BEN GURION UNIVERSITY OF THE NEGEV
BE'ER SHEVA, ISRAEL

אוניברסיטת בן גוריון בנגב
באר שבע

מבוא לאלגברה ליניארית

אמנון יקותיאלי
המחלקה למתמטיקה
אוניברסיטת בן גוריון
amyekut@math.bgu.ac.il

חוברת זו מיועדת לקורס "מבוא לאלגברה ליניארית" לתלמידי הנדסה באוניברסיטת בן גוריון. זה קורס של סמסטר אחד (בדרך כלל 26 הרצאות של שעתיים כל אחת). החומר כולל פתרון מערכות משוואות ליניאריות, מרחבים וקטוריים, חשבון מטריצות, טרנפורמציות ליניאריות, ליכסון אופרטורים ומרחבי מכפלה פנימית. החומר מוגש בצורה מדוייקת מבחינה מתמטית, ורוב המשפטים מוכחים. יוצא דופן הוא הטיפול בדטרמיננטות, שם דילגתי על ההוכחות (מקוצר זמן) והסתפקתי בהפנייה לספרים אחרים. בכל פרק משולבות דוגמאות רבות.

החוברת מבוססת על רשימות בכתב יד שהכנתי בעת שלימדתי את הקורס בשנים 1999 – 2002. בהכנת הרשימות נעזרתי ברשימותיו של עידו אפרת, ושאלתי מהן הרבה מן החומר התאורטי, הסימונים והדוגמאות המספריות. ברצוני להודות לעידו אפרת על הסכמתו לשימוש ברשימותיו.

במהדורה הרביעית נוסף הפרק על מרחבי מכפלה פנימית, וכן נעשו שיפורים רבים בטקסט.

הערות, תיקונים והצעות לשיפור יתקבלו בברכה.

ברצוני להודות לאנדריי מלניקוב על העבודה המסורה בהכנת גירסת ה־ LaTeX הראשונה. תודה לרמה פורת ואמנון בסר על הסיוע הטכני ב־ LaTeX בעברית.

תוכן העניינים

מהדורה רביעית 15.10.2006

מבוא

נקודת המוצא של הקורס היא פתרון **מערכת של משוואות ליניאריות**, כלומר משוואות ממעלה ראשונה במספר נעלמים. כולנו יודעים לפתור מערכת כדוגמת

$$2x + 4y = 0$$

$$5x + 12y = 8 \; .$$

אולם נשאלות השאלות הבאות:

- כיצד פותרים מערכת שבה הרבה נעלמים?
- כיצד פותרים מערכת שבה הרבה משוואות?
- האם בכלל קיימים פתרונות למערכת המשוואות? אם כן אז כמה?

במהלך הדיון בכיתה במערכות משוואות ליניאריות יופיעו באופן טבעי המושגים **מרחב וקטורי** ו־ **מטריצה**. אלו נושאים חשובים בפני עצמם. המרחב הוקטורי הוא מושג מרכזי בגיאומטריה, בפיסיקה (למשל המרחב הוקטורי $\mathbb{R}^3$, שהוא המרחב של המכניקה הניוטונית), בתורת הפונקציות (מרחבי הילברט) ובתורת ההצפנות (מרחבים וקטוריים מעל שדות סופיים). חשבון מטריצות משמש בין היתר לטרנספורם פורייה מהיר (FFT), טכניקה נפוצה בעיבוד אותות. ליכסון מטריצות חשוב מאוד להרבה מטרות, למשל לפתרון משוואות דיפרנציאליות או לניתוח תהליכי מרקוב.

ספרים נוספים לעיון ותירגול:

1. Hoffman and Kunze, "Linear Algebra", Prentice-Hall 1971.
2. ברמן וקון, "אלגברה ליניארית", הוצאת בק 1999.
3. ליפשוץ, "אלגברה ליניארית", סדרת שאום 1991; מהדורה עברית 1993.
4. עמיצור, "אלגברה א'", הוצאת אקדמון 1970.
5. גולן, "יסודות האלגברה הליניארית", הוצאת דקל 2000.

א. שדות

במשוואות שלנו יופיעו מספרים מסוגים שונים, ונתחיל את הקורס בנושא זה. ראשית הנה רשימה של כמה קבוצות מספרים וסימוליהן המקובלים.

1) המספרים הטבעיים $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$.
2) המספרים השלמים $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$.
3) המספרים הרציונליים $\mathbb{Q} = \{\frac{m}{n} \mid m, n \in \mathbb{Z};\ n \neq 0\}$.
4) המספרים הממשיים $\mathbb{R}$.

נזכיר שכל מספר ממשי אי־שלילי a ניתן לייצוג ע״י פיתוח עשרוני

$$a = d_n \cdots d_1 d_0 . d_{-1} d_{-2} \cdots$$

כאשר n מספר טבעי ו־

$$d_n, \ldots d_1, d_0, d_{-1}, d_{-2}, \ldots \in \{0, 1, \ldots, 9\}$$

הן הספרות העשרוניות. הפיתוח העשרוני הוא יחיד, מלבד האפסים שניתן לכתוב מצד שמאל, ומלבד המקרה של 9 במחזור, כמו למשל

$$0.999\ldots = 1.000\ldots\ .$$

המשמעות של הפיתוח העשרוני היא ש־

$$a = \lim_{j \to \infty} a_j$$

כאשר a_j הוא המספר הרציונלי

$$a_j := d_n \cdots d_1 d_0 . d_{-1} d_{-2} \cdots d_{-j} = d_n \cdot 10^n + \cdots + d_1 \cdot 10 + d_0 + d_{-1} \cdot 10^{-1} + \cdots + d_{-j} \cdot 10^{-j}\ .$$

כעת נגדיר קבוצה חדשה של מספרים.

הגדרה 1. **מספר מרוכב** הינו זוג (a, b) של מספרים ממשיים. נסמן ב־ $\mathbb{C}$ את קבוצת המספרים המרוכבים, כלומר

$$\mathbb{C} = \mathbb{R} \times \mathbb{R} = \mathbb{R}^2 = \{\text{זוגות סדורים של מספרים ממשיים}\}\ .$$

נגדיר פעולות חיבור וכפל על הקבוצה $\mathbb{C}$.

חיבור:

$$(a_1, b_1) + (a_2, b_2) := (a_1 + a_2, b_1 + b_2)\ .$$

כפל:

$$(a_1, b_1) \cdot (a_2, b_2) := (a_1 a_2 - b_1 b_2, a_1 b_2 + a_2 b_1)\ .$$

נשים לב כי שני מספרים מרוכבים $z_1 = (a_1, b_1)$ ו־ $z_2 = (a_2, b_2)$ הם שווים אם״ם $a_1 = a_2$ ו־ $b_1 = b_2$.

יהיו a_1 ו־ a_2 מספרים ממשיים. נתבונן במספרים המרוכבים $(a_1, 0)$ ו־ $(a_2, 0)$. רואים כי

$$(a_1, 0) + (a_2, 0) = (a_1 + a_2, 0)$$

ו־

$$(a_1, 0) \cdot (a_2, 0) = (a_1 a_2, 0)\ .$$

זאת אומרת שההתאמה $a \mapsto (a, 0)$ שומרת על פעולות החיבור והכפל. כמו כן לכל מספר מרוכב z מתקיים $z + (0, 0) = z$ ו־ $z \cdot (1, 0) = z$. משום כך **נזהה** את המספר הממשי a עם המספר המרוכב $(a, 0)$, ובצורה זו נקבל הכלה $\mathbb{R} \subset \mathbb{C}$. תהליך זיהוי זה דומה לאופן שבו מזהים את המספר השלם n

עם המספר הרציונלי $\frac{n}{1}$ ואשר באמצעותו מקבלים את ההכלה $\mathbb{Z}\subset\mathbb{Q}$. בתור תת־קבוצה של המספרים המרוכבים, המספרים הממשיים הם בדיוק המספרים $z=(a,b)$ כך ש־ $b=0$.

הגדרה 2. המספר המרוכב $(0,1)$ יסומן באות $\mathbf{i}$.

התכונה המיוחדת של המספר $\mathbf{i}$ היא

$$\mathbf{i}^2=\mathbf{i}\cdot\mathbf{i}=(0,1)\cdot(0,1)=(0\cdot 0-1\cdot 1,0\cdot 1+1\cdot 0)=(-1,0)=-1\ .$$

כפי שנהוג עבור מספרים ממשיים גם כאן המוסכמה היא שפעולת הכפל קודמת לחיבור, ולכן ניתן להשמיט סוגריים לפעמים; למשל $z_1+(z_2\cdot z_3)=:z_1+z_2\cdot z_3$.
חישוב קצר עבור $a,b\in\mathbb{R}$ מראה ש־

$$\begin{aligned}a+b\cdot\mathbf{i}&=(a,0)+(b,0)\cdot(0,1)=(a,0)+(b\cdot 0-0\cdot 1,b\cdot 1+0\cdot 0)\\&=(a,0)+(0,b)=(a,b)\end{aligned}$$

לכן נהוג לכתוב מספר מרוכב $z=(a,b)\in\mathbb{C}$ בצורה הבאה: $z=a+b\cdot\mathbf{i}$. בסימון זה פעולות החשבון הן

$$\begin{aligned}(a_1+b_1\cdot\mathbf{i})+(a_2+b_2\cdot\mathbf{i})&=(a_1+a_2)+(b_1+b_2)\cdot\mathbf{i}\\(a_1+b_1\cdot\mathbf{i})\cdot(a_2+b_2\cdot\mathbf{i})&=(a_1a_2-b_1b_2)+(a_1b_2+a_2b_1)\cdot\mathbf{i}\end{aligned}$$

עבור $a_1,b_1,a_2,b_2\in\mathbb{R}$.

הגדרה 3. יהי $z=(a,b)=a+b\cdot\mathbf{i}$ מספר מרוכב. **החלק הממשי** של z הוא $\mathrm{Re}(z):=a$ ו־ **החלק המדומה** של z הוא $\mathrm{Im}(z):=b$.

לכל מספר מרוכב z מתקיים השוויון

$$z=\mathrm{Re}(z)+\mathrm{Im}(z)\cdot\mathbf{i}\ .$$

נשים לב כי $\mathrm{Im}(z)$ הוא מספר ממשי.
מאחר שכל מספר מרוכב הוא זוג מספרים ממשיים, הרי כל מספר מרוכב מייצג נקודה במישור. לכן משתמשים בביטוי "המישור המרוכב", וזה התיאור הגיאומטרי של $\mathbb{C}$.

טענה 1. לכל שלושה מספרים מרוכבים z_1,z_2,z_3 מתקיימות התכונות הבאות.

1. קומוטטיביות החיבור: $z_1+z_2=z_2+z_1$.
2. אסוציאטיביות החיבור: $z_1+(z_2+z_3)=(z_1+z_2)+z_3$.
3. תכונת ה־ 0: $z_1+0=z_1$.
4. קיום הפכי חיבורי: קיים $w\in\mathbb{C}$ כך ש־ $z_1+w=0$.
5. קומוטטיביות הכפל: $z_1\cdot z_2=z_2\cdot z_1$.
6. אסוציאטיביות הכפל: $z_1\cdot(z_2\cdot z_3)=(z_1\cdot z_2)\cdot z_3$.
7. תכונת ה־ 1: $z_1\cdot 1=z_1$.
8. דיסטריבוטיביות: $z_1\cdot(z_2+z_3)=(z_1\cdot z_2)+(z_1\cdot z_3)$.
9. קיום הפכי כפלי: אם $z_1\neq 0$ אז קיים w כך ש־ $z_1\cdot w=1$.

הוכחה חלקית. נרשום $z_n=a_n+b_n\cdot\mathbf{i}$ עבור $n=1,2,3$.

תכונה 1.

$$\begin{aligned} z_1+z_2 &= (a_1+b_1\cdot\mathbf{i})+(a_2+b_2\cdot\mathbf{i}) \\ &= (a_1+a_2)+(b_1+b_2)\cdot\mathbf{i} && \text{הגדרת החיבור ב־ } \mathbb{C} \\ &= (a_2+a_1)+(b_2+b_1)\cdot\mathbf{i} && \text{קומוטטיביות החיבור ב־ } \mathbb{R} \\ &= (a_2+b_2\cdot\mathbf{i})+(a_1+b_1\cdot\mathbf{i}) && \text{הגדרת החיבור ב־ } \mathbb{C} \\ &= z_2+z_1 \end{aligned}$$

תכונה 3.

$$\begin{aligned} z_1+0 &= (a_1+b_1\cdot\mathbf{i})+(0+0\cdot\mathbf{i}) \\ &= (a_1+0)+(b_1+0)\cdot\mathbf{i} && \text{הגדרת החיבור ב־ } \mathbb{C} \\ &= a_1+b_1\cdot\mathbf{i} && \text{תכונת ה־ 0 ב־ } \mathbb{R} \\ &= z_1 \end{aligned}$$

תכונה 4. ניקח $w:=(-a_1)+(-b_1)\cdot\mathbf{i}$. אז

$$\begin{aligned} z_1+w &= (a_1+b_1\cdot\mathbf{i})+((-a_1)+(-b_1)\cdot\mathbf{i}) \\ &= (a_1+(-a_1))+(b_1+(-b_1))\cdot\mathbf{i} && \text{הגדרת החיבור ב־ } \mathbb{C} \\ &= 0+0\cdot\mathbf{i} && \text{תכונת ה־ 0 ב־ } \mathbb{R} \\ &= 0 \end{aligned}$$

התכונות 7,6,5,2 ו־ 8 מוכחות באופן דומה.

תכונה 9. לצורך פשטות נרשום $z_1 := z = (a,b)$. מאחר ש־ $z \neq 0$ הרי $a^2+b^2>0$ ובפרט $a^2+b^2\neq 0$. נגדיר

$$w := \frac{a}{a^2+b^2}+\frac{-b}{a^2+b^2}\cdot\mathbf{i}\in\mathbb{C}$$

אז

$$\begin{aligned} z\cdot w &= (a+b\cdot\mathbf{i})\cdot\left(\frac{a}{a^2+b^2}+\frac{-b}{a^2+b^2}\cdot\mathbf{i}\right) \\ &= \frac{a^2+b^2}{a^2+b^2}+\frac{-ab+ba}{a^2+b^2}\cdot\mathbf{i}=1 \end{aligned}$$

מש״ל.

נשים לב כי המספר w בתכונה 9 הוא יחיד, שהרי אם גם המספר v מקיים $z_1\cdot v=1$ אז

$w=1\cdot w=(z_1\cdot v)\cdot w=v\cdot(z_1\cdot w)=v\cdot 1=v$.

בדומה מראים כי המספר w בתכונה 4 הוא יחיד. זה מאפשר את ההגדרה הבאה.

הגדרה 4.

א. יהי z מספר מרוכב. **ההפכי החיבורי** של z הוא המספר w כך ש־ $z+w=0$, והוא יסומן ע״י $-z$.

ב. יהי z מספר מרוכב שונה מ־ 0. **ההפכי הכפלי** של z הוא המספר w כך ש־ $z\cdot w=1$, והוא יסומן ע״י $\frac{1}{z}$ או z^{-1}.

דוגמה 1.

א. ההפכי הכפלי של $2+\mathbf{i}$ הוא $\frac{2}{5}+\frac{-1}{5}\cdot\mathbf{i}$.

ב. ההפכי הכפלי של $\mathbf{i}$ הוא $-\mathbf{i}$.

כנהוג לעתים נשמיט את סימן הכפל, ונרשום z_1z_2 במקום $z_1 \cdot z_2$. בשל תכונות האסוציאטיביות מותר במקרים מסוימים להשמיט סוגריים, למשל

$$z_1z_2z_3 := (z_1 \cdot z_2) \cdot z_3 = z_1 \cdot (z_2 \cdot z_3) \, .$$

עוד סימונים נוחים הם $z_1 - z_2 := z_1 + (-z_2)$ ו־ $\frac{z_1}{z_2} := z_1 \cdot z_2^{-1}$.

בהנתן מספר ממשי אי־שלילי a נסמן ב־ $\sqrt{a} = a^{1/2}$ את השורש הריבועי האי־שלילי של a.

הגדרה 5.

א. **הצמוד** של המספר המרוכב $z = a + b\mathbf{i}$ הוא המספר המרוכב $\bar{z} := a - b\mathbf{i}$.

ב. **הערך המוחלט** של $z = a + b\mathbf{i}$ הוא המספר הממשי $|z| := \sqrt{a^2 + b^2}$.

נשים לב כי $a^2 + b^2 \geq 0$, ולכן $|z|$ מוגדר ו־ $|z| \geq 0$.

טענה 2. יהי z מספר מרוכב.

1. $|z| = 0$ אס"ם $z = 0$.
2. $|z| = \sqrt{z\bar{z}}$.
3. אם $z \neq 0$ הרי $\frac{1}{z} = \frac{\bar{z}}{|z|^2}$.

הוכחה. 1. $z \neq 0$ אס"ם $a^2 + b^2 > 0$, אס"ם $\sqrt{a^2 + b^2} > 0$.

2. נרשום $z = a + b\mathbf{i}$, ואז

$$z\bar{z} = (a + b\mathbf{i})(a - b\mathbf{i}) = (a^2 + b^2) + (-ab + ba)\mathbf{i} = a^2 + b^2 \, .$$

עתה נוציא שורש ונקבל $|z| = \sqrt{z\bar{z}}$.

3. יהי $w := \frac{\bar{z}}{|z|^2}$. לפי חלק 1 ידוע כי $z\bar{z} = |z|^2 > 0$, ולכן

$$zw = z \cdot \frac{\bar{z}}{z\bar{z}} = z\bar{z} \cdot \frac{1}{z\bar{z}} = 1 \, .$$

טענה 3. הנה כמה תכונות של הצמוד.

1. $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$.
2. $\overline{z_1z_2} = \bar{z}_1 \cdot \bar{z}_2$.
3. $\bar{\bar{z}} = z$.
4. $z = \bar{z}$ אס"ם $z \in \mathbb{R}$.
5. אם $z = b\mathbf{i}$ כאשר $b \in \mathbb{R}$, אז $\bar{z} = -z$.

הוכחה חלקית. נוכיח את תכונה 2. נרשום $z_n = a_n + b_n\mathbf{i}$ עבור $n = 1, 2$. מקבלים

$$\begin{aligned}\overline{z_1z_2} &= \overline{(a_1 + b_1\mathbf{i}) \cdot (a_2 + b_2\mathbf{i})} \\ &= \overline{(a_1a_2 - b_1b_2) + (a_1b_2 + b_1a_2) \cdot \mathbf{i}} \\ &= (a_1a_2 - b_1b_2) - (a_1b_2 + b_1a_2) \cdot \mathbf{i}\end{aligned}$$

ו־

$$\begin{aligned}\bar{z}_1 \cdot \bar{z}_2 &= (a_1 - b_1\mathbf{i}) \cdot (a_2 - b_2\mathbf{i}) \\ &= (a_1a_2 - b_1b_2) + (-a_1b_2 - b_1a_2) \cdot \mathbf{i} \\ &= (a_1a_2 - b_1b_2) - (a_1b_2 + b_1a_2) \cdot \mathbf{i}\end{aligned}$$

מש"ל.

לשם מה יש לנו צורך במספרים מרוכבים? נתבונן במשוואה

$$x^2+2=0 .$$

אין לה פתרון ב־ $\mathbb{R}$. אולם אם נעבור למשוואה השקולה $x^2=-2$ רואים שיש פתרונות מרוכבים $x:=\sqrt{2}\cdot\mathbf{i}$ ו־ $x:=-\sqrt{2}\cdot\mathbf{i}$. בדומה קל לראות שלכל משוואה ריבועית

$$x^2+a_1x+a_0=0$$

עם מקדמים ממשיים יש פתרונות מרוכבים

$$x:=\frac{-a_1\pm\sqrt{a_1^2-4a_0}}{2} .$$

למעשה קיים משפט כללי (אשר לא נוכיח בקורס שלנו):

משפט 1 (המשפט היסודי של האלגברה). בהנתן פולינום

$$f(x)=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$

ממעלה $n\geq 1$ עם מקדמים מרוכבים $a_0,\ldots,a_{n-1}$ קיים למשוואה $f(x)=0$ פתרון מרוכב.

כדי לתת מושג על ההוכחה של המשפט נדון במקרה פרטי. נניח כי n מספר איזוגי וכל המקדמים a_i הם ממשיים. מקבלים הפונקציה רציפה $f:\mathbb{R}\to\mathbb{R}$, והגבולות $\lim_{x\to\infty}f(x)=\infty$ ו־ $\lim_{x\to-\infty}f(x)=-\infty$ קיימים. משפט ערך הביניים אומר שישנה נקודה $b\in\mathbb{R}$ כך ש־ $f(b)=0$.

אנו נכליל את המושג "מספר". במקום מספרים נעבוד עם **סקלרים**, שהם איברים **בשדה**. הפעולות חיבור, חיסור, כפל וחילוק תהיינה מוגדרות עבור סקלרים. היתרון במושג הכללי של שדה הוא שהתוצאות שנוכיח (למשל לגבי פתרון משוואות ליניאריות) ינוסחו עבור שדה כלשהו; וכך בבת אחת נקבל תוצאות התקפות הן עבור משוואות עם מקדמים מרוכבים, הן עבור משוואות עם מקדמים רציונליים, והן עבור כל שדה אחר שאנו עשויים להתקל בו.

הגדרה 6. שדה הינו מערכת $(F,+,\cdot,0,1)$ שבה F קבוצה, $+$ ו־ $\cdot$ הן פעולות דו־מקומיות על F, ו־ 0 ו־ 1 הם שני איברים ב־ F. התכונות הבאות צריכות להתקיים לכל $a,b,c\in F$.

1. קומוטטיביות החיבור: $a+b=b+a$.
2. אסוציאטיביות החיבור: $a+(b+c)=(a+b)+c$.
3. תכונת ה־ 0: $a+0=a$.
4. קיום הפכי חיבורי: קיים d ב־ F כך ש־ $a+d=0$.
5. קומוטטיביות הכפל: $a\cdot b=b\cdot a$.
6. אסוציאטיביות הכפל: $a\cdot(b\cdot c)=(a\cdot b)\cdot c$.
7. תכונת ה־ 1: $a\cdot 1=a$.
8. דיסטריבוטיביות: $a\cdot(b+c)=a\cdot b+a\cdot c$.
9. קיום הפכי כפלי: אם $a\neq 0$ אז קיים d ב־ F כך ש־ $a\cdot d=1$.
10. $0\neq 1$.

גם בשדה כללי נשתמש במוסכמות הרגילות לגבי השמטת סוגריים וסימן הכפל. לרוב נאמר בקיצור "יהי F שדה" במקום "יהי $(F,+,\cdot,0,1)$ שדה".

טענה 4. יהי F שדה.

1. ההפכי החיבורי בתכונה 4 הוא יחיד.
2. ההפכי הכפלי בתכונה 9 הוא יחיד.

הוכחה. א. יהי $a \in F$ ונניח כי b_1 ו־ b_2 מקיימים $a + b_1 = 0 = a + b_2$. אז

תכונת ה־ 0	$b_2 = b_2 + 0$
נתון	$= b_2 + (a + b_1)$
אסוציאטיביות החיבור	$= (b_2 + a) + b_1$
קומוטטיביות החיבור	$= (a + b_2) + b_1$
נתון	$= 0 + b_1$
קומוטטיביות החיבור	$= b_1 + 0$
תכונת ה־ 0	$= b_1$

כלומר $b_1 = b_2$.

ב. נניח $a \neq 0$ ו־ $ab_1 = 1 = ab_2$. בדומה למה שעשינו למעלה, אבל תוך שימוש בתכונת ה־ 1 במקום בתכונת ה־ 0, מקבלים ש־ $b_1 = b_2$.

דוגמה 2. הקבוצות $\mathbb{Q}$, $\mathbb{R}$ ו־ $\mathbb{C}$ עם פעולות החשבון הרגילות הן שדות.

דוגמה 3. המערכת $(\mathbb{N}, +, \cdot, 0, 1)$ איננה שדה. כדי להראות זאת די למצוא דוגמה נגדית לאחת התכונות. ניקח את המספר הטבעי 3. לא קיים ל־ 3 הפכי חיבורי. לכן תכונה 4 איננה מתקיימת.

דוגמה 4. המערכת $(\mathbb{Z}, +, \cdot, 0, 1)$ איננה שדה. דוגמה נגדית לתכונה 9: לא קיים הפכי כפלי למספר השלם 3.

דוגמה 5. ישנם גם שדות סופיים. לכל מספר ראשוני p ומספר שלם חיובי n ישנו שדה (יחיד) $\mathbb{F}_{p^n}$ שבו p^n איברים. עבור $n = 1$ המערכת $\mathbb{F}_p$ קלה לתיאור. הקבוצה $\mathbb{F}_p$ היא קבוצת הסימנים

$$\{[0], [1], [2], \ldots, [p-1]\}.$$

פעולות החשבון מוגדרות כך. יהיו i ו־ j שני מספרים מבין $0, 1, \ldots, p-1$. החיבור מוגדר ע״י

$$[i] + [j] := [k]$$

כאשר k היא השארית של $i + j$ אחרי חלוקה ב־ p. הכפל מוגדר ע״י

$$[i] \cdot [j] := [l]$$

כאשר l היא השארית של $i \cdot j$ אחרי חלוקה ב־ p. מסתבר שהמערכת $(\mathbb{F}_p, +, \cdot, [0], [1])$ היא שדה. במקרה $p = 3$ טבלאות החיבור והכפל נראות כך.

$+$	$[0]$	$[1]$	$[2]$
$[0]$	$[0]$	$[1]$	$[2]$
$[1]$	$[1]$	$[2]$	$[0]$
$[2]$	$[2]$	$[0]$	$[1]$

$\cdot$	$[0]$	$[1]$	$[2]$
$[0]$	$[0]$	$[0]$	$[0]$
$[1]$	$[0]$	$[1]$	$[2]$
$[2]$	$[0]$	$[2]$	$[1]$

אנו רואים שב־ $\mathbb{F}_3$ מתקיים $[2] \cdot [2] = [1]$, ולכן $[2]^{-1} = [2]$. בדומה $[2] + [1] = [0]$, ולכן $[1] = -[2]$.

טענה 5. יהי F שדה ו־ a, b איברים.

א. $a \cdot 0 = 0$.

ב. $a \cdot (-1) = -a$.

ג. אם $ab = 0$ ו־ $a \neq 0$ אז $b = 0$.

הוכחה. א. $0 = 0 + 0$, לכן בעזרת הדיסטריבוטיביות נקבל

$$a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0\ .$$

תכונת ההפכי החיבורי ותכונת האפס נותנות לנו

$$0 = -(a \cdot 0) + a \cdot 0 = -(a \cdot 0) + a \cdot 0 + a \cdot 0 = 0 + a \cdot 0 = a \cdot 0\ .$$

ב. תכונת האחד והדיסטריבוטיביות נותנות

$$a + a \cdot (-1) = a \cdot 1 + a \cdot (-1) = a \cdot (1 + (-1))\ .$$

על פי חלק א׳ ידוע ש־

$$a \cdot (1 + (-1)) = a \cdot 0 = 0\ .$$

ג. נתון כי $a \neq 0$, ולכן קיים a^{-1}. בעזרת חלק א׳ מקבלים

$$0 = a^{-1} \cdot 0 = a^{-1} \cdot (a \cdot b) = (a^{-1} \cdot a) \cdot b = 1 \cdot b = b\ .$$

ב. משוואות ליניאריות

בפרק זה אנו נעסוק במשוואות ליניאריות מעל שדה כלשהו F. בדוגמאות מספריות בדרך כלל השדה יהיה $\mathbb{Q}$, $\mathbb{R}$ או $\mathbb{C}$.

דוגמה 1.

$$\begin{cases} 2x_1 - x_2 + \frac{3}{2}x_3 = 9 \\ x_1 + 3x_2 + 4x_3 = 5 \end{cases}$$

הינה מערכת של שתי משוואות ליניאריות בשלושה נעלמים מעל השדה $\mathbb{Q}$.

הגדרה 1. מערכת של m **משוואות ליניאריות ב־** n **משתנים** מעל השדה F היא מערכת משוואות מהצורה הבאה

$$(*) \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \quad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

כאן $a_{11}, \ldots, a_{mn}$ הינם סקלרים ב־ F שיקראו **המקדמים** של המערכת; $x_1, \ldots, x_n$ הם **המשתנים** (או **הנעלמים**) ו־ $b_1, \ldots, b_m$ הינם סקלרים ב־ F שיקראו **הקבועים**.

בדוגמה 1 המקדמים היו $a_{11} = 2$, $a_{12} = -1$ וכו'; והקבועים היו $b_1 = 9$ ו־ $b_2 = 5$.

הגדרה 2. פתרון של המערכת $(*)$ הינו סדרת סקלרים $(c_1, \ldots, c_n)$ ב־ F כך שבהצבת $x_1 := c_1$, ..., $x_n := c_n$ במערכת המשוואות מתקיימים כל השוויונות:

$$\begin{cases} a_{11}c_1 + a_{12}c_2 + \cdots + a_{1n}c_n = b_1 \\ a_{21}c_1 + a_{22}c_2 + \cdots + a_{2n}c_n = b_2 \\ \quad\vdots \\ a_{m1}c_1 + a_{m2}c_2 + \cdots + a_{mn}c_n = b_m \end{cases}$$

בדוגמה 1 אחד הפתרונות הוא

$$(c_1, c_2, c_3) = (\tfrac{32}{7}, \tfrac{1}{7}, 0) .$$

נשאלות השאלות הבאות:

- האם תמיד קיים פתרון?
- האם הפתרון יחיד?
- כיצד נמצא את הפתרונות?
- מה מבנה קבוצת הפתרונות?

נסתכל בדוגמה נוספת. כאן השדה הוא $F := \mathbb{R}$.

$$(\sharp) \begin{cases} 2x_1 - x_2 + x_3 = 0 \\ x_1 + 3x_2 + 4x_3 = 0 \end{cases}$$

במצב בו הקבועים כולם 0 קוראים למערכת המשוואות **מערכת הומוגנית**.

ניתן לגשת לפתרון המערכת הזו בצורה נאיבית. תחילה נחסר כפולה מתאימה של משוואה אחת מהשניה

$$\begin{array}{rrrl} & 2x_1 & -x_2 & +x_3 = 0 \\ -2(& x_1 & +3x_2 & +4x_3 = 0\ ) \\ \hline & & -7x_2 & -7x_3 = 0 \end{array}$$

וע״י פישוט ונקבל את המשוואה $x_2 = -x_3$. נציב זאת באחת המשוואות המקוריות ונקבל $x_1 = -x_3$. כעת רואים שהמערכת ($\sharp$) שקולה למערכת המשוואות

$$(\sharp\sharp) \begin{cases} x_2 = -x_3 \\ x_1 = -x_3 \end{cases}$$

לכן כל פתרון הוא מהצורה $(-c, -c, c)$ כאשר c סקלר כלשהו ב־ $\mathbb{R}$. קבוצת הפתרונות היא

$$\{(-c, -c, c) \mid c \in \mathbb{R}\} \ .$$

נסכם את מה שעשינו: באמצעות פעולות פשוטות על המשוואות קיבלנו מערכת משוואות חדשה, שקולה למערכת המשוואות המקורית, אשר אותה ניתן לפתור בהתבוננות. המטרה בשיעורים הקרובים היא ללמוד שיטה לפתרון משוואות אשר מבוססת על הרעיון הזה.

תחילה נכניס סימנים יעילים יותר לכתיבת מערכת המשוואות. בהגדרה הבאה ״מטריצה״ היא מלה נרדפת ל־ ״טבלה״ ו־ ״וקטור״ היא מלה נרדפת ל־ ״עמודה״.

הגדרה 3. בהנתן מערכת משוואות ($*$) כמו בהגדרה 1 נרשום את המקדמים ב־ **מטריצת המקדמים**:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \ .$$

וקטור הנעלמים הוא $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, ו־ **וקטור הקבועים** הוא $B = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$. מערכת המשוואות ($*$) תסומן בקיצור ע״י

$$AX = B \ .$$

מערכת הומוגנית תסומן ע״י $AX = O$, כאשר O הוא הוקטור $\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$.

בהמשך הקורס כאשר נלמד חשבון מטריצות הסימון $AX = B$ יקבל משמעות נוספת.

בדרך כלל כאשר נרצה להתייחס למטריצה A בגודל $m \times n$ שרכיביה הם $a_{11}, \dots, a_{mn}$ נרשום $A = [a_{ij}]$, כאשר האינדקס i מציין את מספר השורה, והאינדקס j מציין את מספר העמודה. (בהתאם לנוהג לא רושמים פסיק בין האינדקסים של רכיבי המטריצה.) מטריצה תמיד תהיה בגודל חיובי, כלומר $m, n \geq 1$. עבור וקטור B שרכיביו הם $b_1, \dots, b_m$ נרשום $B = [b_i]$.

הגדרה 4. תהי $A=[a_{ij}]$ מטריצה בגודל $m\times n$ עם רכיבים בשדה F. נסמן ב־ $L_1,\dots,L_m$ את שורות המטריצה, כלומר $L_i=[a_{i1}\ \cdots\ a_{in}]$ ו־ $A=\begin{bmatrix}L_1\\ \vdots\\ L_m\end{bmatrix}$. להלן שלושה סוגי פעולות על שורות המטריצה A, הנקראות **פעולות שורה אלמנטריות**.

1. "$cL_i\to L_i$": כופלים את השורה L_i בסקלר c. כאן $c\in F$, $c\neq 0$.
2. "$L_i+cL_j\to L_i$": לוקחים את את השורה L_j, כופלים בסקלר c ומוסיפים את התוצאה לשורה L_i. השורה L_j נותרת ללא שינוי. כאן $c\in F$ כלשהו ו־ $i\neq j$.
3. "$L_i\leftrightarrow L_j$": מחליפים את השורות L_i ו־ L_j. כאן $i\neq j$.

דוגמה 2. ניקח את מערכת המשוואות $(\sharp)$ שבדוגמה 1. נרשום את מטריצת המקדמים A ונפעיל עליה סדרת פעולות שורה אלמנטריות.

$$A=\begin{bmatrix}2&-1&1\\1&3&4\end{bmatrix}\xrightarrow{L_1\leftrightarrow L_2}\begin{bmatrix}1&3&4\\2&-1&1\end{bmatrix}\xrightarrow{L_2-2L_1\to L_2}\begin{bmatrix}1&3&4\\0&-7&-7\end{bmatrix}$$

$$\xrightarrow{\frac{1}{-7}L_2\to L_2}\begin{bmatrix}1&3&4\\0&1&1\end{bmatrix}\xrightarrow{L_1-3L_2\to L_1}\begin{bmatrix}1&0&1\\0&1&1\end{bmatrix}$$

המטריצה האחרונה מייצגת את מערכת המשוואות

$$\begin{cases}x_2+x_3=0\\x_1+x_3=0\end{cases}$$

בהעברת המשתנה x_3 לאגף ימין נקבל את מערכת המשוואות

$$(\sharp\sharp)\ \begin{cases}x_2=-x_3\\x_1=-x_3\end{cases}$$

אשר אותה כזכור אנו יכולים לפתור בהתבוננות.

עבור השיטה הכללית נרצה לדעת כי פעולות שורה אלמנטריות אינן משנות את קבוצת הפתרונות של המערכת. בטענה הבאה A היא מטריצה, e היא פעולת שורה אלמנטרית ו־ $e(A)$ היא המטריצה המתקבלת כתוצאה מהפעלת e על A.

טענה 1. תהי e פעולת שורה אלמנטרית ו־ A מטריצה. אז קיימת פעולת שורה אלמנטרית f כך ש־ $e(f(A))=A$ ו־ $f(e(A))=A$.

הוכחה.

1. אם e היא "$cL_i\to L_i$" ניקח את f להיות "$\frac{1}{c}L_i\to L_i$".
2. אם e היא "$L_i+cL_j\to L_i$" אז ניקח את f להיות "$L_i-cL_j\to L_i$".
3. אם e היא "$L_i\leftrightarrow L_j$" אז $f=e$.

מש"ל.

משפט 1. תהי A מטריצה בגודל $m\times n$ עם ערכים בשדה F. אם המטריצה A' מתקבלת מהמטריצה A ע"י הפעלת סדרה סופית של פעולות שורה אלמנטריות, הרי למערכות המשוואות ההומוגניות $AX=O$ ו־ $A'X=O$ יש בדיוק אותם פתרונות.

הוכחה. תהיינה $e_1,\dots,e_r$ פעולות שורה אלמנטריות כך שהפעלתן בזו אחר זו נותנת

$$A=A_0\xrightarrow{e_1}A_1\xrightarrow{e_2}A_2\xrightarrow{e_3}\cdots\xrightarrow{e_r}A_r=A'\ .$$

מספיק להוכיח כי לכל k בטווח $1,\ldots,r$ למערכות המשוואות $A_{k-1}X=O$ ו־ $A_kX=O$ יש אותם פתרונות. לכן די להוכיח את המקרה $r=1$, כלומר $A'=e(A)$ לאיזו פעולת שורה אלמנטרית e.
נרשום $A=[a_{ij}]$.

שלב א. תחילה נניח כי $(d_1,\ldots,d_n)$ פתרון של המערכת $AX=O$, ונראה כי זה גם פתרון של $A'X=O$. נבחין בין שלושה מקרים אפשריים.

1. הפעולה e היא "$cL_i\to L_i$". משוואה מס' i במערכת $AX=O$ היא

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

נתון ש־ $(d_1,\ldots,d_n)$ פתרון של $AX=O$, כלומר

$$a_{i1}d_1+a_{i2}d_2+\cdots+a_{in}d_n=0\ .$$

ע"י כפל המשוואה ב־ c נקבל

$$ca_{i1}d_1+ca_{i2}d_2+\cdots+ca_{in}d_n=0\ .$$

לכן $(d_1,\ldots,d_n)$ פתרון של המשוואה

$$ca_{i1}x_1+ca_{i2}x_2+\cdots+ca_{in}x_n=0\ ,$$

שהיא משוואה מס' i במערכת $A'X=O$. יתר המשוואות הן ללא שינוי.

2. הפעולה e היא "$L_i+cL_j\to L_i$", כאשר $i\neq j$. נתון כי

$$a_{i1}d_1+\cdots+a_{in}d_n=0$$

ו־

$$a_{j1}d_1+\cdots+a_{jn}d_n=0\ .$$

לכן

$$(a_{i1}+ca_{j1})d_1+\cdots+(a_{in}+ca_{jn})d_n=0\ ,$$

כלומר $(d_1,\ldots,d_n)$ פתרון של המשוואה

$$(a_{i1}+ca_{j1})x_1+\cdots+(a_{in}+ca_{jn})x_n=0\ ,$$

שהי משוואה מס' i במערכת $A'X=O$. יתר המשוואות נותרות ללא שינוי.

3. e היא הפעולה "$L_i\leftrightarrow L_j$" כאשר $i\neq j$. כאן המשוואות במערכת $AX=O$ הן אותן משוואות כמו ב־ $A'X=O$, רק בסדר אחר. לכן הסדרה $(d_1,\ldots,d_n)$ היא פתרון של $A'X=O$.

שלב ב. כעת נניח כי $(d_1,\ldots,d_n)$ פתרון של $A'X=O$. ע"פ טענה 1 ישנה פעולת שורה אלמנטריות f כך ש־ $A=f(A')$. ההוכחה שבשלב א' (תוך חילוף תפקידים בין A ו־ A') מראה כי הסדרה $(d_1,\ldots,d_n)$ היא פתרון גם של המערכת $AX=O$. מש"ל.

הגדרה 5. שתי מטריצות A ו־ A' מאותו גודל עם רכיבים בשדה F תקראנה **שקולות שורה** אם ניתן לעבור מ־ A ל־ A' ע"י סדרה סופית של פעולות שורה אלמנטריות.

$$A=A_0\xrightarrow{e_1}A_1\xrightarrow{e_2}A_2\xrightarrow{e_3}\cdots\xrightarrow{e_r}A_r=A'$$

תהליך הדרוג

נתחיל בשתי דוגמאות שימחישו כיצד פעולות שורה אלמנטריות מפשטות את מערכת המשוואות.

דוגמה 3. ניקח את השדה $F := \mathbb{Q}$ ואת מערכת המשוואות ההומוגנית

$$\begin{cases} 2x_1 - x_2 + 3x_3 + 2x_4 = 0 \\ x_1 + 4x_2 - x_4 = 0 \\ 2x_1 + 6x_2 - x_3 + 5x_4 = 0 \end{cases}$$

מטריצת המקדמים היא

$$A := \begin{bmatrix} 2 & -1 & 3 & 2 \\ 1 & 4 & 0 & -1 \\ 2 & 6 & -1 & 5 \end{bmatrix}.$$

נפעיל על A את הפעולות הבאות.

$$\begin{bmatrix} 2 & -1 & 3 & 2 \\ 1 & 4 & 0 & -1 \\ 2 & 6 & -1 & 5 \end{bmatrix} \xrightarrow{L_1 \leftrightarrow L_2} \begin{bmatrix} 1 & 4 & 0 & -1 \\ 2 & -1 & 3 & 2 \\ 2 & 6 & -1 & 5 \end{bmatrix} \xrightarrow{L_2 - 2L_1 \to L_2} \begin{bmatrix} 1 & 4 & 0 & -1 \\ 0 & -9 & 3 & 4 \\ 2 & 6 & -1 & 5 \end{bmatrix}$$

$$\xrightarrow{L_3 - 2L_1 \to L_3} \begin{bmatrix} 1 & 4 & 0 & -1 \\ 0 & -9 & 3 & 4 \\ 0 & -2 & -1 & 7 \end{bmatrix} \xrightarrow{-\frac{1}{2}L_3 \to L_3} \begin{bmatrix} 1 & 4 & 0 & -1 \\ 0 & -9 & 3 & 4 \\ 0 & 1 & \frac{1}{2} & -\frac{7}{2} \end{bmatrix}$$

$$\xrightarrow{L_2 \leftrightarrow L_3} \begin{bmatrix} 1 & 4 & 0 & -1 \\ 0 & 1 & \frac{1}{2} & -\frac{7}{2} \\ 0 & -9 & 3 & 4 \end{bmatrix} \xrightarrow{L_3 + 9L_2 \to L_3} \begin{bmatrix} 1 & 4 & 0 & -1 \\ 0 & 1 & \frac{1}{2} & -\frac{7}{2} \\ 0 & 0 & \frac{15}{2} & -\frac{55}{2} \end{bmatrix}$$

$$\xrightarrow{\frac{2}{15}L_3 \to L_3} \begin{bmatrix} 1 & 4 & 0 & -1 \\ 0 & 1 & \frac{1}{2} & -\frac{7}{2} \\ 0 & 0 & 1 & -\frac{11}{3} \end{bmatrix} \xrightarrow{L_1 - 4L_2 \to L_1} \begin{bmatrix} 1 & 0 & -2 & 13 \\ 0 & 1 & \frac{1}{2} & -\frac{7}{2} \\ 0 & 0 & 1 & -\frac{11}{3} \end{bmatrix}$$

$$\xrightarrow{L_1 + 2L_3 \to L_1} \begin{bmatrix} 1 & 0 & 0 & \frac{17}{3} \\ 0 & 1 & \frac{1}{2} & -\frac{7}{2} \\ 0 & 0 & 1 & -\frac{11}{3} \end{bmatrix} \xrightarrow{L_2 - \frac{1}{2}L_3 \to L_2} \begin{bmatrix} 1 & 0 & 0 & \frac{17}{3} \\ 0 & 1 & 0 & -\frac{5}{3} \\ 0 & 0 & 1 & -\frac{11}{3} \end{bmatrix}$$

מערכת המשוואות שמתאימה למטריצה האחרונה היא:

$$\begin{cases} x_1 + \frac{17}{3}x_4 = 0 \\ x_2 - \frac{5}{3}x_4 = 0 \\ x_3 - \frac{11}{3}x_4 = 0 \end{cases}$$

אחרי העברת המשתנה x_4 לאגף ימין נקבל מערכת שקולה

$$\begin{cases} x_1 = -\frac{17}{3}x_4 \\ x_2 = \frac{5}{3}x_4 \\ x_3 = \frac{11}{3}x_4 \end{cases}$$

מייד רואים כי אין כל הגבלה על הערכים שמציבים במקום המשתנה "החופשי" x_4; ואחרי שהצבנו $x_4 := c$, המשתנים "התלויים" x_1, x_2 ו־ x_3 חייבים לקבל את הערכים $x_1 := -\frac{17}{3}c$, $x_2 := \frac{5}{3}c$ ו־

$x_3 := \frac{11}{3}c$. לכן קבוצת הפתרונות היא
$$\{(-\tfrac{17}{3}c, \tfrac{5}{3}c, \tfrac{11}{3}c, c) \mid c \in \mathbb{Q}\}.$$

דוגמה 4. כעת ניקח את השדה $F := \mathbb{C}$ ואת מערכת המשוואות
$$\begin{cases} -x_1 + \mathbf{i}x_2 = 0 \\ -\mathbf{i}x_1 + 3x_2 = 0 \\ x_1 + 2x_2 = 0 \end{cases}$$
מטריצת המקדמים היא
$$A := \begin{bmatrix} -1 & \mathbf{i} \\ -\mathbf{i} & 3 \\ 1 & 2 \end{bmatrix}.$$
נפעיל על A סדרה של פעולות שורה אלמנטריות.
$$A = \begin{bmatrix} -1 & \mathbf{i} \\ -\mathbf{i} & 3 \\ 1 & 2 \end{bmatrix} \xrightarrow{L_1 \leftrightarrow L_3} \begin{bmatrix} 1 & 2 \\ -\mathbf{i} & 3 \\ -1 & \mathbf{i} \end{bmatrix} \xrightarrow{L_2 + \mathbf{i}L_1 \to L_2} \begin{bmatrix} 1 & 2 \\ 0 & 3+2\mathbf{i} \\ -1 & \mathbf{i} \end{bmatrix}$$
$$\xrightarrow{L_3 + L_1 \to L_3} \begin{bmatrix} 1 & 2 \\ 0 & 3+2\mathbf{i} \\ 0 & 2+\mathbf{i} \end{bmatrix} \xrightarrow{\frac{1}{3+2\mathbf{i}}L_2 \to L_2} \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 2+\mathbf{i} \end{bmatrix}$$
$$\xrightarrow{L_3 - (2+\mathbf{i})L_2 \to L_3} \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \xrightarrow{L_1 - 2L_2 \to L_1} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = A'$$
מערכת המשוואות שקבלנו $A'X = O$ היא
$$\begin{cases} x_1 = 0 \\ x_2 = 0 \\ 0 = 0 \end{cases}$$
הפתרון היחיד למערכת המשוואות הוא $(0,0)$. פתרון זה נקרא **הפתרון הטריוויאלי**. תמיד קיים הפתרון הטריוויאלי למערכת משוואות הומוגנית.

אחרי שראינו כמה דוגמאות נתחיל בלימוד השיטה לפתרון מערכות של משוואות ליניאריות. כאשר אנו מדברים על מטריצה בגודל $m \times n$ הכוונה היא תמיד ש־ $m, n \geq 1$.

הגדרה 6. מטריצה A בגודל $m \times n$ מעל השדה F תקרא **מטריצה מדורגת** אם מתקיימים ארבעת התנאים הבאים.

א. בכל שורה שאיננה כולה 0 הרכיב הראשון השונה מ־ 0 הוא 1. רכיב זה נקרא **ה־ 1 המוביל** של השורה.

ב. השורות שכולן 0 מופיעות אחרי השורות שאינן כולן 0.

ג. יהי r מספר השורות ב־ A שאינן כולן 0, ונניח שבשורה ה־ i ה־ 1 המוביל הוא במקום k_i. אז $k_1 < k_2 < \cdots < k_r$.

ד. עבור כל i מ־ 1 עד r הרכיב היחיד בעמודה k_i השונה מ־ 0 נמצא בשורה i. (רכיב זה הוא כמובן ה־ 1 המוביל של שורה i.)

הנה תרשים של מטריצה מדורגת. הסימן $*$ מייצג סקלר כלשהו.

$$\begin{array}{r c}
\begin{array}{r} 1\rightarrow \\ 2\rightarrow \\ \\ r\rightarrow \\ r+1\rightarrow \\ \\ m\rightarrow \end{array} &
\begin{bmatrix}
0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * & 0 & * & \cdots & * \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0
\end{bmatrix} \\
& \begin{array}{ccccc} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ 1 & k_1 & k_2 & k_r & n \end{array}
\end{array}$$

ננסה לתת תיאור מילולי של ההגדרות האלו. במטריצה מדורגת בכל מלבן שהפינה העליונה־ימנית שלו היא איזה 1 מוביל, כל יתר הרכיבים הם 0. כמו כן בכל עמודה שיש בה 1 מוביל, כל יתר הרכיבים הם 0.

הערה. במהדורות קודמות של הספר השתמשנו בביטוי "מטריצה מדורגת קאנונית". ביטוי זה הושמט החל במהדורה הרביעית, והוחלף ב־ "מטריצה מדורגת".

דוגמה 5. השדה בדוגמה זו הוא $\mathbb{Q}$.

א. המטריצה $\begin{bmatrix} 1 & 2 & 7 & 0 & 3 & 5 \\ 0 & 0 & 0 & 1 & 4 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ היא מדורגת. כאן $r=2$.

ב. המטריצה $\begin{bmatrix} 1 & 0 & \frac{7}{2} & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ איננה מדורגת (תנאי ד' אינו מתקיים).

ג. המטריצה $\begin{bmatrix} 0 & 1 & 0 & -3 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ איננה מדורגת (תנאי א' אינו מתקיים).

ד. המטריצה $\begin{bmatrix} 1 & 0 & -3 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ איננה מדורגת (תנאי ב' אינו מתקיים).

ה. מטריצת היחידה בגודל $n \times n$, אשר נסמן ע"י $I_{n\times n}$, היא מדורגת.

$$I_{n\times n} := \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

ו. מטריצת האפס בגודל $m \times n$, אשר נסמן ע"י $O_{m\times n}$, היא מדורגת.

$$O_{m\times n} := \begin{bmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{bmatrix}$$

במטריצה זו $r=0$, ואין אף 1 מוביל.

ראינו בדוגמאות 3 ו־4 כי את מערכת המשוואות ההומוגנית $AX = O$ ניתן לפתור באמצעות הבאת מטריצת המקדמים A לצורה מדורגת. זה נכון באופן כללי. תחילה נראה כי כל מטריצה ניתנת לדרוג.

משפט 2 (תהליך הדרוג של גאוס־ז'ורדן). תהי A מטריצה בגודל $m \times n$ מעל השדה F. אז ישנה מטריצה מדורגת A' אשר הינה שקולת שורה ל־ A.

הוכחה. נוכיח את המשפט באינדוקציה על m (מספר השורות במטריצה A). נסמן $A = [a_{ij}]$.

התחלת האינדוקציה: נניח כי $m = 1$. אם $A = O_{1\times n}$ אז היא כבר מדורגת, וניקח $A' := A$. אחרת יהי k_1 המספר המינימלי כך ש־ $a_{k_1 1} \neq 0$. תהי $e := (\frac{1}{a_{k_1 1}} L_1 \to L_1)$. אז המטריצה $A' := e(A)$ היא מדורגת.

שלב האינדוקציה: כאן $m \geq 2$, ומניחים שכל מטריצה B שיש בה $m - 1$ שורות ניתנת לדרוג. כלומר ישנה מטריצה B' מדורגת ושקולת־שורה ל־ B.

אם $A = O_{m\times n}$ אז היא כבר מדורגת, וניקח $A' := A$. אחרת (אם $A \neq O_{m\times n}$) נעשה חמישה צעדים כדי למצוא את A'. בתום כל צעד נקרא למטריצה שנקבל בשם $A' = [a'_{ij}]$, ולשורותיה נקרא $L'_1, \ldots, L'_n$ (כדי לפשט את הסימונים).

צעד 1. יהי k_1 המספר המזערי כך שהעמודה מספר k_1 במטריצה A איננה כולה 0. יהי i המספר המזערי כך ש־ $a_{ik_1} \neq 0$. אם $i = 1$ נגדיר $A' := A$. אם $i > 1$ תהי $e := (L_1 \leftrightarrow L_i)$, נגדיר $A' := e(A)$.

צעד 2. כעת $a'_{1k_1} \neq 0$. נבצע את פעולת השורה $\frac{1}{a'_{1k_1}} L'_1 \to L'_1$. נקבל מטריצה A' שבה השורה הראשונה היא

$$\begin{array}{c} L'_1 = \begin{bmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \end{bmatrix} \\ \uparrow \\ k_1 \end{array}$$

צעד 3. לכל i מ־ 2 עד m נבצע את הפעולה $L'_i - a'_{ik_1} L'_1 \to L'_i$, לאיפוס עמודה מס' k_1 מתחת ל־ 1 המוביל של השורה הראשונה. בסוף שלב זה המטריצה A' נראית כך:

$$A' = \begin{array}{c} \begin{bmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & * & \cdots & * \end{bmatrix} \\ \uparrow \\ k_1 \end{array}$$

צעד 4. תהי B' המטריצה המתקבלת מ־ A' ע"י השמטת השורה הראשונה שלה. זאת אומרת ש־

$$. \ A' = \begin{bmatrix} L'_1 \\ B' \end{bmatrix}$$

כעת B' היא מטריצה בגודל $(m-1) \times n$, ולפי הנחת האינדוקציה ישנה מטריצה B'' מדורגת ושקולת־שורה ל־ B'. מאחר שב־ B' היו רק אפסים בעמודות $1, 2, \ldots, k_1$, הרי זה המצב גם במטריצה B''. נגדיר מטריצה חדשה

$$. \ A'' := \begin{bmatrix} L'_1 \\ B'' \end{bmatrix}$$

היא נראית כך:

$$A'' = \begin{bmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & * & \cdots & * \end{bmatrix}$$
$$\begin{array}{c} \uparrow \\ k_1 \end{array}$$

תהיינה $e_1, \ldots, e_r$ פעולות שורה אלמנטריות המעבירות מהמטריצה B' למטריצה B''. אז אותן פעולות, אך בהוספת 1 לאינדקס השורות, הן פעולות שורה אלמנטריות המעבירות מהמטריצה A' למטריצה A''. לכן A' ו־ A'' הן שקולות שורה. לסיום צעד זה נגדיר $A' := A''$.

צעד 5. נותר לאפס את הערכים בשורה 1 של המטריצה A' הנמצאים מעל 1־ים מובילים. נניח כי במטריצה המדורגת A' יש r שורות שאינן כולן 0, ובשורה L'_i ה־ 1 המוביל הוא במקום k_i. לכל i מ־ 2 עד m נבצע את פעולת השורה $L'_1 - a'_{1k_i} L'_i \to L'_1$. למטריצה שמקבלים נקרא בשם A'. זוהי מטריצה מדורגת ושקולת־שורה ל־ A. משי"ל.

את ההוכחה של המשפט ניתן בקלות לתרגם לאלגוריתם (מתכון) לדרוג מטריצות, כפי שנדגים בהמשך.

הגדרה 8. תהי A' מטריצה מדורגת עם ערכים בשדה F. מספר השורות שאינן כולן 0 ב־ A', אשר מסומן בדרך כלל באות r, נקרא **הדרגה** של A'.

הערה. תהי A מטריצה כלשהי ותהיינה A' ו־ A'' מטריצות מדורגות ושקולות־שורה ל־ A. ניתן להוכיח כי $A' = A''$; אולם אנו לא נזדקק לכך.

פתרון מערכת משוואות הומוגנית

הגדרה 9. תהי A' מטריצה מדורגת בגודל $m \times n$ ומדרגה r. נניח כי בשורה L_i ה־ 1 המוביל הוא במקום k_i. המשתנים $x_{k_1}, x_{k_2}, \ldots, x_{k_r}$ נקראים **המשתנים התלויים** של מערכת המשוואות $A'X = O$. יתר המשתנים נקראים **המשתנים החופשיים**, והם יסומנו $x_{l_1}, x_{l_2}, \ldots, x_{l_{n-r}}$, כאשר $l_1 < l_2 < \cdots < l_{n-r}$.

דוגמה 6. ניקח $F := \mathbb{R}$ ו־

$$A' := \begin{bmatrix} 0 & 1 & -3 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

הדרגה היא $r = 2$. העמודות שבהן יש 1־ים מובילים הן $k_1 := 2$ ו־ $k_2 := 4$. המשתנים התלויים הם x_2 ו־ x_4. המשתנים החופשיים הם x_1, x_3 ו־ x_5. מערכת המשוואות $A'X = O$ היא

$$\begin{cases} x_2 - 3x_3 + \frac{1}{2}x_5 = 0 \\ x_4 + 2x_5 = 0 \end{cases}$$

אחרי העברת המשתנים החופשיים לאגף ימין מקבלים

$$\begin{cases} x_2 = -(-3x_3 + \frac{1}{2}x_5) \\ x_4 = -(2x_5) \end{cases}$$

רואים שקבוצת הפתרונות היא

$$\{(c, 3d - \tfrac{1}{2}e, d, -2e, e) \mid c, d, e \in \mathbb{R}\} .$$

בניסוח המשפט הבא נזדקק למושג תבנית ליניארית הומוגנית ב־ n משתנים עם מקדמים ב־ F. זה ביטוי מהצורה
$$f(x_1,\dots,x_n)=d_1x_1+\cdots+d_nx_n$$
כאשר $d_1,\dots,d_n\in F$. בהנתן סדרת סקלרים $(c_1,\dots,c_n)$ ניתן להציב $x_i:=c_i$ ולקבל ערך
$$f(c_1,\dots,c_n):=d_1c_1+\cdots+d_nc_n\in F\ .$$

משפט 3. תהי A מטריצה בגודל $m\times n$ עם ערכים בשדה F. הפתרונות של מערכת המשוואות $AX=O$ ניתנים לתיאור באופן הבא. תהי A' מטריצה מדורגת שקולת־שורה ל־ A. יהיו $x_{k_1},\dots,x_{k_r}$ המשתנים התלויים במערכת המשוואות $A'X=O$, ויהיו $x_{l_1},\dots,x_{l_{n-r}}$ המשתנים החופשיים. לכל i מ־ 1 עד r ישנה תבניות ליניארית הומוגנית $f_i(x_{l_1},\dots,x_{l_{n-r}})$ עם מקדמים ב־ F. עבור כל סדרה $(c_1,\dots,c_{n-r})$ של סקלרים ב־ F ישנו פתרון יחיד למערכת המשוואות $AX=O$, שבו ההצבות למשתנים הן
$$\begin{cases} x_{l_1}:=c_1\\ \vdots\\ x_{l_{n-r}}:=c_{n-r}\\ x_{k_1}:=f_1(c_1,\dots,c_{n-r})\\ \vdots\\ x_{k_r}:=f_r(c_1,\dots,c_{n-r})\end{cases}$$
אלו הם כל הפתרונות של מערכת המשוואות $AX=O$.

הוכחה. ע״פ משפט 1 למערכות המשוואות $AX=O$ ו־ $A'X=O$ יש בדיוק אותם פתרונות. אנו נמצא את הפתרונות למערכת המשוואות השנייה.

נסמן $A'=[a'_{ij}]$. מערכת המשוואות $A'X=O$ נראית כך. אם $i\le r$ אז משוואה מס׳ i היא
$$a'_{i1}x_1+\cdots+a'_{in}x_n=0\ ,$$
כאשר $a'_{ik_i}=1$ (זהו ה־ 1 המוביל של שורה 1), והמקדמים של יתר המשתנים התלויים הם כולם 0. נעביר את המשתנים החופשיים לאגף ימין ונקבל משוואה שקולה
$$x_{k_i}=\sum_{j=1}^{n-r}-a'_{il_j}x_{l_j}\ .$$
נגדיר תבנית ליניארית הומוגנית
$$f_i(x_{l_1},\dots,x_{l_{n-r}}):=\sum_{j=1}^{n-r}-a'_{il_j}x_{l_j}\ .$$
מאחר שהמשוואות מס׳ $m,\dots,r+1$ הן $0=0$ הרי ניתן להתעלם מהן, ולכן המערכת $A'X=O$ שקולה למערכת המשוואות
$$(*)\ \begin{cases} x_{k_1}=f_1(x_{l_1},\dots,x_{l_{n-r}})\\ \vdots\\ x_{k_r}=f_r(x_{l_1},\dots,x_{l_{n-r}})\end{cases}$$

בהנתן סדרה כלשהי $(c_1,\dots,c_{n-r})$ של סקלרים, ההצבות $x_{l_j} := c_j$ למשתנים החופשיים ו־ $x_{k_i} := f_i(c_1,\dots,c_{n-r})$ למשתנים התלויים מהוות פתרון של מערכת המשוואות $(*)$. ולהפך, ברור שכל פתרון של מערכת המשוואות $(*)$ הוא מהצורה הזו. מש״ל.

פתרון מערכת משוואות לא הומוגנית

נחזור למקרה של מערכת משוואות לא הומוגנית $AX = B$. תחילה נשים לב שלא תמיד יש פתרון למערכת כזו.

דוגמה 7. ניקח את השדה $F := \mathbb{Q}$ ואת מערכת המשוואות

$$\begin{cases} x_1 + x_2 = 9 \\ 2x_1 + 2x_2 = 2 \end{cases}$$

למערכת זו אין אף פתרון!

למערכת המשוואות הלא הומוגנית $AX = B$ מתאימה מערכת משוואות הומוגנית $AX = O$, ויש קשר הדוק בין הפתרונות שלהן, כפי שהמשפט הבא מראה.

משפט 4. תהי $AX = B$ מערכת משוואות לא הומוגנית עם m משוואות ו־ n נעלמים, ונניח כי $(d_1,\dots,d_n)$ פתרון של מערכת משוואות זו. תהי $(c_1,\dots,c_n)$ סדרה כלשהי של סקלרים. אז $(c_1,\dots,c_n)$ פתרון של מערכת המשוואות ההומוגנית $AX = O$ אס״ם $(c_1+d_1,\dots,c_n+d_n)$ פתרון של מערכת המשוואות $AX = B$.

הוכחה. נסתכל על משוואה מס׳ i. נתון כי

$$a_{i1}d_1 + \cdots + a_{in}d_n = b_i \ .$$

אם $(c_1,\dots,c_n)$ פתרון של המערכת ההומוגנית הרי

$$a_{i1}c_1 + \cdots + a_{in}c_n = 0 \ .$$

נחבר את המשוואות ונקבל

$$a_{i1}(c_1+d_1) + \cdots + a_{in}(c_n+d_n) = b_i + 0 = b_i \ ,$$

כלומר $(c_1+d_1,\dots,c_n+d_n)$ פתרון של משוואה מס׳ i במערכת המשוואות $AX = B$.
לכוון ההפוך ההוכחה היא ע״י חיסור המשוואות. מש״ל.

אנו רואים שהעניין העיקרי בפתרון מערכת משוואות לא הומוגנית $AX = B$ הוא למצוא פתרון מסוים שלה. שאר הפתרונות הם סכומים של פתרון מסוים זה והפתרונות של המערכת ההומוגנית המתאימה.

הגדרה 10. תהי $AX = B$ מערכת של m משוואות ב־ n משתנים. **המטריצה המורחבת** של מערכת המשוואות $AX = B$ היא המטריצה בגודל $m \times (n+1)$

$$[A|B] := \left[\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array}\right]$$

הקו האנכי אמור להזכיר לנו שזו מטריצה מורחבת. פעולות השורה האלמנטריות פועלות כמובן גם המטריצה המורחבת, ואם $[A'|B']$ מתקבלת מ־ $[A|B]$ ע״י פעולות שורה הרי יש מערכת משוואות לא הומוגנית מתאימה $A'X = B'$.

משפט 5. אם המטריצות המורחבות $[A|B]$ ו־ $[A'|B']$ הן שקולות־שורה אז למערכות המשוואות $AX = B$ ו־ $A'X = B'$ יש אותם פתרונות.

הוכחה. כמו בהוכחה של משפט 1 די להוכיח את המקרה שבו $[A'|B'] = e([A|B])$ כאשר e פעולת שורה אלמנטרית.

שלב א. תחילה נניח כי $(d_1, \ldots, d_n)$ פתרון של המערכת $AX = B$, ונראה כי זה גם פתרון של $A'X = B'$. נבחין בין שלושה מקרים אפשריים.

1. הפעולה e היא "$cL_i \to L_i$". משוואה מס׳ i במערכת $AX = B$ היא

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

נתון ש־ $(d_1, \ldots, d_n)$ פתרון של $AX = B$, כלומר

$$a_{i1}d_1 + a_{i2}d_2 + \cdots + a_{in}d_n = b_i \ .$$

ע״י כפל המשוואה ב־ c נקבל

$$ca_{i1}d_1 + ca_{i2}d_2 + \cdots + ca_{in}d_n = cb_i \ .$$

לכן $(d_1, \ldots, d_n)$ פתרון של המשוואה

$$ca_{i1}x_1 + ca_{i2}x_2 + \cdots + ca_{in}x_n = cb_i \ ,$$

שהיא משוואה מס׳ i במערכת $A'X = B'$. יתר המשוואות הן ללא שינוי.

וכך הלאה כמו במקרה ההומוגני (משפט 1) עוברים על שני סוגי הפעולות האלמנטריות הנוספים, ואח״כ בשלב ב׳ משתמשים בפעולת השורה f ההפכית ל־ e. משׁ״ל.

דוגמה 8. השדה הוא $F := \mathbb{Q}$ ומערכת המשוואות היא

$$\begin{cases} x_1 - 2x_2 + x_3 = 1 \\ 2x_1 + x_2 + x_3 = 2 \\ 5x_2 - x_3 = 3 \end{cases}$$

נבצע דירוג של $[A|B]$.

$$[A|B] = \left[\begin{array}{ccc|c} 1 & -2 & 1 & 1 \\ 2 & 1 & 1 & 2 \\ 0 & 5 & -1 & 3 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & -2 & 1 & 1 \\ 0 & 5 & -1 & 0 \\ 0 & 5 & -1 & 3 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|c} 1 & -2 & 1 & 1 \\ 0 & 5 & -1 & 0 \\ 0 & 0 & 0 & 3 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & -2 & 1 & 1 \\ 0 & 1 & -\frac{1}{5} & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & \frac{3}{5} & 1 \\ 0 & 1 & -\frac{1}{5} & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & \frac{3}{5} & 0 \\ 0 & 1 & -\frac{1}{5} & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] = [A'|B']$$

מערכת המשוואות החדשה $A'X = B'$ היא

$$\begin{cases} x_1 + \frac{3}{5}x_3 = 0 \\ x_2 - \frac{1}{5}x_3 = 0 \\ 0 = 1 \end{cases}$$

ואין לה אף פתרון.

דוגמה 9. נשנה את המשוואה השלישית במערכת המשוואות בדוגמה הקודמת ל־ $5x_2 - x_3 = 0$ ונקבל:

$$[A|B] = \left[\begin{array}{ccc|c} 1 & -2 & 1 & 1 \\ 2 & 1 & 1 & 2 \\ 0 & 5 & -1 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & -2 & 1 & 1 \\ 0 & 5 & -1 & 0 \\ 0 & 5 & -1 & 0 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|c} 1 & -2 & 1 & 1 \\ 0 & 5 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & -2 & 1 & 1 \\ 0 & 1 & -\frac{1}{5} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & \frac{3}{5} & 1 \\ 0 & 1 & -\frac{1}{5} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] = [A'|B']$$

מערכת המשוואות $A'X = B'$ היא

$$\begin{cases} x_1 + \frac{3}{5}x_3 = 1 \\ x_2 - \frac{1}{5}x_3 = 0 \end{cases}$$

ע״י העברת המשתנה החופשי לאגף ימין נקבל

$$\begin{cases} x_1 = 1 - \frac{3}{5}x_3 \\ x_2 = \frac{1}{5}x_3 \end{cases}$$

קבוצת הפתרונות היא הקבוצה האינסופית

$$.\ \{(1 - \tfrac{3}{5}c,\ \tfrac{1}{5}c,\ c) \mid c \in \mathbb{Q}\}$$

נשים לב כי אם המטריצה $[A'|B']$ מדורגת אז גם A' מדורגת (השמטת עמודות מצד ימין של המטריצה לא מקלקלת תכונה זו).

משפט 6. תהי $[A|B]$ המטריצה המורחבת של מערכת המשוואות $AX = B$ ותהי $[A'|B']$ מטריצה מדורגת שקולת־שורה ל־ $[A|B]$. התנאים הבאים שקולים.

א. הדרגה של $[A'|B']$ שווה לדרגה של A'.
ב. קיים פתרון למערכת המשוואות $AX = B$.

הוכחה. תחילה נוכיח כי אם תנאי א׳ איננו מתקיים אז גם תנאי ב׳ איננו מתקיים. נניח שהדרגה של $[A'|B']$ גדולה מזו של A'. אז בהכרח יש 1 מוביל בעמודה B'. לכן השורה האחרונה השונה מ־ 0 במטריצה $[A'|B']$ היא $[0 \cdots 0|1]$. שורה זו מתאימה למשוואה

$$0 \cdot x_1 + \cdots + 0 \cdot x_n = 1$$

אשר אין לה פתרון. לכן למערכת המשוואות $A'X = B'$ אין פתרון. ע״פ משפט 5 למערכת המשוואות $AX = B$ אין פתרון.

כעת נוכיח שקיום תנאי א׳ גורר קיום תנאי ב׳. אם הדרגות של $[A'|B']$ ו־ A' שוות אז כל ה־ 1 המובילים במטריצה $[A'|B']$ הם בהכרח במטריצה A'. נניח שהדרגה היא r. יהיו $x_{k_1}, \ldots, x_{k_r}$ המשתנים התלויים של המערכת $A'X = O$, ויהיו $x_{l_1}, \ldots, x_{l_{n-r}}$ המשתנים החופשיים. יהיו $b'_1, \ldots, b'_m$ הקבועים המופיעים בעמודה B'. אז מערכת המשוואות $A'X = B'$ היא

$$\begin{cases} x_{k_1} - f_1(x_{l_1}, \ldots, x_{l_{n-r}}) = b'_1 \\ \vdots \\ x_{k_r} - f_r(x_{l_1}, \ldots, x_{l_{n-r}}) = b'_r \end{cases}$$

כאשר $f_i(x_{l_1}, \ldots, x_{l_{n-r}})$ הן התבניות הליניאריות ההומוגניות שהופיעו במשפט 3. (השמטנו את $m-r$ המשוואות האחרונות שכולן $0 = 0$.) בתור פתרון ל־ $A'X = B'$, ולכן גם ל־ $AX = B$, ניתן לבחור את ההצבה

$$\begin{cases} x_{l_1} := 0 \\ \vdots \\ x_{l_{n-r}} := 0 \\ x_{k_1} := b'_1 \\ \vdots \\ x_{k_r} := b'_r \end{cases}$$

מש״ל.

ג. מרחבים וקטוריים

במהלך הפרק F הינו שדה כלשהו (בדוגמאות F יהיה $\mathbb{Q}$, $\mathbb{R}$ או $\mathbb{C}$).

נתחיל באזכור כמה מושגים לגבי קבוצות. תהי V קבוצה ויהי n מספר טבעי. n ־יה של איברים ב־ V היא סדרה $(v_1, \dots, v_n)$ של איברים V. (מקור המילה: זוג, שלישיה, רביעיה, ..., n ־יה.) חשוב לזכור שבסדרה סדר האיברים משמעותי; כלומר שתי n ־יות $(v_1, \dots, v_n)$ ו־ $(w_1, \dots, w_n)$ הן שוות אם"ם $v_i = w_i$ לכל i. בכך הסדרה $(v_1, \dots, v_n)$ נבדלת מהקבוצה $\{v_1, \dots, v_n\}$.

בהנתן קבוצות $V_1, \dots, V_n$ **המכפלה הקרטזית** שלהן היא הקבוצה

$$. V_1 \times \cdots \times V_n := \{(v_1, \dots, v_n) \mid v_i \in V_i\}$$

כאשר $V_1 = \cdots = V_n = V$ כותבים בקיצור

$$. V^n := \underbrace{V \times \cdots \times V}_{n}$$

דוגמה 1. פעולת החיבור בשדה היא פונקציה $F^2 \to F$ המתאימה לזוג הסדור $(a, b) \in F^2$ את האיבר $a + b \in F$. מקובל לכתוב זאת כך: $(a, b) \mapsto a + b$.

הגדרה 1. **מרחב וקטורי** מעל השדה F הינו מערכת $(V, +, \cdot, \vec{0})$ שבה V קבוצה שאיבריה נקראים **וקטורים**; $+$ היא פעולה דו־מקומית

$$V \times V \to V, \quad (v, w) \mapsto v + w$$

הנקראת **חיבור**; $\cdot$ היא פעולה דו־מקומית

$$F \times V \to V, \quad (a, v) \mapsto a \cdot v$$

הנקראת **כפל בסקלר**; ו־ $\vec{0}$ הוא איבר מיוחד ב־ V הנקרא **וקטור האפס**. התכונות הבאות חייבות להתקיים לכל $u, v, w \in V$ ו־ $a, b \in F$.

1. קומוטטיביות החיבור: $u + v = v + u$.
2. אסוציאטיביות החיבור: $(u + v) + w = u + (v + w)$.
3. תכונת האפס: $v + \vec{0} = v$.
4. קיום הפכי חיבורי: קיים $z \in V$ כך ש־ $v + z = \vec{0}$.
5. תכונת האחד: $1 \cdot v = v$.
6. אסוציאטיביות הכפל בסקלר: $(a \cdot b) \cdot v = a \cdot (b \cdot v)$.
7. דיסטריבוטיביות מימין: $a \cdot (v + w) = (a \cdot v) + (a \cdot w)$.
8. דיסטריבוטיביות משמאל: $(a + b) \cdot v = (a \cdot v) + (b \cdot v)$.

דוגמה 2. יהי n מספר שלם חיובי. מרחב ה־ n ־יות הוא הקבוצה

$$F^n = \{(a_1, \dots, a_n) \mid a_i \in F\}$$

עם פעולת החיבור

$$, \ (a_1, \dots, a_n) + (b_1, \dots, b_n) := (a_1 + b_1, \dots, a_n + b_n)$$

עם פעולת הכפל בסקלר

$$a \cdot (b_1, \dots, b_n) := (a \cdot b_1, \dots, a \cdot b_n)$$

ועם איבר האפס

$$. \vec{0} := (0, \dots, 0)$$

נבדוק את קיום כמה מן התכונות. ניקח את תכונה מס׳ 1 (קומוטטיביות החיבור).

הגדרת החיבור	$(a_1,\dots,a_n)+(b_1,\dots,b_n)=(a_1+b_1,\dots,a_n+b_n)$
קומוטטיביות החיבור ב־ F	$=(b_1+a_1,\dots,b_n+a_n)$
הגדרת החיבור	$=(b_1,\dots,b_n)+(a_1,\dots,a_n)$

נבדוק את קיום תכונה 3 (תכונת האפס).

הגדרת החיבור	$(a_1,\dots,a_n)+(0,\dots,0)=(a_1+0,\dots,a_n+0)$
תכונת האפס ב־ F	$=(a_1,\dots,a_n)$

בצורה דומה רואים שכל יתר התכונות של מרחב וקטורי מתקיימות עבור המערכת $(F^n,+,\cdot,\vec{0})$, בשל קיום התכונות המקבילות בשדה F.

לבסוף נציין שגם עבור $n=0$ המרחב F^0 מוגדר. הקבוצה F^0 מכילה איבר יחיד; זוהי הסדרה הריקה $():=\emptyset$, שהיא הסדרה היחידה באורך 0. בהכרח מגדירים $\vec{0}:=\emptyset$. הפעולות הן כמובן $\vec{0}:=\vec{0}+\vec{0}$ ו־ $a\cdot\vec{0}:=\vec{0}$. כל תכונות המרחב הוקטורי מתקיימות גם במקרה זה.

דוגמה 3. יהיו m ו־ n שני מספרים שלמים חיוביים. נסמן ב־ $\mathrm{M}_{m\times n}(F)$ את קבוצת המטריצות בגודל $m\times n$ עם רכיבים ב־ F. הפעולות מוגדרות כך.

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} := \begin{bmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{bmatrix}$$

ו־

$$a\cdot\begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} := \begin{bmatrix} ab_{11} & \cdots & ab_{1n} \\ \vdots & \ddots & \vdots \\ ab_{m1} & \cdots & ab_{mn} \end{bmatrix}$$

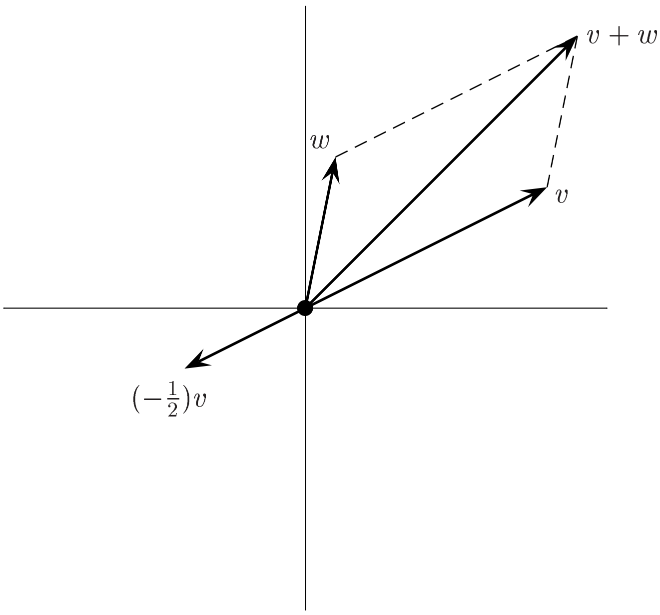
איבר האפס הוא

$$\vec{0}:=O_{m\times n}=\begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$$

בדומה לבדיקה שעשינו בדוגמה 2 רואים שגם כאן מתקיימות כל התכונות של מרחב וקטורי.

בעצם אם ניקח מטריצה $A\in\mathrm{M}_{m\times n}(F)$ ונסדר את שורותיה זו אחר זו נקבל שורה אחת באורך mn, כלומר איבר ב־ F^{mn}. רואים בקלות שפעולות החיבור והכפל בסקלר ב־ $\mathrm{M}_{m\times n}(F)$ מתאימות לפעולות החיבור והכפל בסקלר במרחב F^{mn}, ואיברי האפס מתאימים בשני המרחבים. לכן כמרחבים וקטוריים אין הבדל של ממש בין $\mathrm{M}_{m\times n}(F)$ ו־ F^{mn}; ההבדל היחיד הוא צורת כתיבת האיברים.

דוגמה 4. ניקח את השדה $F:=\mathbb{R}$ ואת מרחב הזוגות $V=\mathbb{R}^2$. למרחב $\mathbb{R}^2$ יש פירוש הגיאומטרי: זהו המישור הממשי. פעולת החיבור היא ע״י כלל המקבילית, וכפל בסקלר הוא מתיחה. (ראה ציור.)



טענה 1. ההפכי החיבורי הוא יחיד.

הוכחה. בהנתן וקטור $v \in V$ נניח שהוקטורים $w_1, w_2 \in V$ מקיימים

$$v + w_1 = v + w_2 = \vec{0}\ .$$

אז

$$w_1 + (v + w_2) = w_1 + \vec{0} = w_1$$

וגם

$$w_1 + (v + w_2) = (w_1 + v) + w_2 = (v + w_1) + w_2 = \vec{0} + w_2 = w_2 + \vec{0} = w_2\ .$$

רואים ש־ $w_1 = w_2$. משל.

טענה 2. יהי V מרחב וקטורי מעל F.

א. לכל $v \in V$ מתקיים $0 \cdot v = \vec{0}$.
ב. לכל $a \in F$ מתקיים $a \cdot \vec{0} = \vec{0}$.
ג. יהיו $a \in F$ ו־ $v \in V$. אם $a \cdot v = \vec{0}$ ו־ $a \neq 0$ אז $v = \vec{0}$.
ד. לכל $a \in F$ ו־ $v \in V$ מתקיים $(-a) \cdot v = -(a \cdot v)$.

הוכחה. א. $0 = 0 + 0$ בשדה, לכן

$$0 \cdot v = (0 + 0) \cdot v = (0 \cdot v) + (0 \cdot v)\ .$$

ע״י חיבור $-(0\cdot v)$ לשני האגפים נקבל

$$\vec{0} = 0\cdot v + (-(0\cdot v)) = (0\cdot v) + (0\cdot v) + (-(0\cdot v)) = 0\cdot v \,.$$

ב. מתכונת האפס נובע $\vec{0}+\vec{0}=\vec{0}$. מדיסטריבוטיביות מימין מקבלים

$$a\cdot\vec{0} = a\cdot(\vec{0}+\vec{0}) = (a\cdot\vec{0}) + (a\cdot\vec{0}) \,.$$

כעת נוסיף $-(a\cdot\vec{0})$ לשני האגפים במשוואה ונקבל

$$\vec{0} = (a\cdot\vec{0}) + (-(a\cdot\vec{0})) = (a\cdot\vec{0}) + (a\cdot\vec{0}) + (-(a\cdot\vec{0})) = a\cdot\vec{0} \,.$$

ג. מאחר ש־ $a\neq 0$ קיים $a^{-1}\in F$.

מחלק ב׳	$\vec{0} = a^{-1}\cdot\vec{0}$
נתון	$= a^{-1}\cdot(a\cdot v)$
אסוציאטיביות	$= (a^{-1}\cdot a)\cdot v$
הגדרת a^{-1}	$= 1\cdot v$
תכונת ה־ 1	$= v$

ד. נעשה את החישוב הבא:

דיסטריבוטיביות משמאל	$(a\cdot v) + ((-a)\cdot v) = (a+(-a))\cdot v$
הגדרת $-a$	$= 0\cdot v$
חלק א׳	$= \vec{0}$

לכן $(-a)\cdot v = -(a\cdot v)$. משי״ל.

כמו בפעולות חשבון בין סקלרים, גם במקרה של וקטורים נהוג לקצר. לדוגמה כותבים av במקום $a\cdot v$, וכן משמיטים סוגריים היכן שניתן.

תת־מרחבים

הגדרה 2. יהי V מרחב וקטורי מעל F. תת־קבוצה W של V תקרא **תת־מרחב** אם מתקיימים שלושת התנאים הבאים.

א. $\vec{0}\in W$.

ב. לכל $v,w\in W$ מתקיים $v+w\in W$.

ג. לכל $v\in W$ ו־ $a\in F$ מתקיים $a\cdot v\in W$.

טענה 3. יהי V מרחב וקטורי מעל F ו־ $W\subset V$ תת־מרחב. אז W, עם הפעולות של V, מהווה מרחב וקטורי.

הוכחה. נשים לב תחילה שב־ W ישנו האיבר $\vec{0}$ ושהפעולות $+$ ו־ $\cdot$ מוגדרות על W. כל התכונות בהגדרה 1 מתקיימות באופן אוטומטי עבור המערכת $(W,+,\cdot,\vec{0})$, מלבד תכונה 4 אותה יש צורך להוכיח. יהי

$w \in W$ וקטור כלשהו. עלינו להוכיח כי $-w \in W$. אולם זה נובע מחלק ד׳ של טענה 2, שהרי

$$-w = -(1 \cdot w) = (-1) \cdot w \in W\ .$$

מש״ל.

דוגמה 5. ניקח $F := \mathbb{Q}$, $V := \mathbb{Q}^2$ ו־

$$W := \{(0,a) \mid a \in \mathbb{Q}\} \subset Q^2\ .$$

זהו תת־מרחב.

דוגמה 6. ניקח $F := \mathbb{R}$, $V := \mathbb{R}^2$ ו־

$$W := \{(a,b) \mid a \in \mathbb{Q},\ b \in \mathbb{R}\} \subset V\ .$$

האם W תת־מרחב של V? נבדוק.

א. $\vec{0} = (0,0) \in W$; בסדר.

ב. W סגור תחת חיבור; בסדר.

ג. ניקח $a := \sqrt{2} \in \mathbb{R}$ ו־ $v := (1,0) \in W$. אז $a \cdot v = (\sqrt{2},0) \notin W$. תכונה ג׳ איננה מתקיימת.

אם כן W איננו תת־מרחב של V.

דוגמה 7. כאן $F := \mathbb{R}$ ו־ $V := \mathbb{R}^2$. הנה כמה סוגי תת־מרחבים: הראשית $\{(0,0)\}$; ישר דרך הראשית; המישור כולו. בהמשך נוכיח שאלו כל האפשרויות.

דוגמה 8. שוב $F := \mathbb{R}$ ו־ $V := \mathbb{R}^2$. הקבוצות הבאות אינן תת־מרחבים של V:

$$W := \{(a,b) \mid a^2 + b^2 \leq 1\}$$

$$W := \{(a,b) \mid b \leq a\}$$

טענה 4. נתונה מערכת משוואות הומוגנית $AX = O$ ב־ n משתנים עם מקדמים בשדה F. ניקח $V := F^n$ ו־

$$W := \{AX = O \text{ קבוצת הפתרונות של}\} \subset F^n\ .$$

אז W תת־מרחב של V.

הוכחה. א. $\vec{0} \in W$. זהו הפתרון הטריוויאלי.

ב. יהיו $(c_1, \ldots, c_n)$ ו־ $(d_1, \ldots, d_n)$ שני פתרונות. כלומר לכל משוואה i מתקיים

$$a_{i1}c_1 + \cdots + a_{in}c_n = 0$$

ו־

$$a_{i1}d_1 + \cdots + a_{in}d_n = 0\ .$$

נחבר את המשוואות ונקבל

$$a_{i1}(c_1 + d_1) + \cdots + a_{in}(c_n + d_n) = 0 + 0 = 0\ .$$

לכן גם $(c_1+d_1,\ldots,c_n+d_n)$ פתרון.

ג. יהי $(c_1,\ldots,c_n)$ פתרון ויהי d סקלר. ע"י כפל משוואה מס' i ב־ d מקבלים

$$a_{i1}dc_1+\cdots+a_{in}dc_n=d\cdot 0=0$$

ולכן גם $(dc_1,\ldots,dc_n)$ הוא פתרון. משל.

פרישה

הגדרה 3. יהי V מרחב וקטורי מעל F ותהי $\mathbf{v}=(v_1,\ldots,v_n)$ סדרה סופית של אברי V. **צרוף ליניארי** (או קומבינציה ליניארית) של הסדרה $\mathbf{v}$ הוא וקטור v מהצורה

$$v=a_1v_1+\cdots+a_nv_n$$

עבור סקלרים $a_1,\ldots,a_n\in F$ כלשהם. כאשר $n=0$ מדובר בסדרה הריקה $\mathbf{v}=\emptyset$, ועל פי הגדרה הצרוף הליניארי היחיד של הסדרה $\emptyset$ הוא הוקטור $\vec{0}$.

קבוצת כל הצרופים הליניאריים של הסדרה $\mathbf{v}$ נקראת **מרחב הפרישה** של $\mathbf{v}$, והיא מסומנת ע"י $\mathrm{Sp}(\mathbf{v})$.

דרך אחרת לבטא זאת היא

$$\mathrm{Sp}(\mathbf{v})=\{a_1v_1+\cdots+a_nv_n\mid a_1,\ldots,a_n\in F\}\ .$$

הגדרה 4. יהי V מרחב וקטורי מעל F ותהי S תת־קבוצה של V. מרחב הפרישה של S הוא קבוצת כל הצרופים הליניאריים של סדרות סופיות ב־ S. הסימון הוא $\mathrm{Sp}(S)$.

במלים אחרות, וקטור v מקיים $v\in\mathrm{Sp}(S)$ אם"ם ישנה סדרה סופית $\mathbf{v}=(v_1,\ldots,v_n)$ של איברי S, וסקלרים $a_1,\ldots,a_n\in F$, כך ש־

$$v=\sum_{i=1}^{n}a_iv_i=a_1v_1+\cdots+a_nv_n\ .$$

טענה 5. יהי V מרחב וקטורי ו־ $S\subset V$ תת־קבוצה. אז $\mathrm{Sp}(S)$ הוא תת־מרחב של V.

הוכחה. נעבור על שלושת התכונות מהגדרה 2.

א. $\vec{0}\in\mathrm{Sp}(S)$ ע"פ הגדרה (לוקחים $\mathbf{v}=\emptyset$ בהגדרה 4).

ב. יהיו $v,w\in\mathrm{Sp}(S)$. אז יש וקטורים

$$v_1,\ldots,v_m,w_1,\ldots,w_n\in S$$

וסקלרים

$$a_1,\ldots,a_m,b_1,\ldots,b_n\in F$$

כך ש־

$$v=a_1v_1+\cdots+a_mv_m$$

ו־

$$w=b_1w_1+\cdots+b_nw_n\ .$$

הסכום הוא

$$v + w = a_1v_1 + \cdots + a_mv_m + b_1w_1 + \cdots + b_nw_n$$

ולכן זה איבר של $\mathrm{Sp}(S)$.

ג. יהיו $v \in \mathrm{Sp}(S)$ ו־ $c \in F$. נרשום את v כצרוף ליניארי של אברי S כמו בפסקה הקודמת. מקבלים ש־

$$cv = (ca_1)v_1 + \cdots + (ca_m)v_m$$

שייך ל־ $\mathrm{Sp}(S)$. משי"ל.

דוגמה 9. נתבונן במרחב $V := F^2$ ובקבוצה $S := \{(1,1)\}$. יהיו $v_1, \ldots, v_n \in S$. אז $v_i = (1,1)$, ולכן לכל n־יה של סקלרים $(a_1, \ldots, a_n)$ מקבלים

$$a_1v_1 + \cdots + a_nv_n = (a_1 + \cdots + a_n) \cdot (1,1) = (a,a) \ ,$$

כאשר $a := a_1 + \cdots + a_n$. ז"א

$$\mathrm{Sp}(S) = \{(a,a) \mid a \in F\} \ .$$

דוגמה 10. כאן $V = F^3$ ו־

$$S := \{(0,0,1), (0,1,0), (1,0,0)\} \ .$$

אז

$$\mathrm{Sp}(S) = \{(a_1, a_2, a_3) \mid a_1, a_2, a_3 \in F\} = F^3 \ .$$

טענה 6. יהי V מרחב וקטורי, יהי W תת־מרחב של V, ותהי S תת־קבוצה של V. אז $S \subset W$ אם"ם $\mathrm{Sp}(S) \subset W$.

הוכחה. תחילה נניח ש־ $\mathrm{Sp}(S) \subset W$. יהי $v \in S$. אז $v = 1 \cdot v \in \mathrm{Sp}(S)$. לכן $v \in W$.

עכשיו נניח כי $S \subset W$. יהי $v \in \mathrm{Sp}(S)$. ע"פ הגדרה ישנם $v_1, \ldots, v_m \in S$ ו־ $a_1, \ldots, a_m \in F$ כך ש־ $v = a_1v_1 + \cdots + a_mv_m$. מההנחה נובע כי $v_i \in W$ לכל i. מאחר ש־ W תת־מרחב הרי גם $a_1v_1 + \cdots + a_mv_m \in W$. לכן $v \in W$. משי"ל.

מסקנה 1. תהיינה S ו־ T תת־קבוצות של V. אז $\mathrm{Sp}(S) = \mathrm{Sp}(T)$ אם"ם $S \subset \mathrm{Sp}(T)$ ו־ $T \subset \mathrm{Sp}(S)$.

הוכחה. נשתמש בטענה פעם עם $W := \mathrm{Sp}(S)$ ופעם עם $W := \mathrm{Sp}(T)$. משי"ל.

דוגמה 11. ניקח $F := \mathbb{Q}$, $V := \mathbb{Q}^3$,

$$S := \{(1,0,1), (0,1,1)\}$$

ו־

$$T := \{(1,1,2), (1,-1,0)\} \ .$$

החישובים

$$(1,0,1) = \tfrac{1}{2}\cdot(1,1,2) + \tfrac{1}{2}\cdot(1,-1,0)$$
$$(0,1,1) = \tfrac{1}{2}\cdot(1,1,2) - \tfrac{1}{2}\cdot(1,-1,0)$$

מראים כי $S \subset \mathrm{Sp}(T)$. מצד שני החישובים

$$(1,1,2) = (1,0,1) + (0,1,1)$$
$$(1,-1,0) = (1,0,1) - (0,1,1)$$

מראים כי $T \subset \mathrm{Sp}(S)$. לכן $\mathrm{Sp}(S) = \mathrm{Sp}(T)$.

הגדרה 5. יהי V מרחב וקטורי. תת־קבוצה $S \subset V$ תיקרא **קבוצה פורשת** של V אם $V = \mathrm{Sp}(S)$.

הגדרה 6. מרחב הוקטורי V יקרא **מרחב נפרש סופית** אם יש לו קבוצה פורשת סופית.

דוגמה 12. יהי $V := F^n$. לכל i מ־ 1 עד n נגדיר וקטור

$$\vec{e}_i := \quad \begin{array}{c}(0,\dots,1,\dots,0)\\ \uparrow \\ i\end{array}$$

בהנתן וקטור $v = (a_1,\dots,a_n) \in V$ מתקיים

$$v = a_1\vec{e}_1 + \cdots + a_n\vec{e}_n \,.$$

לכן $F^n = \mathrm{Sp}(\vec{e}_1,\dots,\vec{e}_n)$, ובפרט F^n מרחב וקטורי נפרש סופית.

מרחב הפולינומים $F[x]$

הגדרה 7. פולינום במשתנה אחד עם מקדמים בשדה F הוא סדרה

$$v = (a_0, a_1, a_2, \dots)$$

של איברים ב־ F, שבה $a_i = 0$ פרט למספר סופי של אינדקסים i.

תהי V קבוצת הפולינומים במשתנה אחד עם מקדמים ב־ F.

הגדרה 8. יהיו $v = (a_0, a_1, \dots)$ ו־ $w = (b_0, b_1, \dots)$ שני פולינומים, ויהי $a \in F$. נגדיר

$$v + w := (a_0 + b_0, a_1 + b_1, \dots)$$
$$a \cdot v := (aa_0, aa_1, \dots)$$
$$\vec{0} := (0, 0, \dots)$$

נשים לב כי הסדרות $v + w$ ו־ $a \cdot v$ הן פולינומים, כלומר מלבד מספר סופי כל רכיביהן הם 0.

טענה 7. המערכת $(V, +, \cdot, \vec{0})$ הינה מרחב וקטורי מעל השדה F.

ההוכחה כמעט זהה למקרה של F^n; ראה דוגמה 2.

הגדרה 9. יהי $v=(a_0,a_1,\dots)$ פולינום השונה מ־ $\vec{0}$. **המעלה** של v מוגדרת להיות

$$\deg(v):=\max\{i \mid a_i\neq 0\}\ .$$

המעלה של הפולינום $\vec{0}$ מוגדרת להיות $\deg(\vec{0}):=-1$.

הגדרה 10. עבור $0\leq n$ נגדיר

$$x^n:=(0,0,\dots,0,1,0,0,\dots)\ ,$$

הסדרה שבה 1 מופיע במקום ה־ n. הביטוי x נקרא **משתנה**, והביטוי x^n נקרא **מונום**. במקום x^0 נרשום לעתים 1.

בהנתן פולינום $v=(a_0,a_1,\dots)$ ממעלה $\leq n$ מתקיים השוויון

$$v=\sum_{i=0}^{n} a_i\cdot x^i$$

במרחב הוקטורי V. יתר על כן, כתיבה זו הינה יחידה: יש בדיוק דרך אחת לבטא את v כצרוף ליניארי של תת־סדרה סופית של סדרת המונומים $(x^0,x^1,\dots)$, אם נדרוש שהמקדמים יהיו שונים מ־ 0.

מטעמי נוחות אנו נעדיף מעתה והלאה להשתמש בהצגה הזו של הפולינומים. כלומר בדרך כלל נסמן פולינום ע״י ביטוי כמו f או $f(x)$. אם $\deg(f(x))\leq n$ אז נתאר את $f(x)$ כסכום

$$f(x)=\sum_{i=0}^{n} a_i x^i$$

עם מקדמים $a_i\in F$. כמו כן נסמן מעתה ב־ $F[x]$ את מרחב הפולינומים.

הערה. שימו לב לדמיון לאופן שבו אנו כותבים מספר מרוכב כ־ $a+b\mathbf{i}$ עם $a,b\in\mathbb{R}$.

טענה 8. $F[x]$ איננו מרחב וקטורי נפרש סופית.

הוכחה. נשים לב כי לכל מספר טבעי n הקבוצה

$$V_n:=\{f\in F[x]\mid \deg(f)\leq n\}$$

היא תת־מרחב של $F[x]$. נניח בשלילה כי ישנה קבוצה פורשת סופית $S=\{f_1,\dots,f_m\}$ של $F[x]$. ניקח n מספיק גדול כך ש־ $\deg(f_i)\leq n$ לכל $1\leq i\leq m$. לכן $f_1,\dots,f_m\in V_n$. ע״פ טענה 6 מקבלים $F[x]=\mathrm{Sp}(S)\subset V_n$, כלומר $V_n=F[x]$. אבל $x^{n+1}\notin V_n$, וזו סתירה. מש״ל.

הגדרה 11. יהי $f(x)=\sum_{i=0}^{n} a_i x^i$ פולינום עם מקדמים ב־ F ויהי $c\in F$. ההצבה של c ב־ $f(x)$ היא הסקלר

$$f(c)=\sum_{i=0}^{n} a_i c^i\in F\ .$$

מקבלים פונקציה $F\to F$, $c\mapsto f(c)$, שגם היא תסומן באות f.

טענה 9. יהיו $f(x),g(x)\in F[x]$ ו־ $a,c\in F$. אז

$$(a\cdot f)(c)=a\cdot f(c)$$

ו־

$$. (f+g)(c) = f(c) + g(c)$$

הוכחה. נרשום את הפולינומים כסכומים: $f(x) = \sum_{i=0}^n a_i x^i$ ו־ $g(x) = \sum_{i=0}^n b_i x^i$. אז $f+g$ הוא הפולינום $\sum_{i=0}^n (a_i+b_i)x^i$ ו־ $a \cdot f$ הוא הפולינום $\sum_{i=0}^n a a_i x^i$. לכן בהצבת c מקבלים

$$(f+g)(c) = \sum_{i=0}^n (a_i+b_i)c^i = \sum_{i=0}^n a_i c^i + \sum_{i=0}^n b_i c^i = f(c) + g(c)$$

ו־

$$. (a \cdot f)(C) = \sum_{i=0}^n a a_i c^i = a \sum_{i=0}^n a_i c^i = a \cdot f(c)$$

מש״ל.

פרישה ומשוואות לא הומוגניות

מעתה ואילך הסימן F^n יציין את **מרחב העמודות** באורך n, כלומר

$$. F^n := \{ \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \mid a_1, \ldots, a_n \in F \}$$

נתבונן במערכת משוואות ליניאריות לא הומוגנית $AX = B$, כאשר $A = [a_{ij}]$ מטריצה בגודל $m \times n$. על פי המוסכמה החדשה שאימצנו העמודה B היא וקטור ב־ F^m. תהיינה $C_1, \ldots, C_n \in F^m$ העמודות של המטריצה A, כלומר $A = [C_1, \ldots, C_n]$ ו־ $C_j = \begin{bmatrix} a_{1,j} \\ \vdots \\ a_{m,j} \end{bmatrix}$. אם הוקטור $\begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix}$ הוא פתרון של $AX = B$ הרי

$$\begin{aligned} a_{11}d_1 + \cdots + a_{1n}d_n &= b_1 \\ \vdots \qquad\qquad & \vdots \quad \vdots \\ a_{m1}d_1 + \cdots + a_{mn}d_n &= b_m \end{aligned}$$

או בצורה שקולה

$$d_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + d_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

ז״א הוקטור B הוא צרוף ליניארי של $C_1, \ldots, C_n$:

$$. d_1 C_1 + \cdots + d_n C_n = B$$

גם ההפך נכון. הוכחנו את התוצאה הבאה:

משפט 1. תהי $AX = B$ מערכת משוואות ליניאריות לא הומוגנית עם m משוואות ו־ n משתנים. למערכת $AX = B$ קיים פתרון אס״ם $B \in \mathrm{Sp}(C_1, \dots, C_n)$, כאשר $C_1, \dots, C_n \in F^m$ הן העמודות של המטריצה A.

תלות ליניארית

הגדרה 12. יהי V מרחב וקטורי מעל השדה F, יהי $n \geq 1$ ותהי $\mathbf{v} = (v_1, \dots, v_n)$ סדרה באורך n של וקטורים ב־ V. אם קיימים סקלרים $c_1, \dots, c_n \in F$ שאינם כולם 0, כך ש־

$$c_1 v_1 + \cdots + c_n v_n = \vec{0} \ ,$$

אז אומרים כי הסדרה $\mathbf{v}$ היא **תלויה ליניארית**. אחרת אומרים כי הסדרה $\mathbf{v}$ היא **בלתי תלויה ליניארית**.

עבור $n = 0$ הסדרה הריקה $\mathbf{v} = \emptyset$ היא בלתי תלויה ליניארית.

דרך אחרת לבטא את מושג התלות הליניארית מופיעה בטענה הבאה (אשר ההוכחה שלה היא מיידית).

טענה 10. תהי $\mathbf{v} = (v_1, \dots, v_n)$ סדרת וקטורים במרחב V. הסדרה $\mathbf{v}$ היא בלתי תלויה ליניארית אס״ם הפתרון היחיד ב־ F למשוואה

$$x_1 v_1 + \cdots x_n v_n = \vec{0}$$

הוא הפתרון הטריוויאלי $x_1 = \cdots = x_n := 0$.

דוגמה 13. השדה הוא $F := \mathbb{Q}$ והמרחב הוא $V := \mathbb{Q}^2$.

א. הסדרה (v_1, v_2, v_3) כאשר

$$v_1 := \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad v_2 := \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad v_3 := \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

היא תלויה ליניארית, משום ש־

$$v_1 - v_2 + v_3 = \vec{0} \ .$$

ב. הסדרה (v_1, v_2), כאשר

$$v_1 := \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad v_2 := \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

היא בת״ל, שהרי הפתרון היחיד למשוואה

$$x_1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

הוא $x_1 = x_2 := 0$.

ג. ניקח את הסדרה (v_1, v_2) כאשר

$$v_1 := \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad v_2 := \begin{bmatrix} 2 \\ 0 \end{bmatrix} \ .$$

אז

$$v_1 - \tfrac{1}{2} v_2 = \vec{0}$$

ולכן הסדרה תלויה ליניארית.

אנו רואים שסדרה עם חזרות, או סדרה שיש בה וקטורים פרופורציונליים זה לזה, תמיד תהיה תלויה ליניארית.

כיצד לבדוק אם סדרת וקטורים $\mathbf{v}$ במרחב F^m היא בלתי תלויה ליניארית? נניח שנתונים וקטורים

$$v_1 = \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}, \ldots, v_n = \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

הסדרה $\mathbf{v} := (v_1, \ldots, v_n)$ בלתי תלויה ליניארית אס"ם הפתרון היחיד ב־ F למשוואה

$$x_1 v_1 + \cdots + x_n v_n = \vec{0}$$

הוא הפתרון הטריוויאלי $x_1, \ldots, x_n := 0$. תהי A המטריצה $A = [a_{ij}]$ בגודל $m \times n$. כפי שראינו כבר, אנו שואלים האם יש פתרון לא טריוויאלי למערכת ההומוגנית $AX = O$. זאת נבדוק ע"י דירוג A.

דוגמה 14. ניקח $F := \mathbb{R}$ ו־ $V := \mathbb{R}^3$. נגדיר וקטורים

$$v_1 := \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, v_2 := \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, v_3 := \begin{bmatrix} -2 \\ 3 \\ 0 \end{bmatrix}, v_4 := \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}.$$

האם הסדרה $\mathbf{v} := (v_1, v_2, v_3, v_4)$ בלתי תלויה ליניארית? נגדיר $A := \begin{bmatrix} v_1 & v_2 & v_3 & v_4 \end{bmatrix}$, שהיא מטריצה בגודל 3×4. נתבונן במערכת המשוואות $AX = O$ שבה 3 משוואות ו־ 4 נעלמים. במערכת זו יש לכל היותר 3 משתנים תלויים, ולכן לפחות משתנה חופשי אחד. משום כך ישנם פתרונות לא טריוויאליים, והסדרה $\mathbf{v}$ הינה תלויה ליניארית (טענה 9).

חדי העין הבחינו שיש בדוגמה האחרונה רמז לתוצאה כללית: כל סדרה של וקטורים ב־ F^m שאורכה יותר מ־ m היא תלויה ליניארית. בהמשך נוכיח תוצאה יותר מלאה.

משפט 2. יהי V מרחב וקטורי ותהי $\mathbf{v} = (v_1, \ldots, v_n)$ סדרת וקטורים ב־ V באורך 1 לפחות. אז התנאים הבאים שקולים:

1. הסדרה $\mathbf{v}$ תלויה ליניארית.
2. לפחות אחד מבין הוקטורים $v_1, \ldots, v_n$ הוא צירוף ליניארי של קודמיו.
3. לפחות אחד מבין הוקטורים $v_1, \ldots, v_n$ הוא צירוף ליניארי של הוקטורים האחרים בסדרה.

הוכחה. נראה כי $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 1$.

$1 \Rightarrow 2$: נתון שיש סקלרים $a_1, \ldots, a_n$ לא כולם 0 כך ש־

$$a_1 v_1 + \cdots + a_n v_n = \vec{0}.$$

יהי i המספר המקסימלי כך ש־ $a_i \neq 0$. לכן

$$a_1 v_1 + \cdots + a_i v_i = \vec{0}.$$

ע"י העברת אגף נקבל

$$a_1 v_1 + \cdots + a_{i-1} v_{i-1} = -a_i v_i.$$

כפל בסקלר $-a_i^{-1}$ נותן

$$-\frac{a_1}{a_i}v_1 - \cdots - \frac{a_{i-1}}{a_i}v_{i-1} = v_i \,.$$

נשים לב שאם $i=1$ מקבלים $v_1=\vec{0}$.

$3 \Rightarrow 2$: מיידי.

$1 \Rightarrow 3$: נניח כי v_i צירוף ליניארי של הוקטורים $v_1,\ldots,v_{i-1},v_{i+1},\ldots,v_n$. ז"א

$$v_i = \sum_{j=1,\ldots,i-1,i+1,\ldots,n} a_j v_j = a_1v_1 + \cdots + a_{i-1}v_{i-1} + a_{i+1}v_{i+1} + \cdots + a_nv_n$$

לסקלרים כלשהם a_j. נגדיר $a_i := -1$, ואז

$$a_1v_1 + \cdots + a_iv_i + \cdots + a_nv_n = \vec{0} \,.$$

רואים שהסדרה $(v_1,\ldots,v_n)$ תלויה ליניארית. משל.

דוגמה 15. ניקח $F := \mathbb{R}$ ו־ $V := \mathbb{R}^2$. יהי $\mathbf{v} = (v_1, v_2)$ זוג וקטורים ב־ V. ננסה לראות מהו תת־המרחב $\mathrm{Sp}(\mathbf{v})$. יש שלושה מקרים.

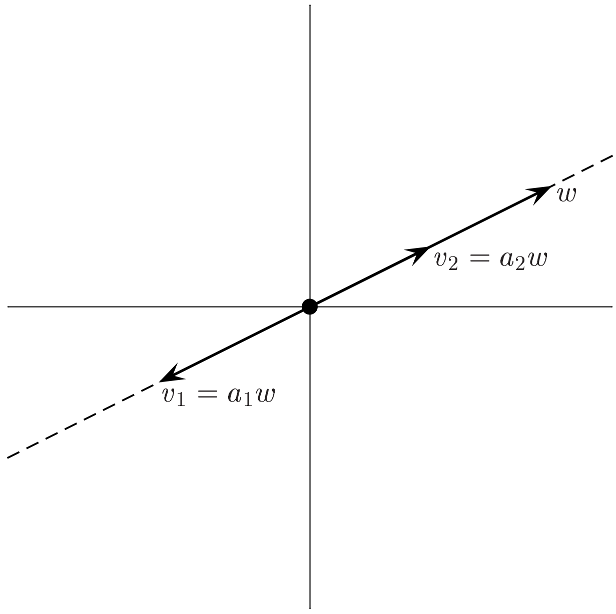
א. הסדרה $\mathbf{v}$ בלתי תלויה ליניארית. תהי A המטריצה $\begin{bmatrix} v_1 & v_2 \end{bmatrix}$, ותהי A' מטריצה מדורגת שקולת שורה ל־ A. מאחר שהפתרון היחיד למערכת המשוואות $AX=O$ הוא הפתרון הטריוויאלי הרי בהכרח $A' = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$.
כעת ניקח וקטור $w \in \mathbb{R}^2$ כלשהו. המטריצה המורחבת של מערכת המשוואות $AX=w$ היא $[A|w]$, וע"י אותו תהליך דרוג שעשינו קודם מקבלים מטריצה שקולת שורה $[A'|w'] = \left[\begin{smallmatrix} 1 & 0 & | & w_1' \\ 0 & 1 & | & w_2' \end{smallmatrix}\right]$. המסקנה היא שקיים פתרון (יחיד) למערכת המשוואות $AX=w$, ולכן $w \in \mathrm{Sp}(\mathbf{v})$. לסיכום, במקרה זה $\mathrm{Sp}(\mathbf{v}) = \mathbb{R}^2$.

ב. המקרה השני הוא ש־ $\mathbf{v}$ תלויה ליניארית, אולם $\mathbf{v} \neq (\vec{0},\vec{0})$. אז הוקטורים v_1 ו־ v_2 פרופורציוניים זה לזה; כלומר יש סקלרים a_1,a_2 ווקטור $w \neq \vec{0}$ כך ש־ $v_1 = a_1w$ ו־ $v_2 = a_2w$, אבל $(a_1,a_2) \neq (0,0)$. אם כן במקרה זה

$$\mathrm{Sp}(\mathbf{v}) = \mathrm{Sp}(w) = \{aw \mid a \in \mathbb{R}\}$$

שהוא הישר העובר דרך הראשית $\vec{0}$ והוקטור w. (ראה ציור.)

ג. המקרה השלישי הוא ש־ $\mathbf{v} = (\vec{0},\vec{0})$, ואז $\mathrm{Sp}(\mathbf{v}) = \{\vec{0}\}$.



בסיס של מרחב וקטורי

משפט 3. יהי V מרחב וקטורי מעל השדה F ותהי $\mathbf{v}$ סדרה סופית של וקטורים ב־ V. אז קיימת תת־סדרה $\mathbf{v}'$ של $\mathbf{v}$ אשר הינה בלתי תלויה ליניארית ומקיימת $\mathrm{Sp}(\mathbf{v}') = \mathrm{Sp}(\mathbf{v})$.

הוכחה. נתבונן באוסף תת־הסדרות $\mathbf{v}'$ של $\mathbf{v}$ אשר מקיימות $\mathrm{Sp}(\mathbf{v}') = \mathrm{Sp}(\mathbf{v})$. זהו אוסף לא ריק, משום שהוא כולל את $\mathbf{v}$ עצמה. תהי $\mathbf{v}'$ סדרה באוסף זה בעלת אורך מינימלי. אנו נוכיח כי $\mathbf{v}'$ היא בת״ל.

נניח על דרך השלילה כי $\mathbf{v}'$ תלויה ליניארית. נסמן $\mathbf{v}' = (v_1, \ldots, v_n)$, כאשר $n \geq 1$. ע״פ משפט 2 ישנו וקטור v_i בסדרה שהוא צרוף ליניארי של האחרים. נגדיר

$$\mathbf{v}'' := (v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n).$$

כלומר $\mathbf{v}''$ היא תת־הסדרה של $\mathbf{v}'$ המתקבלת ע״י השמטת v_i. מאחר ש־ $v_i \in \mathrm{Sp}(\mathbf{v}'')$ הרי

$$\mathrm{Sp}(\mathbf{v}'') = \mathrm{Sp}(\mathbf{v}') = \mathrm{Sp}(\mathbf{v}) .$$

אבל אורך הסדרה $\mathbf{v}''$ קטן מאשר אורכה של $\mathbf{v}'$, וזו סתירה למינימליות של $\mathbf{v}'$. מש״ל.

הגדרה 13. יהי V מרחב וקטורי. **בסיס** של V הוא סדרה פורשת בלתי תלויה ליניארית ב־ V.

מסקנה 2. יהי V מרחב וקטורי נפרש סופית. אז ל־ V יש בסיס.

הוכחה. על פי הגדרה ישנה קבוצה פורשת סופית S של V. יהי $\mathbf{v} = (v_1, \ldots, v_n)$ סידור כלשהו של S. לפי משפט 3 ישנה תת־סדרה $\mathbf{v}'$ של $\mathbf{v}$ אשר הינה בת״ל ומקיימת $V = \mathrm{Sp}(\mathbf{v}) = \mathrm{Sp}(\mathbf{v}')$. הסדרה $\mathbf{v}'$ היא בסיס של V. מש״ל.

הערה. אנו לא נעסוק בבסיסים של מרחבים וקטוריים שאינם נפרשים סופית.

הערה. במהדורות קודמות של הספר הבחנו בין ״בסיס״ ל־ ״בסיס סדור״. החל מהמהדורה הרביעית שינינו את ההגדרות והשמטנו את הביטוי ״בסיס סדור״.

דוגמה 16. יהי F שדה כלשהן. ניקח $n \geq 1$ ו־ $V := F^n$. נגדיר

$$\vec{e}_1 := \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ \vec{e}_2 := \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ \ldots, \ \vec{e}_n := \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

ראינו כבר שהסדרה $\mathbf{e} := (\vec{e}_1, \ldots, \vec{e}_n)$ פורשת את F^n. מאחר שהפתרון היחיד ב־ F למשוואה

$$x_1\vec{e}_1 + \cdots + x_n\vec{e}_n = \vec{0}$$

הוא $x_1, \ldots, x_n := 0$ הרי הסדרה $\mathbf{e}$ היא בסיס, הנקרא **הבסיס הסטנדרטי**.

דוגמה 17. יהי n מספר טבעי ויהי V מרחב הפולינומים ממעלה $\leq n$ מעל השדה F במשתנה x. הסדרה $(1, x, \ldots, x^n)$ היא בסיס של V.

דוגמה 18. השדה הוא $F := \mathbb{R}$ והמרחב הוא $V := \mathbb{R}^2$. נתבונן בסדרה $\mathbf{v} = (v_1, v_2)$ כאשר $v_1 := \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ ו־ $v_2 := \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. האם $\mathbf{v}$ בסיס של־ V?

א. האם $\mathbf{v}$ בלתי תלויה ליניארית? כלומר האם יש פתרון לא טריוויאלי למשוואה $AX = O$ כאשר $A := \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$. נדרג את A:

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

הדרגה היא 2. לכן אין פתרון לא טריוויאלי, ו־ $\mathbf{v}$ בת״ל.

ב. האם $\mathbf{v}$ פורשת את V? כלומר האם בהנתן וקטור $w \in \mathbb{R}^2$ יש פתרון למשוואה $AX = w$? נרשום $w = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ וננסה לדרג את המטריצה המורחבת $[A \mid w]$:

$$\left[\begin{array}{cc|c} 1 & 1 & b_1 \\ 1 & 2 & b_2 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 1 & b_1 \\ 0 & 1 & b_2 - b_1 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 0 & 2b_1 - b_2 \\ 0 & 1 & b_2 - b_1 \end{array}\right] = [A' \mid w']$$

ע״פ משפט שהוכחנו, מאחר שדרגת A' שווה לדרגת $[A' \mid w']$ הרי יש פתרון. לכן $\mathbf{v}$ אמנם סדרה פורשת.

מסקנה: $\mathbf{v}$ בסיס של V.

דוגמה 19. בדרך כלל יש הרבה בסיסים שונים למרחב וקטורי נתון. בתנאים של הדוגמה הקודמת יהיו c,d שני מספרים ממשיים שונים מ־ 0. אז הסדרה $\left(\begin{bmatrix} c \\ c \end{bmatrix}, \begin{bmatrix} d \\ 2d \end{bmatrix} \right)$ היא בסיס של $\mathbb{R}^2$.

ננסה עתה לענות של השאלה הבאה. נניח שנתונה סדרת וקטורים $\mathbf{v} = (v_1, \ldots, v_n)$ במרחב F^m. כיצד למצוא בסיס ל־ $\mathrm{Sp}(\mathbf{v})$? התשובה ניתנת במשפט הבא.

משפט 4. נתונה סדרת וקטורים $\mathbf{v} = (v_1, \ldots, v_n)$ במרחב הוקטורי F^m מעל השדה F. תהי A המטריצה בגודל $m \times n$ שעמודותיה $v_1, \ldots, v_n$, ותהי A' מטריצה מדורגת שקולת־שורה ל־ A. יהיו $x_{k_1}, \ldots, x_{k_r}$ המשתנים **התלויים** של מערכת המשוואות $A'X = O$. אז הסדרה $\mathbf{v}' := (v_{k_1}, \ldots, v_{k_r})$ היא בסיס של המרחב $\mathrm{Sp}(\mathbf{v})$.

הוכחה. א. נוכיח אי־תלות של הסדרה $(v_{k_1}, \ldots, v_{k_r})$. עמודה מס' k_j במטריצה A' היא הוקטור $\vec{e}_j \in F^m$. כלומר המטריצה $B := [v_{k_1} \ \cdots \ v_{k_r}]$ שקולת־שורה למטריצה $B' := [\vec{e}_1 \ \cdots \ \vec{e}_r]$. לכן המשוואה

$$x_{k_1} v_{k_1} + \cdots + x_{k_r} v_{k_r} = \vec{0}$$

שקולה למשוואה

$$x_{k_1} \vec{e}_1 + \cdots + x_{k_r} \vec{e}_r = \vec{0} \ ,$$

אשר אין לה פתרון לא טריוויאלי.

ב. פרישה: לכל $l \notin \{k_1, \ldots, k_r\}$, כלומר לכל משתנה חופשי x_l, רוצים לקבל את הוקטור v_l כצרוף ליניארי של הוקטורים $v_{k_1}, \ldots, v_{k_r}$. נציב $x_l := -1$, וביתר המשתנים החופשיים $x_{l'}$ נציב $x_{l'} := 0$. יהי $(d_1, \ldots, d_n)$ הפתרון המתקבל למערכת המשוואות $AX = O$. זאת אומרת $\sum_{j=1}^{n} d_j v_j = \vec{0}$, $d_l = -1$, ו־ $d_{l'} = 0$ לאינדקסים l' המתאימים ליתר המשתנים החופשיים. ע״י העברת אגף מקבלים

$$d_{k_1} v_{k_1} + \cdots + d_{k_r} v_{k_r} = v_l \ .$$

מש״ל.

בסיס למרחב הפתרונות של מערכת משוואות

בהנתן מערכת משוואות הומוגניות $AX = O$ אנו יודעים כי קבוצת הפתרונות היא מרחב וקטורי. מטרתנו כעת היא למצוא בסיס למרחב זה.

משפט 5. תהי A מטריצה בגודל $m \times n$ מעל השדה F, ויהי W מרחב הפתרונות של מערכת המשוואות ההומוגנית $AX = O$. תהי A' מטריצה מדורגת שקולת־שורה ל־ A, ויהיו $x_{l_1}, \ldots, x_{l_{n-r}}$ המשתנים החופשיים של מערכת המשוואות $A'X = O$. עבור i בטווח $1, \ldots, n-r$ יהי $w_i \in F^n$ הפתרון היחיד למערכת המשוואות $AX = O$ שבו $x_{l_i} := 1$, ו־ $x_{l'} := 0$ לכל המשתנים החופשיים האחרים (ראה משפט 3 בפרק ב'). אז הסדרה $(w_1, \ldots, w_{n-r})$ היא בסיס של W.

הוכחה. בהינתן סקלרים $c_1, \ldots, c_{l_{n-r}}$ יהי $w := \sum_{i=1}^{n-r} c_i w_i \in F^n$. נשים לב כי הרכיב ה־ l_i של העמודה w הוא בדיוק הסקלר c_i, ולכן w הוא הפתרון היחיד למערכת המשוואות $AX = O$ שבו $x_{l_i} := c_i$ לכל המשתנים החופשיים. אם $w = \vec{0}$ הרי $c_1, \ldots, c_{l_{n-r}} = 0$; לכן הסדרה $(w_1, \ldots, w_{n-r})$ בלתי תלויה ליניארית. מאחר שכל הפתרונות למערכת המשוואות $AX = O$ מתקבלים בצורה כזו הרי $W = \mathrm{Sp}(w_1, \ldots, w_{n-r})$. מש״ל.

דוגמה 20. ניקח $F := \mathbb{Q}$ ו־ $A := \begin{bmatrix} 1 & -2 & 1 & 1 \\ 2 & 1 & 1 & 2 \\ 0 & 5 & -1 & 0 \end{bmatrix}$. מחפשים בסיס למרחב הפתרונות W של מערכת המשוואות $AX = O$. נדרג את המטריצה:

$$A = \begin{bmatrix} 1 & -2 & 1 & 1 \\ 2 & 1 & 1 & 2 \\ 0 & 5 & -1 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & \frac{3}{5} & 1 \\ 0 & 1 & -\frac{1}{5} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = A'$$

המשתנים החופשיים הם x_3 ו־ x_4, האינדקסים הם $l_1 = 3$ ו־ $l_2 = 4$, והפתרונות המתאימים הם

$$w_1 := \begin{bmatrix} -\frac{3}{5} \\ \frac{1}{5} \\ 1 \\ 0 \end{bmatrix}, \; w_2 := \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

הסדרה (w_1, w_2) היא בסיס של W.

מימד של מרחב וקטורי

הנושא שנלמד כעת הוא מושג המימד של מרחב וקטורי. זו דרך אלגברית לבטא את מושג המימד המוכר מהגיאומטריה האוקלידית (נקודה, ישר, מישור...).

תהי $\mathbf{v}$ סדרה סופית. נסמן ב־ $|\mathbf{v}|$ את האורך של $\mathbf{v}$.

משפט 6. יהי V מרחב וקטורי מעל שדה F, ותהיינה $\mathbf{v}$ ו־ $\mathbf{w}$ שתי סדרות סופיות של וקטורים ב־ V. נניח כי $\mathbf{v}$ פורשת את V ו־ $\mathbf{w}$ בלתי תלויה ליניארית. אז $|\mathbf{w}| \leq |\mathbf{v}|$.

הוכחה. אם $V = \{\vec{0}\}$ הרי בהכרח $\mathbf{w} = \emptyset$, ולכן $|\mathbf{w}| = 0 \leq |\mathbf{v}|$.

עתה נניח ש־ $V \neq \{\vec{0}\}$. מאחר שהסדרה $\mathbf{v}$ פורשת את V נובע ש־ $|\mathbf{v}| \geq 1$. נרשום $\mathbf{v} = (v_1, \ldots, v_m)$ ו־ $\mathbf{w} = (w_1, \ldots, w_n)$. נניח על דרך השלילה כי $n > m$.

מאחר שהסדרה $\mathbf{v}$ פורשת את V הרי ניתן למצוא סקלרים a_{ij} כך ש־

$$w_j = \sum_{i=1}^{m} a_{ij} v_i$$

לכל j. נתבונן במטריצה $A := [a_{ij}]$ שגודלה $m \times n$. מאחר ש־ $m < n$ נובע שקיים פתרון לא טריוויאלי $\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ למערכת המשוואות $AX = O$. זה אומר שלכל i מתקיים

$$\sum_{j=1}^{n} a_{ij} c_j = 0 \,.$$

אבל אז

$$\sum_{j=1}^{n} c_j w_j = \sum_{j=1}^{n} c_j \left(\sum_{i=1}^{m} a_{ij} v_i \right) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} c_j \right) v_i = \vec{0} \,.$$

מאחר שלא כל הסקלרים $c_1, \ldots, c_n$ הם 0 נובע שהסדרה $\mathbf{w}$ היא תלויה ליניארית. זו סתירה. מש״ל.

מסקנה 3. יהיו $\mathbf{v}$ ו־ $\mathbf{w}$ שני בסיסים (סופיים) של המרחב הוקטורי V. אז $|\mathbf{w}| = |\mathbf{v}|$.

הוכחה. ע״פ משפט 6 ידוע לנו ש־ $|\mathbf{w}| \leq |\mathbf{v}|$ וגם $|\mathbf{v}| \leq |\mathbf{w}|$. מש״ל.

הודות למסקנה יש משמעות להגדרה הבאה.

הגדרה 14. יהי V מרחב וקטורי עם בסיס $\mathbf{v}$ באורך n. המספר n נקרא **המימד** של V, ומסומן $\dim(V)$.

דוגמה 21. ניקח $V := F^n$, $n \geq 1$. למרחב זה יש בסיס $\mathbf{e} = (\vec{e}_1, \ldots, \vec{e}_n)$ שאורכו n. במקרה $n = 0$ הבסיס של $F^0 = \{\vec{0}\}$ הוא הסדרה הריקה $\emptyset$ שאורכה 0. לכן תמיד $\dim(F^n) = n$.

לעתים נשתמש בקיצורים בת״ל ו־ ת״ל עבור 'בלתי תלויה ליניארית' ו־ 'תלויה ליניארית'.

מסקנה 4. יהי V מרחב וקטורי ממימד n ותהי $\mathbf{v}$ סדרה סופית ב־ V.

א. אם $\mathbf{v}$ פורשת את V אז $|\mathbf{v}| \geq n$.
ב. אם $\mathbf{v}$ בת״ל אז $|\mathbf{v}| \leq n$.
ג. אם $\mathbf{v}$ פורשת את V ו־ $|\mathbf{v}| \leq n$ אז $\mathbf{v}$ בסיס של V ו־ $|\mathbf{v}| = n$.
ד. אם $\mathbf{v}$ בת״ל ו־ $|\mathbf{v}| \geq n$ אז $\mathbf{v}$ בסיס של V ו־ $|\mathbf{v}| = n$.

הוכחה. ניקח בסיס כלשהו $\mathbf{w}$ של V. ע״פ הגדרת המימד מתקיים $|\mathbf{w}| = n$.

א. $\mathbf{w}$ בת״ל, לכן על פי המשפט $|\mathbf{v}| \geq |\mathbf{w}| = n$.

ב. $\mathbf{w}$ פורשת את V, ולכן על פי המשפט $|\mathbf{v}| \leq |\mathbf{w}| = n$.

ג. לפי משפט 3 ישנה תת־סדרה $\mathbf{v}'$ של $\mathbf{v}$ שהיא בסיס של V. מחלק א' נובע ש־ $n \leq |\mathbf{v}'|$. לכן $n = |\mathbf{v}| = |\mathbf{v}'|$ ו־ $\mathbf{v} = \mathbf{v}'$.

ד. נניח על דרך השלילה ש־ $\mathbf{v}$ איננה פורשת את V. אז ישנו וקטור $v \in V - \mathrm{Sp}(\mathbf{v})$. הסדרה $\mathbf{v}' := (\mathbf{v}, v)$ היא באורך $n + 1$ לפחות, והיא עדיין בת״ל. זו סתירה לחלק ב' של הטענה. מש״ל.

מסקנה 5. יהי V מרחב וקטורי ממימד n ויהי $W \subset V$ תת־מרחב. אז $\dim(W) \leq n$. אם $\dim(W) = n$ אז $W = V$.

הוכחה. כל סדרה בת״ל $\mathbf{w}$ ב־ W היא גם סדרה בת״ל ב־ V, ובעזרת מסקנה 4(ב) מקבלים $|\mathbf{w}| \leq n$. תהי $\mathbf{w}$ סדרה בת״ל ב־ W בעלת האורך המקסימלי m. ברור ש־ $m \leq n$. אנו טוענים ש־ $\mathbf{w}$ פורשת את W. אחרת כמו בהוכחת מסקנה 4(ד) ישנו וקטור $w \in W - \mathrm{Sp}(\mathbf{w})$. הסדרה $\mathbf{w}' := (\mathbf{w}, w)$ היא באורך $m + 1$ והיא עדיין בת״ל. זו סתירה למקסימליות של m.
אם $m = n$ אז לפי מסקנה 4(ד) הסדרה $\mathbf{w}$ פורשת גם את V, ולכן $W = V$. מש״ל.

דוגמה 22. השדה הוא $\mathbb{R}$ והמרחב הוא המישור $\mathbb{R}^2$. יהי $W \subset \mathbb{R}^2$ תת־מרחב. לפי מסקנה 5 הערכים האפשריים ל־ $\dim(W)$ הם 0, 1 או 2.

- אם $\dim(W) = 0$ אז $W = \{\vec{0}\}$. בסיס הוא $\mathbf{w} = \emptyset$.

- אם $\dim(W) = 1$ אז $W = \mathrm{Sp}(w)$, כאשר w הוא וקטור כלשהו ב־ W השונה מ־ $\vec{0}$. כלומר W הוא הישר העובר דרך w ו־ $\vec{0}$. בסיס ל־ W הוא $\mathbf{w} = (w)$.
- אם $\dim(W) = 2$ אז $W = \mathbb{R}^2$. בתור בסיס אפשר לקחת $\mathbf{e} = (\vec{e}_1, \vec{e}_2)$.

דוגמה 23. נתבונן במערכת משוואות הומוגנית $AX = O$ שבה m משוואות ב־ n משתנים. תהי A' מטריצה מדורגת שקולת שורה ל־ A, ונניח שבמערכת המשוואות $A'X = O$ יש r משתנים תלויים. יהי W מרחב הפתרונות. במשפט 5 ראינו בסיס ל־ W באורך $n - r$. לכן $\dim(W) = n - r$.

דוגמה 24. כאן $F := \mathbb{R}$,

$$V := \{\text{פולינומים ממעלה } \leq 3 \text{ במשתנה } x \text{ מעל } \mathbb{R}\}$$

ו־

$$. \ W := \{f(x) \in V \mid f(-1) = 0\}$$

בשל טענה 9 (עם $c := -1$) רואים כי W הוא תת־מרחב של V. נמצא בסיס ל־ W. נרשום

$$. \ f(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3$$

התנאי ש־ $f(x) \in W$ הוא

$$. \ 0 = f(-1) = c_0 - c_1 + c_2 - c_3$$

אם כן אנו מחפשים פתרון המשוואה ההומוגנית

$$(*) \quad y_1 - y_2 + y_3 - y_4 = 0$$

במשתנים y_1, y_2, y_3, y_4. בסיס למרחב הפתרונות W' של המשוואה $(*)$ הוא הסדרה (w'_1, w'_2, w'_3), כאשר

$$. \ w'_1 := \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \ w'_2 := \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \ w'_3 := \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

התרגום לפולינומים נותן בסיס $(f_1(x), f_2(x), f_3(x))$ ל־ W, כאשר

$$. \ f_1(x) := 1 + x, \ f_2(x) := -1 + x^2, \ f_3(x) := 1 + x^3$$

המימד הוא $\dim(W) = 3$.

מסקנה 6. תהי $AX = B$ מערכת של n משוואות ב־ n נעלמים (ז״א המטריצה A ריבועית). נתון כי הפתרון היחיד למערכת המשוואות ההומוגנית $AX = O$ הוא הפתרון הטריוויאלי. אז קיים פתרון למערכת המשוואות $AX = B$.

הוכחה. תהיינה $v_1, \ldots, v_n$ העמודות של A. מכך שאין פתרון לא טריוויאלי ל־ $AX = O$ נובע שהסדרה $(v_1, \ldots, v_n)$ היא בת״ל. לכן לפי מסקנה 4(ד) רואים ש־ $(v_1, \ldots, v_n)$ בסיס של F^n. בפרט $F^n = \mathrm{Sp}(v_1, \ldots, v_n)$, ולכן יש פתרון ל־ $AX = B$. מש״ל.

משפט 7. תהי $(v_1, \ldots, v_n)$ סדרת וקטורים במרחב הוקטורי V מעל השדה F. התנאים הבאים שקולים:

1. $(v_1, \ldots, v_n)$ היא בסיס של V.

2. כל וקטור v ב־ V ניתן לכתיבה יחידה כצרוף ליניארי

$$v = a_1 v_1 + \cdots + a_n v_n$$

עם סקלרים $a_1, \ldots, a_n \in F$.

הוכחה. $1 \Leftarrow 2$: ידוע כי $V = \mathrm{Sp}(v_1, \ldots, v_n)$. לכן בהנתן $v \in V$ קיימים סקלרים $a_1, \ldots, a_n$ כך ש־ $v = \sum_{i=1}^n a_i v_i$. אם גם $v = \sum_{i=1}^n b_i v_i$ הרי

$$\vec{0} = v - v = (a_1 - b_1)v_1 + \cdots + (a_n - b_n)v_n \ .$$

בשל כך שהסדרה $(v_1, \ldots, v_n)$ בת״ל מקבלים $a_i - b_i = 0$ לכל i. כלומר $a_i = b_i$.

$2 \Leftarrow 1$: העובדה שהוקטורים $v_1, \ldots, v_n$ פורשים את V ברורה. נוכיח אי תלות של הסדרה. ניקח $v := \vec{0}$. היחידות בתנאי 2 אומרת שהפתרון היחיד למשוואה $x_1 v_1 + \cdots + x_n v_n = \vec{0}$ הוא הפתרון הטריוויאלי. משׁ״ל.

הגדרה 15. יהי $\mathbf{v} = (v_1, \ldots, v_n)$ בסיס של מרחב וקטורי V מעל שדה F, ויהי v וקטור כלשהו ב־ V. משפט 7 אומר שיש וקטור יחיד $\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \in F^n$ כך ש־ $v = \sum_{i=1}^n a_i v_i$. הוקטור $\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$ נקרא **וקטור הקואורדינטות** של v ביחס לבסיס $\mathbf{v}$, והוא יסומן $[v]_{\mathbf{v}}$.

דוגמה 25. יהי F שדה כלשהו ו־ $V := F^n$. הבסיס הסטנדרטי $\mathbf{e} = (\vec{e}_1, \ldots, \vec{e}_n)$ הוא בסיס של F^n. לכל וקטור $v \in F^n$ מתקיים $[v]_{\mathbf{e}} = v$.

דוגמה 26. ניקח $F := \mathbb{R}$ ו־ $V := \mathbb{R}^2$. נגדיר וקטורים $v_1 := \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ ו־ $v_2 := \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. הסדרה $\mathbf{v} := (v_1, v_2)$ היא בסיס. נחשב את הקואורדינטות של אברי הבסיס הסטנדרטי $\mathbf{e}$ ביחס לבסיס $\mathbf{v}$. מאחר ש־ $\vec{e}_1 = \frac{1}{2} v_1$ ו־ $\vec{e}_2 = v_2 - \frac{1}{2} v_1$ מקבלים $[\vec{e}_1]_{\mathbf{v}} = \begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix}$ ו־ $[\vec{e}_2]_{\mathbf{v}} = \begin{bmatrix} -\frac{1}{2} \\ 1 \end{bmatrix}$.

טענה 11. יהי V מרחב וקטורי ממימד n ויהי W תת־מרחב של V ממימד m. נניח כי $(v_1, \ldots, v_m)$ בסיס של W. אז קיימים וקטורים $v_{m+1}, \ldots, v_n \in V$ כך שהסדרה

$$(v_1, \ldots, v_m, v_{m+1}, \ldots, v_n)$$

הינה בסיס למרחב V.

הוכחה. נגדיר באינדוקציה על i בטווח $0, \ldots, n-m$ סדרה בת״ל

$$(v_1, \ldots, v_m, \ldots, v_{m+i}) \ .$$

עבור $i = 0$ זה נתון. אם $i < n - m$ אז $m + i < n$ ולכן

$$\mathrm{Sp}(v_1, \ldots, v_m, \ldots, v_{m+i}) \neq V \ .$$

קיים אם כן וקטור $v_{m+i+1} \in V$ שאיננו צרוף ליניארי של $v_1, \ldots, v_m, \ldots, v_{m+i}$, והסדרה

$$(v_1, \ldots, v_m, \ldots, v_{m+i}, v_{m+i+1})$$

היא בת״ל.

כאשר $i = n - m$ נקבל סדרה בת״ל $(v_1, \dots, v_n)$, שהיא בהכרח בסיס של V. משל״ל.

טענה 12. יהי V מרחב וקטורי ויהיו W_1 ו־ W_2 שני תת מרחבים של V. אז החיתוך $W_1 \cap W_2$ הוא תת מרחב.

הוכחה. מאחר ש־ $\vec{0} \in W_1$ וגם $\vec{0} \in W_2$ הרי $\vec{0} \in W_1 \cap W_2$. יהיו $w_1, w_2 \in W_1 \cap W_2$. אז $w_1 + w_2 \in W_1$ וגם $w_1 + w_2 \in W_2$; לכן $w_1 + w_2 \in W_1 \cap W_2$. לבסוף לכל $a \in F$ מתקיים $aw_1 \in W_1$ וגם $aw_1 \in W_2$; לכן $aw_1 \in W_1 \cap W_2$. משל״ל.

דוגמה 27. איחוד של שני תת מרחבים בדרך כלל איננו תת מרחב. ניקח לדוגמה את המרחב $\mathbb{R}^2$ מעל השדה $\mathbb{R}$, ואת תת המרחבים $W_1 := \{(a, 0) \mid a \in \mathbb{R}\}$ ו־ $W_2 := \{(0, a) \mid a \in \mathbb{R}\}$. אז $W_1 \cup W_2$ איננו תת מרחב.

הגדרה 16. יהי V מרחב וקטורי ויהיו W_1 ו־ W_2 שני תת־מרחבים של V. הסכום $W_1 + W_2$ הוא התת־מרחב הבא של V:

$$W_1 + W_2 := \mathrm{Sp}(W_1 \cup W_2).$$

משפט 8. יהי V מרחב וקטורי סוף מימדי מעל שדה F ויהיו W_1 ו־ W_2 תת־מרחבים של V. אז מתקיים השוויון:

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2).$$

הוכחה. נבחר בסיס $\mathbf{w}_0 = (w_1, \dots, w_l)$ ל־ $W_1 \cap W_2$. עפ״י טענה 11 ניתן להוסיף וקטורים ל־ $\mathbf{w}_0$ כך ש־

$$\mathbf{w}_1 := (w_1, \dots, w_l, u_{l+1}, \dots, u_m)$$

בסיס של W_1 ו־

$$\mathbf{w}_2 := (w_1, \dots, w_l, v_{l+1}, \dots, v_n)$$

בסיס של W_2. נגדיר

$$\mathbf{w} := (w_1, \dots, w_l, u_{l+1}, \dots, u_m, v_{l+1}, \dots, v_n).$$

זו סדרת וקטורים באורך

$$m + n - l = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2).$$

אנו נוכיח כי $\mathbf{w}$ בסיס של $W_1 + W_2$.

מאחר ש־ $\mathbf{w}_1$ ו־ $\mathbf{w}_2$ הן תת־סדרות של $\mathbf{w}$ נובע ש־ $\mathbf{w}$ פורשת את $W_1 + W_2$. נוכיח כי סדרה זו בת״ל. נניח כי נתונים סקלרים a_i, b_i, c_i כך ש־

$$a_1 w_1 + \cdots + a_l w_l + b_{l+1} u_{l+1} + \cdots + b_m u_m + c_{l+1} v_{l+1} + \cdots + c_n v_n = \vec{0}.$$

נגדיר וקטור

$$w := a_1 w_1 + \cdots + a_l w_l + b_{l+1} u_{l+1} + \cdots + b_m u_m.$$

אז $w \in W_1$. אבל גם

$$, \ w = 0 \cdot w_1 + \cdots + 0 \cdot w_l + (-c_{l+1})v_{l+1} + \cdots + (-c_n)v_n$$

ולכן $w \in W_2$. אם כן $w \in W_1 \cap W_2$. מאחר ש־ $\mathbf{w}_0$ בסיס של $W_1 \cap W_2$ הרי ישנם סקלרים $d_1, \ldots, d_l$ כך ש־

$$. \ w = d_1 w_1 + \cdots + d_l w_l$$

מהיחידות של הכתיבה של w בבסיס $\mathbf{w}_1$ של W_1 מקבלים $a_1 = d_1, \ldots, a_l = d_l$, $b_{l+1} = 0, \ldots$, $b_m = 0$. מהיחידות של הכתיבה של w בבסיס $\mathbf{w}_2$ של W_2 מקבלים $d_1 = 0, \ldots, d_l = 0$, $c_{l+1} = 0, \ldots, c_n = 0$. מכאן נובע ש־ $a_i = 0$, $b_i = 0$ ו־ $c_i = 0$ לכל האינדקסים i. משי״ל.

דוגמה 28. במרחב $\mathbb{Q}^3$ מעל השדה $\mathbb{Q}$ ניקח את הוקטורים

$$. \ v_1 := \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \ v_2 := \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \ v_3 := \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \ v_4 := \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

נגדיר $W_1 := \mathrm{Sp}(v_1, v_2)$ ו־ $W_2 := \mathrm{Sp}(v_3, v_4)$. ברצוננו לחשב את $W_1 + W_2$ ו־ $W_1 \cap W_2$. מאחר ש־ $W_1 + W_2 = \mathrm{Sp}(v_1, v_2, v_3, v_4)$ הרי בסיס למרחב זה אפשר למצוא ע״י דרוג המטריצה הבאה.

$$\begin{bmatrix} v_1 & v_2 & v_3 & v_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & -\frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & -\frac{1}{2} \end{bmatrix}$$

רואים ש־ $W_1 + W_2 = \mathbb{Q}^3$. מאחר ש־

$$\dim(W_1) = \dim(W_2) = 2$$

מקבלים ממשפט 8 ש־ $\dim(W_1 \cap W_2) = 1$.

עתה ננסה למצוא בסיס של $W_1 \cap W_2$. נשים לב שוקטור $w \in W_1 \cap W_2$ הוא וקטור מהצורה

$$w = a_1 v_1 + a_2 v_2 = a_3 v_3 + a_4 v_4$$

עבור סקלרים a_1, a_2, a_3, a_4. רביעיית סקלרים זו היא פתרון של המשוואה

$$. \ x_1 v_1 + x_2 v_2 + x_3(-v_3) + x_4(-v_4) = \vec{0}$$

נפתור משוואה זו ע״י דרוג:

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & -1 & 0 \\ 1 & 3 & 0 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & -\frac{1}{2} \end{bmatrix}$$

בסיס למרחב הפתרונות הוא הוקטור $\begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{bmatrix}$, והבסיס המתאים ל־ $W_1 \cap W_2$ הוא הוקטור

$$. \ w := -\tfrac{1}{2} v_1 + \tfrac{1}{2} v_2 = \tfrac{1}{2} v_3 + v_4 = \begin{bmatrix} 0 \\ \frac{1}{2} \\ 1 \end{bmatrix}$$

ד. חשבון מטריצות

בפרק זה נלמד כיצד לכפול מטריצות. נזכיר כי לכל $m, n \geq 1$, $\mathrm{M}_{m\times n}(F)$ הוא מרחב המטריצות בגודל $m \times n$ מעל השדה F.

הגדרה 1. בהינתן מטריצות $A \in \mathrm{M}_{m\times n}(F)$ ו־ $B \in \mathrm{M}_{n\times p}(F)$ נגדיר מטריצה $A \cdot B \in \mathrm{M}_{m\times p}(F)$ באופן הבא. נרשום $A = [a_{ij}]$ ו־ $B = [b_{jk}]$, ונגדיר

$$c_{ik} := \sum_{j=1}^{n} a_{ij} b_{jk}$$

לכל i, k. אז המכפלה היא

$$A \cdot B := [c_{ik}] \, .$$

בצורה גראפית זה נראה כך:

$$\begin{array}{ccccc} [a_{ij}] & \cdot & [b_{jk}] & = & [c_{ik}] \\ \begin{bmatrix} * & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ * & * & \cdots & * \end{bmatrix} & \cdot & \begin{bmatrix} * & \cdots & b_{1k} & \cdots & * \\ * & \cdots & b_{2k} & \cdots & * \\ \vdots & & \vdots & & \vdots \\ * & \cdots & b_{nk} & \cdots & * \end{bmatrix} & = & \begin{bmatrix} * & \cdots & * & \cdots & * \\ \vdots & & \vdots & & \vdots \\ * & \cdots & c_{ik} & \cdots & * \\ \vdots & & \vdots & & \vdots \\ * & \cdots & * & \cdots & * \end{bmatrix} \end{array}$$

יש לשים לב שהמכפלה $A \cdot B$ מוגדרת רק אם אורך השורות ב־ A שווה לאורך העמודות ב־ B!

בשתי הדוגמאות הבאות השדה הוא $F := \mathbb{R}$.

דוגמה 1. עבור $A := \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ ו־ $B := \begin{bmatrix} 5 & 6 & 7 \\ 8 & 9 & 10 \end{bmatrix}$ המכפלה היא

$$A \cdot B = \begin{bmatrix} 21 & 24 & 27 \\ 47 & 54 & 61 \end{bmatrix} \, .$$

נשים לב כי לא מוגדרת המכפלה $B \cdot A$.

דוגמה 2. ניקח $A := \begin{bmatrix} -1 \\ 3 \end{bmatrix}$ ו־ $B := \begin{bmatrix} 2 & 4 \end{bmatrix}$. כאן מוגדרות שתי המכפלות:

$$A \cdot B = \begin{bmatrix} -2 & -4 \\ 6 & 12 \end{bmatrix}$$

ו־

$$B \cdot A = [10] \, .$$

נשים לב כי $A \cdot B \neq B \cdot A$.

דוגמה 3. כאן F הוא שדה כלשהו. תהי $A := I_{m\times m} = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}$, ותהי $B \in \mathrm{M}_{m\times n}(F)$ מטריצה כלשהי. נחשב את הרכיב ה־ (i,k) של המטריצה $A \cdot B$. תחילה נעשה זאת בעזרת הייצוג הגראפי:

$$\begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \cdot \begin{bmatrix} * & \cdots & b_{1k} & \cdots & * \\ \vdots & & \vdots & & \vdots \\ * & \cdots & b_{ik} & \cdots & * \\ \vdots & & \vdots & & \vdots \\ * & \cdots & b_{mk} & \cdots & * \end{bmatrix} = \begin{bmatrix} * & \cdots & * & \cdots & * \\ \vdots & & \vdots & & \vdots \\ * & \cdots & b_{ik} & \cdots & * \\ \vdots & & \vdots & & \vdots \\ * & \cdots & * & \cdots & * \end{bmatrix}$$

עתה נעשה זאת בנוסחה. מהגדרת המטריצה A ידוע ש־

$$a_{ij} = \begin{cases} 1; & i = j \\ 0; & i \neq j \end{cases}.$$

לכן

$$\sum_{j=1}^{m} a_{ij} b_{jk} = \left(\sum_{j=i} a_{ij} b_{jk} \right) + \left(\sum_{j\neq i} a_{ij} b_{jk} \right) = 1 \cdot b_{ik} + 0 = b_{ik}.$$

מצאנו שהרכיב ה־ (i,k) של $I_{m\times m} \cdot B$ זהה להרכיב ה־ (i,k) של B. לכן

$$I_{m\times m} \cdot B = B.$$

בדומה מראים כי

$$B \cdot I_{n\times n} = B.$$

דוגמה 4. מטריצה סקלרית היא מטריצה מהצורה

$$a \cdot I_{n\times n} = \begin{bmatrix} a & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a \end{bmatrix}$$

כאשר $a \in F$. תהי $A := a \cdot I_{n\times n}$ מטריצה סקלרית ותהי $B \in \mathrm{M}_{n\times p}(F)$ מטריצה כלשהי. אותו חישוב מהדוגמה הקודמת מראה ש־

$$A \cdot B = a \cdot B.$$

בדומה אם $B \in \mathrm{M}_{m\times n}(F)$ מקבלים

$$B \cdot A = a \cdot B.$$

בפרט כאשר B מטריצה ריבועית בגודל $n \times n$ הרי

$$A \cdot B = a \cdot B = B \cdot A.$$

טענה 1. כפל מטריצות הוא אסוציאטיבי. זאת אומרת שבהנתן $A \in \mathrm{M}_{m\times n}(F)$, $B \in \mathrm{M}_{n\times p}(F)$ ו־ $C \in \mathrm{M}_{p\times q}(F)$ מתקיים

$$A \cdot (B \cdot C) = (A \cdot B) \cdot C.$$

הוכחה. לכל זוג אינדקסים (j,l) הרכיב ה־ (j,l) במטריצה $B\cdot C$ הוא $\sum_{k=1}^{p} b_{jk}c_{kl}$. לכן לכל זוג אינדקסים (i,l) הרכיב ה־ (i,l) במטריצה $A\cdot(B\cdot C)$ הוא

$$(*)\qquad \sum_{j=1}^{n} a_{ij}\left(\sum_{k=1}^{p} b_{jk}c_{kl}\right)$$

מצד שני הרכיב ה־ (i,k) במטריצה $A\cdot B$ הוא $\sum_{j=1}^{n} a_{ij}b_{jk}$. מקבלים שהרכיב ה־ (i,l) במטריצה $(A\cdot B)\cdot C$ הוא

$$(**)\qquad \sum_{k=1}^{p}\left(\sum_{j=1}^{n} a_{ij}b_{jk}\right)c_{kl}$$

בשל תכונות החשבון בשדה הסכומים $(*)$ ו־ $(**)$ שווים. משי"ל.

טענה 2. כפל מטריצות הוא דיסטריבוטיבי. כלומר אם $A,B\in \mathrm{M}_{m\times n}(F)$ ו־ $C,D\in \mathrm{M}_{n\times p}(F)$ אז

$$A\cdot(C+D)=(A\cdot C)+(A\cdot D)$$

ו־

$$(A+B)\cdot C=(A\cdot C)+(B\cdot C)\ .$$

הוכחה. נרשום $A=[a_{ij}]$, $C=[c_{ij}]$ ו־ $D=[d_{ij}]$. הרכיב ה־ (i,k) במטריצה $A\cdot(C+D)$ הוא

$$\sum_{j=1}^{n} a_{ij}(c_{jk}+d_{jk})\ .$$

הרכיב ה־ (i,k) במטריצה $(A\cdot C)+(A\cdot D)$ הוא

$$\left(\sum_{j=1}^{n} a_{ij}c_{jk}\right)+\left(\sum_{j=1}^{n} a_{ij}d_{jk}\right)\ .$$

אולם בשל כללי החשבון בשדה זה בדיוק אותו סקלר. לכן

$$A\cdot(C+D)=(A\cdot C)+(A\cdot D)\ .$$

בצורה דומה מוכיחים ש־

$$(A+B)\cdot C=(A\cdot C)+(B\cdot C)\ .$$

משי"ל.

טענה 3. יהיו $A\in \mathrm{M}_{m\times n}(F)$, $B\in \mathrm{M}_{n\times p}(F)$ ו־ $c\in F$. אז

$$A(cB)=(cA)B=c(AB)\ .$$

הוכחה. זה משום שלכל זוג אינדקסים (i,k) מתקיים

$$\sum_{j=1}^{n} a_{ij} c b_{jk} = \sum_{j=1}^{n} c a_{ij} b_{jk} \text{ .}$$

מש״ל.

מעתה נקצר בכתיבת כפל מטריצות: נשמיט סוגריים ואת סימן הכפל היכן שאין מקום לבלבול, בדיוק כשם שעשינו עבור סקלרים בשדה.

טענה 4. תהיינה $A = [a_{ij}] \in \mathrm{M}_{m\times n}(F)$ ו־ $B \in \mathrm{M}_{n\times p}(F)$ שתי מטריצות. נסמן ב־ $A_1, \dots, A_m$ את השורות של A, ונסמן ב־ $B_1, \dots, B_n$ את השורות של B. אז לכל i השורה ה־ i במטריצה AB היא

$$A_i B = a_{i1} B_1 + \cdots + a_{in} B_n \text{ .}$$

הוכחה. על פי הגדרת הכפל הסקלר שמופיע במקום k בשורה i במטריצה AB הוא $\sum_{j=1}^{n} a_{ij} b_{jk}$. זהו בדיוק הסקלר שמופיע במקום k במטריצה $A_i B$. מש״ל.

באופן גראפי זה נראה כך:

$$\begin{bmatrix} * & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ * & * & \cdots & * \end{bmatrix} \cdot \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_n \end{bmatrix} = \begin{bmatrix} * \\ \vdots \\ a_{i1}B_1 + a_{i2}B_2 + \cdots + a_{in}B_n \\ \vdots \\ * \end{bmatrix}$$

וגם כך:

$$\begin{bmatrix} A_1 \\ \vdots \\ A_n \end{bmatrix} \cdot B = \begin{bmatrix} A_1 B \\ \vdots \\ A_n B \end{bmatrix} \text{ .}$$

דוגמה 5. ניקח $A := \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ ו־ $B := \begin{bmatrix} 7 & 8 \\ 9 & 10 \end{bmatrix}$ מעל $\mathbb{Q}$. אז

$$A_1 \cdot B = \begin{bmatrix} 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 7 & 8 \\ 9 & 10 \end{bmatrix} = \begin{bmatrix} 25 & 28 \end{bmatrix}$$

$$A_2 \cdot B = \begin{bmatrix} 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} 7 & 8 \\ 9 & 10 \end{bmatrix} = \begin{bmatrix} 57 & 64 \end{bmatrix}$$

$$A \cdot B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} 7 & 8 \\ 9 & 10 \end{bmatrix} = \begin{bmatrix} 25 & 28 \\ 57 & 64 \end{bmatrix}$$

כפל מטריצות ומערכות משוואות ליניאריות

עובדה שתשמש אותנו בהמשך היא שניתן לזהות את המרחב F^n של עמודות בגובה n עם מרחב המטריצות $\mathrm{M}_{n\times 1}(F)$. בעזרת זיהוי זה אנו יכולים להגדיר את המכפלה AC עבור $A \in \mathrm{M}_{m\times n}(F)$ ו־ $C \in F^n$.

כעת נניח שנתונה מטריצה $A = [a_{ij}]$ כנ״ל ווקטור $B = [b_i] \in F^m$. יהי $C = [c_j] \in F^n$. לפי הגדרת כפל מטריצות, הוקטור C מקיים את המשוואה הוקטורית $AC = B$ אם״ם מתקיימות m המשוואות הסקלריות הבאות:

$$\begin{aligned} a_{11}c_1 + \cdots + a_{1n}c_n = &\quad b_1 \\ \vdots \qquad\qquad \vdots \quad & \quad \vdots \\ a_{m1}c_1 + \cdots + a_{mn}c_n = &\quad b_m \end{aligned}$$

לכן המשוואה $AX = B$, כאשר X הוא משתנה המקבל ערכים ב־ F^n, שקולה למערכת המשוואות ב־ n משתנים סקלריים $AX = B$ שהגדרנו בתחילת הקורס! יתר על כן, ניתן לעבור מהצורה הוקטורית של המשוואה למערכת המשוואות הסקלריות ע״י כך שנרשום $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, כלומר נחשוב על X כעל ״עמודה של משתנים סקלריים״.

כאשר מערכת המשוואות היא הומוגנית וקטור הקבועים הוא

$$. B = O = \vec{0} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \in F^m$$

הסימן O הוא קיצור של $O_{m\times 1}$, מטריצת האפס בגודל $m \times 1$.

הדבר הבא שנעשה יהיה לתאר את פעולות השורה האלמנטריות ככפל מטריצות מסוג מסוים.

הגדרה 2. מטריצה אלמנטרית בגודל $n \times n$ מעל השדה F היא מטריצה E המתקבלת מהמטריצה $I_{n\times n}$ ע״י פעולת שורה אלמנטרית אחת.

דוגמה 6. אם $n = 2$ יש 5 משפחות של מטריצות אלמנטריות.

- מטריצות מהצורה $\begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix}$ כאשר $c \in F$, $c \neq 0$. הן מתאימות לפעולה $cL_1 \to L_1$.
- מטריצות מהצורה $\begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix}$ כאשר $c \in F$, $c \neq 0$. הן מתאימות לפעולה $cL_2 \to L_2$.
- מטריצות מהצורה $\begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix}$ כאשר $c \in F$. הן מתאימות לפעולה $L_1 + cL_2 \to L_1$.
- מטריצות מהצורה $\begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix}$ כאשר $c \in F$. הן מתאימות לפעולה $L_2 + cL_1 \to L_2$.
- המטריצה $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, שמתאימה לפעולה $L_1 \leftrightarrow L_2$.

משפט 1. תהי e פעולת שורה אלמנטרית ותהי $E := e(I_{m\times m})$ המטריצה האלמנטרית המתאימה. אז לכל מטריצה A בגודל $m \times n$ מתקיים $e(A) = EA$.

הוכחה. נרשום $E = [e_{ij}]$, ותהיינה $L_1, \ldots, L_m$ השורות של A. בעזרת טענה 4 רואים ש־

$$EA = E \cdot \begin{bmatrix} L_1 \\ \vdots \\ L_m \end{bmatrix} = \begin{bmatrix} \vdots \\ e_{i1}L_1 + \cdots + e_{im}L_m \\ \vdots \end{bmatrix} .$$

כרגיל נפריד בין שלושה מקרים לפי סוג הפעולה e.

1) המקרה $e = (cL_p \to L_p)$. כאן $e_{pp} = c$, $e_{ii} = 1$ עבור $i \neq p$, ויתר הרכיבים הם 0. לכן שורה p ב־ EA היא cL_p, ויתר השורות זהות לאלו של A. רואים כי $EA = e(A)$.

2) המקרה $e = (L_p + cL_q \to L_p)$. כאן $e_{ii} = 1$ לכל i, $e_{pq} = c$, ויתר הרכיבים הם 0. לכן שורה p ב־ EA היא $L_p + cL_q$, ויתר השורות זהות לאלו של A. רואים כי $EA = e(A)$.

3) המקרה $e = (L_p \leftrightarrow L_p)$. כאן $e_{pq} = e_{qp} = 1$, $e_{ii} = 1$ עבור $i \neq p, q$, ויתר הרכיבים הם 0. לכן שורה p ב־ EA היא L_q, שורה q ב־ EA היא L_p, ויתר השורות זהות לאלו של A. רואים כי $EA = e(A)$. מש״ל.

דוגמה 7. השדה הוא $\mathbb{Q}$. נתחיל מהמטריצה $A := \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ ונדרג אותה ע״י כמה פעולות שורה אלמנטריות:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \xrightarrow{e_1} \begin{bmatrix} 1 & 2 \\ 0 & -2 \end{bmatrix} \xrightarrow{e_2} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \xrightarrow{e_3} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

כאשר

$$\begin{aligned} e_1 &= (L_2 - 3L_1 \to L_2) \\ e_2 &= (-\tfrac{1}{2}L_2 \to L_2) \\ e_3 &= (L_1 - 2L_2 \to L_1) \end{aligned} .$$

המטריצות האלמנטריות $E_i := e_i(I_{2\times 2})$ הן

$$E_1 = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} .$$

מקבלים

$$E_3E_2E_1 = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$$

ו־

$$E_3E_2E_1A = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

כצפוי.

מסקנה 1. תהיינה A ו־ B מטריצות בגודל $m \times n$ מעל שדה F. התנאים הבאים שקולים:

א. A ו־ B הן שקולות שורה.

ב. קיימות מטריצות אלמנטריות $E_1, \cdots, E_s$ בגודל $m \times m$ כך ש־
$B = E_s \cdots E_2 E_1 A$.

הוכחה. א ⇐ ב: בהינתן סדרת פעולת שורה אלמנטריות

$$A = A_0 \xrightarrow{e_1} A_1 \xrightarrow{e_2} \cdots \xrightarrow{e_s} A_s = B$$

נגדיר מטריצות $E_i := e_i(I_{m\times m})$. ע״פ המשפט $A_i = e_i(A_{i-1}) = E_i A_{i-1}$. משום כך

$$B = A_s = E_s \cdots E_1 A_0 = E_s \cdots E_1 A\ .$$

ב ⇐ א: נתון כי $B = E_s \cdots E_2 E_1 A$. תהי e_i הפעולה האלמנטרית כך ש־ $E_i = e_i(I_{m\times m})$. אז ע״פ המשפט המטריצות $A_i := E_i \cdots E_2 E_1 A$ מקיימות $A_i = e_i(A_{i-1})$, ולכן

$$B = e_s(\cdots e_2(e_1(A))\cdots)\ .$$

מש״ל.

מטריצות הפיכות

טענה 5. נתונה מטריצה $A \in \mathrm{M}_{n\times n}(F)$. אם שתי המטריצות $B, C \in \mathrm{M}_{n\times n}(F)$ מקיימות $AB = BA = I_{n\times n}$ וגם $AC = CA = I_{n\times n}$ אז בהכרח $B = C$.

הוכחה. לפי הטריק המוכר:

$$B = I_{n\times n}B = (CA)B = C(AB) = CI_{n\times n} = C\ .$$

מש״ל.

הגדרה 3. מטריצה A בגודל $n \times n$ מעל שדה F תיקרא **הפיכה** אם קיימת מטריצה B מאותו גודל מעל F כך ש־ $AB = BA = I_{n\times n}$. המטריצה B תיקרא **ההופכית** של A, והיא תסומן A^{-1}.

בהמשך לרוב נקצר ונרשום I במקום $I_{n\times n}$, כאשר גודל המטריצה ברור מן ההקשר.

טענה 6.

א. אם A הפיכה אז גם A^{-1} הפיכה, ומתקיים $(A^{-1})^{-1} = A$.

ב. אם $A, B \in \mathrm{M}_{n\times n}(F)$ הפיכות אז גם AB הפיכה, ומתקיים $(AB)^{-1} = B^{-1}A^{-1}$.

הוכחה. א. מיידי.

ב. החישובים הם

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I$$

ו־

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = A^{-1}A = I\ .$$

מש״ל.

כיצד נזהה אם מטריצה A היא הפיכה? איך מחשבים את A^{-1}? אחת הדרכים היא בעזרת דרוג.

טענה 7. מטריצה אלמנטרית E היא הפיכה, וגם E^{-1} היא אלמנטרית.

הוכחה. נניח ש־ $E = e(I_{n\times n})$. תהי f הפעולה האלמנטרית ההפוכה ל־ e, ונגדיר $F := f(I_{n\times n})$. אז ע״פ משפט 1, עם $I := I_{n\times n}$, מתקיים

$$EF = EFI = e(f(I)) = I$$

ו־

$$FE = FEI = f(e(I)) = I\ .$$

מש״ל.

משפט 2. התנאים הבאים שקולים עבור מטריצה $A \in \mathrm{M}_{n\times n}(F)$.

א. המטריצה A הפיכה.
ב. למשוואה $AX = O$ אין פתרון לא טריוויאלי ב־ F^n.
ג. המטריצה A שקולת שורה ל־ $I_{n\times n}$.
ד. המטריצה A היא מכפלה של מטריצות אלמנטריות.

הוכחה. א $\Leftarrow$ ב: נניח כי $C = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \in F^n$ פתרון של $AX = O$. ידוע כי A הפיכה, לכן קיימת המטריצה ההופכית A^{-1}. אז

$$C = IC = (A^{-1}A)C = A^{-1}(AC) = A^{-1}O = O\ .$$

ב $\Leftarrow$ ג: תהי A' מטריצה מדורגת שקולת שורה ל־ A. מאחר שאין פתרון לא טריוויאלי ל־ $AX = O$ הרי אין משתנים חופשיים ב־ A'. לכן יש n 1־מובילים במטריצה A' ו־ $A' = I$.

ג $\Leftarrow$ ד: תהי

$$A = A_0 \xrightarrow{e_1} A_1 \xrightarrow{e_2} A_2 \xrightarrow{e_3} \cdots \xrightarrow{e_s} A_s = I$$

סדרת פעולות שורה אלמנטריות. נגדיר $E_i := e_i(I)$. ע״פ משפט 1 ידוע ש־ $I = E_s \cdots E_2 E_1 A$. לכן בעזרת טענה 7 אנו מקבלים

$$A = E_1^{-1} E_2^{-1} \cdots E_s^{-1} I = E_1^{-1} E_2^{-1} \cdots E_s^{-1}\ ,$$

וכל המטריצות E_i^{-1} הן אלמנטריות.

ד $\Leftarrow$ א: נובע מטענות 7 ו־ 6(ב). מש״ל.

מן המשפט הקודם אפשר לקבל שיטה לחישוב המטריצה ההופכית. בהוכחה של ״ג $\Leftarrow$ ד״ מקבלים את הנוסחה

$$A = E_1^{-1} E_2^{-1} \cdots E_s^{-1}\ .$$

משום כך

$$A^{-1} = E_s \cdots E_2 E_1 = e_s(\cdots e_2(e_1(I))\cdots)\ .$$

אלגוריתם 1. (חישוב מטריצה הופכית) נתונה מטריצה A בגודל $n \times n$ מעל שדה F.

1. נדרג את A ע״י סדרת פעולות שורה אלמנטריות $e_1, \dots, e_s$ לקבלת מטריצה מדורגת $A' = e_s(\cdots e_1(A) \cdots)$.
2. אם $A' \neq I_{n\times n}$ אז A איננה הפיכה.
3. אם $A' = I_{n\times n}$ אז A הפיכה ו־ $A^{-1} = e_s(\cdots e_1(I) \cdots)$.

דוגמה 8. ננסה להפוך את המטריצה $A := \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \in \mathrm{M}_{2\times 2}(\mathbb{Q})$. אנו נפעיל את פעולות השורה בו־זמנית על A ועל I.

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right] \xrightarrow{e_1=(L_2-3L_1\to L_2)} \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array}\right]$$

$$\xrightarrow{e_2=(-\frac{1}{2}L_2\to L_2)} \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array}\right]$$

$$\xrightarrow{e_3=(L_1-2L_2\to L_1)} \left[\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array}\right]$$

קיבלנו $A^{-1} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$.

רצוי לבדוק את התוצאה.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

ו־

$$\begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

הערה: מטעמי חיסכון בעבודה לעתים רצוי להשאיר את ״אנטי־דרוג״ המטריצה I עד לשלב בו A דורגה...

משוואות ליניאריות במטריצות

טענה 8. תהיינה $A \in \mathrm{M}_{m\times n}(F)$ ו־ $D \in \mathrm{M}_{n\times p}(F)$. נרשום $A = \begin{bmatrix} A_1 & \cdots & A_n \end{bmatrix}$ ו־ $D = [d_{jk}] = \begin{bmatrix} D_1 & \cdots & D_p \end{bmatrix}$, כלומר A_i היא העמודה ה־ i של A וכו׳. אז העמודה ה־ k במטריצה AD היא

$$AD_k = d_{1k}A_1 + \cdots + d_{nk}A_n \,.$$

לכן

$$AD = \begin{bmatrix} AD_1 & \cdots & AD_p \end{bmatrix}.$$

הוכחה. כמו טענה 4. משל.

באופן גראפי זה נראה כך:

$$\begin{bmatrix} A_1 & \cdots & A_n \end{bmatrix} \cdot \begin{bmatrix} * & \cdots & d_{1k} & \cdots & * \\ \vdots & & \vdots & & \vdots \\ * & \cdots & d_{nk} & \cdots & * \end{bmatrix} = \begin{bmatrix} * & \cdots & d_{1k}A_1 + \cdots + d_{nk}A_n & \cdots & * \end{bmatrix}.$$

דוגמה 9. נתונה המשוואה

$$(*) \qquad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} X = \begin{bmatrix} 8 & 9 \\ 10 & 11 \end{bmatrix}$$

כאשר $X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{bmatrix}$ משתנה המקבל ערכים ב־ $\mathrm{M}_{3\times 2}(\mathbb{Q})$. כיצד פותרים את המשוואה?

נשתמש בטענה 8 לפתרון הבעיה. ע״פ הטענה המשוואה המטריציאלית $(*)$ שקולה למערכת המשוואות הוקטוריות

$$(**) \qquad \begin{cases} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} X_1 = \begin{bmatrix} 8 \\ 10 \end{bmatrix} \\ \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} X_2 = \begin{bmatrix} 9 \\ 11 \end{bmatrix} \end{cases}$$

כאשר $X_1 = \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix}$ ו־ $X_2 = \begin{bmatrix} x_{12} \\ x_{22} \\ x_{32} \end{bmatrix}$ משתנים המקבלים ערכים ב־ $\mathbb{Q}^3$. הפתרון עכשיו מוכר. נשים לב שניתן בבת אחת לדרג את המטריצה המורחבת בגודל 2×5:

$$\left[\begin{array}{ccc|cc} 1 & 2 & 3 & 8 & 9 \\ 4 & 5 & 6 & 10 & 11 \end{array}\right] \to \left[\begin{array}{ccc|cc} 1 & 0 & -1 & -\frac{20}{3} & -\frac{23}{3} \\ 0 & 1 & 2 & \frac{22}{3} & \frac{25}{3} \end{array}\right]$$

(זהירות: אלמלא היו שני 1־ים מובילים משמאל לקו האנכי היה צריך להפריד את שני המקרים X_1 ו־ X_2.)

תחילה נפתור עבור המשתנה הוקטורי $X_1 = \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix}$. יש משתנה סקלרי חופשי אחד x_{31}. קבוצת הפתרונות היא

$$\left\{ \begin{bmatrix} -\frac{20}{3} \\ \frac{22}{3} \\ 0 \end{bmatrix} + c_1 \cdot \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \;\middle|\; c_1 \in \mathbb{Q} \right\}.$$

כעת נפתור עבור X_2. המשתנה החופשי הוא x_{32}. הפתרונות ההומוגניים הם ללא שינוי, אבל הפתרון המסוים אחר:

$$\left\{ \begin{bmatrix} -\frac{23}{3} \\ \frac{25}{3} \\ 0 \end{bmatrix} + c_2 \cdot \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \;\middle|\; c_2 \in \mathbb{Q} \right\}.$$

לסיום קבוצת הפתרונות למשוואה $(*)$ היא

$$\left\{ \begin{bmatrix} -\frac{20}{3} & -\frac{23}{3} \\ \frac{22}{3} & \frac{25}{3} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} c_1 & c_2 \\ -2c_1 & -2c_2 \\ c_1 & c_2 \end{bmatrix} \;\middle|\; c_1, c_2 \in \mathbb{Q} \right\}.$$

דוגמה 10. נתונה המשוואה

$$AX + B = C$$

עבור מטריצות $A \in \mathrm{M}_{m\times m}(F)$ ו־ $B, C \in \mathrm{M}_{m\times n}(F)$. המשתנה X מקבל ערכים ב־ $\mathrm{M}_{m\times n}(F)$. ידוע ש־ A הפיכה. אז יש פתרון יחיד למשוואה והוא

$$X := A^{-1}(C - B) .$$

מרחב העמודות והגרעין של מטריצה

הגדרה 4. תהי $A \in \mathrm{M}_{m\times n}(F)$. מרחב הפתרונות ב־ F^n של המשוואה ההומוגנית $AX = O$ הוא תת מרחב של F^n. אנו נקרא למרחב זה בשם **הגרעין** של A ונסמן אותו $\mathrm{Ker}(A)$.

הגדרה 5. תהי $A \in \mathrm{M}_{m\times n}(F)$. **מרחב העמודות** של A הוא המרחב הנפרש ע״י עמודותיה של A. זהו תת מרחב של F^m אשר נסמנו ב־ $\mathrm{Im}(A)$. שם אחר למרחב העמודות הוא **התמונה של** A.

בצורה אחרת,

$$\mathrm{Ker}(A) = \{v \in F^n \mid Av = \vec{0}\}$$

וכן, בעזרת טענה 8,

$$\mathrm{Im}(A) = \{Av \mid v \in F^n\} .$$

נזכיר כי הדרגה $\mathrm{rank}(A')$ של מטריצה מדורגת A' הוא מספר ה־ 1 ־ים המובילים.

משפט 3. תהי $A \in \mathrm{M}_{m\times n}(F)$, ותהי A' מטריצה מדורגת שקולת־שורה ל־ A עם דרגה $\mathrm{rank}(A') = r$. אז

$$\dim(\mathrm{Im}(A)) = r$$

ו־

$$\dim(\mathrm{Ker}(A)) = n - r .$$

הוכחה. בפתרון מערכת המשוואות ההומוגנית $AX = O$ אנו מוצאים בסיס $(w_1, \ldots, w_{n-r})$ למרחב הפתרונות, בהתאמה עם המשתנים החפשיים $x_{l_1}, \ldots, x_{l_{n-r}}$. ראה משפט 5 בפרק ג׳. אנו מסיקים ש־ $\dim(\mathrm{Ker}(A)) = n - r$.

כעת יהיו $x_{k_1}, \ldots, x_{k_r}$ המשתנים התלויים ותהיינה $v_1, \ldots, v_n$ העמודות של A. משפט 4 בפרק ג׳ אומר שסדרת העמודות $(v_{k_1}, \ldots, v_{k_r})$ היא בסיס של $\mathrm{Im}(A)$. לכן $\dim(\mathrm{Im}(A)) = r$. מש״ל.

הערה. בדרך כלל $\mathrm{Im}(A') \neq \mathrm{Im}(A)$, כלומר מרחב העמודות איננו נשמר ע״י פעולות שורה!

כעת ניתן להגדיר דרגה של מטריצה כלשהי, באופן שמתיישב עם ההגדרה הקודמת עבור מטריצה מדורגת.

הגדרה 6. תהי $A \in \mathrm{M}_{m\times n}(F)$. **הדרגה** של A היא

$$\operatorname{rank}(A) := \dim(\operatorname{Im}(A)) \ .$$

בעזרת המושג של דרגה ניתן לאפיין מטריצה הפיכה בצורות נוספות.

משפט 4. תהי $A \in \mathrm{M}_{n\times n}(F)$. התנאים הבאים שקולים.
א. A מטריצה הפיכה.
ב. קיימת מטריצה $B \in \mathrm{M}_{n\times n}(F)$ כך ש־ $AB = I_{n\times n}$.
ג. לכל $v \in F^n$ קיים פתרון ב־ F^n למשוואה $AX = v$.
ד. הדרגה של A היא n.
ה. למשוואה $AX = O$ אין פתרון לא טריוויאלי ב־ F^n.

הוכחה. א $\Leftarrow$ ב: ניקח $B := A^{-1}$.

ב $\Leftarrow$ ג: הוקטור $w := Bv$ מקיים

$$Aw = ABv = I_{n\times n}v = v \ .$$

ג $\Leftarrow$ ד: מתנאי ג׳ נובע שכל וקטור $v \in F^n$ הוא צרוף ליניארי של העמודות של A, כלומר $\operatorname{Im}(A) = F^n$. לכן

$$\operatorname{rank}(A) = \dim(\operatorname{Im}(A)) = n \ .$$

ד $\Leftarrow$ ה: תהיינה $v_1, \ldots, v_n$ העמודות של A. מאחר ש־ $\operatorname{rank}(A) = n$ הרי

$$\operatorname{Sp}(v_1, \ldots, v_n) = \operatorname{Im}(A) = F^n \ .$$

יוצא שהסדרה $(v_1, \ldots, v_n)$ היא בסיס של F^n, ובפרט היא בלתי תלויה ליניארית. לכן הפתרון היחיד למשוואה $AX = O$ הוא הפתרון הטריוויאלי.

ה $\Leftarrow$ א: זה במשפט 2 בפרק זה. משי״ל.

הגדרה 7. תהי A מטריצה בגודל $m \times n$ מעל השדה F. **מרחב השורות** של A הוא תת המרחב של $\mathrm{M}_{n\times n}(F)$ הנפרש ע״י השורות של A.

משפט 5. מימד מרחב השורות של מטריצה A שווה לדרגה שלה.

הוכחה. קל לראות כי פעולות שורה אלמנטריות אינן משנות את מרחב השורות. תהי A' מטריצה מדורגת שקולת־שורה ל־ A. אז מרחבי השורות של שתי המטריצות הם שווים; נקרא למרחב זה W. לפי משפט 3 ב־ A' יש r שורות שונות מאפס, כאשר $r := \operatorname{rank}(A)$. בשל המיקום של ה־ 1 המובילים בשורות של A' שורות אלו מהוות סדרה בלתי תלויה ליניארית, כלומר בסיס של W. משי״ל.

ה. דטרמיננטות

בשיעורים הבאים נלמד על פונקציה חשובה מאוד הנקראת **הדטרמיננטה**. בעצם המדובר באוסף של פונקציות

$$\det : \mathrm{M}_{n\times n}(F) \to F$$

כאשר n הוא מספר שלם חיובי כלשהו. נתחיל במקרים $n = 1, 2$.

דוגמה 1. עבור $n = 1$ מגדירים $\det(\begin{bmatrix} a \end{bmatrix}) := a$.

דוגמה 2. עבור $n = 2$ נרשום $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. במקרה זה ההגדרה היא

$$\det(A) := ad - bc\ .$$

טענה 1. תהי $A \in \mathrm{M}_{2\times 2}(F)$. המטריצה A היא הפיכה אס"ם $\det(A) \neq 0$.

הוכחה. נסמן $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ ו־ $u := \det(A)$. נגדיר מטריצה $B := \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. חישוב קצר מראה כי

$$AB = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} ad - bc & 0 \\ 0 & ad - bc \end{bmatrix} = uI\ .$$

א. אם $u \neq 0$ אז $A(u^{-1}B) = I$, ולכן A הפיכה ע"פ משפט 4 בפרק ד'.
ב. אם $u = 0$ נבחין בין שני מקרים. אם $A = O$ אז ברור ש־ A איננה הפיכה. אם $A \neq O$ אז גם $B \neq O$. נסמן $B = [B_1, B_2]$. אז

$$[AB_1, AB_2] = AB = uI = O\ ,$$

ולכן $AB_1 = AB_2 = O$. מאחר שלפחות אחת מן העמודות B_1 או B_1 שונה מאפס הרי A איננה הפיכה.

משי"ל.

אנו מסיקים:

מסקנה 1. אם $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ הפיכה אז

$$A^{-1} = \frac{1}{\det(A)} \cdot \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\ .$$

הגדרה 1. תהי A מטריצה בגודל $n \times n$ מעל F כאשר $n \geq 2$. המינור ה־ (i, j) של A הוא המטריצה M_{ij} בגודל $(n-1) \times (n-1)$ המתקבלת מ־ A ע"י השמטת שורה i ועמודה j.

דוגמה 3. ניקח $n := 3$ ו־ $A := \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$. אז $M_{12} = \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix}$ ו־ $M_{33} = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}$.

הגדרה 2. תהי A מטריצה בגודל $n \times n$ מעל השדה F. נגדיר את **הדטרמיננטה** $\det(A) \in F$ ברקורסיה על n.

- עבור $n = 1$ המטריצה היא $A := [a]$ ומגדירים $\det(A) := a$.
- עבור $n > 1$ נרשום $A = [a_{ij}]$, ויהי M_{ij} המינור ה־ (i, j) של A. אז

$$\det(A) := \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det(M_{1j}) \ .$$

דוגמה 4. נשווה את ההגדרה האחרונה לדוגמאות 1 ו־ 2. המקרה $n = 1$ ברור. כעת ניקח $A := \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. אז $M_{11} = [d]$, $M_{12} = [c]$, ולכן $\det(M_{11}) = d$ ו־ $\det(M_{12}) = c$. מקבלים

$$\det(A) = a \cdot \det([d]) - b \cdot \det([c]) = ad - bc \ .$$

עכשיו נלמד כמה תכונות של הדטרמיננטה. את הוכחות ניתן למצוא בספרים רבים, למשל "אלגברה ליניארית" של ברמן וקון, פרק V ; בספר "יסודות האלגברה הליניארית" של גולן, פרק ט' (שם הדטרמיננטה מכונה "קוצב") ; או בספר של Hoffman and Kunze.

משפט 1. (פיתוח לפי שורה i) נתונה מטריצה $A = [a_{ij}]$ בגודל $n \times n$ מעל השדה F. אז לכל i מתקיים

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(M_{ij}) \ .$$

נשים לב שהגדרה 2 היא פיתוח לפי שורה 1.

דוגמה 5. ניקח את המטריצה $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 4 & 5 & 6 \end{bmatrix}$ מעל השדה $\mathbb{Q}$. פיתוח לפי שורה 2 של $\det(A)$ נותן

$$0 + 0 + (-1)^5 \cdot \det(\begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}) = 3 \ .$$

פיתוח לפי שורה 1 נותן

$$1 \cdot \det(\begin{bmatrix} 0 & 1 \\ 5 & 6 \end{bmatrix}) - 2 \cdot \det(\begin{bmatrix} 0 & 1 \\ 4 & 6 \end{bmatrix}) + 3 \cdot \det(\begin{bmatrix} 0 & 0 \\ 4 & 5 \end{bmatrix}) = -5 - 2 \cdot (-4) + 3 \cdot 0 = 3 \ .$$

דוגמה 6. נניח כי A מטריצה משולשית עליונה, זאת אומרת

$$A = \begin{bmatrix} a_{11} & * & \dots & * \\ 0 & a_{22} & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}$$

כאשר $*$ מייצג סקלר כלשהו. נראה ש־

$$\det(A) = a_{11}a_{22}\cdots a_{nn} ,$$

כלומר מכפלת אברי האלכסון. נעשה זאת באינדוקציה על n. עבור $n = 1$ זה ברור. עבור $n > 1$ נשתמש בפיתוח לפי שורה n:

$$\det(A) = 0 + \cdots + 0 + (-1)^{2n} a_{nn} \cdot \det(M_{nn}) = a_{nn} \cdot \det(M_{nn}) .$$

אולם גם המינור M_{nn} הוא מטריצה משולשית עליונה, ולכן

$$\det(M_{nn}) = a_{11} \cdots a_{n-1,n-1} .$$

משפט 2. (פיתוח לפי עמודה j) נתונה מטריצה $A = [a_{ij}]$ בגודל $n \times n$ מעל השדה F. אז לכל j מתקיים

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(M_{ij}) .$$

משפט 3. (ליניאריות בעמודה j) בהנתן עמודות

$$A_1, \dots, A_{j-1}, B, C, A_{j+1}, \dots, A_n \in F^n$$

וסקלר d מתקיימים

$$\begin{aligned} &\det([A_1, \dots, A_{j-1}, B + C, A_{j+1}, \dots, A_n]) \\ &\quad = \det([A_1, \dots, A_{j-1}, B, A_{j+1}, \dots, A_n]) \\ &\qquad + \det([A_1, \dots, A_{j-1}, C, A_{j+1}, \dots, A_n]) \end{aligned}$$

ו־

$$\begin{aligned} &\det([A_1, \dots, A_{j-1}, d \cdot B, A_{j+1}, \dots, A_n]) \\ &\quad = d \cdot \det([A_1, \dots, A_{j-1}, B, A_{j+1}, \dots, A_n]) \end{aligned}$$

דוגמה 7. נניח שעמודה j של המטריצה A היא אפס. מהי הדטרמיננטה? ניקח $B = C := \vec{0} \in F^n$. אז כמובן גם $B + C = \vec{0}$. לפי משפט 3 מקבלים

$$\det(A) = \det(A) + \det(A)$$

ולכן $\det(A) = 0$.

הגדרה 3. תהי $A = [a_{ij}]$ מטריצה בגודל $m \times n$. **המטריצה המוחלפת** A^{t} היא המטריצה בגודל $n \times m$ אשר לכל (i, j) הרכיב ה־ (i, j) שלה הוא a_{ji}.

דוגמה 8. ניקח $A := \begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}$. כאן $F = \mathbb{Q}$ ו־ $m = n = 2$. המטריצה המוחלפת היא $A^{\mathrm{t}} = \begin{bmatrix}1 & 3\\2 & 4\end{bmatrix}$.

משפט 4. (מטריצה מוחלפת) הדטרמיננטה של מטריצה מוחלפת לא משתנה:

$$\det(A^{\mathrm{t}}) = \det(A) \ .$$

דוגמה 9. נניח ששורה i של המטריצה A היא אפס. מהי הדטרמיננטה? נתבונן במטריצה המוחלפת A^{t}. מאחר שבמטריצה זו עמודה מס׳ i היא אפס ידוע ש־ $\det(A^{\mathrm{t}}) = 0$. כעת לפי משפט 4 מקבלים $\det(A) = 0$.

משפט 5. (פעולות שורה אלמנטריות) תהי A מטריצה ריבועית ו־ e פעולת שורה אלמנטרית.

1) אם $e = (cL_i \to L_i)$ אז $\det(e(A)) = c \cdot \det(A)$.
2) אם $e = (L_i + cL_j \to L_i)$, כאשר $i \neq j$, אז $\det(e(A)) = \det(A)$.
3) אם $e = (L_i \leftrightarrow L_j)$, כאשר $i \neq j$, אז $\det(e(A)) = -\det(A)$.

דוגמה 10. תהי A מטריצה בגודל $n \times n$. יהיו i ו־ j שני אינדקסים שונים, ונניח כי העמודות מס׳ i ו־ j במטריצה A שוות. מהי $\det(A)$? נגדיר $B := A^{\mathrm{t}}$. לפי משפט 4 ידוע כי $\det(A) = \det(B)$. במטריצה B השורות מס׳ i ו־ j שוות. נפעיל את פעולת השורה $e = (L_i - L_j \to L_i)$ על B. לפי משפט 5 מקבלים $\det(B) = \det(e(B))$. אבל במטריצה $e(B)$ שורה מס׳ i היא אפס, ולכן $\det(e(B)) = 0$. לסיכום קיבלנו $\det(A) = 0$.

דוגמה 11. הרי דרך לחישוב דטרמיננטות בעזרת דרוג. ניקח

$$A := \begin{bmatrix}2 & 3 & -1\\3 & 5 & 2\\1 & -2 & -3\end{bmatrix} \in \mathrm{M}_{3\times 3}(\mathbb{Q}) \ .$$

נפעיל על A את תהליך הדרוג ונרשום את הפעולות האלמנטריות.

$$A = \begin{bmatrix}2 & 3 & -1\\3 & 5 & 2\\1 & -2 & -3\end{bmatrix} \xrightarrow{L_1 \leftrightarrow L_3} \begin{bmatrix}1 & -2 & -3\\3 & 5 & 2\\2 & 3 & -1\end{bmatrix} \xrightarrow{L_2 - 3L_1 \to L_2} \begin{bmatrix}1 & -2 & -3\\0 & 11 & 11\\2 & 3 & -1\end{bmatrix}$$

$$\xrightarrow{L_3 - 2L_1 \to L_3} \begin{bmatrix}1 & -2 & -3\\0 & 11 & 11\\0 & 7 & 5\end{bmatrix} \xrightarrow{\frac{1}{11}L_2 \to L_2} \begin{bmatrix}1 & -2 & -3\\0 & 1 & 1\\0 & 7 & 5\end{bmatrix}$$

$$\xrightarrow{L_3 - 7L_2 \to L_3} \begin{bmatrix}1 & -2 & -3\\0 & 1 & 1\\0 & 0 & -2\end{bmatrix} = B$$

ע״פ משפט 5 ידוע לנו ש־

$$\det(B) = 1 \cdot \tfrac{1}{11} \cdot 1 \cdot 1 \cdot (-1) \cdot \det(A) \ .$$

מצד שני מאחר ש־ B מטריצה משולשית מקבלים $\det(B) = -2$. לכן $\det(A) = 22$.

משפט 6. (כפל מטריצות) בהנתן מטריצות ריבועיות A ו־ B בגודל $n \times n$ מתקיים

$$\det(AB) = \det(A) \cdot \det(B) \ .$$

משפט 7. תהי A מטריצה ריבועית מעל שדה F. אז A הפיכה אס"ם $\det(A) \neq 0$.

הוכחה. אנו נוכיח את המשפט בעזרת התוצאות שציטטנו בלא הוכחה קודם.

א. נניח כי A הפיכה. אז $AA^{-1} = I$. ע"פ משפט 6 מקבלים

$$\det(A) \cdot \det(A^{-1}) = \det(I) \ .$$

מאחר ש־ I מטריצה משולשית, עם 1 על האלכסון, הרי $\det(I) = 1$. לכן $\det(A) \neq 0$.

ב. נניח $\det(A) \neq 0$. תהי A' מטריצה מדורגת אשר מתקבלת מ־ A ע"י סדרת פעולות שורה אלמנטריות

$$A = A_0 \xrightarrow{e_1} A_1 \xrightarrow{e_2} \cdots \xrightarrow{e_s} A_s = A' \ .$$

על פי משפט 5 יש סקלרים $c_1, \ldots, c_s \neq 0$ כך ש־

$$\det(A') = c_s \cdots c_1 \cdot \det(A) \ .$$

לכן $\det(A') \neq 0$. מאחר שכך השורה התחתונה של A' חייבת להיות שונה מאפס. זה מחייב שבכל שורה של A' יש 1 מוביל. רואים ש־ $A' = I_{n \times n}$. ע"פ משפט 2 בפרק ד' המטריצה A היא הפיכה.
מש"ל.

כלל קרמר

ידוע לנו שאם $A \in \mathrm{M}_{n\times n}(F)$ הפיכה אז למשוואה $AX = B$ תמיד יש פתרון יחיד ב־ F^n. כמובן הפתרון הוא $X := A^{-1}B$. להלן נוסחה לפתרון המשתמשת בדטרמיננטות.

משפט 8. (כלל קרמר) תהי $A \in \mathrm{M}_{n\times n}(F)$ מטריצה הפיכה שעמודותיה $A_1, \ldots, A_n$. בהנתן וקטור $B \in F^n$ נגדיר מטריצות

$$\tilde{B}_j := \begin{bmatrix} A_1 & \cdots & A_{j-1} & B & A_{j+1} & \cdots & A_n \end{bmatrix}$$

עבור $j = 1, \ldots, n$. אז הפתרון $D = \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} := A^{-1}B$ למשוואה $AX = B$ מקיים:

$$d_j = \frac{\det(\tilde{B}_j)}{\det(A)} \ .$$

הוכחה. על פי דוגמה 10 אנו יודעים כי לכל j, l מתקיים

$$\det(\begin{bmatrix} A_1 & \cdots & A_{j-1} & A_l & A_{j+1} & \cdots & A_n \end{bmatrix}) = \begin{cases} \det(A) \ ; & l = j \\ 0 \ ; & l \neq j \end{cases} \ .$$

מאחר ש־

$$B = AD = \sum_{l=1}^{n} d_l A_l$$

ותוך שימוש במשפט 3 אני מסיקים

$$\begin{aligned}\det(\tilde{B}_j) &= \det(\begin{bmatrix} A_1 & \cdots & A_{j-1} & B & A_{j+1} & \cdots & A_n \end{bmatrix}) \\ &= \det(\begin{bmatrix} A_1 & \cdots & A_{j-1} & \sum_{l=1}^{n} d_l A_l & A_{j+1} & \cdots & A_n \end{bmatrix}) \\ &= \sum_{l=1}^{n} d_l \cdot \det(\begin{bmatrix} A_1 & \cdots & A_{j-1} & A_l & A_{j+1} & \cdots & A_n \end{bmatrix}) \\ &= d_j \cdot \det(A)\end{aligned}$$

אבל $\det(A) \neq 0$, וע״י חלוקה מקבלים

$$. \; d_j = \det(\tilde{B}_j) \cdot \det(A)^{-1}$$

מש״ל.

בעזרת כלל קרמר ניתן לקבל נוסחה למטריצה ההופכית A^{-1}, המכלילה את הנוסחה למקרה $n=2$ שבמסקנה 1.

משפט 9. תהי A מטריצה הפיכה בגודל $n \times n$. לכל (i,j) יהי M_{ij} המינור ה־ (i,j) של A. נגדיר מטריצה $C = [c_{ij}]$ ע״י הנוסחה

$$. \; c_{ij} := (-1)^{i+j} \frac{\det(M_{ji})}{\det(A)}$$

אז $C = A^{-1}$.

שים לב לסדר האינדקסים בנוסחה: בהגדרת c_{ij} משתמשים במינור M_{ji}.

הוכחה. לכל i מ־ 1 עד n תהי B_i העמודה ה־ i של מטריצת היחידה $I = I_{n\times n}$. (כלומר בעצם $B_i = \vec{e}_i$.) נסמן ב־ A_j את העמודה ה־ j של המטריצה A. נגדיר מטריצות

$$. \; \tilde{B}_{ij} := \begin{bmatrix} A_1 & \cdots & A_{j-1} & B_i & A_{j+1} & \cdots & A_n \end{bmatrix}$$

בעזרת פיתוח לפי עמודה j מקבלים

$$, \; \det(\tilde{B}_{ij}) = (-1)^{i+j} \det(M_{ij})$$

ולכן

$$. \; \frac{\det(\tilde{B}_{ij})}{\det(A)} = (-1)^{i+j} \frac{\det(M_{ij})}{\det(A)} = c_{ji}$$

כעת נשתמש במשפט 8, האומר כי

$$A^{-1} B_i = \begin{bmatrix} c_{1i} \\ \vdots \\ c_{ni} \end{bmatrix}$$

לכל i. לבסוף

$$. \; A^{-1} = A^{-1} I = A^{-1} \begin{bmatrix} B_1 & \cdots & B_n \end{bmatrix} = \begin{bmatrix} A^{-1}B_1 & \cdots & A^{-1}B_n \end{bmatrix} = C$$

מש״ל.

ו. טרנספורמציות ליניאריות

נתחיל בתזכורת על פונקציות. תהיינה X ו־ Y שתי קבוצות. פונקציה $f : X \to Y$ היא כלל המתאים לכל איבר $x \in X$ איבר יחיד $f(x) \in Y$. הקבוצה X נקראת התחום של f והקבוצה Y נקראת הטווח. הפונקציה f היא **חד־חד־ערכית** (חח״ע) אם לכל שני איברים שונים $x_1, x_2 \in X$ מתקיים $f(x_1) \neq f(x_2)$. הפונקציה f היא **על** אם לכל $y \in Y$ קיים $x \in X$ כך ש־ $y = f(x)$. **התמונה** של f היא הקבוצה

$$\mathrm{Im}(f) := \{f(x) \mid x \in X\} \subset Y \ .$$

נשים לב כי f היא על אם״ם $\mathrm{Im}(f) = Y$.

בהנתן פונקציות $f : X \to Y$ ו־ $g : Y \to Z$ ניתן להרכיב אותן לקבלת הפונקציה $g \circ f : X \to Z$. הכלל הוא $(g \circ f)(x) := g(f(x))$. פונקציה $f : X \to Y$ נקראת **הפיכה** אם ישנה פונקציה $g : Y \to X$ כך ש־ $(g \circ f)(x) = x$ ו־ $(f \circ g)(y) = y$ לכל $x \in X$ ו־ $y \in Y$. אם ישנה כזו g אז היא יחידה, היא נקראת **הפונקציה ההפכית** של f, ומסומנת ב־ f^{-1}. ידוע כי f הפיכה אם״ם היא חח״ע ועל.

הגדרה 1. יהיו V ו־ W מרחבים וקטוריים מעל שדה F. **טרנספורמציה ליניארית** היא פונקציה $T : V \to W$ המקיימת את התנאים הבאים:

א. $T(v_1 + v_2) = T(v_1) + T(v_2)$ לכל $v_1, v_2 \in V$.

ב. $T(av) = aT(v)$ לכל $v \in V$ ו־ $a \in F$.

כאשר $V = W$ בדרך כלל קוראים ל־ $T : V \to V$ **אופרטור ליניארי**.

דוגמה 1. $V = W$ ו־ $I : V \to V$ היא פונקצית הזהות $I(v) = v$. זו טרנספורמציה ליניארית.

דוגמה 2. $O : V \to W$ היא פונקצית האפס $O(v) = \vec{0}$. זו טרנספורמציה ליניארית.

דוגמה 3. $V := F^n$, $W := F^m$ ו־ $A \in \mathrm{M}_{m\times n}(F)$. נגדיר $T : V \to W$ ע״י $T(v) := Av$. אז

$$T(v_1 + v_2) = A(v_1 + v_2) = Av_1 + Av_2 = T(v_1) + T(v_2)$$

ו־

$$T(av) = A(av) = aAv = aT(v) \ .$$

רואים שזו טרנספורמציה ליניארית.

דוגמה 4. $F := \mathbb{R}$ ו־ $T : \mathbb{R}^2 \to \mathbb{R}^2$ היא הפונקציה

$$T(\begin{bmatrix} a \\ b \end{bmatrix}) := \begin{bmatrix} a^3 \\ b \end{bmatrix} \ .$$

זו איננה טרנספורמציה ליניארית משום שלדוגמה

$$T(2 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix}) = \begin{bmatrix} 8 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 2 \\ 0 \end{bmatrix} = 2 \cdot T(\begin{bmatrix} 1 \\ 0 \end{bmatrix}) \ .$$

דוגמה 5. בהנתן פולינום $p(x) = a_0 + a_1 x + \ldots + a_n x^n$ עם מקדמים ממשיים נגדיר את הנגזרת שלו להיות הפולינום

$$\frac{\mathrm{d}p}{\mathrm{d}x} := a_1 + 2a_2 x + \cdots + n a_n x^{n-1} \ .$$

כמובן שהגדרה זו מתלכדת עם הגדרת הנגזרת של הפונקציה $p:\mathbb{R}\to\mathbb{R}$ שלומדים בקורס חדו"א.
כעת נתבונן במרחב הפולינומים $\mathbb{R}[x]$ מעל השדה $\mathbb{R}$. מקבלים פונקציה $D:\mathbb{R}[x]\to\mathbb{R}[x]$ עם הנוסחה

$$D(p) := \frac{\mathrm{d}p}{\mathrm{d}x} \,.$$

כפי שידוע לנו מחשבון דיפרנציאלי, וניתן לבדוק גם ישירות מההגדרה, לכל שני פולינומים p ו־ q ולכל סקלר c מתקיים:

$$D(p+q) = D(p) + D(q)$$

ו־

$$D(cp) = cD(p).$$

לכן D הוא אופרטור ליניארי, אשר נקרא אופרטור הגזירה.

דוגמה 6. שוב השדה הוא $\mathbb{R}$, אבל המרחב הוא

$$V := \{f:\mathbb{R}\to\mathbb{R} \text{ פונקציות רציפות}\} \,.$$

לכל פונקציה $f\in V$ נגדיר פונקציה $T(f)$ ע"י:

$$T(f)(x) := \int_0^x f(t)\,\mathrm{d}t \,.$$

לפי הידוע מחשבון דיפרנציאלי הפונקציה $T(f)$ גם היא רציפה, ומתקיים:

$$T(f+g)(x) = \int_0^x (f(t)+g(t))\,\mathrm{d}t = \int_0^x f(t)\,\mathrm{d}t + \int_0^x g(t)\,\mathrm{d}t = T(f)(x) + T(g)(x)$$

ו־

$$T(cf)(x) = \int_0^x cf(t)\,\mathrm{d}t = c\int_0^x f(t)\,\mathrm{d}t = cT(f)(x)$$

לכל שתי פונקציות $f,g\in V$ ולכל סקלר c. לכן

$$T(f+g) = T(f) + T(g)$$

ו־

$$T(cf) = cT(f) \,,$$

כלומר T אופרטור ליניארי.

דוגמה 7. כעת לדוגמה גיאומטרית. השדה הוא $\mathbb{R}$ והמרחב הוא $V := \mathbb{R}^2$. תהי

$$A := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \,.$$

נגדיר אופרטור $T:V\to V$ ע"י הנוסחה $T(v) := Av$. מהי המשמעות הגיאומטרית של T?
כדי להבין זאת נכליל את הדוגמה קצת. בהנתן מספר ממשי θ נגדיר מטריצה

$$A_\theta := \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

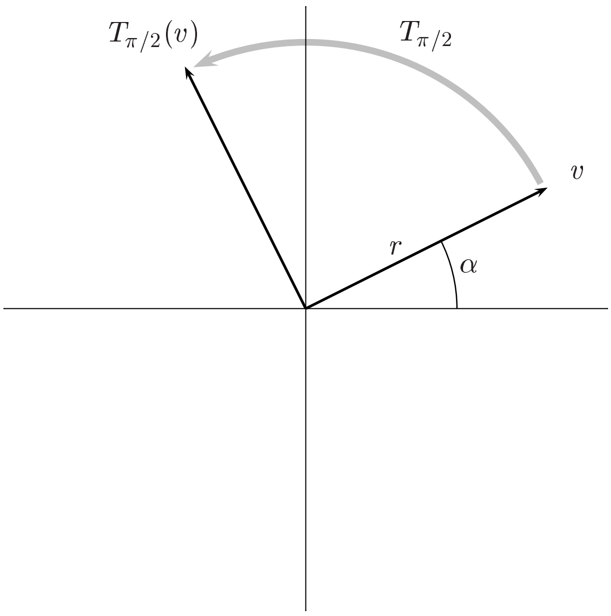
ואופרטור $T_\theta:V\to V$ ע"י הנוסחה $T_\theta(v) := A_\theta v$. נרשום את הוקטור v בקואורדינטות קוטביות:

$$v = \begin{bmatrix} r\cos(\alpha) \\ r\sin(\alpha) \end{bmatrix}$$

כאשר r הוא המרחק מהראשית ו־ α היא הזוית מהציר האופקי. נקבל

$$\begin{aligned}T_\theta(v) &= A_\theta v\\ &= \begin{bmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{bmatrix}\begin{bmatrix}r\cos(\alpha)\\ r\sin(\alpha)\end{bmatrix}\\ &= \begin{bmatrix}r(\cos(\theta)\cos(\alpha)-\sin(\theta)\sin(\alpha))\\ r(\sin(\theta)\cos(\alpha)+\cos(\theta)\sin(\alpha))\end{bmatrix}\\ &= \begin{bmatrix}r\cos(\alpha+\theta)\\ r\sin(\alpha+\theta)\end{bmatrix}\end{aligned}$$

(השתמשנו כאן בזהויות טריגונומטריות). המסקנה היא ש־ T_θ הוא סיבוב בזוית θ. בפרט האופרטור שהתחלנו איתו $T = T_{\pi/2}$ הינו סיבוב בזוית של $\pi/2$.



הגדרה 2. תהי $T : V \to W$ טרנספורמציה ליניארית. **הגרעין** של הטרנספורמציה T הוא הקבוצה

$$\mathrm{Ker}(T) := \{v \in V \mid T(v) = \vec{0}\} \subset V\ .$$

התמונה של הטרנספורמציה T היא הקבוצה

$$\mathrm{Im}(T) := \{T(v) \mid v \in V\} \subset W\ .$$

נשים לב שמושג התמונה מוגדר לכל פונקציה (ראה תחילת הפרק), אולם הגרעין מוגדר אך ורק לטרנספורמציה ליניארית.

טענה 1. תהי $T : V \to W$ טרנספורמציה ליניארית.

א. $\mathrm{Ker}(T)$ הינו תת־מרחב של V.
ב. $\mathrm{Im}(T)$ הינו תת־מרחב של W.

הוכחה. א. נראה כי $\vec{0} \in T$:
$$T(\vec{0}) = T(0 \cdot \vec{0}) = 0 \cdot T(\vec{0}) = \vec{0} .$$
(ראה הערה אחרי ההוכחה.)
כעת יהיו $v_1, v_2 \in \mathrm{Ker}(T)$ ו־ $a \in F$. אז:
$$T(v_1 + v_2) = T(v_1) + T(v_2) = \vec{0} + \vec{0} = \vec{0}$$
ו־
$$T(a \cdot v_1) = a \cdot T(v_1) = a \cdot \vec{0} = \vec{0} .$$
כלומר $v_1 + v_2 \in \mathrm{Ker}(T)$ וגם $a \cdot v_1 \in \mathrm{Ker}(T)$. ז״א $\mathrm{Ker}(T)$ תת־מרחב של V.

ב. מאחר ש־ $T(\vec{0}) = \vec{0}$ הרי $\vec{0} \in \mathrm{Im}(T)$. כעת יהיו $w_2, w_2 \in \mathrm{Im}(T)$ ו־ $a \in F$. לפי הגדרת $\mathrm{Im}(T)$ קיימים $v_1, v_2 \in V$ כך ש־ $w_1 = T(v_1)$ ו־ $w_2 = T(v_2)$. מקבלים
$$T(v_1 + v_2) = T(v_1) + T(v_2) = w_1 + w_2 ,$$
כלומר $w_1 + w_2 \in \mathrm{Im}(T)$. כמו כן
$$T(a \cdot v_1) = a \cdot T(v_1) = a \cdot w_1 ,$$
כלומר $a \cdot w_1 \in \mathrm{Im}(T)$. לכן $\mathrm{Im}(T)$ תת־מרחב של W. מש״ל.

הערה. הסימן $\vec{0}$ בהוכחה הקודמת מייצג שני וקטורים שונים: הן את $\vec{0} \in V$ והן את $\vec{0} \in W$. יכולנו לבדל ביניהם ע״י כך שנכתוב $T(\vec{0}_V) = \vec{0}_W$, אולם זה מסורבל למדי ולא נעשה זאת.

טענה 2. טרנספורמציה ליניארית $T : V \to W$ הינה חד־חד ערכית אס״ם $\mathrm{Ker}(T) = \{\vec{0}\}$.

הוכחה. תחילה נניח כי T הינה חח״ע. מאחר ש־ $T(\vec{0}) = \vec{0}$ נובע כי
$$\mathrm{Ker}(T) = \{v \mid T(v) = \vec{0}\} = \vec{0} .$$
לכיוון השני, נניח כי $\mathrm{Ker}(T) = \{\vec{0}\}$. יהיו $v_1, v_2 \in V$ כך ש־ $T(v_1) = T(v_2)$. אז מתקיים
$$T(v_1 - v_2) = T(v_1) - T(v_2) = \vec{0}$$
ולכן $v_1 - v_2 \in \mathrm{Ker}(T)$. מכאן נובע כי $v_1 - v_2 = \vec{0}$ ו־ $v_1 = v_2$. אם כן T חח״ע. מש״ל.

טענה 3. תהיינה $S, T : V \to W$ טרנספורמציות ליניאריות. נגדיר פונקציה
$$S + T : V \to W$$
ע״י הנוסחה
$$(S + T)(v) := S(v) + T(v) .$$
אז $S + T$ היא טרנספורמציה ליניארית.

הוכחה. עבור $v_1, v_2 \in V$ ו־ $a_1, a_2 \in F$ מקבלים
$$\begin{aligned}(S + T)(a_1 v_1 + a_2 v_2) &= S(a_1 v_1 + a_2 v_2) + T(a_1 v_1 + a_2 v_2) \\ &= a_1 S(v_1) + a_2 S(v_2) + a_1 T(v_1) + a_2 T(v_2) \\ &= a_1 (S(v_1) + T(v_1)) + a_2 (S(v_2) + T(v_2)) \\ &= a_1 (S + T)(v_1) + a_2 (S + T)(v_2)\end{aligned}$$

מש״ל.

טענה 4. תהיינה $S : V \to W$ ו־ $T : W \to U$ טרנספורמציות ליניאריות. אז ההרכבה

$$T \circ S : V \to U$$

היא טרנספורמציה ליניארית.

הוכחה. עבור $v_1, v_2 \in V$ ו־ $a_1, a_2 \in F$ מקבלים

$$\begin{aligned}(T \circ S)(a_1v_1 + a_2v_2) &= T(S(a_1v_1 + a_2v_2)) \\ &= T(a_1S(v_1) + a_2S(v_2)) \\ &= a_1T(S(v_1)) + a_2T(S(v_2)) \\ &= a_1(T \circ S)(v_1) + a_2(T \circ S)(v_2)\end{aligned}$$

מש״ל.

טענה 5. תהי $T : V \to W$ טרנספורמציה ליניארית ו־ $a \in F$. נגדיר פונקציה

$$aT : V \to W$$

ע״י הנוסחה

$$(aT)(v) := a \cdot T(v) .$$

אז aT היא טרנספורמציה ליניארית.

הוכחה. עבור $v_1, v_2 \in V$ ו־ $a_1, a_2 \in F$ מקבלים

$$\begin{aligned}(aT)(a_1v_1 + a_2v_2) &= a \cdot T(a_1v_1 + a_2v_2) \\ &= a \cdot (a_1T(v_1) + a_2T(v_2)) \\ &= a_1 \cdot (a \cdot T(v_1)) + a_2 \cdot (a \cdot T(v_2)) \\ &= a_1 \cdot (aT)(v_1) + a_2 \cdot (aT)(v_2)\end{aligned}$$

מש״ל.

מסקנה 1. יהי $f(x) = \sum_{i=0}^{n} a_ix^i$ פולינום עם מקדמים בשדה F ויהי $T : V \to V$ אופרטור ליניארי. נגדיר

$$T^i := \underbrace{T \circ T \circ \cdots \circ T}_{i}$$

עבור $0 < i$, $T^0 := I$ ו־

$$f(T) := \sum_{i=0}^{n} a_iT^i .$$

אז $f(T) : V \to V$ אופרטור ליניארי.

הוכחה. בעזרת אינדוקציה על i וטענה 3 מקבלים ש־ T^i אופרטור ליניארי. לפי טענה 5 הפונקציה a_iT^i היא אופרטור ליניארי. לבסוף בעזרת אינדוקציה על n וטענה 4 מקבלים ש־ $f(T)$ אופרטור ליניארי. מש״ל.

הגדרה 3. טרנספורמציה לינארית $T : V \to W$ נקראת **איזומורפיזם לינארי**, או **איזומורפיזם של מרחבים וקטוריים**, אם היא חד־חד ערכית ועל.

הגדרה 4. שני מרחבים וקטוריים V ו־ W מעל שדה F נקראים **איזומורפיים** אם יש איזומורפיזם לינארי $T : V \to W$.

טענה 6. יהי $T : V \to W$ איזומורפיזם לינארי. אז הפונקציה ההופכית $T^{-1} : W \to V$ גם היא איזומורפיזם לינארי.

הוכחה. ברור כי T^{-1} היא חח"ע ועל, בהיותה הפונקציה ההופכית של T. צריך להוכיח כי T^{-1} היא טרנספורמציה לינארית. יהיו $w_1, w_2 \in W$ ו־ $c \in F$. נגדיר $v_i := T^{-1}(w_i) \in V$ עבור $i = 1, 2$. אז

$$T(v_1 + v_2) = T(v_1) + T(v_2) = w_1 + w_2 \ .$$

נפעיל T^{-1} על משוואה זו ונקבל:

$$T^{-1}(w_1 + w_2) = T^{-1}(T(v_1 + v_2)) = v_1 + v_2 = T^{-1}(w_1) + T^{-1}(w_2) \ .$$

בדומה:

$$T^{-1}(cw_1) = T^{-1}(T(cv_1)) = cv_1 = cT^{-1}(w_1) \ .$$

משי"ל.

משפט 1. יהי $T : V \to W$ איזומורפיזם של מרחבים וקטוריים, ותהי $\mathbf{v} = (v_1, \ldots, v_n)$ סדרת וקטורים ב־ V. נגדיר $\mathbf{w} := (T(v_1), \ldots, T(v_n))$.

א. $\mathbf{v}$ סדרה בת"ל אם"ם $\mathbf{w}$ סדרה בת"ל.
ב. $\mathbf{v}$ סדרה פורשת של V אם"ם $\mathbf{w}$ סדרה פורשת של W.
ג. $\mathbf{v}$ בסיס של V אם"ם $\mathbf{w}$ בסיס של W.

הוכחה. א. נניח $\mathbf{v}$ בת"ל. יהיו $a_1, \ldots, a_n$ סקלרים כך ש־ $\sum_{i=0}^{n} a_i w_i = \vec{0}$. אז

$$\vec{0} = T^{-1}(\sum_{i=0}^{n} a_i w_i) = \sum_{i=0}^{n} a_i v_i \ .$$

לכן $a_1 = \cdots = a_n = 0$, ורואים ש־ $\mathbf{w}$ סדרה בת"ל. הכיוון השני מוכח בצורה בדומה.

ב. נניח ש־ $\mathbf{v}$ פורשת את W. יהי $w \in W$ כלשהו. אז יש $a_1, \ldots, a_n \in F$ כך ש־ $T^{-1}(w) = \sum_{i=0}^{n} a_i v_i$. לכן

$$w = T(\sum_{i=0}^{n} a_i v_i) = \sum_{i=0}^{n} a_i w_i \ .$$

מכך מסיקים ש־ $\mathbf{w}$ פורשת את W. הכיוון השני בדומה.

ג. צירוף של א' ו־ ב'. משי"ל.

מסקנה 2. נניח V ו־ W הם שני מרחבים וקטוריים איזומורפיים מעל השדה F. אם $\dim(V) = n < \infty$ אז גם $\dim(W) = n$.

הוכחה. ניקח בסיס $(v_1,\dots,v_n)$ של V. יהי $T:V\to W$ איזומורפיזם. אז הסדרה $(T(v_1),\dots,T(v_n))$ היא בסיס של W. משל.

יהי V מרחב וקטורי ממימד n מעל שדה F. נבחר בסיס $\mathbf{v}=(v_1,v_2,\dots,v_n)$ ל־ V. ראינו כבר כי ניתן לייצג כל וקטור $v\in V$ בצורה יחידה כצרוף ליניארי של איברי $\mathbf{v}$:

$$v=a_1v_1+a_2v_2+\ldots+a_nv_n$$

כאשר $a_1,a_2,\dots,a_n\in F$. לכן לכל וקטור v מתאים וקטור הקואורדינטות $[v]_{\mathbf{v}}:=\begin{bmatrix}a_1\\ \vdots\\ a_n\end{bmatrix}\in F^n$.

משפט 2. יהי $\mathbf{v}$ בסיס של המרחב הוקטורי ה־n מימדי V מעל השדה F. הפונקציה $T:V\to F^n$ המוגדרת ע"י הנוסחה

$$T(v):=[v]_{\mathbf{v}}$$

היא איזומורפיזם ליניארי.

הוכחה. 1) נראה כי T טרנספורמציה ליניארית. יהיו $u,v\in V$ ו־ $c\in F$. נסמן את וקטורי הקואורדינטות ע"י

$$[v]_{\mathbf{v}}:=\begin{bmatrix}b_1\\ \vdots\\ b_n\end{bmatrix}\quad \text{ו־}\quad [u]_{\mathbf{v}}:=\begin{bmatrix}a_1\\ \vdots\\ a_n\end{bmatrix}.$$

כלומר

$$v=b_1v_1+\cdots+b_nv_n$$

ו־

$$u=a_1v_1+\cdots+a_nv_n.$$

אז

$$\begin{aligned}u+v&=(a_1v_1+\cdots+a_nv_n)+(b_1v_1+\cdots+b_nv_n)\\&=(a_1+b_1)v_1+\cdots+(a_n+b_n)v_n\end{aligned}$$

ולכן

$$T(u+v)=[u+v]_{\mathbf{v}}=\begin{bmatrix}a_1+b_1\\ \vdots\\ a_n+b_n\end{bmatrix}=\begin{bmatrix}a_1\\ \vdots\\ a_n\end{bmatrix}+\begin{bmatrix}b_1\\ \vdots\\ b_n\end{bmatrix}=[u]_{\mathbf{v}}+[v]_{\mathbf{v}}=T(u)+T(v).$$

בדומה עבור כפל בסקלר:

$$cu=c(a_1v_1+\cdots+a_nv_n)=ca_1v_1+\cdots+ca_nv_n$$

ולכן

$$T(cu)=[cu]_{\mathbf{v}}=\begin{bmatrix}ca_1\\ \vdots\\ ca_n\end{bmatrix}=c\begin{bmatrix}a_1\\ \vdots\\ a_n\end{bmatrix}=c[u]_{\mathbf{v}}=cT(u).$$

2) בשלב זה נוכיח כי T היא חח״ע ועל. בהנתן וקטור כלשהו $\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \in F^n$ נתבונן בוקטור

$$v := a_1 v_1 + \cdots + a_n v_n \ .$$

אז

$$T(v) = [v]_{\mathbf{v}} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \ ,$$

כלומר T היא על. לבסוף נניח $T(v) = \vec{0}$. זאת אומרת $[v]_{\mathbf{v}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$, ומשום כך

$$v = 0 \cdot v_1 + \cdots + 0 \cdot v_n = \vec{0} \ .$$

לפי טענה 2, T היא חח״ע. משי״ל.

מהמשפט אשר הוכחנו נובעת מיד המסקנה הבאה:

מסקנה 3. כל מרחב וקטורי V ממימד n מעל השדה F איזומורפי למרחב F^n.

דוגמה 8. יהי V המרחב

$$V := \{\text{פולינומים } p(x) \text{ מעל } \mathbb{Q} \text{ ממעלה } \leq n\} \ .$$

אז V הינו מרחב וקטורי עם בסיס $\mathbf{v} = (1, x, x^2, \ldots, x^n)$, ולכן V איזומורפי למרחב $\mathbb{Q}^{n+1}$.

משפט 3. יהיו V ו־ W מרחבים וקטוריים ממימדים סופיים מעל שדה F, ותהי $T : V \to W$ טרנספורמציה ליניארית. אז

$$\dim(\mathrm{Ker}(T)) + \dim(\mathrm{Im}(T)) = \dim(V) \ .$$

הוכחה. ניקח בסיס $(v_1, \ldots, v_m)$ לתת־מרחב $\mathrm{Ker}(T)$ ונשלים אותו לבסיס

$$(v_1, \ldots, v_m, v_{m+1}, \ldots, v_n)$$

של V. עבור כל j בטווח $m+1, \ldots, n$ נגדיר $w_j := T(v_j)$. מספיק להוכיח כי הסדרה

$$\mathbf{w} := (w_{m+1}, \ldots, w_n)$$

היא בסיס של $\mathrm{Im}(T)$, שהרי אז נקבל

$$\dim(\mathrm{Im}(T)) = n - m = \dim(V) - \dim(\mathrm{Ker}(T)) \ .$$

א. נוכיח כי הסדרה $\mathbf{w}$ פורשת את $\mathrm{Im}(T)$. יהי $w \in \mathrm{Im}(T)$. אז $w = T(v)$ עבור איזה וקטור

$$v = a_1 v_1 + \cdots + a_n v_n \in V \ .$$

אבל $T(v_j)=0$ לכל $j \le m$, ולכן

$$w=T(v)=\sum_{j=1}^{n} a_jT(v_j)=\sum_{j=m}^{n} a_jT(v_j)=\sum_{j=m}^{n} a_jw_j \ .$$

מצאנו כי $\mathrm{Im}(T)=\mathrm{Sp}(\mathbf{w})$.

ב. עתה נוכיח כי $\mathbf{w}$ סדרה בת״ל. יהיו $a_{m+1},\ldots,a_n$ סקלרים כך ש־

$$\sum_{j=m+1}^{n} a_jw_j=\vec{0} \ .$$

נגדיר

$$v:=\sum_{j=m+1}^{n} a_jv_j \in V \ .$$

אז

$$T(v)=\sum_{j=m+1}^{n} a_jw_j=\vec{0} \ ,$$

ולכן $v \in \mathrm{Ker}(T)$. מאחר ש־ $\mathrm{Ker}(T)=\mathrm{Sp}(v_1,\ldots,v_m)$ הרי ישנם סקלרים $a_1,\ldots,a_m$ כך ש־ $v=-\sum_{j=1}^{m} a_jv_l$. מקבלים

$$\vec{0}=-v+v=\sum_{j=1}^{m} a_jv_j+\sum_{j=m+1}^{n} a_jv_j=\sum_{j=1}^{n} a_jv_j \ .$$

אבל $(v_1,\ldots,v_n)$ בסיס של V ומשום כך $a_1=\cdots=a_n=0$. לכן $\mathbf{w}$ סדרה בת״ל. משי״ל.

דוגמה 9. ניקח $V:=F^n$, $W:=F^m$ ו־ $A \in \mathrm{M}_{m\times n}(F)$. נגדיר $T:F^n \to F^m$ ע״י $T(v):=Av$. מספר המשתנים התלויים במערכת המשוואות $AX=O$ הוא $r=\dim(\mathrm{Im}(A))$, ומספר המשתנים החופשיים הוא $n-r=\dim(\mathrm{Ker}(A))$. אכן רואים כי במקרה זה

$$\dim(\mathrm{Im}(T))+\dim(\mathrm{Ker}(T))=r+(n-r)=n=\dim(F^n) \ .$$

המטריצה של טרנספורמציה ביחס לבסיס

יהיו V ו־ W מרחבים וקטוריים ממימדים סופיים מעל שדה F, ונניח כי נתונים לנו בסיס $\mathbf{v}=(v_1,\ldots,v_n)$ של V ובסיס $\mathbf{w}=(w_1,\ldots,w_m)$ של W. תהי $T:V\to W$ טרנספורמציה ליניארית. לכל אינדקס j הוקטור $T(v_j)\in W$ ניתן לכתיבה יחידה כצרוף ליניארי של איברי הבסיס $\mathbf{w}$; כלומר ישנם סקלרים $a_{1j},\ldots,a_{mj}$ יחידים כך ש־

$$T(v_j)=\sum_{i=1}^{m} a_{ij}w_i .$$

במלים אחרות

$$\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}=[T(v_j)]_{\mathbf{w}} \in F^m \ ,$$

וקטור הקואורדינטות של $T(v_j)$ בבסיס $\mathbf{w}$.

הגדרה 5. המטריצה של הטרנספורמציה הליניארית $T : V \to W$ ביחס לבסיסים $\mathbf{v} = (v_1, \dots, v_n)$ של V ו־ $\mathbf{w} = (w_1, \dots, w_m)$ של W היא המטריצה

$$[T]^{\mathbf{v}}_{\mathbf{w}} = [a_{ij}] \in \mathrm{M}_{m \times n}(F)$$

כך ש־ $T(v_j) = \sum_{i=1}^{m} a_{ij} w_i$ לכל j.

נשים לב כי עמודה j של המטריצה $[T]^{\mathbf{v}}_{\mathbf{w}}$ היא בדיוק $[T(v_j)]_{\mathbf{w}}$, כלומר

$$[T]^{\mathbf{v}}_{\mathbf{w}} = \begin{bmatrix} [T(v_1)]_{\mathbf{w}} & \cdots & [T(v_n)]_{\mathbf{w}} \end{bmatrix} .$$

דוגמה 10. ניקח את מרחבי העמודות $V := F^n$ ו־ $W := F^m$, ויהיו $\mathbf{v}$ ו־ $\mathbf{w}$ הבסיסים הסטנדרטיים $\mathbf{v} := (\vec{e}_1, \dots, \vec{e}_n)$ ו־ $\mathbf{w} := (\vec{e}_1, \dots, \vec{e}_m)$. נשים לב שהסימן $\vec{e}_i$ מתייחס הן לוקטור הסטנדרטי ה־ i ב־ F^m והן לזה שב־ F^n, ונשתדל לא להתבלבל ביניהם. תהיה $A \in \mathrm{M}_{m \times n}(F)$ מטריצה. נגדיר כרגיל טרנספורמציה ליניארית $T : V \to W$ ע״י הנוסחה $T(v) := Av$. אז

$$T(\vec{e}_j) = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mm} \end{bmatrix} \cdot \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} = a_{1j}\vec{e}_1 + \cdots + a_{mj}\vec{e}_m = \sum_{j=1}^{m} a_{ij}\vec{e}_i$$

(ראה טענה 8 בפרק ד׳). כלומר

$$[T(\vec{e}_j)]_{\mathbf{w}} = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} ,$$

ואנו מסיקים כי

$$[T]^{\mathbf{v}}_{\mathbf{w}} = A \; .$$

נאמר את זה במלים: מטריצת הייצוג של T ביחס לבסיסים הסטנדרטיים היא המטריצה A שאיתה התחלנו.

דוגמה 11. יהי V המרחב הוקטורי

$$V := \{\text{פולינומים } p(x) \text{ מעל } \mathbb{R} \text{ ממעלה } \leq n\}$$

עבור איזה $n \geq 0$, ויהי $D : V \to V$ האופרטור הליניארי המוגדר ע״י הנוסחה $D(p) := \frac{\mathrm{d}p}{\mathrm{d}x}$. יהי $\mathbf{v} = (x^0, x^1, \dots, x^n)$, שהוא הבסיס הסטנדרטי של מרחב הפולינומים. לכל $j \in \{1, \dots, n\}$ מקבלים $D(x^j) = jx^{j-1}$, וכן $D(x^0) = \vec{0}$. לכן המטריצה המייצגת היא

$$[D]^{\mathbf{v}}_{\mathbf{v}} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 2 & \dots & 0 \\ 0 & 0 & 0 & \dots & n \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

כאשר האינדקסים בשורות ובעמודות רצים מ־ 0 עד n.

איך להשתמש במטריצת הייצוג $[T]^{\mathbf{v}}_{\mathbf{w}}$? אנו נראה בהמשך שניתן להעזר במטריצת הייצוג לחישוב $\mathrm{Ker}(T)$ ו־ $\mathrm{Im}(T)$.

משפט 4. תהי $T : V \to W$ טרנספורמציה ליניארית, יהי $\mathbf{v}$ בסיס ל־ V ויהי $\mathbf{w}$ בסיס ל־ W. אז לכל $v \in V$ מתקיים

$$[T(v)]_{\mathbf{w}} = [T]^{\mathbf{v}}_{\mathbf{w}} \cdot [v]_{\mathbf{v}} \ .$$

הוכחה. נרשום $\mathbf{v} = (v_1, \ldots, v_n)$, $\mathbf{w} = (w_1, \ldots, w_m)$ ו־ $A = [a_{ij}] := [T]^{\mathbf{v}}_{\mathbf{w}}$. ניקח וקטור $v \in V$ כלשהו, ונגדיר $[v]_{\mathbf{v}} := \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$. עפ״י הגדרת $[T]^{\mathbf{v}}_{\mathbf{w}}$ יש השוויון

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i \ .$$

לכן

$$T(v) = T\Big(\sum_{j=1}^{n} b_j v_j\Big) = \sum_{j=1}^{n} b_j T(v_j) = \sum_{j=1}^{n} b_j \sum_{i=1}^{m} a_{ij} w_i = \sum_{i=1}^{m} \Big(\sum_{j=1}^{n} a_{ij} b_j\Big) w_i \ .$$

עתה נגדיר $c_i := \sum_{j=1}^{n} a_{ij} b_j$. אז $T(v) = \sum_{i=1}^{m} c_i w_i$, ומכאן

$$[T(v)]_{\mathbf{w}} = \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} \ .$$

אולם מהגדרת כפל מטריצות מקבלים

$$\begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} = \begin{bmatrix} a_1 & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \ ,$$

ולכן הגענו ל־ $[T(v)]_{\mathbf{w}} = [T]^{\mathbf{v}}_{\mathbf{w}} \cdot [v]_{\mathbf{v}}$. משל.

מסקנה 4. בתנאי משפט 4 נרשום $A := [T]^{\mathbf{v}}_{\mathbf{w}}$.

א. יהי $v \in V$. אז $v \in \mathrm{Ker}(T)$ אס״ם $[v]_{\mathbf{v}} \in \mathrm{Ker}(A)$ (מרחב הפתרונות של $AX = \vec{0}$).
ב. יהי $w \in W$. $w \in \mathrm{Im}(T)$ אס״ם $[w]_{\mathbf{w}} \in \mathrm{Im}(A)$ (מרחב העמודות של A).

הוכחה. א. $T(v) = 0$ אס״ם $[T(v)]_{\mathbf{w}} = 0$ משום שהפונקציה $w \mapsto [w]_{\mathbf{w}}$ היא איזומורפיזם $W \to F^m$. אולם עפ״י המשפט

$$[T(v)]_{\mathbf{w}} = [T]^{\mathbf{v}}_{\mathbf{w}} \cdot [v]_{\mathbf{v}} \ .$$

אגף ימין הוא $\vec{0}$ בדיוק כאשר $[v]_{\mathbf{v}}$ במרחב הפתרונות של המטריצה $A = [T]^{\mathbf{v}}_{\mathbf{w}}$.

ב. נניח $w \in \mathrm{Im}(T)$. אז יש $v \in V$ כך ש־ $w = T(v)$. עפ״י המשפט $[w]_{\mathbf{w}} = [T]^{\mathbf{v}}_{\mathbf{w}} \cdot [v]_{\mathbf{v}}$, ולכן $[w]_{\mathbf{w}}$ במרחב העמודות של המטריצה $A = [T]^{\mathbf{v}}_{\mathbf{w}}$.

לכוון השני, נניח ש־ $[w]_{\mathbf{w}}$ במרחב העמודות של המטריצה A. כלומר קיימת עמודה $u \in F^n$ כך ש־ $[w]_{\mathbf{w}} = [T]^{\mathbf{v}}_{\mathbf{w}} \cdot u$. יהי $v \in V$ הוקטור כך ש־ $u = [v]_{\mathbf{v}}$. אז עפ״י המשפט

$$[T(v)]_{\mathbf{w}} = [T]^{\mathbf{v}}_{\mathbf{w}} \cdot u = [w]_{\mathbf{w}} \ .$$

לכן $w = T(v)$. משל.

דוגמה 12. נחזור לדוגמה 11, שבה השדה הוא $\mathbb{R}$, V הוא מרחב הפולינומים ממעלה $\geq n$, $\mathbf{v} = (1, x, \ldots, x^n)$ הוא הבסיס הסטנדרטי של V ו־ $D : V \to V$ הוא אופרטור הגזירה. המטריצה המייצגת היא כזכור

$$A = [D]^{\mathbf{v}}_{\mathbf{v}} = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 2 & \ldots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & n \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix} \ .$$

רואים מיד שבסיס של $\mathrm{Ker}(A)$ הוא $(\vec{e}_1)$, ובסיס של $\mathrm{Im}(A)$ הוא $(\vec{e}_1, \ldots, \vec{e}_n)$. נתרגם לוקטורים המתאימים ב־ V ונקבל כי הסדרה (x^0) היא בסיס של $\mathrm{Ker}(D)$ ו־ $(x^0, \ldots, x^{n-1})$ היא בסיס של $\mathrm{Im}(D)$.

דוגמה 13. יהיו $F := \mathbb{R}$,

$$V := \{3 \geq \text{פולינומים מעל } \mathbb{R} \text{ במשתנה } x \text{ ממעלה}\}$$

ו־ $W := \mathbb{R}^2$. נגדיר פונקציה $T : V \to W$ ע״י הנוסחה

$$T(p(x)) := \begin{bmatrix} p(0) \\ p(2) \end{bmatrix} \ .$$

האם T טרנספורמציה ליניארית? יהיו $p(x), q(x) \in V$ ו־ $c \in \mathbb{R}$. אז

$$T(p(x) + q(x)) = \begin{bmatrix} p(0) + q(0) \\ p(2) + q(2) \end{bmatrix} = \begin{bmatrix} p(0) \\ p(2) \end{bmatrix} + \begin{bmatrix} q(0) \\ q(2) \end{bmatrix} = T(p(x)) + T(q(x))$$

ו־

$$T(cp(x)) = \begin{bmatrix} cp(0) \\ cp(2) \end{bmatrix} = cT(p(x)) \ .$$

המסקנה היא שאמנם T טרנספורמציה ליניארית.

עתה נחשב את $\mathrm{Ker}(T)$ ו־ $\mathrm{Im}(T)$. לשם כך נבחר בסיסים, ואנו נבחר בבסיסים הסטנדרטיים $\mathbf{v} := (x^0, \ldots, x^3)$ ב־ V ו־ $\mathbf{w} := (\vec{e}_1, \vec{e}_2)$ ב־ W. המטריצה המייצגת היא

$$A := [T]^{\mathbf{v}}_{\mathbf{w}} = \begin{bmatrix} [T(x^0)]_{\mathbf{w}} & \cdots & [T(x^3)]_{\mathbf{w}} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 4 & 8 \end{bmatrix} \ .$$

בסיס למרחב הפתרונות של המשוואה $AX = 0$ הוא

$$\left(\begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -4 \\ 0 \\ 1 \end{bmatrix} \right) ,$$

ובסיס למרחב העמודות של A הוא $\left(\begin{bmatrix}0\\2\end{bmatrix},\begin{bmatrix}1\\1\end{bmatrix}\right)$. לכן בסיס של $\operatorname{Ker}(T)$ הוא הסדרה $(p_1(x),p_2(x))$, כאשר $p_1(x) := -2x+x^2$ ו־ $p_2(x) := -4x+x^3$. בסיס של $\operatorname{Im}(T)$ הוא $\left(\begin{bmatrix}0\\2\end{bmatrix},\begin{bmatrix}1\\1\end{bmatrix}\right)$. בעצם אנו רואים כי $\operatorname{Im}(T)=W$.

לבסוף נעשה בדיקה: $T(p_1(x)) = \begin{bmatrix}0\\0\end{bmatrix}$ ו־ $T(p_2(x)) = \begin{bmatrix}0\\0\end{bmatrix}$, לכן פולינומים אלו הם אכן בגרעין של T; ומאחר שמעלותיהם שונות הם בלתי תלויים ליניארית.

משפט 5. יהיו V, W ו־ U מרחבים וקטוריים סוף מימדיים מעל השדה F, תהיינה $T: V\to W$ ו־ $S: W\to U$ טרנספורמציות ליניאריות, ויהיו $\mathbf{v}$, $\mathbf{w}$ ו־ $\mathbf{u}$ בסיסים של V, W ו־ U בהתאמה. אז

$$[S\circ T]^{\mathbf{v}}_{\mathbf{u}} = [S]^{\mathbf{w}}_{\mathbf{u}}\cdot[T]^{\mathbf{v}}_{\mathbf{w}} \ .$$

הוכחה. נרשום $\mathbf{v}=(v_1,\ldots,v_n)$. אז תוך שימוש כפול במשפט 4 מקבלים

$$[(S\circ T)(v_j)]_{\mathbf{u}} = [S(T(v_j))]_{\mathbf{u}} = [S]^{\mathbf{w}}_{\mathbf{u}}\cdot[T(v_j)]_{\mathbf{w}} = [S]^{\mathbf{w}}_{\mathbf{u}}\cdot[T]^{\mathbf{v}}_{\mathbf{w}}\cdot[v_j]_{\mathbf{v}} \ .$$

נגדיר מטריצה $A := [S]^{\mathbf{w}}_{\mathbf{u}}\cdot[T]^{\mathbf{v}}_{\mathbf{w}}$. מאחר ש־ $[v_j]_{\mathbf{v}} = \vec{e}_j$ הרי

$$[(S\circ T)(v_j)]_{\mathbf{u}} = A\cdot[v_j]_{\mathbf{v}} = A\vec{e}_j \ .$$

לבסוף

$$[(S\circ T)]^{\mathbf{v}}_{\mathbf{u}} = \begin{bmatrix}[(S\circ T)(v_1)]_{\mathbf{u}} & \cdots & [(S\circ T)(v_n)]_{\mathbf{u}}\end{bmatrix} = \begin{bmatrix}A\vec{e}_1 & \cdots & A\vec{e}_n\end{bmatrix} = A \ .$$

מש"ל.

טענה 7. תהינה $S,T: V\to W$ טרנספורמציות ליניאריות, יהי c סקלר ויהיו $\mathbf{v}$ ו־ $\mathbf{w}$ בסיסים של V ו־ W בהתאמה. אז

$$[S+T]^{\mathbf{v}}_{\mathbf{w}} = [S]^{\mathbf{v}}_{\mathbf{w}} + [T]^{\mathbf{v}}_{\mathbf{w}}$$

ו־

$$[cT]^{\mathbf{v}}_{\mathbf{w}} = c[T]^{\mathbf{v}}_{\mathbf{w}} \ .$$

הוכחה. נרשום $\mathbf{v}=(v_1,\ldots,v_n)$. לכל אינדקס j עמודה מס' j במטריצה $[S+T]^{\mathbf{v}}_{\mathbf{w}}$ היא

$$[(S+T)(v_j)]_{\mathbf{w}} = [S(v_j)+T(v_j)]_{\mathbf{w}} \ .$$

כעת לפי משפט 2

$$[S(v_j)+T(v_j)]_{\mathbf{w}} = [S(v_j)]_{\mathbf{w}} + [T(v_j)]_{\mathbf{w}} \ .$$

אולם זוהי עמודה מס' j במטריצה $[S]^{\mathbf{v}}_{\mathbf{w}} + [T]^{\mathbf{v}}_{\mathbf{w}}$.

בדומה מוכיחים את החלק השני. מש"ל.

בהנתן מטריצה $A\in \mathrm{M}_{n\times n}(F)$ ופולינום $f(x)=\sum_{i=0}^m c_i x^i$ עם מקדמים ב־ F נגדיר

$$f(A) := \sum_{i=0}^m c_i A^i \in \mathrm{M}_{n\times n}(F) \ .$$

מסקנה 5. יהי $T : V \to V$ אופרטור ליניארי, יהי $f(x)$ פולינום עם מקדמים ב־ F, יהי $\mathbf{v}$ בסיס של V ותהי $A := [T]_{\mathbf{v}}^{\mathbf{v}}$ מטריצת הייצוג. אז

$$[f(T)]_{\mathbf{v}}^{\mathbf{v}} = f(A)\ .$$

הוכחה. בעזרת משפט 5, טענה 7 ואינדוקציה על m. משי״ל.

שינוי בסיס ומטריצות מעבר

הגדרה 6. נתון מרחב וקטורי V ממימד n מעל השדה F עם שני בסיסים $\mathbf{v}$ ו־ $\mathbf{v}'$. יהי $I : V \to V$ אופרטור הזהות. המטריצה $[I]_{\mathbf{v}'}^{\mathbf{v}} \in \mathrm{M}_{n\times n}(F)$ נקראת **מטריצת המעבר מהבסיס** $\mathbf{v}$ **לבסיס** $\mathbf{v}'$.

טענה 8. יהיו $\mathbf{v}$ ו־ $\mathbf{v}'$ שני בסיסים של המרחב הוקטורי V. אז

$$[I]_{\mathbf{v}}^{\mathbf{v}'} \cdot [I]_{\mathbf{v}'}^{\mathbf{v}} = I_{n\times n}\ .$$

כלומר אם נסמן $P := [I]_{\mathbf{v}'}^{\mathbf{v}}$ אז P היא הפיכה ו־ $P^{-1} = [I]_{\mathbf{v}}^{\mathbf{v}'}$.

הוכחה. נסמן $\mathbf{v} = (v_1, \dots, v_n)$. אז

$$[I]_{\mathbf{v}}^{\mathbf{v}} = \begin{bmatrix}[v_1]_{\mathbf{v}} & \cdots & [v_n]_{\mathbf{v}}\end{bmatrix} = \begin{bmatrix}\vec{e}_1 & \cdots & \vec{e}_n\end{bmatrix} = I_{n\times n}\ .$$

אולם על פי משפט 5 מתקיים

$$[I]_{\mathbf{v}}^{\mathbf{v}'} \cdot [I]_{\mathbf{v}'}^{\mathbf{v}} = [I]_{\mathbf{v}}^{\mathbf{v}}\ .$$

משי״ל.

דוגמה 14. ניקח את המרחב הוקטורי $V := \mathbb{R}^2$ מעל השדה $F := \mathbb{R}$, עם הבסיסים $\mathbf{e} := (\vec{e}_1, \vec{e}_2)$ ו־ $\mathbf{v} := (v_1, v_2)$, כאשר $v_1 := \begin{bmatrix}1\\2\end{bmatrix}$ ו־ $v_2 := \begin{bmatrix}3\\4\end{bmatrix}$. מטריצת המעבר מ־ $\mathbf{v}$ ל־ $\mathbf{e}$ היא

$$P := [I]_{\mathbf{e}}^{\mathbf{v}} = \begin{bmatrix}v_1 & v_2\end{bmatrix} = \begin{bmatrix}1 & 3\\2 & 4\end{bmatrix}\ .$$

לכיוון השני מחשבים ש־ $\vec{e}_1 = -2v_1 + v_2$ ו־ $\vec{e}_2 = \frac{3}{2}v_1 - \frac{1}{2}v_2$. לכן

$$P^{-1} = [I]_{\mathbf{v}}^{\mathbf{e}} = \begin{bmatrix}[\vec{e}_1]_{\mathbf{v}} & [\vec{e}_2]_{\mathbf{v}}\end{bmatrix} = \begin{bmatrix}-2 & \frac{3}{2}\\1 & -\frac{1}{2}\end{bmatrix}\ .$$

הגדרה 7. תהיינה $A, B \in \mathrm{M}_{n\times n}(F)$. A ו־ B נקראות **מטריצות דומות** אם ישנה מטריצה הפיכה $P \in \mathrm{M}_{n\times n}(F)$ כך ש־ $B = PAP^{-1}$.

טענה 9. יהי $T : V \to V$ אופרטור ליניארי, יהיו $\mathbf{v}$ ו־ $\mathbf{v}'$ שני בסיסים של V, ותהי $P := [I]^{\mathbf{v}}_{\mathbf{v}'}$ מטריצת המעבר. אז

$$[T]^{\mathbf{v}'}_{\mathbf{v}'} = P \cdot [T]^{\mathbf{v}}_{\mathbf{v}} \cdot P^{-1} \ .$$

הוכחה. ע״פ משפט 5 (שימוש כפול)

$$[T]^{\mathbf{v}'}_{\mathbf{v}'} = [I \circ T \circ I]^{\mathbf{v}'}_{\mathbf{v}'} = [I]^{\mathbf{v}}_{\mathbf{v}'} \cdot [T \circ I]^{\mathbf{v}'}_{\mathbf{v}} = [I]^{\mathbf{v}}_{\mathbf{v}'} \cdot [T]^{\mathbf{v}}_{\mathbf{v}} \cdot [I]^{\mathbf{v}'}_{\mathbf{v}} = P \cdot [T]^{\mathbf{v}}_{\mathbf{v}} \cdot P^{-1} \ .$$

משל.

דוגמה 15. השדה הוא $F := \mathbb{R}$ והמרחב הוא המישור $V := \mathbb{R}$. האופרטור $T : V \to V$ הוא ״הטלה על ציר x״, כלומר $T(\begin{bmatrix} a \\ b \end{bmatrix}) = \begin{bmatrix} a \\ 0 \end{bmatrix}$. פעולת T על הבסיס הסטנדרטי $\mathbf{e} = (\vec{e}_1, \vec{e}_2)$ היא $T(\vec{e}_1) = \vec{e}_1$ ו־ $T(\vec{e}_2) = 0$; לכן מטריצת הייצוג הינה

$$[T]^{\mathbf{e}}_{\mathbf{e}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \ .$$

נמצא את מטריצת הייצוג ביחס לבסיס אחר. נגדיר $v_1 := \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $v_2 := \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ ו־ $\mathbf{v} := (v_1, v_2)$. הפעולה של T על אברי בסיס זה היא

$$T(v_1) = T(\begin{bmatrix} 1 \\ 2 \end{bmatrix}) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = -2v_1 + v_2$$

ו־

$$T(v_2) = T(\begin{bmatrix} 3 \\ 4 \end{bmatrix}) = \begin{bmatrix} 3 \\ 0 \end{bmatrix} = -6v_1 + 3v_2 \ .$$

לכן

$$[T]^{\mathbf{v}}_{\mathbf{v}} = \begin{bmatrix} -2 & -6 \\ 1 & 3 \end{bmatrix} \ .$$

מטריצות המעבר הן

$$P^{-1} = [I]^{\mathbf{e}}_{\mathbf{v}} = \begin{bmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{bmatrix}$$

ו־

$$P = [I]^{\mathbf{v}}_{\mathbf{e}} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \ .$$

ואמנם

$$P \cdot [T]^{\mathbf{v}}_{\mathbf{v}} \cdot P^{-1} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \cdot \begin{bmatrix} -2 & -6 \\ 1 & 3 \end{bmatrix} \cdot \begin{bmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = [T]^{\mathbf{e}}_{\mathbf{e}} \ .$$

ז. ערכים עצמיים וליכסון אופרטורים

יהי V מרחב וקטורי ממימד n מעל השדה F ויהי $T : V \to V$ אופרטור ליניארי. אנו נראה כי לאופרטור T מתאימים סקלרים $\lambda_1, \lambda_2, \ldots, \lambda_m$, כאשר $m \leq n$, המתארים במידה רבה את התכונות של T. סקלרים אלו יקראו הערכים העצמיים של T.

הגדרה 1. יהי V מרחב וקטורי סוף מימדי מעל השדה F, יהי $T : V \to V$ אופרטור ליניארי, ויהיו $v \in V$ ו־ $\lambda \in F$. נניח ש־ $v \neq \vec{0}$ ומתקיים

$$T(v) = \lambda v\ .$$

אז λ נקרא **ערך עצמי** של T, ו־ v נקרא **וקטור עצמי** של T, השייך לערך העצמי λ.

הערה. נשים לב כי וקטור עצמי v תמיד שונה מ־ $\vec{0}$!

דוגמה 1. ניקח את המרחב $V := F^n$. בהנתן מטריצה $A := \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$ נגדיר אופרטור $T : V \to V$ ע״י $T(v) := Av$. אז לכל i מתקיים

$$T(\vec{e_i}) = A\vec{e_i} = \lambda_i \vec{e_i}\ .$$

לכן λ_i ערך עצמי ו־ $\vec{e_i}$ וקטור עצמי השייך ל־ λ_i.

דוגמה 2. כאן השדה הוא $F := \mathbb{R}$ והמרחב הוא המישור $V := \mathbb{R}^2$. האופרטור T הוא סיבוב בזוית $\frac{\pi}{2}$. נחפש ערכים עצמיים של T. תחילה נחקור את השאלה באופן גיאומטרי. אם $T(v) = \lambda v$ עבור $\lambda \in \mathbb{R}$ הרי $T(v)$ ו־ v הם על אותו ישר, וזה אפשרי רק אם $v = \vec{0}$. לכן אין ל־ T וקטורים עצמיים ולא ערכים עצמיים.

נראה זאת עכשיו באופן אלגברי. כזכור המטריצה של T בבסיס הסטנדרטי היא $A := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

יהי $\vec{0} \neq \begin{bmatrix} a \\ b \end{bmatrix} = v$. אז

$$T(v) = Av = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -b \\ a \end{bmatrix}\ .$$

אם $\begin{bmatrix} \lambda a \\ \lambda b \end{bmatrix} = \lambda v = T(v)$ אז $-b = \lambda a$ ו־ $a = \lambda b$. לכן $a = \lambda b = -\lambda^2 a$ ו־ $-1 = \lambda^2$. אבל אין מספר ממשי λ המקיים $\lambda^2 = -1$. משום כך אין ל־ T ערכים עצמיים.

דוגמה 3. שוב השדה הוא $F := \mathbb{R}$. המרחב הוא

$$V := \{n \geq \text{פולינומים } p(x) \text{ מעל } \mathbb{R} \text{ ממעלה}\}\ .$$

האופרטור D הוא אופרטור הגזירה $D(p) := \frac{\mathrm{d}p}{\mathrm{d}x}$. ניקח $v := x^0 \in V$. אז $D(v) = \vec{0} = 0 \cdot v$. לכן v וקטור עצמי של D השייך לערך העצמי $\lambda = 0$.

כיצד נמצא את הערכים עצמיים של T? ניקח בסיס $\mathbf{v} = (v_1, \ldots, v_n)$ של V, ותהי $A := [T]^{\mathbf{v}}_{\mathbf{v}} \in \mathrm{M}_{n\times n}(F)$ המטריצה המייצגת את T. בהנתן וקטור $v \in V$ נרשום $w := [v]_{\mathbf{v}} \in F^n$.

ידוע כי

$$[T(v)]_{\mathbf{v}} = [T]^{\mathbf{v}}_{\mathbf{v}} \cdot [v]_{\mathbf{v}} = A \cdot w$$

וכן

$$[\lambda v]_{\mathbf{v}} = \lambda [v]_{\mathbf{v}} = \lambda w \ .$$

אנו רואים ש־ $T(v) = \lambda v$ אס״ם $Aw = \lambda w$. כעת $Aw = \lambda w$ אס״ם $(\lambda I - A)w = 0$, כאשר $I \in \mathrm{M}_{n\times n}(F)$ היא מטריצת היחידה. התנאי שיש וקטור $\vec{0} \neq w \in F^n$ כך ש־ $(\lambda I - A)w = \vec{0}$ הוא ש־ $\det(\lambda I - A) = 0$. הוכחנו את המשפט הבא:

משפט 1. יהי T אופרטור ליניארי על מרחב סוף מימדי V, יהי $\mathbf{v}$ בסיס של V, ותהי $A := [T]^{\mathbf{v}}_{\mathbf{v}}$.

א. סקלר $\lambda \in F$ הוא ערך עצמי של T אס״ם $\det(\lambda I - A) = 0$.
ב. נניח ש־ λ הוא ערך עצמי של T. הוקטורים העצמיים של T השייכים ל־ λ הם הוקטורים $v \in V$ כך ש־ $v \neq \vec{0}$ ו־ $[v]_{\mathbf{v}}$ פתרון של המשוואה ההומוגנית $(\lambda I - A)X = O$.

חשבון פולינומים

יהי F שדה. עד כה קבוצת הפולינומים $F[x]$ היתה רק מרחב וקטורי. אולם ניתן גם לכפול פולינומים. בהנתן פולינומים $f(x) = \sum_{i=0}^{m} a_i x^i$ ו־ $g(x) = \sum_{i=0}^{n} b_i x^i$ מכפלתם היא הפולינום

$$(fg)(x) := \left(\sum_{i=0}^{m} a_i x^i\right) \cdot \left(\sum_{j=0}^{n} b_j x^j\right) = \sum_{k=0}^{m+n} \left(\sum_{l=0}^{k} a_l b_{k-l}\right) \cdot x^k \ .$$

כפל פולינומים הוא קומוטטיבי, ודיסטריבוטיבי ביחס לחיבור. הכפל גם מכבד הצבה: לכל סקלר λ מתקיים

$$(fg)(\lambda) = f(\lambda) g(\lambda) \ .$$

אם $f, g \neq 0$ גם מתקיים

$$\deg(f) + \deg(g) = \deg(fg) \ .$$

משפט 2. (חילוק עם שארית) יהי $f(x)$ פולינום כלשהו ויהי $g(x)$ פולינום ממעלה $d \geq 1$. אז ישנם פולינומים יחידים $q(x)$ ו־ $r(x)$ כך ש־

$$f(x) = g(x)q(x) + r(x)$$

ו־

$$\deg(r) \leq d - 1 \ .$$

רעיון ההוכחה: באינדוקציה על מעלת הפולינום $f(x)$. הפולינום $q(x)$ נקרא **המנה** ו־ $r(x)$ נקרא **השארית**.

הרי דוגמה לחילוק פולינומים. האלגוריתם עובד בצורה דומה לחילוק מספרים שלמים גדולים כפי שלומדים בבית הספר היסודי. במקום החזקות של 10 בכתיבה העשרונית של מספר, כאן יש חזקות של המשתנה x.

דוגמה 4. נתון הפולינום $f(x) := x^3 - 1$ מעל $\mathbb{R}$. נחלק את $f(x)$ ב־ $g(x) := x - 1$:

$$\begin{array}{llll|l}
x^2 & +x & +1 & & \\
\hline
x^3 & & & -1 & x-1 \\
\cline{5-5}
x^3 & -x^2 & & & \\
\cline{1-4}
 & x^2 & & -1 & \\
 & x^2 & -x & & \\
\cline{2-4}
 & & x & -1 & \\
 & & x & -1 & \\
\cline{3-4}
 & & & 0 &
\end{array}$$

התוצאה: השארית היא $r(x) := 0$. לכן $f(x) = (x-1)q(x)$, כאשר $q(x) := x^2 + x + 1$.

מסקנה 1. יהי $f(x)$ פולינום כלשהו ויהי λ סקלר. אז $x - \lambda$ מחלק את $f(x)$ אם״ם $f(\lambda) = 0$.

הוכחה. תחילה נניח ש־ $x - \lambda$ מחלק את $f(x)$. זאת אומרת שקיים פולינום $q(x)$ כך ש־

$$f(x) = q(x)(x - \lambda) \ .$$

נציב λ ונקבל

$$f(\lambda) = q(\lambda) \cdot 0 = 0 \ .$$

לכוון השני נניח ש־ $f(\lambda) = 0$. ע״פ משפט החילוק יש $q(x)$ ו־ $r(x)$ כך ש־

$$f(x) = q(x)(x - \lambda) + r(x)$$

ו־ $\deg(r) \leq 0$. לכן $r(x) = \mu \in F$. נציב λ ונקבל

$$0 = f(\lambda) = q(\lambda) \cdot 0 + \mu = \mu \ .$$

לכן

$$f(x) = q(x)(x - \lambda) \ .$$

מש״ל.

מסקנה 2. בהנתן פולינום $f(x)$ ממעלה $n \geq 0$ קיימים סקלרים $\lambda_1, \ldots, \lambda_m \in F$ כך ש־

$$f(x) = (x - \lambda_1) \cdots (x - \lambda_m) \cdot q(x)$$

ו־ $q(\lambda) \neq 0$ לכל $\lambda \in F$.

הוכחה. באינדוקציה על n. אם $n = 0$ הרי ע״פ הגדרה $f(x)$ הוא סקלר שונה מ־ 0, וניקח $m := 0$ ו־ $q(x) := f(x)$.

כעת נניח ש־ $n \geq 1$. אם ל־ $f(x)$ אין שורשים ניקח $m := 0$ ו־ $q(x) := f(x)$. אחרת יהי λ_1 שורש של $f(x)$. ע״פ מסקנה 1 מתקיים $f(x) = (x - \lambda_1)g(x)$ עבור פולינום $g(x)$ ממעלה $n - 1$. מהנחת האינדוקציה מקבלים ש־

$$g(x) = (x - \lambda_2) \cdots (x - \lambda_m) \cdot q(x)$$

ו־ $q(\lambda) \neq 0$ לכל $\lambda \in F$. מש״ל.

הסקלרים $\lambda_1, \ldots, \lambda_m$ במסקנה האחרונה נקראים **השורשים** של $f(x)$ (בשדה F).

דוגמה 5. בדוגמה 4 קיבלנו
$$x^3-1=f(x)=(x-1)(x^2+x+1)=(x-1)q(x)\ .$$
האם ל־ $q(x)$ יש שורשים ב־ $\mathbb{R}$? (רמז: מצא את כל השורשים המרוכבים של $f(x)$.)

מסקנה 3. בהנתן פולינום $f(x)$ ממעלה $n\geq 0$ קיימים לכל היותר n שורשים ל־ $f(x)$.

הוכחה. זה משום ש־ $m\leq n$ במסקנה 2. משי"ל.

דוגמה 6. ניקח $F:=\mathbb{C}$ ו־ $f(x):=x^2+bx+c$ עבור $b,c\in\mathbb{C}$. אז
$$f(x)=(x-\lambda_+)(x-\lambda_-)$$
כאשר
$$\lambda_\pm:=\frac{-b\pm\sqrt{b^2-4c}}{2}\ .$$
הביטוי b^2-4c נקרא **הדיסקרימיננטה** של $f(x)$, והוא מתאפס בדיוק כאשר $\lambda_+=\lambda_-$.

לפי המשפט היסודי של האלגברה (משפט 1 בפרק א') כל פולינום $f(x)$ מעל $\mathbb{C}$ מתפרק לגורמים ליניאריים. אולם המשפט אינו אומר מהם השורשים של $f(x)$; רק שקיימים שורשים. אם $n=2$ יש לנו נוסחה מוכרת (ראה למעלה). יש נוסחאות (מסובכות יותר) למעלות $n=3,4$. עבור $n\geq 5$ אין נוסחה (זהו משפט גאלואה).

אפשר למצוא **קירובים** של השורשים בשיטות נומריות שונות (בעזרת מחשב). למשל **שיטת ניוטון** למציאת שורשים ממשיים של פולינום עם מקדמים ממשיים. שיטה זו נלמדת בחדו"א ומשתמשת במשיקים לגרף של $f(x)$.

דוגמה 7. הנה עוד שימוש בחשבון פולינומים. יהי V מרחב הפולינומים ממעלה $\leq n$ מעל $\mathbb{R}$. נבחר $n+1$ נקודות שונות $a_0,a_1,\ldots,a_n\in\mathbb{R}$. ידוע לנו כי יש טרנספורמציה ליניארית $T:V\to\mathbb{R}^{n+1}$ שהנוסחה שלה
$$T(f):=\begin{bmatrix}f(a_0)\\ \vdots\\ f(a_n)\end{bmatrix}.$$
אנו נראה ש־ T היא איזומורפיזם ליניארי.

יהי $f(x)\in\operatorname{Ker}(T)$. נניח על דרך השלילה ש־ $f(x)\neq\vec{0}$. אז $n\geq\deg(f)\geq 0$, ועל פי מסקנה 3 ידוע שיש ל־ f לכל היותר n שורשים. מצד שני $f(a_0)=\cdots=f(a_n)=0$; וזו סתירה. אם כן רואים ש־ $\operatorname{Ker}(T)=\{\vec{0}\}$. מאחר ש־ $\dim(V)=n+1=\dim(\mathbb{R}^{n+1})$ נובע כי T איזומורפיזם.

המסקנה המעניינת היא שבהנתן מספרים $b_0,\ldots,b_r\in\mathbb{R}$ קיים פולינום יחיד $f(x)$ ממעלה $\geq n$ כך ש־ $f(a_i)=b_i$ לכל i.

תרגיל: תהי $B=[b_{ij}]$ המטריצה בגודל $(n+1)\times(n+1)$ שרכיביה $b_{ij}:=a_{i-1}^{j-1}$ עבור $1\leq i,j\leq n+1$. השתמש במה שהוכחנו זה עתה להראות כי $\det(B)\neq 0$. B נקראת **מטריצת ונדרמונדה**. בדרך אחרת ניתן אף להראות כי
$$\det(B)=\prod_{0\leq i<j\leq n}(a_j-a_i)\ .$$

הפולינום האופייני

הגדרה 2. תהי נתונה מטריצה $A = [a_{ij}] \in \mathrm{M}_{n\times n}(F)$. נתבונן במטריצה

$$xI - A = \begin{bmatrix} x - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & x - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \dots & x - a_{nn} \end{bmatrix}$$

כאשר x משתנה. זוהי מטריצה ריבועית, וניתן להשתמש בהגדרת הדטרמיננטה (הגדרה 2 בפרק ה׳) לקבל את הפולינום

$$p_A(x) := \det(xI - A) .$$

זהו פולינום עם מקדמים ב־ F הנקרא **הפולינום האופייני** של A.

טענה 1. תהי $A = [a_{ij}] \in \mathrm{M}_{n\times n}(F)$.

א. הפולינום האופייני הוא ממעלה n, וצורתו

$$p_A(x) = x^n - \mathrm{tr}(A)x^{n-1} + \cdots + (-1)^n \det(A) ,$$

כאשר $\mathrm{tr}(A) := \sum_{i=1}^n a_{ii}$.

ב. בהנתן $\lambda \in F$ מתקיים $p_A(\lambda) = \det(\lambda I - A)$; כלומר אם תחילה מציבים $x := \lambda$ במטריצה $xI - A$ ואח״כ מחשבים את הדטרמיננטה, התוצאה זהה להצבת $x := \lambda$ בפולינום $p_A(x)$.

לא נוכיח טענה זו. הביטוי $\mathrm{tr}(A)$ נקרא העקבה של A.

טענה 2. יהי V מרחב וקטורי סוף מימדי, יהי $T : V \to V$ אופרטור ליניארי ויהיו $\mathbf{v}, \mathbf{w}$ שני בסיסים של V. נגדיר $A := [T]_{\mathbf{v}}^{\mathbf{v}}$ ו־ $B := [T]_{\mathbf{w}}^{\mathbf{w}}$. אז $p_A(x) = p_B(x)$.

הוכחה. תהי $P := [I]_{\mathbf{w}}^{\mathbf{v}}$ מטריצת המעבר. אז $B = PAP^{-1}$. מאחר שכללי חשבון מטריצות תופסים גם למטריצות עם איברים שהם פולינומים (ראה למשל בספר Hoffman and Kunze), ומאחר ש־ $\det(P) \cdot \det(P^{-1}) = 1$, הרי

$$\begin{aligned} p_B(x) &= \det(xI - B) = \det(xI - PAP^{-1}) = \det(P(xI)P^{-1} - PAP^{-1}) \\ &= \det(P(xI - A)P^{-1}) = \det(P) \cdot \det(xI - A) \cdot \det(P^{-1}) \\ &= \det(xI - A) = p_A(x) \end{aligned}$$

מש״ל.

הגדרה 3. יהי T אופרטור ליניארי על מרחב סוף מימדי V. **הפולינום האופייני** של T מוגדר להיות

$$p_T(x) := p_A(x) ,$$

כאשר $\mathbf{v}$ הוא בסיס כלשהו של V ו־ $A := [T]_{\mathbf{v}}^{\mathbf{v}}$.

טענה 2 מראה שההגדרה הזו טובה, ז״א $p_T(x)$ איננו תלוי בבחירת הבסיס $\mathbf{v}$.

משפט 3. יהי T אופרטור ליניארי על מרחב סוף מימדי V מעל שדה F.

א. הערכים העצמיים של T הם השורשים ב־ F של הפולינום האופייני $p_T(x)$.

ב. יהי λ ערך עצמי של T. הוקטורים העצמיים של T השייכים ל־ λ הם הוקטורים השונים מאפס בתת־המרחב $\mathrm{Ker}(\lambda I - T)$, כאשר $I : V \to V$ הוא אופרטור הזהות.

הוכחה. זה שילוב של משפט 1 וטענה 1(ב). משי"ל.

דוגמה 8. ניקח את המרחב $V := F^n$, המטריצה $A := \begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{bmatrix}$ והאופרטור $T(v) := Av$.

מהו הפולינום האופייני?

$$p_T(x) = p_A(x) = \det(xI - A) = \det(\begin{bmatrix} x-\lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & x-\lambda_n \end{bmatrix}) = (x-\lambda_1)\cdots(x-\lambda_n)\,.$$

רואים שהערכים העצמיים הם $\lambda_1, \dots, \lambda_n$.

דוגמה 9. השדה הוא $\mathbb{R}$ והמרחב הוא $V := \mathbb{R}^2$. ניקח את המטריצה $A := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ ואת האופרטור $T(v) := Av$. אז

$$p_T(x) = \det(xI - A) = \det(\begin{bmatrix} x & 1 \\ -1 & x \end{bmatrix}) = x^2 + 1\,.$$

לפולינום $p_T(x)$ אין שורשים ב־ $\mathbb{R}$, ולכן אין ערכים עצמיים לאופרטור T. ראינו זאת כבר בדוגמה 2.

דוגמה 10. השדה הוא $\mathbb{R}$, המרחב הוא

$$V := \{\text{פולינומים } f(t) \text{ מעל } \mathbb{R} \text{ ממעלה} \leq n\}$$

והאופרטור הוא $D(f) := \frac{\mathrm{d}f}{\mathrm{d}t}$. ניקח את הבסיס הסטנדרטי $\mathbf{v} = (t^0, \dots, t^n)$. המטריצה המייצגת היא

$$[D]^{\mathbf{v}}_{\mathbf{v}} = A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & n \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

והפולינום האופייני הוא

$$p_D(x) = \det(\begin{bmatrix} x & -1 & 0 & \dots & 0 \\ 0 & x & -2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -n \\ 0 & 0 & 0 & \dots & x \end{bmatrix}) = x^{n+1}\,.$$

רואים שהערך העצמי היחיד הוא $\lambda = 0$.

מציאת וקטורים עצמיים

יהי T אופרטור ליניארי על מרחב V, ונניח כי מצאנו ערך עצמי λ של T (כלומר שורש של $p_T(x)$). כיצד נמצא וקטור עצמי השייך ל־ λ? יהי $\vec{0} \neq v$. אז v וקטור עצמי השייך ל־ λ אס"ם $T(v) = \lambda v$, כלומר אס"ם $v \in \operatorname{Ker}(\lambda I - T)$.

הגדרה 4. יהי T אופרטור ליניארי על מרחב וקטורי V ויהי $\lambda \in F$ ערך עצמי של T. **המרחב העצמי** השייך ל־ λ הוא תת המרחב

$$V_\lambda := \operatorname{Ker}(\lambda I - T) \subset V$$

כאשר I הוא אופרטור הזהות.

כלומר הוקטורים העצמיים של T השייכים ל־ λ הם בדיוק הוקטורים השונים מ־ $\vec{0}$ ב־ V_λ. אנו נתעניין בבסיס של V_λ.

דוגמה 11. נתבונן במקרה הבא: $F := \mathbb{R}$, $V := \mathbb{R}^3$, $A := \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}$ ו־ $T(v) := Av$. נחשב את הפולינום האופייני $p_T(x)$.

$$p_T(x) = \det(xI - A) = \det \begin{bmatrix} x-5 & 6 & 6 \\ 1 & x-4 & -2 \\ -3 & 6 & x+4 \end{bmatrix} = x^3 - 5x^2 + 8x - 4 .$$

ננסה לנחש שורש של $p_A(x)$, ונתחיל במספרים שלמים $\lambda = 0, \pm 1, \pm 2, \ldots$. עם $\lambda = 1$ יש הצלחה:

$$p(1) = 1 - 5 + 8 - 4 = 0 .$$

כעת נחלק את $p_T(x)$ ב־ $x - 1$.

$$\begin{array}{rrrr|l} x^2 & -4x & +4 & & \\ \hline x^3 & -5x^2 & +8x & -1 & x-1 \\ x^3 & -x^2 & & & \\ \hline & -4x^2 & +8x & -4 & \\ & -4x^2 & +4x & & \\ \hline & & 4x & -4 & \\ & & 4x & -4 & \\ \hline & & & 0 & \end{array}$$

מקבלים ש־

$$p_T(x) = (x-1)(x^2 - 4x + 4) = (x-1)(x-2)^2 ,$$

ולכן הערכים העצמיים של T הם $\lambda_1 = 1$ ו־ $\lambda_2 = 2$.

נמשיך במציאת בסיס ל־ V_{λ_1}. דירוג המטריצה

$$\lambda_1 I - A = \begin{bmatrix} -4 & 6 & 6 \\ 1 & -3 & -2 \\ -3 & 6 & 5 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 \end{bmatrix}$$

מראה ש־

$$V_{\lambda_1} = \operatorname{Ker}(\lambda_1 I - A) = \operatorname{Sp}(v_1)$$

כאשר $\begin{bmatrix} 1 \\ -\frac{1}{3} \\ 1 \end{bmatrix} =: v_1$. בדומה מחשבים עבור V_{λ_2} ומוצאים ש־

$$V_{\lambda_2} = \mathrm{Ker}(\lambda_2 I - A) = \mathrm{Sp}(v_2, v_3)$$

כאשר $\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} =: v_2$ ו־ $\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} =: v_3$.

לסיום נעשה בדיקה (חלקית). נחשב את $T(v_2)$:

$$T(v_2) = Av_2 = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 0 \end{bmatrix} = 2 \cdot v_2 = \lambda_2 v_2 .$$

ריבוי גיאומטרי וריבוי אלגברי של ערך עצמי

הגדרה 5. יהי T אופרטור ליניארי על מרחב וקטורי סוף מימדי V מעל השדה F, ויהי $\lambda \in F$ ערך עצמי של T.

א. נפרק את הפולינום האופייני למכפלה

$$p_T(x) = (x-\lambda)^{e_\lambda} q(x)$$

כאשר e_λ מספר שלם חיובי ו־ $q(\lambda) \neq 0$. המספר e_λ נקרא **הריבוי האלגברי** של הערך העצמי λ.

ב. **הריבוי הגיאומטרי** של λ הוא $\dim(V_\lambda)$, כאשר $V_\lambda \subset V$ הוא המרחב העצמי השייך ל־ λ.

משפט 4. בתנאי הגדרה 5 מתקיים אי השוויון $\dim(V_\lambda) \leq e_\lambda$.

לא נוכיח משפט זה, וגם לא נזדקק לו בהמשך.

דוגמה 12. ניקח $F := \mathbb{Q}$, $V := \mathbb{Q}^2$ ו־ $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} =: A$. הפולינום האופייני הוא $p_A(x) = (x-1)^2$. לכן הערך העצמי היחיד של A הוא $\lambda = 1$, והריבוי האלגברי הוא $e_\lambda = 2$. מאחר ש־ $\lambda I - A = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}$ מקבלים שהריבוי הגיאומטרי הוא $\dim(V_\lambda) = 1$.

טענה 3. נניח כי $\lambda_1, \ldots, \lambda_m$ הם ערכים עצמיים **שונים** של T, ולכל i נתון וקטור עצמי v_i של λ_i. אז הסדרה $(v_1, \ldots, v_m)$ בלתי תלויה ליניארית.

הוכחה. נניח בדרך השלילה שהסדרה $(v_1, \ldots, v_m)$ תלויה ליניארית. אז יש וקטור v_i שהוא צירוף ליניארי של קודמיו בסדרה. יהי i_0 האינדקס המינימלי כך ש־ v_{i_0} צירוף ליניארי של קודמיו. ז״א

$(v_1,\dots,v_{i_0-1})$ סדרה בלתי תלויה ליניארית, ויש סקלרים $a_1,\dots,a_{i_0-1}$ כך ש־ $v_{i_0}=\sum_{j=1}^{i_0-1}a_jv_j$.
מאחר ש־ $v_{i_0}\neq\vec{0}$ הרי $i_0\geq 2$ ויש j_0 בטווח $1,\dots,i_0-1$ כך ש־ $a_{j_0}\neq 0$. כעת

$$\begin{aligned}\vec{0}&=\lambda_{i_0}v_{i_0}-T(v_{i_0})=(\lambda_{i_0}I-T)(v_{i_0})\\&=(\lambda_{i_0}I-T)\left(\sum_{j=1}^{i_0-1}a_jv_j\right)=\sum_{j=1}^{i_0-1}a_j(\lambda_{i_0}I-T)(v_j)=\\&=\sum_{j=1}^{i_0-1}a_j(\lambda_{i_0}v_j-T(v_j))=\sum_{j=1}^{i_0-1}a_j(\lambda_{i_0}v_j-\lambda_jv_j)\\&=\sum_{j=1}^{i_0-1}a_j(\lambda_{i_0}-\lambda_j)v_j\end{aligned}$$

מאחר ש־ $\lambda_{j_0}\neq\lambda_{i_0}$ הרי $a_{j_0}(\lambda_{i_0}-\lambda_{j_0})\neq 0$. מצאנו ש־ $\vec{0}$ הוא צירוף ליניארי לא טריוויאלי של אברי הסדרה $(v_1,\dots,v_{i_0-1})$. זאת סתירה משום שהסדרה הזו בת״ל. משי״ל.

משפט 5. יהי V מרחב וקטורי סוף מימדי מעל השדה F, יהי $T:V\to V$ אופרטור ליניארי, ויהיו $\lambda_1,\dots,\lambda_m$ ערכים עצמיים **שונים** של T. נניח כי לכל i מ־ 1 עד m נתונה סדרה בלתי תלויה ליניארית $\mathbf{v}_i=(v_{i1},v_{i2},\dots)$ של וקטורים במרחב העצמי V_{λ_i}. אז הסדרה המשורשרת

$$\mathbf{v}:=(\mathbf{v}_1,\mathbf{v}_2,\dots,\mathbf{v}_m)=(v_{11},v_{12},\dots,v_{21},v_{22},\dots)$$

היא סדרה בלתי תלויה ליניארית ב־ V.

הוכחה. נניח על דרך השלילה כי הסדרה $\mathbf{v}$ תלויה ליניארית. ז״א ישנם סקלרים a_{ij}, לא כולם 0, כך ש־

$$\vec{0}=a_{11}v_{11}+a_{12}v_{12}+\cdots+a_{21}v_{21}+a_{22}v_{22}+\cdots\ .$$

נגדיר וקטורים חדשים

$$w_i:=a_{i1}v_{i1}+a_{i2}v_{i2}+\cdots\in V_{\lambda_i}\ .$$

לכן הסכום הוא

$$\vec{0}=w_1+w_2+\cdots+w_m\ .$$

יהי (i,j) זוג אינדקסים כך ש־ $a_{ij}\neq 0$. מאחר ש־ $\mathbf{v}_i$ סדרה בת״ל נובע ש־ $w_i\neq\vec{0}$. כלומר לא כל הוקטורים $w_1,\dots,w_m$ הם אפס. ע״י מיספור מחדש אפשר להניח כי $w_1,\dots,w_k$ שונים מאפס, ו־ $w_{k+1}=\cdots=w_m=\vec{0}$ ו־ $k\geq 1$. כעת $w_1+\cdots+w_k=\vec{0}$, ו־ $w_1,\dots,w_k$ הם וקטורים עצמיים השייכים לערכים העצמיים השונים $\lambda_1,\dots,\lambda_k$ בהתאמה. זו סתירה לטענה 3, אשר לפיה הסדרה $(w_1,\dots,w_k)$ בת״ל. משי״ל.

מסקנה 4. יהיו $\lambda_1,\dots,\lambda_m$ ערכים עצמיים שונים של אופרטור ליניארי $T:V\to V$, ויהי $n:=\dim(V)$. אז

$$\sum_{i=1}^{m}\dim(V_{\lambda_i})\leq n\ .$$

הוכחה. בהנתן סדרה סופית $\mathbf{w}$ נסמן ב־ $|\mathbf{w}|$ את האורך שלה. לכל i ניקח בסיס $\mathbf{v}_i$ לתת המרחב V_{λ_i}. אז $|\mathbf{v}_i| = \dim(V_{\lambda_i})$. על פי המשפט הסדרה המשורשרת $\mathbf{v} := (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$ היא בת״ל. מקבלים

$$\sum_{i=1}^{m} \dim(V_{\lambda_i}) = \sum_{i=1}^{m} |\mathbf{v}_i| = |\mathbf{v}| \le \dim(V) = n \; .$$

מש״ל.

מסקנה 5. אם $\lambda_1, \ldots, \lambda_m$ ערכים עצמיים שונים של T ו־

$$\sum_{i=1}^{m} \dim(V_{\lambda_i}) = \dim(V)$$

אז ל־ V יש בסיס המורכב מוקטורים עצמיים של T.

הוכחה. ניקח $\mathbf{v}_1, \ldots, \mathbf{v}_m$ בסיסים כמו בהוכחת המסקנה הקודמת. אז $\mathbf{v} := (\mathbf{v}_1, \ldots \mathbf{v}_m)$ סדרה בת״ל באורך $\dim(V)$, כלומר זה בסיס של V. מש״ל.

מסקנה 6. יהי T אופרטור ליניארי על מרחב n־ מימדי V עם n ערכים עצמיים שונים $\lambda_1, \ldots, \lambda_n$. אז ל־ V יש בסיס המורכב מוקטורים עצמיים של T.

הוכחה. מאחר ש־ $\dim(V_{\lambda_i}) \ge 1$ לכל i הרי

$$\sum_{i=1}^{m} \dim(V_{\lambda_i}) \ge n \; .$$

ע״פ מסקנה 4 יש שוויון. כעת נשתמש במסקנה 5. מש״ל.

ליכסון מטריצות ואופרטורים

מטריצה $D = [d_{ij}] \in \mathrm{M}_{n\times n}(F)$ נקראת **אלכסונית** אם $d_{ij} = 0$ לכל $i \ne j$. כלומר אם

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

כאשר $d_{ii} = \lambda_i$. מקובל לרשום את אברי האלכסון כך, כמובן משום שאלו הערכים העצמיים של D.

נוח מאוד לעשות פעולות חשבוניות עם מטריצות אלכסוניות. נדגים זאת במקרה $n = 2$. החיבור הוא

$$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} + \begin{bmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 + \mu_1 & 0 \\ 0 & \lambda_2 + \mu_2 \end{bmatrix} \; .$$

הכפל הוא

$$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \cdot \begin{bmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{bmatrix} = \begin{bmatrix} \lambda_1\mu_1 & 0 \\ 0 & \lambda_2\mu_2 \end{bmatrix} \; .$$

כך גם לגבי פולונומים. בהנתן פולינום

$$f(x) = c_m x^m + \cdots + c_1 x + c_0$$

עם מקדמים בשדה F, אז

$$f(\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}) = \sum_{i=0}^{m} c_i \begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}^i = \begin{bmatrix}\sum_{i=0}^m c_i\lambda_1^i & 0\\ 0 & \sum_{i=0}^m c_i\lambda_2^i\end{bmatrix} = \begin{bmatrix}f(\lambda_1) & 0\\ 0 & f(\lambda_2)\end{bmatrix}\ .$$

הגדרה 6. מטריצה $A \in \mathrm{M}_{n\times n}(F)$ נקראת **מטריצה ניתנת ללכסון** (מעל F) אם יש מטריצה הפיכה $P \in \mathrm{M}_{n\times n}(F)$ כך ש- $D := P^{-1}AP$ היא מטריצה אלכסונית. P נקראת **מטריצה מלכסנת** של A.

הגדרה 7. יהי V מרחב וקטורי סוף מימדי מעל שדה F ויהי $T : V \to V$ אופרטור ליניארי. T נקרא **אופרטור ניתן לליכסון** אם יש בסיס $\mathbf{w}$ של V המורכב מוקטורים עצמיים של T. הבסיס $\mathbf{w}$ נקרא **בסיס מלכסן** של V ביחס ל- T.

טענה 4. יהי V מרחב וקטורי סוף מימדי מעל שדה F, יהי $T : V \to V$ אופרטור ליניארי, ויהי $\mathbf{w} = (w_1, \ldots, w_n)$ בסיס של V. התנאים הבאים שקולים:

א. $\mathbf{w}$ הוא בסיס מלכסן של V ביחס ל- T.

ב. המטריצה $[T]_{\mathbf{w}}^{\mathbf{w}}$ היא אלכסונית.

כאשר זה קורה יהיו $\lambda_1, \ldots, \lambda_n$ הערכים העצמיים כך ש- $T(w_i) = \lambda_i w_i$. אז

$$[T]_{\mathbf{w}}^{\mathbf{w}} = \begin{bmatrix}\lambda_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_n\end{bmatrix}\ .$$

הוכחה. ע״פ הגדרה

$$[T]_{\mathbf{w}}^{\mathbf{w}} = \begin{bmatrix}[T(w_1)]_{\mathbf{w}} & \cdots & [T(w_n)]_{\mathbf{w}}\end{bmatrix}\ .$$

משום כך

$$[T]_{\mathbf{w}}^{\mathbf{w}} = \begin{bmatrix}\lambda_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_n\end{bmatrix}$$

אס״ם $[T(w_i)]_{\mathbf{w}} = \lambda_i \vec{e}_i$ לכל i, אס״ם $T(w_i) = \lambda_i w_i$ לכל i. כלומר $[T]_{\mathbf{w}}^{\mathbf{w}}$ אלכסונית אס״ם $\mathbf{w}$ הוא בסיס מלכסן. מש״ל.

משפט 6. יהי V מרחב וקטורי סוף מימדי מעל השדה F, יהי $T : V \to V$ אופרטור ליניארי, ויהי $\mathbf{v}$ בסיס של V. התנאים הבאים שקולים:

א. T הוא האופרטור ניתן לליכסון.

ב. המטריצה $A := [T]_{\mathbf{v}}^{\mathbf{v}}$ ניתנת לליכסון.

אם התנאים הללו מתקיימים יהי $\mathbf{w} = (w_1, \ldots, w_n)$ בסיס מלכסן של V ביחס ל־ T, ויהיו $\lambda_1, \ldots, \lambda_n$ הערכים העצמיים המתאימים, כלומר $T(w_i) = \lambda_i w_i$. נגדיר

$$D := \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

ו־ $P := [I]^{\mathbf{w}}_{\mathbf{v}}$. אז

$$P^{-1}AP = D = [T]^{\mathbf{w}}_{\mathbf{w}} \ .$$

הוכחה. א $\Leftarrow$ ב: נניח כי T ניתן לליכסון, כלומר יש ל־ V בסיס מלכסן $\mathbf{w}$ שביחס אליו $D := [T]^{\mathbf{w}}_{\mathbf{w}}$ אלכסונית. תוך שימוש במשפט 5 מפרק ו׳ מקבלים כי

$$P^{-1}AP = [I]^{\mathbf{v}}_{\mathbf{w}} \cdot [T]^{\mathbf{v}}_{\mathbf{v}} \cdot [I]^{\mathbf{w}}_{\mathbf{v}} = [T]^{\mathbf{w}}_{\mathbf{w}} = D \ ,$$

כאשר $P := [I]^{\mathbf{w}}_{\mathbf{v}}$. לכן A ניתנת לליכסון. לפי טענה 4 רואים ש־ $D = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$ כאשר $T(w_i) = \lambda_i w_i$.

ב $\Leftarrow$ א: ידוע שיש מטריצה הפיכה P כך ש־ $D := P^{-1}AP$ היא אלכסונית. יהי $w_j \in V$ הווקטור כך ש־ $[w_j]_{\mathbf{v}}$ הוא עמודה מס׳ j במטריצה P, ונגדיר $\mathbf{w} := (w_1, \ldots, w_n)$. אז $P = [I]^{\mathbf{w}}_{\mathbf{v}}$, וע״פ החישוב הקודם $D = [T]^{\mathbf{w}}_{\mathbf{w}}$. על פי טענה 4, $\mathbf{w}$ הוא בסיס מלכסן של T. משי״ל.

מסקנה 7. תהי $A \in \mathrm{M}_{n\times n}(F)$. נניח כי $\mathbf{w} := (w_1, \ldots, w_n)$ בסיס של F^n, ולכל i מתקיים ש־ w_i הוא וקטור עצמי של A השייך לערך עצמי λ_i. נגדיר

$$P := [w_1, \ldots, w_n] \in \mathrm{M}_{n\times n}(F) \ .$$

אז P הפיכה ו־

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \ .$$

הוכחה. במשפט ניקח $V := F^n$, $\mathbf{v} := (\vec{e}_1, \ldots, \vec{e}_n)$ ו־ $T(v) := Av$. משי״ל.

דרך טובה לזכור את הנוסחה שבמסקנה האחרונה היא הצורה השקולה

$$AP = PD$$

אשר ניתן לקבל באופן הבא:

$$\begin{aligned} AP &= A \begin{bmatrix} w_1 & \cdots & w_n \end{bmatrix} = \begin{bmatrix} Aw_1 & \cdots & Aw_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 w_1 & \cdots & \lambda_n w_n \end{bmatrix} = \begin{bmatrix} w_1 & \cdots & w_n \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} = PD \end{aligned}$$

טענה 5. תהי A מטריצה ריבועית מעל השדה F ויהי $f(x)$ פולינום עם מקדמים ב־ F. נניח כי A ניתנת ללכסון, כלומר $A = PDP^{-1}$ ו־ $D = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$. אז

$$f(A) = P \begin{bmatrix} f(\lambda_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & f(\lambda_n) \end{bmatrix} P^{-1}\,.$$

הוכחה. לכל i מתקיים

$$A^i = (PDP^{-1})^i = \underbrace{(PDP^{-1})(PDP^{-1})\cdots(PDP^{-1})}_{i} = PD^iP^{-1}\,.$$

לכל שתי מטריצות B ו־ C בגודל מתאים מתקיים

$$PBP^{-1} + PCP^{-1} = P(BP^{-1} + CP^{-1}) = P(B+C)P^{-1}\,.$$

נניח כי $f(x) = c_m x^m + \cdots + c_1 a + c_0$. אז:

$$\begin{aligned} f(A) = c_m A^m + \cdots + c_1 A + c_0 I &= \\ &= c_m PD^mP^{-1} + \cdots + c_1 PDP^{-1} + c_0 PIP^{-1} \\ &= P(c_m D^m)P^{-1} + \cdots + P(c_1 D)P^{-1} + P(c_0 I)P^{-1} \\ &= P(c_m D^m + \cdots + c_1 D + c_0 I)P^{-1} \\ &= P \cdot f(D) \cdot P^{-1} \\ &= P \begin{bmatrix} f(\lambda_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & f(\lambda_n) \end{bmatrix} P^{-1} \end{aligned}$$

משי"ל.

דוגמה 13. ניקח את השדה $F := \mathbb{Q}$ והמטריצה $A := \begin{bmatrix} -1 & 5 \\ -10 & 14 \end{bmatrix}$. רוצים לחשב את A^{100}. ננסה תחילה ללכסן את A. הפולינום האופייני הוא

$$p_A(x) = \det(\begin{bmatrix} x+1 & -5 \\ 10 & x-14 \end{bmatrix}) = x^2 - 13x + 36 = (x-9)(x-4)\,.$$

יש אם כן שני ערכים עצמיים שונים $\lambda_1 = 4$ ו־ $\lambda_2 = 9$ והמטריצה A ניתנת ללכסון.
עכשיו נמצא וקטורים עצמיים. ע"י דרוג

$$\lambda_1 I - A = \begin{bmatrix} 5 & -5 \\ 10 & -10 \end{bmatrix} \to \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$$

רואים כי $V_{\lambda_1} = \mathrm{Sp}(v_1)$ כאשר $v_1 := \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. בדומה דרוג

$$\lambda_2 I - A = \begin{bmatrix} 10 & -5 \\ 10 & -5 \end{bmatrix} \to \begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{bmatrix}$$

נותן ש־ $V_{\lambda_2} = \mathrm{Sp}(v_2)$ כאשר $v_2 := \begin{bmatrix}1\\2\end{bmatrix}$.
אם כן A ניתנת לליכסון ע״י המטריצות

$$P := [v_1, v_2] = \begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix}$$

ו־

$$D := \begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix} = \begin{bmatrix}4 & 0\\ 0 & 9\end{bmatrix}.$$

חישוב מראה כי $P^{-1} = \begin{bmatrix}2 & -1\\ -1 & 1\end{bmatrix}$. נעצור בשלב זה לבצע בדיקה:

$$PDP^{-1} = \begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix}\begin{bmatrix}4 & 0\\ 0 & 9\end{bmatrix}\begin{bmatrix}2 & -1\\ -1 & 1\end{bmatrix} = \begin{bmatrix}-1 & -10\\ 5 & 14\end{bmatrix} = A.$$

לבסוף נחשב את A^{100}:

$$\begin{aligned} A^{100} = PD^{100}P^{-1} &= \begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix}\begin{bmatrix}4^{100} & 0\\ 0 & 9^{100}\end{bmatrix}\begin{bmatrix}2 & -1\\ -1 & 1\end{bmatrix}\\ &= \begin{bmatrix}2\cdot 4^{100} - 9^{100} & -4^{100} + 9^{100}\\ 2\cdot 4^{100} - 2\cdot 9^{100} & -4^{100} + 2\cdot 9^{100}\end{bmatrix}\end{aligned}$$

דוגמה 14. עבור המטריצה A מדוגמה 13 מצא מטריצה $C \in \mathrm{M}_{2\times 2}(\mathbb{Q})$ כך ש־ $C^2 = A$. המטריצה C היא "$\sqrt[2]{A}$". ננסה למצוא פתרון מהצורה "$C = P\sqrt[2]{D}P^{-1}$". יש ארבע אפשרויות שמייד רואים עבור "$\sqrt{D}$", והן:

$$\begin{bmatrix}\pm\sqrt{\lambda_1} & 0\\ 0 & \pm\sqrt{\lambda_2}\end{bmatrix} = \begin{bmatrix}\pm 2 & 0\\ 0 & \pm 3\end{bmatrix}.$$

נבחר באחת האפשרויות ונקבל

$$C = \begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix}\cdot\begin{bmatrix}2 & 0\\ 0 & 3\end{bmatrix}\cdot\begin{bmatrix}2 & -1\\ -1 & 1\end{bmatrix} = \begin{bmatrix}1 & 1\\ -2 & 4\end{bmatrix}.$$

נסיים בבדיקת התוצאה:

$$C^2 = \begin{bmatrix}1 & 1\\ -2 & 4\end{bmatrix}^2 = \begin{bmatrix}1 & 1\\ -2 & 4\end{bmatrix}\cdot\begin{bmatrix}1 & 1\\ -2 & 4\end{bmatrix} = \begin{bmatrix}-1 & 5\\ -10 & 14\end{bmatrix} = A.$$

(הערה: ניתן להראות כי במקרה זה יש בדיוק ארבעה פתרונות ב־ $\mathrm{M}_{n\times n}(\mathbb{Q})$ למשוואה $X^2 = A$.)

משפט 7. תהי A מטריצה בגודל $n \times n$ מעל השדה F. המטריצה A ניתנת לליכסון מעל F אם״ם מתקיימים שני התנאים הבאים.

1. הפולינום האופייני $p_A(x)$ מתפרק באופן הבא:

$$p_A(x) = (x-\lambda_1)^{e_1}\cdots(x-\lambda_m)^{e_m}$$

כאשר $\lambda_1, \ldots, \lambda_m$ הם סקלרים שונים ב־ F ו־ $e_1, \ldots, e_m$ הם מספרים שלמים חיוביים.

2. לכל i הריבוי האלגברי של הערך העצמי λ_i שווה לריבוי הגיאומטרי שלו.

הוכחה. א. נניח כי A ניתנת ללכסון. תהיינה P, D מטריצות כך ש־ P הפיכה,

$$D = \begin{bmatrix} \mu_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mu_n \end{bmatrix}$$

אלכסונית ו־ $A = PDP^{-1}$. אז

$$p_A(x) = p_D(x) = (x-\mu_1)\cdots(x-\mu_n) \,.$$

יהיו $\lambda_1, \ldots, \lambda_m$ הסקלרים השונים המופיעים בסדרה $\boldsymbol{\mu} := (\mu_1, \ldots, \mu_n)$, ונסמן ב־ e_i את מספר הפעמים ש־ λ_i מופיע בסדרה $\boldsymbol{\mu}$. אז

$$p_A(x) = (x-\lambda_1)^{e_1}\cdots(x-\lambda_m)^{e_m} \,.$$

ו־ $n = \sum_{i=1}^m e_i$. לכל j תהי v_j עמודה מס' j במטריצה P. אז v_j הוא וקטור עצמי של A השייך לערך העצמי μ_j. נגדיר $\mathbf{v} := (v_1, \ldots, v_n)$, ולכל i נגדיר את $\mathbf{v}_i$ להיות תת־הסדרה של $\mathbf{v}$ המורכבת מהוקטורים העצמיים השייכים ל־ λ_i. אז $\mathbf{v}_i$ סדרה בת״ל בתת המרחב V_{λ_i} ו־ $e_i = |\mathbf{v}_i|$. לכן $\dim(V_{\lambda_i}) \geq |\mathbf{v}_i| = e_i$. ע״פ מסקנה 4 מקבלים

$$n \geq \sum_{i=1}^m \dim(V_{\lambda_i}) \geq \sum_{i=1}^m e_i = n \,.$$

זה מחייב שלכל i יש שוויון $\dim(V_{\lambda_i}) = e_i$.

ב. לכיוון השני, נניח שמתקיימים התנאים 1 ו־2. נשים לב שהפרוק של $p_A(x)$ גורר ש־ $n = \sum_{i=1}^m e_i$. לכל i נבחר בסיס $\mathbf{v}_i$ לתת־ המרחב V_{λ_i}. ע״פ משפט 5 הסדרה המשורשרת $\mathbf{v} := (\mathbf{v}_1, \ldots, \mathbf{v}_m)$ היא בת״ל. אבל

$$|\mathbf{v}| = \sum_{i=1}^m |\mathbf{v}_i| = \sum_{i=1}^m \dim(V_{\lambda_i}) = \sum_{i=1}^m e_i = n \,.$$

לכן $\mathbf{v}$ בסיס של V המורכב מוקטורים עצמיים של A. נרשום $\mathbf{v} := (v_1, \ldots, v_n)$ ו־

$$(\mu_1, \mu_2, \ldots, \mu_n) := (\underbrace{\lambda_1, \ldots, \lambda_1}_{e_1}, \ldots, \underbrace{\lambda_m, \ldots, \lambda_m}_{e_m}) \,;$$

כלומר v_i וקטור עצמי השייך לערך העצמי μ_i. נגדיר מטריצות $P := \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$ ו־

$$D := \begin{bmatrix} \mu_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mu_n \end{bmatrix} .$$

אז

$$AP = \begin{bmatrix} Av_1 & \cdots & Av_n \end{bmatrix} = \begin{bmatrix} \mu_1 v_1 & \cdots & \mu_n v_n \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \begin{bmatrix} \mu_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mu_n \end{bmatrix} = PD$$

ולכן $A = PDP^{-1}$. משל.

אלגוריתם 1. (ליכסון של מטריצה) נתונה מטריצה $A \in M_{n\times n}(F)$.

1. חשב את הפולינום האופייני $p_A(x)$.
2. מצא את שורשי $p_A(x)$ ב־ F, ז״א את הערכים העצמיים של A. נניח שיש m ערכים עצמיים שונים $\lambda_1, \ldots, \lambda_m$, עם ריבויים אלגבריים $e_1, \ldots, e_m$ בהתאמה.
3. אם $\sum_{i=1}^m e_i < n$ אז A איננה ניתנת ללכסון. עוצרים.

4. לכל i מצא בסיס $\mathbf{v}_i$ למרחב העצמי V_{λ_i}, ע״י פתרון המשוואה ההומוגנית $(\lambda_i I - A)X = O$.
5. אם $|\mathbf{v}_i| < e_i$ לאיזה i אז A אינה ניתנת ללכסון. עוצרים.
6. כעת $\sum_{i=1}^{m} |\mathbf{v}_i| = n$. המטריצה P שעמודותיה הם הוקטורים בסדרה המשורשרת $(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ היא הפיכה, והמטריצה $D := P^{-1}AP$ היא אלכסונית.

אלגוריתם 2. (ליכסון של אופרטור) נתונים מרחב וקטורי V ממימד n מעל השדה F ואופרטור ליניארי $T : V \to V$.

1. מצא בסיס $\mathbf{v}$ של V.
2. תהי $A := [T]_{\mathbf{v}}^{\mathbf{v}}$ המטריצה המייצגת את T ביחס לבסיס $\mathbf{v}$.
3. אם A איננה ניתנת לליכסון אז גם T איננו ניתן לליכסון. עוצרים.
4. יהי $(u_1, \ldots, u_n)$ בסיס של F^n המורכב מוקטורים עצמיים של A. נגדיר סדרת וקטורים $\mathbf{w} = (w_1, \ldots, w_n)$ ב־ V ע״י הנוסחה $[w_i]_{\mathbf{v}} = u_i$. אז $\mathbf{w}$ בסיס מלכסן של V ביחס ל־ T.

דוגמה 15. ניקח $F := \mathbb{R}$, המרחב

$$V := \{\text{פולינומים } f(t) \text{ ממעלה } \leq 2 \text{ מעל } \mathbb{R}\}$$

ואופרטור ליניארי $T : V \to V$ המוגדר ע״י

$$T(f) := (t+1)\frac{\mathrm{d}f}{\mathrm{d}t} .$$

כבסיס של V ניקח את $\mathbf{v} := (t^0, t^1, t^2)$. אז

$$\begin{aligned}
T(t^0) &= (t+1) \cdot 0 = 0 \\
T(t^1) &= (t+1) \cdot 1 = t + 1 = t^0 + t^1 \\
T(t^2) &= (t+1) \cdot 2t = 2t + 2t^2 = 2t^1 + 2t^2
\end{aligned}$$

לכן המטריצה המייצגת היא $[T]_{\mathbf{v}}^{\mathbf{v}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix}$. הפולינום האופייני הוא

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

והערכים העצמיים הם $\lambda_1 := 0$, $\lambda_2 := 1$ ו־ $\lambda_3 := 2$. רואים מייד ש־ A ניתנת לליכסון. נמצא וקטורים עצמיים ל־ A השייכים לערכים עצמיים אלו:

$$u_3 := \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad u_2 := \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad u_1 := \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

נתרגם לבסיס $\mathbf{w} = (f_1(t), f_2(t), f_3(t))$ של V:

$$f_3(t) = t^0 + 2t^1 + t^2, \quad f_2(t) = t^0 + t^1, \quad f_1(t) = t^0 .$$

לבסוף נבדוק את התוצאה.

$$\begin{aligned}
T(f_1) &= T(t_0) = 0 = \lambda_1 f_1 \\
T(f_2) &= T(t^0 + t^1) = t^0 + t^1 = \lambda_2 f_2 \\
T(f_3) &= T(t^0 + 2t^1 + t^2) = 2t^0 + 4t^1 + 2t^2 = \lambda_3 f_3 .
\end{aligned}$$

דוגמה 16. נתון הפולינום $f(x) = x^2 + 5$ עם מקדמים ב־ $\mathbb{C}$. פתור את המשוואה

$$(*) \qquad f(X) = \begin{bmatrix} -1 & 5 \\ -10 & 14 \end{bmatrix}$$

ב־ $\mathrm{M}_{2\times 2}(\mathbb{C})$; כלומר מצא מטריצה C המקיימת את המשוואה. נרשום $A := \begin{bmatrix} -1 & 5 \\ -10 & 14 \end{bmatrix}$. ידוע לנו כבר כי A ניתנת לליכסון (ראה דוגמה 12): $A = PDP^{-1}$ כאשר $D := \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}$. נפתור תחילה את המשוואה "האלכסונית"

$$f(Y) = D \; .$$

ננסה פתרון אלכסוני $Y = \begin{bmatrix} y_1 & 0 \\ 0 & y_2 \end{bmatrix}$. זה אומר שיש לפתור 2 משוואות סקלריות: $f(y_1) = 4$ ו־ $f(y_2) = 9$. הפתרונות ידועים: $y_1 := \pm\mathbf{i}$ ו־ $y_2 := \pm 2$, כאשר $\mathbf{i} := \sqrt{-1}$. לכן

$$Y = \begin{bmatrix} \pm\mathbf{i} & 0 \\ 0 & \pm 2 \end{bmatrix} \; .$$

הקשר בין המשתנים X ו־ Y הוא: $X = PYP^{-1}$. לכן מצאנו ארבעה פתרונות למשוואה $(*)$, שהם

$$X := P \begin{bmatrix} \pm\mathbf{i} & 0 \\ 0 & \pm 2 \end{bmatrix} P^{-1} \; .$$

נבחר אחד מהם, נאמר

$$C := P \begin{bmatrix} \mathbf{i} & 0 \\ 0 & 2 \end{bmatrix} P^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} \mathbf{i} & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2\mathbf{i}-2 & 2-\mathbf{i} \\ 2\mathbf{i}-4 & 4-\mathbf{i} \end{bmatrix} \; .$$

עתה נעשה בדיקה.

$$f(C) = C^2 + 5I = \begin{bmatrix} 2\mathbf{i}-2 & 2-\mathbf{i} \\ 2\mathbf{i}-4 & 4-\mathbf{i} \end{bmatrix}^2 + 5I =$$
$$= \begin{bmatrix} 2\mathbf{i}-2 & 2-\mathbf{i} \\ 2\mathbf{i}-4 & 4-\mathbf{i} \end{bmatrix} \cdot \begin{bmatrix} 2\mathbf{i}-2 & 2-\mathbf{i} \\ 2\mathbf{i}-4 & 4-\mathbf{i} \end{bmatrix} + \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} -1 & 5 \\ -10 & 14 \end{bmatrix}$$

לסיום הערה: ניתן להוכיח כי יש בדיוק ארבעה פתרונות למשוואה $(*)$; אולם זה מעבר ליכולת שלנו בקורס זה.

מטריצות מרקוב

נסיים את הפרק בשימוש של ליכסון מטריצות לתורת ההסתברות.

דוגמה 17. מחקר הראה כי מבין שותי המיצים הטריים, מי שקנה "פריגת" בשבוע מסויים יקנה בהסתברות $\frac{1}{2}$ "פריגת" בשבוע הבא, ויקנה בהסתברות $\frac{1}{2}$ "פרימור" בשבוע הבא. מי שקנה "פרימור" בשבוע מסויים יקנה בהסתברות $\frac{3}{4}$ "פריגת" בשבוע הבא, ויקנה בהסתברות $\frac{1}{4}$ "פרימור" בשבוע הבא. בסופרמרקט מסויים הצריכה היתה בשבוע כלשהו 700 בקבוקים "פריגת" ו־ 300 בקבוקים "פרימור". מה תהיה הצריכה אחרי הרבה זמן (נאמר כעבור 10 שבועות)? מניחים כי סה"כ הצריכה השבועית נשמרת: 1000 בקבוקים. (הנתונים מומצאים כמובן.)

נסמן ב־ b_n את מספר בקבוקי "פריגת" כעבור n שבועות, ובדומה נסמן ב־ c_n מספר בקבוקי "פרימור" כעבור n שבועות. תנאי ההתחלה הם $b_0 := 700$ ו־ $c_0 := 300$. הנוסחה הרקורסיבית

לחישוב b_{n+1} ו־ c_{n+1} היא:

$$. \ b_{n+1} = \tfrac{1}{2}b_n + \tfrac{3}{4}c_n, \qquad c_{n+1} = \tfrac{1}{2}b_n + \tfrac{1}{4}c_n$$

נגדיר מטריצה $A := \begin{bmatrix} \frac{1}{2} & \frac{3}{4} \\ \frac{1}{2} & \frac{1}{4} \end{bmatrix}$. אז

$$, \ \begin{bmatrix} b_{n+1} \\ c_{n+1} \end{bmatrix} = A \cdot \begin{bmatrix} b_n \\ c_n \end{bmatrix}$$

ומשום כך

$$. \ \begin{bmatrix} b_n \\ c_n \end{bmatrix} = A^n \begin{bmatrix} b_0 \\ c_0 \end{bmatrix}$$

אנו רוצים לדעת האם קיים הגבול:

$$\lim_{n\to\infty} \begin{bmatrix} b_n \\ c_n \end{bmatrix} := \begin{bmatrix} \lim_{n\to\infty} b_n \\ \lim_{n\to\infty} c_n \end{bmatrix}$$

תחילה נבדוק האם קיים הגבול $\lim_{n\to\infty} A^n$. הכוונה היא זו: נסמן $A^n = \begin{bmatrix} a_{11;n} & a_{12;n} \\ a_{21;n} & a_{22;n} \end{bmatrix}$, כך שיש לנו 4 סדרות של מספרים ממשיים $\{a_{11;n}\}_{n=0}^{\infty}, \ldots, \{a_{22;n}\}_{n=0}^{\infty}$. האם קיימים הגבולות $\lim_{n\to\infty} a_{11;n}$, $\ldots, \lim_{n\to\infty} a_{22;n}$?

הפולינום האופייני של A הוא

$$. \ p_A(x) = x^2 - \tfrac{3}{4}x - \tfrac{1}{4} = (x-1)(x+\tfrac{1}{4})$$

הערכים העצמיים של A הם: $\lambda_1 := 1$ ו־ $\lambda_2 := -\frac{1}{4}$. וקטורים עצמיים של A הם $v_1 := \begin{bmatrix} \frac{3}{2} \\ 1 \end{bmatrix}$ ו־ $v_2 := \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. המטריצות המלכסנות הן $P := \begin{bmatrix} \frac{3}{2} & -1 \\ 1 & 1 \end{bmatrix}$ ו־ $P^{-1} := \frac{2}{5}\begin{bmatrix} 1 & 1 \\ -1 & \frac{3}{2} \end{bmatrix}$. כעת לכל n מקבלים

$$. \ A^n = PD^nP^{-1} = P \begin{bmatrix} 1 & 0 \\ 0 & (-\frac{1}{4})^n \end{bmatrix} P^{-1} = \frac{2}{5}\begin{bmatrix} \frac{3}{2} + (-\frac{1}{4})^n & \frac{3}{2} - \frac{3}{2}(-\frac{1}{4})^n \\ 1 - (-\frac{1}{4})^n & 1 + \frac{3}{2}(-\frac{1}{4})^n \end{bmatrix}$$

לכן הגבול קיים והוא

$$. \ \lim_{n\to\infty} A^n = \begin{bmatrix} \frac{3}{5} & \frac{3}{5} \\ \frac{2}{5} & \frac{2}{5} \end{bmatrix}$$

נחזור לבעיה המקורית שלנו. אחרי "הרבה זמן" **המצב היציב** הוא:

$$\lim_{n\to\infty} \begin{bmatrix} b_n \\ c_n \end{bmatrix} = \lim_{n\to\infty} (A^n \begin{bmatrix} b_0 \\ c_0 \end{bmatrix}) = (\lim_{n\to\infty} A^n) \begin{bmatrix} b_0 \\ c_0 \end{bmatrix} = \begin{bmatrix} \frac{3}{5} & \frac{3}{5} \\ \frac{2}{5} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} 700 \\ 300 \end{bmatrix} = \begin{bmatrix} 600 \\ 400 \end{bmatrix}$$

כלומר הצריכה אחרי זמן רב היא בקירוב 600 בקבוקים "פריגת" ו־ 400 בקבוקים "פרימור".

אנו רואים תופעה מעניינת, שהמצב היציב $\begin{bmatrix} 600 \\ 400 \end{bmatrix}$ איננו תלוי במצב ההתחלתי $\begin{bmatrix} b_0 \\ c_0 \end{bmatrix}$, אלא רק בכך שסה"כ הצריכה היא 1000 בקבוקים בשבוע.

הגדרה 8. מטריצת מרקוב היא מטריצה $A = [a_{ij}] \in \mathrm{M}_{n\times n}(\mathbb{R})$ כך ש־ $0 \le a_{ij} \le 1$, ובכל עמודה j סכום האיברים הוא $\sum_{i=1}^{n} a_{ij} = 1$.

התאוריה של מטריצות מרקוב אומרת שהתופעה הכללית דומה לדוגמה שלנו ; כלומר ל־ A יש ערך עצמי $\lambda_1 = 1$, ויתר הערכים העצמיים מקיימים $|\lambda_i| \leq 1$. יתר על כן אם $a_{ij} > 0$ לכל האינדקסים אז $|\lambda_i| < 1$ עבור $i \geq 2$. לכן הגבול $\lim_{k\to\infty} A^k$ קיים. בנוסף ניתן להוכיח כי

$$\lim_{k\to\infty} A^k = \begin{bmatrix} v_{\text{stab}} & \cdots & v_{\text{stab}} \end{bmatrix}$$

כאשר $v_{\text{stab}} = [b_i]$ הוא וקטור ההתפלגות היציב, שהוא וקטור עצמי השייך לערך העצמי 1 המקיים $b_i \geq 0$ ו־ $\sum_{i=1}^n b_i = 1$.

ח. מרחבי מכפלה פנימית

בפרק זה השדה הוא $\mathbb{R}$, וכל המרחבים הוקטוריים הם סוף מימדיים.

הגדרה 1. יהי V מרחב וקטורי. **תבנית בילינארית סימטרית** על V היא פונקציה

$$\langle -,- \rangle : V \times V \to \mathbb{R}$$

עם התכונות הבאות:

א. $\langle v,w \rangle = \langle w,v \rangle$ לכל $v,w \in V$.
ב. $\langle av,w \rangle = a\langle v,w \rangle$ לכל $v,w \in V$ ו- $a \in \mathbb{R}$.
ג. $\langle v_1+v_2,w \rangle = \langle v_1,w \rangle + \langle v_2,w \rangle$ לכל $v_1,v_2,w \in V$.

טענה 1. תהי $\langle -,- \rangle$ תבנית בילינארית סימטרית על V. יהיו $v_1,v_2,w \in V$ ו- $a \in \mathbb{R}$.

א. $\langle w,v_1+v_2 \rangle = \langle w,v_1 \rangle + \langle w,v_2 \rangle$.
ב. $\langle w,av \rangle = a\langle w,v \rangle$.
ג. $\langle v,\vec{0} \rangle = 0$.

הוכחה. סעיף א׳ נובע מתכונות א׳ ו- ג׳. סעיף ב׳ נובע מתכונות א׳ ו- ב׳. באשר לסעיף ג׳:

$$\langle v,\vec{0} \rangle = \langle v,0\cdot\vec{0} \rangle = 0\cdot\langle v,\vec{0} \rangle = 0 \ .$$

מש״ל.

הגדרה 2. יהי V מרחב וקטורי. **מכפלה פנימית** על V היא תבנית בילינארית סימטרית $\langle -,- \rangle$ שיש לה גם התכונה

ד. אם $v \neq \vec{0}$ אז $\langle v,v \rangle > 0$.

הזוג $(V,\langle -,- \rangle)$ נקרא **מרחב מכפלה פנימית** (או **מרחב אוקלידי**).

דוגמה 1. ניקח את המרחב הוקטורי $V := \mathbb{R}^n$. נשים לב כי כל וקטור $v \in V$ הוא מטריצה בגודל $n \times 1$, ולכן v^{t} הוא מטריצה בגודל $1 \times n$. נגדיר פונקציה $\langle -,- \rangle$ ע״י הנוסחה

$$\langle v,w \rangle := v^{\mathrm{t}} \cdot w \in \mathrm{M}_{1\times 1}(\mathbb{R}) = \mathbb{R} \ .$$

אם נרשום $v = [a_1 \ \cdots \ a_n]^{\mathrm{t}}$ ו- $w = [b_1 \ \cdots \ b_n]^{\mathrm{t}}$ אז

$$\langle v,w \rangle = \sum_{i=1}^{n} a_i b_i \ .$$

מנוסחה זו ברור כי $\langle -,- \rangle$ הינה תבנית בילינארית סימטרית. אם $v \neq \vec{0}$ אז

$$\langle v,v \rangle = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \sum_{i=1}^{n} a_i^2 > 0 \ .$$

אנו רואים שהפונקציה $\langle -,- \rangle$ היא מכפלה פנימית, הנקראת **המכפלה הפנימית הסטנדרטית** על $\mathbb{R}^n$.

דוגמה 2. שוב ניקח את המרחב הוקטורי $V := \mathbb{R}^n$. נבחר סדרת מספרים $d_1, \ldots, d_n$, ונגדיר פונקציה

$$\langle \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}, \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \rangle := \sum_{i=1}^{n} d_i a_i b_i \; .$$

אם $d_1 = \cdots = d_n = 1$ זוהי המכפלה הפנימית הסטנדרטית מהדוגמה הקודמת. אם $d_1, \ldots, d_n > 0$ עדיין יש לנו מכפלה פנימית. אולם אם $d_i \leq 0$ לאיזה i הרי

$$\langle \vec{e}_i, \vec{e}_i \rangle = d_i \leq 0 \; ,$$

ולכן זו איננה מכפלה פנימית (אלא רק תבנית בילינארית סימטרית).

דוגמה 3. יהי n שלם חיובי, ויהי V מרחב הפולינומים $f(x)$ ממעלה $\geq n$. עבור שני פולינימים f, g נגדיר

$$\langle f, g \rangle := \int_0^1 f(x) g(x) \, \mathrm{d}x \; .$$

מתכונות האינטגרל רואים כי זו תבנית בילינארית סימטרית. כעת יהי $f(x)$ פולינום שונה מאפס. אז $f(a)^2 > 0$ מלבד מספר סופי של נקודות a, ולכן

$$\langle f, f \rangle = \int_0^1 f(x)^2 \, \mathrm{d}x > 0 \; .$$

טענה 2. יהי $(V, \langle -, - \rangle)$ מרחב מכפלה פנימית, ויהי W תת־מרחב של V. אז $(W, \langle -, - \rangle)$ גם הוא מרחב מכפלה פנימית.

הוכחה.התכונות א׳ – ד׳ של הגדרות 1 ו־ 2 מתקיימות עבור וקטורים ב־ W. משי״ל.

בהמשך בדרך כלל נקצר ונאמר ״V הוא מרחב מכפלה פנימית״, ללא ציון מפורש של המכפלה הפנימית $\langle -, - \rangle$.

הערה. מרחבי מכפלה פנימית אינסוף מימדיים נקראים **מרחבי הילברט**, ועוסקים בהם בקורס ״אנליזה פונקציונלית״. יש גם גירסה של מכפלה פנימית מעל $\mathbb{C}$; השוני הוא ש־ $\langle w, v \rangle = \overline{\langle v, w \rangle}$, כלומר ההצמדה המרוכבת היא חלק מההגדרה.

בסיסים אורתונורמליים

הגדרה 3. יהי V מרחב מכפלה פנימית, ותהי $\mathbf{v} = (v_1, \ldots, v_n)$ סדרת וקטורים ב־ V. הסדרה $\mathbf{v}$ נקראת **סדרה אורתונורמלית** אם $\langle v_i, v_i \rangle = 1$ לכל i, ו־ $\langle v_i, v_j \rangle = 0$ לכל $i \neq j$.

טענה 3. יהי V מרחב מכפלה פנימית, ותהי $\mathbf{v} = (v_1, \ldots, v_n)$ סדרה אורתונורמלית. אז $\mathbf{v}$ היא סדרה בלתי תלויה ליניארית.

הוכחה. נניח על דרך השלילה כי איזה v_i הוא צרוף ליניארי של קודמיו, ז״א $v_i = \sum_{j=1}^{i-1} a_j v_j$. מקבלים

$$1 = \langle v_i, v_i \rangle = \langle v_i, \sum_{j=1}^{i-1} a_j v_j \rangle = \sum_{j=1}^{i-1} a_j \langle v_i, v_j \rangle = 0 \ .$$

זו סתירה. משל.

הגדרה 4. יהי V מרחב מכפלה פנימית. **בסיס אורתונורמלי** של V הוא סדרה אורתונורמלית $\mathbf{v}$ שהיא גם סדרה פורשת.

דוגמה 4. עבור $\mathbb{R}^n$ עם המכפלה הפנימית הסטנדרטית, הבסיס הסטנדרטי $\mathbf{e} = (\vec{e}_1, \ldots, \vec{e}_n)$ הוא בסיס אורתונורמלי.

טענה 4. יהי V מרחב מכפלה פנימית, יהי $\mathbf{v} = (v_1, \ldots, v_n)$ בסיס אורתונורמלי, ויהי w וקטור כלשהו ב־ V. אז

$$w = \sum_{i=1}^{n} \langle v_i, w \rangle \cdot v_i \ .$$

הוכחה. נרשום $w = \sum_{i=1}^{n} a_i v_i$, ונחשב:

$$\langle v_i, w \rangle = \langle v_i, \sum_{j=1}^{n} a_j v_j \rangle = \sum_{j=1}^{n} a_j \cdot \langle v_i, v_j \rangle = a_i \ .$$

משל.

משפט 1. (תהליך גראם־שמידט) יהי V מרחב מכפלה פנימית ותהי $\mathbf{v} = (v_1, \ldots, v_n)$ סדרה בלתי תלויה ליניארית ב־ V. אז ישנה סדרה אורתונורמלית $\mathbf{w} = (w_1, \ldots, w_n)$ ב־ V עם התכונה הבאה: לכל i מתקיים

$$\mathrm{Sp}(w_1, \ldots, w_i) = \mathrm{Sp}(v_1, \ldots, v_i) \ .$$

הוכחה. ההוכחה היא באינדוקציה על n.

עבור $n = 1$ יהי $a := \langle v_1, v_1 \rangle$. מאחר ש־ $v_1 \neq 0$ הרי a מספר חיובי, ונגדיר $w_1 := a^{-1/2} v_1$. אז

$$\langle w_1, w_1 \rangle = \langle a^{-1/2} v_1, a^{-1/2} v_1 \rangle = a^{-1} \cdot \langle v_1, v_1 \rangle = 1 \ ,$$

כלומר (w_1) סדרה אורתונורמלית. עניין מרחב הפרישה ברור.

עתה נניח כי $n \geq 2$, וכי המשפט נכון לסדרות באורך $n-1$. נתבונן בסדרה $(v_1, \ldots, v_{n-1})$. ע״פ ההנחה ישנה סדרה אורתונורמלית $(w_1, \ldots, w_{n-1})$ כך ש־

$$\mathrm{Sp}(w_1, \ldots, w_i) = \mathrm{Sp}(v_1, \ldots, v_i)$$

לכל $1 \leq i \leq n-1$. נגדיר

$$w'_n := v_n - \sum_{i=1}^{n-1} \langle w_i, v_n \rangle \cdot w_i \ .$$

מאחר ש־
$$v_n \notin \mathrm{Sp}(w_1, \ldots, w_{n-1})$$
הרי $w'_n \neq \vec{0}$, ולכן $a := \langle w'_n, w'_n \rangle$ מספר חיובי. נגדיר $w_n := a^{-1/2} w'_n$. ברור כי
$$\mathrm{Sp}(w_1, \ldots, w_n) = \mathrm{Sp}(v_1, \ldots, v_n)$$
וכי $\langle w_n, w_n \rangle = 1$. נותר להוכיח כי $\langle w_n, w_i \rangle = 0$ לכל $i < n$. נחשב:
$$\begin{aligned}
\langle w_n, w_i \rangle &= \langle a^{-1/2} w'_n, w_i \rangle \\
&= a^{-1/2} \langle v_n - \sum_{j=1}^{n-1} \langle w_j, v_n \rangle \cdot w_j, w_i \rangle \\
&= a^{-1/2} \langle v_n, w_i \rangle - a^{-1/2} \sum_{j=1}^{n-1} \langle w_j, v_n \rangle \langle w_j, w_i \rangle \\
&= a^{-1/2} \langle v_n, w_i \rangle - a^{-1/2} \langle w_i, v_n \rangle = 0
\end{aligned}$$
מש״ל.

מסקנה 1. יהי V מרחב מכפלה פנימית. אז יש ל־ V בסיס אורתונורמלי.

הוכחה. נבחר בסיס כלשהו $\mathbf{v}$ של V, ותהי $\mathbf{w}$ הסדרה המתקבלת מ־ $\mathbf{v}$ בתהליך גראם־שמידט. אז $\mathbf{w}$ היא בסיס אורתונורמלי. מש״ל.

מסקנה 2. יהי V מרחב מכפלה פנימית ממימד n, ויהי W תת־מרחב של V ממימד m. אז ישנו בסיס אורתונורמלי $(w_1, \ldots, w_n)$ של V, כך ש־ $(w_1, \ldots, w_m)$ הוא בסיס אורתונורמלי של W.

הוכחה. נבחר בסיס כלשהו $(v_1, \ldots, v_m)$ של W, ונשלים אותו לבסיס $\mathbf{v} = (v_1, \ldots, v_m, \ldots, v_n)$ של V. תהי $(w_1, \ldots, w_n)$ הסדרה המתקבלת מ־ $\mathbf{v}$ בתהליך גראם־שמידט. לסדרה זו התכונות הדרושות. מש״ל.

דוגמה 5. נתבונן במרחב
$$V := \{\text{פולינומים } f(x) \text{ ממעלה } \leq 1\}$$
עם המכפלה הפנימית
$$\langle f, g \rangle = \int_0^1 f(x) g(x) \, \mathrm{d}x \ .$$
ננסה למצוא בסיס אורתונורמלי ל־ V. נתחיל מהבסיס
$$\mathbf{v} = (v_1, v_2) := (1, x) \ ,$$
ובעזרת תהליך גראם־שמידט נמצא בסיס אורתונורמלי $\mathbf{w} = (w_1, w_2)$.
מאחר ש־
$$\langle v_1, v_1 \rangle = \int_0^1 1 \cdot \, \mathrm{d}x = 1$$
מקבלים $w_1 = v_1$.

בשלב הבא נחשב

$$. \langle w_1, v_2 \rangle = \int_0^1 1 \cdot x \cdot \mathrm{d}x = \tfrac{1}{2}$$

לכן

$$w_2' := v_2 - \langle w_1, v_2 \rangle \cdot w_1 = x - \tfrac{1}{2} \cdot 1 = x - \tfrac{1}{2}$$

עכשיו נחשב

$$. \langle w_1', w_1' \rangle = \int_0^1 (x - \tfrac{1}{2})^2 \,\mathrm{d}x = \tfrac{1}{12}$$

מקבלים $w_2 = \sqrt{12}(x - \tfrac{1}{2})$.

הטלות מאונכות

הגדרה 5. יהי V מרחב מכפלה פנימית, ויהי W תת־מרחב של V. **המרחב המאונך** ל־ W הוא הקבוצה

$$. W^\perp := \{ v \in V \mid \text{לכל } w \in W \ \langle v, w \rangle = 0 \ \}$$

חשבון קצר מראה כי הקבוצה $W^\perp$ היא תת־מרחב של V.

טענה 5. יהי V מרחב מכפלה פנימית, ויהי W תת־מרחב של V. אז יש בסיס אורתונורמלי $(w_1, \ldots, w_n)$ של V, כך ש־ $(w_1, \ldots, w_m)$ הוא בסיס אורתונורמלי של W, ו־ $(w_{m+1}, \ldots, w_n)$ בסיס אורתונורמלי של $W^\perp$.

הוכחה. ניקח את הבסיס האורתונורמלי $(w_1, \ldots, w_n)$ ממסקנה 2. ידוע כי

$$. \ \mathrm{Sp}(w_1, \ldots, w_m) = W$$

צריך להוכיח כי

$$. \ \mathrm{Sp}(w_{m+1}, \ldots, w_n) = W^\perp$$

ההכלה לכוון אחד ברורה: לכל $j \le m$ ולכל $i > m$ מתקיים $\langle w_j, w_i \rangle = 0$, ולכן $\langle w, w_i \rangle = 0$ לכל $w \in W$; כלומר $w_i \in W^\perp$. מכך מסיקים ש־

$$. \ \mathrm{Sp}(w_{m+1}, \ldots, w_n) \subset W^\perp$$

כעת ניקח $v \in W^\perp$. אז $\langle w_i, v \rangle = 0$ לכל $i \le m$. בעזרת טענה 4 אנו מקבלים

$$. \ v = \sum_{i=1}^{n} \langle w_i, v \rangle \cdot w_i = \sum_{i=m+1}^{n} \langle w_i, v \rangle \cdot w_i \in \mathrm{Sp}(w_{m+1}, \ldots, w_n)$$

ז״א ש־

$$. \ W^\perp \subset \mathrm{Sp}(w_{m+1}, \ldots, w_n)$$

משי״ל.

משפט 2. יהי V מרחב מכפלה פנימית, ויהי W תת־מרחב של V. אז ישנו אופרטור ליניארי יחיד $P_W : V \to V$ בעל התכונות הבאות:

א. $\mathrm{Ker}(P_W) = W^\perp$.
ב. $\mathrm{Im}(P_W) = W$.
ג. $P_W(w) = w$ לכל $w \in W$.

האופרטור P_W נקרא **אופרטור ההטלה המאונכת** על W.

הוכחה. נתחיל בהוכחת היחידות. נניח שיש שני אופרטורים $P, Q : V \to V$ שיש להם התכונות א׳־ג׳. אנו נראה כי $P = Q$. יהי $(w_1, \ldots, w_n)$ בסיס אורתונורמלי של V כמו בטענה 5. די להראות כי $(P - Q)(w_i) = \vec{0}$ לכל i. אם $i \leq m$ אז

$$(P - Q)(w_i) = P(w_i) - Q(w_i) = w_i - w_i = \vec{0} .$$

מצד שני אם $i > m$ אז

$$(P - Q)(w_i) = P(w_i) - Q(w_i) = \vec{0} - \vec{0} = \vec{0} .$$

עתה נטפל בסוגיית קיום האופרטור P_W. עבור $v \in V$ נגדיר

$$P_W(v) := \sum_{i=1}^{m} \langle w_i, v \rangle \cdot w_i \in V .$$

בדיקה קלה מראה שזהו אופרטור ליניארי, והתמונה שלו מוכלת ב־ W.
נחשב את הגרעין של P_W. יהי $v \in V$ וקטור כלשהו. מן ההגדרה של P_W נובע ש־ $P_W(v) = 0$ אס״ם $\langle w_i, v \rangle = 0$ לכל $1 \leq i \leq m$. מאחר ש־ $W = \mathrm{Sp}(w_1, \ldots, w_m)$ אנו מסיקים ש־ $\mathrm{Ker}(P_W) = W^\perp$.
לבסוף ניקח $w \in W$. אז $\langle w_i, w \rangle = 0$ לכל $i > m$, ולכן

$$w = \sum_{i=1}^{n} \langle w_i, w \rangle \cdot w_i = \sum_{i=1}^{m} \langle w_i, w \rangle \cdot w_i = P_W(w) .$$

זה גם מוכיח כי $\mathrm{Im}(P_W) = W$. משי״ל.

הערה. אופרטור ההטלה המאונכת P_W ניתן לליכסון. בתור בסיס מלכסן ניתן לקחת את הבסיס האורתונורמלי $(w_1, \ldots, w_n)$ בו השתמשנו בהוכחה. אז $w_1, \ldots, w_m$ הם וקטורים עצמיים של P_W השייכים לערך העצמי 1, בעוד ש־ $w_{m+1}, \ldots, w_n$ הם וקטורים עצמיים של P_W השייכים לערך העצמי 0.

הגיאומטריה של מרחבי מכפלה פנימית

יהי V מרחב מכפלה פנימית. אם v הוא וקטור שונה מ־ $\vec{0}$ ב־ V אז $\langle v, v \rangle > 0$. מצד שני $\langle \vec{0}, \vec{0} \rangle = 0$. רואים שלכל וקטור מתקיים $\langle v, v \rangle \geq 0$.

הגדרה 6. יהי V מרחב מכפלה פנימית. **האורך** של וקטור $v \in V$ הוא

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

ברור כי $\|v\| = 0$ אס״ם $v = \vec{0}$.

הגדרה 7. יהי V מרחב מכפלה פנימית. **המרחק** בין שני וקטורים $v, w \in V$ הוא
$$d(v,w) := \|v-w\| \in \mathbb{R}\ .$$

דוגמה 6. נתבונן במישור $\mathbb{R}^2$ עם המכפלה הפנימית הסטנדרטית. יהי $v = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ ו־ $r := \|v\|$. אז
$$r^2 = \langle v, v\rangle = \begin{bmatrix} a_1 & a_2 \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = a_1^2 + a_2^2\ ,$$
ולכן
$$r = \sqrt{a_1^2 + a_2^2}\ .$$
לפי משפט פיתגורס זהו בדיוק האורך של v, במובן הגיאומטרי.

אפשר להראות כי המרחק $d(v,w)$ זהה למרחק בין שתי הנקודות v, w, שוב במובן הגיאומטרי. וכן
$$\langle v, w\rangle = \|v\| \cdot \|w\| \cdot \cos(\alpha)\ ,$$
כאשר α היא הזוית (בראשית הצירים) בין שני הוקטורים האלו.

טענה 6. יהי V מרחב מכפלה פנימית, יהי $(w_1, \ldots, w_n)$ בסיס אורתונורמלי של V, ויהי v וקטור ב־ V. נרשום $v = \sum_{i=1}^n a_i w_i$. אז
$$\|v\| = \Big(\sum_{i=1}^n a_i^2 \Big)^{1/2}\ .$$

הוכחה. נחשב:
$$\begin{aligned} \|v\|^2 = \langle v, v\rangle &= \langle \sum_{i=1}^n a_i w_i, \sum_{i=1}^n a_i w_i \rangle \\ &= \sum_{i,j=1}^n a_i a_j \cdot \langle w_i, w_j\rangle = \sum_{i=1}^n a_i^2 \end{aligned}$$
מש"ל.

משפט 3. יהי V מרחב מכפלה פנימית, ויהי W תת־מרחב של V, עם אופרטור הטלה מאונכת P_W. יהי v וקטור ב־ V, ונגדיר $w := P_W(v)$. אז w הוא הוקטור ב־ W אשר מרחקו מ־ v מינימלי. ליתר דיוק, לכל וקטור w' ב־ W השונה מ־ w מתקיים
$$d(v, w') > d(v, w)\ .$$

הוכחה. נבחר בסיס אורתונורמלי $(w_1, \ldots, w_n)$ כמו בטענה 5, ונרשום $v = \sum_{i=1}^n a_i w_i$. אז $w = \sum_{i=1}^m a_i w_i$, $v - w = \sum_{i=m+1}^n a_i w_i$, ובעזרת הטענה הקודמת מקבלים
$$d(v,w)^2 = \|v-w\|^2 = \sum_{i=m+1}^n a_i^2\ .$$

כעת נרשום $w' = \sum_{i=1}^m b_i w_i$. מאחר ש־ $w \neq w'$ הרי לפחות לאינדקס אחד מתקיים $a_i \neq b_i$. ההפרש הוא

$$v - w' = \sum_{i=1}^{m}(a_i - b_i)w_i + \sum_{i=m+1}^{n} a_i w_i \ ,$$

ולכן

$$d(v, w')^2 = \|v - w'\|^2 = \sum_{i=1}^{m}(a_i - b_i)^2 + \sum_{i=m+1}^{n} a_i^2 \ .$$

מאחר ש־ $\sum_{i=1}^m (a_i - b_i)^2 > 0$ נובע ש־ $d(v, w') > d(v, w)$. משי"ל.

שיטת הריבועים הפחותים

לסיום הנה שימוש של הטלות מאונכות לפתרון משוואות ליניאריות לא הומוגניות. נתונה מערכת משוואות ליניאריות לא הומוגניות של m משוואות ו־ n משתנים

$$AX = v \ .$$

כלומר A מטריצה בגודל $m \times n$ ו־ $v \in \mathbb{R}^m$. יהי W מרחב העמודות של A. ידוע לנו שקיים פתרון למערכת המשוואות $AX = v$ אם"ם $v \in W$.

עתה נניח כי $v \notin W$. הרעיון הוא לחפש וקטור $u \in \mathbb{R}^n$ שהוא הקירוב הטוב ביותר לפתרון של $AX = v$. כלומר אנו רוצים למצוא וקטור u המקיים

$$d(v, Au) = \min\{d(v, Au') \mid u' \in \mathbb{R}^n\} \ .$$

כאן המרחק $d(-, -)$ הוא ביחס למכפלה הפנימית הסטנדרטית של $\mathbb{R}^m$. נשים לב כי

$$\{Au' \mid u' \in \mathbb{R}^n\} = W \ .$$

לכן התנאי על u הינו שהוקטור $w := Au$ יקיים

$$d(v, w) = \min\{d(v, w') \mid w' \in W\} \ .$$

לפי משפט 3 אין ברירה אלא $w = P_W(v)$, כאשר P_W הוא אופרטור ההטלה המאונכת על W. זאת אומרת ש־ u הוא פתרון של מערכת המשוואות

$$AX = P_W(v) \ .$$

שיטת קירוב זאת נקראת **שיטת הריבועים הפחותים** (least squares approximation). יש לשיטה זו שימושים רבים, במיוחד בסטטיסטיקה.