תרגול #1

ב' - 23/3/20

נרכוש בסיס מתמטי לקורס.

נושאי התרגול: 1- חזרה על אלגברה לינארית: העתקות לינאריות, נורמות ומ"פ.

2- פירוקי SVD, EVD

3- הטלות אורתוגונליות

4- הנורמה הספקטרלית

מטלות: * תרגילים (תאורטי + תכנותי): 4/6 תרגילים טובים קובע

* מטלות מעבדה: ציון בינארי עבור ה-8 מטלות מעבדה.

* בוחן אמצע - 10% מגן

* האקתון - עד 5 נק' בונוס

1. חזרה על אלגברה לינארית

(1) העתקות לינאריות

העתקה לינארית: יהיו $V\in R^n, W\in R^m$ מרחבים וקטוריים. פונקציה $T:V\to W$ נקראת העתקה לינארית של V ל-W אם לכל $u,v\in V$, $c\in R$ **מתקיים:**

- **אדטיביות:** $T(u+v)=T(u)+T(v)$
- **כפל בסקלר:** $T(c\cdot v)=c\cdot T(v)$

★ אם $V,W<\infty$, אז קיימת מטריצה (והפיכה) שמייצגת את ההעתקה: לכל $v\in V$ קיימת מטריצה A כך ש: $T(v)=Av$

★ **נזכיר:**

$$Im(A) := \{w\in W \mid w=Ax,\ x\in V\} = col(A)$$
$$Im(A^T) := \{x\in V \mid x=A^T w,\ w\in W\} = row(A)$$
$$Ker(A) := \{x\in V \mid Ax=0\}$$

העתקה אפינית: העתקה $T:V\to W$ ($V\in R^n, W\in R^m$ מ"ו) המקיימת: $T(u)=Au+w$, כאשר: $u\in V$, $w\in W$.

★ עבור $w=0$: T ה"ל. עבור $w\neq 0$: T אינה ה"ל ולכן לא מייצגת תת-מרחב וקטורי.

דרגת A: יהא $A\in R^{n\times m}$. הדרגה של A שווה למקסימום של מס' השורות הבלתי-תלויות לינארית ב-A. נסמן: $rank(A)$.

טענה: עבור $A\in\mathbb{R}^{n\times m}$ מתקיים: $rank(A)=dim(Row(A))=dim(Col(A))$

- **A מדרגה מלאה/חלקית:** עבור $A\in\mathbb{R}^{n\times m}$, אם $rank(A)=\min(n,m)$ אז A מדרגה מלאה. אחרת נאמר ש-A מדרגה חלקית.

- **מטריצה ריבועית הפיכה (לא-סינגולרית):** תהא $A\in\mathbb{R}^{n\times n}$ מטריצה ריבועית. A נקראת הפיכה אם קיימת מטריצה $B\in\mathbb{R}^{n\times n}$ כך ש: $BA=AB=I_n$. נסמן את B המטריצה ההפוכה ל-A: A^{-1}.

טענה: תהא A מטריצה ריבועית. הבאים שקולים:

1. A הפיכה (לא-סינגולרית)
2. A מדרגה מלאה
3. $det(A)\neq 0$
4. $im(A)=\mathbb{R}^m$ = כל המרחב
5. $Ker(A)=\vec{0}$

שקול ע"י משפט המימדים השלישי

דוגמאות:

(ייצוג מטריציוני של מע' משוואות)

נניח מע' של n משוואות בה"ב, המכילות n נעלמים כך:

$$\forall i\in[n]: \sum_{j=1}^{n}\beta_j\cdot x_{ij}=y_i$$ (i משוואה, j נעלם)

ונרצה למצוא את המשתנים $\{\beta_j\}_{j=1}^{n}$. נסמן: $\underline{\underline{y}}=\begin{bmatrix}y_1\\ \vdots\\ y_n\end{bmatrix}\in\mathbb{R}^n$.

כמו כן, לכל $i\in[n]$, נסמן: $x_i=(x_{i1},\dots,x_{in})$, וכל ה-$\{x_i\}$

נכתוב למטריצה: $\underline{\underline{X}}=\begin{bmatrix}\top\\ (x_{i1},\dots,x_{in})\\ \bot\end{bmatrix}\in\mathbb{R}^{n\times n}$.

ונרצה למצוא $\beta=\begin{bmatrix}\beta_1\\ \vdots\\ \beta_n\end{bmatrix}\in\mathbb{R}^n$ שפותר את מע' המשוואות כך שמתקיים: $y=X\cdot\beta$.

<u>מה יש לנו עד כה?</u> שורות המטריצה X בה"ב $\Leftrightarrow$ $n=rank(X)$

$\Leftrightarrow$ X מדרגה מלאה $\Leftrightarrow$ קיימת X^{-1} $\Leftrightarrow$ $X^{-1}X=I$

<u>ולכן מצאנו פתרון:</u> $y=X\beta\Rightarrow X^{-1}y=X^{-1}X\beta\Rightarrow\beta=X^{-1}y$

☆ הבעיה הכללית נקראת <u>Linear Regression</u>: מטרה: קשר בין משתנים לתשובה והוא מתואר ע"י הווקטור β. בשיעור הבא נעמיק במה שקורה כאשר X לא ריבועית/הפיכה.

נורמות ומ"פ (ב)

• **נורמה:** פונקציה $\|\cdot\|: \mathbb{R}^m \to \mathbb{R}$ כך שלכל $u,v\in\mathbb{R}^m, a\in\mathbb{R}$

מתקיים: 1. אי-שלילי: $\|v\| \geq 0$ ו- $\|v\|=0 \Leftrightarrow v=0\in\mathbb{R}^m$

2. הומוגניות: $\|a\cdot v\| = |a|\cdot\|v\|$

3. אי"ש ה-Δ: $\|u+v\| \leq \|u\|+\|v\|$

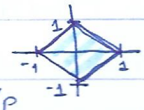
☆ **וקטור יחידה:** $v\in\mathbb{R}^n$ כך ש: $\|v\|=1$. (נורמה היא מדד לגודל הוקטור)

• **כדור יחידה:** $B=\{v \mid \|v\|\leq 1\}$

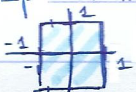
דוג': • **הנורמה האוקלידית:** $L_2 = \|x\|_2 = \sqrt{\sum (x_i)^2} = (\sum |x_i|^2)^{1/2}$

אזי: $B_{\|\cdot\|_2} =$  (מעגל היחידה ב-$\mathbb{R}^2$)

• $L_1 = \|x\|_1 = \sum |x_i| = (\sum |x_i|^1)^{1/1}$

אזי: $B_{\|\cdot\|_1} =$ 

טענה: לכל $p\in\mathbb{N}$: $L_p = \|\cdot\|_p = (\sum_{i=1}^{m} |x_i|^p)^{1/p}$ היא נורמה.

• $L_\infty = \|x\|_\infty = \max_{i\in[m]} |x_i|$, אזי: $B_{\|\cdot\|_\infty} =$ 

• **מרחב מכפלה פנימית:** מרחב מכפלה פנימית הוא הזוג: $\langle V, \langle\cdot,\cdot\rangle\rangle$

כאשר V מ"ו מעל $\mathbb{R}$ ו- $\langle\cdot,\cdot\rangle: V\times V\to\mathbb{R}$ פונק' המקיימת:

לכל $x,y,z\in V$, $\alpha\in\mathbb{R}$:

1. סימטריה: $\langle x,y\rangle = \langle y,x\rangle$

2. לינאריות: $\langle \alpha x+z, y\rangle = \alpha\langle x,y\rangle + \langle z,y\rangle$

3. אי"ש: $\langle x,x\rangle \geq 0$, ו- $\langle x,x\rangle = 0 \Leftrightarrow x=0$

• הפונק' $\langle\cdot,\cdot\rangle$ נקראת **מכפלה פנימית**.

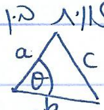
(induced norm)

• **נורמה מוגדרת ממ"פ:** ממ"פ מגדיר נורמה: $\|x\| = \langle x,x\rangle$.

מכפלה פנימית סטנדרטית: $\langle x,x\rangle = \sum x_i^2 = (\|x\|_2)^2$

טענה: יהיו $x,y\in\mathbb{R}^n$. מתקיים: $\langle x,y\rangle = \|x\|\cdot\|y\|\cdot\cos(\theta)$

עבור θ הזווית בין x,y.

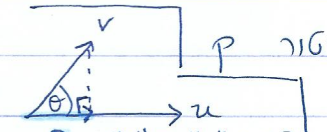
הוכחה: נסמן: . מתקיימת (הנכונה:) $c^2 = a^2+b^2-2ab\cos(\theta)$

עבור: $c=\|x-y\|$, $b=\|y\|$, $a=\|x\|$:

(i) $\|x-y\|^2 = \|x\|^2+\|y\|^2-2\|x\|\|y\|\cos(\theta)$

ה-N induced norm: (ii) $\|x-y\|^2 = \langle x-y, x-y\rangle = \|x\|^2+\|y\|^2-2\langle x,y\rangle$

ולכן מצירוף (i),(ii) מתקיים: $\|x\|\|y\|\cos(\theta) = \langle x,y\rangle$, כנדרש. ■



p , $\|p\| = \|v\|\cos\theta$

$P = \frac{v\cdot u}{|u|^2}u = |v|\cos\theta\frac{u}{|u|}$

• **הטלת וקטור**: הטלת וקטור v על וקטור u היא וקטור p שאורכו: $\|v\|\cos(\theta)$ בכיוון u.

★ נבחין ש: $\langle u,v\rangle=0 \Leftrightarrow \theta=90^\circ \Leftrightarrow u\perp v$: במקרה כזה נאמר ש- u,v אורתוגונליים.

• **מטריצה אורתוגונלית**: מטריצה ריבועית היא אורתוגונלית ⇔ עמודותיה הן וקטורי יחידה שאורתוגונלים זה לזה (כלו' אורתונורמליים), וכן שורותיה הן וקטורים אורתונורמליים.

למה: תהא $A\in M_{d\times d}(\mathbb{R})$ מטריצה אורתוגונלית. מתקיים: $AA^T = A^TA = I$

2. פירוקי EVD, SVD

• **מטריצה אלכסונית**: מטריצה A היא אלכסונית אם לכל $i\neq j$: $A_{ij}=0$.

• **מטריצה לכסינה**: תהא A מטריצה ריבועית. A לכסינה אם קיימת מטריצה הפיכה P כך שהמטריצה $P^{-1}AP$ היא אלכסונית.

• **וקטור עצמי וערך עצמי**: תהא A מטריצה ריבועית. וקטור $0\neq v\in V$ הוא הוקטור העצמי של A המתאים לערך העצמי $\lambda\in\mathbb{R}$ אם: $Av=\lambda v$.

★ כלו' ש-A לא משנה את הזווית של v במרחב, אלא רק מכווצת/מנפחת אותו, בהתאם לגודל λ, ו/או משנה כיוון.

משפט EVD: תהא $A\in M_{n\times n}$ מטריצה סימטרית בממשיים. אזי קיימת מטריצה אורתוגונלית $U\in M_{n\times n}$ ומטריצה אלכסונית D כך שאלכסון המטריצה D מכיל את הערכים העצמיים של A ומתקיים: $A = UDU^T$.

דוגמא: נרצה לחשב מהר יותר את $(A^{10})^6$. פתרון נאיבי ייקח $O(n^3)$. אם A ממשית סימטרית וריבועית, מהמשפט: $(A^{10})^6 = (UDU^T)\cdots(UDU^T)$ ומצמצום U^TU נקבל: $UD^{10^6}U^T$; ונבחין ש: $(D^{10^6})_{i,i} = (d_{i,i})^{10^6}$ ⇐ קל.

וקטורים סינגולריים (ימני ושמאלי): יהיו $A\in\mathbb{R}^{n\times m}$, $v\in\mathbb{R}^m$, $u\in\mathbb{R}^n$.
v הוא הוקטור הסינגולרי הימני של A, ו-u הוא הוקטור הסינגולרי השמאלי של A אם קיים ערך סינגולרי $\sigma\in\mathbb{R}_{\geq 0}$ מתקיים: $Av=\sigma u$.
המטריצה A נקראת סינגולרית.

משפט SVD: כל מטריצה $A\in M_{m\times d}(\mathbb{R})$ ניתנת להצגה כמכפלה של 3 מטריצות, כך: $A=U\Sigma V^T$, כאשר: $U\in M_{m\times m}(\mathbb{R})$, $V\in M_{d\times d}(\mathbb{R})$ מטריצות אורתוגונליות; $\Sigma\in M_{m\times d}(\mathbb{R})$ מטריצה אלכסונית עם ערכים אי״ש. ערכים אלו נקראים הערכים הסינגולריים של A.

טענה: אם $A=U\Sigma V^T$ פרוק SVD של המטריצה A, אז: מתקיים שעמודות U ושורות V^T הם וקטורים סינגולריים ימניים ושמאליים של A, המתאימים לערכים הסינגולריים שנמצאים על האלכסון של Σ.

הוכחה: $A=U\Sigma V^T$, ונניח: $rank(A)=r\in\mathbb{R}$. אז:

$$A=\begin{pmatrix} | & & | \\ u_1 & \cdots & u_m \\ | & & | \end{pmatrix}\begin{pmatrix} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_r & \\ 0 & & & 0 \end{pmatrix}\begin{pmatrix} \vdash v_1^T \dashv \\ \vdots \\ \vdash v_d^T \dashv \end{pmatrix}$$

נבחין שכל זה: אנחנו מכפילים באפסים. לכן למעשה נשאר עם:

$$=\begin{pmatrix} | & & | \\ u_1 & \cdots & u_r \\ | & & | \end{pmatrix}\begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{pmatrix}\begin{pmatrix} \vdash v_1^T \dashv \\ \vdots \\ \vdash v_r^T \dashv \end{pmatrix} \Rightarrow^{\text{לכן:}} Av_i=u_i\sigma_i$$

$\blacksquare$

★ SVD, EVD עוזרים לפתור את בעיית הרגרסיה הלינארית. נראה בשבוע הבא.

3. הטלות אורתוגונליות

בהינתן תת-מרחב וקטורי V ממימד k ב-$\mathbb{R}^d$ ($k\leq d$), נגדיר: $P=\sum_{i=1}^{k} v_i v_i^T$ (כאשר: $\{v_1,\dots,v_k\}\in\mathbb{R}^d$ בסיס אורתונורמלי של V). אז P הוא מטריצת ההטלה האורתוגונלית על תת-המרחב V.

★ נבחין ש: $v_iv_i^T=v_i\otimes v_i^T$ ולכן למעשה מטריצה מדרגה 1 $\left(\begin{bmatrix}\vdash v\dashv\\ \vdots\\ \vdash v\dashv\end{bmatrix}\right)$,
אז' P הוא סכום k מטריצות מדרגה 1. א״ש - P לוקחת וקטור ב-$\mathbb{R}^d$ ומטילה אותו על תת-מרחב ממימד k.

למה: יהיו $\{v_1,..,v_k\}$ אוסף של וקטורים אורתונורמליים, ותהא $P=\sum_{i=1}^{k} v_i v_i^T$. אזי ל-$P$ יש את התכונות הבאות:

1. P סימטרית.
2. $P^2=P$.
3. הע"ע של P הם 0 או 1 ($\{v_i\}_{i=1}^{k}$ הו"ע המתאימים)
4. $(I-P)P=0$
5. לכל $x\in\mathbb{R}^d$, $u\in V$: $|x-u|\geq|x-Px|$
6. $x\in V \Rightarrow Px=x$

4. הנורמה הספקטרלית

העתקה איזומטרית: תהא $T:\mathbb{R}^n\to\mathbb{R}^m$ העתקה לינארית.
T איזומטריה $\Leftrightarrow$ $\forall u\in\mathbb{R}^n: \|T(v)\|=\|v\|$ (משמרת נורמה).

טענה: תהא A מטריצה. A אורתוגונלית $\Leftrightarrow$ הה"ל שמוגדרת אותה איזומטרית.
כלומר לכל x: $\|Ux\|=\|x\|$

נורמת האופרטור/הנורמה הספקטרלית: תהא A מטריצה ממשית.
הנורמה הספקטרלית של A תהא: $\|A\|_{op}=\max_{0\neq x\in\mathbb{R}^n}\frac{\|Ax\|_2}{\|x\|_2}=\max_{\|x\|=1}\|Ax\|_2$
(מהו הגודל ש-A מגדילה הכי שאפשר - כדי להגדיר גודל של מטריצות)

טענה: תהא $A=U\Sigma V^T$ וע"ס $\sigma_1\geq..\geq\sigma_k$. מתקיים: $\|A\|_{op}=\sigma_1$.

אינטואיציה ל-$\|\cdot\|_{op}$:
$\Leftarrow$ U,V אורתוגונליות, ולכן מטענה קודמת הן איזומטריות. אז אם נכפול את A ב וקטור x כלשהו, למעשה נסובב את x (בלי לשנות את גודלו כי V^T אורתוגונלית), ונמתח אותו בכ"א מהקואורדינטות שלו בהתאם לערכים הסינגולריים ב-Σ, ונסובב אותו בכיוון כלשהו חזרה.

☆ איור מתרגול.