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

אלגברה לינארית 2 א

מערכי שעור

תשפ"א

נערך על ידי

דן הרן

עדכון אחרון: 26.10.2020

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 $0 \neq 1$ 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) $a0+a0=a(0+0)=a0$. נחבר את הנגדי $-(a0)$ של $a0$ לשני האגפים ונקבל $a0=0$. באופן דומה $0a=0$.

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:

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

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

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 לכתוב 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 הביטויים שווים. באפן דומה aX אפשר להבין כפולינום $0X^0 + aX^1 + 0X^2 + \cdots$ או כמכפלה של הפולינומים a ו־X; שוב, שני הביטויים 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

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

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 לציין 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 $f = \sum_{i=0}^\infty a_iX^i, g = \sum_{i=0}^\infty b_iX^i \in R[X]$ 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_mb_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־$b_nc = 1$ אז מתקיים שוויון בתנאי (ב).*

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

(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 $a_ib_j = 0$. 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 $a_ib_j = 0$. לכן המקדם של X^k ב־fg הוא a_mb_n.

(ו) $(a_mb_n)c = a_m(b_nc) = a_m1 = a_m \neq 0$, לכן $a_mb_n \neq 0$. 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, $\varphi(0) + 0 = \varphi(0) = \varphi(0+0) = \varphi(0) + \varphi(0)$ ואם נחבר לשני האגפים את הנגדי של $\varphi(0)$, נקבל $\varphi(0) = 0$.

(ב) $\varphi(-a) = -\varphi(a)$, לכל $a \in R$. אכן, $0 = \varphi(0) = \varphi(-a+a) = \varphi(-a) + \varphi(a)$, 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 הצבה** $\varphi\colon R[X] \to R$ 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:

$$X\cdot a=(0X^0+1X^1+0X^2+\cdots)(aX^0+0X^1+0X^2+\cdots)=$$
$$=(0\cdot a)X^0+(0\cdot 0+1\cdot a)X^1+(0\cdot 0+1\cdot 0+0\cdot a)X^2+\cdots=$$
$$=0X^0+aX^1+0X^2+\cdots=aX$$

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!):

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

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.

דוגמאות 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 $S \to R$. 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

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

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:

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

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 $F^\times = F \smallsetminus \{0\}$ 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־*

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

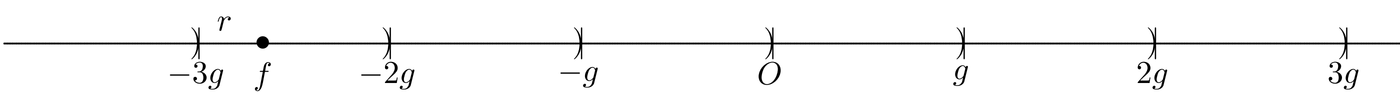
(ב) *קיימים $q', r' \in R[X]$ יחידים כך ש־*

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

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

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

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

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

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

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

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

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:

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

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

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. (בדוק! 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 $0 = ac - bc = (a - b)c$. 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), $uv = 1$. כיוון ש־R קומוטטיבי, גם $vu = 1$. 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 $m := \min(\deg h |\ h \in I \smallsetminus \{0\})$.*

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.*

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 (בדוק!) ולכן יש $c_0\in R$ כך ש־$I=(c_0)$.

$a\in I$, 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 $a=rc_0=(ru^{-1})q$, 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־*

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

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 $f = cq_{i_0}$ 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.

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

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

באשר $d \in R^\times = F^\times$, ו־$n_i \in \mathbb{N}\cup\{0\}$, 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 $i \in I$, $j \neq i$, כך ש־$q_j | q_i$. כלומר $q_i = q_j u$, 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־$F[X]$. הוכח 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־ $b^2 - 4c < 0$. הוכח 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 $(X - \alpha_1)|f$ ב־$\mathbb{R}[X]$ לפי מסקנה 2.6, ומכאן $f = X - \alpha_1$, בגלל ש־f אי פריק ומתוקן (בדוק!).

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

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

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 $q|f$ ב־$\mathbb{C}[X]$, לכן לפי תרגיל 3.9, גם ב־$\mathbb{R}[X]$ ומכאן $f = q$, 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:

אם 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. ∎

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־$F[X]$:

$$
\begin{array}{llll}
(1) & f_1 = f_2 q_1 + f_3, & \deg f_3 < \deg f_2, & f_3 \neq 0 \\
(2) & f_2 = f_3 q_2 + f_4, & \deg f_4 < \deg f_3, & f_4 \neq 0 \\
\vdots & \vdots & \vdots & \vdots \\
(r-2) & f_{r-2} = f_{r-1} q_{r-2} + f_r, & \deg f_r < \deg f_{r-1}, & f_r \neq 0 \\
(r-1) & f_{r-1} = f_r q_{r-1} + f_{r+1}, & & f_{r+1} = 0
\end{array}
$$

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

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

$$
\begin{array}{ll}
\text{מתוך } (1) \text{ נובע:} & d|f_2, f_3 \Leftrightarrow d|f_1, f_2 \\
\text{מתוך } (2) \text{ נובע:} & d|f_3, f_4 \Leftrightarrow d|f_2, f_3 \\
\vdots & \vdots \\
\text{מתוך } (r-1) \text{ נובע:} & d|f_r, 0 \Leftrightarrow d|f_{r-1}, f_r
\end{array}
$$

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

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

$$.d = g f_1 + h f_2$$

הוכחה: נראה (בסימונים דלעיל) שלכל $1 \le i \le r$ 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 $i = r$ נקח $g_r = u^{-1}, h_r = 0$. נניח נכונות עבור $1 \le i \le r$ מסויים. אז לפי המשוואה $(i-1)$ לעיל מתקיים $f_{i+1} = f_{i-1} - f_i q_{i-1}$. אם נציב זאת במשוואה $(*_i)$ ונקבל

$$.d = g_i f_i + h_i (f_{i-1} - f_i q_{i-1}) = h_i f_{i-1} + (g_i - h_i q_{i-1}) f_i$$

לכן אם נסמן $h_i = g_{i-1},\ h_i = g_i - h_i q_{i-1}$, אז

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

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

$$f_3 = (gf_1 + hf_2)f_3 = f_1gf_3 + hf_2f_3 = gf_1f_3 + hqf_1 = (gf_3 + hq)f_1$$

מתחלק ב־f_1. ∎

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

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

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

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

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

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

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

אז לפי תרגיל 3.20

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

אכן, אם $d \in F[X]$ מתוקן, אז

$$d = \prod_{i\in I} q_i^{k_i},$$

באשר $k_i \ge 0$ לכל $i \in I$ וכמעט כל k_i הוא אפס. לפי תרגיל 3.20,

$$d|f_1, f_2 \Leftrightarrow k_i \le m_i, n_i \Leftrightarrow k \le \min(m_i, n_i).$$

מכאן ברור ש־d גדול ביותר כזה אם ורק אם $k = \min(m_i, n_i)$. ∎

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

אם m' מתוקן ו־ $f_1, \dots, f_k | m'$ 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 ווקטורים 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.) הסקלאר 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 $\mathcal{B} = (v_1, \ldots, v_n)$ 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 $(A-\lambda I)X=0$. לכן λ ערך עצמי של A אם ורק אם $\det(A-\lambda I)=0$.

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 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 $\alpha_1,\dots,\alpha_n,\beta_1,\dots\beta_n\in F$. אז

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

ואם $\alpha_1,\dots,\alpha_n\in F^\times$ 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 $\mathcal{B}=(v_1,\dots,v_n)$ 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 $j=1,\dots,n$. אם זה קורה אז $\lambda_1,\dots,\lambda_n$ ערכים עצמיים של T.*

הוכחה:

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

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

לכן $\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = [T]_{\mathcal{B}}^{\mathcal{B}}$ אם ורק אם $[T(v_j)]_{\mathcal{B}} = \lambda_j [v_j]_{\mathcal{B}}$ לכל $1 \le j \le n$ אם ורק אם (בגלל שההעתקה $v \mapsto [v]_{\mathcal{B}}$ היא איזומורפיזם) $T(v_j) = \lambda_j v_j$ לכל $1 \le j \le n$. מכאן השקילות. הטענה האחרונה של המשפט נובעת מכך ש־$v_1, \dots, v_n$ 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 $v_1, \dots, v_n$ בסיס של F^n המורכב מוקטורים עצמיים של A, כאשר v_j שייך ל־λ_j, לכל $1 \le j \le n$. להיפך, אם יש בסיס כזה ונגדיר $P = (v_1, \dots, v_n) \in \mathrm{M}_n(F)$, אז $P^{-1}AP = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$.*

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

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

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

לכן

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \Leftrightarrow AP = P(\lambda_1 \mathbf{e}_1, \dots, \lambda_n \mathbf{e}_n)$$
$$\Leftrightarrow (Av_1, \dots, Av_n) = (\lambda_1 v_1, \dots, \lambda_n v_n)$$
$$\Leftrightarrow Av_j = \lambda_j v_j \text{ לכל } 1 \le j \le n$$

מכאן המסקנה. ∎

משפט 4.10: *יהיו $v_1, \dots, v_m$ וקטורים עצמיים של T* ***השונים מאפס*** *והשייכים לערכים עצמיים $\lambda_1, \dots, \lambda_m$, בהתאמה. נניח ש־$\lambda_1, \dots, \lambda_m$ שונים זה מזה. אז $v_1, \dots, v_m$ 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 $\alpha_1, \dots, \alpha_m \in F$ 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. לפי המשפט הקודם הם בלתי תלויים לינארית. אבל לפי $(3')$

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

סתירה. ∎

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

לקוח מתוך

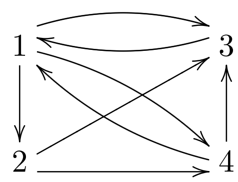
Kurt Bryan, Tanya Leise, *The $25,000,000,000 Eigenvector*, SIAM Review **48** (2006), 569–581.

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 $1,\dots,n$ 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 ההַפניות 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 $L_i\subseteq\{1,\dots,n\}$ 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 ההַפניות 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 $\lambda=1$.

חישוב 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? (כנראה בגלל שהוא מַפנה לאתר 1 וזה עושה את 1 לחשוב מאד.)

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 $A = \mathrm{Diag}(A_1, A_2)$, באשר A_1, A_2 סטוכסטיות לפי עמודות. אז גם A סטוכסטית לפי עמודות. לכן יש $v_i \neq 0$ כך ש־$A_i v_i = v_i$, עבור $i = 1, 2$. אבל אז $\binom{v_1}{0}, \binom{0}{v_2} \in \mathbb{R}^n$ בלתי תלויים לינארית ושניהם וקטורים עצמיים של A השייכים ל־1.

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.*

(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

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

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

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

כעת נראה איך אפשר לחשב (באופן מקורב) 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. אז לכל* $v \in W$ *מתקיים*

(א) $Av \in W$;

(ב) $||Av|| \leq c\,||v||$, באשר $c := \max_j(1 - 2\min_i a_{ij})$.

(ג) $0 \leq c < 1$.

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

(ג) 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) נסמן $w = Av$. אם $w = 0$, הטענה ברורה. לכן נניח כי $w \neq 0$. לפי (א), $\sum_i (w)_i = 0$. 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,

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

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

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

לכן $|\sum_i \varepsilon_i a_{ij}| \leq c$. מכאן $||w|| \leq c\,|\sum_j (v)_j| \leq c \sum_j |(v)_j| = c\,||v||$. ∎

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, $||A^k w|| \leq c^k ||w||$, לכן $A^k w \to 0$. מכאן $A^k v_0 \to v$. ∎

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

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

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

(א') $\det A_{nj} \in F[X]$,

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

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

מכאן, לפי (1):

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

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

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

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

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 $f_A(X) = (X-1)(X-4) - 2 \cdot 3 = X^2 - 5X - 2$. ∎

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), $c_0 = f_A(0) = \det(0I_n - A) = \det(-A) = (-1)^n \det A$.

(1) + (2): באינדוקציה על n.

עבור $n = 1$, $f_A(X) = X - a_{11}$.

עבור $n > 1$ נעשה פיתוח של $B := XI_n - A$ 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_{ij} מתקבלת מ־A על ידי מחיקת השורה ה־i והעמודה ה־j, 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 בדיוק $n-2$ רכיבים ממעלה 1, וכל היתר ממעלה≤ 0. (אכן, ב־$XI_n - A$ יש בדיוק n רכיבים ממעלה 1 ־ באלכסון הראשי; כאשר מחקנו שורה n, מחקנו אחד מהם, וכאשר מחקנו עמודה j, מחקנו אחר.) לכן $\deg \det B_{nj} \leq n-2$ לפי למה 6.1(ב).

מכאן

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

הוכחה נוספת: תהי $B = XI_n - A$. אז

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

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

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

(2) + (1): אם $\sigma = 1$, אז $\operatorname{Sgn}(\sigma)\prod_{i=1}^n (B)_{i\sigma(i)} = 1 \cdot (X - a_{11}) \cdots (X - a_{nn})$. 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 $0 \geq$, וכל השאר ממעלה $1 \geq$. לכן $\deg \prod_{i=1}^n (B)_{i\sigma(i)} \leq n-2$. מכאן

$$f_A(X) = (X - a_{11}) \cdots (X - a_{nn}) + \cdots + (\text{פולינום ממעלה } \leq n-2).$$

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

(3): נציב $X = 0$ ב־(2):

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

$$(-1)^n \sum_{\sigma \in S_n} \operatorname{Sgn}(\sigma) \prod_{i=1}^n a_{i\sigma(i)} = (-1)^n \det A \qquad \blacksquare$$

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 $\lambda_1, \ldots, \lambda_n \in F$ שונים זה מזה, אז A דומה למטריצה אלכסונית.*

הוכחה : לכל $1 \leq i \leq n$, 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, $v_1, \ldots, v_n$ בלתי תלויים לינארית, ולכן בסיס של F^n. לפי מסקנה 4.9, A דומה לאלכסונית. $\blacksquare$

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,

$$P^{-1}(XI - A)P = XP^{-1}IP - P^{-1}AP = XI - P^{-1}AP$$

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

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

$$= \det P^{-1} \det P \det(XI - A) = \det I \; f_A(X) = f_A(X) \qquad \blacksquare$$

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

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

אפשר להבין הגדרה זו כהצבה: F תת חוג של $\mathrm{M}_n(F)$ (כאשר מזהים סקלר α עם המטריצה הסקלרית αI_n) ולכן $g = \sum_{i=0}^{\infty} a_i I_n X^i \in \mathrm{M}_n(F)[X]$. כעת אפשר להציב בפולינום זה את $A \in \mathrm{M}_n(F)$. התוצאה של ההצבה היא $\sum_{i=0}^{\infty} (a_i I_n) A^i = \sum_{i=0}^{\infty} a_i A^i$. 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: $(cg)(A) = c(A)g(A) = cg(A)$.

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 $\left(\varphi(\sum_{\mu=0}^{\infty} A_\mu X^\mu)\right)_{ij} = \sum_{\mu=0}^{\infty} (A_\mu)_{ij} X^\mu$ היא איזומורפיזם של חוגים. קל לראות כי

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

כאשר $f_A(X) \in F[X] \subseteq \mathrm{M}_n(F)[X]$. כיון 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, כיון 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 $f_A(A) = 0$. ∎

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

תרגיל 6.11: *תהי* $A \in \mathrm{M}_n(F)$. *הראה, מבלי להסתמך על משפט* Cayley-Hamilton*, שיש* $0 \neq g \in F[X]$ *ממעלה* $\geq n^2$ *כך ש־* $g(A) = 0$.

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

תרגיל 6.12: *תהי*

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

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

(א) *הוכח ש־* $\det(M) = \det(A_1) \cdots \det(A_n)$.

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

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

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

הגדרה 6.13: **הפולינום האופייני של** T הוא הפולינום האופייני של המטריצה $f_T \in F[X]$ $[T]_{\mathcal{B}}^{\mathcal{B}}$, 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. 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 $[g(T)]_{\mathcal{B}}^{\mathcal{B}} = \sum_{i=0}^{\infty} a_i [T^i]_{\mathcal{B}}^{\mathcal{B}} = \sum_{i=0}^{\infty} a_i ([T]_{\mathcal{B}}^{\mathcal{B}})^i = g([T]_{\mathcal{B}}^{\mathcal{B}})$. ∎

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 $f_T(T) = 0$. ∎

7. פולינום מזערי

יהי F שדה.

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

הגדרה 7.1: תהי $A \in \mathrm{M}_n(F)$. פולינום $m_A(X) \in F[X]$ 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: $f(A) = 0 \Leftrightarrow m_A | f$. בפרט,* $m_A | f_A$.

הערה 7.3: $\deg m_A \geq 1$. ∎

הערה 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. אז $\varphi(X - A) = XI_n - A$, $\varphi\big(m_A(X)\big) = m_A(X)I_n$, ותהי $C = \varphi\big(q(X)\big)$. אזי לפי (7)

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

ניקח דטרמיננטה משני האגפים ונקבל

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

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,

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

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

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 $A_1, \dots, A_n$ מטריצות ריבועיות ויהי $g \in F[X]$.*

(א) אם
$$M = \begin{pmatrix} A_1 & B & \cdots & C \\ 0 & A_2 & \cdots & D \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & A_n \end{pmatrix} \text{ אז } g(M) = \begin{pmatrix} g(A_1) & B' & \cdots & C' \\ 0 & g(A_2) & \cdots & D' \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & g(A_n) \end{pmatrix}$$
עבור מטריצות $B', \dots, C', \dots, D', \dots$ מתאימות.

(ב) אם
$$M = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & A_n \end{pmatrix} \text{ אז } g(M) = \begin{pmatrix} g(A_1) & 0 & \cdots & 0 \\ 0 & g(A_2) & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & g(A_n) \end{pmatrix}$$

משפט 7.10: *תהי M כמו בתרגיל 7.9 ויהי $h = \mathrm{lcm}(m_{A_1}, \dots, m_{A_n})$. 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 $m_{A_1}, \dots, m_{A_n}|m_M$ 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־H, m_M 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 $g = \sum_{i=0}^{\infty} a_i X^i \in F[X]$ מתקיים

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

לכן $g(A) = 0 \Leftrightarrow P^{-1}g(A)P = 0 \Leftrightarrow g(P^{-1}AP) = 0$. 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. לפי 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 $\deg m_A \le \deg f_A = n$. יהי $m_A = \sum_{i=0}^{d} a_i X^i$, באשר $a_d = 1$. אז $\sum_{i=0}^{d} a_i A^i = m_A(A) = 0$ ולכן $\sum_{i=0}^{d} a_i A^i \mathbf{e}_1 = 0$. מכאן $A^d\mathbf{e}_1 = \sum_{i=0}^{d-1} -a_i A^i \mathbf{e}_1$. לפי טענה 2 זה לא יתכן אם $d \le n-1$. לכן $\deg m_A = d \ge n$.

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, $g(T) = 0 \Leftrightarrow [g(T)]_{\mathcal{B}}^{\mathcal{B}} = 0 \Leftrightarrow g(C) = 0$. 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:* $f(T) = 0 \Leftrightarrow m_T | f$.

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, $f(C) = [f(T)]_{\mathcal{B}}^{\mathcal{B}}$. לכן $f(T) = 0 \Leftrightarrow f(C) = 0$. לפי משפט 7.2, $f(C) = 0 \Leftrightarrow m_C | f$. אך לפי משפט 7.14 $m_T = m_C$. לכן $f(T) = 0 \Leftrightarrow m_T | f$. ∎

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 $V_\lambda = \{v \in V |\ T(v) = \lambda v\}$. ∎

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 $v_1, \dots, v_r$ בסיס של V_λ. נשלים אותו לבסיס $v_1, \dots, v_r, w_1, \dots, w_s$ 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 $\lambda_1, \dots, \lambda_r \in F$ 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 $V_{\lambda_1}, \dots, V_{\lambda_r}$ 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 שבבסיס 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 $\lambda_1, \ldots, \lambda_n \in F$. אז יש מטריצה משולשית עליונה $C \in \mathrm{M}_n(F)$ עם אלכסון ראשי $(\lambda_1, \ldots, \lambda_n)$ 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 $(\lambda_1, \ldots, \lambda_n)$. 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. 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 $v_1, v_2, \ldots, v_n$ של F^n ותהי $Q = (v_1, v_2, \ldots, v_n) \in \mathrm{M}_n(F)$. אז Q הפיכה, כיון שעמודותיה בלתי תלויות לינארית, ו־$Q^{-1}v_1 = \mathbf{e}_1$ כי $Q\mathbf{e}_1 = v_1$, ולכן

$$Q^{-1}AQ = Q^{-1}(Av_1, \ldots, Av_n) = Q^{-1}(\lambda_1 v_1, Av_2, \ldots, Av_n) = (\lambda_1 \mathbf{e}_1, * \cdots *) =,$$

$$= \begin{pmatrix} \lambda_1 & * \cdots * \\ 0 & \\ \vdots & A_1 \\ 0 & \end{pmatrix} \tag{8}$$

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־

$$Q_1^{-1} A_1 Q_1 = \begin{pmatrix} \lambda_2 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}. \tag{10}$$

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

$$P^{-1}AP = (Q_2^{-1}Q^{-1})A(QQ_2) = Q_2^{-1}(Q^{-1}AQ)Q_2 =.$$

$$\begin{pmatrix} 1 & 0 \\ 0 & Q_1^{-1} \end{pmatrix} \begin{pmatrix} \lambda_1 & * \\ 0 & A_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & Q_1^{-1} \end{pmatrix} \begin{pmatrix} \lambda_1 & * \\ 0 & A_1 Q_1 \end{pmatrix} =$$

$$\begin{pmatrix} \lambda_1 & * \\ 0 & Q_1^{-1} A_1 Q_1 \end{pmatrix} = \begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \quad \blacksquare$$

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

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. ∎

מסקנה 8.37: *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 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 C עם אלכסון ראשי $(\lambda_1, \ldots, \lambda_n)$. אז יש בסיס $\mathcal{B}$ כך ש־$C = [T]_{\mathcal{B}}^{\mathcal{B}}$. ∎

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 **שָׁמוּר־**T אם $T(W) \subseteq W$, כלומר, אם $T(w) \in W$ לכל $w \in W$. 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 $T|_W$. ∎

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.*

(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}.*

מסקנה 9.4: *יהי $f \in F[X]$. אז $\mathrm{Ker}\, f(T), \mathrm{Im}\, f(T)$ 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 $[T]_{\mathcal{B}} = \begin{pmatrix} A & C \\ 0 & D \end{pmatrix}$ באשר $A = [T|_W]_{\mathcal{B}_0}$.*

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

$$v = w_1 + \cdots + w_s, \qquad w_1 \in W_1, \dots, w_s \in W_s. \quad ∎$$

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: תרגיל. ראה גם "אלגברה לינארית" של ליפשיץ, תרגיל 10.7. ∎

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

(א) *הוכח כי W שמור־$g(T)$.*

(ב) *הוכח 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*

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

משפט 9.9: *נניח כי $V = W_1 \oplus \cdots \oplus W_s$. לכל $1 \leq i \leq s$ 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, $[T]_{\mathcal{B}}^{\mathcal{B}} = (A_{ij})$, באשר $A_{ij} \in \mathrm{M}_{n_i \times n_j}(F)$ לכל $1 \leq i, j \leq s$. נקבע $1 \leq j \leq s$. אז W_j שמור־T אם ורק אם $A_{ij} = 0$ לכל $j \neq i$. 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 $1 \leq j \leq s$.*

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

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 $i \neq j$ אם ורק אם $(A_{ij})_{kl} = 0$ לכל k, l ולכל $i \neq j$ אם ורק אם $T(w_{jl}) \in \mathrm{Sp}(w_{jk}|\ 1 \leq k \leq n_j) = W_j$ לכל l אם ורק אם $T(W_j) \subseteq W_j$.

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, $A_{jj} = [T|_{W_j}]_{\mathcal{B}_j}^{\mathcal{B}_j}$.

(ב), (ג) נובעים מ־(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 $1 \leq i \leq s$. 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 $A = [T]_{\mathcal{B}}^{\mathcal{B}}$. 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 $A = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_s \end{pmatrix}$. ∎

למה 9.11: *יהי $V \subseteq W$ תת מרחב שמור־T. יהי $S = T|_W \colon W \to W$ 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 $i = 0$, $T^0 = 1V, S^0 = 1_W$, 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

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

(ב) לפי תרגיל 7.15 די להוכיח כי $m_T(S) = 0$. ואכן, $m_T(S)(w) = m_T(T)(w) = 0(w) = 0$ לכל $w \in W$, 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, $[T]_{\mathcal{B}} = \begin{pmatrix} A & C \\ 0 & D \end{pmatrix}$, באשר $A = [S]_{\mathcal{B}_0} \in \mathrm{M}_m(F)$. מכאן לפי תרגיל 6.12

$$f_T(X) = \begin{vmatrix} XI_m - A & -C \\ 0 & XI_k - D \end{vmatrix} = |XI_m - A| \cdot |XI_k - D| = f_S(X) \cdot f_D(X)$$

ובפרט $f_S | f_T$. ∎

משפט 9.12: *יהי V מרחב וקטורי מעל שדה F ותהי $T \colon V \to V$ 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־$rg + sh = 1$. 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 $2 \le i \le s$. 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

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

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־$r_1, \ldots, r_s \ge 1$. 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 $[T]_{\mathcal{B}}^{\mathcal{B}} = \operatorname{Diag}(A_1, \ldots, A_s)$, 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־$[T]_{\mathcal{B}}^{\mathcal{B}} = \operatorname{Diag}(A_1, \ldots, A_t)$, 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, $A_i = [T|_{W_i}]_{\mathcal{B}_i}^{\mathcal{B}_i}$, 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, $[T]_{\mathcal{B}}^{\mathcal{B}} = \operatorname{Diag}(A_1, \ldots, A_s)$, 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) לפי משפט 7.10, וכיון ש־$m_{A_1}, \ldots, m_{A_t}$ זרים,

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

לפי יחידות הפירוק לגורמים אי פריקים $t = s$ ו־$m_{A_i} = q_i^{r_i}$, עד כדי הסדר. לפי משפט 9.7 מתקיים $V = \operatorname{Sp}(\mathcal{B}_1) \oplus \cdots \oplus \operatorname{Sp}(\mathcal{B}_t)$. לפי משפט 9.9, $\operatorname{Sp}(\mathcal{B}_i)$ הוא שמור־T ו־$A_i = [T|_{\operatorname{Sp}(\mathcal{B}_i)}]_{\mathcal{B}_i}^{\mathcal{B}_i}$ לכל i. לכן $m_{T|_{\operatorname{Sp}(\mathcal{B}_i)}} = m_{A_i} = q_i^{r_i}$. 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, $v \in W_i$. מכאן $\operatorname{Sp}(\mathcal{B}_i) \subseteq W_i$, ובפרט $|\mathcal{B}_i| \le \dim W_i$. אבל $\sum_i |\mathcal{B}_i| = n = \sum_i \dim W_i$, 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-$A = [T]_{\mathcal{C}}^{\mathcal{C}}$. 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-$D = [T]_{\mathcal{B}}^{\mathcal{B}}$. 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. ∎

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־*$J = [T]_{\mathcal{B}}^{\mathcal{B}}$. *(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 בין השתיים הבאות

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

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

(ב) שתי המטריצות הבאות הן מטריצות ז'ורדן שונות (ממש, לא רק עד כדי סדר המערכים).

$$B = \begin{pmatrix} 1&0&0&0\\ 1&1&0&0\\ 0&0&1&0\\ 0&0&0&1 \end{pmatrix} \qquad A = \begin{pmatrix} 1&0&0&0\\ 1&1&0&0\\ 0&0&1&0\\ 0&0&1&1 \end{pmatrix}$$

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. ∎

הערה 10.5: 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־$N\mathbf{e}_1 = \mathbf{e}_2, N\mathbf{e}_2 = \mathbf{e}_3, \ldots, N\mathbf{e}_{n-1} = \mathbf{e}_n, N\mathbf{e}_n = 0$, 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 10.6: *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 מהערה 10.5.

(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 10.7: *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*

$$\operatorname{rank}(T-\lambda 1_V)^{k-1} - 2\operatorname{rank}(T-\lambda 1_V)^k + \operatorname{rank}(T-\lambda 1_V)^{k+1}.$$

בפרט, J יחידה, עד כדי סדר מערכיה (כפי שנטען במשפט 10.2).

הוכחה: לפי מסקנה 10.6(ב), מספר גושי ז'ורדן ב־J מסדר k השייכים ל־λ הוא

$$\left(\operatorname{rank}(J-\lambda I)^{k-1} - \operatorname{rank}(J-\lambda I)^k\right) - \left(\operatorname{rank}(J-\lambda I)^k - \operatorname{rank}(J-\lambda I)^{k+1}\right) =$$

$$\operatorname{rank}(J-\lambda I)^{k-1} - 2\operatorname{rank}(J-\lambda I)^k + \operatorname{rank}(J-\lambda I)^{k+1}$$

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 $\operatorname{rank}(J-\lambda I)^m = \operatorname{rank}(T-\lambda 1_V)^m$. ∎

משפט 10.8: *תהי J מטריצת ז'ורדן.*

(א) $m_{J_n(\lambda)} = (X-\lambda)^n$.

(ב) *נניח שמערכי J שייכים ל־ $\lambda_1,\dots,\lambda_s \in F$ השונים זה מזה והם בעלי אינדקסים $r_1,\dots,r_s$, 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 $\lambda_1,\dots,\lambda_s \in F$ השונים זה מזה ו־ $r_1,\dots,r_s \geq 1$. אז מערכי J שייכים ל־ $\lambda_1,\dots,\lambda_s$, והם בעלי אינדקסים $r_1,\dots,r_s$, 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־ $\lambda_1 = \bar{\lambda}_1,\dots,\lambda_r = \bar{\lambda}_s$, 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 10.9: סדרה $v = v_1,\dots,v_k \in V$ נקראת **שרשרת ז'ורדן** (**של** N) **מאורך** k אם

$$N(v_1) = v_2, N(v_2) = v_3, \dots, N(v_{k-1}) = v_k, N(v_k) = 0.$$

במלים אחרות: הסדרה היא $v, N(v), N^2(v), \dots, N^{k-1}(v) = \left(N^j(v)\right)_{j=0}^{k-1}$ ו־ $v \in \operatorname{Ker} N^k$.

תרגיל 10.10: *יהי $\mathcal{B}$ בסיס של V. אז*

(א) $[N]_{\mathcal{B}}^{\mathcal{B}} = J_k(0)$ *אם ורק אם $\mathcal{B}$ שרשרת ז'ורדן מאורך k.*

(ב) $[N]_{\mathcal{B}}^{\mathcal{B}}$ *מערך ז'ורדן השייך ל־ 0 ובעל אידקס $\leq r$ אם ורק אם $\mathcal{B}$ צירוף של שרשראות ז'ורדן שסדרת אורכיהם אינה עולה, כלומר, $\mathcal{B} = \bigvee_{m=r}^{1} \bigvee_{v\in\mathcal{C}_m} \left(N^j(v)\right)_{j=0}^{m-1}$ באשר $\mathcal{C}_m \subseteq \operatorname{Ker} N^m$ לכל m, כלומר, (אחרי שינוי משתנה $m = r+1-k$) $\mathcal{B} = \bigvee_{k=1}^{r} \bigvee_{v\in\mathcal{A}_k} \left(N^j(v)\right)_{j=0}^{r-k}$, באשר $\mathcal{A}_k \subseteq \operatorname{Ker} N^{r+1-k}$.*

נשתמש בתרגיל הבא:

תרגיל 10.11: *תהי $T\colon V \to W$ העתקה לינארית של מרחבים בעלי מימד סופי. יהי $u_1, \ldots, u_r$ בסיס של $\operatorname{Ker} T$ ויהיו $v_1, \ldots, v_s \in V$. אז*

(א) *$T(v_1), \ldots, T(v_s)$ בלתי תלויים לינארית אם ורק אם $u_1, \ldots, u_r, v_1, \ldots, v_s$ בלתי תלויים לינארית.*

(ב) *$T(v_1), \ldots, T(v_s)$ בסיס של $T(V)$ אם ורק אם $u_1, \ldots, u_r, v_1, \ldots, v_s$ בסיס של V, כלומר, $v_1, \ldots, v_s$ השלמה של בסיס של $\operatorname{Ker} T$ לבסיס של V. נאמר אז ש־* $v_1, \ldots, v_s$ **בסיס של** $V/\operatorname{Ker} T$.

(ג) *נניח $T(V) = W$ ותהי $S\colon W \to U$ העתקה לינארית. אז* $T^{-1}(\operatorname{Ker} S) = \operatorname{Ker}(S \circ T)$.

משפט 10.12: *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 $\geq r$.*

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 $V' = N(V)$ ויהי $N' = N|_{V'}$. אז $(N')^{r-1} = 0$. לפי הנחת האינדוקציה ותרגיל 10.10(ב) יש ל־V' בסיס

$$\mathcal{B}' = \bigvee_{k=1}^{r-1} \bigvee_{v' \in \mathcal{A}'_k} \left(N^j(v')\right)_{j=0}^{r-1-k},$$

באשר $\mathcal{A}'_k \subseteq \operatorname{Ker}(N')^{r-k}$ לכל k. 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 $\mathcal{B}'$

$$\bigvee_{k=1}^{r-1} \bigvee_{v' \in \mathcal{A}'_k} \left(N^{r-1-k}(v')\right) = \bigvee_{k=1}^{r-1} \bigvee_{v \in \mathcal{A}_k} \left(N^{r-k}(v)\right) = N^{r-1}(\mathcal{A}_1) \vee N^{r-2}(\mathcal{A}_2) \vee \ldots \vee \mathcal{A}_{r-1}$$

היא בלתי תלויה לינארית ומוכלת ב־$\operatorname{Ker} N$. נשלים אותה לבסיס 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 $\operatorname{Ker} N$.

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 10.11(ב), $\mathcal{B}$ בסיס של V. לפי תרגיל 10.10(ב), $[N]_{\mathcal{B}}^{\mathcal{B}}$ היא מערך ז׳ורדן השייך ל־0 ובעל אינדקס $\geq r$. ∎

בניה 10.13: 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 10.11(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 10.11(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 10.11(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 10.11(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 10.11(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, $\mathcal{B} = \bigvee_{k=1}^{r} \bigvee_{v\in\mathcal{A}_k} \big(N^j(v)\big)_{j=0}^{r-k}$ הבסיס המבוקש של V.

הוכחת משפט 10.2: נניח $m_T = (X-\lambda_1)^{r_1}\cdots(X-\lambda_s)^{r_s}$, באשר $\lambda_1,\dots,\lambda_s \in F$ 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 $1 \le k \le s$. נסמן $N = T|_{W_k} - \lambda_k 1_{W_k}\colon W_k \to W_k$. אז $m_N = X^{r_k}$. לפי משפט 10.12 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 10.7, 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 10.8(ג) מערכי J שייכים ל־$\lambda_1,\dots,\lambda_s$ והם מאינדקס $r_1,\dots,r_s$, בהתאמה, עד כדי הסדר. ∎

תרגיל 10.14: *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 $J = \operatorname{Diag}(J_{n_1}(\lambda_1),\dots,J_{n_s}(\lambda_s)) \in \mathrm{M}_n(F)$ (באשר $\lambda_1,\dots,\lambda_s$ לא בהכרח שונים זה מזה). אז הריבוי הגיאומטרי של λ ב־T הוא

$$\dim\operatorname{Ker}(T-\lambda 1_V) = \dim\operatorname{Ker}(J-\lambda I) = n - \operatorname{rank}(J-\lambda I) =$$
$$\sum_{i=1}^{s} n_i - \sum_{i=1}^{s}\operatorname{rank}J_{n_i}(\lambda_i-\lambda) = \sum_{i=1}^{s}\big(n_i - \operatorname{rank}J_{n_i}(\lambda_i-\lambda)\big)$$

ולפי מסקנה 10.6(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. ∎

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.

(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, $C_1(X-\lambda) = (\lambda) \in \mathrm{M}_1(F)$.

(ב) יהי $r \in \mathbb{N}$. המטריצה הבאה מסדר $dr \times dr$ תיקרא **גוש יעקובסון מאורך** r **השייך ל־**q

$$\left(\begin{array}{cccccccccccccc}
0 & \cdots & 0 & -a_0 & & & & & & & & & & \\
1 & \ddots & \vdots & \vdots & & & & & & & \mathbf{0} & & & \\
 & \ddots & 0 & \vdots & & & & & & & & & & \\
0 & \cdots & 1 & -a_{d-1} & & & & & & & & & & \\
 & & & 1 & 0 & \cdots & 0 & -a_0 & & & & & & \\
 & & & & 1 & \ddots & \vdots & \vdots & & & & & & \\
 & & & & & \ddots & 0 & \vdots & & & & & & \\
 & & & & 0 & \cdots & 1 & -a_{d-1} & & & & & & \\
 & & & & & & & 1 & & & & & & \\
 & & & & & & & & \ddots & & & & & \\
 & & & & & & & & & 1 & 0 & \cdots & 0 & -a_0 \\
 & & & & & & & & & & 1 & \ddots & \vdots & \vdots \\
 & & \mathbf{0} & & & & & & & & & \ddots & 0 & \vdots \\
 & & & & & & & & & & 0 & \cdots & 1 & -a_{d-1}
\end{array}\right)$$

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 $m_T = q_1^{r_1} \cdots q_s^{r_s}$, 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 $m_A = q_1^{r_1} \cdots q_s^{r_s}$, 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 $q \in F[X]$ מתוקן ממעלה d ותהי $C = C_r(q) \in \mathrm{M}_{dr}(F)$. אז $q^r = m_C = f_C$.*

הוכחה: תחילה נניח כי $r = 1$. לפי ההגדרה,

$$.f_C = \det(XI_d - C) = \det \begin{pmatrix} X & 0 & 0 & \cdots & 0 & a_0 \\ -1 & X & 0 & \cdots & 0 & a_1 \\ 0 & -1 & X & & & \\ . & . & \ddots & \ddots & . & \vdots \\ & & & -1 & X & a_{d-2} \\ 0 & 0 & 0 & \cdots & -1 & X + a_{d-1} \end{pmatrix}$$

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

$$f_C = \det\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & q(X) \\ -1 & 0 & 0 & \cdots & 0 & X^{d-1}+a_{d-1}X^{d-2}+\cdots+a_1 \\ 0 & -1 & 0 & & & \\ . & . & \ddots & \ddots & . & \vdots \\ & & & -1 & 0 & X^2+a_{d-1}X+a_{d-2} \\ 0 & 0 & 0 & \cdots & -1 & X+a_{d-1} \end{pmatrix}.$$

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 **6.20**. לפי תרגיל **7.12**, $m_C = f_C$. ∎

מסקנה 11.5: *תהי C מטריצת יעקובסון. יהיו $C_1,\dots,C_s$ 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), $q_1,\dots,q_s$ שונים.)

הוכחה: נקבע $1 \le i \le s$. אז $C_i = \mathrm{Diag}(C_{n_1}(q),\dots,C_{n_t}(q))$ באשר $r_i = n_1 \ge \cdots \ge n_t$. לפי תרגיל **6.20** ותרגיל **11.4**, $m_{C_i} = \mathrm{lcm}(m_{C_{n_1}(q)},\dots,m_{C_{n_t}(q)}) = \mathrm{lcm}(q^{n_1},\dots,q^{n_t}) = q^{r_i}$. לפי תרגיל **6.20**, $m_C = \mathrm{lcm}(m_{C_1},\dots,m_{C_s}) = q_1^{r_1}\cdots q_s^{r_s}$. ∎

הערה 11.6: סימון. מעתה יהי $q = \sum_{i=0}^{d-1} a_iX^i + X^d \in F[X]$. תהי $N = q(T)$. אז לכל $v \in V$

$$T^d(v) = -\sum_{i=0}^{d-1} a_iT^i(v) + N(v). \tag{1}$$

עבור $v \in V$ נסמן $\widetilde{v} = \big(v, T(v),\dots,T^{d-1}(v)\big)$ ועבור סדרה $\mathcal{A} = (v_1,\dots,v_k)$ של וקטורים ב־V נסמן $\widetilde{\mathcal{A}} = \big(v_1,\dots,v_k, T(v_1),\dots,T(v_k)\ ,\ \dots\ ,\ T^{d-1}(v_1),\dots,T^{d-1}v_k\big)$. 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.

לכל i,j מתקיים $T^iN^j = N^jT^i$ (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־ $v \in \operatorname{Ker} N^r$.

(ב) *אם $\mathcal{B}'$ מקיים את התנאים השקולים של (א) אז גם* $\mathcal{B} := \overbrace{\big(N^j(v)\big)_{j=0}^{r-1}}^{\sim}$ *בסיס של V, ו־* $[N]_{\mathcal{B}}^{\mathcal{B}} = \mathrm{Diag}\big(\overbrace{J_r(0),\dots,J_r(0)}^{d}\big)$.

הוכחה: (א) בדיקה ישירה, בעזרת (1).

(ב) נתבונן במלבן הבא של $d \cdot r$ וקטורים ב־V:

$$\begin{array}{cccc} v & N(v) & \dots & N^{r-1}(v) \\ T(v) & T(N(v)) & \dots & T(N^{r-1}(v)) \\ T^2(v) & T^2(N(v)) & \dots & T^2(N^{r-1}(v)) \\ & & \dots & \\ T^{d-1}(v) & T^{d-1}(N(v)) & \dots & T^{d-1}(N^{r-1}(v)) \end{array} \tag{2}$$

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 $0 \leq i \leq d-1$ 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 השייכת ל־0. 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) $$[T]^{\mathcal{B}'}_{\mathcal{B}'} = \mathrm{Diag}\big(\overbrace{C_r(q), \dots, C_r(q)}^{\ell_1}, \overbrace{C_{r-1}(q), \dots, C_{r-1}(q)}^{\ell_2}, \dots, \overbrace{C_1(q), \dots, C_1(q)}^{\ell_r}\big)$$
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 $\mathcal{B} \subseteq V$ סדרה כך ש־$\widetilde{\mathcal{B}}$ בלתי תלויה לינארית. אז יש $\mathcal{A} \subseteq V$ כך ש־$\widetilde{\mathcal{A}} \vee \widetilde{\mathcal{B}}$ בסיס של V. בפרט יש $\mathcal{A} \subseteq V$ כך ש־$\widetilde{\mathcal{A}}$ בסיס של V.*

הוכחה: באינדוקציה יורדת על $k = |\mathcal{B}|$. (מתקיים $dk = |\widetilde{\mathcal{B}}| \leq \dim V$.) אם $k = \frac{1}{d}\dim V$, אז $|\widetilde{\mathcal{B}}| = \dim V$, לכן $\widetilde{\mathcal{B}}$ בסיס של V וניקח $\mathcal{A} = \emptyset$.

נניח נכונות עבור k ויהי $\mathcal{B} = (v_1, \dots, v_{k-1})$. אז $|\widetilde{\mathcal{B}}| < \dim V$, לכן $\widetilde{\mathcal{B}}$ בלתי תלויה לינארית לא מרבית, ולכן יש $v = v_k \in V$ כך ש־$\widetilde{\mathcal{B}} \vee (v)$ בלתי תלויה לינארית. נראה ש־$\widetilde{\mathcal{B}} \vee \widetilde{v}$ בלתי תלויה לינארית. אז, לפי הנחת האינדוקציה, יש $\mathcal{A}'$ כך ש־ $\widetilde{\mathcal{B}} \vee \widetilde{v} \vee \widetilde{\mathcal{A}'}$ בסיס של V; ניקח $\mathcal{A} = (v) \vee \mathcal{A}'$ וסיימנו.

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 $i \le k-1$.

נניח שלילה. לכל $1 \le i \le k$ יהי $f_i = \sum_{j=0}^{d-1} a_{ij}X^j \in F[X]$. אז $f_k \ne 0$, $\deg f_k < d = \deg q$. 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 עם שארית לכל $1 \le i \le k-1$ 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

$$0 = h(T)\big(\sum_{i=1}^{k} f_i(T)(v_i)\big) = \sum_{i=1}^{k} h(T)f_i(T)(v_i) =$$
$$\sum_{i=1}^{k-1} g_i(T)q(T)(v_i) + \sum_{i=1}^{k-1}\sum_{j=0}^{d-1} a'_{ij}T^j(v_i) + h(T)f_k(T)(v).$$

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־$1 \ne 0$, סתירה להנחה $\widetilde{\mathcal{B}} \vee (v)$ בלתי תלויה לינארית. ∎

כעת נוכיח מקרה פרטי של משפט 11.2:

משפט 11.10: *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־ $C = [T]_{\mathcal{B}'}^{\mathcal{B}'}$. יתר על כן, C מערך יעקובסון בעל אינדקס r השייך ל־ q.*

הוכחה: נסמן $N = q(T)$. אז $N^r = 0$.

קיום: נחזור על ההוכחה של משפט 10.12 עבור N, רק 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 למה 11.9 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 $r \ge 2$ והמשפט נכון עבור $r - 1$.

יהי $V' = N(V)$ ויהי $N' = N|_{V'}$. אז $(N')^{r-1} = 0$. לפי הנחת האינדוקציה ותרגיל 10.10(ב) יש ל־V' בסיס

$$\mathcal{B}'' = \bigvee_{k=1}^{r-1} \bigvee_{v' \in \widetilde{\mathcal{A}'_k}} \big(N^j(v')\big)_{j=0}^{r-1-k},$$

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 $\widetilde{\mathcal{A}'_k} = \widetilde{N(\mathcal{A}_k)} = N(\widetilde{\mathcal{A}_k})$, 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 למה 11.9 (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 10.12 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 $1 \leq k \leq r$ 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:

$$[T]_{\mathcal{B}'}^{\mathcal{B}'} = \operatorname{Diag}\Big(\overbrace{C_r(q), \ldots, C_r(q)}^{\ell_1}, \overbrace{C_{r-1}(q), \ldots, C_{r-1}(q)}^{\ell_2}, \ \ldots,\ \overbrace{C_1(q), \ldots, C_1(q)}^{\ell_r}\Big).$$

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 10.7) 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 11.10 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 $[T]_{\mathcal{B}}^{\mathcal{B}} = \operatorname{Diag}(C_1, \ldots, C_t)$, באשר $C_1, \ldots, C_t$ מערכי יעקובסון השייכים לחזקות של פולינומים אי פריקים מתוקנים שונים, אז לפי מסקנה 11.5, $m_{C_1}, \ldots, m_{C_t}$ חזקות של פולינומים אי פריקים שונים. לכן לפי מסקנה 9.15(ב), $t = s$ ויש בסיס $\mathcal{B}_i$ 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 11.10, C_i יחידה. ∎

תרגיל 11.11: *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־$\sim$ יחס הדמיון.)

תרגיל 11.12: *תהי* $A \in \mathrm{M}_n(F)$. *אז* $A \sim A^t$. *(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־ $A \sim C$, לפי תרגיל 11.11, $A^t \sim C^t$. אם $C \sim C^t$, אז, כיוון שדמיון מטריצות הוא יחס שקילות, $A \sim A^t$.
די להניח כי A גוש יעקובסון. אכן, $A = \operatorname{Diag}(C_1, \ldots, C_t)$, 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 $P := \operatorname{Diag}(P_1, \ldots, P_t)$ מקיימת

$$P^{-1}AP = \operatorname{Diag}(P_1^{-1} C_1 P_1, \ldots, P_t^{-1} C_t P_t) = \operatorname{Diag}(C_1^t, \ldots, C_t^t) = A^t.$$

נניח, אם כן, כי $A = C_r(q)$. לפי תרגיל 11.4, $m_A = q^r$. לפי תרגיל 11.11, $m_{A^t} = q^r$. תהי B צורת יעקובסון של A^t, אז $B \sim A^t$ 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, $A \sim A^t$. ∎

תרגיל 11.13: *יהיו $F \subseteq F'$ שני שדות. תהיינה $A, B \in \mathrm{M}_n(F)$ דומות מעל F'. אז הן דומות גם מעל F.*

הדרכה: באינדוקציה על n.

טענה א: *אפשר להחליף את A, B במטריצות דומות להן מעל F.*

לכן בלי הגבלת הכלליות A, B מטריצות יעקובסון מעל F וצריך להוכיח $A = B$, עד כדי סדר הגושים. לפי ההנחה $m_A = m_B$, 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. נסמן $C_1 = C_{r_1}(q_1)$.

טענה ב: $A = \mathrm{Diag}(C_1, A_1), B = \mathrm{Diag}(C_1, B_1)$*, באשר A_1, B_1 מטריצות 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: $\mathrm{Diag}(C', A'), \mathrm{Diag}(C', B')$ *הן צורות יעקובסון של A, B מעל F', עד כדי סדר הגושים.*

לפי ההנחה הן זהות, עד כדי סדר הגושים. לכן $A' = B'$. בפרט A_1, B_1 דומות מעל F'. לפי הנחת האינדוקציה הן דומות מעל F. לכן הן שוות. מכאן $A = B$, עד כדי סדר הגושים. ∎

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

$$\overline{x+iy}=x-iy, \quad |x+iy|=\sqrt{x^2+y^2}, \qquad x,y\in\mathbb{R}.$$

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 $F \to V\times V$ 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,

$$\langle\begin{pmatrix}a_1\\a_2\\\vdots\\,a_n\end{pmatrix},\begin{pmatrix}b_1\\b_2\\\vdots\\,b_n\end{pmatrix}\rangle_{\mathrm{st}}=(\overline{a_1},\overline{a_2},\dots,\overline{a_n})\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}=\overline{a_1}b_1+\overline{a_2}b_2+\cdots+\overline{a_n}b_n=\sum_{i=1}^n\overline{a_i}b_i.$$

נבדוק שהתכונה (4) בהגדרה מתקיימת: נניח $a_1,a_2,\dots,a_n\in F$ לא כולם אפס. אז

$$\langle\begin{pmatrix}a_1\\a_2\\\vdots\\,a_n\end{pmatrix},\begin{pmatrix}a_1\\a_2\\\vdots\\,a_n\end{pmatrix}\rangle_{\mathrm{st}}=\sum_{i=1}^n\overline{a_i}a_i=\sum_{i=1}^n|a_i|^2>0,$$

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:

$$\langle\begin{pmatrix}a_1\\a_2\end{pmatrix},\begin{pmatrix}a_1\\a_2\end{pmatrix}\rangle=2\overline{a_1}a_1+\overline{a_1}a_2+\overline{a_2}a_1+\overline{a_2}a_2=(\overline{a_1}+\overline{a_2})(a_1+a_2)+\overline{a_1}a_1$$
$$=|a_1+a_2|^2+|a_1|^2>0$$

כאשר $a_2\neq 0$ 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) $\langle 0,v\rangle=\langle u,0\rangle=0$.

הגדרה 12.4:

(1) $||v||=\sqrt{\langle v,v\rangle}$ נקרא **הנורמה** (**האורך**) של v.

(2) u **נִצָּב ל־**v – נכתוב $u\perp v$ – אם $\langle u,v\rangle=0$. זה שקול ל־$\langle v,u\rangle=0$.

(3) סדרה $u_1,u_2,\dots,u_n\in V$ נקראת **אורתוגונלית** אם $u_i\perp u_j$ לכל $i\neq j$.

(4) סדרה $u_1,u_2,\dots,u_n\in V$ 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 $\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n$ 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, $\langle\vec{OA},\vec{OB}\rangle_{\mathrm{st}}=0$.**

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־$u_1,\dots,u_k$ אז הוא ניצב לכל u ב־$\mathrm{Sp}(u_1,\dots,u_k)$.*

למה 12.8 (למת האיפיון): *תהי $u_1,u_2,\dots,u_n$ סדרה אורתוגונלית ב־V. יהיו*

$$u=a_1u_1+\cdots+a_nu_n,\qquad v=b_1u_1+\cdots+b_nu_n\in V.$$

(א) $\langle u_i,u\rangle=a_i||u_i||^2$. *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* $\langle u,v\rangle=\langle[u]_{\mathcal{B}},[v]_{\mathcal{B}}\rangle_{\text{st}}$

משפט 12.10: *סדרה אורתוגונלית* $u_1,u_2,\ldots,u_n$ *שאבריה שונים מאפס הינה בלתי תלויים לינארית מעל* F.

הוכחה: נניח $a_1u_1+a_2u_2+\cdots+a_nu_n=0$. 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:*

(1^k) $\text{Sp}(v_1,v_2,\ldots,v_k)=\text{Sp}(u_1,u_2,\ldots,u_k)$, *לכל* $1\le k\le n$.

(2^k) $u_k=0$ *אם ורק אם* $v_k\in\text{Sp}(v_1,v_2,\ldots,v_{k-1})$, *לכל* $1\le k\le n$.

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 שיתקיימו תנאים (1^k), (2^k), וכן

(0^k) $u_1,u_2,\ldots,u_k$ אורתוגונלית.

עבור $k=1$ נגדיר $u_1=v_1$. אז (0^1), (1^1) ברורים; (2^1): $v_1\in\text{Sp}(\emptyset)\Leftrightarrow u_1=v_1=0$.

נניח באינדוקציה שכבר בנינו $u_1,u_2,\ldots,u_k$ שמקיימים את (0^k), (1^k), (2^k). נגדיר

$$u_{k+1}=v_{k+1}-(c_1u_1+c_2u_2+\cdots+c_ku_k), \qquad (*)$$

באשר $c_i=\begin{cases}\frac{\langle u_i,v_{k+1}\rangle}{||u_i||^2} & u_i\ne 0\\ 0 & u_i=0\end{cases}$ לכל $1\le i\le k$.

אז $\langle u_i,v_{k+1}\rangle=c_i||u_i||^2$ לכל $1\le i\le k$ (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 $1\le i\le k$. 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):

$$\mathrm{Sp}(u_1, u_2, \dots, u_k, u_{k+1}) = \mathrm{Sp}(u_1, u_2, \dots, u_k, v_{k+1}) = \mathrm{Sp}(v_1, v_2, \dots, v_k, v_{k+1}).$$

(2^{k+1}) אם $u_{k+1} = 0$ אז לפי $(*)$

$$v_{k+1} = c_1u_1 + c_2u_2 + \dots + c_ku_k \in \mathrm{Sp}(u_1, u_2, \dots, u_k) = \mathrm{Sp}(v_1, v_2, \dots, v_k).$$

להיפך, אם $v_{k+1} \in \mathrm{Sp}(v_1, v_2, \dots, v_k) = \mathrm{Sp}(u_1, u_2, \dots, u_k)$ אז $v_{k+1} = \sum_{i=1}^k a_iu_i$ עבור איזה $a_1, a_2, \dots, a_k \in F$. 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) $u_1 = v_1 = (0, 3, 4)$, $\quad ||u_1||^2 = 0^2 + 3^2 + 4^2 = 25$,

(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) $\langle u_2, v_3\rangle_{\mathrm{st}} = 1 \cdot 0 + \frac{4}{5} \cdot 0 + \frac{-3}{5} \cdot 2 = \frac{-6}{5}$, $\quad \langle u_1, v_3\rangle_{\mathrm{st}} = 0 \cdot 0 + 3 \cdot 0 + 4 \cdot 2 = 8$.

$$u_3 = (0, 0, 2) - \tfrac{8}{25}(0, 3, 4) - \tfrac{-\frac{6}{5}}{2}(1, \tfrac{4}{5}, \tfrac{-3}{5}) = \tfrac{1}{25}(15, -12, 9).$$

מסקנה 12.13: *אם V נוצר סופית אז יש לו בסיס אורתונורמלי.*

הוכחה: ל־V יש בסיס $v_1, v_2, \dots, v_n$. לפי המשפט הקודם יש לו בסיס אורתוגונלי $u_1, u_2, \dots, u_n$. קל לבדוק כי $\frac{1}{||u_1||}u_1, \frac{1}{||u_2||}u_2, \dots, \frac{1}{||u_n||}u_n$ סדרה אורתונורמלית. ∎

הערה 12.14: המשמעות הגיאומטרית של תהליך גרם־שמידט. נתבונן בנוסחה $(*)$. נסמן

$$U = \mathrm{Sp}(v_1, v_2, \dots, v_k) = \mathrm{Sp}(u_1, u_2, \dots, u_k), \qquad u = c_1u_1 + c_2u_2 + \dots + c_ku_k.$$

אז $v_{k+1} = u_{k+1} + u$, $u \in U$, ולפי תרגיל 12.7, u_{k+1} ניצב לכל הוקטורים ב־U.

ניח 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* $v_1,\dots,v_r,v_{r+1},\dots,v_n$ *ו־* $v_1,\dots,v_r$ *סדרה אורתוגונלית, אז* $u_1=v_1,u_2=v_2,\dots,u_r=v_r$.

הוכחה: באינדוקציה: לפי הבניה, תמיד $u_1=v_1$.

נניח $u_1=v_1,\dots,u_k=v_k$, באשר $1\le k<r$. 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* $v_1,\dots,v_r$ *סדרה אורתוגונלית, שאיבריה שונים מאפס. אז אפשר להשלימה לבסיס אורתוגונלי של* V.

הוכחה: לפי משפט 12.10, $v_1,\dots,v_r$ בלתי תלויה לינארית. לכן ניתן להשלימה לבסיס $v_1,\dots,v_r,\dots,v_n$ של V. תהליך גרם־שמידט עליו יתן בסיס אורתוגונלי $u_1,\dots,u_r,\dots,u_n$. לפי תרגיל 12.16, $u_1=v_1,\dots,u_r=v_r$.
∎

משפט 12.18 (אי שוויון Cauchy-Schwarz): *יהיו* v_1,v_2 *שני וקטורים ב־*V. *אז מתקיים* $|\langle v_1,v_2\rangle|\le ||v_1||\cdot||v_2||$ *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 $u_1=v_1$, $u_2=v_2-c_1u_1$, 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),

$$||v_2||^2=|c_1|^2\cdot||u_1||^2+1\cdot||u_2||^2=\frac{|\langle v_1,v_2\rangle|^2}{||u_1||^2}+||u_2||^2\ge\frac{|\langle v_1,v_2\rangle|^2}{||u_1||^2}$$

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* $||v_1+v_2||\le||v_1||+||v_2||$.

הוכחה: נשתמש באי שוויון קושי־שוורץ כדי להוכיח $(||v_1+v_2||)^2\le(||v_1||+||v_2||)^2$:

$$(||v_1+v_2||)^2=\langle v_1+v_2,v_1+v_2\rangle=\langle v_1,v_1\rangle+\langle v_1,v_2\rangle+\langle v_2,v_1\rangle+\langle v_2,v_2\rangle=$$

$$||v_1||^2+\langle v_1,v_2\rangle+\overline{\langle v_1,v_2\rangle}+||v_2||^2=||v_1||^2+2\mathrm{Re}(\langle v_1,v_2\rangle)+||v_2||^2\le$$

$$||v_1||^2+2|\langle v_1,v_2\rangle|+||v_2||^2\le||v_1||^2+2||v_1||\cdot||v_2||+||v_2||^2=(||v_1||+||v_2||)^2.$$ ∎

הגדרה 12.20: *יהי U תת מרחב של V.* **המשלים הניצב (האורתוגונלי)** הוא

$$U^\perp = \{v \in V |\ \langle u, v\rangle = 0 \text{ לכל } u \in U\}.$$

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 $\langle u, u\rangle = 0 = ||u||^2$, 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 $\mathrm{Sp}(u_{r+1}, \ldots, u_n) \subseteq U^\perp$. 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* $(U^\perp)^\perp = U$.

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). לפי למת האיפיון 12.8(א), $\langle u_i, w\rangle = \langle T(u_i), v\rangle$ 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, לפי למת האיפיון 12.8(א), $u = \sum_{i=1}^n \langle u_i, u\rangle u_i$, לכן $T(u) = \sum_{i=1}^n \langle u_i, u\rangle T(u_i)$. 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 $T^*(av) = aT^*(v)$.

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 $T^*(v_1+v_2) = T^*(v_1)+T^*(v_2)$. ∎

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 למת האיפיון 12.8(א), הרכיב ה־j של העמודה ה־i של $[T]_{\mathcal{B}}$ הוא $\langle u_j, T(u_i)\rangle$. לכן

$$([T]_{\mathcal{B}}^*)_{ij} = \overline{([T]_{\mathcal{B}})_{ji}} = \overline{\langle u_j, T(u_i)\rangle} = \langle T(u_i), u_j\rangle = \langle u_i, T^*(u_j)\rangle = ([T^*]_{\mathcal{B}})_{ij}. \qquad \blacksquare$$

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')

$$\big((AB)^*\big)_{ij} = \overline{(AB)_{ji}} = \overline{\sum_{k=1}^{n}(A)_{jk}(B)_{ki}} = \sum_{k=1}^{n}\overline{(A)_{jk}(B)_{ki}} = \sum_{k=1}^{n}\overline{(A)_{jk}}\;\overline{(B)_{ki}} =$$
$$= \sum_{k=1}^{n}(B^*)_{ik}(A^*)_{kj} = (B^*A^*)_{ij}$$

לכל i, j ולכן $(AB)^* = B^*A^*$.

אפשר לתת גם הוכחה ישירה, לא דרך המטריצות. למשל, (ד): לכל $u \in V$

$$\langle T^*(u), v\rangle = \overline{\langle v, T^*(u)\rangle} = \overline{\langle T(v), u\rangle} = \langle u, T(v)\rangle$$

ומכאן לפי ההגדרה $(T^*)^*(v) = T(v)$. $\blacksquare$

הגדרה 13.5: תהי V $T\colon V\to V$ 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 $T = T^* \Leftrightarrow \langle v, T(v)\rangle \in \mathbb{R}$ לכל $v \in V$.*

הוכחה: נניח $T = T^*$. אז $\overline{\langle v, T(v)\rangle} = \langle T(v), v\rangle = \langle v, T^*(v)\rangle = \langle v, T(v)\rangle$. 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.*

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 *לְבסיס אורתונורמלי של* V.

(4) T *מעתיקה בסיס אורתונורמלי מסוים של* V *לְבסיס 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) ⇐(2): אם $T^*T = 1_V$, אז T, T^* הפיכות (T חח"ע ו־T^* על. בדוק!) לכן $T^* = T^{-1}$. מכאן $TT^* = 1_V$.

(2) ⇐(1): באותו אופן (החלף בין T לבין T^*).

(1) ⇐(3): יהי $u_1, \ldots, u_n$ בסיס אורתונורמלי של V. אז לכל $1 \leq i, j \leq n$

$$\langle T(u_i), T(u_j)\rangle = \langle u_i, T^*T(u_j)\rangle = \langle u_i, 1_V(u_j)\rangle = \langle u_i, u_j\rangle = \delta_{ij},$$

$T(u_1), \ldots, T(u_n)$ אורתונורמלית. לכן היא בלתי תלויה לינארית: היות ו־$n = \dim V$, היא אף בסיס.

(3) ⇐(4): טריביאלי.

(4) ⇐(5): נניח כי $u_1,\ldots,u_n$ בסיס אורתונורמלי וגם $T(u_1),\ldots,T(u_n)$ 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) ⇐(6): $||T(u)||^2=\langle T(u),T(u)\rangle=\langle u,u\rangle=||u||^2$, לכן $||T(u)||=||u||$.

(6) ⇐(7): $||T(u)-T(v)||=||T(u-v)||=||u-v||$.

(7) ⇐(1): יהי $u\in V$. לפי ההנחה (עם $v=0$): $||T(u)||=||u||$. לכן

$$\langle u,T^*T(u)\rangle=\langle T(u),T(u)\rangle=||T(u)||^2=||u||^2=\langle u,u\rangle=\langle u,1_V(u)\rangle.$$

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)⇔(ב):

תהיינה $u_1,\ldots,u_n$ העמודות של A. אז $\overline{u_1^t},\ldots,\overline{u_n^t}$ הן השורות של A^*. מכאן

$$(A^*A)_{ij}=\sum_{k=1}^n\overline{(u_i^t)}_k\,(u_j)_k=\sum_{k=1}^n\overline{(u_i)}_k\,(u_j)_k=\langle u_i,u_j\rangle_{\text{st}}.$$

לכן $u_1,\ldots,u_n$ אורתונורמלית $\Leftrightarrow$ $\langle u_i,u_j\rangle_{\text{st}}=\delta_{ij}$ לכל i,j $\Leftrightarrow A^*A=I$.

(א)⇔(ג): תהיינה $v_1,\ldots,v_n$ השורות של A. אז $\overline{v_1^t},\ldots,\overline{v_n^t}$ הן העמודות של A^*. מכאן

$$(AA^*)_{ij}=\sum_{k=1}^n(v_i)_k\,\overline{(v_j^t)}_k=\sum_{k=1}^n\overline{(v_j)}_k\,(v_i)_k=\langle v_j,v_i\rangle_{\text{st}}.$$

לכן $AA^*=I\Leftrightarrow\langle v_j,v_i\rangle_{\text{st}}=\delta_{ji}\Leftrightarrow$ לכל i,j $v_1,\ldots,v_n$ 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 $\mathcal{B} = (u_1, \ldots, u_n)$, $\mathcal{B}' = (u'_1, \ldots, u'_n)$. אז $A = \big([u'_1]_{\mathcal{B}}, \ldots, [u'_n]_{\mathcal{B}}\big)$, 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 $i, j \Leftrightarrow u'_1, \ldots, u'_n$ 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.) ∎

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 $\langle w, T^*(v) \rangle = \langle T(w), v \rangle = 0$, 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 $1\leq i\leq n$. 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 $\deg f_S\geq 1$. 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, $D = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$, באשר $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ השרשים של f_A. אז $D^* = P^*A^*P$ ומכאן בקלות:

A צמודה לעצמה אם ורק אם D צמודה לעצמה; A אוניטרית אם ורק אם D אוניטרית .

כמו כן $f_A = f_D$ כי A, D דומות. לכן בלי הגבלת הכלליות $A = D$ אלכסונית.

(א) D צמודה לעצמה $\Leftrightarrow \mathrm{Diag}(\lambda_1, \dots, \lambda_n) = \mathrm{Diag}(\overline{\lambda_1}, \dots, \overline{\lambda_n}) \Leftrightarrow \overline{\lambda_i} = \lambda_i$ לכל $1 \le i \le n$ $\Leftrightarrow \lambda_1, \dots, \lambda_n$ ממשיים.

(ב) D אוניטרית $\Leftrightarrow D^*D = I_n \Leftrightarrow \mathrm{Diag}(\overline{\lambda_1}, \dots, \overline{\lambda_n})\mathrm{Diag}(\lambda_1, \dots, \lambda_n) = \mathrm{Diag}(1, \dots, 1) \Leftrightarrow |\lambda_i|^2 = \overline{\lambda_i}\lambda_i = 1$ 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־ $[T]_{\mathcal{B}}$ מהצורה*

$$\begin{array}{c} \overbrace{\hphantom{1 \; \ddots \; 1}}^{p} \quad \overbrace{\hphantom{-1 \; \ddots \; -1}}^{q} \hphantom{\cos\theta_1 \; -\sin\theta_1 \; \ddots \; \cos\theta_r \; -\sin\theta_r} \\ \begin{pmatrix} 1 & & & & & & & & & & & \\ & \ddots & & & & & & & & & \mathbf{0} & \\ & & 1 & & & & & & & & & \\ & & & -1 & & & & & & & & \\ & & & & \ddots & & & & & & & \\ & & & & & -1 & & & & & & \\ & & & & & & \cos\theta_1 & -\sin\theta_1 & & & & \\ & & & & & & \sin\theta_1 & \cos\theta_1 & & & & \\ & & & \mathbf{0} & & & & & \ddots & & & \\ & & & & & & & & & & \cos\theta_r & -\sin\theta_r \\ & & & & & & & & & & \sin\theta_r & \cos\theta_r \end{pmatrix} \end{array}$$

באשר $p, q, r \geq 0$.

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

$$f_T = f_A = X^2 - \mathrm{tr}(A)X + \det A = X^2 - 1 = (X-1)(X+1),$$

ו־T נורמלית, לכן לפי למה 13.19 יש בסיס אורתונורמלי אחר $\mathcal{B}'$ כך ש־ $[T]_{\mathcal{B}'}^{\mathcal{B}'} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

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 $W = \mathrm{Sp}(w)$ 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 $W = \mathrm{Sp}(v, T(v))$ 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־$[T]_{\mathcal{B}_i}^{\mathcal{B}_i}$ מהצורה המבוקשת. אז $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ הוא בסיס של V, הוא אורתונורמלי, ו־ $[T]_{\mathcal{B}} = \begin{pmatrix} [T]_{\mathcal{B}_1}^{\mathcal{B}_1} & 0 \\ 0 & [T]_{\mathcal{B}_2}^{\mathcal{B}_2} \end{pmatrix}$ היא מהצורה המבוקשת (עד כדי הסדר של אברי $\mathcal{B}$).

∎

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 $f(v,w)=\varphi(v)\varphi(w)$, 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 $f(v,w)=v^tAw\in \mathrm{M}_1(F)=F$ (מזהים $\mathrm{M}_1(F)$ עם F). במפורש: אם $A=(a_{ij})$, $v=(x_1,\dots,x_m)^t$, $w=(y_1,\dots,y_n)^t$, אז

$$f(v,w)=(x_1,\dots,x_m)\begin{pmatrix} a_{11} & \cdots & a_{1n}\\ -- & \cdots & --\\ a_{m1} & \cdots & a_{mn}\end{pmatrix}\begin{pmatrix} y_1\\ \vdots\\ y_n\end{pmatrix}=$$

$$(x_1,\dots,x_m)\begin{pmatrix} \sum_{j=1}^n a_{1j}y_j\\ \vdots\\ \sum_{j=1}^n a_{mj}y_j\end{pmatrix}=\sum_{i=1}^m\sum_{j=1}^n x_ia_{ij}y_j \qquad \blacksquare$$

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 $\mathrm{Bil}(V,W)$.

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* $A = \theta_{\mathcal{C}}^{\mathcal{B}}(f) \in \mathrm{M}_{m\times n}(F)$ *כך שמתקיים*

$$(1) \qquad f(v,w) = \big([v]_{\mathcal{B}}\big)^t A[w]_{\mathcal{C}} \qquad \text{לכל } v \in V,\ w \in W.$$

(אומרים ש- A **מייצגת את** f *לפי* $\mathcal{B},\mathcal{C}$*.)*

(ב) *המטריצה* $A = \theta_{\mathcal{C}}^{\mathcal{B}}(f)$ *מוגדרת על ידי*

$$(2) \qquad 1 \le i \le m,\ 1 \le j \le n \qquad (A)_{ij} = f(v_i, w_j).$$

(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-* $A = \theta_{\mathcal{C}}^{\mathcal{B}}(f)$.

(ד) *התאמה* $A \mapsto f$ *הנ"ל היא איזומורפיזם של מרחבים וקטוריים* $\theta_{\mathcal{C}}^{\mathcal{B}}\colon \mathrm{Bil}(V,W) \to \mathrm{M}_{m\times n}(F)$.

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: נגדיר את A על ידי (2). נניח $v = \sum_{i=1}^m x_i v_i,\ w = \sum_{j=1}^n y_j w_j$. אז

$$f(v,w) = f\Big(\sum_{i=1}^m x_i v_i, \sum_{j=1}^n y_j w_j\Big) = \sum_{i=1}^m\sum_{j=1}^n x_i y_j f(v_i,w_j),$$

$$\big([v]_{\mathcal{B}}\big)^t A[w]_{\mathcal{C}} = (x_1,\dots,x_m)A\begin{pmatrix} y_1\\ \vdots \\ y_n\end{pmatrix} = \sum_{i=1}^m\sum_{j=1}^n x_i (A)_{ij} y_j = \sum_{i=1}^m\sum_{j=1}^n x_i y_j f(v_i,w_j).$$

(ג) קל לראות ש-f תבנית בילינארית. לפי היחידות בחלק (א), $A = \theta_{\mathcal{C}}^{\mathcal{B}}(f)$.

(ד) לפי (ג), $\theta_{\mathcal{C}}^{\mathcal{B}}$ על. נראה שהיא לינארית וחד חד ערכית.

טענה: $\theta_{\mathcal{C}}^{\mathcal{B}}$ *לינארית.* תהיינה $f,g \in \mathrm{Bil}(V,W)$, ותהיינה $A = \theta_{\mathcal{C}}^{\mathcal{B}}(f), B = \theta_{\mathcal{C}}^{\mathcal{B}}(g)$. יהי $\alpha \in F$. אז

$$(f+g)(v_i,w_j) = f(v_i,w_j) + g(v_i,w_j) = (A)_{ij} + B_{ij} = (A+B)_{ij},$$

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

כלומר, $\theta_{\mathcal{C}}^{\mathcal{B}}(f+g) = A+B$, $\theta_{\mathcal{C}}^{\mathcal{B}}(\alpha f) = \alpha A$.

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 $v\in V, w \in W$

$$f(v,w) = \big([v]_{\mathcal{B}}\big)^t A[w]_{\mathcal{C}} = g(v,w).$$

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 $A=\theta^{\mathcal{B}}_{\mathcal{C}}(f)$, $A'=\theta^{\mathcal{B}'}_{\mathcal{C}'}(f)$. אז $A'=P^tAQ$, 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)

$$f(v,w)=\big([v]_{\mathcal{B}}\big)^tA[w]_{\mathcal{C}}=\big(P[v]_{\mathcal{B}'}\big)^tA\big(Q[w]_{\mathcal{C}'}\big)=\big([v]_{\mathcal{B}'}\big)^tP^tAQ[w]_{\mathcal{C}'}.$$

מכאן, לפי היחידות במשפט 14.3(א), $A'=P^tAQ$. ∎

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־$A'=P^tAQ$.

מטריצות ריבועיות $A,A'\in\mathrm{M}_n(F)$ נקראות **חופפות** אם יש $P\in\mathrm{M}_n(F)$ הפיכה כך ש־$A'=P^tAP$.

∎

תרגיל 14.7: *חפיפה ושקילות הם יחסי שקילות. בפרט: אם A חופפת ל־B ו־B חופפת ל־C אז A חופפת ל־C.*

מסקנה 14.8: *(א) $A,A'\in\mathrm{M}_{m\times n}(F)$ שקולות אם ורק אם הן מייצגות אותה תבנית בילינארית (כ"א ביחס לבסיס מתאים).*

(ב) $A,A'\in\mathrm{M}_{m\times n}(F)$ שקולות אם ורק אם $\mathrm{rk}\,A=\mathrm{rk}\,A'$.

(ג) $A,A'\in\mathrm{M}_n(F)$ חופפות אם ורק אם הן מייצגות אותה תבנית בילינארית על מרחב וקטורי V ממימד n (כ"א ביחס לבסיס מתאים).

הוכחה: נוכיח רק (ג): לפי משפט 14.5, אם $A=\theta^{\mathcal{B}}_{\mathcal{B}}(f)$ ו־$A'=\theta^{\mathcal{B}'}_{\mathcal{B}'}(f)$, אז $A'=P^tAP$, 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־$A=\theta^{\mathcal{B}}_{\mathcal{B}}$. לפי משפט 14.5, $A'=\theta^{\mathcal{B}'}_{\mathcal{B}'}(f)$. ∎

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 $\mathcal{B} = (v_1, \ldots, v_n)$.

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)

$$f(w,v) = \left([w]_\mathcal{B}\right)^t A[v]_\mathcal{B} = \left(\left([w]_\mathcal{B}\right)^t A[v]_\mathcal{B}\right)^t = \left([v]_\mathcal{B}\right)^t A^t[w]_\mathcal{B} = \left([v]_\mathcal{B}\right)^t A[w]_\mathcal{B} = f(v,w).$$

■

למה 14.12: *תהי* $f \in \mathrm{Bil}(V,V)$ *סימטרית. אזי*

(א) *(למת הקיטוב:)* $2f(v,w) = f(v+w, v+w) - f(v,v) - f(w,w)$ *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* $\mathcal{B} = (u_1, \ldots, u_n)$ *של* V*, כך ש־*$f(u_i, u_j) = 0$ *לכל* $i \neq j$*. בפרט*

(א) *המטריצה של* f *לפי* $\mathcal{B}$ *היא אלכסונית* $\mathrm{Diag}(c_1, \ldots, c_n)$*, 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 (בדוק!).

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: נגדיר v_1, v_2 על פי (3). אז $v_1 \in V_1$, $v_2 \in V_2$ (בדוק!), ומתקיים $v = v_1 + v_2$.

כעת $\dim V = \dim V_1 + \dim V_2$, 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 $u_2, \ldots, u_n$ של V_2 כך ש־$f(u_i, u_j) = 0$ לכל $i \neq j$. לפי הטענה, $u_1, u_2, \ldots, u_n$ 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 $D_r = \mathrm{Diag}(\overbrace{1, \ldots, 1}^{r}, 0, \ldots, 0)$.*

(ב) $r = \mathrm{rk}\, f$ *(ולכן נקבע על ידי f).*

הוכחה: (א) לפי משפט 11.13 יש בסיס $u_1, \ldots, u_n$ 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, $c_{r+1}, \ldots, c_n = 0$, $c_1, \ldots, c_r \neq 0$.

נגדיר בסיס אחר $u_1', \ldots, u_n'$ 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 $u_1', \ldots, u_n'$ 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 $\Leftrightarrow \operatorname{rk} A = \operatorname{rk} B$.*

הוכחה: (ב) $\Leftarrow$: טענה 14.9; $\Rightarrow$: נניח $\operatorname{rk} A = \operatorname{rk} B = r$. אז A, B חופפות ל־D_r, ולכן חופפות. ∎

משפט 14.18: *יהי V מרחב וקטורי מעל $\mathbb{R}$ ותהי $f \in \operatorname{Bil}(V, V)$ 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*

$$D_{p,q} = \operatorname{Diag}(\overbrace{1, \ldots, 1}^{p}, \overbrace{-1, \ldots, -1}^{q}, 0, \ldots, 0).$$

הוכחה: לפי משפט 11.13 יש בסיס $u_1, \ldots, u_n$ ל־V כך ש־$f(u_i, u_j) = 0$. נסמן $c_i := f(u_i, u_i)$. המטריצה של f לפי בסיס זה היא אלכסונית: $\operatorname{Diag}(c_1, \ldots, c_n)$. בלי הגבלת הכלליות, אחרי סידור מחדש של אברי הבסיס,

$$c_1, \ldots, c_p > 0, \quad c_{p+1}, \ldots, c_{p+q} < 0, \quad c_{p+q+1} = \cdots = c_n = 0.$$

נגדיר בסיס אחר $u'_1, \ldots, u'_n$ 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 $i \ne j$.
- אם $1 \le i \le p$ אז $f(u'_i, u'_i) = (\frac{1}{\sqrt{c_i}})^2 f(u_i, u_i) = 1$;
- אם $p < i \le p+q$ אז $f(u'_i, u'_i) = (\frac{1}{\sqrt{-c_i}})^2 f(u_i, u_i) = -1$;
- 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 $f \in \operatorname{Bil}(V, V)$ סימטרית. נניח ש־f מיוצגת לפי בסיס $u_1, \ldots, u_n$ על ידי מטריצה אלכסונית $D'_{p,q}$ בעלת p רכיבים חיוביים ו־q רכיבים שליליים, ולפי בסיס $v_1, \ldots, v_n$ על ידי מטריצה אלכסונית $D'_{s,t}$ בעלת s רכיבים חיוביים ו־t רכיבים שליליים. אז $p = s, q = t$.*

הוכחה: $p + q = \operatorname{rk} D'_{p,q} = \operatorname{rk} f = \operatorname{rk} D'_{s,t} = s + t$. 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 $u_1, \ldots, u_p, v_{s+1}, \ldots, v_n$ סדרה ב־V בת יותר מ־n אברים, ולכן תלויה לינארית מעל $\mathbb{R}$. לכן יש $a_1, \ldots, a_p, b_{s+1}, \ldots, b_n$, לא כולם 0, כך ש־

$$\sum_{i=1}^{p} a_i u_i + \sum_{i=s+1}^{n} b_i v_i = 0. \tag{4}$$

יתר על כן, מכאן נובע שיש $1 \leq i \leq p$ כך ש־$a_i \neq 0$. אכן, אחרת לפי (4), $\sum_{i=s+1}^{n} b_i v_i = 0$, ומכאן $b_{s+1} = \cdots = b_n = 0$, כי $v_{s+1}, \ldots, v_n$ בלתי תלויים לינארית, סתירה. כעת,

$$f(\sum_{i=1}^{p} a_i u_i, \sum_{i=1}^{p} a_i u_i) = \sum_{i=1}^{p} a_i^2 f(u_i, u_i) > 0,$$

כי $f(u_i, u_i) > 0$ לכל $1 \leq i \leq p$ ויש $1 \leq i \leq p$ כך ש־$a_i \neq 0$.

מצד שני, לפי (4),

$$f(\sum_{i=1}^{p} a_i u_i, \sum_{i=1}^{p} a_i u_i) = f(-\sum_{i=s+1}^{m} b_i v_i, -\sum_{i=s+1}^{n} b_i v_i) = \sum_{i=s+1}^{n} b_i^2 f(v_i, v_i) \leq 0,$$

כי $f(v_i, v_i) \leq 0$ לכל $s+1 \leq i \leq n$. 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 הדרגה ואותה החתימה.*

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) ⇐(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) ⇐(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, $v=\sum_{i=1}^n a_i u_i$. אז יש i כך ש־$a_i \neq 0$. מכאן $f(v,v)=\sum_{i,j} a_i a_j \delta_{ij}=\sum_{i=1}^n a_i^2 > 0$.

(3) ⇔(4): ברור, כי $P^tP=P^tI_nP$.

(3) ⇔(5): לפי הגדרת החתימה, $\sigma(A)=n$ אם ורק אם A חופפת ל־$D_{n,0}=I_n$.

(1) ⇔(6): לפי הגדרת מכפלה פנימית.

לבסוף, כדי להוכיח (2) ⇔(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) ⇐(7): D, D' חופפות, ולכן לפי משפט סילבסטר $\lambda_1,\dots,\lambda_n>0$.

(7) ⇐(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 סדרה 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 הכפולות 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 הכפולות של השורה השניה מהשורות שמתחתיה של A. פעולות אלה אינן משנות את

$\det A_1, \dots, \det A_n$ (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, $A = \mathrm{Diag}(a_{11}, \dots, a_{nn})$, באשר $a_{11}, \dots, a_{nn} \neq 0$. לכן p הוא מספר החיובים ו־q מספר השליליים בסדרה $a_{11}, \dots, a_{nn}$. אבל

$$k = 1, \dots, n \qquad a_{kk} > 0 \Leftrightarrow \frac{\det A_k}{\det A_{k-1}} > 0 \Leftrightarrow \det A_{k-1}, \det A_k$$ בעלי סימן זהה

(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) $f_{\alpha A}(X) = \det(XI - \alpha A) = \alpha^n \det(\alpha^{-1} XI - A) = \alpha^n f_A(\alpha^{-1} X)$. ∎

הגדרה 15.1: **שניונית** (quadric) X במרחב $\mathbb{R}^n$ היא קבוצה X של $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ $v =$ שהרכיבים $x_1, \dots, x_n$ 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־$x_1, \dots, x_n$. (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).

(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 $1 \le i < j \le n$, $\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{R}^n$ $u =$, $A = (a_{ij}) \in \mathrm{M}_n(\mathbb{R})$. 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 $\tilde{A} = \begin{array}{c} \\ 1 \\ n \end{array}\!\!\begin{array}{c} \begin{array}{cc} 1 & n \end{array} \\ \begin{pmatrix} a & u^t \\ u & A \end{pmatrix} \end{array} \in \mathrm{M}_{n+1}(\mathbb{R})$ תקרא **המטריצה המורחבת** של השניונית. גם היא סימטרית. את (2) אפשר לרשום בעזרתה כך:

$$.\begin{pmatrix} 1 \\ v \end{pmatrix}^t \tilde{A} \begin{pmatrix} 1 \\ v \end{pmatrix} = 0 \tag{3}$$ ∎

דיון 15.4: פעולות על שניוניות. מטרתנו להגיע למשוואה פשוטה יותר של X על ידי סדרה של הפעולות הבאות:

(I) *שינוי בסיס אורתונורמלי*(**שב"א**): 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־ $a + 2u^t P v' + (Pv')^t A (Pv') = 0$, כלומר

$$.a + 2(P^t u)^t v' + (v')^t (P^t A P) v' = 0 \tag{2'}$$

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* (**הזזה**): 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* (**הכפלה**): 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: *שב"א או הזזה 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 ב־ $c \neq 0$ *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.*

למה 15.6: *תהיינה* $X, X' \in \mathbb{R}^n$ *שניוניות ותהיינה* $\tilde{A}, \tilde{A}'$ *המטריצות המורחבות שלהן, בהתאמה. אז* X' *מתקבלת מ־* X *על ידי סדרה של שב"א או הזזות אם ורק אם* $\tilde{A}' = \tilde{P}^t\tilde{A}\tilde{P}$*, באשר* $\tilde{P} = \begin{pmatrix} 1 & 0 \\ w & P \end{pmatrix} \in \mathrm{M}_{n+1}(\mathbb{R})$ *מטריצת גושים, בה* $P \in \mathrm{M}_n(\mathbb{R})$ *אורתוגונלית ו־* $w \in \mathbb{R}^n$ *עמודה.*

הוכחה: "$\Leftarrow$": בדיון לעיל ראינו שהתנאי מתקיים אם X' מתקבלת מ־X על ידי פעולה אחת. אכן, אם הפעולה היא שב"א, נקח P להיות מטריצת המעבר ו־$w = 0$; אם הפעולה היא הזזה, ניקח $P = I_n$ ו־w וקטור ההזזה של הראשית.

נניח, באינדוקציה שהתנאי נכון אם שניונית אחת מתקבלת מאחרת על ידי k פעולות, ונניח ש־X' מתקבלת מ־X על ידי $k+1$ פעולות. תהי Y השניונית שמתקבלת מ־X על ידי k הפעולות הראשונות, ותהי $\tilde{B}$ המטריצה המורחבת שלה. אז X' מתקבלת מ־Y על ידי פעולה אחת. לפי הנחת האינדוקציה יש $\tilde{P}_1 = \begin{pmatrix} 1 & 0 \\ w_1 & P_1 \end{pmatrix}$ עם P_1 אורתוגונלית כך ש־$\tilde{B} = \tilde{P}_1^t \tilde{A} \tilde{P}_1$. לפי המקרה הקודם (של פעולה אחת) יש $\tilde{P}_2 = \begin{pmatrix} 1 & 0 \\ w_2 & P_2 \end{pmatrix}$ עם P_2 אורתוגונלית כך ש־$\tilde{A}' = \tilde{P}_2^t \tilde{B} \tilde{P}_2$. מכאן $\tilde{A}' = \tilde{P}_2^t \tilde{P}_1^t \tilde{A} \tilde{P}_1 \tilde{P}_2 = \tilde{P}^t \tilde{A} \tilde{P}$ באשר

$$\tilde{P} = \tilde{P}_1 \tilde{P}_2 = \begin{pmatrix} 1 & 0 \\ w_1 & P_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ w_2 & P_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ w_1 + P_1 w_2 & P_1 P_2 \end{pmatrix}$$

עם $P_1 P_2$ אורתוגונלית (כי היא מכפלה של אורתוגונליות).

"$\Rightarrow$": נעשה הזזת הראשית ל־w ואחר כך שינוי בסיס לפי P. לפי הדיון לעיל $\tilde{A}' = \tilde{P}^t \tilde{A} \tilde{P}$, באשר

$$\tilde{P} = \begin{pmatrix} 1 & 0 \\ w & I_n \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ w & P \end{pmatrix}. \quad \blacksquare$$

משפט 15.7: *תהי X שניונית. נסמן $\tilde{\sigma} = \sigma(\tilde{A}), \tilde{r} = \mathrm{rk}(\tilde{A}), \sigma = \sigma(A), r = \mathrm{rk}(A)$ באשר A המטריצה מצומצמת ו־$\tilde{A}$ המטריצה המורחבת של X.*

(א) *על ידי סדרה של שב"א והזזות 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:** יש P אורתוגונלית כך ש־$P^t A P = D = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$ אלכסונית ובלי הגבלת הכלליות מתקיים (5) עם $s = r$ ו־$\lambda_{r+1} = \cdots = \lambda_n = 0$. לכן, לפי דיון 15.4, בלי הגבלת הכלליות $A = D$.

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 $\bar{u} = \begin{pmatrix} b_{r+1} \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{R}^{n-r}$. $0 \neq \bar{u}$. יהי $\beta = ||\bar{u}|| > 0$. אז $||-\frac{1}{\beta}\bar{u}|| = 1$, לכן אפשר להשלים את $-\frac{1}{\beta}\bar{u}$ לבסיס אורתונורמלי של $\mathbb{R}^{n-r}$. לכן לפי משפט 13.13 יש $\bar{P} \in \mathrm{M}_{n-r}(\mathbb{R})$ אורתוגונלית כך

ש־$-\frac{1}{\beta}\bar{u}$ היא העמודה הראשונה שלה, כלומר, $\bar{u} \frac{1}{\beta}- = \bar{P}\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}$. מכאן $\bar{P}^t\bar{u} = \bar{P}^{-1}\bar{u} = \begin{pmatrix}-\beta\\0\\\vdots\\0\end{pmatrix}$. אז

$P = \begin{pmatrix}I_r & 0\\0 & \bar{P}\end{pmatrix}$ אורתוגונלית ומתקיים

$$P^tu = \begin{pmatrix}I_r & 0\\0 & \bar{P}^t\end{pmatrix}\begin{pmatrix}0\\\bar{u}\end{pmatrix} = \begin{pmatrix}0\\\bar{P}^t\bar{u}\end{pmatrix} = -\beta\mathbf{e}_{r+1}, \qquad P^tAP = A.$$

לכן, אחרי שב״א, 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־

$$a' = a + 2u^tw + w^tAw = a + 2(-\beta\mathbf{e}_{r+1})^t\frac{a}{2\beta}\mathbf{e}_{r+1} = a + 2(-\beta)\frac{a}{2\beta}\mathbf{e}_{r+1}^t\mathbf{e}_{r+1} = 0,$$

לכן, אחרי הזזה ב־w זה, בלי הגבלת הכלליות $a = 0$ (ושאר המרכיבים של המטריצה המורחבת של השניונית לא ישתנו).

(ב) המטריצה המצומצמת של (4) היא $\mathrm{Diag}(\lambda_1, \dots, \lambda_s, 0, \dots, 0)$. היא דומה ל־$A$. לכן: $s = r$ ו־$\lambda_1, \dots, \lambda_r, 0, \dots, 0$ הם הערכים העצמיים של A (על ריבוייהם).

(ג) המטריצה המורחבת של (4) היא, לפי המקרים השונים של Φ,

$$\begin{pmatrix}0 & 0 & \cdots & & & 0\\0 & \lambda_1 & & & & \\\vdots & & \ddots & & & \\ & & & \lambda_s & & \\ & & & & 0 & \\ & & & & & \ddots \\0 & & & & & 0\end{pmatrix} \quad \begin{pmatrix}-\alpha & 0 & \cdots & & & 0\\0 & \lambda_1 & & & & \\\vdots & & \ddots & & & \\ & & & \lambda_s & & \\ & & & & 0 & \\ & & & & & \ddots \\0 & & & & & 0\end{pmatrix} \quad \begin{pmatrix}-\alpha & 0 & \cdots & & & 0\\0 & \lambda_1 & & & & \\\vdots & & \ddots & & & \\ & & & \lambda_s & & \\ & & & & 0 & \\ & & & & & \ddots \\0 & & & & & 0\end{pmatrix} \quad \begin{pmatrix}0 & 0 & & -\beta & & \\0 & \lambda_1 & & & & \\ & & \ddots & & & \\ & & & \lambda_s & & \\-\beta & & & & 0 & \\ & & & & & \ddots \\ & & & & & 0\end{pmatrix}$$

$\Phi = 0$	$\Phi = \alpha > 0$	$\Phi = \alpha < 0$	$\Phi = 2\beta x_{r+1}$
$\tilde{r} = r, r = s$	$\tilde{r} = r + 1, r = s$	$\tilde{r} = r + 1, r = s$	$\tilde{r} = r + 2, r = s$
$\tilde{\sigma} = ?$ (לא חשוב)	$\tilde{\sigma} = \sigma - 1$	$\tilde{\sigma} = \sigma + 1$	$\tilde{\sigma} = ?$ (לא חשוב)
	אם $\sigma \neq 0$ אז	אם $\sigma \neq 0$ אז	
	$\lvert\tilde{\sigma}\rvert = \lvert\sigma\rvert - 1$	$\lvert\tilde{\sigma}\rvert = \lvert\sigma\rvert + 1$	

ומכאן, לפי מסקנה 15.5, (ג). (שתי השורות האחרונות נועדו בשביל המשפט הבא.)

(א) **יחידות:** לפי (ב) נקבעים באופן יחיד r, $\lambda_1, \dots, \lambda_r$. 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 $\Phi = 2\beta' x_{r+1}$, אז $\beta = \beta'$.

תהיינה $\tilde{A}, \tilde{A}'$ המטריצות המורחבות של שתי השניוניות האלה. כיון שאפשר לעבור על ידי סדרה של שב"א והזזות מאחת לשניה, לפי למה 15.6 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 $Aw = 0$ (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 $\beta \leq \beta'$. 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 $u' = P^t(u + Aw)$. כיון ש־P^t אורתוגונלית, ההעתקה $v \mapsto P^t v$ היא אוניטרית (מסקנה 13.6). לפי משפט 13.12 היא משמרת נורמה. לכן

$$(\beta')^2 = ||u'||^2 = ||u + Aw||^2 = || -\beta \mathbf{e}_{r+1}^t + (\lambda_1 w_1, \ldots, \lambda_r w_r, 0, \ldots, 0)^t ||^2 =$$
$$= ||(\lambda_1 w_1, \ldots, \lambda_r w_r, -\beta, 0, \ldots, 0)^t||^2 = \lambda_1^2 w_1^2 + \ldots + \lambda_r^2 w_r + \beta^2 \geq \beta^2$$

לכן $\beta \geq \beta'$. מכאן $\beta = \beta'$. ∎

אם נרשה גם הכפלה (הכפלה בסקלר שונה מאפס) הערכים העצמיים של המטריצה המצומצמת לא יישמרו, אך אפשר לקבל $\alpha = \pm 1$ או $\alpha = 0$, $\beta = 1$ (תיאור "מנורמל"):

משפט 15.8: *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 שב"א, הזזות והכפלות אפשר להעביר את המשוואה של X למשוואה מהצורה*

$$(4') \quad \lambda_1 x_1^2 + \cdots + \lambda_p x_p^2 \quad + \quad \lambda_{p+1} x_{p+1}^2 + \cdots + \lambda_{p+q} x_{p+q}^2 = \Phi,$$

$(5')$ באשר $\lambda_1, \ldots, \lambda_p > 0 > \lambda_{p+1}, \ldots, \lambda_{p+q}$, $\quad p \geq 1, p \geq q \geq 0$

ו־$\Phi = 0$ או $\Phi = \pm 1$ או $\Phi = 2x_{p+q+1}$.

יתר עלן, משוואה $(4')$ – תחת תנאי $(5')$ – היא יחידה במובן הבא:

(ב1) $p + q = r$, $p - q = |\sigma|$;

(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 $\sigma = 0$, $\tilde{r} = r+1$, אפשר להגיע לצורה $(4')$ עם $\Phi = 1$ וגם 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:

$$(-\lambda_{p+1}) x_{p+1}^2 + \cdots + (-\lambda_{2p}) x_{2p}^2 + (-\lambda_1) x_1^2 + \cdots + (-\lambda_p) x_p^2 = -1$$

ולאחר שינוי הסדר של אברי הבסיס היא מהצורה $(4')$. ∎

הוכחת המשפט: (א) על ידי הכפלת המשוואה ב־-1, אם צריך, אפשר להניח בלי הגבלת הכלליות ש־$\sigma(A) \geq 0$. (דבר זה אינו משנה את r ומחליף את σ ב־$|\sigma|$.) לכן לפי המשפט הקודם אפשר לעבור על ידי שב"א והזזות 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 $r = s$ ואם נסמן ב־p את מספר איברים החיובים 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 $p \geq 1$. נכפיל את (4) בסקלר מתאים $\gamma \neq 0$ ($\frac{1}{|\alpha|}$, אם $\Phi = \alpha \neq 0$ או $\frac{1}{\beta}$, אם $\Phi = 2\beta x_{p+q+1}$) ונקבל Φ כמבוקש.

(ב) שב"א, הזזות, והכפלות אינן משנות את $|\sigma|$, r 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. למעשה כבר עשינו זאת שם. נזכור ש־$|\tilde{\sigma}|, |\sigma|$ נשמרים. ∎

תרגיל 15.9: *בסימונים של משפט 15.8 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 $\sigma = \sigma(A) > 0$. *הוכח:*

(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: (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. כיון 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 $\det A, \det \tilde{A}'$ שוות סימן, ו־$\sigma(\tilde{A}') = \sigma - 1$ אם $\det A, \det \tilde{A}'$ מנוגדות סימן. מכאן המסקנה לפי משפט 15.8. ∎

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-** $\mathbb{R}^3$:

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	$\alpha^2x^2 = 2y$

- **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	0	1	1	$\alpha^2x^2 = 1$
קבוצה ריקה	2	2	1	1	$\alpha^2x^2 = -1$
פרבולה	1	3	1	1	$\alpha^2x^2 = 2y$

דוגמה 15.10 **:** 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 15.9, $\Phi=+1$, 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 15.11 **:** 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 15.9, $\Phi=-1$, 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 יריעות.

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 מסמנים 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:*

$$\begin{array}{c} \\ m_1 \\ m_2 \\ \vdots \\ m_r \end{array} \begin{array}{c} \begin{array}{cccc} p_1 & p_2 & \dots & p_s \end{array} \\ \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1s} \\ A_{21} & A_{22} & \dots & A_{2s} \\ \dots & \dots & \dots & \dots \\ A_{r1} & A_{r2} & \dots & A_{rs} \end{pmatrix} \end{array} \quad \begin{array}{c} \\ p_1 \\ p_2 \\ \vdots \\ p_s \end{array} \begin{array}{c} \begin{array}{cccc} n_1 & n_2 & \dots & n_t \end{array} \\ \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1t} \\ B_{21} & B_{22} & \dots & B_{2t} \\ \dots & \dots & \dots & \dots \\ B_{s1} & B_{s2} & \dots & B_{st} \end{pmatrix} \end{array} =$$

$$= \begin{array}{c} \\ m_1 \\ m_2 \\ \vdots \\ m_r \end{array} \begin{array}{c} \begin{array}{cccc} n_1 & n_2 & \dots & n_t \end{array} \\ \begin{pmatrix} C_{11} & C_{12} & \dots & C_{1t} \\ C_{21} & C_{22} & \dots & C_{2t} \\ \dots & \dots & \dots & \dots \\ C_{r1} & C_{r2} & \dots & C_{rt} \end{pmatrix} \end{array}$$

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:

$$\begin{array}{c} \\ m_1 \\ m_2 \end{array} \begin{array}{c} \begin{array}{cc} p_1 & p_2 \end{array} \\ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \end{array} \quad \begin{array}{c} \\ p_1 \\ p_2 \end{array} \begin{array}{c} \begin{array}{cc} n_1 & n_2 \end{array} \\ \begin{pmatrix} P & Q \\ R & S \end{pmatrix} \end{array} \quad = \quad \begin{array}{c} \\ m_1 \\ m_2 \end{array} \begin{array}{c} \begin{array}{cc} n_1 & n_2 \end{array} \\ \begin{pmatrix} AP+BR & AQ+BS \\ CP+DR & CQ+DS \end{pmatrix} \end{array}$$

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

$$\begin{gathered}(C)_{i1}(P)_{1j}+(C)_{i2}(P)_{2j}+\cdots+(C)_{i\,p_1}(P)_{p_1\,j}+\\ +(D)_{i1}(R)_{1j}+(D)_{i2}(R)_{2j}+\cdots+(D)_{i\,p_2}(R)_{p_2\,j}=(CP)_{ij}+(DR)_{ij}=(CP+DR)_{ij}\end{gathered}$$

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

$$\begin{aligned}&(A_{21})_{i1}(B_{11})_{1j}+(A_{21})_{i2}(B_{11})_{2j}+\cdots+(A_{21})_{i\,p_1}(B_{11})_{p_1\,j}\ +\ldots+\\ &(A_{2s})_{i1}(B_{s1})_{1j}+(A_{2s})_{i2}(B_{s1})_{2j}+\cdots+(A_{2s})_{i\,p_s}(B_{s1})_{p_s\,j}=\\ &\sum_{k=1}^{p_1}(A_{21})_{ik}(B_{11})_{kj}+\cdots+\sum_{k=1}^{p_s}(A_{2s})_{ik}(B_{s1})_{kj}\end{aligned}$$

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*

$$AB = (\, A_1, \quad \dots \quad , A_p \,)\begin{pmatrix}B_1\\ \vdots\\ B_p\end{pmatrix} = (\, \textstyle\sum_{k=1}^{p} A_kB_k \,) \in \mathrm{M}_{m\times n}(F).$$

הוכחה: הפעל את המשפט עם $r=1, s=p, t=1, p_1=\cdots=p_s=1$. ∎

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*

$$A\begin{pmatrix}c_1\\ \vdots\\ c_n\end{pmatrix} = c_1v_1+\cdots+c_nv_n.$$

הוכחה: לפי המסקנה, אגף ימין שווה למטריצה $v_1(c_1)+\cdots+v_n(c_n)$. 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*

$$(P)_{i1}A_1 + (P)_{i_2}A_2 + \cdots + (P)_{im}A_m = \sum_{k=1}^{m}(P)_{ik}A_k, \qquad i=1,2,\dots,m$$

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: *הוכח כי מטריצת הגושים* $\begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}\end{array} = M$ *הפיכה אם ורק אם* A, D *הפיכות ואם זה קורה אז* $\begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{pmatrix} A^{-1} & -A^{-1}BD^{-1} \\ 0 & D^{-1} \end{pmatrix}\end{array} = N$ *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

$$AQ+BS=0,\quad DR=0,\quad DS=I_n,\quad AP+BR=I_m.$$

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: *הוכח כי מטריצת הגושים* $M=\begin{array}{cc} & \begin{array}{cc} n & m\end{array}\\ \begin{array}{c} m\\ n\end{array} & \begin{pmatrix}A&B\\C&0\end{pmatrix}\end{array}$ *הפיכה אם ורק אם* B,C *הפיכות ואם זה קורה אז* $N=\begin{array}{cc} & \begin{array}{cc} m & n\end{array}\\ \begin{array}{c} n\\ m\end{array} & \begin{pmatrix}0&C^{-1}\\B^{-1}&-B^{-1}AC^{-1}\end{pmatrix}\end{array}$ *ההופכי שלה.*

בפרט, $M_0=\begin{array}{cc} & \begin{array}{cc} n & m\end{array}\\ \begin{array}{c} m\\ n\end{array} & \begin{pmatrix}0&B\\C&0\end{pmatrix}\end{array}$ *הפיכה אם ורק אם* B,C *הפיכות ואז* $M_0^{-1}=\begin{array}{cc} & \begin{array}{cc} m & n\end{array}\\ \begin{array}{c} n\\ m\end{array} & \begin{pmatrix}0&C^{-1}\\B^{-1}&0\end{pmatrix}\end{array}$.

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: *הוכח כי* $\begin{array}{cc} & \begin{array}{cc} m & n\end{array}\\ \begin{array}{c} m\\ n\end{array} & \begin{pmatrix}A&0\\0&D\end{pmatrix}\end{array}$ *דומה למטריצה* $\begin{array}{cc} & \begin{array}{cc} n & m\end{array}\\ \begin{array}{c} n\\ m\end{array} & \begin{pmatrix}D&0\\0&A\end{pmatrix}\end{array}$.

הוכחה: לפי תרגיל 16.6, $\begin{array}{cc} & \begin{array}{cc} n & m\end{array}\\ \begin{array}{c} m\\ n\end{array} & \begin{pmatrix}0&I_m\\I_n&0\end{pmatrix}\end{array}$ הפיכה ו־ $\begin{array}{cc} & \begin{array}{cc} m & n\end{array}\\ \begin{array}{c} n\\ m\end{array} & \begin{pmatrix}0&I_n\\I_m&0\end{pmatrix}\end{array}$ ההופכי שלה. לפי המשפט

$$\begin{array}{cc} & \begin{array}{cc} m & n\end{array}\\ \begin{array}{c} n\\ m\end{array} & \begin{pmatrix}0&I_n\\I_m&0\end{pmatrix}\end{array}\begin{array}{cc} & \begin{array}{cc} m & n\end{array}\\ \begin{array}{c} m\\ n\end{array} & \begin{pmatrix}A&0\\0&D\end{pmatrix}\end{array}\begin{array}{cc} & \begin{array}{cc} n & m\end{array}\\ \begin{array}{c} m\\ n\end{array} & \begin{pmatrix}0&I_m\\I_n&0\end{pmatrix}\end{array}=\begin{array}{cc} & \begin{array}{cc} n & m\end{array}\\ \begin{array}{c} n\\ m\end{array} & \begin{pmatrix}D&0\\0&A\end{pmatrix}\end{array}$$

∎

תרגיל 16.8: *תהי* $M = \begin{array}{cc} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \end{array}$ *מטריצת גושים. הוכח:* $|M| = |A| \cdot |D|$.

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 מטריצת מעבר.

יהי V מרחב וקטורי מעל שדה F ויהי $\mathcal{B}=(v_1,\dots,v_n)$ בסיס שלו.

הגדרה 17.1: תהי $\mathcal{B}'=(v'_1,\dots,v'_n)$ סדרה של אברי V. אז $P=\big([v'_1]_{\mathcal{B}},\dots,[v'_n]_{\mathcal{B}}\big)\in \mathrm{M}_n(F)$ 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 $\mathcal{B}'=(v'_1,\dots,v'_n)$ של אברי V כך ש־P היא מטריצת המעבר מ־$\mathcal{B}$ ל־$\mathcal{B}'$.*

פתרון: ההעתקה $[v]_{\mathcal{B}}\mapsto v$ 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 $\mathcal{B}'=(v'_1,\dots,v'_n)$, 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 $\Leftrightarrow \dim(Sp([v'_1]_{\mathcal{B}},\dots,[v'_n]_{\mathcal{B}}))=\dim(C(P))=\mathrm{rk}\,P=n$

$\Leftrightarrow [v'_1]_{\mathcal{B}},\dots,[v'_n]_{\mathcal{B}}$ 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 $[v'_1]_{\mathcal{B}},\dots,[v'_n]_{\mathcal{B}}$ בסיס של $F^n \Leftrightarrow v'_1,\dots,v'_n$ בסיס של V. ∎

טענה 17.5: *יהי $\mathcal{B}'=(v'_1,\dots,v'_n)$ 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

$$P[v'_j]_{\mathcal{B}'}=P\mathbf{e}_j=\text{העמודה ה־}j\text{ של }P=[v'_j]_{\mathcal{B}}.$$

במקרה הכללי יש $a_1,\dots,a_n\in F$ כך ש־$v=\sum_{j=1}^n a_jv'_j$. אז, לפי המקרה הקודם,

$$P[\sum_{j=1}^n a_jv'_j]_{\mathcal{B}'}=P\big(\sum_{j=1}^n a_j[v'_j]_{\mathcal{B}'}\big)=\sum_{j=1}^n a_jP[v'_j]_{\mathcal{B}'}=\sum_{j=1}^n a_j[v'_j]_{\mathcal{B}}=[\sum_{j=1}^n a_jv'_j]_{\mathcal{B}},$$

כלומר, $P[v]_{\mathcal{B}'}=[v]_{\mathcal{B}}$. ∎

18. נספח: מטריצות אלמנטריות

נספח זה מכיל חומר מאלגברה לינארית 1 אודות מטריצת 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 $E(A) = \sum_{i=1}^m a_i \mathbf{e}_i A \mathbf{e}_i$, 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: $\mathcal{E}'(B) = \big(\mathcal{E}(B^t)\big)^t = (EB^t)^t = BE^t$. ∎

משפטים למבחן

במבחן הסופי של שנת תשע"ז (במועד א וגם במועד ב) תהיינה שתי שאלות של הוכחה של חומר שהיה בהרצאה ("משפטים"). השאלות תהיינה מתוך הרשימה הבאה:

- 2.3 [חילוק עם שארית]
- 2.4 [$f(\alpha)=0$ אם ורק אם $(X-\alpha)|f$.]
- 3.4 [כל אידאל ב־$F[X]$ הוא ראשי]
- 3.6 +3.7 [איבר הינו אי פריק אם ורק אם הוא ראשוני.]
- 3.8 [פירוק לראשוניים]
- 4.10 [אי תלות ו"ע השייכים לע"ע שונים]
- 6.9 [Cayley-Hamilton]
- 7.5 [$f_A|m_A^n$]
- 7.12 [תרגיל: $\deg m_A = n$]
- 8.2 [ריבוי אלגברי $\leq$ ריבוי גיאומטרי]
- 8.3 [לכסון לפי ריבוי]
- 9.13+9.12 [פירוק פרימרי, $n=2$]
- 10.6 [$\mathrm{rank}(J)$]
- 10.12 [קיום בסיס שהמטריצה לפיו היא מטריצת ז'ורדן נילפוטנטית]
- 11.10 [צורת יעקובסון]
- 12.11 [תהליך גרם שמידט]
- 12.18 [אי שוויון קושי שוורץ]
- 12.22 [בנית מרחב משלים]
- 13.1 [הגדרת $T*$]
- 13.12 [שקילויות של העתקה אוניטרית]
- 13.18 + 13.17 [$T^*(v)=\bar{\lambda}v$ ו"ע ניצבים זה ללזה.]
- 13.19 [משפט הלכסון האוניטרי להעתקות נורמליות]
- 13.22 [ע"ע של מטריצה צמודה לעצמה ואוניטרית]
- 13.25 [מטריצה של העתקה אורתוגונלית]
- 14.13 [לכסון תבנית בילינארית]
- 14.19 [משפט סילבסטר]
- 14.24 [חתימת התבנית לפי הדטרמיננטות]

משפטים נוספים, מהפרקים שעוד לא נלמדו, יכולים להתווסף לרשימה בהמשך.