אלגברה ליניארית 2

סיכם: נריה אור

ע"פ הרצאות של ד"ר דן רומיק וצוות המתרגלים. אין המרצים קשורים לסיכום זה בשום אופן.

גירסא 1.0

סוכמו כל ההרצאות (ללא תרגולים כמעט).

חסרים כל מיני דברים שנלמדו רק בתרגול (שיקופים וכו').

ותודה רבה לכל מי ששלח תיקונים והערות במייל!

אין אחריות לתוכן הסיכום ולדיוקו - ייתכנו טעויות. השימוש על אחריות הקורא בלבד.

הקטעים שהוכרזו כ"לא למבחן", וכתוב בסיכום שהם כאלו, הם פרשנות אישית שלי של ההודעה הרשמית, ולכן ייתכן שטעיתי בהבנה של אילו טענות לא כלולות בחומר למבחן. אין אחריות לכך שאולי חסר פה חומר, בכל אופן.

תוכן עניינים

1 חזרה והרחבה של מושגים (+דברים מהתרגול) 6

1.0.1 חוגים 6

1.0.2 דטרמיננטות 6

1.0.3 ייצוג של העתקות ליניאריות ע"י מטריצות 7

1.0.4 המטריצה המצורפת (adj) 7

1.0.5 נוסחת קרמר 9

1.0.6 היטל ומרחק של וקטור מתת־מרחב 9

1.0.7 הטלות 9

2 ו"ע ע"ע ועוד ראשי תיבות מבלבלים 9

2.1 מטריצות אלכסוניות 10

2.1.1 וקטור עצמי \ ערך עצמי 11

2.1.2 משפטים בסיסיים על ו"ע וע"ע 11

2.1.3 תזכורת קצרה לגבי החלפת בסיסים 13

2.2 דמיון מטריצות 14

2.2.1 דמיון וערכים עצמיים ־ טענה 14

2.2.2 הקשר בין דמיון לליכסון 15

2.3 דוגמאות (כולל פיבונצ'י) 15

3 פולינום אופייני, פולינום מינימלי 18

3.0.1 פולינומים ־ הגדרה ותכונות 18

3.0.2 חלוקת פולינומים עם שארית 19

3.1 מטריצות שאיבריהן הם פולינומים 20

3.2 הפולינום האופייני 21

3.2.1 הגדרה, והיות ע"ע שורשים של f_A 21

3.2.2 תכונות הפולינום האופייני ־ $-Trace(A),\ (-1)^n det(A)$ והיותו מתוקן 22

3.2.3 מסקנה לגבי מטריצות דומות A, B: $f_A(t) = f_B(t)$ 24

3.3 הפולינום המינימלי 24

3.3.1 קצת טענות מקדימות 24

3.3.2 הפולינום המינימלי 26

3.3.3 אם $A, B \in M_n(F)$ דומות אז $m_A = m_B$ 26

4 מרחבים עצמיים, ריבוי אלגברי וריבוי גיאומטרי 26

4.1 מרחב עצמי, ריבוי גיאומטרי 26

4.1.1 טענות ומשפטים 27

4.2 ריבוי אלגברי 28

4.2.1 הגדרה 28

4.2.2 טענות ומשפטים 28

5 עוד על פולינומים ומטריצות 29
5.1 הצבות של פולינומים במטריצות של פולינומים שאיבריהם מטריצות ש....?! 29
5.2 קצת הסברים, טענות לגבי $M_n(\mathbb{F})[t]$ 29
5.2.1 טענה חשובה, שלפיה אם $h(t) = (tI - A)g(t)$ אז $h(A) = 0$ 30
5.3 הקשר בין $M_n(\mathbb{F})[t]$ ו־ $M_n(\mathbb{F}[t])$ 31
6 משפט קיילי ־ המילטון 32
6.0.1 מסקנות מהמשפט 33
7 סיכום ליכסון מטריצות 37
8 מרחבי מכפלה פנימית 37
8.1 הגדרות 37
8.1.1 תכונות 38
8.2 קושי שוורץ וחברים. 39
8.2.1 אי שויון קושי־שוורץ 39
8.2.2 מסקנה מקושי שוורץ: אי־שויון המשולש 40
8.3 אורתוגונליות, קבוצות אורתוגונליות ואורתונורמליות 41
8.3.1 טענה לגבי המקדמים של וקטור שהוא צ"ל של קבוצה אורתונורמלית, ומסקנות 41
8.3.2 משפט פיתגורס 42
8.3.3 תהליך האורתוגונליזציה של גראם־שמידט 42
8.3.4 דוגמא לשימוש בגראם־שמידט 43
8.3.5 משלים אורתוגונלי וסכומים ישרים 44
8.4 פונקציונלים ליניאריים 46
9 העתקות ליניאריות וההעתקה הצמודה 47
9.1 ההעתקה הצמודה 47
9.1.1 הגדרה והוכחת הקיום\יחידות 47
9.1.2 תכונות ההעתקה הצמודה 49
9.1.3 מטריצה צמודה 50
9.2 העתקות צמודות לעצמן: הרמיטיות ($\mathbb{C}$), סימטריות ($\mathbb{R}$) 50
9.2.1 הגדרה 50
9.2.2 תכונות 51
9.2.3 טענה, שאומרת שמעל $\mathbb{C}$, T הרמיטית $\Leftrightarrow$ $\langle Tv, v\rangle \in \mathbb{R}$ $\forall v \in V$ 52
9.2.4 למה: תנאים שקולים להיות העתקה זהותית 0, מעל $\mathbb{C}$ 52
9.2.5 קצת על העתקות אנטי־הרמיטיות, וטענה על יחידות הפיתוח $T = R + iS$ (R, S הרמיטיות). 53
9.2.6 למה: תנאים שקולים להיות העתקה *סימטרית* זהותית 0, מעל $\mathbb{R}$ 53
9.3 העתקות אוניטאריות: אוניטאריות ($\mathbb{C}$), אורתוגונליות ($\mathbb{R}$) 54
9.3.1 הגדרה 54
9.3.2 תנאים שקולים להיות T אוניטארית 54

9.3.3 קשרים בין העתקה אוניטארית למטריצה שלה 56
9.3.4 תכונות של מטריצה אוניטארית - $|det(A)| = 1$, ומשפט על ע"ע, ו"ע. . 57
9.4 העתקות נורמליות: $T^*T = TT^*$ 58
9.4.1 הגדרה 58
9.4.2 טענות על ו"ע וע"ע של העתקה נורמלית. ($Tv = \lambda v \Leftarrow T^*v = \overline{\lambda}v$, וגם היות ו"ע של ע"ע שונים מאונכים זה לזה) 59
9.4.3 הקשר בין ת"מ T-אינוריאנטי, והמשלים האורתוגונלי של הת"מ ו-T . 60
9.4.4 לכסון אורתונורמלי של מטריצות נורמליות (מעל $\mathbb{C}$) 60
9.4.5 פירוק קוטבי: הצגה של מטריצה נורמלית בתור $A = HU = UH$ (U אוניטרית, H הרמיטית) 61
9.4.6 פירוק לחלק ממשי ומדומה 63
9.4.7 שני תנאי אם"ם לקביעה אם מט' נורמלית היא אוניטרית, הרמיטית . 63
9.5 מטריצות נורמליות מעל $\mathbb{R}$ 64
9.5.1 ליכסון אורתוגונלי (מעל $\mathbb{R}$) 64
9.5.2 מירכוב של מרחב אוקלידי 65
9.5.3 משפט המבנה להעתקות נורמליות מעל $\mathbb{R}$ 68
10 צורת ז'ורדן 68
10.1 הקדמה - סכומים ישרים 69
10.1.1 שדה סגור אלגברית 69
10.1.2 סכומים ישרים 69
10.2 מרחבים אינווריאנטיים, ציקליים 70
10.2.1 מטריצות בלוקים ומרחבים אינווריאנטים 70
10.2.2 תת-מרחבים ציקליים 71
10.2.3 מטריצת ייצוג של צמצום העתקה לבסיס של ת"מ T-ציקלי 71
10.3 העתקות נילפוטנטיות 71
10.3.1 הגדרה 71
10.3.2 תכונות וטענות 72
10.4 צורת ז'ורדן של העתקה נילפוטנטית 73
10.4.1 בלוק ז'ורדן 73
10.4.2 משפט ההשלמה לסכום ישר עבור ת"מ ציקלי של העתקה נילפוטנטית $V = U \oplus W$ 73
10.4.3 צורת ז'ורדן של העתקה נילפוטנטית 74
10.4.4 הוכחת היחידות (כולל הנוסחה למספר הבלוקים בגודל j) 74
10.5 צורת ז'ורדן של העתקה כללית 76
10.5.1 קיום 76
10.5.2 סיכום מה שראינו 77
10.5.3 הוכחת היחידות 78
10.5.4 סיכום צורת ז'ורדן 78

10.6 שימושים ומסקנות מצורת ז'ורדן 79
10.6.1 הוכחה למשפט קיילי־המילטון 79
10.6.2 עוד מסקנות 80
10.7 אלגוריתמים (מהתרגול) 80
10.7.1 אלגוריתם למציאת בסיס מז'רדן למטריצה נילפוטנטית 80
10.7.2 אלגוריתם למציאת צורת ז'ורדן של מטריצה כללית 81
11 תבניות ביליניאריות 82
11.1 הקדמה 82
11.1.1 הגדרות ובסיס 82
11.1.2 עוד על ייצוג בעזרת מטריצות 83
11.1.3 מטריצות חופפות 85
11.1.4 מטריצות שקולות 85
11.2 ליכסון ותבניות ביליניאריות סימטריות 86
11.2.1 תבניות ביליניאריות סימטריות 86
11.2.2 ליכסון ותנאים לליכסון 86
11.2.3 דוגמא 87
11.2.4 הוכחה נוספת למשפט הליכסון 88
11.2.5 מעל $\mathbb{C}$ - לכל ת. ביליניארית סימטרית יש צורה אלכסונית מצומצמת יותר 89
11.2.6 משפט ההתמדה של סילבסטר 89
11.2.7 אלגוריתם ללכסון תבניות ריבועיות ע"י חפיפה (מהתרגול) 91
11.3 תבניות ריבועיות 92
11.3.1 הגדרה ודברים בסיסיים 92
11.3.2 תבניות ריבועיות מעל $\mathbb{R}$ - חיוביות 93
11.3.3 כיצד לבדוק אם A חיובית בלי לחשב את הע"ע? 94

1 חזרה והרחבה של מושגים (+דברים מהתרגול)

1.0.1 חוגים

הגדרה 1.1 חוג ($ring$**)**

חוג הוא קבוצה R ביחד עם פעולות חיבור וכפל המקיימות:

חיבור:

1. קיום **איבר ניטרלי** שייקרא 0.
2. לכל x קיים **הופכי** $(-x)$ כך ש־ $x+(-x)=0$.
3. **אסוציאטיביות:** $(x+y)+z=x+(y+z)$
4. **קומוטטיביות:** $x+y=y+x$

כפל:

1. **אסוציאטיביות:** $x\cdot(y\cdot z)=(x\cdot y)\cdot z$
2. **דיסטריביוטיביות:** $a(x+y)=ax+ay$ **וגם** $(x+y)a=xa+ya$.

1.0.2 דטרמיננטות

נתחיל ברענון קצר מאלגברה ליניארית 1:

סימונים:

$\mathbb{F}$ שדה, V,W מרחבים וקטוריים,

$M_{n\times m}(\mathbb{F})$ אוסף המטריצות מסדר $n\times m$ מעל $\mathbb{F}$,

$M_n(\mathbb{F})=M_{n\times n}(\mathbb{F})$

כמו כן על מנת לחסוך לי המון הקלדה,

מעכשיו ועד קיץ העולם, נסמן: $\mathbb{F}=F$. (אלא אם כן יהיה נחוץ אחרת).

הגדרה 1.2 דטרמיננטה

אם $A\in M_n(F)$, **הדטרמיננטה** שלה $det(A)$ היא:

(כאשר $A=(a_{ij})_{i,j=1}^n$)

$$det(A)=\sum_{\sigma\in S_n} sgn(\sigma)\prod_{i=1}^{n} a_{i\sigma(i)}$$

כאשר S_n אוסף התמורות מסדר n, וגם:

$$sgn(\sigma)=\begin{cases}1 & \text{if } \sigma \text{is ZUGIT}\\ -1 & \text{if } \sigma \text{is E-ZUGIT}\end{cases}$$

והזוגיות של תמורה היא מספר החילופים שנדרשים בשביל להציג אותה.

לדוגמא:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 5 & 4 & 3 \end{pmatrix} = (3\ 5)(1\ 2)$$

הגדרה שקולה: פיתוח לפי שורה\עמודה.

$$det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \cdot det(A_i^j)$$

כאשר A_i^j המטריצה המינורית A כשמוחקים ממנה את שורה i ועמודה j.

משפט 1.3 A הפיכה אם"ם $det(A) \neq 0$.

1.0.3 ייצוג של העתקות ליניאריות ע"י מטריצות

V, W מ"ו. $T : V \to W$ ה"ל.

רוצים לתת תיאור מפורש של T.

לוקחים בסיס (סדור) $B = (v_1, ..., v_n)$ של V.

ולוקחים בסיס $B' = (w_1, ..., w_m)$ של W.

לכל $1 \leq j \leq n$ נתבונן בפיתוח $T(v_j) = \sum_{i=1}^{m} a_{ij} w_i$.

מטריצת המקדמים $\underset{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}{(a_{ij})}$ תיקרא

מטריצת הייצוג של T בבסיסים B', B , ותסומן: $[T]_{B'}^{B}$.

במקרה שבו $V = W$, בד"כ ניקח $B = B'$ ונסמן $[T]_B = [T]_B^B$.

(הסבר מקיף וברור היה בליניארית 1, אפשר להסתכל שם אם זה לא מספיק).

1.0.4 המטריצה המצורפת (adj)

הגדרה 1.4 המטריצה המצורפת $adj(A)$:

אם $A \in M_n(R)$, כאשר R חוג קומוטטיבי (לכפל), אז לכל $1 \leq i, j \leq n$

נסמן ב־$M_{i,j}$ את **המינור ה־**(i, j) **של** A, שהוא הדטרמיננטה של המטריצה המינורית,

שהיא המטר' A כאשר מוחקים ממנה את השורה ה־i והעמודה ה־j.

ונסמן: $A_{i,j} = (-1)^{i+j} M_{i,j}$ ונסמן: $A = (a_{i,j})_{i,j=1}^{n}$.

אז מתקיים: (פיתוח לפי שורה i):

$$\sum_{j=1}^{n} a_{i,j} A_{i,j} = det(A)$$

כמו כן, לכל $k \neq i$:

$$\sum_{j=1}^{n} a_{i,j} A_{k,j} = 0$$

(זאת הדטרמיננטה של המטריצה A' שמתקבלת מ־A ע"י החלפת השורה ה־k בשורה ה־i, כלומר כך שיש פעמיים את השורה ה־i).

הבהרה לגבי זה:

נגדיר את A' הנ"ל:

$$A' = \begin{pmatrix} -\alpha_1- \\ -\alpha_i- \\ -\alpha_i- \\ -\alpha_n- \end{pmatrix}$$

מתכונות הדטרמיננטה, נובע ש־$|A'| = 0$. כעת, נפתח לפי השורה ה־k:

$$0 = |A'| = \sum_{j=1}^{n} a_{k,j} A_{k,j} = \sum_{j=1}^{n} a_{i,j} A_{k,j}$$

והשויון האחרון נכון בגלל שהגדרנו ששתי השורות i ו־k שוות ב־A'.

ולכן אם נסמן: $adj(A) = (A_{j,i})_{i,j=1}^{n}$ אז המשוואות למעלה אומרות ש־

$$A \cdot adj(A) = \begin{pmatrix} det(A) & & \cdots & 0 \\ & det(A) & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & det(A) \end{pmatrix} = det(A)I$$

או בקיצור, לסיכום, הגדרנו:

$$(adj(A))_{ij} = (-1)^{i+j} \cdot M_{ji}$$

ומתקיים:

$$A \cdot adj(A) = det(A) \cdot I$$

ומזה אפשר להסיק שאם A הפיכה, אז $A \cdot (\frac{1}{det(A)} \cdot adj(A)) = I$ ולכן $A^{-1} = \frac{1}{det(A)} \cdot adj(A)$.

1.0.5 נוסחת קרמר

תהא $A \in M_n(\mathbb{F})$, עם עמודות $A = \begin{pmatrix} | & & | \\ c_1 & \cdots & c_n \\ | & & | \end{pmatrix}$.

ותהא המערכת הלא הומוגנית: $Ax = b$.

כידוע, למערכת לא הומוגנית יש פתרון יחיד אם"ם A הפיכה. (כלומר $det(A) \neq 0$).

נגדיר: $A_k = \begin{pmatrix} | & | & | \\ c_1 & b & c_n \\ | & | & | \end{pmatrix}$ (כלומר A, כאשר שמנו במקום העמודה ה־k את b).

אז כלל קרמר אומר את הדבר הבא: אם נסמן $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, אז לכל $1 \leq i \leq n$, מתקיים

$$x_i = \frac{det(A_k)}{det(A)}$$

(תודה לויקיפדיה על ההסבר הברור).

1.0.6 היטל ומרחק של וקטור מתת־מרחב

(אין לי הגדרה רשמית, זה מהראש שלי\הסברים מילוליים, אז אתנצל מראש אם יש פה טעות)

אם נרצה לחשב את ההיטל של וקטור $v \in U$ כאשר U תת־מרחב,

ניקח בסיס אורתונורמלי $\{u_1, ..., u_n\}$ של U, ונחשב:

$$proj_U(v) = \sum_{i=1}^{n} \langle v, u_i \rangle \cdot u_i$$

זהו ההיטל. (נשים לב - מחשבים היטל גם בתהליך גראם־שמידט!)

כעת, אם נרצה לחשב את **המרחק** של וקטור v מתת־המרחב, נחשב את $\|v - proj_U(v)\|$, וזהו המרחק המבוקש.

1.0.7 הטלות

הגדרה 1.5 טרנספורמציה ליניארית $P : V \to V$ תיקרא **הטלה** אם $P = P^2$.

(כאשר הסימון הוא $P^2(v) = P(P(v))$ כמובן).

טענה 1.6 (הוכחנו בתרגיל 7): לכל הטלה $P : V \to V$, מתקיים $V = Ker\, P \oplus Im\, P$.

2 ו"ע ע"ע ועוד ראשי תיבות מבלבלים

המטרה שלנו, היא למצוא בסיסים "טובים" לייצוג העתקות ליניאריות.

בסיס "טוב" - בסיס שבו מטריצת הייצוג תהיה "פשוטה".

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} \in M_n(\mathbb{F})$$

הגדרה 2.1 ה"ל $T: V \to V$ נקראת **ניתנת לליכסון** (או **לכסינה**),
אם קיים בסיס B של V כך ש $[T]_B$ היא אלכסונית.

מה זה אומר?

אם נסמן $B = \{v_1, ..., v_n\}$ אז

$$T(v_i) = \lambda_i v_i \qquad \Leftrightarrow \qquad [T]_B = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

כלומר המשמעות הגיאומטרית היא שיש מערכת צירים כך ש־ T מנפחת\מכווצת\משקפת את הציר ה־i בפקטור קבוע λ_i. (????)

הגדרה 2.2 מטריצה $A \in M_n(F)$ נקראת **ניתנת לליכסון** אם היא ניתנת לליכסון כהעתקה ליניארית הפועלת על F^n, על ידי כפל מטריצות.

תזכורת: כפל מטריצות
אם $A \in M_{m\times n}(F)$ ו־ $B \in M_{n,p}(F)$, אז $C = AB \in M_{m\times p}(F)$ המטריצה המוגדרת ע"י:

$$(C)_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$$

בפרט, כל מטריצה $A \in M_{m\times n}(F)$ פועלת כהעתקה ליניארית $A: F^n \to F^m$ ע"י כפל,
כאשר חושבים על איברי F^n כוקטורי עמודות.

$$A(v) = Av$$

מסקנה:

מטריצה $A \in M_n(F)$ **ניתנת לליכסון אם"ם**

קיים בסיס $B = \{v_1, ..., v_n\}$ של F^n וסקלרים $\lambda_1, ..., \lambda_n \in F$ כך ש־ $Av_i = \lambda_i v_i$

2.1.1 וקטור עצמי \ ערך עצמי

הגדרה 2.3 וקטור עצמי \ ערך עצמי - ו"ע \ ע"ע

אם $T:V\to V$ ה"ל, אז וקטור $v\in V$ ייקרא **וקטור עצמי** (ו"ע) של T,

אם $v\neq 0$ וקיים סקלר $\lambda\in F$ כך ש- $T(v)=\lambda v$.

הסקלר λ ייקרא **הערך העצמי** (ע"ע) המתאים ל-v.

הגדרה עבור מטריצות:

$A\in M_n(F)$. אומרים ש $v\in F^n$ הוא **וקטור עצמי** של A עם **ערך עצמי** $\lambda\in F$,

אם $v\neq 0$ ומתקיים $Av=\lambda v$.

2.1.2 משפטים בסיסיים על ו"ע וע"ע

טענה 2.4 אם v הוא ו"ע של A אז לכל $0\neq\alpha\in F$, גם αv הוא ו"ע עם אותו ע"ע λ.

הוכחה: אם $Av=\lambda v$ אז:

$$A(\alpha v)=\alpha(Av)=\alpha(\lambda v)=\lambda(\alpha v)$$

■

טענה 2.5 $\lambda\in F$ הוא ע"ע של A **אם"ם** $det(\lambda I-A)=0$.

הוכחה: λ הוא ע"ע $\Leftrightarrow$קיים וקטור $0\neq v\in F^n$ כך ש-$Av=\lambda v=\lambda Iv\Leftrightarrow$

קיים פתרון $v\neq 0$ למשוואה $(\lambda I-A)v=0\Leftrightarrow ker(\lambda I-A)\neq\{0\}\Leftrightarrow det(\lambda I-A)=0$.

(נשים לב ש-$\lambda v=\lambda Iv$ מכיוון שכפל של I ב-v נותן וקטור עמודה ב-F^n ולכן זה לא משנה). ■

טענה 2.6 אם V מ"ו, $T:V\to V$ ה"ל, ו- $v_1,...,v_k\in V$ ו"ע של T עם ע"ע $\lambda_1,...,\lambda_k$ בהתאמה,

ונניח ש-$\lambda_i\neq\lambda_j$ לכל $i\neq j$.

אז $v_1,...,v_k$ הם בת"ל.

הוכחה: באינדוקציה על k.

עבור $k=1$: זה ברור: ו"ע הוא שונה מ-0 ולכן בת"ל.

נראה את $k=2$, כדי לקבל תחושה לגבי ההוכחה:

נניח ש-

$$av_1+bv_2=0$$

אז נפעיל את T:

$$0=T(0)=T(av_1+bv_2)=aT(v_1)+bT(v_2)=a\lambda_1v_1+b\lambda_2v_2$$

נחסיר מהמשוואה השניה λ_1 כפול המשוואה הראשונה, ונקבל:

$$b(\lambda_2-\lambda_1)v_2=0$$

ומכיון ש־ $\lambda_1 \neq \lambda_2$**, וגם** $v_2 \neq 0$**, נקבל ש־** $b=0$**. ולכן מהמשוואה הראשונה גם** $a=0$.
ולכן הם בת"ל.

שלב האינדוקציה: נניח והוכחנו ל־$k-1$, ונוכיח ל־k:
אם

$$\alpha_1 v_1+\alpha_2 v_2+...+\alpha_k v_k=0$$

אז נפעיל את T על שני האגפים:

$$T(\alpha_1 v_1+...+\alpha_k v_k)=\sum_{i=1}^{k} T(\alpha_i v_i)=\sum_{i=1}^{k}\alpha_i T(v_i)=\sum_{i=1}^{k}\alpha_i\lambda_i v_i=T(0)=0$$

ונחסיר ממשוואה זו λ_k כפול המשוואה הראשונה, ונקבל:

$$(\lambda_1-\lambda_k)\alpha_1 v_1+(\lambda_2-\lambda_k)\alpha_2 v_2+...+(\lambda_k-\lambda_k)\alpha_k v_k=0$$

$$(\lambda_1-\lambda_k)\alpha_1 v_1+(\lambda_2-\lambda_k)\alpha_2 v_2+...+0\cdot\alpha_k v_k=0$$

ולכן לפי הנחת האינדוקציה, מקבלים $(\lambda_i-\lambda_k)\alpha_i=0$ לכל $1\leq i\leq k-1$**, אבל** $\lambda_i \neq \lambda_k$ **לכל** i **כזה,**
ולכן $\alpha_i=0$ לכל $1\leq i\leq k-1$. ומהמשוואה הראשונה, זה נותן גם: $\alpha_k=0$. ■

מסקנה:

טענה 2.7 אם $T:V\to V$ ה"ל, ו־ $dimV=n$, אז:

- ל־T יש לכל היותר n ע"ע שונים.
- אם ל־T יש n ע"ע שונים, אז יש בסיס ל־V המורכב מו"ע.

הוכחה: א. אם ל־T יש k ע"ע שונים, אז יש k ו"ע מתאימים.
לפי הטענה הקודמת, הם בת"ל ולכן $k\leq n$.

ב. אם יש n ע"ע שונים $\lambda_1,...,\lambda_n$, אז לכל $1\leq i\leq n$ יהי $v_i\in V$ ו"ע המתאים ל־λ_i.
ה־v_i **הם שונים**, כי לא ייתכן ש־v יהיה ו"ע עבור שני ע"ע שונים,
ולכן לפי הטענה הקודמת, הם בת"ל ולכן בסיס. ■

2.1.3 תזכורת קצרה לגבי החלפת בסיסים

יהי $B=\{v_1,...,v_n\}$ בסיס למ"ו V.

הגדרה 2.8 וקטור הקואורדינטות של וקטור $v\in V$ לפי הבסיס B מוגדר להיות :

$[v]_B=\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$, כאשר $v=\sum_{i=1}^n a_i v_i$.

כמו כן,

הגדרה 2.9 מטריצת הייצוג של ה"ל $T:V\to V'$,
לפי בסיסים $B=\{v_1,...,v_n\}$ של V, B' של V' מוגדרת ע"י:

$$[T]_{B'}^{B}=\begin{pmatrix} | & | & \\ [T(v_1)]_{B'} & [T(v_2)]_{B'} & \cdots \\ | & | & \end{pmatrix}$$

טענה 2.10 מה קורה אם מחליפים בסיסים?
אם C,C' בסיסים אחרים של V,V' בהתאמה, אז מתקיים:

$$[T]_{C'}^{C}=M_{C'}^{B'}[T]_{B'}^{B}M_{B}^{C}$$

כאשר $M_B^C=[Id]_B^C$ מטריצת המעבר מבסיס C ל־B. ונזכור ש־ $M_{B'}^{C'}=(M_{C'}^{B'})^{-1}$.

הוכחה: נכפול את שני האגפים משמאל ב־ $M_{B'}^{C'}$ ואז צריך להוכיח ש־

$$M_{B'}^{C'}[T]_{C'}^{C}=[T]_{B'}^{B}M_{B}^{C}$$

נראה ששתי המטריצות זהות בכל עמודה: (כאשר מתייחסים לתוצאת המכפלה כמובן)
אם $1\le i\le n$ אז: (כאשר מסמנים: $C=\{u_1,...,u_n\}$)

$$M_{B'}^{C'}[T]_{C'}^{C}e_i=M_{B'}^{C'}([T]_{C'}^{C}e_i)=M_{B'}^{C'}[T(u_i)]_{C'}=[T(u_i)]_{B'}$$

ומצד שני,

$$[T]_{B'}^{B}M_{B}^{C}e_i=[T]_{B'}^{B}(M_{B}^{C}e_i)=[T]_{B'}^{B}[u_i]_B=[T(u_i)]_{B'}$$

ולסיכום הראינו שיויון ואת מה שצריך להוכיח. ■

במקרה הפרטי ש־$V = V'$, $B = B'$, $C = C'$, מקבלים את הנוסחא:

$$[T]_C = M_C^B [T]_B M_B^C = P[T]_B P^{-1}$$

הגדרה 2.11 מטריצות דומות

אומרים ששתי מטריצות $A, B \in M_n(F)$ הן **דומות,**

אם קיימת מטריצה הפיכה $P \in M_n(F)$ כך ש $A = PBP^{-1}$.

יחס הדמיון הוא יחס שקילות. כלומר:

- רפלקסיבי ־ כל מטריצה דומה לעצמה: $A = IAI^{-1}$.
- סימטרי ־ אם A דומה ל־B אז B דומה ל־A,
  כי: $A = PBP^{-1}$ ולכן $B = P^{-1}AP = QAQ^{-1}$ כאשר מסמנים $Q = P^{-1}$.
- טרנזיטיבי ־ אם A דומה ל־B, ו־B דומה ל־C, אז A דומה ל־C.
  הוכחה: נניח $A = PBP^{-1}$, $B = QCQ^{-1}$ אז
  $A = PQCQ^{-1}P^{-1} = (PQ)C(Q^{-1}P^{-1}) = TCT^{-1}$ כאשר מסמנים $T = PQ$.

המשמעות של הדמיון הוא ששתי המטריצות מייצגות את אותה העתקה.

הערה ־ עיקרון כללי במתמטיקה:

אם יש משפחה מאוד גדולה של עצמים,

מגדירים יחס שקילות על אוסף העצמים וחוקרים את אוסף מחלקות השקילות.

זה מתקשר לבעיית הליכסון ־ בהינתן מטריצה, ננסה למצוא מטריצה אלכסונית שדומה לה, כלומר למצוא נציג פשוט במחלקת השקילות. עוד בעיה היא לאפיין את כל מחלקות השקילות.

2.2.1 דמיון וערכים עצמיים ־ טענה

טענה 2.12 אם A ו־B דומות, אז כל ע"ע λ של A הוא ע"ע של B, ולהיפך.

(האם זה נכון שמט' בעלות אותם ע"ע הן בהכרח דומות?) (לא).

הוכחה: אם $\lambda \in F$ הוא ע"ע של B, כלומר קיים $0 \neq v \in F^n$ כך ש־ $Bv = \lambda v$,

ומתקיים $A = PBP^{-1}$, אז נראה ש־$u = Pv$ הוא ו"ע של A עם ע"ע λ:

$$Au = A(Pv) = (PBP^{-1})(Pv) = PB(P^{-1}P)v = PBv = P(\lambda v) = \lambda(Pv) = \lambda u$$

הוכחה נוספת:

λ ע"ע של $A \Leftrightarrow$ $\lambda I - A$ אינה הפיכה$\Leftrightarrow$

$\Leftrightarrow$ $\lambda I - PBP^{-1} = \lambda PP^{-1} - PBP^{-1} = P(\lambda I - B)P^{-1}$ אינה הפיכה$\Leftrightarrow$

$\Leftrightarrow$ $\lambda I - B$ אינה הפיכה$\Leftrightarrow$ λ ע"ע של B. ∎

2.2.2 הקשר בין דמיון לליכסון

כזכור, אם $A \in M_n(F)$ ניתנת לליכסון, אז יש בסיס $B = \{v_1, ..., v_n\}$ של ו"ע של A,

כלומר $Av_i = \lambda_i v_i$ לכל $1 \leq i \leq n$, כלומר $[A]_B = \begin{pmatrix} \lambda_1 & & \cdots & 0 \\ & \lambda_2 & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & \lambda_n \end{pmatrix}$.

מצד שני, $A = [A]_E$ כאשר E הבסיס הסטנדרטי. ולכן מתקיים:

$$[A]_E = P[A]_B P^{-1}$$

כאשר $P = M_E^B$.

כלומר:

$$P = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix} \quad \text{and also:} \quad A = P \begin{pmatrix} \lambda_1 & & \cdots & 0 \\ & \lambda_2 & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & \lambda_n \end{pmatrix} P^{-1}$$

כאשר $P = M_E^B$ בצד שמאל היא כזו מכיוון שזה ייצוג לפי E של B.

לסיכום:

A ניתנת לליכסון אם"ם קיים בסיס ו"ע של A,

אם"ם קיימת מטריצה הפיכה P כך ש־ $D = PAP^{-1}$ ("הצמדה של A ע"י P") היא אלכסונית.
ובמקרה זה, עמודות P הן בסיס של ו"ע.

2.3 דוגמאות (כולל פיבונצ'י)

דוגמא ראשונה:

נלכסן את: $A = \begin{pmatrix} 13 & -4 \\ 30 & -9 \end{pmatrix}$. ראשית נמצא בסיס של ו"ע:

שלב I**:** מציאת הע"ע:

$$0 = det(\lambda I - A) = \begin{vmatrix} \lambda - 13 & 4 \\ -30 & \lambda + 9 \end{vmatrix} = (\lambda - 13)(\lambda + 9) + 30 \cdot 4 = \lambda^2 - 4\lambda + 3$$

פותרים: $\lambda^2 - 4\lambda + 3 = 0$:

$$\lambda_{1,2} = \frac{4 \pm \sqrt{16 - 12}}{2} = 1 \ , \ 3$$

שלב II**:** מציאת הו"ע:

$\lambda_1=1$, $\lambda_2=3$. נמצא ו"ע: צריך לפתור את המשוואה: (נסמן $v_1=\begin{pmatrix} a \\ b \end{pmatrix}$)

$$(\lambda_1 I-A)v_1=0 \ \Rightarrow \ \begin{pmatrix} -12 & 4 \\ -30 & 10 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

כלומר

$$\begin{cases} -12a+4b=0 \\ -30a+10b=0 \end{cases}$$

הפתרון: למשל $a=1\ ,\ b=3$ (במקרה יצא כמו הע"ע), $v_1=\begin{pmatrix} 1 \\ 3 \end{pmatrix}$.

נמצא עוד ו"ע $v_2=\begin{pmatrix} c \\ d \end{pmatrix}$ עבור $\lambda_2=3$. נפתור:

$$(\lambda_2 I-A)v_2=0 \ \Rightarrow \ \begin{pmatrix} -10 & 4 \\ -30 & 12 \end{pmatrix}\begin{pmatrix} c \\ d \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

נקבל (למשל):

$$-10c+4d=0 \ \Rightarrow \ v_2=\begin{pmatrix} 2 \\ 5 \end{pmatrix}$$

סה"כ מצאנו בסיס $\left\{\begin{pmatrix} 1 \\ 3 \end{pmatrix}\begin{pmatrix} 2 \\ 5 \end{pmatrix}\right\}$ של ו"ע.

קיבלנו מטריצה מלכסנת: $P=\begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$

ומתקיים: $A=P\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}P^{-1}$.

דוגמא שניה: נוסחה מפורשת למספרי פיבונאצ'י

מספרי פיבונאצ'י מוגדרים ע"י: $f_0=f_1=1$, ובנוסף $f_{n+1}=f_n+f_{n-1}$.
נמצא נוסחה מפורשת ולא רקורסיבית ל-f_n.

נקשר את הבעיה לאלגברה ליניארית: $v_n=\begin{pmatrix} f_n \\ f_{n-1} \end{pmatrix}$.

אז מתקיים: $v_1=\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, וגם:

$$v_{n+1}=\begin{pmatrix} f_{n+1} \\ f_n \end{pmatrix}=\begin{pmatrix} f_n+f_{n-1} \\ f_n \end{pmatrix}=\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} f_n \\ f_{n-1} \end{pmatrix}=\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}v_n$$

כלומר, $v_{n+1} = Av_n$ כאשר $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$. לכן,

$$v_n = Av_{n-1} = A \cdot Av_{n-2} = A \cdot A \cdot Av_{n-3} = ... = A^{n-1}v_1 = A^{n-1}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

ולכן מספיק לחשב את A^{n-1}.

אם נמצא ליכסון של A: $A = PDP^{-1}$, $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ אז:

$$A^n = (PDP^{-1})(PDP^{-1})...(PDP^{-1}) = PD(P^{-1}P)D(P^{-1}P)...DP^{-1} = PD^nP^{-1}$$

ומכיוון שהיא אלכסונית, $D^n = \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix}$.

כלומר, D^n ו־ P הן ליכסון עבור A^n.

נלכסן את A:

$$0 = det(\lambda I - A) = \begin{vmatrix} \lambda - 1 & -1 \\ -1 & \lambda \end{vmatrix} = (\lambda-1)\lambda - 1 = \lambda^2 - \lambda - 1$$

פתירת המשוואה הנ"ל נותנת לנו את **"חתך הזהב"**, מספר ממש מעניין שמופיע בכל מיני מקומות לא צפויים. למשל, אם מפחיתים ממנו 1 מקבלים את אחד חלקי עצמו. בכל אופן,

$$\lambda_1 = \frac{1+\sqrt{5}}{2}\ ,\ \lambda_2 = \frac{1-\sqrt{5}}{2}$$

נמצא את הו"ע: תחילה ו"ע ל־λ_1:

$$\lambda_1 I - A = \begin{pmatrix} \frac{-1+\sqrt{5}}{2} & -1 \\ -1 & \frac{1+\sqrt{5}}{2} \end{pmatrix}$$

ו"ע הוא פתרון $\begin{pmatrix} a \\ b \end{pmatrix} = v_1 \neq 0$ למשוואות:

$$\begin{cases} \frac{-1+\sqrt{5}}{2}a - b = 0 \\ -a + \frac{1+\sqrt{5}}{2}b = 0 \end{cases}$$

למשל: $v_1 = \begin{pmatrix} 1 \\ \frac{-1+\sqrt{5}}{2} \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}$

כעת נמצא ו"ע ל־λ_2:

$$\lambda_2 I - A = \begin{pmatrix} -(\frac{1+\sqrt{5}}{2}) & -1 \\ -1 & \frac{1-\sqrt{5}}{2} \end{pmatrix}$$

ו"ע הוא פתרון $0 \neq v_2 = \begin{pmatrix} c \\ d \end{pmatrix}$ למשוואות:

$$\begin{cases} -\frac{1+\sqrt{5}}{2}c - d = 0 \\ -c + \frac{1-\sqrt{5}}{2}d = 0 \end{cases}$$

למשל: $v_2 = \begin{pmatrix} 1 \\ -\frac{1+\sqrt{5}}{2} \end{pmatrix} = \begin{pmatrix} c \\ d \end{pmatrix}$.

סה"כ מצאנו ליכסון $A = PDP^{-1}$, כאשר: (השתמשנו בנוסחה: $P^{-1} = \frac{1}{|P|}adj(P)$).

$$P = \begin{pmatrix} 1 & 1 \\ \frac{-1+\sqrt{5}}{2} & -\frac{1+\sqrt{5}}{2} \end{pmatrix}, \ D = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix}, \ P^{-1} = \frac{1}{\sqrt{5}} \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 1 \\ \frac{-1+\sqrt{5}}{2} & -1 \end{pmatrix}$$

נותר לחשב את: $\begin{pmatrix} f_n \\ f_{n-1} \end{pmatrix} = v_n = PD^{n-1}P^{-1}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

אחרי מספר חישובים מגיעים לתוצאה:

$$f_n = \frac{1}{\sqrt{5}}\left[(\frac{1+\sqrt{5}}{2})^{n+1} - (\frac{1-\sqrt{5}}{2})^{n+1}\right]$$

(הערה: ייתכן שיש בחישוב זה טעויות - כך השתמע מהדיון בהרצאה, כנראה).

3 פולינום אופייני, פולינום מינימלי

3.0.1 פולינומים - הגדרה ותכונות

הגדרה 3.1 פולינום - מהו.

פולינום p מעל שדה $\mathbb{F}$ הוא **ביטוי פורמלי:**

$$p(t) = a_0 + a_1 t + a_2 t^2 + ... + a_n t^n$$

כאשר $a_0, ..., a_n \in \mathbb{F}$ "מקדמי הפולינום". a_0 ייקרא "המקדם הקבוע", ו-a_n "המקדם העליון".

סדר המקדם העליון $a_n \neq 0$ ייקרא "מעלת הפולינום" ויסומן $deg(p)$.

במקרה של פולינום האפס, המעלה לא מוגדרת, ולפעמים אומרים שהיא $-\infty$.

נסמן: $\mathbb{F}[t]$ - "אוסף של פולינומים מעל $\mathbb{F}$" (המבין יבין...)

זהו **חוג.**

סכום, מכפלה של פולינומים

- אם $p=\sum_{i=0}^{n}a_it^i$ פולינום, ו־$\lambda\in F$ סקלר, אז λp הפולינום $\lambda p=\sum_{i=0}^{n}(\lambda a_i)t^i$.
- אם $q=\sum_{i=0}^{m}b_it^i$, $p=\sum_{i=0}^{n}a_it^i$ אז מגדירים: $p+q:=\sum_{i=0}^{max(m,n)}(a_i+b_i)t^i$, כאשר מקדמים מעל המקדם העליון מחשיבים אותם כ־0.
- מכפלת פולינומים p,q כנ"ל מוגדרת כפולינום: $p\cdot q=\sum_{i=0}^{m+n}c_it^i$, כאשר $c_i=\sum_{k=0}^{i}a_kb_{i-k}$

תכונות של פולינומים

- $deg(fg)=deg(f)+deg(g)$
- $deg(f+g)\leq max(deg(f),deg(g))$

חלוקת פולינומים

הגדרה 3.2 ריבוי, מחלקים

אם $p,q\in F[t]$, אומרים ש־q **מחלק את** p (סימון: $q|p$)

אם קיים פולינום $r\in F[t]$ כך ש־ $p(t)=q(t)r(t)$.

אם $\alpha\in F$ שורש של פולינום $p\in F[t]$, כלומר $p(\alpha)=0$,

אז **הריבוי** של α כשורש של p הוא המספר המקסימלי k כך ש־ $(t-\alpha)^k|p(t)$.

אם הוא מריבוי 1, אז אומרים שהוא **שורש פשוט**.

3.0.2 חלוקת פולינומים עם שארית

הערה: הוכחנו את זה רק בתרגול, יכול להיות שזה לא חלק מחומר הקורס הרשמי?

טענה 3.3 אם $f,g\in F[t]$ פולינומים ($g\neq 0$), אז קיימים פולינומים $q,r\in F[t]$ **יחידים** כך ש $f(t)=g(t)q(t)+r(t)$ **וגם** $deg(r)<deg(g)$.

הוכחה: ראשית,

אם $deg(f)<deg(g)$,

אז ניקח $f=0\cdot g+f$ **ואז** $deg(r)<deg(g)$. **(לקחנו** $r=f,\ q\equiv 0$**).**

ולכן נניח ש־$deg(f)\geq deg(g)$ **(נזכור ש**$g\neq 0$**).**

נוכיח באינדוקציה על $deg(f)$ שקיים q,r כנ"ל.

בסיס האינדוקציה ־ אם $deg(f)=0$:

אז נסמן $f=a_0$, $g=b_0$ **ואז** $f=a_0=\frac{a_0}{b_0}b_0=\frac{a_0}{b_0}g$.

נסמן $q=\frac{a_0}{b_0}$ ונקבל $r=0\Leftrightarrow f-gq=r=0$. **ולכן,**

$$deg(r)=deg(f-gq)=deg(0)=-\infty<deg(g)=0$$

כעת, נניח ש־ $deg(g)=n\leq deg(f)=n+d$, **ו־** $0<n+d$, **ו־** $0\leq d$.

נסמן:

$$g(t)=\sum_{i=0}^{n}b_it^i \qquad f(x)=\sum_{i=0}^{n+d}a_it^i$$

אז:

$$deg\left(f(t)-\frac{a_{n+d}}{b_n}t^dg(t)\right)<n+d$$

הסבר: גם $f(t)$ וגם $\frac{a_{n+d}}{b_n}x^dg(t)$ הם פולינומים ממעלה $n+d$ כאשר המקדם העליון של שניהם הוא a_{n+d}, ולכן החיסור שלהם "משמיד" את המקדם העליון, ומוריד את דרגת פולינום התוצאה).

ולכן, מהנחת האינדוקציה, קיים q,r כך ש־

$$f(t)-\frac{a_{n+d}}{b_n}x^dg(t)=g(t)q(t)+r(t)$$

נעביר אגפים ונקבל המבוקש:

$$f(t)=\left(\frac{a_{n+d}}{b_n}t^d+q(t)\right)g(t)+r(t)$$

■

מסקנה 3.4 אם $\alpha\in F$ הוא שורש של p אז $(t-\alpha)|p(t)$.

הוכחה: \ הסבר:

נחלק עם שארית: $p(t)=(t-\alpha)q(t)+r(t)$, **ואז,**

$deg(r)<deg(t-\alpha)=1$ **כלומר** $r(t)=\beta\in\mathbb{F}$ **קבוע.**

נציב את α במשוואה: $0=p(\alpha)=(\alpha-\alpha)f(\alpha)+\beta$ **ולכן** $\beta=0$.

(כלומר, "אין שארית"). ■

3.1 מטריצות שאיבריהן הם פולינומים

המטרה: לכל מטריצה $A\in M_n(F)$ להגדיר פולינום $f_A(t)=a_0+...+a_{n-1}t^{n-1}+t^n$,

כך שלכל $\lambda\in F$ יתקיים $det(\lambda I-A)=f_A(\lambda)\in F$. **(זה יהיה הפולינום האופייני).**

בשלב זה נשים לב ליצור חדש ומוזר:

מטריצה שאיבריה הם פולינומים: ונסמן את אוסף המטריצות האלה ב־ $M_n(\mathbb{F}[t])$.

(עוד בנושא יהיה בהמשך, ב"עוד על פולינומים ומטריצות") .

לדוגמא: $\begin{pmatrix} t^2+2 & -3 \\ 0 & 3t^2+4t \end{pmatrix}=(f_{ij})_{\substack{i=1,...,n\\ j=1,...,n}}\in M_n(F[t])$

הגדרה 3.5 אם $(a_{ij}) = A \in M_n(F)$ אז נגדיר:

$$tI - A = \begin{pmatrix} t & & 0 \\ & \ddots & \\ 0 & & t \end{pmatrix} - \begin{pmatrix} & & \\ & a_{ij} & \\ & & \end{pmatrix} = \begin{pmatrix} t - a_{11} & -(a_{12}) & \cdots & -(a_{1n}) \\ -(a_{21}) & t - a_{22} & & \\ \vdots & & \ddots & \\ -(a_{n1}) & & & t - a_{nn} \end{pmatrix}$$

כלומר $tI - A = (f_{ij})$ כך ש־

$$\boxed{f_{ij} = \begin{cases} t - a_{ii} & if\ i = j \\ -(a_{ij}) & if\ i \neq j \end{cases}}$$

(נשתמש ב־(f_{ij}) הזה בסעיפים הבאים הרבה, נא לשים לב...)

כדוגמא, נתבונן עבור $n = 2$:

$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, אז: $tI - A = \begin{pmatrix} t - a & -b \\ -c & t - d \end{pmatrix}$

הדטרמיננטה של מטריצה שכזאת מוגדרת באופן זהה לזו של מטריצה "רגילה", אבל לפי ההגדרות המתאימות לפולינומים (כפל, חיבור וכו'):

הגדרה 3.6 דטרמיננטה של מטריצה ב־ $M_n(\mathbb{F}[t])$
לכל $A = (f_{ij}) \in M_n(\mathbb{F}[t])$ נגדיר:

$$det(A) = \sum_{\sigma \in Sn} sgn(\sigma) \cdot f_{1\sigma(1)} \cdot \ldots \cdot f_{n\sigma(n)}$$

נחזור לדוגמא של $n = 2$:

$$det(tI - A) = (t - a)(t - d) - ((-b)(-c)) = t^2 - t(a + d) + (ad - bc)$$

נשים לב שקיבלנו את $Tr(A) = a + d$ ואת $det(A) = ad - bc$ כמקדמים!! זוהי תכונה שנוכיח אח"כ.
מה שמביא אותנו ל־

3.2 הפולינום האופייני

3.2.1 הגדרה, והיות ע"ע שורשים של f_A

הגדרה 3.7 הפולינום האופייני של A
זהו הפולינום:

$$f_A(t) = det(tI - A) \in \mathbb{F}[t]$$

למה 3.8 לכל $\lambda \in F$, מתקיים: $f_A(\lambda) = det(\lambda I - A)$.
זה לכאורה טריוויאלי, אבל משום מה זה דורש הוכחה. (!?)

הוכחה: ההוכחה נכונה בגלל איך שמגדירים חיבור\כפל פולינומים:
לכל $f, g \in F[t]$:

$$(f+g)(\lambda) = f(\lambda) + g(\lambda)$$
$$(fg)(\lambda) = f(\lambda)g(\lambda)$$

ולכן, (הביטוי הבא מסמל "הצבה", כמו "$f(x)$")

$$f_A(\lambda) = (det(tI - A))\,(\lambda) =$$

$$\sum_{\sigma \in S_n} \left(sgn(\sigma) f_{1\sigma(1)} \cdot ... \cdot f_{n\sigma(n)}\right)(\lambda) = \sum_{\sigma \in S_n} \left[sgn(\sigma)\left(f_{1\sigma(1)} \cdot ... \cdot f_{n\sigma(n)}\right)(\lambda)\right] =$$

$$= \sum_{\sigma \in S_n} sgn(\sigma) \cdot f_{1\sigma(1)}(\lambda) \cdot ... \cdot f_{n\sigma(n)}(\lambda) = det(f_{ij}(\lambda)) = det(\lambda I - A)$$

■

מסקנה 3.9 ערכים עצמיים של A הם שורשים של f_A, ב־$\mathbb{F}$.
(כי הרי $\lambda \in F$ הוא ע"ע של A אם"ם $det(\lambda I - A) = 0$)

3.2.2 תכונות הפולינום האופייני - $-Trace(A),\ (-1)^n det(A)$ והיותו מתוקן

משפט 3.10 לכל $A \in M_n(\mathbb{F})$, הפולינום $f_A(t)$ הוא מתוקן מדרגה n, כלומר

$$f_A(t) = a_0 + a_1 t + ... + a_{n-1} t^{n-1} + 1 \cdot t$$

ובנוסף, מתקיים:

$$a_0 = (-1)^n det(A)$$
$$a_{n-1} = -Tr(A)$$

הוכחה: (ההוכחה היתה ממש לא ברורה, אתנצל מראש).

ראשית נזכור ש־ $(tI-A)_{ij} = f_{ij} = \begin{cases} t-a_{ii} & if\ i=j \\ -(a_{ij}) & if\ i \neq j \end{cases}$. ולכן,

$$deg(f_{ij}) = \begin{cases} 1 & if\ i=j \\ \leq 0 & if\ i \neq j \end{cases}$$

כאשר $deg \leq 0$ הוא 0 או $-\infty$.

(0 אם $a_{ij} \neq 0$ או $-\infty$ אם $a_{ij} = 0$ "פולינום האפס").

כמו כן נזכור שמתקיים: $\begin{matrix} deg(fg) = deg(f) + deg(g) \\ deg(f+g) \leq max(deg(f), deg(g)) \end{matrix}$ (נשתמש בזה עוד רגע).

וכעת, מתקיים:

$$\bigstar \qquad f_A(t) = \sum_{\sigma \in S_n} sgn(\sigma) f_{1\sigma(1)} \cdot ... \cdot f_{n\sigma(n)}$$

נרצה לראות מה ה־deg של זה עבור σ כלשהו:

$$deg\left(sgn(\sigma) f_{1\sigma(1)} \cdot ... \cdot f_{n\sigma(n)}\right) = \sum_{i=1}^{n} deg(f_{i\sigma(i)}) = \begin{cases} (1+...+1) = n & if\ \sigma(i) = i \quad (\sigma\text{=Id}) \\ < n & if\ \sigma(i) \neq i \quad (\sigma\neq\text{Id}) \end{cases}$$

כלומר אם $\sigma = Id$ אז $deg = n$, (זהו חיבור ה"אלכסון" של ה־ $(t-a_{ii})$ במטריצה $(tI-A)$).

ואם $\sigma \neq Id$ אז $deg \leq n-2$, כי במצב זה יש לפחות 2 i־ים כך ש־$\sigma(i) \neq i$.

ולכן, נחזור אל $\bigstar$:

$$f_A(t) = \sum_{\sigma \in S_n} sgn(\sigma) f_{1\sigma(1)} \cdot ... \cdot f_{n\sigma(n)} = (t-a_{11}) \cdot ... \cdot (t-a_{nn}) + \sum_{\sigma \neq Id} (\text{something})$$

כאשר המכפלה משמאל היא תרומת $\sigma = Id$, וכל השאר לא מעניין אותנו,

חוץ מזה שהוא (הביטוי "$\sum$something") בהכרח פולינום מדרגה $\leq (n-2)$.

כעת, ניעזר **בלמה**:

$$(t-a_{11}) \cdot ... \cdot (t-a_{nn}) = t^n - (a_{11} + ... + a_{nn})t^{n-1} + [\text{polynomial of deg} \leq \text{n-2}]$$

(לא ממש הוכחנו את זה, כנראה שאפשר להניח שזה נכון).

מסקנה: (משני החישובים האחרונים)

$$f_A(t) = t^n - (a_{11} + ... + a_{nn})t^{n-1} + [\text{polynomial of deg} \leq \text{n-2}]$$

ובפרט, $deg(f_A)=n$, וגם $a_n=1$ (כלומר f_A הוא מתוקן) וגם $a_{n-1}=-Tr(A)$.

•לגבי a_0, צריך להראות שהוא שווה ל $(-1)^n det(A)$: (המעבר האחרון נובע ממולטי־ליניאריות הדטרמיננטה)

$$a_0 = f_A(0) = det(0I - A) = det(-A) = (-1)^n det(A)$$

ובכך הוכחנו את הנדרש. ■

3.2.3 מסקנה לגבי מטריצות דומות A, B: $f_A(t) = f_B(t)$

משפט 3.11 אם $A, B \in M_n(F)$ מטריצות דומות, אז $f_A(t) = f_B(t)$

בפרט, $det(A) = det(B)$ ו־ $Tr(A) = Tr(B)$

כלומר, f_A הוא שמורה (אינווריאנט) של יחס הדמיון.

הוכחה: נניח ש־ $A = PBP^{-1}$. אזי,

$$P(tI - B)P^{-1} = (tP - PB)P^{-1} = tI - PBP^{-1} = tI - A$$

ולכן,

$$f_A(t) = det(tI - A) = det(P(tI - B)P^{-1}) = det(P)det(tI - B)det(P^{-1}) =$$

$$= det(P)det(P^{-1})det(tI - B) = det(PP^{-1})det(tI - B) = 1 \cdot det(tI - B) = f_B$$

■

נשים לב: אנו מניחים שמתקיים $detAB = detA \cdot detB$ גם ב־$M_n(F[t])$.
ההוכחה לכך לא ניתנה, עד כמה שידוע לי, אבל אפשר לראות שזה נכון מכך שפולינום הוא איבר בשדה הפונקציות הרציונליות: $\{\frac{f}{g} \mid f, g \in \mathbb{F}[t]\}$, וידוע לנו שב־$M_n(\mathbb{F})$ עבור שדה $\mathbb{F}$ מתקיימת הכפליות הזו.

3.3 הפולינום המינימלי

3.3.1 קצת טענות מקדימות

הגדרה 3.12 הצבת מטריצה בפולינום

לכל $f(t) = a_0 + a_1 t + ... + a_n t^n \in F[t]$ ו־ $A \in M_n(F)$ נגדיר:

$$f(A) = a_0 I + a_1 A + a_2 A^2 + ...$$

המוטיבציה לפולינום המינימלי היא כזו:

בהינתן $A \in M_n(F)$, מהם כל $f \in F[t]$ כך ש־$f(A) = 0 \in M_n(F)$?

תחילה נוכיח למה:

למה 3.13 קיים $0 \neq f \in F[t]$ מדרגה $\geq n^2$ כך ש־$f(A) = 0$.

הוכחה: אנו רוצים למצוא:

$$f(t) = a_0 + a_1 t + ... + a_{n^2} t^{n^2}$$

כך ש:

$$0 = f(A) = a_0 I + a_1 A + ... + a_{n^2} A^{n^2}$$

כעת,

קיים $f(t) \neq 0$ מדרגה $\geq n^2$ כנדרש $\Leftrightarrow$

$\Leftrightarrow$ המטריצות $I, A, A^2, ..., A^{n^2}$ הן תלויות ליניארית ב־$M_n(F)$.

(לפי התנאי הבסיסי לבדיקת תלות ליניארית מאלגברה ליניארית 1).

אבל, מתקיים $M_n(\mathbb{F}) \simeq \mathbb{F}^{n^2}$ (איזומורפיזם) כמרחב וקטורי מעל $\mathbb{F}$, ולכן $dim_{\mathbb{F}} M_n(\mathbb{F}) = n^2$.

ולכן, אם נשים לב שיש בקבוצה $I, A, A^2, ..., A^{n^2}$ בדיוק $n^2 + 1$ איברים, שזה יותר ממימד המ"ו, הרי שבהכרח היא קבוצה תלויה ליניארית. ולכן ממה שאמרנו, קיים $f(t)$ כנדרש. ■

דוגמא:

תהי $A = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}$. אז $f_A(t) = det \begin{pmatrix} t - a_1 & & 0 \\ & & \\ 0 & & t - a_n \end{pmatrix} = (t - a_1) \cdot ... \cdot (t - a_n)$

ולכן,

$$f_A(A) = (A - a_1 I)(A - a_2 I) \cdot ... \cdot (A - a_n I) =$$

$$= \begin{pmatrix} 0 & & & & 0 \\ & (a_2 - a1) & & & \\ & & \ddots & & \\ & & & \ddots & \\ 0 & & & & (a_n - a_1) \end{pmatrix} \begin{pmatrix} * & & & & 0 \\ & 0 & & & \\ & & * & & \\ & & & \ddots & \\ 0 & & & & * \end{pmatrix} \cdot ... \cdot \begin{pmatrix} * & & & & 0 \\ & * & & & \\ & & \ddots & & \\ & & & \ddots & \\ 0 & & & & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} 0 & & & & 0 \\ & 0 & & & \\ & & \ddots & & \\ & & & \ddots & \\ 0 & & & & 0 \end{pmatrix}$$

3.3.2 הפולינום המינימלי

הגדרה 3.14 \ טענה: פולינום מינימלי

עבור $A \in M_n(F)$, קיים פולינום **יחיד, מתוקן,** ממעלה מינימלית $g \in F[t]$ **כך ש־** $g(A) = 0$.

הוא נקרא **הפולינום המינימלי** של A, **ויסומן** m_A.

הסבר ללמה הוא **יחיד:** אם g_1, g_2 שני פולינומים כאלה, הם מאותה מעלה ומתוקנים. **ולכן** $g_1 - g_2$ **יהיה פולינום ש־**A **מאפסת,** ממעלה יותר נמוכה (כי המקדם העליון התאפס בעקבות היותם מתוקנים), **ולכן בהכרח** $g_1 - g_2 \equiv 0$.

3.3.3 אם $A, B \in M_n(F)$ דומות אז $m_A = m_B$

טענה 3.15 אם $A, B \in M_n(F)$ דומות אז $m_A = m_B$.

הוכחה: אם $A = P^{-1}BP$ **ו־** $g(t) = \sum c_i t^i \in F[t]$ **וגם** $deg(g) = m$ **אז:**

$$g(A) = c_0 I + c_1 A + c_2 A^2 + ... + c_m A^m =$$

$$= c_0 P^{-1}P + c_1 P^{-1}BP + c_2 P^{-1}B^2 P + ... + c_m P^{-1}B^m P =$$

$$= P^{-1}\left(c_0 I + c_1 B + c_2 B^2 + ... + c_m B^m\right)P = P^{-1}g(B)P$$

(נשים לב שהשתמשנו בתכונה שראינו בהוכחה של פיבונאצ'י,

לגבי העלאת מטריצה דומה בחזקה: למשל $P^{-1}B^2P = P^{-1}BPP^{-1}BP$.)

ובפרט, $g(A) = 0$ **אם ורק אם** $g(B) = 0$. **ולכן** $m_A = m_B$. ■

4 מרחבים עצמיים, ריבוי אלגברי וריבוי גיאומטרי

4.1 מרחב עצמי, ריבוי גיאומטרי

הגדרה 4.1 מרחב עצמי \ ריבוי גיאומטרי

תהי $A \in M_n(F)$. אם $\lambda \in F$ ע"ע של A, **המרחב העצמי** של A **המתאים לע"ע** λ

הוא $V_\lambda = Ker(A - \lambda I)$. **זהו תת־מרחב ליניארי של** F^n. (הערה: לא משנה אם זה $A - \lambda I$ או $\lambda I - A$)

המימד $dim(V_\lambda)$ נקרא **הריבוי הגיאומטרי** של הע"ע λ.

כלומר,

V_λ - אוסף כל הו"ע המתאימים לע"ע λ, ביחד עם וקטור האפס (שלא נחשב ו"ע...).

הריבוי הגיאומטרי הוא מספר הו"ע הבלתי תלויים ליניארית המקסימלי שמתאימים לע"ע λ.

הוא גם **אינווריאנט ליחס הדמיון:** אם A, B דומות אז הריבוי הגיאומטרי של כל ע"ע λ יהיה אותו דבר ל־A ול־B. (זה נובע מכך שע"ע של דומות הם זהים. הושאר כתרגיל, ללא הוכחה).

טענה 4.2 אם $\lambda_1, ..., \lambda_k$ הע"ע של $A \in M_n(F)$,
עם ריבויים גיאומטריים $d_1, ..., d_k$ בהתאמה, אז $\sum_{i=1}^{k} d_i \le n$.

הוכחה: לכל $1 \le i \le k$, יהי $\{v_{i,1}, v_{i,2}, ..., v_{i,d_i}\}$ בסיס למרחב העצמי V_{λ_i}.
נראה ש־ $\{v_{i,j}\}_{\substack{1 \le i \le k \\ for\ each\ i:\ 1 \le j \le d_i}}$ קבוצה בת"ל, ולכן עוצמתה $\sum d_i$ היא $\le n$.
(כי כל $v_{i,j}$ הוא וקטור ב־F^n ולכן כל קבוצה בת"ל של כאלה היא $\le n$)(הערה שלי).
נניח אם כן, שקיים צ"ל של וקטורים אלה ששווה 0. הוא יהיה מהצורה:

$$\underbrace{c_{1,1}v_{1,1} + c_{1,2}v_{1,2} + ... + c_{1,d_1}v_{1,d_1}}_{V'_{\lambda_1} s\ eigenspace} \underbrace{+c_{2,1}v_{2,1} + c_{2,2}v_{2,2} + ... + c_{2,d_2}v_{2,d_2}}_{V'_{\lambda_2} s\ eigenspace} \underbrace{+...+}_{all\ the\ rest...}$$

$$\underbrace{+c_{k,1}v_{k,1} + c_{k,2}v_{k,2} + ... + c_{k,d_k}v_{k,d_k}}_{V'_{\lambda_k} s\ eigenspace} = 0$$

אם נסמן $u_i = c_{i,1}v_{i,1} + c_{i,2}v_{i,2} + ... + c_{i,d_i}v_{i,d_i}$ אז קיבלנו ש־

$$u_1 + u_2 + ... + u_k = 0$$

וכל $u_i \in V_{\lambda_i}$. נשים לב שu_i הוא צ"ל של בסיס למרחב עצמי ולכן הוא ו"ע. (כנראה ־ זו דעתי...)
ולכן, לפי טענה שהראינו:
־ "אם $u_1, ..., u_k$ ו"ע של A עם ע"ע **שונים** $\lambda_1, ..., \lambda_k$ בהתאמה, אז הם בת"ל" ־
אז קיבלנו שכל $u_i = 0$. כלומר:

$$u_i = c_{i,1}v_{i,1} + c_{i,2}v_{i,2} + ... + c_{i,d_i}v_{i,d_i} = 0$$

אבל הוקטורים $v_{i,1}, ..., v_{i,d_i}$ הם בסיס ל־V_{λ_i}, ובפרט בת"ל, ולכן $c_{i,j} = 0$ לכל $1 \le j \le d_i$. ■

מסקנה:

משפט 4.3 A ניתנת ללכסון אם"ם סכום הריבויים הגיאומטריים של כל הע"ע שלה הוא בדיוק n.

הוכחה: כיוון אחד: נניח A ניתנת ללכסון.
נסמן: $\lambda_1, ..., \lambda_k$ את הע"ע השונים שלה. ויהיו $d_1, ..., d_k$ הריבויים הגיאומטריים שלהם.
וכעת, ידוע שמטריצה ניתנת ללכסון אם קיים בסיס של ו"ע בגודל n , ולכן,
יהי $\{v_1, ..., v_n\}$ בסיס של ו"ע של A, אותו נכתוב מחדש בצורה: $\{v_{ij}\}_{\substack{1 \le i \le k \\ 1 \le j \le a_i}}$ כלומר,

לכל $1 \le i \le k$, $v_{i,1}, ..., v_{i,a_i}$ הם הו"ע מתוך הבסיס עם ע"ע λ_i.

מכיוון שאלה בסיס, נובע ש־ $a_i \le d_i$ לכל $1 \le i \le k$.

ולכן, $\sum_{i=1}^k d_i \ge \sum_{i=1}^k a_i = n$, כלומר הראינו $\sum d_i \ge n$.

($\sum a_i = n$ בגלל שבסה"כ יש לנו n וקטורים, שחילקנו ל־k קבוצות לפי ע"ע λ_i. כל אחת בגודל a_i שונה, אבל איחודן נותן לנו את n הוקטורים המקוריים).

מצד שני, בטענה הקודמת הוכיחה ש־ $\sum d_i \le n$, ולכן בסה"כ קיבלנו את הנדרש:

$$\sum_{i=1}^{n} d_i = n$$

כיוון שני: נניח שסכום הריבויים הגיאומטריים של כל הע"ע של A הוא בדיוק n.

כעת, נניח ש־ $\lambda_1, ..., \lambda_k$ הע"ע של A.

לכל $1 \le i \le k$, יהי d_i הריבוי הגיאומטרי של λ_i,

ויהי $\{v_{i,1}, v_{i,2}, ..., v_{i,d_i}\}$ בסיס למרחב העצמי V_{λ_i}.

לפי מה שהוכחנו בטענה הקודמת, הוקטורים $\{v_{i,j}\}_{\substack{1 \le i \le k \\ 1 \le j \le d_i}}$ הם קבוצה בת"ל.

נתון שגודלה n, כלומר זהו בסיס המורכב מו"ע של A, ולכן A ניתנת ללכסון. ■

4.2 ריבוי אלגברי

4.2.1 הגדרה

הגדרה 4.4 אם $\lambda \in F$ ע"ע של $A \in M_n(F)$, **הריבוי האלגברי** שלו הוא הריבוי של λ כשורש של הפולינום האופייני f_A.

(תזכורת: אם $\alpha \in F$ שורש של פולינום $p \in F[t]$, כלומר $p(\alpha) = 0$,

אז **הריבוי** של α כשורש של p הוא המספר המקסימלי k כך ש־ $(t-\alpha)^k | p(t)$).

4.2.2 טענות ומשפטים

טענה 4.5 הריבוי האלגברי של $\lambda \le$ הריבוי הגיאומטרי.

הוכחה: נניח λ הוא מריבוי גיאומטרי d.

אז יש בסיס $\{v_1, ..., v_d\}$ למרחב העצמי $V_\lambda = Ker(A - \lambda I)$.

נשלים קבוצה זו לבסיס $B = \{v_1, ..., v_d, v_{d+1}, ..., v_n\}$ של $\mathbb{F}^n$. אז:

נתבונן ב־A לפי הבסיס B. לצורך שינוי הבסיס נתבונן ב PAP^{-1}:

$$PAP^{-1} = [A]_B = \left(\begin{array}{cccc|c} \lambda & & & & \\ & \lambda & & & * \\ & & \ddots & & \\ & & & \lambda & \\ \hline & \ddots & & 0 & \\ 0 & & \ddots & & * \end{array}\right)$$

(סליחה על האיור הגרוע. **יש** d **עמודות של** λ **בבלוק השמאלי עליון**, אפסים מתחתיו, ומשהו אחר מימין.)

ולכן,

$$f_A(t) = f_{PAP^{-1}}(t) = det(tI - PAP^{-1}) = det \left(\begin{array}{cccc|c} t-\lambda & & & & \\ & t-\lambda & & & * \\ & & \ddots & & \\ & & & t-\lambda & \\ \hline & \ddots & & 0 & \\ 0 & & \ddots & & * \end{array} \right)$$

כשנכתוב דטרמיננטה זו כסכום על תמורות,

התרומות השונות מ־0 יהיו כולן מהצורה: $(t-\lambda)^d[\text{mashu}]$ ("משהו" באנגלית. תפתחו מילון.)

וסה"כ נקבל ש־ $f_A(t) = (t-\lambda)^d \cdot g(t)$ עבור $g \in F[t]$ כלשהו.

ולכן הריבוי האלגברי הוא $d \leq$.

■

5 עוד על פולינומים ומטריצות

5.1 הצבות של פולינומים במטריצות של פולינומים שאיבריהם מטריצות ש....!?

קצת סדר:

- $\mathbb{F}[t]$ ־ הפולינומים מעל $\mathbb{F}$.
- **הצבה של סקלר בפולינום:** אם $\alpha \in \mathbb{F}$, ו־ $f(t) = \sum_{i=0}^{n} c_i t^i \in \mathbb{F}[t]$,
  אז מגדירים $f(\alpha) = \sum_{i=0}^{n} c_i \alpha^i$ "ההצבה של α ב־f" או "הערך של α ב־f".
- **הצבה של מטריצה בפולינום:** אם $f(t) = \sum_{i=0}^{m} c_i t^i$, $A \in M_n(\mathbb{F})$ אז
  מגדירים: $f(A) = \sum_{i=0}^{m} c_i A^i = c_0 I + c_1 A + c_2 A^2 + ... + c_m A^m$,
  כאשר הגדרנו $A^0 = I$, וזה ייקרא "ההצבה של A ב־f".
- $M_n(\mathbb{F}[t])$ ־ מטריצות שאיבריהן פולינומים.
- $M_n(\mathbb{F})[t]$ ־ פולינומים שהמקדמים שלהם מטריצות.

5.2 קצת הסברים, טענות לגבי $M_n(\mathbb{F})[t]$

הגדרה 5.1 פולינום ב־$M_n(\mathbb{F})[t]$,

הוא ביטוי פורמלי מהצורה $\sum_{i=0}^{m} t^i A_i$,

כאשר $A_0, A_1, ..., A_m \in M_n(\mathbb{F})$ "מקדמי הפולינום".

נשים לב שהכפל משמאל ־ זוהי החלטה שרירותית, אבל נקיים אותה תמיד,

כדי שיעבוד טוב עם "$tI - A$".

לדוגמא,

$$f(t)=\begin{pmatrix}1&2\\0&1\end{pmatrix}+t\begin{pmatrix}3&0\\0&4\end{pmatrix}+t^2\begin{pmatrix}0&-1\\0&1\end{pmatrix}$$

האם זה אותו הדבר כמו $\begin{pmatrix}1+3t&2-t^2\\0&1+4t+t^2\end{pmatrix}$? לא, כי אנחנו לא יודעים אם t הוא סקלר.

ב־$M_n(\mathbb{F})[t]$ אפשר להגדיר פעולות חיבור, כפל בסקלר וכפל.

אם $f,g\in M_n(\mathbb{F})[t]$, מגדירים את $f\cdot g$ ע"י פתיחת סוגריים ו"דחיפת" כל החזקות של t לצד שמאל. כלומר:

$$(\sum_{i=0}^{m}t^iA_i)(\sum_{j=0}^{k}t^jB_j)=\sum_{d=0}^{m+k}t^d(\sum_{i=0}^{d}A_iB_{d-i})$$

הערה: כפל מטריצות אינו חילופי,
ולכן אם מתקיימת המשוואה $f(t)g(t)=h(t)$ כאשר $f,g,h\in M_n(\mathbb{F})[t]$,
זה לא הבכרח אומר ש־ $f(B)g(B)=h(B)$ לכל מטריצה $B\in M_n(\mathbb{F})$ שנציב.
לדוגמא, ניקח $g(t)=f(t)=tI-A$ כאשר $A\in M_n(F)$ נתונה. אז:

$$h(t)=(tI-A)(tI-A)=t^2I+t(-2A)+A^2$$

אם נציב מטריצה B במקום t:

$$f(B)g(B)=(BI-A)(BI-A)=B^2-AB-BA+A^2$$

ומצד שני,

$$h(B)=B^2I-2BA+A^2$$

וביטויים אלה הם שונים כאשר A,B לא מתחלפות.

5.2.1 טענה חשובה, שלפיה אם $h(t)=(tI-A)g(t)$ אז $h(A)=0$

הטענה הנ"ל היא מקרה פרטי של הטענה הבאה:

טענה 5.2 אם $h(t)=(tI-A)g(t)$ ב־$M_n(\mathbb{F})[t]$,
אז $h(B)=(B-A)g(B)$ לכל מטריצה B, כך ש־$AB=BA$.

הוכחה: נסמן: $g(t)=C_0+tC_1+t^2C_2+...+t^mC_m$. אז:

$$h(t)=(tI-A)(C_0+tC_1+t^2C_2+...+t^mC_m)=$$

$$= tC_0 + t^2C_1 + t^3C_2 + ... + t^{m+1}C_m - AC_0 - tAC_1 - ... - t^mAC_m$$

נציב את B:

$$-AC_0 + B(C_0 - AC_1) + B^2(C_1 - AC_2) + ... + B^m(C_{m-1} - AC_m) + B^{m+1}C_m$$

ומצד שני,

$$(B-A)g(B) = (B-A)(C_0 + BC_1 + B^2C_2 + ... + B^mC_m) =$$

$$BC_0 + B^2C_1 + ... + B^{m+1}C_m - AC_0 - ABC_1 - AB^2C_2 - ... - AB^mC_m =$$

$$-AC_0 + B(C_0 - AC_1) + B^2(C_1 - AC_2) + ... + B^m(C_{m-1} - AC_m) + B^{m+1}C_m$$

כאשר המעבר האחרון בגלל ש־ $BA = AB$. ולכן הוכחנו שוויון. ■

מסקנה 5.3 אם $h(t) = (tI - A)g(t)$ אז $h(A) = 0$.

5.3 הקשר בין $M_n(\mathbb{F})[t]$ ו־ $M_n(\mathbb{F}[t])$

שתי הקבוצות $M_n(\mathbb{F})[t]$ ו־ $M_n(\mathbb{F}[t])$ הן שונות, אך יש התאמה חח"ע ועל ביניהן:

הגדרה 5.4 אם $f(t) \in M_n(F)[t]$, $f(t) = \sum_{i=0}^{m} t^iA_i$, אז
נגדיר $\tilde{f}(t) \in M_n(F[t])$ המתאימה לה ע"י:

$$(\tilde{f}(t))_{j,k} = \sum_{i=0}^{m} (A_i)_{j,k}t^i$$

קל לראות שזו התאמה חח"ע ועל.

בנוסף,

טענה 5.5 אם $f, g \in M_n(\mathbb{F})[t]$ אז: $\widetilde{(f+g)} = \tilde{f} + \tilde{g}$ וכן לכל $\lambda \in F$, $\widetilde{(\lambda f)} = \lambda\tilde{f}$.
ללא הוכחה ־ מושאר כתרגיל ("קל").

כלומר, זהו **איזומורפיזם** של מרחבים וקטוריים.

טענה 5.6 בנוסף, מתקיים גם: $\widetilde{(f\cdot g)} = \tilde{f}\cdot\tilde{g}$.

הוכחה: יהיו

$$g(t) = \sum_{j=0}^{k} t^j B_j \qquad f(t) = \sum_{i=0}^{m} t^i A_i$$

אז, $\tilde{f}(t) = \sum_{i=0}^{m} t^i A_i$ כאשר פה המשמעות היא שעושים סכום של מטריצות ב־ $M_n(\mathbb{F}[t])$.
ובאופן דומה, $\tilde{g}(t) = \sum_{j=0}^{k} t^j B_j$. **(??????????)**
ולכן,

$$\tilde{f}(t)\cdot\tilde{g}(t) = (\sum_i t^i A_i)(\sum_j t^j B_j) = \ldots = \sum_{d=0}^{m+k} t^d (\sum_{i=0}^{d} A_i B_{d-i}) = \widetilde{(fg)}$$

כאשר בשלב ה־..., משתמשים בתכונות החיבור והכפל בסקלר ב־$M_n(\mathbb{F}[t])$.
(דרושה הוכחה פורמלית יותר, זה לא מספיק, לדעתי.) ■

לסיכום, הוכחנו:

משפט 5.7 יש איזומורפיזם טבעי ("קנוני") בין החוגים $M_n(\mathbb{F})[t]$ ו־ $M_n(\mathbb{F}[t])$.

6 משפט קיילי - המילטון

הערה: ההוכחה הזאת היא לא למבחן. כך נאמר בהודעה הרשמית.

מומלץ להתרענן בנושא המטריצה המצורפת לפני קריאת הוכחה זו. (כתוב על זה בפרק הראשון בסיכום).

כמו כן זהו המשך ישיר של הסעיף הקודם על האיזומורפיזם בין $M_n(\mathbb{F})[t]$ ו־ $M_n(\mathbb{F}[t])$.

משפט 6.1 קיילי - המילטון
לכל $A \in M_n(F)$, מתקיים: $f_A(A) = 0$

הוכחה: בהנתן $A \in M_n(F)$, נסמן: $A(t) = tI - A \in M_n(\mathbb{F})[t]$, ותהי $\tilde{A}(t) \in M_n(\mathbb{F}[t])$.
נסמן: $\tilde{C}(t) = adj(\tilde{A}(t))$, ו־ $C(t) \in M_n(\mathbb{F})[t]$ המתאימה לה.
ונסמן: $f_A(t) = \sum_{i=0}^{m} b_i t^i$.
אז מתקיים: (בעקבות המשפט: $A \cdot adj(A) = det(A)\cdot I$):

$$\widetilde{(A(t)C(t))} = \tilde{A}(t)\tilde{C}(t) = \begin{pmatrix} det(\tilde{A}(t)) & & \cdots & 0 \\ & det(\tilde{A}(t)) & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & det(\tilde{A}(t)) \end{pmatrix} = f_A(t)\cdot I = \sum_{i=0}^{m} b_i t^i I$$

מכאן נובע לפי האיזומורפיזם, שמתקיים: (בסוף מסמנים $\sum_{i=0}^{m} t^i(b_iI) = h(t)$)

$$(tI - A) \cdot C(t) = A(t)C(t) = \sum_{i=0}^{m} t^i(b_iI) = h(t) \in M_n(\mathbb{F})[t]$$

וכעת, ראינו בעבר ש־ $0 = h(A) = \sum_{i=0}^{m} A^i(b_iI) = \sum_{i=0}^{m} b_iA^i = f_A(A)$.
(בגלל המסקנה הממוסגרת כעמוד או שניים אחורה מכאן) ■

6.0.1 מסקנות מהמשפט

1. אם m_A הפולינום המינימלי של A, אז $deg(m_A) \leq n$.
2. הפולינום המינימלי m_A מחלק את הפולינום האופייני f_A ללא שארית.
   כלומר, $f_A(t) = m_A(t)q(t)$ עבור איזשהו $q \in F[t]$.

הסבר ל־2:

באופן יותר כללי, אם $g \in F[t]$ פולינום כך ש־$g(A) = 0$, נחלק אותו ב־m_A עם שארית, $g(t) = m_A(t)q(t) + r(t)$. אבל, $0 = g(A) = m_A(A)q(A) + r(A)$ כלומר $r(A) = 0$. כמו כן $deg(r) < deg(m_A)$ ולכן בהכרח $r \equiv 0$.

טענה 6.2 אם $g(t) \in M_n(F)[t]$ כך ש־$g(A) = 0$,
אז קיים $h \in M_n(F)[t]$ כך ש־$g(t) = (tI - A)h(t)$.
נשים לב, זה הפוך מהמסקנה שראינו קודם.

הוכחה: נסמן $g(t) = \sum_{i=0}^{m} t^iC_i$. אז:

$$g(t) = g(t) - 0 = g(t) - g(A) = \sum_{i=0}^{m} t^iC_i - \sum_{i=0}^{m} A^iC_i = \sum_{i=0}^{m} (t^iI - A^i)C_i$$

אבל לכל i מתקיים:

$$t^iI - A^i = (tI - A)(A^{i-1} + tA^{i-2} + t^2A^{i-3} + ... + t^{i-2}A + t^{i-1}I)$$

נסמן את האיבר בסוגריים הימניים ב־$h_i(t)$. ולכן סה"כ נקבל:

$$g(t) = \sum_{i=0}^{m} (tI - A)h_i(t)C_i = (tI - A)\left(\sum_{i=0}^{m} h_i(t)C_i\right) = (tI - A) \cdot (h(t))$$

■

טענה 6.3 f_A מחלק את m_A^n, כאשר $A \in M_n(\mathbb{F})$
(משפט זה הוא **לא למבחן** - הוא אחד מהמשפטים שירדו בהודעה הרשמית).

הוכחה: נסמן $m_A(t) = \sum_{i=0}^{k} b_i t^i$, ונגדיר פולינום $M(t) \in M_n(\mathbb{F})[t]$
ע"י $M(t) = m_A(t)I = \sum_{i=0}^{k} t^i (b_i I)$.
מכיוון ש־$m_A(A) = 0$, גם $M(A) = 0$. ולכן לפי הטענה הקודמת, קיים
$h \in M_n(F)[t]$, כך ש־ $M(t) = (tI - A)h(t)$.
נעבור למטריצות ב־$M_n(\mathbb{F}[t])$, ונקבל:

$$M(t) = (tI - A)h(t) \qquad \Leftrightarrow$$

$$\begin{pmatrix} m_A(t) & & \cdots & 0 \\ & m_A(t) & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & m_A(t) \end{pmatrix} = \begin{pmatrix} t - a_{11} & -(a_{12}) & \cdots & -(a_{1n}) \\ -(a_{2,1}) & t - a_{22} & & \\ \vdots & & \ddots & \\ -(a_{n,1}) & & & t - a_{nn} \end{pmatrix} \cdot \tilde{h}(t)$$

ניקח דטרמיננטה של שני האגפים (הדטרמיננטה כפלית - בחוג חילופי [מסתבר?]), ונקבל:
(ממולטי ליניאריות ומהגדרת $f_A(t)$)

$$m_A(t)^n = f_A(t) \cdot det(\tilde{h}(t)) = f_A(t) \cdot "g(t)"$$

(הצעד האחרון הוא סתם סימון). וזה בדיוק מה שרצינו להוכיח. ■

מסקנה 6.4 אם $A \in M_n(F)$, $\lambda \in F$ אז **התנאים הבאים שקולים**:

- λ הוא ע"ע של A
- $f_A(\lambda) = 0$
- $m_A(\lambda) = 0$

הוכחה: כיוון אחד:
$f_A(t) = m_A(t)g(t) \;\Rightarrow\; f_A(\lambda) = m_A(\lambda)g(\lambda)$ ולכן,
אם $m_A(\lambda) = 0$ אז גם $f_A(\lambda) = 0$.
כיוון שני:
$m_A(t)^n = f_A(t)h(t) \;\Rightarrow\; m_A(\lambda)^n = f_A(\lambda)h(\lambda)$ ולכן,
אם $f_A(\lambda) = 0$ אז גם $m_A(\lambda)^n = 0$ ואז גם $m_A(\lambda) = 0$. ■

משפט 6.5 מטריצה $A \in M_n(\mathbb{F})$ ניתנת לליכסון אם"ם הפולינום m_A הוא מכפלה של גורמים ליניאריים שונים, כלומר:

$$m_A(t) = (t-\lambda_1)(t-\lambda_2)...(t-\lambda_k)$$

כאשר $\lambda_1, ..., \lambda_k$ **שונים**.

הוכחה: כיוון א: נניח A ניתנת לליכסון, ויהיו $\lambda_1, ..., \lambda_k$ הע"ע השונים שלה.

כלומר A דומה למטריצה אלכסונית $D = \begin{pmatrix} a_1 & & & 0 \\ & a_2 & & \\ & & \ddots & \\ 0 & & & a_n \end{pmatrix}$ כאשר כל a_i הוא אחד מה־λ_i־ים.
(הם עלולים להופיע יותר מפעם אחת בגלל הריבוי הגיאומטרי).

נראה ש־ $m_A(t) = m_D(t) = \prod_{i=1}^{k}(t-\lambda_i) = g(t)$.

(כאשר $m_A = m_D$ כי זהו אינוריאנט לדמיון).

ראשית, מתקיים: (זה מה שקורה בהצבת מט' אלכסונית בפולינום):

$$g(D) = \begin{pmatrix} g(a_1) & & & 0 \\ & g(a_2) & & \\ & & \ddots & \\ 0 & & & g(a_n) \end{pmatrix} = 0 \in M_n(F)$$

והשיון למטריצת האפס נכון בגלל ש־g מתאפס על $\lambda_1, ..., \lambda_k$.

כמו כן, מהגדרת m_D מתקיים:

$$m_D(D) = \begin{pmatrix} m_D(a_1) & & & 0 \\ & m_D(a_2) & & \\ & & \ddots & \\ 0 & & & m_D(a_n) \end{pmatrix} = 0$$

ולכן m_D מתאפס ב־$\lambda_1, ..., \lambda_k$.

ולכן $\deg(m_D) \geq k$, כלומר $\deg(m_D) \geq \deg(g) = k$. ולכן g הוא הפולינום המינימלי.

כיוון ב: נניח ש־$m_A(t) = \prod_{i=1}^{k}(t-\lambda_i)$. ונראה ש־$A$ ניתנת לליכסון:

נראה שהמרחבים העצמיים $V_{\lambda_i} = ker(A-\lambda_i I) \quad (1 \leq i \leq k)$ פורשים את $\mathbb{F}^n$, כלומר כל וקטור $v \in F^n$ ניתן להצגה כסכום של ו"ע. (וזה יהיה מספיק, לפי מה שכבר הראינו.)

נגדיר לכל $1 \leq i \leq k$ פולינום:

$$g_i(t) = \left(\frac{1}{\prod_{\substack{1 \leq j \leq k \\ j \neq i}} (\lambda_i - \lambda_j)} \right) \cdot \prod_{\substack{1 \leq j \leq k \\ j \neq i}} (t-\lambda_j)$$

נשים לב: הביטוי בסוגריים הוא סקלר, ובצד ימין, j רץ על $1,2,...,i-1,i+1,...k$.

לדוגמא: עבור $k=4$: $m_A(t)=(t-\lambda_1)(t-\lambda_2)(t-\lambda_3)(t-\lambda_4)$ אז $g_1(t)=c_1(t-\lambda_2)(t-\lambda_3)(t-\lambda_4)$ $g_2(t)=c_2(t-\lambda_1)(t-\lambda_3)(t-\lambda_4)$ $g_3(t)=c_3(t-\lambda_1)(t-\lambda_2)(t-\lambda_4)$ $g_4(t)=c_4(t-\lambda_1)(t-\lambda_2)(t-\lambda_3)$

כעת, אם נתון $v\in F^n$, נסמן: $v_i=g_i(A)\cdot v$ $(1\leq i\leq k)$ כלומר הגדרנו וקטורים $v_1,..,v_k$.

ניעזר בשתי טענות עזר:

טענה 1: v_i הוא ו"ע של A עם ע"ע λ_i.

טענה 2: $v=v_1+v_2+...+v_k$.

הוכחת טענה 1:

נתבונן: (ונסמן: $h_i(t)=t-\lambda_i$)

$$(A-\lambda_i I)v_i=(A-\lambda_i I)\cdot g_i(A)\cdot v=(h_i g_i)(A)\cdot v=$$

כאשר הביטוי האחרון ("$(h_i g_i)(A)$") הוא סימון להצבה בפונקציה, לא כפל. נמשיך:

$$=\left(\frac{1}{\prod_{\substack{1\leq j\leq k\\ j\neq i}}(\lambda_i-\lambda_j)}\right)\cdot m_A(A)\cdot v=0\cdot v=0$$

(האפס השמאלי יותר הוא מטריצת האפס, והאפס הימני הוא וקטור האפס).

כלומר, קיבלנו ש- $Av_i=\lambda_i v_i$.

הוכחת טענה 2:

נסמן: $g(t)=g_1(t)+g_2(t)+...+g_k(t)$.

אז $deg(g)\leq k-1$ (כי g הוא סכום של g_i, וכ"א מהם מוגדר ממעלה $\leq k-1$).

כמו כן, נערוך חישוב קצר:

$$g_j(\lambda_i)=\left(\frac{1}{\prod_{\substack{1\leq p\leq k\\ p\neq j}}(\lambda_j-\lambda_p)}\right)\cdot\prod_{\substack{1\leq p\leq k\\ p\neq j}}(\lambda_i-\lambda_p)=\begin{cases}0 & if\, i\neq j\\ 1 & if\, i=j\end{cases}$$

וסה"כ מקבלים ש-

$$g(\lambda_i)=\sum_{j=1}^{k}g_j(\lambda_i)=1$$

ולכן $g(t) = 1$ **(זהותית).** (כנראה שרק עבור $\{\lambda_i\}$? משהו לא ברור פה).

ולכן,

$$\sum_{i=1}^{k} v_i = \sum_{i=1}^{k} g_i(A) \cdot v = g(A) \cdot v = Iv = v$$

(לא ברור לי אישית למה $g(A) = I$, ייתכן שיש פה טעות או חסר הסבר).

■

7 סיכום ליכסון מטריצות

תנאים לליכסון:

התנאים הבאים שקולים:

1. קיים ל־A בסיס של ו"ע (A ניתנת לליכסון).
2. A דומה למטריצה אלכסונית $A = PDP^{-1}$.
3. סכום הריבויים הגיאומטריים שווה לגודל המטריצה.
4. הפולינום המינימלי m_A מתפרק לגורמים ליניאריים שונים.

איך מלכסנים מטריצה?

1. מחשבים את הפולינום האופייני f_A.
2. מוצאים את השורשים שלו $\lambda_1, .., \lambda_k$ הע"ע.
3. לכל ע"ע λ_i מוצאים בסיס למרחב העצמי $V_{\lambda_i} = ker(A - \lambda_i I)$.

מצאנו אינווריאנטים חשובים של יחס הדמיון:

דרגה, דטרמיננטה, עקבה Tr, הפולינום האופייני, הפולינום המינימלי, הערכים העצמיים, הריבויים הגיאומטריים והאלגבריים.

8 מרחבי מכפלה פנימית

8.1 הגדרות

מרחב מכפלה פנימית (**ממ"פ**) ־ מרחב וקטורי עם מבנה נוסף ("דרך למדוד זויות וגדלים של וקטורים").

הגדרה 8.1 מרחב מכפלה פנימית ממשי

יהי V מ"ו מעל $\mathbb{R}$. **ממ"פ ממשי** הוא מ"ו V מעל $\mathbb{R}$ ביחד עם מכפלה פנימית,

שהיא פונקציה $\langle \cdot, \cdot \rangle : V^2 \to \mathbb{R}$ המקיימת את התכונות הבאות:

1. **סימטריה:** $\langle u,v\rangle=\langle v,u\rangle$
2. **ליניאריות** (במשתנה הראשון): $\langle u+u',v\rangle=\langle u,v\rangle+\langle u',v\rangle$
3. **הומוגניות** (במשתנה הראשון): $\forall\alpha\in\mathbb{R},\ \langle\alpha u,v\rangle=\alpha\langle u,v\rangle$
4. $\langle u,u\rangle\geq 0$, ושויון מתקיים אס"ם $u=0$.

הערה 8.2 עקב הסימטריה, תכונות 2,3 מתקיימות גם ביחס למשתנה השני (ב־$\mathbb{R}$...)

הגדרה 8.3 מרחב מכפלה פנימית מרוכב

ממ"פ מרוכב הוא מ"ו V מעל $\mathbb{C}$ ביחד עם פונקציה $\langle\cdot,\cdot\rangle : V^2\to\mathbb{C}$ המקיימת:

1. $\langle u,v\rangle=\overline{\langle v,u\rangle}$ (כלומר צמוד מרוכב: $z=x+iy\to\overline{z}=x-iy$)
2. $\langle\alpha u+\beta u',v\rangle=\alpha\langle u,v\rangle+\beta\langle u',v\rangle$
3. $\langle u,u\rangle$ הוא מספר ממשי לכל $u\in V$, ומתקיים $\langle u,u\rangle\geq 0$, ושויון אס"ם $u=0$.

הערה 8.4 מתוך תכונות 1,2 נובע:

$$\langle u,\alpha v+\beta v'\rangle=\overline{\langle\alpha v+\beta v',u\rangle}=\overline{\alpha\langle v,u\rangle+\beta\langle v',u\rangle}=\overline{\alpha}\overline{\langle v,u\rangle}+\overline{\beta}\overline{\langle v',u\rangle}=$$

$$=\overline{\alpha}\langle u,v\rangle+\overline{\beta}\langle u,v'\rangle$$

הערה 8.5 ממ"פ ממשי נקרא לפעמים **מרחב אוקלידי**.
ממ"פ מרוכב נקרא גם **מרחב אוניטארי**.

דוגמאות:

- ב־$\mathbb{R}^n$ המכפלה הפנימית הסטנדרטית: $\langle u,v\rangle=\sum_{i=1}^n u_iv_i$
- ב־$\mathbb{C}^n$ המכפלה הפנימית הסטנדרטית: $\langle u,v\rangle=\sum_{i=1}^n u_i\overline{v_i}$
- ניקח $V=C[0,1]$ מרחב הפונקציות הרציפות $f:[0,1]\to\mathbb{R}$, ונגדיר מ"פ:
  $\langle f,g\rangle=\int_0^1 f(x)g(x)\,dx$

8.1.1 תכונות

.1

$$\langle\sum_{i=1}^m\alpha_iu_i,v\rangle=\sum_{i=1}^m\alpha_i\langle u_i,v\rangle$$

.2

$$\langle u,\sum_{i=1}^m\alpha_iv_i\rangle=\sum_{i=1}^m\overline{\alpha}_i\langle u,v_i\rangle$$

.3

$$\langle\overline{0},u\rangle=\langle v,\overline{0}\rangle=0\in\mathbb{R},\mathbb{C}$$

הגדרה 8.6 נורמה **("המונח הפלצני לגודל של וקטור")**

הנורמה של וקטור $u\in V$ בממ"פ V ממשי או מרוכב מוגדרת ע"י:

$$\|u\|=\sqrt{\langle u,u\rangle}$$

תכונות הנורמה:

1. $\|u\|\geq 0$, וכן $\|u\|=0$ אם"ם $u=0$.
2. $\|\alpha u\|=|\alpha|\cdot\|u\|$. זה נכון כי: $\|\alpha u\|^2=\langle\alpha u,\alpha u\rangle=\alpha\overline{\alpha}\langle u,u\rangle=|\alpha|^2\|u\|^2$.[1]
3.

$$\|u+v\|^2=\langle u+v,u+v\rangle=\langle u,u\rangle+\langle u,v\rangle+\langle v,u\rangle+\langle v,v\rangle=$$

$$=\|u\|^2+\|v\|^2+\langle u,v\rangle+\langle v,u\rangle=\|u\|^2+\|v\|^2+2\Re e(\langle u,v\rangle)$$

[2]

8.2 קושי שוורץ וחברים.

8.2.1 אי שויון קושי־שוורץ

משפט 8.7 אם V ממ"פ (ממשי או מרוכב) אז מתקיים:

$$|\langle u,v\rangle|\leq\|u\|\cdot\|v\|$$

ומתקיים שויון אם"ם הוקטורים u,v הם תלויים ליניארית (באופן גרפי, הם מקבילים).

הוכחה: נוכיח את אי השויון:

יהי t מספר מרוכב כלשהו. נתבונן ב־$\|u-tv\|^2$.

$$\|u-tv\|^2=\langle u-tv,u-tv\rangle=\langle u,u\rangle-t\langle v,u\rangle-\overline{t}\langle u,v\rangle+|t|^2\langle v,v\rangle=$$

[1] נשים לב: אם $z=x+iy$, אז $z\overline{z}=(x+iy)(x-iy)=x^2+y^2=|z|^2$.

[2] $z+\overline{z}=(x+iy)+(x-iy)=2x=2\Re e(z)$.

$$= \|u\|^2 - t\langle v,u\rangle - \bar{t}\langle u,v\rangle + |t|^2\|v\|^2$$

נשים לב שהשתמשנו בנוסחה $t\bar{t} = |t|^2$. נשתמש בה גם בשלב הבא, אז לשים לב.

הנ"ל נכון לכל t, ובפרט עבור $t = \frac{\langle u,v\rangle}{\|v\|^2}$. (מניחים $v \neq 0$, אחרת אין מה להוכיח). ואז,

$$0 \leq \|u-tv\|^2 = \|u\|^2 - \frac{\langle u,v\rangle\langle v,u\rangle}{\|v\|^2} - \frac{\overline{\langle u,v\rangle}\langle u,v\rangle}{\|v\|^2} + \frac{|\langle u,v\rangle|^2}{\|v\|^4}\|v\|^2 = \|u\|^2 - \frac{|\langle u,v\rangle|^2}{\|v\|^2}$$

נעביר אגף ונכפול ב־$\|v\|^2$ ונקבל את הנדרש. נשים לב ששלושת המחוברים הימניים בביטוי שלפני האחרון הם שווים (בגלל ש $\overline{\langle u,v\rangle}\langle u,v\rangle = |\langle u,v\rangle|^2$), ולכן מקבלים את המעבר האחרון.

נוכיח שיש שויון אס"ם u,v תלויים ליניארית:

אם יש שויון ממש, אז עבור $t = \frac{\langle u,v\rangle}{\|v\|^2}$ ממקודם, מתקיים $\|u-tv\|^2 = 0$, כלומר $u = tv$.

(נזכור שנורמה=0 אס"ם הוקטור שבה הוא 0. וכמו כן t הוא סקלר, ולכן נקבל תלות ליניארית).

בכיוון השני, אם u,v ת"ל, אז $u = \alpha v$ עבור איזשהו $\alpha \in \mathbb{C}$ (או ש$v = 0$ ואז אין מה להוכיח).

ואז,

$$|\langle u,v\rangle| = |\langle \alpha v,v\rangle| = |\alpha|\cdot\langle v,v\rangle = |\alpha|\cdot\|v\|^2 = (|\alpha|\cdot\|v\|)\|v\| = \|\alpha v\|\cdot\|v\| = \|u\|\cdot\|v\|$$

וזהו. ■

מקרים פרטיים של קושי־שוורץ:

- ב־$\mathbb{R}^n$ עם המ"פ הסטנדרטית, מקבלים:

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

- ב־$C[0,1]$ מקבלים:

$$\int_0^1 f(x)g(x)dx \leq \sqrt{\int_0^1 f(x)^2dx}\cdot\sqrt{\int_0^1 g(x)^2dx}$$

8.2.2 מסקנה מקושי שוורץ: אי־שויון המשולש

טענה 8.8 בממ"פ V,

$$\|u+v\| \leq \|u\| + \|v\|$$

הוכחה: מתקיים (אי השויון הראשון הוא כי לכל $z \in \mathbb{C}$, $|\Re e(z)| \le |z|$).

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2 + 2\Re e(\langle u,v\rangle) \le \|u\|^2 + \|v\|^2 + 2|\langle u,v\rangle| \le$$

ונשתמש בקושי־שוורץ לקבל:

$$\le \|u\|^2 + \|v\|^2 + 2\|u\| \cdot \|v\| = (\|u\| + \|v\|)^2$$

נוציא שורש ונקבל את המבוקש. ∎

8.3 אורתוגונליות, קבוצות אורתוגונליות ואורתונורמליות

נתון V ממ"פ.

הגדרה 8.9 נגיד ש־ $u, v \in V$ **ניצבים** (או אורתוגונליים) אם $\langle u,v\rangle = 0$. ונסמן: $u \perp v$.
קבוצת וקטורים $u_1, ..., u_k \in V$ נקראת **קבוצה אורתוגונלית** אם $u_i \perp u_j$ לכל $1 \le i \ne j \le k$,
ונקראת **קבוצה אורתונורמלית** אם היא אורתוגונלית, וכן $\|u_i\| = 1$ לכל $1 \le i \le k$.

- בהנתן קבוצה אורתוגונלית של וקטורים $u_1, ..., u_k$ **שונים מ־0**, אפשר להפוך אותה לאורתונורמלית ע"י החלפת כל u_i ב־ $\frac{u_i}{\|u_i\|}$.

8.3.1 טענה לגבי המקדמים של וקטור שהוא צ"ל של קבוצה אורתונורמלית, ומסקנות

טענה 8.10 אם $\{u_1, ..., u_k\}$ קבוצה אורתונורמלית ונתון $v \in V$ כך ש־ $v = \sum_{i=1}^k a_i u_i$,
אז $a_i = \langle v, u_i\rangle$ לכל $1 \le i \le k$.

הוכחה: מתקיים:

$$\langle v, u_i\rangle = \langle \sum_{j=1}^k a_j u_j, u_i\rangle = \sum_{j=1}^k a_j \langle u_j, u_i\rangle = \sum_{j=1}^k a_j \cdot \delta_{ij} = a_i$$

∎

מסקנה 8.11 אם $u_1, ..., u_k \in V$ קבוצה אורתונורמלית, אז היא בלתי תלויה ליניארית.

הוכחה: אם $\sum_{i=1}^k a_i u_i = 0$, אז $a_i = \langle 0, u_i\rangle = 0$ לכל $1 \le i \le k$. ∎

מסקנה 8.12 אם $V \in span\{u_1, ..., u_k\}$ (קב' א"נ), אז $\|v\|^2 = \sum_{i=1}^k |\langle v, u_i\rangle|^2$

הוכחה: כי:

$$\|v\|^2 = \langle \sum_{i=1}^k a_i u_i, \sum_{j=1}^k a_j u_j\rangle = \sum_{i=1}^k \sum_{j=1}^k a_i \overline{a}_j \langle u_i, u_j\rangle = \sum_{i=1}^k \sum_{j=1}^k a_i \overline{a}_j \cdot \delta_{ij} = \sum_{i=1}^k |a_i|^2$$

∎

8.3.2 משפט פיתגורס

אם $u \perp v$ בממ"פ V, אז $\|u+v\|^2 = \|u\|^2 + \|v\|^2$.

(באופן כללי הראינו ש־ $\|u+v\|^2 = \|u\|^2 + \|v\|^2 + 2\Re e(\langle u,v\rangle)$, **אבל אם** $u \perp v$ **אז** $\langle u,v\rangle = 0$**)**.

8.3.3 תהליך האורתוגונליזציה של גראם־שמידט

אנו מעוניינים במציאת **בסיס אורתונורמלי**, כלומר בסיס שהוא קבוצה אורתונורמלית. האם תמיד קיים בסיס א"נ, ואיך מוצאים אחד כזה?

<u>תהליך האורתוגונליזציה של גראם־שמידט:</u>

בהנתן ממ"פ V וקבוצה בת"ל $v_1, ..., v_m$, נבנה קבוצת וקטורים $u_1, ..., u_m$ כך שמתקיים:

- $u_1, ..., u_m$ קבוצה אורתונורמלית.
- לכל $1 \le k \le m$, $span\{u_1, ..., u_k\} = span\{v_1, ..., v_k\}$.

הבניה היא באינדוקציה:

ראשית מגדירים $w_1 = v_1\ ,\ u_1 = \frac{w_1}{\|w_1\|}$.

בשלב הבא, נגדיר:

$$w_2 = v_2 - \langle v_2, u_1\rangle \cdot u_1 \qquad\qquad u_2 = \frac{w_2}{\|w_2\|}$$

ולאחר מכן,

$$w_3 = v_3 - \langle v_3, u_1\rangle \cdot u_1 - \langle v_3, u_2\rangle \cdot u_2 \qquad\qquad u_3 = \frac{w_3}{\|w_3\|}$$

$$\vdots$$

ובאופן כללי, לכל $1 \le k \le m$, אם הגדרנו את $u_1, ..., u_{k-1}$ ואת $w_1, ..., w_{k-1}$, נגדיר u_k, w_k ע"י:

$$w_k = v_k - \sum_{i=1}^{k-1} \langle v_k, u_i\rangle u_i \qquad\qquad u_k = \frac{w_k}{\|w_k\|}$$

ונשים לב ש־$w_k \neq 0$ בגלל האי־תלות הליניארית.

(אם נניח בשלילה ש־$w_k = 0$, ונעביר אגף, נקבל: $v_k = \sum_{i=1}^{k-1} \langle v_k, u_i\rangle u_i$ כלומר ש־v_k תלוי ליניארית ב־$u_1, ..., u_{k-1}$. אבל $\{u_i\}_{i=1}^{k-1}$ הם צ"ל של $v_1, ..., v_{k-1}$, אבל $\{v_i\}$ הם בת"ל ולכן v_k לא תלוי ליניארית בשאר ה־v_i. ולכן נקבל סתירה).

הוכחה: בדיקה שזה עובד:

עבור $k = 1$ הטענה מתקיימת. ($\{u_1\}$ היא קב' א"נ, וברור ש $span\{u_1\} = span\{v_1\}$).

עבור $k = 2$: ראשית, $w_2 \neq 0$ בגלל האי־תלות, ולכן u_2 מוגדר, ומתקיים:

$$\langle w_2, u_1\rangle = \langle v_2 - \langle v_2, u_1\rangle u_1, u_1\rangle = \langle v_2, u_1\rangle - \langle\langle v_2, u_1\rangle u_1, u_1\rangle =$$

$$= \langle v_2, u_1\rangle - \langle v_2, u_1\rangle \cdot \langle u_1, u_1\rangle = 0$$

(נשים לב ש־$\langle u_1, u_1\rangle = 1$ כי $\|u_1\| = 1$).

ולכן גם $\langle u_2, u_1\rangle = \langle \frac{w_2}{\|w_2\|}, u_1\rangle = 0$.

כמו כן, $span\{u_1, u_2\} \subseteq span\{v_1, v_2\}$ כי u_1, u_2 הוגדרו כצ"ל של v_1, v_2.

אבל $\{u_1, u_2\}$ קבוצה אורתונורמלית, ולכן בת"ל, ולכן $dim\, span\{u_1, u_2\} = 2 = dim\, span\{v_1, v_2\}$ ולכן שני ה־$span$־ים שווים.

- **שלב האינדוקציה הכללי**: נניח שהטענות מתקיימות עבור $\begin{matrix} u_1, ..., u_{k-1} \\ w_k, ..., w_{k-1} \end{matrix}$ ונוכיח עבור k.

$w_k = v_k - \sum_{i=1}^{k-1} \langle v_k, u_i\rangle u_i$ אינו 0, כי לפי הנחת האינדוקציה $\sum_{i=1}^{k-1} \langle v_k, u_i\rangle u_i \in span\{v_1, ..., v_{k-1}\}$,

אבל $v_k \notin span\{v_1, ..., v_{k-1}\}$ לפי הנחת האי־תלות של ה־v_i־ים.

נחשב את $\langle u_k, u_j\rangle$ עבור $1 \le j \le k-1$. ראשית,

$$\langle w_k, u_j\rangle = \langle v_k - \sum_{i=1}^{k-1} \langle v_k, u_i\rangle u_i, u_j\rangle = \langle v_k, u_j\rangle - \sum_{i=1}^{k-1} \langle v_k, u_i\rangle \langle u_i, u_j\rangle =$$

ולפי הנחת האינדוקציה, זה שווה ל־

$$= \langle v_k, u_j\rangle - \langle v_k, u_j\rangle = 0$$

(כי בסכום שאנו מחסירים, כל האיברים שבהם $i \neq j$ התאפסו, ובאיבר שבו $i = j$, נקבל $\langle v_k, u_j\rangle \langle u_j, u_j\rangle = \langle v_k, u_i\rangle \cdot 1$)

ושנית, מתקיים $\langle u_k, u_i\rangle = \langle \frac{w_k}{\|w_k\|}, u_i\rangle = 0$. וכמו כן, $\|u_k\| = \|\frac{w_k}{\|w_k\|}\| = 1$, ולכן $u_1, ..., u_k$ קב' אורתונורמלית.

ולבסוף, $span\{u_1, ...u_k\} \subseteq span\{v_1, ..., v_k\}$ מכיוון ש $u_1, ..., u_{k-1} \in span\{v_1, ..., v_k\}$ לפי הנחת האינדוקציה, וכמו כן $u_k \in span\{v_k, u_1, ..., u_{k-1}\} \subseteq span\{v_1, ..., v_k\}$ (שוב לפי הנחת האינדוקציה).

אבל משיקולי מימד יש למעשה שויון בין שני המרחבים.

($u_1, ..., u_k$ קבוצה אורתונורמלית, ולכן $dim\, span\{u_1, ..., u_k\} = k$) ■

8.3.4 דוגמא לשימוש בגראם־שמידט

ב־$\mathbb{R}^3$ עם המ"פ הסטנדרטית נבצע את תהליך גראם־שמידט,

על הוקטורים $v_1 = (1, -1, 1),\ v_2 = (1, 0, 1),\ v_2 = (1, 1, 2)$. נחשב:

$$w_1 = v_1 = (1, -1, 1)$$

$$u_1 = \frac{w_1}{\|w_1\|} = \frac{1}{\sqrt{1^2+1^2+1^2}} w_1 = \frac{1}{\sqrt{3}}(1,-1,1)$$

$$w_2 = v_2 - \langle v_2, u_1 \rangle u_1 = (1,0,1) - \langle (1,0,1), \frac{1}{\sqrt{3}}(1,-1,1) \rangle \cdot \frac{1}{\sqrt{3}}(1,-1,1) = ... = (\frac{1}{3}, \frac{2}{3}, \frac{1}{3})$$

$$u_2 = \frac{w_2}{\|w_2\|} = \frac{1}{\sqrt{(\frac{1}{3})^2 + (\frac{2}{3})^2 + (\frac{1}{3})^2}} (\frac{1}{3}, \frac{2}{3}, \frac{1}{3}) = \frac{1}{\sqrt{6}}(1,2,1)$$

$$w_3 = v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2 = ... = (-\frac{1}{2}, 0, \frac{1}{2})$$

$$u_3 = \frac{w_3}{\sqrt{(\frac{1}{2})^2 + (\frac{1}{2})^2}} = (\frac{-\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2})$$

וזהו.

8.3.5 משלים אורתוגונלי וסכומים ישרים

הגדרה 8.13 אם V מ"ו, $U, W \subseteq V$ תתי־מרחבים, אז $U + W = span(U \cup W)$ הוא המרחב הנפרש ע"י U, W.

אומרים ש־V הוא **סכום ישר** של U, W (סימון: $V = U \oplus W$),

אם $V = U + W$ וגם $U \cap W = \{0\}$.

טענה 8.14 $V = U \oplus W$ אם"ם כל וקטור $v \in V$ ניתן להצגה **יחידה** בצורה $v = u + w$ כאשר $u \in U, w \in W$.

הוכחה: כיוון א': יש הצגה כזו כי זה סכום רגיל. כדי להראות יחידות, נניח: $u + w = u' + w'$

ואז נקבל ש $u - u' = w - w'$. אבל החיתוך הוא רק $\overline{0}$, ולכן $u = u'$, $w = w'$.

כיוון ב': הושאר כתרגיל. ■

טענה 8.15 אם $V = U \oplus W$ אז $dim\, V = dim\, U + dim\, W$.

הוכחה: הושארה כתרגיל. ■

הגדרה 8.16 משלים אורתוגונלי

יהי V ממ"פ ויהי $W \subseteq V$ ת"מ.

המשלים האורתוגונלי של W, שיסומן ב־ $W^\perp$, הוא קבוצת הוקטורים $v \in V$ שניצבים לכל $w \in W$:

$$W^\perp = \{v \in V \ : \forall w \in W \ w \perp v\}$$

תכונות:

- $W^\perp$ הוא תת־מרחב ליניארי:
  אם $u, u' \in W^\perp$ אז לכל $w \in W$,
  $$\langle u+u', w\rangle = \langle u, w\rangle + \langle u', w\rangle = 0+0=0$$
  ולכן $u+u' \in W^\perp$. ובאופן דומה, $\alpha u \in W^\perp$ לכל $\alpha \in \mathbb{F}$.

טענה 8.17 $V = W \oplus W^\perp$ אם V ממ"פ **ממימד סופי.** (כאשר $W \subseteq V$ ת"מ).

הוכחה: ניקח בסיס אורתונורמלי $u_1, ..., u_k$ של W (קיים לפי גראם־שמידט).

אם $v \in V$ אז נגדיר $v' = \sum_{i=1}^k \langle v, u_i\rangle u_i$, וגם $v'' = v - v'$.

אז קיבלנו הצגה $v = v' + v''$.

$v' \in W$ כצירוף ליניארי של $u_1, ...u_k$ (נשים לב שהממ"פ בהגדרת v' היא סקלר בלבד).

וכמו כן, לכל $1 \leq i \leq k$,

$$\langle v'', u_i\rangle = \langle v - v', u_i\rangle = \langle v - \sum_{j=1}^k \langle v, u_j\rangle u_j, u_i\rangle =$$

$$= \langle v, u_i\rangle - \sum_{i=1}^k \langle v, u_j\rangle \langle u_j, u_i\rangle = \langle v, u_i\rangle - \langle v, u_i\rangle = 0$$

ולכן $v'' \in W^\perp$ והראינו ש $V = W + W^\perp$.

הערה: נשים לב שזה זהה לחלוטין לבניה שעשינו ולהוכחה של גראם־שמידט.

לבסוף, אם $u \in W \cap W^\perp$, אז $u \perp u$ ולכן $\langle u, u\rangle = 0$ ולכן $u = 0$. ולכן הסכום הוא ישר. ■

מסקנה 8.18 אם Vממ"פ, $U \subseteq V$ תת־מרחב, אז $(U^\perp)^\perp = U$.

הוכחה: כיוון א': אם $u \in U$ אז לכל $v \in U^\perp$ מתקיים $u \perp v$ ולכן $u \in (U^\perp)^\perp$.

כיוון ב': אם $v \in (U^\perp)^\perp$,

לפי הטענה שראינו קודם קיימים u, u' כך ש $u \in U$, $u' \in U^\perp$, ומתקיים $v = u + u'$.

מכיוון ש־ $v \in (U^\perp)^\perp$, מתקיים $v \perp u'$.

כמו כן, $u \in U$ ולכן $u \perp u'$.

ולכן גם $v - u = u' \perp u'$. (כי $\langle v-u, u'\rangle = \langle v, u\rangle - \langle u, u'\rangle = 0$).

קיבלנו ש u' ניצב לעצמו, ולכן $u' = 0$. כלומר $v = u \in U$. ■

8.4 פונקציונלים ליניאריים

הגדרה 8.19 יהי V מ"ו מעל שדה $\mathbb{F}$.

פונקציונל ליניארי הוא העתקה $\varphi : V \to \mathbb{F}$.

סימון: $Hom(V,W)$ - אוסף ההעתקות הליניאריות ממ"ו V למ"ו W.

$$dim\, Hom(V,W) = dimV \cdot dimW$$

סימון: $V^* = Hom(V,\mathbb{F})$ - המרחב הדואלי של V - אוסף הפונקציונלים הליניאריים על V.

בפרט, מתקיים $dimV^* = dimV$ (כי הרי $dim\mathbb{F} = 1$).

כעת, נניח ש-V ממ"פ מעל $\mathbb{F}$ (כאשר $\mathbb{F} = \mathbb{R}$ או $\mathbb{F} = \mathbb{C}$).

לכל וקטור $y \in V$ אפשר להגדיר **פונקציונל** $\varphi = \varphi_y \in V^*$ ע"י:

$$\varphi_y(u) = \langle u, y\rangle$$

זהו פונקציונל ליניארי:

$$\varphi_y(u+u') = \langle u+u', y\rangle = \langle u,y\rangle + \langle u', y\rangle = \varphi_y(u) + \varphi_y(u')$$

ובאופן דומה,

$$\alpha \in \mathbb{F}, \qquad \varphi_y(\alpha u) = \langle \alpha u, y\rangle = \alpha\langle u,y\rangle = \alpha\varphi_y(u)$$

משפט 8.20 בממ"פ ממימד סופי V,

לכל פונקציונל $\varphi \in V^*$ קיים וקטור **יחיד** $y \in V$ כך ש $\varphi(v) = \varphi_y(v)$ **לכל** $v \in V$.

הוכחה: ראשית נראה **יחידות**: אם $y, y' \in V$ כך שלכל $v \in V$ מתקיים $\varphi_y(v) = \varphi_{y'}(v)$,

כלומר $\langle v,y\rangle = \langle v,y'\rangle$, אז $\langle v, y-y'\rangle = 0$. זה נכון לכל $v \in V$, ובפרט עבור $v = y - y'$.

ומכאן, $y-y' \perp y-y'$, ולכן $y - y' = 0 \Leftrightarrow y = y'$.

הוכחת הקיום:

ניתנו בהרצאה שתי הוכחות. הראשונה מעל $\mathbb{R}$ בלבד ("$abstract\, nonsense$" לפי רומיק), והשניה כללית. אכלול כאן רק את ההוכחה הכללית. כלומר זאת שתקפה גם מעל $\mathbb{C}$.

הקדמה להוכחה: נשים לב שעבור φ_y מתקיים: $Ker\,\varphi_y = (span\{y\})^\perp$. (כי $v \in Ker\,\varphi_y \;\Leftrightarrow\; \varphi_y(v) = \langle v,y\rangle = 0$). בעזרת אבחנה זו, אפשר לשחזר את y בהנתן φ_y.

באופן פורמלי: נניח נתון $\varphi \in V^*$.

מתקיים: $dim\,Im\varphi + dim\,Ker\varphi = dim\,V = n$ (כאשר סימנו את מספר זה ב-n).

אבל, $Im\,\varphi \subseteq \mathbb{F}$ ת"מ, ולכן $dim\,Im\varphi = \begin{cases} 0 \\ -or- \\ 1 \end{cases}$. (כי $dim\mathbb{F} = 1$).

אם $dim\,Im\varphi = 0$ אז $\varphi(v) = 0$ לכל $v \in V$, כלומר $\varphi \equiv 0$ פונקציונל האפס. במקרה זה, $\varphi = \varphi_0$.

במקרה השני, $dim\,Im\varphi = 1$. מכאן, $dim\,Ker\varphi = n - 1$.

ולכן גם $dim\,(Ker\varphi)^{\perp} = 1$.

זהו תת מרחב ממימד 1, ולכן נפרש ע"י וקטור אחד שנסמנו ב־z.

כלומר, $(Ker\varphi)^{\perp} = span\{z\}$. כעת, נסמן: (נשים גם לב ש $\frac{\overline{\varphi(z)}}{\|z\|^2}$ הוא סקלר, ו־z הוא כמובן וקטור).

$$y = \frac{\overline{\varphi(z)}}{\|z\|^2} \cdot z$$

ונבדוק ש־ $\varphi = \varphi_y$.

אנו יודעים ש־ $V = Ker\varphi \oplus span\{z\}$, ולכן אם $v \in V$ כלשהו, אז קיימת לו הצגה בצורה $v = u + tz$, כאשר מתקיים ש $\varphi(u) = 0$, וגם $t \in \mathbb{F}$.

נחשב את $\varphi(v)$ ואת $\varphi_y(v)$ ונראה שהם שווים:

$$\varphi(v) = \varphi(u) + t\varphi(z) = 0 + t\varphi(z) = t\varphi(z)$$

ומצד שני, (נשים לב שלמטה מתקיים $\langle u, y\rangle = 0$ כי $y \in span\{z\}$ ולכן ניצב ל־$Ker\varphi$).

$$\varphi_y(v) = \langle v, y\rangle = \langle u + tz, y\rangle = \langle u, y\rangle + t\langle z, y\rangle = 0 + t\langle z, \frac{\overline{\varphi(z)}}{\|z\|^2} z\rangle =$$

$$= t \cdot \overline{\frac{\overline{\varphi(z)}}{\|z\|^2}} \cdot \langle z, z\rangle = t \cdot \frac{\varphi(z)}{\|z\|^2} \cdot \|z\|^2 = t\varphi(z)$$

הראינו, אם כן, ש־$\varphi(v) = \varphi_y(v)$ לכל $v \in V$, ולכן $\varphi = \varphi_y$. ■

9 העתקות ליניאריות וההעתקה הצמודה

9.1 ההעתקה הצמודה

9.1.1 הגדרה והוכחת הקיום\יחידות

יהי V ממ"פ מעל $\mathbb{F}$ ($\mathbb{R}$ או $\mathbb{C}$), ותהי $T : V \to V$ העתקה ליניארית.

משפט 9.1 (הגדרה)

קיימת ה"ל **יחידה** $T^* : V \to V$ שתיקרא **ההעתקה הצמודה** ל־T, כך שלכל $v, w \in V$ מתקיים:

$$\langle T(v), w\rangle = \langle v, T^*(w)\rangle$$

הוכחה: נוכיח **יחידות:**

ראשית, אם עבור w נתון קיימים שני וקטורים u, u' כך ש $\langle T(v), w\rangle = \langle v, u\rangle = \langle v, u'\rangle$

לכל $v \in V$, אז מכאן, כמו שכבר ראינו (בהוכחה הקודמת), $u = u'$.

מכאן מקבלים יחידות: אם יש T_1^*, T_2^*, ניקח $u = T_1^*,\ u' = T_2^*$.

נוכיח **קיום**:

עבור $w \in V$, נגדיר פונקציונל $\psi_w : V \to \mathbb{F}$ ("זה פסִי. תרגעו.") ע"י הנוסחה:

$$\psi_w(v) = \langle T(v), w\rangle$$

זהו אכן פונקציונל ליניארי, כי $\psi_w = \varphi_w \circ T$ והרכבת ה"ל היא ה"ל (כנראה?).

לפי המשפט הקודם, קיים וקטור u כך ש $\psi_w = \varphi_u$.

נסמן $u = T^*(w)$. במילים אחרות, מתקיים לכל $v \in V$,

$$\langle T(v), w\rangle = \psi_w(v) = \varphi_{T^*(w)}(v) = \langle v, T^*(w)\rangle$$

נותר לבדוק ש־T^* היא העתקה ליניארית.

נשים לב שמתקיים (זה נובע מכך ש $\psi_w = \varphi_w \circ T$):

$$\begin{cases} \psi_{w+w'} = \psi_w + \psi_{w'} \\ \psi_{\alpha w} = \overline{\alpha}\psi_w \end{cases}$$

כעת, **נסמן:** אם $\varphi = \varphi_y$ אז נסמן $y = y_\varphi$.

ניתן לבדוק שבאופן דומה (הושאר כתרגיל...), מתקיים

$$\begin{cases} y_{\varphi+\varphi'} = y_\varphi + y_{\varphi'} \\ y_{\alpha\varphi} = \overline{\alpha}y_\varphi \end{cases}$$

ולבסוף, מתקיים:

$$T^*(w+w') = y_{\psi_{w+w'}} = y_{\psi_w+\psi_{w'}} = y_{\psi_w} + y_{\psi_{w'}} = T^*(w) + T^*(w')$$

ובאופן דומה,

$$T^*(\alpha w) = y_{\psi_{\alpha w}} = y_{\overline{\alpha}\psi_w} = \overline{\overline{\alpha}}y_{\psi_w} = \alpha y_{\psi_w} = \alpha T^*(w)$$

ולכן T^* ליניארית. ■

9.1.2 תכונות ההעתקה הצמודה

טענה 9.2 יהי V ממ"פ, $T,S:V\to V$ ה"ל. אז:

1.

$$(T+S)^*=T^*+S^*$$

2.

$$(\alpha T)^*=\overline{\alpha}T^*$$

3.

$$(T\circ S)^*=S^*\circ T^*$$

4.

$$(T^*)^*=T$$

הוכחה: אם $v,w\in V$ נתונים, אז:

1) מתקיים:

$$\langle (T+S)v,w\rangle=\langle Tv,w\rangle+\langle Sv,w\rangle=\langle v,T^*w\rangle+\langle v,S^*w\rangle=\langle v,(T^*+S^*)w\rangle$$

ולכן $(T+S)^*=T^*+S^*$.

2) מתקיים:

$$\langle \alpha Tv,w\rangle=\alpha\langle Tv,w\rangle=\alpha\langle v,T^*w\rangle=\langle v,\overline{\alpha}T^*w\rangle$$

ולכן $(\alpha T)^*=\overline{\alpha}T^*$.

3) מתקיים:

$$\langle TSv,w\rangle=\langle Sv,T^*w\rangle=\langle v,S^*T^*w\rangle$$

זה נכון לכל v,w ולכן $(TS)^*=S^*T^*$.

4) מתקיים:

$$\langle T^*v,w\rangle=\overline{\langle w,T^*v\rangle}=\overline{\langle Tw,v\rangle}=\overline{\overline{\langle v,Tw\rangle}}=\langle v,Tw\rangle$$

ולכן $(T^*)^*=T$. ■

9.1.3 מטריצה צמודה

הגדרה 9.3 אם $A \in M_n(\mathbb{C})$, נסמן $A^* = \overline{A^t}$, **המטריצה הצמודה** ל־A.
(מעל $\mathbb{R}$ מתקיים פשוט $A^* = A^t$).

טענה 9.4 יהי Vממ"פ. $T : V \to V$ ה"ל, ונתון **בסיס אורתונורמלי** $B = \{v_1, ..., v_n\}$.
אז אם $[T]_B = (a_{ij})_{i,j=1}^n$, אז $[T^*] = (\overline{a}_{ji})_{i,j=1}^n$.
כלומר, $[T^*]_B = (\overline{[T]_B})^t$.

הוכחה: $[T]_B = (a_{ij})_{i,j=1}^n$, כלומר $Tv_j = \sum_{i=1}^n a_{ij}v_i$ לכל $1 \leq j \leq n$. (זוהי העמודה במטריצת הייצוג).
מכאן נקבל: (משתמשים בטענה שראינו לגבי מקדמים של צ"ל של בסיס א"נ. כנראה.):

$$a_{ij} = \langle Tv_j, v_i \rangle = \langle v_j, T^*v_i \rangle = \overline{\langle T^*v_i, v_j \rangle}$$

ואם נסמן: $[T^*]_B = (b_{ij})_{i,j=1}^n$ אז משמעות הדבר היא שמתקיים:

$$T^*v_j = \sum_{i=1}^n b_{ij}v_i$$

ואם נחליף את i, j, נקבל:

$$T^*v_i = \sum_{j=1}^n b_{ji}v_j$$

וזה נותן: $b_{ji} = \langle T^*v_i, v_j \rangle = \overline{a}_{ij}$, שזה מה שרצינו להראות. (עוד פעם השתמשנו בטענה ההיא). ■

9.2 העתקות צמודות לעצמן: הרמיטיות ($\mathbb{C}$), סימטריות ($\mathbb{R}$)

9.2.1 הגדרה

הגדרה 9.5 העתקה ליניארית $T : V \to V$ (V ממ"פ) נקראת **צמודה לעצמה** אם $T = T^*$.
מעל $\mathbb{C}$, העתקה צמודה לעצמה נקראת גם **הרמיטית** (ע"ש שארל הרמיט).
מעל $\mathbb{R}$, העתקה צמודה לעצמה נקראת גם **סימטרית**.
מטריצה $A \in M_n(\mathbb{C})$ נקראת צמודה לעצמה אם $A^* = A$.
מעל $\mathbb{C}$ (מטריצה הרמיטית), $A^t = \overline{A}$.
מעל $\mathbb{R}$ (מטריצה סימטרית), $A = A^t$.

הערה 9.6 T צמודה לעצמה $\Leftrightarrow$ $\langle Tv, w \rangle = \langle v, Tw \rangle \quad \forall v, w \in V$

9.2.2 תכונות

טענה 9.7 תהי $T:V\to V$ העתקה ליניארית. אז:

1. $T+T^*$ צמודה לעצמה
2. $T\circ T^*$ צמודה לעצמה
3. אם V ממ"פ מעל $\mathbb{C}$, אז $\frac{T-T^*}{2i}$ הרמיטית.
4. אם T צמודה לעצמה, ו $\alpha\in\mathbb{R}$ (ממשי!!), אז αT צמודה לעצמה.
5. מעל $\mathbb{C}$, אם T כלשהי אז קיימות העתקות ליניאריות $R,S:V\to V$ כך ש R,S הרמיטיות, ומתקיים

$$T=R+iS \qquad , \qquad T^*=R-iS$$

הוכחה:

1) מתקיים:

$$(T+T^*)^*=T^*+T^{*^*}=T^*+T$$

2) מתקיים:

$$(TT^*)^*=T^{*^*}T^*=TT^*$$

3) מתקיים:

$$(\frac{1}{2i}(T-T^*))^*=\frac{-1}{2i}(T^*-T^{*^*})=\frac{-1}{2i}(T^*-T)=\frac{1}{2i}(T-T^*)$$

4) אם $T^*=T$ ו־$\alpha\in\mathbb{R}$ אז ($\alpha=\overline{\alpha}$ כי הוא ממשי):

$$(\alpha T)^*=\overline{\alpha}T^*=\alpha T^*$$

5) נגדיר: $R=\frac{1}{2}(T+T^*)$, $S=\frac{1}{2i}(T-T^*)$. אז R,S הרמיטיות ומתקיים:

$$R+iS=\frac{1}{2}(T+T^*)+i\frac{1}{2i}(T-T^*)=T$$

$$R-iS=\frac{1}{2}(T+T^*)-i\frac{1}{2i}(T-T^*)=T^*$$

■

אנלוגיה מעניינת (או שלא) **מעל** $\mathbb{C}$:

העתקות ליניאריות $\leftrightarrow$ מספרים מרוכבים

העתקות הרמיטיות $\leftrightarrow$ מספרים ממשיים

למה? כי, עבור R,S הרמיטיות ניתן לראות את ההקבלות הבאות:

$T=R+iS$	$z=x+iy$
$T^*=R-iS$	$\overline{z}=x-iy$
$R=\frac{T+T^*}{2}$	$\Re e(z)=x=\frac{z+\overline{z}}{2}$
$S=\frac{T-T^*}{2i}$	$\Im m(z)=y=\frac{z-\overline{z}}{2i}$

9.2.3 טענה, שאומרת שמעל $\mathbb{C}$, T הרמיטית $\Leftrightarrow \langle Tv,v\rangle \in \mathbb{R} \quad \forall v\in V$

טענה 9.8 T הרמיטית (בממ"פ מעל $\mathbb{C}$) אם"ם לכל $v\in V$ מתקיים $\langle Tv,v\rangle \in \mathbb{R}$.

הוכחה: כיוון א': נניח T הרמיטית. אז,

$$\langle Tv,v\rangle = \langle v,Tv\rangle = \overline{\langle Tv,v\rangle}$$

וכעת, מספר הצמוד לעצמו הוא ממשי, ולכן $\langle Tv,v\rangle \in \mathbb{R}$.
כיוון ב':
אם לכל $v\in V$, $\langle Tv,v\rangle \in \mathbb{R}$, אז (המעבר הימני ביותר הוא ההנחה):

$$\langle v,T^*v\rangle = \langle Tv,v\rangle = \overline{\langle v,Tv\rangle} = \langle v,Tv\rangle$$

כלומר מתקיים $\langle (T-T^*)v,v\rangle = 0$ לכל v, ולכן $T-T^*=0$, לפי הלמה הבאה. ■

9.2.4 למה: תנאים שקולים להיות העתקה זהותית 0, מעל $\mathbb{C}$

למה 9.9 בממ"פ מעל $\mathbb{C}$, התנאים הבאים שקולים:

1. $\langle Tv,v\rangle = 0$ לכל v. (כלומר $Tv\perp v$)
2. $\langle Tv,w\rangle = 0$ לכל v,w.
3. $T=0$.

הערה 9.10 מעל $\mathbb{R}$ זה לא נכון. למשל, ב־$\mathbb{R}^2$, ההעתקה $T\begin{pmatrix} x\\ y\end{pmatrix} = \begin{pmatrix} -y\\ x\end{pmatrix}$ (סיבוב ב־90°) מקיימת $Tv\perp v$ לכל v, אבל אינה העתקת האפס.

הוכחה: • $1 \Leftarrow 2$: נניח ש $\langle Tv,v\rangle = 0$ לכל v.
אם $v,w\in V$ אז: (נשים לב ש $\langle Tv,v\rangle = \langle Tw,w\rangle = 0$ מההנחה. וגם השויון השמאלי ביותר מההנחה).

$$0=\langle T(v+w),v+w\rangle = \langle Tv,v\rangle + \langle Tv,w\rangle + \langle Tw,v\rangle + \langle Tw,w\rangle = \langle Tv,w\rangle + \langle Tw,v\rangle$$

וכעת, מתקיים גם:

$$0=\langle T(iv+w),iv+w\rangle = i^2\langle Tv,v\rangle + \langle Tw,w\rangle + i\langle Tv,w\rangle - i\langle Tw,v\rangle =$$

$$=i(\langle Tv,w\rangle - \langle Tw,v\rangle)$$

ולכן $\langle Tv,w\rangle - \langle Tw,v\rangle = 0$. בצירוף עם התוצאה מקודם, שאומרת $\langle Tv,w\rangle + \langle Tw,v\rangle = 0$,
נקבל ש $\langle Tv,w\rangle = \langle Tw,v\rangle = 0$.
• $2 \Leftarrow 3$: נניח ש $\langle Tv,w\rangle = 0$ לכל v,w.
יהי $v\in V$. אז זה נכון בפרט עבור $w=Tv$, ומכאן, $\langle Tv,Tv\rangle = 0$ כלומר $\|Tv\|^2=0$ ולכן $T=0$.
• $3 \Leftarrow 1$ ברור. ■

9.2.5 קצת על העתקות אנטי־הרמיטיות, וטענה על יחידות הפיתוח $T = R + iS$ (R, S הרמיטיות).

הגדרה 9.11 T נקראת **אנטי־הרמיטית** אם $T^* = -T$.

תכונות:

1. T הרמיטית $\Leftrightarrow$ iT אנטי־הרמיטית.
2. אם T הרמיטית וגם אנטי־הרמיטית אז $T = 0$.

הוכחה: 1 הושאר כתרגיל.

בדיקה ל־2: מתקיים $T = T^* = -T$ ולכן, $T = 0$. ■

טענה 9.12 בפיתוח $T = R + iS$ כאשר R, S הרמיטיות, R, S נקבעות **ביחידות**.

הוכחה: נניח $T = R + iS = R' + iS'$ כאשר R, R', S, S' הרמיטיות. אז:

$$R - R' = i(S - S')$$

כעת, $R - R'$ הרמיטית, וגם $S - S'$. אבל ראינו שאז זה אומר ש $i(S - S')$ אנטי הרמיטית.
קיבלנו העתקה הרמיטית ששווה לאנטי הרמיטית, אז נשתמש בטענה הקודמת ונקבל שאז,
$i(S - S') = R - R' = 0$, כלומר $S - S'$ וגם $R = R'$. ■

9.2.6 למה: תנאים שקולים להיות העתקה *סימטרית* זהותית 0, מעל $\mathbb{R}$

נשים לב שאלו התנאים שהוכחנו קודם עבור $\mathbb{C}$ להעתקה כללית ־ הפעם אנו מעל $\mathbb{R}$ וההעתקה סימטרית.

למה 9.13 תהי $T : V \to V$ העתקה **סימטרית**. אז התנאים הבאים שקולים:

1. $\langle Tv, v\rangle = 0$ לכל $v \in V$.
2. $\langle Tv, w\rangle = 0$ לכל $v, w \in V$.
3. $T = 0$.

הוכחה: $3 \Leftarrow 2$ ו־ $1 \Leftarrow 3$, כמו בהוכחה הקודמת (שעשינו מעל $\mathbb{C}$).

נוכיח $2 \Leftarrow 1$:

אנו מניחים ש $\langle Tv, v\rangle = 0$ לכל $v \in V$. אז,

$$0 = \langle T(v+w), v+w\rangle = \langle Tv, v\rangle + \langle Tw, w\rangle + \langle Tv, w\rangle + \langle Tw, v\rangle =$$

$$= 0 + 0 + \langle v, T^*w\rangle + \langle Tw, v\rangle == \overline{\langle T^*w, v\rangle} + \langle Tw, v\rangle = \langle Tw, v\rangle + \langle Tw, v\rangle = 2\langle Tw, v\rangle$$

נשים לב שהשתמשנו בכך שמעל $\mathbb{R}$ הצמוד שווה ללא־צמוד, וגם בכך ש $T = T^*$ כי היא סימטרית.
(כנראה ־ שיפצתי קצת את ההוכחה הזאת, היא היתה פחות מפורטת בהרצאה).
ולכן $\langle Tv, w\rangle = 0$. ■

9.3 העתקות אוניטאריות: אוניטאריות ($\mathbb{C}$), אורתוגונליות ($\mathbb{R}$)

9.3.1 הגדרה

הגדרה 9.14 נגיד ש־ T **אוניטארית** אם T שומרת על מכפלה פנימית.
זאת אומרת, $\langle Tv, Tw\rangle = \langle v, w\rangle$ לכל $v, w \in V$.

הגדרה 9.15 מטריצה אוניטארית

נגיד ש־ $A \in M_n(\mathbb{F})$ (כאשר $\mathbb{F}$ הוא $\mathbb{R}$ או $\mathbb{C}$) היא אוניטארית, אם $A^*A = I$.

(זה מסתדר עם ההגדרה להעתקות, בגלל המשפט הבא:)

9.3.2 תנאים שקולים להיות T אוניטארית

משפט 9.16 בממ"פ V סוף־מימדי, התנאים הבאים שקולים:

1. T אוניטארית
2. $\|Tv\| = \|v\|$ לכל $v \in V$. (כאשר $\|v\| = \sqrt{\langle v, v\rangle}$)
3. T הפיכה, ו־ $T^{-1} = T^*$. זאת אומרת, $TT^* = I = T^*T$.
4. לכל בסיס אורתונורמלי $\Gamma = (v_1, ..., v_n)$ של V, מתקיים ש $T(\Gamma) = (T(v_1), ..., T(v_n))$ גם בסיס אורתונורמלי.
5. קיים בסיס אורתונורמלי $\Gamma = (v_1, ..., v_n)$ של V כך ש $T(\Gamma)$ אורתונורמלי.

הוכחה: החחה:

- $1 \Leftarrow 2$:

נניח ש־T אוניטארית, כלומר $\langle Tv, Tw\rangle = \langle v, w\rangle$ לכל $v, w \in V$. אז לכל $v \in V$ מתקיים:

$$\|Tv\|^2 = \langle Tv, Tv\rangle = \langle v, v\rangle = \|v\|^2$$

ולכן $\|Tv\| = \|v\|$.

- $2 \Leftarrow 3$:

נניח ש־ $\langle Tv, Tv\rangle = \langle v, v\rangle$. וצריך להראות ש $T^*T = I$.
לכל $v \in V$,

$$\langle v, v\rangle = \langle Tv, Tv\rangle = \langle v, T^*Tv\rangle$$

ולכן, נחסיר אגפים ונקבל:

$$\langle v, (T^*T - I)v\rangle = \langle v, T^*Tv\rangle - \langle v, v\rangle = 0$$

נתבונן. נגדיר $S := T^*T - I$. אז ראינו שלכל $v \in V$, $\langle v, Sv\rangle = 0$.

אם היינו מניחים שאנו מעל $\mathbb{C}$, אז ראינו שבמקרה זה, אם $\langle v, Sv\rangle = 0$ לכל v, אז $S = 0$. כלומר $T^*T = I$. אבל כזכור, מעל $\mathbb{R}$ זה לא נכון. (העתקת הסיבוב ב־$90°$).

אז נניח ש־V מ"ו מעל $\mathbb{R}$. ונראה ש $S = S^*$:

$$S^* = (T^*T - I)^* = (T^*T)^* - I = T^*T - I = S$$

ולכן S סימטרית. אבל ראינו בעבר, שאם $S^* = S$ וגם $\langle Sv, v\rangle = 0$ לכל $v \in V$ אז $S = 0$.

ולכן $T^*T = I$.

(כנראה שהמקרה של $\mathbb{R}$ מכסה כאן גם את המקרה של $\mathbb{C}$).

- $3 \Leftarrow 4$:

נניח ש $T^*T = I$. ויהי $\Gamma = (v_1, ..., v_n)$ בסיס אורתונורמלי של V.

צריך להראות שגם $T(\Gamma) = (T(v_1), ..., T(v_n))$ הוא בסיס אורתונורמלי.

לכל $1 \le i, j \le n$: מתקיים

$$\langle Tv_i, Tv_j\rangle = \langle v_i, T^*Tv_j\rangle \Rightarrow (T^*T = I) \Rightarrow \langle v_i, v_j\rangle = \delta_{ij}$$

זאת אומרת, $Tv_1, ..., Tv_n$ היא מערכת אורתונורמלית. ובפרט, $Tv_1, ..., Tv_n$ בת"ל, והם בגודל מימד המרחב $dimV = n$ (כי הרי $|\Gamma| = n$ והוא היה בסיס גם), ולכן זהו בסיס.

- $4 \Leftarrow 5$:

ברור. (כי ב־V קיים בסיס אורתונורמלי).

- $5 \Leftarrow 1$:

יהי $\Gamma = (v_1, ..., v_n)$ בסיס אורתונורמלי של V, כך ש־ $T(\Gamma) = (T(v_1), ..., T(v_n))$ גם כזה.

עלינו להראות ש T אוניטארית, כלומר ש $\langle Tv, Tw\rangle = \langle v, w\rangle$ לכל $v, w \in V$.

טענת עזר: יהי $(v_1, ...v_n)$ בסיס א"נ. אזי,

$$\langle\sum_{i=1}^n a_i v_i, \sum_{i=1}^n b_i v_i\rangle = \sum_{i=1}^n a_i \bar{b}_i$$

הוכחת הטענה: ראינו בעבר, ש־ $a_j = \langle\sum a_i v_i, v_j\rangle$. ולכן,

$$\langle\sum_{i=1}^n a_i v_i, \sum_{i=1}^n b_i v_i\rangle = \sum_{i=1}^n \bar{b}_i \langle\sum_{i=1}^n a_i v_i, v_i\rangle = \sum_{i=1}^n \bar{b}_i a_i = \sum_{i=1}^n a_i \bar{b}_i$$

כעת, $v_1, ...v_n$ בסיס, ולכן קיימים $\begin{matrix} a_1, ...a_n \\ b_1, ..., b_n \end{matrix} \in \mathbb{F}$ כך ש־ $\begin{matrix} v = \sum a_i v_i \\ w = \sum b_i v_i \end{matrix}$. ואז,

$$\langle v, w\rangle = \langle\sum_{i=1}^n a_i v_i, \sum_{i=1}^n b_i v_i\rangle = \sum_{i=1}^n a_i \bar{b}_i$$

וגם,

$$\langle Tv, Tw\rangle = \langle T(\sum_{i=1}^n a_i v_i), T(\sum_{i=1}^n b_i v_i)\rangle = \langle\sum_{i=1}^n a_i Tv_i, \sum_{i=1}^n b_i Tv_i\rangle = \sum_{i=1}^n a_i \bar{b}_i$$

ולכן יש שויון. נשים לב שהשתמשנו בטענת העזר בשתי המשוואות, במעבר האחרון בשתיהן. ■

9.3.3 קשרים בין העתקה אוניטארית למטריצה שלה

טענה 9.17 יהי V ממ"פ, $T:V\to V$ ה"ל, ו־ $\Gamma=(v_1,...,v_n)$ **בסיס א"נ.**
אזי, T אוניטארית אס"ם $A=[T]_\Gamma^\Gamma$ היא אוניטארית (כלומר המטריצה המייצגת את T).

הוכחה: T אוניטארית $\Leftrightarrow\ T^*T=I \Leftrightarrow [T^*T]_\Gamma^\Gamma=I \Leftrightarrow [T^*]_\Gamma^\Gamma[T]_\Gamma^\Gamma=Id$
ראינו, שאם $A=[T]_\Gamma^\Gamma$, **וגם** Γ **אורתונורמלי,** אזי $A^*=[T^*]_\Gamma^\Gamma$.
ולכן, נקבל ש T אוניטארית $\Leftrightarrow A^*A=I \Leftrightarrow A$ אוניטארית. ■

מסקנה 9.18 :1 יהי $A\in M_n(\mathbb{F})$ (כאשר $\mathbb{F}$ הוא $\mathbb{C}$ או $\mathbb{R}$), ותהי $T_A:\mathbb{F}^n\to\mathbb{F}^n$ העתקה מתאימה, $T_A(v)=Av$.
אזי A אוניטארית אס"ם T_A אוניטארית עבור מכפלה פנימית סטנדרטית.

הוכחה: בבסיס סטנדרטי E, $[T_A]_E^E=A$.
ולכן לפי הטענה הקודמת, A אוניטארית אס"ם T_A אוניטארית.
(כי E הוא בסיס א"נ עבור מכפלה פנימית סטנדרטית). ■

מסקנה 9.19 :2 תהי $A\in M_n(\mathbb{F})$. אז התנאים הבאים שקולים:

1. A אוניטארית
2. העמודות של A הן בסיס אורתונורמלי של $\mathbb{F}^n$
3. השורות של A הן בסיס אורתונורמלי של $\mathbb{F}^n$
4. A^* אוניטארית

הוכחה: נראה $2\Leftrightarrow1\Leftrightarrow4\Leftrightarrow3$.
- $2\Leftrightarrow1$:

לפי מסקנה 1, A אוניטארית $\Leftrightarrow T_A$ אוניטארית. (עבור מ"פ סטנדרטית).
ולכן לפי שקילות $1\Leftrightarrow5\Leftrightarrow4$ של המשפט "תנאים שקולים" הקודם, נקבל ש:
T_A אוניטארית $\Leftrightarrow T_A(e_1),...,T_A(e_n)$ **בסיס א"נ.**
אבל, $A(e_i)=T_A(e_i)=$העמודה ה־i.
ולכן זה מתקיים אס"ם העמודות של A הן בסיס א"נ.
הוכחה שניה: חישוב ישיר.
- $1\Leftrightarrow4$:

A אוניטארית $\Leftrightarrow A^*A=I \Leftrightarrow AA^*=I \Leftrightarrow (A^*)^*A^*=I \Leftrightarrow A^*$ אוניטארית.
- $4\Leftrightarrow3$:

לפי $1\Leftrightarrow2$ שהוכחנו, מתקיים A^* אוניטארית $\Leftrightarrow$ העמודות של A^* הן בסיס א"נ.

- אם $\mathbb{F}=\mathbb{R}$, אז העמודות של A^* = השורות של A. (כי $A^*=A^t$ במקרה זה).

• אם $\mathbb{F}=\mathbb{C}$, אז העמודות של A^* = הצמוד המרוכב של השורות של A. (כי $A^*=\overline{A}^t$ במקרה זה).

במקרה זה, ניעזר בטענה הבאה: $v_1,..,v_n\in\mathbb{C}^n$ א"נ $\Leftrightarrow$ $\overline{v}_1,...,\overline{v}_n$ הם א"נ.
באופן יותר כללי, נראה ש $\langle\overline{w},\overline{v}\rangle=\overline{\langle w,v\rangle}$.
הוכחה: אם $v=\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}$, $w=\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}$ ו־ $\overline{w}=\begin{pmatrix}\overline{b}_1\\ \vdots\\ \overline{b}_n\end{pmatrix}$, $\overline{v}=\begin{pmatrix}\overline{a}_1\\ \vdots\\ \overline{a}_n\end{pmatrix}$, אז:

$$\langle v,w\rangle=\sum b_i\overline{a}_i$$

וגם,

$$\langle\overline{v},\overline{w}\rangle=\sum\overline{b}_i a_i=\overline{\sum b_i\overline{a}_i}$$

(הערה: ייתכן שה"צמוד" האחרון צריך להיות רק על תוכן ה־$\sum$).

ובזאת סיימנו. ∎

9.3.4 תכונות של מטריצה אוניטארית ־ $|det(A)|=1$, ומשפט על ע"ע, ו"ע.

למה 9.20 אם A אוניטארית, אז $|det(A)|=1$.

הוכחה: אם $A^*A=I$, אז:

$$1=det(I)=det(A^*A)=det(A)det(A^*)$$

כמו כן,

$$A^*=\overline{A^t}\Rightarrow det(A^*)=det(\overline{A^t})=\overline{det(A^t)}=\overline{det(A)}$$

ולכן,

$$1=det(A^*)det(A)=\overline{det(A)}det(A)=|det(A)|^2$$

(כי לכל $z\in\mathbb{C}$, $z\cdot\overline{z}=|z|^2$).
ולכן $|det(A)|=1$. ∎

משפט 9.21 תהי $T:V\to V$ אוניטארית.

1. אם $\lambda\in\mathbb{C}$ הוא ע"ע של T, אזי $\lambda\overline{\lambda}=1$.
2. אם $v,w\in V$ ו"ע של T **השייכים לע"ע שונים**, אזי $v\perp w$.

הוכחה: 1) נניח שקיים $v \neq 0$ כך ש־ $Tv = \lambda v$. אזי,

$$\langle v,v\rangle = \langle Tv,Tv\rangle = \langle \lambda v,\lambda v\rangle = \lambda\overline{\lambda}\langle v,v\rangle$$

כעת, היות ו־ $v \neq 0$, מקבלים ש־ $\langle v,v\rangle \neq 0$, ולכן $\lambda\overline{\lambda} = 1$.

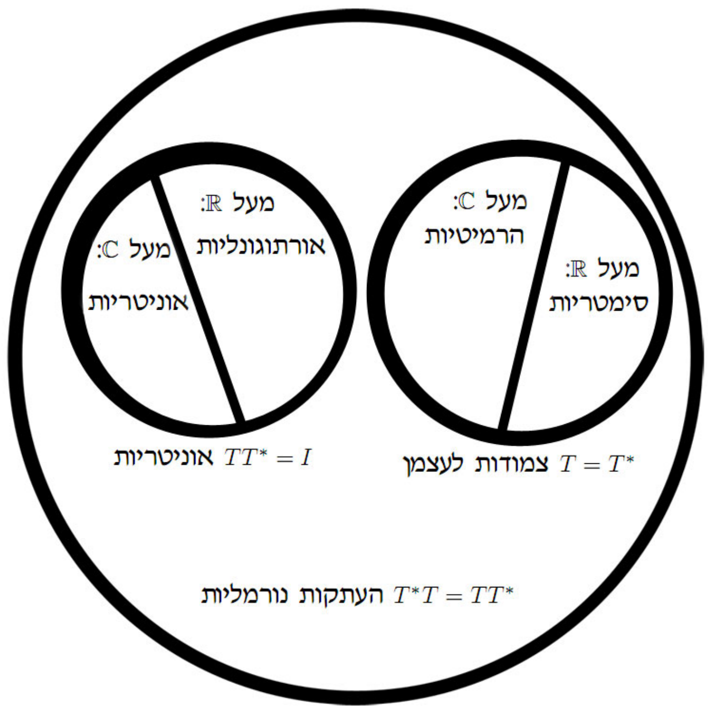
2) נניח ש־ $Tv = \lambda v$, $Tw = \mu w$. וגם $v, w \neq 0$, $\mu \neq \lambda$. צריך להוכיח ש־ $\langle v,w\rangle = 0$.

מתקיים:

$$\langle v,w\rangle = \langle Tv,Tw\rangle = \langle \lambda v,\mu w\rangle = \lambda\overline{\mu}\langle v,w\rangle$$

כעת, $\mu \neq \lambda$ ולכן $\lambda\overline{\mu} \neq \mu\overline{\mu} = 1$ (לפי החלק הראשון במשפט זה). ולכן, $\langle v,w\rangle = 0$. ■

9.4 העתקות נורמליות: $T^*T = TT^*$



9.4.1 הגדרה

הגדרה 9.22 העתקה ליניארית $T : V \to V$ תיקרא **נורמלית** אם $T^*T = TT^*$

מטריצה $A \in M_n(\mathbb{F})$ ($\mathbb{F}$ הוא $\mathbb{R}$ או $\mathbb{C}$) תיקרא **נורמלית** אס"ם $A^*A = AA^*$.

דוגמאות

- אם T הרמיטית, אז $T = T^*$, ובפרט $TT^* = T^*T$. ולכן כל העתקה הרמיטית היא נורמלית.

- אם T אוניטארית, אז $TT^* = I = T^*T$, ולכן כל העתקה אוניטארית היא נורמלית.

המטרות שלנו בנושא הזה הן בעיקר להראות שהעתקה נורמלית ניתנת ללכסון אורתונורמלי, וגם להגיע ל"משפט מבנה" על העתקות נורמליות.

9.4.2 טענות על ו"ע וע"ע של העתקה נורמלית. ($T^*v = \overline{\lambda}v \Leftarrow Tv = \lambda v$, וגם היות ו"ע של ע"ע שונים מאונכים זה לזה)

טענה 9.23 בממ"פ V סוף־מימדי, אם T נורמלית, λ ע"ע של T, ו־v ו"ע המתאים לע"ע λ, אז $\overline{\lambda}$ הע"ע של T^* המתאים ל־v. (במילים אחרות, אם $Tv = \lambda v$ אז $T^*v = \overline{\lambda}v$).

הוכחה: נשים לב שאם T נורמלית, אז לכל v, $\|T^*v\| = \|Tv\|$ ובפרט, ל־T ול־T^* יש את אותו הגרעין.

מדוע זה נכון? כי $\|T^*v\|^2 = \langle T^*v, T^*v\rangle = \langle v, TT^*v\rangle = \langle v, T^*Tv\rangle = \langle Tv, Tv\rangle = \|Tv\|^2$.

כזכור, הנחנו ש־ $Tv = \lambda v$.

נגדיר העתקה $S = T - \lambda I$. אז S נורמלית.

איך אנו יודעים זאת? כי:

$$S^* = T^* - \overline{\lambda}I$$

ואז,

$$SS^* = (T - \lambda I)(T^* - \overline{\lambda}I) = TT^* - \lambda T^* - \overline{\lambda}T + |\lambda|^2 I$$

$$S^*S = (T^* - \overline{\lambda}I)(T - \lambda I) = TT^* - \lambda T^* - \overline{\lambda}T + |\lambda|^2 I$$

ולכן $S^*S = SS^*$.

S נורמלית, ו־ $Sv = Tv - \lambda v = 0$ ולכן $S^*v = 0$ (יש להן את אותו הגרעין ־ הוכחנו קודם).

ומכאן, $T^*v - \overline{\lambda}v = 0$ ו־ $T^*v = \overline{\lambda}v$. ■

טענה 9.24 בממ"פ סוף־מימדי V, תהי T העתקה נורמלית, ויהיו v_1, v_2 ו"ע של T, ביחס לע"ע λ_1, λ_2 **ששונים זה מזה**. אז $v_1 \perp v_2$.

הוכחה: אנו רוצים להראות ש־ $\langle v_1, v_2\rangle = 0$. מתקיים:

$$\lambda_1\langle v_1, v_2\rangle = \langle \lambda_1 v_1, v_2\rangle = \langle Tv_1, v_2\rangle = \langle v_1, T^*v_2\rangle = \langle v_1, \overline{\lambda}v_2\rangle = \lambda_2\langle v_1, v_2\rangle$$

ולכן, $(\lambda_1 - \lambda_2)\langle v_1, v_2\rangle = 0$.

ומכיוון ש־ $\lambda_1 - \lambda_2 \neq 0$, קיבלנו ש־ $\langle v_1, v_2\rangle = 0$. ■

9.4.3 הקשר בין ת"מ T־אינוריאנטי, והמשלים האורתוגונלי של הת"מ ו־T^*

הגדרה 9.25 יהי V מ"ו מעל שדה $\mathbb{F}$, ותהי $T : V \to V$ ה"ל.
תת־מרחב $U \subseteq V$ נקרא **אינוריאנטי ביחס ל־T** (או T־אינוריאנטי), אם $Tu \in U$ לכל $u \in U$.

טענה 9.26 אם T העתקה (לא בהכרח נורמלית!), ו־U אינוריאנטי ביחס ל־T, אז $U^\perp$ אינוריאנטי ביחס ל־T^*.

הוכחה: יהי $v \in U^\perp$. עלינו להראות ש־ $T^*v \in U^\perp$.
כלומר, שלכל $u \in U$, מתקיים $\langle u, T^*v\rangle = 0$.
אז יהי $u \in U$, $v \in U^\perp$. מתקיים:

$$\langle u, T^*v\rangle = \langle Tu, v\rangle = 0$$

והשויון האחרון נכון כי $Tu \in U$ (שהרי U הוא T־אינוריאנטי). ■

9.4.4 לכסון אורתונורמלי של מטריצות נורמליות (מעל $\mathbb{C}$)

משפט 9.27 יהי V ממ"פ ממימד סופי (מעל $\mathbb{C}$) , $T : V \to V$ **נורמלית**.
אז קיים בסיס אורתונורמלי ל־V שמורכב מוקטורים עצמיים של T.

הוכחה: יהיו $\lambda_1, ..., \lambda_n$ הע"ע השונים של T. קיימים כאלה כי אנחנו מעל $\mathbb{C}$.
לכל $k = 1, ..., n$, **יהי** $U_k = \{v \,:\, Tv = \lambda_k v\}$.
ולכל k, יהי B_k בסיס אורתונורמלי ל־U_k.
ויהי $B = \bigcup_{k=1}^n B_k$.
כל איברי B הם ו"ע של T, וכולם ניצבים זה לזה.
מדוע? כי, יהיו $v_1, v_2 \in B$. אם $v_1, v_2 \in U_k$ אז הם ניצבים מבחירת B_k.
אם $v_1 \in B_k$, $v_1 \in B_j$ כאשר $k \neq j$, אז הם ניצבים: $v_1 \perp v_2$ מטענה קודמת.
נותר להראות: ש $V = span(B)$.
יהי $U = span(B)$. אז $U = U_1 \oplus U_2 \oplus ... \oplus U_n$.
נניח בשלילה ש־$U \neq V$. במקרה זה, $U^\perp \neq \{0\}$.
(כנראה, כי $V = U \oplus U^\perp$ ולכן זה נובע. זה השערה שלי, לא "רשמי").
וכעת, U אינוריאנטי ביחס ל־T.

מדוע? יהי $u \in U$. אז $u = \alpha_1 b_1 + ... + \alpha_n b_n$ כאשר $b_k \in U_k$. ואז, $Tu = \alpha_1 Tb_1 + ... + \alpha_n Tb_n = \lambda_1\alpha_1 b_1 + ... + \lambda_n \alpha_n b_n \in U$

ולכן $U^\perp$ הוא אינוריאנטי ביחס ל־T^*. (מהטענה הקודמת).
ל־T^* חייב להיות ע"ע α ב־$U^\perp$ (כנראה שהכוונה לα ב־$\mathbb{F}$, זו טעות?) עם ו"ע $v \in U^\perp$ (כי הכל מעל ל־$\mathbb{C}$). (למה? לא ברור לי).
אז, $T^*v = \alpha v$.
ולכן, עבור $\lambda' = \overline{\alpha}$, מתקיים $Tv = \lambda' v$. (מהטענה שראינו).
בפרט, λ' הוא ע"ע של T, ולכן, $\lambda' = \lambda_j$ לאיזשהו $1 \leq j \leq n$.
ולכן $v \in U_j \subseteq U$, **בסתירה** לכך ש־ $v \in U^\perp$.
ולכן $V = span(B)$. ■

מסקנה 9.28 אם T נורמלית, אז קיים בסיס אורתונורמלי Γ כך ש־ $[T]_\Gamma$ היא מטריצה אלכסונית.
(עוד פעם רק מעל $\mathbb{C}$ כנראה).

הערה 9.29 היתה עוד הוכחה למשפט הזה, על רגל אחת ובאינדוקציה, ב18.5.09.

טענה 9.30 יהי V ממ"פ ממימד סופי, ו־Γ בסיס אורתונורמלי.
אז להעתקה ליניארית T, מתקיים ש־ T נורמלית $\Leftrightarrow$ $[T]_\Gamma$ נורמלית.

הוכחה: נזכור שמכיוון ש־Γ אורתונורמלי, אז מתקיים $[T]_\Gamma^* = [T^*]_\Gamma$.
(הכוונה כאן היא ש־ $[T]_\Gamma^* = \overline{([T]_\Gamma)^t}$ כנראה)
⇐נניח T נורמלית. ונסמן: $A = [T]_\Gamma$. אז,

$$AA^* = [T]_\Gamma[T]_\Gamma^* = [T]_\Gamma[T^*]_\Gamma = [TT^*]_\Gamma = [T^*T]_\Gamma = [T^*]_\Gamma[T]_\Gamma = A^*A$$

כאשר המעבר השני מימין והשני משמאל הם מאורתונורמליות הבסיס, ובאמצע השתמשנו בכך שT נורמלית.
⇒נניח ש־ A נורמלית. אז לכל v,

$$[TT^*v]_\Gamma = [T]_\Gamma[T_\Gamma^*][v]_\Gamma = [T]_\Gamma[T]_\Gamma^*[v]_\Gamma = AA^*[v]_\Gamma = A^*A[v]_\Gamma =$$

$$= [T^*]_\Gamma[T]_\Gamma[v]_\Gamma = [T^*Tv]_\Gamma$$

ולכן, לכל v, $TT^*v = T^*Tv$ ומכאן, $TT^* = T^*T$. ■

מסקנה 9.31 לכסון אוניטארי:
אם A מטריצה נורמלית, אז קיימת מטריצה אוניטארית U ומטריצה אלכסונית B כך ש־$A = UBU^*$.
(מעל $\mathbb{R}$ זה לא תמיד נכון. רק מעל $\mathbb{C}$, כי השתמשנו במשפט ממקודם...)

הוכחה: הושארה כתרגיל!! (ניתן רמז: U תהיה מטריצת מעבר לבסיס האורתונורמלי של ו"ע של A).
(**כנראה** שU תהיה אוניטרית בגלל המשפט שלמדנו שאומר שU אוניטרית⇔העמודות שלה הן בסיס א"נ של $\mathbb{F}^n$).

■

9.4.5 פירוק קוטבי: הצגה של מטריצה נורמלית בתור $A = HU = UH$ (U אוניטרית, H הרמיטית)

משפט 9.32 תהי A מטריצה נורמלית.
אז קיימות מטריצה אוניטרית U ומטריצה הרמיטית H כך ש־ $A = HU = UH$.
(כנראה רק מעל $\mathbb{C}$. זה לא נאמר בהרצאה, אך הגיוני).

הוכחה: קודם כל נוכיח למקרה שבו B מטריצה אלכסונית, ואח"כ ניעזר בזה.

תהי B מטריצה אלכסונית. נראה שקיימות H,U כנ"ל כך ש־ $B=HU=UH$:

נניח ש־ $B=\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}$ (וייתכן שיש $\lambda_i=\lambda_j$).

ונסמן: $z_k=\begin{cases}\frac{\lambda_k}{|\lambda_k|} & if\,|\lambda_k|\neq 0\\ 1 & if\,|\lambda_k|=0\end{cases}$.

אז יהיו $H=\begin{pmatrix}|\lambda_1| & & 0\\ & \ddots & \\ 0 & & |\lambda_n|\end{pmatrix}$, $U=\begin{pmatrix}z_1 & & 0\\ & \ddots & \\ 0 & & z_n\end{pmatrix}$.

אז H הרמיטית כי היא מט' אלכסונית ממשית.

U אוניטרית, כי היא אלכסונית שכל איברי האלכסון הם בעלי ערך מוחלט 1.

ומתקיים $B=HU=UH$.

כעת נוכיח את המשפט.

תהי A מטריצה נורמלית. אז קיימות U_2 אוניטרית ו־B אלכסונית כך ש־ $A=U_2BU_2^*$.

(ליכסון אוניטרי, הוכחנו כבר שקיים).

כמו כן, קיימות U_1 אוניטרית ו־H_1 הרמיטית כך ש־ $B=U_1H_1=H_1U_1$ (חלק א' של הוכחה זאת).

אז, (המעבר השמאלי ביותר הוא סימון ־ כל סוגריים סימנו באות):

$$A=U_2U_1H_1U_2^*=U_2U_1(U_2^*U_2)H_1U_2^*=(U_2U_1U_2^*)(U_2H_1U_2^*)=UH$$

ומצד שני,

$$A=U_2H_1U_1U_2^*=U_2H_1(U_2^*U_2)U_1U_2^*=HU$$

אז קיבלנו ש־ $A=HU=UH$.

נותר רק לבדוק ש־H הרמיטית ו־U אוניטרית.

נזכור ש־H_1 הרמיטית, ולכן $H_1^*=H_1$. אז:

$$H^*=(U_2H_1U_2^*)^*=U_2^{*^*}H_1^*U_2^*=U_2H_1U_2^*=H$$

ולכן H הרמיטית.

כמו כן נזכור ש־U_1,U_2 אוניטריות ולכן $U_1U_1^*=I=U_2U_2^*$. אז:

$$UU^*=(U_2U_1U_2^*)(U_2U_1U_2^*)^*=U_2U_1U_2^*U_2U_1^*U_2^*=U_2U_1U_1^*U_2^*=U_2U_2^*=I$$

ולכן U אוניטרית, וסיימנו. ■

9.4.6 פירוק לחלק ממשי ומדומה

טענה 9.33 אם $A \in M_n(\mathbb{C})$ **נורמלית**, אז קיימות מטריצות הרמיטיות R, S כך שמתקיים $A = R + iS\,,\ A^* = R - iS$, **וגם** $\underline{RS = SR}$.

הוכחה: ראינו כבר בעבר, שלכל מטריצה $A \in M_n(\mathbb{C})$ מתקיים $A = R + iS,\ A^* = R - iS$ כאשר R, S הרמיטיות, המוגדרות ע"י $R = \frac{A+A^*}{2}\ \ ,\ \ S = \frac{A-A^*}{2i}$.
ואם A נורמלית (כמו בטענה הנוכחית...), אז גם מתקיים **בנוסף**:

$$RS = \frac{A+A^*}{2} \cdot \frac{A-A^*}{2i} = ...something.$$

$$SR = \frac{A-A^*}{2i} \cdot \frac{A+A^*}{2} = ...the\,same\,thing.$$

■

9.4.7 שני תנאי אם"ם לקביעה אם מט' נורמלית היא אוניטרית, הרמיטית

טענה 9.34

1. מט' נורמלית $A \in M_n(\mathbb{C})$ היא הרמיטית אם"ם כל הע"ע שלה ממשיים.
2. מט' נורמלית היא אוניטרית אם"ם כל הע"ע שלה הם מס' מרוכבים בעלי ערך מוחלט 1.(כנראה גם רק מעל $\mathbb{C}$)

הוכחה: לפי משפט הליכסון האוניטרי, $A = U^*DU$, $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, U אוניטרית.

1) אם A הרמיטית, אז אם $Av = \lambda v$, אז $\lambda v = Av = A^*v = \overline{\lambda} v \ \Leftarrow\ \lambda = \overline{\lambda} \Leftarrow \lambda \in \mathbb{R}$
 בכיוון השני, אם $\lambda_1, ..., \lambda_n$ ממשיים, אז $D = D^*$, ואז:

$$A^* = (UDU^*)^* = U^*D^*U^{**} = U^*DU = A$$

ולכן A הרמיטית.

2) אם A אוניטרית, אז אם $Av = \lambda v$, אז (נשים לב ש $\langle v, v\rangle = \langle Av, Av\rangle$ מהאוניטריות):

$$\|v\|^2 = \langle v, v\rangle = \langle Av, Av\rangle = \langle \lambda v, \lambda v\rangle = \lambda\overline{\lambda}\langle v, v\rangle = |\lambda|^2\|v\|^2 \quad \Rightarrow |\lambda| = 1$$

בכיוון השני, אם $|\lambda_i| = 1$ לכל $1 \le i \le n$, אז $DD^* = I$ (כי $\lambda_i\overline{\lambda_i} = |\lambda_i|^2 = 1$), ואז:

$$AA^* = (U^*DU)(U^*D^*U) = U^*DD^*U = I$$

ולכן A אוניטרית. ■

9.5 מטריצות נורמליות מעל $\mathbb{R}$

הערה: כדאי לשים לב שחלק מהטענות בפרק הקודם גם היו תקפות מעל $\mathbb{R}$.

9.5.1 ליכסון אורתוגונלי (מעל $\mathbb{R}$)

מעל $\mathbb{R}$, לא לכל מטריצה נורמלית קיים ליכסון אורתוגונלי.

למשל, $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ (סיבוב ב־$90°$).

מה שכן אפשר להגיד הוא המשפט הבא:

משפט 9.35 משפט הליכסון האורתוגונלי

לכל העתקה **סימטרית** $T : V \to V$ בממ"פ ממשי V ממימד סופי, קיים ליכסון אורתוגונלי, כלומר בסיס אורתונורמלי של ו"ע.

במילים אחרות, לכל מטריצה סימטרית מעל $\mathbb{R}$ קיימת מטריצה אלכסונית D ומטריצה אורתוגונלית P כך ש־ $A = P^t DP$. (נזכור שהשם "מט' אורתוגונלית" הוא עבור מט' אוניטרית מעל $\mathbb{R}$).

הוכחה: נוכיח עבור מטריצות.

אם $A \in M_n(\mathbb{R})$ סימטרית, אז אפשר לחשוב על A כמטריצה מרוכבת. לפי משפט הליכסון האוניטרי, קיים לה בסיס אורתונורמלי (ביחס למכפלה הפנימית הסטנדרטית על $\mathbb{C}^n$) של וקטורים עצמיים (אולי עם קואורדינטות לא ממשיות).

בפרט זה אומר שסכום הריבויים הגיאומטריים שווה ל־n, כלומר:

$$n = \sum_{j=1}^{k} dim_{\mathbb{C}}\, Ker_{\mathbb{C}}\, (A - \lambda_j I)$$

כאשר מסמנים ב־ $\lambda_1, ..., \lambda_k$ את הע"ע השונים.

אבל, לכל $1 \leq j \leq k$, λ_j הוא ממשי (כי A סימטרית), ולכן $A - \lambda_j I$ היא מטריצה ממשית, ולכן

$$dim_{\mathbb{R}}\, Ker_{\mathbb{R}}\, (A - \lambda_j I) = dim_{\mathbb{C}}\, Ker_{\mathbb{C}}\, (A - \lambda_j I)$$

זה נכון בגלל תהליך הדירוג של גאוס:

אם מטריצה היא ממשית אז כשמבצעים את תהליך הדירוג, בכל שלב כל המספרים במטריצה יהיו ממשיים, ולכן הצורה הקנונית תהיה אותה מטריצה מעל $\mathbb{R}$ או $\mathbb{C}$.

(ובפרט מספר שורות האפסים שלה יהיה שווה בשני המקרים, ולכן המימד של הגרעין גם).

המסקנה היא ש־$n = dim_{\mathbb{R}}\, Ker_{\mathbb{R}}\, (A - \lambda_j I)$ ולכן A ניתנת ללכסון מעל $\mathbb{R}$.

אך למעשה ניתן למצוא ליכסון אורתוגונלי:

לכל $1 \leq j \leq n$, ניקח בסיס אורתונורמלי $\{v_{j,1}, v_{j,2}, ..., v_{j,m_j}\}$.

אז הקבוצה: $\left\{ v_{j,i} \,,\, \begin{matrix} 1 \leq j \leq k \\ 1 \leq i \leq m_j \end{matrix} \right\}$ היא בסיס אורתונורמלי של ו"ע של A (מעל $\mathbb{R}$).

(משתמשים בכך שו"ע של מט' נורמלית המתאימים לע"ע שונים הם ניצבים). ■

9.5.2 מירכוב של מרחב אוקלידי

אנו רוצים להוכיח כעת את משפט המבנה להעתקות נורמליות מעל $\mathbb{R}$. לצורך כך בילינו הרצאה שלמה בהסברים על מה שנקרא מירכוב של מרחב אוקלידי.

הגדרה 9.36 יהי V ממ"פ מעל $\mathbb{R}$. נגדיר מרחב $\hat{V}$ של כל הביטויים הפורמליים מהצורה $u+iv$ כאשר $u,v \in V$.

$$\hat{V} = \{u+iv \,:\, u,v \in V\}$$

(בעצם איברי $\hat{V}$ הם זוגות (u,v) כאשר $u,v \in V$ ומסמנים "$u+iv$" עבור הזוג (u,v)).

על $\hat{V}$ נגדיר פעולות:

חיבור:

$$(u+iv)+(u'+iv') = (u+u')+i(v+v')$$

כפל בסקלר מרוכב:

$$(a+ib)(u+iv) = (au-bv)+i(av+bu)$$

קל לבדוק ש־$\hat{V}$ עם פעולות אלו הוא מ"ו מעל $\mathbb{C}$.

הערה 9.37 א. אנו אומרים ש־ $u+iv = u'+iv'$ אם"ם $u=u',\, v=v'$.

ב. סימון: אם $w = u+iv \in \hat{V}$, $u,v \in V$, אז נסמן $\begin{matrix} u = \Re e\, w \\ v = \Im m\, w \end{matrix}$ (חלק "ממשי" ו"מדומה").

עוד סימון: $\overline{w} = u - iv$ ("הצמוד של w").

ג. אם נזהה את $v \in V$ עם הביטוי הפורמלי $v+i0$ אז V נהיית תת־קבוצה של $\hat{V}$:

$V \subseteq \hat{V}$ ("בעייתי מבחינת תורת הקבוצות, אבל לא נורא").

זהו **לא תת־מרחב**! (כי הסקלרים ב־$\hat{V}$ הם מ־$\mathbb{C}$).

טענה 9.38 אם $v_1,\ldots,v_k \in V$ קבוצה בת"ל אז היא גם בת"ל ב־$\hat{V}$ (מעל $\mathbb{C}$).

הוכחה: אם $\sum_{j=1}^k \alpha_j v_j = 0$ כאשר $\alpha_j = a_j + ib_j$ ($a_j, b_j \in \mathbb{R}$ לכל $1 \le j \le k$), אז

$$0+i0 = \sum_{j=1}^k \alpha_j v_j = \sum_{j=1}^k (a_j+ib_j)(v_j+i0) = \left(\sum_{j=1}^k a_j v_j\right) + i\left(\sum_{j=1}^k b_j v_j\right)$$

ומכאן מקבלים $\sum_{j=1}^k b_j v_j = 0$, $\sum_{j=1}^k a_j v_j = 0$. ולכן לכל j, מתקיים $a_j = b_j = 0$, כלומר $\alpha_j = 0$. ■

טענה 9.39 אם $v_1,\ldots,v_k \in V$ הם בסיס, אז הם גם בסיס ל־$\hat{V}$ (מעל $\mathbb{C}$).

הוכחה: צריך להראות שהם פורשים את $\hat{V}$. (הראינו כבר בת"ל).

אם $w = u+iv \in \hat{V}$, נכתוב: $u = \sum_{j=1}^k a_j v_j$, $v = \sum_{j=1}^k b_j v_j$ כאשר $a_i, b_i \in \mathbb{R}$.

ואז, $w = u+iv = \sum_{j=1}^k (a_j+ib_j) v_j$. ■

הגדרה 9.40 כעת נגדיר מבנה של ממ"פ על $\hat{V}$:
אם $w = u + iv \in \hat{V}$, $w' = u' + iv' \in \hat{V}$ נגדיר:

$$\langle w, w' \rangle_{\mathbb{C}} = (\langle u, u' \rangle_{\mathbb{R}} + \langle v, v' \rangle_{\mathbb{R}}) + i\left(-\langle u, v' \rangle_{\mathbb{R}} + \langle v, u' \rangle_{\mathbb{R}}\right)$$

תרגיל: לבדוק שזה מקיים את האקסיומות של מ"פ מעל $\mathbb{C}$.

הערה 9.41 מתקיים: $\|w\|^2 = \|u\|_V^2 + \|v\|_V^2$ כאשר $w = u + iv \in \hat{V}$.

הערה 9.42 $\hat{V}$ נקרא המירכוב של V.

תהי $T : V \to V$ העתקה ליניארית. נגדיר העתקה: $\hat{T} : \hat{V} \to \hat{V}$ ע"י $\hat{T}(u + iv) = T(u) + iT(v)$.
קל לבדוק שזו העתקה ליניארית.
היא מרחיבה את הגדרת T, כלומר $\hat{T}(u) = T(u)$ עבור $u \in V$.

טענה 9.43 טענה זו הובאה ללא הוכחה:

1. אם $u, v \in V$ אז $\langle u, v \rangle_{\mathbb{C}} = \langle u, v \rangle_{\mathbb{R}}$
2. לכל $w_1, w_2 \in \hat{V}$, מתקיים $\overline{w_1 + w_2} = \overline{w}_1 + \overline{w}_2$
3. אם $w \in \hat{V}$, $a \in \mathbb{C}$, אז $\overline{aw} = \overline{a} \cdot \overline{w}$
4. אם $w \in \hat{V}$ אז $\Re e\, w = \frac{w + \overline{w}}{2}$, $\Im m\, w = \frac{w - \overline{w}}{2}$ (כנראה $2i$ במכנה של $\Im m$, אולי לא העתקתי טוב).
5. $\langle \overline{w}_1, \overline{w}_2 \rangle_{\mathbb{C}} = \overline{\langle w_1, w_2 \rangle}$
6. $w \in V \Leftrightarrow w = \overline{w}$

טענה 9.44 אם $T, S : V \to V$ העתקות ליניאריות, אז:

1. $\widehat{T + S} = \hat{T} + \hat{S}$
2. אם $\alpha \in \mathbb{R}$ אז $\widehat{\alpha T} = \alpha \hat{T}$
3. $\widehat{T \circ S} = \hat{T} \circ \hat{S}$
4. $(\widehat{T^*}) = (\hat{T})^*$
5. אם $I_V : V \to V$ היא העתקת הזהות, אז $\widehat{I_V} = I_{\hat{V}}$

הוכחה: נוכיח רק את (4):
("לקרוא לזה הוכחה זה עלבון למילה 'הוכחה'").
ניקח $w = u + iv\,,\ w' = u' + iv'$. אז,

$$\langle w, (\hat{T})^*(w') \rangle = \langle \hat{T}(w), w' \rangle = \langle T(u) + iT(v), u' + iv' \rangle =$$

$$= \langle T(u), u' \rangle + i\langle T(v), u' \rangle - i\langle T(u), v' \rangle + \langle T(v), v' \rangle =$$

$$= \langle u, T^*(u')\rangle + i\langle v, T^*(u')\rangle - i\langle u, T^*(v')\rangle + \langle v, T^*(v')\rangle$$

רוצים להראות שזה שווה ל־ $\langle w, \widehat{T^*}(w')\rangle$. נחשב באופן דומה:

$$\langle w, \widehat{T^*}(w')\rangle = \langle u+iv, T^*(u') + iT^*(v')\rangle = ... = the\, same\, thing.$$

(התעצלנו בהרצאה). ■

מסקנה 9.45

1. אם T נורמלית אז $\hat{T}$ נורמלית.
2. אם T אורתוגונלית אז $\hat{T}$ היא אוניטארית
3. אם T סימטרית אז $\hat{T}$ היא הרמיטית.

הוכחה: בדיקה של א':

$$\hat{T}(\hat{T})^* = \hat{T}(\widehat{T^*}) = (\widehat{TT^*}) = (\widehat{T^*T}) = (\widehat{T^*})\hat{T} = (\hat{T})^*\hat{T}$$

■

משפט 9.46 תהי $T : V \to V$ העתקה ליניארית, ויהי $\alpha \in \mathbb{R}$ ע"ע של $\hat{T}$ (אם יש כזה).
אז α הוא ע"ע של T, וכן קיים בסיס אורתונורמלי של $\hat{V}_\alpha = Ker_{\hat{V}}(\hat{T} - \alpha I)$ המורכב כולו מוקטורים של V.

הוכחה: α הוא ע"ע של $\hat{T}$, כלומר קיים $w = u + iv \in \hat{V}$, $w \neq 0$, כך ש־ $\hat{T}(w) = \alpha w$.

כלומר $T(u) + iT(v) = \alpha u + i\alpha v$, או באופן שקול: $\begin{matrix} T(u) = \alpha u \\ T(v) = \alpha v \end{matrix}$. אז לפחות אחד מבין u, v (זה שאינו אפס) הוא ו"ע של T עם ע"ע α.

יהי $\{v_1, ..., v_k\}$ בסיס א"נ ל־ $V_\alpha = Ker_V(T - \alpha I)$. נראה שזהו בסיס ל־ $\hat{V}_\alpha$:

ה־v_i הם ו"ע של $\hat{T}$ עם ע"ע α. הם אורתונורמליים, ולכן בת"ל.

הם גם פורשים את $\hat{V}_\alpha$, כי אם $w = u + iv \in \hat{V}_\alpha$, אז $u, v \in \hat{V}_\alpha$. (ולכן ב־V_α).

את u, v אפשר להציג בצורה: $u = \sum_{j=1}^k a_j v_j\ ,\ v = \sum_{j=1}^k b_j v_j$

ומכאן מקבלים $w = \sum_{j=1}^k (a_j + ib_j) v_j$. ■

משפט 9.47 תהי $T : V \to V$ ה"ל. אם $\lambda \notin \mathbb{R}$ הוא ע"ע של $\hat{T}$, אז $\overline{\lambda}$ הוא גם ע"ע של $\hat{T}$, וכן אם $\{v_1, ..., v_k\}$ בסיס א"נ של $\hat{V}_\lambda$, אז $\overline{v}_1, ..., \overline{v}_k$ בסיס א"נ ל־ $\hat{V}_{\overline{\lambda}}$.

הוכחה: (לא ממש פורמלית, לא מלאה, כך היה בהרצאה).

אם $\lambda = a+ib$, $a,b\in\mathbb{R}$, $b\neq 0$, **אז לפי הנתון קיים** $w=u+iv$ **כך ש־** $\hat{T}(w)=\lambda w$.

לכן,

$$\hat{T}(\overline{w}) = T(u) - iT(v) = \overline{\hat{T}(w)} = \overline{\lambda w} = \overline{\lambda}\cdot\overline{w}$$

כלומר $\overline{w}$ הוא ו"ע של $\hat{T}$ עם ע"ע $\overline{\lambda}$.

המשך ההוכחה הושאר כתרגיל. ■

9.5.3 משפט המבנה להעתקות נורמליות מעל $\mathbb{R}$

אם $A\in M_n(\mathbb{R})$, **נורמלית**, רוצים למצוא מטריצה אורתוגונלית P ומטריצה B שתהיה פשוטה ככל האפשר כך ש־$A=P^tBP$.

משפט 9.48 תהי $T:V\to V$ העתקה נורמלית מעל ממ"פ V ממימד סופי מעל $\mathbb{R}$.

אז קיים בסיס אורתונורמלי $C=\{v_1,...,v_n\}$ של V כך שמטריצת הייצוג של T לפי C תהיה מהצורה:

$$[T]_C = \begin{pmatrix} \lambda_1 & & & & & & & 0 \\ & \lambda_2 & & & & & & \\ & & \ddots & & & & & \\ & & & \lambda_k & & & & \\ & & & & A_1 & & & \\ & & & & & A_2 & & \\ & & & & & & \ddots & \\ 0 & & & & & & & A_m \end{pmatrix}$$

כאשר $\lambda_1,...,\lambda_k\in\mathbb{R}$,

ו־$A_1,...,A_m$ הם **מטריצות** ("בלוקים") מסדר 2×2, מהצורה: $A_j = \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}$

(יכול להיות יותר מבלוק אחד עם אותם ערכים. וכנראה שגם $\lambda_i=\lambda_j$ ייתכן).

הוכחה: רעיון כללי: ל־$\hat{T}$ קיים ליכסון אוניטרי. יהיו $\lambda_1,...,\lambda_n$ הע"ע שלה (כולל ריבוי, כלומר עלולים להופיע יותר מפעם אחת). נסדר אותם כך ש־ $\lambda_1,...\lambda_k$ הם הע"ע הממשיים. נותר להתמודד עם $\lambda_{k+1},...,\lambda_n$.

ההוכחה ארוכה ונוראית ואיומה, אך למזלנו היא לא למבחן!! אז לא אכלול אותה כאן. ■

10 צורת ז'ורדן

מעתה ואילך V הוא מרחב וקטורי, לא ממ"פ.

10.1 הקדמה ־ סכומים ישרים

10.1.1 שדה סגור אלגברית

שדה $\mathbb{F}$ נקרא **סגור אלגברית** אם לכל פולינום $f \in \mathbb{F}[x]$ יש שורש $\alpha \in \mathbb{F}$.

עובדה חשובה (לא נלמד בקורס הזה): כל שדה $\mathbb{F}$ מוכל בשדה סגור אלגברית. השדה המינימלי הסגור אלגברית שמכיל שדה נתון $\mathbb{F}$ נקרא הסגור האלגברי של $\mathbb{F}$.

לדוגמא: הסגור האלגברי $\mathbb{Q}$: אוסף המספרים האלגבריים. (מספר מרוכב נקרא מספר אלגברי אם הוא שורש של פולינום עם מקדמים רציונליים).

המטרה:

למצוא **צורה קנונית** למחלקות דמיון של העתקות ליניאריות $T : V \to V$ כאשר V מ"ו מעל שדה סגור אלגברית $\mathbb{F}$.

במילים אחרות, למצוא צורה פשוטה ככל האפשר של מטריצות, כך שלכל העתקה ליניארית קיים בסיס כך שמטריצת הייצוג שלה תהיה מהצורה הפשוטה.

10.1.2 סכומים ישרים

הגדרה 10.1 יהי V מרחב וקטורי מעל שדה $\mathbb{F}$, ויהיו $U_1, ..., U_k \subseteq V$ תתי־מרחבים.

נאמר שת"מ $W \subseteq V$ הוא סכום של $U_1, ..., U_k$ (סימון: $W + U_1 + ... + U_k$ או $W = \sum_{i=1}^{k} U_i$),

אם לכל $w \in W$ קיימת הצגה בצורה $w = u_1 + ... + u_k$ כאשר $u_i \in U_i$.

נאמר שהסכום הוא **סכום ישר** (סימון: $W = U_1 \oplus ... \oplus U_k$ או $W = \bigoplus_{i=1}^{k} U_i$),

אם לכל $w \in W$ יש הצגה **יחידה** כזו.

טענה 10.2 $W = \bigoplus_{i=1}^{m} U_i$ אם"ם מתקיים:

1.

$$W = \sum_{i=1}^{m} U_i$$

<u>**ובנוסף,**</u>

2.

$$\forall 1 \leq i \leq m \qquad : \qquad U_i \cap \sum_{\substack{j=1 \\ j \neq i}}^{m} U_j = \{0\}$$

הוכחה: דומה למה שהוכחנו בעבר, הושאר כתרגיל. ∎

טענה 10.3 יהיו $U_1, ..., U_k \subseteq V$ תתי־מרחבים, כך ש־ $U_i \cap U_j = \{0\}$ לכל $i \neq j$, ולכל i, יהי B_i בסיס ל־ U_i.

אז עבור ת"מ $W \subseteq V$, מתקיים ש $W = \bigoplus_{i=1}^{k} U_i$ אם"ם איחוד הבסיסים $\bigcup_{i=1}^{k} B_i$ הוא בסיס ל־W.

הוכחה: כיוון א':

אם $W=\bigoplus_{i=1}^k U_i$ אז ברור ש־ $\bigcup_{i=1}^k B_i$ קב' פורשת ל־W.

נבדוק שהיא בת"ל: אם יש צירוף ליניארי ששווה 0, אז תחילה ע"י קיבוץ איברים מקבלים צירוף מהצורה $\sum_{i=1}^k u_i = 0$ כאשר $u_i \in U_i$. (כל u_i הוא החלק בצירוף הליניארי שמורכב מאיברי B_i). מהגדרת סכום ישר מקבלים ש־ $u_i = 0$ לכל i.

כל u_i הוא צ"ל של איברי B_i, אבל B_i בת"ל ולכן כל המקדמים גם שווים 0.

כיוון ב':

אם איחוד הבסיסים הוא בסיס ל־W, אז בוודאי $W=\sum_{i=1}^k U_i$.

הסכום הוא ישר, כי אם קיים וקטור $u \neq 0$ ב־ $U_i \cap (\sum_{j\neq i} U_j)$ אז אפשר להציג את u כצ"ל של איברי B_i (כי $u\in U_i$), וגם כצ"ל של איברי $\bigcup_{j\neq i} B_j$ (כי $u \in \sum_{j\neq i} U_j$).

אבל אמרנו ש־$U_i \cap U_j = \{0\}$ לכל $i\neq j$ (זה נתון במשפט), ולכן $u=0$, בסתירה לכך שהוא $\neq 0$. ■

מסקנה 10.4 אם $W=\bigoplus_{i=1}^k U_i$ אז $dim\, W = \sum_{i=1}^k dim\, U_i$.

טענה 10.5 אם $W=\sum_{i=1}^k U_i$ וגם $dim\, W = \sum_{i=1}^k dim\, U_i$, אז $W=\bigoplus_{i=1}^k U_i$.

הוכחה: יהי B_i בסיס ל־U_i בגודל $dim\, U_i$.

הקבוצה $\bigcup_{i=1}^k B_i$ היא קבוצה בגודל $dim\, W = \sum_{i=1}^k dim\, U_i$. היא פורשת את W (כי $W=\sum U_i$) ולכן היא בסיס. ולכן $W=\bigoplus_{i=1}^k U_i$.

(זאת הוכחה מספיק טובה? נראה לי שחסר משהו.)

■

10.2 מרחבים אינווריאנטיים, ציקליים

10.2.1 מטריצות בלוקים ומרחבים אינווריאנטים

אבחנה:

תהי $T:V\to V$ העתקה ליניארית, ויהיו $U_1,...,U_k \subseteq V$ תתי־מרחבים T־אינווריאנטים כך ש־ $V=\bigoplus_{i=1}^k U_i$.

אזי אם B בסיס המתקבל מאיחוד של בסיסים $B_1,..,B_k$ של $U_1,...,U_k$, אז מטריצת הייצוג של T לפי B תהיה מהצורה:

$$[T]_B = \begin{pmatrix} A_1 & & & & 0 \\ & A_2 & & & \\ & & A_3 & & \\ & & & \ddots & \\ 0 & & & & A_k \end{pmatrix}$$

כאשר לכל $1\leq i \leq k$, מתקיים $A=[T_{|U_i}]_{B_i}$.

A_i היא מטריצה ריבועית מסדר $dim\, U_i$. (ולכן, הן לא כולן בהכרח באותו גודל).

10.2.2 תת־מרחבים ציקליים

הגדרה 10.6 יהיו V מ"ו מעל שדה $\mathbb{F}$, $T:V\to V$ העתקה ליניארית, ו־ $v\in V$.
המרחב:

$$U=span\{v,T(v),T^2(v),T^3(v),...\}$$

נקרא המרחב ה־T־ציקלי שנוצר ע"י v.
(או התת־מרחב הציקלי של T שנוצר ע"י v).
זהו ת"מ T־אינווריאנטי.
במקרה ש־V ממימד סופי, קיים k מינימלי כך ש־ $T^k(v)=span\{v,T(v),...,T^{k-1}(v)\}$,
ובמקרה זה, בעצם $U=span\{v,T(v),...,T^{k-1}(v)\}$.
הסבר:
אם $T^k(v)=\sum_{j=0}^{k-1}c_jT^j(v)$ אז,

$$T^{k+1}(v)=\sum_{j=0}^{k-1}c_jT^{j+1}(v)=\left(\sum_{j=1}^{k-1}c_{j-1}T^j(v)\right)+c_{k-1}T^k(v)$$

ונשים לב ששני המחוברים בביטוי הימני שייכים ל־ $span\{v,...,T^{k-1}(v)\}$.
ובאופן דומה נובע באינדוקציה כי $T^m(v)\in span\{v,...,T^{k-1}(v)\}$ לכל $m\geq k$.

10.2.3 מטריצת ייצוג של צמצום העתקה לבסיס של ת"מ T־ציקלי

אם $U=span\{v,T(v),T^2(v),...,T^{k-1}(v)\}$ הת"מ ה־T־ציקלי שנוצר ע"י v,
נסמן $B=\{T^{k-1}(v),T^{k-2}(v),...,T(v),v\}$. זהו בסיס ל־$U$. (הוא בת"ל בגלל הגדרת k כמינימלי).
אז מטריצת הייצוג של $T_{|_U}$ תהיה מהצורה:

$$[T_{|_U}]_B=\begin{pmatrix} c_1 & 1 & 0 & 0 & & & 0\\ c_2 & 0 & 1 & 0 & & & \\ c_3 & 0 & 0 & 1 & & & \\ \vdots & & & \ddots & \ddots & & \\ \vdots & & & & \ddots & \ddots & \\ \vdots & & & & & \ddots & 1\\ c_k & 0 & 0 & 0 & \cdots & & 0 \end{pmatrix}$$

(כל ה־c_i בצד שמאל הם בגלל שראינו קודם ש$T^k(v)$ הוא צ"ל של הבסיס)

10.3 העתקות נילפוטנטיות

10.3.1 הגדרה

הגדרה 10.7 העתקה ליניארית $T:V\to V$ נקראת **נילפוטנטית** אם קיים k כך ש־ $T^k\equiv 0$.
ה־k המינימלי הזה נקרא **אינדקס הנילפוטנטיות**.

טענה 10.8 אם $T:V\to V$ ה"ל, V מ"ו ממימד סופי מעל שדה **סגור אלגברית*** $\mathbb{F}$,
אז התנאים הבאים שקולים:

1. T נילפוטנטית
2. 0 הוא הע"ע היחיד של T
3. הפולינום האופייני של T הוא $f_T(t)=t^n$ כאשר $n=dimV$
4. הפולינום המינימלי של T הוא $m_T(t)=t^k$ כאשר k אינדקס הנילפוטנטיות.

*נשים לב: התנאים 1,3,4 שקולים גם **ללא** הדרישה לסגירות אלגברית.

הוכחה:

- $1 \Leftrightarrow 4$:

אם $m_T(t)=t^k$ אז $T^k=0$. אבל $T^{k-1}\neq 0$ (כי אחרת היינו מקבלים סתירה למינימליות הדרגה של m_T).

ולכן T נילפוטנטית, מאינדקס k.

מצד שני, אם $T^k=0$ אז $m_T|t^k$

(כי אחרת נכתוב $t^k=q(t)m_T(t)+r(t)$ כאשר $\deg r<k$, ואז $r(T)=T^k-q(T)m_T(T)=0$ בסתירה למינימליות). ולכן $m_T(t)=t^j$ עבור איזשהו j, אבל אז לפי הכיוון שכבר ראינו נובע ש־j הוא אינדקס הנילפוטנטיות.

- $3 \Leftrightarrow 4$:

אם $m_T(t)=t^k$, ראינו ש־ $f_T|m_T^n$ כאשר $n=dim\,V$.

כלומר $f_T(t)|t^{k\cdot n}$. ולכן $f_T(t)=t^j$ עבור איזשהו j. אבל f_T הוא פולינום מתוקן ממעלה n, ולכן $f_T(t)=t^n$.

בכיוון השני, ידוע ש־ $m_T|f_T$ ולכן אם $f_T(t)=t^n$ אז $m_T(t)=t^k$ עבור איזשהו k, ולכן k הוא אינדקס הנילפוטנטיות לפי האמור לעיל.

- $2 \Leftrightarrow 4$:

הע"ע הם השורשים של m_T, f_T.

ולכן אם $f_T=t^n$ אז 0 הוא הע"ע היחיד.

בכיוון השני, אם $\mathbb{F}$ שדה סגור אלגברית אז כל פולינום מתוקן ממעלה n ניתן לכתיבה בצורה:

$$(t-\alpha_1)(t-\alpha_2)\cdot\ldots\cdot(t-\alpha_n)$$

("כל פולינום מתפרק לגורמים ליניאריים").

בפרט, גם $f_T(t)$ ניתן לכתיבה בצורה זו, ואם נניח ש־0 הוא הע"ע היחיד, אז $\alpha_1=\ldots=\alpha_n=0$.

ולכן $f_T(t)=t^n$. ■

טענה 10.9 אם T נילפוטנטית מאינדקס k, ו־$v\in V$ וקטור כך ש־ $T^{k-1}v\neq 0$, אז הת"מ $U=span\{v,T(v),...,T^{k-1}(v)\}$ הוא ת"מ T־ציקלי **<u>ממימד k</u>**. (זה שהוא ציקלי זה ברור, אנו מראים את המימד...)

הוכחה: צריך להוכיח ש־ $v, T(v), ..., T^{k-1}(v)$ הם בת"ל.

אם $\sum_{i=0}^{k-1} a_i T^i(v) = 0$, נפעיל את T^{k-1} על המשוואה ונקבל $a_0 T^{k-1}(v) = 0$ (החזקות שגדולות או שוות ל־k התאפסו). ומכאן, $a_0 = 0$. (כי הנחנו ש־$T^{k-1}v \neq 0$).

באופן דומה, נפעיל T^{k-2} על המשוואה ונקבל $a_0 T^{k-2}v + a_1 T^{k-1}v = 0$ ומכאן $a_1 = 0$.

ממשיכים באינדוקציה ומראים ש־ $a_i = 0$ לכל i. ∎

מסקנה 10.10 אם T נילפוטנטית מאינדקס k אז קיים ת"מ T־ציקלי ממימד k.

טענה 10.11 בתנאי הטענה הקודמת, אם נסמן $B = \{T^{k-1}v, T^{k-2}v, ..., T^2v, Tv, v\}$ (זה בסיס ל־U), אז

$$[T_{|U}]_B = \begin{pmatrix} 0 & 1 & 0 & 0 & & & 0 \\ 0 & 0 & 1 & 0 & & & \\ 0 & 0 & 0 & 1 & & & \\ \vdots & & & \ddots & \ddots & & \\ \vdots & & & & \ddots & \ddots & \\ \vdots & & & & & \ddots & 1 \\ 0 & 0 & 0 & 0 & \cdots & & 0 \end{pmatrix}$$

(זה מקרה פרטי של העובדה שניסחנו קודם לגבי T כללית. פשוט כש־ T נילפוטנטית אז במקום צ"ל של הבסיס אנו מקבלים אפסים, בעמודה השמאלית ביותר).

10.4 צורת ז'ורדן של העתקה נילפוטנטית

10.4.1 בלוק ז'ורדן

הגדרה 10.12 המטריצה: $M_{k\times k}(\mathbb{F}) \ni$ $$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & & & & 0 \\ & \lambda & 1 & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & & & & \lambda & 1 \\ 0 & & & & & \lambda \end{pmatrix}$$ נקראת **בלוק ז'ורדן**

מסדר k עם ע"ע λ.

10.4.2 משפט ההשלמה לסכום ישר עבור ת"מ ציקלי של העתקה נילפוטנטית $V = U \oplus W$

משפט 10.13 יהי V מ"ו ממימד סופי מעל $\mathbb{F}$, $T : V \to V$ נילפוטנטית מאינדקס k.

אם $U \subseteq V$ ת"מ T־ציקלי, אז קיים $W \subseteq V$ ת"מ T־אינווריאנטי כך ש־ $V = U \oplus W$.

הוכחה: מורכבת מ־3 טענות עזר והוכחה באינדוקציה. זה לקח שתי הרצאות, ובסוף נאמר שההוכחה לא כלולה בחומר למבחן. לדעתי שלושת טענות העזר נחשבות חלק מההוכחה, ולכן לא אכלול אותן גם בסיכום.

(זה היה ב־1.6.09) ∎

10.4.3 צורת ז'ורדן של העתקה נילפוטנטית

משפט 10.14 V מ"ו מעל שדה $\mathbb{F}$, $dimV = n$, $T : V \to V$ נילפוטנטית מאינדקס k.
אז קיים פירוק $V = \bigoplus_{i=1}^{m} U_i$ כאשר $U_i \subseteq V$ ת"מ T-ציקליים, $dim\, U_i = r_i$.
וגם: $k = r_1 \geq r_2 \geq ... \geq r_m \geq 1$ (ככה מסדרים אותם).
אי לכך קיים בסיס $B \subseteq V$ כך שמטריצת הייצוג של T תהיה מהצורה:

$$[T]_B = \begin{pmatrix} J_{r_1}(0) & & & & 0 \\ & J_{r_2}(0) & & & \\ & & J_{r_3}(0) & & \\ & & & \ddots & \\ 0 & & & & J_{r_m}(0) \end{pmatrix}$$

(אלה בלוקי ז'ורדן).
מטריצה זו תיקרא **צורת ז'ורדן** של T, והיא נקבעת ביחידות ע"י T.

הוכחה: נוכיח **קיום**:
באינדוקציה על המימד $n = dim\, V$.
$n = 1$: אין מה להוכיח.
שלב האינדוקציה: יהי $U \subseteq V$ ת"מ T-ציקלי ממימד k.
(כנראה שקיים בגלל המסקנה שראינו: "אם T נילפוטנטית מאינדקס k אז קיים ת"מ T-ציקלי ממימד k".)
אם $k = n$ אז $V = U$ וזו צורת ז'ורדן הדרושה (יש רק בלוק אחד).
אחרת, לפי המשפט הקודם קיים ת"מ $W \subseteq V$ שהוא T-אינווריאנטי כך ש- $V = U \oplus W$.
$dim\, W < n$, ולכן לפי הנחת האינדוקציה,
קיים פירוק $W = \bigoplus_{i=2}^{m} U_i$ כאשר U_i הם ת"מ $T_{|W}$-ציקליים.
$dim\, U_i = r_i$ - סדרה יורדת, ומתקיים:
r_2=אינדקס הנילפוטנטיות של $T_{|W}$, שהוא $\leq k = r_1 = dim\, U$.
סה"כ קיבלנו את הפירוק הדרוש: (אם מסמנים $U = U_1$)

$$V = U_1 \oplus U_2 \oplus ... \oplus U_m$$

כאשר U_i T-ציקלי ממימד r_i ומתקיים: $k = r_1 \geq r_2 \geq ... \geq r_m$ ■

10.4.4 הוכחת היחידות (כולל הנוסחה למספר הבלוקים בגודל j)

נוכיח את **היחידות**: **הוכחה:** מספיק להראות כיצד לשחזר את $r_1, ..., r_m$ באמצעות T.
נסמן: $P = [T]_B$. ונחשב את $rank(T^j) = rank(P^j)$ (הדרגה היא אינווריאנט).
נשים לב שעבור מט' בלוקים מהצורה $E = \begin{pmatrix} D_1 & & & 0 \\ & D_2 & & \\ & & \ddots & \\ 0 & & & D_k \end{pmatrix}$, מתקיים: $rank(E) = \sum_{i=1}^{k} rank(D_i)$.

לכל בלוק ז'ורדן $J_r(0)$ מתקיים:

$$J_r(0)^j = \begin{cases} \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ & & & & \ddots & & \vdots \\ & & & & & \ddots & 0 \\ & & & & & & 1 \\ & & & & & & 0 \\ & & & & & & \vdots \\ 0 & & & & & & 0 \end{pmatrix} & if\ 0 \le j \le r-1 \\ 0 & if\ j \ge r \end{cases}$$

(כאשר במטריצה למעלה, יש j פעמים אפסים משמאל לפני ה־1, בשורה העליונה).

ולכן,

$$rank(J_r(0)^j) = \begin{cases} r-j & if\ 0 \le j \le r-1 \\ 0 & if\ j \ge r \end{cases} = (r-j)^+$$

החלק החיובי של מספר ממשי x מוגדר ע"י:

$$x^+ = max(x, 0)$$

ומכאן נובע ש־

$$rank(T^j) = \sum_{i=1}^{m} rank(J_{r_i}(0)^j) = \sum_{i=1}^{m} (r_i - j)^+ = \sum_{\substack{1 \le i \le m \\ r_i - j > 0}} (r_i - j)$$

ובאופן דומה, (לשים לב לאינדקסים!)

$$rank(T^{j+1}) = \sum_{\substack{1 \le i \le m \\ r_i - j > 0}} (r_i - j - 1)$$

ולכן מקבלים ש־ $rank(T^j) - rank(T^{j+1}) =$(מס' ה־$r_i$ כך ש־ $r_i > j$).

למה זה נכון? (הסבר שלי) כי:

$$rank(T^j) - rank(T^{j+1}) = \sum_{\substack{1 \le i \le m \\ r_i - j > 0}} (r_i - j) - (r_i - j - 1) = \left(\sum_{\substack{1 \le i \le m \\ r_i - j > 0}} 1 \right) = |\{1 \le i \le m \ : \ r_i - j > 0\}|$$

כלומר אנו מקבלים את הגודל של הקבוצה של כל ה $1 \le i \le m$ כך ש־$r_i - j > 0$, או ליתר דיוק $r_i > j$.
כלומר כמה בלוקים יש שגודלם גדול מ־j.

נסמן מספר זה ב־c_j.

נסמן ב־d_j את מספר ה־r_i־ים **ששווים ל־**j, אז ברור ש־

$$d_j = c_{j-1} - c_j = \big(rank(T^{j-1}) - rank(T^j)\big) - \big(rank(T^j) - rank(T^{j+1})\big) =$$

$$= rank(T^{j-1}) - 2rank(T^j) + rank(T^{j+1})$$

(כדאי לזכור את הנוסחה, היא חשובה!)

ברור שהמס' $d_1, d_2, d_3, \ldots$ קובעים את $r_1, r_2, \ldots$ ולכן הראינו שה־r_i נקבעים ביחידות. ■

10.5 צורת ז'ורדן של העתקה כללית

10.5.1 קיום

תזכורת: המטריצה: $M_{r\times r}(\mathbb{F}) \ni \begin{pmatrix} \lambda & 1 & & & & 0 \\ & \lambda & 1 & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & & & & \lambda & 1 \\ 0 & & & & & \lambda \end{pmatrix} = J_r(\lambda)$ נקראת **בלוק ז'ורדן** מסדר r עם ע"ע λ.

עבור העתקות נילפוטנטיות, ראינו שכל הע"ע הם 0. הראינו שלכל העתקה כזאת יש מטריצת ייצוג מהצורה:

$$\begin{pmatrix} J_{r_1}(0) & & & & 0 \\ & J_{r_2}(0) & & & \\ & & J_{r_3}(0) & & \\ & & & \ddots & \\ 0 & & & & J_{r_k}(0) \end{pmatrix}$$

וצורה זו יחידה עד כדי קביעת סדר הבלוקים.

(לשם נוחות דורשים בד"כ $r_1 \geq r_2 \geq \ldots \geq r_k$).

כעת נעבור למטריצה כללית, לא בהכרח נילפוטנטית:

משפט 10.15 לכל העתקה ליניארית $T: V \to V$, כאשר V מ"ו ממימד סופי מעל שדה סגור אלגברית $\mathbb{F}$, קיימת צורת ז'ורדן:

$$[T]_B = \begin{pmatrix} \boxed{\begin{matrix} J_{r_{1,1}}(\lambda_1) & & & \\ & J_{r_{1,2}}(\lambda_1) & & \\ & & \ddots & \\ & & & J_{r_{1,b_1}}(\lambda_1) \end{matrix}} & & & \\ & \boxed{\begin{matrix} J_{r_{2,1}}(\lambda_2) & & \\ & \ddots & \\ & & J_{r_{2,b_2}}(\lambda_2) \end{matrix}} & & \\ & & \ddots & \\ & & & \boxed{\begin{matrix} J_{r_{k,1}}(\lambda_k) & & \\ & \ddots & \\ & & J_{r_{k,b_k}}(\lambda_k) \end{matrix}} \end{pmatrix}$$

כאשר $\lambda_1, ..., \lambda_k$ הם הע"ע **השונים** של T.

אם דורשים שמתקיים $r_{i,1} \geq r_{i,2} \geq ... \geq r_{i,b_i}$ לכל $1 \leq i \leq k$, אז צורה זו יחידה עד כדי בחירת סדר הע"ע $\lambda_1, ..., \lambda_k$.

הוכחה: בערך 4 טענות עזר ועוד כשני דפים של הוכחה. כל זה לא למבחן (לדעתי זה כולל את טענות העזר), ולכן לא אכלול את זה בסיכום. זה היה ב־8.6.09. ■

10.5.2 סיכום מה שראינו

$T : V \to V$ (V מ"ו מעל $\mathbb{F}$ סגור אלגברית).

$\lambda_1, ..., \lambda_k$ הע"ע השונים של T.

כל ע"ע λ_i הוא מריבוי אלגברי n_i, כלומר $f_T(t) = \prod_{i=1}^{k} (t - \lambda_i)^{n_i}$.

לכל $1 \leq i \leq k$, נסמן ב־ a_i את האינדקס **המינימלי**

כך שמתקיים $Ker(T - \lambda_i I)^j = Ker(T - \lambda_i I)^{a_i}$ לכל $j \geq a_i$.

ומסמנים: $U_i = Ker(T - \lambda_i I)^{a_i}$.

אז יש פירוק לסכום ישר $V = \bigoplus_{i=1}^{k} U_i$. כל U_i הוא T־אינווריאנטי.

לכל $1 \leq i \leq k$, $(T - \lambda_i I)_{|U_i}$ היא נילפוטנטית (מסדר a_i).

(במילים אחרות, $T_{|U_i} = \lambda_i I + N_i$ כאשר $N_i : U_i \to U_i$ נילפוטנטית).

כמו כן, לכל $1 \leq i \neq j \leq k$, מתקיים ש־ $(T - \lambda_i I)_{|U_j}$ הפיכה.

מכאן מקבלים את קיום צורת ז'ורדן. לכל $1 \leq i \leq k$, **יהי** B_i **בסיס של** U_i, כך שמטריצת הייצוג של ההעתקה הנילפוטנטית $N_i = T_{|U_i} - \lambda_i I$ **בבסיס** B_i תהיה מטריצה נילפוטנטית מהצורה:

$$\begin{pmatrix} J_{r_{i,1}}(0) & & & \\ & J_{r_{i,2}}(0) & & \\ & & \ddots & \\ & & & J_{r_{i,d_i}}(0) \end{pmatrix}.$$

אז מט' הייצוג של $T_{|U_i}$ היא: $[T_{|U_i}]_{B_i} = \begin{pmatrix} J_{r_{i,1}}(\lambda_i) & & & \\ & J_{r_{i,2}}(\lambda_i) & & \\ & & \ddots & \\ & & & J_{r_{i,d_i}}(\lambda_i) \end{pmatrix}$.

ולכן מטריצת הייצוג של T בבסיס $B=\bigcup_{i=1}^{k}B_i$ של V היא:

$$[T]_B=\begin{pmatrix}[T_{|_{U_1}}]_{B_1} & & & \\ & [T_{|_{U_2}}]_{B_2} & & \\ & & \ddots & \\ & & & [T_{|_{U_k}}]_{B_k}\end{pmatrix}$$

וזו צורת ז'ורדן.

10.5.3 הוכחת היחידות

נוכיח את היחידות של צורת ז'ורדן: **הוכחה:** עבור מטריצה A (או העתקה ליניארית) נסמן:

$$n(A,r)=rank(A^{r-1})-2rank(A^r)+rank(A^{r+1})$$

(זוהי הנוסחה שראינו מקודם, עבור מספר הבלוקים מגודל j של צורת ז'ורדן של מט' נילפוטנטית).

אז מס' הבלוקים מהצורה $J_r(\lambda_i)$ בצורת ז'ורדן שמצאנו שווה ל־ $n\left(\left[(T-\lambda_i I)_{|_{U_i}}\right]_{B_i}\ ,\ r\right)$.

אבל זה גם שווה ל־ $n\left([T-\lambda_i I]_B\ ,\ r\right)$, בגלל שהדרגה של מט' בלוקים $\begin{pmatrix}D_1 & & & \\ & D_2 & & \\ & & \ddots & \\ & & & D_m\end{pmatrix}$

שווה לסכום הדרגות.

ועבור כל אחד מהבלוקים $\left[(T-\lambda_i I)_{|_{U_j}}\right]_{B_j}$ (עבור $j\neq i$) שהוא הפיך,

מתקיים $n\left(\text{the block}\ ,\ r\right)=0$.

אבל, הדרגה היא אינווריאנט ליחס הדמיון, ולכן $n\left([T-\lambda_i I]_B\ ,r\right)$ זה בעצם $n\left(T-\lambda_i I\ ,\ r\right)$.

זה לא תלוי בבחירת הבסיס ולכן מוכיח את היחידות. (!!!) ∎

10.5.4 סיכום צורת ז'ורדן

נתונה העתקה $T:V\to V$, V מ"ו ממימד סופי מעל סדה סגור אלגברית $\mathbb{F}$.

$\lambda_1,...,\lambda_k$ הע"ע השונים של T.

הפולינום האופייני $f_T(t)=\prod_{i=1}^{k}(t-\lambda_i)^{n_i}$, כאשר n_i הריבוי האלגברי.

הפולינום המינימלי $m_T(t)=\prod_{i=1}^{k}(t-\lambda_i)^{a_i}$ (כאשר a_i הוגדר בסעיף הקודם).

<u>אז מתקיימים הקשרים הבאים:</u>

- הריבוי האלגברי n_i שווה ל־ $r_{i,1}+r_{i,2}+...+r_{i,d_i}$ סכום הגדלים של בלוקי ז'ורדן המתאימים לע"ע λ_i.
- החזקה a_i בפולינום המינימלי $= r_{i,1}$ גודל הבלוק הגדול ביותר המתאים ל־λ_i.
  זה אינדקס הנילפוטנטיות של $(T-\lambda_i I)_{|Ker(T-\lambda_i I)^{n_i}}$. (הסבר: $(J_r(\lambda)-\lambda I)^j=0$ אם"ם $j\geq r$)
- באופן כללי, מס' הבלוקים בגודל r המתאימים לע"ע λ_i הוא $n(T-\lambda_i I\ ,\ r)$, כאשר
  $$n(A,r)=rank(A^{r-1})-2rank(A^r)+rank(A^{r+1})$$

10.6 שימושים ומסקנות מצורת ז'ורדן

10.6.1 הוכחה למשפט קיילי־המילטון

משפט 10.16 לכל $A \in M_n(F)$, מתקיים: $f_A(A) = 0$ (קיילי־המילטון).

הוכחה: (זה עובד רק לשדות סגורים אלגברית...)
אם צורת ז'ורדן של T היא:

$$[T]_B = \begin{pmatrix} J_{r_1}(\lambda_1) & & & & \\ & J_{r_2}(\lambda_2) & & & \\ & & J_{r_3}(\lambda_3) & & \\ & & & \ddots & \\ & & & & J_{r_m}(\lambda_m) \end{pmatrix}$$

(כאשר λ_i לא בהכרח שונים ־ נשים לב שזה לא כמו בהצגה שראינו קודם ־ **הסימון כאן שונה**).
ו־$\lambda_1, ..., \lambda_m$ הע"ע של T (כולל ריבוי, כנראה, כלומר יש חזרות).
נסמן $P = [T]_B$. אז $f_T(t) = f_P(t)$. (הפולינום האופייני הוא אינווריאנט לדמיון).
מספיק לבדוק ש־ $f_P(P) = 0$.
אבל,

$$f_P(t) = \prod_{i=1}^{m} (t - \lambda_i)^{r_i}$$

(זה כי $f_{J_r(\lambda)} = (t - \lambda)^r$).
את זה אפשר לכתוב בצורה:

$$= \prod_{j=1}^{k} (t - \alpha_j)^{b_j}$$

כאשר $\alpha_1, ..., \alpha_k$ הם הע"ע <u>**השונים**</u> של T.
ואז, b_j=(סכום הגדלים של בלוק ז'ורדן עם ע"ע α_j).
אבל אז, ("קל לבדוק ש")

$$f_P(P) = \begin{pmatrix} f_P(J_{r_1}(\lambda_1)) & & & \\ & f_P(J_{r_2}(\lambda_2)) & & \\ & & \ddots & \\ & & & f_P(J_{r_k}(\lambda_k)) \end{pmatrix}$$

ומתקיים: לכל $1 \le i \le m$,
$f_P(J_{r_i}(\lambda_i)) = 0$ בגלל הגורם $(t - \alpha_j)^{b_j}$ עבור אותו j כך ש־ $\lambda_i = \alpha_j$. (כי $b_j \ge r_i$).

למה $f_P(J_{r_i}(\lambda_i)) = 0$? (הסבר שלי) כמובן שעבור הגורם $(t - \alpha_j)^{b_j}$ עבור אותו j כך ש$\lambda_i = \alpha_j$, נקבל שהאלכסון הראשי בבלוק $J_{r_i}(\lambda_i)$ מתאפס. כמו כן, מכיוון ש־$b_j \ge r_i$, נקבל שאנו מעלים את כל הבלוק הזה בחזקה גדולה מספיק שהוא יתאפס (כי הוא הפך לנילפוטנטי ־ האלכסון המשני הוא אחדים והראשי אפסים, ולכן זוהי מט' נילפוטנטית). לא בטוח שזה נכון, זה הסבר שלי.

■

10.6.2 עוד מסקנות

טענה 10.17 מטריצה ניתנת לליכסון אס"ם הפולינום המינימלי שלה מתפרק לגורמים ליניאריים שונים.

הוכחה: בזכות צורת ז'ורדן קיבלנו הוכחה חדשה לזה:

זה פשוט אומר שכל הבלוקים הם מגודל 1, כלומר שכל ה־a_i־ים שווים 1. ■

מסקנות נוספות:

- הבנה של מחלקות הדמיון של מטריצות:

  מצאנו נציג בכל מחלקת שקילות, וזה מאפשר להכריע מתי שתי מטריצות הן דומות ־ **אם ורק אם** יש להן אותה צורת ז'ורדן.

  במילים אחרות, אוסף האינווריאנטים $n(T-\lambda I, r)$ הוא אוסף **שלם** של אינווריאנט.

- חישוב חזקות של מטריצה: $A = PJP^{-1} \Leftarrow A^n = PJ^nP^{-1}$.

  את J^n אפשר לחשב בעזרת פיתוח הבינום של ניוטון.

- שורש של מטריצה:

  אם A מטריצה ריבועית מעל $\mathbb{C}$ אזי קיימת מטריצה B כך ש־ $A = B^2$ (כלומר $B = "\sqrt{A}"$).

  (בונים שורש לבלוק ז'ורדן $J_r(\lambda)$).

משפט 10.18 כל מטריצה ריבועית A דומה למטריצה המשוחלפת A^t.

הוכחה: ההוכחה שהיתה בהרצאה לא היתה ברורה לי, אז היא לא פה. ■

10.7 אלגוריתמים (מהתרגול)

10.7.1 אלגוריתם למציאת בסיס מז'רדן למטריצה נילפוטנטית

אנחנו מעל $\mathbb{C}$, במ"ו V.

כמו כן, נסמן: $d_i = rank(T^{i-1}) - 2rank(T^i) + rank(T^{i+1})$. זהו מספר הבלוקים מגודל i שנקבל.

נניח T ט"ל נילפוטנטית מאינדקס נילפוטנטיות k.

נסמן: $W_i = ker(T^i)$. (אז כמובן $W_k = V$ כי T נילפוטנטית מאינדקס k).

B_1 יהיה בסיס ל־W_1.

נרחיב אותו עם תוספת שנסמן ב־B_2, לבסיס ל־W_2. כלומר $B_1 \cup B_2$ (איחוד זר) בסיס ל־W_2.

וכך נמשיך: $B_1 \cup B_2 \cup ... \cup B_i$ (איחוד זר) יהיה בסיס ל־W_i.

אז $B_1 \cup ... \cup B_k$ הוא בסיס ל־V.

נבנה ממנו בסיס מז'רדן ע"י כך שנחליף כל B_i ב־C_i מהסוף להתחלה: (כאשר C_i מוגדר להלן:)

- בהתחלה נגדיר: $C_k := B_k$. ואח"כ נמשיך באופן הבא:

כעת אנו רוצים לייצר את C_{k-1}. ראשית, ניקח את $T(C_k)$, כלומר נפעיל T על כל C_k. ואז, ייתכן שצריך להוסיף לזה עוד וקטורים, בהתאם למצב:

- אם $d_{k-1}=0$ אז נגדיר: $C_{k-1}:=T(C_k)$.

(כלומר לא נוסיף עוד וקטורים לC_{k-1} מעבר ל־$T(C_k)$ שכבר חישבנו בתחילת השלב הזה)

ואז, $B_1\cup B_2\cup...\cup B_{k-2}\cup C_{k-1}$ הוא בסיס ל־W_{k-1}.

- אם $d_{k-1}>0$ אז נשלים את $B_1\cup...\cup B_{k-2}\cup T(C_k)$ לבסיס ל־W_{k-1} (צריך להוסיף d_{k-1} וקטורים).

C_{k-1} יהיה הוקטורים ב־$T(C_k)$ יחד עם d_{k-1} הוקטורים של ההשלמה. (את ההשלמה אפשר לקחת מתוך B_{k-1}).

ואז $B_1\cup...\cup B_{k-2}\cup C_{k-1}$ הוא בסיס ל־W_{k-1}.

וכך ממשיכים. כלומר הגדרנו את C_{k-1}, ועכשיו עלינו להגדיר את C_{k-2} וכן הלאה.

בסוף מקבלים בסיס $C=C_1\cup...\cup C_k$ ל־V.

$C_1\cup...\cup C_i$ הוא בסיס ל־W_i.

כל וקטור $v\in C_i$ מקיים $v\in W_i\backslash W_{i-1}$.

ולפי האלגוריתם, $T(v),T^2(v),...,T^{i-1}(v)$ גם כן נמצאים בבסיס C.

אם נשנה ל־C את הסדר, נקבל ש־C הוא בסיס מז'רדן.

איך נסדר את הבסיס? (כדי לקבל מטריצת בלוקים יפה)

1. עבור כל $v\in C_k$, ניקח את $T^{k-1}v,...,Tv,v$
2. עבור כל $v\in C_{k-1}$ שעוד לא הופיע, לוקחים את $T^{k-2}v,...,Tv,v$
3. וכן הלאה...

רושמים מטריצה P שעמודותיה הם הוקטורים הנ"ל, באופן הבא:

$$\begin{pmatrix}\left|\begin{array}{c|c|c} | & | & | \\ C_k's\,vectors\,(in\,that\,order) & C_{k-1}'s\,vectors\,(in\,that\,order) & \cdots \\ | & | & | \end{array}\right|\end{pmatrix}$$

(הכוונה היא לרשום את הוקטורים מהשלב האחרון, באותו סדר).

(בתכל'ס אם יש לנו r פעמים הפעלת T על וקטור, אז פשוט נשים אותו לפי חזקות, משמאל לימין: $T^rv,...Tv,v$).

נקבל ש־ $P^{-1}AP$ היא מטריצה בצורת ז'ורדן.

10.7.2 אלגוריתם למציאת צורת ז'ורדן של מטריצה כללית

(בסימונים שהשתמשנו בסעיף "סיכום מה שראינו").

צריך למצוא בסיס מז'רדן לכל $U_i=Ker(T-\lambda_iI)^{a_i}$.

$(T-\lambda_iI)_{|Ker(T-\lambda_iI)^{a_i}}$ היא נילפוטנטית מאינדקס נילפוטנטיות a_i. נמצא לה בסיס מז'רדן B.

כלומר, $\left[(T-\lambda_iI)_{|Ker(T-\lambda_iI)^{a_i}}\right]_B$ היא בצורת ז'ורדן (עם 0 על האלכסון).

ועכשיו, $T_{|Ker(T-\lambda_iI)^{a_i}}=(T-\lambda_iI)_{|Ker(T-\lambda_iI)^{a_i}}+(\lambda_iI)_{|Ker(T-\lambda_iI)^{a_i}}$,

ולכן, בבסיס B, נקבל ש־ $\left[T_{|Ker(T-\lambda_iI)^{a_i}}\right]_B$ היא מטריצה בצורת ז'ורדן (עם λ_i על האלכסון).

<u>**האלגוריתם למציאת בסיס מז'רדן ל־ $(T-\lambda_iI)_{|Ker(T-\lambda_iI)^{a_i}}$:**</u>

ראשית נמצא את כל הλ_i־ים, כלומר הע"ע של המטריצה,

ואז נצטרך למצוא בסיס מז'רדן ל$(T-\lambda_i I)_{|Ker(T-\lambda_i I)^{a_i}}$.

זה זהה (עד כמה שאני הבנתי, לפחות!) למציאת בסיס מז'רדן של מטריצה נילפוטנטית.

כלומר, נסמן $W_1 = Ker(T-\lambda_i I)$, $W_2 = (Ker(T-\lambda_i I)^2$, ... , $W_{a_i} = Ker(T-\lambda_i I)^{a_i}$.

אז $W_1 \subsetneqq W_2 \subsetneqq ... \subsetneqq W_{a_i}$.

(בפועל: אפשר לדעת מהו a_i פשוט ע"י כך שנראה: $W_{a_i} = W_{a_i+1}$,

כלומר $Ker(T-\lambda_i I)^{a_i} = Ker(T-\lambda_i I)^{a_i+1}$).

וממשיכים באלגוריתם למציאת בסיס מז'רדן למט' נילפוטנטית.

נשים לב שבניגוד למציאת בסיס מז'רדן של נילפוטנטית, באלגוריתם הזה אנו לא מאפסים בהכרח את המטריצה, אלא פשוט מגיעים לשלב שבו הגרעין לא משתנה. זה עלול קצת לבלבל בהתחלה.

11 תבניות ביליניאריות

11.1 הקדמה

11.1.1 הגדרות ובסיס

הגדרה 11.1 יהיו V, W מ"ו מעל שדה $\mathbb{F}$.

פונקציה $f : V \times W \to \mathbb{F}$ תיקרא **תבנית ביליניארית** אם היא מקיימת את התכונות:

1.
$$f(v+v', w) = f(v,w) + f(v', w)$$

2.
$$f(\alpha v, w) = \alpha f(v,w)$$

3.
$$f(v, w+w') = f(v,w) + f(v, w')$$

4.
$$f(v, \alpha w) = \alpha f(v,w)$$

לכל $v \in V$, $w \in W$, $\alpha \in \mathbb{F}$.

דוגמאות:

- מכפלה פנימית במרחב אוקלידי (ממ"פ ממשי).
- ההעתקה $f : V \times V^* \to \mathbb{F}$ המוגדרת ע"י $f(v, \varphi) = \varphi(v)$. זו תבנית ביליניארית.

- אם V,W מ"ו, ונתונים פונקציונלים $\varphi \in V^*$, $\psi \in W^*$,
אז אפשר להגדיר $f : V \times W \to \mathbb{F}$ ע"י $f(v,w) = \varphi(v)\psi(w)$.
- בהינתן מטריצה $A \in M_{m\times n}(\mathbb{F})$,
אפשר להגדיר תבנית בילינארית $f : \mathbb{F}^m \times \mathbb{F}^n \to \mathbb{F}$ ע"י $f(v,w) = v^t A w$.

באופן מפורש, אם $\begin{matrix} i = 1...m \\ j = 1...n \end{matrix}$ $A = (a_{ij})$, $v = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}$, $w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$ אז:

$$f(v,w) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} v_i w_j$$

ולמעשה כל תבנית בילינארית ניתנת ל"הצגה" בצורה כזו.

הגדרה 11.2 מטריצת ייצוג של תבנית בילינארית:
יהיו V,W מרחבים וקטוריים ממימדים סופיים $dimV = m$, $dimW = n$.
ויהיו $B = \{v_1,...,v_m\}$, $C = \{w_1,...,w_n\}$ בסיסים של V,W בהתאמה.
בהינתן תבנית בילינארית $f : V \times W \to \mathbb{F}$, **מטריצת הייצוג** שלה בבסיסים B,C
היא המטריצה $A = (a_{ij}) \in M_{m\times n}(\mathbb{F})$, המוגדרת ע"י: $a_{ij} = f(v_i, w_j)$.
מסמנים: $A = [f]_{B,C}$.

11.1.2 עוד על ייצוג בעזרת מטריצות

טענה 11.3 בסימונים הנ"ל, $f(v,w) = [v]_B^t \cdot [f]_{B,C} \cdot [w]_C$.

הוכחה: אם $v = \sum_{i=1}^m b_i v_i$ ו־ $w = \sum_{i=1}^n c_i w_i$, אז

$$f(v,w) = f(\sum_{i=1}^{m} b_i v_i, \sum_{j=1}^{n} c_j w_j) = \sum_{i=1}^{m}\sum_{j=1}^{n} b_i c_j f(v_i,w_j) = \sum_{i=1}^{m}\sum_{j=1}^{n} b_i c_j \cdot a_{ij} =$$

$$= \begin{pmatrix} b_1 & b_2 & \cdots & b_m \end{pmatrix} \cdot \begin{pmatrix} & & \\ & a_{ij} & \\ & & \end{pmatrix}_{i,j} \cdot \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = [v]_B^t \cdot [f]_{B,C} \cdot [w]_C$$

■

מסקנה 11.4 :1
הערכים $f(v_i,w_j)$ $\left(\begin{matrix} 1 \le i \le m \\ 1 \le j \le n \end{matrix} \right)$ קובעים את התבנית f על כל זוג וקטורים v,w.

מסקנה 11.5 :2
אוסף התבניות הבילינאריות מעל $V \times W$ מהווה מרחב וקטורי. ההתאמה $f \to [f]_{B,C}$ מגדירה איזומורפיזם בין מרחב התבניות הבילינאריות מעל V,W ובין $M_{m\times n}(\mathbb{F})$.
(המ"ו הנ"ל מוגדר עם פעולות חיבור: $(f+g)(v,w) = f(v,w) + g(v,w)$ וכפל בסקלר: $(\alpha f)(v,w) = \alpha f(v,w)$).

הסבר:

ההתאמה $f \to [f]_{B,C}$ היא חח"ע בגלל מסקנה 1, ואם $[f]_{B,C} = [g]_{B,C}$ אז $f(v_i, w_j) = g(v_i, w_j)$ לכל i, j ולכן $f = g$.

ההתאמה היא על - אם $A \in M_{m\times n}(\mathbb{F})$ אז נגדיר $f(v, w) = [v]_B^t \cdot A \cdot [w]_C$. זוהי תבנית בילינארית:

מתקיים $[v]_B^t \cdot A \cdot [w]_C = [v]_B^t \cdot [f]_{B,C} \cdot [w]_C$ לכל v, w, ולכן $[f]_{B,C} = A$.

נותר לבדוק ליניאריות - הושאר כתרגיל.

שאלה: מה קורה כשמחליפים בסיסים?

משפט 11.6 יהיו V, W מ"ו מעל שדה $\mathbb{F}$ ממימדים m, n בהתאמה.

תהי $f : V \times W \to \mathbb{F}$ תבנית בילינארית. יהיו B, B' שני בסיסים של V, C, C' שני בסיסים של W.

אזי מתקיים (כאשר M היא מט' מעבר בין בסיסים):

$$[f]_{B',C'} = (M_B^{B'})^t \cdot [f]_{B,C} \cdot M_C^{C'}$$

הוכחה: מתקיים:

$$f(v, w) = [v]_B^t \cdot [f]_{B,C} \cdot [w]_C = \left(M_B^{B'} \cdot [v]_{B'}\right)^t \cdot [f]_{B,C} \cdot \left(M_C^{C'} [w]_{C'}\right) =$$

$$= [v]_{B'}^t \cdot \left((M_B^{B'})^t \cdot [f]_{B,C} \cdot M_C^{C'}\right) \cdot [w]_{C'}$$

ומצד שני, גם מתקיים:

$$f(v, w) = [v]_{B'}^t \cdot [f]_{B',C'} \cdot [w]_{C'}$$

ומכיוון שיש את השויון הזה לכל v, w, נקבל שהטענה נכונה. ∎

מסקנה 11.7 המקרה שבו $V = W$, $B = C$, $B' = C'$ הוא מעניין במיוחד.

במקרה זה, נסמן $[f]_B = [f]_{B,B}$ ואז מתקיים:

$$[f]_{B'} = P^t [f]_B P$$

כאשר $P = M_B^{B'}$.

11.1.3 מטריצות חופפות

הגדרה 11.8 אומרים ששתי מטריצות ריבועיות $A,B\in M_n(\mathbb{F})$ הן **חופפות** אם קיימת מטריצה P **הפיכה** כך ש־ $A=P^tBP$.

טענה 11.9 יחס החפיפה הוא ***יחס שקילות***. כלומר מתקיים:

<u>רפלקסיבי:</u> כל מטריצה A חופפת לעצמה: $A=I^tAI$.

<u>סימטרי:</u> אם A חופפת ל־B אז B חופפת ל־A. כי אם $A=P^tBP$ אז $B=(P^{-1})^tAP^{-1}$.

<u>טרנזיטיבי:</u> אם A חופפת ל־B, B חופפת ל־C, אז A חופפת ל־C.

כי אם $A=P^tBP$, $B=Q^tCQ$, אז $A=(QP)^tC(QP)$.

המטרה: להבין את יחס החפיפה ־ מתי שתי מטריצות הן חופפות? מה הם אינווריאנטים מעניינים? איך למצוא נציג "פשוט" בכל מחלקת שקילות?

11.1.4 מטריצות שקולות

הערה 11.10 אפשר להגדיר יחס שקילות יותר פשוט אם לא דורשים שהבסיסים $B=C$ ו־ $B'=C'$.

הגדרה 11.11 אם $A,B\in M_{m\times n}(\mathbb{F})$, נגיד שהן **שקולות**

אם קיימות מטריצות הפיכות $P\in M_{m\times m}(\mathbb{F})$, $Q\in M_{n\times n}(\mathbb{F})$, כך ש־ $A=P^tBQ$.

(נשים לב שהדרישה לשיחלוף של P היא לא נחוצה ולא משנה).

טענה 11.12 $A,B\in M_{m\times n}(\mathbb{F})$ שקולות אם"ם הן מאותה דרגה.

הוכחה: אם A,B מאותה הדרגה, אז כפל במטריצה הפיכה שומר על הדרגה, ולכן זה נותן כיוון אחד.

בכיוון השני, נניח $rank(A)=k$, ונניח גם ש־ $m\le n$.

ננסה להביא את A לצורה: $\left(\begin{array}{c|c} I_k & 0 \\ \hline 0 & 0\end{array}\right)$ (כלומר בלוק אחד שהוא I_k ואפסים בשאר המטריצה),

ע"י פעולות שורה ועמודה אלמנטריות.

ראשית, ע"י פעולות שורה אלמנטריות, מביאים את A לצורת מדרגות:
$$\left(\begin{array}{cccccccc} 0 & 0 & 1 & & & & * & * \\ & 0 & 0 & 1 & & & & * \\ & & 0 & 0 & 0 & 1 & & \\ \vdots & & & & & & \ddots & \\ 0 & \cdots & & & & & & 1 \\ \hline 0 & \cdots & & & & & \cdots & 0 \\ 0 & & & & & & & 0 \end{array}\right)$$

כאשר בלוק האפסים התחתון הוא בגובה $m-k$. (איפה שיש קו אופקי, זה "בלוק" של אפסים מתחת...)

עכשיו, בעזרת פעולות עמודה דוחפים את ה־1־ים לצד שמאל, ומקבלים:
$$\left(\begin{array}{cccc|ccc} 1 & & & & & & \\ & 1 & & & & & * \\ & & \ddots & & & & \\ & & & 1 & & & \\ \hline 0 & \cdots & & 0 & \cdots & & 0 \\ 0 & & & & & & 0 \end{array}\right)$$

ולבסוף ע"י ביצוע עוד פעולות עמודה,

ניתן לאפס את האיברים מימין לאחדים ולקבל $\left(\begin{array}{cccc|cc} 1 & & & & 0 & 0 \\ & 1 & & & \vdots & \vdots \\ & & \ddots & & \vdots & \vdots \\ & & & 1 & 0 & 0 \\ \hline 0 & \cdots & & 0 & \cdots & 0 \\ 0 & & & & & 0 \end{array}\right).$

ואז, P מתארת את פעולות השורה, ו־Q מתארת את פעולות העמודה. ■

מטריצות שקולות לא כל כך מעניינות אותנו, אז נמשיך להתעסק במטריצות חופפות מעתה ואילך.

11.2 ליכסון ותבניות בילינאריות סימטריות

11.2.1 תבניות בילינאריות סימטריות

$A, B \in M_n(\mathbb{F})$ חופפות אם קיימת P הפיכה כך ש־ $A = P^tBP$.
זה אומר ש־A ו־B מייצגות את אותה תבנית בילינארית בשני בסיסים שונים.

הגדרה 11.13 תבנית בילינארית $f : V \times V \to \mathbb{F}$ נקראת **סימטרית אם** $f(u,v) = f(v,u)$ לכל $v, u \in V$.

למה 11.14 תבנית f היא סימטרית
אם"ם קיים בסיס B כך ש־ $[f]_B$ היא סימטרית,
אם"ם **לכל** בסיס B, $[f]_B$ היא סימטרית.

הוכחה: מתקיים:

$$f(u,v) = [u]_B^t \cdot [f]_B \cdot [v]_B$$

ומצד שני, (ונזכור שסקלר שווה למט' המשוחלפת של עצמו)

$$f(v,u) = [v]_B^t \cdot [f]_B[u]_B = [u]_B^t \cdot [f]_B^t[v]_B$$

וזה נותן את שתי הגרירות ־
אם f סימטרית נקבל ש־ $[f]_B = [f]_B^t$ לכל בסיס B.
ולהיפך, אם יש בסיס אחד B כך ש־ $[f]_B = [f]_B^t$ אז נקבל ש־ $f(u,v) = f(v,u)$ לכל $v, u \in V$. ■

הערה 11.15 יחס החפיפה שומר על סימטריה של מטריצות.
אם B סימטרית, ו־$A = P^tBP$ אז גם A סימטרית, כי $A^t = P^tB^tP = P^tBP$.

11.2.2 ליכסון ותנאים לליכסון

הגדרה 11.16 תבנית $f : V \times V \to \mathbb{F}$ נקראת **ניתנת לליכסון**
אם קיים בסיס $B = \{v_1, ..., v_n\}$ כך ש־ $[f]_B$ אלכסונית,
או במילים אחרות כך ש־ $f(v_i, v_j) = 0$ אם $i \neq j$.

מסקנה 11.17 תנאי הכרחי לליכסון הוא ש־f סימטרית.

הגדרה 11.18 מציין של שדה

המציין של שדה $\mathbb{F}$ הוא המספר הטבעי המינימלי k כך ש־ $\underbrace{1+1+...+1}_{k\,times}=0$, או 0 אם אין כזה.

משפט 11.19 אם המציין של השדה $\mathbb{F}$ אינו 2, אז כל תבנית בילינארית $f:V\times V\to\mathbb{F}$ סימטרית ניתנת ללכסון.

הוכחה: באינדוקציה על $dimV=n$. עבור $n=1$, אין מה להוכיח.

אם הוכחנו עבור $n-1$, נוכיח עבור n.

ראשית, אם התבנית היא זהותית 0, אז אין מה להוכיח.

אחרת, קיים וקטור $v_1\in V$ כך ש־ $f(v_1,v_1)\neq 0$.

הסבר: אם נניח בשלילה שלכל $x\in V$ מתקיים $f(x,x)=0$, אז יתקיים לכל $v,w\in V$:

$$0=f(v+w,v+w)=f(v,v)+f(w,w)+2f(v,w)$$

(המעבר של $f(v,w)+f(w,v)=2f(v,w)$ הוא כי הנחנו ש־f סימטרית).
ולכן, היינו מקבלים ש־$f(v,w)=0$ לכל v,w, אלא אם כן המציין של השדה הוא 2.
בסתירה לכך שהנחנו ש־f היא לא זהותית 0.

נסמן: $W=\{u\in V\ :\ f(v_1,u)=0\}=\text{"}v_1^{\perp}\text{"}$.

זה שווה ל־$Ker(f_{v_1})$ אם מגדירים $f_{v_1}(u)=f(v_1,u)$.

W הוא תת־מרחב של V. הוא ממימד n או $n-1$ (כי תמונה של פונקציונל היא 0 או 1, וממשפט המימדים).

אבל בעצם הוא ממימד $n-1$, כי $f_{v_1}(v_1)\neq 0$.

לפי הנחת האינדוקציה, התבנית $f_{|W\times W}$ היא ניתנת לליכסון,

ולכן קיים בסיס $B'=\{v_2,v_3,...,v_n\}$ (כנראה של W, זה לא נאמר בהרצאה אבל מתבקש),

כך ש־ $[f_{|W\times W}]_{B'}$ אלכסונית, כלומר $f(v_i,v_j)=0$ לכל $2\leq i\neq j\leq n$.

עכשיו נסמן $B=\{v_1,v_2,...,v_n\}$. זהו בסיס, כי אם לא,

אז היינו מקבלים ש־ v_1 ניתן להצגה כצירוף ליניארי $v_1=\sum_{i=2}^n a_iv_i$, אבל אז,

$$0\neq f(v_1,v_1)=f(v_1,\sum_{i=2}^n a_iv_i)=0$$

כאשר אי השויון משמאל הוא ידוע לנו מהגדרת v_1, והשויון לאפס מימין הוא כי $\sum_{i=2}^n a_iv_i\in W$.

כמו כן, ראינו ש־ $f(v_i,v_j)=0$ עבור $2\leq i\neq j\leq n$, וגם $f(v_1,v_i)=0$ לכל $2\leq i\leq n$, כי $v_i\in W$. ■

11.2.3 דוגמא

הערה: יש דרך הרבה יותר ידידותית לליכסון ע"י חפיפה ־ זה בסוף הפרק, לקוח מהתרגול.

$S=\begin{pmatrix}2&1&3\\1&0&1\\3&1&3\end{pmatrix}$ מטריצה סימטרית המגדירה תבנית בילינארית סימטרית: $(x,y)\to x^tSy$ על $\mathbb{R}^3$.

נלכסן אותה (במובן של חפיפה).

נתחיל מוקטור $v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ כך ש־ $v_1^t S v_1 = 2 \neq 0$.

אחרי זה נתבונן ב־

$$W = "v_1^{\perp}" = \{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \cdot S \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 2x + y + 3z = 0 \}$$

נמצא בסיס ל־W, למשל $\{ \begin{pmatrix} 3 \\ 0 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} \}$. ונבחר את v_2 להיות $v_2 = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}$.

מתקיים $v_2 S v_2 = -2 \neq 0$,

ולכן מספיק למצוא וקטור $v_3 \in W$ כך ש־ $v_3 S v_2 = 0$, כלומר $v_3 = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ צריך לקיים $\begin{cases} 2x + y + 3z = 0 \\ y + z = 0 \end{cases}$.

לדוגמא ניקח $v_3 = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$. אז $v_3^t S v_3 = 3$.

סה"כ מצאנו בסיס $B = \{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \}$ המלכסן את S.

כלומר $P^t S P = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ כאשר $P = \begin{pmatrix} 1 & 1 & -1 \\ 0 & -2 & -1 \\ 0 & 0 & 1 \end{pmatrix}$.

11.2.4 הוכחה נוספת למשפט הליכסון

(לא בדיוק הוכחה, יותר "רעיון").

ראינו ש־A שקולה ל־B אם ניתן להגיע מ־A ל־B ע"י סדרה של פעולות שורה ופעולות עמודה אלמנטריות.

באופן דומה, A חופפת ל־B אם אפשר להגיע מ־A ל־B ע"י סדרה של פעולות שורה\עמודה אלמנטריות.

(כלומר מבצעים את אותה פעולה על השורות ועל העמודות).

אם מתחילים ממטריצה סימטרית A, ננסה לעשות "דירוג בשורות ובעמודות בו־זמנית" -

נמצא איבר שונה מ־0 באלכסון.

אם לא קיים כזה, אז או שהתבנית זהותית 0, או שאפשר למצוא איבר $\neq 0$ במטריצה, נניח במקום (i,j), ואז נבצע את הפעולות: מוסיפים לשורה j את שורה i, ומוסיפים לעמודה j את שורה i. ואז מקבלים מטריצה שהאיבר ה־(j,j) יהיה $\neq 0$. (כי המציין $\neq 2$).

כעת כשמצאנו איבר שונה מ־0 באלכסון, נעביר אותו לפינה השמאלית־עליונה, ונשתמש בו כציר על מנת לאפס את האיברים מתחתיו ובו זמנית את אלה שמימינו.

זה עובד כי מבצעים את אותן פעולות על השורות ועל העמודות, ומצד שני המטריצה סימטרית.

ממשיכים באינדוקציה.

יש אלגוריתם מדוייק בסוף הפרק, מהתרגול.

11.2.5 מעל $\mathbb{C}$ - לכל ת. בילינארית סימטרית יש צורה אלכסונית מצומצמת יותר

כעת נניח ש- $\mathbb{F} = \mathbb{C}$. ולכן אם נסמן $rank(f) = rank([f]_B)$ (בבסיס כלשהו B, ולכן בכל בסיס - זהו אינווריאנט), אז מתקיים:

משפט 11.20 כל תבנית בילינארית **סימטרית** $f : V \times V \to \mathbb{C}$ כאשר $dim\, V = n$, יש לה צורה אלכסונית:

$$\left(\begin{array}{c|c} I_k & 0 \\ \hline 0 & 0 \end{array}\right)$$

(כלומר בלוק יחידה בגודל k משמאל למעלה ועוד שלושה בלוקים של אפסים).

k שווה לדרגה, ולכן הצורה נקבעת ביחידות. כלומר זיהינו את מחלקות השקילות תחת חפיפה עבור תבינות סימטריות.

הוכחה: היחידות נובעת מכך שהדרגה היא אינווריאנט.

קיום: נניח $B = \{v_1, ..., v_n\}$ הוא בסיס מלכסן ל-f. כלומר, $f(v_i, v_j) = 0$ אם $i \neq j$.

נסמן: $c_i = f(v_i, v_i)$. אנו יודעים ש-k מתוך ה-c_i הם שונים מ-0 (אם מסמנים $k = rank(f)$).

בה"כ, נניח $c_1, ..., c_k \neq 0$, והאחרים הם 0. נעשה שינוי בסיס:

נגדיר $B' = \{v'_1, ..., v'_n\}$ כאשר

$$v'_i = \begin{cases} \frac{v_i}{\sqrt{c_i}} & if\ 1 \leq i \leq k \\ v_i & if\ k < i \end{cases}$$

($\sqrt{c_i}$ הוא בחירה שרירותית של אחד משני השורשים).

אז מתקיים, עבור $1 \leq i \leq k$:

$$f(v'_i, v'_i) = f(\frac{v_i}{\sqrt{c_i}}, \frac{v_i}{\sqrt{c_i}}) = \frac{1}{c_i} f(v_i, v_i) = \frac{c_i}{c_i} = 1$$

ועבור $k < i$: $\qquad f(v'_i, v'_i) = f(v_i, v_i) = c_i = 0$

כמו כן, $f(v'_i, v'_j) = 0$ עבור $i \neq j$. ■

11.2.6 משפט ההתמדה של סילבסטר

משפט 11.21 <u>**א.**</u> לכל תבנית בילינארית סימטרית מעל $\mathbb{R}$ במ"ו ממימד n,

יש שני מספרים $p, m \geq 0$ כך ש- $p + m \leq n$, וכך שלתבנית יש מטריצת ייצוג:

$$\left(\begin{array}{ccc|ccc|ccc}
1 & & & & & 0 & & & 0\\
& \ddots & & & & & & & \\
& & 1 & 0 & & & 0 & & \\ \hline
& & 0 & -1 & & & & & 0\\
& & & & \ddots & & & & \\
0 & & & & & -1 & 0 & & \\ \hline
& & 0 & & & & 0 & & \\
& & & & & & & \ddots & \\
0 & & & & & & & & 0
\end{array}\right)$$

(כלומר בלוק בגודל p של אלכסון 1, אז בלוק בגודל m של אלכסון -1, ואז בלוק אפסים).

(וכל מה שמסביב זה אפסים כמובן).

ב. p,m נקבעים ביחידות ע"י f.

הוכחה: נוכיח את **א'**: אם $B'=\{v'_1,...,v'_n\}$ בסיס, כך ש־ $[f]_{B'}=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}$,

(קיים כזה, כי ראינו שכל תבנית בילינארית סימטרית ניתנת ללכסון מעל שדה כך ש־$char\mathbb{F}\neq 2$. וכמובן $char\mathbb{R}\neq 2$...)

ונניח ש־ $\lambda_1,...,\lambda_p>0\ ,\ \lambda_{p+1},...,\lambda_{p+m}<0\ ,\ \lambda_{p+m+1},...,\lambda_n=0$ (סידרנו אותם ככה וסימנו p,m).

אז כעת נגדיר $B=\{v_1,...,v_n\}$ ע"י:

$$v_i=\begin{cases}\frac{1}{\sqrt{\lambda_i}}v'_i & if\ \lambda_i>0\\ \frac{1}{\sqrt{-\lambda_i}}v'_i & if\ \lambda_i<0\\ v'_i & if\ \lambda_i=0\end{cases}$$

אז B הוא בסיס, ומתקיים:

$$f(v_i,v_i)=\begin{cases}f(\frac{1}{\sqrt{\lambda_i}}v'_i,\frac{1}{\sqrt{\lambda_i}}v'_i)=\frac{1}{\lambda_i}f(v'_i,v'_i)=1 & if\ \lambda_i>0\\ f(\frac{1}{\sqrt{-\lambda_i}}v'_i,\frac{1}{\sqrt{-\lambda_i}}v'_i)=\frac{1}{-\lambda_i}f(v'_i,v'_i)=-1 & if\ \lambda_i<0\\ f(v'_i,v'_i)=\lambda_i=0 & if\ \lambda_i=0\end{cases}$$

הערה שלי: נשים לב שזו הוכחה מאוד דומה למשפט הקודם! מכיוון ששם היינו מעל $\mathbb{C}$, השורש היה מוגדר תמיד. אבל כאן אנחנו מעל $\mathbb{R}$ אז צריך להתייחס לסימן. חוץ מהפרט הזה, עשינו כאן בדיוק אותו דבר.

נוכיח את **ב'**, יחידות:

נניח שבשני בסיסים שונים, יש ל־f ייצוג עם זוג (p,m) וזוג (p',m'). צריך להראות: $p=p',m=m'$.

אנו יודעים ש־ $p+m=p'+m'$ (כי זאת הדרגה), ולכן מספיק להראות ש־ $p=p'$.

נניח בשלילה ש־ $p\neq p'$. ובלי הגבלת הכלליות, $p>p'$.

ונניח ששני הבסיסים שנותנים צורות אלו הם $B'=\{v'_1,...,v'_n\}\ ,\ B=\{v_1,...,v_n\}$.

נתבונן בוקטורים $v_1,...,v_p,v'_{p'+1},v'_{p'+2},...,v'_n$. יש כאן $p+(n-p')>n$ וקטורים, ולכן יש תלות ליניארית. כלומר חייב להתקיים שויון מהצורה:

$$u=\sum_{i=1}^{p}a_iv_i=\sum_{j=p'+1}^{n}b_jv'_j$$

כאשר לא כל ה־a_i־ים הם 0. (כנראה שהכוונה היא שחייב להיות קיים וקטור u כזה) אבל מכאן מקבלים ש־

$$f(u,u)=f(\sum_{i=1}^{p}a_iv_i,\sum_{i=1}^{p}a_iv_i)=\sum_{i=1}^{p}a_i^2f(v_i,v_i)=\sum_{i=1}^{p}\lambda_ia_i^2>0$$

(ונשים לב שזה גדול מאפס כי $\lambda_i=1$ (כי לכל $1\le i\le p$, מתקיים $\lambda_i=1$)).

ומצד שני,

$$f(u,u)=f(\sum_{j=p'+1}^{n}b_jv'_j,\sum_{j=p'+1}^{n}b_jv'_j)=\sum_{j=p'+1}^{n}b_j^2f(v'_j,v'_j)=\sum_{j=p'+1}^{n}\lambda'_jb_j^2\leq 0$$

(כי $\lambda'_j\leq 0$).

וקיבלנו סתירה. ■

ניסוח אחר של מה שהוכחנו:

הזוג (p,m) הם אינווריאנט שלם עבור מטריצות סימטריות מעל $\mathbb{R}$ ביחס לחפיפה.

$p=$ מס' הע"ע החיוביים

$m=$ מס' הע"ע השליליים

$p+m=$ הדרגה

$p-m=$ "הסיגנטורה" של התבנית. (סימון היסטורי).

11.2.7 אלגוריתם ללכסון תבניות ריבועיות ע"י חפיפה (מהתרגול)

תהי $A=(a_{ij})$ מטריצה סימטרית מעל שדה $\mathbb{F}$, כך ש־ $char(\mathbb{F})\neq 2$.

<u>האלגוריתם:</u>

מקרה (i): $a_{11}\neq 0$.

לכל $2\leq i\leq n$ נבצע את פעולת השורה:

$$R_i:=R_i-\frac{a_{i1}}{a_{11}}R_1$$

ואז את אותה פעולת עמודה:

$$C_i:=C_i-\frac{a_{i1}}{a_{11}}C_1$$

או בקיצור, לאפס את העמודה, שורה הראשונים: $\left(\begin{array}{c|ccc} a_{11} & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & & \\ 0 & & & * \end{array}\right)$ ונקבל ש־A הפכה להיות: $\left(\begin{array}{c|ccc} a_{11} & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & & \\ 0 & & & B \end{array}\right)$.

(אם רוצים להמנע משברים, אפשר לבצע: $R_i:=a_{11}R_i-a_{i1}R_1$ ו־ $C_i:=a_{11}C_i-a_{i1}C_1$).

מקרה (ii): $a_{11}=0$, אבל קיים $2\leq i\leq n$ כך ש־ $a_{ii}\neq 0$.

נבצע את פעולת השורה $R_1\leftrightarrow R_i$, ואז את פעולת העמודה $C_1\leftrightarrow C_i$ כדי להביא את a_{ii} למקום הראשון באלכסון, וחוזרים למקרה (i).

מקרה (iii): כל איברי האלכסון שווים 0, וקיים $i\neq j$ כך ש־ $a_{ij}\neq 0$.

נבצע:

$$R_i:=R_i+R_j$$

$$C_i:=C_i+C_j$$

זה יביא את $2a_{ij} \neq 0$ למקום ה־i באלכסון. וחזרנו למקרה (ii).

מקרה (iv): כל המטריצה היא אפסים, ואז היא כבר אלכסונית.

ובסה"כ:

הבאנו את A לצורה $\left(\begin{array}{c|ccc} a_{11} & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & & \\ 0 & & & B \end{array}\right)$ כאשר B היא סימטרית, ונמשיך באינדוקציה.

בכדי למצוא את המלכסנת ע"י חפיפה,

נבצע את פעולות השורה (**בלבד!!!**) על I, ונקבל מטריצה P שמקיימת $PAP^t = D$ (D אלכסונית).

בפועל, הכי קל לרשום את I בצד של המטריצה, ולבצע עליה כל פעם יחד עם המטריצה שלנו את פעולות השורה. כמו שעושים באלגוריתם למציאת מטריצה הופכית שלמדנו בלינארית 1...

11.3 תבניות ריבועיות

11.3.1 הגדרה ודברים בסיסיים

הגדרה 11.22 אם $f : V \times V \to \mathbb{F}$ תבנית בילינארית,

אז הפונקציה $q : V \to \mathbb{F}$ המוגדרת ע"י $q(v) = f(v, v)$, נקראת **התבנית הריבועית** שמתאימה ל־f.

דוגמאות:

- הנורמה (בריבוע) במרחב מכפלה פנימית מעל $\mathbb{R}$.
- אם $A \in M_n(\mathbb{F})$ אז $q(v) = v^t A v$ היא $\sum_{i=1}^n \sum_{j=1}^n a_{ij} \cdot v_i \cdot v_j$. (כמו שלמדנו באינפי 2.... בערך)

מטריצת הייצוג של תבנית ריבועית בבסיס נתון B היא מטריצת הייצוג של התבנית הבילינארית המתאימה.

טענה 11.23 כל תבנית ריבועית נקבעת ע"י תבנית בילינארית **סימטרית, יחידה**.

הוכחה: נניח תבנית בילינארית f קובעת את התבנית הריבועית q.

כלומר $q(u) = f(u, u)$. אז,

$$q(u+v) = f(u+v, u+v) = f(u,u) + f(v,v) + f(u,v) + f(v,u) = q(u) + q(v) + f(u,v) + f(v,u)$$

או במילים אחרות,

$$\frac{1}{2}(f(u,v) + f(v,u)) = \frac{1}{2}(q(u+v) - q(u) - q(v))$$

מכאן נובעים שני הכיוונים:

אם f סימטרית אז היא מקיימת: $\frac{1}{2}(2f(u,v)) = f(u,v) = \frac{1}{2}(q(u+v) - q(u) - q(v))$,

ולכן נקבעת ביחידות מ־q.

כמו כן, לגבי הקיום, נגדיר פונקציה $h(u,v) = \frac{1}{2}(q(u+v) - q(u) - q(v))$. היא סימטרית ב־$u, v$ (כי הביטוי סימטרי), וקל לבדוק שהיא תבנית בילינארית (הושאר כתרגיל).

■

הערה 11.24 בשפה של מטריצות, החלפנו את המטריצה A בתבנית $v \longmapsto v^t A v$.

במקרה ש־A לא סימטרית, בתבנית $v \longmapsto v^t \left(\frac{A+A^t}{2}\right) v$. זה נותן את אותה פונקציה של v.

לדוגמא:

$$A = \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}$$

$$v = \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} x & y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix} = x^2 + 2xy + 4yx + 5y^2 =$$

$$= x^2 + 2 \cdot 3xy + 5y^2 = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

11.3.2 תבניות ריבועיות מעל $\mathbb{R}$ - חיוביות

מעכשיו, $\mathbb{F} = \mathbb{R}$.

הגדרה 11.25 תבנית ריבועית q נקראת **חיובית** אם $q(v) > 0$ לכל $v \neq 0$.

כלומר אם מסתכלים על מטריצת הייצוג $A = [q]_B$, צריך לבדוק שמתקיים $v^t A v > 0$ לכל $v \neq 0 \in \mathbb{R}^n$.

התבנית q נקראת **אי־שלילית** אם $q(v) \geq 0$ לכל v.

(על מטריצה אומרים שהיא חיובית\אי־שלילית אם היא מט' ייצוג של תבנית ריבועית מסוג זה).

משפט 11.26 מטריצה סימטרית A היא חיובית אם"ם כל הע"ע שלה חיוביים.

A היא אי־שלילית אם"ם כל הע"ע שלה אי־שליליים.

הוכחה: אם $\lambda_1, ..., \lambda_n$ הע"ע, ו־$\{v_1, ..., v_n\}$ בסיס אורתונורמלי של ו"ע,

אז לכל $v \in V$, נכתוב $v = \sum_{i=1}^n a_i v_i$. ואז,

$$v^t A v = \left(\sum_{i=1}^n a_i v_i^t\right) \cdot A \cdot \left(\sum_{j=1}^n a_j v_j\right) = \sum_{i=1}^n \sum_{j=1}^n a_i a_j \left(v^t A v\right) = \sum_{i=1}^n a_i^2 \lambda_i$$

המעבר האחרון הוא כי $\sum_{i=1}^n \sum_{j=1}^n a_i a_j \left(v^t A v\right) = \sum_{i=1}^n \sum_{j=1}^n a_i a_j \cdot \left(\lambda_j v_i^t v_j\right)$.

ומתקיים $\left(\lambda_j v_i^t v_j\right) = \begin{cases} \lambda_i & i = j \\ 0 & i \neq j \end{cases}$. ככה לפחות כתוב לי במחברת...

זה חיובי לכל $(a_1, ..., a_n) \neq (0, ..., 0)$ אם"ם כל ה־λ_i חיוביים, ואי־שלילי לכל $(a_1, ..., a_n)$ אם"ם כל ה־λ_i אי־שליליים. ∎

11.3.3 כיצד לבדוק אם A חיובית בלי לחשב את הע"ע?

תנאי הכרחי: $det(A) > 0$. (מכפלת הע"ע).

עוד תנאי הכרחי: אם $A = (a_{ij})$ אז $v^tAv{=}a_{11} > 0$ כאשר $= e_1$ $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ $v =$.

באופן יותר כללי, לכל $1 \leq k \leq n$,

אם נגדיר

$$A_k = \left(\begin{array}{c|c} (a_{ij}) & \begin{matrix} & 0 \\ 0 & \end{matrix} \\ \hline \begin{matrix} & 0 \\ 0 & \end{matrix} & \begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix} \end{array}\right)$$

(כאשר הבלוק (a_{ij}) הוא בגודל $k \times k$)

אם נבדוק את התנאי $v^tAv > 0$ על כל וקטור $v \in span\{e_1, ..., e_k\}$ אז עבור וקטורים כאלה $v^tAv = v^tA_kv$,

ולכן צריך שיתקיים $det(A_k) > 0$.

כלומר הדטרמיננטה של "המינור העיקרי מסדר k" של A היא $0 <$, לכל $1 \leq k \leq n$.

משפט 11.27 קריטריון החיוביות של סילבסטר

אם $det(A_k) > 0$ לכל k, אז A (סימטרית) חיובית.

הוכחה: למזלנו נגמר הסמסטר לפני שהוכחנו את הנ"ל.

■