רשימות אלגברה לינארית 2א

שחר פרץ $\sim$ 2025B

גרסה 1.1.0

0.1 Introduction

0.1.1 ~ License ~ רישיון

0.1.2 ~ הרבה מילים שאפשר לדלג עליהן

הסיכום להלן נוצר משילוב של ארבעת המקורות הבאים:

- הרצאות בפקולטה
- הרצאות באודיסאה של בן בסקין
- מסקנות מהספר "Linear Algebra Done Right" (עם הוכחות שאני כתבתי)
- תרגולים של עומרי שדה־אור

כנגד שלושה נושאים דיברה אלגברה לינארית 2א –

1. **אופרטורים ליניארים**, הן העתקות ממרחב לעצמו.
2. **תבניות בי־ליניאריות**, אוביקט מתמטי נוסף שניתן לייצג ע"י מטריצה.
3. **מרחבי מכפלה פנימית**, מרחבים בהם מוגדרת מעין תבנית ססקווי בי־לינארית שמאפשרת תיאור "גודל", ובהם יש ערך לפירוק מטריצות לכפל של מספר מטריצות שונות.

הגרסה האחרונה של הסיכום תהיה זמינה בקישור הבא כל עוד מיקרוסופט לא פשטו את הרגל. אם מצאתם בסיכום טעויות (החל בתקלדות, כלה בשגיהוט חטיב, וכמובן טעויות מתמטיות) אשמח אם תפנו אלי במייל (perets.shahar@gmail.com), בטלפון (אם אתם מכירים אותי ויש לכם אותו), או באמצעות GitHub Issues (קישור בתחילת המשפט).

מקווה שתהנו מהסיכום ותמצאו אותו מועיל;

שחר פרץ, 19.7.2025

אזהרה. הסיכום הזה מכיל בחלקו הוכחות שאני כתבתי ולא הופיעו בהרצאה. השימוש בסיכום על אחריות המשתמש ואני לא ערב לנכונות המידע.

0.1.3 ~ סימונים

בסיכום הבא נניח את הסימונים הבאים:

- $[n] := \mathbb{N} \cap [0, n]$
- בהינתן $T : V \to W$ העתקה ו־$U \subseteq V$ תמ"ו, נסמן $T(U) := \{Tu \mid u \in U\}$
- בהינתן $T : V \to W$ העתקה ו־$v \in V$, נסמן $Tv := T(v)$
- בהינתן A קבוצה עם יחס שקילות $\sim$, נסמן את קבוצת המנה ב־$A/_\sim$
- בפקולטה למתמטיקה בת"א מקובל להשתמש ב־(v, w) בשביל מכפלה פנימית. בסיכום הזה אשתמש ב־$\langle v \mid w \rangle$, גם כן סימון מקובל (בעיקר בפיזיקה), שאני חושב שנראה מגניב הרבה יותר.
- נסמן שחלוף (transpose) ב־A^T ולא A^t.
- הטבעיים כוללים את 0, ו־$\mathbb{N}_+$ ("הטבעיים החיוביים") אינם.
- ט"ל הוא קיצור לטרנספורמציה לינארית.

0.1.4 ~ רשימת נושאים שהוספתי לסיכום נוסף על החומר של הקורס

- מציאת צורת ג'ורדן באמצעות מרחבים עצמיים מורחבים (עוזר מאוד להבין מה צורת ג'ורדן עושה).
- תוצאות מצורת ג'ורדן (מופיע בסמסטרים קודמים וברמה הפרקטית חומר למבחן).

- הרחבה על הראדיקלים של תבניות בי־לינאריות (סתם כי זה מגניב).
- הרחבת פירוק SVD להעתקות שאינן אופרטורים (מועיל למדמ"חיסטים).
- מסקנות מ־SVD ושימוש הקונספט של ערכים סינגולריים.

תוכן העניינים

0.1 מבוא 2
0.1.1 רישיון 2
0.1.2 הרבה מילים שאפשר לדלג עליהן 2
0.1.3 סימונים 2
0.1.4 רשימת נושאים שהוספתי לסיכום נוסף על החומר של הקורס 2

1 חקר אופרטורים לינאריים וצורת ג'ורדן 5
1.1 לכסון 6
1.1.1 מבוא לפרק 6
1.1.2 ערכים עצמיים ווקטורים עצמיים לאופרטורים לינארים 6
1.1.3 ערכים עצמיים ווקטורים עצמיים למטריצות 7
1.1.4 פולינום אופייני 8
1.1.5 על הקשר בין ריבוי גיאומטרי ואלגברי 10
1.1.5.1 פיבונאצ'י בשדה סופי 11
1.1.6 שילוש 11
1.1.7 על ההבדל בין פולינום לפולינום 13
1.1.8 משפט קיילי־המילטון 13
1.2 תורת החוגים 15
1.2.1 מבוא והגדרות בסיסיות 15
1.2.2 ראשוניות ואי־פריקות 15
1.2.3 הרחבת שדות 18
1.2.4 חוג הפולינומים 19
1.2.4.1 פונקציות רציונליות ומספרים אלגבריים 20
1.3 פירוק פרימרי 22
1.3.1 מרחבים T־שמורים וציקליים 22
1.3.2 הפולינום המינימלי 23
1.3.3 ניסוח והוכחת משפט הפירוק הפרימרי 25
1.4 צורת ג'ורדן 28
1.4.1 מציאת שורשי פולינום אופייני ממעלה חמישית ואילך 28
1.4.2 צורת ג'ורדן לאופרטור לינארי נילפוטנטי 29
1.4.2.1 נילפוטנטיות 29
1.4.2.2 שרשאות וציקליות 29
1.4.2.3 ניסוח צורת מייקל ג'ורדן לאופרטור נילפוטנטי 31
1.4.3 צורת ג'ורדן לאופרטור לינארי כללי 33
1.4.3.1 בעזרת פירוק פרימרי 33
1.4.3.2 בעזרת מרחביים עצמיים מוכללים 34
1.4.4 תוצאות מצורת ג'ורדן 36

2 הגדרה וחקר מרחבי מכפלה פנימית 38
2.1 תבניות בי־לינאריות 39
2.1.1 הגדרות בסיסיות בעבור תבניות בי־לינאריות כלליות 39
2.1.2 חפיפה וסימטריות 41
2.1.3 תבניות ריבועיות 43
2.1.4 משפט ההתאמה של סילבסטר 43
2.2 מרחבי מכפלה פנימית 46
2.2.1 הגדרה כללית 46

2.2.1.1 מעל $\mathbb{R}$ 46
2.2.1.2 מעל $\mathbb{C}$ 46
2.2.2 אורתוגונליות, זהויות ואי־שוויונות של המכפלה הפנימית 47
2.2.2.1 משפט פיתגורס ותוצאותיו 47
2.2.3 מרחבים ניצבים והיטלים 49
2.2.3.1 אלגוריתם גרהם־שמידט 50
2.2.4 צמידות ודואליות 51
2.2.4.1 העתקות צמודות לעצמן, והעתקות חיוביות 51
2.2.4.2 ההעתקה הצמודה 54
2.3 פירוקים 56
2.3.1 המשפט הספקטרלי להעתקות 56
2.3.1.1 ניסוח המשפט הספקטרלי להעתקות צמודות לעצמן 56
2.3.1.2 ניסוח המשפט הספקטרלי בעבור העתקה כללית 57
2.3.1.3 תוצאות ממשפט הפירוק הספקטרלי 58
2.3.2 מטריצות אוניטריות 60
2.3.2.1 צורה נורמלית למטריצה אורתוגונלית 62
2.3.2.2 המשפט הספקטלי בניסוח מטריציוני 64
2.3.3 פירוק פולארי 64
2.3.3.1 מבוא, וקישור לתבניות בי־לינאריות 64
2.3.3.2 ניסוח הפירוק הפולארי 66
2.3.4 פירוק SVD 67
2.3.4.1 ניסוח והוכחת SVD 67
2.3.4.2 הרחבת SVD להעתקות שאינן אופרטורים 68
2.3.4.3 נורמה של העתקה 70

3 נספחים 71
3.1 מרחבים דואלים 72
3.1.1 הגדרות בסיסיות 72
3.1.2 איזומורפיות למרחבי מכפלה פנימית 73
3.1.2.1 העתקה צמודה (דואלית) 73
3.1.2.2 המאפס הדואלי ומרחב אורתוגונלי 74
3.2 סיכום תוצאות מרכזיות 76
3.2.1 סיכום פירוק פרימרי וצורת ג'ורדן 76
3.2.2 סיכום תבניות בי־לינאריות 76
3.2.3 סיכום נושא הפירוקים 77
3.3 אלגוריתמים 78
3.4 תרגילים מומלצים 80

בתיאבון

פרק 1

חקר אופרטורים לינאריים וצורת ג'ורדן

Diagonalization 1.1

1.1.1 ~ מבוא לפרק

הגדרה 1. נאמר ש־A מטריצה אלכסונית, אם:

$$A = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} =: \operatorname{diag}(\lambda_1, \lambda_2 \ldots \lambda_n)$$

נאמר שישנה פעולה כשהי שנרצה להפעיל. נרצה לקרות מה יקרה לפעולה לאחר שנפעיל אותה מספר פעמים. כפל מטריצות היא פעולה מסדר גודל של $\mathcal{O}(n^3)$. אך, ישנן מטריצות שקל מאוד להעלות בריבוע, ובכך נוכל להפוך את ההליך לפשוט בהרבה, ואף לנסח אותו בצורה של נוסחה סגורה פשוטה. דוגמה מטריצה כזו היא מטריצה אלכסונית. ננסה למצוא דרך ,להמיר" בין מטריצה "רגילה" למטריצה אלכסונית.

הגדרה 2. ו־T ההעתקה אלכסונית:

$$T_A^m(A) = A^m(v) = \begin{pmatrix} \lambda_1^m & & \\ & \ddots & \\ & & \lambda_n^m \end{pmatrix}$$

למה זה מועיל? נזכר בסדרת פיבונצ'י. נתבונן בהעתקה הבאה:

$$\begin{pmatrix} a_{n+2} \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} \implies \begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{n-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

(בהנחת איברי בסיס $a_0 = 0, a_1 = 1$).

ואכן, מה חבל שלא כיף להכפיל את המטריצה $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ בעצמה המון פעמים. מה נוכל לעשות? ננסה למצוא בסיס שבו $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)_B =$ (v_1, v_2) ו־$P^{-1}\Lambda P = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. [המשמעות של Λ היא מטריצה אלכסונית כלשהית] אז נקבל:

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n = \left(P^{-1}\Lambda P\right)^n = P^{-1}\Lambda^n P$$

(די קל להראות את השוויון האחרון באינדוקציה). במקרה כזה יהיה נורא נחמד כי אין בעיה להעלות לכסינה בחזקה.

הדבר הנחמד הבא שנוכל ליצור הוא צורת ג'ורדן – מטריצת בלוקים אלכסונית, ככה שנעלה בחזקה את הבלוקים במקום את כל המטריצה. נעשה זאת בהמשך הקורס.

הגדרה 3. אופרטור ליניארי (א"ל) הוא ה"ל/ט"ל ממרחב וקטורי V לעצמו.

1.1.2 ~ ערכים עצמיים ווקטורים עצמיים לאופרטורים לינארים

הגדרה 4. יהי $T\colon V \to V$ א"ל. אז $0 \neq v \in V$ נקרא וקטור עצמי של T (ו"ע) אם קיים $\lambda \in \mathbb{F}$ כך ש־$Tv = \lambda v$.

הגדרה 5. λ מההגדרה הקודמת נקרא ערך עצמי (ע"ע) של T, המתאים לו"ע v.

שאלה. יהי $T\colon \mathbb{F}^n \to \mathbb{F}^n$. נניח ש־$V = (v_1 \ldots v_n)$ בסיס של ו"ע של $\mathbb{F}^n$ [תיאורטית יכול להתקיים באופן ריק כי עדיין לא הראינו שקיים בסיס כזה] אז קיימת $P \in M_n(\mathbb{F})$ הפיכה כך ש־A המקיימת $Tv = Av$ לפי הבסיס הסטנדרטי, אז $[T]_B^B = \operatorname{diag}(\lambda_1 \ldots \lambda_n)$, כאשר $\lambda_1 \ldots \lambda_n$ ע"ע המתאימים לו"ע $v_1 \ldots v_n$.

כדאי לדעת כי $\operatorname{Hom}(\mathbb{F}^n, \mathbb{F}^n) \cong M_{m\times n}(\mathbb{F})$. מה המשמעות של איזומורפי ($\cong$)? בהינתן A, B מבנים אלגבריים כלשהם, נסמן $A \cong B$ אם קיימת $\varphi\colon A \to B$ העתקה חח"ע ועל שמשמרת את המבנה (כאשר המבנה שלנו מורכב מפעולות חיבור וכפל, העתקה כזו תהיה ליניארית).

דוגמה. אם V, U מ"ו מעל $\mathbb{F}$, הם נקראים איזומורפים אם קיימת $\varphi\colon V \to U$ חח"ע ועל המקיימת

$$\forall \lambda_1, \lambda_2 \in \mathbb{F}\, \forall v_1, v_2 \in V\colon \varphi(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 \varphi(v_1) + \lambda_2 \varphi(v_2)$$

כלומר "המרנו" שני מבנים אלגבריים, אבל לא באמת עשינו שום דבר – כל מבנה עדיין שומר על התכונות שלו.

סימון 1. בסוף הסיכום מופיעה הרחבה על תופעות מעין אלו.

הגדרה 6. יהי $T\colon V\to V$ א"ל, נניח $\lambda\in\mathbb{F}$ ע"ע, אז המרחב העצמי (מ"ע) של λ הוא:

$$V_\lambda := \{v\in V \mid Tv=\lambda v\}$$

משפט 1. V_λ תמ"ו של V.

הגדרה 7. יהי $T\colon V\to V$ א"ל, ויהי $\lambda\in\mathbb{F}$ ע"ע של T. נגדיר את הריבוי הגיאומטרי של λ (ביחס ל־T) הוא $\dim V_\lambda$.

דוגמה. יהי V מ"ו ממימד n, $T\colon V\to V$ א"ל. נניח קיום $v\in V$ המקיים $T^n v=v$, ונניח $\{v,Tv,T^2v,\dots,T^{n-1}v\}$ בסיס של V. ננסה להבין מהם הע"ע.

יהי $0\neq u\in V$ ו"ע כך ש־$Tu=\lambda u$. נראה כי $T^n u=u$. ידוע קיום $\alpha_0,\dots,\alpha_{n-1}\in\mathbb{F}$ כך ש־$u=\sum\alpha_i T^i(v)$. אז:

$$\lambda^n u = T^n(u) = \sum_{i=0}^{n-1}\alpha_i \underbrace{T^{n+i}(v)}_{=T^i(T^n v)=T^i v} = u$$

נבחין שהוקטורים העצמיים הם שורשי היחידה. מי הם שורשי היחידה – זה תלוי שדה.

מסקנה 1. ערכים עצמיים תלויים בשדה. ערכים עצמיים של מטריצה מעל $\mathbb{R}$ יכולים להיות שונים בעבור אותה המטריצה מעל $\mathbb{C}$.
דוגמה יותר פשוטה לכך היא העתקת הסיבוב ב־$\mathbb{R}^2$, שאין לה ו"עים מעל $\mathbb{R}$ אך יש כאלו מעל $\mathbb{C}$.

משפט 2. תהי $T\colon V\to V$ א"ל, ונניח $A\subseteq V$ קבוצה של ו"ע של T עם ע"ע שונים, אז A בת"ל. הוכחה בתרגול.

הגדרה 8. יהי $T\colon V\to V$ א"ל. נאמר ש־T ניתן לכסון/לכסין אם קיים ל־V בסיס של ו"ע של T.

מסקנה 2. אם $\dim V=n$ ול־T יש n ע"ע שונים אז T לכסין.

הערה 1. שימו לב – ייתכן מצב בו קיימים פחות מ־n ע"ע שונים אך T עדיין לכסין. דוגמה: $0, id$.

מסקנה 3. תהי $T\colon V\to V$ א"ל. נניח שלכל λ ע"א, ישנה $B_\lambda\subseteq V_\lambda$ בת"ל. אז $B=\bigcup_\lambda B_\lambda$ בת"ל.

הוכחה. ניקח צירוף לינארי כלשהו שווה ל־0:

$$\begin{aligned}
&\sum_{v_i\in B}\alpha_i v_i = 0\\
&\quad = \sum_\lambda\sum_{\lambda_i}\alpha_i v_{\lambda,i}\\
&\quad \implies \sum_{\lambda_j}\alpha_i v_{\lambda_{ji}} =: u_j\in V_{\lambda_j}\\
&\quad \implies \sum_j u_j = 0
\end{aligned}$$

קיבלנו צירוף ליניארי לא טרוויאלי של איברים במ"ע שונים (=עם ע"א שונים). אם אחד מהם אינו 0, קיבלנו סתירה למשפט. סה"כ קיבלנו שלכל j מתקיים $\sum\alpha_{ji}v_{ji}=0$. בגלל ש־$v_{ji}\in B_j$ אז בת"ל ולכן כל הסקלרים 0. ■

הערה 2. ההוכחה הזו עובדת בעבור ההכללה לממדים שאינם נוצרים סופית

מסקנה 4. יהי $T\colon V\to V$ א"ל כך ש־$\dim V=n$. אז:

$$\sum_\lambda \dim V_\lambda \leqslant n$$

שוויון אמ"מ T לכסין.

הוכחה. לכל λ יהא B_λ בסיס. אז $B=\sum_\lambda B_\lambda$ בת"ל. אז $n\geqslant |B|=\sum_\lambda \dim V_\lambda$.

אם T לכסין אז קיים בסיס של ו"ע כך שאכל אחד מהם מבין V_λ ושוויון.

מצד שני, אם יש שוויון אז B קבוצה בת"ל של n ו"ע ולכן בסיס ולכן T לכסין. ■

1.1.3 $\sim$ ערכים עצמיים ווקטורים עצמיים למטריצות

הגדרה 9. תהי $A\in M_n(\mathbb{F})$. נאמר ש־$0\neq v\in\mathbb{F}^n$ הוא ו"ע של A עם ע"ע λ אם $Av=\lambda v$.

משפט 3. תהי $T\colon V\to V$ א"ל ויהי B בסיס סדור, ו־V נוצר סופי (לעיתים יקרא: סוף־ממדי). נניח $A=[T]_B$. אז $v\neq 0$ וקטור עצמי של T עם ערך עצמי λ אמ"מ $[v]_B$ וקטור עצמי של A עם ע"ע λ.

הוכחה. גרירה דו־כיוונית. נניח V ו"ע של T. אז $A[v]_B = [Tv]_B = [\lambda_v]_B \lambda[v]_B$. מהכיוון השני "לכו הפוך". ■

הגדרה 10. מטריצה $A \in M_n(\mathbb{F})$ תקרא לכסינה/נתנת ללכסון אם היא דומה למטריצה $\Lambda \in M_n(\mathbb{F})$ אלכסונית, כלומר קיימת $P \in M_n(\mathbb{F})$ הפיכה שעבורה $\Lambda = P^{-1}AP$.

משפט 4. יהיו $A, P \in M_n(\mathbb{F})$. נניח P הפיכה. אז אם $P^{-1}AP = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ אמ"מ עמודות P הן ו"ע של A עם ע"ע $\lambda_1 \dots \lambda_n$ בהתאמה.

הוכחה. נסמן $P = (P_1 \dots P_n)$ עמודותיה. אז:

$$AP = (AP_1 \dots AP_n) = (\lambda_1 P_1 \dots \lambda_n P_n) = P \underbrace{\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}}_{\Lambda}$$

ולכן:

$$P^{-1}AP = P^{-1}P\Lambda$$

ההוכחה מהכיוון השני היא לקרוא את זה מהצד השני. ■

"אני מקווה שראיתם שכפל מטריצה באלכסונית מתחלף" (הוא לא בהכרח). "אני אמרתי שטות." ~ בן

משפט 5. בהינתן העתקות T, S שתיהן לכסינות לפי אותו הבסיס B (לא בהכרח אותם הע"עים), אז $TS = ST$ מתחלפות.

משפט 6. המטריצה λI עבור $\lambda \in \mathbb{F}$ דומה רק לעצמה.

הוכחה. בהינתן P הפיכה, הכפל של P עם λI מתחלף בהכרח, ולכן $P^{-1}\lambda IP = PP^{-1}\lambda I = \lambda I$ לכל מטריצה $P^{-1}\lambda IP$ דומה. ■

1.1.4 ~ פולינום אופייני

תרגיל. תהי $A = \left(\begin{smallmatrix} -7 & 8 \\ 6 & 7 \end{smallmatrix}\right)$. מצאו ו"ע וע"ע של A ולכסנו אם אפשר.

פתרון. מחפשים $\mathbb{R}^2 \ni \binom{x}{y} \neq \binom{0}{0}$ ו־$\lambda \in \mathbb{R}$ כך ש־:

$$\begin{aligned} A\begin{pmatrix} x \\ y \end{pmatrix} &= \lambda \begin{pmatrix} x \\ y \end{pmatrix} \\ A\begin{pmatrix} x \\ y \end{pmatrix} &= \lambda I \begin{pmatrix} x \\ y \end{pmatrix} \\ \lambda I \begin{pmatrix} x \\ y \end{pmatrix} - A \begin{pmatrix} x \\ y \end{pmatrix} &= 0 \\ (\lambda I - A)\begin{pmatrix} x \\ y \end{pmatrix} &= 0 \end{aligned}$$

סה"כ $\binom{x}{y}$ ו"ע עם ו"ע λ אמ"מ $\binom{x}{y} \in \ker(\lambda I - A)$, אמ"מ $\lambda I - A$ לא הפיכה, אמ"מ $\det(\lambda I - A) = 0$ (AKA "הפולינום האופייני"). במקרה הזה:

$$\lambda I - A = \begin{pmatrix} \lambda + 7 & -8 \\ -6 & \lambda - 7 \end{pmatrix} \implies \det(\lambda I - A) = (\lambda + 7)(\lambda - 7) = \lambda^2 - 1$$

לכן הע"ע הם ± 1. נמצא את הו"ע. עבור $\lambda = 1$, מתקיים:

$$\begin{pmatrix} 8 & 8 \\ -6 & -6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

יש לנו חופש בחירה (ופתרון יחיד עד לכדי כפל בסקלר). במקרה הזה, נבחר $\binom{x}{y} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

עבור $\lambda = -1$, יתקיים:

$$\begin{pmatrix} 6 & 8 \\ -6 & 8 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 0 \impliedby b\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 \\ -3 \end{pmatrix}$$

ראינו שהמלכסנת היא העמודות של הו"ע. אז:

$$P = \begin{pmatrix} 1 & 4 \\ -1 & -3 \end{pmatrix}$$

וסה"כ $P^{-1}AP = I$. מכאן צריך למצוא את P^{-1}.

משפט 7. תהי $A \in M_n(\mathbb{F})$ אז $\lambda \in \mathbb{F}$ ע"ע של A אמ"מ $|\lambda I - A| = 0$.

הגדרה 11. תהי $A \in M_n(\mathbb{F})$. הפולינום האופייני של A מוגדר להיות:

$$f_A(x) = |xI - A|$$

משפט 8. תהי $A \in M_n(\mathbb{F})$. אז $f_A(x)$ הוא פולינום מתוקן [=מקדם מוביל הוא 1] ממעלה n, המקדם של x^{n-1} הוא $-\operatorname{tr} A$ והמקדם החופשי הוא $(-1)^n|A|$.

הגדרה 12. בעבור $A \in M_n(\mathbb{F})$ הפולינום האופייני של A הוא $f_A(x) = \det(Ix - A)$.

ראינו ש־v ו"ע של A עם ערך עצמי λ אמ"מ $v \in \ker(\lambda I - A)$, וכן λ ע"ע אמ"מ $\dim\ker \lambda - A > 0$.

משפט 9. $f_A(x)$ פולינום מתוקן (מקדם מוביל 1) מדרגה n, המקדם של x^{n-1} הוא $-\operatorname{tr} A$, המקדם החופשי הוא $(-1)^n \det A$.

הוכחה.

- **תקינות הפולינום.** מבין $n!$ המחוברים, ישנו אחד יחיד שדרגתו היא n. הסיבה היא שמדטרמיננטה לפי תמורות, התמורה היחידה שתיצור איבר מסדר x^n היא תמורת הזהות שתעבור על האלכסון. באינדוקציה על n, ונקבל:

$$f_A(x) = (x - a_{11})\cdots(x - a_{nn}) = x^n + \cdots$$

באופן דומה אפשר להוכיח באינדוקציה באמצעות פיתוח לפי שורות:

$$f_A(x) = (x - a_{11})|A_11| + \underbrace{a_{21}|A_21| - a_{31}|A_31| + \cdots + (-1)^{n+1}a_{n1}|A_{n1}|}_{\text{מה.א. דרגה קטנה מ־}n}$$

סה"כ גם כאן הראינו שהדרגה מתקבלת מהפולינום $\prod_{i=1}^n (x - a_{ii})$, כלומר הפולינום האופייני מתוקן.

- **המקדם של x^{n-1} הוא $-\operatorname{tr} A$.** מקדמי x^{n-1} מגיעים גם הם רק מ־$\prod_{i=1}^n (x - a_{ii})$ (הפולינום למעלה) שהם $-\operatorname{tr} A =$ המקדם החופשי $\sum_{i=1}^n -a_{ii}$.

- **המקדם החופשי.** מתקבל מהצבת 0. $f_A(0) = \det(I \cdot 0 - A) = \det(-A) = (-1)^n \det A$.

■

דוגמאות.

א) אם $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ אז $f_A(x) = x^2 - (a+d)x + ad - bc$ (נטו מהמשפט הקודם).

ב) אם $A = \operatorname{diag}(\lambda_1 \dots \lambda_n)$ אז $f_A(x) = \prod_{i=1}^n (x - \lambda_i)$.

ג) אם $A = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ אז גם כאן $f_A(x) = \prod_{i=1}^n (x - \lambda_i)$ אך כדאי לשים לב שמשולשית עליונה לא בהכרח דומה לאלכסונית עם אותם הקבועים.

ד) אם $A = \begin{pmatrix} B & * \\ 0 & C \end{pmatrix}$ כאשר B, C בלוקים ריבועיים אז $f_A(x) = f_B(x) \cdot f_C(x)$.

הגדרה 13. בהינתן $T\colon V \to V$ ט"ל נגדיר את הפולינום האופייני שלה (פ"א) באופן הבא: נבחר בסיס B למ"ו V, ונתבונן ב־$A = [T]_B$ ונגדיר את $f_T(x) := f_A(x)$.

"אתה פותר עכשיו שאלה משיעורי הבית" "אל תדאג הבודק כבר שלח פתרון" "מה?😨"

משפט 10. הפ"א של ט"ל מוגדר היטב. למטריצות דומות אותו פ"א. הוכחה. בשיעורי הבית. ויש סיכוי שגם בדף הסיכום.

דוגמה. נתבונן בהעתקה $T(f)=f',\ \mathbb{R}_n[x]\to\mathbb{R}_n[x]$. נבחר בסיס $B=(1,x,\dots,x^n)$. אז:

$$[T]_B=\begin{pmatrix}0&1&0&0&\dots&0\\0&0&2&0&\dots&0\\0&0&0&3&\ddots&\vdots\\\vdots&\vdots&\vdots&\ddots&\ddots&\\0&0&0&0&&\end{pmatrix}$$

אז:

$$f_T(x)=\det\begin{pmatrix}x&-1&0&\dots&&\\&x&-2&0&\dots&\\&&\ddots&\ddots&&\end{pmatrix}=x^{n+1}$$

משפט 11. $T\colon V\to V$ ט"ל, אז λ ע"ע של T אמ"מ $f_T(\lambda)=0$.

הוכחה. יהא $B\subseteq V$ בסיס של V. אז $A=[T]_B$ ואז $f_T(\lambda)=0$ אמ"מ $f_A(\lambda)=0$ אמ"מ λ ע"ע של A. ■

הגדרה 14. יהי $\lambda\in\mathbb{F}$ ע"ע של T (או A). הריבוי האלגברי של λ הוא החזקה המקסימלית d כך ש־$(x-\lambda)^d\mid f_T(x)$ (חלוקת פולינומים).

דוגמה. בעבור T היא העתקת גזירת פולינום, הפ"ע $f_T(x)=x^{n+1}$ ולכן ע"ע יחיד הוא 0. הריבוי האלגברי של 0 הוא $n+1$. הריבוי הגיאומטרי של 0 הוא 1.

סימון 2. נניח ש־λ ע"ע של T (או A) אז d_λ הריבוי האלגברי של λ ו־r_λ הריבוי הגיאומטרי של λ.

1.1.5 ∼ על הקשר בין ריבוי גיאומטרי ואלגברי

הערה 3. במקרים רבים $\sum d_i=n$ כאשר n דרגת הפולינום. זה לא תמיד המצב.

דוגמה למצב בו זה לא קורה: $x^2(x^2+1)\in\mathbb{R}[x]$. סכום הריבויים האלגבריים הוא 2, אבל דרגת הפולינום היא 4. זה נכון מעל שדות סגורים אלגברית.

משפט 12. תהי $T\colon V\to V$ ט"ל. אזי לכל ע"ע λ מתקיים $r_\lambda\leqslant d_\lambda$.

הוכחה. יהי λ ע"ע. אז $V_\lambda=\{v\in V\mid Tv=\lambda v\}$. יהי $B_\lambda\subseteq V_\lambda$ בסיס עבור V_λ. נשלים אותו לבסיס B של V.

$$[T]_B=\begin{pmatrix}\lambda&&0&*\\&\lambda&&\\0&&\ddots&\\ *&&&C\end{pmatrix}$$

ואז:

$$f_T(x)=(x-\lambda)^{r_\lambda}C(x)\implies r_\lambda\leqslant d_\lambda$$

■

משפט 13. תהי $T\colon V\to V$ ט"ל עם פ"א $f_T(x)$. אז T לכסינה אמ"מ שתי הטענות הבאות מתקיימות:

1. בעבור k הע"ע שונים, $f_T(x)=\prod_{i=1}^k(x-\lambda_i)^{d_\lambda}$
2. לכל λ ע"ע של T מתקיים $r_\lambda=d_\lambda$

(הבהרה: 1 לא גורר את 2. צריך את שניהם).

הוכחה.

$\Leftarrow$ T לכסינה ראינו ש־1 מתקיים. במקרה שלכסינה ראינו ש־$n=\sum r_{\lambda_i}\leqslant\sum d_{\lambda_i}=n$ ולכן אם לאחד מבין הערכים העצמיים מתקיים $r_\lambda\neq d_k$ אז מתקיים $r_k<d_k$ ונקבל סתירה לשוויונות לעיל.

$\Longrightarrow$

$$
\begin{aligned}
1 &\implies \sum d_{\lambda_i} = n \\
2 &\implies \sum r_{\lambda_i} = \sum d_{\lambda_i} = n
\end{aligned}
$$

וסה"כ $\sum r_{\lambda_i} = n$ אמ"מ T לכסינה. ■

(1.1.5.1) פיבונאצ'י בשדה סופי

סדרת פיבונאצ'י:

$$\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}}_{=0}^{n} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

נניח שאנו מסתכלים מעל $\mathbb{F}_p$ כלשהו. אז הסדרה חייבת להיות מחזורית. **שאלה:** מתי מתקיים ש־$A^m = I$ (בעבור m מינימלי)? במילים אחרות, מתי מתחילים מחזור.

היות שמספר הזוגות השונים עבור $\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix}$ הוא p^2, אז $m \leqslant p^2$. עבור $p = 7$: $0, 1, 1, 2, 3, 4, 5, 1, 6, 0, 6, 6, 5, 4, 2, 6, 1, 0, 1$ – כלומר עבור $p = 7$ יש מחזור באורך $m = 16$.

הערה 4. תירואטית עם המידע הנוכחי ייתכן ויהפוך למחזורי ולא יחזור להתחלה

טענה. אם p ראשוני אז $p \equiv 1 \pmod 5$ אז אורך המחזור חסום מלעיל ע"י $p - 1$.

הוכחה. תנאי מספיק (אך לא הכרחי) לקבלת מחזור באורך k הוא $A^k = I$. אז:

$$f_A(x) = x^2 - x - 1$$

יש דבר שנקרא "הדדיות ריבועית" (חומר קריאה רשות במודל) שמבטיחה שורש לפולינום להלן עבור p כנ"ל. אכן יש לנו שני ע"ע שונים (אם קיים רק אחד אז סתירה מהיות הדיסקרימיננטה $5 = 0$ אך $p \not\equiv 1 \pmod 5$). לכן קיימת P הפיכה כך ש־:

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

כך ש־$\lambda_1, \lambda_2 \neq 0$. משפט פרמה הקטן אומר ש־$\lambda_1^{p-1} = \lambda_2^{p-1} = 1$. ואז $A^{p-1} = I$. ■

1.1.6 $\sim$ שילוש

הגדרה 15. $T\colon V \to V$ ט"ל ניתנת לשילוש אם קיים בסיס B ל־V כך ש־$[T]_B$ משולשית.

הערה 5. אם T ניתנת לשילוש אז הפולינום האופייני שלה מתפרק לגורמים ליניארים (האם איברי האלכסון של הגרסה המשולשית). יהיה מעניין לשאול אם הכיוון השני מתקיים.

משפט 14. $T\colon V \to V$ ט"ל. נניח ש־$f_T(x) = \prod_{i=1}^{n}(x - \lambda_i)$ (ניתנת לפירוק לגורמים ליניאריים) אז T ניתנת לשילוש.

הוכחה. בסיס. $n = 1$ היא כבר משולשית וסיימנו.

צעד. נניח שהטענה נכונה בעבור n טבעי כלשהו, ונראה נכונות עבור $n{+}1$. אז f_T מתפרק לגורמים ליניאריים, לכן יש לו שורש. יהי λ ע"ע של T. בסיס B של V מקיים ש־$[T]_B$ משולשית עליונה (נסמן $B = (w_1 \dots w_{n+1})$) $\iff$ אז $T(w_i) \in \operatorname{span}(w_1 \dots w_i)$. נגדיר את w_1 להיות ו"ע של λ. נשלימו לבסיס B^1.

$$[T]_B = \begin{pmatrix} \lambda & & * & \\ 0 & & \vdots & \\ \vdots & \cdots & C & \cdots \\ 0 & & \vdots & \end{pmatrix}$$

אז ניתן לומר כי:

$$f_T(x) = (x - \lambda) f_C(x)$$

נסמן $w = \operatorname{span}(w_2 \dots w_{n+1})$. קיימת העתקה ליניארית $S\colon W \to W$ כך ש־$f_S(x) = f_C(x)$. לפי ה״א קיים בסיס ל־W הוא B'' שעבורו S משולשית עליונה. נטען ש־$B = B'' \cup \{w_1\}$ ייתן את הדרוש.

$$\forall w \in B''\colon (T-S)(w) = Tw - Sw = aw_1 + S(w) - S(w) = aw_1$$

(כלומר, השורה העליונה של $[T]_B$ ״תרמה״ את aw_1 בלבד) לכן:

$$(T-S)w \subseteq \operatorname{span}(w_1)$$

זה גורר שלכל $W \in w$ מליניאריות מתקיים ש־$(T-S)w \subseteq \operatorname{span}(w_1)$. סה״כ לכל $\{w_1\} \cup B'' \in w$ מתקיים $T(w_i) \in \operatorname{span}(w_1 \dots)$. ■

בהוכחה הזו, בנינו בסיס כך ש־:

$$[T(w) - S(w)]_B = ae_1$$

הגדרה 16. מטריצה ניתנת לשילוש אם היא דומה למשולשית.

משפט 15. מטריצה A ניתנת לשילוש, אמ״מ הפ״א האופייני שלה מתפצל לגורמים ליניארים.

המשך בעמוד הבא

1.1.7 ~ על ההבדל בין פולינום לפולינום

נבחין ש־$\mathbb{F}[x]$ הוא מ"ו מעל $\mathbb{F}$. וכן $\mathbb{F}[x]$ הוא חוג חילופי עם יחידה. בחוג כפל לא חייב להיות קומטטיבי (נאמר, חוג המטריצות הריבועיות). אומנם קיימת יחידה (פולינום קבוע ב־1) אך אין הופכיים לשום דבר חוץ מלפונ' הקבועות. שזה מאוד חבל כי זה כמעט שדה. בהמשך, נגדיר את אוסף הפונקציות הרציונליות כדי להתגבר על כך.

אם נתבונן במטריצות דומות, יש הבדל בין להגיד $f_A(x) \in \mathbb{F}[x]$, אך אפשר לטעון $f_A(x) = |B|$ כש־$B \in M_n(\mathbb{F}(x))$. למה? כי $xI - A \in M_n(\mathbb{F}(x))$ (זה קצת מנוון כי איברי המטריצה הם או פולינומים קבועיים או ממעלה 1). משום שדטרמיננטה שולחת איבר לשדה, אז $|B| \in \mathbb{F}(x)$. כך למעשה נגיע לכך שפולינומים אופייניים שווים כשני איברים בתוך השדה, ולא רק באיך שהם מתנהגים ביחס לקבועיים.

דוגמה:

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in M_2(\mathbb{F}_2),\ f(x) = x^3,\ g(x) = x,\ f, g \in \mathbb{F}_2 \to \mathbb{F}_2 \implies f = g$$

אך:

$$f(A) = A^3 = 0,\ g(A) = A \neq 0$$

זה לא רצוי. נבחין בשני שוויונות שונים – שוויון פונקציות, בהם $f = g$ מעל $\mathbb{F}_2$, ושוויון בשדה – בו $f - g \neq 0$ (כי $-x^2$ לא פולינום האפס, ואף מעל $\mathbb{F}_2$) ולכן ב־$\mathbb{F}_2(x)$ מתקיים $f \neq g$.

1.1.8 ~ משפט קיילי־המילטון

הגדרה 17. שדה $\mathbb{F}$ נקרא סגור אלגברית אם כל פולינום f ב־$\mathbb{F}[x]$ ניתן לבטא כמכפלה של גורמים לינארים $(x-a)$ כאשר $a \in \mathbb{F}$, עד לכדי כפל בסקלר.

הגדרה 18. יהי $f(x) = \sum_{i=0}^{d} a_i x^i \in \mathbb{F}[x]$, V מ"ו מעל $\mathbb{F}$ נ"ס (נוצר סופית) וכן $T \colon V \to V$ ט"ל. נגדיר:

$$f(T) = \sum_{i=0}^{d} a_i T^i,\ T^0 = id,\ T^n = T \circ T^{n-1}$$

כנ"ל עם מטריצות (ראה תרגול)

טענה. אם $[T]_B = A$ ו־$f(x) \in \mathbb{F}[x]$, אז $[f(T)]_B = f(A)$. הוכחה נובעת מהתכונות $[TS]_B = AC,\ [T+S]_B = A + C,\ [\alpha T]_B = \alpha A,\ [S]_B = C < [T]_B = A$.

טענה. אם $f, g \in \mathbb{F}[x]$ ו־$T \colon V \to V$ ט"ל, אז $(f \cdot g)(T) = f(T) \cdot g(T)$. באופן דומה $(f+g)(T) = f(T) + g(T)$.

לכן קל לראות ש־$f(A) = 0 \iff f(T) = 0$.

מסקנה 5. אם A, C דומות אז $f(C) = 0 \iff f(A) = 0$.

דוגמה. (מנוונת) נתבונן ב־$D \colon \mathbb{F}_n[x] \to \mathbb{F}_n[x]$ אופרטור הגזירה. ראינו $f_D(x) = x^{n+1}$ (הפולינום האופייני). אז נקבל:

$$f_D(D)(p) = p^{(n+1)} = 0 \implies f_D(D) = 0$$

משפט 16 (משפט קיילי־המילטון). תהי $T \colon V \to B$ ט"ל (V נ"ס) או $A \in M_n(\mathbb{F})$, ו־$f_A(x) = f_T(x)$ הפ"א, אז $f_T(T) = 0,\ f_A(A) = 0$.

הערה 6. באנגלית, Cayley–Hamilton

הוכחה. נוכיח את המשפט בשלושה שלבים –

- נניח ש־T ניתנת לשילוש. אזי, קיים בסיס $B =: (v_1 \dots v_n)$ כך ש־$[T]_B$ משולשית (עליונה). זאת מתקיים אמ"מ $\forall i \in [n] \colon Tv_i \in \operatorname{span}(v_1 \dots v_i)$. נפנה להוכיח את משפט קיילי־המילטון למקרה זה.

  תת־הוכחה.

  - בסיס: בעבור $n = 1$, אז קיים $\lambda \in \mathbb{F}$ כך ש־$f_T(T) = T - \lambda I = 0$ (העתקה לינארית חד ממדית היא כפל בסקלר). בפרט $\forall v \in V \colon (T - \lambda)v = 0$.
  - צעד: נניח ש־$B = (v_1 \dots v_n, v_{n+1})$ שעבורו $[T]_B$ משולשית. נגדיר תמ"ו $W = \operatorname{span}(v_1 \dots v_n)$ כך ש־$\dim W \leqslant \dim V$, ועבורו נכון $\forall w \in W \colon Tw \in W$ (ניתן להראות שזה נכון עבור וקטורי הבסיס, ונכון לכל $w \in W$ מלינאריות). נגדיר $T|_W \colon W \to W$ את הצמצום של T ל־W. ידוע ש־$T|_w$ ניתנת לשילוש ולכן מקיימת את תנאי האינדוקציה.

לכן, $\forall w \in W\colon f_{T|_W}(T)(w) = 0$. אזי $f_{T|_W}(x) = \prod_{i=1}^{n}(x-\lambda_i)$ וסה"כ $f_T(x) = (x-\lambda_{n+1})f_{T|_W}(x)$ וקיבלנו $\forall w \in W\colon f_T(T)(w) = 0$.

מספיק להראות ש־$\forall v \in V\colon (T-\lambda_{n+1})v \in W$. למה? כי:

$$f_T(T)(v) = \left(\prod_{i=1}^{n}(T-\lambda_i)\right)(T-\lambda_{n+1})(v)$$

מלינאריות, מספיק להראות ש־$(T-\lambda_{n+1})(v_{n+1}) \in W$, שכן זה מתקיים על כל בסיס אחר. אך זה ברור – עבור $[T]_B$ העמודה האחרונה היא:

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \\ \lambda_{n+1} \end{pmatrix}$$

ולכן:

$$T(v_{n+1}) = \alpha_1 v_1 + \cdots + \alpha_n v_n + \lambda_{n+1}v_{n+1} \implies (T-\lambda_{n+1})(v_{n+1}) = \alpha_1 v_1 + \cdots + \alpha_n v_n \in W$$

■

- נוכיח בעבור מטריצה משולשית/ניתנת לשילוש.

תת־הוכחה. אם A משולשית, אז $f_A(x) = f_{T_A}(x)$ כאשר $T_A\colon \mathbb{F}^n \to \mathbb{F}^n$ המוגדרת ע"י $T_A(v) = Av$, ואז T_A ניתנת לשילוש וסיימנו.

אם A ניתנת לשילוש, אז היא דומה למשולשית, והן בעלות אותו הפולינום האופייני. אז חזרה לתחילת ההוכחה. ■

- עבור T כללית או A כללית.

תת־הוכחה. נניח $A = [T]_B$ עבור בסיס B, וידוע $f_T(x) = f_A(x)$. ידוע ש־A ניתנת לשילוש אמ"מ $f_A(x)$ מתפצל. טענה מהעתיד הלא רחוק: לכל שדה $\mathbb{F}$ קיים שדה $\mathbb{K} \supseteq \mathbb{F}$ סגור אלגברית (כל פולינום מעל שדה סגור אלגברית מפתצל). על כן, ניתן לחשוב על $A \in M_n(\mathbb{F})$ כמו $A \in M_n(\mathbb{K})$. הפולינום האופייני מעל K הוא אותו הפולינום האופייני מעל $\mathbb{F}$. לכן הוא מתפצל (מעל $\mathbb{K}$), ולכן הוא דומה למשולשית, ומהמקרה הראשון $f_A(A) = 0$. זאת כי $f_A(A)$ לא תלוי בשדה עליו אנו עובדים, וסה"כ הוכחנו בעבור מטריצה כללית, ולכן לכל ט"ל. ■

■

משפט 17. אם A מייצגת של העתקה T, ו־$f \in \mathbb{F}[x]$, אז $f(T) = 0 \iff f(A) = 0$.

הערה 7 (בנוגע לשדות סגורים אלגברית). הטענה שלכל שדה יש שדה שסגור אלגברית – טענה שתלויה באקסיומת הבחירה. הסגור האלגברי הוא יחיד. **הטענה הזו לא נאמרת באופן רשמי בקורס** על אף שהרחבה לשדה סגור אלגברית מועילה מאוד בלינארית 2א באופן כללי.

המשך בעמוד הבא

1.2 Ring Theory

1.2.1 ~ מבוא והגדרות בסיסיות

אז, מה זה אובייקט אלגברי? הרעיון – "Data עם אקסיומות". אנו כבר מכירים רבים מהם: תמורות, חבורות, שדות, מרחבים וקטורים, ועוד. עתה נכיר אובייקט אלגברי בשם חוג. מקובל לסמן חוג בתור השלישייה הסדורה $(R,+,\cdot)$ כאשר $\cdot: R\times R\to R$ ו־$+: R\times R\to R$ הפעולות הבינאריות בחוג R, כאשר $+$ קומטטיבי וקיים נגדי, וכן הפעולות הבינאריות דיסטרבוטיביות.

הגדרה 19. חוג עם יחידה הוא קבוצה עם שתי פעולות, כפל וחיבור, ניטרלים לפעולות (0, 1) כך שמתקיימות כל אקסיומות השדה למעט (פוטנציאלית) קיום איבר הופכי, וקומטטיביות הכפל.

אנחנו נתעניין ספציפית בחוגים קומטטיבים, כלומר, בהם הכפל כן קומטטיבי. המטריצות הריבועיות מעל אותו הגודל, לדוגמה, הוא חוג שאיננו קומטטיבי. החוג ה"בסיסי ביותר" – חוג השלמים (אין הופכי) הוא חוג קומוטטיבי. ישנם חוגים בלי יחידה (לדוגמה הזוגיים בלי יחידה), שלא נדבר עליהם כלל.

הגדרה 20. תחום שלמות הוא חוג קומוטטיבי עם יחידה ללא מחלקי 0.

הערה 8. באנגלית, Integral Domain

הגדרה 21. חוג ייקרא ללא מחלקי 0 אם:
$$\forall a,b\in R\colon ab=0\implies a=0\lor b=0$$

דוגמאות לחוגים עם מחלקי 0:

- $M_2(\mathbb{R})$: הוכחה $a\cdot b=0$, $a=b=\left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right)$
- $\mathbb{Z}/_{6\mathbb{Z}}$ הוכחה $2\cdot 3=0$.

משפט 18. בתחום שלמות יש את כלל הצמצום בכפל: אם $ab=ac\land a\neq 0$ אז $b=c$.

הוכחה.
$$ab\cdot ac=0\implies a(b\cdot c)=0\implies a=0\lor b-c=0$$

בגלל ש־$a\neq 0$, אז $b-c=0$. נוסיף את c הנגדי של $-c$ ונקבל $b=c$. ■

דוגמאות לתחום שלמות:

- שדות
- השלמים
- חוג הפולינומים

1.2.2 ~ ראשוניות ואי־פריקות

הגדרה 22. יהי R תחום שלמות, $a,b\in R$. נאמר ש־$a\mid b$ אם קיים $c\in R$ כך ש־$b=ac$.

הגדרה 23. $u\in R$ נקרא הפיך אם קיים $\alpha\in R$ כך ש־$\alpha u=1$.

משפט 19. יהי R תחום שלמות, $u\in R$ הפיך. יהי $a\in R$. אז $a\mid u$.

הוכחה. $1\mid a,\ u\mid 1$. יחס החלוקה טרנזטיבי ולכן $a\mid u$. ■

סימון 3. קבוצת ההפיכים מוסמנת ב־R^x.

דוגמאות.

1. אם $R=\mathbb{F}$, אז $\mathbb{F}^x=\mathbb{F}\setminus\{0\}$
2. אם $R=\mathbb{Z}$ אז $\mathbb{Z}^2=\{\pm 1\}$
3. אם $R=\mathbb{F}[x]$ אז $R^x=\mathbb{F}^x$ (ההתייחסות לסקלרים $\mathbb{F}$ היא כאל פונקציות קבועות)

הגדרה 24. $a,b\in R$ נקראים חברים אם קיים $u\in R^x$ הפיך כך ש־$a=ub$, ומסמנים $a\sim b$

משפט 20. יחס החברות הוא יחס שקילות.

הוכחה.

א. $a\sim a$ כי $1\in R^x$

ב. אם $a\sim b$ אז קיים $u\in R^x$ כך ש־$a=ub$. קיים ל־u הופכי α אז $\alpha a+\alpha ub=b$ ולכן $b\sim a$.

ג. נניח $a\sim b\land b\sim c$, כי מכפלת ההופכיים הפיכה $a\sim c$ וסיימנו.

■

"אני אני חבר של עומר, ועומר חבר של מישהו בכיתה שאני לא מכיר, אני לא חבר של מי שאני לא מכיר." המרצה: "למה לא? תהיה חבר שלו".

משפט 21. הופכי הוא יחיד.

(אותה ההוכחה כמו בשדות)

הוכחה. יהי $a \in R^x$ ו־u, u' הופכיים שלו, אז:

$$u = u \cdot 1 = u \cdot a \cdot u' = 1 \cdot u' = u'$$

■

הערה 9. המשפט להלן נכון לא רק בתחום שלמות, אלא בכל חוג

משפט 22. בהינתן תחום שלמות R ו־$a, b \in R$, אז אם $b \mid a$ וגם $a \mid b$ אז $a \sim b$ (ביחס החברות).

הוכחה.

$$a \mid b \implies \exists c \in R\colon ac = b$$
$$b \mid a \implies \exists d \in R\colon bd = a$$

לכן:

$$ac = b \implies acd = a \implies a(cd-1) = 0 \implies a = 0 \vee cd = 1$$

■

אם $a = 0$ אז $b = 0$ (ממש לפי הגדרה) ו־$\sim$ שקילות (רפליקסיביות). אחרת, $cd = 1$ ולכן c הפיך, סה"כ $a \sim b$.

"אני חושב שב[אוניברסיטה ה]עברית קראו להם ידידים, לא רצו להתחייב לחברות ממש".

הגדרה 25. איבר $p \in R$ נקרא אי־פריק אם מתקיים $p = ab \implies a \in R^x \vee b \in R^x$.

הגדרה 26. איבר $p \in R$ יקרא ראשוני אם $p \mid (a \cdot b) \implies p \mid a \vee p \mid b$.

הערה 10. איברים הפיכים לא נחשבים אי־פריקים או ראשוניים. הסיבה להגדרה: בשביל נכונות המשפט היסודי של האריתמטיקה (יחידות הפירוק לראשוניים).

משפט 23. בתחום שלמות כל ראשוני הוא אי פריק.

הערה 11. שקילות לאו דווקא.

הוכחה. יהי $p \in R$ ראשוני. יהיו $a, b \in R$ כך ש־$p = ab$. בה"כ $a \mid p$. אז קיים c כך ש־$pc = a$ ולכן $pcb = p$. סה"כ $p \neq 0$ ולכן $cb = 1$ (ראה לעיל) ו־b הפיך. ■

הגדרה 27. R תחום פריקות יחידה אם $\prod_{i=1}^{n} p_i = \prod_{j=1}^{m} q_i$ עבור p_i, q_j ראשוניים, אז $n = m$, ועד לכדי סידור מחדש, לכל $i \in [n]$ $p_i \sim q_i$.

הערה 12. באנגלית, Unique Factorization Domain.

משפט 24. נניח שבתחום שלמות R, כל אי־פריק הוא גם ראשוני. אז R תחום פריקות יחידה.

ההוכחה: זהה לחלוטין לזו של המשפט היסודי.

הוכחה. באינדוקציה על $n + m$. בסיס: $n + m = 2$ ולכן $n = m = 1$ (כי מעפלה ריקה לא רלוונטית מאוד) אז $p = q$. נעבור לצעד. נניח שהטענה נכונה לכל $n + m < k$. נניח ש־$n + m = k$. אז $p_1 \mid \prod_{j=1}^{m} q_j$. בה"כ $p_1 \mid q_1$. q_1 אי־פריק ולא הפיך. p_1 לא הפיך. לכן $p_1 \sim q_1$. אז עד כדי כפל בהופכי נקבל ש־$\prod_{i=2}^{n} p_i = \prod_{j=2}^{n} q_j$. הערה: ראשוני כפול הפיך נשאר ראשוני. מכאן הקענו לדרוש וסיימנו (הערה שלי: כאילו תכפילו בחברים ותקבלו את מה שצריך). ■

הגדרה 28. יהי R תחום שלמות. תת־קבוצה $0 \neq I \subseteq R$ נקראת אידיאל אם:

א. $\forall a, b \in I\colon a + b \in I$ – סגירות לחיבור.

ב. $\forall a \in I\, \forall b \in R\colon ab \in I$ – תכונת הבליעה. [בפרט $0 \in I$]

דוגמאות:

1. 0 תמיד אידיאל, וכן החוג כולו תמיד אידיאל.

2. הזוגיים ב־$\mathbb{Z}$.
3. לכל $n \in \mathbb{Z}$, $n\mathbb{Z}$ אידיאל (n כפול השלמים). הזוגיים לדוגמה, מקרה פרטי הוא $2\mathbb{Z}$.
4. $\langle f \rangle \subseteq \mathbb{F}[x]$ המוגדר לפי $\langle f \rangle := \{g \in \mathbb{F}[x] \mid f|g\}$
5. הכללה של הקודמים: עבור $a \in R$ נסמן $\langle a \rangle := \{a \cdot b \mid b \in R\}$ הוא אידאל.
6. $I = \{f \in \mathbb{F}[x] \mid f(0) = 0\}$ (לעיתים מסומן $\forall a \in R \colon aR = \langle a \rangle$)
7. נוכל להכליל את 4 עוד: ("הכללה של הכללה היא הכללה. זה סגור להכללה. זה קורה הרבה במתמטיקה")

$$I = aR + bR = \{ar + bs \mid r, s \in R\}$$

וניתן להכליל עוד באינדוקציה.

הגדרה 29. אידיאל I נקרא ראשי אם הוא מהצורה aR עבור $a \in R$ כלשהו.

סימון 4. $Ra =: (a) =: \langle a \rangle =: \{ar \mid r \in R\}$

הגדרה 30. תחום שלמות נקרא תחום ראשי אם כל אידיאל שלו ראשי.

הערה 13. באנגלית, Principal Ideal Domain או בקיצור PID.

הערה 14. אנחנו סימנו אידיאל שנוצר ע"י $a \in R$ ב־aR ובקורס מסמנים Ra, באופן כללי אפשר לדבר על אידיאל שמאלי ואידיאל ימני. תחת ההנחה שהחוג קומטטיבי שני הסימונים שקולים בכל מקרה.

משפט 25. ב־$R \neq \{0\}$ תחום ראשי אז כל אי פריק הוא ראשוני.

(את הכיוון השני כבר הוכחנו בעבור תחומי שלמות באופן כללי)

הוכחה. יהי o אי פריק (א"פ). יהיו $a, b \in R$ כך ש־$p \mid ab$. ניעזר באידיאל $I = Ra + Rp$. בכלל ש־R תחום ראשי, קיים $c \in R$ כך ש־$I = Rc$, ו־$a, p \in I$ כלומר $p \mid c \wedge c \mid a$. p א"פ ולכן $p \sim c$ או c הפיך.

- c הפיך $\Longleftarrow\ \in I \Longleftarrow R = R \cdot 1 \in I \subseteq R \Longleftarrow I = R$. קיימים $r, s \in R$ כך ש־$ra + sp = 1$. נכפיל ב־b ונקבל $rab + spb = b$ וסה"כ $p \mid b$.
- אם $p \sim c$, אז $a \mid c \wedge c \mid p$ ולכן $p \mid a$. ■

מסקנה 6. אם R תחום שלמות ראשי אזי יש פריקות יחידה למכפלה של אי פריקים עד כדי חברות.

משפט 26. יהיו $a, b \in R$, אז a, b ייקראו זרים אם $\forall c \in R \colon c \mid a \wedge c \mid b \implies c \in R^x$

הגדרה 31. יהי $g \in R$ כך ש־:

1. $g \mid a \wedge g \mid b$
2. $\forall \ell \in R \colon \ell \mid a \wedge \ell \mid b \implies \ell \mid g$

אז g כנ"ל הוא הגורם המשותף המקסימלי של a, b, הוא $\gcd(a, b)$.

משפט 27. יהי R תחום שלמות ויהיו $a, b \in R$. נניח שקיימים $r, s \in R$ כך ש־$g = ra + sb$ אשר מחלק את a, b. אז:

1. $\gcd(a, b) = g$
2. ה־$\gcd$ מוגדר ביחידות עד לכדי חברות.
3. בתחום ראשי, לכל a, b קיים g כנ"ל.

הוכחה.

1. יהי $\ell \mid a, b$ אז $\ell \mid ra, sb$ וסה"כ $\ell \mid g$.
2. מ־1 (בערך) אם g, g' מקיימים את היותם $\gcd$ אז $g \mid g' \wedge g' \mid g$ ולכן $g \sim g'$.
3. נסמן $I = Ra + Rb$. אז $I = Rg$, וקיימים $r, s \in R$ כך ש־$g = ra + sb$ ולכן $a, b \in I$ אז $g \mid a, b$ וסיימנו מ־1. ■

מסקנה 7 (אלגוריתם אוקלידס המורחב). בתחום ראשי, אם a, b זרים אז $\exists r, s \in R \colon ra + sb = 1$.

משפט 28. $\mathbb{F}[x]$ תחום ראשי.

הוכחה. יהי $I \subseteq \mathbb{F}[x]$ אידיאל. אם $I = \{0\}$, הוא ראשי. אחרת, $I \neq \{0\}$, ואז: יהי $0 \neq p \in I$ פולינום מדרגה מינימלית, ויהי $f \in I$. אז קיימים $q, r \in \mathbb{F}[x]$ כך ש־$f = qp + r$. ידוע $\deg r < \deg p$. בגלל ש־$p \in I \wedge f \in I$ אז $f - qp \in I$. אם r אינו 0, קיבלנו סתירה למינימליות הדרגה של p. ■

הוכחה זהה עובדת בשביל להראות ש־$\mathbb{Z}$ תחום ראשי, אך עם דרגה במקום ערך מוחלט.

הגדרה 32. תחום שלמות נקרא אוקלידי אם קיימת $N\colon R\setminus\{0\}\to\mathbb{N}_+$ כך ש־$\forall a,b\in R\setminus\{0\}\colon \exists u,r\in R\colon a=ub+r$ כאשר $r\neq 0$ או $N(b)>N(r)$, ו־N סאב־כפלית כלומר $\forall 0\neq a,b\in R\colon N(a)\leqslant N(ab)$.

ברגע שיש לנו את ההגדרה של תחום אוקלידי, N הפונקציה שתשתמש אותנו בשביל להראות את ההוכחה שכל תחום אוקלידי הוא תחום ראשי (בדומה לערך מוחלט או deg בהוכחות קודמות). ההפך נכון תחת השערת רימן המוכללת (לא ראיתם את זה צץ, נכון?).

אינטואציה לחוג אוקלידי היא "חלוקה עם שארית", כאשר פונקצית הגודל N דורשת שהשארית תהיה "אופטימלית". בחוג הפולינומים $N=\deg$ (פרטים בהמשך), ובחוג המספרים השלמים $N=|\cdot|$.

דוגמה לחוג שאינו אוקלידי: $\mathbb{Z}[\sqrt{-5}]$ הוא $\{a+b\sqrt{-5}\mid a,b\in\mathbb{Z}\}$.

משפט 29. חוג אוקלידי $\Longleftarrow$ תחום פריקות יחידה (גרסה מוכללת של המשפט היסודי של האריתמטיקה).

משפט 30. חוג אוקלידי $\Longleftarrow$ תחום שלמות.

(הוכחה בויקיפדיה)

לדוגמה בחוג לעיל $6=2\cdot 3=(1+\sqrt{-5})(1-\sqrt{-5})$ על אף ש־$2,3$ אי־פריקים וכן $(1+\sqrt{-5}),(1-\sqrt{-5})$ אי פריקים.

דוגמה (חוג השלמות של גאוס).

$$R=\mathbb{Z}[i]=\{a+bi\mid a,b\in\mathbb{Z}\}$$

הנורמה:

$$N\colon\mathbb{R}\to\mathbb{Z}_{\geqslant 0},\ \ N(a+bi)=a^2+b^2=|a+bi|^2$$

בדומה להוכחה לפיה הערך המוחלט של מורכב הוא כפלי, ניתן להראות ש־N כפלית. מי הם ההפיכים ב־$\mathbb{Z}[i]$? מי שמקיים $\alpha\beta=1$, כלומר:

$$N(\alpha)N(\beta)=1=N(1),\ \ \alpha=a+bi,\ \ a^2+b^2=1\implies a+bi=\pm1,\pm i$$

בהנחה שמוגדרת נורמה כזו, החוג הוא אוקלידי (תנאי זה הכרחי אך לא מספיק).

הערה 15. למספרים הראשוניים בחוג השלמות של גאוס קוראים "ראשוניים גאוסייאנים", והם מקיימים תכונות מעניינות. בפרט אפשר להוכיח ש־p הוא ראשוני בחוג השלמות של גאוס אמ"מ $N(p)$ הוא ראשוני ב־$\mathbb{Z}$, או $p\equiv 4n+3$ כאשר $\equiv$ יחס החברות.

משפט 31. יהי $p\in\mathbb{Z}$ ראשוני. התנאים הבאים שקולים:

- p פריק ב־$\mathbb{Z}[i]$
- $p=m^2+n^2$ עבור $n,m\in\mathbb{Z}$
- $p=2$ או $p\equiv 1\pmod 4$
- קיימים $r,s\in R$ כך ש־$ra+sb=1$

שימו לב ש־$\mathbb{Z}$ בתוך $\mathbb{Z}[i]$ לא סגורים לבליעה.

הגדרה 33. $I\subseteq R$ אידיאל נקרא ראשוני אם $\forall a,b\in R\colon (a\cdot b)\subseteq P$, אז $(a)\in P\vee(b)\in P$.

הגדרה 34. אידיאל $I\subseteq R$ נקרא אי־פריק אם $\forall a,b\in R$ אם $I=(a\cdot b)$ אז $I=(a)$ או $I=(b)$.

ראינו, שבתחום ראשי אי פריק אמ"מ ראשוני. ניתן להראות דומה ניתן לטעון ש־:

משפט 32. R תחום ראשי, אז I ראשוני אמ"מ I אי־פריק.

הגדרה 35. יהי R תחום שלמות [אפשר להתעסק גם עם אידיאל ימני ושמאלי] ונניח ש־$I\subseteq R$ אידיאל. אז $R/_I:=\{a+I\mid a\in R\}$ הוא חוג (בהגדרת $a+I=\{a+i\mid i\in I\}$ חיבור מנות), כאשר הפעולות:

- $(a+I)+(b+I)=(a+b)+I$
- $(a+I)(b+I)=ab+I$

עקרונית צריך להוכיח שהחיבור/כפל לא תלוי בנציגים a,b כדי שהחוג יהיה מוגדר היטב (זה כמובן מתקיים בתחום ראשי).

1.2.3 ∼ הרחבת שדות

משפט 33. בתחום ראשי, R, אם I אידיאל אי־פריק, אז $R/_I$ שדה.

דוגמאות.

- $\mathbb{Z}/_{\langle p\rangle}$ שדה.

- $\mathbb{R}[x]$ תחום ראשי, ידוע x^2+1 אי־פריק. לכן $\mathbb{R}[x]/_{\langle x^2+1\rangle} \cong \mathbb{C}$. הרעיון: נוכל להסתכל על p פולינום המבוטא כמו:

$$p(x) = q(x)\cdot(x^2+1) + ax + b = ax + b + \langle x^2+1\rangle$$

ואם נכפיל שני יצורים כאלו:

$$((ax+b)+I) + (cx+d+I)) = acx^2 + (ad+bc)x + bd + I$$

אך ידוע ש־$x^2+1=0$ (כי זה האידיאל שלנו) עד לכדי נציג, כלומר מתקיים שוויון ל־$bd-ac+(ad+bc)x+I$ וזהו כפל מרוכבים. אפשר גם להראות קיום הופכי וכו'.

הוכחה. יהי $0 \neq a + I \in R/_I$. אם $a \neq 0$, אז ב־R מתקיים $p \nmid a$ אי p א"פ (אם הוא היה מחלק את a אז $a=0$) ולכן p, a זרים (כי האידיאל אי פריק וכו'). אז קיימים $r, s \in R$ כך ש־$ar+ps=1$. סה"כ:

$$ar + ps + I = 1 + I \implies ar + I = 1 + I \quad \top$$

לכן $r+I$ הופכי של $a+I$ וסיימנו. ■

(למעשה זה אמ"מ – הכיוון השני תרגיל בעבור הקורא).

הגדרה 36. יהי R תחום שלמות, $a_1 \dots a_n \in R$ ו־$\ell = \operatorname{lcm}(a_1 \dots a_n)$ אמ"מ:

1. $\forall i \in [n] \colon a_i \mid \ell$
2. $\forall b \in R \colon \forall i \in [n] \colon a_i \mid b \longrightarrow \ell \mid b$

דוגמה. $R = \mathbb{Z},\ \operatorname{lcm}(2,6,5) = 30$.

משפט 34. יהי $\mathbb{F}$ שדה ויהי $f \in \mathbb{F}[x]$ פולינום אי־פריק ממעלה $\deg f > 1$. אז קיים $\mathbb{K} \supseteq \mathbb{F}$ כך שב־$\mathbb{K}$ יש ל־f שורש.

ההוכחה למשפט קונסטקרטיבית, ובה צריך להראות שהקבוצה:

$$\mathbb{K} = \{p(A_f) \mid p \in \mathbb{F}[x]\}$$

עם חיבור וכפל מטריצות, היא שדה. השיכון $\alpha \mapsto \alpha I$ משכן את $\mathbb{K} \mapsto \mathbb{F}$.

משפט 35. (ללא הוכחה בקורס, תלוי באקסיומת הבחירה) לכל שדה $\mathbb{F}$ קיים ויחיד שדה $\mathbb{K} \subseteq \mathbb{F}$ סגור אלגברית.

דוגמה: $\mathbb{R}$ ו־$\mathbb{C}$.

1.2.4 ~ חוג הפולינומים

(תת־פרק זה לקוח מתרגול בקורס)

הגדרה 37. הדרגה של הפולינום היא $\deg(f) := \max\{n \in \mathbb{N} \mid a_n \neq 0\}$, ומגדירים $\deg(0) = -\inf$

משפט 36.

$$\deg(fg) = \deg f + \deg g \quad \deg(d+g) = \max\{\deg f, \deg g\}$$

הערה 16. חוג הפולינומים הוא חוג אוקלידי כי פונקציית הגודל $N = \deg f$ מקיימת את התנאי של חוג אוקלידי. לכן ממשפט הוא תחום ראשי.

מסקנה 8. לכל $f, g \in \mathbb{F}[x]$, אם $g \neq 0$ אז קיימים ויחידים פולינומים $q, r \in \mathbb{F}[x]$ כך ש־$f = qg + r \wedge \deg r < \deg g$.

הגדרה 38. נאמר שפולינום q מחלק את f אם $r=0$ ומסמנים $q \mid f$.

מסקנה 9.

1. $(x-a) \mid f \iff f(a) = 0$ (משפט בזו)
2. אם $\deg f = n > -\inf$, ל־f לכל היותר n שורשים כולל ריבוי.
3. נניח ש־$f, g \in \mathbb{F}[x]$ ו־$\mathbb{K} \subseteq \mathbb{F}$, כאשר K שדה. אם $f \mid g$ מעל K אז $f \mid g$ מעל $\mathbb{F}$.

הוכחה.

1. הוכחה למשפט בזו:

$\Longrightarrow$ נניח $(x-a) \mid f$. אז קיים פולינום g כך ש־$f = (x-a)g$, קרי $f(a) = (a-a)g(a) = 0$.

$\Longleftarrow$ נניח $f(a) = 0$. אז קיימים $q, r \in \mathbb{F}[x]$ כך ש־$f = q(x-a)$ ועל כן $0 = f(a) = q(a)(a-a) + r(a) = 0$ ולכן $r(a) = 0$. משום ש־r פולינום קבוע (דרגתו קטנה מ־1, כי חילקנו ב־$(x-a)$ מדרגה 1), אז $r(x) = 0$.

2. אינדוקציה

3. נוכיח ב־"contrapositive": אנו יודעים ש־$P \to \neg Q \iff \neg P \to Q$. נניח ש־$f \nmid g$ מעל $\mathbb{F}$. קיימים $q, r \in \mathbb{F}[x]$ כך ש־$r \neq 0,\ f = qg + r$. הפירוק הזה הוא גם ב־$\mathbb{K}[x]$. מיחידות r, נקבל ש־$f \nmid g$ מעל כל K.

■

"לא הנחתי בשלילה, הוכחתי בקונטראפוזיטיב"

משפט 37. בהינתן $f \in \mathbb{F}[x]$ ו־$r \in \mathbb{N}, \lambda \in \mathbb{F}$ אז λ יקרא שורש מריבוי r של f אם $(x-\lambda)^n \mid f \wedge (x-\lambda)^{n+1} \nmid f$.

משפט 38.

$$\forall f \in \mathbb{F}[x] \colon f(\lambda) = 0 \implies \exists g \in \mathbb{F}[x] \colon f(x) = (x-\lambda)g(x)$$

משפט 39. (באינדוקציה על הטענה הקודמת) בהינתן $\mathbb{F}$ שדה סגור אלגברית:

$$\forall f \in \mathbb{F}[x] \colon \exists (\alpha_i)_{i=1}^{n} \in \mathbb{F}, a_n \in \mathbb{F} \colon a_n \cdot f = \prod_{i=1}^{n} (x - \alpha_i)$$

משפט 40. (מסקנה מהטענה הקודמת שניתן להוכיח באינדוקציה ללא הרחבת שדות) לפולינום $f \in \mathbb{F}[x]$ שאינו אפס יש לכל היותר $\deg f$ שורשים.

הערה 17. שימו לב! כל המסקנות שלנו על תחומים ראשיים תקפים גם על פולינומים. בפרט, ניתן לכתוב כל פולינום $\mathbb{F}[x]$ כמכפלה של גורמים אי־פריקים ב־$\mathbb{F}[x]$ (אם $\mathbb{F}$ סגור אלברית, אלו גורמים לינאריים) עד לכדי סדר וחברות (קבועים).

הערה 18. שימו לב שחלק ניכר מהמשפטים לעיל נכונים בעבור פולינומים מעל שדה ולא מעל כל חוג (בפרט, המשפט לפיו חוג הפולינומים תחום אוקלידי).

עתה נציג משפט פשוט אך מועיל ממתמטיקה B, שלעיתים משמש לניחוש שורשי פולינום ע"מ לפרקו.

משפט 41. יהי $p = \sum_{i=1}^{n} \alpha_i x_i \in \mathbb{Z}[x]$ פולינום עם מקדמים שלמים. יהי $\frac{a}{b} \in \mathbb{Q}$ כך ש־$p\left(\frac{a}{b}\right) = 0$ שורש, ובה"כ $\gcd(a,b) = 1$ (אחרת ניתן לצמצם). אזי $a \mid \alpha_0 \wedge b \mid \alpha_n$.

מסקנה 10. $\forall A \in M_n(\mathbb{F})\ \forall k \geqslant n\ \exists p(c) \in \mathbb{F}_{n-1}[x] \colon A^k = p(A)$.

מסקנה זו נובעת מאלגוריתם לביטוי A^{n+c} כקומבינציה לינארית של $I \cdots A^{n-1}$ שמופיע בסוף הסיכום.

(1.2.4.1) פונקציות רציונליות ומספרים אלגבריים

אינטואיציה: הרעיון של פונקציה רציונלית היא להיות "פולינום חלקי פולינום". נפרמל את הדבר הזה בעבור מרחב פולינומים מעל כל שדה.

משפט 42. בהינתן $\mathbb{F}$ שדה הקבוצה $\{(f,g) \mid f,g \in \mathbb{F}[x], g \neq 0\}$ משרה את יחס השקילות הבא:

$$(f,g) \sim (\tilde{f}, \tilde{g}) \iff f \cdot \tilde{g} = \tilde{f} \cdot g$$

סימון 5. נסמן כל איבר במחלקת השקילות ע"י $\frac{f}{g}$ שמייצגים אותו.

הגדרה 39. שדה הפונקציות הרציונליות הוא הקבוצה $Q[x]$ היא אוסף מחלקות השקילות של $\sim$ מהמשפט הקודם, עם פעולות החיבור והכפל הבאות:

$$\frac{f}{g} \cdot \frac{\tilde{f}}{\tilde{g}} = \frac{f\tilde{f}}{g\tilde{g}} \wedge \frac{f}{g} + \frac{\tilde{f}}{\tilde{g}} = \frac{f\tilde{g} + g\tilde{f}}{g\tilde{g}}$$

למה 1. הגדרות הפעולות לעיל מוגדרות היטב (כלומר הן לא תלויות בנציגים)

משפט 43. $Q[x]$ שדה, כאשר $\frac{0}{1}$ הניטרלי לחיבור ו־$\frac{1}{1}$ הניטרלי לכפל.

המלצה. לקרוא שוב את פרק 2.1, "על ההבדל בין פולינום לפולינום", בו נבחין שלמרות ש־$|\mathbb{F}_2 \to \mathbb{F}_2| = 4$, ישנם אינסוף פולינומים מעל השדה הזה.

אינטואציה. למעשה, נרצה להגיד שדה הפונקציות הרציונליות הוא איזומורפית (קאנונית, ולכן נתייחס אליו כאילו הוא שווה) ל־:

$$Q[x] \cong \left\{ \frac{f(x)}{g(x)} \mid f(x), \underbrace{g(x)}_{\neq 0} \in \mathbb{F}[x] \right\}$$

כאשר $\mathbb{F}[x]$ חוג הפולינומים מעל השדה $\mathbb{F}$. עוד כדאי לציין ש־$Q[x]$ מכיל עותק של $\mathbb{F}[x]$ (עד לכדי איזומורפיזם) בעבור $g = 1$ פולינום היחידה. כמובן ש"איזומורפיזם" בהקשר הזה מדבר על העתקה (לא בהכרח לינארית) שמשמרת את פעולות החוג.

משפט 44. לכל p ראשוני $\forall x \in \mathbb{F}_p \colon x^p = x$.

הערה 19. זוהי מסקנה ישירה מהמשפט הקטן של פרמה.

הגדרה 40. מספר מרוכב $\alpha \in \mathbb{C}$ יקרא מספר אלגברי אם קיים פולינום $0 \neq f \in \mathbb{Q}[x]$ כך ש־$f(\alpha) = 0$.

הגדרה 41. מספר מרוכב שאינו אלגברי יקרא מספר טרנסצנדנטי.

דוגמאות. נבחין ש־$\sqrt{\alpha}$ הוא אלגברי כי הוא שורש של $x^2 - \alpha$. קיימות הוכחות לפיהן e ו־π הם מספרים טרנסצדנטיים.

משפט 45. בהניתן $0 \neq V \subseteq \mathbb{C}$ תמ"ו מעל $\mathbb{C}$, אם $\forall x \in \mathbb{C} \colon xV \subseteq V$ אז x אלגברי.

הוכחה. נגדיר $T_x \colon V \to V$ כך ש־$T_x(v) = xv$ (ההעתקה מוגדרת היטב מהנתון). אזי $f_T(T) = 0$. אזי:

$$f_T(t) =: \sum_{i=1}^{n} a_n t^n \implies 0_V = f(T)v = \sum_{i=1}^{n} a_n T^n v = \left(\sum_{i=1}^{n} a_n x^n\right) v = f(x)v$$

בפרט עבור $v \in V \setminus \{0\}$ יתקיים $f(x) = 0$ ולכן x אלגברי. ■

המשך בעמוד הבא

1.3 Primary Decomposition

1.3.1 ~ מרחבים T-שמורים וציקליים

הגדרה 42. נניח ש-V מ"ו מעל $\mathbb{F}$, ו-$T\colon V\to V$ ט"ל. אז $U\subseteq V$ תמ"ו נקרא T-אינווריאנטי/T-שמור/ה אם לכל $u\in U$ מתקיים $T(u)\in U$.

דוגמאות. $V,\{0\}$ הם T-אינווריאנטים. גם המ"ע (המרחבים העצמיים) הם T-אינוואריאנטי.

הערה 20. שימו לב: אם $U\subseteq V$ תמ"ו אינווריאנטי, אז $T|_U\colon U\to U$ ט"ל.

הערה 21. נניח ש-$u_1\dots u_k$ בסיס ל-U כנ"ל, ו-$W\subseteq V$ תמ"ו כך ש-$V=U\oplus W$, ונגיד ש-$w_{k+1}\dots w_n$ בסיס ל-W, אז $B=(u_1\dots u_k,w_{k+1}\dots w_n)$ מקיים:

$$[T]_B=\begin{pmatrix}[T|_U] & A\\ 0 & B\end{pmatrix}$$

(כאשר $[T|_U]\in M_k$ ו-$B\in M_k$). ותחת ההנחה שאכן T הוא U-איוואריאנטי ו-W-איוואריאנטי, אפשר לייצג אותו באמצעות שתי מטריצות מייצגות על האלכסון (ראה הוכחת המשפט הבא)

משפט 46. יהי V מ"ו, U,W תמ"וים ונניח $V=U\oplus W$ וגם U,W הם T-איוואריאנטים. אז $p_T(x)=p_{T|_U}(x)\cdot p_{T|_W}(x)$

הוכחה. משום ש-$V=U\oplus W$, קיים בסיס $B=(u_1\dots u_k,w_{k+1}\dots w_n)$ כך ש-$u_1\dots u_k$ בסיס ל-U ו-$w_{k+1}\dots w_n$ בסיס ל-W. נבחין, שבייצוג תחת הבסיס הזה:

$$[T]_B=\begin{pmatrix}[T|_U]_B & 0_{n\times(n-k)}\\ 0_{(n-k)\times n} & [T|_W]_B\end{pmatrix}$$

זאת כי לכל $v\in V$ ניתן לייצגו בצורה יחידה כסכום של $u\in U,w\in W$ כך ש-$v=u+w$, כלומר $Tv=Tu+Tw$. ואכן תחת העתקת הקורדינאטות מהגדרת כפל וקטור במטריצה הטענה לעיל מתקיימת. כלומר:

$$p_T(x)=\begin{vmatrix}Ix-[T|_U]_B & 0\\ 0 & Ix-[T|_W]_B\end{vmatrix}=|Ix-[T|_U]_B|\cdot|Ix-[T|_W]_B|=p_{T|_U}(x)\cdot p_{T|_W}(x)$$

כדרוש. ■

משפט 47. בהינתן $U_1\dots U_k$ מרחבים T-איוואריאנטים כך ש-$V=\bigoplus_{i=1}^k U_i$, מתקיים $p_T(x)=\prod_{i=1}^k p_{T|_{U_i}}$.

הוכחה. באינדוקציה על המשפט הקודם. ■

הגדרה 43. יהי V מ"ו מעל $\mathbb{F}$, $T\colon V\to V$ ט"ל ו-$v\in V$ וקטור. אז תת־המרחב־הציקלי הנוצר מ-T ע"י T הוא

$$\mathcal{Z}(T,v):=\operatorname{span}\{T^k(v)\mid k\in\mathbb{N}_0\}$$

משפט 48.

- $\mathcal{Z}(T,v)$ תמ"ו של V – טרוויאלי.
- $\mathcal{Z}(T,v)$ תמ"ו T-איוואריאנטי – טרוויאלי גם.

עתה נציג משהו נחמד. אם V נוצר סופית, גם $\mathcal{Z}(T,v)$ נ"ס. נגיד שיהיה $k\in\mathbb{N}_0$ מינימלי, כך שמתקיים $\mathcal{Z}(T,v)=\operatorname{span}\{v,Tv,\dots T^{k-1}v\}$. אז $T^kv\in\mathcal{Z}(T,v)$. לכן קיימים $a_0\dots a_{k-1}$ כך ש-$T^kv+a_{k-1}T^{k-1}v+\dots+a_0v=0$. ניתן לקחת כבסיס את $v,Tv\dots T^{k-1}v$ של $\mathcal{Z}(T,v)$. אז:

$$[T]_B=\begin{pmatrix}0&0&\cdots&0&-a_0\\ 1&0&\cdots&0&-a_1\\ 0&1&\vdots&0&\vdots\\ 0&0&\cdots&1&-a_{n-1}\end{pmatrix}$$

כאשר השורה האחרונה כי:

$$T(T^{n-1}v)=-\sum_{k=0}^{n-1}a_kT^kv$$

הגדרה 44. $A_f=[T]_B$ היא המטריצה המצורפת לפולינום $f(x)=x^n+a_{n-1}x^{n-1}+\dots+a_0$.

הערה 22. באנגלית: "Companion Matrix", ולעיתים קרויה בעברית "מטריצה מלווה".

1.3.2 ~ הפולינום המינימלי

דיברנו על הפולינום האופייני $f_A = f_T = \det(Ix - A)$. עוד ציינו בהינתו מטריצה, המטריצה המצורפת A_f מקיימת $f_{A_f} = f(x)$.

משפט 49. תהי $A \in M_n(\mathbb{F})$, נביט בקבוצה $I_A = \{p \in \mathbb{F}[x] \colon p(A) = 0\}$, אז $I_A \subseteq \mathbb{F}[x]$ אידיאל, קיים ויחיד ב־I_A פולינום מתוקן בעל דרגה מינימלית.

הגדרה 45. I_A לעיל יקרא הפולינום המיינימלי.

הוכחה. נבחין כי $0 \in I_A$. סגירות לחיבור – ברור. תכונת הבליעה – גם ברור. סה"כ אידיאל. $\mathbb{F}[x]$ תחום שלמות ולכן נוצר ע"י פולינום יחיד $I_A = (p)$. אם $I_A = (p) = (p')$ אז $p \sim p'$. אם נקבע אותו להיות מתוקן אז הוא יחיד (חברות בשדה הפולינומים נבדלת ע"י כפל בפולינום קבוע). לפולינום הנ"ל נקרא הפולינום המינימלי של A הוא m_A. באותו האופן, עבור $T \colon V \to V$ ט"ל ניתן להגדיר את M_T. ■

סימון 6. m_A יהיה הפולינום המינימלי של המטריצה A.

הערה 23. אם $A \in M_n(\mathbb{F})$ ו־$p \in \mathbb{F}[x]$ כך ש־$p(A) = 0$, אז $p \in I_A$ ומתקיים $m_A \mid p$.

הערה 24. אנו יודעים ש־$m_A \mid f_A$ כי ממשפט קיילי המילטון $f_A(A) = m_A(A) = 0$, ו־$f_A \in I_A$ כאשר I_A האידיאל של המאפסים של A. מהיות מרחב הפולינומים תחום ראשי, $m_A \mid f_A$ כדרוש.

דוגמה. עבור $A = I_n$ אז $f_A = (x-1)^n$ ו־$m_a = (x-1)$. לא בהכרח $m_a = f_a$, אל לפעמים כן – לדוגמה בעבור $D \colon \mathbb{F}[x] \to \mathbb{F}[x]$ אופרטור הגזירה מתקיים $f_D = x^{n+1}$ וכן $m_D = x^{n+1}$ כי יש פולינומים שנדרש לכזור n פעמים ע"מ לקבל 0, לדוגמה x^n.

משפט 50. תהא $A = A_f$ המטריצה המצורפת ל־A. אז $f_A = m_A$.

משפט 51. אם A מייצגת את $T \colon V \to V$ אז $m_A = m_T$ (כלומר, הפולינום המינימלי לא תלוי בבחירת בסיס).

הוכחה. נבחר בסיס ל־V, B. יהי $p \in \mathbb{F}[x]$. אז $[p(T)]_B = p([T]_B)$. שני האגפים מתאפסים ביחד, ולכן $I_A = I_T$. ■

הערה 25. נניח ש־A לכסינה והע"ע השונים הם $\lambda_1 \dots \lambda_k$ (כלומר $f_A = \prod_{i=1}^k (x-\lambda_i)^{r^i}$) אז $m_A = \prod_{i=1}^k (x-\lambda_i)$.

הוכחה. בה"כ A אלכסונית, $A = \operatorname{diag}(\lambda_1 \dots \lambda_k)$ עם חזרות. נבחין ש־$\prod_{i=1}^k (x-\lambda_i) = 0$ (הסברים בהמשך). A מייצגת העתקה $T \colon V \to V$ ול־V יש בסיס של ו"עים $B = (v_1 \dots v_n)$. אז $\left(\prod_{i=1}^k (T-\lambda_i)\right)(v_j) = 0$ כי v_j מתאים ל־λ_i כלשהו וכך זה מתאפס. ידוע $m_a \mid \prod_{i=1}^k (x-\lambda_i)$. אם נוריד את אחד המכופלים אז הע"ע שירד לא יתאפס/לא יאפס את הוקטור העצמי המצאים, כלומר כל הגורמים הלינארים דרושים כדי לאפס את T, ומכאן המינימליות והשוויון ל־m_A. ■

הערה 26. אם $A \in M_n(\mathbb{F})$, ו־$\mathbb{F} \subseteq \mathbb{K}$, אז ניתן לחשוב על $A \in M_n(\mathbb{K})$ ו־m_A לא משתנה ללא תלות בשדה.

משפט 52. אם $g, h \in \mathbb{F}[x]$ ו־$T \colon V \to V$ ט"ל אז $g(T), h(T)$ מתחלפות.

הוכחה.

$$\big(g(T) \circ h(T)\big)(v) = (g \cdot h)(T)(v) = (h \cdot g)(T)(v) = \big(h(T) \circ g(T)\big)(v)$$

■

למה 2 (למת המחלק של פולינום מינימלי). יהי m_T הפולינום המינימלי של ט"ל $T \colon V \to V$. אם $f(x) \mid m_T(x)$ וגם $\deg f > 0$ אז $f(T)$ אינו הפיך.

הוכחה. בכלל ש־$m_T \mid f$ אז קיים $g \in \mathbb{F}[x]$ כך ש־$f \cdot g = m_T$. נניח בשלילה ש־$f(T)$ הפיכה. אז:

$$f(T) \circ g(T) = \underbrace{m_T(T)}_{0} \implies \underbrace{f(T)^{-1} \circ (0)}_{0} = g(T)$$

ידוע:

$$\deg m_T = \underbrace{\deg f}_{>0} + \deg g \implies \deg g < \deg m_T$$

בה"כ g מתוקן וקיבלנו סתירה למינימליות של m_T, אלא אם כן $g(x)$ פולינום ה־0 אבל אז $m_T = 0$ בסתירה להגדרתו של פולינום מינימלי. ■

הוכחה זהה עבור מחלק של m_A, עבור A מטריצה.

משפט 53. אם λ ע"ע של T אז בהינתן $p(T) = 0$ מתקיים $p(\lambda) = 0$.

הוכחה. קיים $v \neq 0$ ו"ע כלומר $Tv = \lambda$, ולכן:

$$p(x) = \sum_{i=0}^{n} a_i x^i \quad 0 = 0v = p(T)(v) = \left(\sum_{i=0}^{n} a_i T^i\right)(v) = \sum_{i=0}^{n} a_i T^i(v) = \left(\sum_{i=0}^{n} a_i \lambda^i\right) v = p(\lambda) v$$

מהיות $v \neq 0$ נקבל $p(\lambda) = 0$ כדרוש. ■

"זה טבעוני, זה טבעוני וזה ממשש טבעוני". "מה זה אומר שזה לא טבעוני? יש בזה קצת ביצה".

משפט 54. λ ע"ע של T אמ"מ $m_T(\lambda) = 0$.

הוכחה. כיוון אחד הוא מקרה פרטי של המשפט הקודם. מהכיוון השני, ידוע $m_T(\lambda) = 0$. לפי משפט בזו $(x - \lambda)|m_T(x)$. ידוע $m_T|f_T$ וסה"כ $(x - \lambda)|f_T$ וסה"כ λ ע"ע של T. ■

משפט 55.

$$m_A(x) \mid f_A(x) \mid (m_A(x))^n$$

הוכחה. נותר להוכיח $f_A(x)|(m_A(x))^n$ (השאר ממשפטים קודמים). ידוע שפולינום מינימלי/אופייני נשארים זהים מעל כל שדה שמכיל את $\mathbb{F}$. לכן, ניתן להניח שהוא מתפרק לגורמים לינאריים. ראינו שאם $f, g \in \mathbb{F}[x]$, $\mathbb{K} \subseteq \mathbb{F}$ ומתקיים $g \mid f$ מעל $\mathbb{K}$, אז $g \mid f$ מעל $\mathbb{F}$. אז:

$$\left(\sum n_i = n\right) \qquad f_A = \prod_{i=1}^{k} (x - \lambda_i)^{n_i},\ m_A(x) = \prod_{i=1}^{k} (x - \lambda_i)^{m_i} \quad (1 \leqslant m_i \leqslant n_i)\ (m_a(x))^n = \prod_{i=1}^{k} (x - \lambda_i)^{n|m_i}$$

בגלל ש־$n \leqslant m_i \cdot n \implies 1 \leqslant m_i$ אז מצאנו $f_A \mid m_A^n$. ■

הוכחה זהה עבור $T \colon V \to V$ עם $\dim V = n$.

מסקנה 11 (שימושית!). נניח ש־$f_A \mid g$. נניח ש־g אי פריק. אז $g \mid m_A$.

הוכחה.

$$g \mid f_A \mid (m_A)^n$$

ידוע g אי פריק, ולכן ראשוני (כי $\mathbb{F}[x]$ תחום ראשי) ולכן $g \mid m_A$. ■

משפט 56. נניח ש־A בלוקים עם בלוקים על האלכסון, $A = \operatorname{diag}(A_1 \dots A_k)$ כך ש־$\sum n_i = n$, $\forall i \in [k] \colon A_i \in M_{n_i}(\mathbb{F})$, אז מתקיים $m_A = \operatorname{lcm}(m_{A_1} \dots m_{A_k})$.

במקרה שלנו, ה־lcm הנ"ל הוא הפולינום בעל הדרגה המינימלית שמתחלק בכל ה־$m_{A_i}(x)$־ים. באופן כללי, $\operatorname{lcm}(m_{a_1} \dots m_{a_n})$ מתקבל כיוצר של אידיאל החיתוך בתחום הראשי. כלומר:

$$I = (\operatorname{lcm}(A_1 \dots A_k)) = \bigcap_{i=1}^{n} R m_{a_i}$$

(הבהרת הסימון: $Ra = (a) = \langle a \rangle$).

הוכחה (למשפט לעיל). לכל $g \in \mathbb{F}[x]$ מתקיים:

$$g(A) = \begin{pmatrix} g(A_1) & & \\ & \ddots & \\ & & g(A_k) \end{pmatrix}$$

בבירור מתקיים $g(A) = 0$ אמ"מ $\forall i \in [k] \colon g(A_i) = 0$. לכן $\forall i \in [k] \colon m_{A_i}|g$. מהגדרת ה־lcm סיימנו. ■

מסקנה 12. תהי ט"ל $T \colon V \to V$ ו־V מונ"ס, אז בהינתן $U_1 \dots U_k$ מרחבים T־שמורים כך ש־$V = \bigoplus_{i=1}^{k} U_i$, אזי $m_T = \operatorname{lcm}(\{m_{T|_{U_i}} : i \in [k]\})$.

משפט 57. נניח ש־$T, S \colon V \to V$ ט"לים. אז:

1. אם T, S מתחלפות, אז $\ker S,\ \operatorname{Im} S$ הם T־אינווריאנטים (ולהפך).
2. אם T, S מתחלפות ו־$W \subseteq S$ תמ"ו הוא T־אינוואריאנטי, אז גם $S(W)$ הוא T־איוואריאנטי.

3. אם $W_1, W_2 \subseteq V$ הם T-איווריאנטי אז גם $W_1 \cap W_2,\ W_1 + W_2$ הם T-איווריאנטי.
4. אם $f \in \mathbb{F}[x]$ ו-$W \subseteq V$ תמ"ו T-איווריאנטי, אז W גם $f(T)$-איווריאנטי.

הוכחה.

1. יהא $v \in \operatorname{Im} S$, אז קיים $u \in V$ כך ש-$S(u) = v$:

$$Tv = T(S(u)) = (T \circ S)(u) = (S \circ T)(u) = S(T(u)) \in \operatorname{Im} S \implies Tv \in \operatorname{Im} S$$

ועבור $v \in \ker S$:

$$S(T(v)) = (ST)v = (TS)v = T(S(v)) = T(0) = 0 \implies Tv \in \ker S$$

2. יהי $v \in S(W)$. קיים $w \in W$ כך ש-$v = S(w)$

$$Tv = T(S(w)) = S(T(w)) \in S(W)$$

כי $Tw \in W$.

3. ראינו בתרגול הקודם
4. יהי $w \in W$.

$$f = \sum_{i=1}^{n} a_i x^i,\ \ f(T)w = \left(\sum_{i=0}^{n} a_i T^i\right)(w) = \sum_{i=0}^{n} a_i T^i(w)$$

באינדוקציה $T^i(w) \in W$. W תמ"ו ולכן סגור וסיימנו. ■

1.3.3 ~ ניסוח והוכחת משפט הפירוק הפרימרי

משפט 58 (מקרה הבסיס של משפט הפירוק הפרימרי). ("מאוד חשוב") יהי V מ"ו מעל $\mathbb{F}$. נניח $T \colon V \to V$ ט"ל. נניח $f(T) = 0$. נניח ש-$f = g \cdot h$ עבור $\gcd(g, h) = 1$. אז:

$$V = \ker g(T) \oplus \ker h(T)$$

ואם $f = m_T$, אז g, h הם הפולינומים המינימליים לצמצום T על תת־המרחבים לעיל בהתאמה.

הבהרת הכוונה ב"פולינום המינימלי לצמצום T על תתי המרחבים": בהינתן $T = U \oplus W$, $T_u = T_{|U} \colon U \to U$ ובאופן דומה T_w, אז $m_T = m_{T_U} \cdot m_{T_W}$.

הוכחה.

- ידוע $h = g \cdot h$ ולכן $\exists a(x), b(x) \in \mathbb{F}[x]$ כך ש-$a(x)g(x) + b(x)h(x) = 1$, כך ש-:

$$\underbrace{(a(T) \circ g(T))(v)}_{\in \ker h(T)} + \underbrace{(b(T) \circ h(T))(v)}_{\in \ker g(T)} = V$$

הטענה ש-$(aT \circ gT)v \in \ker hT$ נובעת מכך ש-:

$$(hT)((aT \circ gT)v) = hT((ag(T))v) = (hag)Tv = ((agh)T)v = ((af)T)v = (aT)(fT)v = (a(T) \cdot 0)v = 0v = 0$$

(זאת כי כפל פולינומים קומוטטיבי, כל עולמות הדיון אסוציאטיביים, וכאשר ההעתקה aT תקבל את $fT = 0$ היא תחזיר אפס וסה"כ $0v = 0$ כדרוש). מהכיוון השני:

$$(gT)((bT \circ hT)v) = gT((bh(T))v) = (gbh)Tv = ((bgh)T)v = ((bf)T)v = (bT)(fT)v = (b(T) \cdot 0)v = 0v = 0$$

כלומר אכן $(aT \circ gT) \subseteq \ker hT$ ו-$(bT \circ hT) \subseteq \ker gT$. מהשוויון לעיל סה"כ אכן $V = \ker h(T) + \ker g(T)$. הסכום אכן ישר שכן:

$$\forall v \in \ker gT \cap \ker hT \colon 0 + 0 = (aT \circ gT)v + (bT + hT)v = v$$

דהיינו, $V = \ker h(T) \oplus \ker g(T)$ כדרוש מהחלק הראשון של המשפט.

- עתה נוכיח את החלק השני של המשפט. נניח $f = m_T$, ונסמן:

$$W_1 = \ker g(T) \qquad W_2 = \ker h(T)$$
$$T_1 = T|_{W_1} \qquad T_2 = T|_{W_2}$$

וכן B_1 בסיס ל־W_1, B_2 ל־W_2. לכן $B = B_1 \uplus B_2$ בסיס ל־V. משום שהראינו ש־W_1, W_2 הם T-אינוואריאנטי (כי gT, hT מתחלפות):

$$[T]_B = \begin{pmatrix} [T_1] & 0 \\ 0 & [T_2] \end{pmatrix}$$

מהמשפט שראינו, $m_T = \operatorname{lcm}(m_{T_1}, m_{T_2})$. ברור ש־$m_{T_1}|g$ וגם $m_{T_2}|h$. אז:

$$\deg m_{T_1} = \deg g + \deg h \geqslant \deg m_{T_1} + \deg m_{T_2} = \deg(m_{T_1} \cdot m_{T_2}) \geqslant \deg(\operatorname{lcm}(m_{T_1}, m_{T_2})) = \deg m_T$$

ולכן כל ה"אשים לעיל הדוקים ושוויון בכל מקום.

$$\deg m_{T_1} \leqslant \deg g \wedge \deg m_{T_2} \leqslant \deg h$$

אם אחד מהשוויונות לא הדוקים, אז:

$$\deg m_{T_1} + \deg m_{T_2} < \deg g + \deg h$$

וסתירה למה שהראינו. לכן:

$$(m_{T_1}|g \wedge \deg m_{T_1} = \deg g) \implies m_{T_1} \sim g$$

אבל שניהם מתוקנים ולכן שווים. כנ"ל עבור $m_{T_2} = h$.

סה"כ הוכחנו את כל חלקי המשפט, כדרוש. ■

דוגמה. נסמן $f(x) = x^2(x-1)^3$, $f(T) = 0$. החלק הראשון של המשפט אומר $V = \ker T^2 \oplus \ker(T-I)^3$. החלק השני אומר שאם $f = m_T$ אז x^2 הוא הפולינום המינימלי של $T|_{\ker T^2}$ וכן $(x-1)^3$ המינילי של $T|_{T-I^3}$.

משפט 59 (משפט הפירוק הפרימרי). יהיו $T \colon V \to V$, m_T הפולינום המינימלי של T, ונניח ש־:

$$m_T = g_1 \cdots g_s \quad \forall i \neq j \colon \ \gcd(g_i, g_j) = 1$$

אז:

$$V = \bigoplus_{i=1}^{s} \ker g_i(T)$$

ובנוסף g_i הוא הפולינום המינימלי של $T|_{\ker g_i(T)}$.

"יש לו שם מפוצץ אז הוא כנראה חשוב"

הוכחה. באינדוקציה על s

- **בסיס:** עבור $s = 2$ המשפט שהוכחנו.
- **צעד:** נסמן:

$$h(x) = g_s(x), \ \ g(x) = \prod_{i=1}^{s-1} g_i(x)$$

ואז:

$$\forall i \neq j \colon \ \gcd(g_i, g_j) = 1 \implies \gcd(g, h) = 1$$

מהמשפט שקיבלנו:

$$V = \ker g(T) \oplus \ker h(T) \overset{\text{ה.א.}}{\implies} \bigoplus_{i=1}^{s} \ker g_i(T)$$

וכדי להוכיח את החלק השני של המשפט, נגדיר $m_{T|_{\ker g_i}} = g_i$:

$$\ker h(T) = \bigoplus_{i=1}^{s} \ker g_i(T_{|\ker h(T)}) = \bigoplus_{i=1}^{s} \ker g_i(T)$$

■

הערה 27. בהתאם למקרה הבסיס, מספיק היה להניח $f = g_1 \cdots g_s$, ולא היה באמת צורך להניח $f = m_T$ ספציפית, אם רק רוצים להראות קיום פירוק (ולא צריך להראות ש־g_i הם הפולינומים המינימליים לצמצום T על התמ"וים). למעשה נשתמש בגרסה מוחלשת זו של משפט הפירוק הפרימרי.

משפט 60 (תוצאה 1 ממשפט הפירוק הפרימרי). T לכסינה אמ"מ $m_T = \prod_{i=1}^{s}(x-\lambda_i)$ מתפרק לגורמים לינארים $i \neq j \implies \lambda_i \neq \lambda_j$ שונים זה מזה.

הוכחה.

$\implies$ לפי המשפט, אם נסמן $g_i = (x-\lambda_i)$:

$$V = \bigoplus_{i=1}^{s} \ker(T-\lambda_i I)$$

כלומר V סכום ישר של המ"ע של $\lambda_1 \ldots \lambda_s$. לכל מרחב עצמי מממד k_i קיימים $v_1 \ldots v_{k_i}$ בסיס כך ש־$Tv_j = \lambda_i v_j$ ($j \in [k_i]$), ומהסכום הישר ידוע $\sum_{i=1}^{s} k_i = n$ ומהיות איחוד בסיסים של מ"ע גם בסיס (כי המ"ע זרים) מצאנו בסיס מלכסן הוא אוסף הבסיסים של המ"עים.

$\impliedby$ אם T לכסינה, אז הפולינום המינימלי הוא ה־lcm של הפולינומים המינימליים של הבלוקים על האלכסון. הבלוקים על האלכסון הם λ_i הע"ע מגודל 1, ולכן ה־lcm שלהם הוא מכפלת $x-\lambda_i$ כאשר λ_i הע"עים השונים, וסה"כ m_T מכפלת גורמים לינאריים שונים.

■

משפט 61 (תוצאה 2 ממשפט הפירוק הפרימרי). נניח $T\colon V \to V$ לכסינה, וקיים $W \subseteq V$ תמ"ו T־שמור. אז $T|_W$ לכסינה.

הוכחה. נסמן $S = T|_W$. אנחנו יודעים $m_T(T) = 0$ ולכן $m_T(S) = 0$. ידוע $m_T = \prod_{i=1}^{r}(x-\lambda_i)$ ולכן $m_S \mid m_T$ מתפרק לגורמים לינארים זרים, סה"כ S לכסינה. ■

סיכום. $T\colon V \to V$ ט"ל, ו־:

$$\forall i \neq j\colon\ \gcd(g_i, g_j) = 1 \wedge m_T(x) = \prod_{i=1}^{s} g_i$$

אז:

$$V = \bigoplus_{i=1}^{s} \ker(g_i(T)) \wedge \forall i\colon\ m_{T|_{\ker g_i(T)}} = g_i$$

המשך בעמוד הבא

1.4 Jordan Form

1.4.1 ~ מציאת שורשי פולינום אופייני ממעלה חמישית ואילך

נבחין בבעיה: $A = M_5(\mathbb{Z})$, קבעו אם היא לכסינה מעל $\mathbb{C}$.

- נחשב את $f_A(x)$
- נמצא שורשים, אלו הם הע"ע
- לכל ע"ע נחשב את v_λ
- אם סכום הממדים מסעיף ג' הוא 5, אז היא לכסינה
- T לכסינה אמ"מ קיים בסיס ו"ע אמ"מ ריבוי גיאומטרי = ריבוי אלגברי

אבל (המתמטיקאי, לא מילת הניגוד ולא מילה נרדפת ליגון) הוכיח שאין פתרונות לפולינומים ממעלה חמישית ויותר, וגלואה מצא דוגמאות לפולינומים שאי אפשר לבצע עליהם נוסחאת שורשים ופיתח את תורה להרחבת שדות לשם כך (תורת גלואה).

היוונים העתיקים ביססו את הגיאומטריה שלהם באמצעות דברים שאפשר לבדוק עם שדה ומחוגה. באמצעות כלים של תורת גלואה אפשר לראות מה אפשר לעשות עם האמצעים האלו, ולהוכיח שלוש בעיות שהיוונים לא הצליחו לפתור – האם אפשר לרבע את המעגל (האם אפשר לבנות באמצעות שדה ומחוגה ריבוע ששטחו שווה לשטח המעגל), או במילים אחרות, האם אפשר למצוא את $\sqrt{\pi}$ – אי אפשר כי זה לא מספר אלגברי. שאלות אחרות היו, בהינתן קובייה, האם אני יכול למצוא קובייה בנפח כפול? באותה המידה אי אפשר למצוא את $\sqrt{3}$. שאלה אחרת הייתה האם אפשר לחלק זווית ל־3 חלקים שווים.

גלואה הראה שכדי לעשות את זה צריך למצוא שורשים שלישיים של כל מני דברים, ושבאמצעות סרגל ומחוגה אי אפשר לעשות זאת. בעיות שהיו פתוחות לעולם המתמטי במשך אלפי שנים נפתרו בעזרת אותן התורות.

בגלל שאין אלגוריתם למציאת פולינום ממעלה חמישית ואילך, ננסה לפתח כלים נוספים שיעזרו לנו למצוא שורשים לפולינומים הללו במקרים פרטיים.

אבל ניאלץ להאביל את משפחתו עליו כשמת משחפת בגיל 26. גלואה מת בגיל 21 מדו־קרב.
מסקנה 13 (מסקנת הבדיעבד של גלואה)**.** לא ללכת לדו־קרב.

הוכחה. ההוכחה מתקדמת ועוסקת בתורת גלואה. ∎

הגדרה 46. בהינתן $\forall i \neq j: \lambda_i \neq \lambda_j$ $f(x) = \prod_k (x-\lambda_k)^{r_k}$. אז $f^{\text{red}} := \prod_k (x-\lambda_k)$
משפט 62.

$$f^{\text{red}} = \frac{f}{\gcd(f, f')}$$

הוכחה. נשאר כתרגיל בעבור הקורא. (נתנו לנו את זה בשיעורי הבית) ∎

משפט 63. A לכסינה אמ"מ $f_A^{\text{red}}(A) = 0$
למה 3. $f_A^{\text{red}} \mid m_A$ ושוויון אמ"מ A לכסינה.

הוכחת הלמה. יהיו $\lambda_1 \dots \lambda_r$ הע"ע של A (אפשר בה"כ להרחיב שדה כדי שהם יהיו קיימים). אז אם $f_A(x) = \prod_{i=1}^r (x-\lambda_i)^{s_i}$ ומתקיים $m_A(x) = \prod_{i=1}^r (x-\lambda_i)^{r_i}$ וידוע $1 \leqslant r_i \leqslant s_i$ ולכן $f_A^{\text{red}} \mid m_A$.

עתה נוכיח את החלק השני של הלמה (השוויון). אם A לכסינה, אז יש בסיס של ו"עים ועם λ הוא ע"ע של ו"ע בבסיס B אז $Av_\lambda - \lambda v_\lambda = 0$, ולכן $f_A^{\text{red}}(A) = 0$ וסה"כ $f_A \mid m_A$.

אם $f_A^{\text{red}} = m_A$ אז m_A מכפלה של גורמים לינארית זרים, וראינו גרירה ללכסינות. ∎

הוכחת המשפט באמצעות הלמה. A לכסינה אמ"מ $m_A = f_A^{\text{red}}$, ואנחנו יודעים כי $m_A(A) = 0$ ולכן A לכסינה אמ"מ $f_A^{\text{red}}(A) = 0$. ∎

משום ש־f_A^{red} כולל את כל הגורמים הלינאריים של f_A, עבור $\deg f_A > 4$ נוכל למצוא את f^{red} (באמצעות משפט 62, אלגו' אויקלדס, וחלוקת פולינומים) ולקוות שהוא ממעלה קטנה יותר, ואז נפרק גורמים לינארים ל־f^{red} במקום, ואז כבר יהיה קל למצוא את הריבוי כי נוכל להוציא מ־f_A גורמים לינאריים כגורם משותף החוצה.

1.4.2 ~ צורת ג'ורדן לאופרטור לינארי נילפוטנטי

(1.4.2.1) נילפוטנטיות

מטרה: בהינתן $T\colon V\to V$ נרצה לפרק את V לסכומים ישרים של מרחבים T־אינווריאנטים, קטנים ככל האפשר.

הגדרה 47. יהי $T\colon V\to V$ ט"ל. נאמר ש־V פריק ל־T אם קיימים $U,W\subseteq V$ תמ"וים כך ש:

$$V=U\oplus W\quad\wedge\quad\dim U,\dim W>0\quad\wedge\quad U,W\ \text{are}\ T\text{-invariant}$$

מעתה ואילך (עד סוף הנושא)**,** נניח ש־$f_T(x)$ מתפצל מעל f לגורמים לינארים (כלומר, נרחיב לשדה סגור אלגברית).

הגדרה 48. $T\colon V\to V$ ט"ל. T נקראת העתקה נילפוטנטית אם קיים $n\in\mathbb{N}$ כך ש־$T^n=0$. באופן דומה A תקרא מטריצה נילפוטנטית אם $\exists n\in\mathbb{N}\colon A^n=0$.

הגדרה 49. עבור n המינימלי שעבורו $T^n=0/A^n=0$, אז n הנ"ל נקרא דרגת הנילפוטנטיות של T/A, ומסמנים $n(T)/n(A)$.

"נילפוטנטית בא מלשון null. הרעיון: דבר מה שמתבטל.

משפט 64 (תוצאה 3 ממשפט הפירוק הפרימרי)**.** בהינתן V אי־פריק ביחס ל־T, ובהנחה ש־$f_T(x)$ מתפצל לגורמים לינאריים, אז $m_T(x)=(x-\lambda)^r$. נוסף על כך $T-\lambda I$ נילפוטנטית ו־$n(T-\lambda I)=r$.

הוכחה. נפרק למקרים.

- אם $m_T(x)$ לא מתפרק, הוא בהכרח לא קבוע אחרת $m_T(T)\neq 0$ וסתירה, לכן $m_T(x)=(x-\lambda)$ לינארי כלשהו (אם לא לינארי ניתן לפרק לגורמים לינארים ואז m_T מתפרק וסתירה).
- אם $m_T(x)$ מתפרק, אז נוציא גורם לינארי אחד ונקבל $m_T=g_1\cdots g_i$ כאשר g_i לינאריים, דהיינו ממשפט הפירוק הפרימרי, נניח בשלילה $g_i\neq g_j$ ומהיות m_T מתוקן נקבל $\gcd(g_i,g_j)=1$ כלומר $V=\ker\bigoplus_{i=1}^r\ker g_i(T)$ ולכן V פריק וסתירה. דהיינו $g_i=g_j$ וסה"כ $m_T(x)$ הוא מהצורה $m_T(x)=g_i^r=(x-\lambda)^r$ כדרוש.

עתה ניגש להוכיח את החלק השני של ההוכחה (ש־$T-\lambda I$ ניל' מדרגה r). משום ש־$m_T(x)=(x-\lambda)^r$, אזי $0=M_T(T)=(T-\lambda I)^r$ ולכן $n(T-\lambda I)\leqslant r$, ומהמינימליות של m_T נסיק $n(T-\lambda I)=r$ כדרוש. ∎

נסמן $S=T-\lambda I$ בהקשר לעיל. עוד כדאי להבחין ש־V הוא S־איווריאנטי (**אך לא בהכרח אי־פריק ביחס ל־S**) שכן $S(V)=T(V)-\lambda V\in V$ מסגירות לכפל בסקלר λ ולחיבור נגדי.

מה למדנו? שמשום שאנו יכולים לפרק (ממשפט הפירוק הפרימרי) את T למרחבים T־איווריאנטיים פריקים מינימליים, אז לכל U_i כזה נוכל להגדיר $S_i=T-\lambda_i I$ כזו כך שהיא נילפוטנטית. אם נוכל להבין טוב מה S_i עושה למרחב שהיא שמורה עליו, נוכל להבין באופן כללי מה ההעתקה T עושה לכל אחד מהמרחבים אליהם פריקנו אותה.

למה 4. תהי T העתקה כללית, אז אם $\ker T^i=\ker T^{i+1}$ לכל $j\geqslant i$ מתקיים $\ker T^i=\ker T^j$.

למה 5. תהי T העתקה כללית, אז $\forall i>j\colon\ \ker T^i\supseteq\ker T^j\wedge\operatorname{Im}T^i\subseteq\operatorname{Im}T^j$.

משפט 65. תהי $T\colon V\to V$ העתקה מעל מ"ונסים, $\dim V=n$, אז קיים $\mathcal{F}(T)\in[n]$ כך ש־ $\forall i\in\mathbb{N}\colon\ \ker T^{\mathcal{F}(T)}=\ker T^{\mathcal{F}(T)+i}\wedge\operatorname{Im}T^{\mathcal{F}(T)}=\operatorname{Im}T^{\mathcal{F}(T)+i}$.

הוכחה. מלמה 5, בהכרח:

$$\ker T\subseteq\ker T^2\subseteq\ker T^3\subseteq\cdots\subseteq T^i\subseteq\cdots\subseteq V$$

נניח בשלילה שכל ההכלות עד $i=n$ חלשות, ממשפט נסיק:

$$\dim\ker T<\dim\ker T^2<\cdots<\dim\ker T^i\leqslant n$$

כלומר יש n מספרים טבעיים שונים בין $\ker T$ ובין n (לא כולל) ולכן $\dim\ker T<0$ וסתירה. דהיינו קיים $\mathcal{F}(T)$ כך ש־$\ker T^{\mathcal{F}(T)}=\ker T^{\mathcal{F}(T)+1}$ ומלמה 4 נקבל ש־$\ker T^{\mathcal{F}(T)}=\ker T^i\colon\forall i\geqslant\mathcal{F}(T)$. ניכר ש־$\operatorname{Im}T^i\supseteq\operatorname{Im}T^{\mathcal{F}(T)}$ לכל $i\geqslant\mathcal{F}(T)$, וממשפט הממדים $\dim\operatorname{Im}T^{\mathcal{F}(T)}=\dim\operatorname{Im}T^i$ ולכן $\forall i\geqslant\mathcal{F}(T)\colon\ \operatorname{Im}T^{\mathcal{F}(T)}=\operatorname{Im}T^i$ כדרוש. ∎

משפט 66. בהינתן T העתקה נילפוטנטית, אז $\mathcal{F}(T)=n(T)$.

סימון 7. $\mathcal{F}(T)$ לעיל סימון (שמקובל אך ורק בסיכום הזה), וקרוי ה־"fitting index" של T.

(1.4.2.2) שרשאות וציקליות

הגדרה 50. קבוצה מהצורה $\{v,Tv\cdots T^kv\}$ כאשר $T^{k+1}v=0$ והוא המינימלי, נקראת שרשרת.

משפט 67. $T\colon V\to V$ נילפוטנטית, אז כל שרשרת היא בת"ל.

הוכחה. יהיו $\alpha_0 \dots \alpha_k \in \mathbb{F}$ כך ש־$\sum_{i=0}^{k} \alpha_i T^{(i)}(v) = 0$. נניח בשלילה שהצירוף אינו טרוויאלי. אז קיים j מינימלי שעבורו $\alpha_j \neq 0$. נניח n המקסימלי ש־T^n לא מאפס את v. אז:

$$T^{n-j}\left(\sum \alpha_i T^{(i)}(v)\right) = T^{n-j}\left(\sum_{i=j}^{k} \alpha_i T^i(v)\right) = \alpha_j T^{n-1}(v) + 0 = 0$$

אבל $\alpha_j, T^{n-1} \neq 0$ וזו סתירה. ■

תזכורת. תמ"ו שקיים לו בסיס שהוא שרשרת, נקרא ציקלי.

אנטי־דוגמה: ישנם מ"וים שאינם T־ציקליים. למשל:

$$V = \left\{ f\colon \mathbb{R}^2 \to \mathbb{R} \mid f\begin{pmatrix} x \\ y \end{pmatrix} - p(x) + h(y) \mid \leqslant n \text{ פולינומים ממעלה } p, h \right\}$$

ו־T אופרטור הגזירה הפורמלית. כדי ש־V יהיה ציקלי, צריך למצוא בסיס ציקלי שממדו הוא לכל היותר דרגת הנילפוטנטיות. נבחין ש־$n(T) = n+1$, וידוע ש־$\dim V = 2n+1$, ולכן שרשרת מקסימלית באורך $n+1$ ולכן לא יכול להיות בסיס שרשרת. לכן V אינו T־ציקלי.

הערה 28. יהי $T\colon V \to V$ נילפוטנטית ו־$\dim V = n$ אז $n(T) \leqslant n$ וישנו שוויון אמ"מ V ציקלי.

הערה 29. אם $T\colon V \to V$ נילפוטנטית ו־V ציקלי אז V אי־פריק ל־T.

הוכחה. נניח בשלילה שישנו פירוק לא טרוויאלי של V ל־T. אז $V = U \oplus W$ לא טרוויאלים. נסמן $\dim U = k, \dim W = \ell$ וידוע $k, \ell < n$. בה"כ $k \geqslant \ell$. נסמן $B_v = (v, Tv \dots T^{n-1}v)$. קיימים (ויחידים) $u \in U, w \in W$ כך ש־$v = u + w$. אז:

$$T^k v = T^k u + T^k w$$

אך משום ש־T נילפוטנטית אז $T|_U, T|_W$ נילפוטנטית גם כן. ידוע $n(T|_U), n(T|_W) \leqslant k$ ולכן בפרט $T^k(u) = T^k(w) = 0$ ולכן $T^k v = 0$ אבל $T^k(v) \in B_v$ $\wedge$ $k < n$. ■

משפט 68. תהי $T\colon V \to V$ נילפוטנטית ונניח U תמ"ו של V הוא T־אינוואריאנטי וציקלי, אז עבור $S =: T_U$:

1. $\dim U \leqslant n(T)$
2. $\operatorname{Im}(T_U) = T(U)$ ציקלי ו־$\dim T(U) = \dim U - 1$

הוכחה.

1. $n(T) \geqslant n(T|_U)$ וגם $\dim U = n(T_W)$
2. $T(u) = T(\operatorname{span}(v, \dots T^k v)) = \operatorname{span}(Tv \dots T(T^k v))$ ז"א $T(U) = \operatorname{span}(Tv \dots T^k v)$, זו קבוצה בת"ל ופורש את $T(U)$ ולכן $\dim T(U) = \dim U - 1$. ■

הגדרה 51. $U \subseteq V$ תמ"ו ציקלי ייקרא ציקלי מקסימלי אם $\dim U = n(T)$.

משפט 69. לכל V מ"ו, $T\colon V \to V$ נילפוטנטית קיים תמ"ו ציקלי מקסימלי.

הוכחה. קיים $v \in V$ כך ש־$T^{n(T)-1}v \neq 0$. אז $v \neq 0$ ו־$v, Tv, \dots T^{n(T)-1}$ ומטעה מקודם בת"ל ולכן $\operatorname{span}(v \dots T^{n(T)-1})$ תמ"ו ציקלי מקסימלי. ■

משפט 70. נניח $U \subseteq V$ תמ"ו ציקלי מקסימלי. אזי:

1. אם $T(U) \subseteq T(V)$ הוא גם ציקלי מקסימלי.
2. $U \cap T(V) = T(U)$

הוכחה.

1. U – ציקלי. לכן $\dim T(U) = \dim U - 1$. טענה:

$$\dim T(U) = n\left(T|_{T(V)}\right) = n(T) - 1$$

וסיימנו.

2. ידוע $T(U) \subseteq U$ כי U ציקלי ולכן שמור, וכן $U \subseteq V$ והסקנו $T(U) \subseteq T(V)$, לכן $T(U) \subseteq U \cap T(V)$

עתה נוכיח שוויון באמצעות שיקולי ממד. אם לא היה שוויון אז:

$$T(U) \subsetneq U \cap T(V) \subseteq U \implies \dim T(U) < \dim(U \cap T(V)) \leqslant \dim(T(U)) + 1 \implies U \cap T(V) = U$$

זו סתירה כי:

$$U \cap T(V) \subseteq T(V) \implies n(T|_{T(v)}) = n(T) - 1 \implies n(T) = \dim U \leqslant n(T) - 1$$

■

(1.4.2.3) ניסוח צורת מייקל ג'ורדן לאופרטור נילפוטנטי

משפט 71 (המשלים הישר לתמ"ו ציקלי מקסימלי)**.** נניח $T\colon V \to V$ ט"ל לינארית נילפוטנטית, $U \subseteq V$ תמ"ו ציקלי מקסימלי אז קיים $W \subseteq V$ תמ"ו T־איווריאנטי כך ש־$V = U \oplus W$.

הוכחה. נוכיח באינדוקציה על $n = n(T)$.

בסיס: אם $n(T) = 1$ אז $T = 0$ אז כל $W \subseteq V$ הוא T־איווריאנטי. והיות שכל קבוצה בת"ל ניתנת להשלמה לבסיס, אז $U = \operatorname{span}(v)$ אז $W = \operatorname{span}(v_2 \dots v_m)$ כאשר $B_V = (v := v_1 \dots v_m)$.

צעד: ("צעד, מעבר, אותו דבר, תקראו לזה איך שבא לכם") נניח שאנו יודעים את נכונות הטענה עבור $n = n(T) - 1$. נוכיח עבור $n = n(T)$. נצמצם את T ל־$T|_{T(V)}$. ידוע $T(U) \subseteq T(V)$ ציקלי מקסימלי. לכן, לפי ה.א. קיים W_1 הוא T־איווריאנטי כך ש־$T(V) = T(U) \oplus W_1$.

נגדיר $W_2 = \{v \in V \mid Tv \in W_1\}$. אז
למה 6. ("למה א") $U + W_2 = V$ (לאו דווקא סכום ישר) וגם $U \cap W_1 = \{0\}$.
למה 7. ("למה ב") בהינתן $W_1 \subseteq W_2$ ו־$V \subseteq U$ תמ"ו כך ש־$U + W_2 = V$ וגם $U \cap W_1 = \{0\}$ אז קיים $W' \subseteq V$ כך ש־$W_1 \subseteq W' \subseteq W_2$ וגם $U \oplus W' = V$

נניח שהוכחנו את הלמות. יהי $w \in W_1$ אז $T(w) \in W_1$ ולכן $w \in W_2$ ולכן $W_1 \subseteq W_2$. אז מצאנו W' תמ"ו של V כך ש־$W_1 \subseteq W' \subseteq W_2$. יהי $w \in W'$ בפרט $w \in W_2$ ולכן $T(w) = W_1 \subseteq W'$. ■

הוכחת למה ב' היא תרגיל רגיל בלינארית 1א שאין ערך להביא את הוכחתו.

הוכחת למה א' גם היא לא מעניינת במיוחד, אבל אותה המרצה כן הוכיח:

הוכחה. יהי $v \in V$, נביט ב־$T(v)$. קיימים $u \in U, w_1 \in W_1$ כך ש־:

$$Tv = Tu + w_1 \implies Tv - Tu = w_1 \implies T(v-u) = w_1 \in W_1$$

ידוע $v = v - u + u$. לכן $T(v-u) \in W_1 \implies v - u \in W_2$.

אזי משהו $V = U + W_2$ ו־$W_1 \subseteq T(V)$ ולכן:

$$U \cap W_1 = U \cap (T(V) \cap W_1)$$

ידוע ש־:

$$U \cap W_1 = (U \cap T(V)) \cap W_1 \implies U \cap T(V) = T(U) \implies T(V) = T(U) \oplus W_1 \implies U \cap W_1 = T(U) \cap W_1 = \{0\}$$

■

מסקנה 14. $T\colon V \to V$ ט"ל נילפוטנטית אז V אי־פריק ל־T אמ"מ V ציקלי.

הוכחה.

$\implies$ זהו משפט שכבר הוכחנו

$\impliedby$ נניח V אי־פריק. אז קיים $U \subseteq V$ תמ"ו ציקלי מקסימלי. לפי המשפט קיים $W \subseteq V$ תמ"ו T־איווריאנטי כך ש־$T = U \oplus W$. ידוע U, W תמ"וים איווריאנטי. אם $U = \{0\}$ אז $V = 0$ ובפרט ציקלי. אחרת, מאי־פריקות V ל־T, נסיק ש־$W = \{0\}$ ולכן $V = U$ ציקלי.

■

משפט 72 (משפט ג'ורדן בעבור T נילפוטנטית 1). תהי $T\colon V\to V$ נילפוטנטית אז קיים פירוק של V לסכום ישר של $V=\bigoplus U_i$ כאשר U_i הם T־ציקליים.

הוכחה. באינדוקציה על המשפט הקודם: נמצא ב־V ציקלי מקסימלי כלשהו. אז קיים $W\subseteq V$ תמ"ו T־שמור כך ש־$V=U_1\oplus W$. ידוע $T|_W\colon W\to W$ נילפוטנטית, וכעת באינדוקציה שלמה על $\dim V$. ■

משפט 73 (משפט ג'ורדן בעבור ט"ל נילפוטנטית 2). עבור $T\colon V\to V$ נילפוטנטית, קיים בסיס B של V שהוא איחוד של שרשראות.

מסקנה 15. בעבור B בסיס מג'רדן, נסיק:

$$[T]_B=\begin{pmatrix}\square & 0 & \dots & 0\\ 0 & \square & \ddots & \vdots\\ \vdots & \ddots & \ddots & 0\\ 0 & \dots & 0 & \square\end{pmatrix}$$

כאשר הבלוקים הם בלוקי ג'ורדן אלמנטרי. בלוק ג'ורדן מהצורה הבאה יקרא אלמנטרי ויסומן:

$$J_n(0)=\begin{pmatrix}| & & |\\ T(v) & \cdots & T(T^kv)\\ | & & |\end{pmatrix}=\begin{pmatrix}0 & 0 & \cdots & 0\\ 1 & \ddots & \ddots & \vdots\\ \vdots & \ddots & 0 & 0\\ 0 & 0 & 1 & 0\end{pmatrix}$$

(יש ספרים בהם זה ה־transpose של זה).

משפט 74 (יחידות צורת ג'ורדן בעבור ט"ל נילפוטנטית). עבור $T\colon V\to V$ נילפוטנטית, אז בכל הפירוקים של $V=\bigoplus U_i$ עבור U_i ציקליים (אי־פריקים) אז מספר תתי־המרחב ממימד נתון הוא זהה עבור כל פירוק.

הוכחה. באינדוקציה על $n=n(T)$.

- עבור $n=1$, העתקת ה־0. V מתפרק לסכום ישר של מרחבים ממימד 1.
- צעד, נניח נכונות עבור $n\in\mathbb{N}$. נניח ש־$n(T)=n+1$. נסמן פירוק:

$$V=\bigoplus_{i=1}^{k}U_i=\bigoplus_{i=1}^{\ell}W_i$$

נסדר את $(u_i)_{i=1}^k$ לפי גודל מימד, ונניח שרשימת הגדלים היא:

$$(\underbrace{1,1,\dots 1}_{\times s}<a_1\leqslant\dots\leqslant a_p)\implies s+p:=k$$

רשימת הממדים מגודל 1 ועוד כל השאר. נעשה כנ"ל עבור $(w_i)_{i=1}^{\ell}$ ונקבל:

$$(\underbrace{1,1,\dots 1}_{\times t}<b_1\leqslant\dots\leqslant b_r)\implies t+r:=\ell$$

ידוע:

$$T(v)=\bigoplus_{i=1}^{k}T(U_i)=\bigoplus_{i=1}^{k}T(W_i),\quad n\left(T|_{T(v)}\right)=n,\quad p=r,\quad \forall i\colon a_i-1=b_i-1\implies a_i=b_i$$

(הפירוק ל־s ו־t דרוש כדי שהפירוק לעיל לא יכלול אפסים כאשר מפעילים את T) ידוע $a_i-1=b_i-1$ כי אינדקס הנילפוטנטיות קטן ב־1 בהחלת T)

משפט הממדים השני אומר ש־:

$$a_i=\dim U_i=\dim\ker T|_{U_i}+\underbrace{\dim\operatorname{Im}T|_{U_i}}_{a_i-1}\implies\dim\ker T|_{U_i}=1$$

מהטענה השנייה בלמה:

$$\ker T = \bigoplus_{i=1}^{k} \ker T|_{U_i} \implies \dim\ker T = \sum_{i=1}^{k} \dim\ker T|_{U_i} = k \implies k = \ell \implies s = t$$
$$= \sum_{i=1}^{k} \dim\ker T_{W_i} = \ell$$

■

במילים אחרות: כל מטריצה מייצגת של ט"ל נילפוטנטית דומה למטריצת ג'ורדן יחידה, עד כדי שינוי סדר הבלוקים.

(למה זה שקול? כי הגודל של בלוק הוא הממד של התמ"ו שנפרש ע"י וקטורי הבסיס שמתאימים לעמודות הללו)

למעשה, בכך הבנו לחלוטין כיצד העתקות נילפוטנטיות מתנהגות. עשינו רדוקציה למקרה הפרטי של נילפוטננטית, ועתה ננסה להבין את המקרה הכללי. ניעזר בתוצאה 3 ממשפט הפירוק הפרימרי לשם כך.

למה 8. נניח $V = \bigoplus_{i=1}^{k} U_i$ כאשר U_i הוא T-איווריאנטי (אין צורך להניח נילפוטנטיות), אז:

א. $$T(V) = \bigoplus_{i=1}^{k} T(U_i)$$

ב. $$\ker T = \bigoplus_{i=1}^{k} \ker T(U_i)$$

הוכחה: נותר כתרגיל בעבור הקורא.

1.4.3 ~ צורת ג'ורדן לאופרטור ליניארי כללי

הגדרה 52. בלוק ג'ורדן אלמנטרי עם ערך λ הוא בלוק מהצורה:

$$J_n(\lambda) = J_n(0) + \lambda I_n = J_n(0) = \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \lambda & 0 \\ 0 & 0 & 1 & \lambda \end{pmatrix}$$

הגדרה 53. בהינתן $T\colon V \to V$, בסיס B נקרא בסיס מג'רדן אם $[T]_B$ היא מטריצה עם בלוקי ג'ורדן מינימליים על האלכסון.

משפט 75 (משפט ג'ורדן). לכל העתקה $T\colon V \to V$ כאשר V מונ"ס מעל שדה סגור אלגברית $\mathbb{K}$, קיים בסיס מג'רדן.

מה עומד לקרות?

1. נפרק את המרחב V לתתי־מרחבים, שכל אחד מהם משוייך לערך עצמי λ_i. נעשה זאת בשתי גישות – הראשונה באמצעות משפט הפירוק הפרימרי, והשנייה באמצעות פירוק למרחבים עצמיים מוכללים (שני הפירוקים מניבים את אותם המרחבים).
2. נתבונן על המרחבים האלו, ונסיק שיש העתקה ציקלית עליהם, שאנחנו כבר מכירים את צורת הג'ורדן שלה. היא תאפשר לנו לפרק את המרחבים שקיבלנו לתתי־מרחבים ציקליים, עם בסיס שרשרת שנותן לנו צורת ג'ורדן.

(1.4.3.1) בעזרת פירוק פרימרי

ראשית כל, נוכיח את משפט ג'ורדן באמצעות משפט הפירוק הפרימרי שכבר ראינו.

הוכחה באמצעות פירוק פרימרי. נניח ש־$f_T(x)$ מתפצל לחלוטין. מהגרסה החלשה של משפר הפירוק הפרימרי (ראה הערה תחתיו), ממשפט קיילי־המילטון f_T מאפס את T, ותחת הסימון $f_T(x) =: \prod_{i=1}^{s}(x-\lambda_i)^{d_\lambda}$ מתקיים:

$$V = \bigoplus_{i=1}^{s} \underbrace{\ker((T-\lambda_i)^{d_\lambda})}_{U_i}$$

כאשר $U_1 \dots U_n$ הם T-איווארינאטים. משום ש־U_i האי־פריקים ביחס ל־T, ו־T שמורים. היות שהם אי פריקים $f_{T|_{U_i}} = (x-\lambda)^n$. נגדיר $S = T - \lambda I$. אז U_i הוא T-איווריאנטי אמ"מ הוא S-איווריאנטי (טענה שראינו בעבר). ראינו ש־$S|_{U_i}$ היא נילפוטנטית שכן ממשפט הפירוק $(T-\lambda_i)^{r_i}$ מאפס את $T|_{U_i}$ (אך לא בהכרח מינימלי, שכן f_T לא בהכרח מינימלי) ולכן

$S|_{U_i}$,$(T-\lambda_i)^{r_i}=S^{r_i}=f|_{U_i}(T)=0$ כלומר $S|_{U_i}$ נילפוטנטית. ל־$S|_{U_i}$ הוכחנו קיום צורת ג'ורדן, משמע קיים לה בסיס מג'רדן $\mathcal{B}_i$ כך ש־:

$$[S|_{U_i}]_{\mathcal{B}_i}=[T|_{U_i}-\lambda I_V]=[T|_{U_i}]_{\mathcal{B}_i}-\lambda I\implies [T|_{U_i}]_{B_i}=\operatorname{diag}(J_{a_1}(0)\dots J_{a_n}(0))+\lambda_i I=\operatorname{diag}(J_{\lambda_i}(0)\dots J_{a_n}(\lambda_i))$$

לכן, נוכל לשרשר את הבלוקים הללו ולקבל $\mathcal{B}=\bigcup_{i=1}^{s}\mathcal{B}_i$, המקיים:

$$[T|_{U_i}]_{\mathcal{B}}=\operatorname{diag}\left\{[T|_{U_1}]_{\mathcal{B}_1}\dots[T|_{U_s}]_{\mathcal{B}_s}\right\}$$

משום שכל אחד מ־$[T|_{U_i}]_{B_i}$ הוא בלוק ג'ורדן בעצמו, סה"כ נקבל:

$$[T]_B=\operatorname{diag}(J_1(\lambda_1)\dots J_n(\lambda_j))$$

זוהי צורת הג'ורדן של מטריצה כללית. ■

במילים אחרות – נעזרנו בפירוק פרימרי ע"מ לפרק את המרחב למרחבים T־איווראינטים פריקים מינימליים (בהמשך נראה שאלו המרחבים העצמיים המוכללים של T, שמקיימים כל מיני תכונות נחמדות) ואת המרחבים אליהם פירקנו, ניתחנו בעזרת צורת ג'ורדן להעתקות נילפוטנטיות.

משפט 76. צורת ג'ורדן היא יחידה עד לכדי סדר בלוקים.

(1.4.3.2) בעזרת מרחביים עצמיים מוכללים

בגישה הזו נוכל לפתח את צורת ג'ורדן למטריצה כללית ללא צורך בפירוק פרימרי, פולינום מינימלי, משפט קיילי־המילטון וכו'. זו גישה יותר אלמנטרית ופשוטה, ואם מבינים אותה האלגוריתם המסורבל למציאת צורת ג'ורדן הופך לאינטואיטיבי בהרבה.

הגדרה 54. המרחב העצמי המוכלל של λ הוא מ"ו:

$$\tilde{\mathcal{V}}_\lambda:=\bar{\mathcal{V}}_\lambda:=\{v\in V\mid \exists n\in\mathbb{N}\colon (T-\lambda I)^n v=0\}$$

משפט 77. המרחב העצמי המוכלל הוא מ"ו.

מסקנה 16. באופן מידי נסיק $V_\lambda\subseteq\tilde{\mathcal{V}}_\lambda$.

הגדרה 55. וקטור עצמי מוכלל הוא וקטור $v\in V$ כך ש־$\exists i\in[n]\colon T^{(i)}v=\lambda v$.

הערה 30. החלק הזה ואילך, אז סוף הפרק, הינו הרחבה שלי בלבד ואילו אינם משפטים המופיעים בקורס. עם זאת, המשפטים להלן מאפשרים להבין בצורה הרבה יותר טובה את צורת ג'ורדן, ולעיתים קרובות תצטרכו להוכיח אותם בעצמכם.

הערה 31. מרגישים אבודים? אני ממש ממליץ על הסדרה הבאה (פרקים 36-42) (סליחה למי שהמליץ לי על זה במקור, אני לא זוכר מי זה היה אז אני לא אוכל לתת קרדיט).

משפט 78. תהי העתקה T כללית ו־$\lambda\in\mathbb{F}$ סקלר, אז $\tilde{\mathcal{V}}_\lambda=\mathcal{N}(T-\lambda I)^{\dim V}$ (כאשר $\mathcal{N}$ המרחב המאפס/הקרנל של המטריצה)

הוכחה. נוכיח באמצעות הכלה דו כיוונית. הכיוון $\mathcal{N}(T-\lambda I)^{\dim V}\subseteq\tilde{\mathcal{V}}_\lambda$ טרוויאלי. יהי $v\in\mathcal{N}(T-\lambda I)^j$, אם $j<\dim V$ אז $(T-\lambda)^j v=0$ ולכן $(T-\lambda I)^{\dim V}v=0$ וסה"כ $v\in\mathcal{N}(T-\lambda I)^{\dim V}$, אחרת $j>\dim V$ ואז הוכחנו $\mathcal{N}(T-\lambda I)^j=\mathcal{N}(T-\lambda I)^{\dim V}$. נסיק מעקרון ההחלפה:

$$\tilde{\mathcal{V}}_\lambda=\bigcup_{j=1}^{\inf}\mathcal{N}(T-\lambda I)^j=\bigcup_{j=1}^{\dim V}\mathcal{N}(T-\lambda I)^j\ \cup\bigcup_{j=\dim V}^{\inf}(T-\lambda I)^j=\mathcal{N}(T-\lambda I)^{\dim V}$$

■

משפט 79. בהינתן v ו"ע מוכלל של T, קיים (מהגדרה) ויחיד λ_i כך ש־$v\in\tilde{\mathcal{V}}_{\lambda_i}$.

הוכחה. ההוכחה בעיקר אלגברית ולא מעניינת במיוחד, יש צורך לפתח את הבינום של ניוטון. ■

מסתבר, שאפשר לפרק את המרחב למרחביים עצמיים מוכללים, ומשם אפשר להסיק מה קורה בהם ביתר פרטים בעזרת העתקות נילפוטנטיות.

משפט 80. הטענות הבאות מתקיימות:

1. $\tilde{\mathcal{V}}_{\lambda_i}$ הוא T־איווריאנטי.
2. $(T-\lambda_i I)|_{\tilde{\mathcal{V}}_{\lambda_i}}$ נילפוטנטית.
3. מעל שדה סגור אלגברית, הריבוי האלגברי d_{λ_i} הוא $\dim\tilde{\mathcal{V}}_{\lambda_i}$

הוכחה.

1. יהי $v \in \tilde{\mathcal{V}}_{\lambda_i}$, אז קיים $k \leqslant \mathcal{F}(T)$ כך ש־$(T-\lambda I)^k v = 0$. נפעיל את T על שני האגפים ונקבל $(T-\lambda I)^{k+1}v = T(0) = 0$ ולכן $Tv \in \tilde{\mathcal{V}}_{\lambda_i}$. סה"כ $\tilde{\mathcal{V}}_{\lambda_i}$ הוא T־אינווריאנטי.
2. נגדיר $S = (T-\lambda_i I)|_{\tilde{\mathcal{V}}_{\lambda_i}}$. לכן לכל $v \in \operatorname{dom} S = \tilde{\mathcal{V}}_{\lambda_i}$ מתקיים $\exists k_v \colon (T-\lambda_i I)^{k_v} = S^{k_v} = 0$. משום ש־$k_v \leqslant \mathcal{F}(T) \leqslant n$ נוכל לטעון ש־$\forall v \in \operatorname{dom} S \colon S^n v = 0$, ומהגדרה S נילפוטנטית כדרוש.
3. (הוכחה זו נכתבה בעזרתו האדיבה של chatGPT) נסמן $U = \tilde{\mathcal{V}}_{\lambda_i}$, ונרחיב את הבסיס של U לבסיס של V כך שנוצר מ"ו W כך ש־$U \oplus W = V$. ממשפט $p_T(x) = p_{T|_W}(x) \cdot p_{T|_U}(x)$. מסעיף קודם ידוע ש־$S := (T-\lambda_i I)$ ש־$S|_U$ נילפוטנטית, לכן n כך ש־$S|_U^n = 0$. נכתוב את T באופן הבא: $T|_U = S|_U + \lambda I \implies S|_U = T|_U - \lambda I$. נקבל שתי הבחנות:

- λ_i הוא הע"ע היחיד של $T|_U$, והוא ע"ע $\lambda_i \in U = \tilde{\mathcal{V}}_{\lambda_i}$ ולכן λ_i ע"ע של $T|_U$, והיחידות נובעת מכך שכל ע"ע מוכלל שייך לע"ע יחיד של T.
- $S|_W$ הפיכה, שכן בבירור $\ker S|_W \subseteq \ker(T-\lambda_i I) = V_{\lambda_i} \subseteq \tilde{\mathcal{V}}_{\lambda_i} = U$ ולכן $W \supseteq \ker S|_W$ ולכן $\ker S \subseteq W \cap U = \{0\}$.

נסיק משתי הטענות הללו שתי מסקנות:

- מהיות λ_i הע"ע היחיד של $T|_U$, ומהיות $\deg p_{T|_U} = \dim U$, ויחדיו עם ההנחה שאנחנו בשדה סגור אלגברית, $p_{T|_U}$ בהכרח מורכב מ־$\dim U$ גורמים לינאריים שהם $(x-\lambda_i)$.
- λ_i איננו ע"ע של $T|_W$, בגלל שאם (בשלילה) λ_i ע"ע של $T|_W$ עם ו"ע v אז $\lambda_i v = T|_W(v) = Sv + \lambda_i v$ ומחיסור אגפים נקבל $S|_W v = 0$, כלומר $v = 0$ (כי $S|_W$ הפיכה) ואז v לא ו"ע וסתירה.

סה"כ, מהיות $p_T(x) = p_{T|_W}(x) \cdot p_{T|_U}(x)$, נקבל שהריבוי האלגברי של $(x-\lambda_i)$ בא אך ורק מ־$p|_{T|_U}$ ושם הריבוי הוא $\dim U$, כלומר סה"כ הריבוי האלגברי של λ_i בהעתקה T הוא $\dim U = \dim \tilde{\mathcal{V}}_{\lambda_i}$ כדרוש.

■

הגדרה 56. v הוא ו"ע עצמי מורחב של λ_i מדרגה k אם הוא ו"ע עצמי מורחב של λ_i כך ש־$v \in \ker(T-\lambda I)^k \setminus \ker(T-\lambda I)^{k-1}$, כאשר בסיס $k = 1$ מוגדר להיות $v \in V_{\lambda_i}$.

משפט 81 (פירוק המרחב למרחבים עצמיים מוכללים). נניח שאנחנו במ"ו סגור אלגברית (אפשר להרחיב לכזה במידת הצורך). אז ל־T יש ע"עים $\lambda_1 \dots \lambda_k$ כלשהם. בהינתן V מ"ו ו־T העתקה לינארית, מההרחבה יש לה ערכים עצמיים $\lambda_1 \dots \lambda_k$ כלשהם. אזי:

$$V = \bigoplus_{i=1}^{k} \tilde{\mathcal{V}}_{\lambda_i}$$

הוכחה. נתחיל מלהוכיח שהחיתוך בין שני מרחבים עצמיים מוכללים ריק. זה נובע ישירות מכך שכל שני ע"ע עצמיים מוכללים שייכים לע"ע רגיל יחיד של T. ניעזר בכך ש־$\dim \tilde{\mathcal{V}}_{\lambda_i} = d_{\lambda_i}$, ונקבל:

$$\begin{cases} \sum_{i=1}^{n} \dim \tilde{\mathcal{V}}_{\lambda_i} = \sum_{i=1}^{n} d_{\lambda_i} = n \\ \forall i \in [k] \colon \tilde{\mathcal{V}}_{\lambda_i} \subseteq V \\ \forall i,j \in [k] \colon \tilde{\mathcal{V}}_{\lambda_i} \cap \tilde{\mathcal{V}}_{\lambda_j} = \{0\} \end{cases}$$

כאשר d_{λ_i} הריבוי האלגברי של λ_i, וידוע סכום הריבויים האלגבריים הוא n שכן $p_T(x)$ פולינום ממעלה n. לכן ממשפט יש סכום ישר כדרוש. ■

עתה נוכיח מחדש את משפט ג'ורדן, אך הפעם ללא תלות בפולינום מינימלי ופירוק פרימרי.

הוכחה באמצעות מרחבים עצמיים מוכללים. תהי העתקה T. מפריקות הפולינום האופייני יש לה $\lambda_1 \dots \lambda_k$ ע"עים כלשהם. ממשפט:

$$V = \bigoplus_{i=1}^{k} \tilde{\mathcal{V}}_{\lambda_i}$$

עוד ידוע שההעתקה $S_i = (T-\lambda_i)_{\tilde{\mathcal{V}}_{\lambda_i}}$ נילפוטנטית. כבר הוכחנו את צורת ג'ורדן עבור העתקות נילפוטנטיות ולכן ל־S_i קיים בסיס מג'ורדן $\mathcal{B}_i$. נבחין ש־:

$$T|_{\tilde{\mathcal{V}}_{\lambda_i}} = S_i + \lambda_i \implies \left[T|_{\tilde{\mathcal{V}}_{\lambda_i}}\right]_{\mathcal{B}_i} = \underbrace{\operatorname{diag}\{J_{a_1}(0) \dots J_{a_\ell}(0)\}}_{[S_i]_{\mathcal{B}_i}} + \lambda I = \operatorname{diag}\left(J_{a_1}(\lambda_i) \dots J_{a_\ell}(\lambda_i)\right)$$

ולכן אפשר לשרשר את הבסיסים לכדי בסיס מג'רדן: $\mathcal{B}=\bigcup_{i=1}^k \mathcal{B}_i$, ואכן:

$$[T]_\mathcal{B}=\operatorname{diag}\left(\left[T|_{\tilde{V}_{\lambda_i}}\right]_{\mathcal{B}_i}\mid i\in[k]\right)=\operatorname{diag}\left(J(\lambda_1)\dots J(\lambda_1)\dots J(\lambda_k)\dots J(\lambda_k)\right)$$

שרשור של בלוקי ג'ורדן. ■

הערה 32. מיחידות צורת ג'ורדן, הצורה המתקבלת מפירוק פרימרי ומפירוק למרחבים עצמיים מוכללים היא זהה. דרך אחרת לראות את זה, היא שהמרחבים אליהם פירקנו פרימרית שהם $(T-\lambda_i)^{d_\lambda}=\tilde{V}_{\lambda_i}$ בכל מקרה.

1.4.4 ~ תוצאות מצורת ג'ורדן

משפט 82. כמות בלוקי הג'ורדן לע"ע λ היא הריבוי הגיאומטרי.

הוכחה. נראה שהריבוי הגיאומטרי r_λ שווה לכמות בלוקי הג'ורדן השייכים ל־λ. בהינתן בלוקי ג'ורדן $J_{k_1}(\lambda)\dots J_{k_\ell}(\lambda)$ השייכים ל־λ, ידוע שלכל אחד מהם קיים בסיס שרשרת $B_i=\{v,(\tilde{T})v,\dots(\tilde{T})^{k_i}v\}$ כאשר $\tilde{T}=T-\lambda I$. כמו כן ידוע קיום $w_1\dots w_{r_\lambda}$ בלתי תלויים לינארית כך ש־$Tw_i=\lambda w_i$ כלומר $w_i\in\ker\tilde{T}$, ומכאן $\tilde{T}w_i=0$.

- **חסם עליון:** לכל k_i ידוע ש־$(\tilde{T})^{k_i}$ הוא המקסימלי שלא מאפס את v, כלומר $T((\tilde{T})^{k_i}v_i)=(\tilde{T})^{k_i+1}v=0$ ולכן $(\tilde{T})^{k_i}v_i\in\ker\tilde{T}=V_\lambda$. משום שכל השרשאות בלתי תלויות לינארית (אחרת השרשור שלהן לא יהווה בסיס), בהכרח $\{(\tilde{T})^{k_i}w_i\}_{i=1}^\ell\subseteq V_\lambda$ קבוצה בלתי תלוייה לינארית וה־span שלה תמ"ו של V_λ, ולכן $\ell\leqslant r_\lambda$.
- **חסם תחתון:** מהחסם העליון, כל w_i יכול להוות סיום של שרשרת, ולכן יש לפחות r_λ שרשראות שונות (שימו לב: יש לנו חופש בבחירת הבסיס ${w_i}_{i=1}^{r_\ell}$, ומכאן החופש בבחירת הבסיס המג'רדן), ומכאן החסם התחתון. ■

משפט 83. כמות הוקטורים בבסיס המג'רדן המשוייכים ל־λ הוא הריבוי האלגברי d_λ (ניסוח אחר: סכום גדלי הבלוקים השייכים ל־λ בצורת הג'ורדן הוא d_λ).

הוכחה. ראינו בצורת ג'ורדן בעזרת פירוק למרחבים עצמיים מוכללים, שמספר הוקטורים השייכים ל־λ הוא $\dim\tilde{V}_\lambda$ וידוע שזה מ"ו ממימד d_λ. סה"כ הראינו את הדרוש. ■

משפט 84. בלוק הגו'רדן המשויך ל־λ הגדול ביותר, הוא הריבוי של $(x-\lambda)$ בפולינום $m_T(x)$.

הוכחה. ראינו שבלוק הג'ורדן $J_a(\lambda)$ מגיע מפירוק ג'ורדן של $S=(T-\lambda)_{\tilde{V}_\lambda}$. הבלוק הכי גדול בצורת הגו'רדן של S נילפוטנטית, היא השרשרת הכי ארוכה של S ב־$\tilde{V}_\lambda$. משום ש־$\tilde{V}_{\lambda_i}=\{\ker S^k\mid k\in[n]\}=S^{n(S)}$, השרשרת הארוכה ביותר האפשרית היא $v,Sv\dots S^{n(S)}v$ והיא קיימת כי הסדרה הזו בת"ל עבור v כלשהו (אחרת $n(S)$ לא החזקה המינימלית שמאפסת את S וסתירה).

ראינו ש־m_T הוא ה־lcm של הצמצום של T למרחבים T־אינווראינטים, ומשום שכל $\tilde{V}_{\lambda_k}$ בעל פולינום אופייני $(x-\lambda_k)^{d_{\lambda_i}}$, ו־$\gcd((x-\lambda_i)^k,(x-\lambda_j)^m)=1$, ובגלל ש־$m_{T|_{\tilde{V}_\lambda}}(x)$ מחלק את $p_{T|_{\tilde{V}_\lambda}}(x)$, אז $m_{T|_{\tilde{V}_{\lambda_i}}}(x)=(x-\lambda)^k$ עבור $k\in[r_\lambda]$ כלשהו. אז:

$$\forall i\neq j\in[k]\colon\ \gcd\left(T|_{\tilde{V}_{\lambda_i}}(x),T|_{\tilde{V}_{\lambda_j}}(x)\right)=1$$

דהיינו, ה־lcm זה פשוט כפל של הפולינומים המינימליים של $T|_{\tilde{V}_\lambda}$. לכן, תחת הסימון m_λ להיות הריבוי של λ בפולינום m_T, בהכרח $m_{T|_{\tilde{V}_\lambda}}(x)=(x-\lambda)^{m_i}$. מהגדרת פולינום מינימלי, m_λ הוא המינימלי כך ש־$(T-\lambda)^{m_i}=0$ כלומר m_λ המינימלי כך ש־$S^{m_\lambda}=0$. סה"כ m_λ דרגת הנילפוטנטיות של S. הראנו ש־$n(S)$ השרשרת המקסימלית בצורת הג'ורדן של S, וסה"כ בלוק הגו'רדן הגדול ביותר של $J(\lambda)$ הוא m_λ הריבוי של $(x-\lambda)$ ב־$m_T(x)$. ■

משפט 85. תהי $A\in M_n(\mathbb{K})$ מטריצה, כאשר $\mathbb{K}$ סגור אלגברית. אז $A\sim A^T$.

הוכחה. ממשפט ג'ורדן ל־A יש צורת ג'ורדן Λ, כלומר קיימת P הפיכה כך ש־$P^{-1}\Lambda P=A$ ו־Λ מטריצה אלכסונית עם בלוקי ג'ורדן. נבחין בכך ש־$A^T=(P^{-1}\Lambda P)^T=P^T\Lambda^T(P^{-1})^T$, כלומר $A^T\sim\Lambda^T\wedge A\sim\Lambda$. נותר להוכיח $\Lambda\sim\Lambda^T$, כלומר, כל בלוק ג'ורדן $J_i(\lambda)\sim J_i(\lambda)^T$. טענה זו אכן מתקיימת בעבור מעבר לבסיס הסדור $(e_1\dots e_n)\to(e_n\dots e_1)$. סה"כ אכן כל מטריצה דומה לשחלוף שלה. ■

משפט 86. תהי $A\in M_n(\mathbb{F})$ מטריצה מעל $\mathbb{F}$ שדה. אז בהינתן $\lambda_1\dots\lambda_k$ הערכים העצמיים של A מעל $\mathbb{K}$ הרחבת $\mathbb{F}$ לסגור אלגברית, אז $\det A=\prod_{i=1}^k\lambda_i^{d_{\lambda_i}}$.

הוכחה. ידוע של־$\mathbb{F}$ קיימת הרחבה ל־$\mathbb{K}$. מעל $\mathbb{K}$, ל־A יש צורת ג'ורדן $A = P^{-1}NP$ כך ש־N מטריצת בלוקים הכוללת לפחות k בלוקים, כאשר הבלוק ה־$\square_i$ יסומן להיות הבלוק הכולל את בלוקי הג'ורדן המשוייכים לע"ע λ_i. אזי $\square_i$ מטריצה משולשית עליונה מגודל d_i (ממשפט קודם, לפיו כמות הוקטורים המשוייכים לע"ע λ היא d_λ) עם λ_i על האלכסון ולכן $\det \square_i = \lambda_i^{d_i}$. מדטרמיננטה של מטריצת בלוקים נסיק:

$$\det A = \det P^{-1}NP = \underbrace{\det \overbrace{P^{-1}P}^{I}}_{1} \cdot \det N = \prod_{i=1}^{k} \det \square_i = \prod_{i=1}^{k} \lambda_i^{d_i}$$

כדרוש. ■

המשך בעמוד הבא

פרק 2

הגדרה וחקר מרחבי מכפלה פנימית

2.1 Bi-Linear Forms

2.1.1 ~ הגדרות בסיסיות בעבור תבניות בי־לינאריות כלליות

הגדרה 57. יהי V מ"ו מעל $\mathbb{F}$. פונקציונל לינארי φ מעל V הוא $\varphi\colon V \to \mathbb{F}$.

הערה 33. ראה הרחבה על פונקציונלים לינאריים ומרחבים דואלים בסוף הסיכום.

הגדרה 58. יהיו V, W מ"ויים מעל $\mathbb{F}$. תבנית בי־לינארית על $V \times W$ הינה העתקה $f\colon V \times W \to \mathbb{F}$ כך ש־$\forall v_0 \in V\ \forall w_0 \in W$ כך שהעתקות $(v, w_0) \mapsto v,\ \ f(v_0, w) \mapsto w$ הן פונקציונליים לינאריים.

אינטואיטיבית, זו ההעתקה לינארית בכל אחת מהקורדינאטות בנפרד (בדומה לדוגמה לדטרמיננטה, שהיא העתקה מולטי־לינארית ולינארית בכל אחת מהשורות בנפרד)

משפט 87. הטענה הבאה שקולה לכך ש־f בי־לינארית. יהיו $\forall v \in V,\ w \in W,\ \alpha \in \mathbb{F}$:

$$\begin{aligned}\forall v_1, v_2 \in V\colon f(v_1 + v_2, w) &= f(v, w) + f(v_2, w)\\ \forall w_1, w_2 \in W\colon f(v, w_1, w_2) &= f(v, w_1) + f(v, w_2)\\ f(\alpha v, w) = \alpha f(v, w) &= f(v, \alpha w)\end{aligned}$$

בשביל העתקות n־לינאריות צריך טנזור n ממדי. זה לא נעים ויודעים מעט מאוד על האובייקטים הללו. בפרט, בעבור העתקה בי־לינארית נראה שנוכל לייצג אותה באמצעות מטריצות, בלי טנזור ובלגנים – שזה נחמד, וזו אחת הסיבות שאנו מתעסקים ספציפית עם העתקות בי־לינאריות (פרט לכך שמאוחר יותר נעסוק גם במכפלות פנימיות, וחלק מהתוצאות על ההעתקות בי־לינאריות יעזרו לנו להגיד דברים על מטריצות).

דוגמאות.

1. תבנית ה־0: $\forall v, w\colon f(v, w) = 0$
2. נגדיר $V = W = \mathbb{R}^2$, אז $f\left(\binom{x}{y}, \binom{u}{v}\right) = 2xu + 5xv - 12yu$
3. (חשוב) על $\mathbb{F}^n$:

   הגדרה 59. לכל שדה $\mathbb{F}$ מוגדרת התבנית הבי־לינארית הסטנדרטית היא:

$$f\left(\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}, \begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}\right) = \sum_{i=1}^{n} x_i y_i$$

4. יהיו $\varphi\colon V \to \mathbb{F},\ \psi\colon W \to \mathbb{F}$ פונקציונליים לינאריים: $f(v, w) = \varphi(v) \cdot \psi(w)$
5. הכללה של 4: יהיו $\varphi_1 \dots \varphi_k\colon V \to \mathbb{F}$ פונקציונליים לינאריים וכן $\psi_1 \dots \psi_k\colon W \to \mathbb{F}$ פונקציונליים לינאריים. אז ההעתקה הבאה בי־לינארית: $f(v, w) = \sum_{i=1}^{k} \varphi_i(v)\psi_i(w)$

הרעיון: ברגע שנקבע וקטור ספציפי נקבל לינאריות של הוקטור השני.

הערה 34. במקרה ש־$\mathbb{F} = \mathbb{R}$ לעיל, התבנית הבי־לינארית הסטנדרטית משרה את הגיאומטריה האוקלידית. כלומר $v \perp u \iff f(v, u) = 0$.

הערה 35. בעתיד נראה שכל תבנית בי־לינארית נראית כמו מקרה 5.

משפט 88. נסמן את מרחב התבניות הבי־לינאריות על $V \times W$ בתור $B(V, W)$. זהו מ"ו מעל $\mathbb{F}$.

אני ממש לא עומד להגדיר את החיבור והכפל בסקלר של המשפט הקודם כי זה טרוויאלי והמרצה כותב את זה בעיקר בשביל להטריל אותנו.

דוגמה חשובה אחרת.

משפט 89. נסמן ש־$\dim V = n,\ \dim W = m$ ותהי $A \in M_{n\times m}(\mathbb{F})$. יהי $\mathcal{A}$ בסיס ל־B, $\mathcal{B}$ בסיס ל־W. אז:

$$f(u, w) = [v]_{\mathcal{A}}^T \cdot A \cdot [w]_{\mathcal{B}}$$

העתקה בי־לינארית.

הוכחה. נְקַבֵּעַ v כלשהו:

$$[v]_{\mathcal{A}}^T \cdot A =: B \in M_{1\times m},\ \ g(w) := f(v, w) = B[w]_{\mathcal{B}}$$

נוכיח ש־g לינארית:

$$\forall w_1, w_2 \in V,\ \ \lambda_1, \lambda_2 \in \mathbb{F}\colon g(\lambda_1 w_1 + \lambda_2 w_2) = B[\lambda_1 w_1 + \lambda_2 w_2]_{\mathcal{B}} = \lambda_1(B[w_1]_{\mathcal{B}}) + \lambda_2(B[w_2]_{\mathcal{B}}) = \lambda_1 g(w_1) + \lambda_2 g(w_2)$$

נְקַבֵּע w, ובאופן דומה נגדיר $C = A[w]_{\mathcal{B}} \in M_{n\times 1}(\mathbb{F})$ ו־$h(v) := f(v,w) = [v]_{\mathcal{B}}^T C$:

$$\forall v_1, v_2 \in V,\ \lambda_1, \lambda_2 \in \mathbb{F}\colon h(\lambda_1 v_1 + \lambda_2 v_2) = [\lambda_1 v_1 + \lambda_2 v_2]_{\mathcal{B}}^T = \lambda_1([v_1]_{\mathcal{B}}^T C) + \lambda_2([v_2]_{\mathcal{B}}^T C) = h(v_1) + h(v_2)$$

■

("זה $\mathcal{A}$ mathcal, אתם תסתדרו" – המרצה ברגע שיש לו שני A־ים על הלוח)

הגדרה 60. בהינתן תבנית בי־לינארית $f\colon V\times W \to \mathbb{F}$ ונניח ש־$\mathcal{A}$ בסיס ל־V, $\mathcal{B}$ בסיס ל־W. נגדיר את המטריצה המייצגת את f ביחס לבסיסים $\mathcal{A}, \mathcal{B}$ ע"י $A \in M_{n\times m}(\mathbb{F})$ כאשר $(A)_{ij} = f(v_i, w_j)$ (תחת הסימונים $\mathcal{B} = (w_i)_{i=1}^m$ $\mathcal{A} = (v_i)_{i=1}^n$)

משפט 90. $f(v,w) = [v]_{\mathcal{A}}^T A [w]_{\mathcal{B}}$

הוכחה. קיימים ויחידים $\alpha_1 \dots \alpha_n,\ \beta_1 \dots \beta_m \in \mathbb{F}$ כך ש־$v = \sum \alpha_i v_i,\ w = \sum b_i w_i$. כלומר:

$$[v]_{\mathcal{A}}^T = (\alpha_1 \dots \alpha_n),\ [w]_{\mathcal{B}} = \begin{pmatrix}\beta_1 \\ \vdots \\ \beta_m\end{pmatrix}$$

ומכאן פשוט נזרוק אלגברה:

$$\begin{aligned}
f(v,w) =& f\left(\sum_{i=1}^n \alpha_i v_i,\ w\right) = \sum_{i=1}^n \alpha_i f(v_i,\ w) \\
&= \sum_{i=1}^n \alpha_i f\left(v,\ \sum_{j=1}^n \beta_j w_j\right) = \sum_{i=1}^n \alpha_i \left(\sum_{j=1}^m \beta_j f(v_i, w_j)\right) \\
&= \sum_{i,j\in[n]\times[m]} \alpha_i f(v_i, w_j)\beta_j \\
&= \left(\sum_{i=1}^n \alpha_i a_{i1},\ \sum_{i=1}^n \alpha_i a_{i2},\ \dots,\ \sum_{i=1}^n \alpha_i a_{im}\right)\begin{pmatrix}\beta_1 \\ \vdots \\ \beta_m\end{pmatrix} \\
&= (\alpha_1 \dots \alpha_n)\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm}\end{pmatrix}\begin{pmatrix}\beta_1 \\ \vdots \\ \beta_m\end{pmatrix}
\end{aligned}$$

■

סימון 8. נאמץ לסיכום הזה את הסימון $[f]_{\mathcal{A},\mathcal{B}}$ עבור המטריצה המייצגת של f בי־לינארית.

(זהו **אינו** סימון רשמי בקורס אם כי בהחלט צריך להיות)

משפט 91. עם אותם הסימונים כמו קודם:

$$\psi\colon B(v,w) \to M_{n\times m}(\mathbb{F}),\ f \mapsto [f]_{\mathcal{A},\mathcal{B}}$$

אז ψ איזו'.

הוכחה. נסמן את $[f]_{\mathcal{A},\mathcal{B}} = A$ ואת $[g]_{\mathcal{A},\mathcal{B}} = B$. אז:

- **לינאריות.**

$$\begin{aligned}
(\mathcal{P}(f+g))_{ij} &= (f+g)(v_i, w_j) \\
&= f(v_i, w_j) + g(v_i, w_j) \\
&= (A)_{ij} + (B)_{ij} \\
&= (\psi(f))_{ij} + (\psi(g))_{ij} \implies \psi(f+g) \\
&= \psi(f) + \psi(g)
\end{aligned}$$

באופן דומה בעבור כפל בסקלר:

$$(\psi(\alpha f))_{ij} = \alpha f(v_i, w_j) = \alpha(\psi(f))_{ij} \implies \psi(\alpha f) = \alpha\psi(f)$$

- **חח"ע.** תהי $f \in \ker\psi$, אז: $\forall i,j \in [n]\times[m]\colon f(v_i,w_j)=0 \implies \psi(f)=0\in M_{n\times m}$ ולכן $\forall v\in V, w\in W\colon f(v,w)=\sum_{i,j\in[n]\times[m]}\alpha_i f(v_i,w_j)\beta_j=0$ (עם אותם הסימונים כמו קודם)
- **על.** תהי $A\in M_{n\times m}(\mathbb{F})$. נגדיר $f(v,w)=[v]_{\mathcal{A}}^T A[w]_{\mathcal{B}}$ ואכן $f(v_i,w_j)=e_i^T Ae_j=(A)_{ij}$. ■

תזכורת (מלינארית 1). מטריצת המעבר מבסיס $\mathcal{B}$ לבסיס $\mathcal{C}$ מוגדרת להיות $[I]_{\mathcal{C}}^{\mathcal{B}}$, היא מטריצה הפיכה, ומתקיים השוויון $[I]_{\mathcal{B}}^{\mathcal{C}}[T]_{\mathcal{B}}^{\mathcal{B}}[I]_{\mathcal{C}}^{\mathcal{B}}=[T]_{\mathcal{C}}^{\mathcal{C}}$.

משפט 92. יהיו V,W מ"ום מעל $\mathbb{F}$ נניח $\mathcal{A},\mathcal{A}'\subseteq V$ בסיסים של V וכן $\mathcal{B},\mathcal{B}'\subseteq W$ בסיסים של W. תהי $f\in B(V,W)$. תהי המייצגת של f לפי $\mathcal{A},\mathcal{B}$ היא A ותהי A' המייצגת בבסיסים $\mathcal{A}',\mathcal{B}'$. תהי P מטריצת המעבר מ־$\mathcal{A}$ ל־$\mathcal{A}'$ ו־Q מטריצת המעבר מ־$\mathcal{B}$ ל־$\mathcal{B}'$, אז $A'=P^TAQ$.

הוכחה. ידוע:

$$P[v]_{\mathcal{A}}=[v]_{\mathcal{A}'},\ \ Q[w]_{\mathcal{B}}=[w]_{\mathcal{B}'}$$

ואכן:

$$f(v,w)=[v]_{\mathcal{A}}^TA[w]_{\mathcal{B}}=(P[v]_{\mathcal{A}'})^TA(Q[w]_{\mathcal{B}'})=[v]_{\mathcal{A}'}^T\left(P^TAQ\right)[w]_{\mathcal{B}'}\implies A'=P^TAQ$$

כדרוש. ■

הגדרה 61. עבור $f\in B(V,W)$ נגדיר את $\operatorname{rank} f=\operatorname{rank} A$ כאשר A מייצגת אותה ביחס לבסיסים כלשהם.

משפט 93. $\operatorname{rank} f$ מוגדר היטב.

הוכחה. כפל בהפיכה לא משנה את דרגת המטריצה (ו־transpose של מטריצה הוא הפיך), ומטריצת שינוי הבסיס הפיכה, דהיינו כפל מטריצות שינוי הבסיס לא משנות את דרגת המטריצה ולכן לכל שני נציגים אותה הדרגה. ■

מסקנה 17. תהא $f\in B(V,W)$ ונניח $\operatorname{rank} f=r$. אז קיימים בסיסים $\mathcal{A},\mathcal{B}$ של W,V בהתאמה כך ש־$[f]_{\mathcal{A},\mathcal{B}}=\begin{pmatrix}I_R&0\\0&0\end{pmatrix}$

הרעיון הוא לדרג את כל כיוון, שורות באמצעות transpose ועמודות באמצעות המטריצה השנייה. אפשר גם לקבע בסיס, ולדרג שורות ועמודות עד שיוצאים אפסים (הוכחה לא נראתה בכיתה).

"חצי השעה הזו גרמה לי לשנוא מלבנים בצורה יוקדת" – מעתה ואילך נתעסק במקרה בו $V=W$. נשתמש בבסיס יחיד.

2.1.2 ∼ חפיפה וסימטריות

הגדרה 62. יהיו $A,A'\in M_n(\mathbb{F})$, נאמר שהן חופפות אם קיימת הפיכה $P\in M_n(\mathbb{F})$ כך ש־$A'=P^TAP$.

משפט 94. מטריצות חופפות אמ"מ הן מייצגות את אותה התבנית הבי־לינארית.

משפט 95. אם A,A' חופפות, אז:

1. $\operatorname{rank} A=\operatorname{rank} A^T$
2. $\exists 0\neq c\in\mathbb{F}\colon\ \det A'=c^2\det A$

הוכחה. הגדרנו $\operatorname{rank} f$ כאשר f בי־לינארית להיות הדרגה של המייצגת את התבנית, וראינו שהיא לא תלויה בבסיס. בכך למעשה כבר הוכחנו את 1. עבור 2, מתקיים $A'=P^TAP$ ו־P הפיכה (ולמעשה מטריצת מעבר בסיס) ולכן אם נסמן $c=|P|=|P^T|$ מתקיים:

$$|A'|=|P^TAP|=|P^T|\cdot|A|\cdot|P|=|P|^2|A|=c^2|A|$$

■

הערה 36. יש שדות שמעליהם טענה 2 לא מעניינת במיוחד (שדות עבורם יש שורש לכל מספר, כמו $\mathbb{C}$).

הגדרה 63. תבנית f מעל V נקראת סימטרית אם: $\forall v,w\in V\colon f(v,w)=f(w,v)$

הגדרה 64. תבנית f מעל V נקראת אנטי־סימטרית אם: $\forall v,w\in V\colon f(v,w)=-f(w,v)$

משפט 96 (פירוק תבנית בי־לינארית לחלק סימטרי וחלק אנטי־סימטרי). אם $\operatorname{char}\mathbb{F}\neq 2$, בהינתן תבנית $f\colon V\times V\to\mathbb{F}$ בי־לינארית, קיימות $\varphi,\psi\colon V\times V\to\mathbb{F}$ בי־לינאריות כך ש־φ סימטרית, ψ אנטי־סימטרית ו־$f=\varphi+\psi$.

הוכחה. נבחין שאם $\operatorname{char}\mathbb{F}\neq 2$, ניתן להגדיר את:

$$\varphi(v,w)=\frac{f(v,w)+f(w,v)}{2},\ \ \psi(v,w)=\frac{f(v,w)-f(w,v)}{2}$$

מתקיים ש־φ סימטרית ו־ψ אנטי־סימטרית וכן $f=\varphi+\psi$. ■

משפט 97. תהי f תבנית בי־לינארית על V, ו־$B = (v_i)_{i=1}^n$ בסיס ל־B. נניח $A = (a_{ij})_{i,j=1}^n$ המייצגת את f ביחס ל־B. אז f סימטרית/אנטי־סימטרית אמ"מ A סימטרית/אנטי־סימטרית.

הוכחה.

$\Longrightarrow$ אם f סימטרית/אנטי־סימטרית, אז:

$$a_{ij} = f(v_i, v_j) \overset{\text{Sym}}{=} f(v_j, v_i) = a_{ji}$$
$$a_{ij} = f(v_i, v_j) \overset{\text{ASym}}{=} -f(v_j, v_i) = -a_{ji}$$

$\Longleftarrow$ אם A סימטרית אז:

$$f(v, w) = [u]_B^T A[w]_B \overset{(1)}{=} ([u]_B^T A[w]_B)^T = [w]_B^T A^T([u]_B^T)^T = [w]_B^T A[u]_B = f(w, v)$$

כאשר (1) מתקיים כי transpose למטריצה מגודל 1×1 מחזיר אותו הדבר. וכן במקרה האנטי־סימטרי:

$$f(u, w) = [w]_B^T(-A)[u]_B = -[w]_B^T A[u]_B = -(w, u)$$

■

הגדרה 65. בעבור תבנית בי־לינארית, הראדיקאל הימני שלה מוגדר להיות $\text{rad}_r(f) = \{x \in V \mid \forall w \in W\colon f(x, w) = 0\}$.
הגדרה 66. בעבור תבנית בי־לינארית, הראדיקאל השמאלי שלה מוגדר להיות $\text{rad}_\ell(f) = \{x \in W \mid \forall v \in V\colon f(v, x) = 0\}$.
משפט 98. הראדיקלים מרחבים וקטוריים.

הוכחה. יהי $x, y \in \text{rad}_r(f)$ וכן $\lambda \in \mathbb{F}$. נראה ש־$\lambda x + y \in \text{rad}_r(f)$. אכן, מלינאריות, מתקיים:

$$\forall v \in V\colon f(\lambda x + y, v) = f(\lambda x, v) + f(y, v) = \lambda \underbrace{f(x, v)}_{0} + \underbrace{f(y, x)}_{0} = \lambda 0 + 0 = 0$$

כדרוש. ההוכחה זהה לראדיקל השמאלי. ■

משפט 99. בהינתן תבנית סימטרית, $\text{rad}_r(f) = \text{rad}_\ell(f)$.

הוכחה. יהי x. לכל $v \in V$, מסימטריות מתקיים $f(x, v) = 0 \iff f(v, x) = 0$. סה"כ $x \in \text{rad}_\ell(f) \iff x \in \text{rad}_r(f)$ דהיינו $\text{rad}_r(f) = \text{rad}_\ell(f)$ כדרוש. ■

משפט 100. לכל תבנית $f\colon V \times W$, מתקיים $\dim \text{rad}_r(f) = \dim \text{rad}_\ell(f)$.

הוכחה. בהינתן בסיסים $\mathcal{B}, \mathcal{C}$ נגדיר $A = [f]_{\mathcal{B}}^{\mathcal{C}}$. נבחין ש־:

$$\begin{aligned}
\text{rad}_r(f) &\cong \{x \in \mathbb{F}^n \mid \forall v \in \mathbb{F}^m\colon v^T Ax = 0\} = \{x \in \mathbb{F}^n \mid \forall i \in [n]\colon \overbrace{e_i^T Ax}^{(Ax)_i} = 0\} = \{x \in \mathbb{F}^n \mid Ax = 0\} = \mathcal{N}(A) \\
\text{rad}_\ell(f) &\cong \{x \in \mathbb{F}^m \mid \forall v \in \mathbb{F}^n\colon x^T Av = 0\} = \{x \in \mathbb{F}^n \mid \forall i \in [m]\colon \underbrace{x^T Ae_i}_{x^T \text{Row}_i(A)} = 0\} = \{x \in \mathbb{F}^n \mid A^T x = 0\} = \mathcal{N}(A^T)
\end{aligned}$$

ידוע ממשפט הדרגה ש־$\text{rank}\, A = \text{rank}\, A^T$ וממשפט הממדים $\dim \mathcal{N}(A) = \dim \mathcal{N}(A^T)$. סה"כ $\dim \text{rad}_\ell(f) = \dim \text{rad}_r(f)$ כדרוש. ■

הערה 37. ניתן להוכיח את הטענה באופן כללי למרחבים לא נוצרים סופית באמצעות שימוש במרחבים דואליים ומרחבי מנה.
הגדרה 67. תבנית בי־לינארית f נקראת לא־מנוונת או רגולרית או לא־סינגולרית אם $\text{rad}(f) = \{0\}$.
הגדרה 68. תבנית בי־לינארית f נקראת מנוונת או סינגולרית או לא־רגולרית אם היא לא לא־מנוונת.
הערה 38. אין צורך לציין איזה ראדיקל שווה לאפס, שכן הם שווי ממד.

2.1.3 ~ תבניות ריבועיות

הגדרה 69. תהא f תבנית על V. התבנית הריבועית:

$$Q_p\colon V\to\mathbb{F},\ Q_f(v)=f(v,v)$$

היא לא העתקה לינארית. היא תבנית ריבועית. **דוגמאות:**

- $$f\left(\begin{pmatrix}x\\y\end{pmatrix},\ \begin{pmatrix}u\\v\end{pmatrix}\right)=xv+uy\implies Q\left(\begin{pmatrix}x\\y\end{pmatrix}\right)=xy+yx=2xy$$
- $$f\left(\begin{pmatrix}x\\y\end{pmatrix},\ \begin{pmatrix}u\\v\end{pmatrix}\right)=xv-uy\implies Q\left(\begin{pmatrix}x\\y\end{pmatrix}\right)=xy-yx=0$$
- התבנית הסטנדרטית:

$$f\left(\begin{pmatrix}x\\y\end{pmatrix},\begin{pmatrix}u\\v\end{pmatrix}\right)=xu+yv\implies Q_f\begin{pmatrix}x\\y\end{pmatrix}=x^2+y^2$$

סימון 9. עבור תבנית בי־לינארית f על V, נגדיר את $\hat{f}(u,v)=f(v,u)$

אם f סימטרית נבחין ש־$Q_f=Q_{\hat{f}}$

משפט 101 (שחזור תבנית בי־לינארית מתבנית ריבועית). תהי f תבנית בילי' סימטרית על V, ונניח ש־$\operatorname{char}\mathbb{F}\neq 2$, אז:

1. $$f(v,w)=\frac{Q_f(v+w)-Q_f(v)-Q_f(w)}{2}$$
2. אם f איינה תבנית ה־0 אז קיים $0\neq v\in V$ כך ש־$Q_f(v)\neq 0$.

הוכחה.

$$\begin{aligned}Q_f(v+w)-Q_f(v)-Q_f(w)=&f(v+w,v+w)-f(v,v)-f(w,w)\\=&\quad f(v,v)+f(v,w)\\&-f(w,v)+f(w,w)\\&-f(v,v)-f(w,w)\\\overset{\text{Sym}}{=}&2f(v,w)\end{aligned}$$

ובכך הוכחנו את 1. עתה נוכיח את 2. נניח $\forall v\in V\colon Q_f(v)=0$ אז

$$\forall v,w,\in V\colon f(v,w)=\frac{Q_f(v+w)-Q_f(v)-Q_f(w)}{2}=0$$

אז

$$f(\binom{x}{y},\ \binom{u}{v})=xv+yu\implies Q_f=0\land f\neq 0$$

■

הערה 39. אין ממש טעם להגדיר תבנית ריבועית על תבנית בי־לינארית שאיננה סימטרית. הסיבה? כל מטריצה יכולה להיות מפורקת לחלק סימטרי וחלק אנטי־סימטרי, החלק האנטי־סימטרי לא ישפיע על התבנית הריבועית (כי אלכסון אפס במטריצה המייצגת) ורק החלק הסימטרי ישאר בכל מקרה. זה לא שאי־אפשר, זה פשוט לא מעניין במיוחד.

2.1.4 ~ משפט ההתאמה של סילבסטר

משפט 102. נניח $\operatorname{char}F\neq 2$, ו־$f$ סימטרית על V. אז קיים בסיס ל־V הוא $B=(v_i)_{i=1}^n$ כך ש־$[f]_B$ אלכסונית. אם $\mathbb{F}=\mathbb{R}$, אז האיברים על האלכסון יהיו $\{1,-1,0\}$ ולא רק $\{1,0\}$.

תזכורת: $[f]_B$ סימון המוגדר בסיכום זה בלבד. בקורס מדברים על "המטריצה המייצגת של בי־לינארית" במילים מפורשות.

הוכחה. באינדוקציה על n. בסיס $n=1$ ברור. אם f תבנית ה־0, אז כל בסיס שנבחר מתאים. אחרת, קיים $0\neq v\in V$ כך ש־$Q_f(v)\neq 0$. נגדיר $U=\{u\in V\mid f(u,v)=0\}$. תמ"ו כי גרעין של ה"ל (כי קיבענו את v). מה התמונה של ההעתקה?

$f(v,v)=Q_f(v)\neq 0$. לכן תמונת ההעתקה היא כל $\mathbb{F}$, וממדה 1. ידוע U תמ"ו מממד $n-1$. אז $f_{|U}\colon U\times U\to\mathbb{F}$ לכסינה ולכן קיים בסיס B_U כך ש־$[f_{|U}]$ אלכסונית. נגדיר את $B=\{v\}\cup B_U$ נבחין שהיא מהצורה:

$$\begin{pmatrix} * & 0 & \cdots \\ 0 & [f_{|U}]_B & - \\ \vdots & | & \end{pmatrix}$$

סה"כ מהנחת האינדוקציה צעד האינדוקציה הושלם כדרוש. ■

הערה 40. מטריצה הפיכה לעיתים קרויה "לא־סינגולרית"

מסקנה 18. תבנית בי־לינארית היא לא־סינגולרית אמ"מ המטריצה המייצגת שלה (בבסיס כלשהו) היא לא־סינגולרית.

משפט 103. לכל f תבנית סימטרית קיימת מטריצה מייצגת מהצורה $\left(\begin{smallmatrix} I_r & 0 \\ 0 & 0\end{smallmatrix}\right)$ כאשר $\mathbb{F}=\mathbb{C}$ (או סגור אלגברית כלשהו).

אינטואיציה להוכחה. ננרמל את המטריצה, נבחין שחלוקה ב־c של השורה ה־i ניאלץ להפעיל גם עם העמודה ה־i, כלומר את $a_{i,i}$ נחלק ב־c^2 בצורה הזו (זאת כי כאשר P^TAP הגדרת חפיפה, ו־P מדרגת שורות, P^T מדרגת עמודות).

הוכחה. נסמן את $\dim f=r$. עד כדי שינוי סדר איברי הבסיס, המטריצה המייצגת אלכסונית היא:

$$[f]_B=\begin{pmatrix} \mathrm{diag}(c_1\dots c_r) & 0 \\ 0 & 0\end{pmatrix}$$

כאשר $c_1\dots c_r\neq 0$, ביחס לבסיס $B=(v_1\dots v_r,\dots v_n)$. באופן כללי לכל $i\in\mathbb{R}$ נוכל להגדיר את $v_i'=\frac{v_i}{\sqrt{c_i}}$ כך ש־$f(v_i',v_i')=1$ כי $f(v_i,v_i)=c_i$ ומליניאריות בכל אחת מהקורדינאטות. בשל כך $B'=(v_1'\dots v_r',v_{r+1}\dots v_n)$ בסיס המקיים את הדרוש.

באותו האופן, אם $\mathbb{F}=\mathbb{R}$ (ולא $\mathbb{C}$) אז קיים בסיס שהמטריצה המייצגת לפיו היא:

$$\begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0\end{pmatrix}$$

בלוקים, כך ש־$p+q=r$. כאן נגדיר:

$$f(v,v)=c<0,\ v'=\frac{v}{\sqrt{|c|}},\ f(v',v')=\frac{c}{|c|}=-1$$

■

הגדרה 70. יהי V מ"ו מעל $\mathbb{R}$ ו־f תבנית בי־לינארית מעל V. נאמר ש־f מוגדרת חיובית/אי־שלילית/שלילית/אי־חיובית אם $\forall 0\neq v\in V$ מתקיים ש־$f(v,v)>0$ / $f(v,v)\geqslant 0$ / $f(v,v)<0$ / $f(v,v)\leqslant 0$

הערה 41. באנגלית: "Definite matrix"

משפט 104. תהא A מטריצה מייצגת של תבנית בי־לינארית סימטרית, עם ערכים $0,-1,1$ בלבד על האלכסון, מקיימת:

- f מוגדרת חיובית אמ"מ ישנם רק 1־ים.
- f מוגדרת אי־שלילית אמ"מ ישנם רק 1־ים ואפסים.
- f מוגדרת שלילית אמ"מ ישנם רק -1־ים
- f מוגדרת חיובית אמ"מ ישנם רק -1־ים ואפסים.

הוכחה.

$\Leftarrow$ טרוויאלי

$\Longrightarrow$ לכל $0\neq v\in V$ קיימים ויחידים $\alpha_1\dots\alpha_n\in\mathbb{R}$ כך ש־$v=\sum_{i=1}^n\alpha_iv_i$ ומתקיים $f(v,v)=\alpha_i^2f(v_{i,i})$ ולפי המקרה זה יסתדר יפה. ■

משפט 105 (משפט ההתאמה של סילבסטר). p,q עבורם המטריצה חופפת ל־$\mathrm{diag}(I_p,I_{-q},0)$ נקבעים ביחידות.

(תחזרו כמה משפטים למעלה למקרה בו $\mathbb{F}=\mathbb{R}$)

גישה שגויה להוכחה. הוכחה באמצעות tr לא עובדת. בניגוד ליחס הדמיון להעתקות לינאריות, ביחס החפיפה להעתקות בי־לינאריות ה־tr לא נשמר. ■

הוכחה תקינה. נסמן $B = (v_1 \dots v_p, u \dots u_q, w_1 \dots w_k)$ וכן $B' = (v'_1 \dots v'_t, u'_1 \dots u'_s, w_1 \dots w_k)$ כי $t + s = p + q$. בה"כ $t \leqslant p$, נניח בשלילה ש־$t < p$. נסמן $U = \operatorname{span}(v_1 \dots v_p)$. ידוע f חיובית על U, וכן $\dim U = p$. נתבונן ב־$W = \operatorname{span}(u'_1 \dots u'_s, w_1 \dots w_k)$. אזי גם f חיובית על W, ו־$\dim W = s + k$. בגלל ש־$U \cap W = \{0\}$ (כי אם לא, אז עבור $0 \neq v \in U \cap W$ נקבל $f(v,v) > 0$ כי $v \in U$ וכן $f(v,v) \leqslant 0$ כי $v \in W$ וסתירה). ידוע ש־$U \oplus W \subseteq V$ תמ"ו וכן $\dim U + \dim W \leqslant \dim V$. נציב ונקבל $\dim V = t + s + k > p + s + k$, סתירה. לכן p, q נקבעים ביחידות. ■

סימון 10. ה־(p, q) לעיל נקראים הסיגנטורה של f.

(תזהרו, הסיגנטורה תתקוף אותנו אח"כ)

המשך בעמוד הבא

Inner Product Vector Spaces 2.2

2.2.1 ~ הגדרה כללית

(2.2.1.1) מעל $\mathbb{R}$

מעתה ועד סוף הקורס, מתקיים $\mathbb{F} = \mathbb{R}, \mathbb{C}$. כל עוד נאמר "$\mathbb{F}$", זה נכון בעבור שני המקרים. אחרת, נפצל.

הגדרה 71. יהי V מ"ו, מכפלה פנימית מעל $\mathbb{R}$ היא תבנית בי־לינארית סימטרית חיובית מעל V, ומסומנת $f(v, u) = \langle v, u \rangle$ (ויש ספרים שמסמנים $\langle v \mid u \rangle$, $\langle \cdot \mid \cdot \rangle : V \times V \to \mathbb{R}$).

מסקנה 19. התבנית f היא מ"פ אמ"מ המייצגת שלה A בבסיס כלשהו, היא סימטרית חיובית.

סימון 11. בקורס מסמנים $\langle \cdot, \cdot \rangle$ אבל אני מגניב אז אני משתמש ב־$\langle \cdot \mid \cdot \rangle$.

למה 9. $\forall v \in V \colon \ \langle v \mid v \rangle \geqslant 0$ ו־$\langle v, v \rangle$ אמ"מ $v = 0$.

הוכחה מסימטריה.

דוגמה. (המכפלה הפנימית הסטנדרטית על $\mathbb{R}^n$, AKA כפל סקלרי):

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \middle| \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \sum_{i=1}^{n} x_i y_i$$

הגדרה 72. אם V מ"ו וקיימת $\langle \cdot \mid \cdot \rangle : V \times V \to \mathbb{F}$ מכפלה פנימית אז $(V, \langle \cdot \mid \cdot \rangle)$ נקרא מרחב מכפלה פנימית, ממ"פ.

משפט 106. $V = M_n(\mathbb{R})$, אז $\langle A \mid B \rangle = \operatorname{tr}(A \cdot B^T)$ אז $(V, \langle \cdot \mid \cdot \rangle)$ ממ"פ.

דוגמה מגניבה. בהינתן $V = [0, 1]$, מ"ו הפונקציות הממשיות הרציפות על $[0, 1]$, ו־$\langle f \mid g \rangle = \int_0^1 f(x) \cdot g(x) \, \mathrm{d}x$

משפט 107. (שהפליצו מחדו"א) אם $f \geqslant 0$ אינטרבילית על קטע $[a, b]$ וגם ישנה נקודה חיובית $c \in [a, b]$ שעבורה $f(x) \geqslant 0$ וגם f רציפה ב־c, אז $\int_a^b f(x) \, \mathrm{d}x > 0$.

(2.2.1.2) מעל $\mathbb{C}$

ישנה בעיה עם חיובית: אם $v \in V$ כך ש־$\langle v \mid v \rangle \geqslant 0$ אך $\langle iv \mid iv \rangle = -1 \langle v \mid v \rangle < 0$ סתירה. לכן, במקום זאת, נשתמש בהגדרה הבאה:

הגדרה 73. יהי V מ"ו מעל $\mathbb{C}$. מכפלה פנימית $\langle \cdot \mid \cdot \rangle : V \times V \to \mathbb{C}$ מקיימת:

- לינאירות ברכיב הראשון: אם נקבע v, אז $u \mapsto \langle v \mid u \rangle$ לינארית.
- ססקווי־לינאריות/אנטי־לינארית ברכיב השני (במקום לינאריות): $\langle u_1 + u_2 \mid v \rangle = \langle u_1 \mid v \rangle + \langle u_2 \mid v \rangle \wedge \langle u \mid \alpha v \rangle = \bar{\alpha} \langle u \mid v \rangle$

  כאשר $\bar{\alpha}$ הצמוד המרוכב של α.
- הרמטיות (במקום סימטריות): $\langle v \mid u \rangle = \overline{\langle u \mid v \rangle}$
- חיוביות ואנאיזוטרופיות: $\forall 0 \neq v \in V \colon \ \langle v \mid v \rangle > 0 \wedge \langle 0 \mid 0 \rangle = 0$

למעשה – נבחין שאין צורך בממש ססקווי־ליניאיריות ברכיב השני וכן לא בתנאי $\langle 0 \mid 0 \rangle = 0$, וההגדרה שקולה בעבור חיבוריות ברכיב השני בלבד, זאת כי:

$$\langle u \mid \alpha v \rangle = \overline{\langle \alpha v \mid u \rangle} = \overline{\alpha \langle v \mid u \rangle} = \bar{\alpha} \cdot \overline{\langle v \mid u \rangle} = \bar{\alpha} \langle v \mid u \rangle$$

ומכאן נגרר ססקווי־ליניאריות, וכן $\langle 0 \mid 0 \rangle = 0$ נובע ישירות מליניאריות ברכיב השני.

הערה 42. באוניברסיטאות אחרות מקובל להגדיר לינאריות ברכיב השני ולא בראשון. זה לא באמת משנה.

מסקנה 20. נוכל להגדיר מטריצה A מייצגת של תבנית f ססקווי־בילניארית, באופן דומה להגדרה והרגילה, ואז להבחין שגם כאן A סימטרית חיובית אמ"מ f מ"פ. עוד נבחין ש־:

$$\langle v \mid u \rangle = f(v, u) = vAu^* = vA\bar{u}^T$$

הגדרה 74. למטריצה המייצגת של המכפלה הפנימית קוראים מטריצת גראם או גראמיאן.

הגדרה 75. $\overline{B^T} = B^*$

הגדרה 76 (הגדרה נחמדה)**.** יהי ממ"פ V מ"ו מעל $\mathbb{F}$. לכל $v \in V$ מגדירים את הנורמה של v להיות $||v|| = \sqrt{\langle v \mid v \rangle}$.

משפט 108. הנורמה כפלית וחיובית.

הוכחה. מאקסיומת החיוביות:

$$||v|| \geqslant 0 \wedge (||v|| = 0 \iff v = 0)$$

וכן:

$$\left|\left|t \cdot v^2\right|\right| = \langle tv \,|\, tu \rangle = t\bar{t} \langle v \,|\, v \rangle = |t| \,||v|| \implies ||t \cdot v|| = |t| \cdot ||v||$$

■

הגדרה 77. יהי V מ"ו מעל $\mathbb{F}$, ו־$||\cdot|| : V \to \mathbb{R}_{\geqslant 0}$, אז $(V, ||\cdot||)$ יקרא מרחב נורמי.

משפט 109. ("נוסחאת הפולריזציה") בהינתן $(V, ||\cdot||)$ מרחב נורמי, ניתן לשחזר את המכפלה הפנימית, באמצעות הנוסחה הבאה:

גרסה מעל $\mathbb{R}$:

$$\forall v, u \in V: \ \langle v \,|\, u \rangle = \frac{1}{4}(||u+v||^2 + ||u-v||^2)$$

גרסה מעל $\mathbb{C}$:

$$\langle u \,|\, v \rangle = \frac{1}{4}\Big(\, ||u+v||^2 - ||u-v||^2 + i \,||u+iv|| - i \,||u+iv|| \, \Big)$$

הוכחה (ל־$\mathbb{C}$).

$$\begin{aligned}
\langle u+v \,|\, u+v \rangle &= ||u||^2 + \langle u \,|\, v \rangle + \langle v \,|\, u \rangle + ||v||^2 \\
&= ||u||^2 + ||v||^2 + 2\Re(\langle v \,|\, u \rangle) \\
\langle v-u \,|\, v-u \rangle &= ||u||^2 + ||v||^2 - 2\Re(\langle v \,|\, u \rangle) \\
\langle u+iv \,|\, u+iv \rangle &= ||u||^2 + ||v||^2 + \langle u \,|\, iv \rangle + \langle iv \,|\, u \rangle \\
&= ||u||^2 + ||v||^2 - i \langle u \,|\, v \rangle + i\overline{\langle u \,|\, v \rangle} \\
&= ||u||^2 + ||v||^2 - i(2\Im \langle u \,|\, v \rangle) \\
||u-iv|| &= ||u|| + ||v|| - \langle u \,|\, iv \rangle - \langle iv \,|\, u \rangle \\
&= ||u|| + ||v|| - 2\Im(\langle v \,|\, u \rangle)
\end{aligned}$$

וסה"כ אם נציב בנוסחה, אחרי שחישבנו את כל אבירה, הכל יצטמצם וש־$\langle u \,|\, v \rangle$ אכן שווה לדרוש. ■

במילים אחרות, באותה המידה שתבניות שמתבניות בי־לינאריות ותבניות ריבועיות אפשר להסיק אחת מהשניה, אפשר גם ממכפלה פנימית להסיק נורמה ולהפך. אזי, ממ"פ ומרחב נומרי הם די שקולים. זה לא מפתיע, בהתחשב בזה שכל תבנית סימטרית לא־מנוונת משרה באופן יחיד תבנית ריבועית, ולהפך (במאפיין שונה מ־2).

2.2.2 ∼ אורתוגונליות, זהויות ואי־שוויונות של המכפלה הפנימית

הגדרה 78. בהינתן $(v, \langle \cdot \,|\, \cdot \rangle)$ ממ"פ, לכל $v \in V$ נאמר ש־u מאונך ל־v (או אורתוגולי ל־ אם אנחנו מרגישים מפונפנים) ונסמן $v \perp u$ אם $\langle u \,|\, v \rangle = 0$.

הערה 43. אם $v \perp u$ אז $u \perp v$. (כי צמוד של 0 הוא 0).

(2.2.2.1) משפט פיתגורס ותוצאותיו

משפט 110 (משפט פיתגורס). (מאוד מועיל) יהי V ממ"פ כך ש-$v, u \in V$ אורתוגונלים, אז $||v+u||^2 = ||v||^2 + ||u||^2$

הוכחה. משום שהם מאונכים מתקיים $\langle v \,|\, u \rangle = 0$. נפתח אלגברה:

$$||v+u||^2 = \langle v+u \,|\, v+u \rangle = ||v||^2 + \cancel{\langle v \,|\, u \rangle} + \cancel{\langle u \,|\, v \rangle} + \left|\left|u^2\right|\right| = ||v||^2 + ||u||^2 \quad \top$$

■

הערה 44. בתוך $\mathbb{R}^n$ הוקטורים הסטנדרטיים מאונכים אחד לשני (במכפלה הפנימית הסטנדרטית) ולכן $\langle e_i \,|\, e_j \rangle = \delta_{ij}$ כאשר δ_{ij} הדלתא של כרוניקר. באינדוקציה על משפט פיתגורס נקבל ש־:

$$v = \sum_{i=1}^{n} \alpha_i e^i \implies ||v|| = \sum_{i=1}^{n} \alpha_i^2$$

שזה בדיוק מושג הגודל בגיאומטריה אוקלידית.

הערה 45. מעל $\mathbb{R}$ מקבלים אמ"מ למשפט פיתגורס, מעל $\mathbb{C}$ לאו דווקא.
משפט 111. (אי שוויון קושי־שוורץ)

$$\forall v, u \in V\colon\ |\langle u \,|\, v\rangle| \leqslant ||u|| \cdot ||v||$$

ושוויון אמ"מ u, v ת"ל.

הוכחה. אם v או u הם 0, אז מתקקבל שוויון. טענת עזר: קיים איזשהו $\alpha \in \mathbb{F}$ כך ש־$u - \alpha v \perp v$. נסמן $v_u = \alpha v$ כאשר נמצא אותו. הוכחת טענת העזר. נחפש כזה:

$$\langle u - \alpha v \,|\, v\rangle = 0 \iff \langle v \,|\, u\rangle - \alpha\, ||v||^2 = 0 \iff \alpha = \frac{\langle v \,|\, u\rangle}{||v||^2}$$

כדרוש. (מותר לחלק בנורמה כי הם לא 0). ניעזר במשפט פיתגורס:

$$\begin{aligned}
&\begin{cases} u - \alpha v \perp v \\ u - \alpha v \perp v \end{cases} \\
\implies\ & ||u||^2 = ||(u - \alpha v + \alpha v)||^2 = \overbrace{||u - \alpha v||^2}^{\geqslant 0} + |\alpha|^2\, ||v||^2 \\
& \geqslant |\alpha| \cdot ||v||^2 = \frac{|\,\langle v \,|\, u\rangle\,|^2}{(||v||^2)^2} = ||v||^2 = \frac{|\,\langle u \,|\, v\rangle\,|^2}{||v||^2} \\
\implies\ & |\langle v \,|\, u\rangle|^2 \leqslant ||v|| \cdot ||u||
\end{aligned}$$

בפרט $||u - \alpha v||^2 = 0$ אמ"מ הם תלויים ליניארית ומכאן הכיוון השני של המשפט. ∎

הערה 46. זה לא מדויק להגיד שזה נגרר ממשפט הקוסינוסים מעל $\mathbb{R}^n$, משום שהגדרת הזווית בין u, v בגיאומטריה האוקלידית מבוצעת כדלקמן:

$$(\cos \widehat{u, v}) := \frac{\langle u \,|\, v\rangle}{||u||\,||v||} \qquad \widehat{u, v} = \arccos(\cos(\widehat{u, v}))$$

דוגמאות.

1. ממכפלה פנימית סטנדרטית:

$$\left|\sum_{i=1}^{n} \alpha_i \beta_i\right|^2 \leqslant \left(\sum_{i=1}^{n} \alpha_i^2\right)\left(\sum_{i=1}^{n} \beta_i^2\right)$$

2. נניח $f, g[0,1] \to \mathbb{R}$ רציפות אז:

$$\left|\int_0^1 f(t) g(t)\,\mathrm{d}t\right|^2 \leqslant \int_0^1 f^2(t)\,\mathrm{d}t \cdot \int_0^1 g^2(t)\,\mathrm{d}t$$

כאשר $f^2 = f \cdot f$ (לא הרכבה).

3. אי־שוויון המשולש:

$$\forall u, v \in V\colon\ ||u + v|| \leqslant ||u|| + ||v||$$

ושוויון אמ"מ אחד מהם הוא 0 או אם הם כפולה חיובית אחד של השני (לא שקול לתלויים ליניארית – יכולה להיות כפולה שלילית).

הוכחה (לאי שוויון המשולש). תזכורת: עבור $\mathcal{Z} \in \mathbb{C}$ מתקיים $|\mathcal{Z}|^2 = (\Re \mathcal{Z})^2 + (\Im \mathcal{Z})^2$

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

ושוויון אמ"מ u הוא אפס או כפולה חיובית של v. מקושי־שוורץ:

$$\leqslant ||u||^2 + 2\,||u||\,||v|| + ||v||^2 = (||u|| + ||v||)^2$$

∎

משפט 112 (משפט הקוסינוסים). בהינתן $\langle\cdot\,|\,\cdot\rangle : V\times V\to\mathbb{R}$ מ"פ, מתקיים:

$$||u+v||^2 = ||u||^2 + ||v||^2 + 2\,||u||\,||v||\cos(\widehat{u,v})$$

הוכחה. פשוט נפתח אלגברה:

$$\begin{aligned}||u||^2+||v||^2+2\,||u||\,||v||\cos(\widehat{u,v}) &= \langle u\,|\,u\rangle^2+\langle v\,|\,v\rangle^2+2\cancel{||u||\,||v||}\frac{\langle u\,|\,v\rangle}{\cancel{||u||\,||v||}}\\ &= \langle u\,|\,u\rangle+2\,\langle u\,|\,v\rangle+\langle v\,|\,v\rangle\\ &= \langle u\,|\,u\rangle+\langle u\,|\,v\rangle+\langle v\,|\,u\rangle+\langle v\,|\,v\rangle\\ &= \langle u\,|\,u+v\rangle+\langle v\,|\,u+v\rangle\\ &= \langle u+v\,|\,u+v\rangle = ||u+v||^2\end{aligned}$$

■

הערה 47. מעל המרוכבים אין לנו את הסימטריות הדרושה. במקום זאת, מתקיים $||u+v|| = ||u||^2+||v||^2+2\Re(u+iv)$.

2.2.3 ~ מרחבים ניצבים והיטלים

סימון 12. יהי $(V,\langle\cdot\,|\,\cdot\rangle)$ ממ"פ. יהיו $S,T\subseteq V$. נסמן:

א. $u\in V\colon\ (u\perp S\iff(\forall v\in S\colon u\perp v))$

ב. $S\perp T\iff\forall v\in S\ \forall u\in T\colon v\perp u$

ג. $S^\perp := \{v\in V\mid v\perp S\}$

הגדרה 79. $T^\perp$ הוא תת־המרחב הניצב ל־T.

משפט 113. יהיו $S,T\subseteq V$ קבוצות, ו־$U,W\subseteq V$ תמ"וים. אז:

א. $v\perp S$ אמ"מ $v\perp\operatorname{span}(S)$

ב. $S^\perp\subseteq V$ תמ"ו

ג. אם $S\subseteq T$ אז $T^\perp\subseteq S^\perp$

ד. $U\oplus U^\perp = V$

ה. $\left(S^\perp\right)^\perp = \operatorname{span}S$

ו. $(U+W)^\perp = U^\perp\cap W^\perp$

ז. $(U\cap W)^\perp = U^\perp+W^\perp$

הוכחה (לג').

$$\forall v\perp T\colon c\perp S\implies v\in S^\perp$$

■

הערה 48. שוויון בג' מתקיים אמ"מ $\operatorname{span}S=\operatorname{span}T$.

הגדרה 80. משפחה של וקטורים $A\subseteq V$ נקראת אורתוגונלית אם $\forall u\neq v\in V\colon u\perp v$

הערה 49. אם A משפחה אורתוגונלית וגם $0\notin A$ אז ניתן לייצור ממנה משפחה של וקטורים אורתוגונלים שהם גם וקטורי יחידה, ע"י נרמול.

הגדרה 81. משפחה של וקטורים $A\subseteq V$ נקראת אורתונורמלית, אם היא אורתוגונלית ובנוסף כל הוקטורים הם וקטורי יחידה.

הגדרה 82. יהי $U\subseteq V$ תמ"ו. יהא $v\in V$. אז ההטלה האורתוגונלית של V על U היא $p_U(v)$ הוא וקטור המקיים:

- $p_U(v)\in U$
- $v-p_U(v)\in U^\perp$

משפט 114. בסימונים לעיל, $\forall u\in U\colon\ ||v-u||\geqslant||v-p_U(v)||$ ושוויון אמ"מ $u=p_U(v)$.

הוכחה. יהי $u\in U$. ידוע $p_U(v)\in U$. אזי $u-P_u(v)\in U$. כמו כן ידוע $u\perp p_U(v)-v$. אזי בפרט $\langle u-p_U(v)\,|\,p_U(v)-v\rangle$. נתבונן ב־:

$$||u-v||^2 = ||(u-p_U(v))+(p_U(v)-v)||^2 \overset{\text{פית'}}{=} ||u-p_U(v)||^2+||v-p_U(v)||^2$$

וסה"כ $||v-u||^2\geqslant||v-p_U(v)||^2$. ושוויון אמ"מ $||u-p_U(v)||=0$ אמ"מ $u=p_U(v)$. ■

עתה נוכיח את יחידות ההטלה האורתוגונלית (קיום נוכיח בהמשך באופן קונסטרקטיבי)
משפט 115. ההטלה הניצבת, היא יחידה.

הוכחה. יהיו $p_U(v)$ וכן $p'_U(v)$ הטלות של v על U. מהטענה:

$$||v - p_U(v)|| \leqslant ||v - p'_U(v)||$$

אבל בהחלפת תפקידים מקבלים את אי־השוויון ההפוך. לכן יש שוויון נורמות. מהמשפט לעיל $p_U(v) = p'_U(v)$. ■

משפט 116. תהי $A \subseteq V$ משפחה אורתוגונלית ללא 0. אז היא בת"ל.

הוכחה. יהיו $v_1 \dots v_n \in A$ וכן $\alpha_1 \dots \alpha_n \in \mathbb{F}$, כך ש־$\sum_{i=0}^n \alpha_i v_i = 0$. יהי $i \in [n]$. אז:

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

כאשר השוויון האחרון מהיות הקבוצה אורתוגונלית. ■

משפט 117 (קיום היטל אורתוגונלי). נניח ש־$u \subseteq V$ תמ"ו. נניח U נ"ס וכן $B = (e_1 \dots e_n)$ בסיס אורתונורמלי של U (כלשהם, לא בהכרח סטנדרטיים כי גם לא בהכרח $\mathbb{F}^n$). אז

$$\forall v \in V \colon p_U(v) = \sum_{i=1}^m \langle v \,|\, e_i \rangle e_i$$

הוכחה. צ.ל. $p_U(v) \in U$ וגם $\forall u \in U \colon \langle v - p_U(v) \,|\, u \rangle = 0$ אך לגבי התנאי האחרון די להוכיח $\forall j \in [n] \colon \langle v_i p_U(v) \,|\, e_j \rangle = 0$. החלק הראשון ברור, נותר להוכיח:

$$\langle v - p_U(v) \,|\, e_j \rangle = \langle v \,|\, e_j \rangle - \langle p_u(v) \,|\, e_j \rangle =: *$$

ידוע:

$$\langle p_U(v) \,|\, e_j \rangle = \left\langle \sum_{i=1}^m \langle v_i \,|\, e_i \rangle e_i \,\middle|\, e_j \right\rangle = \sum_{i=1}^m \langle v_i \,|\, e_i \rangle \cdot \langle e_i \,|\, e_j \rangle = \sum_{i=1}^m \langle v \,|\, e_i \rangle \delta_{ij} = \langle v \,|\, e_j \rangle$$

נחזור לשוויון לעיל:

$$* = \langle v \,|\, e_j \rangle - \langle v \,|\, e_j \rangle = 0$$

כדרוש. ■

(בכך הוכחנו את קיום $p_U(v)$ לכל מ"ו נ"ס, אם נשלב את זה עם המשפט הבא)

(2.2.3.1) אלגוריתם גרהם־שמידט

משפט 118 (אלגוריתם גרהם־שמידט). תהי $(b_1 \dots b_k)$ קבוצה סדורה בת"ל של וקטורים בממ"ס V. אז בכל משפחה א"נ $(u_1 \dots u_k)$ כך ש־$\operatorname{span}(b_1 \dots b_k) = \operatorname{span}(u_1 \dots u_k)$.

מסקנות מהמשפט. לכל ממ"ס נ"ס קיים בסיס א"נ (=אורתונורמלי). יתרה מזאת, בהינתן בסיס $B = (b_1 \dots b_n)$ ניתן להופכו לבסיס א"נ $(u_1 \dots u_n)$ המקיים $\forall k \in [n] \colon \operatorname{span}(b_1 \dots b_k) = \operatorname{span}(u_1 \dots u_k)$.

הוכחה. בנייה באינדוקציה. נגדיר עבור $k = 1$ את $u_1 = b''_1$. מתקיים $\operatorname{span} u_1 = \operatorname{span} b_1$ וכן $\{u_1\}$ קבוצה א"נ. נניח שבנינו את k האיברים הראשונים, נבנה את האיבר ה־$k+1$ (כלומר את u_{k+1}). במילים אחרות, הנחנו $u_1 \dots u_k$ אורתונורמלית וגם $\operatorname{span}(u_1 \dots u_k) = \operatorname{span}(b_1 \dots b_k) = U$.

מהסעיף הקודם $p_U(b_{k+1})$ קיים, וגם $b_{k+1} - p_U(b_{k+1}) \neq 0$ מהבנייה. נגדיר $u_{k+1} = (b_{k+1} - p_U(b_{k+1}))$. בצורה מפורשת:

$$u_{k+1} = \frac{b_{k+1} - \sum_{i=1}^k \langle b_{k+1} \,|\, u_i \rangle u_i}{\left|\left| b_{k+1} - \sum_{i=1}^k \langle b_{k+1} \,|\, u_i \rangle u_i \right|\right|}$$

מהגדרת $p_U(b_{k+1})$, מתקיים $b_{k+1} - p_U(b_{k+1}) \in U^\perp$ ולכן גם $u_{k+1} \in U^\perp$ ולכן $(u_1 \dots u_{k+1})$ משפחה א"נ.

$$b_1 \dots b_k = \overbrace{\operatorname{span}(u_1 \dots u_{k+1})}^{\text{בת"ל}}$$

נשאר להוכיח ש־$b_{k+1} \in \operatorname{span}(u_1 \dots u_{k+1})$. זה מספיק משום שאז נקבל $\operatorname{span}(u_1 \dots u_{k+1}) \subseteq \operatorname{span}(b_1 \dots b_{k+1})$. אבל הם שווי ממד ולכן שווים. סה"כ:

$$b_{k+1} = ||b_{k+1} - p_U(b_{b+1})|| \cdot u_{k+1} + \sum_{i=1}^{k} \langle b_{k+1} \,|\, u_i \rangle \, u_i \implies b_{k+1} \in \operatorname{span}(u_1 \dots u_{k+1})$$

מש"ל. ■

משפט 119. יהי V מ"ו $U \subseteq V$. נניח שלכל $v \in V$ מוגדר $p_U(v)$ (בפרט כל מ"ו נ"ס). אז $p_U \colon V \to V$ המוגדרת לפי $v \mapsto p_U(v)$ העתקה לינארית.

הוכחה. יהיו $v, v' \in V, \alpha \in \mathbb{F}$. ידוע $v - p_U(v), v' - p_U(v') \in U^\perp$ ועל כן:

$$(v - p_U(v)) + \alpha(v' - p_u(v')) \in U^T \implies (v + \alpha v') - (p_U(v) + \alpha p_U(v')) \in U^T \implies (v + \alpha v') - p_U(v + \alpha v') \in U^T$$

מה מקיים היטל וקטור? ראשית ההיטל ב־U, ושנית v פחות ההיטל מאונך. הוכחנו שבהינתן היטל, הוא יחיד. והראינו ש־$(v + \alpha v') - p_U(v + \alpha v')$ מקיים את זה, ולכן אם יש וקטור אחד אז הוא יחיד, וסה"כ שווים וליניארית. ■

משפט 120. $\min_{u \in U} ||v - u|| = ||v - p_U(v)||$

בניסוח אחר: ההיטל $p_U(v)$ הוא הוקטור הכי קרוב ל־v ב־U. בתרגול צוין שזוהי דרך למצוא את הפתרון "הכי קרוב" למערכת משוואות לינארית שאין לה פתרון.

הגדרה 83. הפתרון האופטימלי למערכת משוואות $(A \mid b)$ הוא $p_{\operatorname{Col} A}(b)$ (כאשר $\operatorname{Col} A$ מ"ו העמודות).

2.2.4 ∼ צמידות ודואליות

(2.2.4.1) העתקות צמודות לעצמן, והעתקות חיוביות

הגדרה 84. V ממ"פ ו־$T \colon V \to V$ ט"ל. אז T נקראת סימטרית ($\mathbb{F} = \mathbb{R}$) או הרמטית ($\mathbb{F} = \mathbb{C}$) אם $\forall u, v \in V \langle Tu \mid v \rangle = \langle u, Tv \rangle$
באופן כללי, העתקה כזו תקרא צמודה לעצמה.

דוגמה. (המקרה בפרטי בממ"פ המשרה את הגיאו' האוקלידית) עבור $V = \mathbb{R}^n$, $\langle \cdot \mid \cdot \rangle$ מ"פ סטנדרטית, ו־ $A \in M_n(\mathbb{R})$ מתקיים $T_A \colon V \to V$ ט"ל, היא צמודה לעצמה אם: ידוע $\langle v \,|\, u \rangle = v^T u$:

$$\langle T_A v \,|\, u \rangle = (Av)^T u = v^T A^T u = \left\langle v \,\middle|\, A^T u \right\rangle$$

ז"א אם $A = A^T$ אז T_A סימטרית, כלומר A מטריצה סימטרית. גם הכיוון השני נכון: אם $T \colon V \to V$ סימטרית אז ע"י בחירת בסיס נקבל $[T]_B^B$ גם היא סימטרית.

דוגמה נוספת (בדמות משפט)

משפט 121. ההעתקה $v \mapsto p_U(v)$ עבור U תמ"ו כלשהו, היא ההיטל, צמודה לעצמה.

הוכחה. יהי V תמ"ו. ניעזר בעובדה לכל U תמ"ו ו־$u \in V$ ניתן לפרק את u לפי $u = p_U(u) + p_{U^\perp}(u)$. עוד נבחין שמכיוון ש־$\operatorname{Im} p_{U^\perp} = U^\perp \wedge \operatorname{Im} p_U = U$ אז לכל $v, u \in V$ מתקיים $p_U(u)$ ניצב ל־$p_{U^\perp}(v)$.

$$\begin{aligned}
\langle p_U(v) \,|\, u \rangle &= \langle p_U(v) \,|\, p_U(u) + p_{U^\perp}(u) \rangle \\
&= \langle p_U(v) \,|\, p_U(u) \rangle + \underbrace{\langle p_U(v) \,|\, p_{U^\perp}(u) \rangle}_{0} \\
&= \langle p_U(v) \,|\, p_U(u) \rangle + \overbrace{\langle p_{U^\perp}(v) \,|\, p_U(u) \rangle} \\
&= \langle p_U(v) + p_{U^\perp}(v) \,|\, p_U(u) \rangle \\
&= \langle v \,|\, p_U(u) \rangle \quad \top
\end{aligned}$$

■

משפט 122. נניח V ממ"פ מעל $\mathbb{C}$. אז T הרמיטית אמ"מ $\forall v \in V \colon \ \langle Tv \,|\, v \rangle \in \mathbb{R}$

משפט 123. יהיו $T, S \colon V \to V$ צמודות לעצמן. אז:

1. $\alpha T, T+S$ צמודות לעצמן.
2. המכפלה $S\circ T$ צמודה לעצמה אמ"מ $ST=TS$
3. אם p פולינום מעל $\mathbb{F}$ אז $p(T)$ צמודה לעצמה.

קל לראות ש־$3 \impliedby 1+2$. 1 נובע ישירות מהגדרה. 1 טרוויאלי. נוכיח את 2.

הוכחה ל־2. נניח $S\circ T$ צמודה לעצמה. בהנחות המשפט ידוע S,T צמודות לעצמן. נקבל:

$$\langle (S\circ T)v \,|\, u\rangle = \langle v \,|\, STu\rangle = \langle Sv \,|\, Tu\rangle = \langle TSv \,|\, u\rangle \implies \langle (ST-TS)v \,|\, u\rangle = 0 \quad \forall v,u$$

נסיק:

$$\implies \forall v \, \langle (ST-TS)v \,|\, (ST-TS)v\rangle = 0 \implies (ST-TS)v=0 \implies STv=TSv \implies \top$$

מהכיוון השני, אם $ST=TS$ אז מהיות S,T צמודות לעצמן:

$$\langle STv \,|\, u\rangle = \langle S(Tv) \,|\, u\rangle = \langle Tv \,|\, Su\rangle = \langle v \,|\, TSu\rangle = \langle v \,|\, STu\rangle$$

■

הגדרה 85. $T\colon V\to V$ תקרא חיובית/אי־שלילית/שלילית/אי־חיובית אם לכל $v\in V$:

חיובית: $\langle Tv \,|\, v\rangle > 0$ אי־שלילית: $\langle Tv \,|\, v\rangle \geqslant 0$
שלילית: $\langle Tv \,|\, v\rangle < 0$ אי־חיובית: $\langle Tv \,|\, v\rangle \leqslant 0$

הערה 50. באופן כללי, אין קשר הדוק בין ההגדרה הזו לבין הגדרת חיוביות של תבניות בי־לינאריות. המושגים יתקשרו אחד לשני בהמשך רק בהקשר של העתקות צמודות לעצמן.
משפט 124. אם T חיובית/שלילית, אז היא הפיכה.

הוכחה. נניח ש־T לא הפיכה, נניח בשלילה שהיא חיובית. קיים $0\neq v\in V$. אז $v\in\ker T$, אז $\langle Tv \,|\, v\rangle = \langle 0 \,|\, v\rangle = 0$, בסתירה לכך ש־$T$ חיובית. ■

משפט 125. נניח ש־S צמודה לעצמה, אז S^2 צמודה לעצמה ואי־שלילית.

הוכחה. ממשפט קודם S^2 צמודה לעצמה. נוכיח אי־שלילית:

$$\forall 0\neq v\in V\colon \; \left\langle S^2v \,|\, v\right\rangle = \langle Sv \,|\, Sv\rangle = ||Sv||^2 \geqslant 0$$

■

הגדרה 86. פולינום $p\in\mathbb{R}[x]$ יקרא חיובי אם $\forall x\in\mathbb{R}\colon p(x)>0$.
משפט 126. נניח $p(x)\in\mathbb{R}[x]$ חיובי, ו־$T\colon V\to V$ צמודה לעצמה, אז $p(T)$ חיובית גם־כן, וצמודה לעצמה.
למה 10. אם $p\in\mathbb{R}[x]$ אי־שלילי, אז קיימים $g_1\dots g_k\in\mathbb{R}[x]$ וכן $c\in\mathbb{R}\geqslant 0$ כך ש־$p(x)=\sum_{i=1}^k g_i^2(x)+c$, ו־$c\neq 0$ אמ"מ p חיובי.

רעיון להוכחת הלמה: מעל $\mathbb{C}$ זה מתפרק, ונוכל לכתוב $p(x)=a_n\prod_{j=1}^s(x-i\alpha_j)(x+i\alpha_j)$ (מעל $\mathbb{R}$ כל פולינום מתפרק לגורמים ריבועיים, ואם כל שורשיו מרוכבים, כל גורמיו ריבועיים). הרעיון הוא להוכיח את הטענה ש־$g^2h\bar{h}=g_1^2+g_2^2$.

הוכחה (של המשפט, לא של הלמה). יהי $0\neq v\in V$. אז:

$$\langle p(T)v \,|\, v\rangle = \underbrace{\left\langle \sum_{i=1}^k g_i^2(T)v \,\middle|\, v\right\rangle}_{\sum_{i=1}^k\langle g_i^2(T)v \,|\, v\rangle \geqslant 0} + \overbrace{c\,\langle v \,|\, v\rangle}^{c||v||^2>0} \geqslant 0$$

■

מסקנה 21. אם $T\colon V\to V$ צמודה לעצמה ו־$p(x)\in\mathbb{R}[x]$ פולינום חיובי, אז $p(T)$ הפיכה.
משפט 127. נניח ש־$T\colon V\to V$ סימטרית (צמודה לעצמה מעל $\mathbb{R}$/המטריצה המייצגת סימטרית) ויהי $m_T(x)$ הפולינום המינימלי של T. אז m_T מתפרק לגורמים לינאריים. בנוסף, הם שונים זה מזה.

הוכחה. נניח בשלילה קיום $p \mid m_T$ ו־$\deg p \geqslant 2$, p אי־פריק. בה"כ נניח ש־p חיובי (אין לו שורש ב־$\mathbb{R}$, לכן נמצא כולו מעל/מתחת לציר ה־x). אז אפשר לכתוב את m_T כ־$m_T = p \cdot g$ כלשהו. ידוע $p(T) \neq 0$ כי m_T מינימלי מדרגה גבוהה יותר. אזי:

$$0 = m_T(T) = \underbrace{p(T)}_{\neq 0} \cdot g(T) \implies g(T) = 0$$

בסתירה למינימליות של m_T. סה"כ m_T אכן מתפרק לגורמים לינארים. עתה יש להראות שהגורמים הלינארים שלו זרים. נניח ש־T סימטרית. ניעזר בלמה המופיע מיד אחרי ההוכחה הזו. נניח בשלילה שהם לא כולם שונים, אז $m_T(x) = (x-\lambda)^2 g(x)$ ואז:

$$0 = m_T(T)v = (T-\lambda I)^2 g(T) \implies \omega = g(T)v,\ (T-\lambda I)^2\omega = 0$$

לכן בפרט $(T-\lambda I)\omega = 0$ מהסעיף הקודם. סה"כ $\forall v \in V\colon (T-\lambda I)g(T) = 0$ וסתירה למינימליות. ■

מסקנה 22. T סימטרית היא לכסינה.

זכרו מסקנה זו להמשך. היא תהפוך להיות להגיונית כאשר נדבר על המשפט הספקטרלי מעל $\mathbb{R}$.

למה 11. נניח T סמטרית ו־$\lambda \in \mathbb{R}$, אם $(T-\lambda I)^2 = 0$ אז $T - \lambda I = 0$.

הוכחה. ידוע:

$$\forall v\colon 0 = \left\langle (T-\lambda I)^2 v \,\middle|\, v\right\rangle = \langle (T-\lambda I)v \,|\, (T-\lambda I)v\rangle = ||(T-\lambda I)v||^2 \implies (T-\lambda I)v = 0$$ ■

משפט 128. אם V ממ"פ ו־$T\colon V \to V$ ט"ל צמודה לעצמה, אז הע"ע של T ממשיים.

הוכחה. יהי $0 \neq v \in V$ ו"ע של T שמתאים לע"ע λ. נחשב:

$$\lambda v\, ||v||^2 = \langle \lambda v \,|\, v\rangle = \langle Tv \,|\, v\rangle = \langle v \,|\, Tv\rangle = \langle v \,|\, \lambda v\rangle = \overline{\lambda}\, ||v||$$

ידוע $v \neq 0$ ולכן $||v|| \neq 0$ ונסיק $\lambda v = \bar{\lambda}$ ולכן $\lambda \in \mathbb{R}$. ■

משפט 129. אם V ממ"פ ו־$T\colon V \to V$ ט"ל צמודה לעצמה, אז כל זוג $0 \neq u, v \in V$ ע"ע שונים, המתאימים לערכים $\alpha, \beta \in \mathbb{C}$, מאונכים זה לזה.

הוכחה. למעשה, מהטענה הקודמת $\alpha, \beta \in \mathbb{R}$. כאן $Tu = \alpha u,\ Tv = \beta v$, כאשר $\alpha = \beta$. נחשב:

$$\alpha \langle u \,|\, v\rangle = \langle \alpha u \,|\, v\rangle = \langle Tu \,|\, v\rangle = \langle u \,|\, Tv\rangle = \langle u \,|\, \beta v\rangle = \bar{\beta}\langle u \,|\, v\rangle = \beta \langle u \,|\, v\rangle$$

בגלל ש־$\beta \in \mathbb{R}$ מתקיים $\beta = \bar{\beta}$. ולכן $(\alpha - \beta)\langle v \,|\, u\rangle = 0$ מהעברת אגף וסה"כ $\langle u \,|\, v\rangle = 0$ ואכן $u \perp v$. ■

הערה 51. בחלק הבא יש שימוש קל במרחבים דואלים. בעבור סטונדטים שבעבורם מרחבים דואלים לא נכלל כחלק מלינארית 1א, אני ממליץ לקרוא את החלק הראשון של מרחבים דואלים בסוף הסיכום.

משפט 130 (משפט ריס)**.** יהי V ממ"פ סופי ויהי $\varphi \in V^*$. אז קיים ויחיד וקטור $u \in V$ שמקיים $\forall v \in V\colon \varphi(v) = \langle v \,|\, u\rangle$.

הוכחה.

קיום. יהי $B = (b_i)_{i=1}^n$ בסיס אורתונורמלי של V (הוכחנו קיום בהרצאות קודמות). נסמן $u = \sum_{i=1}^n \overline{\varphi(b_i)} b_i$. בכדי להראות $\forall v \in V\colon \varphi(v) = \langle v \,|\, u\rangle$ מספיק להראות תכונה זו לאברי הבסיס B, כלומר נראה ש־$\forall 1 \leqslant j \leqslant n\colon \varphi(b_j) = \langle b_j \,|\, u\rangle$. ואכן:

$$\langle b_j \,|\, u\rangle = \left\langle b_j \,\middle|\, \sum_{i=0}^n \overline{\varphi(b_i)} b_i \right\rangle = \sum_{i=1}^n \underbrace{\overline{\overline{\varphi(b_i)}}}_{b_i} \underbrace{\langle b_j \,|\, b_i\rangle}_{\delta_{ij}} = b_j \quad \top$$

יחידות: אם קיים וקטור נוסף שעבורו $\forall v \in V\colon \varphi(v) = \langle v \,|\, w\rangle$ אז בפרט עבור $v = u - w$ נקבל:

$$\begin{aligned} &\varphi(v) = \langle v \,|\, w\rangle = \langle v \,|\, u\rangle \\ &\implies \langle v \,|\, u - w\rangle = 0 \\ &\implies 0 = \langle u - w \,|\, u - w\rangle = ||v - w||^2 = 0 \\ &\implies v - w = 0 \\ &\implies v = w \end{aligned}$$

סה"כ הוכחנו קיום ויחידות כדרוש. ■

הגדרה 87. $A \in M_n(\mathbb{F})$ תקרא חיובית/אי־שלילית/שלילית/אי־חיובית אם לכל $v \in \mathbb{F}^n$:

חיובית: $\langle Av \,|\, v\rangle > 0$ אי־שלילית: $\langle Av \,|\, v\rangle \geqslant 0$
שלילית: $\langle Av \,|\, v\rangle < 0$ אי־חיובית: $\langle Av \,|\, v\rangle \leqslant 0$

כאשר $\langle \cdot \,|\, \cdot\rangle$ המ"פ הסטנדטית מעל $\mathbb{F}^n$.

הערה 52. זהו אינה הגדרה נפוצה. לרוב מדברים רק על מטריצה מוגדרת חיובית.

משפט 131. נניח ש־$A = A^* \in M_n(\mathbb{F})$, אז התנאים הבאים שקולים (קיצור מוכר: TFAE, the following are equaivlent):

1. $T_A \colon \mathbb{F}^n \to \mathbb{F}^n$ חיובית.
2. לכל $T \colon V \to V$ ולכל בסיס אורתונורמלי B כך ש־$A = [T]_B$, A חיובית.
3. קיימים $T \colon V \to V$ חיובית/אי שלילית ו־B בסיס אורתונורמלי, כך ש־$A = [T]_B$ חיובית.
4. הע"ע של A חיוביים (הם בהכרח ממשיים כי היא צמודה לעצמה).

הוכחה. מספיק לטעון זאת כדי להוכיח את השקילויות של 1, 2, 3:

$$\langle Tv \,|\, v\rangle_V = \langle [Tv]_B \,|\, [v]_B\rangle_{\mathbb{F}^n} = \langle A[v]_B \,|\, [v]_B\rangle_{\mathbb{F}^n}$$

בשביל $2 \to 1$, ידוע שהאגף הימני גדול מ־0 מההנחה שהיא חיובית/אי שלילית על $\mathbb{F}^n$, ומכאן הראנו שהמיוצגת בכל בסיס חיובית כדרוש. בשביל $1 \to 3$, נפעיל טיעונים דומים מהאגף השמאלי במקום. הגרירה $3 \to 2$ ברורה. סה"כ הראינו את $3 \leftrightarrow 2 \leftrightarrow 1$.

עתה נוכיח שקילות בין 1 ל־4.

$1 \to 4$ יהי $\lambda \in \mathbb{R}$ ע"ע של A (נוכל להניח λ ממשי כי A צמודה לעצמה)

$$\langle Av \,|\, v\rangle = \lambda \left|\left|v\right|\right|^2 > 0 \implies \lambda > 0$$

$4 \to 1$ יהי $B = (v_1 \ldots v_n)$ בסיס א"נ של ו"ע, ויהי $V \ni v = \sum \alpha_i v_i$. נקבל:

$$\langle T_A v \,|\, v\rangle = \langle Av \,|\, v\rangle = \left\langle A\sum_{i=1}^{n} \alpha_i v_i \,\middle|\, \sum_{i=1}^{n} \alpha_i v_i \right\rangle = \sum \lambda_i |\alpha_i|^2 > 0$$

■

הערה 53. ההוכחה עובדת באותה הצורה בעבור A אי־שלילית, שלילית או אי־חיובית, ולמען נוחות הוכחנו בעבור העתקה חיובית בלבד.

(2.2.4.2) ההעתקה הצמודה

משפט 132. יהי V ממ"פ מנ"ס ותהי $T \colon V \to V$ לינארית. אז קיימת ויחידה $T^* \colon V \to V$ ומקיימת $\forall u, v \in V \colon \langle Tu \,|\, v\rangle = \langle u \,|\, T^* v\rangle$.

הוכחה. לכל $v \in V$, נתבונן בפונקציונל הלינארי $\varphi_V \in V^*$ המוגדר ע"י $\forall u \in V \colon \varphi_V(u) = \langle Tu \,|\, v\rangle$. ממשפט ריס קיים ויחיד $T^* v \in V$ שעבורו $\forall u \in V \colon \langle Tu \,|\, v\rangle = \varphi_V(u) = \langle u \,|\, T^* v\rangle$. כלומר, ההעתקה $T^* \colon V \to V$, קיימת ויחידה, ונותר להראות שהיא לינארית. עבור $v, w \in V$ ועבור $\alpha, \beta \in \mathbb{F}$ מתקיים:

$$\begin{aligned}
\forall u \in V \colon \quad & \langle u \,|\, T^*(\alpha v + \beta w)\rangle \\
&= \langle Tu \,|\, \alpha v + \beta w\rangle \\
&= \bar{\alpha} \langle Tu \,|\, v\rangle + \bar{\beta} \langle Tu \,|\, v\rangle \\
&= \bar{\alpha} \langle u \,|\, T^* v\rangle + \bar{\beta} \langle u \,|\, T^* w\rangle \\
&= \langle u \,|\, \alpha T^* u + \beta T^* w\rangle
\end{aligned}$$

מסך נסיק ש־$T^*(\alpha v + \beta w) = \alpha T^* u + \beta T^* w$ מנימוקים דומים. ■

הגדרה 88. ההעתקה T^* לעיל נקראת ההעתקה הצמודה ל־T.

דוגמאות. מעל $\mathbb{C}^n$, עם המ"פ הסטנדרטי, נגדיר ט"ל $T_A \colon \mathbb{C}^n \to \mathbb{C}^n$ עבור $A \in M_n(\mathbb{C})$ מוגדרת ע"י $T_A(x) = Ax$. אז:

$$\forall x, y \in \mathbb{C}^n \colon \langle T_A(x) \,|\, y\rangle = \langle Ax \,|\, y\rangle = \overline{(Ax)^T} \cdot y = \overline{x^T} \cdot \overline{A^T} y \cdot = \overline{x^T} T_{\overline{A^T}}(y) = \left\langle x \,\middle|\, T_{\overline{A^T}} y \right\rangle$$

כלומר, $(T_A)^* = T_{A^*}$ כאשר $A^* = \overline{A^T}$, וקראנו לה המטריצה הצמודה.

נבחין שהעתקה נקראת צמודה לעצמה אמ"מ $T^* = T$.

עוד נבחין שעבור העתקה הסיבוב $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ בזווית θ, מתקיים ש־T^* היא הסיבוב ב־$-\theta$, וכן היא גם ההופכית לה. כלומר $(T_\theta)^* = T_{-\theta} = (T_\theta)^{-1}$. זו תכונה מאוד מועילה וגם נמציא לה שם במועד מאוחר יותר.

משפט 133 (תכונות ההעתקה הצמודה). יהי V ממ"פ ותהיינה $T, S\colon V \to V$ זוג העתקות ליניאריות. נבחין ש־:

(א) $$(T^*)^* = T$$

(ב) $$(T \circ S)^* = S^* \circ T^*$$

(ג) $$(T + S)^* = T^* + S^*$$

(ד) $$\forall \lambda \in \mathbb{F}\colon (\lambda T)^* = \bar{\lambda}(T^*)$$

הוכחה.

א $$\forall u, v \in V\colon \langle T^*u \,|\, v\rangle = \overline{\langle v \,|\, T^*u\rangle} = \overline{\langle Tv \,|\, u\rangle} = \langle u \,|\, Tv\rangle \implies (T^*)^* = T$$

ב $$\langle (T \circ S)u \,|\, v\rangle = \langle Su \,|\, T^*v\rangle = \langle u \,|\, S^*T^*\rangle \implies (TS)^* = T^*S^*$$

ג $$\langle (T+S)u \,|\, v\rangle = \langle Tu \,|\, v\rangle + \langle Su \,|\, v\rangle = \langle u \,|\, T^*v\rangle + \langle u \,|\, S^*v\rangle = \langle u \,|\, T^*v + S^*v\rangle$$

ד $$\langle (\lambda T)u \,|\, v\rangle = \lambda \langle Tu \,|\, v\rangle = \lambda \langle u \,|\, Tv\rangle = \langle u \,|\, (\bar{\lambda}T)v\rangle$$ ■

סימון 13. העתקה צמודה לעצמה לעיתים קרובות (בעיקר בפיזיקה) מסמנים ב־$T^\dagger$. באופן דומה גם מטריצה צמודה מסמנים ב־$A^\dagger$.

משפט 134. בהינתן B **אורתונורמלי** של V אז $[T^*]_B = [T]_B^*$ (שימו לב: האחד צמוד מטריציוני, והשני העתקה צמודה)

משפט 135. T צמודה לעצמה אמ"מ $\forall v \in V\colon \langle Tv \,|\, v\rangle \in \mathbb{R}$.

הוכחה.

$$\begin{aligned}
& T \text{ צמודה לעצמה} \\
\iff & T = T^* \iff T - T^* = 0 \\
\iff & \forall v \in V\colon \langle (T - T^*)v \,|\, v\rangle = 0 \\
\iff & \forall v \in V\colon \langle Tv \,|\, v\rangle - \langle T^*v \,|\, v\rangle = 0 \\
\iff & \forall v \in V\colon \langle Tv \,|\, v\rangle - \overline{\langle Tv \,|\, v\rangle} = 0 \\
\iff & \forall v \in V\colon \Re(\langle Tv \,|\, v\rangle) + \Im(\langle Tv \,|\, v\rangle) - \Re(\langle Tv \,|\, v\rangle) + \Im(\langle Tv \,|\, v\rangle) = 0 \\
\iff & \forall v \in V\colon 2\Im(\langle Tv \,|\, v\rangle) = 0 \iff \forall v \in V\colon \Im(\langle Tv \,|\, v\rangle) = 0 \iff \forall v \in V\colon \langle Tv \,|\, v\rangle \in \mathbb{R} \quad \top
\end{aligned}$$ ■

משפט 136. התנאים הבאים שקולים:

1. T צמודה לעצמה.
2. לכל בסיס $\mathcal{E}$ אורתונורמלי מתקיים $[T]_\mathcal{E} = [T]_\mathcal{E}^*$.
3. ל־T קיים בסיס אורתונורמלי $\mathcal{E}$ כך ש־$[T]_\mathcal{E} = [T]_\mathcal{E}^*$.

הוכחה. את השקילות $2 \leftrightarrow 1$ הוכחנו במשפט הקודם. עתה נראה $1 \to 3 \to 2 \to 1$. ניעזר בכך שידוע שלכל $\mathcal{E}$ אורתוגונלי מתקיים $[T^*]_\mathcal{E} = [T]_\mathcal{E}^*$.

$1 \to 2$ ידוע שלכל $\mathcal{E}$ אורתוגונלי מתקיים $[T^*]_\mathcal{E} = [T]_\mathcal{E}^*$ וכן $T = T^*$ מהנתון, כלומר $[T]_\mathcal{E} = [T^*]_\mathcal{E}$. סה"כ $[T]_\mathcal{E} = [T^*]_\mathcal{E}$ לכל $\mathcal{E}$ אורתוגונלי.

$2 \to 3$ מאלגו' גראם־שמידט, בהכרח קיים $\mathcal{E}$ אורתוגונלי כלשהו. לכן מ־4 בפרט הוא מקיים $[T]_\mathcal{E} = [T]_\mathcal{E}^*$ כדרוש.

$3 \to 1$ ידוע קיום $\mathcal{E}$ כך ש־$[T]_\mathcal{E} = [T]_\mathcal{E}^* = [T^*]_\mathcal{E}$. לכן מהיות $[\cdots]_\mathcal{E}$ העתקת הקורדינאטות לבסיס איזומורפיזם, בהכרח $T = T^*$ כדרוש. ■

הערה 54. כאן מתקשר השם "סימטרי" לאופרטור סימטרי, כי למעשה משום שמטריצה סימטרית מקיימת $A = A^T = \overline{A^T} = A^*$ מעל $\mathbb{R}$, העתקה היא סימטרית אמ"מ המייצגת (תחת בסיס אורתוגונלי) סימטרית.

המשך בעמוד הבא

2.3 Decompositions

אינטואציה. ראינו לכסון – ניסיון למצוא מטריצה אלכסונית דומה. ה"בעיה" בלכסון, ובמטריצות מעבר באופן כללי, זה שהן לא משמרת את כל תכונות המרחב הוקטורי – הן לא משמרות את הנורמה. לכן, בחלק הזה של הקורס, נגדיר "מטריצת מעבר מיוחד" שמשמרת נורמה, כלומר $||Av||=||v||$. בניסוח אחר, ננסה למצוא בסיס אורתונורמלי שבו הלכסון מתקיים. על התנאי הזה בדיוק נלמד כאשר נדבר על משפט הפירוק הספקטרלי.

לצערנו, בדיוק כמו שלא יכולנו ללכסן כל מטריצה, נוכל ללכסן אורתונורמלית עוד פחות מטריצות. לכן, לאחר מכן נעסוק במושג "התאמת המטריצות" – בהינתן מטריצה A, נאמר שהיא מתאימה למטריצה B אמ"מ קיימות מטריצות מעבר בסיסים P,Q כך ש־$A=PBQ^{-1}$. לכאורה, זה נראה תנאי חלש נורא. ואכן, אפשר להראות שהוא באמת חלש, ו־A מתאימה ל־B אמ"מ $\operatorname{rank} A=\operatorname{rank} B$. אך מסתבר שאם נגביל את B להיות אלכסונית, וממטריצות המעבר שלנו נדרוש לא לשנות את הנורמה (דהיינו $||Bv||=||Av||=||v||$), אז מצאנו פירוק מאוד משמעותי. לפירוק כזה נקרא "פירוק לערכים סינגולריים", והוא מאפשר לנו להגיד הרבה על ה"גדלים" שהמטריצה משנה, כי אנחנו מתאימים אותה למטריצה אלכסונית (שמאוד נוח לעבוד איתה) בעוד מטריצות המעבר לא באמת משפיעות על הנורמות (הגדלים) של הדברים. על המצב הזה נדבר בהקשר של פירוק SVD, המקרה המוכלל של פירוק ספקטרלי.

ישנו פירוק נוסף, הפירוק הפולארי. מסתבר, שבאופן דומה לכך שאפשר לפרק וקטור פולארית (לזווית ולגודל) מעל ממ"פים, אפשר לפרק העתקה "פולארית" – להרכבת העתקות, כך שהראשונה "חיובית" ומשנה את הגדלים, והשנייה רק מסובבת (או באופן שקול, לא משנה גודל).

2.3.1 ~ המשפט הספקטרלי להעתקות

(2.3.1.1) ניסוח המשפט הספקטרלי להעתקות צמודות לעצמן

משפט 137 (המשפט הספקטרלי להעתקה לינארית צמודה לעצמה). יהי V ממ"פ ממימד סופי, ותהי $T\colon V\to V$ ט"ל צמודה לעצמה. אז קיים ל־V בסיס אורתוגונלי (או אורתונורמלי) שמורכב מו"ע של T.

הוכחה. יהי $m_T(x)$ הפולינום המינימלי של T. נציג $m_T(x)=\prod_{i=1}^m (x-\lambda_i)^{d_i}$ כאשר $\lambda_1\dots\lambda_n$ הע"ע השונים של T. מהטענה הקודמת $\lambda_1\dots\lambda_n\in\mathbb{R}$. [הערה: התמשתנו במשפט היסודי של האלגברה מעל המרוכבים, והסקנו פירוק מעל $\mathbb{R}$]. בכדי להראות ש־T לכסינה, עלינו להוכיח ש־$\forall 1\leqslant i\leqslant m\colon d_i=1$. נניח בשלילה שזה לא מתקיים, אזי $m_T(x)=(x-\lambda)^2\cdot p(x)$ כאשר λ ע"ע כלשהו. כעת, לכל $v\in V$ מתקיים מהיות T צמודה לעצמה (כלומר גם $p(T)$ צמוד לעצמו):

$$0=\underbrace{m_T(T)}_{=0}(v)\implies 0=\langle m_T(T)(v)\,|\,p(T)(v)\rangle=\langle (T-\lambda I)(p(T)v)\,|\,p(T)v\rangle=$$

$$\langle (T-\lambda I)(p(T)v)\,|\,(T-\lambda I)(p(T)v)\rangle=\left|\left|(T-\lambda I)^2(p(T)v)\right|\right|^2=0$$

ולכן $\forall v\in V\colon (T-\lambda I)(p(T)v)=0$ ולכן $((x-\lambda)(p(x))(T)=0$ בסתירה למינימליות של $m_T(x)$. נאמר, מכפלת גורמים לינארים שונים, ולכן T לכסינה, ונוכל לפרק את V באמצעות:

$$V=\bigoplus_{i=1}^m \ker(T-\lambda_i I)$$

והמרחבים העצמיים הללו אורתוגונליים זה לזה, מטענה שהוכחנו. נבנה בסיס B_i של $\ker(T-\lambda_i)$ וסה"כ $\bigcup_{i=1}^m B_i$ בסיס אורתוגונלי מלכסן של T. ∎

משפט 138. יהי V נ"ס מעל $\mathbb{R}$ ותהי $T\colon V\to V$ ט"ל. אז T סימטרית אמ"מ קיים לה בסיס אורתוגונלי מלכסן.

הוכחה. מכיוון אחד, הוכחנו באמצעות המשפט הספקטרלי להעתקות לינאריות צמודות לעצמן. מהכיוון השני, נניח שקיים ל־V בסיס אורתוגונלי מלכסן של ו"ע של T. ננרמל לבסיס אורתונורמלי $B=(b_i)_{i=1}^n$ של ו"ע של T, המתאימים ל־$\lambda_1\dots\lambda_n$. עבור $v,u\in V$, נציג:

$$u=\sum_{i=1}^n \alpha_i b_i,\ \ v=\sum_{i=1}^n \beta_i b_i$$

ונחשב:

$$\langle Tu\,|\,v\rangle=\left\langle T\left(\sum_{i=1}^m \alpha_i b_i\right)\,\middle|\,\sum_{i=1}^m \beta_i b_i\right\rangle=\sum_{i=1}^n\sum_{j=1}^n \alpha_i\beta_j\left\langle Tb_i\,|\,b_j\right\rangle=\sum_{i=1}^n\sum_{j=1}^n \alpha_i\beta_j\lambda_i\underbrace{\langle b_i\,|\,b_j\rangle}_{\delta_{ij}}=\sum_{i=0}^n \alpha_i\beta_i\lambda_i$$

מהצד השני:

$$\langle u\,|\,Tv\rangle = \left\langle \sum_{i=0}^{n}\alpha_i b_i \,\middle|\, T\left(\sum_{i=0}^{n}\beta_i b_i\right)\right\rangle = \sum_{i=0}^{n}\sum_{j=1}^{n}\alpha_i\beta_i \langle b_i\,|\,Tb_j\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\beta_j\lambda_j \underbrace{\langle b_i\,|\,b_j\rangle}_{\delta_{ij}} = \sum_{i=1}^{n}\alpha_i\beta_j\lambda_i$$

מטרנזטיביות שוויון, הראינו ש־$\langle Tu\,|\,v\rangle = \langle u\,|\,Tv\rangle$ ולכן T צמודה לעצמה. השוויון לדלתא של כקוניקר נכונה מאורתוגונליות איברי הבסיס, והבי־לינאריות כי אנחנו מעל הממשיים. המשפט לא נכון מעל מהרוכבים. ■

הוכחה שהמשפט מתקיים בהכרח מעל המרוכבים: ההעתקה $T(x) = ix$ היא העתקה סקלרית לינארית, לכן כל וקטור הוא ו"ע וכל בסיס מלכסן, בסיס אורתונורמלי כלשהו יהיה בסיס מלכסן על אף שההעתקה לא צמודה לעצמה, אלא אנטי־הרמיטית.

(2.3.1.2) ניסוח המשפט הספקטרלי בעבור העתקה כללית

המטרה: להבין לאילו העתקות בדיוק מתקיים המשפט הספקטרלי. מעל הממשיים, הבנו שאילו העתקות צמודות לעצמן. אז מה קורה מעל המרוכבים?

משפט 139. יהי V ממ"פ נ"ס ותהי $T\colon V\to V$ לינאריות. אם $B=(b_i)_{i=1}^{n}$ בסיס אורתוגונלי לו"ע של T, אז $\forall 1\leqslant i\leqslant n$ ו"ע של ההעתקה הצמודה.

כלומר: אם מתקיים המשפט הספקטרלי, אז הבסיס שמלכסן אורתוגונלית את T מלכסן אורתוגונלית את הצמודה.

הוכחה. יהי $i\in[n]$ ונסמן בעבורו את λ_i הע"ע המתאים לו"ע b_i. עבור $i\neq j\in[n]$ נחשב את $\langle b_i\,|\,T^*b_j\rangle$:

$$\langle b_i\,|\,T^*b_j\rangle = \overline{\langle Tb_i\,|\,b_j\rangle} = \overline{\langle \lambda_i b_i\,|\,b_j\rangle} = \lambda_i\langle b_i\,|\,b_j\rangle = 0$$

לכן $T^*b_j\in(\operatorname{span}\{b_i\}_{i=1}^{n})^{\perp}\overset{!}{=}\operatorname{span}\{b_j\}$. משיקולי ממדים, הפריסה מממד $n-1$ ולכן המשלים האורתוגונלי שלו מממד 1 ולכן השוויון. סה"כ $T^*b_j\in\operatorname{span}\{b_j\}$ ולכן b_j ו"ע של T^* כדרוש. ■

מסקנה. אאם V ממ"פ נ"ס ו־$T\colon V\to V$ ט"ל עם בסיס מלכסן אורתוגונלי, אז T,T^* מתחלפות כלומר $T^*T = TT^*$.

הוכחה. לפי הטענה הקודמת כל b_i הוא ו"ע משותף ל־T ול־T^*, ולכן:

$$TT^*(b_i) = T(T^*(b_i)) = \beta_i T^*(b_i) = \beta_i\delta_i b_i = \alpha_i\beta_i b_i = \alpha_i T^(b_i) = T^*T(b_i)$$

העתקה מוגדרת לפי מה שהיא עושה לבסיס ולכן $TT^* = T^*T$. ■

הגדרה 89. העתקה כזו המקיימת $AA^* = A^*A$ נקראת נורמלית (או "נורמאלית" בעברית של שנות ה־60).

למה 12. T היא העתקה נורמלית אמ"מ $\forall v\in V\colon\ ||Tv|| = ||T^*v||$.

הוכחה.

$\Longrightarrow$ אם T נורמלית, אז $TT^* = T^*T$ ואז:

$$||Tv||^2 = \langle Tv\,|\,Tv\rangle = \langle T^*Tv\,|\,v\rangle = \langle TT^*v\,|\,v\rangle = \langle T^*v\,|\,T^*v\rangle = ||T^*v||^2$$

$\Longleftarrow$ אם $\forall v\in V\colon\ ||Tv|| = ||T^*v||$, אז:

$$\forall v\in V\colon\quad 0 = ||Tv||^2 - ||T^*v||^2 = \langle Tv\,|\,Tv\rangle - \langle T^*v\,|\,T^*v\rangle = \langle T^*Tv\,|\,v\rangle - \langle TT^*v\,|\,v\rangle = \langle (T^*T - TT^*)v\,|\,v\rangle$$

נבחין ש־:

$$(T^*T - TT^*)^* = T^*(T^*)^* - (T^*)^*T^* = T^*T - TT^* =: \varphi$$

כלומר, φ צמודה לעצמה. ממשפט שהוכחנו $\varphi = 0$, כלומר $T^*T = TT^*$ כדרוש. ■

מעתה ואילך, ננסה להראות שכל העתקה נורמלית מקיימת את התנאי של המשפט הספקטרלי (כלומר ניתן ללכסנה אורתוגונלית)

משפט 140 (המשפט הספקרלי). יהי V ממ"פ נוצר סופית מעל $\mathbb{C}$, ותהי $T\colon V\to V$ לינארית. אז קיים בסיס אורתוגונלי של ו"ע של T אמ"מ T נורמלית.

למה 13. יהי V ממ"פ ותהיינה $S_1,S_2\colon V\to V$ זוג ט"ל צמודות ולעצמן ומתחלפות (כלומר $S_1S_2 = S_2S_1$). אז קיים בסיס אורתוגונלי של V שמורכב מו"עים משופים ל־S_1 ול־S_2.

הוכחה. ידוע ש־S_1 צמודה לעצמה, לכן לפי המשפט הספקטרלי להעתקות צמודות לעצמן (לא מעגלי כי הוכח בנפרד בהרצאה הקודמת), קיים לה לכסון אורתוגונלי ובפרט S_1 לכסינה. נציג את V כ־$V = \bigoplus_{i=1}^m \ker(S_1 - \lambda_i I)$, כאשר $\lambda_1 \dots \lambda_m$ הע"עים השונים של S_1. לכל $1 \leqslant i \leqslant m$ מתקיים ש־V_{λ_i} (המרחב העצמי) הוא S_1־אינווריאנטי שהרי אם $v \in V_{\lambda_i}$ ונחשב:

$$S_1(S_2 v) = S_2(S_1 v) = S_2(\lambda_i v) = \lambda_i S_2 v \implies S_2 v \in V_{\lambda_i}$$

כאשר $S_2|_{V_{\lambda_i}} : V_{\lambda_i} \to V_{\lambda_i}$ צמודה לעצמה, ולכן המפשט הספקטרלי לצמודות לעצמן אומר שבתוך V_{λ_i} ישנו בסיס אורתוגונלי של ו"עים מ־S_2. האיחוד של כל הבסיסים הללו מכל מ"ע של S_1 יהיה בסיס אורתוגונלי של ו"עים משותפים ל־S_1 ול־S_2. ■

הוכחת המשפט הספקטרלי.

$\implies$ לפי המסקנה הקודמת, אם ישנו לכסון אורתוגונלי T בהכרח נורמלית.

$\impliedby$ נגדיר $S_1 = \frac{T+T^*}{2},\ S_2 = \frac{T-T^*}{2i}$. הן וודאי צמודות לעצמן מהלינאריות וכל השטויות ממקודם, והן גם מתחלפות אם תטרחו להכפיל אותן. מהטענה קיים ל־V בסיס אורתוגונלי של ו"עים משותפים ל־S_1, S_2 ונסמנו $\{b_i\}_{i=1}^n$ וגם $S_1 b_i = \alpha_i b_i,\ S_2 b_i = \beta_i b_i$. אפשר גם לטעון ש־$\alpha_i, \beta_i \in \mathbb{R}$ אבל זה לא מועיל לנו. נשים לב ש־$T = S_1 + iS_2$, כלומר $\forall i \in [n] \colon T(b_i) = S_1(b_i) + iS_2(b_i) = \alpha_i b_i + i\beta_i b_i = (\alpha + i\beta_i) b_i$ וזהו בסיס אורתוגונלי של ו"עים של T. ■

למעשה, הבנו מהפירוק של S_1, S_2 ש־S_1 נותנת את החלק הממשי של הע"ע ו־S_2 את החלק המדומה. זהו פירוק מועיל שכדאי לזכור.

נסכם: יש לנו שתי גרסאות של המשפט הספקטרלי:

משפט (המשפט הספקטרלי מעל $\mathbb{R}$)**.** T סימטרית אמ"מ קיים בסיס א"נ של ו"ע.

משפט (המשפט הספקטרלי מעל $\mathbb{C}$)**.** T נורמלית אמ"מ קיים בסיס א"נ של ו"ע.

משום שמטריצה הרמיטית (וצמודה לעצמה באופן כללי) היא בפרט נורמלית כי מטריצה מתחלפת עם עצמה, נסיק שלצמודה לעצמה קיים בסיס אורתוגונלי מלכסן (בעמוד הכיוון ההפוך לא נכון מעל המרוכבים, שם ההעתקה יכולה להיות נורמלית ולא סתם הרמיטית).

הערה 55. המשפט הספקטרלי מעל $\mathbb{R}$ לא אומר שהעתקה/מטריצה סימטרית היא לכסינה מעל $\mathbb{R}$, משום שהבסיס האורתונורמלי המלכסן המדובר הוא בסיס מעל $\mathbb{C}$ (בעוד ההעתקה/מטריצה מעל $\mathbb{R}$). לדוגמה, סיבוב ב־$90°$ לא לכסין ב־$\mathbb{R}$ אך צמוד לעצמו.

(2.3.1.3) תוצאות ממשפט הפירוק הספקטרלי

משפט 141. תהי $T \colon V \to V$ ט"ל, ו־V ממ"פ מעל $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, ויהי B בסיס א"נ של V. אזי אם $A = [T]_B$:

$$[T^*]_B = A = ([T]_B)^*$$

הוכחה. נזכר ש־:

$$[T]_B = \begin{pmatrix} | & & | \\ [Te_1]_B & \cdots & [Te_n]_B \\ | & & | \end{pmatrix}$$

נסמן $B = \{e_i\}_{i=1}^n$ בסיס. נבחין ש־:

$$Te_j = \sum_{i=0}^n a_{ij} e_i,\ a_{ij} = \langle Te_j \,|\, e_i \rangle$$

נסמן ב־C את המטריצה המייצגת $[T^*]_B$:

$$c_{ij} = \langle T^* e_j \,|\, e_i \rangle$$

ונחשב:

$$c_i j = \langle T^* e_j \,|\, e_i \rangle = \langle e_j \,|\, Te_i \rangle = \overline{\langle Te_i \,|\, e_j \rangle} = a_{ij}$$

■

מסקנה: אם A נורמלית אז T_A נורמלית מעל $\mathbb{F}^n$ אם הסטנדרטית. בפרט מתקיים עליה המשפט הספקטרלי. גם אם A ממשית, הע"ע עלולים להמצא מעל $\mathbb{C}$ (אלא אם היא צמודה לעצמה, ואז הם מעל $\mathbb{R}$).

משפט 142. יהיו $x_1 \dots x_n, y_1 \dots y_n \in \mathbb{R}$. נניח $\forall i, j \in [n] \colon i \neq j \implies x_i \neq x_j$. אז $\exists! p \in \mathbb{R}_{\leqslant n-1}[x] \colon \forall i \in [n] \colon p(x_i) = y_i$ עד לכדי חברות (באופן שקול: נניח p מתוקן)

הוכחה. ידוע שהפולינום מהצורה $p(x) = \sum_{k=0}^{n-1} a_k x^k = (1, x, x^2, \dots x^{n-1})(a_0 \dots a_{n-1})^T$. למעשה, נקבל את מטריצת ונדרמונד:

$$\underbrace{\begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix}}_{\mathcal{V}} \underbrace{\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{pmatrix}}_{a} = \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}}_{y}$$

וידוע שהדטרמיננטה של $\mathcal{V}$ היא מטריצת ונדרמונד היא $\prod_{i<j}(x_i - x_j)$, שאיננה אפס מההנחה ש־$\forall i \neq j \colon x_i \neq x_j$, ולכן למערכת המשוואות $(\mathcal{V} \mid y)$ קיים ויחיד פתרון, הוא a, שמגדיר באופן יחיד את מקדמי הפולינום.

אם $x_i = y_i$ בפולינום לעיל, אז $f \in \mathbb{R}[x] \implies \forall a \in \mathbb{C} \colon f(\bar{a}) = \overline{f(a)}$. הוכחה: נניח בשלילה, אז $\exists \alpha \in \mathbb{C} \setminus \mathbb{R} \colon f(\alpha) = 0$ כך ש־$f(\bar{\alpha}) = 0$. אזי $0 \neq f(\bar{\alpha}) = \overline{f(\alpha)} = 0$ וזו סתירה. ∎

הערה 56. הפולינום שמקיים זאת נקרא פולינום לגראנג' והוא בונה אינטרפולציה די נחמדה אך יקרה חישובית. ניתן לחשב את הפולינום מפורשות באופן הבא:

$$p(x) = \sum_{i=1}^{n} \left(y_i \prod_{j=1}^{n} \frac{x - x_j}{x_i - x_j} \right)$$

משפט 143. תהי $A \in M_n(\mathbb{F})$ נורמלית, אז קיים פולינום $\exists f(x) \in \mathbb{R}[x] \colon A^* = f(A)$

הערה: באופן כללי התנאי ש־$\exists f \in \mathbb{R}[x] \colon f(A) = B$ הזה מספיק **אך לא הכרחי** לכך ש־A, B מתחלפות.

הוכחה. עבור A נורמלית מהמשפט הספקטרלי קיים בסיס אורתונורמלי מלכסן ולכן קיימת P הפיכה כך ש־$P^{-1}AP =$ $\operatorname{diag}(\lambda_1 \dots \lambda_n)$. לכן $P^{-1}A^*P = \operatorname{diag}(\bar{\lambda}_1 \dots \bar{\lambda}_n)$. נשתמש במשפט לפיו יש פולינום $f \in \mathbb{R}[x]$ כך ש־$f(x_i) = \bar{x}_i$ ובפרט בעבור $x_i = \lambda_i$ קיים פולינום עבורו $f(\lambda_i) = \bar{\lambda}_i$. אזי

$$f(\operatorname{diag}(\lambda_1 \dots \lambda_n)) = \operatorname{diag}(f(\lambda_1) \dots f(\lambda_n)) = \operatorname{diag}(\bar{\lambda}_1 \dots \bar{\lambda}_n)$$

$$f(P^{-1}AP) = P^{-1}f(A)P = P^{-1}A^*P \implies f(A) = A^*$$

עוד נבחין ש־$\deg f = n - 1$. ∎

משפט 144. אם $T \colon V \to V$ נורמלית, אז $\exists f \in \mathbb{R}[x] \colon f(T) = T^*$.

הוכחה. נבחר בסיס א"נ $A = [T]_B \implies A^* = [T^*]_B$. כבר הוכחנו שאם T נורמלית אז A נורמלית ולכן מהמשפט הקודם קיים f מתאים כך ש־$[T^*]_B = A^* = f(A) = f([T]_B) = [f(T)]_B$. סה"כ $[T^*]_B = [f(T)]_B$ ומחח"ע העברת בסיס $T^* = f(T)$ כדרוש. ∎

אם $T \colon V \to V$ ט"ל, $U, W \subseteq V$ תמ"וים T־איוונריאנטי כך ש־$B = U \oplus W$. אם $\mathcal{B}$ בסיס של V, כאשר קישא של הבסיס הוא הבסיס של U אז:

$$[T]_{\mathcal{B}} = \begin{pmatrix} [T|_U]_{\mathcal{B}} & \\ & [T|_W]_{\mathcal{B}} \end{pmatrix}$$

בפרט בעבור ניצבים $V = U \oplus U^\perp \implies U \subseteq V$. ניעזר בכך כדי להוכיח את המשפט הבא:

משפט 145. אם $U \subseteq V$ תמ"ו אינוואריאנטי ביחס ל־T אז $U^\perp$ הוא T^*־אינוואריאנטי.

הוכחה. יהי $w \in U^\perp$. רוצים להראות $T^*w \in U^\perp$. יהי $u \in U$ אז:

$$\langle T^*w \,|\, u \rangle = \langle w \,|\, Tu \rangle = \langle w \,|\, u' \rangle \,,\ u' \in U \implies \langle w \,|\, u' \rangle = 0 \quad \top$$ ∎

משפט 146. בעבור $T \colon V \to V$ נורמלית, אם היא U־אינוואריאנטי אז גם T^* הוא U־אינווארינטי

הוכחה. נבחין ש־$f(T) = T^*$ כלשהו, וכן U הוא T־אינוואריאנטי ולכן U הוא $f(T)$־איוו' וכאן די גמרנו את ההוכחה. ■

מסימטריות $U^\perp$ הוא T^*, מהמשפט גם $(T^*)^*$ איוו' ולכן T־אינוואריאנטי.

משפט 147. יהי V מעל $\mathbb{R}$ מ"ו וכן $T\colon V \to V$ ט"ל. אז קיים $U \subseteq V$ שהוא T־אינוואריאנטי וממדו לכל היותר 2.

משפט 148. מעל $M_2(\mathbb{R})$, קיימת צורה כללית למטריצות לא לכסינות נורמליות.

הוכחה. ננסה להבין מי הן $A \in M_2(\mathbb{R})$ שהן נורמליות. מעל $\mathbb{C}$ הן פשוט לכסינות. נבחין ש־:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \; A^* = f(A) \in \mathbb{R}_{\leqslant 1}[x] = \alpha A + BI, \; A = \alpha A^T + \beta I = \alpha(\alpha A + \beta I) + \beta I \tag{2.1}$$

$$\implies \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \alpha^2 a + \beta(\alpha+1) & \alpha^2 b \\ \alpha^2 c & \alpha^2 d + \beta(\alpha+1) \end{pmatrix} \tag{2.2}$$

$$\begin{cases} (b \wedge c \neq 0) \quad \alpha = 1 \implies A = A + 2\beta I \implies \beta = 0, \; A = A^T \cancel{+\beta I} \\ (b \wedge c \neq 0) \quad \alpha = -1 \implies A^T = -A + \beta I \implies A + A^T = \begin{pmatrix} \beta & 0 \\ 0 & \beta \end{pmatrix} \implies b = -c, \; A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \\ (b \vee c = 0) \implies A = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} \end{cases}$$

המקרה השני – זה פשוט סיבובים, אבל בניפוח (כי הדטרמיננטה היא $a^2 - b^2$). ■

הערה: מעל $\mathbb{C}$ "זה מטופש" כי הפולינום מתפרק (ואז המרחב העצמי של ע"ע כלשהו יקיים את זה).

הוכחה. נפרק ל־$m_T(x)$ מינימלי ו־$g(x)$ גורם אי־פריק כך ש־$m_T(x) = g(x)h(x)$. לכל g אי פריק ב־$\mathbb{R}$ מתקיים $\deg g \leqslant 2$, כי אם g ממעלה אחת סיימנו, אחרת הוא ממעלה 2 לפחות כי בהינתן שורש מרוכב a לפולינום $m_T(x)$ גם $\bar{a}$ שורש, ואז $m_T(x) = (x-a)(x-\bar{a}) = (x^2 - |a|^2)$ דהיינו כל שורש מרוכב משוייך לגורם ממשי ריבועי לכל היותר, ומשום ש־$m_T(x)$ מתפרק מעל המרוכבים, ניתן לסכם ש־g מדרגה 2 לכל היותר.

- אם g ממםד 1 אז $g = x - \lambda$ כלשהו ואז V_λ המ"ו העצמי של λ הממשי, מרחב מממד $1 \leqslant 2$ המקיים את הדרוש.
- אם $\deg g = 2$ בה"כ ניתן להניח g מתוקן (נעביר את הקבוע ל־h). ז"א $g(x) = x^2 + ax + b$ אז $g(T)$ אינו הפיך (מלמת החלוקה לפולינום מינימלי) כלומר $\exists 0 \neq v \in \ker g(T)$. לכן:

$$(T^2 + aT + bI)v = 0 \implies T^2 v = -aTv - bv$$

ולכן $U = \operatorname{span}(v, Tv)$ תמ"ו עם ממד לכל היותר 2 וגם T־איוואריאנטי.

סה"כ בשני המקרים מצאנו תמ"ו המקיים את הדרוש. ■

הערה 57. בעבור T נורמלית (ולא כללית) הטענה נכונה ללא תלות במשפט היסודי של האלגברה.

בעבור נורמליות, מהמשפט הספקטרלי ומטענות קודמות, עבור $T\colon V \to V$ ממשית קיים בסיס א"נ $\mathcal{B}$ של V שבעבורו המטריצה המייצגת של T היא מטריצת בלוקים 2×2 מצורה של $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$:

$$[T]_{\mathcal{B}} = \operatorname{diag}\left(\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix} \cdots \begin{pmatrix} a_k & b_k \\ -a_k & b_k \end{pmatrix}, \; \lambda_1 \cdots \lambda_m \right)$$

כאשר כמובן $2k + m = n$.

2.3.2 ∼ מטריצות אוניטריות

הגדרה 90. יהי V ממ"פ. אז $T\colon V \to V$ תקרא אוניטרית (אם $\mathbb{F} = \mathbb{C}$) או אורתוגונלית (אם $\mathbb{F} = \mathbb{R}$) אם $T^*T = I$ או במילים אחרות $T^* = T^{-1}$ (מהגדרת הפיכה).

ברור שט"ל כזו היא נורמלית. **דוגמה.** עבור T_θ הסיבוב ב־θ־מעלות, במישור $\mathbb{R}^2$, אז $(T_\theta)^* = T_{-\theta} = T_\theta^{-1}$. **דוגמה.** עבור T שיקוף מתקיים $T^2 = I$ וכן $T^* = T$ וסה"כ $T^* = T = T^{-1}$.

משפט 149. T איזומטריה אמ"מ מתקיים אחד מבין הבאים:

1. (ההגדרה) $T^*=T^{-1}$
2. $TT^*=T^*T=I$
3. $\forall u,v\in V\colon\ \langle Tu\,|\,Tv\rangle=\langle u\,|\,v\rangle$
4. T מעבירה כל בסיס א"נ של V לבסיס א"נ של V
5. T מעבירה בסיס א"נ אחד של V לבסיס א"נ של V [מקרה פרטי של 4 בצורה טרוויאלית, אך גם שקול!]
6. $\forall v\in V\colon\ ||Tv||=||v||$
7. $\widehat{u,v}=\widehat{Tu,Tv}$

כלומר: היא משמרת מכפלה פנימית, גודל וזווית.

הערה 58. את קבוצת המטריצות האורתוגונליות מסמנים ב־$O_n(\mathbb{F})$, ומקובל להתייחס אליה כאל חבורה אבלית ביחס לפעולת ההרכבה. ישנן סוג מיוחד של מטריצות אוניטריות, כאלו שלא רק מקיימות $|\det A|=1$, אלא ממש $\det A=1$. חבורת האובייקטים האלו קרויה $SO_n(\mathbb{F})$, קיצור של Special Orthogonal Matrix. יש שתי קבוצות של מטריצות סיבוב מיוחדות שמעיינות אותנו – $SO_2(\mathbb{R})\cong\{c\in C\colon\ |c|=1\}$, איזומורפיזם שראינו בעבר כשלכסנו מעל המרוכבים מטריצות סיבוב, ו־SO_3 שמעניינת בגלל הקשר שלה לאלגברת קווטרניונים.

הגדרה 91. העתקה $T\colon V\to V$ (כאשר V ממ"פ) תקרא איזומטריה אם $\forall v\in V\colon\ ||v||=||Tv||$

באופן כללי אוניטרית/אורתוגונלית שקולות לאיזומטריה לינארית (כלומר שם כללי לאורתוגונליות/אוניטריות יהיה איזומטריות).

הערה 59. איזומטריה, גם מחוץ לאלגברה לינארית, היא פונקציה שמשמרת נורמה/גודל.

הערה 60. אפשר להסתכל על מתכונה 4 על איזומטריות לינאריות כעל איזומורפיזם של ממ"פים.

הוכחה. נפרק לרצף גרירות

$1\to2$ אם $T^*=T^{-1}$ אז מהגדרה $TT^*=I$ ומהיות הופכית יחידה משני הצדדים $TT^*=T*^T=I$.

$2\to3$
$$T^*=T^{-1}\implies\langle Tv\,|\,Tu\rangle=\langle v\,|\,T^*Tu\rangle=\langle v\,|\,u\rangle$$

$3\to4$ נאמר ש־$(v_1\dots v_n)$ א"נ. צ.ל. $(Tv_i)_{i=1}^n$ א"נ. לשם כך נצטרך להוכיח את שני התנאים – החלק של האורתו והחלק של הנורמלי. בשביל שניהם מספיק להוכיח ש־: $\langle Tv_i\,|\,Tv_j\rangle=\langle v_i\,|\,v_j\rangle=\delta_{ij}$

$4\to5$ טרוויאלי

$5\to6$ יהי $(v_1\dots v_n)$ בסיס א"נ כך ש־$(Tv_1\dots Tv_n)$ א"נ. אז:

$$v=\sum_{i=0}^n\alpha_iv_i\implies||v||^2=\left\langle\sum_{i=0}^n\alpha_iv_i\,\middle|\,\sum_{i=0}^n\alpha_iv_i\right\rangle=\sum_{i=0}^n|\alpha_i|^2$$
$$||Tv||^2=\left\langle T\left(\sum_{i=0}^n\alpha_iv_i\right)\middle|\,T\left(\sum_{i=0}^n\alpha_iv_i\right)\right\rangle=\left\langle\sum_{i=0}^n\alpha_iT(v_i)\,\middle|\,\sum_{i=0}^n\alpha_iT(v_i)\right\rangle=\sum|\alpha_i|^2$$

$5\to1$ מניחים $\forall v\in V\colon\ ||Tv||=||v||$. ידועות השקילויות הבאות:

$$T^*=T^{-1}\iff T^*T=I\iff T^*T-I=0$$

בעבר ראינו את הטענה הבאה: נניח ש־S צמודה לעצמה וכן ש־$\forall v\colon\ \langle Sv\,|\,v\rangle=0$, אז $S=0$. במקרה הזה:

$$S:=T^*T-I\implies S^*=(T^*T-I)^*=(T^*T)^*-I^*=T^*(T^*)^*-I=S\implies S^*=S$$

והיא אכן צמודה לעצמה. עוד נבחין ש־:

$$\langle Sv\,|\,v\rangle=\langle(T^*T-I)v\,|\,v\rangle=\langle T^*Tv\,|\,v\rangle-\langle v\,|\,v\rangle=\langle Tv\,|\,Tv\rangle-\langle v\,|\,v\rangle=||Tv||^2-||v||^2=0$$

השוויון האחרון נכון מההנחה היחידה שלנו ש־$||Tv||=||v||$. סה"כ $TT^*-I=0$. סה"כ הוכחנו $TT^*-I=0$ שזה שקול ל־$T^*=T^{-1}$ מהשקילויות לעיל כדרוש. ■

משפט 150. תהי $T\colon V\to V$ איזומטריה, ו־λ ע"ע של T. אז $|\lambda|=1$.

הוכחה. יהי v ו"ע של הע"ע λ. אז:

$$|\lambda|^2\langle v\,|\,v\rangle=\lambda\bar{\lambda}\langle v\,|\,v\rangle=\langle Tv\,|\,Tv\rangle=\langle v\,|\,v\rangle$$ ■

הערה 61. מעל המרוכבים לא מתקיים בהכרח $\lambda \in \{1,-1\}$, בעוד מעל הממשיים כן. שימו לב שיש אינסוף מספרים המקיימים $|\lambda| = 1$ מעל המרוכבים, הם התמונה של e^{ix} .$\lambda x \in \mathbb{R}$.
הגדרה 92. תהי $A \in M_n(\mathbb{F})$. אז A תיקרא אוניטרית/אורתוגונלית אם $A^* = A^{-1}$.
משפט 151. אוניטרית אמ"מ $A\overline{A^T} = I$.
משפט 152. אורתוגונלית אמ"מ $AA^T = I$.
הערה 62. אוניטרית בה מלשון unit – היא שומרת על הגודל, על וקטורי היחידה (ה־unit vectors).
משפט 153. יהי $\mathcal{B}$ בסיס א"נ של V ו־$T\colon V \to V$ אז T אוניטרית/אורתוגונלית אמ"מ $A = [T]_B$ אוניטרית/אורתוגונלית.

הוכחה.

$$AA^* = [T]_\mathcal{B}[T^*]_\mathcal{B} = [TT^*]_\mathcal{B}, I = AA^* \iff [TT^*]_\mathcal{B} = I \iff TT^* = I$$ ■

"היה לי מרצה בפתוחה שכתב דבר לא מדויק בסיכום, ואז הוריד נקודות לסטודנטים שהסתמכו על זה. הוא אמר שזה מתמטיקה, אתם אחראים להבין מה נכון או לא – גם אם כתבתי שטויות".
סימון 14. א"נ = אוניטרית בהקשר של מטריצות (בהקשר של מרחבים – אורתונורמלי)
משפט 154. התנאים הבאים שקולים על $A \in M_n(\mathbb{F})$.

1. A אוניטרית
2. שורות A מהוות בסיס א"נ של $\mathbb{F}^n$ (ביחס למכפלה הפנימית הסטנדרטית)
3. עמודות A מהוות בסיס א"נ של $\mathbb{F}^n$.
4. $\forall u, v \in \mathbb{F}^n\colon \langle Au \,|\, Av\rangle = \langle u \,|\, v\rangle$ (ביחס למכפלה הפנימית הסטנדרטית)
5. $\forall v \in \mathbb{F}^n\colon ||Av|| = ||v||$ (ביחס למכפלה הפנימית הסטנדרטית)

הערה שלא קשורה למשפט: נאמר ש־$[T^*]_B = [T]_B^*$ אמ"מ B בסיס **אורתונורמלי**. עבור בסיס שאינו א"נ זה לא בהכרח מתקיים.

הערה נוספת: זה בערך אמ"מ כי יש כמה מקרי קצה כמו מטריצת האפס.

הוכחה.

$1 \leftrightarrow 2$ נוכיח את הגרירה הראשונה

$$\begin{pmatrix} - & v_1 & - \\ & \vdots & \\ - & v_n & - \end{pmatrix}\begin{pmatrix} | & & | \\ \bar{v}_1^T & \cdots & \bar{v}_n^T \\ | & & | \end{pmatrix} = AA^* = I \iff A \text{ א"נ} \implies v_i\bar{v}_j^T = \delta_{ij}$$

הטענה האחרונה שקולה לכך ש־$v_1 \dots v_n$ בסיס א"נ (ביחס למ"פ הסטנדטית של $\mathbb{F}^n$)

$1 \leftrightarrow 3$ מספיק להוכיח A אוניטרית אמ"מ A^T אוניטרית (בגלל השקילות $2 \leftrightarrow 1$). מסימטריה ($(A^T)^T = A$) למעשה מספיק להוכיח A א"נ גורר A^T א"נ. נוכיח:

$$A^*A = I \implies A^T\bar{A} = I \implies (A^T)^* = \bar{A} \implies A^T(A^T)^* = I$$

$4 \leftrightarrow 1$ נתבונן ב־$T_A\colon \mathbb{F}^n \to \mathbb{F}^n$ כאשר $\mathcal{E}$ הבסיס הסטנדרטי, ו־$T_A(v) = Av$. לכן $[T_A]_\mathcal{E} = A$. אז T_A אוניטרית אמ"מ $[T_A]_\mathcal{E}$ אוניטרית. אז:

$$\langle Au \,|\, Av\rangle = \langle T_Au \,|\, T_Av\rangle = \langle u \,|\, v\rangle$$

$5 \leftrightarrow 1$ תוצאה ישירה מ־$4 \leftrightarrow 1$, שכן מנוסחת הפולריזציה A משמרת מכפלה פנימית אמ"מ היא משמרת נורמה. ■

(2.3.2.1) צורה נורמלית למטריצה אורתוגונלית

סימון 15. נסמן ב־A_{θ_i} את מטריצת הסיבוב של $\mathbb{R}^2$ ב־θ מעלות, היא:

$$A_{\theta_i} = \begin{pmatrix} \cos\theta_i & \sin\theta_i \\ -\sin\theta_i & \cos\theta_i \end{pmatrix}$$

שאלה. מהן המטריצות $A \in M_2(\mathbb{R})$ האורתוגונליות?

התשובה. בהינתן $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ מהיות העמודות והשורות מהוות בסיס א"נ:

$$\begin{cases} a^2+b^2=1 \\ c^2+d^2=1 \\ a^c+c^2=1 \end{cases} \implies a=\cos\theta,\ b=\sin\theta$$

עוד נבחין ש־$ac+bd=0$ כי:

$$AA^T = I \implies \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{pmatrix} = I$$

סה"כ מכך ש־$a^c+c^2=1$ ו־$b^2+d^2=1$ נקבל שתי צורות אפשריות:

$$A_1 := A = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \vee A_2 := A = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = A_\theta$$

נבחין ש־A_2 הוא סיבוב ב־θ, ו־A_1 שיקוף ניצב ביחס ל־$\frac{\theta}{2}$. זה לא מפתיע שכן $\det A_1 = -1,\ \det A_2 = 1$. יתרה מכך, A_1 לכסינה עם ע"ע שני ע"ע – 1 ו־-1. ■

"אם הייתם רוצים תקופות מבחנים נורמליות הייתם צריכים להיוולד בזמן אחר"

הערה 63. לבדיקת שפיות, ננסה לפרק מעל המרוכבים את הצורה שקיבלנו, ואכן:

$$\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \sim \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

בהתאם לכך ש־$\left|e^{i\theta}\right| = \left|e^{-i\theta}\right| = 1$ כמצופה מע"עים של מטריצה אוניטרית.

מסקנה 23 (הצורה הנורמלית של ט"ל אורתוגונלית). תהי $T\colon V \to V$ אורתוגונלית. אז קיים בסיס א"נ של V, שביחס אליו קיימות $\theta_1 \dots \theta_k$ זווית כך שהמטריצה המייצגת את T היא מהצורה:

$$\operatorname{diag}(A_{\theta_1}, \cdots, A_{\theta_k}, 1, \cdots, 1, -1, \cdots, -1)$$

(לכאורה אוניטרית לא מעיינת כי היא נורמלית ולכן לכסינה אורתונורמלית מהמשפט הספקטרלי, וכל הע"עים המרוכבים שלה מגודל 1 בכל מקרה)

הוכחה. ידוע עבור נורמלית:

$$A = \begin{pmatrix} \begin{bmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{bmatrix} & & & & & \\ & \ddots & & & & \\ & & \begin{bmatrix} a_m & b_m \\ -b_m & a_m \end{bmatrix} & & & \\ & & & \lambda_1 & & \\ & & & & \ddots & \\ & & & & & \lambda_k \end{pmatrix}$$

נסמן $\square_i = \left(\begin{smallmatrix} a_i & b_i \\ -b_i & a_i \end{smallmatrix}\right)$. במקרה הזה משום שהיא אורתוגונלית על $\mathbb{R}$ אז $\lambda_i = \pm 1$ כי $|\lambda_i| = 1$. נתבונן במטריצה $\square_i$ כלשהי, אז $\square_i$ הנפרש ע"י $U =: u_k, u_{k+1}$ מקיים:

$$[T_{|U}]_{B_U} = \square_i = \begin{pmatrix} a_i & b_i \\ -b_i & a_i \end{pmatrix} \quad [Tu_k]_{U_B} = \begin{pmatrix} a \\ -b \end{pmatrix}, \quad [Tu_{k+1}] = \begin{pmatrix} b \\ a \end{pmatrix}$$

ומשום שהצמצום של אורתוגונלית על מ"ו T־אינוואריאנטי היא עדיין אורתוגונלית, והראנו שהאורתוגונליות ב־$M_2(\mathbb{R})$ הן מטריצות הסיבוב/השיקוף+סיבוב לעיל. המטריצה של שיקוף+סיבוב ב־$\frac{\theta}{2}$ (זו שסומנה לעיל ב־A_1) לכסינה ולכן תהפוך לע"ע $\lambda_1 \dots \lambda_n$ (עד לכדי סדר איברי בסיס) שהם בהכרח מגודל ± 1 בכל מקרה, ויבלעו בשאר הע"ע, ובכך סיימנו. ■

אבל האם הייצוג יחיד? ננסה להבין את יחידות הייצוג עבור נורמלית כללית, ומשם לגזור על אורתוגונלית.

משפט 155. כל שתי מטריצות בצורה לעיל שמייצגות את אותה $T\colon V \to V$ נורמלית, שוות עד כדי סדר הבלוקים על האלכסון.

(יש כאן מה להוכיח רק בעבור $\mathbb{R}$, שכן מעל $\mathbb{C}$ לכסינה בכל מקרה, והע"עים וריבויים לא משתנים כתלות בייצוג).

הוכחה. ידוע שבעבור $\lambda_1 \dots \lambda_k$ ע"עים:

$$f_T(x) = \left(\prod (x-\lambda_i)\right)\left(\prod (x^2 - 2a_i x + a_i^2 + b_i^2)\right)$$

כאשר המכפלה הראשונה באה מהע"עים והשניה מהריבועים $\square_i$. נבחין שלכל תמ"ו a_i נקבע ביחידות, ולכן b_i נקבל ביחידות עד כדי סימן (נסיק זאת מהפולינום האופייני). ברור שהע"עים נקבעים ביחידות עוד מההרצאות הראשונות. ■

אז מאיפה בה שינוי הכיוון של b, בעבור מטריצות אורתוגונליות? כלומר, מדוע A_{θ_i} שקולה ל־$A_{-\theta_i}$? זאת כי הן דומות באמצעות ההעתקה שהופכת את הצירים, מה ששקול להחליף את עמודות A_{θ_i}.

תרגיל. חשבו: (כולל פתרון)

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} a & b \\ -b & a \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

מכאן נסיק שאכן המטריצות להלן דומות עד לכדי שינוי בסיס, וזו הסיבה שלא איכפת לנו מהסימן של b.

הערה 64. למעשה, משום שהמטריצות $\square_i$ אינן פריקות למרחבים אינווראינטים קטנים יותר, ולכן נוכל להפוך את כל הבלוקים על המטריצה ולקבל בלוקי ג'ורדן, שכבר אנחנו יודעים שהם יחידים.

(2.3.2.2) המשפט הספקטלי בניסוח מטריציוני

משפט 156 (המשפט הספקטרלי "בשפה קצת מטרציונית"). תהי $A \in M_n(\mathbb{F})$ מטריצה צמודה לעצמה. אז קיימת מטריצה P אורתוגונלית/אוניטרית (בהתאם לשדה ממנו יצאנו), ומטריצה אלכסונית D כך ש־ $A = P^{-1}DP$

כלומר – מטריצת מעבר הבסיס של המשפט הספקטרלי, שמעביר אותנו לפירוק הספקטרלי, היא איזומטריה. למעשה חיזקנו את המשפט הספקטרלי – המעבר לבסיס המלכסן, מסתבר להיות מיוצג ע"י מטריצות איזומטריות.

המרצה מדגיש שלא השתמשנו במשפט הזה בכלל בבסיסים ווקטורים – אפשר לתאר את עולם הדיון של המטריצות, מעצם היותו עולם דיון איזומורפי להעתקות ולמרחבים וקטורים, בלי לדבר בכלל על העתקות ומרחבים וקטורים. המשפט מתאר באופן טהור מטריצות בלבד.

למה 14. תהי $A \in M_n(\mathbb{F})$ מטריצה ריבועית, וכן $\{e_1 \dots e_n\}$ בסיס א"נ של V. נניח ש־A היא מטריצת המעבר מבסיס $\{e_1 \dots e_n\} \to \{v_1 \dots v_n\}$. אז A איזומטריה אמ"מ $\{v_1 \dots v_n\}$ בסיס אורתונורמלי.

הוכחת המשפט. תהי $T_A \colon \mathbb{F}^n \to \mathbb{F}^n$ כך ש־$T_A(x) = Ax$. אז $A = [T_A]_{\mathcal{E}}$ כאשר $\mathcal{E} = \{e_1 \dots e_n\}$ הבסיס הסטנדרטי. ידוע של־T_A יש בסיס אורתונורמלי מלכסן, כלומר קיים בסיס א"נ $\mathcal{B}$ כך ש־$[T_A]_{\mathcal{B}} = D$ כאשר D אלכסונית כלשהית. נבחין ש־$[T_A]_{\mathcal{B}} = [Id]^{\mathcal{B}}_{\mathcal{E}}[T_A]_{\mathcal{E}}[Id]^{\mathcal{E}}_{\mathcal{B}}$, נסמן $P = [Id]^{\mathcal{B}}_{\mathcal{E}}$ ונבחין ש־$[T_A]_{\mathcal{B}} = P^{-1}AP$ ומהלמה P מטריצת מעבר מבסיס א"נ לבסיס א"נ ולכן איזומטריה. סה"כ הראנו את הדרוש. ■

"יאללה הפסקה? לאם'

2.3.3 ~ פירוק פולארי

(2.3.3.1) מבוא, וקישור לתבניות בי־לינאריות

הערה 65. עבור מטריצות אורתוגונליות, במקרה של $\mathbb{F} = \mathbb{R}$ נקבל ש־

$$A = P^{-1}DP \implies PP^T = I \implies P^{-1} = P^T \implies A = P^T DP$$

מה שמחזיר אותנו לתבניות בי־לינאריות. נוכל לקשר את זה לסינגטורה. זאת כי A לא רק דומה, אלא גם חופפת ל־D. גם מעל $\mathbb{C}$ נקבל דברים דומים, אך לא במדויק, שכן מכפלה פנימית מעל $\mathbb{C}$ היא ססקווי־בי־לינארית פשוט בי־לינארית.

משפט 157. עבור $A \in M_n(\mathbb{C})$ נורמלית, אז

- $A^* = A$ (צמודה לעצמה) אמ"מ כל הע"עים שלה ממשיים.
- $A^* = A^{-1}$ אמ"מ כל הע"ע שלה מנורמה 1.

הוכחה. את הכיוון $\Longleftarrow$ כבר הוכחנו במשפט קודם. נותר להוכיח את הכיוון השני.

- נניח שכל הערכים העצמיים ממשיים, ו־A נורמלית. נוכל להשתמש במשפט הספקטרלי עליה: לכן קיימת מטריצה אוניטרית P ואלכסונית Λ כך ש־$A = P^{-1}\Lambda P$. ידוע $\Lambda \in M_n(\mathbb{R})$ כי אלו הע"ע מההנחה. נבחין ש־:

$$A^* = P^*\Lambda^*(P^{-1})^* = P^{-1}\Lambda P = A$$

כי $PP^* = I$ ו־Λ אוניטרית (אז ה־transpose לא עושה שום דבר) מעל $\mathbb{R}$ (אז ההצמדה לא עושה שום דבר).

- נניח A נורמלית וכל הע"ע מנורמה 1. נוכיח A אוניטרית. בעבור הפירוק הספקטרלי לעיל $A = P^{-1}\Lambda P$ נקבל כאן ש־Λ אוניטרית, ומהמשפט הספקטרלי P אונטרית גם כן. A מכפלה של 3 אוניטריות ולכן אוניטרית.

(הסיבה שמכפלה של אוניטריות היא אונטרית: בעבור A, B א"נ מתקיים

$$\forall v \in V\colon\ \langle ABv \,|\, ABv\rangle = \langle Bv \,|\, Bv\rangle = \langle v \,|\, v\rangle$$

משמרת מכפלה פנימית, וזה שקול להיותה אוניטרית ממשפט לעיל) ■

תזכורת: אם V ממ"פ מעל $\mathbb{F}$, אז $T\colon V \to V$ תקרא חיובית או אי־שלילית (וכו') אם $T = T^*$ וגם $\forall v \neq 0\colon\ \langle Tv \,|\, v\rangle \geqslant / > 0$.

תזכורת: מעל $\mathbb{R}$, הוכחנו שלכל תבנית סימטרית, יש ייצוג יחיד באמצעות מטריצה אלכסונית עם $-1, 1, 0$ על האלכסון.

סימון 16. הסיגנטורה של f תסומן ע"י $\sigma_-(f), \sigma_0(f), \sigma_+(f)$ כמספר האפסים, האחדים וה־-1 ב־f.

המשך תזכורת: כל מטריצה סימטרית חופפת למטריצה יחידה מהצורה לעיל.

משפט 158. נניח ש־A מייצגת את התבנית הסימטרית f (עולם הדיון מעל $\mathbb{R}$). אז, הסיגנטורה שווה ל־$\sigma_+ = \#(\lambda \mid \lambda > 0)$ עבור λ ע"ע עם חזרות (במידה ושייך ליותר מו"ע יחיד). באופן דומה $\sigma_- = \#(\lambda \mid \lambda < 0)$ כאשר λ ע"ע.

הוכחה. משום ש־A מייצגת סימטרית אז A סימטרית. לפי המשפט הספקטרלי קיימת P אורתוגונלית ו־Λ אלכסונית כך ש־$A = P^{-1}\Lambda P = P^T\Lambda P$. A דומה לאלכסונית ממשית (כי A סימטרית ולא סתם נורמאלית) וחופפת אליה. בעזרת נרמול המטריצה Λ האלכסונית (ניתן לבצע תהליך נרמול באמצעות פעולות שקולות תחת חפיפה), היא חופפת למטריצה מהצורה $\operatorname{diag}(1\ldots1, -1\ldots-1, 0\ldots0)$ כאשר הסימן נקבע לפני הנרמול. ■

הערה 66. בניסוחים אחרים, מדברים על A הרמיטית, במקום על f סימטרית. שימו לב שבכל מקרה אין משמעות למשפט מעל המרוכבים (שכן במקרה הזה $\sigma_- = 0 \wedge \sigma_+ = \operatorname{rank} f$, וכן כל מספר מרוכב נוכל לנרמל ל־1) ולכן שני הניסוחים חזקים באותה המידה.

מסקנה 24. מכאן, שבהינתן A מטריצה הרמיטית חיובית, היא מייצגת תבנית בי־לינארית חיובית וגם מייצגת העתקה חיובית. למעשה, אפשר להוכיח שבהינתן A הרמיטית, היא חיובית (בהבט של המכפלה הפנימית) אמ"מ היא חיובית (בהבט של תבניות בי־לינאריות).

משפט 159 (קיום שורש לצמודה לעצמה אי־שלילית). תהי $T\colon V \to V$ צמודה לעצמה ואי שלילית $\langle Tv \,|\, v\rangle \geqslant 0$, אז קיימת ויחידה $R\colon V \to V$ אי־שלילית צמודה לעצמה כך ש־$R^2 = T$.

הוכחה. **קיום.** מהמשפט הספקטרלי קיים בסיס א"נ של ו"ע להעתקה אי־שלילית כל הע"ע הם אי־שליליים.

$$[T]_B^B = \operatorname{diag}(\lambda_1 \ldots \lambda_n) \quad [R]_B^B = \operatorname{diag}(\sqrt{\lambda}_1 \ldots \sqrt{\lambda}_n)$$

(ראינו זאת בתרגול). עוד נבחין ש־R צמודה לעצמה כי ע"ע ממשיים.

יחידות. נבחין שכל ו"ע של T הוא ו"ע של R: יהי $i \in [n]$, ו־$B = (e_1 \ldots e_n)$ בסיס מלכסן, ואז עבור R צמודה לעצמה כלשהי מתקיים: אז ו"ע של R עם ע"ע $\sqrt{\lambda}$ הוא ו"ע של T עם ע"ע λ כי:

$$\lambda v = R^2 v = Tv \implies Rv = \sqrt{\lambda}$$

הגרירה נכונה מאי־שליליות R שהמשפט מניח עליה יחידות. כלומר הערכים העצמיים של R כלשהי (לא בהכרח זו שברחנו בהוכחת הקיום) נקבעים ביחידות מע"ע של T. בסיס של ו"ע של T הוא בסיס ו"ע של R, סה"כ ראינו איך R פועלת על בסיס ו"ע כלשהו של T מה שקובע ביחידות את R. ■

סימון 17. את ה־R לעיל נסמן $R := \sqrt{T}$.

מסקנה 25 (פירוק שולסקי). לכל A צמודה לעצמה ואי־שלילית חיובית קיים פירוק יחיד של מטריצה R משולשית עליונה כך ש־$A = RR^*$.

משפט 160 (לכסון סימולטני). מעל $\mathbb{R}$, בהינתן A מוגדרת חיובית ממש ו־B מטריצה, שתיהן סימטריות, קיים בסיס $\mathcal{P}$ בו $[A]_\mathcal{P}$ אלכסונית וגם $[B]_\mathcal{P}$ אלכסונית.

הוכחה. נפרק ספקטרלית של A ונקבל $A = P\Lambda_A P^T$. ממשפט סילבסטר Λ_A מוגדרת ביחידות ועל איבריה הסינגטורה של A, שהם כולם 1 מהיותה מוגדרת חיובית, כלומר $\Lambda_A = I$. באופן דומה נוכל לפרק ספקטרלית את PBP^T ולקבל $PBP^T = Q^T\Lambda_B Q$ ומכאן $QPBP^TQ^T = \Lambda_B$. נסמן $M = QP$ ואז $MBM^T = \Lambda_B$. בסימון $\mathcal{P} = \mathrm{Col}(M)$ נקבל $[B]_{\mathcal{P}} = \Lambda_B$, וכמו כן:

$$[A]_{\mathcal{P}} = MAM^T = Q\overbrace{PAP^T}^{I}Q^T = QIQ^T = I$$

כלומר $[A]_{\mathcal{P}}$ כדרוש. ■

(2.3.3.2) ניסוח הפירוק הפולארי

משפט 161 (פירוק פולארי בעבור העתקות). תהי $T\colon V \to V$ הפיכה, אז קיימות $R\colon V \to V$ חיובית וצמודה לעצמה ו־$U\colon V \to V$ אוניטרית כך ש־$T = RU$.

הערה 67. לא הנחנו T צמודה לעצמה. הפירוק נכון להעתקה הפיכה כללית.

הערה 68. לעיתים נקרא "פירוק UH" או "פירוק UP".

הוכחה. נגדיר $S = TT^*$. נבחין ש־S צמודה לעצמה וחיובית:

$$\forall V \ni v \neq 0\colon\ \langle Sv \,|\, v\rangle = \langle TT^*v \,|\, v\rangle = \langle T^*v \,|\, T^*v\rangle = ||T^*v|| > 0$$

האי־שוויון האחרון נכון כי $\ker T = \{0\}$, ממשפט קודם $\ker T^* = \ker T = \{0\}$, ו־$v \neq 0$. יצא שזה חיובי ולכן בפרט ממשי, כלומר היא צמודה לעצמה וחיובית.

קיימת ויחידה $R\colon V \to V$ צמודה וחיובית כך ש־$S = R^2$. כל ערכיה העצמיים של $R = \sqrt{S}$ אינם 0, ולכן היא הפיכה (ראינו בהוכחה של קיומה שהיא לכסינה יחדיו עם S).

נגדיר $U = R^{-1}T$. נותר להראות ש־U אוניטרית.

$$U^*U = (R^{-1}T)^*(R^{-1}T) = T^*\underbrace{(R^{-1})^*}_{R^{-1}}R^{-1}T = T^*(R^{-1})^2T = T^*S^{-1}T = T^*(TT^*)^{-1}T = I$$

כדרוש. הטענה $(R^{-1})^* = R^{-1}$ נכונה משום ש־R צמודה לעצמה. ■

הערה 69 (לגבי יחידות). אם T אינה הפיכה, מקבלים ש־R יחידה אבל U אינה. בשביל לא הפיכות נצטרך להצטמצם לבסיס של התמונה ועליו לפרק כמתואר לעיל. במקרה של הפיכות אז $T = RU = R\tilde{U}$ ואז נקבל R הפיכה כלומר $U = \tilde{U}$ וגם U הפיכה.

עתה נראה ש־R נקבעת ביחידות (בניגוד ליחידות U – יחידות R נכונה גם בעבור פירוק פולארי של העתקה שאיננה הפיכה):

משפט 162. בפירוק פולארי $T = RU$, כאשר R הרמיטית חיובית, אז בהכרח $R = \sqrt{TT^*}$ ולכן יחידה.

הוכחה.

$$TT^* = RU(RU)^* = RUU^*R^* = R^2$$

כלומר R היא בכל פירוק שורש של TT^*, והראינו קודם לכן את יחידות השורש. ■

מסקנה 26. קיים גם פירוק כנ"ל מהצורה $T = UR$.

הוכחה. באותו האופן שפירקנו את T, נוכל לפרק את $T^* = \tilde{R}\tilde{U}$ פירוק פולארי. נפעיל * על שני האגפים ונקבל:

$$T^* = \tilde{R}\tilde{U} \implies T = (T^*)^* = \tilde{U}^*\tilde{R}^* = \tilde{U}^{-1}\tilde{R}$$

נסמן $\tilde{R} =: R,\ \tilde{U}^{-1} =: U$ וסה"כ $T = UR$ כדרוש. ■

למה 15. עבור $T\colon V \to V$ אז ל־T^*T, TT^* נגדיר $S = TT^*$. נבחין ש־S צמודה לעצמה וחיובית:

$$\forall V \ni v \neq 0\colon\ \langle Sv \,|\, v\rangle = \langle TT^*v \,|\, v\rangle = \langle T^*v \,|\, T^*v\rangle = ||T^*v|| > 0$$

יש אותם הערכים העצמיים.

הוכחה. ניעזר בפירוק הפולארי:

$$\begin{aligned} TT^* &= RUU^*R^* \\ &= R^2 \\ TT^* &= U^{-1}R^2U \end{aligned}$$

סה"כ TT^*, T^*T הן העתקות דומות ולכן יש להן את אותם הערכים העצמיים. ■

הערה 70. אז איך זה קשור לפולארי? R האי־שלילית היא "הגודל", בעוד U האוניטרית לא משנה גודל – היא ה"זווית". ניתן לראות זאת גם באופן הבא: בהינתן $A = RU$ פירוק פולארי ל־U אורתוגונלית ו־R מוגדרת חיובית הרמיטית, אז $\det A = |\det R| =: r$ וגם $|\det U| = 1$ כלומר $\det U = e^{i\theta}$, ואז מקבלים $\det A = \det U \det R = re^{i\theta}$.

משפט 163 (פירוק פולארי בעבור מטריצות)**.** תהי $A \in M_n(\mathbb{F})$ הפיכה, אז קיימות $U, R \in M_n(\mathbb{F})$ כאשר U א"נ ו־R חיובית צמודה לעצמה כך ש־$A = UR$.

הוכחה. נסתכל על A^*A. היא חיובית וצמודה לעצמה (בדומה לעיל). אז $A^*A = P^{-1}DP$, כאשר D אלגסונית חיובית. כאשר $R^2 = AA^*$, $R = P^{-1}\sqrt{D}P$. היא קיימת ויחידה מאותה הוכחה בדיוק להעתקות. ■

2.3.4 ~ פירוק SVD

(2.3.4.1) ניסוח והוכחת SVD

הערה 71. SVD הינו קיצור של Singular Value Decomposition.

משפט 164 (גרסה מצומצמת של פירוק לערכים סינגולריים)**.** לכל מטריצה $A \in M_n(\mathbb{F})$ קיימות מטריצות אוניטריות U, V ומטריצה אלכסונית עם ערכים אי־שליליים כך ש־$A = UDV$.

הוכחה. ידוע שניתן לכתוב $A = \tilde{U}R$ פירוק פולארי. משום ש־R צמודה לעצמה, ניתן לפרקה ספקטרלית ל־V אוניטרית ו־D אלכסונית אי־שלילית (כי R אי־שלילית) כך ש־$R = V^{-1}DV$. סה"כ:

$$A = \underbrace{\tilde{U}V^{-1}}_{=:U}DV = UDV \quad \top$$

כי $\tilde{U}V^{-1}$ מכפלה של אוניטריות ולכן U אוניטרית כנדרש. ■

הערה 72. משום ש־V, U איזומטריות אז $U^* = U^{-1}, V^* = V^{-1}$ ובגלל ש־D אלכסונית אז $D^* = D$. לכן:

$$\begin{aligned} AA^* &= (UDV)(V^*D^*U^*) = UD^2U^{-1} \\ A^*A &= (V^*D^*U^*)(UDV) = V^{-1}D^2V \end{aligned}$$

הגדרה 93 (ערך סינגולרי של מטריצה)**.** הערכים העצמיים האי־שליליים של A^*A נקראים הערכים הסינגולריים והם נקבעים ביחידות ע"י A.

הגדרה 94 (ערך סינגולרי של העתקה)**.** σ הוא ערך סינגולרי של העתקה T הוא אמ"מ $\sigma \in \mathbb{R} \wedge \sigma \geqslant 0$ וגם σ^2 הוא ע"ע של TT^*.

סימון 18. את הערכים הסינגולרים של העתקה/מטריצה A כלשהי נסמן ב־$\sigma_1 \dots \sigma_n$ כאשר $\forall i \geqslant j \colon \sigma_i \geqslant \sigma_j$.

משפט 165. פירוק SVD הוא יחיד (גם למטריצה שאיננה ריבועית/הפיכה), בהנחה שהערכים הסינגולרים שונים.

הוכחה. יהיו שני פירוקי SVD של מטריצה A הפיכה כלשהי, נסמנם:

$$A = \bar{U}\bar{D}\bar{V}^T \wedge A = UDV^T$$

אזי:

$$AA^* = UD^2U^{-1} = \bar{U}^*\bar{D}^2\bar{U}^{-1} \ \wedge \ A^*A = V^{-1}D^2V = \bar{V}^{-1}\bar{D}^2\bar{V}$$

בגלל ש־$D^2, \bar{D}^2$ אלכסוניות, ומיחידות הפירוק הספקטרלי, $U = \bar{U}, V = \bar{V}, D = \bar{D}$. ■

(2.3.4.2) הרחבת SVD להעתקות שאינן אופרטורים

הערה 73. במסדרת הקורס הזה, ראינו פירוק SVD של מטריצה ריבועית בלבד. חלק מהחוזקה של פירוק SVD נובע מקיומו למטריצות שאינן בהכרח ריבועיות, דהיינו, לכל מטריצה סופית. יש לו מגוון רחב של שימושים במדעי המחשב. כדי להבין לעומק יותר כיצד פירוק SVD עובד, כתבתי את תת־הפרק הזה.

הגדרה 95. מטריצה $\Lambda \in M_{m\times n}(\mathbb{F})$ (לא בהכרח ריבועית) מוגדרת להקרא אלכסונית אמ"מ $i = j \impliedby a_{ij} \neq 0$.

משפט 166 (גרסה מורחבת של פירוק לערכים סינגולריים). תהי $A \in M_{m\times n}(\mathbb{F})$ מטריצה ($\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$) שאיננה מטריצת האפס, אז קיים פירוק למטריצות $U \in M_{n\times n}(\mathbb{F}), V \in M_{m\times m}(\mathbb{F})$ אורתונורמליות ו־$\Sigma \in M_{m\times n}$ אלכסונית, כך ש־$A = U\Sigma V^T$.

הגדרה 96. מטריצה $A \in M_{m\times n}$ מתאימה למטריצה $B \in M_{m\times n}$ אמ"מ קיימות מטריצות $U \in M_{m\times m}, V \in M_{n\times n}$ הפיכות כך ש־$A = UBV^{-1}$.

למעשה, פירוק SVD הוא התאמה אורתונורמלית ללכסינה, בדיוק כמו שפירוק ספקטרלי הוא דמיון אורתונורמלי ללכסינה (לכסון אורתונורמלי).

הוכחה. בויקיפדיה האנגלית ■

משפט 167. פירוק SVD יחיד אמ"מ הערכים הסינגולרים שונים.

מסקנה 27. תהי $T: Q \to W$. אז קיימים $\mathcal{B}, \mathcal{C}$ בסיסים אורתונורמליים כך ש־$[T]^{\mathcal{B}}_{\mathcal{C}}$ אלכסונית.

כדי להוכיח מסקנה זו, נשתמש בשורות U, V שיהוו את הבסיס האורתונרמלי הנדרש (עד לכדי העתקת קורדינאטות).

משפט 168. בהינתן B בסיס אורתונורמלי של V, והעתקה $T: V \to W$ כלשהי, σ ערך סינגולרי של T אמ"מ σ על האלכסון של Σ כאשר Σ המטריצה האלכסונית בפירוק SVD של $[T]_B$.

הוכחה. נסמן את פירוק ה־SVD של $[T]_B$ ב־$[T]_B = U\Sigma V^T$. אז ידוע $[TT^*]_B = U\Sigma^2 U^{-1}$ עבור U אורתונורמלית, ולכן $[TT^*]_B$ דומה ל־Σ^2. עוד נבחין שכל ע"ע של Σ^2 אמ"מ מופיע על אלכסון Σ^2 אמ"מ השורש שלו מופיע על אלכסון Σ. עתה נוכיח גרירה דו־כיוונית. אם σ ערך סינגולרי של T אז σ^2 הוא ע"ע של $[TT^*]$, ומהדמיון שהראנו הוא ע"ע של Σ^2 כלומר הוא מופיע על אלכסון Σ כדרוש. מהצד השני, אם σ מופיע על אלכסון Σ אז הוא ע"ע של Σ^2 ואז הוא ע"ע של $[TT^*]$, ומשום ש־Σ מוגדרת חיובית אז $\sigma \in \mathbb{R} \wedge \sigma \geqslant 0$ כדרוש. ■

מסקנה 28. מספר הערכים הסינגולריים הוא הממד של Σ והדרגה של $\operatorname{rank} A$.

הערה 74. לבדיקת שפיות, נבחין שהערכים העצמיים של Σ הם אכן "מועמדים" להיות ערכים סינגולריים, שכן היא מטריצה נורמאלית ולכן הערכים העצמיים שלה ממשיים, וכן היא מוגדרת חיובית ולכן הערכים העצמיים שלה חיוביים.

משפט 169. בהינתן $T: V \to W$ העתקה, מתקיים:

$$\sqrt{\det(T^*T)} = \prod_{i=1}^{n} \sigma_i$$

כאשר $\sigma_1 \dots \sigma_n$ הערכים הסינגולריים של T.

הוכחה. ידוע של־T קיים פירוק SVD $T = U\Sigma V^T$ ממנו נסיק את הפירוק הספקטרלי הבא ל־T^*T הצמודה לעצמה:

$$T^*T = U\Sigma^2 U^T$$

אם T איננה הפיכה אז יש לה ערך סינגולרי 0, ו־T^*T איננה הפיכה (כי מכפלת לא הפיכות איננה הפיכה) וסיימנו. אם T הפיכה, את U הפיכה בהכרח. משום ש־U אוניטרית, $U^T = U^{-1}$. נפעיל את $\det$ על שני האגפים ונקבל:

$$\det(TT^*) = \det(U)\det(\Sigma^2)\det(U^{-1}) = \det(UU^{-1})\det(\Sigma)^2 = \det(\Sigma)^2 =: *$$

בגלל שהוכחנו ש־Σ מטריצה אלכסונית שעל האלכסונה הערכים הסינגולריים של T, אז נקבל שוויון:

$$* = \left(\prod_{i=1}^{n} \sigma_i\right)^2$$

נוציא שורש ונקבל את הנדרש. ■

מסקנה 29. עבור T ריבועית, נוכל לטעון:

$$\left(\prod_{i=1}^{n} \sigma_i\right) = \det(TT^*) = \det(T)\det(\bar{T}^T) = \det(T)\det(\bar{T}) = \det T\overline{\det T} = \det T^2$$

נוציא שורש ונקבל שהדטרמיננטה של T היא מכפלת הערכים הסינגולריים:

$$\prod_{i=1}^{n}\sigma_i=\det T$$

משפט 170 (פירוק העתקה לערכים סינגולריים). בהינתן $T\colon V\to W$ כלשהי, וערכים סינגולריים $\sigma_1\dots\sigma_r$ כלשהם, אז קיימים $\mathbf{u}_1\dots\mathbf{u}_r\in V$ ו־$\mathbf{v}_1\dots\mathbf{v}_r\in W$ כך ש־:

$$Tv=\sum_{i=1}^{r}\sigma_i\left\langle v\,|\,\mathbf{u}_i\right\rangle\mathbf{v}_i$$

הוכחה. נסמן $\dim V=n,\dim W=m$ בהינתן בסיסים $\mathcal{B},\mathcal{C}$ אורתונורמליים ל־V,W בהתאמה כך ש־$|\mathcal{B}|=n\wedge|\mathcal{C}|=m$, ידוע קיום פירוק של $[T]_{\mathcal{C}}^{\mathcal{B}}$ לערכים סינגולריים כך ש־:

$$[T]_{\mathcal{C}}^{\mathcal{B}}=V\Sigma U^T$$

כאשר V,U אוניטריות ו־Σ אלכסונית. ממשפט ידוע שעל אלכסון Σ מופיעים $\sigma_1\dots\sigma_n$. בגלל ש־V מטריצה עם r שורות ב־$\mathbb{R}^m$ מהן נצמצם שורות בת"ל $V_1\dots V_r$ ובאופן דומה U מטריצה עם שורות $W_1\dots W_r$ ב־$\mathbb{R}^n$. נוכל להניח שהשורות הבת"ל במטריצות האוניטריות יהיו הראשונות, שכן הערכים הסינגולריים על המטריצה האלכסונית Σ מופיעים לפני שורות/עמודות האפסים (אם יש) ב־Σ. כעת נוכל להגדיר (כאשר $[]_B^{-1}$ ההעתקה ההופכית לייצוג בבסיס B)

$$\mathbf{u}_i=[U_i]_{\mathcal{B}}^{-1}\quad\mathbf{v}_i=[V_i]_{\mathcal{C}}^{-1}$$

ועתה נשאר להראות שהבחירה שלנו אכן עובדת. יהי $v\in V$, ונסמן $[v]_{\mathcal{B}}=(a_1\dots a_n)$. כאשר $\mathcal{E}=(e_1\dots e_n)$ הבסיס הסטנדרטי ל־$\mathbb{F}^n$ ו־$\mathcal{E}=(e_1\dots e_m)$ הסטנדרטי ל־$\mathbb{F}^m$, נקבל:

$$\begin{aligned}
[Tv]_{\mathcal{C}}=[T]_{\mathcal{C}}^{\mathcal{B}}[v]_{\mathcal{B}}&=V\Sigma U^T\cdot(a_1\dots a_n)\\
&=\sum_{i=1}^{n}V\Sigma\overbrace{U^Te_i}^{U_i}a_i=V\Sigma\sum_{i=1}^{r}a_1U_i\\
&=V\Sigma\sum_{i=1}^{r}\sum_{j=1}^{n}a_i\left\langle e_j\,|\,U_i\right\rangle e_j&&\text{ייצוג של } U_i \text{ ב־} \mathbb{R}^n \text{ בבסיס } \mathcal{E}\\
&=V\Sigma\sum_{i=1}^{r}\left\langle[v]_{\mathcal{B}}\,|\,U_i\right\rangle e_i&&\text{לינאריות ברכיב הראשון}\\
&=\sum_{i=1}^{r}\left\langle[v]_{\mathcal{B}}\,|\,U_i\right\rangle\underbrace{V\overbrace{\Sigma e_i}^{\sigma_ie_i}}_{(Ve_i)\sigma_i=V_i\sigma_i}=\sum_{i=1}^{r}\sigma_i\left\langle[v]_{\mathcal{B}}\,|\,U_i\right\rangle V_i
\end{aligned}$$

משום ש־$\mathcal{B}$ בסיס אורתונורמלי, אז מעבר מבסיס $\mathcal{E}$ ל־$\mathcal{B}$ ולהפך הוא אוניטרי, כלומר $\left\langle[v]_{\mathcal{B}}\,|\,U_i\right\rangle=\left\langle[[v]_{\mathcal{B}}]_{\mathcal{B}}^{-1}\,\middle|\,[U_i]_{\mathcal{B}}^{-1}\right\rangle=\left\langle v\,|\,\mathbf{u}_i\right\rangle$. כעת נפעיל את $[]_B^{-1}$ על שני האגפים, ונקבל:

$$Tv=\left[\sum_{i=1}^{r}\sigma_i\left\langle v\,|\,\mathbf{u}_i\right\rangle V_i\right]_{\mathcal{B}}^{-1}=\sum_{i=1}^{r}\sigma_i\left\langle v\,|\,\mathbf{u}_i\right\rangle[V_i]_{\mathcal{B}}^{-1}=\sum_{i=1}^{r}\sigma_i\left\langle v\,|\,\mathbf{u}_i\right\rangle\mathbf{v}_i$$

כדרוש. ■

מסקנה 30. בהינתן $g_1\dots g_r,f_1\dots f_r$ לעיל, אז:

$$T^*v=\sum_{i=1}^{r}\sigma_i\left\langle v\,|\,\mathbf{v}_i\right\rangle\mathbf{u}_i$$

הוכחה. ניעזר פעמיים באדטיביות רכבי המכפלה הפנימית:

$$\left\langle Tv\,|\,w\right\rangle=\left\langle\sum_{i=1}^{r}\sigma_i\left\langle v\,|\,\mathbf{u}_i\right\rangle\mathbf{v}_i\,\middle|\,w\right\rangle=\sum_{i=1}^{r}\sigma_i\left\langle v\,|\,\mathbf{u}_i\right\rangle\left\langle\mathbf{v}_i\,|\,w\right\rangle=\left\langle v\,\middle|\,\underbrace{\sum_{i=1}^{r}\sigma_i\left\langle\mathbf{v}_i\,|\,w\right\rangle\mathbf{u}_i}_{T^*w}\right\rangle\quad\top$$

■

(2.3.4.3) נורמה של העתקה

הערה 75. גם תת־הפרק להלן לא בחומר הרשמי של הקורס. אבל חשבתי שזה מגניב אז הוספתי את זה.
הגדרה 97. הנורמה של העתקה $T\colon V\to W$ ממכ"פים מוגדרת להיות:

$$||T||=\max\{||Tv||:v\in V\wedge ||v||\leqslant 1\}$$

הערה 76. אינטואציה גיאומטרית טובה היא לחשוב על $||T||$ כעל "הכדור המינימלי שחוסם את Tu" כאשר u נורמלי.
למה 16. כזכור, σ_1 הערך הסינגולרי המקסימלי של T. אז:

$$||Tv||\leqslant \sigma_1\,||v||$$

הוכחה. מפירוק העתקה לערכים סינגולרים:

$$Tv=\sum_{i=1}^{r}\sigma_i\left\langle \sigma_i\,|\,\mathbf{u}_i\right\rangle\mathbf{v}_i$$

משום ש־$\mathbf{v}_i$ אורתונורמליים, אז $||\mathbf{v}_i||=1$. לכן:

$$||Tv||=\sum_{i=1}^{n}\sigma_i\left\langle v\,|\,\mathbf{u}_i\right\rangle\leqslant\sum_{i=1}^{n}\sigma_1\left\langle v\,|\,\mathbf{u}_i\right\rangle=\sigma_1\left(\sum_{i=1}^{n}\left\langle v\,|\,\mathbf{u}_i\right\rangle\right)=:*$$

ממשפט, בגלל ש־g_i בסיס אורתונורמלי אז $\sum_{i=1}^{n}\left\langle v\,|\,\mathbf{u}_i\right\rangle\mathbf{u}_i=v$. דהיינו בגלל ש־$||g_i||=1$:

$$*=\sigma_1\left|\left|\sum_{i=1}^{n}\left\langle v\,|\,\mathbf{u}_i\right\rangle\mathbf{u}_i\right|\right|=\sigma_1\,||v||$$

וסה"כ אכן $||Tv||\leqslant\sigma_1\,||v||$ כדרוש. ■

משפט 171. הנורמה של ההעתקה היא פונקציה חיובית ופחות או יותר לינארית:

1. $||T||\geqslant 0$
2. $||T||=0\iff T=0$
3. $||\lambda T||=|\lambda|\,||T||$
4. $||S+T||=||S||+||T||$
5. $||T||=||T^*||$

משפט 172. כאשר σ_1 הערך הסינגולרי הגדול ביותר של T, אז $||T||=\sigma_1$.
הערה 77. שני המשפטים הבאים לא טרוויאלים אך מובאים כאן ללא הוכחה, לידע כללי בלבד.
משפט 173. בהינתן $T\colon V\to W$ ו־$\sigma_1\dots\sigma_n$ ערכים סינגולריים, אז:

$$\min\{||T-S||:S\in V\to W\wedge \operatorname{rank}S\leqslant k\}=\sigma_{k+1}$$

משפט 174 (משפט המינ־מקס). לכל $k\in[n]$, כאשר S מ"ו:

$$\sigma_k=\max_{\dim S=k}\ \min_{x\in S\wedge||x||=1}||Tx||$$

ובאופן שקול (ודי הגיוני):

$$\sigma_k=\min_{\dim S=n-k+1}\ \underbrace{\max_{x\in S\wedge||x||=1}||Tx||}_{||T|_S||}$$

באופן כללי, ערכים סינגולרים משמשים כדי להגדיר נורמות רבות על העתקות.

המשך בעמוד הבא

פרק 3

נספחים

3.1 Dual Spaces

את הפרק להלן המרצה של אודיסאה, בן בסקין, החליט להעביר, כדי לתת ראיה נרחבה יותר על לינארית - מנקודת מבט של תורת הקטגוריות. הרעיון הוא להבחין בכך ש־(א) בין כל קטגוריה לדואל שלה קיים פונקטור קונטראינווריאנטי, ו־(ב) הן הצמדת העתקה, והן פונקציונל, הן קטגוריות דואליות למרחב הוקטורי, ולכן איזומורפיות אחת לשנייה (שכן הדואל יחיד עד לכדי איזומורפיזם).

3.1.1 ∼ הגדרות בסיסיות

הגדרה 98. בהינתן V מ"ו מעל $\mathbb{F}$, נגדיר $V^*=\hom(V,F)$.

הבנה. אם $\dim V=n$ אז $\dim V^*=n$. לכן $V\cong V^*$. לא נכון במקרה הסוף ממדי.

למה 17. יהי $B=(v_i)_{i=1}^n$ בסיס ל־V. אז $\forall i\in[n]\colon \exists \psi_i\in V^*\colon \forall j\in[n]\colon \psi_i(v_j)=\delta_{ij}$

משפט 175. יהי V נ"ס ו־$B=(v_i)_{i=1}^n$ אז קיים ויחיד בסיס $B^*=(\psi_i)_{i=1}^n$ המקיים $\forall i,j\in[n]\colon \psi_i(v_j)=\delta_{ij}$.

הוכחה. נבחין שהבדרנו העתקה לינארית $\varphi\colon B\to V^*$ והיא מגדירה ביחידות ψ לינארית $\psi\colon V\to V^*$ המקיימת את הנרש. ברור שהבנייה של φ_i קיימת ויחידה כי היא מוגדר לפי מה קורה לבסיס. נותר להוכיח שזה אכן בסיס. יהיו $\alpha_1\dots\alpha_n\in\mathbb{F}$ כך ש־$\sum\alpha_i\psi_i=0$. (האפס הזה הוא פונקציונל האפס). יהי $j\in[n]$. אז $0(v_j)=0=\left(\sum_{i=0}^n\alpha_i\psi_i\right)v_j=\sum_{i=0}^n\alpha_i\psi_i(v_j)=\alpha_j\sum\alpha_i\delta_{ij}=\alpha_j$ וסה"כ $\alpha_j=0$. ■

נבחין שאפשר להגדיר:

הגדרה 99. $V^{**}=\hom(V^*,\mathbb{F})$

ואכן $\dim V<\inf$ אז:

$$V\cong V^*\cong V^{**}$$

במקרה הזה, בניגוד לאיזו' הקודם, יש איזו' "טבעי" (קאנוני), כלומר לא תלוי באף בסיס.

משפט 176. קיים איזומורפיזם קאנוני בין V ל־V^{**}.

הוכחה. נגדיר את האיזו' הבא:

$$\psi\colon V\to V^{**}\quad \psi(v)=:\bar{v}\quad \forall\psi\in V^*\colon \bar{v}(\psi)=\psi(v)$$

נוכיח שהוא איזו':

- **ט"ל:** יהיו $v,u\in V$, $\alpha,\beta\in\mathbb{F}$. אז:

$$\psi(\alpha v+\beta u)=\overline{\alpha v+\beta u}\stackrel{!}{=}\alpha\bar{v}+\beta\bar{u}$$

נוכיח זאת:

$$\overline{\alpha v+\beta u}(\varphi)=\varphi(\alpha v+\beta u)=\alpha\varphi(v)+\beta\varphi(u)=\alpha v(\varphi)+\beta\bar{u}(\varphi)=(\alpha\bar{v}+\beta\bar{u})(\varphi)=(\alpha\psi(v)+\beta\psi(u))(\varphi)$$

- **חח"ע:** יהי $v\in\ker\psi$. רוצים להראות $v=0$.

$$\forall\varphi\in V^*\colon \bar{(}\varphi)=0\implies \forall\varphi\in V^*\colon \varphi(v)=0$$

אם v אינו וקטור האפס, נשלימו לבסיס $V=(v_i)_{i=1}^n$ ואם $\varphi_1\dots\varphi_n$ בסיס הדואלי אז $\varphi_1(v)=1$ אבל אז $0=\bar{v}(\varphi_1)=1$ וסתירה.

- **על:** משוויון ממדים $\dim V^{**}=\dim V$.

■

כלומר, הפונקציונלים בדואלי השני הם למעשה פונקציונלים שלוקחים איזשהו פונקציונל בדואלי הראשון ומציבים בו וקטור קבוע.

3.1.2 ~ איזומורפיות למרחבי מכפלה פנימית

(3.1.2.1) העתקה צמודה (דואלית)

סימון 19. לכל $v \in V$ ו־$\varphi \in V^*$ נסמן $\varphi(v) = (\varphi, v)$
הערה 78. סימון זה הגיוני משום שהכנסת וקטור לפונרציונל דואלי איזומורפי למכפלה פנימית.
משפט 177. יהיו V, W מ"וים נוצרים סופית מעל $\mathbb{F}$, $T\colon V \to W$. אז קיימת ויחידה $T^*\colon W^* \to V^*$ כך ש־$(\psi, T(v)) = (T^*(\psi), v)$.

אם לצייר דיאגרמה:

$$V \xrightarrow{T} W \cong W^* \xrightarrow{T^*} V^* \cong V$$

(תנסו לצייר את זה בריבוע, ש־V, W למעלה ו־V^*, W^* למטה, כדי להבין ויזולאית למה זה הופך את החצים)

ברמה המטא־מתמטית, בתורת הקטגוריות, יש דבר שנקרא פנקטור – דרך לזהות בין אובייקטים שונים במתמטיקה. מה שהוא עושה, לדוגמה, זה להעביר את $\hom(V, W)$ – מרחבים וקטרים סוף ממדיים – למרחב המטריצות. הדבר הזה נקרא פנקטור קו־וראיינטי. במקרה לעיל, זהו פנקטור קונטרא־ווריאנטי – שימוש ב־T^* הופך את החצים. (הרחבה של המרצה)

אז אפשר להגדיר פנקטור אבל במקום זה נעשה את זה בשפה שאנחנו מכירים – לינארית 1א. בהינתן $\psi \in W^*$, נרצה למצוא $T^*(\psi) \in V^*$. נגדיר:

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

ברור מדוע $T^*\colon W^* \to V^*$. בעצם, זהו איזומורפיזם ("בשפת הפנקטורים") קאנוני. עוד קודם לכן ידענו (בגלל ממדים) שהם איזומורפים, אך לא מצאנו את האיזומורפיזם ולא ראינו שהוא קאנוני.

$$\tau\colon \hom(V, W) \to \hom(W^*, V^*) \quad \tau(T) = T^*$$

היא איזומורפיזם.

(הערה: תודה למרצה שנענה לבקשתי ולא השתמש ב־/phi אחרי שעשיתי {/renewcommand/phi{/varphi)

הוכחת לינאריות. יהיו $T, S \in \hom(V, W)$, $\alpha \in \mathbb{F}$. אז:

$$\tau(\alpha T + S) = (T + \alpha S)^*$$

יהי $\psi \in W^*$, אז:

$$(T + \alpha S)^*(\psi) = \psi \circ (T + \alpha S)$$

יש למעלה פונקציונל ב־V^*. ננסה להבין מה הוא עושה על V. יהי $v \in V$:

$$\begin{aligned}
[\psi(T + \alpha S)](v) &= \psi((T + \alpha S)v) \\
&= \psi(T)v + \alpha\psi(S)(v) = (\psi \circ T + \alpha\psi \circ S)(v) \\
&= ((T^* + \alpha S^*) \circ (\psi))v \\
&= (\tau(T) + \alpha\tau(S))(\psi)(v)
\end{aligned}$$

סה"כ קיבלנו לינאריות:

$$\tau(T + \alpha S) = \tau(T) + \alpha\tau(S)$$

■

נוכל להוכיח זאת יותר בפשטות עם הנוטציה של "המכפלה הפנימית" שהגדרנו לעיל, (φ, v). עתה נוכיח ש־τ לא רק לינארית, אלא מוגדרת היטב.

הוכחה.

- **חח"ע:** תהי $T \in \ker\tau$, אז $\tau(T) = T^* = 0_{\hom(W^*,V^*)}$. נרצה להראות ש־$T$ העתקה האפס. נניח בשלילה ש־$T \neq 0$. אז קיים $v' \in V$ כך ש־$T(v') \neq 0$. נשלימו לבסיס – $(T(v) = w_1, w_2 \dots w_n)$ בסיס ל־W. יהי $(\psi_1 \dots \psi_n)$ הבסיס הדואלי. אז:

$$\tau(T)(\psi_1) = T^*(\psi_1) = \psi_1$$

אז:

$$0 = \tau(T)(\psi_1)(v') = T^*(\psi_1)(v') = \psi_1 \circ T(v') = \psi_1(w_1) = 1$$

סתירה. לכן $\ker\tau = \{0\}$ ולכן τ חח"ע.

- **על:** גם כאן משוויון ממדים

■

שאלה ממבחן שבן עשה. ("את השאלה הזו לא פתרתי בזמן המבחן, ואני די מתבייש כי אפשר לפתור אותה באמצעות כלים הרבה יותר פשוטים" "חה חה" "לא חח"ע זה חד־חד ערכי") יהיו V, W מ"ו מעל $\mathbb{F}$ ו־$(w_1 \dots w_n)$ בסיס של W. תהי $T\colon V \to V$. הוכיחו שקיימים $\varphi_1 \dots \varphi_n \in V^*$ כך שלכל $v \in V$ מתקיים:

$$T(v) = \sum_{i=0}^{n} \varphi_i(v) w_i$$

שימו לב: בניגוד למה שבן עשה במבחן, V לא בהכרח נוצר סופית.

הוכחת ראש בקיר. לכל $v \in V$ קיימים ויחידים $\alpha_1 \dots \alpha_n$ כך ש־$T(v) = \sum_{i=0}^{n} \alpha_i w_i$. נגדיר $\forall i \in [n]\colon \varphi_i(v) = \alpha_i$. זה ליניארי. ■

הוכחה "מתוחכמת". נתבונן בבסיס הדואלי $B^* = (\psi_1 \dots \psi_n)$ שמקיים את הדלתא של קרונקר והכל. נגדיר $\varphi_i =: T^*(\psi_i)$. יהי $v \in V$. קיימים ויחידים $\alpha_1 \dots \alpha_n$ כך ש־$T(v) = \sum_{i=0}^{n} \alpha_i w_i$. אז:

$$\sum_{i=0}^{n} \varphi_i(v) w_i = \sum T^*(\psi_i)(v) w_i$$

צ.ל. $\alpha_i = T^*(\psi_i)(v)$. אך נבחין שהגדרנו:

$$T^*(\psi_i)(v) = \psi_i(T(v)) = \psi_i\left(\sum_{i=0}^{n} \alpha_j w_j\right) = \alpha_j$$

■

"הפכת למרצה במתמטיקה כדי להתנקם באחותך?" "כן."

(3.1.2.2) המאפס הדואלי ומרחב אורתוגונלי

הגדרה 100. יהי V מ"ו נוצר סופית. יהי $S \subseteq V$ קבוצה. נגדיר $S^0 \subseteq V^* =: \{\varphi \in V^* \mid \forall v \in S\colon \varphi(v) = 0\}$.

דוגמאות.

$$\{0\}^0 = V^*,\ V^0 = \{0\}$$

משפט 178.

1. S^0 תמ"ו של V^*.
2. $$(\operatorname{span} S)^0 = S^0$$
3. $$S \subseteq S' \implies S'^0 \subseteq S^0$$

משפט 179. יהי V נ"ס, $U \subseteq V$ תמ"ו. אז $\dim U + \dim U^0 = n$

באופן דומה אפשר להמשיך ולעשות:

$$\dim U^0 + \dim U^{**} = n$$

וכן:

$$U \cong U^{**}$$

איזומורפיזם קאנוני. זאת כי:

$$\forall\varphi \in U^0 \, \forall u \in U : \varphi(u) = 0$$

ומי אלו הוקטורים שיאפסו את φ שמאפס את u? הוקטורים ב־U עד לכדי האיזומורפיזם הקאנוני מ־U ל־U^{**}.
נבחין ש־:

$$[T^*]^{\mathcal{B}^*}_{\mathcal{A}^*} = ([T]^{\mathcal{A}}_{\mathcal{B}})^T$$

כאשר $\mathcal{A}$ בסיס ל־V, $\mathcal{A}^*$ ל־V^*, $\mathcal{B}$ ל־W, $\mathcal{B}^*$ ל־W^*.

המשך בעמוד הבא

Summary of Notable Result 3.2

3.2.1 ~ סיכום פירוק פרימרי וצורת ג'ורדן

התחלנו בלנסות ללכסן מטריצות. הבחנו שלא כל מטריצה היא לכסינה, ובהן מתקיים $r_\lambda < d_\lambda$ עבור ע"ע כלשהו. את הבעיה הזו תקפנו בשני כיוונים:

- מצאנו את **משפט הפירוק הפרימרי**, שאומר שבהינתן פירוק של הפולינום המינימלי למכפלת פולינומים זרים $m_T(x) = g_1 \dots g_r$, אז $V = \bigoplus_{i=1}^r \ker g_i(T)$ וגם $m_{T|_{(\ker g_i(T))}} = g_i$ כלומר g_i הפולינום המינימלי של צמצום T על המ"ו $\ker g_i(T)$. גם ראינו שאם במקום $m_T(x)$ משתמשים בפולינום כלשהו f המאפס את T (כלומר $f(T) = 0$) החלק הראשון של המשפט עדיין מתקיים.

  אז פשוט זרקנו את המשפט הזה על העתקה כללית, מעל שדה סגור אלגברית, ואז ראינו שבכל תמ"ו שפירקנו אליו אפשר להגדיר העתקה נילפוטנטית מהצורה $T - \lambda I$. את המקרה של נילפוטנטיות חקרנו בנפרד, וגילינו שאפשר לייצג העתקה נילפוטנטית כמטריצת בלוקים $\operatorname{diag}(J_{x_1}(0) \dots J_{x_k}(0))$. כששירשרנו את הבסיסים והוספנו את ה־λI בחזרה, קיבלנו את צורת ג'ורדן המתבקשת.
- בצורה אחרת עשינו בדיוק את אותו הדבר. אך במקום להבטיח את משפט הפירוק הפרימרי ולגלות שדברים עובדים, ניסינו להבין לאילו בדיוק תתי־מרחבים המרחב מתפרק. מצאנו שהמרחבים האלו הם **המרחבים העצמיים המורחבים** של T (והוכחנו את זה ללא תלות במשפט הפירוק הפרימרי), ועליהם כבר יכולנו להגיד הרבה יותר דברים. לדוגמה:
  - הגודל של מ"ו מורחב המשוייך לע"ע λ הוא הריבוי האלגברי, ולכן זוהי כמות הו"עים המשוייך לאותו הע"ע.
  - העתקה הנילפוטנטית $T - \lambda I$ בצמצום על המ"ו הזה, בעלת דרגת נילפונטיות שהיא הריבוי של λ ב־m_T, ולכן זהו גודל בלוק הג'ורדן המירבי עם ע"ע λ.
  - כל וקטור ב־$V(\lambda)$ המ"ו העצמי הלא מורחב פותח שרשרת אחרת, ולכן כמות בלוקי הג'ורדן לע"ע λ הוא $\dim V(\lambda) = r_\lambda$ הריבוי הגיאומטרי.

הקטע הכיפי, הוא שצורת הג'ורדן היא יחידה עד כדי סדר בלוקים. לכן, כל המסקנות שלנו לגבי איך נראית צורת ג'ורדן שפיתחנו בשיטה כזו או אחרת, תקפות למעשה לכל צורת ג'ורדן של ההעתקה/מטריצה.

"על הדרך", קיבלנו כל מני תוצאות מעניינות:

- אם הפולינום האופייני מתפרק, האופרטור ניתן לשילוש (בפרט כל אופרטור ניתן לשילוש מעל שדה סגור אלגברית).
- הפולינום המינימלי מתפרק לגורמים לינאריים שונים, אם ורק אם המטריצה לכסינה, אם ורק אם $m_T = f_T^{\text{red}}$.
- הבחנו בקיום המטריצה המצורפת A_f, שהראתה לנו שלכל פולינום מתוקן קיימת מטריצה שהוא הפולינום האופייני שלה (הגדרתה מופיע בהמשך הסיכום).
- ג'ירדון ולכסון הוכחו ככלים יעלים לפתירת נוסחאות נסיגה לינאריות.

בדרך, עברנו דרך תורת החוגים בעיקר כדי לצאת עם שתי הטענות הבאות:

- חוג הפולינומים הוא תחום אוקלידי (ובפרט ראשי), מה שמאפשר לנו לחלק פולינומים עם שארית.
- קבוצת הפולינומים המאפסים של T היא אידיאל, ומהיות חוג הפלינומים תחום ראשי, הוא נוצר על ידי פולינום מסויים שסימנו ב־m_T (שמהגדרה הוא המינימלי ביחס ההכלה).

3.2.2 ~ סיכום תבניות בי־לינאריות

התעניינו באופן מיוחד בשלושה סוגים של תבניות בי־לינאריות:

- **תבנית חיובית,** (או אי־שלילית וכו'), כזו המקיימת $f(v,v) \geqslant 0$ לכל $v \in V$, מה ששקול להיות התבנית הריבועית שהיא מגדירה, חיובית גם היא. תבנית היא חיובית אם הסינגטורה $\sigma_+ = n$.
- **תבנית סימטרית.** הבחנו שכל תבנית אפשר לפרק לחלק סימטרי וחלק אנטי־סימטרי, ותבנית ריבועית מתייחסת לחלק הסימטרי בלבד (ואף שיש זיווג בין תבניות סימטריות לריבועיות). הבחנו שהמייצגת של תבנית כזו, סימטרית גם. ראינו שאם נשלב את ההנחה של סימטריות עם תבנית מוגדרת חיובית, אז נקבל מכפלה פנימית.
- **תבנית לא־מנוונת,** שמתקבלת ישירות מהגדרת הראדיקל של התבנית. ראינו שתבנית היא לא־מנוונת אמ"מ המטריצה המייצגת שלה הפיכה.

הבחנו שבמידה והמטריצה המייצגת הרמיטית, אז הסימן של הערכים העצמיים קובע את הסינגטורה (זאת כי פירוק ספקטרלי הוא לא רק דמיון, אלא גם חפיפה!).

3.2.3 $\sim$ סיכום נושא הפירוקים

יש לנו מספר סוגי העתקות שעניינו אותנו באופן מיוחד:

$\mathbb{R}/\mathbb{C}$	נורמאלית	אורתוגונלית/אוניטרית	הרמיטית/סימטרית
הגדרה	$TT^* = T^*T$	$T^* = T^{-1}$	$T^* = T$
מ"פ	$\langle Tv \mid Tu \rangle = \langle T^*v \mid T^*u \rangle$	$\langle Tv \mid Tu \rangle = \langle v \mid u \rangle$	$\langle Tv \mid u \rangle = \langle v \mid Tu \rangle$
ע"עים	$Tv = \lambda v \iff T^*v = \lambda v$	$\lvert\lambda\rvert = 1$	$\lambda \in \mathbb{R}$

כאשר העתקה אוניטרית/אורתוגונלית נקראת באופן כללי **איזומטריה לינארית**. להגדרות אילו, נלווים המשפטים הבאים:

- **משפט הפירוק הספקטרלי ב־$\mathbb{R}$:** העתקה היא סימטרית אמ"מ היא לכסינה אורתונורמלית.
- **משפט הפירוק הספקטרלי ב־$\mathbb{C}$:** אם העתקה היא נורמאלית, אז היא לכסינה אורתונורמלית.
- T לכסינה אותורנורמלית אמ"מ T לכסינה אורתוגונלית (תוצאה ישירה מנרמול).

וההבחנה שאם A מטריצה לכסינה אורתונורמלית (או מייצגת העתקה לכסינה אורתונורמלית), אז קיימת מטריצת מעבר בסיס U אוניטרית/אורתוגונלית ו־Λ אלכסונית כך ש־$A = U\Lambda U^*$. בפרט הפירוק הספקטרלי של מטריצה הניתנת ללכסון אונטירי, הוא פירוק ה־SVD שלה.

יש לנו שתי הגדרות לחיוביות (ושליליות, וכיו"ב):

- **מטריצה מוגדרת חיובית:** אמ"מ היא מייצגת תבנית בי־לינארית חיובית, כלומר $\forall x \in \mathbb{F}^n : x^T Ax > 0$.
- **מטריצה חיובית:** הגדרה מצחיקה שלא מקובלת בשום מקום אחר חוץ מבקורס הזה, ודורשת ש־$\langle Av \mid v \rangle > 0$. כל העתקה היא חיובית אמ"מ תחת ייצוג בבסיס אורתונורמלי, המטריצה המיייצגת חיובית.

למזלנו, ההגדרות מתלכדות במקרה של ההעתקה או מטריצה צמודה לעצמה. זהו המקרה הרלוונטי ל**פירוק פולארי** שמפרק מטריצה כללית A לכפל של מטריצה אוניטרית U ומטריצה הרמיטית מוגדרת־חיובית P כך ש־$A = UP$ ("לסובב ואז לשנות גודל"). למעשה, לא דיברנו בקורס כלל על מטריצה חיובית (בהבט של $\langle Av \mid v \rangle > 0$) שאיננה צמודה לעצמה, וכאמור במקרה זה היא בכל מקרה מוגדרת־חיובית.

כבר ידוע שהע"עים של מטריצה צמודה לעצמה (כלומר הרמיטית/סימטרית) הם בהכרח ממשיים. אבל במקרה של מטריצה מוגדרת, נוכל לטעון שמטריצה מוגדרת חיובית אמ"מ הע"עים אף חיוביים (ובאופן דומה לגבי א־שלילית, שלילית, וא־חיובית)!

בגלל שכל מכפלה פנימית היא סימטרית ובפרט צמודה לעצמה, אז כדי לקבוע שהמכפלה הפנימית אכן חיובית, יש רק צורך ללכסן את התכנית הריבועית כדי לוודא שהיא אכן מכפלה פנימית.

המשך בעמוד הבא

3.3 Algorithms

הנושא הזה מסכם בקצרה אלגוריתמים מועילים שרואים בתרגולים וכדאי לזכור (**אין** כאן סיכום מלא של התרגולים).

א' **לכסון:** בהינתן $A \in M_n(\mathbb{F})$ מטריצה.

- נחשב את f_A
- נמצא את שורשי f_A. אם אנו מתקשים למצוא את שורשי הפולינום, נמצא את f^{red}.
- לכל ע"ע λ_i, נמצא בסיס למרחב העצמי באמצעות חישוב $\mathcal{N}(\lambda_i I - A)$. איברי הבסיס יהיו הו"עים בעבור הע"ע λ_i.
- סה"כ $\operatorname{diag}(\lambda_1 \dots \lambda_n)$ המטריצה האלכסונית המתקבלת ע"י מטריצת מעבר הבסיס הנתונה ע"י הו"עים מהשלב הקודם.

ב' **ג'רדון מטריצה כללית:** תהי $A \in M_n(\mathbb{F})$ מטריצה שהפולינום האופייני שלה $f_A(x) = \prod_{j=1}^{m}(x-\lambda_j)^{r_j}$ (בה"כ מתקיים מעל הרחבה לשדה סגור אלגברית). לכל $j \in [m]$ נבצע את הפעולות הבאות:

- נמצא את הפולינום $f_A(x)$ האופייני ונפרק אותו לכדי גורמים לינאריים.
- נחשב את $V_{\lambda_j}^{(i)} := \mathcal{N}((A-\lambda_j)^{\ell_j})$ עד שנקבל $\dim\left(V_{\lambda_j}^{(\ell_j)}\right) = m_i$ (המרחב העצמי המוכלל).

  הערה: אפשר באופן חלופי לחשב את הפולינום המינימלי, שכן ראינו ש־m_i הריבוי של λ_i ב־$m_T(x)$.
- נחזור על האלגו' למציאת צורת ג'ורדן למטריצה נילפוטנטית:
  - נגדיר $B_{\lambda_j} = \varnothing$
  - לכל $i \in [\ell_j]$ נבצע:
    * נמצא בסיס כלשהו $C_{\lambda_j}^{(i)-}$ של $V_{\lambda_j}^{(i-1)}$.
    * נוסיף ל־$C_{\lambda_j}^{(i)-}$ את $B \cap (V_{\lambda_j}^{(i)} \setminus V_{\lambda_j}^{(i-1)})$
    * נשלים את $C_{\lambda_j}^{(i)-}$ לבסיס של $V_{\lambda_j}^{(i)}$. נסמן ב־$C_{\lambda_j}^{(i)+}$.
    * נוסיף ל־B_{λ_j} את $\left\{(A-\lambda_j I)^k v \mid 0 \leqslant k < i,\ v \in C_{\lambda_j}^{(i)+}\right\}$.
  - נגדיר $B = \bigcup_{j=1}^{m} B_{\lambda_j}$ הבסיס המג'רדן.

ג' **מציאת $\boldsymbol{J_n(\lambda)^m}$:** ידוע $J_n(\lambda) = \lambda I_n + J_n(0)$ ולכן מהיות $\lambda I, J_n(0)$ מתחלפות נקבל מנוסחת הבינום של ניוטון:

$$(J_n(\lambda)^m)_{ij} = \begin{cases} 0 & j > i \\ 0 & j < i - m \\ \binom{m}{i-j}\lambda^{m-(i-g)} & \text{else} \end{cases}$$

דהיינו:

$$J_n(\lambda)^m = \begin{pmatrix} \binom{m}{0}\lambda^m & & & & & \\ \binom{m}{1}\lambda^{m-1} & \binom{m}{0}\lambda^m & & & & \\ \vdots & \ddots & \ddots & & & \\ \binom{m}{m}\lambda^0 & \binom{m}{m-1}\lambda^1 & \ddots & \ddots & & \\ 0 & \binom{m}{m}\lambda^0 & \cdots & \ddots & \binom{m}{0}\lambda^m & \\ \vdots & 0 & \binom{m}{m}\lambda^0 & \cdots & \binom{m}{m-1}\lambda^1 & \binom{m}{0}\lambda^m \\ 0 & \vdots & \ddots & \ddots & & \end{pmatrix}$$

ד' **גראם־שמידט:** נרצה למצוא בסיס אורתונורמלי/אורתוגונלי לממ"פ כלשהו. יהי בסיס $B = v_1 \dots v_n$ של V.

- **למציאת בסיס אורתוגונלי:** נגדיר לכל $i \in [n]$

$$\tilde{v}_i = v_i - \sum_{j=1}^{i-1} \frac{\langle v_i \mid \tilde{v}_j \rangle}{\left\|v_k\right\|^2} \cdot \tilde{v}_j$$

ואז $(\tilde{v}_1 \dots \tilde{v}_n)$ בסיס אורתוגונלי (הבחנה: התהליך רקורסיבי, נתחיל מ־$i=1$ ונסיים ב־$i=n$). במידת הצורך נוכל לנרמל בסוף ע"י הגדרת:

$$\bar{v}_i = \frac{\tilde{v}_i}{||\tilde{v}_i||}$$

ואז $(\bar{v}_1 \dots \bar{v}_n)$ אורתונורמלי מסיבות ברורות.

- **מציאת בסיס אורתונורמלי:** (פחות יציב נומרית מאשר למצוא אורתונורמלי ואז לנרמל, אך יותר קל חישובית) נגדיר לכל $i \in [n]$:

$$\bar{v}_i = \frac{1}{||\tilde{v}_i||}\underbrace{\left(v_i - \sum_{j=1}^{i-1} \langle v_i \mid \bar{v}_j \rangle \cdot \bar{v}_j\right)}_{\tilde{v}_i}$$

בצורה זו נוכל לנרמל תוך כדי התהליך.

ה' **אלגוריתם אוקלידס לתחום אוקלידי** (בפרט בעבור פולינומים ומספרים שלמים): ניעזר בזהות $\gcd(a,b) = \gcd(a, b+qa)$. כדי למצוא את $\gcd(a,b)$ נחזור על הפעולה הבאה: בה"כ $a > b$, בתחום אוקלידי מוגדרת השארית r כך ש־$a = bq + r$. לכן $\gcd(a,b) = \gcd(bq+r, b) = \gcd(b,r)$, ומהגדרה $N(r) < N(b)$ כאשר N הנורמה האוקלידית. לכן נוכל להמשיך בתהליך עד שנגיע לזוג $\gcd(a',b')$ כך שאחד מהם (בה"כ b') מקיים $b'=0$ ואז $\gcd(a',b') = a'$. בחוג המספרים השלמים לאלגו' סיבוכיות $\mathcal{O}(\log_\varphi(\max\{a,b\}))$.

ו' **נרמול וקטור:** נגדיר $v = \frac{v}{||v||}$ הוא v מנורמל.

ז' **בדיקת T־איווריאנטיות:** בהינתן B בסיס של $W \subseteq V$ נחשב את $T(B)$ ונבדוק האם $T(B) \subseteq W$ ע"י מעבר על כל איבר בסיס ודירוג.

ח' **חישוב A^{-1}, A^{n+c} באמצעות משפט קיילי־המילטון:** ידוע $f_A(A) = 0$, ואם נשאר גורם חופשי $\alpha_n A^n + \alpha_0 A^0 = 0$ אז נוכל להעביר אגפים ולקבל: $I = A^0 = A\left(\frac{\sum_{k=1}^n \alpha_k A^{k-1}}{\alpha_0}\right)$ ולכן $A^{-1} = \frac{1}{\alpha_0}\left(\sum_{k=1}^n \alpha_k A^{k-1}\right)$. כדי לחשב את A^{n+c} תחילה נחשב את A^n באמצעות העברת אגפים וקבלת $A^n = -\sum_{k=0}^{n-1} \alpha_k A^k$ (כי $\alpha_n = 1$ כי הפולינום מתוקן). עתה, נכפול ב־A בדיוק c פעמים, ומשום שידוע A^n, בכל חלוקה שבא נקבל A^{n+1} נוכל להוציא גורם משותף ולקבל $A(A^n) = \sum_{k=1}^n \beta_k A^k$ ביטוי שהכפל הגבוהה ביותר בו תמיד A^{n-1}. סה"כ נוכל לבטא את A^{n+c} כקומבינציה לינארית של $I \cdots A^{n-1}$, שעבור מספרי n קטנים קל לחשב.

ט' **ייצוג בבסיס אורתוגונלי:** לכל $u \in V$ בהינתן $(v_1 \dots v_n)$ בסיס אורתוגונלי, מתקיים $u = \sum_{i=1}^n \frac{\langle u \mid v_i \rangle}{||v_i||^2} \cdot v_i$ (אין צורך לחלק בנורמה בעבור אורתונורמלי).

י' **מציאת היטל אורתוגונלי:** בהינתן $(u_1 \dots u_n)$ בסיס אורתוגונלי של U תמ"ו, אז $p_U(v) = \sum_{i=1}^k \frac{\langle v \mid u_i \rangle}{||u_i||^2} u_i$ (גם כאן אין צורך לחלק בנורמה בעבור בסיס אורתונורמלי).

לחילופין, אפשר להיעזר בעובדה שבהינתן v המוטל על $U = (u_1 \dots u_n)$, נאמר ו־u היא התוצאה, אז ניתן לבטאו כ־$v = u + u^\perp$ כאשר $u \in U, u^\perp \in U^\perp$. לשם כך, נסמן ב־$u$ את התוצאה, ואז ידוע $v - u = u^\perp$ כלומר $v - u \perp u_i$ לכל $i \in [n]$. קיבלנו מערכת לינארית משוואות ב־n נעלמים שאפשר לדרג ולפתור.

יא' **מציאת ללכסון אוניטרי/אורתוגונלי** (אם קיים ממשפט הפירוק הספקטרלי):

- נמצא את הע"ע של ההעתקה.
- לכל ע"ע, נמצא בסיס עצמי של ו"ע ואז נבצע עליו גראם־שמידט כדי לקבל וקטורים אורתוגונליים/אורתונורמליים.
- נשרשר את הבסיסים לקבלת בסיס אורתוגונלי/אורתונורמלי מלכסן.

בניסוח אחר: נלכסן את ההעתקה T, אבל נעשה גראם־שמידט על כל ו"ע. כדי להוכיח שאלגו' זה אכן עובד, יש להוכיח את הטענה הנפוצה לפיה כל שני מרחבים עצמיים אורתוגונליים בהינתן העתקה נורמלית.

יב' **מציאת פירוק SVD:** בהינתן T העתקה, נמצא את הפירוק הספקטרלי של T^*T, TT^* ומהזהויות $TT^* = U\Sigma^2 U^*$ ו־$T^*T = V\Sigma^2 V^*$, נקבל $T = U\Sigma V^*$.

המשך בעמוד הבא

3.4 Recommended Exercises

תת־הנושא הבא כולל תרגילים נפוצים במיוחד, או תרגילים קשים ומעניינים שאספתי ממבחני עבר. אני ממליץ בחום לעבור על כולם לקראת המבחן.

תרגיל 1 (נפוץ). תהי T מעל ממ"פ מרוכב, ו־$f_T = g \cdot h$ כאשר $g, h \in \mathbb{F}[x]$ פולינומים זרים,

א' הוכח שאם T לכסינה, מתקיים $\ker(gT) = \ker(hT)^\perp$.

ב' הוכח שלכל T, מתקיים $\ker gT \oplus \ker hT = V$ (בסעיף זה אין ערך להיות ההעתקה מעל ממ"פ).

הדרכה: נסו להתחיל מהמקרה של שני ערכים עצמיים מורחבים, ואז להכליל באמצעות פירוק למרחבים עצמיים.

תרגיל 2 (נפוץ). הוכח שלכל T נורמאלית, עם ע"עים $\lambda_1 \dots \lambda_k$, מתקיים לכל $n, m \in [k]$ ש־$\forall v \in V_{\lambda_n}, \tilde{u} \in V_{\lambda_m} : \langle v \mid u \rangle = 0$ (דהיינו V_{λ_n} ניצב ל־V_{λ_m}).

תרגיל 3 (נפוץ). הוכח שלכל $T : V \to V$ מתקיים $\ker(T^*) = \operatorname{Im}(T)^\perp$ וגם $\operatorname{Im}(T^*) = \ker(T)^\perp$

תרגיל 4 (נפוץ). הוכח שלכל $T : V \to V$ נורמלית, מתקיים:

א' $$\ker T = \ker T^*$$

ב' $$\operatorname{Im} T = \operatorname{Im} T^*$$

ג' $$V = \ker T \oplus \operatorname{Im} T$$

תרגיל 5 (נפוץ). ללא שימוש בפירוק SVD, הראה שהערך הסינגולרי הגדול ביותר והקטן ביותר חוסמים את הנורמה $||Tv||$ לכל $||v|| = 1$.

סוף הקורס ~ 2025B

מאת שחר פרץ

קומפל ב־LaTeX ונוצר באמצעות תוכנה חופשית בלבד