אלגברה לינארית 2א

שחר פרץ $\sim$ 2025B

Introduction 1

סיכום זה לאלגברה לינארית 2א, נעשה במסגרת תוכנית אודיסאה, עם בן בסקין כמרצה. עקב מבצע "עם כלביא" בוטלו חלק מההרצאות, שהושלמו באמצעות סיכומים חיצוניים, והקלטות של הקורס (בפרט, הפרקים על המשפט הספקטרלי מסתמכים על הרצאות של ענת). כל המרצים הסתמכו על סיכום אחר של סטונדטים אחרים, אך הם גם הוסיפו הערות בע"פ לגבי אינטואציה, הוכחות נוספות, והרחבות קלות של החומר (לדוגמה בהקשר של תורת החוגים, הוכחות המשפט הספקטרלי, או ציקליות) שבסיכום זה אני שואף להעביר. פרט לכך הוספתי ציטוטים מן ההרצאה שמצאתי משעשעים.

כנגד שלושה נושאים דיברה אלגברה לינארית 2א –

1. **אופרטורים ליניארים** שיובילו אותנו לצורת ג'ורדן.
2. **תבניות בי־ליניאריות**, אובייקט 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.

נוסף על שלושת הנושאים ה"רגילים" של הקורס, מופיעה בסוף הרחבה של בן בסקין לגבי מרחבים דואלים. אני ממליץ בחום גם למי שלמד את הנושא בלינארית 1א לקרוא את הפרק עם מרחבים דואלים, משום שהוא קצר, ומראה קשרים חזקים (ומרתקים!) בין החומר הנלמד באלגברה לינארית 2א (כמו מרחבי מכפלה פנימית והעתקות צמודות) למרחבים דואלים.

באופן אישי, אני מוצא את הקורס די מעניין, ובמיוחד את הפרק האחרון בנוגע לפירוקים של מטריצות/העתקות.

אם מצאתם בסיכום טעויות (החל בתקלדות, כלה בשגיהוט חטיב, ובטח טעויות מתמטיות) אשמח אם תפנו אלי בטלפון או במייל (perets.shahar@gmail.com). הגרסה האחרונה של הסיכום תמיד זמינה בקישור הבא.

מקווה שתהנו מהסיכום ותמצאו אותו מועיל;

שחר פרץ, 19.7.2025

אזהרה! נכון למצב הנוכחי, הסיכום הזה עדיין בעבודה לצריך לעבור הגהה. אתם מוזמנים להשתמש בו, אך אין לראות בו כרגע את הגרס הסופית.

כמה הערות טכניות

- 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 (במיוחד בפיזיקה), שאני חושב שנראה מגניב הרבה יותר.
- אני לא אחראי בשום צורה על דברים לא נכונים שכתבתי בטעות. פעמים רבות מרצים מדלגים עם שלבים ואני משלים מההבנה שלי, ולמרות שעברתי על הסיכום ואני משתדל שיהיה מדויק ככל האפשר, ייתכן שישנן טעויות.
- היו משפטים שהוכחנו בתרגילי הבית, או שהוכחתם טרוויאלית. פתרונות לתרגילי הבית זמינים ב־github שלי, למי שמעוניין.

תוכן העניינים

בתיאבון

2 Diagonalization

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.

2.1 ~ מבוא לפרק

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

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

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 פיבונ'צי. נתבונן בהעתקה הבאה: (בהנחת איברי בסיס $a_0 = 0, a_1 = 1$).

מה חבל שלא כיף להכפיל את המטריצה הזו בעצמה המון פעמים. מה נוכל לעשות? ננסה למצוא בסיס שבו $(v_1, v_2) = \left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)_B$ 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 צורה ג'ורדנית – 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/טל ממרחב וקטורי V לעצמו.

2.2 ~ ערכים עצמיים ווקטורים עצמיים לאופרטורים ליניארים

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

"ראיתם את המרחב הומו?". כדאי לדעת כי $\operatorname{Hom}(\mathbb{F}^n, \mathbb{F}^n) \approxeq M_{m \times n}(\mathbb{F})$. מה המשמעות של איזומורפי ($\approxeq$)? בהינתן A, B מבנים אלגברים כלשהם, נסמן $A \approxeq B$ 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 $T^nv=v$, 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 $T^nu=u$. ידוע קיום $\alpha_0,\dots,\alpha_{n-1}\in\mathbb{F}$ כך ש־$u=\sum\alpha_iT^i(v)$. אז:

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

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: $id, 0$.

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 צ"ל כלשהו שווה ל־0:

$$\begin{aligned}\sum_{v_i\in B}\alpha_iv_i=0\\ &=\sum_\lambda\sum_{\lambda_i}\alpha_iv_{\lambda,i}\\ &\implies\sum_{\lambda_j}\alpha_iv_{\lambda_{ji}}=:u_j\in V_{\lambda_j}\\ &\implies\sum_ju_j=0\end{aligned}$$

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:

$$\sum_\lambda\dim V_\lambda\le n$$

שוויון אמ"מ T לכסין.

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

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. ∎

2.3 $\sim$ 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 $v\neq0$ 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". $\sim$ בן

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

2.4 $\sim$ פולינום אופייני

תרגיל. תהי $A=\binom{-7\ 8}{6\ 7}$. מצאו ו"ע וע"ע של A ולכסנו אם אפשר.

פתרון. מחפשים $\binom{0}{0}\neq\binom{x}{y}\in\mathbb{R}^2$ 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 6. 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 7. 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 8. $f_A(x)$ 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 ראינו 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 $\sum_{i=1}^n -a_{ii} = -\operatorname{tr} A$ המקדם החופשי.

- **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 9. 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 10. $T\colon V \to V$ 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 הגזירה, ממקודם: $f_T(x) = x^{n+1}$ 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.

2.5 ~ על הקשר בין ריבוי גיאומטרי ואלגברי

הערה 3. בדוגמה שבטענה ראינו שמתקיים $\sum d_i = \sum n_i = n$ 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 11. תהי $T\colon V \to V$ ט"ל. אזי לכל ע"ע λ מתקיים $r_\lambda \le d_\lambda$.

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:

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

■

משפט 12. 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, $f_T(x) = \prod_{i=1}^k (x-\lambda_i)^{n_i}$
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־$n = \sum r_{\lambda_i} \le \sum d_{\lambda_i} = n$ ולכן אם לאחד מבין הערכים העצמיים מתקיים $r_\lambda \neq d_k$ אז מתקיים $r_k < d_k$ ונקבל סתירה לשוויונות לעיל.

$\Longrightarrow$

$$1 \implies \sum d_{\lambda_i} = n$$
$$2 \implies \sum r_{\lambda_i} = \sum d_{\lambda_i} = n$$

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

2.5.1 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 $m \le p^2$. 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 $A^{p-1} = I$. ■

2.6 ~ 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 13. $T\colon V \to V$ 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 $w \in B'' \cup \{w_1\}$ 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 14. מטריצה A ניתנת לשילוש, אמ"מ הפ"א האופייני שלה מתפצל לגורמים לינארים.

המשך בעמוד הבא

3 Cayley–Hamilton Theorem

3.1 ∼ 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.

לכן, נגדיר את $\mathbb{F}(x)$ – אוסף הפונקציות הרציונליות:

$$\mathbb{F}(x)=\left\{\frac{f(x)}{g(x)}\mid g(x)\neq 0\right\}$$

זהו שדה. 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 $f\neq g$.

3.2 ∼ מבוא למשפט קיילי־המילטון

הגדרה 17. יהי $f(x)=\sum_{i=0}^{d}a_ix^i\in\mathbb{F}[x]$, V מ"ו מעל $\mathbb{F}$ נ"ס (נוצר סופית) וכן $T\colon V\to V$ ט"ל. נגדיר:

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

כנ"ל עם מטריצות (ראה תרגול)

טענה. אם $A=[T]_B$ 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 $f(C)=0\iff f(A)=0$.

3.3 ∼ משפט קיילי־המילטון

משפט קיילי־המילטון. לכל $T\colon V\to V$ ט"ל (V נוצר סופית) ולכל $A\in M_n(\mathbb{F})$ מתקיים:

$$f_T(T)=0,\ f_A(A)=0$$

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 15. **(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 $f_T(T)=0,\ f_A(A)=0$.

משפט 16. אם A מייצגת של העתקה T. ו־$f\in\mathbb{F}[x]$, אז $f(T)=0\iff f(A)=0$.

(באופן כללי, עבור כל פולינום).

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־$\dim W \le \dim V$, 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:

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

■

- 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 $\mathbb{F} \subseteq \mathbb{K}$ 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. ■

■

הערה על שדות סגורים אלגברית. (לא נאמר בקורס) העובדה שלכל שדה יש שדה שסגור אלגברית – טענה שתלויה באקסיומת הבחירה. הסגור האלגברי הוא יחיד.

המשך בעמוד הבא

4 Ring Theory

4.1 ~ מבוא והגדרות בסיסיות

מה זה אובייקט אלגברי? דוגמאות: תמורות, חבורות, שדות, מרחבים וקטורים, ועוד. הרעיון – "Data עם אקסיומות".

הגדרה 18. חוג עם יחידה הוא קבוצה עם שתי פעולות, כפל וחיבור, ניטרלים לפעולות (0, 1) כך שמתקיימות כל אקסיומות השדה למעט (אולי) קיום איבר הופכי, וקומטטיביות הכפל.

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

הגדרה 19. תחום שלמות הוא חוג קומוטטיבי עם יחידה ללא מחלקי 0.

הגדרה 20. חוק ייקרא ללא מחלקי 0 אם: $\forall a, b \in \mathbb{R}\colon ab = 0 \implies a = 0 \vee b = 0$

דוגמאות לחוגים עם מחלקי 0:

- $M_2(\mathbb{R})$: הוכחה $a = b = \binom{0\,1}{0\,0},\ a \cdot b = 0$
- $\mathbb{Z}/_{6\mathbb{Z}}$ הוכחה $2 \cdot 3 = 0$.

משפט 17. בתחום שלמות יש את כלל הצמצום בכפל: אם $ab = ac \wedge a \neq 0$ אז $b = c$.

הוכחה.

$$ab \cdot ac = 0 \implies a(b \cdot c) = 0 \implies a = 0 \vee b - c = 0$$

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

4.2 ~ ראשוניות ואי־פריקות

הגדרה 21. יהי R תחום שלמות, $a, b \in R$. נאמר ש־$a \mid b$ אם קיים $c \in \mathbb{R}$ כך ש־$ac = b$.

הגדרה 22. $u \in R$ נקרא הפיך אם קיים $\alpha \in R$ כך ש־$\alpha u = 1$.

משפט 18. יהי R תחום שלמות, $u \in R$ הפיך. יהי $a \in \mathbb{R}$. אז $u \mid a$.

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

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

משפט 19. 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 $a \sim b \wedge b \sim c$, 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 20. הופכי הוא יחיד

(אותה ההוכחה כמו בשדות. לא בהכרח בתחום שלמות, מעל כל חוג)

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

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

■

משפט 21. אם $a \mid b$ וכם $b \mid a$ אז $a \mid b$ (בתחום שלמות).

הוכחה.

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

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 $a \mid b$.

"אני חושב שבעברית קראו להם ידידים, לא רצו להתחייב לחברות ממש".

הגדרה 24. איבר $p \in R$ נקרא אי־פריק אם מתקיים $p = ab \implies a \in R^x \vee b \in R^x$.

הגדרה 25. איבר $p \in R$ יקרא ראשוני אם $p \mid (a \cdot b) \implies p \mid a \vee p \mid b$.

הערה 6. איברים הפיכים לא נחשבים אי־פריקים או ראשוניים. הסיבה להגדרה: בשביל נכונות המשפט היסודי של האריתמטיקה (יחידות הפירוק לראשוניים).

משפט 22. בתחום שלמות כל ראשוני הוא אי פריק.

הערה: שקילות לאו דווקא.

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

הגדרה 26. R תחום פריקות יחידה אם $\prod_{i=1}^{n} p_i = \prod_{j=1}^{m} q_i$ עבור p_i, q_j ראשוניים, אז $m = n$, ועד לכדי סידור מחדש, לכל $i \in [n]$ $p_i \sim q_i$

משפט 23. 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 27. יהי R תחום שלמות. תת־קבוצה $0 \neq I \subseteq R$ נקראת אידיאל אם:

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

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

דוגמאות:

1. 0 תמיד אידיאל, כך החוג תמדי אידיאל.
2. הזוגיים ב־$\mathbb{Z}$.
3. לכל $n \in \mathbb{Z}$, $n\mathbb{Z}$ אידיאל (n כפול השלמים). הזוגיים מקרה פרטי.
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 $\langle a \rangle := \{a \cdot b \mid b \in R\}$
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 28. אידיאל I נקרא ראשי אם הוא מהצורה aR עבור $a \in R$ כלשהו.

הגדרה 29. תחום שלמות נקרא ראשי אם כל אידיאל שלו ראשי.

הערה 7. אנחנו סימנו אידיאל ב־aR ובקורס מסמנים Ra, באופן כללי אפשר לדבר על אידיאל שמאלי ואידיאל ימני. תזכורת: $I \subseteq R$ היא אידיאל אם היא סגורה לחיבור ומקיימת את תכונת הבליעה. בתחום ראשי כל אידיאל הוא אידיאל ראשי.

משפט 24. ב־$R \neq \{0\}$ תחום ראשי אז כל אי פריק הוא ראשוני.

(תנאי מספיק אך לא הכרחי)

הוכחה. יהי o אי פריק (א"פ). יהיו $a, b \in \mathbb{R}$ כך ש־$p \mid ab$. תיקון [משבוע שעבר]: במקום $I = Ra + Rb$, נשתמש ב־$I = Ra + Rp$. בכלל ש־R תחום ראשי, קיים $c \in R$ כך ש־$I = Rc$, ו־$a, p \in I$ כלומר $p \mid c \wedge a \mid c$. p א"פ ולכן $p \sim c$ או c הפיך.

- c הפיך $\impliedby \in I \impliedby R = R \cdot 1 \in I \subseteq R \impliedby I = R$. 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 $c \mid a \wedge p \mid c$ ולכן $p \mid a$.

■

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

משפט 25. 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 30. יהי $g \in R$ כך ש־:

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

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

משפט 26. 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.

(הערה: רק 3 באמת דורש תחום ראשי)

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־$ra + sb = g$ ולכן $a, b \in I$ אז $g \mid a, b$ וסיימנו מ־1.

■

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

משפט 27. $\mathbb{F}[x]$ 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־$f \in I \wedge p \in I$ 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 31. 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 $\forall 0 \neq a, b \in R \colon N(a) \leq N(ab)$.

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 28. חוג אוקלידי $\Longleftarrow$ פריקות יחידה (דומה למשפט היסודי של האריתמטיקה).

משפט 29. 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:

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

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:

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

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

משפט 30. 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 32. $I \subseteq R$ אידיאל נקרא ראשוני אם $\forall a, b \in R\colon (a \cdot b) \subseteq P$, אז $(a) \in P \vee (b) \in P$.

הגדרה 33. 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=ראשוני. ניתן להראות באופן שקול כי:

משפט 31. R תחום ראשי, אז I ראשוני אמ"מ I אי־פריק.

הגדרה 34. 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:

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

צריך להוכיח שזה לא תלוי בנציגים (הנציגים a, b) והכל אבל בן לא עומד לעשות את זה. נשאר כתרגיל בעבור הקורא.

4.3 $\sim$ הרחבת שדות

משפט 32. 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־$bd - ac + (ad + bc)x + I$ זהו 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 אממ - הכיוון השני תרגיל בעבור הקורא).

הגדרה 35. 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 33. 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 קונסטרקטיבית, 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 $\mathbb{F} \mapsto \mathbb{K}$.

משפט 34. (ללא הוכחה בקורס) לכל שדה $\mathbb{F}$ קיים ויחיד שדה $\mathbb{F} \subseteq \mathbb{K}$ סגור אלגברית.

דוגמה: $\mathbb{R}$ ו־$\mathbb{C}$.

4.4 ~ חוג הפולינומים

(תת־פרק זה לקוח מתרגול בקורס)

הגדרה 36. הדרגה של הפולינום היא $\deg(f) := \max\{n \in \mathbb{N} \mid a_n \neq 0\}$, ומגדירים $\deg(0) = -\infty$

משפט 35.

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

הערה 8. חוג הפולינומים הוא חוג אוקלידי כי $N = \deg f$ 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 37. נאמר שפולינום q מחלק את f אם $r = 0$ ומסמנים $q \mid f$.

מסקנה 9.

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

הוכחה.

1. הוכחה למשפט בזו:

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

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

■

"לא הנחתי בשלילה, הוכחתי בקונטראפוסיטיב"

משפט 36. 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 37.

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

משפט 38. (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 39. (מסקנה מהטענה הקודמת שניתן להוכיח באינדוקציה ללא הרחבת שדות) לפולינום $f \in \mathbb{F}[x]$ שאינו אפס יש לכל היותר $\deg f$ שורשים.

הערה 9. 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 10. 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 40. 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. $\forall A \in M_n(\mathbb{F})\ \forall k \geq n\ \exists p(c) \in \mathbb{F}_{n-1}[x]\colon A^k = p(A)$.

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

4.4.1 פונקציות רציונליות ומספרים אלגבריים

אינטואציה: הרעיון של פונקציה רציונלית היא להיות "פולינום חלקי פולינום". נפרמל את הדבר הזה בעבור כל שדה מעל נוכל להגדיר את מרחב הפולינומים.

משפט 41. 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 4. נסמן כל איבר במחלקת השקילות ע"י $\frac{f}{g}$ שמייצגים אותו.

הגדרה 38. 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 תללויות בנציגים)

משפט 42. $Q[x]$ 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.

משפט 43. לכל p ראשוני $\forall x \in \mathbb{F}_p\colon x^p = x$.

הערה: זוהי מסקנה ישירה מהמשפט הקטן של פרמה.

הגדרה 39. 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 40. מספר מרוכב שאינו אגלברי יקרא טרנסצנדנטי.

משפט 44. 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

5 Primary Decomposition

5.1 ~ מרחבים T־שמורים וציקליים

הגדרה 41. 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־אינוו'.

הערה 11. שימו לב: אם $U \subseteq V$ תמ"ו אינווריאנטי, אז $T|_U \colon U \to U$ ט"ל.

הערה 12. 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־$B \in M_k$).

הגדרה 42. יהי V מ"ו מעל $\mathbb{F}$, V $T\colon V \to V$ 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 45.

- $\mathcal{Z}(T, v)$ תמ"ו של V – טרוויאלי.
- $\mathcal{Z}(T, v)$ תמ"ו T־אינ' – 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 $T^k v \in \mathcal{Z}(T, v)$. לכן קיימים $a_0 \dots a_{k-1}$ כך ש־$T^k v + a_{k-1} T^{k-1} v + \dots + a_0 v = 0$. 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:

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

הגדרה 43. $A_f = [T]_B$ היא המטריצה המצורפת לפולינום $f(x) = x^n + a_{n-1}x^{n-1} + \dots + a_0$.

5.2 ~ הפולינום המינימלי

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

משפט 46. 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 44. I_A לעיל יקרא הפולינום המינימלי.

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 5. m_A יהיה הפולינום המינימלי של המטריצה A.

הערה 13. 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 14. 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 שלמות, $m_A \mid f_A$ 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 - לדוגומה 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 לגזור n פעמים ע"מ לקבל 0, לדוגמה x^n.

משפט 47. תהא $A=A_f$ המטריצה המצורפת ל־A. אז $f_A=m_A$.

משפט 48. 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 15. נניח ש־A לכסינה והע"ע השונים הם $\lambda_1\ldots\lambda_k$ (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, $A=\operatorname{diag}(\lambda_1\ldots\lambda_k)$ 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 $B=(v_1\ldots v_n)$. 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 $\prod_{i=1}^k(x-\lambda_i)\mid m_a$. 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 16. 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 49. 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־$f\mid m_T$ 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 50. 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:

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

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

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

משפט 51. 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 $(x-\lambda)\mid m_T(x)$. ידוע $m_T\mid f_T$ וסה"כ $(x-\lambda)\mid f_T$ וסה"כ λ ע"ע של T. ∎

משפט 52.

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

הוכחה. נותר להוכיח $f_A(x) \mid (m_A(x))^n$ (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 $f,g \in \mathbb{F}[x]$, $\mathbb{F} \subseteq \mathbb{K}$ ומתקיים $g \mid f$ מעל $\mathbb{K}$, אז $g \mid f$ מעל $\mathbb{F}$. אז:

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

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

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

מסקנה (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 53. 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־$m_A(x)$. באופן כללי, $\operatorname{lcm}(a_1 \dots a_n)$ מתקבל כיוצר של אידיאל החיתוך בתחום הראשי. כלומר:

$$I = (\ell) = \bigcap_{i=1}^{n} Ra_i$$

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

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

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

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

משפט 54. נניח ש־$T, S \colon V \to V$ ט"ל מתחלפות. אז:

1. $\operatorname{Im} S,\ \ker S$ הם T־אינווריאנטים (ולהפך).
2. אם $S \subseteq W$ תמ"ו הוא T־אינוו' אז גם $S(W)$ הוא T־אינ'.
3. אם $W_1, W_2 \subseteq V$ הם T־אינ' אז גם $W_1 + W_2,\ W_1 \cap W_2$ הם T־אינ'.
4. אם $f \in \mathbb{F}[x]$ ו־$W \subseteq V$ תמ"ו T־אינ', אז W גם $f(T)$־אינ'.

הוכחה.

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

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

ועבור $v \in \ker S$:

$$S(T(v)) = \cdots = T(S(v)) = T(0) = 0 \implies T(v) \in \ker S$$

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

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

כי $T(w) \in W$.

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 וצ"ל וסיימנו. [בסיכום כתוב הוכחה: קל]

$\blacksquare$

(הערה: 3, 4 לא תלויים בהיות הטרנספורמציות מתחלפות)

5.3 $\sim$ ניסוח והוכחה של משפט הפירוק הפרימרי

משפט 55. ("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־:

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

ולכן $V=\ker h(T)+\ker g(T)$. וכן הראינו שזהו סכום ישר. עתה, נסמן:

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

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־אינוו' בשיעור קודם:

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

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

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

ולכן כל ה"אשים לעיל הדוקים ושוויון בכל מקום.

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

אם אחד מהשווינות לא הדוקים, אז:

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

וסתירה למה שהראינו. לכן:

$$m_{T_1}\mid g\wedge\deg m_{T_1}=\deg g\implies m_{T_1}\sim g$$

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

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 $T_{|^3_{T-I}}$.

משפט 56 (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 $T_{|_{\ker g_i(T)}}$. זה פשוט אינדוקציה על המשפט הקודם.

המרצה גם מוכיח את זה על הלוח אבל לא מתחשק לי לכתוב את זה. טוב, אני אכתוב את זה. "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:

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

מהמשפט שקיבלנו:

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

והמשך דומה עבור $m_{T|_{\ker g_i}}=g_i$:

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

■

הערה 17. (מקרה פרטי חשוב) נניח כי m_T מפרק לגורמים לינאריים שונים זה מזה. כלומר:

$$m_T=\prod_{i=1}^{s}(x-\lambda_i)$$

ו־$\lambda_i\neq\lambda_j$ לכל $i\neq j$, אז T לכסינה.

הוכחה. לפי המשפט:

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

V סכום ישר של מ"ע $\lambda_1\dots\lambda_2$ הם כולם ע"ע שונים, אז יש ל־V בסיס של ו"א ולכן T לכסינה. ■

סיכום. $T\colon V\to V$ ט"ל, ו־:

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

אז:

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

המשך בעמוד הבא

6 Jordan Form

6.1 ~ 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 התורה של הרחבת שדות לשם כך.

היוונים העתיקים ביססו את הגיאומטריה שלהם באמצעות דברים שאפשר לבדוק עם שדה ומחוגה. באמצעות כלים של 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, שבאמצעות סרגל ומחוגה אי אפשר לעשות זאת. בעיות שהיו פתוחות לעולם המתמטי במשך אלפי שנים נפתרו בעזרת אותן התורות.

אבל, בהתאם לשמו, מת משחפת בגיל 26. גלואה מת בגיל 21 מדו־קרב.

מסקנה 11. לא ללכת לדו־קרב.

הוכחה. ההוכחה מתקדמת ועוסקת בתורת גלואה. ■

הגדרה 45. בהינתן $\forall i \neq j \colon \lambda_i \neq \lambda_j$ $\quad f(x) = \prod_k (x-\lambda_k)^{r_k}$. אז $f^{\text{red}} := \prod_k (x-\lambda_k)$

טענה:

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

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

6.2 ~ צורת ג'ורדן לאופרטור לינארי נילפוטנטי

6.2.1 f^{red} ונילפוטנטיות

משפט 57. 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 $1 \leq r_i \leq s_i$ 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 מהלמה. A לכסינה אמ"מ $m_A = f_A^{\text{red}}$, ואנחנו יודעים כי $m_A(A) = 0$ ולכן A לכסינה אמ"מ $f_A^{\text{red}}(A) = 0$. ■

משפט 58. 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 $m_S \mid m_T = \prod_{i=1}^r (x-\lambda_i)$ ולכן m_S מתפרק לגורמים לינארים זרים, סה"כ S לכסינה. ■

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

הגדרה 46. 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, נניח ש־$f_T(x)$ מתפצל מעל f לגורמים לינארים (כלומר, נרחיב לשדה סגור אלגברית).

הערה 18. (נובע ממשפט הפירוק הפרימרי) אם $S\geq 2$, ידוע־$V=\bigoplus_{i=1}^{s}\ker g_i(T)$ ולכן V פריק ביחס ל־T. בהינתן ההנחה כי f_T מתפצל לחלוטין, ונניח V אי־פריק ביחס ל־T, אז $m_T(x)=(x-\lambda)^r$.

הגדרה 47. $T\colon V\to V$ ט"ל. T נקראת נילפוטנטית אם קיים $n\in\mathbb{N}$ כך ש־$T^n=0$. באופן דומה A מטריצה נילפוטנטית אם $\exists n\in\mathbb{N}\colon A^n=0$.

הגדרה 48. עבור n המינימלי שעבורו $T^n=0/A^n=0$, אז n הנ"ל נקרא דרגת הנילפוטנטיות של T/A, ומסמנים $n(T)/n(A)$.

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

דוגמה. בסיטואציה ש־$m_T(x)=(x-\lambda)^r$ נסיק ש־:

$$\left((T-\lambda I)^r=0\wedge S:=T-\lambda I\right)\implies n(S)=r$$

סימון 6. כל פירוק של V ל־T נותן פירוק שלו ל־$T-\lambda I$ ולהיפך.

הוכחה. ההערה נכונה כי אם $V=U\oplus W$ כאשר U,W הם T־שמורים, לא טרוויאלים, אז הם גם $T-\lambda I$־שמורים. זאת כי אם U הוא T שמור אז:

$$\forall u\in U\colon T(u)\in U\implies (T-\lambda)(u)=\underset{\in U}{T(u)}-\lambda\underset{\in U}{u}\in U$$

■

המשך ההערה. כדי להבין איך נראים תת־מרחבים אי־פריקים, עשינו רדוקציה לט"ל ניל' [רדוקציה=מספיק לי להבין את המקרה הזה בשביל להבין את המקרה הכללי].

6.2.2 ציקליות

משפט 59. $T\colon V\to V$ ניל' אז לכל $0\neq v\in V$ הקבוצה $v,Tv,\ldots,T^kv\neq 0$ היא בת"ל.

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

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

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

הגדרה 49. קבוצה מהצורה $\{v,Tv\cdots T^kv\}$ כאשר $T^{k+1}v=0$ והוא המינימלי, נקרתא שרשרת.

הגדרה 50. תמ"ו שקיים לו בסיס שהוא שרשרת, נקרא ציקלי.

(ראה לינארית 2א סיכום 8)

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

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

ו־T אופרטור הגזירה הפורמלית. כדי ש־V יהיה ציקלי, צריך למצוא בסיס ציקלי שממדו הוא 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 19. יהי $T\colon V\to V$ ניל' ו־$\dim V=n$ אז $n(T)\leq n$ וישנו שוויון אמ"מ V ציקלי.

הערה 20. אם $T\colon V\to V$ ניל' 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 $k\geq\ell$. נסמן $B_v=(v,Tv\ldots T^{n-1}v)$. קיימים (ויחידים) $u\in U,w\in W$ כך ש־$v=u+w$. אז:

$$T^kv=T^ku+T^kw$$

אך משום ש־T ניל' אז $T_{|_U},T_{|_W}$ ניל' גם כן. ידוע $n(T_{|_U}),n(T_{|_W})\leq k$ ולכן בפרט $T^k(u)=T^k(w)=0$ ולכן $T^kv=0$ אבל $k<n\wedge T^k(v)\in B_v$. ■

משפט 60. תהי $T\colon V\to V$ ניל' ונניח U תמ"ו של V הוא T-אינו' וציקלי, אז עבור $S=:T_{_U}$:

1. $\dim U\le n(T)$
2. $\operatorname{Im}(T_{_U})=T(U)$ ציקלי ו-$\dim T(U)=\dim U-1$

הוכחה.

1. $n(T)\ge n(T_{|_U})$ וגם $\dim U=n(T_{_W})$
2. $T(u)=T(\operatorname{span}(v,\dots T^kv))=\operatorname{span}(Tv\dots T(T^kv))$ ז"א $T(U)=\operatorname{span}(Tv\dots T^kv)$, זו קבוצה בת"ל ופורתש 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 61. לכל V מ"ו, $T\colon V\to V$ ניל' 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 62. נניח $U\subseteq V$ תמ"ו ציקלי מקסימלי. אזי:

1. אם $T(U)\subseteq T(V)$ הוא גם ציקלי מקסימלי. (הערה: הורדת הממד באחד מועילה מאוד באינדוקציה)
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:

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

זו סתירה כי:

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

■

6.2.3 ניסוח צורת מייקל ג'ורדן לאופרטור נילפוטנטי

משפט 63 (המשלים הישר לתמ"ו ציקלי מקסימלי)**.** נניח $T\colon V\to V$ ט"ל לינ' ניל' (ניל"י), $U\subseteq V$ תמ"ו ציקלי מקסימלי אז קיים $W\subseteq V$ תמ"ו T-אינ' כך ש-$V=U\oplus W$.

הוכחה. נוכיח באינדוקציה על $n=n(T)$.

בסיס: אם $n(T)=1$ אז $T=0$ "יש מה להוכיח בכלל" אז כל $W\subseteq V$ הוא T-אינ'. 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-אינ' כך ש-$T(V)=T(U)\oplus W_1$.

נגדיר $W_2=\{v\in V\mid Tv\in W_1\}$. אז

למה 4. ("למה א")

- $U+W_2=V$ (לאו דווקא סכום ישר)
- $U\cap W_1=\{0\}$

למה 5. ("למה ב") בהינתן $W_1\subseteq W_2$ ו-$U\subseteq V$ 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 $T(w) = W_1 \subseteq W'$.

ולכן מש"ל משפט. ■

ציור של למה 2: אני לא יודע לעבוד עם tikz מספיק טוב, ואני בטוח ש־chatGPT יוכל לעשות tikz עבורי, אבל אני גם רוצה להיות מרוכז בהרצאה. אז בבקשה פשוט תעשו דיאגרמת ואן ללמה ב'. גם המרצה לא הוכיח, זה משחקים על הרחבות בסיס וממדים בצורה כזו שאתם מכירים מלינארית 1א. הוכחת הלמה נשארה כתרגיל בעבור הקורא.

נוכיח את למה א'.

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 12. $T \colon V \to V$ ט"ל ניל' אז V אי־פריק ל־T אמ"מ V ציקלי.

הוכחה.

- $\implies$ הוכחנו בשיעור הקודם
- $\impliedby$ נניח V אי־פריק. אז קיים $U \subseteq V$ תמ"ו ציקלי מקסימלי. לפי המשפט קיים $W \subseteq V$ תמ"ו T־אינ' כך ש־$T = U \oplus W$. ידוע U, W תמ"ןים אינ'. אם $U = \{0\}$ אז $V = 0$ ובפרט ציקלי. אחרת, מאי־פריקות V ל־T, נסיק ש־$W = \{0\}$ ולכן $V = U$ ציקלי.

■

משפט 64 (משפט ג'ורדן למקרה של T ניל')**.** תהי $T \colon V \to V$ ניל' אז קיים פירוק של V לסכום ישר של $V = \bigoplus U_i$ כאשר U_i הם T־ציקליים.

הוכחה. נמצא ב־V ציקלי מקסימילי. אז קיים $W \subseteq V$ תמ"ו T־שמור כך ש־$V = U_1 \oplus W$. ידוע $T_{|_W} \colon W \to W$ ניל', וכעת באינדוקציה שלמה על $\dim V$. ■

גרסה שקולה למשפט ג'ורדן:

משפט 65. עבוד $T \colon V \to V$ ניל', קיים בסיס B של V שהוא איחור של שרשראות.

בסיס כזה נקרא בסיס מג'רדן ופירושו של דבר היא ש־:

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

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

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

(יש ספרים בהם זה ה־transpose של זה).

משפט 66. עבור $T\colon V\to V$ ניל', 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.

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

השקילות בין המשפט לבין ה"במילים אחרות" נובע מזה שגודל בלוק פחות 1 הוא הממד של התמ"ו שנפרש ע"י הוקטורים בעמודו תהללו.

למה 6. נניח $V=\bigoplus_{i=1}^k U_i$ כאשר U_i הוא T־אינ' (לא נניח אפילו ניל'), 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.

6.3 $\sim$ צורת מייקל ג'ורדן לאופרטור לינארי כללי

הגדרה 52. מרחב העצמי המוכלל של λ הוא מ"ו:

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

הגדרה 53. וקטור עצמי מוכלל הוא וקטור $v\in V$ כך ש־$\exists i\in[n]\colon T^{(i)}v=\lambda v$.

במקרה הכללי:

הגדרה 54. בלוק ג'ורדן אלנטרי עם ערך λ הוא בלוק מהצורה:

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

למה זה הגיוני? כי הסיבה שעשינו מתכתחילה רדוקציה ל־T ניל' היא כי $T-\lambda I$ היא T־אינוו' וניל', ועתה רק נותר להוסיף את ה־λI חזרה לקבלת המקרה הכללי.

הוכחה. באינדוקציה על $n=n(T)$.

- עבור $n=1$, העתקת ה0. V מתפרק לסכום ישר של מרחבים מממד 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:

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

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

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

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 מחילים 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:

$$\left.\begin{aligned} \ker T = \bigoplus_{i=1}^{k} \ker T_{|_{U_i}} \implies \dim\ker T &= \sum_{i=1}^{k} \dim\ker T_{|_{U_i}} = k \\ &= \sum_{i=1}^{k} \dim\ker T_{W_i} = \ell \end{aligned}\right\} \implies k = \ell \implies s = t$$

■

"נראה לי שמי שסיכם את ההרצאה קצת חירטט את הסטודנטים ומי שסיכם את ההרצאה לא הבין את החרטוט" - בן על הסיכום של הסטודנטים. (אני מקווה שאני עושה עבודה יותר טובה).

צורת ג'ורדן לט"ל כללית: נניח ש־$f_T(x)$ מתפצל לחלוטין. כלומר

$$f_T(x) = \prod_j (x - \lambda_j)^{n_j} = \prod_{i=1}^{k} f_{|_{n_i}}(x)$$

כאשר 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 – T־איני' אממ הוא S־איני'. . אז $S_{|_{u_i}}$ היא ניל'. לכן ל־u_i יש בסיס שרשראות B שעבורו $[S_{|U_i}]_B$ מורכבת מבלוקי ג'ורדן ניל' כלומר:

$$\left[S_{|U_i}\right]_B = \begin{pmatrix} \square & & & \\ & \square & & \\ & & \ddots & \\ & & & \square \end{pmatrix}$$

כאשר כל $\square$ מהצורה $J_n(0) \in M_n(\mathbb{F})$. אז:

$$\left[T_{|U_i}\right]_B = \operatorname{diag}\{\square \dots \square\}$$

כאשר כל בלוק מהצורה $J_n(\lambda)$.

(המרצה לא כתב את זה אז אני מוסיף משהו משלי): ומכאן צורת הג'ורדן של המטריצה הזו זה פשוט בלוקים של הצמצומים בבסיס B על גבי עוד מטריצת בלוקים.

משפט 67. צורת ג'ורדן היא יחידה עבור סדר הבלוקים.

אסטרטגיית הוכחה: ניקח צורת ג'ורדן עבור $T\colon V \to V$ ונראה שהיא נקבעת מ־T, V בלבד.

הוכחה. תהא צורת ג'ורדן עבור T תהא צורת ג'ורדן עבור T. קיים בסיס B שעבורו:

$$[T]_B = \operatorname{diag}\{\square_{\lambda_1} \dots \square_{\lambda_k}\}$$

כאשר $\square_{\lambda_i}$ זו דרך מוזרה לכתוב $J_k(\lambda)$. אז:

$$V = \bigoplus_{i=1}^{k} U_i = \bigoplus \bar{v}_\lambda,\ \bar{v}_\lambda = \bigoplus_{i=s}^{\ell} U_i$$

כאשר $\bar{v}_\lambda$ הוא סכום של אי־פריקים שעבורם $T - \lambda I$ ניל'.

מה ניתן להגיד על הממדים של ה־u_i־ים שמרכיבים את $\tilde{v}_\lambda$? הממדים שלהם נקבעים ביחידות, עד כדי סדר, כי היות ש־u_i הם T־איני' הם גם $(T-\lambda I)$־איני', ולכן $[S_{|\tilde{V}_\lambda}]_{B_\lambda}$ היא כ'ורדן ניל' ואז:

$$\left[T_{|\tilde{V}_\lambda}\right]_{B_\lambda} = \left[S_{|\tilde{V}_\lambda}\right]_{B_\lambda}$$

■

הגיון: המרחבים v_λ נקבעים ביחידות ללא תלות בפירוק שבחרנו.

הגיון אחר: כל בלוק מורכבת מהעתקות שבהן $T - \lambda I$ ניל' (פירוק פרימרי).

הערה 21. בהוכחה בסיכום צריך להראות שה־span של הבלוקים הוא באמת $\tilde{v}_\lambda$.

הערה שלי ביחס ללמה צריך את ההוכחה הזאת: כי באיזשהו מקום אם נבחר בסיסים שונים לפירוק אז יכול להיות שדברים מתחרבשים.

6.4 $\sim$ דיבורים ואינטואציה לסוף הנושא

הסיפור של מה שעשינו עד עכשיו: אנחנו חוקרים אופרטורים לינאריים, בצורה שתהיה נוחה להעלות את האופרטור בחזקה. הגענו למסקנה שהכי נוח כשזה לכסין. כשזה קורה, אנחנו יודעים איך לפרק. ראינו כמה אפיונים לזה – גיאומטרי, אלגברי וכו'. ניסינו לעשות מטריצה עם בלוקים על האלכסון במקום, לשם כך, נסתכל על המרחבים שרלוונטיים לבלוקים האלו בלבד. הבנו שבמקום לחקור את ה־T־אינ', נחקור את ה־S־אינ' (הניל' כמו שהגדרתי למעלה). הבנו שהם מורכבים מבלוקי ג'ורדן ניל' אלמטריים, עד לכדי סדר, ואז הרחבנו לצורה הכללית. עברנו דרך חוגים רק כדי להגיד שחוג הפולינום הוא תחום ראשי, ע"מ שנוכל להגדיר פולינום מינימלי המחלק כל פולינום אחר. לא באמת היה צריך חוגים. סתם המרצה רצה לרצוח אותנו. כל הדיבורים על פולינום מינימלי בזכות משפט קיילי־המילטון.

בסיכום אחר שיעלה למודל, [הזהרת הרבה דברים שהמרצה אמר בעפ ולא באמת הבנתי] מתחילים מלפרק את המרחב למרחבים T־ציקליים שלכולם יש פולינום אופייני משל עצמם. הראינו שאם נציב את האופרטור בפולינום האופייני של המטריצה המצורפת זה יתאפס (מה? איפה עשיתי את זה?). ומכאן הפולינום המינימלי של אופרטור הצמצום על המרחב הציקלי מחלק את הפולינום האופייני של ההעתקה שלו.

עכשיו הוא אומר להראות דרך אחרת לפתח צורת ג'ורדן: בגלל ש־$\mathcal{Z}(T,V) \subseteq V$ תמ"ו, ונוכל לקחת $\mathcal{Z}(T,V) \oplus W = V$ (נסמן $\mathcal{Z}(T,V) = U$) אז $f_T(x) = f_{T_{|U}} \cdot f_{T_{|W}}$ (סוף סוף משפט טרוויאלי) והמטריצה המצורפת האקראית ההיא ש־$f_{T_{|U}}$ הפולינום המינילי גם מאפס את $T_{|U}$ והוא שווה ל־$\prod f_{T_{|U_i}}$ כלשהם. מהכיוון הזה אפשר להראות גם את קיילי המילטון, בלי לעבור דרך פיצול מקרים למשולשית/לא משולשית ומשום מה הרחבת שדות באמצע שאיכשהו גם את זה הוכחנו.

המשך בעמוד הבא

7 Bilinear Forms

7.1 ~ הגדרות בסיסיות בעבור תבניות בי־לינאריות כלליות

הגדרה 55. יהי V מ"ו מעל $\mathbb{F}$. פונקציונל לינארי φ מעל V הוא $\varphi\colon V\to\mathbb{F}$.

הערה 22. ראה הרחבה על פונקציונלים לינארים ומרחבים דואלים בסוף הסיכום.

הגדרה 56. יהיו V,W מ"וים 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 $w\mapsto f(v_0,w),\ \ v\mapsto(v,w_0)$ הן פונקציונליים לינאריים.

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

משפט 68. באופן שקול: יהיו $\forall v\in V,\ \ w\in W,\ \ \alpha\in\mathbb{F}$:

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

בסופו של דבר נתמקד בסוג מסויים של העתקות בילינאריות, הן מכפלות פנימיות.

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 הסיבה שאנו מתעסקים ספציפית עם העתקות בילינאריות.

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

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

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 לינאיריים וכן $\psi_1\dots\psi_k\colon W\to\mathbb{F}$ פונקציונליים לינארים. אז ההעתקה הבאה בילינארית: $f(v,w)=\sum_{i=1}^k\varphi_i(v)\psi_i(w)$

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

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

הערה 23. בעתיד נראה שכל תבנית בילינארית נראית כמו מקרה 5.

משפט 69. 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 70. 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

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

העתקה בילינארית.

הוכחה. נקבע v כלשהו:

$$[v]_{\mathcal{A}}^T\cdot A=B\in M_{1\times m},\ \ g(w)=f(v,w),\ \ g(w_1+w_2)=B[w_1+w_2]_{\mathcal{B}}=B[w_1]_{\mathcal{B}}+B[w_2]_{\mathcal{B}}$$

כנ"ל עבור כפל בסקלר. נקבע w, אז $C=A[w]_{\mathcal{B}}\in M_{n\times 1}(\mathbb{F})$:

$$h(v)=f(v,w)\ h(v)=[v]_B^T\cdot C,\ \ h(v_1+v_2)=[v_1+v_2]_{\mathcal{B}}^T=([v_1]_{\mathcal{B}}^T+[v_2]_{\mathcal{B}}^T)C=h(v_1)+h(v_2)$$

■

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

הגדרה 58. בהינתן תבנית בי־לינ׳ $f\colon V\times W\to\mathbb{F}$ 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 $(A)_{ij}=f(v_i,w_j)$ (נסמן $\mathcal{B}=(w_i)_{i=1}^m$ $\mathcal{A}=(v_i)_{i=1}^n$)

משפט 71. $f(v,w)=[v]_{\mathcal{A}}^TA[w]_{\mathcal{B}}$

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

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

ומכאן פשוט נזרוק אלגברה:

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

■

סימון 7. נאמץ לסיכום הזה את הסימון $[f]_{\mathcal{A},\mathcal{B}}$ עבור המטריצה המייצגת של f בי־לינארית.

(זהו **אינו** סימון רשמי בקורס אבל בהחלט צרדי להיות)

משפט 72. 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: $\psi(f)=0\in M_{n\times m}\implies\forall i,j\in[n]\times[m]\colon f(v_i,w_j)=0$ ולכן $\forall v\in V,w\in W\colon f(v,w)=\sum_{i,j\in[n]\times[m]}\alpha_if(v_i,w_j)\beta_j=0$ (עם אותם הסימונים כמו קודם)
- **על.** תהי An$M_{n\times m}(\mathbb{F})$. נגדיר $f(v,w)=[v]_{\mathcal{A}}^TA[w]_{\mathcal{B}}$ ואכן $f(v_i,w_j)=e_i^TAe_j=(A)_{ij}$.

■

משפט 73. יהיו V,W מ"וים 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:

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

מצד אחד:

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

כדרוש. ■

הגדרה 59. עבור $f\in B(V,W)$ נגדיר את $\operatorname{rank} f=\operatorname{rank} A$ כאשר A מייצגת אותהת ביחס לבסיסים כלשהם.

משפט 74. $\operatorname{rank} f$ מוגדר היטב

הוכחה. כפל בהפיכה לא משנה את דרגת המטריצה (ו־transpose של מטריצה הוא הפיך), דהיינו מטריצות שינוי הבסיס לא משנות את דרגת המטריצה. ■

מסקנה 13. תהא $f\in B(V,W)$ ונניח $\operatorname{rank} f=r$. אז קיימים בסיסים $\mathcal{A},\mathcal{B}$ של W,V בהתאמה כך ש־$[f]_{\mathcal{A},\mathcal{B}}=\left(\begin{smallmatrix}I_R&0\\0&0\end{smallmatrix}\right)$

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.

7.2 $\sim$ חפיפה וסימטריות

הגדרה 60. 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 75. מטריצות חופפות אמ"מ הן מייצגות את אותה התבנית הבילינארית.

משפט 76. אם A,A' חופפות, אז:

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

הוכחה. הגדרנו $\operatorname{rank} f$ כאשר f בילינ' 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 24. יש שדות שמעליהם טענה 2 לא מעניינת במיוחד

הגדרה 61. תבנית f מעל V נקראת סימטרית אם: $\forall v,w\in V:\ f(v,w)=f(w,v)$

הגדרה 62. תבנית f מעל V נקראת אנטי־סימטרית אם: $\forall v,w\in V:\ f(v,w)=-f(w,v)$

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 סימ' ו־ψ א־סימ' וכן $f=\varphi+\psi$.

משפט 77. תהי f תבנית בילינ' 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 סימ'/אסימ' אממ A סימ'/אסימ'.

הוכחה. $\Longrightarrow$ אם f סימ'/אסימ', 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 סימ' אז:

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

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

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

■

7.3 ~ תבנית ריבועית

הגדרה 63. 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 8. עבור תבנית בילינארית f על V, נגדיר את $\hat{f}(u,v)=f(v,u)$

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

משפט 78. תהי f תבנית בילי' סימ' על V, ונניח ש־$\operatorname{char}\mathbb{F}\neq 2$, אז:

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

הוכחה.

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

■

עבור 1, עתה נוכיח את 2: נניח $\forall v\in V: Q_f(v)=0$ אז

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

למה שונה ממציין 2 חשוב:

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

הערה 25. אין ממש טעם להגדיר תבנית ריבועית על תבנית בילינארית שאיננה סימטרית. הסיבה? כל מטריצה יכולה להיות מפורקת לחלק סימטרי וחלק אנטי־סימטרי, החלק האנטי־סימטרי יתאפס (אלכסון אפס) ורק החלק הסימטרי ישאר בכל מקרה. זה לא שאי־אפשר, זה פשוט לא מעניין במיוחד.

משפט 79. 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.

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

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 ממימד $n-1$. 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:

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

■

7.4 ~ הסיגנטורה ומשפט ההתאמה של סילבסטר

משפט 80. לכל f תבנית סימ' קיימת מטריצה מייצגת מהצורה $\left(\begin{smallmatrix} I_r & 0 \\ 0 & 0\end{smallmatrix}\right)$ כאשר $\mathbb{F}=\mathbb{C}$ (או סגור אלגברית כלשהו)
.

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

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

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. ולכן $B'=(v_1'\dots v_r',v_{r+1}\dots v_n)$ 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:

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

■

בשיעורי הבית נראה ש־: נניח ש־f אנטי־סימטרית לא מנוונת (לא תבנית ה־0), אז תמיד ישנה מטריצה מייצגת מהצורה (תחפשו "מטריצה סימפלקטית" בגוגל, זה קצת סיוט לעשות את זה בלאטך). הרעיון הוא אם:

$$\hat{I}_n=\begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix},\ J=\begin{pmatrix} 0 & & -\hat{I}_n \\ & \ddots & \\ \hat{I}_n & & 0\end{pmatrix}$$

אז J סימפלקטית.

הגדרה 64. יהי V מ"ו מעל $\mathbb{R}$ ו־f תבנית בילינ' מעל V. נאמר ש־f חיובית/אי־שלילית/שלילית/אי־חיובית אם $\forall 0\neq v\in V$ מתקיים ש־$f(v,v)\le 0$/$f(v,v)<0$/$f(v,v)\ge 0$/$f(v,v)>0$

משפט 81. תהא A מטריצה מייצגת של תבנית בי־לינארית סימ', עם ערכים $0,-1,1$ בלבד על האלכסון, מקיימת:

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

הוכחה.

$\Longleftarrow$ ברור

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

■

משפט 82. משפט ההתאמה של סילבסטר. p, q הנ"ל נקבעים ביחידות.

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

ההוכחה של קרני. נכתבה ונמחקה מהלוח. שימו לב שה־tr לא נשמר בשינוי בסיס של תבניות בילינאריות, זה לא העתקות. ההוכחה שגויה. ■

הוכחה. 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 $t \leq p$, 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 $f(v,v) \leq 0$ כי $v \in W$ וסתירה). ידוע ש־$U \oplus W \subseteq V$ תמ"ו וכן $\dim U + \dim W \leq \dim V$. נציב ונקבל $\dim V = t + s + k < p + s + k$, סתירה. לכן p, q נקבעים ביחידות. ■

סימון 9. ה־(p,q) לעיל נקראים הסיגנטורה של f.

(תזהרו, הסיגנטורה תתקוף אותנו אח"כ)

המשך בעמוד הבא

Inner Product Vector Spaces 8

8.1 ~ הגדרה כללית

8.1.1 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 65. יהי V מ"ו, מכפלה פניימית מעל $\mathbb{R}$ היא תבנית בילינ' סימטרית חיובית מעל V, ומסומנת $f(v,u) = \langle v, u\rangle$ (ויש ספרים שמסמנים $\langle v \mid u\rangle$, $\langle \cdot \mid \cdot \rangle : V \times V \to \mathbb{R}$.

סימון 10. בקורס מסמנים $\langle \cdot, \cdot \rangle$ אבל אני מגניב אז אני משתמש ב־$\langle \cdot \mid \cdot \rangle$.

למה 7. $\forall v \in V\colon\ \langle v \mid v\rangle \geq 0$ 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 66. 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 83. $V = M_n(\mathbb{R})$, אז $\langle A \mid B\rangle = \operatorname{tr}(A \cdot B^T)$ אז $(V, \langle \cdot \mid \cdot \rangle)$ ממ"פ.

דוגמה מגניבה. בהינתן $V = c[0,1]$, 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 84. (שהפליצו מחדו"א) אם $f \geq 0$ אינטרבילית (זה נשמע כמו מפלצת) על קטע $[a,b]$ וגם ישנה נקודה חיובית $c \in [a,b]$ שעבורה $f(x) \geq 0$ וגם f רציפה ב־c, אז $\int_a^b f(x)\, \mathrm{d}x > 0$.

8.1.2 מעל $\mathbb{C}$

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

הגדרה 67. יהי V מ"ו מעל $\mathbb{C}$. מכפלה פנימית $\langle \cdot \mid \cdot \rangle : V \times V \to \mathbb{C}$ מקיימת:

- ליניאירות ברכיב הראשון: אם נקבע v, אז $u \mapsto \langle v \mid u\rangle$ לינארית.
- ססקווי־ליניאריות ברכיב השני: $\langle u_1 + u_2 \mid v\rangle = \langle u_1 \mid v\rangle + \langle u_2 \mid v\rangle \wedge \langle u \mid \alpha v\rangle = \bar{\alpha} \langle u \mid v\rangle$
  כאשר $\bar{\alpha}$ הצמוד המרוכב של α.
- הרמטיות: $\langle v \mid u\rangle = \overline{\langle u \mid v\rangle}$
- $\forall 0 \neq v \in V\colon\ \langle v \mid v\rangle > 0 \wedge \langle 0 \mid 0\rangle = 0$

למעשה – נבחין שאין צורך בממש ססקווי־ליניאירות 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.

(אופס! בן הגדיר את זה לליניאירות ברכיב השני, כלומר הפוך, כי ככה עושים את זה בפתוחה. תיקנתי בסיכום אבל יכול להיות שיש משהו הפוך כי פספסתי. זה אמור להיות ליניארי ברכיב השני).

הגדרה 68. $\bar{B}^T = B^*$

הגדרה 69. יהי ממ"פ V מ"ו מעל $\mathbb{F}$. לכל $v \in V$ מגדירים את הנורמה של v להיות $||v|| = \sqrt{\langle v \mid v\rangle}$

מאקסיומות החיוביות:

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

וכן:

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

8.2 ∼ הקשרים גיאומטריים של מכפלה פנימית

הגדרה 70. יהי V מ"ו מעל $\mathbb{F}$, ו־$||\cdot|| : V \to \mathbb{R}_{\geq 0}$, אז $(V, ||\cdot||)$ יקרא מרחב נורמי.

משפט 85. ("נוסחאת הפולריזציה") בהינתן $(V, ||\cdot||)$ מרחב נורמי;

גרסה מעל $\mathbb{R}$:

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

גרסה מעל $\mathbb{C}$:

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

הוכחה (ל־$\mathbb{C}$).

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

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

מנוסחאת הפולריזציה, נוכל לשחזר באמצעות נורמה את המכפלה הפנימית. במילים אחרות, ממ"פ ומרחב נורמי הם אותו הדבר.

הגדרה 71. בהינתן $(v, \langle \cdot \mid \cdot \rangle)$ ממ"פ, לכל $v \in V$ נאמר ש־u מאונך ל־v (או אורתוגולי ל־ אם אנחנו מרגישים מפונפנים) ונסמן $v \perp u$ אם $\langle u \mid v \rangle = 0$.

הערה 26. אם $v \perp u$ אז $u \perp v$. (כי צמוד של 0 הוא 0).

משפט 86 (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 $\langle v \mid u \rangle = 0$. נפתח אלגברה:

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

■

הערה 27. בעבור $v = \mathbb{R}^n$ מ"פ סטנדרטית אז $||v||$ מזדהה עם מושג הגודל של וקטור בגיאומטריה רגילה.

הערה 28. בתוך $\mathbb{R}^n$ הוקטורים הסטנדרטיים מאונכים אחד לשני (במכפלה הפנימית הסטנדרטית) ולכן $\langle e_i \mid e_j \rangle = \delta_{ij}$ 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 29. מעל $\mathbb{R}$ מקבלים אמ"מ למשפט פיתגורס, מעל $\mathbb{C}$ לאו דווקא.

משפט 87. (אי שוויון קושי־שוורץ)

$$\forall v, u \in V: \ |\langle u \mid v \rangle| \leq ||u|| \cdot ||v||$$

ושוויון אמ"מ u, v ת"ל.

הערה 30. זה בפרט נכון בגיאומטריה סטנדרטית ממשפט הקוסינוסים.

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

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

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

$$\begin{aligned}
&\begin{cases} u - \alpha v \perp v \\ u - \alpha v \perp v \end{cases} \\
\implies & \left|\left|u\right|\right|^2 = \left|\left|(u - \alpha v + \alpha v)\right|\right|^2 = \overbrace{\left|\left|u - \alpha v\right|\right|^2}^{\geq 0} + \left|\alpha\right|^2 \left|\left|v\right|\right|^2 \\
& \geq \left|\alpha\right| \cdot \left|\left|v\right|\right|^2 = \frac{\left|\langle v \mid u \rangle\right|^2}{(\left|\left|v\right|\right|^2)^2} = \left|\left|v\right|\right|^2 = \frac{\left|\langle u \mid v \rangle\right|^2}{\left|\left|v\right|\right|^2} \\
\implies & \left|\langle v \mid u \rangle\right|^2 \leq \left|\left|v\right|\right| \cdot \left|\left|u\right|\right|
\end{aligned}$$

בפרט $\left|\left|u - \alpha v\right|\right|^2 = 0$ אמ"מ הם תלויים לינארית ומכאן הכיוון השני של המשפט. ■

הערה 31. המשפט למעלה לא אומר כלום כי מגדירים קוסינוס לפי המכפלה הפנימית הסטנדרטית.

דוגמאות.

1. ממכפלה פנימית סטנדרטית:

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

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

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

כאשר $f^2 = f \cdot f$ (לא הרכבה).

3. אי־שוויון המשולש:

$$\forall u, v \in V\colon \ \left|\left|u + v\right|\right| \leq \left|\left|u\right|\right| + \left|\left|v\right|\right|$$

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

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

$$\left|\left|u + v\right|\right|^2 = \left|\left|u\right|\right|^2 + \left|\left|v\right|\right|^2 + 2\Re(\langle u \mid v \rangle) \leq \left|\left|u\right|\right|^2 + \left|\left|v\right|\right|^2 + 2\left|\langle u \mid v \rangle\right|$$

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

$$\leq \left|\left|u\right|\right|^2 + 2\left|\left|u\right|\right|\left|\left|v\right|\right| + \left|\left|v\right|\right|^2 = (\left|\left|u\right|\right| + \left|\left|v\right|\right|)^2$$

■

8.3 ~ אורתוגונליות

סימון 11. יהי $(V, \langle \cdot \mid \cdot \rangle)$ 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 72. $T^{\perp}$ הוא תת־המרחב הניצב ל־T.

משפט 88. 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 32. שוויון בג' מתקיים אמ"מ $\operatorname{span} S = \operatorname{span} T$.

הגדרה 73. משפחה של וקטורים $A \subseteq V$ נקראת אורתוגונלית אם $\forall u \neq v \in V \colon u \perp v$

הערה 33. אם A משפחה אורתוגונלית וגם $0 \notin A$ אז ניתן לייצור ממנה משפחה של וקטורים אורתוגונלים שהם גם וקטורי יחידה, ע"י נרמול.

הגדרה 74. משפחה של וקטורים $A \subseteq V$ נקראת אורתונורמלית, אם היא אורתוגונלית ובנוסף כל הוקטורים הם וקטורי יחידה.

הגדרה 75. 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 89. בסימונים לעיל, $\forall u \in U \colon \; ||v-u|| \geq ||v - p_U(v)||$ 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 $u \perp v - p_U(v) - v$. 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 $||v-u||^2 \geq ||v - p_U(v)||^2$. ושוויון אמ"מ $||u - p_U(v)|| = 0$ אמ"מ $u = p_U(v)$. ■

עתה נוכיח את יחידות ההטלה האורתוגונלית (קיום נוכיח בהמשך באופן קונסטקרטיבי)

משפט 90. ההטלה הניצבת (אם קיימת), היא יחידה.

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

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

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

משפט 91. תהי $A \subseteq V$ משפחה אורתוגונלית ללא 0. אז היא בת"ל.

הוכחה. יהיו $v_1 \ldots v_n \in A$ וכן $\alpha_1 \ldots \alpha_n \in \mathbb{F}$, 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 92 (קיום היטל אורתוגונלי)**.** נניח ש־$u \subseteq V$ תמ"ו. נניח U נ"ס וכן $B = (e_1 \ldots e_n)$ 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 $\forall u\in U:\ \langle v-p_U(v)\,|\,u\rangle=0$ אך לגבי התנאי האחרון די להוכיח $\forall j\in[n]:\ \langle v_i p_U(v)\,|\,e_j\rangle=0$. 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)

משפט 93 (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 $\forall k\in[n]:\ \operatorname{span}(b_1\dots b_k)=\operatorname{span}(u_1\dots u_k)$.

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 $\operatorname{span}(b_1\dots b_{k+1})\subseteq\operatorname{span}(u_1\dots u_{k+1})$. 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 94. 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 ולינארית. ■

משפט 95. $\min_{u\in U}||v-u||=||v-p_U(v)||$

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 76. הפתרון האופטימלי למערכת משוואות $(A\mid b)$ הוא $p_{\operatorname{Col}A}(b)$ (כאשר $\operatorname{Col}A$ מ"ו העמודות).

8.4 ~ צמידות

הגדרה 77. 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 $\langle v\mid u\rangle=v^Tu$:

$$\langle T_Av\mid u\rangle=(Av)^Tu=v^TA^Tu=\left\langle v\mid A^Tu\right\rangle$$

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 96. ההעתקה $v\mapsto p_U(v)$ עבור U תמ"ו כלשהו, היא ההיטל, צמודה לעצמה.

משפט 97. העתקה סימטרית אמ"מ היא דומה למטריצה סימטרית.

משפט 98. 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־$3\implies 2+1$. 1 נובע ישירות מהגדרה. נוכיח את 2.

הוכחה ל־2. נניח $S\circ T$ צמודה לעצמה. בהנחות המשפט ידוע S,T צמודות לעצמן. נקבל:

$$\langle(S\circ T)v\mid u\rangle=\langle v\mid STu\rangle=\langle Sv\mid Tu\rangle=\langle TSv\mid u\rangle\implies\langle(ST-TS)v\mid u\rangle=0\quad\forall v,u$$

נסיק:

$$\implies\forall v\,\langle(ST-TS)v\mid(ST-TS)v\rangle=0\implies(ST-TS)v=0\implies STv=TSv\implies\top$$

מהכיוון השני:

$$\langle STv\mid u\rangle=\langle S(Tv)\mid u\rangle=\langle v\mid TSu\rangle=\langle v\mid STu\rangle$$

■

הגדרה 78. $T\colon V\to V$ תקרא חיובית/אי־שלילית/שלילית/אי־חיובית אם:

חיובית: $\langle Tv\mid v\rangle>0$	אי־שלילית: $\langle Tv\mid v\rangle\geq 0$
שלילית: $\langle Tv\mid v\rangle<0$	אי־חיובית: $\langle Tv\mid v\rangle\leq 0$

משפט 99. אם T חיובית, אז היא הפיכה (כנ"ל לשלילית)

הוכחה. נניח ש־T לא הפיכה, נקרא שהיא לא חיובית. קיים $0\neq v\in V$. אז $v\in\ker T$, אז $\langle Tv\mid v\rangle=\langle 0\mid v\rangle=0$, בסתירה לכך ש־$T$ חיובית. ■

משפט 100. נניח ש־S צמודה לעצמה, אז S^2 צמודה לעצמה ואי־שלילית.

הוכחה. ממשפט קודם S^2 צמודה לעצמה. נוכיח אי־שלילית:

$$\forall 0\neq v\in V\colon\ \left\langle S^2v\,\middle|\,v\right\rangle=\langle Sv\mid Sv\rangle=||Sv||^2\geq 0$$

■

הגדרה 79. פולינום $p\in\mathbb{R}[x]$ יקרא חיובי אם $\forall x\in\mathbb{R}p(x)>0$.

מסקנה. 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 8. אם $p\in\mathbb{R}[x]$ חיובי, אז קיימים $g_1\dots g_k\in\mathbb{R}[x]$ וכן $0<c\in\mathbb{R}$ כך ש־$p(x)=\sum_{i=1}^k g_i^2(x)+c$

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:

$$\langle p(T)v \mid v\rangle = \underbrace{\left\langle \sum_{i=1}^{k} g_i^2(T)v \,\middle|\, v \right\rangle}_{\sum_{i=1}^{k}\langle g_i^2(T)v \mid v\rangle \geq 0} + \overbrace{c\,\langle v \mid v\rangle}^{c||v||^2>0} \geq 0$$

■

מסקנה. אם $T\colon V \to V$ צמודה לעצמה ו־$p(x) \in \mathbb{R}[x]$ פולינום חיובי, אז $p(T)$ הפיכה.

הוכחה. "תסתכלו על צד ימין של הלוח" $\sim$ המרצה ■

משפט 101. נניח ש־$T\colon V \to V$ סימטרית (צמודה לעצמה מעל $\mathbb{R}$/המייצגת סימטרית) ויהי $m_T(x)$ הפולינום המינימלי של T. אז m_T מתפרק לגורמים לינארים. בנוסף, הם שונים זה מזה.

מסקנה. T סימטרית היא לכסינה.

הוכחה. נניח בשלילה קיום $p \mid m_T$ ו־$\deg p \geq 2$, pאי־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. ■

למה 9. נניח T סמטרית ו־$\lambda \in \mathbb{R}$, אם $(T-\lambda I)^2 = 0$ אז $T - \lambda I = 0$.

הוכחה. ידוע:

$$\forall v\colon 0 = \left\langle (T-\lambda I)^2 v \,\middle|\, v \right\rangle = \langle (T-\lambda I)v \mid (T-\lambda I)v\rangle = ||(T-\lambda I)v||^2 \implies (T-\lambda I)v = 0$$

■

משפט 102. 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:

$$\lambda v\,||v||^2 = \langle \lambda v \mid v\rangle = \langle Tv \mid v\rangle = \langle v \mid Tv\rangle = \langle v \mid \lambda v\rangle = \overline{\lambda}\,||v||$$

ידוע $v \neq 0$ ולכן $||v|| \neq 0$ ונסיק $\lambda v = \bar{\lambda}$ ולכן $\lambda \in \mathbb{R}$. ■

משפט 103. 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:

$$\alpha \langle u \mid v\rangle = \langle \alpha u \mid v\rangle = \langle Tu \mid v\rangle = \langle u \mid Tv\rangle = \langle u \mid \beta v\rangle = \bar{\beta}\langle u \mid v\rangle = \beta \langle u \mid v\rangle$$

בגלל ש־$\beta \in \mathbb{R}$ מתקיים $\beta = \bar{\beta}$. ולכן $(\alpha - \beta)\langle v \mid u\rangle = 0$ מהעברת אגף וסה"כ $\langle u \mid v\rangle = 0$ ואכן $u \perp v$. ■

הערה 34. 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 104. (משפט ריס). יהי V ממ"פ סופי ויהי $\varphi \in V^*$. אז קיים ויחיד וקטור $u \in V$ שמקיים $\forall v \in V\colon \varphi(v) = \langle v \mid u\rangle$.

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 $\forall v\in V: \varphi(v)=\langle v\,|\,u\rangle$ מספיק להראות תכונה זו לאברי הבסיס B, כלומר נראה ש־$\forall 1\le j\le n: \varphi(b_j)=\langle b_j\,|\,u\rangle$.
ואכן:

$$\langle b_j\,|\,u\rangle=\left\langle b_j\,\middle|\,\sum_{i=0}^n \overline{\varphi(b_i)}b_i\right\rangle=\sum_{i=1}^n \underbrace{\overline{\overline{\varphi(b_i)}}}_{b_i}\underbrace{\langle b_j\,|\,b_i\rangle}_{ij}=b_j \quad \top$$

יחידות: אם קיים וקטור נוסף שעבורו $\forall v\in V: \varphi(v)=\langle v\,|\,w\rangle$ 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. ■

משפט 105. יהי V ממ"פ מנ"ס ותהי $T\colon V\to V$ לינארית. אז קיימת ויחידה $T^*\colon V\to V$ ומקיימת $\forall u,v\in V\colon\ \langle Tu\,|\,v\rangle=\langle u\,|\,T^*v\rangle$.

הגדרה 80. ההעתקה T^* לעיל נקראת ההעתקה הצמודה ל־T.

הוכחה. לכל $v\in V$, נתבונן בפונקציונל הלינארי $\varphi_V\in V^*$ המוגדר ע"י $\forall u\in V: \varphi_V(u)=\langle Tu\,|\,v\rangle$. 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 $\alpha,\beta\in\mathbb{F}$ ועבור $v,w\in V$ 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. ■

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 106. (תכונות ההעתקה הצמודה) יהי V ממ"פ ותהיינה $T,S\colon V\to V$ זוג העתקות לינאריות. נבחין ש־:

(א) $(T^*)^*=T$

(ב) $(T\circ S)^*=S^*\circ T^*$

(ג) $(T+S)^*=T^*+S^*$

(ד) $\forall\lambda\in\mathbb{F}\colon(\lambda T)^*=\bar\lambda(T^*)$

"זה אחד וחצי לינאריות"

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) כנ"ל

■

סימון 12. העתקה צמודה לעצמה לעיתים קרובות מסמנים ב־T

משפט 107. בהינתן B אורתונורמלי של V אז $[T^*]_B=[T]_B^*$ (שימו לב: האחד צמוד מטריציוני, והשני העתקה צמודה)

משפט 108. T צמודה לעצמה אמ"מ $\forall v\in V:\ \langle Tv\,|\,v\rangle\in\mathbb{R}$.

הוכחה.

$$
\begin{aligned}
&T \text{ צמודה לעצמה}\\
\iff & T=T^* \iff T-T^*=0\\
\iff & \forall v\in V:\ \langle (T-T^*)v\,|\,v\rangle=0\\
\iff & \forall v\in V:\ \langle Tv\,|\,v\rangle-\langle T^*v\,|\,v\rangle=0\\
\iff & \forall v\in V:\ \langle Tv\,|\,v\rangle-\overline{\langle Tv\,|\,v\rangle}=0\\
\iff & \forall v\in V:\Re(\langle Tv\,|\,v\rangle)+\Im(\langle Tv\,|\,v\rangle)-\Re(\langle Tv\,|\,v\rangle)+\Im(\langle Tv\,|\,v\rangle)=0\\
\iff & \forall v\in V:2\Im(\langle Tv\,|\,v\rangle)=0 \iff \forall v\in V:\Im(\langle Tv\,|\,v\rangle)=0 \iff \forall v\in V:\ \langle Tv\,|\,v\rangle\in\mathbb{R}\quad\top
\end{aligned}
$$

■

המשך בעמוד הבא

9 Decompositions

9.1 ~ המשפט הספקטרלי להעתקות

9.1.1 ניסוח להעתקות צמודות לעצמן

משפט 109. (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־$\forall 1\le i\le m\colon d_i=1$. 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, מהטענה השנייה שהוכחנו. נבנה בסיס $B_i\subseteq\ker(T-\lambda_i)$ וסה"כ $\bigcup_{i=1}^m B_i$ בסיס אורתוגונלי של T. ■

משפט 110. יהי V נ"ס מעל $\mathbb{R}$ ותהי $T\colon V\to V$ ט"ל. אז T צמודה לעצמה 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:

$$\langle Tu\,|\,v\rangle=\left\langle T\left(\sum_{i=1}^m \alpha_i b_i\right)\middle|\sum_{i=1}^m \beta_i b_i\right\rangle=\sum_{i=1}^n\sum_{j=1}^n \alpha_i\beta_j\langle Tb_i\,|\,b_j\rangle=\sum_{i=1}^n\sum_{j=1}^n \alpha_i\beta_j\lambda_i\underbrace{\langle b_i\,|\,b_j\rangle}_{\delta_{ij}}=\sum_{i=0}^n \alpha_i\beta_i\lambda_i$$

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 לא נכון 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.

מכאן ואילך המרצה מוכיחה את המשפט הספקטרלי ללא המשפט היסודי של האלגברה. לשם כך, צריך להראות שהפולינום המינימלי מתפצל למכפלה של גורמים לינארים מעל המרוכבים.

משפט 111. אם $p(x)$ פולינום אי־שלילי, אז נוכל להציגו כ־$p(x)=\sum_{i=1}^k g_i(x)^2+c$, כאשר $c\ge 0$, וכמו כן $c>0$ אמ"מ $p(x)$ פולינום חיובי.

9.1.2 מבוא למשפט 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 112. 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 $\forall 1\le i\le n$ 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 $i\ne j\in[n]$ 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 $TT^*=T^*T$.

הוכחה. לפי הטענה הקודמת כל b_i הוא ו"ע משותף ל־T ול־T^*, ולכן:

$$TT^*(b_i)=T(T^*(b_i))=\beta_iT^*(b_i)=\beta_i\delta_ib_i=\alpha_i\beta_ib_i=\alpha_iT^(b_i)=T^*T(b_i)$$

העתקה מוגדרת לפי מה שהיא עושה לבסיס ולכן $TT^*=T^*T$. ■

הגדרה 81. העתקה כזו המקיימת $AA^*=A^*A$ נקראת נורמלית (או "נורמאלית" בעברית של שנות ה־60).

עתה, ננסה להראות שכל העתקה נורמלית מקיימת את התנאי של המשפט הספקטרלי (כלומר ניתן ללכסנה אורתוגונלית)

9.1.3 הוכחת המשפט הספקטרלי בעבור העתקה כללית

משפט 113. (המשפט הספקטרלי) 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 10. יהי V ממ"פ ותהיינה $S_1,S_2\colon V\to V$ זוג ט"ל צמודות ולעמן 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־$V=\bigoplus_{i=1}^m\ker(S_1-\lambda_iI)$, כאשר $\lambda_1\dots\lambda_m$ הע"עים השונים של S_1. לכל $1\le i\le m$ מתקיים ש־V_{λ_i} (המרחב העצמי) הוא S_1־אינווריאנטי שהרי אם $v\in V_{\lambda_i}$ ונחשב:

$$S_1(S_2v)=S_2(S_1v)=S_2(\lambda_iv)=\lambda_iS_2v\implies S_2v\in V_{\lambda_i}$$

כאשר $S_2|_{V_{\lambda_i}}\colon V_{\lambda_i}\to V_{\lambda_i}$ 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 המפשט הספקטרלי.

- $\Longrightarrow$ לפי המסקנה הקודמת, אם ישנו לכסון אורתוגונלי T בהכרח נורמלית.
- $\Longleftarrow$ 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 $S_1b_i=\alpha_ib_i,\ S_2b_i=\beta_ib_i$. 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 $\forall i\in[n]\colon T(b_i)=S_1(b_i)+iS_2(b_i)=\alpha_ib_i+i\beta_ib_i=(\alpha+i\beta_i)b_i$ 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.

"אגב – לא השתמשתי במשפט היסודי של האלגברה"

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, ולכן נסיק שלצמודה לעצמה קיים בסיס אורתוגונלי מלכסן (הכיוון ההפוך לא נכון מעל המרוכבים).

9.2 $\sim$ צורה קאנונית למטריצות נורמליות מעל הממשיים

תזכורת: $A\in M_n(\mathbb{C})$ נקראית סימטרית אממ $A=A^T$ והרמיטית אם $A=A^*$, ונורמלית אממ $AA^*=A^*A$.

משפט 114. 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:

$$c_ij=\langle T^*e_j\,|\,e_i\rangle=\langle e_j\,|\,Te_i\rangle=\overline{\langle Te_i\,|\,e_j\rangle}=a_{ij}$$

■

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 $\mathbb{R}$).

משהו על אינטרפולציות:

משפט 115. יהיו $x_1\dots x_n,y_1\dots y_n\in\mathbb{R}$. נניח $x_i\neq x_j \implies \forall i,j\in[n]\colon i\neq j$. אז $\exists! p\in\mathbb{R}_{\leq n-1}[x]\colon \forall i\in[n]\colon p(x_i)=y_i$ עד לכדי חברות (באופן שקול: נניח p מתוקן)

(הערה מהידע הכללי שלי: זהו פולינום לגראנג' והוא בונה אינטרפולציה די נחמדה).

הוכחה. ידוע שהפולינום מהצורה $p(x)=\sum_{k=0}^{n-1}a_kx^k=(1,x,x^2,\dots x^{n-1})(a_0\dots a_{n-1})^T$. למעשה, נקבל את מטריצת ונדרמונד:

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1}\end{pmatrix}\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1}\end{pmatrix}=\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n\end{pmatrix}$$

וידוע שהדטרמיננטה של ונדרמונד היא $\prod_{i<j}(x_i-x_j)$ וזה איכשהו אמור לגמור את ההוכחה.

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. ■

משפט 116. תהי $A\in M_n(\mathbb{F})$ נורמלית, אז קיים פולינום $f(x)\in\mathbb{R}[x]$ כך ש־$A^*=f(A)$.

הערה: באופן כללי **לא** נכון שאם A,B מתחלפות אז $\exists f(x)\in\mathbb{F}[x]\colon f(A)=B$.

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־$\deg f=n-1$. ■

ננסה להבין מי הן $A \in M_2(\mathbb{R})$ שהן נורמליות. מעל $\mathbb{C}$ הן פשוט לכסינות. נבחין ש־:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},\ A^* = f(A) \in \mathbb{R}_{\le 1}[x] = \alpha A + BI,\ A = \alpha A^T + \beta I = \alpha(\alpha A + \beta I) + \beta I$$

$$\implies \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \alpha^2 a + \beta(\alpha+1) & \alpha^2 b \\ \alpha^2 c & \alpha^2 d + \beta(\alpha+1) \end{pmatrix}$$

$$\begin{cases} (b \wedge c \neq 0) \quad \alpha = 1 \implies A = A + 2\beta I \implies \beta = 0,\ A = A^T \cancel{+\beta I} \\ (b \wedge c \neq 0) \quad \alpha = -1 \implies A^T = -A + \beta I \implies A + A^T = \begin{pmatrix} \beta & 0 \\ 0 & \beta \end{pmatrix} \implies b = -c,\ A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \\ (b \vee c = 0) \implies A = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} \end{cases}$$

המקרה השני – זה פשוט סיבובים, אבל בניפוח (כי הדטרמיננטה היא $a^2 - b^2$).

בכל מקרה, מסקנה מהמשפט הקודם.

משפט 117. אם $T\colon V \to V$ נורמלית, אז $\exists f \in \mathbb{R}[x]\colon f(T) = T^*$.

הוכחה. נבחר בסיס א"נ $A = [T]_B \iff A^* = [T^*]_B$. 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־$U \oplus W = B$. 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 $U \subseteq V \implies V = U \oplus U^\perp$. ניעזר בכך כדי להוכיח את המשפט הבא:

משפט 118. אם $U \subseteq V$ תמ"ו אינו' ביחס ל־T אז $U^\perp$ הוא T^*־אינו'.

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 119. בעבור $T\colon V \to V$ נורמלית, אם היא U־אינו' אז גם T^* הוא U־אינו'

הוכחה. נבחין ש־$T^* = f(T)$ כלשהו, וכן U הוא T־איוו' ולכן U הוא $f(T)$־איוו' וכאן די גמרנו את ההוכחה. ■

מסימטריות $U^\perp$ הוא T^*, מהמשפט גם $(T^*)^*$ איונ' ולכן T־אינו'.

משפט 120. יהי V מעל $\mathbb{R}$ מ"ו וכן $T\colon V \to V$ ט"ל. אז קיים $U \subseteq V$ שהוא T־איונ' וממדו לכל היותר T.

הערה: מעל $\mathbb{C}$ "זה מטופש" כי הפולינום מתפרק (ואז המרחב העצמי 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־$\mathbb{R}$ הוא לינארי הוא ממעלה 2, מהמשפט היסודי של האלגברה ומהעובדה ש־$m_T(\bar{x}) = 0 \implies m_T(x) = 0$.

- אם g לינארי אז יש ע"ע ממשי של T מה שנותא U (שנפרש ע"י הו"ע) ממד 1.
- אם $\deg g = 2$ כמובן שניתן להניח g מתוקן. 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 נשמר תחת T.

■

הערה: בעבור נורמלית הטענה נכונה ללא תלות במשפט היסודי של האלגברה.

לכן, 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 $2k+m=n$.

9.3 $\sim$ מטריצות אוניטריות

הגדרה 82. 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 $T^*=T^{-1}$.

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. 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. $$\forall v\in V\colon\ ||Tv||=||v||$$

כלומר: היא משמרת זווית (העתקה פנימית) וגודל. במילים אחרות, היא משמרת העתקה פנימית.

באופן כללי זה שקול לאיזומטריה ליניארית (כלומר שם כללי לאורתוגונליות/אוניטריות יהיה איזומטריות).

הערה 35. איזומטריה, גם מחוץ ללינארית, היא פונקציה שמשרת גודל.

הערה 36. אפשר להסתכל על מתכונה 4 על איזומטריות לינאריות כעל איזומורפיזם של ממ"פים.

הוכחה. נפרק לרצף גרירות

$1\to 2$
$$T^*=T^{-1}\implies \langle Tv\,|\,Tu\rangle=\langle v\,|\,T^*Tu\rangle=\langle v\,|\,u\rangle$$

$2\to 3$ 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־: $\langle Tv_i\,|\,Tv_j\rangle=\langle v_i\,|\,v_j\rangle=\delta_{ij}$

$3\to 4$ טרוויאלי

$4\to 5$ יהי $(v_1\dots v_n)$ בסיס א"נ כך ש־$(Tv_1\dots Tv_n)$ א"נ. אז:

$$v=\sum_{i=0}^n\alpha_iv_i\implies ||v||^2=\left\langle\sum_{i=0}^n\alpha_iv_i\,\middle|\,\sum_{i=0}^n\alpha_iv_i\right\rangle=\sum_{i=0}^n|\alpha_i|^2$$
$$||Tv||^2=\left\langle T\left(\sum_{i=0}^n\alpha_iv_i\right)\middle|\,T\left(\sum_{i=0}^n\alpha_iv_i\right)\right\rangle=\left\langle\sum_{i=0}^n\alpha_iT(v_i)\,\middle|\,\sum_{i=0}^n\alpha_iT(v_i)\right\rangle=\sum|\alpha_i|^2$$

$5\to 1$ 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 121. תהי $T\colon V\to V$ אוני'אורתו', ו־λ ע"ע של T. אז $|\lambda|=1$

הוכחה. יהי v ו"ע של הע"ע λ. אז:

$$|\lambda|^2\langle v\,|\,v\rangle=\lambda\bar{\lambda}\langle v\,|\,v\rangle=\langle Tv\,|\,Tv\rangle=\langle v\,|\,v\rangle$$

■

הגדרה 83. תהי $A\in M_n(\mathbb{F})$. אז A תיקרא אוניטרית/אורתוגונלית אם $A^*=A^{-1}$.

משפט 122. אוניטרית אמ"מ $A\overline{A^T}=I$.

משפט 123. אורתוגונלית אמ"מ $AA^T=I$.

הערה 37. אוניטרית בה מלשון unit – היא שומרת על הגודל, על וקטורי היחידה (ה־unit vectors).

משפט 124. 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.

$$AA^*=[T]_{\mathcal{B}}[T^*]_{\mathcal{B}}=[TT^*]_{\mathcal{B}},I=AA^*\iff[TT^*]_{\mathcal{B}}=I\iff TT^*=I$$

■

הערה 38. איזומטריה היא העתקה שמשמרת גדלים, ואיזומטריה אורתוגונלית היא פשוט אוניטרית. משום מה זה שם שמדברים עליו בע"פ אבל לא הגדירו מסודר.

"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 13. א"נ = אוניטרית בהקשר של מטריצות (בהקשר של מרחבים – אורתונורמלי)

משפט 125. התאים הבאים שקולים על $A\in M_n(\mathbb{F})$.

1. A א"נ
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. (ביחס למכפלה הפנימית הסטנדרטית) $\forall u,v\in\mathbb{F}^n\colon\ \langle Au\,|\,Av\rangle=\langle u\,|\,v\rangle$
5. (ביחס למכפלה הפנימית הסטנדרטית) $\forall v\in\mathbb{F}^n\colon\ ||Av||=||n||$

הערה שלא קשורה למשפט: נאמר ש־$[T^*]_B=[T]_B^*$ אמ"מ B **בסיס א"נ**. עבור בסיס שאינו א"נ זה לא בהכרח מתקיים.

הערה נוספת: זה בערך אמ"מ כי יש כמה מקרי קצה כמו מטריצת האפס.

הוכחה.

$2\leftrightarrow 1$ 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־$v_1\ldots v_n$ בסיס א"נ (ביחס למ"פ הסטנדטית של $\mathbb{F}^n$)

$3\leftrightarrow 1$ מספיק להוכיח A א"נ אמ"מ A^T א"נ. 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. אז $A=[T_A]_{\mathcal{E}}$. אז T_A א"נ אמ"מ $A=:[T_A]_{\mathcal{E}}$. אז:

$$\langle Au\,|\,Av\rangle=\langle T_Au\,|\,T_Av\rangle=\langle u\,|\,v\rangle$$

$5\leftrightarrow 1$ אותה הדרך כמו קודם.

■

9.3.1 צורה קאנונית למטריצה אוניטרית

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:

$$A_1:=A=\begin{pmatrix}\cos\theta&\sin\theta\\\sin\theta&-\cos\theta\end{pmatrix}\vee A_2:=A=\begin{pmatrix}\cos\theta&\sin\theta\\-\sin\theta&\cos\theta\end{pmatrix}$$

נבחין ש־A_2 הוא סיבוב ב־θ, ו־A_1 שיקוף ניצב ביחס ל־$\frac{\theta}{2}$. זה לא מפתיע שכן $\det A_1=-1,\ \det A_2=1$. ■

"דרך נוספת לראות את זה":

$$a=\cos\theta\implies b=\sin\theta,\ c=\sin\varphi\implies d=\cos\varphi$$

אז (עד לכדי סיבוב)

$$\cos\theta\sin\varphi+\sin\theta\cos\varphi\implies\sin(\theta+\varphi)=0\implies\theta+\varphi=0\vee\theta+\varphi=\pi$$

במקרה הראשון ש־$\theta=\varphi$ קיבלנו סיבוב, ובמקרה השני נקבל ש־$\varphi=\pi-\theta$ ואז $\sin(\pi-\theta)=\sin\theta$ כדרוש.

ננסה להבין יותר טוב למה הן מסובבות בצורה הזו. A_2 מטריצה מוכרת אך A_1 פחות. נתבונן בפולינום האופייני שלה:

$$f_{A_1}(x)=\begin{vmatrix}x-\cos\theta&-\sin\theta\\-\sin\theta&x+\cos\theta\end{vmatrix}=x^2-\cos^2\theta-\sin^2\theta=(x+1)(x-1)$$

אזי הע"ע $-1,+1$ (שימו לב ש־A_2 לא לכסינה מעל $\mathbb{R}$).

$$\begin{aligned}A\begin{pmatrix}\cos\frac{\theta}{2}\\\sin\frac{\theta}{2}\end{pmatrix}&=\begin{pmatrix}\cos\theta&\sin\theta\\\sin\theta&-\cos\theta\end{pmatrix}\begin{pmatrix}\cos\frac{\theta}{2}\\\sin\frac{\theta}{2}\end{pmatrix}=\begin{pmatrix}\cos\theta\cos\frac{\theta}{2}+\sin\theta\sin\frac{\theta}{2}\\\sin\theta\cos\frac{\theta}{2}-\cos\theta\sin\frac{\theta}{2}\end{pmatrix}=\begin{pmatrix}\cos\left(\theta-\frac{\theta}{2}\right)\\\sin\left(\theta-\frac{\theta}{2}\right)\end{pmatrix}\\&=\begin{pmatrix}\cos\theta/2\\\sin\theta/2\end{pmatrix}\quad\left(\frac{\theta}{2}\mapsto\frac{\theta}{2}+\pi\right)=\begin{pmatrix}\cos\left(\frac{\theta}{2}-\frac{\pi}{2}\right)\\\sin\left(\frac{\theta}{2}-\frac{\pi}{2}\right)\end{pmatrix}=\begin{pmatrix}-\cos\left(\frac{\theta}{2}+\frac{\pi}{2}\right)\\-\sin\left(\frac{\theta}{2}+\frac{\pi}{2}\right)\end{pmatrix}\end{aligned}$$

[אני ממש חלש בטריגו ואני מקווה שאני לא מסכם דברים לא נכונים. תבדקו אותי פעמיים כאן בחלק הזה. גם המרצה עשה את ההחלפה המוזרה של $\theta/2\to\theta/2+\pi/2$].

"אם הייתם רוצים תקופות מבחנים נורמליות הייתם צריכים להיוולד בזמן אחר".

"ומה, אתם חושבים שאחרי שהפקולטה דחתה בשבוע היא תיאמה את זה עם הפקולטות האחרות? הם דיברו איתם כמה ימים אח"כ"

מסקנה. (הצורה הנורמלית של ט"ל אורתוגונלית) תהי $T\colon V\to V$ אורתוגונלית. אז קיים בסיס א"נ של V, שביחס אליו המטריצה המייצגת את T היא מהצורה:

$$\begin{pmatrix}A_{\theta_1}&&&&&&&&\\&\ddots&&&&&&&\\&&A_{\theta_n}&&&&&&\\&&&1&&&&&\\&&&&\ddots&&&&\\&&&&&1&&&\\&&&&&&-1&&\\&&&&&&&\ddots&\\&&&&&&&&-1\end{pmatrix}$$

כאשר:

$$A_{\theta_i}=\begin{pmatrix}\cos\theta_i & \sin\theta_i\\ -\sin\theta_i & \cos\theta_i\end{pmatrix}$$

(אוניטרית לא מעיינת כי היא לכסינה)

הגיון:

אורתוגונלית, לכן נורמלית, לכן נראית בצורה של בלוקים 2×2 של ע"ע. הע"ע מגודל 1 כי היא אורתוגונלית, והם חייבים להיות ממשיים על מעגל היחידה הממשי. המטריצה A_θ חייבת להיות אורתוגונלית מגודל 2×2 כי כל תמ"ו שם הוא T-אינוו', כלומר אפשר לחלק את T לבלוקים מתאימים ובפרט T המצומצמת גם אורתוגונלית, ולכן A_θ סה"כ אורתוגונלית. כאשר $A_{\theta_i}=\begin{pmatrix}a & b\\ -b & a\end{pmatrix}$ ו־$b\neq 0$ נשארנו עם המטריצות הללו.

הוכחה. ידוע עבור נורמלית:

$$A=\begin{pmatrix}\square_1 &&&&&\\ &\ddots&&&&\\ &&\square_m&&&\\ &&&\lambda_1&&\\ &&&&\ddots&\\ &&&&&\lambda_k\end{pmatrix}$$

כאשר

$$\square_i=\begin{pmatrix}a_i & b_i\\ -b_i & a_i\end{pmatrix}$$

ובמקרה 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-אינו' היא עדיין אורתוגונלית, והיא בהכרח מהצורה של מטריצת הסיבוב לעיל. המטריצה של שיקוף וסיבוב ב־$\frac{\theta}{2}$ לכסינה ולכן להפוך 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 126. 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 $\mathbb{C}$ לכסין).

הוכחה. ידוע שבעבור $\lambda_1\dots\lambda_k$ ע"עים:

$$f_T(x)=\left(\prod(x-\lambda_i)\right)\left(\prod(x^2-2a_ix+a_i^2+b_i^2)\right)$$

כאשר המכפלה הראשונה באה מהע"עים והשניה מהריבועים $\square_i$. נבחין שלכל תמ"ו a_i נקבבע 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־$A_{-\theta_i}$ (תפתחו את האלגברה/טריגו, זה מה שזה אומר)? זאת כי הן דומות באמצעות ההעתקה שהופכת את הצירים, מה ששקול להחליף את עמודות A_{θ_i}.

9.3.2 המשפט הספקטרלי בניסוח מטריציוני

משפט 127. (המשפט הספקטרלי "בשפה קצת מטרציונית") תהי $A\in M_n(\mathbb{F})$ מטריצה סימטרית (מעל $\mathbb{R}$)/נורמלית (מעל $\mathbb{C}$). 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 על בסיסים ועל וקטורים – אפשר לתאר עולם הדיון של המטריצות, משום שהוא עולם דיון הומורפי להעתקות ולמרחבים וקטורים, בלי לדבר בכלל על העתקות ומרחבים וקטורים. המשפט מתאר באופן טהור מטריצות בלבד.

למה 11. 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 $T_A \colon \mathbb{F}^n \to \mathbb{F}^n$ באופן הרגיל. 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 B כך ש־$[T_A]_B = D$ כאשר D אלכסונית כלשהית. נבחין ש־$[T_A]_B = [Id]_B^{\mathcal{E}}[T_A]_{\mathcal{E}}[Id]_{\mathcal{E}}^B$, נסמן $P = [Id]_B^{\mathcal{E}}$ ונבחין ש־$[T_A]_B = PAP^{-1}$ ומהלמה P מטריצת מעבר מבסיס א"נ לבסיס א"נ ולכן איזומטריה. נכפיל בהופכיות ונקבל $A = P^{-1}DP$. ■

באמצעות כלים של אנליזה פונקציונלית אפשר להגדיר נורמה גם על פונקציות, ואיכשהו להגדיר את העובדה שההעתקה שמעבירה בסיס (בעולם הדיון של ההעתקות) היא אוניטרית/אורתוגונלית.

"אני יודע איך מגדירים נורמה של טרנספומציה. יופי של שאלות – לא לעכשיו"

"יאללה הפסקה? לא!"

9.4 ~ פירוק פולארי

9.4.1 מבוא, וקישור לתבניות בי־לינאריות

הערה 39. במקרה של $\mathbb{F} = \mathbb{R}$ נקבל ש־

$$A = P^{-1}DP \implies PP^T = I \implies P^{-1} = P^T \implies A = P^TDP$$

מה שמחזיר אותנו לתבניות בילינאריות. 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־בילינארית ולא בילינארית רגילה.

משפט 128. 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.

את הכיוון $\Longleftarrow$ כבר הוכחנו. 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

$$\forall v \in V \colon \ \langle ABv \mid ABv \rangle = \langle Bv \mid Bv \rangle = \langle v \mid v \rangle$$

  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 $\forall v \neq 0 \colon \ \langle Tv \mid Tv \rangle \geq / > 0$.

משפט 129. נניח ש־$A = A^* \in M_n(\mathbb{F})$, אז התנאים הבאים שקולים (קיצור מוכר: TFAE, the following are equaivlent):

1. T_A חיובית/אי שלילית על $\mathbb{F}^n$.
2. לכל $T \colon V \to V$ ובסיס א"נ B כך ש־$A = [T]_B$, T חיובית/אי שלילית.
3. קיימים $T \colon V \to V$ חיובית/אי שלילית ו־B בסיס, כך ש־$A = [T]_B$.
4. הע"ע של A (יודעים ממשיים כי צמודה לעצמה) חיובים/אי שליליים.

הוכחה. מספיק לטעון זאת כדי להוכיח את השקילויות של 1, 2, 3:

$$\langle Tv \mid v \rangle_V = \langle [Tv]_B \mid [v]_B \rangle_{\mathbb{F}^n} = \langle A[v]_B \mid [v]_B \rangle_{\mathbb{F}^n}$$

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 $1 \leftrightarrow 2 \leftrightarrow 3$.

עתה נוכיח שקילות בין 1 ל־4.

$4 \to 1$ יהי $\lambda \in \mathbb{R}$ ע"ע של A (נוכל להניח ממשי כי A צמודה לעצמה)

$$\langle Av \,|\, v\rangle = \lambda \,||v||^2 > 0 \implies \lambda > 0$$

$1 \to 4$ יהי $B = (v_1 \dots v_n)$ בסיס א"נ של ו"ע, ויהי $V \ni v = \sum \alpha_i v_i$. נקבל:

$$\langle T_A v \,|\, v\rangle = \langle Av \,|\, v\rangle = \left\langle A\sum_{i=1}^{n} \alpha_i v_i \,\middle|\, \sum_{i=1}^{n} \alpha_i v_i \right\rangle = \sum \lambda_i |\alpha_i|^2 > 0$$

■

תזכורת: מעל $\mathbb{R}$, הוכחנו שלכל תבנית סימטרית, יש ייצוג יחיד באמצעות מטריצה אלכסונית עם $-1, 1, 0$ על האלכסון. **סימון.** 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 130. נניח ש־A מייצגת את התבנית הסימטרית f (עולם הדיון מעל $\mathbb{R}$). אז, אם הסיגנטורה $\sigma_+ = \#(\lambda \mid \lambda > 0)$ עבור λ ע"ע עם חזרות. באופן דומה $\sigma_- = \#(\lambda \mid \lambda > 0)$ וכו'.

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 וחופפת אליה. בעזרת נרמול היא חופפת למטריצה מהצורה $\operatorname{diag}(1 \dots 1, -1 \dots -1, 0 \dots 0)$ כאשר הסימן נקבע לפני הנרמול. ■

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 כמו שראינו בהרצאה 22.

משפט 131. תהי $T\colon V \to V$ צמודה לעצמה ואי שלילית $\langle Tv \,|\, v\rangle \geq 0$, 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.

$$[T]_B^B = \operatorname{diag}(\lambda_1 \dots \lambda_n) \quad [R]_B^B = \operatorname{diag}(\sqrt{\lambda}_1 \dots \sqrt{\lambda}_n)$$

(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־$B = (e_1 \dots e_n)$ 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 14. את ה־R לעיל נסמן $R := \sqrt{T}$.

מסקנה 14 (פירוק שולסקי)**.** לכל A צמודה לעצמה ואי־שלילית חיובית קיים פירוק של מטריצה R משולשית עליונה כך ש־$A = RR^*$. בעזרת פירוק RU שנראה בהמשך, נוכל להראות ש־R משולשית עליונה.

9.4.2 ניסוח פירוק פולארי בעבור העתקות

משפט 132. (פירוק הפולארי) תהי $T\colon V \to V$ הפיכה, אז קיימות $R \to V \to V$ חיובית וצמודה לעצמה ו־$U\colon V \to V$ אוניטרית כך ש־$T = RU$.

הערה 40. לא הנחנו T צמודה לעצמה. הפירוק נכון להעתקה הפיכה כללית.

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 לגבי יחידות. אם T אינה הפיכה, מקבלי סש־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):

הוכחה.

$$TT^* = RU(RU)^* = RUU^*R^* = R^2$$

כלומר R היא בכל פירוק שורש, והראינו קודם את יחידות השורש. ■

הערה 41. 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 $\tilde{U}^{-1} =: U$, $\tilde{R} =: R$ וסה"כ $T = UR$ כדרוש. ■

למה 12. 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 T^*T, TT^* הן העתקות דומות ולכן יש להן את אותם הערכים העצמיים. ■

הערה 42. אז איך זה קשור לפולארי? R האי־שלילית היא "הגודל", בעוד U האוניטרית לא משנה גודל – היא ה"זווית".

9.4.3 ניסוח פירוק פולארי בעבור מטריצות

משפט 133. (פירוק פולארי עבור 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. ■

9.5 ∼ פירוק SVD

הערה 43. SVD הינו קיצור של Singular Value Decomposition.

משפט 134. (פירוק לערכים סינגולריים למטריצה – SVD) 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 44.

$$AA^* = (UDV)V^*D^*U^* = UD^2U^{-1}$$
$$A^*A = V^{-1}D^2V$$

הגדרה 84. הערכים העצמיים האי־שליליים של A^*A נקראים הערכים הסינגולריים והם נקבעים ביחידות ע"י A.

הערכים הסינגולרים הם גם הע"ע של R^2 הפירוק בפולארי וכן הע"ע של D^2 בפירוק SVD.

הערה 45. פירוק SVD יחיד למטריצה הפיכה.

הערה 46. 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.

הערה 47. כאן בערך נגמר החומר בקורס. עם זאת, חסר מאוד הסברים על השימושים של פירוק SVD ועל מסקנות די חשובות ממנו. לכן הוספתי קצת הרחבות עליו, תוך הסתמכות על הספר "Linear Algebra Done Right".

המשך בעמוד הבא

10 Common Algorithms

בניגוד לפרק הבא, הפרק הזה לא מעניין בכלל. הוא מסכם בקצרה אלגוריתמים מועילים שרואים בתרגולים וכדאי לזכור (**אין** כאן סיכום מלא של התרגולים).

10.1 ~ אלגוריתמים מרכזיים

10.1.1 לכסון

שלבים: 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 $\text{diag}(\lambda_1 \dots \lambda_n)$ המטריצה האלכסונית המתקבלת ע"י מטריצת מעבר הבסיס הנתונה ע"י הו"עים מהשלב הקודם.

10.1.2 ג'ירדון

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 $\dim\left(V_{\lambda_j}^{(\ell_j)}\right) = r_i$ (המרחב העצמי המוכלל).
- נחזור על האלגו' למציאת צורת ג'ורדן למטריצה ניל':
  - 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 $\left\{(A-\lambda_j I)^k v \mid 0 \le k < i,\ v \in C_{\lambda_j}^{(i)+}\right\}$.
  - נגדיר $B = \bigcup_{j=1}^m B_{\lambda_j}$ הבסיס המג'רדן.

מציאת $J_n(\lambda)^m$**:** 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:

$$J_n(\lambda)^m = \begin{pmatrix} \binom{m}{0}\lambda^m \\ \binom{m}{1}\lambda^{m-1} & \binom{m}{0}\lambda^m \\ \vdots & \ddots & \ddots \\ \binom{m}{m}\lambda^0 & \binom{m}{m-1}\lambda^1 & \ddots & \ddots \\ 0 & \binom{m}{m}\lambda^0 & \cdots & \ddots & \binom{m}{0}\lambda^m \\ \vdots & 0 & \binom{m}{m}\lambda^0 & \cdots & \binom{m}{m-1}\lambda^1 & \binom{m}{0}\lambda^m \\ 0 & \vdots & \ddots & \ddots \end{pmatrix}$$

10.1.3 אלגוריתם 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 $i \in [n]$

$$\tilde{v}_i = v_i - \sum_{j=1}^{i-1} \frac{\langle v_i \mid \tilde{v}_i \rangle}{\langle \tilde{v}_i \mid \tilde{v}_j \rangle} \cdot \tilde{v}_j$$

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 $i \in [n]$:

$$\bar{v}_i = \frac{1}{||v_i||} \left(v_i - \sum_{j=1}^{i-1} \langle v_i \mid \bar{v}_i \rangle \cdot \bar{v}_i \right)$$

בצורה זו נוכל לנרמל תוך כדי התהליך.

10.2 ~ מספר אלגוריתמים נוספים

- **אלגוריתם אוקלידס לתחום ראשי** (בפרט בעבור פולינומים):
- **נרמול וקטור:** נגדיר $v = \frac{v}{||v||}$ הוא v מנורמל.
- **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.
- **חישוב A^{-1}, A^{n+c} באמצעות משפט קיילי־המילטון:** ידוע $f_A(A) = 0$, ואם נשאר גורם חופשי $\alpha_n A^n + \cdot \alpha_0 A^0 = 0$ 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 $A^{n-1} \cdots I$, שעבור מספרי n קטנים קל לחשב.
- **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).
- **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).
- **מציאת ללכסון אוניטרי/רוגתורונלי** (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.

. . . .

סוף הקורס ~ 2025B

הסיכום לא נגמר – יש הרחבה על דואלים בעמוד הבא

קומפל ב־LaTeX ונוצר באמצעות תוכנה חופשית בלבד

11 Dual Spaces

11.1 ~ הגדרות בסיסיות

הגדרה 85. 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 13. 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 135. 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 $0(v_j) = 0 = \left(\sum_{i=0}^n \alpha_i \psi_i\right) v_j = \sum_{i=0}^n \alpha_i \psi_i(v_j) = \sum \alpha_i \delta_{ij} = \alpha_j$ וסה"כ $\alpha_j = 0$. ■

נבחין שאפשר להגדיר:

הגדרה 86. $V^{**} = \hom(V^*, \mathbb{F})$

ואכן $\dim V < \infty$ אז:

$$V \cong V^* \cong V^{**}$$

במקרה הזה, בניגוד לאיזו' הקודם, יש איזו' "טבעי" (קאנוני), כלומר לא תלוי באף בסיס.

משפט 136. 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 $\alpha, \beta \in \mathbb{F},\ v, u \in V$. 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 $1 = \varphi_1(v)$ 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.

11.2 ~ איזומורפיות למרחבי מכפלה פנימית

11.2.1 העתקה צמודה (דואלית)

סימון 15. לכל $v \in V$ ו־$\varphi \in V^*$ נסמן $\varphi(v) = (\varphi, v)$

הערה 48. סימון זה הגיוני משום שהכנסת וקטור לפונרציונל דואלי איזומורפי למכפלה פנימית.

משפט 137. 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 ויזואלית 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 $T^* : W^* \to V^*$. בעצם, זהו איזומורפיזם ("בשפת הפנקטורים") קאנוני. עוד קודם לכן ידענו (בגלל ממדים) שהם איזומורפים, אך לא מצאנו את האיזומורפיזם ולא ראינו שהוא קאנוני.

$$\tau : \ \hom(V, W) \to \hom(W^*, V^*) \quad \tau(T) = T^*$$

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 לינארי. ■

הוכחה "מתוחכמת". "אני אהבתי את ההוכחה שלי": 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."

11.2.2 המאפס הדואלי ומרחב אורתוגונלי

הגדרה 87. יהי V מ"ו נוצר סופית. יהי $S \subseteq V$ קבוצה. נגדיר $S^0 \subseteq V^*$ $S^0 =: \{\varphi \in V^* \mid \forall v \in S\colon \varphi(v) = 0\}$.

דוגמאות. $\{0\}^0 = V^*,\ V^0 = \{0\}$

משפט 138.

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 139. 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:

$$\forall \varphi \in U^0\, \forall u \in U\colon \varphi(u) = 0$$

ומי אלו הוקטורים שיאפסו את φ שמאפס את u? הוקטורים ב־U עד לכדי האיזומורפיזם הקאנוני מ־U ל־U^{**}.

נבחין ש־:

$$[T^*]_{\mathcal{A}^*}^{\mathcal{B}^*} = ([T]_{\mathcal{B}}^{\mathcal{A}})^T$$

כאשר $\mathcal{A}$ בסיס ל־V, $\mathcal{A}^*$ ל־V^*, $\mathcal{B}$ ל־W, $\mathcal{B}^*$ ל־W^*.

"כוס אמא של קושי" – בן על זה שקושי גילה משפט כלשהו לפניו.

שחר פרץ, 2025

קומפל ב־LaTeX ונוצר באמצעות תוכנה חופשית בלבד