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

אלגברה לינארית להנדסה

מערכי שעור

תשס"ב

נערך על ידי

דן הרן

תוכן העניינים

0. וקטורים במרחב ובמישור 1
1. שדות 4
2. משוואות לינאריות 6
3. מטריצות 17
4. מרחבים וקטוריים 29
5. דטרמיננטות 45
6. העתקות לינאריות 54
7. וקטורים עצמיים וערכים עצמיים 69
8. פולינום אופייני, פולינום מזערי וצורת ז'ורדן 73
8א. מציאת צורת ז'ורדן למטריצות מסדר 80
9. מרחבי מכפלה פנימית 82
10. שקילות קוניונקטיבית וחפיפה 90
11. לכסון מטריצה סימטרית 93
דוגמה של מבחן 100

ספרות מומלצת

הקורס יתנהל לפי חוברת שהכנתי לקורס, אותה ניתן למצוא (בפורמט pdf) באתר

http://www.math.tau.ac.il/~haran/lin-alg-eng

החוברת מבוססת פחות או יותר על הספר הבא:

• Hans Schneider, George Phillip Barker, *Matrices and Linear Algebra*, Dover Publications Inc., 2nd edition, 1989

בספריה נמצאים עותקים רבים של ספר זה, חלקם ממהדורה ישנה יותר: אין הבדל משמעותי בין המהדורות. אין חובה לרכוש את הספר הזה או להחזיק עותק שלו!

• *אלגברה לינארית* של האוניברסיטה הפתוחה.

חוברות אלו הנן ברמה קצת גבוהה יותר ממה שנדרש לנו, (לא בהרבה!) אך הנן קריאות מאד ומתאימות ללימוד עצמי.

• 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 במהדורות אחדות. לצערי, במהדורות מסויימות יש פתרונות שגויים לתרגילים.

הסילבוס שלהלן הוא הסילבוס הרשמי של הפקולטה להנדסה. הוא קצת יומרני מדי, וחלק מהנושאים לא יכוסה בקורס. הנושאים בפרק 0 (גיאומטריה) לא יינתנו בנפרד, אלא כתת נושאים בתוך הסעיפים של יתר הפרקים.

סילבוס מפורט לקורס אלגברה ליניארית

0509.1824

מתכונת הקורס: 5 שעות הרצאה + 2 שעות תרגול (סה"כ 65 שעות הרצאה). סמסטר אחד.

ספר הלימוד:

1. Introduction to Linear Algebra, S. Lang, Springer, 2nd Ed.
2. H. Schneider and G.P. Barker: Matrices and Linear Algebra, Dover

1. גיאומטריה.

שיעורים (6־10 שעות הרצאה): *הגדרה גיאומטרית והצגה קרטזית של וקטורים במישור ובמרחב.* הצגה גרפית, אורך וכוון, וקטורים קשורים וחופשיים, סקלר וכפל בסקלר, חיבור וחיסור גיאומטרי על פי חוק המקבילית ובקואורדינטות, תנאים גיאומטריים ואלגבריים לקו־ליניאריות ולקו־מישוריות, תלות של וקטורים במישור ובמרחב, וקטורי יחידה על הצירים והצגת כל וקטור בעזרתם, היטלים על ישר ומישור, פירוק וקטור לרכיבים בכוונים בלתי תלויים. הגדרה גיאומטרית וקרטזית של מכפלה סקלרית ומכפלה וקטורית, זהויות וקטוריות, חוקי פילוג וקיבוץ (דיסטריבוטיבי ואסוציאטיבי), חישוב המכפלות הנ"ל לוקטורי היחידה על הצירים, תנאי לניצבות, מכפלה משולשת, שטח מקבילית ונפח מקבילון. הצגה וקטורית וקרטזית של ישר ומישור באמצעות המכפלות הנ"ל, נורמל למישור, נוסחה למרחק בין נקודה לישר ולמישור, ישרים מצטלבים במרחב והמרחק ביניהם.

2. אלגברה לינארית.

שיעורים 1, 2: *האלגברה של המטריצות.* שדות מספרים (רציונאליים, ממשיים, מרוכבים, דוגמאות לשדות סופיים), הגדרת מטריצה מעל שדה, האלגברה של המטריצות. מטריצות ריבועיות, הופכית, מחלקי אפס (פרק 1).

שיעורים 3, 4, 5: *משוואות ליניאריות.* מערכות שקולות של משוואות ליניאריות, פעולות על שורות של מטריצות, צורת row echelon form. מערכת הומוגנית, תנאי קונסיסטנטיות, פתרון כללי, היפוך של מטריצות לא סינגולריות (פרק 2).

שיעורים 6, 7, 8, 9: *מרחבים וקטוריים.* הוקטורים ומרחבים וקטוריים, תת־מרחב וצירופים ליניאריים, תלות ואי־תלות ליניארית. בסיסים, בסיסים והצגות, מרחב השורות של מטריצה, שקילות בעמודות, שקילות שורות־עמודות, יחס שקילות וצורות קנוניות של מטריצות (פרק 3).

שיעורים 12 ,11 ,10: *דטרמיננטים.* נפח ב־$\mathbb{R}^3$ וב־$\mathbb{R}^n$, תמורות ומטריצות תמורה, קיום ויחידות של פונקצית הדטרמיננט, חישוב של דטרמיננט ודטרמיננט של המטריצה המוחלפת. קו־פקטורים, מינורים, ו־adjoints. הדטרמיננט ודרגת מטריצה (פרק 4).

שיעור 14 ,13: *העתקות ליניאריות.* הגדרות, הצגה של העתקה ליניארית, הצגות תחת החלפת בסיס. (פרק 5) מטריצת סיבוב במישור ובמרחב כאופרטורים.

שיעורים 15–18: *ערכים ווקטורים עצמיים.* הגדרות, הקשר בין ע"ע ומינורים, דמיון, ריבוי אלגברי וגיאומטרי, תת־מרחבים אינווריאנטיים של אופרטור. הצורה הקנונית של ז'ורדן (ללא הוכחה), פונקציה של מטריצה, יישום למשוואות דיפרנציאליות לינאריות עם מקדמים קבועים (יציבות) (פרק 6, 8.2).

שיעורים 19–23: *מרחבי מכפלה פנימית.* מכפלות פנימיות, הצגה של מכפלה פנימית, בסיסים אורתוגונליים, שקילות אוניטרית ומטריצות הרמיטיות, קונגרואנציה ושקילות קוניונקטיבית.

שיעורים 24–28: ליכסון מטריצה סימטרית ע"י דמיון אורתוגונאלי, מציאת ציר הסיבוב, מיון שניוניות במישור ובמרחב וצורות קנוניות (7.6 - 7.1).

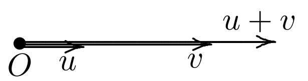
0. וקטורים במרחב ובמישור

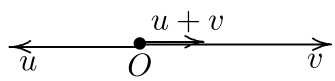
הגדרה 1.0 (זמנית): **וֶקטור** הוא חץ במרחב (ובפרט במישור, שהינו חלק מן המרחב) מהראשית O. חץ שקצהו בנקודה A יסומן $\overrightarrow{OA}$. לפעמים נשתמש גם באותיות לטיניות קטנות לסימון וקטורים, כגון $v, w, u, \ldots$. **וקטור האפס**, $\overrightarrow{OO}$ יסומן ב־0.

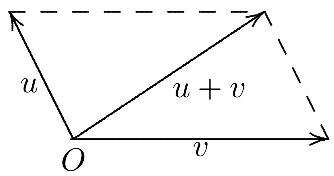
סקלר הוא מספר ממשי (איבר ב־$\mathbb{R}$).

פעולות על הוקטורים:

(א) חיבור: אם u, v שני וקטורים, אז $u+v$ הוא אלכסון המקבילית הנקבעת על ידי u, v:

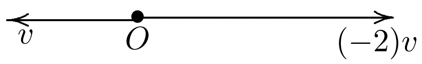


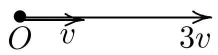




(יתכן שהמקבילית מנוונת ־ כמו בציורים משמאל ־ הזוית בין u ל־v יכולה להיות 0 או 180 מעלות.)

(ב) כפל וקטור בסקלר. אם v וקטור ו־a סקלר, אז av הוא הוקטור בעל אותו הכיוון כמו v, אך ארכו גדול פי a. (אם $a < 0$, פירוש המשפט הקודם הוא ש־av בעל אורך פי $|a|$ מארכו של v, וכיוונו מנוגד לכיוון של v.) בפרט $0u = 0$.





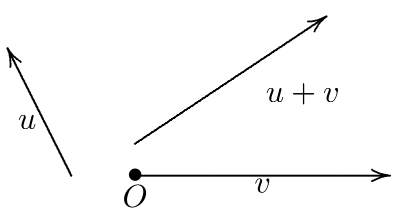
סימון: $-u = (-1)u$.

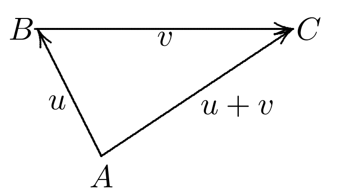
תרגיל: אם $u + v = 0$ אז $u = -v$.

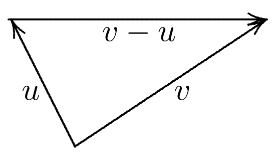
הגדרה 2.0 (זמנית): **וקטור** הוא חץ במרחב או במישור. חץ מנקודה A (נקודת **המוצא**) לנקודה B (נקודת **הקצה**) יסומן $\overrightarrow{AB}$. שני ווקטורים הנם **שווים**, אם יש להם אותו אורך ואותו כיוון (גם אם אינם יוצאים מאותה נקודת המוצא).

על כן וקטור ניתן לייצג על ידי חצים רבים. ביניהם יש אחד ויחיד שמוצאו בראשית.

הפעולות על וקטורים מוגדרות כך: לוקחים מייצגים שהם חצים מן הראשית ומשתמשים בהגדרה הקודמת לגביהם. (כמובן, את חץ התוצאה של הפעולה אפשר להזיז במרחב או במישור ־ מבלי לשנות את כיוונו או את ארכו ־ והוא עדיין ייצג את תוצאת הפעולה.) אך ניתן לבטא זאת גם על פי *כלל המשולש*.







החיבור לפי כלל המשולש: הזז את v, אם צריך, כך שנקודת המוצא שלו היא נקודת הקצה B של u. תהי A נקודת המוצא של u ותהי C נקודת הקצה של v. אז $u + v = \overrightarrow{AC}$.

הגדרה 3.0 (זמנית): וקטור הוא נקודה במרחב או במישור.

הגדרה זו מתקשרת להגדרה 1.0: אם $v = \overrightarrow{OA}$ שמוצאו בראשית, אז הוא קובע את הנקודה A. להיפך, כל נקודה A במרחב או במישור קובעת את החץ $\overrightarrow{OA}$ שמוצאו בראשית.

חשוב להדגיש: אם הוקטור נתון כחץ $\overrightarrow{AB}$ שמוצאו לא מן הראשית, במובן של הגדרה 2.0, אז B איננה הנקודה המתאימה לו כוקטור במובן של הגדרה 3.0.

כל נקודה A במרחב אפשר לייצג על ידי שלֹשה של מספרים ממשיים (x, y, z), הקואורדינטות של A. באופן דומה, כל נקודה A במישור אפשר לייצג על ידי זוג של מספרים ממשיים (x, y), הקואורדינטות של A.

מכאן:

הגדרה 4 (זמנית): וקטור הוא שלשה או זוג של מספרים ממשיים.

הפעולות על הוקטורים לפי הגדרה 4:

(א) $(x, y, z) + (x', y', z') = (x + x', y + y', z + z')$

(ב) $a(x, y, z) = (ax, ay, az)$

לפיכך וקטור האפס הוא $(0, 0, 0)$.

משפט: *יהי V המרחב או המישור. אז מתקיימים החוקים הבאים:*

(ח1) *אסוציאטיביות החיבור:* $(u + v) + w = u + (v + w)$, *לכל* $u, v, w \in V$.

(ח2) *תכונת איבר האפס:* $v + 0 = v = 0 + v$, *לכל* $v \in V$.

(ח3) *תכונת איבר נגדי:* $v + (-v) = 0 = (-v) + v$.

(ח4) *קומוטטיביות החיבור:* $u + v = v + u$, *לכל* $u, v \in V$.

(כ1) *חוק הפילוג:* $a(u + v) = au + av$, *לכל* $u, v \in V$ *ולכל* $a \in \mathbb{R}$.

(כ2) *חוק הפילוג:* $(a + b)v = av + bv$, *לכל* $v \in V$ *ולכל* $a, b \in \mathbb{R}$.

(כ3) *אסוציאטיביות הכפל בסקלר:* $(ab)v = a(bv)$, *לכל* $v \in V$ *ולכל* $a, b \in \mathbb{R}$.

(כ4) $1v = v$ *לכל* $v \in V$.

תרגיל: נניח $u = \overrightarrow{OA}$, $v = \overrightarrow{OB}$.

(א) מצא את $\overrightarrow{OD}$, באשר D נקודת האמצע של הקטע AB.

(ב) (הכללה:) מצא את $\overrightarrow{OD}$, באשר D מחלקת את הקטע AB ביחס $\alpha : \beta$.

פתרון: (א) $\overrightarrow{AD}$ באותו כיוון כמו $\overrightarrow{AB} = v - u$ וארכו מחצית מזה של $\overrightarrow{AD}$, לכן $\overrightarrow{AD} = \frac{1}{2}(v - u)$. מכאן $\overrightarrow{OD} = \overrightarrow{OA} + \overrightarrow{AD} = u + \frac{1}{2}(v - u) = \frac{1}{2}(u + v)$.

(ב) דומה לחלק הקודם, רק שעכשיו היחס בין הארכים של AD ושל AB הוא $\frac{\alpha}{\alpha+\beta}$. לכן $\overrightarrow{AD} = \frac{\alpha}{\alpha+\beta}(v-u)$ ומכאן

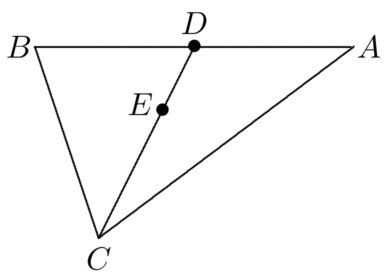
$$\overrightarrow{OD} = u + \frac{\alpha}{\alpha+\beta}(v-u) = \frac{\alpha+\beta}{\alpha+\beta}u + \frac{\alpha}{\alpha+\beta}(v-u) = \frac{\beta}{\alpha+\beta}u + \frac{\alpha}{\alpha+\beta}v.$$

תרגיל: הוכח שהתיכונים בכל משולש נפגשים באותה הנקודה, אשר מחלקת אותם ביחס $1:2$.

פתרון: יהיו A, B, C קדקדי המשולש. נראה שהנקודה שמחלקת כל תיכון ביחס $1:2$ היא אותה הנקודה. לכן היא נקודת המפגש המשותפת של כולם ומחלקת אותם ביחס המבוקש.

יהיו $u = \overrightarrow{OA}, v = \overrightarrow{OB}, w = \overrightarrow{OC}$. תהי D נקודת האמצע של הצלע AB. לפי התרגיל הקודם $\overrightarrow{OD} = \frac{1}{2}(u+v)$. כעת CD הוא תיכון במשולש. תהי E הנקודה המחלקת אותו ביחס $1:2$. לפי התרגיל הקודם,

$$\overrightarrow{OE} = \frac{1}{1+2}\overrightarrow{OC} + \frac{2}{1+2}\overrightarrow{OD} = \frac{1}{3}w + \frac{2}{3}\cdot\frac{1}{2}(u+v) = \frac{1}{3}(u+v+w).$$



באופן דומה ניתן לחשב את הנקודות שמחלקות ביחס $1:2$ את שני התיכונים האחרים. אך כדי לעשות זאת, די להחליף בחישוב לעיל את האותיות A, B, C ביניהן ואת האותיות u, v, w ביניהן. כיוון שהתוצאה $\frac{1}{3}(u+v+w)$ אינה תלויה בחילוף זה, שלוש נקודות החלוקה זהות. ∎

1. שדות

הגדרה 1.1 (זמנית): **שדה** F הוא קבוצה חלקית של המספרים המרוכבים $\mathbb{C}$ שמכילה את $0,1$ וסגורה תחת 4 פעולות החשבון, כלומר, אם $a,b \in F$ אז גם $a \pm b, ab \in F$ וגם $\frac{a}{b} \in F$, בתנאי ש־$b \neq 0$.

דוגמאות 1.2: דוגמאות לשדות. שדה המספרים המרוכבים $\mathbb{C}$, שדה המספרים הממשיים $\mathbb{R}$, שדה המספרים הרציונליים $\mathbb{Q}$.

טענה 1.3: *הקבוצה הבאה היא שדה:*

$$F = \{a + b\sqrt{2} |\ a, b \in \mathbb{Q}\}.$$

הוכחה: נשתמש בעובדה הידועה ש־$\sqrt{2} \notin \mathbb{Q}$.

כעת, $0,1 \in F$ ולמעשה, $a \in F$ לכל $a \in \mathbb{Q}$, כי $a = a + 0\sqrt{2} \in F$.

יהיו $a + b\sqrt{2}, c + d\sqrt{2} \in F$. אז

$$(a + b\sqrt{2}) \pm (c + d\sqrt{2}) = (a \pm c) + (b \pm d)\sqrt{2} \in F,$$
$$(a + b\sqrt{2})(c + d\sqrt{2}) = (ac + 2bd) + (ad + cb)\sqrt{2} \in F.$$

אם $c + d\sqrt{2} \neq 0$ אז גם $c - d\sqrt{2} \neq 0$. (אכן, אם $c - d\sqrt{2} = 0$, אז או ש־$d = 0$ ואז גם $c = 0$ ולכן $c + d\sqrt{2} = 0$, סתירה, או ש־$d \neq 0$ ואז $\sqrt{2} = \frac{c}{d} \in \mathbb{Q}$, סתירה.) לכן

$$\frac{a + b\sqrt{2}}{c + d\sqrt{2}} = \frac{(a + b\sqrt{2})(c - d\sqrt{2})}{(c + d\sqrt{2})(c - d\sqrt{2})} = \frac{(ac - 2bd) + (-ad + cb)\sqrt{2}}{c^2 - 2d^2} =$$
$$= \frac{ac - 2bd}{c^2 - 2d^2} + \frac{-ad + cb}{c^2 - 2d^2}\sqrt{2} \in F. \quad \blacksquare$$

קל לראות שאם F תת קבוצה של $\mathbb{C}$ הסגורה תחת 4 פעולות החשבון אז $\mathbb{Q} \subseteq F$.

כעת נביא את ההגדרה הכללית של שדה. אך לפני כן נזכיר **שפעולה** על קבוצה F היא איזשהו כלל, שמתאים לכל זוג של איברי F איבר של F. לדוגמה, אם F היא קבוצת המספרים הממשיים הגדולים מ־10, הכלל שמתאים לכל זוג (x,y) של מספרים כאלה את הממוצע שלהם הוא פעולה. דוגמה נוספת: אם F היא קבוצת כל המספרים הטבעיים, הכלל שמתאים לכל זוג (m,n) את המספר m^n הוא פעולה. גם הכלל שמתאים לכל זוג (m,n) את המספר n^m הוא פעולה, שהינה שונה מהפעולה הקודמת.

הגדרה 1.4: **שדה** הנו קבוצה F יחד עם שתי פעולות עליו, חיבור (+) וכפל (·), המקיימות את הכללים הבאים:

(1א) כלל החילוף של החיבור: $a + b = b + a$,

(1ב) כלל הצירוף לחיבור: $(a + b) + c = a + (b + c)$

(1ג) קיום איבר אפס: קיים איבר יחיד ב־F, המסומן ב־0, שמקיים $a + 0 = a$ לכל $a \in F$,

(1ד) קיום איבר נגדי: לכל $a \in F$ קיים איבר יחיד $-a \in F$ כך ש $a + (-a) = 0$,

(2א) כלל החילוף של הכפל: $a \cdot b = b \cdot a$,

(2ב) כלל הצירוף לכפל: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$,

(2ג) קיום איבר יחידה: קיים איבר יחיד ב־F, המסומן 1, כך ש־$0 \neq 1$ ו־$a \cdot 1 = a$ לכל $a \in F$.

(2ד) קיום איבר הופכי: לכל איבר $a \in F$ השונה מאיבר האפס 0, קיים איבר יחיד ב־F, המסומן a^{-1}, המקיים $a \cdot a^{-1} = 1$.

(3) כלל הפילוג: $a \cdot (b + c) = a \cdot b + a \cdot c$.

דוגמאות 1.5:

(א) כל שדה כפי שהגדרנו אותו בתחילת הפרק הוא גם שדה לפי ההגדרה החדשה.

(ב) תהי F קבוצה בת שני איברים a, b ונגדיר עליה פעולות $+, \cdot$ בעזרת שתי הטבלאות הבאות

$\cdot$	a	b
a	a	a
b	a	b

$+$	a	b
a	a	b
b	b	a

בדיקה מיגעת תראה שכל התנאים של שדה מתקיימים. במקרה זה איבר האפס הוא a ואיבר היחידה הוא b.

שים לב ש־$1 + 1 = 0$, שלא כמו בדוגמאות (א)!

(ג) יהי p מספר ראשוני (למשל $p = 5$) ותהי $F = \{0, 1, 2, \ldots, p-1\}$ קבוצה בת p איברים. הסכום $a + b$ של זוג (a, b) של איברי F (לפי הכללים של מספרים שלמים) אינו בהכרח איבר של F. אך השארית של $a + b$ לאחר חילוק ב־p היא איבר של F; נגדיר אותה כסכום של a, b. באופן דומה נגדיר את הכפל על F. ניתן להוכיח ש־F הוא שדה. שים לב ש־$\overbrace{1 + 1 + \cdots + 1}^{p\times} = 0$.

(ד) $F = \{a, b, c, d\}$ עם הפעולות הבאות

$\cdot$	a	b	c	d
a	a	a	a	a
b	a	b	c	d
c	a	c	d	b
d	a	d	b	c

$+$	a	b	c	d
a	a	b	c	d
b	b	a	d	c
c	c	d	a	b
d	d	c	b	a

הוא שדה (a איבר האפס, b איבר היחידה).

2. משוואות לינאריות

בפרק זה יהי F שדה כלשהו.

עבור מספר טבעי n נסמן ב־F^n את אוסף כל העמודות $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ בעלות n קואורדינטות $x_1, x_2, \ldots, x_n$ ב־F. לכתיבה בעמודה תהיה חשיבות מאוחר יותר; כרגע כתיב זה די מפריע לסידור הטקסט, לכן לפעמים נכתוב $(x_1, x_2, \ldots, x_n)^t$ במקום $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$.

הגדרה 2.1: **משוואה לינארית** (מעל F) הוא ביטוי מהצורה

$$a_1X_1 + a_2X_2 + \cdots + a_nX_n = b, \tag{1}$$

באשר $a_1, a_2, \ldots, a_n, b \in F$. הם נקראים **המקדמים** של (1), כאשר b נקרא **המקדם החפשי**. בביטוי זה $X_1, X_2, \ldots, X_n$ הם סמלים סתמיים, שנקראים **נעלמים** או **משתנים**. המשוואה נקראת **הומוגנית** אם $b = 0$.
עמודה $(x_1, x_2, \ldots, x_n)^t \in F^n$ היא **פתרון** של (1) אם מתקיים

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b.$$

קבוצה הפתרון של משוואה (1) היא אוסף כל הפתרונות שלה.

דוגמה 2.2: קבוצת הפתרון של $X_1 - 2X_2 + X_4 + 8X_5 = 2$ היא

$$P = \{(2 + 2x_2 - x_4 - 8x_5, x_2, x_3, x_4, x_5)^t |\ x_2, x_3, x_4, x_5 \in F\}.$$

הגדרה 2.3: **מערכת של** m **משוואה לינאריות ב־**n **נעלמים** (מעל F) ־ כשמה כן היא:

$$\begin{aligned} a_{11}X_1 + a_{12}X_2 + \cdots + a_{1n}X_n &= b_1 \\ a_{21}X_1 + a_{22}X_2 + \cdots + a_{2n}X_n &= b_2 \\ &\cdots \\ a_{m1}X_1 + a_{m2}X_2 + \cdots + a_{mn}X_n &= b_m \end{aligned} \tag{2}$$

באשר $a_{ij}, b_i \in F$ לכל $1 \le i \le m$ ולכל $1 \le j \le n$. **פתרון** של (2) הוא פתרון של כל m המשוואות ב־(2); **קבוצה הפתרון** של (2) הוא אוסף כל הפתרונות שלה.

כיצד פותרים מערכת משוואות לינאריות (כלומר, מוצאים את קבוצת הפתרון שלה)?

דוגמה 2.4 (איך לא לפתור!):

$$\begin{aligned} &(I) && X_1+X_2+X_3=1\\ &(II) && X_1-X_2-X_3=1\\ &(III) && -X_1-X_2+X_3=-1\\ &(IV) && -X_1+X_2+X_3=1 \end{aligned}$$

על ידי חיבור המשוואות נקבל:

$$\begin{aligned} &(I+II) && 2X_1=2\\ &(II+III) && -2X_2=0\\ &(I+II+III+IV) && 2X_3=2 \end{aligned}$$

ולמערכת הזאת פתרון יחיד $(1,0,1)^t$. אך הוא איננו פתרון של המערכת המקורית. (לזו אין פתרון, כי $(II),(IV)$ סותרות זו את זו.)

דוגמה זו ממחישה שאנחנו זקוקים לשיטה שתוביל אותנו לקבוצת הפתרון בכל מקרה. נציג שיטה כזאת – על יד פעולות אלמנטריות על מערכות המשוואות ושיטת החילוץ של Gauss.

כפי שנראה מיד, פעולה אלמנטרית היא שינוי פשוט יחסית של מערכת משוואות נתונה שכתוצאתו מתקבלת מערכת משוואות אחרת, אך *בעלת אותה קבוצת הפתרון.* על ידי סדרה של פעולות אלמנטריות ניתן להגיע מהמערכת הנתונה למערכת חדשה מאד שונה ממנה, אך, עדיין, בעלת אותה קבוצת הפתרון. אם נבחר את סדרת הפעולות האלמנטריות באופן מושכל, נגיע למערכת פשוטה, אותה כבר נדע לפתור.

הגדרה 2.5: **פעולה אלמנטרית** על מערכת (2) היא אחת מבין השלוש הבאות:

(א) החלפת שתי משוואות במערכת ביניהן; סימון: החלפת המשוואות ה־i וה־k ביניהן תסומן על ידי $\mathcal{P}_{ik}$ או $R_i \leftrightarrow R_k$.

(ב) הכפלת משוואה אחת במערכת ב־α, באשר $0 \neq \alpha \in F$. סימון: הכפלת המשוואה ה־i ב־α תסומן על ידי $\mathcal{P}_i(\alpha)$ או $R_i \to \alpha R_i$.

(ג) הוספת הכפולה ב־λ של משוואה אחת במערכת לאחרת, באשר $\lambda \in F$. סימון: הוספת הכפולה ב־λ של המשוואה ה־i למשוואה ה־k תסומן על ידי $P_{ik}(\lambda)$ או $R_k \to R_k + \lambda R_i$. (בפעולה זו אין משנים את המשוואה ה־i.)

למה 2.6: *פעולה אלמנטרית אינה משנה את קבוצת הפתרון, כלומר, למערכת נתונה ולמערכת שמתקבלת ממנה על ידי פעולה אלמנטרית אחת אותה קבוצת פתרון.*

הוכחה: נניח למשל שהפעולה היא $\mathcal{P}_{ik}(\lambda)$. תהי (2) המערכת המקורית ותהי $(2')$ המערכת המתקבלת על ידי הפעולה. צריך להוכיח שכל פתרון של (2) הוא גם פתרון של $(2')$ ולהיפך, כל פתרון של $(2')$ הוא גם פתרון של (2).

נשים לב ש־$(2')$ נבדלת מ־(2) רק במשוואה ה־k, שהינה:

$$(\lambda a_{i1}+a_{k1})X_1+(\lambda a_{i2}+a_{k2})X_2+\cdots+(\lambda a_{in}+a_{kn})X_n=\lambda b_i+b_k.$$

יהי $(x_1,x_2,\ldots,x_n)^t$ פתרון של (2). אז

$$a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n=b_i$$

$$a_{k1}x_1+a_{k2}x_2+\cdots+a_{kn}x_n=b_k.$$

נכפיל את השוויון הראשון ב־λ ונוסיפו לשני. אז

$$(\lambda a_{i1}+a_{k1})x_1+(\lambda a_{i2}+a_{k2})x_2+\cdots+(\lambda a_{in}+a_{kn})x_n=\lambda b_i+b_k.$$

כלומר, $(x_1,x_2,\ldots,x_n)^t$ הוא פתרון של המערכת החדשה.

להיפך, יהי $(x_1,x_2,\ldots,x_n)^t$ פתרון של המערכת החדשה. הוא בפרט פתרון של המשוואה ה־i והמשוואה ה־k של מערכת זו. לכן

$$a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n=b_i$$

$$(\lambda a_{i1}+a_{k1})x_1+(\lambda a_{i2}+a_{k2})x_2+\cdots+(\lambda a_{in}+a_{kn})x_n=\lambda b_i+b_k.$$

נכפיל את השוויון הראשון ב־$(-\lambda)$ ונוסיפו לשני. אז נקבל כי $(x_1,x_2,\ldots,x_n)^t$ פתרון של המשוואה ה־k במערכת המקורית (2), ולכן פתרון של (2).

באופן דומה לגבי פעולות אלמטריות מהסוגים האחרים. ∎

שאלה 2.7: מדוע בפעולה $\mathcal{P}_i(\alpha)$ נדרש $\alpha\neq 0$?

דוגמה 2.8: מערכת משוואות מעל $F=\mathbb{R}$

$$\begin{array}{cccccc} X & +Y & +Z & -W & = & 1\\ 2X & +3Y & +4Z & +aW & = & 3\\ X & +aY & +(2a-1)Z & +3W & = & 2 \end{array}$$

$$\mathcal{P}_{12}(-2),\mathcal{P}_{13}(-1)$$

$$\begin{array}{cccccc} X & +Y & +Z & -W & = & 1\\ & Y & +2Z & +(a+2)W & = & 1\\ & (a-1)Y & +(2a-2)Z & +4W & = & 1 \end{array}$$

$$\mathcal{P}_{21}(-1),\mathcal{P}_{23}(1-a)$$

$$\begin{array}{cccccc} X & & -Z & -(a+3)W & = & 0\\ & +Y & +2Z & +(a+2)W & = & 1\\ & & & bW & = & 2-a \end{array}$$

באשר

$$b=4+(1-a)(a+2)=4-a^2-a+2=6-a-a^2=(3+a)(2-a).$$

אם $a=-3$, למערכת זו אין פתרון, כי משוואתה האחרונה היא $0W=5$, ולה אין פתרון.

אם $a=2$, המערכת היא

$$\begin{array}{cccccc} X & & -Z & -5W & = & 0 \\ & +Y & +2Z & +4W & = & 1 \\ & & & 0W & = & 0 \end{array}$$

וקבוצת הפתרון שלה היא

$$.P=\{(c+5d,-2c-4d+1,c,d)^t |\ c,d\in F\}$$

אם $a\neq 2,3$, על ידי שלוש פעולות אלמנטריות נגיע למערכת

$$\begin{array}{ccccc} X & & -Z & & = & 1 \\ & Y & +2Z & & = & \frac{1}{a+3} \\ & & & W & = & \frac{1}{a+3} \end{array}$$

(בדוק! פסחנו על הסבר מפורט) וקבוצת הפתרון שלה היא

$$.P=\{(c+1,-2c+\frac{1}{a+3},c,\frac{1}{a+3})^t |\ c\in F\}$$

תרגיל 2.9: *פתור:*

$$\begin{array}{cccccccc} X_1 & +a_{12}X_2 & & & +a_{15}X_5 & +a_{16}X_6 & = & b_1 \\ & & +X_3 & & +a_{25}X_5 & +a_{26}X_6 & = & b_2 \\ & & & X_4 & +a_{35}X_5 & +a_{36}X_6 & = & b_3 \end{array}$$

(שים לב לצורת המדרגות של המערכת.)

קבוצת הפתרון היא

$$.P=\{(b_1-a_{12}c_2-a_{15}c_5-a_{16}c_6,c_2,b_2-a_{25}c_5-a_{26}c_6,b_3-a_{35}c_5-a_{36}c_6,c_5,c_6)^t$$
$$|\ c_2,c_5,c_6\in F\}$$

כדי לחסוך כתיבה, נהוג להחליף מערכת משוואות במקדמים שלה:

הגדרה 2.10: **מטריצה מסדר** $m\times n$ **מעל** F היא מלבן של איברי F מסדר $m\times n$ מעל F. רישום:

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \qquad (3)$$

באשר $a_{ij}\in F$ לכל $1\le i\le m,\ 1\le j\le n$.

אוסף כל המטריצות מסדר $m\times n$ מעל F יסומן $\mathrm{M}_{m\times n}(F)$. נכתוב בדרך כלל $\mathrm{M}_n(F)$ במקום $\mathrm{M}_{n\times n}(F)$. נשים לב ש-$\mathrm{M}_{n\times 1}(F)=F^n$.

הגדרה 2.11: **מטריצה המורחבת של המערכת** (2) היא המטריצה מסדר $m \times (n+1)$ הבאה

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}$$

מטריצה המצומצמת של (2) היא המטריצה מסדר $m \times n$ שמתקבלת מהמורחבת על ידי השמטת העמודה האחרונה.

מטריצה אינה חייבת לבוא מאיזה מערכת משוואות! נדון במטריצות באופן כללי, ולעתים (אך לא תמיד) נייִשם זאת למערכות של משוואות.

הגדרה 2.12: **פעולה אלמנטרית על** (**שורות**) המטריצה (3) היא אחת מבין השלוש הבאות:

(א) החלפת שתי שורות במטריצה ביניהן; סימון: החלפת השורה ה־i וה־k ביניהן תסומן על ידי $\mathcal{P}_{ik}$ או $R_i \leftrightarrow R_k$.

(ב) הכפלת שורה אחת במטריצה ב־α, באשר $0 \neq \alpha \in F$. סימון: הכפלת השורה ה־i ב־α תסומן על ידי $\mathcal{P}_i(\alpha)$ או $R_i \to \alpha R_i$.

(ג) הוספת הכפולה ב־λ של שורה אחת במטריצה לאחרת, באשר $\lambda \in F$. סימון: הוספת הכפולה ב־λ של השורה ה־i לשורה ה־k תסומן על ידי $P_{ik}(\lambda)$ או $R_k \to R_k + \lambda R_i$.

הגדרה 2.13: שתי מטריצות מאותו סדר נקראות **שקולות** (**־שורות**) אם אחת מתקבלת מהשניה על ידי סדרה של פעולות אלמנטריות.

הערה 2.14: מטריצות שקולות הן תמיד מאותו סדר. אין להשמיט שורות מהן, גם אם הן שורות של אפסים!

הערה 2.15: לכל פעולה אלמנטרית $\mathcal{P}$ מתאימה פעולה אלמנטרית **הפוכה** $\mathcal{P}'$ במובן הבא: אם A' מתקבלת מ־A על ידי $\mathcal{P}$ אז A מתקבלת מ־A' על ידי $\mathcal{P}'$, ולהיפך. אכן,

(א) אם $\mathcal{P} = \mathcal{P}_{ik}$ אז $\mathcal{P}' = \mathcal{P}_{ik} = \mathcal{P}$.

(ב) אם $\mathcal{P} = \mathcal{P}_i(\alpha)$ אז $\mathcal{P}' = \mathcal{P}_i(\alpha^{-1})$.

(ג) אם $\mathcal{P} = \mathcal{P}_{ik}(\lambda)$ אז $\mathcal{P}' = \mathcal{P}_{ik}(-\lambda)$.

מכאן יוצא שיחס השקילות בין מטריצות הוא סימטרי. קל לראות שהוא גם טרנזיטיבי, דהיינו: אם A, A' שקולות וגם A', A'' שקולות אז גם A, A'' שקולות.

כדי לפתור מערכת משוואות לינאריות, אנחנו בעצם מבצעים פעולות אלמנטריות על השורות של המטריצה המורחבת של המערכת, בשאיפה להגיע למטריצה "פשוטה" יותר. מהי מטריצה "פשוטה" זו, נבהיר בהגדרה להלן.

הגדרה 2.16: מטריצה A מסדר $m \times n$ (ראה (3)) נקראת **מדורגת קנונית** אם:

(א) שורות האפסים שבה באות בסוף, כלומר, יש איזה $0 \leq r \leq m$ כך ש־r השורות הראשונות של A אינן שורות אפסים, ו־$m - r$ השורות האחרונות הן שורות של אפסים.

(ב) בכל שורה שאיננה שורת אפסים הרְכיב הראשון השונה מאפס הוא 1. כלומר, לכל $1 \leq i \leq r$ יהי $1 \leq j_i \leq n$ הקטן ביותר כך ש־$a_{ij_i} \neq 0$. רכיב a_{ij_i} נקרא **הרכיב המוביל** (**פִּיבוֹט**) **של השורה ה־**i. אז $a_{ij_i} = 1$.

(ג) $j_1 < j_2 < \cdots < j_r$.

(ד) בעמודה של רכיב מוביל כל הרכיבים הם 0, מלבד הרכיב המוביל (שהינו 1).

דוגמאות 2.17:

$$\begin{pmatrix} 0 & 0 & 1 & 2 & 5 & 0 & 8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, (0), (1)$$

העמודות $j_1, j_2, \ldots, j_r$ של מטריצה מדורגת קנונית שמכילות איבר מוביל נקראות **עמודות מובילות.** במטריצה השמאלית לעיל העמודות המובילות הן השלישית, השישית והשמינית. במטריצה שלידה כל העמודות הן מובילות.

הערה 2.18: (א) ניתן לסכם, בצורה קצת פחות מדויקת, את ההגדרה של מטריצה מדורגת קנונית כך: ניתן לצייר במטריצה צורת מדרגות (אולי ריקה) כך שמתחתיה יש רק אפסים, לפני כל מדרגה עומד 1 ומעליו יש רק אפסים.

(ב) מספר העמודות המובילות במטריצה מדורגת קנונית הוא מספר השורות שאינן שורות של אפסים (כי בכל שורה כזאת יש בדיוק איבר מוביל אחד).

תרגיל 2.19: האם, לכל מטריצה מדורגת קנונית A, המטריצה A' הבאה גם מדורגת קנונית?

(i) A' מתקבלת מ־A על ידי השמטת שורות אחדות.

(ii) A' מתקבלת מ־A על ידי השמטת עמודות אחדות.

(iii) A' מתקבלת מ־A על ידי השמטת עמודות אחרונות אחדות.

(iv) A' מתקבלת מ־A על ידי השמטת עמודות לא מובילות אחדות.

פתרון: בכל סעיף צריך לבדוק שהתנאים של ההגדרה לעיל מתקיימים:

(i) כן. (בדיקה ישירה.)

(ii) לא. דוגמה נגדית: נשמיט מהמטריצה $\begin{pmatrix} 1 & 2 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}$ את העמודה הראשונה (או שלוש הראשונות).

(iii) כן. רק הדרישה (א) איננה ברורה מיידית. אם נסמן את השורות של A ב־$R_1, \ldots, R_m$ ואת השורות של A' ב־$R'_1, \ldots, R'_m$, צריך להוכיח שאם $i \leq k$ ו־R'_k איננה שורת אפסים, אז גם R'_i אינה שורת אפסים.

נניח ש־$R'_k \neq 0$. אז $R_k \neq 0$ והמוביל בשורה ה־k נמצא בעמודה שלא הושמטה. כיוון ש־$i \leq k$, גם $R_i \neq 0$ והמוביל בשורה זו נמצא יותר שמאלה מהמוביל שבשורה R_k, ולכן גם עמודתו לא הושמטה. לכן $R'_i \neq 0$.

(iv) כן, מספרי השורות השונות מאפס אינם משתנים. ∎

תרגיל 2.20: מהן המטריצות המדורגות קנונית מסדר

(א) $m \times 1$?

(ב) $m \times 2$?

פתרון:

$$a\in F\qquad \begin{pmatrix}0&0\\0&0\\0&0\\\vdots&\vdots\\0&0\end{pmatrix},\begin{pmatrix}0&1\\0&0\\0&0\\\vdots&\vdots\\0&0\end{pmatrix},\begin{pmatrix}1&0\\0&1\\0&0\\\vdots&\vdots\\0&0\end{pmatrix},\begin{pmatrix}1&a\\0&0\\0&0\\\vdots&\vdots\\0&0\end{pmatrix},\qquad \begin{pmatrix}0\\0\\0\\\vdots\\0\end{pmatrix},\begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix}\qquad\blacksquare$$

אלגוריתם 2.21: פתרון של מערכת שמטריצת המקדמים המורחבת שלה מדורגת קנונית. נניח שנתונה מערכת של m משוואות לינאריות ב־n נעלמים שמטריצת המקדמים המורחבת שלה (מסדר $m\times(n+1)$) הינה מדורגת קנונית. מהי אז קבוצת הפתרון שלה?

(א) אם העמודה האחרונה היא מובילה, אז יש שורה במטריצה מהצורה $(0,\dots,0,1)$, שמתאימה למשוואה $0X_1+\dots+0X_n=1$. למשוואה זו אין פתרון ולכן למערכת אין פתרון ($P=\emptyset$).

(ב) אם העמודה האחרונה אינה מובילה, נשאיר את r המשתנים שמתאימים לעמודות מובילות באגף שמאל, ונעביר את יתר המשתנים לאגף ימין. כעת קל לרשום את קבוצת הפתרון. למשל, אם עמודות המובילות הן $1,2,\dots,r$, והמטריצה היא

$$\begin{pmatrix}
1&0&\dots&0&a_{1\,r+1}&\dots&a_{1\,n}&b_1\\
0&1&\dots&0&a_{2\,r+1}&\dots&a_{2\,n}&b_2\\
\dots\\
0&0&\dots&1&a_{r\,r+1}&\dots&a_{r\,n}&b_r\\
0&0&\dots&0&0&\dots&0&0\\
\dots\\
0&0&\dots&0&0&\dots&0&0
\end{pmatrix}$$

ואז קבוצת הפתרון $P\neq\emptyset$ היא

$$\left\{\begin{pmatrix}
-a_{1\,r+1}x_{r+1}-\dots-a_{1\,n}x_n+b_1\\
-a_{2\,r+1}x_{r+1}-\dots-a_{2\,n}x_n+b_2\\
\vdots\\
-a_{r\,r+1}x_{r+1}-\dots-a_{r\,n}x_n+b_r\\
x_{r+1}\\
\vdots\\
x_n
\end{pmatrix}\;\middle|\; x_{r+1},\dots,x_n\in F\right\}$$

קבוצת הפתרון אינה ריקה ותלויה ב־$(n-r)$ פרמטרים (לפחות לפי איך שכתבנו אותה).

משפט 2.22: *כל מטריצה שקולת שורות למטריצה מדורגת קנונית אחת ויחידה, כלומר, מכל מטריצה ניתן להגיע על יד סדרה של פעולות אלמנטריות למטריצה מדורגת קנונית אחת ויחידה.*

הוכחה: תהי $A=(a_{ij})$ מטריצה מסדר $m\times n$. נציג את אלגוריתם (שיטת החילוץ של גאוס) שמוביל למטריצה מדורגת קנונית. (נדגיש שזאת לא בהכרח הדרך היחידה לקבל מטריצה מדורגת קנונית על ידי סדרה של פעולות אלמנטריות.)

(א) אם כל הרכיבים ב־A הם 0, המטריצה כבר מדורגת קנונית, ולכן סיימנו. נניח שזה לא כך. אז

(ב) יש עמודה ב־A שאיננה עמודה של אפסים. יהי j_1 המספר הקטן ביותר כך שהעמודה ה־j_1 איננה עמודה של אפסים (ואז כל העמודות לפניה הן עמודות של אפסים). ניתן להניח ש־$a_{1j_1} \neq 0$; אם לא, נחליף את השורה הראשונה עם שורה מתאימה אחרת שמתחתיה, בה יש רכיב שונה מאפס בעמודה ה־j_1.

נכפיל את שורה הראשונה ב־$a_{1j_1}^{-1}$ ואז נוכל להניח כי $a_{1j_1} = 1$.

על ידי הוספת כפולות מתאימות של השורה הראשונה ליתר השורות נוכל להניח שבעמודה ה־j_1 כל הרכיבים הם אפסים, מלבד $a_{1j_1} = 1$.

שים לב שכל הפעולות האלמנטריות שעשינו עד כה לא שינו את העמודות שלפני העמודה ה־j_1.

תהי A' המטריצה המתקבלת מ־A על ידי השמטת השורה הראשונה. אם כל השורות ב־A' הן שורות אפסים, אז A כבר מדורגת קנונית. נניח שזה לא כך. אז

(ג) יש עמודה ב־A' שאיננה עמודה של אפסים. יהי j_2 המספר הקטן ביותר כך שהעמודה ה־j_2 של A' איננה עמודה של אפסים (ואז כל העמודות לפניה ב־A' הן עמודות של אפסים). נשים לב ש־$j_1 < j_2$, כי כל העמודות לפני העמודה j_1 הן עמודות אפסים ב־A, ואילו בעמודה ה־j_1 יש רק רכיב אחד שונה מאפס, אך הוא בשורה הראשונה המושמטת. ניתן להניח ש־$a_{2j_2} \neq 0$; אם לא, נחליף את השורה השניה של A עם שורה מתאימה אחרת שמתחתיה, בה יש רכיב שונה מאפס בעמודה ה־j_1.

נכפיל את שורה השניה של A ב־$a_{2j_2}^{-1}$ ואז נוכל להניח כי $a_{2j_2} = 1$.

על ידי הוספת כפולות מתאימות של השורה השניה של A ליתר השורות (גם לראשונה!) נוכל להניח שבעמודה ה־j_2 כל הרכיבים הם אפסים, מלבד $a_{2j_2} = 1$.

שים לב שכל הפעולות האלמנטריות שעשינו עד כה לא שינו את העמודות שלפני העמודה ה־j_2.

וכן הלאה.

(למי שזה בכל זאת אינו מספיק, נפרט כאן את השלב הכללי:

נוכל להניח, אחרי פעולות אלמנטריות שכבר עשינו, שבמטירצה A יש עמודות $j_1 < j_2 < \cdots < j_{k-1}$, כך שלכל $1 \leq i < k$ מתקיים:

(i) בעמודה ה־j_i יש 1 בשורה ה־i ואפסים ביתר השורות,

(ii) ובשורה ה־i יש רק אפסים לפני העמודה ה־j_i.

כמו כן,

(iii) בשורות $k, k+1, \ldots, m$ יש רק אפסים בעמודות שלפני העמודה ה־j_{k-1} (וגם בעמודה ה־j_{k-1}, לפי (i)).

תהי A' המטריצה המתקבלת מ־A על ידי השמטת k השורה הראשונה. אם כל השורות ב־A' הן שורות אפסים, אז A כבר מדורגת קנונית. נניח שזה לא כך. אז

(ד) יש עמודה ב־A' שאיננה עמודה של אפסים. יהי j_k המספר הקטן ביותר כך שהעמודה ה־j_k של A' איננה עמודה של אפסים (ואז כל העמודות לפניה ב־A' הן עמודות של אפסים). נשים לב ש־$j_{k-1} < j_k$, כי כל העמודות ב־A' לפני העמודה j_{k-1} וכולל אותה הן עמודות אפסים לפי (iii).

ניתן להניח ש־$a_{kj_k} \neq 0$; אם לא, נחליף את השורה ה־k של של A עם שורה מתאימה אחרת שמתחתיה, בה יש רכיב שונה מאפס בעמודה ה־j_1. נכפיל את שורה ה־k של A ב־ $a_{kj_k}^{-1}$ ואז נוכל להניח כי $a_{kj_k} = 1$.

על ידי הוספת כפולות מתאימות של השורה ה־k של A ליתר השורות (גם לראשונה!) נוכל להניח שבעמודה ה־j_k כל הרכיבים הם אפסים, מלבד $a_{kj_k} = 1$.

לפי הבניה:

(i') בעמודה ה־j_k יש 1 בשורה ה־k ואפסים ביתר השורות,

(ii') ובשורה ה־k יש רק אפסים לפני העמודה ה־j_k.

(iii') בשורות $k+1, \ldots, m$ יש רק אפסים בעמודות שלפני העמודה ה־j_k.

ולכן אפשר להמשיך הלאה באותו אופן. שים לב שכל הפעולות האלמנטריות שעשינו עד כה לא שינו את העמודות שלפני העמודה ה־j_k.)

את היחידות נוכיח אחרי המשפט הבא. ∎

משפט 2.23: *אם A, B מדורגות קנונית ושקולות שורות (כלומר ניתן להגיע על ידי סדרה של פעולות אלמנטריות מ־A ל־B) אז $A = B$.*

הוכחה: באינדוקציה על מספר העמודות.

(א) נניח כי ב־A, B רק עמודה אחת. אם $A = 0$, אז כל פעולה אלמנטרית לא תשנה אותה, ולכן גם $B = 0$. אם $A \neq 0$, אז גם $B \neq 0$ (אחרת, אם $B = 0$, אפשר להגיע מ־B ל־A על ידי סדרה של פעולות אלמנטריות ולכן $A = 0$, סתירה). לכן, לפי תרגיל 2.20(א), $B = (1, 0, 0, \ldots, 0)^t = A$.

(ב) נניח שהטענה נכונה למטריצות מסדר $m \times n$, כאשר A, B מסדר $m \times (n+1)$. תהיינה A', B' המטריצות מסדר $m \times n$ המתקבלות מן A, B על ידי השמטת העמודה האחרונה. לפי תרגיל 2.19(iii) הן מדורגות קנונית.

אותה סדרה של פעולות אלמנטריות שמעבירה את A ל־B, מעבירה את A' ל־B'. לכן לפי הנחת האינדוקציה $A' = B'$. לכן אפשר לכתוב

$$A' = B' = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdots & & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix} \qquad B = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & c_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & c_2 \\ \cdots & & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} & c_m \end{pmatrix}$$

צריך להוכיח: גם העמודות האחרונות של A, B שוות זו לזו.

יהי r' מספר העמודות המובילות (כלומר, מספר השורות השונות מאפס) ב־$A' = B'$. מערכות המשוואות הלינאריות שמתאימות ל־A ול־B (כמטריצות המורחבות) הן שקולות ולכן יש להן אותה קבוצת הפתרון P.

אם P ריקה, אז, כפי שראינו, העמודה האחרונה של A, B היא מובילה, כלומר יש בה אפסים פרט ל־1 בשורה ה־$(r'+1)$. בפרט העמודות האחרונות ב־A, B שוות.

אם P אינה ריקה, יש פתרון $(x_1, \ldots, x_n)^t \in P$. נציב אותו בשתי המערכות: לכל $1 \leq i \leq m$

$$a_{i1}x_1 + \cdots + a_{in}x_n = c_i, \qquad a_{i1}x_1 + \cdots + a_{in}x_n = b_i,$$

ומכאן $b_i = c_i$. לכן העמודות האחרונות ב־A, B שוות. ∎

מסקנה 2.24: *תהיינה B, C מדורגות קנונית. אם ממטריצה A ניתן להגיע על ידי פעולות אלמנטריות ל־B וגם ל־C, אז $B = C$.*

הוכחה: ניתן להגיע על ידי פעולות אלמנטריות מ־A ל־B, לכן ניתן להגיע על ידי פעולות אלמנטריות (הפוכות) מ־B ל־A. כמו כן ניתן להגיע על ידי פעולות אלמנטריות מ־A ל־C. לכן ניתן להגיע על ידי פעולות אלמנטריות מ־B (דרך A) ל־C. לפי המשפט הקודם, $B = C$. ∎

ממסקנה זו נובעת היחידות במשפט 2.22.

בעקבות המסקנה נוכל להגדיר לכל מטריצה A את **המטריצה המדורגת קנונית של A** כמטריצה (היחידה) המדורגת קנונית שקולה לה, כלומר, שניתן להגיע אליה מ־A על ידי פעולות אלמנטריות.

הגדרה 2.25: תהי A מטריצה. **הדרגה** של A היא מספר השורות השונות מאפס במטריצה המדורגת קנונית של A. זהו גם מספר העמודת המובילות במטריצה המדורגת קנונית של A. הדרגה של A תסומן $\mathrm{rk}(A)$.

מסקנה 2.26: *תהי $A \in \mathrm{M}_{m\times n}(F)$. אז* $0 \leq \mathrm{rk}(A) \leq m, n$.

הוכחה: המטריצה המדורגת קנונית של A היא מסדר $m \times n$, לכן מספר השורות השונות מאפס בה $\leq m$ ומספר העמודות המובילות בה $\leq n$. ∎

מסקנה 2.27: *אם A, B מטריצות שקולות, יש להן אותה הדרגה.*

הוכחה: ניתן להגיע על ידי פעולות אלמנטריות מ־A ל־B, וממנה למטריצה המדורגת קנונית C של B. לכן ל־A, B אותה מטריצה מדורגת קנונית C. ∎

מסקנה 2.28: *תהי A מטריצת המקדמים המורחבת של מערכת משוואות לינאריות ב־n נעלמים ותהי A' מטריצת המקדמים המצומצמת שלה. אז*

(א) $\mathrm{rk}(A') = \mathrm{rk}(A)$ או $\mathrm{rk}(A') = \mathrm{rk}(A) - 1$.

(ב) *אם $\mathrm{rk}(A') < \mathrm{rk}(A)$ (כלומר $\mathrm{rk}(A') = \mathrm{rk}(A) - 1$) אז אין פתרון למערכת.*

(ג) *אם $\mathrm{rk}(A') = \mathrm{rk}(A) = r$, יש פתרון (ואנחנו יכולים לרשום את קבוצת הפתרון כתלויה ב־$(n-r)$ פרמטרים). בפרט*

(ד) *יש פתרון יחיד אם ורק אם* n $\mathrm{rk}(A') = \mathrm{rk}(A) = n$.

הוכחה: תהי B המטריצה המדורגת קנונית של A ותהי B' המטריצה מתקבלת ממנה על ידי השמטת העמודה האחרונה. אז גם B' מדורגת קנונית. אותן הפעולות האלמנטריות שמעבירות את A ל־B מעבירות את A' ל־B'. לכן (על ידי החלפת A ב־B ו־A' ב־B', החלפה שאינה משנה את קבוצת הפתרון) נוכל להניח כי A, A' מדורגות קנונית.

כעת,

- אם העמודה האחרונה של A מובילה, אז $\mathrm{rk}(A) = \mathrm{rk}(A') + 1$ וראינו שאין פתרון.
- אם העמודה האחרונה של A אינה מובילה, אז $\mathrm{rk}(A) = \mathrm{rk}(A')$ וראינו שקבוצת הפתרון תלויה ב־$(n-r)$ פרמטרים.

מכאן (בדוק!) (א), (ב), (ג).

חלק (ד) נובע מ־(ג). ∎

אם המערכת היא הומוגנית, אז ב־A –ולכן גם במטריצה המדורגת קנונית שלה– העמודה האחרונה היא עמודת אפסים, ולכן אינה מובילה. לכן במקרה זה $\mathrm{rk}(A) = \mathrm{rk}(A')$ ויש פתרון. אך קל לראות זאת גם אחרת: $(0, 0, \ldots, 0)^t$ הוא פתרון. הוא נקרא **הפתרון הטריוויאלי**.

מסקנה 2.29: *למערכת הומוגנית יש תמיד פתרון. יש לה פתרון טריוויאלי בלבד אם ורק אם דרגת מטריצת המקדמים המצומצמת שווה למספר הנעלמים.*

3. מטריצות

בפרק זה יהי F שדה כלשהו. סקלר פירושו איבר של F.

כאמור,

הגדרה 3.1: **מטריצה מסדר** $m \times n$ **מעל** F היא מלבן של איברי F מסדר $m \times n$ מעל F. רישום:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

באשר $a_{ij} \in F$ לכל $1 \leq j \leq n$, $1 \leq i \leq m$. רישום מקוצר: $A = (a_{ij})$.

אוסף כל המטריצות מסדר $m \times n$ מעל F יסומן $\mathrm{M}_{m\times n}(F)$. נכתוב בדרך כלל $\mathrm{M}_n(F)$ במקום $\mathrm{M}_{n\times n}(F)$. נשים לשב ש־$F^n = \mathrm{M}_{n\times 1}(F)$.

רישום נוסף: אם A מטריצה אז $(A)_{ij}$ מסמן את הרכיב שלה בשורה ה־i ובעמודה ה־j (בסימון לעיל, $(A)_{ij} = a_{ij}$).

הגדרה 3.2: חיבור של שתי מטריצות וכפל מטריצה בסקלר: אם $A, B \in \mathrm{M}_{m\times n}(F)$ אז $A+B \in \mathrm{M}_{m\times n}(F)$ מוגדרת על ידי: $(A+B)_{ij} = (A)_{ij} + (B)_{ij}$ לכל $1 \leq i \leq m$ ולכל $1 \leq j \leq n$.

אם $\alpha \in F$ אז $\alpha A \in \mathrm{M}_{m\times n}(F)$ מוגדרת על ידי: $(\alpha A)_{ij} = \alpha(A)_{ij}$

נגדיר גם $-A = (-1)A$.

תכונות:

משפט 3.3: *מתקיימים החוקים הבאים:*

(ח1) *אסוציאטיביות החיבור:* $(A+B)+C = A+(B+C)$, *לכל* $A, B, C \in \mathrm{M}_{m\times n}$.

(ח2) *תכונת איבר האפס:* $B+0 = B = 0+B$, *לכל* $B \in \mathrm{M}_{m\times n}$.

(ח3) *תכונת איבר נגדי:* $B+(-B) = 0 = (-B)+B$.

(ח4) *קומוטטיביות החיבור:* $A+B = B+A$, *לכל* $A, B \in \mathrm{M}_{m\times n}$.

(כ1) *חוק הפילוג:* $\alpha(A+B) = \alpha A + \alpha B$, *לכל* $A, B \in \mathrm{M}_{m\times n}$ *ולכל* $\alpha \in F$.

(כ2) *חוק הפילוג:* $(\alpha+\beta)B = \alpha B + \beta B$, *לכל* $B \in \mathrm{M}_{m\times n}$ *ולכל* $\alpha, \beta \in F$.

(כ3) *אסוציאטיביות הכפל בסקלר:* $(\alpha\beta)B = \alpha(\beta B)$, *לכל* $B \in \mathrm{M}_{m\times n}$ *ולכל* $\alpha, \beta \in F$.

(כ4) $1B = B$ *לכל* $B \in \mathrm{M}_{m\times n}$.

נגדיר כפל בין מטריצות:

הגדרה 3.4: תהיינה $A \in \mathrm{M}_{m\times p}(F), B \in \mathrm{M}_{p\times n}(F)$. נגדיר $AB \in \mathrm{M}_{m\times n}(F)$ על ידי

$$(AB)_{ij} = (A)_{i1}(B)_{1j} + (A)_{i2}(B)_{2j} + \cdots + (A)_{ip}(B)_{pj} = \sum_{k=1}^{p}(A)_{ik}(B)_{kj},$$

לכל $1 \le i \le m,\ 1 \le j \le n$

(ובמלים: הרכיב בשורה ה־i ובעמודה ה־j של המכפלה AB הוא סכום של מכפלות רכיבי השורה ה־i של A ברכיבי עמודה ה־j של B, בהתאמה.)

דוגמאות3.5:

$$\begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix}\begin{pmatrix} \alpha & \alpha' \\ \beta & \beta' \\ \gamma & \gamma' \end{pmatrix} = \begin{pmatrix} a\alpha + b\beta + c\gamma & a\alpha' + b\beta' + c\gamma' \\ a'\alpha + b'\beta + c'\gamma & a'\alpha' + b'\beta' + c'\gamma' \end{pmatrix}$$

$$\begin{pmatrix} \alpha & \alpha' \\ \beta & \beta' \\ \gamma & \gamma' \end{pmatrix}\begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} = \begin{pmatrix} \alpha a + \alpha' a' & \alpha b + \alpha' b' & \alpha c + \alpha' c' \\ \beta a + \beta' a' & \beta b + \beta' b' & \beta c + \beta' c' \\ \gamma a + \gamma' a' & \gamma b + \gamma' b' & \gamma c + \gamma' c' \end{pmatrix}$$

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

לא כל שתי מטריצות ניתן להכפיל. אך גם כאשר שתי מטריצות ריבועיות A, B מאותו סדר, לא בהכרח מתקיים $AB = BA$, ויתכן ש־$A, B \neq 0$ אך $AB = 0$ (לפי הדוגמאות לעיל). אנו רואים מהדוגמאות שהכפל אינו חילופי.

סימון 3.6: אם $B_1, B_2, \ldots, B_n \in F^m = \mathrm{M}_{m\times 1}(F)$ עמודות, אז $(B_1, \ldots, B_n)$ מסמן את המטריצה מסדר $m \times n$ ש־$B_1, \ldots, B_n$ הן העמודות שלה.

נסמן גם

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_m = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \in F^m.$$

(סימון זה תלוי ב־m, שצריך להיות ברור מההקשר.)

תרגיל 3.7: תהיינה $A \in \mathrm{M}_{m\times p}(F), (B_1, \ldots, B_n) \in \mathrm{M}_{p\times n}(F)$, אז

$$A(B_1, \ldots, B_n) = (AB_1, \ldots, AB_n).$$

הוכחה: יהיו $1 \le i \le m,\ 1 \le j \le n$. אז

$$(A(B_1, \ldots, B_n))_{ij} = \sum_{k=1}^{p}(A)_{ik}((B_1, \ldots, B_n))_{kj} = \sum_{k=1}^{p}(A)_{ik}(B_j)_{k1} = (AB_j)_{i1}$$
$$= ((AB_1, \ldots, AB_n))_{ij} \qquad \blacksquare$$

תרגיל 3.8: תהי $A \in \mathrm{M}_{m\times n}(F)$. אז Ae_j (באשר $e_j \in F^n$ הוגדרה לעיל) היא העמודה ה־j של A.

משפט 3.9: *מתקיימים החוקים הבאים:*

(כו) *אסוציאטיביות הכפל:* $(AB)C = A(BC)$, לכל $C \in \mathrm{M}_{q\times n}(F)$, $B \in \mathrm{M}_{p\times q}(F)$, $A \in \mathrm{M}_{m\times p}(F)$.

(2) *חוק הפילוג:* $A(B+C) = AB + AC$ לכל $B, C \in \mathrm{M}_{p\times n}(F)$ ולכל $A \in \mathrm{M}_{m\times p}(F)$.

(3) *חוק הפילוג:* $(A+B)C = AC + BC$ לכל $C \in \mathrm{M}_{p\times n}(F)$ ולכל $A, B \in \mathrm{M}_{m\times p}(F)$.

(4) $(\alpha A)B = A(\alpha B) = \alpha(AB)$, לכל $B \in \mathrm{M}_{p\times n}(F)$, $A \in \mathrm{M}_{m\times p}(F)$, ולכל $\alpha \in F$.

(5) *המטריצה* $(e_1, e_2, \ldots, e_m) \in \mathrm{M}_m(F)$ $= \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} =: I_m$ *מקיימת* $I_m A = A$ *לכל* $A \in \mathrm{M}_{m\times n}(F)$ ו־$BI_m = B$ *לכל* $B \in \mathrm{M}_{n\times m}(F)$. *נקראת* ***מטריצת היחידה*** *מסדר* $m \times m$.

הוכחה: (1) נשים לב שכל המכפלות מוגדרות (למשל, AB מסדר $m \times q$ ולכן $(AB)C$ מוגדרת) ושני האגפים מסדר $m \times n$.

יהיו $1 \le i \le m,\ 1 \le j \le n$. נראה ש־$\big((AB)C\big)_{ij} = \big(A(BC)\big)_{ij}$.

$$\big((AB)C\big)_{ij} = \sum_{\ell=1}^{q}(AB)_{i\ell}(C)_{\ell j} = \sum_{\ell=1}^{q}\Big[\sum_{k=1}^{p}(A)_{ik}(B)_{k\ell}\Big](C)_{\ell j} = \sum_{\ell=1}^{q}\Big[\sum_{k=1}^{p}(A)_{ik}(B)_{k\ell}(C)_{\ell j}\Big]$$

ואילו

$$\big(A(BC)\big)_{ij} = \sum_{k=1}^{p}(A)_{ik}(BC)_{kj} = \sum_{k=1}^{p}(A)_{ik}\Big[\sum_{\ell=1}^{q}(B)_{k\ell}(C)_{\ell j}\Big] = \sum_{k=1}^{p}\Big[\sum_{\ell=1}^{q}(A)_{ik}(B)_{k\ell}(C)_{\ell j}\Big]$$

ושני האגפים הימניים שווים: שניהם הסכום של כל $(A)_{ik}(B)_{k\ell}(C)_{\ell j}$, באשר $1 \le k \le p, 1 \le \ell \le q$, אך בכל אחד סדר המחוברים וקיבוציהם לסוגריים שונים.

(2) שוב, נשים לב שמבחינת הסדרים של המטריצות כל המכפלות והסכומים מוגדרים.

יהיו $1 \le i \le m,\ 1 \le j \le n$. נראה ש־$\big(A(B+C)\big)_{ij} = (AB)_{ij} + (AC)_{ij}$.

$$\begin{aligned}\big(A(B+C)\big)_{ij} &= \sum_{k=1}^{p}(A)_{ik}(B+C)_{kj} = \sum_{k=1}^{p}(A)_{ik}[(B)_{kj} + (C)_{kj}] \\ &= \sum_{k=1}^{p}(A)_{ik}(B)_{kj} + (A)_{ik}(C)_{kj} = \sum_{k=1}^{p}(A)_{ik}(B)_{kj} + \sum_{k=1}^{p}(A)_{ik}(C)_{kj} \\ &= (AB)_{ij} + (AC)_{ij} = (AB + AC)_{ij} \qquad \blacksquare\end{aligned}$$

הערה 3.10: רישום מערכת משוואות לינאריות בעזרת מטריצות. תהי נתונה מערכת של m משוואות לינאריות ב־n נעלמים. יהיו

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

עמודת האיברים החפשיים שלה והמטריצה המצומצמת שלה, בהתאמה. אז עמודה $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in F^n$ היא פתרון אם ורק אם $A\mathbf{x} = \mathbf{b}$. לכן רושמים לפעמים את המערכת באופן סמלי כך: $A\mathbf{X} = \mathbf{b}$.

הגדרה 3.11: תהי $A \in \mathrm{M}_m(F)$. מטריצה $B \in \mathrm{M}_m(F)$ תקרא **ההופכי של** A אם $AB = BA = I_m$. סימון: $B = A^{-1}$. מטריצה A תקרא **הפיכה** (**רגולרית, לא סינגולרית**) אם יש לה הופכי.

טענה 3.12: *אם למטריצה יש הופכי, הוא יחיד.*

הוכחה: יהיו B, B' שני הופכיים של A. אז $B' = I_m B' = (BA)B' = B(AB') = BI_m = B$. ∎

דוגמה 3.13: I_m הפיכה: $I_m^{-1} = I_m$. מטריצה A שיש לה שורת אפסים אינה הפיכה (כי למטריצה AB שורת אפסים, לכל B). מטריצה A שיש לה עמודת אפסים אינה הפיכה (כי למטריצה BA עמודת אפסים, לכל B).

למה 3.14: *(א) אם $A \in \mathrm{M}_m(F)$ הפיכה אז A^{-1} הפיכה ו־$(A^{-1})^{-1} = A$.*

(ב) *אם $A_1, \ldots, A_k \in \mathrm{M}_m(F)$ הפיכות אז $A_1 \cdots A_k$ הפיכה ו־$(A_1 \cdots A_k)^{-1} = A_k^{-1} \cdots A_1^{-1}$.*

הוכחה: (א) $AA^{-1} = I_m = A^{-1}A$. **(ב)**

$$(A_1 \cdots A_k)(A_k^{-1} \cdots A_1^{-1}) = A_1 \cdots A_{k-1}\ A_{k-1}^{-1} \cdots A_1^{-1} = \ldots = A_1 A_1^{-1} = I_m$$

$$(A_k \cdots A_1)(A_1^{-1} \cdots A_k^{-1}) = A_k \cdots A_2\ A_2^{-1} \cdots A_k^{-1} = \ldots = A_k A_k^{-1} = I_m$$

∎

סימון 3.15: אם $\mathcal{P}$ פעולה אלמנטרית ו־A מטריצה, אז $\mathcal{P}(A)$ מסמן את התוצאה של של פעולת $\mathcal{P}$ על A.

הגדרה 3.16: תהי $\mathcal{P}$ פעולה אלמנטרית על מטריצות בנות m שורות (ולא חשוב כמה עמודות). אז $E_{\mathcal{P}} = \mathcal{P}(I_m) \in \mathrm{M}_m(F)$ נקראת **המטריצה האלמנטרית המתאימה ל־$\mathcal{P}$**.

אם כן, מטריצות אלמנטריות הן משלושה סוגים (הרכיבים הלא רשומים הם 0):

$$\mathcal{P}_{ij}(\lambda)\qquad \begin{array}{cc} & \begin{array}{cccccccc} & & & i & & & & \end{array} \\ \begin{array}{c} \\ \\ \\ i \\ \\ j \\ \\ \\ \end{array} & \begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & 1 & & & & \\ & & & \vdots & \ddots & & & \\ & & & \lambda & \cdots & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix} \end{array}$$

$$\mathcal{P}_{i}(\alpha)\qquad \begin{array}{cc} & \begin{array}{ccccc} & & i & & \end{array} \\ \begin{array}{c} \\ \\ i \\ \\ \\ \end{array} & \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \alpha & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \end{array}$$

$$\mathcal{P}_{ij}\qquad \begin{array}{cc} & \begin{array}{ccccccc} & & i & & j & & \end{array} \\ \begin{array}{c} \\ \\ i \\ \\ j \\ \\ \\ \end{array} & \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & \cdots & 1 & & \\ & & \vdots & \ddots & \vdots & & \\ & & 1 & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \end{array}$$

משפט 3.17: *תהי $\mathcal{P}$ פעולה אלמנטרית. תהי $A \in \mathrm{M}_{m\times n}$. אז $E_{\mathcal{P}}A = \mathcal{P}(A)$.*

הוכחה: תהיינה $A_1, \dots A_n$ העמודות של A. אז

$$E_{\mathcal{P}}A = E_{\mathcal{P}}(A_1, \dots A_n) = (E_{\mathcal{P}}A_1, \dots E_{\mathcal{P}}A_n)$$
$$\mathcal{P}(A) = \mathcal{P}(A_1, \dots A_n) = (\mathcal{P}(A_1), \dots \mathcal{P}(A_n))$$

לכן די להוכיח $E_{\mathcal{P}}A_j = \mathcal{P}(A_j)$, לכל j, כלומר, די להוכיח את המשפט עבור עמודה $A = (a_1, \dots, a_m)^t \in \mathrm{M}_{m\times 1}(F)$.

במקרה זה נעשה בדיקה ישירה. (עשה זאת!) ∎

מסקנה 3.18: *מטריצה אלמנטרית $E_{\mathcal{P}}$ היא הפיכה, $E_{\mathcal{P}}^{-1} = E_{\mathcal{P}'}$, באשר $\mathcal{P}'$ הפעולה ההפוכה ל־$\mathcal{P}$.*

הוכחה: (1) בדיקה ישירה. או:

(2) $E_{\mathcal{P}}E_{\mathcal{P}'} = E_{\mathcal{P}}\mathcal{P}'(I_m) = \mathcal{P}(\mathcal{P}'(I_m)) = I_m$. באופן דומה $E_{\mathcal{P}'}E_{\mathcal{P}} = I_m$. ∎

משפט 3.19: *תהי $P \in \mathrm{M}_m(F)$, ותהי P' המטריצה המדורגת קנונית שלה. התנאים הבאים שקולים זה לזה:*

(1) *P הפיכה;*

(2) $\mathrm{rk}(P) = m$;

(3) $P' = I_m$;

(4) *P מכפלה של מטריצות אלמנטריות.*

הוכחה: לפי משפט 3.17: $P' = E_{\mathcal{P}_k} \cdots E_{\mathcal{P}_2}E_{\mathcal{P}_1}P$ באשר $E_{\mathcal{P}_1}, E_{\mathcal{P}_2}, \dots, E_{\mathcal{P}_k}$ מטריצות אלמנטריות.

(1) $\Leftarrow$(2) : $E_{\mathcal{P}_1}, E_{\mathcal{P}_2}, \dots, E_{\mathcal{P}_k}, P$ הפיכות, לכן $P' = E_{\mathcal{P}_k} \cdots E_{\mathcal{P}_1}E_{\mathcal{P}_2}P$ הפיכה. בפרט אין בה שורות אפסים. לכן $\mathrm{rk}(P) = m$. (לפי ההגדרה, $\mathrm{rk}(P)$ הוא מספר השורות השונות מ־0 ב־P'.)

(2) $\Leftarrow$(3) : ב־P' יש m עמודות מובילות $e_1, \dots, e_m$. היות ויש לה בדיוק m עמודות, אלה כל עמודותיה. לכן $P' = (e_1, \dots, e_m) = I_m$.

(3) ⇐(4) : מתקיים $E_{\mathcal{P}_k}\cdots E_{\mathcal{P}_2}E_{\mathcal{P}_1}P = I_m$, לכן $P = E_{\mathcal{P}_1}^{-1}E_{\mathcal{P}_2}^{-1}\cdots E_{\mathcal{P}_k}^{-1}$. לפי מסקנה 3.18, המטריצות באגף ימין אלמנטריות.

(4) ⇐(1) : מטריצות אלמנטריות הפיכות, לכן מכפלתן P הפיכה לפי למה 3.14(ב). ∎

מסקנה 3.20: *תהיינה* $A, A' \in \mathrm{M}_{m\times n}(F)$. *אפשר לעבור מ־* A *ל־* A' *על ידי סדרה של פעולות אלמנטריות אם ורק אם* $A' = PA$*, באשר* $P \in \mathrm{M}_m(F)$ *הפיכה.*

הוכחה: $A' = \mathcal{P}_k(\cdots(\mathcal{P}_2(\mathcal{P}_1(A)))\cdots)$ אם ורק אם $A' = E_{\mathcal{P}_k}E_{\mathcal{P}_2}E_{\mathcal{P}_1}A$ אם ורק אם $A' = PA$, באשר P מכפלה של מטריצות אלמנטריות. התנאי האחרון שקול לפי משפט 3.19 לכך ש־P הפיכה. ∎

מסקנה 3.21: *תהי* $A \in \mathrm{M}_{m\times n}(F)$ *ותהי* $P \in \mathrm{M}_m(P)$ *הפיכה. אז* $\mathrm{rk}(PA) = \mathrm{rk}(A)$.

הוכחה: לפי המסקנה הקודמת PA מתקבלת מ־A על ידי סדרה של פעולות אלמנטריות. לכן לשתיהן אותה מטריצה מדורגת קנונית ולכן דרגותיהן שוות. ∎

מסקנה 3.22: *תהיינה* $B \in \mathrm{M}_{p\times n}(F)$, $A \in \mathrm{M}_{m\times p}(F)$.

(א) $\mathrm{rk}(AB) \le \mathrm{rk}(A)$.

(ב) *אם* B *הפיכה אז* $\mathrm{rk}(AB) = \mathrm{rk}(A)$.

הוכחה: (א) יש $P \in \mathrm{M}_m(F)$ הפיכה כך ש־PA מדורגת קנונית. לפי המסקנה הקודמת $\mathrm{rk}(PA) = \mathrm{rk}(A)$, $\mathrm{rk}(PAB) = \mathrm{rk}(AB)$. לכן די להוכיח $\mathrm{rk}(PAB) \le \mathrm{rk}(PA)$, כלומר, די להוכיח את הטענה המקורית עבור PA במקום עבור A. לכן ניתן להניח ש־A מדורגת קנונית.

יהי $r = \mathrm{rk}(A)$. אז ב־A יש $m-r$ שורות של אפסים. לפי כלל הכפל גם ב־AB יש $m-r$ שורות של אפסים. מכאן שבמטריצה המדורגת קנונית של AB לפחות $m-r$ שורות של אפסים, ולכן $\mathrm{rk}(AB) \le r$.

(ב) לפי (א), $\mathrm{rk}(AB) \le \mathrm{rk}(A)$ וגם $\mathrm{rk}(A) = \mathrm{rk}((AB)B^{-1}) \le \mathrm{rk}(AB)$. לכן יש שוויון. ∎

מסקנה 3.23: *תהי* $P \in \mathrm{M}_m(F)$. *התנאים הבאים שקולים זה לזה:*

(1) P *הפיכה (כלומר יש* $Q \in \mathrm{M}_m(F)$ *כך ש־* $PQ = I_m = QP$ *).*

(2) *יש* $Q \in \mathrm{M}_m(F)$ *כך ש־* $PQ = I_m$.

(3) *יש* $Q \in \mathrm{M}_m(F)$ *כך ש־* $QP = I_m$.

הוכחה: (1) ⇐(2), (3) ־ ברור.

(2) ⇐(1): $m = \mathrm{rk}(I_m) = \mathrm{rk}(PQ) \le \mathrm{rk}(P) \le m$, לכן $\mathrm{rk}(P) = m$. מכאן ש־P הפיכה.

(3) ⇐(1): לפי (2) ⇐(1), Q הפיכה. לכן $P = Q^{-1}(QP) = Q^{-1}I_m = Q^{-1}$, ובפרט P הפיכה.

∎

אלגוריתם 3.24: *מציאת ההופכי של מטריצה* $A \in \mathrm{M}_m(F)$. —על ידי דירוג:

(1) רשום את A, I_m זו ליד זו, כמטריצה מסדר $m \times 2m$, כך: $(A|I_m)$.

(2) בצע פעולות אלמנטריות על מטריצה זו, עד שתגיע למטריצה $(A'|B)$, באשר A' המטריצה המדורגת קנונית של A.

(3) אם $A' \neq I_m$, ל־A אין הופכי.

(4) אם $A' = I_m$, ל־A יש הופכי: $A^{-1} = B$.

הוכחה: יש פעולות אלמנטריות $\mathcal{P}_1, \ldots, \mathcal{P}_k$, כך שמתקיים $A' = \mathcal{P}_k(\cdots(\mathcal{P}_2(\mathcal{P}_1(A)))\cdots)$. לפי הבניה, $B = \mathcal{P}_k(\cdots(\mathcal{P}_2(\mathcal{P}_1(I_m)))\cdots)$. לכן

$$E_{\mathcal{P}_k}\cdots E_{\mathcal{P}_1}A = A' = I_m \qquad E_{\mathcal{P}_k}\cdots E_{\mathcal{P}_1}I_m = B$$

מכאן $BA = I_m$. ∎

הגדרה 3.25: תהי $A \in \mathrm{M}_{m\times n}(F)$. **המטריצה המוחלפת** A^t **של** A היא המטריצה שמתקבלת מ־A על ידי הפיכת השורות של A לעמודות. כלומר, אם

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in \mathrm{M}_{m\times n}(F)$$

אז

$$A^t = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix} \in \mathrm{M}_{n\times m}(F)$$

ובסימון אחר: אם $A \in \mathrm{M}_{m\times n}(F)$ אז $A^t \in \mathrm{M}_{n\times m}(F)$ מוגדרת על ידי $(A^t)_{ij} = (A)_{ji}$ לכל i, j.

למה 3.26: *(א) תהיינה* $A, B \in \mathrm{M}_{m\times n}(F)$. *אז* $(A+B)^t = A^t + B^t$.

(ב) *תהי* $A \in \mathrm{M}_{m\times n}(F)$ *ו־*$\alpha \in F$. *אז* $(\alpha A)^t = \alpha A^t$.

(ג) *תהי* $A \in \mathrm{M}_{m\times n}(F)$. *אז* $(A^t)^t = A$.

(ד) *תהיינה* $B \in \mathrm{M}_{p\times m}(F), A \in \mathrm{M}_{m\times p}(F)$. *אז* $(AB)^t = B^tA^t$.

הוכחה: (ד) תחילה נשים לב ש־ $A^t \in \mathrm{M}_{p\times m}(F), B^t \in \mathrm{M}_{m\times p}(F)$, לכן B^tA^t מוגדרת ומסדר $m \times n$, שהוא הסדר של $(AB)^t$. כעת

$$((AB)^t)_{ij} = (AB)_{ji} = \sum_{k=1}^{p}(A)_{jk}(B)_{ki} = \sum_{k=1}^{p}(A^t)_{kj}(B^t)_{ik} = \sum_{k=1}^{p}(B^t)_{ik}(A^t)_{kj} = (B^tA^t)_{ij}$$

לכל i, j. ∎

תרגיל 3.27: אם $P \in \mathrm{M}_m(F)$ הפיכה אז גם P^t הפיכה ו-$(P^t)^{-1} = (P^{-1})^t$.

הוכחה: מתקיים $PP^{-1} = I_m = P^{-1}P$. מכאן $(P^{-1})^t P^t = I_m^t = P^t(P^{-1})^t$. אך $I_m^t = I_m$. מכאן המסקנה. ∎

משפט 3.28: *תהי* $A \in \mathrm{M}_{m\times n}(F)$. *אז* $\mathrm{rk}(A^t) = \mathrm{rk}(A)$.

הוכחה: די להוכיח כי $\mathrm{rk}(A^t) \le \mathrm{rk}(A)$. אכן, אם זה נכון לכל A, אז גם נכון $\mathrm{rk}(A) = \mathrm{rk}((A^t)^t) \le \mathrm{rk}(A^t)$, ומכאן השוויון.

תהי B המטריצה המדורגת קנונית של A. אז $\mathrm{rk}(B) = \mathrm{rk}(A)$. כמו כן יש $P \in \mathrm{M}_m(F)$ הפיכה כך ש-$A = PB$. לפי למה 3.26(ד) $A^t = B^tP^t$, לכן לפי מסקנה 3.22(ב) $\mathrm{rk}(A^t) = \mathrm{rk}(B^t)$. די, איפוא, להוכיח $\mathrm{rk}(B^t) \le \mathrm{rk}(B)$. (במלים אחרות, די להוכיח $\mathrm{rk}(A^t) \le \mathrm{rk}(A)$ עבור A מדורגת קנונית.)

יהי $r = \mathrm{rk}(B)$. ב-B יש בדיוק r שורות שונות מאפס, לכן ב-B^t יש בדיוק r עמודות שונות מאפס. אם נעשה פעולות אלמנטריות על B^t, עמודות אפסים תשארנה עמודות אפסים. לכן במטריצה המדורגת קנונית של B^t יש לכל היותר r עמודות שונות מאפס, ובפרט לכל היותר r עמודות מובילות. מכאן $\mathrm{rk}(B^t) \le r$. ∎

כפל של 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 $A_{ij} \in \mathrm{M}_{m_i\times n_j}(F)$ לכל i, j. (המספרים משמאל ומעל למטריצה מסמנים את גדלי הגושים) מטריצה שרשומה כך נקראת **מטריצה של גושים**.

משפט 3.29 (מכפלה של מטריצות של גושים): *מכפלה של מטריצות גושים היא מטריצת גושים:*

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

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

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: כיוון שהסימונים מסובכים (בגלל הרבה אינדקסים) נטפל קודם במקרה פרטי של מטריצות בעלות 2×2 גושים. במקרה זה אפשר לנסח את המשפט כך:

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

נוכיח נוסחה זו: תחילה נשים לב שהמכפלות $AP, BR, AQ, BS, CP, DR, CQ, DS$ מוגדרות וכן מוגדרים הסכומים $AP+BR, AQ+BS, CP+DR, CQ+DS$, והמטריצה בשני האגפים מאותו הסדר $(m_1+m_2) \times (n_1+n_2)$. נותר לבדוק שהרכיבים המתאימים בשני האגפים שווים.

למשל, נחשב את הרכיב ה־(m_1+i, j) בשני האגפים, באשר $1 \le i \le m_2,\ 1 \le j \le n_1$.

באגף שמאל רכיב זה מתקבל על ידי הכפלת השורה ה־m_1+i של $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ שהיא

$$\big((C)_{i1}, (C)_{i2}, \ldots, (C)_{i\,p_1},\ (D)_{i1}, (D)_{i2}, \ldots, (D)_{i\,p_2}\big)$$

בעמודה ה־j של $\begin{pmatrix} P & Q \\ R & S \end{pmatrix}$ שהיא

$$\big((P)_{1j}, (P)_{2j}, \ldots, (P)_{p_1\,j},\ (R)_{1j}, (R)_{2j}, \ldots, (R)_{p_2\,j}\big)^t.$$

מכפלתן היא

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

וזה בדיוק הרכיב ה־(m_1+i, j) באגף ימין. בזאת הסתיימה ההוכחה של המקרה הפרטי.

המקרה הכללי דומה: תחילה נשים לב שהמכפלות $A_{ik}B_{kj}$ מוגדרות וכן מוגדרים הסכומים $\sum_{k=1}^{s} A_{ik}B_{kj}$ והמטריצות בשני האגפים מאותו הסדר $(m_1+m_2+\ldots+m_r) \times (n_1+n_2+\ldots+n_t)$. נותר לבדוק שהרכיבים המתאימים בשני האגפים שווים.

למשל, נחשב את הרכיב ה־(m_1+i,j) בשני האגפים, באשר $1 \le j \le n_1,\ 1 \le i \le m_2$. באגף שמאל רכיב זה מתקבל על ידי הכפלת השורה ה־m_1+i של המטריצה $(\ A_{ij}\ )$ שהיא

$$\big((A_{21})_{i1},(A_{21})_{i2},\dots,(A_{21})_{i\,p_1},\ \ \dots\ \ ,(A_{2s})_{i1},(A_{2s})_{i2},\dots,(A_{2s})_{i\,p_s}\big)$$

בעמודה ה־j של המטריצה $(\ B_{ij}\ )$ שהיא

$$\big((B_{11})_{1j},(B_{11})_{2j},\dots,(B_{11})_{p_1\,j},\ \ \dots\ \ ,(B_{s1})_{1j},(B_{s1})_{2j},\dots,(B_{s1})_{p_s\,j}\big)^t.$$

מכפלתן היא

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

וזה בדיוק הרכיב ה־(m_1+i,j) של המטריצה $(\ C_{ij}\ )$. ∎

מסקנה 3.30: *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 3.31: *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 3.32: *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*

$$(P)_{i1}A_1+(P)_{i2}A_2+\cdots+(P)_{im}A_m=\sum_{k=1}^{m}(P)_{ik}A_k,\qquad i=1,2,\dots,m.$$

הוכחה: לפי המשפט, $PA = \begin{pmatrix} (P)_{11} & \dots & (P)_{1m} \\ \dots & \dots & \dots \\ (P)_{m1} & \dots & (P)_{mm} \end{pmatrix} \begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^m (P)_{1k} A_k \\ \vdots \\ \sum_{k=1}^m (P)_{mk} A_k \end{pmatrix}$.

הוכחה נוספת: לכל $1 \leq j \leq n$ הרכיב ה־j של A_k הוא $(A)_{kj}$. לכן הרכיב ה־j של $\sum_{k=1}^m (P)_{ik} A_k$ הוא $\sum_{k=1}^m (P)_{ik}(A)_{kj}$, שהינו $(PA)_{ij}$, לפי הגדרה 3.4. זהו הרכיב ה־j של השורה ה־i של PA. ∎

תרגיל 3.33: *הוכח כי מטריצת הגושים* $M = \begin{array}{cc} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \end{array}$ *הפיכה אם ורק אם* A, D *הפיכות ואם זה קורה אז*

$N = \begin{array}{cc} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{pmatrix} A^{-1} & -A^{-1}BD^{-1} \\ 0 & D^{-1} \end{pmatrix} \end{array}$ *ההופכי שלה.*

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

הוכחה: אם M הפיכה אז $M^{-1} = \begin{array}{cc} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{pmatrix} P & Q \\ R & S \end{pmatrix} \end{array}$ ומתקיים $\begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \begin{pmatrix} P & Q \\ R & S \end{pmatrix} = I_{m+n}$, כלומר, לפי המשפט,

$$\begin{pmatrix} AP + BR & AQ + BS \\ DR & DS \end{pmatrix} = \begin{pmatrix} I_m & 0 \\ 0 & I_n \end{pmatrix}.$$

מכאן

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

מתוך $DS = I_n$ נסיק ש־D הפיכה ו־$S = D^{-1}$. מכאן, מתוך $DR = 0$, נסיק ש־$R = D^{-1}(DR) = 0$. לכן $AP + BR = I_m$ הופך להיות ל־$AP = I_m$ ומכאן נסיק ש־A הפיכה ו־$P = A^{-1}$. לבסוף, מתוך $AQ + BS = 0$ נסיק כי $Q = -A^{-1}BS = -A^{-1}BD^{-1}$. בסיכום: A, D הפיכות ו־$M^{-1} = N$.

להיפך, אם A, D מטריצות הפיכות, אז לפי המשפט, $\begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \begin{pmatrix} A^{-1} & -A^{-1}BD^{-1} \\ 0 & D^{-1} \end{pmatrix} = \begin{pmatrix} I_m & 0 \\ 0 & I_n \end{pmatrix} = I_{m+n}$, לכן M הפיכה ו־$M^{-1} = N$. ∎

תרגיל 3.34: *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, $M_0 = \begin{array}{c} \\ m \\ n \end{array}\!\!\begin{array}{c} \begin{array}{cc} n & m \end{array} \\ \begin{pmatrix} 0 & B \\ C & 0 \end{pmatrix} \end{array}$ *הפיכה אם ורק אם* B, C *הפיכות ואז* $M_0^{-1} = \begin{array}{c} \\ n \\ m \end{array}\!\!\begin{array}{c} \begin{array}{cc} m & n \end{array} \\ \begin{pmatrix} 0 & C^{-1} \\ B^{-1} & 0 \end{pmatrix} \end{array}$.

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

הוכחה א': לפי תרגיל 3.34, $\begin{array}{c} \\ m \\ n \end{array}\!\!\begin{array}{c} \begin{array}{cc} n & m \end{array} \\ \begin{pmatrix} 0 & I_m \\ I_n & 0 \end{pmatrix} \end{array}$ הפיכה ו־$\begin{array}{c} \\ n \\ m \end{array}\!\!\begin{array}{c} \begin{array}{cc} m & n \end{array} \\ \begin{pmatrix} 0 & I_n \\ I_m & 0 \end{pmatrix} \end{array}$ ההופכי שלה. לפי המשפט

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

הוכחה ב': (הוכחה זו מסתמכת על החומר על העתקות לינאריות שנביא בהמשך.) נבחר מרחב וקטורי V ממימד $m+n$ ונבחר בסיס סדור שלו $\mathcal{B} = (v_1, \ldots, v_m, w_1, \ldots, w_n)$. לפי משפט 6.27 קיימת העתקה לינארית (יחידה) $T\colon V \to V$ כך ש־$[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}$.

הסדרה $\mathcal{B}' = (w_1, \ldots, w_n, v_1, \ldots, v_m)$ גם היא בסיס של V. קל לראות ש־$[T]_{\mathcal{B}'}^{\mathcal{B}'} = \begin{pmatrix} D & 0 \\ 0 & A \end{pmatrix}$.

לכן לפי משפט 6.44 $\begin{pmatrix} D & 0 \\ 0 & A \end{pmatrix}, \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}$ דומות. ∎

4. מרחבים וקטוריים

מרחבים ותת־מרחבים.

יהי F שדה.

הגדרה 4.3: **מרחב וקטורי** מעל שדה F היא קבוצה V, עליה מוגדרות שתי פעולות, חיבור $(+)$ וכפל (ללא סימון) של איברי V (שייקראו וקטורים) בסקלרים, אשר מקיימות את החוקים הבאים:

(ח0) קשירות החיבור:

לכל $u, v \in V$ קיים $w \in V$ יחיד כך ש־$u + v = w$.

(ח1) אסוציאטיביות החיבור:

$(u + v) + w = u + (v + w)$, לכל $u, v, w \in V$.

(ח2) קיום איבר האפס:

קיים איבר ב־V, המסומן 0, המקיים $v + 0 = v = 0 + v$, לכל $v \in V$.

(ח3) קיום איבר נגדי:

לכל $v \in V$ קיים איבר $-v \in V$ המקיים $v + (-v) = 0 = (-v) + v$.

(ח4) קומוטטיביות החיבור:

$u + v = v + u$, לכל $u, v \in V$.

(כ0) קשירות הכפל:

לכל $v \in V$ ולכל $a \in F$ קיים $w \in V$ יחיד כך ש־$av = w$.

(כ1) חוק פילוג:

$a(u + v) = au + av$, לכל $u, v \in V$ ולכל $a \in F$.

(כ2) חוק פילוג:

$(a + b)v = av + bv$, לכל $v \in V$ ולכל $a, b \in F$.

(כ3) אסוציאטיביות הכפל בסקלר:

$(ab)v = a(bv)$, לכל $v \in V$ ולכל $a, b \in F$.

(כ4) $1v = v$ לכל $v \in V$. (כאן $1 \in F$ איבר היחידה של F.)

חשוב להדגיש: מרחב וקטורי אינו, לפי ההגדרה, אוסף של n־יות, ולכן אין משמעות לשאלה "מהו המימד של V?". (לשאלה זו תהיה משמעות מאוחר יותר, אך משמעות אחרת.) זוהי קבוצה כלשהי שעליה הוגדרה פעולת החיבור ופעולת הכפל בסקלר, כך שמתקיימות האקסיומות לעיל. לכאורה אפשר למצוא דוגמאות מאד מוזרות למרחבים וקטורים; אך הדבר אינו קל, בגלל שצריך לספק אקסיומות רבות.

דוגמאות 4.4:

(א) $\mathrm{M}_{m\times n}(F)$ הוא מרחב וקטורי מעל F.

(ב) (מקרה פרטי של (א)): $F^m = \mathrm{M}_{m\times 1}(F)$.

(ג) (מקרה פרטי של (ב)): $F = F^1$.

(ד) (יהי $F = \mathbb{R}$.) אוסף כל החצים מהראשית (או: כל הנקודות) במרחב התלת ממדי.

(ה) אוסף $F[X]$ של כל הפולינומים עם מקדמים ב־F.

(ו) (יהי $F = \mathbb{R}$.) אוסף כל הפונקציות הממשיות הרציפות על הקטע $[0,1]$.

(ז) אוסף כל המשוואות הלינאריות ב־n נעלמים מעל F.

(ח) $\{0\}$.

למה 4.5: *יהי V מרחב וקטורי מעל שדה F.*

(א) *איבר האפס של V הוא יחיד.*

(ב) *נגדי של איבר הוא יחיד.*

(ג) *כלל הצמצום: יהיו $v,u,u' \in V$. אם $v+u = v+u'$ אז $u = u'$.*

(ד) $0v = 0$ *לכל* $v \in V$

(ה) $a0 = 0$ *לכל* $a \in F$.

(ו) *אם* $av = 0$ *אז* $v = 0$ *או* $a = 0$.

(ז) $(-a)v = a(-v) = -(av)$.

הוכחה: (א) יהיו $0, 0'$ שני איברי אפס. אז $0 = 0+0' = 0'$. (השוויון השמאלי הוא בגלל ש־$0'$ הוא איבר האפס; השוויון הימני הוא בגלל ש־0 הוא איבר האפס.)

(ב) יהיו v', v'' שני נגדיים של $v \in V$. אז

$$v' = v' + 0 = v' + (v + v'') = (v' + v) + v'' = 0 + v'' = v''.$$

(ג) $u = 0 + u = ((-v)+v)+u = (-v)+(v+u)$ ובאותו אופן $u' = (-v)+(v+u')$. לפי הנתון, $u = u'$.

(ד) $0v + 0v = (0+0)v = 0v = 0v + 0$. לפי (ג), $0v = 0$.

(ה) $a0 + a0 = a(0+0) = a0 = a0 + 0$. לפי (ג), $a0 = 0$.

(ו) נניח $a \neq 0$. (צריך להוכיח כי $v = 0$.) אז $v = 1v = (a^{-1}a)v = a^{-1}(av) = a^{-1}0 = 0$.

(ז) מתקיים $av + (-a)v = -(a+a)v = 0v = 0$, לכן $(-a)v$ הוא הנגדי של av. כמו כן $av + a(-v) = a\big(v+(-v)\big) = a0 = 0$, לכן $a(-v)$ הוא הנגדי של av. ∎

הגדרה 4.6: קבוצה חלקית U של מרחב וקטורי V (מעל שדה F) נקראת **תת־מרחב** של V אם

(א) $0 \in U$;

(ב) U סגורה תחת החיבור, כלומר: אם $u_1, u_2 \in U$ אז גם $u_1 + u_2 \in U$;

(ג) U סגורה תחת הכפל בסקלר, כלומר: אם $u \in U$ ו־$a \in F$ אז גם $au \in U$.

משפט 4.7: *יהי U תת־מרחב של מרחב וקטורי V מעל שדה F. אז U הוא בעצמו מרחב וקטורי מעל F, וזאת ביחס לחיבור ולכפל בסקלר שמושרים מ־ V.*

הוכחה: נבהיר תחילה למה הכוונה בפעולה מושרית: אם $u_1, u_2 \in U$, אז החיבור ביניהם מוגדר כחיבור ב־V. אם $u \in U$, $a \in F$ אז הכפל au מוגדר ככפל בסקלר ב־V. הקשירות של הפעולת החיבור נובעת מתנאי (ב); הקשירות של הכפל בסקלר נובעת מתנאי (ג).

קיום איבר האפס נובע מתנאי (ג).

קיום איבר נגדי: יהי $u \in U$. אז יש לו נגדי $-u \in V$ (אך מראש לא ברור ש־$u \in U$). אך $-u = (-1)u \in U$, לפי תנאי (ג).

יתר הדרישות של הגדרה 4.3 מתקיימות ב־U כי הן מתקיימות ב־V (כל התכונות שמתקיימות **לכל** איברי V, מתקיימות בפרט לכל איברי U; הבעיה היתה רק עם אותן התכונות שדורשות **קיום** של איבר ב־U, כי מתוך כך שקיים איבר כזה ב־V אי אפשר להסיק שהוא קיים גם ב־U.) ∎

דוגמאות 4.8: דוגמאות לתת־מרחבים.

(א) $\{0\}$, V הם תת־מרחבים של מרחב וקטורי V.

(ב) אוסף כל הפתרונות של מערכת משוואות לינאריות הומוגנית $A\mathbf{X} = 0$ ב־n נעלמים מעל שדה F הוא תת מרחב של F^n. נקרא **מרחב הפתרונות של** $A\mathbf{X} = 0$ או גם **הגרעין של** A; יסומן $\operatorname{Ker} A$. אם כן, $\operatorname{Ker} A = \{v \in F^n |\ Av = 0\}$.

(ג) אוסף כל הפולינומים עם מקדם חפשי 0 הוא תת מרחב של $F[X]$.

(ד) אוסף כל הפולינומים ממעלה $\geq n$ הוא תת מרחב של $F[X]$.

(ה) אם U_1, U_2 שניהם תת מרחבים של V אז גם החיתוך שלהם $U_1 \cap U_2$ תת־מרחב של V.

צירופים לינאריים.

הגדרה 4.9: יהי V מרחב וקטורי מעל שדה F ויהיו $v_1, v_2, \ldots, v_k \in V$. וקטור $v \in V$ ייקרא **צירוף לינארי של** $v_1, v_2, \ldots, v_k$ אם יש $a_1, a_2, \ldots, a_k \in F$ כך ש־$v = a_1v_1 + a_2v_2 + \cdots + a_kv_k$.

אוסף כל הצירופים הלינאריים של $v_1, v_2, \ldots, v_k$ יסומן $\operatorname{Sp}(v_1, v_2, \ldots, v_k)$. כלומר,

$$\operatorname{Sp}(v_1, v_2, \ldots, v_k) = \{a_1v_1 + a_2v_2 + \cdots + a_kv_k |\ a_1, a_2, \ldots, a_k \in F\}..$$

דוגמה 4.10: אם V הוא המרחב התלת ממדי, אז $(1, 2, 3)$ הוא צירוף לינארי של $(1, 1, 1), (0, 1, 2)$, אך $(1, 1, 0)$ אינו צירוף לינארי שלהם.

דוגמאות 4.11: (א) אם $v \neq 0$, אז $\operatorname{Sp}(v) = \{av |\ a \in F\}$ מורכב מכל הכפולות של v. בפרט, אם v חץ מהראשית במרחב התלת ממדי, אז $\operatorname{Sp}(v)$ הוא הישר שמונח על חץ זה.

(ב) יהיו u, v שני חצים מהראשית במרחב התלת ממדי, שניהם שונים מ־0 ואינם על אותו ישר דרך הראשית. אז $\mathrm{Sp}(u,v)$ הוא המישור דרך שני החצים (שעובר דרך הראשית).

(ג) להגדרה לעיל יש גם משמעות במקרה $k=0$. אז $\mathrm{Sp}(\emptyset)=\{0\}$.

משפט 4.12: *יהי V מרחב וקטורי מעל שדה F ויהיו $v_1, v_2, \ldots, v_k \in V$. נסמן $S = \mathrm{Sp}(v_1, v_2, \ldots, v_k)$. אז S הוא תת־המרחב הקטן ביותר שמכיל את $v_1, v_2, \ldots, v_k$. כלומר,*

(א) *S הוא תת־המרחב של V;*

(ב) *$v_1, v_2, \ldots, v_k \in S$;*

(ג) *אם U תת־מרחב של V ו־ $v_1, v_2, \ldots, v_k \in U$ אז $S \subseteq U$.*

הוכחה: (א)

(1) $0 = 0v_1 + 0v_2 + \cdots + 0v_k \in S$

(2) אם $v = a_1v_1 + a_2v_2 + \cdots + a_kv_k \in S$ ו־$v' = b_1v_1 + b_2v_2 + \cdots + b_kv_k \in S$ אז $v + v' = (a_1+b_1)v_1 + (a_2+b_2)v_2 + \cdots + (a_k+b_k)v_k \in S$.

(3) אם $v = a_1v_1+a_2v_2+\cdots+a_kv_k \in S$ ו־$\alpha \in F$ אז $\alpha v = (\alpha a_1)v_1+(\alpha a_2)v_2+\cdots+(\alpha a_k)v_k \in S$.

לכן S תת־מרחב.

(ב) $v_j = 0v_1 + \cdots + 0v_{j-1} + 1v_j + 0v_{j+1} \cdots + 1v_k \in S$.

(ג) צריך להוכיח שלכל $v \in S$ מתקיים $v \in U$.

יהי $v \in S$. אז יש $a_1, a_2, \ldots, a_k \in F$ כך ש־$v = a_1v_1 + a_2v_2 + \cdots + a_kv_k$. כיוון ש־$U$ תת־מרחב של V ו־$v_j \in U$ אז גם $a_jv_j \in U$, לכל $1 \le j \le k$. שוב, כיוון ש־U תת־מרחב של V ו־$a_jv_j \in U$ לכל j, אז גם $v = a_1v_1 + a_2v_2 + \cdots + a_kv_k \in U$. ∎

תרגיל 4.13: אם $u_1, \ldots, u_r$ תת סדרה של $v_1, \ldots, v_k$ אז $\mathrm{Sp}(u_1, \ldots, u_r) \subseteq \mathrm{Sp}(v_1, v_2, \ldots, v_k)$.

הוכחה: אגף ימין הוא תת מרחב שמכיל את $v_1, v_2, \ldots, v_k$, ובפרט את $u_1, \ldots, u_r$. לכן לפי המשפט הוא מכיל את אגף שמאל. ∎

תרגיל 4.14: יהי V מרחב וקטורי מעל F. נניח ש־$w_1, \ldots, w_m$ הם צירופים לינאריים של $v_1, v_2, \ldots, v_n$. אז $\mathrm{Sp}(v_1, \ldots, v_n) = \mathrm{Sp}(v_1, \ldots, v_n, w_1, \ldots, w_m)$.

בפרט, $\mathrm{Sp}(v_1, \ldots, v_n) = \mathrm{Sp}(v_1, \ldots, v_n, 0, \ldots, 0)$.

הוכחה: ההכלה "$\subseteq$" נובעת מהתרגיל הקודם. אגף שמאל הוא תת מרחב שמכיל את $v_1, v_2, \ldots, v_n$; לפי ההנחה הוא מכיל את נניח ש־$w_1, \ldots, w_m$. לכן, לפי המשפט, הוא מכיל את אגף ימין. ∎

אין חשיבות לסדר של $v_1, v_2, \ldots, v_k$.

תרגיל 4.15: נתונים $v, v_1, \ldots, v_n \in F^m$. קבע האם v צירוף לינארי של $v_1, \ldots, v_n$ ואם כן, מצא את המקדמים $c_1, \ldots, c_n \in F$ כך ש־$v = c_1v_1 + \cdots + c_nv_n$.

פתרון (תיאורטי): תהי A המטריצה ש־$v_1, \ldots, v_n$ עמודותיה. לפי תרגיל 3.29, $v = c_1v_1 + \cdots + c_nv_n$ אם ורק אם $c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$ פתרון של המערכת משוואות לינאריות $AX = v$, כלומר, המערכת שמטריצת המקדמים המורחבת שלה היא $(v_1, v_2, \ldots, v_n, v)$.

בפרט, v צירוף לינארי של $v_1, \ldots, v_n$ אם ורק אם למערכת זו יש פתרון, כלומר, אם ורק אם

$$\operatorname{rk}\big((v_1, \ldots, v_n)\big) = \operatorname{rk}((v_1, \ldots, v_n, v)). \qquad \blacksquare$$

דוגמה 4.16: מספרית (חסרה עדיין).

מסקנה 4.17: *יהיו $v, v_1, \ldots, v_n \in F^m$, $v', v'_1, \ldots, v'_n$, ויהיו $c_1, \ldots, c_n \in F$. נניח ששתי המטריצות הבאות*

$$(v_1, v_2, \ldots, v_n, v),\ (v'_1, v'_2, \ldots, v'_n, v') \in \mathrm{M}_{m\times(n+1)}(F)$$

שקולות שורות (אחת מתקבלת מהשניה על ידי סדרה של פעולות אלמנטריות).

(א) *יהיו $c_1, \ldots, c_n \in F$. אז $v = c_1v_1 + \cdots + c_nv_n$ אם ורק אם $v' = c_1v'_1 + \cdots + c_nv'_n$.*

(ב) *v צירוף לינארי של $v_1, \ldots, v_n$ אם ורק אם v' צירוף לינארי של $v'_1, \ldots, v'_n$.*

הוכחה: (א) נובע מהתרגיל לעיל, כי למערכות שקולות אותם הפתרונות. (ב) נובע מ־(א). $\blacksquare$

תרגיל 4.18: תהי $A \in \mathrm{M}_{m\times n}(F)$ (כך שעמודותיה הן איברי F^m). העמודה ה־j של A היא צירוף לינארי של קודמותיה (ז.א., העמודות שמשמאלה) אם ורק אם העמודה ה־j של המטריצה המדורגת קנונית A' של A אינה מובילה.

הוכחה: נוכל להניח שהעמודה ה־j של A היא העמודה האחרונה, אחרת נשמיט את העמודות שאחריה. (אם נסמן ב־$B, B' \in \mathrm{M}_{m\times j}(F)$ את המטריצות המתקבלות מ־A, A' על ידי ההשמטה, אז B' שקולה ל־B. ראינו ש־B' מדורגת קנונית. לכן היא המטריצה המדורגת קנונית של B.) אז A היא מטריצת מקדמים המורחבת של איזושהי מערכת משוואות לינאריות. לפי התרגיל הקודם, העמודה האחרונה היא צירוף לינארי של האחרות אם ורק אם למערכת זו יש פתרון. זה קורה, כידוע, אם ורק אם העמודה האחרונה של A' אינה מובילה. $\blacksquare$

הגדרה 4.19: יהי V מרחב וקטורי ותהי $v_1, v_2, \ldots, v_n$ סדרה של איברים ב־V. סדרה זו תיקרא **סדרת יוצרים של** V אם $\mathrm{Sp}(v_1, v_2, \ldots, v_n) = V$, כלומר, כל $v \in V$ הוא צירוף לינארי של $v_1, v_2, \ldots, v_n$.

תרגיל 4.20: $e_1, e_2, \ldots, e_m$ סדרת יוצרים של F^m, כי $\mathrm{Sp}(e_1, e_2, \ldots, e_m) = F^m$.

מסקנה 4.21: *סדרה* $v_1,\ldots,v_n\in F^m$ *היא סדרת יוצרים של* F^m *אם ורק אם* $m=\operatorname{rk}\left((v_1,\ldots,v_n)\right)$.

הוכחה: נסמן $A=(v_1,\ldots,v_n)\in \mathrm{M}_{m\times n}(F)$.

נניח $\operatorname{rk}(A)=m$ ויהי $v\in F^m$. אז $m\le \operatorname{rk}\left((A,v)\right)\le \operatorname{rk}(A)=m$ לכן $\operatorname{rk}(A)=\operatorname{rk}\left((A,v)\right)$. לפי תרגיל 4.15, v צירוף לינארי של $v_1,\ldots,v_n$.

נניח $\operatorname{rk}(A)<m$. תהי A' המטריצה המדורגת קנונית של A. אז השורה האחרונה של A' היא שורת אפסים. לכן למערכת, שמטריצת מקדמים המורחבת שלה היא (A',e_m), אין פתרון. אפשר להגיע מ־(A',e_m) על ידי פעולות אלמנטריות למטריצה (A,v), עבור איזה $v\in F^m$. אז גם למערכת, שמטריצת מקדמים המורחבת שלה היא (A,v), אין פתרון. לפי התרגיל, v אינו צירוף לינארי של $v_1,\ldots,v_n$. ∎

הגדרה 4.22: **מרחב השורות של** $\operatorname{Row}(A)$ $A\in \mathrm{M}_{m\times n}(F)$ הוא תת־המרחב של $F^n=\mathrm{M}_{1\times n}(F)$ הנפרש על ידי השורות של A. **מרחב העמודות של** $\operatorname{Col}(A)$ $A\in \mathrm{M}_{m\times n}(F)$ הוא תת־המרחב של $F^m=\mathrm{M}_{m\times 1}(F)$ הנפרש על ידי העמודות של A.

משפט 4.23: *תהיינה* $A,B\in \mathrm{M}_{m\times n}(F)$. *תהיינה* $A',B'\in \mathrm{M}_{m\times n}(F)$ *המטריצות המדורגות קנונית שלהן, בהתאמה. אז* $\operatorname{Row}(A)=\operatorname{Row}(B)$ *אם ורק אם* $A'=B'$.

הוכחה: נניח $A'=B'$. אז A,B שקולות שורות, ולכן יש $P\in \mathrm{M}_m(F)$ הפיכה כך ש־$B=PA$. לפי תרגיל 3.32 השורות של B הן

$$B_i=(P)_{i1}A_1+(P)_{i2}A_2+\cdots+(P)_{im}A_m\in \operatorname{Row}(A),\qquad i=1,2,\ldots,m$$

לכן $\operatorname{Row}(B)=\operatorname{Sp}(B_1,\ldots,B_m)\subseteq \operatorname{Row}(A)$. באופן דומה נסיק מתוך $A=P^{-1}B$ שמתקיים $\operatorname{Row}(A)\subseteq \operatorname{Row}(B)$. מכאן $\operatorname{Row}(A)=\operatorname{Row}(B)$.

להיפך, נניח $\operatorname{Row}(A)=\operatorname{Row}(B)$. כיוון ש־$A,A'$ שקולות שורות, לפי הפסקה הראשונה של ההוכחה $\operatorname{Row}(A)=\operatorname{Row}(A')$. באותו אופן $\operatorname{Row}(B)=\operatorname{Row}(B')$. לכן $\operatorname{Row}(A')=\operatorname{Row}(B')$.

בלי הגבלת הכלליות $r:=\operatorname{rk}(B')\ge \operatorname{rk}(A')$. אז $m-r$ השורות האחרונות ב־A',B' הן שורות אפסים. נסמן ב־A'',B'' את המטריצות מסדר $r\times n$ המתקבלות מ־A',B' על ידי השמטת שורות אלה. אז $\operatorname{Row}(B'')=\operatorname{Row}(B')$, $\operatorname{Row}(A'')=\operatorname{Row}(A')$, ולכן $\operatorname{Row}(A'')=\operatorname{Row}(B'')$ ודי להוכיח כי $A''=B''$.

תהיינה $A_1,\ldots,A_r$ השורות של A'' ו־$B_1,\ldots,B_r$ השורות של B''. כיוון ש־$B_i\in \operatorname{Sp}(A_1,\ldots,A_r)$, לכל i, יש $c_{ij}\in F$ כך שמתקיים $B_i=c_{i1}A_1+\cdots+c_{ir}A_r$. נגדיר $P\in \mathrm{M}_r(F)$ על ידי $(P)_{ij}=c_{ij}$. לפי תרגיל 3.32, $B''=PA''$. לפי מסקנה 3.22, $r=\operatorname{rk}(B'')\le \operatorname{rk}(P)\le r$, לכן $\operatorname{rk}(P)=r$, כלומר, P הפיכה. זה אומר ש־A'',B'' שקולות שורות. כיוון שהן מדורגות קנונית, הן שוות (לפי משפט 2.23). ∎

אי תלות לינארית.

הגדרה 4.24: יהי V מרחב וקטורי מעל שדה F. סדרה $v_1, v_2, \dots, v_n \in V$ הנה **תלויה לינארית** (**מעל** F) אם קיימים $a_1, a_2, \dots, a_n \in F$, לא כולם 0, כך ש־$a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0$.

לפיכך, $v_1, v_2, \dots, v_n \in V$ **בלתי תלויה לינארית** (**מעל** F) אם לכל $a_1, a_2, \dots, a_n \in F$ מתקיים: אם $a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0$ אז $a_1 = a_2 = \cdots = a_n = 0$.

דוגמה 4.25:

(א) $\begin{pmatrix} 3 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix} \in F^2$ היא תלויה לינארית, כי $0 = (-2)\begin{pmatrix} 3 \\ 3 \end{pmatrix} + 3\begin{pmatrix} 2 \\ 2 \end{pmatrix}$.

(ב) $e_1, e_2, \dots, e_m \in F^m$ היא בלתי תלויה לינארית.

אכן, יהיו $a_1, a_2, \dots, a_m \in F$ כך ש־$a_1e_1 + a_2e_2 + \cdots + a_me_m = 0$. אגף שמאל הוא $(a_1, a_2, \dots, a_m)^t$. לכן $a_1 = a_2 = \cdots = a_m = 0$.

הערה 4.26: סדרה ריקה $\emptyset$ נחשבת לבלתי תלויה לינארית.

תרגיל 4.27: יהי V מרחב וקטורי מעל שדה F.

(א) הסדרה v_1 (בת איבר אחד) תלויה לינארית אם ורק אם $v_1 = 0$.

(ב) הסדרה v_1, v_2 (בת שני איברים) תלויה לינארית אם ורק אם v_1 כפולה של v_2 או v_2 כפולה של v_1.

(ג) אם יש j כך ש־$v_j = 0$ אז $v_1, v_2, \dots, v_n \in V$ תלויה לינארית.

משפט 4.28: *יהי V מרחב וקטורי מעל שדה F. סדרה $v_1, v_2, \dots, v_n \in V$ תלויה לינארית אם ורק אם יש $1 \le i \le n$ כך ש־$v_i \in \mathrm{Sp}(v_1, v_2, \dots, v_{i-1})$.* (עבור $i = 1$ תנאי זה אומר $v_1 \in \mathrm{Sp}(\emptyset) = \{0\}$, כלומר, $v_1 = 0$.)

הוכחה: נניח $v_i \in \mathrm{Sp}(v_1, v_2, \dots, v_{i-1})$. אז יש $a_1, \dots, a_{i-1} \in F$ כך ש־$v_i = a_1v_1 + \cdots + a_{i-1}v_{i-1}$. מכאן $(-a_1)v_1 + \cdots (-a_{i-1})v_{i-1} + 1v_i = 0$. היות ו־$1 \neq 0$, הסדרה $v_1, v_2, \dots, v_n \in V$ תלויה לינארית.

להיפך, נניח כי $v_1, v_2, \dots, v_n \in V$ תלויה לינארית. אז קיימים $a_1, a_2, \dots, a_n \in F$, לא כולם 0, כך ש־$a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0$. יהי i הגדול ביותר עבורו $a_i \neq 0$. אז מתקיים $a_iv_i = (-a_1)v_1 + (-a_2)v_2 + \cdots + (-a_{i-1})v_{i-1}$ ומכאן

$$v_i = (-a_i^{-1})a_1v_1 + (-a_i^{-1})a_2v_2 + \cdots + (-a_i^{-1})a_{i-1}v_{i-1} \in \mathrm{Sp}(v_1, v_2, \dots, v_{i-1}). \quad \blacksquare$$

מסקנה 4.29: *סדרה $v_1, v_2, \dots, v_n \in F^m$ בלתי תלויה לינארית אם ורק אם $\mathrm{rk}\left((v_1, v_2, \dots, v_n)\right) = n$.*

הוכחה: תהי $(v'_1, v'_2, \dots, v'_n)$ המטריצה המדורגת קנונית של $(v_1, v_2, \dots, v_n)$. לפי המשפט, $v_1, v_2, \dots, v_n$ בלתי תלויה לינארית אם ורק אם $v_i \notin \mathrm{Sp}(v_1, v_2, \dots, v_{i-1})$ לכל i. לפי תרגיל 4.18 זה שקול לכך ש־v'_i עמודה מובילה, לכל i. זה אומר שכל העמודות מובילות, כלומר, $\mathrm{rk}\left((v_1, v_2, \dots, v_n)\right) = n$. $\blacksquare$

תרגיל 4.30: יהי V מרחב וקטורי מעל שדה F. תהי $v_1, v_2, \dots, v_n \in V$ סדרה ותהי $w_1, w_2, \dots, w_k$ תת סדרה שלה. אם $w_1, w_2, \dots, w_k$ תלויה לינארית אז $v_1, v_2, \dots, v_n$ תלויה לינארית. כלומר, אם $v_1, v_2, \dots, v_n$ בלתי תלויה לינארית אז $w_1, w_2, \dots, w_k$ בלתי תלויה לינארית.

הוכחה: יהיו $a_1, \ldots, a_k \in F$ כך ש־$a_1w_1 + \cdots + a_kw_k = 0$. נשלים את אגף שמאל לצירוף לינארי של $v_1, \ldots, v_n$ על ידי הוספת מחוברים מהצורה $0v_j$, באשר v_j אינו בתת הסדרה $w_1, w_2, \ldots, w_k$. לפי ההנחה כל המקדמים באגף שמאל הם 0, בפרט $a_1, \ldots, a_k$. ∎

תרגיל 4.31: תהי $A \in \mathrm{M}_{m\times n}(F)$ מדורגת קנונית. אז השורות השונות מאפס של A בלתי תלויות לינארית (כאיברי $F^n = \mathrm{M}_{1\times n}(F)$).

הוכחה: תהיינה $A_1, A_2, \ldots, A_r$ השורות השונות מאפס של A ויהיו $a_1, a_2, \ldots, a_r \in F$ סקלרים כך ש־$a_1A_1 + a_2A_2 + \cdots + a_rA_r = 0$. (שני האגפים הם שורות מאורך n.) יהי $1 \le j \le n$. הרכיב ה־j בשני האגפים שווה: $a_1(A)_{1j} + a_2(A)_{2j} + \cdots + a_r(A)_{rj} = 0$. בפרט זה נכון, אם j הוא מספר העמודה של העמודה המבילה ה־i, באשר $1 \le i \le r$. אבל אז $(A)_{kj} = 0$ לכל k שונה מ־i ואילו $(A)_{ij} = 1$. לכן $a_i = 0$. ∎

בסיסים וממד.

משפט 4.32 (משפט ההחלפה): *יהי V מרחב וקטורי מעל שדה F. תהי $v_1, v_2, \ldots, v_n \in V$ סדרת יוצרים של V, ותהי $w_1, w_2, \ldots, w_m \in V$ סדרה בלתי תלויה לינארית. אז $m \le n$ וניתן להחליף m איברים מתוך $v_1, v_2, \ldots, v_n$ ב־$w_1, w_2, \ldots, w_m$ כך שהסדרה המתקבלת (בת n איברים) היא סדרת יוצרים של V.*

הוכחה: באינדוקציה על m.

אם $m = 0$, אין מה להוכיח.

נניח $m \ge 1$ ונניח שהמשפט נכון עבור $m - 1$.

נתון: $w_1, w_2, \ldots, w_m$ בלתי תלויה לינארית. לכן לפי תרגיל 4.30 גם $w_1, w_2, \ldots, w_{m-1}$ בלתי תלויה לינארית. לפי הנחת האינדוקציה $m - 1 \le n$ ובלי הגבלת הכלליות (אחרי סידור מחדש) $w_1, w_2, \ldots, w_{m-1}, v_m, \ldots, v_n$ היא סדרת יוצרים של V. (שים לב: אם $m - 1 = n$, אז האיבר האחרון בסדרה זו הוא w_{m-1}.) אז גם $w_1, w_2, \ldots, w_{m-1}, v_m, \ldots, v_n, w_m$ היא סדרת יוצרים של V. היא תלויה לינארית לפי משפט 4.28, כי

$$w_m \in V = \mathrm{Sp}(w_1, w_2, \ldots, w_{m-1}, v_m, \ldots, v_n),$$

כלומר, w_m צירוף לינארי של קודמיו. לכן גם $w_1, w_2, \ldots, w_{m-1}, w_m, v_m, \ldots, v_n$ סדרת יוצרים תלויה לינארית של V. לכן (שוב לפי משפט 4.28) יש בה איבר שהוא צירוף לינארי של קודמיו. איבר זה איננו אחד מתוך $w_1, w_2, \ldots, w_m$, כי היא בלתי תלויה לינארית. לכן זה אחד מבין $v_m, \ldots, v_n$. בפרט (כיוון שיש איבר כזה) $m \le n$. האיבר הנ"ל הוא גם צירוף לינארי של כל יתר איברי הסדרה $w_1, w_2, \ldots, w_{m-1}, w_m, v_m, \ldots, v_n$. בלי הגבלת הכלליות איבר זה הוא v_m. לפי תרגיל 4.14,

$$\mathrm{Sp}(w_1, w_2, \ldots, w_{m-1}, w_m, v_{m+1}, \ldots, v_n) = \mathrm{Sp}(w_1, w_2, \ldots, w_{m-1}, w_m, v_m, \ldots, v_n) = V$$

כלומר, $w_1, w_2, \dots, w_m, v_{m+1}, \dots, v_n$ סדרת יוצרים של V. ∎

הגדרה 4.33: יהי V מרחב וקטורי מעל שדה F. סדרה $v_1, v_2, \dots, v_n \in V$ נקראת **בסיס** של V אם היא בלתי תלויה לינארית וגם סדרת יוצרים של V.

דוגמה 4.36: (א) $e_1, e_2, \dots, e_m$ היא בסיס של F^m.

(היא סדרת יוצרים של F^m לפי תרגיל 4.20; היא בלתי תלויה לינארית לפי דוגמה 4.24(ב).)

(ב) יהי V מרחב השורות של $A \in \mathrm{M}_{m\times n}(F)$. אז השורות השונות מאפס של המטריצה המדורגת קנונית A' של A מהוות בסיס של V.

(לפי תרגיל 4.31 הן בלתי תלויות לינארית; לפי תרגיל 4.14 הן פורשות את $\mathrm{Row}(A')$; לפי משפט 4.23 $\mathrm{Row}(A) = \mathrm{Row}(A')$.)

(ג) בסיס למרחב הפולינומים ממעלה $\leq d$ הוא (למשל), $1, X, X^2, \dots, X^d$.

מסקנה 4.35: *אם למרחב וקטורי V יש בסיס, אז לכל בסיס של V אותו מספר איברים. מספר זה ייקרא* **המימד** *של V ויסומן $\dim(V)$.*

דוגמה 4.36: (א) $\dim\big(\mathrm{Row}(A)\big) = \mathrm{rk}(A)$.

(ב) $\dim(F^n) = n$.

משפט 4.37: *יהי V מרחב וקטורי מעל שדה F. סדרה $v_1, v_2, \dots, v_n \in V$ היא בסיס אם ורק אם לכל $v \in V$ יש $a_1, a_2, \dots, a_n \in F$ יחידים כך ש־ $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$.*

הוכחה: ברור ש"לכל $v \in V$ יש $a_1, a_2, \dots, a_n$ כך ש־$v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$" שקול לכך ש־$v_1, v_2, \dots, v_n$ סדרת יוצרים של V. לכן די להראות:

לכל $v \in V$ יש לכל היותר סדרה אחת $a_1, a_2, \dots, a_n$ כך ש־$v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$ אם ורק אם $v_1, v_2, \dots, v_n$ בלתי תלויה לינארית.

נניח שהתנאי הראשון מתקיים. יהיו $a_1, a_2, \dots, a_n \in F$ כך ש־$0 = a_1v_1 + a_2v_2 + \cdots + a_nv_n$. אך גם $0 = 0v_1 + 0v_2 + \cdots + 0v_n$. לפי היחידות $a_1 = a_2 = \cdots a_n = 0$.

להיפך, נניח ש־$v_1, v_2, \dots, v_n$ בלתי תלויה לינארית. יהי $v \in V$ ויהיו $a_1, a_2, \dots, a_n \in F$, $b_1, b_2, \dots, b_n \in F$, כך שמתקיים

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n = v = b_1v_1 + b_2v_2 + \cdots + b_nv_n.$$

אז $(a_1-b_1)v_1 + (a_2-b_2)v_2 + \cdots + (a_n-b_n)v_n = 0$. לכן $a_1-b_1 = a_2-b_2 = \cdots = a_n-b_n = 0$, ומכאן $a_1 = b_1, a_2 = b_2, a_n = b_n$. ∎

משפט 4.38: *יהי V מרחב וקטורי מעל שדה F. התנאים הבאים על סדרה $v_1, v_2, \ldots, v_m \in V$ שקולים זה לזה:*

(1) *$v_1, v_2, \ldots, v_m$ בסיס.*

(2) *$v_1, v_2, \ldots, v_m$ בלתי תלויה לינארית מקסימלית, כלומר, היא בלתי תלויה לינארית ולכל $v \in V$ הסדרה $v_1, v_2, \ldots, v_m, v$ תלויה לינארית.*

(3) *$v_1, v_2, \ldots, v_m$ סדרת יוצרים מינימלית, כלומר, היא סדרת יוצרים, ולכל $1 \leq i \leq n$ הסדרה $v_1, \ldots, \widehat{v_i}, \ldots, v_m$ איננה סדרת יוצרים.*

הוכחה: (1) $\Leftarrow$(2): המקסימליות נובעת ממשפט ההחלפה.

(2) $\Leftarrow$(1): צריך להראות ש־$v_1, v_2, \ldots, v_m$ סדרת יוצרים. יהי $v \in V$. בגלל המקסימליות $v_1, v_2, \ldots, v_m, v$ תלויה לינארית. לכן יש בה איבר שהוא צירוף לינארי של קודמיו. זה אינו אחד מבין $v_1, v_2, \ldots, v_m$ כי הם בלתי תלויים לינארית. לכן זה v.

(1) $\Leftarrow$(3): המינימליות נובעת ממשפט ההחלפה.

(3) $\Leftarrow$(1): צריך להראות ש־$v_1, v_2, \ldots, v_m$ בלתי תלויה לינארית. אילו היא היתה תלויה לינארית, היה אחד מאיבריה v_i צירוף לינארי של קודמיו. בפרט, v_i צירוף לינארי של יתר איברי הסדרה. לכן (תרגיל 4.14) $V = \mathrm{Sp}(v_1, \ldots, v_m) = \mathrm{Sp}(v_1, \ldots, \widehat{v_i}, \ldots, v_m)$, סתירה. ∎

הגדרה 4.39: מרחב וקטורי V נקרא **נוצר סופית** אם יש לו סדרת יוצרים.

משפט 4.40: *יהי V מרחב וקטורי נוצר סופית. אז יש לו בסיס. יתר על כן,*

(1) *כל סדרת יוצרים של V מכילה תת סדרה שהיא בסיס.*

(2) *כל סדרה בלתי תלויה לינארית ב־V ניתן להשלים לבסיס של V.*

יהי $n = \dim(V)$.

(3) *כל סדרת יוצרים בת n איברים ב־V היא בסיס של V.*

(4) *כל סדרה בלתי תלויה לינארית בת n איברים ב־V היא בסיס של V.*

הוכחה: (1) נשמיט איברים מסדרת היוצרים עד שנקבל סדרת יוצרים מינימלית. לפי המשפט הקודם היא בסיס.

(2) לפי ההנחה יש ל־V סדרת יוצרים בת n איברים. נוסיף איברים לסדרה הבלתי תלויה לינארית, עד שנקבל סדרה בלתי תלויה מקסימלית. (תהליך ההוספה חייב להסתיים לפני שבסדרה יהיו יותר מ־n איברים, לפי משפט ההחלפה.) לפי המשפט הקודם היא בסיס.

(3) נבחר תת סדרה של הסדרה המקורית שהיא בסיס, לפי (1). לפי מסקנה 4.35 יש בו n איברים, כמו בסדרה המקורית. לכן הסדרה המקורית היא הבסיס הזה.

(4) נשלים את הסדרה הנתונה לבסיס, לפי (2). לפי מסקנה 4.35 יש בו n איברים, כמו בסדרה המקורית. לכן הסדרה המקורית היא הבסיס הזה. ∎

תרגיל 4.41: מציאת בסיס של תת מרחב U של F^m אם נתונה סדרת יוצרים שלו $v_1, v_2, \ldots, v_n$. כמובן, $U = \mathrm{Sp}(v_1, v_2, \ldots, v_n)$.

(1) רשום את הוקטורים כעמודות של מטריצה $A \in \mathrm{M}_{m\times n}(F)$. תהי A' המטריצה המדורגת קנונית של A. אם מספרי העמודות המובילות ב־A' הם $j_1 < j_2 < \cdots < j_r$, אז $v_{j_1}, v_{j_2}, \ldots, v_{j_r}$ הוא בסיס של U.

אכן, אף אחד מבין איברי סדרה זו אינו צירוף לינארי של קודמיו (תרגיל 4.18), לכן סדרה זו בלתי תלויה לינארית (משפט 4.28). קל לראות שבמטריצה מדורגת קנונית כל עמודה לא מובילה היא צירוף לינארי של העמודות המובילות שלפניה. לכן לפי מסקנה 4.17(ב) כל העמודות של A שמספריהן מם מספרי העמודות הלא מובילות של A' הם צירופים לינאריים של $v_{j_1}, v_{j_2}, \ldots, v_{j_r}$. לפי תרגיל 4.14, $U = \mathrm{Sp}(v_1, v_2, \ldots, v_n) = \mathrm{Sp}(v_{j_1}, v_{j_2}, \ldots, v_{j_r})$.

(2) רשום את הוקטורים כשורות של מטריצה $B \in \mathrm{M}_{n\times m}(F)$. תהי B' המטריצה המדורגת קנונית של B. אז $U = \mathrm{Row}(B) = \mathrm{Row}(B')$, לכן השורות השונות מאפס מהוות בסיס של U. כמובן, בסיס זה איננו בהכרח מתוך הסדדה הנתונה. ∎

מסקנה 4.42: $\dim \mathrm{Col}(A) = \dim \mathrm{Row}(A) = \mathrm{rk}(A)$.

תרגיל 4.43: מציאת בסיס של $\mathrm{Ker}\, A$.

בלי הגבלת הכלליות $A \in \mathrm{M}_{m\times n}(F)$ מדורגת קנונית. יהי $r = \mathrm{rk}(A)$ ויהי $s = n - r$. יהיו $i_1 < i_2 < \cdots < i_r$ מספרי העמודות המובילות ויהיו $j_1 < j_2 < \cdots < j_s$ מספרי העמודות הלא מובילות. כדי לפשט את הסימון, נרשום את העמודות המובילות של A כראשונות, כלומר נחליף את סדר העמודות מ־$1, 2, \ldots, n$ ל־$i_1, \ldots, i_r, j_1, \ldots, j_s$. במלים אחרות, נחליף את סדר המשתנים מ־$X_1, X_2, \ldots, X_n$ ל־$X_{i_1}, \ldots, X_{i_r}, X_{j_1}, \ldots, X_{j_s}$. בפתרונות שנמצא (איברי בסיס של $\mathrm{Ker}\, A$), ייצגו r הרכיבים הראשונים את הרכיבים $i_1, \ldots, i_r$, ו־s הרכיבים האחרונים את הרכיבים $j_1, \ldots, j_s$.

$$A = \begin{array}{c} \begin{array}{ccccccc} i_1 & i_2 & \ldots & i_r & j_1 & \ldots & j_s \end{array} \\ \begin{pmatrix} 1 & 0 & \ldots & 0 & a_{1j_1} & \ldots & a_{1j_s} \\ 0 & 1 & \ldots & 0 & a_{2j_1} & \ldots & a_{2j_s} \\ & \ldots & & & & & \\ 0 & 0 & \ldots & 1 & a_{rj_1} & \ldots & a_{rj_s} \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ & \ldots & & & & & \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{pmatrix} \end{array} \qquad \begin{array}{l} X_{i_1} = -a_{1j_1}X_{j_1} - \cdots - a_{1j_s}X_{j_s} \\ X_{i_2} = -a_{2j_1}X_{j_1} - \cdots - a_{2j_s}X_{j_s} \\ \quad \ldots \\ X_{i_r} = -a_{rj_1}X_{j_1} - \cdots - a_{rj_s}X_{j_s} \end{array}$$

ואז קבוצת הפתרון היא, כפי שלמדנו (סדר הרכיבים בעמודות הוא $(i_1, i_2, \ldots, i_r, j_1, j_2, \ldots, j_s$

$$\operatorname{Ker} A = \left\{ \begin{pmatrix} -a_{1j_1}c_1 - a_{1j_2}c_2 - \cdots - a_{1j_s}c_s \\ -a_{2j_1}c_1 - a_{2j_2}c_2 - \cdots - a_{2j_s}c_s \\ \vdots \\ -a_{rj_1}c_1 - a_{rj_2}c_2 - \cdots - a_{rj_s}c_s \\ c_1 \\ c_2 \\ \vdots \\ c_s \end{pmatrix} \middle| \; c_1, c_2, \ldots, c_s \in F \right\}$$

$$= \left\{ c_1 \begin{pmatrix} -a_{1j_1} \\ -a_{2j_1} \\ \vdots \\ -a_{rj_1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} -a_{1j_2} \\ -a_{2j_2} \\ \vdots \\ -a_{rj_2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + c_s \begin{pmatrix} -a_{1j_s} \\ -a_{2j_s} \\ \vdots \\ -a_{rj_s} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \middle| \; c_1, c_2, \ldots, c_s \in F \right\}$$

$$= \operatorname{Sp}\left(\begin{pmatrix} -a_{1j_1} \\ -a_{2j_1} \\ \vdots \\ -a_{rj_1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} -a_{1j_2} \\ -a_{2j_2} \\ \vdots \\ -a_{rj_2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} -a_{1j_s} \\ -a_{2j_s} \\ \vdots \\ -a_{rj_s} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right)$$

קל לראות ש־$s = n - r$ העמודות האלה בלתי תלויות לינארית. לכן הן בסיס של P.

איך אפשר לאפיין את איברי הבסיס הזה?

לעמודה הראשונה יש 1 במקום המשתנה הלא מוביל הראשון ו־0 במקום יתר המשתנים הלא מובילים; המשתנים המובילים נקבעים על ידי בחירה זו, בגלל שעמודה היא פתרון של המערכת.

לעמודה השניה יש 1 במקום המשתנה הלא מוביל השני ו־0 במקום יתר המשתנים הלא מובילים; המשתנים המובילים נקבעים על ידי בחירה זו, בגלל שעמודה היא פתרון של המערכת.

וכן הלאה...

קל לראות שאם נשמיט מאיברי הבסיס את הרכיבים שמתאימים לעמודות מובילות, נקבל את העמודות $e_1, e_2, \ldots, e_s \in F^s$. אם נשמיט מהם את הרכיבים שמתאימים לעמודות הלא מובילות, ונכפיל אותם ב־(-1), נקבל את עמודות הלא מובילות של A, ללא שורות האפסים. ∎

דוגמה 4.44: מצא בסיס של $\mathrm{Ker}\, A$ אם $A = \begin{pmatrix} 0 & 1 & -2 & -5 & 0 & -8 & 0 & -3 \\ 0 & 0 & 0 & 0 & 1 & -4 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$.

המטריצה המצומצמת של המערכת כבר מדורגת קנונית; העמודות הלא מובילות הן חמש העמודות $1, 3, 4, 6, 8$.

נבחר, אם כן, חמשה פתרונות פרטיים u_1, u_2, u_3, u_4, u_5 מהצורה $(c_1, c_2, \ldots, c_9)^t$ שיהוו בסיס ל־$\mathrm{Ker}\, A$:

בחירה $c_1 = 1, c_3 = 0, c_4 = 0, c_6 = 0, c_8 = 0$ נותנת $u_1 = (1, 0, 0, 0, 0, 0, 0, 0)^t$.

בחירה $c_1 = 0, c_3 = 1, c_4 = 0, c_6 = 0, c_8 = 0$ נותנת $u_2 = (0, 2, 1, 0, 0, 0, 0, 0)^t$.

בחירה $c_1 = 0, c_3 = 0, c_4 = 1, c_6 = 0, c_8 = 0$ נותנת $u_3 = (0, 5, 0, 1, 0, 0, 0, 0)^t$.

בחירה $c_1 = 0, c_3 = 0, c_4 = 0, c_6 = 1, c_8 = 0$ נותנת $u_4 = (0, 8, 0, 0, 4, 1, 0, 0)^t$.

בחירה $c_1 = 0, c_3 = 0, c_4 = 0, c_6 = 0, c_8 = 1$ נותנת $u_5 = (0, 3, 0, 0, 2, 0, 7, 1)^t$.

חמשת הוקטורים האלה בלתי תלויים לינארית, כפי שקל לראות מהתבוננות ברכיבים $1, 3, 4, 6, 8$ שלהם. ∎

מסקנה 4.45: *תהי* $A \in \mathrm{M}_{m \times n}(F)$. *אז* $\dim \mathrm{Ker}\, A = n - \mathrm{rk}(A)$.

משפט 4.46: *יהי V מרחב וקטורי נוצר סופית מעל שדה F ויהי U תת מרחב שלו. אז*

(א) *U נוצר סופית, ו־* $\dim U \le \dim V$.

(ב) *אם* $\dim U = \dim V$ *אז* $U = V$.

הוכחה: נסמן $n = \dim V$.

(א) כל סדרה בלתי תלויה לינארית ב־U היא גם סדרה בלתי תלויה לינארית ב־V, לכן בת $\le n$ איברים (משפט ההחלפה). מכאן שיש ב־U סדרה בלתי תלויה מקסימלית. לפי משפט 4.38 היא בסיס. בפרט U נוצר סופית. יש בבסיס הנ"ל $\le n$ איברים, לכן $\dim U \le n = \dim V$.

(ב) יהי $u_1, u_2, \ldots, u_n$ בסיס של U. אז $u_1, u_2, \ldots, u_n$ סדרה בלתי תלויה לינארית ב־U, ולכן גם ב־V. היא בת n איברים, לכן לפי משפט 4.40 היא בסיס של V. מכאן $U = \mathrm{Sp}(u_1, u_2, \ldots, u_n) = V$. ∎

תרגיל 4.47: $F[X]$ אינו נוצר סופית.

אכן, לו היתה לו סדרת יוצרים סופית, נאמר, בת n איברים, אז כל סדרה בלתי תלויה לינארית בו היתה בת n איברים לכל היותר. אך $1, X, X^2, \ldots, X^n$ בלתי תלויה לינארית בת $n+1$ איברים, סתירה.

דוגמה 4.48: מהם כל תת המרחבים של המרחב התלת ממדי $\mathbb{R}^3$?

הם מארבעה סוגים:

(א) ממימד 0: מרחב האפס $\{0\}$ (נקודת הָראשית).

(ב) ממימד 1: ישרים דרך הראשית.

(ג) ממימד 2: מישורים דרך הראשית.

(ד) ממימד 3: כל המרחב (לפי משפט 4.46(ב)).

סכומים של תת־מרחבים.

הגדרה 4.49: יהי V מרחב וקטורי מעל שדה F ויהיו U, W שני תת־מרחבים שלו. נגדיר

$$.U + W = \{u + w|\ u \in U, w \in W\}$$

משפט 4.50: *יהי V מרחב וקטורי מעל שדה F ויהיו U, W שני תת־מרחבים שלו.*

(א) *$U + W$ תת מרחב של V.*

(ב) $.U, W \subseteq U + W$

(ג) *אם V' תת מרחב של V כך ש־ $U, W \subseteq V'$, אז $.U + W \subseteq V'$*

(ד) *אם $U = \mathrm{Sp}(u_1, u_2, \ldots, u_m), W = \mathrm{Sp}(w_1, w_2, \ldots, w_n)$, אז*

$$.U + W = \mathrm{Sp}(u_1, u_2, \ldots, u_m, w_1, w_2, \ldots, w_n)$$

הוכחה: (א) $0 \in U, W$ לכן $.0 = 0 + 0 \in U + W$

יהיו $u + w, u' + w' \in U + W$ (באשר $u, u' \in U, w, w' \in W$). אז $u + u' \in U, w + w' \in W$, (כי U, W תת מרחבים), לכן $.(u + w) + u('+w') = (u + u') + (w + w') \in U + W$

יהי $u + w \in U + W$ (באשר $u \in U, w \in W$) ויהי $a \in F$. אז $au \in U, aw \in W$, (כי U, W תת מרחבים), לכן $.a(u + w) = au + aw \in U + W$

(ב) יהי $u \in U$. אז $u = u + 0 \in U + W$. לכן $U \subseteq U + W$. באופן דומה $.W \subseteq U + W$

(ג) יהי $u + w \in U + W$ (באשר $u \in U, w \in W$). אז $u, w \in V'$, לכן $u + w \in V'$. לכן $.U + W \subseteq V'$

(ד) אגף ימין הוא תת מרחב של V. הוא מכיל את $u_1, u_2, \ldots, u_m$, לכן את U. באופן דומה הוא מכיל את W. לפי (ג) הוא מכיל את $.U + W$

להיפך, $U + W$ הוא תת מרחב של V. לפי (ב) הוא מכיל את U לכן את $u_1, u_2, \ldots, u_m$. באופן דומה הוא מכיל את $w_1, w_2, \ldots, w_n$. לכן, משפט 4.12, לפי הוא מכיל את אגף ימין. ∎

משפט 4.51 (משפט המימד): *יהי V מרחב וקטורי נוצר סופית מעל שדה F ויהיו U, W שני תת־מרחבים שלו. אז*

$$.\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W)$$

הוכחה: לפי המשפט הקודם $U, W \subseteq U + W \subseteq V$ תת מרחבים של מרחב נוצר סופית. לפי משפט 4.46 הם נוצרים סופית. יהיו $r = \dim(U \cap W)$, $m = \dim(U)$, $.n = \dim(W)$ נבחר בסיס $v_1, v_2, \ldots, v_r$ של $U \cap W$ ונשלים אותו

(א) לבסיס $v_1, v_2, \ldots, v_r, u_{r+1}, \ldots, u_m$ של $;U$

(ב) לבסיס $v_1, v_2, \ldots, v_r, w_{r+1}, \ldots, w_n$ של W. אז לפי משפט 4.50(ד)

$$\begin{aligned} U+W =& \mathrm{Sp}(v_1, v_2, \ldots, v_r, u_{r+1}, \ldots, u_m, v_1, v_2, \ldots, v_r, w_{r+1}, \ldots, w_n) \\ =& \mathrm{Sp}(v_1, v_2, \ldots, v_r, u_{r+1}, \ldots, u_m, w_{r+1}, \ldots, w_n) \end{aligned}$$

לכן די להוכיח: $v_1, v_2, \ldots, v_r, u_{r+1}, \ldots, u_m, w_{r+1}, \ldots, w_n$ בלתי תלויה לינארית.

נוכיח זאת: יהיו $a_1, a_2, \ldots, a_r, b_{r+1}, \ldots, b_m, c_{r+1}, \ldots, c_n \in F$ כך ש־

$$a_1v_1 + a_2v_2 \cdots + a_rv_r + b_{r+1}u_{r+1} \cdots + b_mu_m + c_{r+1}w_{r+1} + \cdots + c_nw_n = 0. \tag{1}$$

אז

$$a_1v_1 + a_2v_2 \cdots + a_rv_r + b_{r+1}u_{r+1} \cdots + b_mu_m = (-c_{r+1})w_{r+1} + \cdots + (-c_n)w_n. \tag{2}$$

נסמן את שני האגפים של (1) (ששוים זה לזה) ב־v. אגף שמאל הוא צירוף לינארי של איברי U, לכן $v \in U$. אגף ימין הוא צירוף לינארי של איברי W, לכן $v \in W$. לכן $v \in U \cap W$.

כיוון ש־$v_1, \ldots, v_r$ בסיס של $U \cap W$, יש $a'_1, a'_2, \ldots, a'_r \in F$ כך ש־ $v = a'_1v_1 + a'_2v_2 \cdots + a'_rv_r$.
מכאן ומאגף שמאל של (2) מקבלים

$$\begin{aligned} v =& a'_1v_1 + a'_2v_2 \cdots + a'_rv_r + 0u_{r+1} \cdots + 0u_m \\ v =& a_1v_1 + a_2v_2 \cdots + a_rv_r + b_{r+1}u_{r+1} \cdots + b_mu_m \end{aligned}$$

אך לפי יחידות ההצגה של וקטור לפי הבסיס $v_1, v_2, \ldots, v_r, u_{r+1}, \ldots, u_m$ של U (משפט 4.37) מקבלים $b_{r+1} = 0, \ldots, b_m = 0$.

נציב זאת במשוואה (1) ונקבל

$$a_1v_1 + a_2v_2 \cdots + a_rv_r + c_{r+1}w_{r+1} + \cdots + c_nw_n = 0.$$

היות ו־ $v_1, v_2, \ldots, v_r, w_{r+1}, \ldots, w_n$ בסיס של W ולכן בלתי תלויה לינארית,

$$a_1 = a_2 = \cdots = a_r = c_{r+1} = \cdots = c_n = 0 \quad \blacksquare$$

תרגיל 4.52: יהיו U, W שני תת מרחבים של $\mathbb{R}^{10}$ ממימד 9, שונים זה מזה. אז $\dim(U \cap W) = 8$.

הוכחה: $U + W$ תת מרחב של $\mathbb{R}^{10}$, לכן לפי משפט 4.46(א), $\dim(U+W) \leq \dim(\mathbb{R}^{10}) = 10$. לפי משפט המימד, $\dim(U+W) = \dim(U) + \dim(W) - \dim(U \cap W)$, לכן $9 + 9 - \dim(U \cap W) \leq 10$, ומכאן $\dim(U \cap W) \geq 8$.

מצד שני $U\cap W$ תת מרחב של U, לכן לפי משפט 4.46(א), $\dim(U\cap W)\le\dim(U)=9$. אילו היה כאן שוויון, אז לפי משפט 4.46(ב), היה $U\cap W=U$, ולכן $U\subseteq U\cap W\subseteq W$. היות ו־$\dim(U)=\dim(W)=9$, נובע מכאן לפי משפט 4.46(ב), $U=W$, בסתירה לנתון.

מתוך $\dim(U\cap W)\ge 8$, $\dim(U\cap W)<9$ נובע $\dim(U\cap W)=8$. ∎

יהיו $U,W\subseteq F^n$ שני תת מרחבים של F^n. כיצד מוצאים בסיס לחיתוך $U\cap W$?

נניח ש־$U=\operatorname{Ker}A_1$ ו־$W=\operatorname{Ker}A_2$, באשר $A_1\in \mathrm{M}_{m_1\times n}(F)$ ו־$A_2\in \mathrm{M}_{m_2\times n}(F)$. נצרף את השורות של A_1,A_2 לשורות של מטריצה אחת $\begin{pmatrix}A_1\\A_1\end{pmatrix}\in \mathrm{M}_{(m_1+m_2)\times n}(F)$. אז ברור ש־$U\cap W=\operatorname{Ker}\begin{pmatrix}A_1\\A_1\end{pmatrix}$. למדנו כיצד למצוא בסיס למרחב כזה, וזה נותן תשובה לבעיה.

על כן עלינו להסביר כיצד פותרים את התרגיל הבא:

תרגיל 4.53: יהיו $v_1,\dots,v_m\in F^n$. נסמן $U=\operatorname{Sp}(v_1,\dots,v_m)$. מצא מטריצה B כך ש־$U=\operatorname{Ker}B$.

פתרון: תהי $A=(v_1,\dots,v_n)\in \mathrm{M}_{n\times m}(F)$. למדנו למצוא בסיס ל־$\operatorname{Ker}A^t$, נאמר, $w_1,\dots,w_r\in F^n$.
תהי $B^t=(w_1,\dots,w_r)\in \mathrm{M}_{n\times r}(F)$. אז $B\in \mathrm{M}_{r\times n}(F)$.

טענה: $U=\operatorname{Ker}B$.

הוכחה: נראה כי (א) $U\subseteq\operatorname{Ker}B$; (ב) $\dim U=\dim\operatorname{Ker}B$. מכאן, לפי משפט 4.46, $U=\operatorname{Ker}B$.

(א) היות ו־$w_1,\dots,w_r\in\operatorname{Ker}A^t$, מתקיים $A^tw_1=0,\dots,A^tw_r=0$. מכאן $A^tB^t=0$. לכן $BA=(A^tB^t)^t=0$. זה אומר ש־$Bv_1=0,\dots,Bv_m=0$, כלומר, $v_1,\dots,v_m\in\operatorname{Ker}B$. מכאן $U=\operatorname{Sp}(v_1,\dots,v_m)\subseteq\operatorname{Ker}B$.

(ב) מתקיים $U=\operatorname{Col}(A)$, לכן $\dim U=\dim\operatorname{Col}(A)=\operatorname{rk}A$. לכן $r=\dim\operatorname{Ker}A^t=n-\operatorname{rk}A^t=n-\operatorname{rk}A=n-\dim U$.

העמודות $w_1,\dots,w_r$ של B^t הן בלתי תלויות לינארית (כי הן בסיס של $\operatorname{Ker}A^t$), לכן הן בסיס של $\operatorname{Col}(B^t)$. לכן $\operatorname{rk}(B)=\operatorname{rk}(B^t)=\dim\operatorname{Col}(B^t)=r$. מכאן

$\dim\operatorname{Ker}B=n-\operatorname{rk}B=n-r=n-(n-\dim U)=\dim U$. ∎

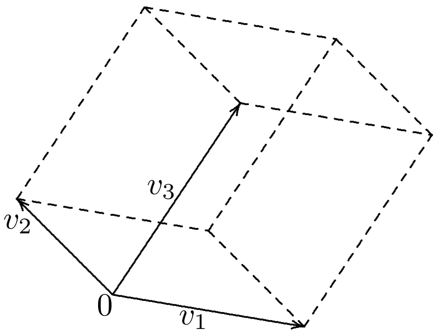
5. דטרמיננטות

פונקציות הנפח.

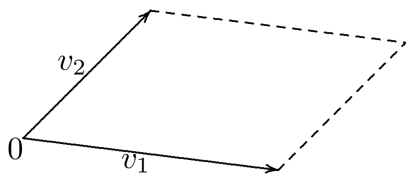
מוטיבציה:

(1) עבור זוג וקטורים v_1, v_2 במישור נסמן ב־$\mathcal{N}(v_1, v_2)$ את השטח (המסומן) של המקבילית הנקבעת על ידי v_1, v_2, כחצים מהראשית. (השטח נחשב לחיובי אם הזוית בין v_1 ל־v_2 היא חיובית, כלומר, נגד כיוון השעון.)

(2) באופן דומה עבור שלשה של וקטורים v_1, v_2, v_3 במרחב נסמן ב־$\mathcal{N}(v_1, v_2, v_3)$ את הנפח (המסומן) של המקבילון הנקבע על ידי v_1, v_2, v_3, כחצים מהראשית.



(2)



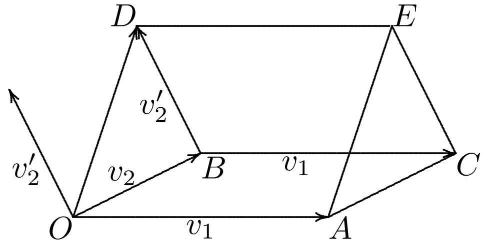
(1)

(3) כיצד לחשב אותם? האם יש הכללות של המושגים האלה ל־$\mathbb{R}^n$? ואולי אף ל־F^n, כאשר F שדה כלשהו?

הערה 5.1: תכונות של פונקצית הנפח. (נדון רק במקרה של וקטורים במישור.)

(א) $\mathcal{N}(v_1, v_2) + \mathcal{N}(v_1, v_2') = \mathcal{N}(v_1, v_2 + v_2')$

אכן, בציור הבא (בו כל הוקטורים הם במישור דף הנייר) $\overrightarrow{OD} = v_2 + v_2'$



ומתקיים (כאשר $\square$ מציין את שטח המקבילית שמימינו ו־$\triangle$ את שטח המשולש שמימינו)

$$\mathcal{N}(v_1, v_2) + \mathcal{N}(v_1, v_2') = \square(OACB) + \square(BCED) =$$

$$\square(OACB) + \square(BCED) + \triangle(OBD) - \triangle(ACE) = \square(OAED) = \mathcal{N}(v_1, v_2 + v_2')$$

באופן דומה $\mathcal{N}(v_1, v_2) + \mathcal{N}(v_1', v_2) = \mathcal{N}(v_1 + v_1', v_2)$.

(ב) $\mathcal{N}(av_1, v_2) = a\mathcal{N}(v_1, v_2) = \mathcal{N}(v_1, av_2)$ **לכל** $a \in F$.

(ג) **אם** $v_1 = v_2$ **אז** $\mathcal{N}(v_1, v_2) = 0$.

(ד) $\mathcal{N}(e_1, e_2) = 1$ **כאן** $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $e_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. **(כלומר, שטח ריבוע היחידה הוא 1.)**

הגדרה 5.2: **יהי F שדה. פונקציה $\mathcal{N}$ שמתאימה לכל n וקטורים $v_1, v_2, \ldots, v_n \in F^n$ ערך בשדה $\mathcal{N}(v_1, v_2, \ldots, v_n) \in F$ נקראת פונקצית נפח על F^n אם**

(א) $\mathcal{N}$ **מולטילינארית:**

(א') **לכל** $1 \leq i \leq n$ **ולכל** $v_1, v_2, \ldots, v_n, v_i' \in F^n$

$$\begin{aligned}&\mathcal{N}(v_1, \ldots, v_{i-1}, v_i + v_i', v_{i+1}, \ldots, v_n) = \\ &\mathcal{N}(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n) + \mathcal{N}(v_1, \ldots, v_{i-1}, v_i', v_{i+1}, \ldots, v_n)\end{aligned}$$

(א'') **לכל** $1 \leq i \leq n$, **לכל** $v_1, v_2, \ldots, v_n \in F^n$ **ולכל** $a \in F$

$$\mathcal{N}(v_1, \ldots, v_{i-1}, av_i, v_{i+1}, \ldots, v_n) = a\mathcal{N}(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n)$$

(ב) **אם שניים (לפחות) מבין הוקטורים** $v_1, v_2, \ldots, v_n$ **שווים, אז** $\mathcal{N}(v_1, v_2, \ldots, v_n) = 0$

(ג) $\mathcal{N}$ **מנורמלת:** $\mathcal{N}(e_1, e_2, \ldots, e_n) = 1$

עבור $A \in \mathrm{M}_n(F)$ **נגדיר** $\mathcal{N}(A) = \mathcal{N}(v_1, v_2, \ldots, v_n)$**, באשר** $v_1, v_2, \ldots, v_n$ **שורותיה של** A**. אז, למשל, תנאי (ג) הוא** $\mathcal{N}(I_n) = 1$

דוגמה 5.3: **נניח** $n = 2$**. נגדיר** $\mathcal{N}\big((a,b),(c,d)\big) = ad - bc$**. בכתיב של מטריצות:** $\mathcal{N}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$ **זוהי פונקצית נפח.**

אכן (נבדוק כאן רק את התכונה (א'); השאר קל יותר לבדיקה):

$$\begin{aligned}&\mathcal{N}\big((a,b)+(a',b'),(c,d)\big) = \mathcal{N}\big((a+a',b+b'),(c,d)\big) = (a+a')d - (b+b')c = \\ &(ad-bc) + (a'd - b'c) = \mathcal{N}\big((a,b),(c,d)\big) + \mathcal{N}\big((a',b'),(c,d)\big)\end{aligned}$$ ∎

מעתה תהי $\mathcal{N}$ פונקצית נפח, $A \in \mathrm{M}_n(F)$ **מטריצה,** $v_1, v_2, \ldots, v_n$ **שורותיה.**

תרגיל 5.4: **אם ב־A שורת אפסים אז** $\mathcal{N}(A) = 0$

הוכחה: **נניח** $v_i = 0$**. אז**

$$\mathcal{N}(A) = \mathcal{N}(v_1, \ldots, v_{i-1}, 0v_i, v_{i+1}, \ldots, v_n) = 0\mathcal{N}(v_1, \ldots, v_i, \ldots, v_n) = 0$$ ∎

משפט 5.5: *תהי $\mathcal{P}$ פעולה אלמנטרית. אז $\mathcal{N}(\mathcal{P}(A)) = \beta \cdot \mathcal{N}(A)$, באשר* $\beta = \begin{cases} -1 & \mathcal{P} = \mathcal{P}_{kl} \text{ אם,} \\ \alpha & \mathcal{P} = \mathcal{P}_k(\alpha) \text{ אם,} \\ 1 & \mathcal{P} = \mathcal{P}_{kl}(\lambda) \text{ אם} \end{cases}$.

הוכחה: (א) נניח $\mathcal{P} = \mathcal{P}_{kl}$. נעשה חישוב שלכאורה אינו קשור להוכחה:

$$\begin{array}{rllllll} 0. & = & \mathcal{N}(v_1, \ldots, v_{k-1}, & v_k + v_l, & v_{k+1}, \ldots, v_{l-1}, & v_k + v_l, & v_{l+1}, \ldots, v_n) \\ & = & \mathcal{N}(v_1, \ldots, v_{k-1}, & v_k, & v_{k+1}, \ldots, v_{l-1}, & v_k, & v_{l+1}, \ldots, v_n) \\ & + & \mathcal{N}(v_1, \ldots, v_{k-1}, & v_k, & v_{k+1}, \ldots, v_{l-1}, & v_l, & v_{l+1}, \ldots, v_n) \\ & + & \mathcal{N}(v_1, \ldots, v_{k-1}, & v_l, & v_{k+1}, \ldots, v_{l-1}, & v_k, & v_{l+1}, \ldots, v_n) \\ & + & \mathcal{N}(v_1, \ldots, v_{k-1}, & v_l, & v_{k+1}, \ldots, v_{l-1}, & v_l, & v_{l+1}, \ldots, v_n) \end{array}$$

המחובר הראשון והרביעי הם 0, השני הוא $\mathcal{N}(A)$ והשלישי $\mathcal{N}(\mathcal{P}_{kl}(A))$. לכן $\mathcal{N}(A) + \mathcal{N}(\mathcal{P}_{kl}(A)) = 0$.

(ב) נניח $\mathcal{P} = \mathcal{P}_k(\alpha)$. אז $\mathcal{N}(\mathcal{P}_k(\alpha)(A) = \alpha\mathcal{N}(A)$ מהתכונה (א") של $\mathcal{N}$ בהגדרה 5.2.

(ג) נניח $\mathcal{P} = \mathcal{P}_{kl}(\lambda)$. נשתמש בתכונות (א'), (א"), (ב) של $\mathcal{N}$ בהגדרה 5.2:

$$\begin{array}{rcllllll} \mathcal{N}(\mathcal{P}_{kl}(\lambda)(A)) & = & \mathcal{N}(v_1, \ldots, v_{k-1}, & v_k, & v_{k+1}, \ldots, v_{l-1}, & v_l + \lambda v_k, & v_{l+1} \ldots, v_n) \\ & = & \mathcal{N}(v_1, \ldots, v_{k-1}, & v_k, & v_{k+1}, \ldots, v_{l-1}, & v_l, & v_{l+1} \ldots, v_n) \\ & + & \lambda\mathcal{N}(v_1, \ldots, v_{k-1}, & v_k, & v_{k+1}, \ldots, v_{l-1}, & v_k, & v_{l+1} \ldots, v_n) \\ & = & \mathcal{N}(A) + \lambda 0 & = & \mathcal{N}(A) & & \end{array} \quad \blacksquare$$

מסקנה 5.6: *אם $\mathcal{P}$ פעולה אלמנטרית אז* $\mathcal{N}(A) \neq 0 \Leftrightarrow \mathcal{N}(\mathcal{P}(A)) \neq 0$.

הוכחה: לפי המשפט, $\mathcal{N}(\mathcal{P}(A)) = \beta\mathcal{N}(A)$, באשר β הוא -1 או α או 1. בכל מקרה, $\beta \neq 0$. ∎

משפט 5.7: *A הפיכה (כלומר $\mathrm{rk}(A) = n$) אם ורק אם* $\mathcal{N}(A) \neq 0$.

הוכחה: תהי A' המטריצה המדורגת קנונית של A. היא התקבלה מ־A על ידי סדרה של פעולות אלמנטריות.

אם A הפיכה אז $A' = I_n$, לכן $\mathcal{N}(A') = 1 \neq 0$, ולכן, לפי מסקנה 5.6, $\mathcal{N}(A) \neq 0$.

אם A אינה הפיכה אז ב־A' יש שורת אפסים, לכן, לפי תרגיל 5.4, $\mathcal{N}(A') = 0$, ולכן, לפי מסקנה 5.6, $\mathcal{N}(A) = 0$. ∎

מסקנה 5.8: *תהי E מטריצה אלמנטרית המתאימה לפעולה אלמנטרית $\mathcal{P}$. אז $\mathcal{N}(E) = \beta$, באשר β כמו במשפט 5.5.*

הוכחה: לפי ההגדרה, $E = \mathcal{P}(I_n)$. נציב $A = I_n$ במשפט 5.5: $\mathcal{N}(E) = \beta \cdot \mathcal{N}(I_n) = \beta$ ∎

מסקנה 5.9: *אם E מטריצה אלמנטרית, אז* $\mathcal{N}(EA) = \mathcal{N}(E)\mathcal{N}(A)$

הוכחה: אם E מתאימה לפעולה אלמנטרית $\mathcal{P}$, אז $EA = \mathcal{P}(A)$. לפי משפט 5.5 $\mathcal{N}(\mathcal{P}(A)) = \beta\mathcal{N}(A)$, באשר $\beta \in F$ רשום במשפט (תלוי ב־$\mathcal{P}$). לפי מסקנה 5.8, $\beta = \mathcal{N}(E)$. ∎

משפט 5.10 (משפט המכפלה): *אם $A, B \in \mathrm{M}_n(F)$, אז* $\mathcal{N}(AB) = \mathcal{N}(A)\mathcal{N}(B)$

הוכחה: (1) אם $\mathrm{rk}(A) < n$, אז $\mathrm{rk}(AB) \leq \mathrm{rk}(A) < n$ ולכן $\mathcal{N}(AB) = 0$, $\mathcal{N}(A) = 0$, ומכאן הנוסחה.

(2) אם $\mathrm{rk}(A)=n$, אז A הפיכה, ולכן מכפלה של מטריצות אלמנטריות, נאמר $A=E_k\cdots E_1$. נראה $\mathcal{N}(E_k\cdots E_1B)=\mathcal{N}(E_k\cdots E_1)\mathcal{N}(B)$ באינדוקציה על k.

עבור $k=1$ זוהי מסקנה 5.9.

נניח נכונות עבור $k-1$. לפי מסקנה 5.9 והנחת האינדוקציה

$$\mathcal{N}(E_kE_{k-1}\cdots E_1B)=\mathcal{N}(E_k)\mathcal{N}(E_{k-1}\cdots E_1B)=\mathcal{N}(E_k)\mathcal{N}(E_{k-1}\cdots E_1)\mathcal{N}(B)=$$
$$\mathcal{N}(E_kE_{k-1}\cdots E_1)\mathcal{N}(B) \qquad \blacksquare$$

משפט 5.11: *לכל n יש לכל היותר פונקצית נפח אחת $\mathcal{N}$.*

הוכחה: תהיינה $\mathcal{N}_1,\mathcal{N}_2$ שתי פונקציות נפח. נראה שלכל $A\in \mathrm{M}_n(A)$ מתקיים $\mathcal{N}_1(A)=\mathcal{N}_2(A)$

(1) אם A אינה הפיכה אז $\mathcal{N}_1(A)=0=\mathcal{N}_2(A)$ לפי משפט 5.7.

(2) אם A אלמנטרית אז $\mathcal{N}_1(A)=\mathcal{N}_2(A)$ לפי מסקנה 5.8.

(3) אם A הפיכה, אז היא מכפלה של מטריצות אלמנטריות $A=E_k\cdots E_1$, ולכן לפי משפט המכפלה

$$\mathcal{N}_1(A)=\mathcal{N}_1(E_k)\cdots\mathcal{N}_1(E_1)=\mathcal{N}_2(E_k)\cdots\mathcal{N}_2(E_1)=\mathcal{N}_2(A) \qquad \blacksquare$$

משפט 5.12: $\mathcal{N}(A^t)=\mathcal{N}(A)$.

הוכחה: (1) אם A אינה הפיכה אז גם A^t אינה הפיכה, ולכן שני האגפים הם 0, לפי משפט 5.7.

(2) נניח ש-A אלמנטרית, כלומר $A=\mathcal{P}(I_n)$, באשר $\mathcal{P}$ פעולה אלמנטרית.

אם $\mathcal{P}=\mathcal{P}_{kl}$ או $\mathcal{P}=\mathcal{P}_k(\alpha)$ אז $A^t=A$ ולכן $\mathcal{N}(A^t)=\mathcal{N}(A)$

אם $\mathcal{P}=\mathcal{P}_{kl}(\lambda)$ אז גם A^t מטריצה אלמנטרית; היא מתאימה לפעולה $\mathcal{P}_{lk}(\lambda)$. לפי מסקנה 5.8 $\mathcal{N}(A^t)=1=\mathcal{N}(A)$

(3) אם A הפיכה, אז היא מכפלה של מטריצות אלמנטריות $A=E_k\cdots E_1$, ולכן $A^t=E_1^t\cdots E_k^t$. לכן לפי (2),

$$\mathcal{N}(A^t)=\mathcal{N}(E_1^t)\cdots\mathcal{N}(E_k^t)=\mathcal{N}(E_1)\cdots\mathcal{N}(E_k)=\mathcal{N}(E_k\cdots E_1)=\mathcal{N}(A) \qquad \blacksquare$$

משפט 5.13: *לכל n קיימת פונקצית נפח יחידה (על מטריצות מסדר $n\times n$). היא נקראת* ***הדטרמיננטה****. סימון:* $\det(A)$ *או* $|A|$*, באשר* $A\in \mathrm{M}_n(F)$.

הוכחה: היחידות הוכחה במשפט 5.11, לכן די למצוא פונקצית נפח אחת לכל n. דבר זה נעשה באינדוקציה על n.

עבור $n=1$ נגדיר $|(a)|=a$.

(עבור $n=2$ נגדיר $\mathcal{N}\begin{pmatrix} a & b \\ c & d\end{pmatrix}=ad-bc$; ראה דוגמה 5.3.)

נניח, באינדוקציה, שהגדרנו פונקצית נפח (עם הסימון $|\,|$) למטריצות מסדר $(n-1)\times(n-1)$.

תהי $A \in \mathrm{M}_n(F)$. עבור $1 \le i, j \le n$ נסמן ב־A_{ij} את **המינור ה־**(i,j) של A, דהיינו, המטריצה מסדר $(n-1)\times(n-1)$ שמתקבלת מ־A על ידי מחיקת השורה ה־i והעמודה ה־j. (להבדיל מ־$(A)_{ij}$, שמסמן את הרכיב של A בשורה ה־i בעמודה ה־j.) אז נגדיר

$$|A| = \sum_{i=1}^{n} (-1)^{i+j} (A)_{ij} \, |A_{ij}| \, , \tag{1}$$

באשר j כלשהו.

נותר להוכיח, שזוהי פונקצית נפח. (הוכחה זו לא הובאה בשיעור וידיעתה לא תידרש במבחן.)

(א') יהי $1 \le k \le n$. תהיינה $A, B, C \in \mathrm{M}_n(F)$ מטריצות שנבדלות זו מזו רק בשורה ה־k, כאשר השורה ה־k של C היא סכום השורות ה־k של A, B. צריך להוכיח: $|C| = |A| + |B|$.

נראה תחילה שלכל $1 \le i \le n$ מתקיים

$$(C)_{ij}|C_{ij}| = (A)_{ij}|A_{ij}| + (B)_{ij}|B_{ij}| . \tag{2}$$

אם $i \ne k$, אז $(A)_{ij} = (B)_{ij} = (C)_{ij}$. כמו כן, גם המינורים $A_{ij}, B_{ij}, C_{ij} \in \mathrm{M}_{n-1}(F)$ נבדלים זה מזה רק בשורה אחת, ושורה זו ב־C_{ij} היא סכום השורות המתאימות ב־A_{ij}, B_{ij}. לכן לפי הנחת האינדוקציה, $|C_{ij}| = |A_{ij}| + |B_{ij}|$. מכאן (2).

אם $i = k$, אז $A_{ij} = B_{ij} = C_{ij}$ ואילו $(C)_{ij} = (A)_{ij} + (B)_{ij}$ (רכיבים בשורות ה־k). מכאן (2).

לכן

$$\begin{aligned} |C| &= \sum_{i=1} (-1)^{i+j} (C)_{ij}|C_{ij}| = \sum_{i=1} (-1)^{i+j} \big((A)_{ij}|A_{ij}| + (B)_{ij}|B_{ij}| \big) \\ &= \sum_{i=1} (-1)^{i+j} (A)_{ij}|A_{ij}| + \sum_{i=1} (-1)^{i+j} (B)_{ij}|B_{ij}| = |A| + |B| . \end{aligned}$$

(א'') יהי $1 \le k \le n$ ויהי $a \in F$. תהיינה $A, B \in \mathrm{M}_n(F)$ מטריצות שנבדלות זו מזו רק בשורה ה־k, כאשר השורה ה־k של B היא הכפולה ב־a של השורה ה־k של A. צריך להוכיח: $|B| = a|A|$.

נראה תחילה שלכל $1 \le i \le n$ מתקיים

$$(B)_{ij}|B_{ij}| = a(A)_{ij}|A_{ij}| . \tag{3}$$

אם $i \ne k$, אז $(A)_{ij} = (B)_{ij}$. כמו כן, גם המינורים $A_{ij}, B_{ij} \in \mathrm{M}_{n-1}(F)$ נבדלים זה מזה רק בשורה אחת, ושורה זו ב־B_{ij} היא הכפולה ב־a של השורה המתאימה ב־A_{ij}. לכן לפי הנחת האינדוקציה, $|B_{ij}| = a|A_{ij}|$. מכאן (3).

אם $i = k$, אז $A_{ij} = B_{ij}$ ואילו $(B)_{ij} = a(A)_{ij}$ (כי אלה רכיבים בשורה ה־k). מכאן (3).

לכן

$$|B| = \sum_{i=1}(-1)^{i+j}(B)_{ij}|B_{ij}| = \sum_{i=1}(-1)^{i+j}a(A)_{ij}|A_{ij}| = a\sum_{i=1}(-1)^{i+j}(A)_{ij}|A_{ij}| = a|A|.$$

(ב) תהי $A \in \mathrm{M}_n(F)$ ונניח שיש לה שתי שורות זהות: השורה ה־k והשורה ה־l, באשר $k < l$. נראה ש־$|A| = 0$.

יהי $1 \le i \le n$. אם $i \neq k, l$ אז גם במינור A_{ij} שתי שורות זהות, ולכן $|A_{ij}| = 0$. כמו כן $(A)_{kj} = (A)_{lj}$ (בגלל שאלה רכיבים בשורות הזהות). קל לראות שאת המינור A_{kj} אפשר לקבל מהמינור A_{lj} על ידי מחיקת השורה ה־k והוספתה אחרי השורה ה־$(l-1)$. מחיקה והוספה אלה שקולות להחלפת השורה ה־k ב־A_{lj} עם $(l-1-k)$ שורות שאחריה (שמספריהן $k+1, k+2, \dots, l-1$). כל החלפה כזו מכפילה ב־(-1) את הדטרמיננטה של המינור. לכן $|A_{kj}| = (-1)^{l-1-k}|A_{lj}|$. לכן

$$\begin{aligned}|A| &= \sum_{i=1}(-1)^{i+j}(A)_{ij}|A_{ij}| = (-1)^{k+j}(A)_{kj}|A_{kj}| + (-1)^{l+j}(A)_{lj}|A_{lj}| = \\ &= (-1)^{k+j}(-1)^{l-1-k}(A)_{lj}|A_{lj}| + (-1)^{l+j}(A)_{lj}|A_{lj}| = \\ &= ((-1)^{l+j-1} + (-1)^{l+j})(A)_{lj}|A_{lj}| = 0.\end{aligned}$$

(ג) נראה שאם $A = I_n$ אז $|A| = 1$. אכן, $|A_{jj}| = 1$. כמו כן $(A)_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$. לכן

$$|A| = \sum_{i=1}(-1)^{i+j}(A)_{ij}|A_{ij}| = (-1)^{j+j}(A)_{jj}|A_{jj}| = (-1)^{2j}1 \cdot 1 = 1.$$

הוכחנו שלפונקציה המוצעת כל התכונות של פונקצית נפח, לכן היא פונקצית נפח. ∎

הערה 5.14: לפי היחידות של פונקצית הנפח, ההגדרה (1) איננה תלויה בבחירת j. לכן (1) נותן n אפשרויות שונות לחישוב הדטרמיננטה. קוראים לה **נוסחת הפיתוח של הדטרמיננטה לפי העמודה ה־j**.

דוגמה 5.15: *חישוב דטרמיננטות לפי ההגדרה במשפט.*

(א) נִקח $j = 1$:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = (-1)^{1+1}a\,|d| + (-1)^{2+1}c\,|b| = ad - bc$$

(ב) נִקח, למשל, $j = 2$:

$$\begin{aligned}\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} &= (-1)^{1+2}b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + (-1)^{2+2}e\begin{vmatrix} a & c \\ g & i \end{vmatrix} + (-1)^{3+2}h\begin{vmatrix} a & c \\ d & f \end{vmatrix} = \\ &= -b(di - gf) + e(ai - gc) - h(af - dc) = \\ &= aei + dhc + gbf - ceg - fha - ibd\end{aligned}$$

(ג) אם $n=4$ אז

$$|A|=\sum_{i=1}^{4}(-1)^{i+1}(A)_{i1}\,|A_{i1}|$$

וכל אחד מארבעת המחוברים באגף ימין אפשר לכתוב, לפי דוגמה (ב) כסכום של 6 מכפלות של הרכיבים של A. מכאן ש־$|A|$ הוא סכום של 24 מחוברים, כל אחד מכפלה של 4 רכיבים של A. ראה גם דוגמה **5.18** להלן.

(שים לב שארבעת הרכיבים האלה הם תמיד משורות שונות ומעמודות שונות של A. מספר המכפלות של ארבעה רכיבים שיש להן תכונה זו הוא בדיוק 24. לכן כל המכפלות האלה מופיעות בפיתוח הדטרמיננטה. מחציתן מופיעות עם סימן $+$ ומחציתן עם סימן $-$. לא נסביר כאן מהו הכלל לקביעת הסימן.)

תרגיל 5.16: $A^t_{ij}=(A_{ji})^t$ לכל $1\le i,j\le n$.

הוכחה: אגף שמאל נוצר מ־A על ידי הפיכת שורות לעמודות ומחיקת השורה ה־i והעמודה ה־j. אגף ימין נוצר מ־A על ידי מחיקת השורה ה־j והעמודה ה־i והפיכת שורות לעמודות. ברור ששניהם שווים. ∎

משפט 5.17 (פיתוח הדטרמיננטה לפי השורה ה־i): *יהי* $1\le i\le n$. *אז לכל* $A\in \mathrm{M}_n(F)$

$$|A|=\sum_{j=1}^{n}(-1)^{i+j}(A)_{ij}\,|A_{ij}|\,. \qquad (4)$$

הוכחה: נשתמש בנוסחת הפיתוח של $|A^t|$ לפי העמודה ה־i ובתרגיל **5.16**:

$$|A|=|A^t|=\sum_{j=1}^{n}(-1)^{j+i}(A^t)_{ji}|A^t_{ji}|=\sum_{j=1}^{n}(-1)^{i+j}(A)_{ij}|A_{ij}|.$$

דוגמה 5.18: נִקח $i=4$:

$$\begin{aligned}
&\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14}\\ a_{21} & a_{22} & a_{23} & a_{24}\\ a_{31} & a_{32} & a_{33} & a_{34}\\ a_{41} & a_{42} & a_{43} & a_{44}\end{vmatrix}\\
={}&-a_{41}(a_{12}a_{23}a_{34}+a_{22}a_{33}a_{14}+a_{32}a_{13}a_{24}-a_{14}a_{23}a_{32}-a_{24}a_{33}a_{12}-a_{34}a_{13}a_{22})\\
&+a_{42}(a_{11}a_{23}a_{34}+a_{21}a_{33}a_{14}+a_{31}a_{13}a_{24}-a_{14}a_{23}a_{31}-a_{24}a_{33}a_{11}-a_{34}a_{13}a_{21})\\
&-a_{43}(a_{11}a_{22}a_{34}+a_{21}a_{32}a_{14}+a_{31}a_{12}a_{24}-a_{14}a_{22}a_{31}-a_{24}a_{32}a_{11}-a_{34}a_{12}a_{21})\\
&+a_{44}(a_{11}a_{22}a_{33}+a_{21}a_{32}a_{13}+a_{31}a_{12}a_{23}-a_{13}a_{22}a_{31}-a_{23}a_{32}a_{11}-a_{33}a_{12}a_{21})\\
={}&-a_{12}a_{23}a_{34}a_{41}-a_{14}a_{22}a_{33}a_{41}-a_{13}a_{24}a_{32}a_{41}+a_{14}a_{23}a_{32}a_{41}+a_{12}a_{24}a_{33}a_{41}\\
&+a_{13}a_{22}a_{34}a_{41}+a_{11}a_{23}a_{34}a_{42}+a_{14}a_{21}a_{33}a_{42}+a_{13}a_{24}a_{31}a_{42}-a_{14}a_{23}a_{31}a_{42}\\
&-a_{11}a_{24}a_{33}a_{42}-a_{13}a_{21}a_{34}a_{42}-a_{11}a_{22}a_{34}a_{43}-a_{14}a_{21}a_{32}a_{43}-a_{12}a_{24}a_{31}a_{43}\\
&+a_{14}a_{22}a_{31}a_{43}+a_{11}a_{24}a_{32}a_{43}+a_{12}a_{21}a_{34}a_{43}+a_{11}a_{22}a_{33}a_{44}+a_{13}a_{21}a_{32}a_{44}\\
&+a_{12}a_{23}a_{31}a_{44}-a_{13}a_{22}a_{31}a_{44}-a_{11}a_{23}a_{32}a_{44}-a_{12}a_{21}a_{33}a_{44}
\end{aligned}$$

הגדרה 5.19: תהי $A \in \mathrm{M}_n(F)$. **המטריצה המצורפת ל-**A $\mathrm{adj}\, A \in \mathrm{M}_n(F)$ מוגדרת על ידי

$$((\mathrm{adj}\, A)^t)_{ij} = (-1)^{i+j}|A_{ij}| \qquad \text{כלומר} \qquad (\mathrm{adj}\, A)_{ij} = (-1)^{i+j}|A_{ji}|.$$

דוגמה 5.20: אם $A = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$ אז

$$\mathrm{adj}\, A = \begin{vmatrix} |A_{11}| & -|A_{21}| \\ -|A_{12}| & |A_{22}| \end{vmatrix} = \begin{vmatrix} d & -b \\ -c & a \end{vmatrix}.$$

אם $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ אז $(\mathrm{adj}\, A)^t = \begin{pmatrix} ei-fh & fg-di & dh-eg \\ hc-bi & ai-cg & bg-ah \\ bf-ec & dc-af & ae-bd \end{pmatrix}$.

לכן $\mathrm{adj}\, A = \begin{pmatrix} ei-fh & hc-bi & bf-ec \\ fg-di & ai-cg & dc-af \\ dh-eg & bg-ah & ae-bd \end{pmatrix}$.

תרגיל 5.21: $\mathrm{adj}\, A^t = (\mathrm{adj}\, A)^t$.

הוכחה: $(\mathrm{adj}\, A^t)_{ij} = (-1)^{i+j}|{A^t}_{ji}| = (-1)^{i+j}|(A_{ij})^t| = (-1)^{i+j}|A_{ij}| = ((\mathrm{adj}\, A)^t)_{ij}$. ∎

משפט 5.22: *תהי* $A \in \mathrm{M}_n(F)$.

(א) $A\, \mathrm{adj}\, A = |A| I_n = \begin{pmatrix} |A| & & 0 \\ & \ddots & \\ 0 & & |A| \end{pmatrix}$.

(ב) $\mathrm{adj}\, A\; A = |A| I_n$.

הוכחה: (א) יהיו $1 \le i,j \le n$. צריך להוכיח: $(A\, \mathrm{adj}\, A)_{ij} = \begin{cases} |A| & i=j \\ 0 & i \ne j \end{cases}$.

$$(A\, \mathrm{adj}\, A)_{ij} = \sum_{k=1}^{n} (A)_{ik} (\mathrm{adj}\, A)_{kj} = \sum_{k=1}^{n} (-1)^{j+k} (A)_{ik} |A_{jk}|.$$

אם $i = j$, הביטוי באגף ימין הוא בדיוק הפיתוח של $|A|$ לפי השורה ה-j, לכן שווה ל-$|A|$.

אם $i \ne j$, תהי A' המטריצה המתקבלת מ-A, אם רושמים את השורה ה-i של A במקום השורה ה-j של A. אז ב-A' שתי שורות זהות, לכן $|A'| = 0$. כמו כן, לכל k,

$$(A')_{jk} = (A)_{ik}, \qquad A'_{jk} = A_{jk}.$$

אם נציב זאת בחישוב לעיל, נקבל (פיתוח של A' לפי השורה ה-j)

$$(A\, \mathrm{adj}\, A)_{ij} = \sum_{k=1}^{n} (-1)^{j+k} (A')_{jk} |A'_{jk}| = |A'| = 0.$$

(ב) לפי תרגיל 5.21 ולפי (א) מיושם על A^t,

$$(\mathrm{adj}\, A\; A)^t = A^t (\mathrm{adj}\, A)^t = A^t (\mathrm{adj}\, A^t) = |A^t| I_n = |A| I_n = (|A| I_n)^t.$$

ומכאן הטענה. ∎

תוצאה 5.23: אם $|A| \neq 0$ אז $A^{-1} = \frac{1}{|A|}\text{adj}\, A$.

דוגמה 5.24: $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

משפט 5.25 (Cramer): *תהי $AX = b$ מערכת של n משוואות לינאריות ב־n נעלמים מעל שדה F. נניח $|A| \neq 0$. אז למערכת פתרון יחיד $c = (c_1, c_2, \ldots, c_n)^t$ הנתון על ידי הנוסחה $c_i = \frac{|A_i|}{|A|}$ באשר $A_i \in \mathrm{M}_n(F)$ מתקבלת מ־A על ידי החלפת העמודה ה־i של A בעמודה b.*

הוכחה: אם $|A| \neq 0$ אז A הפיכה ולכן למערכת יש פתרון יחיד c. אז $Ac = b$ ולכן $c = A^{-1}b$. לפי תוצאה 5.23, $c = \frac{1}{|A|}(\text{adj}\, A)b$. נניח $b = (b_1, b_2, \ldots, b_n)^t \in F^n$ ונסמן $B = A_j$. לפי כלל הכפל של מטריצות,

$$c_i = \frac{1}{|A|}\sum_{k=1}^{n}(\text{adj}\, A)_{ik}b_k = \frac{1}{|A|}\sum_{k=1}^{n}(-1)^{i+k}|A_{ki}|b_k = \frac{1}{|A|}\sum_{k=1}^{n}(-1)^{i+k}|B_{ki}|(B)_{ki} = \frac{1}{|A|}|B|.$$

השוויון האחרון נובע מפיתוח של $|B|$ לפי העמודה ה־i. ∎

דוגמה 5.26: למערכת

$$2X + 3Y = 4$$

$$5X + 6Y = 1$$

פתרון יחיד $(c_1, c_2)^t$, באשר

$$c_1 = \frac{\begin{vmatrix} 4 & 3 \\ 1 & 6 \end{vmatrix}}{\begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix}} = \frac{4\cdot 6 - 3\cdot 1}{2\cdot 6 - 3\cdot 5} = -7, \qquad c_2 = \frac{\begin{vmatrix} 2 & 4 \\ 5 & 1 \end{vmatrix}}{\begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix}} = \frac{2\cdot 1 - 4\cdot 5}{2\cdot 6 - 3\cdot 5} = 6.$$

6. העתקות לינאריות

דברים כלליים אודות העתקות (= פונקציות, טרנספורמציות) בין קבוצות:

סימון 6.1: $T\colon V \to W$ או $V \xrightarrow{T} W$ מסמן העתקה בשם T שמעתיקה איברי קבוצה V ("התחום") לתוך הקבוצה W ("הטווח"). אפשר גם לכתוב $V \to W$ אם אין רוצים לציין את שם ההעתקה במפורש. לכל $x \in V$, הסימון $T(x)$ מסמן את האיבר בקבוצה W אליו מועתק x. איבר זה נקרא **התמונה של** x **תחת** T.

גם הסימון $\mapsto$ נועד להגדיר העתקה, בלי לציין במפורש את שמה, את תחומה ואת טווחה. למשל, ההעתקה $x \mapsto x^2$ היא ההעתקה שמעתיקה כל איבר לריבוע שלו.

שים לב לכתיב: מימין ומשמאל ל־$\to$ עומדות קבוצות, בעוד שמימין ומשמאל ל־$\mapsto$ עומדים איברים!

עבור תת קבוצה U של V, הסימון $T(U)$ מסמן את הקבוצה $\{T(u) |\ u \in U\}$.

העתקת הזהות 1_V של קבוצה V מוגדרת על ידי $1_V(v) = v$ לכל $v \in V$.

הרכבה של העתקות: תהיינה $S\colon U \to V,\ T\colon V \to W$ שתי העתקות. נגדיר העתקה $T \circ S\colon U \to W$ על ידי $T \circ S(u) = T(S(u))$ לכל $u \in U$.

נעיר שלא בהכרח $T \circ S = S \circ T$.

שתי העתקות $S, T\colon V \to W$ הן שוות אם $S(v) = T(v)$ לכל $v \in V$.

הגדרה 6.2: העתקה $T\colon V \to W$ נקראת

(א) **חד חד ערכית** (חח"ע) אם לכל $v, v' \in V$ מתקיים: אם $T(v) = T(v')$ אז $v = v'$.

(ב) **על**, אם לכל $w \in W$ יש $v \in V$ כך ש־$T(v) = w$.

אם $T\colon V \to W$ חח"ע ועל, יש לה העתקה **הופכית** $T^{-1}\colon W \to V$ אשר מוגדרת באופן הבא: לכל $w \in W$, $T^{-1}(w)$ הוא אותו האיבר היחיד של V עבורו $T(v) = w$. (איבר v כזה קיים כי T על; הוא יחיד כי T חח"ע.) קל לראות שגם T^{-1} חח"ע ועל. כמו כן $T^{-1} \circ T = 1_V$ ו־$T \circ T^{-1} = 1_W$.

הגדרה 6.3: יהיו V, W שני מרחבים וקטוריים מעל אותו שדה F. העתקה $T\colon V \to W$ נקראת **לינארית** אם היא

(1) שומרת חיבור: $T(v + v') = T(v) + T(v')$, לכל $v, v' \in V$.

(2) שומרת כפל בסקלר: $T(\alpha v) = \alpha T(v)$, לכל $v \in V$ ולכל $\alpha \in F$.

דוגמאות 6.4: (א) נגדיר $T\colon F^2 \to F^3$ על ידי $T(\begin{pmatrix} x \\ y \end{pmatrix}) = \begin{pmatrix} 2x+3y \\ x-y \\ -x \end{pmatrix}$. אז T לינארית.

אכן, היא שומרת חיבור:

$$T(\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x' \\ y' \end{pmatrix}) = T(\begin{pmatrix} x+x' \\ y+y' \end{pmatrix}) = \begin{pmatrix} 2(x+x')+3(y+y') \\ (x+x')-(y+y') \\ -(x+x') \end{pmatrix} =$$

$$= \begin{pmatrix} 2x+3y \\ x-y \\ -x \end{pmatrix} + \begin{pmatrix} 2x'+3y' \\ x'-y' \\ -x' \end{pmatrix} = T(\begin{pmatrix} x \\ y \end{pmatrix}) + T(\begin{pmatrix} x' \\ y' \end{pmatrix})$$

והיא שומרת כפל בסקלר:

$$.T(\alpha\begin{pmatrix}x\\y\end{pmatrix})=T(\begin{pmatrix}\alpha x\\\alpha y\end{pmatrix})=\begin{pmatrix}2\alpha x+3\alpha y\\\alpha x-\alpha y\\-\alpha x\end{pmatrix}=\alpha\begin{pmatrix}2x+3y\\x-y\\-x\end{pmatrix}=\alpha T(\begin{pmatrix}x\\y\end{pmatrix})$$

(ב) העתקת הזהות 1_V של מרחב וקטורי V היא לינארית.

(ג) אם V,W מרחבים וקטוריים כלשהם מעל שדה F, **העתקת האפס** $S\colon V\to W$ המוגדרת על ידי $S(v)=0$ לכל $v\in V$, היא לינארית.

(ד) תהי $A\in \mathrm{M}_{m\times n}(F)$. נגדיר $T_A\colon F^n\to F^m$ על ידי $T_A(v)=Av$. אז T_A לינארית. אכן, לכל $v,v'\in F^n$ ולכל $\alpha\in F$

(1) $.T_A(v+v')=A(v+v')=Av+Av'=T_A(v)+T_A(v')$

(2) $.T_A(\alpha v)=A(\alpha v)=\alpha Av=\alpha T_A(v)$

(ה) הסיבוב $S\colon\mathbb{R}^2\to\mathbb{R}^2$ במישור סביב הראשית 0 בזוית נתונה θ היא העתקה לינארית.

למה 6.5: *תהי $T\colon V\to W$ העתקה לינארית של מרחבים וקטוריים מעל שדה F. אז*

(א) $.T(0)=0$

(ב) $T(-v)=-T(v)$ *לכל* $.v\in V$

(ג) $T(a_1v_1+\cdots+a_nv_n)=a_1T(v_1)+\cdots+a_nT(v_n)$ *לכל* $v_1,\dots v_n\in V$ *ולכל* $.a_1,\dots a_n\in F$

הוכחה: (א) $0+T(0)=T(0)=T(0+0)=T(0)+T(0)$, ולפי כלל הצמצום $.T(0)=0$

(ב) $T(-v)+T(v)=T(-v+v)=T(0)=0$, לכן $.T(-v)=-T(v)$

(ג) ברור. ∎

למה 6.6: *אם $T\colon V\to W$ העתקה לינארית חח"ע ועל אז גם $T^{-1}\colon W\to V$ לינארית וחח"ע ועל.*

הוכחה: ההעתקה ההופכית תמיד חח"ע ועל. נראה את הלינאריות. יהיו $w,w'\in W$ ויהי $\alpha\in F$. יהיו $v=T^{-1}(w),v'=T^{-1}(w')$. אז $T(v)=w,T(v')=w'$, לפי הגדרת T^{-1}. לכן, כיוון ש־T לינארית,

$$T(v+v')=T(v)+T(v')=w+w',\quad T(\alpha v)=\alpha T(v)=\alpha w$$

ולכן, לפי הגדרת T^{-1},

$$.T^{-1}(w+w')=v+v'=T^{-1}(w)+T^{-1}(w'),\quad T^{-1}(\alpha w)=\alpha v=\alpha T^{-1}(w)$$ ∎

העתקה לינארית חח"ע ועל נקראת **איזומורפיזם**. נביא להלן דוגמה לאיזומורפיזם.

הגדרה 6.7: יהי V מרחב וקטורי מעל שדה F ממימד n ויהי $\mathcal{B} = (v_1, v_2, \ldots, v_n)$ בסיסו. כזכור (משפט 4.38), לכל $v \in V$ יש $a_1, a_2, \ldots, a_n \in F$ יחידים כך ש־$v = a_1 + a_2v_2 + \cdots + a_nv_n$. נסמן $[v]_{\mathcal{B}} = (a_1, a_2, \ldots, a_n)^t \in F^n$ ונקרא ל־$[v]_{\mathcal{B}}$ **וקטור הקואורדינטות של** v **לפי** $\mathcal{B}$. (שים לב: הגדרה זו תלויה בסדר איברי $\mathcal{B}$. כדי להדגיש זאת נדבר על $\mathcal{B}$ כבסיס סדור.)

דוגמה 6.8:

(א) אם $V = F^2$ ו־$\mathcal{B}$ מורכב מ־$\begin{pmatrix}2\\3\end{pmatrix}, \begin{pmatrix}1\\2\end{pmatrix}$, אז $\left[\begin{pmatrix}4\\7\end{pmatrix}\right]_{\mathcal{B}} = \begin{pmatrix}2\\1\end{pmatrix}$, כי $\begin{pmatrix}4\\7\end{pmatrix} = 2\begin{pmatrix}1\\2\end{pmatrix} + 1\begin{pmatrix}2\\3\end{pmatrix}$.

(ב) אם $V = F^m$ ו־$\mathcal{S} = (e_1, \ldots, e_m)$, אז $[(a_1, \ldots, a_m)^t]_{\mathcal{S}} = (a_1, \ldots, a_m)^t$. כלומר, $[v]_S = v$ לכל $v \in V$.

(ג) אם V מרחב הפולינומים ממעלה $\geq n$ עם הבסיס $\mathcal{B} = (1, X, X^2, \ldots, X^n)$, אז

$$[a_nX^n + \cdots + a_1X + a_0]_{\mathcal{B}} = (a_0, a_1, \ldots, a_n)^t.$$

המשפט הבא מהווה דוגמה לאיזומורפיזם.

משפט 6.9: *יהי V מרחב וקטורי ממימד n מעל שדה F, ויהי $\mathcal{B} = (v_1, \ldots, v_n)$ בסיס סדור של V. אז ההעתקה $V \to F^n$ המוגדרת על ידי $v \mapsto [v]_{\mathcal{B}}$ היא איזומורפיזם.*

הוכחה: ההעתקה על: אם $(a_1, \ldots, a_n)^t \in F^n$, נגדיר $v = a_1v_1 + \cdots + a_nv_n$. אז $[v]_{\mathcal{B}} = (a_1, \ldots, a_n)^t$.

ההעתקה חח"ע: אם $[v]_{\mathcal{B}} = (a_1, \ldots, a_n)^t = [v']_{\mathcal{B}}$, אז $v = a_1v_1 + \cdots + a_nv_n = v'$.

כדי להוכיח לינאריות, צריך להראות לכל $v, v' \in V$ ולכל $\alpha \in F$, כי

$$[v + v']_{\mathcal{B}} = [v]_{\mathcal{B}} + [v']_{\mathcal{B}}, \quad [\alpha v]_{\mathcal{B}} = \alpha[v]_{\mathcal{B}} \tag{4}$$

ואכן, נניח כי

$$v = a_1v_1 + \cdots + a_nv_n, \quad v' = a_1'v_1 + \cdots + a_n'v_n.$$

אז

$$\alpha v = \alpha a_1v_1 + \cdots + \alpha a_nv_n, \quad v + v' = (a_1 + a_1')v_1 + \cdots + (a_n + a_n')v_n.$$

מכאן

$$[v]_{\mathcal{B}} = \begin{pmatrix}a_1\\\vdots\\a_n\end{pmatrix}, [v']_{\mathcal{B}} = \begin{pmatrix}a_1'\\\vdots\\a_n'\end{pmatrix}, [v + v']_{\mathcal{B}} = \begin{pmatrix}a_1 + a_1'\\\vdots\\a_n + a_n'\end{pmatrix}, [\alpha v]_{\mathcal{B}} = \begin{pmatrix}\alpha a_1\\\vdots\\\alpha a_n\end{pmatrix}.$$

מכאן נובע (4) בקלות. ∎

משפט 6.10: *יהיו V, W מרחבים וקטוריים מעל שדה F, תהי $T\colon V \to W$ העתקה לינארית, ויהיו $v_1, \dots v_m \in V$. אז*

(א) *$T(\mathrm{Sp}(v_1, \dots, v_m)) = \mathrm{Sp}(T(v_1), \dots, T(v_m))$. בפרט, $T(\mathrm{Sp}(v_1, \dots, v_m))$ הוא תת מרחב של W.*

(ב) *אם $v_1, \dots, v_m$ סדרת יוצרים של V אז $T(v_1), \dots, T(v_m)$ סדרת יוצרים של $T(V)$.*

(ג) *אם $T(v_1), \dots, T(v_m)$ בלתי תלויים לינארית, אז גם $v_1, \dots v_m$ בלתי תלויים לנארית.*

הוכחה: (א) $\mathrm{Sp}(v_1, \dots, v_m) = \{a_1v_1 + \cdots + a_mv_m |\, a_1, \dots, a_m \in F\}$, לכן

$$\begin{aligned} T(\mathrm{Sp}(v_1, \dots, v_m)) &= \{T(a_1v_1 + \cdots + a_mv_m) |\, a_1, \dots, a_m \in F\} = \\ &= \{a_1T(v_1) + \cdots + a_mT(v_m) |\, a_1, \dots, a_m \in F\} = \\ &= \mathrm{Sp}(T(v_1), \dots, T(v_m)) \end{aligned}$$

(ב) נתון ש־$V = \mathrm{Sp}(v_1, \dots, v_m)$. לפי (א), $T(V) = \mathrm{Sp}(T(v_1), \dots, T(v_m))$, כלומר, $T(v_1), \dots, T(v_m)$ סדרת יוצרים של $T(V)$.

(ג) יהיו $a_1, \dots, a_n \in F$ כך ש־$a_1v_1 + \cdots + a_mv_m = 0$. נפעיל T על שני האגפים ונקבל, בגלל הלינאריות,

$$a_1T(v_1) + \cdots + a_mT(v_m) = 0.$$

כיוון ש־$T(v_1), \dots, T(v_m)$ בלתי תלויים לינארית, $a_1 = \cdots = a_n = 0$. ∎

מסקנה 6.11: *יהיו V, W מרחבים וקטוריים מעל שדה F, יהי $T\colon V \to W$ איזומורמפיזם, ויהיו $v_1, \dots, v_m \in V$. אז $v_1, \dots, v_m$ בסיס של V אם ורק אם $T(v_1), \dots, T(v_m)$ בסיס של W.*

בפרט, אם אחד מבין V, W נוצר סופית, אז גם השני נוצר סופית ו־$\dim V = \dim W$.

הוכחה: די להוכיח שאם $v_1, \dots, v_m$ בסיס של V אז $T(v_1), \dots, T(v_m)$ בסיס של W, כי הכיוון ההפוך נובע מהפעלת T^{-1} על הבסיס $T(v_1), \dots, T(v_m)$.

כיוון ש־T על, $W = T(V)$. לפי (ב) של המשפט הקודם, $T(v_1), \dots, T(v_m)$ סדרת יוצרים של W.

כיוון ש־$v_1 = T^{-1}(T(v_1)), \dots, v_n = T^{-1}(T(v_n))$ היא בלתי תלויה לינארית לפי (ב) של המשפט הקודם (מיושם על העתקה T^{-1}), גם $T(v_1), \dots, T(v_m)$ בלתי תלויה לינארית. ∎

הגדרה 6.12: יהיו V, W מרחבים וקטוריים מעל שדה F ותהי $T\colon V \to W$ העתקה לינארית.

(א) **הגרעין** $\operatorname{Ker} T$ של T היא הקבוצה $\operatorname{Ker} T = \{v \in V |\, T(v) = 0\}$

(ב) **התמונה** $\operatorname{Im} T$ של T היא הקבוצה

$$\operatorname{Im} T = \{w \in W |\, \text{קיים } v \in V \text{ כך ש־} T(v) = w\} = \{T(v) |\, v \in V\} = T(V)$$

למה 6.13: *תהי* $T\colon V\to W$ *העתקה לינארית.*

(א) $\operatorname{Ker} T$ *הוא תת מרחב של* V.

(ב) $\operatorname{Im} T$ *הוא תת מרחב של* W.

הוכחה: בדיקה ישירה, לפי ההגדרות. ∎

משפט 6.14: *תהי* $T\colon V\to W$ *העתקה לינארית. אז* T *חח"ע אם ורק אם* $\operatorname{Ker} T=\{0\}$.

הוכחה: נניח כי T חח"ע. יהי $v\in \operatorname{Ker} T$. אז $T(0)=0=T(v)$ ולכן $v=0$. מכאן $\operatorname{Ker} T\subseteq\{0\}$. ההכלה ההפוכה נכונה לכל העתקה לינארית (גם אם אינה חח"ע), לכן $\operatorname{Ker} T=\{0\}$.

להיפך, נניח כי $\operatorname{Ker} T=\{0\}$. יהיו $v,v'\in V$ כך ש־$T(v)=T(v')$. אז

$$T(v-v')=T(v)-T(v')=0,$$

ולכן $v-v'\in \operatorname{Ker} T=\{0\}$, כלומר $v-v'=0$. מכאן $v=v'$. לכן T חח"ע. ∎

הערה 6.15: יהי F שדה ותהי $A\in \mathrm{M}_{m\times n}(F)$ מטריצה. נגדיר $V=F^n, W=F^m$ ותהי $T_A\colon V\to W$ ההעתקה הלינארית הנתונה על ידי $T_A(v)=Av$. אז

$$\operatorname{Ker} T_A=\{v\ \in V|\ Av=0\}=\operatorname{Ker} A\,.$$

בפרט, $\dim\operatorname{Ker} T_A=n-\operatorname{rk} A$.

כמו כן

$$\begin{aligned}\operatorname{Im} T_A&=T_A(V)=T_A(\operatorname{Sp}(e_1,\dots,e_n))\\&=\operatorname{Sp}(T_A(e_1),\dots,T_A(e_n))=\operatorname{Sp}(Ae_1,\dots,Ae_n)=\operatorname{Col}(A)\end{aligned}$$

כי Ae_j היא העמודה ה־j של A, לכל $1\le j\le n$. בפרט, $\dim\operatorname{Im} T_A=\operatorname{rk} A$.

אנו רואים ש־

$$\dim\operatorname{Ker} T_A+\dim\operatorname{Im} T_A=\dim V\,.$$

דבר זה נכון באופן כללי:

משפט 6.16 (משפט המימד): *יהיו* V,W *מרחבים וקטוריים מעל שדה* F *ותהי* $T\colon V\to W$ *העתקה לינארית. נניח כי* V *נוצר סופית (כלומר, יש לו בסיס). אז* $\dim\operatorname{Ker} T+\dim\operatorname{Im} T=\dim V$.

הוכחה: $\operatorname{Ker} T$ הוא תת מרחב של V, לכן גם הוא נוצר סופית. יהי $v_1,\dots,v_m$ בסיסו. זוהי בפרט סדרה בלתי תלויה לינארית ב־V, ולכן ניתן להשלים אותה לבסיס של V, נאמר

$$v_1,\dots,v_m,v_{m+1}\dots,v_n.$$

היות ו־$\dim V=n,\ \dim\operatorname{Ker} T=m$, עלינו להוכיח: $\dim\operatorname{Im} T=n-m$. לשם כך די אם נוכיח:

טענה: $T(v_{m+1}),\dots,T(v_n)$ הוא בסיס של $\operatorname{Im} T$.

ואכן,

$$\begin{aligned}\operatorname{Im} T = T(V) = T(\operatorname{Sp}(v_1,\dots,v_m,\dots,v_n)) &= \operatorname{Sp}(T(v_1),\dots,T(v_m),\dots,T(v_n))\\ &= \operatorname{Sp}(\overbrace{0,\dots,0}^{m},T(v_{m+1}),\dots,T(v_n)) = \operatorname{Sp}(T(v_{m+1}),\dots,T(v_n))\end{aligned}$$

(כאשר השויון האחרון נכון, כי מתוך סדרה של וקטורים אפשר להשמיט אפסים מבלי לשנות את תת המרחב הנוצר).
לכן $T(v_{m+1}),\dots,T(v_n)$ סדרת יוצרים של $\operatorname{Im} T$.

נראה שהיא גם בלתי תלויה לינארית: יהיו $a_{m+1},\dots,a_n \in F$ כך שמתקיים

$$a_{m+1}T(v_{m+1}) + \cdots + a_nT(v_n) = 0.$$

בגלל הלינאריות, אגף שמאל שווה ל- $T(a_{m+1}v_{m+1}+\cdots+a_nv_n)$. לכן $a_{m+1}v_{m+1}+\cdots+a_nv_n \in \operatorname{Ker} T$.
מכאן שקיימים $a_1,\dots,a_m \in F$ כך ש-

$$a_1v_1 + \cdots + a_mv_m = a_{m+1}v_{m+1} + \cdots + a_nv_n,$$

כלומר

$$(-a_1)v_1 + \cdots + (-a_m)v_m + a_{m+1}v_{m+1} + \cdots + a_nv_n = 0.$$

אך $v_1,\dots,v_n$ בלתי תלויה לינארית, לכן $a_{m+1} = \cdots = a_n = 0$. ∎

מסקנה 6.17: *יהיו V,W מרחבים וקטוריים מאותו מימד מעל שדה F ותהי $T\colon V \to W$ העתקה לינארית. אז T חח"ע אם ורק אם T על.*

הוכחה: ההעתקה T חח"ע אם ורק אם $\operatorname{Ker} T = \{0\}$. היות ותמיד $\{0\} \subseteq \operatorname{Ker} T$ תת מרחבים, זה שקול לשוויון הממדים $0 = \dim\{0\} = \dim \operatorname{Ker} T$.

ההעתקה T על אם ורק אם $\operatorname{Im} T = W$. היות ותמיד $\operatorname{Im} T \subseteq W$ תת מרחבים, זה שקול לשוויון הממדים $\dim \operatorname{Im} T = \dim W = \dim V$.

לכן צריך להוכיח: $\dim \operatorname{Ker} T = 0$ אם ורק אם $\dim \operatorname{Im} T = \dim V$. וזה אכן נובע מהנוסחה $\dim \operatorname{Ker} T + \dim \operatorname{Im} T = \dim V$ של משפט 6.16. ∎

הגדרה 6.18: יהיו V,W שני מרחבים וקטוריים מעל אותו שדה F, יהי $\mathcal{B} = (v_1,\dots,v_n)$ בסיס של V, ויהי $\mathcal{C} = (w_1,\dots,w_m)$ בסיס של W. לכל העתקה לינארית $T\colon V \to W$ נגדיר מטריצה $[T]_{\mathcal{C}}^{\mathcal{B}} \in \mathrm{M}_{m\times n}(F)$ על ידי

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

במלים אחרות: אם $A = (a_{ij}) \in \mathrm{M}_{m\times n}(F)$ אז $[T]_{\mathcal{C}}^{\mathcal{B}} = A$ אם ורק אם

$$T(v_j) = a_{1j}w_1 + \cdots + a_{mj}w_m, \qquad j = 1,\dots,n.$$

דוגמאות 6.19:

(א) נגדיר $T\colon \mathbb{R}^3 \to \mathbb{R}^2$ על ידי $T(\begin{pmatrix} x \\ y \\ z \end{pmatrix}) = \begin{pmatrix} 2x-y \\ x+3z \end{pmatrix}$. אז T לינארית. (בדוק!) יהיו $\mathcal{B} = (v_1, v_2, v_3)$ בסיס של $\mathbb{R}^3$, ו־$\mathcal{C} = (w_1, w_2)$ בסיס של $\mathbb{R}^2$, באשר

$$w_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, w_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad v_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

אז

$$T(v_1) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, T(v_2) = \begin{pmatrix} 2 \\ 4 \end{pmatrix}, T(v_3) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

בדיקה קלה מראה כי

$$T(v_1) = -w_1, \quad T(v_2) = -2w_1 + 4w_2, \quad T(v_3) = -w_1 + w_2,$$

לכן

$$[T]^{\mathcal{B}}_{\mathcal{C}} = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 4 & 1 \end{pmatrix}.$$

(ב) יהי $S\colon \mathbb{R}^2 \to \mathbb{R}^2$ הסיבוב במישור סביב הראשית 0 בזוית נתונה θ, ויהי $\mathcal{B} = (e_1, e_2)$ הבסיס הסטנדרטי של $\mathbb{R}^2$. נזכור שלכל $v \in \mathbb{R}^2$ מתקיים: $[v]_{\mathcal{B}} = v$. (זה לא יהיה נכון בבסיס אחר!) לכן

$$[S]^{\mathcal{B}}_{\mathcal{B}} = \big([S(e_1)]_{\mathcal{B}}, [S(e_2)]_{\mathcal{B}}\big) = \big(S(e_1), S(e_2)\big) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

(ג) תהי $A \in \mathrm{M}_{m\times n}(F)$ ותהי $T_A\colon F^n \to F^m$ ההעתקה $v \mapsto Av$. יהיו $\mathcal{B}, \mathcal{C}$ הבסיסים הסטנדרטיים של F^n, F^m, בהתאמה. אז, היות ו־Ae_j היא העמודה ה־j של A,

$$[T]^{\mathcal{B}}_{\mathcal{C}} = \big([Ae_1]_{\mathcal{C}}, \ldots, [Ae_n]_{\mathcal{C}}\big) = \big(Ae_1, \ldots, Ae_n\big) = A$$

(ד) תהי $0\colon V \to W$ העתקת האפס ויהיו $\mathcal{B}, \mathcal{C}$ בסיסים כלשהם. אז $[T]^{\mathcal{B}}_{\mathcal{C}} = 0 = (0)$.

(ה) יהי V מרחב וקטורי ממימד n, תהי $1_V\colon V \to V$ העתקת הזהות ויהי $\mathcal{B}$ בסיס כלשהו של V. אז $[1_V]^{\mathcal{B}}_{\mathcal{B}} = I_n$.

משפט 6.20: *בתנאים ובסימונים של ההגדרה 6.18, לכל $v \in V$ מתקיים*

$$[T(v)]_{\mathcal{C}} = [T]^{\mathcal{B}}_{\mathcal{C}} [v]_{\mathcal{B}}.$$

הוכחה: נניח

$$[v]_{\mathcal{B}} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \in F^n, \qquad [T]^{\mathcal{B}}_{\mathcal{C}} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \in \mathrm{M}_{m\times n}(F)$$

אז $v = c_1v_1 + c_2v_2 + \cdots + c_nv_n$ ולכן $T(v) = c_1T(v_1) + c_2T(v_2) + \cdots + c_nT(v_n)$, ואילו

$$[T(v_1)]_{\mathcal{C}} = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, [T(v_2)]_{\mathcal{C}} = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} \ldots, [T(v_n)]_{\mathcal{C}} = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

ולכן

$$T(v_1) = a_{11}w_1 + a_{21}w_2 + \cdots + a_{m1}w_m$$
$$T(v_2) = a_{12}w_1 + a_{22}w_2 + \cdots + a_{m2}w_m$$
$$\cdots$$
$$T(v_n) = a_{1n}w_1 + a_{2n}w_2 + \cdots + a_{mn}w_m$$

נכפיל משוואות אלה ב־$c_1, c_2, \ldots, c_n$ בהתאמה ונחבר אותן:

$$\begin{aligned} T(v) = c_1T(v_1) + c_2T(v_2) + \cdots + c_nT(v_n) = \\ &= (a_{11}c_1 + a_{12}c_2 + \cdots + a_{1n}c_n)w_1 \\ &+ (a_{21}c_1 + a_{22}c_2 + \cdots + a_{2n}c_n)w_2 \\ &\cdots\cdots\cdots \\ &+ (a_{m1}c_1 + a_{m2}c_2 + \cdots + a_{mn}c_n)w_m \end{aligned}$$

מכאן שהעמודה $[T(v)]_{\mathcal{C}}$ היא

$$\begin{pmatrix} a_{11}c_1 + a_{12}c_2 + \cdots + a_{1n}c_n \\ a_{21}c_1 + a_{22}c_2 + \cdots + a_{2n}c_n \\ \cdots\cdots\cdots \\ a_{m1}c_1 + a_{m2}c_2 + \cdots + a_{mn}c_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = [T]^{\mathcal{B}}_{\mathcal{C}}[v]_{\mathcal{B}} \quad \blacksquare$$

דוגמה 6.21: יהי $S\colon \mathbb{R}^2 \to \mathbb{R}^2$ הסיבוב במישור סביב הראשית 0 בזוית θ, ויהי $\mathcal{C}$ הבסיס הסטנדרטי של $\mathbb{R}^2$. נזכור שלכל $v \in \mathbb{R}^2$ מתקיים: $[v]_{\mathcal{C}} = v$. אז

$$S\begin{pmatrix} x \\ y \end{pmatrix} = \left[S\begin{pmatrix} x \\ y \end{pmatrix}\right]_{\mathcal{B}} = [S]^{\mathcal{B}}_{\mathcal{B}}\left[\begin{pmatrix} x \\ y \end{pmatrix}\right]_{\mathcal{B}} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

תרגיל 6.22: *תהי $T\colon V \to W$ העתקה לינארית, ויהיו $\mathcal{B}$ בסיס של V ו־$\mathcal{C}$ בסיס של W.*

(א) $\mathrm{rk}[T]^{\mathcal{B}}_{\mathcal{C}} = \dim \mathrm{Im}\, T$.

(ב) *T איזומורפיזם אם ורק אם $[T]^{\mathcal{B}}_{\mathcal{C}}$ הפיכה (ובפרט ריבועית).*

הוכחה: יהי $\mathcal{B} = (v_1, \ldots, v_n)$, ו־$\mathcal{C} = (w_1, \ldots, w_m)$. אז $[T]^{\mathcal{B}}_{\mathcal{C}} \in \mathrm{M}_{m\times n}(F)$. נזכור שההעתקה $\varphi\colon W \to F^m$ הנתונה על ידי $w \mapsto [w]_{\mathcal{C}}$ היא איזומורפיזם. מתקיים

$$\begin{aligned} \mathrm{Col}([T]^{\mathcal{B}}_{\mathcal{C}}) = \mathrm{Sp}([T(v_1)]_{\mathcal{C}}, \ldots, [T(v_n)]_{\mathcal{C}}) = \mathrm{Sp}(\varphi(T(v_1), \ldots, \varphi(T(v_n))) = \\ \varphi(\mathrm{Sp}(T(v_1), \ldots, T(v_n))) = \varphi(T(\mathrm{Sp}(v_1, \ldots, v_n))) = \varphi(T(V)) = \varphi(\mathrm{Im}\, T) \end{aligned}$$

כלומר, φ מעתיקה את $\operatorname{Im} T$ על מרחב העמודות של $[T]_{\mathcal{C}}^{\mathcal{B}}$. הצמצום של φ ל־$\operatorname{Im} T$ הוא, אם כן, איזומורפיזם $\operatorname{Im} T \to \operatorname{Col}([T]_{\mathcal{C}}^{\mathcal{B}})$. לכן לפי מסקנה 6.11, $\operatorname{rk}[T]_{\mathcal{C}}^{\mathcal{B}} = \dim \operatorname{Col}([T]_{\mathcal{C}}^{\mathcal{B}}) = \dim \operatorname{Im} T$. מכאן (א).

(ב) תהי $A = [T]_{\mathcal{C}}^{\mathcal{B}}$. אם T איזומורפיזם, אז $m = \dim W = \dim V = n$, ולכן $A \in \mathrm{M}_n(F)$, ומתקיים $\operatorname{rk} A = \dim \operatorname{Im} T = \dim W = n$, לכן A הפיכה. להיפך, אם A הפיכה, אז $A \in \mathrm{M}_n(F)$, ו־$\dim W = n = \operatorname{rk} A = \dim \operatorname{Im} T$, ולכן T על. כיוון ש־$\dim W = \dim V$, היא גם חח"ע. ∎

הגדרה 6.23: יהיו V, W מרחבים וקטוריים מעל שדה F. תהיינה $T, S\colon V \to W$ העתקות לינאריות ויהי $\alpha \in F$. נגדיר העתקות $T + S,\ \alpha T\colon V \to W$ על ידי $(T + S)(v) = T(v) + S(v)$, לכל $v \in V$. $(\alpha T)(v) = \alpha T(v)$, לכל $v \in V$.

טענה: $T + S,\ \alpha T$ הן העתקות לינאריות.

הוכחה: נוכיח, לגבי שתי ההעתקות, כי הן שומרות כפל בסקלר (באופן דומה מוכיחים גם שהן שומרות חיבור): יהיו $v \in V$ ו־$\beta \in F$. (שים לב שהאות α תפוסה כבר!) אז

$$\begin{aligned}(T+S)(\beta v) &=^1 T(\beta v) + S(\beta v) =^2 \beta T(v) + \beta S(v) =^3\\ &=^3 \beta(T(v) + S(v)) =^4 \beta(T+S)(v)\end{aligned}$$

כאשר השוויונות מוסברים כך:

(1) , (4) לפי הגדרת $T + S$.

(2) S, T שומרות כפל בסקלר.

(3) אחד מחוקי הפילוג במרחבים וקטוריים.

כמו כן

$$\begin{aligned}(\alpha T)(\beta v) &=^1 \alpha T(\beta v) =^2 \alpha(\beta T(v)) =^3 (\alpha\beta)T(v) =^4 (\beta\alpha)T(v) =^5\\ &=^5 \beta(\alpha T(v)) =^6 \beta(\alpha T)(v)\end{aligned}$$

כאשר השוויונות מוסברים כך:

(1) , (6) לפי הגדרת αT.

(2) T שומרת כפל בסקלר.

(3) ,(5) חוק האסוציאטיביות של הכפל בסקלר ב־V.

(4) חוק החילוף של הכפל ב־F. ∎

משפט 6.24: *יהיו V, W שני מרחבים וקטוריים מעל אותו שדה F. יהי $v_1, \ldots, v_n$ בסיס של V, ותהי $w_1, \ldots, w_n$ סדרה של איברי W. אז קיימת העתקה לינארית יחידה $T\colon V \to W$ המקיימת*

$$T(v_1) = w_1, \ldots, T(v_n) = w_n. \tag{1}$$

הוכחה: נניח לרגע ש־T כזו אכן קיימת. יהי $v \in V$ כלשהו. אז, כיוון ש־$v_1, \ldots, v_n$ בסיס של V, קיימים $a_1, \ldots, a_n \in F$ יחידים כך ש־

$$v = a_1 v_1 + \cdots + a_n v_n. \tag{2}$$

כיוון ש־T לינארית, $T(v) = a_1 T(v_1) + \cdots + a_n T(v_n)$. וכיוון שהיא מקיימת (1),

$$T(v) = a_1 w_1 + \cdots + a_n w_n. \tag{3}$$

אגף ימין במשוואה זו אינו תלוי כלל ב־T! (הוא אמנם תלוי בבסיס הנתון וב־v, אך לא ב־T.) מכאן היחידות של T.

(ביתר פירוט, אם גם T' העתקה לינארית שמקיימת (1), אז עבור $v \in V$ שמקיים (2), $T'(v)$ חייב להיות אגף ימין של (3) ולכן $T(v) = T'(v)$. זה נכון לכל $v \in V$, לכן $T = T'$.)

קיום: נגדיר T על ידי (3), לכל $v \in V$, באשר $a_1, \ldots, a_n$ באים מההצגה (2) של v.

אז T מקיימת (1): אכן, יהי $1 \leq j \leq n$, אז

$$v_j = 0v_1 + \cdots + 0v_{j-1} + 1v_j + 0v_{j+1} + \cdots + 0v_n.$$

לכן לפי ההגדרה (3) של T,

$$T(v_j) = 0w_1 + \cdots + 0w_{j-1} + 1w_j + 0w_{j+1} + \cdots + 0w_n = w_j.$$

כמו כן T לינארית. אכן, נראה כאן רק שהיא שומרת כפל בסקלר (שמירת החיבור מראים באופן דומה). יהי $v \in V$ ויהי $\alpha \in F$. יהיו $a_1, \ldots, a_n \in F$ כך שמתקיים (2). אז

$$\alpha v = (\alpha a_1 v_1) + \cdots + (\alpha a_n) v_n.$$

לכן לפי ההגדרה (3) של T,

$$\begin{aligned} T(\alpha v) &= \alpha a_1 T(v_1) + \cdots + \alpha a_n T(v_n) = \alpha a_1 w_1 + \cdots + \alpha a_n w_n = \\ &= \alpha(a_1 w_1 + \cdots + a_n w_n) = \alpha T(v) \end{aligned}$$

∎

הגדרה 6.25: יהיו V, W מרחבים וקטוריים מעל שדה F. נסמן ב־$L(V, W)$ את קבוצה כל ההעתקות הלינאריות מ־V ל־W.

משפט 6.26: *$L(V, W)$ הוא מרחב וקטורי, ביחס לפעולות חיבור והכפל בסקלר שהגדרנו לעיל.*

הוכחה: צריך להוכיח את כל האקסיומות של מרחב וקטורי. נראה כאן רק אקסיומה אחת. למשל אסוציאטיביות של החיבור: צריך להוכיח כי לכל $S, T, R \in L(V, W)$ מתקיים

$$(S + T) + R = S + (T + R).$$

ואכן, יהי $v \in V$. אז

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

ושני הביטויים שווים בגלל האסוציאטיביות של החיבור ב־W. ∎

משפט 6.27: *יהיו V, W שני מרחבים וקטוריים מעל שדה F. נניח $\dim V = n,\ \dim W = m$, ויהיו $\mathcal{B}$ בסיס של V ו־$\mathcal{C}$ בסיס של W. ההעתקה $\psi\colon L(V,W) \to \mathrm{M}_{m\times n}(F)$ הנתונה על ידי $T \mapsto [T]_{\mathcal{C}}^{\mathcal{B}}$ הינה איזומורפיזם.*

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

(א) ψ לינארית: תהיינה $T, T' \in L(V,W)$ ויהי $\alpha \in F$. אז

$$\begin{aligned}[T+T']_{\mathcal{C}}^{\mathcal{B}} &= \big([(T+T')(v_1)]_{\mathcal{C}}, \ldots, [(T+T')(v_n)]_{\mathcal{C}}\big) = \\ &= \big([T(v_1)+T'(v_1)]_{\mathcal{C}}, \ldots, [T(v_n)+T'(v_n)]_{\mathcal{C}}\big) = \\ &= \big([T(v_1)]_{\mathcal{C}} + [T'(v_1)]_{\mathcal{C}}, \ldots, [T(v_n)]_{\mathcal{C}} + [T'(v_n)]_{\mathcal{C}}\big) = \\ &= \big([T(v_1)]_{\mathcal{C}}, \ldots, [T(v_n)]_{\mathcal{C}}\big) + \big([T'(v_1)]_{\mathcal{C}}, \ldots, T'(v_n)]_{\mathcal{C}}\big) = [T]_{\mathcal{C}}^{\mathcal{B}} + [T']_{\mathcal{C}}^{\mathcal{B}}\end{aligned}$$

לכן ψ שומרת חיבור. השתמשנו במשפט 6.9, לפיו ההעתקה $w \mapsto [w]_{\mathcal{C}}$ שומרת חיבור.

באופן דומה מוכיחים ש־ψ שומרת כפל בסקלר.

(ב) ψ על: תהי $A \in \mathrm{M}_{m\times n}(F)$. לפי משפט 6.9, ההעתקה $w \mapsto [w]_{\mathcal{C}}$ היא על F^m. לכן יש $w_1, \ldots, w_n \in W$ כך ש־$[w_1]_{\mathcal{C}}, \ldots, [w_n]_{\mathcal{C}}$ הן העמודות של A. לפי משפט 6.24 יש $T \in L(V,W)$ כך ש־$T(v_1) = w_1, \ldots, T(v_n) = w_n$. מכאן

$$\psi(T) = [T]_{\mathcal{C}}^{\mathcal{B}} = \big([T(v_1)]_{\mathcal{C}}, \ldots, [T(v_n)]_{\mathcal{C}}\big) = \big([w_1]_{\mathcal{C}}, \ldots, [w_n]_{\mathcal{C}}\big) = A.$$

(ג) ψ חח"ע: נניח $[T]_{\mathcal{C}}^{\mathcal{B}} = [T']_{\mathcal{C}}^{\mathcal{B}}$. צריך להוכיח שלכל $v \in V$ מתקיים $T(v) = T'(v)$.

ואכן, $[T(v)]_{\mathcal{C}} = [T]_{\mathcal{C}}^{\mathcal{B}}[v]_{\mathcal{B}} = [T']_{\mathcal{C}}^{\mathcal{B}}[v]_{\mathcal{B}} = [T'(v)]_{\mathcal{C}}$ לפי משפט 6.9 ההעתקה $w \mapsto [w]_{\mathcal{C}}$ היא חח"ע, לכן $T(v) = T'(v)$. ∎

מסקנה 6.28: $\dim L(V,W) = (\dim V) \times (\dim W)$.

הוכחה: נניח $\dim V = n,\ \dim W = m$. לפי מסקנה 6.11, $\dim L(V,W) = \dim \mathrm{M}_{m\times n}(F) = m \times n$.

∎

למה 6.29: *תהיינה $U \xrightarrow{S} V \xrightarrow{T} W$ העתקות לינאריות של מרחבים וקטוריים מעל שדה F. אז $T \circ S\colon U \to W$ הינה העתקה לינארית.*

הוכחה: נראה רק ש־$T \circ S$ שומרת כפל בסקלר (שמירת החיבור ־ באופן דומה).

יהי $u \in U$ ויהי $\alpha \in F$. אז

$$(T \circ S)(\alpha u) =^1 T(S(\alpha u)) =^2 T(\alpha S(u)) =^3 \alpha T(S(u)) =^1 \alpha (T \circ S)(u),$$

כאשר (1) נובע מההגדרה של $T \circ S$; (2) נובע מהלינאריות של S; (3) נובע מהלינאריות של T. ∎

משפט 6.30: *תהיינה* $U \xrightarrow{S} V \xrightarrow{T} W$ *העתקות לינאריות של מרחבים וקטוריים מעל שדה* F. *יהיו* $\mathcal{A}$ *בסיס של* U; $\mathcal{B}$ *בסיס של* V; $\mathcal{C}$ *בסיס של* W. *אז*

$$[T \circ S]^{\mathcal{A}}_{\mathcal{C}} = [T]^{\mathcal{B}}_{\mathcal{C}} [S]^{\mathcal{A}}_{\mathcal{B}}.$$

הוכחה: נניח כי $\mathcal{A} = (u_1, \dots, u_n)$. נשים לב ש־$[u_j]_{\mathcal{A}} = e_j$, לכל $1 \le j \le n$, כי

$$u_j = 0u_1 + \dots + 0u_{j-1} + 1u_j + 0u_{j+1} + \dots + 0u_n.$$

כעת

$$\begin{aligned}
[T \circ S]^{\mathcal{A}}_{\mathcal{C}} &=^1 \left([T(S(u_1))]_{\mathcal{C}}, \dots, [T(S(u_n))]_{\mathcal{C}}\right) =^2 \left([T]^{\mathcal{B}}_{\mathcal{C}}[S(u_1)]_{\mathcal{B}}, \dots, [T]^{\mathcal{B}}_{\mathcal{C}}[S(u_n)]_{\mathcal{B}}\right) \\
&=^3 \left([T]^{\mathcal{B}}_{\mathcal{C}}[S]^{\mathcal{A}}_{\mathcal{B}}[u_1]_{\mathcal{A}}, \dots, [T]^{\mathcal{B}}_{\mathcal{C}}[S]^{\mathcal{A}}_{\mathcal{B}}[u_n]_{\mathcal{A}}\right) =^4 \left([T]^{\mathcal{B}}_{\mathcal{C}}[S]^{\mathcal{A}}_{\mathcal{B}}e_1, \dots, [T]^{\mathcal{B}}_{\mathcal{C}}[S]^{\mathcal{A}}_{\mathcal{B}}e_n\right) \\
&=^5 [T]^{\mathcal{B}}_{\mathcal{C}}[S]^{\mathcal{A}}_{\mathcal{B}}
\end{aligned}$$

כאשר

(1) נובע מההגדרה של $T \circ S$ ומהגדרת המטריצה של העתקה לינארית לפי בסיסים $\mathcal{A}, \mathcal{C}$;

(2) נובע מהכלל $[T(v)]_{\mathcal{C}} = [T]^{\mathcal{B}}_{\mathcal{C}}[v]_{\mathcal{B}}$ לכל $v \in V$;

(3) נובע מהכלל $[S(u)]_{\mathcal{B}} = [S]^{\mathcal{A}}_{\mathcal{B}}[u]_{\mathcal{A}}$ לכל $u \in U$;

(4) נובע מההערה לפני המשוואה;

(5) נובע מכך שלכל מטריצה M, הוקטור Me_j הוא העמודה ה־j של M (תרגיל 3.8). ∎

הגדרה 6.31: יהי V מרחב וקטורי מעל F. יהי $\mathcal{B} = (v_1, \dots, v_n)$ בסיס שלו ותהי $\mathcal{B}' = (v'_1, \dots, v'_n)$ סדרה של איברי V. **מטריצת המעבר** מ־$\mathcal{B}$ ל־$\mathcal{B}'$ היא המטריצה $P = (p_{ij}) \in \mathrm{M}_n(F)$ עבורה מתקיים

$$1 \le j \le n \qquad v'_j = p_{1j}v_1 + p_{2j}v_2 + \dots + p_{nj}v_n. \qquad (*)$$

במלים אחרות

$$P = \left([v'_1]_{\mathcal{B}}, \dots, [v'_n]_{\mathcal{B}}\right) \in \mathrm{M}_n(F).$$

דוגמה 6.32: מטריצת המעבר מ־$\left(\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\right)$ ל־$\left(\begin{pmatrix}1\\-1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\right)$ היא $\begin{pmatrix}1 & 1\\0 & -2\end{pmatrix}$.

תרגיל 6.33: *יהי V מרחב וקטורי מעל F. יהי $\mathcal{B} = (v_1, \ldots, v_n)$ בסיס שלו ותהי $P = (p_{ij}) \in \mathrm{M}_n(F)$ המטריצה. אז קיימת סדרה $\mathcal{B}' = (v'_1, \ldots, v'_n)$ של איברי V כך ש־P היא מטריצת המעבר מ־$\mathcal{B}$ ל־$\mathcal{B}'$.*

פתרון: נגדיר את $v'_1, \ldots, v'_n$ על ידי (*). ∎

משפט 6.34: *בסימונים של ההגדרה, P הפיכה אם ורק אם $\mathcal{B}'$ בסיס.*

הוכחה: P הפיכה $\Leftrightarrow$ $\mathrm{rk}\, P = n$ $= \dim(\mathrm{Col}(P))$ $= \dim(Sp([v'_1]_{\mathcal{B}}, \ldots, [v'_n]_{\mathcal{B}}))$ $\Leftrightarrow$ $[v'_1]_{\mathcal{B}}, \ldots, [v'_n]_{\mathcal{B}}$ $\Leftrightarrow$ בסיס של F^n.

אך ההעתקה $V \to F^n$ הנתונה על ידי $v \mapsto [v]_{\mathcal{B}}$ היא האיזומורפיזם $V \to F^n$, לכן לפי מסקנה 6.11, $[v'_1]_{\mathcal{B}}, \ldots, [v'_n]_{\mathcal{B}}$ בסיס של F^n $\Leftrightarrow$ $v'_1, \ldots, v'_n$ בסיס של V. ∎

מעתה נניח כי $\mathcal{B}, \mathcal{B}'$ שניהם בסיסים ו־P מטריצת המעבר מ־$\mathcal{B}$ ל־$\mathcal{B}'$.

הערה 6.35: מטריצת המעבר מ־$\mathcal{B}$ ל־$\mathcal{B}'$ היא $[1_V]^{\mathcal{B}'}_{\mathcal{B}}$, ומטריצת המעבר מ־$\mathcal{B}'$ ל־$\mathcal{B}$ היא ההופכי שלה $[1_V]^{\mathcal{B}}_{\mathcal{B}'}$.

אכן, הטענה הראשונה נובעת מההגדרות, והשנייה מ־ $[1_V]^{\mathcal{B}'}_{\mathcal{B}}[1_V]^{\mathcal{B}}_{\mathcal{B}'} = [1_V]^{\mathcal{B}}_{\mathcal{B}} = I_n$. ∎

מסקנה 6.36: *יהי $v \in V$. אז*

(א) $[v]_{\mathcal{B}} = P[v]_{\mathcal{B}'}$

(ב) $[v]_{\mathcal{B}'} = P^{-1}[v]_{\mathcal{B}}$

הוכחה: : (א) לפי משפט 6.20, $[v]_{\mathcal{B}} = [1_V(v)]_{\mathcal{B}} = [1_V]^{\mathcal{B}'}_{\mathcal{B}}[v]_{\mathcal{B}'} = P[v]_{\mathcal{B}'}$. (ב) – דומה. ∎

דוגמה 6.37: יהי $V = \mathbb{R}^2$, ותהי P מטריצת המעבר מהבסיס הסטנדרטי $\mathcal{B} = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$ לבסיס $\mathcal{B}' = \left(v'_1 = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, v'_2 = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} \right)$ שמתקבל מ־$\mathcal{B}$ על ידי הסיבוב בזוית θ. אז מטריצת המעבר מ־$\mathcal{B}$ ל־$\mathcal{B}'$ היא $P = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. יהי v וקטור בעל קואורדינטות x', y' לפי $\mathcal{B}'$. אז

$$v = [v]_{\mathcal{B}} = P \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x'\cos\theta - y'\sin\theta \\ x'\sin\theta + y'\cos\theta \end{pmatrix}.$$

משפט 6.38 (נוסחת המעבר): *תהי $T\colon V \to W$ העתקה לינארית. יהיו:*

$\mathcal{B}, \mathcal{B}'$ בסיסים סדורים של V, ותהי P מטריצת המעבר מ־$\mathcal{B}$ ל־$\mathcal{B}'$;

$\mathcal{C}, \mathcal{C}'$ בסיסים סדורים של W. ותהי Q מטריצת המעבר מ־$\mathcal{C}$ ל־$\mathcal{C}'$. אז

$$[T]^{\mathcal{B}'}_{\mathcal{C}'} = Q^{-1}[T]^{\mathcal{B}}_{\mathcal{C}}P$$

הוכחה: $T = 1_W \circ T \circ 1_V$, לכן לפי משפט 6.30, $[T]^{\mathcal{B}'}_{\mathcal{C}'} = [1_W]^{\mathcal{C}}_{\mathcal{C}'}[T]^{\mathcal{B}}_{\mathcal{C}}[1_V]^{\mathcal{B}'}_{\mathcal{B}}$. ∎

דוגמה 6.39: יהי $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ השיקוף ביחס לציר דרך הראשית שעובר דרך v_1' מהדוגמה הקודמת. אז

$$T(v_1') = v_1', \quad T(v_2') = -v_2'$$

ולכן $[T]_{\mathcal{B}'}^{\mathcal{B}'} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

יהיו $\mathcal{C} = \mathcal{B}, \mathcal{C}' = \mathcal{B}'$.

אז $P = Q = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. מכאן $P^{-1} = Q^{-1} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$. לכן

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = [T]_{\mathcal{B}'}^{\mathcal{B}'} = P^{-1}[T]_{\mathcal{B}}^{\mathcal{B}}P$$

ולכן

$$\begin{aligned}[T]_{\mathcal{B}}^{\mathcal{B}} =& P\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}P^{-1} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \\ =& \begin{pmatrix} \cos^2\theta - \sin^2\theta & 2\sin\theta\cos\theta \\ 2\sin\theta\cos\theta & \sin^2\theta - \cos^2\theta \end{pmatrix} = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}\end{aligned}$$

מכאן נובע

$$\begin{aligned} T\begin{pmatrix} x \\ y \end{pmatrix} &= [T\begin{pmatrix} x \\ y \end{pmatrix}]_{\mathcal{B}} = [T]_{\mathcal{B}}^{\mathcal{B}}[\begin{pmatrix} x \\ y \end{pmatrix}]_{\mathcal{B}} \\ &= \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos 2\theta + y\sin 2\theta \\ x\sin 2\theta - y\cos 2\theta \end{pmatrix}\end{aligned}$$

מסקנה 6.40: *(בתנאי המשפט)* $\mathrm{rk}[T]_{\mathcal{C}'}^{\mathcal{B}'} = \mathrm{rk}[T]_{\mathcal{C}}^{\mathcal{B}}$.

הוכחה: הכפלה במטריצות הפיכות (מימין או משמאל) אינה משנה את דרגת המטריצה. ∎

הגדרה 6.41: מטריצות $A, A' \in \mathrm{M}_{m\times n}(F)$ נקראות **שקולות** (**־שורות ועמודות**) אם יש $P \in \mathrm{M}_n(F)$, $Q \in \mathrm{M}_m(F)$ הפיכות כך ש־$A' = Q^{-1}AP$.

המשפט הקודם אומר: אם A, A' מטריצות של אותה העתקה לינארית T, כל אחת לפי זוג בסיסים משלה, אז A, A' שקולות. נראה שגם ההיפך נכון:

משפט 6.42: *תהי $T\colon V \to W$ העתקה לינארית. יהיו:*
$\mathcal{B}, \mathcal{C}$ בסיסים סדורים של V, W, בהתאמה, ותהי $A = [T]_{\mathcal{C}}^{\mathcal{B}}$. תהי $A' \in \mathrm{M}_{m\times n}(F)$. אם A, A' שקולות, אז יש בסיסים סדורים $\mathcal{B}', \mathcal{C}'$ בסיסים של V, W, בהתאמה, כך ש־ $A' = [T]_{\mathcal{C}'}^{\mathcal{B}'}$.

הוכחה: נתון ש־$A' = Q^{-1}AP$, באשר $P \in \mathrm{M}_n(F)$, $Q \in \mathrm{M}_m(F)$ הפיכות. לפי המשפט הקודם די למצוא בסיסים $\mathcal{B}', \mathcal{C}'$ של V, W, בהתאמה, כך ש־P היא מטריצת המעבר מ־$\mathcal{B}$ ל־$\mathcal{B}'$ ו־Q היא מטריצת המעבר מ־$\mathcal{C}$ ל־$\mathcal{C}'$.
אבל זה נכון לפי תרגיל 6.33. ∎

משפט 6.43: *מטריצות $A, A' \in \mathrm{M}_{m\times n}(F)$ שקולות אם ורק אם $\mathrm{rk}\, A = \mathrm{rk}\, A'$. בפרט, אם $\mathrm{rk}\, A = r$, אז A שקולה ל-* $\begin{pmatrix} I_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix} = (e_1, \ldots, e_r, 0 \ldots, 0)$.

הוכחה: נניח $A' = Q^{-1}AP$, באשר $P \in \mathrm{M}_n(F)$, $Q \in \mathrm{M}_m(F)$ הפיכות. הכפלה במטריצה הפיכה אינה משנה את הדרגה, לכן $\mathrm{rk}\, A = \mathrm{rk}\, A'$.

להיפך, נניח כי $r = \mathrm{rk}\, A = \mathrm{rk}\, A'$. נראה תחילה ש-$A$ שקולה ל-$\begin{pmatrix} I_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}$.

נתבונן ב-$A^t \in \mathrm{M}_{n\times m}(F)$. קיימת $P_1 \in \mathrm{M}_n(F)$ הפיכה כך ש-P_1A^t מדורגת קנונית. אז $\mathrm{rk}\, P_1A^t = \mathrm{rk}\, A^t = \mathrm{rk}\, A$, ולכן $n - r$ השורות האחרונות של P_1A^t הן 0. לכן $n - r$ העמודות האחרונות של $AP_1^t = (P_1A^t)^t$ הן 0.

כעת קיימת $Q_1 \in \mathrm{M}_m(F)$ הפיכה כך ש-$Q_1AP_1^t$ מדורגת קנונית. גם היא מדרגה r וגם $n - r$ העמודות האחרונות שלה הן 0. לכן r העמודות הראשונות שלה הן המבילות. מכאן $Q_1AP_1^t = (e_1, \ldots, e_r, 0 \ldots, 0)$.
נסמן $P = P_1^t$ ו-$Q = Q_1^{-1}$. אז Q, P הפיכות ו-$\begin{pmatrix} I_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix} = (e_1, \ldots, e_r, 0 \ldots, 0) = Q^{-1}AP$.

באותו אופן קיימות Q', P' הפיכות כך ש-$\begin{pmatrix} I_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix} = Q'^{-1}AP'$. לכן $Q'^{-1}AP' = Q^{-1}AP$.
נכפיל שוויון זה מימין ב-P'^{-1} ומשמאל ב-Q'. נקבל $A' = Q'Q^{-1}APP'^{-1} = (QQ'^{-1})A(PP'^{-1})$. כיוון ש-$QQ'^{-1}, PP'^{-1}$ הפיכות, A, A' שקולות. ∎

הגדרה 6.44: מטריצות $A, A' \in \mathrm{M}_n(F)$ נקראות **דומות** אם יש $P \in \mathrm{M}_n(F)$ הפיכה כך ש-$A' = P^{-1}AP$.

בדומה למשפטים לעיל אפשר להוכיח:

משפט 6.45: *יהי V מרחב וקטורי עם בסיס $\mathcal{B}$, תהי $T\colon V \to V$ העתקה לינארית ותהי $A = [T]_{\mathcal{B}}^{\mathcal{B}}$.*

(א) *יהי $\mathcal{B}'$ בסיס של V. אז $[T]_{\mathcal{B}'}^{\mathcal{B}'}$ דומה ל-A.*

(ב) *תהי $A' \in \mathrm{M}_{m\times n}(F)$ דומה ל-A. אז יש בסיס $\mathcal{B}'$ של V כך ש-$A' = [T]_{\mathcal{B}'}^{\mathcal{B}'}$.*

את דמיון המטריצות נחקור בפרק הבא.

7. וקטורים עצמיים וערכים עצמיים

הגדרה 7.1: יהי V מרחב וקטורי מעל שדה F. תהי $T\colon V \to V$ העתקה לינארית, ויהי $\lambda \in F$ סקלר. הקבוצה

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

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

דוגמה 7.2: 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 7.3: יהי F שדה ותהי $A \in \mathrm{M}_n(F)$, ויהי $\lambda \in F$ סקלר. אז

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

המרחב עצמי של A השייך ל־λ. וקטור $v \in F^n$ הוא **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־$Av = \lambda v$.

במלים אחרות, V_λ הוא המרחב עצמי ו־v וקטור עצמי של T_A השייכים ל־λ, ו־λ הוא ערך עצמי של T_A.

למה 7.4: *יהי V מרחב וקטורי מעל F, ותהי $T\colon V \to V$ העתקה לינארית. יהי $\lambda \in F$. 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 $[T(v)]_\mathcal{B} = A[v]_\mathcal{B}$. כמו כן $[\lambda v]_\mathcal{B} = \lambda[v]_\mathcal{B}$ (כי $v \mapsto [v]_\mathcal{B}$ שומרת כפל בסקלר) ולכן

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

ו־$[v]_\mathcal{B} \neq 0 \;\Leftrightarrow\; v \neq 0$ (כי $v \mapsto [v]_\mathcal{B}$ חד חד ערכית). מכאן בקלות המסקנה. ∎

משפט 7.5:

(א) *יהי V מרחב וקטורי מעל שדה F, $T\colon V \to V$ העתקה לינארית, $\lambda \in F$. אז $V_\lambda = \mathrm{Ker}(\lambda 1_V - T)$. בפרט V_λ תת מרחב של V. לכן λ הוא ערך עצמי של T אם ורק אם ההעתקה $\lambda 1_V - T$ אינה חד־חד־ערכית.*

(ב) *יהי F שדה, $A \in \mathrm{M}_n(F)$, $\lambda \in F$. אז $V_\lambda = \mathrm{Ker}(\lambda I - A)$. לכן λ ערך עצמי של A אם ורק אם $\det(\lambda I - A) = 0$.*

הוכחה: (א) יהי $v \in V$. אז

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

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

(ב) באופן דומה λ הוא ערך עצמי של A אם ורק אם $\mathrm{Ker}(\lambda I_n - A) \neq \{0\}$, כלומר, $\mathrm{rk}(\lambda I_n - A) < n$, כלומר $|\lambda I_n - A| = 0$. ∎

מטריצה $A \in \mathrm{M}_n(F)$ נקראת **אלכסונית** אם $(A)_{ij} = 0$ לכל $i \neq j$, כלומר $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

תרגיל 7.6: חשיבות המטריצות האלכסוניות.

$$\mathrm{diag}(\alpha_1, \ldots, \alpha_n)\mathrm{diag}(\beta_1, \ldots, \beta_n) = \mathrm{diag}(\alpha_1\beta_1, \ldots, \alpha_n\beta_n)$$

$$\mathrm{diag}(\alpha_1^{-1}, \ldots, \alpha_n^{-1}) = \mathrm{diag}(\alpha_1, \ldots, \alpha_n)^{-1}, \quad \text{אם } \alpha_1, \ldots, \alpha_n \in F^\times.$$

למה 7.7: *יהי V מרחב וקטורי מעל שדה F ויהי $\mathcal{B} = (v_1, \ldots, v_n)$ בסיס סדור שלו. תהי $T\colon V \to V$ העתקה לינארית. יהי $1 \leq j \leq n$ ויהי $\lambda_j \in F$. אז העמודה ה־ j של המטריצה $[T]_\mathcal{B}^\mathcal{B}$ היא $\lambda_j e_j$ אם ורק אם v_j וקטור עצמי של T השייך ל־ λ_j. אם זה קורה אז λ_j הוא ערך עצמי של T.*

הוכחה: לפי ההגדרה של $[T]_\mathcal{B}^\mathcal{B}$, העמודה ה־$j$ שלה היא $[T(v_j)]_\mathcal{B}$. לפי ההגדרה של $[-]_\mathcal{B}$, $[T(v_j)]_\mathcal{B} = \lambda_j e_j$ אם ורק אם $T(v_j) = \lambda_j v_j$. אם זה קורה, אז λ_j ערך עצמי של A, כי v_j, כאיבר של בסיס, שונה מ־0. ∎

משפט 7.8: *להעתקה לינארית $T\colon V \to V$ יש הצגה אלכסונית אם ורק אם ל־ V יש בסיס המורכב מוקטורים עצמיים של T.*

הוכחה: יהי $\mathcal{B} = (v_1, \ldots, v_n)$ בסיס של V. ההצגה של T לפי $\mathcal{B}$ היא אלכסונית אם ורק אם יש סקלרים $\lambda_1, \ldots, \lambda_n \in F$ כך ש־ $[T]_\mathcal{B}^\mathcal{B} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) = (\lambda_1 e_1, \ldots, \lambda_n e_n)$. לפי למה 7.7 זה קורה אם ורק אם v_j וקטור עצמי של T השייך ל־λ_j, לכל $1 \leq j \leq n$. ∎

משפט 7.9: *יהי F שדה, תהי $A \in \mathrm{M}_n(F)$, ותהי $P = (v_1, \ldots, v_n) \in \mathrm{M}_n(F)$ הפיכה.*

(א) *יהי $1 \leq j \leq n$ ויהי $\lambda_j \in F$. אז העמודה ה־ j של $P^{-1}AP$ היא $\lambda_j e_j$ אם ורק אם v_j וקטור עצמי של A השייך ל־ λ_j. אם זה קורה אז λ_j הוא ערך עצמי של A.*

(ב) *A דומה למטריצה אלכסונית אם ורק אם ל־ F^n בסיס המורכב מוקטורים עצמיים של A.*

הוכחה: (א) העמודה ה־j של $P^{-1}AP$ היא $P^{-1}Av_j$. היא שווה ל־$\lambda_j e_j$ אם ורק אם $Av_j = P\lambda_j e_j$, כלומר, $Av_j = \lambda_j v_j$. כיוון ש־P הפיכה, $v_j \neq 0$. לכן, אם $Av_j = \lambda_j v_j$, אז λ_j ערך עצמי של A.

(ב) נניח ש־A דומה למטריצה אלכסונית, כלומר, יש $P \in \mathrm{M}_n(F)$ הפיכה כך ש־$P^{-1}AP$ אלכסונית. כיוון ש־P הפיכה, עמודותיה בלתי תלויות לינארית, ולכן מהוות בסיס של F^n. לפי (א), איברי בסיס זה הם וקטורים עצמיים של A. להיפך, נניח שיש בסיס של F^n שמורכב מוקטורים עצמיים. תהי P המטריצה שאיברי בסיס זה הם עמודותיה. כיוון שהעמודות בלתי תלויות לינארית, P הפיכה. לפי (א), $P^{-1}AP$ אלכסונית. ∎

תרגיל 7.10: תהי $A = \begin{pmatrix} 33 & 24 \\ -40 & -29 \end{pmatrix}$. חשב את A^{100}.

פתרון: מספר λ הוא ערך עצמי של A אם ורק אם $0 = \begin{vmatrix} \lambda - 33 & -24 \\ 40 & \lambda + 29 \end{vmatrix}$, כלומר $\lambda^2 - 4\lambda + 3 = 0$. למשוואה זו שרשים $1, 3$, לכן הערכים העצמיים של A הם $1, 3$. נמצא לכל אחד מהם וקטורים עצמיים:

$$V_3 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \;\middle|\; (3I - A)\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \;\middle|\; \begin{pmatrix} -30 & -24 \\ 40 & 32 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} =$$

$$= \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \;\middle|\; \begin{pmatrix} 5 & 4 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} = \mathrm{Sp}(\begin{pmatrix} 4 \\ -5 \end{pmatrix})$$

$$V_1 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \;\middle|\; (1I - A)\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \;\middle|\; \begin{pmatrix} -32 & -24 \\ 40 & 30 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} =$$

$$= \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \;\middle|\; \begin{pmatrix} 4 & 3 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} = \mathrm{Sp}(\begin{pmatrix} -3 \\ 4 \end{pmatrix})$$

קל לראות ש־$\begin{pmatrix} 4 \\ -5 \end{pmatrix}, \begin{pmatrix} -3 \\ 4 \end{pmatrix}$ בלתי תלויים לינארית. (ראה גם את המשפט הבא.) לפי המסקנה לעיל, אם נגדיר $P = \begin{pmatrix} 4 & -3 \\ -5 & 4 \end{pmatrix}$ אז $P^{-1}AP = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$. מכאן

$$P^{-1}A^{100}P = (P^{-1}AP)^{100} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}^{100} = \begin{pmatrix} 3^{100} & 0 \\ 0 & 1 \end{pmatrix},$$

ולכן

$$A^{100} = P \begin{pmatrix} 3^{100} & 0 \\ 0 & 1 \end{pmatrix} P^{-1} = \begin{pmatrix} 4 & -3 \\ -5 & 4 \end{pmatrix} \begin{pmatrix} 3^{100} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 4 & 3 \\ 5 & 4 \end{pmatrix}.$$

$$= \begin{pmatrix} 16 \cdot 3^{100} - 15 & 12 \cdot 3^{100} - 12 \\ -20 \cdot 3^{100} + 20 & -15 \cdot 3^{100} + 16 \end{pmatrix}$$ ∎

כדי למצוא בסיס מורכב מוקטורים עצמיים, אם הוא קיים, ניעזר במשפט הבא:

משפט 7.11: *יהי V מרחב וקטורי, $T\colon V \to V$ העתקה לינארית, $\lambda_1, \ldots, \lambda_m$ ערכים עצמיים של T, שונים זה מזה. לכל $1 \le i \le m$ תהי $v_{i1}, v_{i2}, \ldots, v_{ik_i}$ סדרה של וקטורים בלתי תלויים לינארית ב־V_{λ_i}. אזי הסדרה*

$$v_{11}, v_{12}, \ldots, v_{1k_1}, \quad v_{21}, v_{22}, \ldots, v_{2k_2}, \quad \ldots \quad , v_{m1}, v_{m2}, \ldots, v_{mk_m}$$

בלתי תלויה לינארית.

הוכחה: באינדוקציה על m. עבור $m = 1$ זה נובע מההנחה: $v_{11}, v_{12}, \ldots, v_{1k_1}$ בלתי תלויים לינארית. נניח נכונות עבור $m - 1$ סדרות. יהיו

$$\alpha_{11}, \alpha_{12}, \ldots, \alpha_{1k_1}, \quad \alpha_{21}, \alpha_{22}, \ldots, \alpha_{2k_2}, \quad \ldots \quad , \alpha_{m1}, \alpha_{m2}, \ldots, \alpha_{mk_m} \in F$$

כך ש־

$$\left(\alpha_{11}v_{11}+\cdots+\alpha_{1k_1}v_{1k_1}\right) \quad +\cdots+ \quad \left(\alpha_{m1}v_{m1}+\cdots+\alpha_{mk_m}v_{mk_m}\right)=0 \tag{1}$$

צריך להוכיח $\alpha_{ij}=0$ לכל i,j. נפעיל T על שני האגפים ונקבל, בגלל ש־$T(v_{ij})=\lambda_i v_{ij}$,

$$\left(\alpha_{11}\lambda_1 v_{11}+\cdots+\alpha_{1k_1}\lambda_1 v_{1k_1}\right) \quad +\cdots+ \quad \left(\alpha_{m1}\lambda_m v_{m1}+\cdots+\alpha_{mk_m}\lambda_m v_{mk_m}\right)=0 \tag{2}$$

נכפיל את (1) ב־λ_m ונחסיר אותה מ־(2):

$$\begin{aligned}\left(\alpha_{11}(\lambda_1-\lambda_m)v_{11}+\cdots+\alpha_{1k_1}(\lambda_1-\lambda_m)v_{1k_1}\right) \quad &+ \quad \cdots \\ +\left(\alpha_{m1}(\lambda_{m-1}-\lambda_m)v_{m-1\,1}+\cdots+\alpha_{m-1\,k_{m-1}}(\lambda_{m-1}-\lambda_m)v_{m-1\,k_{m-1}}\right)&=0\end{aligned} \tag{3}$$

לפי הנחתה האידנוקציה (על $m-1$ סדרות) כל המקדמים במשוואה זו הם 0. כיוון ש־$\lambda_i-\lambda_m\neq 0$ לכל $1\leq i\leq m-1$, קיבלנו

$$\alpha_{11}=\cdots=\alpha_{1k_1} \quad =\cdots= \quad \alpha_{m-1\,1}=\cdots=\alpha_{m-1\,k_{m-1}}=0$$

אם נציב זאת ב־(1), נקבל

$$\alpha_{m1}v_{m1}+\cdots+\alpha_{mk_m}v_{mk_m}=0$$

וכיוון ש־$v_{m1},\ldots,v_{mk_m}$ בלתי תלויים לינארית, גם $\alpha_{m1}=\cdots=\alpha_{mk_m}=0$. ∎

8. פולינום אופייני, פולינום מזערי וצורת ז'ורדן

הגדרה 8.1: יהי F שדה ותהי $A \in \mathrm{M}_n(F)$, נאמר

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

הפולינום

$$f_A(X) := |(XI_n - A)| = \begin{vmatrix} X - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & X - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & X - a_{nn} \end{vmatrix} \in F[X]$$

ייקרא **הפולינום האופייני** של A.

דוגמה 8.2: תהי $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. אז

$$f_A(X) = \begin{vmatrix} X-1 & -2 \\ -3 & X-4 \end{vmatrix} = (X-1)(X-4) - (-2)\cdot(-3) = X^2 - 5X - 2.$$

משפט 8.3: *יהי F שדה ותהי $A \in \mathrm{M}_n(F)$, ויהי $\lambda \in F$. אז λ ערך עצמי של A אם ורק אם $f_A(\lambda) = 0$.*

הוכחה: $f_A(\lambda) = |(\lambda I_n - A)|$, לכן לפי משפט 7.5, λ ערך עצמי של A אם ורק אם $f_A(\lambda) = 0$. ∎

למה 8.4: *$f_A(X) = c_0X^n + c_1X^{n-1} + \cdots + c_n$, ממעלה n, באשר*

(1) $c_0 = 1$;

(2) $c_1 = -\operatorname{tr} A = -(a_{11} + a_{22} + \cdots + a_{nn})$;

(3) $c_n = (-1)^n |A|$.

הוכחה: לא נוכיח ש־$f_A(X)$ ממעלה n וגם לא את (1), (2). כדי להוכיח (3), הצב 0 ב־f_A. ∎

מסקנה 8.5: *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: $\lambda_1, \ldots, \lambda_n$ ערכים עצמיים של A, שונים זה מזה. יהיו $0 \neq v_i \in V_{\lambda_i}$, עבור $1 \leq i \leq n$. לפי משפט 7.11, $v_1, \ldots, v_n$ בלתי תלוים לינארית, ולכן בסיס של F^n. ומכאן המסקנה לפי מסקנה 7.9. ∎

משפט 8.6: *(ללא הוכחה)*

(א) *(המשפט היסודי של אלגברה) יהי* $f(X) = X^n + c_1X^{n-1} + \cdots + c_n \in \mathbb{C}[X]$. *אז קיימים* $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$, *לא בהכרח שונים זה מזה, כך ש־* $f(X) = (X - \lambda_1) \cdots (X - \lambda_n)$.

(ב) $\lambda_1, \ldots, \lambda_n$ *כאלה יחידים, עד כדי הסדר.*

מעתה נניח כי השדה F מוכל בשדה המרוכבים $\mathbb{C}$.

מסקנה 8.7: *למטריצה* $A \in \mathrm{M}_n(F)$ *יש לכל היותר* n *ערכים עצמיים שונים ב־* F.

הוכחה: לפי המשפט היסודי של אלגברה, $f_A(X) = (X - \lambda_1) \cdots (X - \lambda_n)$, באשר $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$. ברור שאם $\lambda \in F \subseteq \mathbb{C}$ שורש של f_A אז $\lambda = \lambda_i$ עבור איזה i. לכן ל־f_A לכל היותר n שורשים ב־F. היות ולפי משפט 8.2 ערכים עצמיים של A הם שרשי f_A, מכאן המסקנה. ∎

משפט 8.8: *למטריצות דומות אותו פולינום אופייני.*

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

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

מכאן

$$\begin{aligned} f_{P^{-1}AP}(X) &= \left|(XI_n - P^{-1}AP)\right| = \left|P^{-1}\right| \left|(XI_n - A)\right| \left|P\right| \\ &= \left|P^{-1}\right| \left|P\right| \left|(XI_n - A)\right| = \left|I_n\right| f_A(X) = f_A(X) \end{aligned}$$ ∎

אם F שדה, $g = \sum_{i=0}^{k} a_iX^i = a_0 + a_1X + a_2X^2 + \cdots + a_kX^k$, ו־$A \in \mathrm{M}_n(F)$, נגדיר

$$g(A) := \sum_{i=0}^{k} a_iA^i = a_0I_n + a_1A + a_2A^2 + \cdots + a_kA^k \in \mathrm{M}_n(F).$$

תרגיל 8.9: *יהי* F *שדה ותהי* $A \in \mathrm{M}_n(F)$. *אז יש* $0 \neq g \in F[X]$ *(ממעלה* $\leq n^2$*) כך ש־* $g(A) = 0$.

הוכחה: צ"ל: יש $a_0, a_1, \ldots, a_{n^2} \in F$, כך שהפולינום $g(X) = a_0 + a_1X + \cdots + a_{n^2}X^{n^2}$ מקיים $g \neq 0$, $g(A) = 0$, כלומר, כך ש־$a_0, a_1, \ldots, a_{n^2}$ לא כולם אפס ו־$a_0I_n + a_1A + \cdots + a_{n^2}A^{n^2} = 0$. במלים אחרות, צריך להוכיח: $I_n, A, \ldots, A^{n^2} \in \mathrm{M}_n(F)$ תלויות לינארית מעל F. וזה נכון, כי $\dim \mathrm{M}_n(F) = n^2$. ∎

משפט 8.10 (Cayley-Hamilton): *תהי* $A \in \mathrm{M}_n(F)$. *אז* $f_A(A) = 0$.

לא נוכיח את המשפט. נעיר רק ש"ההוכחה" הבאה **איננה נכונה:**

הוכחה שגויה: הצב A במקום X בשוויון $f_A(X) = |XI_n - A|$. אז $f_A(A) = |AI_n - A| = 0$. (שים לב: $f_A(A)$ היא מטריצה ו־$|AI_n - A|$ הוא סקלר.) ∎

הגדרה 8.11: יהיו $f(X), g(X) \in F[X]$ שני פולינומים שונים מ־0. נאמר ש־$g(X)$ **מחלק את** $f(X)$ אם יש פולינום $h(X)$ כך ש־$f(X) = g(X)h(X)$. ידוע שאם יש $h(X)$ כזה ב־$\mathbb{C}[X]$, אז $h(X) \in F[X]$.

איך יודעים אם $g(X)$ מחלק את $f(X)$? לפי המשפט היסודי של אלגברה אפשר לכתוב

$$f(X) = c(X-\lambda_1)^{n_1}\cdots(X-\lambda_k)^{n_k}, \qquad 0 \neq c \in F$$
$$g(X) = c'(X-\lambda_1)^{m_1}\cdots(X-\lambda_k)^{m_k}, \qquad 0 \neq c' \in F$$

באשר $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ שונים זה מזה, ו־$0 \leq m_i, n_i, \ell_i$ לכל i. (אם $n_i = 0$, אז $(X-\lambda_i)$ כלל אינו מופיע בפירוק של $f(X)$; אך הוא יכול להופיע בפירוק של $g(X)$.)

נניח שיש $h(X)$ כך ש־$f(X) = g(X)h(X)$. אז, בלי הגבלת הכלליות

$$h(X) = c''(X-\lambda_1)^{\ell_1}\cdots(X-\lambda_k)^{\ell_k}, \qquad \ell_1, \ldots, \ell_k \geq 0 \quad ,0 \neq c'' \in F.$$

מכאן $f(X) = g(X)h(X) = c' \cdot c''(X-\lambda_1)^{m_1+\ell_1}\cdots(X-\lambda_k)^{m_k+\ell_k}$ ולכן $n_i = m_i + \ell_i$ לכל i. בפרט $m_i \leq n_i$ לכל i.

להיפך, אם $m_i \leq n_i$ לכל i, נגדיר $h(X)$ כמו לעיל, באשר $c'' = \frac{c}{c'}$ ו־$\ell_i = n_i - m_i \geq 0$. אז $f(X) = g(X)h(X)$, כלומר, $g(X)$ מחלק את $f(X)$. ∎

תרגיל 8.12: $f(X)$ מתחלק ב־$(X-\lambda)$ אם ורק אם $f(\lambda) = 0$.

משפט 8.13: *תהי $A \in \mathrm{M}_n(F)$. אז קיים פולינום מתוקן (כלומר, בעל מקדם עליון 1) יחיד $m_A(X) \in F[X]$ כך ש־*

(א) $m_A(A) = 0$,

(ב) *אם $g(X)$ פולינום כך ש־$g(A) = 0$ אז $g(X)$ מתחלק ב־$m_A(X)$.*

פולינום זה נקרא ***הפולינום המזערי*** *של A.*

דוגמה 8.14: (א) תהי

$$N = N_n = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & 0 \\ 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \ldots & & & & & \\ 0 & 0 & 0 & \ldots & 1 & 0 \end{pmatrix} = (e_2, e_3, \ldots, e_n, 0) \in \mathrm{M}_n(F).$$

אז $N^2 = N(e_2, e_3, \ldots, e_n, 0) = (Ne_2, Ne_3, \ldots, Ne_n, N0)$, לכן

$$N^2 = (e_3, e_4, \ldots, e_{n-2}, 0, 0), N^3 = (e_4, e_5, \ldots, 0, 0, 0), \ldots,$$
$$N^{n-1} = (e_n, 0, \ldots, 0, 0, 0), N^n = 0 = N^{n+1} = \cdots$$

לכן אם $g=\sum_{i=0}^{k}a_iX^i\in F[X]$, אז

$$g(N)=a_0I_n+a_1N+a_2N^2+\cdots+a_{n-1}N^{n-1}=\begin{pmatrix} a_0 & 0 & 0 & \dots & 0\\ a_1 & a_0 & 0 & \dots & 0\\ a_2 & a_1 & a_0 & \dots & 0\\ \dots & & & & \\ a_{n-1} & a_{n-2} & a_{n-3} & \dots & a_0\end{pmatrix}.$$

לכן $g(N)=0$ אם ורק אם $a_0=\cdots=a_{n-1}=0$. קל לראות שזה שקול לכך ש־g מתחלק ב־X^n. לכן $m_N(X)=X^n$.

(ב) תהי $A=I_n$ ויהי $g=\sum_{i=0}^{k}a_iX^i\in F[X]$. אז $g(I_n)=\sum_{i=0}^{k}a_iI_n=(\sum_{i=0}^{k}a_i)I_n=g(1)I_n$. לכן $g(I_n)=0$ אם ורק אם $g(1)=0$ אם ורק אם $(X-1)$ מופיע בפירוק של $g(X)$, כלומר, $(X-1)$ מחלק את $g(X)$. לכן $m_A=X-1$.

משפט 8.15: *תהי* $A\in \mathrm{M}_n(F)$.

(א) *f_A מתחלק ב־* m_A.

(ב) *כל שורש של m_A הוא שורש של f_A (כלומר ע"ע של A)* **וגם להיפך**.

(ג) *למטריצות דומות אותו פולינום מזערי.*

(ד) *A דומה לאלכסונית אם ורק אם $m_A(X)=(X-\lambda_1)\cdots(X-\lambda_k)$, באשר $\lambda_1,\dots,\lambda_k$ שונים זה מזה.*

(ה) *$m_A(X)$ הוא הפולינום המתוקן מהמעלה הקטנה ביותר כך ש־* $m_A(A)=0$.

דוגמה 8.16: *חישוב $m_A(X)$ של A אלכסונית בלי (ד).* תהי $A=\mathrm{diag}(\overbrace{\lambda_1,\dots,\lambda_1}^{n_1},\overbrace{\lambda_2,\dots,\lambda_2}^{n_2},\overbrace{\lambda_k,\dots,\lambda_k}^{n_k})$, באשר $\lambda_1,\lambda_2,\dots,\lambda_k$ שונים זה מזה. אז $f_A(X)=(X-\lambda_1)^{n_1}(X-\lambda_2)^{n_2}\cdots(X-\lambda_k)^{n_k}$. כיוון של־$f_A(X),m_A(X)$ אותם השרשים, $m_A(X)=(X-\lambda_1)^{m_1}(X-\lambda_2)^{m_2}\cdots(X-\lambda_k)^{m_k}$, עבור $m_1,m_2,\dots,m_k\ge 1$ מתאימים. כיוון ש־$m_A(X)$ הוא הפולינום המתוקן מהמעלה הקטנה ביותר שמאפס את A, צריך לבחור $m_1,m_2,\dots,m_k\ge 1$ הקטנים ביותר כך שאגף ימין של המשוואה האחרונה יאפס את A. ניקח $m_1=m_2=\cdots=m_k=1$, ונבדוק שהפולינום $h(X)=(X-\lambda_1)(X-\lambda_2)\cdots(X-\lambda_k)$ מאפס את A ואז $m_A(X)=h(X)$:

$$\begin{aligned}h(A)&=(A-\lambda_1I_n)(A-\lambda_2I_n)\cdots(A-\lambda_kI_n)=\\ &\mathrm{diag}(\overbrace{0\quad,\dots,\quad 0}^{n_1},\overbrace{\lambda_2-\lambda_1,\dots,\lambda_2-\lambda_1}^{n_2},\overbrace{\lambda_k-\lambda_1,\dots,\lambda_k-\lambda_1}^{n_k})\\ &\mathrm{diag}(\overbrace{\lambda_1-\lambda_2,\dots,\lambda_1-\lambda_2}^{n_1},\overbrace{0\quad,\dots,\quad 0}^{n_2},\overbrace{\lambda_k-\lambda_1,\dots,\lambda_k-\lambda_1}^{n_k})\cdots\\ &\mathrm{diag}(\overbrace{\lambda_1-\lambda_k,\dots,\lambda_1-\lambda_k}^{n_1},\overbrace{\lambda_2-\lambda_k,\dots,\lambda_2-\lambda_k}^{n_2},\overbrace{0\quad,\dots,\quad 0}^{n_k})=0\end{aligned}$$

לפי תרגיל 7.6 ∎

דוגמה 8.17: תהי $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. מהו $m_A(X)$? כיוון ש־$f_A(X) = |XI_3 - A| = (X-1)^3$, הפולינום המזערי הוא $(X-1)^r$, באשר $1 \le r \le 3$ ו־r הוא הקטן ביותר כך שפולינום זה מאפס את A. בדיקה מראה ש־$A - I \neq 0$ (ולכן $r > 1$) אך $(A-I)^2 = 0$. לכן $r = 2$. לפי המשפט, A אינה דומה לאלכסונית.

∎

נזכור שפולינום $g(X) \in F[X]$ מתחלק בפולינום $X - \lambda$ אם ורק אם $g(\lambda) = 0$. לכן כל פולינום $0 \neq g(X) \in F[X]$ אפשר להציג באופן הבא $g(X) = (X-\lambda)^m h(X)$, באשר $h(\lambda) \neq 0$, ו־$m \geq 0$. מספר m הוא **החזקה של** $(X - \lambda)$ **ב־**$g(X)$.

הגדרה 8.18: תהי $A \in \mathrm{M}_n(F)$. יהי $\lambda \in F$ ערך עצמי שלה.

(א) **הריבוי האלגברי של** λ הוא החזקה של $(X-\lambda)$ ב־$f_A(X)$.

(ב) **הריבוי הגיאומטרי של** λ הוא $\dim V_\lambda$.

משפט 8.19: *תהי $A \in \mathrm{M}_n(F)$ ויהי $\lambda \in F$ ערך עצמי שלה.*

(א) *הריבוי האלגברי$\geq$הריבוי הגיאומטרי≥ 1.*

(ב) *A דומה לאלכסונית אם ורק אם $f_A(X) = (X-\lambda_1)^{n_1} \cdots (X-\lambda_k)^{n_k}$, באשר $\lambda_1, \ldots, \lambda_k \in F$ שונים זה מזה והריבוי האלגברי n_i הוא הריבוי הגיאומטרי של λ_i, לכל i.*

הוכחה: (א) יהי $d = \dim V_\lambda$ הריבוי הגיאומטרי של λ. יהי $v_1, \ldots, v_d$ בסיס של V_λ. נשלים אותו לבסיס $v_1, \ldots, v_d, \ldots, v_n$ של F^n. תהי $P = (v_1, \ldots, v_n) \in \mathrm{M}_n(F)$. אז P הפיכה ולכל $1 \le i \le d$ מתקיים $Pe_i = v_i$ ולכן $P^{-1}v_i = e_i$. כמו כן $Av_i = \lambda v_i$ עבור $i = 1, \ldots, d$. לכן

$$P^{-1}AP = P^{-1}A(v_1, \ldots, v_d, *, \ldots, *) = (P^{-1}Av_1, \ldots, P^{-1}Av_d, *, \ldots, *) =$$
$$(\lambda P^{-1}v_1, \ldots, \lambda P^{-1}v_d, *, \ldots, *) = (\lambda e_1, \ldots, \lambda e_d, *, \ldots, *) = \begin{pmatrix} \lambda I_d & B \\ \mathbf{0} & C \end{pmatrix}$$

עבור איזה מטריצות $B \in \mathrm{M}_{d\times(n-d)}(F), C \in \mathrm{M}_{n-d}(F)$. פיתוח לפי d העמודות הראשונות נותן

$$f_A(X) = f_{P^{-1}AP}(X) = \begin{vmatrix} (X-\lambda)I_d & -B \\ \mathbf{0} & XI_{n-d} - C \end{vmatrix} = (X-\lambda)^d \cdot |XI_{n-d} - C| =$$
$$= (X-\lambda)^d \cdot f_C(X).$$

לכן החזקה של $(X-\lambda)$ בפולינום האופייני $f_A(X)$ הינה לפחות d.

(ב) יהיו $\lambda_1, \ldots, \lambda_k$ כל הערכים העצמיים השונים של A ב־F. יהיו $d_1, \ldots, d_k$ הריבויים הגיאומטריים שלהם, ויהיו $n_1, \ldots, n_k$ הריבויים האלגבריים שלהם. אז $f_A(X) = (X-\lambda_1)^{n_1} \cdots (X-\lambda_k)^{n_k} g(X)$ באשר $g(X) \in F[X]$ ללא שרשים ב־F ו־$\deg g = n - (n_1 + \cdots + n_k)$. לפי (א), $1 \le d_i \le n_i$ לכל i.

לפי משפט 7.9, A דומה לאלכסונית אם ורק אם יש ל־F^n בסיס המורכב מוקטורים עצמיים של A.

נניח קודם כי $f_A(X) = (X-\lambda_1)^{n_1}\cdots(X-\lambda_k)^{n_k}$ ו־$d_i = n_i$ לכל i. לכל i יש בסיס $\mathcal{B}_i$ של V_{λ_i} בן $d_i = n_i$ איברים. לפי משפט 7.11 צירוף הבסיסים האלה היא סדרה $\mathcal{B}$ בלתי תלויה לינארית. מספר איבריה הוא

$$\sum_i d_i = \sum_i n_i = \deg f_A = n = \dim F^n,$$

לכן $\mathcal{B}$ בסיס של F^n, שמורכב מוקטורים עצמיים של A.

להיפך, נניח שיש בסיס $\mathcal{B}$ של F^n שמורכב מוקטורים עצמיים של A. כל וקטור עצמי שייך לאיזה ערך עצמי, לכן $\mathcal{B}$ צירוף של התת סדרות $\mathcal{B}_1,\dots,\mathcal{B}_k$, באשר $\mathcal{B}_i$ מורכבת מאותם איברי $\mathcal{B}$ שנמצאים ב־V_{λ_i}. לכן

$$n = \dim F^n = \#\mathcal{B} = \sum_i \#\mathcal{B}_i \le \sum_i d_i \le \sum_i n_i \le \deg f_A = n.$$

מכאן $d_i = n_i$ לכל i ו־$\sum_i n_i = n$, לכן $g(X) = 1$, כלומר $f_A(X) = (X-\lambda_1)^{n_1}\cdots(X-\lambda_k)^{n_k}$.

∎

הגדרה 8.20: יהי $\lambda \in F$. מטריצה

$$J_n(\lambda) = \begin{pmatrix} \lambda & 0 & 0 & \cdots & 0 \\ 1 & \lambda & 0 & \cdots & 0 \\ 0 & 1 & \lambda & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & \lambda \end{pmatrix} = \lambda I_n + N_n \in \mathrm{M}_n(F)$$

נקראת **גוש ז'ורדן** מסדר n. מקרים פרטיים: $J_1(\lambda) = (\lambda) \in \mathrm{M}_1(F)$, $J_2(\lambda) = \begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix} \in \mathrm{M}_2(F)$.

תהי $A = J_n(\lambda)$. אז $f_A(X) = (X-\lambda)^n$. לכן נקרא ל־λ **הערך העצמי** של A. נשים לב ש־$(A-\lambda I_n)^k = N_n^k$ היא אפס אם ורק אם $k \ge n$. לכן $m_A(X) = (X-\lambda)^n = f_A(X)$.

הגדרה 8.21: (א) יהי $\lambda \in F$. **מטריצת ז'ורדן השייכת ל־λ** היא מטריצה מהצורה

$$J(\lambda) = \mathrm{diag}(J_{n_1}(\lambda),\dots,J_{n_k}(\lambda)), \qquad n_1 \ge \cdots \ge n_k \ge 1.$$

(ב) **מטריצת ז'ורדן** היא מטריצה מהצורה $J = \mathrm{diag}(J(\lambda_1),\dots,J(\lambda_r))$, באשר $J(\lambda_i)$ מטריצת ז'ורדן השייכת ל־λ_i, לכל i, ו־$\lambda_1,\dots,\lambda_r$ שונים זה מזה.

משפט 8.22: *בסימונים לעיל*

(א) $f_J(X) = f_{J(\lambda_1)}(X)\cdots f_{J(\lambda_r)}(X)$, $m_J(X) = m_{J(\lambda_1)}(X)\cdots m_{J(\lambda_r)}(X)$.

(ב) $f_{J(\lambda)}(X) = (X-\lambda)^{n_1+\cdots+n_k}$, $m_{J(\lambda)}(X) = (X-\lambda)^{n_1}$.

משפט 8.23: *תהי $A \in \mathrm{M}_n(F)$. נניח שהפולינום האופייני שלה הוא מכפלה של פולינומים ממעלה ראשונה (זה תמיד נכון, אם $F = \mathbb{C}$). אז קיימת מטריצת ז'ורדן $J = \mathrm{diag}(J(\lambda_1),\dots,J(\lambda_r))$ הדומה ל־A, והיא יחידה, עד כי הסדר של $J(\lambda_1),\dots,J(\lambda_r)$. הסקלרים $\lambda_1,\dots,\lambda_r$ הם הערכים העצמיים של A. מטריצה J נקראת* **צורת ז'ורדן של A**.

מסקנה 8.24: *יהי V מרחב וקטורי נוצר סופית מעל $\mathbb{C}$. תהי $T\colon V \to V$ העתקה לינארית. אז יש בסיס $\mathcal{B}'$ של V כך ש־ $[T]^{\mathcal{B}'}_{\mathcal{B}'}$ בצורת ז'ורדן.*

הוכחה: נבחר בסיס $\mathcal{B}$ כלשהו של V. לפי המשפט, $[T]^{\mathcal{B}}_{\mathcal{B}}$ דומה למטריצת ז'ורדן J, כלומר, יש P הפיכה כך ש־$P^{-1}[T]^{\mathcal{B}}_{\mathcal{B}}P = J$. לפי תרגיל 6.33, P היא מטריצת מעבר מ־$\mathcal{B}$ לאיזושהי סדרה $\mathcal{B}'$ של איברי V. לפי משפט 6.34, $\mathcal{B}'$ בסיס של V. לפי משפט 6.37, $[T]^{\mathcal{B}'}_{\mathcal{B}'} = P^{-1}[T]^{\mathcal{B}}_{\mathcal{B}}P$. לכן $[T]^{\mathcal{B}'}_{\mathcal{B}'} = J$. ∎

מסקנה 8.25: *תהיינה $A, B \in \mathrm{M}_n(F)$ שיש להן צורות ז'ורדן. אז הן דומות אם ורק אם יש להן אותה צורת ז'ורדן, עד כדי סדר מטריצות־ז'ורדן השייכות לסקלרים.*

תרגיל 8.26: המטריצות הבאות אינן דומות מעל שום שדה F (המוכל ב־$\mathbb{C}$):

$$B = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \qquad A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix}.$$

אכן, לו היתה קיימת $P \in \mathrm{M}_n(F)$ הפיכה כך ש־$P^{-1}AP = B$, כיוון ש־$P \in \mathrm{M}_n(\mathbb{C})$, היו A, B דומות מעל $\mathbb{C}$. אך A, B מטריצות ז'ורדן שונות (בסימונים של ההגדרה 8.21 $r = 1$ ו־$\lambda = \lambda_1 = 2$), ולכן אין דומות מעל $\mathbb{C}$.

תרגיל 8.27: מהי צורת ז'ורדן J של A אם $f_A(X) = (X-1)^3(X-2)^2$ ו־$m_A(X) = (X-1)^2(X-2)^2$?

לפי משפט 8.8, $f_A(X) = f_J(X)$. לפי משפט 8.15(ג), $m_A(X) = m_J(X)$. כיוון ש־$f_J(X)$ ממעלה 5, $J \in \mathrm{M}_5(F)$. מתקיים $J = \mathrm{diag}(J(1), J(2))$, באשר $J(1)$ מטריצת ז'ורדן השייכת ל־1 ו־$J(2)$ מטריצת ז'ורדן השייכת ל־2. לפי משפט 8.22(א), $f_{J(1)}(X) = (X-1)^3$ ו־$f_{J(2)}(X) = (X-2)^2$. לכן $J(1)$ מסדר 3×3 ו־$J(2)$ מסדר 2×2. כמו כן, לפי אותו המשפט, $m_{J(1)}(X) = (X-1)^2$ ו־$m_{J(2)}(X) = (X-2)^2$.

אם $J(1) = \mathrm{diag}(J_{n_1}(1), \ldots, J_{n_r}(1))$, באשר $n_1 \geq \cdots \geq n_r \geq 1$, אז $\sum n_i = 3$ ולפי משפט 8.22(ב), $n_1 = 2$. מכאן $r = 2$, $n_1 = 2, n_2 = 1$. לכן $J(1) = \mathrm{diag}(J_2(1), J_1(1))$.

אם $J(2) = \mathrm{diag}(J_{\ell_1}(2), \ldots, J_{\ell_s}(2))$, באשר $\ell_1 \geq \cdots \geq \ell_s \geq 1$, אז $\sum \ell_i = 2$ ולפי משפט 8.22(ב), $\ell_1 = 2$. מכאן $s = 1$, $\ell_1 = 2$. לכן $J(2) = J_2(2)$.

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

נעיר שהפולינום האופייני והמזערי של מטריצה לא תמיד קובעים את צורת ז'ורדן שלה.

8א. מציאת צורת ז'ורדן למטריצות מסדר 2×2

תרגיל 8א.1: *תהיינה $A, B \in \mathrm{M}_n(F)$ דומות, ויהי $\lambda \in F$ ערך עצמי שלהן. אז הריבוי הגיאומטרי של λ ב־A, B שווה.*

הוכחה: יהיו V_λ^A, V_λ^B המרחבים העצמיים המתאימים ל־A, B. לפי ההנחה יש $P \in \mathrm{M}_n(F)$ הפיכה כך ש־$B = P^{-1}AP$.

אם $v \in V_\lambda^A$, כלומר $PBP^{-1}v = Av = \lambda v$, אז $BP^{-1}v = \lambda P^{-1}v$, לכן $P^{-1}v \in V_\lambda^B$.

יהי $v_1, \ldots, v_r$ בסיס של V_λ^A. לפי האמור לעיל $P^{-1}v_1, \ldots, P^{-1}v_r \in V_\lambda^B$. ההעתקה הלינארית $Pv \mapsto v$ מעתיקה אותם על $v_1, \ldots, v_r$ ולכן הם בלתי תלויים לינארית. מכאן $\dim V_\lambda^B \geq r = \dim V_\lambda^A$.

באופן דומה מראים את אי השוויון ההפוך. ∎

תרגיל 8א.2: *תהי $A \in \mathrm{M}_n(F)$ ויהי $\lambda \in F$ ערך עצמי של A. אז הריבוי הגיאומטרי של λ ב־A הוא מספר גושי ז'ורדן מהצורה*

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

במטריצת ז'ורדן של A.

הוכחה: לפי התרגיל הקודם אפשר להניח ש־A בצורת ז'ורדן. יהי k מספר גושי ז'ורדן הנזכרים לעיל. אז ב־$A - \lambda I_n$ יש (כאן רצוי להתבונן בדוגמה של מטריצה כזו) בדיוק k שורות של אפסים ־ בכל גוש ז'ורדן השורה הראשונה היא שורת אפסים ־ ואחרי פעולות אלמנטריות פשוטות מקבלים מ־$A - \lambda I_n$ מטריצה מדורגת קנונית עם k שורות של אפסים בדיוק. לכן V_λ, שהינו מרחב הפתרונות של המערכת ההומוגנית $(A - \lambda I_n)X = 0$, הוא ממימד k. ∎

תהי $A \in \mathrm{M}_2(F)$. נניח שהפולינום האופייני שלה $f_A(X)$ מהצורה $f_A(X) = (X - \lambda_1)(X - \lambda_2)$, באשר $\lambda_1, \lambda_2 \in F$. (זה קורה אם, למשל, $F = \mathbb{C}$, לפי המשפט היסודי של אלגברה.)

למדנו שאז יש $P \in \mathrm{M}_2(F)$ הפיכה כך ש־$P^{-1}AP$ מטריצת ז'ורדן. כעת נסביר כיצד למצוא P כזאת.

(א) אם $\lambda_1 \neq \lambda_2$, אז A דומה לאלכסונית, ולמדנו: אם $0 \neq v_1 \in F^2$ וקטור עצמי של A השייך ל־λ_1, ו־$0 \neq v_2 \in F^2$ וקטור עצמי של A השייך ל־λ_2, ואם נגדיר $P = (v_1, v_2) \in \mathrm{M}_2(F)$ אז $P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ אלכסונית, ובפרט מטריצת ז'ורדן.

נניח, אם כן, כי $\lambda_1 = \lambda_2$; נסמן $\lambda = \lambda_1 = \lambda_2$. אז $f_A(X) = (X - \lambda)^2$ והריבוי האלגברי של λ הוא 2. לכן הריבוי הגיאומטרי $\dim V_\lambda$ של λ הוא 2 או 1.

(ב) אם הריבוי הגיאומטרי של λ הוא 2, אז $V_\lambda = F^2$, ולכן בפרט e_1, e_2 הם וקטורים עצמיים של A. מכאן $A = (Ae_1, Ae_2) = (\lambda e_1, \lambda e_2) = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$ כבר אלכסונית בעצמה. (נקח $P = I_2$.)

(ג) אם הריבוי הגיאומטרי של λ הוא 1, אז $V_\lambda \subsetneqq F^2$. נבחר $v_1 \in F^2$ כך ש־$v_1 \notin V_\lambda$. נגדיר

$$v_2 = (A-\lambda I_2)v_1.$$

אז

(1) $v_2 \neq 0$, כי $v_1 \notin V_\lambda$.

(2) $v_2 \in V_\lambda$: אכן, לפי משפט קיילי־המילטון, $(A-\lambda I_2)^2$ היא מטריצת האפס, לכן $(A-\lambda I_2)^2 v_1 = 0$. מכאן $(A-\lambda I_2)v_2 = 0$.

(3) v_1, v_2 בלתי תלוים לינארית מעל F: לפי (1) $v_2 \neq 0$. אילו היה v_1 כפולה שלו, לפי (2) הוא היה שייך ל־V_λ, בסתירה לבחירתו.

(4) $Av_1 = \lambda v_1 + v_2$.

נגדיר $P = (v_1, v_2) \in \mathrm{M}_2(F)$. לפי (3), P הפיכה.

מתקיים $Pe_i = v_i$, ולכן $e_i = P^{-1}v_i$, עבור $i = 1,2$. לכן

$$P^{-1}AP = P^{-1}A(v_1,v_2) = P^{-1}(Av_1, Av_2) = P^{-1}(\lambda v_1 + v_2, \lambda v_2) =$$
$$(P^{-1}\lambda v_1 + P^{-1}v_2, P^{-1}\lambda v_2) = (\lambda e_1 + e_2, \lambda e_2) = \begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix}$$

מטריצת ז'ורדן. ∎

דוגמה 8.3א: $A = \begin{pmatrix} -18 & -25 \\ 16 & 22 \end{pmatrix}$. אז $f_A(X) = X^2 - 4X + 4 = (X-2)^2$.

נבחר $v_1 = \begin{pmatrix} 4 \\ -3 \end{pmatrix}$; אז $v_2 = \begin{pmatrix} -20 & -25 \\ 16 & 20 \end{pmatrix}\begin{pmatrix} 4 \\ -3 \end{pmatrix} = \begin{pmatrix} -5 \\ 4 \end{pmatrix}$. לכן $P = \begin{pmatrix} 4 & -5 \\ -3 & 4 \end{pmatrix}$.

ומתקיים $P^{-1}AP = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}$.

נמצא את A^m כאשר m מספר טבעי (גדול).

נסמן $N = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. אז $P^{-1}AP = 2I_2 + N$. היות ו־$(2I_2)N = N(2I_2)$, מקבלים לפי נוסחת הבינום של ניוטון

$$P^{-1}A^mP = (P^{-1}AP)^m = (2I_2+N)^m = \sum_{i=1}^m \binom{m}{i} N^i (2I_2)^{m-i}.$$

אבל $N^2 = N^3 = \cdots = 0$, לכן

$$P^{-1}A^mP = \binom{m}{0}N^0(2I_2)^m + \binom{m}{1}N^1(2I_2)^{m-1} = 2^mI_2 + 2^{m-1}N = \begin{pmatrix} 2^m & 0 \\ 2^{m-1} & 2^m \end{pmatrix}.$$

מכאן

$$A^m = P(2^mI_2 + 2^{m-1}N)P^{-1} = \begin{pmatrix} 4 & -5 \\ -3 & 4 \end{pmatrix}\begin{pmatrix} 2^m & 0 \\ 2^{m-1} & 2^m \end{pmatrix}\begin{pmatrix} 4 & 5 \\ 3 & 4 \end{pmatrix}. \quad ∎$$

9. מרחבי מכפלה פנימית.

בפרק זה יהיה $F = \mathbb{R}$ או $F = \mathbb{C}$.

נזכור שלכל $z \in \mathbb{C}$ קיים **הצמוד** שלו $\overline{z}$ המוגדר על ידי

$$\overline{x + iy} = x - iy, \qquad x, y \in \mathbb{R}$$

התכונות הבסיסיות:

(א) $\overline{z_1 \pm z_2} = \overline{z_1} \pm \overline{z_2}$;

(ב) $\overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}$;

(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 9.1: יהי V מרחב וקטורי מעל F. **מכפלה פנימית על** V היא העתקה שמתאימה לכל זוג $u, v \in V$ של וקטורים 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), $\langle v, v\rangle \in \mathbb{R}$.)

דוגמה 9.2: (א) **המכפלה הסטנדרטית**, נקראת גם **סקלרית**: $V = F^n$, ונגדיר $\langle u, v\rangle_{\mathrm{st}} = \overline{u}^t v$, כלומר,

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

נבדוק שהתכונה (4) בהגדרה מתקיימת: נניח $a_1, a_2, \ldots, a_n \in F$ 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:

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

כאשר $a_2 \neq 0$ או $a_1 + a_2 \neq 0$, כלומר, כאשר $a_2 \neq 0$ או $a_1 \neq 0$.

למה 9.3:

(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 9.4:

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

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

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

(4) סדרה $u_1, u_2, \ldots, u_n \in V$ נקראת **אורתונורמלית** אם היא אורתוגונלית ו־$||u_i|| = 1$ לכל i. במלים אחרות, $\langle u_i, u_j\rangle = \delta_{ij} := \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$.

דוגמה 9.5: הבסיס הסטנדרטי $e_1, e_2, \ldots, e_n$ של F^n הינו אורתונורמלי ביחס למכפלה הסטנדרטית.

הערה 9.6: יהי $V = \mathbb{R}^2$ עם המכפלה הסטנדרטית. נראה את איברי V כחצים מהראשית. אם $A = (a_1, a_2)$ אז אורכו של $\overrightarrow{OA}$ הוא $\sqrt{a_1^2 + a_2^2}$, כלומר, הנורמה של $\overrightarrow{OA}$. כמו כן, תהי גם $B = (b_1, b_2)$. אז, שוב לפי משפט פיתגורס, $\overrightarrow{OA}, \overrightarrow{OB}$ 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, $\langle\overrightarrow{OA}, \overrightarrow{OB}\rangle_{\mathrm{st}} = 0$.

באופן דומה האורך והניצבות ב־$V = \mathbb{R}^3$ ביחס למכפלה הסטנדרטית הם האורך והניצבות המוכרים לנו מגיאומטריה.

למה 9.7 (למת האפיון)**:** *יהי V מרחב מכפלה פנימית. תהי $u_1, u_2, \ldots, u_n$ סדרה אורתוגונלית ב־V. יהיו $u = a_1u_1 + \cdots + a_nu_n, \quad v = b_1u_1 + \cdots + b_nu_n \in V$. אז*

(א) $\langle u_i, u\rangle = a_i||u_i||^2$. *בפרט, אם* $u_i \neq 0$, *אז* $a_i = \frac{\langle u_i, u\rangle}{||u_i||^2}$. *בפרט, אם* $||u_i|| = 1$ *לכל* i, *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) $a_i||u_i||^2 = a_i\langle u_i, u_i\rangle = a_1\langle u_i, u_1\rangle + a_2\langle u_i, u_2\rangle + \cdots + a_n\langle u_i, u_n\rangle = \langle u_i, u\rangle$.

(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 9.8: *יהי $\mathcal{B}$ בסיס אורתונורמלי של מרחב מכפלה פנימית V. יהיו $u,v \in V$. אז* $\langle u,v\rangle = \langle [u]_{\mathcal{B}}, [v]_{\mathcal{B}}\rangle_{\text{st}}$
(באגף ימין המכפלה היא הסטנדרטית.)

משפט 9.9: *תהי $u_1,u_2,\dots,u_n$ סדרה אורתוגונלית, שאיבריה שונים מאפס. אז $u_1,u_2,\dots,u_n$ בלתי תלויים לינארית מעל F.*

הוכחה: נניח $a_1u_1+a_2u_2+\cdots+a_nu_n=0$. לפי למת האפיון $0=\frac{0}{||u_i||^2}=\frac{\langle u_i,0\rangle}{||u_i||^2}=a_i$, לכל i. ∎

תרגיל 9.10: *אם v ניצב ל $v_1,\dots,v_k$ אז הוא ניצב לכל v ב־ $\mathrm{Sp}(v_1,\dots,v_k)$.*

משפט 9.11 (תהליך Gram-Schmidt): *יהי V מרחב מכפלה פנימית מעל F.*

תהי $v_1,v_2,\dots,v_n$ סדרה ב־ V. אז יש סדרה אורתוגונלית $u_1,u_2,\dots,u_n$ ב־ V שמקיימת:

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

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

בפרט

(3) $u_1,u_2,\dots,u_n$ *שונים מאפס אם ורק אם* $v_1,v_2,\dots,v_n$ *בלתי תלוים לינארית מעל* F.

(4) *אם* $v_1,v_2,\dots,v_n$ *בסיס של* V *אז* $u_1,u_2,\dots,u_n$ *בסיס של* V.

הוכחה: תחילה נשים לב ש־(3) נובע מ־(2) ו־(4) נובע מ־(1^n).

נבנה באינדוקציה על k את הסדרה כך שיתקיימו תנאים (1^k), (2^k), וכן

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

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

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

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

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

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

(1^{k+1}): לפי $(*)$ ולפי (1^k):

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

(0^{k+1}): $u_1,u_2,\dots,u_k,u_{k+1}$ אורתוגונלית: די להוכיח $\langle u_i,u_{k+1}\rangle=0$ לכל $1\le i\le k$. ואכן

$$\langle u_i,u_{k+1}\rangle=\langle u_i,v_{k+1}\rangle-\sum_{j=1}^k c_j\langle u_i,u_j\rangle=\langle u_i,v_{k+1}\rangle-c_i||u_i||^2=0.$$

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

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

להיפך, אם $v_{k+1} \in \mathrm{Sp}(v_1, v_2, \ldots, v_k) = \mathrm{Sp}(u_1, u_2, \ldots, u_k)$ אז $v_{k+1} = \sum_{i=1}^{k} a_i u_i$ עבור איזה $a_1, a_2, \ldots, a_k \in F$. לפי למת האפיון, $a_i = c_i$ אם $u_i \neq 0$, לכל i. לכן לפי $(*)$, $u_{k+1} = 0$ ∎

דוגמה 9.12: נעשה תהליך גרם־שמידט על $v_1 = (0,3,4), v_2 = (1,2,1), v_3 = (0,0,2)$.

(א) $u_1 = v_1 = (0,3,4)$, $||u_1||^2 = 0^2 + 3^2 + 4^2 = 25$,

(ב) $\langle u_1, v_2 \rangle_{\mathrm{st}} = 0 \cdot 1 + 3 \cdot 2 + 4 \cdot 1 = 10$,

$$u_2 = (1,2,1) - \tfrac{10}{25}(0,3,4) = (1,2,1) - (0, \tfrac{6}{5}, \tfrac{8}{5}) = (1, \tfrac{4}{5}, \tfrac{-3}{5}) = \tfrac{1}{5}(5,4,-3),$$
$$||u_2||^2 = \tfrac{25}{25} + \tfrac{16}{25} + \tfrac{9}{25} = 2.$$

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

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

מסקנה 9.13: *יהי V מרחב מכפלה פנימית נוצר סופית. אז יש לו בסיס אורתונורמלי.*

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

הערה 9.14: *המשמעות הגיאומטרית של תהליך גרם־שמידט.* נתבונן בנוסחה $(*)$. נסמן

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

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

נניח כי $V = \mathbb{R}^3$ ו־U הוא מישור או ישר. אז

(א) u הוא ההיטל של v_{k+1} על U,

(ב) u_{k+1} הוא החץ מאיזושהי נקודה ב־U, ניצב ל־U אל הקצה של v_{k+1}. בפרט אורכו $||u_{k+1}||$ הוא המרחק בין U לנקודת הקצה של v_{k+1}.

תרגיל 9.15: *מצא את מרחק הנקודה* $(0,0,2)$ *מהמישור* $U = \mathrm{Sp}((0,3,4),(1,2,1))$

פתרון: בדוגמה 9.12 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 9.16: *אם מבצעים תהליך גרם־שמידט על $v_1, \ldots, v_r, v_{r+1}, \ldots, v_n$ ו־ $v_1, \ldots, v_r$ סדרה אורתוגונלית, אז* $u_1 = v_1, u_2 = v_2, \ldots, u_r = v_r$.

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

נניח $u_1 = v_1, \ldots, u_k = v_k$, באשר $1 \le k < r$. 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 9.17: *יהי V מרחב מכפלה פנימית נוצר סופית. תהי $v_1, \ldots, v_r$ סדרה אורתוגונלית, שאיבריה שונים מאפס. אז אפשר להשלימה לבסיס אורתוגונלי של V. אם $v_1, \ldots, v_r$ אורתונורמלית אז אפשר להשלימה לבסיס אורתונורמלי של V.*

הוכחה: לפי משפט 9.9, $v_1, \ldots, v_r$ בלתי תלויה לינארית. לכן ניתן להשלימה לבסיס $v_1, \ldots, v_r, \ldots, v_n$ של V. תהליך גרם־שמידט עליו יתן בסיס אורתוגונלי $u_1, \ldots, u_r, \ldots, u_n$. לפי תרגיל 9.16, $u_1 = v_1, \ldots, u_r = v_r$. אם $v_1, \ldots, v_r$ אורתונורמלית ונחליף את u_i ב־$\frac{1}{||u_i||}u_i$ לכל $r < i \leq n$, נקבל בסיס אורתונורמלי. ∎

משפט 9.18 (אי שוויון Cauchy-Schwarz): *יהי V מרחב מכפלה פנימית ויהיו v_1, v_2 שני וקטורים ב־V. אז מתקיים $||v_1|| \cdot ||v_2|| \leq |\langle v_1, v_2\rangle|$ 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 מתקיים $v_2 = c_1u_1 + u_2$. לפי למת האפיון,

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

ויש שוויון אם ורק אם $||u_2|| = 0$, כלומר, $u_2 = 0$, כלומר (לפי התנאים של תהליך גרם־שמידט), $v_2 \in \mathrm{Sp}(v_1)$, כלומר, v_1, v_2 תלויים לינארית. נכפיל את האי שוויון ב־$||u_1||^2$ כדי לקבל את מבוקש. ∎

משפט 9.19 (אי שוויון המשולש): *יהי V מרחב מכפלה פנימית ויהיו $v_1, v_2 \in V$. אז* $||v_1+v_2|| \leq ||v_1||+||v_2||$.

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

$$\begin{aligned}(||v_1 + v_2||)^2 &= \langle v_1 + v_2, v_1 + v_2\rangle = \langle v_1, v_1\rangle + \langle v_1, v_2\rangle + \langle v_2, v_1\rangle + \langle v_2, v_2\rangle = \\ &||v_1||^2 + \langle v_1, v_2\rangle + \overline{\langle v_1, v_2\rangle} + ||v_2||^2 = ||v_1||^2 + 2\mathrm{Re}(\langle v_1, v_2\rangle) + ||v_1||^2 \leq \\ &||v_1||^2 + 2|\langle v_1, v_2\rangle| + ||v_1||^2 \leq ||v_1||^2 + 2||v_1|| \cdot ||v_2|| + ||v_1||^2 = (||v_1|| + ||v_2||)^2\end{aligned}$$

∎

הגדרה 9.20: יהי V מרחב מכפלה פנימית ויהי U תת מרחב שלו. **המשלים הניצב** (**האורתוגונלי**) הוא

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

למה 9.21:

(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 $\alpha v \in U^\perp$.

(ב) יהי $u \in U \cap U^\perp$. אז $||u||^2 = \langle u, u\rangle = 0$, ולכן $u = 0$. ∎

מציאת המשלים הניצב במרחב נוצר סופית:

יהי $v_1, \ldots, v_r$ בסיס של U. נשלים אותו לבסיס $v_1, \ldots, v_r, \ldots, v_n$ של V. תהליך גרם־שמידט יתן בסיס אורתוגונלי $u_1, \ldots, u_r, \ldots, u_n$ כך ש־$U = \mathrm{Sp}(v_1, \ldots, v_r) = \mathrm{Sp}(u_1, \ldots, u_r)$. ואז:

טענה 9.22: *(א)* $U^\perp = \mathrm{Sp}(u_{r+1}, \ldots, u_n)$, *בפרט* $\dim U^\perp = \dim V - \dim U$.

(ב) $U + U^\perp = V$.

הוכחה: (א) תחילה נראה $U^\perp \supseteq \mathrm{Sp}(u_{r+1}, \ldots, u_n)$. לשם כך די להוכיח שלכל $r+1 \le k \le n$ מתקיים $u_k \in U^\perp$, כלומר, u_k ניצב לכל וקטור ב־U. אבל $U = \mathrm{Sp}(u_1, \ldots, u_r)$ ו־u_k ניצב ל־$u_1, \ldots, u_r$, לכן ניצב לכל וקטור ב־U לפי תרגיל 9.10.

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 האפין, $a_i = 0$. מכאן $v = a_{r+1}u_{r+1} + \cdots + a_nu_n \in \mathrm{Sp}(u_{r+1}, \ldots, u_n)$.

(ב) לפי (א) ולפי משפט 4.50, $U + U^\perp = \mathrm{Sp}(u_1, \ldots, u_r, u_{r+1}, \ldots, u_n) = V$. ∎

מסקנה 9.23: *בתנאים לעיל (V נוצר סופית)* $(U^\perp)^\perp = U$.

בסוף הפרק הזה נבדוק מהן כל המכפלות הפנימיות על מרחב וקטורי נוצר סופית.

לכל $A \in \mathrm{M}_{m\times n}(F)$ נגדיר $\overline{A} \in \mathrm{M}_{m\times n}(F)$ על ידי $(\overline{A})_{ij} = \overline{(A)_{ij}}$.

מתקיים: $\overline{\overline{A}} = A$, $\overline{A+B} = \overline{A} + \overline{B}$, $\overline{\alpha A} = \overline{\alpha}\,\overline{A}$, $\overline{AB} = \overline{A}\,\overline{B}$.

הגדרה 9.24: לכל $A \in \mathrm{M}_{m\times n}(F)$ נגדיר $A^* = \overline{A}^t = \overline{A^t}$.

דוגמה 9.25: (א) $\begin{pmatrix} 2+i & i & 1 \\ 1 & 1-i & 0 \end{pmatrix}^* = \begin{pmatrix} 2-i & 1 \\ -i & 1+i \\ 1 & 0 \end{pmatrix}$.

(ב) המכפלה הסקלרית על F^n נתונה על ידי $\langle u, v\rangle_{\mathrm{st}} = \overline{u}^t v = u^* v$.

(ג) אם A ממשית אז $A^* = A^t$.

תכונות 9.26: $(A^*)^* = A$, $(A+B)^* = A^* + B^*$, $(\alpha A)^* = \overline{\alpha} A^*$, $(AB)^* = B^*A^*$.

הגדרה 9.27: $A \in \mathrm{M}_n(F)$ נקראת **הרמיטית** או **צמודה לעצמה** אם $A^* = A$; אם $F = \mathbb{R}$ ו־$A^t = A^* = A$, אז A נקראת גם **סימטרית**.

דוגמה 9.28: $\begin{pmatrix} 2 & i & 1+i \\ -i & 1 & 2-3i \\ 1-i & 2+3i & 8 \end{pmatrix}$ הרמיטית; I_n הנה צמודה לעצמה, סימטרית.

הערות 9.29: (א) הרכיבים באלכסון הראשי של מטריצה הרמיטית הם ממשיים.

(ב) $A = (a) \in \mathrm{M}_1(F)$ הרמיטית אם ורק אם $a \in \mathbb{R}$.

(ג) אם $A \in \mathrm{M}_n(F)$ הרמיטית ו־$B \in \mathrm{M}_{n\times m}(F)$ כלשהי אז $B^*AB \in \mathrm{M}_m(F)$ הרמיטית (לפי תכונות 9.26). בפרט, אם $w \in F^n$ (ונזהה את w^*Aw עם הרכיב היחיד שלה) אז $w^*Aw \in \mathbb{R}$.

הגדרה 9.30: $A \in \mathrm{M}_n(F)$ הרמיטית נקראת **חיובית גמורה** אם $w^*Aw > 0$ לכל $0 \neq w \in F^n$.

דוגמה 9.31: I_n הנה חיובית גמורה.

נביא ללא הוכחה את האיפיון המעשי הבא של חיוביות גמורה.

משפט 9.32: *תהי $A \in \mathrm{M}_n(F)$ הרמיטית. לכל $1 \le k \le n$ תהי $A^{(k)} \in \mathrm{M}_k(F)$ המטריצה מתקבלת מ־A על ידי השמטת $n-k$ השורות האחרונות ו־$n-k$ העמודות האחרונות. אז A חיובית גמורה אם ורק אם $|A^{(k)}| > 0$ לכל $1 \le k \le n$.*

כך, למשל, $\begin{pmatrix} 1 & 2 & 0 \\ 2 & 5 & 2 \\ 0 & 2 & 8 \end{pmatrix}$ חיובית גמורה, כי $\begin{vmatrix} 1 & 2 & 0 \\ 2 & 5 & 2 \\ 0 & 2 & 8 \end{vmatrix}$, $\begin{vmatrix} 1 & 2 \\ 2 & 5 \end{vmatrix}$, $|\,1\,|$ חיוביים.

הגדרה 9.33: יהי V מרחב וקטורי עם בסיס $\mathcal{B} = (v_1, \dots, v_n)$. תהי $\langle\ ,\ \rangle$ מכפלה פנימית על V. המטריצה $A \in \mathrm{M}_n(F)$ הנתונה על ידי

$$(A)_{ij} = \langle v_i, v_j \rangle$$

נקראת **מטריצת גרם של** $\langle\ ,\ \rangle$ **לפי** $\mathcal{B}$.

דוגמה 9.34: מטריצת גרם לפי בסיס אורתונורמלי $\mathcal{B}$ היא I_n.

למה 9.35: *תהי A מטריצת גרם של מכפלה פנימית $\langle\ ,\ \rangle$ לפי בסיס $\mathcal{B}$. אז לכל $u, v \in V$, $\langle u, v \rangle = ([u]_{\mathcal{B}})^* A [v]_{\mathcal{B}}$.*

הוכחה: יהיו $u, v \in V$. אז יש להן הצגות

$$u = a_1 v_1 + \cdots + a_n v_n, \qquad v = b_1 v_1 + \cdots + b_n v_n.$$

מכאן

$$\langle u, v \rangle = \langle \sum_{i=1}^n a_i v_i, \sum_{i=1}^n b_j v_j \rangle = \sum_{i=1}^n \sum_{i=1}^n \overline{a_i} b_j \langle v_i, v_j \rangle = \sum_{i=1}^n \sum_{j=1}^n \overline{a_i} b_j (A)_{ij}$$

$$([u]_{\mathcal{B}})^* A [v]_{\mathcal{B}} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}^* A \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = (\overline{a_1}, \dots, \overline{a_n}) A \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} =$$

$$\sum_{i=1}^n \overline{a_i} \Big(A \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \Big)_{i1} = \sum_{i=1}^n \overline{a_i} \sum_{j=1}^n (A)_{ij} b_j = \sum_{i=1}^n \sum_{j=1}^n \overline{a_i} b_j (A)_{ij}$$

משפט 9.36: *מטריצת גרם A של מכפלה פנימית $\langle\ ,\ \rangle$ לפי בסיס $\mathcal{B}$ היא*

(א) *הרמיטית;*

(ב) *חיובית גמורה;*

הוכחה: (א) $(A^*)_{ij} = \overline{(A)_{ji}} = \overline{\langle v_j, v_i\rangle} = \langle v_i, v_j\rangle = (A)_{ij}$.

(ב) יהי $0 \neq w \in F^n$. היות ו־$v \mapsto [v]_{\mathcal{B}}$ הוא איזומורפיזם $V \to F^n$, יש $0 \neq v \in V$ כך ש־$w = [v]_{\mathcal{B}}$. לפי למה 9.35, $w^*Aw = [v]_{\mathcal{B}}^*A[v]_{\mathcal{B}} = \langle v, v\rangle > 0$. ∎

משפט 9.37: *יהי V מרחב וקטורי עם בסיס $\mathcal{B} = (v_1, \ldots, v_n)$. תהי $A \in \mathrm{M}_n(F)$ חיובית גמורה. אז קיימת מכפלה פנימית יחידה על V שמטריצת גרם שלה היא A.*

הוכחה: קיום: נגדיר את המכפלה הפנימית על ידי

$$\langle u, v\rangle = [u]_{\mathcal{B}}^*A[v]_{\mathcal{B}}.$$

קל לבדוק שזאת מכפלה פנימית על V. למשל,

$$\langle v, u\rangle = [v]_{\mathcal{B}}^*A[u]_{\mathcal{B}} = \overline{([v]_{\mathcal{B}}^*A[u]_{\mathcal{B}})^*} = \overline{[u]_{\mathcal{B}}^*A^*[v]_{\mathcal{B}}} = \overline{\langle u, v\rangle}.$$

מתקיים $\langle v_i, u_j\rangle = (e_i)^*Ae_j = (A)_{ij}$. לכן A היא מטריצת גרם של מכפלה פנימית זו.

היחידות: לפי למה 9.35 מכפלה פנימית נקבעת על ידי מטריצת גרם שלה. ∎

משפט 9.38: *יהי V מרחב מכפלה פנימית, יהיו $\mathcal{B}, \mathcal{B}'$ שני בסיסים שלו, והתיינה A, A' מטריצות גרם של המכפלה הפנימית לפי $\mathcal{B}, \mathcal{B}'$, בהתאמה. תהי P מטריצת המעבר מ־$\mathcal{B}$ ל־$\mathcal{B}'$. אז $A' = P^*AP$.*

הוכחה: יהיו $u_1', \ldots, u_n'$ איברי $\mathcal{B}'$. לכל $u, v \in V$ מתקיים

$$\langle u, v\rangle = [u]_{\mathcal{B}}^*A[v]_{\mathcal{B}} = (P[u]_{\mathcal{B}'})^*A(P[v]_{\mathcal{B}'}) = [u]_{\mathcal{B}'}^*(P^*AP)[v]_{\mathcal{B}'}$$

לכן $(A')_{ij} = \langle u_i', u_j'\rangle = [u_i']_{\mathcal{B}'}^*(P^*AP)[u_j']_{\mathcal{B}'} = e_i^*(P^*AP)e_j = (P^*AP)_{ij}$. ∎

10. שקילות קוניונקטיבית וחפיפה

גם בפרק זה יהי $F = \mathbb{R}$ או $F = \mathbb{C}$.

הגדרה 10.1: מטריצות $A, B \in \mathrm{M}_n(F)$ נקראות **שקולות קוניונקטיבית** אם יש $P \in \mathrm{M}_n(F)$ הפיכה כך ש־$B = P^*AP$. מטריצות $A, B \in \mathrm{M}_n(\mathbb{R})$ נקראות **חופפות** אם יש $P \in \mathrm{M}_n(\mathbb{R})$ הפיכה כך ש־$B = P^tAP$.

תרגיל 10.2: *תהיינה* $A, B, C, P \in \mathrm{M}_n(F)$.

(א) *אם* P *הפיכה אז* P^* *הפיכה ו־* $(P^*)^{-1} = (P^{-1})^*$.

(ב) *אם* A *שקולה קוניוקנטיבית ל־* B *אז* B *שקולה קוניוקנטיבית ל־* A.

(ג) *אם* A *שקולה קוניוקנטיבית ל־* B *ו־* B *שקולה קוניוקנטיבית ל־* C *אז* A *שקולה קוניוקנטיבית ל־* C.

הוכחה: (א) $PP^{-1} = I_n$, לכן $(P^{-1})^*P^* = I_n^* = I_n$,

(ב) נניח $B = P^*AP$, באשר P הפיכה. אז $A = ((P^*)^{-1}BP^{-1} = (P^{-1})^*BP^{-1}$. לפי (א) B שקולה קוניוקנטיבית ל־A.

(ג) נניח $B = P^*AP$, $C = Q^*BQ$, באשר P, Q הפיכות. אז $C = Q^*P^*APQ = (PQ)^*A(PQ)$, כלומר, C שקולה קוניוקנטיבית ל־A. ∎

משפט 10.3: *מטריצה* $A \in \mathrm{M}_n(F)$ *היא חיובית גמורה אם ורק אם היא שקולה קוניוקטיבית ל־* I_n.

הוכחה: נניח ש־A שקולה קוניונקטיבית ל־I_n. אז יש P הפיכה כך ש־$A = P^*I_nP = P^*P$.

A צמודה לעצמה: $A^* = (P^*P)^* = P^*P = A$.

A חיובית גמורה: יהי $0 \neq w \in F^n$. אז $Pw \neq 0$. לכן

$$w^*Aw = w^*P^*Pw = (Pw)^*(Pw) = \langle Pw, Pw\rangle_{\mathrm{st}} > 0$$

באשר $\langle -, -\rangle_{\mathrm{st}}$ היא המכפלה הפנימית הסטנדרטית על F^n.

להיפך, נניח ש־A חיובית גמורה. נבחר מרחב וקטורי V כלשהו ממימד n מעל F ונבחר בסיס כלשהו $\mathcal{B}'$ שלו. לפי משפט 9.33 יש מכפלה פנימית על V שמטריצת גרם שלה ביחס ל־$\mathcal{B}'$ היא A. לפי מסקנה 9.13 יש ל־V בסיס אורתונורמלי $\mathcal{B}$. לפי דוגמה 9.30 מטריצת גרם של המכפלה הפנימית לפיו היא I_n. תהי P מטריצת המעבר מ־$\mathcal{B}$ ל־$\mathcal{B}'$. לפי משפט 9.34, $A = P^*I_nP$. ∎

סימון 10.4: נסמן ב־$I_{n,s,t}$ את המטריצה האלכסונית מסדר $n \times n$ שבאלכסון הראשי שלה יש s "1"־ים, אחריהם t "(-1)"־ים, ואחריהם $n - s - t$ אפסים.

משפט 10.5: *כל* $A \in \mathrm{M}_n(F)$ *הרמיטית שקולה קוניונקטיבית למטריצה יחידה מהצורה* $I_{n,s,t}$.

הוכחה: אם P מטריצה הפיכה אז גם P^* הפיכה ולכן מכפלה של מטריצות אלמנטריות $P^* = E_k \cdots E_2E_1$. לכן $B = P^*AP$ אם ורק אם

$$B = E_k \cdots E_2E_1AE_1^*E_2^* \cdots E_k^*.$$

זה שקול לכך ש־B מתקבלת מ־A על ידי סדרה של פעולות אלמנטריות $\mathcal{P}_1, \mathcal{P}_2, \dots \mathcal{P}_k$ על **השורות** של A, (לפעולה $\mathcal{P}_i$ מתאימה המטריצה E_i) כאשר, לאחר כל פעולה $\mathcal{P}_i$ מבצעים את הפעולה האלמנטרית "הצמודה המתאימה לה" $\mathcal{P}_i^*$ (לה מתאימה המטריצה $\overline{E_i}^t$) על **העמודות** של המטריצה המתקבלת. אכן, $(\overline{E}A^t)^t = A\overline{E}^t = AE_i^*$.

לא קשה להראות שמכל A סימטרית אפשר להגיע על ידי סדרה של פעולות כאלה למטריצה $I_{n,s,t}$. (התהליך דומה לתהליך הדירוג של גאוס. ראה הדוגמה למטה.)

היחידות (משפט Sylvester): נניח ש־A שקולה קוניונקטיבית ל־$I_{n,s,t}$ וגם ל־$I_{n,s',t'}$, כלומר, יש מטריצות הפיכות P, Q כך ש־$P^*AP = I_{n,s,t}$ ו־$Q^*AQ = I_{n,s',t'}$. אז

$$s + t = \mathrm{rk}(I_{n,s,t'}) = \mathrm{rk}(A) = \mathrm{rk}(I_{n,s',t'}) = s' + t',$$

ולכן די להוכיח $s = s'$.

נניח, למשל, $s > s'$. אז בסדרה $Pe_1, \dots, Pe_s, Qe_{s'+1}, \dots, Qe_n \in F^n$ יש יותר מ־n וקטורים. לכן הם תלויים לינארית, כלומר, יש $a_1, \dots, a_s, b_{s'+1}, \dots, b_n \in F$, לא כולם 0, כך ש־

$$a_1Pe_1 + \cdots + a_sPe_s + b_{s'+1}Qe_{s'+1} + \cdots + b_nQe_n = 0$$

כלומר

$$P(a_1e_1 + \cdots + a_se_s) + Q(b_{s'+1}e_{s'+1} + \cdots + b_ne_n) = 0.$$

נסמן $u = -a_1e_1 - \cdots - a_se_s$, $v = b_{s'+1}e_{s'+1} + \cdots + b_ne_n$. אז $u, v \in F^n$ ו־$Pu = Qv$. נשים לב שלא כל ה־a_i הם אפס (אחרת $u = 0$, ואז $Qv = 0$ ומכאן $v = Q^{-1}Qv = 0$, מה שנותן שכל ה־b_i הם אפס, סתירה). לכן

$$u^*I_{n,s,t}u = |a_1|^2 + \cdots + |a_s|^2 > 0$$
$$v^*I_{n,s',t'}v = |b_{s'+1}|^2 + \cdots + |b_n|^2 \leq 0$$

כלומר,

$$(Pu)^*APu > 0, (Qv)^*AQv \leq 0$$

וזה מהווה סתירה ל־$Pu = Qv$. ∎

נגסח את המשפט הנ"ל במפורש עבור המקרה $F = \mathbb{R}$:

מסקנה 10.6: *כל $A \in \mathrm{M}_n(\mathbb{R})$ סימטרית חופפת למטריצה יחידה מהצורה $I_{n,s,t}$.*

דוגמה 10.7: מהי הצורה הקנונית $I_{2,s,t}$ של $A=\begin{pmatrix}0&-2+2i\\-2-2i&0\end{pmatrix}$?

$$\begin{pmatrix}0&-2+2i\\-2-2i&0\end{pmatrix}\xrightarrow{\mathcal{P}_{2,1}(1)}\begin{pmatrix}-2-2i&-2+2i\\-2-2i&0\end{pmatrix}\xrightarrow{\mathcal{P}_{2,1}(1)^*}\begin{pmatrix}-4&-2+2i\\-2-2i&0\end{pmatrix}\xrightarrow{\mathcal{P}_1(\frac{1}{2})}$$
$$\begin{pmatrix}-2&-1+i\\-2-2i&0\end{pmatrix}\xrightarrow{\mathcal{P}_1(\frac{1}{2})^*}\begin{pmatrix}-1&-1+i\\-1-i&0\end{pmatrix}\xrightarrow{\mathcal{P}_{1,2}(-1-i)}\begin{pmatrix}-1&-1+i\\0&(-1+i)(-1-i)\end{pmatrix}$$
$$\xrightarrow{\mathcal{P}_{1,2}(-1-i)^*}\begin{pmatrix}-1&-1+i+(-1)(-1+i)\\0&(-1+i)(-1-i)\end{pmatrix}=\begin{pmatrix}-1&0\\0&2\end{pmatrix}\xrightarrow{\mathcal{P}_2(\frac{1}{\sqrt{2}})}\begin{pmatrix}-1&0\\0&\frac{2}{\sqrt{2}}\end{pmatrix}\xrightarrow{\mathcal{P}_2(\frac{1}{\sqrt{2}})^*}$$
$$\begin{pmatrix}-1&0\\0&1\end{pmatrix}\xrightarrow{\mathcal{P}_{1,2}}\begin{pmatrix}0&1\\-1&0\end{pmatrix}\xrightarrow{\mathcal{P}_{1,2}^*}\begin{pmatrix}1&0\\0&-1\end{pmatrix}=I_{2,1,1}$$

מהי הצורה הקנונית $I_{2,s,t}$ של $B=\begin{pmatrix}0&i\\-i&0\end{pmatrix}$?

$$\begin{pmatrix}0&i\\-i&0\end{pmatrix}\xrightarrow{\mathcal{P}_1(-i)}\begin{pmatrix}0&1\\-i&0\end{pmatrix}\xrightarrow{\mathcal{P}_1(-i)^*}\begin{pmatrix}0&1\\1&0\end{pmatrix}\xrightarrow{\mathcal{P}_{2,1}(1)}\begin{pmatrix}1&1\\1&0\end{pmatrix}\xrightarrow{\mathcal{P}_{2,1}(1)^*}\begin{pmatrix}2&1\\1&0\end{pmatrix}\xrightarrow{\mathcal{P}_{1,2}(-\frac{1}{2})}$$
$$\begin{pmatrix}2&1\\0&-\frac{1}{2}\end{pmatrix}\xrightarrow{\mathcal{P}_{1,2}(-\frac{1}{2})^*}\begin{pmatrix}2&0\\0&-\frac{1}{2}\end{pmatrix}\xrightarrow{\mathcal{P}_1(\frac{1}{\sqrt{2}})}\xrightarrow{\mathcal{P}_1(\frac{1}{\sqrt{2}})^*}\xrightarrow{\mathcal{P}_2(\sqrt{2})}\xrightarrow{\mathcal{P}_2(\sqrt{2})^*}\begin{pmatrix}1&0\\0&-1\end{pmatrix}=I_{2,1,1}$$

מסקנה 10.8: *שתי מטריצות הרמיטיות $A,B\in \mathrm{M}_n(F)$ שקולות קוניונקטיבית אם ורק אם שתיהן שקולות קוניונקטיבית לאותה מטריצה מהצורה $I_{n,s,t}$.*

הוכחה: נניח ש־B שקולה קוניונקטיבית ל־A. לפי משפט 10.5, A שקולה קוניונקטיבית לאיזה $I_{n,s,t}$. לפי תרגיל 10.2 גם B שקולה קוניונקטיבית ל־$I_{n,s,t}$.

להיפך, נניח ש־A,B שקולות קוניונקטיבית ל־$I_{n,s,t}$. לפי תרגיל 10.2, A שקולה קוניונקטיבית ל־B. ∎

דוגמה 10.9: המטריצות A,B בדוגמה 10.7 שקולות קוניונקטיבית, לפי המסקנה הקודמת, כי שתיהן שקולות קוניונקטיבית ל־$I_{2,1,1}$.

11. לכסון מטריצה סימטרית

גם בפרק זה יהי $F = \mathbb{R}$ או $F = \mathbb{C}$.

הגדרה 11.1: $A \in \mathrm{M}_n(F)$ נקראת **אוניטרית** אם $A^*A = I_n$, כלומר, $A^* = A^{-1}$; אם A ממשית ואוניטרית (כלומר, $A^t = A^{-1}$), אז A נקראת גם **אורתוגונלית**.

המטריצה I_n הנה ואורתוגונלית ובפרט אוניטרית.

למה 11.2: *תהי $P \in \mathrm{M}_n(F)$ ותהיינה $u_1, \ldots, u_n$ עמודותיה. אז P אוניטרית אם ורק אם $u_1, \ldots, u_n$ בסיס אורתונורמלי של F^n (ביחס למכפלה הסטנדרטית).*

הוכחה: השורה ה־i של P^* היא u_i^*. לכן $(P^*P)_{ij} = u_i^* u_j = \langle u_i, u_j \rangle_{\mathrm{st}}$. לכן

$P^*P = I_n \;\Leftrightarrow\; \langle u_i, u_j \rangle_{\mathrm{st}} = (P^*P)_{ij} = \delta_{ij} \;\Leftrightarrow\; u_1, \ldots, u_n$ סדרה אורתונורמלית.

אך לפי משפט 9.9 סדרה אורתונורמלית בת n איברים בלתי תלויה לינארית ולכן, לפי משפט 4.40(3), היא בסיס. לכן $u_1, \ldots, u_n$ סדרה אורתונורמלית אם ורק אם $u_1, \ldots, u_n$ בסיס אורתונורמלי. ∎

דוגמה 11.3: $\begin{pmatrix} 0 & 1 \\ i & 0 \end{pmatrix}$ אוניטרית; $\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$ אורתוגונלית.

מטריצות אוניטריות קשורות למכפלות פנימיות באופן הבא:

למה 11.4: *יהי V מרחב מכפלה פנימית. יהיו $\mathcal{B}, \mathcal{B}'$ שני בסיסים שלו ותהי P מטריצת מעבר מ־$\mathcal{B}$ ל־$\mathcal{B}'$. נניח ש־$\mathcal{B}$ אורתונורמלי. אז $\mathcal{B}'$ אורתונורמלי אם ורק אם P אוניטרית.*

הוכחה: יהיו $u_1', \ldots, u_n'$ האיברים של $\mathcal{B}'$. אז $P = ([u_1']_{\mathcal{B}}, \ldots, [u_n']_{\mathcal{B}})$. לפי הלמה הקודמת P אוניטרית אם ורק אם $\langle [u_i']_{\mathcal{B}}, [u_j']_{\mathcal{B}} \rangle_{\mathrm{st}} = \delta_{ij}$. לפי ההגדרה, $\mathcal{B}'$ אורתונורמלי אם ורק אם $\langle u_i', u_j' \rangle = \delta_{ij}$. אבל לפי מסקנה 9.8, $\langle u_i', u_j' \rangle = \langle [u_i']_{\mathcal{B}}, [u_j']_{\mathcal{B}} \rangle_{\mathrm{st}}$. לכן שני התנאים שקולים. ∎

תרגיל 11.5: *(א) אם P אוניטרית אז $P^{-1} = P^*$ אוניטרית.*

(ב) *אם P_1, P_2 אוניטריות אז P_1P_2 אוניטרית.*

הוכחה: (א) נתון $P^*P = I_n$, לכן $PP^* = I_n$, כלומר $(P^*)^*P^* = I_n$.

(ב) $(P_1P_2)^*(P_1P_2) = P_2^*P_1^*P_1P_2 = P_2^*I_nP_2 = I_n$. ∎

משפט 11.6: *תהי $A \in \mathrm{M}_n(\mathbb{C})$. אז קיימת $P \in \mathrm{M}_n(\mathbb{C})$ אוניטרית כך ש־$P^*AP = P^{-1}AP$ משולשית עליונה.*

הוכחה: באינדוקציה על n.

אם $n = 1$, אז A כבר משולשית עליונה; נקח $P = I_1$ שהנה אוניטרית.

נניח נכונות למטריצות מסדר $(n-1) \times (n-1)$. תהי $A \in \mathrm{M}_n(\mathbb{C})$. לפולינום האופייני שלה $f_A(X)$ יש שורש $\lambda_1 \in \mathbb{C}$, לפי המשפט היסודי של אלגברה, והוא ערך עצמי של A, לפי משפט 8.2. לכן יש לו וקטור עצמי $0 \neq u_1 \in \mathbb{C}^n$. בלי הגבלת הכלליות $||u_1|| = 1$ (ביחס למכפלה הסטנדרטית) אחרת נחליף את u_1 ב־$\frac{1}{||u_i||}u_i$.

נשלים את u_1 לבסיס אורתונורמלי $u_1, u_2, \dots, u_n$ של $\mathbb{C}^n$. לפי למה 11.2, $P_1 = (u_1, \dots, u_n) \in \mathrm{M}_n(\mathbb{C})$ אוניטרית.

מתקיים $Au_1 = \lambda_1 u_1 = \lambda_1 P_1 e_1 = P_1(\lambda_1 e_1)$, לכן

$$\begin{aligned} P_1^* A P_1 &= P_1^{-1} A (u_1, \dots, u_n) = (P_1^{-1} A u_1, \dots, P_1^{-1} A u_n) = \\ &= (P_1^{-1} P_1 \lambda_1 u_1, *, \dots, *) = \begin{pmatrix} \lambda_1 & * \cdots * \\ 0 & \\ \vdots & A_1 \\ 0 & \end{pmatrix} \end{aligned}$$

באשר $A_1 \in \mathrm{M}_{n-1}(\mathbb{C})$.

לפי הנחת האינדוקציה יש $Q \in \mathrm{M}_{n-1}(\mathbb{C})$ אוניטרית כך ש־$O^* A_1 Q$ משולשית עליונה. נגדיר

$$P_2 = \begin{pmatrix} 1 & 0 \cdots 0 \\ 0 & \\ \vdots & Q \\ 0 & \end{pmatrix}. \quad \text{אז} \quad P_2^* = \begin{pmatrix} 1 & 0 \cdots 0 \\ 0 & \\ \vdots & Q^* \\ 0 & \end{pmatrix} \quad \text{ו־} \quad P_2^* P_2 = \begin{pmatrix} 1 & 0 \cdots 0 \\ 0 & \\ \vdots & Q^* Q \\ 0 & \end{pmatrix} = I_n.$$

לכן P_2 אוניטרית. לפי התרגיל גם $P_1 P_2$ אוניטרית.

$$\begin{aligned} (P_1 P_2)^* A (P_1 P_2) &= P_2^* (P_1^* A P_1) P_2 = \\ &= \begin{pmatrix} 1 & 0 \cdots 0 \\ 0 & \\ \vdots & Q^* \\ 0 & \end{pmatrix} \begin{pmatrix} \lambda_1 & * \cdots * \\ 0 & \\ \vdots & A_1 \\ 0 & \end{pmatrix} \begin{pmatrix} 1 & 0 \cdots 0 \\ 0 & \\ \vdots & Q \\ 0 & \end{pmatrix} = \begin{pmatrix} \lambda_1 & * \cdots * \\ 0 & \\ \vdots & Q^* A_1 Q \\ 0 & \end{pmatrix} \end{aligned}$$

כאשר המטריצה הימנית משולשית עליונה. ■

מסקנה 11.7: *אם A הרמיטית, אז יש $P \in \mathrm{M}_n(\mathbb{C})$ אוניטרית כך ש־* $P^* A P = P^{-1} A P = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ *ו־ $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. בפרט, $f_A(X) = (X - \lambda_1) \cdots (X - \lambda_n)$, באשר $\lambda_1, \dots, \lambda_n$ ממשיים.*

הוכחה: יש P אוניטרית כך ש־$B = P^* A P$ משולשית עליונה. לפי הערה 9.29(ג), B הרמיטית. לכן היא אלכסונית, עם רכיבים ממשיים באלכסון הראשי, נאמר, $\lambda_1, \dots, \lambda_n$. כיוון ש־$A$ ו־$B = P^{-1} A P$ דומות, יש להן אותו פולינום אופייני. לכן $f_A(X) = f_B(X) = (X - \lambda_1) \cdots (X - \lambda_n)$. ■

משפט 11.8: *תהי A סימטרית. אז יש $P \in \mathrm{M}_n(\mathbb{R})$ אורתוגונלית כך ש־$P^t A P = P^{-1} A P$ אלכסונית.*

הוכחה: לפי המסקנה הקודמת,

(1) $f_A(X) = (X - \lambda_1) \cdots (X - \lambda_n)$, באשר $\lambda_1, \dots, \lambda_n \in \mathbb{R}$.

נחזור על ההוכחה של המשפט הקודם (משפט השילוש, אשר לפי המסקנה לעיל נותן כבר הליכסון במקרה של מטריצה צמודה לעצמה), רק שנשתמש ב־(1) במקום במשפט היסודי של אלגברה.

(קצת) ביתר פירוט: בשלב האינדוקציה נמצא תחילה P_1 אורתוגונלית כך שמתקיים

$$P_1^tAP_1 = \begin{pmatrix} \lambda_1 & * \cdots * \\ 0 & \\ \vdots & A_1 \\ 0 & \end{pmatrix}.$$

אך $P_1^tAP_1 = P_1^tA^t(P_1^t)^t = (P_1^tAP_1)^t$, לכן

$$P_1^tAP_1 = \begin{pmatrix} \lambda_1 & 0 \cdots 0 \\ 0 & \\ \vdots & A_1 \\ 0 & \end{pmatrix}.$$

היות ואגף שמאל מטריצה סימטרית ($(P_1^tAP_1)^t = P_1^tAP_1$), בהכרח גם A_1 סימטרית. לכן אפשר להפעיל עליה הנחת האינדוקציה ולהמשיך כמו בהוכחת משפט 11.6. ∎

כיצד ניתן למצוא, עבור A סימטרית נתונה, P אורתוגונלית כך ש־P^tAP אלכסונית?

ההוכחה לעיל למעשה נותנת את הבניה. אך נשתמש בשיטה פשוטה יותר, שמשתמשת בלמה הבאה:

למה 11.9: *תהי A הרמיטית. יהיו $\lambda_1 \neq \lambda_2$ שני ערכים עצמיים של A, ויהי v_i וקטור עצמי של A השייך ל־ λ_i, עבור $i = 1, 2$. אז $v_1 \perp v_2$, כלומר, $\langle v_1, v_2\rangle_{\text{st}} = 0$.*

הוכחה: לפי דוגמה 9.25(ב), $\langle v_1, v_2\rangle_{\text{st}} = v_1^*v_2$. לפי מסקנה 11.7 $\lambda_1, \lambda_2 \in \mathbb{R}$. לכן

$$\lambda_1v_1^*v_2 = \overline{\lambda_1}v_1^*v_2 = (\lambda_1v_1)^*v_2 = (Av_1)^*v_2 = v_1^*A^*v_2 = v_1^*Av_2 = v_1^*\lambda_2v_2 = \lambda_2v_1^*v_2$$

וכיוון ש־$\lambda_1 \neq \lambda_2$, יוצא ש־$v_1^*v_2 = 0$. ∎

בניה 11.10: *תהי $A \in \mathrm{M}_n(\mathbb{R})$ סימטרית. מצא P אורתוגונלית כך ש־ P^tAP אלכסונית.*

פתרון: מצא את הפולינום האופייני של A. לפי מסקנה 11.7, $f_A(X) = (X-\lambda_1)^{n_1}\cdots(X-\lambda_r)^{n_r}$, באשר $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ שונים זה מזה. כמו כן, כיוון ש־A דומה לאלכסונית, הריבוי האלגברי n_i של λ_i הוא גם הריבוי הגיאומטרי $\dim V_{\lambda_i}$ שלו, לכל i.

לכל $1 \leq i \leq r$ נמצא בסיס $\mathcal{B}_i$ של המרחב העצמי V_{λ_i}; מספר איבריו הוא n_i. בלי הגבלת כלליות $\mathcal{B}_i$ הוא אורתונורמלי, לפי מסקנה 9.13. אם נצרף את איברי $\mathcal{B}_1, \ldots, \mathcal{B}_r$ לסדרה אחת $\mathcal{B}$, אז מספר איבריה $n_1 + \cdots + n_r = \deg f_A(X) = n$. כולם בעלי נורמה 1. כל שנים מביניהם ניצבים זה לזה, לפי הבניה (אם הם שייכים לאותו ערך עצמי) או לפי הלמה הקודמת (אם הם שייכים לערכים עצמיים שונים). לכן $\mathcal{B}$ סדרה אורתונורמלית בת n איברים ב־$\mathbb{R}^n$. לפי משפט 9.9 היא בלתי תלויה לינארית, ולכן בסיס של $\mathbb{R}^n$.

תהי P המטריצה שאיברי $\mathcal{B}$ עמודותיה. לפי למה 11.2, P אורתוגונלית. לפי משפט 7.9, $P^tAP = P^{-1}AP$ אלכסונית. ∎

דוגמה 11.11: דוגמה ללכסון אורתוגונלי. עבור $A=\begin{pmatrix}2&1&1\\1&2&1\\1&1&2\end{pmatrix}$ מצא P אורתוגונלית כך ש־P^tAP אלכסונית.

פתרון: נשים לב ש־A סימטרית. תחילה נחשב את הפולינום האופייני:

$$f_A(X)=\begin{vmatrix}X-2&-1&-1\\-1&X-2&-1\\-1&-1&X-2\end{vmatrix}\overset{\mathcal{P}_{21}(-1),\mathcal{P}_{32}(-1)}{=}\begin{vmatrix}X-1&-X+1&0\\0&X-1&-X+1\\-1&-1&X-2\end{vmatrix}=$$

$$\overset{\mathcal{P}_{12}(-1)^t}{=}\begin{vmatrix}X-1&0&0\\0&X-1&-X+1\\-1&-2&X-2\end{vmatrix}=(X-1)^2\begin{vmatrix}1&-1\\-2&X-2\end{vmatrix}=(X-1)^2(X-4)$$

לכן הערכים העצמיים של A הם $1,4$ ואנו יודעים שהריבויים הגיאומטריים שלהם הם $2,1$, בהתאמה.

נחשב את המרחב העצמי $V_1=\{v\in\mathbb{R}^3|\ Av=1v\}=\{v\in\mathbb{R}^3|\ (A-I)v=0\}$.

$V_1=\left\{\begin{pmatrix}x\\y\\z\end{pmatrix}|\ x+y+z=0\right\}=\mathrm{Sp}(v_1,v_2)$, באשר $v_1=\begin{pmatrix}-1\\1\\0\end{pmatrix}, v_2=\begin{pmatrix}-1\\0\\1\end{pmatrix}$. נבצע

תהליך גרם־שמידט על v_1,v_2 כדי לקבל בסיס אורתוגונלי של V_1:

$$u_1=v_1,\qquad u_2=v_2-\frac{\langle u_1,v_2\rangle}{||u_1||^2}u_1=\begin{pmatrix}-1\\0\\1\end{pmatrix}-\frac{1}{2}\begin{pmatrix}-1\\1\\0\end{pmatrix}=\frac{1}{2}\begin{pmatrix}-1\\-1\\2\end{pmatrix}$$

וננרמל אותו כדי לקבל בסיס אורתונורמלי של V_1:

$$u_1'=\begin{pmatrix}-\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\\0\end{pmatrix}, u_2'=\begin{pmatrix}-\frac{1}{\sqrt{6}}\\-\frac{1}{\sqrt{6}}\\\frac{2}{\sqrt{6}}\end{pmatrix}.$$

כעת נחשב את המרחב העצמי $V_4=\{v\in\mathbb{R}^3|\ Av=4v\}=\{v\in\mathbb{R}^3|\ (A-4I)v=0\}$. לשם חישובו נדרג את $A-4I$:

$$A-4I=\begin{pmatrix}-2&1&1\\1&-2&1\\1&1&-2\end{pmatrix}\longrightarrow\begin{pmatrix}0&0&0\\0&-3&3\\1&1&-2\end{pmatrix}\longrightarrow\begin{pmatrix}1&0&-1\\0&1&-1\\0&0&0\end{pmatrix},$$

לכן בסיס של V_4 בן איבר אחד, למשל, $\begin{pmatrix}1\\1\\1\end{pmatrix}$. ננרמל אותו כדי לקבל בסיס אורתונורמלי, בן האיבר היחיד, $u_3'=\begin{pmatrix}\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{3}}\end{pmatrix}$. כעת, אם נגדיר

$$P=(u_1',u_2,',u_3')=\begin{pmatrix}-\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{3}}\\0&\frac{2}{\sqrt{6}}&\frac{1}{\sqrt{3}}\end{pmatrix},$$

אז P אורתוגונלית ו־ $P^tAP = P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$. ∎

תרגיל 11.12: *תהי* $P \in \mathrm{M}_n(F)$ *אוניטרית, ויהי* $\lambda \in F$ *ערך עצמי שלה. אז* $|\lambda| = 1$. *בפרט, אם* $\lambda \in \mathbb{R}$ *אז* $\lambda = \pm 1$.

הוכחה: יש $0 \neq v \in F^n$ כך ש־$Pv = \lambda v$. אז

$$|\lambda|^2 \cdot ||v||^2 = \overline{\lambda}\lambda v^* v = (\lambda v)^* \lambda v = (Pv)^*(Pv) = v^* P^* P v = v^* I_n v = v^* v = ||v||^2,$$

וכיוון ש־$||v||^2 \neq 0$, נחלק בו ונקבל $|\lambda|^2 = 1$. ∎

משפט 11.13: *יהי* V *מרחב מכפלה פנימית מעל* $\mathbb{R}$ *ויהי* $\mathcal{B}$ *בסיס אורתונורמלי שלו. תהי* $T\colon V \to V$ *העתקה לינארית. התנאים הבאים שקולים זה לזה:*

(א) T *שומרת מכפלה פנימית, כלומר* $\langle T(u), T(v)\rangle = \langle u, v\rangle$ *לכל* $u, v \in V$.

(ב) T *שומרת אורך, כלומר* $||T(u)|| = ||u||$ *לכל* $u \in V$.

(ג) T *שומרת מרחקים, כלומר* $||T(u) - T(v)|| = ||u - v||$ *לכל* $u, v \in V$.

(ד) *המטריצה* $[T]_{\mathcal{B}}^{\mathcal{B}}$ *אורתוגונלית.*

הוכחה: (א)⇐(ב): הצב $v = u$; לפי (א), $||T(u)||^2 = ||u||^2$.

(ב) ⇐(ג): $||T(u) - T(v)|| = ||T(u - v)|| = ||u - v||$.

(ג)⇐(ב): הצב $v = 0$.

(ב)⇐(א): $||u - v||^2 = \langle u - v, u - v\rangle = ||u||^2 - 2\langle u, v\rangle + ||v||^2$,

$$||T(u - v)||^2 = ||T(u) - T(v)||^2 = \langle T(u) - T(v), T(u) - T(v)\rangle =$$
$$= ||T(u)||^2 - 2\langle T(u), T(v)\rangle + ||T(v)||^2.$$

ומכאן לפי (ב), אם נשווה את כל הארכים המתאימים, נקבל $\langle u, v\rangle = \langle T(u), T(v)\rangle$.

בהמשך נשתמש במסקנה 9.8, לפיה $\langle u, v\rangle = \langle [u]_{\mathcal{B}}, [v]_{\mathcal{B}}\rangle_{\mathrm{st}} = [u]_{\mathcal{B}}^t [v]_{\mathcal{B}}$ לכל $u, v \in V$.

(א)⇐(ד) יהיו $u_1, u_2, \ldots, u_n$ איברי $\mathcal{B}$. אז

$$\delta_{ij} = \langle u_i, u_j\rangle = \langle T(u_i), T(u_j)\rangle = ([T(u_i)]_{\mathcal{B}})^t [T(u_j)]_{\mathcal{B}} = ([T]_{\mathcal{B}}^{\mathcal{B}} [u_i]_{\mathcal{B}})^t [T]_{\mathcal{B}}^{\mathcal{B}} [u_j]_{\mathcal{B}} =$$
$$= ([u_i]_{\mathcal{B}})^t ([T]_{\mathcal{B}}^{\mathcal{B}})^t [T]_{\mathcal{B}}^{\mathcal{B}} [u_j]_{\mathcal{B}} = e_i^t \big(([T]_{\mathcal{B}}^{\mathcal{B}})^t [T]_{\mathcal{B}}^{\mathcal{B}}\big) e_j = \big(([T]_{\mathcal{B}}^{\mathcal{B}})^t [T]_{\mathcal{B}}^{\mathcal{B}}\big)_{ij}.$$

לכן $([T]_{\mathcal{B}}^{\mathcal{B}})^t [T]_{\mathcal{B}}^{\mathcal{B}} = I_n$.

(ד) ⇐(א)

$$\langle T(u), T(v)\rangle = ([T(u)]_{\mathcal{B}})^t [T(v)]_{\mathcal{B}} = ([T]_{\mathcal{B}}^{\mathcal{B}} [u]_{\mathcal{B}})^t [T]_{\mathcal{B}}^{\mathcal{B}} [v]_{\mathcal{B}} = [u]_{\mathcal{B}}^t ([T]_{\mathcal{B}}^{\mathcal{B}})^t [T]_{\mathcal{B}}^{\mathcal{B}} [v]_{\mathcal{B}} =$$
$$= [u]_{\mathcal{B}}^t I_n [v]_{\mathcal{B}} = [u]_{\mathcal{B}}^t [v]_{\mathcal{B}} = \langle u, v\rangle$$ ∎

הגדרה 11.14: העתקה שמקיימת את התנאים השקולים של המשפט נקראת **אורתוגונלית**.

דוגמה 11.15: *מהן ההעתקות האורתוגונליות* $\mathbb{R}^2 \to \mathbb{R}^2$?

פתרון: תהי T אורתוגונלית ותהי $A \in \mathrm{M}_2(\mathbb{R})$ המטריצה שלה לפי הבסיס הסטנדרטי. אז A אורתוגונלית. העמודה הראשונה (וגם השניה) של A היא וקטור מאורך 1, כלומר, מהצורה $\binom{a}{b}$, באשר $a^2 + b^2 = 1$. אפשר לכתוב אותו בצורה $\binom{\cos\theta}{\sin\theta}$ עבור איזה θ. כדי שהעמודה השניה תהיה ניצבת לראשונה וגם מאורך 1, היא חייבת להיות $\binom{-\sin\theta}{\cos\theta}$ או $\binom{\sin\theta}{-\cos\theta}$. לכן A היא

$$A_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \qquad \text{או} \qquad B_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$

יש הבדל עקרוני בין שתי הצורות: $|A_\theta| = 1,\ |B_\theta| = -1$. יתר על כן:

(א) בדוגמה 6.19(ב) ראינו ש־A_θ מייצגת, לפי הבסיס הסטנדרטי, את הסיבוב סביב הראשית בזוית θ. בפרט אין לה וקטורים עצמיים שונים מאפס, ולכן אין לה ערכים עצמיים (ממשיים; כמובן, לפי המשפט היסודי של אלגברה יש לפולינום האופייני שלה שרשים מרוכבים), אלא אם θ כפולה של π, ואז $A_0 = I_2, A_\pi = -I_2$. קל לראות זאת גם אלגברית: הפולינום האופייני של A_θ הוא

$$X^2 - 2(\cos\theta)X + 1 = (X-\cos\theta)^2 + \sin^2\theta = (X - \cos\theta + i\sin\theta)(X - \cos\theta - i\sin\theta).$$

(ב) הפולינום האופייני של B_θ הוא $X^2 - 1 = (X-1)(X+1)$. קל לחשב ש־$\binom{\cos\frac{\theta}{2}}{\sin\frac{\theta}{2}}$ הוא וקטור עצמי של B_θ השייך לערך עצמי 1. היות ו־B_θ סימטרית, וקטור עצמי שלה השייך לערך עצמי -1 חייב להיות ניצב לראשון. לכן, אם T ההעתקה ש־B_θ המטריצה שלה לפי הבסיס הסטנדרטית אז T היא השיקוף ביחס לציר דרך הראשית בעל שיפוע של זוית $\frac{\theta}{2}$. (ראה גם דוגמה 6.39.) ∎

דוגמה 11.15: *מהן ההעתקות האורתוגונליות* $T\colon \mathbb{R}^3 \to \mathbb{R}^3$?

פתרון: תהי T אורתוגונלית ותהי $A \in \mathrm{M}_3(\mathbb{R})$ המטריצה שלה לפי הבסיס הסטנדרטי. אז A אורתוגונלית. הפולינום האופייני של A כזאת ממעלה 3, לכן (כידוע) יש לו שורש ממשי λ. לפי תרגיל 11.12, $\lambda = \pm 1$. אם גם 1 וגם -1 ערכים עצמיים של A, נבחר λ בעל אותו סימן כמו $|A|$. (למעשה, $|A| = \pm 1$ כי $|A|^2 = |A^t|\cdot|A| = |I_3| = 1$.)

נבחר וקטור עצמי u_1 ל־λ. בלי הגבלת הכלליות $||u_1|| = 1$. נשלים את u_1 לבסיס אורתונורמלי u_1, u_2, u_3 של $\mathbb{R}^3$ ותהי $P_1 = (u_1, u_2, u_3) \in \mathrm{M}_3(\mathbb{R})$. אז P_1 אורתוגונלית וכמו במשפט השילוש האוניטרי,

$$P_1^{-1}AP_1 = P_1^tAP_1 = \begin{array}{c|cc} & 1 & 2 \\ \hline 1 & \lambda & v \\ 2 & 0 & A_1 \end{array} \tag{1}$$

עבור איזה $v \in \mathrm{M}_{1\times 2}(\mathbb{R})$ ו־$A_1 \in M_2(\mathbb{R})$. כמו כן P_1 היא מטריצת המעבר מהבסיס הסטנדרטי לבסיס האורתונורמלי $\mathcal{B} = (u_1, u_2, u_3)$, לכן (1) מייצגת את T לפי $\mathcal{B}$.

אנו טוענים ש־$v = 0$ ו־A_1 אורתוגונלית. אכן, $P_1^t A P_1$ אורתוגונלית:

$$(P_1^t A P_1)^t (P_1^t A P_1) = P_1^t \big(A^t (P_1 P_1^t) A\big) P_1 = I_n.$$

לכן $\begin{pmatrix} \lambda & v \\ 0 & A_1 \end{pmatrix}$ אורתוגונלית. מכאן

$$\begin{pmatrix} 1 & 0 \\ 0 & I_2 \end{pmatrix} = I_3 = \begin{pmatrix} \lambda & 0 \\ v^t & A_1^t \end{pmatrix} \begin{pmatrix} \lambda & v \\ 0 & A_1 \end{pmatrix} = \begin{pmatrix} \lambda^2 & \lambda v \\ \lambda v^t & v^t v + A_1^t A_1 \end{pmatrix}.$$

מכאן $\lambda = \pm 1$, $v = 0$, $A_1^t A = I_2$. לכן A_1 אורתוגונלית.

למטריצות דומות אותה דטרמיננטה ואותו פולינום אופייני. לכן $|A| = \lambda \cdot |A_1|$ ו־$f_A(X) = (X - \lambda) f_{A_1}(X)$. אילו היה $|A_1| = -1$, אז A_1 היה שיקוף ולכן גם 1 וגם -1 ערכים עצמיים שלו. לכן הם גם ערכים עצמיים של A. לפי בחירת λ מתקיים $|A_1| > 0$, סתירה. לכן A_1 מייצגת סיבוב.

לכן $[T]_{\mathcal{B}}^{\mathcal{B}}$ מהצורה $A' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$ או $A' \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. הימנית מייצגת סיבוב סביב u_1 (וקטור עצמי של T) בזוית θ. השמאלית היא אותו הסיבוב שאחריו בא השיקוף ביחס למישור $\mathrm{Sp}(u_2, u_3)$ הניצב ל־u_1. ∎

המשפט הבא שייך יותר לפרק על מכפלה פנימית, אך משתמש בתוצאות של הפרק הנוכחי:

משפט 9.38: *תהי $A \in \mathrm{M}_n(F)$ הרמיטית. אז A חיובית גמורה אם ורק אם כל הערכים העצמיים של A חיוביים.*

הוכחה: (לפי מסקנה 11.7 הערכים העצמיים של מטריצה הרמיטית הם ממשיים.)

יהי $\lambda \in F$ ערך עצמי של A. אז יש וקטור עצמי $0 \neq v \in F^n$ כך ש־$Av = \lambda v$. מכאן

$$0 < v^* A v = v^* \lambda v = \lambda v^* v = \lambda \langle v, v \rangle_{\mathrm{st}}.$$

כיוון ש־$\langle v, v \rangle_{\mathrm{st}}$ ממשי חיובי, גם λ ממשי חיובי.

להיפך, נניח שכל הערכים העצמיים של A ממששיים חיוביים. לפי מסקנה 11.7 יש $P \in \mathrm{M}_n(F)$ אוניטרית כך ש־$P^{-1} A P = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ אלכסונית. לפי למה 11.2 עמודותיה $u_1, \ldots, u_n$ של P הן בסיס אורתונורמלי של F^n. לפי משפט 7.9 מתקיים $A u_i = \lambda_i u_i$ לכל $1 \le i \le n$.

יהי $0 \neq w \in F^n$. אז יש $a_1, \ldots, a_n \in F$ כך ש־$w = a_1 u_1 + \cdots + a_n u_n$. כיוון ש־$w \neq 0$, לא כן a_i הם 0. כעת, לפי למה 9.7(ב)

$$w^* A w = (\sum_{i=1}^n a_i u_i)^* A (\sum_{i=1}^n a_i u_i) = (\sum_{i=1}^n a_i u_i)^* (\sum_{i=1}^n \lambda_i a_i u_i) = \langle \sum_{i=1}^n a_i u_i), \sum_{i=1}^n \lambda_i a_i u_i \rangle_{\mathrm{st}} =$$

$$\sum_{i=1}^n \overline{a_i} a_i \lambda_i = \sum_{i=1}^n |a_i|^2 \lambda_i > 0$$ ∎

TEL AVIV UNIVERSITY אוניברסיטת תל-אביב
THE IBY AND ALADAR FLEISCHMAN FACULTY OF ENGINEERING הפקולטה להנדסה ע"ש איבי ואלדר פליישמן

מבחן באלגברה לינארית להנדסה

0509.1824

מועד א'
תשס"ד, סמסטר א'

י"ח בשבט, תשס"ד
10 בפברואר 2004

המורים: אשר בן ארצי, דוד גינזבורג, דן הרן, שמואל רוסט, יבגני שוסטין.
משך המבחן: שלוש שעות.
אין להשתמש בכל חומר עזר. אין להשתמש במחשבונים.

ענה על חמש מתוך שבע השאלות הבאות. (רק חמש התשובות הראשונות תבדקנה!)

שאלה 1: נתונה מטריצה $A = \begin{pmatrix} 3 & -1 & 2 \\ 9 & -3 & 6 \\ 12 & -4 & 8 \end{pmatrix}$ מעל $\mathbb{R}$. בכל אחד מהמקרים הבאים קבע האם A דומה ל-B:

(א) $B = \begin{pmatrix} 2 & 8 & -2 \\ 2 & 8 & -2 \\ 2 & 8 & -2 \end{pmatrix}$

(ב) $B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 8 & 1 \\ 0 & 0 & 0 \end{pmatrix}$

(ג) $B = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

(ד) $B = \begin{pmatrix} 8 & 1 & 2 \\ 1 & 8 & 3 \\ 2 & 3 & 8 \end{pmatrix}$

(ה) $B = \begin{pmatrix} 3 & 3 & 3 \\ 9 & 9 & 9 \\ 12 & 12 & 12 \end{pmatrix}$

שאלה 2: יהי V מרחב וקטורי ויהיו $v_1, v_2, v_3, v_4 \in V$ וקטורים בלתי תלויים לינארית. קבע אם $v_1 + v_2, v_2 + v_3, v_3 + v_4, v_4 + v_1$ תלויים לינארית. הוכח טענתך.

שאלה 3: תהי $A = \begin{pmatrix} 1 & 2 & p & -2 & 1 \\ 1 & p & 2 & -p & 1 \\ 1 & 2 & -p & -2 & 1 \end{pmatrix}$. נגדיר $T_A\colon \mathbb{R}^5 \to \mathbb{R}^3$ על ידי $T_A(v) = Av$. מצא בסיס ל-$\mathrm{Im}\ T$.

שאלה 4: יהי V מרחב הפולינומים מעל שדה F ממעלה≤ 3. יהי $\mathcal{B} = \{x^3, x^2, x, 1\}$ בסיס של V (בסדר הזה). נגדיר העתקה לינארית $T\colon V \to V$ על ידי

$$\big(T(P)\big)(x) = P(0)x^3 + P'(x),$$

באשר $P'(x)$ היא הנגזרת של $P(x)$. מצא את המטריצה $[T]_{\mathcal{B}} = [T]_{\mathcal{B}}^{\mathcal{B}}$ של T לפי $\mathcal{B}$.

שאלה 5: יהי V מרחב המטריצות מסדר 2×2 מעל $\mathbb{R}$. נגדיר עליו מכפלה פנימית על ידי

$$\langle A, B\rangle = \operatorname{tr}\left(B^t \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} A\right).$$

מצא בסיס אורתונורמלי של V לפי מכפלה זו. (כאן $\operatorname{tr}(A)$ הוא הסכום של אברי האלכסון הראשי של A.)

שאלה 6: תהי A מטריצה מסדר 2×2 מעל $\mathbb{R}$ אשר מקיימת

$$A\begin{pmatrix} 2 \\ 5 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \end{pmatrix}, \qquad A\begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} -2 \\ -6 \end{pmatrix}.$$

מצא את A, את הערכים העצמיים של A^{-1} ואת הוקטורים העצמיים של A^{-1}.

שאלה 7: תהי A מטריצה מסדר 6×6 מעל $\mathbb{C}$. נתון שהפולינום האופייני שלה הוא $x^2(x-2)(x+1)^3$.

(א) חשב את הדטרמיננטה $|A^2 + A|$.

(ב) רשום את צורות ז'ורדן האפשריות של A.

בהצלחה!