במצב i, ההסתברות שהיא תעבור למצב j נתונה על ידי $p_{ij} \geq 0$.
המערכת היא חסרת זכרון ־ השאלה באיזה מצב נהיה ביחידת הזמן הבאה תלויה רק בi וj, ולא במצבים שהיינו בהם ביחידות זמן אחרות.

בהינתן שרשרת מרקוב, נוכל לקבל **מטריצת מעבר** של השרשרת. זו המטריצה P שאיבריה הם p_{ij}. P תקיים את התכונות הבאות:

1. לכל i,j, $p_{ij} \geq 0$, כלומר היא מטריצה אי שלילית.
2. לכל i, $\sum_j p_{ij} = 1$

באופן כללי, מטריצה שתקיים את התכונות 1 ו2 תיקרא מטריצה סטוכסטית.

נשים לב שהמטריצה P היא לא בהכרח סימטרית, ולכן לא מובטח לנו שכל הערכים העצמיים שלה יהיו ממשיים ־ יתכן ויש לה ערכים עצמיים מרוכבים.

יש כל מני שאלות שנרצה לשאול עבור המטריצה שהתקבלה כך, נרצה להבין אותה יותר.

באופן כללי נציין שיש יותר דברים מעניינים שאפשר לומר על מטריצה חיובית מאשר על מטריצה אי שלילית (שזה הגיוני, כיוון שמטריצה חיובית נותנת לנו יותר מידע על הקואורדינטות).

במשפט פרון־פרובניוס, פרון אמר משהו על מטריצה שכל איבריה חיוביים, ואז פרובניוס הכליל את האמירה למטריצה אי שלילית.

איך יכולה לעבוד רדוקציה כזו? נניח שאנחנו יודעים להפעיל את משפט פרון (שלא אמרנו עדיין מה הוא אומר) על מטריצה חיובית. כעת נתונה לנו מטריצה אי שלילית. נשאלת השאלה, כיצד נוכל להגיע למטריצה חיובית ולהפעיל עליה את המשפט? כמובן שנרצה לדעת גם **האם** בכלל אפשר להגיע למטריצה חיובית על מנת להפעיל עליה את המשפט.
המשפט ידבר על וקטורים עצמיים וערכים עצמיים, לכן דבר ראשון שנשים לב אליו הוא שאם u וקטור עצמי של A המתאים לע"ע λ, אז u הוא ו"ע של A^k המתאים לע"ע λ^k, לכל $k \geq 1$. הדבר הזה מוביל לשאלה ־ אם A היא אי שלילית, האם יכול להיות שA^k היא חיובית? באיזה תנאים נוכל להבטיח זאת?

לכן המטרה שלנו תהיה למצוא $k \geq 1$ כך שA^k תהיה חיובית, אם קיים כזה כמובן.

קודם כל נשים לב שכאשר מכפילים את A בעצמו, נקבל בקואורדינטה מסויימת ערך 0 או ערך חיובי, לכן לא באמת אכפת לנו מה הערכים המספריים עצמם. כלומר, נוכל להחליף את A במטריצה M כלשהי, שמתקבלת ממנה באופן הבא:

$$M_{ij} = \begin{cases} 1 & A_{ij} > 0 \\ 0 & A_{ij} = 0 \end{cases}$$

בגלל שהדבר היחיד שיכול להשפיע על השאלה מה יהיה בA_{ij}^k הוא האם האיברים המתאימים הם 0 או לא, כאמור לא אכפת לנו מהערך המדוייק שלהם, אפשר לראות ש$A^k > 0 \iff M^k > 0$ (כאשר הכוונה בסימון הזה היא שהמטריצה חיובית).

נשים לב שהמטריצה M היא מטריצה ריבועית, שבכל מקום יש לה 0 או 1. לכן היא למעשה מתארת גרף מכוון (היא לא בהכרח סימטרית). נסמן את הגרף המתקבל מM בG, ואז אנחנו יכולים להעביר שאלות על M ללהיות שאלות על G.
למשל, אנחנו יודעים ש$(M^k)_{ij} > 0$ אם ורק אם יש מסלול בין i לj באורך k בגרף G. לכן בעצם אנחנו רוצים לשאול האם קיים k עבורו יהיה מסלול באורך k בין כל שני קודקודים בגרף.

ניתן לראות מתי זה **לא** יתקיים:
טענה: אם G איננו קשיר חזק, לא נוכל למצוא k כזה
זה נכון בגלל שאם G איננו קשיר חזק, אז יש קודקודים i, j כך שאין מסלול בין i לj, ובפרט אין מסלול מכל אורך, ולכן לכל k, $(M^k)_{ij} = 0$.
במקרה כזה לא נוכל לעבור למטריצה חיובית.

מצב אחר בו זה לא מתקיים הוא קצת יותר מורכב:
נאמר שלגרף G יש **מבנה ציקלי** אם קיימת חלוקה של הקודקודים, כלומר קיימות קבוצות $C_1, .., C_l$ של קודקודים כך שכל קודקוד נמצא בבדיוק קבוצה C_i יחידה, עבורה מתקיים הדבר הבא: אם עבור $1 \leq i < l$, יש צלע היוצאת מהרכיב C_i, אז היא בהכרח מקשרת אותו לC_{i+1}, ואם יש צלע היוצאת מהרכיב C_l, היא בהכרח מקשרת אותו לC_1.
כלומר, גרף ציקלי יכול להיות קשיר, אבל יש קבוצות קודקודים שלא מחוברות ישירות אחת לשנייה.

טענה: אם לG יש מבנה ציקלי, לא נוכל למצוא k כזה.
נשים לב שאם לG יש מבנה ציקלי, מסלול באורך 1 בגרף יוכל לקשר קודקוד בC_i רק לקודקוד C_{i+1} (ואולי בין קודקודים בתוך אותו הרכיב) - למשל, לא נוכל לקשר בין קודקוד בC_2 לקודקוד ב C_4 עם מסלול באורך 1.
אותו הדבר יקרה עם מסלול באורך k לכל k: הוא יוכל לקשר רק בין קודקודים הנמצאים ברכיבים מסויימים, ובפרט לא נוכל למצוא k המחבר בין כל הרכיבים. לכן גם במקרה הזה לא נוכל לעבור למטריצה חיובית.

כלומר, ראינו ששני תנאים הכרחיים לכך שיהיה אפשר לעבור למטריצה חיובית הם שהגרף שמתואר על ידי M יהיה קשיר חזק ולא יהיה לו מבנה ציקלי. מסתבר שאלו הן ההפרעות היחידות, כלומר:
טענה: אם A היא מטריצה ריבועית אי שלילית כך שהגרף G המתקבל ממנה הוא קשיר חזק ואין לו מבנה ציקלי, קיים $k \geq 1$ כך שA^k היא מטריצה חיובית.

נאמר ששרשרת מרקוב היא **ארגודית** אם למטריצת המעבר שלה A יש את שתי התכונות האלה.

המשפט העיקרי שנוכיח יהיה לשרשראות מרקוב ארגודיות. חשוב לציין שהמשפט המלא מתמודד גם עם שרשראות שאינן ארגודיות, אבל זה מקרה מסובך ומצד שני פחות מעניין.

כלומר, עד כה קיבלנו את היכולת לעשות רדוקציה מהמקרה של מטריצה אי שלילית למקרה של מטריצה חיובית עליה נוכל להפעיל את משפט פרון, במידה שהמטריצה מקיימת את שתי התכונות הנ"ל.

משפט פרון: תהי $A_{n\times n}$ מטריצה חיובית. אזי:

1. לA יש ערך עצמי חיובי, ובפרט ממשי λ_{max}, כך שלכל ערך עצמי אחר λ, $|\lambda| < \lambda_{max}$ (במקרה הכללי ערך עצמי λ שכזה יהיה מרוכב).
2. λ_{max} הוא ערך עצמי פשוט, כלומר יש לו ריבוי אלגברי וגאומטרי של 1. כמו כן, מתאים לו וקטור עצמי חיובי (כאשר וקטור הוא חיובי אם כל הקואורדינטות שלו חיוביות)

3. אם x הוא הוקטור העצמי שמתאים לλ_{max}, אז כל וקטור עצמי חיובי אחר של A הוא כפולה של x

לא נראה את המשפט במלואו כרגע. נמשיך בשיעור הבא, כרגע נראה חלקים מבין הטענות הללו:
טענה: אם $u \geq 0$, $u \neq 0$, אז $Au > 0$.
לפי הנתונים, $(Au)_i$ הוא מכפלה של שורה מA עם u. השורה של A מכילה איברים חיוביים בלבד, נתון שu מכיל איברים אי שליליים בלבד וגם מכיל איבר חיובי ממש באחת מהקואורדינטות שלו, לכן הטענה מתקבלת.

טענה: לA יש וקטור עצמי חיובי עם ערך עצמי חיובי
נתאר את הרעיון הכללי כרגע: מכפלה בסקלר לא משפיעה על השאלה האם וקטור הוא וקטור עצמי, ולא על השאלה לאיזה ערך עצמי הוא שייך, לכן נוכל להסתכל רק על וקטורים עצמיים עם נורמה 1. כמו כן, יעניינו אותנו רק וקטורים עצמיים אי שליליים. נגדיר את S להיות קבוצת הוקטורים שהנורמה שלהם 1 וכל הקואורדינטות שלהם אי שליליות.

אנחנו לא ממש יודעים איך למצוא וקטור עצמי לA, לכן נגדיר קירוב מסויים לכך. נגדיר את הפונקציה L על S באופן הבא:

$$L(x) = \min_i\{\frac{(Ax)_i}{x_i} : x_i \neq 0\}$$

כלומר, בהינתן וקטור אי שלילי x נסתכל עבור כל הקואורדינטות החיוביות ממש שלו, ונסתכל על היחס שבין $(Ax)_i$ לx_i. נתאים לx את הערך המינימלי שהתקבל מכל הקואורדינטות שלו.
אילו היינו יודעים שיש וקטור עצמי חיובי שמתאים לערך עצמי חיובי, היינו מקבלים שקיים וקטור עבורו כל המנות האלה הן אותו הדבר, כלומר לכל i $\lambda = \frac{(Ax)_i}{x_i}$, ונרצה לקבל שהערך הזה יהיה λ_{max}.
יש לנו כרגע שתי בעיות עם הבנייה הזו: אנחנו מתעלמים מקואורדינטות שהן 0 בוקטור, ואנחנו לא יודעים שכל המספרים יצאו אותו הדבר - L מחזירה לנו רק את המינימלי.
מה שכן, הפונקציה L היא רציפה, ולכן תקבל מקסימום על S. בשיעור הבא נראה שהמקסימום הזה הוא וקטור עצמי.

שבוע 8

הגדרה: תהי A מטריצה ריבועית. **הרדיוס הספקטרלי** של A הוא הערך המוחלט המקסימלי של ערך עצמי של A. כלומר, אם $\lambda_1, .., \lambda_n$ הם הערכים העצמיים של A (נשים לב שהם לא בהכרח ממשיים כיוון שלא הנחנו שA סימטרית), אז הערך המקסימלי שמקיים: $\lambda = \max_{1\leq i\leq n} |\lambda_i|$ יקרא הרדיוס הספקטרלי של A.

בשיעור שעבר ניסחנו את משפט פרון:
משפט פרון: תהיה $A_{n\times n}$ מטריצה חיובית. אז הרדיוס הספקטרלי של A $\rho > 0$ (ובפרט ממשי), וכן מתקיים:

1. ρ הוא ערך עצמי פשוט (כלומר יש לו ריבוי אלגברי וגאומטרי של 1)
2. הוקטור העצמי המתאים לρ הוא חיובי
3. לכל ערך עצמי $\rho \neq \mu$, $|\mu| < \rho$ (כלומר הוא מקסימום חזק)
4. הוקטור העצמי החיובי היחיד (עד כדי כפל בסקלר) הוא הוקטור העצמי של ρ

בהמשך נבין למה המשפט הזה הוא חשוב. בעיקר נראה שימושים שלו בהקשרים של שרשראות מרקוב.

לפני הדיון במשפט, נרצה להכליל אותו גם למקרים כלליים יותר, ונוכל לעשות זאת עם משפט פרובניוס:
משפט פרובניוס: תהי $A_{n\times n}$ מטריצה אי שלילית. נייחס לA גרף מכוון עם n קודקודים, כך שהצלע $(i,j) \in E$ אם ורק אם $a_{ij} > 0$. אז אם הגרף שהתקבל הוא קשיר חזק ואפריודי, כל הנאמר במשפט פרון תקף גם לA.

הגדרה: יהי G גרף מכוון. לכל קודקוד x כך שניתן להגיע במסלול מx לעצמו, נגדיר את הקבוצה:

$$\{i : \text{there is a path of length i from x to itself}\}$$

כלומר, הקבוצה הזו מכילה את האורכים של המסלולים מx לעצמו. נאמר שהקודקוד x הוא **אי מחזורי** אם הGCD של קבוצת המספרים הזו הוא 1.
נאמר שהגרף G הוא **אי מחזורי**, או **אפריודי**, אם כל קודקוד x שלו הוא אי מחזורי.

בפעם שעברה דיברנו על המושג ההפוך: גרף ציקלי. המושגים אכן מתלכדים: אם הגרף איננו אי מחזורי, אז יש קודקוד כלשהו כך שכל מסלול ממנו לעצמו הוא באורך שהוא כפולה של k כלשהו - מה שנותן את התיאור מהפעם הקודמת של החלוקה של הגרף לקבוצות.
אפשר להבין למה מחזוריות של הגרף יכולה להפריע: כפי שאמרנו שיעור שעבר, אם M היא מטריצת השכנויות, אז על מנת שM^k תהיה מטריצה חיובית, חייב להיות מסלול באורך k בדיוק בין כל שני קודקודים. אם זה מתקיים, אפשר לראות שיש גם מסלול באורך $k+1$ בין כל שני קודקודים (על מנת לקבל מסלול באורך $k+1$ מx לy, נבחר שכן של y z כלשהו. יש מסלול באורך k מx לz, ואז נוכל להרחיב אותו למסלול באורך $k+1$ מx לy), וכך גם לכל $k' > k$. אבל אם הגרף הוא מחזורי, נניח לקודקוד x יש מחזור d, אז גם אם יש מסלול באורך k בין x לעצמו, לא יהיה מסלול באורך $k+1$ בין x לעצמו, כיוון שלפי ההנחה המסלול באורך הבא בגודלו בין x לעצמו הוא באורך $k+d$. לכן ניתן לראות שזה לא יתקיים.

כעת נעבור להוכיח חלקים ממשפט פרון.
טענה: אם G גרף מכוון קשיר חזק ואי מחזורי וM מטריצת השכנויות של G, אז יש k טבעי כך ש$M^k > 0$.
כלומר, יש חזקה שבה מטריצת השכנויות היא גרף חיובי.
הסברנו למה אם G איננו קשיר חזק, או G הוא מחזורי, לא יהיה קיים k שכזה, אך לא נראה למה אם G אכן קשיר חזק ואי מחזורי קיים k כזה.

לפני שנתחיל בהוכחה, נציין שכאשר אנחנו פוגשים במשפט מתמטי שאומר "קיים", מיד צריכה לעלות לנו השאלה - איך למצוא אותו?

2.8.3 הוכחת משפט פרון

הוכחה של קיום וקטור עצמי חיובי:
תחילה נשים לב שאם $w \neq 0$ וקטור כך ש$w \geq 0$ וA מטריצה חיובית, אז $Aw > 0$. זה נכון כיוון ש$w \neq 0$, יש j כלשהו כך ש$w_j > 0$, וכל שאר הקואורדינטות אי שליליות, $a_{ij} > 0$ לפי הנתון, ולכן:

$$(Aw)_i = \sum_{k=1}^{n} a_{ik}w_k \geq a_{ij}w_j > 0$$

המטרה הראשונה תהיה למצוא וקטור עצמי חיובי ומקסימלי. נסמן בS את קבוצת הוקטורים האי שליליים שנורמת l_2 שלהם היא 1.
נגדיר פונקציה כך שלכל $x \in S$,

$$L(x) = \min\{\frac{(Ax)_i}{x_i} : x_i \neq 0\}$$

הרעיון בבנייה הזו הוא כזה: אם $u \in S$ הוא וקטור עצמי של A שמתאים לערך העצמי המקסימלי λ_{max}, אז $\frac{(Ax)_i}{x_i} = \lambda_{max}$ לכל i. בL אנחנו מסתכלים על הקואורדינטה שהכי "לא הצלחנו" איתה, כלומר על הקואורדינטה שנמתחה הכי פחות על ידי המטריצה A.

הקבוצה S היא קומפקטית, וL היא רציפה לכן L מקבלת עליה מקסימום. נניח ש$v \in S$ הוא הוקטור כך ש$L(v)$ מקסימלי, ונסמן $\rho = L(v)$.
נטען שv הוא וקטור עצמי המתאים לערך עצמי ρ, כלומר $Av = \rho v \iff Av - \rho v = 0$. לפי ההגדרה של L, לכל $1 \leq i \leq n$, $(Av)_i \geq \rho v_i$, ויש i כלשהו בו יש שיוויון. לכן אם נסתכל על הוקטור $Av - \rho v$, נדע שהוא וקטור אי שלילי, ושיש לו קואורדינטה אחת שהיא אפס.
נניח בשלילה $Av - \rho v \neq 0$. כיוון ש$Av - \rho v \geq 0$, לפי מה שהראנו בהתחלה,

$$A(Av - \rho v) > 0$$

Av הוא גם כן וקטור חיובי, לכן נוכל לכווץ אותו כך שיהיה קטן ממש מ$A(Av - \rho v)$. כלומר, קיים $\epsilon > 0$ כך ש:

$$A(Av - \rho v) > \epsilon Av \Rightarrow A(Av) > \epsilon Av + \rho Av = (\rho + \epsilon)Av$$

נסמן: $Av=z$. אז $z\geq 0$, ומתקיים: $Az>(\rho+\epsilon)z$. כלומר, לכל $1\leq i\leq n$, $(Az)_i>(\rho+\epsilon)z_i$, ולכן נסיק $L(z)>\rho+\epsilon$. אבל זאת סתירה לכך שהמקסימום של L על S הוא ρ ולכן לא יתכן (*הערה: התעלמנו כאן מכך ש$\|Av\|$ היא לא בהכרח 1, אבל זו לא בעיה, שכן אפשר לנרמל את Av להיות עם נורמה 1, והוא עדיין יקיים את התכונה ש$L(Av)>\rho+\epsilon$ - ברור שאם $Az>(\rho+\epsilon)z$, אז גם לכל קבוע c, $cAz>c(\rho+\epsilon)z$ ולכן לא אכפת לנו איך מנרמלים אותה).

כלומר, קיבלנו סתירה ולכן הנחת השלילה לא מתקיימת, וv הוא בהכרח וקטור עצמי.

טענה: ρ הוא הרדיוס הספקטרלי של A. כלומר, אם μ הוא ערך עצמי אחר של A, אז $|\mu|\leq\rho$.

בהמשך נראה שאי השיוויון הוא חריף.

הוכחה: $Ay=\mu y \Leftarrow$ לכל i, $(Ay)_i=\mu y_i$, כלומר $\sum_j a_{ij}y_j=\mu y_i$.

נפעיל את אי שיוויון המשולש ונקבל:

$$|\mu|\cdot|y_i|\leq\sum|a_{ij}|\cdot|y_j|=\sum_j a_{ij}|y_j|$$

כיוון שA מטריצה חיובית.

לכן בפרט,

$$L(|y|)\geq|\mu|$$

לפי ההגדרה של L. כעת, בגלל שρ הוא המקסימום של L, $|\mu|\leq\rho$ כנדרש.

הערה: אמנם הגדרנו את L רק על כדור היחידה S, כדי להשתמש בקומפקטיות, אבל לכל וקטור כללי y אפשר לנרמל ועדיין נקבל את אותה התוצאה, לכן לא הקפדנו על להפעיל את L רק על וקטורי יחידה.

טענה: לρ יש ריבוי גאומטרי של 1

הוכחה: נשתמש בעקרון חשוב במציאת פתרון אופטימלי: נדע שפתרון הוא אופטימלי כאשר נראה שכל שינוי שנעשה לו לא יוכל לשפר אותו. לכן על דרך השלילה, אם ידוע שפתרון הוא אופטימלי, כל שינוי שנעשה לו לא יתן לנו משהו טוב יותר.

כך נפעל כאן: ρ התקבל כפתרון לבעיית אופטימיזציה, ולכן נראה שאם היינו מניחים בשלילה שהריבוי הגאומטרי שלו איננו 1, היה אפשר למצוא ערך עצמי טוב יותר.

נניח בשלילה שקיימים u,v וקטורים עצמיים המתאימים לערך עצמי ρ, והם בת"ל. בלי הגבלת הכלליות, נניח שu חיובי (ראינו שקיים וקטור עצמי חיובי לρ לכן אין בעיה לבחור אותו בתור u), ושv ממשי (המטריצה ממשית, לכן וקטור מרוכב הוא וקטור עצמי אם ורק אם החלק הממשי והחלק המדומה וקטורים עצמיים או אפס. כיוון שv וקטור עצמי הוא לא אפס, לכן החלק הממשי או המדומה שלו חייב להיות שונה מאפס, ואותו נבחר).

כיוון שv הוא ממשי, נוכל "להזיז" קצת ולקבל קואורדינטה שהיא אפס. כלומר, קיים $\epsilon\neq 0$ כך ש$u+\epsilon v$ הוא וקטור שיש לו קואורדינטה שהיא אפס. כעת:

$$0<A(u+\epsilon v)=Au+\epsilon Av=\rho(u+\epsilon v)$$

כאשר $A(u+\epsilon v)>0$ כיוון שהכפלנו מטריצה חיובית בוקטור אי שלילי. אבל $\rho(u+\epsilon v)$ הוא וקטור שיש לו אפסים, והוא לא וקטור חיובי, לכן זו סתירה ■

טענה: $\rho>0$

הוכחה: לפי מה שראינו קודם:

$$\rho=\max_{\|x\|_2=1,x\geq 0}\min_{i:x_i\neq 0}\frac{(Ax)_i}{x_i}$$

כל איבר בסכום הזה הוא חיובי: $x_i > 0$, וA מכילה רק איברים חיוביים. לכן נקבל שρ הוא מקסימום של איברים חיוביים ובפרט חיובי ממש גם כן.

תוספת למשפט פרון פרובניוס: תהי A כמו במשפט פרון פרובניוס. אזי:

$$\lim_{t \to \infty} (\frac{A}{\rho})^t = \frac{x_r \otimes x_l}{< x_r, x_l >}$$

כאשר x_r, x_l הם הוקטורים העצמיים החיוביים השמאלי והימני בהתאמה שמתאימים לערך העצמי ρ.

3 שרשראות מרקוב

יש לנו מערכת עם מספר סופי של מצבים, כלומר היא יכולה להיות באחד מn מצבים. אנחנו מסתכלים על המערכת ביחידות זמן בדידות. לכל זוג מצבים i, j, יש לנו את ההסתברות לכך שאם המערכת היתה במצב i בזמן t, היא תהיה במצב j בזמן $t+1$, נסמן הסתברות זו בp_{ij}.

מבנה כזה נקרא **שרשרת מרקוב סופית**. נרצה לנתח מערכות כאלה, יש הרבה שאלות מעניינות וחשובות שאפשר לשאול. בפרט, נראה שהמצב של המערכת לטווח ארוך בכלל לא תלוי במצב ההתחלתי, כלומר בסוף נתיצב במקום מסויים תמיד.

הערה: שרשרת מרקוב היא דוגמא ל**תהליך סטוכסטי**, כיוון שההסתברות למעבר בין מצבים במערכת תלויה רק במצב הנוכחי והמצב החדש. כלומר, ההסתברות למעבר ממצב i למצב j היא אותה ההסתברות, ללא תלות במצבים שהיו קודם או בנקודת הזמן בה אנחנו מסתכלים.

כל המידע על השרשרת מקודד על ידי **מטריצת המעבר**. זו מטריצה P כך ש$P_{ij} = p_{ij}$, כלומר במקום הij נמצאת ההסתברות למעבר ממצב i למצב j.

הגדרה: מטריצה P תיקרא **מטריצה סטוכסטית** אם מטריצה אי שלילית כך שהסכום של האיברים בכל שורה הוא 1, כלומר לכל i, $\sum_j p_{ij} = 1$

למשל, מטריצת המעבר של שרשרת היא מטריצה סטוכסטית. נשים לב שמובטח לנו שהסכום של איברים בכל שורה יהיה 1, אך לא מובטח לנו כלום לגבי הסכום של עמודה.

הערה: חשוב לשים לב שבמטריצה סטוכסטית **לא** מובטח לנו כלום לגבי הסכום של העמודות (אלא אם כן ידוע שהיא סימטרית למשל).

הערה: זו מטריצה סטוכסטית ביחס להכפלה משמאל. באופן דומה אפשר לקבל גם מטריצה סטוכסטית להכפלה מימין, אבל אז נדרוש שהסכום של כל עמודה יהיה 1 ולא הסכום של כל שורה.

בניתוח שלנו, נתחיל מ**וקטור התפלגות**, שהוא וקטור $x = (x_1, .., x_n)$, כך שלכל i, $x_i \geq 0$ ו$\sum_{i=1}^n x_i = 1$. הוא מתאר את ההסתברויות למצב של המערכת: x_i מכיל את ההסתברות לכך שהמערכת תימצא במצב i (ברגע נתון כלשהו בזמן בו אנחנו מסתכלים). בהינתן וקטור התפלגות x שמתאר את המערכת בזמן t, הוקטור התפלגות y שיתאר את המערכת בזמן $t+1$ יתקבל על ידי:

$$y = xP$$

ונשים לב שזה כפל משמאל ולא מימין.

נרצה להשתמש במסקנות ממשפט פרון־פרובניוס. אם נשכח מהדרישה של ארגודיות, או לחילופין נניח שאפשר להפעיל את המשפט על מטריצת המעבר של המערכת, נדע לומר שצריך להיות וקטור עצמי חיובי עם ערך עצמי שהוא גדול יותר מכל השאר.

ניתן לראות שהוקטור $\mathbf{1}$, כלומר הוקטור שכל הקואורדינטות שלו הם 1, הוא וקטור עצמי ימני השייך לערך עצמי 1 - הוא סוכם כל שורה, והסכום של כל שורה הוא 1. כעת, לפי פרון פרובניוס כיוון שמצאנו ערך עצמי עם וקטור עצמי חיובי, הוא בהכרח הערך העצמי המקסימלי.
כיוון ש1 הוא ערך עצמי, צריך להיות גם וקטור עצמי שמאלי. עבור x שהוא וקטור עצמי שמאלי, $1x = xP$, כלומר $x = xP$. x הוא חיובי, ואפשר לנרמל אותו כך שהוא יהיה וקטור התפלגות. נשים לב שx כזה מתאר את **ההתפלגות הגבולית**, או **ההתפלגות הסטציונרית**, שזהו המצב שבו המערכת כבר נשארת, משם היא לא זזה יותר.

את הוקטור העצמי השמאלי שמתאים ל1 מסמנים לרוב בπ.

עקרונית אפשר להסתכל על המערכת גם במעברים של יותר מזמן אחד: נוכל להסתכל על נקודת זמן מסויימת, ולשאול אם היינו בה במצב i, מה ההסתברות שבעוד 2 יחידות זמן נהיה במצב j. באותו אופן, אפשר להסתכל גם על k יחידות זמן. בהינתן וקטור התפלגות x, וקטור ההתפלגות של המערכת אחרי k יחידות זמן נתון על ידי xP^k. קל לראות שלכל k, P^k היא גם כן מטריצה סטוכסטית: היא בבירור אי שלילית, ו1 הוא עדיין ערך עצמי עם וקטור עצמי $\mathbf{1}$, לכן אפשר לקבל שהסכום של כל שורה עדיין יהיה 1.

3.1 מהלכים מקריים על גרפים

יהי G גרף קשיר ולא דו צדדי. **מהלך מקרי** על G היא שרשרת מרקוב שקבוצת המצבים שלה הוא V, וההסתברות להגיע מקודקוד אחד לקודקוד אחר היא אחידה. כלומר, בכל נקודת זמן ה"מצב" הוא הקודקוד שבו אנחנו נמצאים במהלך המקרי.
אם G לא קשיר, אז המערכת משמעותית פחות מעניינת: היא מתחלקת לרכיבי קשירות ורק ביניהם אפשר לזוז, לכן נרצה להצטמצם לרכיבי קשירות. אם G הוא דו צדדי, הוא יהיה גרף לא ארגודי ולכן לא נוכל לנתח אותו עם פרון פרובניוס.

השאלה המעניינת ביותר ביחס לשרשרת מרקוב היא מה התפלגות הגבול, כלומר על מה המערכת מתייצבת.

מטריצת המעבר של השרשרת תיראה כך: בשורה הv יהיו $d_v = deg(v)$ מקומות בהם מופיע $\frac{1}{d_v}$, ואפס במקומות האחרים. אם הגרף איננו דו צדדי, יש לנו גרף סטוכסטי ולכן יש התפלגות גבול.

מה היינו מצפים שתהיה התפלגות הגבול למהלך המקרי על גרף? ככל שלקודקוד יש יותר שכנים, יש יותר מקומות מהם אפשר להגיע אליו. לכן אפשר לצפות שבסוף בהסתברות גבוהה נהיה אצל הקודקודים עם דרגות גבוהות.

התפלגות הגבול היא פרופורציונלית לוקטור הדרגות: נגדיר את הוקטור π להיות כך ש$\pi_v = \frac{d_v}{2|E|}$: כלומר, ההסתברות להגיע לקודקוד מסויים היא הפרופורציה של הדרגה שלו מתוך סכום הדרגות הכללי.

טענה: $\pi P = \pi$, כלומר הוקטור π שתיארנו הוא אכן ההתפלגות הסטציונארית.

מבחינתנו, לומר שלשרשרת מרקוב יש "גבול" זה לא מספיק טוב. מאוד יעניין אותנו לא רק מה הגבול, כלומר לא רק איפה המערכת תהיה ב"אינסוף", אלא גם מה קצב ההתכנסות אליו. נרצה לשאול - אם הזמן הוא מאוד גדול, כמה קרובים נהיה לגבול? זו שאלה מהסוג שלא התעסקנו בה כמעט באינפי אבל במדעי המחשב מאוד נרצה לענות עליה.

3.2 אקספנדרים (גרפים מרחיבים)

בצורה לא מדוייקת, אלה גרפים שהמהלך המקרי עליהם מתכנס מהר ככל האפשר להתפלגות הגבול. מתברר שאלה גרפים "ללא צווארי בקבוק": כלומר אין מצב בו אפשר לחלק את הקודקודים של הגרף לשני חלקים גדולים עם מעט צלעות באמצע.

אינטואיטיבית, אפשר לחשוב על זה כמו על כביש מהיר, בו אם יש מעבר מאוד צר יווצר פקק והמכוניות לא יוכלו לעבור בקצב המירבי.

הגדרה: אומרים שלG יש **הרחבה צלעית** לפחות δ אם לכל $S \subseteq V$ המקיימת $|S| \leq \frac{|V|}{2}$ מתקיים $e(S,\overline{S}) \geq \delta|S|$, כאשר $e(S,\overline{S})$ הוא מספר הצלעות שמחברות בין S למשלים שלו $\overline{S}$.

כלומר, זו דרך מדוייקת לתאר שאין צוואר בקבוק.

הגדרה: הδ המקסימלי עבורו לG יש הרחבה צלעית יקרה **קבוע ההרחבה הצלעית.**

אנחנו מבינים שהתכונה של "קיום הרחבה צלעית" היא תכונה "טובה", כלומר טוב לנו למצוא גרפים שאין בהם צווארי בקבוק. טבעי לשאול **מהו** הδ הטוב ביותר שאפשר למצוא, אך מסתבר שזו בעיה $co-NP$ קשה ובפרט לא יהיה הגיוני מבחינתינו לנסות לפתור אותה. השאלה הבאה תהיה האם אפשר להתקרב לδ הטוב ביותר, והתשובה היא שכן:

משפט: יהי G גרף קשיר (לא מכוון) d־רגולרי ולא דו צדדי, ותהי A_G מטריצת השכנויות שלו (היא כמובן סימטרית). יהיו $d=\lambda_1 > \lambda_2 \geq .. \geq \lambda_n$ הערכים העצמיים של A_G (ראינו בתרגיל ש$\lambda_1 = d$ וכן שהוא גדול ממש מכל האחרים). אז אם δ הוא קבוע ההרחבה הצלעית של G,

$$\frac{d-\lambda_2}{2} \leq \delta \leq \sqrt{d^2-\lambda_2^2}$$

כלומר, אנחנו יודעים לתת חסם לקבוע ההרחבה הצלעית.

שבוע 9

משפט: תהי P מטריצת המעבר של שרשרת מרקוב ארגודית סופית. אז:

$$P^t \longrightarrow \begin{matrix} -\pi- \\ .. \\ -\pi- \end{matrix}$$

כאשר $t \to \infty$ וπ הוא וקטור ההתפלגות הסטציונרית של השרשרת.

כלומר, כפי שאמרנו קודם, המצב של המערכת לטווח הארוך לא תלוי במצב ההתחלתי שלה.

הוכחה: השרשרת ארגודית, לכן קיים k כך ש$P^k_{>0}$ - נזכיר שלא ראינו את המשפט הזה אבל אמרנו שהוא נכון ואנחנו הולכים להשתמש בו. נסמן את המטריצה שהשורות שלה הן π בZ. המטרה שלנו תהיה להראות ש$P^t - Z$ הוא קטן, כלומר שככל שt שואף לאינסוף, המרחק של P^t מZ הולך וקטן.

טענת עזר: לכל מטריצה סטוכסטית עם התפלגרות סטציונארית של π N, מתקיים $NZ=ZN=Z$

הוכחת טענת העזר: עבור ZN, אנחנו כופלים את N בπ משמאל, ולכן כיוון שπ הוא כפי שהראנו קודם וקטור עצמי של מטריצה סטוכסטית, נקבל שלכל שורה, $\pi N = \pi$, ולכן $ZN=Z$.

עבור הכפל של העמודה הi של Z, נשים לב שz_{ji} מכיל את ההסתברות של המצב i, כלומר זהו וקטור שהוא קבוע, מהצורה:

$$\epsilon \cdot \begin{matrix} 1 \\ .. \\ 1 \end{matrix}$$

כיוון ש$(1,..,1)$ הוא וקטור עצמי שמאלי של מטריצה סטוכסטית המתאים לערך עצמי 1, נקבל שוב את Z.

כיוון שP^k היא חיובית וZ אי שלילית, קיים קבוע $0 < c < 1$ כך שמתקיים:

$$P^k \geq cZ$$

אם $c = 1$, סיימנו: במקרה זה, $P^k = Z$, ולכן גם $P^{k+1} = PP^k = PZ = Z$, ובאינדוקציה אפשר לקבל שזה נכון לכל $t \geq k$.
לכן נניח $c \neq 0$. במקרה זה, נגדיר:

$$N = \frac{1}{1-c}(P^k - cZ)$$

וזו גם כן מטריצה סטוכסטית. לפי הגדרתה היא אי שלילית. נראה ש$\mathbf{1}$ הוא וקטור עצמי המתאים לערך עצמי 1:

$$N \cdot \mathbf{1} = \frac{1}{1-c}(P^k - cZ)\mathbf{1} = \frac{1}{1-c}(\mathbf{1} - c\mathbf{1}) = \frac{1-c}{1-c}\mathbf{1} = \mathbf{1}$$

ולכן טענת העזר נכונה גם עבורה. לא הראנו שלN יש התפלגות סטציונרית של π, אבל בכל זאת טענת העזר תהיה נכונה גם עבורה.

טענת עזר: עבור המטריצה N שהגדרנו, מתקיים:

$$(N - Z)^l = N^l - Z$$

הוכחת טענת העזר: נראה באינדוקציה על l. המקרה $l = 1$ הוא טריוויאלי. נניח נכונות עבור l. אז:

$$(N-Z)^{l+1} = (N-Z)^l(N-Z) = (N^l - Z)(N-Z) = N^{l+1} - N^l Z - ZN + Z^2$$

לפי הפעלה של טענת העזר הראשונה, נוכל לקבל $N^l Z = ZN = Z^2 = Z$, ולכן:

$$= N^{l+1} - Z$$

כנדרש.

מסטוכסטיות, נשים לב שכל איבר של המטריצה $N^l - Z$ הוא בתחום $[-1, 1]$.

טענת עזר: $P^{kl} - Z = (P^k - Z)^l = (1-c)^l (N - Z)^l$
לא הוכחנו את טענת העזר, אבל היא גם בסגנון של טענת העזר הקודמת: אינדוקציה על החזקה.

טענת עזר: $\|P^{kl} - Z\|_\infty \leq (1-c)^l$

כעת, ההעתקה $l \mapsto \|P^l - Z\|_\infty$ היא העתקה שהיא לא עולה (מונוטונית יורדת, עם אי שיוויון חלש בין האיברים), כיוון ש$P^{l+1} - Z = P(P^l - Z)$ וכל איבר ב$P^{l+1} - Z$ הוא צירוף ליניארי במקדמים אי שליליים שמסתכמים ב1 של המטריצה $P^l - Z$, לכן היא לא יכולה לגדול, לכן נוכל להראות:

$$\|P^l - Z\|_\infty \leq (1-c)^{\lfloor \frac{l}{k} \rfloor}$$

כעת, זו סדרה מונוטונית יורדת וחסומה, ולכן היא מתכנסת. כמו כן, מצאנו תת סדרה שאנחנו יודעים מה הגבול שלה, ולכן נוכל להסיק שכל הסדרה מתכנסת לגבול של תת הסדרה. ■

3.2.1 חזרה לאקספנדרים

נחזור לנושא העיקרי: דיברנו על אקספנדרים, גרפים מרחיבים. אמרנו שאלה הם גרפים "ללא צווארי בקבוק", וראינו כיצד אפשר לתאר זאת.
נסמן את הδ המקסימלי עבורו לG יש הרחבה צלעית ב$h(G)$, והוא נקרא **קבוע ההרחבה הצלעית** של G, כלומר:

$$h(G) = \min_{S \subseteq V, |S| \leq \frac{|V|}{2}} \frac{e(S, \overline{S})}{|S|}$$

הרעיון הוא שאם קבוע ההרחבה הצלעית של הגרף הוא גדול, אז הגרף יותר רחוק משיהיה לו צוואר בקבוק. כלומר, התכונה שאנחנו נרצה היא שקבוע ההרחבה הצלעית יהיה גדול ככל האפשר.

הערה: המושג של קבוע הרחבה צלעית מזכיר את המושג של חתכים בגרף: בשניהם מחלקים את קודקודי הגרף לשני חלקים ומסתכלים על הצלעות שמחברות בין שני החלקים. ההבדל ביניהם הוא שבהגדרה של קבוע ההרחבה הצלעית הסתכלנו על גודל יחסי למספר הקודקודים, כלומר הסתכלנו על מספר הצלעות שמחברות **ביחס** למספר הקודקודים סך הכל, בעוד שבהסתכלות על חתכים לא אכפת לנו כמה קודקודים יש בכל חתך, מסתכלים רק על מספר הצלעות שנמצאות בחתך.

הערה: כפי שאמרנו בשיעור שעבר, אחת מהשאלות המעניינות עבורנו הן השאלות של מהלך מקרי על גרף. כאמור, יעניינו אותנו הרבה פעמים שאלות שהן לא רק איפה נהיה במהלך המקרי בזמן "אינסוף", אלא אם ניקח כמות זמן גדולה, t כלשהי, כמה קרובים נהיה למצב הגבול.
אפשר לראות שכדי להגיע להתפלגות הגבול, צריך זמן לפחות כמו הקוטר של הגרף, כאשר **הקוטר** זה המרחק המקסימלי בין שני קודקודים בגרף. אם G הוא גרף d רגולרי עם n קודקודים, נטען ש$diam(G) \geq log_{d-1}n$:
כפי שאמרנו, נצטרך זמן לפחות להגיע מכל קודקוד לכל קודקוד אחר. אם התחלנו מקודקוד מסויים, בצעד אחד אפשר להגיע לd קודקודים (d השכנים שלו). מכל שכן שלו, אפשר להגיע לכל היותר $d-1$ קודקודים חדשים. נמשיך כך, בכל צעד אנחנו עולים במכפלה של $O(d)$. מספר המכפלות המינימלי שנצטרך הוא $log_{d-1}n$, אחרי מספר כזה לכל הפחות נכסה את כל הגרף.
כלומר, כדי להיות קרובים להתפלגות הגבול נצטרך לפחות $log_{d-1}n$. תכונה חשובה של אקספנדרים היא שהם משיגים את ההתכנסות המהירה ביותר - מסתבר שאצלם זה בדיוק זמן לוגריתמי.

נזכיר את המשפט שניסחנו שיעור שעבר, ונוכיח את החסם התחתון:
משפט: יהי G גרף קשיר d-רגולרי ולא דו צדדי, A_Gמטריצת השכנויות שלו. כפי שראינו בתרגיל, הערכים העצמיים של A_Gהם $\lambda_1 = d \geq \lambda_2 \geq ... \geq \lambda_n$ אז קבוע ההרחבה הצלעית של G $h(G)$ מקיים:

$$\frac{d - \lambda_2}{2} \leq h(G) \leq \sqrt{d^2 - \lambda_2^2}$$

הדבר המעניין שרואים מהמשפט הזה הוא שקבוע ההרחבה הצלעית והמרחק $d - \lambda_2$ קשורים אחד לשני.
הגודל $d - \lambda_2$ נקרא **הפער הספקטרלי** של A_G.
כפי שראינו בתרגיל, $d = \lambda_2$ אם ורק אם הגרף איננו קשיר. הנחנו במשפט שהגרף קשיר, אבל אפשר לראות שאם הוא לא קשיר נקבל $h(G) = 0$, כלומר קבוע רע מאוד.

הוכחה: נוכיח רק את החסם התחתון, כלומר נראה $\frac{d-\lambda_2}{2} \leq h(G)$. אין לנו יותר מדי מה לומר על $h(G)$, אבל אנחנו כן מכירים כמה משפטים שיכולים לתאר את λ_2. לכן נארגן מחדש את אי השיווין, כך שנוכל לחשוב עליו כעל חסם לλ_2:

$$\lambda_2 \geq d - 2h(G)$$

כזכור, $\mathbf{1}$ הוא וקטור עצמי לd, ולפי משפט קודם, $\lambda_2 = \max_{x \perp \mathbf{1}} \frac{xAx^T}{\|x\|^2}$.
כיוון שλ_2 הוא מקסימום של איברים, מספיק להראות שעבור x כלשהו, $\frac{xAx^T}{\|x\|^2} \geq d - 2h(G)$, ואז נוכל לקבל שאי השיוויון נכון גם עבור λ_2.
בהינתן חלוקה כלשהי של הקודקודים לS ו$\overline{S}$, נמצא וקטור x מתאים. המשמעות של $x \perp \mathbf{1}$ היא $<x, \mathbf{1}> = 0$ כלומר $\sum_{i=1}^{n} x_i = 0$. נגדיר את x באופן הבא:

$$x_i = \begin{cases} |\overline{S}| & i \in S \\ -|S| & i \in \overline{S} \end{cases}$$

כלומר, ב$|S|$ קואורדינטות יופיע לנו $|\overline{S}|$, וב$|\overline{S}|$ קואורדינטות של הוקטור יופיע לנו $-|S|$, לכן סך הכל סכום המקדמים הוא $|S|\cdot|\overline{S}|-|S|\cdot|\overline{S}|=0$, לכן הוא אכן מאונך ל$\mathbf{1}$.
נחשב:

$$\|x\|_2^2=\sum_{i=1}^{n}x_i^2=\sum_{i\in S}|\overline{S}|^2+\sum_{i\in\overline{S}}|S|^2=|S|\cdot|\overline{S}|^2+|\overline{S}|\cdot|S|^2=|S|\cdot|\overline{S}|(|\overline{S}|+|S|)=n|S|\cdot|\overline{S}|$$

כאשר המעבר האחרון נכון כיוון שS, $\overline{S}$ הם חלוקה של קודוקדי G.
כמו כן:

$$xAx^T=\sum_{i,j}A_{ij}x_ix_j=2\sum_{\{i,j\}\in E}x_ix_j$$

כיוון שA_{ij} יכול להיות 1 או 0, והוא 1 אם ורק אם $\{i,j\}\in E$, וגם אנחנו סופרים כל צלע פעמיים.
יש לנו 3 סוגים של צלעות: $i,j\in S$, $i,j\in\overline{S}$ ו$i\in S$, $j\in\overline{S}$ (בלי הגבלת הכלליות).
כעת, אם $i,j\in S$ אז $x_i=x_j=|\overline{S}|$ ולכן קיבלנו את המחובר $|\overline{S}|^2$. כמו כן, אם $i,j\in\overline{S}$ אז $x_i=x_j=-|S|$ ולכן קיבלנו את המחובר $|S|^2$.
את מספר הצלעות בS וב$\overline{S}$ נסמן ב$e(S)$ ו$e(\overline{S})$ בהתאמה.
אם $i\in S$ ו$j\in\overline{S}$ (או להיפך), נקבל $x_i=|\overline{S}|$ ו$x_j=-|S|$, ולכן נקבל את המחובר $-|S||\overline{S}|$. כלומר:

$$xAx^T=2e(S)|\overline{S}|^2+2e(\overline{S})|S|^2-2e(S,\overline{S})|S|\cdot|\overline{S}|$$

אם נסתכל על כל הxים המתקבלים באופן הזה מחלוקות $S,\overline{S}$, זה רק חלק מהוקטורים האפשריים ולכן:

$$\lambda_2=\max_{x\perp\mathbf{1}}\frac{xAx^T}{\|x\|^2}\geq\max_S\frac{2e(S)|\overline{S}|^2+2e(\overline{S})|S|^2-2e(S,\overline{S})|S|\cdot|\overline{S}|}{n|S|\cdot|\overline{S}|}$$

$$=2\max_S\frac{e(S)|\overline{S}|^2+e(\overline{S})|S|^2-e(S,\overline{S})|S|\cdot|\overline{S}|}{n|S|\cdot|\overline{S}|}$$

הגדלים $e(S)$ ו$e(\overline{S})$ לא הופיעו באי השיוויון שאנחנו רוצים להוכיח, לכן נרצה להיפתר מהם. נשים לב שמתקיים:

$$d\cdot|S|=2e(S)+e(S,\overline{S})\Rightarrow e(S)=\frac{d\cdot|S|-e(S,\overline{S})}{2}$$

כיוון שיש $|S|$ קודקודים בS ולכל אחד מהם יש דרגה d. יש 2 סוגים של צלעות: כאלה שמחברות בין שני קודקודים בS ואלה נספרות פעמיים, וכאלה שמחברות בין קודקוד בS לקודקוד ב$\overline{S}$, הן נספרות פעם אחת.
באותו אופן בדיוק:

$$d\cdot|\overline{S}|=2e(\overline{S})+e(S,\overline{S})\Rightarrow e(\overline{S})=\frac{d\cdot|\overline{S}|-e(S,\overline{S})}{2}$$

מכאן שמתקיים:

$$xAx^T=(d\cdot|S|-e(S,\overline{S}))|\overline{S}|^2+(d\cdot|\overline{S}|-e(S,\overline{S}))|S|^2-2e(S,\overline{S})|S|\cdot|\overline{S}|$$

$$=d(|\overline{S}|^2|S|+|\overline{S}|\cdot|S|^2)-e(S,\overline{S})(|\overline{S}|^2+|S|^2+2|S|\cdot|\overline{S}|)=d\cdot|S|\cdot|\overline{S}|\cdot(|S|+|\overline{S})-e(S,\overline{S})((|S|+|\overline{S}|)^2)$$

$$=d\cdot|S|\cdot|\overline{S}|\cdot n-e(S,\overline{S})n^2$$

ולכן:

$$\frac{xAx^T}{\|x\|^2} = \frac{d\cdot|S|\cdot|\overline{S}|\cdot n - e(S,\overline{S})n^2}{n|S|\cdot|\overline{S}|} = d - \frac{n\cdot e(S,\overline{S})}{|S|\cdot|\overline{S}|}$$

עדיין נשאר לנו $|\overline{S}|$ - בחרנו את S כך ש$|\overline{S}| \geq \frac{n}{2}$, ולכן $\frac{n}{|\overline{S}|} \leq \frac{1}{2}$, נציב ונקבל שלכל S,

$$\frac{xAx^T}{\|x\|^2} \geq d - \frac{2e(S,\overline{S})}{|S|}$$

כאשר x הוא הוקטור שמתקבל מS. לפי ההגדרה,

$$h(G) = \min_{S\subseteq V, |S|\leq\frac{|V|}{2}} \frac{e(S,\overline{S})}{|S|}$$

ולכן סך הכל קיבלנו:

$$\lambda_2 \geq d - 2h(G)$$

כנדרש ■

4 אופטימיזציה

באופן כללי, **בעיית אופטימיזציה** היא בעיה שבה יש לנו תחום הגדרה D ופונקציה ממשית $f: D\to\mathbb{R}$, ונרצה למצוא את $\max_{x\in D} f(x)$ או באופן מקביל את $\min_{x\in D} f(x)$.
כלומר, נרצה למצוא את הערך ה"טוב ביותר" של הפונקציה, כאשר הערך הטוב ביותר יכול להיות המקסימלי או המינימלי לפי המקרה.

בעיות אופטימיזציות הן שאלות כלליות מאוד, ונפגשים איתן במקומות רבים במחקר ובחיים.
נציג סוג מסויים של בעיות אופטימיזציה:

בעיית התכנון הליניארי, או **Linear programming (LP)** היא מקרה פרטי (חשוב מאוד) של בעיות אופטימיזציה. בבעיית תכנון ליניארי, התחום הוא קבוצה $D\subseteq\mathbb{R}^n$ שמוגדרת על ידי רשימה סופית של משוואות ואי שיוויון ליניאריים, הנקראים **האילוצים** של הבעיה. הפונקציה f גם היא ליניארית, והיא נקראת **פונקציית המטרה**.
(משוואה\אי שיוויון ליניארי משמעם שהם מהצורה $\sum_{i=1}^k \alpha_i x_i$, כאשר α_i מקדמים וx_i משתנים - אין מכפלות או חזקות בין המשתנים).
קבוצה D שכזו תיקרא **פוליהדר**, ואם היא חסומה היא תיקרא **פאון**, או **פוליטופ**.

למרות שבעיות LP הן יותר פשוטות מהמקרה הכללי, מסתבר שזהו עולם מאוד עשיר.

בתחום הזה לרוב יהיו לנו שתי נקודות מבט: מבחינה אלגברית ומבחינה גאומרטית.
מבחינה גאומטרית, אם מדובר ב$\mathbb{R}^2$ אז הקבוצה D היא קבוצה במישור שמוגבלת על ידי ישרים (יכול להיות שהיא לא חסומה, למשל אם היא מוגדרת על ידי האילוץ $x\geq 0$). באותו אופן, גם במרחב היא קבוצה שמוגבלת על ידי מישורים וכך הלאה.

חשוב לציין שבבעיית LP העולם שלנו הוא רציף, אך יש גם מקרה פרטי של בעיות עבורן הפתרונות הם דיסקרטיים, שהם מהווים סוג חשוב מאוד של בעיות.
נראה בהמשך שהאופטימום (נקודת המקסימום או מינימום) מתקבלים תמיד בקודקוד, וכיוון שיש מספר סופי של קודקודים זה נותן קישור למקרה הדיסקרטי.

4.0.1 דוגמא ־ ניהול משק

נניח שיש לנו משק עם פרות. ידוע שכל פרה צריכה לקבל ערכים תזונתיים מסויימים על מנת להיות בריאה: כמות מסויימת של חלבון, של פחמימות, של ויטמינים וכך הלאה. יש לנו סוגים שונים של מזון שאפשר לתת לפרות: חציר, תחמיץ, קש, וכך הלאה, ולכל מאכל ידוע מה הערך התזונתי שהוא מספק ליחידה. בנוסף, לכל סוג של מזון יש מחיר ליחידה.

המטרה שלנו תהיה לתכנן את התזונה של פרה: נרצה לדעת כמה יחידות מכל סוג של מזון לתת לה. נדרוש שהתזונה תענה על הצרכים שלה, ונרצה כמובן שהמחיר שנצטרך לשלם יהיה מינימלי.

אפשר לתאר את הבעיה הזו כבעיית תכנון ליניארי. נוכל לתת לכל סוג של מזון אינדקס $1 \leq j \leq n$. נתאר את המזון שנקנה בתור וקטור שורה $x \in \mathbb{R}^n$, כך שלכל j, x_j יכיל את מספר יחידות המזון שנקנה מהסוג ה־i.
כמו כן, נתאר את הערכים התזונתיים על ידי מטריצה $A \in \mathbb{R}_{n\times m}$, כאשר יש m אבות מזון שהפרה צריכה, במקום A_{ji} רשומה הכמות של הערך התזונתי ה־i שנמצא ביחידה של מזון מסוג j.
נתאר את הערכים התזונתיים של פרה צריכה לקבל על ידי וקטור $b \in \mathbb{R}^m$, כאשר לכל $1 \leq i \leq m$, b_i מכיל את הכמות שפרה צריכה לקבל מאב המזון ה־i.
את המחירים נתאר בוקטור $c \in \mathbb{R}^n$, כך ש־c_j מכיל את המחיר של המזון ה־j.

מהם האילוצים שלנו? אנחנו מחפשים וקטורים $x \in \mathbb{R}^n$ כך ש־x ייצג תזונה תקינה עבור פרה. ראשית, נדרוש $x \geq 0$ (לא יתכן לקנות כמות שלילית של מזון).
כמו כן, נדרוש שכל פרה תקבל לפחות את הערך התזונתי הדרוש לה מכל אב מזון (נאפשר מצב בו היא מקבלת יותר). המטריצה A והוקטור b למעשה מספקים לנו אוסף של אי שיוויונים ליניאריים:

$$xA \geq b$$

כלומר, לכל i נדרוש $(xA)_i \geq b_i$: $(xA)_i$ זה הכמות שהפרה מקבלת מהאב המזון ה־i, ונדרוש שהיא תקבל לפחות b_i.

כעת תיארנו את האילוצים. פונקציית המטרה שלנו היא המחיר המינימלי, והמחיר נתון על ידי $< c, x >$. לכן הפונקציה עבורה מנסים למצוא אופטימום היא $x \mapsto < c, x >$.

4.0.2 קישור לבעיית הספיקות

בעיה הדומה לבעיית ה־LP היא **בעיית הספיקות**: נתונה מערכת של משוואות ליניאריות ואי שיוויונים, ונרצה לדעת האם יש פתרון למערכת. כלומר, האם יש x כלשהו שמקיים את כל האילוצים הנתתונים על ידי המשוואות ואי השיוויונים.

הערה: בעיית הספיקות מזכירה אולי את בעיית האופטימיזציה SAT: בבעיה SAT, נתונה מערכת של משוואות ליניאריות ואי שיוויונים, ונרצה למצוא x שמספק מספר מירבי של אילוצים. חשוב להבין את ההבדל: SAT היא בעיית אופטימיזציה, בה נרצה למצוא פתרון שמספק כמה שיותר משוואות ־ התשובה בה היא פתרון שמספק מספר מירבי של משוואות. לעומת זאת, בעיית הספיקות היא שאלה (אין אופטימיזציה) ־ האם ניתן לספק מערכת נתונה ־ התשובה אליה יכולה להיות כן או לא.

נשים לב שיש קשר הדוק בין בעיית הספיקות: למשל, בהינתן פתרון לבעיית הספיקות נוכל לעשות רדוקציה ממנו לפתרון לבעיית האופטימזציה באופן הבא:
נניח שבעיית האופטימיזציה שלנו נתונה על ידי מטריצה A, ווקטורים c, b, כך שהאילוצים הם $x \geq 0$, $xA \geq b$, ופונקציית המטרה היא $< c, x >$.
בהינתן ערך T כלשהו, נוכל לשאול האם יש פתרון למערכת:

$$xA \geq b, x \geq 0, < c, x > \geq T$$

למעשה נוכל לקבל כך פתרון גם לבעיית האופטימיזציה: נריץ מעין חיפוש בינארי על ערכי T, עד שנמצא את הערך הקטן\גדול ביותר.
עדיין יש כאן נקודה עדינה שכיוון ויש לנו אינסוף איברים החיפוש אולי לא יהיה מוגדר היטב, אך לרוב זה לא יפריע לנו (למשל, נוכל להחליט שרק מסתכלים על מספרים ברזולוציה של ספרה אחת אחרי הנקודה).

שבוע 10

בשיעור הקודם הגדרנו בעיות אופטימזציה באופן כללי, והגדרנו את בעיית התכנון הליניארי. הצגנו בעיה מסויימת - בעיית האכלת הפרות. בסופו של דבר, בתיאור של הבעיה בתור בעיית תכנון ליניארי היו לנו וקטורים b, c ומטריצה A, ווקטור משתנים x, כאשר רצינו למצוא $\min < c, x >$, כאשר מתקיים $x \geq 0$ ו-$Ax \geq b$ (אין משמעות כמובן לשאלה האם מדובר בוקטור שורה או וקטור עמודה).
זו הצורה הכללית של בעיית תכנון ליניארי.

הערה: נשים לב שפונקציה ליניארית מהצורה $a_1x_1 + ... + a_nx_n$ היא בעצם מכפלה פנימית בין הוקטור $(a_1, .., a_n)$ והוקטור $(x_1, .., x_n)$.
כלומר, כל פונקציה ליניארית היא מכפלה פנימית בין וקטור המקדמים ווקטור המשתנים, וכך נרצה לרוב לכתוב אותה.

כפי שאמרנו, בעיית התכנון הליניארי היא בעייה חשובה: יש הרבה מאוד תחומים בחיים שבהם מופיעות בעיות אופטימיזציה הנתונות על ידי אילוצים ליניאריים ועם פונקציית מטרה ליניארית.
בדוגמא של האכלת הפרות, ברגע שניסחנו את הבעיה על ידי משוואות היה ברור שמדובר בשאלה של תכנון ליניארי - מיד קיבלנו אוסף של משוואות ואי שיוויונים ליניאריים. לא תמיד זה יהיה המצב, ישנן בעיות שבמבט ראשון לא נראות כמו בעיות תכנון ליניארי, אך אפשר להציג אותן גם כן בתור בעיות תכנון ליניארי. נציג כמה כאלה כעת.
כמו כן, יש גם בעיות שהן בעצמן לא בעיות תכנון ליניאריות, אך ניתן "להנמיך" אותן לבעיות תכנון ליניאריות. ידוע כיצד לפתור בעיות תכנון ליניארי ביעילות. לאחר שיש לנו פתרון, נוכל לפעמים להגיע לקירוב או להסיק ממנו משהו לגבי הפתרון המקורי.

אמרנו שתחום החסום על ידי משוואות ואי שיוויונים ליניאריים יכול להיות פאון (במקרה החסום) או פוליהדרון (במקרה הלא חסום). אם מדובר בפוליהדרון, ישנה אפשרות שהמקסימום לא חסום. זה כמובן המקרה הפחות מעניין. אך אם המקסימום חסום, מספיק להסתכל על תחום חסום. בתור דוגמא, אם מחפשים במישור מספרים כך ששתי הקואורדינטות שליליות והקואורדינטה המקסימלית בערך מוחלט היא כמה שיותר קטנה, אין צורך להסתכל בכל הרביע השלישי - מספיק להסתכל רק על נקודות שהן בטווח של 1 למשל מהראשית - כיוון שהאופטימום יתקבל בסביבה של הראשית.
לכן נוכל להניח בלי הגבלת הכלליות שמדובר על פאונים. רוב הדוגמאות שלנו יהיו עם פאונים.

טענה: פאון הוא גוף קמור

הוכחה: אם נסמן את הפאון בתור P, נצטרך להראות שלכל $x_1, x_2 \in P$, ולכל $0 \leq \alpha \leq 1$, $\alpha x_1 + (1-\alpha)x_2 \in P$.
P מוגדרת על ידי אוסף של משוואות ואי שיוויונים ליניאריים. לכן נראה שאם x_1 ו-x_2 קיימו את האילוצים, גם $\alpha x_1 + (1-\alpha)x_2$ תקיים את האילוצים.
אם מדובר באילוץ מהצורה $< b, x >= \beta$, אז נקבל:

$$< b, \alpha x_1 + (1-\alpha)x_2 >= \alpha < b, x_1 > +(1-\alpha) < b, x_2 = \alpha\beta + (1-\alpha)\beta = \beta$$

ואם מדובר באילוץ מהצורה $< b, x >> \delta$, כיוון ש-$0 \leq \alpha \leq 1$, נסיק $\alpha, 1-\alpha \geq 0$ ולכן נוכל להכפיל את אי שיוויונים בהם ולקבל:

$$\alpha < b, x_1 >> \alpha\delta, (1-\alpha) < b, x_2 >> (1-\alpha)\delta$$

ואז:

$$< b, \alpha x_1 + (1-\alpha)x_2 >= \alpha < b, x_1 > +(1-\alpha)x_2 > \alpha\delta + (1-\alpha)\delta = \delta$$

כלומר, $\alpha x_1 + (1-\alpha)x_2 \in P$ ■

בעיית תכנון קמור היא בעיית אופטימיזציה שהתחום שלה הוא קבוצה קמורה כלשהי Ω (לאו דווקא פאון), ופונקציית המטרה f היא פונקציה קמורה (לאו דווקא ליניארית).
נזכיר שפונקציה $f : \mathbb{R}^n \rightarrow \mathbb{R}$ היא קמורה אם לכל $x_1, x_2 \in \Omega$ ולכל $0 \leq \alpha \leq 1$ מתקיים:

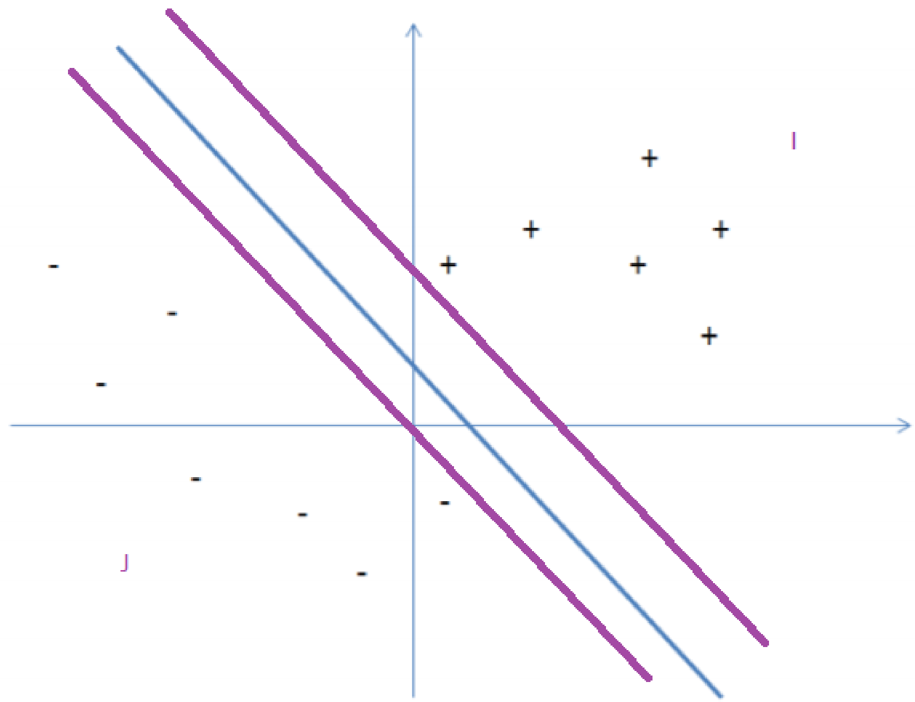
$$\alpha f(x_1) + (1-\alpha)f(x_2) \geq f(\alpha x_1 + (1-\alpha)x_2)$$

כלומר הישר שמחבר בין שתי נקודות על הגרף של הפונקציה נמצא מעל הגרף.

4.0.3 בעיית הפרדת הנקודות

נתונים אוספים I וJ של נקודות ב$\mathbb{R}^n$. אוסף הנקודות שבI יקראו נקודות "חיוביות" והנקודות בJ יקראו נקודות "שליליות" (אין קשר לסימנים של הקואורדינטות).

נרצה למצוא מישור שמפריד בין הנקודות, כך שהשוליים שלו יהיו רחבים ככל האפשר. כלומר, נניח שאפשר להגדיר ישר כך שכל הנקודות שבI הן מעל הישר וכל הנקודות שבJ הן מתחת לישר. נרצה שה"אזור המפורז" שבין הנקודות שבI לנקודות שבJ יהיה גדול ככל האפשר:



את הישר נוכל לתאר על ידי המשוואה $y = ax + b$, ונניח שרוחבו ϵ. נרצה שϵ יהיה גדול ככל האפשר, לכן נחפש את $\max \epsilon$. נדרוש שכל נקודה בI תימצא במרחק של לפחות אפסילון מהישר ומעליו, כלומר שלכל $i \in I$, הנקודה (x_i, y_i) תקיים:

$$y_i - ax_i - b > \epsilon$$

ולכל $j \in J$, הנקודה (x_j, y_j) תימצא במרחק של לפחות אפסילון מהישר ומתחתיו, כלומר:

$$y_j - ax_j - b < -\epsilon$$

אנחנו מחפשים את הישר, שהוא נקבע על ידי הבחירה של a וb, ואת הרוחב המקסימלי ϵ. כלומר, במערכת שלנו המשתנים הם a, b, ϵ (נשים לב שבמערכת שלנו הxים והyים הם לא המשתנים, בשונה מבדרך כלל). זו מערכת אילוצים ליניארית ולכן הכל מסתדר.

מה היה קורה אילו הנקודות I וJ היו מאורגנות במישור כך שלא היה קיים מישור המפריד, אלא פרבולה? נכתוב את המשוואות המתארות את הבעיה עם השינוי. פרבולה מוגדרת על ידי המשוואה $y = ax^2 + bx + c$. נסמן בϵ את הרוחב של השטח המפורז סביב לפרבולה. עדיין מחפשים את $\max \epsilon$, אך הפעם נדרוש שלכל $i \in I$,

$$ax_i^2 + bx_i + c > y_i + \epsilon$$

ולכל $j \in J$,

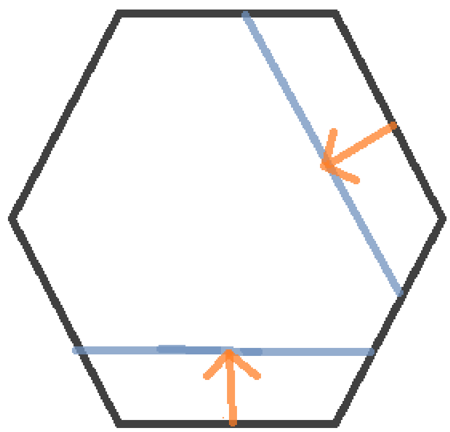
$$ax_j^2 + bx_j + c < y_j - \epsilon$$

המשוואות האלה הן לא ליניאריות בx_i, **אבל** כיוון שהמשתנים שלנו הם a, b, c, ϵ, הם מי שאנחנו מחפשים, והxים והyים הם קבועים, זו משוואה ליניארית בa, b, c, ϵ ־ הם מופיעים כמשתנים ליניאריים. גם זו מערכת תכנון ליניארי (למרות שאולי במבט ראשון היא לא נראית כזו).

4.0.4 רדיוס מקסימלי של מעגל חסום במצולע

נתון מצולע במישור, נרצה למצוא את הרדיוס המקסימלי האפשרי עבור מעגל שנמצא בתוך המצולע.
מעגל נתון על ידי משוואה לא ליניארית, לכן אולי במבט ראשון זה לא נראה כמו בעיית תכנון ליניארי, אך נראה שהיא כן.

הרעיון הוא כזה: בהינתן מעגל שרדיוסו מקסימלי, המרכז שלו תהיה הנקודה שמרחקה מכל דופן מקסימלי. כל דופן מגדירה אילוץ: כביכול עבור כל דופן מתחילים בדופן ומתחילים להתקדם פנימה לתוך המצולע. המקום שבו כל הישרים האלה יפגשו יהיה מרכז המעגל המדובר:



בהינתן נקודה $(u,v)\in\mathbb{R}^2$, המרחק שלה מהישר $y=ax+b$ נתון על ידי

$$\frac{|v+au-b|}{\sqrt{1+a^2}}$$

כאשר הביטוי שבתוך הערך המוחלט יכול להיות חיובי או שלילי, כתלות במיקום של הנקודה ביחס לישר.
מעגל מוגדר על ידי נקודת המרכז והרדיוס, כאשר נקודת המרכז צריכה לקיים שמרחקה מכל צלע הוא לכל היותר R.
נרצה למצוא את $\max R$, כאשר כל צלע $y_i=a_ix+b_i$ מגדירה אילוץ:

$$-R\leq\frac{v-a_iu-b_i}{\sqrt{1+a_i^2}}\leq R$$

נשים לב שעבור ישר יחיד, זה כולל גם נקודות במרחק לכל היותר R ממנו שהן בתוך המצולע, אך גם נקודות שהן מחוץ למצולע. הרעיון הוא שבכל זאת נקודה שמקיימת את כל האילוצים, כלומר היא במרחק לכל היותר R מכל אחת מהצלעות, תהיה חייבת להיות בתוך המצולע.
אפשר גם להתיחס בנפרד לצלעות עבורן הנקודות שבתוך המצולע מקבלות ערך חיובי, כלומר נצפה שזה יהיה קטן־שווה לR, והצלעות שעבורן הנקודות שבתוך המצולע מקבלות ערך שלילי, כלומר נצפה שזה יהיה גדול שווה ל$-R$.
בסופו של דבר, הפתרון למערכת היא נקודה (u,v) ורדיוס R מקסימלי $\Leftarrow$ זו מערכת ליניארית בu,v וR.

4.0.5 ריבועים פתוחים

נתון אוסף של נקודות במישור $\{(x_i,y_i)\}_{i\in I}$, ונרצה לשאול מהו הישר כך שכל הנקודות "כמה שיותר קרובות" ללשבת עליו. אם הנקודות לא יושבות על ישר, יתכן שאין ישר שעובר בדיוק בכל הנקודות. לכן נרצה לקבל ישר שהוא "כמה שיותר קרוב" ללעבור דרך כל הנקודות.
כמובן שהמושג "כמה שיותר קרוב" הוא לא מושג מתמטי, ונצטרך להגדיר במדוייק באיזה סוג של קירוב אנחנו מעוניינים, כלומר כיצד מודדים את המרחק של הנקודות מהישר.

קירוב על ידי L_2:
הגישה של אוקלידס לבעיה: לכל נקודה, נסתכל על המרחק של הנקודה מהישר בנורמת L_2, וננסה למזער את המרחק הכולל. כלומר, נרצה למצוא ישר $y=ax+b$ שמקבל את:

$$\min\sum_{i\in I}(ax_i+b-y_i)^2$$

כאשר x_i, y_i נתונים - זו משוואה בa וb. יש לנו פונקציה ריבועית ואנחנו מחפשים נקודת מינימום שלה - לכן טבעי לגזור ולהשוות לאפס: נגזור לפי a ונגזור לפי b, נקבל שתי משוואות ליניאריות בשתי משתנים, ונוכל בקלות למצוא פתרון.
הישר שימצא בשיטה הזו הוא הקירוב הטוב ביותר במובן של L_2.

יש גם עוד תפיסות של מרחק:
קירוב על ידי L_∞:
עבור L_∞, נרצה למצוא ישר $y = ax + b$ כך שהמרחק המקסימלי של נקודה כלשהי יהיה מינימלי. המרחק של נקודה x_i, y_i מהישר נתון על ידי $|ax_i + b - y_i|$. נסמן בתור ϵ את המרחק המקסימלי שמתקבל, לכן נרצה ישר שעבורו מתקבל $\min \epsilon$, כאשר לכל i,

$$-\epsilon \leq ax_i + b - y_i \leq \epsilon$$

כלומר, נקודה נמצאת בטווח של ϵ מהישר הנתון, ונרצה למצוא את הϵ הטוב ביותר. זו בבירור בעיית תכנון ליניארי.

קירוב על ידי L_1:
עבור L_1, נרצה למצוא ישר $y = ax + b$ שעבורו סכום המרחקים בערך מוחלט מינימלי. כלומר, נרצה למצוא ישר עבורו מתקבל $\min_{a,b} \sum_{i \in I} |ax_i + b - y_i|$.
הבעיה היא שערך מוחלט איננו פונקציה ליניארית, כלומר במקרה הזה פונקציית המטרה כפשוטה היא לא פונקציה ליניארית.
נוכל בכל זאת לארגן מחדש את הבעיה ולקבל פונקציה ליניארית: לכל i, נסמן את המרחק של הנקודה (x_i, y_i) מהישר בϵ_i. נרצה שסכום המרחקים ϵ_i יהיה מינימלי $\Longleftarrow$ נגדיר מערכת מעל המשתנים $a, b, \epsilon_1, ..., \epsilon_n$. נחפש את $\min_{a,b} \sum_{i \in I} \epsilon_i$, תחת האילוצים שלכל $i \in I$,

$$-\epsilon_i \leq ax_i + b - y_i \leq \epsilon$$

וזו אכן מערכת שמתאימה לבעיית תכנון ליניארי.

הערה: בבעיות מהסוג הזה, הרבה פעמים נקבל את אותה התבנית - קירוב בL_∞ זו בעיית תכנון ליניארי קלה, קירוב בL_2 זה חיפוש נקודת קיצון לפונקציה, וקירוב בL_1 נעשה על ידי הטריק שלמעלה.

4.0.6 בעיות תכנון ליניארי בשלמים

יש מקרים מסויימים בהם נרצה שהפתרון האופטימלי יהיה מעל השלמים. כלומר, יש מצבים שבהם פתרון שאיננו נתון על ידי מספרים שלמים לא מתאים לעולם הבעיה.
מסתבר שהבעיה של פתרון בעיות תכנון ליניאריות בשלמים, ILP **(integer linear programming)** היא בעיה NP-קשה. כלומר, אם נחפש אופטימום כללי לפונקציה שמקיים אילוצים ליניאריים זו בעיה שניתן לפתור ביעילות, אך אם נוסיף את הדרישה שנרצה אופטימום כך שכל האיברים שלו שלמים, כבר קיבלנו בעיה קשה.

למשל, בעיית הספיקות $3CNF$ שראינו בחישוביות ניתנת לתיאור כבעיית תכנון ליניארי מעל השלמים, כאשר $x_i \in \{0, 1\}$, ואז אילו היינו יודעים לפתור אותה היינו יודעים לפתור את $3CNF$, שהיא בעיה NP-קשה.

יש מקרים מיוחדים בהם הפתרון האופטימלי יצא מעל השלמים בכל מקרה.
למשל, **בעיית הזיווג הדו צדדי האופטימלי**: נתון גרף דו צדדי, ונרצה למצוא אוסף של צלעות מקסימלי שמהווה זיווג. לכל צלע (i, j), נוכל להגדיר משתנה x_{ij} שמקבל 1 אם (i, j) נמצא בזיווג, ו0 אחרת.
נרצה למקסם את מספר הצלעות בזיווג:

$$\max \sum_{i,j} x_{ij}$$

התכונה של זיווג היא שכל קודקוד יגע בצלע יחידה, לכן:

$$\sum_j x_{ij} = 1$$

כעת, אם $x_{ij} \in \{0,1\}$ נקבל בעיית תכנון ליניארי בשלמים. אך מסתבר שאם נפתור את בעיית התכנון הליניארי הדומה שדורשת רק $0 \le x_{ij} \le 1$, **עדיין** נקבל שקיים פתרון אופטימלי מעל השלמים.

זו תכונה מאוד מעניינת וחשובה, היא קורת לכל מני בעיות. לא נעמיק כאן בסוג הזה של שאלות.
גם בבעיית הזרימה שראינו באלגו קורה משהו דומה: ראינו שאם המשקלים שעל הצלעות הם שלמים, קיימת זרימה אופטימלית שגם היא בשלמים.

עבור בעיות ILP הרבה פעמים משתמשים ב־LP כדי לקבל קירוב: פותרים קודם את הבעיה כבעיית LP, ואז מוצאים דרך לעגל או לעבור לבעייה המקורית מבלי לפגוע יותר מדי באופטימליות של הפתרון.

4.1 מציאת הפתרון האופטימלי

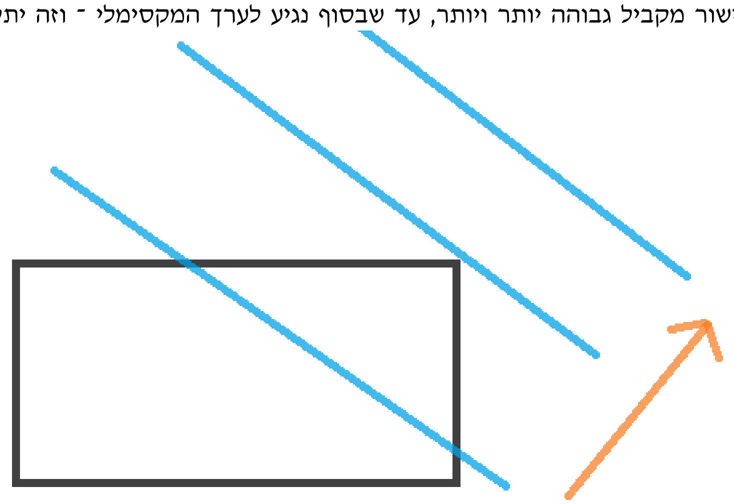
הגדרה: על־מישור, או $\boldsymbol{Hyperplane}$ ב־$\mathbb{R}^n$ מוגדר על ידי וקטור c ומספר ממשי γ באופן הבא:

$$H = \{x :< c,x >= \gamma\}$$

למשל, על־מישר ב־$\mathbb{R}^2$ הוא ישר, על־מישור ב־$\mathbb{R}^3$ הוא מישור וכך הלאה.

נשים לב שאם נקבע את c ונתחיל להזיז את γ, נקבל אוסף של על־מישורים מקבילים.

כאמור בעיית תכנון ליניארי נתונה על ידי האילוצים $Ax \le b$, $x \ge 0$ שמגדירים פאון במרחב, ונרצה למקסם את $< c,x >$. מבחינה גאומטרית, אפשר לחשוב על הבעיה באופן הבא: וקטור x הוא פתרון לבעיה אם הוא נמצא בתוך הפאון שנוצר על ידי האילוצים. עבור ערכים שונים של γ, נסתכל על העל־מישור שמוגדר על ידי $< c,x >= \gamma$. על־מישור שכזה יכיל פתרון שמשיג ערך γ אם ורק אם יש לו חיתוך לא ריק עם הפאון.
כלומר, אנחנו מחפשים γ מקסימלי כך שלעל־מישור המוגדר על ידי γ ו־c יהיה חיתוך לא ריק עם הפאון.
אם נחשוב על הבעיה מבחינה גאומטרית, האינטואיציה תגיד לנו שהערך המקסימלי מתקבל תמיד בקודקוד (יכול להיות שיש יותר מערך מקסימלי יחיד, אבל בהכרח יש ערך מקסימלי המתקבל על קודקוד). כדי להבין זאת, אפשר לחשוב על מישור שעובר בחדר למשל בזווית כלשהי. כל פעם נעבור למישור מקביל גבוהה יותר ויותר, עד שבסוף נגיע לערך המקסימלי - וזה יתקבל בקודקוד. בציור:



בשיעור הבא נוכיח במדוייק את התכונה הזו.

שבוע 11

דיברנו על בעיית התכנון הליניארי, והסברנו למה היא מעניינת, אך בינתיים לא דיברנו בכלל באופן פורמלי על איך לפתור אותה.

הגדרה: נאמר שבעיית LP היא **בייצוג סטנדרטי** אם היא מהצורה:

$$\max < c, x >, Ax = b, x \geq 0$$

כלומר, נדרוש שוקטור המשתנים יהיה אי שלילי, ונדרוש שהאילוצים יהיו שיוויונות.

הגדרה: נאמר שבעיית LP היא **בייצוג קנוני** אם היא מהצורה:

$$\max < c, x >, Ax \leq b$$

כלומר, ויתרנו על משוואות בכלל, והשתמשנו באי שיוויונים בלבד.

לכאורה צמצמנו את עולם הבעיות שלנו, אך מסתבר שלא:
טענה: בהינתן בעיית LP כלשהי, אפשר להציג אותה בייצוג סטנדרטי, ואפשר להציג אותה בייצוג קנוני.

הוכחה: כדי להביא תוכנית ליניארית לצורה סטנדרטית, עלינו להיפתר מכל האי שיוויונים הליניאריים למעט אלה מהצורה $x_i \geq 0$.
נוכל לעשות זאת כך: אם בתיאור המקורי של הבעיה מופיע אי השיוויון $< a, x > \geq \alpha$, נוסיף עוד משתנה סלקרי $s \geq 0$ (עוד קואורדינטה בx), ונחליף את אי השיוויון במשוואה

$$< a, x > -s = \alpha$$

כלומר, קראנו להפרש בשם ואמרנו שהוא אחד מהקואורדינטות של וקטור המשתנים.

כדי לעבור לייצוג קנוני, נרצה להחליף משוואות ליניאריות מהסוג

$$< a, x >= \alpha$$

בזוג אי שיוויונים:

$$< a, x > \geq \alpha$$

$$< a, x > \leq \alpha$$

$\blacksquare$

יהיו מקרים בהם יהיה לנו נוח יותר לעבוד עם צורה כזו או אחרת.

נשים לב שבמחינת סיבוכיות, הגדלנו את הבעיה לכל היותר פי שתיים, לכן אין הבדל משמעותי מבחינת סיבוכיות זמן הריצה.

משהו שחוזר בבעיות אלה הוא ההתיחסות המקבילה שלנו: מצד אחד חושבים על הבעיה מבחינה גאומטרית, ומצד שני חושבים עליה מבחינה אלגברית. האינטואיציה הגאומטרית אומרת שהמכפלה הפנימית מגדירה על מישור, וכשאנחנו מגדילים אותו אנחנו מסתכלים על העל מישורים המקבילים, ומנסים למצוא אזור כך שיש חיתוך עם הפאון והערך מקסימלי. כפי שאמרנו מקודם, במקרה הכללי נקודת האופטימום הזו מתקבלת בקודקוד (בין היתר - יכול להיות שהוא מתקבל גם בעוד מקומות).
הצד השני הוא האינטואיציה האלגברית: מטריצות, מספרים, פעולות אלגבריות. נצטרך דרך לעבור מהשפה הזו לשפה הזו.

4.1.1 מציאת הפתרון האופטימלי ־ חלק ראשון: מהצד האלגברי

כפי שאמרנו קודם, בתחילת הדיון נרצה בכלל להשתכנע שקיים אלגוריתם סופי (כלומר אלגוריתם שעוצר תמיד) לפתרון הבעיה. על פניו, לא ברור בכלל שאפשר לנסח אלגוריתם שיוכל לפתור את הבעיה, גם אם אין לנו דרישות של סיבוכיות זמן ריצה בכלל: יש אינסוף xים פוטנציאליים שיכולים לקיים את המשוואה. ברור שהפתרון הנאיבי ־ לעבור על כל הxים המקיימים את המשוואה ולבחור את הטוב ביותר ־ לא יעבוד כאן. לכן להראות שאפשר למצוא אלגוריתם שפותר את הבעיה בזמן סופי תמיד זה דבר לא קטן בכלל.

בדיון שלנו עכשיו אנחנו נעבוד עם LP בייצוג סטנדרטי, כלומר עם בעיות מהצורה $\max < c, x >$, $Ax = b, x \geq 0$

הגדרה: אומרים שx **פתרון מותר** אם $Ax = b$ ו$x \geq 0$.
כלומר, נאמר שוקטור x הוא פתרון מותר לבעיה אם הוא עומד באילוצים של הבעיה.

הגדרה: בהינתן וקטור x, נגדיר את **התומך** שלו להיות קבוצת האינדקסים בהם הוא לא מתאפס:

$$supp(x) = \{i : x_i \neq 0\}$$

כאשר נחשוב על x בתור וקטור שהוא פתרון לבעיית LP, נשים לב לכך שהתומך מגדיר קבוצת עמודות בA. באופן כללי, Ax מגדיר צירוף ליניארי של עמודות המטריצה A. כיוון שx מתאפס מחוץ ל$supp(x)$, אז למעשה יש לנו צירוף ליניארי של בדיוק אותן העמודות המופיעות בתוך $supp(x)$.

הגדרה: אומרים שx הוא **פתרון בסיסי** אם קבוצת העמודות בA המתאימות ל$supp(x)$ היא קבוצה בלתי תלויה ליניארית.

הגדרה: נאמר שx הוא **פתרון בסיסי מותר** אם הוא גם פתרון בסיסי וגם מותר, ונסמן אותו בbfs ($\boldsymbol{basic\ feasible\ solution}$). (נשים לב שאין קשר לאלגוריתם BFS).

משפט: תהי LP בעיה המוצגת בצורה הסטנדרטית, כך שיש לה לפחות פתרון מותר אחד ופונקציית המטרה חסומה מלעיל. אז יש לLP פתרון אופטימלי שהוא פתרון בסיסי מותר.

כלומר, אנחנו מניחים שנתונה בעיה מהצורה $\max < c, x >$ כך ש$x \geq 0$ וגם $Ax = b$. אנחנו מניחים שקיים x כלשהו שמקיים את אוסף האילוצים, ושהפונקציה $\max < c, x >$ חסומה בתחום־ אחרת השאלה לא מעניינת ונוכל לבחור ערכים גדולים כרצוננו.
המשפט נותן לנו מקום מסויים לחפש את הפתרון האופטימלי ־ הוא אומר לנו שעלינו לחפש אותו בין הפתרונות הבסיסיים.

הוכחה: נראה שלכל פתרון מותר, ניתן למצוא פתרון בסיסי מותר שהוא טוב לפחות כמוהו, ואז נסיק שכיוון שקיים אופטימום (למשל כיוון שמדובר בפונקציה חסומה על קבוצה קומפקטית), קיים אופטימום שהוא פתרון בסיסי מותר.
יהי x פתרון מותר. כלומר, $x \geq 0$ ו$Ax = b$. נביט בעמודות של A המתאימות ל$supp(x)$. אם זו קבוצה של וקטורים בת"ל ־ סיימנו, כי x הוא פתרון בסיסי מותר, ובבירור הוא לא גרוע מעצמו.

אחרת, יש תלות ליניארית בין עמודות A שנבחרו על ידי x. כלומר, יש צירוף ליניארי של העמודות הנ"ל שמתאפס, כלומר נוכל למצוא וקטור y כך ש$supp(y) \subseteq supp(x)$ (כלומר הוא משפיע על אותן עמודות כמו x), ו$Ay = 0$.
נפעל על פי בנייה רגילה: נמצא ϵ כלשהו כך ש$x + \epsilon y$ משפר את x.
מטרותינו הן:

1. שהוא יקיים $A(x + \epsilon y) = b$
2. $x + \epsilon y \geq 0$
3. $supp(x + \epsilon y) \subsetneq supp(x)$
4. $< c, x + \epsilon y > \geq < c, x >$

תכונות 1 ו2 אומרות ש$x+\epsilon y$ הוא פתרון חוקי לבעיה, 4 אומר שהוא משיג ערך טוב לפחות כמו מה שx השיג, ו3 אומר לנו שהצלחנו להקטין ממש את כמות הוקטורים שבפרוש.

תנאי 3 הוא שיבטיח לנו שהפעלה של התהליך מספר פעמים סופי יתן לנו פתרון בסיסי: כל פעם מורידים וקטור מאוסף הוקטורים, לכן בסופו של דבר נגיע לקבוצה בת"ל, ואז נקבל וקטור שהוא פתרון בסיסי, ושאר התכונות מבטיחות לנו שהוא יהיה פתרון מותר ולא גרוע מx.

1 מיידי עבור כל בחירה של ϵ, כיוון שמליניאריות נקבל $A(x+\epsilon y)=Ax+\epsilon Ay=Ax+0=Ax=b$.

את ϵ נבחר באופן הבא: נכפיל את y בערך הקטן ביותר שיגרום להתאפסות הראשונה בקואורדינטה כלשהי בתומך של x. הבנייה הזו תבטיח את קיומם של 3 (איפסנו קואורדינטה שנמצאת ב$supp(x)$), ואת 2 (השארנו כל קואורדינטה בx חיובית והפכנו אחת מהן לאפס).

נותר לנמק מדוע פונקציית המטרה לא התקלקלה. נשים לב שמתקיים:

$$< c, x+\epsilon y > \geq < c, x > \iff < c, x > + \epsilon < c, y > - < c, x > \geq 0 \iff \epsilon < c, y > \geq 0$$

אם בy יש גם קואורינטות חיוביות וגם שליליות, אז אפשר לבחור את הסימן של ϵ כך $sign\epsilon = sign < c, y >$. כלומר, אם יש קואורדינטות חיוביות וקואורדינטות שליליות, נוכל לבחור ϵ לפי מה שמסתדר לנו, ואז הדרישה תתקיים.

נעבור למצב שבו y וקטור אי שלילי. במקרה הזה ϵ חייב להיות שלילי. אם $0 >< c, y >$ - סיימנו, כיוון ש$0 \geq \epsilon < c, y >$.

נשאר להתיחס למקרה שבו $0 << c, y >$, $0 < \epsilon$. נראה שבמקרה זה פונקציית המטרה אינה חסומה - בניגוד להנחה. נשים לב שלכל $k > 0$, $x+ky$ הוא גם כן וקטור אי שלילי שמקיים $(x+ky) \geq 0$, $A(x+ky)=b$, ומקיים $< c, x+ky >>< c, x >$. כלומר, הוא מהווה שיפור ממש על x $\Leftarrow$ קיבלנו שעבור כל k חיובי, נוכל להמשיך לקבל ערכים גדולים כרצוננו בשביל הפונקציה - סתירה לכך שהפונקציה חסומה. $\blacksquare$

נשים לב שבשלב הזה הפכנו את בעיית LP לבעיה סופית: בהינתן מטריצה A, יש מספר סופי של תת קבוצות של עמודות של A, ולכן מספר סופי של תת קבוצות של עמודות שהן בת"ל. בהינתן קבוצה בת"ל של עמודות, אם קיים צירוף ליניארי שלהן שיתן את הוקטור b, צירוף שכזה הוא יחיד (זה בעצם הייצוג של b לפי בסיס).

לכן קיבלנו את האלגוריתם הבא:

1. נעבור על כל תת הקבוצות של עמודות A.
2. בהינתן אוסף עמודות, אם הוא בת"ל, נחפש את הצירוף הליניארי (אם קיים כזה) שיתן $Ax=b$ (אם הוא קיים, אז הוא יחיד)
3. נבחר את הוקטור x שנותן את הערך המקסימלי
   (א) המשפט מבטיח לנו שאם למערכת יש פתרון כלשהו אז יש לה גם פתרון אופטימלי שימצא בדרך הזו

4.1.2 מציאת הפתרון האופטימלי: חלק שני - מהצד הגאומטרי

בצד הגאומטרי רק פנינו לאינטואיציה בינתיים, עוד לא הגדרנו כלום. יכולנו לדמיין את הצורות במרחב, וטענו שהפתרון האופטימלי מתקבל תמיד בקודקוד. נרצה לטעון לקשר שבין bfs לקודקוד של פאון. תחילה נפתח את המושגים הדרושים על מנת לדון באופן פורמלי בצדדים הגאומטריים של הבעיה.

הגדרה: על מישור $H \subseteq \mathbb{R}^n$ הוא קבוצה מהצורה:

$$\{x \in \mathbb{R}^n :< a, x >= \alpha\}$$

זו קבוצה ממימד אחד פחות מ$\mathbb{R}^n$. כל H כנ"ל מגדיר חלוקה של המרחב לשלושה חלקים: H, ועוד שני חצאי מרחב:

$$H^+ = \{x \in \mathbb{R}^n :< a, x >> \alpha\}$$

$$H^- = \{x \in \mathbb{R}^n :< a, x >< \alpha\}$$

הערה: הרבה פעמים נכתוב H^+ כאשר אנחנו מתכוונים בפועל ל$H^+ \cup H$, וכך גם לגבי H^-, צריך להבין מההקשר.

הגדרה: **פוליהדר** ב$\mathbb{R}^n$ הוא חיתוך סופי של חצאי מרחב.
פוליהדר חסום נקרא **פאון**.

כלומר, אפשר לתאר פאון על ידי רשימה סופית של **אי שיוויונים מגדירים**.

יש גם דרך אחרת לחשוב על פאונים יש משפט האומר שהגדרה שקולה לפאון ב$\mathbb{R}^n$ היא בתור הקמור של קבוצה סופית של נקודות.
כלומר, אם $X = \{x_1, .., x_p\} \subseteq \mathbb{R}^n$, אז

$$convX = \{\sum_{i=1}^{p} \alpha_i x_i : \forall i, \alpha_i \geq 0, \sum_{i=1}^{p} \alpha_i = 1\}$$

(לא נוכיח את המשפט).

למשל, הפאון הn מימדי המקיים שלכל i, $0 \leq x_i \leq 1$ הוא הקובייה הn מימדית, והיא הקמור של כל הוקטורים מהצורה $(\epsilon_1, .., \epsilon_n)$ כאשר כל $\epsilon_i \in \{0,1\}^n$. כלומר, הקמור של נקודות שהקודקודים שלהם זה סדרה של 0 ו1.

נזכור שקבוצה S ב$\mathbb{R}^n$ היא קמורה אם לכל $x, y \in S$, גם כל הקטע $[x, y] \subseteq S$, כאשר

$$[x, y] = \{\alpha x + (1 - \alpha)y : \alpha \in [0, 1]\}$$

למה: חיתוך של קבוצות קמורות הוא קבוצה קמורה.

הגדרה: **תת מרחב אפיני** הוא הזזה של תת מרחב ליניארי.
כלומר, אם W הוא מרחב ליניארי, אז $W + k$, כאשר k הוא וקטור כלשהו, הוא תת מרחב אפיני. למשל, במישור, תת מרחב ליניארי זה ישר העובר דרך הראשית, ואז תת מרחב אפיני זה הזזה שלו.

הגדרה: אם $S = \{x_1, .., x_k\} \in \mathbb{R}^n$ קבוצה סופית, **המרחב האפיני הנפרש על ידי** S הוא:

$$aff(S) = \{\sum_{i=1}^{k} \alpha_i x_i : \sum_{i=1}^{k} \alpha_i = 1\}$$

נשים לב שבשונה מקמור, לא דרשנו שהα_iים יהיו אי שליליים.

כמו כן, נגדיר את **המימד** של תת מרחב אפיני להיות המימד של תת המרחב הליניארי שמתאים לו.

הערה: באופן כללי יותר, אם $T \subseteq \mathbb{R}^n$, אז הקמור שלה $convT$ הוא החיתוך של כל הקבוצות הקמורות המכילות את T.

הגדרה: יהי P פאון וH על מישור ב$\mathbb{R}^n$. נאמר שH הוא **על מישור תומך** של P אם $P \subseteq H^+$ וגם $P \cap H \neq \phi$.
כלומר, כל הפאון נמצא מעל העל מישור, וגם החיתוך שלהם לא ריק.
במקרה כזה, נאמר ש$P \cap H$ הוא **פאה** $(face)$ של P מימד

$$affdim(P \cap H)$$

לפאה 0 מימדית קוראים **קודקוד**.
לפאה חד מימדית קוראים **צלע**.
לפאה $d - 1$ מימדית כש$affdimP = d$ קוראים **דופן** $(facet)$.

משפט: יהי $P = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$ הפאון המתאים לתוכנית ליניארית LP בצורה סטנדרית, ויהי $v \in P$. אז התנאים הבאים שקולים:

1. v הוא קודקוד של P
2. v הוא bfs של LP

הוכחה:
$2 \Leftarrow 1$: כיוון שv קודקוד, יש על מישור תומך $H = \{x :< c, x >= \gamma\}$, $P \subseteq H^-$, $H \cap P = v$.
זאת אומרת, $v \in P$ לפי ההגדרה, $< c, v >= \gamma$, ולכל $z \in P$, $z \neq v$. יוצא שv הוא האופטימום (היחיד) לLP. כפי שראינו מקודם, האופטימום מתקבל בbfs, וכיוון שיש אופטימום יחיד נסיק שv הוא bfs.

$1 \Rightarrow 2$: v הוא bfs , כלומר $Av = b$, $v \geq 0$ ו$supp(v)$ מתאים לקבוצת עמודות בת"ל בA.
המטרה היא למצוא על מישור תומך לP שפוגש את P בv, ובה בלבד. בעצם אנחנו מחפשים וקטור c כך ש$< c, v >\geq< c, z >$ לכל $z \in P$, $z \neq v$ (כך נקבל שהעל מישור שמקבל את הערך $\gamma =< c, v >$ יחתך עם P רק בv, וכך נקבל שv קודקוד של P כנדרש).
נניח שv כתוב כך שהקואורדינטות החיוביות מופיעות בהתחלה, וכל השאר אפסים. נניח $v = (x_1, .., x_k, 0, ...0)$, ואז נגדיר $c = (0, .., 0, -1, ..., -1)$.
כלומר, c מקבל 0 בכל קואורינטה בה v שונה מאפס, ו-1 בכל קואורדינטה בה v מקבל 0. אז $< c, v >= 0$.
נרצה להראות שלכל $z \in P$, $0 >< c, z >$. z הוא וקטור אי שלילי. מעניין רק החלק השני של z, כיון שהחלק הראשון מתאפס במכפלה הפנימית. אם z חיובי - סיימנו. אם z כולו אפס, אז $supp(z) \subseteq supp(v)$ - אבל ברגע שקבענו את התומך אז v נקבע ביחידות, כלומר $z = v$, סתירה לבחירתו. כלומר קיבלנו שv הוא אכן קודקוד ולכן סיימנו ■

4.2 אלגוריתם הסימפלקס

אלגוריתם הסימפלקס הוא הפתרון שיאפשר לנו לפתור את השאלה באופן יעיל. הרעיון: נתחיל מקודקוד כלשהו של הפאון, כלומר מפתרון bfs כלשהו של המערכת, ומבין השכנים שלו נעבור לשכן שמשפר את פונקציית המטרה.
עד כאן מה שאמרנו הוא מאוד מעורפל - יש הרבה מאוד מה להוסיף ולהגדיר במדוייק, מה שכתוב למעלה זה עוד לא ממש תיאור של אגלוריתם.

הגדרה: אם u וv הם קודקודים של פאון P, ואם הקטע $[u, v]$ הוא צלע (כלומר פאה חד מימדית) של P, אז אומרים שהקודקודים u, v הם **שכנים**.
כלומר גאומטרית, אנחנו נמצאים בקודקוד, מסתכלים על כל הקודקודים השכנים שלו ובוחרים את הקודקוד השכן שנותן לנו את ה"שיפור" הטוב ביותר. יש משפחה שלמה של אלגוריתמי סימפלקס, לפי הבחירה של השכן: כפי שאמרנו בכל שלב מסתכלים על קודקוד כלשהו, ועוברים לשכן כלשהו שהוא "לא גרוע ממנו", כלומר שהערך של פונקציית המטרה עבורו היא לכל הפחות מה שהקודקוד המקורי השיג. השאלה היא, אם יש כמה קודקודים פוטנציאליים, כיצד נחליט לאיזה קודקוד לעבור? השיקול הזה מגדיר משפחה של אלגוריתמים, שכולם נקראים אלגוריתמי סימפלקס.

איך זה מתבטא ברמה האלגברית? בעצם נחשוב כאילו אנחנו עובדים עם בסיס - את הקבוצה הבלתי תלויה הליניארית שקיבלנו אפשר להרחיב אותה לבסיס. כעת, אפשר להוסיף עמודה ולהיפתר מהוקטור שיוצר תלות ליניארית. צריך להוכיח שהפעולה האלגברית של לעבור בין בסיסים למרחב שתיארנו מקבילה לפעולה של המעבר על הקודקודים שתיארנו.

שאלה - מדוע אופטימליות מקומית גוררת אופטימליות גלובלית? כלומר, נגיד שנמצאים בקודקוד שידוע שאף אחד מהשכנים שלו לא טוב ממנו (כלומר המעבר לכל שכן מקטין את פונקציית המטרה). על לפניו נראה שהיה יכול להיות קודקוד שאיננו שכן שלו שמנצח אותו.
בשביל אינטואיציה, נניח שיש קודקוד א' כלשהו, שיש לו 2 שכנים שהוא טוב מהם, ויש קודקוד אחר "רחוק", קודקוד ד' שהוא טוב ממנו. נוסיף צלע, בין קודקוד ד' לקודקוד א'. פונקציית המטרה היא פונקציה במשתנה אחד כשמצטמצים לישר הזה, ואז נקבל שהיא יכולה להיות מונטונית עולה או מונוטונית יורדת - ומהנתון היא בהכרח עולה. אבל אז בסביבה של הקודקודים אפשר לכתוב את הערך של הפונקציה בתור צירוף קמור של הערכים הקרובים, אבל הם יורדים ולכן זו סתירה.

כמה זמן לוקח לבצע את האלגוריתם הזה? מסתבר שלא ממש יודעים. הראו שיש גרסא של אלגוריתם סימפלקס שלוקח זמן מעריכי. משתמשים באלגוריתמי סימפלקס כל הזמן, והם עובדים יפה, ולמעשה לא ממש מבינים למה. עם הזמן הראו עבור עוד ועוד גרסאות (גרסאות נבדלות על פי ההחלטה בין הקודקודים הלא מקלקלים לאן ללכת).

שבוע 12

דיברנו פעם קודמת על אלגוריתם הסימפלקס. אנחנו מבינים את המושג של קודקוד גם במובן הגאומטרי וגם במובן האלגברי.

אמרנו בשיעור שעבר שלתכנון ליניארי יש את התכונה החשובה שאופטימום מקומי שווה לאופטימום גלובלי (למעשה יש את התכונה הזו גם לתכנון קמור ולא רק לתכנון ליניארי). כלומר, כפי שאמרנו שיעור שעבר, אם מצאנו קודקוד שהוא טוב יותר מכל שאר השכנים שלו, אנחנו יכולים לדעת שהוא יהיה טוב יותר מכל שאר הקודקודים.

אלגוריתם סימפלקס עובר בזה אחר זה מקודקוד של הפאון לקודקוד שכן, בלי להפסיד בפונקצית המטרה (כלומר בכל מעבר שהוא עושה, הוא לא מקטין את פונקציית המטרה).
כעת נרצה להבין מה זה שכנות, מה זה צלע, ואיך נראה המעבר בין קודקודים שכנים בעולם האלגברי. מבחינה גאומטרית קל לנו לתאר איך זה נראה ־ אבל בשביל אלגוריתם צריך לתאר מבחינה אלגברית.

כפי שראינו, צלע זה חיתוך של המישור התומך עם הפאון, שהממד האפיני שלה הוא 1.

איך רואים שכנות במערכת המושגים האלגברית? כפי שאמרנו בשיעור שעבר, זה מאופיין על ידי המעבר בין קבוצה בת״ל אחת לקבוצה בת״ל אחרת בעמודות של המטריצה.

4.2.1 שאלות (חלקן פתוחות) על תכנון ליניארי

ראינו שיש פתרון לבעיית LP ־ אלגוריתם הסימפלקס.
שאלה מאוד מתבקשת היא האם יש אלגוריתם פולינומי לLP? התשובה היא שכן. הוא נקרא אלגוריתם אליפסואידי, ואנחנו נלמד אותו בהמשך.
האלגוריתם האליפסואידי הוא פולינומיאלי באופן שהוגדר בסיבוכיות: הוא פולינומיאלי בגודל הקלט, כאשר גודל הקלט הוא כמות הביטים הנדרשים כדי לייצג את הקלט לבעיה. כלומר, כמות הביטים הנדרשים כדי לייצג את A, b, c.

האם יש גרסא לאלגוריתם הסימפלקס שרץ בזמן פולינומי?
אמרנו שאלגוריתמי סימפלקס הם משפחה של אלגוריתמים, ולא מדובר באלגוריתם יחיד, כאשר הם נבדלים אחד מהשני לפי האופן שבו הם מחליטים לאן להתקדם כאשר יש מספר קודקודים אפשריים.
לא ידוע בשלב זה מהי התשובה לשאלה הזו. ידוע שלגרסאות רבות התשובה שלילית, כלומר יש גרסאות שיש להן קלטים עליהם הזמן יהיה לא פולינומיאלי. למשל, יש גרסא של האלגוריתם שכאשר הוא רץ על מאין קובייה תלת מימדים עם פונקציית מטרה מסוימת הוא סורק את כל הקודקודים ־ זמן אקספוננציאלי. במהלך הזמן הראו לכל מני גרסאות נוספות של אלגוריתמי סימפלקס שקיימים קלטים רעים עבורם ־ אך לא הוכח שלא קיים אלגוריתם הרץ בזמן פולינומיאלי, ויתכן מאוד שיש כזה.

העובדה הזו מובילה לשאלה עוד יותר מפתיעה ־ למה אלגוריתם סימפלקס עובד מצויין בפרקטיקה? הרי בעולם האמיתי משתמשים בו כל הזמן.

שאלה נוספת היא האם יש אלגוריתם הרץ בזמן **פולינומי חזק**? תהי $A_{m\times n}$ מטריצת הקלט לבעיה, b, c הוקטורים המתאימים. נסמן בk את מספר הביטים הנדרשים על מנת לייצג את A, b, c. נאמר שיש אלגוריתם הרץ בזמן פולינומי חזק הפותר את הבעיה אם קיים פולינום $p(m, n)$ כך שהוא מסיים ב$p(m, n)$ פעולות העובדות על מספרים המיוצגים בk ביטים.
זו שאלה פתוחה, ולא ידוע האם יש אלגוריתם הרץ בזמן פולינומי חזק.

האם תיתכן גרסא של אלגוריתם הסימפלקס שתרוץ בזמן פולינומיאלי?
במקרים מסויימים אפשר להבטיח שלא: אנחנו בעצם מסתכלים על פאון. אם לפאון יש קוטר "גדול", כיוון שאלגוריתם הסימפלקס עובר מקודקוד לקודקוד, לא משנה איזה בחירות נעשה לא נוכל להגיע למשהו יעיל. נוכל לשאול עבור פאון d מימדי המוגדר על ידי n דפנות, כמה גדול יהיה הקוטר של הגרף? האם הוא חסום על ידי פולינום בn וd?
זו גם כן שאלה פתוחה. היתה השערה (השערת הירש) שהיתה שהקוטר חסום על ידי משהו ליניארי בd וn ־ אך היא הופרכה. עדיין יכול להיות שהקוטר חסום על ידי משהו פולינומיאלי (אך לא ליניארי) בd וn.
נשים לב שאם התשובה שלילית, כלומר הקוטר של הפולינום לא בהכרח חסום על ידי d וn, לא תיתכן גרסא של אלגוריתם סימפלקס שתעבוד בזמן פולינומיאלי תמיד.

תחום מאוד מעניין הוא $smoothed\ analysis$. באופן הסתכלות זו, מסתכלים על כל עולם הקלטים לבעיית LP, ושואלים היכן נמצאים הקלטים הלא טובים לבעיה ־ איך הם נראים במרחב עולם הקלטים?
ספלמן וטנג, שהגו את שיטת ההסתכלות הזו, הראו שהקלטים הרעים לבעיה הם מאוד מבודדים. כלומר, הם נראים כמו נקודות במרחב ־ ומוקפים בקלטים טובים. המשמעות היא שאם ניקח קלט לבעיה ונבלגן אותו קצת, בהסתברות 1 נגיע לקלט קל. יש לניתוח הזה משמעויות חשובות

באופן כללי גם לבעיות NP-קשות: אפשר לנסות להסתכל על הקלטים לבעיות הקשות, ולהבין איך הם נראים. המיקום של הקלטים הקשים במרחב יכול להגיד לנו הרבה על הבעיה.

4.3 בעיות דואליות בLP

בהינתן אוסף של משוואות ליניאריות, נוכל לעשות צירוף ליניארי שלהן ולקבל עוד משוואה ליניארית.
בהינתן אוסף של אי שיוויונים ליניאריים, נוכל לעשות צירוף ליניארי עם סקלרים חיוביים ולקבל עוד אי שיוויון ליניארי.
בהינתן אוסף של משוואות ואי שיוויונים, נקרא לצירוף כזה **צירוף ליניארי מותר** (צירוף ליניארי כלשהו למשוואות, וצירוף ליניארי על ידי סקלרים חיוביים לאי שיוויונים). קיבלנו בעצם שצירוף ליניארי מותר של משוואות ואי שיוויונים ליניאריים נותנים עוד משוואות ואי שיוויונים.

עקרון הדואליות אומר שכל מה שניתן להסיק מאוסף של משוואות ואי שיוויונים ליניאריים, ניתן להסיק על ידי צירוף ליניארי מותר.
כלומר, כל המידע נמצא כבר בתוך המערכת, ואם נרצה להסיק משהו לגביה נוכל לקבל את מה שרצינו להסיק על ידי צירופים ליניאריים מותרים של המערכת.

בהשלכה לתכנון ליניארי, נניח שאנחנו רוצים למצוא את $\max < x, c >$ כאשר x מקיים את האילוצים $Ax = b$ ו$x \geq 0$ (בעיית תכנון ליניארי כלשהי). נוכל לשאול עבור γ כלשהו, האם קיים x כך שיתקיים $Ax = b, x \geq 0$ וגם $< c, x >\geq \gamma$. אם התשובה חיובית ־ נוכל למצוא x כזה. אם התשובה שלילית, עקרון הדואליות אומר לשנו שיהיה צירוף ליניארי מותר של אוסף אי השיוויונים הנ"ל שיביא לסתירה (למשל, $1 < 0$). זהו כלי מאוד חזק.
נשים לב שאם אנחנו יודעים עבור כל γ לענות על השאלה ־ האם קיים x המקיים את האילוצים וגם $< c, x >\geq \gamma$, נוכל גם לפתור את בעיית התכנון הליניארי (למשל על ידי חיפוש בינארי בערכי γ).

משפט הדואליות חלשה: יהיו x, y וקטורים העומדים באילוצים הבאים:

$$Ax = b,\ \ x \geq 0\ yA \geq c$$

אז מתקיים:

$$< c, x >\leq< b, y >$$

הוכחה: נסתכל על $yAx \in \mathbb{R}$: כיוון ש$Ax = b$, אז מצד אחד נקבל $yAx = y(Ax) =< y, b >$.
מצד שני, $yA \geq c$ ולכן נקבל $yAx = (yA)x \geq< c, x >$.
נחבר יחד ונקבל:

$$< c, x >\leq yAx =< y, b >$$

ונסיים ■

בהינתן בעיית LP המוגדרת על ידי האילוץ $Ax = b$, $x \geq 0$ כאשר מנסים למצוא את $\max < c, x >$, נוכל להסתכל על הבעיה הנתונה על ידי האילוץ $yA \geq c$ כאשר מנסים למצוא את $\min < b, y >$.
בעיה זו נקראת **בעיה הדואלית** לבעיית הLP הנתונה, כפי שראינו באלגו, והיא תסומן בDLP.

משפט הדואליות החלשה מבטיח לנו שכל פתרון לבעיה המקורית יושב מתחת לכל פתרון של בעיית הDLP. בהמשך נראה משפט האומר שבתנאים (קלים מאוד) יש לשתי הבעיות את אותו האופטימום, כלומר הפתרונות ממש נפגשים באמצע.

גם דואליות חלשה יכולה להיות מאוד מועילה: היא מאפשרת לנו למצוא חסם תחתון לבעיית המינימיזציה, וחסם עליות לבעיית המקסימיזציה.

נשים לב שיש לנו דרישה $x \geq 0$ אבל אין לנו דרישה מקבילה עבור y. אפשר להבין זאת כיוון שבמובן מסויים, y יגדיר את הצירוף הליניארי המותר על המשוואה $Ax = b$, בעוד שx יגדיר את הצירוף הליניארי המותר על $yA \geq c$. כיוון ש$yA \geq c$, נהיה חייבים שהצירוף יכלול מקדמים אי שליליים בלבד, ומכאן האילוץ $x \geq 0$. אך עבור המשוואה $Ax = b$, אין מגבלות על המקדמים בצירוף.

נמשיך לנסח משפטים שיעזרו לנו לפתח טענות לתחום הזה.

למת פרקש: תהי $A_{m\times n}$ מטריצה ממשית, $b\in\mathbb{R}^n$ וקטור. אז למערכת $Ax=b, x\geq 0$ **אין** פתרון אם ורק אם קיים $y\in\mathbb{R}^n$ כך ש$yA\geq 0$ וגם $<y,b><0$.

סקיצה של הוכחה:

$\Rightarrow$: נניח שקיים y כך ש$yA\geq 0$ וגם $<y,b>\leq 0$. נניח בשלילה שקיים x שפותר את המערכת. כלומר, $Ax=b \Leftarrow yAx=y(Ax)=<y,b><0$ על פי הנתון. אך מצד שני, $yAx=<yA,x>\geq 0$ לפי הנתון, וזו סתירה. לכן לא קיים x כזה.

$\Leftarrow$: נניח שלמערכת $Ax=b$, $x\geq 0$ אין פתרון. נשים לב שאם $x\geq 0$, אז Ax הוא צירוף ליניארי של עמודות A על ידי סקלרים אי שליליים $\Leftarrow$ למעשה, אפשר להכיל את כל הצירופים האלה מוכלים בתוך קונוס: נזכיר שקונוס ב$\mathbb{R}^n$ זו קבוצה שהיא סגורה לכפל בסקלר חיובי ולחיבור. חיתוך של קונוסים הוא קונוס, לכן אפשר לדבר על הקונוס הכי קטן המכיל קבוצה של וקטורים. אפשר להראות שהקונוס הנוצר על ידי אוסף של וקטורים הוא אוסף כל הצירופים הליניאריים שלהם על ידי מקדמים אי שליליים . נסמן בC את הקונוס הנוצר על ידי עמודות C, אז C מכיל בדיוק את כל האיברים מהצורה Ax כאשר $x\geq 0$.

לכן לומר שאין למערכת $Ax=b$ כך ש$x\geq 0$ פתרון זה בדיוק כמו לומר $b\notin C$. לכן נוכל להסתכל על העל־מישור שעובר בבסיס של C, נניח שהוא מוגדר על ידי $<y,x>=0$, ונסמן אותו H.

כיוון ש$yA\geq 0$, נקבל שלכל עמודה aב של A $<y,a>\geq 0$ ולכן זה נכון גם לכל צירוף ליניארי אי שלילי של עמודות $A \Leftarrow C\subseteq H^+$.

מצד שני, לפי הנתון $<y,b><0$, ולכן $b\in H^-$ ■

הערה: על מנת להוכיח במדוייק את למת פרקש יש צורך בכל מני משפטים גאומטריים שלא הוכחנו עדיין. בהמשך נראה הוכחה מדוייקת, שאכן תשתמש במשפטים המתאימים.

4.3.1 דוגמא בבעיה דואלית

נזכיר את אחת מהבעיות הליניאריות הראשונות שראינו: בעיית ניהול המשק. נניח שאנחנו מנהלים משק, ונרצה לדעת כמה להאכיל את הפרות - איזה סוגי מזון וכמה מכל אחד. יש לנו n סוגים של מזון, וm אבות מזון. יש לנו מטריצה A שממפה בין כל סוג מזון j לכמות האב מזון i שנמצא ביחידה שלו. בנוסף יש לנו וקטור c שמייצג את העלות של המזון, ווקטור b שמייצג את הצרכים התזונתיים של הפרה.

ראינו שהבעיה הליניארית האמורה היא בעצם מציאת $\min <c,x>$ תוך קיום האילוצים $Ax\geq b$ ו$x\geq 0$. כלומר, נדרוש שכל הצרכים התזונתיים של הפרות יהיו, כך שהמחיר יהיה מינימלי.

הבעיה הדואלית לבעיה הזו היא הבעיה המתקבלת מהיפוך התפקידים:

$$yA\leq c,\ y\geq 0$$

כך שאנחנו מחפשים את $\max <b,y>$.

מסתבר שבעיה זו היא גם כן בעיה מעניינת בפני עצמה. היא נקראת "הבעיה של הכימאי": הסיפור שמאחוריה הוא שנניח שיש כימאי שרוצה לפתוח מפעל המייצר מזון סינטטי לפרות. הוא רוצה לדרוש כמה שיותר עבור המזון שלו, ומצד שני המזון שלו חייב להיות מסוגל להתחרות בסוגי המזון האמיתיים. בבעיה הזו, y מייצג את וקטור המחירים. האילוץ $yA\leq c$ אומר שהמזון הסינטטי צריך להיות מסוגל להתחרות במחירי השוק של המזון הטבעי, ואנחנו רוצים למקסם את $<b,y>$, שזה העלות של להזין פרה עם המזון הסינטטי.

זו דוגמא למצב שבו הבעיה הדואלית היא גם כן מעניינת ובעלת משמעות.

נמשיך עם מושגים בגאומטריה קמורה:

למת ההטלה: תהי $C \subseteq \mathbb{R}^n$ קבוצה קמורה וסגורה. אז לכל $b \in \mathbb{R}^n$ קיימת ויחידה נקודה $z \in C$ שהיא הנקודה הקרוב ביותר לb. נקודה זו מאופיינת על ידי התנאי:

$$\forall x \in C, \quad < x - z, b - z > \leq 0$$

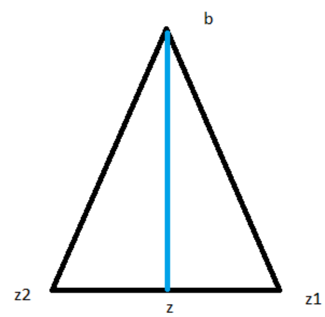
z שכזו נקראת **ההיטל**, או **ההטלה** של b על C.

נשים לב שהלמה מעניינת רק עבור $b \notin C$. אם $b \in C$, נבחר את $z = b$ ונסיים.

סקיצה של ההוכחה:

קיום: נסתכל על B - כדור המכיל את b ויש לו חיתוך לא ריק עם C. נסתכל על $B \cap C$ במקום על C. זו קבוצה קומפקטית כחיתוך של קבוצה קומפקטית עם קבוצה סגורה. הפונקציה שמתאימה לכל וקטור x את $\|b - x\|$ (מרחקו מb) היא רציפה, לכן קיים $z \in C$ עבורו מתקבל הערך המינימלי של פונקציה זו.

יחידות: אם נניח בשלילה שיש שתי נקודות שמשיגות את המינימום, נוכל לקחת את הנקודה שביניהם. אפשר לראות את זה בצורה גאומטרית: נסתכל על משולש שווה שוקיים, ואז על הנקודה שבין השוקיים.



על מנת להראות שיש אפיון על ידי התנאי הנתון, יש להראות שני דברים: z נתונה על ידי הביטוי הנ"ל, וגם כל נקודה המקיימת את התנאי הנ"ל היא z.

עבור z, נשים לב שב$\mathbb{R}^n$, עבור שני וקטורים x, y, הסימן של המכפלה הפנימית $< x, y >$ מתארת את הזוית ביניהם: $< x, y >= 0$ אם ורק אם x וy מאונכים, $< x, y >> 0$ אם ורק אם הזוית ביניהם חדה, ו$< x, y >< 0$ אם ורק אם הזוית ביניהם כהה. נניח בשלילה שקיים $x \in C$ כך שעבורו $< x - z, b - z >> 0$. אבל אז נקבל זוית חדה ביניהם, ונוכל להעביר גובה ולקבל נקודה קרוב יותר מz.

בכיוון השני,נניח שעבור $z \in C$ כלשהו, לכל $x \in C$ $< x - z, b - z > \leq 0$. נרצה להראות שz היא הנקודה הקרובה ביותר לb בC, כלומר שלכל $x \in C$,

$$\|b - x\| \geq \|b - z\|$$

נחשב:

$$\|b - x\|^2 = \|(x - z) + (z - b)\|^2 = \|x - z\|^2 + \|z - b\|^2 + 2 < x - z, z - b >$$

נשים לב ש$\|x - z\|^2 \geq 0$. לפי הנתון $< x - z, b - z > \leq 0 \iff - < x - z, b - z > \geq 0 \iff < x - z, z - b > \geq 0$ מליניאריות. כלומר, הוספנו גדולים אי שליליים ל$\|z - b\|^2$ ולכן נסיק

$$\|b - x\|^2 \geq \|b - z\|^2$$

ובכך נסיים ■

הערה: את הטענה שהתכונה מתקיימת עבור z הראנו באופן לא פורמלי, נראה אותה באופן פורמלי יותר בשיעור הבא.

שבוע 13

העולם של אופטימיזציה עבר כמה שלבים. באופן כללי בעיות אופטימיזציה ליניארית זה מה שאנחנו עוסקים בו כרגע, וזה כלי מאוד חזק. השלב הבא הולך לשני כיוונים: תכנון קמור (כאשר התחום הוא גוף קמור כלשהו ולא בהכרח פוליהדרון), את כל הדברים העיקריים שאנחנו רואים כאן אפשר להפעיל גם שם. זה כבר לוקח למרחק מאוד גדול. בבעיות שמנסים לפתור היום, יש גם תחומים לא קמורים - זה מה שהשנים הבאות של מחקר כנראה יקדמו.

יש בתחום הזה פער בין הפרקטיקה לתאוריה: במובנים מסויימים הפרקטיקה מובילה על התאוריה, כלומר יש דברים שעובדים שלא ידוע עדיין בדיוק למה.

אמרנו שאם אנחנו יודעים לפתור את בעיית הספיקות אנחנו יודעים לפתור את בעיית התכנון הליניארי. אנחנו יודעים שהפתרון האופטימלי מתקבל בקודקוד, ואנחנו יודעים כיצד למצוא קודקוד: קודקוד הוא המפגש של כמה דפנות, לכן חלק מאי השיוויונים שמגדירים את הפאון יהיו שיוויונות שם. לכן כאשר אנחנו מגיעים כבר לקרבה מסויימת בחיפוש נוכל כבר למצוא את הקודקוד הקרוב ביותר לאיפה שאנחנו נמצאים ובכך לסיים.

אם בעיית הספיקות פתירה - אפשר למצוא x שמעיד על כך. אם לא, אפשר כאמור למצוא צירוף ליניארי של משוואות שיראה זאת. לכן נקבל $LP \in NP \cap coNP$ (נזכיר ששפה היא בNP אם בהינתן "עד", קל לוודא שהיא בשפה - במקרה הזה הx שמעיד על כך, וששפה היא ב$coNP$ אם בהינתן "עד" קל לוודא שהיא לא בשפה - במקרה הזה הצירוף הליניארי המעיד על כך שהבעיה לא ספיקה. כבר אמרנו שאם יש פתרון לבעיית הספיקות אפשר בקלות להגיע לפתרון לבעיית התכנון הליניארי, ולכן נובע מכך שגם בעיית LP נמצאת ב$NP \cap coNP$).

השלמה להוכחת למת ההטלה: פעם שעברה תיארנו מבחינה גאומטרית מדוע, אם z הוא ההיטל אז לכל $x \in C$, בהכרח $< x - z, b - z >> 0$. נתאר באופן אלגברי למה זה נכון:

אם נניח בשלילה שקיים $x \in C$ כך ש$< x - z, b - z >> 0$, אז לכל $0 \leq t \leq 1$ גם

$$z + t(x - z) \in C$$

כי $x, z \in C$ וC קמורה. נחשב:

$$\|z + t(x - z) - b\|_2^2 = \|z - b\|^2 + t^2\|x - z\|^2 - 2t < b - z, x - z >$$

נשים לב שכאשר t מאוד קטן, אז $|t|$ גדול משמעותית מt^2. הגודל $< b - z, x - z >> 0$, ולכן אנחנו מחסירים כאן גודל חיובי. זה נכון שאנחנו גם מוסיפים את $t^2\|x - z\|^2$, אבל כאמור כיוון שהגודל t הרבה יותר דומיננטי מt^2, הגודל שאנחנו מוסיפים זניח לעומת מה שאנחנו מורידים. מכאן שניתן למצוא t עבורו $z + t(x - z)$ קרוב ממש לb$ מz, וזו סתירה לבחירת z ■

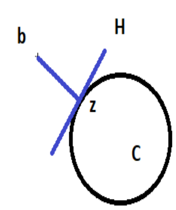
מסקנה: אם C קמורה וסגורה, $b \notin C$ וz היטל של b על C, אז העל-מישור

$$H = \{x :< x - z, b - z >= 0\}$$

תומך בC ומפריד את b מC.

זהו העל-מישור שעובר בz וניצב לC.

בציור:



המסקנות יאפשרו לנו לנסח את משפט ההפרדה:
משפט ההפרדה: תהי C קבוצה קמורה וסגורה ב־$\mathbb{R}^n$, ו־$b \notin C$, אז יש על־מישור H המפריד את b מ־C. בפרט, ניתן לקחת:

$$H = \{x :< x - z, b - z >= 0\}$$

כאשר z ההיטל של b על C.

כעת, כאשר יש לנו את משפטי ההפרדה האלה, נוכל להוכיח את למת פרקש שוב באופן מדוייק

הוכחה מדוייקת של למת פרקש: את הכיוון בו קיים y כך ש־$yA \geq 0$ וגם $< y, b >\leq 0$ אז לא קיים פתרון הראנו כבר (לא היה צורך בשימוש במשפטי הפרדה).
נראה את הכיוון השני:
ממשפט ההפרדה שכתוב למעלה, נובע שהעל מישור

$$H = \{x :< x, y >=< z, y >\}$$

כאשר z **מפריד חזק** את b מ־C, כאשר z ההיטל של b על C, ו־$y = z - b$ (העברנו אגפים במכפלה הפנימית). נרצה להראות ש־y הזה עונה לתנאים של הלמה. כלומר, נרצה להראות $yA \geq 0$ וגם $< y, b >< 0$.
נראה תחילה ש־$< y, z >= 0$. לפי למת ההטלה, $< y, z >\geq 0$: C הוא קונוס, $0 \in C$, ולפי למת ההטלה, לכל $x \in C$ $< x - z, b - z >\leq 0$.
נציב $b - z = -y$ ו־$x = 0$, ואז נקבל $< y, z >\geq 0$.
מדוע לא יתכן $< y, z >> 0$: מכיוון ש־C קונוס, ו־$z \in C$, אז גם $(1 + t)z \in C$ עבור $|t|$ קטן (קונוס סגור למכפלה חיובית). כעת:

$$\|(1 + t)z - b\|^2 = \|z - b\|^2 + t^2\|z\|^2 + 2t < z - b, z >$$

וזו נקודה קרובה יותר: t דומיננטי יותר מ־t^2 כאשר t קטן, ואז כיוון ש־$< b - z, z >\neq 0$ נוכל לבחור t כך שהגודל הזה יהיה קטן ממש מהמרחק בין z ל־b.
זאת אומרת שהעל מישור המפריד

$$H = \{x :< x, y >= 0\}$$

אם x הוא אחת העמודות של A, מקבלים

$$< x, y >\geq 0$$

ו־$yA \geq 0$ והראנו מקודם $< y, b >< 0$ ■

4.3.2 סקיצה של ההוכחה למשפט הדואליות החזקה

נראה כמה צעדים ונסביר מה אנחנו לא עושים. נסתכל על LP. במקום לדבר על אופטימזיציה נעבור לדבר על ספיקות: אנחנו שואלים האם יש $x \geq 0$ כך ש־$Ax = b$ וגם $< c, x >= \gamma$.
כל עוד התשובה חיובית, אפשר למצוא x כלשהו שיקיים תכונה זו. עקרון הדואליות כפי שאנחנו מבינים אותו אומר שכאשר התשובה שלילית, התוכנית הדואלית תספר לנו על כך.
נסתכל על מטריצה

$$\tilde{A} = \binom{A}{c}$$

השאלה היא האם יש $x \geq 0$ כך ש

$$\tilde{A}x = \binom{b}{\gamma}$$

כלומר b עם קואורדיטנה נוספת. כשהתשובה היא שלילית, כלומר γ גדול ממש מהמקסימום שהוא משיג, אז יש וקטור $\tilde{y}$ שמוכיח זאת:

$$\tilde{y}\tilde{A} \geq 0$$

ו

$$< \tilde{y}, \tilde{b} >< 0$$

לפי למת פרקש.
לקחת $\tilde{y}$ כזה ולהפכיל אותו במשהו חיובי לא משנה את התכונות של $\tilde{y}$.
אם הקואורדינטה האחרונה של $\tilde{y}$ היא שלילית. ננרמל אותו כך שהקו' האחרונה תהיה -1.

$$(y, -1)\binom{A}{c} \geq 0$$

וגם

$$< (y, -1), (b, \gamma) >< 0$$

כלומר, $yA \geq c$ וגם המכפלה הפנימית מקיימת $< y, b >< \gamma$. אבל כאן מתקבלת סתירה לדואליות חלשה.
אם הקואורדינטה האחרונה היא לא שלילית, אפשר להראות שמשהו בהנחות נופל: אחת התוכניות לא ספיקה או לא חסומה. זה הרבה עבודה להראות ולא נראה את זה כאן.

4.3.3 האלגוריתם האליפסואידי והרחבת הבעיה לתכנון קמור

אמרנו שיש קשר בין תכנון קמור לתכנון ליניארי. ההבדל היחיד שבין תכנון ליניארי לתכנון קמור הוא שבתכנון קמור אנחנו מאפשרים לפונקציית המטרה להיות על גבי גוף קמור כללי, ולא בהכרח פאון. נשים לב שפאון מוגדר על ידי רשימה סופית של אי שיוויונים ליניאריים. אפשר להראות (לא נראה זאת) שגוף קמור מוגדר גם כן על ידי רשימה של אי שיוויונים ליניאריים, אלא שרשימה זו לאו דווקא סופית.

האלגוריתם האליפסואידי הוא אלגוריתם שפותר את בעיית LP בזמן פולינומיאלי. מבחינה מעשית, לא משתמשים בו בגרסא הפשוטה שלו כל כך בעולם המעשי, אבל יש לו חשיבות תאורתית גדולה מאוד.

נזכיר שאם ידוע לנו כיצד לפתור את בעיית הספיקות, ידוע כיצד לפתור את בעיית האופטימיזציה. לכן נתמקד בתיאור פתרון בעיית הספיקות.

תיאור הבעיה:
יש לנו גוף קמור $\mathcal{P} \subseteq \mathbb{R}^m$, ונרצה לברר האם הוא ריק או לא.
נניח תמיד ש־$\mathcal{P}$ מוכל בתוך כדור סביב הראשית, כלומר קיים $R > 0$ כך ש־$\mathcal{P} \subseteq B(0, R)$. כמו כן, נניח שאם $\mathcal{P} \neq \phi$ אז קיים r כך ש־$B(0, r) \subseteq \mathcal{P}$ (כלומר אנחנו לא נתעסק בגופים קמורים "מוזרים מדי", ולא בגופים שאינם חסומים). נניח ש־R ו־r נתונים לנו כשאנחנו באים לפתור את הבעיה.

רעיון האלגוריתם יהיה להסתכל על תחום שידוע ש־$\mathcal{P}$ מוכל בתוכו, וכל פעם ננסה לצמצם אותו. אם יגיע שלב שבו כבר אי אפשר לצמצם, נסיק ש־$\mathcal{P}$ איננו ריק. אבל אם הצלחנו לצמצם לתחום שהוא מוכל ממש בתוך $B(0, r)$, נסיק ש־$\mathcal{P} = \phi$.

לא נעשה הכל באופן פורמלי לחלוטין, כיוון שיש הרבה פרטים טכניים שלא נרצה להעמיק בהם.

השאלה הראשונה היא כיצד נתון לנו הגוף הקמור? כלומר, במקרה של פאון היה קל: הוא היה נתון לנו על ידי אוסף סופי של אי שיוויונים. אך אמרנו שגוף קמור כללי מוגדר על ידי אוסף אינסופי של אי שיוויונים ליניאריים, וזה כבר בעייתי לספק בתור קלט לבעיה. לכן נניח שיש לנו אלגוריתמים יעילים (במונחי המימד m), להם נקרא **אוב שייכות** ו**אוב הפרדה**. כלומר, נניח שהגוף הקמור שלנו $\mathcal{P}$ מגיע יחד עם R, r ושני

האלגוריתמים האלה. אוב השייכות יהיה אלגוריתם שמקבל נקודה $x \in \mathbb{R}^m$, ומחזיר תשובה האם $x \in \mathcal{P}$ או לא. אוב ההפרדה יהיה אלגוריתם המקבל נקודה $x \in \mathbb{R}^m$, ובמידה ו$x \notin \mathcal{P}$ מחזיר על־מישור המפריד בין x ל$\mathcal{P}$. אם $x \in \mathcal{P}$ הוא מחזיר שx שייך ל$\mathcal{P}$.

נשים לב שלא מאוד קשה לתאר למשל איך נראים האלגוריתמים האלה במקרה ש$\mathcal{P}$ הוא כדור: כדי לבדוק שייכות, נוכל לחשב את המרחק מהראשית. כדי למצוא על־מישור מפריד, נוריד אנך לכדור.

תחילה ננסח מספר הגדרות:\

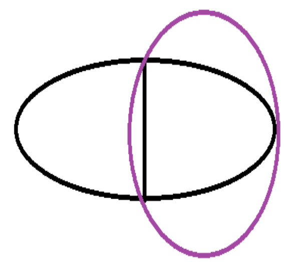
הגדרה: **אליפסואיד** ב$\mathbb{R}^m$ זו התמונה של כדור היחידה תחת העתקה אפינית הפיכה

הגדרה: בהינתן אליפסואיד, **המרכז** של האליפסואיד זה התמונה של הראשית (נשים לב שזה מושג שונה ממוקד של אליפסואיד).

טענה: לכל חצי אליפסואיד E ב$\mathbb{R}^m$ יש אליפסואיד F כך ש$E^+ \subseteq F$ וגם:

$$volF < (1-\frac{1}{m})volE$$

כלומר, זו טענה גאומטרית שאומרת לנו שאם יש לנו אליפסואיד שמחולק לחצי, נוכל להקיף את אחד מהחצאים על ידי אליפסואיד חדש, שהנפח שלו קטן ממש בסדר גודל מסויים מהנפח של האליפסואיד המקורי.
בציור:



<u>תיאור האלגוריתם:</u>
בשלב ההתחלתי, אנחנו יודעים $\mathcal{P} \subseteq E_0 = B(0,R)$, ונבחר $x_0 = 0$.
בכל שלב k, יש לנו אליפסואיד E_k ונקודה x_k שהיא המרכז של האליפסואיד. נשאל את אוב ההכלה האם $x_k \in \mathcal{P}$. אם כן ־ סיימנו, כיוון שהראנו ש$\mathcal{P}$ לא ריק.
אחרת, נפנה לאוב ההפרדה ונקבל על־מישור שמפריד בין x_k ל$\mathcal{P}$. אם נסמן על־מישור זה בתור $\tilde{H}$, אנחנו יכולים לדעת $\mathcal{P} \subseteq \tilde{H}^+$ ־ ובכך צמצמנו את שטח המחייה של $\mathcal{P}$. נזיז את $\tilde{H}$ לעל־מישור מקביל שעובר דרך x_k. נסמן על מישור זה בתור H. נגדיר $E^+ = H^+ \cap E_k$ ־ וזהו חצי אליפסואיד המכיל את $\mathcal{P}$.
נפעיל את הטענה הגאומטרית שכתובה למעלה, ונקבל אליפסואיד F כך ש$E^+ \subseteq F$, ומתקיים:

$$volF < (1-\frac{1}{n})volE$$

בסופו של דבר (אם לא סיימנו לפני כן), אחרי מספר פולינומיאלי של צעדים אנחנו נקטין בגורם של e^{-m} את תחום המחייה ־ מתישהו הוא יהיה קטן ממש מr. בשלב הזה נסיים ונסיק ש$\mathcal{P}$ ריק.