TEL AVIV UNIVERSITY אוניברסיטת תל-אביב
RAYMOND AND BEVERLY SACKLER FACULTY OF EXACT SCIENCES
SCHOOL OF MATHEMATICAL SCIENCES
הפקולטה למדעים מדויקים ע"ש ריימונד ובברלי סאקלר
בית הספר למדעי המתמטיקה

אלגברה לינארית 2 א

מערכי שעור

תשפ"א

נערך על ידי

דן הרן

עדכון אחרון: 26.1.2022

תוכן העניינים

מבוא

אלגברה לינארית 2 כוללת בדרך כלל את החלקים הבאים:

(א) פירוק של פולינומים מעל שדות (פרקים 1-3).

(ב) לכסון של מטריצות והצגה אלכסונית של העתקות לינאריות (פרקים 4, 6).

(ג) צורות קנוניות של מטריצות ושל העתקות (פרקים 7-9).

(ד) מרחבי מכפלה פנימית (פרק 10).

(ה) העתקות לינאריות במרחבי מכפלה פנימית ולכסון אוניטרי (פרק 11).

(ו) תבניות בילינאריות ומשפט סילבסטר (פרק 12).

(ז) שניוניות (פרק 13).

למרות השמות השונים, כל הנושאים האלה קשורים ברעיון מרכזי אחד אותו ננסה להסביר כעת.

יהי V מרחב וקטורי ממימד סופי ותהי $T\colon V \to V$ העתקה לינארית. אם $\mathcal{B}$ בסיס של V, אז ל־T מתאימה מטריצה $[T]_{\mathcal{B}}^{\mathcal{B}}$. תכונות רבות של T נוכל לחשב מתוך המטריצה הזאת. אך המטריצה תלויה בבחירת הבסיס. לכן אם נבחר בסיס טוב, נקבל מטריצה פשוטה יחסית שקל לעשות חישובים אתה, למשל, מטריצה אלכסונית.

חלק **ב**′ מתעסק בדיוק בשאלה מתי ואיך ניתן לעשות זאת ואיזו מטריצה אלכסונית תתקבל.

חלק **ג**′ עונה בעצם על השאלה, מה קורה באופן כללי, כאשר לא בהכרח אפשר לקבל מטריצה אלכסונית. נראה שיש משפחה מסוימת של מטריצות (צורות קנוניות) שאפשר, על ידי בחירת בסיס מתאים, להגיע לאחת ורק אחת (במובן מסוים) מהן. ברצוני להודות לפרופ′ אשר בן־ארצי על תרומתו לפישוט חלק זה.

חלק **ד**′ אינו בדיוק שייך לכאן מבחינה רעיונית (ולפעמים הוא נלמד בחלקו באלגברה לינארית 1); הוא הכנה לחלק הבא. בחלק זה נרחיב את המושג של מרחב וקטורי באשר נוספים אליו מושגים כמו אורך הוקטור וניצבות של וקטורים, כפי שהם מוכרים לנו מן הגיאומטריה.

כעת נוכל לשאול לאיזה מטריצה $[T]_{\mathcal{B}}^{\mathcal{B}}$ נוכל להגיע אם נדרוש שהבסיס $\mathcal{B}$ מורכב מוקטורים בעלי אורך 1 ניצבים זה לזה. בזה עוסק חלק **ה**′.

חלקים **ו**′,**ז**′ אינם עוסקים בהעתקות לינאריות, אלא בעצמים אלגבריים או גיאומטריים שגם אותם אפשר לייצג על ידי מטריצות, והמטריצה תלויה בבחירת בסיס. שוב נשאל לאיזו מטריצה נוכל להגיע ועד כמה מטריצה זו יחידה. במיוחד בחלק **ז**′ נענה על השאלה, מהי הצורה הגיאומטרית המיוצגת במישור או במרחב התלת ממדי על ידי משואה ממעלה 2.

אך חלקים **ב**′,**ג**′ תלויים בידע טכני על פירוק פולינומים מעל שדות, ולכן נפתח את הלימוד בחלק **א**′, שלכאורה אינו שייך לאלגברה לינארית. (למען הסר ספק, גם חומר זה הינו למבחן...)

ספרים מומלצים

כאמור, הקורס יתנהל לפי חוברת זו. מלבדה אין ספר אשר מתאים במדויק לקורס, אך יש ספרים רבים שמתאימים לחלקים שונים של הקורס:

- ש. עמיצור, *אלגברה א'*, אקדמון. מכיל את רוב החומר, אך בסימונים מיוחדים לספר זה (בהרכבת העתקות הראשונה שפועלת היא השמאלית, וכו').
- *אלגברה לינארית* של האוניברסיטה הפתוחה.

חוברות אלו הנן ברמה קצת נמוכה יותר ממה שנדרש לנו, אך הנן קריאות מאד ומתאימות ללימוד עצמי.

- Hoffman, Kunze, *Linear Algebra*, Prentice Hall, 2nd edition, 1971

ספר קלאסי בנושא, רוב הספרים האחרים מבוססים במידה זו או אחרת עליו. לא קל לקריאה.

- פרנק איירס, *המטריצות*, סדרת שאום, Frank Ayres, *Matrices*, Schaum Series,
- סיימור ליפשיץ, *אלגברה לינארית*, סדרת שאום. Seymour Lipschuts, *Linear Algebra*, Schaum Series.

שני הספרים האלה מכילים הרבה תרגילים (פתורים). החומר התיאורטי מאורגן בנספחים ובתרגילים. יצאו בהוצאות אחדות. לצערי, בהוצאות מסויימות יש פתרונות שגויים לתרגילים.

1. חוגים

בפרק זה אנו מתחילים ללמוד על פולינומים. אוסף כל הפולינומים מעל שדה מסוים הוא מבנה אלגברי שנקרא **חוג**. טענות מסוימות על פולינומים אפשר להכליל לחוגים או לחוגים עם תכונות מסוימות. לכן נלמד קצת על חוגים באופן כללי.

הגדרה 1.1: **חוג** (ring) הנו קבוצה R, עליה מוגדרות שתי פעולות: חיבור $(+)$ וכפל $(\cdot$ או ללא סימן$)$, אשר מקיימים את החוקים הבאים:

(ח0) קשירות החיבור: לכל $a,b\in R$ קיים $c\in R$ יחיד כך ש־$a+b=c$.

(ח1) כלל הצירוף לחיבור: $(a+b)+c=a+(b+c)$ לכל $a,b,c\in R$.

(ח2) כלל החילוף של החיבור: $a+b=b+a$, לכל $a,b\in R$.

(ח3) קיום איבר אפס: קים איבר ב־R המסומן ב־0 והמקיים $a+0=a$ לכל $a\in R$.

(ח4) קיום איבר נגדי: לכל $a\in R$ קים איבר $-a\in R$ כך ש־$a+(-a)=0$.

(כ0) קשירות הכפל: לכל $a,b\in R$ קיים $c\in R$ יחיד כך ש־$ab=c$.

(כ1) כלל הצירוף לכפל: $(ab)c=a(bc)$, לכל $a,b,c\in R$.

(כ2) כללי פילוג: $a(b+c)=ab+ac$ ו $(a+b)c=ac+bc$, לכל $a,b,c\in R$.

הגדרה 1.2:

(א) חוג R ייקרא **חוג עם יחידה** אם יש בו איבר 1 המקיים $a1=a=a1$, לכל $a\in R$.

(ב) חוג R ייקרא **חילופי** אם $ab=ba$ לכל $a,b\in R$.

(ג) חוג חילופי עם יחידה בו $1\neq 0$ ייקרא **תחום שלמות** אם לכל $a,b\in R$ שונים מ־0 מתקיים $ab\neq 0$. ∎

תרגיל: *יהי R חוג.*

(א) *איבר 0 שמקיים את (ח4), יחד עם התנאים הקודמים לו, הוא יחיד. ייקראה* **האפס** *של R.*

(ב) *לכל $a\in R$ איבר $-a$ שמקיים את (ח5), יחד עם התנאים הקודמים לו, הוא יחיד. ייקראה* **הנגדי** *של a.*

(ג) *איבר $1\in R$ שמקיים את תנאי 1.2(א) –אם קיים –הוא יחיד. ייקראה* **היחידה** *של R.*

(ד) *יהי $a\in R$. אז $a0=0=0a$.*

הוכחה: (ד) $a0=a(0+0)=a0+a0$. נחבר את הנגדי $-(a0)$ של $a0$ לשני האגפים ונקבל $0=a0$.
באופן דומה $0=0a$.

דוגמאות 1.3: (א) $\mathbb{Z}$ הוא תחום שלמות. $(\mathbb{Z}=\{0,\pm1,\pm2,\pm3,\ldots\})$.

(ב) $2\mathbb{Z}$ (=קבוצת המספרים הזוגיים) הוא חוג חילופי ללא יחידה.

(ג) כל שדה הוא תחום שלמות. בפרט $\mathbb{Q}$ (שדה המספרים הרציונליים), $\mathbb{R}$ (שדה המספרים הממשיים), $\mathbb{C}$ (שדה המספרים המרוכבים) ו־$\mathbb{F}_p$ (השדה בעל p איברים, באשר p ראשוני; $\mathbb{F}_p=\mathbb{Z}/p\mathbb{Z}$) הם תחומי שלמות.

(ד) עבור כל מספר טבעי n, $\mathbb{Z}/n\mathbb{Z}$ הוא חוג חילופי עם יחידה. זהו שדה אם ורק אם n הנו מספר ראשוני.

(ה) יהי R חוג (לאו דוקא חילופי). **חוג הפולינומים** מעל R מסומן ב $R[X]$ ומוגדר כקבוצת כל הביטויים הפורמליים

$$a_0X^0 + a_1X^1 + a_2X^2 + \dots \qquad (\text{או בכתיב מקוצר: } \sum_{i=0}^{\infty} a_iX^i)$$

שבהם שייכים ה־a_i־ים ל־R וכמעט כולם (=פרט למספר סופי) שווים לאפס. חיבור וכפל מוגדרים על ידי הנוסחאות:

$$(\sum_{i=0}^{\infty} a_iX^i) + (\sum_{i=0}^{\infty} b_iX^i) = \sum_{i=0}^{\infty}(a_i + b_i)X^i,$$

$$(\sum_{i=0}^{\infty} a_iX^i)(\sum_{j=0}^{\infty} b_jX^j) = \sum_{k=0}^{\infty}\Big(\sum_{i+j=k} a_ib_j\Big)X^k = \sum_{k=0}^{\infty}\Big(\sum_{i=0}^{k} a_ib_{k-i}\Big)X^k.$$

איבר האפס של $R[X]$ הוא $0 = \sum_{i=0}^{\infty} 0X^i$.

רישום מקוצר של פולינומים. נהוג להזניח $+0X^i$ מהסכומים האינסופיים ולרשום את הפולינומים כסכומים סופיים: $\sum_{i=0}^{n} a_iX^i$. כמו כן נהוג לכתוב X במקום X^1, ו־a_0 במקום a_0X^0. אם יש ב־R יחידה, נהוג לכתב X^i במקום $1X^i$. הדבר עלול לכאורה לגרום לבילבול, כי, למשל, לביטוי $X + X^2$ שתי משמעויות: האחת: סכום ב־$R[X]$ של הפולינומים X ו־X^2, והשניה: הפולינום $0X^0 + 1X^1 + 1X^2 + 0X^3 + \cdots$. אך שני הביטוים שווים. באפן דומה aX אפשר להבין כפולינום $0X^0 + aX^1 + 0X^2 + \cdots$ או כמכפלה של הפולינומים a ו־X; שוב, שני הביטוים שווים.

האיברים a_i נקראים **מקדמי הפולינום**. אם $a_n \neq 0$, המקדם a_n של $\sum_{i=0}^{n} a_iX^i$ נקרא **המקדם העליון**. אם $a_n = 1$ אומרים ש**הפולינום מתוקן**. שים לב שהפולינום $X = 0X^0 + 1X^1 + 0X^2 + 0X^3 + \cdots$ מתחלף בכפל עם כל איברי $R[X]$. את איברי $R[X]$ נהוג לסמן על ידי אותיות לטיניות קטנות, למשל f, או גם על ידי $f(X)$, אם רוצים להדגיש את התלות של f ב־X.

המעלה $\deg(f)$ של פולינום $f = \sum_{i=1}^{n} a_iX^i$ מוגדרת כך: $\deg(f) = \max(i|\ a_i \neq 0)$, אם $f \neq 0$, ואילו $\deg 0 = -\infty$.

(ו) אם R חוג (לא בהכרח חילופי!), יהי $\mathrm{M}_n(R)$ **חוג המטריצות מסדר** $n \times n$ **מעל** R, עם החיבור: $(A+B)_{ij} = (A)_{ij} + (B)_{ij}$ והכפל: $(AB)_{ij} = \sum_{k=1}^{n}(A)_{ik}(B)_{kj}$, לכל $1 \leq i,j \leq n$. (כאן $(A)_{ij}$ מסמן את הרכיב ה־(i,j) של A ב־R.) בדרך כלל $\mathrm{M}_n(R)$ איננו חילופי, אפילו אם R חילופי.

אם R חוג עם יחידה, גם $\mathrm{M}_n(R)$ חוג עם יחידה; היחידה בו היא

$$I = I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

בפרט אפשר להתבונן בחוג המטריצות $\mathrm{M}_n(R[X])$ מעל חוג הפולינומים $R[X]$. הרכיבים של כל מטריצה כזו הם פולינומים עם מקדמים ב־R.

באפן דומה אפשר להתבונן בחוג $\mathrm{M}_n(R)[X]$. איברי חוג זה הנם פולינומים שמקדמיהם מטריצות.

(ז) יהי V מרחב וקטורי מעל שדה F. נסמן ב $L_F(V)$ את קבוצת כל ההעתקות הלינאריות $T\colon V \to V$ (הנקראות גם **אנדמורפיזמים** של V). קבוצה זו מהווה חוג ביחס לחיבור, $(T+S)(v) = Tv + Sv$, ולכפל המוגדר כהרכבה של העתקות, $(TS)(v) = T(S(v))$. לחוג זה יש יחידה (העתקת הזהות של V) שנסמנה ב־1 או ב־1_V, אם רוצים לצין את V. אם $\dim(V) \geq 2$ אז $L_F(V)$ אינו חילופי.

גם כאן אנו יכולים לבנות את החוג $L_F(V)[X]$ של פולינומים ב־X שמקדמיהם העתקות לינאריות מ־V ל־V. ∎

תרגיל 1.4: *יהי R חוג ויהיו $f = \sum_{i=0}^{\infty} a_i X^i, g = \sum_{i=0}^{\infty} b_i X^i \in R[X]$ שני פולינומים מעליו. יהיו $m = \deg(f)$, $n = \deg(g)$. הוכח:*

(א) $\deg(f+g) \leq \max(m,n)$.

(ב) $\deg(fg) \leq m+n$.

(ג) *נניח $m,n \geq 0$. אז המקדם של X^{m+n} ב־fg הוא $a_m b_n$.*

(ד) *אם R הוא תחום שלמות, מתקיים שוויון בתנאי (ב).*

(ה) *אם R תחום שלמות, גם $R[X]$ הוא תחום שלמות.*

(ו) *אם R חוג עם יחידה וקיים $c \in R$ כך ש־$b_n c = 1$ אז מתקיים שוויון בתנאי (ב).*

הוכחה: המקדם של X^k ב־fg הוא $\sum_{i+j=k} a_i b_j$.

(ב) אם $k > m+n$ ו־$i+j=k$ אז $j > n$ או $i > m$ (אחרת $j \leq n$ וגם $i \leq m$, ולכן $i+j \leq m+n < k$, סתירה). לכן $b_j = 0$ או $a_i = 0$ ומכאן $a_i b_j = 0$. לכן המקדם של X^k ב־fg הוא 0.

(ג) אם $k = m+n$ ו־$i+j=k$ אז $j > n$ או $i > m$ או $i = m, j = n$ (אחרת $j \leq n$ וגם $i \leq m$, ולפחות אחד משני האי שוויונים הוא חריף; לכן $i+j < m+n = k$, סתירה). בשני המקרים הראשונים $a_i b_j = 0$. לכן המקדם של X^k ב־fg הוא $a_m b_n$.

(ו) $(a_m b_n)c = a_m(b_n c) = a_m 1 = a_m \neq 0$, לכן $a_m b_n \neq 0$. לפי (ב), (ג), $\deg(fg) = m+n$.

∎

הגדרה 1.5: הומומורפיזם. יהיו R ו־S שני חוגים. העתקה $\varphi\colon R \to S$ תקרא **הומומורפיזם** אם היא שומרת על החיבור ועל הכפל. כלומר $\varphi(a+b) = \varphi(a) + \varphi(b)$ ו־$\varphi(ab) = \varphi(a)\varphi(b)$. בפרט מקבלים:

(א) $\varphi(0) = 0$. אכן, $0 + \varphi(0) = \varphi(0) = \varphi(0+0) = \varphi(0) + \varphi(0)$ ואם נחבר לשני האגפים את הנגדי של $\varphi(0)$, נקבל $0 = \varphi(0)$.

(ב) $\varphi(-a) = -\varphi(a)$, לכל $a \in R$. אכן, $\varphi(-a) + \varphi(a) = \varphi(-a+a) = \varphi(0) = 0$, ומכאן המסקנה.

אם R ו־S הם חוגים עם יחידה, נניח תמיד ש־$\varphi(1) = 1$. (מדוע צריך להניח זאת? זה לא מתקיים תמיד?)

הומומורפיזם חד חד ערכי ועל ייקרא **איזומורפיזם**.

דוגמה 1.6: (א) יהי V מרחב וקטורי ממימד סופי n מעל שדה F ויהי $\mathcal{B}$ בסיסו. ההעתקה $L_F(V) \to \mathrm{M}_n(F)$, אשר מתאימה להעתקה לינארית $T\colon V \to V$ את המטריצה שלה $[T]_{\mathcal{B}}^{\mathcal{B}}$ ביחס לבסיס $\mathcal{B}$, היא איזומורפיזם של חוגים. (זה נלמד באלגברה לינארית 1, גם אם לא הוגדרו שם חוגים.)

(ב) יהי R חוג עם יחידה ויהי $\alpha \in R$. אם $f = \sum_{i=0}^{\infty} a_i X^i \in R[X]$, נגדיר $f(\alpha) = \sum_{i=0}^{\infty} a_i \alpha^i$ (באשר $\alpha^0 = 1$). בכך הגדרנו **העתקת ההצבה** $\varphi\colon R[X] \to R$ על ידי $\varphi(f) = f(\alpha)$. האם φ הומומורפיזם?

טענה 1.7: *יהי R חוג עם יחידה, יהי $\alpha \in R$ ותהי $\varphi\colon R[X] \to R$ ההצבה $f \mapsto f(\alpha)$. יהיו $f, g \in R[X]$.*

(1) $\varphi(f+g) = \varphi(f) + \varphi(g)$.

(2) *אם α מתחלף בכפל עם כל מקדמי g אז* $\varphi(fg) = \varphi(f)\varphi(g)$.

(3) *אם α מתחלף בכפל עם כל אברי R אז φ הומומורפיזם.*

הוכחה: (1) נניח $f = \sum_{i=0}^{\infty} a_i X^i,\ g = \sum_{i=0}^{\infty} b_i X^i \in R[X]$.

$$\begin{aligned}\varphi(f+g) &= \varphi\Big(\sum_{i=0}^{\infty}(a_i+b_i)X^i\Big) = \sum_{i=0}^{\infty}(a_i+b_i)\alpha^i = \sum_{i=0}^{\infty} a_i\alpha^i + \sum_{i=0}^{\infty} b_i\alpha^i \\ &= f(\alpha) + g(\alpha) = \varphi(f) + \varphi(g)\end{aligned}$$

(בעיקר) בגלל חוק הפילוג ב־R.

(2) נניח $f = \sum_{i=0}^{\infty} a_i X^i,\ g = \sum_{i=0}^{\infty} b_i X^i \in R[X]$.

$$\begin{aligned}\varphi(fg) &= \varphi\Big(\Big(\sum_{i=0}^{\infty} a_i X^i\Big)\Big(\sum_{j=0}^{\infty} b_j X^j\Big)\Big) = \varphi\Big(\sum_{k=0}^{\infty}\Big(\sum_{i+j=k} a_i b_j\Big)X^k\Big) = \sum_{k=0}^{\infty}\Big(\sum_{i+j=k} a_i b_j\Big)\alpha^k \\ &= \sum_{k=0}^{\infty}\Big(\sum_{i+j=k} a_i b_j \alpha^k\Big)\end{aligned}$$

ואילו

$$\varphi(f)\varphi(g) = \Big(\sum_{i=0}^{\infty} a_i\alpha^i\Big)\Big(\sum_{j=0}^{\infty} b_j\alpha^j\Big) = \sum_{k=0}^{\infty}\Big(\sum_{i+j=k} a_i\alpha^i b_j\alpha^j\Big) = \sum_{k=0}^{\infty}\Big(\sum_{i+j=k} a_i b_j \alpha^k\Big).$$

(3) נובע מ־(2). נשים לב ש־$\varphi(1) = 1$. ∎

דוגמה 1.8: יהי R חוג לא חילופי עם יחידה. יש $\alpha, a \in R$ כך ש־$a\alpha \neq \alpha a$. אז

$$\varphi(X \cdot a) = \varphi(aX) = a\alpha \neq \alpha a = \varphi(X)\varphi(a).$$

כלומר, במקרה זה φ אינה שומרת כפל ובפרט איננה הומומורפיזם. ∎

הערה:

$$\begin{aligned} X\cdot a &= (0X^0+1X^1+0X^2+\cdots)(aX^0+0X^1+0X^2+\cdots) = \\ &= (0\cdot a)X^0+(0\cdot 0+1\cdot a)X^1+(0\cdot 0+1\cdot 0+0\cdot a)X^2+\cdots = \\ &= 0X^0+aX^1+0X^2+\cdots = aX \end{aligned}$$

נציג עוד דוגמה להומומורפיזם (קצת מסובכת, אך נחוצה בהמשך).

אבל לפני כן נתבונן בדוגמה של שני פולינומים מעל $\mathrm{M}_2(\mathbb{Z})$, למשל,

$$f=\begin{pmatrix}2&1\\0&1\end{pmatrix}+\begin{pmatrix}1&3\\1&1\end{pmatrix}X+\begin{pmatrix}0&1\\2&1\end{pmatrix}X^2,\quad g=\begin{pmatrix}2&1\\3&1\end{pmatrix}+\begin{pmatrix}0&1\\1&2\end{pmatrix}X+\begin{pmatrix}1&0\\0&1\end{pmatrix}X^2.$$

מבקשים להכפיל אותם.

מישהו מציע את הדרך הבאה:

$$f=\begin{pmatrix}2+X&1+3X+X^2\\X+2X^2&1+X+X^2\end{pmatrix},\quad g=\begin{pmatrix}2+X^2&1+X\\3+X&1+2X+X^2\end{pmatrix}.$$

ועכשיו נכפיל את שתי המטריצות (כמטריצות!):

$$\begin{aligned} fg &= \begin{pmatrix}7+12X+8X^2+2X^3&3+8X+9X^2+5X^3+X^4\\3+6X+8X^2+2X^3+2X^4&1+4X+7X^2+5X^3+X^4\end{pmatrix}. = \\ &= \begin{pmatrix}7&3\\3&1\end{pmatrix}+\begin{pmatrix}12&8\\6&4\end{pmatrix}X+\begin{pmatrix}8&9\\8&7\end{pmatrix}X^2+\begin{pmatrix}2&5\\2&5\end{pmatrix}X^3+\begin{pmatrix}0&1\\2&1\end{pmatrix}X^4 \end{aligned}$$

האם שיטה זו נכונה??

טענה 1.9: *יהי R חוג. נגדיר $\varphi\colon \mathrm{M}_n(R)[X]\to \mathrm{M}_n(R[X])$ על ידי*

$$\big(\varphi(\sum_{\mu=0}^{\infty}A_\mu X^\mu)\big)_{ij}=\sum_{\mu=0}^{\infty}(A_\mu)_{ij}X^\mu$$

לכל $1\le i,j\le n$. אז φ איזומורפיזם חוגים.

כך, למשל, אם $R=\mathbb{Z}$ ו־$n=2$,

$$\varphi\Big(\begin{pmatrix}2&1\\-3&0\end{pmatrix}+\begin{pmatrix}7&0\\1&1\end{pmatrix}X+\begin{pmatrix}0&1\\0&0\end{pmatrix}X^2\Big)=\begin{pmatrix}2+7X&1+X^2\\-3+X&X\end{pmatrix}.$$

הוכחה:

(i) φ חח"ע: אכן, אם

$$\varphi(\sum_{\mu=0}^{\infty}A_\mu X^\mu)=\varphi(\sum_{\mu=0}^{\infty}B_\mu X^\mu).$$

אז

$$\sum_{\mu=0}^{\infty}(A_\mu)_{ij}X^\mu = \sum_{\mu=0}^{\infty}(B_\mu)_{ij}X^\mu$$

לכל i,j. מכאן לכל μ

$$(A_\mu)_{ij} = (B_\mu)_{ij} \quad \text{לכל} \quad i,j.$$

כלומר לכל μ מתקיים $A_\mu = B_\mu$. ובפרט

$$\sum_{\mu=0}^{\infty} A_\mu X^\mu = \sum_{\mu=0}^{\infty} B_\mu X^\mu.$$

(ii) φ על: אכן, תהי $C \in \mathrm{M}_n(R[X])$ ויהי $(C)_{ij} = c_{ij}^{(0)} + c_{ij}^{(1)}X + \cdots$. זהו איבר של $R[X]$, לכל $1 \le i,j \le n$. נגדיר, לכל $0 \le \mu < \infty$, $A_\mu \in \mathrm{M}_n(R)$ על ידי $(A_\mu)_{ij} = c_{ij}^{(\mu)}$, לכל $1 \le i,j \le n$. אם μ גדול מספיק, אז $c_{ij}^{(\mu)} = 0$ לכל $1 \le i,j \le n$, ולכן $A_\mu = 0$. ומכאן ש־$\sum_{\mu=0}^{\infty} A_\mu X^\mu$ איבר של $\mathrm{M}_n(R)[X]$. מתקיים

$$\big(\varphi(\sum_{\mu=0}^{\infty} A_\mu X^\mu)\big)_{ij} = \sum_{\mu=0}^{\infty}(A_\mu)_{ij}X^\mu = \sum_{\mu=0}^{\infty} c_{ij}^{(\mu)}X^\mu = (C)_{ij}$$

ומכאן $\varphi(\sum_{\mu=0}^{\infty} A_\mu X^\mu) = C$.

(iii) φ שומרת חיבור: אם $\sum_{\mu=0}^{\infty} A_\mu X^\mu, \sum_{\mu=0}^{\infty} B_\mu X^\mu \in \mathrm{M}_n(R)[X]$ אז

$$\begin{aligned}
&\big(\varphi(\sum_{\mu=0}^{\infty} A_\mu X^\mu + \sum_{\mu=0}^{\infty} B_\mu X^\mu)\big)_{ij} = \big(\varphi(\sum_{\mu=0}^{\infty}(A_\mu + B_\mu)X^\mu)\big)_{ij} = \sum_{\mu=0}^{\infty}(A_\mu + B_\mu)_{ij}X^\mu \\
&= \sum_{\mu=0}^{\infty}\big((A_\mu)_{ij} + (B_\mu)_{ij}\big)X^\mu = \sum_{\mu=0}^{\infty}(A_\mu)_{ij}X^\mu + \sum_{\mu=0}^{\infty}(B_\mu)_{ij}X^\mu \\
&= \big(\varphi(\sum_{\mu=0}^{\infty} A_\mu X^\mu)\big)_{ij} + \big(\varphi(\sum_{\mu=0}^{\infty} B_\mu X^\mu)\big)_{ij} = \big(\varphi(\sum_{\mu=0}^{\infty} A_\mu X^\mu) + \varphi(\sum_{\mu=0}^{\infty} B_\mu X^\mu)\big)_{ij}
\end{aligned}$$

לכל i,j ולכן

$$\varphi(\sum_{\mu=0}^{\infty} A_\mu X^\mu + \sum_{\mu=0}^{\infty} B_\mu X^\mu) = \varphi(\sum_{\mu=0}^{\infty} A_\mu X^\mu) + \varphi(\sum_{\mu=0}^{\infty} B_\mu X^\mu).$$

(iv) φ שומרת כפל: אם $\sum_{\mu=0}^{\infty} A_\mu X^\mu, \sum_{\mu=0}^{\infty} B_\mu X^\mu \in \mathrm{M}_n(R[X])$ אז

$$\varphi\big((\sum_{\mu=0}^{\infty} A_\mu X^\mu)(\sum_{\nu=0}^{\infty} B_\nu X^\nu)\big) =^{(1)} \varphi(\sum_{\mu=0}^{\infty}\sum_{\nu=0}^{\infty} A_\mu B_\nu X^{\mu+\nu})$$

$$=^{(2)} \sum_{\mu=0}^{\infty}\sum_{\nu=0}^{\infty} \varphi(A_\mu B_\nu X^{\mu+\nu})$$

$$\varphi(\sum_{\mu=0}^{\infty} A_\mu X^\mu)\varphi(\sum_{\nu=0}^{\infty} B_\nu X^\nu) =^{(2)} \big(\sum_{\mu=0}^{\infty} \varphi(A_\mu X^\mu)\big)\big((\sum_{\nu=0}^{\infty} \varphi(B_\nu X^\nu)\big)$$

$$=^{(1)} \sum_{\mu=0}^{\infty}\sum_{\nu=0}^{\infty} \varphi(A_\mu X^\mu)\varphi(B_\nu X^\nu)$$

[(1) בגלל חוק הפילוג; (2) - בגלל ש־φ שומרת חיבור] ולכן די להוכיח, לכל $\mu, \nu \geq 0$ כי

$$\varphi(A_\mu B_\nu X^{\mu+\nu}) = \varphi(A_\mu X^\mu)\varphi(B_\nu X^\nu).$$

לשם פשטות נכתוב A במקום A_μ ו־B במקום B_ν. נראה את המשוואה האחרונה:

$$\varphi(ABX^{\mu+\nu}) = \varphi(AX^\mu)\varphi(BX^\nu).$$

ואכן,

$$\big(\varphi(AX^\mu)\varphi(BX^\nu)\big)_{ij} = \sum_{k=1}^{n} \big(\varphi(AX^\mu)\big)_{ik}\big(\varphi(BX^\nu)\big)_{kj} = \sum_{k=1}^{n}(A)_{ik}X^\mu(B)_{kj}X^\nu =$$

$$\sum_{k=1}^{n}(A)_{ik}(B)_{kj}X^{\mu+\nu} = \big(\sum_{k=1}^{n}(A)_{ik}(B)_{kj}\big)X^{\mu+\nu} = (AB)_{ij}X^{\mu+\nu} = \big(\varphi(ABX^{\mu+\nu})\big)_{ij}$$

לבסוף נשים לב שאם יש יחידה $1 \in R$, אז $\varphi(1) = 1$, כלומר, $\varphi(I_n) = I_n$. ∎

הגדרה 1.10: תת חוג. תת קבוצה S של חוג R תכונה **תת חוג** אם היא מכילה את איבר האפס, סגורה תחת חיבור, לקיחת איבר נגדי וכפל. אם R חוג עם יחידה, נדרוש גם ש־S מכילה את היחידה של R.

במקרה זה מהווה S חוג (עם יחידה) ביחס לחיבור והכפל המושרים מ־R.

דגמאות 1.11: (א) $\mathbb{Z}$ הוא תת חוג של $\mathbb{Q}$.

(ב) $2\mathbb{Z}$ הוא תת חוג של $\mathbb{Z}$ שאינו מכיל את איבר היחידה.

(ג) אם R חוג, אז $S = \{a_0 + 0X + 0X^2 + \cdots \mid a_0 \in R\}$ הוא תת חוג של $R[X]$. ברור שההעתקה $a_0 \mapsto a_0 + 0X + 0X^2 + \cdots$ היא איזומורפיזם $R \to S$. אנו נזהה את S עם R ונאמר כי R תת חוג של $R[X]$.

(ד) אם R חוג, אז $\left\{ \begin{pmatrix} \alpha & & 0 \\ & \ddots & \\ 0 & & \alpha \end{pmatrix} = \alpha I_n \middle| \alpha \in R \right\}$ הוא תת חוג של $\mathrm{M}_n(R)$. שוב, הוא איזומורפי ל־R, על ידי $\alpha \mapsto \begin{pmatrix} \alpha & & 0 \\ & \ddots & \\ 0 & & \alpha \end{pmatrix} = \alpha I_n$, ונאמר כי R תת חוג של $\mathrm{M}_n(R)$.

(ה) אם R הוא חוג חילופי עם יחידה ו־$A \in \mathrm{M}_n(R)$ מטריצה, אזי

$$R[A] = \{a_0 I_n + a_1 A + a_2 A^2 + \cdots \mid 0 \text{ כמעט כולם}, a_i \in R\}$$

הוא תת חוג של $\mathrm{M}_n(R)$. אין לבלבל את $R[A]$ עם חוג הפולינומים מעל R, כי A אינו משתנה. הואיל וחזקות שונות של A מתחלפות זו עם זו בכפל, $R[A]$ הנו חוג חילופי.

תרגיל 1.12: *תהי φ כמו בטענה 1.9. תהי $A \in \mathrm{M}_n(R)$. אז (תחת זיהויים מתאימים)*

(א) $\varphi(X - A) = XI_n - A$.

(ב) *לכל $g \in R[X]$ מתקיים $\varphi(g) = gI_n$.*

הוכחה: (א) אם $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \in \mathrm{M}_n(R)$,

אז $XI_n - A = \begin{pmatrix} X - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & X - a_{22} & \dots & -a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ -a_{n1} & -a_{n2} & \dots & X - a_{nn} \end{pmatrix} \in \mathrm{M}_n(R[X])$

ואילו

$$\begin{aligned} X - A &= -AX^0 + I_n X^1 + 0X^2 + \cdots = \\ &= \begin{pmatrix} -a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & -a_{22} & \dots & -a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ -a_{n1} & -a_{n2} & \dots & -a_{nn} \end{pmatrix} + \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} X \in \mathrm{M}_n(R)[X] \end{aligned}$$

לכן

$$\varphi(X - A) = \begin{pmatrix} -a_{11} + X & -a_{12} & \dots & -a_{1n} \\ -a_{21} & -a_{22} + X & \dots & -a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ -a_{n1} & -a_{n2} & \dots & -a_{nn} + X \end{pmatrix}$$

(ב) יהי $g = \sum_{i=0}^{n} a_i X^i$. לפי דוגמה 1.11(ד), אפשר לראות כל $a_i \in R$ כמטריצה $a_i I_n \in \mathrm{M}_n(R)$. לכן $g = \sum_{i=0}^{n} (a_i I_n) X^i \in \mathrm{M}_n(R)[X]$. בפרט $\varphi(g)$ מוגדר. מתקיים

$$\varphi(g) = \sum_{i=0}^{n} (a_i X^i) I_n = \Big(\sum_{i=0}^{n} a_i X^i\Big) I_n = g(X) I_n. \qquad \blacksquare$$

משפט 1.13: *יהי R תחום שלמות. אז קיים שדה כך ש־R תת חוג של שדה זה. בפרט אם F שדה, אז $F[X]$ תחום שלמות ולכן תת חוג של איזשהו שדה.*

הוכחה: (הוכחה זו לא תובא בהרצאה וניתנת כאן רק לשם השלמת התמונה.)

נגדיר יחס שקילות על אסף הזוגות $\{(a,b) \in R^2 |\ b \neq 0\}$:

$$(a,b) \sim (a',b') \iff ab' = a'b.$$

נסמן את מחלקת השקילות של (a,b) ב־$\frac{a}{b}$. נסמן את אסף כל מחלקות השקילות ב־$\mathrm{Quot}(R)$. נגדיר חיבור וכפל על $\mathrm{Quot}(R)$ בדרך הרגילה:

$$\frac{a}{b}\frac{a'}{b'} = \frac{aa'}{bb'}, \qquad \frac{a}{b} + \frac{a'}{b'} = \frac{ab'+a'b}{bb'}.$$

הגדרות אלו אינם תלויות במייצגים והופכות את $\mathrm{Quot}(R)$ לשדה. שדה זה נקרא **שדה המנות** של R. ההעתקה $a \mapsto \frac{a}{1}$ היא הומומורפיזם חד חד ערכי של R לתוך $\mathrm{Quot}(R)$. נזהה את a עם $\frac{a}{1}$ כדי לראות את R כתת חוג של $\mathrm{Quot}(R)$. אם L הוא שדה כלשהו המקיף את R אזי L מקיף גם את $\mathrm{Quot}(R)$. במלים אחרות, $\mathrm{Quot}(R)$ הנו השדה הקטן ביותר המקיף את R.

בפרט שדה המנות של $\mathbb{Z}$ הוא $\mathbb{Q}$. עבור שדה כלשהו F, שדה המנות של $F[X]$ הוא **שדה הפונקציות הרציונליות** מעל F:

$$F(X) = \left\{ \frac{f(X)}{g(X)} \middle| f,g \in F[X],\ g \neq 0 \right\}. \qquad \blacksquare$$

2. חילוק בחוגים ובחוגי פולינומים

הגדרה 2.1: יהי R חוג עם יחידה, ויהי $a \in R$. איבר $b \in R$ ייקרא **הופכי** של a אם

$$ab = 1 = ba.$$

a ייקרא **הפיך** אם יש לו הופכי. נסמן ב־$R^\times$ את קבוצת האיברים ההפיכים של R. ∎

הערה 2.2: אם ל־a יש הופכי אז הופכי זה הוא יחיד (ואז הוא ייקרא **ההופכי** ויסומן a^{-1}). אכן, אם $b, c \in R$ מקיימים $ab = 1 = ba$, $ac = 1 = ca$, אז $c = 1c = (ba)c = b(ac) = b1 = b$. ∎

דוגמאות 2.2: $\mathbb{Z}^\times = \{\pm 1\}$; אם F שדה אז $F^\times = F \setminus \{0\}$ ו־$(\mathrm{M}_n(F))^\times$ היא קבוצת המטריצות ההפיכות.

משפט 2.3 (חילוק עם שארית): *יהי R חוג עם יחידה ויהיו $f, g \in R[X]$,*

$$g = b_m X^m + \cdots + b_1 X + b_0, \quad b_m \in R^\times.$$

אזי

(א) *קיימים $q, r \in R[X]$ יחידים כך ש־*

$$f = qg + r, \qquad \deg r < \deg g = m.$$

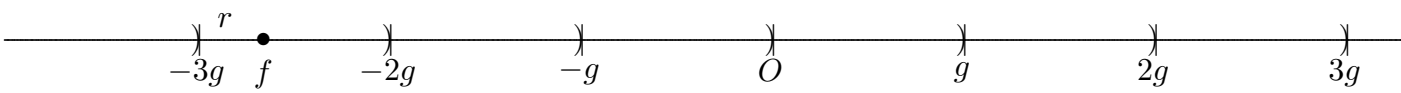
(ב) *קיימים $q', r' \in R[X]$ יחידים כך ש־*

$$f = gq' + r', \qquad \deg r' < \deg g = m.$$

מוכר לנו ודאי המשפט המקביל במספרים שלמים:

משפט (חילוק עם שארית מעל $\mathbb{Z}$): *יהיו $f, g \in \mathbb{Z}$, $g > 0$. אזי קיימים $q, r \in \mathbb{Z}$ יחידים כך ש־*

$$f = qg + r, \qquad 0 \le r < g.$$



הוכחה של (א): **קיום.** באינדוקציה על $\deg f$. אם $\deg f < m$ (ובפרט אם $f = 0$), ניקח $q = 0,\ r = f$. נניח כעת כי $\deg f = n \ge m$, נאמר,

$$f = a_n X^n + a_{n-1} X^{n-1} + \ldots + a_0, \qquad a_n \neq 0$$

וטענת הקיום נכונה לכל הפולינומים ממעלה $< n$. אז

$$f = (a_n b_m^{-1} X^{n-m}) \cdot g + f_1,$$

באשר

$$\begin{aligned} f_1 =& f - (a_n b_m^{-1} X^{n-m}) \cdot g \\ =& (a_n X^n + a_{n-1} X^{n-1} + \ldots) - (a_n b_m^{-1} X^{n-m})(b_m X^m + b_{m-1} X^{m-1} + \ldots) \\ =& (a_n X^n + a_{n-1} X^{n-1} + \ldots \cdot \cdot \cdot) - (a_n X^n + a_n b_m^{-1} b_{m-1} X^{n-1} + \ldots) \\ =& (a_n - a_n) X^n + (a_{n-1} - a_n b_m^{-1} b_{m-1}) X^{n-1} + \ldots \end{aligned}$$

ומכאן $\deg f_1 < n$. לכן לפי הנחת האינדוקציה יש $q_1, r \in R[X]$ כך ש־

$$f_1 = q_1 g + r, \qquad \deg r < m.$$

לכן $f = (a_n b_m^{-1} X^{n-m}) \cdot g + q_1 g + r = qg + r$, באשר $q = a_n b_m^{-1} X^{n-m} + q_1$.

יחידות. נניח כי בנוסף ל־q, r יש גם $r_0, q_0 \in R[X]$ כך ש־

$$f = q_0 g + r_0, \qquad \deg r_0 < \deg g = m.$$

אז $(q - q_0) g = r_0 - r$. אם $q = q_0$ אז מכאן נובע שגם $r = r_0$. אילו היה $q - q_0 \neq 0$, אז לפי תרגיל 1.4

$$\deg g \leq \deg(q - q_0) + \deg g = \deg((q - q_0) \cdot g) = \deg(r - r_0) \leq \max(\deg r_0, \deg r) < \deg g,$$

סתירה. לכן $q = q_0$. ∎

דוגמה: $R = \mathbb{Z}$, $f = X^3 - 7X - 6$, $g = X - 4$.

נעשה את החישוב לפי האלגוריתם שבהוכחת משפט 2.3:

$$\begin{array}{rrrrrl} & X^3 & +0X^2 & -7X & -6 & : \; X - 4 = X^2 + 4X + 9 \\ -\Big(X^2(X-4) = & X^3 & -4X^2 & & \Big) & \\ \hline & & 4X^2 & -7X & -6 & \\ -\Big(4X(X-4) = & & 4X^2 & -16X & \Big) & \\ \hline & & & 9X & -6 & \\ -\Big(9(X-4) = & & & 9X & -36\Big) & \\ \hline & & & & 30 & \end{array}$$

לכן

$$X^3 - 7X - 6 = (X - 4)(X^2 + 4X + 9) + 30.$$

מסקנה 2.4: *יהי R חוג עם יחידה, יהי $\alpha \in R$ ויהי $f \in R[X]$. אז $f(\alpha) = 0$ אם ורק אם קיים $q \in R[X]$ כך ש־$f = q(X) \cdot (X - \alpha)$.*

הוכחה: לפי משפט 2.3(א), עם $g = X - \alpha$, יש $q, r \in R[X]$ יחידים כך ש־

$$f = q(X) \cdot (X - \alpha) + r, \qquad \deg r < \deg(X - \alpha) = 1.$$

כעת $\deg r \leq 0$ פירושו ש־$r \in R$ (ולכן $r(\alpha) = r$, כמובן). נציב α בשני האגפים של המשוואה $f = q(X) \cdot (X - \alpha) + r$ ונקבל $f(\alpha) = q(\alpha) \cdot (\alpha - \alpha) + r = r$. (כאן משתמשים בכך ש־$\alpha$ מתחלף בכפל עם מקדמי $X - \alpha$, הם $1, -\alpha$, ולכן לפי טענה 1.7(2) ההצבה שומרת כפל!) מכאן

$$f = q(X) \cdot (X - \alpha) + f(\alpha). \tag{1}$$

כעת, אם $f(\alpha) = 0$, אז (1) היא ההצגה המבוקשת.

להיפך, אם יש $q' \in R[X]$ כך ש־$f = q'(X) \cdot (X - \alpha)$, אז מהיחידות בחילוק עם שארית נובע $q' = q$ וגם $f(\alpha) = 0$. ∎

הגדרה 2.5: יהי R חוג קומוטטיבי ויהיו $a, b \in R$. נאמר ש־"a מחלק b ב־R", ונסמן $a|b$, אם יש $c \in R$ כך ש־$b = ac$. ∎

מסקנה 2.6: *יהי F שדה, יהי $\alpha \in F$ ויהי $f \in F[X]$. אזי: α שורש של f אם ורק אם $(X - \alpha)|f$ ב־$F[X]$.*

משפט 2.7: *יהי F שדה ויהי $0 \neq f \in F[X]$ ממעלה n. אז ל־f לכל היותר n שורשים שונים ב־F.*

הוכחה: באינדוקציה על n. אם $n = 0$ אז $f \in F^\times$ ומתקיים $f(\alpha) = f \neq 0$ לכל $\alpha \in F$, כלומר ל־f אין שורש ב־F. נניח $n > 0$.

נניח כי הטענה נכונה לכל פולינום ממעלה $n - 1$. אם ל־f אין שורש ב־F, אז סיימנו. אם יש לו שורש $\alpha \in F$ אז לפי המסקנה לעיל $f = (X - \alpha)g$, באשר $g \in F[X]$. לפי תרגיל 1.4(ו), g ממעלה $n - 1$. לפי הנחת האינדוקציה יש ל־g לכל היותר $n - 1$ שורשים שונים ב־F. קל לראות ששורשי f הם בדיוק שורשי g ו־α, לכן ל־f לכל היותר n שורשים שונים ב־F. ∎

מסקנה 2.8: *יהי F שדה, יהי $f \in F[X]$ ממעלה $n \geq 0$, ויהי $c \in F^\times$ המקדם העליון שלו. נניח כי ל־f יש n שורשים שונים $\alpha_1, \ldots, \alpha_n$ ב־F. אזי*

$$f = c(X - \alpha_1) \cdots (X - \alpha_n).$$

הוכחה: יהי

$$g := c(X - \alpha_1) \cdots (X - \alpha_n) \in F[X].$$

ברור ש־g פולינום ממעלה n, עם מקדם עליון c. לכן $f-g \in F[X]$ ממעלה $n-1 \geq$. אולם $(f-g)(\alpha_i) = f(\alpha_i) - g(\alpha_i) = 0$ לכל $1 \leq i \leq n$, כלומר לפולינום $f-g$ לפחות n שורשים שונים ב־F. לפי המשפט זה יתכן רק אם $f-g=0$, כלומר $f=g$. ∎

הגדרה 2.9: פולינום ממעלה ראשונה נקרא **פולינום לינארי**. ∎

הגדרה 2.10: שדה F נקרא **סגור אלגברית** אם לכל $f \in F[X]$ ממעלה $0 <$ יש שורש (לפחות אחד) ב־F. ∎

דוגמה 2.11 : $\mathbb{C}$ סגור אלגברית. (ללא הוכחה! "המשפט היסודי של אלגברה".)

$\mathbb{Q}, \mathbb{R}$ אינם סגורים אלגברית: ל־X^2+1 אין שורש בהם.

משפט 2.12 : *(ללא הוכחה) לכל שדה F קיים שדה סגור אלגברית כך ש־F הוא שדה חלקי שלו.*

משפט 2.13 : *יהי F שדה סגור אלגברית ויהי $f \in F[X]$ ממעלה $n \geq 0$. אזי קיימים $\alpha_1, \ldots, \alpha_n \in F$ (לאו דווקא שונים זה מזה) וקיים $c \in F^\times$ כך ש־*

$$f = c(X-\alpha_1)\cdots(X-\alpha_n).$$

הוכחה: באינדוקציה על n. אם $n=0$ אז $f=c \in F^\times$ מהצורה המבוקשת. נניח $n>0$ ונניח נכונות לכל פולינום ממעלה $n-1$. ל־f יש שורש $\alpha_n \in F$ (כי F סגור אלגברית). לכן $(X-\alpha_n)|f$ ב־$F[X]$, כלומר, יש $g \in F[X]$ כך ש־

$$f = g \cdot (X-\alpha_n).$$

ברור ש־g ממעלה $n-1$, ולכן לפי הנחת האינדוקציה יש $\alpha_1, \ldots, \alpha_{n-1} \in F$, $c \in F^\times$, כך ש־

$$g = c(X-\alpha_1)\cdots(X-\alpha_{n-1}).$$

מכאן $f = c(X-\alpha_1)\cdots(X-\alpha_{n-1})(X-\alpha_n)$. ∎

3. פריקות בחוגי פולינומים

הגדרה 3.1: יהי R חוג עם יחידה. קבוצה $I \subseteq R$ נקראת **אידאל** של R אם

(א) $0 \in I$;

(ב) אם $a, b \in I$ אז גם $a + b \in I$.

(ג) אם $a \in I$ ו־$r \in R$ אז $ra, ar \in I$. ∎

שים לב: אם $a \in I$ אז $-a = (-1)a \in I$.

דוגמאות 3.2:

(1) אם R חוג קומוטטיבי עם יחידה ו־$a \in R$ אז קל לראות ש־

$$(a) := \{ra|\ r \in R\} = \{b \in R|\ \ a|b\}$$

הוא אידאל של R אשר מכיל את a. אידאל מהצורה (a) נקרא **ראשי**.

(2) אם $\varphi\colon R \to S$ הומומורפיזם של חוגים, אז

$$\operatorname{Ker}\varphi := \{r \in R|\ \varphi(r) = 0\}$$

אידאל ב־R. (בדקו! צריך לדעת ש־$\varphi(0) = 0$.)

(3) יהי R חוג עם יחידה. אז $\{0\}$ אידאל של R.

תרגיל 3.3: *יהי R תחום שלמות.*

(א) *יש כלל הצמצום ב־R: אם $ac = bc$ ו־$c \neq 0$ אז $a = b$.*

(ב) *יהיו $a, b \in R$ שונים מ־0. אם $u, v \in R$ מקיימים $b = ua$, $a = vb$ אז $uv = 1$ ובפרט $u, v \in R^\times$,*

הוכחה: (א) אם $ac = bc$ אז $(a - b)c = ac - bc = 0$. כיוון ש־R תחום שלמות, אם גם $c \neq 0$, אז $a - b = 0$. מכאן $a = b$.

(ב) $1 \cdot b = b = ua = u(vb) = (uv)b$. לפי (א), $1 = uv$. כיוון ש־R קומוטטיבי, גם $1 = vu$. מכאן $u, v \in R^\times$, ∎

נזכור שאם F שדה, אז $F[X]$ תחום שלמות, לפי תרגיל 1.4(ה).

משפט 3.4: *יהי F שדה. אז כל אידאל I ב־$R = F[X]$ הוא ראשי. יתר על כן: אם $I \neq \{0\}$, אז*

(א) *קיים $g \in F[X]$ מתוקן* **יחיד** *כך ש־$I = (g)$.*

(ב) *g הוא הפולינום המתוקן היחיד ב־I ממעלה $m := \min(\deg h|\ h \in I \setminus \{0\})$.*

הוכחה: יהי $I \subseteq R$ אידאל. אם $I = \{0\}$ אז $I = (0)$ ראשי. אחרת יש $0 \neq g \in I$ ממעלה m.

טענה: $I = (g) := \{rg|\ r \in R\}$.

ואכן, ההכלה "$\supseteq$" נובעת מכך ש־$g \in I$, ו־I סגור תחת הכפל באברי R. להיפך, יהי $f \in I$. לפי משפט החילוק עם שארית יש $q, r \in R$ כך ש־

$$f = gq + r, \qquad \deg r < \deg g = m.$$

כעת, $f, g \in I$, לכן $r = f - gq = f + (-q)g \in I$. לפי הגדרת m יוצא: $r = 0$. לכן $f = gq \in (g)$.

בה"כ g מתוקן: אם $c \in F^\times$ המקדם העליון של g, אז $c^{-1}g \in I$ מתוקן, ומאותה המעלה.

היחידות ב־(א): אם גם g' מתוקן, $I = (g) = (g')$, אז $g' \in (g)$ ולכן $g|g'$. באופן דומה $g'|g$. לכן לפי תרגיל 3.3 יש $c \in (F[X])^\times = F^\times$ כך ש־$g' = cg$. מכאן, היות ו־g, g' מתוקנים, $g = g'$.

היחידות ב־(ב): אם גם $g' \in I$ ממעלה m, אז לפי ההוכחה $I = (g')$, ולכן לפי (א) $g' = g$. ∎

הגדרה 3.5: יהי R תחום שלמות. יהי $0 \neq q \in R$ לא הפיך.

(א) q נקרא **אי פריק** אם לכל $a, b \in R$ מתקיים: אם $q = ab$ אז $b \in R^\times$ או $a \in R^\times$.

(ב) q נקרא **ראשוני** אם לכל $a, b \in R$ מתקיים: אם $q|ab$ אז $q|b$ או $q|a$. ∎

תרגיל 3.6: *יהי R תחום שלמות, $q \in R$.*

(א) *אם q ראשוני אז הוא אי פריק.*

(ב) *אם q אי פריק, $u \in R^\times$, אז גם qu אי פריק.*

הוכחה: (א) יהיו $a, b \in R$ כך ש־$q = ab$. אז $q|ab$ (כי $ab = q \cdot 1$), לכן, כיוון ש־q ראשוני, $q|b$ או $q|a$.

בה"כ $q|b$ (המקרה השני דומה). כלומר, יש $u \in R$ כך ש־$b = uq$. אבל גם $q = ab$. לכן, לפי תרגיל 3.3(ב), $a \in R^\times$.

(ב) מושאר לקורא. ∎

משפט 3.7: *יהי R תחום שלמות בו כל אידאל ראשי. יהי $q \in R$ אי פריק. אז q ראשוני.*

הוכחה: יהיו $a, b \in R$ כך ש־$q|ab$. הקבוצה

$$I = \{c \in R|\ \ q|cb\}$$

היא אידאל ב־R.

אכן,

(א) $0 \in I$, כלומר, $q|0 \cdot b = 0$, כי $0 = 0 \cdot q$.

(ב) אם $c_1, c_2 \in I$, כלומר, יש $d_1, d_2 \in R$ כך ש־ $c_1b = d_1q,\ c_2b = d_2q$, אז

$$(c_1 + c_2)b = c_1b + c_2b = d_1q + d_2q = (d_1 + d_2)q,$$

כלומר, $q|(c_1 + c_2)b$. לכן $c_1 + c_2 \in I$.

(ג) אם $c \in I$ ו־$r \in R$ אז cb כפולה של q, ולכן $(rc)b$, שהינו כפולה של cb, הוא כפולה של q. לכן $rc \in I$.

הוכחנו ש־I הוא אידאל ב־R. לכן יש $c_0 \in R$ כך ש־$I = (c_0)$.

כעת, $a \in I$, כי $q|ab$. לכן יש $r \in R$ כך ש־$a = rc_0$.

$q \in I$ (כי $q|qb$), לכן יש $u \in R$ כך ש־ $q = uc_0$. היות ו־q אי פריק, c_0 הפיך או u הפיך. במקרה הראשון $1 = c_0^{-1}c_0 \in I$, כלומר $q|b$. במקרה השני $c_0 = u^{-1}q$, לכן $a = rc_0 = (ru^{-1})q$, ולכן $q|a$. ∎

משפט 3.8: *יהי F שדה ויהי $R = F[X]$. יהי $\{q_i\}_{i\in I}$ האוסף של כל הפולינומים האי פריקים המתוקנים (=בעלי מקדם עליון 1) ב־R. אז לכל $0 \neq f \in R$ יש $c \in R^\times = F^\times$ יחיד ו־$m_i \in \mathbb{N} \cup \{0\}$, לכל $i \in I$, כמעט כולם 0, יחידים, כך ש־*

$$f = c\prod_{i\in I} q_i^{m_i}. \qquad (*)$$

הוכחה: (א) קיום ההצגה $(*)$: קודם שני מקרים פרטיים. אם f הפיך, אז $f = f\prod_{i\in I} q_i^0$ היא ההצגה שלו. אם f אי פריק, ו־$c \in F^\times$ המקדם העליון שלו, אז $c^{-1}f$ מתוקן וגם אי פריק, לפי תרגיל 3.6(ב). לכן $c^{-1}f = q_{i_0}$, עבור איזה $i_0 \in I$, ואז $f = cq_{i_0} = cq_{i_0}\prod_{i\neq i_0} q_i^0$ היא ההצגה $(*)$ שלו.

נניח בשלילה שיש $0 \neq f \in R$ שאין לו הצגה $(*)$, ונבחר אותו ממעלה מזערית. לפי האמור לעיל, f אינו הפיך ואינו אי פריק. לכן $f = f_1f_2$ באשר $f_1, f_2 \in R$ לא הפיכים. בפרט $\deg f_1, \deg f_2 \geq 1$ ומכאן $\deg f_1, \deg f_2 < \deg f$ (כי $\deg f_1 + \deg f_2 = \deg f$). לכן גם ל־$f_1$ וגם ל־f_2 הצגות $(*)$; אבל מכפלת ההצגות האלה נותנת הצגה $(*)$ עבור f, סתירה.

(אם $f_2 = c_2\prod_{i\in I} q_i^{n_i}$, $f_1 = c_1\prod_{i\in I} q_i^{m_i}$, אז $f_1f_2 = c_1c_2\prod_{i\in I} q_i^{m_i+n_i}$.)

(ב) יחידות ההצגה $(*)$: נניח כי בנוסף ל־$(*)$ מתקיים גם $(**)$

$$f = c\prod_{i\in I} q_i^{m_i}, \qquad (*)$$

$$f = d\prod_{i\in I} q_i^{n_i} \qquad (**)$$

באשר $c, d \in R^\times = F^\times$, ו־$m_i, n_i \in \mathbb{N} \cup \{0\}$, כמעט כולם 0.

מתוך $(*)$ ברור ש־c הוא המקדם העליון של f. מתוך $(**)$ ברור ש־d הוא המקדם העליון של f. לכן $c = d$ ומכאן גם $\prod_{i\in I} q_i^{m_i} = \prod_{i\in I} q_i^{n_i}$.

נניח בשלילה שיש $j \in I$ כך ש־$m_j \neq n_j$, ובה"כ (מטעמי סימטריה) $m_j < n_j$. נצמצם את השיוויון האחרון ב־$q_j^{m_j}$:

$$\prod_{\substack{i\in I\\ i\neq j}} q_i^{m_i} = q_j^{n_j-m_j}\prod_{\substack{i\in I\\ i\neq j}} q_i^{n_i}.$$

באגף ימין מופיע q_j כגורם, לכן $q_j|\prod_{\substack{i\in I\\ i\neq j}} q_i^{m_i}$. אבל לפי משפט 3.7, q_j ראשוני, לכן יש $i \in I$, $i \neq j$, כך ש־$q_j|q_i$. כלומר $q_i = q_ju$, באשר $u \in R$. כעת q_i אי פריק, לכן $u \in R^\times = F^\times$. אבל q_i, q_j מתוקנים, לכן $u = 1$. מכאן $q_i = q_j$, בסתירה ל־$i \neq j$. ∎

תרגיל 3.9: *יהי E תת שדה של שדה F ויהיו $f,g \in E[X]$. נניח כי $f|g$ ב־$F[X]$. הוכיחו ש־$f|g$ ב־$E[X]$.*

תרגיל 3.10: *יהי F שדה סגור אלגברית ויהי $f \in F[X]$ מתוקן (בפרט שונה מ־0). אז f אי פריק $\iff f = X - \alpha$, באשר $\alpha \in F$.*

תרגיל 3.11: *יהיו $b,c \in \mathbb{R}$ כך ש־ $b^2 - 4c < 0$. הוכיחו ש־$f = X^2 + bX + c$ אי פריק ב־$\mathbb{R}[X]$.*

משפט 3.12: *יהי $f \in \mathbb{R}[X]$ מתוקן. אז f אי פריק אם ורק אם*

(א) $f = X - \alpha$, *באשר* $\alpha \in \mathbb{R}$; *או*

(ב) $f = X^2 + bX + c$, *באשר* $b,c \in \mathbb{R}$ *כך ש־* $b^2 - 4c < 0$.

הוכחה: אם f מהצורה (א), ברור שהוא אי פריק. אם הוא מהצורה (ב) הוא אי פריק לפי תרגיל 3.11.

להיפך, נניח כי f מתוקן ואי פריק. אז $f \in \mathbb{C}[X]$, לכן לפי משפט 2.13,

$$f = c(X - \alpha_1)\cdots(X - \alpha_n)$$

באשר $c \in \mathbb{C}^\times$ ו־$\alpha_1,\ldots,\alpha_n \in \mathbb{C}$; כאן $c = 1$, כי f מתוקן. נעיר ש־$n \geq 1$, כי f אינו הפיך.

מתקיים $f(\alpha_1) = 0$. נראה שגם $f(\overline{\alpha_1}) = 0$ (הגג מסמן את ההצמדה המרוכבת). ואכן, אם $f = \sum_{i=0}^{\infty} a_i X^i$, באשר $a_0, a_1, \ldots \in \mathbb{R}$, אז

$$f(\overline{\alpha_1}) = \sum_{i=0}^{\infty} a_i \overline{\alpha_1}^i = \sum_{i=0}^{\infty} \overline{a_i}\,\overline{\alpha_1^i} = \sum_{i=0}^{\infty} \overline{a_i \alpha_1^i} = \overline{\sum_{i=0}^{\infty} a_i \alpha_1^i} = \overline{f(\alpha_1)} = \overline{0} = 0$$

כי ההצמדה שומרת כפל וחיבור.

לכן יש $1 \leq j \leq n$ (לא בהכרח יחיד) כך ש־$\overline{\alpha_1} = \alpha_j$. כעת תיתכנה שתי אפשרויות:

(1) $\overline{\alpha_1} = \alpha_1$, כלומר $\alpha_1 \in \mathbb{R}$. אז, לפי מסקנה 2.6, יש $q \in \mathbb{R}[X]$ כך ש־$f = q \cdot (X - \alpha_1)$. אבל f אי פריק, ו־$X - \alpha_1$ אינו הפיך ב־$\mathbb{R}[X]$, לכן $q \in (\mathbb{R}[X])^\times = \mathbb{R}^\times$. כיוון ש־$f$ מתוקן, $q = 1$ ו־$f = X - \alpha_1$.

(2) $\overline{\alpha_1} \neq \alpha_1$, (ובפרט $\operatorname{Im}\alpha_1 \neq 0$). אז $\overline{\alpha_1} = \alpha_j$, באשר $j > 1$. מכאן

$$g := (X - \alpha_1)(X - \alpha_j) = (X - \alpha_1)(X - \overline{\alpha_1}) = X^2 + bX + c$$

באשר $b = -(\alpha_1 + \overline{\alpha_1}), c = \alpha_1\overline{\alpha_1} \in \mathbb{R}$, ולכן

$$b^2 - 4c = (\alpha_1 + \overline{\alpha_1})^2 - 4\alpha_1\overline{\alpha_1} = (\alpha_1 - \overline{\alpha_1})^2 = (2i\operatorname{Im}\alpha_1)^2 = -4(\operatorname{Im}\alpha_1)^2 < 0$$

כעת $g|f$ ב־$\mathbb{C}[X]$, לכן לפי תרגיל 3.9, גם ב־$\mathbb{R}[X]$ ומכאן, כמו ב־(1), $f = g$, כי f אי פריק ומתוקן. ∎

המשפט הבא שימושי כאשר רוצים לפרק פולינום ב־$\mathbb{Q}[X]$ לגורמים:

למה 3.13 (הלמה של גאוס): *אם $f \in \mathbb{Z}[X]$ מתוקן, ויש $g, h \in \mathbb{Q}[X]$ מתוקנים כך ש־$f = gh$, אז $g, h \in \mathbb{Z}[X]$.*

מסקנה 3.14: *יהי $f \in \mathbb{Z}[X]$ מתוקן, ויהי $\alpha \in \mathbb{Q}$ שרשו. אז $\alpha \in \mathbb{Z}$.*

הוכחה: $f(\alpha) = 0$, לכן $(X - \alpha)|f$ ב־$\mathbb{Q}[X]$, כלומר יש $h \in \mathbb{Q}[X]$ כך ש־$f = (X - \alpha)h$. ברור ש־h מתוקן, לכן, לפי למה 3.13, $X - \alpha \in \mathbb{Z}[X]$, כלומר, $\alpha \in \mathbb{Z}$. ∎

הגדרה 3.15: יהי F שדה ויהיו $f_1, \ldots, f_k \in F[X]$, לא כולם אפס. פולינום מתוקן $d \in F[X]$ ייקרא **מחלק משותף** של $f_1, \ldots, f_k$, אם $d|f_1, \ldots, f_k$ ב־$F[X]$. מחלק משותף d של $f_1, \ldots, f_k$ ייקרא **הגדול ביותר** ויסומן $\gcd(f_1, \ldots, f_k)$, אם הוא מתחלק בכל מחלק משותף d' של $f_1, \ldots, f_k$, כלומר, d מוגדר על ידי:

d מתוקן ו־$d|f_1, \ldots, f_k$ ב־$F[X]$, ואם d' מתוקן ו־$d'|f_1, \ldots, f_k$ אז $d'|d$. ∎

דוגמה 3.16: $k = 1$, $f_1 \neq 0$. אז $\gcd(f_1) = c^{-1}f_1$, באשר c המקדם העליון של f_1.

אכן, פולינום זה מתוקן, הוא מחלק את f_1, כי $f_1 = c(c^{-1}f_1)$, ואם d' מחלק את f_1, אז הוא מחלק את $c^{-1}f_1$. ∎

בניה 3.17: האלגוריתם של אוקלידס (Euclides). זהו אלגוריתם למציאת המחלק המשותף הגדול ביותר d של שני פולינומים $f_1, f_2 \in F[X]$ מעל שדה F, לא שניהם אפס.

נניח $f_2 \neq 0$. נבצע סדרה של חילוקים עם שארית ב־$F[X]$:

$$
\begin{array}{llll}
(1) & f_1 = f_2q_2 + f_3, & \deg f_3 < \deg f_2, & f_3 \neq 0 \\
(2) & f_2 = f_3q_3 + f_4, & \deg f_4 < \deg f_3, & f_4 \neq 0 \\
\vdots & \vdots & \vdots & \vdots \\
(i) & f_i = f_{i+1}q_{i+1} + f_{i+2}, & \deg f_{i+2} < \deg f_{i+1}, & f_{i+2} \neq 0 \\
\vdots & \vdots & \vdots & \vdots \\
(r-1) & f_{r-1} = f_rq_r + f_{r+1}, & \deg f_{r+1} < \deg f_r, & f_{r+1} \neq 0 \\
(r) & f_r = f_{r+1}q_{r+1} + f_{r+2}, & & f_{r+2} = 0
\end{array}
$$

כאן f_{r+2} היא השארית הראשונה בסדרה זו השווה לאפס. התהליך חייב להסתיים, כלומר, יש r כך ש־$f_{r+2} = 0$, (ויתכן כי $r = 1$ וכך יש להבין את סדרת המשוואות הנ"ל!), כי הסדרה $\{\deg f_i\}_{i \geq 2}$ יורדת ממש.

כעת יהי $d \in F[X]$ מתוקן כלשהו. נשים לב כי

מתוך (1) נובע: $d|f_2, f_3 \Leftrightarrow d|f_1, f_2$

מתוך (2) נובע: $d|f_3, f_4 \Leftrightarrow d|f_2, f_3$

$\vdots$

מתוך (r) נובע: $d|f_{r+1}, 0 \Leftrightarrow d|f_r, f_{r+1}$

לכן $d|f_{r+1} \Leftrightarrow d|f_{r+1}, 0 \Leftrightarrow d|f_1, f_2$. מכאן $\gcd(f_1, f_2) = \gcd(f_{r+1})$. לפי דוגמה 3.16, $\gcd(f_1, f_2) = u^{-1}f_{r+1}$, באשר $u \in F^\times$ המקדם העליון של f_{r+1}. ∎

למה 3.18: *יהי $d = \gcd(f_1, f_2)$. אזי קיימים $g, h \in F[X]$ (לא בהכרח יחידים) כך ש־*

$$d = gf_1 + hf_2.$$

הוכחה: נראה (בסימונים דלעיל) שלכל $1 \le i \le r+1$ קיימים $g_i, h_i \in F[X]$ כך ש־

$$d = g_i f_i + h_i f_{i+1} \qquad (*_i)$$

(ואז נקח $g = g_1, h = h_1$). ההוכחה היא באינדוקציה **יורדת** על i.

עבור $i = r+1$ נקח $g_{r+1} = u^{-1}, h_{r+1} = 0$. נניח נכונות עבור $1 \le i+1 \le r+1$ מסויים, כלומר,

$$d = g_{i+1}f_{i+1} + h_{i+1}f_{i+2}. \qquad (*_{i+1})$$

אז לפי המשוואה (i) לעיל מתקיים $f_{i+2} = f_i - f_{i+1}q_{i+1}$. אם נציב זאת ב־$(*_{i+1})$, נקבל

$$d = g_{i+1}f_{i+1} + h_{i+1}(f_i - f_{i+1}q_{i+1}) = h_{i+1}f_i + (g_{i+1} - h_{i+1}q_{i+1})f_{i+1}.$$

לכן אם נסמן $g_i = h_{i+1},\ h_i = g_{i+1} - h_{i+1}q_{i+1}$, נקבל את $(*_i)$. ∎

פולינומים f_1, f_2 נקראים **זרים** אם $\gcd(f_1, f_2) = 1$.

תרגיל 3.19: *יהי F שדה, ויהיו $f_1, f_2, f_3 \in F[X]$ שונים מאפס. אם $f_1|f_2f_3$ ו־f_1, f_2 זרים אז $f_1|f_3$.*

הוכחה: לפי ההנחה יש $q \in F[X]$ כך ש־$qf_1 = f_2f_3$. לפי למה 3.18 יש $g, h \in F[X]$ כך ש־$1 = gf_1 + hf_2$. לכן

$$f_3 = (gf_1 + hf_2)f_3 = gf_1f_3 + hf_2f_3 = gf_3f_1 + hqf_1 = (gf_3 + hq)f_1$$

מתחלק ב־f_1. ∎

תרגיל 3.20: *יהי F שדה, ויהיו $f, g \in F[X]$ שונים מאפס,*

$$g = d\prod_{i\in I} q_i^{m_i}, \qquad d \in F^\times \qquad\qquad f = c\prod_{i\in I} q_i^{n_i}, \qquad c \in F^\times$$

הפירוקים שלהם לגורמים אי פריקים ב־$F[X]$. אזי

$$g|f \text{ ב־} F[X] \Leftrightarrow m_i \le n_i \text{ לכל } i \in I.$$

הוכחה: "$\Rightarrow$": יהי $h = cd^{-1}\prod_{i\in I} q_i^{n_i - m_i}$, אז $h \in F[X]$ ו־$hg = f$. לכן $g|f$.

"$\Leftarrow$": יהי $h \in F[X]$ כך ש־$hg = f$. אז $h \neq 0$, לכן יש לו הצגה

$$h = u\prod_{i\in I} q_i^{k_i}, \qquad u \in F^\times,\ k_i \geq 0$$

וברור ש־ $f = hg = ud\prod_{i\in I} q_i^{k_i+m_i}$. מהיחידות של ההצגה נובע $k_i + m_i = n_i$ לכל $i \in I$, בפרט $n_i - m_i = k_i \geq 0$. ∎

הערה 3.21: אם יודעים לפרק שני פולינומים $f_1, f_2 \in F[X]$ שונים מאפס לגורמים אי פריקים ב־$F[X]$, נאמר

$$\begin{aligned} f_1 &= c_1\prod_{i\in I} q_i^{m_i}, \qquad c_1 \in F^\times,\ m_i \geq 0, \\ f_2 &= c_2\prod_{i\in I} q_i^{n_i}, \qquad c_2 \in F^\times,\ n_i \geq 0, \end{aligned} \tag{1}$$

אז לפי תרגיל 3.20

$$\gcd(f_1, f_2) = \prod_{i\in I} q_i^{\min(m_i,n_i)}.$$

אכן, אם $d', d \in F[X]$ מתוקים, אז

$$d' = \prod_{i\in I} q_i^{k'_i}, \qquad d = \prod_{i\in I} q_i^{k_i},$$

באשר $k'_i, k_i \geq 0$ לכל $i \in I$ וכמעט כולם אפסים. לפי תרגיל 3.20,

$$d', d|f_1, f_2 \Leftrightarrow k'_i, k_i \leq m_i, n_i \Leftrightarrow k'_i, k_i \leq \min(m_i, n_i), \quad \text{לכל } i.$$

$$d'|d \Leftrightarrow k'_i \leq k_i \quad \text{לכל } i$$

מכאן ברור ש־$d = \gcd(f_1, f_2)$ אם ורק אם לכל $i \in I$ מתקיים: $k_i \leq \min(m_i, n_i)$ וגם $k'_i \leq k_i$ לכל k'_i שמקיים $k'_i \leq \min(m_i, n_i)$, כלומר $k_i = \min(m_i, n_i)$ לכל $i \in I$. ∎

הגדרה 3.22: יהי F שדה ויהיו $f_1, \ldots, f_k \in F[X]$ שונים מאפס. פולינום מתוקן $m \in F[X]$ ייקרא **כפולה משותפת** של $f_1, \ldots, f_k$, אם $f_1, \ldots, f_k|m$ ב־$F[X]$. כפולה משותפת של $f_1, \ldots, f_k$ תיקרא **הקטנה ביותר**, ותסומן $\mathrm{lcm}(f_1, \ldots, f_k)$, אם היא מחלקת כל כפולה משותפת m' של $f_1, \ldots, f_k$, כלומר,

אם m' מתוקן ו־ $f_1, \ldots, f_k|m'$ אז $m|m'$. ∎

הערה 3.23: יהיו f_1, f_2 כמו במשוואה (1) של הערה 3.21. אז $m = \prod_{i\in I} q_i^{\max(m_i,n_i)}$. ההוכחה – כמו בהערה הקודמת. ∎

הערה 3.24: יהי F שדה ויהיו $f_1, f_2 \in F[X]$ מתוקנים. יהי d המחלק המשותף הגדול ביותר ותהי m הכפולה המשותפת הקטנה ביותר שלהם. אז $dm = f_1 f_2$.

אכן, יהיו f_1, f_2 נתונים על ידי (1). הם מתוקנים, לכן $c_1 = c_2 = 1$. אז $d = \prod_{i\in I} q_i^{\min(m_i,n_i)}$ ו־$m = \prod_{i\in I} q_i^{\max(m_i,n_i)}$. אבל $\max(m_i, n_i) + \min(m_i, n_i) = m_i + n_i$, ומכאן המסקנה. ∎

4. ערכים עצמיים וקטורים עצמיים

יהי F שדה. יהי V מרחב וקטורי מעל F, תהי $T\colon V \to V$ העתקה לינארית, תהי $A \in \mathrm{M}_n(F)$, ויהי $\lambda \in F$.

הגדרה 4.1: הקבוצה

$$V_\lambda = V_\lambda(T) = \{v \in V \mid T(v) = \lambda v\}$$

תקרא **המרחב העצמי** של T (eigenspace) **השייך ל־λ**. וקטור $v \in V$ ייקרא **וקטור עצמי** של T (eigenvector) **השייך ל־λ**, אם $T(v) = \lambda v$, כלומר $v \in V_\lambda$. (בד"כ דורשים גם $v \neq 0$, אך אנו לא נעשה זאת כאן.) הסקלר λ ייקרא **ערך עצמי** (eigenvalue) של T אם יש $0 \neq v \in V$ כך ש־$T(v) = \lambda v$, כלומר, אם $V_\lambda \neq 0$. ∎

דוגמה 4.2: תהי $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ נתונה על ידי $T((a,b)) = (a,-b)$. אז 1 ערך עצמי של T ו־$(x,0)$ וקטור עצמי ששייך לו, לכל $x \in \mathbb{R}$. כמו כן, -1 ערך עצמי של T, ו־$(0,y)$ וקטור עצמי ששייך לו, לכל $y \in \mathbb{R}$. ∎

באופן דומה:

הגדרה 4.3: הקבוצה

$$V_\lambda = V_\lambda(A) = \{v \in F^n \mid Av = \lambda v\}$$

תקרא **המרחב עצמי של A השייך ל־λ**. וקטור $v \in F^n$ **וקטור עצמי של A השייך ל־λ** אם $Av = \lambda v$. כמו כן $\lambda \in F$ **ערך עצמי של A**, אם יש $0 \neq v \in F^n$ כך ש־$Av = \lambda v$. ∎

למה 4.4: *יהי $\mathcal{B} = (v_1, \dots, v_n)$ בסיס סדור של V. תהי $A = [T]_{\mathcal{B}}^{\mathcal{B}} \in \mathrm{M}_n(F)$. אזי*

(א) *$v \in V$ הוא וקטור עצמי של T השייך ל־λ אם ורק אם $[v]_{\mathcal{B}} \in F^n$ הוא וקטור עצמי של A השייך ל־λ.*

(ב) *λ הוא ערך עצמי של T אם ורק אם λ ערך עצמי של A.*

הוכחה: נזכור שההעתקה $V \to F^n$ הנתונה על ידי $v \mapsto [v]_{\mathcal{B}}$ היא איזומורפיזם. יהי $\lambda \in F$ ויהי $v \in V$. אז $[\lambda v]_{\mathcal{B}} = \lambda [v]_{\mathcal{B}}$ (לפי האמור לעיל). כמו כן $[T(v)]_{\mathcal{B}} = A[v]_{\mathcal{B}}$ ולכן

$$A[v]_{\mathcal{B}} = \lambda [v]_{\mathcal{B}} \Leftrightarrow [T(v)]_{\mathcal{B}} = [\lambda v]_{\mathcal{B}} \Leftrightarrow T(v) = \lambda v \tag{1}$$

ומכאן (א).

(ב) אם λ ערך עצמי של T אז יש $0 \neq v \in V$ כך ש־$T(v) = \lambda v$. כיוון ש־$v \mapsto [v]_{\mathcal{B}}$ איזומורפידם, $[v]_{\mathcal{B}} \neq 0$. לפי (1), λ ערך עצמי של A.

להיפך, אם λ ערך עצמי של A, יש $0 \neq w \in F^n$ כך ש־$Aw = \lambda w$. כיוון ש־$v \mapsto [v]_{\mathcal{B}}$ היא על F^n, יש $v \in V$ כך ש־ $w = [v]_{\mathcal{B}}$. כיוון ש־$w \neq 0$, גם $v \neq 0$. לפי (1), λ ערך עצמי של T. ∎

בפרט λ הוא ערך עצמי של A אם ורק אם λ ערך עצמי של T_A, באשר $T_A\colon F^n \to F^n$ מוגדרת על ידי $T_A(v) = Av$, לכל $v \in F^n$. (כי A היא המטריצה של T_A לפי הבסיס הסטנדרטי.)

משפט 4.5:

(א) $V_\lambda(T) = \mathrm{Ker}(T - \lambda 1_V)$. *בפרט* $V_\lambda(T)$ *תת מרחב של* V. *לכן* λ *הוא ערך עצמי של* T *אם ורק אם ההעתקה* $T - \lambda 1_V$ *אינה חח"ע.*

(ב) *יהיו* $A \in \mathrm{M}_n(F)$, $\lambda \in F$. *אז* $V_\lambda(A)$ *הוא מרחב הפתרונות של המערכת ההומוגנית של משוואות ליניאריות* $(A - \lambda I_n)X = 0$. *לכן* λ *ערך עצמי של* A *אם ורק אם* $\det(A - \lambda I_n) = 0$.

הוכחה: (א) יהי $v \in V$. אז

$$v \in V_\lambda(T) \Leftrightarrow T(v) = \lambda v \Leftrightarrow (T - \lambda 1_V)(v) = 0 \Leftrightarrow v \in \mathrm{Ker}(T - \lambda 1_V)$$

לכן $V_\lambda(T) = \mathrm{Ker}(T - 1_V)$. כעת λ הוא ערך עצמי של T אם ורק אם $V_\lambda(T) \neq \{0\}$ אם ורק אם ההעתקה $T - \lambda 1_V$ אינה חח"ע.

(ב) באפן דומה. יהי $v \in F^n$. אז

$$v \in V_\lambda(A) \Leftrightarrow Av = \lambda v \Leftrightarrow (A - \lambda I_n)v = 0$$

לכן λ הוא ערך עצמי של A אם ורק אם למטריצה $A - \lambda I_n$ מאפס לא טריביאלי, כלומר $\det(A - \lambda I_n) = 0$.

∎

מטריצה $C \in \mathrm{M}_n(F)$ נקראת **אלכסונית** אם $(C)_{ij} = 0$ לכל $i \neq j$, כלומר $C = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$.

פעולת החיבור והכפל בין מטריצות אלכסוניות פשוטות יותר:

תרגיל 4.6: יהיו $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots \beta_n \in F$. אז

$$\begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix} \underset{\cdot}{+} \begin{pmatrix} \beta_1 & & 0 \\ & \ddots & \\ 0 & & \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_1 \underset{\cdot}{+} \beta_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \underset{\cdot}{+} \beta_n \end{pmatrix},$$

ואם $\alpha_1, \ldots, \alpha_n \in F^\times$ אז $\begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}^{-1} = \begin{pmatrix} \alpha_1^{-1} & & 0 \\ & \ddots & \\ 0 & & \alpha_n^{-1} \end{pmatrix}$.

משפט 4.7: *יהי* $\mathcal{B} = (v_1, \ldots, v_n)$ *בסיס סדור של* V. *אז* $[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ *אם ורק אם* v_j *וקטור עצמי של* T *השייך ל־* λ_j, *לכל* $j = 1, \ldots, n$. *אם זה קורה אז* $\lambda_1, \ldots, \lambda_n$ *ערכים עצמיים של* T.

הוכחה:

$$[T]_{\mathcal{B}}^{\mathcal{B}} = ([T(v_1)]_{\mathcal{B}}, \ldots, [T(v_n)]_{\mathcal{B}}),$$

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = (\lambda_1 \mathbf{e}_1, \ldots, \lambda_n \mathbf{e}_n) = (\lambda_1 [v_1]_{\mathcal{B}}, \ldots, \lambda_n [v_n]_{\mathcal{B}}),$$

לכן $[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ אם ורק אם $[T(v_j)]_{\mathcal{B}} = \lambda_j [v_j]_{\mathcal{B}}$ לכל $1 \leq j \leq n$ אם ורק אם (בגלל שההעתקה $v \mapsto [v]_{\mathcal{B}}$ היא איזומורפיזם) $T(v_j) = \lambda_j v_j$ לכל $1 \leq j \leq n$. מכאן השקילות. הטענה האחרונה של המשפט נובעת מכך ש־$v_1, \ldots, v_n$ שונים מאפס, כי הם אברי בסיס. ∎

מסקנה 4.8: *ל־T יש הצגה אלכסונית אם ורק אם ל־V יש בסיס המורכב מוקטורים עצמיים של T.*

מסקנה 4.9: *אם יש $P \in \mathrm{M}_n(F)$ הפיכה כך ש־$P^{-1}AP = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, אז עמודותיה $v_1, \ldots, v_n$ בסיס של F^n המורכב מוקטורים עצמיים של A, כאשר v_j שייך ל־λ_j, לכל $1 \leq j \leq n$.*

להיפך, אם יש בסיס כזה ונגדיר $P = (v_1, \ldots, v_n) \in \mathrm{M}_n(F)$, אז $P^{-1}AP = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$.

לכן A דומה למטריצה אלכסונית אם ורק אם יש בסיס של F^n המורכב מוקטורים עצמיים של A.

תהי $P \in \mathrm{M}_n(F)$ הפיכה, ויהיו $v_1, \ldots, v_n \in F^n$ עמודותיה. אז $v_1, \ldots, v_n$ בלתי תלויים לינארית מעל F, ובפרט בסיס של F^n. מתקיים (לכל $\lambda_1, \ldots, \lambda_n \in F$)

$$AP = A(v_1, \ldots, v_n) = (Av_1, \ldots, Av_n),$$

$$P \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = P(\lambda_1 \mathbf{e}_1, \ldots, \lambda_n \mathbf{e}_n) = (\lambda_1 P\mathbf{e}_1, \ldots, \lambda_n P\mathbf{e}_n) = (\lambda_1 v_1, \ldots, \lambda_n v_n).$$

לכן

$$\begin{aligned} P^{-1}AP = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} &\Leftrightarrow AP = P \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \\ &\Leftrightarrow (Av_1, \ldots, Av_n) = (\lambda_1 v_1, \ldots, \lambda_n v_n) \\ &\Leftrightarrow 1 \leq j \leq n \text{ לכל } Av_j = \lambda_j v_j \end{aligned}$$

ומכאן המסקנה. ∎

משפט 4.10: *יהיו $v_1, \ldots, v_m$ וקטורים עצמיים של T **השונים מאפס** והשייכים לערכים עצמיים $\lambda_1, \ldots, \lambda_m$, בהתאמה. נניח ש־$\lambda_1, \ldots, \lambda_m$ שונים זה מזה. אז $v_1, \ldots, v_m$ בלתי תלויים לינארית מעל F.*

הוכחה: באינדוקציה על m. עבור $m = 1$ זה ברור: $v_1 \neq 0$, לכן v_1 בת"ל. נניח נכונות עבור $m - 1$ וקטורים. יהיו $\alpha_1, \ldots, \alpha_m \in F$ כך ש־

$$\alpha_1 v_1 + \cdots + \alpha_{m-1} v_{m-1} + \alpha_m v_m = 0. \tag{1}$$

נפעיל T על שני האגפים ונקבל

$$\alpha_1\lambda_1 v_1 + \cdots + \alpha_{m-1}\lambda_{m-1}v_{m-1} + \alpha_m\lambda_m v_m = 0. \tag{2}$$

נכפיל את (1) ב־λ_m ונחסיר אותה מ־(2):

$$\alpha_1(\lambda_1 - \lambda_m)v_1 + \cdots + \alpha_{m-1}(\lambda_{m-1} - \lambda_m)v_{m-1} = 0.$$

לפי הנחת האינדוקציה $v_1, \ldots, v_{m-1}$ בלתי תלויים לינארית, לכן, לכל $1 \le i \le m-1$, $\alpha_i(\lambda_i - \lambda_m) = 0$; היות ו־ $\lambda_i - \lambda_m \neq 0$, יוצא $\alpha_i = 0$. אם נציב זאת ב־(1), נקבל $\alpha_m v_m = 0$, ומכאן $\alpha_m = 0$, כי $v_m \neq 0$.

∎

מסקנה 4.11: *יהיו $\lambda_1, \ldots, \lambda_n$ ערכים עצמיים של T, שונים זה מזה. לכל $1 \le i \le n$ תהי $v_{i1}, v_{i2}, \ldots, v_{ik_i}$ סדרה של וקטורים בלתי תלויים לינארית ב־ V_{λ_i}. אזי הסדרה*

$$v_{11}, v_{12}, \ldots, v_{1k_1},\ v_{21}, v_{22}, \ldots, v_{2k_2}\ , \ldots,\ v_{n1}, v_{n2}, \ldots, v_{nk_n}$$

בלתי תלויה לינארית.

הוכחה : יהיו $\alpha_{11}, \alpha_{12}, \ldots, \alpha_{1k_1},\ \alpha_{21}, \alpha_{22}, \ldots, \alpha_{2k_2},\ \ldots,\ \alpha_{n1}, \alpha_{n2}, \ldots, \alpha_{nk_n} \in F$ כך ש־

$$(\alpha_{11}v_{11} + \cdots + \alpha_{1k_1}v_{1k_1}) + \cdots + (\alpha_{n1}v_{n1} + \cdots + \alpha_{nk_n}v_{nk_n}) = 0. \tag{3}$$

נסמן

$$w_i = \alpha_{i1}v_{i1} + \cdots + \alpha_{ik_i}v_{ik_i}, \qquad i = 1, \ldots, n. \tag{4}$$

אז $w_i \in V_{\lambda_i}$, כי V_{λ_i} תת מרחב של V ו־$v_{i1}, \ldots, v_{ik_i} \in V_{\lambda_i}$. את (3) אפשר לרשום כך

$$w_1 + w_2 + \cdots + w_n = 0. \tag{3'}$$

אם $w_1 = w_2 = \cdots = w_n = 0$ אז לפי (4)

$$\alpha_{i1}v_{i1} + \cdots + \alpha_{ik_i}v_{ik_i} = 0, \qquad i = 1, \ldots, n$$

ומכאן $\alpha_{i1} = \cdots = \alpha_{ik_i} = 0$, כי $v_{i1}, \ldots, v_{ik_i}$ בלתי תלויים לינארית, ואז סיימנו.

נניח בשלילה ש־$w_1, \ldots, w_n$ לא כולם אפס. יהיו $w_{j_1}, \ldots, w_{j_m}$ השונים מאפס מתוכם. הם וקטורים עצמיים של T השייכים לערכים עצמיים שונים. לפי המשפט הקודם הם בלתי תלויים לינארית. אבל לפי $(3')$

$$1w_{j_1} + \cdots + 1w_{j_m} = 0,$$

סתירה. ∎

דוגמה 4.12: העלאה בחזקה של מטריצה, עם מעריך גדול, היא פעולה נפוצה בשימושים של אלגברה לינארית. נביא כאן דוגמה לבעיה המובילה לפעולה כזאת ונצביע על דרך פשוטה של ביצוע הפעולה.

במדינה מסויימת פרצה מגפה. אם ביום מסוים x_0 מתושבי העיר בריאים ו y_0 חולים, אזי למחרת 10% מהבריאים נדבקים במחלה, ואילו 50% מהחולים מבריאים. במלים אחרות, אם נסמן ב־x_n את מספר הבריאים וב־y_n את מספר החולים אחרי n ימים, אז:

$$\begin{aligned} x_1 &= 0.9x_0 + 0.5y_0 \\ y_1 &= 0.1x_0 + 0.5y_0 \end{aligned} \tag{1}$$

נשאלת השאלה, מה יקרה במשך הזמן, האם המגפה תעלם או חס וחלילה יחלו בסופו של דבר כולם במחלה?

כדי לענות על שאלה זו, קודם נרשום את (1) בעזרת מטריצות:

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = A \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

באשר $A = \begin{pmatrix} 0.9 & 0.5 \\ 0.1 & 0.5 \end{pmatrix}$. מכאן, באינדוקציה על n,

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = A^n \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}. \tag{1}$$

אילו A היתה אלכסונית, $A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, אז את A^n קל לחשב: $A^n = \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix}$. אבל A אינה אלכסונית.

אם נמצא $P \in \mathrm{M}_2(\mathbb{R})$ כך ש־$D = P^{-1}AP$ אלכסונית, אז $A = PDP^{-1}$ ומכאן

$$A^n = (PDP^{-1})^n = \overbrace{(PDP^{-1})(PDP^{-1})(PDP^{-1})\cdots(PDP^{-1})}^{n\times} = PD^nP^{-1} \tag{3}$$

ואת זה קל לחשב.

נמצא P כזאת:

(א) נחשב את הערכים העצמיים של A: כזכור, λ ערך עצמי של A אם ורק אם $\det(A - \lambda I_2) = 0$, כלומר

$$0 = \begin{vmatrix} 0.9-\lambda & 0.5 \\ 0.1 & 0.5-\lambda \end{vmatrix} = (0.9-\lambda)(0.5-\lambda) - 0.5\cdot 0.1 = \lambda^2 - 1.4\lambda + 0.4 = (\lambda-1)(\lambda-0.4)$$

לכן ל־A שני ערכים עצמיים, $\lambda_1 = 1,\ \lambda_2 = 0.4$.

(ב) נחשב עבור כל אחד מהערכים העצמיים וקטורים עצמיים:

$$A - \lambda_1 I_2 = \begin{pmatrix} 0.9-1 & 0.5 \\ 0.1 & 0.5-1 \end{pmatrix} = \begin{pmatrix} -0.1 & 0.5 \\ 0.1 & -0.5 \end{pmatrix},$$

לכן למערכת $(A-\lambda_1 I_2)v_1=0$ יש פתרון (יחיד עד כדי כפל בסקלר) $v_1=\begin{pmatrix}5\\1\end{pmatrix}$.

$$A-\lambda_2 I_2=\begin{pmatrix}0.9-0.4 & 0.5\\ 0.1 & 0.5-0.4\end{pmatrix}=\begin{pmatrix}0.5 & 0.5\\ 0.1 & 0.1\end{pmatrix},$$

לכן למערכת $(A-\lambda_2 I_2)v_2=0$ יש פתרון (יחיד עד כדי כפל בסקלר) $v_2=\begin{pmatrix}1\\-1\end{pmatrix}$.

(ג) לפי משפט 4.10, v_1,v_2 בלתי תלויה לינארית, לכן זהו בסיס של $\mathbb{R}^2$. לכן $P=(v_1,v_2)=\begin{pmatrix}5 & 1\\1 & -1\end{pmatrix}$ הפיכה. לפי מסקנה 4.9, $P^{-1}AP=\begin{pmatrix}\lambda_1 & 0\\0 & \lambda_2\end{pmatrix}=\begin{pmatrix}1 & 0\\0 & 0.4\end{pmatrix}=D$. כמו כן,

$$P^{-1}=\frac{1}{-6}\begin{pmatrix}-1 & -1\\-1 & 5\end{pmatrix}=\frac{1}{6}\begin{pmatrix}1 & 1\\1 & -5\end{pmatrix}$$

מכאן, לפי (3),

$$A^n=(PDP^{-1})^n=PD^nP^{-1}=\begin{pmatrix}5 & 1\\1 & -1\end{pmatrix}\begin{pmatrix}1 & 0\\0 & (0.4)^n\end{pmatrix}\frac{1}{6}\begin{pmatrix}1 & 1\\1 & -5\end{pmatrix}=$$

$$=\frac{1}{6}\begin{pmatrix}5+(0.4)^n & 5-5(0.4)^n\\1-(0.4)^n & 1+5(0.4)^n\end{pmatrix}\longrightarrow\begin{pmatrix}\frac{5}{6} & \frac{5}{6}\\ \frac{1}{6} & \frac{1}{6}\end{pmatrix}$$

$$\begin{pmatrix}x_\infty\\y_\infty\end{pmatrix}=\begin{pmatrix}\frac{5}{6} & \frac{5}{6}\\ \frac{1}{6} & \frac{1}{6}\end{pmatrix}\begin{pmatrix}x_0\\y_0\end{pmatrix}=\begin{pmatrix}\frac{5}{6}x_0+\frac{5}{6}y_0\\ \frac{1}{6}x_0+\frac{1}{6}y_0\end{pmatrix}=\begin{pmatrix}\frac{5}{6}(x_0+y_0)\\ \frac{1}{6}(x_0+y_0)\end{pmatrix}$$

מסקנה: לאחר זמן רב, שישית מהאוכלוסיה תהיה חולה.

5. אלגוריתם החיפוש של Google - דוגמה לשימוש של אלגברה לינארית

לקוח מתוך

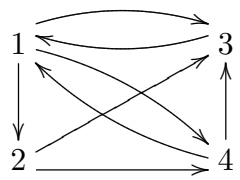
Kurt Bryan, Tanya Leise, *The $25,000,000,000 Eigenvector*, SIAM Review **48** (2006), 569–581.

כאשר מחפשים ביטוי באינטרנט, המחשב מוצא את כל האתרים שמכילים אותו ואז צריך להציג אותם. סוד ההצלחה של מנוע החיפוש של Google הוא בכך שהוא היה הראשון להציג את האתרים לפי *סדר החשיבות* שלהם. אם נניח שחשיבות האתר היא מספר ממשי אי שלילי (גדול יותר, ככל שהאתר חשוב יותר), כיצד לייחס לאתר מספר זה?

חשיבות של אתר תיקבע לפי הַפניות (לינקים) אליו מאתרים אחרים. (כיצד בדיוק – בהמשך.)

נראה את האינטרנט כגרף מכוון: קדקדיו $1,\ldots,n$ הם אתרים. קשתות הם זוגות (i,j) עבור כל הפניה מאתר i לאתר j. חשיבות אתר i תסומן על ידי מספר ממשי $x_i \geq 0$.

דוגמה 5.1:



אתר 1 מפנה לאתרים $2,3,4$;
אתר 2 מפנה לאתרים $3,4$;
אתר 3 מפנה לאתרים 1;
אתר 4 מפנה לאתרים $1,3$.

הגדרה 5.2: ננסה להגדיר x_i. (בכל הגישות מזניחים הַפניות מאתר לעצמו!)

(א) הגדרה פשוטה מאד: x_i= מספר הַפניות אל i. בדוגמה לעיל, $x_1=2, x_2=1, x_3=3, x_4=2$.

חסרון: הַפניות מכל האתרים שוות ערך; רצוי שהפניה מאתר חשוב ל־i תתרום יותר ל־x_i. נשפר:

(ב) $x_i=\sum_{j\in L_i} x_j$, באשר $L_i \subseteq \{1,\ldots,n\}$ קבוצת האתרים המפנים אל i.

חסרונות: אגף ימין תלוי באגף שמאל; אתר משפיע על ידי יצירת הַפניות לאתרים רבים – לא הוגן. נשפר:

(ג) $x_i=\sum_{j\in L_i} x_j/n_j$, באשר n_j מספר הַפניות שיוצאות מאתר j (לרשת כולה). אמנם גם כאן אגף ימין תלוי באגף שמאל, אך זוהי בעצם מערכת של n משוואות ב־n נעלמים שאולי אפשר לפתור. בדוגמה לעיל:

$$\begin{aligned}
x_1 &= x_3/1 + x_4/2\\
x_2 &= x_1/3\\
x_3 &= x_1/3 + x_2/2 + x_4/2\\
x_4 &= x_1/3 + x_2/2
\end{aligned}$$

כלומר, $v=Av$, באשר $v=\begin{pmatrix} x_1\\ x_2\\ x_3\\ x_4\end{pmatrix}$ ו־ $A=\begin{pmatrix} 0 & 0 & 1 & 1/2\\ 1/3 & 0 & 0 & 0\\ 1/3 & 1/2 & 0 & 1/2\\ 1/3 & 1/2 & 0 & 0\end{pmatrix}$.

כלומר, $v\in V_1(A)$ הוא וקטור עצמי של A השייך לערך עצמי $\lambda=1$.

המשוואה $Av=v$ שקולה למערכת $(A-I_4)v=0$, אותה אנחנו יודעים לפתור.

החישוב נותן פתרון $v = (12, 4, 9, 6)^t$ יחיד עד כדי כפל בסקלר.

שים לב שאתר 3 חשוב פחות (למרות שיש אליו יותר הַפניות) מאתר 1. מדוע? (בגלל שהוא מַפנה רק לאתר 1 וזה עושה את 1 לחשוב מאד. זה מתבטא במשוואה הראשונה $x_1 = x_3/1 + x_4/2$.)

נכליל מהדוגמה למקרה הכללי:

הגדרה 5.3: (א) **מטריצת הקישוריות** של הרשת היא מטריצה $A = (a_{ij}) \in \mathrm{M}_n(\mathbb{R})$ המוגדרת על ידי

$$a_{ij} = \frac{\text{מספר ההפניות מאתר } j \text{ לאתר } i}{\text{מספר ההפניות מאתר } j \text{ לשאר האתרים}}. \tag{1}$$

(ב) **וקטור החשיבות** היא עמודה $v = (x_1, \ldots, x_n)^t \in \mathbb{R}^n$, כך ש־$v \neq 0$ וקטור עצמי של A השייך ל־$\lambda = 1$, ו־$x_i \geq 0$ לכל i. נהוג לנרמל אותו כך ש־$\sum_i x_i = 1$.

הערות 5.4: (א) כדי שב־(1) לא נחלק באפס, נניח שמכל אתר יוצאות איזשהן הַפניות.

(ב) כזכור, מזניחים הַפניות מאתר לעצמו. לכן $a_{ii} = 0$ לכל i.

(ג) האם v קיים ויחיד? כלומר: האם $\dim V_1(A) = 1$? ואם כן, האם יש $v \in V_1(A)$ בעל רכיבים אי שליליים?

∎

הגדרה 5.5: מטריצה $A = (a_{ij}) \in \mathrm{M}_n(\mathbb{R})$ נקראת

(א) **סטוכסטית לפי עמודות** [סל"ע] אם $a_{ij} \geq 0$ לכל i, j ו־$\sum_{i=1}^n a_{ij} = 1$ לכל j;

(ב) **סטוכסטית לפי עמודות חיובית** אם $a_{ij} > 0$ לכל i, j ו־$\sum_{i=1}^n a_{ij} = 1$ לכל j. ∎

תרגיל 5.6: *מטריצת הקישוריות היא סטוכסטית לפי עמודות.*

טענה 5.7: *אם $A = (a_{ij}) \in \mathrm{M}_n(\mathbb{R})$ סטוכסטית לפי עמודות, אז $V_1(A) \neq 0$, ובפרט $\dim V_1(A) \geq 1$.*

הוכחה: די להוכיח כי 1 ערך עצמי של A, כלומר, $\det(A - 1 \cdot I) = 0$, כלומר, (למשל) שורות של $A - I$ תלויות לינארית. ואכן, סכום הרכיבים בעמודה ה־j של $A - I$ הוא $\sum_{i=1}^n a_{ij} - 1 = 0$, כלומר, סכום השורות של $A - I$ הוא 0. ∎

האם $\dim V_1(A) \leq 1$? לא בהכרח:

דוגמה 5.8: נניח $A = \mathrm{Diag}(A_1, A_2) = \begin{array}{c c} & \begin{array}{cc} n_1 & n_2 \end{array} \\ \begin{array}{c} n_1 \\ n_2 \end{array} & \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \end{array}$, באשר $A_1 \in \mathrm{M}_{n_1}(\mathbb{R}), A_2 \in \mathrm{M}_{n_2}(\mathbb{R})$ סטוכסטיות לפי עמודות. אז גם A סטוכסטית לפי עמודות. בגלל ש־A_1, A_2 סטוכסטיות לפי עמודות, יש $0 \neq v_1 \in \mathbb{R}^{n_1}$ כך ש־$A_1 v_1 = v_1$ ו־$0 \neq v_2 \in \mathbb{R}^{n_2}$ כך ש־$A_2 v_2 = v_2$. אבל אז $\begin{array}{c} n_1 \\ n_2 \end{array}\begin{pmatrix} v_1 \\ 0 \end{pmatrix}, \begin{array}{c} n_1 \\ n_2 \end{array}\begin{pmatrix} 0 \\ v_2 \end{pmatrix} \in \mathbb{R}^n$

בלתי תלויים לינארית (כי אף אחד מהם אינו כפולה של השני) ושניהם וקטורים עצמיים של A השייכים ל־1. אכן,

$$\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}\begin{pmatrix} v_1 \\ 0 \end{pmatrix} = \begin{pmatrix} v_1 \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}\begin{pmatrix} 0 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ v_2 \end{pmatrix}$$

לכן, במקרה זה, $\dim V_1(A) \geq 2$.

האם יתכן ש־A כזאת תהיה מטריצת קישוריות?

כן, זה קורה כאשר האינטרנט מחולק לשתי רשתות נפרדות שאין הַפניות הדדיות ביניהן. וזה יתכן.

מצד שני, אם A חיובית, זה אינו קורה:

משפט 5.9: *תהי $A = (a_{ij}) \in \mathrm{M}_n(\mathbb{R})$ סטוכסטית לפי עמודות חיובית. אז*

(א) $\dim V_1(A) = 1$.

(ב) *אם $0 \neq v \in V_1(A)$ אז כל רכיביו בעלי אותו סימן (כולם חיוביים או כולם שליליים).*

הוכחה: (ב) יהי $0 \neq v = (x_1, \ldots, x_n)^t \in V_1(A)$. אז יש k כך ש־$x_k \neq 0$. בלי הגבלת הכלליות, $x_k > 0$, אחרת נכפיל v ב־(-1). צריך להוכיח כי $x_i > 0$ לכל i. מתוך $v = Av$ נקבל $x_i = \sum_{j=1}^n a_{ij}x_j$, ולכן

$$|x_i| = |\sum_{j=1}^n a_{ij}x_j| \leq \sum_{j=1}^n |a_{ij}x_j| = \sum_{j=1}^n a_{ij}|x_j|$$

לכל i, ויש כאן שוויון אם ורק אם $x_j \geq 0$ לכל j. אבל

$$\sum_{i=1}^n |x_i| \leq \sum_{i=1}^n\sum_{j=1}^n a_{ij}|x_j| = \sum_{j=1}^n(\sum_{i=1}^n a_{ij})|x_j| = \sum_{j=1}^n |x_j|$$

ולכן במשוואה הקודמת $\leq$ הוא $=$ לכל i, כלומר, $x_j \geq 0$ לכל j. אבל $x_i = \sum_{j=1}^n a_{ij}x_j \geq a_{ik}x_k > 0$.

(א) לפי (ב) די להוכיח:

טענה: *יהיו $u, v \in \mathbb{R}^n$ בלתי תלויים לינארית. אז יש $\alpha, \beta \in \mathbb{R}$ כך שלוקטור $\alpha u + \beta v$ יש רכיבים חיוביים וגם שליליים.*

הוכחה: נסמן $W = \{v \in \mathbb{R}^n |\ \sum_i (v)_i = 0\}$. זהו תת מרחב של $\mathbb{R}^n$. אם $0 \neq w \in W$, אז ל־w יש בהכרח רכיבים חיוביים וגם שליליים. לכן די למצוא $\alpha, \beta \in \mathbb{R}$ כך ש־$0 \neq \alpha u + \beta v \in W$.

אם $u \in W$, נקח $\alpha = 1, \beta = 0$, כי $\alpha u + \beta v = u$.

אחרת יהי β סכום רכיביו של u ו־$-\alpha$ סכום רכיביו של v אז $\beta \neq 0$, לכן $\alpha u + \beta v \neq 0$ ומתקיים

$$\sum_{i=1}^n(\alpha u + \beta v)_i = \sum_{i=1}^n \alpha(u)_i + \beta(v)_i = \alpha\sum_{i=1}^n(u)_i + \beta\sum_{i=1}(v)_i = \alpha\beta + \beta(-\alpha) = 0. \qquad \blacksquare$$

אמנם מטריצת הקישוריות אינה חיובית (כי באלכסון יש לה אפסים) אך שינוי קטן בה הופך אותה לכזאת:

משפט 5.10: *תהי $A = (a_{ij}) \in \mathrm{M}_n(\mathbb{R})$ סטוכסטית לפי עמודות. תהי $B = (b_{ij}) \in \mathrm{M}_n(\mathbb{R})$ כך ש־$b_{ij} = 1/n$ לכל i, j ויהי $0 < m < 1$. אז*

(א) *B סטוכסטית לפי עמודות.*

(ב) *$M = (1-m)A + mB$ סטוכסטית לפי עמודות חיובית.*

הוכחה: (א) ברור.

(ב) $(M)_{ij} = (1-m)a_{ij} + m/n \geq m/n > 0$, ואילו

$$\sum_i (M)_{ij} = \sum_i \big((1-m)a_{ij} + m/n\big) = (1-m)\sum_i a_{ij} + n \cdot m/n = (1-m) + m = 1. \qquad \blacksquare$$

כעת נראה איך אפשר לחשב (באופן מקורב, אבל מהר) את וקטור החשיבות:

עבור וקטור $v \in \mathbb{R}^n$ נסמן $\|v\| = \sum_{i=1}^n |(v)_i|$. אז $\|v\| \geq 0$ ומתקיים $v = 0 \Leftrightarrow \|v\| = 0$.

למה 5.11: *יהי $n \geq 2$ ותהי $A = (a_{ij}) \in \mathrm{M}_n(\mathbb{R})$ מטריצה סטוכסטית לפי עמודות חיובית.*

נסמן $c := \max_j (1 - 2\min_i a_{ij})$. אז

(א) $0 \leq c < 1$.

כמו כן, לכל $v \in W = \{v \in \mathbb{R}^n |\ \sum_i (v)_i = 0\}$ מתקיים

(ב) $Av \in W$;

(ג) $\|Av\| \leq c\,\|v\|$.

הוכחה: (א) נקבע $1 \leq j \leq n$. אז $\sum_{i=1}^n a_{ij} = 1$. אם $a_{kj} = \min_i a_{ij}$, אז

$$0 \leq \sum_{i \neq k} a_{ij} - a_{kj} = 1 - 2a_{kj} = 1 - 2\min_i a_{ij} < 1.$$

מכאן $0 \leq c < 1$.

(ב) $\sum_{i=1}^n (Av)_i = \sum_{i=1}^n \sum_{j=1}^n a_{ij}(v)_j = \sum_{j=1}^n (\sum_{i=1}^n a_{ij})(v)_j = \sum_{j=1}^n (v)_j = 0$.

(ג) נסמן $w = Av$. אם $w = 0$, הטענה ברורה. לכן נניח כי $w \neq 0$. לפי (ב), $\sum_i (w)_i = 0$. לכן יש k כך ש־$(w)_k > 0$ ויש ℓ כך ש־$(w)_\ell < 0$. יהי $\varepsilon_i = \begin{cases} +1 & (w)_i \geq 0 \\ -1 & (w)_i < 0 \end{cases}$. אז $\varepsilon_k = 1$ ו־$\varepsilon_\ell = -1$. כעת,

$$\|w\| = \sum_{i=1}^n |(w)_i| = \sum_{i=1}^n \varepsilon_i \cdot (w)_i = \sum_{i=1}^n \varepsilon_i \sum_{j=1}^n a_{ij}(v)_j = \sum_{j=1}^n \big(\sum_{i=1}^n \varepsilon_i a_{ij}\big)(v)_j. \tag{2}$$

נקבע $1 \leq j \leq n$. אז

$$\sum_i \varepsilon_i a_{ij} = \sum_{i \neq \ell} \varepsilon_i a_{ij} - a_{\ell j} \leq \sum_{i \neq \ell} a_{ij} - a_{\ell j} = 1 - 2a_{\ell j} \leq 1 - 2\min_i a_{ij} \leq c,$$

ובאופן דומה

$$\sum_i \varepsilon_i a_{ij} = \sum_{i\neq k} \varepsilon_i a_{ij} + a_{kj} \geq \sum_{i\neq k} -a_{ij} + a_{kj} = -1 + 2a_{kj} \geq -1 + 2\min_i a_{ij} \geq -c .$$

משתי המשוואות האלה נובע ש־ $|\sum_i \varepsilon_i a_{ij}| \leq c$. מכאן, לפי (2),

$$\|w\| \leq \sum_{j=1}^{n} \Big|\Big(\sum_{i=1}^{n} \varepsilon_i a_{ij}\Big)(v)_j\Big| = \sum_{j=1}^{n} \Big|\sum_{i=1}^{n} \varepsilon_i a_{ij}\Big| \cdot |(v)_j| \leq \sum_{j=1}^{n} c \cdot |(v)_j| = \\ = c \sum_j |(v)_j| = c\,\|v\| . \qquad \blacksquare$$

משפט 5.12: *יהי $n \geq 2$ ותהי $A = (a_{ij}) \in \mathrm{M}_n(\mathbb{R})$ מטריצה סטוכסטית לפי עמודות חיובית. יהי $v \in \mathbb{R}^n$ כך ש־ $Av = v$ ו־ $\sum_i (v)_i = 1$. יהי $v_0 \in \mathbb{R}^n$ כך ש־ $\sum_i (v_0)_i = 1$. אז $v = \lim_{k\to\infty} A^k v_0$, כלומר, $\lim_{k\to\infty} \|A^k v_0 - v\| = 0$.*

הוכחה: יהי $w = v_0 - v$. אז $\sum_i (w)_i = \sum_i (v_0)_i - \sum_i (v)_i = 1 - 1 = 0$, לכן $w \in W$. כעת, $A^k v_0 = A^k(w+v) = A^k w + v$ ולכן $A^k v_0 - v = A^k w$. לפי למה 5.11, $\|A^k w\| \leq c^k \|w\|$, באשר $0 \leq c < 1$, לכן $A^k w \to 0$. מכאן $A^k v_0 \to v$. $\blacksquare$

הערה 5.13: משפט זה הוא מקרה פרטי של "שיטת החזקות" של Richard von Mises למציאת וקטורים עצמיים וערכים עצמיים (מסוימים) של מטריצות ריבועיות כלשהן. $\blacksquare$

6. פולינום אופייני

יהי F שדה.

עבור מטריצה $C = \big(g_{ij}(X)\big) \in \mathrm{M}_n(F[X])$ ועבור $\lambda \in F$ נגדיר $C(\lambda) = (g_{ij}(\lambda)) \in \mathrm{M}_n(F)$.

למה 6.1: *תהי* $C = (g_{ij}(X)) \in \mathrm{M}_n(F[X])$. *אזי*

(א) $\det(C) \in F[X]$;

(ב) $\deg(\det(C)) \le d_1 + \cdots + d_n$, *באשר* $d_i = \max_{1\le j\le n} \deg g_{ij}$;

(ג) *לכל* $\lambda \in F$: $(\det C)(\lambda) = \det(C(\lambda))$.

הוכחה: (למדנו - אמנם ללא הוכחה - שיש שדה L כך ש־$F[X] \subseteq L$ תת חוג שלו. לכן $(g_{ij}(X)) \in \mathrm{M}_n(L)$, ולכן ל־ $\det(g_{ij}(X))$ יש משמעות, וזהו איבר של L. אך עדיין לא ברור מדוע זהו איבר של $F[X]$!)

(א) + (ב) + (ג): באינדוקציה על n.

עבור $n = 1$, $C = (g)$, לכן $\det C = g \in F[X]$ ו־$g(\lambda) = \det(g(\lambda))$, כך שהכל ברור.

עבור $n > 1$ נעשה פיתוח לפי השורה האחרונה

$$\det(C) = \det(g_{ij}(X)) = \sum_{j=1}^{n} (-1)^{n+j} g_{nj} \det C_{nj}, \tag{1}$$

באשר $C_{nj} \in \mathrm{M}_{n-1}(F[X])$ המינור של המקום ה־(n,j) של C. לפי הנחת האינדוקציה

(א') $\det C_{nj} \in F[X]$,

(ב') $\deg \det C_{nj} \le d_1 + \cdots + d_{n-1}$;

(ג') $\big(\det C_{nj}\big)(\lambda) = \det\big(C_{nj}(\lambda)\big)$ לכל i, j.

מכאן, לפי (1):

(א) $\det(C)$ הוא סכום של מכפלות של פולינומים מעל F, לכן פולינום מעל F;

(ב) $\deg(\det(C)) \le \max_j(\deg g_{n_j} + \deg \det C_{nj}) \le d_n + (d_1 + \cdots + d_{n-1})$;

(ג) $(\det C)(\lambda) = \sum_{j=1}^{n} (-1)^{n+j} g_{nj}(\lambda)\big(\det C_{nj}\big)(\lambda) =$.

$= \sum_{j=1}^{n} (-1)^{n+j} g_{nj}(\lambda) \det\big(C_{nj}(\lambda)\big) = \det(C(\lambda))$ ∎

הוכחה נוספת:

$$\det(C) = \sum_{\sigma \in S_n} \mathrm{Sgn}(\sigma) \prod_{i=1}^{n} g_{i\sigma(i)}(X)$$

הוא סכום של מכפלות של פולינומים מעל F, לכן פולינום מעל F. מכאן (א). לכל $\sigma \in S_n$ מתקיים

$$\deg \mathrm{Sgn}(\sigma) \prod_{i=1}^{n} g_{i\sigma(i)}(X) = \sum_{i=1}^{n} \deg g_{i\sigma(i)}(X) \le d_1 + \cdots + d_n,$$

לכן, לפי המשוואה הקודמת, נובע (ב). כמו כן, לפי אותה המשוואה, (בגלל שהצבה היא הומומורפיזם)

$$(\det C)(\lambda) = \sum_{\sigma\in S_n} \operatorname{Sgn}(\sigma)\prod_{i=1}^{n} g_{i\sigma(i)}(\lambda) = \det(C(\lambda)). \qquad \blacksquare$$

הגדרה 6.2: תהי $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \in \mathrm{M}_n(F)$. אזי המטריצה

$$XI_n - A = \begin{pmatrix} X & 0 & \cdots & 0 \\ 0 & X & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & X \end{pmatrix} - A = \begin{pmatrix} X-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & X-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & X-a_{nn} \end{pmatrix}$$

ב־$\mathrm{M}_n(F[X])$ תקרא **המטריצה האופיינית** של A, ו־

$$f_A(X) := \det(XI_n - A) \in F[X]$$

ייקרא **הפולינום האופייני** של A.

דוגמה 6.3: אם $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ אז $XI_2 - A = \begin{pmatrix} X-1 & -2 \\ -3 & X-4 \end{pmatrix}$ ולכן
$f_A(X) = (X-1)(X-4) - 2\cdot 3 = X^2 - 5X - 2$. $\blacksquare$

למה 6.4: *תהי* $A = (a_{ij}) \in \mathrm{M}_n(F)$. *אז* $f_A(X) = c_nX^n + c_{n-1}X^{n-1} + \cdots + c_0$, *באשר*

(1) $c_n = 1$; *כלומר,* $f_A(X)$ *מתוקן ממעלה* n;

(2) $c_{n-1} = -\operatorname{tr} A = -(a_{11} + a_{22} + \cdots + a_{nn})$;

(3) $c_0 = (-1)^n \det A$.

הוכחה: (3): לפי למה 6.1(ג), $c_0 = f_A(0) = \det(0I_n - A) = \det(-A) = (-1)^n \det A$.

(2) + (1): באינדוקציה על n.

עבור $n = 1$, $f_A(X) = X - a_{11}$. במקרה זה (1), (2) ברורים.

עבור $n > 1$ נעשה פיתוח של

$$B := XI_n - A = \begin{pmatrix} X-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & X-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & X-a_{nn} \end{pmatrix}$$

לפי השורה האחרונה:

$$f_A(X) = \det(XI_n - A) = (-1)^{n+1}(-a_{n1})\det B_{n1} + (-1)^{n+2}(-a_{n2})\det B_{n2} + \cdots$$
$$+(-1)^{n+n-1}(-a_{n\,n-1})\det B_{n\,n-1} + (-1)^{n+n}(X - a_{nn})\det B_{nn},$$

באשר $B_{ij} \in \mathrm{M}_{n-1}(F[X])$ מתקבלת מ־B על ידי מחיקת השורה ה־i והעמודה ה־j. נשים לב:

(א) $B_{nn} = XI_{n-1} - A_{nn}$, באשר A_{nn} מתקבלת מ־A על ידי מחיקת השורה ה־n והעמודה ה־n, ולכן לפי הנחת האינדוקציה

$$\det B_{nn} = X^{n-1} - (a_{11} + a_{22} + \cdots + a_{n-1\,n-1})X^{n-2} + \cdots (\text{פולינום ממעלה } < n-2).$$

(ב) ב־B_{nj}, עבור $j < n$, יש בכל שורה, פרט לשורה ה־j, בדיוק רכיב אחד ממעלה 1, וכל היתר ממעלה ≥ 0; בשורה ה־j כל הרכיבים הם ממעלה ≥ 0. לכן לפי למה 6.1(ב), $\deg\det B_{nj} \leq n-2$. מכאן

$$f_A(X) = (X - a_{nn})\big(X^{n-1} - (a_{11} + \cdots + a_{n-1\,n-1})X^{n-2} + \cdots (\text{פולינום ממעלה } < n-2)\big)$$
$$+ (\text{$n-1$ פולינומים ממעלה } \geq n-2)$$
$$= X^n - (a_{11} + a_{22} + \cdots + a_{n-1\,n-1} + a_{nn})X^{n-1} + \cdots (\text{פולינום ממעלה } \geq n-2) \qquad \blacksquare$$

הוכחה נוספת: תהי $B = XI_n - A$. אז

$$f_A(X) = \sum_{\sigma \in S_n} \mathrm{Sgn}(\sigma) \prod_{i=1}^{n} (B)_{i\sigma(i)} \qquad (2)$$

באשר, $(B)_{ij} = \begin{cases} X - a_{ij} & i = j \\ -a_{ij} & i \neq j \end{cases}$.

כיוון ש־$(B)_{ij} \in F[X]$ לכל i, j, ברור ש־$f_A(X) \in F[X]$.

(2) + (1): אם $\sigma = 1$, אז $\mathrm{Sgn}(\sigma)\prod_{i=1}^{n}(B)_{i\sigma(i)} = 1 \cdot (X - a_{11}) \cdots (X - a_{nn})$. אם $\sigma \neq 1$ אז יש לפחות שני אינדקסים שונים i עבורם $\sigma(i) \neq i$, ולכן לפחות שנים מהגורמים ב־$\prod_{i=1}^{n}(B)_{i\sigma(i)}$ ממעלה ≥ 0, וכל השאר ממעלה ≥ 1. לכן $\deg\prod_{i=1}^{n}(B)_{i\sigma(i)} \leq n-2$. מכאן

$$f_A(X) = (X - a_{11}) \cdots (X - a_{nn}) + \cdots + (\text{פולינום ממעלה } \geq n-2)$$
$$= X^n - (a_{11} + a_{22} + \cdots + a_{nn})X^{n-1} + \cdots (\text{פולינום ממעלה } \geq n-2)$$

(3): נציב $X = 0$ ב־(2):

$$c_0 = f_A(0) = \sum_{\sigma \in S_n} \mathrm{Sgn}(\sigma) \prod_{i=1}^{n} (B(0))_{i\sigma(i)} = \sum_{\sigma \in S_n} \mathrm{Sgn}(\sigma) \prod_{i=1}^{n} \big(-a_{i\sigma(i)}\big) =$$
$$(-1)^n \sum_{\sigma \in S_n} \mathrm{Sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} = (-1)^n \det A \qquad \blacksquare$$

משפט 6.5: *תהי $A \in \mathrm{M}_n(F)$, ויהי $\lambda \in F$. אז λ ערך עצמי של A $\Leftrightarrow f_A(\lambda) = 0$.*

הוכחה: לפי משפט 4.5, λ ערך עצמי $\Leftrightarrow \det(\lambda I_n - A) = 0$. לפי למה 6.1(ג) (לחילופין, על ידי הצבת λ ב־(2)), $f_A(\lambda) = \det(\lambda I_n - A)$. $\blacksquare$

מסקנה 6.6: *אם $f_A(X) = (X-\lambda_1)\cdots(X-\lambda_n)$, באשר $\lambda_1,\dots,\lambda_n \in F$ שונים זה מזה, אז A דומה למטריצה אלכסונית.*

הוכחה : לכל $1 \le i \le n$, λ_i הוא שרש של f_A ולכן ערך עצמי של A. לכן יש $0 \neq v_i \in V_{\lambda_i}$. לפי משפט 4.10, $v_1,\dots,v_n$ בלתי תלויים לינארית, ולכן בסיס של F^n. לפי מסקנה 4.9, A דומה לאלכסונית. ∎

משפט 6.7: *למטריצות דומות אותו פולינום אופייני.*

הוכחה: צריך להוכיח: אם $A, P \in \mathrm{M}_n(F)$, ו־P הפיכה אז $f_{P^{-1}AP} = f_A$.

ואכן,

$$P^{-1}(XI_n - A)P = XP^{-1}I_nP - P^{-1}AP = XI_n - P^{-1}AP$$

(כאשר שוויונות אלה יש להבין כשוויונות של מטריצות מעל שדה L אשר מכיל את $F[X]$). מכאן

$$f_{P^{-1}AP}(X) = \det(XI_n - P^{-1}AP) = \det P^{-1}\det(XI_n - A)\det P$$

$$= \det P^{-1}\det P\det(XI_n - A) = \det I_n\ f_A(X) = f_A(X)$$ ∎

מסקנה: *למטריצות דומות אותה דטרמיננטה ואותה עקבה.*

הגדרה 6.8: הצבת מטריצה בפולינום. אם $g = \sum_{i=0}^{\infty} a_iX^i$ ו־$A \in \mathrm{M}_n(F)$, נגדיר

$$g(A) := \sum_{i=0}^{\infty} a_iA^i \in \mathrm{M}_n(F).$$

אפשר להבין הגדרה זו כהצבה: F תת חוג של $\mathrm{M}_n(F)$ (כאשר מזהים סקלר α עם המטריצה הסקלרית αI_n) ולכן $g = \sum_{i=0}^{\infty} a_iI_nX^i \in \mathrm{M}_n(F)[X]$. כעת אפשר להציב בפולינום זה את $A \in \mathrm{M}_n(F)$. התוצאה של ההצבה היא $\sum_{i=0}^{\infty}(a_iI_n)A^i = \sum_{i=0}^{\infty} a_iA^i$. שים לב שמקדמי g (מטריצות סקלריות) מתחלפים עם A.

מתקיים (למשל בגלל תכונות ההצבה –טענה 1.6(3)): אם $f, g \in F[X]$ אז

$$(f+g)(A) = f(A) + g(A) \qquad ,(fg)(A) = f(A)g(A).$$

בפרט לכל $c \in F$ מתקיים: $(cg)(A) = c(A)g(A) = cg(A)$.

הוכחה ישירה של $(fg)(A) = f(A)g(A)$:

נניח $f = \sum_i a_iX^i,\ g = \sum_j b_jX^j$, כך ש־$fg = \sum_k(\sum_{i+j=k} a_ib_j)X^k$, אז

$$f(A)g(A) = \Big(\sum_i a_iA^i\Big)\Big(\sum_j b_jA^j\Big) = \sum_i a_iA^i\Big(\sum_j b_jA^j\Big) = \sum_i\sum_j(a_iA^i)(b_jA^j) =$$

$$\sum_i\sum_j a_ib_jA^{i+j} = \sum_k\Big(\sum_{i+j=k} a_ib_j\Big)A^k = (fg)(A)$$

באופן דומה מוכיחים $(f+g)(A) = f(A) + g(A)$.

משפט 6.9 (Cayley-Hamilton): *תהי* $A \in \mathrm{M}_n(F)$. *אז* $f_A(A) = 0$.

הוכחה: נתבונן ב־ $B := XI_n - A \in \mathrm{M}_n(F[X])$. אז גם המינורים של B הם מטריצות מעל $F[X]$, ולכן, לפי למה 6.1(א), גם $\mathrm{adj}(B) \in \mathrm{M}_n(F[X])$. כידוע,

$$\mathrm{adj}(B)B = \det(B)I_n = f_A(X)I_n. \tag{5}$$

לפי טענה 1.8 ההעתקה $\varphi\colon \mathrm{M}_n(F)[X] \to \mathrm{M}_n(F[X])$ הנתונה על ידי המשוואה $\big(\varphi(\sum_{\mu=0}^{\infty} A_\mu X^\mu)\big)_{ij} = \sum_{\mu=0}^{\infty}(A_\mu)_{ij}X^\mu$ היא איזומורפיזם של חוגים. קל לראות (תרגיל 1.9) כי

$$f_A(X)I_n = \varphi\big(f_A(X)\big), \qquad B = XI_n - A = \varphi(X - A) = \varphi(I_nX - AX^0),$$

כאשר $f_A(X) \in F[X] \subseteq \mathrm{M}_n(F)[X]$. כיוון ש־$\varphi$ על, יש פולינום $q(X) \in \mathrm{M}_n(F)[X]$ כך ש־ $\mathrm{adj}(B) = \varphi\big(q(X)\big)$. את (5) אפשר לכתוב כך:

$$\varphi\big(q(X)(X - A)\big) = \varphi\big(q(X)\big)\varphi(X - A) = \varphi\big(f_A(X)\big).$$

לכן, כיוון ש־φ חד חד ערכית,

$$q(X)(X - A) = f_A(X). \tag{6}$$

מכאן, לפי מסקנה 2.4 עבור החוג $R = \mathrm{M}_n(F)$, מתקיים $f_A(A) = 0$. ∎

מדוע ההוכחה הבאה של משפט קיילי־המילטון איננה נכונה?

הוכחה לא נכונה: לפי ההגדרה, $f_A = \det(XI_n - A)$. נציב A במקום X. אז $f_A(A) = \det(AI_n - A) = \det(0) = 0$. ∎

מסקנה 6.10: *תהי* $A \in \mathrm{M}_n(F)$. *קיים פולינום* $0 \neq f \in F[X]$ *ממעלה* $\geq n$ *כך ש־* $f(A) = 0$.

תרגיל 6.11: *תהי* $A \in \mathrm{M}_n(F)$. *הראה, מבלי להסתמך על משפט* Cayley-Hamilton, *שיש* $0 \neq g \in F[X]$ *כך ש־* $g(A) = 0$.

רמז: $I_n, A, \ldots, A^{n^2} \in \mathrm{M}_n(F)$ תלויות לינארית מעל F. ∎

תרגיל 6.12: *תהי*

$$M = \begin{pmatrix} A_1 & B & \cdots & C \\ 0 & A_2 & \cdots & D \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & A_r \end{pmatrix} \tag{7}$$

מטריצת גושים מעל שדה F, באשר $A_1, \ldots, A_r$ מטריצות ריבועיות.

(א) *הוכיחו ש־* $\det(M) = \det(A_1) \cdots \det(A_r)$.

(ב) *הוכיחו ש־* $f_M = f_{A_1} \cdots f_{A_r}$.

(ג) *אם M מטריצת גושים באלכסון (כלומר, $B = \cdots = C = \cdots = D = 0$), הוכיחו ש־* $\mathrm{rank} M = \sum_i \mathrm{rank} A_i$.

פתרון: (א) כזכור, אם נעשה פעולה אלמנטרית על השורות של M, המטריצה המתקבלת M' מקיימת $\det(M') = \alpha \det(M)$, באשר $\alpha \in F^\times$ תלוי בסוג הפעולה

(אם הפעולה היא החלפת שתי שורות, אז $\alpha = -1$; אם הפעולה היא הוספה של כפולה של שורה אחת לאחרת, אז $\alpha = 1$; ואם הפעולה היא הכפלת שורה ב־$\beta \in F^\times$, אז $\alpha = \beta$).

אבל אם הפעולה היא כזו שהשורות שמעורבות בה הן רק כאלה שמכילות, עבור איזה $1 \leq i \leq r$, את השורות של A_i (כלומר, את A_i ואת הגושים שמימין לה וגושי האפסים שמשמאל לה), אז

$$M' = \begin{pmatrix} A_1' & B' & \cdots & C' \\ 0 & A_2' & \cdots & D' \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & A_r' \end{pmatrix}$$

באשר $A_j' = A_j$ (ובפרט $\det(A_j') = \det(A_j)$) לכל $j \neq i$ ו־$\det(A_i') = \alpha \det(A_i)$.

לכן, אם נוכיח ש־$\det(M') = \det(A_1') \cdots \det(A_r')$, אז על ידי הכפלת שני האגפים ב־α^{-1} נקבל את (א). לכן בלי הגבלת הכלליות $A_1, \ldots, A_r$ הן מטריצות מדורגות קנונית. בפרט, הן משולשיות עליונות. לכן גם M משולשית עליונה. במקרה זה $\det(M)$ היא מכפלת האיברים באלכסון הראשי של M, $\det(A_i)$ היא מכפלת האיברים באלכסון הראשי של A_i, לכל i, והאלכסון הראשי של M מורכב מהאלכסונים הראשיים של $A_1, \ldots, A_r$. לכן $\det(M) = \det(A_1) \cdots \det(A_r)$.

(ב)

$$XI - M = \begin{pmatrix} XI - A_1 & -B & \cdots & -C \\ 0 & XI - A_2 & \cdots & -D \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & XI - A_r \end{pmatrix}$$

לכן, לפי (א), $\det(XI - M) = \det(XI - A_1) \cdots \det(XI - A_r)$, ומכאן

$$f_M = f_{A_1} \cdots f_{A_r}.$$

(ג) נניח ש־ $M = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & A_r \end{pmatrix}$. אם נעשה על M פעולה אלמנטרית מהסוג המתואר

בחלק (א), נקבל $M' = \begin{pmatrix} A_1' & 0 & \cdots & 0 \\ 0 & A_2' & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & A_r' \end{pmatrix}$, באשר $A_j' = A_j$ לכל $j \neq i$, $\mathrm{rank}A_i' = \mathrm{rank}A_i$, ו־$\mathrm{rank}M' = \mathrm{rank}M$. לכן די להוכיח את הטענה עבור M'. לכן, בלי הגבלת הכלליות, $A_1, \ldots, A_r$ מדורגות קנונית.

אז $\mathrm{rank}A_i$ הוא מספר השורות השונות מאפס ב־A_i, לכל $1 \leq i \leq r$, לכן, $\sum_i \mathrm{rank}A_i$ הוא מספר השורות השונות מאפס ב־M. אבל קל לראות שאם נעביר את שורות האפסים של M לסוף, נקבל מטריצה מדורגת קנונית, לכן מספר השורות השונות מאפס ב־M הוא $\mathrm{rank}M$. מכאן (ג). ∎

בהמשך יהי V מרחב וקטורי ממימד סופי n מעל שדה F ותהי $T\colon V \to V$ העתקה לינארית.

הגדרה 6.13: **הפולינום האופייני** $f_T \in F[X]$ **של** T הוא הפולינום האופייני של המטריצה $[T]_{\mathcal{B}}^{\mathcal{B}}$, באשר $\mathcal{B}$ בסיס סדור של V. (בפרט $\deg f_T = \dim V$.)

הגדרה זו אינה תלויה בבסיס $\mathcal{B}$: אם $\mathcal{A}$ בסיס סדור אחר של V, אז $[T]_{\mathcal{A}}^{\mathcal{A}} = P^{-1}[T]_{\mathcal{B}}^{\mathcal{B}}P$, באשר $P \in \mathrm{M}_n(F)$ הפיכה (היא מטריצת המעבר מ־$\mathcal{B}$ ל־$\mathcal{A}$). לכן $[T]_{\mathcal{B}}^{\mathcal{B}}, [T]_{\mathcal{A}}^{\mathcal{A}}$ דומות; לפי משפט 6.7 יש להן אותו פולינום אופייני.

הגדרה 6.14: עבור כל $g(X) = \sum_{i=0}^{\infty} a_i X^i \in F[X]$ מגדירים העתקה לינארית

$$g(T) = \sum_{i=0}^{\infty} a_i T^i \colon V \to V$$

(באשר T^0 היא הזהות של V ו־$T^i = \overbrace{T \circ T \circ \cdots \circ T}^{i}$).

בדומה להגדרה 6.8 (מחליפים $\mathrm{M}_n(F)$ ב־$L_F(V)$) מתקיים עבור $f, g \in F[X]$

$$(f+g)(T) = f(T) + g(T) \quad (fg)(T) = f(T)g(T) = f(T) \circ g(T).$$

בפרט $f(T) \circ g(T) = g(T) \circ f(T)$, כי $fg = gf$. ובפרט $f(T) \circ T = T \circ f(T)$.

טענה 6.15: *יהי $\mathcal{B}$ בסיס סדור של V. אז* $[g(T)]_{\mathcal{B}}^{\mathcal{B}} = g([T]_{\mathcal{B}}^{\mathcal{B}})$

הוכחה: ההעתקה $S \mapsto [S]_{\mathcal{B}}^{\mathcal{B}}$ מתוך $L(V,V)$ לתוך $\mathrm{M}_n(F)$ שומרת חיבור, כפל וכפל בסקלר, לכן עבור $g = \sum_{i=0}^{\infty} a_i X^i$ מתקיים

$$[g(T)]_{\mathcal{B}}^{\mathcal{B}} = [\sum_{i=0}^{\infty} a_i T^i]_{\mathcal{B}}^{\mathcal{B}} = \sum_{i=0}^{\infty} a_i [T^i]_{\mathcal{B}}^{\mathcal{B}} = \sum_{i=0}^{\infty} a_i ([T]_{\mathcal{B}}^{\mathcal{B}})^i = g([T]_{\mathcal{B}}^{\mathcal{B}}).$$

∎

משפט 6.16 (משפט Cayley-Hamilton להעתקות): $f_T(T) = 0$.

הוכחה: יהי $\mathcal{B}$ בסיס סדור של V ותהי $C = [T]_{\mathcal{B}}^{\mathcal{B}}$. לפי ההגדרה, $f_T = f_C$. לפי טענה 6.15, $[f_T(T)]_{\mathcal{B}}^{\mathcal{B}} = f_T(C) = f_C(C)$ ולפי משפט קיילי־המילטון למטריצות (משפט 6.9) $f_C(C) = 0$. לכן $[f_T(T)]_{\mathcal{B}}^{\mathcal{B}} = 0$ ומכאן $f_T(T) = 0$, בגלל שההעתקה $T \mapsto [T]_{\mathcal{B}}^{\mathcal{B}}$ חד חד ערכית. ∎

7. פולינום מזערי

יהי F שדה.

אם $A \in \mathrm{M}_n(F)$. אז $I = \{f \in F[X] |\ f(A) = 0\}$ הוא אידאל (בדקו!). לפי מסקנה 6.10 (או תרגיל 6.11) $I \neq \{0\}$. לפי משפט 3.3, I אידאל ראשי ויש לו יוצר מתוקן יחיד, הוא הפולינום המתוקן היחיד ב־I מהמעלה $\min(\deg h |\ h \in I \setminus \{0\})$.

הגדרה 7.1: תהי $A \in \mathrm{M}_n(F)$. פולינום מתוקן $m_A(X) \in F[X]$ ייקרא **הפולינום המזערי של** A אם $(m_A) = \{f \in F[X] |\ f(A) = 0\}$. במלים אחרות, m_A מתוקן, מקיים $m_A(A) = 0$, והוא בעל מעלה מזערית מבין כל הפולינומים ב־$F[X]$ המאפסים את A והשונים מאפס.

מההגדרה נובע מיידית:

משפט 7.2: *לכל $f \in F[X]$ מתקיים: $m_A|f \Leftrightarrow f(A) = 0$. בפרט, $m_A|f_A$.*

הערה 7.3: $\deg m_A \geq 1$. ∎

דוגמה 7.4: $m_A = X - \lambda$ אם ורק אם $A = \lambda I_n$. ∎

משפט 7.5: *תהי $A \in \mathrm{M}_n(F)$. אז $f_A|m_A^n$ בחוג $F[X]$.*

הוכחה: $m_A(X) \in F[X] \subseteq \mathrm{M}_n(F)[X]$ ו־$m_A(A) = 0$. לפי מסקנה 2.4 יש $q(X) \in \mathrm{M}_n(F)[X]$ כך ש־

$$m_A(X) = q(X)(X - A). \tag{7}$$

יהי $\varphi\colon \mathrm{M}_n(F)[X] \to \mathrm{M}_n(F[X])$ האיזומורפיזם שהוגדר קודם. לפי תרגיל 1.12, $\varphi(X - A) = XI_n - A$, $\varphi\big(m_A(X)\big) = m_A(X)I_n$. תהי $C = \varphi\big(q(X)\big)$. אזי לפי (7)

$$m_A(X)I_n = C(XI_n - A).$$

ניקח דטרמיננטה משני האגפים ונקבל

$$m_A^n(X) = \det(C) \cdot f_A(X).$$

היות ולפי למה 6.1(א), $\det(C) \in F[X]$, מכאן המסקנה. ∎

משפט 7.6: *תהי $A \in \mathrm{M}_n(F)$. אז ל־f_A ו־m_A אותם גורמים אי פריקים מתוקנים ב־$F[X]$. בפרט, אם*

$$f_A = q_1^{k_1} q_2^{k_2} \cdots q_r^{k_r},$$

באשר $q_1, \ldots, q_r \in F[X]$ אי פריקים מתוקנים שונים זה מזה, $k_1, \ldots, k_r \geq 1$, אז

$$m_A = q_1^{\ell_1} q_2^{\ell_2} \cdots q_r^{\ell_r},$$

באשר $1 \le \ell_i \le k_i$ לכל $1 \le i \le r$.

הוכחה: יהי $q \in F[X]$ אי פריק מתוקן. אם $q|m_A$ אז $q|f_A$ כי $m_A|f_A$, לפי משפט 7.2. אם $q|f_A$ אז $q|m_A^n$, כי $f_A|m_A^n$, לפי משפט 7.5. מכאן (היות ו־q ראשוני) $q|m_A$. אם כן, הוכחנו: $q|f_A \Leftrightarrow q|m_A$.

הטענה שניה נובעת מהראשונה ומכך ש־$m_A|f_A$, לפי קריטריון ההתחלקות של תרגיל 3.20. ∎

מסקנה 7.7: *תהי $A \in \mathrm{M}_n(F)$ ויהי $\lambda \in F$. אז λ ערך עצמי של A אם ורק אם $m_A(\lambda) = 0$.*

הוכחה: בסימונים של המשפט הקודם,

$$\lambda \text{ ע"ע של } A \Leftrightarrow f_A(\lambda) = 0 \Leftrightarrow \text{יש } 1 \le i \le r \text{ כך ש־} q_i(\lambda) = 0 \Leftrightarrow m_A(\lambda) = 0.$$ ∎

תרגיל 7.8: *תהי $A \in \mathrm{M}_3(\mathbb{R})$. הוכיחו: $A^2 \ne -I_3$.*

הוכחה: יהי $f(X) = X^2 + 1$. נניח בשלילה $f(A) = 0$. אז $m_A|f$. אבל f אי פריק, מתוקן, לכן $m_A = f$. מכאן של־m_A אין שרשים ממשיים, לכן ל־A אין ערכים עצמיים ממשיים או, במילים אחרות, ל־f_A אין שרשים ממשיים. אבל $f_A \in \mathbb{R}[X]$ מתוקן ממעלה 3, לכן יש לו שורש ממשי. סתירה. ∎

תרגיל 7.9: *תהיינה $A_1, \dots, A_r$ מטריצות ריבועיות ויהי $g \in F[X]$.*

(א) *אם* $M = \begin{pmatrix} A_1 & B & \cdots & C \\ 0 & A_2 & \cdots & D \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & A_r \end{pmatrix}$, *אז* $g(M) = \begin{pmatrix} g(A_1) & B' & \cdots & C' \\ 0 & g(A_2) & \cdots & D' \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & g(A_r) \end{pmatrix}$

עבור מטריצות $B', \dots, C', \dots, D', \dots$ מתאימות.

(ב) *אם* $M = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & A_r \end{pmatrix}$, *אז* $g(M) = \begin{pmatrix} g(A_1) & 0 & \cdots & 0 \\ 0 & g(A_2) & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & g(A_r) \end{pmatrix}$.

הוכחה: תחילה נוכיח את שני החלקים עבור $g = X^k$, באינדוקציה על k. עבור $k = 0$ המטריצות $g(M), g(A_1), \dots, g(A_r)$ הן מטריצות יחידה (מסדרים מתאימים), ומכאן המסקנה. עבור $k = 1$ מתקיים $g(M) = M$ ו־$g(A_i) = A_i$ לכל i, ומכאן המסקנה.

כעת, אם הטענה נכונה עבור $g = X^k$ אז היא נכונה עבור $h = X^{k+1} = g \cdot X$, כי במקרה (א)

$$h(M) = g(M)M = \begin{pmatrix} g(A_1)A_1 & B'' & \cdots & C'' \\ 0 & g(A_2)A_2 & \cdots & D'' \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & g(A_r)A_r \end{pmatrix}$$

עבור מטריצות $B'', \dots, C'', \dots, D'', \dots$ מתאימות. באופן דומה במקרה (ב), שם $B'', \dots, C'', \dots, D''$ הן מטריצות אפסים.

במקרה הכללי $g=\sum_k a_kX^k$, ולכן ב(א) $g(M)=\sum_k a_kM^k$ היא

$$\sum_k a_k\begin{pmatrix} A_1^k & B_k & \cdots & C_k \\ 0 & A_2^k & \cdots & D_k \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & A_r^k \end{pmatrix}=\begin{pmatrix} \sum_k a_kA_1^k & \sum_k a_kB_k & \cdots & \sum_k a_kC_k \\ 0 & \sum_k a_kA_2^k & \cdots & \sum_k a_kD_k \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \sum_k a_kA_r^k \end{pmatrix}=$$
$$=\begin{pmatrix} g(A_1) & B' & \cdots & C' \\ 0 & g(A_2) & \cdots & D' \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & g(A_r) \end{pmatrix}$$

עבור $B',\dots,C',\dots,D',\dots,B_k,\dots,C_k,\dots,D_k,\dots$ מתאימות. זוהי המסקנה המבוקשת. במקרה (ב) המטריצות האלה הן מטריצות האפסים. זוהי המסקנה המבוקשת. ∎

משפט 7.10: *תהי M כמו בתרגיל 7.9(א) ויהי $h=\operatorname{lcm}(m_{A_1},\dots,m_{A_r})$. אז*

(א) $h|m_M$

(ב) *אם M מטריצת גושים באלכסון –כמו בתרגיל 7.9(ב)– אז $h=m_M$.*

הוכחה: (א) כיוון ש־$m_M(M)=0$, לפי תרגיל 7.9, $m_M(A_i)=0$ לכל i. לכן $m_{A_1},\dots,m_{A_r}|m_M$ ומכאן $h|m_M$.

(ב) לכל i מתקיים $m_{A_i}|h$ ולכן $h(A_i)=0$. לכן לפי תרגיל 7.9, $h(M)=0$. לכן $m_M|h$. לפי (א) $h|m_M$. כיון ש־m_M,h מתוקנים, מכאן נובע ש־$m_M=h$. ∎

משפט 7.11: *למטריצות דומות אותו פולינום מזערי.*

הוכחה: תהיינה $A,P\in \mathrm{M}_n(F)$, באשר P הפיכה. אז

$$(P^{-1}AP)^i=(P^{-1}AP)(P^{-1}AP)\cdots(P^{-1}AP)=P^{-1}A^iP$$

לכן לכל $g=\sum_{i=0}^{\infty}a_iX^i\in F[X]$ מתקיים

$$g(P^{-1}AP)=\sum_{i=0}^{\infty}a_iP^{-1}A^iP=P^{-1}(\sum_{i=0}^{\infty}a_iA^i)P=P^{-1}g(A)P$$

לכן $g(P^{-1}AP)=0\Leftrightarrow P^{-1}g(A)P=0\Leftrightarrow g(A)=0$. מכאן המסקנה. ∎

תרגיל 7.12: *תהי $A\in \mathrm{M}_n(F)$ בעלת רכיבים שונים מאפס באלכסון הראשון שמתחת לאלכסון הראשי ואפסים מתחתיהם. אז $\deg m_A=n$ ולכן $m_A=f_A$.*

הוכחה: יהי $0\le k\le n-1$.

טענה 1: *לעמודה $A^k\mathbf{e}_1$ יש רכיב שונה מאפס בשורה $k+1$ ואפסים מתחתיו.* אכן, עבור $k=0$: $A^0\mathbf{e}_1=\mathbf{e}_1$ מקיים את הטענה. נניח נכונות עבור $k-1$. אז יש $a_1,\ldots,a_{k-1},a_k\in F$, באשר $a_k\neq 0$ כך שמתקיים

$$A^{k-1}\mathbf{e}_1=(a_1,\ldots,a_{k-1},a_k,0,\ldots,0)^t=a_1\mathbf{e}_1+\ldots+a_{k-1}\mathbf{e}_{k-1}+a_k\mathbf{e}_k,$$

לכן $A^k\mathbf{e}_1=a_1A\mathbf{e}_1+\ldots+a_{k-1}A\mathbf{e}_{k-1}+a_kA\mathbf{e}_k$. אך $A\mathbf{e}_1,\ldots,A\mathbf{e}_{k-1},A\mathbf{e}_k$ הן k העמודות הראשונות של A. לפנ הנתון, יש להן אפסים בשורות $k+2,\ldots,n$, וגם אפס בשורה $k+1$, פרט לעמודה $A\mathbf{e}_k$, לה יש שם רכיב שונה מאפס. לכן $A^k\mathbf{e}_1$ מהצורה המבוקשת.

טענה 2: *$A^k\mathbf{e}_1$ אינו צירוף לינארי של $A^0\mathbf{e}_1,A\mathbf{e}_1,\ldots,A^{k-1}\mathbf{e}_1$.* אכן, לעמודה $A^k\mathbf{e}_1$ יש בשורה $k+1$ רכיב שונה מאפס, בעוד שלעמודות $A^0\mathbf{e}_1,A\mathbf{e}_1,\ldots,A^{k-1}\mathbf{e}_1$, ולכן גם לכל צירוף לינארי שלהם, יש שם 0.

טענה 3: $\deg m_A=n$. אכן, $m_A|f_A$, לכן $\deg m_A\leq\deg f_A=n$. יהי $m_A=\sum_{i=0}^d a_iX^i$, באשר $a_d=1$. אז $\sum_{i=0}^d a_iA^i=m_A(A)=0$ ולכן $\sum_{i=0}^d a_iA^i\mathbf{e}_1=0$. מכאן $A^d\mathbf{e}_1=\sum_{i=0}^{d-1}-a_iA^i\mathbf{e}_1$. לפי טענה 2 זה לא יתכן אם $d\leq n-1$. לכן $\deg m_A=d\geq n$.

לבסוף, $m_A|f_A$ מתוקנים, $\deg m_A=n=\deg f_A$, לכן $m_A=f_A$. ∎

בהמשך יהי V מרחב וקטורי ממימד סופי n מעל שדה F ותהי $T\colon V\to V$ העתקה לינארית.

הגדרה 7.13: פולינום $m(X)\in F[X]$ ייקרא **פולינום מזערי של** T אם $m(T)=0$, m הוא מתוקן ובעל מעלה מזערית מבין כל הפולינומים ב־$F[X]$ השונים מ־0 והמאפסים את T. אם הוא קיים ויחיד, נסמנו m_T. ∎

משפט 7.14: *פולינום מזערי של T קיים ויחיד. ביתר דיוק*

(א) *אם $V=\{0\}$, אז $m_T=1$.*

(ב) *אם $V\neq\{0\}$, יהי $\mathcal{B}$ בסיס של V ותהי $C=[T]_{\mathcal{B}}^{\mathcal{B}}$. אז $m_T=m_C$.*

הוכחה: (א) לכל $m\in F[X]$ מתקיים $m(T)=0$, כי אפס היא ההעתקה הלינארית היחידה $V\to V$.

(ב) יהי $g\in F[X]$. לפי טענה 6.15, $g(C)=0\Leftrightarrow[g(T)]_{\mathcal{B}}^{\mathcal{B}}=0\Leftrightarrow g(T)=0$. לכן g מתוקן וממעלה מזערית שמאפס את T אם ורק אם g מתוקן וממעלה מזערית שמאפס את C. מכאן הטענה. ∎

תרגיל 7.15: *לכל $f\in F[X]$ מתקיים: $m_T|f\Leftrightarrow f(T)=0$.*

הוכחה: זהו אנלוג של משפט 7.2, ואפשר להוכיח אותו באותו אופן אך אנו נסיק אותו מהמשפט: אם $V=\{0\}$ אז $f(T)=0$ ו־$1|f$ לכל $f\in F[X]$, לכן הטענה נכונה. אם $V\neq\{0\}$, יהי $\mathcal{B}$ בסיס של V ותהי $C=[T]_{\mathcal{B}}^{\mathcal{B}}$. לפי טענה 6.15, $[f(T)]_{\mathcal{B}}^{\mathcal{B}}=f(C)$. לכן $f(C)=0\Leftrightarrow f(T)=0$. לפי משפט 7.2, $m_C|f\Leftrightarrow f(C)=0$. אך לפי משפט 7.14 $m_T=m_C$. לכן $m_T|f\Leftrightarrow f(T)=0$. ∎

8. ריבויים של ערכים עצמיים וליכסון, שילוש

יהי V מרחב נוצר סופית מעל שדה F ותהי $T\colon V \to V$ העתקה לינארית.

הגדרה 8.1: יהי $\lambda \in F$ ערך עצמי של T. **הריבוי האלגברי של** λ הוא המעריך k עמו מופיע $(X-\lambda)$ בפירוק של f_T לחזקות של גורמים אי פריקים שונים. **הריבוי הגיאומטרי של** λ הוא המימד של $V_\lambda = \{v \in V \mid T(v) = \lambda v\}$.

∎

למה 8.2: *יהי $\lambda \in F$ ערך עצמי של T. אז הריבוי הגיאומטרי של λ קטן או שווה לריבוי האלגברי של λ.*

הוכחה: יהי $r = \dim V_\lambda \geq 1$ הריבוי הגיאומטרי של λ ויהי $v_1, \ldots, v_r$ בסיס של V_λ. נשלים אותו לבסיס $v_1, \ldots, v_r, w_1, \ldots, w_s$ של V. אז $r + s = n$. המטריצה של T לפי בסיס זה היא

$$C = \begin{pmatrix} \lambda & 0 & \cdots & 0 & & \\ 0 & \lambda & \cdots & 0 & & \\ \cdot & \cdot & \cdots & \cdot & & * \\ 0 & 0 & \cdots & \lambda & & \\ 0 & 0 & \cdots & 0 & & \\ \cdot & \cdot & \cdots & \cdot & & * \\ 0 & 0 & \cdots & 0 & & \end{pmatrix} = \begin{pmatrix} \lambda I_r & A \\ 0 & B \end{pmatrix}$$

לכן לפי תרגיל 6.12, $f_T(X) = f_C(X) = f_{\lambda I_r}(X) f_B(X) = (X-\lambda)^r f_B(X)$. ∎

משפט 8.3: *להעתקה $T\colon V \to V$ יש הצגה אלכסונית אם ורק אם f_T הוא מכפלה של גורמים ממעלה אחת ב־$F[X]$ ולכל ערך עצמי של T הריבוי האלגברי שווה לריבוי הגיאומטרי שלו.*

הוכחה: יהיו $\lambda_1, \ldots, \lambda_r \in F$ הערכים העצמיים השונים של T ולכל i יהי k_i הריבוי האלגברי של λ_i.

נניח שהתנאי מתקיים, כלומר, $f_T(X) = \prod_{i=1}^r (X-\lambda_i)^{k_i}$ ו־$\dim V_{\lambda_i} = k_i$ לכל i. אם נצרף בסיסים סדורים של $V_{\lambda_1}, \ldots, V_{\lambda_r}$ לסדרה אחת, לפי מסקנה 4.11 נקבל סדרה $\mathcal{B}$ בלתי תלויה לינארית של $\sum_{i=1}^r k_i = \deg f_T = \dim V$ וקטורים עצמיים של T. לכן $\mathcal{B}$ בסיס של V ולפי מסקנה 4.8 ל־T הצגה אלכסונית.

להיפך, אם $\mathcal{B}$ בסיס כך ש־$[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}$, אז $f_T = f_{[T]_{\mathcal{B}}^{\mathcal{B}}} = \prod_{j=1}^n (X-\alpha_j)$ מכפלה של גורמים ממעלה 1. לכן $f_T(X) = \prod_{i=1}^r (X-\lambda_i)^{k_i}$. לפי משפט 4.7, אברי $\mathcal{B}$ הם וקטורים עצמיים של T ולכן $\mathcal{B} = \bigcup_{i=1}^r \mathcal{B}_i$, באשר $\mathcal{B}_i = \mathcal{B} \cap V_{\lambda_i}$ לכל i. לכל i מתקיים $\mathcal{B}_i \subseteq \mathcal{B}$, לכן $\mathcal{B}_i$ בלתי תלויה לינארית, ולכן $|\mathcal{B}_i| \leq \dim V_{\lambda_i} \leq k_i$, לפי למה 8.2. אילו היה j כך ש־$\dim V_{\lambda_j} < k_j$ אז $|\mathcal{B}| = \sum_{i=1}^r |\mathcal{B}_i| < \sum_{i=1}^r k_i \leq \deg f_T = \dim V$, בסתירה לכך שבְּבסיס יש $\dim V$ אברים. ∎

הגדרה 8.4: מטריצה $A \in \mathrm{M}_n(F)$ נקראת **משולשית (עליונה)** אם $(A)_{ij} = 0$ לכל $i > j$, כלומר, A מהצורה

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$ ∎

משפט 8.5: *תהי $A \in \mathrm{M}_n(F)$ ויהיו $\lambda_1, \dots, \lambda_n \in F$. אז יש מטריצה משולשית עליונה $C \in \mathrm{M}_n(F)$ עם אלכסון ראשי $(\lambda_1, \dots, \lambda_n)$ ש־A דומה לה אם ורק אם $f_A(X) = (X - \lambda_1) \cdots (X - \lambda_n)$.*

הוכחה: $\Leftarrow$: נניח כי A דומה ל־C משולשית עליונה עם אלכסון ראשי $(\lambda_1, \dots, \lambda_n)$. אז

$$f_A(X) = f_C(X) = \det \begin{pmatrix} X - \lambda_1 & & * \\ & \ddots & \\ 0 & & X - \lambda_n \end{pmatrix} = (X - \lambda_1) \cdots (X - \lambda_n).$$

$\Rightarrow$: באינדוקציה על n. אם $n = 1$, כל $A \in \mathrm{M}_1(F)$ משולשית עליונה.

נניח $n > 1$. אז $f_A(\lambda_1) = 0$, לכן λ_1 ערך עצמי של A, כלומר, יש $0 \neq v_1 \in F^n$ כך ש־$Av_1 = \lambda_1 v_1$. נשלים את v_1 לבסיס $v_1, v_2, \dots, v_n$ של F^n ותהי $Q = (v_1, v_2, \dots, v_n) \in \mathrm{M}_n(F)$. אז Q הפיכה, כיון שעמודותיה בלתי תלויות לינארית, ו־$Q^{-1}v_1 = \mathbf{e}_1$ כי $Q\mathbf{e}_1 = v_1$, ולכן

$$\begin{aligned} Q^{-1}AQ &= Q^{-1}(Av_1, \dots, Av_n) = Q^{-1}(\lambda_1 v_1, Av_2, \dots, Av_n) = (\lambda_1 \mathbf{e}_1, * \cdots *) = \\ &= \begin{pmatrix} \lambda_1 & * \cdots * \\ 0 & \\ \vdots & A_1 \\ 0 & \end{pmatrix} \end{aligned} \tag{8}$$

באשר $A_1 \in \mathrm{M}_{n-1}(F)$.

מכאן לפי תרגיל 6.12(ב)

$$(X - \lambda_1) \cdots (X - \lambda_n) = f_A(X) = f_{Q^{-1}AQ}(X) = (X - \lambda_1) f_{A_1}(X),$$

ולכן

$$f_{A_1}(X) = (X - \lambda_2) \cdots (X - \lambda_n). \tag{9}$$

לפי הנחת האינדוקציה יש $Q_1 \in \mathrm{M}_{n-1}(F)$ הפיכה כך ש־

$$Q_1^{-1} A_1 Q_1 = \begin{pmatrix} \lambda_2 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \in \mathrm{M}_{n-1}(F). \tag{10}$$

נגדיר $Q_2 = \begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix} \in \mathrm{M}_n(F)$. נשים לב ש־ $\begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & Q_1^{-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & I_{n-1} \end{pmatrix} = I_n$, לכן Q_2 הפיכה ו־ $Q_2^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & Q_1^{-1} \end{pmatrix}$. מכאן שגם $P = QQ_2$ הפיכה. מתקיים

$$\begin{aligned} P^{-1}AP &= (Q_2^{-1}Q^{-1})A(QQ_2) = Q_2^{-1}(Q^{-1}AQ)Q_2 = \\ &\begin{pmatrix} 1 & 0 \\ 0 & Q_1^{-1} \end{pmatrix} \begin{pmatrix} \lambda_1 & * \\ 0 & A_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & Q_1^{-1} \end{pmatrix} \begin{pmatrix} \lambda_1 & * \\ 0 & A_1 Q_1 \end{pmatrix} = \\ &\begin{pmatrix} \lambda_1 & * \\ 0 & Q_1^{-1} A_1 Q_1 \end{pmatrix} = \begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \end{aligned}$$

∎

מסקנה 8.6: *אם F שדה סגור אלגברית, אז כל $A \in \mathrm{M}_n(F)$ דומה למשולשית עליונה.*

הוכחה: כיון ש־f_A מתוקן, לפי משפט 2.13, $f_A = (X-\lambda_1)\cdots(X-\lambda_n)$. לכן לפי משפט 8.4, A דומה למשולשית עליונה. ∎

מסקנה 8.7: *יהי V מרחב וקטורי ממימד n מעל שדה F, תהי $T\colon V \to V$ העתקה לינארית ויהיו $\lambda_1, \ldots, \lambda_n \in F$. אז קיים בסיס $\mathcal{B}$ של V כך ש־$[T]_{\mathcal{B}}^{\mathcal{B}}$ משולשית עליונה עם אלכסון ראשי $(\lambda_1, \ldots, \lambda_n)$ אם ורק אם $f_T = (X-\lambda_1)\cdots(X-\lambda_n)$.*

הוכחה: $\Leftarrow$: $f_T = f_{[T]_{\mathcal{B}}^{\mathcal{B}}} = (X-\lambda_1)\cdots(X-\lambda_n)$.

$\Rightarrow$: יהי $\mathcal{A}$ בסיס כלשהו של V. אז $f_{[T]_{\mathcal{A}}^{\mathcal{A}}} = f_T$, לכן לפי המשפט $[T]_{\mathcal{A}}^{\mathcal{A}}$ דומה למשולשית עליונה C עם אלכסון ראשי $(\lambda_1, \ldots, \lambda_n)$. אז יש בסיס $\mathcal{B}$ כך ש־$[T]_{\mathcal{B}}^{\mathcal{B}} = C$. ∎

תרגיל 8.8: *יהי V מרחב וקטורי ממימד n מעל שדה F, תהי $T\colon V \to V$ העתקה לינארית ויהי $\mathcal{B} = (v_1, \ldots, v_n)$ בסיס סדור של V. יהי $\mathcal{B}' = (v_n, \ldots, v_1)$. הוכיחו ש־$[T]_{\mathcal{B}}^{\mathcal{B}}$ משולשית עליונה אם ורק אם $[T]_{\mathcal{B}'}^{\mathcal{B}'}$ משולשית תחתונה.*

9. הפירוק הפרימרי

בסעיף זה יהי V מרחב וקטורי מעל שדה F ותהי $T\colon V \to V$ העתקה לינארית. בכל מקום שמוזכר m_T או f_T נניח גם ש־V נוצר סופית (ואז m_T ו־f_T מוגדרים).

הגדרה 9.1: תת מרחב W של V נקרא **שְׁמוּר־T** (**T-invariant**) אם $T(W) \subseteq W$, כלומר, אם לכל $w \in W$ מתקיים $T(w) \in W$. במקרה כזה **הצמצום של T ל־W**, כלומר, ההעתקה $T'\colon W \to W$ המוגדרת על ידי $T'(w) = T(w)$ לכל $w \in W$, היא העתקה לינארית; היא תסומן $T|_W$. ∎

הערה: $T(W) := \{T(w) |\ w \in W\}$. ∎

דוגמה 9.2: (א) $0, V$ הם תת מרחבים שמורי־T לכל T.

(ב) אם λ ערך עצמי של T ו־W תת מרחב של V_λ (למשל, $W = V_\lambda$), אז W הוא שמור־T: אכן, אם $w \in W$ אז $T(w) = \lambda w \in W$.

(ג) אם $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ הוא הסיבוב בזוית θ סביב ציר ה־z, אז ציר ה־z (=$\mathrm{Sp}(\mathbf{e}_3)$) הוא שמור־$T$. גם מישור־$(x, y)$, כלומר, $\mathrm{Sp}(\mathbf{e}_1, \mathbf{e}_2)$, הוא שמור־$T$. ∎

תרגיל 9.3: *תהיינה $T, S\colon V \to V$ שתי העתקות לינאריות כך ש־$TS = ST$. אז*

(א) $\mathrm{Ker}(S), \mathrm{Im}(S)$ *הם שמורי־T.*

(ב) *אם $W \subseteq V$ תת מרחב שמור־T אז גם $S(W)$ תת מרחב שמור־T. (כאן $S(W) = \{S(w) |\ w \in W\}$.)*

(ג) *אם $W_1, W_2 \subseteq V$ תת מרחבים שמורי־T אז גם $W_1 \cap W_2$ ו־$W_1 + W_2$ שמורי־T.*

(ד) *נניח ש־T איזומורפיזם. אם $W \subseteq V$ תת מרחב שמור־T ו־$\dim W < \infty$ אז W שמור־T^{-1}.*

פתרון: (א) נראה ש־$\mathrm{Ker}(S)$ שמור־T:

יהי $v \in \mathrm{Ker}(S)$. אז $S(v) = 0$, לכן $S(T(v)) = T(S(v)) = T(0) = 0$, לכן $T(v) \in \mathrm{Ker}(S)$.

זה ש־$\mathrm{Im}(S)$ הוא השמור־T הוא מקרה פרטי של (ב), עם $W = V$.

(ב) W תת מרחב שמור־T של V. נראה ש־$S(W)$ שמור־T:

יהי $v \in S(W)$. אז יש $w \in W$ כך ש־$v = S(w)$. כיוון ש־W שמור־T, מתקיים $T(w) \in W$. לכן $T(v) = T(S(w)) = S(T(w)) \in S(W)$.

(ג) W_1, W_2 שמורי־T. נראה ש־$W_1 + W_2 = \{w_1 + w_2 |\ w_1 \in W_1,\ w_2 \in W_2\}$ שמור־T:

יהי $w_1 + w_2 \in W_1 + W_2$ (באשר $w_1 \in W_1, w_2 \in W_2$). אז $T(w_1) \in W_1, T(w_2) \in W_2$, לכן $T(w_1 + w_2) = T(w_1) + T(w_2) \in W_1 + W_2$.

נראה ש־$W_1 \cap W_2$ שמור־T:

יהי $w \in W_1 \cap W_2$. אז $w \in W_1$, לכן $T(w) \in W_1$. באופן דומה $w \in W_2$, לכן $T(w) \in W_2$. מכאן $T(w) \in W_1 \cap W_2$.

(ד) $T(W) \subseteq W$ הוא תת מרחב של W. אבל T חד חד ערכית, לכן הצמצום $T|_W \colon W \to W$ גם חד חד ערכי. לכן $\dim T(W) = \dim W$. מכאן $T(W) = W$.

נראה ש־W שמור־T^{-1}:

יהי $w \in W$. אז יש $v \in W$ כך ש־$T(v) = w$. אבל אז $T^{-1}(w) = v \in W$. ∎

מסקנה 9.4: *יהי $f \in F[X]$. אז $\operatorname{Ker} f(T), \operatorname{Im} f(T)$ הם תת מרחבים שמורי־T של V.*

הוכחה: כפי שהערנו בהגדרה 6.14, $f(T) \circ T = T \circ f(T)$. לכן לפי תרגיל 9.3(א) נובעת הטענה. ∎

תרגיל 9.5: *נניח כי V נוצר סופית ויהי W תת מרחב שמור־T של V. נשלים בסיס $\mathcal{B}_0$ של W לבסיס $\mathcal{B}$ של V. אז $[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} A & C \\ 0 & D \end{pmatrix}$ באשר $A = [T|_W]_{\mathcal{B}_0}^{\mathcal{B}_0}$.*

פתרון: נניח $\mathcal{B}_0 = (v_1, \dots, v_r)$ ו־$\mathcal{B} = (v_1, \dots, v_r, v_{r+1}, \dots, v_n)$. יהי $1 \le j \le r$ ונניח כי $[T(v_j)]_{\mathcal{B}_0}^{\mathcal{B}_0} = (a_{1j}, \dots, a_{rj})^t$. אז

$$T(v_j) = a_{1j}v_1 + \cdots + a_{rj}v_r + 0v_{r+1} + \cdots + 0v_n$$

לכן $[T(v_j)]_{\mathcal{B}}^{\mathcal{B}} = (a_{1j}, \dots, a_{rj}, 0, \dots, 0)^t$.

כלומר, אם $1 \le j \le r$, אז העמודה ה־j של $[T]_{\mathcal{B}}^{\mathcal{B}}$ מורכבת מהעמודה ה־j של $[T]_{\mathcal{B}_0}^{\mathcal{B}_0}$ ו־$n-r$ אפסים מתחתיה. מכאן המסקנה. ∎

הגדרה 9.6: יהי V מרחב וקטורי ויהיו $W_1, \dots, W_s$ תת מרחבים שלו. נאמר ש־V הוא **הסכום הישר** של $W_1, \dots, W_s$ ונכתוב זאת כך: $V = W_1 \oplus \cdots \oplus W_s$ אם לכל $v \in V$ יש הצגה יחידה מהצורה

$$v = w_1 + \cdots + w_s, \qquad w_1 \in W_1, \dots, w_s \in W_s \tag{1}$$ ∎

משפט 9.7: *יהי V מרחב וקטורי ויהיו $W_1, \dots, W_s$ תת מרחבים שלו. לכל $1 \le i \le s$ יהי $\mathcal{B}_i = (w_{i1}, \dots, w_{in_i})$ בסיס סדור של W_i (באשר $n_i = \dim W_i$). אזי $V = W_1 \oplus \cdots \oplus W_s$ אם ורק אם צירוף הבסיסים האלה*

$$\mathcal{B} = (w_{11}, \dots, w_{1n_1}\ ,\ \dots,\ w_{s1}, \dots, w_{sn_s})$$

הוא בסיס של V.

הוכחה: "$\Leftarrow$:" נראה ש־$\mathcal{B}$ סדרת יוצרים. יהי $v \in V$. לפי ההנחה יש לו הצגה (1). לכל $1 \le i \le s$ יש הצגה $w_i = \sum_{j=1}^{n_i} a_{ij}w_{ij}$, באשר $a_{ij} \in F$. לכן $v = \sum_{i=1}^{s}\sum_{j=1}^{n_i} a_{ij}w_{ij}$ צירוף לינארי של איברי $\mathcal{B}$.

נראה ש־$\mathcal{B}$ בלתי תלויה לינארית. יהיו $a_{ij} \in F$ כך ש־ $\sum_{i=1}^{s}\sum_{j=1}^{n_i} a_{ij}w_{ij} = 0$. מתקיים

$$\sum_{i=1}^{s}\Big(\sum_{j=1}^{n_i} a_{ij}w_{ij}\Big) = 0 = \sum_{i=1}^{s} 0$$

באשר $\sum_{j=1}^{n_i} a_{ij}w_{ij} \in W_i$ לכל i. לפי היחידות ב־(1), $\sum_{j=1}^{n_i} a_{ij}w_{ij} = 0$ לכל i. בגלל ש־$\mathcal{B}_i$ בסיס, $a_{ij} = 0$ לכל j, לכל i. לכן $\mathcal{B}$ בלתי תלויה לינארית.

הוכחנו ש־$\mathcal{B}$ סדרת יוצרים של V וגם בלתי תלויה לינארית, לכן היא בסיס של V.

"$\Rightarrow$:" נראה שלכל $v \in V$ יש הצגה (1). כיוון ש־$\mathcal{B}$ בסיס של V, יש $a_{ij} \in F$ יחידים כך ש־ $v = \sum_{i=1}^{s}\sum_{j=1}^{n_i} a_{ij}w_{ij}$. נסמן $w_i = \sum_{j=1}^{n_i} a_{ij}w_{ij}$ לכל i. אז $w_i \in W_i$ ומתקיים (1).

נראה את היחידות של הצגה (1). אם מתקיים גם

$$v = w'_1 + \cdots + w'_s, \qquad w'_1 \in W_1, \ldots, w'_s \in W_s \tag{1}$$

אז, לכל w'_i, בגלל ש־$\mathcal{B}_i$ בסיס של W_i, יש הצגה $w'_i = \sum_{j=1}^{n_i} a'_{ij}w_{ij}$, באשר $a'_{ij} \in F$. מכאן $v = \sum_{i=1}^{s}\sum_{j=1}^{n_i} a'_{ij}w_{ij}$. מהיחידות של a_{ij} לעיל נובע $a'_{ij} = a_{ij}$ לכל i, j, ולכן $w'_i = w_i$ לכל i. ∎

תרגיל 9.8: *יהי W תת מרחב שמור־T של V ויהיו $f, g \in F[X]$.*

(א) *הוכיחו כי W שמור־$g(T)$.*

(ב) *הוכיחו כי $f(T)(W)$ שמור־$g(T)$*

(ג) *אם $V = W_1 \oplus \cdots \oplus W_s$, באשר $W_1, \ldots, W_s$ שמורי־T, אז*

$$f(T)(V) = f(T)(W_1) \oplus \cdots \oplus f(T)(W_s).$$

הוכחה: (א)

יהי $w \in W$. באינדוקציה על i, $T^i(w) \in W$ לכל $i \geq 0$. אכן, $T^0(w) = 1_V(w) = w \in W$, ($T(w) \in W$ בגלל ש־W שמור־T) ואם $T^{i-1}(w) \in W$, אז $T^i(w) = T(T^{i-1}(w)) \in W$.

לכן אם $g = \sum_{i=0}^{\infty} a_iX^i$, אז $g(T)(w) = \sum_{i=0}^{\infty} a_iT^i(w) \in W$.

(ב)

לפי (א), W שמור־$g(T)$. מתקיים $g(T)f(T) = f(T)g(T)$, לכן לפי תרגיל 9.3(ב), $f(T)(W)$ שמור־$g(T)$.

(ג)

יהי $u \in f(T)(V)$, כלומר, $u = f(T)(v)$ עבור איזה $v \in V$. אז יש $w_i \in W_i$, לכל $1 \leq i \leq s$ כך ש־$v = w_1 + \cdots + w_s$. מכאן

$$u = f(T)(v) = f(T)(w_1) + \cdots + f(T)(w_s),$$

באשר $f(T)(w_i) \in f(T)(W_i)$ לכל i.

נניח שגם $u = f(T)(v) = u_1 + \cdots + u_s$, באשר $u_i \in f(T)(W_i)$ לכל i. כיוון ש־$f(T)(w_i) \in f(T)(W_i) \subseteq W_i$ וגם $u_i \in W_i$, מהיחידות בהגדרה $V = W_1 \oplus \cdots \oplus W_s$, נובע $f(T)(w_i) = u_i$ לכל i.

∎

משפט 9.9: *נניח כי* $V = W_1 \oplus \cdots \oplus W_s$. *לכל* $1 \le i \le s$ *יהי* $\mathcal{B}_i = (w_{i1}, \ldots, w_{in_i})$ *בסיס של* W_i, *כך ש־* $\mathcal{B} = (w_{11}, \ldots, w_{1n_1}\ ,\ \ldots\ ,\ w_{s1}, \ldots, w_{sn_s})$ *בסיס של* V.

(א) *נרשום את* $[T]_{\mathcal{B}}^{\mathcal{B}}$ *כמטריצת גושים,* $[T]_{\mathcal{B}}^{\mathcal{B}} = (A_{ij}) = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ A_{s1} & A_{s2} & \cdots & A_{ss} \end{pmatrix}$, *באשר* $A_{ij} \in \mathrm{M}_{n_i \times n_j}(F)$ *לכל* $1 \le i, j \le s$. *נקבע* $1 \le j \le s$. *אז* W_j *שמור־* T *אם ורק אם* $A_{ij} = 0$ *לכל* $i \ne j$. *אם תנאים שקולים אלה מתקיימים, אז* $A_{jj} = [T|_{W_j}]_{\mathcal{B}_j}^{\mathcal{B}_j}$.

(ב) $W_1, \ldots, W_s$ *שמורי־* T *אם ורק אם* $[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_s \end{pmatrix}$ *באשר* $A_j \in \mathrm{M}_{n_j}(F)$ *לכל* $1 \le j \le s$.

(ג) *אם התנאי השקולים ב־(ב) מתקיימים, אז* $A_j = [T|_{W_j}]_{\mathcal{B}_j}^{\mathcal{B}_j}$ *לכל* $1 \le j \le s$.

הוכחה: (א) לפי ההגדרה של $[T]_{\mathcal{B}}^{\mathcal{B}}$, $(A_{ij})_{kl}$ הוא המקדם בפיתוח של $T(w_{jl})$, כצירוף לינארי של אברי $\mathcal{B}$, שעומד לפני w_{ik}. לכן:

$A_{ij} = 0$ לכל $i \ne j$

$\Leftrightarrow$ $(A_{ij})_{kl} = 0$ לכל $i \ne j$ ולכל k, l $\quad (1 \le k \le n_i,\ 1 \le l \le n_j)$

$\Leftrightarrow$ לכל l, המקדם בפיתוח של $T(w_{jl})$, כצירוף לינארי של אברי $\mathcal{B}$, שעומד לפני w_{ik} הוא 0, לכל $i \ne j$ ולכל k

$\Leftrightarrow$ לכל l, $T(w_{jl})$ הוא צירוף לינארי של $w_{j1}, \ldots, w_{jn_j}$

$\Leftrightarrow$ לכל l, $T(w_{j1}), \ldots, T(w_{jn_j}) \in \mathrm{Sp}(w_{j1}, \ldots, w_{jn_j})$

$\Leftrightarrow$ $T(\mathrm{Sp}(w_{j1}, \ldots, w_{jn_j})) \subseteq \mathrm{Sp}(w_{j1}, \ldots, w_{jn_j})$

$\Leftrightarrow$ $T(W_j) \subseteq W_j$.

אם W_j שמור־T ולכן $T|_{W_j}$ מוגדר, אז $([T|_{W_j}]_{\mathcal{B}_j}^{\mathcal{B}_j})_{kl}$ הוא המקדם בפיתוח של $T(w_{jl})$ כצירוף לינארי של אברי $\mathcal{B}_j$ שעומד לפני w_{jk}. לפי המשפט הראשון של ההוכחה, $(A_{jj})_{kl} = ([(T|_{W_j}]_{\mathcal{B}_j}^{\mathcal{B}_j})_{kl}$, לכל k, l.

(ב), (ג) הם מקרים פרטיים של (א). ∎

מסקנה 9.10: *נניח כי* $V = W_1 \oplus \cdots \oplus W_s$, *באשר* $W_1, \ldots, W_s$ *הם שמורי־* T. *נסמן* $T_i = T|_{W_i}$ *לכל* $1 \le i \le s$. *אז*

$$f_T(X) = f_{T_1}(X) \cdots f_{T_s}(X)$$

$$m_T(X) = \mathrm{lcm}(m_{T_1}(X) \cdots m_{T_s}(X)).$$

הוכחה: בסימונים של המשפט תהי $A = [T]_{\mathcal{B}}^{\mathcal{B}}$ ותהיינה $A_i = [T|_{W_i}]_{\mathcal{B}_i}^{\mathcal{B}_i}$ לכל $1 \le i \le s$.
אז $f_T = f_A$, $m_T = m_A$, $f_{T_i} = f_{A_i}$, ו־ $m_{T_i} = m_{A_i}$, לכל i. לכן צריך להוכיח

$$f_A(X) = f_{A_1}(X) \cdots f_{A_s}(X)$$

$$m_A(X) = \mathrm{lcm}\left(m_{A_1}(X) \cdots m_{A_s}(X)\right).$$

וזה נכון לפי תרגיל 6.12 ומשפט 7.10, כי לפי המשפט הקודם $A = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_s \end{pmatrix}$. ∎

למה 9.11: *יהי $W \subseteq V$ תת מרחב שמור־T. יהי $S = T|_W : W \to W$ הצמצום של T ל־W. אז*

(א) *יהי $f \in F[X]$. אז לכל $w \in W$ מתקיים $f(S)(w) = f(T)(w)$. כלומר, $f(T|_W) = f(T)|_W$.*

(ב) $m_S | m_T$.

(ג) $f_S | f_T$.

הוכחה: (א) תחילה נראה באינדוקציה כי $S^i(w) = T^i(w) \in W$ לכל $i \geq 0$. ואכן, עבור $i = 0$, $S^0 = 1_W, T^0 = 1_V$, לכן $S^0(w) = w = T^0(w)$. נניח נכונות עבור $i-1$, אז

$$S^i(w) = S(S^{i-1}(w)) = S(T^{i-1}(w)) = T(T^{i-1}(w)) = T^i(w).$$

מכאן שאם $f = \sum_{i=0}^{\infty} a_i X^i$ אז

$$f(S)(w) = \big(\sum_{i=0}^{\infty} a_i S^i\big)(w) = \sum_{i=0}^{\infty} a_i S^i(w) = \sum_{i=0}^{\infty} a_i T^i(w) = \big(\sum_{i=0}^{\infty} a_i T^i\big)(w) = f(T)(w).$$

(ב) לפי תרגיל 7.15 די להוכיח כי $m_T(S) = 0$. ואכן, לכל $m_T(S)(w) = m_T(T)(w) = 0(w) = 0$ $w \in W$, כלומר, $m_T(S) = 0$.

(ג) נניח $\dim W = m$, $\dim V = m+k$. נבחר בסיס $\mathcal{B}_0$ של W ונשלים אותו לבסיס $\mathcal{B}$ של V. לפי תרגיל 9.5, $[T]^{\mathcal{B}}_{\mathcal{B}} = \begin{pmatrix} A & C \\ 0 & D \end{pmatrix}$, באשר $A = [S]^{\mathcal{B}_0}_{\mathcal{B}_0} \in \mathrm{M}_m(F)$. מכאן לפי תרגיל 6.12 $f_T(X) = f_S(X) \cdot f_D(X)$, ובפרט $f_S | f_T$. ∎

משפט 9.12: *יהי V מרחב וקטורי מעל שדה F ותהי $T: V \to V$ העתקה לינארית. יהיו $g, h \in F[X]$ זרים (כלומר המחלק המשותף הגדול ביותר שלהם הוא 1) כך ש־$(gh)(T) = 0$. אז $V = \operatorname{Ker} g(T) \oplus \operatorname{Ker} h(T)$.*

הוכחה: היות ו־g, h זרים, יש $r, s \in F[X]$ כך ש־$rg + sh = 1$ (למה 3.18). מכאן

$$r(T)g(T) + s(T)h(T) = 1_V. \tag{1}$$

יהי $v \in V$. לפי (1)

$$v = 1_V(v) = r(T)g(T)(v) + s(T)h(T)(v).$$

אבל $r(T)g(T)(v) \in \operatorname{Ker} h(T)$, כי

$$h(T)r(T)g(T)(v) = r(T)g(T)h(T)(v) = r(T)(gh)(T)(v) = r(T)(0) = 0.$$

ובאופן דומה $s(T)h(T)(v) \in \operatorname{Ker} g(T)$. לכן ל־$v$ יש הצגה

$$v = u + w, \qquad u \in \operatorname{Ker} g(T),\ w \in \operatorname{Ker} h(T). \tag{2}$$

נניח שהצגה (2) קיימת ונראה שהיא יחידה. לפי (2)

$$r(T)g(T)(v) = r(T)g(T)(u) + r(T)g(T)(w) = 0 + r(T)g(T)(w).$$

ואם נפעיל את שני האגפים של (1) על w, נקבל

$$w = 1_V(w) = r(T)g(T)(w) + s(T)h(T)(w) = r(T)g(T)(w) + 0.$$

משתי המשוואות האחרונות קיבלנו $w = r(T)g(T)(v)$, כלומר, w נקבע באופן יחיד על ידי v. באופן דומה גם u יחיד. ∎

מסקנה 9.13: *יהיו $g, h \in F[X]$ מתוקנים וזרים כך ש־$m_T = gh$. נסמן ב־T_1, T_2 את הצמצומים של T ל־$\operatorname{Ker} g(T), \operatorname{Ker} h(T)$ בהתאמה. אז $g = m_{T_1}, h = m_{T_2}$.*

הוכחה: $g(T_1) = 0$ כי לכל $u \in \operatorname{Ker} g(T)$ מתקיים $g(T_1)(u) = g(T)(u) = 0$. לכן לפי תרגיל 7.15, $m_{T_1}|g$, כלומר $g = rm_{T_1}$, באשר $r \in F[X]$; בהכרח r מתוקן. באותו אופן $h = sm_{T_2}$, באשר $s \in F[X]$ מתוקן. כעת, לפי מסקנה 9.10,

$$rsm_{T_1}m_{T_2} = gh = m_T = \operatorname{lcm}(m_{T_1}, m_{T_2})|m_{T_1}m_{T_2}$$

ומכאן ברור ש־$rs|1$, כלומר $r = s = 1$. לכן המסקנה. ∎

משפט 9.14 (הפירוק הפרימרי): *יהי V מרחב וקטורי מעל שדה F ותהי $T\colon V \to V$ העתקה לינארית כך ש־$m_T = g_1 \cdots g_s$ באשר $g_1, \ldots, g_s$ מתוקנים וזרים זה לזה. נסמן $W_i = \operatorname{Ker} g_i(T)$, עבור $i = 1, \ldots, s$. אזי*

(א) *$V = W_1 \oplus \cdots \oplus W_s$, באשר $W_1, \ldots, W_s$ שמורי־T;*

(ב) *$g_i = m_{T|_{W_i}}$ לכל $i = 1, \ldots, s$.*

הוכחה באינדוקציה על s. אם $s = 1$ אז $W_1 = V$ והכל ברור. נניח נכונות עבור $s - 1$. היות ושני הפולינומים $g_1, g_2 \cdots g_s$ זרים, לפי המשפט הקודם מתקיים $V = W_1 \oplus V'$, באשר $W_1 = \operatorname{Ker} g_1(T)$, $V' = \operatorname{Ker}(g_2 \cdots g_s)(T)$. כמו כן, W_1, V' הם שמורי־T. נסמן $S = T|_{V'}$. לפי המסקנה, $m_S = g_2 \cdots g_s$, $m_{T|_{W_1}} = g_1$.

לפי הנחת האינדוקציה $V' = W_2 \oplus \cdots \oplus W_s$, באשר $W_i = \operatorname{Ker} g_i(S)$ ו־ $g_i = m_{S|_{W_i}}$ לכל $i = 2, \ldots, s$. מכאן לפי משפט 9.7, $V = W_1 \oplus (W_2 \oplus \cdots \oplus W_s) = W_1 \oplus \cdots \oplus W_s$. יהי $2 \leq i \leq s$. ברור ש־$T|_{W_i} = S|_{W_i}$, ולכן $g_i = m_{T|_{W_i}}$. נותר להוכיח כי $\operatorname{Ker} g_i(S) = \operatorname{Ker} g_i(T)$. אך

$$(g_2 \cdots g_s)(T) = (g_2 \cdots g_{i-1} g_{i+1} \cdots g_s g_i)(T) = (g_2 \cdots g_{i-1} g_{i+1} \cdots g_s)(T) g_i(T),$$

לכן $\operatorname{Ker} g_i(T) \subseteq \operatorname{Ker}(g_2 \cdots g_s)(T) = V'$, ולכן לפי למה 9.11

$$\operatorname{Ker} g_i(T) = \{v \in V' |\ g_i(T)(v) = 0\} = \{v \in V' |\ g_i(S)(v) = 0\} = \operatorname{Ker} g_i(S). \qquad \blacksquare$$

השימוש העיקרי של המשפט: אם $m_T = q_1^{r_1} \cdots q_s^{r_s}$, , באשר $q_1, \ldots, g_s$ אי פריקים מתוקנים שונים, נקח $g_i = q_i^{r_i}$. אז $g_1, \ldots, g_s$ זרים.

מסקנה 9.15 (מטריצה לפי פירוק פרימרי): *יהי V מרחב וקטורי מעל שדה F ותהי $T\colon V \to V$ העתקה לינארית. נניח כי $m_T = q_1^{r_1} \cdots q_s^{r_s}$, באשר $q_1, \ldots, q_s$ פולינומים אי פריקים מתוקנים שונים ו־$r_1, \ldots, r_s \geq 1$. נסמן $W_i = \operatorname{Ker} q_i^{r_i}(T)$ לכל i. אזי*

(א) *צירוף $\mathcal{B}$ של בסיסים $\mathcal{B}_1, \ldots, \mathcal{B}_s$ של $W_1, \ldots, W_s$, בהתאמה, הוא בסיס של V שמקיים $[T]_{\mathcal{B}} = \operatorname{Diag}(A_1, \ldots, A_s)$, באשר $[T|_{W_i}]_{\mathcal{B}_i}^{\mathcal{B}_i} = A_i$ ו־ $m_{A_i} = q_i^{r_i}$ לכל i.*

(ב) *יהי $\mathcal{B}$ בסיס של V כך ש־$[T]_{\mathcal{B}} = \operatorname{Diag}(A_1, \ldots, A_t)$, באשר $m_{A_1}, \ldots, m_{A_t}$ חזקות של פולינומים אי פריקים מתוקנים שונים. נניח כי $A_i \in \mathrm{M}_{n_i}(F)$ לכל i. נכתוב את $\mathcal{B}$ כצירוף (שרשור) של סדרות $\mathcal{B}_1, \ldots, \mathcal{B}_t$, באשר $|\mathcal{B}_i| = n_i$ לכל i. אז $t = s$, ומתקיים לכל i, עד כדי הסדר: $\mathcal{B}_i$ בסיס של W_i, $[T|_{W_i}]_{\mathcal{B}_i}^{\mathcal{B}_i} = A_i$, ו־ $m_{A_i} = q_i^{r_i}$.*

הוכחה: (א) זה נובע ממשפט 9.14: לפי משפט 9.7, $\mathcal{B}$ אכן בסיס של V. לפי משפט 9.9, $[T]_{\mathcal{B}} = \operatorname{Diag}(A_1, \ldots, A_s)$, באשר $A_i = [T|_{W_i}]_{\mathcal{B}_i}^{\mathcal{B}_i}$ לכל i. מכאן לפי משפט 7.14, $m_{A_i} = m_{T|_{W_i}} = q_i^{r_i}$.

(ב) כיוון ש־$m_{A_1}, \ldots, m_{A_t}$ חזקות של פולינומים אי פריקים מתוקנים שונים, הם זרים. לכן לפי משפט 7.10,

$$q_1^{r_1} \cdots q_s^{r_s} = m_T = \operatorname{lcm}(m_{A_1}, \ldots, m_{A_t}) = m_{A_1} \cdots m_{A_t}.$$

לפי יחידות הפירוק לגורמים אי פריקים $t = s$ ו־ $m_{A_i} = q_i^{r_i}$, לכל i, עד כדי הסדר. כיוון ש־$\mathcal{B}$ הוא הצירוף של התת סדרות $\mathcal{B}_1, \ldots, \mathcal{B}_s$, לפי משפט 9.7 $V = \operatorname{Sp}(\mathcal{B}_1) \oplus \cdots \oplus \operatorname{Sp}(\mathcal{B}_t)$. כיוון ש־$[T]_{\mathcal{B}} = \operatorname{Diag}(A_1, \ldots, A_t)$, באשר $A_i \in \mathrm{M}_{n_i}(F)$ לכל i, לפי משפט 9.9, $\operatorname{Sp}(\mathcal{B}_i)$ הוא שמור־T ו־$[T|_{\operatorname{Sp}(\mathcal{B}_i)}]_{\mathcal{B}_i}^{\mathcal{B}_i} = A_i$, לכל i.

נותר להראות ש־$\mathcal{B}_i$ בסיס של W_i, לכל i. לפי השוויון האחרון, $m_{T|_{\operatorname{Sp}(\mathcal{B}_i)}} = m_{A_i} = q_i^{r_i}$. לכן אם $v \in \mathcal{B}_i$, אז לפי למה 9.11, $q_i^{r_i}(T)(v) = q_i^{r_i}(T|_{\operatorname{Sp}(\mathcal{B}_i)})(v) = 0$, כלומר, $v \in \operatorname{Ker} q_i^{r_i}(T) = W_i$. מכאן $\operatorname{Sp}(\mathcal{B}_i) \subseteq W_i$, ובפרט, כיוון ש־$\mathcal{B}_i$ סדרה בלתי תלויה לינארית, $|\mathcal{B}_i| \leq \dim W_i$. אבל $\sum_i |\mathcal{B}_i| = |\mathcal{B}| = \dim V = \sum_i \dim W_i$, לכן $|\mathcal{B}_i| = \dim W_i$ לכל i. לכן $\mathcal{B}_i$ בסיס של W_i. $\blacksquare$

משפט 9.16: *יהי V מרחב וקטורי מעל שדה F ותהי $T\colon V \to V$ העתקה לינארית. אזי יש בסיס $\mathcal{B}$ של V כך ש־$[T]_{\mathcal{B}}^{\mathcal{B}}$ אלכסונית אם ורק אם $m_T = (X-\mu_1)\cdots(X-\mu_s)$, באשר $\mu_1,\ldots,\mu_s \in F$ שונים זה מזה.*

הוכחה: נניח $A = [T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$. אז

$$m_T = m_A = \operatorname{lcm}(m_{(\lambda_1)},\ldots,m_{(\lambda_n)}) = \operatorname{lcm}(X-\lambda_1,\ldots,X-\lambda_n),$$

לכן $m_T = (X-\mu_1)\cdots(X-\mu_s)$, באשר $\mu_1,\ldots,\mu_s$ כל האברים השונים בקבוצה $\{\lambda_1,\ldots,\lambda_n\}$.

להיפך, נניח כי $m_T = (X-\mu_1)\cdots(X-\mu_s)$, באשר $\mu_1,\ldots,\mu_s \in F$ שונים זה מזה. לפי משפט 9.14, $V = W_1 \oplus \cdots \oplus W_s$, באשר

$$W_i = \operatorname{Ker}(T-\mu_i 1_V) = \{v \in V \mid T(v) = \mu_i v\} = V_{\mu_i}.$$

זה אומר שאיחוד הבסיסים של $V_{\mu_1},\ldots,V_{\mu_s}$ הוא בסיס של V, לכן ל־V יש בסיס $\mathcal{B}$ מורכב מוקטורים עצמיים של V. מכאן, לפי מסקנה 4.8, $[T]_{\mathcal{B}}^{\mathcal{B}}$ אלכסונית. ∎

מסקנה 9.17: *$A \in \mathrm{M}_n(F)$ דומה למטריצה אלכסונית אם ורק אם $m_A(X) = (X-\mu_1)\cdots(X-\mu_s)$ באשר $\mu_1,\ldots,\mu_s \in F$ שונים זה מזה.*

הוכחה: יש מרחב וקטורי V מעל F, העתקה לינארית $T\colon V \to V$, ובסיס $\mathcal{C}$ של V כך ש־$[T]_{\mathcal{C}}^{\mathcal{C}} = A$. לפי משפט 7.14, $m_T = m_A$.

מטריצה D (כלשהי) דומה ל־A אם ורק אם יש בסיס $\mathcal{B}$ של V כך ש־$[T]_{\mathcal{B}}^{\mathcal{B}} = D$. לכן A דומה למטריצה אלכסונית אם ורק אם יש בסיס $\mathcal{B}$ של V כך ש־$[T]_{\mathcal{B}}^{\mathcal{B}}$ אלכסונית. לפי משפט 9.16, (ובגלל ש־$m_T = m_A$) זה שקול לתנאי המבוקש על m_A. ∎

10. צורת ז'ורדן

בסעיף זה יהי V מרחב וקטורי בעל מימד סופי מעל שדה F ותהי $T\colon V \to V$ העתקה לינארית. (חלק מהחומר נכון גם עבור V בעל מימד אינסופי, אך לא נציין זאת במפורש.)

הגדרה 10.1: יהי $\lambda \in F$.

(א) המטריצה $\begin{pmatrix} \lambda & 0 & \cdots & 0 & 0 \\ 1 & \lambda & \ddots & & \vdots \\ & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & 1 & \lambda & 0 \\ 0 & 0 & \cdots & 1 & \lambda \end{pmatrix} \in \mathrm{M}_n(F)$ נקראת **גוש ז'ורדן מסדר** n **השייך ל־**λ ותסומן $J_n(\lambda)$.

(ב) מטריצה מהצורה $J(\lambda) = \begin{pmatrix} J_{n_1}(\lambda) & & 0 \\ & \ddots & \\ 0 & & J_{n_t}(\lambda) \end{pmatrix}$, באשר $n_1 \geq n_2 \geq \cdots \geq n_t \geq 1$, נקראת **מערך ז'ורדן השייך ל־**λ. המספר n_1 נקרא **האינדקס** של המערך.

(ג) **מטריצת ז'ורדן** היא מטריצה מהצורה $J = \begin{pmatrix} J(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J(\lambda_s) \end{pmatrix}$ באשר $\lambda_1, \ldots, \lambda_s \in F$ שונים זה מזה ולכל $1 \leq i \leq s$ המטריצה $J(\lambda_i)$ היא מערך ז'ורדן השייך ל־λ_i.

בפרט, מטריצת ז'ורדן היא מטריצת גושים באלכסון, אשר הינם גושי ז'ורדן. ∎

בסוף הפרק הזה נוכיח את המשפט המרכזי הבא:

משפט 10.2: *נניח $m_T = (X-\lambda_1)^{r_1} \cdots (X-\lambda_s)^{r_s}$, באשר $\lambda_1, \ldots, \lambda_s \in F$ שונים ו־$r_1, \ldots, r_s \geq 1$. אז קיימת מטריצת ז'ורדן יחידה (עד כדי סדר מערכיה) J עבורה יש בסיס $\mathcal{B}$ של V כך ש־$[T]_{\mathcal{B}}^{\mathcal{B}} = J$. (מטריצה זו נקראת* **צורת ז'ורדן** *של T.) ב־J יש s מערכי ז'ורדן, כאשר (עד כדי הסדר) המערך ה־i־י שייך ל־λ_i ובעל אינדקס r_i.*

ממנו נובעת מיידית:

מסקנה 10.3: *תהי $A \in \mathrm{M}_n(F)$. נניח $m_A = (X-\lambda_1)^{r_1} \cdots (X-\lambda_s)^{r_s}$, באשר $\lambda_1, \ldots, \lambda_s \in F$ שונים ו־$r_1, \ldots, r_s \geq 1$. אז קיימת מטריצת ז'ורדן יחידה (עד כדי סדר מערכיה) J דומה ל־A. (מטריצה זו נקראת* **צורת ז'ורדן** *של A.) ב־J יש s מערכי ז'ורדן, כאשר (עד כדי הסדר) המערך ה־i־י שייך ל־λ_i ובעל אינדקס r_i.*

דוגמה 10.4: (נסתמך על המסקנה הקודמת, אותה נוכיח אחר כך.)

(א) אם $A \in \mathrm{M}_5(F)$ ו־$m_A = X^2(X-1)^2$, אז צורת ז'ורדן שלה (=מטריצת ז'ורדן שדומה לה) היא

בהכרח (עד כדי סדר מערכיה) אחת –ורק אחת –בין השתיים הבאות

$$\begin{pmatrix} J_2(0) & 0 & 0 \\ 0 & J_2(1) & 0 \\ 0 & 0 & J_1(1) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\begin{pmatrix} J_2(0) & 0 & 0 \\ 0 & J_1(0) & 0 \\ 0 & 0 & J_2(1) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

איך יודעים איזו מבין שתי הצורות האלה היא הנכנוה? למשל לפי הפולינום האופייני: למטריצה הראשונה פולינום אופייני $X^2(X-1)^3$, ואילו לשנייה $X^3(X-1)^2$.

(ב) שתי המטריצות הבאות הן מטריצות ז'ורדן שונות (ממש, לא רק עד כדי סדר המערכים).

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(ביתר דיוק, הן מערכי ז'ורדן שונים השייכים ל־1.) לכן אינן דומות, לפי היחידות במסקנה 10.3. אך מתקיים $f_A = f_B = (X-1)^4$, ו־$m_A = m_B = (X-1)^2$. זה מראה שהפולינום האופייני והמזערי אינם קובעים את צורת ז'ורדן. ∎

תרגיל 10.5: *תהי $A = \mathrm{Diag}(A_1, \ldots, A_s)$ מטריצה, באשר $A_i \in \mathrm{M}_{n_i}(F)$ לכל $1 \le i \le s$. תני σ תמורה של הקבוצה $\{1, \ldots, n\}$ ותהי $A' = \mathrm{Diag}(A_{\sigma(1)}, \ldots, A_{\sigma(s)})$. אז A' דומה ל־A.*

הוכחה: יהי V מרחב וקטורי מעל F ממימד $n_1 + \cdots + n_s$ ויהי $\mathcal{B}$ בסיס שלו. אז יש העתקה לינארית $T\colon V \to V$ כך ש־$[T]_{\mathcal{B}}^{\mathcal{B}} = A$. נרשום את $\mathcal{B}$ כצירוף של תת סדרות $\mathcal{B} = \mathcal{B}_1 \vee \cdots \vee \mathcal{B}_s$, באשר $|\mathcal{B}_i| = n_i$ לכל i. יהי $W_i = \mathrm{Sp}(\mathcal{B}_i)$ לכל i.

לפי משפט 9.9, W_i הינו שמור־T ו־$A_i = [T|_{W_i}]_{\mathcal{B}_i}^{\mathcal{B}_i}$ לכל i.

ברור שגם $\mathcal{B}' = \mathcal{B}_{\sigma(1)} \vee \cdots \vee \mathcal{B}_{\sigma(s)}$ בסיס של V. לפי משפט 9.9, $[T]_{\mathcal{B}'}^{\mathcal{B}'} = \mathrm{Diag}(A'_1, \ldots, A'_s)$, באשר $A'_i = [T|_{W_{\sigma(i)}}]_{\mathcal{B}_{\sigma(i)}}^{\mathcal{B}_{\sigma(i)}}$ לכל i. לכן $[T]_{\mathcal{B}'}^{\mathcal{B}'} = A'$,

כיוון ש־A, A' מייצגות אותה העתקה T, הן דומות. ∎

דיון 10.6: העלאה בחזקה של מטריצת ז'ורדן. תחילה תהי

$$N = J_n(0) = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \ddots & & \vdots \\ 0 & 1 & 0 & \ddots & & \vdots \\ & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 \end{pmatrix} = (\mathbf{e}_2, \mathbf{e}_3, \ldots, \mathbf{e}_n, 0).$$

מכיוון ש־$N\mathbf{e}_n = 0$, $N\mathbf{e}_1 = \mathbf{e}_2, N\mathbf{e}_2 = \mathbf{e}_3, \ldots, N\mathbf{e}_{n-1} = \mathbf{e}_n$, קל לראות (באינדוקציה על k) כי

$$N^2 = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \ddots & & \vdots \\ 1 & 0 & 0 & \ddots & & \vdots \\ & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & 1 & 0 & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 & 0 \end{pmatrix}, \quad N^3 = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \ddots & & \vdots \\ 0 & 0 & 0 & \ddots & & \vdots \\ 1 & 0 & 0 & \ddots & & \vdots \\ & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & 1 & 0 & 0 & 0 \end{pmatrix},$$

$$N^{n-2} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \ddots & & \vdots \\ 0 & 0 & 0 & \ddots & & \vdots \\ & \ddots & \ddots & \ddots & 0 & 0 \\ 1 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 \end{pmatrix}, \quad N^{n-1} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \ddots & & \vdots \\ 0 & 0 & 0 & \ddots & & \vdots \\ & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & 0 & 0 & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix},$$

ו־$N^k = 0$ לכל $k \geq n$.

כעת נשים לב ש־$J_n(\lambda) = N + \lambda I_n$, באשר N כמו לעיל. כיוון ש־$(\lambda I_n)N = N(\lambda I_n)$, מתקיים

$$J_n(\lambda)^k = \sum_{i=0}^{k} \binom{k}{i} N^i (\lambda I_n)^{k-i} = \sum_{i=0}^{k} \lambda^{k-i} \binom{k}{i} N^i = \sum_{i=0}^{\infty} \lambda^{k-i} \binom{k}{i} N^i =$$

$$= \begin{pmatrix} \lambda^k & 0 & \cdots & 0 & 0 \\ \binom{k}{1}\lambda^{k-1} & \lambda^k & \ddots & & \vdots \\ \binom{k}{2}\lambda^{k-2} & \binom{k}{1}\lambda^{k-1} & \ddots & & \vdots \\ & \ddots & \ddots & 0 & 0 \\ \binom{k}{n-2}\lambda^{k-(n-2)} & & \binom{k}{1}\lambda^{k-1} & \lambda^k & 0 \\ \binom{k}{n-1}\lambda^{k-(n-1)} & \binom{k}{n-2}\lambda^{k-(n-2)} & \cdots & \binom{k}{1}\lambda^{k-1} & \lambda^k \end{pmatrix}$$

(כאשר יש לזכור ש־$\binom{k}{i} = 0$ עבור $i > k$).

כעת, מטריצת ז'ורדן כלשהי J היא מטריצת גושים באלכסון, כאשר כל גוש מצורה $J_n(\lambda)$ עבור איזה $n \geq 0$ ואיזה $\lambda \in F$. לכן J^k היא מטריצת גושים באלכסון, כאשר כל גוש מצורה $J_n(\lambda)^k$.

לבסוף, תהי $A \in \mathrm{M}_n(F)$ ונניח שאנו יודעים למצוא $P \in \mathrm{M}_n(F)$ הפיכה כך ש־$J = P^{-1}AP$ מטריצת ז'ורדן (ראה מסקנה 10.3). אז $A = PJP^{-1}$, ולכן

$$A^k = (PJP^{-1})^k = PJ^kP^{-1}.$$

מכאן, לפי הפסקה הקודמת, נוכל לחשב את A^k, גם עבור k גדול מאד. ∎

מסקנה 10.7: *תהי J מטריצת ז'ורדן, יהי $\lambda \in F$ ויהי $k \geq 0$ שלם.*

(א) $\mathrm{rank} J_n(0)^k = \begin{cases} n-k & k \leq n \\ 0 & k > n \end{cases}$; *אם* $\lambda \neq 0$ *אז* $\mathrm{rank} J_n(\lambda)^k = n$.

(ב) $\mathrm{rank}(J-\lambda I)^{k-1} - \mathrm{rank}(J-\lambda I)^k$ *הוא מספר גושי ז'ורדן ב־J השייכים ל־λ והינם מסדר $\leq k$.*

הוכחה: (א) נובע מדיון 10.6.

(ב) נניח תחילה כי J גוש ז'ורדן, $J = J_n(\mu)$. אז $J - \lambda I = J_n(\mu - \lambda)$ גם גוש ז'ורדן. לפי (א),

$$\mathrm{rank}(J-\lambda I)^{k-1} - \mathrm{rank}(J-\lambda I)^k = \begin{cases} 1 & k \leq n \quad , \mu = \lambda \\ 0 & \text{אחרת} \end{cases} \tag{1}$$

באופן כללי $J = \mathrm{Diag}(J_1, \ldots, J_m)$, באשר $J_1, \ldots, J_m$ גושי ז'ורדן. אז $J - \lambda I = \mathrm{Diag}(J_1 - \lambda I, \ldots, J_m - \lambda I)$. לכן $(J-\lambda I)^k = \mathrm{Diag}((J_1-\lambda I)^k, \ldots, (J_m - \lambda I)^k)$. לפי תרגיל 6.12(ג), $\mathrm{rank}(J-\lambda I)^k = \sum_i \mathrm{rank}(J_i - \lambda I)^k$ ואותה הנוסחה תקפה עם $k-1$ במקום k. לכן

$$\mathrm{rank}(J-\lambda I)^{k-1} - \mathrm{rank}(J-\lambda I)^k = \sum_i (\mathrm{rank}(J_i - \lambda I)^{k-1} - \mathrm{rank}(J_i - \lambda I)^k).$$

מכאן לפי (1) המסקנה. ∎

מסקנה 10.8: *אם $\mathcal{B}$ בסיס של V כך ש־$J = [T]_{\mathcal{B}}^{\mathcal{B}}$ מטריצת ז'ורדן, אז מספר גושי ז'ורדן $J_k(\lambda)$ של J הוא*

$$\mathrm{rank}(T-\lambda 1_V)^{k-1} - 2\mathrm{rank}(T-\lambda 1_V)^k + \mathrm{rank}(T-\lambda 1_V)^{k+1}.$$

בפרט, J יחידה, עד כדי סדר מערכיה (כפי שנטען במשפט 10.2).

הוכחה: לפי מסקנה 10.7(ב), מספר גושי ז'ורדן ב־J מסדר k השייכים ל־λ הוא

$$= \big(\text{מספר גושי ז'ורדן ב־}J\text{ מסדר } \leq k \text{ השייכים ל־}\lambda\big) - \big(\text{מספר גושי ז'ורדן ב־}J\text{ מסדר } \leq k+1 \text{ השייכים ל־}\lambda\big)$$

$$= \big(\mathrm{rank}(J-\lambda I)^{k-1} - \mathrm{rank}(J-\lambda I)^k\big) - \big(\mathrm{rank}(J-\lambda I)^k - \mathrm{rank}(J-\lambda I)^{k+1}\big)$$

$$\mathrm{rank}(J-\lambda I)^{k-1} - 2\mathrm{rank}(J-\lambda I)^k + \mathrm{rank}(J-\lambda I)^{k+1}$$

אך לכל m שלם מתקיים $[(T-\lambda 1_V)^m]_{\mathcal{B}}^{\mathcal{B}} = ([(T-\lambda 1_V]_{\mathcal{B}}^{\mathcal{B}})^m = ([T]_{\mathcal{B}}^{\mathcal{B}} - \lambda [1_V]_{\mathcal{B}}^{\mathcal{B}})^m = (J-\lambda I)^m$, לכן $\mathrm{rank}(J-\lambda I)^m = \mathrm{rank}(T-\lambda 1_V)^m$. ∎

משפט 10.9: *תהי J מטריצת ז'ורדן.*

(א) $m_{J_n(\lambda)} = (X-\lambda)^n$.

(ב) *נניח שמערכי J שייכים ל־F $\lambda_1, \ldots, \lambda_s$ השונים זה מזה והם בעלי אינדקסים $r_1, \ldots, r_s$, בהתאמה. אז* $m_J = (X-\lambda_1)^{r_1} \cdots (X-\lambda_s)^{r_s}$.

(ג) *נניח ש־$m_J = (X-\lambda_1)^{r_1} \cdots (X-\lambda_s)^{r_s}$, באשר $\lambda_1, \ldots, \lambda_s \in F$ השונים זה מזה ו־$r_1, \ldots, r_s \geq 1$. אז מערכי J שייכים ל־$\lambda_1, \ldots, \lambda_s$, והם בעלי אינדקסים $r_1, \ldots, r_s$, בהתאמה.*

הוכחה: (א) מתקיים $f_{J_n(\lambda)} = (X-\lambda)^n$, ולפי תרגיל 7.12, $m_{J_n(\lambda)} = (X-\lambda)^n$.

(ב) נובע מ־(א) לפי מסקנה 9.10.

(ג) נניח שמערכי J שייכים ל־$\bar{\lambda}_1, \dots, \bar{\lambda}_t \in F$ שונים זה מזה והם בעלי אינדקסים $\bar{r}_1, \dots, \bar{r}_t$, בהתאמה. לפי (ב), $m_J = (X - \bar{\lambda}_1)^{\bar{r}_1} \cdots (X - \bar{\lambda}_t)^{\bar{r}_t}$. בגלל יחידות הפירוק של m_J לגורמים אי פריקים מתקיים $t = s$ ו־$\lambda_1 = \bar{\lambda}_1, \dots, \lambda_s = \bar{\lambda}_s$, ו־$r_1 = \bar{r}_1, \dots, r_s = \bar{r}_s$, עד כדי הסדר. ∎

כעת ניגש להוכחת משפט 10.2. תהי $N\colon V \to V$ העתקה לינארית (במקום T).

הגדרה 10.10: סדרה $v_1, \dots, v_\ell \in V \setminus \{0\}$ נקראת **שרשרת ז'ורדן (של N) מאורך ℓ** אם

$$N(v_1) = v_2, N(v_2) = v_3, \dots, N(v_{\ell-1}) = v_\ell, N(v_\ell) = 0.$$

במלים אחרות: הסדרה היא $v_1, N(v_1), N^2(v_1), \dots, N^{\ell-1}(v_1) = \left(N^j(v_1)\right)_{j=0}^{\ell-1}$, ו־$v_1 \in \operatorname{Ker} N^\ell$.

תרגיל 10.11: *יהי $\mathcal{B}$ בסיס של V. אז*

(א) $[N]_{\mathcal{B}}^{\mathcal{B}} = J_\ell(0)$ *אם ורק אם $\mathcal{B}$ שרשרת ז'ורדן מאורך ℓ.*

(ב) $[N]_{\mathcal{B}}^{\mathcal{B}}$ *מערך ז'ורדן השייך ל־0 ובעל אינדקס $r \geq$ אם ורק אם $\mathcal{B}$ צירוף של שרשראות ז'ורדן שסדרת אורכיהם מתחילה ב־r ואינה עולה, כלומר, $\mathcal{B} = \bigvee_{k=1}^{r} \bigvee_{v \in \mathcal{A}_k} \left(N^j(v)\right)_{j=0}^{r-k}$, באשר $\mathcal{A}_k \subseteq \operatorname{Ker} N^{r+1-k}$, לכל k.*

נשים לב: $\left(N^j(v)\right)_{j=0}^{r-k}$ בחלק (ב) היא שרשרת ז'ורדן מאורך $r + 1 - k$.

הוכחה: (א) נובע מההגדרה של $[N]_{\mathcal{B}}^{\mathcal{B}}$. ביתר פירוט,

העמודה הראשונה של $[N]_{\mathcal{B}}^{\mathcal{B}}$ היא e_2 אם ורק אם $N(v_1) = v_2$,

העמודה השנייה של $[N]_{\mathcal{B}}^{\mathcal{B}}$ היא e_3 אם ורק אם $N(v_2) = v_3$, $, \dots, $

העמודה האחרונה של $[N]_{\mathcal{B}}^{\mathcal{B}}$ היא 0 אם ורק אם $N(v_\ell) = 0$.

(ב) לפי משפט 9.9, $[N]_{\mathcal{B}}^{\mathcal{B}} = \operatorname{Diag}(J_{n_1}(0), \dots, J_{n_t}(0))$ אם ורק אם $\mathcal{B}$ היא צירוף של תת סדרות $\mathcal{B}_1, \dots, \mathcal{B}_t$ מאורכים $n_1, \dots, n_t$, בהתאמה, כך שלכל $1 \leq i \leq t$ התת מרחב $W_i = \operatorname{Sp}(\mathcal{B}_i)$ של V הינו שמור־N והמטריצה של הצמצום של $N|_{W_i}$ לפי $\mathcal{B}_i$ היא $J_{n_i}(0)$, כלומר, לפי (א), $\mathcal{B}_i$ היא שרשרת ז'ורדן מאורך n_i. ∎

נשתמש גם בתרגיל הבא:

תרגיל 10.12: *תהי $T\colon V \to W$ העתקה לינארית של מרחבים בעלי מימד סופי. יהי $u_1, \dots, u_r$ בסיס של $\operatorname{Ker} T$ ויהיו $v_1, \dots, v_s \in V$. אז*

(א) $T(v_1), \dots, T(v_s)$ *בלתי תלויים לינארית אם ורק אם $u_1, \dots, u_r, v_1, \dots, v_s$ בלתי תלויים לינארית.*

(ב) $T(v_1), \dots, T(v_s)$ *סדרת יוצרים של $T(V)$ אם ורק אם $u_1, \dots, u_r, v_1, \dots, v_s$ סדרת יוצרים של V.*

(ג) $T(v_1), \dots, T(v_s)$ *בסיס של $T(V)$ אם ורק אם $u_1, \dots, u_r, v_1, \dots, v_s$ בסיס של V, כלומר, $v_1, \dots, v_s$ השלמה של בסיס של $\operatorname{Ker} T$ לבסיס של V.*

נאמר אז ש־$v_1, \dots, v_s$ **בסיס של** $V/\operatorname{Ker} T$. *[קרי: V מודולו $\operatorname{Ker} T$]*

(ד) *נניח* $T(V)=W$ *ותהי* $S\colon W\to U$ *העתקה לינארית. אז* $T^{-1}(\operatorname{Ker} S)=\operatorname{Ker}(S\circ T)$.

הוכחה: (א), (ב), (ג) הם למעשה חלקים של הוכחת משפט המימד להעתקות לינאריות ($\dim\operatorname{Ker} T+\dim\operatorname{Im} T=\dim V$).

(א) נניח ש־$T(v_1),\dots,T(v_s)$ בלתי תלויים לינארית ויהיו $a_1,\dots,a_r,b_1,\dots,b_s\in F$ כך ש־

$$\sum_{i=1}^{r}a_iu_i+\sum_{j=1}^{s}b_jv_j=0. \tag{2}$$

נפעיל T על שני האגפים ונקבל (כיוון ש־T שומרת צירופים לינאריים) $\sum_{i=1}^{r}a_iT(u_i)+\sum_{j=1}^{s}b_jT(v_j)=0$. אבל $T(u_i)=0$ (כי $u_i\in\operatorname{Ker} T$), לכל i, לכן $\sum_{j=1}^{s}b_jT(v_j)=0$. כיוון ש־$T(v_1),\dots,T(v_s)$ בלתי תלויים לינארית, $b_j=0$, לכל j. נציב זאת במשוואה (2) ונקבל $\sum_{i=1}^{r}a_iu_i=0$. כיוון ש־$u_1,\dots,u_r$ בלתי תלויים לינארית, $a_i=0$, לכל i. בכך הוכחנו ש־ $u_1,\dots,u_r,v_1,\dots,v_s$ בלתי תלויים לינארית.

להיפך, נניח ש־$u_1,\dots,u_r,v_1,\dots,v_s$ בלתי תלויים לינארית ויהיו $b_1,\dots,b_s\in F$ כך ש־

$$\sum_{j=1}^{s}b_jT(v_j)=0.$$

כיוון ש־T שומרת צירופים לינאריים, $T\bigl(\sum_{j=1}^{s}b_jv_j\bigr)=0$. כלומר, $\sum_{j=1}^{s}b_jv_j\in\operatorname{Ker} T$. לכן קיימים $a_1,\dots,a_r\in F$ כך ש־ $\sum_{j=1}^{s}b_jv_j=\sum_{i=1}^{r}a_iu_i$. מכאן $\sum_{i=1}^{r}(-a_i)u_i+\sum_{j=1}^{s}b_jv_j=0$. כיוון ש־$u_1,\dots,u_r,v_1,\dots,v_s$ בלתי תלויים לינארית, $b_j=0$, לכל j (וגם $a_i=0$, לכל i). בכך הוכחנו ש־$T(v_1),\dots,T(v_s)$ בלתי תלויים לינארית.

(ב) נניח ש־$T(v_1),\dots,T(v_s)$ סדרת יוצרים של $T(V)$ ויהי $v\in V$. אז $T(v)\in T(V)$, ולכן יש $b_1,\dots,b_s\in F$ כך ש־

$$T(v)=\sum_{j=1}^{s}b_jT(v_j).$$

אז $T(v-\sum_{j=1}^{s}b_jv_j)=T(v)-\sum_{j=1}^{s}b_jT(v_j)=0$, לכן $v-\sum_{j=1}^{s}b_jv_j\in\operatorname{Ker} T$. לכן קיימים $a_1,\dots,a_r\in F$ כך ש־ $v-\sum_{j=1}^{s}b_jv_j=\sum_{i=1}^{r}a_iu_i$. מכאן $v=\sum_{i=1}^{r}a_iu_i+\sum_{j=1}^{s}b_jv_j$. בכך הוכחנו ש־ $u_1,\dots,u_r,v_1,\dots,v_s$ סדרת יוצרים של V.

להיפך, נניח ש־$u_1,\dots,u_r,v_1,\dots,v_s$ סדרת יוצרים של V ויהי $w\in T(V)$. אז קיים $v\in V$ כך ש־$w=T(v)$. לפי ההנחה קיימים $a_1,\dots,a_r,b_1,\dots,b_s\in F$ כך ש־

$$v=\sum_{i=1}^{r}a_iu_i+\sum_{j=1}^{s}b_jv_j. \tag{3}$$

נפעיל T על שני האגפים ונקבל $w=\sum_{i=1}^{r}a_iT(u_i)+\sum_{j=1}^{s}b_jT(v_j)$. מכאן, כיוון ש־$T(u_i)=0$ לכל i, נקבל $w=\sum_{j=1}^{s}b_jT(v_j)$. בכך הוכחנו ש־$T(v_1),\dots,T(v_s)$ סדרת יוצרים של $T(V)$.

(ג) נובע משני החלקים קודמים, כי בסיס היא סדרת יוצים שהינה בלתי תלויה לינארית.

(ד) זוהי טענה בתורת הקבוצות. נעיר ש־T^{-1} אינו ההעתקה ההופכית של T (שאולי אפילו איננה קיימת) אלא הסימון לתמונה ההפוכה: $T^{-1}(Y) := \{v \in V |\ T(v) \in Y\}$.

יהי $v \in V$. אז

$$v \in T^{-1}(\operatorname{Ker} S) \Leftrightarrow T(v) \in \operatorname{Ker} S \Leftrightarrow S(T(v)) = 0 \Leftrightarrow (S \circ T)(v) = 0 \Leftrightarrow v \in \operatorname{Ker}(S \circ T).$$

∎

הערה: בשביל אלה שלומדים\למדו אלגברה ב1. V הוא חבורה חילופית (ביחס לחיבור) ו־$\operatorname{Ker} T$ תת חבורה (נורמלית) שלה. חבורת המנה $V/\operatorname{Ker} T$ היא גם מרחב וקטורי מעל F, כאשר הכפל בסקלר בו מושרה מהכפל בסקלר ב־V (בדקו!).

לא קשה לראות ש־$v_1, \ldots, v_s \in V$ בסיס של $V/\operatorname{Ker} T$ כפי שזה הוגדר בתרגיל לעיל אם ורק אם התמונות של $v_1, \ldots, v_s \in V$ ב־$V/\operatorname{Ker} T$ בסיס של $V/\operatorname{Ker} T$ במובן הרגיל של בסיס. ∎

משפט 10.13: *נניח $N^r = 0$ באשר $r \geq 1$. אז קיים בסיס $\mathcal{B}$ של V כך ש־$[N]_{\mathcal{B}}^{\mathcal{B}}$ מערך ז'ורדן השייך ל־0 ובעל אינדקס $r \geq$.*

הוכחה: באינדוקציה על r: אם $r = 1$ אז $N = 0$ ולכל בסיס $\mathcal{B}$ המטריצה $[N]_{\mathcal{B}}^{\mathcal{B}} = 0$ היא אכן מערך ז'ורדן השייך ל־0 ובעל אינדקס 1. נניח כי $r \geq 2$ והמשפט נכון עבור $r - 1$.

יהי $V' = N(V) = \operatorname{Im} N$. זהו תת מרחב שמור־$N$ של V (תרגיל 9.3(א)); יהי $N' = N|_{V'}$. אז $(N')^{r-1} = 0$. לפי הנחת האינדוקציה ותרגיל 10.11(ב) יש ל־V' בסיס שהוא צירוף של שרשראות ז'ורדן

$$\mathcal{B}' = \bigvee_{k=1}^{r-1} \bigvee_{v' \in \mathcal{A}'_k} \left(N^j(v')\right)_{j=0}^{r-1-k},$$

באשר $\mathcal{A}'_k \subseteq \operatorname{Ker}(N')^{r-k}$ לכל k (קבוצת הרישות של השרשראות). כיוון ש־$V' = N(V)$, לכל k יש סדרה $\mathcal{A}_k \subseteq V$ כך ש־$N(\mathcal{A}_k) = \mathcal{A}'_k$. אז $\mathcal{A}_k \subseteq \operatorname{Ker} N^{r+1-k}$. התת סדרה הבאה של $\mathcal{B}'$ (של הסיפות של שרשראות ז'ורדן)

$$\bigvee_{k=1}^{r-1} \bigvee_{v' \in \mathcal{A}'_k} \left(N^{r-1-k}(v')\right) = \bigvee_{k=1}^{r-1} \bigvee_{v \in \mathcal{A}_k} \left(N^{r-k}(v)\right) = N^{r-1}(\mathcal{A}_1) \vee N^{r-2}(\mathcal{A}_2) \vee \ldots \vee N(\mathcal{A}_{r-1})$$

היא בלתי תלויה לינארית ומוכלת ב־$\operatorname{Ker} N$. נשלים אותה לבסיס $\mathcal{B}_1$ של $\operatorname{Ker} N$ על ידי איזו סדרה $\mathcal{A}_r$. אז $\mathcal{B}_1 = \bigvee_{k=1}^{r} \bigvee_{v \in \mathcal{A}_k} \left(N^{r-k}(v)\right)$ בסיס של $\operatorname{Ker} N$ (כי $\mathcal{A}_r = \bigvee_{v \in \mathcal{A}_r} \left(N^0(v)\right)$).

טענה: $\mathcal{B} := \bigvee_{k=1}^{r} \bigvee_{v\in\mathcal{A}_k} \left(N^j(v)\right)_{j=0}^{r-k}$ *היא בסיס של* V.

אכן, עד כדי הסדר, $\mathcal{B} = \mathcal{B}_1 \vee \mathcal{B}_2$, באשר $\mathcal{B}_2 = \bigvee_{k=1}^{r-1} \bigvee_{v\in\mathcal{A}_k} \left(N^j(v)\right)_{j=0}^{r-1-k}$.

אז $N(\mathcal{B}_2) = \mathcal{B}'$ בסיס של V' ו־$\mathcal{B}_1$ בסיס של $\operatorname{Ker} N$. לפי תרגיל 10.12(ב), $\mathcal{B}$ בסיס של V. לפי תרגיל 10.11(ב), $[N]_{\mathcal{B}}^{\mathcal{B}}$ היא מערך ז'ורדן השייך ל־0 ובעל אינדקס $\geq r$. ∎

בניה 10.14: נרצה להפוך את ההוכחה באינדוקציה לבניה מעשית של בסיס. נניח $m_N = X^r$. נשתמש בסדרה של העתקות (צמצומים של N)

$$V = N^0(V) \to N(V) \to \cdots \to N^{r-3}(V) \to N^{r-2}(V) \to N^{r-1}(V) \to N^r(V) = \{0\}.$$

נשנה את הסימון: במקום $\mathcal{A}_k \subseteq N^{r-k}(V)$ נכתוב $N^{r-k}(\mathcal{A}_k)$ עבור איזה $\mathcal{A}_k \subseteq V$. אז ההוכחה היא כדלקמן:

1. נמצא $\mathcal{A}_1 \subseteq V$ כזו ש־$N^{r-1}(\mathcal{A}_1)$ בסיס של $N^{r-1}(V)$. לפי תרגיל 10.12(ב) זה שקול ל:
   - $\mathcal{A}_1$ בסיס של $V/\operatorname{Ker} N^{r-1}$.
2. יהי $N' = N|_{N^{r-2}(V)}$. לפי תרגיל 10.12(ג), $(N^{r-2})^{-1}(\operatorname{Ker} N') = \operatorname{Ker} N^{r-1}$. נשלים את

$$N^{r-1}(\mathcal{A}_1) = N^{r-2}\big(N(\mathcal{A}_1)\big)$$

על ידי איזה $N^{r-2}(\mathcal{A}_2)$ לבסיס של $\operatorname{Ker} N'$. לפי תרגיל 10.12(ב) (עבור $N^{r-2}\colon \operatorname{Ker} N^{r-1} \to \operatorname{Ker} N'$) זה שקול ל:

   - $N(\mathcal{A}_1) \vee \mathcal{A}_2$ בסיס של $\operatorname{Ker} N^{r-1}/\operatorname{Ker} N^{r-2}$.
3. יהי $N' = N|_{N^{r-3}(V)}$. לפי תרגיל 10.12(ג), $(N^{r-3})^{-1}(\operatorname{Ker} N') = \operatorname{Ker} N^{r-2}$. נשלים את

$$N^{r-1}(\mathcal{A}_1) \vee N^{r-2}(\mathcal{A}_2) = N^{r-3}\big(N^2(\mathcal{A}_1) \vee N(\mathcal{A}_2)\big)$$

על ידי איזה $N^{r-3}(\mathcal{A}_3)$ לבסיס של $\operatorname{Ker} N'$. לפי תרגיל 10.12(ב) (עבור $N^{r-3}\colon \operatorname{Ker} N^{r-2} \to \operatorname{Ker} N'$) זה שקול ל:

   - $N^2(\mathcal{A}_1) \vee N(\mathcal{A}_2) \vee \mathcal{A}_3$ בסיס של $\operatorname{Ker} N^{r-2}/\operatorname{Ker} N^{r-3}$.

וכך הלאה. לבסוף, $\mathcal{B} = \bigvee_{k=1}^{r} \bigvee_{v\in\mathcal{A}_k} \left(N^j(v)\right)_{j=0}^{r-k}$ הבסיס המבוקש של V.

הוכחת משפט 10.2: נניח $m_T = (X-\lambda_1)^{r_1}\cdots(X-\lambda_s)^{r_s}$, באשר $\lambda_1,\ldots,\lambda_s \in F$ שונים זה מזה.

קיום J: לפי משפט הפירוק הפרימרי (משפט 9.14), $V = \oplus_{k=1}^{s} W_k$, באשר $W_k = \operatorname{Ker}(T-\lambda_k 1_V)^{r_k}$ ו־$m_{T|_{W_k}} = (X-\lambda)^{r_k}$ לכל k. יהי $1 \leq k \leq s$. נסמן $N = T|_{W_k} - \lambda_k 1_{W_k}\colon W_k \to W_k$. אז $m_N = X^{r_k}$. לפי משפט 10.13 יש בסיס $\mathcal{B}_k$ של W_k כך ש־$[N]_{\mathcal{B}_k}^{\mathcal{B}_k} = J(0)$ מערך ז'ורדן השייך ל־0. אז

$$[T|_{W_k}]_{\mathcal{B}_k}^{\mathcal{B}_k} = [N]_{\mathcal{B}_k}^{\mathcal{B}_k} + \lambda_k [1_{W_k}]_{\mathcal{B}_k}^{\mathcal{B}_k} = J(0) + \lambda_k I$$

הוא מערך ז'ורדן השייך ל־λ_k.

נסמן $\mathcal{B} = \bigvee_{k=1}^{s} \mathcal{B}_k$. אז לפי משפט 9.9, $[T]_{\mathcal{B}}^{\mathcal{B}}$ מטריצת ז'ורדן.

יחידות: יהי $\mathcal{B}$ בסיס כך ש־$J := [T]_{\mathcal{B}}^{\mathcal{B}}$ מטריצת ז'ורדן. יהי $\lambda \in F$. לפי מסקנה 10.8, מספר גושי ז'ורדן $J_k(\lambda)$ ב־J תלוי רק ב־T וב־λ, לא ב־$\mathcal{B}$.

לבסוף, כיוון ש־$m_T = m_J$, לפי משפט 10.9(ג) מערכי J שייכים ל־$\lambda_1, \ldots, \lambda_s$ והם מאינדקס $r_1, \ldots, r_s$, בהתאמה, עד כדי הסדר. ∎

תרגיל 10.15: *תהי J צורת ז'ורדן של T. אז מספר גושי ז'ורדן ב־J השייכים ל־λ הוא הריבוי הגיאומטרי של λ ב־T.*

הוכחה: נניח $J = \mathrm{Diag}(J_{n_1}(\lambda_1), \ldots, J_{n_s}(\lambda_s)) \in \mathrm{M}_n(F)$ (באשר $\lambda_1, \ldots, \lambda_s$ לא בהכרח שונים זה מזה). אז הריבוי הגיאומטרי של λ ב־T הוא

$$\dim \mathrm{Ker}(T - \lambda 1_V) = \dim \mathrm{Ker}(J - \lambda I) = n - \mathrm{rank}(J - \lambda I) =$$
$$\sum_{i=1}^{s} n_i - \sum_{i=1}^{s} \mathrm{rank} J_{n_i}(\lambda_i - \lambda) = \sum_{i=1}^{s} \left(n_i - \mathrm{rank} J_{n_i}(\lambda_i - \lambda)\right)$$

ולפי מסקנה 10.7(א), עבור $k = 1$, $n_i - \mathrm{rank} J_{n_i}(\lambda_i - \lambda) = \begin{cases} 1 & \lambda = \lambda_i \\ 0 & \text{אחרת} \end{cases}$. מכאן התוצאה. ∎

הערה 10.16: דרך אחרת לראות זאת: תהי $N = T - \lambda 1_V$. אז $\mathcal{B}_1 = \bigvee_{k=1}^{r} \bigvee_{v \in \mathcal{A}_k} \left(N^{r-k}(v)\right)$ בסיס של $V_\lambda = \mathrm{Ker}\, N$. לכן $\dim V_\lambda = \sum_{k=1}^{r} |\mathcal{A}_k|$ הוא מספר גושי ז'ורדן ששייכים ל־λ. ∎

דוגמה 10.16: נתונה מטריצה $M \in \mathrm{M}_{16}(\mathbb{R})$. נחשב מטריצה הפיכה $P \in \mathrm{M}_{16}(\mathbb{R})$ ומטריצת ז'ורדן $J \in \mathrm{M}_{16}(\mathbb{R})$ כך ש־$P^{-1}MP = J$.

המטריצה בדוגמה זאת מסדר גדול (16×16), כדי שאפשר יהיה להדגים כמה תופעות בחישוב. מאידך גיסא, הדוגמה בנאלית מבחינה מספרית: כאשר נדרשים להשלים בסיס של תת מרחב אחד לבסיס של תת מרחב אחר שמכיל אותו, בדוגמה הזאת שני התת מרחבים נפרשים על ידי בסיסים כאלה שהבסיס של התת מרחב הגדול יותר מכיל כבר את הבסיס של התת מרחב המוכל בו. לכן ההשלמה המבוקשת מורכבת מאותם האיברים של הבסיס הראשון שאינם כלולים בבסיס של התת מרחב הכלול. (ובנוסף לכך, הבסיסים הם חלק מהבסיס הסטנדרטי של $\mathbb{R}^{16}$.) לא תמיד יהיה לנו המזל למצוא בסיסי כאלה מלכתחילה.

תהי

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

(הרכיבים הלא רשומים הם אפסים.)

בגלל ש־M משולשית עליונה, הפולינום המזערי שלה הוא מכפלה של גורמים לינאריים (ולכן יש ל־M צורת ז'ורדן) והערכים העצמיים של M האיברים באלכסון הראשי, כלומר, $3, 2$.

(א) ערך עצמי 3: נסמן $N = M - 3I$.

$$N := M - 3I = \begin{pmatrix} 0 & 1 & 0 & 0 & & & & & & & & & & & \\ 0 & 0 & 1 & 0 & & & & & & & & & & & \\ 0 & 0 & 0 & 1 & & & & & & & & & & & \\ 0 & 0 & 0 & 0 & & & & & & & & & & & \\ & & & & 0 & 1 & 0 & 0 & & & & & & & \\ & & & & 0 & 0 & 1 & 0 & & & & & & & \\ & & & & 0 & 0 & 0 & 1 & & & & & & & \\ & & & & 0 & 0 & 0 & 0 & & & & & & & \\ & & & & & & & & 0 & 1 & & & & & \\ & & & & & & & & 0 & 0 & & & & & \\ & & & & & & & & & & 0 & 1 & & & \\ & & & & & & & & & & 0 & 0 & & & \\ & & & & & & & & & & & & 0 & & \\ & & & & & & & & & & & & & -1 & 1 \\ & & & & & & & & & & & & & 0 & -1 \\ & & & & & & & & & & & & & & & -1 \end{pmatrix}$$

$\operatorname{Ker} N = \operatorname{Sp}(e_1, e_5, e_9, e_{11}, e_{13})$

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

$$\operatorname{Ker} N^2 = \operatorname{Sp}(e_1, e_2, e_5, e_6, e_9, e_{10}, e_{11}, e_{12}, e_{13})$$

$$.N^3 = \begin{pmatrix} 0 & 0 & 0 & 1 & & & & & & & & & \\ 0 & 0 & 0 & 0 & & & & & & & & & \\ 0 & 0 & 0 & 0 & & & & & & & & & \\ 0 & 0 & 0 & 0 & & & & & & & & & \\ & & & & 0 & 0 & 0 & 1 & & & & & \\ & & & & 0 & 0 & 0 & 0 & & & & & \\ & & & & 0 & 0 & 0 & 0 & & & & & \\ & & & & 0 & 0 & 0 & 0 & & & & & \\ & & & & & & & & 0 & 0 & & & \\ & & & & & & & & 0 & 0 & & & \\ & & & & & & & & & & 0 & 0 & \\ & & & & & & & & & & 0 & 0 & \\ & & & & & & & & & & & & 0 \\ & & & & & & & & & & & -1 & 3 \\ & & & & & & & & & & & 0 & -1 \\ & & & & & & & & & & & & & -1 \end{pmatrix}$$

$$\operatorname{Ker} N^3 = \operatorname{Sp}(e_1, e_2, e_3, e_5, e_6, e_7, e_9, e_{10}, e_{11}, e_{12}, e_{13})$$

$$N^4 = \begin{pmatrix} 0 & 0 & 0 & 0 & & & & & & & & & \\ 0 & 0 & 0 & 0 & & & & & & & & & \\ 0 & 0 & 0 & 0 & & & & & & & & & \\ 0 & 0 & 0 & 0 & & & & & & & & & \\ & & & & 0 & 0 & 0 & 0 & & & & & \\ & & & & 0 & 0 & 0 & 0 & & & & & \\ & & & & 0 & 0 & 0 & 0 & & & & & \\ & & & & 0 & 0 & 0 & 0 & & & & & \\ & & & & & & & & 0 & 0 & & & \\ & & & & & & & & 0 & 0 & & & \\ & & & & & & & & & & 0 & 0 & \\ & & & & & & & & & & 0 & 0 & \\ & & & & & & & & & & & & 0 \\ & & & & & & & & & & & 1 & -4 \\ & & & & & & & & & & & 0 & 1 \\ & & & & & & & & & & & & & 1 \end{pmatrix}.$$

$$\operatorname{Ker} N^4 = \operatorname{Sp}(e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8, e_9, e_{10}, e_{11}, e_{12}, e_{13})$$

$r = 4$, $N = M - 3I$,

- נבחר $\mathcal{A}_1 = (e_4, e_8)$ בסיס של $\operatorname{Ker} N^4 / \operatorname{Ker} N^3$.

  נשלים את $N(\mathcal{A}_1) = (e_3, e_7)$ על ידי $\mathcal{A}_2 = \emptyset$ לבסיס של $\operatorname{Ker} N^3 / \operatorname{Ker} N^2$:

- $N(\mathcal{A}_1) \vee \mathcal{A}_2 = (e_3, e_7)$ בסיס של $\operatorname{Ker} N^3 / \operatorname{Ker} N^2$.

  נשלים את $N^2(\mathcal{A}_1) \vee N(\mathcal{A}_2) = (e_2, e_6)$ על ידי $\mathcal{A}_3 = (e_{10}, e_{12})$ לבסיס של $\operatorname{Ker} N^2 / \operatorname{Ker} N^1$:

- $N^2(\mathcal{A}_1) \vee N(\mathcal{A}_2) \vee \mathcal{A}_3 = (e_2, e_6, e_{10}, e_{12})$ בסיס של $\operatorname{Ker} N^2 / \operatorname{Ker} N^1$.

  נשלים את $N^3(\mathcal{A}_1) \vee N^2(\mathcal{A}_2) \vee N(\mathcal{A}_3) = (e_1, e_5, e_9, e_{11})$ על ידי $\mathcal{A}_4 = (e_{13})$ לבסיס של $\operatorname{Ker} N^1 / \operatorname{Ker} N^0$ (כלומר, בסיס של $\operatorname{Ker} N$):

- $N^3(\mathcal{A}_1) \vee N^2(\mathcal{A}_2) \vee N(\mathcal{A}_3) \vee \mathcal{A}_4 = (e_1, e_5, e_9, e_{11}, e_{13})$ בסיס של $\operatorname{Ker} N^1 / \operatorname{Ker} N^0$ (כלומר, של $\operatorname{Ker} N$).

מכאן נקבל בסיס של $\operatorname{Ker} N^4$:

$$\mathcal{B}_3 = \bigvee_{k=1}^{r} \bigvee_{v \in \mathcal{A}_k} \left(N^j(v)\right)_{j=0}^{r-k} = \left(e_4, N(e_4), N^2(e_4), N^3(e_4)\right) \vee \left(e_8, N(e_8), N^2(e_8), N^3(e_8)\right) \vee$$
$$\vee \left(e_{10}, N(e_{10})\right) \vee \left(e_{12}, N(e_{12})\right) \vee \left(e_{13}\right) =$$
$$= (e_4, e_3, e_2, e_1, e_8, e_7, e_6, e_5, e_{10}, e_9, e_{12}, e_{11}, e_{13})$$

בצורת ז'ורדן של M יהיו

$|\mathcal{A}_1| = 2$ גושי ז'ורדן $J_4(3)$;

$|\mathcal{A}_2| = 0$ גושי ז'ורדן $J_4(3)$;

$|\mathcal{A}_3| = 2$ גושי ז'ורדן $J_2(3)$;

$|\mathcal{A}_4| = 1$ גושי ז'ורדן $J_1(3)$.

(ב) ערך עצמי 2: נסמן $L = M - 2I$.

$$L := \begin{pmatrix}
1 & 1 & 0 & 0 & & & & & & & & & & & & \\
0 & 1 & 1 & 0 & & & & & & & & & & & & \\
0 & 0 & 1 & 1 & & & & & & & & & & & & \\
0 & 0 & 0 & 1 & & & & & & & & & & & & \\
 & & & & 1 & 1 & 0 & 0 & & & & & & & & \\
 & & & & 0 & 1 & 1 & 0 & & & & & & & & \\
 & & & & 0 & 0 & 1 & 1 & & & & & & & & \\
 & & & & 0 & 0 & 0 & 1 & & & & & & & & \\
 & & & & & & & & 1 & 1 & & & & & & \\
 & & & & & & & & 0 & 1 & & & & & & \\
 & & & & & & & & & & 1 & 1 & & & & \\
 & & & & & & & & & & 0 & 1 & & & & \\
 & & & & & & & & & & & & 1 & & & \\
 & & & & & & & & & & & & & 0 & 1 & \\
 & & & & & & & & & & & & & 0 & 0 & \\
 & & & & & & & & & & & & & & & 0
\end{pmatrix}$$

$\operatorname{Ker} L = \operatorname{Sp}(e_{14}, e_{16})$

$$L^2 := \begin{pmatrix}
1 & 2 & 1 & 0 & & & & & & & & & & & & \\
0 & 1 & 2 & 1 & & & & & & & & & & & & \\
0 & 0 & 1 & 2 & & & & & & & & & & & & \\
0 & 0 & 0 & 1 & & & & & & & & & & & & \\
 & & & & 1 & 2 & 1 & 0 & & & & & & & & \\
 & & & & 0 & 1 & 2 & 1 & & & & & & & & \\
 & & & & 0 & 0 & 1 & 2 & & & & & & & & \\
 & & & & 0 & 0 & 0 & 1 & & & & & & & & \\
 & & & & & & & & 1 & 2 & & & & & & \\
 & & & & & & & & 0 & 1 & & & & & & \\
 & & & & & & & & & & 1 & 2 & & & & \\
 & & & & & & & & & & 0 & 1 & & & & \\
 & & & & & & & & & & & & 1 & & & \\
 & & & & & & & & & & & & & 0 & 0 & \\
 & & & & & & & & & & & & & 0 & 0 & \\
 & & & & & & & & & & & & & & & 0
\end{pmatrix}$$

$\operatorname{Ker} L^2 = \operatorname{Sp}(e_{14}, e_{15}, e_{16})$

$r = 2$, $L = M - 2I$,

- נבחר $\mathcal{C}_1 = (e_{15})$ בסיס של $\operatorname{Ker} L^2 / \operatorname{Ker} L^1$.

נשלים את $L(\mathcal{C}_1) = (e_{14})$ על ידי $\mathcal{C}_2 = (e_{16})$ לבסיס של $\operatorname{Ker} L^1 / \operatorname{Ker} L^0$ (כלומר, בסיס של $\operatorname{Ker} L$):

- $L(\mathcal{C}_1) \vee \mathcal{C}_2 = (e_{14}, e_{16})$ בסיס של $\operatorname{Ker} L$.

מכאן נקבל בסיס של $\operatorname{Ker} L^2$:

$$\mathcal{B}_2 = \bigvee_{k=1}^{r} \bigvee_{v \in \mathcal{C}_k} \big(L^j(v)\big)_{j=0}^{r-k} = \big(e_{15}, L(e_{15})\big) \vee \big(e_{16}\big) = (e_{15}, e_{14}, e_{16})$$

בצורת ז'ורדן J של M יהיו

$|\mathcal{C}_1| = 1$ גושי ז'ורדן $J_2(2)$;

$|\mathcal{C}_2| = 1$ גושי ז'ורדן $J_1(2)$.

לכן

$$J = \operatorname{Diag}\big(J_4(3), J_4(3), J_2(3), J_2(3), J_1(3), J_2(2), J_1(2)\big) =$$

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

העמודות של מטריצה P שמקימת $P^{-1}MP = J$ הן הצירוף של שני הבסיסים שמצאנו, $\mathcal{B}_3 \vee \mathcal{B}_2$. לכן

$$P = \left(e_4, e_3, e_2, e_1, e_8, e_7, e_6, e_5, e_{10}, e_9, e_{12}, e_{11}, e_{13}, e_{15}, e_{14}, e_{16}\right) =$$

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

∎

11. הצורה הרציונלית (צורת יעקובסון)

בסעיף זה יהי V מרחב וקטורי נוצר סופית מעל שדה F ותהי $T\colon V \to V$ העתקה לינארית.

הגדרה 11.1: יהי $q = a_0 + a_1X + \cdots + a_{d-1}X^{d-1} + X^d \in F[X]$ מתוקן ממעלה $d \geq 1$.

(א) **המטריצה הנִלוֵית ל־q** היא המטריצה הבאה

$$C_1(q) = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & & & \\ \cdot & \cdot & \ddots & \ddots & \cdot & \vdots \\ & & & 1 & 0 & -a_{d-2} \\ 0 & 0 & 0 & \cdots & 1 & -a_{d-1} \end{pmatrix} \in \mathrm{M}_d(F).$$

בפרט, $C_1(a_0 + a_1X + X^2) = \begin{pmatrix} 0 & -a_0 \\ 1 & -a_1 \end{pmatrix}$, $C_1(X - \lambda) = (\lambda) \in \mathrm{M}_1(F)$.

(ב) יהי $r \in \mathbb{N}$. המטריצה הבאה מסדר $dr \times dr$ תיקרא **גוש יעקובסון מאורך r השייך ל־q**

$$\begin{pmatrix} 0 & \cdots & 0 & -a_0 & & & & & & & & & & & \\ 1 & \ddots & \vdots & \vdots & & & & & & & & & \mathbf{0} & & \\ & \ddots & 0 & \vdots & & & & & & & & & & & \\ 0 & \cdots & 1 & -a_{d-1} & & & & & & & & & & & \\ & & & 1 & 0 & \cdots & 0 & -a_0 & & & & & & & \\ & & & & 1 & \ddots & \vdots & \vdots & & & & & & & \\ & & & & & \ddots & 0 & \vdots & & & & & & & \\ & & & & 0 & \cdots & 1 & -a_{d-1} & & & & & & & \\ & & & & & & & 1 & & & & & & & \\ & & & & & & & & \ddots & & & & & & \\ & & & & & & & & & 1 & 0 & \cdots & 0 & -a_0 & \\ & & & & & & & & & & 1 & \ddots & \vdots & \vdots & \\ & & \mathbf{0} & & & & & & & & & \ddots & 0 & \vdots & \\ & & & & & & & & & & 0 & \cdots & 1 & -a_{d-1} & \end{pmatrix}$$

היא תסומן $C_r(q)$. אפשר לכתוב אותה כמטריצת גושים כך:

$$C_r(q) = \begin{pmatrix} C_1(q) & 0 & \cdots & 0 & 0 \\ M & C_1(q) & \ddots & & \vdots \\ & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & M & C_1(q) & 0 \\ 0 & 0 & \cdots & M & C_1(q) \end{pmatrix} \in \mathrm{M}_{dr}(F),$$

באשר

$$M = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \ddots & & \vdots \\ & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \in \mathrm{M}_d(F).$$

שים לב ש־$C_r(X-\lambda)$ היא $J_r(\lambda)$, גוש ז'ורדן מסדר r (הגדרה 10.1).

בדומה למטריצת ז'ורדן נגדיר מערכי יעקובסון ומטריצות יעקובסון:

(ג) מטריצה מהצורה $C(q) = \begin{pmatrix} C_{n_1}(q) & & 0 \\ & \ddots & \\ 0 & & C_{n_t}(q) \end{pmatrix}$, באשר $n_1 \geq n_2 \geq \cdots \geq n_t \geq 1$, נקראת **מערך יעקובסון השייך ל־q**. המספר n_1 נקרא **האינדקס** של המערך.

(ד) **מטריצת יעקובסון** היא מטריצה מהצורה $C = \begin{pmatrix} C(q_1) & & 0 \\ & \ddots & \\ 0 & & C(q_s) \end{pmatrix}$, באשר $q_1, \ldots, q_s \in F[X]$ אי פריקים מתוקנים שונים זה מזה ולכל $1 \leq i \leq s$ המטריצה $C(q_i)$ היא מערך יעקובסון השייך ל־q_i. ∎

המשפט המרכזי של הפרק, אותו נוכיח בהמשך, הוא:

משפט 11.2: *ל־V יש בסיס $\mathcal{B}$ כך ש־$[T]_{\mathcal{B}}^{\mathcal{B}}$ מטריצת יעקובסון. היא יחידה, עד כדי סדר המערכים שלה. (מטריצה זו נקראת* **צורת יעקובסון** *או* **הצורה הרציונלית של** *T.) אם $m_T = q_1^{r_1} \cdots q_s^{r_s}$, באשר $q_1, \ldots, q_s \in F$ אי פריקים מתוקנים זרים זה לזה ו־$r_1, \ldots, r_s \geq 1$, אז במטריצה זו יש s מערכי יעקובסון, כאשר, עד כדי הסדר של המערכים, המערך ה־i־י שייך ל־q_i והוא בעל אינדקס r_i.*

ממנו נובעת מיידית:

מסקנה 11.3: *תהי $A \in \mathrm{M}_n(F)$. אז קיימת מטריצת יעקובסון דומה ל־A יחידה, עד כדי סדר המערכים שלה. (מטריצה זו נקראת* **צורת יעקובסון** *או* **הצורה הרציונלית של** *A.) אם $m_A = q_1^{r_1} \cdots q_s^{r_s}$, באשר $q_1, \ldots, q_s \in F$ אי פריקים מתוקנים זרים זה לזה ו־$r_1, \ldots, r_s \geq 1$, אז במטריצה זו יש s מערכי יעקובסון, כאשר,עד כדי הסדר של המערכים, המערך ה־i־י שייך ל־q_i והוא בעל אינדקס r_i.*

תרגיל 11.4: *יהי $q = a_0 + a_1X + \cdots + a_{d-1}X^{d-1} + X^d \in F[X]$ ותהי $C = C_r(q) \in \mathrm{M}_{dr}(F)$. אז $q^r = m_C = f_C$.*

הוכחה: תחילה נניח כי $r = 1$ ונחשב את f_C. לפי ההגדרה,

$$f_C = \det(XI_d - C) = \det \begin{pmatrix} X & 0 & 0 & \cdots & 0 & a_0 \\ -1 & X & 0 & \cdots & 0 & a_1 \\ 0 & -1 & X & & & \\ \cdot & \cdot & \ddots & \ddots & \cdot & \vdots \\ & & & -1 & X & a_{d-2} \\ 0 & 0 & 0 & \cdots & -1 & X + a_{d-1} \end{pmatrix}.$$

נוסיף את השורה האחרונה X פעמים לשורה הקודמת, אח"כ נוסיף את השורה שלפני האחרונה X פעמים לשורה הקודמת, וכן הלאה, עד שנוסיף את השורה השניה X פעמים לראשונה. אז

$$f_C = \det \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & q(X) \\ -1 & 0 & 0 & \cdots & 0 & X^{d-1} + a_{d-1}X^{d-2} + \cdots + a_1 \\ 0 & -1 & 0 & & & \\ \cdot & \cdot & \ddots & \ddots & \cdot & \vdots \\ & & & -1 & 0 & X^2 + a_{d-1}X + a_{d-2} \\ 0 & 0 & 0 & \cdots & -1 & X + a_{d-1} \end{pmatrix}.$$

פיתוח לפי השורה הראשונה נותן

$$f_C = (-1)^{d+1} q(X) \det(-I_{d-1}) = (-1)^{d+1} q(X)(-1)^{d-1} = q(X).$$

כעת, אם r כלשהו, אז $f_C = q^r$ לפי תרגיל 6.12(ב). לפי תרגיל 7.12, $m_C = f_C$. ∎

מסקנה 11.5: *תהי C מטריצת יעקובסון. יהיו $C_1, \ldots, C_s$ המערכים שלה, באשר C_i שייך ל־ q_i מאינדקס r_i, לכל i. אז $m_C = q_1^{r_1} \cdots q_s^{r_s}$.*

(לפי הגדרה 11.1(ד), $q_1, \ldots, q_s$ שונים.)

הוכחה: נקבע $1 \le i \le s$. אז $C_i = \mathrm{Diag}(C_{n_1}(q), \ldots, C_{n_t}(q))$ באשר $r_i = n_1 \ge \cdots \ge n_t$. לפי תרגיל 11.4, $m_{C_{n_j}} = q_i^{n_j}$, לכל j. לכן לפי משפט 7.10, $m_{C_i} = \mathrm{lcm}(m_{C_{n_1}(q)}, \ldots, m_{C_{n_t}(q)}) = \mathrm{lcm}(q^{n_1}, \ldots, q^{n_t}) = q^{r_i}$. לפי משפט 7.10, $m_C = \mathrm{lcm}(m_{C_1}, \ldots, m_{C_s}) = q_1^{r_1} \cdots q_s^{r_s}$. ∎

סימון 11.6: מעתה יהי $q = \sum_{i=0}^{d-1} a_i X^i + X^d \in F[X]$.

עבור $v \in V$ נסמן $\widetilde{v} = \big(v, T(v), \ldots, T^{d-1}(v)\big)$ ועבור סדרה $\mathcal{A} = (v_1, \ldots, v_k)$ של וקטורים ב־V נסמן $\widetilde{\mathcal{A}} = \big(v_1, \ldots, v_k, T(v_1), \ldots, T(v_k)\ ,\ \ldots,\ T^{d-1}(v_1), \ldots, T^{d-1}v_k\big)$. אז $|\widetilde{\mathcal{A}}| = d|\mathcal{A}|$.

אם $\mathcal{A}, \mathcal{B}$ שתי סדרות, אז הסדרות $\widetilde{\mathcal{A}} \vee \widetilde{\mathcal{B}}$ ו־$\widetilde{\mathcal{A} \vee \mathcal{B}}$ שוות, עד כדי הסדר.

תהי $N = q(T)$. אז לכל $v \in V$

$$T^d(v) = -\sum_{i=0}^{d-1} a_i T^i(v) + N(v). \tag{1}$$

לכל i, j מתקיים $T^i N^j = N^j T^i$ (כי $f(T)g(T) = g(T)f(T)$ באשר $f = X^i, g = q^j$). לכן $N(\widetilde{\mathcal{A}}) = \widetilde{N(\mathcal{A})}$. ∎

תרגיל 11.7: *יהי $\mathcal{B}'$ בסיס של V ויהי v האיבר הראשון של $\mathcal{B}'$.*

(א) $[T]_{\mathcal{B}'}^{\mathcal{B}'} = C_r(q)$ *אם ורק אם* $\mathcal{B}' = \big(\widetilde{v} \vee \widetilde{N(v)} \vee \ldots \vee \widetilde{N^{r-1}(v)}\big) = \big(\widetilde{N^j(v)}\big)_{j=0}^{r-1}$ ו־ $v \in \mathrm{Ker}\, N^r$.

(ב) *אם $\mathcal{B}'$ מקיים את התנאים השקולים של (א) אז גם* $\mathcal{B} := \overbrace{\left(N^j(v)\right)_{j=0}^{r-1}}$ *בסיס של* V,

ו־ $[N]_{\mathcal{B}}^{\mathcal{B}} = \mathrm{Diag}\big(\overbrace{J_r(0),\ldots,J_r(0)}^{d}\big)$.

הוכחה: (א) בדיקה ישירה, בעזרת (1). נפרט:

נניח $\mathcal{B}' = (v_{11}, v_{12}, \ldots, v_{1d}, v_{21}, v_{22}, \ldots, v_{2d}, \ , \ \ldots \ , \ v_{r1}, v_{r2}, \ldots, v_{rd})$, באשר $v_{11} = v$.

אז $[T]_{\mathcal{B}'}^{\mathcal{B}'} = C_r(q)$ אם ורק אם העמודות של שתי המטריצות זהות, בהתאמה, כלומר

ע' 1 : $T(v_{11}) = v_{12}$, כלומר $v_{12} = T(v)$

ע' 2 : $T(v_{12}) = v_{13}$, כלומר $v_{13} = T(T(v)) = T^2(v)$

$\ldots$

ע' $d-1$: $T(v_{1\ d-1}) = v_{1d}$, כלומר $v_{1d} = T^{d-1}(v)$

ע' d : $T(v_{1d}) = -a_0 v_{11} - a_1 v_{12} - \ldots - a_{d-1} v_{1d} + v_{21}$, כלומר

$$v_{21} = a_0 v + a_1 T(v) + \ldots + a_{d-1} T^{d-1}(v) + T^d(v) = N(v)$$

ע' $d+1$: $T(v_{21}) = v_{22}$, כלומר $v_{22} = TN(v)$

ע' $d+2$: $T(v_{22}) = v_{23}$, כלומר $v_{23} = T(T(v)) = T^2N(v)$

$\ldots$

ע' $2d-1$: $T(v_{2\ d-1}) = v_{2d}$, כלומר $v_{2d} = T^{d-1}N(v)$

ע' $2d$: $T(v_{2d}) = -a_0 v_{21} - a_1 v_{22} - \ldots - a_{d-1} v_{2d} + v_{31}$, כלומר $v_{31} = N(N(v)) = N^2(v)$

$, \ \ldots \ ,$

ע' $rd-1$: $T(v_{r\ d-1}) = v_{rd}$, כלומר $v_{rd} = T^{d-1}N^{r-1}(v)$

ע' rd : $T(v_{rd}) = -a_0 v_{r1} - a_1 v_{r2} - \ldots - a_{d-1} v_{rd} + 0$, כלומר $0 = N(N^{r-1}(v)) = N^r(v)$

(ב) נתבונן במלבן הבא של $d \cdot r$ וקטורים ב־V:

$$\begin{array}{cccc} v & N(v) & \ldots & N^{r-1}(v) \\ T(v) & T(N(v)) & \ldots & T(N^{r-1}(v)) \\ T^2(v) & T^2(N(v)) & \ldots & T^2(N^{r-1}(v)) \\ & & \ldots & \\ T^{d-1}(v) & T^{d-1}(N(v)) & \ldots & T^{d-1}(N^{r-1}(v)) \end{array} \tag{2}$$

אז $\mathcal{B}'$ מתקבל אם לוקחים את אברי המלבן עמודה אחרי עמודה, ואילו $\mathcal{B}$ מתקבל אם לוקחים את אברי המלבן שורה אחרי שורה. לכן ברור שאם $\mathcal{B}'$ בסיס אז גם $\mathcal{B}$ בסיס.

עבור $0 \le i \le d-1$ תהי $\mathcal{B}_i$ השורה ה־i במלבן (2) ו־$W_i = \mathrm{Sp}(\mathcal{B}_i)$. אז $\mathcal{B}_i$ היא שרשרת ז'ורדן מאורך r של N (משתמשים בכך ש־$T^iN^j = N^jT^i$ לכל i,j.) מתקיים $N(\mathcal{B}_i) \subseteq W_i$, לכן

$N(W_i) \subseteq W_i$, כלומר W_i הינו שמור־N ו־$[N|_{W_i}]_{\mathcal{B}_i}^{\mathcal{B}_i} = J_r(0)$. לכן $V = \mathrm{Sp}(\mathcal{B}) = \bigoplus_{i=0}^{d-1} W_i$ ו־$[N]_{\mathcal{B}}^{\mathcal{B}} = \mathrm{Diag}\big(\overbrace{J_r(0),\dots,J_r(0)}^{d}\big)$. ∎

נכליל את התרגיל:

מסקנה 11.8: *יהי $\mathcal{B}'$ בסיס של V. אז*

(א) $[T]_{\mathcal{B}'}^{\mathcal{B}'} = \mathrm{Diag}\big(\overbrace{C_r(q),\dots,C_r(q)}^{\ell_1\geq 0},\overbrace{C_{r-1}(q),\dots,C_{r-1}(q)}^{\ell_2\geq 0},\ \dots,\ \overbrace{C_1(q),\dots,C_1(q)}^{\ell_r\geq 0}\big)$
אם ורק אם $\mathcal{B}' = \bigvee_{k=1}^{r}\bigvee_{v\in\mathcal{A}_k}\big(\widetilde{N^j(v)}\big)_{j=0}^{r-k}$, באשר $\mathcal{A}_k \subseteq \mathrm{Ker}\,N^{r+1-k}$ ו־$|\mathcal{A}_k| = \ell_k$ לכל k.

(ב) *אם $\mathcal{B}'$ כמו ב־(א) אז $\mathcal{B} := \bigvee_{k=1}^{r}\bigvee_{v\in\mathcal{A}_k}\overbrace{\big(N^j(v)\big)_{j=0}^{r-k}}$ גם בסיס של V ו־$[N]_{\mathcal{B}}^{\mathcal{B}}$ הוא מערך ז'ורדן:*

$$[N]_{\mathcal{B}}^{\mathcal{B}} = \mathrm{Diag}\big(\overbrace{J_r(0),\dots,J_r(0)}^{d|\mathcal{A}_1|},\overbrace{J_{r-1}(0),\dots,J_{r-1}(0)}^{d|\mathcal{A}_2|},\ \dots,\ \overbrace{J_1(0),\dots,J_1(0)}^{d|\mathcal{A}_r|}\big).$$

למה 11.9: *נניח כי q אי פריק ו־$q(T)=0$. תהי $\mathcal{B}=(v_1,\dots,v_{k-1})\subseteq V$ סדרה ויהי $v=v_k\in V$ כך ש־$\widetilde{\mathcal{B}}\vee(v)$ בלתי תלויה לינארית. אז $\widetilde{\mathcal{B}}\vee\widetilde{v}$ בלתי תלויה לינארית.*

הוכחה: אברי $\widetilde{\mathcal{B}}$ הם $\{T^j(v_i)\}_{i=1\,j=0}^{k-1\,d-1}$ ואברי $\widetilde{\mathcal{B}}\vee\widetilde{v}$ הם $\{T^j(v_i)\}_{i=1\,j=0}^{k\,d-1}$ (עד כדי הסדר).
יהיו $a_{ij}\in F$ כך ש־

$$\sum_{i=1}^{k}\sum_{j=0}^{d-1} a_{ij}T^j(v_i) = 0. \tag{3}$$

עלינו להוכיח כי $a_{ij}=0$ לכל i,j. אך די להוכיח כי $a_{kj}=0$ לכל j, כי אז $\sum_{i=1}^{k-1}\sum_{j=0}^{d-1} a_{ij}T^j(v_i)=0$, ומכאן, כיוון ש־$\widetilde{\mathcal{B}}$ בלתי תלויה לינארית, $a_{ij}=0$ גם לכל j ולכל $i\leq k-1$.

נניח שלילה. לכל $1\leq i\leq k$ יהי $f_i=\sum_{j=0}^{d-1}a_{ij}X^j\in F[X]$. אז $f_k\neq 0$, $\deg f_k<d=\deg q$.
כיוון ש־q אי פריק, f_k,q זרים. אז יש $g,h\in F[X]$ כך ש־$1=gq+hf_k$. מכאן

$$v = 1_V(v) = g(T)q(T)(v)+h(T)f_k(T)(v) = h(T)f_k(T)(v). \tag{4}$$

כמו כן לפי החילוק ב־q עם שארית לכל $1\leq i\leq k-1$ יש $a'_{ij}\in F$ ו־$g_i\in F[X]$ כך שמתקיים $hf_i = g_iq+\sum_{j=0}^{d-1}a'_{ij}X^j$. לפי (3), $\sum_{i=1}^{k}f_i(T)(v_i)=0$, לכן

$$0 = h(T)\big(\sum_{i=1}^{k}f_i(T)(v_i)\big) = \sum_{i=1}^{k}h(T)f_i(T)(v_i) = \sum_{i=1}^{k-1}h(T)f_i(T)(v_i)+h(T)f_k(T)(v_k) =$$
$$= \sum_{i=1}^{k-1}g_i(T)q(T)(v_i)+\sum_{i=1}^{k-1}\sum_{j=0}^{d-1}a'_{ij}T^j(v_i)+h(T)f_k(T)(v)$$

כיוון ש־$q(T)=0$, המחובר הראשון באגף ימין הוא 0; לפי (4), המחובר האחרון הוא v, לכן

$$\sum_{i=1}^{k-1}\sum_{j=0}^{d-1} a'_{ij}T^j(v_i)+1v=0$$

ו־$1\neq 0$, סתירה להנחה ש־$\widetilde{\mathcal{B}}\vee(v)$ בלתי תלויה לינארית. ∎

מסקנה 11.10: *נניח כי q אי פריק ו־$q(T)=0$. תהי $\mathcal{B}\subseteq V$ סדרה כך ש־$\widetilde{\mathcal{B}}$ בלתי תלויה לינארית. אז*

(א) *יש $\mathcal{A}\subseteq V$ כך ש־$\widetilde{\mathcal{B}}\vee\widetilde{\mathcal{A}}$ בסיס של V. בפרט*

(ב) *יש $\mathcal{A}\subseteq V$ כך ש־$\widetilde{\mathcal{A}}$ בסיס של V.*

הוכחה: (א) יהי $d=\deg q$. אז לכל סדרה $\mathcal{B}\subseteq V$ מתקיים $|\widetilde{\mathcal{B}}|=d|\mathcal{B}|$. נוכיח את הטענה באינדוקציה על $\dim V-|\widetilde{\mathcal{B}}|$. כיוון ש־$\widetilde{\mathcal{B}}$ בלתי תלויה לינארית, $\dim V-|\widetilde{\mathcal{B}}|\geq 0$.

אם $\dim V-|\widetilde{\mathcal{B}}|=0$, אז $\widetilde{\mathcal{B}}$ בסיס של V; אז ניקח $\mathcal{A}=\emptyset$ (אז גם $\widetilde{\mathcal{A}}=\emptyset$).

תהי $\mathcal{B}$ סדרה כך ש־$\widetilde{\mathcal{B}}$ בלתי תלויה לינארית וכך ש־$\dim V-|\widetilde{\mathcal{B}}|>0$. נניח שהטענה נכונה לכל $\mathcal{B}'$ כך ש־$\widetilde{\mathcal{B}'}$ בלתי תלויה לינארית ו־$\dim V-|\widetilde{\mathcal{B}'}|<\dim V-|\widetilde{\mathcal{B}}|$, כלומר, $|\widetilde{\mathcal{B}'}|>|\widetilde{\mathcal{B}}|$, כלומר, $|\mathcal{B}'|>|\mathcal{B}|$.

אז $|\widetilde{\mathcal{B}}|<\dim V$, לכן $\widetilde{\mathcal{B}}$ בלתי תלויה לינארית לא מרבית, ולכן יש $v\in V$ כך ש־$\widetilde{\mathcal{B}}\vee(v)$ בלתי תלויה לינארית. לפי למה 11.9, $\widetilde{\mathcal{B}}\vee\widetilde{v}$ בלתי תלויה לינארית. נסמן $\mathcal{B}'=\mathcal{B}\vee(v)$. אז $\widetilde{\mathcal{B}}\vee\widetilde{v}$ שווה ל־$\widetilde{\mathcal{B}'}$, עד כדי סדר האיברים. לכן $\widetilde{\mathcal{B}'}$ בלתי תלויה לינארית $|\widetilde{\mathcal{B}'}|>|\widetilde{\mathcal{B}}|$, כלומר, $\mathcal{B}'$ מקיימת את הנחת האינדוקציה. לכן יש $\mathcal{A}'$ כך ש־$\widetilde{\mathcal{B}'}\vee\widetilde{\mathcal{A}'}$ בסיס של V. יהי $\mathcal{A}=(v)\vee\mathcal{A}'$. אז $\widetilde{\mathcal{B}}\vee\widetilde{\mathcal{A}}$ בסיס של V, כי

$$\widetilde{\mathcal{B}}\vee\widetilde{\mathcal{A}}=\widetilde{\mathcal{B}}\vee(\widetilde{v}\vee\widetilde{\mathcal{A}'})=(\widetilde{\mathcal{B}}\vee\widetilde{v})\vee\widetilde{\mathcal{A}'}=\widetilde{\mathcal{B}'}\vee\widetilde{\mathcal{A}'}$$ (שוויון עד כדי סדר האיברים).

(ב) נובע מ־(א) עם $\mathcal{B}=\emptyset$. ∎

כעת נוכיח מקרה פרטי של משפט 11.2:

משפט 11.11: *נניח כי $m_T=q^r$ באשר $q\in F[X]$ פולינום אי פריק מתוקן ממעלה d. אז יש מטריצת יעקובסון יחידה C עבורה יש בסיס $\mathcal{B}'$ של V (לא בהכרח יחיד) כך ש־$[T]_{\mathcal{B}'}^{\mathcal{B}'}=C$. יתר על כן, C מערך יעקובסון בעל אינדקס r השייך ל־q.*

הוכחה: נסמן $N=q(T)$. אז $N^r=0$.

קיום: נעשה תחילה משהו שלכאורה אינו קשור לנושא: נחזור על ההוכחה של משפט 10.13 עבור N, (כלומר, נמצא בסיס $\mathcal{B}$ של V כך ש־$[N]_{\mathcal{B}}^{\mathcal{B}}$ מערך ז'ורדן השייך ל־0) רק שנחזק קצת את המשפט בכך שנחליף את $\mathcal{A}'_k$, $\mathcal{A}_k$, $\mathcal{A}_r$ ב־$\widetilde{\mathcal{A}'_k}$, $\widetilde{\mathcal{A}_k}$, $\widetilde{\mathcal{A}_r}$ עבור תת קבוצות מתאימות של V.

באינדוקציה על r: אם $r=1$ אז $N=0$ ולפי מסקנה 11.10 יש $\mathcal{A}\subseteq V$ כך ש־$\mathcal{B}=\widetilde{\mathcal{A}}$ בסיס של V. אז $[N]_{\mathcal{B}}^{\mathcal{B}}=0$ היא אכן מערך ז'ורדן השייך ל־0. נניח כי $r\geq 2$ והטענה נכונה עבור $r-1$.

יהי $V' = N(V)$ ויהי $N' = N|_{V'}$. אז $(N')^{r-1} = 0$. לפי הנחת האינדוקציה ותרגיל 10.11(ב) יש ל־V' בסיס (צירוף של שרשראות ז'ורדן)

$$\mathcal{B}'' = \bigvee_{k=1}^{r-1} \bigvee_{v' \in \widetilde{\mathcal{A}'_k}} \left(N^j(v')\right)_{j=0}^{r-1-k},$$

באשר $\widetilde{\mathcal{A}'_k} \subseteq \operatorname{Ker}(N')^{r-k}$ לכל k. כיוון ש־$V' = N(V)$, לכל k יש סדרה $\mathcal{A}_k \subseteq V$ כך ש־$N(\mathcal{A}_k) = \mathcal{A}'_k$. אז $N(\widetilde{\mathcal{A}_k}) = \widetilde{N(\mathcal{A}_k)} = \widetilde{\mathcal{A}'_k}$, לכן $\widetilde{\mathcal{A}_k} \subseteq \operatorname{Ker} N^{r+1-k}$. התת סדרה הבאה של $\mathcal{B}''$

$$\bigvee_{k=1}^{r-1} \bigvee_{v' \in \widetilde{\mathcal{A}'_k}} \left(N^{r-1-k}(v')\right) = \bigvee_{k=1}^{r-1} \bigvee_{v \in \widetilde{\mathcal{A}_k}} \left(N^{r-k}(v)\right) = N^{r-1}(\widetilde{\mathcal{A}_1}) \vee N^{r-2}(\widetilde{\mathcal{A}_2}) \vee \ldots \vee N(\widetilde{\mathcal{A}_{r-1}})$$

היא בלתי תלויה לינארית ומוכלת ב־$\operatorname{Ker} N$. לפי מסקנה 11.10 (הצמצום של $N = q(T)$ ל־$\operatorname{Ker} N$ הוא 0) אפשר להשלים אותה לבסיס של $\operatorname{Ker} N$ על ידי איזו סדרה $\widetilde{\mathcal{A}_r}$.

בהוכחה של משפט 10.13 מוכיחים שאז $\mathcal{B} := \bigvee_{k=1}^{r} \bigvee_{v \in \widetilde{\mathcal{A}_k}} \left(N^j(v)\right)_{j=0}^{r-k}$ בסיס של V ו־$[N]_{\mathcal{B}}^{\mathcal{B}}$ מערך ז'ורדן השייך ל־0. לכל $1 \le k \le r$ יש ב־$\mathcal{B}$ בדיוק $|\widetilde{\mathcal{A}_k}| = d|\mathcal{A}_k|$ שרשראות ז'ורדן מאורך $r+1-k$, לכן ב־$[N]_{\mathcal{B}}^{\mathcal{B}}$ יש $d|\mathcal{A}_k|$ גושי ז'ורדן $J_{r+1-k}(0)$.

נסדר את $\mathcal{B}$ בסדר הבא:

$$\mathcal{B}' := \bigvee_{k=1}^{r} \bigvee_{v \in \mathcal{A}_k} \left(N^j(\widetilde{v})\right)_{j=0}^{r-k} = \bigvee_{k=1}^{r} \bigvee_{v \in \mathcal{A}_k} \left(\widetilde{N^j(v)}\right)_{j=0}^{r-k}.$$

אז לפי מסקנה 11.8(א),

$$[T]_{\mathcal{B}'}^{\mathcal{B}'} = \operatorname{Diag}\Big(\overbrace{C_r(q), \ldots, C_r(q)}^{|\mathcal{A}_1|}, \overbrace{C_{r-1}(q), \ldots, C_{r-1}(q)}^{|\mathcal{A}_2|}, \ \ldots, \ \overbrace{C_1(q), \ldots, C_1(q)}^{|\mathcal{A}_r|}\Big).$$

יחידות: יהי $\mathcal{B}'$ בסיס כך ש־$C := [T]_{\mathcal{B}'}^{\mathcal{B}'}$ מטריצת יעקובסון. אז $m_C = m_T = q^r$, לכן לפי מסקנה 11.5, C מערך יעקובסון מאינדקס r השייך ל־q:

$$[T]_{\mathcal{B}'}^{\mathcal{B}'} = \operatorname{Diag}\Big(\overbrace{C_r(q), \ldots, C_r(q)}^{\ell_1}, \overbrace{C_{r-1}(q), \ldots, C_{r-1}(q)}^{\ell_2}, \ \ldots, \ \overbrace{C_1(q), \ldots, C_1(q)}^{\ell_r}\Big).$$

צריך להוכיח שהמספרים $\ell_1, \ldots, \ell_r$ יחידים. לפי מסקנה 11.8, יש בסיס $\mathcal{B}$ אחֵר כך ש־

$$[N]_{\mathcal{B}}^{\mathcal{B}} = \operatorname{Diag}\Big(\overbrace{J_r(0), \ldots, J_r(0)}^{d\ell_1}, \overbrace{J_{r-1}(0), \ldots, J_{r-1}(0)}^{d\ell_2}, \ \ldots, \ \overbrace{J_1(0), \ldots, J_1(0)}^{d\ell_r}\Big).$$

לפי היחידות של צורת ז'ורדן (מסקנה 10.8) סדרת המספרים $d\ell_1, \ldots, d\ell_r$, ולכן גם הסדרה $\ell_1, \ldots, \ell_r$, נקבעת באופן יחיד. ∎

הוכחת משפט 11.2: **קיום ההצגה:** יהי

$$m_T = q_1^{r_1} \cdots q_s^{r_s}$$

הפירוק של הפולינום המזערי של T לחזקות של גורמים אי פריקים מתוקנים שונים, כאשר $r_i \geq 1$ לכל i.

לפי משפט הפירוק הפרימרי, $V = W_1 \oplus \cdots \oplus W_s$, באשר $W_i = \operatorname{Ker} q_i^{r_i}(T)$ ו־$m_{T|_{W_i}} = q_i^{r_i}$ לכל i. לפי משפט 11.11 יש לכל W_i בסיס $\mathcal{B}_i$ כך ש־$C_i := [T|_{W_i}]_{\mathcal{B}_i}^{\mathcal{B}_i}$, מערך יעקובסון השייך ל־$q_i^{r_i}$. לפי משפט 9.7 הצירוף $\mathcal{B}$ של $\mathcal{B}_1, \ldots, \mathcal{B}_s$ הוא בסיס של V ולפי משפט 9.9, $[T]_{\mathcal{B}}^{\mathcal{B}} = \operatorname{Diag}(C_1, \ldots, C_s)$ מטריצת יעקובסון.

יחידות ההצגה: אם $[T]_{\mathcal{B}}^{\mathcal{B}} = C = \operatorname{Diag}(C_1, \ldots, C_t)$, באשר $C_1, \ldots, C_t$ מערכי יעקובסון השייכים לחזקות של פולינומים אי פריקים מתוקנים שונים, אז לפי מסקנה 11.5, $m_C = m_{C_1} \cdots m_{C_t}$ ו־$m_{C_1}, \ldots, m_{C_t}$ חזקות של פולינומים אי פריקים שונים. אבל $m_C = m_T$, לכן, לפי היחידות של פירוק הפולינומים לגורמים אי פריקים, $t = s$ ו־$m_{C_i} = q_i^{r_i}$, לכל i (עד כדי הסדר). אז C_i מערך יעקובסון השייך ל־q_i, לכל i. לפי מסקנה 9.5(ב), $\mathcal{B} = \mathcal{B}_1 \vee \cdots \vee \mathcal{B}_s$, באשר $\mathcal{B}_i$ בסיס של $W_i = \operatorname{Ker} q_i^{r_i}(T)$ כך ש־$C_i = [T|_{W_i}]_{\mathcal{B}_i}^{\mathcal{B}_i}$ לכל i. לפי משפט 11.11, C_i יחידה. ∎

תרגיל 11.12: *תהי $A \in \mathrm{M}_n(F)$. אז*

(א) $f_A = f_{A^t}$.

(ב) $m_A = m_{A^t}$.

(ג) *אם A דומה ל־B אז A^t דומה ל־B^t.*

הוכחה: (א) $f_{A^t} = \det(XI_n - A^t) = \det\big((XI_n - A)^t\big) = \det(XI_n - A) = f_A$.

(ב) אם $g = \sum_{i=0}^m c_i X^i \in F[X]$ כלשהו, אז $(g(A))^t = g(A^t)$. אכן,

$$\big(g(A)\big)^t = \big(\sum_{i=0}^m c_i A^i\big)^t = \sum_{i=0}^m c_i (A^i)^t = \sum_{i=0}^m c_i (A^t)^i = g(A^t).$$

בפרט, $g(A^t) = 0$ אם ורק אם $g(A) = 0$. מכאן נובע ש־$m_A = m_{A^t}$.

(ג) לפי ההנחה יש P הפיכה כך ש־$P^{-1}AP = B$. אז $P^t A^t (P^{-1})^t = B^t$. לכן די להוכיח ש־P^t הפיכה ו־$(P^t)^{-1} = (P^{-1})^t$. ואכן, $PP^{-1} = I_n$, לכן $(P^{-1})^t P^t = I_n^t = I_n$. ∎

נסמן ב־$\sim_F$ את יחס הדמיון של מטריצות מעל F. כלומר, עבור $A, A' \in \mathrm{M}_n(F)$ נכתוב $A \sim_F A'$ אם יש $P \in \mathrm{M}_n(F)$ הפיכה כך ש־$A' = P^{-1}AP$.

תרגיל 11.13: *תהי $A \in \mathrm{M}_n(F)$. אז $A \sim_F A^t$. (כל מטריצה ריבועית דומה למטריצה המוחלפת שלה.)*

הוכחה: די להניח כי A מטריצת יעקובסון. אכן, לפי מסקנה 11.3 יש מטריצת יעקובסון C כך ש־$A \sim_F C$, לפי תרגיל 11.12, $A^t \sim_F C^t$. אם $C \sim_F C^t$, אז, כיוון שדמיון מטריצות הוא יחס שקילות, $A \sim_F A^t$.

די להניח כי A גוש יעקובסון. אכן, $A = \mathrm{Diag}(C_1, \ldots, C_t)$, באשר $C_1, \ldots, C_t$ גושי יעקובסון. אם לכל i יש P_i הפיכה כך ש־$P_i^{-1} C_i P_i = C_i^t$, אז $P := \mathrm{Diag}(P_1, \ldots, P_t)$ הפיכה ומקיימת

$$P^{-1}AP = \mathrm{Diag}(P_1^{-1}C_1P_1, \ldots, P_t^{-1}C_tP_t) = \mathrm{Diag}(C_1^t, \ldots, C_t^t) = A^t.$$

נניח, אם כן, כי $A = C_r(q)$. לפי תרגיל 11.4, $m_A = q^r$. לפי תרגיל 11.12, $m_{A^t} = q^r$. תהי B צורת יעקובסון של A^t, אז $A^t \sim_F B$ ולכן גם $m_B = q^r$. לפי מסקנה 11.5, ב־B יש גוש $C_r(q)$. אבל A, B מאותו הסדר, ו־$A = C_r(q)$, לכן $A = B$. כלומר, $A \sim_F A^t$. ∎

יהיו $F \subseteq F'$ שני שדות ותהיינה $A, B \in \mathrm{M}_n(F)$. אם $A \sim_F B$ אז גם $A \sim_{F'} B$. האם גם ההיפך נכון?

תרגיל 11.14: *תהיינה* $A_1, A_1' \in \mathrm{M}_{n_1}(F)$, $A_2, A_2' \in \mathrm{M}_{n_2}(F)$. *נניח* $A_1 \sim_F A_1'$, $A_2 \sim_F A_2'$. *אז* $\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \sim_F \begin{pmatrix} A_1' & 0 \\ 0 & A_2' \end{pmatrix}$.

הוכחה: לפי ההנחה קיימות מטריצות הפיכות $P_1 \in \mathrm{M}_{n_1}(F), P_2 \in \mathrm{M}_{n_2}(F)$ כך ש־ $P_2^{-1}A_2P_2 = A_2'$, $P_1^{-1}A_1P_1 = A_1'$. לכן

$$\begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix}^{-1} \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix} = \begin{pmatrix} P_1^{-1}A_1P_1 & 0 \\ 0 & P_2^{-1}A_2P_2 \end{pmatrix} = \begin{pmatrix} A_1' & 0 \\ 0 & A_2' \end{pmatrix},$$

כלומר, $\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \sim_F \begin{pmatrix} A_1' & 0 \\ 0 & A_2' \end{pmatrix}$. ∎

תרגיל 11.15: *יהיו* $F \subseteq F'$ *שני שדות. תהיינה* $A, B \in \mathrm{M}_n(F)$ *דומות מעל* F'. *אז הן דומות גם מעל* F.

הוכחה: באינדוקציה על n.

טענה א: *אפשר להחליף את* A, B *במטריצות דומות להן מעל* F.

אכן, A, B הן מטריצות גם מעל F' וגם $A \sim_{F'} B$. אם יש $A', B' \in \mathrm{M}_n(F)$ כך ש־$B \sim_F B'$, $A \sim_F A'$ אז $A' \sim_{F'} A \sim_{F'} B \sim_{F'} B'$, לכן $A' \sim_{F'} B'$. אם נצליח להוכיח ש־$A' \sim_F B'$, אז $A \sim_F A' \sim_F B' \sim_F B$, ולכן $A \sim_F B$.

לכן בלי הגבלת הכלליות A, B מטריצות יעקובסון מעל F. לפי תרגיל 10.5 די להוכיח $A = B$, עד כדי סדר הגושים באלכסון.

כיוון ש־$A \sim_{F'} B$, מתקיים $m_A = m_B$ (כאן משתמשים בכך שהפולינום המזערי של מטריצה אינו תלוי בשדה), נאמר $m_A = m_B = q_1^{r_1} \cdots q_s^{r_s}$, באשר $q_1, \ldots, q_s \in F[X]$ אי פריקים מתוקנים שונים. בלי הגבלת הכלליות המערך הראשון של A, B שייך ל־q_1 (אחרת נחליף את סדר המערכים, לפי תרגיל 10.5). נסמן $C_1 = C_{r_1}(q_1)$.

טענה ב: $A = \begin{pmatrix} C_1 & 0 \\ 0 & A_1 \end{pmatrix}, B = \begin{pmatrix} C_1 & 0 \\ 0 & B_1 \end{pmatrix}$ *באשר* A_1, B_1 *מטריצות יעקובסון מעל* F.

אכן, זה ברור מההגדרה של מטריצת יעקובסון, כמטריצה של גושים באלכסון שהינם גושי יעקובסון מעל F.

בפרט A_1, B_1 מאותו סדר, קטן מ־n. תהיינה A', B', C' צורות יעקובסון של A_1, B_1, C_1 מעל F'.

טענה ג: $\begin{pmatrix} C' & 0 \\ 0 & B' \end{pmatrix}, \begin{pmatrix} C' & 0 \\ 0 & A' \end{pmatrix}$ *הן צורות יעקובסון של* A, B *מעל* F'*, עד כדי סדר הגושים באלכסון.*

אכן, שתי המטריצות הן מטריצות של גושים באלכסון, שהינם גושי יעקובסון מעל F', לכן, עד כדי סדר הגושים, הן מטריצות יעקובסון מעל F'. לפי תרגיל 11.14 הן דומות ל־A, B מעל F'. לכן הן צורות יעקובסון שלהן מעל F'.

כיוון ש־A, B דומות מעל F', גם $\begin{pmatrix} C' & 0 \\ 0 & B' \end{pmatrix}, \begin{pmatrix} C' & 0 \\ 0 & A' \end{pmatrix}$ דומות מעל F'. לכן הן זהות, עד כדי סדר הגושים באלכסון. מכאן $A' = B'$, עד כדי סדר הגושים באלכסון. בפרט A_1, B_1 דומות מעל F'. לפי הנחת האינדוקציה הן דומות מעל F. לכן הן שוות, עד כדי סדר הגושים באלכסון. מכאן $A = B$, עד כדי סדר הגושים באלכסון. ∎

12. מרחבי מכפלה פנימית

בפרק זה יהיה $F=\mathbb{R}$ או $F=\mathbb{C}$.

נזכור שלכל $z\in\mathbb{C}$ קיים **הצמוד** שלו $\overline{z}$ **והערך המוחלט** שלו $|z|$ המוגדרים על ידי

$$\overline{x+iy}=x-iy\quad,\ |x+iy|=\sqrt{x^2+y^2}\qquad\qquad x,y\in\mathbb{R}.$$

התכונות הבסיסיות:

(א) $\overline{z_1\pm z_2}=\overline{z_1}\pm\overline{z_2}$;

(ב) $\overline{z_1z_2}=\overline{z_1}\,\overline{z_2}$;

(ג) $\overline{z}z=|z|^2$.

(ד) $\mathrm{Re}(z)=\frac{z+\overline{z}}{2}$ החלק הממשי של z;

(ה) z ממשי אם ורק אם $z=\overline{z}$;

(ו) $z=0$ אם ורק אם $|z|=0$;

(ז) $\overline{\overline{z}}=z$.

הגדרה 12.1: יהי V מרחב וקטורי מעל F. **מכפלה פנימית על** V היא העתקה $V\times V\to F$ שמתאימה לכל זוג (u,v) של וקטורים ב־V סקלר $\langle u,v\rangle\in F$ כך שמתקיים לכל $u,v,v'\in V$ ולכל $\alpha\in F$

(1) $\langle u,v+v'\rangle=\langle u,v\rangle+\langle u,v'\rangle$.

(2) $\langle u,\alpha v\rangle=\alpha\langle u,v\rangle$.

(3) $\overline{\langle v,u\rangle}=\langle u,v\rangle$.

(4) $\langle v,v\rangle>0$ לכל $0\neq v\in V$. (שים לב: לפי (3), $\langle v,v\rangle\in\mathbb{R}$.)

מרחב מכפלה פנימית הוא מרחב וקטורי מעל F יחד עם מכפלה פנימית עליו.

דוגמה 12.2: (א) **המכפלה הסטנדרטית**, נקראת גם **סקלרית**: $V=F^n$, ונגדיר $\langle u,v\rangle_{\mathrm{st}}=\overline{u}^tv$, כלומר,

$$\left\langle\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix},\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}\right\rangle_{\mathrm{st}}=(\overline{a_1},\overline{a_2},\dots,\overline{a_n})\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}=\overline{a_1}b_1+\overline{a_2}b_2+\dots+\overline{a_n}b_n=\sum_{i=1}^n\overline{a_i}b_i.$$

קל לראות שתכונות (1)־(3) מתקיימות. נבדוק את (4): נניח $a_1,a_2,\dots,a_n\in F$ לא כולם אפס. אז

$$\left\langle\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix},\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}\right\rangle_{\mathrm{st}}=\sum_{i=1}^n\overline{a_i}a_i=\sum_{i=1}^n|a_i|^2>0,$$

כי $|a_i|^2\geq 0$ ו־$|a_i|^2>0$ אם $a_i\neq 0$.

(ב) $V = F^2$, $\langle \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \rangle = 2\overline{a_1}b_1 + \overline{a_1}b_2 + \overline{a_2}b_1 + \overline{a_2}b_2$.

נבדוק את התכונה (4) בהגדרה:

$$\begin{aligned}\langle \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \rangle &= 2\overline{a_1}a_1 + \overline{a_1}a_2 + \overline{a_2}a_1 + \overline{a_2}a_2 = (\overline{a_1} + \overline{a_2})(a_1 + a_2) + \overline{a_1}a_1 \\ &= |a_1 + a_2|^2 + |a_1|^2 > 0\end{aligned}$$

כאשר $a_1 \neq 0$ או $a_1 + a_2 \neq 0$, כלומר, כאשר $a_2 \neq 0$ או $a_1 \neq 0$.

מעתה ועד סוף הפרק יהי V מרחב מכפלה פנימית.

למה 12.3: *יהיו $u, u', v \in V$ ויהי $\alpha \in F$:*

(1) $\langle u + u', v \rangle = \langle u, v \rangle + \langle u', v \rangle$;

(2) $\langle \alpha u, v \rangle = \overline{\alpha}\langle u, v \rangle$;

(3) $\langle 0, v \rangle = \langle u, 0 \rangle = 0$.

הוכחה: (1) $\langle u + u', v \rangle = \overline{\langle v, u + u' \rangle} = \overline{\langle v, u \rangle + \langle v, u' \rangle} = \overline{\langle v, u \rangle} + \overline{\langle v, u' \rangle} = \langle u, v \rangle + \langle u', v \rangle$.

(2) $\langle \alpha u, v \rangle = \overline{\langle v, \alpha u \rangle} = \overline{\alpha \langle v, u \rangle} = \overline{\alpha}\overline{\langle v, u \rangle} = \overline{\alpha}\langle u, v \rangle$

(3) $\langle u, 0 \rangle = \langle u, 0 \cdot 0 \rangle = 0\langle u, 0 \rangle = 0$. באופן דומה $\langle 0, v \rangle = \langle 0 \cdot 0, v \rangle = \overline{0}\langle 0, v \rangle = 0$. ∎

הגדרה 12.4:

(1) $||v|| = \sqrt{\langle v, v \rangle}$ נקרא **הנורמה (האורך)** של v.

(2) u **נִצב ל־**v – נכתוב $u \perp v$ – אם $\langle u, v \rangle = 0$. זה שקול ל־$\langle v, u \rangle = 0$.

(3) סדרה $u_1, u_2, \ldots, u_n \in V$ נקראת **אורתוגונלית** אם $u_i \perp u_j$ לכל $i \neq j$.

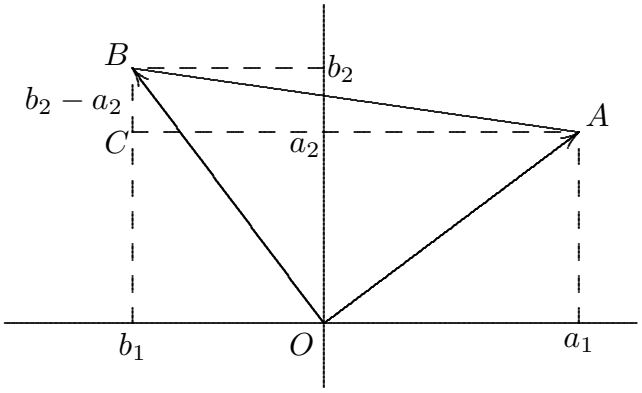
(4) סדרה $u_1, u_2, \ldots, u_n \in V$ נקראת **אורתונורמלית** אם היא אורתוגונלית ו־$||u_i|| = 1$ לכל i. במלים אחרות, $\langle u_i, u_j \rangle = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$.

דוגמה 12.5: הבסיס הסטנדרטי $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ של F^n הינו אורתונורמלי ביחס למכפלה הסטנדרטית.

הערה 12.6: יהי $V = \mathbb{R}^2$ עם המכפלה הסטנדרטית. נראה את אברי V כחצים מהראשית. אם $A = (a_1, a_2)$ אז לפי משפט פיתגורס אורכו של $\vec{OA}$ הוא $\sqrt{a_1^2 + a_2^2}$, כלומר, הנורמה של $\vec{OA}$.

כמו כן, תהי גם $B = (b_1, b_2)$. אז, שוב לפי משפט פיתגורס, $\vec{OA}, \vec{OB}$ מאונכים זה לזה אם ורק אם $AB^2 = OA^2 + OB^2$, כלומר, $(b_1 - a_1)^2 + (b_2 - a_2)^2 = \left(a_1^2 + a_2^2\right) + \left(b_1^2 + b_2^2\right)$. זה שקול ל־$a_1b_1 + a_2b_2 = 0$, כלומר, $\langle \vec{OA}, \vec{OB} \rangle_{\text{st}} = 0$.

באופן דומה האורך והניצבות ב־$V = \mathbb{R}^3$ ביחס למכפלה הסטנדרטית הם האורך והניצבות המוכרים לנו מהגיאומטריה.



תרגיל 12.7: *אם v ניצב ל־ $u_1, \ldots, u_k$ אז הוא ניצב לכל u ב־ $\mathrm{Sp}(u_1, \ldots, u_k)$.*

הוכחה: יהי $u = \sum_{i=1}^k a_i u_i \in \mathrm{Sp}(u_1, \ldots, u_k)$. אז $\langle v, u\rangle = \sum_{i=1}^k a_i \langle v, u_i\rangle = \sum_{i=1}^k a_i 0 = 0$.

∎

למה 12.8 (למת האיפיון): *תהי $u_1, u_2, \ldots, u_n$ סדרה אורתוגונלית ב־ V. יהיו*

$$u = a_1 u_1 + \cdots + a_n u_n, \qquad v = b_1 u_1 + \cdots + b_n u_n \in V.$$

(א) $a_i ||u_i||^2 = \langle u_i, u\rangle$. בפרט, אם $u_i \neq 0$, אז $a_i = \frac{\langle u_i, u\rangle}{||u_i||^2}$. ובפרט, אם $||u_i|| = 1$, אז $a_i = \langle u_i, u\rangle$.

(ב) $\langle u, v\rangle = \sum_{i=1}^n \overline{a_i} b_i ||u_i||^2$. בפרט, אם $||u_i|| = 1$ לכל i, אז $\langle u, v\rangle = \sum_{i=1}^n \overline{a_i} b_i$.

(ג) $||u||^2 = \sum_{i=1}^n |a_i|^2 ||u_i||^2$. בפרט, אם $||u_i|| = 1$ לכל i, אז $||u||^2 = \sum_{i=1}^n |a_i|^2$.

הוכחה: (א) $\langle u_i, u\rangle = a_1\langle u_i, u_1\rangle + a_2\langle u_i, u_2\rangle + \cdots + a_n\langle u_i, u_n\rangle = a_i\langle u_i, u_i\rangle = a_i ||u_i||^2$.

(ב) $\langle u, v\rangle = \sum_{i=1}^n \sum_{j=1}^n \overline{a_i} b_j \langle u_i, u_j\rangle = \sum_{i=1}^n \overline{a_i} b_i ||u_i||^2$.

(ג) הצב $v = u$ בחלק הקודם. ∎

מסקנה 12.9: *יהי $\mathcal{B}$ בסיס אורתונורמלי של V. יהיו $u, v \in V$. אז $\langle u, v\rangle = \langle [u]_{\mathcal{B}}, [v]_{\mathcal{B}}\rangle_{\mathrm{st}}$*

משפט 12.10: *סדרה אורתוגונלית $u_1, u_2, \ldots, u_n$ שאבריה שונים מאפס הינה בלתי תלויים לינארית מעל F.*

הוכחה: נניח $a_1 u_1 + a_2 u_2 + \cdots + a_n u_n = 0$. לפי הלמה $a_i = \frac{\langle u_i, 0\rangle}{||u_i||^2} = \frac{0}{||u_i||^2} = 0$, לכל i. ∎

בניה 12.11 (תהליך Gram-Schmidt למציאת בסיס אורתוגונלי):

נתונה סדרה $v_1, v_2, \ldots, v_n$ ב־ V. נמצא סדרה אורתוגונלית $u_1, u_2, \ldots, u_n$ ב־ V כך שמתקיים לכל $1 \leq k \leq n$:

(1^k) $\mathrm{Sp}(v_1, v_2, \ldots, v_k) = \mathrm{Sp}(u_1, u_2, \ldots, u_k)$.

(2^k) $u_k = 0$ אם ורק אם $v_k \in \mathrm{Sp}(v_1, v_2, \ldots, v_{k-1})$.

בפרט

(3) $u_1, u_2, \ldots, u_n$ *שונים מאפס אם ורק אם* $v_1, v_2, \ldots, v_n$ *בלתי תלויים לינארית מעל* F.

(4) *אם* $v_1, v_2, \ldots, v_n$ *בסיס של* V *אז* $u_1, u_2, \ldots, u_n$ *בסיס של* V.

הבניה: תחילה נשים לב ש־(3) נובע מ־(2) ו־(4) נובע מ־(1).

נבנה באינדוקציה על k את הסדרה כך שיתקיים

(0^k) $u_1, u_2, \ldots, u_k$ אורתוגונלית.

(1^k) $\mathrm{Sp}(v_1, v_2, \ldots, v_k) = \mathrm{Sp}(u_1, u_2, \ldots, u_k)$.

(2^k) $u_k = 0$ אם ורק אם $v_k \in \mathrm{Sp}(v_1, v_2, \ldots, v_{k-1})$.

עבור $k = 1$ נגדיר $u_1 = v_1$. אז (0^1), (1^1) ברורים; (2^1): $u_1 = v_1 = 0 \Leftrightarrow v_1 \in \mathrm{Sp}(\emptyset)$.

נניח באינדוקציה שכבר בנינו $u_1, u_2, \ldots, u_k$ שמקיימים את (0^k), (1^k), (2^k). נגדיר

$$u_{k+1} = v_{k+1} - (c_1 u_1 + c_2 u_2 + \cdots + c_k u_k), \qquad (*)$$

באשר $c_i = \begin{cases} \frac{\langle u_i, v_{k+1}\rangle}{||u_i||^2} & u_i \neq 0 \\ 0 & u_i = 0 \end{cases}$ לכל $1 \leq i \leq k$.

אז $c_i||u_i||^2 = \langle u_i, v_{k+1}\rangle$ לכל $1 \leq i \leq k$ (גם אם $u_i = 0$).

נוכיח את (0^{k+1}), (1^{k+1}), (2^{k+1}):

(0^{k+1}) $u_1, u_2, \ldots, u_k, u_{k+1}$ אורתוגונלית: די להוכיח $\langle u_i, u_{k+1}\rangle = 0$ לכל $1 \leq i \leq k$. ואכן,

$$\langle u_i, u_{k+1}\rangle = \langle u_i, v_{k+1}\rangle - \sum_{j=1}^{k} c_j \langle u_i, u_j\rangle = \langle u_i, v_{k+1}\rangle - c_i||u_i||^2 = 0.$$

(1^{k+1}) לפי $(*)$ ולפי (1^k):

$$\mathrm{Sp}(u_1, u_2, \ldots, u_k, u_{k+1}) = \mathrm{Sp}(u_1, u_2, \ldots, u_k, v_{k+1}) = \mathrm{Sp}(v_1, v_2, \ldots, v_k, v_{k+1}).$$

(2^{k+1}) אם $u_{k+1} = 0$ אז לפי $(*)$

$$v_{k+1} = c_1 u_1 + c_2 u_2 + \cdots + c_k u_k \in \mathrm{Sp}(u_1, u_2, \ldots, u_k) = \mathrm{Sp}(v_1, v_2, \ldots, v_k).$$

להיפך, אם $v_{k+1} \in \mathrm{Sp}(v_1, v_2, \ldots, v_k) = \mathrm{Sp}(u_1, u_2, \ldots, u_k)$ אז $v_{k+1} = \sum_{i=1}^{k} a_i u_i$ עבור איזה $a_1, a_2, \ldots, a_k \in F$. לפי למת האיפיון 12.8(א), $a_i = c_i$ אם $u_i \neq 0$. לפי $(*)$, $u_{k+1} = 0$.

∎

דוגמה 12.12: נעשה תהליך גרם־שמידט ב־$\mathbb{R}^3$ על $v_1 = (0,3,4), v_2 = (1,2,1), v_3 = (0,0,2)$ לפי המכפלה הסטנדרטית.

(א) $u_1 = v_1 = (0,3,4)$, $||u_1||^2 = 0^2 + 3^2 + 4^2 = 25$,

(ב) $\langle u_1, v_2\rangle_{\mathrm{st}} = 0 \cdot 1 + 3 \cdot 2 + 4 \cdot 1 = 10$,

$$u_2 = (1,2,1) - \tfrac{10}{25}(0,3,4) = (1,2,1) - (0,\tfrac{6}{5},\tfrac{8}{5}) = (1,\tfrac{4}{5},\tfrac{-3}{5}) = \tfrac{1}{5}(5,4,-3),$$
$$||u_2||^2 = \tfrac{25}{25} + \tfrac{16}{25} + \tfrac{9}{25} = 2.$$

(ג) $\langle u_1, v_3\rangle_{\text{st}} = 0\cdot 0 + 3\cdot 0 + 4\cdot 2 = 8$, $\langle u_2, v_3\rangle_{\text{st}} = 1\cdot 0 + \frac{4}{5}\cdot 0 + \frac{-3}{5}\cdot 2 = \frac{-6}{5}$.

$$u_3 = (0,0,2) - \tfrac{8}{25}(0,3,4) - \tfrac{-\frac{6}{5}}{2}(1,\tfrac{4}{5},\tfrac{-3}{5}) = \tfrac{1}{25}(15,-12,9).$$

מסקנה 12.13: *אם V נוצר סופית אז יש לו בסיס אורתונורמלי.*

הוכחה: ל־V יש בסיס $v_1, v_2, \ldots, v_n$. לפי המשפט הקודם יש לו בסיס אורתוגונלי $u_1, u_2, \ldots, u_n$. קל לבדוק כי $\frac{1}{||u_1||}u_1, \frac{1}{||u_2||}u_2, \ldots, \frac{1}{||u_n||}u_n$ סדרה אורתונורמלית. ∎

הערה 12.14: המשמעות הגיאומטרית של תהליך גרם־שמידט. נתבונן בנוסחה $(*)$. נסמן

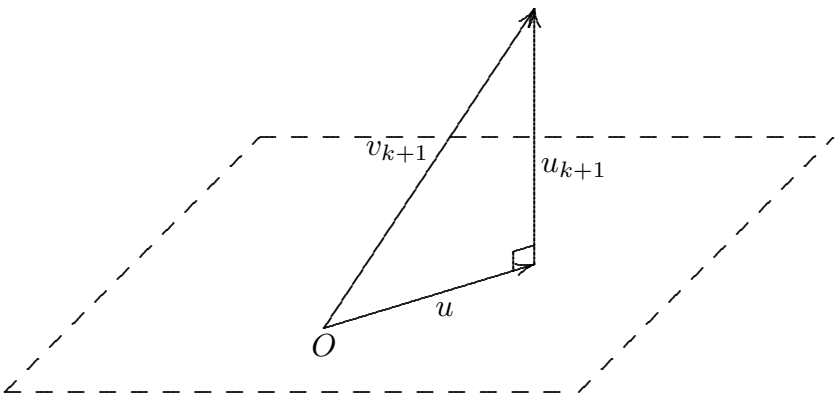
$$U = \mathrm{Sp}(v_1, v_2, \ldots, v_k) = \mathrm{Sp}(u_1, u_2, \ldots, u_k), \qquad u = c_1u_1 + c_2u_2 + \cdots + c_ku_k.$$

אז $v_{k+1} = u_{k+1} + u$, $u \in U$, ולפי תרגיל 12.7, u_{k+1} ניצב לכל הוקטורים ב־U.

נניח כי $V = \mathbb{R}^3$ ו־U הוא מישור או ישר דרך הראשית. אז

(א) u הוא ההיטל של v_{k+1} על U,

(ב) u_{k+1} הוא החץ מאיזושהי נקודה ב־U אל נקודת הקצה של v_{k+1}, אשר ניצב ל־U. בפרט אורכו $||u_{k+1}||$ הוא המרחק בין U לנקודת הקצה של v_{k+1}.



תרגיל 12.15: *מצא את מרחק של הנקודה $(0,0,2)$ מהמישור $U = \mathrm{Sp}((0,3,4),(1,2,1))$*

פתרון: בדוגמה 12.12 עשינו תהליך גרם־שמידט על $v_1 = (0,3,4), v_2 = (1,2,1), v_3 = (0,0,2)$. מצאנו ש־$u_3 = \frac{1}{25}(15,-12,9)$. לכן $||u_3||^2 = \frac{1}{25^2}(225+144+81) = \frac{1}{25^2}(450) = \frac{18}{25}$. המרחק המבוקש הוא $||u_3|| = \sqrt{\frac{18}{25}} = \frac{3}{5}\sqrt{2}$. ∎

תרגיל 12.16: *אם מבצעים תהליך גרם־שמידט על $v_1, \ldots, v_r, v_{r+1}, \ldots, v_n$ ו־ $v_1, \ldots, v_r$ סדרה אורתוגונלית, אז $u_1 = v_1, u_2 = v_2, \ldots, u_r = v_r$.*

הוכחה: באינדוקציה: לפי הבניה, תמיד $u_1 = v_1$.

נניח $u_1 = v_1, \ldots, u_k = v_k$, באשר $1 \leq k < r$. אז $u_{k+1} = v_{k+1} - (c_1u_1 + c_2u_2 + \cdots + c_ku_k)$, באשר $c_i = \begin{cases} \frac{\langle v_i, v_{k+1}\rangle}{||v_i||^2} = 0 & u_i \neq 0 \\ 0 & u_i = 0 \end{cases}$. לכן $u_{k+1} = v_{k+1}$. ∎

מסקנה 12.17: *יהי V נוצר סופית. תהי $v_1, \ldots, v_r$ סדרה אורתוגונלית, שאיבריה שונים מאפס. אז אפשר להשלימה לבסיס אורתוגונלי של V.*

הוכחה: לפי משפט 12.10, $v_1, \ldots, v_r$ בלתי תלויה לינארית. לכן ניתן להשלימה לבסיס $v_1, \ldots, v_r, \ldots, v_n$ של V. תהליך גרם־שמידט עליו יתן בסיס אורתוגונלי $u_1, \ldots, u_r, \ldots, u_n$. לפי תרגיל 12.16, $u_1 = v_1, \ldots, u_r = v_r$.

∎

משפט 12.18 (אי שוויון Cauchy-Schwarz): *יהיו v_1, v_2 שני וקטורים ב־V. אז $|\langle v_1, v_2\rangle| \leq ||v_1|| \cdot ||v_2||$ ויש כאן שוויון אם ורק אם v_1, v_2 תלויים לינארית.*

הוכחה: אם $v_1 = 0$, יש שוויון ו־v_1, v_2 תלויים לינארית. לכן נניח $v_1 \neq 0$.

תהליך גרם־שמידט על v_1, v_2 נותן סדרה אורתוגונלית u_1, u_2, בה $u_1 = v_1 \neq 0$, $u_2 = v_2 - c_1u_1$, באשר $c_1 = \frac{\langle u_1, v_2\rangle}{||u_1||^2}$. נעביר c_1u_1 לאגף שני ונקבל $v_2 = c_1u_1 + u_2$. לפי למת האיפיון 12.8(ג),

$$||v_2||^2 = |c_1|^2 \cdot ||u_1||^2 + 1 \cdot ||u_2||^2 = \frac{|\langle v_1, v_2\rangle|^2}{||u_1||^2} + ||u_2||^2 \geq \frac{|\langle v_1, v_2\rangle|^2}{||u_1||^2}$$

ויש שוויון אם ורק אם $||u_2|| = 0$, כלומר, $u_2 = 0$, כלומר (לפי התנאים של תהליך גרם־שמידט), $v_2 \in \mathrm{Sp}(v_1)$, כלומר, v_1, v_2 תלויים לינארית. נכפיל את האי שוויון ב־$||u_1||^2$ כדי לקבל את מבוקש. ∎

משפט 12.19 (אי שוויון המשולש): *יהיו $v_1, v_2 \in V$. אז $||v_1 + v_2|| \leq ||v_1|| + ||v_2||$.*

הוכחה: נשתמש באי שוויון קושי־שוורץ כדי להוכיח $(||v_1 + v_2||)^2 \leq (||v_1|| + ||v_2||)^2$:

$$\begin{aligned}(||v_1 + v_2||)^2 &= \langle v_1 + v_2, v_1 + v_2\rangle = \langle v_1, v_1\rangle + \langle v_1, v_2\rangle + \langle v_2, v_1\rangle + \langle v_2, v_2\rangle = \\ &= ||v_1||^2 + \langle v_1, v_2\rangle + \overline{\langle v_1, v_2\rangle} + ||v_2||^2 = ||v_1||^2 + 2\mathrm{Re}(\langle v_1, v_2\rangle) + ||v_2||^2 \leq \\ &\leq ||v_1||^2 + 2|\langle v_1, v_2\rangle| + ||v_2||^2 \leq ||v_1||^2 + 2||v_1|| \cdot ||v_2|| + ||v_2||^2 = (||v_1|| + ||v_2||)^2\end{aligned}$$

∎

הגדרה 12.20: יהי U תת מרחב של V. **המשלים הניצב (האורתוגונלי)** הוא

$$U^\perp = \{v \in V |\ u \in U \text{ לכל } \langle u, v\rangle = 0\}.$$

למה 12.21: *(א) $U^\perp$ הוא תת מרחב של V;*

(ב) $U \cap U^\perp = \{0\}$.

הוכחה: (א) יהיו $v, v' \in U^\perp$ ויהי $\alpha \in F$. אז, לכל $u \in U$,

$$\langle u, v + v'\rangle = \langle u, v\rangle + \langle u, v'\rangle = 0 + 0 = 0$$

ולכן $v + v' \in U^\perp$.

כמו כן $\langle u, \alpha v\rangle = \alpha\langle u, v\rangle = \alpha 0 = 0$ ולכן $\alpha v \in U^\perp$.

לבסוף, $\langle u, 0\rangle = 0$ לכל $u \in U$, לכן $0 \in U^\perp$.

(ב) יהי $u \in U \cap U^\perp$. אז $||u||^2 = \langle u, u\rangle = 0$, ולכן $u = 0$. ∎

מציאת המשלים הניצב במרחב נוצר סופית:

יהי $v_1, \ldots, v_r$ בסיס של U. נשלים אותו לבסיס $v_1, \ldots, v_r, \ldots, v_n$ של V. תהליך גרם־שמידט יתן בסיס אורתוגונלי $u_1, \ldots, u_r, \ldots, u_n$ כך ש־$U = \mathrm{Sp}(v_1, \ldots, v_r) = \mathrm{Sp}(u_1, \ldots, u_r)$. ואז:

משפט 12.22: *יהי V נוצר סופית. אז בסימונים לעיל*

(א) $U^\perp = \mathrm{Sp}(u_{r+1}, \ldots, u_n)$, *בפרט* $\dim U^\perp = \dim V - \dim U$.

(ב) $U + U^\perp = V$, *ולכן* $U \oplus U^\perp = V$.

הוכחה: (א) תחילה נראה $U^\perp \supseteq \mathrm{Sp}(u_{r+1}, \ldots, u_n)$. לשם כך די להוכיח ש־$u_k \in U^\perp$ לכל $r + 1 \le k \le n$. יהי $u \in U$. אז יש $a_1, \ldots, a_r \in F$ כך ש־$u = a_1u_1 + \cdots + a_ru_r$. מכאן

$$\langle u_k, u\rangle = a_1\langle u_k, u_1\rangle + \cdots + a_r\langle u_k, u_r\rangle = 0,$$

לכן $u_k \in U^\perp$.

להיפך, יהי $v \in U^\perp$. אז $v \in V$ ולכן יש $a_1, \ldots, a_n \in F$ כך ש־$v = a_1u_1 + \cdots + a_nu_n$. יהי $1 \le i \le r$. אז $v \perp u_i$ ולכן, לפי למת האיפיון 12.8, $a_i = 0$. מכאן $v = a_{r+1}u_{r+1} + \cdots + a_nu_n \in \mathrm{Sp}(u_{r+1}, \ldots, u_n)$.

(ב) לפי (א), $U + U^\perp = \mathrm{Sp}(u_1, \ldots, u_r, u_{r+1}, \ldots, u_n) = V$. ∎

מסקנה 12.23: *אם V נוצר סופית אז $(U^\perp)^\perp = U$.*

13. העתקות במרחבי מכפלה פנימית

בפרק זה יהי $F = \mathbb{R}$ או $F = \mathbb{C}$, ו־V הוא מרחב מכפלה פנימית בעל מימד סופי מעל F. תהי $T\colon V \to V$ העתקה לינארית.

משפט 13.1: *לכל $v \in V$ קיים $w = T^*(v) \in V$ יחיד שמקיים*

$$\langle T(u), v\rangle = \langle u, w\rangle = \langle u, T^*(v)\rangle \qquad \text{לכל } u \in V. \tag{1}$$

יתר על כן, ההעתקה $T^*\colon V \to V$ לינארית.

הוכחה: יהי $u_1, \ldots, u_n$ בסיס אורתונורמלי של V.

יחידות w: יהי $v, w \in V$. לפי למת האיפיון 12.8(א), $w = \sum_{i=1}^n \langle u_i, w\rangle u_i$. אם w מקיים (1), אז

$$w = \sum_{i=1}^n \langle T(u_i), v\rangle u_i \tag{2}$$

ומכאן היחידות.

קיום w: נגדיר w על ידי (2). אז לפי למת האיפיון 12.8(א), $\langle u_i, w\rangle = \langle T(u_i), v\rangle$ ו־$u = \sum_{i=1}^n \langle u_i, u\rangle u_i$ לכל i. לכן לפי למת האיפיון 12.8(ב),

$$\langle u, w\rangle = \sum_{i=1}^n \overline{\langle u_i, u\rangle}\langle u_i, w\rangle = \sum_{i=1}^n \overline{\langle u_i, u\rangle}\langle T(u_i), v\rangle.$$

אך מצד שני, מתוך $u = \sum_{i=1}^n \langle u_i, u\rangle u_i$ (לעיל) נקבל $T(u) = \sum_{i=1}^n \langle u_i, u\rangle T(u_i)$. מכאן

$$\langle T(u), v\rangle = \sum_{i=1}^n \overline{\langle u_i, u\rangle}\langle T(u_i), v\rangle.$$

לכן (1) מתקיים.

T^* לינארית: יהי $a \in F$. אז לכל $u \in V$

$$\langle T(u), av\rangle = a\langle T(u), v\rangle = a\langle u, T^*(v)\rangle = \langle u, aT^*(v)\rangle$$

ומכאן, לפי היחידות ב־(1), $T^*(av) = aT^*(v)$.

יהיו $v_1, v_2 \in V$. אז לכל $u \in V$

$$\langle T(u), v_1+v_2\rangle = \langle T(u), v_1\rangle + \langle T(u), v_2\rangle = \langle u, T^*(v_1)\rangle + \langle u, T^*(v_2)\rangle = \langle u, T^*(v_1)+T^*(v_2)\rangle$$

ומכאן, לפי היחידות ב־(1), $T^*(v_1 + v_2) = T^*(v_1) + T^*(v_2)$. ∎

הגדרה 13.2: (א) T^* נקראת **ההעתקה הצמודה של** T.

(ב) אם $A \in \mathrm{M}_n(F)$, אז **המטריצה הצמודה לה** $A^* \in \mathrm{M}_n(F)$ מוגדרת על ידי

$$(A^*)_{ij} = \overline{(A)_{ji}} \qquad \text{לכל } 1, \leq i, j \leq n$$

(כלומר, $A^* = \overline{A^t}$). אם $F = \mathbb{R}$, אז $A^* = A^t$.

בהמשך נסמן את המטריצה של העתקה T לפי בסיס $\mathcal{B}$ ב־$[T]_{\mathcal{B}}$ (במקום $[T]_{\mathcal{B}}^{\mathcal{B}}$).

משפט 13.3: *יהי $\mathcal{B} = (u_1, \ldots, u_n)$ בסיס אורתונורמלי של V. אז $[T^*]_{\mathcal{B}} = ([T]_{\mathcal{B}})^*$.*

הוכחה: לפי ההגדרה של מטריצה של העתקה, העמודה ה־i של $[T]_{\mathcal{B}}$ היא וקטור הקואורדינטות $[T(u_i)]_{\mathcal{B}}$ של $T(u_i)$ לפי הבסיס $\mathcal{B}$. לפי למת האיפיון 12.8(א), הרכיב ה־j של העמודה הזאת הוא $\langle u_j, T(u_i)\rangle$. לכן

$$((T]_{\mathcal{B}})^*)_{ij} = \overline{([T]_{\mathcal{B}})_{ji}} = \overline{\langle u_j, T(u_i)\rangle} = \langle T(u_i), u_j\rangle = \langle u_i, T^*(u_j)\rangle = ([T^*]_{\mathcal{B}})_{ij}.$$ ∎

למה 13.4: *יהי $a \in F$ ותהיינה $T, S\colon V \to V$ העתקות לינאריות. אז*

(א) $(T+S)^* = T^* + S^*$

(ב) $(aT)^* = \overline{a}T^*$

(ג) $(TS)^* = S^*T^*$

(ד) $(T^*)^* = T$

הוכחה: יהי $\mathcal{B}$ בסיס אורתונורמלי של V ותהיינה $A = [T]_{\mathcal{B}}, B = [S]_{\mathcal{B}}$. אז $[T+S]_{\mathcal{B}} = A + B$, $[aT]_{\mathcal{B}} = aA$, $[TS]_{\mathcal{B}} = [T]_{\mathcal{B}}[S]_{\mathcal{B}} = AB$, ולכן די להוכיח

(א') $(A+B)^* = A^* + B^*$

(ב') $(aA)^* = \overline{a}A^*$

(ג') $(AB)^* = B^*A^*$

(ד') $(A^*)^* = A$

וזה קל. למשל, (ג')

$$\left((AB)^*\right)_{ij} = \overline{(AB)_{ji}} = \overline{\sum_{k=1}^{n}(A)_{jk}(B)_{ki}} = \sum_{k=1}^{n}\overline{(A)_{jk}(B)_{ki}} = \sum_{k=1}^{n}\overline{(A)_{jk}}\,\overline{(B)_{ki}} =$$
$$= \sum_{k=1}^{n}(B^*)_{ik}(A^*)_{kj} = (B^*A^*)_{ij}$$

לכל i, j ולכן $(AB)^* = B^*A^*$.

אפשר לתת גם הוכחה ישירה, לא דרך המטריצות. למשל, (ד) $(T^*)^* = T$:

לפי הגדרה (1), $(T^*)^*(v)$ הוא אותו איבר יחיד w של V שמקיים $\langle T^*(u), v\rangle = \langle u, w\rangle$ לכל $u \in V$.
אבל לכל $u, v \in V$ מתקיים

$$\langle T^*(u), v\rangle = \overline{\langle v, T^*(u)\rangle} = \overline{\langle T(v), u\rangle} = \langle u, T(v)\rangle,$$

לכן $w = T(v)$, כלומר, $(T^*)^*(v) = T(v)$, לכל $v \in V$. ∎

הגדרה 13.5: תהי $T\colon V \to V$ העתקה לינארית ותהי $A \in \mathrm{M}_n(F)$.

T (A) נקראת **צמודה לעצמה** אם $T = T^*$ ($A = A^*$).

T (A) נקראת **אוניטרית** אם $T^*T = 1_V$ ($A^*A = I$).

T (A) נקראת **נורמלית** אם $T^*T = TT^*$ ($A^*A = AA^*$).

אם $F = \mathbb{C}$, אומרים גם **הֶרמיטית** במקום "צמודה לעצמה".

אם $F = \mathbb{R}$, אומרים גם **סימטרית** במקום "צמודה לעצמה".

אם $F = \mathbb{R}$, אומרים גם **אורתוגונלית** במקום "אוניטרית". ∎

מסקנה 13.6: *יהי $\mathcal{B}$ בסיס אורתונורמלי של V ותהי $A = [T]_{\mathcal{B}}$. אז T צמודה לעצמה/אוניטרית/נורמלית אם ורק אם A כזו.*

תרגיל 13.7: *אם $\langle u, T(v)\rangle = 0$ לכל $u, v \in V$ אז $T = 0$.*

הוכחה: יהי $v \in V$. נקח $u = T(v)$. אז $\langle T(v), T(v)\rangle = 0$, לכן $T(v) = 0$. מכאן $T = 0$. ∎

למה 13.8: *נניח $\langle v, T(v)\rangle = 0$ לכל $v \in V$ וכי*

(א) $F = \mathbb{C}$; *או*

(ב) $F = \mathbb{R}$ *ו־T צמודה לעצמה.*

אז $T = 0$.

הוכחה: יהיו $u, v \in V$ ויהי $a \in F$.

$$0 = \langle u + av, T(u + av)\rangle = \langle u, T(u)\rangle + \bar{a}\langle v, T(u)\rangle + a\langle u, T(v)\rangle + \bar{a}a\langle v, T(v)\rangle =$$
$$= \bar{a}\langle v, T(u)\rangle + a\langle u, T(v)\rangle$$

(א) נניח כי $F = \mathbb{C}$. אם נציב בחישוב לעיל $a = 1$ ואח"כ $a = i$, נקבל

$$\langle v, T(u)\rangle + \langle u, T(v)\rangle = 0$$
$$-i\langle v, T(u)\rangle + i\langle u, T(v)\rangle = 0$$

ומכאן $\langle u, T(v)\rangle = 0$ לכל $u, v \in V$. לפי התרגיל, $T = 0$.

(ב) נניח כי $F = \mathbb{R}$ ו־$T = T^*$. אז (כיון ש־$\bar{b} = b$ לכל $b \in \mathbb{R}$)

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

לכן אם נציב בחישוב לעיל $a = 1$, נקבל $2\langle u, T(v)\rangle = 0$, ומכאן $\langle u, T(v)\rangle = 0$ לכל $u, v \in V$. לפי התרגיל, $T = 0$. ∎

דוגמה 13.9: תנאים (א) או (ב) אכן נחוצים: יהי $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ הסיבוב ב־$90°$ סביב הראשית. אז $T \neq 0$, אך $\langle v, T(v)\rangle = 0$ לכל $v \in V$. ∎

משפט 13.10: *נניח כי $F = \mathbb{C}$. אז $\langle v, T(v)\rangle \in \mathbb{R} \Leftrightarrow T = T^*$ לכל $v \in V$.*

הוכחה: נניח $T^* = T$. אז $\langle v, T(v)\rangle = \langle v, T^*(v)\rangle = \langle T(v), v\rangle = \overline{\langle v, T(v)\rangle}$. לכן $\langle v, T(v)\rangle$ ממשי.
להיפך, נניח $\langle v, T(v)\rangle \in \mathbb{R}$ לכל $v \in V$. אז לכל $v \in V$

$$\langle v, (T-T^*)(v)\rangle = \langle v, T(v)\rangle - \langle v, T^*(v)\rangle = \langle v, T(v)\rangle - \langle T(v), v\rangle = \langle v, T(v)\rangle - \overline{\langle v, T(v)\rangle} = 0$$

ולכן לפי הלמה $T - T^* = 0$, כלומר, $T = T^*$. ∎

תרגיל 13.11: *ההעתקות T^*T, TT^*, $T^*T - 1_V$ צמודות לעצמן.*

הוכחה: $(T^*T)^* = T^*(T^*)^* = T^*T$, $\quad (T^*T)^* = T^*(T^*)^* = T^*T$,
$(T^*T - 1_V)^* = (T^*T)^* - (1_V)^* = T^*T - 1_V$. ∎

משפט 13.12: *התנאים הבאים שקולים זה לזה:*

(1) *T אוניטרית, כלומר, $T^*T = 1_V$.*

(2) *$TT^* = 1_V$.*

(3) *T מעתיקה כל בסיס אורתונורמלי של V לבסיס אורתונורמלי של V.*

(4) *T מעתיקה בסיס אורתונורמלי מסוים של V לבסיס אורתונורמלי של V.*

(5) *T משמרת מכפלה פנימית, כלומר: $\langle T(u), T(v)\rangle = \langle u, v\rangle$ לכל $u, v \in V$.*

(6) *T משמרת נורמה, כלומר: $||T(u)|| = ||u||$ לכל $u \in V$.*

(7) *T משמרת מרחק, כלומר: $||T(u) - T(v)|| = ||u - v||$ לכל $u, v \in V$.*

הוכחה:

(1) $\Leftarrow$ (2): אם $T^*T = 1_V$, אז T^*, T הפיכות (T חח"ע ו־T^* על. בדקו!) לכן $T^* = T^{-1}$. מכאן $TT^* = 1_V$.

(2) $\Leftarrow$ (1): באותו אופן (החלף בין T לבין T^*).

(1) $\Leftarrow$(3): יהי $u_1,\dots,u_n$ בסיס אורתונורמלי של V. אז לכל $1 \le i,j \le n$

$$\langle T(u_i),T(u_j)\rangle = \langle u_i, T^*T(u_j)\rangle = \langle u_i, 1_V(u_j)\rangle = \langle u_i,u_j\rangle = \delta_{ij},$$

$T(u_1),\dots,T(u_n)$ אורתונורמלית. לכן היא בלתי תלויה לינארית; היות ו־$n=\dim V$, היא אף בסיס.

(3) $\Leftarrow$(4): טריביאלי.

(4) $\Leftarrow$(5): נניח כי $u_1,\dots,u_n$ בסיס אורתונורמלי וגם $T(u_1),\dots,T(u_n)$ בסיס אורתונורמלי. יהיו $u=\sum_{i=1}^n a_iu_i, v=\sum_{i=1}^n b_iu_i$. אז לפי למת האיפיון 12.8

$$\langle T(u),T(v)\rangle = \langle \sum_{i=1}^n a_iT(u_i), \sum_{i=1}^n b_iT(u_i)\rangle = \sum_{i=1}^n \overline{a_i}b_i,$$
$$\langle u,v\rangle = \langle \sum_{i=1}^n a_iu_i, \sum_{i=1}^n b_iu_i\rangle = \sum_{i=1}^n \overline{a_i}b_i,$$

כלומר, $\langle T(u),T(v)\rangle = \langle u,v\rangle$.

(5) $\Leftarrow$(6): $||T(u)||^2 = \langle T(u),T(u)\rangle = \langle u,u\rangle = ||u||^2$, לכן $||T(u)|| = ||u||$.

(6) $\Leftarrow$(7): $||T(u)-T(v)|| = ||T(u-v)|| = ||u-v||$.

(7) $\Leftarrow$(6): הצב $v=0$.

(6) $\Leftarrow$(1): יהי $u\in V$. אז

$$\langle u, T^*T(u)\rangle = \langle T(u),T(u)\rangle = ||T(u)||^2 = ||u||^2 = \langle u,u\rangle = \langle u, 1_V(u)\rangle.$$

מכאן

$$\langle u, (T^*T-1_V)(u)\rangle = 0 \qquad \text{לכל } u\in V.$$

לפי תרגיל 13.11, T^*T-1_V צמודה לעצמה, לכן לפי למה 13.8(ב), $T^*T-1_V=0$. לכן $T^*T=1_V$. ∎

משפט 13.13: *עבור $A\in \mathrm{M}_n(F)$ שלושת התנאים הבאים שקולים זה לזה:*

(א) *A אוניטרית, כלומר, $A^*A=I$ (כלומר, $AA^*=I$).*

(ב) *העמודות של A מהוות בסיס אורתונורמלי של F^n (ביחס למכפלה הסטנדרטית על F^n).*

(ג) *השורות של A מהוות בסיס אורתונורמלי של F^n (ביחס למכפלה הסטנדרטית על F^n).*

הוכחה: (אם u שורה או עמודה של מטריצה, נסמן ב־$(u)_k$ הרכיב ה־k שלה.)

(א) $\Leftrightarrow$(ב):

תהיינה $u_1,\dots,u_n$ העמודות של A. אז $\overline{u_1^t},\dots,\overline{u_n^t}$ הן השורות של A^*. מכאן

$$(A^*A)_{ij}=\sum_{k=1}^n \overline{(u_i^t)}_k\,(u_j)_k=\sum_{k=1}^n \overline{(u_i)}_k\,(u_j)_k=\langle u_i,u_j\rangle_{\mathrm{st}}.$$

לכן $u_1,\dots,u_n$ אורתונורמלית $\Leftrightarrow$ $\langle u_i,u_j\rangle_{\mathrm{st}}=\delta_{ij}$ לכל i,j $\Leftrightarrow A^*A=I$.

(א) $\Leftrightarrow$(ג): תהיינה $v_1,\dots,v_n$ השורות של A. אז $\overline{v_1^t},\dots,\overline{v_n^t}$ הן העמודות של A^*. מכאן

$$(AA^*)_{ij}=\sum_{k=1}^n (v_i)_k\,\overline{(v_j^t)}_k=\sum_{k=1}^n \overline{(v_j)}_k\,(v_i)_k=\langle v_j,v_i\rangle_{\mathrm{st}}.$$

לכן $AA^*=I\Leftrightarrow$ $\langle v_j,v_i\rangle_{\mathrm{st}}=\delta_{ji}$ לכל i,j $\Leftrightarrow v_1,\dots,v_n$ אורתונורמלית. ∎

תרגיל 13.14: *תהי $A\in \mathrm{M}_n(F)$ מטריצת המעבר מבסיס אורתונורמלי $\mathcal{B}$ של V לבסיס $\mathcal{B}'$ של V. אז $\mathcal{B}'$ אורתונורמלי $\Leftrightarrow A$ אוניטרית.*

הוכחה: נניח $\mathcal{B}=(u_1,\dots,u_n)$, $\mathcal{B}'=(u_1',\dots,u_n')$. אז $A=\big([u_1']_{\mathcal{B}},\dots,[u_n']_{\mathcal{B}}\big)$, כלומר, רכיב $(A)_{ij}$ של A הוא הרכיב ה־i של u_j' לפי בסיס $\mathcal{B}$. לכן, לפי למת האיפיון 12.8(א), $(A)_{ij}=\langle u_i,u_j'\rangle$. מכאן לפי למת האיפיון 12.8(ב)

$$(A^*A)_{ij}=\sum_{k=1}^n (A^*)_{ik}(A)_{kj}=\sum_{k=1}^n \overline{(A)_{ki}}(A)_{kj}=\sum_{k=1}^n \overline{\langle u_k,u_i'\rangle}\,\langle u_k,u_j'\rangle=\langle u_i',u_j'\rangle$$

לכן $A^*A=I\Leftrightarrow$ $(A^*A)_{ij}=\delta_{ij}$ לכל i,j $\Leftrightarrow u_1',\dots,u_n'$ אורתונורמלית. ∎

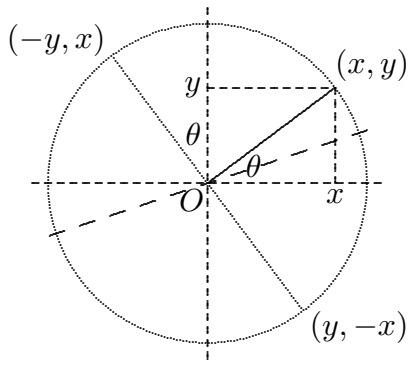
דוגמה 13.15: מטריצה $A=(\lambda)\in \mathrm{M}_1(F)$ אוניטרית אם ורק אם $|\lambda|=1$.

(לפי משפט 13.13) $A\in \mathrm{M}_2(\mathbb{R})$ אורתוגונלית (=אוניטרית מעל $\mathbb{R}$) אם ורק אם היא מהצורה

$$A=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\qquad \text{או}\qquad A=\begin{pmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{pmatrix}$$

עבור איזה $\theta\in\mathbb{R}$. הימנית מייצגת סיבוב סביב הראשית, השמאלית שיקוף ביחס לְציר דרך הראשית.

אכן, העמודה הראשונה של A היא $\begin{pmatrix}x\\y\end{pmatrix}$, באשר $||\begin{pmatrix}x\\y\end{pmatrix}||=1$, כלומר, $x^2+y^2=1$. לכן היא מהצורה $\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}$, עבור איזה $\theta\in\mathbb{R}$. העמודה השניה מייצגת וקטור ניצב לו, בעל נורמה 1, לכן הוא $\begin{pmatrix}\cos(\theta+\frac{\pi}{2})\\ \sin(\theta+\frac{\pi}{2})\end{pmatrix}$ או $\begin{pmatrix}\cos(\theta-\frac{\pi}{2})\\ \sin(\theta-\frac{\pi}{2})\end{pmatrix}$.



ההכפלה במטריצה הימנית מעתיקה את e_1 על $\begin{pmatrix} x \\ y \end{pmatrix}$, ואת e_2 על $\begin{pmatrix} -y \\ x \end{pmatrix}$, כלומר, מסובבת את שניהם בזוית θ סביב הראשית. לכן היא עושה זאת לכל וקטור ב־$\mathbb{R}^2$.

ההכפלה במטריצה השמאלית מעתיקה את e_1 על $\begin{pmatrix} x \\ y \end{pmatrix}$, ואת e_2 על $\begin{pmatrix} y \\ -x \end{pmatrix}$, כלומר, משקפת את שניהם ביחס לציר דרך הראשית בעל זוית $\frac{\theta}{2}$ עם ציר ה־x־ים. לכן היא עושה זאת לכל וקטור ב־$\mathbb{R}^2$. ∎

תרגיל 13.16: *תהי $T\colon V \to V$ העתקה לינארית ויהי W תת מרחב שמור־T. אז $W^\perp$ הוא שמור־T^*.*

הוכחה: יהי $v \in W^\perp$. אז לכל $w \in W$ מתקיים $T(w) \in W$ ולכן $0 = \langle T(w), v\rangle = \langle w, T^*(v)\rangle$, כלומר, $T^*(v)$ ניצב לכל וקטור ב־W. מכאן $T^*(v) \in W^\perp$. ∎

למה 13.17: *תהי $T\colon V \to V$ נורמלית ויהיו $\lambda \in F$, $v \in V$. אם $T(v) = \lambda v$ אז $T^*(v) = \bar{\lambda} v$.*

הוכחה: תחילה נניח כי $\lambda = 0$. אז $T(v) = 0$. מכאן

$$\begin{aligned} ||T^*(v)||^2 &= \langle T^*(v), T^*(v)\rangle = \langle v, (T^*)^* T^*(v)\rangle = \langle v, TT^*(v)\rangle = \\ &= \langle v, T^*T(v)\rangle = \langle T(v), T(v)\rangle = ||T(v)||^2 = 0 \end{aligned}$$

ולכן $T^*(v) = 0$, כנדרש.

המקרה הכללי: נסמן $N = T - \lambda 1_V$. אז $N^* = T^* - \bar{\lambda} 1_V^* = T^* - \bar{\lambda} 1_V$. לכן

$$\begin{aligned} NN^* =& (T - \lambda 1_V)(T^* - \bar{\lambda} 1_V) = TT^* - \lambda T^* - \bar{\lambda} T + \lambda\bar{\lambda} 1_V = \\ =& T^*T - \bar{\lambda} T - \lambda T^* + \bar{\lambda}\lambda 1_V = (T^* - \bar{\lambda} 1_V)(T - \lambda 1_V) = N^*N \end{aligned}$$

כלומר N נורמלית. כעת, $N(v) = T(v) - \lambda v = 0$. לכן לפי המקרה הקודם, $N^*(v) = 0$, כלומר, $T^*(v) - \bar{\lambda} v = 0$, כלומר, $T^*(v) = \bar{\lambda} v$. ∎

למה 13.18: *תהי $T\colon V \to V$ נורמלית. וקטורים עצמיים של T השייכים לערכים עצמיים שונים הינם ניצבים זה לזה.*

הוכחה: יהיו $v, w \in V$ כך ש־$T(v) = \lambda v$, $T(w) = \mu w$, באשר $\lambda, \mu \in F$ שונים. לפי הלמה הקודמת, $T^*(v) = \bar{\lambda} v$. לכן

$$\mu\langle v, w\rangle = \langle v, \mu w\rangle = \langle v, T(w)\rangle = \langle T^*(v), w\rangle = \langle \bar{\lambda} v, w\rangle = \lambda\langle v, w\rangle.$$

מכאן $(\lambda - \mu)\langle v, w\rangle = 0$. אבל $\lambda - \mu \neq 0$, לכן $\langle v, w\rangle = 0$ ∎

שני המשפטים הבאים (13.20 ו־13.23) הם התוצאות המרכזיות של הפרק, האחד עבור $F = \mathbb{C}$ והשני עבור $F = \mathbb{R}$. עיקר ההוכחה נעשית בלמה הבאה – באופן מאוחד לשני השדות:

למה 13.19: *תהי $T\colon V \to V$ העתקה לינארית כך ש־f_T הוא מכפלה של גורמים מהצורה $X-\lambda$, $\lambda \in F$. אז T נורמלית אם ורק אם קיים בסיס אורתונורמלי $\mathcal{B}$ של V המורכב מוקטורים עצמיים של T (כלומר, $[T]_{\mathcal{B}}$ אלכסונית).*

הוכחה: $\Rightarrow$: נניח $\mathcal{B} = (u_1, \ldots, u_n)$, באשר $T(u_i) = \lambda_i u_i$ לכל $1 \le i \le n$. תהי $A = [T]_{\mathcal{B}}$. לפי משפט 4.7, $A = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$. ולכן $A^* = \mathrm{Diag}(\bar{\lambda}_1, \ldots, \bar{\lambda}_n)$. מכאן

$$AA^* = \mathrm{Diag}(\lambda_1\bar{\lambda}_1, \ldots, \lambda_n\bar{\lambda}_n) = \mathrm{Diag}(\bar{\lambda}_1\lambda_1, \ldots, \bar{\lambda}_n\lambda_n) = A^*A,$$

כלומר, A נורמלית. לפי מסקנה 13.6, T נורמלית.

$\Leftarrow$: יהיו $\lambda_1, \ldots, \lambda_k \in F$ הערכים העצמיים השונים של T. לכל מרחב עצמי V_{λ_i} נבחר בסיס אורתונורמלי. לפי למה 13.18, צירוף הבסיסים האלה היא סדרה אורתונורמלית. נסמנה $\mathcal{B}$, ויהי $W = \mathrm{Sp}(\mathcal{B})$. נשים לב ש־$V_{\lambda_i} \subseteq W$ לכל i. נראה ש־$W = V$ ואז $\mathcal{B}$ היא בסיס אורתונורמלי של V, לפי משפט 12.10.

די להראות כי $W^\perp = \{0\}$, כי אז לפי מסקנה 12.23, $W = (W^\perp)^\perp = \{0\}^\perp = V$.

טענה: *$W, W^\perp$ שמורי־T.*

אם $u \in \mathcal{B}$, אז יש i כך ש־$T(u) = \lambda_i u$; לפי למה 13.17, $T^*(u) = \overline{\lambda_i} u$, ובפרט נובע מכאן ש־$T(u), T^*(u) \in \mathrm{Sp}(\mathcal{B}) = W$. לכן $T(W) = T(\mathrm{Sp}(\mathcal{B})) = \mathrm{Sp}(T(\mathcal{B})) \subseteq W$, כלומר, W הוא שמור־T. באותו אופן W שמור־T^*. לפי למה 13.4(ד), $(T^*)^* = T$. לכן לפי תרגיל 13.16, $W^\perp$ שמור־T.

נניח בשלילה כי $W^\perp \neq \{0\}$. יהי S הצמצום של T ל־$W^\perp$. אז $\deg f_S \ge 1$. לפי משפט 12.22, $V = W \oplus W^\perp$, לכן לפי מסקנה 9.10, f_S מחלק את f_T. מכאן שגם f_S הוא מכפלה של גורמים מהצורה $X-\lambda$, $\lambda \in F$. בפרט ל־f_S יש שורש $\lambda \in F$. אז λ ערך עצמי של S, כלומר, יש $0 \neq u \in W^\perp \subseteq V$ כך ש־$T(u) = S(u) = \lambda u$. מכאן ש־λ גם ערך עצמי של T, כלומר יש i כך ש־$\lambda = \lambda_i$. ולכן $u \in V_{\lambda_i} \subseteq W$. אבל אז $0 \neq u \in W \cap W^\perp = \{0\}$, סתירה. ∎

אם $F = \mathbb{C}$, התנאי על f_T בלמה 13.19 מתקיים בגלל המשפט היסודי של אלגברה. לכן הוכחנו:

משפט 13.20: *נניח $F = \mathbb{C}$. העתקה לינארית $T\colon V \to V$ היא נורמלית אם ורק אם קיים בסיס אורתונורמלי $\mathcal{B}$ של V המורכב מוקטורים עצמיים של T (כלומר, $[T]_{\mathcal{B}}$ אלכסונית).*

מסקנה 13.21: *תהי $A \in \mathrm{M}_n(\mathbb{C})$. אז A נורמלית אם ורק אם קיימת $P \in \mathrm{M}_n(\mathbb{C})$ אוניטרית כך ש־$P^{-1}AP = P^*AP$ אלכסונית.*

הוכחה: יהי $V = \mathbb{C}^n$ ויהי St הבסיס הסטנדרטי שלו; אז St אורתונורמלי. תהי $T\colon V \to V$ נתונה על ידי $T(v) = Av$; אז $[T]_{\mathrm{St}}^{\mathrm{St}} = A$. לפי מסקנה 13.6, A נורמלית אם ורק אם T נורמלית. לפי המשפט הקודם T נורמלית אם ורק אם יש בסיס אורתונורמלי $\mathcal{B}$ של V כך ש־$[T]_{\mathcal{B}}$ אלכסונית.

אם זה קורה, תהי P מטריצת המעבר מ־St ל־$\mathcal{B}$; לפי תרגיל 13.14, P אוניטרית. כידוע, $P^{-1}AP = [T]_{\mathcal{B}}$, לכן $P^{-1}AP$ אלכסונית.

להיפך, נניח כי $P \in \mathrm{M}_n(\mathbb{C})$ אוניטרית כך ש־$P^{-1}AP$ אלכסונית. אז P מטריצת המעבר מ־St לסדרה $\mathcal{B}$ כלשהי (ביתר דיוק, $\mathcal{B}$ סדרת העמודות של P); לפי תרגיל 13.14, $\mathcal{B}$ בסיס אורתונורמלי. כידוע, $[T]_{\mathcal{B}} = P^{-1}AP$. לכן $[T]_{\mathcal{B}}$ אלכסונית. ∎

מסקנה 13.22: *תהי $A \in \mathrm{M}_n(F)$ נורמלית. אז*

(א) *A צמודה לעצמה אם ורק אם כל שרשי $f_A(X)$ ב־$\mathbb{C}$ (!) הם ממשיים.*

(ב) *A אוניטרית אם ורק אם כל שרשי $f_A(X)$ ב־$\mathbb{C}$ (!) הם בעלי ערך מוחלט 1.*

הוכחה: נתבונן ב־A כמטריצה מעל $\mathbb{C}$ (גם אם $F = \mathbb{R}$). אז A נורמלית, לכן לפי מסקנה 13.21 יש $P \in \mathrm{M}_n(\mathbb{C})$ אוניטרית כך ש־$D = P^*AP = P^{-1}AP$ אלכסונית, נאמר, $D = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$, באשר $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ השרשים של f_A. אז $D^* = P^*A^*P$ ומכאן:

- D צמודה לעצמה $\Leftrightarrow P^*AP = P^*A^*P \Leftrightarrow A = A^* \Leftrightarrow A$ צמודה לעצמה.
- D אוניטרית $\Leftrightarrow (P^*A^*P)(P^*AP) = I_n \Leftrightarrow P^*(A^*A)P = I_n \Leftrightarrow A^*A = I_n \Leftrightarrow A$ אוניטרית.

כמו כן $f_A = f_D$ כי A, D דומות. לכן בלי הגבלת הכלליות $A = D$ אלכסונית.

(א) D צמודה לעצמה $\Leftrightarrow \mathrm{Diag}(\lambda_1, \ldots, \lambda_n) = \mathrm{Diag}(\overline{\lambda_1}, \ldots, \overline{\lambda_n}) \Leftrightarrow \overline{\lambda_i} = \lambda_i$ לכל $1 \le i \le n$ $\Leftrightarrow \lambda_1, \ldots, \lambda_n$ ממשיים.

(ב) D אוניטרית $\Leftrightarrow D^*D = I_n \Leftrightarrow \mathrm{Diag}(\overline{\lambda_1}, \ldots, \overline{\lambda_n})\mathrm{Diag}(\lambda_1, \ldots, \lambda_n) = \mathrm{Diag}(1, \ldots, 1)$ $\Leftrightarrow |\lambda_i|^2 = \overline{\lambda_i}\lambda_i = 1$ לכל $1 \le i \le n$ $\Leftrightarrow |\lambda_1| = \cdots = |\lambda_n| = 1$. ∎

משפט 13.23: *נניח $F = \mathbb{R}$. העתקה לינארית $T\colon V \to V$ היא צמודה לעצמה אם ורק אם קיים בסיס אורתונורמלי $\mathcal{B}$ של V המורכב מוקטורים עצמיים של T (כלומר, $[T]_{\mathcal{B}}$ אלכסונית).*

הוכחה: $\Rightarrow$: אם $[T]_{\mathcal{B}} \in \mathrm{M}_n(\mathbb{R})$ אלכסונית, היא צמודה לעצמה, ולכן T צמודה לעצמה לפי מסקנה 13.6.

$\Leftarrow$: T צמודה לעצמה, לכן נורמלית. לפי מסקנה 13.22 f_T הוא מכפלה של גורמים מהצורה $X - \lambda$, $\lambda \in \mathbb{R}$. לכן הטענה נובעת מלמה 13.19. ∎

מסקנה 13.24: *מטריצה $A \in \mathrm{M}_n(\mathbb{R})$ היא סימטרית אם ורק אם קיימת $P \in \mathrm{M}_n(\mathbb{R})$ אורתוגונלית כך ש־$P^tAP = P^{-1}AP$ אלכסונית.*

הוכחה: כמו ההוכחה של מסקנה 13.21: החלף $\mathbb{C}$ ב־$\mathbb{R}$ ובמקום נורמלית כתוב סימטרית (= צמודה לעצמה). ∎

משפט 13.25: *נניח $F=\mathbb{R}$. תהי $T\colon V\to V$ העתקה אורתוגונלית. אזי קיים בסיס אורתונורמלי $\mathcal{B}$ של V כך ש־*

$$[T]_{\mathcal{B}} = \begin{array}{c}
\begin{array}{ll}\overbrace{\hspace{5em}}^{p} & \overbrace{\hspace{6em}}^{q}\hspace{20em}\end{array}\\
\begin{pmatrix}
1 & & & & & & & & & & \mathbf{0}\\
& \ddots & & & & & & & & & \\
& & 1 & & & & & & & & \\
& & & -1 & & & & & & & \\
& & & & \ddots & & & & & & \\
& & & & & -1 & & & & & \\
& & & & & & \cos\theta_1 & -\sin\theta_1 & & & \\
& & & & & & \sin\theta_1 & \cos\theta_1 & & & \\
& & & \mathbf{0} & & & & & \ddots & & \\
& & & & & & & & & \cos\theta_r & -\sin\theta_r\\
& & & & & & & & & \sin\theta_r & \cos\theta_r
\end{pmatrix}
\end{array}$$

באשר $p,q,r\geq 0$ ו־$\theta_1,\dots,\theta_r\in\mathbb{R}$.

הוכחה: באינדוקציה על $\dim V$. נניח קודם כי $\dim V=1$ או $\dim V=2$. נבחר בסיס אורתונורמלי כלשהו $\mathcal{B}$ של V. לפי מסקנה 13.6, $[T]_{\mathcal{B}}$ אורתוגונלית. לפי דוגמה 13.15, היא מהצורה $A=(1)$ או $A=(-1)$ או $A=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$ או $A=\begin{pmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{pmatrix}$. אך במקרה האחרון

$$f_T=f_A=X^2-\operatorname{tr}(A)X+\det A=X^2-1=(X-1)(X+1),$$

ו־T נורמלית, לכן לפי למה 13.19 יש בסיס אורתונורמלי אחר $\mathcal{B}'$ כך ש־$[T]_{\mathcal{B}'}=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$.

כעת נניח כי $\dim V\geq 3$ וכי המשפט הוכח עבור כל מרחב ממימד $<\dim V$.

טענה א: *יש תת מרחב $W\subseteq V$ שמור־T כך ש־$0<\dim W<\dim V$.* אכן, אם ל־T יש וקטור עצמי $w\neq 0$, אז $W=\operatorname{Sp}(w)$ הוא שמור־T ממימד 1. נניח, אם כן, של־T אין וקטורים עצמיים שונים מאפס.

יהי $S=T+T^*=T+T^{-1}\colon V\to V$. אז $S^*=(T+T^*)^*=T^*+T=S$, לכן לפי מסקנה 13.22(א), יש ל־S ערך עצמי $\lambda\in\mathbb{R}$. לכן יש $0\neq v\in V$ כך ש־$(T+T^{-1})(v)=S(v)=\lambda v$. נפעיל T על שוויון זה ונקבל

$$T^2(v)+v=\lambda T(v).$$

כעת, $T(v)$ אינו כפולה של v (אחרת v היה וקטור עצמי של T, סתירה), לכן $v,T(v)$ בלתי תלויים לינארית. לכן $W=\operatorname{Sp}(v,T(v))$ בעל מימד 2. הוא שמור־T, כי $T(v)\in W$ וגם

$$T(T(v))=T^2(v)=\lambda T(v)-v\in W.$$

טענה ב: *אם $W\subseteq V$ תת מרחב שמור־T, אז גם $W^{\perp}$ שמור־T.* אכן, $T^*=T^{-1}$, לכן לפי תרגיל 9.3(ד), W שמור־T^*. לפי תרגיל 13.16, $W^{\perp}$ שמור־T.

סיום ההוכחה: יהי W כמו בטענה א. נסמן $W_1 = W, W_2 = W^\perp$. אז W_1, W_2 שמורי־T. לפי משפט 12.22(ב), $V = W_1 \oplus W_2$, ו־$\dim W_1, \dim W_2 < \dim V$ (כי $\dim V = \dim W_1 + \dim W_2$). לפי הנחת האינדוקציה יש בסיס אורתונורמלי $\mathcal{B}_i$ של W_i כך ש־$[T|_{W_i}]_{\mathcal{B}_i}$ מהצורה המבוקשת. אז $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ הוא בסיס של V, הוא אורתונורמלי, ו־ $[T]_{\mathcal{B}} = \begin{pmatrix} [T|_{W_1}]_{\mathcal{B}_1} & 0 \\ 0 & [T|_{W_2}]_{\mathcal{B}_2} \end{pmatrix}$ היא מהצורה המבוקשת (עד כדי הסדר של אברי $\mathcal{B}$). ∎

תרגיל 13.26: הוכיחו שקבוצת המטריצות האוניטריות ב־$\mathrm{M}_n(F)$ היא **חבורה**, כלומר, היא סגורה תחת כפל ולקיחת הופכי ומכילה את I_n.

הוכחה: מתקיים $I_n^* = I_n$, לכן $I_n^* I_n = I_n$. לכן I_n אוניטרית.

כעת נשתמש במשפט 13.14. יהי $\mathcal{B}$ בסיס אורתונורמלי של F^n.

תהיינה $P, P' \in \mathrm{M}_n(F)$ אוניטריות. אז P מטריצת המעבר מ־$\mathcal{B}$ לבסיס אורתונורמלי $\mathcal{B}'$, ואילו P' היא מטריצת המעבר מ־$\mathcal{B}'$ לבסיס אורתונורמלי $\mathcal{B}''$. לכן PP' היא מטריצת המעבר מ־$\mathcal{B}$ ל־$\mathcal{B}''$. לכן היא אוניטרית.

כמו כן, P^{-1} היא מטריצת המעבר מ־$\mathcal{B}'$ ל־$\mathcal{B}$, שניהם בסיסים אורתונורמליים, ולכן היא אוניטרית. ∎

תרגיל 13.27: בתרגיל זה נסביר את הקשר בין העתקה צמודה לבין העתקה דואלית.

נזכור שאם V מרחב וקטורי מעל שדה F, אז **המרחב הדואלי** שלו הוא $V^* = \{\varphi\colon V \to F \mid \varphi \text{ לינארית}\}$. זהו מרחב וקטורי מעל F. נשתמש בסימון הבא: עבור $v \in V$ ו־$\varphi \in V^*$ נכתוב (φ, v) במקום $\varphi(v)$.

אם $T\colon V \to W$ העתקה לינארית בין שני מרחבים וקטוריים מעל F, אז **ההעתקה הדואלית** של T היא ההעתקה $T^o\colon W^* \to V^*$ המוגדרת על ידי ש־$(\psi, T(u)) = (T^o(\psi), u)$ לכל $u \in V$ ולכל $\psi \in W^*$.

כעת נניח ש־$W = V$ הוא מרחב מכפלה פנימית בעל מימד סופי ו־F הוא $\mathbb{R}$ או $\mathbb{C}$. מה הקשר בין T^o לבין $T^*\colon V \to V$ שמוגדרת על ידי $\langle T^*(v), u\rangle = \langle v, T(u)\rangle$?

נגדיר העתקה $\Phi\colon V \to V^*$ על ידי

$$(\Phi(v), u) = (\Phi(v))(u) = \langle v, u\rangle \qquad \text{לכל } u \in V.$$

(א) הראה ש־Φ מוגדרת היטב (כלומר, שאכן $\Phi(v)$ הוא פונקציונל לינארי על V), חד חד ערכית ועל.

(ב) הראה שהתרשים הבא של העתקות חילופי

$$\begin{tikzcd} V \arrow[r, "\Phi"] \arrow[d, "T^*"] & V^* \arrow[d, "T^o"] \\ V \arrow[r, "\Phi"] & V^* \end{tikzcd}$$

במלים אחרות, אם מזהים את V עם V^* בעזרת Φ, אז T^* מזוהה עם T^o. ∎

הוכחה: (א) $\Phi(v)$ הוא פונקציונל לינארי על V כי מכפלה פנימית היא לינארית במשתנה השני.

נראה ש־Φ חד חד ערכית: אם $\Phi(v)=\Phi(v')$ אז $\langle v,u\rangle=\langle v',u\rangle$ לכל $u\in V$. מכאן $\langle v-v',u\rangle$ לכל $u\in V$, ובפרט $\langle v-v',v-v'\rangle$. לכן $v-v'=0$, כלומר, $v=v'$.

נראה ש־Φ על: יהי $\varphi\in V^*$ נבחר בסיס אורתונורמלי $u_1,\dots,u_n$ של V. נגדיר $v=\sum_{i=1}^n\overline{\varphi(u_i)}u_i$.

אז

$$\big(\Phi(v)\big)(u_j)=\langle v,u_j\rangle=\langle\sum_{i=1}^n\overline{\varphi(u_i)}u_i,u_j\rangle=\sum_{i=1}^n\varphi(u_i)\langle u_i,u_j\rangle=\varphi(u_i)$$

לכל i, לכן $\Phi(v)=\varphi$.

(ב) יהיו $u,v\in V$. אז

$$\big((T^o\circ\Phi)(v)\big)(u)=\Big(T^o\big(\Phi(v)\big)\Big)(u)=\Big(T^o\big(\Phi(v)\big),u\Big)=\Big(\Phi(v),T(u)\Big)=$$
$$\langle v,T(u)\rangle=\langle T^*(v),u\rangle=\big(\Phi(T^*(v))\big)(u)=\big((\Phi\circ T^*)(v)\big)(u)$$

לכל $u\in V$, לכן $(T^0\circ\Phi)(v)=(\Phi\circ T^*)(v)$ לכל $v\in V$. מכאן $T^0\circ\Phi=\Phi\circ T^*$. ∎

אם $F=\mathbb{R}$, אז Φ העתקה לינארית. אם $F=\mathbb{C}$, אז Φ איננה לינארית. היא שומרת חיבור ומקיימת $\Phi(av)=\overline{a}v$.

14. תבניות בילינאריות

בפרק זה יהיו V, W מרחבים וקטוריים ממימד סופי מעל שדה F.

הגדרה 14.1: **תבנית בילינארית על** $V \times W$ היא העתקה $f\colon V \times W \to F$ המקיימת

$$f(v+v', w) = f(v,w) + f(v',w),$$
$$f(\alpha v, w) = \alpha f(v,w),$$
$$f(v, w+w') = f(v,w) + f(v,w'),$$
$$f(v, \alpha w) = \alpha f(v,w),$$

לכל $\alpha \in F$, $v, v' \in V$, $w, w' \in W$.

תבנית בילינארית על $V \times V$ נקראת גם, בקיצור, **תבנית בילינארית על** V.

דוגמאות 14.2:

(א) תבנית האפס: $f(v,w) = 0$ לכל $v \in V, w \in W$.

(ב) מכפלה פנימית במרחב מכפלה פנימית מעל $F = \mathbb{R}$ היא תבנית בילינארית.

(ג) אם φ פונקציונל לינארי על V (כלומר, העתקה לינארית $\varphi\colon V \to F$) ו־ψ פונקציונל לינארי על W, אז $f\colon V \times W \to F$, הנתונה על ידי $f(v,w) = \varphi(v)\psi(w)$, היא תבנית בילינארית.

(ד) יהיו $V = F^m, W = F^n$ (מרחבי עמודות). תהי $A \in \mathrm{M}_{m\times n}(F)$. נתאים לה העתקה $f\colon F^m \times F^n \to F$ על ידי $f(v,w) = v^t A w \in \mathrm{M}_1(F) = F$ (מזהים $\mathrm{M}_1(F)$ עם F). במפורש: אם $A = (a_{ij})$, $v = (x_1, \ldots, x_m)^t$, $w = (y_1, \ldots, y_n)^t$, אז

$$f(v,w) = (x_1, \ldots, x_m) \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ -- & \cdots & -- \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} =$$
$$(x_1, \ldots, x_m) \begin{pmatrix} \sum_{j=1}^n a_{1j} y_j \\ \vdots \\ \sum_{j=1}^n a_{mj} y_j \end{pmatrix} = \sum_{i=1}^m x_i \Big(\sum_{j=1}^n a_{ij} y_j \Big) = \sum_{i=1}^m \sum_{j=1}^n x_i a_{ij} y_j \qquad \blacksquare$$

נגדיר על אוסף התבניות הבילינאריות על $V \times W$ חיבור וכפל בסקלר:

$$(f+g)(v,w) = f(v,w) + g(v,w),$$
$$(\alpha f)(v,w) = \alpha f(v,w).$$

קל לבדוק שאם f, g תבניות בילינאריות, אז גם $f+g, \alpha f$ תבניות בילינאריות. יתר על כן, אוסף התבניות הבילינאריות על $V \times W$ הוא מרחב וקטורי מעל F ביחס לפעולות הנ"ל. נסמנו $\mathrm{Bil}(V,W)$.

(למעשה, $\mathrm{Bil}(V,W)$ הוא תת מרחב של המרחב הוקטורי של כל הפונקציות מ־$V \times W$ ל־F.)

משפט 14.3: *יהיו $\mathcal{B}=(v_1,\dots,v_m)$ בסיס של V ו־ $\mathcal{C}=(w_1,\dots,w_n)$ בסיס של W.*

(א) *לכל $f\in \mathrm{Bil}(V,W)$ קיימת מטריצה יחידה $A=\theta^{\mathcal{B}}_{\mathcal{C}}(f)\in \mathrm{M}_{m\times n}(F)$ כך שמתקיים*

$$f(v,w)=\big([v]_{\mathcal{B}}\big)^t A[w]_{\mathcal{C}} \qquad \text{לכל } v\in V,\ w\in W. \tag{1}$$

*(אומרים ש־A **מייצגת את** f לפי $\mathcal{B},\mathcal{C}$.)*

(ב) *המטריצה $A=\theta^{\mathcal{B}}_{\mathcal{C}}(f)$ מוגדרת על ידי*

$$(A)_{ij}=f(v_i,w_j) \qquad 1\le i\le m,\ 1\le j\le n. \tag{2}$$

(ג) *תהי $A\in \mathrm{M}_{m\times n}(F)$. נגדיר $f\colon V\times W\to F$ על ידי (1). אז f תבנית בילינארית ו־ $A=\theta^{\mathcal{B}}_{\mathcal{C}}(f)$.*

(ד) *התאמה $f\mapsto A$ הנ"ל היא איזומורפיזם של מרחבים וקטוריים $\theta^{\mathcal{B}}_{\mathcal{C}}\colon \mathrm{Bil}(V,W)\to \mathrm{M}_{m\times n}(F)$.*

הוכחה: (א), (ב) אם A מקיימת (1), אז $f(v_i,w_j)=\big([v_i]_{\mathcal{B}}\big)^t A[w_j]_{\mathcal{C}}=\mathbf{e}_i^t A\mathbf{e}_j=(A)_{ij}$, לכל i,j. מכאן (ב) והיחידות של A.

קיום A: נגדיר את A על ידי (2). נניח $v=\sum_{i=1}^m x_iv_i,\ w=\sum_{j=1}^n y_jw_j$. אז

$$f(v,w)=f\Big(\sum_{i=1}^m x_iv_i,\sum_{j=1}^n y_jw_j\Big)=\sum_{i=1}^m x_if\Big(v_i,\sum_{j=1}^n y_jw_j\Big)=\sum_{i=1}^m\sum_{j=1}^n x_iy_jf(v_i,w_j),$$

$$\big([v]_{\mathcal{B}}\big)^t A[w]_{\mathcal{C}}=(x_1,\dots,x_m)A\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}=\sum_{i=1}^m\sum_{j=1}^n x_i(A)_{ij}y_j=\sum_{i=1}^m\sum_{j=1}^n x_iy_jf(v_i,w_j).$$

(השוויון האמצעי בשורה השנייה נובע מהחישוב שעשינו בדוגמה 14.2(ד).)

(ג) נקבע $v\in V$. אז f הינה לינארית כפונקציה של w, כי $w\mapsto [w]_{\mathcal{C}}$ היא לינארית (איזומורפיזם של מרחבים וקטוריים) והכפלה במטריצה משמאל היא גם לינארית. באופן דומה f היא גם לינארית במשתנה הראשון v. לכן f תבנית בילינארית.

לפי היחידות בחלק (א), $A=\theta^{\mathcal{B}}_{\mathcal{C}}(f)$.

(ד) לפי (ג), $\theta^{\mathcal{B}}_{\mathcal{C}}$ על. נראה שהיא לינארית וחד חד ערכית.

טענה: *$\theta^{\mathcal{B}}_{\mathcal{C}}$ לינארית.* תהיינה $f,g\in \mathrm{Bil}(V,W)$, ותהיינה $A=\theta^{\mathcal{B}}_{\mathcal{C}}(f),A'=\theta^{\mathcal{B}}_{\mathcal{C}}(g)$. יהי $\alpha\in F$. אז

$$(f+g)(v_i,w_j)=f(v_i,w_j)+g(v_i,w_j)=(A)_{ij}+(A')_{ij}=(A+A')_{ij},$$
$$(\alpha f)(v_i,w_j)=\alpha f(v_i,w_j)=\alpha(A)_{ij}=(\alpha A)_{ij},$$

כלומר, $\theta^{\mathcal{B}}_{\mathcal{C}}(f+g)=A+A'$, $\theta^{\mathcal{B}}_{\mathcal{C}}(\alpha f)=\alpha A$.

טענה: *$\theta_{\mathcal{C}}^{\mathcal{B}}$ חד חד ערכית.* אם $\theta_{\mathcal{C}}^{\mathcal{B}}(f) = \theta_{\mathcal{C}}^{\mathcal{B}}(g) = A$, אז לפי (א) מתקיים לכל $v \in V, w \in W$

$$f(v,w) = \left([v]_{\mathcal{B}}\right)^t A[w]_{\mathcal{C}} = g(v,w).$$

לכן $f = g$. ∎

מסקנה 14.4: *(א)* $\dim \mathrm{Bil}(V,W) = \dim V \cdot \dim W$.

(ב) $\dim \mathrm{Bil}(V,V) = (\dim V)^2$.

משפט 14.5: *יהיו $\mathcal{B}, \mathcal{B}'$ שני בסיסים של V ויהיו $\mathcal{C}, \mathcal{C}'$ שני בסיסים של W. תהיינה $A = \theta_{\mathcal{C}}^{\mathcal{B}}(f)$, $A' = \theta_{\mathcal{C}'}^{\mathcal{B}'}(f)$. אז $A' = P^t A Q$, באשר P מטריצת המעבר מ־$\mathcal{B}$ ל־$\mathcal{B}'$, ו־Q מטריצת המעבר מ־$\mathcal{C}$ ל־$\mathcal{C}'$.*

הוכחה: נזכור שמתקיים $[v]_{\mathcal{B}} = P[v]_{\mathcal{B}'}$ לכל $v \in V$, ובאותו אופן $[w]_{\mathcal{C}} = Q[w]_{\mathcal{C}'}$ לכל $w \in W$. לכן לפי (1)

$$f(v,w) = \left([v]_{\mathcal{B}}\right)^t A[w]_{\mathcal{C}} = \left(P[v]_{\mathcal{B}'}\right)^t A\left(Q[w]_{\mathcal{C}'}\right) = \left([v]_{\mathcal{B}'}\right)^t P^t A Q[w]_{\mathcal{C}'}.$$

מכאן, לפי היחידות במשפט 14.3(א), $A' = P^t A Q$. ∎

הגדרה 14.6: מטריצות $A, A' \in \mathrm{M}_{m\times n}(F)$ נקראות **שקולות** אם יש $P \in \mathrm{M}_m(F), Q \in \mathrm{M}_n(F)$ הפיכות כך ש־$A' = P^t A Q$ (או: כך ש־$A' = PAQ$).

מטריצות ריבועיות $A, A' \in \mathrm{M}_n(F)$ נקראות **חופפות** אם יש $P \in \mathrm{M}_n(F)$ הפיכה כך ש־$A' = P^t A P$.

∎

אזהרה: השקילות כאן **איננה** שקילות־שורות. זוהי שקילות שורות ועמודות. ∎

תרגיל 14.7: *חפיפה ושקילות הם יחסי שקילות. בפרט: אם A חופפת ל־B ו־B חופפת ל־C אז A חופפת ל־C.*

הוכחה: נראה ששקילות היא יחס שקילות. תהיינה $A, B, C \in \mathrm{M}_{m\times n}(F)$.

רפלקסיביות: $A = (I_m)^t A I_n$.

סימטריות: אם $B = P^t A Q$, אז $A = (P^t)^{-1} B Q^{-1} = (P^{-1})^t B Q^{-1}$.

טרנזיטיביות: אם $B = P_1^t A Q_1$, ו־$C = P_2^t B Q_2$, אז $C = P_2^t (P_1^t A Q_1) Q_2 = (P_1 P_2)^t A (Q_1 Q_2)$.

באופן דומה חפיפה היא יחס שקילות (כתוֹב P במקום Q ו־n במקום m בכל מקום). ∎

מסקנה 14.8: *(א) $A, A' \in \mathrm{M}_n(F)$ חופפות אם ורק אם הן מייצגות אותה תבנית בילינארית על מרחב וקטורי V ממימד n (כ"א ביחס לבסיס מתאים).*

(ב) *$A, A' \in \mathrm{M}_{m\times n}(F)$ שקולות אם ורק אם הן מייצגות אותה תבנית בילינארית (כ"א ביחס לזוג בסיסים מתאים).*

(ג) *$A, A' \in \mathrm{M}_{m\times n}(F)$ שקולות אם ורק אם $\mathrm{rk}\, A = \mathrm{rk}\, A'$.*

הוכחה: (א) לפי משפט 14.5, אם $A = \theta_{\mathcal{B}}^{\mathcal{B}}(f)$ ו־$A' = \theta_{\mathcal{B}'}^{\mathcal{B}'}(f)$, אז $A' = P^t A P$, באשר P מטריצת המעבר מ־$\mathcal{B}$ ל־$\mathcal{B}'$. לכן A, A' חופפות.

להיפך, נניח $A' = P^tAP$, באשר $P \in \mathrm{M}_n(F)$ הפיכה. יהי $\mathcal{B}$ בסיס כלשהו של V. אזי קיים בסיס יחיד $\mathcal{B}'$ של V כך ש־P מטריצת המעבר מ־$\mathcal{B}$ ל־$\mathcal{B}'$. לפי משפט 14.3(ג) יש $f \in \mathrm{Bil}(V,V)$ כך ש־$A = \theta_{\mathcal{B}}^{\mathcal{B}}(f)$. לפי משפט 14.5, $A' = \theta_{\mathcal{B}'}^{\mathcal{B}'}(f)$.

(ב) דומה ל־(א).

(ג) $\Leftarrow$: הכפלה מימין או משמאל במטריצה הפיכה אינה משנה את הדרגה.

$\Rightarrow$: נניח כי $r = \mathrm{rk}\, A = \mathrm{rk}\, A'$. נראה תחילה ש־$A$ שקולה ל־$\begin{pmatrix} I_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}$.

נתבונן ב־$A^t \in \mathrm{M}_{n\times m}(F)$. קיימת $Q_1 \in \mathrm{M}_n(F)$ הפיכה כך ש־Q_1A^t מדורגת קנונית. אז $\mathrm{rk}\, Q_1A^t = \mathrm{rk}\, A^t = \mathrm{rk}\, A$, ולכן $n-r$ השורות האחרונות של Q_1A^t הן 0. לכן $n-r$ העמודות האחרונות של $(Q_1A^t)^t = AQ_1^t$ הן 0.

כעת קיימת $P_1 \in \mathrm{M}_m(F)$ הפיכה כך ש־$P_1AQ_1^t$ מדורגת קנונית. גם היא מדרגה r וגם $n-r$ העמודות האחרונות שלה הן 0. לכן r העמודות הראשונות שלה הן המובילות. מכאן $P_1AQ_1^t = (e_1,\dots,e_r,0\dots,0)$.
נסמן $Q = Q_1^t$ ו־$P = P_1^t$. אז P,Q הפיכות ו־$P^tAQ = (e_1,\dots,e_r,0\dots,0) = \begin{pmatrix} I_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}$.

לכן A ו־$\begin{pmatrix} I_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}$ שקולות. באותו אופן A' ו־$\begin{pmatrix} I_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}$ שקולות. לכן A ו־A' שקולות. ∎

טענה 14.9: *תהיינה $A, A' \in \mathrm{M}_n(F)$ חופפות. אז $\mathrm{rk}\, A' = \mathrm{rk}\, A$ ו־$\det A' = c^2 \det A$ עבור איזה $0 \neq c \in F$.*

הוכחה: לפי ההנחה קיימת $P \in \mathrm{M}_n(F)$ הפיכה כך ש־$A' = P^tAP$. אז גם P^t הפיכה וכפל במטריצה הפיכה אינו משנה את הדרגה, לכן $\mathrm{rk}\, A' = \mathrm{rk}\, P^tAP = \mathrm{rk}\, A$.

כמו כן, $\det A' = \det P^t \det A \det P = (\det P)^2 \det A$. ∎

תבניות בילינאריות סימטריות ותבניות ריבועיות.

הגדרה 14.10: תבנית $f \in \mathrm{Bil}(V,V)$ נקראת **סימטרית** אם $f(v,w) = f(w,v)$ לכל $v,w \in V$. עבור f כזאת הפונקציה $q\colon V \to F$, המוגדרת על ידי $q(v) = f(v,v)$, נקראת **תבנית ריבועית**. ∎

למה 14.11: *תהי $f \in \mathrm{Bil}(V,V)$, יהי $\mathcal{B}$ בסיס של V ותהי A המטריצה של f לפי $\mathcal{B}$. אזי f סימטרית אם ורק אם A סימטרית (כלומר, $A = A^t$).*

הוכחה: נזכור שלפי ההגדרה $(A)_{ij} = f(v_i,v_j)$, לכל $1 \le i,j \le n$, באשר $\mathcal{B} = (v_1,\dots,v_n)$.

אם f סימטרית, אז $(A^t)_{ij} = (A)_{ji} = f(v_j,v_i) = f(v_i,v_j) = (A)_{ij}$, לכן $A^t = A$.

אם A סימטרית, אז לפי (1)

$$f(w,v) = \big([w]_{\mathcal{B}}\big)^t A[v]_{\mathcal{B}} = \Big(\big([w]_{\mathcal{B}}\big)^t A[v]_{\mathcal{B}}\Big)^t = \big([v]_{\mathcal{B}}\big)^t A^t[w]_{\mathcal{B}} = \big([v]_{\mathcal{B}}\big)^t A[w]_{\mathcal{B}} = f(v,w).$$

∎

למה 14.12: *תהי $f \in \mathrm{Bil}(V,V)$ סימטרית ותהי q התבנית הריבועית המתאימה לה. אזי*

(א) *(למת הקיטוב:)* $2f(v,w) = f(v+w,v+w) - f(v,v) - f(w,w) = q(v+w) - q(v) - q(w)$ *לכל* $v,w \in V$.

(ב) *אם* $\mathrm{char}\, F \neq 2$ *ו־* $f \neq 0$*, אז יש* $u \in V$ *כך ש־* $f(u,u) \neq 0$.

הוכחה: (א)

$$f(v+w,v+w) = f(v,v) + f(v,w) + f(w,v) + f(w,w) = f(v,v) + 2f(v,w) + f(w,w).$$

(ב) נניח בשלילה ש־$f(u,u) = 0$ לכל $u \in V$. לפי (א), $2f(v,w) = 0$ לכל $v,w \in W$. אבל $2 \neq 0$ ב־F, לכן $f(v,w) = 0$ לכל $v,w \in W$. מכאן $f = 0$, סתירה. ∎

משפט 14.13: *תהי $f \in \mathrm{Bil}(V,V)$ סימטרית. אם $\mathrm{char}\, F \neq 2$, קיים בסיס $\mathcal{B} = (u_1,\dots,u_n)$ של V, כך ש־$f(u_i,u_j) = 0$ לכל $i \neq j$. בפרט*

(א) *המטריצה של f לפי $\mathcal{B}$ היא אלכסונית $\mathrm{Diag}(c_1,\dots,c_n)$, באשר $c_i = f(u_i,u_i)$ לכל i.*

(ב) *אם $v = \sum_{i=1}^n a_i v_i, w = \sum_{i=1}^n b_i v_i$, אז $f(v,w) = \sum_{i=1}^n a_i b_i c_i$.*

הוכחה: טענות (א), (ב) נובעות מהטענה הראשונה לפי משפט 14.3.

נוכיח את קיום $\mathcal{B}$ באינדוקציה על $n = \dim V$.

עבור $n = 1$ אין מה להוכיח. נניח כי $n > 1$. אם $f = 0$, נבחר $\mathcal{B}$ כלשהו. אם $f \neq 0$, לפי הלמה יש $u_1 \in V$ כך ש־$c_1 = f(u_1,u_1) \neq 0$. בפרט $u_1 \neq 0$. נסמן $V_1 = \mathrm{Sp}(u_1)$ ויהי

$$V_2 = \{v \in V \mid f(u_1,v) = 0\}.$$

ברור ש־V_2 תת מרחב של V (בדקו: $0 \in V_2$ ו־V_2 סגורה תחת החיבור ותחת הכפל בסקל!).

טענה: $V = V_1 \oplus V_2$. אכן, יהי $v \in V$ כך ש־$v = v_1 + v_2$, באשר $v_1 \in V_1, v_2 \in V_2$. נראה שהצגה זו של v הינה יחידה: יש $a \in F$ כך ש־$v_1 = au_1$. אז

$$f(u_1,v) = f(u_1,au_1+v_2) = af(u_1,u_1) + f(u_1,v_2) = ac_1 + 0 = ac_1$$

ומכאן $a = c_1^{-1} f(u_1,v)$, ולכן

$$v_1 = c_1^{-1} f(u_1,v) u_1, \qquad v_2 = v - c_1^{-1} f(u_1,v) u_1. \tag{3}$$

מכאן היחידות.

קיום ההצגה: יהי $v \in V$. נגדיר v_1, v_2 על פי (3). אז $v_1 \in V_1$, $v = v_1 + v_2$, ו־$v_2 \in V_2$: אכן, $f(u_1,v_2) = f(u_1,v) - c_1^{-1} f(u_1,v) f(u_1,u_1) = f(u_1,v) - f(u_1,v) = 0$.

לפי הטענה $\dim V = \dim V_1 + \dim V_2$, ו־$\dim V_1 = 1, \dim V = n$, לכן $\dim V_2 = n-1$.
הצמצום של f ל־V_2 היא תבנית בילינארית סימטרית על V_2. לפי הנחת האינדוקציה קיים בסיס $u_2, \dots, u_n$ של V_2 כך ש־$f(u_i, u_j) = 0$ לכל $i \neq j$. לפי הטענה, $u_1, u_2, \dots, u_n$ בסיס של V, ולפי הגדרת V_2 מתקיים גם $f(u_1, u_i) = f(u_i, u_1) = 0$ לכל $i \neq 1$. ∎

מסקנה 14.14: *תהי $A \in \mathrm{M}_n(F)$ סימטרית ו־$\mathrm{char}\, F \neq 2$. אז A חופפת למטריצה אלכסונית.*

הוכחה: יהי V מרחב וקטורי ממימד n ויהי $\mathcal{B}$ בסיסו. לפי משפט 14.3 יש $f \in \mathrm{Bil}(V,V)$ כך ש־A היא המטריצה של f לפי $\mathcal{B}$. לפי למה 14.11, f סימטרית. לפי משפט 14.13 יש בסיס $\mathcal{B}'$ כך שהמטריצה A' לפיו היא אלכסונית. כיון ש־A, A' מייצגות אותה תבנית f, הן חופפות. ∎

הגדרה 14.15: **הדרגה** $\mathrm{rk}\, f$ של $f \in \mathrm{Bil}(V,V)$ סימטרית היא הדרגה של המטריצה המייצגת את f לפי בסיס כלשהו של V. (ההגדרה טובה: המטריצות המייצגות את f לפי בסיסים שונים הן חופפות ולכן שוות דרגה.) ∎

משפט 14.16: *יהי V מרחב וקטורי מעל $\mathbb{C}$ ותהי $f \in \mathrm{Bil}(V,V)$ סימטרית. אז*

(א) *יש בסיס ל־V כך שהמטריצה של f לפיו היא מהצורה $D_r = \mathrm{Diag}(\overbrace{1, \dots, 1}^{r}, 0, \dots, 0)$.*

(ב) *$r = \mathrm{rk}\, f$ (ולכן נקבע על ידי f).*

הוכחה: (א) לפי משפט 14.13 יש בסיס $u_1, \dots, u_n$ ל־V כך ש־$f(u_i, u_j) = 0$, עבור $i \neq j$. נסמן $c_i := f(u_i, u_i)$. בלי הגבלת הכלליות, אחרי סידור מחדש של אברי הבסיס, $c_1, \dots, c_r \neq 0, \quad c_{r+1}, \dots, c_n = 0$.
נגדיר בסיס אחר $u'_1, \dots, u'_n$ של V על ידי $u'_i = \begin{cases} \frac{1}{\sqrt{c_i}} u_i & 1 \leq i \leq r \\ u_i & r < i \leq n \end{cases}$. אז ברור ש־$f(u'_i, u'_j) = 0$ לכל $i \neq j$. אם $1 \leq i \leq r$ אז

$$f(u'_i, u'_i) = f(\frac{1}{\sqrt{c_i}} u_i, \frac{1}{\sqrt{c_i}} u_i) = (\frac{1}{\sqrt{c_i}})^2 f(u_i, u_i) = \frac{1}{c_i} c_i = 1$$

ואם $r < i$ אז $f(u'_i, u'_i) = f(u_i, u_i) = c_i = 0$. לכן f מיוצגת לפי $u'_1, \dots, u'_n$ על ידי D_r.
(ב) $\mathrm{rk}\, f = \mathrm{rk}\, D_r = r$. ∎

מסקנה 14.17: *(א) כל מטריצה סימטרית מעל $\mathbb{C}$ חופפת למטריצה יחידה מהצורה D_r.*

(ב) *תהיינה $A, B \in \mathrm{M}_m(\mathbb{C})$ סימטריות. אז A, B חופפות $\Leftrightarrow \mathrm{rk}\, A = \mathrm{rk}\, B$.*

הוכחה: (ב) $\Leftarrow$: טענה 14.9; $\Rightarrow$: נניח $\mathrm{rk}\, A = \mathrm{rk}\, B = r$. אז A, B חופפות ל־D_r, ולכן חופפות. ∎

משפט 14.18: *יהי V מרחב וקטורי מעל $\mathbb{R}$ ותהי $f \in \mathrm{Bil}(V,V)$ סימטרית. אז יש בסיס ל־V כך שהמטריצה של f לפיו היא אלכסונית. אם במטריצה זו p רכיבים חיוביים ו־q שליליים, אז יש בסיס ל־V כך שהמטריצה של f לפיו היא $D_{p,q} = \mathrm{Diag}(\overbrace{1, \dots, 1}^{p}, \overbrace{-1, \dots, -1}^{q}, 0, \dots, 0)$.*

הוכחה: לפי משפט 14.13 יש בסיס $u_1, \dots, u_n$ ל־V כך ש־$f(u_i, u_j) = 0$ לכל $i \neq j$. נסמן $c_i := f(u_i, u_i)$. המטריצה של f לפי בסיס זה היא אלכסונית: $\mathrm{Diag}(c_1, \dots, c_n)$. בלי הגבלת הכלליות, אחרי סידור מחדש של

אברי הבסיס,

$$c_1,\dots,c_p>0,\quad c_{p+1},\dots,c_{p+q}<0,\quad c_{p+q+1}=\dots=c_n=0.$$

נגדיר בסיס אחר $u'_1,\dots,u'_n$ של V על ידי
$$u'_i=\begin{cases}\frac{1}{\sqrt{c_i}}u_i & 1\le i\le p\\ \frac{1}{\sqrt{-c_i}}u_i & p<i\le p+q\\ u_i & p+q<i\le n\end{cases}$$

אזי

- ברור ש־$f(u'_i,u'_j)=0$ לכל $i\neq j$.
- אם $1\le i\le p$ אז $f(u'_i,u'_i)=(\frac{1}{\sqrt{c_i}})^2f(u_i,u_i)=\frac{1}{c_i}c_i=1$;
- אם $p<i\le p+q$ אז $f(u'_i,u'_i)=(\frac{1}{\sqrt{-c_i}})^2f(u_i,u_i)=\frac{1}{-c_i}c_i=-1$;
- ואם $p+q<i$ אז $f(u'_i,u'_i)=f(u_i,u_i)=0$.

לכן המטריצה של f לפי בסיס זה היא $D_{p,q}$. ∎

משפט 14.19 (משפט Sylvester): *יהי V מרחב וקטורי ממימד n מעל $\mathbb{R}$ ותהי $f\in\mathrm{Bil}(V,V)$ סימטרית. נניח ש־f מיוצגת לפי בסיס $u_1,\dots,u_n$ על ידי מטריצה אלכסונית $D'_{p,q}$ בעלת p רכיבים חיוביים ו־q רכיבים שליליים, ולפי בסיס $v_1,\dots,v_n$ על ידי מטריצה אלכסונית $D'_{s,t}$ בעלת s רכיבים חיוביים ו־t רכיבים שליליים. אז $p=s,q=t$.*

הוכחה: $p+q=\mathrm{rk}\,D'_{p,q}=\mathrm{rk}\,f=\mathrm{rk}\,D'_{s,t}=s+t$. לכן די להוכיח כי $p=s$.

נניח בשלילה כי, למשל, $p>s$. בלי הגבלת הכלליות $f(u_i,u_i)=(D'_{p,q})_{ii}>0$ עבור $1\le i\le p$ ו־$f(v_i,v_i)=(D'_{s,t})_{ii}\le 0$ עבור $s<i\le n$, אחרת נשנה את הסדר של האיברים בשני הבסיסים.

אז $u_1,\dots,u_p,v_{s+1},\dots,v_n$ סדרה ב־V בת יותר מ־n אברים, ולכן תלויה לינארית מעל $\mathbb{R}$.

לכן יש $a_1,\dots,a_p,b_{s+1},\dots,b_n$, לא כולם 0, כך ש־

$$\sum_{i=1}^p a_iu_i+\sum_{i=s+1}^n b_iv_i=0. \tag{4}$$

יתר על כן, מכאן נובע שיש $1\le i\le p$ כך ש־$a_i\neq 0$. אכן, אחרת לפי (4), $\sum_{i=s+1}^n b_iv_i=0$, ומכאן $b_{s+1}=\dots=b_n=0$, כי $v_{s+1},\dots,v_n$ בלתי תלויים לינארית, סתירה. כעת,

$$f(\sum_{i=1}^p a_iu_i,\sum_{i=1}^p a_iu_i)=\sum_{i=1}^p a_i^2f(u_i,u_i)>0,$$

כי $f(u_i,u_i)>0$ לכל $1\le i\le p$ ויש $1\le i\le p$ כך ש־$a_i\neq 0$.

מצד שני, לפי (4),

$$f(\sum_{i=1}^p a_iu_i,\sum_{i=1}^p a_iu_i)=f(-\sum_{i=s+1}^m b_iv_i,-\sum_{i=s+1}^n b_iv_i)=\sum_{i=s+1}^n b_i^2f(v_i,v_i)\le 0,$$

כי $f(v_i,v_i)\le 0$ לכל $s+1\le i\le n$. סתירה. לכן $p=s$. ∎

מסקנה 14.20: *תהי $A \in \mathrm{M}_n(\mathbb{R})$ מטריצה סימטרית ממשית. אז*

(א) *A חופפת למטריצה יחידה מהצורה $D_{p,q}$.*

(ב) *אם A חופפת ל־$D_{p,q}$, נגדיר $\sigma(A) = p - q$ להיות **החתימה** (**הסיגנטורה**) של A. אז $A, B \in \mathrm{M}_n(\mathbb{R})$ סימטריות הינן חופפות אם ורק אם יש להן אותה דרגה ואותה חתימה.*

הוכחה: (א) יהי V מרחב וקטורי ממימד n מעל $\mathbb{R}$ ויהי $\mathcal{B}$ בסיסו. לפי משפט 14.3 יש $f \in \mathrm{Bil}(V,V)$ כך ש־A היא המטריצה של f לפי $\mathcal{B}$. לפי למה 14.11, f סימטרית. לפי משפט 14.18 יש בסיס של V כך שהמטריצה לפיו היא $D_{p,q}$. כיוון ש־$A, D_{p,q}$ מייצגות אותה תבנית f, הן חופפות.

אם A חופפת גם ל־$D_{s,t}$, אז גם $D_{s,t}$ מייצגת את f לפי איזשהו בסיס של V. לפי משפט 14.19, $p = s, q = t$.

(ב) A חופפת ל־$D_{p,q}$, באשר $p + q = \mathrm{rk}(A)$ ו־$p - q = \sigma(A)$. מכאן $p = \frac{\sigma(A)+\mathrm{rk}(A)}{2}$, $q = \frac{\sigma(A)-\mathrm{rk}(A)}{2}$. באותו אופן B חופפת ל־$D_{s,t}$, באשר $s = \frac{\sigma(B)+\mathrm{rk}(B)}{2}$, $t = \frac{\sigma(B)-\mathrm{rk}(B)}{2}$. לכן אם ל־$A, B$ אותה דרגה ואותה חתימה, אז הן חופפות לאותה $D_{p,q} = D_{s,t}$, ולכן חופפות. להיפך, אם A, B חופפות, אז יש להן אותה דרגה לפי טענה 14.9 ואותה חתימה, כי הן חופפות לאותה $D_{p,q}$. ∎

הגדרה 14.21: יהי V מרחב וקטורי ממימד סופי מעל $\mathbb{R}$. תבנית $f \in \mathrm{Bil}(V,V)$ סימטרית נקראת

- **חיובית** (positive) אם $f(v,v) \geq 0$ לכל $v \in V$;
- **חיובית גמורה** (positive definite) אם $f(v,v) > 0$ לכל $0 \neq v \in V$. ∎

משפט 14.22: *יהי V מרחב וקטורי ממימד סופי מעל $\mathbb{R}$ תהי $f \in \mathrm{Bil}(V,V)$ סימטרית ותהי $A \in \mathrm{M}_n(\mathbb{R})$ המטריצה של f לפי איזשהו בסיס $\mathcal{B}$ של V. התנאי הבאים שקולים זה לזה:*

(1) *f חיובית גמורה, כלומר $f(v,v) > 0$ לכל $0 \neq v \in V$.*

(2) *A חופפת ל־$D = \mathrm{Diag}(c_1, \ldots, c_n)$, באשר $c_1, \ldots, c_n > 0$.*

(3) *A חופפת ל־I_n.*

(4) *$A = P^t P$, באשר $P \in \mathrm{M}_n(\mathbb{R})$ הפיכה.*

(5) *$\sigma(A) = n$ (=החתימה של A).*

(6) *f היא מכפלה פנימית על V.*

(7) *כל שרשי f_A ב־$\mathbb{C}$ הם ממשיים חיוביים.*

הוכחה:

(1) $\Leftarrow$(2): יש בסיס $u_1, \ldots, u_n$ של V כך שהמטריצה D של f לפיו היא אלכסונית. נזכור שמתקיים $(D)_{ij} = f(u_i, u_j)$ לכל i, j. לכן $D = \mathrm{Diag}(c_1, \ldots, c_n)$, באשר, לפי (1), $c_i = f(u_i, u_i) > 0$ לכל i. אך A, D מייצגות את f, לכן חופפות.

(2) $\Leftarrow$ (3): f מיוצגת על ידי D לפי איזה בסיס $u_1,\dots,u_n$. גם $\frac{1}{\sqrt{c_1}}u_1,\dots,\frac{1}{\sqrt{c_n}}u_n$ בסיס של V, ו־f מיוצגת על ידי I_n לפיו. לכן D, I_n חופפות.

(3) $\Leftarrow$ (1): יש בסיס $u_1,\dots,u_n$ לפיו f מיוצגת על ידי I_n, כלומר, $f(u_i,u_j)=\delta_{ij}$. יהי $0\neq v\in V$, נאמר, $v=\sum_{i=1}^n a_iu_i$. אז יש i כך ש־$a_i\neq 0$. מכאן $f(v,v)=\sum_{i,j}a_ia_j\delta_{ij}=\sum_{i=1}^n a_i^2>0$.

(3) $\Leftrightarrow$ (4): ברור, כי $P^tP=P^tI_nP$.

(3) $\Leftrightarrow$ (5): לפי הגדרת החתימה, $\sigma(A)=n$ אם ורק אם A חופפת ל־$D_{n,0}=I_n$.

(1) $\Leftrightarrow$ (6): לפי הגדרת מכפלה פנימית.

לבסוף, כדי להוכיח (2) $\Leftrightarrow$ (7), נזכור ש־A סימטרית, לכן לפי מסקנה 13.24 יש $P\in \mathrm{M}_n(\mathbb{R})$ אורתוגונלית כך ש־$P^tAP=P^{-1}AP$ אלכסונית, נאמר, $D'=\mathrm{Diag}(\lambda_1,\dots,\lambda_n)$. כעת, A,D' דומות, לכן יש להן אותו פולינום אופייני, $f_A=f_{D'}=(X-\lambda_1)\cdots(X-\lambda_n)$. אבל A,D' גם חופפות. לכן:

(2) $\Leftarrow$ (7): D,D' חופפות, ולכן לפי משפט סילבסטר $\lambda_1,\dots,\lambda_n>0$.

(7) $\Leftarrow$ (2): נקח $D=D'$. ∎

סימון 14.23: עבור $A=(a_{ij})\in \mathrm{M}_n(\mathbb{R})$ נסמן

$$A_k=\begin{pmatrix} a_{11} & \cdots & a_{1k}\\ - & \cdots & -\\ a_{k1} & \cdots & a_{kk}\end{pmatrix}\in \mathrm{M}_k(\mathbb{R}),\qquad k=1,\dots,n.$$

מטריצות אלה (או הדטרמיננטות שלהן) נקראות **המינורים הראשיים של** A. ∎

משפט 14.24: *תהי $A\in \mathrm{M}_n(\mathbb{R})$ סימטרית נניח כי $\det A_1,\dots,\det A_n\neq 0$. אזי A חופפת ל־$D_{p,q}$, באשר q הוא מספר החלפות הסימן בסדרה $1,\det A_1,\det A_2,\dots,\det A_n$ ו־p הוא מספר אי החלפות הסימן בסדרה זו.*

הוכחה: נכתוב $A=(a_{ij})$.

1. אז $a_{11}=\det A_1\neq 0$. נחסיר כפולות מתאימות של העמודה הראשונה מהעמודות האחרות ואחר כך נחסיר אותן כפולות של השורה הראשונה מהשורות האחרות של A. פעולות אלה אינן משנות את $\det A_1,\dots,\det A_n$, ו־A תעבור למטריצה חופפת (כי פעולה אלמנטרית על העמודות מתאימה להכפלה במטריצה אלמנטרית P מימין, ופעולה מקבילה על השורות מתאימה להכפלה ב־P^t). לכן בלי הגבלת הכלליות

$$A=\begin{pmatrix} a_{11} & 0 & \cdots & 0\\ 0 & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ 0 & a_{n2} & \cdots & a_{nn}\end{pmatrix}.$$

2. אז $a_{11}a_{22}=\det A_2\neq 0$. בפרט $a_{22}\neq 0$. נחסיר כפולות מתאימות של העמודה השניה מהעמודות שאחריה ואחר כך נחסיר אותן כפולות של השורה השניה מהשורות שמתחתיה של A. פעולות אלה אינן משנות את

$\det A_1, \ldots, \det A_n$ (וגם לא את השורה הראשונה ואת העמודה הראשונה של A), ו־A תעבור למטריצה חופפת.

לכן בלי הגבלת הכלליות

$$A = \begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a_{n3} & \cdots & a_{nn} \end{pmatrix}.$$

3. באינדוקציה, בלי הגבלת הכלליות, $A = \mathrm{Diag}(a_{11}, \ldots, a_{nn})$, באשר $a_{11}, \ldots, a_{nn} \neq 0$. לכן p הוא מספר החיוביים ו־q מספר השליליים בסדרה $a_{11}, \ldots, a_{nn}$. אבל $a_{kk} = \frac{\det A_k}{\det A_{k-1}}$, לכן

$$\det A_{k-1}, \det A_k \text{ בעלי סימן זהה} \Leftrightarrow \frac{\det A_k}{\det A_{k-1}} > 0 \Leftrightarrow a_{kk} > 0, \qquad k = 1, \ldots, n$$

(הגדרנו $\det A_0 = 1$, כדי שלמשוואה זו תהיה משמעות גם עבור $k = 1$). מכאן המסקנה. ∎

תרגיל 14.25: *תהי $A \in \mathrm{M}_n(\mathbb{R})$ סימטרית ויהי $0 \neq \alpha \in \mathbb{R}$. אז*

(א) $\sigma(\alpha A) = \begin{cases} \sigma(A) & \alpha > 0 \\ -\sigma(A) & \alpha < 0 \end{cases}$,

(ב) $f_{\alpha A}(X) = \alpha^n f_A(\alpha^{-1} X)$.

בפרט, אם $f_A(X) = \prod_{i=1}^n (X - \lambda_i)$, אז $f_{\alpha A}(X) = \prod_{i=1}^n (X - \alpha\lambda_i)$.

הוכחה: (א) לפי משפט 14.3 יש תבנית בילינארית f על $\mathbb{R}^n$ כך שהמטריצה של f לפי הבסיס הסטנדרטי היא A. לפי למה 14.11, f סימטרית. לפי משפט 14.18 יש בסיס $\mathcal{B}$ של $\mathbb{R}^n$ כך שהמטריצה של f לפיו היא $D_{p,q}$.

כעת, המטריצות של αf לפי הבסיס הסטנדרטי ולפי $\mathcal{B}$ הן αA ו־$\alpha D_{p,q}$, בהתאמה. אם $\alpha > 0$, אז $\alpha D_{p,q}$ אלכסונית, בעלת p רכיבים חיוביים ו־q רכיבים שליליים, לכן לפי משפט סילבסטר חופפת ל־$D_{p,q}$. לכן $\sigma(\alpha A) = p - q = \sigma(A)$. אם $\alpha < 0$, אז $\alpha D_{p,q}$ אלכסונית, בעלת q רכיבים חיוביים ו־p רכיבים שליליים, לכן לפי משפט סילבסטר חופפת ל־$D_{q,p}$. לכן $\sigma(\alpha A) = q - p = -\sigma(A)$.

(ב) $f_{\alpha A}(X) = \det(XI - \alpha A) = \alpha^n \det(\alpha^{-1} XI - A) = \alpha^n f_A(\alpha^{-1} X)$. ∎

15. מיון שניוניות

הגדרה 15.1: **שניונית** (quadric) X במרחב $\mathbb{R}^n$ היא קבוצה X של $v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ שהרכיבים $x_1, \ldots, x_n$ שלהם מקיימים משוואה מהצורה:

$$a + \sum_{i=1}^{n} 2b_i x_i + \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{1 \le i < j \le n} 2a_{ij} x_i x_j = 0 \tag{1}$$

ו־a_{ij} לא כולם 0. אגף שמאל של המשוואה הוא פולינום כלשהו ממעלה שניה ב־$x_1, \ldots, x_n$. (המקדמים 2 שרירותיים, תפקידם יתברר בהמשך.)

דוגמה 15.2: (א) משוואה $x_1^2 + x_2^2 + x_3^2 - 1 = 0$ מתארת שניונות ב־$\mathbb{R}^3$. זוהי סְפֵרָה (פני כדור).

אכן, $v = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ מקיימת את המשוואה אם ורק אם $||v|| = 1$.

(ב) משוואה $(x_1 + x_2 + 1)(x_1 + x_2 - 1) = 0$ מתארת זוג ישרים ב־$\mathbb{R}^2$ (או זוג מישורים ב־$\mathbb{R}^3$).

(ג) משוואה $x_1 x_2 - 1 = 0$ מתארת היפרבולה במישור (או היפרבולואיד ב־$\mathbb{R}^3$). ∎

הערה 15.3: שיטות רישום אחרות של שניוניות. אם X נתונה על ידי (1), נגדיר $a_{ji} = a_{ij}$ לכל $1 \le i < j \le n$, ואז (1) הופכת ל־

$$a + \sum_{i=1}^{n} 2b_i x_i + \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j = 0 \tag{1'}$$

נגדיר $u = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{R}^n$, $A = (a_{ij}) \in \mathrm{M}_n(\mathbb{R})$. אז A סימטרית, $A \neq 0$. את (1) אפשר לרשום כך (נזהה את $\mathrm{M}_1(\mathbb{R})$ עם $\mathbb{R}$):

$$a + 2u^t v + v^t A v = 0. \tag{2}$$

A תקרא **המטריצה המצומצמת** של השניונית.

מטריצת הגושים $\tilde{A} = \begin{array}{c c} & \begin{array}{cc} 1 & n \end{array} \\ \begin{array}{c} 1 \\ n \end{array} & \begin{pmatrix} a & u^t \\ u & A \end{pmatrix} \end{array} \in \mathrm{M}_{n+1}(\mathbb{R})$ תקרא **המטריצה המורחבת** של השניונית. גם היא סימטרית. את (2) אפשר לרשום בעזרתה כך:

$$\begin{pmatrix} 1 \\ v \end{pmatrix}^t \tilde{A} \begin{pmatrix} 1 \\ v \end{pmatrix} = 0. \tag{3}$$

אכן,

$$\begin{pmatrix} 1 \\ v \end{pmatrix}^t \tilde{A} \begin{pmatrix} 1 \\ v \end{pmatrix} = (1 \quad v^t) \begin{pmatrix} a & u^t \\ u & A \end{pmatrix} \begin{pmatrix} 1 \\ v \end{pmatrix} = (1 \quad v^t) \begin{pmatrix} a + u^t v \\ u + Av \end{pmatrix} =$$
$$= (a + u^t v + v^t u + v^t A v) = (a + 2u^t v + v^t A v)$$ ∎

דיון 15.4: פעולות על שניוניות. מטרתנו להגיע למשוואה פשוטה יותר של X על ידי סדרה של הפעולות הבאות:

(I) *שינוי בסיס אורתונורמלי:* נבחר בסיסי אורתונורמלי $\mathcal{B}$ של $\mathbb{R}^n$ ונחליף את מערכת הצירים המקורית בזו שנתונה על ידי $\mathcal{B}$. מטריצת המעבר $P \in \mathrm{M}_n(\mathbb{R})$ מהבסיס הסטנדרטי ל־$\mathcal{B}$ היא אורתוגונלית. אם $v' = [v]_{\mathcal{B}}$ אז $v = Pv'$, ולכן (2) שקולה ל־ $a + 2u^tPv' + (Pv')^tA(Pv') = 0$, כלומר

$$a + 2(P^tu)^tv' + (v')^t(P^tAP)v' = 0. \qquad (2')$$

אגף שמאל הוא פולינום ממעלה שניה בקואורדינטות של v'. המטריצה המצומצמת של השניונית בקואורדינטות החדשות היא $A' = P^tAP$, והמורחבת

$$\tilde{A}' = \begin{pmatrix} a & u^tP \\ P^tu & P^tAP \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & P^t \end{pmatrix} \begin{pmatrix} a & u^t \\ u & A \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix} = \tilde{P}^t\tilde{A}\tilde{P},$$

באשר $\tilde{P} = \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix} \in \mathrm{M}_{n+1}(\mathbb{R})$ הפיכה, ואפילו אורתוגונלית, כי $\tilde{P}^t\tilde{P} = I_{n+1}$ לפי כללי הכפל של מטריצות גושים.

(II) *הזזת צירים:* נזיז את ראשית הצירים ל־$w \in \mathbb{R}^n$. אם v'' הוא וקטור הקואורדינטות של v במערכת המוזזת, אז $v'' = v - w$, לכן (2) שקולה ל־ $a + 2u^t(w + v'') + (w + v'')^tA(w + v'') = 0$, כלומר,

$$(a + 2u^tw + w^tAw) + 2(u + Aw)^tv'' + v''^tAv'' = 0. \qquad (2'')$$

זוהי שוב משוואה של שניונית, מהצורה (2), עם המטריצה המצומצמת A והמטריצה המורחבת

$$\begin{pmatrix} a + 2u^tw + w^tAw & u^t + w^tA \\ u + Aw & A \end{pmatrix} = \begin{pmatrix} 1 & w^t \\ 0 & I_n \end{pmatrix} \begin{pmatrix} a & u^t \\ u & A \end{pmatrix} \begin{pmatrix} 1 & 0 \\ w & I_n \end{pmatrix} = \tilde{P}^t\tilde{A}\tilde{P},$$

באשר $\tilde{P} = \begin{pmatrix} 1 & 0 \\ w & I_n \end{pmatrix} \in \mathrm{M}_{n+1}(\mathbb{R})$ הפיכה (אך לא אורתוגונלית, אם $w \neq 0$).

(III) *הכפלה בקבוע:* נכפיל את המשוואה של השניונית בסקלר $0 \neq c \in \mathbb{R}$. המשוואה המתקבלת מתארת את X, והמטריצה המצומצמת והמורחבת שלה הן cA, $c\tilde{A}$. ∎

מסקנה 15.5: *שינוי בסיס אורתונורמלי או הזזת צירים מעביר את המשוואה של השניונית למשוואה בקואורדינטות אחרות, אבל הגדלים הבאים אינם משתנים:*

(1) $\mathrm{rk}(A), \mathrm{rk}(\tilde{A})$,

(2) $\sigma(A), \sigma(\tilde{A})$,

(3) *הסימן של* $\det A$, $\det \tilde{A}$.

(4) *הערכים העצמיים של* A.

הכפלה בקבוע $0 \neq c$ *שומרת (1) ומכפילה את (4) ב־*c. *אם* $c > 0$, *היא שומרת גם (2), (3). אם* $c < 0$, *היא מכפילה (2) ב־*-1, *ו־(3) ב־* $(-1)^n$, $(-1)^{n+1}$, *בהתאמה.*

הוכחה: לפי תרגיל 14.25. ∎

הגדרה 15.6: מטריצה $\tilde{P} \in \mathrm{M}_{n+1}(\mathbb{R})$ תקרא **אפינית** אם $\tilde{P} = \begin{array}{c c} & \begin{array}{cc} 1 & n \end{array} \\ \begin{array}{c} 1 \\ n \end{array} & \begin{pmatrix} 1 & 0 \\ w & P \end{pmatrix} \end{array} \in \mathrm{M}_{n+1}(\mathbb{R})$ באשר $P \in \mathrm{M}_n(\mathbb{R})$ אורתוגונלית ו־$w \in \mathbb{R}^n$ עמודה.

למה 15.7: *קבוצת המטריצות האפיניות ב־$\mathrm{M}_{n+1}(\mathbb{R})$ היא* **חבורה**, *כלומר, היא סגורה תחת כפל ולקיחת הופכי ומכילה את I_{n+1}.*

הוכחה: ברור ש־I_{n+1} אפינית (הציבו $P = I_n, w = 0$). לפי כללי הכפל של מטריצות גושים

$$\begin{pmatrix} 1 & 0 \\ w & P \end{pmatrix} \begin{pmatrix} 1 & 0 \\ w' & P' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ w + Pw' & PP' \end{pmatrix}$$

ואם P, P' אורתוגונליות, אז PP' אורתוגונלית. לכן קבוצת המטריצות האפיניות סגורה תחת הכפל. כמו כן

$$\begin{pmatrix} 1 & 0 \\ w & P \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -P^{-1}w & P^{-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ w - PP^{-1}w & PP^{-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & I_n \end{pmatrix} = I_{n+1}$$

ואם P אורתוגונליות, אז P^{-1} אורתוגונליות. לכן מטריצה אפינית הפיכה וההופכית שלה היא מטריצה אפינית.

∎

למה 15.8: *תהיינה $X, X' \subseteq \mathbb{R}^n$ שניוניות ותהיינה $\tilde{A}, \tilde{A}'$ המטריצות המורחבות שלהן, בהתאמה. אז X' מתקבלת מ־X על ידי סדרה של שינויי בסיס אורתונורמלי או הזזות צירים אם ורק אם $\tilde{A}' = \tilde{P}^t \tilde{A} \tilde{P}$, באשר $\tilde{P} \in \mathrm{M}_{n+1}(\mathbb{R})$ אפינית.*

הוכחה: "$\Leftarrow$": בדיון לעיל ראינו שהתנאי מתקיים אם X' מתקבלת מ־X על ידי פעולה אחת. אכן, אם הפעולה היא שינוי בסיס אורתונורמלי, נקח P להיות מטריצת המעבר ו־$w = 0$; אם הפעולה היא הזזת צירים, ניקח $P = I_n$ ו־w וקטור ההזזה של הראשית.

נניח ש־X' מתקבלת מ־X על ידי k פעולות. אז יש $\tilde{P}_1, \ldots, \tilde{P}_k \in \mathrm{M}_{n+1}(\mathbb{R})$ אפיניות כך ש־

$$\tilde{A}' = \tilde{P}_k^t(\cdots(\tilde{P}_2^t(\tilde{P}_1^t \tilde{A} \tilde{P}_1)\tilde{P}_2)\cdots)\tilde{P}_k = (\tilde{P}_1\tilde{P}_2\cdots\tilde{P}_k)^t \tilde{A}(\tilde{P}_1\tilde{P}_2\cdots\tilde{P}_k)$$

ובאינדוקציה על k, המטריצה $\tilde{P}_1\tilde{P}_2\cdots\tilde{P}_k$ אפינית, לפי למה 15.7.

"$\Rightarrow$": נניח כי $\tilde{P} = \begin{pmatrix} 1 & 0 \\ w & P \end{pmatrix}$. אז $\tilde{P} = \begin{pmatrix} 1 & 0 \\ w & P \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ w & I_n \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix}$. נעשה הזזת הראשית ל־$w$ ואחר כך שינוי בסיס לפי P. אז, לפי הדיון לעיל, X עוברת לשניונית בעלת מטריצה מורחבת $\tilde{A}'$, כלומר, ל־X'. ∎

משפט 15.9: *תהי X שניונית. תהיינה A המטריצה המצומצמת ו־$\tilde{A}$ המטריצה המורחבת שלה. נסמן $\sigma = \sigma(A), r = \mathrm{rk}(A), \tilde{\sigma} = \sigma(\tilde{A}), \tilde{r} = \mathrm{rk}(\tilde{A})$. אזי*

(א) *על ידי סדרה של שינויי בסיס אורתונורמלי והזזות צירים עוברת X לשניונית בעלת משוואה יחידה מהצורה*

$$\lambda_1 x_1^2 + \cdots + \lambda_s x_s^2 = \Phi \quad , \text{באשר} \tag{4}$$

$$1 \le s \le n \quad , \lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_s \quad \text{שונים מ־}0 , \tag{5}$$

ו־ $\Phi = \alpha \in \mathbb{R}$ או $\Phi = 2\beta x_{s+1}$, $\beta > 0$.

(ב) *מתקיים $s = r$ ו־ $\lambda_1, \ldots, \lambda_r$ הם הערכים העצמיים של A השונים מ־ 0 (על ריבוייהם האלגבריים).*

(ג) $$\Phi = \begin{cases} 0 & \tilde{r} = r \\ \alpha > 0 & \tilde{r} = r+1, \quad \tilde{\sigma} = \sigma - 1 \\ \alpha < 0 & \tilde{r} = r+1, \quad \tilde{\sigma} = \sigma + 1 \\ 2\beta x_{r+1} & \tilde{r} = r+2 \end{cases}.$$

הוכחה: (א) **קיום:** לפי משפט 13.23 יש P אורתוגונלית כך ש־$P^t A P = D = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$ אלכסונית. אז בדיוק r איברים מתוך $\lambda_1, \ldots, \lambda_n$, שונים מאפס. בלי הגבלת הכלליות אלה הם $\lambda_1, \ldots, \lambda_r$ והם מסודרים בסדר יורד (אחרת נשנה את סדר העמודות ב־P). לכן מתקיים (5) עם $s = r$ ו־$\lambda_{r+1} = \cdots = \lambda_n = 0$. לכן, לפי דיון 15.4, בלי הגבלת הכלליות $A = D$. נסמן $A_0 = \mathrm{Diag}(\lambda_1, \ldots, \lambda_r) \in \mathrm{M}_r(\mathbb{R})$.

נבחר $w = \begin{pmatrix} -\frac{b_1}{\lambda_1} \\ \vdots \\ -\frac{b_r}{\lambda_r} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$. אז $u + Aw = \begin{pmatrix} b_1 \\ \vdots \\ b_r \\ b_{r+1} \\ \vdots \\ b_n \end{pmatrix} - \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \begin{pmatrix} \frac{b_1}{\lambda_1} \\ \vdots \\ \frac{b_r}{\lambda_r} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b_{r+1} \\ \vdots \\ b_n \end{pmatrix}$.

לכן, לפי דיון 15.4, בלי הגבלת הכלליות $b_1 = \cdots = b_r = 0$.

אם $u \neq 0$, אז $0 \neq u' = \begin{pmatrix} b_{r+1} \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{R}^{n-r}$. יהי $\beta = ||u'|| > 0$. אז $||-\frac{1}{\beta}u'|| = 1$, לכן אפשר להשלים את $-\frac{1}{\beta}u'$ לבסיס אורתונורמלי של $\mathbb{R}^{n-r}$. לכן לפי משפט 13.13 יש $Q \in \mathrm{M}_{n-r}(\mathbb{R})$ אורתוגונלית כך ש־$-\frac{1}{\beta}u'$ היא העמודה הראשונה שלה, כלומר, $Q \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = -\frac{1}{\beta}u'$. מכאן $Q^t u' = Q^{-1} u' = \begin{pmatrix} -\beta \\ 0 \\ \vdots \\ 0 \end{pmatrix}$. אז

$P = \begin{pmatrix} I_r & 0 \\ 0 & Q \end{pmatrix}$ אורתוגונלית ומתקיים

$$P^t A P = \begin{pmatrix} I_r & 0 \\ 0 & Q^t \end{pmatrix} \begin{pmatrix} A_0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} I_r & 0 \\ 0 & Q \end{pmatrix} = \begin{pmatrix} A_0 & 0 \\ 0 & 0 \end{pmatrix} = A,$$

$$P^t u = \begin{pmatrix} I_r & 0 \\ 0 & Q^t \end{pmatrix} \begin{pmatrix} 0 \\ u' \end{pmatrix} = \begin{pmatrix} 0 \\ Q^t u' \end{pmatrix} = -\beta \mathbf{e}_{r+1}.$$

לכן, אחרי שינוי בסיס אורתונורמלי, בלי הגבלת הכלליות $u = 0$ או $u = -\beta \mathbf{e}_{r+1}$, באשר $\beta > 0$.

- אם $u = 0$, קיבלנו את (4) עם $\Phi = \alpha \in \mathbb{R}$.
- אם $u = -\beta \mathbf{e}_{r+1}$ ו־$a = 0$, קיבלנו את (4) עם $\Phi = 2\beta x_{r+1}$.
- אם $u = -\beta \mathbf{e}_{r+1}$ ו־$a \neq 0$, נגדיר $w = \frac{a}{2\beta}\mathbf{e}_{r+1}$. אז $Aw = 0$ ו־

$$a' = a + 2u^t w + w^t A w = a + 2(-\beta \mathbf{e}_{r+1})^t \frac{a}{2\beta}\mathbf{e}_{r+1} = a + 2(-\beta)\frac{a}{2\beta}\mathbf{e}_{r+1}^t \mathbf{e}_{r+1} = 0,$$

לכן, אחרי הזזת צירים ב־w זה, בלי הגבלת הכלליות $a = 0$ (ושאר המרכיבים של המטריצה המורחבת של השניונית לא ישתנו).

(ב) המטריצה המצומצמת של (4) היא $\mathrm{Diag}(\lambda_1, \ldots, \lambda_s, 0, \ldots, 0)$. היא דומה ל־$A$. לכן:

- לשתיהן אותה דרגה, ולכן $s = r$;
- לשתיהן אותו פולינום אופייני, ולכן $\lambda_1, \ldots, \lambda_r, 0, \ldots, 0$ הם הערכים העצמיים של A (על ריבוייהם).

(ג) המטריצה המורחבת של (4) היא, לפי המקרים השונים של Φ,

$$\begin{pmatrix} 0 & 0 & \cdots & & & 0 \\ 0 & \lambda_1 & & & & \\ \vdots & & \ddots & & & \\ & & & \lambda_s & & \\ & & & & 0 & \\ & & & & & \ddots \\ 0 & & & & & & 0 \end{pmatrix} \quad \begin{pmatrix} -\alpha & 0 & \cdots & & & 0 \\ 0 & \lambda_1 & & & & \\ \vdots & & \ddots & & & \\ & & & \lambda_s & & \\ & & & & 0 & \\ & & & & & \ddots \\ 0 & & & & & & 0 \end{pmatrix} \quad \begin{pmatrix} -\alpha & 0 & \cdots & & & 0 \\ 0 & \lambda_1 & & & & \\ \vdots & & \ddots & & & \\ & & & \lambda_s & & \\ & & & & 0 & \\ & & & & & \ddots \\ 0 & & & & & & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & & -\beta \\ 0 & \lambda_1 & & \\ & & \ddots & \\ & & \lambda_s & \\ -\beta & & & 0 \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}$$

$\Phi = 0$	$\Phi = \alpha > 0$	$\Phi = \alpha < 0$	$\Phi = 2\beta x_{r+1}$
$\tilde{r} = r, r = s$	$\tilde{r} = r+1, r = s$	$\tilde{r} = r+1, r = s$	$\tilde{r} = r+2, r = s$
$\tilde{\sigma} = \sigma$ (לא חשוב)	$\tilde{\sigma} = \sigma - 1$	$\tilde{\sigma} = \sigma + 1$	$\tilde{\sigma} = \sigma$ (לא חשוב)

ומכאן, לפי מסקנה 15.5, נובע (ג).

(דבר זה דורש הסבר נוסף: למעשה הוכחנו, למשל, שאם $\Phi = 0$ אז $\tilde{r} = r$, בעוד שצריך להוכיח שאם $\tilde{r} = r$ אז $\Phi = 0$. אבל הטענה האחרונה בכל זאת נובעת מכאן. אכן, אם $\tilde{r} = r$, אז לא יתכן כי $\Phi = \alpha > 0$ או $\Phi = \alpha < 0$ או $\Phi = 2\beta x_{r+1}$, כי בשלושת המקרים האלה $\tilde{r} \neq r$. לכן, בהכרח, $\Phi = 0$. באופן דומה נובעים שאר המקרים.)

(א) **יחידות:** לפי (ב) נקבעים באופן יחיד s, $\lambda_1, \ldots, \lambda_s$. לפי (ג) נקבע "הסוג" של Φ. נותר רק להוכיח שאם X יכולה לעבור

(i) לשניונית (4) עם $\Phi = \alpha$ וגם לשניונית (4) עם $\Phi = \alpha'$, אז $\alpha = \alpha'$;

(ii) לשניונית (4) עם $\Phi = 2\beta x_{r+1}$ וגם לשניונית (4) עם $\Phi = 2\beta' x_{r+1}$, באשר $\beta, \beta' > 0$, אז $\beta = \beta'$.

תהיינה $\tilde{A}, \tilde{A}'$ המטריצות המורחבות של שתי השניוניות האלה. כיוון שאפשר לעבור על ידי סדרה של שינויי בסיס אורתונורמלי והזזות צירים מאחת לשניה, לפי למה 15.8 יש $\tilde{P} = \begin{pmatrix} 1 & 0 \\ w & P \end{pmatrix}$ עם $w = (w_1, \ldots, w_n)^t \in \mathbb{R}^n$ ו־P אורתוגונלית כך ש־$\tilde{A}' = \tilde{P}^t \tilde{A} \tilde{P}$. נבדיל בין שני המקרים:

(i) $\tilde{A} = \begin{pmatrix} -\alpha & 0 \\ 0 & A \end{pmatrix}, \tilde{A}' = \begin{pmatrix} -\alpha' & 0 \\ 0 & A \end{pmatrix}$, באשר $A = \mathrm{Diag}(\lambda_1, \ldots, \lambda_r, 0, \ldots, 0)$. לכן

$$\begin{pmatrix} -\alpha' & 0 \\ 0 & A \end{pmatrix} = \begin{pmatrix} 1 & w^t \\ 0 & P^t \end{pmatrix} \begin{pmatrix} -\alpha & 0 \\ 0 & A \end{pmatrix} \begin{pmatrix} 1 & 0 \\ w & P \end{pmatrix} = \begin{pmatrix} -\alpha + w^t A w & w^t A P \\ P^t A w & P^t A P \end{pmatrix}.$$

מכאן $P^t A w = 0$ ולכן $A w = 0$ (כי P^t הפיכה). לכן $w^t A w = 0$. מכאן $\alpha = \alpha'$.

(ii) $\tilde{A} = \begin{pmatrix} 0 & u^t \\ u & A \end{pmatrix}, \tilde{A}' = \begin{pmatrix} 0 & (u')^t \\ u' & A \end{pmatrix}$, באשר $u = -\beta \mathbf{e}_{r+1}$, $u' = -\beta' \mathbf{e}_{r+1}$ ו־A כמו

קודם. בלי הגבלת הכלליות $0 < \beta' \leq \beta$. אז

$$\begin{pmatrix} 0 & (u')^t \\ u' & A \end{pmatrix} = \begin{pmatrix} 1 & w^t \\ 0 & P^t \end{pmatrix} \begin{pmatrix} 0 & u^t \\ u & A \end{pmatrix} \begin{pmatrix} 1 & 0 \\ w & P \end{pmatrix} = \begin{pmatrix} 2u^t w + w^t A w & u^t P + w^t A P \\ P^t u + P^t A w & P^t A P \end{pmatrix}.$$

מכאן $u' = P^t(u + Aw)$. כיוון ש־P^t אורתוגונלית, ההעתקה $v \mapsto P^t v$ היא אוניטרית (מסקנה 13.6). לפי משפט 13.12 היא משמרת נורמה. לכן $||u'|| = ||P^t(u + Aw)|| = ||u + Aw||$. מכאן

$$(\beta')^2 = ||u'||^2 = ||u + Aw||^2 = || - \beta \mathbf{e}_{r+1} + (\lambda_1 w_1, \ldots, \lambda_r w_r, 0, \ldots, 0)^t||^2 =$$
$$= ||(\lambda_1 w_1, \ldots, \lambda_r w_r, -\beta, 0, \ldots, 0)^t||^2 = \lambda_1^2 w_1^2 + \ldots + \lambda_r^2 w_r^2 + \beta^2 \geq \beta^2.$$

לכן $\beta' \geq \beta$. אבל מקודם $\beta' \leq \beta$. לכן $\beta' = \beta$. ∎

אם נרשה גם הכפלה בקבוע (סקלר שונה מאפס) הערכים העצמיים של המטריצה המצומצמת לא יישמרו, אך אפשר לקבל $\alpha = \pm 1$ או $\alpha = 0$, $\beta = 1$ (תיאור "מנורמל"):

משפט 15.10: *תהי X שניונית. נסמן $r = \mathrm{rk}(A), \sigma = \sigma(A), \tilde{r} = \mathrm{rk}(\tilde{A}), \tilde{\sigma} = \sigma(\tilde{A})$.*

(א) *על ידי סדרה של שינויי בסיס אורתונורמלי, הזזות צירים והכפלות בקבוע אפשר להעביר את המשוואה של X למשוואה מהצורה*

$$(4') \qquad \lambda_1 x_1^2 + \cdots + \lambda_p x_p^2 \quad + \quad \lambda_{p+1} x_{p+1}^2 + \cdots + \lambda_{p+q} x_{p+q}^2 = \Phi, \qquad \text{באשר}$$

$$(5') \qquad \lambda_1, \ldots, \lambda_p > 0 > \lambda_{p+1}, \ldots, \lambda_{p+q}, \quad p \geq 1, p \geq q \geq 0 \quad \text{ו־} \quad \Phi = 0 \quad \text{או} \quad \Phi = \pm 1 \quad \text{או} \quad \Phi = 2x_{p+q+1}.$$

יתר על כן, משוואה $(4')$ – תחת תנאי $(5')$ – היא יחידה במובן הבא:

(ב1) $p = \frac{r+|\sigma|}{2}, \quad q = \frac{r-|\sigma|}{2}$;

(ב2) *יש $0 \neq \gamma \in \mathbb{R}$ כך ש־ $\gamma\lambda_1, \ldots, \gamma\lambda_{p+q}, 0, \ldots, 0$ הם הערכים העצמיים של A (על ריבוייהם);*

(ג)
$$\Phi = \begin{cases} 0 & \tilde{r} = r \\ +1 & \tilde{r} = r+1, \quad |\sigma| \neq 0, \quad |\tilde{\sigma}| = |\sigma| - 1 \\ -1 & \tilde{r} = r+1, \quad |\sigma| \neq 0, \quad |\tilde{\sigma}| = |\sigma| + 1 \\ \pm 1^{*)} & \tilde{r} = r+1, \quad |\sigma| = 0 \\ 2x_{p+q+1} & \tilde{r} = r+2 \end{cases}.$$

*הערה: כאשר $|\sigma| = 0$ (כלומר, $p = q$) ואפשר להגיע לצורה $(4')$ עם $\Phi = 1$, אז אפשר להגיע לצורה $(4')$ גם עם $\Phi = -1$.

אכן, אם ל־X משוואה

$$\lambda_1 x_1^2 + \cdots + \lambda_p x_p^2 + \lambda_{p+1} x_{p+1}^2 + \cdots + \lambda_{2p} x_{2p}^2 = 1,$$

אז גם המשוואה הבאה מתארת את X:

$$(-\lambda_{2p}) x_{2p}^2 + \cdots + (-\lambda_{p+1}) x_{p+1}^2 + (-\lambda_p) x_p^2 + \cdots + (-\lambda_1) x_1^2 = -1$$

וגם היא מהצורה $(4')$. ∎

הוכחת המשפט: (א) על ידי הכפלת המשוואה ב־-1, אם צריך, אפשר להניח בלי הגבלת הכלליות ש־$\sigma \geq 0$. (דבר זה אינו משנה את r ומחליף את σ ב־$|\sigma|$.) לכן לפי המשפט הקודם אפשר לעבור על ידי שינויי בסיס אורתונורמלי והזזות צירים למשוואה (4) עם תנאי (5) ו־$\Phi = \alpha \in \mathbb{R}$ או $\Phi = 2\beta x_{p+q+1}$, $\beta > 0$. פעולות אלה אינן משנות את r ואת $\sigma \geq 0$. לכן $s = r$ ואם נסמן ב־p את מספר איברים החיוביים והשליליים בסדרת הערכים העצמיים $\lambda_1, \ldots, \lambda_s$, אז $p - q = |\sigma| \geq 0$, לכן $p \geq q$, ו־$p + q = r \geq 1$, לכן $p \geq 1$. אז מתקיים (בו). נכפיל את (4) בסקלר מתאים $\gamma > 0$ ($\gamma = \frac{1}{|\alpha|}$, אם $\Phi = \alpha \neq 0$ או $\gamma = \frac{1}{\beta}$, אם $\Phi = 2\beta x_{p+q+1}$) ונקבל Φ כמבוקש.

(ב) שינויי בסיס אורתונורמלי, הזזות צירים, והכפלות בקבוע אינן משנות את $r, |\sigma|$ ואת סדרת הערכים העצמיים של A, עד כדי כפל בסקלר.

(ג) כמו במשפט הקודם. נבחן את המטריצה המורחבת של $(4')$ לפי המקרים השונים של Φ:

$$\begin{pmatrix} 0 & 0 & \cdots & & & 0 \\ 0 & \lambda_1 & & & & \\ \vdots & & \ddots & & & \\ & & & \lambda_{p+q} & & \\ & & & & 0 & \\ & & & & & \ddots \\ 0 & & & & & & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 & \cdots & & & 0 \\ 0 & \lambda_1 & & & & \\ \vdots & & \ddots & & & \\ & & & \lambda_{p+q} & & \\ & & & & 0 & \\ & & & & & \ddots \\ 0 & & & & & & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & & & 0 \\ 0 & \lambda_1 & & & & \\ \vdots & & \ddots & & & \\ & & & \lambda_{p+q} & & \\ & & & & 0 & \\ & & & & & \ddots \\ 0 & & & & & & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & & & -1 \\ 0 & \lambda_1 & & & \\ & & \ddots & & \\ & & & \lambda_{p+q} & \\ -1 & & & & 0 \\ & & & & & \ddots \\ & & & & & & 0 \end{pmatrix}$$

$\Phi = 0$	$\Phi = 1$	$\Phi = -1$	$\Phi = 2x_{r+1}$
$r = p + q$	$r = p + q$	$r = p + q$	$r = p + q$
$\tilde{r} = r$	$\tilde{r} = r + 1$	$\tilde{r} = r + 1$	$\tilde{r} = r + 2$
$\sigma = p - q \geq 0$	$\sigma = p - q \geq 0$	$\sigma = p - q \geq 0$	$\sigma = p - q \geq 0$
$\tilde{\sigma} = \sigma$ (לא חשוב)	$\tilde{\sigma} = \sigma - 1$	$\tilde{\sigma} = \sigma + 1$	$\tilde{\sigma} = \sigma$ (לא חשוב)
	אם $\sigma \neq 0$ אז	אם $\sigma \neq 0$ אז	
	$\lvert\tilde{\sigma}\rvert = \lvert\sigma\rvert - 1$	$\lvert\tilde{\sigma}\rvert = \lvert\sigma\rvert + 1$	

נזכור ש־$|\sigma|, |\tilde{\sigma}|$ נשמרים. כמו במשפט הקודם נסיק את (ג) מכאן. ∎

תרגיל 15.11: *בסימונים של משפט 15.10 נניח* $\det A \neq 0$, $\det \tilde{A} \neq 0$. *אז*

(א) $r = n$, $\tilde{r} = n + 1$, *ולכן* $\Phi = +1$ *או* $\Phi = -1$.

נניח גם כי $\sigma = \sigma(A) > 0$ *והדטרמיננטות של המינורים הראשיים של* A *שונות מאפס. הוכיחו:*

(ב) *אם* $\det A, \det \tilde{A}$ *שוות סימן, אז* $\Phi = -1$.

(ג) *אם* $\det A, \det \tilde{A}$ *מנוגדות סימן, אז* $\Phi = +1$.

הוכחה: (א) ברור.

(ב), (ג) נחליף את העמודה הראשונה של $\tilde{A}$ עם יתר העמודות ואחר כך נחליף את השורה הראשונה עם יתר השורות. על ידי כך נקבל מ־$\tilde{A} = \begin{pmatrix} a & u^t \\ u & A \end{pmatrix}$ מטריצה חופפת $\tilde{A}' := \begin{pmatrix} A & u \\ u^t & a \end{pmatrix}$ ובפרט בעלת אותה חתימה וגם אותו סימן של דטרמיננטה. כיוון ש־Φ נקבע לפי החתימות, אפשר להחליף את $\tilde{A}$ ב־$\tilde{A}'$.

סדרת המינורים הראשיים של $\tilde{A}'$ היא סדרת המינורים הראשיים של A, בתוספת $\det \tilde{A}'$. לפי משפט 14.24, $\sigma(\tilde{A}') = \sigma + 1$ אם $\det \tilde{A}', \det A$ שוות סימן, ו־$\sigma(\tilde{A}') = \sigma - 1$ אם $\det \tilde{A}', \det A$ מנוגדות סימן. מכאן המסקנה לפי משפט 15.10. ∎

הטבלה הבאה מתארת את הצורה הגיאומטרית של שניונות בעלת משוואה קנונית. באתרים

http://en.wikipedia.org/wiki/Quadric

http://tutorial.math.lamar.edu/Classes/CalcIII/QuadricSurfaces.aspx

מופיעים גם ציורים של חלק מצורות אלה.

טבלה של מיון שניוניות

מקרא: $r = \mathrm{rk}(A),\ \tilde{r} = \mathrm{rk}(\tilde{A}),\ \sigma = \mathrm{Sign}(A),\ \tilde{\sigma} = \mathrm{Sign}(\tilde{A})$, $\alpha, \beta, \gamma \in \mathbb{R}$ שונים מאפס.

- ב־$\mathbb{R}^2$:

שם השניונית	$\lvert\tilde{\sigma}\rvert$	$\tilde{r}$	$\lvert\sigma\rvert$	r	תיאור קנוני (מנורמל)
נקודה	2	2	2	2	$\alpha^2x^2 + \beta^2y^2 = 0$
אליפסה	1	3	2	2	$\alpha^2x^2 + \beta^2y^2 = 1$
קבוצה ריקה	3	3	2	2	$\alpha^2x^2 + \beta^2y^2 = -1$
שני ישרים נחתכים	0	2	0	2	$\alpha^2x^2 - \beta^2y^2 = 0$
היפרבולה	1	3	0	2	$\alpha^2x^2 - \beta^2y^2 = \pm 1$
ישר (כפול)	1	1	1	1	$\alpha^2x^2 = 0$
שני ישרים מקבילים	0	1	1	1	$\alpha^2x^2 = 1$
קבוצה ריקה	2	2	1	1	$\alpha^2x^2 = -1$
פרבולה	1	3	1	1	$\alpha^2x^2 = 2y$

- ב־ $\mathbb{R}^3$:

שם השניונית	$\lvert\tilde{\sigma}\rvert$	$\tilde{r}$	$\lvert\sigma\rvert$	r	תיאור קנוני (מנורמל)
נקודה	3	3	3	3	$\alpha^2x^2 + \beta^2y^2 + \gamma^2z^2 = 0$
אליפסואיד	2	4	3	3	$\alpha^2x^2 + \beta^2y^2 + \gamma^2z^2 = 1$
קבוצה ריקה	4	4	3	3	$\alpha^2x^2 + \beta^2y^2 + \gamma^2z^2 = -1$
חרוט אליפטי	1	3	1	3	$\alpha^2x^2 + \beta^2y^2 - \gamma^2z^2 = 0$
היפרבולואיד אליפטי בעל יריעה אחת	0	4	1	3	$\alpha^2x^2 + \beta^2y^2 - \gamma^2z^2 = 1$
היפרבולואיד אליפטי בעל שתי יריעות	2	4	1	3	$\alpha^2x^2 + \beta^2y^2 - \gamma^2z^2 = -1$
ישר	2	2	2	2	$\alpha^2x^2 + \beta^2y^2 = 0$
גליל אליפטי	1	3	2	2	$\alpha^2x^2 + \beta^2y^2 = 1$
קבוצה ריקה	3	3	2	2	$\alpha^2x^2 + \beta^2y^2 = -1$
פרבולואיד אליפטי	2	4	2	2	$\alpha^2x^2 + \beta^2y^2 = 2z$
שני מישורים נחתכים	0	2	0	2	$\alpha^2x^2 - \beta^2y^2 = 0$
גליל היפרבולי	1	3	0	2	$\alpha^2x^2 - \beta^2y^2 = \pm 1$
פרבולואיד היפרבולי	0	4	0	2	$\alpha^2x^2 - \beta^2y^2 = 2z$
מישור (כפול)	1	1	1	1	$\alpha^2x^2 = 0$
שני מישורים מקבילים	0	0	1	1	$\alpha^2x^2 = 1$
קבוצה ריקה	2	2	1	1	$\alpha^2x^2 = -1$
גליל פרבולי	1	3	1	1	$\alpha^2x^2 = 2y$

דוגמה 15.12: מה מתארת השניונית $-2+4x-4y+3x^2+2y^2=0$ ב־$\mathbb{R}^2$?

המטריצה המצומצמת שלה היא $A=\begin{pmatrix}3&0\\0&2\end{pmatrix}$ והמורחבת $\tilde{A}=\begin{pmatrix}-2&2&-2\\2&3&0\\-2&0&2\end{pmatrix}$. בפרט $\sigma(A)=2>0$. כמו כן $\det A=6>0$, $\det\tilde{A}=-32<0$, לכן לפי תרגיל 15.11, $\Phi=+1$, כלומר, אפשר להביא את השניונית לצורה

$$\lambda_1x^2+\lambda_2y^2=1,\qquad \lambda_1,\lambda_2>0$$

זוהי משוואה של אליפסה. יתר על כן, $(\lambda_1:\lambda_2)=(3:2)$.

דוגמה 15.13: מה מתארת השניונית $1+2z+4xy-4xz-2x^2+3y^2+2z^2=0$ ב־$\mathbb{R}^3$?

המטריצה המצומצמת שלה היא $A=\begin{pmatrix}-2&2&-2\\2&3&0\\-2&0&2\end{pmatrix}$ והמורחבת $\tilde{A}=\begin{pmatrix}1&0&0&1\\0&-2&2&-2\\0&2&3&0\\1&-2&0&2\end{pmatrix}$.

סדרת (הדטרמיננטות של) המינורים הראשיים של A היא $-2,-10,-32$, לכן $\sigma(A)=2-1=1$. כמו כן $\det A=-32<0$, $\det\tilde{A}=-22<0$, לכן לפי תרגיל 15.11, $\Phi=-1$, כלומר, אפשר להביא את השניונית לצורה

$$\alpha x^2+\beta y^2-\gamma z^2=-1,\qquad \alpha,\beta,\gamma>0.$$

זהו היפרבולואיד אליפטי בעל שתי ירעוח.

תרגילים

תרגיל 1: *תהיינה $A, B \in \mathrm{M}_n(\mathbb{R})$ סימטריות וחיוביות גמורות. הוכיחו:* $\det(A+B) \geq \det A + \det B$.

נניח A, B אלכסוניות, $A = \mathrm{Diag}(a_1, \ldots, a_n), B = \mathrm{Diag}(b_1, \ldots, b_n)$.

אז $A + B = \mathrm{Diag}(a_1 + b_1, \ldots, a_n + b_n)$. לכן

$$\det(A) = a_1 \cdots a_n, \quad \det(B) = b_1 \cdots b_n, \quad \det(A+B) = (a_1 + b_1) \cdots (a_n + b_n).$$

כיוון ש־A, B חיוביות גמורות, $b_i > 0$, $a_i > 0$ לכל i. מכאן ברור ש־$\det(A+B)$ הוא סכום של 2^n מחוברים חיוביים, ביניהם $\det(A)$ ו־$\det(B)$. לכן $\det(A+B) \geq \det A + \det B$.

כיוון ש־A חיובית גמורה, יש $P \in \mathrm{M}_n(\mathbb{R})$ הפיכה כך ש־$P^t A P = I_n$.

באותו אופן, יש $Q \in \mathrm{M}_n(\mathbb{R})$ הפיכה כך ש־$Q^t B Q = I_n$.

אם נחליף את A, B במטריצות $P^t A P, P^t B P$, באשר $P \in \mathrm{M}_n(\mathbb{R})$ הפיכה, ונוכיח את הטענה עבורן, אז הוכחנו בכך

$$\det(P^t A P + P^t B P) \geq \det(P^t A P) + \det(P^t B P)$$

$$(\det P)^2 \det(A+B) \geq (\det P)^2 \det A + (\det P)^2 \det B$$

ומכאן, אם נחלק ב־$(\det P)^2$, נובעת הטענה המקורית המבוקשת.

תחילה נמצא P הפיכה כך ש־$P^t A P = I_n$. לכן בלי הגבלת הכלליות $A = I_n$.

כעת, כיוון ש־B סימטרית, יש Q אורתוגונלית כך ש־$Q^t B Q = Q^{-1} B Q$ אלכסונית. אז עדיין $Q^t A Q = Q^t Q = I_n$. לכן בלי הגבלת הכלליות $A = I_n$ ו־B אלכסונית. כפי שראינו קודם, במקרה זה הטענה נכונה. ∎

תרגיל 2: *יהי V מרחב מכפלה פנימית בעל מימד סופי מאל $\mathbb{R}$. תהי $v_1, v_2, \ldots, v_m$ סדרה בלתי תלויה לינארית של איברים ב־V. הוכיחו שקיימת סדרה אורתונורמלית יחידה $u_1, u_2, \ldots, u_m$ ב־V כך ש־*

(א) $\mathrm{Sp}(u_1, \ldots, u_k) = \mathrm{Sp}(v_1, \ldots, v_k)$ *לכל* $1 \leq k \leq m$ ו־

(ב) $\langle u_k, v_k \rangle < 0$ *לכל* $1 \leq k \leq m$.

לפי תהליך גרם־שמידט קיימת סדרה אורתוגונלית (ולא בהכרח אורתונורמלית) $u_1, u_2, \ldots, u_m$ שמקיימת את (א). איבריה מוגדרים על ידי

$$u_k = v_k - (c_1 u_1 + c_2 u_2 + \cdots + c_{k-1} u_{k-1}), \qquad (*)$$

(עבור איזה $c_1, c_2, \ldots, c_{k-1} \in \mathbb{R}$) לכן

$$\langle u_k, v_k \rangle = \langle u_k, u_k + \sum_{i=1}^{k-1} c_i u_i \rangle = \langle u_k, u_k \rangle + \sum_{i=1}^{k-1} c_i \langle u_k, u_i \rangle = \langle u_k, u_k \rangle > 0$$

לכן צריך להכפיל במשהו אחר:

אם נחליף כל u_k ב־$u'_k := \frac{1}{||u_k||}u_k$ אז $u'_1, u'_2, \ldots, u'_m$ אורתונורמלית,
$\mathrm{Sp}(u_1, \ldots, u_k) = \mathrm{Sp}(u'_1, \ldots, u'_k)$ ולכן $\mathrm{Sp}(u'_1, \ldots, u'_k) = \mathrm{Sp}(v_1, \ldots, v_k)$ לכל $1 \leq k \leq m$,
ומתקיים

$$\langle u'_k, v_k \rangle = \frac{1}{||u_k||}\langle u_k, v_k \rangle > 0.$$

אם נחליף כל u_k ב־$u'_k := -\frac{1}{||u_k||}u_k$ אז $u'_1, u'_2, \ldots, u'_m$ אורתונורמלית,
$\mathrm{Sp}(u_1, \ldots, u_k) = \mathrm{Sp}(u'_1, \ldots, u'_k)$ ולכן $\mathrm{Sp}(u'_1, \ldots, u'_k) = \mathrm{Sp}(v_1, \ldots, v_k)$ לכל $1 \leq k \leq m$,
ומתקיים

$$\langle u'_k, v_k \rangle = -\frac{1}{||u_k||}\langle u_k, v_k \rangle < 0.$$

בכך הוכח הקיום.

היחידות:

נניח ש־ $u_1, u_2, \ldots, u_m$ וגם $u'_1, u'_2, \ldots, u'_m$ מקיימות את התנאים של התרגיל. אז, לכל k,

$$\mathrm{Sp}(u_1, \ldots, u_k) = \mathrm{Sp}(v_1, \ldots, v_k) = \mathrm{Sp}(u'_1, \ldots, u'_k)$$

לכן יש $a_1, \ldots, a_{k-1} \in \mathbb{R}$ כך ש־

$$u'_k = a_1u_1 + a_2u_2 + \cdots + a_ku_k.$$

צריך להוכיח: $a_1 = a_2 = \cdots, = a_{k-1} = 0, \quad a_k = 1$.

לפי למת האפיון, $a_i = \langle u_i, u'_k \rangle$, לכל $1 \leq i \leq k$.

כיוון ש־u'_k ניצב ל־$u'_1, u'_2, \ldots, u'_{k-1}$ ו־$u_1, \ldots, u_k \in \mathrm{Sp}(u'_1, \ldots, u'_k)$,
u'_k ניצב גם ל־$u_1, u_2, \ldots, u_{k-1}$. לכן $a_1 = a_2 = \cdots = a_{k-1} = 0$. מכאן

$$u'_k = a_ku_k.$$

$$1 = ||u'_k|| = |a_k| \cdot ||u_k|| = |a_k|$$

לכן $a_k = \pm 1$.

$$0 > \langle u'_k, v_k \rangle = a_k\langle u_k, v_k \rangle$$

אבל $\langle u_k, v_k \rangle < 0$, לכן מכאן $a_k > 0$. יחד עם $a_k = \pm 1$ זה נותן $a_k = 1$, כנדרש. ∎

תרגיל 3: *תהי $T\colon V \to V$ העתקה לינארית ויהי $p(X) \in F[X]$. הוכיחו ש־$p(T)$ איזומורפיזם אם ורק אם p, m_T זרים.*

פתרון: נניח ש־p, m_T זרים. אז יש $g, h \in F[X]$ כך ש־$gp + hm_T = 1$. מכאן

$$, g(T)p(T) + h(T)m_T(T) = 1_V$$

כלומר, $g(T)p(T) = 1_V$. לכן $p(T)$ העתקה לינארית הפיכה (ו־$g(T)$ ההופכי שלה).

נניח ש־p, m_T אינם זרים. אז $d = \gcd(p, m_T)$ ממעלה גדולה מאפס. יש $g \in F[X]$ כך ש־$gd = m_T$.

בהכרח $\deg g < \deg m_T$, לכן $g(T) \neq 0$. אז $g(T)d(T) = 0$, לכן $d(T)$ אינה הפיכה.

יש $h \in F[X]$ כך ש־$hd = p$. אז $h(T)d(T) = p(T)$ ו־$d(T)$ אינה הפיכה, לכן $p(T)$ אינה הפיכה.

∎

תרגיל 4: *(א) $\langle A, B\rangle = \operatorname{tr}(A^*B)$ היא מכפלה פנימית על $V = \mathrm{M}_n(\mathbb{C})$.*

(ב) תהי $M \in \mathrm{M}_n(\mathbb{C}) = V$. נתבונן בהעתקה $T\colon V \to V$ המוגדרת על ידי $T(A) = MA$. הוכיחו שיש $N \in \mathrm{M}_n(\mathbb{C})$ כך ש־$T^*(A) = NA$. מהי N?

פתרון: (א) בדיקה פשוטה:

(1)

$$\langle A, B + B'\rangle = \operatorname{tr}(A^*(B + B')) = \operatorname{tr}(A^*B + A^*B') = \operatorname{tr}(A^*B) + \operatorname{tr}(A^*B') =$$
$$= \langle A, B\rangle + \langle A, B'\rangle$$

(2)

$$\langle A, \alpha B\rangle = \operatorname{tr}(A^*(\alpha B)) = \operatorname{tr}(\alpha A^*B) = \alpha \operatorname{tr}(A^*B) = \alpha\langle A, B\rangle$$

(3)

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

(4)

$$\langle A, A\rangle = \operatorname{tr}(A^*A) = \sum_{i=1}^{n}(A^*A)_{ii} = \sum_{i=1}^{n}\sum_{k=1}^{n}(A^*)_{ik}(A)_{ki} =$$
$$\sum_{i=1}^{n}\sum_{k=1}^{n}\overline{(A)_{ki}}(A)_{ki} = \sum_{i=1}^{n}\sum_{k=1}^{n}|(A)_{ki}|^2 > 0$$

אם יש i, k כך ש־$(A)_{ki} \neq 0$, כלומר, אם $A \neq 0$.

(ב) צריך להוכיח שיש N שמקיימת $\langle MA, B\rangle = \langle A, NB\rangle$ לכל A, B, כלומר,

$\operatorname{tr}\left((MA)^*B\right) = \operatorname{tr}\left(A^*NB\right)$ לכל A, B, וזה אכן מתקיים עבור $N = M^*$, כי

$(MA)^*B = A^*M^*B = A^*NB$. ∎

תרגיל 5: *תהי $A \in \mathrm{M}_n(F)$ כך ש־$\operatorname{rk} A = 1, \operatorname{tr} A = 0$. מצא m_A, f_A.*

פתרון: תהי $T\colon F^n \to F^n$ ההכפלה ב־A. אז $\dim \operatorname{Im} T = \operatorname{rk} A = 1$, לכן $\dim \operatorname{Ker} T = n-1$. נבחר בסיס של $\operatorname{Ker} T$ ונשלים אותו לבסיס $\mathcal{B}$ של F^n. אז

$$C = [T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 0 & \cdots & 0 & a_{1n} \\ \vdots & \cdots & \vdots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

כיוון שלמטריצות דומות אותה העקבה, $0 = \operatorname{tr} A = \operatorname{tr} C = a_{nn}$. לכן $f_A = f_C = X^n$. לכן צורת ז'ורדן J של A היא $J = \operatorname{Diag}(J_{n_1}(0), \ldots, J_{n_r}(0))$. אך $\operatorname{rk} J = \operatorname{rk} A = 1$, לכן $n_1 = 2, n_2 = \cdots = n_r = 1$. מכאן $m_A = m_J = m_{J_2(0)} = X^2$.

אפשר גם לשים לב כי $C^2 = 0$, אחרי שקבענו כי $a_{nn} = 0$. מכאן בקלות ש־$m_C = X^2$. ∎

תרגיל 6: *אילו מבין המטריצות הבאות לכסינות?*

(א) $\begin{pmatrix} 1 & 0 & 2 & 3 & 5 \\ 0 & 4 & 1 & 0 & 1 \\ 2 & 1 & 2 & 3 & 8 \\ 3 & 0 & 3 & 1 & 1 \\ 5 & 1 & 8 & 1 & 1 \end{pmatrix}$ *מעל* $\mathbb{R}$, *(ב)* $\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ \pi & 4 & 0 & 0 & 0 \\ 2 & 1 & 2 & 0 & 0 \\ 3 & 0 & i & 3 & 0 \\ 0 & 1 & 8 & 1 & 5 \end{pmatrix}$ *מעל* $\mathbb{C}$.

(ג) $\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$ *מעל* $\mathbb{Q}$.

פתרון: (א) המטריצה סימטרית, לכן לכסינה לפי מסקנה 10.24.

(ב) הפולינום האופייני הוא מכפלה של גורמים לינאריים שונים, לכן המטריצה לכסינה לפי מסקנה 5.6.

(ג) המטריצה היא מטריצת שלושה גושים באלכסון, $\operatorname{Diag}(\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, (1), \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix})$. אם נחליף את סדר הגושים, נקבל מטריצה דומה.

לכן המטריצה הנתונה דומה למטריצת ז'ורדן שאיננה אלכסונית, ואילו מטריצה אלכסונית דומה למטריצת ז'ורדן אלכסונית. בגלל יחידות צורת ז'ורדן, המטריצה הנתונה אינה דומה לאלכסונית. ∎

תרגיל 7: *(א) האם* $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, B = \begin{pmatrix} -1 & 0 & 0 & 0 \\ -1 & 1 & 1 & -1 \\ -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$ *דומות?*

(ב) *האם* $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, B = \begin{pmatrix} -1 & 0 & 0 & 0 \\ -1 & 1 & 1 & -1 \\ -1 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \end{pmatrix}$ *דומות?*

פתרון: (א) לא, כי $3 = \operatorname{rk} A \neq \operatorname{rk} B = 2$.

(ב) A היא מטריצת ז'ורדן, $A = \mathrm{Diag}\Big((1),(-1),\begin{pmatrix}0&0\\1&0\end{pmatrix}\Big)$,
לכן $f_A = m_A = (X-1)(X+1)X^2$. ואילו

$$f_B = \begin{vmatrix} X+1 & 0 & 0 & 0 \\ 1 & X-1 & -1 & 1 \\ 1 & 0 & X & -1 \\ 1 & 0 & 0 & X \end{vmatrix} = (X+1)(X-1)\begin{vmatrix} X & -1 \\ 0 & X \end{vmatrix} = (X+1)(X-1)X^2$$

לכן $m_B = (X-1)(X+1)X^2$ או $m_B = (X-1)(X+1)X$.

$$(B-I_4)(B+I_4)B = \begin{pmatrix} -2 & 0 & 0 & 0 \\ -1 & 0 & 1 & -1 \\ -1 & 0 & -1 & 1 \\ -1 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 2 & 1 & -1 \\ -1 & 0 & 1 & 1 \\ -1 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 & 0 & 0 \\ -1 & 1 & 1 & -1 \\ -1 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 & 0 & 0 \\ -1 & 1 & 1 & -1 \\ -1 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \neq 0$$

לכן $m_B \neq (X-1)(X+1)X$ ומכאן $m_B = (X-1)(X+1)X^2$. לכן צורת ז'ורדן של B היא $\mathrm{Diag}\Big((1),(-1),\begin{pmatrix}0&0\\1&0\end{pmatrix}\Big)$ שהינה צורת ז'ורדן של A. כיוון ש־A, B אותה צורת ז'ורדן, הן דומות. ∎

תרגיל 8: $A \in \mathrm{M}_3(\mathbb{R})$ *אורתוגונלית,* $Av \neq v$ *לכל* $0 \neq v \in \mathbb{R}^3$. *מצא את* $\det A$.

פתרון: $A \in \mathrm{M}_3(\mathbb{C})$ אוניטרית ו־1 אינו ערך עצמי שלה. לכן A דומה מעל $\mathbb{C}$ ל־$C = \mathrm{Diag}(\lambda_1,\lambda_2,\lambda_3)$, באשר $1 \neq \lambda_i \in \mathbb{C}$, $|\lambda_i| = 1$ לכל i. אך $f_A = (X-\lambda_1)(X-\lambda_2)(X-\lambda_3)$ הוא ממשי, לכן אם λ שרש שלו אז גם $\overline{\lambda}$ שרש שלו. לכן או שני שרשים שלו מרוכבים וצמודים זה לזה ושרש שלישי הוא -1, או ששלושתם -1. בשני המקרים מכפלת שלושת השרשים היא -1. לכן $\det A = -1$. ∎

פתרון אחר: 1 אינו ערך עצמי של A.

יש העתקה לינארית $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ כך שהמטריצה של T לפי הבסיס הסטנדרטי היא A. כיוון ש־A אורתוגונלית, T אורתוגונלית. לכן יש בסיס אורתונורמלי $\mathcal{B}$ כך ש־

$$[T]_{\mathcal{B}} = \mathrm{Diag}(\overbrace{1,\dots,1}^{p},\overbrace{-1,\dots,-1}^{q},A_1,\dots,A_r)$$

באשר $A_1,\dots,A_r \in \mathrm{M}_2(\mathbb{R})$ מטריצות מהצורה $\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$ שהינן בעלות דטרמיננטה $+1$. כיוון ש־A דומה למטריצה זו ולמטריצות דומות יש אותה דטרמיננטה, בלי הגבלת הכלליות A היא מהצורה הזאת.

לכן $\det(A) = (-1)^q$.

אבל A מסדר 3×3, לכן $p+q+2r = 3$. אם $p > 0$, אז 1 ערך עצמי של A, סתירה להנחה. לכן $p = 0$.

לכן $q = 3-2r$ אי זוגי. מכאן $\det(A) = (-1)^q = -1$. ∎

תרגיל 9: *תהי $A \in \mathrm{M}_n(\mathbb{C})$ הפיכה.*

(א) *הוכיחו ש־A^*A מטריצה מרוכבת צמודה לעצמה חיובית גמורה (כלומר, כל הערכים העצמיים של A ב־$\mathbb{C}$ חיוביים).*

(ב) *הוכיחו שיש $U \in \mathrm{M}_n(\mathbb{C})$ אוניטרית ו־$P \in \mathrm{M}_n(\mathbb{C})$ צמודה לעצמה חיובית גמורה כך ש־$A = UP$.*

(רמז: שימו לב שאז $A^*A = P^*P$; הוכיחו את הטענה קודם במקרה ש־A^*A אלכסונית.)

(הגדרנו בשיעור חיובית גמורה רק עבור מטריצות מעל $\mathbb{R}$. כאן זה משהו אחר.)

פתרון:

(א) אנלוג של מה שהיה מעל $\mathbb{R}$: $(A^*A)^* = A^*A$. נקבע מרחב מכפלה פנימי V ממימד n ונקבע בסיס אורתונורמלי $\mathcal{B}$ של V (למשל $\mathbb{C}^n$ עם הבסיס הסטנדרטי). כידוע, $A = [T]_{\mathcal{B}}$ עבור איזה העתקה $T\colon V \to V$ לפי $\mathcal{B}$. אז $A^* = [T^*]_{\mathcal{B}}$ ולכן $A^*A = [T^*T]_{\mathcal{B}}$. לכן כל ערך העצמי λ של A^*A הוא ערך עצמי של T^*T. יהי $0 \neq v \in V$ וקטור עצמי השייך ל־λ. אז $T^*T(v) = \lambda v$, לכן

$$\lambda\langle v, v\rangle = \langle v, \lambda v\rangle = \langle v, T^*T(v)\rangle = \langle T(v), T(v)\rangle > 0$$

כי $T(v) \neq 0$. מכאן $\lambda > 0$, כי $\langle v, v\rangle > 0$.

(ב) לגבי הרמז: אם $A = UP$ אז $A^*A = (UP)^*(UP) = P^*U^*UP = P^*P$.

נניח קודם כי $A = (\alpha) \in \mathrm{M}_1(\mathbb{C})$ (כדי לקבל רעיון). אז $\alpha = \frac{\alpha}{|\alpha|}\alpha$ ונגדיר $U = \frac{\alpha}{|\alpha|}, P = |\alpha|$.

כעת נניח כי A^*A אלכסונית, $A^*A = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$, לפי (א), $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ חיוביים. נגדיר $P = \mathrm{Diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})$. אז $P^2 = A^*A$. נגדיר $U = AP^{-1}$. אז $A = UP$ ו־U אוניטרית:

$$U^*U = (AP^{-1})^*(AP^{-1}) = (P^{-1})^*A^*AP^{-1} = P^{-1}P^2P^{-1} = I_n$$

כי $P^{-1} = \mathrm{Diag}(\sqrt{\lambda_1}^{-1}, \ldots, \sqrt{\lambda_n}^{-1}) = (P^{-1})^*$.

במקרה הכללי יש $Q \in \mathrm{M}_n(\mathbb{C})$ אוניטרית כך ש־

$$Q^*A^*AQ = Q^*A^*QQ^*AQ = (Q^*AQ)^*(Q^*AQ)$$

אלכסונית. נסמן $A' = Q^*AQ$. אז $(A')^*A'$ אלכסונית. לכן לפי המקרה הקודם יש $U' \in \mathrm{M}_n(\mathbb{C})$ אוניטרית ו־$P' \in \mathrm{M}_n(\mathbb{C})$ צמודה לעצמה חיובית גמורה כך ש־$A' = U'P'$. נסמן $U = QU'Q^*, P = QP'Q^*$. אז U אוניטרית ו־P צמודה לעצמה חיובית גמורה ו־$A = UP$. ∎

תרגיל 10: *מהן כל האפשרויות לזוגות (f_A, m_A) עבור $A \in \mathrm{M}_3(\mathbb{R})$ אם ידוע ש־A אינה הפיכה, $\mathrm{tr}(A) \neq 0$, ו־$A^3 - A = 0$.*

פתרון: נתון ש־A מאפסת את $X^3 - X$, לכן $m_A | X^3 - X = X(X-1)(X+1)$. בפרט m_A הוא מכפלה של גורמים ממעלה ראשונה (מתוך הקבוצה $\{X, (X-1), (X+1)\}$), לכן A לכסינה (דומה לאלכסונית). היות

ולמטריצה דומה ל־A אותו פולינום אופיני ומזערי, ובפרט גם היא מאפסת את $X^3 - X$, גם אינה הפיכה ויש לה אותה העקבה, בלי הגבלת כלליות A אלכסונית. באלכסון הראשי מופיעים $0, 1, -1$ (אולי לא כולם ואולי יותר מפעם אחת). בלי הגבלת הכלליות הם מופיעים בסדר הזה (כלומר, קודם אפסים אח"כ אחדים ולבסוף -1ים).

לפי זה ש־A אינה הפיכה, 0 מופיע באלכסון הראשי לפחות פעם אחת. כיוון ש־$\operatorname{tr} A \neq 0$, $+1, -1$ אינם מופיעים אותו מספר הפעמים. לכן $A = \operatorname{Diag}(0, 0, \pm 1)$ או $A = \operatorname{Diag}(0, \pm 1, \pm 1)$ (לפי מספר הפעמים שמופיע 0). מכאן $f_A = X^2(X \mp 1), m_A = X(X \mp 1)$ או $f_A = X(X \mp 1)^2, m_A = X(X \mp 1)$.

∎

16. נספח: מטריצות של גושים

יהי F שדה. מטריצה מסדר $\sum_{k=1}^{r} m_k \times \sum_{k=1}^{s} n_k$ מעל F ניתן לחלק ל־$r \times s$ גושים: המטריצה נראית כך

$$\begin{array}{c} \\ m_1 \\ m_2 \\ \vdots \\ m_r \end{array} \begin{array}{c} \begin{array}{cccc} n_1 & n_2 & \dots & n_s \end{array} \\ \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1s} \\ A_{21} & A_{22} & \dots & A_{2s} \\ \dots & \dots & \dots & \dots \\ A_{r1} & A_{r2} & \dots & A_{rs} \end{pmatrix} \end{array}$$

באשר (המספרים משמאל ומעל למטריצה מסמסנים את גדלי הגושים) $A_{ij} \in \mathrm{M}_{m_i \times n_j}(F)$ לכל i, j. מטריצה שרשומה כך נקראת **מטריצה של גושים**.

משפט 16.1 (מכפלה של מטריצות של גושים): *מכפלה של מטריצות גושים היא מטריצת גושים:*

$$\begin{array}{c} \\ m_1 \\ m_2 \\ \vdots \\ m_r \end{array} \begin{array}{c} \begin{array}{cccc} p_1 & p_2 & \dots & p_s \end{array} \\ \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1s} \\ A_{21} & A_{22} & \dots & A_{2s} \\ \dots & \dots & \dots & \dots \\ A_{r1} & A_{r2} & \dots & A_{rs} \end{pmatrix} \end{array} \begin{array}{c} \\ p_1 \\ p_2 \\ \vdots \\ p_s \end{array} \begin{array}{c} \begin{array}{cccc} n_1 & n_2 & \dots & n_t \end{array} \\ \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1t} \\ B_{21} & B_{22} & \dots & B_{2t} \\ \dots & \dots & \dots & \dots \\ B_{s1} & B_{s2} & \dots & B_{st} \end{pmatrix} \end{array} =$$

$$= \begin{array}{c} \\ m_1 \\ m_2 \\ \vdots \\ m_r \end{array} \begin{array}{c} \begin{array}{cccc} n_1 & n_2 & \dots & n_t \end{array} \\ \begin{pmatrix} C_{11} & C_{12} & \dots & C_{1t} \\ C_{21} & C_{22} & \dots & C_{2t} \\ \dots & \dots & \dots & \dots \\ C_{r1} & C_{r2} & \dots & C_{rt} \end{pmatrix} \end{array}$$

באשר $C_{ij} = A_{i1}B_{1j} + A_{i2}B_{2j} + \ldots + A_{is}B_{sj} = \sum_{k=1}^{s} A_{ik}B_{kj}$ לכל i, j.

הוכחה: כיון שהסימונים מסובכים (בגלל הרבה אינדקסים) נטפל קודם במקרה פרטי של מטריצות בעלות 2×2 גושים. במקרה זה אפשר לנסח את המשפט כך:

$$\begin{array}{c} \\ m_1 \\ m_2 \end{array} \begin{array}{c} \begin{array}{cc} p_1 & p_2 \end{array} \\ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \end{array} \begin{array}{c} \\ p_1 \\ p_2 \end{array} \begin{array}{c} \begin{array}{cc} n_1 & n_2 \end{array} \\ \begin{pmatrix} P & Q \\ R & S \end{pmatrix} \end{array} = \begin{array}{c} \\ m_1 \\ m_2 \end{array} \begin{array}{c} \begin{array}{cc} n_1 & n_2 \end{array} \\ \begin{pmatrix} AP+BR & AQ+BS \\ CP+DR & CQ+DS \end{pmatrix} \end{array}$$

נוכיח נוסחה זו: תחילה נשים לב שהמכפלות $AP, BR, AQ, BS, CP, DR, CQ, DS$ מוגדרות וכן מוגדרים הסכומים $AP+BR, AQ+BS, CP+DR, CQ+DS$, והמטריצה בשני האגפים מאותו הסדר $(m_1+m_2) \times (n_1+n_2)$. נותר לבדוק שהרכיבים המתאימים בשני האגפים שווים.

למשל, נחשב את הרכיב ה־(m_1+i,j) בשני האגפים, באשר $1 \le i \le m_2,\ 1 \le j \le n_1$.
באגף שמאל רכיב זה מתקבל על ידי הכפלת השורה ה־m_1+i של $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ שהיא

$$\big((C)_{i1},(C)_{i2},\ldots,(C)_{i\,p_1},\ (D)_{i1},(D)_{i2},\ldots,(D)_{i\,p_2}\big)$$

בעמודה ה־j של $\begin{pmatrix} P & Q \\ R & S \end{pmatrix}$ שהיא

$$\big((P)_{1j},(P)_{2j},\ldots,(P)_{p_1\,j},\ (R)_{1j},(R)_{2j},\ldots,(R)_{p_2\,j}\big)^t.$$

מכפלתן היא

$$(C)_{i1}(P)_{1j}+(C)_{i2}(P)_{2j}+\cdots+(C)_{i\,p_1}(P)_{p_1\,j}+$$
$$+(D)_{i1}(R)_{1j}+(D)_{i2}(R)_{2j}+\cdots+(D)_{i\,p_2}(R)_{p_2\,j}=(CP)_{ij}+(DR)_{ij}=(CP+DR)_{ij}$$

וזה בדיוק הרכיב ה־(m_1+i,j) באגף ימין. בזאת הסתיימה ההוכחה של המקרה הפרטי.

המקרה הכללי דומה: תחילה נשים לב שהמכפלות $A_{ik}B_{kj}$ מוגדרות וכן מוגדרים הסכומים $\sum_{k=1}^{s} A_{ik}B_{kj}$ והמטריצות בשני האגפים מאותו הסדר $(m_1+m_2+\ldots+m_r)\times(n_1+n_2+\ldots+n_t)$. נותר לבדוק שהרכיבים המתאימים בשני האגפים שווים.

למשל, נחשב את הרכיב ה־(m_1+i,j) בשני האגפים, באשר $1 \le i \le m_2,\ 1 \le j \le n_1$.
באגף שמאל רכיב זה מתקבל על ידי הכפלת השורה ה־m_1+i של המטריצה $(\ A_{ij}\ )$ שהיא

$$\big((A_{21})_{i1},(A_{21})_{i2},\ldots,(A_{21})_{i\,p_1},\ \ \ldots\ ,(A_{2s})_{i1},(A_{2s})_{i2},\ldots,(A_{2s})_{i\,p_s}\big)$$

בעמודה ה־j של המטריצה $(\ B_{ij}\ )$ שהיא

$$\big((B_{11})_{1j},(B_{11})_{2j},\ldots,(B_{11})_{p_1\,j},\ \ \ldots\ ,(B_{s1})_{1j},(B_{s1})_{2j},\ldots,(B_{s1})_{p_s\,j}\big)^t.$$

מכפלתן היא

$$(A_{21})_{i1}(B_{11})_{1j}+(A_{21})_{i2}(B_{11})_{2j}+\cdots+(A_{21})_{i\,p_1}(B_{11})_{p_1\,j}\ +\ldots+$$
$$(A_{2s})_{i1}(B_{s1})_{1j}+(A_{2s})_{i2}(B_{s1})_{2j}+\cdots+(A_{2s})_{i\,p_s}(B_{s1})_{p_s\,j}=$$
$$\sum_{k=1}^{p_1}(A_{21})_{ik}(B_{11})_{kj}+\cdots+\sum_{k=1}^{p_s}(A_{2s})_{ik}(B_{s1})_{kj}$$

וזה בדיוק הרכיב ה־(m_1+i,j) של המטריצה $(\ C_{ij}\ )$. ∎

מסקנה 16.2: *תהיינה $A \in \mathrm{M}_{m\times p}(F)$, $B \in \mathrm{M}_{p\times n}(F)$, תהיינה $A_1,\dots,A_p$ העמודות של A, ו־$B_1,\dots,B_p$ השורות של B. אז*

$$AB = (\,A_1, \quad \dots \quad , A_p\,)\begin{pmatrix} B_1 \\ \vdots \\ B_p \end{pmatrix} = (\,\textstyle\sum_{k=1}^{p} A_k B_k\,) \in \mathrm{M}_{m\times n}(F).$$

הוכחה: הפעל את המשפט עם $r=1, s=p, t=1, p_1=\dots=p_s=1$. ∎

תרגיל 16.3: *תהיינה $v_1,\dots,v_n \in F^m$ העמודות של $A \in \mathrm{M}_{m\times n}(F)$, ויהיו $c_1,\dots,c_n \in F$. אז*

$$A\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = c_1v_1 + \dots + c_nv_n.$$

הוכחה: לפי המסקנה, אגף ימין שווה למטריצה $v_1(c_1)+\dots+v_n(c_n)$. כעת רק צריך לשים לב ש־$c_iv_i = v_i(c_i)$ לכל i ($v_i(c_i)$ הוא מטריצה מסדר $m\times 1$, מכפלת המטריצה v_i מסדר $m\times 1$ במטריצה (c_i) מסדר 1×1). ∎

תרגיל 16.4: *תהיינה $A \in \mathrm{M}_{m\times n}(F)$ ו־$P \in \mathrm{M}_m(F)$. תהיינה $A_1,\dots,A_m \in F^n$ השורות של A. אז השורות של PA הן*

$$(P)_{i1}A_1 + (P)_{i_2}A_2 + \dots + (P)_{im}A_m = \sum_{k=1}^{m}(P)_{ik}A_k, \qquad i=1,2,\dots,m$$

הוכחה: לפי המשפט, $PA = \begin{pmatrix} (P)_{11} & \dots & (P)_{1m} \\ \dots & \dots & \dots \\ (P)_{m1} & \dots & (P)_{mm} \end{pmatrix}\begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{m}(P)_{1k}A_k \\ \vdots \\ \sum_{k=1}^{m}(P)_{mk}A_k \end{pmatrix}$ ∎

תרגיל 16.5: *הוכיחו כי מטריצת הגושים* $M = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \end{array}$ *הפיכה אם ורק אם A, D הפיכות ואם זה קורה אז*

$N = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{pmatrix} A^{-1} & -A^{-1}BD^{-1} \\ 0 & D^{-1} \end{pmatrix} \end{array}$ *ההופכי שלה.*

בפרט, $M_0 = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} \end{array}$ *הפיכה אם ורק אם A, D הפיכות ואז* $M_0^{-1} = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{pmatrix} A^{-1} & 0 \\ 0 & D^{-1} \end{pmatrix} \end{array}$.

הוכחה: אם M הפיכה אז $M^{-1} = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{pmatrix} P & Q \\ R & S \end{pmatrix} \end{array}$ ומתקיים $\begin{pmatrix} A & B \\ 0 & D \end{pmatrix}\begin{pmatrix} P & Q \\ R & S \end{pmatrix} = I_{m+n}$, כלומר, לפי המשפט,

$$\begin{pmatrix} AP+BR & AQ+BS \\ DR & DS \end{pmatrix} = \begin{pmatrix} I_m & 0 \\ 0 & I_n \end{pmatrix}.$$

מכאן

$$.AP+BR=I_m \quad ,DS=I_n \quad ,DR=0 \quad ,AQ+BS=0$$

מתוך $DS=I_n$ נסיק ש־D הפיכה ו־$S=D^{-1}$. מכאן, מתוך $DR=0$, נסיק ש־$R=D^{-1}(DR)=0$. לכן $AP+BR=I_m$ הופך להיות ל־$AP=I_m$ ומכאן נסיק ש־A הפיכה ו־$P=A^{-1}$. לבסוף, מתוך $AQ+BS=0$ נסיק כי $Q=-A^{-1}BS=-A^{-1}BD^{-1}$. בסיכום: A,D הפיכות ו־$M^{-1}=N$.

להיפך, אם A,D מטריצות הפיכות, אז לפי המשפט, $\begin{pmatrix} A & B \\ 0 & D \end{pmatrix}\begin{pmatrix} A^{-1} & -A^{-1}BD^{-1} \\ 0 & D^{-1} \end{pmatrix} = \begin{pmatrix} I_m & 0 \\ 0 & I_n \end{pmatrix} = I_{m+n}$, לכן M הפיכה ו־$M^{-1}=N$. ∎

תרגיל 16.6: *הוכיחו כי מטריצת הגושים* $M = \begin{array}{c c} & \begin{array}{cc} n & m \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{pmatrix} A & B \\ C & 0 \end{pmatrix} \end{array}$ *הפיכה אם ורק אם* B,C *הפיכות ואם זה קורה אז*

$N = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} n \\ m \end{array} & \begin{pmatrix} 0 & C^{-1} \\ B^{-1} & -B^{-1}AC^{-1} \end{pmatrix} \end{array}$ *ההופכי שלה.*

בפרט, $M_0 = \begin{array}{c c} & \begin{array}{cc} n & m \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{pmatrix} 0 & B \\ C & 0 \end{pmatrix} \end{array}$ *הפיכה אם ורק אם* B,C *הפיכות ואז* $M_0^{-1} = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} n \\ m \end{array} & \begin{pmatrix} 0 & C^{-1} \\ B^{-1} & 0 \end{pmatrix} \end{array}$.

הוכחה: כמו בתרגיל הקודם. יש לשים לב שסדרי הגושים של N שונים מהסדרים של הגושים ב־M, אך המכפלה MN מוגדרת ומתאימה לתנאי המשפט. ∎

תרגיל 16.7: *הוכיחו כי* $\begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} \end{array}$ *דומה למטריצה* $\begin{array}{c c} & \begin{array}{cc} n & m \end{array} \\ \begin{array}{c} n \\ m \end{array} & \begin{pmatrix} D & 0 \\ 0 & A \end{pmatrix} \end{array}$.

הוכחה: לפי תרגיל 16.6, $\begin{array}{c c} & \begin{array}{cc} n & m \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{pmatrix} 0 & I_m \\ I_n & 0 \end{pmatrix} \end{array}$ הפיכה ו־ $\begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} n \\ m \end{array} & \begin{pmatrix} 0 & I_n \\ I_m & 0 \end{pmatrix} \end{array}$ ההופכי שלה. לפי המשפט

$$\begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} n \\ m \end{array} & \begin{pmatrix} 0 & I_n \\ I_m & 0 \end{pmatrix} \end{array} \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} \end{array} \begin{array}{c c} & \begin{array}{cc} n & m \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{pmatrix} 0 & I_m \\ I_n & 0 \end{pmatrix} \end{array} = \begin{array}{c c} & \begin{array}{cc} n & m \end{array} \\ \begin{array}{c} n \\ m \end{array} & \begin{pmatrix} D & 0 \\ 0 & A \end{pmatrix} \end{array}$$

∎

תרגיל 16.8: *תהי* $M = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \end{array}$ *מטריצת גושים. הוכיחו:* $|M| = |A| \cdot |D|$.

הוכחה: נוכיח כי $\frac{|A|\cdot|D|}{|M|} = 1$.

כידוע, לכל פעולה אלמנטרית $\mathcal{P}$ קיים סקלר $\lambda_{\mathcal{P}}$ כך ש־$|\mathcal{P}(C)| = \lambda_{\mathcal{P}} \cdot |A|$ לכל מטריצה ריבועית C. (אם $\mathcal{P}$ היא החלפת שורות, אז $\lambda_{\mathcal{P}} = -1$; אם $\mathcal{P}$ היא הכפלת שורה בסקלר, אז $\lambda_{\mathcal{P}}$ הוא הסקלר הזה; ואם $\mathcal{P}$ היא הוספת של כפולה של שורה לשורה אחרת, אז $\lambda_{\mathcal{P}} = 1$.) נסמן $A' = \mathcal{P}(A)$, $B' = \mathcal{P}(B)$, ויהי $M' = \begin{pmatrix} A' & B' \\ 0 & D \end{pmatrix}$. אז $M' = \mathcal{P}(M)$. לכן $\frac{|A'|\cdot|D|}{|M'|} = \frac{\lambda_{\mathcal{P}}|A|\cdot|D|}{\lambda_{\mathcal{P}}|M|} = \frac{|A|\cdot|D|}{|M|}$, כלומר, הפעלת $\mathcal{P}$ על M אינה משנה את $\frac{|A|\cdot|D|}{|M|}$. לכן אפשר להניח ש־A מדורגת ובפרט משולשית עליונה.

באופן דומה ניתן להניח ש־D משולשית עליונה. אז גם M משולשית עליונה. כיון שבמטריצה משולשית עליונה הדטרמיננטה היא מכפלת האיברים באלכסון הראשי, המסקנה נובעת. ∎

17. נספח: מטריצות מעבר

נספח זה מכיל חומר מאלגברה לינארית 1 אודות מטריצות מעבר.

יהי V מרחב וקטורי מעל שדה F ויהי $\mathcal{B} = (v_1, \ldots, v_n)$ בסיס שלו.

הגדרה 17.1: תהי $\mathcal{B}' = (v'_1, \ldots, v'_n)$ סדרה של אברי V. אז $P = \big([v'_1]_{\mathcal{B}}, \ldots, [v'_n]_{\mathcal{B}}\big) \in \mathrm{M}_n(F)$ נקראת **מטריצת המעבר** מ־$\mathcal{B}$ ל־$\mathcal{B}'$. שים לב שאם $\mathcal{B}'$ בסיס של V אז $P = [\mathrm{id}_V]_{\mathcal{B}}^{\mathcal{B}'}$.

דוגמה 17.2: מטריצת המעבר מ־$\left(\binom{1}{1}, \binom{0}{1}\right)$ ל־$\left(\binom{1}{1}, \binom{1}{-1}\right)$ היא $\begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix}$.

תרגיל 17.3: *תהי $P \in \mathrm{M}_n(F)$. אז קיימת סדרה $\mathcal{B}' = (v'_1, \ldots, v'_n)$ של אברי V כך ש־P היא מטריצת המעבר מ־$\mathcal{B}$ ל־$\mathcal{B}'$.*

פתרון: ההעתקה $v \mapsto [v]_{\mathcal{B}}$ היא על F^n. לכן לכל $1 \le j \le n$ קיים $v'_j \in V$ כך ש־$[v'_j]_{\mathcal{B}}$ היא העמודה ה־j של P. נגדיר $\mathcal{B}' = (v'_1, \ldots, v'_n)$, אז P מטריצת המעבר מ־$\mathcal{B}$ ל־$\mathcal{B}'$. ∎

משפט 17.4: *מטריצת המעבר P מ־$\mathcal{B}$ ל־$\mathcal{B}'$ הפיכה אם ורק אם $\mathcal{B}'$ בסיס.*

הוכחה: P הפיכה $\Leftrightarrow \dim(Sp([v'_1]_{\mathcal{B}}, \ldots, [v'_n]_{\mathcal{B}})) = \dim(C(P)) = \mathrm{rk}\, P = n$

$\Leftrightarrow$ $[v'_1]_{\mathcal{B}}, \ldots, [v'_n]_{\mathcal{B}}$ בסיס של F^n. (כאן $C(P)$ הוא מרחב העמודות של P.)

אך האיזומורפיזם $V \to F^n$ הנתון על ידי $v \mapsto [v]_{\mathcal{B}}$ וגם ההופכי שלו $F^n \to V$ מעתיקים בסיסים לבסיסים. לכן $[v'_1]_{\mathcal{B}}, \ldots, [v'_n]_{\mathcal{B}}$ בסיס של F^n $\Leftrightarrow$ $v'_1, \ldots, v'_n$ בסיס של V. ∎

טענה 17.5: *יהי $\mathcal{B}' = (v'_1, \ldots, v'_n)$ בסיס של V ויהי $v \in V$. אז $[v]_{\mathcal{B}} = P[v]_{\mathcal{B}'}$.*

הוכחה: : נניח תחילה ש־$v = v'_j \in \mathcal{B}'$. אז $[v'_j]_{\mathcal{B}'} = \mathbf{e}_j$, לכן

$$[v'_j]_{\mathcal{B}} = \text{העמודה ה־}j\text{ של } P = P\mathbf{e}_j = P[v'_j]_{\mathcal{B}'}.$$

במקרה הכללי יש $a_1, \ldots, a_n \in F$ כך ש־$v = \sum_{j=1}^n a_j v'_j$. אז, לפי המקרה הקודם,

$$[\sum_{j=1}^n a_j v'_j]_{\mathcal{B}} = \sum_{j=1}^n a_j [v'_j]_{\mathcal{B}} = \sum_{j=1}^n a_j P[v'_j]_{\mathcal{B}'} = P\big(\sum_{j=1}^n a_j [v'_j]_{\mathcal{B}'}\big) = P[\sum_{j=1}^n a_j v'_j]_{\mathcal{B}'},$$

כלומר, $P[v]_{\mathcal{B}'} = [v]_{\mathcal{B}}$. ∎

18. נספח: מטריצות אלמנטריות

נספח זה מכיל חומר מאלגברה לינארית 1 אודות מטריצות אלמנטריות.

נקבע שדה F ומספר טבעי m.

הגדרה 18.1: פעולה אלמנטרית $\mathcal{E}$ (על שורות של מטריצות בנות m שורות) היא אחת מהבאות:

(א) הכפלה של שורה בסקלר שונה מאפס.

(ב) הוספה של כפולה של שורה לשורה אחרת.

(ג) החלפה בין שתי שורות.

המטריצה $\mathcal{E}(I_m) \in \mathrm{M}_m(F)$ תקרא **המטריצה האלמנטרית המתאימה ל־$\mathcal{E}$**.

הערה 18.2: מטריצת אלמנטרית הינה הפיכה, כי פעולה אלמנטרית אינה משנה את הדרגה.

טענה 18.3: *תהי $A \in \mathrm{M}_{m\times n}(F)$. תהי $\mathcal{E}$ פעולה אלמנטרית. ותהי E המטריצה האלמנטרית המתאימה לה. אז $\mathcal{E}(A) = EA$.*

הוכחה: (א) בלי הגבלת הכלליות $n = 1$. אכן, תהיינה $A_1, \ldots, A_n$ העמודות של A. אז $\mathcal{E}(A) = \big(\mathcal{E}(A_1), \ldots, \mathcal{E}(A_n)\big)$ ואילו $E(A) = \big(EA_1, \ldots, EA_n\big)$, לכן די להוכיח כי $\mathcal{E}(A_j) = EA_j$ לכל j.

(ב) בלי הגבלת הכלליות $A = \mathbf{e}_i$ לכל $1 \le i \le m$. אכן, $A = \sum_{i=1}^m a_i\mathbf{e}_i$, באשר $a_1, \ldots, a_m \in F$. אז $\mathcal{E}(A) = \sum_{i=1}^m a_i\mathcal{E}(\mathbf{e}_i)$ ואילו $E(A) = \sum_{i=1}^m a_i\mathbf{e}_iA\mathbf{e}_i$, לכן די להוכיח כי $\mathcal{E}(\mathbf{e}_i) = E\mathbf{e}_i$ לכל i.

(ג) $E\mathbf{e}_i$ היא העמודה ה־i של E. אבל $E = \mathcal{E}(I_m) = \mathcal{E}(\mathbf{e}_1, \ldots, \mathbf{e}_m) = \big(\mathcal{E}(\mathbf{e}_1), \ldots, \mathcal{E}(\mathbf{e}_m)\big)$, לכן העמודה ה־$i$ של E היא $\mathcal{E}(\mathbf{e}_i)$. ∎

הגדרה 18.4: תהי $\mathcal{E}$ פעולה אלמנטרית. אז **הפעולה האלמנטרית $\mathcal{E}'$ המתאימה ל־$\mathcal{E}$** (על עמודות של מטריצות בנות m עמודות) היא הפעולה שמתקבלת מ־$\mathcal{E}$ על ידי החלפת המלים "שורה\שורות" במלים "עמודה\עמודות".

מסקנה 18.5: *תהי $\mathcal{E}$ פעולה אלמנטרית ותהי $\mathcal{E}'$ הפעולה האלמנטרית על עמודות המתאימה לה. תהי $B \in \mathrm{M}_{m\times n}(F)$. אז $\mathcal{E}'(B) = BE^t$.*

הוכחה: $\mathcal{E}'(B) = \big(\mathcal{E}(B^t)\big)^t = (EB^t)^t = BE^t$. ∎