אלגברה ליניארית ב' - הרצאות - אביב 2018

מרצה: עדי וולף

20 ביוני 2018

הרצאה 1 - 21 למרץ

מוטיבציה

באלגברה א' עצרנו בשאלה מתי 2 מטריצות הן דומות. ראינו שאם המטריצות לכסינות שתיהן אז הן דומות $\iff$ יש להן את אותה צורה אלכסונית. בקורס הזה נענה על השאלה באופן כללי, גם כשהמטריצות לא לכסינות. נראה שלכל מטריצה יש צורה כמעט אלכסונית, שנקראת צורת ג'ורדן, והיא נראית משהו כזה:

$$\begin{pmatrix} \lambda_1 & 1 & & & \\ & \lambda_2 & 1 & & \\ & & \lambda_3 & 1 & \\ & & & \ddots & 1 \\ & & & & \lambda_n \end{pmatrix}$$

אנו נראה בהמשך ששתי מטריצות הן דומות אם ורק אם יש להן את אותה צורת ג'ורדן. זו בעצם דוגמה למציאת תכונה שמורה (צורת ג'ורדן) של אובייקטים (מטריצות), ואז אנו אומרים ששני אובייקטים הם שקולים אם יש להם את אותה תכונה שמורה.

בנוסף נלמד יותר על הדטרמיננטה, נבין אותה כמעין פונקציית נפח או שטח, ומכאן גם נבין מדוע אופן הפיתוח שלה לא משפיע על הערך הסופי שלה.

נדבר גם על פולינומים שמתאפסים ממטריצה, ובפרט נראה את משפט קיילי-המילטון, על פיו אם $p(x) = \det(x\mathbb{I} - A)$ הינו הפולינום האופייני של $A_{n\times n}$ אזי $P(A) = 0_{n\times n}$. נשים לב שאסור סתם להציב $x = A$ בביטוי הנ"ל, כי x הוא בעצם סקלר.

אנחנו כבר יודעים ש-A מאפסת פולינום ממעלה שהיא לכל היותר n^2 כי בקבוצה $\left\{A^0 = \mathbb{I}, A^1, A^2, ..., A^{n^2}\right\}$ יש $n^2 + 1$ איברים ולכן היא תלויה (כי מימד מרחב המטריצות מגודל $n \times n$ הוא n^2), ולכן קיימת קומבינציה ליניארית של איברים אלו שמתאפסת, והפולינום המתאים לקומבינציה זו הוא זה שמתאפס בהצבת A. החיסרון של טיעון זה הוא בכך שהוא אינו אומר לנו מהו הפולינום שמתאפס. הכוח של משפט קיילי המילטון הוא בכך שהוא נותן לנו פולינום שמתאפס בצורה קונקרטית, וגם בכך שפולינום זה הוא ממעלה n.

בחלק השני של הקורס נעסוק במרחבי מכפלה פנימית, בהם למעשה נכניס גם מבנה גיאומטרי למרחב וקטורי - ונקבל מושגים כמו אורך, מרחק, זווית. נראה שתמיד יש בסיס של וקטורים ניצבים ונלמד תהליך להגיע אליו (תהליך גרהם-שמידט). בנוסף נשאל מה עוד ניתן להכניס למרחב וקטורי - וזה יהיה תבניות ביליניאריות ומכפלות טנזוריות.

השלמות מאלגברה א' - המטריצה הצמודה (Adjoint), כלל קרמר, משפט קיילי-המילטון

יהי $\mathbb{F}$ שדה (למשל, הממשיים), ו-$A \in \mathcal{M}_{n\times n}(\mathbb{F})$ היא מטריצה $n \times n$ שהכניסות שלה היא מעל איברי השדה $\mathbb{F}$. נראה כעת דרך נוספת לחשוב על הדטרמיננטה, שתשרת אותנו בהמשך כשנדון במשפט קיילי-המילטון.

נגדיר מטריצה $n \times n$ שתסומן $\operatorname{adj}(A)$ ותיקרא המטריצה הצמודה (הקלאסית) של A (לפעמים קוראים לה "המצורפת", אנחנו פשוט נקרא לה Adjoint), והיא מוגדרת באופן הבא:

$$[\operatorname{adj}(A)]_{ij} := (-1)^{i+j} M_{ji}$$

כאשר M_{ji} הוא המינור ה-j,i שהוא ה-$\det$ של המטריצה $(n-1) \times (n-1)$ המתקבלת מ-A לאחר מחיקת שורה j ועמודה i.

דוגמה: תהי מטריצה:

$$A=\begin{pmatrix}1 & 0 & 2\\ 2 & 1 & 2\\ 3 & 1 & 2\end{pmatrix}$$

נרצה לחשב את $\mathrm{adj}\left(A\right)$. נקבל:

$$\mathrm{adj}\left(A\right)=\begin{pmatrix}M_{11} & -M_{12} & M_{13}\\ -M_{21} & M_{22} & -M_{23}\\ M_{31} & -M_{32} & M_{33}\end{pmatrix}^{t}=\begin{pmatrix}\det\begin{pmatrix}1 & 2\\ 1 & 2\end{pmatrix} & -\det\begin{pmatrix}2 & 2\\ 3 & 2\end{pmatrix} & \det\begin{pmatrix}2 & 1\\ 3 & 1\end{pmatrix}\\ -\det\begin{pmatrix}0 & 2\\ 1 & 2\end{pmatrix} & \det\begin{pmatrix}1 & 2\\ 3 & 2\end{pmatrix} & -\det\begin{pmatrix}1 & 0\\ 3 & 1\end{pmatrix}\\ \det\begin{pmatrix}0 & 2\\ 1 & 2\end{pmatrix} & -\det\begin{pmatrix}1 & 2\\ 2 & 2\end{pmatrix} & \det\begin{pmatrix}1 & 0\\ 2 & 1\end{pmatrix}\end{pmatrix}^{t}=$$

$$=\begin{pmatrix}0 & 2 & -1\\ 2 & -4 & -1\\ -2 & 2 & 1\end{pmatrix}^{t}=\begin{pmatrix}0 & 2 & -2\\ 2 & -4 & 2\\ -1 & -1 & 1\end{pmatrix}$$

טענה:

$$(*)\quad A\cdot\mathrm{adj}\left(A\right)=\mathrm{adj}\left(A\right)\cdot A=\det\left(A\right)\cdot\mathbb{I}$$

מסקנה מהטענה: אם A הפיכה אז ההופכית שלה היא $A^{-1}=\frac{\mathrm{adj}(A)}{\det(A)}$ (ולא חילקנו באפס, כי היא הפיכה ולכן הדטרמיננטה שלה שונה מאפס). אחת החשיבויות של נוסחה זו היא שרואים ממנה שהפונקציה ההופכית היא רציפה (כמנה של פולינומים). בנוסף זוהי נוסחה מפורשת לפונקציה ההפוכה (בניגוד לשיטת גאוס שלמדנו באלגברה א', שהייתה אלגוריתמית מטבעה), ומכאן חשיבותה הרבה.

הערה: ראינו בעבר שבמקרה הפרטי של $A_{2\times2}=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$ המטריצה ההופכית היא $A^{-1}=\frac{1}{\det(A)}\begin{pmatrix}d & -b\\ -c & a\end{pmatrix}$. אנו רואים שזהו אכן המקרה הפרטי של המסקנה הנ"ל.

דוגמה/תרגיל: תהי $A\in M_{n}\left(\mathbb{Z}\right)$ מטריצה $n\times n$ שכל כניסותיה שלמות. הראו ש-A הפיכה עם $A^{-1}\in M_{n}\left(\mathbb{Z}\right)$ אמ"ם $\det\left(A\right)=\pm1$. נפתור דוגמה זו בתרגול.

הוכחת הטענה: נראה ש-$A\cdot\mathrm{adj}\left(A\right)=\left|A\right|\cdot\mathbb{I}$, השוויון השני מוכח בדיוק באותו האופן. אנחנו צריכים להראות כי:

$$\left[A\cdot\mathrm{adj}\left(A\right)\right]_{ij}=\begin{cases}\left|A\right| & i=j\\ 0 & i\neq j\end{cases}$$

נסמן $A_{ij}=\left(a_{ij}\right)$, אז על פי הגדרת כפל מטריצות נקבל שהביטוי $\left[A\cdot\mathrm{adj}\left(A\right)\right]_{ij}$ הוא למעשה כפל שורה i של A בעמודה j של $\mathrm{adj}\left(A\right)$, לכן נקבל:

$$\left[A\cdot\mathrm{adj}\left(A\right)\right]_{ij}=\sum_{k=1}^{n}a_{ik}\left(\mathrm{adj}\left(A\right)\right)_{kj}=\sum_{k=1}^{n}a_{ik}\left(-1\right)^{k+j}M_{jk}^{A}$$

כעת נפריד למקרים:

אם $i=j$, קיבלנו:

$$\left[A\cdot\mathrm{adj}\left(A\right)\right]_{ii}=\sum_{k=1}^{n}\left(-1\right)^{k+i}a_{ik}M_{ik}^{A}=\left|A\right|$$

כאשר במעבר האחרון הסתמכנו על כך שהביטוי הוא בדיוק הדטרמיננטה של A כשאנחנו מפתחים אותה לפי השורה ה-i!

אם $j \neq i$ נבצע "תעלול", ניקח מטריצה B שזהה ל־A פרט לכך ששורה j של B היא השורה i של A. כלומר:

$$b_{m\ell} = \begin{cases} a_{m\ell} & m \neq j \\ a_{i\ell} & m = j \end{cases}$$

מאחר ולמטריצה B שתי שורות זהות, הרי שהדטרמיננטה שלה מתאפסת, נפתח אותה לפי שורה j ונקבל:

$$0 = \det(B) = \sum_{k=1}^{n} (-1)^{j+k} b_{jk} M_{jk}^{B} = \sum_{k=1}^{n} (-1)^{j+k} a_{ik} M_{jk}^{A} = [A \cdot \operatorname{adj}(A)]_{ij}$$

למעשה הסתמכנו על כך שאיברי המטריצה B בשורה j זהים לאיברי A מהשורה ה־i, ולכן $b_{jk} = a_{ik}$, וכמו כן על כך שהמינורים שווים (כי פיתחנו לפי שורה j, ולכן מחקנו את השורה ה־j במטריצה B, ושורה זו הייתה ההבדל היחיד בין המטריצות A ו־B!). ■

<u>כלל קרמר</u>

תהי $A \in \mathcal{M}_n(\mathbb{F})$ <u>הפיכה</u>, יהי $\vec{b} \in \mathbb{F}^n$ ונביט על המערכת של n משוואות עם נעלמים:

$$A\vec{x} = \vec{b}$$

כאשר $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. כלל קרמר נותן לנו נוסחה סגורה למציאת הפתרון (בניגוד לאלגוריתם גאוס שאינו נוסחה).

כלל קרמר: למערכת הנ"ל פתרון יחיד (ואת זה כבר ראינו באלגברה א') שניתן על ידי:

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

כאשר C_i היא המטריצה $n \times n$ המתקבלת מ־A לאחר החלפת עמודה i בוקטור $\vec{b}$.

<u>דוגמה</u>: נביט במערכת:

$$\begin{cases} 2x_1 - x_2 = 1 \\ 3x_1 + 5x_2 = 7 \end{cases}$$

לכן נקבל:

$$A = \begin{pmatrix} 2 & -1 \\ 3 & 5 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 1 \\ 7 \end{pmatrix}$$

ומכאן:

$$C_1 = \begin{pmatrix} 1 & -1 \\ 7 & 5 \end{pmatrix}, \quad C_2 = \begin{pmatrix} 2 & 1 \\ 3 & 7 \end{pmatrix}$$

ולכן:

$$\det(C_1) = 12, \quad \det(C_2) = 11, \quad \det(A) = 13$$

אזי לפי כלל קרמר הפתרון יהיה:

$$x_1 = \frac{\det(C_1)}{\det(A)} = \frac{12}{13}$$
$$x_2 = \frac{\det(C_2)}{\det(A)} = \frac{11}{13}$$

וקל לראות שזהו אכן הפתרון.

נעיר שכלל קרמר מאוד רגיש לשגיאות בכניסות המטריצה, מספיק שהמקדמים יהיו שונים במעט על מנת שנקבל פתרון שונה משמעותית מהפתרון המקורי, לכן הוא אינו שימושי במיוחד למציאת פתרונות נומריים (והוא גם דורש הרבה חישובים במילא, שזה לא יתרון בחישוב נומרי).

<u>הוכחת כלל קרמר:</u>

<u>דרך א'</u> - חשבון ישיר בעזרת הטענה $(*)$ שראינו מקודם, נסו זאת בתור תרגיל!

<u>דרך ב'</u> - תהי A הפיכה, נגדיר העתקה $T:\mathbb{F}^n\to\mathbb{F}^n$ באופן הבא:

$$T\left(x_1,...,x_n\right)=\left(\frac{\det\left(B_1\right)}{\det\left(A\right)},\frac{\det\left(B_2\right)}{\det\left(A\right)},...,\frac{\det\left(B_n\right)}{\det\left(A\right)}\right)$$

כאשר B_i היא המטריצה המתקבלת מ-A לאחר החלפת עמודה i בוקטור $\vec{x}=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$. T היא פונקציה, כעת נראה שהיא ליניארית וכמו כן ש-T פועלת כמו A^{-1} על הוקטורים ב-$\mathbb{F}^n$. כלומר עלינו להראות כי $T\left(v\right)=A^{-1}v$ לכל $v\in\mathbb{F}^n$. בכך נסיים, כי נקבל שבפרט עבור הוקטור $\vec{b}$ מתקיים $T\left(\vec{b}\right)=A^{-1}\vec{b}=\vec{x}$ (כזכור נתונה המערכת $A\vec{x}=\vec{b}$ ולכן $\vec{x}=A^{-1}\vec{b}$).

<u>הוכחת ליניאריות של T:</u> צריך להראות ש-$T\left(\vec{x}+\vec{y}\right)=T\left(\vec{x}\right)+T\left(\vec{y}\right)$ וכמו כן ש-$T\left(\alpha\vec{x}\right)=\alpha T\left(\vec{x}\right)$. זה נובע מהתכונות הבאות של הדטרמיננטה שראינו באלגברה א' שאם יש לנו מטריצה אזי:

$$\det\begin{pmatrix}\vdots & \vdots & & \vdots & & \vdots\\ \vec{v}_1 & \vec{v}_2 & \dots & \alpha\vec{v}_i & \dots & \vec{v}_n\\ \vdots & \vdots & & \vdots & & \vdots\end{pmatrix}=\alpha\det\begin{pmatrix}\vdots & \vdots & & \vdots & & \vdots\\ \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_i & \dots & \vec{v}_n\\ \vdots & \vdots & & \vdots & & \vdots\end{pmatrix}$$

$$\det\begin{pmatrix}\vdots & \vdots & & \vdots & & \vdots\\ \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}+\vec{u} & \dots & \vec{v}_n\\ \vdots & \vdots & & \vdots & & \vdots\end{pmatrix}=\det\begin{pmatrix}\vdots & \vdots & & \vdots & & \vdots\\ \vec{v}_1 & \vec{v}_2 & \dots & \vec{v} & \dots & \vec{v}_n\\ \vdots & \vdots & & \vdots & & \vdots\end{pmatrix}+\det\begin{pmatrix}\vdots & \vdots & & \vdots & & \vdots\\ \vec{v}_1 & \vec{v}_2 & \dots & \vec{u} & \dots & \vec{v}_n\\ \vdots & \vdots & & \vdots & & \vdots\end{pmatrix}$$

כלומר בגלל שהדטרמיננטה היא העתקה ליניארית, גם T היא העתקה ליניארית.

נותר להראות ש-$T\left(\vec{v}\right)=A^{-1}\vec{v}$ לכל $v\in\mathbb{F}^n$. לשם כך מספיק להראות זאת עבור בסיס של $\mathbb{F}^n$. נתון ש-A הפיכה ולכן עמודותיה הן בסיס של $\mathbb{F}^n$ (הן בלתי תלויות ויש n מהן), שנסמנו $\left\{\varphi_1,...,\varphi_n\right\}$. אזי:

$$A^{-1}\cdot\vec{\varphi}_i=\begin{pmatrix}0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix}$$

כאשר הערך 1 הוא בקואורדינטה ה-i של הוקטור $A^{-1}\cdot\vec{\varphi}_i$.

מצד שני, מתקיים

$$\left[T\left(\vec{\varphi}_i\right)\right]_j=\begin{cases}\frac{\det(B_j)}{\det(A)}=0 & i\neq j\\ \frac{\det(A)}{\det(A)}=1 & i=j\end{cases}$$

כאשר הסיבה ש-$\frac{\det(B_j)}{\det(A)}$ מתאפס היא שעמודה i מופיעה פעמיים - גם בעמודה i וגם במקום עמודה j. בכך סיימנו את ההוכחה. ■

הרצאה 2 - 26 למרץ

מטרת העל: מציאת צורת ג'ורדן, שזו צורה כמעט אלכסונית של מטריצה.

התחלנו להשלים כלים שנזדקק להם מאלגברה א': Adjoint, כלל קרמר, משפט קיילי המילטון.

בהמשך נדון בסכומים ישרים, הטלות, ומרחבים שמורים.

משפט קיילי המילטון

משפט (קיילי המילטון): יהי $\mathbb{F}$ שדה ו־$A \in M_n(\mathbb{F})$ מטריצה $n \times n$. נסמן את הפ"א של A ע"י $p(x) = \det(x\mathbb{I} - A)$. אזי $p(A) = 0_{n\times n}$.

נשים לב שלא ניתן להוכיח את הטענה על ידי הצבה פשוטה של $x = A$, כי x הינו סקלר ואין משמעות להצבה של A במקומו. בנוסף, כפי שציינו בהרצאה הקודמת, אנו יודעים שמימד אוסף המטריצות הוא n^2, ולכן האוסף:

$$A^0, A^1, ..., A^{n^2}$$

שמכיל n^2+1 מטריצות הוא תלוי ליניארית, ולכן קיימת קומבינציה ליניארית של חזקות של A שסכומה אפס, כלומר קיים פולינום ש־A מאפסת. אך למשפט קיילי המילטון שני יתרונות - הוא גם נותן לנו במפורש את הפולינום, והוא גם עושה זאת עם פולינום ממעלה n.

דוגמה: נביט במטריצה $A = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}$, אז הפ"א של A הוא

$$p(x) = (x-1)(x-2) = x^2 - 3x + 2$$

ואכן מתקיים:

$$p(A) = A^2 - 3A + 2\mathbb{I} = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}^2 - 3\begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix} + 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 9 \\ 0 & 4 \end{pmatrix} - \begin{pmatrix} 3 & 9 \\ 0 & 6 \end{pmatrix} + \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = 0_{2\times 2}$$

הערה: המשפט נותן לנו עוד דרך לחשב את A^{-1} (במידה וקיימת), כצירוף ליניארי של חזקות A, כי יכולנו לרשום:

$$A^2 - 3A = -2\mathbb{I} \Rightarrow A\left(-\frac{1}{2}A + \frac{3}{2}\mathbb{I}\right) = \mathbb{I}$$

ולכן $A^{-1} = -\frac{1}{2}A + \frac{3}{2}\mathbb{I}$.

הקדמה להוכחה: במהלך ההוכחה נשתמש בטריק (שימנע סרבול טכני) שמבוסס על סכום סדרה הנדסית, אנו יודעים שעבור $x \in \mathbb{R}$ המקיים $|x| < 1$ מתקיים:

$$1 + x + x^2 + ... = \frac{1}{1-x}$$

אך למעשה ניתן להביט בצד שמאל כטור פורמלי ("פולינום אינסופי") במשתנה x, וניתן לחשוב גם על $1-x$ כטור פורמלי (רק שכל איבריו מתאפסים למעט השניים הראשונים). כעת נוכל לכפול שני טורים אלו:

$$(1-x)\left(1 + x + x^2 + ...\right)$$

כאשר הדרך האינטואיטיבית לכפול אותם היא כפל איבר איבר, לכן נקבל:

$$(1-x)\left(1 + x + x^2 + ...\right) = 1 - x + x - x^2 + x^2 - x^3 + x^3 - ...$$

לכן אנו יכולים לחשוב על הטורים הפורמליים הנ"ל בתור הופכיים אחד של השני.

<u>הערה</u>: באופן כללי עבור שני טורים פורמליים $\sum_{i=0}^{\infty} a_i x^i, \sum_{i=0}^{\infty} b_i x^i$ מגדירים את הכפל באופן הבא:

$$\left(\sum_{i=0}^{\infty} a_i x^i\right)\left(\sum_{i=0}^{\infty} b_i x^i\right) = \sum_{i=0}^{\infty} c_i x^i$$

כאשר

$$c_i = \sum_{k=0}^{i} a_k b_{i-k}$$

<u>המשך</u>: כאמור, קיבלנו ש-$\sum_{i=0}^{\infty} x^i$ הוא ההפכי של $1-x$ ולהפך. עד כה הטורים הפורמליים היו עם מקדמים מעל שדה כלשהו. כעת, נביט על טורים פורמליים במשתנה x עם מקדמים שהם <u>מטריצות</u>. מפורשות נביט על טור מהצורה:

$$\mathbb{I} + Ax + A^2x^2 + ...$$

ובדיוק כמו קודם נקבל:

$$\boxed{(\mathbb{I} - xA)\left(\mathbb{I} + Ax + A^2x^2 + ...\right) = \mathbb{I}}$$

במהלך ההוכחה נשתמש בטריק זה.

<u>הוכחת משפט קיילי המילטון</u>: נסמן את הפ"א של A ע"י:

$$(*) \quad p(x) = \det(x\mathbb{I} - A) = a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0$$

כעת, נרצה לחשב את הדטרמיננטה של $\mathbb{I} - xA$. נזכור את תכונת הדטרמיננטה של כפל מטריצה בסקלר: $\det(cA) = c^n \det(A)$, ונקבל:

$$\det(\mathbb{I} - xA) = \det\left(x\left(\frac{1}{x}\mathbb{I} - A\right)\right) = x^n \det\left(\frac{1}{x}\mathbb{I} - A\right) \underbrace{=}_{(*)} x^n \left(a_n \frac{1}{x^n} + a_{n-1}\frac{1}{x^{n-1}} + ... + a_1 \frac{1}{x} + a_0\right) =$$
$$= a_n + a_{n-1}x + ... + a_1 x^{n-1} + a_0 x^n$$

כאשר במעבר המסומן $(*)$ הצבנו $\frac{1}{x}$ בפולינום $p(x)$. בסה"כ קיבלנו:

$$\boxed{\det(\mathbb{I} - xA) = a_n + a_{n-1}x + ... + a_1 x^{n-1} + a_0 x^n}$$

ניזכר כעת בזהות

$$B \cdot \operatorname{adj}(B) = \operatorname{adj}(B) \cdot B = \det(B) \cdot \mathbb{I}$$

ונציב בפרט $B = \mathbb{I} - xA$, נקבל:

$$(\mathbb{I} - xA)\operatorname{adj}(\mathbb{I} - xA) = \det(\mathbb{I} - xA) \cdot \mathbb{I}$$

כעת נשתמש בטריק עם הטור הפורמלי שהוא:

$$(\mathbb{I} - xA)\left(\sum_{k=0}^{\infty} A^k x^k\right) = \mathbb{I}$$

ונכפול את שני אגפי המשוואה $\left(\mathbb{I}-xA\right)\text{adj}\left(\mathbb{I}-xA\right)=\det\left(\mathbb{I}-xA\right)\cdot\mathbb{I}$ בטור הפורמלי $\left(\sum_{k=0}^{\infty}A^{k}x^{k}\right)$, נשים לב שכל האיברים מתחלפים ולכן נקבל:

$$\text{adj}\left(\mathbb{I}-xA\right)=\left(\sum_{k=0}^{\infty}A^{k}x^{k}\right)\left(a_{n}+a_{n-1}x+...+a_{1}x^{n-1}+a_{0}\right)$$

<u>אבחנה</u>: כניסות המטריצה $\text{adj}\left(\mathbb{I}-xA\right)$ הן פולינומים ב־x ממעלה לכל היותר $n-1$, כי כל כניסה היא דטרמיננטה של בלוק $(n-1)\times(n-1)$.

<u>דוגמה להמחשת האבחנה</u>: ניקח $A=\begin{pmatrix}1 & 3\\ 0 & 2\end{pmatrix}$, אזי:

$$\mathbb{I}-xA=\begin{pmatrix}1-x & -3x\\ 0 & 1-2x\end{pmatrix}$$

ולכן:

$$\text{adj}\left(\mathbb{I}-xA\right)=\begin{pmatrix}1-2x & 3x\\ 0 & 1-x\end{pmatrix}$$

ואכן במקרה זה $n=2$ והמעלה הגבוהה ביותר של הכניסות היא $n-1=1$. נשים לב שניתן לפרק את המטריצה הנ"ל לטור פורמלי:

$$\begin{pmatrix}1-2x & 3x\\ 0 & 1-x\end{pmatrix}=\underbrace{\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}}_{B_{0}}+\underbrace{\begin{pmatrix}-2 & 3\\ 0 & -1\end{pmatrix}}_{B_{1}}x^{1}$$

נחזור להוכחה, מאחר וכניסות המטריצה $\text{adj}\left(\mathbb{I}-xA\right)$ הן פולינומים ממעלה $n-1$, נוכל לכתוב אותה כטור פורמלי

$$\text{adj}\left(\mathbb{I}-xA\right)=\sum_{i=0}^{\infty}B_{i}x^{i}$$

כאשר עבור $i\geq n$ בהכרח יתקיים $B_{i}=0_{n\times n}$.

בסה"כ קיבלנו שוויון של טורים פורמליים (עם מקדמים מטריציוניים):

$$\left(a_{n}+a_{n-1}x+...+a_{1}x^{n-1}+a_{0}x^{n}\right)\left(\sum_{k=0}^{\infty}A^{k}x^{k}\right)=\sum_{k=0}^{\infty}B_{k}x^{k}$$

כאשר $B_{k}=0_{n\times n}$ ל־$k\geq n$.

נשווה מקדמים עבור x^{n}:

$$a_{n}A^{n}+a_{n-1}A^{n-1}+...+a_{1}A+a_{0}\mathbb{I}=0$$

אבל נשים לב שאגף שמאל הוא בדיוק $p\left(A\right)$, ובכך סיימנו! ■

סכומים ישרים, הטלות ומרחבים T-שמורים

הגדרה: יהי V מרחב וקטורי סוף־מימדי, מעל שדה $\mathbb{F}$, ויהיו $W_1, ..., W_t \subset V$ תתי־מרחבים. נאמר שהסכום

$$\sum_{i=1}^{t} W_i = \{w_1 + ... + w_t \,:\, w_i \in W_i\}$$

הוא סכום ישר, אם

$$W_i \cap \left(\sum_{k \neq i} W_k \right) = \{0\}$$

לכל $1 \leq i \leq t$. במקרה זה נסמן

$$\bigoplus_{i=1}^{t} W_i$$

נזכיר שבאלגברה א' למדנו גרסה פשוטה יותר של משפט זה, בה עסקנו רק בשני תתי מרחב $W_1, W_2 \subset V$, ואמרנו שהסכום הוא ישר אם $W_1 \cap W_2 = \{0\}$. אנו רואים שבמקרה הכללי לא מספיק לדרוש שיתקיים $W_i \cap W_j = \{0\}$ לכל i, j, אלא נדרש חיתוך של המרחב W_i עם סכום כל המרחבים האחרים.

טענה (שנוכיח בשיעורי הבית): הטענות הבאות שקולות:

1. $\sum_{i=1}^{t} W_i$ הוא סכום ישר.
2. לכל $v \in \sum_{i=1}^{t} W_i$ יש הצגה יחידה כסכום $v = w_1 + ... + w_t$ כאשר $w_i \in W_i$.
3. $\dim \left(\sum_{i=1}^{t} W_i \right) = \sum_{i=1}^{t} \dim W_i$

רמז: הראו שאם B_i הוא בסיס של W_i אז איחוד הבסיסים $\bigcup B_i$ הוא בסיס של V. אנו גם נשתמש בטענה זו במהלך ההרצאה.

דוגמה: נביט במרחב $\mathbb{R}^2$, ונביט בשני תתי המרחבים $\text{span} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$, $\text{span} \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$. אזי כל וקטור ב-$\mathbb{R}^2$ ניתן לכתיבה בצורה יחידה כסכום של וקטורים מתתי מרחבים אלו:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ y \end{pmatrix}$$

וזהו אכן סכום ישר, וקל לראות שהמימד של המרחב כולו שווה לסכום המימדים של שני תתי המרחב ($1 + 1 = 2$).

הגדרה: יהי V מרחב וקטורי ו-$W \subset V$ תת־מרחב. נאמר שתת־מרחב $U \subset V$ הוא משלים ישר של W אם $V = W \oplus U$.

מוטיבציה לנושא של הטלות: פירוק מרחבים לסכום ישר קשור קשר הדוק להטלות, שיילמד מיד. נניח שקיים לנו פירוק של V לסכום ישר של שני מרחבים (לצורך הפשטות):

$$V = W \oplus U$$

נוכל להגדיר אופרטור $P : V \to V$ באופן הבא:

$$P\left(v = w + u\right) = w$$

הטלה זו למעשה אומרת לנו איזה רכיב של v הוא מתת המרחב W, כאשר אנו יודעים שיש כתיבה יחידה של v בתור סכום $w+u$ כי סכום המרחבים הוא ישר.

וודאו את התכונות הבאות: P היא ליניארית, ומתקיים $\mathrm{Im}P=W$ ו־$\ker P=U$. כמו כן מתקיים $P^2=P$, כי:

$$P^2\left(v\right)=P\left(P\left(v\right)\right)\underbrace{=}_{v=w+u}P\left(w\right)=w=P\left(v\right)$$

$v\in V$ היה שרירותי ולכן $P^2=P$.

<u>דוגמה</u>: נביט ב־$V=\mathbb{R}^2$, ונגדיר תת־מרחב $W=\mathrm{span}\left\{\begin{pmatrix}1\\0\end{pmatrix}\right\}$ (ציר ה־x). נגדיר את ההטלה על ציר ה־x:

$$P\left(\begin{pmatrix}x\\y\end{pmatrix}\right)=\begin{pmatrix}x\\0\end{pmatrix}$$

וכאן $U=\ker P=\mathrm{span}\left\{\begin{pmatrix}0\\1\end{pmatrix}\right\}$ (ציר y).

<u>דוגמה נוספת</u>: הפעם נגדיר $\tilde{P}\left(\begin{pmatrix}x\\y\end{pmatrix}\right)=\begin{pmatrix}x-y\\0\end{pmatrix}$, וגם זו הטלה על ציר x. אך הפעם הגרעין הינו הישר $x=y$, כלומר:

$$\tilde{U}=\ker\tilde{P}=\mathrm{span}\left\{\begin{pmatrix}1\\1\end{pmatrix}\right\}$$

אנו רואים שכדי לתאר הטלה, לא מספיק לומר על מי הטלנו (כלומר מיהי התמונה), אלא גם במקביל למי הטלנו (כלומר מיהו הגרעין). בנוסף אנו רואים שהטלה תמיד תקיים את התכונה $P^2=P$. בכך סיימנו את המוטיבציה, ונעבור להגדרות הפורמליות.

<u>הגדרה</u>: יהי $P:V\to V$ אופרטור ליניארי, נאמר ש־P הוא <u>הטלה על תת־מרחב W במקביל לתת־מרחב U</u> אם:

$$P^2=P$$
$$\mathrm{Im}P=W$$
$$\ker P=U$$

לעיתים לא יעניין אותנו מיהו הגרעין, ואז יאמרו לנו רק שמדובר בהטלה על W, אך באופן כללי ברגע שנתונה הטלה גם הגרעין שלה נקבע (הוא המשלים הישר של התמונה, נראה זאת בהמשך).

<u>טענה</u>: אם $P:V\to V$ הטלה על W במקביל ל־U, אז $V=U\oplus W$ (הסכום הוא ישר).

<u>הוכחה</u>: עלינו להראות שני דברים: (1) $V=U+W$ (2) $U\cap W=\{0\}$.

מאלגברה א' אנו זוכרים את משפט המימדים:

$$\dim\left(U+W\right)=\dim U+\dim W-\dim\left(U\cap W\right)$$

בנוסף הכרנו משפט מימדים הנכון לכל אופרטור ליניארי:

$$\dim V=\dim\ker P+\dim\mathrm{Im}P=\dim U+\dim W$$

לכן מהטענה שנוכיח בשיעורי הבית, מספיק שנוכיח ש־$U\cap W=\{0\}$ כדי לקבל ש־$U+W$ הוא סכום ישר, ומכיוון ש־$U+W\subset V$ אז שוויון המימדים $\dim V=\dim\left(U+W\right)$ יגרור שוויון $V=U+W$ בין המרחבים.

אכן, יהי $x\in U\cap W$, אזי $x\in U$ וגם $x\in W$. מאחר ו־$x\in U$ אזי $P\left(x\right)=0$ (כי $U=\ker P$). מאחר ו־$x\in W$ אזי קיים $y\in V$ כך ש־$P\left(y\right)=x$ (כי $W=\mathrm{Im}P$). נציב ונקבל:

$$0=P\left(x\right)=P\left(P\left(y\right)\right)\underbrace{=}_{P^2=P}P\left(y\right)=x$$

■

<u>תכונות שימושיות של הטלות ליניאריות</u>

1) $P:V\to V$ הטלה על W במקביל ל־U $\iff$ $\mathbb{I}-P:V\to V$ הטלה על U במקביל ל־W. נוכיח תכונה זו בתרגול, החשיבות שלה היא גם שלפעמים נוח יותר לעבוד עם $\mathbb{I}-P$ וגם שתכונה זו עוזרת בהוכחות לעיתים.

2) תהי $P:V\to V$ הטלה על W, ויהי $v\in V$, אז: $v\in W \iff P(v)=v$. כלומר הטלת וקטור על תת־מרחב W נותנת את הוקטור עצמו אמ"ם הוא מלכתחילה היה באותו תת־מרחב. דרך אחרת לחשוב על טענה זו, היא לומר שהצמצום של P למרחב W היא למעשה העתקת הזהות. נוכיח טענה זו בתרגיל הבית.

3) יהי V מרחב וקטורי מעל שדה $\mathbb{F}\subset\mathbb{C}$, ו־$P:V\to V$ אופרטור הטלה, אז P לכסין ומתקיים $\text{rank}\left(P\right)=\dim\left(\text{Im}P\right)=\text{tr}\left(P\right)$. נוכיח טענה זו כעת.

<u>הוכחת טענה 3</u>: ראינו קודם שכיוון ש־P הטלה אז $V=\underbrace{\ker P}_{U}\oplus\underbrace{\text{Im}P}_{W}$. ברצוננו לחשב את העקבה של P, ולשם כך נרצה לבנות מטריצה מייצגת של P, בבסיס מתאים.

נסמן $\dim\text{Im}P=k$, יהי $\{v_1,...,v_k\}$ בסיס לתמונה $\text{Im}P$. אם נסמן $\dim V=n$, אז נשלים לבסיס $\{u_1,...,u_{n-k}\}$ של הגרעין $\ker P$. כעת מהרמז של שיעורי הבית נובע ש:

$$B=\{v_1,...,v_k,u_1,...,u_{n-k}\}$$

הוא בסיס של V. נשים לב שמתקיים:

$$\begin{aligned}
P\left(v_1\right)&=v_1\\
P\left(v_2\right)&=v_2\\
&\vdots\\
P\left(v_k\right)&=v_k\\
P\left(u_1\right)&=0\\
P\left(u_2\right)&=0\\
&\vdots\\
P\left(u_{n-k}\right)&=0
\end{aligned}$$

ולכן המטריצה המייצגת של P בבסיס B היא:

$$[P]_B=\left(\begin{array}{ccccccc}
1 & 0 & \cdots & 0 & 0 & \cdots & 0\\
0 & 1 & & & \vdots & & \vdots\\
\vdots & & \ddots & & & & \\
 & & \ddots & 0 & & & \\
0 & & 0 & 1 & 0 & & 0\\
\hline
0 & & & 0 & 0 & \cdots & 0\\
\vdots & & & \vdots & \vdots & 0 & \\
 & & & & & \ddots & \\
0 & & & 0 & 0 & & 0
\end{array}\right)$$

ואנו רואים שקיבלנו בלוק של $k\times k$ שמהווה מטריצת יחידה, ובלוק של $(n-k)\times(n-k)$ של מטריצת אפסים. לכן קיבלנו

$$\dim\text{Im}P=\text{tr}P=k$$

וברור ש־P לכסין (B בסיס מלכסן).

<u>דוגמה/תרגיל</u>: יהי $V=\mathbb{R}^3$, $W=\operatorname{span}\left\{v_1=\begin{pmatrix}1\\1\\1\end{pmatrix}\right\}$ (ישר), $U=\operatorname{span}\left\{u_1=\begin{pmatrix}1\\1\\0\end{pmatrix},u_2=\begin{pmatrix}1\\0\\1\end{pmatrix}\right\}$ (מישור). חשבו את ההטלה P על W במקביל ל־U.

דרך אחת (נאיבית) לפתור תרגיל זה היא לומר ש־W היא התמונה, ש־U הוא הגרעין, ולעשות חשבונות ישירים ־ כלומר לכתוב ש־$P(u_1)=P(u_2)=0$, ולהשתמש בכך ש־$P(v_1)$ אמור להיות מקביל לוקטור $\begin{pmatrix}1\\1\\1\end{pmatrix}$, אך זו דרך מסורבלת ולא נוחה.

דרך אחרת, יותר אלגוריתמית, היא לעבוד עם הבסיס B, בו אנו יודעים שצורת המטריצה היא $[P]_B=\begin{pmatrix}1&0&0\\0&0&0\\0&0&0\end{pmatrix}$, ולהשתמש במטריצת מעבר כדי לקבל את צורתה של P בבסיס הסטנדרטי $\left\{\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$. נפתור בדרך זו כעת.

<u>פתרון</u>: $B=\{v_1,u_1,u_2\}$ הוא בסיס ל־$\mathbb{R}^3$ (ניתן לבדוק שוקטורים אלו הם בלתי תלויים), ולפי תכונה (3) שראינו קודם מתקיים:

$$[P]_B=\begin{pmatrix}1&0&0\\0&0&0\\0&0&0\end{pmatrix}$$

כעת, מטריצת המעבר Q בין הבסיסים B ו־E (הבסיס הסטנדרטי) מקיימת

$$[P]_E=Q\,[P]_B\,Q^{-1}$$

מתקיים $Q\,[v]_B=[Qv]_E$, כלומר Q מעבירה אותנו מוקטור לפי B, לוקטור לפי E. בצורה דומה, Q^{-1} מעבירה אותנו מוקטור לפי E לוקטור לפי B. לכן

$$Q=\begin{pmatrix}\vdots&\vdots&\vdots\\v_1&u_1&u_2\\\vdots&\vdots&\vdots\end{pmatrix}$$

כלומר וקטורי הבסיס B משמשים בתור העמודות של Q. מחשבים ומקבלים ש־

$$[P]_E=Q\,[P]_B\,Q^{-1}=\begin{pmatrix}-1&1&1\\-1&1&1\\-1&1&1\end{pmatrix}$$

ולכן

$$P\left(\begin{pmatrix}x\\y\\z\end{pmatrix}\right)=\begin{pmatrix}y+z-x\\y+z-x\\y+z-x\end{pmatrix}$$

וסיימנו.

<u>מטרה</u>: עד כה למדנו על סכומים ישרים והטלות, ודנו בשני תתי המרחב שהם התמונה והגרעין של ההטלה. כעת נכליל את התוצאות שראינו כדי לאפיין סכום ישר כלשהו על ידי הטלות.

<u>טענה 1</u>: יהי $\mathbb{F}$ שדה, V מרחב וקטורי מעל $\mathbb{F}$, $W_i\subset V$ ($1\le i\le t$) תתי מרחב, $V=\bigoplus_{i=1}^t W_i$ ונסמן ב־$P_i:V\to V$ את ההטלה על W_i במקביל ל־$\sum_{k\neq i}W_k$. אז

$$\sum_{i=1}^t P_i=\mathbb{I}$$

וכן

$$P_i\circ P_j=\begin{cases}P_i & i=j\\ 0 & i\neq j\end{cases}$$

<u>הוכחה</u>: יהי $v\in V$, מאחר והסכום ישר אז ל־v פירוק יחיד בתור סכום $v=w_1+w_2+...+w_t$ כאשר $w_i\in W_i$ לכל $1\leq i\leq t$.
אז

$$\left(\sum_{i=1}^{t}P_i\right)(v)=\sum_{i=1}^{t}P_i\left(v\right)=\sum_{i=1}^{t}w_i=v=\mathbb{I}\left(v\right)$$

ו־v היה שרירותי ולכן קיבלנו:

$$\boxed{\sum_{i=1}^{t}P_i=\mathbb{I}}$$

בנוסף, אם $i=j$ אז $P_i\left(v\right)\underbrace{=}_{P_i^2=P_i}P_iP_i\left(v\right)=P_iP_j\left(v\right)$ ולכן במקרה זה אכן מתקבל $P_i\circ P_j=P_i$.

אם $i\neq j$: כזכור P_j הטלה על W_j במקביל ל־W_k $\ker P_j=\sum_{k\neq j}W_k$, ובדומה לכך P_i הטלה על W_i במקביל ל־W_k $\ker P_i=\sum_{k\neq i}W_k$.
לכן

$$P_i\circ P_j\left(v\right)\underbrace{=}_{P_j(v)=w_j}P_i\left(w_j\right)\underbrace{=}_{w_j\in\ker P_i}0$$

ולכן קיבלנו:

$$\boxed{P_i\circ P_j=\begin{cases}P_i & i=j\\ 0 & i\neq j\end{cases}}$$

<u>טענה 2 (נוכיח אחרי פסח)</u>: יהיו $P_i:V\to V$ אופרטורים ליניאריים המקיימים $\sum P_i=\mathbb{I}$, ו־$P_i\circ P_j=0$ לכל $i\neq j$. אזי P_i הטלה על $W_i=\mathrm{Im}P_i$ במקביל ל־$\sum_{k\neq i}W_k$ ומתקיים $V=\bigoplus_{i=1}^{t}W_i$.

<u>לאן פנינו מועדות?</u> בהינתן העתקה T, אנו נרצה למצוא לה צורה כמעט אלכסונית (או במקרה הלכסין צורה אלכסונית), לכן אם למשל המרחב הוא סכום ישר $W\oplus U$ נרצה להביט במרחבים שייקראו <u>מרחבים T שמורים</u>, עבורם מתקיים:

$$T\left(W\right)\subset W$$
$$T\left(U\right)\subset U$$

לאחר שנעשה זאת, נלמד שני משפטי פירוק חזקים שישרתו את מטרתנו.

<u>הרצאה 4 - 9 לאפריל</u>

<u>תזכורת</u>: V מ"ו (מימד סופי) מעל שדה $\mathbb{F}$. אופרטור ליניארי $P:V\to V$ נקרא <u>הטלה</u> <u>על</u> תת־מרחב W <u>במקביל</u> ל־U, אם $P=P^2$, $W=\mathrm{Im}P$ ו־$U=\ker P$. במקרה זה $V=U\oplus W$ (סכום ישר). מכאן שכל $v\in V$ ניתן לכתוב כסכום $v=u+w$ כאשר $u\in U$ ו־$w\in W$ בצורה יחידה, ואז ההטלה תפעל באופן הבא:

$$P\left(v\right)=P\left(u+w\right)=w$$

כלומר ההטלה נותנת את "הרכיב ה־w" של v.

אם B_1 הוא בסיס של W, ו־B_2 הוא בסיס של U, אזי $B_1 \cup B_2$ הוא בסיס של V והמטריצה המייצגת של P בבסיס זה היא:

$$[P]_{B_1 \cup B_2} = \left(\begin{array}{ccccccccc}
1 & 0 & \cdots & & 0 & 0 & \cdots & & 0 \\
0 & 1 & & & & \vdots & & & \vdots \\
\vdots & & \ddots & & & & & & \\
 & & & \ddots & 0 & & & & \\
0 & & & 0 & 1 & 0 & & & 0 \\
\hline
0 & & & & 0 & 0 & \cdots & & 0 \\
\vdots & & & & \vdots & \vdots & 0 & & \\
 & & & & & & & \ddots & \\
0 & & & & 0 & 0 & & & 0
\end{array}\right)$$

בנוסף, הכללנו (והוכחנו) את הרעיון הזה באופן הבא:

טענה 1: אם $V = \bigoplus_{i=1}^{t} W_i$, $P_i : V \to V$ הטלה על W_i במקביל ל־$\sum_{k \neq i} W_k$. אז $\sum_{i=1}^{t} P_i = \mathbb{I}$ וכמו כן $P_i \circ P_j = \begin{cases} P_i & i = j \\ 0 & i \neq j \end{cases}$.

הערה: תזכורת מאלגברה א' ־ בטענה 1 רשמנו סכום של אופרטורים, וגם הרכבה שלהם. כדאי לזכור כיצד הם הוגדרו באלגברה א'.

סכום של אופרטורים מתנהג בצורה הבאה:

$$(T + S)(v) = T(v) + S(v)$$

הרכבה של אופרטורים מתנהגת בצורה הבאה:

$$T \circ S(v) = T(S(v))$$

בשיעור הקודם ניסחנו את טענה 2, ונוכיח אותה כעת.

טענה 2: יהיו $P_i : V \to V$ אופרטורים ליניאריים המקיימים $\sum P_i = \mathbb{I}$, ו־$P_i \circ P_j = 0$ לכל $i \neq j$. אזי P_i הטלה על $W_i = \mathrm{Im} P_i$ במקביל ל־$\sum_{k \neq i} W_k$ ומתקיים $V = \bigoplus_{i=1}^{t} W_i$.

הוכחה: עלינו להראות שלושה דברים:

(1) P_i הטלה.

(2) $V = \bigoplus W_i$.

(3) P_i הטלה במקביל ל־$\sum_{k \neq i} W_k$ (ברור שהיא על $\mathrm{Im} P_i$).

יהי $v \in V$. נוכיח את (1):

$$P_i(v) = P_i(\mathbb{I}(v)) \underbrace{=}_{\mathbb{I} = \sum P_i} P_i\left(\sum_{k=1}^{t} P_k(v)\right) \underbrace{=}_{P_i \text{ is linear}} \sum_{k=1}^{t} P_i P_k(v) \underbrace{=}_{P_i P_k = 0 \text{ for } i \neq k} P_i P_i(v) = P_i^2(v)$$

כעת נוכיח את (2): עלינו להראות $V = \sum_{i=1}^{t} W_i$ וכן שהסכום ישר. נסמן $w_i = P_i(v) \in W_i$ (לפי הגדרה). כעת, מהנתון

$$v = \mathbb{I}(v) = \sum_{i=1}^{t} P_i(v) = \sum_{i=1}^{t} w_i$$

נותר להראות שהסכום הוא ישר. נראה כי $W_i \bigcap \sum_{k \neq i} W_k = \{0\}$. יהי $w \in W_i \bigcap \sum_{k \neq i} W_k$, לכן $w \in W_i$ וגם $w \in \sum_{k \neq i} W_k$.

$w \in W_i \Leftarrow P_i(w) = w$.

$w \in \sum_{k\neq i} W_k \Leftarrow w=\sum_{k\neq i} w_k$ כאשר $w_k \in W_k$.

נציב את התוצאה השניה בראשונה ונקבל:

$$w=P_i(w)=P_i\left(\sum_{k\neq i} w_k\right)\underbrace{=}_{\text{linearity}}\sum_{k\neq i} P_i(w_k)=\sum_{k\neq i} P_i(P_k(w_k))\underbrace{=}_{P_iP_k=0\text{ for }i\neq k}\sum 0=0$$

ובכך הוכחנו שהסכום הוא ישר.

כעת נוכיח את (3): החשבון לעיל ממחיש כי:

$$P_i\left(\sum_{k\neq i} w_k\right)=0$$

ולכן $\sum_{k\neq i} W_k \subset \ker P_i$, בנוסף נראה שוויון מימדים בין תתי המרחב $\sum_{k\neq i} W_k$ ו־$\ker P_i$.

$$\dim\left(\sum_{k\neq i} W_k\right)\underbrace{=}_{*}\sum_{k\neq i}\dim W_k=\sum_{k\neq i}\dim W_k+\dim W_i-\dim W_i=\sum_{k=1}^{t}\dim W_k-\dim W_i=\dim V-\dim W_i\underbrace{=}_{**}\dim\ker P_i$$

כאשר במעבר $(*)$ השתמשנו בכך ש־$\sum_{k\neq i} W_k$ הוא סכום ישר (ולכן המימד של הסכום הוא סכום המימדים) ובמעבר $(**)$ השתמשנו במשפט המימדים הבא:

$$\dim V=\dim\underbrace{\ker P_i}_{W_i}+\dim \mathrm{Im}P_i=\dim W_i+\dim\mathrm{Im}P_i$$

■

<u>מרחבים שמורים</u>

<u>מוטיבציה:</u> עד כה למדנו על סכום ישר באופן כללי, כעת נלמד כיצד להשתמש בו כדי להגיע למטרה שלנו ־ שבהינתן אופרטור $T:V\to V$, נוכל למצוא את צורת ז'ורדן שלו.

הרעיון: בהינתן $T:V\to V$, אם $V=U\oplus W$ המקיימים $T(W)\subset W, T(U)\subset U$ אז אם ניקח בסיס של W ובסיס של U (ונבחר "בסיס טוב"), אז כש־T פועל על איברי הבסיס של W נקבל איברים ב־W, וכש־T פועל על איברי הבסיס של U נקבל איברים ב־U, ולכן המטריצה המייצגת של T בבסיס זה תיראה כמו מטריצה אלכסונית־בלוקים. היתרון בכך הוא שהרבה מאוד מאיברי המטריצה יתאפסו, וזה מקרב אותנו למטרה שלנו.

<u>הגדרה:</u> יהי $T:V\to V$ אופרטור ליניארי, ו־$W\subset V$ תת־מרחב. נאמר ש־W הוא T־שמור אם לכל $w\in W$ מתקיים $T(w)\in W$.

<u>דוגמאות למרחבים T שמורים:</u>

- $\ker T$, $\mathrm{Im}T$, $\{0\}$, V
- אם $\lambda\in\mathbb{F}$ הוא ע"ע אז $V_\lambda=\{v\in V\ :\ T(v)=\lambda v\}$ (המרחב העצמי) הוא גם T־שמור.

<u>הרצאה 5 ־ 10 לאפריל</u>

נזכיר:

<u>טענה 1:</u> אם $V=\bigoplus_{i=1}^{t} W_i$ ו־P_i הטלה על W_i במקביל ל־$\sum_{k\neq i} W_k$ אז $\sum_{i=1}^{t} P_i=\mathbb{I}$, ו־$P_iP_j=0$ עבור $i\neq j$.

<u>טענה 2:</u> אם $P_i:V\to V$ ליניאריות כך ש־$\sum_{i=1}^{t} P_i=\mathbb{I}$, ו־$P_iP_j=0$ עבור $i\neq j$, אז $V=\bigoplus_{i=1}^{t} W_i$ (כאשר $W_i=\mathrm{Im}P_i$) ו־P_i הטלה על W_i במקביל ל־$\sum_{k\neq i} W_k$.

<u>תזכורת:</u> תהי $T:V\to V$, תת־מרחב $W\subset V$ נקרא <u>T־שמור</u> אם $T(W)\subset W$.

<u>דוגמאות למרחבים T שמורים:</u>

- $\ker T$, $\mathrm{Im}T$, $\{0\}$, V
- אם $\lambda \in \mathbb{F}$ הוא ע"ע אז $V_\lambda = \{v \in V \,:\, T(v) = \lambda v\}$ (המרחב העצמי) הוא גם T־שמור.

נוכיח עבור V_λ. נניח ש־$v \in V_\lambda$. צ"ל $T(v) \in V_\lambda$, כלומר עלינו להראות שהוקטור $T(v)$ הוא וקטור עצמי של האופרטור T עם ע"ע λ, נבדוק זאת:

$$T(T(v)) \underbrace{=}_{v \in V_\lambda} T(\lambda v) = \lambda T(v) \Rightarrow T(v) \in V_\lambda$$

<u>עוד דוגמה חשובה:</u> אם $S, T : V \to V$ אופרטורים ליניאריים, כך ש־$T \circ S = S \circ T$ אז $\ker S$, $\mathrm{Im}S$ הם T־שמורים.

<u>הוכחה:</u> יהי $v \in \ker S$, ונראה כי $T(v) \in \ker S$, נחשב:

$$S(T(v)) \underbrace{=}_{ST=TS} T(S(v)) \underbrace{=}_{v \in \ker S} T(0) = 0 \Rightarrow T(v) \in \ker S$$

לפיכך $\ker S$ הוא T־שמור.

כעת, יהי $v \in \mathrm{Im}S$ ונראה ש־$T(v) \in \mathrm{Im}S$. עלינו להראות שיש $w \in V$ כך ש־$T(v) = S(w)$. אנו יודעים שיש $u \in V$ כך ש־$v = S(u)$, נפעיל T ונקבל:

$$T(v) = T(S(u)) \underbrace{=}_{TS=ST} S(T(u))$$

ואם נסמן $w = T(u)$ נקבל $T(v) = S(w)$, כנדרש. לכן $\mathrm{Im}S$ הוא T־שמור.

תוצאה זו מאפשרת לנו לקבל שתי תוצאות מעניינות, שנוכיח אחת מהן בתרגול הכיתה ואחת בתרגיל הבית.

<u>תרגיל (שיעורי בית)</u>

יהיו $S, T : V \to V$ אופרטורים ליניאריים <u>לכסינים</u> שניהם. הראו שקיים בסיס B של V כך ש־$[T]_B, [S]_B$ אלכסוניות שתיהן $\iff$ $TS = ST$.

<u>תרגיל (בתרגול)</u>

$T : V \to V$ אופרטור ליניארי, ו־$P : V \to V$ הטלה על W במקביל ל־U. הראו כי U, W הם T־שמורים $\iff$ $PT = TP$.

<u>הערה:</u> תרגיל זה חשוב לנו להמשך, מאחר ואנו רוצים להבין מתי שני מרחבים שסכומם ישר הם גם T־שמורים, כי ראינו שבמקרה זה המטריצה המייצגת של T היא מטריצת בלוקים, וזה מקרב אותנו ל"מטרת העל" הסופית שלנו של הבאת מטריצה לצורת ג'ורדן.

<u>הגדרה:</u> יהי $T : V \to V$ אופרטור ליניארי ויהי $W \subset V$ תת־מרחב T־שמור. אז נגדיר אופרטור $T|_W : W \to W$ שנקרא <u>הצמצום של T ל־W</u> המקיים:

$$T|_W(u) = T(u)$$

כמובן שזהו אופרטור ליניארי (כי T ליניארי), ובתור פונקציות T ו־$T|_W$ הן לא אותה פונקציה, אבל הן פועלות באותה צורה על איברי W.

הסיבה שהגדרנו את האופרטור $T|_W$ היא שבהמשך אנו נפרק את המרחב V לתתי־מרחבים המוכלים בו, ומטבע הדברים המימד שלהם יהיה קטן יותר. המטריצה המייצגת של $T|_W$ תהיה קטנה הרבה יותר מזו של T, ואנו נרצה לבנות את המטריצה המייצגת של T על ידי הבנת המטריצות המייצגות של $T|_W$ עבור תתי המרחבים השונים שסכומם הישר הוא V.

<u>דוגמה:</u> $T : \mathbb{R}^3 \to \mathbb{R}^3$ מוגדר באופן הבא:

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x \\ 2y \\ 3z \end{pmatrix}$$

נסמן את וקטורי הבסיס הסטנדרטי באופן הבא:

$$b_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \; b_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \; b_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

אזי המרחב העצמי של $\lambda = 2$ הוא:

$$W = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\} = \operatorname{span}\{b_1, b_2\}$$

והמרחב העצמי של $\lambda = 3$ הוא:

$$U = \operatorname{span}\left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} = \operatorname{span}\{b_3\}$$

הצמצום של T ל-W:

מתקיים $\begin{pmatrix} 2x \\ 2y \\ 0 \end{pmatrix} = T|_W \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$ ולכן $T|_W : W \to W$ הוא:

$$[T|_W]_{\{b_1, b_2\}} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}_{2\times 2}$$

הצמצום של T ל-U:

מתקיים $\begin{pmatrix} 0 \\ 0 \\ 3z \end{pmatrix} = T|_U \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix}$ ולכן $T|_U : U \to U$ הוא:

$$[T|_U]_{\{b_3\}} = \{3\}_{1\times 1}$$

ונשים לב שמתקיים:

$$[T]_{\{b_1, b_2, b_3\}} = \begin{pmatrix} 2 & 0 & \\ 0 & 2 & \\ & & 3 \end{pmatrix}$$

משפט הפירוק הספקטרלי

משפט הפירוק הספקטרלי: יהי V מרחב וקטורי (סוף מימדי) מעל שדה $\mathbb{F}$, $T:V\to V$ אופרטור ליניארי לכסין עם ערכים עצמיים $\lambda_1,...,\lambda_t\in\mathbb{F}$ (שונים, ייתכן ריבוי שלהם). נסמן ב־$W_i=V_{\lambda_i}$ את המרחב העצמי של λ_i עבור $i=1,...,t$, ונסמן ב־P_i את ההטלה על W_i במקביל ל־$\sum_{k\neq i}W_k$. אזי:

1. $V=\bigoplus_{i=1}^{t}W_i$ (זה נקרא הפירוק הספקטרלי של המרחב V).
2. $\sum_{i=1}^{t}P_i=\mathbb{I}$.
3. $P_iP_j=\begin{cases}P_i & i=j\\ 0 & i\neq j\end{cases}$
4. $T=\sum_{i=1}^{t}\lambda_iP_i$ (זה נקרא הפירוק הספקטרלי של האופרטור T).
5. לכל פולינום $f(x)\in\mathbb{F}[x]$ מתקיים $f(T)=\sum_{i=1}^{t}f(\lambda_i)P_i$.

יתר על כן א': הפירוק הנ"ל הוא יחיד במובן הבא: אם $\alpha_1,...,\alpha_t\in\mathbb{F}$, $\tilde{P}_1,...,\tilde{P}_t$ אופרטורים ליניאריים שונים מ־0 כך שתכונות $2,3,4$ מתקיימות עבורם, אז T לכסין עם הערכים העצמיים α_i, $\tilde{P}_i$ הן ההטלות על המרחבים העצמיים $\mathrm{Im}\tilde{P}_i$ ו־$\tilde{W}_i$ ו־$V=\bigoplus_{i=1}^{t}\tilde{W}_i$.

יתר על כן ב': למעשה P_i היא הפולינום הבא ב־T:

$$P_i=\varphi_i(T)$$

כאשר $\varphi_i(x)$ הוא פולינום המוגדר באופן הבא:

$$\varphi_i(x)=\prod_{k\neq i}\frac{(x-\lambda_k)}{(\lambda_i-\lambda_k)}$$

לפולינום הזה קוראים אינטרפולציית לגרנז'.

הערות

א. זהו משפט פירוק עבור אופרטור לכסין, אך בהמשך נכליל תוצאה זו גם לאופרטורים שאינם לכסינים.

ב. הכוונה בסימון $f(T)$ דומה להצבת מטריצה בפולינום, אם למשל $f(x)=x^3+2x^2$, אזי $f(T)=T^3+2T^2$ (כאשר $T^3=T\circ T\circ T$, כלומר זו הפעלה של T שלוש פעמים).

ג. היותן של ההטלות פולינומים ב־T למעשה מהווה נימוק נוסף לכך שהן מתחלפות עם T.

ד. המשמעות של סעיף 4 היא שאם אנו מסתכלים על T לפי בסיס של וקטורים עצמיים B, נקבל מטריצת בלוקים:

$$[T]_B=\begin{pmatrix}\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_1\end{pmatrix} & & & \\ & \begin{pmatrix}\lambda_2 & & \\ & \ddots & \\ & & \lambda_2\end{pmatrix} & & \\ & & \ddots & \\ & & & \begin{pmatrix}\lambda_t & & \\ & \ddots & \\ & & \lambda_t\end{pmatrix}\end{pmatrix}$$

כאשר כל בלוק הוא המטריצה המייצגת של הצמצום של T למרחב העצמי המתאים, כלומר הבלוק $\begin{pmatrix}\lambda_i & & \\ & \ddots & \\ & & \lambda_i\end{pmatrix}$ הוא בעצם המטריצה המייצגת של:

$$T|_{W_i}=\lambda_iP_i|_{W_i}=\lambda_i\mathbb{I}|_{W_i}$$

הוכחת משפט הפירוק הספקטרלי

סעיף 1: באלגברה א' ראינו שאם T לכסין אז יש בסיס B של V המורכב מוקטורים עצמיים. נחלק בסיס זה לבסיסים B_i של W_i, ולכן ינבע (מתרגיל 3 בגיליון בית 1) ש־$V = \bigoplus_{i=1}^{t} W_i$.

סעיפים 2,3: נובעים ישירות מטענה 1 שניסחנו בתחילת הרצאה זו.

סעיף 4: יהי $v \in V$, מהגדרת P_i בתור ההטלה על W_i (המרחב העצמי המתאים ל־λ_i) נקבל ש:

$$\left(T \circ P_i\right)(v) = T\left(\underbrace{P_i\left(v\right)}_{\in W_i}\right) = \lambda_i P_i\left(v\right)$$

זה נכון לכל $v \in V$, ולכן קיבלנו שוויון בין האופרטורים:

$$T \circ P_i = \lambda_i P_i$$

לפיכך נקבל:

$$T = T \circ \mathbb{I} \underbrace{=}_{(2)} T\left(\sum_{i=1}^{t} P_i\right) \underbrace{=}_{\text{linearity}} \sum_{i=1}^{t} T \circ P_i = \sum_{i=1}^{t} \lambda_i P_i$$

סעיף 5: נביט על T^2, לאור סעיף 4 (אותו כבר הוכחנו) משמעות הביטוי T^2 היא:

$$T^2 = \left(\sum_{i=1}^{t} \lambda_i P_i\right) \circ \left(\sum_{j=1}^{t} \lambda_j P_j\right) \underbrace{=}_{\text{linearity}} \sum_{i=1}^{t} \sum_{j=1}^{t} \lambda_i \lambda_j P_i \circ P_j$$

אנו זוכרים ש־$P_i P_j = 0$ עבור $i \neq j$ ולכן נקבל:

$$= \sum_{i=1}^{t} \lambda_i^2 P_i^2 \underbrace{=}_{P_i^2 = P_i} \sum_{i=1}^{t} \lambda_i^2 P_i$$

כעת, באינדוקציה, עבור T^n נקבל$T^n = \sum_{i=1}^{t} \lambda_i^n P_i$, ומליניאריות ניתן גם לכפול ביטוי זה בסקלר, ולכן נקבל כך את התוצאה עבור פולינום כלשהו.

יתר על כן א': משתמשים בטענה 2 כדי לקבל ש־$V = \bigoplus_{i=1}^{t} W_i$ ואז משתמשים בנתון (4) ובהגדרת ע"ע, כדי לקבל ש־$\tilde{W}_i$ הוא במרחב העצמי המתאים לע"ע α_i (מושאר כתרגיל לא פורמלי).

יתר על כן ב': עלינו להראות ש־$P_i = \varphi_i\left(T\right)$. נשים לב שמתקיים:

$$\varphi_i\left(T\right)(v) = \frac{\left(T - \lambda_1 \mathbb{I}\right)}{\left(\lambda_i - \lambda_1\right)} \frac{\left(T - \lambda_2 \mathbb{I}\right)}{\left(\lambda_i - \lambda_2\right)} \dots \frac{\left(T - \lambda_{i-1} \mathbb{I}\right)}{\left(\lambda_i - \lambda_{i-1}\right)} \frac{\left(T - \lambda_{i+1} \mathbb{I}\right)}{\left(\lambda_i - \lambda_{i+1}\right)} \dots \frac{\left(T - \lambda_t \mathbb{I}\right)}{\left(\lambda_i - \lambda_t\right)} (v)$$

כאשר $v = w_1 + w_2 + \dots + w_t$ עבור $w_k \in W_k$ (ו־W_k הוא המרחב העצמי של λ_k). נשים לב שבמונה יש $t-1$ איברים, כי לא מופיע הגורם $\left(T - \lambda_i \mathbb{I}\right)$. כל האופרטורים המופיעים במונה הם מתחלפים, ולכן ניתן להפעיל אותם על v בכל סדר שנבחר. כמו כן אנו רואים כי:

$$\left(T - \lambda_1 \mathbb{I}\right)\left(w_1\right) = \lambda_1 w_1 - \lambda_1 w_1 = \vec{0}$$

ובאותו אופן $(T-\lambda_k\mathbb{I})(w_k)=0$ לכל $k\neq i$. לכן אם נרשום $v=w_1+w_2+...+w_t$ כל הרכיבים למעט w_i יתאפסו תחת הפעלת $\varphi_i(T)$ על v. לכן נותר רק להבין את פעולת $\varphi_i(T)$ על w_i, אנו רואים שמתקיים:

$$\frac{(T-\lambda_k\mathbb{I})}{(\lambda_i-\lambda_k)}(w_i)\underbrace{=}_{T(w_i)=\lambda_i w_i}\frac{(\lambda_i-\lambda_k)}{(\lambda_i-\lambda_k)}w_i=w_i$$

ולכן קיבלנו $\varphi_i(T)(v)=w_i$. כמובן ש־$P_i(v)=w_i$ מהגדרתה כהטלה על W_i. $v\in V$ היא שרירותי ולכן $\varphi_i(T)=P_i$. ■

<u>דוגמה</u>

פירוק ספקטרלי של $A=\begin{pmatrix}2&3\\1&4\end{pmatrix}$. קל לבדוק שהע"ע של A הם $1,5$, עם מרחבים עצמיים:

$$W_{\lambda=1}=\operatorname{span}\left\{\begin{pmatrix}-3\\1\end{pmatrix}\right\}$$
$$W_{\lambda=5}=\operatorname{span}\left\{\begin{pmatrix}1\\1\end{pmatrix}\right\}$$

פירוק ספקטרלי של A משמעו מציאת A_1,A_5 כך ש־$A_1^2=A_1,\ A_5^2=A_5$ (כי A_1,A_5 הטלות) המקיימים:

$$A=1\cdot A_1+5\cdot A_5$$

נחשב את A_1,A_5 בדרך הישירה (ניתן לעשות זאת גם בדרך אחרת - עם מטריצת מעבר). אנו יודעים שהמרחב V הוא סכום ישר של שני תתי המרחב העצמיים, ולכן ניתן לרשום כל וקטור $\begin{pmatrix}x\\y\end{pmatrix}\in V$ בצורה יחידה כקומבינציה של וקטור מ־$W_{\lambda=1}$ עם וקטור מ־$W_{\lambda=5}$. אכן:

$$\begin{pmatrix}x\\y\end{pmatrix}=\left(\frac{-x+y}{4}\right)\begin{pmatrix}-3\\1\end{pmatrix}+\left(\frac{x+3y}{4}\right)\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}\frac{3x-3y}{4}\\\frac{(-x+y)}{4}\end{pmatrix}+\begin{pmatrix}\frac{(x+3y)}{4}\\\frac{(x+3y)}{4}\end{pmatrix}$$

ולכן $\begin{pmatrix}3/4&-3/4\\-1/4&1/4\end{pmatrix}$ היא המטריצה המייצגת בבסיס הסטנדרטי של ההטלה P_1 על W_1, וכמו כן $\begin{pmatrix}1/4&3/4\\1/4&3/4\end{pmatrix}$ היא המטריצה המייצגת בבסיס הסטנדרטי של ההטלה P_5 על W_5. ניתן לבדוק ש־$P_1^2=P_1$ ו־$P_5^2=P_5$, וכמו כן אנו רואים כי:

$$\begin{pmatrix}2&3\\1&4\end{pmatrix}=1\cdot\begin{pmatrix}3/4&-3/4\\-1/4&1/4\end{pmatrix}+5\cdot\begin{pmatrix}1/4&3/4\\1/4&3/4\end{pmatrix}$$

דרך אחרת: אנו יודעים כיצד נראית המטריצה המייצגת של כל הטלה בבסיס העצמי המתאים. למשל $[P_1]$ בבסיס העצמי של W_1 נראית כך:

$$[P_1]_{B_{\lambda=1}}=\begin{pmatrix}1&0\\0&0\end{pmatrix}$$

ו־$[P_2]$ בבסיס העצמי של W_2 נראית כך:

$$[P_2]_{B_{\lambda=5}}=\begin{pmatrix}0&0\\0&1\end{pmatrix}$$

וניתן למצוא אותן בבסיס הסטנדרטי בעזרת מטריצת המעבר שהיא:

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

ונקבל את $[P_i]_E$:

$$[P_i]_E = Q\,[P_i]_{B_{\lambda_i}}\,Q^{-1}$$

<u>הרצאה 6 ־ 16 לאפריל</u>

<u>תזכורת</u>: בפעם הקודמת שאלנו מהו הפירוק הספקטרלי של $A=\begin{pmatrix}2&3\\1&4\end{pmatrix}$ (אך הרעיון עובד בצורה דומה לכל מטריצה לכסינה או אופרטור לכסין). שיטת העבודה שלנו הייתה למצוא ע"ע, כדי לוודא שהוא אכן לכסין. ראינו שתי דרכים להגיע לפירוק הספקטרלי.

<u>דרך 1</u>: אמרה לנו שאם המטריצה לכסינה (או לחילופין האופרטור $T_A:\mathbb{R}^2\to\mathbb{R}^2$, $T_A(v)=Av$ לכסין) אז ניתן לכתוב $V=\bigoplus W_i$, מוצאים הצגה $v=w_1+...+w_k$ ואז $P_i(v)=w_i$ ולכן:

$$A_i=[P_i]_E$$

כאשר E הבסיס הסטנדרטי. הסיבוכיות בשיטה זו היא בפתרון מערכת המשוואות המתקבלת.

<u>דרך 2</u>: אמרה שעבור בסיס B של וקטורים עצמיים נקבל $[P_i]_B=\begin{pmatrix}0&&&&\\&1&&&\\&&\ddots&&\\&&&1&\\&&&&0\end{pmatrix}$ (בתרגיל שעשינו קיבלנו $[P_1]_B=\begin{pmatrix}1&0\\0&0\end{pmatrix}$ ו־$[P_5]_B=\begin{pmatrix}0&0\\0&1\end{pmatrix}$), כלומר מקבלים בלוקים של מטריצות אפסים ומטריצות יחידה. לאחר מכן אנחנו עוברים לבסיס הסטנדרטי על ידי שימוש במטריצת מעבר. הסיבוכיות בשיטה זו היא במציאת מטריצות המעבר והכפלת המטריצות בסיום כדי להציג את ההטלות בבסיס הסטנדרטי.

ישנה דרך נוספת, שלא ראינו עדיין, נראה אותה כעת.

<u>דרך 3</u>: מהמשפט ראינו ש־$P_i=\varphi_i(T)$, כאשר

$$\varphi_i(x)=\prod_{k\neq i}\frac{x-\lambda_k}{\lambda_i-\lambda_k}$$

אצלנו: יש ערכים עצמיים $\lambda_1=1$ ו־$\lambda_2=5$, ונקבל:

$$A_1=\frac{A-5\mathbb{I}}{(1-5)}=-\frac{1}{4}\left[\begin{pmatrix}2&3\\1&4\end{pmatrix}-5\begin{pmatrix}1&0\\0&1\end{pmatrix}\right]=\frac{1}{4}\begin{pmatrix}3&-3\\-1&1\end{pmatrix}$$
$$A_2=\frac{A-1\mathbb{I}}{(5-1)}=\frac{1}{4}\left[\begin{pmatrix}2&3\\1&4\end{pmatrix}-\begin{pmatrix}1&0\\0&1\end{pmatrix}\right]=\frac{1}{4}\begin{pmatrix}1&3\\1&3\end{pmatrix}$$

ואלו בדיוק התוצאות שראינו בשיעור הקודם.

פולינומים מעל שדה

יהי $\mathbb{F}$ שדה, נעסוק בפולינומים מעל $\mathbb{F}$, כלומר פולינומים שמקדמיהם איברים מ־$\mathbb{F}$.

תכונת החלוקה: בהינתן $a(x) \in \mathbb{F}[x]$ ו־$0 \neq b(x) \in \mathbb{F}[x]$ קיימים פולינומים $q(x) \in \mathbb{F}[x]$ (מנה) ו־$r(x) \in \mathbb{F}[x]$ (שארית) כך ש:

$$a(x) = q(x) b(x) + r(x)$$

כאשר $r(x) = 0$ (פולינום האפס) או $\deg(r(x)) < \deg(b(x))$. הסיבה להפרדה הזו היא שאצלנו בקורס עבור פולינום האפס לא מוגדרת דרגה. כלומר עצם הכתיבה $\deg(p(x))$ משמעה בפרט ש־$p(x)$ אינו פולינום האפס.

הגדרה (פולינום מינימלי ־ פ"מ): יהי $\mathbb{F}$ שדה, נאמר ש־$m(x) \in \mathbb{F}[x]$ (או $m_A(x)$) הוא הפולינום המינימלי של מטריצה $A \in M_n(\mathbb{F})$, אם:

1. $m(x)$ מתוקן (המקדם המוביל הוא 1).
2. $m(A) = 0_{n\times n}$.
3. $\deg(m(x))$ היא המינימלית מבין $\deg(p(x))$ המקיימים $p(x) \in \mathbb{F}[x]$ ו־$p(A) = 0_{n\times n}$.

הגדרה דומה לאופרטור: יהי $\mathbb{F}$ שדה ו־V מרחב וקטורי, נאמר ש־$m(x) \in \mathbb{F}[x]$ (או $m_T(x)$) הוא הפולינום המינימלי של אופרטור $T : V \to V$, אם:

1. $m(x)$ מתוקן (המקדם המוביל הוא 1).
2. $m(T) = 0_{n\times n}$.
3. $\deg(m(x))$ היא המינימלית מבין $\deg(p(x))$ המקיימים $p(x) \in \mathbb{F}[x]$ ו־$p(T) = 0_{n\times n}$.

הערה: מדוע $m(x)$ קיים ויחיד?

קיום: אנו יודעים שהוא קיים כי ראינו שיש פולינומים $p(x)$ כך ש־$p(A) = 0$ (למשל, אנו יודעים ממשפט קיילי־המילטון שהפולינום האופייני מאפס את A).

יחידות: נניח ש־$m(x)$ ו־$\tilde{m}(x)$ שניהם פ"מ של A. אז מתכונת החלוקה קיימים $r(x)$ ו־$q(x)$ כך ש:

$$\tilde{m}(x) = q(x) m(x) + r(x)$$

אם $r(x) = 0$ (פולינום האפס), אז $\tilde{m}(x) = q(x) m(x)$, אך $\deg(m(x)) = \deg(\tilde{m}(x))$ ולכן $q(x)$ הוא פולינום קבוע. בנוסף, $\tilde{m}(x), m(x)$ מתוקנים שניהם ולכן $q(x) = 1$, ולכן $m(x) = \tilde{m}(x)$ כנדרש.

אחרת, מתקיים $\deg(r(x)) < \deg(m(x))$, אך זה מוביל לסתירה למינימליות כי

$$\underbrace{\tilde{m}(A)}_{=0} = q(A) \underbrace{m(A)}_{=0} + r(A)$$

ומכאן נקבל $r(A) = 0$, בסתירה למינימליות.

טענה: באופן דומה, מקבלים שאם $p(x) \in \mathbb{F}[x]$ כך ש־$p(A) = 0$ אז: $m(x) | p(x)$. ההוכחה תעשה בצורה דומה על ידי חילוק עם שארית, כלומר קיימים $q(x), r(x) \in \mathbb{F}[x]$ כך שמתקיים:

$$p(x) = q(x) m(x) + r(x)$$

ואז אם $r(x) = 0$ סיימנו (כי מתקבל $p(x) = q(x) m(x)$, כלומר $m(x)$ מחלק את $p(x)$). אם $r(x)$ אינו פולינום האפס אז מתקבלת סתירה למינימליות כי

$$\underbrace{p(A)}_{=0} = q(A) \underbrace{m(A)}_{=0} + r(A)$$

ומכאן נקבל ש־$r(A) = 0$, בסתירה למינימליות.

מסקנה מקיילי־המילטון: נסמן ב־$p_A(x)$ את הפולינום האופייני של המטריצה A, אזי $m_A(x)\,|\,p_A(x)$. כלומר הפולינום המינימלי מחלק את הפולינום האופייני. זה נותן לנו דרך לחשב את $m_A(x)$, מאחר וברגע שמצאנו את $p_A(x)$ אנו יכולים לבדוק את מחלקיו.

דוגמה:

	$\begin{pmatrix}1 & 3\\ 0 & 2\end{pmatrix}$	$\begin{pmatrix}2 & 0\\ 0 & 2\end{pmatrix}$	$\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$	$\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}$
פולינום אופייני	$(x-1)(x-2)$	$(x-2)^2$	x^2	x^2
פולינום מינימלי	$(x-1)(x-2)$	$(x-2)$	x	x^2

נשים לב שהפולינום המינימלי הבחין בין שתי המטריצות האחרונות, להן אותם ערכים עצמיים (אפס הוא ערך עצמי) עם ריבוי אלגברי 2, המטריצה $\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$ לכסינה ו־$\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}$ לא לכסינה, ואנו נראה את הקשר של הפולינום המינימלי לכך במשפט הבא.

משפט: יהי V מרחב וקטורי סוף מימדי מעל שדה $\mathbb{F}$, ויהי $T:V\to V$ אופרטור ליניארי. אז T לכסין $\iff$ $m_T(x)=(x-\lambda_i)\cdot\ldots\cdot(x-\lambda_t)$, כאשר $\lambda_i\in\mathbb{F}$ שונים זה מזה. דהיינו $m_T(x)$ מתפרק מעל $\mathbb{F}$ לגורמים ליניאריים שונים.

הערה: נשים לב שיש במשפט הזה שתי אמירות לגבי המבנה של הפולינום המינימלי של אופרטור לכסין ־ גם שהוא מתפרק למכפלת גורמים ליניאריים (שזה לא מובן מאליו) וגם שהם שונים זה מזה.

נוכיח את המשפט בשיעור הבא.

הרצאה 7 ־ 23 לאפריל

הוכחת המשפט:

כיוון 1 $\Leftarrow$: נתון ש־T לכסין, נסמן ב־$\lambda_1,...,\lambda_t$ את הע"ע השונים שלו (כמובן שייתכן שיש להם ריבוי אלגברי כלשהו). נסמן:

$$p(x)=(x-\lambda_1)(x-\lambda_2)...(x-\lambda_t)$$

ונראה ש־$p(x)$ הוא הפולינום המינימלי. לשם כך עלינו להראות שלושה דברים: שהוא מתוקן, שהוא מאפס את T, ושהוא הפולינום בעל הדרגה המינימלית שמקיים תנאים אלו.

ראשית, ברור ש־$p(x)$ מתוקן (ניתן לפתוח סוגריים ולראות שהאיבר המוביל הוא x^t).

שנית, נראה ש־$p(T)=0$: מהנתון ש־T לכסין נובע ש־$V=\bigoplus_{i=1}^t W_i$, כאשר W_i הוא המרחב העצמי של λ_i. לכן בהינתן $v\in V$ מתקיים $v=w_1+...+w_t$ כאשר $w_i\in W_i$ לכל $1\le i\le t$. לכן נקבל

$$\begin{aligned}p(T)(v)&=[(T-\lambda_1\mathbb{I})(T-\lambda_2\mathbb{I})...(T-\lambda_t\mathbb{I})]\left(\underbrace{w_1+...+w_t}_{v}\right)=\\&=[(T-\lambda_1\mathbb{I})...(T-\lambda_t\mathbb{I})](w_1)+...+[(T-\lambda_1\mathbb{I})...(T-\lambda_t\mathbb{I})](w_t)\end{aligned}$$

נשים לב שכל האופרטורים $(T-\lambda_i\mathbb{I})$ מתחלפים בהרכבה זה עם זה, הסיבה היא שאופרטור סקלרי $\lambda_i\mathbb{I}$ מתחלף עם כל אופרטור, ו־T מתחלף עם עצמו, ולכן גם קומבינציות שלהם מתחלפות. לכן אנחנו יכולים לשנות את סדר הפעולה שלהם.

לכן כשנרצה להפעיל את האופרטור $(T-\lambda_1\mathbb{I})...(T-\lambda_t\mathbb{I})$ על w_i, נשנה את סדר ההרכבה כך שהאופרטור שיפעל עליו ראשון יהיה $(T-\lambda_i\mathbb{I})$, הסיבה לכך היא ש־$(T-\lambda_i\mathbb{I})(w_i)=Tw_i-\lambda_iw_i=\lambda_iw_i-\lambda_iw_i=0$, ולכן הביטוי יתאפס! לכן נקבל:

$$p(T)(v)=\vec{0}$$

וזה נכון לכל $v\in V$, ולכן $P(T)=0$.

נותר להראות ש־$\deg(P)$ מינימלית מבין כל ה־$\deg(q)$ של פולינומים $q(x)$ עבורם $q(T)=0$. נניח ש־$q(x)\in\mathbb{F}[x]$ כך ש־$q(T)=0$, ממשפט הפירוק הספקטרלי $T=\sum\lambda_iP_i$ (כאשר P_i הטלה על W_i) ומתקיים

$$0=q(T)=\sum_{i=1}^t\underbrace{q(\lambda_i)}_{\text{scalar}}P_i$$

כעת נפעיל P_j על המשוואה לעיל, ומליניאריות נקבל:

$$0=\sum_{i=1}^{t}q\left(\lambda_i\right)P_jP_i$$

כזכור מתקיים $P_jP_i=\begin{cases}0 & i\neq j\\ P_j & i=j\end{cases}$ ולכן נקבל:

$$q\left(\lambda_j\right)P_j=0$$

אבל P_j איננו אופרטור האפס, ולכן בהכרח $q\left(\lambda_j\right)=0$. בסה"כ קיבלנו שאם $q\left(x\right)\in\mathbb{F}\left[x\right]$ כך ש־$q\left(T\right)=0$ אז $q\left(\lambda_i\right)=0$ לכל ע"ע λ_i של T. לפיכך הפולינום בעל $\deg$ מינימלית ש־T מאפס חייב להתחלק בכל אחד מ־$\left(x-\lambda_i\right)$, אך אלו גורמיו היחידים (אחרת נקבל סתירה למינימליות). לכן $p\left(x\right)=m_T\left(x\right)$.

כיוון $2\Rightarrow$: נתון ש־$m_T\left(x\right)=\left(x-\lambda_1\right)...\left(x-\lambda_t\right)$, ועלינו להראות שהאופרטור T לכסין. ממשפט הפירוק הספקטרלי (החלק של "יתר על כן א"), מספיק שנראה ש־T מקיים

$$T=\sum_{i=1}^{t}\lambda_iP_i$$

כאשר $P_i:V\to V$ הם אופרטורים ליניאריים שמקיימים

$$\sum_{i=1}^{t}P_i=\mathbb{I}$$
$$P_iP_j\underbrace{=}_{i\neq j}0$$

נגדיר

$$\varphi_i\left(x\right)=\prod_{k\neq i}\frac{x-\lambda_k}{\lambda_i-\lambda_k}$$

נשים לב ש־$\varphi_i\left(x\right)$ הוא פולינום (לו קראנו פולינום אינטרפולציית לגרנז') בעל $\deg\left(\varphi_i\right)=t-1$, המחזיר 0 עבור $x=\lambda_j$ ($i\neq j$), ומחזיר 1 עבור $x=\lambda_i$. לפיכך אם $y_1,...,y_t\in\mathbb{F}$ אז נביט ב־$y_i\varphi_i\left(x\right)$, זהו פולינום שמחזיר 0 עבור $x=\lambda_j$ ($i\neq j$) ומחזיר y_i עבור $x=\lambda_i$. נגדיר את הפולינום $h\left(x\right)$ על ידי הסכום הבא

$$h\left(x\right)=\sum_{i=1}^{t}y_i\varphi_i\left(x\right)$$

זהו פולינום ממעלה $\leq t-1$, עבורו $h\left(\lambda_i\right)=y_i$.

בפרט, נשים לב שאם נבחר $y_1=y_2=...=y_t=1$ אזי לכל λ_i יתקיים $h\left(\lambda_i\right)=1$, ובעצם נקבל שמתקיים

$$\sum_{i=1}^{t}1\varphi_i\left(x\right)=1$$

ההסבר: בשני האגפים יש לנו פולינומים ממעלה $\leq t-1$ המזדהים על t דגימות <u>שונות</u> (הערכים העצמיים), ולכן הם שווים כפולינומים (ומכאן השם "אינטרפולציה").

<u>הערה:</u> אנו מניחים כאן ש־$t>1$. אם $t=1$ אז הטענה נכונה לפי ההסבר הבא: מהנתון $m_T\left(x\right)=\left(x-\lambda\right)$ נובע $T=\lambda\mathbb{I}$ וזהו אופרטור סקלרי, ולכן לכסין.

המשך ההוכחה: כעת, נשים לב שאם נבחר $y_i = \lambda_i$ אזי נקבל:

$$\sum_{i=1}^{t} \lambda_i \varphi_i (x) = x$$

מאותו נימוק של אינטרפולציה שראינו מקודם.

לכן אם נגדיר $P_i = \varphi_i (T)$ אז נקבל את הדרוש:

$$\sum_{i=1}^{t} \lambda_i P_i = T$$

$\blacksquare$

הרצאה 8 - 24 לאפריל

תזכורת: ראינו שאם $T : V \to V$ אופרטור ליניארי, אז T לכסין $\iff$ $m_T (x) = (x - \lambda_1) (x - \lambda_2) ... (x - \lambda_t)$.

מסקנה: אם W תת־מרחב T-שמור, אז T לכסין $\iff$ $T|_W$ לכסין.

הוכחה: ראינו בתרגול הכיתה ש-$m_{T|_W} (x) | m_T (x)$ ולכן מהמשפט בתזכורת אם T לכסין אז $m_T (x)$ מתפרק לגורמים ליניאריים שונים ולכן $m_{T|_W} (x)$ מתפרק לגורמים ליניאריים שונים, ולכן (שוב מהמשפט) $T|_W$ לכסין.

הערה: איפה השתמשנו בכך ש-W הוא T-שמור? השתמשנו בכך בעצם ההגדרה של $T|_W$, אם W אינו T שמור אז לא ניתן כלל להגדיר את הצמצום שלו ל-W.

עד כה עסקנו באופרטורים לכסינים, כעת נלמד משפט פירוק חדש, שעוסק באופרטורים באופן כללי.

משפט הפירוק הפרימרי

מספר מושגים ועובדות על פולינומים מעל שדה $\mathbb{F}$:

הגדרה: נאמר ש־$p(x) \in \mathbb{F}[x]$ $0 \neq$ הוא פולינום אי פריק אם שוויון מהצורה $p(x) = h(x)g(x)$ (כאשר $h(x), g(x) \in \mathbb{F}[x]$) גורר ש־$h(x)$ פולינום קבוע או $g(x)$ פולינום קבוע.

הערה: זה דומה לתכונה של מספרים ראשוניים ־ אם מספר ראשוני הוא מכפלה של שני מספרים טבעיים, אזי אחד מהם הוא 1.

דוגמה: $p(x) = x^2 + 1$ אי פריק מעל $\mathbb{R}$, כי אם היינו מפרקים אותו למכפלת שני פולינומים, כל אחד מהם היה חייב להיות ממעלה 1, כלומר היו חייבים להיות לפולינום שורשים ממשיים, אך אין לו שורשים כאלו.

מצד שני, $p(x) = x^2 + 1$ כן פריק מעל $\mathbb{C}$, באופן הבא:

$$p(x) = (x - i)(x + i)$$

עובדה: לכל $p(x) \in \mathbb{F}[x]$ יש פירוק יחיד (עד כדי סדר הגורמים וכפל בקבועים הפיכים) למכפלת גורמים אי פריקים (זה אנלוגי לפירוק היחיד של כל מספר לגורמים ראשוניים בשלמים).

מושג: הממג"ב של $f(x), g(x)$ יסומן $d(x) = \gcd(f(x), g(x))$, בדומה לממג"ב במקרה של מספרים שלמים, ניתן למצוא אותו בעזרת אלגוריתם אוקלידס או על ידי פירוק למכפלת אי־פריקים ולקחת את הגורמים המשותפים.

עובדה: קיימים $a(x), b(x) \in \mathbb{F}[x]$ כך ש־$d(x) = a(x)f(x) + b(x)g(x)$, בדומה לממג"ב במקרה של מספרים שלמים, ניתן למצוא אותם על ידי שימוש באלגוריתם אוקלידס ו"הליכה אחורנית".

משפט הפירוק הפרימרי

יהי V מרחב וקטורי ממימד סופי מעל שדה $\mathbb{F}$. יהי $T : V \to V$ אופרטור ליניארי כלשהו, עם פולינום מינימלי

$$m(x) = m_T(x) = f_1(x)^{\beta_1} f_2(x)^{\beta_2} \ldots f_t(x)^{\beta_k}$$

כאשר $f_i(x)$ אי־פריקים, ו־$\beta_i \geq 1$ שלמים.

נסמן

$$W_i = \ker\left(f_i^{\beta_i}(T)\right) = \left\{v \in V \ : \ f_i^{\beta_i}(T)(v) = \vec{0}\right\}$$

אזי $V = \bigoplus_{i=1}^{t} W_i$ פירוק למרחבים T־שמורים.

לפני שנוכיח את המשפט, מספר הערות.

1. אם T היה לכסין, כבר ראינו בעבר שיש דרך פשוטה לפרק את המרחב לסכום ישר ־ למרחבים העצמיים שלו. אך משפט זה מטפל גם במקרה בו האופרטור לא לכסין (ואז אין מספיק מרחבים עצמיים כדי לכסות את המרחב כולו), ונותן לנו תתי־מרחב אחרים שיהוו מועמדים לפירוק המרחב.

2. בדיקת שפיות: אילו T היה לכסין, הרי ש־$\beta_i = 1$ לכל i, כלומר $m_T(x) = f_1(x) f_2(x) \ldots f_t(x) = (x - \lambda_1)(x - \lambda_2) \ldots (x - \lambda_t)$ כאשר λ_i הם הע"ע של T, ואז $W_i = \ker(T - \lambda_i \mathbb{I})$ היה בדיוק המרחב העצמי המתאים ל־λ_i.

הוכחת משפט הפירוק הפרימרי

נרצה להשתמש בטענה 2 מהעבר, לפיה:

אם $P_i : V \to V$ ליניאריים כך ש־$\sum_{i=1}^{t} P_i = \mathbb{I}$, ו־$P_i P_j = 0$ עבור $i \neq j$, אז $V = \bigoplus_{i=1}^{t} \mathrm{Im} P_i$ ו־P_i הטלות. לכן מטרתנו תהיה להגדיר P_i שמקיימות את דרישות טענה 2 כך ש־$W_i = \mathrm{Im} P_i$.

נגדיר

$$q_i(x) = \frac{m_T(x)}{f_i^{\beta_i}(x)} = \prod_{k \neq i} f_k^{\beta_k}(x)$$

אזי, מתקיים ש־

$$\gcd\left(q_1(x),...,q_t(x)\right)=1$$

<u>מדוע? דוגמה להמחשת הרעיון</u>: נניח ש־$t=3$, אזי

$$q_1=f_2^{\beta_2}f_3^{\beta_3}$$
$$q_2=f_1^{\beta_1}f_3^{\beta_3}$$
$$q_3=f_1^{\beta_1}f_2^{\beta_2}$$

וקל לראות כעת שהגורם המשותף היחיד שלהם הוא 1.

<u>המשך ההוכחה</u>: מהעובדה שהזכרנו קודם, קיימים $h_i(x)\in\mathbb{F}[x]$ כך ש־

$$q_1(x)h_1(x)+q_2(x)h_2(x)+...+q_t(x)h_t(x)=\gcd\left(q_1(x),q_2(x),...,q_t(x)\right)=1$$

כלומר:

$$(*)\quad\sum_{i=1}^{t}h_iq_i(x)=1$$

נגדיר $P_i=h_iq_i(T)$, זהו אופרטור ליניארי (כי הצבה של אופרטור ליניארי T בפולינום נותנת אופרטור ליניארי). נראה ש־P_i הנ"ל מקיימים את הנדרש. התנאים שעלינו להראות שמתקיימים הם:

(1) $\sum_{i=1}^{t}P_i=\mathbb{I}$

(2) $P_iP_j=0$ עבור $i\neq j$

(3) $\mathrm{Im}P_i=W_i$

הוכחת (1): נציב T ב־$(*)$ ונקבל: $\sum_{i=1}^{t}\underbrace{h_iq_i(T)}_{P_i}=\mathbb{I}$.

הוכחת (2): נניח ש־$j\neq i$, נשים לב ש־$m_T(x)\,|\,h_iq_ih_jq_j(x)$, לכן קיים $g(x)$ כך ש־$h_iq_ih_jq_j(x)=m_T(x)g(x)$, כזכור $m_T(T)=0$ ולכן:

$$P_iP_j=\left(h_iq_ih_jq_j(T)\right)=g(T)m_T(T)=0$$

נשים לב שכעת, לאחר שהוכחנו את תכונות $(1),(2)$, אזי מטענה 2 אנו יודעים שה־P_i הן הטלות וש־$V=\bigoplus_{i=1}^{t}\mathrm{Im}P_i$. נותר להוכיח את (3).

הוכחת (3): בהכלה כפולה. כיוון ראשון ־ נניח ש־$w\in\mathrm{Im}P_i$, נראה ש־$w\in W_i=\ker f_i^{\beta_i}(T)$, נחשב:

$$f_i^{\beta_i}(T)(w)\underbrace{=}_{*}f_i^{\beta_i}(T)(P_i(w))\underbrace{=}_{**}f_i^{\beta_i}(T)(h_iq_i(T))(w)\underbrace{=}_{***}\underbrace{f_i^{\beta_i}q_i}_{m_T}h_i(T)(w)=\underbrace{m_T(T)}_{=0}h_i(T)=0$$

ולכן קיבלנו ש־$w\in\ker\left(f_i^{\beta_i}(T)\right)$.

הסבר המעברים: במעבר $*$ הסתמכנו על כך ש־P_i הטלה ולכן $P_i(w)=w$ אם $w\in\mathrm{Im}P_i$. במעבר $**$ הסתמכנו על הגדרת P_i בתור $h_iq_i(T)$, ומבעבר $***$ הסתמכנו על כך שפולינומים ב־T מתחלפים בהרכבה כאופרטורים.

כיוון שני ־ נניח ש־$w\in W_i=\ker\left(f_i^{\beta_i}(T)\right)$ ונראה ש־$w\in\mathrm{Im}P_i$. כזכור, אם נראה ש־$P_i(w)=w$ זה יגרור ש־$w\in\mathrm{Im}P_i$. אכן:

$$w=\mathbb{I}(w)\underbrace{=}_{(1)}\sum_kP_k(w)=\sum_kh_kq_k(T)(w)=P_i(w)$$

כאשר המעבר האחרון נובע מכך שעבור $k \neq i$ מופיע הגורם $f_i^{\beta_i}(x)$ ב־$q_k(x)$, וכמו כן נתון ש־$w \in \ker\left(f_i^{\beta_i}(x)\right)$ ולכן נקבל

$$h_k q_k(T)(w) = \text{something} \cdot f_i^{\beta_i}(T)(w) = 0$$

לסיום, נותר רק להסביר מדוע W_i הם T־שמורים. תזכורת: $\ker T$ הוא <u>תמיד</u> T־שמור. תרגיל: הראו ש־$\ker \psi(T)$ גם T־שמור לכל פולינום $\psi(x)$ (וזה יוכיח את הטענה גם במקרה של משפט הפירוק הפרימרי). ■

<u>הערה</u>: <u>**להוסיף הערה על התחלפות של T עם ההטלות P_i.**</u>

<u>משמעות המשפט בלשון של מטריצות מייצגות</u>

אם B_i בסיס של $W_i = \ker\left(f_i^{\beta_i}(T)\right)$, אז $B = B_1 \cup B_2 \cup ... \cup B_t$ הוא בסיס של V, ומתקיים:

$$[T]_B = \begin{pmatrix} A_1 & & & & \\ & A_2 & & & \\ & & A_3 & & \\ & & & \ddots & \\ & & & & A_t \end{pmatrix}$$

כאשר $A_1, A_2, ..., A_t$ הם בלוקים ריבועיים, כלומר קיבלנו מטריצה אלכסונית־בלוקים, ונשים לב שהבלוקים לא בהכרח מאותו גודל. למעשה מתקיים $A_i = [T|_{W_i}]_{B_i}$, כלומר כל בלוק הוא המטריצה המייצגת של הצמצום של האופרטור לתת המרחב המתאים.

<u>שאלה</u>: מהו גודל הבלוקים? שאלה שקולה: מהו $\dim_F W_i$?

<u>תשובה</u>: נניח שהפולינום האופייני של T הוא:

$$\mathbb{N} \ni \alpha_i \geq \beta_i \geq 1 \quad p_T(x) = f_1^{\alpha_1}(x) ... f_t^{\alpha_t}(x)$$

וכזכור הפולינום המינימלי הוא:

$$m_T(x) = f_1^{\beta_1}(x) ... f_t^{\beta_t}(x)$$

<u>אז:</u>

- הפולינום המינימלי של $T|_{W_i}$ הוא $f_i^{\beta_i}(x)$ (וודאו זאת על פי הגדרת הפולינום המינימלי - בדקו שהוא מתוקן, שהוא מאפס את $T|_{W_i}$, ושהוא בעל דרגה מינימלית).
- הפולינום האופייני של $T|_{W_i}$ הוא $f_i^{\alpha_i}(x)$ - זאת משום $m_{T|_{W_i}} | p_{T|_{W_i}}(x)$, ובנוסף ראינו בתרגול שמכפלת הפ"א של $T|_{W_i}$ היא בדיוק $p_T(x)$, ביחד עם העובדה שכל f_i אי־פריקים שונים זה מזה, התוצאה נובעת.

כעת נוכל לענות על השאלה ששאלנו: באופן כללי המעלה של פולינום אופייני של מטריצה היא בדיוק המימד של המרחב המתאים, לכן:

$$\dim W_i = \deg\left(p_{T|_{W_i}}\right) = \alpha_i \deg(f_i)$$

וזהו סדר הבלוק של A_i.

<u>מסקנה למקרה בו $\mathbb{F} = \mathbb{C}$</u>: נניח ש־V מרחב וקטורי סוף מימדי מעל $\mathbb{C}$, ו־$T : V \to V$ אופרטור ליניארי. אזי

$$p_T(x) = (x - \lambda_1)^{\alpha_1} (x - \lambda_2)^{\alpha_2} ... (x - \lambda_t)^{\alpha_t}$$

$$m_T(x) = (x-\lambda_1)^{\beta_1}(x-\lambda_2)^{\beta_2}...(x-\lambda_t)^{\beta_t}$$

כאשר $\lambda_i \in \mathbb{C}$ שונים ו־$1 \geq \beta_i \geq \alpha_i \ni \mathbb{N}$ (וכמובן אם $\beta_i = 1$ לכל i אז אנחנו במקרה הלכסין).

אז, ממשפט הפירוק הפרימרי נסיק כי $W_i = \ker\left((T-\lambda\mathbb{I})^{\beta_i}\right)$ (כאן $\dim W_i = \alpha_i$), ואם B_i בסיס של W_i, $B = B_1 \cup ... \cup B_t$ בסיס של V, אז

$$[T]_B = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_t \end{pmatrix}$$

כאשר $A_i = [T|_{W_i}]_{B_i}$ עם פולינום מינימלי $(x-\lambda_i)^{\beta_i}$, ולכן $(A_i - \lambda_i\mathbb{I})^{\beta_i} = 0_{\alpha_i\times\alpha_i}$.

<u>הגדרה</u>: $A \in M_n(\mathbb{F})$ נקראת <u>נילפוטנטית</u>, אם קיים m טבעי כך ש־$A^m = 0$.

בחזרה לענייננו: אם נסמן $N_i = A_i - \lambda_i\mathbb{I}$ אז קיבלנו ש־N_i נילפוטנטית.

<u>מסקנה</u>: לכל i יש מטריצה N_i נילפוטנטית מסדר $\alpha_i \times \alpha_i$ ומתקיים $A_i = N_i + \lambda_i\mathbb{I}$, כלומר

$$[T]_B = \begin{pmatrix} (A_1)_{\alpha_1\times\alpha_1} & & 0 \\ & \ddots & \\ 0 & & (A_t)_{\alpha_t\times\alpha_t} \end{pmatrix} = \begin{pmatrix} N_1+\lambda_1\mathbb{I} & & 0 \\ & \ddots & \\ 0 & & N_t+\lambda_t\mathbb{I} \end{pmatrix} = \underbrace{\begin{pmatrix} N_1 & & 0 \\ & \ddots & \\ 0 & & N_t \end{pmatrix}}_{N} + \underbrace{\begin{pmatrix} \lambda_1\mathbb{I} & & 0 \\ & \ddots & \\ 0 & & \lambda_t\mathbb{I} \end{pmatrix}}_{S}$$

ובכך פירקנו את המטריצה המייצגת $[T]_B$ לשתי מטריצות בלוקים: המטריצה הראשונה היא נילפוטנטית, כי כל בלוק שלה נילפוטנטי, ולכן נעלה את המטריצה בחזקת $\beta_1\beta_2...\beta_t$ ואז כל הבלוקים יתאפסו, והמטריצה השניה היא בעלת בלוקים סקלריים, ולכן היא אלכסונית!

נסמן ב־S, N את האופרטורים המתאימים למטריצה האלכסונית והמטריצה הנילפוטנטית שקיבלנו, אז:

$$T = N + S$$

כאשר N נילפוטנטי ו־S לכסין. לטענה זו קוראים <u>משפט ג'ורדן שבלייה (Jordan-Chevalley)</u>. המשפט אומר את הדבר הבא: אם V מרחב וקטורי סוף מימדי, מעל שדה $\mathbb{F}$, אז יש S לכסין ו־N נילפוטנטי המקיימים $NS = SN$ כך ש־$T = N + S$.

<u>הרצאה 9 - 30 לאפריל</u>

<u>תזכורת</u>: ראינו שאם $T : V \to V$ אופרטור ליניארי, כאשר V מ"ו סוף־מימדי <u>מעל $\mathbb{C}$</u>, אזי הפולינום המינימלי של T הוא

$$m_T(x) = (x-\lambda_1)^{\beta_1}...(x-\lambda_t)^{\beta_t}$$

(נזכיר שאם הפולינום $m_T(x)$ מתפרק לגורמים ליניאריים (כלומר $\beta_1 = ... = \beta_t = 1$) אזי T לכסין).

אזי $T = S + N$ כאשר S אופרטור לכסין, N אופרטור נילפוטנטי, N, S פולינומים ב־T ולכן בפרט $SN = NS$.

איך מצאנו אותם? ממשפט הפירוק הפרימרי - אם B_i בסיס של $W_i = \ker(T-\lambda_i\mathbb{I})^{\beta_i}$ אז $B = \bigcup B_i$ בסיס של V ו־

$$[T]_B = \begin{pmatrix} (A_1)_{\alpha_1\times\alpha_1} & & 0 \\ & \ddots & \\ 0 & & (A_t)_{\alpha_t\times\alpha_t} \end{pmatrix}$$

כאשר $A_i = [P_i|_{W_i}]_{B_i}$, נזכור שההטלות הן פולינומים ב־T, והפולינום המינימלי של A_i הוא $(x-\lambda_i)^{\beta_i}$ ולכן $N_i = A_i - \lambda_i\mathbb{I}$ נילפוטנטית. ולכן נקבל

$$[N]_B = \begin{pmatrix} N_1 & & 0 \\ & \ddots & \\ 0 & & N_t \end{pmatrix}$$

$$[S]_B = \underbrace{\begin{pmatrix} \lambda_1 \mathbb{I} & & 0 \\ & \ddots & \\ 0 & & \lambda_t \mathbb{I} \end{pmatrix}}_{S}$$

ובכך קיבלנו דרך למצוא את N, S.

<u>דרך אחרת למצוא את N, S</u>: נרשום את S בצורה

$$S = \sum_{i=1}^{t} \lambda_i P_i$$

כאשר P_i הטלה על W_i במקביל ל־$\sum_{k \neq i} W_k$, אז ממשפט הפירוק הספקטרלי S הוא לכסין. כעת נגדיר $N = T - S$, הסיבה שהוא אכן נילפוטנטי היא שהע"ע של T הם λ_i, ושהע"ע של S הם λ_i, ולכן הע"ע של N הם רק אפס, <u>ומעל $\mathbb{C}$</u> זה גורר ש־N נילפוטנטי כנדרש.

תוצאה זו מקרבת אותנו מאוד למטרת העל שלנו: מציאת צורת ג'ורדן של אופרטור, אנו רואים שכל אופרטור ניתן לכתיבה כסכום של אופרטור אלכסוני ואופרטור נילפוטנטי, ולכן המסקנה הנ"ל היא המוטיבציה שלנו למציאת <u>צורת ג'ורדן לאופרטורים נילפוטנטיים</u>.

<u>הגדרות</u>

1) <u>בלוק ג'ורדן</u> $J_m(\lambda)$ עבור $m \geq 1$ טבעי ו־$\lambda \in \mathbb{F}$ הינו המטריצה $m \times m$ הבאה:

$$J_m(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & & \\ \vdots & & \ddots & \ddots & \\ & & & \ddots & 1 \\ 0 & & & & \lambda \end{pmatrix}$$

2) נאמר שמטריצה $J \in M_n(\mathbb{F})$ היא <u>בצורת ג'ורדן</u> אם J היא מטריצת בלוקים אלכסונית כאשר כל בלוק הוא בלוק ג'ורדן:

$$J = \operatorname{diag}\left(J_{m_1}(\lambda_1), J_{m_2}(\lambda_2), ..., J_{m_k}(\lambda_k)\right)$$

<u>דוגמאות:</u>

$$J_2(0) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$J_3(0) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

<u>מטרה</u>: להראות שכל $A \in M_n(\mathbb{C})$ <u>דומה</u> לאיזושהי מטריצה J בצורת ג'ורדן. (זה דומה למה שראינו באלגברה א' ־ שכל מטריצה לכסינה דומה למטריצה אלכסונית). למעשה תוצאה זו מחלקת את כל מרחב המטריצות למחלקות של מטריצות הדומות זו לזו ־ בכך שיש להן אותה צורת ג'ורדן.

<u>דיון לא פורמלי</u>: עבור המקרה של $A \in M_2(\mathbb{C}), M_3(\mathbb{C})$ <u>נילפוטנטית</u>.

רוצים למצוא J בצורת ג'ורדן ו־P הפיכה כך ש־$P^{-1}AP = J$.

<u>עבור 2×2</u>: מאחר והמטריצה נילפוטנטית, הע"ע היחיד שלה הוא אפס, לכן יש שתי אפשרויות למטריצת ג'ורדן. המקרה הראשון (והלכסין) הוא

$$J = \begin{pmatrix} J_1(0) & \\ & J_1(0) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

וזהו המקרה הפחות מעניין. המקרה השני הוא

$$J=J_2\left(0\right)=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}$$

נסמן $P=\begin{pmatrix}\vdots & \vdots\\ \vec{u} & \vec{v}\\ \vdots & \vdots\end{pmatrix}_{2\times2}$ (תזכורת מאלגברה א' - P הפיכה $\iff$ וקטורי העמודה שלה $\vec{u},\vec{v}$ הם בת"ל). אזי $P^{-1}AP=J$ גורר $AP=PJ$. מתקיים:

$$PJ=\begin{pmatrix}\vdots & \vdots\\ \vec{u} & \vec{v}\\ \vdots & \vdots\end{pmatrix}_{2\times2}\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}=\begin{pmatrix}0 & \vdots\\ & \vec{u}\\ 0 & \vdots\end{pmatrix}$$

לכן קיבלנו $A\vec{u}=0$ ו־$A\vec{v}=\vec{u}$. מהתוצאה הראשונה נסיק

$$A\vec{u}=\left(A-0\mathbb{I}\right)\vec{u}=0\cdot\vec{u}$$

כלומר $\vec{u}$ וקטור עצמי של A עם ע"ע 0, כלומר $\vec{u}\in\ker A$.

לפיכך, אם נמצא $\vec{v}\in\ker\left(A^{2}\right)$ כך ש־$\vec{v}\notin\ker A$, ונגדיר $\vec{u}=A\vec{v}$ אזי $\vec{u}\in\ker A$, כי אם $A^{2}\vec{v}=\vec{0}$ אזי $A\left(A\vec{v}\right)=A\vec{u}=\vec{0}$ כלומר $\vec{u}\in\ker A$. אם נוכל בנוסף להבטיח ש־$\vec{v},\vec{u}=A\vec{v}$ בת"ל, הרי שתהליך הג'ירדון יצליח.

<u>עבור 3×3</u>:

נתעניין בתת־המקרה בו $J=J_3\left(0\right)$ (כי בשניים האחרים כבר טיפלנו), נסמן $P_{3\times3}=\begin{pmatrix}\vdots & \vdots & \vdots\\ \vec{u} & \vec{v} & \vec{w}\\ \vdots & \vdots & \vdots\end{pmatrix}$, אז מהחשבון $P^{-1}AP=J$ נסיק שוב ש־$AP=PJ$ ולכן:

$$PJ=\begin{pmatrix}0 & \vdots & \vdots\\ 0 & \vec{u} & \vec{v}\\ 0 & \vdots & \vdots\end{pmatrix}$$

כלומר: עלינו למצוא שרשרת של וקטורים בלתי תלויים, כך ש־$A\vec{u}=\vec{0}$, $A\vec{v}=\vec{u}$ ו־$A\vec{w}=\vec{u}$. כלומר נרצה

$$\vec{w}\in\ker\left(A^{3}\right)\setminus\ker\left(A^{2}\right),\quad\vec{v}=A\vec{w},\ \vec{u}=A\vec{v}=A^{2}\vec{w}$$

מי שיבוא לעזרתנו הוא משפט הפירוק הציקלי (שנלמד בהמשך). נזכיר שעד כה הדיון אינו פורמלי, אלא רק נועד להראות לנו את הכיוון אליו אנו הולכים.

<u>הגדרה</u>: יהי V מ"ו סוף־מימדי מעל $\mathbb{F}$, $T:V\to V$ אופרטור ליניארי ו־$v\in V$. נסמן ב־$\left\langle v\right\rangle =Z\left(v,T\right)$ את <u>תת־המרחב ה־T ציקלי הנוצר על ידי v</u> המוגדר באופן הבא:

$$\left\langle v\right\rangle =Z\left(v,T\right)=\mathrm{span}_{\mathbb{F}}\left\{ v,T\left(v\right),T^{2}\left(v\right),...\right\}$$

נשים לב שזהו מרחב ממימד סופי, בגלל משפט קיילי המילטון.

<u>דוגמה</u>: אם $T=\mathbb{I}$ אז $\left\langle \vec{v}\right\rangle =\mathrm{span}_{\mathbb{F}}\left\{ \vec{v}\right\}$.

דוגמה נוספת: אם $T:\mathbb{R}^2\to\mathbb{R}^2$ הוא האופרטור $T\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}y\\0\end{pmatrix}$, ו־$v=\begin{pmatrix}0\\1\end{pmatrix}$, אזי:

$$\langle v\rangle=\text{span}_{\mathbb{F}}\left\{\underbrace{v}_{\begin{pmatrix}1\\0\end{pmatrix}},\underbrace{T\left(v\right)}_{\begin{pmatrix}0\\1\end{pmatrix}}\right\}=\mathbb{R}^2$$

כלומר במקרה זה המרחב הציקלי כבר נותן לנו את כל המרחב.

משפט הפירוק הציקלי (לאופרטור נילפוטנטי ־ הוא נכון גם לאופרטור כללי): יהי $N:V\to V$ אופרטור נילפוטנטי עם פולינום מינימלי $m_N\left(x\right)=x^r$. אזי קיימים $r=r_1\geq r_2\geq...\geq r_t\geq 1$ ווקטורים $v_1,v_2,...,v_t\in V$, כך ש־

$$V=\langle v_1\rangle\oplus\langle v_2\rangle\oplus...\oplus\langle v_t\rangle$$

כך שהפולינום המינימלי של $N|_{\langle v_i\rangle}$ הוא x^{r_i}.

מסקנה: קיום צורת ג'ורדן לאופרטור נילפוטנטי. מדוע? כי משפט הפירוק הציקלי למעשה אומר לנו שקיים בסיס

$$B=\left\{N^{r_1-1}v_1,...,Nv_1,v_1,N^{r_2-1}v_2,...,Nv_2,v_2,...,N^{r_t-1}v_t,...,Nv_t,v_t\right\}$$

לכן

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

הרצאה 10 ־ 1 למאי

תזכורת: משפט הפירוק הציקלי (לאופרטור נילפוטנטי)

יהי $N:V\to V$ אופרטור נילפוטנטי עם פולינום מינימלי $m_N\left(x\right)=x^r$, אז קיימים $r=r_1\geq r_2\geq...\geq r_t\geq 1$ ווקטורים $v_1,v_2,...,v_t\in V$ כך ש:

$$V=\langle v_1\rangle\oplus\langle v_2\rangle\oplus...\oplus\langle v_t\rangle$$

והפולינום המינימלי של $N|_{\langle v_i\rangle}$ הוא x^{r_i}.

נזכיר מהו המרחב הציקלי $\langle v\rangle$:

$$\langle v\rangle=\text{span}\left\{v,Nv,N^2v,...\right\}$$

מסקנה ־ קיום צורת ג'ורדן לאופרטור נילפוטנטי

קיים בסיס B מהצורה הבאה:

$$B=\left\{N^{r_1-1}v_1,...,N^2v_1,Nv_1,v_1,N^{r_2-1}v_2,...,v_2,...,N^{r_t-1}v_t,...,v_t\right\}$$

נשים לב שאם נפעיל את N על הוקטור הראשון, נקבל $N^{r_1-1}v_1$, שהוא אפס (ומכאן שנקבל עמודת אפסים במטריצה), אם נפעיל את N על האיבר השני, נקבל $N^{r_1-2}v_1$, שהוא האיבר הראשון, ולכן נקבל את העמודה $\begin{pmatrix}1\\0\\\vdots\\\vdots\\0\end{pmatrix}$, וכן הלאה, וכך מתקבל בלוק ג'ורדן $J_{r_1}(0)=\begin{pmatrix}0&1&&&0\\0&0&1&&\\\vdots&\vdots&0&\ddots&\\&&&\ddots&1\\0&&&&0\end{pmatrix}$. בעצם הסתמכנו על כך שהצגה מטריצית של אופרטור בבסיס מסוים נקבעת על פי פעולתו על איברי הבסיס, ובמקרה שלנו

$$N^{r_1}v_1=\vec{0},\ N\left(N^{r_2-2}\right)v_1=N^{r_1-1}v_1...$$

בצורה דומה נקבל בלוקים נוספים, ונוכל להציג את האופרטור N (בבסיס B) בצורת ג'ורדן:

$$[N]_B=\begin{pmatrix}J_{r_1}(0)&&&&\\&J_{r_2}(0)&&&\\&&\ddots&&\\&&&\ddots&\\&&&&J_{r_t}(0)\end{pmatrix}$$

נשים לב, שהוקטור $N^{r_1-1}v_1$ הוא בגרעין של N, כלומר $N^{r_1-1}v_1\in\ker N$, ולכן זהו ו"ע של N עם ע"ע אפס, בצורה דומה $N^{r_2-1}v_2$ הוא ו"ע של N, וממשיכים כך עד $N^{r_t-1}v_t$, לכן ישנם t ו"ע בת"ל של הע"ע 0.

<u>בפרט</u>: מספר הבלוקים בצורת ג'ורדן = t = מספר הו"ע הבת"ל של הע"ע 0 = הר"ג של 0 = $\dim\ker(N)$.

גודל הבלוק המקסימלי = $r=r_1$ = מעלת הפולינום המינימלי של N.

סכום גדלי הבלוקים = $\dim V$ = מעלת הפולינום האופייני של N = הר"א של 0.

<u>תרגיל</u>

נתון אופרטור ליניארי נילפוטנטי $N:\mathbb{R}^6\to\mathbb{R}^6$ עם פולינום מינימלי $m_N(x)=x^3$ וכך ש-$\dim\mathrm{Im}N=4$. מהי צורת ג'ורדן של N? וכיצד ייראה בסיס מג'רדן?

<u>פתרון</u>

גודל הבלוק המקסימלי הוא 3 (מעלת הפולינום המינימלי), מספר הבלוקים הוא $\dim\ker N$, אותו נקבל ממשפט המימדים:

$$\dim\ker N=\dim V-\dim\mathrm{Im}N=6-4=2$$

סה"כ צורת ג'ורדן היא מטריצה 6×6 (כי זה המימד של V), ולכן נובע שיש 2 בלוקים מגודל 3×3 כל אחד:

$$J=\begin{pmatrix}J_3(0)&\\&J_3(0)\end{pmatrix}$$

בסיס מג'רדן ייראה כך:

$$B=\left\{N^2v_1,Nv_1,v_1,N^2v_2,Nv_2,v_2\right\}$$

כאשר קיומם של v_1,v_2 מובטח ממשפט הפירוק הציקלי, והם מקיימים $V=\langle v_1\rangle\oplus\langle v_2\rangle$. כמובן שתשובה זו לא מספקת ועלינו ממש למצוא אותם (נראה בתרגול 5 דוגמאות לכך). נסביר כעת כיצד מוצאים אותם בתכלס: אנו רוצים שני וקטורים בלתי תלויים (כי

שניהם בבסיס), ומאחר וכל אחד מהם יוצר שרשרת שאורכה 3 נדרוש ששניהם יהיו ב־$\ker\left(N^{3}\right)$ כך שאינם שייכים ל־$\ker\left(N^{2}\right)$ (כי אם הם היו ב־$\ker\left(N^{2}\right)$ השרשרת שהם יוצרים הייתה מתכווצת, היא לא הייתה מאורך 3).

דוגמה אחרת: נניח ש־

$$J=\begin{pmatrix}J_{5}\left(0\right) & & & 0\\ & J_{3}\left(0\right) & 0\\ & 0 & J_{2}\left(0\right)\\ 0 & & & J_{1}\left(0\right)\end{pmatrix}$$

איך ייראה בסיס מג'רדן? המרחב של כל הוקטורים העצמיים של N הוא $\ker N$, נשים לב שמתקיים

$$\ker\left(N\right)\subsetneqq\ker\left(N^{2}\right)\subsetneqq\ker\left(N^{3}\right)\subsetneqq\ker\left(N^{4}\right)\subsetneqq\ker\left(N^{5}\right)=V$$

כאשר הסימן $\subsetneqq$ מייצג הכלה ממש.

נסביר: $\ker\left(N\right)$ הוא המרחב העצמי של הע"ע 0, ואילו N היה לכסין היה מתקיים $\ker\left(N\right)=V$, אבל הוא לא לכסין, כי אם היה לכסין צורת ג'ורדן שלו הייתה מכילה בלוקים $J_{1}\left(0\right)$ בלבד.

הסיבה לכך ש־$\ker\left(N^{5}\right)=V$ היא שמהנתון הבלוק המקסימלי הוא 5×5 ולכן הפולינום המינימלי של N הוא x^{5}.

כעת, עבור $J_{5}\left(0\right)$ נחפש שרשרת באורך 5, ולכן נדרוש $v_{1}\in\ker\left(N^{5}\right)$ אבל $v_{1}\notin\ker\left(N^{4}\right)$ (ולכן הוא גם לא בשאר הגרעינים). זה ייתן לנו שרשרת מאורך 5

$$N^{4}v_{1},N^{3},v_{1},...,v_{1}$$

כעת, עבור $J_{3}\left(0\right)$ נחפש שרשרת באורך 3, לכן נחפש $v_{2}\in\ker\left(N^{3}\right)$ כך ש־$v_{2}\notin\ker\left(N^{2}\right)$, כך ש־$v_{2}$ בת"ל בוקטורים הקודמים (v_{1}) וכן הלאה.

נראה עוד דוגמאות מעניינות בתרגול 5, כעת עלינו להשלים חוב: הוכחת משפט הפירוק הציקלי, לשם כך, לרוב משתמשים בהוכחה באינדוקציה (על המימד) כדי לטפל במימדים הקטנים יותר ומהם להסיק את נכונות המשפט עבור המרחב כולו. אנו נראה כלי אחר כדי להוכיח את משפט הפירוק הציקלי ־ מרחבי מנה.

מרחבי מנה

מוטיבציה: באלגברה א' למדנו על מרחבים ועל תתי־מרחבים, וראינו שתמיד תתי המרחבים עוברים בראשית (כי איבר האפס תמיד שייך לכל תת־מרחב), לכן למשל ב־$\mathbb{R}^2$ ראינו שכל תת־מרחב ממימד 1 הוא ישר העובר דרך הראשית, וב־$\mathbb{R}^3$ כל תת־מרחב ממימד 2 (למשל) הוא מישור העובר דרך הראשית. נשאל את השאלה הבאה: מדוע אי אפשר לדבר גם על "על מישור", שהוא בעצם מעין תת־מרחב שלא עובר דרך הראשית, ולכן למעשה מהווה הזזה של תת־מרחב שכן עובר דרך הראשית? הרעיון של מרחבי מנה מטפל בדיוק בסוגיה זו.

יהי V מרחב וקטורי ממימד סופי, ו־$W \subset V$ תת־מרחב. מגדירים יחס שקילות $v \sim u \iff v - u \in W$. אנו רואים שהוא מקיים את כל הדרישות של יחס שקילות:

1) רלפקסיביות: $v \sim v$.

2) סימטריות: $v \sim u$ גורר $u \sim v$.

3) טרנזיטיביות: $v \sim u$ וגם $u \sim z$ גורר $v \sim z$.

מחלקות השקילות הן:

$$\begin{aligned}[v] &= \{u \in V \,:\, u \sim v\} = \\ &= \{u \in V \,:\, u - v \in W\} = \\ &= \{u \in V \,:\, \exists w \in W \;\; u - v = w\} = \\ &= \{u \in V \,:\, \exists w \in W \;\; u = v + w\} = v + W\end{aligned}$$

ולאוסף $v + W$ קוראים <u>מרחב אפיני</u> או <u>קוסט</u>. שימו לב: באופן כללי $v + W$ לא תת־מרחב. (אם $v = \vec{0}$ או באופן כללי יותר אם $v \in W$ אז $v + W = W$ ואז הוא כן תת־מרחב, אבל אחרת הוא לא תת־מרחב).

<u>דוגמה</u>: $V = \mathbb{R}^2$, $W = \operatorname{span}\left\{\begin{pmatrix}1\\1\end{pmatrix}\right\}$, זהו ישר העובר דרך ראשית הצירים, וכל הקוסטים האחרים הם ישרים המקבילים לו, ומהווים הזזה שלו בוקטור $v = \begin{pmatrix}a\\b\end{pmatrix}$ כלשהו.

<u>המשך הגדרות</u>:

<u>הגדרה</u>: <u>קבוצת המנה</u> היא קבוצת מחלקות השקילות (קוסטים) ומסומנת על ידי ${}^{V}\!/_{W}$:

$${}^{V}\!/_{W} = \{v + W \,:\, v \in V\}$$

ניתן לה מבנה של מרחב וקטורי, לשם כך עלינו להגדיר פעולות של חיבור איברים במרחב המנה וכפל בסקלר. לכן נגדיר חיבור וקטורים:

$$(v + W) + (u + W) = (v + u) + W$$

חשוב לשים לב מה כל סימן $+$ מייצג, סימן הפלוס בביטויים $(v + W)$ ו־$(u + W)$ הוא רק סימון (של קוסט), לעומת זאת, סימן הפלוס ביניהם הוא הסימן לחיבור קוסטים שהגדרנו כרגע. בביטוי $(v + u) + W$ סימן הפלוס הראשון מייצג חיבור וקטורים רגיל ב־V, וכמקודם סימן הפלוס השני הוא רק סימון של קוסט.

וכעת נגדיר כפל בסקלר:

$$\alpha \cdot (v + W) = \alpha \cdot v + W$$

יש להראות מדוע החיבור והכפל בסקלר <u>מוגדרים היטב</u>, כלומר שאין תלות בנציג.

<u>תשובה (עבור כפל בסקלר, המקרה השני מושאר כתרגיל)</u>: נניח ש־$v + W = u + W$ (כלומר v, u נציגים שונים של אותו קוסט), ו־α סקלר. כלומר $v - u \in W$, אבל W תת־מרחב ולכן $\alpha\,(v - u) \in W$, כלומר $\alpha v - \alpha u \in W$, ולכן $\alpha v + W = \alpha u + W$, כנדרש.

כעת, לא קשה להראות שעם פעולות אלו ${}^{V}\!/_{W}$ הוא מרחב וקטורי, זו לא תוצאה מפתיעה כי כל הנימוקים נובעים מכך ש־V עצמו הוא מרחב וקטורי (בדקו זאת כתרגיל!). בפרט, וקטור האפס של מרחב המנה הוא הקוסט W.

דוגמה: $V=\mathbb{R}^3$ ו־$W=\operatorname{span}\left\{\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix}\right\}$ (המישור x,y). מהו V/W? אנו מבינים שהקוסטים השונים הם מישורים המקבילים למישור זה, ומה שמאפיין כל אחד מהם הוא ערך ה־z שלו, לכן למעשה $V/W\cong\operatorname{span}\left\{\begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$. זהו טיעון גיאומטרי ולא ריגורוזי, אבל כעת נראה משפט שמסביר סוגיה זו.

משפט האיזומופיזם: אם $T:V\to V$ אופרטור ליניארי, אז $V/\ker T\cong\operatorname{Im}T$.

הוכחה: נסמן $K=\ker T$, נגדיר $S:V/K\to\operatorname{Im}T$ באופן הבא:

$$S\left(v+K\right)=T\left(v\right)$$

מאחר והארגומנט של S מוגדר בעזרת נציגים, עלינו להראות שהיא מוגדרת היטב, כמו כן נראה שהיא העתקה ליניארית (צריך, כי אנו עוסקים באיזומורפיזם בין מרחבים וקטוריים, ולכן אנחנו רוצים לשמור על מבנה של מרחב וקטורי), חח"ע ועל.

מוגדרת היטב: ניקח $v+K=u+K$ (כלומר שני נציגים של אותה מחלקה), עלינו להראות ש־$S\left(v+K\right)=S\left(u+K\right)$ כלומר להראות ש־$T\left(v\right)=T\left(u\right)$. נתון ש־$v-u\in K=\ker T$ ולכן $T\left(v-u\right)=0$ $\Leftarrow$מליניאריות נקבל $T\left(v\right)=T\left(u\right)$.

העתקה ליניארית: נובע מיידית מכך ש־T ליניארית.

חח"ע ועל ־ מושאר כתרגיל.

לסיכום: קיבלנו ש־$V/K\cong\operatorname{Im}T$.

כעת נחזור לדוגמה הקודמת: $V=\mathbb{R}^3$, נגדיר $T:\mathbb{R}^3\to\operatorname{span}\left\{\begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$ באופן הבא $T\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\0\\z\end{pmatrix}$. ברור שזו העתקה ליניארית, חח"ע ועל, עם גרעין = המישור $x,y=W$. ולכן מהמשפט מקבלים ש־$V/W\cong\mathbb{R}^1$ ($\mathbb{R}^1$ = ציר ה־z).

דוגמה נוספת: נסתכל על $V=\mathbb{R}\left[x\right]$, ו־$W=\left\langle x^2+1\right\rangle=\left\{\left(x^2+1\right)\cdot p\left(x\right)\ :\ p\left(x\right)\in\mathbb{R}\left[x\right]\right\}$ = המרחב הציקלי ביחס לאופרטור $p\left(x\right)\longmapsto xp\left(x\right)$. מהו מרחב המנה V/W? נראה ש־$\mathbb{R}\left[x\right]/\left\langle x^2+1\right\rangle=V/W\cong\mathbb{C}$.

נגדיר $T:\mathbb{R}\left[x\right]\to\mathbb{C}$ באופן הבא: $T\left(p\left(x\right)\right)=p\left(i\right)$ (וודאו שזו העתקה ליניארית), היא על (כי ניתן לקבל כל מספר $a+bi$ בעזרת הפולינום $a+bx$), ובנוסף מתקיים $\ker T=\left\langle x^2+1\right\rangle$.

לכן ממשפט האיזומורפיזם נקבל $\mathbb{R}\left[x\right]/\left\langle x^2+1\right\rangle\cong\mathbb{C}$. שימו לב שהמרחבים הנ"ל מוגדרים מעל $\mathbb{R}$.

בשיעור הבא נראה ש־$\dim V/W=\dim V-\dim W$, ותכונות נוספות, ואף נוכיח את משפט הפירוק הציקלי.

הרצאה 11 ־ 7 למאי

תזכורת ־ מרחבי מנה

אמרנו שאם יש לנו מ"ו V, $W\subset V$ ת"מ, אז הגדרנו יחס שקילות $\vec{v}\sim\vec{u}$ אמ"ם $v-u\in W$. בכך לקחנו וקטורים השונים זה מזה, אך הפכנו אותם ל"אותו דבר", במובן שאם הם נבדלים במשהו ב־W אז מבחינתנו הם אותו דבר. זה דומה לשעון מחוגים, בו השעה 1 והשעה 13 הן אותו דבר מבחינתנו, למרות ש־13 ו־1 שונים מספרית זה מזה.

ראינו שקבוצת המנה

$$V/W=\left\{\left[v\right]\right\}_{v\in V}=\left\{v+W\ :\ v\in V\right\}$$

כאשר הסימון $v+W$ מייצג קוסט. בשלב הזה לא הסתפקנו בכך ש־V/W היא קבוצה, אלא רצינו מבנה של מרחב וקטורי, ולכן הגדרנו עליה פעולת חיבור קוסטים:

$$\left(v+W\right)+\left(u+W\right)=v+u+W$$

ופעולת כפל בסקלר:

$$\alpha\left(v+W\right)=\alpha v+W$$

ביחד עם פעולות אלא, קיבלנו ש־V/W הוא מרחב וקטורי (הוא למעשה ירש את הפעולות שהגיעו מ־V).

ראינו את המשפט: אם $T : V \to V$ העתקה ליניארית, אז $V/\ker T \cong \mathrm{Im}T$. זה נותן לנו דרך קלה יחסית להבין מרחבים וקטוריים, מאחר וכל שעלינו לעשות הוא לממש אותם כגרעין של איזושהי העתקה. כעת נראה כמה תכונות חדשות.

משפט: $\dim V/W = \dim V - \dim W$.

הוכחה: נסתכל על העתקה $T : V \to V/W$, המוגדרת על ידי $T(v) = v + W$. קל לראות שהיא ליניארית (ודאו זאת!), $\mathrm{Im}T$ היא כל המרחב V/W (כי לכל קוסט ב־V/W יש וקטור ב־V שיצר אותו) , וכמו כן $\ker T = W$. ולכן ממשפט המימדים

$$\dim V = \dim \ker T + \dim \mathrm{Im}T = \dim W + \dim V/W$$

ולכן מהעברת אגפים התוצאה נובעת:

$$\boxed{\dim V/W = \dim V - \dim W}$$

טענה:

1) אם $V = W \bigoplus U$, אז ממשפט המימדים אנו יודעים כי $\dim V = \dim W + \dim U$, ולכן $\dim U = \dim V - \dim W = \dim V/W$, ומכאן ש־$V/W \cong U$ (הם איזומורפיים כי יש להם אותו מימד, ראינו באלגברה א' שאם V_1, V_2 מרחבים וקטוריים מעל $\mathbb{F}$ בעלי מימד שווה, אז $V_1 \cong V_2$).

הערת העשרה: למעשה אנו יודעים מעבר לכך ־ הם איזומורפיים באופן קנוני ־ כלומר באופן שלא תלוי בנציג, "טבעי". כלומר הוא ניתן על ידי $T : U \to V/W$ המוגדר על ידי $T(u) = u + W$, כאשר אין תלות בנציג בהגדרה זו.

2)

א) אם $v_1 + W, v_2 + W, ..., v_k + W$ בת"ל ב־V/W, אזי $v_1, ..., v_k$ בת"ל ב־V.

ב) יתר על כן, אם $v_1 + W, v_2 + W, ..., v_k + W$ הם בסיס ל־V/W ו־$w_1, ..., w_m$ בסיס של W, אז $\{v_1, ..., v_k, w_1, ..., w_m\}$ בסיס של V.

הוכחה

א) בשלילה, אם יש α_i לא כולם 0, כך ש־$\alpha_1 v_1 + \alpha v_2 + ... + \alpha_k v_k = \vec{0}$ ב־V, אז בפרט נקבל במרחב המנה ש־

$$\alpha_1 v_1 + ... + \alpha_k v_k + W = \vec{0} + W = W$$

אבל נזכור כיצד הגדרנו חיבור קוסטים וכפל של סקלר בקוסט, ונקבל:

$$\alpha_1 (v_1 + W) + \alpha_2 (v_2 + W) + ... + \alpha_k (v_k + W) = W$$

אבל α_i לא כולם אפס, וזה אומר ש־$v_i + W$ תלויים, בסתירה לנתון.

ב) נשים לב שמהנתון ו"משפט המימדים למנה" שראינו קודם, מתקבל

$$\dim V = \dim V/W + \dim W = k + m$$

אם נדע שבקבוצה $B = \{v_1, ..., v_k, w_1, ..., w_m\}$ יש בדיוק $k + m$ איברים, אזי נדע שכדי להראות ש־B בסיס מספיק להראות שהקבוצה B פורשת את V. מדוע ישנם $k + m$ איברים בקבוצה B? כלומר, כיצד אנחנו יודעים שאין חזרות על איברים בקבוצה זו? ראשית, מובן שכל ה־w_i שונים זה מזה כי נתון שהם מהווים בסיס למרחב W. כמו כן, מובן שכל אחד מה־v_iים שונה מכל אחד מה־w_iים, כי אם למשל v_1 היה ב־W אז $v_1 + W$ היא בעצם קוסט האפס, אבל הקוסטים הנתונים מהווים בסיס ולכן לא יכול להיות בהם איבר האפס. לסיום, נותר רק להסביר מדוע כל ה־v_i שונים זה מזה, אבל זה שוב ברור כי אם למשל שניים מהם זהים, גם הקוסטים המתקבלים מהם יהיו זהים, בסתירה לנתון שהקוסטים מהווים בסיס במרחב המנה.

לכן יש בדיוק $k+m$ איברים בקבוצה B, ונותר להראות שהיא פורשת את V. אכן, בהינתן $v\in V$ נביט על הקוסט $v+W\in {}^{V}/_{W}$. מהנתון קיימים $\alpha_1,...,\alpha_k$ כך ש־

$$v+W=\alpha_1\left(v_1+W\right)+...+\alpha_k\left(v_k+W\right)=\sum_{i=1}^{k}\alpha_i v_i+W$$

כאשר במעבר האחרון השתמשנו בהגדרה של חיבור וכפל בסקלר במרחב המנה. המשמעות של שוויון שני קוסטים עם נציגים שונים, היא שהפרש הנציגים ב־W, מכאן קיבלנו

$$v-\sum_{i=1}^{k}\alpha_i v_i\in W$$

ולכן קיימים β_i כך ש־

$$v-\sum_{i=1}^{k}\alpha_i v_i=\beta_1 w_1+...+\beta_m w_m$$

ולכן

$$v=\alpha_1 v_1+...+\alpha_k v_k+\beta_1 w_1+...+\beta_m w_m$$

ובכך קיבלנו ש־v הוא צירוף ליניארי של איברי B ולכן B פורש את V כנדרש.

3) נניח ש־$W\subset V$ הוא תת־מרחב T שמור עבור איזשהו אופרטור $T:V\to V$. אזי נסמן $\overline{T}:{}^{V}/_{W}\to {}^{V}/_{W}$ המקיים

$$\overline{T}\left(v+w\right)=T\left(v\right)+W$$

לאופרטור $\overline{T}$ נקרא <u>האופרטור המושרה על ידי T על מרחב המנה ${}^{V}/_{W}$</u>.

$\overline{T}$ כמוגדר לעיל הוא אופרטור ליניארי, מוגדר היטב (לא תלוי בנציג), ומתקיים:

$$\ker\overline{T}=\left\{v+W\ :\ T\left(v\right)\in W\right\}$$
$$\mathrm{Im}\overline{T}=\left\{v+W\ :\ v\in\mathrm{Im}\left(T\right)\right\}$$

נוכיח שהוא אכן מוגדר היטב: נניח שיש לנו שני נציגים של אותו קוסט, כלומר $v+W=u+W$, נרצה להראות שתחת ההעתקה $\overline{T}$ הם נשלחים לאותו מקום. מתקיים

$$\overline{T}\left(v+W\right)=T\left(v\right)+W$$
$$\overline{T}\left(u+W\right)=T\left(u\right)+W$$

מהנתון ש־u,v נציגים של אותה מחלקה, נסיק ש־$v-u\in W$. אבל W הוא T־שמור, ולכן $T\left(v-u\right)\in W$ ומהליניאריות של T נקבל ש־$T\left(v\right)-T\left(u\right)\in W$, כלומר $T\left(v\right)+W=T\left(u\right)+W \Leftarrow \overline{T}\left(u+W\right)=\overline{T}\left(v+W\right) \Leftarrow \overline{T}$ מוגדר היטב.

ליניאריות $\overline{T}$ נובעת מליניאריות T, בדקו זאת!

<u>הרצאה 12 ־ 8 למאי</u>

תזכורת: בשיעור הקודם הראינו טענה:

$(*)$ אם V מ"ו ו־$W\subset V$ תת־מרחב, ו־$\{w_1,...,w_m\}$ בסיס ל־W ו־$\{v_1+W,v_2+W,...,v_k+W\}$ בסיס ל־${}^{V}/_{W}$, אזי $\{w_1,...,w_m,v_1,...,v_k\}$ בסיס ל־V.

באופן שקול, אם נסמן ב־$\overline{u_1}$ את הקוסט $u_1 + W$, ניתן לכתוב את הטענה באופן הבא: אם $V/W = \overline{u_1} \oplus \overline{u_2} \oplus ... \oplus \overline{u_k}$ אז $V = W \oplus (u_1 \oplus u_2 \oplus ... \oplus u_k)$.

$(**)$ אם $T : V \to V$ אופרטור ליניארי ו־$W \subset V$ הוא T־שמור אז $\overline{T} : V/W \to V/W$ המוגדר באופן הבא:

$$\overline{T}\left(v + w\right) = T\left(v\right) + W$$

הוא אופרטור ליניארי מוגדר היטב.

<u>הערה</u>: מתקיים $m_{\overline{T}}\left(x\right) | m_T\left(x\right)$ **(להוסיף נימוק)**.

תזכורת: בהינתן אופרטור $N : V \to V$ ו־$v \in V$, הגדרנו את המרחב ה־N־ציקלי הנפרש על ידי v באופן הבא:

$$\langle v \rangle = \text{span}\left\{v, Nv, N^2 v, ...\right\}$$

חשוב לשים לב שרשימת הוקטורים הללו היא סופית, ולכן המרחב הנפרש על ידיה הוא ממימד סופי (הסבר: משפט קיילי־המילטון).

נזכיר כעת את משפט הפירוק הציקלי.

<u>**משפט הפירוק הציקלי (לאופרטור נילפוטנטי)**</u>

יהי $N : V \to V$ אופרטור נילפוטנטי עם פולינום מינימלי $m_N\left(x\right) = x^r$ (לעיתים אומרים שהאופרטור הוא "נילפוטנטי מאינדקס r"), אז קיימים $r = r_1 \geq r_2 \geq ... \geq r_t \geq 1$ ווקטורים $v_1, v_2, ..., v_t \in V$ כך ש:

$$V = \langle v_1 \rangle \oplus \langle v_2 \rangle \oplus ... \oplus \langle v_t \rangle$$

והפולינום המינימלי של $N|_{\langle v_i \rangle}$ הוא x^{r_i} (כלומר $N|_{\langle v_i \rangle}$ נילפוטנטי מאינדקס r_i).

<u>הוכחה</u>: נוכיח באינדוקציה על $\dim V$.

<u>בסיס</u>: עבור $\dim V = 1$ הטענה מתקיימת באופן טריוויאלי.

<u>הנחת האינדוקציה</u>: אם $\dim V < n$ ו־$N : V \to V$ נילפוטנטי מאינדקס r, אז קיימים $r = r_1 \geq r_2 \geq ... \geq r_t \geq 1$ ווקטורים $v_1, v_2, ..., v_t \in V$ כנדרש במשפט.

<u>צעד האינדוקציה</u>: נניח ש־$\dim V = n$ ויהי $N : V \to V$ נילפוטנטי מאינדקס r. מהנתון, אם N נילפוטנטי מאינדקס r אז $N^r \equiv 0$, אבל $N^{r-1} \neq 0$, כלומר קיים $v_1 \in V$ כך ש־$N^{r-1}\left(v_1\right) \neq \vec{0}$, כלומר $v_1 \notin \ker N^{r-1}$. נגדיר $W = \langle v_1 \rangle$, נשים לב ש־$N|_{\langle v_1 \rangle}$ נילפוטנטי מאינדקס $r_1 = r$ (כי v_1 מתאפס רק לאחר r הפעלות של N עליו). כמו כן, מ"משפט המימדים למנה" ומכך ש־$\dim W \geq 1$ נקבל $\dim V < \dim\left(V/W\right) = \dim V - \underbrace{\dim W}_{\geq 1}$. בנוסף W הוא N־שמור, ולכן נוכל להביט באופרטור $\overline{N} : V/W \to V/W$ שכזכור מוגדר באופן הבא

$$\overline{N}\left(v + w\right) = N\left(v\right) + W$$

אז $\overline{N}$ נילפוטנטי עם פולינום מינימלי $m_{\overline{N}}\left(x\right) = x^{r_2}$ כאשר $r_2 \leq r_1$. לפי הנחת האינדוקציה קיימים $r_2 \geq r_3 \geq ... \geq r_t$ ווקטורים $\overline{v_2}, \overline{v_3}, ..., \overline{v_t} \in V/W$ (כאשר השתמשנו בסימון חדש לקוסט: $\overline{v_i} = v_i + W$) כך ש:

$$V/W = \langle \overline{v_2} \rangle \oplus \langle \overline{v_3} \rangle \oplus ... \oplus \langle \overline{v_t} \rangle$$

כאשר הפולינום המינימלי של $\overline{N}|_{\langle \overline{v_i} \rangle}$ הוא x^{r_i}. בפרט, $\overline{N}^{r_i}\left(\overline{v_i}\right) = 0_{V/W} = W$ ו־r_i הוא המינימלי הנ"ל, כלומר $N^{r_i}\left(v_i\right) + W = W = 0_{V/W}$ ו־r_i הוא המינימלי כנ"ל, כלומר $N^{r_i}\left(v_i\right) \in W$ ו־r_i הוא המינימלי כנ"ל. אבל אנחנו רוצים יותר מזה, המטרה שלנו: אנו רוצים לבחור נציג v_i' לקוסט $v_i + W$ כך שיתקיים $N^{r_i}\left(v_i'\right) = \vec{0}$ (ו־r_i המינימלי כנ"ל). כלומר אנחנו לא מסתפקים בקבלת וקטור ב־W, אלא רוצים שהוא יהיה וקטור האפס.

בבירור, $N^{r_i - 1}\left(v_i'\right) \neq \vec{0}$, ולכן זה "נציג טוב" כי עבורו מתקיים ש־$N|_{\langle v_i \rangle}$ בעל פולינום מינימלי x^{r_i} כדרוש. בהינתן שמצאנו נציגים טובים לכל $2 \leq i \leq t$ אז סיימנו את ההוכחה ־ כי מהטענה $(*)$ בתזכורת שבתחילת השיעור נובע ש:

$$V = \underbrace{\langle v_1 \rangle}_{W} \oplus \langle v_2' \rangle \oplus \langle v_3' \rangle \oplus ... \oplus \langle v_t' \rangle$$

כעת, בלי הגבלת הכלליות נראה איך למצוא נציג טוב עבור $\langle\overline{v_2}\rangle$ (בצורה דומה מוצאים עבור שאר המרחבים). ראינו ש־$N^{r_2}(v_2)\in$ $W=\langle v_1\rangle=\operatorname{span}\left\{v_1,Nv_1,...,N^{r_1-1}v_1\right\}$, ו־$r_2$ הוא המינימלי כנ"ל. לכן

$$N^{r_2}(v_2)=f(N)(v_1)=N^{\ell}g(N)(v_1)$$

כאשר $f(N)$ הוא פולינום ב־N ממעלה r_1-1 לכל היותר, והוצאנו מהצירוף הליניארי של חזקות N ב־f את N^{ℓ} ככל שניתן (ייתכן ש־$\ell=0$), וכך קיבלנו את הפולינום $g(N)$, עבורו אנו יודעים בוודאות שהאיבר החופשי אינו אפס. כלומר g הוא פולינום כך ש־$g(0)\neq 0$. כעת

- אם $\ell\geq r$ אז סיימנו כי $N^{\ell}\equiv 0$ ולכן $N^{r_2}(v_2)=0_V$ ונגדיר $v_2'=v_2$.
- אחרת, אם $\ell<r$ אז $N^{\ell}g(N)(v_1)$ לא מתאפס בוודאות (כי ל־g יש איבר חופשי שאינו אפס), לכן הוא יתאפס על ידי הפעלת N $r-\ell$ פעמים (ולא לפני!), נפעיל $N^{r-\ell}$ על שני אגפי המשוואה ונקבל:

$$N^{r-\ell}N^{r_2}(v_2)=\underbrace{N^{r-\ell}N^{\ell}}_{=0}g(N)(v_1)=0_V$$

כלומר קיבלנו $N^{r-\ell+r_2}(v_2)=0_V$, ו־$r-\ell+r_2$ הוא המינימלי כנ"ל $\Leftarrow$ $r\leq r-\ell+r_2$, כלומר $\ell\leq r_2$. לפיכך נוכל להגדיר $v_2'=v_2-N^{\ell-r_2}g(N)(v_1)$, וזה "הנציג הטוב". מדוע?

ראשית, נשים לב שהם שייכים לאותו קוסט במנה V/W, כלומר $\overline{v_2}=\overline{v_2'}$, הסיבה לכך היא ש־$N^{\ell-r_2}g(N)(v_1)\in W$. בנוסף, נשים לב ש

$$N^{r_2}(v_2')=N^{r_2}(v_2)-N^{r_2}\left(N^{\ell-r_2}g(N)(v_1)\right)=N^{r_2}(v_2)-\underbrace{N^{\ell}g(N)(v_1)}_{N^{r_2}(v_2)}=0_V$$

ולכן זהו נציג טוב, שהפעלת N^{r_2} עליו מאפסת אותו. ■

נזכיר שהמטרה שלנו הייתה לדעת למצוא צורת ג'ורדן, לפני שנראה כיצד משפט הפירוק הציקלי מביא אותנו למטרה זו, ננסח את הלמה הבאה.

<u>למה</u>: אם $N:V\to V$ נילפוטנטי מאינדקס r ו־$v\notin\ker N^{r-1}$, אז $v,Nv,...,N^{r-1}v$ הם בת"ל.

<u>הוכחה</u>: נניח שעבור $\alpha_1,\alpha_2,...,\alpha_r\in\mathbb{F}$ מתקיים

$$\alpha_1v+\alpha_2Nv+...+\alpha_rN^{r-1}v=0$$

עלינו להראות ש־$\alpha_1=\alpha_2=...=\alpha_r=0$. נפעיל את N:

$$\alpha_1Nv+\alpha_2N^2v+...+\underbrace{\alpha_rN^rv}_{=0}=\vec{0}$$

נפעיל N שוב ושוב עד שנקבל

$$\alpha_1N^{r-1}v+\underbrace{\alpha_2N^rv}_{=0}=\vec{0}$$

אבל $v\notin\ker N^{r-1}$ ולכן בהכרח $\alpha_1=0$. כעת נחזור על התהליך ונקבל $\alpha_2=0$ וכו'. ■

<u>מסקנה ממשפט הפירוק הציקלי והלמה</u>: קיום צורת ג'ורדן.

עבור $i=1,2,...,t$ מתקיים ש־$B_i=\left\{N^{r_i-1}v_i,...,Nv_i,v_i\right\}$ הוא בסיס של $\langle v_i\rangle$ ו־

$$\left[N|_{\langle v_i\rangle}\right]_{B_i}=\begin{pmatrix}0&1&0&\cdots&&0\\0&0&1&&&\\0&0&0&\ddots&&\\\vdots&\vdots&\vdots&&&\\&&&&0&1\\0&&&&&0\end{pmatrix}_{r_i\times r_i}=J_{r_i}(0)$$

כלומר קיבלנו בלוק ג'ורדן (עליון, כלומר ה־1ים מעל האלכסון הראשי. אם היינו כותבים את וקטורי הבסיס בסדר הפוך, היינו מקבלים בלוק ג'ורדן תחתון ־ בו ה־1ים מתחת לאלכסון הראשי).

בנוסף, אנו יודעים ממשפט הפירוק הציקלי ש־$B = B_1 \cup B_2 \cup ... \cup B_t$ הוא בסיס של V. ומתקיים

$$[N]_B = \begin{pmatrix} J_{r_1}(0) & & & \\ & J_{r_2}(0) & & \\ & & \ddots & \\ & & & J_{r_t}(0) \end{pmatrix}$$

וקיבלנו מטריצה בצורת ג'ורדן.

<u>הערות חשובות</u>

1. גודל מקסימלי של בלוק ג'ורדן הוא $r = r_1$ (כי ראינו ש־$r = r_1 \geq r_2 \geq ...$), r הוא אינדקס הנילפוטנטיות (החזקה בפולינום המינימלי של N).

2. מספר הבלוקים הוא t, שזה בדיוק הריבוי הגיאומטרי של 0, זהו $\dim\ker N$. הסיבה היא שבכל בסיס B_i יש בדיוק וקטור אחד עם ע"ע 0, והוא הוקטור $N^{r_i-1}v_i$.

3. מספר הבלוקים מגודל <u>לפחות</u> k הוא בדיוק $\dim\ker\left(N^k\right) - \dim\ker\left(N^{k-1}\right)$. מדוע? כי כדי ליצור בלוק בגודל k, נדרשת שרשרת באורך k, שרשרת כזו מתחילה מאיבר השייך ל־$\dim\ker N^k$ אבל לא שייך ל־$\dim\ker N^{k-1}$. מתוך כך ניתן גם למצוא כמה בלוקים יש מגודל <u>בדיוק k</u>, מצאו את הנוסחה כתרגיל.

4. מסקנה מסעיפים $1,2,3$: יחידות צורת ג'ורדן, עד כדי סדר הבלוקים. נשים לב שהבסיס המג'רדן איננו יחיד, אך צורת ג'ורדן עצמה כן יחידה (עד כדי סדר הבלוקים), שכן מספר הבלוקים וגודלם נקבעים באופן יחיד על ידי תכונות N כמפורט בסעיפים $1,2,3$.

<u>מסקנה סופית</u>: קיום ויחידות צורת ג'ורדן לאופרטור <u>כללי</u>.

יהי $T : V \to V$ אופרטור ליניארי, כאשר V מרחב וקטור <u>מעל $\mathbb{C}$</u>. ראינו ממשפט ג'ורדן־שבלייה שקיים בסיס $\mathcal{F}$ של V כך ש־

$$[T]_{\mathcal{F}} = \begin{pmatrix} \lambda_1\mathbb{I} & & 0 \\ & \ddots & \\ 0 & & \lambda_m\mathbb{I} \end{pmatrix} + \begin{pmatrix} N_1 & & 0 \\ & \ddots & \\ 0 & & N_m \end{pmatrix}$$

כאשר $N_1,..,N_t$ נילפוטנטיות, $\lambda_1\mathbb{I},...,\lambda_m\mathbb{I}$ סקלריות, ו־$\lambda_1,...,\lambda_m$ הם הע"ע השונים של T. כעת, לכל אחת מ־N_i קיימת P_i הפיכה (שעמודותיה הן השרשראות הציקליות הבת"ל) כך ש־

$$P_i^{-1}N_iP_i = J_i$$

כאשר J_i היא צורת ג'ורדן של N_i <u>המורכבת מבלוקי ג'ורדן</u>.

נשים לב ש־$P = \begin{pmatrix} P_1 & & \\ & \ddots & \\ & & P_m \end{pmatrix}$ היא הפיכה והיא מעבירה את $[T]_{\mathcal{F}}$ למטריצה המייצגת בבסיס המג'רדן $[T]_B$:

$$[T]_B = P^{-1}[T]_{\mathcal{F}}P = P^{-1}\begin{pmatrix} \lambda_1\mathbb{I} & & 0 \\ & \ddots & \\ 0 & & \lambda_m\mathbb{I} \end{pmatrix}P + P^{-1}\begin{pmatrix} N_1 & & 0 \\ & \ddots & \\ 0 & & N_m \end{pmatrix}P = \begin{pmatrix} \lambda_1\mathbb{I} & & 0 \\ & \ddots & \\ 0 & & \lambda_m\mathbb{I} \end{pmatrix} + \begin{pmatrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_m \end{pmatrix} =$$

$$= \begin{pmatrix} \begin{pmatrix} J_{\ell_1}^1(\lambda_1) & & \\ & \ddots & \\ & & J_{\ell_k}^1(\lambda_1) \end{pmatrix} & & \\ & \ddots & \\ & & \begin{pmatrix} J_{\ell_1}^m(\lambda_m) & & \\ & \ddots & \\ & & J_{\ell_j}^m(\lambda_m) \end{pmatrix} \end{pmatrix}$$

דוגמה:

$$\begin{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} & & \\ & (2) & \\ & & \begin{pmatrix} 3 & 1 & \\ & 3 & 1 \\ & & 3 \end{pmatrix} \end{pmatrix}$$

חשוב להדגיש שבפועל כשנרצה לג'רדן לא נעשה את כל התהליך הזה, רשמנו את התהליך של כתיבת $[T]_{\mathcal{F}}$ כסכום של מטריצת בלוקים נילפוטנטיים ומטריצת בלוקים סקלריים רק כדי להצדיק את האמירה שקיימת צורת ג'ורדן לאופרטור כללי.

<u>חשוב לזכור</u>: לכל ע"ע λ

- מספר בלוקי ג'ורדן של λ הוא הריבוי הגיאומטרי של λ.
- גודל בלוק ג'ורדן המקסימלי של λ יהיה החזקה k בפולינום המינימלי של T: $(x-\lambda)^k$.
- סכום כל הגדלים של בלוקי ג'ורדן של λ הוא הריבוי האלגברי - שהוא החזקה בפולינום האופייני.

<u>הערת העשרה</u>: לצורת ג'ורדן יתרונות רבים, מבחינה חישובית כמו שנוח יותר לחשב חזקות בצורה אלכסונית, כך גם נוח למדי לחשב אותן בצורת ג'ורדן. בנוסף, משתמשים בצורת ג'ורדן הרבה בתחום של משוואות דיפרנציאליות, בשביל חישובים של בינה מלאכותית, ועוד.

<u>תרגיל</u>: מצאו את כל המטריצות $A \in M_4(\mathbb{C})$ (עד כדי דמיון) כך ש-$\left(A^2-5A+6\mathbb{I}\right)^2=0$.

<u>פתרון</u>: נשים לב ש-$\left((A-2\mathbb{I})(A-3\mathbb{I})\right)^2=0$, ולכן $\lambda=2$ הוא ע"ע עם ר"א 2, ו-$\lambda=3$ הוא ע"ע עם ר"א 2. המטריצה היא 4×4 ולכן זהו הפולינום האופייני שלה בהכרח. נרשום טבלה של כל האפשרויות כעת:

פולינום אופייני	פולינום מינימלי	צורת ג'ורדן
$(x-2)^2(x-3)^2$	$(x-2)^2(x-3)^2$	$\operatorname{diag}\left(\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}\right)$
$(x-2)^2(x-3)^2$	$(x-2)(x-3)^2$	$\operatorname{diag}\left(\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}\right)$
$(x-2)^2(x-3)^2$	$(x-2)^2(x-3)$	$\operatorname{diag}\left(\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}\right)$
$(x-2)^2(x-3)^2$	$(x-2)(x-3)$	$\operatorname{diag}\left(\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}\right)$

והתשובה לשאלה היא: כל מטריצה הדומה מעל $\mathbb{C}$ לאחת מארבעת מטריצות ג'ורדן הנ"ל.

<u>הרצאה 13 - 14 למאי</u>

עד כה סיימנו את מטרת העל שהצבנו לעצמנו: מציאת צורת ג'ורדן. בשיעור זה נסכם את מה שעשינו עד כה בצורה כללית. הרעיון בשיעור זה הוא לראות את היער ופחות להתעסק בעצים.

<u>סיכום החלק הראשון בקורס</u>

<u>המטרה של החלק הראשון</u>: לתת תשובה מלאה לשאלה של דמיון מטריצות/אופרטורים, או במילים אחרות: רצינו למצוא צורת ג'ורדן למטריצה/אופרטור. בעצם מצאנו אינוואריאנטה לדמיון, וצורות ג'ורדן הן נציגים של מחלקות הדמיון השונות.

<u>משפטים</u>	<u>כלים</u>
טרלי: אם S <u>לכסין</u>, אז $S=\sum\lambda_i P_i$ כאשר P_i הטלות על $\ker(S-\lambda_i\mathbb{I})$	פולינום אופייני,פולינום מינימלי, סכום ישר, הטלות, מרחבים שמורים
ימרי - מעל $\mathbb{C}$ - $T=\sum\lambda_i P_i$ כאשר P_i הטלות על $\ker\left((T-\lambda_i\mathbb{I})^{\beta_i}\right)$	כנ"ל + מרחבי מנה
משפט ג'ורדן-שבלייה - מעל $\mathbb{C}$ - $T=S+N$	משפט הפירוק הציקלי לאופרטור נילפוטנטי - קיום של צורת ג'ורדן

משפט הפירוק הספקטרלי שימש אותנו לאופרטור לכסין, משפט הפירוק הפרימרי נתן לנו הכללה שלו, שעובדת גם למקרה הלא לכסין (בו המרחבים העצמיים לא מכסים את כל המרחב ולכן משתמשים בהטלות על "מרחבים עצמיים מוכללים").

תרגיל: יהי $T:\mathbb{R}^3\to\mathbb{R}^3$ הנתון בבסיס הסטנדרטי E על ידי

$$[T]_E=\begin{pmatrix}1&2&3\\0&1&2\\0&0&2\end{pmatrix}$$

מצאו בסיס B של $\mathbb{R}^3$ כך ש־$[T]_B$ היא בצורת ג'ורדן (לחילופין יכלו לבקש שנמצא מטריצה הפיכה P כך ש־$P^{-1}TP$ היא בצורת ג'ורדן).

פתרון: המטריצה משולשת עליונה, לכן הע"ע הם 1 עם ר"א 2, ו־2 עם ר"א 1. לכן הפולינום האופייני הוא:

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

ניתן לנסות לחשב את הפולינום המינימלי על ידי הצבה של T במחלקים של $p_T(x)$, אבל זו לא שיטה נוחה (בגלל החזקה). לכן נחפש את הריבוי הגיאומטרי של הע"ע 1 וכך נדע מהם בלוקי ג'ורדן של המטריצה.

$$A-\mathbb{I}=\begin{pmatrix}0&2&3\\0&0&2\\0&0&1\end{pmatrix}$$

אנו יודעים שהריבוי הגיאומטרי = דרגות החופש של המערכת = מספר הנעלמים פחות הדרגה = $3-r(A-\mathbb{I})$ = $3-2$ = 1.

מכאן שר"א $\neq$ ר"ג, ולכן T לא לכסין, ולכן הפולינום המינימלי הוא $m_T(x)=(x-1)^2(x-2)$ (כי ראינו שאם כל החזקות בפולינום המינימלי הן 1 אז האופרטור לכסין). מכאן שצורת ג'ורדן חייבת להכיל בלוק 2×2 עבור הע"ע 1, ובלוק 1×1 עבור הע"ע 2. כלומר נקבל

$$J=\begin{pmatrix}1&1&0\\0&1&0\\0&0&2\end{pmatrix}$$

עלינו למצוא גם את הבסיס. בשביל כל בלוק עלינו למצוא שרשרת ציקלית שאורכה מתאים לגודל הבלוק. הבסיס יהיה מהצורה:

$$B=\{(A-\mathbb{I})v_1,\ v_1\ :\ v_2\}$$

כאשר $v_1\in\ker\left((A-\mathbb{I})^2\right)$ ו־$v_1\notin\ker(A-\mathbb{I})$, ו־$v_2$ הוא כל ו"ע עם ע"ע של 2. על מנת למצוא v_1 מתאים, עלינו להשתמש בשיטות של אלגברה א' (הוא אמור להיות ב־$\ker\left((A-\mathbb{I})^2\right)$ אבל לא ב־$\ker(A-\mathbb{I})$, רושמים משוואות ומוצאים אותו).

תרגיל: נתון $A\in M_{4\times4}(\mathbb{C})$ לא לכסינה, ונתון $A^3=A^2$. מהן צורות ג'ורדן האפשריות ל־A?

פתרון: נתון $A^3-A^2=0_{4\times4}$ ⇐ מהגדרת הפולינום המינימלי נסיק ש־$m_A(x)|x^3-x^2=x^2(x-1)$, מהנתון ש־A לא לכסינה נובע שהפולינום המינימלי אינו מכפלה של גורמים ליניאריים שונים, ולכן יש שתי אפשרויות: $m_A(x)=x^2(x-1)$ או $m_A(x)=x^2$.
לכן כל האפשרויות הן:

צורת ג'ורדן	פולינום אופייני	פולינום מינימלי
$\begin{pmatrix} (1) & & \\ & (0) & \\ & & \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \end{pmatrix}$	$x^3(x-1)$	$x^2(x-1)$
$\begin{pmatrix} (1) & & \\ & (1) & \\ & & \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \end{pmatrix}$	$x^2(x-1)^2$	$x^2(x-1)$
$\left\{\begin{matrix} \begin{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} & \\ & \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \end{pmatrix} \\ \begin{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} & & \\ & (0) & \\ & & (0) \end{pmatrix} \end{matrix}\right.$	x^4	x^2

בכך סיימנו את החלק הראשון של הקורס, כעת נעבור לחלק השני.

מוטיבציה לחלק השני של הקורס

אנו רוצים להוסיף למבנה של מרחב וקטורי, גם תכונות גיאומטריות. לשם כך נרצה להגדיר משהו בעל תכונות של מרחק. מהן התכונות שהיינו מצפים שיהיה למרחק? אנחנו יודעים שלמרחק התכונות הבאות:

1) אי שלילי - והוא מתאפס רק כאשר מודדים מרחק מהנקודה עצמה.

2) סימטריות: מרחק מ-p ל-q שווה למרחק מ-q ל-p.

3) ליניאריות: $d\left(\alpha q,p\right)=\alpha d\left(q,p\right)$.

זה מביא אותנו לנושא העיקרי הבא בקורס, מרחבי מכפלה פנימית.

מרחבי מכפלה פנימית

מעתה (עד שיוכרז אחרת) השדה $\mathbb{F}$ משמעותו $\mathbb{R}$ (הממשיים) או $\mathbb{C}$ (המרוכבים).

יהי V מרחב וקטורי סוף-מימדי מעל $\mathbb{F}$. נאמר ש-V הוא מרחב מכפלה פנימית (ממ"פ) אם קיימת פונקציה $\langle\cdot,\cdot\rangle\ :\ V\times V\longrightarrow\mathbb{F}$ המקיימת:

1) אי שליליות: $\langle v,v\rangle\geq0$ לכל $v\in V$ (בפרט $\langle v,v\rangle$ ממשי, זה ינבע מהתכונה השניה). בנוסף $\langle v,v\rangle=0$ אמ"ם $v=\vec{0}$.

2) סימטריות עד כדי הצמדה:

$$\forall v,u\in V\quad\langle v,u\rangle=\overline{\langle u,v\rangle}$$

נשים לב שאם $\mathbb{F}=\mathbb{R}$ אז מתקבלת סימטריות ממש, אך אם $\mathbb{F}=\mathbb{C}$ מתקבלת סימטריות בהצמדה קומפלקסית.

3) ליניאריות ברכיב הראשון:

$$\begin{aligned}\forall v,u,w\in V\quad&\langle v+u,w\rangle=\langle v,w\rangle+\langle u,w\rangle\\ \forall v,w\in V,\,\alpha\in\mathbb{F}\quad&\alpha\langle v,w\rangle=\langle\alpha v,w\rangle\end{aligned}$$

אנו נראה שליניאריות ברכיב השני "כמעט" נובעת מהנחות אלו, ליניאריות בחיבור תתקיים ברכיב השני, אך ליניאריות לכפל בסקלר תתקיים עד כדי הצמדה של הסקלר.

הערה: אם $\mathbb{F}=\mathbb{R}$ אז ממ"פ V נקרא מרחב אוקלידי. אם $\mathbb{F}=\mathbb{C}$ אז ממ"פ V נקרא אוניטרי.

הרצאה 14 - 16 למאי

תזכורת: הנושא החדש שלנו הוא מרחבי מכפלה פנימית.

מרחב וקטורי סוף-מימדי V מעל שדה $\mathbb{F}$ ($\mathbb{C}$ או $\mathbb{R}$) הוא מרחב מכפלה פנימית אם קיימת פונקציה $\langle\cdot,\cdot\rangle : V\times V\to\mathbb{F}$ (הנקראת מכפלה פנימית) כך שמתקיימות התכונות הבאות:

1) אי שליליות: $\forall v\in V\quad\langle v,v\rangle\geq0$, וכמו כן $\langle v,v\rangle=0\iff v=0_V$.

2) סימטריות (עד כדי הצמדה): $\forall v,u\in V\quad\langle v,u\rangle=\overline{\langle u,v\rangle}$.

3) ליניאריות ברכיב הראשון:

$$\begin{aligned}\forall v,u,w\in V\quad&\langle v+u,w\rangle=\langle v,w\rangle+\langle u,w\rangle\\ \forall v,u\in V,\ \ \alpha\in\mathbb{F}\quad&\alpha\langle v,u\rangle=\langle\alpha v,u\rangle\end{aligned}$$

הערות:

א. אם $\mathbb{F}=\mathbb{R}$ אז משמעותה של תכונה (2) היא סימטריות:

$$\langle v,u\rangle=\langle u,v\rangle$$

ואז ליניאריות מתקיימת גם לרכיב השני.

אם $\mathbb{F}=\mathbb{C}$ אז מתכונות $(3),(2)$ נובע שמתקיים

$$\langle v,\alpha w\rangle=\overline{\langle \alpha w,v\rangle}=\overline{\alpha\,\langle w,v\rangle}\underbrace{=}_{\overline{z_1z_2}=\overline{z_1}\,\overline{z_2}}\overline{\alpha}\overline{\langle w,v\rangle}=\overline{\alpha}\,\langle v,w\rangle$$

לכן במקרה זה הליניאריות בכפל בסקלר ברכיב השני היא עד כדי הצמדה. לגבי חיבור נקבל

$$\langle v,u+w\rangle=\overline{\langle u+w,v\rangle}=\overline{\langle u,v\rangle+\langle w,v\rangle}\underbrace{=}_{\overline{z_1+z_2}=\overline{z_1}+\overline{z_2}}\overline{\langle u,v\rangle}+\overline{\langle w,v\rangle}=\langle v,u\rangle+\langle v,w\rangle$$

לסיכום: הליניאריות לחיבור <u>תמיד</u> מתקיימת ברכיב השני, הליניאריות לכפל בסקלר ברכיב השני מתקיימת אם $\mathbb{F}=\mathbb{R}$, ומתקיימת עד כדי הצמדה אם $\mathbb{F}=\mathbb{C}$.

ב. למרחבי מכפלה פנימית יש שימושים רבים בפיסיקה, הנדסה, וכו'...

ג. ישנן הכללות של מושג המכפלה הפנימית, דוגמאות: תבניות ריבועיות מעל כל שדה, פיתוח תורת פורייה מעל חבורות, אנליזה פונקציונלית (מרחבי הילברט), ועוד ועוד.

<u>דוגמאות</u>

1) נביט ב-$V=\mathbb{R}^2$, ונגדיר מכפלה פנימית באופן הבא:

$$\left\langle \begin{pmatrix}x_1\\x_2\end{pmatrix},\begin{pmatrix}y_1\\y_2\end{pmatrix}\right\rangle=x_1y_1+x_2y_2$$

בדקו שפונקציה זו אכן מקיימת את שלושת הדרישות של מכפלה פנימית!

<u>הערת העשרה</u>: זו למעשה הגדרת מכפלה סקלרית המוכרת מפיזיקה

$$\vec{v}\cdot\vec{u}=|\vec{v}|\,|\vec{u}|\cos(\alpha)$$

ניתן להכליל זאת, באופן כללי אם $V=\mathbb{R}^n$ מעל $\mathbb{R}$ נגדיר את <u>המכפלה הפנימית הסטנדרטית על $\mathbb{R}^n$</u>:

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

אם $V=\mathbb{C}^n$ מעל $\mathbb{C}$ אז ההגדרה הזו לא תעבוד, ובמקומה נגדיר את <u>המכפלה הפנימית הסטנדרטית על $\mathbb{C}^n$</u>:

$$\left\langle \begin{pmatrix}z_1\\\vdots\\z_n\end{pmatrix},\begin{pmatrix}w_1\\\vdots\\w_n\end{pmatrix}\right\rangle=\sum_{i=1}^{n}z_i\overline{w_i}$$

בתרגילים, ברירת המחדל היא המכפלה הפנימית הסטנדרטית, לכן אם לא ציינו במפורש מהי המכפלה הפנימית נוכל להניח שהיא הסטנדרטית.

2) נביט במרחב $V=M_{m\times n}(\mathbb{F})$ - מרחב המטריצות $m\times n$. נגדיר את המכפלה הפנימית הבאה

$$\langle A,B\rangle=\operatorname{tr}\left(\overline{B}^tA\right)$$

הסבר: בסימון $\overline{B}$ כוונתנו הצמדה של כל כניסות המטריצה (וכמובן שאם $\mathbb{F}=\mathbb{R}$ ניתן להתעלם משלב זה), הסיבה שלקחנו עקבה היא שבסופו של דבר המכפלה הפנימית אמורה לתת מספר. נשים לב שמבחינת מימדי המטריצה הכפל מוגדר:

$$\underbrace{\overline{B}^t}_{n\times m}\underbrace{A}_{m\times n}$$

כלומר מתקבלת מטריצה $n \times n$ והמכפלה הפנימית היא העקבה שלה. בדקו שפונקציה זו אכן מקיימת את שלושת התכונות של מכפלה פנימית!

למכפלה הפנימית הזו קוראים המכפלה הפנימית הסטנדרטית על V. נעיר גם שלעיתים משתמשים בסימון $B^* = \overline{B}^t$.

דוגמאות העשרה (לא בסילבוס)

1) נביט במרחב V=מרחב כל הפונקציות הממשיות הרציפות ב־$[a,b]$. ניתן להגדיר עליו מכפלה פנימית

$$\langle g(t), f(t) \rangle = \int_a^b g(t) f(t)\, dt$$

ניתן גם להגדיר מכפלה פנימית על מרחב הפונקציות המרוכבות הרציפות, ואז תהיה הצמדה על $f(t)$.

2) $V = \ell^2$= מרחב הסדרות הממשיות $(a_1, a_2, ...)$ כך שהטור $\sum_{n=1}^{\infty} a_n^2$ מתכנס, כלומר

$$\sum_{n=1}^{\infty} a_n^2 < \infty$$

ואז ניתן להגדיר מכפלה פנימית באופן הבא

$$\langle (a_n), (b_n) \rangle = \sum_{n=1}^{\infty} a_n b_n$$

ונשים לב שהטור $\sum_{n=1}^{\infty} a_n b_n$ מתכנס וששלושת התכונות של מכפלה פנימית מתקיימות.

3) V= מרחב של משתנים מקריים, ואז מגדירים מכפלה פנימית באופן הבא

$$\langle X, Y \rangle = \mathbb{E}(XY)$$

כלומר המכפלה הפנימית היא התוחלת של מכפלת המשתנים המקריים.

תרגיל ־ מוטיבציה לבאות

יהי $V = \mathbb{R}^2$ ויהי a פרמטר ממשי. נגדיר:

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle = x_1 y_1 + a x_1 y_2 - x_2 y_1 + 2 x_2 y_2$$

לאילו ערכי a נקבל מכפלה פנימית?

דרך פתרון א': לבדוק ישירות לפי ההגדרה את שלושת האקסיומות.

דרך פתרון ב': נשים לב שהביטוי הנ"ל הוא למעשה:

$$\underbrace{\begin{pmatrix} x_1 & x_2 \end{pmatrix}}_{1\times 2} \underbrace{\begin{pmatrix} 1 & a \\ -1 & 2 \end{pmatrix}}_{2\times 2} \underbrace{\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}}_{2\times 1}$$

אכן:

$$\begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 1 & a \\ -1 & 2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 & a x_1 + 2 x_2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = x_1 y_1 - x_2 y_1 + a x_1 y_2 + 2 x_2 y_2 = \left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle$$

ולכן השאלה האם מתקיימות דרישות המכפלה הפנימית תלויה בתכונות של המטריצה $A=\begin{pmatrix}1 & a\\ -1 & 2\end{pmatrix}$.

שפי שנראה בטענה הבאה, $\langle\cdot,\cdot\rangle$ יהיה מכפלה פנימית אמ"ם המטריצה A סימטרית ובנוסף $\vec{x}^tA\vec{x}>0$ לכל $\vec{0}\neq\vec{x}\in\mathbb{R}^2$ (כאשר $\vec{x}=\begin{pmatrix}x_1\\ x_2\end{pmatrix}$ ולכן $\vec{x}^t=\begin{pmatrix}x_1 & x_2\end{pmatrix}$), תנאי זה למעשה מבטיח את האי־שליליות. מטריצה המקיימת תנאים אלו נקראת <u>חיובית לחלוטין</u> (הגדרה מדויקת בהמשך).

בעתיד נראה כלים שיחסכו לנו את החישוב הזה, למשל אנו נראה שמטריצה סימטרית וממשית (בנוסף להיותה לכסינה) היא חיובית לחלוטין אמ"ם כל הערכים העצמיים שלה חיוביים ממש.

בדוגמה שלנו תנאי זה יתקיים כאשר $a=-1$ (כדי שהמטריצה תהיה סימטרית), ואז

$$x^tAx=x_1^2-x_1x_2-x_2x_1+2x_2^2=x_1^2-2x_1x_2+2x_2^2=(x_1-x_2)^2+x_2^2>0$$

כלומר קיבלנו ש־$x^tAx>0$ לכל $\vec{x}\neq 0$, ולכן המטריצה חיובית לחלוטין.

<u>הגדרה:</u>

א. מטריצה $A\in M_n(\mathbb{R})$ תיקרא <u>חיובית לחלוטין</u> אם A סימטרית (כלומר $A^t=A$) ובנוסף $x^tAx>0$ לכל $\vec{0}\neq\vec{x}\in\mathbb{R}^n$, כאשר $\vec{x}=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$ ולכן $\vec{x}^t=\begin{pmatrix}x_1 & \dots & x_n\end{pmatrix}$.

ב. מטריצה $A\in M_n(\mathbb{C})$ תיקרא <u>חיובית לחלוטין</u> אם A <u>הרמיטית</u>, כלומר $A=A^*=\overline{A}^t$, ובנוסף $x^*Ax>0$ לכל $\vec{0}\neq\vec{x}\in\mathbb{C}^n$, כאשר כאמור $x^*=\overline{x}^t$ והצמדה מרוכבת היא איבר־איבר.

<u>טענה:</u>

א. תהי $A\in M_n(\mathbb{R})$ חיובית לחלוטין, אז

$$\langle\vec{x},\vec{y}\rangle_A=x^tAy$$

היא מכפלה פנימית $\langle\cdot,\cdot\rangle:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$.

ב. אם $\langle\cdot,\cdot\rangle:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ איזושהי מכפלה פנימית, אז קיימת $A\in M_n(\mathbb{R})$ חיובית לחלוטין כך ש־$\langle x,y\rangle=\langle x,y\rangle_A=x^tAy$.

ג. א. תהי $A\in M_n(\mathbb{C})$ חיובית לחלוטין, אז

$$\langle\vec{x},\vec{y}\rangle_A=x^*Ay$$

היא מכפלה פנימית $\langle\cdot,\cdot\rangle:\mathbb{C}^n\times\mathbb{C}^n\to\mathbb{C}$.

ד. אם $\langle\cdot,\cdot\rangle:\mathbb{C}^n\times\mathbb{C}^n\to\mathbb{C}$ איזושהי מכפלה פנימית, אז קיימת $A\in M_n(\mathbb{C})$ חיובית לחלוטין כך ש־$\langle x,y\rangle=\langle x,y\rangle_A=x^*Ay$.

<u>הוכחת סעיפים א',ב' (הוכחת סעיפים ג' ו־ד' דומה)</u>

א. נוכיח על פי הגדרה.

- <u>אי שליליות</u>: ברורה מהנתון.
- <u>סימטריות</u>:

$$\langle x,y\rangle_A=\underbrace{x^t}_{1\times n}\underbrace{A}_{n\times n}\underbrace{y}_{n\times 1}=\underbrace{x^t}_{1\times n}\underbrace{(Ay)}_{n\times 1}$$

אבל נשים לב שמתקיים

$$\begin{pmatrix}\alpha_1 & \dots & \alpha_n\end{pmatrix}\begin{pmatrix}\beta_1\\ \vdots\\ \beta_n\end{pmatrix}=\sum\alpha_i\beta_i=\sum\beta_i\alpha_i=\begin{pmatrix}\beta_1 & \dots & \beta_n\end{pmatrix}\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix}$$

לכן נקבל

$$\langle x,y\rangle_A = x^tAy = x^t(Ay) = (Ay)^t\left(x^t\right)^t = y^tA^tx \underbrace{=}_{A=A^t} y^tAx = \langle y,x\rangle_A$$

- ליניאריות: נובעת מתכונות כפל מטריצות, כפל בסקלר, וטרנספוז:

$$\langle x+z,y\rangle_A = (x+z)^tAy = \left(x^t+z^t\right)Ay = x^tAy + z^tAy = \langle x,y\rangle_A + \langle z,y\rangle_A$$
$$\langle \alpha x,y\rangle_A = (ax)^tAy = \alpha x^tAy = \alpha\langle x,y\rangle_A$$

ב. תהא $\langle\cdot,\cdot\rangle:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ מכפלה פנימית, בפרט נוכל להגדיר סקלרים ממשיים

$$1\le i,j\le n \quad \mathbb{R}\ni a_{ij} = \langle e_i,e_j\rangle$$

כאשר $\{e_i\}_{i=1}^n$ הוא הבסיס הסטנדרטי. נקבל את המטריצה $M_n(\mathbb{R})\ni A=(a_{ij})=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & & & \vdots \\ \vdots & & & \\ a_{n1} & \dots & & a_{nn}\end{pmatrix}$. כיוון ש־$\{e_i\}_{i=1}^n$ בסיס של $\mathbb{R}^n$ אז מתכונות הליניאריות והסימטריות של מכפלה פנימית נקבל שעבור

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n\end{pmatrix} = \sum_{i=1}^n x_ie_i$$
$$\vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n\end{pmatrix} = \sum_{j=1}^n y_je_j$$

מתקבל:

$$\langle\vec{x},\vec{y}\rangle = \sum_{i=1}^n\sum_{j=1}^n x_iy_j\langle e_i,e_j\rangle \underbrace{=}_{\text{definition of }A} \sum_{i,j=1}^n a_{ij}x_iy_j = x^tAy$$

וכמו כן, לאור החשבונות מסעיף א', A חיובית לחלוטין. ■

טענה: יהיו V,W מרחבים וקטוריים ממימד סופי מעל $\mathbb{F}$ ($\mathbb{R}$ או $\mathbb{C}$), ונניח של־W יש מכפלה פנימית $\langle\cdot,\cdot\rangle_W$, אז בהינתן $T:V\to W$ העתקה ליניארית, הפיכה (במילים אחרות, חח"ע ועל, או לחילופין איזומורפיזם) מתקיים ש־

$$\langle\vec{v},\vec{u}\rangle := \langle T\vec{v},T\vec{u}\rangle_W$$

מכפלה פנימית של V. הוכיחו טענה זו כתרגיל!

מסקנה: לכל מרחב וקטורי V סוף־מימדי מעל $\mathbb{F}$ יש מכפלה פנימית, כי כזכור מאלגברה א' $T:V\to\mathbb{F}^n$ המוגדר כך ש־$T(v)=[v]_B$ (כאשר $[v]_B$ הוא וקטור קואורדינטות לפי בסיס B) הוא העתקה ליניארית הפיכה, ולכן משתי הטענות הקודמות התוצאה נובעת.

הגדרה (נורמה): נניח ש־V מרחב מכפלה פנימית סוף מימדי מעל $\mathbb{F}$ ($\mathbb{R}$ או $\mathbb{C}$) עם המכפלה הפנימית $\langle\cdot,\cdot\rangle$. לכל $v\in V$ נגדיר את הנורמה של v או האורך של v באופן הבא:

$$\|v\| = \sqrt{\langle v,v\rangle}\in\mathbb{R}$$

כמובן שהגודל $\sqrt{\langle v,v\rangle}$ הוא תמיד ממשי ואי שלילי (מתכונות המכפלה הפנימית), והוקטור היחיד בעל אורך אפס הוא וקטור האפס של V. אלו בדיוק התכונות שהיינו רוצים שפונקציית אורך תקיים.

דוגמה: עבור $V=\mathbb{R}^2$ עם המכפלה הפנימית הסטנדרטית נקבל

$$\|\vec{x}\| = \sqrt{\langle \vec{x},\vec{x}\rangle} \underbrace{=}_{\vec{x}=\begin{pmatrix}x_1\\x_2\end{pmatrix}} \sqrt{x_1^2+x_2^2}$$

וזהו בדיוק האורך האוקלידי המוכר לנו.

הערת העשרה: ניתן להגדיר מרחק עבור $\vec{v},\vec{u}\in V$ על ידי

$$d(\vec{v},\vec{u}) = \|\vec{v}-\vec{u}\|$$

למרחק הנ"ל קוראים מטריקה, ואנו רואים שהוא מקיים את התכונות שמרחק בין שתי נקודות אמור לקיים: אי שלילי ומתאפס רק כאשר $\vec{u}=\vec{v}$ (כלומר מודדים את המרחק מהנקודה לעצמה), סימטרי, ואי שוויון המשולש. עוד על כך ניתן ללמוד בקורס במרחבים מטריים וטופולוגיים.

טענה: יהי V ממ"פ עם מ"פ $\langle\cdot,\cdot\rangle$ ונורמה $\|\cdot\|$.

1) בהינתן $\alpha\in\mathbb{F}$ ו־$v\in V$ מתקיים $\|\alpha v\| = |\alpha|\,\|v\|$.

2) נאמר ש־v הוא וקטור יחידה אם $\|v\|=1$. הראו שבהינתן $0\neq v\in V$ אז $\hat{v}=\frac{v}{\|v\|}$ הוא וקטור יחידה.

הוכחה:

1)

$$\|\alpha v\| = \sqrt{\langle \alpha v,\alpha v\rangle} = \sqrt{\alpha\overline{\alpha}\,\langle v,v\rangle} \underbrace{=}_{\alpha\overline{\alpha}=|\alpha|^2} |\alpha|\sqrt{\langle v,v\rangle} = |\alpha|\,\|v\|$$

2)

$$\|\hat{v}\| = \left\|\frac{v}{\|v\|}\right\| \underbrace{=}_{(1)} \left|\frac{1}{\|v\|}\right|\|v\| \underbrace{=}_{\|v\|>0} \frac{1}{\|v\|}\|v\| = 1$$

הרצאה - 21 למאי

תזכורת: אנו מדברים כעת על מרחבי מכפלה פנימית, וכעת בנוסף לפעולת החיבור של וקטורים והכפל של סקלר, נרצה להוסיף תכונות גיאומטריות למרחב: מרחקים וזוויות.

אנו עוסקים ב־V מרחב וקטורי (ממימד סופי בלבד) מעל $\mathbb{F}$ (כאשר $\mathbb{F}$ הוא $\mathbb{R}$ או $\mathbb{C}$) עם מכפלה פנימית $\langle\cdot,\cdot\rangle$ (נזכיר שלמדנו שכל מכפלה פנימית ניתן להציג כמטריצה סימטרית וחיובית לחלוטין) ונורמה המוגדרת באופן הבא

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

וראינו שהנורמה מקיימת את התכונות שהיינו מצפים שפונקציה המייצגת מרחק תקיים (למשל, האורך של וקטור האפס הוא אפס).

כעת אנו רוצים להעמיק את ההיכרות שלנו עם מרחבים אלו, אנו יודעים למשל שב־$\mathbb{R}^3$ ניתן לקחת את הבסיס הסטנדרטי ולקבל שלושה וקטורים הניצבים זה לזה. מאחר ואנו יודעים שבחירת בסיס כזה היא נוחה מאוד, נרצה להכליל זאת למרחב מכפלה פנימית V כלשהו.

משפט (אי שוויון קושי־שווארץ - Cauchy-Schwarz): יהי V ממ"פ (מעל $\mathbb{R}$ או $\mathbb{C}$), ויהיו $a,b\in V$, אזי

$$|\langle a,b\rangle| \leq \|a\|\cdot\|b\|$$

ומתקיים שוויון כאשר a, b תלויים ליניארית.

<u>הערת העשרה</u>: בקורס במרחבים מטריים משתמשים באי־שוויון זה כדי להראות שפונקציית המכפלה הפנימית היא פונקציה רציפה.

<u>הוכחה</u>: מאקסיומת אי־שליליות אנחנו יודעים שלכל $t \in \mathbb{F}$ מתקיים

$$\langle a - tb, a - tb \rangle \geq 0$$

מליניאריות וסימטריות (עד כדי הצמדה) נקבל

$$(*) \quad \langle a, a \rangle - t \underbrace{\overline{\langle a, b \rangle}}_{\langle b, a \rangle} - \bar{t} \langle a, b \rangle + t\bar{t} \langle b, b \rangle \geq 0$$

אם $b = 0_V$ אז $\langle a, b \rangle = \langle b, a \rangle = \langle b, b \rangle = 0$ ונקבל שא"ש קושי שווארץ מתקיים לכל $a \in V$ כי $\langle a, a \rangle \geq 0$. אחרת, $b \neq 0_V$ ואז $\|b\| \neq 0$ ולכן נוכל להציב ב־$(*)$ בפרט $t = \frac{\langle a, b \rangle}{\|b\|^2}$ (נשים לב ש־t מוגדר כי המכנה אינו אפס), לאחר ההצבה גם נכפיל את $(*)$ ב־$\langle b, b \rangle = \|b\|^2$ ונקבל:

$$\langle a, a \rangle \langle b, b \rangle - \overline{\langle a, b \rangle} \langle a, b \rangle - \langle a, b \rangle \overline{\langle a, b \rangle} + \langle a, b \rangle \overline{\langle a, b \rangle} \geq 0$$

כלומר:

$$|\langle a, b \rangle|^2 = \overline{\langle a, b \rangle} \langle a, b \rangle \leq \langle a, a \rangle \langle b, b \rangle = \|a\|^2 \|b\|^2$$

קיבלנו אי שוויון ששני אגפיו אי־שליליים, לכן ניתן להוציא שורש ולקבל:

$$\boxed{|\langle a, b \rangle| \leq \|a\| \cdot \|b\|}$$

כלומר קיבלנו את א"ש קושי־שווארץ. נותר להראות מדוע יש שוויון אמ"ם a, b תלויים ליניארית. כיוון ראשון: נשים לב שאם a, b תלויים ליניארית אז קיים סקלר $\alpha \in \mathbb{F}$ כך ש־$a = \alpha b$. ובמקרה זה נקבל

$$|\langle a, b \rangle| = |\langle \alpha b, b \rangle| = |\alpha| \left| \underbrace{\langle b, b \rangle}_{\geq 0} \right| = |\alpha| \langle b, b \rangle = |\alpha| \cdot \|b\|^2 = \underbrace{|\alpha| \|b\|}_{\|a\|} \|b\| = \|a\| \|b\|$$

את הכיוון השני ודאו לבד - כלומר בדקו מדוע שוויון באי־שוויון קושי שווארץ גורר ש־a, b תלויים.

<u>דוגמה</u>: נמשיך את "דוגמת ההעשרה" שלנו, במרחב הפונקציות שראינו פעם קודם נסיק מאי־שוויון קושי שווארץ שמתקיים

$$\left| \int_a^b f(t) g(t) \, dt \right|^2 \leq \int_a^b f(t)^2 \, dt \int_a^b g(t)^2 \, dt$$

וזו תכונה שלקבל ישירות בעזרת חישוב האינטגרלים יכול להיות קשה, אך קיבלנו אותה מיידית מאי־שוויון קושי שווארץ.

<u>מסקנה (אי שוויון המשולש)</u>: יהי V מרחב מכפלה פנימית עם נורמה $\|\cdot\|$, ויהיו $a, b \in V$, אז

$$\|a + b\| \leq \|a\| + \|b\|$$

ומתקיים שוויון אמ"ם a, b תלויים ליניארית עם מקדם ממשי וחיובי. כלומר $a = \alpha b$ כאשר $0 \leq \alpha \in \mathbb{R}$.

הוכחת המסקנה:

$$\|a+b\|^2=\langle a+b,a+b\rangle=\langle a,a\rangle+\underbrace{\langle a,b\rangle+\overline{\langle a,b\rangle}}_{2\Re\langle a,b\rangle}+\langle b,b\rangle=\langle a,a\rangle+2\Re\langle a,b\rangle+\langle b,b\rangle$$
$$\underbrace{\leq}_{(*)}\langle a,a\rangle+2\left|\langle a,b\rangle\right|+\langle b,b\rangle\underbrace{\leq}_{\text{c.s.}}\|a\|^2+2\|a\|\|b\|+\|b\|^2=\left(\|a\|+\|b\|\right)^2$$

כאשר במעבר $(*)$ השתמשנו בתכונה שלמדנו באלגברה א' לפיה $\Re(z)\leq|z|$. לסיום נוציא שורש ונקבל את א"ש המשולש:

$$\boxed{\|a+b\|\leq\|a\|+\|b\|}$$

את התנאי לשוויון הוכיחו בעצמכם בתור תרגיל. ■

תזכורת (מחדו"א או פיזיקה): המשמעות של המכפלה הפנימית/מכפלה סקלרית ב־$\mathbb{R}^2$ היא שבהינתן שני וקטורים $\vec{u},\vec{v}$ אז המכפלה הסקלרית היא

$$\vec{v}\cdot\vec{u}=|\vec{v}|\,|\vec{u}|\cos(\alpha)$$

פירוש אחר הוא מכפלה פנימית לפי קואורדינטות

$$\vec{v}\cdot\vec{u}=v_1u_1+v_2u_2$$

זה מוביל אותנו להגדרה של זווית במרחב מכפלה פנימית כללי.

הגדרה: יהי V מרחב אוקלידי (כלומר מרחב מכפלה פנימית מעל $\mathbb{R}$) ויהיו $a,b\in V$ כך ש־$a\neq 0_V,\ b\neq 0_V$. נגדיר:

$$\cos\varphi=\frac{\langle a,b\rangle}{\|a\|\cdot\|b\|}$$

מא"ש קושי־שווארץ נובע ש־φ קיימת, ומקיימת $0\leq\varphi\leq\pi$.

הערה: את כל מה שעשינו היום ניתן להכליל באינדוקציה ל־n וקטורים, למשל ניתן להכליל את א"ש קושי־שווארץ בצורה כזו (אולי יינתן כתרגיל רשות לבית).

הקדמה לשיעור הבא: כעת, כשיש לנו מושג של זווית, נוכל להגדיר בסיס של וקטורים הניצבים זה לזה (בסיס אורתוגונלי).

הרצאה ־ 22 למאי

נמשיך עם מרחבי מכפלה פנימית ־ ראינו שברגע שיש לנו מכפלה פנימית ניתן לקבל ממנו "אורך" ־ נורמה ־ וראינו שהנורמה אכן מקיימת את התכונות שנצפה מאורך לקיים. לאחר מכן ראינו שבעזרת המכפלה הפנימית והנורמה ניתן להגדיר זוויות במרחב שלנו. כעת נוכל להגדיר ניצבות של וקטורים, ומכאן שנוכל להגדיר מהו בסיס של וקטורים הניצבים זה לזה.

סימונים לאורך שיעור זה: V ממ"פ מעל שדה $\mathbb{F}$ ($\mathbb{C}$ או $\mathbb{R}$), המכפלה הפנימית היא $\langle\cdot,\cdot\rangle$, הנורמה היא $\|v\|=\sqrt{\langle v,v\rangle}$ עבור $v\in V$, והמרחק בין שני וקטורים הוא $d(u,v)=\|u-v\|$.

אורתוגונליות

הגדרה: יהיו $a,b\in V$. נאמר שהם אורתוגונליים (א"ג)/ניצבים אם מתקיים $\langle a,b\rangle=0$.

הערה: וקטור האפס הוא היחיד שא"ג לכל $v\in V$.

דוגמאות:

1) ב־$\mathbb{R}^2/\mathbb{R}^3$ עם המכפלה הפנימית הסטנדרטית, u,v הם א"ג אמ"ם הזווית ביניהם היא 90°. לכן אנו רואים שההגדרה שלנו לניצבות עקבית עם המקרה המוכר לנו עד כה.

2) המשך דוגמת ההעשרה שלנו: במרחב הפונקציות הממשיות ורציפות ב־$[-\pi,\pi]$, הפונקציות $\sin(mx),\sin(nx)$ ($m\neq n$) הן אורתוגונליות. נזכיר שבמרחב זה המכפלה הפנימית בין שתי פונקציות f,g מוגדרת באופן הבא:

$$\int_{-\pi}^{\pi}fgdx$$

ואכן נקבל אורתוגונליות:

$$\begin{aligned}\int_{-\pi}^{\pi}\sin(nx)\sin(mx)\,dx & \underbrace{=}_{\text{trig identity}}\frac{1}{2}\int_{-\pi}^{\pi}\cos\left[(m-n)x\right]-\cos\left[(m+n)x\right]dx\\ & \underbrace{=}_{\text{parts}}\frac{1}{m-n}\sin\left[(m-n)x\right]|_{-\pi}^{\pi}-\frac{1}{m+n}\sin\left[(m+n)x\right]|_{-\pi}^{\pi}\\ & =0\end{aligned}$$

תכונה זו שימושית מאוד בקורסים רבים: אנליזה פונקציונלית, מתמטיקה שימושית, אלקטרודינמיקה, קוונטים, ועוד. בעצם מצאנו בסיס אורתוגונלי של מרחב הפונקציות.

3) הוקטורים $b=\begin{pmatrix}1-i\\1\\2i-1\end{pmatrix}$, $a=\begin{pmatrix}1+i\\1\\1\end{pmatrix}$ אורתוגונליים ב־$\mathbb{C}^3$ עם המכפלה הפנימית הסטנדרטית. נזכיר שהמכפלה הפנימית הסטנדרטית במרחב $\mathbb{C}^n$ היא מהצורה

$$\langle\vec{x},\vec{y}\rangle=\sum_{i=1}^{n}x_i\overline{y_i}$$

ואכן מתקבל

$$\langle a,b\rangle=\sum_{j=1}^{3}a_j\overline{b_j}=(1+i)\overline{(1-i)}+1\cdot1+1\cdot\overline{2i-1}=(1+i)^2+1-1-2i=1+2i-1-2i=0$$

<u>הגדרה</u>: קבוצת וקטורים $S\subset V$ תיקרא אורתוגונלית (א"ג) אם כל זוג של וקטורים שונים ב־S הוא אורתוגונלי. כלומר לכל $a\neq b\in S$ מתקיים $\langle a,b\rangle=0$.

<u>טענה</u>: אם S קבוצה אורתוגונלית של וקטורים השונים מוקטור האפס ($\neq 0_V$) אז S בת"ל.

<u>הוכחה</u>: נסמן $S=\{a_1,a_2,...,a_n\}$, יהיו $\alpha_i\in\mathbb{F}$ כך ש:

$$\alpha_1a_1+\alpha_2a_2+...+\alpha_na_n=0$$

נכפיל את שני האגפים פנימית מול הוקטור a_i כלשהו, תוך כדי שימוש בליניאריות. נקבל:

$$\alpha_1\underbrace{\langle a_1,a_i\rangle}_{=0}+\alpha_2\underbrace{\langle a_2,a_i\rangle}_{=0}+...\alpha_i\langle a_i,a_i\rangle+...+\alpha_n\underbrace{\langle a_n,a_i\rangle}_{=0}=\langle 0,a_i\rangle=0$$

כאשר השתמשנו באורתוגונליות. כלומר קיבלנו

$$\alpha_i\langle a_i,a_i\rangle=0$$

אבל $\langle a_i,a_i\rangle>0$ כי $a_i\neq\vec{0}$, ולכן $\alpha_i=0$. i היה שרירותי, ולכן תוצאה זו נכונה לכל i, ולכן הוקטורים בלתי תלויים. ■

<u>מסקנה</u>: אם V ממ"פ מעל $\mathbb{F}$ בעל מימד n אז קבוצה אורתוגונלית בעלת n וקטורים $\neq\vec{0}$ היא <u>בסיס</u>.

משפט (תהליך האורתוגונליזציה של גרם־שמידט - Gram-Schmidt process): בכל ממ"פ V (ממימד n) מעל $\mathbb{F}$ ($\mathbb{R}$ או $\mathbb{C}$) קיים בסיס אורתוגונלי. יתר על כן, לכל קבוצה בת"ל הפורשת תת־מרחב $W \subset V$ יש קבוצה אורתוגונלית הפורשת את W.

הוכחה: נניח ש־$S = \{a_1, a_2, ..., a_n\}$ איזשהו בסיס ל־V (לחילופין - איזושהי קבוצה בת"ל). נבנה קבוצה $\{b_1, b_2, ..., b_n\}$ של וקטורים אורתוגונליים שונים מ־0_V כך ש־

$$\mathrm{span}S = \mathrm{span}\{b_1, ..., b_n\}$$

נגדיר $b_1 = a_1$. עבור כל סקלר α מתקיים ש־$\mathrm{span}\{b_1, \alpha b_1 + a_2\} = \mathrm{span}\{a_1, a_2\}$, נרצה למצוא α כך ש־$b_2 = \alpha b_1 + a_2$ יהיה אורתוגונלי ל־b_1, כלומר נדרוש:

$$0 = \langle b_2, b_1 \rangle = \langle \alpha b_1 + a_2, b_1 \rangle = \alpha \langle b_1, b_1 \rangle + \langle a_2, b_1 \rangle$$

לכן

$$\alpha \langle b_1, b_1 \rangle = -\langle a_2, b_1 \rangle \;\Rightarrow\; \alpha = -\frac{\langle a_2, b_1 \rangle}{\langle b_1, b_1 \rangle}$$

כאשר נשים לב שלא חילקנו באפס כי נתון ש־$b_1 \neq \vec{0}$, ולכן $\langle b_1, b_1 \rangle > 0$.

כלומר עבור ה־α שמצאנו נקבל:

$$b_2 = \alpha b_1 + a_2 = a_2 - \frac{\langle a_2, b_1 \rangle}{\langle b_1, b_1 \rangle} b_1 = a_2 - \frac{\langle a_2, b_1 \rangle}{\|b_1\|^2} b_1 = a_2 - \frac{\langle a_2, b_1 \rangle}{\|b_1\|} \frac{b_1}{\|b_1\|}$$

זו תוצאה הגיונית, כי בעצם אנחנו לוקחים את a_2 ומחסרים ממנה את ההיטל של a_2 על b_1, ובכך "נפטרים" מהרכיב של a_2 שאינו ניצב ל־b_1.

נשים לב ש־$b_2 \neq \vec{0}$ כי נתון ש־a_2, $a_1 = b_1$ הם בת"ל, ואם $b_2 = 0$ נקבל $\alpha b_1 + a_2 = 0$, בסתירה להיותם בת"ל.

כעת, עלינו למצוא את b_3: נשים לב שלכל סקלרים ℓ_1, ℓ_2 מתקיים

$$\mathrm{span}\{b_1, b_2, \ell_1 b_1 + \ell_2 b_2 + a_3\} = \mathrm{span}\{a_1, a_2, a_3\}$$

עלינו לבחור ℓ_1, ℓ_2 כך ש־$b_3 = \ell_1 b_1 + \ell_2 b_2 + a_3$ יהיה אורתוגונלי עם b_1 ו־b_2. כלומר

$$0 = \langle b_3, b_1 \rangle = \langle \ell_1 b_1 + \ell_2 b_2 + a_3, b_1 \rangle = \ell_1 \langle b_1, b_1 \rangle + \ell_2 \underbrace{\langle b_2, b_1 \rangle}_{=0} + \langle a_3, b_1 \rangle =$$

$$= \langle a_3, b_1 \rangle + \|b_1\|^2 \ell_1$$

כלומר $\ell_1 = -\frac{\langle a_3, b_1 \rangle}{\|b_1\|^2}$. באופן דומה, מהדרישה $\langle b_3, b_2 \rangle = 0$ נקבל $\ell_2 = -\frac{\langle a_3, b_2 \rangle}{\|b_2\|^2}$. לכן בסך הכל קיבלנו:

$$b_3 = a_3 - \frac{\langle a_3, b_1 \rangle}{\|b_1\|^2} b_1 - \frac{\langle a_3, b_2 \rangle}{\|b_2\|^2} b_2$$

בצורה דומה נקבל את b_4

$$b_4 = a_4 - \frac{\langle a_4, b_1 \rangle}{\|b_1\|^2} b_1 - \frac{\langle a_4, b_2 \rangle}{\|b_2\|^2} b_2 - \frac{\langle a_4, b_3 \rangle}{\|b_3\|^2} b_3$$

וכך נמשיך עד n ונקבל שלכל $1 \leq k \leq n$:

$$b_k = a_k - \sum_{i=1}^{k-1} \frac{\langle a_k, b_i \rangle}{\|b_i\|^2} b_i$$

וכך $\{b_1, ..., b_n\}$ קבוצה אורתוגונלית שלא מכילה את וקטור האפס, גודלה n ולכן היא בסיס אורתוגונלי. ■

<u>מסקנות/תרגילים</u>:

1) אם $\{v_1, ..., v_n\}$ בסיס אורתוגונלי לממ"פ V, ו-$v \in V$, אז

$$v = \sum_{i=1}^{n} \frac{\langle v, v_i \rangle}{\|v_i\|^2} v_i$$

כלומר אנו יודעים מהם המקדמים בהצגה של v כק"ל של איברי הבסיס!

2) משפט פיתגורס: אם $\{v_1, ..., v_n\}$ קבוצה אורתוגונלית אז

$$\|v_1 + ... + v_n\|^2 = \|v_1\|^2 + ... + \|v_n\|^2$$

ועבור $n = 2$ מתקבל משפט פיתגורס המוכר מ-$\mathbb{R}^2$.

3) בעזרת משפט פיתגורס ותהליך גרם־שמידט ניתן להוכיח את אי־שוויון קושי־שווארץ.

<u>הגדרות</u>:

1) נאמר שקבוצה אורתוגונלית של וקטורים היא <u>אורתונורמלית (א"נ)</u>, אם כל וקטור v בה הוא מנורמל, כלומר בעל אורך 1, כלומר נורמה 1:

$$\|v\| = 1$$

נעיר שתמיד ניתן לנרמל את הוקטורים האורתוגונליים המתקבלים בתהליך גרם־שמידט, על ידי הגדרת $\tilde{b}_k = \frac{b_k}{\|b_k\|}$, ואכן מתקיים $\left\|\tilde{b}_k\right\| = 1$. מכאן המסקנה: שתמיד קיים בסיס אורתונורמלי למרחב.

2) בהינתן תת־מרחב $W \subset V$ נגדיר את המרחב הניצב (או המרחב האורתוגונלי) ל-W באופן הבא

$$W^\perp = \{v : V \ : \ \langle v, w \rangle = 0 \quad \forall w \in W\}$$

מדוע $W^\perp$ הוא מרחב? עלינו להראות סגירות לכפל בסקלר וגם לחיבור, אך זה מתקבל מיידית מהליניאריות של המכפלה הפנימית. בנוסף, ברור ש-$W^\perp$ לא ריקה כי איבר האפס שייך לה (כי הוא ניצב לכל וקטור במרחב).

<u>טענה</u>: $V = W \oplus W^\perp$. כלומר בהינתן $W, W^\perp$ החיתוך שלהם ריק וסכומם מכסה את כל המרחב.

<u>הוכחה</u>: תרגיל (תמצית ההוכחה: ניקח בסיס ל-W, נשלים לבסיס של V, נפעיל תהליך גרם־שמידט, ונקבל שה"גרמשמודט" של ההשלמה הוא בסיס ל-$W^\perp$).

<u>הגדרה</u>: להטלה על W במקביל ל-$W^\perp$ קוראים <u>הטלה אורתוגונלית על W</u>. כלומר אם אנו רואים את המילים "הטלה אורתוגונלית על W" אנו יודעים שזה במקביל לניצב $W^\perp$.

<u>דוגמה</u>: נביט ב-$V = \mathbb{R}^3$ עם המכפלה הפנימית הסטנדרטית. נגדיר תת־מרחב

$$W = \text{span}\left\{\begin{pmatrix}1\\2\\3\end{pmatrix}\right\}$$

חשבו בסיס אורתונורמלי ל-$W^\perp$.

פתרון: מטרה: למצוא שני וקטורים אורתוגונליים ל־$\begin{pmatrix}1\\2\\3\end{pmatrix}$ ואז לנרמל אותם. במילים אחרות, לגרם־שמדט (ואז לנרמל). נתחיל בקבוצה

$$\left\{\underbrace{\begin{pmatrix}1\\2\\3\end{pmatrix}}_{=a_1},\underbrace{\begin{pmatrix}0\\1\\0\end{pmatrix}}_{=a_2},\underbrace{\begin{pmatrix}0\\0\\1\end{pmatrix}}_{=a_3}\right\}$$

קבוצה זו מהווה בסיס למרחב, כעת נגרם־שמדט אותה:

נגדיר $b_1=a_1$, אז נקבל את b_2:

$$b_2=a_2-\frac{\langle a_2,b_1\rangle}{\|b_1\|^2}b_1=\begin{pmatrix}0\\1\\0\end{pmatrix}-\frac{2}{14}\begin{pmatrix}1\\2\\3\end{pmatrix}=\begin{pmatrix}0\\1\\0\end{pmatrix}-\frac{1}{7}\begin{pmatrix}1\\2\\3\end{pmatrix}=\frac{1}{7}\begin{pmatrix}-1\\5\\-3\end{pmatrix}$$

ונקבל את b_3:

$$b_3=a_3-\frac{\langle a_3,b_1\rangle}{\|b_1\|^2}b_1-\frac{\langle a_3,b_2\rangle}{\|b_2\|^2}b_2=\begin{pmatrix}0\\0\\1\end{pmatrix}-\frac{3}{14}\begin{pmatrix}1\\2\\3\end{pmatrix}-\frac{-\frac{3}{7}}{25}\cdot\frac{1}{7}\begin{pmatrix}-1\\5\\-3\end{pmatrix}=\begin{pmatrix}*\\ *\\ *\end{pmatrix}$$

תשובה סופית: בסיס אורתונורמלי ל־$W^\perp$ הינו: $\left\{\frac{b_2}{\|b_2\|},\frac{b_3}{\|b_3\|}\right\}$.

דרך נוספת: לחילופין, יכולנו לומר שברצוננו למצוא שני וקטורים $\begin{pmatrix}x\\y\\z\end{pmatrix}$ המאונכים ל־$\begin{pmatrix}1\\2\\3\end{pmatrix}$, ולכן עליהם לקיים:

$$x+2y+3z=0$$

שני וקטורים אפשריים הם $\begin{pmatrix}0\\1.5\\-1\end{pmatrix}$, $\begin{pmatrix}1\\1\\-1\end{pmatrix}$, אנו רואים שהם בלתי תלויים, ולכן רק נותר לבצע תהליך גרם־שמידט על שניהם בלבד, ולנרמל אותם.

<u>הרצאה ־ 28 למאי</u>

תזכורת/סיכום: V ממ"פ מעל $\mathbb{F}$ ($\mathbb{R}$ או $\mathbb{C}$) עם מכפלה פנימית $\langle\cdot,\cdot\rangle$, שהיא בעצם מטריצה סימטרית וחיובית לחלוטין כלומר

$$\langle x,y\rangle=x^*Ay$$

עם A חיובית לחלוטין. המכפלה הפנימית גם נותרת לנו נורמה $\|x\|=\sqrt{\langle x,x\rangle}$, וראינו (מאי שוויון קושי־שווארץ) שניתן להגדיר זווית בין v,w:

$$\cos\varphi=\frac{\langle v,w\rangle}{\|v\|\,\|w\|}$$

בכך בעצם הכללנו את המקרה המוכר לנו (מפיזיקה) של אורך וזווית. הכללה זו שימושית לכל מרחב מכפלה פנימית: בין אם האובייקטים בו הם וקטורים, פונקציות, מטריצות, וכו'.

ברגע שהיה לנו מושג של זווית, הגדרנו ש־v,w הם אורתוגונליים/ניצבים אם $\langle v,w\rangle=0$. בהינתן $W\subset V$ תת־מרחב, הגדרנו את תת המרחב הניצב

$$W^\perp=\{v\in V\ :\ \forall w\in W\quad\langle v,w\rangle=0\}$$

ראינו ש־$V = W \oplus W^{\perp}$ (מסקנה מתהליך גרם־שמידט).

תהליך גרם־שמידט: אם $\{a_1, ..., a_n\}$ הוא בסיס של V, אז תהליך גרם־שמידט מאפשר לנו למצוא $\{b_1, ..., b_n\}$ שהוא בסיס אורתוגונלי של V, ואז $\left\{\frac{b_1}{\|b_1\|}, ..., \frac{b_n}{\|b_n\|}\right\}$ בסיס אורתונורמלי של V. כיצד בונים את ה־b_kים? הנוסחה הכללית היא

$$b_k = a_k - \sum_{i=1}^{k-1} \frac{\langle a_k, b_i \rangle}{\|b_i\|^2} b_i$$

אך במקום לשנן נוסחה כדאי גם להבין אותה. הרעיון הוא שבהינתן שני וקטורים a_1, a_2, ניתן לבחור $b_1 = a_1$ ואז נותר רק להטיל את a_2 עליו. אם הזווית בין a_1, a_2 היא φ, אז גודל ההיטל של a_2 על b_1 הוא $\|a_2\| \cos\varphi$, כלומר ההיטל הוא

$$\vec{P}(a_2) = \frac{\|a_2\| \langle a_2, b_1 \rangle}{\|a_2\| \|b_1\|} \frac{b_1}{\|b_1\|}$$

לכן נגדיר $b_2 = b_1 - \vec{P}(a_2)$, בכך אנו בעצם לוקחים את b_1, ומחסרים ממנו את החלק של a_2 שלא ניצב לו, ומקבלים וקטור b_2 שכן ניצב לו. באופן כללי עבור b_k עובדים בצורה דומה: לוקחים את a_k ומחסרים את כל הרכיבים שלו שלא ניצבים ל־b_k, ומכאן הנוסחה הכללית שכתבנו מקודם.

הערה: ההטלה על W במקביל ל־$W^{\perp}$ נקראת ההטלה האורתוגונלית על W.

טענה: אם $\{v_1, ..., v_n\}$ בסיס אורתונורמלי של V, ו־$v \in V$ אז $v = \sum_{i=1}^{n} \langle v, v_i \rangle v_i$. למקדמים $\langle v, v_i \rangle$ קוראים מקדמי פורייה.

הוכחה: נסמן $v = \sum_{i=1}^{n} \alpha_i v_i$ (ניתן, כי $\{v_i\}$ בסיס). יש להוכיח שלכל $1 \le i \le n$ מתקיים $\alpha_i = \langle v, v_i \rangle$. נכפיל פנימית עם v_j ונקבל:

$$\langle v, v_j \rangle = \left\langle \sum_{i=1}^{n} \alpha_i v_i, v_j \right\rangle \underbrace{=}_{\text{linearity}} \sum_{i=1}^{n} \alpha_i \langle v_i, v_j \rangle \underbrace{=}_{\langle v_i, v_j \rangle = \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}} \alpha_j$$

מיחידות ההצגה של איבר במרחב כצירוף ליניארי של איברי בסיס, נקבל את הדרוש. למעשה הערה זו אינה הכרחית, כי הוכחנו שכל הצגה של v כק"ל של v_i מקיימת $\alpha_i = \langle v, v_i \rangle$.

הערה: השתמשנו בסימון δ_{ij}, הנקרא הדלתא של קרוניקר, והוא מקיים:

$$\delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

כלומר אם האינדקסים זהים מתקבל ערך 1, ואם הם שונים מתקבל הערך 0.

משפט (אי שוויון בסל, שוויון פרסבל): תהי $S = \{v_1, ..., v_m\}$ קבוצה אורתונורמלית ב־V ממ"פ ממימד סופי n, יהי $v \in V$ עם $\alpha_i = \langle v, v_i \rangle$ עבור $i = 1, ..., m$. אזי

$$\|v\|^2 \ge |\alpha_1|^2 + ... + |\alpha_m|^2$$

במקרה בו $n = m$ (כלומר במקרה בו S היא בסיס אורתונורמלי) מתקיים שוויון

$$\|v\|^2 = |\alpha_1|^2 + ... + |\alpha_n|^2$$

הוכחה: אם $m = n$ אז S כבר בסיס, אחרת (אם $m < n$) נשלים אותה לבסיס אורתונורמלי (מובטח שאפשר בגלל גרם־שמידט) $\{v_1, ..., v_m, v_{m+1}, ..., v_n\}$ ונקבל מהטענה הקודמת ש־

$$v = \sum_{i=1}^{n} \alpha_i v_i$$

כאשר $\alpha_i = \langle v, v_i \rangle$ הם מקדמי פורייה. נחשב את ריבוע הנורמה:

$$\|v\|^2 = \langle v, v \rangle = \left\langle \sum_{i=1}^{n} \alpha_i v_i, \sum_{j=1}^{n} \alpha_j v_j \right\rangle = \sum_{i,j=1}^{n} \alpha_i \overline{\alpha_j} \langle v_i, v_j \rangle \underbrace{=}_{\langle v_i, v_j \rangle = \delta_{ij}} \sum_{i=1}^{n} \alpha_i \overline{\alpha_i} = \sum_{i=1}^{n} |\alpha_i|^2$$

בכך קיבלנו את שוויון פרסבל, כאשר השתמשנו בליניאריות, בהרמטיות (=סימטריות עד כדי הצמדה), ובכך ש־$\{v_i\}_{i=1}^{n}$ הוא בסיס אורתונורמלי. למעשה קיבלנו גם את אי־שוויון בסל, כי אם $m < n$ אז

$$\|v\|^2 = \sum_{i=1}^{n} |\alpha_i|^2 \geq \sum_{i=1}^{m} |\alpha_i|^2$$

כאשר השתמשנו בכך שכל $|\alpha_i|^2$ הוא אי־שלילי. ■

<u>הערה</u>: שוויון פרסבל בעצם מהווה מעין הכללה של משפט פיתגורס.

<u>תרגיל (יופיע בשיעורי הבית)</u>: נראה שאם $S = \{v_1, ..., v_m\}$ היא קבוצה אורתונורמלית, ו־$v \in V$ ו־β_i סקלרים כלשהם, אז

$$\left\| v - \sum_{i=1}^{m} \beta_i v_i \right\| \geq \left\| v - \sum_{i=1}^{m} \alpha_i v_i \right\|$$

כאשר α_i הם מקדמי פורייה. המשמעות של טענה זו היא שאם יש לנו קבוצה אורתונורמלית עם m איברים (ולכן אינה בסיס אם $m < n$), אז מקדמי פורייה נותנים את הקירוב הטוב ביותר ל־v, מבין כל האפשרויות האחרות.

<u>מסקנה מהתרגיל</u>: $W \subset V$ תת־מרחב עם בסיס אורתונורמלי $\{w_1, ..., w_m\}$. תהי P_W ההטלה האורתוגונלית על W. יהי $v \in V$, אז

$$P_W(v) = \sum_{i=1}^{m} \langle v, w_i \rangle w_i$$

ו־$P_W(v)$ הזה הינו הקירוב הטוב ביותר ל־v מבין כל הוקטורים ב־W, כלומר לכל $w \in W$ מתקיים

$$\|v - w\| \geq \|v - P_W(v)\|$$

ובאותו אופן, אם נשלים את הבסיס לבסיס אורתונורמלי $\{w_1, ..., w_m, w_{m+1}, ..., w_n\}$, אז

$$P_{W^\perp}(v) = \sum_{i=m+1}^{n} \langle v, w_i \rangle w_i$$

ו־$P_{W^\perp}(v)$ הוא הקירוב הטוב ביותר ל־v ב־$W^\perp$.

מסקנה זו מובילה אותנו להגדרה הבאה.

<u>הגדרה</u>: <u>המרחק בין v ל־W</u> מוגדר על ידי

$$d(v, W) = \|v - P_W(v)\| = \|P_{W^\perp}(v)\|$$

כאשר השתמשנו בכך ש־$v = P_W(v) + P_{W^\perp}(v)$.

<u>דוגמת העשרה (במרחב אינסוף מימדי)</u>: נביט בפונקציה $f(x) = \frac{x}{\pi}$, כאשר $-\pi < x < \pi$, אז טור פורייה שלה הוא:

$$f(x) = \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin(nx)$$

הקבוצה $\{\sin nx\}_{n=1}^{\infty}$ היא בסיס אורתונורמלי, ובכך הצגנו את האיבר $f(x)$ בתור ק"ל של איברי בסיס זה, כאשר מקדמי פורייה הם $\frac{(-1)^{n+1}}{n}$ והם חושבו על ידי:

$$\langle f(x), \sin(nx)\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)\,dx$$

זוהי תוצאה מאוד יפה, בכך בעצם פרשנו את הפונקציה $f(x)$ בעזרת פונקציות מחזוריות ויפות שאנו מכירים היטב.

<u>הרצאה - 29 למאי</u>

עד עכשיו דיברנו על מרחבים וקטוריים עם מכפלה פנימית, והשלב הבא הוא להשוות בין מרחבים שונים - כלומר אנו נדבר על העתקות בין מרחבים, נדון באיזומורפיזם, וכו'.

<u>הגדרה</u>: יהיו $(V_1, \langle\cdot,\cdot\rangle_1)$, $(V_2, \langle\cdot,\cdot\rangle_2)$ שני ממ"פים מעל $\mathbb{F}$ ($\mathbb{R}$ או $\mathbb{C}$, שניהם מעל אותו $\mathbb{F}$), נאמר שהם <u>איזומורפיים כממ"פ</u> אם קיימת $T: V_1 \to V_2$ העתקה ליניארית חח"ע ועל של מ"ו כך ש-

$$\forall v, w \quad \langle v, w\rangle_1 = \langle Tv, Tw\rangle_2$$

<u>טענה</u>: בסימונים לעיל

$$(V_1, \langle\cdot,\cdot\rangle_1) \cong (V_2, \langle\cdot,\cdot\rangle_2) \iff \dim V_1 = \dim V_2$$

כאשר $\cong$ מסמן איזומורפיזם כממ"פים.

<u>הוכחה</u>:

כיוון $\Rightarrow$: איזומורפיזם כממ"פ גורר גם איזומורפיזם כמ"ו, ואז מאלגברה א' יש שוויון מימדים.

כיוון $\Leftarrow$: מהנתון + גרם שמידט, קיימים $\{x_1, ..., x_n\}$ בסיס אורתונורמלי של V_1 ביחס למכפלה הפנימית $\langle\cdot,\cdot\rangle_1$, וקיימים $\{y_1, ..., y_n\}$ בסיס אורתונורמלי של V_2 ביחס למכפלה הפנימית $\langle\cdot,\cdot\rangle_2$.

נגדיר $T: V_1 \to V_2$ על ידי $T(x_i) = y_i$ לכל $1 \le i \le n$. ראינו באלגברה א' שמספיק להגדיר העתקה על איברי בסיס, ואז ניתן להרחיב את T ליניארית כך ש-T איזומורפיזם של מ"ו ובנוסף

$$\langle x_i, x_j\rangle_1 = \langle y_i, y_j\rangle_2 = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

ולכן מליניאריות והרמיטיות (=סימטריות עד כדי הצמדה) של המכפלות הפנימיות ינבע $\langle v, w\rangle_1 = \langle Tv, Tw\rangle_2$ לכל $v, w \in V_1$ כנדרש.

פונקציונלים ליניאריים והאופרטור הצמוד

מוטיבציה:

1) כשדיברנו על מכפלות פנימיות, ראינו שכל מכפלה פנימית ניתן לכתוב כמטריצה המקיימת $A = A^*$. ראינו שיש קשר הדוק בין העתקות ליניאריות למטריצות המייצגות שלהם. לכן הגיוני לשאול מיהו האופרטור שמתאים ל-A^*, כלומר אנו רוצים לדעת מהו האנלוג האופרטורי למטריצה A^*.

2) אנו רואים שלשאול על איזומורפיזם בין שני מרחבים זה לא כל כך מעניין, כי מה שקובע זאת הוא המימד. נשאל שאלה שונה: כיצד נראית העתקה ליניארית מ-V ל-$\mathbb{F}$ (שהוא ממימד 1)?

כעת נתחיל את הדיון בנושא החדש.

יהי V ממ"פ מעל $\mathbb{F}$ ($\mathbb{R}$ או $\mathbb{C}$) ויהי $v_0 \in V$ קבוע. נגדיר העתקה $f_{v_0} : V \to \mathbb{F}$ באופן הבא

$$f_{v_0}(v) = \langle v, v_0 \rangle$$

אזי f_{v_0} היא העתקה ליניארית (כי הממ"פ ליניארית ברכיב הראשון).

הגדרה: העתקה ליניארית $f : V \to \mathbb{F}$ נקראת פונקציונל ליניארי (פ"ל).

משפט: יהי V ממ"פ סוף מימדי, ויהי $f : V \to \mathbb{F}$ פונקציונל ליניארי, אז קיים ויחיד $v_0 \in V$ כך ש-$f = f_{v_0}$.

הערה: זה שהמשפט עוסק בממ"פ לא פוסל לנו חלק מהמרחבים הוקטוריים האפשריים, זאת מאחר וראינו שבכל מ"ו ניתן להגדיר מכפלה פנימית (לוקחים מטריצה ומתקבלת ממנה מכפלה פנימית).

הערה: המשפט נכון רק למרחב ממ"פ סוף מימדי, עבור מימד אינסופי הטענה איננה נכונה בהכרח.

טיפ: ברגע שנותנים לנו מרחב מכפלה פנימית, הדבר הראשון שמומלץ לעשות הוא לבחור בסיס אורתונורמלי. כך נעשה גם בהוכחה.

הוכחה: V ממ"פ, יהי $\{v_1, ..., v_n\}$ בסיס א"נ של V (מובטח שיש כזה מתהליך גרם שמידט). נגדיר

$$v_0 = \sum_{j=1}^{n} \overline{f(v_j)} v_j$$

ונראה שמתקיים $f = f_{v_0}$. מספיק להראות שקיים שוויון בין f ל-f_{v_0} על איברי הבסיס (כי שניהם פונקציונלים ליניאריים, ובפרט העתקות ליניאריות, ולכן מספיק לעבוד על איברי בסיס). אכן

$$f_{v_0}(v_i) = \langle v_i, v_0 \rangle = \left\langle v_i, \sum_{j=1}^{n} \overline{f(v_j)} v_j \right\rangle \underbrace{=}_{\text{linearity, hermitian}} \sum_{j=1}^{n} f(v_j) \underbrace{\langle v_i, v_j \rangle}_{=\delta_{ij}} = f(v_i)$$

בכך הוכחנו קיום v_0 כך ש-$f_{v_0} = f$, נותר להראות יחידות v_0. לכאורה v_0 תלוי בבסיס האורתונורמלי שבחרנו, אך הוכחת היחידות תראה שבעצם הוא לא תלוי בבחירת הבסיס.

נניח שבנוסף ל-v_0 הנ"ל גם $u \in V$ כך ש-$f_{v_0} = f = f_u$ ונראה ש-$v_0 = u$. מההנחה נובע

$$\forall v \in V \quad \langle v, v_0 \rangle = \langle v, u \rangle$$

נעביר אגף ונשתמש בליניאריות:

$$\forall v \in V \quad \langle v, v_0 - u \rangle = 0$$

מכאן נובע ש-$v_0 - u$ ניצב לכל $v \in V$, ולכן הוא וקטור האפס, כלומר $v_0 - u = 0_V \Leftarrow v_0 = u$, והוכחנו יחידות. ■

הערת העשרה: כאמור הטענה הנ"ל כבר לא נכונה למקרה של V ∞-מימדי. כלומר עבור מימד סופי, כל פונקציונל ליניארי הוא מהצורה של $\langle v, v_0 \rangle$, אבל ברגע שהמימד הוא אינסופי כבר יש לנו פונקציונלים ליניאריים מצורות נוספות. זו תכונה חשובה בתורת ההצגות, שם מסווגים חבורות לפי סוגים שונים בעזרת פונקציונלים ליניאריים, ולכן זה משפיע על ההצגה של חבורות ממימד אינסופי.

תרגיל בונוס: לשלוח לעדי דוגמה נגדית.

הערה: בהינתן פונקציונל ליניארי $f:V\to\mathbb{F}$, אם נסמן את הגרעין $W=\ker f$, אז או ש-$f\equiv 0$ (ואז $W=V$) או ש-$\dim \mathrm{Im} f=1$ (כלומר $\mathrm{Im} f=\mathbb{F}$) ואז $\dim W^{\perp}=1$ כי

$$\dim\ker f+\dim \mathrm{Im} f=\dim V=\dim W+\dim W^{\perp}$$

לכן למעשה אם $0_V\neq u\in W^{\perp}$ אז $W^{\perp}=\mathrm{span}\left\{\frac{u}{\|u\|}\right\}$. נסמן $v_1=\frac{u}{\|u\|}$ ונשלים למרחב א"נ למרחב כולו $\{v_1,...,v_n\}$. לכן $v_2,..,v_n$ כולם בסיס של W, כלומר הם בגרעין. ואז לפי הגדרת v_0 בהוכחה נקבל

$$v_0=\frac{\overline{f(u)}}{\|u\|^2}u$$

כעת ניתן לראות ש-

$$f_{v_0}(v)=f(v)\underbrace{=}_{v=\underbrace{P_W(v)}_{\in\ker f}+P_{W^{\perp}}(v)}f\left(P_{W^{\perp}}(v)\right)=f\left(\left\langle v,\frac{u}{\|u\|}\right\rangle\frac{u}{\|u\|}\right)=f\left(\frac{\langle v,u\rangle}{\|u\|^2}u\right)$$

כלומר לא מעניין אותנו החלק של v שנמצא בגרעין.

הערה: נתון V מ"ו סוף מימדי. נסמן ב-V^* את אוסף כל הפונקציונלים הליניאריים $f:V\to\mathbb{F}$, כלומר

$$V^*=\{f:V\to\mathbb{F}\}$$

זהו מרחב וקטורי מעל $\mathbb{F}$ (נקרא המרחב הדואלי) ביחס לחיבור נקודתי וכפל בסקלר נקודתי:

$$\begin{gathered}(f+g)(v)=f(v)+g(v)\\(\alpha f)(v)=\alpha f(v)\end{gathered}$$

זהו מקרה פרטי של מה שראינו באלגברה א' ($\mathrm{Hom}(V,W)$). אנו יכולים לחשוב על כל העתקה כמטריצה $\dim V\times 1$, ולכן מתקיים ש-$\dim V=\dim V^*$ ולכן $V\cong V^*$, אבל האיזומורפיזם הזה תלוי בבחירת הבסיס, כלומר הוא איזומורפיזם לא קנוני. וודאו שאם $\{v_1,...,v_n\}$ בסיס א"נ של V אז $\{f_{v_1},...,f_{v_n}\}$ בסיס של V^*.

לעומת זאת, $V\cong(V^*)^*$ באופן קנוני (לא תלוי בבחירת הבסיס), מי שילמד קורס בתורת ההצגות יראה הגדרה פורמלית של איזומורפיזם קנוני. כיצד מתקבל איזומורפיזם קנוני זה? נשים לב ש-

$$(V^*)^*=\{g:V^*\to\mathbb{F}\}$$

האיזומורפיזם מתקבל על ידי העתקה $T:V\to(V^*)^*$ המקיימת $T(v)=Tv$, ואז מגדירים $Tv:V^*\to\mathbb{F}$ המקיימת $Tv(f)=f(v)$. הראו ש-T העתקה ליניארית חח"ע ועל.

<u>האופרטור הצמוד</u>

<u>משפט/הגדרה</u>: יהי V ממ"פ סוף מימדי, ויהי $T:V\to V$ אופרטור ליניארי. אזי קיים ויחיד אופרטור ליניארי $T^*:V\to V$ כך ש־

$$\forall u,v\in V \quad \langle Tv,u\rangle=\langle v,T^*u\rangle$$

ל־T^* קוראים <u>האופרטור הצמוד (Adjoint Operator)</u>.

<u>הוכחה</u>: יהי $u\in V$ ונביט על הפ"ל $f:V\to\mathbb{F}$ הבא

$$f(v)=\langle Tv,u\rangle$$

זהו פונקציונל ליניארי כי T ליניארי והמכפלה הפנימית ליניארית במשתנה הראשון. לפי טענה קודמת, קיים ויחיד $u_0\in V$ כך ש־$f=f_{u_0}$ כלומר

$$f(v)=\langle Tv,u\rangle=\langle v,u_0\rangle$$

נגדיר פונקציה $T^*:V\to V$ באופן הבא:

$$T^*(u)=u_0$$

בעצם f תלוי ב־u, ולכן לכל u קיים u_0. נשים לב שזו אכן פונקציה כי לכל $u\in V$ <u>קיים ויחיד $u_0\in V$</u> שמתאים לו. הפונקציה הזו גם מקיימת

$$(*)\quad \langle Tv,u\rangle=\langle v,T^*u\rangle$$

עלינו להראות ש־T^* ליניארית. אכן

$$\begin{aligned}\langle v,T^*(\alpha u+w)\rangle \underbrace{=}_{(*)} \langle Tv,\alpha u+w\rangle \underbrace{=}_{\text{linearity}} \langle Tv,\alpha u\rangle+\langle Tv,w\rangle=\\ =\overline{\alpha}\langle Tv,u\rangle+\langle Tv,w\rangle\underbrace{=}_{(*)}\overline{\alpha}\langle v,T^*u\rangle+\langle v,T^*w\rangle=\\ =\langle v,\alpha T^*u\rangle+\langle v,T^*w\rangle=\langle v,\alpha T^*u+T^*w\rangle\end{aligned}$$

אם נעביר אגפים נקבל

$$\langle v,T^*(\alpha u+w)-(\alpha T^*u+T^*w)\rangle=0$$

כלומר $T^*(\alpha u+w)-(\alpha T^*u+T^*w)$ ניצב <u>לכל $v\in V$</u>, ומכאן שהוא וקטור האפס, כלומר קיבלנו:

$$T^*(\alpha u+w)=\alpha T^*u+T^*w$$

ובכך הוכחנו ליניאריות. עד כה הוכחנו קיום של T^* כנדרש, וודאו כתרגיל את היחידות של T^*. ■

<u>משפט</u>: יהי $B=\{v_1,...,v_n\}$ בסיס **<u>אורתונורמלי</u>** של V ממ"פ סוף־מימדי מעל $\mathbb{F}$ ($\mathbb{C}$ או $\mathbb{R}$) ויהי $T:V\to V$ אופרטור ליניארי. נסמן את המטריצה המייצגת $A=[T]_B$. אז

$$[T^*]_B=A^*=\overline{A}^t$$

במילים אחרות, $[T^*]_B = [T]_B^*$. חשוב לזכור שטענה זו נכונה רק עבור בסיס אורתונורמלי!!!

הערה: כעת למעשה יש לנו שתי דרכים למצוא את T^*, דרך אחת לפי ההגדרה $\langle Tv,u\rangle = \langle v,T^*u\rangle$ - רושמים את מערכת המשוואות ופותרים. דרך שניה היא לפי בסיס אורתונורמלי ושימוש במטריצה המייצגת. אם נתון הבסיס האורתונורמלי הדרך השניה נוחה יותר, אך ללא בסיס זה ייתכן שהדרך הראשונה תהיה נוחה יותר (כי למצוא את הבסיס האורתונורמלי יכול לקחת זמן מה).

הוכחה: לצורך ההוכחה, ננסח ונוכיח את הלמה הבאה:

למה: אם $T:V\to V$ אופרטור ליניארי כלשהו, ו-$B=\{v_1,...,v_n\}$ בסיס אורתונורמלי, אז

$$([T]_B)_{ij} = \langle Tv_j, v_i\rangle$$

הוכחת הלמה: בשיעור הבא.

איך הלמה עוזרת לנו? נסמן $[T]_B = A = (a_{ij})$, לפי הגדרת A^* מתקיים $(A^*)_{ij} = \overline{a_{ji}}$, ולפי הלמה נקבל

$$\overline{a_{ji}} \underbrace{=}_{\text{Lemma for } T} \overline{\langle Tv_i, v_j\rangle} \underbrace{=}_{\text{Def of } T^*} \overline{\langle v_i, T^*v_j\rangle} \underbrace{=}_{\text{Hermitian}} \langle T^*v_j, v_i\rangle \underbrace{=}_{\text{Lemma for } T^*} ([T^*]_B)_{ij}$$

ובכך סיימנו את ההוכחה.

הרצאה - 4 ליוני

תזכורת: ראינו שאם V ממ"פ ממימד סופי ו-$T:V\to V$ אופרטור ליניארי, אז קיים ויחיד T^* (האופרטור הצמוד) המקיים

$$\forall u,v\in V \quad \langle Tv,u\rangle = \langle v,T^*u\rangle$$

הערה: במרחב אינסופי ישנם אופרטורים שאין להם צמוד, אך אנו עוסקים במרחב ממימד סופי בלבד בקורס זה ולכן תמיד קיים האופרטור הצמוד.

ראינו את המשפט: אם B בסיס אורתונורמלי (ביחס לממ"פ הנתונה של V) ו-$T:V\to V$ אז

$$[T^*]_B = [T]_B^*$$

חשוב לציין שתכונה זו נכונה רק בבסיס אורתונורמלי! אם הממ"פ היא המכפלה הפנימית הסטנדרטית אז אנו יודעים שהבסיס הסטנדרטי הוא אורתונורמלי, אבל הוא לא בהכרח אורתונורמלי עבור מכפלה פנימית כלשהי!

לצורך הוכחת המשפט, ראינו את הלמה הבאה:

למה: אם $T:V\to V$ אופרטור ליניארי כלשהו, ו-$B=\{v_1,...,v_n\}$ בסיס אורתונורמלי, אז

$$([T]_B)_{ij} = \langle Tv_j, v_i\rangle$$

ובעזרת הלמה הוכחנו את המשפט, וכעת נותר להוכיח את הלמה. נשים לב ש-ij מופיעים בסדר הפוך בכל אגף, זה נובע בגלל הטרנספוז של המטריצה המייצגת.

הוכחת הלמה: כיוון ש-B בסיס אורתונורמלי אז לכל $v\in V$

$$v = \sum_{i=1}^n \langle v, v_i\rangle v_i$$

בפרט

$$Tv_j = \sum_{i=1}^n \langle Tv_j, v_i\rangle v_i$$

בכך סיימנו, כי אנו רואים כיצד T פועל על איברי הבסיס! לכן לפי הגדרת מטריצה מייצגת

$$[T]_B = \begin{pmatrix} \langle Tv_1, v_1\rangle & \cdots \\ \langle Tv_1, v_2\rangle & \cdots \\ \vdots & \\ & \\ \langle Tv_1, v_n\rangle & \cdots \end{pmatrix}$$

כלומר $\left([T]_B\right)_{ij} = \langle Tv_j, v_i\rangle$. ■

<u>מסקנה</u>: יהי V ממ"פ ממימד סופי מעל שדה $\mathbb{F}$ ($\mathbb{R}$ או $\mathbb{C}$) ויהיו $S,T : V \to V$ אופרטורים ליניארים ו־$\alpha \in \mathbb{F}$ סקלר, אז:

$$(1) \quad (T+S)^* = T^* + S^*$$
$$(2) \quad (\alpha T)^* = \overline{\alpha} T^*$$
$$(3) \quad (T \circ S)^* = S^* \circ T^*$$
$$(4) \quad (T^*)^* = T$$

ואם T הפיך אז

$$(5) \quad \left(T^{-1}\right)^* = (T^*)^{-1}$$

ניתן להסיק מסקנה זו הן על ידי מעבר למטריצה מייצגת ושימוש בכך שהצמוד ההרמיטי הוא למעשה הצמוד הקופלקסי + טרנספוז, או על ידי שימוש בהגדרה של הצמוד ההרמיטי $\langle Tv, u\rangle = \langle v, T^*u\rangle$ ללא שימוש במטריצה מייצגת (נסו את שתי הדרכים כתרגיל).

<u>דוגמה</u>: נתון $V = \mathbb{R}^2$ עם המכפלה הפנימית הסטנדרטית, ונתון האופרטור $T : \mathbb{R}^2 \to \mathbb{R}^2$ המקיים

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix}$$

זהו למעשה אופרטור המבצע סיבוב ב־90° נגד כיוון השעון. חשבו את T^*.

<u>דרך א' ־ לפי הגדרה</u>: נסמן $\begin{pmatrix} z \\ w \end{pmatrix} = T^*\begin{pmatrix} a \\ b \end{pmatrix}$, עלינו למצוא את $\begin{pmatrix} z \\ w \end{pmatrix}$, אלא הנעלמים שלנו ויש למצוא אותם כפונקציה של a, b (ובכך למעשה נדע מהו T^*). לפי ההגדרה

$$\left\langle T\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} a \\ b \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} x \\ y \end{pmatrix}, T^*\begin{pmatrix} a \\ b \end{pmatrix} \right\rangle$$

כלומר

$$\left\langle \underbrace{\begin{pmatrix} -y \\ x \end{pmatrix}}_{T\begin{pmatrix} x \\ y \end{pmatrix}}, \begin{pmatrix} a \\ b \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} z \\ w \end{pmatrix} \right\rangle$$

מאחר וזו המכפלה הפנימית הסטנדרטית, נקבל:

$$-ya + xb = xz + yw$$

<u>נשווה מקדמים</u>:

מקואורדינטת x:

$$b=z$$

מקואורדינטת y:

$$-a=w$$

ובכך מצאנו את z,w כפונקציה של a,b. קיבלנו

$$T\begin{pmatrix}a\\b\end{pmatrix}=\begin{pmatrix}b\\-a\end{pmatrix}$$

כלומר זהו סיבוב עם השעון ב־90°.

הערה: נשים לב ש־$TT^{*}=T^{*}T=\mathbb{I}$

דרך ב' ־ לפי המשפט: נשים לב ש־$B=\left\{\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right\}$ הוא בסיס אורתונורמלי ביחס למכפלה הפנימית הנתונה, ולכן

$$[T^{*}]_{B}=[T]_{B}^{*}=\begin{pmatrix}0&-1\\1&0\end{pmatrix}^{*}=\begin{pmatrix}0&1\\-1&0\end{pmatrix}$$

ולכן

$$T^{*}\begin{pmatrix}a\\b\end{pmatrix}=\begin{pmatrix}b\\-a\end{pmatrix}$$

הערה: אם המכפלה הפנימית היא לא הסטנדרטית, אז בדרך הראשונה היינו משתמשים במכפלה המוגדרת בשאלה, ובדרך השניה היינו צריכים למצוא בסיס אורתונורמלי (עם תהליך גרם־שמידט).

הגדרות

V ממ"פ סוף־מימדי מעל שדה $\mathbb{F}$, $T:V\to V$ אופרטור ליניארי ו־$T^{*}:V\to V$ הצמוד שלו (תמיד קיים ויחיד, כי המרחב סוף־מימדי).

עבור $\mathbb{F}=\mathbb{C}$ נאמר ש:

* נקרא ל־T צמוד עצמית (או צ"ע) או הרמיטי אם $T=T^{*}$.
* T נקרא אוניטרי (Unitary, מהמילה Unit) אם $TT^{*}=T^{*}T=\mathbb{I}$. לכן אופרטור אוניטרי הוא בפרט הפיך.
* T נקרא נורמלי אם $TT^{*}=T^{*}T$. נשים לב שאם T אוניטרי הוא נורמלי, אבל ההפך אינו נכון, למשל $T\equiv 0$ הוא נורמלי ולא אוניטרי (כי אינו הפיך).

נשים לב שכל ההגדרות האלו תלויות במכפלה הפנימית הנתונה, וזאת מאחר וההגדרות מכילות את הצמוד, והצמוד מוגדרת ביחס למכפלה פנימית.

$A\in M_{n}\left(\mathbb{C}\right)$ נקראת:

* צמודה עצמית/הרמיטית אם $A=A^{*}$
* אוניטרית אם $AA^{*}=A^{*}A=\mathbb{I}$
* נורמלית אם $AA^{*}=A^{*}A$

באופן שקול: אם $T_{A}:\mathbb{C}^{n}\to\mathbb{C}^{n}$ המוגדר על ידי $T_{A}\left(x\right)=Ax$ הוא צ"ע/הרמיטי או אוניטרי או נורמלי ביחס למכפלה הפנימית הסטנדרטית של $\mathbb{C}^{n}$. למעשה כאן ה־$*$ (שהוא בעצם סימן לצמוד קומפלקסי + טרנספוז, כלומר $A^{*}=\overline{A}^{t}$) הוא ביחס למכפלה הפנימית הסטנדרטית על $\mathbb{C}^{n}$.

נרחיב שוב: כשאנו מסתכלים על אופרטור T, ההגדרות של הרמיטי, אוניטרי, ונורמלי הן ביחס למכפלה הפנימית הנתונה. לכן לא ברור מאליו שאם הוא אוניטרי (למשל) במכפלה פנימית אחת, הוא גם יהיה אוניטרי במכפלה פנימית אחרת. אבל כשמדובר במטריצות, הן בעצם מערך של מספרים, ובעצם ה־$*$ הוא ביחס למכפלה הפנימית הסטנדרטית.

עבור $\mathbb{F}=\mathbb{R}$ נאמר ש:

* T נקרא צמוד עצמית או סימטרי אם $T=T^{*}$
* T נקרא אורתוגונלי אם $TT^{*}=T^{*}T=\mathbb{I}$
* T נקרא נורמלי אם $TT^{*}=T^{*}T$

באופן דומה, מגדירים לגבי מטריצות:

$A\in M_{n}\left(\mathbb{R}\right)$ נקראת:

* צמודה עצמית/סימטרית אם $A=A^{t}$
* אורתוגונלית אם $AA^{t}=A^{t}A=\mathbb{I}$
* נורמלית אם $AA^{t}=A^{t}A$

באופן שקול: אם $T_{A}:\mathbb{R}^{n}\to\mathbb{R}^{n}$ המוגדר על ידי $T_{A}\left(x\right)=Ax$ הוא צ"ע/סימטרי או אורתוגונלי או נורמלי ביחס למכפלה הפנימית הסטנדרטית של $\mathbb{R}^{n}$.

הרצאה ־ 5 ליוני

ראינו שאם יש לנו ממ"פ סוף־מימדי אז לכל אופרטור ליניארי קיים ויחיד הצמוד שלו. הגדרנו מושגים רבים (אופרטור הרמיטי, אוניטרי, נורמלי, וכנ"ל עבור מטריצות). אחד הדברים שנעשה הוא להראות דמיון מסוים בין אופרטורים/מטריצות לבין מספרים מרוכבים, אנו יודעים שניתן אנחנו נראה שגם כל מטריצה הפיכה ניתן לכתוב בתור מטריצה חיובית לחלוטין כפול מטריצה אוניטרית.

דבר נוסף שראינו בקורס הוא הכללה של מושג הלכסינות: ראינו שאם אופרטור הוא לכסין אז יש לו בסיס מלכסן, ואחרת יש לו בסיס מג'רדן. אנו נראה כיום דבר נוסף: מתי לאופרטור יש בסיס מלכסן אורתונורמלי (ואנו נראה שזה שקול לכך שהאופרטור נורמלי).

הערות:

1) אם T צמוד עצמית או אוניטרי אז T נורמלי.

2) דוגמה לאופרטור צמוד עצמית שאינו אוניטרי: $T\equiv0$.

3) דוגמה לאורתוגונלי שאינו צמוד עצמית: $T=$סיבוב ב־90° (ראינו זאת בשיעור הקודם).

4) דוגמה לאופרטור נורמלי שאיננו צמוד עצמית: אם נעבוד מעל $\mathbb{R}$ ובעזרת מטריצות, אז בעצם אנו רוצים מטריצה שאיננה סימטרית אבל שכן תתחלף עם הטרנספוז שלה, דוגמה לכך היא המטריצה:

$$A=\begin{pmatrix}2 & -3\\
3 & 2
\end{pmatrix}$$

ולכן האופרטור המתאים הוא T_{A}.

תרגיל: יהי $T:V\to V$, כאשר V ממ"פ סוף־מימדי מעל $\mathbb{C}$. אזי הבאים שקולים:

א) T אוניטרי (כלומר $TT^{*}=T^{*}T=\mathbb{I}$)

ב) T הוא איזומורפיזם של V כממ"פ (כלומר T העתקה ליניארית, חח"ע ועל, ומקיימת $\left\langle Tv,Tu\right\rangle =\left\langle v,u\right\rangle$ לכל $v,u\in V$). את החח"ע ועל מוכיחים מיידית מכך ש־T הוא הפיך.

ג) T מעביר בסיס אורתונורמלי לבסיס אורתונורמלי.

מסקנה: $A\in M_{n}\left(\mathbb{C}\right)$ אוניטרית ($AA^{*}=A^{*}A=\mathbb{I}$) אמ"מ עמודות A הן בסיס אורתונורמלי ביחס למכפלה הפנימית הסטנדרטית ב־$\mathbb{C}^{n}$.

הרעיון: הבסיס הסטנדרטי $\left\{ \begin{pmatrix}1\\
0\\
\vdots\\
0
\end{pmatrix},\begin{pmatrix}0\\
1\\
\vdots\\
0
\end{pmatrix},\ldots,\begin{pmatrix}0\\
0\\
\vdots\\
1
\end{pmatrix}\right\}$ הוא בסיס אורתונורמלי של $\mathbb{C}^{n}$, ומתקיים ש־Ae_{i} הוא העמודה ה־i של A.

ניתן לנסח את התרגיל והמסקנה גם מעל השדה $\mathbb{R}$ באופן הבא:

תרגיל: יהי $T:V\to V$, כאשר V ממ"פ סוף־מימדי מעל $\mathbb{R}$. אזי הבאים שקולים:

א) T אורתוגונלי (כלומר $TT^*=T^*T=\mathbb{I}$)

ב) T הוא איזומורפיזם של V כממ"פ (כלומר T העתקה ליניארית, חח"ע ועל, ומקיימת $\langle Tv,Tu\rangle=\langle v,u\rangle$ לכל $v,u\in V$). את החח"ע ועל מוכיחים מיידית מכך ש־T הוא הפיך.

ג) T מעביר בסיס אורתונורמלי לבסיס אורתונורמלי.

מסקנה: $A\in M_n\left(\mathbb{R}\right)$ אורתוגונלית ($AA^t=A^tA=\mathbb{I}$) אמ"מ עמודות A הן בסיס אורתונורמלי ביחס למכפלה הפנימית הסטנדרטית ב־$\mathbb{R}^n$.

הערת העשרה: אם לוקחים את כל האופרטורים האוניטריים, מתקבלת חבורה. חבורה היא קבוצה שיש פעולה בין איבריה, ומקיימת סגירות, אסוציאטיביות, קיום הפכי, ואיבר נייטרלי. מטריצות אוניטריות הן בפרט הפיכות, ולכן מגדירים את $\mathcal{U}\left(n\right)$=חבורת מטריצות אוניטריות $n\times n$, והיא תת־חבורה של כל המטריצות ההפיכות $GL_n\left(\mathbb{C}\right)$. בצורה דומה מגדירים את חבורת המטריצות האורתוגונליות $n\times n$ אשר תסומן $\mathcal{O}\left(n\right)$, והיא תת חבורה של $GL_n\left(\mathbb{R}\right)$.

הערה אזהרה: יהא T אופרטור אוניטרי ־ זה ביחס למכפלה פנימית מסוימת של V. אם B בסיס אורתונורמלי של V אז המטריצה $[T]_B$ היא אוניטרית, אבל זה לא בהכרח נכון עבור בסיס שרירותי B, כלומר אם B סתם בסיס אז $[T]_B$ היא לאו דווקא אוניטרית כמטריצה, למרות ש־T אוניטרי כאופרטור.

תכונות אופרטור נורמלי

טענה: יהי V ממ"פ סוף־מימדי מעל $\mathbb{F}$ ($\mathbb{R}$ או $\mathbb{C}$), יהי $T:V\to V$ אופרטור ליניארי, אזי הבאים שקולים:

א) T נורמלי (כלומר $TT^*=T^*T$).

ב) T^* נורמלי.

ג) $\|T^*x\|=\|Tx\|$ לכל $x\in V$. תכונה זו שימושית מאוד, כפי שנראה בהמשך.

הוכחה:

ברור ש־א' $\Longleftrightarrow$ ב', נראה ש־א' $\Leftarrow$ ג': יהי $x\in V$, נחשב

$$\|Tx\|^2=\langle Tx,Tx\rangle=\langle x,T^*Tx\rangle=\langle x,TT^*x\rangle=\langle T^*x,T^*x\rangle=\|T^*x\|^2$$

נראה ש־ג' $\Leftarrow$ א': עם אותו חשבון נקבל

$$\forall x\in V\quad\langle x,T^*Tx\rangle=\langle x,TT^*x\rangle$$

נעביר אגף ונקבל

$$\forall x\in V\quad\langle x,T^*Tx-TT^*x\rangle$$

לכן קיבלנו שכל $x\in V$ ניצב לוקטור T^*Tx-TT^*x, ולכן זהו וקטור האפס, ומכאן שלכל x מתקיים

$$T^*Tx=TT^*x$$

ולכן מתקיים גם שוויון כאופרטורים

$$T^*T=TT^*$$

כלומר T נורמלי. ■

מסקנה: אם T אופרטור נורמלי אז

א) $\ker\left(T\right)=\ker\left(T^*\right)$

ב) $\text{Im}\,(T)=[\ker(T)]^{\perp}$ (ולכן בפרט $V=\ker T\oplus \text{Im}T$).

ג) $\text{Im}T=\text{Im}\,(T^*)$

מסקנה מהמסקנה: אם T נורמלי אז לכל $n\in\mathbb{N}$ מתקיים

$$\ker(T^n)=\ker((T^*)^n)$$
$$\text{Im}\,(T^n)=\text{Im}\,((T^*)^n)$$

הוכחת המסקנה:

א) עלינו להוכיח ש־$\ker(T)=\ker(T^*)$, אכן: $x\in\ker(T)\iff T(x)=0\iff \|Tx\|=0\iff \|T^*x\|=0\iff T^*x=0\iff x\in\ker(T^*)$.

ב) עלינו להוכיח ש־$\text{Im}\,(T)=[\ker(T)]^{\perp}$, נשים לב ש־$\dim\text{Im}T=\dim(\ker T)^{\perp}$, כי שניהם שווים ל־$\dim V-\dim\ker T$. ולכן מספיק להראות ש־$\text{Im}T\subset(\ker T)^{\perp}$. ניקח $y\in\text{Im}T$, כלומר יש $x\in V$ כך ש־$Tx=y$. ניקח $v\in\ker T$ ונראה ש־$\langle y,v\rangle=0$, ומכאן נסיק ש־$y\in(\ker T)^{\perp}$. אכן

$$\langle y,v\rangle=\langle Tx,v\rangle=\left\langle x,\underbrace{T^*v}_{=0}\right\rangle=0$$

כאשר השתמשנו בסעיף א' כדי להסיק מכך ש־$v\in\ker T$ שגם $v\in\ker T^*$, ולכן $T^*v=0$.

ג) מהסעיפים הקודמים נקבל

$$\text{Im}\,(T^*)=(\ker T^*)^{\perp}=(\ker T)^{\perp}=\text{Im}T$$

הוכחת המסקנה מהמסקנה: בגלל ש־$(ST)^*=S^*T^*$ אנו יודעים שמתקיים גם $(T^*)^n=(T^n)^*$, מכך ש־T נורמלי אנו מבינים ש־T ו־T^* מתחלפים, ולכן גם חזקותיהם מתחלפות, לכן T^n נורמלי ולכן סעיפים א',ב',ג' תקפים גם עבורו.

כעת אנו הולכים לשלב את שני החלקים של הקורס (בערך), ראינו כמה תכונות של אופרטור נורמלי, ואנו נאפיין אותו כעת.

הגדרה: יהי $T:V\to V$, כאשר V ממ"פ מעל $\mathbb{C}$. נאמר ש־T לכסין אוניטרית אם קיים בסיס B אורתונורמלי של וקטורים עצמיים של T, כך ש־$[T]_B=D$ אלכסונית.

בדומה, עבור מטריצה: תהי $A\in M_n(\mathbb{C})$, נאמר ש־A לכסינה אוניטרית אם קיימת P הפיכה, אוניטרית (כלומר $P^*=P^{-1}$) כך ש־$P^*AP=P^{-1}AP=D$ אלכסונית (זה נקרא דמיון אוניטרי).

ניתן להגדיר בצורה דומה גם מעל השדה $\mathbb{R}$, באופן הבא:

הגדרה: יהי $T:V\to V$, כאשר V ממ"פ מעל $\mathbb{R}$. נאמר ש־T לכסין אורתוגונלית אם קיים בסיס B אורתונורמלי של וקטורים עצמיים של T, כך ש־$[T]_B=D$ אלכסונית.

בדומה, עבור מטריצה: תהי $A\in M_n(\mathbb{R})$, נאמר ש־A לכסינה אורתוגונלית אם קיימת P הפיכה, אורתוגונלית (כלומר $P^t=P^{-1}$) כך ש־$P^tAP=P^{-1}AP=D$ אלכסונית (זה נקרא דמיון אורתוגונלי).

משפט: יהי V ממ"פ סוף מימדי מעל $\mathbb{C}$, יהי $T:V\to V$ אופרטור ליניארי. אז T לכסין אוניטרית $\iff$ T נורמלי.

הערה: ברגע שנוכיח משפט זה, נוכל להשתמש בכל המסקנות שהכרנו עבור אופרטור לכסין (למשל משפט הפירוק הספקטרלי) עבור אופרטור נורמלי. היתרון הפעם הוא שקל יהיה לחשב את ההטלות, כי יש לנו מכפלה פנימית.

הוכחה:

כיוון ראשון $\Leftarrow$: מהנתון יש בסיס $B=\{v_1,...,v_n\}$ אורתונורמלי ביחס למכפלה הפנימית הנתונה, של וקטורים עצמיים v_i. נסמן ב־λ_i את הערך העצמי המתאים ל־v_i (ייתכנו חזרות), ונקבל ש־

$$[T]_B=\begin{pmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_n\end{pmatrix}\Rightarrow [T^*]_B\underbrace{=}_{B\text{ is orthonormal}}[T]_B^*=\begin{pmatrix}\overline{\lambda_1}&&&\\&\overline{\lambda_2}&&\\&&\ddots&\\&&&\overline{\lambda_n}\end{pmatrix}$$

קיבלנו ש־2 המטריצות המייצגות מתחלפות (כי שתיהן אלכסוניות) ולכן T ו־T^* מתחלפים, כלומר T נורמלי.

כיוון שני $\Leftarrow$: נוכיח באינדוקציה על $n = \dim V$. בסיס האינדוקציה $n = 1$ הוא טריוויאלי.

<u>הנחת האינדוקציה</u>: נניח שלכל ממ"פ V מעל $\mathbb{C}$ עם $\dim V < n$ מתקיים שאם $T : V \to V$ נורמלי, אז T לכסין אוניטרית.

<u>צעד האינדוקציה</u>: יהי V ממ"פ מעל $\mathbb{C}$ עם $\dim V = n$, ויהי $T : V \to V$ אופרטור נורמלי, יש להראות ש־T לכסין אוניטרית. יהי λ ערך עצמי של T (יש כזה כי אנו עובדים מעל $\mathbb{C}$ - שהוא סגור אלגברית), נסמן את המרחב העצמי של λ:

$$W = \ker\left(T - \lambda \mathbb{I}\right)$$

ניקח בסיס של W ונפעיל עליו את אלגוריתם גרם־שמידט (כולל הנרמול), אז נקבל בסיס B_1 של ו"ע (כולם של λ) שהוא אורתונורמלי. אם $V = W$, אז סיימנו, כי אז קיבלנו בסיס אורתונורמלי מלכסן, ולמעשה המסקנה במקרה זה היא שהאופרטור הוא סקלרי.

<u>הערת ביניים חשובה</u>: למעשה אם $V = W \oplus W^{\perp}$ אז זהו פירוק למרחבים T שמורים, כי המרחב העצמי W תמיד T שמור ועבור אופרטור נורמלי אם W הוא T-שמור אז גם $W^{\perp}$ הוא T-שמור. נוכיח זאת: T מתחלף עם T^* ולכן מתרגיל בית נקבל ש־W הוא T^*-שמור, ואז בהינתן $w \in W$ ו־$v \in W^{\perp}$ נקבל

$$\langle Tv, w\rangle = \left\langle \underbrace{v}_{\in W^{\perp}}, \underbrace{T^*w}_{\in W} \right\rangle = 0$$

לכן לקחנו $v \in W^{\perp}$ וקיבלנו ש־$Tv \in W^{\perp}$, ולכן $W^{\perp}$ הוא T-שמור.

<u>המשך הוכחה</u>: בהנחה ש־$W \neq V$, נביט על $V = W \oplus W^{\perp}$, ונביט על הצמצום $S = T|_{W^{\perp}}$ (אז S אכן מוגדר כי $W^{\perp}$ הוא T-שמור). וודאו ש־$S^* = T^*|_{W^{\perp}}$ (נובע ישירות מההגדרה של צמצום), לפיכך S מתחלף עם S^* (כי T התחלף עם T^*), כלומר S נורמלי, $\dim W^{\perp} < n$ ולכן לפי הנחת האינדוקציה יש בסיס B_2 אורתונורמלי של וקטורים עצמיים של S. כל וקטור עצמי של S הוא גם וקטור עצמי של T, ולכן בסך הכל מאחר ו־$V = W \oplus W^{\perp}$ נקבל ש־$B = B_1 \cup B_2$ הוא בסיס אורתונורמלי למרחב V של וקטורים עצמיים של T (נשים לב שכל האיברים ב־B_2 תמיד ניצבים לאיברים ב־B_1, לכן הבסיס אורתונורמלי). $\Leftarrow$ T לכסין אוניטרית. ■

<u>מסקנה</u>: ניתן להפעיל את משפט הפירוק הספקטרלי שראינו בעבר ל־T נורמלי. נסו זאת, שימו לב שניתן לעבוד בצורה יותר קלה כעת, מאחר וחישוב ההטלות קל יותר בממ"פ בגלל שיש לנו מכפלה פנימית. נזכיר את המשפט:

$$\begin{aligned} V &= \ker\left(T - \lambda_1 \mathbb{I}\right) \oplus \ldots \oplus \ker\left(T - \lambda_n \mathbb{I}\right) \\ T &= \sum \lambda_i P_i \end{aligned}$$

בעבר היינו צריכים לחשב כל הטלה בנפרד, למשל בעזרת פולינומי אינטרפולציה של לגרנז'. אך עבור אופרטור נורמלי, נשים לב שניתן להביט על ההטלות כאורתוגונליות. כפי שראינו בהוכחה, יש מרחב עצמי W ומרחב $W^{\perp}$, ולכן אם P ההטלה על W במקביל ל־$W^{\perp}$ הרי שהיא הטלה אורתוגונלית, וניתן להשתמש בכך.

<u>מסקנות</u>:

1) אם T נורמלי אז וקטורים עצמיים השייכים לערכים עצמיים שונים הם אורתוגונליים. (ראינו זאת בהוכחה, אם W הוא מרחב עצמי של ערך עצמי λ, אז וקטורים עצמיים של ע"ע אחר הם ב־$W^{\perp}$, ולכן ניצבים לכל וקטור ב־W).

2) אם $A \in M_n\left(\mathbb{C}\right)$ נורמלית, אז A לכסינה אוניטרית.

<u>טענה</u>: יהי V ממ"פ מעל $\mathbb{C}$ ו־$T : V \to V$ צמוד עצמית, אז הערכים העצמיים של T ממשיים.

<u>הוכחה</u>: יהי λ ערך עצמי עם וקטור עצמי $0 \neq v \in V$, כלומר $Tv = \lambda v$. לכן

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

לכן $\lambda = \overline{\lambda}$ ומכאן ש־λ ממשי.

<u>הרצאה -11 ליוני</u>

מסקנה: אם V מרחב מכפלה פנימית סוף מימדי מעל $\mathbb{R}$ ו־$T: V \to V$ אופרטור ליניארי, אז T לכסין אורתוגונלית אמ"מ T צמוד עצמית.

בפרט, אם $A \in M_n(\mathbb{R})$ מטריצה סימטרית אז A לכסינה אורתוגונלית, כלומר דומה אורתוגונלית למטריצה אלכסונית. כלומר קיימת P אורתוגונלית כך ש־

$$P^t AP = P^{-1}AP = D$$

כאשר עמודותיה של P הן בסיס אורתונורמלי של ו"ע, המורכב מבסיס אורתונורמלי של המצבים העצמיים (מרחב עצמי), ו־D היא אלכסונית עם ערכים עצמיים באלכסונה.

מתקיים עבור A משפט הפירוק הספקטרלי:

$$A = \sum_{i=1}^{t} \lambda_i Q_i$$

כאשר Q_i היא ההטלה האורתוגונלית על המרחב העצמי V_{λ_i} ובפרט $Q_i = Q_i^*$ (קחו בסיס אורתונורמלי וחשבו).

במפורש ־ כיצד נמצא את Q_i? נניח ש־$\{v_1, ..., v_{m_i}\}$ בסיס אורתונורמלי של V_{λ_i} (ביחס למכפלה הסטנדרטית על $\mathbb{R}^n$), אזי:

$$Q_i = v_1 v_1^t + v_2 v_2^t + ... + v_{m_i} v_{m_i}^t$$

זאת משום ש־

$$Q_i(e_1) = \sum_{i=1}^{m_i} \langle e_1, v_i \rangle v_i$$

ונשים לב ש־$Q_i(e_1)$ היא בעצם העמודה ה־1 של Q_i, ו־$\langle e_1, v_i \rangle$ הוא הקואורדינטה ה־1 של v_i, וכן הלאה עבור העמודות האחרות.

דוגמה: נביט במטריצה $A = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}$, אזי $\lambda_1 = 5, \lambda_2 = -1$ הם הערכים העצמיים, הוקטורים העצמיים (המנורמלים) הם $v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ בהתאמה, ולכן המרחבים העצמיים הם:

$$V_{\lambda_1} = \mathrm{span}\left\{\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$$
$$V_{\lambda_2} = \mathrm{span}\left\{\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}\right\}$$

וההטלות הן:

$$Q_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$
$$Q_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \end{pmatrix} = \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix}$$

והפירוק הספקטרלי הינו:

$$A = 5 \cdot Q_1 - 1 \cdot Q_2$$

תזכורת: יהי $\mathbb{F}$ ($\mathbb{C}$ או $\mathbb{R}$), עבור $A \in M_n(\mathbb{F})$ הגדרנו ש־A חיובית לחלוטין אם $A = A^*$ ו־$x^*Ax > 0$ לכל $\mathbb{F}^n \ni x \neq \vec{0}$.

הגדרה: נאמר ש־$A \in M_n(\mathbb{F})$ אי־שלילית אם $A = A^*$ ו־$x^*Ax \geq 0$ לכל $\mathbb{F}^n \ni x \neq \vec{0}$.

<u>טענה</u>: תהי $A=A^*$, אז הבאים שקולים:

א. A אי־שלילית (חיובית לחלוטין)

ב. הערכים העצמיים של A הם אי־שליליים (חיוביים ממש)

ג. יש מטריצה הרמיטית C (הפיכה) כך ש־$A=C^2$.

ד. יש מטריצה B (הפיכה) כך ש־$A=B^*B$.

<u>רעיון ההוכחה</u>:

א' ⇐ ב': יהי λ ערך עצמי, אז יש $x\neq\vec{0}$ עבורו $Ax=\lambda x$. נתון ש־$0\le x^*Ax$ ולכן נקבל:

$$0\le x^*Ax=x^*\lambda x=\lambda x^*x=\lambda\left\|x\right\|^2$$

ולכן $\lambda\ge 0$. במעבר האחרון השתמשנו במכפלה הפנימית הסטנדרטית ב־$\mathbb{F}^n$ ובכך ש־$\|x\|$ חיובי ממש (כי $x\neq\vec{0}$).

ב' ⇐ ג': יהי $A=\sum_{i=1}^t\lambda_iQ_i$ הפירוק הספקטרלי של A, כאשר מהנתון λ_i אי־שליליים וכמו כן Q_i הן הטלות אורתוגונליות והרמיטיות ($Q_i^*=Q_i$). נגדיר

$$C=\sum_{i=1}^t\sqrt{\lambda_i}Q_i$$

אזי C הרמיטית כסכום של כאלו ומתקיים $A=C^2$ כי $Q_iQ_j=\delta_{ij}Q_i$, כאשר נזכיר כי $\delta_{ij}=\begin{cases}1 & i=j\\0 & i\neq j\end{cases}$, להלן החישוב:

$$C^2=C\cdot C=\left(\sum_{i=1}^t\sqrt{\lambda_i}Q_i\right)\left(\sum_{j=1}^t\sqrt{\lambda_j}Q_j\right)=\sum_{i,j=1}^t\sqrt{\lambda_i\lambda_j}Q_iQ_j=$$
$$=\sum_{i,j=1}^t\sqrt{\lambda_i\lambda_j}\delta_{ij}Q_i=\sum_{i=1}^t\sqrt{\lambda_i\lambda_i}Q_i=\sum_{i=1}^t\lambda_iQ_i=A$$

ג' ⇐ד': נגדיר $B=C$, מאחר ו־C הרמיטית התוצאה מתקבלת מיידית:

$$A=C^2=C^*C=B^*B$$

ד' ⇐א': מהמכפלה הפנימית הסטנדרטית ב־$\mathbb{F}^n$ נקבל

$$x^*Ax=\underbrace{x^*B^*}_{y^*}\underbrace{Bx}_{y}=y^*y=\left\|y\right\|^2\ge 0$$

ולכן A אי־שלילית.

<u>הערה</u>: ניתן לראות ש־C בסעיף ג' נקבעת באופן יחיד, ואי־שלילית בעצמה. קוראים לה <u>השורש של A</u>.

<u>הערה נוספת</u>: תהי $A=A^*$. אז A אי־שלילית (חיובית לחלוטין) $\iff$ המינורים הראשיים שלה הם אי־שליליים (חיוביים ממש)

<u>סיכומון</u>: תהי $A\in M_n\left(\mathbb{F}\right)$ מטריצה <u>נורמלית</u>. אזי

- A הרמיטית $\iff$ הערכים העצמיים שלה ממשיים
- A אי־שלילית $\iff$ הערכים העצמיים ממשיים ואי־שליליים
- A אוניטרית $\iff$ הערכים העצמיים שלה על מעגל היחידה ($|z|=1$)

<u>משפט (פירוק פולרי)</u>: תהי $A\in M_n(\mathbb{F})$, אז קיימות U אוניטרית ו־R אי־שלילית כך ש־$A=UR$. כמו כן, אם A הפיכה אז R חיובית לחלוטין.

<u>הוכחה</u>: נוכיח ל־A הפיכה, ראינו כי A^*A תמיד אי־שלילית, ולכן קיימת R אי־שלילית שהיא השורש שלה: $R^2=A^*A$. אם A הפיכה אז R הפיכה ובפרט חיובית לחלוטין ונוכל להגדיר $U=AR^{-1}$.

נראה כעת ש־U אוניטרית, ראשית נחשב את U^*:

$$U^*=\left(AR^{-1}\right)^*=\left(R^{-1}\right)^*A^*=\left(R^*\right)^{-1}A^*\underbrace{=}_{R^*=R}R^{-1}A^*$$

ואז

$$U^*U=R^{-1}A^*AR^{-1}=R^{-1}R^2R^{-1}=\underbrace{R^{-1}R}_{\mathbb{I}}\underbrace{RR^{-1}}_{\mathbb{I}}=\mathbb{I}$$

$$UU^*=AR^{-1}R^{-1}A^*=A\left(R^2\right)^{-1}A^*=A\left(A^*A\right)^{-1}A^*=\underbrace{AA^{-1}}_{\mathbb{I}}\underbrace{\left(A^*\right)^{-1}A^*}_{\mathbb{I}}=\mathbb{I}$$

ולכן U אוניטרית, ובכך סיימנו את ההוכחה למקרה של A הפיכה. ■

<u>הרצאה - 12 ליוני</u>

תבניות ביליניאריות

<u>הגדרה</u>: יהי $\mathbb{F}$ שדה כלשהו, ויהי V מרחב וקטורי מעל $\mathbb{F}$. פונקציה $f:V\times V\to\mathbb{F}$ נקראת <u>תבנית ביליניארית</u> אם מתקיימות התכונות הבאות:

1. $f(v+w,u)=f(v,u)+f(w,u)$
2. $f(\alpha v,u)=\alpha f(v,u)$
3. $f(v,w+u)=f(v,w)+f(v,u)$
4. $f(v,\alpha w)=\alpha f(v,w)$

כלומר זוהי פונקציה המקיימת ליניאריות בכל רכיב בנפרד, ומכאן השם "תבנית ביליניארית".

<u>דוגמה</u>: אם V מרחב מכפלה פנימית מעל $\mathbb{R}$ אז המכפלה הפנימית היא תבנית ביליניארית.

<u>טענה</u>: יהי V מרחב וקטורי סוף־מימדי מעל $\mathbb{F}$, יהי $B=\{b_1,b_2,...,b_n\}$ בסיס של V ותהי $M\in M_n(\mathbb{F})$, אז הפונקציה $f:V\times V\to\mathbb{F}$ המוגדרת באופן הבא:

$$\forall v,w\in V\quad f(v,w)=\underbrace{[v]_B^t}_{1\times n}\underbrace{M}_{n\times n}\underbrace{[w]_B}_{n\times 1}$$

היא תבנית ביליניארית. יתר על כן, בהינתן תבנית ביליניארית כלשהי $g:V\times V\to\mathbb{F}$ נגדיר

$$A_{ij}=g(b_i,b_j)$$

ואז

$$g(v,w)=[v]_B^tA[w]_B$$

כלומר $A\in M_n(\mathbb{F})$, והיא תלויה בבסיס הנתון בתבנית g הנתונה. נסמן לכן $A=[g]_B$, שהיא המטריצה המייצגת של התבנית g לפי הבסיס B.

דוגמה: נביט במרחב הוקטורי $V=\mathbb{R}^2$ מעל השדה $\mathbb{F}=\mathbb{R}$, ובבסיסים הבאים ל-V:

$$B_1=\left\{\underbrace{\begin{pmatrix}1\\0\end{pmatrix}}_{e_1},\underbrace{\begin{pmatrix}0\\1\end{pmatrix}}_{e_2}\right\},\quad B_2=\left\{\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}1\\1\end{pmatrix}\right\}$$

ונתונה התבנית הבילינארית $g:V\times V\to\mathbb{F}$ הבאה

$$g\left(\begin{pmatrix}x_1\\x_2\end{pmatrix},\begin{pmatrix}y_1\\y_2\end{pmatrix}\right)=2x_1y_1+3x_1y_2+4x_2y_2$$

אזי

$$A=[g]_{B_1}=\begin{pmatrix}2&3\\0&4\end{pmatrix},\quad B=[g]_{B_2}=\begin{pmatrix}2&5\\2&9\end{pmatrix}$$

כעת נשאלת השאלה: האם יש קשר בין המטריצות? שאלה זו מובילה אותנו להגדרה הבאה.

הגדרה: תהיינה $A,B\in M_n(\mathbb{F})$, נאמר ש-A חופפת ל-B אם קיימת P הפיכה כך ש-$B=P^tAP$.

שימו לב: באופן כללי חפיפה $\neq$ דמיון מטריצות. אם A דומה אורתוגונלית ל-B אז A חופפת ל-B (תזכורת: A דומה אורתוגונלית ל-B אם קיימת P אורתוגונלית, כלומר P המקיימת $P^t=P^{-1}$, כך ש-$P^{-1}AP=B$).

אצלנו בדוגמה: P המקיימת $B=P^tAP$ היא מטריצת המעבר בין בסיסים

$$P=\begin{pmatrix}1&1\\0&1\end{pmatrix}$$

משפט: יהי מרחב וקטורי $V=\mathbb{F}^n$ מעל שדה $\mathbb{F}$, יהיו $A,B\in M_n(\mathbb{F})$, אז A,B חופפות $\iff$ הן מייצגות את אותה תבנית בילינארית בבסיסים שונים.

הוכחה:

כיוון ראשון $\Leftarrow$: נתון ש-A,B חופפות, כלומר קיימת P הפיכה כך ש-$B=P^tAP$. נסמן ב-B_1 את הבסיס הסטנדרטי, וב-B_2 את הבסיס המורכב מעמודות P (זהו אכן בסיס כי P הפיכה). נגדיר $f:V\times V\to\mathbb{F}$ באופן הבא:

$$f(v,w)=v^tAu$$

זו תבנית בילינארית, ומתקיים

$$\begin{aligned}[f]_{B_1}&=A\\ [f]_{B_2}&=B\end{aligned}$$

וודאו את החשבון המראה מדוע $[f]_{B_2}=B$.

כיוון שני $\Rightarrow$: נניח ש-$B_1=\{x_1,x_2,...,x_n\}$, $B_2=\{y_1,y_2,...,y_n\}$ הם בסיסים של V, ו-f תבנית בילינארית כך ש-$[f]_{B_1}=A,\ [f]_{B_2}=B$. אז נגדיר את P להיות מטריצת המעבר בין הבסיסים, כלומר

$$\forall v\in V\quad [v]_{B_1}=P[v]_{B_2}$$

כלומר

$$y_j=\sum_{i=1}^{n}P_{ij}x_i$$

ומתקיים

$$B=P^tAP$$

■

<u>מסקנה:</u> $A\in M_n\left(\mathbb{R}\right)$ חופפת למטריצת היחידה $\iff$ A חיובית לחלוטין (כלומר $A=A^t$ ו־$x^t\vec{A}x>0$ $,\forall\vec{x}\neq 0$).

<u>הוכחה:</u>

כיוון ראשון $\Leftarrow$: מהנתון, יש P הפיכה כך ש־$A=P^t\mathbb{I}P=P^tP$. ראינו בשיעור הקודם בתנאים השקולים שזהו בדיוק תנאי שקול למטריצה חיובית לחלוטין.

כיוון שני $\Rightarrow$: ראינו שאם A חיובית לחלוטין אז A מגדירה מכפלה פנימית על $\mathbb{R}^n$. מכפלה פנימית זו היא בפרט תבנית בילינארית. כעת ניקח בסיס אורתונורמלי B עבור מכפלה פנימית זו ונקבל שהמטריצה המייצגת של המכפלה הפנימית ביחס ל־B היא $\mathbb{I}_{n\times n}$, ולכן על פי המשפט הקודם A ו־$\mathbb{I}$ חופפות.

<u>מטרה:</u> נרצה לענות על השאלה: מתי שתי מטריצות הן חופפות? נענה עבור המקרה שבו התבניות הבילינאריות הן סימטריות.

<u>הגדרה:</u> נאמר שתבנית בילינארית $f:V\times V\to\mathbb{F}$ היא <u>סימטרית</u> אם $f\left(v,u\right)=f\left(u,v\right)$.

<u>תנאי שקול:</u> לכל בסיס B, $\left[f\right]_B$ היא מטריצה סימטרית.

<u>הסכם:</u> מעתה נעבוד עם מרחב וקטורי V ממימד n מעל שדה $\mathbb{F}$, אשר לו קריקטריסטה שונה מ־2, כלומר $\mathrm{Char}\left(\mathbb{F}\right)\neq 2$. הקריקטריסטה של שדה מציינת את מספר הפעמים המינימלי שעלינו לחבר את איבר היחידה לכפל, על מנת לקבל את האיבר הנייטרלי לחיבור. למשל, ב־$\mathbb{Z}_5$ מתקיים $1+1+1+1+1=0$, ולכן $\mathrm{Char}(\mathbb{Z}_5)=5$. מאידך, בשדה הממשיים הסכום $1+1+1+....$ לעולם לא יתאפס, ולכן מגדירים $\mathrm{Char}\left(\mathbb{R}\right)=0$.

<u>הגדרה:</u> נאמר שפונקציה $q:V\to\mathbb{F}$ היא תבנית ריבועית אם q היא פולינום הומוגני ב־n משתנים ממעלה 2, כלומר:

$$a_{ij}\in\mathbb{F}\quad q\left(x_1,...,x_n\right)=\sum_{1\leq i,j\leq n}a_{ij}x_ix_j$$

נשים לב שכיוון ש־$\mathrm{Char}\mathbb{F}\neq 2$ אז ניתן לרשום

$$q\left(x_1,...,x_n\right)=\sum_{1\leq i,j\leq n}a_{ij}x_ix_j=\sum_{i=1}^{n}\underbrace{a_{ii}}_{=b_{ii}}x_i^2+\sum_{1\leq i<j\leq n}\underbrace{\left(a_{ij}+a_{ji}\right)}_{b_{ij}}x_ix_j=$$

$$=\begin{pmatrix}x_1, & \dots & ,x_n\end{pmatrix}\begin{pmatrix}b_{11} & \frac{b_{12}}{2} & \frac{b_{13}}{2} & \dots & \frac{b_{1n}}{2}\\ \frac{b_{12}}{2} & b_{22} & & \dots & \\ \frac{b_{13}}{2} & & b_{33} & \dots & \\ \vdots & \vdots & \vdots & \ddots & \\ \frac{b_{1n}}{2} & & & & b_{nn}\end{pmatrix}\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=x^tBx\quad(*)$$

כאשר B מטריצה סימטרית.

<u>דוגמה:</u> נביט בתבנית הריבועית $q:\mathbb{R}^3\to\mathbb{R}$ הבאה:

$$q\left(x_1,x_2,x_3\right)=2x_1^2+3x_1x_2+4x_1x_3+5x_3^2$$

אז ניתן לרשום אותה בצורה הבאה:

$$q\left(x_1,x_2,x_3\right)=\begin{pmatrix}x_1 & x_2 & x_3\end{pmatrix}\begin{pmatrix}2 & \frac{3}{2} & 2\\ \frac{3}{2} & 0 & 0\\ 2 & 0 & 5\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}$$

הקשר בין תבניות בילינאריות סימטריות לתבניות ריבועיות

- בהינתן $f:V\times V\to\mathbb{F}$ תבנית בילינארית סימטרית, נגדיר $q_f:V\to\mathbb{F}$ על ידי

$$q_f\left(\vec{x}\right)=f\left(\vec{x},\vec{x}\right)$$

וזוהי תבנית ריבועית.

- בהינתן $q:V\to\mathbb{F}$ תבנית ריבועית, נגדיר $f_q:V\times V\to\mathbb{F}$ על ידי

$$f_q\left(\vec{x},\vec{y}\right)=\frac{1}{4}q\left(x+y\right)-\frac{1}{4}q\left(x-y\right)$$

זוהי תבנית בילינארית סימטרית, והיא נקראת זהות הפולריזציה. מתקיים ששתי התאמות אלו הן הופכיות אחת לשניה.

הערה: מטריצה מייצגת של תבנית ריבועית q היא המטריצה המייצגת של התבנית הבילינארית f_q.

לצורך מעשי: להשתמש בדרך שראינו בדוגמה האחרונה או באופן כללי ב־$(*)$.

הרצאה ־ 18 ליוני

תזכורת: $\mathbb{F}$ שדה עם $\operatorname{char}\mathbb{F}=0$ (למשל $\mathbb{Q}$, $\mathbb{R}$, וכו'), המרחב הוקטורי הוא $V=\mathbb{F}^n$. תבנית ריבועית היא פונקציה $q:V\to\mathbb{F}$ מהצורה

$$q\left(\vec{x}\right)=x^tAx$$

כאשר $A\in M_n\left(\mathbb{F}\right)$ היא מטריצה סימטרית. סימון נוסף שהיה לנו הוא $[q]_E=A$, כלומר A היא המטריצה המייצגת לפי הבסיס הסטנדרטי E.

הגדרנו: $A,B\in M_n\left(\mathbb{F}\right)$ חופפות אם יש P הפיכה כך ש־$B=P^tAP$, והוכחנו: A,B חופפות $\Longleftrightarrow$ הן מייצגות אותה תבנית בילינארית בבסיסים שונים.

המסקנה: A,B סימטריות חופפות $\Longleftrightarrow$ הן מייצגות את אותה תבנית בילינארית סימטרית בבסיסים שונים, כלומר הן מייצגות את אותה תבנית ריבועית בבסיסים שונים.

טרמינולוגיה: נאמר ששתי תבניות ריבועיות q_1,q_2 הן שקולות אם קיים ביניהן שינוי משתנים ליניארי והפיך, כלומר קיים $T:V\to V$ אופרטור ליניארי הפיך כך ש־$\vec{x}=T\vec{y}$ ומתקיים

$$q_1\left(\vec{x}\right)=q_2\left(\vec{y}\right)$$

דוגמה לשינוי משתנים ליניארי: ניקח

$$q_1\left(x_1,x_2,x_3\right)=x_1^2+2x_2x_3+x_2^2$$

ונבצע שינוי משתנה ליניארי והפיך

$$\begin{aligned}x_1&=2y_1\\x_2&=y_2\\x_3&=y_1+y_3\end{aligned}$$

וכך נקבל את $q_2\left(\vec{y}\right)$.

הערה: נשים לב ש־q_1,q_2 הן שקולות אמ"ם המטריצות המייצגות $[q_1]_E,[q_2]_E$ חופפות. נוכיח זאת, נסמן $[q_1]_E=A,[q_2]_E=B$, ואז מתקיים

$$\begin{aligned}q_1\left(x\right)&=x^tAx\\q_2\left(y\right)&=y^tBy\end{aligned}$$

מאחר ויש שינוי משתנים הפיך $x=Py$, נקבל:

$$q_1(x)=x^tAx=(Py)^t APy=y^tP^tAPy=y^tBy=q_2(y)$$

מכאן ש־A,B חופפות אמ"ם יש P הפיכה כך ש־$B=P^tAP$ אמ"ם יש P הפיכה כך ש־$x=Py$ ו־$q_1(x)=q_2(y)$ אמ"ם q_1,q_2 שקולות.

<u>מטרה</u>: אנו רוצים לדעת מתי שתי תבניות ריבועיות שקולות, כלומר מתי שתי מטריצות <u>סימטריות</u> הן חופפות. לשם כך עלינו להשוות ביניהן איכשהו, בהינתן שתי מטריצות סימטריות, מה ניתן לעשות כדי להשוות ביניהן? אם אנחנו מעל הממשיים, אנו יודעים שכל מטריצה סימטרית היא לכסינה אורתוגונלית, אבל זה לא נכון עבור כל שדה (אפילו לא עבור שדה המרוכבים).

<u>דוגמה ־ שיטה 1 כדי להגיע לצורה אלכסונית ־ השלמה לריבוע</u>: נביט בתבנית הריבועית

$$q_1(x_1,x_2,x_3)=-3x_1^2+x_2^2-2x_2x_3+3x_3^2$$

המטריצה המייצגת שלה היא

$$[q_1]_E=\begin{pmatrix}-3 & 0 & 0\\ 0 & 1 & -1\\ 0 & -1 & 3\end{pmatrix}$$

נרצה להביא תבנית ריבועית/מטריצה סימטרית לצורה אלכסונית, ובמקרה זה משמעות הדבר היא להיפטר מהאיבר המעורב x_2x_3 ולהישאר רק עם ריבועים של המשתנים x_1,x_2,x_3. נעשה זאת על ידי השלמה לריבוע

$$q_1(x_1,x_2,x_3)==-3x_1^2+(x_2-x_3)^2+2x_3^2$$

וכעת ניתן להגדיר החלפת משתנים ליניארית והפיכה

$$\begin{aligned}x_1&=y_1\\ x_2-x_3&=y_2\\ x_3&=y_3\end{aligned}$$

ולקבל צורה אלכסונית. אם נניח שאנחנו מעל $\mathbb{R}$ אז ניתן לשפר מעט את הצורה באופן הבא:

$$q_1(x_1,x_2,x_3)=-\left(\sqrt{3}x_1\right)^2+(x_2-x_3)^2+\left(\sqrt{2}x_3\right)^2$$

ואז להגדיר החלפת משתנה ליניארית והפיכה שונה:

$$\begin{aligned}y_1&=\sqrt{3}x_1\\ y_2&=x_2-x_3\\ y_3&=\sqrt{2}x_3\end{aligned}$$

ואז נקבל:

$$q_2(y_1,y_2,y_3)=-y_1^2+y_2^2+y_3^2$$

ובכך קיבלנו צורה אלכסונית, המטריצה המייצגת היא

$$[q_2]_E = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

כמובן ששיטה זו (השלמה לריבוע, או Lagrange Reduction) היא די פרטית, ולא ברור האם היא תעבוד תמיד, או מה צריך להשלים לריבוע. כלומר אנו רוצים איזושהי שיטה אלגוריתמית שנדע כיצד להשתמש בה תמיד, וזה מוביל אותנו לשיטה השניה.

<u>שיטה 2 - חפיפת מטריצות - ביצוע פעולה אלמנטרית על השורות ואז ביצוע אותה פעולה על העמודות כדי להגיע לצורה אלכסונית</u>

למשל, נביט במטריצה

$$A = \begin{pmatrix} -1 & 1 & 2 \\ 1 & 1 & -1 \\ 2 & -1 & 3 \end{pmatrix}$$

אנו רוצים להביאה לצורה אלכסונית, נבצע פעולות אלמנטריות על שורה ועמודה לסירוגין, חשוב לבצע על העמודה את אותה פעולה שביצענו על השורה, כדי לשמור על הסימטריות.

$$\begin{pmatrix} -1 & 1 & 2 \\ 1 & 1 & -1 \\ 2 & -1 & 3 \end{pmatrix} \xrightarrow{R_2 \to R_2 + R_1} \begin{pmatrix} -1 & 1 & 2 \\ 0 & 2 & 1 \\ 2 & -1 & 3 \end{pmatrix} \xrightarrow{C_2 \to C_2 + C_1} \begin{pmatrix} -1 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 1 & 3 \end{pmatrix}$$

$$\xrightarrow{R_3 \to R_3 + 2R_1} \begin{pmatrix} -1 & 0 & 2 \\ 0 & 2 & 1 \\ 0 & 1 & 7 \end{pmatrix} \xrightarrow{C_3 \to C_3 + 2C_1} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 7 \end{pmatrix}$$

$$\xrightarrow{R_3 \to R_3 - \frac{1}{2}R_2} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 6.5 \end{pmatrix} \xrightarrow{C_3 \to C_3 - \frac{1}{2}C_2} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6.5 \end{pmatrix}$$

בכך הגענו כבר לצורה אלכסונית, ניתן גם לנרמל את איברי האלכסון באופן הבא

$$\xrightarrow{R_2 \to \frac{1}{\sqrt{2}}R_2} \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{2}{\sqrt{2}} & 0 \\ 0 & 0 & 6.5 \end{pmatrix} \xrightarrow{C_2 \to \frac{1}{\sqrt{2}}C_2} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 6.5 \end{pmatrix}$$

$$\xrightarrow{R_3 \to \frac{1}{\sqrt{6.5}}R_3} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{6.5}{\sqrt{6.5}} \end{pmatrix} \xrightarrow{C_3 \to \frac{1}{\sqrt{6.5}}C_3} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

נשים לב שמתוך החפיפה ניתן למצוא גם את המטריצה "המחפפת", על ידי כך שנבצע על היחידה $\mathbb{I}$ את פעולות <u>העמודה</u> בלבד:

$$\mathbb{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{C_2 \to C_2 + C_1} \ldots = P$$

ואז ניתן לוודא שאכן $P^t AP = D$ כאשר D היא המטריצה האלכסונית שקיבלנו.

התהליך הזה מסורבל וארוך אך ניתן להשתמש בו. מצד שני, כזכור מאלגברה א' אנו יודעים שפעולה אלמנטרית על שורה היא כמו כפל <u>משמאל</u> (במטריצה האלמנטרית המתאימה) ופעולה אלמנטרית על עמודה היא כמו כפל <u>מימין</u>. מאחר וביצענו את אותן פעולות על השורות והעמודות, הרי שבעצם אנו מבצעים תהליך מהצורה:

$$A \xrightarrow{E_1} E_1 A \xrightarrow{E_1^t} E_1 A E_1^t \xrightarrow{E_2} E_2 E_1 A E_1^t \xrightarrow{E_2^t} E_2 E_1 A E_1^t E_2^t \ldots$$

לפיכך, אם D אלכסונית אז מתקיים

$$\underbrace{E_k...E_2E_1}_{P^t}A\underbrace{E_1^tE_2^t...E_k^t}_{P}=D$$

כאשר P הפיכה.

<u>שיטה 3 - רק עבור $\mathbb{F}=\mathbb{R}$ - לכסון אורתוגונלי</u>

ראינו שאם A סימטרית ממשית אז A לכסינה אורתוגונלית, כלומר דומה למטריצה אלכסונית על ידי מטריצה P שהיא אורתוגונלית. כלומר

$$D=P^{-1}AP$$

כאשר על אלכסונה של D נמצאים הערכים העצמיים של A. נשים לב שזה לא מביא אותנו לצורה בה איברי האלכסון הם $0,\pm1$, אלא לצורה בה איברי האלכסון הם הערכים העצמיים של A. A דומה למטריצה D ולכן בפרט גם חופפת לה. A גם חופפת למטריצה שאיברי האלכסון שלה הם $0,\pm1$, אבל <u>לא דומה לה</u>.

<u>הרצאה - 19 ליוני</u>

תזכורת: בשלב זה אנו מגבילים את הדיון שלנו למטריצות סימטריות/תבניות ביליניאריות סימטריות/תבניות ריבועיות. שאלנו מתי שתי מטריצות סימטריות הן חופפות, ואמרנו שהשלב הראשון אמור להיות להביא אותה לצורה נוחה, וראינו שלוש דרכים להביא מטריצה סימטרית לצורה אלכסונית.

<u>משפט:</u>

א. תהי $A\in M_n(\mathbb{R})$ סימטרית (שימו לב, אנחנו בממשיים!). אזי A חופפת למטריצה אלכסונית כאשר על האלכסון שלה מופיעים 0, -1, או 1ים.

ב. תהי $A\in M_n(\mathbb{C})$ סימטרית (שימו לב, היא לא חייבת להיות צמודה לעצמה, או נורמלית, או לכסינה אוניטרית). אז A חופפת למטריצה אלכסונית כאשר על האלכסון מופיעים 0 או 1ים.

<u>הוכחה:</u>

ראשית הן מעל $\mathbb{R}$ והן מעל $\mathbb{C}$, ראינו ש-A חופפת למטריצה אלכסונית (באחת משלוש השיטות שראינו).

כעת, ניתן ללכסן על ידי פעולות חפיפה (על שורה ועמודה לסירוגין) ולקבל:

$$A\to...\to D=\begin{pmatrix}\alpha_1&&&\\&\alpha_2&&\\&&\ddots&\\&&&\alpha_n\end{pmatrix}$$

<u>מעל $\mathbb{R}$</u>: לכל $\alpha_i\neq0$ - אם $\alpha_i>0$ אז נבצע את פעולת החפיפה:

$$D\xrightarrow{R_i\to\frac{1}{\sqrt{\alpha_i}}R_i}\begin{pmatrix}\ddots&&\\&\frac{\alpha_i}{\sqrt{\alpha_i}}&\\&&\ddots\end{pmatrix}\xrightarrow{C_i\to\frac{1}{\sqrt{\alpha_i}}C_i}\begin{pmatrix}\ddots&&\\&\frac{\alpha_i}{\sqrt{\alpha_i}}\frac{1}{\sqrt{\alpha_i}}&\\&&\ddots\end{pmatrix}=\begin{pmatrix}\ddots&&\\&1&\\&&\ddots\end{pmatrix}$$

כלומר, על ידי פעולת חפיפה החלפנו את α_i החיובי ב-1. בצורה דומה ניתן להחליף כל α_i שלילי ב-(-1), וזאת על ידי כפל השורה והעמודה ב-$\frac{1}{\sqrt{-\alpha_i}}$ (לחילופין יכולנו לטפל בשני המקרים בו זמנית על ידי כפל ב-$\frac{1}{\sqrt{|\alpha_i|}}$). בכך הוכחנו את סעיף א'.

<u>מעל $\mathbb{C}$</u>: לכל $\alpha_k\in\mathbb{C}$ מרוכב קיים פתרון z למשוואה $z^2=\alpha_k$. נסמן פתרון כזה על ידי $z=\sqrt{\alpha_k}$, ואז באופן דומה נחליף כל $\alpha_k\neq0$ ב-1. בכך הוכחנו את סעיף ב'. ■

<u>שימו לב</u>: המשפט אומר שהמטריצות חופפות למטריצות עם המספרים הנ"ל על אלכסונן, אבל זה לא אומר שהן <u>דומות</u> להן. המטריצות דומות למטריצה האלכסונית שעל אלכסונה מופיעים הערכים העצמיים, ופעולות ה-Scaling שעשינו כדי שיופיעו ±1 על האלכסון לא שומרות על הדמיון.

דיון: אנו רוצים להבין מתי 2 מטריצות סימטריות חופפות זו לזו, או מתי 2 תבניות ריבועיות שקלות. תזכורת: A ו־B חופפות אם יש P הפיכה כך ש־$B = P^t AP$. כשעסקנו בדמיון מטריצות, ראינו שיש תכונות שזהות לשתי מטריצות דומות: למשל העקבה, הדרגה, הפולינום האופייני, וכו'. אבל ראינו שלא ניתן להסיק את הכיוון ההפוך: שאם יש להן (למשל) אותו פולינום אופייני, לא ניתן להסיק מכך שהן דומות. ראינו שהשמורה השקולה לדמיון היא שיש לשתי המטריצות אותה צורת ג'ורדן. באופן דומה נרצה להבין מהן השמורות (אינווריאנטות) של מטריצות חופפות.

הגדרה: בהינתן $M \in M_n(\mathbb{R})$ אלכסונית, נסמן

$$\begin{aligned} n_0(M) &= \text{number of zeros on the diagonal} = n - \operatorname{rank}(M) \\ n_+(M) &= \text{number of positive eigenvalues} \\ n_-(M) &= \text{number of negative eigenvalues} \end{aligned}$$

משפט (משפט ההתמדה של סילבסטר): תהיינה $A, B \in M_n(\mathbb{R})$ אלכסוניות, חופפות, אז

$$\begin{aligned} n_0(A) &= n_0(B) \\ n_+(A) &= n_+(B) \\ n_-(A) &= n_-(B) \end{aligned}$$

הוכחה: ראשית, ברור ש־$n_0(A) = n_0(B)$, כי מהנתון ש־A, B חופפות נובע שה־rank שלהן שווה (כי פעולות אלמנטריות על שורה או עמודה לא משפיעות על הדרגה, ואם הן חופפות יש שרשרת של פעולות חפיפה מהאחת לשניה). נזכיר שהדרגה היא בעצם מספר האפסים על האלכסון.

שנית, מספיק להראות ש־$n_+(A) = n_+(B)$, ואז נובע גם ש־$n_-(A) = n_-(B)$ (כי $n = n_0(A) + n_+(A) + n_-(A) = n_0(B) + n_+(B) + n_-(B)$). מהנתון, A, B חופפות, ומכך נובע שיש q ריבועית ובסיסים F, L כך ש־

$$[q]_L = B, \quad [q]_F = A$$

ללא הגבלת הכלליות, נוכל לסדר את איברי הבסיסים

$$\begin{aligned} L &= \{u_1, u_2, ..., u_k, u_{k+1}, ..., u_n\} \\ F &= \left\{v_1, v_2, ..., v_{\tilde{k}}, v_{\tilde{k}+1}, ..., v_n\right\} \end{aligned}$$

כך ש־$q(u_1), ..., q(u_k) > 0$ ו־$q(u_{k+1}), ..., q(u_n) \leq 0$, וכמו כן $q(v_1), ..., q\left(v_{\tilde{k}}\right) > 0$ ו־$q\left(v_{\tilde{k}+1}\right), ..., q(v_n) \leq 0$. כלומר בעצם סידרנו את האיברים החיוביים כך שיופיעו ראשונים על האלכסון, ועלינו בעצם להראות שמספר האיברים החיוביים יהיה זהה בשני המקרים, כלומר עלינו להראות ש־$k = \tilde{k}$. למעשה אנו יכולים בצורה כזו להביא את המטריצות לצורה הבאה:

$$A = \begin{pmatrix} \mathbb{I}_{k\times k} & & \\ & -\mathbb{I}_{m\times m} & \\ & & 0 \end{pmatrix}, \quad B = \begin{pmatrix} \mathbb{I}_{\tilde{k}\times \tilde{k}} & & \\ & -\mathbb{I}_{\tilde{m}\times \tilde{m}} & \\ & & 0 \end{pmatrix}$$

המשך ההוכחה: נגדיר

$$\begin{aligned} U &= \operatorname{span}\{u_1, ..., u_k\} \\ W &= \operatorname{span}\left\{v_{\tilde{k}+1}, ..., v_n\right\} \end{aligned}$$

מאחר והוקטורים $u_1, ..., u_k$ הם חלק מבסיס, אז $\dim U = k$, ובצורה דומה $\dim W = n - \tilde{k}$.

אנו יודעים שהפעלת q על כל אחד מהאיברים $u_1, ...u_k$ תיתן תוצאה חיובית. מאחר ו־q תבנית ריבועית, אנו יודעים שכפל של וקטור בסקלר α ייצא החוצה בריבוע, כלומר אם $q(x) = f(x, x)$ אז $q(\alpha x) = f(\alpha x, \alpha x) = \alpha^2 f(x, x) = \alpha^2 q(x)$, לכן הפעלת q על כל כפולה סקלרית של אחד מהאיברים $u_1, ..., u_k$ תיתן תוצאה חיובית. בצורה דומה, גם הפעלת q על כל קומבינציה ליניארית של $u_1, ..., u_k$ תיתן תוצאה חיובית (מהביליניאריות).

לכן נקבל שלכל $0 \neq u \in U$ מתקיים $q(u) > 0$, ובצורה דומה לכל $w \in W$ מתקיים $q(w) \leq 0$, ולכן $U \cap W = \left\{\vec{0}\right\}$. ממשפט המימדים מאלגברה א' נקבל

$$n \underbrace{\geq}_{U+W\subset\mathbb{R}^n} \dim(U+W) = \underbrace{\dim U}_{=k} + \underbrace{\dim W}_{n-\tilde{k}} - \underbrace{\dim(U\cap W)}_{=0}$$

לכן $k \geq \tilde{k}$. משיקולי היפוך תפקידים בין W ל-U (כלומר על ידי הגדרת U בתור $\mathrm{span}\{u_{k+1},...,u_n\}$ והגדרת W בתור $\mathrm{span}\left\{v_1,...,v_{\tilde{k}}\right\}$) נקבל באופן דומה $k \geq \tilde{k}$ וסיימנו. ■

<u>מסקנה:</u>

- בהינתן תבנית ריבועית $q(x) = x^t Ax$ כאשר $A \in M_n(\mathbb{R})$ סימטרית, נוכל להגדיר את <u>הסיגנטורה (חותם)</u> בתור הזוג הבא:

$$\mathrm{sgn}(A) = \mathrm{sgn}(q) = (n_+, n_-)$$

כאשר n_+, n_- הם למעשה $n_+(D), n_-(D)$ כאשר D היא איזו צורה אלכסונית של q. לפי המשפט אנו מבינים שהסיגנטורה מוגדרת היטב. נעיר שיש ספרים שמגדירים את הסיגנטורה בתור ההפרש $n_+ - n_-$, ואז נדרשת מעט עבודה כדי להסיק מהו הזוג הסדור (n_+, n_-), אבל ניתן לעשות זאת מאחר ואנו יודעים את n ואת $\mathrm{rank}(A)$.

- אם $A, B \in M_n(\mathbb{R})$ סימטריות חופפות, אז $\mathrm{sgn}(A) = \mathrm{sgn}(B)$.

<u>מיון תבניות ריבועיות/מטריצות סימטריות</u>

1) מעל $\mathbb{C}$: 2 תבניות ריבועיות q_1, q_2 שקולות $\iff$ $\mathrm{rank}(q_1) = \mathrm{rank}(q_2)$ (כאשר הדרגה היא בעצם של מטריצות מייצגות של q_1 ו-q_2).

לחילופין: 2 מטריצות סימטריות $A, B \in M_n(\mathbb{C})$ חופפות $\iff$ $r(A) = r(B)$. כלומר: <u>מעל $\mathbb{C}$</u> ה-rank הוא אינווריאנט מלא לחפיפת מטריצות.

<u>הוכחה:</u>

כיוון ראשון $\Leftarrow$: ראינו שפעולת חפיפה משמרות את הדרגה.

כיוון שני $\Rightarrow$: ראינו שכל אחת מ-A, B חופפת למטריצה אלכסונית D שיש בה $r(A) = r(B)$ 1ים והשאר 0ים. ולכן מסימטריות וטרנזיטיביות החפיפה נקבל ש-A חופפת ל-B.

2) <u>מעל $\mathbb{R}$</u>: 2 תבניות ריבועיות q_1, q_2 שקולות $\iff$ $\mathrm{rank}(q_1) = \mathrm{rank}(q_2)$ וגם $\mathrm{sgn}(q_1) = \mathrm{sgn}(q_2)$.

לחילופין: 2 מטריצות סימטריות $A, B \in M_n(\mathbb{R})$ חופפות $\iff$ $r(A) = r(B)$ וגם $\mathrm{sgn}(A) = \mathrm{sgn}(B)$.

<u>הוכחה:</u>

כיוון ראשון $\Leftarrow$: ראינו שפעולות חפיפה משמרות את הדרגה, וממשפט ההתמדה של סילבסטר גם הסיגנטורה זהה.

כיוון שני $\Rightarrow$: מהנתון נובע ש-

$$\begin{aligned} n_0(q_1) &= n_0(q_2) \\ n_+(q_1) &= n_+(q_2) \\ n_-(q_1) &= n_-(q_2) \end{aligned}$$

כלומר בכל אחד מהם, בצורה אלכסונית מתאימה, יש אותו מספר של 0ים, 1ים, ו-(-1)ים. מכאן שכל אחת מ-q_1 ו-q_2 שקולות ל-

$$x_1^2 + ... + x_k^2 - x_{k+1}^2 - ... - x_m^2 + 0 \cdot x_{m+1}^2 + ... + 0 \cdot x_n^2$$

ושוב מסימטריות וטרנזיטיביות של שקילות תבניות נסיק ש-$q_1 \sim q_2$.

3) (הערה) <u>מעל $\mathbb{Q}$</u>: כעת אנו לא יכולים לעשות את הטריק של הבאת הצורה האלכסונית כך שהיא תכיל ± 1, כי לא ניתן לחלק בשורש של כל מספר שהוא (הוא לא בהכרח יהיה רציונלי). לכן נדרשת שמורה אחרת במקרה זה, לשם כך מסתכלים על הדטרמיננטה.
נשים לב שאם A, B סימטריות חופפות מעל $\mathbb{Q}$, אז ה־$\det$ שלהם נבדל בריבוע, כי אם $B = P^t AP$ אז

$$\det(B) = \det\left(P^t AP\right) = \underbrace{\det\left(P^t\right)}_{\det(P)} \det(A)\det(P) = \det(A)\left[\det(P)\right]^2$$

ואז נביט ב־

$$\mathbb{Q}^x/(\mathbb{Q}^x)^2 = \{[a] \,:\, a \in \mathbb{Q}\}$$

כלומר אלו מחלקות השקילות תחת יחס השקילות הבא עבור a, b רציונליים

$$a \sim b \iff \exists c \in \mathbb{Q} \,:\, a = c^2 b$$

מגדירים דיסקרמיננטה של תבנית ריבועית באופן הבא:

$$\text{disc}(q) = [\det(A)] \in \mathbb{Q}^x/(\mathbb{Q}^x)^2$$

כאשר A היא איזושהי מטריצה מייצגת של q. הדיסקרמיננטה נשמרת בין תבניות חופפות, אבל מסתבר שגם זו לא שמורה מלאה. למעשה מתקיים:

שתי תבניות ריבועיות הן שקולות מעל $\mathbb{Q}$ $\iff$ יש להן אותה rank, אותה סיגנטורה, אותה דיסקרמיננטה, וגם אותם Hasse-Witt Invariants, שכדי לתאר אותם צריך להבין את $\mathbb{Q}_p$, שהם המספרים הפיאדיים (כאשר p ראשוני). למעשה הרעיון בהם הוא להשלים את $\mathbb{Q}$ עם מטריקה אחרת, ממנה מתקבלים גדלים שונים ממה שאנו רגילים במרחב אוקלידי, גבולות שונים, וכו'.

חפיפה סימולטנית

<u>משפט</u>: אם $A, B \in M_n(\mathbb{R})$ כאשר B סימטרית, A חיובית לחלוטין, אז ל־A, B יש חפיפה סימולטנית, כלומר יש P הפיכה ויש D אלכסונית כך ש־

$$P^t AP = \mathbb{I} \quad P^t BP = D$$

<u>הוכחה</u>: A חיובית לחלוטין ולכן חופפת ל־$\mathbb{I}$ (שימו לב, היא לא דומה ל־$\mathbb{I}$, היא דומה למטריצה אלכסונית בעלת ערכים עצמיים חיוביים, אבל אם אז עושים את ה־Scaling כך שעל האלכסון יופיעו רק 1ים הדמיון לא נשמר). לכן יש Q הפיכה כך ש־$Q^t AQ = \mathbb{I}$.
נסמן $C = Q^t BQ$. נשים לב שגם C סימטרית כי

$$C^t = \left(Q^t BQ\right)^t = Q^t B^t \left(Q^t\right)^t = Q^t BQ = C$$

מכאן ש־C סימטרית ממשית ולכן לכסינה אורתוגונלית, כלומר יש R אורתוגונלית (כלומר $R^{-1} = R^t$) ו־D אלכסונית (עם הע"ע של C) כך ש־$D = R^{-1}CR = R^t CR$. נביט על $P = QR$, ברור ש־P הפיכה כמכפלה של שתי מטריצות הפיכות, ומתקיים

$$P^t AP = (QR)^t AQR = R^t \underbrace{Q^t AQ}_{\mathbb{I}} R = R^t \mathbb{I} R = R^{-1} R \mathbb{I} = \mathbb{I}$$
$$P^t BP = (QR)^t BQR = R^t \underbrace{Q^t BQ}_{C} R = R^t CR = D$$

■

<u>תרגיל לבית</u>: מצאו P, D כאלו עבור המטריצות

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 5 & 5 \\ 1 & 5 & 14 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 3 \\ 0 & 4 & 7 \\ 3 & 7 & 10 \end{pmatrix}$$

בכך סיימנו את החומר למבחן, הנושאים הנותרים שנלמד יהיו או העשרה, או חזרה למבחן. אחד מנושאי ההעשרה שנלמד הוא מכפלה טנזורית, שהיא תהיה גם קשורה לנושא של מהי תבנית בילינארית באופן הכי כללי. זהו נושא חשוב ושימושי מאוד. בנוסף נלמד על המספרים ה־pאדיים כנושא העשרה.