כמה מההיבטים החישוביים של אלגברה ליניארית

בפרק זה נעמוד על כמה מהבעיות האלגוריתמיות המתעוררות בלימוד האלגברה הליניארית. הרבה מהנושאים הללו הופיעו בלימודיכם בקורסים הרלוונטיים אולם כאן נעמיק את הדיון האלגוריתמי. נעיר תחילה על היבט חשוב נוסף של בעיות אלה- המדובר ביציבות נומרית. כאשר יש לבצע חישובים במספרים ממשיים אנו נתקלים בקושי הנובע מכך שבמחשב מיוצגים המספרים בדרגת דיוק חסומה הנובעת מאורך המילה (מספר הביטים) שבו מיוצגים המספרים במחשב שלנו. כתוצאה מזה החישובים שאנו מבצעים במחשב הם רק בגדר קירוב לתוצאה התיאורטית. שיקולים אלה מובילים מצד אחד לתורה של אנליזה נומרית. במסגרת זו מחפשים שיטות חישוב שיושפעו מעט ככל האפשר מבעיות הדיוק המוגבל במחשב. זהו תחום שלם שלא נדון בו כאן. גם על מבנה החומרה של מחשבים משפיעים השיקולים האלה באופן מהותי. ייצוגם של מספרים ממשיים בשיטת נקודה צפה ואחרות ואופן המימוש של הפעולות האריתמטיות היסודיות הם בין השיקולים המרכזיים בתכנון מחשבים. אלו נושאים המכוסים בקורס "מבנה מחשב".

עוד הסתייגות אחת לפני שניכנס בעובי הקורה:

הדיון בהיבטים החישוביים של אלגברה ליניארית הוא רחב ועמוק. נושא זה בפני עצמו יכול לספק בקלות חומר לקורס אוניברסיטאי שלם ואנו נאלץ רק לטעום מקצת ממנו.

להלן סיכום העניינים העיקריים שבהם נדון בפרק זה:

א. נזכיר כמה בעיות אלגוריתמיות/אלגבריות בסיסיות שבהן האלגוריתם כהגדרתו אופטימלי (בזה לא נדון לעומק).

ב. מאידך גיסא, נראה שתי בעיות מרכזיות שבהן יכול אלגוריתם מתוחכם להשיג חסכון של ממש: כפל מטריצות (ע"פ Strassen) וכפל פולינומים (בעזרת טרנספורטים פורייה הבדיד DFT)

ג. גם כאשר איננו יודעים מהי הסיבוכיות של בעיות אלגבריות מסויימות עדיין ניתן לומר דבר מה על קושיין היחסי. נערוך כמה השוואות מסוג זה.

ד. נכניס מערכת מושגים המאפשרת לדון בחישוב אלגברי מקורב ונדגים כמה משימושיה.

ה. נציין את החשיבות שבמציאת פירוקים למטריצות: בפרט PLU ו- SVD.

ו. נסביר את הרקע הרחב יותר שמאחורי טרנספורים פורייה הבדיד (DFT) וכמה משימושיו.

ז. נראה איך במגוון של הקשרים מועיל להציג (במדוייק או בקירוב) אובייקט מתימטי מסובך כצירוף של "אבני בניין" פשוטות יותר.

יש כמה מהפעולות הבסיסיות של אלגברה ליניארית שבהן האלגוריתם הנאיבי הנובע מן ההגדרה הוא אופטימלי. כך למשל חישוב המכפלה הסקלרית ("מכפלה פנימית") של שני וקטורים n- מימדיים $\langle x,y\rangle=\sum x_iy_i$ אכן מצריך n פעולות כפל. זו עובדה לא טריביאלית שלא נוכיח כאן ומכוסה בקורסים בסיבוכיות החישוב. כיוצא בו גם כפל של מטריצה ממשית $A_{n\times n}$ בווקטור n- מימדי x, כלומר,
$(Ax)_i=\sum_j a_{ij}x_j$ אכן מחייבת n^2 פעולות כפל, כלומר אין לבעיה זו אלגוריתם יעיל יותר מאשר האלגוריתם הנאיבי הנובע מן ההגדרה. זוהי שוב תוצאה מתקדמת למדי בתורה של סיבוכיות החישוב והיא אינה מוכחת בקורס זה. לעומת זאת החישוב ע"פ ההגדרה של כפל מטריצות ממשיות $n\times n$ $AB=C$ מצריך n^3 פעולות כפל. בהמשך נכיר אלגוריתם יעיל יותר (של Strassen) המצריך רק $O\left(n^{2.81}\right)$ פעולות אריתמטיות. מרגע שמתבררת נקודת המבט שלנו, מתעוררת באחת שאלות אלגוריתמיות רבות, מהי הסיבוכיות של חישוב הדטרמיננט של מטריצה $n\times n$? של פתרון מערכת משוואות ליניאריות $Ax=b$? מהו המחיר החישובי של הצבת ערך בפולינום? של כפל פולינומים? ועוד ועוד.

נפתח בדיון קצר בפתרון מערכות של משוואות ליניאריות, נזכיר ראשית את שיטת האלימינציה של גאוס. בשיטה זו פותרים את מערכת המשוואות הליניאריות $Ax=b$ ע"י כך שמציגים את A בצורה $A=PLU$. כאן P היא מטריצת תמורה, L היא מטריצה

משולשת תחתונה ו- U מטריצה משולשת עליונה. יוצא שלמערכת הליניארית $Ax = b$ יש הצורה $LUx = P^{-1}b = u$. מערכת מהצורה $Ly = u$ ניתן לפתור באופן סידרתי כך:
$y_1 = \frac{u_1}{\ell_{11}}$ אז מציבים ופותרים את המשוואה ל- y_2 וכן הלאה. אז עלינו לפתור מערכת מהסוג $Ux = y$ ואותה ניתן לפתור באופן דומה, הפעם כשמתחילים ב- $x_n = \frac{y_n}{u_{nn}}$.

אנו נרצה להרחיב את היריעה ולדון במצבים שבהם מערכת המשוואות $Ax = b$ אינה אלא קירוב לבעייה האמיתית. כך למשל אנו יכולים לתאר מצבים רבים שבהם איברי המטריצה A כמו גם איברי הווקטור b מתקבלים ע"י מדידה של מערכת טבע כלשהי ולכן נגועים ברעש. מתברר גם כי אותו הדיון עוזר לטפל בבעיות הנובעות משגיאות עיגול במשתנים שבווקטור x. שגיאות אלה מקורן, כאמור, במגבלות הדיוק של המחשבים שלנו.

על מנת להמחיש את הרעיון נתאר מצב כזה: המשתנים הנעלמים שהם הקואורדינטות של הווקטור x הם n גדלים פיזיקליים שאותם ברצוננו לקבוע. אם כל ניסוי שאנו מבצעים מגדיר משוואה ליניארית אחת, אז באופן תיאורטי n ניסויים בלתי תלויים הכרחיים ומספיקים על מנת לקבוע את הווקטור ה- n מימדי x ביחידות.

נניח שבהיותנו מודעים לבעיית הרעש בניסויים, אנחנו מקפידים ומבצעים מספר גדול באופן ניכר $n \gg m$ של ניסויים. אנו מנסים, אם כן לפתור מערכת משוואות ליניאריות $Ax = b$ שבה A מטריצה ממשית $m \times n$, x ווקטור n מימדי ו- b ווקטור m מימדי. אילו היו המדידות שלנו נטולות רעש, היתה המערכת הנ"ל קובעת את x באופן חד ערכי. דא עקא שהמדידות שלנו מושפעות מרעש והמערכת הנ"ל כרגיל כלל אינה פתירה (באופן כללי, אם במערכת משוואות ליניאריות יש יותר משוואות מנעלמים, אנו איננו מצפים שתהיה פתירה כלל).

יוצא שהפרשנות הסבירה היחידה למערכת המשוואות $Ax = b$ היא "מצא ווקטור x כך שהווקטור Ax יהיה קרוב ככל האפשר לווקטור הנתון b". על מנת למלא את השאלה הזו בתוכן עלינו לפתח תחילה מושג של "קירבה בין ווקטורים", מה שמוביל אותנו לדיוננו הבא.

נורמות ווקטוריות: יהיה V מרחב ווקטורי מעל הממשיים. אנו רוצים לייחס לווקטורים $u \in V$ "אורך". את האורך של v נסמן ב- $\|v\|$ ונקרא לו הנורמה של v. נצפה ממושג הנורמה לשתי תכונות:

א. אי שיוון המשולש: $\forall_{u,u \in V} \|u+v\| \le \|u\| + \|v\|$.

ב. הומוגניות: לכל $u \in V$ ולכל $\alpha \in R$ $\|\alpha v\| = |\alpha| \|v\|$

האם יש נורמות כנ"ל?

מסתבר שיש אינסוף נורמות שונות ויש גם נורמות רבות ושונות שהן מעניינות. שלוש הנורמות החשובות ביותר נקראות נורמת ℓ_2, נורמת ℓ_1 ונורמת ℓ_∞. להלן הגדרותיהן והסימון המקובל להן. נניח ש- V הוא מרחב n מימדי

$$x = (x_1, \ldots, x_n)$$

הוא ווקטור אופייני ב- V .

את נורמת ℓ_2 אתם מכירים בעצם מקדמת דנא $\|x\|_2 := \sqrt{\sum_{i=1}^{n} x_i^2}$

נורמת ℓ_1 מוגדרת כך: $\|x\|_1 := \sum_1^n |x_i|$

ונורמת ℓ^∞: $\|x\|_\infty := max\,|x_i|$

נעיר ששלושתן הן מקרים פרטיים של משפחה אינסופית של נורמות הקרויות נורמות ℓ_p (כאן $\infty > p \ge 1$) $\|x\|_p := \left(\sum |x_i|^p\right)^{\frac{1}{p}}$

שימו לב ש- $lim_{p \to \infty} \|x\|_p = \|x\|_\infty$

(הוכיחו)

עתה נוכל לנסח את שאלתנו הקודמת בצורה מלאה: נקבע איזושהי נורמה ווקטורית $\|\bullet\|$ על ווקטורים m מימדים.

קלט: מטריצה ממשית $A_{m\times n}$ כש $m>n$ ודרגתה $rank\left(A\right)=n$ ווקטור m מימדי b.

פלט: ווקטור n- מימדי x כך ש $\|Ax-b\|$ מזערי (הווקטור Ax קרוב ל- b ככל האפשר).

כפי שכבר רמזנו הנורמה החשובה ביותר וכרגיל (אך לא תמיד) גם הנוחה ביותר לחישוב היא נורמת ℓ_2. זהו אכן המקרה שבו נטפל. כאשר הנורמה שבשימוש היא נורמת ℓ_2 נקראת הבעייה הנ"ל בעיית הקירוב בריבועים פחותים (least squares approximation).

בהינתן A,b כנ"ל מחפשים אם כן x כך ש: $min\sum_i\left(\sum_j a_{ij}x_j-b_i\right)^2$

נגזור ע"פ x_k ונקבל כי לכל k:

$\sum_i a_{ik}\left(\sum_j a_{ij}x_j-b_i\right)=0$

נסדר מחדש את המחוברים:

$$\sum_j x_j\sum_i a_{ik}a_{ij}=\sum_i a_{ik}b_i$$

נרשום זאת כך:

$$\sum_j x_j\sum\left(a_{ki}^T\right)a_{ij}=\sum_i a_{ik}b_i$$

כלומר

$$\sum_i\left(A^TA\right)_{kj}x_j=\left(bA\right)_k$$

כלומר $A^TAx=A^Tb$. נעיר ש- A^TA לא סינגולרית, מפני שע"פ ההנחה $rank\left(A\right)=n$ אם $A^TAu=0$ אז גם $u^TA^TAu=0$ כלומר $\|Au\|^2=0$ ולכן $u=0$. קיבלנו אם כן $x=\left(A^TA\right)^{-1}A^Tb$.

הערה: כאן ראינו איך לפתור במקורב מערכת משוואות ליניאריות כשהקירוב נעשה ב- ℓ_2. באמצעות תכנון ליניארי ניתן לפתור את הבעייה של פתרון מקורב ב- l_∞ אולם אין עוד מקרים מעניינים רבים שבהם יש לנו אלגוריתמים יעילים לפתרון הבעייה.

דרך חליפית לביצוע החשבון הנ"ל:

$\parallel Ax-b\parallel_2^2=\left(x^TA^T-b^T,Ax-b\right)=x^TA^TAx-2b^TAx+\|b\|^2$

נסמן $Z=A^TA$ אז הנ"ל שווה ל:

$$\sum_{k,l}x_kx_lz_{kl}-2\sum_k x_k\left(A^Tb\right)_k+\|b\|^2$$

נגזור ע"פ x_k ונאפס אז נקבל $2\sum z_{kl}x_l=2\left(A^Tb\right)_k$ כלומר $Zx=A^Tb$

$x=Z^{-1}A^Tb=\left(A^TA\right)^{-1}A^Tb$. למטריצה $\left(A^TA\right)^{-1}A^T$ קוראים הפסאודו הופכית של A או הופכית $Moore-Penrose$ של A.

תיאור חלופי לבעיית הפתרון המקורב של מערכת משוואות:

אוסף הווקטורים $L=\{Ax\mid x\in\Re^n\}$ הוא תת מרחב n מימדי של $\Re^m$. הבעייה היא למצוא את הנקודה הקרובה ביותר ל- b וזהו כמובן ההיטל של b על L. הנוסחה שפיתחנו לפסאודו הופכי נותנת ביטוי מפורש להיטל.

להלן שימוש יפה וחשוב לבעייה שפתרנו: התאמת עקומות (Curve fitting). נפתח בעובדה מתימטית בסיסית שאולי כבר מוכרת לכם מהקורסים באנליזה מתימטית:

יהיו $i=1,.....,n$ (x_i,y_i) n נקודות במישור כשה - x_i שונים זה מזה. אז יש פולינום אחד ויחיד P ממעלה $<n$ כך שלכל i מתקיים $P\left(x_i\right)=y_i$. תזכורת מהירה

על ההוכחה: יחידות נובעת מכך שאם P,Q שני פתרונות אז $P-Q=0$ ב- n הנקודות $x_1,.....x_n$ אבל לפולינום P-Q יש לכל היותר n-1 שורשים אלא אם $P\equiv Q$.

קיום: ניתן להוכיח א' ע"י תרגום הבעייה למערכת משוואות ליניאריות במקדמי הפולינום P. מטריצת המערכת היא מטריצת ונדרמונדה שאינה סינגולרית ולכן המערכת פתירה ופתרונה מספק את ה- P המבוקש. לחילופין: ניתן לרשום את P במפורש ע"י מה שנקרא נוסחת האינטרפולציה של לגרנז': $P\left(x\right)=\sum_i y_i\prod_{j\neq i}\frac{x-x_j}{x_i-x_j}$.

הסבר: כאשר $x=x_r$, כל המחוברים שבהם $i\neq r$ מתאפסים כי הגורם $x_r-x_r=0$ מופיע במונה. מאידך המחובר ה-r נותן לנו y_r כי הגורם הכופל את y_r הוא 1.

שוב תהיינה -$(a_1,b_1)....(a_p,b_p)$ נקודות במישור ונרצה למצוא פולינום ממעלה $p>q\geq$ המקרב את הנקודות האלה היטב ככל האפשר במובן של l_2. כלומר מחפשים $P\left(x\right)=\sum_{j=0}^r\alpha_jx^j$ כך ש $min\sum_i\left(P\left(a_i\right)-b_i\right)^2$ אם נגדיר $A=\left(\begin{array}{cccc}1 & a_1 & .. & a_1^q\\ 1 & a_2 & ... & a_2^q\\ & & \vdots & \\ 1 & a_p & .. & a_p^q\end{array}\right)$

$b=\left(\begin{array}{c}b_1\\ \vdots\\ b_p\end{array}\right)$ אז רצוננו למצוא ווקטור $z=\left(\alpha_o,\ldots,\alpha_q\right)$ כך ש $min_z\left\|Az-b\right\|_2$ הביטוי שפיתחנו לעיל מספק אם כן תשובה לבעיית האינטרפולציה המקורבת.

נורמות מטריציוניות

איך נמדוד את המרחק בין שני מטריצות ואיך נייחס "גודל" למטריצה? ראשית ניתן להתייחס למטריצה כאל ווקטור וכך למשל ניתן להגדיר $\left\|A\right\|_2:=\sqrt{\sum a_{ij}^2}$ (זו קרויה נורמת l_2, אך גם נורמת פרובניוס או נורמת הילברט שמיט) או $\left\|A\right\|_\infty=max\left|a_{ij}\right|$ וכו'. אולם מטריצות הן גם אופרטורים לינאריים ומעניין להבין איך מתקשר מושג הנורמה עם היבט זה של מטריצות. ואכן אומרים שנורמה $\|\cdot\|$המוגדרת על מטריצות ריבועיות היא נורמה מטריציונית במובן הצר אם $\left\|A\cdot B\right\|\leq\left\|A\right\|\cdot\left\|B\right\|$(נדרש כמובן שניתן להכפיל את A,B זו בזו). לא קשה להראות למשל שנורמת $\left\|\bullet\right\|_2$ מקיימת תנאי זה.

דוגמאות חשובות של נורמה מטריציונית המקיימת תנאי זה, הן הנורמות האופרטוריות. הן עונות על השאלה "עד כמה יכולה הנורמה של ווקטור לגדול כשמופעל עליו האופרטור A". כלומר, $\left\|A\right\|:=\max_x\frac{\left\|Ax\right\|}{\left\|x\right\|}$יש טעם בהגדרה זו לכל בחירה של נורמה במונה ונורמה במכנה. כך למשל: $\left\|A\right\|_{2\to2}=\max\frac{\left\|Ax\right\|_2}{\left\|x\right\|_2}$

$\left\|A\right\|_{1\to\infty}=\max_x\frac{\left\|Ax\right\|_\infty}{\left\|x\right\|_1}$וכו'.

לפעמים נוח להשתמש בניסוח חליפי כבדוגמה הבאה: $\left\|A\right\|_{\infty\to1}=\max\left\|Ax\right\|_1$ כשהמקסימום הוא ע"פ כל הווקטורים כך ש $\left\|x\right\|_\infty=1$. כמובן, הנ"ל תקף לכל זוג נורמות.

יש נורמות אופרטוריות שיש להן תיאור חילופי מענין. כך למשל,

$\left\|A\right\|_{\infty\to\infty}=\max_i\sum_j\left|a_{ij}\right|$

באופן דומה$\left\|A\right\|_{1\to1}=\max_j\sum_i\left|a_{ij}\right|$ (אלו תרגילים לא קשים - הוכיחו זאת).

טענה חשובה נוספת המצריכה הוכחה היא: $\left\|A\right\|_{2\to2}=\sqrt{\lambda_{max}\left(A^TA\right)}$. כפי שעוד נראה בהמשך לגודל זה קוראים הערך הסינגולרי הגדול ביותר של A.

הערה: לפעמים מקצרים ורושמים את $\left\|A\right\|_{2\to2}$ כ- $\left\|A\right\|_2$. שימו לב שלסימון זה גם משמעות אחרת כמתואר לעיל ויש לברר מהי הכוונה.

הערה נוספת שכדאי להעיר היא זו: הנורמה האופרטורית $\left\|A\right\|$ היא המספר החיובי הקטן ביותר c כך שלכל ווקטור x מתקיים $\left\|Ax\right\|\leq c\left\|x\right\|$. רוב השימושים שלנו בנורמה האופרטורית יעברו דרך צורה זו של הגדרה.

עתה נראה איך להשתמש במושגים אלה כדי לנתח את הרגישות של מערכת מש-

וואות לינאריות לרעש. נניח שלפנינו מערכת משוואות $Ax = b$ המייצגת איזושהי מערכת מהטבע או מערכת הנדסית. אנו יודעים לפתור מערכת כזו. אולם רצוננו להבין מה יקרה לפתרון אם חלה טעות, בין אם באיברי A ובין אם בקואורדינטות של b. כלומר: $Ax = b$, היא המשוואה האמיתית ופתרונה x הוא הפתרון שאותו אנחנו רוצים למצוא. המערכת שנפתור בפועל היא $(A+\Delta)\hat{x} = b+\beta$ מפני שלאיברי המטריצה A נוסף רעש בצורת המטריצה Δ ולאיברי הווקטור b נוסף ווקטור הרעש β ופתרון המערכת כפי שהיא מוגשת לנו הוא הווקטור $\hat{x}$. אנו מעוניינים להעריך את הטעות היחסית בחישוב הנובעת מהרעש. כלומר אם $\hat{x} - x = \delta$, רצוננו להבין את הגודל $\frac{\|\delta\|}{\hat{x}}$(ואנו מקווים שהטעות היחסית הזו קטנה).

לכן $(A+\Delta)(x+\delta) = b+\beta$

$Ax + A\delta + \Delta x + \Delta\delta = b + \beta$ אבל $Ax = b$לכן $A\delta + \Delta\hat{x} = \beta$

נובע ש $\|\delta\| \le \|A^{-1}\| (\|\beta\| + \|\Delta\| \|\hat{x}\|)$

$\frac{\|\delta\|}{\|\hat{x}\|} \le \|A^{-1}\| \|A\| \left(\frac{\|\Delta\|}{\|A\|} + \frac{\|\beta\|}{\|A\|\|\hat{x}\|}\right)$. מתברר שהשגיאה היחסית בפתרון נשלטת ע"י $\|A\| \|A^{-1}\|$שנקרא גורם ההצגה (Condition Number) של A.

שימו לב שלכל מטריצה A מתקיים $k(A) = \|A^{-1}\|\|A\| \ge \|I\| = 1$. למטריצות בעלות גורם הצגה נמוך (קרוב ל-1) מתאימה מערכת משוואות "מוצגת היטב" ואילו כאשר $k(A)$ גדול אנו אומרים שהמערכת "אינה מוצגת היטב", כלומר שהפיתרון רגיש אפילו לטעויות קלות בנתוני הבעייה.

גירסה מדוייקת יותר של ההערכות נותנת:

משפט (ללא הוכחה): אם $\|\Delta\| \le \frac{1}{A^{-1}}$ אז $\frac{\|\delta\|}{\|x\|} \le \frac{k(A)}{1-k(A)\frac{\|\Delta\|}{\|A\|}} \left(\frac{\|\Delta\|}{\|A\|} + \frac{\|\beta\|}{\|b\|}\right)$.

מעניין ש $k(A)$ מודד באיזו מידה קרובה המטריצה A להיות סינגולרית. זה נעשה במשפט הבא:

משפט: $\frac{1}{k(A)} = \inf\{\frac{\|\Delta\|}{\|A\|} | \det(A+\Delta) = 0\}$.

אנו נראה רק כי $\frac{1}{k(A)} \ge \inf \frac{\|\Delta\|}{\|A\|} | \det(A+\Delta) = 0$.

נמצא מטריצה Δ כך ש $A + \Delta$ סינגולרית ו-$\|\Delta\| \le \frac{\|A\|}{k(A)} = \frac{1}{\|A^{-1}\|}$. יהיה x ווקטור יחידה שבו מושגת הנורמה $\|A^{-1}\|$ כלומר $\|A^{-1}x\| = \|A^{-1}\|$. נסמן $y = \frac{A^{-1}x}{\|A^{-1}\|}$. נשים לב כי y הוא ווקטור יחידה ונגדיר את Δ כמטריצה (מדרגה 1) $\Delta := -\frac{xy^T}{\|A^{-1}\|}$. נשים לב תחילה ש $A + \Delta$ סינגולרית, ואכן:

$$(A+\Delta)y = Ay - \frac{xy^Ty}{\|A^{-1}\|} = \frac{x}{\|A^{-1}\|} - \frac{x}{\|A^{-1}\|} = 0$$

יש עוד להראות כי $\|\Delta\| \le \frac{1}{\|A^{-1}\|}$, כלומר שהנורמה האופרטורית של המטריצה xy^T היא ≤ 1 ואכן, אם $\|z\| = 1$ אז

$$\|xy^Tz\| = \|x\| < y, z > \le 1$$

כי x, y, z ווקטורי יחידה.

ההגדרה הוריאציונית של ערכים עצמיים:

משפט: תהיה A מטריצה ממשית סימטרית $n \times n$ ויהיו $\lambda_1 \ge ..... \ge \lambda_n$ הערכים העצמיים של A אז: $\lambda_1 = max\frac{x^TAx}{\|x\|^2}$

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

הוכחה: נוכיח את הטענה לגבי λ_1 ההוכחה ל-λ_n דומה. נשים לב תחילה שאם v הו"ע המתאים לע"ע λ_1 אז $\frac{v^TAv}{\|v\|^2} = \frac{\lambda_1\|v\|^2}{\|v\|^2} = \lambda_1$ ולכן $\max_x \frac{x^tAx}{\|x\|^2} \ge \lambda_1$.

כדי להוכיח את האי שוויון הנגדי נזכיר את העובדה הבאה שלמדתם באלגברה לינ-יארית: אם A מטריצה ממשית סימטרית, אז יש למרחב בסיס אורתוגונלי של ווקטור-ים עצמיים של A. יהיו $u_1, \ldots, u_n$ ווקטורים עצמיים מנורמלים $(\|u_i\| = 1)$המתאימים

לע"ע $\lambda_n \geq .... \geq \lambda_1$. ניתן להציג כל ווקטור x באופן יחיד כ: $x = \sum \alpha_i u_i$. בגלל האורתוגונליות של ה- u_i מתקיים $\|x\|^2 = \sum \alpha_i^2 \|u_i\|^2$ באופן דומה גם

$$x^T A x = \sum \lambda_i \alpha_i^2 \|u_i\| \leq \lambda_1 \sum \alpha_i^2 \|u_i\|^2 = \lambda_1 \|x\|^2$$

כפי שטענו.

האפיון הזה יאפשר לנו להסיק את הנוסחה המועילה הבאה לנורמה האופרטורית $2 \to 2$, כדלקמן:

$\|A\|_{2\to 2} \sqrt{\lambda_{max}\left(A^T A\right)}$

הוכחה: נפעיל את הנוסחה הוריאציונית על הערך העצמי הגדול ביותר של $A^T A$ דהיינו, $\lambda_{max}\left(A^T A\right) = max \frac{x^T A^T A x}{\|x\|^2} = max \frac{\|Ax\|^2}{\|x\|^2} = \left(max \frac{\|Ax\|}{\|x\|}\right)^2 = \|A\|_{2\to 2}^2$.

בהמשך נקרא ל- $\sqrt{\lambda_{max}\left(A^T A\right)}$ הערך הסינגולרי המקסימלי של A.

תזכורת: מטריצה W נקראת אורתוגונלית אם $WW^T = I$. לא קשה לראות שהתנאים הבאים שקולים: 1. W אורתוגונלית, 2. השורות של W הן מערכת אורתונורמלית, 3. העמודות של W הן מערכת אורתונורמלית.

למטריצה אורתוגונלית יש התכונות הבאות כטרנספורמציה ליניארית: היא שומרת על אורכים (במובן של ℓ_2) ושומרת על זויות. השמירה על האורך פירושה שלכל ווקטור x מתקיים

$$\|Wx\| = \|x\|$$

ואכן נחשב $\|Wx\|_2^2 = x^T W^T W x = x^T I x = \|x\|_2^2$ מכיוון שהאורך נשמר, שמירת הזויות משמעותה שלכל שני ווקטורים x,y מתקיים $\langle Wx, Wy\rangle = \langle x, y\rangle$ ואכן $y^T W^T W x = y^T I x = < x, y >$. נוסיף גם שאם $\langle Wx, Wy\rangle = \langle x, y\rangle$ לכל זוג ווקטורים נובע גם ש- W אורתוגונלית אבל לא נזדקק כאן לעובדה זו.

הערה נוספת (להשכלה כללית, לא נחוץ בקורס הזה): כאשר השדה הוא שדה המספרים המרוכבים ולא הממשיים מוגדרת המכפלה הפנימית כ: $\langle x, y\rangle = \sum x_i \bar{y_i}$ ולכן $\langle Wx, y\rangle = \langle x, W^* y\rangle$ כאשר $W^* = \left(\bar{W}\right)^T$ (צמוד ומוחלף). למטריצה מרוכבת המקיימת $XX^* = I$ קוראים מטריצה אוניטרית ויש לה תכונות אנלוגיות לאלה של מטריצה אורתוגונלית.

שימו לב שאם X, Y מטריצות $n \times n$ אורתוגונליות אז גם XY אורתוגונלית, שהרי $(XY)(XY)^T = XYY^T X^T = XIX^T = XX^T = I$.

SVD- פירוק ערכים סינגולרים

יש לבעייה הזו מספר תיאורים אפשריים. נתחיל בנקודת המבט הגיאומטרית: אליפסואיד זו התמונה של כדור היחידה תחת העתקה ליניארית. לחליפין: יוצאים מכדור היחידה, בוחרים מערכת קואורדינטות אורתוגונלית ובכ"א מהכיוונים מבצעים מתיחה בגורם σ_i. תהיה S ספירת היחידה ב- $\mathbb{R}^m$.

$S = \{x \in \mathbb{R}^m \mid \|x\|_2 = 1\}$ ותהיה $A_{m\times n}$ מטריצה ממשית כש- $m \geq n$ והדרגה של A היא n. נביט ב- AS זהו אליפסואיד n- מימדי המשוכן ב- $\mathbb{R}^m$. יהיו $\sigma_1 \geq ...... \geq \sigma_n > 0$ אורכי הצירים של האליפסואיד. אם $u_1, ..., u_n$ ווקטורי היחידה בכיווני הצירים של AS אז $u_1, ..., u_n$ נקראים הווקטורים הסינגולריים השמאליים של A. הווקטורים הסינגולריים הימניים של A אלו n ווקטורי היחידה $v_1, ..., v_n$ כך ש- $Av_i = \sigma_i u_i$ נסכם זאת:

$$A \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix}$$

ובסימון מטריציוני: $AV = UD$ כאן D מטריצה אלכסונית $n \times n$. $V_{n\times n}$ מטריצה אורתוגונלית (כלומר $VV^T = V^T V = I$). U מטריצה $m \times n$ עם עמודות אורתונורמליות ($U^T U = I$) ולכן $A = UDV^T$ להצגה זו קוראים SVD מצומצם. פירוק SVD מלא מושג ע"י

הרחבת U למטריצה אורתוגונלית $m \times m$ וריפוד D ב- $m - n$ שורות של אפסים כלומר, $A = UDV^T$. כש- $U_{m\times m}$ אורתוגונלית $D_{m\times n}$ "אלכסונית" ו- $\mathrm{V}_{\mathrm{n}\times\mathrm{n}}$ אורתו-גונלית. עכשיו קל לראות מדוע AS הוא אליפסואיד: הכפל ב-V^T מסובב את הספירה, הכפל ב- D מותח את n הקואורדינטות ומוסיף עוד $m - n$ קואורדינטות של אפסים. לבסוף, U מסובבת את האליפסואיד ה- n מימדי ב- $\mathbb{R}^m$.

מבט נוסף על פירוק **SVD** של מטריצות

נתחיל בדיון כללי וקצר במושג "סיבוכיות". קשה לתת הגדרה ממצה ומדובר בדיון אינטואיטיבי בלבד. יחד עם זאת ניתן להצביע על תבנית כללית: אנחנו מזהים אוסף בסיסי של "אבני בניין" וכן מספר פעולות המאפשרות לצרף אותן למבנה מורכב יותר. השאלה היא, בהינתן מבנה כלשהו, כמה אבני יסוד וכמה פעולות נדרשות ליצירתו. בהקשר של סיבוכיות החישוב פעולות היסוד יכולות להיות למשל פעולות אריתמטיות והמבנה הגדול יכול להיות חישוב כגון חישוב דטרמיננטה והשאלה היא מהו המספר המזערי של פעולות אלמנטריות בחישוב כלשהו של דטרמיננטה.

אולם למוטיב הכללי הזה יש דוגמאות מסוגים אחרים. כך, למשל, נראה שדרגה של מטריצה היא מין מדד סיבוכיות שכזה. מדוע? נזכיר את העובדה הבאה מאלגברה ליניארית.

משפט: הדרגה של מטריצה A היא המספר הקטן ביותר r כך שניתן לבטא את A בצורה $A = M_1 + \dots\dots + M_r$ כש -M_i מטריצות בדרגה 1. (תזכורת: למטריצה M יש דרגה 1 אם יש שני וקטורים x,y כך ש- $M_{ij} = x_i y_j$). המשפט הזה מסביר מדוע ניתן לראות את הדרגה של מטריצה כמדד לסיבוכיות. אבני הבניין שלנו הן המטריצות מדרגה 1 - אובייקטים פשוטים לכל הדעות. הדרך המותרת לצירוף מטריצות בדרגה 1 היא ע"י סיכום ומדד הסיבוכיות הוא "מהו המספר המזערי של אבני בניין הדרושות כדי ליצור את המטריצה הנתונה A"? הרקע שכבר הכנו מביא אותנו להתעניין בוריאציה הבאה על שאלה זו:

בעייה: תהיה נתונה מטריצה A ויהיה נתון גם $k < rank\,(A)$. מהו הקירוב הטוב ביותר ל- A ע"י מטריצה B מדרגה $\le k$. כלומר מחפשים את $min\{\|A - B\| \mid rank\,(B) \le k\}$. כאן הנורמה היא הנורמה האופרטורית $2 \to 2$, אבל כפי שיסתבר, אותה ה- B נותנת גם את הפתרון הטוב ביותר כשמדובר בנורמת Frobenius. הטיפול בשאלה זו מוביל אותנו גם הוא למושג של פירוק ערכים סינגולרים Singular Value Decomposition=SVD.

משפט ה-SVD: לכל $A_{m\times n}$ ממשית, $m \ge n$, יש הצגה $A = UDV^T$, כש-$U_{m\times m}$ אורתוגונלית, $V_{n\times n}$ אורתוגונלית ו-$D_{m\times n}$ "אלכסונית". האיברים $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_n \ge 0$ שבאלכסון נקבעים ביחידות. נדחה את הוכחת המשפט ונציין תחילה מסקנה חשובה הנובעת ממנו.

הערה (ללא הוכחה): אם A ריבועית וה-σ_i שונים זה מזה אז גם u_i, v_j נקבעים ביחידות עד כדי סימן.

מסקנה: יהיה $k < m,n$ ותהיה D_k המטריצה שבה נותרים רק k האיברים הראשונים שב-D. תהיה $A_k = UD_kV$. אז

$$\sigma_{k+1} = \|A - A_k\|_{2\to 2} = \min_{rank(B)\le k} \|A - B\ |_{2\to 2}$$

.

ועוד הערה אחת בטרם נמשיך: תהיה A_kהמטריצה המקרבת ביותר את A מבין המטריצות מדרגה $\le k$. כמו כל מטריצה מדרגה k, ניתן לבטא את A_k כסכום של מטריצות מדרגה 1. העובדה המפתיעה היא שכאשר מביטים בייצוג זה, מתברר ש $k - 1$ המטריצות הראשונות מבין המטריצות מדרגה 1 שסכומן הוא A_k מסתכמות ל A_{k-1}. במילים אחרות A_k מתקבלת מ A_{k-1} ע"י הוספת מטריצה אחת מדרגה 1. זו כאמור עובדה מפתיעה המצדיקה את טענתנו שאלגוריתם ה-SVD מוצא את "אבני הבניין" שמהן בנוייה A (שהן מטריצות מדרגה 1). כשמסכמים אותן בזו אחרי זו מתקבלים קירובים הולכים ומשתפרים ל-A. ביישומים רבים מופיעה התופעה הבאה:

ל A יש מספר קטן, נאמר t של ערכים סינגולריים גדולים וכל הערכים הסינגולריים שאחריהם קטנים במידה ניכרת. במקרה כזה מהווה כרגיל A_t קירוב מעשי מוצלח למטריצה A.

לפני שנוכיח את המסקנה נספק לה אינטרפרטציה כפי שהיא מופיעה ביישומים מעשיים רבים: את ההצגה $A = UDV^T$ (כאן A מטריצה $m \times n$ ו $m > n$) אפשר גם לרשום כך:

$$A = \sum_{j=1}^{n} \sigma_j u_j v_j^T$$

כאשר u_j, v_j היא העמודה ה j ב U, V בהתאמה. זאת רואים ע"י כך שמפרקים את D לסכום של n מטריצות שה j מביניהן יש בה σ_j במקום ה (j,j) וכל שאר האיברים הם 0.

שימו לב שכ"א מהמטריצות $\sigma_j u_j v_j^T$ היא מדרגה 1. הטענה היא אם כן שהקירוב הטוב ביותר מדרגה $\leq k$ ל A היא $UD_kV^T = \sum_{j=1}^{k} \sigma_j u_j v_j^T$.

הוכחת המסקנה: נעיר כי $\|A - A_k\|_{2\to 2} = \sigma_{k+1}$ נובע מהאפיון שנתנו לנורמת $2 \to 2$ באופן כללי. נניח כי $rank(B) \leq k$ וכי $\|A - B\|_{2\to 2} < \sigma_{k+1}$. אם כן, $W = \ker(B)$ תת מרחב $n-k$ מימדי $W \subset \mathbb{R}^n$. לכל $w \in W$ מתקיים $Aw = (A-B)w$. אם כן:

$$\|Aw\|_2 = \|(A-B)w\|_2 \leq \|A-B\|_2 \|w\|_2 < \sigma_{k+1}\|w\|_2$$

מאידך גיסא, נתבונן בתת מרחב W' הנפרש ע"י $k+1$ הוקטורים הסינגולריים הראשונים של A. זהו תת מרחב $k+1$ מימדי, ומתקיים $\|Av\| \geq \sigma_{k+1}\|v\| \quad \forall v \in W'$.

מדוע? יהיו $v_1, \ldots, v_{k+1}$ $(k+1)$ הווקטורים הסינגולריים הימניים הראשונים. הם וו-קטורי יחידה ניצבים זל"ז ו $Av_i = \sigma_i u_i$. יהיה $x = \sum_{i=1}^{k+1} \alpha_i v_i$ ווקטור כלשהו במרחב W' הנפרש ע"י $v_1, \ldots, v_{k+1}$. אז $\|x\|^2 = \sum_{i=1}^{k+1} \alpha_i^2$. נזכור שגם ה u_i הם ווקטורי יחידה ניצבים זל"ז ולכן

$$\|Ax\|^2 = \|\sum \alpha_i A v_i\|^2 = \|\sum \alpha_i \sigma_i u_i\|^2 = \sum_{i=1}^{k+1} \alpha_i^2 \sigma_i^2 \geq \sigma_{k+1}^2 \sum_{i=1}^{k+1} \alpha_i^2 = \sigma_{k+1}^2 \|x\|^2$$

אבל ל W' התת מרחב ממימד $k+1$ ול W המרחב ממימד $n-k$ יש חיתוך הכלל לא רק את ווקטור ה-0 וכל ווקטור $0 \neq$ בחיתוך z חייב לקיים $\|Az\| < \sigma_{k+1}\|z\|$ וכן $\|Az\| \geq \sigma_{k+1}\|z\|$ וזו סתירה.

ועוד מילה על היישום האופייני של משפט ה SVD: בסיטואציות מעשיות רבות מתברר שכבר קירוב מדרגה נמוכה של מטריצה A יכול להוות תחליף נאות ל A. לעובדה זו יש שימושים לדחיסה של תמונות וגם לסינון רעשים מתמונות.

נחזור עתה להוכחת משפט ה SVD: יהיה $\sigma_1 = \|A\|_{2\to 2}$. אז יש ווקטור v_1 כך ש: $\|v_1\|_2 = 1$ ו- $\|Av_1\|_2 = \sigma_1$. נסמן $Av_1 = \sigma_1 u_1$. נשלים את v_1 ל- $v_1, v_2, \ldots, v_n$ בסיס אורתונורמלי כלשהו ל- $\mathbb{R}^n$. באופן דומה נשלים את u_1 לבסיס אורתונורמלי $u_1, \ldots u_m$ ל- $\mathbb{R}^m$. תהיה V_1 המטריצה האורתוגונלית שעמודותיה הן $v_1, \ldots, v_n$. וכיו"ב U_1 המטריצה האורתוגונלית שעמודותיה הן $u_1, \ldots, u_m$. אז

$$U_1^T A V_1 = S = \begin{pmatrix} \sigma_1 & w^T \\ 0 & B \end{pmatrix}$$

. כאן B מטריצה $(n-1)\times(m-1)$ ו w^T וקטור שורה $(n-1)$-מימדי. פירוש השויון הזה הוא ש-

$$U_1^T A v_1 = \begin{pmatrix} \sigma_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

(שימו לב שאיננו טוענים בעצם דבר על העמודה השנייה ואילך).
אבל $Av_1 = \sigma_1 u_1$ והטענה היא לכן ש:

$$U_1^T u_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

ואכן זה נובע מהעובדה ש $\|u_1\| = 1$ ו u_1 ניצב לשאר השורות של U_1.
ואולם

$$\left\| \begin{pmatrix} \sigma_1 & w^T \\ 0 & B \end{pmatrix} \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} \right\|_2 \geq \sigma_1^2 + \|w\|_2^2 = \sqrt{\sigma_1^2 + \|w\|_2^2} \left\| \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} \right\|_2$$

ולכן $\|S\|_{2\to2} \geq \sqrt{\sigma_1^2 + \|w\|_2^2}$. אבל U_1, V_1 הן מטריצות אורתוגונליות ולכן $\|S\|_{2\to2} = \|A\|_{2\to2} = \sigma_1$. יוצא ש $w = 0$. כלומר

$$U_1^T A V_1 = \begin{pmatrix} \sigma_1 & 0 \\ 0 & B \end{pmatrix}$$

. באינדוקציה, יש ל B פירוק ערכים סינגולריים $B = U_2 D' V_2^T$ ולכן

$$A = U_1 \begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & D' \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & V_2 \end{pmatrix} V_1^T$$

והמשפט מוכח, מפני ש

$$U_1 \begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix}$$

היא מטריצה אורתוגונלית וכן גם

$$V_1 \begin{pmatrix} 1 & 0 \\ 0 & V_2 \end{pmatrix}$$

(כ"א מהן היא מכפלה של שתי מטריצות אורתוגונליות).
איך מוצאים פירוק SVD באופן יעיל? לא ניכנס לפרטי העניין אך נסביר את הרעיון העיקרי. נניח כי $A = UDV^T$ פירוק ערכים סינגולריים של A. בפרט $AV = UD$ וכן $A^T U = VD$. לכן

$$\begin{pmatrix} 0 & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} V & V \\ U & -U \end{pmatrix} = \begin{pmatrix} V & V \\ U & -U \end{pmatrix} \begin{pmatrix} D & 0 \\ 0 & -D \end{pmatrix}$$

. יוצא שהעמודות במטריצה

$$\begin{pmatrix} V & V \\ U & -U \end{pmatrix}$$

הן הו"ע של

$$\begin{pmatrix} 0 & A^T \\ A & 0 \end{pmatrix}$$

ולכן אם נחשב פירוק ערכים עצמיים של

$$\begin{pmatrix} 0 & A^T \\ A & 0 \end{pmatrix}$$

נקבל באחת את U, D, V ולכן את פירוק הערכים הסינגולריים של A.