אלגברה לינארית

ד"ר יונתן שלח, ד"ר נדב מאיר

11 בינואר 2026

תוכן העניינים

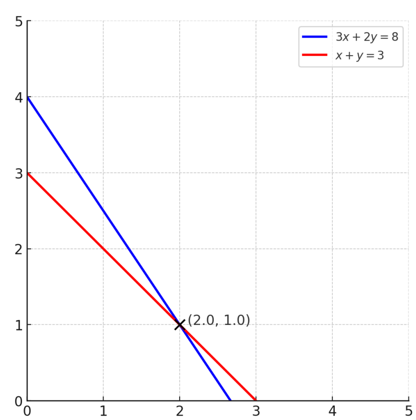
פרק 1

הקדמה

אלגברה לינארית היא אחת מאבני היסוד של המתמטיקה המודרנית. בתור התחלה, היא קשורה באופן הדוק לגיאומטריה של המישור וגם לגיאומטריה של המרחב. בחטיבת הביניים לומדים על משוואה בנעלם אחד, וגם על מערכת של שתי משוואות בשני נעלמים. למשל:

$$\begin{cases} 3x + 2y = 8 \\ x + y = 3 \end{cases}$$

יש יותר מדרך אחת לפתור מערכת כזו, כמו להכפיל את המשוואה השנייה ב-2 ואז להחסיר אותה מהמשוואה הראשונה. כך מקבלים $x = 2$ ולבסוף $y = 1$ ע"י הצבה באחת המשוואות. מבחינה גיאומטרית, כל משוואה מייצגת קו ישר במישור, ופתרון המערכת מוביל לנקודת החיתוך של שני הישרים.



איור 1.1: שני הישרים ונקודת החיתוך

בבסיסה אלגברה לינארית עוסקת בפתרון מערכת של משוואות לינאריות, כלומר משוואות ממעלה ראשונה במספר נעלמים. בתחומים רבים במדע ובכלל יכולים להופיע הרבה נעלמים והרבה אילוצים (משוואות), בהתאם למה שידוע לנו על הבעיה הרלוונטית.

דוגמה 1.1. תומר נוסע מאילת צפונה, ואלה נוסעת מתל אביב דרומה. המרחק ההתחלתי ביניהם הוא 350 ק"מ. תומר נוסע במהירות 60 קמ"ש במשך t_1 שעות, וממשיך במהירות 90 קמ"ש במשך t_2 שעות עד שהוא חולף על פני המכונית של אלה ומזהה אותה. עד לרגע זה אלה נסעה במהירות 90 ק"מ במשך t_3 שעות ואחר כך במהירות 110 ק"מ במשך t_4 שעות. אם נשווה בין סכומי זמני התנועה של שתי המכוניות וגם נדרוש שסכום הדרכים יהיה המרחק ההתחלתי, נקבל שתי משוואות בארבעה נעלמים:

$$\begin{cases} t_1 + t_2 = t_3 + t_4 \\ 60t_1 + 90t_2 + 90t_3 + 110t_4 = 350 \end{cases}$$

למערכת משוואות לינארית (ממ"ל בראשי תיבות) זו יש אינסוף פתרונות בתחום ההגדרה הרלוונטי של מספרים חיוביים. אם למשל נציב $t_3 = t_4 = 1$ בשתי המשוואות, נקבל ממ"ל חדשה של שתי משוואות בשני נעלמים:

$$\begin{cases} t_1 + t_2 = 2 \\ 60t_1 + 90t_2 = 150 \end{cases}$$

אפשר לפתור את הממ"ל הזו כרגיל, אבל קל לבדוק שהפתרון הוא $t_1 = t_2 = 1$. לכן, אחד הפתרונות של הממ"ל המקורית (בארבעה נעלמים) הוא $t_1 = t_2 = t_3 = t_4 = 1$. אבל זה לא הפתרון היחיד כי בחרנו את הערכים של t_3, t_4 באופן שרירותי (נוח לחישובים, אך לא יותר מזה). באותה מידה אפשר להציב $t_3 = \frac{2}{3},\ t_4 = \frac{4}{3}$ ולקבל ממ"ל קצת שונה מהקודמת:

$$\begin{cases} t_1 + t_2 = 2 \\ 60t_1 + 90t_2 = \frac{430}{3} \end{cases}$$

הפעם אפשר לבדוק שהפתרון הוא $t_1 = \frac{11}{9},\ t_2 = \frac{7}{9}$. כלומר יש פתרון נוסף לממ"ל המקורית, שנתון ע"י $t_1 = \frac{11}{9},\ t_2 = \frac{7}{9}, t_3 = \frac{2}{3},\ t_4 = \frac{4}{3}$.

נראה בפרק 5 איך אפשר לכתוב את קבוצת הפתרונות באופן כללי, אבל כבר אפשר להבין שהיא אינסופית כי יש לנו חופש לבחור את ערכי t_3, t_4. אמנם לא מדובר בחופש מוחלט כי כל ארבעת הזמנים צריכים להיות חיוביים (מה שמקטין את קבוצת הפתרונות), אבל עדיין יש פה מספיק חופש לאינסוף פתרונות.

במובן מסוים אפשר לומר שאלה "מחליטה" בשביל תומר על זמני הנסיעה שלו. זו דרך הסתכלות שרירותית ומאוד טכנית - אפשר לחשוב על פתרון המערכת גם באופן הפוך ולתת את החופש לתומר. בפועל (מחוץ לעולם האלגברי), תומר ואלה לא שמו לב אחד לשנייה עד לרגע

הפגישה המקרית. להם יש מידע מלא על זמני הנסיעה, אבל אנחנו מסתפקים במידע חלקי. אם היו לנו מספיק משוואות נוספות (לפחות שתיים), היה ניתן להגיע לפתרון המדויק שמתאר את תנועת המכוניות. אבל לא תמיד יש לנו מספיק מידע ונלמד להתייחס לכך בהתאם.

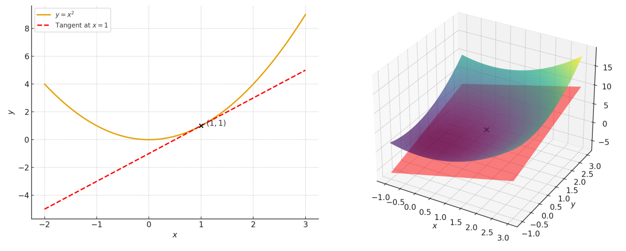
באופן כללי, נשאל את השאלות הבאות:

- כיצד פותרים ממ״ל שבה הרבה נעלמים?
- כיצד פותרים ממ״ל שבה הרבה משוואות?
- האם בכלל קיימים פתרונות לממ״ל? אם כן, אז כמה?

כדי לענות על שאלות כאלו באופן מלא ומסודר, ישמשו אותנו שני מושגים יסודיים: וקטורים ומטריצות. נדחה את ההגדרות שלהם לפרקים הרלוונטיים, אבל לעת עתה מספיק לחשוב עליהם כאובייקטים מתמטיים שבהם מופיעים כמה מספרים. למשל, הזוג הסדור (x,y) של שני מספרים x ו-y הוא וקטור עם שני רכיבים (קוארדינטות). בחטיבת הביניים ובתיכון ראינו שאפשר לחשוב על זוג כזה כעל נקודה סטטית במישור, אבל בלימודי הפיזיקה יש הסתכלות דינמית על וקטור כאובייקט בעל גודל וכיוון (הדוגמה הקלאסית היא כוח שפועל על גוף). נראה שאפשר לאמץ את שתי הגישות (סטטית ודינמית) במקביל, כאשר האינטואיציה הפיזיקלית מועילה מאוד אך ממש לא הכרחית להבנת הקורס.

לאחר שנתרגל לוקטורים ומטריצות, נראה שאפשר להסתכל עליהם באופן מופשט (תיאורטי) ולשאול עליהם כל מיני שאלות שלא בהכרח קשורות לממ״ל כזו או אחרת. במתמטיקה, דבר אחד מוביל למשנהו וזה טוב לשמור על ראש פתוח כשלומדים מושגים חדשים. הגישה המופשטת של אלגברה לינארית מאפשרת יישומים מגוונים, גם מחוץ לגיאומטריה ופיזיקה. לטובת הסקרנים, נעסוק קצת בפיזיקה דרך הנדסה בחלק מהיישומים בסוף הספר שחורגים מהקורס עצמו. אבל גם נעסוק ביישומים שקשורים למאגר גדול של נתונים כמו למידת מכונה (AI).

רבים מכם לומדים חשבון דיפרנציאלי ואינטגרלי במקביל. אפשר לומר שאלגברה לינארית מופיעה בחשבון דיפרנציאלי ואינטגרלי הרבה יותר מאשר להיפך. מושג הנגזרת מוביל לקירוב לינארי של פונקציה נתונה ע״י פונקציה לינארית שמתארת את הישר המשיק לגרף הפונקציה בנקודה נתונה. הרעיון הזה תקף גם במימדים יותר גבוהים, בפרט עבור פונקציה של שני משתנים שבה הקירוב הלינארי מתאר את המישור המשיק לגרף הפונקציה בנקודה נתונה. בנוסף, הקשר לאלגברה לינארית מתבטא בחישוב שטחים (במישור) ונפחים (במרחב) בעזרת כלי שנקרא דטרמיננטה. נפתח אותו בפרק 7.



איור 1.2: מישור משיק בצד ימין, ישר משיק צד שמאל

בספר מופיעות דוגמאות רבות, תרגילים פתורים וגם קישורים לסרטונים. תוכלו לחזור אליו בהמשך התואר ככל שתצטרכו להשתמש באלגברה לינארית.

פרק 2

קבוצות ומספרים

2.1 קבוצת המספרים הטבעיים והגדרת קבוצה

המתמטיקה מתחילה בחשבון, וחשבון מבוסס על ספירה. המספר 1 מתאר את היחידה הבסיסית, לצורך העניין אצבע אחת (שבעזרתה אפשר לספור כל מיני דברים). ניתן להוסיף 1 לספירה כאוות נפשנו, ואם נמשיך כך לנצח (בדמיון שלנו לפחות) נייצר אינסוף מספרים שלא בהכרח נדע איך לקרוא להם כי השמות יהיו ארוכים מאוד. אבל יש שם לקבוצה של כל המספרים האלה: מספרים טבעיים. הסימון המקובל הוא $\mathbb{N}$, ואפשר למנות את איברי הקבוצה באופן הבא:

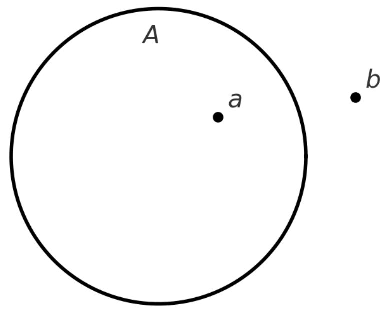
$$\mathbb{N} = \{\, 1, 1+1, 1+1+1, ... \,\} = \{\, 1, 2, 3, ... \,\}$$

הסוגריים המסולסלים מתארים קבוצה שאיבריה מופיעים בין הסוגריים ומופרדים ע"י פסיקים. מאחר שהקבוצה אינסופית, אנחנו נאלצים לכתוב "..." מתוך הנחה שברור איך להמשיך ע"י הוספת 1 בכל שלב.

הערה 2.1. לפעמים גם 0 נחשב מספר טבעי, אבל לא בקורס שלנו וזה לא באמת חשוב. מבחינה היסטורית ופילוסופית, 0 הוא מספר מוזר ומיוחד כי לא רואים אותו בטבע. הוא מתאר את מה שאינו.

הגדרה 2.2. קבוצה היא אוסף כלשהו של איברים (לא בהכרח מספרים) ללא חשיבות לסדר הופעתם.

זו הגדרה מאוד כללית, ובלבד שיהיה ברור מהגדרת הקבוצה אילו איברים שייכים לה ואילו לא. אם הקבוצה היא A, אז נכתוב $a \in A$ כדי לומר שהאיבר a שייך לקבוצה A. אם הוא לא שייך לה, נכתוב $a \notin A$.



איור 2.1: העיגול הוא הקבוצה A ומתקיים $a \in A, b \notin A$

דוגמה 2.3. *משפחת לוי כוללת את דרור (האב), אורית (האם) ואיתי (הבן). נקרא לקבוצת אנשים זו L ונתאר אותה באופן מפורש תוך שימוש בעברית:*

$$L = \{ \text{דרור, איתי, אורית} \}$$

בפרט, מתקיים

$$\text{איתי} \in L$$

אך

$$\text{עופר} \notin L$$

כי מבין השניים האלה רק איתי שייך למשפחה. נדגיש שאין חשיבות לסדר האיברים בתוך הקבוצה, ובדרך כלל יש יותר מדרך אחת להציג את הקבוצה. למשל, כאן גם מתקיים:

$$L = \{ \text{דרור, אורית, איתי} \} = \{ \text{דרור, איתי, אורית} \}$$

תרגיל 2.4. האם מתקיים $100^{100} \in \mathbb{N}$? $100^{-100} \in \mathbb{N}$? $(-100)^{100} \in \mathbb{N}$?

פתרון. מתקיים $100^{100} \in \mathbb{N}$ כי מכפלה של מספרים טבעיים היא תמיד מספר טבעי. בנוסף, $(-100)^{100} \in \mathbb{N}$ כי $(-100)^{100} = 100^{100}$ לפי הזוגיות של המעריך 100.
אבל $100^{-100} \notin \mathbb{N}$ כי $100^{-100} = \frac{1}{10^{100}}$, וכל הופכי של מספר טבעי (שאינו 1) אינו מספר טבעי.

ניתן להגדיר תת-קבוצות של $\mathbb{N}$ ע"י שימוש בתכונות. למשל, את קבוצת המספרים הזוגיים החיוביים $2\mathbb{N}$ ניתן להגדיר באופן הבא:

$$2\mathbb{N} = \{ n \in \mathbb{N} \mid \text{מתחלק ב-2 } n \} = \{ 2, 4, 6, ... \}$$

הסימן $\in$ מתאר שייכות, ולכן הנוסחה $n \in \mathbb{N}$ פירושה שהמספר n שייך לקבוצה $\mathbb{N}$. הקו $|$ שמפריד בין הנוסחה לתנאי, פירושו "כך ש-". לכן, הגדרת הקבוצה אומרת לנו שמדובר בקבוצת

כל המספרים מהצורה $n \in \mathbb{N}$ כך ש-n מתחלק ב-2. אפשר גם להגדיר את הקבוצה הזו ע"י נוסחה מפורשת $2n$ שתפיק את כל המספרים הזוגיים החיוביים כאשר נציב בה כל מספר מהצורה $n \in \mathbb{N}$. הפעם הכתיבה היא כדלקמן:

$$2\mathbb{N} = \{\, 2n \mid n \in \mathbb{N} \,\}$$

שימו לב ששתי צורות הכתיבה דומות (נוסחה בצד שמאל ותנאי בצד ימין, עם קו מפריד באמצע). ההבדל הוא שבצורה הראשונה הנוסחה מתייחסת לקבוצה ידועה ($\mathbb{N}$) והתנאי דרוש כדי לקבוע את השייכות לקבוצה החדשה. בצורה השנייה יש נוסחה של משתנה n, והתנאי מתייחס לערכי המשתנה -שיש להציב בנוסחה (כאן התנאי הוא זה שמתייחס ל $\mathbb{N}$). בהמשך הפרק נוכיח שאכן שתי ההגדרות של קבוצת הזוגיים הן הגדרות שקולות, כלומר שתיהן אכן מתארות את המספרים $2, 4, 6, ...$ ושום מספר אחר. זה אולי כבר נראה ברור, אבל נשאלת השאלה איך כותבים הוכחה מסודרת.

2.2 קבוצות מספרים נוספות

ב-$\mathbb{N}$ יש פעולות חיבור וכפל. הפעולות ההפוכות, חיסור וחילוק, דורשות הרחבה של $\mathbb{N}$ לקבוצות יותר גדולות. למשל, למשוואה $x + 2 = 1$ אין פתרון טבעי ואנחנו יודעים שאפשר לפתור אותה ע"י חיסור 2 מכל אגף ואז הפתרון יהיה $x = -1$, שהוא מספר שלילי. בעצם, מגדירים את -1 כפתרון של המשוואה $x + 1 = 0$ וזה מוביל להגדרה של המספרים השלמים:

$$\mathbb{Z} = \{\, ..., -3, -2, -1, 0, 1, 2, 3, ... \,\} = \{\, n - m \mid n, m \in \mathbb{N} \,\}$$

שימו לב לכתיבה בצד ימין. זו דרך להגדיר את קבוצת השלמים בעזרת קבוצת הטבעיים, כאשר הכוונה היא לקבוצת כל המספרים מהצורה $n - m$ כאשר n, m הם מספרים טבעיים (פה יש שני משתנים). יש יותר מדרך אחת לכתוב מספר שלם כהפרש של מספרים טבעיים (למעשה אינסוף).

ניתן גם להרחיב את $\mathbb{Z}$ כך שיהיה ניתן לבצע חילוק. בתור התחלה מגדירים לכל $m \neq 0$ שלם את המספר ההופכי $\frac{1}{m}$, וכדי לאפשר כפל גם מוסיפים את כל השברים מהצורה $\frac{n}{m}$ כאשר $n \in \mathbb{Z}$. כך מקבלים את קבוצת המספרים הרציונליים:

$$\mathbb{Q} = \left\{\, \frac{n}{m} \;\middle|\; m, n \in \mathbb{Z},\, m \neq 0 \,\right\}$$

גם כאן יש אינסוף דרכים להציג מספר רציונלי כמנה של מספרים שלמים, ואין עם זה בעיה מבחינת הגדרת הקבוצה. אנחנו רגילים להצגה הפשוטה ביותר, לאחר צמצום המחלקים המשותפים של המונה והמכנה.

זה מביא אותנו לקבוצת המספרים הממשיים $\mathbb{R}$, שהיא הקבוצה העיקרית שנתמקד בה בקורס. קבוצה זו מכילה את קבוצת המספרים הרציונליים (כל מספר רציונלי הוא ממשי), אבל יש בה מספרים נוספים שנקראים אי-רציונליים. יש הרבה מה לומר על מספרים ממשיים, אבל הדיון

המלא מתאים לקורס בחשבון דיפרנציאלי ואינטגרלי. אז נסתפק באפיון הבא: ניתן להציג כל $x \in \mathbb{R}$ בהצגה עשרונית מהצורה

$$x = \pm d_n...d_2 d_1 d_0.d_{-1} d_{-2} d_{-3}...$$

כאשר $d_n, d_{n-1}, ..., d_0, d_{-1}, d_{-2}, ...$ הן ספרות עשרוניות בין 0 ל-9 (כולל), והנקודה העשרונית מפרידה בין החלק השלם משמאל לחלק השברי מימין. x הוא רציונלי כאשר סדרת הספרות העשרוניות היא מחזורית החל משלב מסוים. למשל:

$$\frac{1}{5} = 0.200000000000...$$
$$\frac{1}{6} = 0.166666666666...$$
$$\frac{1}{7} = 0.142857142857...$$

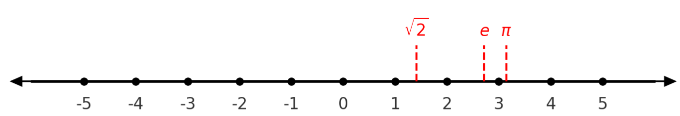
במקרה הראשון הספרה 0 חוזרת על עצמה (אפשר להשמיט אותה ולקבל הצגה סופית), במקרה השני הספרה 6 חוזרת על עצמה, ובמקרה השלישי הרצף 142857 חוזר על עצמו. המחזוריות נובעת מאופן החישוב של הספרות העשרוניות (לפי חילוק ארוך).

המספר הוא אי-רציונלי כאשר אין מחזוריות באף שלב של ההצגה העשרונית, ואז החוקיות של סדרת הספרות העשרוניות עלולה להיות מסובכת מאוד. למשל:

$$\sqrt{2} = 1.4142135623730950488016887242096980785696718753 7...$$
$$\pi = 3.1415926535897932384626433832795028841971693993 7...$$

נדגיש שאלה מספרים אי-רציונליים מיוחדים כי יש להם משמעות גיאומטרית. $\sqrt{2}$ הוא אורך היתר של משולש ישר-זווית עם שני ניצבים באורך 1, לפי משפט פיתגורס. π הוא היקף מעגל שקוטרו באורך 1.

מבחינת הקורס, המספרים האי-רציונליים הרלוונטיים הם בעיקר שורשים כמו $\sqrt{2}$ שהם יחסית נוחים לחישובים. אבל טוב לזכור שיש המון מספרים אי-רציונליים (יותר מאשר מספרים רציונליים במובן מסוים), והם משלימים את המספרים הרציונליים למה שנקרא הישר הממשי. ניתן לחשוב על המספרים כנקודות על ישר עם ראשית 0. המספרים החיוביים מופיעים בצד ימין, ואילו המספרים השליליים מופיעים בצד שמאל.



איור 2.2: הישר הממשי וחלק מאיבריו

נשארה עוד קבוצה אחת, שהיא הגדולה ביותר מבין קבוצות המספרים שנעסוק בהן. באופן דומה להגדרת -1 כשורש (פתרון) של המשוואה $x+1=0$, אפשר להגדיר את i כשורש של המשוואה $x^2+1=0$. זהו מספר מדומה, שהרי אין פתרון ממשי למשוואה זו (ערך המינימום של הפונקציה הוא 1). המספר i לא ניתן למדידה במציאות אך הוא שימושי מאוד במתמטיקה, פיזיקה וחלק מההנדסות. לפי ההגדרה מתקיים $i^2=-1$ וזה מספיק כדי להגדיר את קבוצת המספרים המרוכבים ואת פעולות החשבון המתאימות לה.

$$\mathbb{C}=\{\,a+bi \mid a,b\in\mathbb{R}\,\}$$

ההצגה $a+bi$ עבור $a,b\in\mathbb{R}$ הינה יחידה. כלומר, אם מתקיים $a+bi=c+di$ עבור $c,d\in\mathbb{R}$, אז בהכרח $a=c,\ b=d$.

הגדרה 2.5. עבור $z=x+yi$ מרוכב עם $x,y\in\mathbb{R}$ נגדיר את החלק הממשי $\mathrm{Re}(z)=x$ ואת החלק המדומה $\mathrm{Im}(z)=y$. בנוסף, נגדיר את המספר הצמוד $\bar{z}=x-yi$.

שימו לב כי בניגוד לשמו, החלק המדומה הוא מספר ממשי (זהו המקדם של i, שהוא עצמו באמת מדומה). מספר מרוכב נקרא מדומה אם החלק הממשי שלו הוא 0, כלומר זה מספר מהצורה yi כאשר $y\in\mathbb{R}$.

דוגמה 2.6.

א. $\mathrm{Re}(3+5i)=3,\ \mathrm{Im}(3+5i)=5,\ \overline{3+5i}=3-5i$

ב. $\mathrm{Re}(4i)=0,\ \mathrm{Im}(4i)=4,\ \overline{4i}=-4i$

ג. $\mathrm{Re}(3)=3,\ \mathrm{Im}(3)=0,\ \overline{3}=3$ כאשר זיהינו את 3 כמספר מרוכב עם חלק מדומה 0 לפי ההצגה $3=3+0\cdot i$.

תרגיל 2.7. לכל $z\in\mathbb{C}$ הראו כי המספר הצמוד ל-$\bar{z}$ הוא z.

פתרון. בהינתן $z=x+iy$ כאשר $x,y\in\mathbb{R}$, מתקיים לפי ההגדרה $\bar{z}=x-yi$. אם נשתמש בהגדרה שוב נקבל

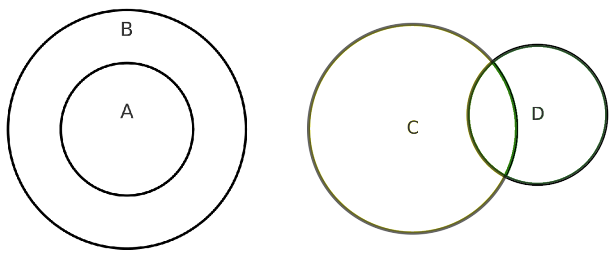
$$\bar{\bar{z}}=\overline{x-yi}=x-(-yi)=x+yi=z.$$

בפרק הבא נדון בפעולות, תכונות וחלק מהשימושים של המספרים המרוכבים. את סוף הפרק הזה נקדיש לקשר בין כל קבוצות המספרים.

2.3 תורת הקבוצות על קצה המזלג

ראינו את יחס השייכות בין איבר a לקבוצה A, וסימנו $a \in A$. בקורס שלנו הקבוצות יהיו יחסית פשוטות, ולכן לא ניתקל בקבוצה שאיבריה הם גם קבוצות. זה בהחלט תרחיש אפשרי במתמטיקה (למשל קבוצה של שני ישרים, כאשר כל ישר הוא קבוצת נקודות), אבל בקורס נעסוק בקבוצות מספרים וקבוצות וקטורים (וקטורים אינם קבוצות). בכל אופן, ייתכן קשר יותר טבעי בין שתי קבוצות A, B.

הגדרה 2.8. נאמר ש-A מוכלת ב-B ונסמן $A \subseteq B$ אם כל איבר של A הוא גם איבר של B.
באופן קצת יותר מתמטי: $A \subseteq B$ אם לכל $a \in A$ מתקיים $a \in B$.
אם ההיפך הוא הנכון, כלומר קיים $a \in A$ עבורו $a \notin B$, אז נסמן $A \nsubseteq B$ ונאמר ש-A לא מוכלת ב-B.



איור 2.3: כאן מתקיים $A \subseteq B$, אך $C \nsubseteq D$.

דוגמה 2.9. עבור $A = \{ 1, 2 \}, B = \{ 1, 2, 3 \}$ מתקיים $A \subseteq B$ כי גם $1 \in B$ וגם $2 \in B$. אבל $B \nsubseteq A$ כי $3 \in B$, אך $3 \notin A$.

תרגיל 2.10. נגדיר $A = \{ 1, 3 \}, B = \{ -1, 1, 2, 3 \}, C = \{ -1, 2 \}$. מצאו את כל זוגות הקבוצות שמקיימות קשר של הכלה.

פתרון. מתקיים $A \subseteq B$ וגם $C \subseteq B$.
אין שום הכלות אחרות. למעשה, לקבוצות A, C אין אף איבר משותף.

הגדרה 2.11. נאמר ששתי קבוצות A, B הן שוות ונסמן $A = B$ אם יש בהן בדיוק אותם האיברים, כלומר מתקיים $x \in A$ אם ורק אם $x \in B$.

הגדרה 2.12. כדי להוכיח כי $A = B$, הדרך המקובלת היא להוכיח כי $A \subseteq B$ וגם $B \subseteq A$. דרך הוכחה זו נקראת הכלה דו-צדדית.

דוגמה 2.13. הגדרנו את קבוצת המספרים הזוגיים החיוביים בשתי דרכים שונות. נוכיח שאכן מתקיים

$$\{ 2n \mid n \in \mathbb{N} \} = \{ n \in \mathbb{N} \mid \text{מתחלק ב-}2\ n \} .$$

נסמן $A = \{\, 2n \mid n \in \mathbb{N} \,\}, B = \{\, n \in \mathbb{N} \mid$ n מתחלק ב-$2 \,\}$.

נוכיח תחילה כי $A \subseteq B$: יהי $a \in A$. אז לפי ההגדרה קיים $n \in \mathbb{N}$ כך ש-$a = 2n$. מספר זה מתחלק ב-2 כי $\frac{a}{2} = n$ הוא מספר טבעי. לכן $a \in B$ כנדרש, ומאחר ש-a נבחר באופן שרירותי נובע כי $A \subseteq B$.

נעבור לכיוון השני, ההכלה $B \subseteq A$: יהי $b \in B$. אז לפי ההגדרה b מתחלק ב-2, כלומר $\frac{b}{2} \in \mathbb{N}$. נסמן $n = \frac{b}{2}$ והרי שקיבלנו כי $b = 2n$ כאשר $n \in \mathbb{N}$, ולכן $b \in A$. אז נובע כי $B \subseteq A$.

משתי ההכלות נובע כי $A = B$.

נסכם את רעיון ההוכחה: מוכיחים שתי הכלות. לכל הכלה מתחילים את הטיעון ב-"יהי" כדי להצהיר שבחרנו איבר כללי מתוך הקבוצה הנתונה. אחר כך משתמשים בהגדרת הקבוצה כדי להראות שהאיבר גם מקיים את ההגדרה של הקבוצה השנייה.

טענה 2.14. $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$ *וכל הקבוצות שונות זו מזו.*

הוכחה. ברור כי $\mathbb{Z} \subseteq \mathbb{N}$ לפי הגדרת $\mathbb{Z}$ כהרחבה של $\mathbb{N}$, ורואים שאין שוויון כי $-1 \notin \mathbb{N}$. גם דיברנו על ההכלה $\mathbb{R} \subseteq \mathbb{Q}$ שאינה שוויון, כי המספרים הרציונליים הם בדיוק המספרים הממשיים שיש להם הצגה עשרונית שהיא מחזורית החל ממקום מסוים. יש הרבה מספרים ממשיים שאינם כאלה, למשל $\sqrt{2} \notin \mathbb{Q}$.

נוכיח כי $\mathbb{Z} \subseteq \mathbb{Q}$: יהי $n \in \mathbb{Z}$. מתקיים $n = \frac{n}{1}$ וזו מנה של מספרים שלמים, ולכן $n \in \mathbb{Q}$. אז $\mathbb{Z} \subseteq \mathbb{Q}$, ואין שוויון כי למשל $\frac{1}{2} \in \mathbb{Q}$ אך $\frac{1}{2} \notin \mathbb{Z}$.

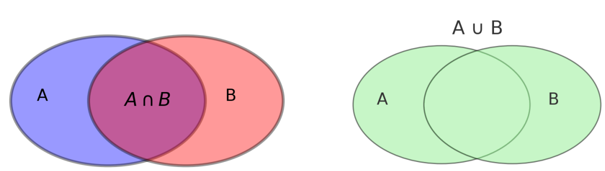
נוכיח כי $\mathbb{R} \subseteq \mathbb{C}$: יהי $x \in \mathbb{R}$. מתקיים $x = x + 0 \cdot i$ וזו הצגה של מספר מרוכב, ולכן $x \in \mathbb{C}$. אז $\mathbb{R} \subseteq \mathbb{C}$, ואין שוויון כי למשל $i \in \mathbb{C}$ אך $i \notin \mathbb{R}$. $\square$

2.4 פעולות בין קבוצות

הגדרה 2.15. בהינתן שתי קבוצות A, B נגדיר את הקבוצות הבאות:

א. החיתוך $A \cap B$ הוא קבוצת כל האיברים השייכים גם ל-A וגם ל-B.

ב. האיחוד $A \cup B$ הוא קבוצת כל האיברים השייכים לפחות לאחת משתי הקבוצות A, B.



איור 2.4: בצד ימין מופיע האיחוד בירוק, ובצד שמאל החיתוך בסגול

הערה 2.16. במקרה של איחוד זה לא משנה אם האיבר שייך רק לקבוצה אחת או לשתיהן. בכל מקרה הוא נספר רק פעם אחת באיחוד.

דוגמה 2.17.

א. עבור $A = \{ 1,2,3 \}, \ B = \{ 2,3,4 \}$ מתקיים

$$A \cap B = \{ 2,3 \}, \ A \cup B = \{ 1,2,3,4 \}.$$

ב. מתקיים $\mathbb{R} \cap \mathbb{C} = \mathbb{R}, \ \mathbb{R} \cup \mathbb{C} = \mathbb{C}$ כי $\mathbb{R} \subseteq \mathbb{C}$.

ג. נגדיר קבוצות

$$C = \{ n \in \mathbb{N} \mid \text{מתחלק ב-3 } n \}, D = \{ n \in \mathbb{N} \mid \text{מתחלק ב-5 } n \}.$$

אז לפי הגדרת החיתוך והעובדה ש-$3,5$ הם מספרים זרים (המחלק המשותף היחיד שלהם הוא 1) נובע כי החיתוך נתון באופן הבא:

$$\begin{aligned} C \cap D &= \{ n \in \mathbb{N} \mid \text{מתחלק ב-3 וגם ב-5 } n \} = \{ n \in \mathbb{N} \mid \text{מתחלק ב-15 } n \} \\ &= \{ 15, 30, 45, ... \} \end{aligned}$$

האפיון של האיחוד פחות פשוט: מדובר בקבוצת כל המספרים שמתחלקים ב-3, 5 או בשניהם (כלומר ב-15, המקרה של החיתוך). כך נקבל

$$C \cup D = \{ 3,5,6,9,10,12,15,18,20,21,24,25,27,30... \}.$$

תרגיל 2.18. חשבו את $A \cap B, \ A \cup B$ עבור $A = \{ -1,0,1 \}, \ B = \{ -2,0,1,2 \}$.

פתרון. לפי ההגדרות מתקיים

$$A \cap B = \{ 0,1 \}, \quad A \cup B = \{ -2,-1,0,1,2 \}.$$

טענה 2.19. *לכל שתי קבוצות A, B מתקיים*

$$A \cap B \subseteq A \subseteq A \cup B.$$

הוכחה. יש כאן שתי הכלות. נוכיח תחילה כי $A \cap B \subseteq A$: יהי $a \in A \cap B$. אז מיידית לפי הגדרת החיתוך מתקיים $a \in A$ ולכן $A \cap B \subseteq A$.
כעת נוכיח כי $A \subseteq A \cup B$. גם כאן זה מיידי כי כל $x \in A$ מקיים את הגדרת האיחוד (בין אם $x \in B$ ובין אם לאו). □

הערה 2.20. באותו אופן (או משיקולי סימטריה) גם מתקיים

$$A \cap B \subseteq B \subseteq A \cup B.$$

פרק 3

תכונות ושימושים של המספרים המרוכבים

3.1 פעולות חשבון

נסמן $z_1 = a + bi,\ z_2 = c + di$. נרחיב את הגדרות פעולות החשבון המוכרות ל-$\mathbb{C}$:

חיבור:

$$z_1 + z_2 = a + c + (b + d)i$$

חיסור:

$$z_1 - z_2 = a - c + (b - d)i$$

כפל:

$$z_1 z_2 = ac - bd + (ad + bc)i$$

חילוק: עבור $z_2 \neq 0$

$$\frac{z_1}{z_2} = \frac{(a + bi)(c - di)}{(c + di)(c - di)} = \frac{ac + bd}{c^2 + d^2} + \frac{(bc - ad)}{c^2 + d^2} i$$

הכפל מתאים לפתיחת סוגריים לפי חוקי הפילוג והחילוף של מספרים ממשיים. בשביל חילוק מכפילים ומחלקים במספר הצמוד $\overline{z_2} = c - di$ כדי המכנה החדש יהיה מספר ממשי.

דוגמה 3.1. ניקח $z_1 = 1 + i,\ z_2 = 2 + 3i$. כאן נקבל:

חיבור: $z_1 + z_2 = 3 + 4i$

חיסור: $z_1 - z_2 = -1 - 2i$

כפל: $z_1 z_2 = (1 + i)(2 + 3i) = 2 - 3 + (3 + 2)i = -1 + 5i$

חילוק: $\frac{1+\mathrm{i}}{2+3i} = \frac{(1+i)(2-3i)}{(2+3i)(2-3i)} = \frac{5}{13} - \frac{1}{13} i$

תרגיל 3.2. חשבו את ארבע הפעולות עבור $z_1 = 1 + 2i,\ z_2 = 3 - 4i$.

פתרון. נחשב:

$$z_1 + z_2 = 4 - 2i$$
$$z_1 - z_2 = -2 + 6i$$
$$z_1 z_2 = (1+2i)(3-4i) = 11 + 2i$$
$$\frac{z_1}{z_2} = \frac{(1+2i)(3+4i)}{(3-4i)(3+4i)} = \frac{-5+10i}{25} = \frac{-1+2i}{5}$$

הערה 3.3. נראה בהמשך שחיבור של מספרים מרוכבים שקול לחיבור בין שני וקטורים במישור, ואכן ניתן לייצג כל מספר מרוכב $z = x + yi$ כנקודה/וקטור במישור עם קואורדינטות (x, y). קואורדינטות אלו נקראות קרטזיות.

טענה 3.4. *לכל* $z_1, z_2, z_3 \in \mathbb{C}$ *מתקיים:*

א. חוקי החילוף לחיבור וכפל: $z_1 + z_2 = z_2 + z_1$ **וגם** $z_1 z_2 = z_2 z_1$.

ב. חוקי הקיבוץ לחיבור וכפל: $(z_1+z_2)+z_3 = z_1+(z_2+z_3)$ **וגם** $(z_1 z_2) z_3 = z_1 (z_2 z_3)$.

ג. חוק הפילוג: $(z_1 + z_2) z_3 = z_1 z_3 + z_2 z_3$

ד. 0 *ניטרלי ביחס לחיבור:* $z + 0 = z$ *לכל* $z \in \mathbb{C}$.

ה. 1 *ניטרלי ביחס לכפל:* $z \cdot 1 = z$ *לכל* $z \in \mathbb{C}$.

הוכחה. ישירות מן ההגדרות של חיבור וכפל תוך שימוש בחוקים המוכרים למספרים ממשיים. נסתפק בהוכחת חוק החילוף לכפל: ראשית נסמן $z_1 = a + bi,\ z_2 = c + di$. לפי הגדרת הכפל מתקיים

$$\begin{cases} z_1 z_2 = (a+bi)(c+di) = ac - bd + (ad+bc)i \\ z_2 z_1 = (c+di)(a+bi) = ca - db + (cb+da)i \end{cases}.$$

נזכור כי $a, b, c, d \in \mathbb{R}$ ולכן כבר ידוע שחוק החילוף תקף לגביהם (גם לכפל וגם לחיבור). לכן $ac = ca,\ bd = db,\ ad = da,\ bc = cb$ ומכאן נובע כי $z_1 z_2 = z_2 z_1$. $\square$

נסכם את רעיון ההוכחה: השתמשנו בנוסחה של פעולת הכפל כדי להבין כיצד חוק החילוף לכפל של מספרים מרוכבים, נובע מחוק החילוף המוכר לכפל של מספרים ממשיים.

תרגיל 3.5. הוכיחו את חוק החילוף לחיבור של מספרים מרוכבים.

פתרון.

$$\begin{cases} z_1 + z_2 = (a+bi) + (c+di) = a + c + (b+d)i \\ z_2 + z_1 = (c+di) + (a+bi) = c + a + (d+b)i \end{cases}.$$

חוק החילוף מתקיים לחיבור בין מספרים ממשיים. בפרט, מתקיים

$$a + c = c + a, \quad b + d = d + b$$

ולכן $z_1 + z_2 = z_2 + z_1$ כנדרש.

הטענה הבאה קשורות להגדרה של מספר צמוד.

טענה 3.6. *לכל $z_1, z_2 \in \mathbb{C}$ מתקיים:*

א. $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$

ב. $\overline{z_1 z_2} = \overline{z_1} \cdot \overline{z_2}$

הוכחה. נסמן $z_1 = a + bi,\ z_2 = c + di$. נחשב:

א. $\overline{z_1 + z_2} = \overline{a + c + (b + d)i} = a + c - (b + d)i = a - bi + c - di = \overline{z_1} + \overline{z_2}$

ב.

$$\begin{aligned}\overline{z_1 z_2} &= \overline{ac - bd + (ad + bc)i} \\ &= ac - bd - (ad + bc)i = (a - bi)(c - di) = \overline{z_1} \cdot \overline{z_2}\end{aligned}$$

□

טענה 3.7. *לכל $z \in \mathbb{C}$ מתקיים:*

א. $z + \overline{z} = 2Re(z)$

ב. $z - \overline{z} = 2iIm(z)$

ג. $z\overline{z} \in \mathbb{R}$

הוכחה. נסמן $z = x + yi$ כאשר $\mathrm{Re}(z) = x,\ \mathrm{Im}(z) = y$. נחשב:

א. $z + \overline{z} = x + yi + x - yi = 2x = 2\mathrm{Re}(z)$

ב. $z - \overline{z} = x + yi - (x - yi) = 2yi = 2i\mathrm{Im}(z)$

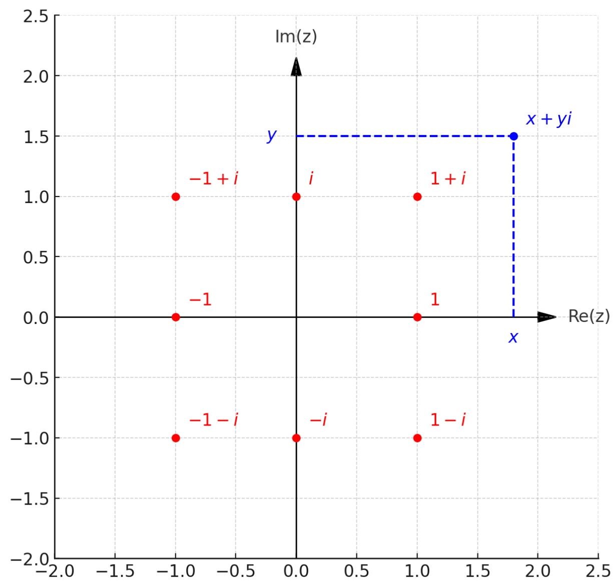
ג. $z\overline{z} = (x + yi)(x - yi) = x^2 - y^2 i^2 + (-xy + yx)i = x^2 + y^2 \in \mathbb{R}$

□

זה מראה מדוע מכפילים ומחלקים ב-$\overline{z_2}$ כדי לחשב את $\frac{z_1}{z_2}$. קל לחלק במספר ממשי, אז צריך להפוך את המכנה לממשי בעזרת הצמוד שלו.

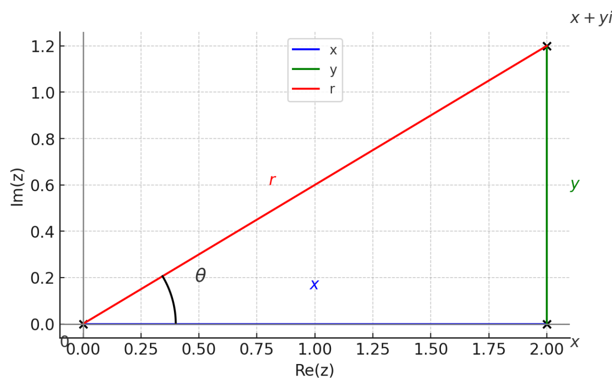
3.2 קואורדינטות פולריות

הזכרנו את הקואורדינטות הקרטזיות (x, y) עבור מספר מרוכב $z = x + yi$. בדרך זו ניתן לחשוב על מישור, שנקרא המישור המרוכב, שבו הנקודה (x, y) מייצגת את z.



איור 3.1: המישור המרוכב

יש דרך אחרת לתאר נקודה במישור. במקום להסתכל על ההיטלים x, y על הצירים, אפשר להסתכל על המרחק r מהראשית ועל הזווית θ שנוצרת בין הוקטור שיוצא אל הנקודה לבין הכיוון החיובי של הציר הממשי (ציר x). הקואורדינטות (r, θ) נקראות קואורדינטות פולריות (קוטביות), ונדגיש שהזווית θ נמדדת ברדיאנים. בפרט, נכתוב π במקום $180°$.



איור 3.2: קואורדינטות פולריות

הערה 3.8. r נקרא הערך המוחלט של z ומתקיים $r = \sqrt{z\bar{z}}$. עבור $z \in \mathbb{R}$ מדובר על ההגדרה הרגילה של ערך מוחלט.

אם הקואורדינטות הפולריות ידועות, ניתן לחשב את הקואורדינטות הקרטזיות לפי ההגדרות של סינוס וקוסינוס:

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases}$$

דוגמה 3.9. בהינתן $r = 2,\ \theta = \frac{\pi}{4}$ נוכל לחשב את הקואורדינטות הקרטזיות לפי המשוואות לעיל. נקבל $x = 2\cos\frac{\pi}{4} = 2\cdot\frac{1}{\sqrt{2}} = \sqrt{2},\ y = 2\sin\frac{\pi}{4} = 2\cdot\frac{1}{\sqrt{2}} = \sqrt{2}$. לכן המספר המרוכב המתאים הוא $z = \sqrt{2} + \sqrt{2}i$.

בכיוון ההפוך, נניח שהקואורדינטות הקרטזיות (x, y) ידועות. איך נחשב את הקואורדינטות הפולריות? ניתן להשתמש במשפט פיתגורס ולקבל

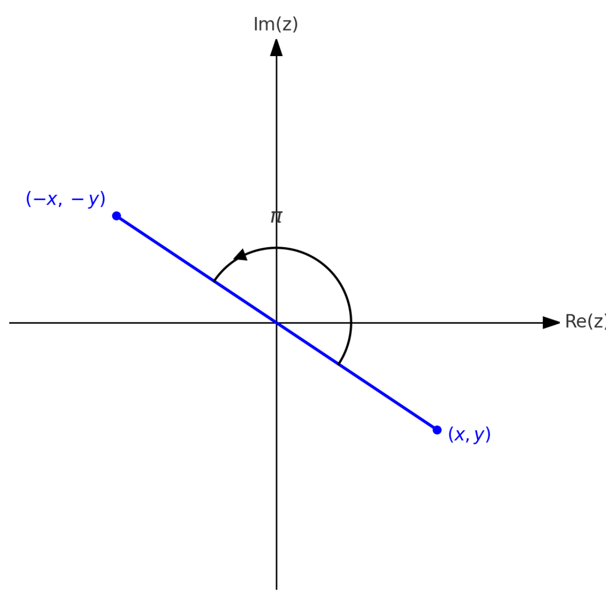
$$x^2 + y^2 = r^2 \Rightarrow r = \sqrt{x^2 + y^2}.$$

את הזווית θ ניתן לחשב לפי המשוואה $\tan\theta = \frac{y}{x}$, אבל קודם כל צריך לבחור את תחום הזוויות (ברדיאנים). התחום המקובל הוא $[-\pi, \pi)$ כאשר זה שרירותי אם לכלול את הקצה השמאלי או הימני של הקטע (בחרנו את השמאלי). הסיבה שנוח להשתמש בתחום זה היא שההפונקציה ההופכית ל- $\tan x$ היא $\arctan x$ או $\tan^{-1} x$ במחשבון, שמחזירה זוויות בתחום $(-\frac{\pi}{2}, \frac{\pi}{2})$. כלומר המחשבון עצמו עובד עם זוויות חיוביות וגם שליליות. הוא יודע לחשב את θ באופן מדויק עבור z ברביע הראשון או הרביעי, אבל בשביל שאר המקרים דרוש תיקון שקשור למחזוריות של

$\tan x$. נשתמש בהגדרה מפוצלת של פונקציה לפי מקרים:

$$\theta = \begin{cases} \arctan\frac{y}{x} & x > 0 \\ \arctan\frac{y}{x} + \pi & x < 0,\ y > 0 \\ \arctan\frac{y}{x} - \pi & x < 0,\ y < 0 \\ \frac{\pi}{2} & x = 0,\ y > 0 \\ -\frac{\pi}{2} & x = 0,\ y < 0 \end{cases}$$

זה עלול להיראות מסובך, אבל בפועל צריך לחשב את $\arctan\frac{y}{x}$ ולתקן במידת הצורך כדי לקבל זווית שמתאימה לרביע הרלוונטי (מוסיפים π כדי לעבור מהרביע הרביעי לרביע השני, מחסירים π כדי לעבור מהרביע הראשון לרביע השלישי). מספרים מדומים, עבורם $x = 0$, הם מקרה מיוחד כי מלכתחילה לא ניתן לחלק ב-0. אבל אין צורך במחשבון במקרה זה, ותמיד אפשר להשתמש בציור ולנסות לחשב את הזווית לבד.



איור 3.3: תיאור של סיבוב ב-π כדי לעבור מהרביע הרביעי לרביע השני

הערה 3.10. הזווית לא מוגדרת עבור $z = 0$. זהו מקרה יוצא דופן אך פשוט, כך שגם אין צורך בקוארדינטות פולריות בשבילו.

דוגמה 3.11. נחשב את הקוארדינטות הפולריות של $z = -2 + 2i$. נקבל

$$r = \sqrt{2^2 + 2^2} = \sqrt{8} = 2\sqrt{2}.$$

כאן הנקודה ברביע השני ולכן יש לתקן את חישוב המחשבון עבור הזווית, מה שנותן

$$\theta = \arctan\frac{2}{-2} + \pi = -\frac{\pi}{4} + \pi = \frac{3\pi}{4}.$$

3.3 משפט דה-מואבר

הגדרה 3.12. נגדיר $e^{i\theta} = \cos\theta + i\sin\theta$.

מבחינתנו זה רק סימון נוח, אבל זו בעצם נוסחה שנקראת נוסחת אוילר. יש לה משמעות יותר עמוקה שדורשת הסבר לגבי הפונקציה $f(z) = e^z$ של משתנה מרוכב. זה כבר חורג מאוד מאלגברה לינארית וגולש לתחום שנקרא אנליזה מרוכבת, שבבסיסה היא הגרסה המרוכבת של חשבון דיפרנציאלי ואינטגרלי.

משפט 3.13. *לכל* $\theta \in \mathbb{R}, n \in \mathbb{Z}$ *מתקיים* $(\cos\theta + i\sin\theta)^n = \cos(n\theta) + i\sin(n\theta)$.

שימו לב שהניסוח השקול הוא $(e^{i\theta})^n = e^{in\theta}$, שעלול להיראות מובן מאליו לפי חוקי חזקות. אבל חוקי החזקות הידועים הם למעריכים ממשיים, וגם לא הצדקנו את הסימון. אם נשים את ההוכחה בצד, המשפט מראה מדוע הסימון נוח. לא נראה את ההוכחה (שדורשת כלי שנקרא אינדוקציה), אך נציין את הקשר לטענה הבאה שמבוססת על זהויות טריגונומטריות של סכום זוויות:

טענה 3.14. *לכל* $\alpha, \beta \in \mathbb{R}$ *מתקיים* $e^{i\alpha}e^{i\beta} = e^{i(\alpha+\beta)}$.

הוכחה.

$$\begin{aligned} e^{i\alpha}e^{i\beta} &= (\cos(\alpha) + i\sin(\alpha))(\cos(\beta) + i\sin(\beta)) \\ &= \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) + i(\cos(\alpha)\sin(\beta) + \sin(\alpha)\cos(\beta)) \\ &= \cos(\alpha+\beta) + i\sin(\alpha+\beta) = e^{i(\alpha+\beta)} \end{aligned}$$

השתמשנו בזהויות הבאות:

$$\begin{cases} \cos(\alpha+\beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) \\ \sin(\alpha+\beta) = \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta) \end{cases}$$

□

תרגיל 3.15. השתמשו בטענה כדי להוכיח את המשפט במקרה $n = 2$.

פתרון. לכל $\theta \in \mathbb{R}$ מתקיים

$$(e^{i\theta})^2 = e^{i\theta} \cdot e^{i\theta} = e^{i(\theta+\theta)} = e^{i2\theta}.$$

דוגמה 3.16. נחשב את $(1+\sqrt{3}i)^{100}$ בעזרת משפט דה-מואבר (בלעדיו החישוב ארוך מאוד). ראשית נחשב קוארדינטות פולריות של $1+\sqrt{3}i$:

$$\begin{cases} r=\sqrt{1+3}=2 \\ \theta=\arctan\frac{\sqrt{3}}{1}=\frac{\pi}{3} \end{cases}$$

לכן, לפי המשפט נובע כי

$$(1+\sqrt{3}i)^{100}=(2e^{i\frac{\pi}{3}})^{100}=2^{100}e^{i\frac{100\pi}{3}}.$$

הזווית $\frac{100\pi}{3}$ חורגת מהתחום המקובל $(-\pi,\pi]$. לפי המחזוריות של קוסינוס וסינוס ניתן להחסיר מהזווית כל כפולה שלמה של 2π בלי לשנות את התוצאה. מתקיים $\frac{100\pi}{3}=(33+\frac{1}{3})\pi$, ולכן נחסיר 34π כדי לקבל זווית בתחום: $-\frac{2\pi}{3}$. מכאן:

$$(1+\sqrt{3}i)^{100}=2^{100}e^{-i\frac{2\pi}{3}}=2^{100}(\cos(-\frac{2\pi}{3})+i\sin(-\frac{2\pi}{3}))=2^{99}(-1-\sqrt{3}i).$$

אין צורך להחסיר מהזווית 34π אם המטרה היא ההצגה הקרטזית בלבד. לצורך הדוגמה, רצינו גם להראות את ההצגה הפולרית של החזקה.

תרגיל 3.17. חשבו את $(1+\sqrt{3}i)^{10}$.

פתרון. כבר חישבנו את הקוארדינטות הפולריות:

$$r=\sqrt{1+3}=2,\quad \theta=\arctan(\frac{\sqrt{3}}{1})=\frac{\pi}{3}$$

נשתמש במשפט דה-מואבר ונקבל

$$(1+\sqrt{3}i)^{10}=(2e^{i\frac{\pi}{3}})^{10}=2^{10}e^{i\frac{10\pi}{3}}=2^{10}e^{i\frac{4\pi}{3}}.$$

החסרנו מהזווית 2π כדי לעבור לתחום הזוויות הרצוי. נחשב הצגה קרטזית ונקבל

$$2^{10}e^{i\frac{4\pi}{3}}=2^{10}(-1-\sqrt{3}i)=-2^{10}-2^{10}\sqrt{3}i.$$

3.4 נוסחת השורשים

טענה 3.18. *יהי $w\in\mathbb{C}$ מספר נתון השונה מ-0, ונניח כי $w=re^{i\theta}$ בהצגה פולרית. אז למשוואה $z^n=w$ יש בדיוק n שורשים (פתרונות) מרוכבים בנעלם z הנתונים ע"י*

$$z_k=\sqrt[n]{r}e^{i\frac{\theta+2\pi k}{n}}.$$

עבור $k\in\{0,1,...,n-1\}$.

הוכחה. זו משוואה ממעלה n ולכן יש לה לכל היותר n שורשים שונים. נציב את את הנוסחה במשוואה כדי לוודא שאכן z_k הוא שורש לכל $k \in \{0, 1, ..., n-1\}$. לפי משפט דה-מואבר מתקיים

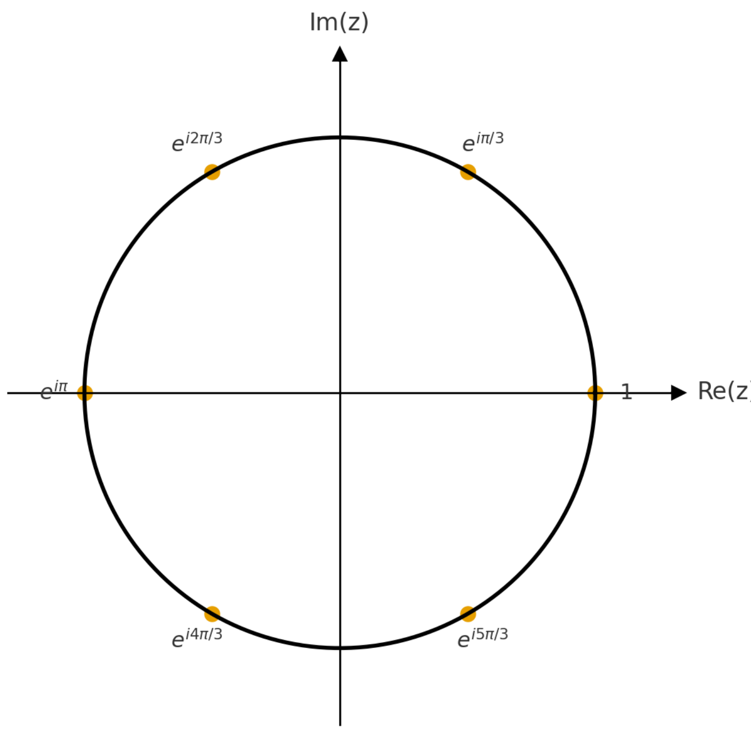
$$z_k^n = (\sqrt[n]{r} e^{i\frac{\theta+2\pi k}{n}})^n = re^{i(\theta+2\pi k)} = re^{i\theta} = w.$$

אז עברנו על כל השורשים השונים, ויש בדיוק n כאלה. □

הערה 3.19. שימו לב לשימוש במחזוריות של סינוס וקוסינוס. למעשה, z_k הוא שורש לכל $k \in \mathbb{Z}$, אבל אם נצא מהקבוצה $\{0, 1, ..., n-1\}$ נחזור על השורשים שכבר חישבנו. למשל:

$$z_n = \sqrt[n]{r} e^{i\frac{\theta+2\pi n}{n}} = \sqrt[n]{r} e^{i(\frac{\theta}{n}+2\pi)} = \sqrt[n]{r} e^{i\frac{\theta}{n}} = z_0.$$

באופן דומה, מתקיים $z_{n+1} = z_1$ וכן הלאה באופן מחזורי (עם מחזור n). לכן מספיק להציב מספרים מתוך הקבוצה $\{0, 1, ..., n-1\}$, שהיא קבוצת השאריות שניתן לקבל בחלוקה ב-n.



איור 3.4: קבוצת שורשי המשוואה $z^6 = 1$

נשים לב כי יש הפרש קבוע בין הזוויות של השורשים. הפרש זה הוא $\frac{2\pi}{n}$, ולכן אפשר לעבור משורש אחד לשורש הבא ע"י סיבוב נגד כיוון השעון בזווית זו. כאשר עושים זאת n פעמים, משלימים סיבוב שלם וחוזרים לנקודת ההתחלה. באופן כללי, לכל זווית $\alpha \in \mathbb{R}$ כפל ב- $e^{i\alpha}$ מתאים לסיבוב בזווית α נגד כיוון השעון.

תרגיל 3.20. במקרה של $n = 2$, הראו כי שני השורשים מקיימים $z_1 = -z_0$.

פתרון. זה נובע מכך ששינוי סימן מתאים לסיבוב ב-π ($180°$) נגד כיוון השעון, כי

$$e^{i\pi} = \cos(\pi) + i\sin(\pi) = -1$$

לכן, נובע מ- **3.18** ו- **3.14** כי

$$z_1 = \sqrt{r}e^{i\frac{\theta+2\pi}{2}} = \sqrt{r}e^{i\frac{\theta}{2}} \cdot e^{i\pi} = -\sqrt{r}e^{i\frac{\theta}{2}} = -z_0.$$

כנדרש.

הערה 3.21. עבור $w = 0$ יש שורש יחיד, שהוא $z = 0$. זהו מקרה מיוחד שבו השורשים מתלכדים ומתקבל שורש יחיד.

דוגמה 3.22. נפתור את המשוואה $z^4 = 1 - i$. ראשית נחשב קואורדינטות פולריות של המספר באגף ימין:

$$\begin{cases} r = \sqrt{1+1} = \sqrt{2} \\ \theta = \arctan\frac{-1}{1} = -\frac{\pi}{4} \end{cases}$$

נציב בנוסחת השורשים ונקבל את ארבעת השורשים הבאים:

$$\begin{cases} z_0 = (\sqrt{2})^{\frac{1}{4}}e^{-i\frac{\pi}{16}} = \sqrt[8]{2}e^{-i\frac{\pi}{16}} \\ z_1 = \sqrt[8]{2}e^{i(-\frac{\pi}{16}+\frac{2\pi}{4})} = \sqrt[8]{2}e^{i\frac{7\pi}{16}} \\ z_2 = \sqrt[8]{2}e^{i(-\frac{\pi}{16}+\frac{4\pi}{4})} = \sqrt[8]{2}e^{i\frac{15\pi}{16}} \\ z_3 = \sqrt[8]{2}e^{i(-\frac{\pi}{16}+\frac{6\pi}{4})} = \sqrt[8]{2}e^{i\frac{23\pi}{16}} = \sqrt[8]{2}e^{-i\frac{9\pi}{16}} \end{cases}$$

במעבר האחרון החסרנו מהזווית 2π כדי לעבור לתחום המקובל. אפשר לחשב את ההצגה הקרטזית של כל שורש באופן מקורב בעזרת מחשבון (שיודע לקרב ערכי סינוס וקוסינוס, שהם לרוב אי-רציונליים), אך אין צורך. נשים לב כי $z_2 = -z_0$ שכן ההפרש בין הזוויות הוא π, ובאופן דומה $z_3 = -z_1$. אבל אם נסתכל על הפרש הזוויות בין שורשים סמוכים, למשל z_0, z_1, נקבל $\frac{\pi}{2}$. זה בעצם אומר שאפשר לעבור משורש אחד לשורש הבא ע"י כפל במספר $e^{i\frac{\pi}{2}} = i$.

פרק 4

גיאומטריה במישור ובמרחב

בפרק זה נרצה להגדיר ולהבין וקטורים מבחינה גיאומטרית ואלגברית. נתמקד תחילה במישור הדו-מימדי (עם דגש על הישרים המוכלים בו) ואחר כך נעבור למרחב התלת-מימדי (שם נדבר גם על מישורים).

4.1 המישור $\mathbb{R}^2$

בפרקים הקודמים דיברנו על הישר הממשי וגם על המישור המרוכב. אפשר גם לתאר את המישור (הממשי) בלי מספרים מרוכבים, אבל הרעיון דומה מאוד.

הגדרה 4.1. המישור $\mathbb{R}^2$ הוא קבוצת כל הנקודות בעלות שתי קוארדינטות ממשיות, כלומר

$$\mathbb{R}^2 = \{ \, (x,y) \mid x,y \in \mathbb{R} \, \}.$$

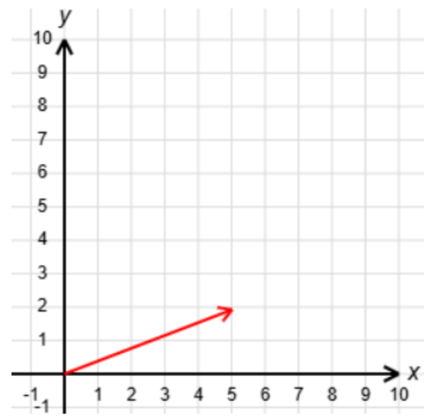
דוגמה 4.2.

א. $(0,0),(1,-2),(-3,-5),(\frac{3}{5},\sqrt{2}),(\pi,-10) \in \mathbb{R}^2$

ב. $1,(i,0),(0,1,2) \notin \mathbb{R}^2$

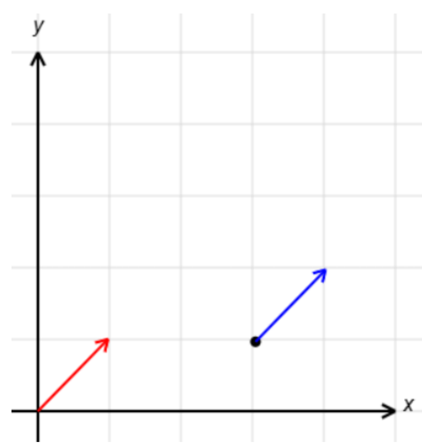
הגדרה 4.3. וקטור $\vec{v}=(x,y)$ במישור הוא זוג סדור של מספרים ממשיים, כלומר איבר של $\mathbb{R}^2$.

הערה 4.4. מבחינה אלגברית לא נבחין בין וקטור לנקודה המתאימה, ולכן גם נשתמש באותו סימון (x,y). אבל מבחינה גיאומטרית, נחשוב על וקטור כחץ שיוצא מהראשית $(0,0)$ ומסתיים בנקודה (x,y).



איור 4.1: הוקטור $(5,2)$

אמרנו שוקטורים ניתנים לתיאור חצים שיוצאים מהראשית, אבל הם יכולים לצאת מכל נקודה. מה שקובע את הוקטור הם גודלו וכיוונו, לא נקודת ההתחלה. וקטורים זהים יופיעו כחצים מקבילים ושווי גודל במישור.



איור 4.2: שני הוקטורים זהים ושווים ל-$(1,1)$

הערה 4.5. וקטורים שווים אם ורק אם יש שוויון בכל קוארדינטה, כלומר $(a,b)=(c,d)$ אם ורק אם $a=c$ וגם $b=d$. למשל, $(1,2)=(1,2)$ אך $(1,2)\neq(1,-2)$. שני הוקטורים האחרונים שווים בגודלם אך הכיוונים שונים. בנוסף, גם מתקיים $(1,2)\neq(2,1)$ ובאופן כללי יש חשיבות לסדר הקוארדינטות.

4.1.1 פעולות על וקטורים

שתי הפעולות הבסיסיות שניתן לבצע על וקטורים הן חיבור וכפל בסקלר. סקלר הוא מילה נרדפת למספר, וכאן נסמנו $c \in \mathbb{R}$.

חיבור: $(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 + b_2)$

כפל בסקלר: $c \cdot (a_1, a_2) = (ca_1, ca_2)$

דוגמה 4.6.

א. $(1, 2) + (10, 20) = (1 + 10, 2 + 20) = (11, 22)$

ב. $2 \cdot (5, 3) = (10, 6)$

ג. $-5 \cdot (2, 5) = (-10, -25)$

ד. $(9, 2) - (11, 8) = (9 - 11, 2 - 8) = (-2, -6)$

הערה 4.7. הוקטור הנגדי (שווה גודל אך הפוך בכיוון) מתקבל ע"י כפל בסקלר -1. חיסור בין שני וקטורים זה כמו חיבור בין הראשון לגרסה הנגדית של השני, כלומר:

$$(a_1, a_2) - (b_1, b_2) = (a_1, a_2) + (-1) \cdot (b_1, b_2) = (a_1 - b_1, a_2 - b_2)$$

תרגיל 4.8. בצעו את הפעולות הבאות:

א. $(-3, 5) + (20, 36)$

ב. $-7 \cdot (-2, 9)$

ג. $(11, 27) - (13, 42)$

ד. $3(8, 2) + 5(3, -6)$

פתרון.

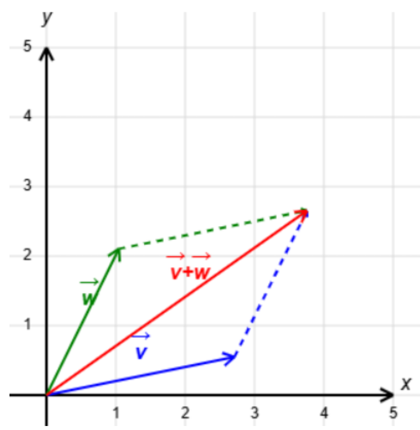
א. $(-3, 5) + (20, 36) = (17, 41)$

ב. $-7 \cdot (-2, 9) = (14, -63)$

ג. $(11, 27) - (13, 42) = (-2, -15)$

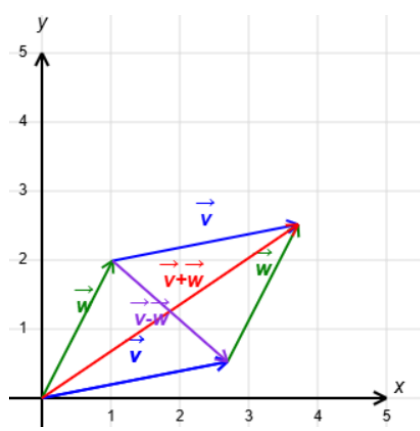
ד. $3(8, 2) + 5(3, -6) = (24 + 15, 6 - 30) = (39, -24)$

אפשר להבין חיבור וקטורי באופן גיאומטרי ע"י כלל המקבילית. כלל זה קובע שאם נצייר שני וקטורים $\vec{v},\vec{w}$ שיוצאים מהראשית ונשלים אותם למקבילית ע"י הוספת שני וקטורים מקבילים ושווי גודל כצלעות נגדיות, אז האלכסון שיוצא מהראשית יתאר את $\vec{v}+\vec{w}$.



איור 4.3: החיבור הוקטורי מתאים לאלכסון המוצג

ההפרש $\vec{v}-\vec{w}$ מיוצג ע"י האלכסון השני.



איור 4.4: ההפרש הוקטורי מופיע בסגול לצד החיבור הוקטורי באדום

זה גם מוביל אותנו לנוסחה שמתארת וקטור $\vec{u}$ שמתחיל בנקודה (a_1,a_2) ומסתיים ב-(b_1,b_2): מקבלים את ההפרש הוקטורי

$$\vec{u}=(b_1-a_1,b_2-a_2). \tag{4.1}$$

תרגיל 4.9. הראו שהוקטור שמתחיל ב-$(1,3)$ ומסתיים ב-$(5,0)$, שווה לוקטור שמתחיל ב-$(5,-1)$ ומסתיים ב- $(9,-4)$.

פתרון. הוקטור הראשון שווה ל-

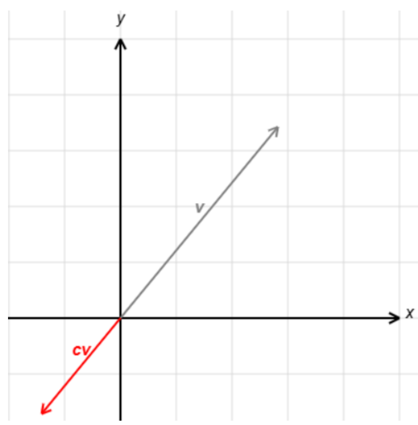
$$(5,0)-(1,3)=(4,-3),$$

ואילו הוקטור השני שווה ל-

$$(9,-4)-(5,-1)=(4,-3).$$

לכן יש שוויון ביניהם.

כפל בסקלר זו פעולה שאפשר לפרש גיאומטרית כמתיחה של הוקטור אם $c>1$, או כיווצו אם $0<c<1$, בלי לשנות כיוון. כאשר הסקלר שלילי, יש היפוך כיוון ובנוסף מתיחה/כיווץ בהתאם לערך של $|c|$.



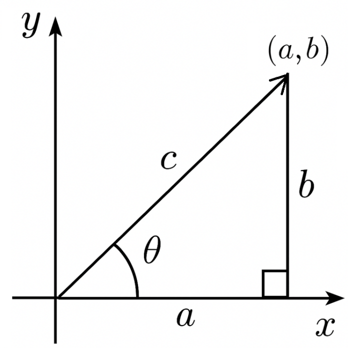
איור 4.5: כפל בסקלר $c=-\frac{1}{2}$

4.1.2 גדלים וזוויות ב-$\mathbb{R}^2$

כאמור, לכל וקטור יש גודל וכיוון. אם נניח שהוא מתחיל בראשית, אז גודלו/אורכו הוא המרחק מנקודת הסיום לראשית. הכיוון נקבע לפי הזווית θ עם הכיוון החיובי של ציר x בדיוק כמו בקואורדינטות פולריות. עבור וקטור (a,b) נקבל את הנוסחאות הבאות:

גודל: $\|(a,b)\|=\sqrt{a^2+b^2}$

זווית: $\tan\theta=\frac{b}{a}$



איור 4.6: משולש ישר-זווית שבו היתר מתאים לוקטור הנתון

החישוב של גודל הוקטור נובע ממשפט פיתגורס. בציור c הוא הסימון של $\|(a,b)\|$ ולפי משפט פיתגורס מתקיים

$$a^2+b^2=c^2 \implies \|(a,b)\|=c=\sqrt{a^2+b^2}.$$

כבר דיברנו על החישוב המדויק של θ בסוף פרק 3. לא ממש נצטרך לחשב זוויות בקורס, אבל זה טוב לידע כללי.

הערה 4.10. כל וקטור נקבע ביחידות ע"י הגודל והכיוון שלו. קואורדינטות פולריות מתארות גודל וכיוון, אבל בשביל פעולות בין וקטורים יותר נוח להשתמש בקואורדינטות קרטזיות. כך נעשה בקורס.

דוגמה 4.11.

א. $\|(2,1)\|=\sqrt{2^2+1^2}=\sqrt{5}$

ב. $\|(3,-4)\|=\sqrt{3^2+(-4)^2}=5$

ג. $\|(-3,-7)\|=\sqrt{(-3)^2+(-7)^2}=\sqrt{58}$

אם כפל בסקלר מתאים למתיחה/כיווץ, זה אמור להתבטא בגודל. מכאן הטענה הבאה:

טענה 4.12.

$$\|c\cdot(a,b)\|=|c|\cdot\|(a,b)\|$$

לכל וקטור $(a,b)\in\mathbb{R}^2$ *ולכל סקלר* $c\in\mathbb{R}$.

הוכחה. נשתמש בהגדרות ונחשב:

$$\begin{aligned}\|c\cdot(a,b)\| &= \|(ca,cb)\| = \sqrt{(ca)^2+(cb)^2} = \sqrt{c^2(a^2+b^2)} = \sqrt{c^2}\sqrt{a^2+b^2}\\ &= |c|\sqrt{a^2+b^2} = |c|\cdot\|(a,b)\|\end{aligned}$$

נציין שמתקיים $\sqrt{c^2}=|c|$ כי מדובר בשורש חיובי. $\square$

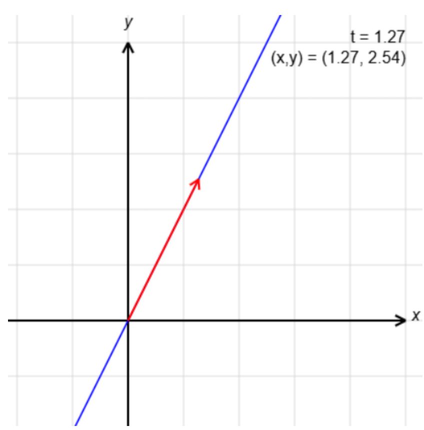
הערה 4.13. בפרט, כפל ב--1 לא משנה את גודל הוקטור. רק הכיוון מתהפך.

4.1.3 ישרים ב-$\mathbb{R}^2$

נרצה להבין את הקשר בין ישרים לוקטורים. יותר קל להבין את זה במקרה של ישרים שעוברים דרך הראשית. ניקח ישר כזה ונבחר נקודה כלשהי שאינה הראשית, מה שייתן לנו וקטור. הישר הוא אוסף כל הנקודות שמתקבלות ע״י כפל בסקלר של וקטור זה.

דוגמה 4.14. הישר שמשוואתו $y=2x$ עובר דרך הראשית והנקודה $(1,2)$. דרך נוספת לתאר את הישר היא כקבוצת כל הכפולות (בסקלר ממשי) של הוקטור $(1,2)$. כל הוקטורים האלה מתאימים לנקודות על הישר, וכותבים את הנקודות כמו הוקטורים (כי הם יוצאים מהראשית). כך נקבל את ההצגה הבאה לישר כקבוצת נקודות:

$$\{\, t(1,2) \mid t\in\mathbb{R} \,\} = \{\, (t,2t) \mid t\in\mathbb{R} \,\}$$



איור 4.7: הישר הנ״ל עם דגש על הנקודה המתאימה להצבה $t=1.27$

אכן, זה מיידי שכל נקודה מהצורה $(t,2t)$ מקיימת את המשוואה $y=2x$.
באופן כללי, הישר שעובר דרך הראשית בכיוון הוקטור $\vec{v}$ נתון ע״י ההצגה $\{\, t\vec{v} \mid t\in\mathbb{R} \,\}$.
נשים לב שעבור $t<0$ מקבלים וקטורים עם כיוון מנוגד (הצד השני של הישר).

נכליל: ישר שעובר דרך (x_0, y_0) בכיוון הוקטור $\vec{v}$ נתון ע"י ההצגה

$$\{ (x_0, y_0) + t\vec{v} \mid t \in \mathbb{R} \}.$$

אם נתון שהישר גם עובר דרך (x_1, y_1), אז נקבל

$$\{ (x_0, y_0) + t(x_1 - x_0, y_1 - y_0) \mid t \in \mathbb{R} \}.$$

תרגיל 4.15. איזו נקודה מתקבלת עבור $t = 0$? $t = 1$? $t = 2$?

פתרון. עבור $t = 0$ מתקבלת נקודת ההתחלה (x_0, y_0).
עבור $t = 1$ מתקבלת נקודת הסיום (x_1, y_1).
לבסוף, עבור $t = 2$ מתקבלת הנקודה

$$(x_0 + 2x_1 - 2x_0, y_0 + 2y_1 - 2y_0) = (2x_1 - x_0, 2y_1 - y_0).$$

הערה 4.16. ההצגה לעיל נקראת הצגה פרמטרית של ישר. לעומת זאת, הצגה מהצורה

$$L = \left\{ (x, y) \in \mathbb{R}^2 \;\middle|\; ax + by = c \right\}$$

נקראת הצגה אלגברית של ישר. הצגה אלגברית אינה יחידה, כי למשל המשוואה $y = 2x$ שקולה למשוואה $2y = 4x$. באופן דומה, גם הצגה פרמטרית אינה יחידה משתי סיבות: אפשר לבחור את נקודת ההתחלה (x_0, y_0) באופן שרירותי כל עוד היא על הישר, וגם אפשר לבחור את (x_1, y_1) כרצוננו ובהתאם את וקטור הכיוון $\vec{v}$ (אפשר לקחת כל כפולה שלו בסקלר שאינו 0). כך מקבלים אינסוף הצגות פרמטריות שונות לאותו הישר. האות של הפרמטר היא סימון שרירותי ולכן ניתן להשתמש באותה האות להצגות שונות, אך לפעמים נרצה להשתמש באותיות שונות כדי למנוע בלבול. למשל:

$$\{ t(1,2) \mid t \in \mathbb{R} \} = \{ (1,2) + s(2,4) \mid s \in \mathbb{R} \} = \{ (2,4) + r(-3,-6) \mid r \in \mathbb{R} \}$$

תרגיל 4.17. עבור הישר שמשוואתו $y = 3x + 1$, מצאו שלוש הצגות פרמטריות שונות. נסו לשנות גם את נקודת ההתחלה וגם את וקטור הכיוון.

פתרון. נבחר שתי נקודות על הישר. תחילה נציב $x = 0$ ונקבל את הנקודה $(0, 1)$. בשביל הנקודה השנייה נציב $x = 1$ ונקבל את הנקודה $(1, 4)$. וקטור הכיוון המתאים לשתי הנקודות האלו הוא $(1, 4) - (0, 1) = (1, 3)$.

אז ההצגה הפרמטרית הראשונה שמתאימה לשתי הנקודות האלו היא

$$L = \{ (0,1) + t(1,3) \mid t \in \mathbb{R} \}.$$

בשביל ההצגה השנייה נחליף את נקודת ההתחלה בנקודת הסיום (אפשר להפוך את סימן וקטור הכיוון, אך אין צורך):

$$L = \{ (1,4) + s(1,3) \mid s \in \mathbb{R} \}.$$

בשביל ההצגה השלישית נכפיל את וקטור הכיוון ב- 2 בלי לשנות את נקודת ההתחלה המקורית:

$$L = \{ (0,1) + r(2,6) \mid r \in \mathbb{R} \}.$$

4.1.4 מכפלה סקלרית

מכפלה סקלרית היא פעולה נוספת בין וקטורים, אבל התוצאה שלה היא סקלר ומכאן שמה.

הגדרה 4.18. עבור וקטורים $(a_1, a_2), (b_1, b_2) \in \mathbb{R}^2$ נגדיר $(a_1, a_2)\cdot(b_1, b_2) = a_1b_1 + a_2b_2$.

דוגמה 4.19.

א. $(1,3)\cdot(2,5) = 1\cdot 2 + 3\cdot 5 = 17$

ב. $(3,-2)\cdot(6,5) = 3\cdot 6 + (-2)\cdot 5 = 8$

ג. $(1,5)\cdot(-5,1) = 1\cdot(-5) + 5\cdot 1 = 0$

למכפלה סקלרית יש קשר לגודל. יש מספר תכונות שכדאי לזכור:

טענה 4.20. *יהיו* $\vec{v_1} = (x_1, y_1),\ \vec{v_2} = (x_2, y_2),\ \vec{v_3} = (x_3, y_3) \in \mathbb{R}^2$ *וקטורים. אז מתקיים:*

א. $\vec{v_1}\cdot\vec{v_2} = \vec{v_2}\cdot\vec{v_1}$.

ב. $(\vec{v_1} + \vec{v_3})\cdot\vec{v_2} = \vec{v_1}\cdot\vec{v_2} + \vec{v_3}\cdot\vec{v_2}$.

ג. $\vec{v_1}\cdot\vec{v_1} = \|\vec{v_1}\|^2$.

ד. $\vec{v_1}\cdot(c\vec{v_2}) = c(\vec{v_1}\cdot\vec{v_2})$ *לכל סקלר* $c \in \mathbb{R}$.

ה. $\vec{v_1} = (0,0) \iff \|\vec{v_1}\| = 0$.

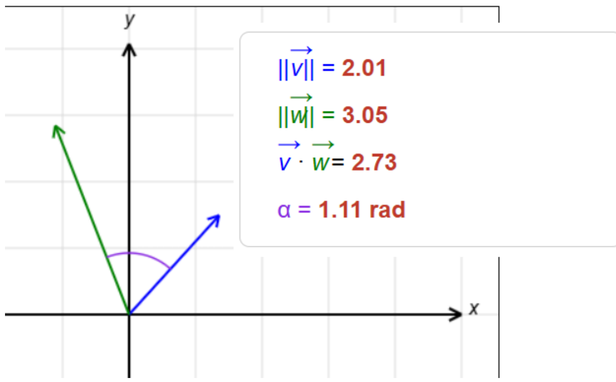
ו. $\|\frac{\vec{v_1}}{\|\vec{v_1}\|}\| = 1$ *אם* $\vec{v_1} \neq (0,0)$.

הוכחה. נסתפק בסעיפים ג׳ ו-ה׳. סעיפים אחרים יופיעו כתרגילים. לפי ההגדרה של מכפלה סקלרית, מתקיים

$$\vec{v_1}\cdot\vec{v_1} = x_1^2 + y_1^2 = \left(\sqrt{x_1^2 + y_1^2}\right)^2 = \|\vec{v_1}\|^2.$$

סכום ריבועים (של מספרים ממשיים) הוא אי-שלילי, והוא מתאפס אם ורק אם כל מחובר מתאפס. במקרה זה $x_1^2 = y_1^2 = 0$ גורר $x_1 = y_1 = 0$, או באופן שקול $\vec{v_1} = (0,0)$. □

מכפלה סקלרית גם קשורה לזווית בין שני וקטורים. ליתר דיוק, הכוונה היא לזווית הקטנה ביניהם (בין 0 ל-π ברדיאנים).



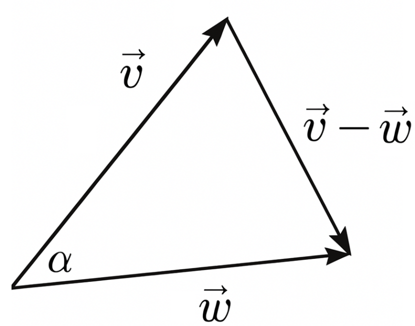
איור 4.8: שני וקטורים והזווית ביניהם, עם חישובים נלווים

טענה 4.21. *יהיו $\vec{v},\vec{w}$ שני וקטורים ב-$\mathbb{R}^2$ השונים מ-$(0,0)$, ותהי α הזווית הקטנה ביניהם. אז מתקיים*

$$\cos\alpha = \frac{\vec{v}\cdot\vec{w}}{\|\vec{v}\|\cdot\|\vec{w}\|} . \tag{4.2}$$

ההוכחה תדגים היטב את השימוש בתכונות של מכפלה סקלרית.

הוכחה. נסתמך על משפט הקוסינוסים מטריגונומטריה. נסתכל על המשולש שנוצר ע"י שני הוקטורים:



איור 4.9: המשולש המתאים לשני הוקטורים יחד עם וקטור ההפרש

$$\|\vec{v}\|^2+\|\vec{w}\|^2=\|\vec{v}-\vec{w}\|^2+2\|\vec{v}\|\cdot\|\vec{w}\|\cos\alpha$$

נשתמש בתכונות של מכפלה סקלרית כדי למצוא את הקשר בין $\|v-w\|^2$ ל-$v\cdot w$:

$$\|\vec{v}-\vec{w}\|^2=(\vec{v}-\vec{w})\cdot(\vec{v}-\vec{w})=\vec{v}\cdot\vec{v}-\vec{v}\cdot\vec{w}-\vec{w}\cdot\vec{v}+\vec{w}\cdot\vec{w}=\|\vec{v}\|^2-2\vec{v}\cdot\vec{w}+\|\vec{w}\|^2$$

נציב את המשוואה השנייה בראשונה ונקבל

$$\|\vec{v}\|^2+\|\vec{w}\|^2=\|\vec{v}\|^2-2\vec{v}\cdot\vec{w}+\|\vec{w}\|^2+2\|\vec{v}\|\cdot\|\vec{w}\|\cos\alpha.$$

נקזז $\|\vec{v}\|^2+\|\vec{w}\|^2$ מכל אגף ונקבל

$$0=-2\vec{v}\cdot\vec{w}+2\|\vec{v}\|\cdot\|\vec{w}\|\cos\alpha \implies \cos\alpha=\frac{\vec{v}\cdot\vec{w}}{\|\vec{v}\|\cdot\|\vec{w}\|}.$$

כנדרש. □

מסקנה 4.22. *וקטורים $\vec{v},\vec{w}$ השונים מ-$(0,0)$ הם מאונכים/ניצבים אם ורק אם $\vec{v}\cdot\vec{w}=0$. במקרה זה נסמן $\vec{v}\perp\vec{w}$.*

הוכחה. תהי α הזווית הקטנה ביניהם. אז מתקיים

$$\vec{v}\cdot\vec{w}=0 \iff \cos\alpha=\frac{\vec{v}\cdot\vec{w}}{\|\vec{v}\|\cdot\|\vec{w}|}=0 \iff \alpha=\frac{\pi}{2}.$$

□

דוגמה 4.23.

א. $(1,2)\perp(-2,1)$ כי $(1,2)\cdot(-2,1)=0$.

ב. באופן כללי, לכל $(a,b)\in\mathbb{R}^2$ מתקיים $(a,b)\perp(-b,a)$ כי $(a,b)\cdot(-b,a)=0$.

הערה 4.24. הוקטור $(0,0)$ הוא מקרה חריג כי הוא חסר כיוון. אבל מאחר שמתקיים $\vec{v}\cdot(0,0)=0$ לכל וקטור $\vec{v}\in\mathbb{R}^2$, עדיין נגדיר $\vec{v}\perp(0,0)$. זהו הוקטור היחיד שמאונך לכל וקטור במישור, כי כל וקטור אחר לא מאונך לעצמו.

4.1.5 מעבר מהצגה פרמטרית של ישר להצגה אלגברית

ההצגה האלגברית של ישר קשורה באופן הדוק למכפלה סקלרית. נבין את הקשר דרך דוגמה.

דוגמה 4.25. נסתכל על הישר $L = \{\, t(2,-1) \mid t \in \mathbb{R} \,\}$. רואים לפי הצגה הפרמטרית שוקטור הכיוון הוא $(2,-1)$. לפי המכפלה הסקלרית אפשר לקבוע שהוקטור $(1,2)$ מאונך לו. הוא לא רק מאונך ל- $(2,-1)$ אלא לכל וקטור שמתאים לנקודה על הישר, כי לכל $t \in \mathbb{R}$ מתקיים

$$.(1,2)\cdot(2t,-t) = 2t - 2t = 0$$

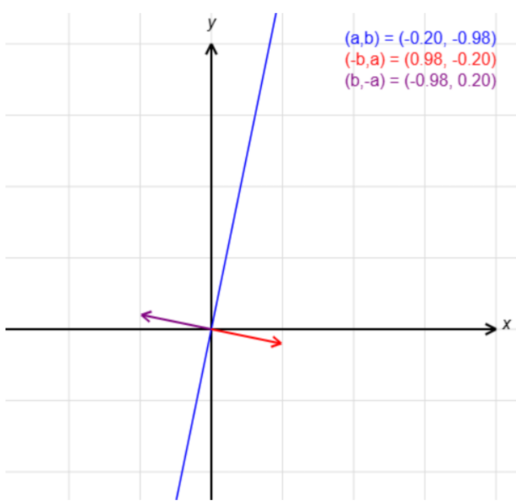
להיפך, אפשר לתאר את הישר כאוסף כל הוקטורים שמאונכים לוקטור $(1,2)$. כלומר כל וקטור $(x,y) \in L$ נתון ע"י המשוואה

$$.(x,y)\cdot(1,2) = 0 \iff x + 2y = 0$$

או בכתיב קבוצות: $.L = \{\, (x,y) \in \mathbb{R}^2 \mid x + 2y = 0 \,\}$

הגדרה 4.26. וקטור המאונך לכל וקטור שמתאים לשתי נקודות על ישר (התחלה וסיום) נקרא נורמל לישר.

באופן כללי, לישר עם הצגה פרמטרית $\{\, (x_0,y_0) + t(a,b) \mid t \in \mathbb{R} \,\}$ יש וקטור נורמל $(-b,a)$ המאונך לוקטור הכיוון (a,b). אבל וקטור הנורמל אינו יחיד, כי אם נכפיל אותו בסקלר $c \neq 0$ הוא יישאר מאונך לוקטור הכיוון של הישר. כל שני וקטורים מאונכים יישארו מאונכים גם אם נכפיל את אחד מהם (או שניהם) בסקלר.



איור 4.10: ישר ושני וקטורי נורמל לדוגמה

כדי לעבור להצגה אלגברית נשתמש בוקטור שמתחיל ב-(x_0,y_0) ומסתיים ב-(x,y), כלומר הוקטור $(x-x_0,y-y_0)$. וקטור זה בהכרח מאונך לוקטור הנורמל $(-b,a)$, ולכן צריך להתקיים:

$$\begin{aligned}(x-x_0,y-y_0)\cdot(-b,a)=0 &\iff -b(x-x_0)+a(y-y_0)=0\\ &\iff -bx+ay=-bx_0+ay_0\end{aligned}$$

כלומר קיבלנו הצגה אלגברית $L=\{\,(x,y)\in\mathbb{R}^2 \mid -bx+ay=-bx_0+ay_0\,\}$. נדגיש שהבחירה של נקודת ההתחלה (x_0,y_0) היא שרירותית. המשוואה מראה שהביטוי $-bx+ay$ קבוע לאורך הישר, וקבוע זה מופיע באגף ימין של המשוואה.

דוגמה 4.27. נסתכל על הישר $L=\{\,(2,3)+t(1,-1) \mid t\in\mathbb{R}\,\}$. יש לו וקטור נורמל $(1,1)$ ולכל $(x,y)\in L$ הוקטור $(x-2,y-3)$ מאונך ל- $(1,1)$, ולכן מתקיים

$$(x-2,y-3)\cdot(1,1)=0 \iff x-2+y-3=0 \iff x+y=5.$$

קיבלנו הצגה אלגברית $L=\{\,(x,y)\in\mathbb{R}^2 \mid x+y=5\,\}$.

תרגיל 4.28. מצאו הצגה אלגברית של $L=\{\,(1,-2)+t(3,1) \mid t\in\mathbb{R}\,\}$.

פתרון. כאן יש וקטור נורמל $(-1,3)$. לכל $(x,y)\in L$ הוקטור

$$(x-1,y-(-2))=(x-1,y+2)$$

מאונך ל- $(-1,3)$, ולכן מתקיים

$$(x-1,y+2)\cdot(-1,3)=0 \iff -x+1+3y+6=0 \iff x-3y=7.$$

ההצגה האלגברית המתאימה היא $L=\{\,(x,y)\in\mathbb{R}^2 \mid x-3y=7\,\}$.

4.1.6 מעבר מהצגה אלגברית של ישר להצגה פרמטרית

כדי למצוא הצגה פרמטרית של ישר, צריך למצוא שתי נקודות עליו. האחת בשביל נקודת ההתחלה, והשנייה בשביל נקודת הסיום של וקטור הכיוון. הבחירה של נקודות אלו היא שרירותית לחלוטין, אבל יחסית נוח להשתמש בנקודות חיתוך עם הצירים.

דוגמה 4.29. נסתכל על הישר $L=\{\,(x,y)\in\mathbb{R}^2 \mid 2x+3y=1\,\}$. אם נציב $y=0$ נקבל $x=\frac{1}{2}$, ולכן נבחר את $(\frac{1}{2},0)$ כנקודת התחלה. אם נציב $x=0$ נקבל $y=\frac{1}{3}$, ולכן נבחר את $(0,\frac{1}{3})$ כנקודת סיום של וקטור הכיוון. כך נקבל וקטור כיוון $(-\frac{1}{2},\frac{1}{3})=(0,\frac{1}{3})-(\frac{1}{2},0)$.

מכאן $L=\left\{\,(\frac{1}{2},0)+t(-\frac{1}{2},\frac{1}{3}) \;\middle|\; t\in\mathbb{R}\,\right\}$ היא הצגה פרמטרית אחת מתוך רבות. מי שלא אוהב שברים יכול להכפיל ב-6 את וקטור הכיוון, וגם אפשר להחליף את נקודת ההתחלה ב-$(2,-1)$, למשל. אז גם מתקיים $L=\{\,(2,-1)+t(-3,2) \mid t\in\mathbb{R}\,\}$ ושתי התשובות נכונות באותה המידה. נשים לב שלפי הנוסחה $(-b,a)$, וקטור הנורמל שמתאים להצגה הפרמטרית הראשונה הוא $(-\frac{1}{3},-\frac{1}{2})$. באופן דומה, וקטור הנורמל שמתאים להצגה הפרמטרית השנייה הוא $(-2,-3)$. שני הוקטורים האלה הם כפולות של הוקטור $(2,3)$ שמופיע בתוך ההצגה האלגברית המקורית בתור המקדמים של x,y בהתאמה.

באופן כללי, עבור הצגה אלגברית $L = \{ (x,y) \in \mathbb{R}^2 \mid ax+by=c \}$ נבחר נקודת התחלה (x_0, y_0) כלשהי. זו נקודה שמקיימת את משוואת הישר ולכן בהכרח $ax_0 + by_0 = c$, ולאחר הצבה במשוואה נקבל:

$$\begin{aligned}(x,y) \in L &\iff ax+by = ax_0 + by_0 \iff a(x-x_0) + b(y-y_0) = 0 \\ &\iff (x-x_0, y-y_0) \perp (a,b)\end{aligned}$$

אז (a,b) הוא וקטור נורמל של L. כל וקטור כיוון שנמצא ע"י בחירה נקודה נוספת של נקודה על L בהכרח מאונך לוקטור זה. זו בדיוק המשמעות של התנאי $(x-x_0, y-y_0) \perp (a,b)$.

תרגיל 4.30. מצאו הצגה פרמטרית של $L = \{ (x,y) \in \mathbb{R}^2 \mid 3x+y=5 \}$. האם וקטור הכיוון מאונך ל- $(3,1)$?

פתרון. נוח להציב $x=0$ ולקבל $y=5$. אז נבחר את $(0,5)$ להיות נקודת ההתחלה. כעת נציב (בחירה שרירותית אך נוחה) $x=1$ ונקבל $y=2$. אז נקודת הסיום של וקטור הכיוון היא $(1,2)$, ולכן הוקטור עצמו נתון ע"י $\vec{v} = (1,2) - (0,5) = (1,-3)$.

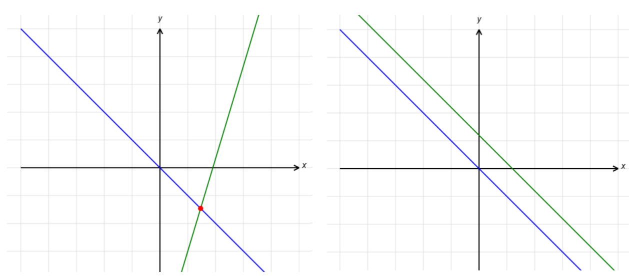
אז ההצגה הפרמטרית המתאימה היא

$$L = \{ (0,5) + t(1,-3) \mid t \in \mathbb{R} \}.$$

אכן מתקיים $(1,-3) \cdot (3,1) = 3-3 = 0$, אז וקטור הכיוון מאונך לוקטור הנורמל כצפוי.

4.1.7 מצבים הדדיים בין ישרים במישור

קיימים שלושה מצבים הדדיים אפשריים בין שני ישרים במישור: נחתכים, מקבילים ומתלכדים.



איור 4.11: ישרים מקבילים בצד ימין, ישרים נחתכים בצד שמאל

נתחיל בדוגמאות לפני שנעבור לאפיון.

דוגמה 4.31.

א. $L_1 = \{ (1,1)+t(1,2) \mid t \in \mathbb{R} \}$ ו-$L_2 = \{ (0,-1)+t(2,4) \mid t \in \mathbb{R} \}$ הם ישרים מתלכדים, כלומר $L_1 = L_2$. ניתן לראות זאת כי $(2,4) = 2 \cdot (1,2)$ ולכן שני וקטורי הכיוון מקבילים. זה אומר שהישרים או מקבילים או מתלכדים (בהכרח לא נחתכים). בנוסף, $(0,-1) \in L_1$ כי $(0,-1) = (1,1) - 1 \cdot (1,2)$. אז יש לפחות נקודת חיתוך אחת, ולכן שני הישרים אינם מקבילים אלא מתלכדים.

ב. $L_1 = \{ (1,1)+t(1,2) \mid t \in \mathbb{R} \}$ ו-$L_3 = \{ (0,1)+t(-1,-2) \mid t \in \mathbb{R} \}$ הם ישרים מקבילים. שני וקטורי הכיוון הם מנוגדים, אבל כל ישר מכיל נקודות שמתאימות גם לוקטור הכיוון $(1,2)$ וגם לכיוון הנגדי. אז הישרים או מקבילים או מתלכדים, וניתן להראות שאין נקודת חיתוך. לצורך כך נכתוב $L_3 = \{ (0,1)+s(-1,-2) \mid s \in \mathbb{R} \}$ עם פרמטר s כדי להדגיש שזה פרמטר שונה. נחפש נקודת חיתוך $(x,y) \in L_1 \cap L_3$. אפשר להציג את הקואורדינטות שלה בשתי דרכים שונות לפי ההצגות הפרמטריות של הישרים. כך נקבל:

$$\begin{cases} x = 1 + t = 0 - s \\ y = 1 + 2t = 1 - 2s \end{cases}$$

כעת אפשר להתעלם מ-x, y ולפתור מערכת של שתי משוואות בשני נעלמים. מהמשוואה הראשונה נובע כי $s = -1 - t$ וכאשר מציבים את זה במשוואה השנייה נקבל

$$1 + 2t = 1 - 2(-1-t) \implies 1 = 3.$$

זה לא ייתכן (פסוק שקר) ולכן קיבלנו סתירה. מה המשמעות של סתירה? עצם ההנחה שיש נקודת חיתוך התבררה כשגויה, ולכן שני הישרים מקבילים.

ג. $L_1 = \{ (1,1)+t(1,2) \mid t \in \mathbb{R} \}$ ו-$L_4 = \{ (3,2)+t(1,-1) \mid t \in \mathbb{R} \}$ הם ישרים נחתכים. שוב נשנה את הפרמטר ונכתוב $L_4 = \{ (3,2)+s(1,-1) \mid s \in \mathbb{R} \}$. נחפש נקודת חיתוך $(x,y) \in L_1 \cap L_4$ ונקבל

$$\begin{cases} 1 + t = 3 + s \\ 1 + 2t = 2 - s \end{cases}.$$

נחבר את שתי המשוואות ונקבל

$$2 + 3t = 5 \implies 3t = 3 \implies t = 1.$$

נציב $t = 1$ במשוואה הראשונה (אפשר גם בשנייה) ונקבל

$$2 = 3 + s \implies s = -1.$$

זה מראה כי $(2,3) \in L_1 \cap L_4$ לפי הצבת $t = 1$ בהצגה הפרמטרית של L_1, או לפי הצבת $s = -1$ בהצגה הפרמטרית של L_4. יותר מכך, זו נקודת החיתוך היחידה לפי החישוב ולכן שני הישרים אכן נחתכים (לא מקבילים).

הגדרה 4.32. שני וקטורים $\vec{v}, \vec{w} \in \mathbb{R}^2$ נקראים קו-לינאריים אם קיים $c \in \mathbb{R}$ כך ש-$\vec{v} = c\vec{w}$ או $\vec{w} = c\vec{v}$.

הערה 4.33. אם שני הוקטורים שונים מ-$(0,0)$, אז המשמעות הגיאומטרית של קו-לינאריות היא ששני הוקטורים מקבילים. במקרה זה $c \neq 0$ והתנאי $\vec{v} = c\vec{w}$ שקול לתנאי $\vec{w} = \frac{1}{c}\vec{v}$.
אבל אם למשל $\vec{v} = (0,0)$, אז לכל $\vec{w} \in \mathbb{R}^2$ מתקיים $\vec{v} = 0 \cdot \vec{w}$ ולכן שני הוקטורים נחשבים קו-לינאריים אך לא מקבילים (כי וקטור האפס חסר כיוון).

דוגמה 4.34.

א. $(2,-4), (-3,6)$ הם קו-לינאריים כי $(-3,6) = -\frac{3}{2}(2,-4)$.

ב. $(1,5), (0,0)$ הם קו-לינאריים כי $(0,0) = 0 \cdot (1,5)$.

ג. $(1,2), (3,1)$ אינם קו-לינאריים. כדי לראות זאת, נניח בשלילה שהם קו-לינאריים וננסה לקבל סתירה. אז לפי ההגדרה קיים $c \in \mathbb{R}$ כך ש-$(1,2) = c(3,1)$, או להיפך אבל שני התנאים שקולים כי אף וקטור אינו $(0,0)$. אז מצד אחד קיבלנו

$$.1 = 3c \implies c = \frac{1}{3}$$

ומצד שני $c = 2$ לפי המשוואה השנייה. זו סתירה ולכן שני הוקטורים אינם קו-לינאריים.

הערה 4.35. הוכחה בדרך השלילה היא שיטת הוכחה כללית במתמטיקה. כשמה כן היא: מניחים ההיפך ממה שרוצים להוכיח, ואם מקבלים סתירה זה אומר שהנחה זו שגויה ולכן ההיפך הוא הנכון. נוח להשתמש בשיטה זו כדי להראות שהגדרה כלשהי לא מתקיימת, כמו בדוגמה האחרונה.

הערה 4.36. שני ישרים בעלי וקטורי נורמל מקבילים (קו-לינאריים), הם בעצמם או מקבילים או מתלכדים. אכן, נסמן את וקטורי הכיוון של שני הישרים ב-$\vec{v_1}, \vec{v_2}$. אם וקטור הכיוון של הישר הראשון וקטור הנורמל של הישר הראשון הוא $\vec{n_1}$, אז לפי ההנחה וקטור הנורמל של הישר השני הוא $\vec{n_2} = c\vec{n_1}$ עבור $c \neq 0$ כלשהו.

$$\begin{aligned} \vec{v_2} \cdot \vec{n_2} = 0 \iff \vec{v_2} \cdot (c\vec{n_1}) = 0 \iff c(\vec{v_2} \cdot \vec{n_1}) = 0 \\ \iff \vec{v_2} \cdot \vec{n_1} = 0 \iff \vec{v_2} \perp \vec{n_1} \end{aligned}$$

הוקטורים היחידים שמאונכים ל-$\vec{n_1}$ הם $\vec{v_1}$ וכל הכפולות שלו בסקלר. לכן $\vec{v_1}, \vec{v_2}$ קו-לינאריים וזה אומר ששני הישרים או מקבילים או מתלכדים.

נסכם בטבלה את שלושת המצבים ההדדיים של שני ישרים

$$L_1 = \{\, (x_1, y_1) + t\vec{v_1} \mid t \in \mathbb{R} \,\},\ L_2 = \{\, (x_2, y_2) + t\vec{v_2} \mid t \in \mathbb{R} \,\}.$$

טבלה 4.1: מצבים הדדיים בין ישרים במישור

מצב	פירוש בהצגה הפרמטרית	פירוש בהצגה האלגברית	מספר נקודות חיתוך
נחתכים	$\vec{v_1},\vec{v_2}$ אינם קו-לינאריים	למערכת יש פתרון יחיד כי וקטורי הנורמל אינם קו-לינאריים	**1**
מקבילים	$\vec{v_1},\vec{v_2}$ הם קו-לינאריים, אך $(x_1,y_1) \notin L_2$	וקטורי הנורמל הם קו-לינאריים, אך המשוואות סותרות	**0**
מתלכדים	$\vec{v_1},\vec{v_2}$ הם קו-לינאריים וגם $(x_1,y_1) \in L_2$	וקטורי הנורמל הם קו-לינאריים והמשוואות שקולות	אינסוף (ישר)

דוגמה 4.37.

א. $L_1 = \{\, (x, y) \in \mathbb{R}^2 \mid 2x + 5y = 7 \,\}$ ו-$L_2 = \{\, (x, y) \in \mathbb{R}^2 \mid 5x - 4y = 1 \,\}$ הם ישרים נחתכים. רואים זאת כי וקטורי הנורמל, שהם $(2,5)$ ו-$(5,-4)$, אינם קו-לינאריים. אם הם היו קו-לינאריים, אז היה קיים $c \neq 0$ כך ש-$(5,-4) = c \cdot (2,5)$. זו סתירה כי מצד אחד נובע כי $c = \frac{5}{2}$ ומצד שני נובע כי $c = -\frac{4}{5}$. אז שני הישרים אכן נחתכים, ואפשר לחשב את נקודת החיתוך ע"י פתרון מערכת של שתי משוואות בשני נעלמים:

$$\begin{cases} 2x + 5y = 7 \\ 5x - 4y = 1 \end{cases}$$

אפשר להשתמש בשיטת ההצבה, או לחילופין להכפיל את המשוואה השנייה ב-$\frac{2}{5}$ ולהחסיר אותה מהמשוואה הראשונה (כדי להיפטר מ-x):

$$0 \cdot x + (5 + \frac{8}{5})y = 7 - \frac{2}{5} \implies \frac{33}{5}y = \frac{33}{5} \implies y = 1$$

נציב $y = 1$ במשוואה השנייה ונקבל

$$2x + 5 = 7 \implies 2x = 2 \implies x = 1.$$

אז נקודת החיתוך היחידה היא $(1,1)$, כלומר $L_1 \cap L_2 = \{\, (1,1) \,\}$.

ב. $L_1 = \{ (x,y) \in \mathbb{R}^2 \mid 2x+5y=7 \}$ ו-$L_3 = \{ (x,y) \in \mathbb{R}^2 \mid 4x+10y=14 \}$ הם ישרים מתלכדים. וקטורי הנורמל הם קו-לינאריים כי $(4,10) = 2 \cdot (2,5)$, ולכן הישרים בהכרח לא נחתכים. נסתכל על מערכת המשוואות:

$$\begin{cases} 2x+5y=7 \\ 4x+10y=14 \end{cases}$$

שתי המשוואות שקולות כי אפשר לקבל את המשוואה השנייה מהראשונה ע"י כפל ב-2.

ג. $L_1 = \{ (x,y) \in \mathbb{R}^2 \mid 2x+5y=7 \}$ ו-$L_3 = \{ (x,y) \in \mathbb{R}^2 \mid 6x+15y=20 \}$ הם ישרים מקבילים. וקטורי הנורמל הם קו-לינאריים כי $(6,15) = 3 \cdot (2,5)$, ולכן שוב הישרים לא נחתכים. אבל הפעם שתי המשוואות לא שקולות:

$$\begin{cases} 2x+5y=7 \\ 6x+15y=20 \end{cases}$$

אם נכפיל את המשוואה הראשונה ב-3 ונחסיר ממנה את המשוואה השנייה, נקבל את המשוואה השגויה $0 = 21 - 20$ וזו סתירה.

תרגיל 4.38. מהו המצב ההדדי בין הישר $L_1 = \{ (x,y) \in \mathbb{R}^2 \mid 3x-2y=1 \}$ והישר $L_2 = \{ (x,y) \in \mathbb{R}^2 \mid 15x-10y=5 \}$?

פתרון. הישרים מתלכדים כי משוואותיהם שקולות. אם נכפיל ב- 5 את המשוואה

$$3x-2y=1$$

נקבל את המשוואה

$$15x-10y=5.$$

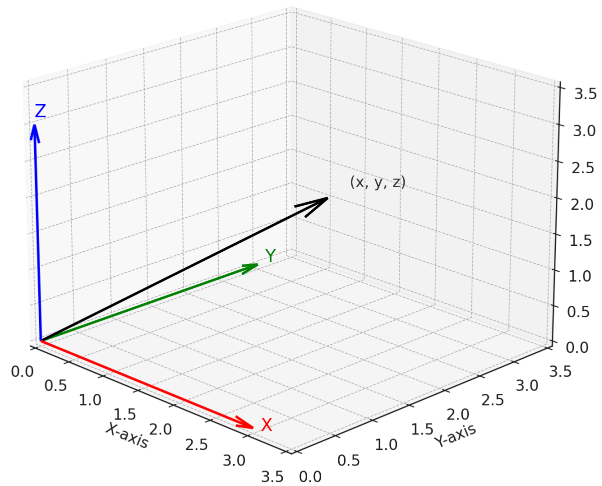
מבחינה גיאומטרית, שני וקטורי הנורמל $(3,-2), (15,-10)$ הם קו-לינאריים ולשני הישרים יש לפחות נקודה אחת משותפת, למשל הנקודה $(\frac{1}{3},0)$. אז זו דרך נוספת לראות שהם מתלכדים.

4.2 המרחב $\mathbb{R}^3$

עד עכשיו דיברנו על המישור $\mathbb{R}^2$. נרצה להרחיב את הדיון למרחב התלת-מימדי, שמוגדר באמצעות שלוש קוארדינטות.

הגדרה 4.39. המרחב $\mathbb{R}^3$ הוא קבוצת כל הנקודות בעלות שלוש קוארדינטות ממשיות, כלומר

$$\mathbb{R}^3 = \{ (x,y,z) \mid x,y,z \in \mathbb{R} \}.$$



איור 4.12: וקטור במרחב שיוצא מהראשית אל נקודה (x,y,z)

דוגמה 4.40.

א. $(0,0,0),(1,-2,5),(-3,-5,\pi),(\frac{3}{5},\sqrt{2},\sqrt{3}),(\pi^2,-10,\sqrt{5})\in\mathbb{R}^3$

ב. $1,(1,1),(i,0,0),(0,1,2,3)\notin\mathbb{R}^3$.

הרעיון של וקטור לא משתנה במרחב. אפשר לשנות את כל ההגדרות הרלוונטיות כך שתהיה התייחסות גם לקוארדינטה השלישית.

הגדרה 4.41. וקטור $\vec{v}=(x,y,z)$ במרחב הוא שלשה סדורה של מספרים ממשיים, כלומר איבר ב- $\mathbb{R}^3$.

הערה 4.42. גם כאן נוסיף להגדרה האלגברית את הפירוש הגיאומטרי של וקטור כחץ שיוצא מהראשית $(0,0,0)$ ומסתיים בנקודה (x,y,z). לכל וקטור יש גודל וכיוון והוא נקבע על ידם ללא קשר לנקודת ההתחלה.

הערה 4.43. שוב וקטורים שווים אם ורק אם יש שוויון בכל קוארדינטה, ללא קשר לנקודת ההתחלה. כלומר $(a_1,a_2,a_3)=(b_1,b_2,b_3)$ אם ורק אם $a_i=b_i$ לכל $1\le i\le 3$. בפרט, יש חשיבות לסדר ולכן $(3,2,1)\neq(1,3,2)\neq(1,2,3)$ וכן הלאה.

4.2.1 פעולות על וקטורים

גם במרחב שתי הפעולות הבסיסיות שניתן לבצע על וקטורים הן חיבור וכפל בסקלר.

חיבור: $(a_1, a_2, a_3) + (b_1, b_2, b_3) = (a_1 + b_1, a_2 + b_2, a_3 + b_3)$
כפל בסקלר: $c \cdot (a_1, a_2, a_3) = (ca_1, ca_2, ca_3)$

הערה 4.44. חיסור בין שני וקטורים זה שוב כמו חיבור בין הראשון לגרסה הנגדית של השני, כלומר

$$(b_1, b_2, b_3) - (a_1, a_2, a_3) = (b_1 - a_1, b_2 - a_2, b_3 - a_3). \qquad \textbf{(4.3)}$$

זה הוקטור שמתחיל בנקודה (a_1, a_2, a_3) ומסתיים ב-(b_1, b_2, b_3).

דוגמה 4.45.

א. $(2, 4, 6) + (1, 3, 5) = (2 + 1, 4 + 3, 6 + 5) = (3, 7, 11)$

ב. $2 \cdot (5, 3, 1) = (10, 6, 2)$

ג. $-3 \cdot (2, 5, 7) = (-6, -15, -21)$

ד. $(9, 2, 3) - (5, 8, 2) = (9 - 5, 2 - 8, 3 - 2) = (4, -6, 1)$

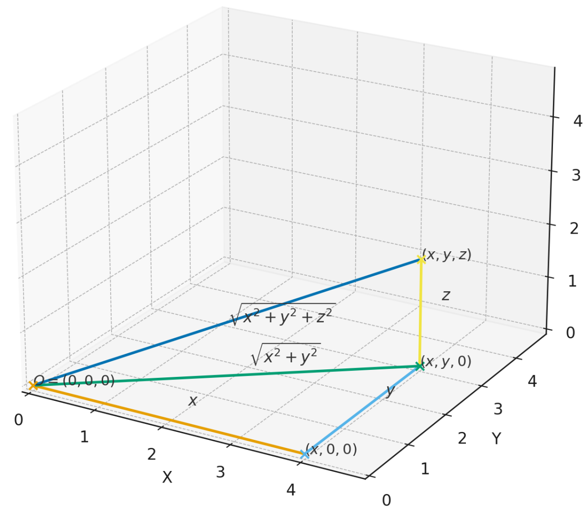
תרגיל 4.46. חשבו את הוקטור שיוצא מ- $(3, 1, 5)$ אל $(2, -6, 11)$.

פתרון. נציב את הנקודות ב- 4.3 ונקבל

$$(2, -6, 11) - (3, 1, 5) = (2 - 3, -6 - 1, 11 - 5) = (-1, -7, 6).$$

4.2.2 גדלים ב-$\mathbb{R}^3$

לכל וקטור יש גודל וכיוון. אם נניח שהוא מתחיל בראשית, אז גודלו הוא המרחק מנקודת הסיום לראשית שניתן לחשב לפי משפט פיתגורס:
גודל: $\|(x, y, z)\| = \sqrt{x^2 + y^2 + z^2}$



איור 4.13: שני משולשים ישרי-זווית במרחב

למעשה, החישוב מבוסס על שימוש במשפט פיתגורס פעמיים. היתר הירוק שמוטל על מישור xy הוא באורך $\sqrt{x^2+y^2}$ כי אורכי הניצבים הם x, y (בציור, או בערך מוחלט באופן כללי). לכן, אורך היתר הכחול של המשולש השני שאורכי ניצביו הם $|z|$, $\sqrt{x^2+y^2}$, הוא אכן $\sqrt{x^2+y^2+z^2}$.

הערה 4.47. אפשר גם לדבר על כיוון במרחב במונחים של שתי זוויות, אך לא נעשה זאת בקורס שלנו. נסתפק בוקטור כיוון.

דוגמה 4.48.

א. $\|(2,1,2)\| = \sqrt{2^2+1^2+2^2} = \sqrt{9} = 3$

ב. $\|(3,-4,2)\| = \sqrt{3^2+(-4)^2+2^2} = \sqrt{29}$

ג. $\|(-3,-6,2)\| = \sqrt{(-3)^2+(-6)^2+2^2} = \sqrt{49} = 7$

הטענה הבאה היא הכללה למרחב של טענה **4.12**.

טענה 4.49. *לכל וקטור* $(x,y,z) \in \mathbb{R}^3$ *ולכל סקלר* $c \in \mathbb{R}$ *מתקיים*

$$\|c\cdot(x,y,z)\| = |c|\cdot\|(x,y,z)\|.$$

תרגיל 4.50. בדקו שההוכחה של הטענה המקורית ניתנת להכללה למרחב.

פתרון. נשתמש בהגדרות ונחשב:

$$\begin{aligned}\|c\cdot(x,y)\| &= \|(cx,cy,cz)\| = \sqrt{(cx)^2+(cy)^2+(cz)^2} = \sqrt{c^2(x^2+y^2+z^2)}\\ &= \sqrt{c^2}\sqrt{x^2+y^2+z^2} = |c|\sqrt{x^2+y^2+z^2} = |c|\cdot\|(x,y,z)\|\end{aligned}$$

4.2.3 מכפלה סקלרית

גם במרחב אפשר להגדיר מכפלה סקלרית, ושוב יש קשר לזווית בין שני וקטורים.

הגדרה 4.51. עבור וקטורים $(a_1,a_2,a_3),(b_1,b_2.b_3)\in\mathbb{R}^3$ נגדיר

$$(a_1,a_2,a_3)\cdot(b_1,b_2,b_3) = a_1b_1+a_2b_2+a_3b_3.$$

דוגמה 4.52.

א. $(1,3,2)\cdot(2,5,-1) = 1\cdot2+3\cdot5+2\cdot(-1) = 15$

ב. $(3,-2,1)\cdot(6,5,-3) = 3\cdot6+(-2)\cdot5+1\cdot(-3) = 5$

ג. $(1,4,1)\cdot(-5,1,1) = 1\cdot(-5)+4\cdot1+1\cdot1 = 0$

התכונות של המכפלה הסקלרית שראינו בטענה 4.20 נשמרות במרחב, בלי שינוי מהותי להוכחות (רק הוספת קוארדינטה שלישית).

טענה 4.53. *יהיו* $\vec{v_1}=(x_1,y_1,z_!),\ \vec{v_2}=(x_2,y_2,z_2),\ \vec{v_3}=(x_3,y_3,z_3)\in\mathbb{R}^3$ *וקטורים. אז מתקיים:*

א. $\vec{v_1}\cdot\vec{v_2}=\vec{v_2}\cdot\vec{v_1}$.

ב. $(\vec{v_1}+\vec{v_3})\cdot\vec{v_2}=\vec{v_2}\cdot\vec{v_1}$.

ג. $\vec{v_1}\cdot\vec{v_1}=\|\vec{v_1}\|^2$.

ד. $\vec{v_1}\cdot(c\vec{v_2})=c(\vec{v_1}\cdot\vec{v_2})$.

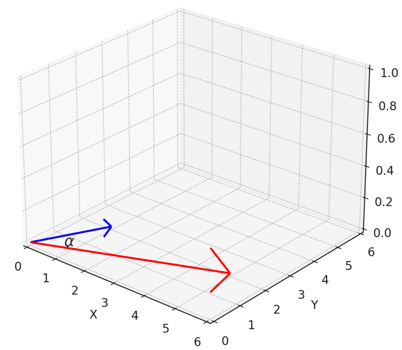
ה. $\|\vec{v_1}\|=0 \iff \vec{v_1}=(0,0)$.

ו. $\|\frac{\vec{v_1}}{\|\vec{v_1}\|}\|=1$ *אם* $\vec{v_1}\neq(0,0)$.

טענה 4.54. *יהיו* $\vec{v},\vec{w}$ *שני וקטורים ב-*$\mathbb{R}^3$ *השונים מ-*$(0,0,0)$*, ותהי* α *הזווית הקטנה ביניהם. אז מתקיים*

$$\cos\alpha = \frac{\vec{v}\cdot\vec{w}}{\|\vec{v}\|\cdot\|\vec{w}\|}.$$

הערה 4.55. כאן ההוכחה היא לא הכללה מיידית. נדלג עליה, אבל נציין שכל שני וקטורים שאינם מקבילים קובעים מישור (יחיד) וניתן לחשב את הזווית ביניהם כוקטורים באותו המישור. קל להבין זאת במקרה הפרטי שבו שני הוקטורים שייכים למישור xy שנתון ע"י המשוואה $z = 0$.



איור 4.14: שני וקטורים השייכים למישור xy

מבחינה גיאומטרית (לא אלגברית), את מישור זה ניתן להבין כאילו הוא $\mathbb{R}^2$. אכן, אם $\vec{v} = (x_1, y_1, 0)$ ו-$\vec{w} = (x_2, y_2, 0)$, אז מתקיים

$$\begin{cases} \vec{v} \cdot \vec{w} = x_1 x_2 + y_1 y_2 + 0^2 = x_1 x_2 + y_1 y_2 \\ \|\vec{v}\|| = \sqrt{x_1^2 + x_2^2 + 0^2} = \sqrt{x_1^2 + x_2^2} \\ \|\vec{w}\|| = \sqrt{y_1^2 + y_2^2 + 0^2} = \sqrt{y_1^2 + y_2^2} \end{cases}.$$

כלומר, במקרה זה אין הבדל בין הנוסחאות של המרחב לנוסחאות של המישור (במובן מסוים, אין חשיבות לקוארדינטה השלישית). בפרט, מתקיים $\cos\alpha = \frac{\vec{v} \cdot \vec{w}}{\|\vec{v}\|| \cdot \|\vec{w}\||}$.

המקרה הכללי יותר מסובך (יש חשיבות לקוארדינטה השלישית), אבל אין הבדל מהותי והנוסחה עדיין מתקיימת.

מסקנה 4.56. *נניח כי $\vec{v}, \vec{w}$ שני וקטורים ב-$\mathbb{R}^3$ השונים מ-$(0,0,0)$. אז הם מאונכים זה לזה אם ורק אם $\vec{v} \cdot \vec{w} = 0$.*

תרגיל 4.57. בדקו שההוכחה למקרה של המישור לא משתנה פה.

פתרון. תהי α הזווית הקטנה ביניהם. הנוסחה לא השתנתה ולכן שוב נקבל

$$\vec{v} \cdot \vec{w} = 0 \iff \cos\alpha = \frac{\vec{v} \cdot \vec{w}}{\|\vec{v}\| \cdot \|\vec{w}|} = 0 \iff \alpha = \frac{\pi}{2}.$$

4.2.4 ישרים ב-$\mathbb{R}^3$

בדומה לישרים במישור, ישר במרחב ניתן לתיאור פרמטרי

$$L = \{\ (x_0, y_0, z_0) + t \cdot \vec{v} \mid t \in \mathbb{R}\ \}$$

כאשר $\vec{v} \in \mathbb{R}^3$ נקרא וקטור כיוון ומקיים $\vec{v} \neq (0,0,0)$. $(x_0, y_0, z_0) \in \mathbb{R}^3$ היא נקודת התחלה שרירותית, ושוב אפשר לשנות את ההצגה הפרמטרית ע"י שינוי הנקודה (כל עוד היא על הישר) וכפל בסקלר של וקטור הכיוון (כל עוד הסקלר אינו 0).

דוגמה 4.58. $L = \{\ (1,1,1) + t \cdot (1,2,3) \mid t \in \mathbb{R}\ \}$ ניתן להצגה חלופית כמו

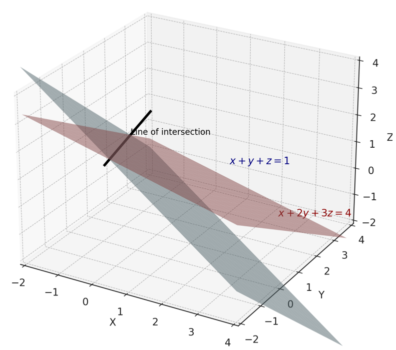
$$L = \{\ (2,3,4) + t \cdot (2,4,6) \mid t \in \mathbb{R}\ \}$$

כי $L \ni (1,2,3) \cdot 1 + (1,1,1) = (2,3,4)$ וגם $(2,4,6) = 2 \cdot (1,2,3)$.

ההבדל העיקרי לעומת ישרים במישור זה שבמרחב ההצגה האלגברית של הישר היא יותר מסורבלת (מערכת של שתי משוואות במקום אחת), ולכן ההצגה הפרמטרית היא יותר שימושית. במרחב, משוואה אחת מתארת מישור ולא ישר. למשל, ראינו כבר כי המשוואה $z = 0$ מתארת את מישור xy ולא ישר.

אינטואיטיבית: בדרך כלל משוואה מורידה מימד. לכן מתוך המרחב שמימדו 3, קבוצת הפתרונות של משוואה היא מישור ממימד 2. אם נוסיף עוד משוואה (חדשה, לא שקולה למשוואה הקודמת), המימד יירד ל-1 ונקבל ישר.

גיאומטרית: כאשר פותרים מערכת של שתי משוואות (לינאריות), מחשבים את החיתוך של שני מישורים. לרוב, שני מישורים נחתכים לאורך ישר ולכן ניתן לתאר ישר כחיתוך של שני מישורים שמכילים אותו. יש אינסוף מישורים כאלה.



איור 4.15: שני מישורים נחתכים לאורך ישר (שמוצג באופן חלקי)

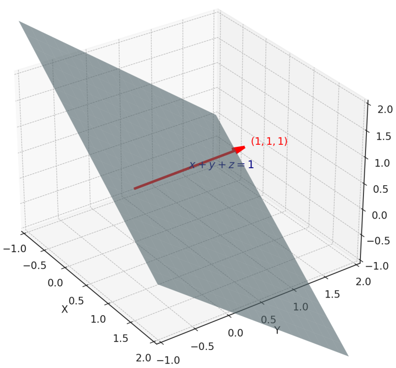
נעדיף לעבור מהצגה אלגברית של ישר (שהיא רחוקה מאוד מלהיות יחידה) להצגה פרמטרית, ולא להיפך. אבל ראשית צריך להבין מישורים במרחב יותר לעומק.

4.2.5 מישורים ב-$\mathbb{R}^3$

מישור כללי ב-$\mathbb{R}^3$ נתון בהצגה אלגברית כקבוצת פתרונות של משוואה אחת, כלומר

$$H = \left\{ (x,y,z) \in \mathbb{R}^3 \;\middle|\; ax + by + cz = d \right\}$$

כאשר $(a,b,c) \neq (0,0,0)$. נבין בקרוב שגם כאן (בדומה לישר במישור שמתואר ע"י משוואה אחת) המשמעות הגיאומטרית של (a,b,c) היא וקטור נורמל שמאונך לכל וקטור שמוכל במישור.



איור 4.16: מישור עם וקטור נורמל

דוגמה 4.59. נרצה למצוא נקודה על $H = \{ (x,y,z) \in \mathbb{R}^3 \mid 2x + 3y - z = 5 \}$. זו בחירה שרירותית, אבל אפשר למשל להציב $x = y = 0$ ולקבל $-z = 5 \implies z = -5$. אז $(0,0,-5) \in H$ ולכל $(x,y,z) \in H$ מתקיים

$$\begin{aligned} &(2,3,-1) \cdot (x-0, y-0, z-(-5)) = 2x + 3y - z - 5 = 0 \\ &\implies (2,3-1) \perp (x-0, y-0, z-(-5)) \end{aligned}$$

זה מראה כי הוקטור $(2,3,-1)$ מאונך לכל וקטור שמתחיל בנקודה $(0,0,-5)$ ומסתיים בנקודה על המישור. אבל $(0,0,-5)$ היא שרירותית לחלוטין, והיינו מקבלים מסקנה דומה לכל וקטור שמוכל במישור.

באופן כללי, עבור מישור $H = \{ (x,y,z) \in \mathbb{R}^3 \mid ax + by + cz = d \}$ ונקודה כלשהי $(x_0,y_0,z_0) \in H$ נקבל:

$$\begin{aligned} H &= \left\{ (x,y,z) \in \mathbb{R}^3 \;\middle|\; ax + by + cz = ax_0 + by_0 + cz_0 \right\} \\ &= \left\{ (x,y,z) \in \mathbb{R}^3 \;\middle|\; a(x-x_0) + b(y-y_0) + c(z-z_0) = 0 \right\} \\ &= \left\{ (x,y,z) \in \mathbb{R}^3 \;\middle|\; (a,b,c) \cdot (x-x_0, y-y_0, z-z_0) = 0 \right\} \\ &= \left\{ (x,y,z) \in \mathbb{R}^3 \;\middle|\; (a,b,c) \perp (x-x_0, y-y_0, z-z_0) \right\} \end{aligned} \tag{4.4}$$

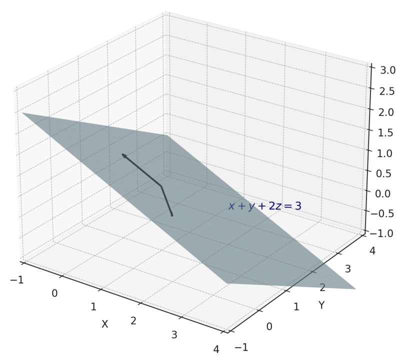
אז (a,b,c) הוא וקטור נורמל של H. הוא לא יחיד כי ההצגה האלגברית אינה יחידה, שכן ניתן להכפיל את המשוואה בכל סקלר $\alpha \neq 0$ ובהתאם לקבל $\alpha \cdot (a,b,c)$ כוקטור נורמל.

4.2.6 מעבר מהצגה אלגברית של מישור להצגה פרמטרית

למישור יש גם הצגה פרמטרית, אלא שבניגוד לישר דרושים שני פרמטרים כדי לתאר מישור. מישור נקבע ביחידות ע"י שלוש נקודות (באופן שקול: שלושה וקטורים) $\vec{v_0}, \vec{w_1}, \vec{w_2} \in \mathbb{R}^3$ שאינן שייכות לישר אחד. מבחינה וקטורית, המשמעות היא שהוקטורים $\vec{w_1} - \vec{v_0}, \vec{w_2} - \vec{v_0}$ אינם קו-לינאריים (מקבילים), כלומר אף אחד אינו כפל בסקלר של השני. אז המישור H שעובר דרך שלוש הנקודות, מתקבל מאוסף כל הוקטורים שיוצאים מ- $\vec{v_0}$ ונפרשים ע"י הוקטורים $\vec{w_1} - \vec{v_0}, \vec{w_2} - \vec{v_0}$. באופן מתמטי:

$$H = \{\, \vec{v_0} + t(\vec{w_1} - \vec{v_0}) + s(\vec{w_2} - \vec{v_0}) \mid t,s \in \mathbb{R} \,\}$$

תרגיל 4.60. איזו נקודה מתאימה להצבה $(t,s) = (0,0)$? להצבה $(t,s) = (1,0)$? להצבה $(t,s) = (0,1)$?



איור 4.17: מישור נוצר ע"י שני וקטורים שיוצאים מאותה הנקודה

פתרון. הנקודה המתאימה ל- $(t,s) = (0,0)$ היא $\vec{v_0}$. הנקודה המתאימה ל- $(t,s) = (1,0)$ היא

$$\vec{v_0} + 1 \cdot (\vec{w_1} - \vec{v_0}) = \vec{w_1}.$$

באופן דומה, הנקודה המתאימה ל- $(t,s) = (0,1)$ היא $\vec{w_2}$.

דוגמה 4.61. נמצא הצגה פרמטרית של המישור שעובר דרך הנקודות $(1,2,3), (1,-1,0), (2,3,1)$. נבחר נקודת התחלה $\vec{v_0} = (1,2,3)$ ונקבל:

$$\begin{aligned} H &= \{\, (1,2,3) + t(1-1,-1-2,0-3) + s(2-1,3-2,1-3) \mid t,s \in \mathbb{R} \,\} \\ &= \{\, (1,2,3) + t(0,-3,-3) + s(1,1,-2) \mid t,s \in \mathbb{R} \,\} \end{aligned}$$

הערה 4.62. כבר הדגשנו שהצגה פרמטרית של ישר אינה יחידה, ולכן זה לא מפתיע שגם הצגה פרמטרית של מישור אינה יחידה. כאן זה עוד יותר בולט: עבור מישור נתון (נניח בהצגה אלגברית), יש לנו הרבה חופש לבחור שלוש נקודות על המישור כל עוד הן לא על ישר אחד. אם אין הצגה אלגברית ורק נתונות שלוש נקודות (כי קיבלנו מידע חלקי), עדיין יש לנו חופש לבחור איזו מהן תהיה נקודת ההתחלה. בנוסף, ניתן להחליף את כל אחד משני וקטורי הכיוון בכפולה בסקלר השונה מ-0 כמו במקרה של ישר. אז למשל, גם מתקיים

$$H = \{\ (1,2,3) + t(0,1,1) + s(2,2,-4) \mid t,s \in \mathbb{R}\ \}.$$

דוגמה 4.63. נמצא הצגה פרמטרית למישור

$$H = \left\{\ (x,y,z) \in \mathbb{R}^3\ \middle|\ x + 2y + z = 4\ \right\}$$

לפי שלוש נקודות החיתוך עם הצירים (אפשר לבחור נקודות אחרות ובלבד שלא יהיו על ישר אחד). כדי למצוא את נקודת החיתוך עם ציר x נציב $y = z = 0$ ונקבל $x = 4$, כלומר את הנקודה $(4,0,0)$ שנבחר כנקודת ההתחלה. באופן דומה, שתי נקודות החיתוך הנוספות הן $(0,2,0), (0,0,4)$. לכן נקבל את ההצגה הפרמטרית הבאה:

$$\begin{aligned} H &= \{\ (4,0,0) + t(0-4, 2-0, 0-0) + s(0-4, 0-0, 4-0) \mid t,s \in \mathbb{R}\ \} \\ &= \{\ (4,0,0) + t(-4,2,0) + s(-4,0,4) \mid t,s \in \mathbb{R}\ \} \end{aligned}$$

אפשר, אך זה לא הכרחי, לכווץ את שני וקטורי הכיוון ולקבל:

$$H = \{\ (4,0,0) + t(-2,1,0) + s(-1,0,1) \mid t,s \in \mathbb{R}\ \}.$$

4.2.7 מעבר מהצגה פרמטרית של מישור להצגה אלגברית

אם נתונה הצגה פרמטרית, אז נתונים שני וקטורי כיוון. כל וקטור נורמל $\vec{n} = (a,b,c)$ מאונך לשניהם, וכדי לחשב אותו ניתן לפתור מערכת של שתי משוואות שמבטאת את העובדה שהמכפלות הסקלריות שלו עם שני וקטורי הכיוון מתאפסות. בדרך זו לא נקבל וקטור נורמל יחיד (כצפוי), אבל התשובה תהיה יחידה עד כדי כפל בסקלר וניתן לבחור את הסקלר כרצוננו. נלמד בפרק הבא איך לפתור מערכות משוואות לינאריות, אבל כאן מספיק לעשות פעולה אחת בין שתי המשוואות כדי לקבל משוואה עם שני משתנים בלבד. אם נבחר את הערך של אחד המשתנים באופן שרירותי, זה יקבע את הערכים של שני המשתנים אחרים.

דוגמה 4.64. נתון מישור בהצגה פרמטרית:

$$H = \{\ (1,-1,0) + t(2,3,1) + s(1,4,-2) \mid t,s \in \mathbb{R}\ \}$$

נדרוש כי $\vec{n} = (a,b,c)$ יהיה מאונך לוקטורי הכיוון, ונקבל

$$\begin{cases} 2a + 3b + c = 0 \\ a + 4b - 2c = 0 \end{cases}$$

נרצה לקבל משוואה עם שני משתנים במקום שלושה. אפשר לבודד את a במשוואה השנייה ולהציב אותו במשוואה הראשונה, או לחילופין להכפיל את המשוואה השנייה ב-2 ואז להחסיר ממנה את המשוואה הראשונה. כך נקבל $5b-5c=0$ או באופן שקול $b=c$. נציב את מה שקיבלנו באחת מהמשוואות המקוריות, למשל המשוואה השנייה, ונקבל $a+2c=0$ או באופן שקול $a=-2c$. אז כל וקטור נורמל הוא מהצורה $c(-2,1,1)$ ונוח לבחור $c=1$. כדי למצוא הצגה אלגברית צריך להציב את נקודת ההתחלה ואת a,b,c שחישבנו ב- **(4.4)**:

$$\begin{aligned} H &= \left\{ (x,y,z)\in\mathbb{R}^3 \;\middle|\; -2x+y+z=-2\cdot 1+(-1)+0 \right\} \\ &= \left\{ (x,y,z)\in\mathbb{R}^3 \;\middle|\; -2x+y+z=-3 \right\} \end{aligned}$$

תרגיל 4.65. פתרו את מערכת המשוואות לעיל בדרך אחרת, כך שתקבלו משוואה אחת שלא כוללת את c. בדקו שעדיין מקבלים אותו וקטור נורמל עד כדי כפל בסקלר.

פתרון. נחזור אל המשוואות עבור וקטור הנורמל:

$$\begin{cases} 2a+3b+c=0 \\ a+4b-2c=0 \end{cases}$$

כדי להיפטר מהמשתנה c, נכפיל את המשוואה הראשונה ב- 2 ונחבר אליה את המשוואה השנייה. כך נקבל את המשוואה

$$.5a+10b=0 \iff a=-2b$$

כעת נציב $a=-2b$ באחת המשוואות המקוריות, למשל הראשונה, ונקבל:

$$.-4b+3b+c=0 \iff c=b$$

אז קיבלנו משוואות שקולות למשוואות שקיבלנו בדוגמה, אבל הפעם a,c תלויים ב- b. נבחר $b=1$ וזה שוב יוביל לערכים $a=-2, c=1$ ולכן וקטור הנורמל וההצגה האלגברית זהים לאלה מהדוגמה. אם נבחר $c=2$ או כל ערך השונה מ- 0, נקבל הצגה אלגברית שקולה (כפל בסקלר של וקטור הנורמל ולכן גם של המשוואה).

הערה 4.66. יש דרך אחרת לחשב וקטור נורמל, שנקראת מכפלה וקטורית. זו מכפלה שלוקחת שני וקטורים ב- $\mathbb{R}^3$ ומחזירה וקטור שלישי שמאונך לשניהם. נושא זה לא נלמד בקורס, אבל נתייחס אליו בקצרה (לידע כללי) בהמשך כי הוא קשור לדטרמיננטה. במסגרת הקורס יש לחשב וקטורי נורמל ע"י פתרון מערכת משוואות לינאריות כמו בדוגמה.

4.2.8 מצבים הדדיים בין מישורים במרחב

קיימים שלושה מצבים הדדיים אפשריים בין שני מישורים $H_1, H_2 \subseteq \mathbb{R}^3$. לצורך הדיון נניח כי $(x_1, y_1, z_1) \in H_1$. נתאר את שלושת המצבים בטבלה:

טבלה 4.2: מצבים הדדיים בין מישורים במרחב

מצב	משמעות גיאומטרית	משמעות אלגברית	מספר נקודות חיתוך
נחתכים	וקטורי הנורמל אינם קו-לינאריים	פתרון המערכת כולל משתנה חופשי	אינסוף (ישר)
מקבילים	וקטורי הנורמל קו-לינאריים, אך $(x_1, y_1, z_1) \notin H_2$	שתי המשוואות סותרות	0
מתלכדים	וקטורי הנורמל קו-לינאריים וגם $(x_1, y_1, z_1) \in H_2$	שתי המשוואות שקולות	אינסוף (מישור)

דוגמה 4.67.

א. המישורים

$$H_1 = \left\{ (x,y,z) \in \mathbb{R}^3 \;\middle|\; x+y+z=1 \right\}$$
$$H_2 = \left\{ (x,y,z) \in \mathbb{R}^3 \;\middle|\; 3x+2y+z=3 \right\}$$

נחתכים. הסיבה הגיאומטרית היא שוקטורי הנורמל $(1,1,1), (3,2,1)$ אינם קו-לינאריים. נוודא זאת: נניח בשלילה שקיים $\alpha \in \mathbb{R}$ כך ש- $(3,2,1) = \alpha \cdot (1,1,1)$. אז קיבלנו $\alpha = 3$ לפי הקוארדינטה הראשונה, ומצד שני $\alpha = 2$ לפי הקוארדינטה השנייה. זו סתירה.

אם ננסה לפתור את מערכת המשוואות (נדון בכך בהרחבה בפרק הבא), נראה שהפתרון כולל משתנה חופשי. נחסיר את המשוואה $x+y+z=1$ מהמשוואה $3x+2y+z=3$, ונקבל

$$2x+y=2 \iff y=2-2x.$$

x הוא משתנה חופשי שניתן להציב בו ערך שרירותי, שהוא הפרמטר שיופיע בהצגה הפרמטרית של הישר. נציב $x=t$ ובהתאם $y=2-2t$ במשוואה $x+y+z=1$, ונקבל

$$t+2-2t+z=1 \iff z=t-1.$$

לכן מתקיים

$$H_1 \cap H_2 = \left\{ (t, 2-2t, t-1) \mid t \in \mathbb{R} \right\}.$$

ב. נסתכל על המישורים

$$H_1 = \left\{ (x, y, z) \in \mathbb{R}^3 \;\middle|\; x - 2y + 5z = 4 \right\}$$
$$H_2 = \left\{ (x, y, z) \in \mathbb{R}^3 \;\middle|\; 3x - 6y + 15z = 13 \right\}$$

וקטורי הנורמל הם קו-לינאריים כי $(3, -3, 15) = 3 \cdot (1, -1, 5)$. לכן המישורים או מקבילים או מתלכדים. ניתן לראות ששתי המשוואות סותרות. נכפיל את המשוואה הראשונה ב- 3 ונחסיר אותה מהמשוואה השנייה. אז נקבל $0 = 13 - 12$ וזו אכן סתירה.

לעומת זאת, המישור

$$H_3 = \left\{ (x, y, z) \in \mathbb{R}^3 \;\middle|\; 3x - 6y + 15z = 12 \right\}$$

מתלכד עם H_1 כי המשוואות שקולות. כפל ב- 3 מראה כי

$$3x - 6y + 15z = 12 \iff x - 2y + 5z = 4.$$

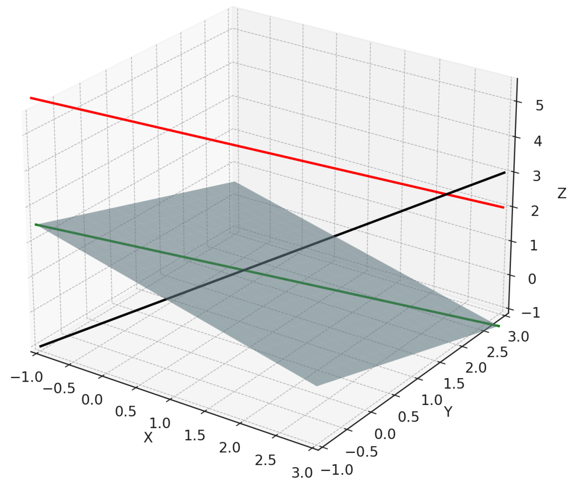
מכאן גם נובע כי H_3 מקביל ל- H_2, וזה בעצם מיידי מהמשוואות עצמן. הן בפירוש סותרות.

4.2.9 מצבים הדדיים בין ישר למישור במרחב

קיימים שלושה מצבים הדדיים אפשריים בין ישר L ומישור H במרחב. נניח כי $\vec{v}$ הוא וקטור כיוון של L, ו- $\vec{n}$ הוא וקטור נורמל של H. בנוסף, נניח כי $(x_0, y_0, z_0) \in L$.

טבלה 4.3: מצבים הדדיים בין ישר למישור במרחב

מצב	משמעות גיאומטרית	משמעות אלגברית	מספר נקודות חיתוך
נחתכים	$\vec{v}$ לא מאונך ל-$\vec{n}$	הצבת ההצגה הפרמטרית של L במשוואה של H מובילה לפתרון יחיד	1
מקבילים	$\vec{v}$ מאונך ל-$\vec{n}$, אך $(x_0, y_0, z_0) \notin H$	ההצבה מובילה לסתירה (פסוק שקר)	0
המישור מכיל את הישר	$\vec{v}$ מאונך ל-$\vec{n}$ וגם $(x_0, y_0, z_0) \in H$	ההצבה מובילה לזהות (פסוק אמת)	אינסוף (ישר)



איור 4.18: מישור וישר החותך אותו, ישר המקביל לו וישר המוכל בו

דוגמה 4.68.

א. נסתכל על הישר $L = \{ (1,0,3) + t(2,-4,5) \mid t \in \mathbb{R} \}$ והמישור

$.H = \left\{ (x,y,z) \in \mathbb{R}^3 \;\middle|\; 3x + 7y + 2z = 4 \right\}$

נחשב את המכפלה הסקלרית של וקטור הכיוון של הישר עם וקטור הנורמל של המישור:

$$(2,-4,5) \cdot (3,7,2) = 6 - 28 + 10 = -12 \neq 0$$

לכן הישר והמישור בהכרח נחתכים. נחשב את נקודת החיתוך ע"י הצבת ההצגה הפרמטרית של הישר במשוואת המישור:

$$.3(1+2t) + 7(-4t) + 2(3+5t) = 4 \iff -12t + 9 = 4 \iff t = \frac{5}{12}$$

נציב $t = \frac{5}{12}$ בהצגה הפרמטרית ונקבל את נקודת החיתוך:

$$(1 + 2 \cdot \frac{5}{12}, -4 \cdot \frac{5}{12}, 3 + 5 \cdot \frac{5}{12}) = (\frac{11}{6}, -\frac{5}{3}, \frac{61}{12})$$

ב. נסתכל על הישר $L = \{ (1,0,3) + t(2,-4,5) \mid t \in \mathbb{R} \}$ והמישור

$$H_1 = \left\{ (x,y,z) \in \mathbb{R}^3 \;\middle|\; 5x - 2z = 1 \right\}.$$

וקטור הכיוון $(2,-4,5)$ מאונך לוקטור הנורמל $(5,0,-2)$ כי

$$(2,-4,5) \cdot (5,0,-2) = 10 + 0 - 10 = 0.$$

כדי לקבוע אם L מקביל או מוכל ב- H_1, אפשר לבדוק אם $(1,0,3) \in H_1$. מתקיים $5 \cdot 1 - 2 \cdot 3 = -1 \neq 1$, ולכן הישר מקביל למישור. זה מספיק לקביעת המצב ההדדי, אבל אפשר גם להציב את ההצגה הפרמטרית של הישר במשוואת המישור ולקבל

$$5(1+2t) - 2(3+5t) = 1 \iff -1 = 1.$$

הסתירה מראה באופן מפורש שאין נקודת חיתוך.

לעומת זאת, עבור המישור

$$H_2 = \left\{ (x,y,z) \in \mathbb{R}^3 \;\middle|\; 5x - 2z = -1 \right\}$$

החישוב לעיל (עם תיקון אגף ימין) מראה כי L מוכל בו. ההצבה מובילה לזהות $-1 = -1$ ולכן כל נקודה על L גם נמצאת על H_2.

פרק 5

מערכות משוואות לינאריות

בפרק הקודם עסקנו במשוואות לינאריות עם שני משתנים x, y לתיאור המישור, וגם בשלושה משתנים x, y, z לתיאור המרחב. אנחנו רוצים להכליל את הדיון למספר (טבעי) כלשהו של משתנים. מהי המוטיבציה? משוואות לינאריות מופיעות בפיזיקה כבר בחוקי ניוטון. למשל, החוק השני מתאר תלות לינארית מהצורה $\vec{F} = m\vec{a}$ בין וקטור הכוח $\vec{F}$ הפועל על גוף לבין לוקטור התאוצה $\vec{a}$ של הגוף, כאשר מסת הגוף m היא סקלר חיובי. למעשה, המשוואה הוקטורית $\vec{F} = m\vec{a}$ מבטאת שלוש משוואות - אחת לכל קוארדינטה של המרחב.

כאשר יש הרבה גופים בבעיה עם תלות ביניהם (אילוצים), יכולות להתקבל מספר משוואות לינאריות במספר משתנים ויחד הן נקראות מערכת משוואות לינאריות. זה המקרה גם בתחומים רבים אחרים, כמו למשל רשתות נוירונים בהן כל נוירון מקבל קלטים ומפיק מהם פלט לפי חישוב שמתאים למשוואה לינארית. לא תמיד יש הקשר גיאומטרי/פיזיקלי לבעיה שאותה מבקשים להבין באמצעות מערכת משוואות לינאריות.

תחילה נבהיר איך נראית משוואה לינארית במספר משתנים ואיך מחשבים את קבוצת הפתרונות שלה.

5.1 משוואה לינארית במספר משתנים

איך נראית משוואה לינארית כללית? נשתמש במשתנים $x_1, x_{2,}..., x_n$ במקום אותיות רגילות, כי לצורך העניין n יכול להיות גדול בהרבה ממספר האותיות באלף בית הלועזי (26) או בכל שפה אחרת. גם נראה בהמשך שצריך לבחור איזשהו סדר בין המשתנים, אז אם מספרם גדול זה נוח להשתמש באינדקס מספרי כדי להדגיש את הסדר ביניהם. זה נוח גם למספר משתנים לא גדול במיוחד.

הגדרה 5.1. משוואה לינארית במשתנים $x_1, x_{2,}..., x_n$ היא משוואה מהצורה

$$a_1x_1 + a_2x_2 + a_3x_3 + ... + a_nx_n = b \qquad (5.1)$$

כאשר $a_1, ..., a_n$ הם סקלרים נתונים שנקראים מקדמים, ו-b הוא סקלר נתון שנקרא קבוע או איבר חופשי. אם כל הסקלרים והמשתנים ממשיים, נאמר שזו משוואה לינארית מעל $\mathbb{R}$. אם לפחות סקלר אחד הוא מרוכב ו/או המשתנים מרוכבים, נאמר שזו משוואה לינארית מעל $\mathbb{C}$. אם נרצה להכליל בלי לציין שמדובר ב-$\mathbb{R}$ או $\mathbb{C}$, נכתוב $\mathbb{F}$ ונזכור שקבוצה זו היא $\mathbb{R}$ או $\mathbb{C}$.

דוגמה 5.2.

א. $x_1 - x_2 + x_3 - x_4 = 100$ זו משוואה לינארית בארבעה משתנים מעל $\mathbb{R}$. דוגמה לפתרון היא $(x_1, x_2, x_3, x_4) = (50, 0, 50, 0)$ כי $50 - 0 + 50 - 0 = 100$.

ב. $x_1 + ix_2 - x_3 - ix_4 + x_5 = 1 + i$ זו משוואה לינארית בחמישה משתנים מעל $\mathbb{C}$. דוגמה לפתרון היא $(x_1, x_2, x_3, x_4, x_5) = (1, 1, 0, 0, 0)$ כי $1 + i \cdot 1 - 0 - 0 + 0 = 1 + i$.

ג. $0 = 0$ זו משוואה לינארית במספר כלשהו של משתנים, מעל $\mathbb{R}$ או $\mathbb{C}$. בכתיבה מלאה עבור n משתנים מדובר על המשוואה

$$0 \cdot x_1 + 0 \cdot x_2 + 0 + ... + 0 \cdot x_n = 0.$$

כל בחירה של סקלרים $(x_1, x_2, ..., x_n) = (c_1, c_2, ..., c_n)$ היא פתרון למשוואה זו, שהיא פסוק אמת שלא תלוי בערכי המשתנים.

הגדרה 5.3. סדרת סקלרים $(c_1, c_2, ..., c_n)$ נקראת n-יה, או וקטור באורך n. נגדיר את מרחב ה-n-יות מעל $\mathbb{R}$:

$$\mathbb{R}^n = \{ (c_1, c_2, ..., c_n) \mid c_1, c_2, ..., c_n \in \mathbb{R} \}$$

וגם את מרחב ה-n-יות מעל $\mathbb{C}$:

$$\mathbb{C}^n = \{ (c_1, c_2, ..., c_n) \mid c_1, c_2, ..., c_n \in \mathbb{C} \}$$

לעיתים נכתוב $\mathbb{F}^n$ במקום לפרט אם הכוונה היא ל-$\mathbb{R}^n$ או ל-$\mathbb{C}^n$.

יש לנו אינטואיציה גיאומטרית טובה לגבי המישור הדו-מימדי $\mathbb{R}^2$ והמרחב התלת-מימדי $\mathbb{R}^3$. החל מ- $\mathbb{R}^4$ ו- $\mathbb{C}^2$ כבר אין לנו יכולת לדמיין את המרחבים האלה באופן מלא, אבל נתרגל גם אליהם מבחינה אלגברית.

הגדרה 5.4. פתרון של משוואה לינארית הינו n-יה $(c_1, c_2, ..., c_n) \in \mathbb{F}^n$ כך שבהצבת

$$x_1 = c_1, x_2 = c_2, ..., x_n = c_n$$

המשוואה מתקיימת.

דוגמה 5.5.

א. משוואה לינארית ב-5 משתנים מעל $\mathbb{R}$:

$$x_1 + x_2 + x_3 + x_4 + x_5 = 100$$

אחד הפתרונות של המשוואה הוא $(x_1, x_2, x_3, x_4, x_5) = (20, 20, 20, 20, 20)$. פתרון נוסף הוא $(x_1, x_2, x_3, x_4, x_5) = (60, 10, 10, 10, 10)$, ויש עוד אינסוף אחרים. לעומת זאת, $(x_1, x_2, x_3, x_4, x_5) = (0, 0, 0, 0, 0)$ אינו פתרון כי ההצבה נותנת $0 = 100$ וזה כמובן לא מתקיים.

כדי למצוא את קבוצת כל הפתרונות נבודד את אחד המשתנים. נוח לבחור את x_1 כי הוא הראשון. נקבל

$$x_1 = 100 - x_2 - x_3 - x_4 - x_5.$$

זה למעשה מראה שניתן לבחור את ערכיהם של x_2, x_3, x_4, x_5 באופן שרירותי. זו משוואה מעל $\mathbb{R}$ ולכן נבחר $c_2, c_3, c_4, c_5 \in \mathbb{R}$ שרירותיים ונציב במשוואה

$$(x_2, x_3, x_4, x_5) = (c_2, c_3, c_4, c_5).$$

כך נקבל

$$x_1 = 100 - c_2 - c_3 - c_4 - c_5.$$

לכן, קבוצת הפתרונות נתונה באופן פרמטרי (עם 4 פרמטרים c_2, c_3, c_4, c_5) ע"י

$$\{\ (100 - c_2 - c_3 - c_4 - c_5, c_2, c_3, c_4, c_5) \mid c_2, c_3, c_4, c_5 \in \mathbb{R}\ \}.$$

ב. משוואה לינארית ב-4 משתנים מעל $\mathbb{C}$:

$$x_1 + ix_2 + 2ix_3 + x_4 = 1 + 4i$$

אחד הפתרונות של המשוואה הוא $(x_1, x_2, x_3, x_4) = (1, 1, 1, i)$. נבדוק זאת:

$$1 \cdot 1 + i \cdot 1 + 2i \cdot 1 + 1 \cdot i = 1 + 4i$$

לעומת זאת, $(x_1, x_2, x_3, x_4) = (1, 0, 1, 0)$ אינו פתרון. הפעם בדיקה מראה כי

$$1 \cdot 1 + i \cdot 0 + 2i \cdot 1 + 1 \cdot 0 = 1 + 2i \neq 1 + 4i.$$

נבודד את x_1 כדי למצוא את קבוצת הפתרונות:

$$x_1 = 1 + 4i - ix_2 - 2ix_3 - x_4$$

זו משוואה מעל $\mathbb{C}$ ולכן נבחר $c_2, c_3, c_4 \in \mathbb{C}$ שרירותיים ונציב אותם במשוואה. כך נקבל

$$x_1 = 1 + 4i - ic_2 - 2ic_3 - c_4.$$

לכן, קבוצת הפתרונות נתונה באופן פרמטרי (עם 3 פרמטרים c_2, c_3, c_4) ע"י

$$\{\, (1 + 4i - ic_2 - 2ic_3 - c_4, c_2, c_3, c_4) \mid c_2, c_3, c_4 \in \mathbb{C} \,\}.$$

איך פותרים משוואה לינארית כללית מהצורה 5.1? אפשר למצוא את כל הפתרונות ע"י בידוד אחד המשתנים. מקובל לעשות זאת עבור x_1 בהנחה ש- $a_1 \neq 0$ (אם $a_1 = 0$, נבחר את המשתנה הראשון עם מקדם השונה מ-0). ראשית מעבירים לאגף ימין את כל הביטויים של המשתנים האחרים:

$$a_1 x_1 = b - a_2 x_2 - a_3 x_3 - ... - a_n x_n.$$

נחלק בסקלר a_1 ונקבל

$$x_1 = \frac{1}{a_1}(b - a_2 x_2 - a_3 x_3 - ... - a_n x_n).$$

זה מראה שלכל בחירה של סקלרים $c_2, ..., c_n \in \mathbb{F}$ נקבל פתרון

$$(x_1, x_2, x_3, ..., x_n) = (\frac{1}{a_1}(b - a_2 c_2 - a_3 c_3 - ... - a_n c_n), c_2, c_3, ..., c_n).$$

לכן, קבוצת הפתרונות של המשוואה נתונה בהצגה פרמטרית ע"י

$$\left\{ \, (\frac{1}{a_1}(b - a_2 c_2 - a_3 c_3 - ... - a_n c_n), c_2, c_3, ..., c_n) \;\middle|\; c_1, c_2, ..., c_n \in \mathbb{F} \right\}.$$

נדגיש כי $a_1, a_2, a_3, ..., a_n$ נתונים במשוואה כסקלרים קבועים. $c_1, c_2, ..., c_n$ הם פרמטרים שרצים על פני הקבוצה $\mathbb{F}$ ($\mathbb{R}$ או $\mathbb{C}$) כדי ליצור את כל הפתרונות. זו הכללה (כלומר צורה יותר כללית של מקרה מוכר) של הצגה פרמטרית של ישר (פרמטר אחד) ושל מישור (שני פרמטרים) למספר כללי של פרמטרים, שהוא בעצם $n-1$. בפרט, לכל $n \geq 2$ יש אינסוף פתרונות למשוואה 5.1 כי ניתן להציב אינסוף אפשרויות במשתנה אחד או יותר.

שימו לב שהתייחסנו אל x_1 כמשתנה תלוי ואל $x_2, x_3, ..., x_n$ כמשתנים חופשיים (הצבנו בהם ערכים שרירותיים). אם היינו מבודדים את x_2 או משתנה אחר במקום x_1, אז הוא היה המשתנה התלוי וההצגה הפרמטרית הייתה משתנה בהתאם. עדיין, זו רק הצגה חלופית לקבוצת הפתרונות של אותה המשוואה. טבעי לבחור את x_1, אך אין חובה לעשות זאת.

דוגמה 5.6. את המשוואה

$$x_1 + 2x_2 + 3x_3 + 4x_4 = 4$$

אפשר לפתור ע"י בידוד אחד מארבעת המשתנים (זה נתון לבחירתנו). מקובל לבודד את x_1 ולקבל

$$.x_1 = 4 - 2x_2 - 3x_3 - 4x_4$$

בהצבת $(x_2, x_3, x_4) = (c_2, c_3, c_4)$ מקבלים את קבוצת הפתרונות

$$.\{\ (4 - 2c_2 - 3c_3 - 4c_4, c_2, c_3, c_4) \mid c_2, c_3, c_4 \in \mathbb{R}\ \}$$

אבל באותה המידה אפשר לבודד משתנה אחר, למשל x_2. כך נקבל $2x_2 = 4 - x_1 - 3x_3 - 4x_4$, ולאחר חלוקה ב- 2:

$$.x_2 = \frac{1}{2}(4 - x_1 - 3x_3 - 4x_4)$$

בהצבת סקלרים $(x_1, x_3, x_4) = (d_1, d_3, d_4)$ נקבל הצגה פרמטרית חדשה לקבוצת הפתרונות, שנתונה ע"י

$$.\left\{\ (d_1, \frac{1}{2}(4 - d_1 - 3d_3 - 4d_4), d_3, d_4)\ \middle|\ d_1, d_3, d_4 \in \mathbb{R}\ \right\}$$

בפרט, נובע כי מתקיים שוויון בין שתי הקבוצות. מדובר בשתי הצגות שקולות.

תרגיל 5.7. מצאו את קבוצת הפתרונות של המשוואה $2x_1 + x_2 + 3x_3 + 4x_4 = 5$. בדקו איך ההצגה הפרמטרית משתנה כאשר מבודדים את x_2 במקום x_1.

פתרון. נבודד תחילה את x_1 ונקבל

$$.x_1 = \frac{1}{2}(5 - x_2 - 3x_3 - 4x_4)$$

נציב $(x_2, x_3, x_4) = (c_2, c_3, c_4)$ ונקבל את ההצגה הפרמטרית הבאה לקבוצת הפתרונות:

$$\left\{\ (\frac{1}{2}(5 - c_2 - 3c_3 - 4c_4), c_2, c_3, c_4)\ \middle|\ c_2, c_3, c_4 \in \mathbb{R}\ \right\}$$

כעת נבודד את x_2 וזה יוביל להצגה פרמטרית נוספת של אותה קבוצת הפתרונות. ראשית נקבל

$$.x_2 = 5 - 2x_1 - 3x_3 - 4x_4$$

נציב $(x_1, x_3, x_4) = (d_1, d_3, d_4)$ ונקבל את ההצגה הפרמטרית הבאה:

$$\{\ (d_1, 5 - 2d_1 - 3d_3 - 4d_4, d_3, d_4) \mid d_1, d_3, d_4 \in \mathbb{R}\ \}$$

5.2 מערכות משוואות לינאריות

בפרק הקודם ראינו הרבה דוגמאות למערכות משוואות לינאריות (ממ"ל) מעל $\mathbb{R}$ עם שתי משוואות בשני משתנים (לתיאור חיתוך שני ישרים במישור) וגם לממ"ל עם שתי משוואות בשלושה משתנים (לתיאור חיתוך שני מישורים במרחב). באופן כללי, אפשר להתייחס לכל אוסף של משוואות לינאריות כמערכת אחת.

בהקשר של החוק השני של ניוטון שהוזכרו בתחילת הפרק כמשוואה וקטרית מהצורה $\vec{F}=m\vec{a}$, אפשר לחשוב על ממ"ל במספר כלשהו של משתנים שקשורים לוקטורי התאוצה של מספר גופים ולכוחות הפועלים עליהם. גודל הממ"ל מקשה על האינטואיציה הגיאומטרית ומוסיף חישובים (ייתכן באופן משמעותי), אך עדיין יש גישה אלגברית אחת לפתרון של כל ממ"ל. נפתח אותה בפרק זה.

דוגמה 5.8.

א. ממ"ל של שתי משוואות בארבעה משתנים מעל $\mathbb{R}$:

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

ב. ממ"ל של שלוש משוואות בשלושה משתנים מעל $\mathbb{R}$:

$$\begin{cases} x+y+z=3 \\ x+2y+z=4 \\ x+y+3z=5 \end{cases}$$

כאן חזרנו להשתמש באותיות המוכרות x,y,z במקום x_1,x_2,x_3. הסימונים לא באמת חשובים כל עוד יש עקביות (נוח להשתמש באותיות רגילות עבור מספר קטן של משתנים).

ג. ממ"ל של שתי משוואות בארבעה משתנים מעל $\mathbb{C}$:

$$\begin{cases} iz_1+z_2 \qquad\qquad +iz_3+z_4 \;\;=2i \\ z_1\;\;+(1+i)z_2+z_3\;\;+2iz_4=2 \end{cases}$$

הגדרה 5.9. מערכת משוואות לינאריות (ממ"ל) במשתנים $x_1,x_{2,\ldots},x_n$ היא אוסף של מספר כלשהו (m) של משוואות לינאריות מהצורה

$$\begin{cases} a_{11}x_1 + \ a_{12}x_2 + \ a_{13}x_3 + \ \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \ a_{22}x_2 + \ a_{23}x_3 + \ \cdots + a_{2n}x_n = b_2 \\ a_{31}x_1 + \ a_{32}x_2 + \ a_{33}x_3 + \ \cdots + a_{3n}x_n = b_3 \\ \quad\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{5.2}$$

אם כל המקדמים והקבועים הם מספרים ממשיים, נאמר שהממ"ל מעל $\mathbb{R}$. אחרת, נאמר שהממ"ל מעל $\mathbb{C}$.

שימו לב לאינדקס הכפול: a_{ij} הוא המקדם של המשתנה x_j במשוואה ה- i (כלומר בשורה ה-i). b_i הוא הקבוע (איבר חופשי) במשוואה ה- i. צריך להתרגל לכתיבה עם שני אינדקסים, אבל זו הדרך הטבעית להתייחס גם למספר המשתנה וגם למספר המשוואה. נדגיש ש- ij אינו מכפלה וגם לא מספר דו-ספרתי בהקשר זה. בנוסף, כאן i אינו המספר המדומה אלא אינדקס שמתאר את מספר השורה, ואילו j הוא אינדקס שמתאר את מספר העמודה. אלה הסימונים המקובלים בקורס, והכוונה אמורה להיות ברורה כי אינדקס מופיע בכתב תחתי.

הגדרה 5.10. פתרון של ממ"ל הינו סדרת סקלרים $(c_1, c_2, ..., c_n)$ כך שבהצבת

$$x_1 = c_1, x_2 = c_2, ..., x_n = c_n$$

<u>כל</u> המשוואות בממ"ל מתקיימות.

דוגמה 5.11.

א. נחזור לדוגמה קודמת:

$$\begin{cases} x + \ \ y + \ \ z = 3 \\ x + 2y + \ \ z = 4 \\ x + \ \ y + 3z = 5 \end{cases}$$

כאן המקדמים הם

$$a_{11} = a_{12} = a_{13} = a_{21} = a_{23} = a_{31} = a_{32} = 1,\ a_{22} = 2,\ a_{33} = 3.$$

הקבועים הם $b_1 = 3,\ b_2 = 4,\ b_3 = 5$.

ב.

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 4 \\ \qquad\; x_2 + x_3 + x_4 = 0 \\ x_1 + \qquad\; x_3 + x_4 = 0 \end{cases}$$

כאן הממ"ל לא כתובה באופן מלא, כי בכל משוואה שבה חסר משתנה המשמעות היא שהמקדם המתאים הוא 0 (הרווחים נועדו להדגשה). בכתיבה מלאה, הממ"ל נראית כך:

$$\begin{cases} 1 \cdot x_1 + 1 \cdot x_2 + 1 \cdot x_3 + 1 \cdot x_4 = 4 \\ 0 \cdot x_1 + 1 \cdot x_2 + 1 \cdot x_3 + 1 \cdot x_4 = 0 \\ 1 \cdot x_1 + 0 \cdot x_2 + 1 \cdot x_3 + 1 \cdot x_4 = 0 \end{cases}$$

כלומר מתקיים

$$\begin{cases} a_{11} = a_{12} = a_{13} = a_{14} = 1 \\ a_{21} = 0, a_{22} = a_{23} = a_{24} = 1 \\ a_{32} = 0, a_{31} = a_{33} = a_{34} = 1 \end{cases}$$

הקבועים הם $b_1 = 4, b_2 = b_3 = 0$.

הרבה יותר קשה לדמיין את הגיאומטריה כאשר יש מספר משתנים גבוה מ-3, או אפילו משוואה אחת בשני משתנים מרוכבים. אבל מבחינה אלגברית, כל המקרים האלה שייכים לקטגוריה אחת ושיטת הפתרון משותפת לכולם. באופן כללי, נשאל את השאלות הבאות על ממ"ל:

1. כמה פתרונות יש?
2. כיצד נמצא את כל הפתרונות?
3. מהו מבנה קבוצת הפתרונות?

לפני כן, נראה עוד דוגמאות לממ"ל וגם נחזור אל חלק מהדוגמאות הקודמות. קל לבדוק אם פתרון נתון הוא אכן פתרון ע"י בדיקה שהצבתו מקיימת את כל המשוואות. עבור ממ"ל פשוטה גם לא דרוש הרבה מאמץ כדי למצוא את קבוצת הפתרונות.

דוגמה 5.12.

א.

$$\begin{cases} x & & & = 1 \\ & y & & = 2 \\ & & z & = 3 \end{cases}$$

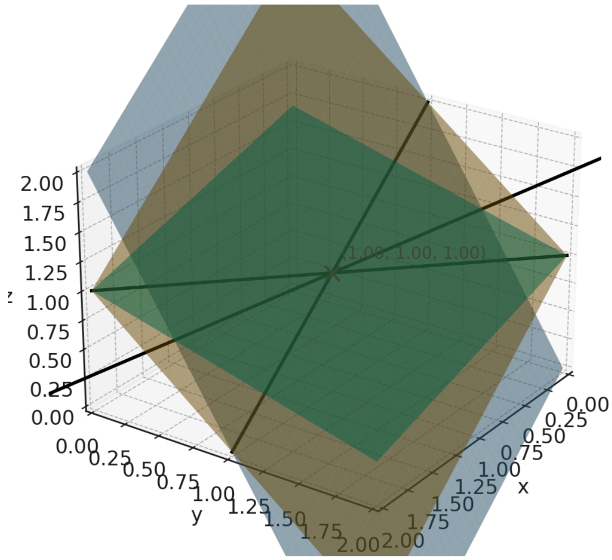
זו ממ"ל פשוטה שיש לה פתרון יחיד $(x,y,z) = (1,2,3)$, שכבר מופיע באופן מפורש.

ב. נחזור לדוגמה קודמת:

$$\begin{cases} x + y + z = 3 \\ x + 2y + z = 4 \\ x + y + 3z = 5 \end{cases}$$

ניתן לבדוק ע"י הצבה כי $(x,y,z) = (1,1,1)$ הוא פתרון של הממ"ל, כי כל שלוש המשוואות מתקיימות, לעומת זאת $(x,y,z) = (2,1,0)$ אינו פתרון כי הוא מקיים רק את שתי המשוואות הראשונות.

האם $(x,y,z) = (1,1,1)$ הוא פתרון יחיד? נראה בהמשך שהתשובה היא כן, אבל בינתיים נשים לב שהממ"ל מתארת חיתוך של שלושה מישורים במרחב. לפי הפרק הקודם, החיתוך בין כל שניים מהמישורים שנתונים בממ"ל - הוא ישר. אז החיתוך בין שלושת המישורים הוא בעצם חיתוך בין שני ישרים. למשל: ישר החיתוך של שני המישורים הראשונים, וישר החיתוך של שני המישורים האחרונים. אפשר לקבוע את המצב ההדדי בין שני הישרים בעזרת ההצגות הפרמטריות שלהם, אבל נראה שיש דרך יותר מסודרת ויעילה לחישוב קבוצת הפתרונות בדוגמה זו ובאופן כללי.



איור **5.1**: שלושת המישורים, ישרי החיתוך ונקודת החיתוך

ג.

$$\begin{cases} x \qquad\qquad + w = 1 \\ \qquad y \qquad + w = 2 \\ \qquad\qquad z + w = 3 \end{cases}$$

זו ממ"ל יחסית פשוטה בארבעה משתנים x, y, z, w. אפשר להתייחס ל-w כמשתנה חופשי ואז כל שאר המשתנים תלויים בו. נשתמש בפרמטר t, נציב $w = t$ במשוואות ואז נחסיר t מכל אגף בכל משוואה. כך נקבל

$$\begin{cases} x = 1 - t \\ y = 2 - t \\ z = 3 - t \end{cases}$$

לכן קבוצת הפתרונות של הממ"ל היא

$$\{\, (1-t, 2-t, 3-t, t) \mid t \in \mathbb{R} \,\} = \{\, (1,2,3,0) + t(-1,-1,-1,1) \mid t \in \mathbb{R} \,\}.$$

אפשר לנסות לדמיין קבוצה זו כישר במרחב $\mathbb{R}^4$ הארבע-מימדי. זה בהחלט לא משהו שרואים ביום-יום כי העולם המוכר הוא תלת-מימדי, אבל זו עדיין הצגה אלגברית של ישר חד-מימדי. זה נובע מכך שיש משתנה חופשי יחיד, ומכאן הפרמטר t. אם לצורך העניין

היה פרמטר נוסף בהצגה הפרמטרית (בקבוצת הפתרונות של ממ"ל אחרת), הקבוצה הייתה "נראית" כמו מישור ב- $\mathbb{R}^4$ ככל שאפשר לדמיין אותו. נראה מקרה כזה בדוגמה הבאה.

ד.

$$\begin{cases} x \qquad + z + w = 1 \\ \qquad y + z + w = 2 \end{cases}$$

כאן המשתנים החופשיים הם z, w והמשתנים התלויים הם x, y. נציב ערכים שרירותיים $z = t, w = s$ ונקבל אחרי העברת אגפים

$$\begin{cases} x = 1 - t - s \\ y = 2 - t - s \end{cases}$$

כלומר קיבלנו את ההצגה הפרמטרית הבאה לקבוצת הפתרונות:

$$\{\ (1 - s - t, 2 - s - t, s, t) \mid s, t \in \mathbb{R}\ \}$$

לפעמים נרצה לכתוב אותה בצורה קצת שונה. נכתוב את הוקטור של ההצגה הפרמטרית כסכום וקטורים: הוקטור של הקבועים $(1, 2, 0, 0)$, הוקטור של כפולות של s $(-s, -s, s, 0)$ והוקטור של כפולות של t $(-t, -t, 0, t)$. במילים אחרות, נעשה חישוב "הפוך" וזאת כדי לראות את הקשר להצגה פרמטרית של מישור. בנוסף, כל פרמטר הוא סקלר ולכן ניתן להוציא אותו מחוץ לוקטור המתאים לו. כך נקבל את ההצגה הבאה:

$$\{\ (1, 2, 0, 0) + s(-1, -1, 1, 0) + t(-1, -1, 0, 1) \mid s, t \in \mathbb{R}\ \}$$

זו כתיבה שמזכירה הצגה פרמטרית של מישור במרחב, אלא שכאן המרחב הוא $\mathbb{R}^4$. בהמשך הקורס נלמד על המושגים של מימד ותת-מרחב (כמו ישר במישור, ומישור במרחב) באופן מדויק.

ה. עוד דוגמה קודמת:

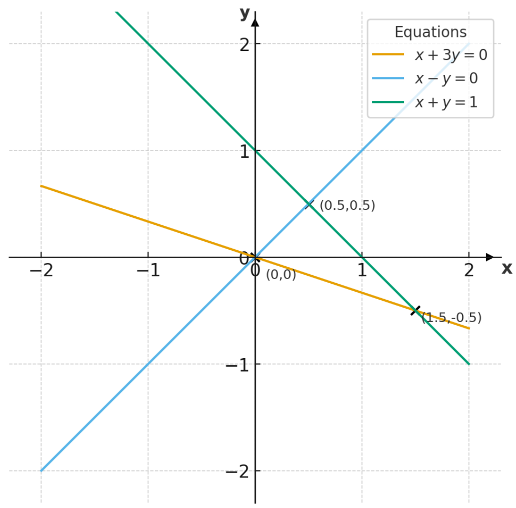
$$\begin{cases} iz_1 + z_2 \qquad\quad + iz_3 + z_4 \ \ = 2i \\ z_1 \ + (1+i)z_2 + z_3 \ + 2iz_4 = 2 \end{cases}$$

ניתן לבדוק ע"י הצבה כי $(z_1, z_2, z_3, z_4) = (1, 0, 1, 0)$ הוא פתרון של הממ"ל. אינטואיטיבית, אמורים להיות אינסוף פתרונות כי מספר המשתנים גדול ממספר המשוואות. נצדיק את האינטואיציה הזו בהמשך.

ו. ממ"ל של שלוש משוואות בשני משתנים מעל $\mathbb{R}$:

$$\begin{cases} x+3y=0 \\ x-y=0 \\ x+y=1 \end{cases}$$

כאן מספר המשתנים גדול ממספר המשוואות, ולמעשה אין פתרון. ניתן לראות זאת ע"י חיבור שתי המשוואות הראשונות, מה שנותן $2x+2y=0$ או באופן שקול $x+y=0$. זו סתירה למשוואה האחרונה. המשמעות הגיאומטרית היא שאמנם כל שניים משלושת הישרים נחתכים, אבל נקודות החיתוך שונות ולכן אין נקודת חיתוך משותפת לכל השלושה.



איור 5.2: שלושת הישרים ונקודות החיתוך

תרגיל 5.13. נסתכל על הממ"ל

$$\begin{cases} x_1+x_2+2x_3-x_4+x_5=2 \\ 5x_1-3x_2-2x_3+x_4 \qquad =0 \end{cases}.$$

האם $(x_1,x_2,x_3,x_4,x_5)=(1,1,1,2,0)$ הוא פתרון? אם לא, לאיזה ערך של $a\in\mathbb{R}$ הוקטור $(x_1,x_2,x_3,x_4,x_5)=(1-a,1,1,2,a)$ הוא פתרון של הממ"ל?

פתרון. נציב בממ"ל $(x_1,x_2,x_3,x_4,x_5)=(1,1,1,2,0)$ ונקבל

$$\begin{cases} 1+1+2-2+0=2 \\ 5-3-2+2=2\neq 0 \end{cases}.$$

אז זה לא פתרון לממ"ל (אלא רק למשוואה הראשונה). נחפש פתרון מהצורה

$$(x_1, x_2, x_3, x_4, x_5) = (1-a, 1, 1, 2, a)$$

ונקבל לאחר הצבה

$$\begin{cases} 1 - a + 1 + 2 - 2 + a = 2 \\ 5(1-a) - 3 - 2 + 2 = 0 \end{cases}.$$

המשוואה הראשונה מתקיימת לכל ערך של a, ואילו המשוואה השנייה מתקיימת אם ורק אם

$$.2 - 5a = 0 \iff a = \frac{2}{5}$$

אז הפתרון מהצורה המבוקשת (יש עוד אינסוף פתרונות מסוג אחר) הוא

$$.(x_1, x_2, x_3, x_4, x_5) = (\frac{3}{5}, 1, 1, 2, \frac{2}{5})$$

5.3 מטריצות

ראינו בפרק הקודם, למשל עבור ממ"ל של שתי משוואות בשני נעלמים, שניתן לחבר בין שתי משוואות כדי לנסות לקבל משוואה יותר פשוטה שנובעת מהממ"ל. למעשה, פעולה זו משפיעה רק על המקדמים והקבועים של הממ"ל. כאשר מחברים את שתי המשוואות הראשונות של 5.2 מקבלים

$$(a_{11} + a_{21})x_1 + (a_{12} + a_{22})x_2 + (a_{13} + a_{23})x_3 + ... + (a_{1n} + a_{2n})x_n = b_1 + b_2$$

לצורך העניין, אפשר לשים את המשתנים בצד ולחזור אליהם בסוף. המידע של הממ"ל אמנם קשור למשתנים, אבל הוא מגולם בסקלרים (מקדמים וקבועים). אז זה שימושי לייצג את הממ"ל כטבלה שבה מופיעים רק הסקלרים. טבלה כזו נקראת מטריצה.

הגדרה 5.14. מטריצה מסדר $m \times n$ מעל $\mathbb{F}$ היא טבלה בת m שורות ו-n עמודות כך שבכל משבצת מופיע סקלר ב-$\mathbb{F}$. בהקשר של הממ"ל

$$\begin{cases} a_{11}x_1 + \ a_{12}x_2 + \ a_{13}x_3 + \ \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \ a_{22}x_2 + \ a_{23}x_3 + \ \cdots + a_{2n}x_n = b_2 \\ a_{31}x_1 + \ a_{32}x_2 + \ a_{33}x_3 + \ \cdots + a_{3n}x_n = b_3 \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n = b_m \end{cases},$$

נגדיר את מטריצת המקדמים (שנקראת גם מטריצת מקדמים מצומצמת):

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

נגדיר גם את מטריצת המקדמים המורחבת מסדר $m\times(n+1)$ שכוללת עמודה בצד ימין (אחרי קו מפריד) בשביל הקבועים:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right)$$

נגדיר בנוסף את וקטור המשתנים $x=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ וגם את וקטור הקבועים $b=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$. דרך מקוצרת לכתיבת הממ"ל 5.2 היא המשוואה הוקטורית $Ax=b$.

בפרק הבא נבין יותר לעומק את הסימון Ax. בינתיים זה סימון מקוצר, וגם x הוא סימון נוח עבור וקטור (אין צורך בחץ מעליו).

דוגמה 5.15.

א. עבור הממ"ל

$$\begin{cases} x+y+z=3 \\ x+2y+z=4 \\ x+y+3z=5 \end{cases}$$

מטריצת המקדמים המצומצמת היא

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 3 \end{pmatrix}.$$

בשביל מטריצת המקדמים המורחבת נוסיף את עמודת הקבועים ונקבל

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 1 & 2 & 1 & 4 \\ 1 & 1 & 3 & 5 \end{array}\right).$$

ב. נסתכל על הממ"ל

$$\begin{cases} 2x + \ \ y + z \qquad = 1 \\ x + 2y + z + w = -1 \\ x + \ \ y \qquad + w = 2 \end{cases}.$$

נשים לב שמסתתרים שני אפסים: המקדם של w במשוואה הראשונה, והמקדם של z במשוואה השלישית. הנה כתיבה יותר מפורשת:

$$\begin{cases} 2x + \ \ y + \ \ z + 0\cdot w = 1 \\ x + 2y + \ \ z + \ \ w = -1 \\ x + \ \ y + 0\cdot z + \ \ w = 2 \end{cases}.$$

לכן מטריצת המקדמים המצומצמת היא

$$\begin{pmatrix} 2 & 1 & 1 & 0 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}.$$

מטריצת המקדמים המורחבת היא

$$\left(\begin{array}{cccc|c} 2 & 1 & 1 & 0 & 1 \\ 1 & 2 & 1 & 1 & -1 \\ 1 & 1 & 0 & 1 & 2 \end{array}\right).$$

הערה 5.16. שימו לב שאם המשתנה ה- j לא מופיע במשוואה ה- i, זה אומר שהמקדם המתאים הוא 0 ולכן $a_{ij} = 0$ במטריצה.

תרגיל 5.17. כתבו את מטריצות המקדמים (מצומצמת ומורחבת) ועמודת הקבועים של הממ"ל

$$\begin{cases} x + \ \ y + \ \ z + \ \ w = 0 \\ \qquad\qquad z - \ \ w = 2 \\ x + 2y + 2z + 3w = 8 \end{cases}.$$

פתרון. מטריצת המקדמים המצומצמת היא

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1 \\ 1 & 2 & 2 & 3 \end{pmatrix}.$$

מטריצת המקדמים המורחבת היא

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 2 \\ 1 & 2 & 2 & 3 & 8 \end{array}\right).$$

וקטור הקבועים הוא וקטור העמודה הימנית של המטריצה האחרונה, כלומר $\begin{pmatrix} 0 \\ 2 \\ 8 \end{pmatrix}$.

5.4 מטריצות מדורגות ומדורגות קנונית

כבר נתקלנו בדוגמאות בהן הממ"ל יחסית פשוטה לפתרון, במובן שקל לזהות את המשתנים החופשיים, להציב בהם ערכים שרירותיים ולקבל את הערכים המתאימים של המשתנים התלויים ע"י קצת אלגברה (פעולות חשבון ואולי הצבות). איך מטריצת המקדמים (מצומצמת או מורחבת) נראית במקרים כאלה? לשם כך נועדה ההגדרה הבאה.

הגדרה 5.18. מטריצה נקראת <u>מדורגת</u> אם:

א. אם במטריצה יש שורות אפסים, אז הן מופיעות מתחת לכל שאר השורות.

ב. בכל שורה שאינה שורת אפסים, האיבר הראשון (משמאל) שאינו 0 נקרא <u>האיבר המוביל</u> של שורה זו, והוא מופיע משמאל לאיבר המוביל של השורה מתחתיו (אם קיימת כזו). בפרט, מתחת לכל איבר מוביל יש אפסים בלבד.

דוגמה 5.19.

א. במטריצה

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

האיברים המובילים הם $1,1,4$ לפי סדר השורות והם מופיעים בסדר הנכון (זזים ימינה כאשר יורדים שורה). בנוסף, אין בכלל שורות אפסים ולכן זו מטריצה מדורגת.

ב. המטריצה

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 5 & 4 \end{pmatrix}$$

אינה מדורגת כי האיבר המוביל של השורה השלישית (5) מופיע באותה העמודה של האיבר המוביל של השורה השנייה (1).

ג. המטריצה

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}$$

מדורגת. שורה האפסים מופיעה בסוף, והאיבר המוביל של השורה הראשונה (1) מופיע משמאל לזה של השורה השנייה (4).

ד. המטריצה

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

אינה מדורגת כי יש שורת אפסים מעל לשורה שאינה כולה אפסים.

מטריצות מדורגות הן יחסית נוחות, אך יש מטריצות מדורגות מסוג יותר מסוים שהוא עוד יותר נוח בפתרון ממ"ל.

הגדרה 5.20. מטריצה נקראת מדורגת קנונית אם:

א. היא מדורגת, כלומר מקיימת את שני תנאי ההגדרה הקודמת.

ב. האיבר המוביל בכל שורה (אם הוא קיים) הוא 1. אפשר גם לקרוא לו אחד מוביל.

ג. בכל עמודה של אחד מוביל, כל שאר האיברים הם 0 (מתחת וגם מעל לאחד המוביל).

דוגמה 5.21.

א. ראינו כבר את המטריצה

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

שנקראת מטריצת היחידה מסדר 3×3. היא מדורגת קנונית כי אין בה שורת אפסים, כל האיברים המובילים הם 1 ומופיעים בסדר הנכון, וכל שאר האיברים הם 0.

ב. המטריצה

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

מדורגת אך לא מדורגת קנונית, כי האיבר המוביל התחתון אינו 1. בנוסף, לא כל האיברים מעליו הם 0.

ג. המטריצה

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

מדורגת קנונית כי שני האיברים המובילים הם 1 ומופיעים בסדר הנכון. חוץ מהם יש רק אפסים בעמודות שלהם, וגם שורת האפסים מופיעה בסוף.

מהי הצורה הכללית של מטריצה מדורגת קנונית מסדר $m \times n$?

$$\begin{array}{r} 1 \rightarrow \\ 2 \rightarrow \\ \\ r \rightarrow \\ r+1 \rightarrow \\ \\ m \rightarrow \end{array}
\begin{bmatrix}
0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * & 0 & * & \cdots & * \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0
\end{bmatrix}$$

$$\begin{array}{ccccc} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ 1 & k_1 & k_2 & k_r & n \end{array}$$

איור 5.3: הצורה הכללית של מטריצה מדורגת קנונית

האיברים המובילים מופיעים בעמודות $k_1, k_2, \ldots, k_r$. r הוא גם מספר השורות שאינן כולן אפסים (בכל שורה כזו יש איבר מוביל אחד).

הגדרה 5.22. נניח שקיבלנו צורה מדורגת קנונית של מטריצת מקדמים מצומצמת A של ממ"ל $Ax = b$.

א. כל משתנה שמתאים לעמודה בה מופיע איבר מוביל נקרא משתנה תלוי. שאר המשתנים נקראים חופשיים.

ב. למספר המשתנים התלויים נקרא הדרגה של הממ"ל/מטריצה. נסמן $\mathrm{rank}(A)$.

הערה 5.23. נראה בקרוב (שיטת הדירוג של גאוס) שניתן לבצע על כל מטריצה סדרת פעולות (שנקראת דירוג) בין שורותיה, כך שבסוף מתקבלת מטריצה מדורגת קנונית. מטריצה זו היא יחידה (כלומר לא תלויה בדרך הדירוג שניתנת לבחירה) וניתן להשתמש בה כדי לקבוע את קבוצת הפתרונות של הממ"ל המקורית. הדרגה, שהיא כאמור מספר המשתנים התלויים, שווה למספר השורות בצורה המדורגת קנונית שאינן שורת אפסים. זהו גם מספר האיברים המובילים.

הרעיון של משתנה תלוי זה שיש משוואה שמתארת את התלות שלו במשתנים החופשיים (או חלקם). ניתן להציב סקלרים שרירותיים במשתנים החופשיים בכל משוואה, ולקבל את הערכים המתאימים של המשתנים התלויים.

דוגמה 5.24. נפתור את הממ"ל

$$\begin{cases} x_1 + ix_4 = 1 + i \\ x_3 + x_4 = 2 \end{cases}.$$

מטריצת המקדמים המורחבת היא

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & i & 1+i \\ 0 & 0 & 1 & 1 & 2 \end{array}\right)$$

היא כבר מדורגת קנונית ויש בה שני אחדות מובילים. לכן הדרגה היא 2 והמשתנים התלויים הם x_1, x_3. נציב $(x_2, x_4) = (s, t) \in \mathbb{C}^2$ בשתי המשוואות ונבודד את המשתנים התלויים:

$$\begin{cases} x_1 = 1 + i - it \\ x_3 = 2 - t \end{cases}.$$

לכן, קבוצת הפתרונות נתונה ע"י

$$\{ (1 + i - it, s, 2 - t, t) \mid s, t \in \mathbb{C} \}.$$

יש שני פרמטרים בהצגה הפרמטרית כי יש שני משתנים חופשיים.

5.5 פעולות שורה אלמנטריות

נרצה לפשט ממ"ל כללית, כלומר להעביר את מטריצת המקדמים לצורה מדורגת קנונית בלי לשנות את קבוצת הפתרונות של הממ"ל. זה לא רעיון חדש: כבר במקרה של שתי משוואות בשני משתנים ראינו שאפשר לחבר/לחסר בין המשוואות וגם להשתמש בכפל בסקלר. כל פעולה כזו לא משנה את קבוצת הפתרונות, כי מדובר בפעולה הפיכה (אפשר לשחזר את הממ"ל המקורית מתוך צורתה החדשה). במילים אחרות, פעולות כאלו יוצרות ממ"ל שקולה לממ"ל המקורית, ויש בזה טעם אם היא יותר פשוטה לפתרון כמו במקרה של ממ"ל המתאימה למטריצה מדורגת קנונית.

הגדרה 5.25. תהי

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

מטריצה מסדר $m \times n$ מעל $\mathbb{F}$. נסמן ב-$R_1, R_2, ..., R_m$ את שורות המטריצה A, כלומר $R_i = (a_{i1}, a_{i2}, ..., a_{in})$ לכל $1 \leq i \leq m$. להלן שלושה סוגי פעולות על שורות A הנקראות <u>פעולות שורה אלמנטריות</u>:

א. כפל השורה R_i בסקלר $c \in F$ שאינו 0. נסמן פעולה זו כך: $cR_i \to R_i$.

ב. הוספה לשורה R_i את הכפולה של השורה R_j בסקלר $c \in \mathbb{F}$ כאשר $i \neq j$. נסמן $R_i + cR_j \to R_i$. השורה R_j נשארת ללא שינוי.

ג. החלפת השורות R_i ו-R_j עבור $i \neq j$. נסמן $R_i \leftrightarrow R_j$.

דוגמה 5.26. נדגים כל סוג פעולה על המטריצה

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$$

א. נכפיל את השורה הראשונה ב-2:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \underset{R_1 \leftarrow \frac{1}{2}R_1}{\overset{2R_1 \to R_1}{\longleftrightarrow}} \begin{pmatrix} 2 & 4 & 6 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

את הפעולה משמאל לימין ציינו מעל לחץ, ומתחתיו ציינו את הפעולה ההפוכה מימין לשמאל. פעולה זו היא כפל השורה הראשונה ב-$\frac{1}{2}$.

ב. נוסיף לשורה השלישית כפולה של השורה הראשונה בסקלר -3:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \underset{R_3 \leftarrow R_3 + 3R_1}{\overset{R_3 - 3R_1 \to R_3}{\longleftrightarrow}} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 4 & 2 & 0 \end{pmatrix}$$

שימו לב שהשורה הראשונה לא משתנה. החישוב שעשינו עבור השורה השלישית הוא $(7,8,9) - 3(1,2,3) = (4,2,0)$. הפעולה ההפוכה היא הוספה לשורה השלישית כפולה של השורה הראשונה בסקלר 3, כלומר הסקלר הנגדי לסקלר המקורי (שינוי סימן).

ג. נחליף בין השורה השנייה לשורה השלישית:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \underset{R_3 \leftrightarrow R_2}{\overset{R_2 \leftrightarrow R_3}{\longleftrightarrow}} \begin{pmatrix} 1 & 2 & 3 \\ 7 & 8 & 9 \\ 4 & 5 & 6 \end{pmatrix}$$

כאן הפעולה ההפוכה היא אותה הפעולה: החלפה בין השורה השנייה לשורה השלישית.

הערה 5.27. באופן כללי, כל פעולות השורה האלמנטריות הן הפיכות. נבהיר למה הכוונה בכל מקרה:

א. לכל $c \neq 0$ ולכל $1 \leq i \leq m$ הפעולה ההפוכה ל-$cR_i \to R_i$ היא $\frac{1}{c}R_i \to R_i$.

ב. לכל $c \in F$ ולכל $i \neq j$ הפעולה ההפוכה ל-$R_i + cR_j \to R_i$ היא $R_i - cR_j \to R_i$. אכן, אם מבצעים את הפעולה הראשונה ואחריה את השנייה (או בסדר הפוך), התוצאה היא $R_i + cR_j - cR_j \to R_i$ או בקיצור $R_i \to R_i$. החזרה למקור היא המשמעות של פעולה הפוכה.

ג. לכל $i \neq j$ הפעולה ההפוכה להחלפה $R_i \leftrightarrow R_j$ היא אותה ההחלפה $R_i \leftrightarrow R_j$.

הגדרה 5.28. מטריצות A, A' נקראות <u>שקולות שורה</u> אם ניתן להגיע מ- A ל- A' ע״י סדרה סופית של פעולות שורה אלמנטריות.

הערה 5.29. זו הגדרה סימטרית. אם ניתן להגיע מ- A ל- A' ע״י סדרה סופית של פעולות שורה אלמנטריות, אז גם ניתן להגיע מ- A' ל- A ע״י סדרת הפעולות ההפוכות.

טענה 5.30. *נסתכל על הממ״ל $Ax = b$ עם מטריצת מקדמים מורחבת $(A|b)$, וגם על הממ״ל $A'x = b'$ עם מטריצת מקדמים מורחבת $(A'|b')$. אם $(A|b), (A'|b')$ הן מטריצות שקולות שורה, אז קבוצת הפתרונות של הממ״ל $Ax = b$ שווה לקבוצת הפתרונות של הממ״ל $A'x = b'$.*

במילים אחרות, כל n-יה $(c_1, c_2, ..., c_n)$ היא פתרון של הממ״ל החדשה (שמתקבלת אחרי סדרת פעולות שורה אלמנטריות) אם ורק אם היא פתרון של הממ״ל המקורית.

הוכחה. לכל ממ״ל $Ax = b$ ולכל פעולת שורה אלמנטרית, כל פתרון של הממ״ל לפני הפעולה נשאר פתרון של הממ״ל שמתקבלת בעקבות הפעולה. ולהיפך, ההפיכות של הפעולה מראה שכל פתרון של הממ״ל החדשה הוא גם פתרון של הממ״ל הקודמת לה. אז קבוצת הפתרונות נשארת ללא שינוי.

נמשיך כך ונבצע מספר סופי של פעולות שורה אלמנטריות כדי לקבל את הממ״ל $A'x = b'$. קבוצת הפתרונות תישמר בכל שלב ולכן גם בסוף התהליך היא תישאר זהה לקבוצת הפתרונות המקורית. □

הערה 5.31. כפל שורה ב-0 אינה פעולה הפיכה. למעשה, הוא הופך את המשוואה המתאימה בממ״ל למשוואה $0 = 0$ שמתקיימת ללא קשר למשתנים. בדרך כלל זה מוביל לאובדן מידע על המשתנים, מה שישפיע על קבוצת הפתרונות (יגדיל אותה) ולכן יש להימנע מכפל שורה ב-0. הנה דוגמה פשוטה במקרה של שתי משוואות בשני משתנים:

$$\begin{cases} x = 1 \\ y = 0 \end{cases}$$

לממ"ל יש פתרון יחיד $(x,y)=(1,0)$ ושני המשתנים תלויים. נכפיל את השורה השנייה ב-0 במטריצת המקדמים המורחבת:

$$\left(\begin{array}{cc|c}1 & 0 & 1\\ 0 & 1 & 0\end{array}\right) \xrightarrow{0\cdot R_2\to R_2} \left(\begin{array}{cc|c}1 & 0 & 1\\ 0 & 0 & 0\end{array}\right)$$

הממ"ל החדשה היא

$$\begin{cases}x=1\\ 0=0\end{cases}.$$

עכשיו y משתנה חופשי. נציב $y=t$ ונקבל קבוצת פתרונות $\{\,(1,t)\mid t\in\mathbb{R}\,\}$, כלומר התווספו אינסוף פתרונות. הממ"ל החדשה לא שקולה לממ"ל המקורית כי הכפלנו ב-0 ובכך מחקנו משוואה.

הגדרה 5.32. כאשר וקטור הקבועים של הממ"ל $Ax=b$ הוא $b=\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}$, הממ"ל נקראת <u>הומוגנית</u> וניתן לכתוב $Ax=0$. הכוונה באגף ימין היא לוקטור האפס באורך m (מספר השורות של A).

הערה 5.33. במקרה של ממ"ל הומוגנית, וקטור הקבועים מיותר מבחינת פעולות השורה האלמנטריות. הסיבה לכך היא שכל פעולת חיבור בין אפסים תיתן **0**, וכנ"ל לכפל בסקלר (כי $a\cdot 0=0$ לכל סקלר a ממשי או מרוכב). לכן זה מוצדק להסתפק במטריצת המקדמים המצומצמת במקרה זה (ורק במקרה זה), וכמובן הממ"ל החדשה גם תהיה הומוגנית.

דוגמה 5.34. נסתכל על ממ"ל הומוגנית:

$$\begin{cases}x+\ \ y+\ \ z=0\\ 2x\qquad\ \ +2z=0\\ 3x+5y+7z=0\end{cases}$$

נבצע סדרה של פעולות על מטריצת המקדמים המצומצמת:

$$\begin{pmatrix}1&1&1\\2&0&2\\3&5&7\end{pmatrix}\xrightarrow{\frac{1}{2}R_2\to R_2}\begin{pmatrix}1&1&1\\1&0&1\\3&5&7\end{pmatrix}\xrightarrow{R_1-R_2\to R_1}\begin{pmatrix}0&1&0\\1&0&1\\3&5&7\end{pmatrix}$$
$$\xrightarrow{R_1\leftrightarrow R_2}\begin{pmatrix}1&0&1\\0&1&0\\3&5&7\end{pmatrix}\xrightarrow{R_3-3R_1\to R_3}\begin{pmatrix}1&0&1\\0&1&0\\0&5&4\end{pmatrix}\xrightarrow{R_3-5R_2\to R_3}\begin{pmatrix}1&0&1\\0&1&0\\0&0&4\end{pmatrix}$$
$$\xrightarrow{\frac{1}{4}R_3\to R_3}\begin{pmatrix}1&0&1\\0&1&0\\0&0&1\end{pmatrix}\xrightarrow{R_1-R_3\to R_1}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$$

נסביר בהמשך איך בחרנו את הפעולות האלו. בינתיים נציין שסדרה כזו של פעולות נקראת דירוג מטריצה, וכאן היא מראה כי הממ"ל המקורית שקולה לממ"ל ההומוגנית שמתאימה למטריצה האחרונה שקיבלנו:

$$\begin{cases}x=0\\y=0\\z=0\end{cases}$$

לכן גם לממ"ל המקורית יש פתרון יחיד, שהוא הפתרון הטריוויאלי שבו כל המשתנים מתאפסים. שימו לב לצורתה של המטריצה האחרונה שהתקבלה בתהליך הדירוג. זו צורה נוחה מאוד בגלל ריבוי האפסים, ובנוסף כל שאר האיברים הם 1. לכן, בכל משוואה בממ"ל האחרונה יש משתנה תלוי שהוא כבר מבודד (ובמקרה זה אין אף משתנה חופשי).

באופן כללי, צריך להתייחס לשתי שאלות כאשר מנסים לפשט ממ"ל:

1. איזו פעולת שורה אלמנטרית בוחרים בכל שלב?
2. מתי עוצרים את הדירוג? כלומר, מתי המטריצה פשוטה מספיק כדי לפתור את הממ"ל?

5.6 שיטת הדירוג של גאוס

כיצד נבחר פעולות שורה אלמנטריות לבצע כדי להעביר מטריצה נתונה

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

לצורה מדורגת קנונית? יש שיטה מסודרת (אלגוריתם) שנקראת שיטת הדירוג של גאוס. בהינתן מטריצה, נפעל לפי השלבים הבאים:

א. נחפש בעמודה הראשונה איבר $a_{i1} \neq 0$ ונהפוך אותו ל-1 ע"י הפעולה $\frac{1}{a_{i1}}R_i \to R_i$ (אם יש רק אפסים, אפשר לדלג לשלב 3). אם $i \neq 1$, נבצע גם את ההחלפה $R_i \leftrightarrow R_1$ כדי שיופיע 1 בפינה השמאלית העליונה.

ב. נשתמש ב-1 שקיבלנו בשורה הראשונה כדי לאפס את כל האיברים מתחתיו. כלומר: נבצע את הפעולה $R_j - a_{j1}R_1 \to R_j$ לכל $j > 1$ כך ש-$a_{j1} \neq 0$. כך נקבל עמודה שכמעט כולה אפסים חוץ מהאחד המוביל בשורה הראשונה.

ג. ממשיכים באופן דומה עם העמודה השנייה וחוזרים על אותם השלבים כדי לאפס את כל האיברים חוץ מאיבר אחד שיופיע בשורה השנייה כ-1 (אם מלכתחילה יש רק אפסים מתחת לשורה הראשונה, ממשיכים הלאה לעמודה הבאה). יש להימנע מפעולה שתשנה את העמודה הראשונה.

ד. ממשיכים באופן דומה עם שאר העמודות לפי הסדר, בלי לשנות את העמודות הקודמות.

ה. עוצרים אחרי שעברנו על כל העמודות, או כאשר כל השורות מתחת לאחד המוביל האחרון הן שורות אפסים (ואז אין מה לעשות איתן).

דוגמה 5.35. נחזור לדוגמה 5.34, אבל הפעם נעבוד קצת שונה וגם נסביר את בחירת הפעולות. ניתן לבצע מספר פעולות במקביל אם הן בלתי תלויות, כמו להחסיר כפולות של שורה כלשהי משורות אחרות (לכל שורה כפולה משלה).

$$\begin{pmatrix}1&1&1\\2&0&2\\3&5&7\end{pmatrix} \xrightarrow[R_2-2R_1\to R_2]{R_3-3R_1\to R_3} \begin{pmatrix}1&1&1\\0&-2&0\\0&2&4\end{pmatrix} \xrightarrow{-\frac{1}{2}R_2\to R_2} \begin{pmatrix}1&1&1\\0&1&0\\0&2&4\end{pmatrix}$$

$$\xrightarrow[R_1-R_2\to R_1]{R_3-2R_2\to R_3} \begin{pmatrix}1&0&1\\0&1&0\\0&0&4\end{pmatrix} \xrightarrow{\frac{1}{4}R_3\to R_3} \begin{pmatrix}1&0&1\\0&1&0\\0&0&1\end{pmatrix}$$

$$\xrightarrow{R_1-R_3\to R_1} \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$$

בשלב הראשון כבר היה 1 בשורה הראשונה. השתמשנו בו כדי לאפס את האיברים מתחתיו. בשלב השני הפכנו את -2 ל-1. בשלב השלישי איפסנו את שאר האיברים בעמודה השנייה. בשלב הרביעי הפכנו את 4 ל-1. בשלב החמישי איפסנו את שאר האיברים בעמודה השלישית.

התוצאה היא מטריצה מדורגת קנונית (למעשה מטריצת היחידה) עם אחדות מובילים ואפסים בלבד.

הערה 5.36. כמו בדוגמה לעיל, לרוב יש יותר מדרך אחת לפעול לפי שיטת הדירוג של גאוס. זה גם לגיטימי לעשות פעולה מיותרת או לא אופטימלית (מבחינת מספר הפעולות) כל עוד היא חוקית. כל דרך היא נכונה ובלבד שאין בה כפל שורה ב-0 (או טעות חישוב) ומתקבלת בסוף מטריצה מדורגת קנונית. מטריצה זו היא יחידה ולכן אינה תלויה בדרך הדירוג.

דוגמה 5.37. נפתור את הממ"ל הבאה:

$$\begin{cases} 2x+y-\ \ z+\ \ w=8 \\ -3x-y+2z+3w=-11 \\ -2x+y+2z \qquad\ \ =-3 \end{cases}$$

נדרג את מטריצת המקדמים המורחבת:

$$\left(\begin{array}{cccc|c} 2 & 1 & -1 & 1 & 8 \\ -3 & -1 & 2 & 3 & -11 \\ -2 & 1 & 2 & 0 & -3 \end{array}\right) \xrightarrow[R_3+R_1\to R_3]{R_2+\frac{3}{2}R_1\to R_2} \left(\begin{array}{cccc|c} 2 & 1 & -1 & 1 & 8 \\ 0 & \frac{1}{2} & \frac{1}{2} & \frac{9}{2} & 1 \\ 0 & 2 & 1 & 1 & 5 \end{array}\right)$$

$$\xrightarrow{2R_2\to R_2} \left(\begin{array}{cccc|c} 2 & 1 & -1 & 1 & 8 \\ 0 & 1 & 1 & 9 & 2 \\ 0 & 2 & 1 & 1 & 5 \end{array}\right) \xrightarrow{R_3-2R_2\to R_3} \left(\begin{array}{cccc|c} 2 & 1 & -1 & 1 & 8 \\ 0 & 1 & 1 & 9 & 2 \\ 0 & 0 & -1 & -17 & 1 \end{array}\right)$$

$$\xrightarrow{-R_3\to R_3} \left(\begin{array}{cccc|c} 2 & 1 & -1 & 1 & 8 \\ 0 & 1 & 1 & 9 & 2 \\ 0 & 0 & 1 & 17 & -1 \end{array}\right) \xrightarrow[R_2-R_3\to R_2]{R_1+R_3\to R_1} \left(\begin{array}{cccc|c} 2 & 1 & 0 & 18 & 7 \\ 0 & 1 & 0 & -8 & 3 \\ 0 & 0 & 1 & 17 & -1 \end{array}\right)$$

$$\xrightarrow{\frac{1}{2}R_1\to R_1} \left(\begin{array}{cccc|c} 1 & \frac{1}{2} & 0 & 9 & \frac{7}{2} \\ 0 & 1 & 0 & -8 & 3 \\ 0 & 0 & 1 & 17 & -1 \end{array}\right) \xrightarrow{R_1-\frac{1}{2}R_2\to R_1} \left(\begin{array}{cccc|c} 1 & 0 & 0 & 13 & 2 \\ 0 & 1 & 0 & -8 & 3 \\ 0 & 0 & 1 & 17 & -1 \end{array}\right)$$

הממ"ל השקולה שמתאימה למטריצה המדורגת קנונית, נתונה ע"י:

$$\begin{cases} x+13w=2 \\ \ \ y-8w=3 \\ z+17w=-1 \end{cases}$$

המשתנה החופשי הוא w. נציב $w=t$ ונבודד את שאר המשתנים כדי לקבל את קבוצת הפתרונות:

$$\{\, (2-13t, 3+8t, -1-17t, t) \mid t \in \mathbb{R} \,\}$$

דוגמה 5.38. נפתור את הממ"ל הבאה מעל $\mathbb{C}$:

$$\begin{cases} x+y+z=i \\ x+iy+z=1+i \\ x+y+iz=2 \end{cases}$$

נדרג את מטריצת המקדמים המורחבת:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & i \\ 1 & i & 1 & 1+i \\ 1 & 1 & i & 2 \end{array}\right) \xrightarrow[R_3-R_1\to R_3]{R_2-R_1\to R_2} \left(\begin{array}{ccc|c} 1 & 1 & 1 & i \\ 0 & i-1 & 0 & 1 \\ 0 & 0 & i-1 & 2-i \end{array}\right)$$

$$\xrightarrow{\frac{1}{i-1}R_2\to R_2} \left(\begin{array}{ccc|c} 1 & 1 & 1 & i \\ 0 & 1 & 0 & \frac{1}{i-1} \\ 0 & 0 & i-1 & 2-i \end{array}\right) \xrightarrow{\frac{1}{i-1}R_3\to R_3} \left(\begin{array}{ccc|c} 1 & 1 & 1 & i \\ 0 & 1 & 0 & \frac{1}{i-1} \\ 0 & 0 & 1 & \frac{2-i}{i-1} \end{array}\right)$$

$$\xrightarrow{R_1-R_2-R_3\to R_1} \left(\begin{array}{ccc|c} 1 & 0 & 0 & i-\frac{3-i}{i-1} \\ 0 & 1 & 0 & \frac{1}{i-1} \\ 0 & 0 & 1 & \frac{2-i}{i-1} \end{array}\right)$$

זה מראה שהפתרון הוא יחיד (כי אין משתנה חופשי) ונתון ע"י:

$$\begin{cases} x = i-\frac{3-i}{i-1} = i-\frac{(3-i)(-i-1)}{2} = \frac{2i+(3-i)(1+i)}{2} = \frac{4+4i}{2} = 2+2i \\ y = \frac{1}{i-1} = -\frac{1+i}{2} \\ z = \frac{2-i}{i-1} = \frac{(2-i)(-i-1)}{2} = -\frac{3+i}{2} \end{cases}$$

כלומר $(x,y,z) = (2+2i, -\frac{1+i}{2}, -\frac{3+i}{2})$.

5.7 מספר הפתרונות

ראינו בדוגמאות השונות שמספר הפתרונות לממ"ל $Ax=b$ נקבע לפי שני קריטריונים:

1. האם יש שורת סתירה מהצורה $\left(\begin{array}{cccc|c} 0 & 0 & \cdots & 0 & 1 \end{array}\right)$ בצורה המדורגת קנונית? אם כן, אין פתרון לממ"ל. אם לא, נעבור לקריטריון הבא.

2. האם יש משתנה חופשי (אחד או יותר)? אם כן, יש אינסוף פתרונות לממ"ל. אם לא, יש פתרון יחיד.

ניתן לאפיין את הקריטריונים לעיל באמצעות הדרגה $\mathrm{rank}(A)$ של מטריצת המקדמים המצומצמת, והדרגה $\mathrm{rank}(A|b)$ של מטריצת המקדמים המורחבת. דירוג קנוני של A מתאים לדירוג קנוני של $(A|b)$, כאשר ההבדל היחיד הוא העמודה הימנית (צריך לבצע עליה את כל פעולות השורה שבחרנו עבור A).

ראשית נוכיח כמה תכונות של דרגה.

טענה 5.39. *א. לכל A מסדר $m \times n$ ולכל $b \in \mathbb{F}^n$ מתקיים $\mathrm{rank}(A) \leq \mathrm{rank}(A|b)$.*

ב. מתקיים $0 \leq \mathrm{rank}(A) \leq \min\{m,n\}$ כאשר $\min\{m,n\}$ הוא המספר הקטן מבין m,n.

הוכחה.

א. נניח כי הצורה המדורגת קנונית של $(A|b)$ היא $(A'|b')$. אז בפרט, A' היא הצורה המדורגת קנונית של A.

בהכרח מתקיים $\mathrm{rank}(A) \leq \mathrm{rank}(A|b)$ כי מספר האיברים המובילים ב- $(A'|b')$ הוא לפחות מספר האיברים המובילים ב- A'. או שהדרגות שוות, או שהדרגה של מטריצת המקדמים המורחבת גדולה יותר (ב-1) כתוצאה משורת סתירה שמוסיפה איבר מוביל.

ב. הדרגה היא מספר טבעי (היא 0 אם ורק אם כל איברי המטריצה הם 0). מאחר שזהו מספר האיברים המובילים בצורה המדורגת קנונית של A, מספר זה הוא לכל היותר m כי כל איבר מוביל מתאים לשורה אחת בלבד (ייתכן שיש גם שורות אפסים ואז הדרגה קטנה מ-m). באופן דומה, כל איבר מוביל לעמודה אחת בלבד (ושוב ייתכן שיש עמודות נוספות). אז הדרגה היא לכל היותר n, ובסך הכל מתקיים

$$0 \leq \mathrm{rank}(A) \leq \min\{m,n\}.$$

□

טענה 5.40. *לכל מטריצה A מסדר $m \times n$ ולכל $b \in \mathbb{R}^n$, מספר הפתרונות לממ"ל $Ax = b$ נקבע באופן הבא:*

א. אם $\mathrm{rank}(A) < \mathrm{rank}(A|b)$, אז אין לממ"ל פתרון.

ב. אם $\mathrm{rank}(A) = \mathrm{rank}(A|b) < n$, אז לממ"ל יש אינסוף פתרונות.

ג. אם $\mathrm{rank}(A) = \mathrm{rank}(A|b) = n$, אז לממ"ל יש פתרון יחיד.

הוכחה. בכל מקרה נתרגם את הנתון על הדרגות לשפה של שורת סתירה ומשתנים חופשיים.

א. במקרה זה יש שורת סתירה ולכן אין פתרון.

ב. כאן אין שורת סתירה ויש לפחות משתנה חופשי אחד, כי מספר המשתנים התלויים קטן ממספר המשתנים הכולל. אז יש אינסוף פתרונות.

ג. כאן אין שורת סתירה ואין משתנה חופשי, אז יש פתרון יחיד.

□

דוגמה 5.41. נסתכל על ממ״ל שבה חלק מהמקדמים תלויים בפרמטר a, כלומר קבוע (לא משתנה) שערכו אינו ידוע מראש. אפשר לחשוב על ממ״ל עם פרמטר כמשפחה של ממ״ליות - לכל הצבה של ערך הפרמטר נקבל ממ״ל מסוימת במשפחה.

$$\begin{cases} ax + \ \ y + \ \ z = 1 \\ \ x + ay + \ \ z = 1 \\ \ x + \ \ y + az = 1 \end{cases}$$

נמצא לאילו ערכי a מספר הפתרונות הוא אפס, אחד או אינסוף. נדרג את מטריצת המקדמים המורחבת, כאשר הפעולות עצמן יהיו תלויות ב- a. כל פעולה של הוספת כפולה של שורה לשורה אחרת היא חוקית, אבל צריך להיזהר עם פעולה של כפל בסקלר (או חילוק בסקלר). אסור להכפיל או לחלק 0, ולכן נצטרך לפצל למקרים לפחות בשלב אחד של הדירוג.

$$\left(\begin{array}{ccc|c} a & 1 & 1 & 1 \\ 1 & a & 1 & 1 \\ 1 & 1 & a & 1 \end{array}\right) \xrightarrow{R_2 \leftrightarrow R_1} \left(\begin{array}{ccc|c} 1 & a & 1 & 1 \\ a & 1 & 1 & 1 \\ 1 & 1 & a & 1 \end{array}\right)$$
$$\xrightarrow[R_3 - R_1 \to R_3]{R_2 - aR_1 \to R_2} \left(\begin{array}{ccc|c} 1 & a & 1 & 1 \\ 0 & 1-a^2 & 1-a & 1-a \\ 0 & 1-a & a-1 & 0 \end{array}\right)$$

נרצה לחלק את השורה האחרונה בסקלר $a - 1$, אבל זה כרוך בהנחה $a \neq 1$. במקרה המיוחד $a = 1$ שתי השורות האחרונות הן שורות אפסים, והשורה הראשונה מתאימה למשוואה $x + y + z = 1$. אז במקרה זה יש שני משתנים חופשיים (והדרגה היא 1), ולכן יש לממ״ל אינסוף פתרונות. קבוצת הפתרונות במקרה זה נתונה ע״י

$$\{\, (1-s-t, s, t) \mid s, t \in \mathbb{R} \,\}.$$

נניח כי $a \neq 1$ ונמשיך את הדירוג:

$$\begin{pmatrix}1 & a & 1 & | & 1\\ 0 & 1-a^2 & 1-a & | & 1-a\\ 0 & 1-a & a-1 & | & 0\end{pmatrix} \xrightarrow{\frac{1}{1-a}R_3\to R_3} \begin{pmatrix}1 & a & 1 & | & 1\\ 0 & 1-a^2 & 1-a & | & 1-a\\ 0 & 1 & -1 & | & 0\end{pmatrix}$$

$$\xrightarrow{R_2\leftrightarrow R_3} \begin{pmatrix}1 & a & 1 & | & 1\\ 0 & 1 & -1 & | & 0\\ 0 & 1-a^2 & 1-a & | & 1-a\end{pmatrix} \xrightarrow{R_3-(1-a^2)R_2\to R_3} \begin{pmatrix}1 & a & 1 & | & 1\\ 0 & 1 & -1 & | & 0\\ 0 & 0 & 2-a-a^2 & | & 1-a\end{pmatrix}$$

אם מתקיים $2-a-a^2\neq 0$, אז הדרגה של מטריצת המקדמים המצומצמת שווה ל- 3 וזו גם הדרגה של מטריצת המקדמים המורחבת. לכן יש פתרון יחיד לממ"ל המקורית.

לעומת זאת, שורשי המשוואה $2-a-a^2=0$ הם $a_1=1, a_2=-2$. את המקרה הראשון כבר ניתחנו וראינו שמתקבלים אינסוף פתרונות. במקרה השני נקבל שהשורה הראשונה היא $(0 \;\; 0 \;\; 0 \mid 3)$ ולכן אין פתרון.

נמשיך בדירוג קנוני כדי לחשב את הפתרון היחיד במקרה של $a\notin\{-2,1\}$. ניתן לחלק בסקלר $(1-a)(2+a)\neq 0$.

$$\begin{pmatrix}1 & a & 1 & | & 1\\ 0 & 1 & -1 & | & 0\\ 0 & 0 & 2-a-a^2 & | & 1-a\end{pmatrix} \xrightarrow{\frac{1}{(1-a)(2+a)}R_3\to R_3} \begin{pmatrix}1 & a & 1 & | & 1\\ 0 & 1 & -1 & | & 0\\ 0 & 0 & 1 & | & \frac{1}{a+2}\end{pmatrix}$$

$$\xrightarrow{R_2+R_3\to R_2} \begin{pmatrix}1 & a & 1 & | & 1\\ 0 & 1 & 0 & | & \frac{1}{a+2}\\ 0 & 0 & 1 & | & \frac{1}{a+2}\end{pmatrix} \xrightarrow{R_1-aR_2-R_3\to R_1} \begin{pmatrix}1 & 0 & 0 & | & \frac{1}{a+2}\\ 0 & 1 & 0 & | & \frac{1}{a+2}\\ 0 & 0 & 1 & | & \frac{1}{a+2}\end{pmatrix}$$

בסוף השתמשנו בחישוב

$$1-\frac{a}{a+2}-\frac{1}{a+2}=\frac{1}{a+2}.$$

נסכם: כאשר $a\notin\{-2,1\}$ מתקבל פתרון יחיד

$$(x,y,z)=\left(\frac{1}{a+2},\frac{1}{a+2},\frac{1}{a+2}\right).$$

כאשר $a=1$ מתקבלת קבוצת פתרונות אינסופית

$$\{\,(1-s-t,s,t)\mid s,t\in\mathbb{R}\,\}.$$

כאשר $a=-2$ אין פתרון.

הערה 5.42. בשאלות שקשורות רק למספר הפתרונות, אין צורך בדירוג קנוני. אפשר להסתפק בדירוג רגיל (חלקי) לחישוב הדרגות.

בניגוד לממ״ל כללית, לממ״ל הומוגנית תמיד קיים לפחות פתרון אחד. הטענה הבאה מראה זאת ומאפיינת את שני המקרים (פתרון יחיד ואינסוף פתרונות).

טענה 5.43.

א. לכל ממ״ל הומוגנית $Ax = 0$ קיים פתרון טריוויאלי שהוא וקטור האפס $x = 0$, כלומר הפתרון בו כל המשתנים מתאפסים.

ב. אם $\text{rank}(A) = n$, אז הפתרון יחיד.

ג. אם $\text{rank}(A) < n$, אז יש אינסוף פתרונות. זה בהכרח המקרה אם מתקיים $m < n$.

הוכחה.

א. נציב $x_1 = x_2 = ... = x_n = 0$ בממ״ל ונקבל

$$a_{i1} \cdot 0 + a_{i2} \cdot 0 + ... + a_{in} \cdot 0 = 0$$

כנדרש לכל $1 \leq i \leq n$.

ב. אם נחזור למטריצת המקדמים המורחבת (עם אפסים בעמודה הימנית), הדרגה שלה בהכרח שווה לזו של מטריצת המקדמים המצומצמת כי אין אפשרות לאיבר מוביל בעמודה הימנית. נקבל

$$\text{rank}(A|0) = \text{rank}(A) = n\,.$$

לכן, לפי טענה 5.40 יש פתרון יחיד.

ג. הפעם מתקיים

$$\text{rank}(A|0) = \text{rank}(A) < n\,.$$

לכן, לפי טענה 5.40 יש אינסוף פתרונות.

בהינתן $m < n$, נובע מ- 5.39 כי $\text{rank}(A) \leq m < n$. לכן בהכרח יש אינסוף פתרונות במקרה זה.

□

תרגיל 5.44. נחזור לממ"ל מדוגמה **5.41** ונחליף את כל הקבועים ב- 0, כלומר נסתכל על הממ"ל הבאה:

$$\begin{cases} ax + \ \ y + \ \ z = 0 \\ \ x + ay + \ \ z = 0 \\ \ x + \ \ y + az = 0 \end{cases}$$

לאילו ערכי $a \in \mathbb{R}$ יש פתרון יחיד? אינסוף פתרונות?

פתרון. נחזור על הדירוג שביצענו בדוגמה. ההבדל היחיד הוא שהפעם ניתן להשמיט את העמודה הימנית כי הממ"ל הומוגנית (צריך לזכור שמסתתרת שם עמודת אפסים בכל שלב של הדירוג). אז לפי תחילת הדירוג שעשינו, מטריצת המקדמים המצומצמת שקולת שורה למטריצה

$$\begin{pmatrix} 1 & a & 1 \\ 0 & 1-a^2 & 1-a \\ 0 & 1-a & a-1 \end{pmatrix}.$$

כבר רואים שאם $a = 1$ אז הדרגה היא 1, ולכן יש אינסוף פתרונות.
אם $a \neq 1$ נמשיך את הדירוג כמו בדוגמה. נקבל

$$\begin{pmatrix} 1 & a & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 2-a-a^2 \end{pmatrix}$$

הדרגה היא לפחות 2 ונקבע אותה לפי הערך בפינה התחתונה. אם מתקיים $2-a-a^2 \neq 0$, אז הדרגה היא 3 ולכן הפתרון יחיד (והוא הפתרון הטריוויאלי). באופן מפורש, יש פתרון יחיד לכל $a \notin \{-2, 1\}$ כי מתקיים

$$2-a-a^2 \neq 0 \iff a \notin \{-2,1\}.$$

אם $a \in \{-2,1\}$ יש אינסוף פתרונות כי הדרגה קטנה מ- 3 בשני המקרים.

פרק 6

מטריצות

בפרק הקודם ראינו מטריצות מקדמים בהקשר של ממ"ל (מצומצמות או מורחבות). אבל הרעיון של מטריצה הוא כללי יותר, ולא תמיד חייב להיות קשר לממ"ל. מטריצה אינה אלא טבלה של מספרים, וטבלה כזו יכולה לייצג כל מיני דברים.

דוגמה 6.1. נניח שאנחנו רוצים לחשב את מספר האוטובוסים (של חברה מסוימת) שעוברים בין אילת, באר שבע ות"א ביום ראשון. לצורך הדוגמה, לכל אוטובוס אין תחנות ביניים - רק תחנת יציאה ותחנת הגעה. אז אפשר להכין מטריצה מסדר 3×3 שבה כל שורה תציין תחנת יציאה, וכל עמודה תציין תחנת הגעה. נשתמש בסימונים E, B, T לת"א, באר שבע ואילת בהתאמה כדי שיהיה ברור איזו עמודה/שורה מתאימה לאיזו עיר (סימונים אלה אינם איברים במטריצה). אז לפי הנתונים מקבלים מטריצה כזו:

$$\left(\begin{array}{c|ccc} & E & B & T \\ \hline E & 0 & 10 & 5 \\ B & 8 & 0 & 12 \\ T & 4 & 6 & 0 \end{array}\right)$$

משמעות המטריצה היא שאין אוטובוס מאף תחנה לעצמה. יש עשרה אוטובוסים שיוצאים מאילת לבאר שבע, ויש שמונה אוטובוסים שעושים את המסלול ההפוך. וכן הלאה. נוותר על הסימונים של הערים ונכתוב את המטריצה בצורה הרגילה, ונסמנה ע"י S:

$$S = \begin{pmatrix} 0 & 10 & 5 \\ 8 & 0 & 12 \\ 4 & 6 & 0 \end{pmatrix}$$

זו המטריצה עבור הנתונים של יום ראשון. עכשיו נניח שיש נתונים שונים ליום שני, ונכין מטריצה חדשה עבורם:

$$M = \begin{pmatrix} 0 & 12 & 7 \\ 11 & 0 & 10 \\ 3 & 8 & 0 \end{pmatrix}$$

כעת אפשר לשאול כמה אוטובוסים יש לכל מסלול בשני הימים יחד. אז צריך לעשות חיבור בכל רכיב של המטריצה (שמתאים למסלול עם נקודת יציאה ונקודת הגעה) לחוד. אפשר לכתוב מטריצה חדשה שאיבריה יהיו סכומי האיברים של המטריצות S, M לפי כל רכיב לחוד. זה מתבקש לקרוא למטריצה החדשה $S + M$ בדומה לחיבור וקטורי. כך נקבל:

$$S + M = \begin{pmatrix} 0+0 & 10+12 & 5+7 \\ 8+11 & 0+0 & 12+10 \\ 4+3 & 6+8 & 0+0 \end{pmatrix} = \begin{pmatrix} 0+0 & 22 & 12 \\ 19 & 0 & 22 \\ 7 & 14 & 0 \end{pmatrix}$$

בפרט, משמעות המטריצה לעיל היא שאין אוטובוס מאף תחנה לעצמה באף יום. יש 22 אוטובסים שיוצאים מאילת לבאר שבע ביומיים של ראשון ושני, ו- 19 שעושים את המסלול ההפוך ביומיים האלה. וכן הלאה.

הגדרה 6.2. נגדיר את קבוצת כל המטריצות מסדר $m \times n$ מעל $\mathbb{F}$ (כאשר $\mathbb{F} = \mathbb{R}$ או $\mathbb{F} = \mathbb{C}$) ע"י

$$\mathbb{M}_{m\times n}(\mathbb{F}) = \left\{ \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \middle| \ \forall 1 \le i \le m \quad \forall 1 \le j \le n \quad a_{ij} \in \mathbb{F} \right\}.$$

הערה 6.3. הסימן $\forall$ פירושו "לכל". אז התנאי המתמטי בהגדרת $\mathbb{M}_{m\times n}(\mathbb{F})$ אומר שכל איבר a_{ij} במטריצה שייך לקבוצה $\mathbb{F}$, כי עוברים על כל האיברים בכל השורות ובכל העמודות.

דוגמה 6.4.

א. מתקיים

$$A = \begin{pmatrix} 10 & 9 & 1 & -7 \\ 1 & 2 & 3 & 4 \\ 0 & 4 & -3 & -5 \end{pmatrix} \in \mathbb{M}_{3\times 4}(\mathbb{R}).$$

ב. מתקיים

$$B = \begin{pmatrix} i & 1 \\ 1 & 1+i \\ 0 & 2i \end{pmatrix} \in \mathbb{M}_{3\times 2}(\mathbb{C}).$$

הערה 6.5. לכל $m,n \in \mathbb{N}$ מתקיים $\mathbb{M}_{m\times n}(\mathbb{R}) \subseteq \mathbb{M}_{m\times n}(\mathbb{C})$ כי $\mathbb{R} \subseteq \mathbb{C}$. אז למעשה כל המטריצות בקורס הן מעל $\mathbb{C}$, אבל אם לדייק רובן (לא כולן) יהיו מעל $\mathbb{R}$. השימוש בסימון $\mathbb{F}$ נועד לרוב הטענות, שנכונות גם מעל $\mathbb{R}$ וגם מעל $\mathbb{C}$. בתחום של אלגברה לינארית אפשר גם לעסוק בקבוצות מספרים נוספות שנקראות שדות (fields), ומכאן הסיבה לאות $\mathbb{F}$. אבל בקורס שלנו נסתפק ב- $\mathbb{R}, \mathbb{C}$ (השדות החשובים ביותר).

6.1 פעולות על מטריצות

6.1.1 חיבור וכפל בסקלר

ראינו בדוגמה 6.1 שיש טעם בהגדרת פעולת חיבור בין מטריצות, בדיוק כמו חיבור וקטורי (עושים חיבור בכל קוארדינטה/רכיב לחוד). קודם כל נכליל חיבור וקטורי וכפל בסקלר (שהגדרנו בפרק 4 במישור ובמרחב) ל- $\mathbb{F}^n$.

הגדרה 6.6. עבור וקטורים $= (a_1, ..., a_n) \in \mathbb{F}^n, w = (b_1, ..., b_n) \in \mathbb{F}^n$ וסקלר $c \in \mathbb{F}$ נגדיר את הפעולות הבאות:

א. פעולת החיבור:

$$v + w = (a_1 + b_1, ..., a_n + b_n)$$

ב. פעולת כפל בסקלר:

$$c \cdot v = (c \cdot a_1, \ldots, c \cdot a_n)$$

הגדרה 6.7. עבור $A, B \in \mathbb{M}_{m\times n}(\mathbb{F})$ וסקלר $c \in \mathbb{F}$ נגדיר את הפעולות הבאות:

א. פעולת החיבור:

$$\begin{aligned} A + B &= \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ b_{21} & \cdots & b_{2n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} \\ &= \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & \cdots & a_{2n} + b_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix} \end{aligned}$$

ב. פעולת כפל בסקלר:

$$cA = \begin{pmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{pmatrix}$$

בפרט, נגדיר $-A = (-1) \cdot A$ **ובהתאמה** $B - A = B + (-1) \cdot A$.

דוגמה 6.8. **א. חיבור של מטריצות מסדר** 2×2 **מעל** $\mathbb{R}$:

$$\begin{aligned} &\begin{pmatrix} 3 & 7 \\ 5 & -2 \end{pmatrix} + \begin{pmatrix} -10 & 5 \\ -4 & -3 \end{pmatrix} \\ &= \begin{pmatrix} 3-10 & 7+5 \\ 5-4 & -2-3 \end{pmatrix} = \begin{pmatrix} -7 & 12 \\ 1 & -5 \end{pmatrix} \end{aligned}$$

ב. חיבור של מטריצות מסדר 4×3 **מעל** $\mathbb{R}$:

$$\begin{aligned} &\begin{pmatrix} 10 & 9 & 1 & -7 \\ 1 & 2 & 3 & 4 \\ 0 & 4 & -3 & -5 \end{pmatrix} + \begin{pmatrix} 3 & -2 & 0 & 6 \\ 5 & -2 & 4 & -4 \\ 10 & -6 & 3 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 10+3 & 9-2 & 1+0 & -7+6 \\ 1+5 & 2-2 & 3+4 & 4-4 \\ 0+10 & 4-6 & -3+3 & -5+1 \end{pmatrix} \\ &= \begin{pmatrix} 13 & 7 & 1 & -1 \\ 6 & 0 & 7 & 0 \\ 10 & -2 & 0 & -4 \end{pmatrix} \end{aligned}$$

ג. חיבור של מטריצות מסדר 2×2 **מעל** $\mathbb{C}$: **אם**

$$A = \begin{pmatrix} 1 & i \\ 2-3i & 3-2i \end{pmatrix}, \quad B = \begin{pmatrix} 7-5i & 2+7i \\ -6 & 3+10i \end{pmatrix}$$

אז

$$A + B = \begin{pmatrix} 1+7-5i & i+2+7i \\ 2-3i-6 & 3-2i+3+10i \end{pmatrix} = \begin{pmatrix} 8-5i & 2+8i \\ -4-3i & 6+8i \end{pmatrix}$$

הגדרה 6.9. יש דרך מקוצרת לכתיבת חיבור. במקום לכתוב את החיבור בכל רכיב, נסתכל על האיבר הכללי של כל מטריצה. נסמן את האיבר בשורה ה- i ובעמודה ה- j ע"י $(P)_{ij}$ לכל $P \in \mathbb{M}_{m\times n}(\mathbb{F})$. אז למעשה הגדרנו לכל $1 \leq i \leq m$ ולכל $1 \leq j \leq n$:

$$(A+B)_{ij} = (A)_{ij} + (B)_{ij}$$

$$(cA)_{ij} = c(A)_{ij}$$

למשל, בדוגמה האחרונה היה לנו $(A)_{12} = i, (B)_{21} = -6, (A+B)_{22} = 6+8i$.

הערה 6.10. שימו לב שהחיבור מוגדר רק למטריצות מאותו הסדר (כנ"ל לוקטורים - רק אם הם שווי אורך). גם הסדר של מטריצת הסכום שווה לסדר המשותף של המטריצות.

דוגמה 6.11. החיבור הבא אינו מוגדר:

$$\begin{pmatrix} 3 & 7 \\ 5 & -2 \end{pmatrix} + \begin{pmatrix} -10 & 5 & 0 \\ -4 & -3 & 0 \end{pmatrix}$$

המטריצה השמאלית היא מסדר 2×2 והימנית היא מסדר 2×3.

דוגמה 6.12 (דוגמאות נוספות).

א. כפל בסקלר של מטריצה מסדר 3×3 מעל $\mathbb{C}$:

$$i\begin{pmatrix} 1 & i & 1+i \\ 1-i & 2i & 3 \\ 2+3i & 3-2i & 5 \end{pmatrix} = \begin{pmatrix} i\cdot 1 & i\cdot i & i(1+i) \\ i(1-i) & i\cdot 2i & i\cdot 3 \\ i(2+3i) & i(3-2i) & i\cdot 5 \end{pmatrix}$$
$$= \begin{pmatrix} i & -1 & -1+i \\ 1+i & -2 & 3i \\ -3+2i & 2+3i & 5i \end{pmatrix}$$

ב. שילוב של כפל בסקלר וחיבור מטריצות מסדר 2×2 מעל $\mathbb{R}$:

$$2\begin{pmatrix} 11 & 2 \\ -2 & 4 \end{pmatrix} + 3\begin{pmatrix} -4 & 5 \\ 6 & -3 \end{pmatrix} = \begin{pmatrix} 22 & 4 \\ -4 & 8 \end{pmatrix} + \begin{pmatrix} -12 & 15 \\ 18 & -9 \end{pmatrix}$$
$$= \begin{pmatrix} 22-12 & 4+15 \\ -4+18 & 8-9 \end{pmatrix}$$

הגדרה 6.13. לכל $m,n \in \mathbb{N}$ נגדיר את מטריצת האפס מסדר $m \times n$ כמטריצה שכל איבריה הם 0:

$$\mathbf{0}_{m\times n} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

טענה 6.14. *יהיו* $A,B,C \in \mathbb{M}_{m\times n}(\mathbb{F})$ *ו-* $\alpha,\beta \in \mathbb{F}$*. אז מתקיימים הכללים הבאים:*

א. חוק החילוף (קומוטטיביות):

$$A+B=B+A$$

ב. חוק הקיבוץ (אסוציאטיביות):

$$(A+B)+C=A+(B+C)$$

ג. חוק הקיבוץ לכפל בסקלר:

$$(\alpha\beta)A=\alpha(\beta A)$$

ד. חוק הפילוג (דיסטריביוטיביות):

$$\alpha(A+B)=\alpha A+\alpha B$$

ה. ניטרליות היחידה:

$$1\cdot A=A$$

ו.

$$0\cdot A=\mathbf{0}_{m\times n}$$

ז. ניטרליות מטריצת האפס:

$$A+\mathbf{0}_{m\times n}=A$$

ח.

$$A-A=0$$

הוכחה. כל התכונות עוברות בירושה מחיבור וכפל ב- $\mathbb{F}$ (מדובר בתכונות מוכרות של $\mathbb{R}$ וראינו בפרק 3 שהן עוברות בירושה ל- $\mathbb{C}$). נוכיח למשל את חוק הפילוג: נתמקד באיבר הכללי של כל מטריצה בשוויון שצריך להוכיח. לכל $1\le i\le m, 1\le j\le n$ מתקיים

$$.(\alpha(A+B))_{ij} = \alpha(A+B)_{ij} = \alpha((A)_{ij} + (B)_{ij}) = \alpha(A)_{ij} + \alpha(B)_{ij}$$

שימו לב שהמעבר האחרון נכון לפי חוק הפילוג ב-$\mathbb{F}$. לכן נובע כי

$$.\alpha(A+B) = \alpha A + \beta A$$

נוכיח גם את ניטרליות מטריצת האפס:

$$(A + \mathbf{0}_{m\times n})_{ij} = (A)_{ij} + (\mathbf{0}_{m\times n})_{ij} = (A)_{ij} + 0 = (A)_{ij}$$

לכל $1 \le i \le m, 1 \le j \le n$, ולכן

$$,A + \mathbf{0}_{m\times n} = A$$

ההוכחות של שאר הסעיפים דומות, כאשר הרעיון הכללי הוא שימוש ב- 6.9 והתכונות המתאימות של חיבור/כפל ב- $\mathbb{F}$.

□

תרגיל 6.15. הוכיחו את חוק הקיבוץ לכפל בסקלר.

פתרון. עבור חוק הקיבוץ: לכל $1 \le i \le m, 1 \le j \le n$ מתקיים

$$.((A+B)+C)_{ij} = (A+B)_{ij} + (C)_{ij} = ((A)_{ij} + (B)_{ij}) + (C)_{ij}$$

באופן דומה, מתקיים

$$.(A+(B+C))_{ij} = A_{ij} + (B+C)_{ij} = (A)_{ij} + ((B)_{ij} + (C)_{ij})$$

שני הביטויים באגף ימין של שתי המשוואות שווים לפי חוק הקיבוץ ב- $\mathbb{F}$. לכן

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

6.1.2 כפל מטריצה בוקטור עמודה

בפרק הקודם דיברנו על ממ"ל מהצורה $Ax = b$ כאשר

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

היא מטריצת המקדמים המצומצמת, $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ הוא וקטור המשתנים ו- $b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$ הוא וקטור הקבועים. נרצה להגדיר באופן מדויק את פעולת הכפל Ax של מטריצה $A \in \mathbb{M}_{m\times n}(\mathbb{F})$ בוקטור עמודה $x \in \mathbb{F}^n$ שמספר שורותיו שווה למספר העמודות של המטריצה.

לפני כן, נגדיר את סימון Σ לסכימה (שאולי כבר מוכר).

הגדרה 6.16. בהינתן $a_1, a_2, ..., a_n \in \mathbb{F}$ נסמן את סכומם בעזרת כתיבה עם Σ ואינדקס l.

$$\sum_{l=1}^{n} a_l = a_1 + a_2 + a_3 + \cdots + a_n .$$

האינדקס l רץ על פני הקבוצה $\{ 1, 2, 3, \cdots, n \}$ וכל הצבה תורמת איבר לסכום.

דוגמה 6.17. במקרה $a_l = l$ נקבל

$$\sum_{l=1}^{5} l = 1 + 2 + 3 + 4 + 5 = 15$$

הגדרה 6.18. בהינתן $A \in \mathbb{M}_{m\times n}(\mathbb{F})$ ו-

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{F}^n ,$$

נגדיר את המכפלה

$$Ax = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \in \mathbb{F}^m$$

עם קוארדינטות הנתונות ע"י

$$y_i = \sum_{j=1}^{n} (A)_{ij} x_j$$

לכל $1 \leq i \leq m$. נסמן $y_i = (Ax)_i$ עבור הקוארדינטות.

הערה 6.19. לפי הגדרה זו, לכל $1 \le i \le m$ מתקיים

$$.(Ax)_i = b_i \iff \sum_{j=1}^{n} (A)_{ij} x_j = b_i$$

לכן, השוויון הוקטורי $Ax = b$ אכן שקול לממ"ל.

הערה 6.20. למעשה יש זיהוי בין $\mathbb{M}_{1\times n}(\mathbb{F})$ לבין $\mathbb{F}^n$, לפי הגדרה 5.3 שהתייחסה לוקטורי שורה. אבל נרשה לעצמנו להשתמש בסימון $\mathbb{F}^n$ גם עבור וקטורי עמודה באורך n, כלומר עבור הקבוצה $\mathbb{M}_{n\times 1}(\mathbb{F})$. לפעמים ההבדל בין וקטור שורה לוקטור עמודה הוא אסתטי בלבד, אבל כאשר מכפילים מטריצה בוקטור בהחלט יש חשיבות לסוג הוקטור (הוא צריך להיות וקטור עמודה). אז ככלל נזכור לשים לב להבדל כאשר מופיעה מכפלה של מטריצה בוקטור, ובשאר המקרים אפשר לכתוב את הוקטור כשורה או עמודה לפי טעם אישי.

דוגמה 6.21.

א. ניקח

$$A = \begin{pmatrix} 2 & -1 \\ 3 & 4 \end{pmatrix}, \quad x = \begin{pmatrix} 1 \\ 5 \end{pmatrix}$$

ונקבל

$$.\begin{pmatrix} 2 & -1 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 5 \end{pmatrix} = \begin{pmatrix} 2\cdot 1 + (-1)5 \\ 3\cdot 1 + 4\cdot 5 \end{pmatrix} = \begin{pmatrix} -3 \\ 23 \end{pmatrix}$$

ב.

$$\begin{pmatrix} a_1 & b_1 & c_1 \end{pmatrix} \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix} = a_1a_2 + b_1b_2 + c_1c_2$$

לכל $a_1, b_1, c_1, a_2, b_2, c_2 \in \mathbb{F}$. במקרה של $\mathbb{F} = \mathbb{R}$ מדובר על המכפלה הסקלרית של

$$.\begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix}, \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix} \in \mathbb{R}^3$$

ג.

$$\begin{pmatrix} 1 & 0 & 2 \\ -2 & 3 & 1 \\ 4 & -1 & 0 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 1\cdot 3 + 0\cdot 2 + 2(-1) \\ -2\cdot 3 + 3\cdot 2 + 1(-1) \\ 4\cdot 3 + (-1)2 + 0(-1) \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 10 \end{pmatrix}$$

ד.

$$\begin{pmatrix} 1+i & 2 \\ 3i & 4-i \end{pmatrix}\begin{pmatrix} 2 \\ 1-i \end{pmatrix} = \begin{pmatrix} (1+i)2+2(1-i) \\ (3i)2+(4-i)(1-i) \end{pmatrix} = \begin{pmatrix} 4 \\ 3+i \end{pmatrix}$$

תרגיל 6.22. חשבו את Av עבור

$$.A = \begin{pmatrix} 1 & i & 2 \\ -1 & 3i & 0 \\ 2+i & 1 & -i \end{pmatrix}, \quad v = \begin{pmatrix} 1+i \\ 2 \\ -1 \end{pmatrix}$$

פתרון.

$$Ax = \begin{pmatrix} 1(1+i)+i\cdot 2+2(-1) \\ (-1)(1+i)+3i\cdot 2+0(-1) \\ (2+i)(1+i)+1\cdot 2+(-i)(-1) \end{pmatrix} = \begin{pmatrix} 1+i+2i-2 \\ -1-i+6i \\ 1+3i+2+i \end{pmatrix} = \begin{pmatrix} -1+3i \\ -1+5i \\ 3+4i \end{pmatrix}$$

טענה 6.23. *יהיו* $A,B\in\mathbb{M}_{m\times n}(\mathbb{F})$*,* $v,w\in\mathbb{F}^n$ *ו-* $\alpha\in\mathbb{F}$*.*

א.

$$A(v+w) = Av + Aw$$

ב.

$$(A+B)v = Av + Bv$$

ג.

$$A(\alpha v) = \alpha(Av)$$

הוכחה. בכל סעיף נראה שוויון בין הקוארדינטות ה- i של שני האגפים, לכל $1\le i\le m$. ראשית נסמן

$$.v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}, \quad w = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}$$

א.

$$(A(v+w))_i = \sum_{j=1}^{n}(A)_{ij}(v+w)_j = \sum_{j=1}^{n}(A)_{ij}v_j + \sum_{j=1}^{n}(A)_{ij}w_j = (Av)_i + (Aw)_i$$

בשביל השוויון השני השתמשנו בחוק הפילוג ב- $\mathbb{F}$, וגם פיצלנו את הסכום לפי חוק החילוף (אין חשיבות לסדר הסכימה).

ב.

$$\begin{aligned}((A+B)v)_i &= \sum_{j=1}^{n}((A)_{ij}+(B)_{ij})v_j = \sum_{j=1}^{n}(A)_{ij}v_j + \sum_{j=1}^{n}(B)_{ij}v_j \\ &= (Av)_i + (Bv)_i = (Av+Bv)_i\end{aligned}$$

ג.

$$(A(\alpha v))_i = \sum_{j=1}^{n}(A)_{ij}(\alpha v_j) = \alpha\sum_{j=1}^{n}(A)_{ij}v_j = \alpha(Av)_i = (\alpha Av)_i$$

בשוויון השני השתמשנו בחוק הפילוג ב- $\mathbb{F}$, לפיו אפשר להוציא גורם משותף מחוץ לסכימה.

□

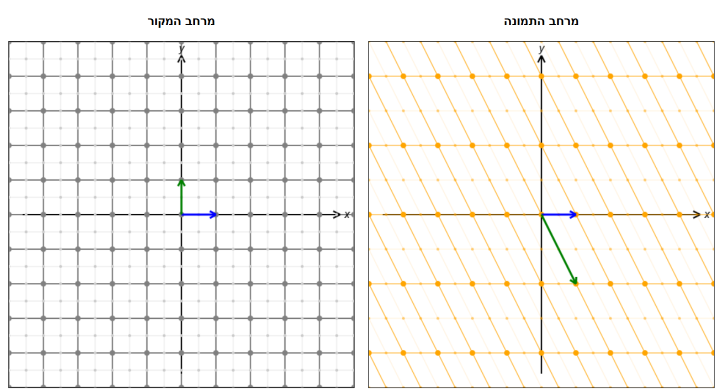
בהינתן מטריצה A, אפשר לחשוב על הכפל Av כפעולה על v. זהו רעיון שנחזור אליו בהמשך הקורס, אבל כבר ניתן להתרשם מהפעולה הזו באופן גיאומטרי, למשל עבור

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

אפשר לצייר רשת במרחב המקור (עבור v) בהתאם לוקטורים

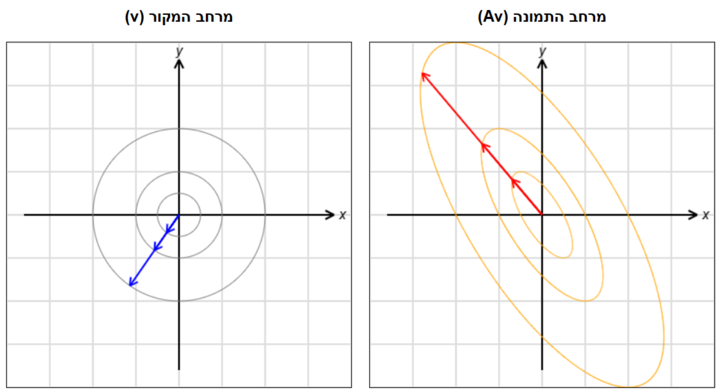
$$\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}$$

ולראות איזו רשת נוצרת במרחב התמונה (עבור Av):



איור 6.1: הכפל במטריצה משפיע על שני הוקטורים באופן שונה, אבל כל ריבוע במרחב המקור הופך למקבילית במרחב התמונה

במקום רשת של ריבועים, אפשר להסתכל על מעגלים סביב הראשית ולראות איזה מסלול מתאים לכל מעגל במרחב התמונה:



איור 6.2: כל מעגל במרחב המקור נשלח לאליפסה במרחב התמונה, ומתיחה של המעגל מתאימה למתיחה של האליפסה

טענה 6.24. *תהי $Ax = 0$ ממ״ל הומוגנית.*

א. קבוצת הפתרונות סגורה לחיבור, כלומר לכל שני פתרונות $v, w \in \mathbb{F}^n$ גם $v + w$ הוא פתרון.

ב. קבוצת הפתרונות סגורה לכפל בסקלר, כלומר לכל פתרון $v \in \mathbb{F}^n$ ולכל $\alpha \in \mathbb{F}$ גם αv הוא פתרון.

הוכחה. יהיו $v, w \in \mathbb{F}^n$ פתרונות של הממ״ל.

א. לפי טענה 6.23 מתקיים

$$A(v + w) = Av + Aw = 0 + 0 = 0.$$

כלומר $v + w$ הוא פתרון של הממ״ל.

ב. שוב, לפי טענה 6.23 לכל $\alpha \in \mathbb{F}$ מתקיים

$$A(\alpha v) = \alpha Av = \alpha \cdot 0 = 0.$$

כלומר αv הוא פתרון של הממ״ל.

□

טענה 6.25. *תהי* $Ax=b$ *ממ"ל לא הומוגנית. נניח שקיים לה פתרון (לא בהכרח יחיד)* $v_0\in\mathbb{F}^n$. *אז מתקיים הקשר הבא בין קבוצת הפתרונות של הממ"ל הלא הומוגנית לקבוצת הפתרונות של הממ"ל ההומוגנית* $Ax=0$:

$$\{\,x\in\mathbb{F}^n \mid Ax=b\,\}=\{\,v_0+y \mid y\in\mathbb{F}^n, Ay=0\,\}.$$

הערה 6.26. הרעיון הזה כבר מוכר לנו מההצגה הפרמטרית של ישר/מישור. למשל, ראינו שלישר במישור יש הצגה פרמטרית מהצורה

$$\left\{\begin{pmatrix}x_0\\y_0\end{pmatrix}+t\begin{pmatrix}\alpha\\\beta\end{pmatrix}\;\middle|\; t\in\mathbb{R}\right\}$$

כאשר $\begin{pmatrix}\alpha\\\beta\end{pmatrix}\in\mathbb{R}^2$ הוא וקטור כיוון (בפרק 4 השתמשנו בוקטורי שורה, אבל בהקשר של כפל צריך לקחת וקטורי עמודה). אם הישר מתאר קבוצת פתרונות של ממ"ל $\begin{pmatrix}a_1 & a_2\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=b$, אז וקטור הכיוון הוא פתרון של הממ"ל ההומוגנית. כלומר מתקיים $\begin{pmatrix}a_1 & a_2\end{pmatrix}\begin{pmatrix}\alpha\\\beta\end{pmatrix}=0$, או באופן שקול $a_1\alpha+a_2\beta=0$. יותר מכך, קבוצת הפתרונות של הממ"ל ההומוגנית היא קבוצת הכפולות בסקלר של וקטור הכיוון.

נוכיח את הטענה:

הוכחה. נוכיח שוויון בין שתי הקבוצות ע"י הכלה דו-כיוונית.

הכיוון $\supseteq$: יהי $y\in\mathbb{F}^n$ כך ש- $Ay=0$. אז מתקיים

$$A(v_0+y)=Av_0+Ay=b+0=b,$$

ולכן $v_0+y\in\{\,x\in\mathbb{F}^n \mid Ax=b\,\}$.

הכיוון $\subseteq$: יהי $x\in\mathbb{F}^n$ כך ש- $Ax=b$. אז מתקיים $x=v_0+y$ עבור $y=x-v_0$. נבדוק ש-y הוא פתרון של הממ"ל ההומוגנית:

$$Ay=A(x-v_0)=Ax-Av_0=b-b=0.$$

לכן

$$x=v_0+y\in\{\,x_0+y \mid y\in\mathbb{F}^n, Ay=0\,\}.$$

□

6.1.3 כפל מטריצות

הגדרה 6.27. בהינתן $A \in \mathbb{M}_{m\times k}(\mathbb{F}), B \in \mathbb{M}_{k\times n}(\mathbb{F})$ נגדיר את מטריצת הכפל AB ע"י

$$(AB)_{ij} = \sum_{l=1}^{k}(A)_{il}(B)_{lj} = (A)_{i1}(B)_{1j} + (A)_{i2}(B)_{2j} + ... + (A)_{ik}(B)_{kj}$$

לכל $1 \leq i \leq m, 1 \leq j \leq n$.

הערה 6.28. שימו לב שכפל בין מטריצות מוגדר רק כאשר מספר העמודות של המטריצה השמאלית שווה למספר השורות של המטריצה הימנית, וסימנו את מספר זה ע"י k. זה אומר שבהרבה מקרים פעולת הכפל אינה מוגדרת, למשל כאשר A, B שתיהן מסדר 2×3.

דוגמה 6.29.

א. עבור $B = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}, A = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$ נקבל

$$AB = \begin{pmatrix} 2\cdot 5 + 3\cdot 7 & 2\cdot 6 + 3\cdot 8 \\ 1\cdot 5 + 4\cdot 7 & 1\cdot 6 + 4\cdot 8 \end{pmatrix} = \begin{pmatrix} 31 & 36 \\ 33 & 38 \end{pmatrix}$$

לעומת זאת, אם נחשב את המכפלה בסדר הפוך נקבל

$$BA = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}\begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix} = \begin{pmatrix} 5\cdot 2 + 6\cdot 1 & 5\cdot 3 + 6\cdot 4 \\ 7\cdot 2 + 8\cdot 1 & 7\cdot 3 + 8\cdot 4 \end{pmatrix} = \begin{pmatrix} 16 & 39 \\ 22 & 53 \end{pmatrix}$$

אז במקרה זה $AB \neq BA$.

ב. עבור $D = \begin{pmatrix} 1 & 0 & 2 \\ -1 & 3 & 1 \\ 0 & 4 & 5 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ נקבל

$$\begin{aligned} CD &= \begin{pmatrix} 1\cdot 1 + 2(-1) + 3\cdot 0 & 1\cdot 0 + 2\cdot 3 + 3\cdot 4 & 1\cdot 2 + 2\cdot 1 + 3\cdot 5 \\ 4\cdot 1 + 5(-1) + 6\cdot 0 & 4\cdot 0 + 5\cdot 3 + 6\cdot 4 & 4\cdot 2 + 5\cdot 1 + 6\cdot 5 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 18 & 19 \\ -1 & 39 & 43 \end{pmatrix} \end{aligned}$$

ג. עבור $F = \begin{pmatrix} 2 & i \\ 1-i & 3 \end{pmatrix}, \quad E = \begin{pmatrix} 1+i & 2 \\ 3 & 4-i \end{pmatrix}$ נקבל

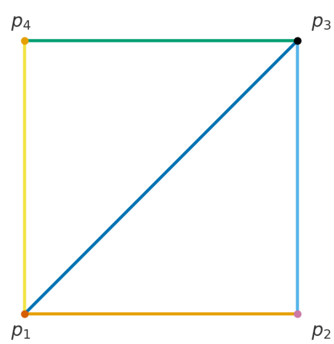
$$\begin{aligned} EF &= \begin{pmatrix} (1+i)2+2(1-i) & (1+i)i+2\cdot 3 \\ 3\cdot 2+(4-i)(1-i) & 3\cdot i+(4-i)3 \end{pmatrix} \\ &= \begin{pmatrix} 2+2i+2-2i & i-1+6 \\ 6+4-4i-i-1 & 3i+12-3i \end{pmatrix} = \begin{pmatrix} 4 & 5+i \\ 9-5i & 12 \end{pmatrix} \end{aligned}$$

ד. המכפלה

$$\begin{pmatrix} 2 & 3 & 1 \\ 0 & -1 & -2 \end{pmatrix} \begin{pmatrix} 10 & -5 \\ 3 & -1 \end{pmatrix}$$

אינה מוגדרת, כי מספר העמודות של המטריצה השמאלית (3) שונה ממספר השורות של המטריצה הימנית (2).

ה. נסתכל על דוגמה שיש לה קשר לתורת הגרפים (שחורגת מהקורס). גרף הוא קבוצת קודקודים שחלקם מחוברים ביניהם ע"י צלעות. לדוגמה:



איור 6.3: ריבוע עם אחד מאלכסוניו

ניתן לקודד את כל המידע הרלוונטי לגרף זה במטריצה M (שנקראת מטריצת השכנויות). אם קיימת צלע בין p_i ל- p_j, נכתוב $(M)_{ij} = 1$. אם לא קיימת צלע ביניהם, נכתוב $(M)_{ij} = 0$. כך נקבל

$$M = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}.$$

מה המשמעות של כפל? נחשב את $M^2 = M \cdot M$ ונקבל

$$.M^2 = \begin{pmatrix} 3 & 1 & 2 & 1 \\ 1 & 2 & 1 & 2 \\ 2 & 1 & 3 & 1 \\ 1 & 2 & 1 & 2 \end{pmatrix}$$

אם המטריצה M מתארת מספרי מסלולים באורך 1, אז המטריצה M^2 מתארת מספרי מסלולים באורך 2. למשל, בין p_1 לעצמו יש 3 מסלולים באורך 2, שהם:

$$p_1 \to p_2 \to p_1, \quad p_1 \to p_3 \to p_1, \quad p_1 \to p_4 \to p_1$$

זה מתאים לחישוב הבא:

$$.(M^2)_{11} = \sum_{k=1}^{4} (M)_{1k}(M)_{k1} = 0 \cdot 0 + 1 \cdot 1 + 1 \cdot 1 + 1 \cdot 1 = 3$$

כל מחובר $(M)_{1k}(M)_{k1}$ שתורם 1 לסכום, מתאר מסלול $p_1 \to p_k \to p_1$ עם תחנת ביניים p_k. לכן, הסכום מתאר את מספר המסלולים מ- p_1 לעצמו באורך 2.

ובהכללה: לכל $k \in \mathbb{N}$ אפשר לחשב את החזקה $M^k = \underbrace{M \cdot M \cdot \ldots \cdot M}_{\text{פעמים } k}$ ואז $(M^k)_{ij}$ שווה למספר המסלולים באורך k מ- p_i ל- p_j, כאשר אין הגבלה על תחנות הביניים (יכולה להיות חזרה על קודקוד אחד או יותר). רעיון זה תקף לכל גרף, והוא דוגמה טובה לשימוש של כפל מטריצות.

הערה 6.30. ראינו כי אין חוק חילוף למטריצות. בהרבה מקרים מתקיים $AB \neq BA$ גם אם שתי פעולות הכפל מוגדרות. לפעמים כן מתקיים $AB = BA$, ואז נאמר שהמטריצות A, B מתחלפות. ניתקל במקרים כאלה בהמשך, אבל בינתיים חשוב להבין שרוב המטריצות אינן מתחלפות.

תרגיל 6.31. חשבו את AB עבור

$$.A = \begin{pmatrix} 1 & -2 & 3 \\ -4 & 5 & -6 \end{pmatrix}, \quad \begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 0 \\ 3 & -1 & 7 \end{pmatrix}$$

פתרון. נחשב לפי ההגדרה:

$$AB = \begin{pmatrix} 1 & -2 & 3 \\ -4 & 5 & -6 \end{pmatrix} \begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 0 \\ 3 & -1 & 7 \end{pmatrix} = \begin{pmatrix} 8 & -7 & 22 \\ -20 & 13 & -46 \end{pmatrix}$$

טענה 6.32. *יהיו $A \in \mathbb{M}_{m\times k}(\mathbb{F}), B \in \mathbb{M}_{k\times n}(\mathbb{F})$. נניח שהעמודות של B הם הוקטורים $v_1, v_2, ..., v_n \in \mathbb{F}^k$. אז מתקיים*

$$AB = A\begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix} = \begin{pmatrix} | & | & & | \\ Av_1 & Av_2 & \cdots & Av_n \\ | & | & & | \end{pmatrix}$$

כלומר התוצאה של כפל ב- A משמאל הוא שכל וקטור עמודה מוכפל ב- A לחוד.

הוכחה. ישירות לפי ההגדרות של כפל מטריצות וכפל מטריצה בוקטור עמודה (שהוא מטריצה עם עמודה אחת). לכל $1 \le j \le n$ נכתוב $(B)_{ij} = b_{ij}$, כלומר

$$v_j = \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{kj} \end{pmatrix}.$$

אז לפי הגדרות הכפל נובע כי

$$(AB)_{ij} = \sum_{l=1}^{k}(A)_{il}(B)_{lj} = \sum_{l=1}^{k} A_{il}b_{lj} = (Av_j)_i.$$

לכן וקטור העמודה ה- j של AB הוא אכן Av_j. □

טענה 6.33. *יהיו $A_1, A_2 \in \mathbb{M}_{m\times n}(\mathbb{F})$, $B_1, B_2 \in \mathbb{M}_{n\times p}(\mathbb{F})$, $C \in \mathbb{M}_{p\times q}(\mathbb{F})$ ו- $\alpha \in \mathbb{F}$. אז מתקיים*

א.

$$(A_1 + A_2)B_1 = A_1B_1 + A_2B_1$$

ב.

$$A_1(B_1 + B_2) = A_1B_1 + A_1B_2$$

ג.

$$A_1(B_1C) = (A_1B_1)C$$

ד.

$$(\alpha A_1)B_1 = \alpha(A_1B_1) = A_1(\alpha B_1)$$

ה.

$$\mathbf{0}_{m\times n}B_1 = \mathbf{0}_{m\times p} = A_1\mathbf{0}_{n\times p}$$

הוכחה. רוב הסעיפים נובעים משילוב בין טענה 6.23 לטענה 6.32. לכן נסתפק בהוכחת סעיף ג' (חוק הקיבוץ לכפל מטריצות), שהוא היוצא מן הכלל. נניח כי

$$(A_1)_{ik} = a_{ik}, \quad (B_1)_{kl} = b_{kl}, \quad (C)_{lj} = c_{lj}.$$

אם כן:

$$(B_1C)_{kj} = \sum_{l=1}^{p} b_{kl}c_{lj},$$

ולכן

(6.1)

$$(A_1(B_1C))_{ij} = \sum_{k=1}^{n} a_{ik}(B_1C)_{kj} = \sum_{k=1}^{n} a_{ik}\left(\sum_{l=1}^{p} b_{kl}c_{lj}\right) = \sum_{k=1}^{n}\sum_{l=1}^{p} a_{ik}b_{kl}c_{lj}.$$

כעת נחשב את $((A_1B_1)C)_{ij}$:

$$(A_1B_1)_{il} = \sum_{k=1}^{n} a_{ik}b_{kl},$$

ולכן

(6.2)

$$((A_1B_1)C)_{ij} = \sum_{l=1}^{p}(A_1B_1)_{il}c_{lj} = \sum_{l=1}^{p}\left(\sum_{k=1}^{n} a_{ik}b_{kl}\right)c_{lj} = \sum_{l=1}^{p}\sum_{k=1}^{n} a_{ik}b_{kl}c_{lj}.$$

נשים לב כי יש שוויון בין שני הביטויים בצד ימין של 6.1 ו- 6.2, כי לפי חוק החילוף אין חשיבות לסדר הסכימה (אפשר לסכום קודם לפי l או קודם לפי k בלי לשנות את הסכום). לכן, לכל i, j מתקיים $(A_1(B_1C))_{ij} = ((A_1B_1)C)_{ij}$, ומכאן

$$A_1(B_1C) = (A_1B_1)C.$$

$\square$

הגדרה 6.34. מטריצת היחידה מסדר $n \times n$ (אפשר גם לומר מסדר n) מוגדרת להיות

$$I_n = I_{n\times n} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

זוהי המטריצה שבה כל איבר על האלכסון הראשי הוא 1, וכל שאר האיברים הם 0. או באופן מתמטי:

$$\begin{aligned}\forall 1 \leq i \leq n \quad (I_n)_{ii} = 1\\ \forall 1 \leq i,j \leq n \quad i \neq j \implies (I_n)_{ij} = 0\end{aligned}$$

דוגמה 6.35.

א.

$$I_1 = \begin{pmatrix} 1 \end{pmatrix}$$

זהו בעצם הסקלר 1, עם הבדל פורמלי של סוגריים.

ב.

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

ג.

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

מטריצת היחידה (מכל סדר) משחקת את התפקיד של הסקלר 1 במובן של ניטרליות ביחס לכפל. ליתר דיוק:

טענה 6.36. *לכל* $A \in \mathbb{M}_{m\times n}(\mathbb{F})$ *מתקיים* $AI_n = A = I_mA$.

הוכחה. לכל $1 \leq i,j \leq n$ מתקיים

$$(AI_n)_{ij} = \sum_{l=1}^{n} (A)_{il}(I_n)_{lj}.$$

נשים לב כי לפי הגדרת מטריצת היחידה, מתקיים $(I)_{lj} = 0$ לכל $l \neq j$. אז רק איבר אחד בסכום באגף ימין לא מתאפס, והוא מתאים להצבה $l = j$. לכן

$$(AI_n)_{ij} = 0 + ... + 0 + (A)_{ij}(I_n)_{jj} + 0 + ... + 0 = (A)_{ij}.$$

אז $AI_n = A$ כי יש שוויון בין כל האיברים המתאימים. באופן דומה:

$$(I_mA)_{ij} = \sum_{l=1}^{m} (I_m)_{il}A_{lj} = 0 + ... + 0 + (I)_{ii}(A)_{ij} + 0 + ... + 0 = (A)_{ij}.$$

ולכן $I_mA = A$.

$\square$

הערה 6.37. אם $m = n$, אז לפי הטענה מתקיים $AI_n = A = I_nA$. במילים: I_n מתחלפת עם כל מטריצה מסדר $n \times n$.

6.1.4 שחלוף מטריצות

ניתן לעשות המרה בין שורות לעמודות. פעולה זו נקראת שחלוף.

הגדרה 6.38. תהי $A \in \mathbb{M}_{m\times n}(\mathbb{F})$. נגדיר את המטריצה המשוחלפת $A^t \in \mathbb{M}_{n\times m}(\mathbb{F})$ ע"י

$$(A^t)_{ij} = A_{ji}$$

לכל $1 \le i \le n, 1 \le j \le m$.

הערה 6.39. A^t זה סימון מיוחד, לא סימון חזקה. כאן t זה מלשון transpose (שחלוף באנגלית).

דוגמה 6.40.

א. עבור וקטור שורה $v = (\alpha_1, \alpha_2, ..., \alpha_n)$ נקבל

$$v^t = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

ב. עבור

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

נקבל

$$A^t = \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

ג. עבור

$$B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$$

נקבל

$$B^t = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}.$$

טענה 6.41. *לכל* $A, B \in \mathbb{M}_{m\times n}(\mathbb{F}), C \in \mathbb{M}_{n\times k}(\mathbb{F})$ *ולכל* $\alpha \in \mathbb{F}$ *מתקיים:*

א.

$$(A^t)^t = A$$

ב.

$$(A+B)^t = A^t + B^t$$

ג.

$$(\alpha A)^t = \alpha A^t$$

ד.

$$(AC)^t = C^t A^t$$

הוכחה. נראה שוויון בין האיברים הכלליים של המטריצות בשני האגפים של כל סעיף.

א. שתי המטריצות מסדר $m \times n$ ולכל $1 \leq i \leq m, 1 \leq j \leq n$ מתקיים

$((A^t)^t)_{ij} = (A^t)_{ji} = (A)_{ij}$.

ב. שתי המטריצות מסדר $n \times m$ ולכל $1 \leq i \leq n, 1 \leq j \leq m$ מתקיים

$(A+B)^t_{ij} = (A+B)_{ji} = (A)_{ji} + (B)_{ji} = (A^t)_{ij} + (B^t)_{ij} = (A^t + B^t)_{ij}$.

ג. שתי המטריצות מסדר $n \times m$ ולכל $1 \leq i \leq n, 1 \leq j \leq m$ מתקיים

$((\alpha A)^t)_{ij} = (\alpha A)_{ji} = \alpha (A)_{ji} = \alpha (A^t)_{ij} = (\alpha A^t)_{ij}$.

ד. כאן AC מסדר $m \times k$ ולכן $(AC)^t$ מסדר $k \times m$. באופן דומה, C^t מסדר $k \times n$ ו-A^t מסדר $n \times m$. אז גם $C^t A^t$ מסדר $k \times m$, ולכל $1 \leq i \leq k, 1 \leq j \leq m$ מתקיים:

$$((AC)^t)_{ij} = (AC)_{ji} = \sum_{l=1}^{n} (A)_{jl}(C)_{li} = \sum_{l=1}^{n} (A^t)_{lj}(C^t)_{il} = \sum_{l=1}^{n} (C^t)_{il}(A^t)_{lj}$$
$$= (C^t A^t)_{ij}$$

לקראת הסוף השתמשנו בחוק החילוף כדי להתאים את סדר הכפל בתוך הסכום לתבנית של $C^t A^t$.

□

הערה 6.42. בתנאי הטענה, המטריצה A^tC^t מוגדרת אם ורק אם מספר העמודות של A^t שווה למספר השורות של C^t, כלומר אם ורק אם $m = k$. במקרה זה הסדר של A^tC^t הוא $n \times n$. לעומת זאת, הסדר של $(AC)^t$ הוא $k \times m$, או $m \times m$ תחת ההנחה ש- A^tC^t אכן מוגדרת. גם במקרה $m = k = n$ אין הכרח שיתקיים $A^tC^t = C^tA^t$. ולסיכום: אין הכרח שיתקיים $(AC)^t = A^tC^t$.

דוגמה 6.43. ניקח

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \\ -1 & 2 \\ 0 & 3 \end{pmatrix}.$$

נחשב גם את $(AC)^t$ וגם את C^tA^t ונראה שאכן יש שוויון (כצפוי). ראשית, נחשב את AC:

$$AC = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 2 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} -1 & 13 \\ -1 & 28 \end{pmatrix}$$

לכן

$$(AC)^t = \begin{pmatrix} -1 & -1 \\ 13 & 28 \end{pmatrix}.$$

נעבור לחישוב C^tA^t:

$$C^tA^t = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 13 & 28 \end{pmatrix}$$

אז אכן קיבלנו $(CA)^t = C^tA^t$.

תרגיל 6.44. עבור

$$A = \begin{pmatrix} 1+i & 2 \\ 3 & 4-i \end{pmatrix}, \quad B = \begin{pmatrix} 2 & i \\ 1-i & 3 \end{pmatrix}$$

הראו ע"י חישוב ישיר כי $(AB)^t = B^tA^t \neq A^tB^t$.

פתרון. נחשב את AB ונקבל

$$AB = \begin{pmatrix} 4 & 5+i \\ 9-5i & 12 \end{pmatrix}.$$

לכן

$$.(AB)^t = \begin{pmatrix} 4 & 9-5i \\ 5+i & 12 \end{pmatrix}$$

נעבור לחישוב השני:

$$B^tA^t = \begin{pmatrix} 2 & 1-i \\ i & 3 \end{pmatrix} \begin{pmatrix} 1+i & 3 \\ 2 & 4-i \end{pmatrix} = \begin{pmatrix} 4 & 9-5i \\ 5+i & 12 \end{pmatrix}$$

אז אכן יש שוויון. לעומת זאת:

$$A^tB^t = \begin{pmatrix} 1+i & 3 \\ 2 & 4-i \end{pmatrix} \begin{pmatrix} 2 & 1-i \\ i & 3 \end{pmatrix} = \begin{pmatrix} 2+5i & 11 \\ 5+4i & 14-5i \end{pmatrix} \neq B^tA^t$$

6.2 מטריצות מיוחדות

נרצה להגדיר קבוצות של מטריצות מיוחדות, שיופיעו בהמשך הקורס.

6.2.1 מטריצות ריבועיות

מטריצות ריבועיות הן סוג נפוץ של מטריצות שניתקל בהן לאורך הקורס. כשמן כן הן - מטריצות שנראות כמו ריבוע עם מספרים בתוכו. הרבה מטריצות בקורס יהיו ריבועיות.

הגדרה 6.45. מטריצה A נקראת ריבועית אם מספר השורות שלה שווה למספר העמודות, כלומר אם קיים $n \in \mathbb{N}$ עבורו $A \in \mathbb{M}_{n\times n}(\mathbb{F})$.

דוגמה 6.46. המטריצות הבאות הן ריבועיות:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad \begin{pmatrix} i & 2 & 3 \\ 4 & 0 & 6 \\ 7 & 8 & i \end{pmatrix}$$

המטריצות הבאות אינן ריבועיות:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 \\ -2 & -1 \\ 0 & 5 \\ 1 & 10 \end{pmatrix}, \quad \begin{pmatrix} i \\ 1 \\ 0 \end{pmatrix}$$

הערה 6.47. אם $A, B \in \mathbb{M}_{n\times n}(\mathbb{F})$ אז גם $A\pm B, AB, BA \in \mathbb{M}_{n\times n}(\mathbb{F})$. במילים: הקבוצה $\mathbb{M}_{n\times n}(F)$ סגורה לחיבור, חיסור וכפל.

הגדרה 6.48. תהי $A \in \mathbb{M}_{n\times n}(\mathbb{F})$. אז נגדיר

$$A^0 = I_n, \quad A^1 = A, \quad A^2 = A\cdot A, \quad A^3 = A\cdot A\cdot A$$

ולכל $k \in \mathbb{N}$

$$.A^k = \underbrace{A\cdot A\cdot \ldots \cdot A}_{k \text{ פעמים}}$$

עבור פולינום

$$p(x) = a_m x^m + a_{m-1}x^{m-1} + \ldots + a_1 x + a_0$$

נגדיר את ההצבה של A בו ע"י

$$.p(A) = a_m A^m + a_{m-1}A^{m-1} + \ldots + a_1 A + a_0 I_n$$

דוגמה 6.49. נחשב את $p(A)$ עבור

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

והפולינום

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

לצורך כך צריך לחשב את החזקות A^3, A^2:

$$\begin{aligned} A^2 &= \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2 \\ \implies A^3 &= A^2\cdot A = I_2 A = A \\ \implies p(A) &= A^3 + 2A^2 + A + I_2 = 2A + 3I_2 = 2\begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} + 3\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 5 & 4 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

הערה 6.50. הדוגמה האחרונה גם מראה שייתכן $A^2 = I_n$ גם אם $A \neq \pm I_n$. אמנם לסקלר יש שני שורשים (או אחד במקרה של 0), אבל למטריצת יחידה מסדר לפחות 2 יש אינסוף שורשים. בקורס שלנו אסור להוציא שורש של מטריצה כי זו לא פעולה שנגדיר (לפעמים יש אינסוף שורשים, לפעמים יש מספר סופי כלשהו של שורשים, ולפעמים אין שורש).

באופן דומה, מתקיים $B^2 = 0$ עבור

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

אז לא רק שקיימת למטריצת האפס $\mathbf{0}_{2\times 2}$ שורש שאינו עצמה, זה גם מראה שכלל הצמצום למטריצות לא מתקיים באופן כללי.

$$B \cdot B = B \cdot 0 \not\Rightarrow B = 0$$

לא ניתן "לחלק" במטריצה, אלא אם כן היא מסוג מסוים שעוד לא הגדרנו (מטריצה הפיכה). זה לא המקרה פה.

6.2.2 מטריצות סימטריות ואנטי-סימטריות

כאן הכוונה לסימטריה בין שורות לעמודות. לכן יש קשר לפעולת השחלוף.

הגדרה 6.51. תהי $A \in \mathbb{M}_{n\times n}(\mathbb{F})$ מטריצה ריבועית.

א. A נקראת סימטרית אם מתקיים $A^t = A$, או באופן שקול:

$$\forall 1 \leq i,j \leq n \quad (A)_{ji} = (A)_{ij}$$

ב. נקראת אנטי-סימטרית אם מתקיים $A^t = -A$, או באופן שקול:

$$\forall 1 \leq i,j \leq n \quad (A)_{ji} = -(A)_{ij}$$

דוגמה 6.52.

א. המטריצות הבאות הן סימטריות:

$$\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & i & 0 \\ i & 3 & -1 \\ 0 & -1 & 5 \end{pmatrix}$$

ב. המטריצות הבאות הן אנטי-סימטריות:

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & i & -2 \\ -i & 0 & -1 \\ 2 & 1 & 0 \end{pmatrix}$$

טענה 6.53. *תהי $P \in \mathbb{M}_{n\times n}(\mathbb{F})$ מטריצה ריבועית.*

א. אם P אנטי-סימטרית, אז כל איבריה על האלכסון הראשי שווים ל- 0.

ב. אם P סימטרית וגם אנטי-סימטרית, אז $P = 0$.

ג. לכל $B \in \mathbb{M}_{n\times n}(\mathbb{F})$ קיימות $S \in \mathbb{M}_{n\times n}(\mathbb{F})$ סימטרית ו- $A \in \mathbb{M}_{n\times n}(\mathbb{F})$ אנטי-סימטרית כך שמתקיים $B = S + A$.

הוכחה.

א.

$$\forall 1 \le i \le n \quad (P)_{ii} = -(P)_{ii} \implies (P)_{ii} = 0$$

כלומר יש רק אפסים על האלכסון הראשי של P.

ב. מצד אחד $P^t = P$ ומצד שני $P^t = -P$. לכן

$$P = -P \implies P + P = -P + P \implies 2P = 0 \implies P = 0$$

למעשה, השתמשנו באלגברה רגילה כדי לקבל $P = 0$ כמו עבור משוואה במשתנה סקלרי. מותר לחלק בסקלר 2, שזה כמו להכפיל בסקלר $\frac{1}{2}$.

ג. ניקח

$$S = \frac{1}{2}(B + B^t), \quad A = \frac{1}{2}(B - B^t).$$

אז לפי טענה **6.41** מתקיים

$$\begin{aligned} S^t &= \frac{1}{2}(B + B^t)^t = \frac{1}{2}(B^t + B) = S \\ A^t &= \frac{1}{2}(B - B^t)^t = \frac{1}{2}(B^t - B) = -A \\ B &= \frac{1}{2}(B + B^t) + \frac{1}{2}(B - B^t) = S + A \end{aligned}$$

לכן S סימטרית ו- A אנטי-סימטרית כנדרש.

$\square$

דוגמה 6.54. נראה כי מכפלה של מטריצות סימטריות A,B אינה בהכרח סימטרית. אם היינו מנסים להוכיח שהמכפלה היא סימטרית, היינו נתקעים:

$$(AB)^t = B^tA^t = BA$$

הבעיה היא שלא בהכרח מתקיים $BA = AB$. אז בתור דוגמה נגדית נמצא שתי מטריצות סימטריות מסדר 2×2 שאינן מתחלפות. למשל:

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

נקבל

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

וזו לא מטריצה סימטרית.

תרגיל 6.55. הסבירו מדוע כל מטריצה אנטי-סימטרית $A \in \mathbb{M}_{2\times 2}(\mathbb{F})$ היא מהצורה

$$\begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix}$$

עבור $a \in \mathbb{F}$. הסיקו שלכל שתי מטריצות אנטי-סימטריות $A,B \in \mathbb{M}_{2\times 2}(\mathbb{F})$ שאינן $\mathbf{0}_{2\times 2}$ מתקיים $AB = BA$ ומכפלה זו אינה אנטי-סימטרית.

פתרון. האיברים על האלכסון הראשי של כל מטריצה אנטי-סימטרית הם 0. נסמן $(A)_{12} = a$ ונקבל $(A)_{21} = -a$. לכן

$$A = \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix},$$

ובאופן דומה קיים $b \in \mathbb{F}$ כך שמתקיים

$$B = \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix}.$$

נחשב מכפלות:

$$AB = \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix} \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix} = \begin{pmatrix} -ab & 0 \\ 0 & -ab \end{pmatrix} = \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix} \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix} = BA.$$

זו לא מטריצה אנטי-סימטרית כי $ab \neq 0$ לפי ההנחה $A,B \neq \mathbf{0}$.

תרגיל 6.56.

א. לכל $A, B \in \mathbb{M}_{n\times n}(\mathbb{F})$ סימטריות ולכל $\alpha \in \mathbb{F}$ מתקיים שהמטריצות $A + B, \alpha A$ הן סימטריות.

ב. לכל $A, B \in \mathbb{M}_{n\times n}(\mathbb{F})$ אנטי-סימטריות ולכל $\alpha \in \mathbb{F}$ מתקיים שהמטריצות $A+B, \alpha A$ הן אנטי-סימטריות.

פתרון. א. לכל $A, B \in \mathbb{M}_{n\times n}(\mathbb{F})$ סימטריות ולכל $\alpha \in \mathbb{F}$ נקבל לפי טענה 6.41

$$(A + B)^t = A^t + B^t = A + B$$
$$(\alpha A)^t = \alpha A^t = \alpha A$$

ולכן אלו מטריצות סימטריות.

ב. הפעם נקבל

$$(A + B)^t = A^t + B^t = -A - B = -(A + B)$$
$$(\alpha A)^t = \alpha A^t = \alpha(-A) = -(\alpha A)$$

ולכן אלו מטריצות אנטי-סימטריות.

6.2.3 מטריצות אלכסוניות ומטריצות משולשיות

הגדרה 6.57. מטריצה $A \in \mathbb{M}_{n\times n}(\mathbb{F})$ נקראת:

א. אלכסונית אם לכל $i \neq j$ מתקיים $(A)_{ij} = 0$, כלומר כל האיברים מחוץ לאלכסון הראשי שווים ל- 0.

ב. משולשית עליונה אם לכל $i > j$ מתקיים $(A)_{ij} = 0$, כלומר כל האיברים מתחת לאלכסון הראשי שווים ל- 0.

ג. משולשית תחתונה אם לכל $i < j$ מתקיים $(A)_{ij} = 0$, כלומר כל האיברים מעל לאלכסון הראשי שווים ל- 0.

דוגמה 6.58.

א. המטריצות הבאות הן אלכסוניות:

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & i \end{pmatrix}, \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

ב. המטריצות הבאות הן משולשיות עליונות:

$$\begin{pmatrix}1 & 3\\ 0 & 2\end{pmatrix},\begin{pmatrix}-1 & -i & 3\\ 0 & 2 & 7\\ 0 & 0 & i\end{pmatrix},\begin{pmatrix}4 & 5 & 20 & 0\\ 0 & 6 & 7 & 13\\ 0 & 0 & 8 & 5\\ 0 & 0 & 0 & 0\end{pmatrix}$$

ג. המטריצות הבאות הן משולשיות תחתונות:

$$\begin{pmatrix}1 & 0\\ -4 & 2\end{pmatrix},\begin{pmatrix}-1 & 0 & 0\\ 1 & 2 & 0\\ 3 & 2i & i\end{pmatrix},\begin{pmatrix}4 & 0 & 0 & 0\\ -2 & 6 & 0 & 0\\ 3 & 7 & 8 & 0\\ 4 & -5 & 6 & 0\end{pmatrix}$$

הערה 6.59. נשים לב שלפי ההגדרות, לכל A ריבועית מתקיים:

א. A משולשית עליונה וגם משולשית תחתונה $\iff$ A אלכסונית.

ב. A משולשית עליונה $\iff$ A^t משולשית תחתונה.

ג. A משולשית תחתונה $\iff$ A^t משולשית עליונה.

טענה 6.60. *יהיו $A,B\in\mathbb{M}_{n\times n}(\mathbb{F})$.*

א. אם A,B משולשיות עליונות, אז AB גם משולשית עליונה.

ב. אם A,B משולשיות תחתונות, אז AB גם משולשית תחתונה.

ג. אם A,B אלכסוניות, אז AB גם אלכסונית.

הוכחה.

א. לכל $j<i$ נסתכל על האיבר הכללי של המכפלה:

$$(AB)_{ij}=\sum_{l=1}^{n}(A)_{il}(B)_{lj}$$

עבור $l<i$ מתקיים $(A)_{il}=0$ כי A משולשית עליונה. אז $i-1$ המכפלות הראשונות בסכום מתאפסות. עבור $j<i\geq l$ מתקיים $(B)_{lj}=0$ כי B משולשית עליונה. אז גם המכפלות הנותרות בסכום מתאפסות, ולכן $(AB)_{ij}=0$ כנדרש.

ב. אפשר לחקות את ההוכחה של הסעיף הקודם, אבל יש קיצור דרך: אם A, B משולשיות תחתונות, אז A^t, B^t משולשיות עליונות. לפי הסעיף הקודם, נובע כי $B^tA^t = (AB)^t$ משולשית עליונה ולכן AB משולשית תחתונה.

ג. שוב אפשר להשתמש בסעיפים הקודמים. אם A, B אלכסוניות, אז הן גם משולשיות עליונות וגם משולשיות תחתונות. לכן, לפי הסעיפים הקודמים AB גם משולשית עליונה וגם משולשית תחתונה, כלומר היא אלכסונית.

$\square$

הערה 6.61. לאחר שהבנו שהמכפלה של שתי מטריצות אלכסוניות (מאותו הסדר) היא בהכרח אלכסונית, קל להשתכנע שהנוסחה הבאה נכונה:

נסמן

$$B = \begin{pmatrix} b_1 & 0 & 0 & \cdots & 0 \\ 0 & b_2 & 0 & \cdots & 0 \\ 0 & 0 & b_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & b_n \end{pmatrix}, \quad A = \begin{pmatrix} a_1 & 0 & 0 & \cdots & 0 \\ 0 & a_2 & 0 & \cdots & 0 \\ 0 & 0 & a_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_n \end{pmatrix}.$$

אז לפי חישוב איברי האלכסון הראשי של המכפלה, נקבל

$$AB = \begin{pmatrix} a_1b_1 & 0 & 0 & \cdots & 0 \\ 0 & a_2b_2 & 0 & \cdots & 0 \\ 0 & 0 & a_3b_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_nb_n \end{pmatrix}.$$

בפרט, אם $B = A$ נקבל

$$A^2 = \begin{pmatrix} a_1^2 & 0 & 0 & \cdots & 0 \\ 0 & a_2^2 & 0 & \cdots & 0 \\ 0 & 0 & a_3^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_n^2 \end{pmatrix}.$$

נכפיל שוב ב- A ונקבל:

$$.A^3 = \begin{pmatrix} a_1^3 & 0 & 0 & \cdots & 0 \\ 0 & a_2^3 & 0 & \cdots & 0 \\ 0 & 0 & a_3^3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_n^3 \end{pmatrix}$$

ובהכללה (או באינדוקציה למי שמכיר) לכל $k \in \mathbb{N}$ נובע כי

$$.A^k = \begin{pmatrix} a_1^k & 0 & 0 & \cdots & 0 \\ 0 & a_2^k & 0 & \cdots & 0 \\ 0 & 0 & a_3^k & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_n^k \end{pmatrix}$$

אז קל יחסית לחשב חזקות של מטריצה אלכסונית, בניגוד למטריצה כללית עבורה יש הרבה יותר חישובים (בוודאי אם n והמעריך k הם מספרים גדולים).

הגדרה 6.62. מטריצה $A \in \mathbb{M}_{n\times n}(\mathbb{F})$ נקראת סקלרית אם קיים $\alpha \in \mathbb{F}$ כך ש- $A = \alpha I_n$.

תרגיל 6.63. הראו כי לכל מטריצה סקלרית $A = \alpha I_n$ ולכל מטריצה $B \in \mathbb{M}_{n\times n}(\mathbb{F})$ מתקיים $AB = BA = \alpha B$.

פתרון. יהיו A, B כנ"ל. אז מתקיים

$$,AB = (\alpha I_n)B = \alpha(I_n B) = \alpha B$$

ובנוסף

$$.BA = B(\alpha I_n) = \alpha(B I_n) = \alpha B$$

לכן יש שוויון.

6.2.4 מטריצות אלמנטריות

כאן יש קשר הדוק לפעולות שורה אלמנטריות. נשתמש בראשי התיבות פש"א עבור "פעולת שורה אלמנטרית".

הגדרה 6.64. מטריצה $A \in \mathbb{M}_{n\times n}(\mathbb{F})$ נקראת אלמנטרית אם היא מתקבלת ממטריצת היחידה I_n ע"י פש"א אחת.

עבור פש"א S, נסמן ב- $S(I_n)$ את המטריצה האלמנטרית המתאימה שמתקבלת מ- I_n ע"י S. באופן כללי, לכל מטריצה A נסמן ב- $S(A)$ את המטריצה שמתקבלת מ- A ע"י S.

דוגמה 6.65.

א. המטריצות הבאות הן אלמנטריות כי הן מתאימות להחלפה בין שורות של מטריצת יחידה:

$$S = (R_1 \leftrightarrow R_2) \implies S(I_2) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$
$$S = (R_1 \leftrightarrow R_3) \implies S(I_3) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$
$$S = (R_2 \leftrightarrow R_4) \implies S(I_4) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

ב. המטריצות הבאות הן אלמנטריות כי הן מתאימות לכפל בסקלר של שורה אחת של מטריצת יחידה:

$$S = (2R_1 \to R_1) \implies S(I_2) = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$
$$S = (3R_2 \to R_2) \implies S(I_3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$S = (4R_3 \to R_3) \implies S(I_4) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

ג. המטריצות הבאות הן אלמנטריות כי הן מתאימות להוספת כפולה בסקלר של שורה אחת לשורה אחרת במטריצת יחידה:

$$S = (R_2 + 2R_1 \to R_2) \implies S(I_2) = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$$

$$S = (R_2 + 3R_1 \to R_2) \implies S(I_3) = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$S = (R_3 + 4R_2 \to R_3) \implies S(I_4) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 4 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

ד. המטריצות הבאות אינן אלמנטריות כי הן מתקבלות ממטריצת יחידה ע"י יותר מפש"א אחת:

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

למעשה, במטריצה אלמנטרית רק שורה אחת שונה מהשורה המתאימה ב- I_n, אלא אם כן זו מטריצה שמתאימה להחלפת שורות. אבל גם אז יש הבדל רק בשתי שורות וקל לזהות את ההחלפה, בניגוד למטריצה השמאלית לעיל שהיא בפירוש לא קשורה להחלפה.

טענה 6.66. *יהיו $A \in \mathbb{M}_{m\times n}(\mathbb{F})$ ו- S פש"א. אז מתקיים $S(I_m)A = S(A)$.*

לפני החישובים הכלליים הדרושים להוכחה, נסתכל על דוגמאות:

דוגמה 6.67.

א.

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & r \end{pmatrix} = \begin{pmatrix} g & h & r \\ d & e & f \\ a & b & c \end{pmatrix}$$

ב.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & r \end{pmatrix} = \begin{pmatrix} a & b & c \\ 3d & 3e & 3f \\ g & h & r \end{pmatrix}$$

ג.

$$\begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & r \end{pmatrix} = \begin{pmatrix} a & b & c \\ 3a+d & 3b+e & 3c+f \\ g & h & r \end{pmatrix}$$

הוכחה. נסמן $E = S(I_m)$ עם איברים $e_{ij} = (E)_{ij}$. נראה ש- EA מתקבלת מ- A בדיוק ע"י הפעלת S על שורות A. נחלק לשלושה מקרים:

א. $S = (R_i \leftrightarrow R_j)$

E מזכירה את מטריצת היחידה, אבל יש ארבעה הבדלים:

$$e_{ii} = e_{jj} = 0, \quad e_{ij} = e_{ji} = 1$$

נחשב:

$$(EA)_{rk} = \sum_{l=1}^{m} e_{rl} a_{lk}.$$

אם $r \notin \{i, j\}$, אז e_{rk} הוא בדיוק כמו במטריצת היחידה: 1 עבור $r = k$, אחרת 0. לכן השורה ה- r נשארת כפי שהיא. לעומת זאת, אם $r = i$ נקבל $(EA)_{ik} = a_{jk}$. אם $r = j$ נקבל $(EA)_{jk} = a_{ik}$. כלומר החלפנו בין השורות ה- i, j של A, ולכן קיבלנו את $S(A)$.

ב. $S = (\alpha R_i \to R_i)$ עבור $\alpha \neq 0$.

כאן E היא כמעט כמו מטריצת היחידה, עם הבדל אחד: $e_{ii} = \alpha$.

נשתמש בנוסחה מפוצלת כי יש פה שני מקרים:

$$(EA)_{rk} = \sum_{l=1}^{m} e_{rl} a_{lk} = \begin{cases} a_{rk}, & r \neq i \\ \alpha a_{ik}, & r = i \end{cases}$$

כלומר השורה ה- i של A הוכפלה ב- α ושאר השורות לא השתנו. קיבלנו את $S(A)$.

ג. $S = (R_i + \alpha R_j \to R_i)$ עבור $i \neq j$. כאן

$$E = I_m + \alpha P(i, j),$$

כאשר $P(i, j)$ היא המטריצה שכל ערכיה הם 0 פרט ל- $(P(i, j))_{ij} = 1$. נחשב:

$$EA = (I_m + \alpha P(i, j))A = A + \alpha(P(i, j)A).$$

אבל

$$(P(i, j)A)_{rk} = \sum_{l=1}^{m} (P(i, j))_{rl} a_{lk} = \begin{cases} 0, & r \neq i \\ a_{jk}, & r = i \end{cases}$$

ולכן

$$(EA)_{rk} = \begin{cases} a_{rk}, & r \neq i \\ a_{ik} + \alpha a_{jk}, & r = i \end{cases}$$

כלומר השורה R_i הוחלפה ב- $R_i + \alpha R_j$. שוב קיבלנו את $S(A)$.

$\square$

מסקנה 6.68. *יהיו* $A, A' \in \mathbb{M}_{m\times n}(\mathbb{F})$ *שקולות שורה. אז קיימת* $B \in \mathbb{M}_{m\times m}(\mathbb{F})$ *כך ש-* $A' = BA$.

הוכחה. לפי הנתון קיימות פש"אות $S_1, S_2, ..., S_k$ כך שמתקיים $A' = S_k(...(S_1(A))...)$, כלומר A' מתקבלת מ- A בעקבות הפעלת S_1, אחריה S_2 וכן הלאה עד S_k.

לפי טענה 6.66, ניתן לפתוח את הסוגריים בהדרגה כאשר בכל שלב הפעולה הרלוונטית הופכת לכפל במטריצה האלמנטרית המתאימה לה:

$$\begin{aligned} A' &= S_k(I_m)S_{k-1}(...(S_1(A))...) = S_k(I_m)S_{k-1}(I_m)S_{k-2}(...(S_1(A))...) = ... \\ &= S_k(I_m)S_{k-1}(I_m)S_{k-2}(I_m)\cdots S_1(I_m)A \end{aligned}$$

נסמן $B = S_k(I_m)S_{k-1}(I_m)S_{k-2}(I_m)\cdots S_1(I_m)$. זו אכן מטריצה מסדר $m \times m$ וקיבלנו $A' = BA$. $\square$

דוגמה 6.69. נסתכל על המטריצה

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

נבצע דירוג קנוני, אבל הפעם נציין בכל שלב את המטריצה האלמנטרית המתאימה E_i ונעקוב אחר המכפלה המצטברת.

שלב 1: $R_2 - 2R_1 \to R_2$

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad E_1A = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & 0 & -1 & -4 \\ 1 & 3 & 4 & 0 \end{pmatrix}$$

שלב 2: $R_3 - R_1 \to R_3$

$$E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \qquad E_2E_1A = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & 0 & -1 & -4 \\ 0 & 1 & 1 & -1 \end{pmatrix}$$

שלב 3: $R_3 \leftrightarrow R_2$

$$E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad E_3E_2E_1A = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & -1 & -4 \end{pmatrix}$$

שלב 4: $-R_3 \to R_3$

$$E_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad E_4E_3E_2E_1A = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 4 \end{pmatrix}$$

שלב 5: $R_2 - R_3 \to R_2$

$$E_5 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}, \qquad E_5E_4E_3E_2E_1A = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 1 & 4 \end{pmatrix}$$

שלב 6: $R_1 - 3R_3 \to R_1$

$$E_6 = \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad E_6E_5E_4E_3E_2E_1A = \begin{pmatrix} 1 & 2 & 0 & -11 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 1 & 4 \end{pmatrix}$$

שלב 7: $R_1 - 2R_2 \to R_1$

$$E_7 = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad E_7E_6E_5E_4E_3E_2E_1A = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 1 & 4 \end{pmatrix}$$

המטריצה האחרונה שקיבלנו בצד ימין היא מדורגת קנונית:

$$A' = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 1 & 4 \end{pmatrix}$$

המכפלה הכוללת של המטריצות האלמנטריות היא $E = E_7E_6E_5E_4E_3E_2E_1$, ובמפורש:

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

בדיקה:

$$EA = \begin{pmatrix} 1 & 1 & -2 \\ -3 & 1 & 1 \\ 2 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 1 \\ 2 & 4 & 5 & -2 \\ 1 & 3 & 4 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 1 & 4 \end{pmatrix} = A'.$$

6.3 מטריצות הפיכות והקשר לממ"ל

לפני שנמשיך, יש סימון נוח לוקטורים שיהיו שימושיים בקורס:

הגדרה 6.70. נגדיר את וקטורי היחידה $e_1, e_2, \cdots, e_n \in \mathbb{F}^n$ ע"י

$$e_i = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

כאשר כל הקואורדינטות הן 0 חוץ מהקואורדינטה ה- i השווה ל- 1. נציין שאלה וקטורי העמודה של I_n, וגם וקטורי השורה עד כדי שחלוף (ההבדל בין שורה לעמודה).

תרגיל 6.71. לכל

$$x = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

מתקיים

$$.x = \sum_{j=1}^{n} a_j e_j$$

פתרון. בסכום $\sum_{j=1}^{n} a_j e_j$ רק המחובר ה- i תורם לקואורדינטה ה- i, כי כל השאר תורמים 0. התרומה עבור $j = i$ היא $a_i \cdot 1 = a_i$ כנדרש.

לפני שנגדיר מטריצות הפיכות, נוכיח טענה שמראה את חשיבותה של מטריצת היחידה בהקשר של קיום ויחידות הפתרון לממ"ל כאשר מספר המשתנים שווה למספר המשוואות.

טענה 6.72. *אם $A' \in \mathbb{M}_{n\times n}(\mathbb{F})$ מדורגת קנונית, אז התנאים הבאים שקולים:*

(א) $A' = I_n$

(ב) $\mathrm{rank}(A') = n$

(ג) לממ"ל $A'x=0$ יש פתרון יחיד.

(ד) לכל $b'\in\mathbb{F}^n$ יש פתרון יחיד לממ"ל $A'x=b'$.

(ה) לכל $b'\in\mathbb{F}^n$ יש פתרון לממ"ל $A'x=b'$.

הוכחה. נוכיח כי (א) שקול לכל אחד מהתנאים האחרים. תחילה נראה שהוא גורר את האחרים:
(א) $\Longleftarrow$ (ב): למטריצה I_n יש דרגה n כמספר האיברים המובילים שלה.
(א) $\Longleftarrow$ (ג): ברור שיש פתרון יחיד, שהוא הפתרון הטריוויאלי.
(א) $\Longleftarrow$ (ד): ברור שיש פתרון יחיד, שהוא $x=b'$.
(א) $\Longleftarrow$ (ה): כנ"ל.
כעת נוכיח את הכיוון ההפוך של כל גרירה.
(א) $\Longrightarrow$ (ב): אם הדרגה היא n, אז יש 1 מוביל בכל עמודה. כל שאר האיברים הם 0 (כי הדירוג קנוני) ולכן $A'=I_n$.
(א) $\Longrightarrow$ (ג): מיחידות הפתרון נובע (ב) ומכאן גם (א).
(א) $\Longrightarrow$ (ד): כנ"ל.
(א) $\Longrightarrow$ (ה): אם תמיד קיים פתרון לממ"ל $A'x=b$, אז אין שורת אפסים ב- A' כי אחרת מתקבלת שורה סתירה ב- $(A'|b')$ עבור

$$.b'=\begin{pmatrix}0\\0\\\vdots\\0\\1\end{pmatrix}$$

לכן נובע (ב) ומכאן (א).

□

כמסקנה מהטענה הקודמת למטריצות ריבועיות שהן מדורגות קנונית, יש לנו משפט למטריצות ריבועיות כלליות:

טענה 6.73. *לכל $A\in\mathbb{M}_{n\times n}(\mathbb{F})$ התנאים הבאים שקולים:*

(א) A שקולת שורה ל- I_n.

(ב) $\operatorname{rank}(A)=n$

(ג) לממ"ל $Ax=0$ יש פתרון יחיד.

(ד) לכל $b\in\mathbb{F}^n$ יש פתרון יחיד לממ"ל $Ax=b$.

(ה) לכל $b \in \mathbb{F}^n$ *יש פתרון לממ"ל* $Ax = b$.

הוכחה. תהי $A \in \mathbb{M}_{n\times n}(\mathbb{F})$. היא שקולת שורה למטריצה מתאימה A' שהיא מדורגת קנונית, ונשים לב כי כל אחד מהתנאים (ב), (ג), (ד), (ה) שקול לתנאי המקביל לו בטענה **6.72** עם A' במקום A. הסיבה לכך היא שמתקיים $\operatorname{rank}(A) = \operatorname{rank}(A')$ לפי הגדרת דרגה, וגם קבוצת הפתרונות של הממ"ל $Ax = b$ שווה לקבוצת הפתרונות של הממ"ל $A'x = b'$. לכן, כל אחד מהתנאים (ב), (ג), (ד), (ה) שקול לתנאי $A' = I_n$, שזה בדיוק (א).

□

משפט זה מראה שיש משהו מיוחד לגבי מטריצות ריבועיות שהן שקולות שורה למטריצת יחידה. נראה שיש דרך חלופית להגדיר אותן, שקשורה לכפל מטריצות. בהתחלה ההגדרה תיראה שונה, ובסוף נוכיח שזו הגדרה שקולה.

הגדרה 6.74. מטריצה $A \in \mathbb{M}_{n\times n}(\mathbb{F})$ נקראת הפיכה משמאל אם קיימת $B \in \mathbb{M}_{n\times n}(\mathbb{F})$ כך ש- $BA = I_n$. A נקראת הפיכה מימין אם קיימת $C \in \mathbb{M}_{n\times n}(\mathbb{F})$ כך ש- $AC = I_n$. אם שני התנאים מתקיימים, A נקראת הפיכה.

נראה שהפיכות מימין והפיכות משמאל, הן למעשה שני תנאים שקולים (עבור מטריצה ריבועית). יותר מכך:

משפט 6.75. *לכל* $A \in \mathbb{M}_{n\times n}(\mathbb{F})$ *התנאים הבאים שקולים:*

(א) A *שקולת שורה ל-* I_n.

(ב) $\operatorname{rank}(A) = n$

(ג) לממ"ל $Ax = 0$ *יש פתרון יחיד.*

(ד) לכל $b \in \mathbb{F}^n$ *יש פתרון יחיד לממ"ל* $Ax = b$.

(ה) לכל $b \in \mathbb{F}^n$ *יש פתרון לממ"ל* $Ax = b$.

(ו) A *הפיכה משמאל.*

(ז) A *הפיכה מימין.*

(ח) A *הפיכה.*

הוכחה. כבר הוכחנו שחמשת התנאים הראשונים שקולים. נראה כי

$$(\text{א}) \Longrightarrow (\text{ו}) \Longrightarrow (\text{ג})$$

וגם

$$(\text{ה}) \Longleftrightarrow (\text{ז})$$

ומכאן נסיק שכל אחד מהתנאים (ו), (ז), (ח) שקול לכל שאר התנאים.

(א) $\Longleftarrow$ (ו) : זה נובע ישירות ממסקנה **6.68** עם $A' = I_n$.

(ו) $\Longleftarrow$ (ג) : נניח כי קיימת $B \in \mathbb{M}_{n\times n}(\mathbb{F})$ כך ש- $BA = I_n$. אז כדי לפתור את הממ"ל $Ax = 0$ נכפיל את שני האגפים ב- B משמאל:

$$B(Ax) = B \cdot 0 \implies I_n x = 0 \implies x = 0$$

לכן יש פתרון יחיד.

(ה) $\Longleftarrow$ (ז) : המטרה היא למצוא $C \in \mathbb{M}_{n\times n}(\mathbb{F})$ כך ש- $AC = I_n$. נסמן את העמודות של C ע"י $v_1, v_2, ..., v_n \in \mathbb{F}^n$, ואז לפי **6.32** מתקיים

$$.AC = \begin{pmatrix} | & | & & | \\ Av_1 & Av_2 & \cdots & Av_n \\ | & | & & | \end{pmatrix}$$

לכן

$$.AC = I_n \iff \forall 1 \le j \le n \quad Av_j = e_j$$

אז קיבלנו n ממ"ליות, ולכל אחת מהן יש פתרון v_j לפי (ה). אז המטריצה המתאימה

$$C = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix}$$

מקיימת $AC = I_n$ כנדרש.

(ז) $\Longleftarrow$ (ה) : נניח כי קיימת $C \in \mathbb{M}_{n\times n}(\mathbb{F})$ כך ש- $AC = I_n$. יהי

$$b = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}$$

ונראה שקיים פתרון לממ"ל $Ax = b$. בסימונים הקודמים לעמודות C, שוב מתקיים לכל $1 \le j \le n$

$$.Av_j = e_j$$

אז ניקח

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

עם $x = \sum_{j=1}^{n} \beta_j v_j$. לכן נקבל

$$.Ax = A(\sum_{j=1}^{n} \beta_j v_j) = \sum_{j=1}^{n} \beta_j A v_j = \sum_{j=1}^{n} \beta_j e_j = b$$

ניתן להכניס את A לתוך הסכימה לפי טענה 6.23. $\square$

לכאורה, הפיכות מימין והפיכות משמאל מתייחסות לשתי מטריצות B, C שיכולות להיות שונות, אבל ראינו כי הפיכות משמאל שקולה להפיכות מימין. אז אמור להיות קשר בין המטריצות, ואכן:

טענה 6.76. *יהיו* $A, B, C \in \mathbb{M}_{n\times n}(\mathbb{F})$ *כך ש-* $BA = I_n = AC$. *אז* $B = C$.

הוכחה. לפי התכונות של I_n וחוק הקיבוץ לכפל מטריצות, נובע כי

$$.B = BI_n = B(AC) = (BA)C = I_nC = C$$

$\square$

מסקנה 6.77. *יהיו* $A, B \in \mathbb{M}_{n\times n}(\mathbb{F})$. *אז מתקיים*

$$.AB = I_n \iff BA = I_n$$

הוכחה. אם $BA = I_n$, אז A הפיכה משמאל. לפי משפט 6.75 היא גם הפיכה מימין, ולכן קיימת $C \in \mathbb{M}_{n\times n}(\mathbb{F})$ כך ש- $AC = I_n$. אבל אז נובע מטענה 6.76 כי $B = C$, ולכן $AB = I_n$.
הכיוון השני דומה, למשל משיקולי סימטריה (משחק תפקידים בין A ל- B). $\square$

האם בהינתן מטריצה הפיכה A, המטריצה B שמקיימת $AB = I_n = BA$ נקבעת ביחידות ע"י A? התשובה חיובית. לפני שנתאר את החישוב, נוכיח את היחידות:

מסקנה 6.78. *יהיו* $A, B_1, B_2 \in \mathbb{M}_{n\times n}(\mathbb{F})$ *כך ש-* $B_1A = B_2A = I_n$. *אז* $B_1 = B_2$.

הוכחה. A הפיכה משמאל ולכן גם מימין לפי משפט 6.75. כלומר קיימת $C \in \mathbb{M}_{n\times n}(\mathbb{F})$ כך ש- $AC = I_n$. לכן, נובע $B_1 = C = B_2$ מטענה 6.76. $\square$

הגדרה 6.79. תהי $A \in \mathbb{M}_{n\times n}(\mathbb{F})$ הפיכה. נסמן ב- A^{-1} את המטריצה שמקיימת

$$.AA^{-1} = A^{-1}A = I_n$$

נקרא לה המטריצה ההופכית ל- A.

הערה 6.80. נשים לב כי A, A^{-1} מתחלפות. בהינתן B, כדי לבדוק שמתקיים $B = A^{-1}$ מספיק להראות $AB = I_n$ או לחילופין $BA = I_n$, כי התנאים שקולים.

תרגיל 6.81. יהיו $A,B\in\mathbb{M}_{n\times n}(\mathbb{F})$ הפיכות ו- $0\neq\alpha\in\mathbb{F}$. אז מתקיים:

א. A^{-1} הפיכה ומתקיים

$$(A^{-1})^{-1}=A.$$

ב. αA הפיכה ומתקיים

$$(\alpha A)^{-1}=\frac{1}{\alpha}A^{-1}.$$

ג. AB הפיכה ומתקיים

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

ד. A^t הפיכה ומתקיים

$$(A^t)^{-1}=(A^{-1})^t.$$

פתרון. בכל סעיף נחשב את אחת משתי המכפלות ונראה שהיא שווה ל- I_n.

א. בהינתן $AA^{-1}=A^{-1}A=I_n$ נובע כי גם A^{-1} הפיכה וההופכית שלה היא A.

ב. מתקיים

$$(\frac{1}{\alpha}A^{-1})(\alpha A)=(\frac{1}{\alpha}\alpha)A^{-1}A=1\cdot I_n=I_n.$$

לכן αA הפיכה ומתקיים $(\alpha A)^{-1}=\frac{1}{\alpha}A^{-1}$.

ג. מתקיים

$$(AB)(B^{-1}A^{-1})=A(BB^{-1})A^{-1}=A(I_nA^{-1})=AA^{-1}=I_n.$$

לכן AB הפיכה ומתקיים $(AB)^{-1}=B^{-1}A^{-1}$.

ד. מתקיים

$$A^t(A^{-1})^t=(A^{-1}A)^t=I_n^t=I_n.$$

לכן A^t הפיכה ומתקיים $(A^t)^{-1}=(A^{-1})^t$.

הערה 6.82. לא בהכרח מתקיים $(AB)^{-1}=A^{-1}B^{-1}$ כי אין הכרח ש- A^{-1},B^{-1} יתחלפו.

הערה 6.83. אם A הפיכה, אז הפתרון היחיד לממ"ל $Ax=b$ נתון ע"י

$$x=A^{-1}b$$

לכל $b\in\mathbb{F}^n$. אכן, ניתן להכפיל את שני אגפי המשוואה ב- A^{-1} בהינתן שהיא קיימת. לעומת זאת, אי אפשר "לחלק" במטריצה שאינה הפיכה.

6.3.1 חישוב מטריצה הופכית

הבנו שכל מטריצה $A \in \mathbb{M}_{n\times n}(\mathbb{F})$ הפיכה שקולת שורה ל- I_n. יותר מכך, לפי ההוכחה של מסקנה 6.68 יש התאמה בין A^{-1} לבין הפש"אות שמובילות ל- I_n בדירוג של A. נחדד את נקודה זו כמסקנה נוספת:

מסקנה 6.84.

א. *תהי $A \in \mathbb{M}_{n\times n}(\mathbb{F})$ הפיכה. אז מתקיים*

$$.A^{-1} = E_k E_{k-1} \cdots E_1$$

כאשר $E_1, E_2, ..., E_k$ הן המטריצות האלמנטריות המתאימות לפש"אות $S_1, S_2, ..., S_k$ בדירוג הקנוני של A.

ב. *A היא מכפלה של מטריצות אלמנטריות:*

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

הוכחה.

א. זהו מקרה פרטי של מסקנה 6.68 עבור $A' = I_n$. כאן $B = A^{-1}$.

ב. לפי תרגיל 6.81 הפעולה של הפיכת מטריצה הופכת את סדר המטריצות במכפלה. אפשר להבין זאת מחישוב הדרגתי:

$$\begin{aligned} & A^{-1} = E_k E_{k-1} \cdots E_1 \\ \Longrightarrow\ & A = ((E_k E_{k-1} \cdots E_2)E_1)^{-1} = E_1^{-1}((E_k E_{k-1} \cdots E_3)E_2)^{-1} \\ & \quad = E_1^{-1}E_2^{-1}((E_k E_{k-1} \cdots E_4)E_3)^{-1} = ... = E_1^{-1}E_2^{-1} \cdots E_k^{-1} \end{aligned}$$

לחילופין, אפשר לוודא שמתקיים

$$.(E_1^{-1}E_2^{-1} \cdots E_k^{-1})(E_k E_{k-1} \cdots E_1) = I_n$$

נדגיש שלכל מטריצה אלמנטרית, גם ההופכית שלה היא מטריצה אלמנטרית המתאימה לפש"א ההפוכה.

□

יש דרך נוחה ומסודרת לחשב את המטריצות האלמנטריות ואת מכפלתן שמובילה למטריצה ההופכית. הרעיון הוא לדרג את I_n לצד A, כמו במטריצת מקדמים מורחבת - אבל עם מספר עמודות מימין לקו המפריד. קובעים את הפש"אות לפי A, אבל מבצעים אותן גם על I_n (כלומר מכפילים גם אותה במטריצות האלמנטריות המתאימות):

$$(A \mid I_n) \to (E_1A \mid E_1) \to (E_2E_1A \mid E_2E_1) \to \cdots \to (E_k\cdots E_1A \mid E_k\cdots E_1)$$

המטריצה האחרונה בצד ימין היא למעשה $(I_n \mid A^{-1})$, כלומר המטריצה ההופכית מופיעה מימין ל- I_n בסוף הדירוג הקנוני של מטריצה הפיכה.

דוגמה 6.85.   א. ניקח את

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 2 & 3 & 4 \end{pmatrix}$$

ונחשב את ההופכית ע"י דירוג קנוני של $(A \mid I_3)$:

$$(A \mid I_3) = \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 2 & 3 & 4 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{R_3-2R_1\to R_3} \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & -1 & 2 & -2 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{R_2\leftrightarrow R_3} \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -1 & 2 & -2 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \end{array}\right)$$

$$\xrightarrow{-R_2\to R_2}\left(\begin{array}{ccc|ccc}1&2&1&1&0&0\\0&1&-2&2&0&-1\\0&1&1&0&1&0\end{array}\right)$$

$$\xrightarrow{R_1-2R_2\to R_1}\left(\begin{array}{ccc|ccc}1&0&5&-3&0&2\\0&1&-2&2&0&-1\\0&1&1&0&1&0\end{array}\right)$$

$$\xrightarrow{R_3-R_2\to R_3}\left(\begin{array}{ccc|ccc}1&0&5&-3&0&2\\0&1&-2&2&0&-1\\0&0&3&-2&1&1\end{array}\right)$$

$$\xrightarrow{\frac{1}{3}R_3\to R_3}\left(\begin{array}{ccc|ccc}1&0&5&-3&0&2\\0&1&-2&2&0&-1\\0&0&1&-\frac{2}{3}&\frac{1}{3}&\frac{1}{3}\end{array}\right)$$

$$\xrightarrow{R_1-5R_3\to R_1}\left(\begin{array}{ccc|ccc}1&0&0&\frac{1}{3}&-\frac{5}{3}&\frac{1}{3}\\0&1&-2&2&0&-1\\0&0&1&-\frac{2}{3}&\frac{1}{3}&\frac{1}{3}\end{array}\right)$$

$$\xrightarrow{R_2+2R_3\to R_2}\left(\begin{array}{ccc|ccc}1&0&0&\frac{1}{3}&-\frac{5}{3}&\frac{1}{3}\\0&1&0&\frac{2}{3}&\frac{2}{3}&-\frac{1}{3}\\0&0&1&-\frac{2}{3}&\frac{1}{3}&\frac{1}{3}\end{array}\right)$$

אז קיבלנו

$$A^{-1}=\begin{pmatrix}\frac{1}{3}&-\frac{5}{3}&\frac{1}{3}\\\frac{2}{3}&\frac{2}{3}&-\frac{1}{3}\\-\frac{2}{3}&\frac{1}{3}&\frac{1}{3}\end{pmatrix}.$$

ב. ניקח את

$$B=\begin{pmatrix}1+i&2\\3&4-i\end{pmatrix}$$

ונחשב את ההופכית ע"י דירוג קנוני של $(B \mid I_2)$:

$$
\begin{aligned}
(B \mid I_2) &= \left(\begin{array}{cc|cc} 1+i & 2 & 1 & 0 \\ 3 & 4-i & 0 & 1 \end{array}\right) \\
&\xrightarrow{R_2-\frac{3}{1+i}R_1\to R_2} \left(\begin{array}{cc|cc} 1+i & 2 & 1 & 0 \\ 0 & 4-i-\frac{6}{1+i} & -\frac{3}{1+i} & 1 \end{array}\right) \\
&\xrightarrow{\frac{1}{1+i}R_1\to R_1} \left(\begin{array}{cc|cc} 1 & \frac{2}{1+i} & \frac{1}{1+i} & 0 \\ 0 & \frac{-1+3i}{1+i} & -\frac{3}{1+i} & 1 \end{array}\right) \\
&\xrightarrow{\frac{1+i}{-1+3i}R_2\to R_2} \left(\begin{array}{cc|cc} 1 & 1-i & \frac{1-i}{2} & 0 \\ 0 & 1 & \frac{3}{1-3i} & \frac{1+i}{-1+3i} \end{array}\right) \\
&\xrightarrow{R_1-(1-i)R_2\to R_1} \left(\begin{array}{cc|cc} 1 & 0 & \frac{1-i}{2}-\frac{(3+9i)(1-i)}{10} & -\frac{(1-i)(1-2i)}{5} \\ 0 & 1 & \frac{3+9i}{10} & \frac{1-2i}{5} \end{array}\right)
\end{aligned}
$$

נפשט את המטריצה האחרונה שקיבלנו בצד ימין:

$$B^{-1} = \begin{pmatrix} -\frac{7+11i}{10} & \frac{1+3i}{5} \\ \frac{3+9i}{10} & \frac{1-2i}{5} \end{pmatrix}$$

תרגיל 6.86.

א. חשבו את A^{-1} עבור

$$.A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

ב. עבור A הקודמת, מצאו $B \in \mathbb{M}_{2\times 2}(\mathbb{R})$ עבורה מתקיים $AB = C$ כאשר

$$.C = \begin{pmatrix} 3 & -5 \\ -9 & 1 \end{pmatrix}$$

פתרון.

א. נדרג את $(A|I_2)$:

$$(A \mid I_2) = \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{R_2-3R_1\to R_2} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array}\right)$$

$$\xrightarrow{-\frac{1}{2}R_2\to R_2} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array}\right)$$

$$\xrightarrow{R_1-2R_2\to R_1} \left(\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array}\right)$$

אז קיבלנו

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

ב. נכפיל משמאל ב- A^{-1} את שני אגפי המשוואה $AB = C$ ונקבל:

$$.B = A^{-1}C = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 3 & -5 \\ -9 & 1 \end{pmatrix} = \begin{pmatrix} -15 & 11 \\ 9 & -8 \end{pmatrix}$$

פרק 7

דטרמיננטה

בפרק הקודם הגדרנו מטריצות הפיכות, וראינו שיש תנאים שונים השקולים להגדרה זו. נרצה לפתח תנאי נוסף שאפשר לבטא באמצעות פונקציה שלוקחת מטריצה ריבועית ומחזירה סקלר (ע״י חישוב שתלוי בערכי המטריצה). סקלר זה ייקרא הדטרמיננטה של המטריצה, וזה גם השם של הפונקציה עצמה.

לדטרמיננטה יש חשיבות בגיאומטריה וחדו״א לצורך חישובים נפחים ושטחים (תחילה של צורות מוכרות, ואחר כך של צורות יותר מסובכות בעזרת אינטגרל). המושג של מכפלה וקטורית, שלא ניכנס אליו בקורס, גם מבוסס על דטרמיננטה וזו עוד דרך לחשב וקטור נורמל למישור במרחב. אבל אנחנו נתמקד בהיבט העיקרי של דטרמיננטה כמדד להפיכות של המטריצה. מה המשמעות שלה במובן הזה? ככל שהדטרמיננטה של A גדולה יותר, כך הממ״ל $Ax = b$ יציבה יותר במובן הבא: הפתרון לממ״ל (הפלט $A^{-1}b$) פחות רגיש לשגיאה בקלט b. דטרמיננטה קטנה פירושה שיש רגישות גדולה לשגיאה, והמידע שמבוטא בממ״ל פחות אמין. אם הדטרמיננטה מתאפסת, אז A אינה הפיכה ומלכתחילה לא קיים פתרון יחיד לממ״ל (יש אובדן של מידע). ניגע בהיבטים המעשיים של הדטרמיננטה ביישומים בסוף הספר, אבל בקורס עצמו נתמקד בתכונות הדטרמיננטה וחישובים.

הנוסחה לדטרמיננטה היא די ארוכה למטריצות ריבועיות מסדר $n \times n$ עבור $n \geq 3$, ולא מומלץ לשנן אותה למעט המקרה $n = 2$. אז נפעל בכיוון הפוך: נקבע את התכונות הרצויות עבור הדטרמיננטה, שלמעשה מגדירות אותה. בהמשך נראה שיטות שונות לחישוב.

נציין שהסימון לדטרמיננטה של מטריצה A הוא $\det(A)$ ללא קשר לסדר המטריצה. הנוסחה עצמה תלויה בסדר.

7.1 הגדרת הדטרמיננטה והקשר לדירוג

הגדרה 7.1. עבור $A \in \mathbb{M}_{n\times n}(\mathbb{F})$ נגדיר את $\det(A)$ ע״י התכונות הבאות:

א. $\det(I_n) = 1$.

ב. אם $A' \xleftarrow{\alpha R_i \to R_i} A$ עבור $1 \leq i \leq n$, אז $\det(A') = \alpha \det(A)$.

ג. אם $A' \xleftarrow{R_i + \alpha R_j \to R_i} A$ עבור $i \neq j$, אז $\det(A') = \det(A)$.

ד. אם $A' \xleftarrow{R_i \leftrightarrow R_j} A$ עבור $i \neq j$, אז $\det(A') = -\det(A)$.

הערה 7.2. דרך שקולה להסתכל על התכונה השנייה היא שאם כבר מופיע במטריצה הנתונה כפל בסקלר של שורה כלשהי, אפשר להוציא את הסקלר מחוץ לדטרמיננטה. זה דומה להוצאת גורם משותף מחוץ לחישוב רגיל. בפרט, לכל $A = (a) \in \mathbb{M}_{1\times 1}(\mathbb{F})$ מתקיים

$$\det(A) = \det(a) = a \det(1) = a \det(I_1) = a \cdot 1 = a\,.$$

הערה 7.3. כבר אפשר לראות דרך אחת לחישוב דטרמיננטה של מטריצה הפיכה: מבצעים דירוג קנוני (שיסתיים ב-I_n) ועוקבים אחרי השינוי בערך הדטרמיננטה בעקבות ביצוע פש"אות, שהן כפל שורה בסקלר או החלפה בין שתי שורות. נניח כי $\alpha_1, \alpha_2, ..., \alpha_k \in \mathbb{F}$ הם הסקלרים בהם מכפילים את שורות המטריצה לאורך הדירוג. בנוסף, נניח כי יש p החלפות שורות. אז מתכונות הדטרמיננטה נובע כי

$$1 = \det(I_n) = (-1)^p \alpha_1 \alpha_2 \cdots \alpha_k \det(A) \implies \det(A) = \frac{(-1)^p}{\alpha_1 \alpha_2 \cdots \alpha_k}$$

החסרונות של נוסחה זו היא שהדירוג עלול להיות ארוך, ולא ברור מראש מהם הסקלרים שמופיעים במכנה. לא תמיד זו הדרך הנוחה ביותר לחישוב דטרמיננטה, אבל זו דרך שתמיד יכולה לעבוד. נרצה שיהיו לנו כלים נוספים כדי שנוכל להסתפק בדירוג חלקי אם בכלל.

דוגמה 7.4. נחשב את הדטרמיננטה של

$$A = \begin{pmatrix} 1 & -3 \\ 4 & 7 \end{pmatrix}.$$

נקבל

$$\det(A) = \det\begin{pmatrix} 1 & -3 \\ 4 & 7 \end{pmatrix} \overset{R_2 - 4R_1 \to R_2}{=} \det\begin{pmatrix} 1 & -3 \\ 0 & 19 \end{pmatrix} \overset{\frac{1}{19}R_2 \to R_2}{=} 19 \cdot \det\begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix}$$
$$\overset{R_1 + 3R_2 \to R_1}{=} 19 \cdot \det(I_2) = 19 \cdot 1 = 19$$

טענה 7.5. *יהיו* $A \in \mathbb{M}_{n\times n}(\mathbb{F})$ *ו-* $\alpha \in \mathbb{F}$. *אז מתקיים*

$$\det(\alpha A) = \alpha^n \det(A)\,.$$

הוכחה. המשמעות של הכפל αA היא שכל שורה מוכפלת באותו סקלר α. לכן, עבור כל שורה ניתן להוציא את α מחוץ לדטרמיננטה. לאחר שנעשה זאת לכל השורות נקבל גורם α^n מחוץ לדטרמיננטה, ואת המטריצה המקורית A בתוכה. לכן

$$.\det(\alpha A)=\alpha^n\det(A)$$

□

טענה 7.6. *אם למטריצה $A\in\mathbb{M}_{n\times n}(\mathbb{F})$ יש שורת אפסים, אז $\det(A)=0$.*

הוכחה. נניח שיש רק אפסים בשורה R_i של A. אז גם לאחר הכפלת שורה זו בכל סקלר $c\neq 0$, המטריצה תישאר ללא שינוי. ניקח $c=\frac{1}{2}$ ונקבל

$$.\det(A)=2\det(A)\implies\det(A)=0$$

□

טענה 7.7. *יהיו $D,U\in\mathbb{M}_{n\times n}(\mathbb{F})$.*

א. אם D אלכסונית מהצורה

$$,D=\begin{pmatrix}d_1&0&\cdots&0\\0&d_2&\cdots&0\\\vdots&\vdots&\ddots&0\\0&0&\cdots&d_n\end{pmatrix}$$

אז מתקיים

$$.\det(D)=d_1d_2\cdots d_n$$

ב. אם U משולשית עליונה מהצורה

$$U=\begin{pmatrix}d_1&*&\cdots&*\\0&d_2&\cdots&*\\\vdots&\vdots&\ddots&*\\0&0&\cdots&d_n\end{pmatrix}$$

כאשר הכוכביות מציינות איברים כלשהם ובנוסף $d_i\neq 0$ לכל $1\leq i\leq n$, אז מתקיים

$$.\det(U)=d_1d_2\cdots d_n$$

הוכחה.

א. אם לפחות אחד מאיברי האלכסון הראשי הוא 0, אז יש שורת אפסים ולפי טענה **7.6** נובע כי

$$\det(A) = 0 = a_1 a_2 \cdots a_n .$$

אז נניח כי כל איברים האלכסון הראשי שונים מ- 0. בשביל הדירוג הקנוני צריך לבצע פש"אות מהצורה $\frac{1}{d_i}R_i \to R_i$ לכל $1 \le i \le n$. כך נקבל

$$1 = \det(I_n) = \frac{1}{d_1} \cdot \frac{1}{d_2} \cdots \frac{1}{d_n} \det(A) \implies \det(A) = d_1 d_2 \cdots d_n .$$

ב. נראה שניתן להגיע מ- U למטריצה האלכסונית D מהסעיף הקודם ע"י פש"אות שלא משנות את הדטרמיננטה. נתון שלכל $2 \le i \le n$ מתקיים $d_i \neq 0$, ולכן זהו איבר מוביל וניתן להשתמש בו כדי לאפס את האיברים מעליו ע"י פש"אות מהסוג של החסרת כפל בסקלר של R_i משורה מעליה. כל פש"א כזו לא משנה את ערך הדטרמיננטה, ולכן לאחר ביצוע כל הפש"אות האלו נקבל

$$\det(U) = \det(D) = d_1 d_2 \cdots d_n .$$

□

הערה 7.8. בהמשך נראה שהנוסחה מתקיימת גם למטריצה משולשית עליונה עם 0 אחד או יותר על האלכסון הראשי (ואז הדטרמיננטה שווה ל- 0). באופן דומה, גם נראה שלכל מטריצה משולשית תחתונה הדטרמיננטה נתונה כמכפלת איברי האלכסון הראשי.

דוגמה 7.9. ניקח

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix} .$$

תחילה נשתמש בפעולות שלא משנות את הדטרמיננטה:

$$\det(A) \overset{\substack{R_2 - 4R_1 \to R_2 \\ R_3 - 7R_1 \to R_3}}{=} \det \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -11 \end{pmatrix} \overset{R_3 - 2R_2 \to R_3}{=} \det \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 1 \end{pmatrix}$$

קיבלנו מטריצה משולשית עליונה. אז מטענה **7.7** נובע כי

$$\det(A) = 1(-3)1 = -3 .$$

7.2 תכונות נוספות של הדטרמיננטה

7.2.1 מטריצות מסדר 2 על 2

כעת נתמקד במקרה הפרטי של $n = 2$, שהוא יותר קל לניתוח וחישוב. כאן כדאי לזכור את נוסחת הדטרמיננטה שנקבל.

טענה 7.10. *עבור* $n = 2$ *הפונקציה* $\det : \mathbb{M}_{2\times 2}(\mathbb{F}) \to \mathbb{F}$ *מהגדרה 7.1 מקיימת:*

א. *לכל*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

מתקיים

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

ב. A *הפיכה אם ורק אם* $\det(A) \neq 0$. *במקרה זה מתקיים*

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

הוכחה.

א. נפריד בין שלושה מקרים. נניח תחילה כי $a \neq 0$. אז מתקיים

$$\det(A) = \det\begin{pmatrix} a & b \\ c & d \end{pmatrix} \overset{R_2 - \frac{c}{a}R_1 \to R_1}{=} \det\begin{pmatrix} a & b \\ 0 & d - \frac{bc}{a} \end{pmatrix}.$$

קיבלנו מטריצה משולשית עליונה, ולכן לפי טענה 7.7 נובע כי

$$\det(A) = a(d - \frac{bc}{a}) = ad - bc \,.$$

אם $a = 0$ אך $c \neq 0$, אז נוכל לבצע החלפה בין השורות ולהשתמש בנוסחה שהוכחנו למקרה הקודם:

$$\det(A) = -\det\begin{pmatrix} c & d \\ a & b \end{pmatrix} = -(bc - ad) = ad - bc$$

אם $a = c = 0$, אז יש שני מקרים: או שיש שורת אפסים, או שבלי הגבלת הכלליות $d \neq 0$ ואפשר לאפס את האיבר מעליו ע"י הפעולה $R_1 - \frac{b}{d}R_2 \to R_1$. אז בכל מקרה מתקבלת שורת אפסים ולפי טענה 7.6, נובע כי

$$\det(A) = \det\begin{pmatrix} 0 & b \\ 0 & d \end{pmatrix} = \det\begin{pmatrix} 0 & 0 \\ 0 & d \end{pmatrix} = 0 = 0 \cdot d - b \cdot 0 = ad - bc \,.$$

ב. כיוון ראשון: אם $\det(A) \neq 0$, נראה כי הנוסחה עבור A^{-1} אכן מקיימת $A^{-1}A = I_2$. נכפיל ונקבל

$$A^{-1}A = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{ad-bc}\begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix} = I_2.$$

אז A אכן הפיכה והנוסחה מתקיימת.

כיוון שני: נניח כי $\det(A) = 0$. ראינו בסעיף הקודם שהנחה זו מובילה לשורת אפסים לאחר דירוג, וזה אומר ש- $\operatorname{rank}(A) < 2$ ולכן A אינה הפיכה.

□

דוגמה 7.11.

א. המטריצה

$$A = \begin{pmatrix} 1 & 4 \\ 3 & 15 \end{pmatrix}$$

הפיכה כי

$$\det(A) = 1 \cdot 15 - 4 \cdot 3 = 3 \neq 0 \, .$$

בנוסף, מתקיים

$$A^{-1} = \frac{1}{3}\begin{pmatrix} 15 & -4 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} 5 & -\frac{4}{3} \\ -1 & \frac{1}{3} \end{pmatrix}.$$

ב. המטריצה

$$A = \begin{pmatrix} 1 & 5 \\ 3 & 15 \end{pmatrix}$$

אינה הפיכה כי

$$\det(A) = 1 \cdot 15 - 5 \cdot 3 = 0 \, .$$

אכן, גם רואים שדרגתה היא 1 כי השורה השנייה מתקבלת מהראשונה ע"י כפל ב- 3.

התרגיל הבא מראה שלדטרמיננטה של מטריצה מסדר 2×2 יש משמעות גיאומטרית של שטח מקבילית, עד כדי ערך מוחלט.

תרגיל 7.12. יהיו $v_1 = (a,b), v_2 = (c,d)$ וקטורים ב- $\mathbb{R}^2$. נסמן ב- A את המטריצה ששורותיה הן v_1, v_2. הראו כי $|\det(A)|$ שווה לשטח המקבילית עם וקטורי צלעות v_1, v_2.

רמז: השטח נתון ע"י $S = \|v_1\|\|v_2\|\sin(\theta)$, כאשר θ היא הזווית הקטנה בין וקטורי הצלעות ונתונה ע"י

$$\cos(\theta) = \frac{v_1 \cdot v_2}{\|v_1\|\|v_2\|} \, .$$

פתרון. נשתמש בזהות הטריגונומטרית

$$\sin^2(\theta) + \cos^2(\theta) = 1 \implies \sin(\theta) = \sqrt{1-\cos^2(\theta)}\ .$$

נציב את הביטוי לעיל בנוסחת השטח וגם נשתמש בנוסחה למכפלה סקלרית, ונקבל

$$\begin{aligned} S &= \sqrt{\|v_1\|^2\|v_2\|^2 - (v_1\cdot v_2)^2} = \sqrt{(a^2+b^2)(c^2+d^2) - (ac+bd)^2} \\ &= \sqrt{(ad)^2 + (bc)^2 - 2ad\cdot bc} = \sqrt{(ad-bc)^2} = |ad-bc| = |\det(A)| \end{aligned}$$

דוגמה 7.13.

א. יהיו $a, c, d \in \mathbb{R}$. עבור $v_1 = (a,0), v_2 = (c,d)$, אורך בסיס המקבילית הוא $|a|$ ואורך הגובה הוא $|d|$. לכן שטח המקבילית הוא $|ad|$, ואכן מתקיים

$$\left|\det\begin{pmatrix} a & 0 \\ c & d \end{pmatrix}\right| = |ad|\ .$$

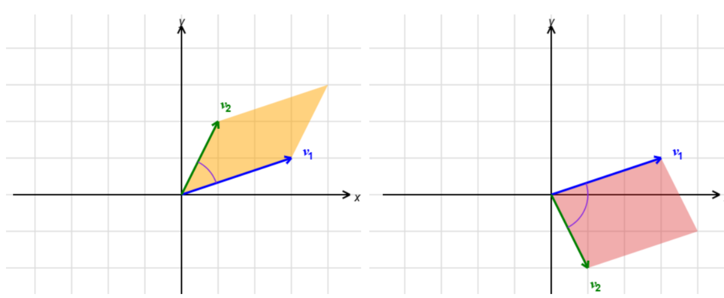
מהי משמעות הסימן של הדטרמיננטה? אם $a, d > 0$, אז $ad > 0$ והוקטורים v_1, v_2 מקיימים את התכונה הבאה: כדי לסובב את v_1 לכיוון v_2, הסיבוב הוא נגד כיוון השעון. תכונה זו מתקיימת גם אם $a, d < 0$ (עדיין $ad > 0$).

לעומת זאת, אם יש הבדל סימן בין a ל-d, כלומר $ad < 0$, אז התכונה ההפוכה מתקיימת: כדי לסובב את v_1 לכיוון v_2, הסיבוב הוא עם כיוון השעון.

המשמעות הגיאומטרית של סימן הדטרמיננטה נכונה באופן כללי למטריצה

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\ ,$$

גם אם $b \neq 0$. סימן חיובי פירושו סיבוב נגד כיוון הראשון כדי לסובב את וקטור השורה הראשון לוקטור השורה השני, וסימן שלילי פירושו סיבוב עם כיוון השעון. הדטרמיננטה מתאפסת אם ורק אם שני הוקטורים קו-לינאריים, ואז ה"מקבילית" מנוונת (שטח 0).



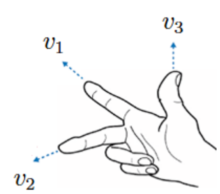
איור 7.1: בצד ימין הדטרמיננטה שלילית, ובצד שמאל היא חיובית

ב. יהיו $b_1, b_2, b_3 \in \mathbb{R}$. נסמן $v_1 = (b_1, 0, 0,), v_2 = (0, b_2, 0), v_3 = (0, 0, b_3)$, נסתכל על המטריצה שאלה הם וקטורי השורה שלה:

$$B = \begin{pmatrix} b_1 & 0 & 0 \\ 0 & b_2 & 0 \\ 0 & 0 & b_3 \end{pmatrix}$$

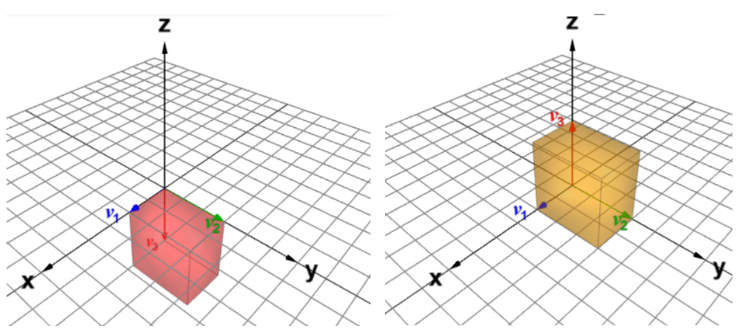
נפח התיבה שצלעותיה מתאימות לשלושת הוקטורים v_1, v_2, v_3 הוא $|b_1 b_2 b_3|$, וזה בדיוק $|\det(B)|$ לפי טענה 7.7.

כאשר $b_1, b_2, b_3 > 0$ הדטרמיננטה חיובית, והמשמעות הגיאומטרית היא כלל יד ימין: הוקטורים v_1, v_2 יוצרים "יד ימין" - כלומר אם מציבים את האצבע המורה לאורך v_1 ואת האמה לאורך v_2, אז האגודל יפנה פחות או יותר לכיוון v_3 (הזווית ביניהם תהיה חדה).



איור 7.2: כלל יד ימין

להבדיל, אם $b_1, b_2 > 0$ אך $b_3 < 0$, אז במקרה זה הדטרמיננטה שלילית ואכן כלל יד ימין לא מתקיים בגלל ההיפוך של אחד הוקטורים. זה חל גם על שאר המקרים שבהם יש היפוך בוקטור אחד בדיוק, או אפילו בכל השלושה (מספר אי-זוגי של היפוכים שבעקבותיו הדטרמיננטה שלילית).



איור 7.3: הדטרמיננטה חיובית רק בצד ימין

בהמשך ניתן דוגמה חישובית יותר כללית למשמעות של דטרמיננטה במרחב כנפח של מקבילון (לא בהכרח תיבה), עד כדי סימן בהתאם לכלל יד ימין.

7.2.2 מטריצות ריבועיות כלליות

טענה 7.14. *יהיו $E, A \in \mathbb{M}_{n\times n}(\mathbb{F})$ כך ש- E מטריצה אלמנטרית המתאימה לפש"א S, כלומר $E = S(I_n)$. אז מתקיים*

$$\det(EA) = \det(E)\det(A).$$

הוכחה. לפי טענה 6.66 מתקיים $EA = S(A)$. לכן

$$\det(EA) = \det(S(A)) = \begin{cases} \alpha\det(A), & S = (\alpha R_i \to R_i) \\ \det(A), & S = (R_i + \alpha R_j \to R_i) \\ -\det(A), & S = (R_i \leftrightarrow R_j) \end{cases}.$$

באופן דומה:

$$\det(E) = \det(S(I_n)) = \begin{cases} \alpha, & S = (\alpha R_i \to R_i) \\ 1, & S = (R_i + \alpha R_j \to R_i) \\ -1, & S = (R_i \leftrightarrow R_j) \end{cases}.$$

לכן, בכל מקרה מתקיים

$$\det(EA) = \det(E)\det(A).$$

□

יש לנו כבר מספיק כלים כדי להוכיח את אחד השימושים העיקריים של הדטרמיננטה, שהוא תנאי שקול להפיכות עבור מטריצות ריבועיות מסדר כללי.

משפט 7.15. *תהי $A \in \mathbb{M}_{n\times n}(F)$. אז A הפיכה אם ורק אם $\det(A) \neq 0$.*

הוכחה. בכיוון הראשון, נניח כי A הפיכה. אז לפי מסקנה 6.84 היא מכפלה של מטריצות אלמנטריות. לכן מתקיים

$$A = \tilde{E}_1\tilde{E}_2\cdots\tilde{E}_k$$

כאשר $\tilde{E}_1, \tilde{E}_2, ..., \tilde{E}_k$ הן מטריצות אלמנטריות (למעשה, הן מתאימות לפש"אות ההפוכות בדירוג הקנוני). אז לפי שימוש חוזר ב- 7.14 נובע כי

$$\det(A) = \det(\tilde{E}_1)\det(\tilde{E}_2)\cdots\det(\tilde{E}_k) \neq 0.$$

נדגיש שהדטרמיננטה של כל מטריצה אלמנטרית שונה מ- 0 כי היא שווה ל- ± 1 או $\alpha \neq 0$.

בכיוון השני, נניח כי A אינה הפיכה. אז לפי משפט 6.75, היא אינה שקולת שורה ל- I_n אלא למטריצה מדורגת קנונית A' עם שורת אפסים (לפחות אחת). אז הפעם מתקיים

$$A = E_1^{-1}E_2^{-1}\cdots E_k^{-1}A'$$

כאשר $E_1, E_2, ..., E_k$ הן המטריצות האלמנטריות שמתאימות לפש"אות בדירוג הקנוני. לכן

$$\det(A) = \det(E_1^{-1})\det(E_2^{-1})\cdots\det(E_k^{-1})\det(A') = 0 \,.$$

הסיבה לשוויון ל- 0 היא העובדה שמתקיים $\det(A') = 0$ לפי טענה 7.6. $\square$

תרגיל 7.16. לאילו ערכי $t \in \mathbb{C}$ המטריצה

$$A = \begin{pmatrix} t & 1 \\ -1 & t \end{pmatrix}$$

הפיכה? חשבו את A^{-1} עבור ערכים אלו.

פתרון. נחשב את הדטרמיננטה לפי הנוסחה בטענה 7.10 ונקבל $\det(A) = t^2+1$. אז לפי משפט 7.15 נובע כי A הפיכה אם ורק אם

$$t^2 + 1 \neq 0 \iff t^2 \neq -1 \iff t \neq \pm i\,.$$

בנוסף, לכל $t \neq \pm i$ מתקיים

$$A^{-1} = \frac{1}{t^2+1}\begin{pmatrix} t & -1 \\ 1 & t \end{pmatrix}.$$

כעת נוכיח תכונה חשובה שנקראת כפליות הדטרמיננטה:

טענה 7.17. *יהיו* $A, B \in \mathbb{M}_{n\times n}(\mathbb{F})$. *אז מתקיים*

$$\det(AB) = \det(A)\det(B)\,.$$

הוכחה. נחלק למקרים:

א. AB אינה הפיכה: לפי תרגיל 6.81 לפחות אחת מ- A ו- B אינה הפיכה. אז לפי משפט 7.15 נובע כי

$$\det(AB) = 0 = \det(A)\det(B)\,,$$

וזאת משום שלפחות אחת מהדטרמיננטות באגף ימין שווה ל- 0.

ב. AB הפיכה: תחילה נראה שגם A וגם B הפיכות. נסמן $C = (AB)^{-1}$. אז מתקיים

$$A(BC) = (AB)C = I_n\,,$$

ולכן A הפיכה מימין ובכלל. באופן דומה:

$$(CA)B = C(AB) = I_n\,,$$

ולכן B הפיכה משמאל ובכלל.

אז לפי טענה 6.84 גם A וגם B הן מכפלות של מטריצות אלמנטריות, כלומר קיימות מטריצות אלמנטריות $E_1, E_2, ..., E_k, E_{k+1}, ..., E_l$ כך ש-

$$B = E_k \cdots E_1 \quad , \quad A = E_l \cdots E_{k+1}\,.$$

לכן

$$AB = E_l \cdots E_{k+1} E_k \cdots E_1\,,$$

ולפי שימוש חוזר בטענה 7.14 נובע כי

$$\begin{aligned}\det(AB) &= \det(E_l) \cdots \det(E_{k+1}) \det(E_k) \cdots \det(E_1) \\ &= \det(E_l \cdots E_{k+1}) \det(E_k \cdots E_1) = \det(A)\det(B)\end{aligned}$$

□

מסקנה 7.18. *תהי* $A \in \mathbb{M}_{n\times n}(\mathbb{F})$ *הפיכה. אז מתקיים*

$$\det(A^{-1}) = \frac{1}{\det(A)}\,.$$

הוכחה. לפי 7.17 מתקיים

$$\det(A)\det(A^{-1}) = \det(AA^{-1}) = \det(I_n) = 1\,.$$

נחלק ב- $\det(A)$ ונקבל

$$\det(A^{-1}) = \frac{1}{\det(A)}\,.$$

□

תרגיל 7.19. יהיו $A, P \in \mathbb{M}_{n\times n}(\mathbb{F})$ כך ש- P הפיכה. הראו כי

$$\det(PAP^{-1}) = \det(A)\,.$$

פתרון. לפי טענה 7.17 ומסקנה 7.18 מתקיים

$$\det(PAP^{-1}) = \det(P)\det(A)\det(P^{-1}) = \det(P)\det(A)\cdot\frac{1}{\det(P)} = \det(A)\ .$$

השתמשנו בחוק החילוף לסקלרים שמתקיים ב- $\mathbb{F}$ גם אם A, P אינן מתחלפות.

הערה 7.20. המטריצה PAP^{-1} נקראת דומה ל- A. נבין את החשיבות של דמיון מטריצות בהמשך הקורס, אבל כבר ראינו שלמטריצות דומות יש דטרמיננטות שוות.

טענה 7.21. *תהי* $A \in \mathbb{M}_{n\times n}(\mathbb{F})$*. אז מתקיים*

$$\det(A^t) = \det(A)\ .$$

הוכחה. שוב נפצל למקרים לפי הפיכות A.

א. נניח כי A אינה הפיכה. לא ייתכן כי A^t הפיכה, כי אז לפי 6.81 גם $A = (A^t)^t$ הפיכה וזו סתירה. לכן, לפי 7.15 נובע כי

$$\det(A^t) = 0 = \det(A)\ .$$

ב. נניח כי A הפיכה. אז לפי 6.84 היא מכפלה של מטריצות אלמנטריות $E_1, \cdots, E_k$. לכן

$$A = E_k\cdots E_1 \implies A^t = (E_k\cdots E_1)^t = E_1^t\cdots E_k^t$$

ואז לפי טענה 7.17 נובע כי

$$\det(A) = \det(E_k)\cdots\det(E_1) \quad , \quad \det(A^t) = \det(E_1^t)\cdots\det(E_k^t)\ .$$

לאור חוק החילוף ב- $\mathbb{F}$, מה שנשאר לסיום ההוכחה זה להראות כי $\det(E^t) = \det(E)$ לכל $E \in \mathbb{M}_{n\times n}(\mathbb{F})$ אלמנטרית.

- אם E מתאימה לפש"א $S = (R_i \leftrightarrow R_j)$, אז $E^t = E$ והדטרמיננטות גם שוות (לערך -1).
- אם E מתאימה לפש"א $S = (\alpha R_i \to R_i)$, אז E היא מטריצה אלכסונית (עם 1 לאורך האלכסון הראשי חוץ מ- α בשורה ה- i) ושוב $E^t = E$.
- אם E מתאימה לפש"א $S = (R_i + \alpha R_j \to R_i)$, אז הפעם E^t מתאימה לפש"א $S = (R_j + \alpha R_i \to R_j)$. במקרה זה $E^t \neq E$, אבל מתקיים

$$\det(E^t) = 1 = \det(E)\ .$$

$\square$

נחזור לחלק השני של טענה 7.7 והפעם לא נניח שאיברי האלכסון הראשי שונים מ- 0.

מסקנה 7.22. *תהי* $U, L \in \mathbb{M}_{n\times n}(\mathbb{F})$.

א. *אם* U *משולשית עליונה מהצורה*

$$U = \begin{pmatrix} d_1 & * & \cdots & * \\ 0 & d_2 & \cdots & * \\ \vdots & \vdots & \ddots & * \\ 0 & 0 & \cdots & d_n \end{pmatrix}$$

כאשר הכוכביות מציינות איברים כלשהם, אז מתקיים

$$\det(U) = d_1 d_2 \cdots d_n\,.$$

ב. *אם* L *משולשית תחתונה מהצורה*

$$L = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ * & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ * & * & \cdots & d_n \end{pmatrix}$$

כאשר הכוכביות מציינות איברים כלשהם, אז מתקיים

$$\det(L) = d_1 d_2 \cdots d_n\,.$$

הוכחה.

א. ראינו כבר שהטענה נכונה כאשר כל איברי האלכסון הראשי שונים מ- 0. לכן, נניח שקיים $1 \le i \le n$ כך ש- $d_i = 0$, אז U מדורגת ודרגתה (מספר האיברים המובילים) קטנה מ- n. אז לפי משפט 6.75 היא אינה הפיכה, ולפי משפט 7.15 בהכרח מתקיים

$$\det(U) = 0 = d_1 d_2 \cdots d_n\,.$$

ב. נסתכל על L^t, שהיא מטריצה משולשית עליונה עם בדיוק אותם האיברים על האלכסון הראשי כמו של L. אז לפי הסעיף הקודם וטענה 7.21 מתקיים

$$\det(L) = \det(L^t) = d_1 d_2 \cdots d_n\,.$$

□

תרגיל 7.23. תהי $A \in \mathbb{M}_{n\times n}(\mathbb{F})$ כאשר n אי-זוגי. השתמשו בדטרמיננטה כדי להוכיח את הטענות הבאות:

א. אם $\mathbb{F} = \mathbb{R}$, לא ייתכן כי $A^2 = -I_n$.

ב. אם A אנטי-סימטרית, אז היא אינה הפיכה.

פתרון.

א. נניח בשלילה כי $A^2 = -I_n$ ונפעיל דטרמיננטה. אז לפי טענה **7.5** מתקיים

$$\det(A^2) = \det(-I_n) = (-1)^n \det(I_n) = -1$$

כי n אי-זוגי. נשתמש בכפליות הדטרמיננטה (טענה **7.17**) ונקבל

$$(\det(A))^2 = \det(A^2) = -1\,,$$

אך זו סתירה כי $\det(A) \in \mathbb{R}$, שכן כל ערכי A ממשיים.

ב. נתון כי $A^t = -A$ ולכן הפעם נקבל

$$\det(A^t) = \det(-A) = (-1)^n \det(A) = -\det(A)\,.$$

מצד שני, לפי טענה **7.21** מתקיים

$$\det(A) = \det(A^t) \implies \det(A) = -\det(A) \implies \det(A) = 0\,.$$

לכן, A אינה הפיכה לפי משפט **7.15**.

7.3 פיתוח דטרמיננטה לפי שורה או עמודה

ראינו שיש נוסחה פשוטה לדטרמיננטה עבור מטריצות מסדר 2×2, ותמיד אפשר להשתמש בדירוג (לא בהכרח קנוני) כדי לקבל מטריצה משולשית עליונה שגם עבורה הנוסחה ידועה. נשאלת השאלה אם יש קיצור דרך, והתשובה חיובית בהרבה מקרים.

הגדרה 7.24. תהי $A \in \mathbb{M}_{n\times n}(\mathbb{F})$. נגדיר את המינור $M_{i,j}(A)$ להיות המטריצה שמתקבלת מ-A ע"י מחיקת השורה ה-i והעמודה ה-j.

הערה 7.25. כאן $M_{i,j}(A) \in \mathbb{M}_{(n-1)\times(n-1)}(\mathbb{F})$. הסדר של המינור יותר קטן מזה של המטריצה המקורית, וזה מקל על חישוב הדטרמיננטה.

דוגמה 7.26. ניקח

$$.A = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 3 & 5 & 7 & 9 \\ -1 & 2 & 5 & 8 \\ 0 & 3 & 0 & -3 \end{pmatrix}$$

נמחק את השורה הרביעית והעמודה השנייה, ונקבל

$$.M_{4,2}(A) = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 7 & 9 \\ -1 & 5 & 8 \end{pmatrix}$$

נמחק את השורה השנייה והעמודה הרביעית, ונקבל

$$.M_{2,4}(A) = \begin{pmatrix} 1 & 0 & 2 \\ -1 & 2 & 5 \\ 0 & 3 & 0 \end{pmatrix}$$

טענה 7.27. *תהי*

$$.A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

אז מתקיים:

א. *לכל $1 \leq i \leq n$, ניתן לפתח את הדטרמיננטה לפי השורה ה- i:*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(M_{i,j}(A))$$

ב. *לכל $1 \leq j \leq n$, ניתן לפתח את הדטרמיננטה לפי העמודה ה- j:*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(M_{i,j}(A))$$

הערה 7.28. שתי הנוסחאות עלולות להיראות זהות במבט ראשון, אבל בפיתוח לפי שורה הסכימה היא על פני האינדקס j ואילו בפיתוח לפי עמודה הסכימה היא על פני האינדקס i.

נדחה את עיקר ההוכחה לפרק על העתקות לינאריות, אבל בינתיים נראה איך פיתוח לפי עמודה נובע מפיתוח לפי שורה (בעזרת שחלוף):

הוכחה.

א. נוכיח את נוסחה זו בהמשך הקורס. לעת עתה נציין שהביטוי $(-1)^{i+j}$ קשור לשינוי הסימן של הדטרמיננטה בעקבות החלפה בין שורות.

ב. נניח שהנוסחה לפיתוח לפי שורה נכונה, לכל שורה.

יהי $1 \le j \le n$. לפי טענה **7.21** מתקיים $\det(A) = \det(A^t)$. נפתח את אגף ימין לפי השורה ה- j (בהתאם לעמודה ה- j של המטריצה המקורית) ונשתמש באינדקס סכימה i:

$$\det(A) = \det(A^t) = \sum_{i=1}^{n} (-1)^{j+i} (A^t)_{ji} \det(M_{j,i}(A^t)) \, .$$
$$= \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det((M_{i,j}(A))^t) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(M_{i,j}(A))$$

השתמשנו בעובדה שהמינור $M_{j,i}(A^t)$ מתקבל ממחיקת השורה ה- j והעמודה ה- i של A^t, כאשר אין הבדל בין זה לבין השחלוף של המינור $M_{i,j}(A)$ (אפשר לבצע את המחיקה לפני או אחרי השחלוף).

□

דוגמה 7.29.

א. נראה שעבור מטריצה

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

פיתוח לפי כל שורה/עמודה משחזר את הנוסחה הידועה מטענה **7.10**.

לפי השורה הראשונה נקבל:

$$\begin{aligned} \det(A) &= (-1)^{1+1} a \cdot \det(M_{1,1}(A)) + (-1)^{1+2} b \cdot \det(M_{1,2}(A)) \\ &= a \det(d) - b \det(c) = ad - bc \end{aligned}$$

לפי השורה השנייה נקבל:

$$\begin{aligned} \det(A) &= (-1)^{2+1} c \cdot \det(M_{2,1}(A)) + (-1)^{2+2} d \cdot \det(M_{2,2}(A)) \\ &= -c \det(b) + d \det(a) = ad - bc \end{aligned}$$

לפי העמודה הראשונה נקבל:

$$\begin{aligned}\det(A) &= (-1)^{1+1}a\cdot\det(M_{1,1}(A)) + (-1)^{2+1}c\cdot\det(M_{2,1}(A))\\ &= a\det(d) - c\det(b) = ad - bc\end{aligned}$$

לפי העמודה השנייה נקבל:

$$\begin{aligned}\det(A) &= (-1)^{1+2}b\cdot\det(M_{1,2}(A)) + (-1)^{2+2}d\cdot\det(M_{2,2}(A))\\ &= -b\det(c) + d\det(a) = ad - bc\end{aligned}$$

ב. עבור

$$B = \begin{pmatrix}1 & i & 1+i\\ 2 & 3 & -i\\ 0 & 2+i & 4\end{pmatrix}$$

נבחר לפתח לפי השורה השלישית, כי אחד מאיבריה הוא 0 (ולכן מספיק לחשב דטרמיננטות של שני מינורים). נקבל

$$\begin{aligned}\det(B) &= 0 + (-1)^{3+2}(2+i)\det(M_{3,2}(B)) + (-1)^{3+3}4\det(M_{3,3}(B))\\ &= -(2+i)\det\begin{pmatrix}1 & 1+i\\ 2 & -i\end{pmatrix} + 4\det\begin{pmatrix}1 & i\\ 2 & 3\end{pmatrix}\\ &= -(2+i)\big(-i-2(1+i)\big) + 4(3-2i) = (2+i)(2+3i) + 12 - 8i\\ &= 1 + 8i + 12 - 8i = 13\end{aligned}$$

לחילופין, נוח גם לפתח לפי העמודה הראשונה. כך נקבל

$$\begin{aligned}\det(B) &= (-1)^{1+1}1\cdot\det\begin{pmatrix}3 & -i\\ 2+i & 4\end{pmatrix} + (-1)^{1+2}2\cdot\det\begin{pmatrix}i & 1+i\\ 2+i & 4\end{pmatrix} + 0\\ &= 1\big((3\cdot 4 - (-i)(2+i)\big) - 2\big(4i - (1+i)(2+i)\big)\\ &= 1(11+2i) - 2(i-1) = 11 + 2i - 2i + 2 = 13\end{aligned}$$

חוץ מפיתוח לפי שורה/עמודה תמיד כדאי לזכור שדירוג יכול להיות שימושי כחלק מהחישוב, גם אם הוא חלקי. בדוגמה הבאה נשתמש קודם בדירוג חלקי כדי לאפס את איברי העמודה הראשונה חוץ מהראשון, כך שנוכל לפתח את הדטרמיננטה לפי העמודה הראשונה יחסית בקלות.

דוגמה 7.30.

$$\det\begin{pmatrix}1&2&3&4\\0&1&5&6\\2&0&1&3\\1&1&0&1\end{pmatrix}\overset{\substack{R_3-2R_1\to R_3\\R_4-R_1\to R_1}}{=}\det\begin{pmatrix}1&2&3&4\\0&1&5&6\\0&-4&-5&-5\\0&-1&-3&-3\end{pmatrix}$$
$$=1\cdot\det\begin{pmatrix}1&5&6\\-4&-5&-5\\-1&-3&-3\end{pmatrix}$$

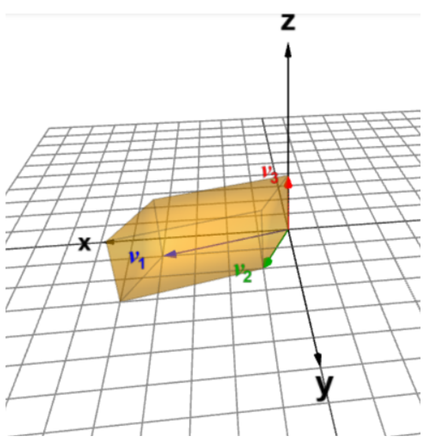
עשינו פיתוח לפי העמודה הראשונה. מכאן אפשר לעשות פיתוח לפי כל שורה או עמודה, אבל לא יזיק שוב לאפס את איברי העמודה הראשונה חוץ מהראשון:

$$\det\begin{pmatrix}1&5&6\\-4&-5&-5\\-1&-3&-3\end{pmatrix}\overset{\substack{R_2+4R_1\to R_2\\R_3+R_1\to R_3}}{=}\det\begin{pmatrix}1&5&6\\0&15&19\\0&2&3\end{pmatrix}=1\cdot\det\begin{pmatrix}15&19\\2&3\end{pmatrix}$$
$$=45-38=7$$

החישוב הבא מתאים לנפח של מקבילון שבסיסו מונח על מישור xy. זהו מקרה יותר כללי מהתיבה שראינו בדוגמה 7.13.

דוגמה 7.31. נסתכל על הוקטורים הבאים ב- $\mathbb{R}^3$:

$$.v_1=(a,b,0),v_2=(c,d,0),v_3=(0,0,h)$$



איור 7.4: המקבילון המתואר

הוקטור v_3 מאונך לוקטורים v_1, v_2 ולכן נפח המקבילון מתקבל משטח המקבילית (שנפרשת ע"י v_1, v_2) ע"י כפל בגובה $|h|$, כלומר הנפח הוא $|(ad-bc)h|$. נוודא שזהו אכן הערך המוחלט של הדטרמיננטה של המטריצה ששורותיה הן v_1, v_2, v_3, ולצורך כך נעשה פיתוח לפי השורה השלישית (גם העמודה השלישית נוחה):

$$\det\begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & h \end{pmatrix} = 0 + 0 + (-1)^{3+3} h \cdot \det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = (ad-bc)h.$$

אז הנפח אכן שווה לערך המוחלט של הדטרמיננטה (והסימן שלה מקיים את כלל יד ימין, כאמור).

הערה 7.32. עוד לא ראינו הסבר כללי לכך שהערך המוחלט של הדטרמיננטה של מטריצה מסדר 3×3 שווה לנפח המקבילון המתאים, גם אם אין וקטור שורה שמאונך לשניים האחרים. נוכל להוכיח זאת כשנחזור לגיאומטריה בסוף הקורס.

תרגיל 7.33. לאילו ערכי $a \in \mathbb{R}$ קיים פתרון יחיד לממ"ל $Ax = 0$ כאשר

$$A = \begin{pmatrix} a & 2a & 3 \\ 2a & a & 2a \\ 3 & 2a & a \end{pmatrix}?$$

פתרון. לפי משפט 7.15 קיים לממ"ל פתרון יחיד (הטריוויאלי) אם ורק אם $\det(A) \neq 0$. נעשה חישוב, כאשר תחילה נוציא את הסקלר a מהשורה השנייה אל מחוץ לדטרמיננטה:

$$\begin{aligned} \det(A) &= \det\begin{pmatrix} a & 2a & 3 \\ 2a & a & 2a \\ 3 & 2a & a \end{pmatrix} = a \cdot \det\begin{pmatrix} a & 2a & 3 \\ 2 & 1 & 2 \\ 3 & 2a & a \end{pmatrix} \\ &\overset{R_3 - R_1 \to R_3}{=} a \cdot \det\begin{pmatrix} a & 2a & 3 \\ 2 & 1 & 2 \\ 3-a & 0 & a-3 \end{pmatrix} = a(a-3)\det\begin{pmatrix} a & 2a & 3 \\ 2 & 1 & 2 \\ -1 & 0 & 1 \end{pmatrix} \end{aligned}$$

עכשיו נוח, אך לא חובה, לשחלף את המטריצה בהתאם לטענה 7.21 ואז לבצע עוד פש"א:

$$\det(A) = a(a-3)\det\begin{pmatrix} a & 2 & -1 \\ 2a & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix} \overset{R_1 + R_3 \to R_1}{=} a(a-3)\det\begin{pmatrix} a+3 & 4 & 0 \\ 2a & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix}$$

נפתח לפי העמודה השלישית ונקבל

$$\det(A) = a(a-3)\cdot 1\cdot \det\begin{pmatrix} a+3 & 4 \\ 2a & 1 \end{pmatrix} = a(a-3)(a+3-8a) = a(a-3)(3-7a) \ .$$

לכן, לממ"ל $Ax = 0$ קיים פתרון יחיד אם ורק אם

$$a(a-3)(3-7a) \neq 0 \iff a \notin \left\{ 0, \frac{3}{7}, 3 \right\}.$$

תוכן העניינים

פרק 8

מרחבים וקטוריים

8.1 הגדרת מרחבים וקטוריים

עד עכשיו עסקנו בוקטורי שורה/עמודה ב- $\mathbb{R}^n$ ו- $\mathbb{C}^n$, וגם במטריצות ב- $\mathbb{M}_{m\times n}(\mathbb{R})$ ו- $\mathbb{M}_{m\times n}(\mathbb{C})$. נמשיך להתמקד בשתי המסגרות האלו, אבל נרצה להכליל את הדיון כך שיהיה ברור שניתן להשתמש בוקטורים כדי לייצג סוגים שונים של נתונים שמתארים את העולם שלנו (או אפילו אובייקטים יותר מופשטים שיכולים להיות מעניינים במתמטיקה). מבחינה פיזיקלית, וקטורים מייצגים מקום, מהירות, תאוצה ואת כל סוגי הכוחות. אפשר להגדיר כל מיני מרחבים וקטוריים בהתאם להקשר שמעניין אותנו, אבל לכל המרחבים האלה יש תכונות בסיסיות משותפות (אקסיומות).

לפני ההגדרה, נסתכל על דוגמה הנדסית:

דוגמה 8.1. נניח שיש לנו מערכת בעלת שלושה כפתורים עצמאיים: חימום, קירור ואוורור. כל מצב של המערכת מתואר על־ידי שלשה (h, c, v), כאשר כל רכיב מייצג את עוצמת ההפעלה של כפתור אחד.

במערכת כזו ניתן לבצע שתי פעולות טבעיות: חיבור שני מצבים (שילוב השפעות של מערכות שונות), וכפל מצב במספר ממשי (הגברה או החלשה אחידה של כל ההשפעות). קיים גם מצב אפס $(0, 0, 0)$, המייצג מערכת כבויה, ולכל מצב יש מצב נגדי המבטל אותו.

אוסף כל המצבים האפשריים יחד עם פעולות אלו מקיים את כל האקסיומות של מרחב וקטורי, שנגדיר כעת.

הגדרה 8.2. מרחב וקטורי (מ"ו) V מעל $\mathbb{F}$ הוא קבוצה של איברים (שנקראים וקטורים) עם פעולות של חיבור וקטורי $+$ וכפל בסקלר $\cdot$ כך שמתקיימות האקסיומות הבאות לכל $v_1, v_2, v_3, v \in V$ ולכל $\alpha, \beta \in \mathbb{F}$:

א. סגירות: $v_1 + v_2 \in V$ וגם $\alpha v \in V$.

ב. אסוציאטיביות (חוק הקיבוץ) של חיבור: $v_1 + (v_2 + v_3) = (v_1 + v_2) + v_3$.

ג. קומוטטיביות (חוק החילוף) של חיבור: $v_1 + v_2 = v_2 + v_1$.

ד. קיום איבר ניטרלי לחיבור: קיים איבר $0_V \in V$ כך שלכל $v \in V$ מתקיים $v + 0_V = v$.

ה. קיום איבר נגדי: לכל $v \in V$ קיים $-v \in V$ כך ש $v + (-v) = 0$.

ו. אסוציאטיביות של כפל בסקלר: לכל $v \in V$ מתקיים $\alpha(\beta v) = (\alpha\beta)v$.

ז. לכל $v \in V$ מתקיים $1v = v$.

ח. דיסטריביוטיביות (חוק הפילוג): $\alpha(v_1 + v_2) = \alpha v_1 + \alpha v_2$.

ט. לכל $v \in V$ מתקיים $(\alpha + \beta)v = \alpha v + \beta v$.

דוגמה 8.3.

א. $\mathbb{R}^2$ הוא מ"ו מעל $\mathbb{R}$.

ב. $\mathbb{R}^3$ הוא מ"ו מעל $\mathbb{R}$.

ג. $\mathbb{R}^n$ הוא מ"ו מעל $\mathbb{R}$ לכל $n \in \mathbb{N}$.

ד. $\mathbb{C}^n$ הוא מ"ו מעל $\mathbb{C}$ לכל $n \in \mathbb{N}$.

ה. $\mathbb{M}_{m\times n}(\mathbb{R})$ הוא מ"ו מעל $\mathbb{R}$ לכל $m, n \in \mathbb{N}$.

ו. $\mathbb{M}_{m\times n}(\mathbb{C})$ הוא מ"ו מעל $\mathbb{C}$ לכל $m, n \in \mathbb{N}$.

הערה 8.4. כל מ"ו מעל $\mathbb{C}$ הוא גם מ"ו מעל $\mathbb{R}$, כי $\mathbb{R} \subseteq \mathbb{C}$ ותמיד אפשר לצמצם את הגדרת הכפל בסקלר רק לסקלרים ממשיים. בפרט, $\mathbb{C}^2$ הוא מ"ו מעל $\mathbb{R}$ אם מניחים שהסקלרים הם ממשיים בלבד (אבל הוקטורים עצמם לא בהכרח ב- $\mathbb{R}^2$). למען הסר ספק, לרוב נעסוק ב- $\mathbb{R}^n$ מעל $\mathbb{R}$, וב- $\mathbb{C}^n$ מעל $\mathbb{C}$.

ניתן להוכיח תכונות נוספות של מרחבים וקטוריים בעזרת האקסיומות:

טענה 8.5. *יהי V מ"ו מעל $\mathbb{F}$. אז לכל $v, w_1, w_2 \in V$ ולכל $\alpha \in \mathbb{F}$ מתקיים:*

א. $w_1 + v = w_2 + v \implies w_1 = w_2$

ב. $0 \cdot v = 0_V$

ג. $(-1) \cdot v = -v$

הוכחה.

א. נניח כי מתקיים $w_1 + v = w_2 + v$. נשתמש באיבר הנגדי $-v$ ונחבר אותו לשני אגפי המשוואה:

$$w_1+(v+(-v)) = w_2+(v+(-v)) \implies w_1+0_V = w_2+0_V \implies w_1 = w_2$$

תחילה השתמשנו באסוציאטיביות ואחר כך בניטרליות של 0_V.

ב. לכל $v \in V$ מתקיים

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

קיים איבר נגדי $-(0\cdot v)$ וניתן לחבר אותו לשני אגפי המשוואה:

$$0\cdot v+0\cdot v+(-(0\cdot v)) = 0\cdot v+(-(0\cdot v)) \implies 0\cdot v+0_V = 0_V \implies 0\cdot v = 0_V$$

ג. לכל $v \in V$ מתקיים

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

לפי האקסיומה האחרונה ולפי סעיף ב' נובע כי

$$1\cdot v + (-1)v = 0_V.$$

לפי עוד אקסיומה מתקיים $1\cdot v = v$, ובנוסף קיים איבר נגדי $-v$. נחבר אותו לשני אגפי המשוואה משמאל ונקבל

$$-v + v + (-1)v = -v + 0_V \implies 0_V + (-1)v = -v \implies (-1)v = -v$$

$\square$

הערה 8.6. פעולת החיסור בין שני וקטורים $v_1, v_2 \in V$ מוגדרת באופן הבא:

$$v_1 - v_2 = v_1 + (-v_2).$$

זהו סימון יותר נוח, אבל כדאי לזכור שחיסור הוא למעשה חיבור האיבר הנגדי.

תרגיל 8.7. יהי V מ"ו. הראו כי לכל $v \in V$ ולכל $\alpha \in \mathbb{F}$ מתקיים

$$\alpha \cdot 0_V = 0_V.$$

רמז: השתמשו בשוויון $0_V + 0_V = 0_V$.

הוכחה. יהיו $v \in V$ ו- $\alpha \in \mathbb{F}$. אז מתקיים

$$.\alpha \cdot (0_V + 0_V) = \alpha \cdot 0_V$$

לפי דיסטריביוטיביות נובע כי

$$.\alpha \cdot 0_V + \alpha \cdot 0_V = \alpha \cdot 0_V$$

נחבר את האיבר הנגדי $-(\alpha \cdot 0_V)$ לשני אגפי המשוואה, ונקבל

$$\alpha \cdot 0_V + \alpha \cdot 0_V - (\alpha \cdot 0_V) = \alpha \cdot 0_V - (\alpha \cdot 0_V) \implies \alpha \cdot 0_V = 0_V$$

$\square$

ניתן דוגמאות למרחבים וקטוריים נוספים הקשורים לחדו"א, אך נציין מראש שלא נעסוק בהם בקורס. טוב לדעת שמרחבים וקטוריים יכולים לצוץ במקומות לכאורה לא צפויים.

דוגמה 8.8.

א. מרחב הסדרות הממשיות:

$$\mathbb{R}^{\mathbb{N}} = \{ \{ a_n \}_{n=1}^{\infty} \mid a_1, a_2, \cdots \in \mathbb{R} \}$$

עם חיבור וקטורי שמוגדר ע"י

$$,\{ a_n \}_{n=1}^{\infty} + \{ b_n \}_{n=1}^{\infty} = \{ a_n + b_n \}_{n=1}^{\infty}$$

כפל בסקלר שמוגדר ע"י

$$\alpha \cdot \{ a_n \}_{n=1}^{\infty} = \{ \alpha a_n \}_{n=1}^{\infty}$$

ואיבר ניטרלי

$$.0_V = \{ 0 \}_{n=1}^{\infty}$$

ב. מרחב הפונקציות הממשיות:

$$\mathbb{R}^{\mathbb{R}} = \{ f : \mathbb{R} \to \mathbb{R} \}$$

עם חיבור וקטורי מוגדר ע"י

$$,\forall x \in \mathbb{R} \quad (f+g)(x) = f(x) + g(x)$$

כפל בסקלר המוגדר ע"י

$$\forall x \in \mathbb{R} \quad (\alpha f)(x) = \alpha \cdot f(x)$$

ואיבר ניטרלי

$$.\forall x \in \mathbb{R} \quad 0_V(x) = 0$$

הערה 8.9. אחת הסיבות שלא נעסוק במרחבים אלה היא שהמימד שלהם הוא אינסופי. עוד לא הגדרנו מימד, אבל אפשר אפשר לחשוב עליו כעל מספר הפרמטרים הדרושים לתיאור המרחב. במקרה של המישור מדובר ב- 2, ובהכללה המימד של $\mathbb{R}^n$ הוא n (מספר הקוארדינטות). האינטואיציה רומזת שמרחבי הסדרות והפונקציות הרבה יותר מסובכים, כלומר הם ממימד אינסופי. זה חורג ממסגרת הקורס.

8.2 תת-מרחבים

הגדרה 8.10. יהי V מ"ו מעל $\mathbb{F}$. תת-קבוצה $W \subseteq V$ תיקרא תת-מרחב של V אם W היא מ"ו מעל $\mathbb{F}$ בפני עצמה, עם פעולות החיבור והכפל בסקלר מ- V.

הערה 8.11. בהינתן ש- V הוא מ"ו, רוב התכונות של הגדרה 8.2 עוברות בירושה אל W באופן אוטומטי. למעשה, עבור תת-קבוצה W מספיק לבדוק שיש בה איבר ניטרלי, כלומר $0_V \in W$, ובנוסף היא מקיימת את תנאי הסגירות:

$$\forall w_1, w_2 \in W \quad \forall \alpha \in \mathbb{F} \quad w_1 + w_2 \in W \quad , \quad \alpha w_1 \in W$$

נוכיח זאת בהמשך. ראשית נראה דוגמאות.

דוגמה 8.12.

א. לכל מ"ו V מעל $\mathbb{F}$ יש תת-מרחב $\{ 0_V \}$. זהו תת-המרחב הקטן ביותר של V. הוא מקיים את תנאי הסגירות כי $0_V + 0_V = 0_V$ ולכל $\alpha \in \mathbb{F}$ מתקיים $\alpha \cdot 0_V = 0_V$.

ב. לכל $v \in \mathbb{R}^2$ כך ש- $v \neq 0$, הישר $L = \{ tv \mid t \in \mathbb{R} \}$ הוא תת-מרחב של $\mathbb{R}^2$.

ג. לכל $v_1, v_2 \in \mathbb{R}^3$ שאינם קו-לינאריים, המישור $H = \{ sv_1 + tv_2 \mid s, t \in \mathbb{R} \}$ הוא תת-מרחב של $\mathbb{R}^3$. אם הוקטורים הם קו-לינאריים ולמשל $v_2 = \alpha v_1$ עבור $\alpha \in \mathbb{R}$, אז למעשה מתקיים $H = \{ tv_1 \mid t \in \mathbb{R} \}$. זהו ישר, אך עדיין תת-מרחב.

ד. $W_1 = \{ (x, y, z, 0) \mid x, y, z \in \mathbb{R} \}$ הוא תת-מרחב של $\mathbb{R}^4$. הוא "דומה" ל- $\mathbb{R}^3$, אך הוא מוכל ב- $\mathbb{R}^4$.

ה. $W_2 = \{ (x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0 \}$ היא תת-מרחב של $\mathbb{R}^4$. נבדוק זאת:

קיום איבר ניטרלי: מתקיים $(0, 0, 0) \in W_2$ כי $0 + 0 + 0 = 0$.

תנאי הסגירות: יהיו $w_1 = (x_1, y_1, z_1), w_2 = (x_2, y_2, z_2) \in W_2$.

נבדוק אם מתקיים $w_1 + w_2 = (x_1 + x_2, y_1 + y_2, z_1 + z_2) \in W_2$. אכן:

$$(x_1+x_2)+(y_1+y_2)+(z_1+z_2) = (x_1+y_1+z_1)+(x_2+y_2+z_2) = 0+0 = 0$$

כנדרש.

באופן דומה: יהי $\alpha \in \mathbb{R}$. עבור $\alpha w_1 = (\alpha x_1, \alpha y_1, \alpha z_1)$ מתקיים

$$\alpha x_1 + \alpha y_1 + \alpha z_1 = \alpha(x_1 + y_1 + z_1) = \alpha \cdot 0 = 0$$

כנדרש.

טענה 8.13. *יהי V מ"ו מעל $\mathbb{F}$. תהי $W \subseteq V$. התנאים הבאים שקולים:*

- W *תת-מרחב של* V.
- $0_V \in W$ ***וגם לכל*** $w_1, w_2 \in W$ ***ולכל*** $\alpha \in \mathbb{F}$ ***מתקיים*** $w_1+w_2 \in W$ ***וגם*** $\alpha w_1 \in W$.

הוכחה. בכיוון הראשון: אם W תת-מרחב של V, אז הוא מקיים את כל התכונות של הגדרה 8.2 למרחב וקטורי. לא רק שיש איבר ניטרלי ב- W, מדובר על 0_V כי לכל $w \in W$ מתקיים $0_V = 0 \cdot w \in W$ לפי טענה 8.5.

בכיוון השני: נתון שאקסיומת הסגירות מתקיימת עבור W, וגם האקסיומה של קיום איבר ניטרלי. כמעט כל שאר האקסיומות מתקיימות באופן אוטומטי לכל תת-קבוצה של V, אבל צריך לוודא קיום איבר נגדי לכל $w \in W$. זה נובע מסגירות לכפל בסקלר ומטענה 8.5 כי $-w = (-1)w \in W$.

□

הגדרה 8.14. תהי $A \in \mathbb{M}_{m\times n}(\mathbb{F})$. קבוצת הפתרונות לממ"ל ההומוגנית $Ax = 0$ נקראת מרחב הפתרונות ונסמנה ע"י $\mathrm{N}(A)$.

הסיבה להגדרה המיוחדת (במקרה של ממ"ל הומוגנית) היא הטענה הבאה:

טענה 8.15. $\mathrm{N}(A)$ *הוא תת-מרחב של* $\mathbb{F}^n$.

הוכחה. ראשית, וקטור האפס $0 \in \mathbb{F}^n$ מקיים $0 \in \mathrm{N}(A)$ כי $A0 = 0$.

יהיו $w_1, w_2 \in \mathrm{N}(A)$. אז לפי ההגדרה מתקיים $Aw_1 = 0 = Aw_2$, ולכן

$$A(w_1 + w_2) = Aw_1 + Aw_2 = 0 + 0 = 0 \implies w_1 + w_2 \in \mathrm{N}(A).$$

בנוסף, לכל $\alpha \in \mathbb{F}$ מתקיים

$$A(\alpha w_1) = \alpha A w_1 = \alpha \cdot 0 = 0 \implies a w_1 \in \mathrm{N}(A).$$

לכן $\mathrm{N}(A)$ הוא תת-מרחב של $\mathbb{F}^n$ לפי טענה 8.13. □

הערה 8.16. עבור $b \neq 0$ קבוצת הפתרונות של הממ״ל $Ax = b$ אינה תת-מרחב של $\mathbb{F}^n$. אין בה איבר ניטרלי לחיבור כי $b \neq 0 = A0$. יותר מכך, היא לא סגורה לחיבור וגם לא לכפל בסקלר. אם w_1, w_2 שייכים לקבוצת הפתרונות, אז

$$A(w_1 + w_2) = Aw_1 + Aw_2 = b + b = 2b \neq b$$

לכן לממ״ל הומוגנית יש חשיבות בהקשר של מרחב וקטוריים, להבדיל מממ״ל לא הומוגנית.

תרגיל 8.17. הראו כי המישור

$$H = \{ (1 + s, t, s + t) \mid s, t \in \mathbb{R} \}$$

אינו תת-מרחב של $\mathbb{R}^3$.

פתרון. אפשר לזהות קבוצה זו כקבוצת הפתרונות של המשוואה הלינארית $x + y - z = 1$, שאינה הומוגנית ולכן קבוצת הפתרונות אינה תת-מרחב. אבל גם בלי המשוואה אפשר לבדוק שוקטור האפס לא שייך ל- H. נשווה את ההצגה הפרמטרית לוקטור האפס ונקבל:

$$(1 + s, t, s + t) = (0, 0, 0)$$

כך מקבלים $(s, t) = (-1, 0)$ מצד אחד, ואת המשוואה $s + t = 0$ מצד שני. סתירה.

מה לגבי מטריצות? הבנו שאפשר להתייחס ל- $\mathbb{M}_{m\times n}(\mathbb{F})$ כמרחב וקטורי (זה מעט מבלבל בהתחלה), אז אילו תת-קבוצות הן ? הדוגמאות החשובות הן עבור מטריצות ריבועיות ($m = n$), ולמעשה כבר ראינו אותן.

דוגמה 8.18.

א. נסמן ב- W_1 את קבוצת המטריצות המשולשיות עליונות. אז W_1 היא תת-מרחב של $\mathbb{M}_{n\times n}(\mathbb{F})$.

ב. נסמן ב- W_2 את קבוצת המטריצות המשולשיות תחתונות. אז W_2 היא תת-מרחב של $\mathbb{M}_{n\times n}(\mathbb{F})$.

ג. נסמן ב- W_3 את קבוצת המטריצות האלכסוניות. אז W_3 היא תת-מרחב של $\mathbb{M}_{n\times n}(\mathbb{F})$.

ד. נסמן ב- W_4 את קבוצת המטריצות הסימטריות. אז W_4 היא תת-מרחב של $\mathbb{M}_{n\times n}(\mathbb{F})$.

ה. נסמן ב- W_5 את קבוצת המטריצות האנטי-סימטריות. אז W_5 היא תת-מרחב של $\mathbb{M}_{n\times n}(\mathbb{F})$.

ו. $W_6 = \{ \mathbf{0}_{n\times n} \}$ היא תת-מרחב של $\mathbb{M}_{n\times n}(\mathbb{F})$.

תרגיל 8.19. בדקו שכל הדוגמאות הן אכן תת-מרחבים של $\mathbb{M}_{n\times n}(\mathbb{F})$.

פתרון. קל לראות שמטריצת האפס שייכת לכל אחת מהקבוצות לעיל. היא אלכסונית (ולכן גם משולשית עליונה וגם משולשית תחתונה), סימטרית ואנטי-סימטרית. נסתפק בבדיקת הסגירות עבור W_1 כי שאר הבדיקות דומות.

יהיו $A, B \in W_1$. אז לפי ההגדרה, לכל $i > j$ מתקיים $(A)_{ij} = 0 = (B)_{ij}$. לכן

$$\forall i > j \quad (A+B)_{ij} = (A)_{ij} + (B)_{ij} = 0 + 0 = 0$$

וזה בדיוק אומר שמתקיים $A + B \in W_1$. באופן דומה, לכל $\alpha \in \mathbb{F}$ מתקיים

$$\forall i > j \quad (\alpha A)_{ij} = \alpha (A)_{ij} = \alpha \cdot 0 = 0$$

ולכן $\alpha A \in W_1$.

הערה 8.20. האינטואיציה המומלצת לתת-מרחב היא כזו: אם התנאים שמגדירים את W הם למעשה ממ"ל הומוגנית, אז W אמורה להיות תת-מרחב. למשל, עבור W_1 ההגדרה מתייחסת לכל ערכי המטריצה מתחת לאלכסון הראשי. הם אמורים להתאפס, ולכן בעצם יש פה ממ"ל הומוגנית.

עבור W_4 התנאים למטריצה סימטרית הם אוסף משוואות מהצורה $a_{ij} = a_{ji}$, אז גם כאן מסתתרת ממ"ל הומוגנית. וכן הלאה.

אבל בפועל, הבדיקה המלאה לצורך הפתרון היא לפי טענה 8.13.

דוגמה 8.21. נראה כי $W = \{\, A \in \mathbb{M}_{2\times 2}(\mathbb{R}) \mid \det(A) = 0 \,\}$ אינה תת-מרחב של $\mathbb{M}_{2\times 2}(\mathbb{R})$. אמנם מתקיים $\mathbf{0}_{2\times 2} \in W$, אבל אין סגירות לחיבור. נמצא דוגמה נגדית: עבור

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

מתקיים

$$.\det(A) = 0 = \det(B) \quad , \quad \det(A+B) = \det(I_2) = 1 \neq 0$$

לכן $A, B \in W$ אך $A + B \notin W$.

אז W אינו תת-מרחב. נציין שדווקא מתקיימת סגירות לכפל בסקלר, אבל זה לא מספיק.

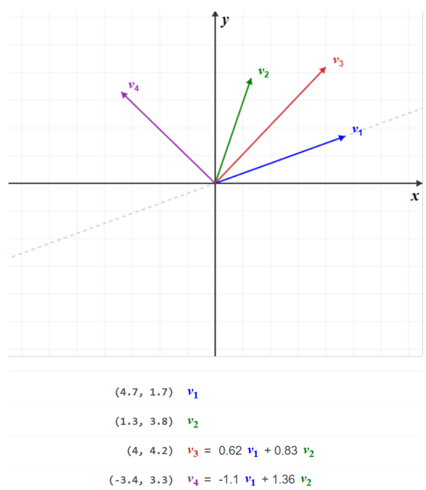
הערה 8.22. הדוגמה הנגדית לעיל גם מראה שלא בהכרח מתקיים

$$.\det(A+B) = \det(A) + \det(B)$$

למעשה, נדיר למצוא מטריצות (השונות ממטריצת האפס) שעבורן יש שוויון.

8.3 פרישה

נניח כי W היא תת-מרחב של מ"ו V מעל $\mathbb{F}$. בהינתן $w_1, w_2 \in W$ נובע מטענה 8.13 שמתקיים $\alpha w_1, \beta w_2 \in W$ לכל $\alpha, \beta \in \mathbb{F}$, ולכן נובע כי $\alpha w_1 + \beta w_2 \in W$. כך שני הוקטורים פורשים אינסוף וקטורים ב- W (לא בהכרח את כולם), והוקטורים האלה נקראים צירופים לינאריים של w_1, w_2.



איור 8.1: פרישה של שני וקטורים לדוגמה במישור ע"י שני וקטורים נתונים, כאשר בכחול מסומן הישר הנפרש ע"י וקטור אחד

הרעיון הזה נקרא פרישה, והוא עובד למספר כללי של וקטורים ב- V. אפשר להשתמש בפרישה כדי להגדיר תת-מרחב חדש.

הגדרה 8.23. יהי V מ"ו מעל $\mathbb{F}$, ויהיו $v_1, v_2, ..., v_k \in V$. אז נגדיר את קבוצת הצירופים הלינאריים של וקטורים אלה:

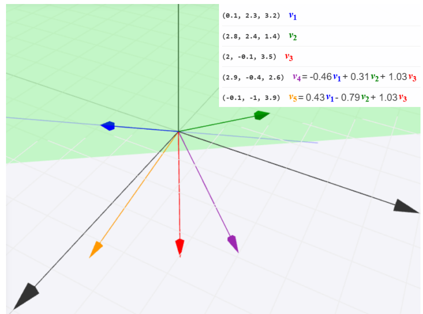
$$\mathrm{Span}(v_1, ..., v_k) = \{\, \alpha_1 v_1 + ... + \alpha_k v_k \mid \alpha_1, ..., \alpha_k \in \mathbb{F} \,\}$$

הערה 8.24. לפי ההגדרה, מתקיים

$$v \in \mathrm{Span}(v_1, ..., v_k) \iff \exists \alpha_1, ..., \alpha_k \in \mathbb{F} \quad v = \alpha_1 v_1 + ... + \alpha_k v_k$$

כאן השתמשנו בסימן $\exists$ שפירושו "קיימים", כלומר הטענה בצד ימין היא שקיימים סקלרים עבורם ("כך ש") מתקיים השוויון של תנאי הפרישה.

רצוי לזכור את $\forall$ (לכל) ו- $\exists$ (קיים/קיימים) כי לפעמים הם מקצרים את הכתיבה. הם נקראים כמתים, במובן של "כימות" הטענה לפי התייחסות לכל הקבוצה או רק לאיבר/איברים מסוימים.



איור 8.2: פרישה של שני וקטורים לדוגמה במרחב ע"י שלושה וקטורים נתונים, כאשר בכחול מסומן הישר הנפרש ע"י וקטור אחד ובירוק מסומן המישור הנפרש ע"י שני וקטורים

טענה 8.25. *$\mathrm{Span}(v_1, ..., v_k)$ הוא תת-מרחב של V.*

הוכחה. נבדוק את התנאים של טענה 8.13.

- קיום איבר ניטרלי: מתקיים

$$0_V = 0 \cdot v_1 + 0 \cdot v_2 + ... + 0 \cdot v_k \in Span(v_1, ..., v_k).$$

- סגירות לחיבור: יהיו $w_1, w_2 \in \mathrm{Span}(v_1, ..., v_k)$.
  אז קיימים סקלרים $\alpha_1, ..., \alpha_k, \beta_1, ..., \beta_k \in \mathbb{F}$ כך ש-

$$w_1 = \alpha_1 v_1 + ... + \alpha_k v_k \quad , \quad w_2 = \beta_1 v_1 + ... \beta_k v_k.$$

  מתקיים:

$$\begin{aligned} w_1 + w_2 &= (\alpha_1 v_1 + ... + \alpha_k v_k) + (\beta_1 v_1 + ... + \beta_k v_k) \\ &= (\alpha_1 + \beta_1) v_1 + ... + (\alpha_k + \beta_k) v_k \in \mathrm{Span}(v_1, ..., v_k) \end{aligned}$$

- סגירות לכפל בסקלר: לכל $\gamma \in \mathbb{F}$ מתקיים

$$\gamma w_1 = \gamma(\alpha_1 v_1 + ... + \alpha_k v_k) = (\gamma\alpha_1) v_1 + ... + (\gamma\alpha_k) v_k \in \mathrm{Span}(v_1, ..., v_k).$$

  לכן, $\mathrm{Span}(v_1, ..., v_k)$ הוא תת-מרחב של V.

□

טענה 8.26. *יהי W תת-מרחב של מ"ו V מעל $\mathbb{F}$. אז לכל $v_1, ..., v_k \in V$ מתקיים*

$$v_1, ..., v_k \in W \iff \mathrm{Span}(v_1, ..., v_k) \subseteq W.$$

הוכחה. הכיוון $\Longrightarrow$: נניח כי $v_1, ..., v_k \in W$. אז לכל $\alpha_1, ..., \alpha_k \in \mathbb{F}$ מתקיים

$$\forall 1 \leq j \leq n \quad \alpha_j v_j \in W$$

מסגירות לכפל בסקלר. לכן, נובע מסגירות לחיבור כי

$$\alpha_1 v_1 + ... + \alpha_k v_k \in W.$$

מכאן:

$$\mathrm{Span}(v_1, ..., v_k) \subseteq W$$

הכיוון $\Longleftarrow$: נניח כי $\mathrm{Span}(v_1, ..., v_k) \subseteq W$. לכל $1 \leq i \leq n$ נראה כי $v_i \in W$. אכן, ניקח $\alpha_j = 0$ לכל $j \neq i$, יחד עם $\alpha_i = 1$. אז נקבל

$$v_i = 0 \cdot v_1 + ... + 0 \cdot v_{i-1} + 1 \cdot v_i + 0 \cdot v_{i+1} + ... + 0 \cdot v_k \in \mathrm{Span}(v_1, ..., v_k) \subseteq W$$

לכן, מתקיים $v_1, ..., v_k \in W$. $\square$

הגדרה 8.27. יהי V מ"ו מעל $\mathbb{F}$. נאמר ש- V נוצרת סופית אם קיימים $v_1, ..., v_k \in V$ כך ש- $V = \mathrm{Span}(v_1, ..., v_k)$. במקרה זה גם נאמר שהקבוצה $\{ v_1, ..., v_k \}$ פורשת את V.

דוגמה 8.28 א ב- $\mathbb{F}^n$ קבוצת וקטורי היחידה היא קבוצה פורשת. מתקיים

$$\mathrm{Span}(e_1, ..., e_n) = \mathbb{F}^n$$

כי לכל $v = (\alpha_1, ..., \alpha_n) \in \mathbb{F}^n$ מתקיים

$$v = \alpha_1 e_1 + ... + \alpha_n e_n.$$

ב ב- $\mathbb{M}_{2\times 2}(\mathbb{F})$ נגדיר

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

הקבוצה $\{ A_1, A_2, A_3, A_4 \}$ פורשת את $\mathbb{M}_{2\times 2}(\mathbb{F})$ כי לכל $a, b, c, d \in \mathbb{F}$ מתקיים

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = aA_1 + bA_2 + cA_3 + dA_4.$$

ג ב- $\mathbb{M}_{m\times n}(\mathbb{F})$ נסמן לכל $1 \le i \le m, 1 \le j \le n$ ב- $P(i,j)$ את המטריצה שכולה אפסים פרט ל- 1 בשורה ה- i ובעמודה ה- j, כלומר

$$P(i,j) = \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}.$$

אז $\{\, P(i,j) \mid 1 \le i \le m, 1 \le j \le n \,\}$ **פורשת את** $M_{m\times n}(\mathbb{F})$ **כי לכל** $A \in M_{m\times n}(\mathbb{F})$ מתקיים

$$A = \sum_{1\le i\le m}\sum_{1\le j\le n} (A)_{ij}P(i,j) = (A)_{11}P(1,1) + ... + (A)_{1n}P(1,n)$$
$$+ ... + (A)_{m1}P(m,1) + ... + (A)_{mn}P(m,n)$$

תרגיל 8.29. נסתכל על $\mathbb{M}_{2\times 2}(\mathbb{F})$. כל תת-מרחב הוא מ"ו בפני עצמו וניתן למצוא לו קבוצה פורשת.

א. מצאו קבוצה פורשת עבור תת-המרחב של מטריצות סימטריות.

ב. מצאו קבוצה פורשת עבור תת-המרחב של מטריצות אנטי-סימטריות.

פתרון.

א. כל מטריצה סימטרית היא מהצורה

$$\begin{pmatrix} a & b \\ b & d \end{pmatrix} = a\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + d\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

אז קיבלנו קבוצה פורשת:

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

ב. כל מטריצה אנטי-סימטרית היא מהצורה

$$\begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix} = b\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

אז קיבלנו קבוצה פורשת:

$$\left\{ \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\}$$

ההגדרה הבאה מתייחסת לתת-מרחבים שקשורים למטריצה נתונה.

הגדרה 8.30. תהי $A \in \mathbb{M}_{m\times n}(\mathbb{F})$ עם שורות $R_1, ..., R_m \in \mathbb{F}^n$ ועמודות $C_1, ..., C_n \in \mathbb{F}^m$. נגדיר את מרחב השורות

$$\mathrm{Row}(A) = \mathrm{Span}(R_1, ..., R_m) \subseteq \mathbb{F}^n ,$$

ואת מרחב העמודות

$$\mathrm{Col}(A) = \mathrm{Span}(C_1, ..., C_n) \subseteq \mathbb{F}^m .$$

דוגמה 8.31.

א. עבור

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

מתקיים

$$\mathrm{Row}(A) = \mathrm{Span}((1,2,3),(-1,0,1)) ,$$
$$\mathrm{Col}(A) = \mathrm{Span}\left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}\right) .$$

ב.

$$B = \begin{pmatrix} i & 0 \\ 2 & 1+i \\ 3 & -i \end{pmatrix}$$

מתקיים

$$\mathrm{Row}(B) = \mathrm{Span}((i,0),(2,1+i),(3,-i)) ,$$
$$\mathrm{Col}(B) = \mathrm{Span}\left(\begin{pmatrix} i \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1+i \\ -i \end{pmatrix}\right) .$$

הטענה הבאה מראה את הקשר בין מרחב עמודות לממ"ל.

טענה 8.32. *תהי $A \in \mathbb{M}_{m\times n}(\mathbb{F})$. אז לכל וקטור עמודה $w \in \mathbb{F}^m$ מתקיים $w \in \mathrm{Col}(A)$ אם ורק אם קיים פתרון לממ"ל $Ax = w$.*

הוכחה. נכתוב את A בהתאם לוקטורי העמודות שלה:

$$A = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix}$$

יהי $w \in \mathbb{F}^m$. אז לפי ההגדרה, מתקיים

$$\begin{aligned} w \in \mathrm{Col}(A) &\iff \exists \alpha_1, \ldots, \alpha_n \in \mathbb{F} \quad w = \alpha_1 v_1 + \cdots + \alpha_n v_n \\ &\iff \exists x = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{F}^n \quad Ax = w \end{aligned}$$

השתמשנו בחישוב

$$Ax = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \alpha_1 v_1 + ... + \alpha_n v_n.$$

$\square$

דוגמה 8.33. תהי

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

נבדוק איזה וקטור (אם בכלל) מהוקטורים הבאים שייך ל- $\mathrm{Col}(A)$:

$$w_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \quad w_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

לפי טענה 8.32 צריך לפתור שתי ממ"ליות: $Ax = w_1$ מחד ו- $Ax = w_2$ מאידך. ניתן לפתור את שתיהן במקביל ע"י מטריצה מורחבת $(A|w_1|w_2)$ ודירוג לפי A:

$$\left(\begin{array}{ccc|c|c} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 2 & 2 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right) \xrightarrow{R_3-R_1\to R_3} \left(\begin{array}{ccc|c|c} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right)$$

$$\xrightarrow{R_3-R_2\to R_3} \left(\begin{array}{ccc|c|c} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right) \xrightarrow{R_3\leftrightarrow R_4} \left(\begin{array}{ccc|c|c} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{R_1-R_3\to R_1} \left(\begin{array}{ccc|c|c} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow{R_2-R_3\to R_2} \left(\begin{array}{ccc|c|c} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

המסקנה היא שלא קיים פתרון לממ"ל $Ax = w_2$ כי יש שורת סתירה, ולכן $w_2 \notin \mathrm{Col}(A)$.
מנגד, קיים פתרון לממ"ל $Ax = w_1$. לכן, מתקיים $w_1 \in \mathrm{Col}(A)$ ואף מצאנו את הסקלרים בתנאי הפרישה:

$$\begin{pmatrix}1\\1\\2\\0\end{pmatrix} = 1\cdot\begin{pmatrix}1\\0\\1\\0\end{pmatrix} + 1\cdot\begin{pmatrix}0\\1\\1\\0\end{pmatrix} + 0\cdot\begin{pmatrix}1\\1\\2\\1\end{pmatrix}$$

תרגיל 8.34. **מצאו קבוצה פורשת עבור $\mathrm{N}(A)$ כאשר**

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

פתרון. **נפתור את הממ"ל ההומוגנית $Ax = 0$ ע"י דירוג, ונראה שהוקטורים הפורשים מסתתרים בתוך ההצגה הפרמטרית.**

$$\begin{pmatrix}1 & 1 & 0 & -1\\0 & 1 & 1 & 1\\1 & 2 & 1 & 0\end{pmatrix} \xrightarrow{R_3-R_1-R_2\to R_3} \begin{pmatrix}1 & 1 & 0 & -1\\0 & 1 & 1 & 1\\0 & 0 & 0 & 0\end{pmatrix} \xrightarrow{R_1-R_2\to R_1} \begin{pmatrix}1 & 0 & -1 & -2\\0 & 1 & 1 & 1\\0 & 0 & 0 & 0\end{pmatrix}$$

$$\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix} \in \mathrm{N}(A) \iff \begin{cases} x_1 - x_3 - 2x_4 = 0 \\ x_2 + x_3 + x_4 = 0 \end{cases}$$

נציב $(x_3,x_4)=(s,t)$ ונקבל

$$\mathrm{N}(A)=\left\{\begin{pmatrix}s+2t\\-s-t\\s\\t\end{pmatrix}\;\middle|\; s,t\in\mathbb{R}\right\}=\left\{s\begin{pmatrix}1\\-1\\1\\0\end{pmatrix}+t\begin{pmatrix}2\\-1\\0\\1\end{pmatrix}\;\middle|\; s,t\in\mathbb{R}\right\}$$
$$=\mathrm{Span}\left(\begin{pmatrix}1\\-1\\1\\0\end{pmatrix},\begin{pmatrix}2\\-1\\0\\1\end{pmatrix}\right)$$

השוויון השני קל לבדיקה, אבל הדרך להגיע אליו היא "לפרק" את הוקטור $\begin{pmatrix}s+2t\\-s-t\\s\\t\end{pmatrix}$ לשני וקטורים כך שבכל אחד יופיע רק פרמטר אחד, ואז מכל וקטור אפשר להוציא את הפרמטר המתאים החוצה לפי כפל בסקלר. הנה ההסבר המלא:

$$\begin{pmatrix}s+2t\\-s-t\\s\\t\end{pmatrix}=\begin{pmatrix}s\\-s\\s\\0\end{pmatrix}+\begin{pmatrix}2t\\-t\\0\\t\end{pmatrix}=s\begin{pmatrix}1\\-1\\1\\0\end{pmatrix}+t\begin{pmatrix}2\\-1\\0\\1\end{pmatrix}$$

אז מצאנו קבוצה פורשת עבור $\mathrm{N}(A)$:

$$\left\{\begin{pmatrix}1\\-1\\1\\0\end{pmatrix},\begin{pmatrix}2\\-1\\0\\1\end{pmatrix}\right\}$$

8.4 תלות ואי-תלות לינארית

אנחנו מכירים את ההגדרה של שני וקטורים קו-לינאריים, שהיא למעשה מקרה פרטי של הגדרה כללית יותר שנקראת תלות לינארית.

הגדרה 8.35. יהי V מ"ו מעל $\mathbb{F}$. קבוצת וקטורים $\{v_1,...,v_k\}\subseteq V$ נקראת:

- תלויה לינארית (ת"ל) אם קיימים $\alpha_1,...,\alpha_k\in\mathbb{F}$ לא כולם 0 כך ש-

$$\alpha_1v_1+...+\alpha_kv_k=0_V.$$

- בלתי תלויה לינארית (בת"ל) אם לכל $\alpha_1, ..., \alpha_k \in \mathbb{F}$ מתקיים

$$\alpha_1 v_1 + ... + \alpha_k v_k = 0_V \implies \alpha_1 = ... = \alpha_k = 0.$$

הערה 8.36. שימו לב שהשלילה של "תלויה לינארית" היא אכן "בלתי תלויה לינארית". או שקיים צירוף לינארי לא טריוויאלי של $v_1, ..., v_k$ השווה ל- 0_V, או שהצירוף הלינארי היחיד השווה ל- 0_V הוא הצירוף הלינארי הטריוויאלי (כל הסקלרים מתאפסים).

הערה 8.37. החידוש של תלות לינארית היא לגבי שלושה וקטורים ומעלה. לגבי שני וקטורים, תלות לינארית שקולה לקו-לינאריות: לכל $v_1, v_2 \in V$ ולכל $\alpha_1, \alpha_2 \in \mathbb{F}$ כך שלא כולם 0, נניח תחילה $\alpha_2 \neq 0$ ונקבל

$$\alpha_1 v_1 + \alpha_2 v_2 = 0_V \iff v_2 = -\frac{\alpha_1}{\alpha_2} v_1,$$

או לחילופין, אם $\alpha_1 \neq 0$ נקבל

$$\alpha_1 v_1 + \alpha_2 v_2 = 0_V \iff v_1 = -\frac{\alpha_2}{\alpha_1} v_2.$$

אז בכל מקרה: $\{ v_1, v_2 \}$ תלויה לינארית אם ורק אם הוקטורים הם קו-לינאריים.

דוגמה 8.38.

א. ב- $\mathbb{R}^3$ הקבוצה $\{ (1,2,3), (3,6,9) \}$ היא ת"ל כי הוקטורים קו-לינאריים, או באופן מפורש:

$$3 \cdot (1,2,3) + (-1) \cdot (3,6,9) = (0,0,0)$$

ב. ב- $\mathbb{R}^3$ נסתכל על הקבוצה $\{ (1,1,0), (0,1,1), (1,0,1), (1,1,1) \}$ נראה כי היא ת"ל: נניח שקיימים $\alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathbb{R}$ כך ש-

$$\alpha_1(1,1,0) + \alpha_2(0,1,1) + \alpha_3(1,0,1) + \alpha_4(1,1,1) = (0,0,0,0)$$

$$\iff \begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

נדרג את המטריצה:

$$\begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix} \xrightarrow{R_2 - R_1 \to R_2} \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix} \xrightarrow{R_3 - R_2 \to R_2} \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 2 & 1 \end{pmatrix}$$

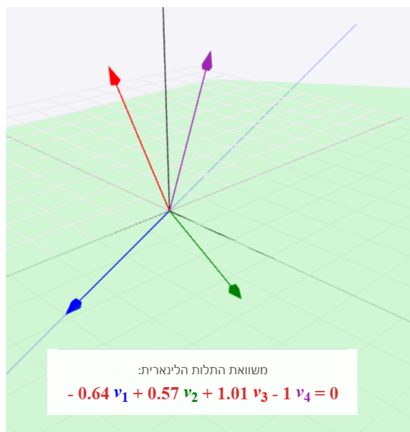
$$\xrightarrow{\frac{1}{2}R_3 \to R_3} \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & \frac{1}{2} \end{pmatrix} \xrightarrow[R_2 \to R_2 + R_3]{R_1 \to R_1 - R_3} \begin{pmatrix} 1 & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} \end{pmatrix}$$

אין צורך בקבוצת כל הפתרונות כדי לזהות תלות לינארית. כבר רואים שהפתרון אינו יחיד, ואם למשל נציב $\alpha_4 = -2$ נקבל משוואה המתארת תלות לינארית:

$$\alpha_1 = \alpha_2 = \alpha_3 = 1 \implies (1,1,0)+(0,1,1)+(1,0,1)-2(1,1,1) = (0,0,0).$$

לעומת זאת, הדירוג לעיל מראה שתת-הקבוצה $\{ (1,1,0),(0,1,1),(1,0,1) \}$ היא בת"ל. הסיבה לכך היא שאם מתעלמים מהוקטור הרביעי בקבוצה המקורית, אז לפי הדירוג מתקבלת מטריצת היחידה I_3. לכן:

$$\alpha_1(1,1,0)+\alpha_2(0,1,1)+\alpha_3(1,0,1) = (0,0,0) \iff \alpha_1 = \alpha_2 = \alpha_3 = 0.$$



איור 8.3: תלות לינארית בין ארבעה וקטורים במרחב, כאשר אף שלושה אינם על מישור אחד

תרגיל 8.39. במ"ו V מעל $\mathbb{F}$, האם ייתכן שקבוצה של וקטור יחיד היא ת"ל?

פתרון. הקבוצה $\{ 0_V \}$ היא ת"ל כי מתקיים

$$1 \cdot 0_V = 0_V$$

וזהו צירוף לינארי לא טריוויאלי. לעומת זאת, לכל $v \neq 0_V$ הקבוצה $\{ v \}$ היא בת"ל כי מתקיים

$$\alpha v = 0_V \implies \alpha = 0.$$

ניתן להסיק $\alpha = 0$ כי אחרת ניתן לחלק ב- α ולקבל סתירה.

טענה 8.40. *יהי V מ"ו מעל $\mathbb{F}$ ויהיו $v_1,...,v_k \in V$. אז $\{ v_1,...,v_k \}$ ת"ל אם ורק אם קיים $1 \leq i \leq k$ כך ש-*

$$v_i \in \text{Span}(v_1,...,v_{i-1}).$$

הערה 8.41. המשמעות היא שתלות לינארית "מקטינה את המימד" כך שהקבוצה מוכלת בתת-מרחב ממימד נמוך מגודל הקבוצה. במישור ובמרחב, קבוצה של שני וקטורים שונים היא ת"ל אם ורק אם קבוצה זו פורשת ישר (ולא מישור). במרחב, קבוצה של שלושה וקטורים שונים היא ת"ל אם ורק אם קבוצה זו פורשת מישור או ישר (ולא את כל המרחב). הרעיון הזה עובד גם במרחב וקטורי ממימד 4 ומעלה, אבל יותר קשה לדמיין את זה.

הוכחה. בכיוון הראשון, נניח שקיים $1 \leq i \leq k$ כך ש- $v_i \in \mathrm{Span}(v_1, ..., v_{i-1})$. אז קיימים $\alpha_1, ..., \alpha_{i-1} \in \mathbb{F}$ כך ש-

$$v_i = \alpha_1 v_1 + ... + \alpha_{i-1} v_{i-1} \implies \alpha_1 v_1 + ... + \alpha_{i-1} v_{i-1} + (-1) v_i + 0 \cdot v_{i+1} + ... + 0 \cdot v_k = 0_V.$$

קיבלנו צירוף לינארי השווה ל- 0_V שבו לא כל הסקלרים 0, ולכן $\{ v_1, ..., v_k \}$ ת"ל.
בכיוון השני, נניח כי $\{ v_1, ..., v_k \}$ ת"ל. אז קיימים $\alpha_1, ..., \alpha_k \in \mathbb{F}$ לא כולם 0 כך ש-

$$\alpha_1 v_1 + ... + \alpha_k v_k = 0_V.$$

נבחר את ה- i המקסימלי שמקיים $\alpha_i \neq 0$. אם $i < k$ אז $\alpha_j = 0$ לכל $i < j \leq k$. בכל מקרה מתקיים

$$\alpha_1 v_1 + ... + \alpha_i v_i = 0_V \implies \alpha_i v_i = -\alpha_1 v_1 - ... - \alpha_{i-1} v_{i-1}$$
$$\implies v_i = -\frac{\alpha_1}{\alpha_i} v_1 - ... - \frac{\alpha_{i-1}}{\alpha_i} v_{i-1} \implies v_i \in \mathrm{Span}(v_1, ..., v_{i-1})$$

□

מסקנה 8.42. *יהי V מ"ו מעל $\mathbb{F}$, ויהיו* $v_1, ..., v_k, v_{k+1}, ..., v_l \in V$.

א. אם $\{ v_1, ..., v_k \}$ ת"ל, אז $\{ v_1, ..., v_k, v_{k+1}, ..., v_l \}$ ת"ל.

ב. אם $\{ v_1, ..., v_k, v_{k+1}, ..., v_l \}$ בת"ל, אז $\{ v_1, ..., v_k \}$ בת"ל.

הוכחה. שני הסעיפים שקולים לוגית, כי "בלתי תלויים לינארית" זו השלילה של "תלויים לינארית". נוכיח את הסעיף הראשון: נניח כי $\{ v_1, ..., v_k \}$ ת"ל. אז לפי טענה **8.40** קיים $1 \leq i \leq k$ כך ש- $v_i \in \mathrm{Span}(v_1, ..., v_{i-1})$. ושוב לפי הטענה, זה גם מראה כי $\{ v_1, ..., v_k, v_{k+1}, ..., v_l \}$ ת"ל. □

הערה 8.43. במילה פשוטות: המשמעות של תלות לינארית היא שלפחות אחד הוקטורים נפרש ע"י הוקטורים האחרים. אם נגדיל קבוצת וקטורים ת"ל, היא תישאר ת"ל. אם נקטין קבוצת וקטורים בת"ל, היא תישאר בת"ל.

מסקנה 8.44. *יהי V מ"ו מעל $\mathbb{F}$, ויהיו $\{ v_1, ..., v_k \} \in V$ בת"ל ו- $w \in V$ כך שמתקיים $w \notin \mathrm{Span}(v_1, ..., v_k)$. אז $\{ v_1, ..., v_k, w \}$ בת"ל.*

הוכחה. נניח בשלילה כי $\{ v_1, ..., v_k, w \}$ ת"ל. אז לפי טענה **8.40**, אחד הוקטורים נפרש ע"י הוקטורים הקודמים לו. נתון כי $w \notin \operatorname{Span}(v_1, ...v_k)$, ולכן קיים $1 \leq i \leq k$ כך ש-$v_i \in \operatorname{Span}(v_1, ..., v_{i-1})$. אז $\{ v_1, ..., v_k \}$ ת"ל וזו סתירה להנחה. $\square$

המסקנה הבאה מסבירה איך ניתן לחלץ מקבוצה הפורשת מ"ו תת-קבוצה שהיא בת"ל וגם פורשת.

מסקנה 8.45. *אם $\{ v_1, ..., v_k \}$ קבוצה הפורשת מ"ו V מעל $\mathbb{F}$, אז תת-הקבוצה*

$$\{ v_i \mid 1 \leq i \leq k, v_i \notin \operatorname{Span}(v_1, ..., v_{i-1}) \}$$

היא בת"ל ופורשת את V.

הערה 8.46. המסקנה תקפה גם לגבי קבוצה הפורשת תת-מרחב W, שהוא מ"ו בפני עצמו.

הוכחה. בלי הגבלת הכלליות, קיים $1 \leq l \leq k$ כך ש-

$$\{ v_1, ..., v_l \} = \{ v_i \mid 1 \leq i \leq k, v_i \notin \operatorname{Span}(v_1, ..., v_{i-1}) \}$$

(הנחנו שהוקטורים התלויים בקודמיהם מופיעים בסוף הרשימה, וניתן לדאוג לכך ע"י שינוי האינדקסים שלהם). אז לפי טענה **8.40**, זו קבוצה בת"ל כי אחרת קיים $1 \leq i \leq l$ כך ש-$v_i \in \operatorname{Span}(v_1, ..., v_{i-1})$ וזו סתירה להגדרה של תת-הקבוצה.
כדי להראות כי $V = \operatorname{Span}(v_1, ..., v_l)$, די לבדוק שמתקיים $v_{l+1}, ..., v_k \in \operatorname{Span}(v_1, ..., v_l)$ כי אז ינבע לפי טענה **8.26**

$$V = \operatorname{Span}(v_1, ..., v_k) \subseteq \operatorname{Span}(v_1, ..., v_l) \subseteq V$$

ומכאן השוויון. נשים לב שלפי ההגדרה לכל $l+1 \leq i \leq k$ מתקיים $v_i \in \operatorname{Span}(v_1, ..., v_{i-1})$. ולכן

$$\begin{aligned} &v_{l+1} \in \operatorname{Span}(v_1, ..., v_l) \implies v_{l+2} \in \operatorname{Span}(v_1, ..., v_l, v_{l+1}) = \operatorname{Span}(v_1, ..., v_l) \\ &\implies \cdots \implies v_k \in \operatorname{Span}(v_1, ..., v_{k-1}) = \operatorname{Span}(v_1, ..., v_l) \end{aligned}$$

$\square$

תרגיל 8.47. ב- $\mathbb{R}^4$ נסתכל על הקבוצה $\{ (1,1,0,1), (0,1,1,1), (1,2,1,2), (2,1,-1,1) \}$ הפורשת תת-מרחב W. מצאו תת-קבוצה שהיא בסיס ל- W.

פתרון. נרכיב את המטריצה שעמודותיה הן הוקטורים:

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

נדרג אותה כדי לזהות תלות לינארית:

$$\begin{pmatrix}1&0&1&2\\1&1&2&1\\0&1&1&-1\\1&1&2&1\end{pmatrix}\xrightarrow[R_4-R_1\to R_4]{R_2-R_1\to R_2}\begin{pmatrix}1&0&1&2\\0&1&1&-1\\0&1&1&-1\\0&1&1&-1\end{pmatrix}\xrightarrow[R_4-R_2\to R_4]{R_3-R_2\to R_3}\begin{pmatrix}1&0&1&2\\0&1&1&-1\\0&0&0&0\\0&0&0&0\end{pmatrix}$$

הדרגה היא 2 והדירוג מראה ששתי העמודות הימניות נפרשות ע"י שתי העמודות השמאליות. למעשה, לפי הדירוג מתקיים

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

קבוצת שתי העמודות השמאליות (שמתאימות לאיברים מובילים בסוף הדירוג) היא בת"ל, ולפי מסקנה 8.45 נובע כי $\{\,(1,1,0,1),(0,1,1,1)\,\}$ היא בסיס ל- W.

8.5 בסיסים

יש חשיבות לקבוצת וקטורים בת"ל שגם פורשת את המרחב הוקטורי הנתון. באמצעות קבוצה כזו ניתן לפרוש כל וקטור במרחב הוקטורי, ובנוסף אין אף וקטור "מיותר" בקבוצה (כי אין תלות לינארית). מכאן ההגדרה הבאה:

הגדרה 8.48. יהי V מ"ו מעל $\mathbb{F}$. תת-קבוצה $\{\,v_1,...,v_n\,\}\subseteq V$ נקראת בסיס ל- V אם היא פורשת את V וגם בת"ל.

הערה 8.49. נראה בהמשך שאם קיים ל- V בסיס אחד, אז למעשה קיימים לו אינסוף בסיסים ולכולם אותו הגודל (n בסימון של ההגדרה). גודל זה נקרא המימד של V, ונגדיר אותו באופן מסודר בהמשך.

דוגמה 8.50. נחזור לדוגמה 8.28 ונראה שהקבוצות הפורשות הן גם בת"ל, ולכן הן בסיסים.

א. ראינו כי הקבוצה $\{\,e_1,...,e_n\,\}$ פורשת את $\mathbb{F}^n$. נראה כי היא גם בת"ל: נניח שקיימים $\alpha_1,...,\alpha_n\in\mathbb{F}$ כך ש- $\alpha_1e_1+...+\alpha_ne_n=0$. נשים לב כי

$$\alpha_1e_1+...+\alpha_ne_n=(\alpha_1,...,\alpha_n)=(0,...,0)$$

ולכן $\{\,e_1,...,e_n\,\}$ היא גם בת"ל, כלומר היא בסיס ל- $\mathbb{F}^n$. היא נקראת הבסיס הסטנדרטי שלו, ונשים לב שגודלה הוא n (ולכן בהמשך נסיק שהמימד של $\mathbb{F}^n$ הוא n, כצפוי).

ב. ראינו שהקבוצה $\{ P(i,j) \mid 1 \le i \le m, 1 \le j \le n \}$ פורשת את $\mathbb{M}_{m\times n}(\mathbb{F})$. נראה שהיא גם בת"ל: נניח שקיימים $\alpha_{11},...,\alpha_{1n},...,\alpha_{m1},...,\alpha_{mn} \in \mathbb{F}$ כך ש-

$$\alpha_{11}P(1,1) + ... + \alpha_{mn}P(m,n) = \mathbf{0}_{m\times n}$$

לפי ההגדרה של $P(i,j)$ נובע כי

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}.$$

אז כל הסקלרים בהכרח מתאפסים כנדרש. לכן $\{ P(i,j) \mid 1 \le i \le m, 1 \le j \le n \}$ היא למעשה בסיס ל- $\mathbb{M}_{m\times n}(\mathbb{F})$ והיא נקראת הבסיס הסטנדרטי שלו. גודלה הוא mn (ולכן נסיק בהמשך שזהו המימד של $\mathbb{M}_{m\times n}(\mathbb{F})$).

נרצה גם להגדיר מרחב 0- מימדי. הרעיון הוא "נקודה", אבל צריך להסביר את הקשר לבסיס.

הגדרה 8.51. יהי $V = \{ 0_V \}$. עבור קבוצת וקטורים ריקה $\emptyset = \{\ \}$ נגדיר

$$\operatorname{Span}(\emptyset) = \{ 0_V \}.$$

הערה 8.52. הרעיון הוא שאם אין אף וקטור בקבוצה, אז ברירת המחדל של הצירוף הלינארי הריק היא 0_V כי אין מה לסכום. שימו לב שהקבוצה הריקה היא בת"ל כי בהעדר וקטורים לא ניתן למצוא סקלרים (לא כולם 0 ובכלל) כך שהצירוף הלינארי יהיה שווה ל- 0_V. אז במובן הזה (שהוא די חסר תוכן) הקבוצה הריקה היא בסיס ל- $\{ 0_V \}$ וגודלה הוא 0.

שימו לב ש- $\{ 0_V \}$ אינה בסיס לעצמה כי היא ת"ל.

טענה 8.53. *יהי V מ"ו מעל $\mathbb{F}$. אז התנאים הבאים שקולים:*

- *ל- V קיים בסיס.*
- *V נוצר סופית, כלומר קיימת תת-קבוצה סופית הפורשת אותו.*

הוכחה. בכיוון הראשון, נניח כי $\{ v_1,...,v_n \}$ היא בסיס ל- V. אז בפרט $V = \operatorname{Span}(v_1,...,v_n)$, ולכן V נוצר סופית.

בכיוון השני, נניח כי V נוצר סופית ולכן קיימת $\{ v_1,...,v_k \}$ כך ש- $V = \operatorname{Span}(v_1,...,v_k)$. לפי מסקנה 8.45 קיימת תת-קבוצה בת"ל של $\{ v_1,...,v_k \}$ הפורשת את V, כלומר היא בסיס. □

טענה 8.54. *יהי V מ"ו מעל $\mathbb{F}$. יהיו $v_1,...,v_n \in V$. אז התנאים הבאים שקולים:*

- $\{\, v_1, ..., v_n \,\}$ *היא בסיס ל-* V.
- *לכל* $v \in V$ *קיימים* $\alpha_1, ..., \alpha_n \in \mathbb{F}$ *יחידים כך ש-* $v = \alpha_1 v_1 + ... + \alpha_n v_n$.

הערה 8.55. המשמעות של התנאי השני בטענה היא שקיים פתרון יחיד למשוואה הוקטורית $v = x_1 v_1 + ... + x_n v_n$. אנחנו כבר יודעים לקבוע מתי יש לממ"ל פתרון יחיד, וכאן הרעיון דומה.

הוכחה. בכיוון הראשון, נניח כי $\{\, v_1, ..., v_n \,\}$ היא בסיס ל- V. יהי $v \in V$. הקבוצה פורשת כל וקטור ב- V, ולכן קיימים $\alpha_1, ..., \alpha_n \in \mathbb{F}$ כך ש- $v = \alpha_1 v_1 ... + \alpha_n v_n$. כדי להראות שהם יחידים, נניח שקיימים בנוסף $\beta_1, ..., \beta_n \in \mathbb{F}$ כך ש- $v = \beta_1 v_1 ... + \beta_n v_n$. אז נובע כי

$$\alpha_1 v_1 + ... + \alpha_n v_n = \beta_1 v_1 ... + \beta_n v_n.$$

נחסיר את הביטוי של אגף ימין משני האגפים ונקבל

$$(\alpha_1 - \beta_1) v_1 ... + (\alpha_n - \beta_n) v_n = 0_V.$$

נתון כי $v_1, ..., v_n$ גם בת"ל ולכן

$$\forall 1 \le j \le n \quad \alpha_j - \beta_j = 0 \implies \alpha_j = \beta_j$$

כלומר הסקלרים $\alpha_1, ..., \alpha_n$ אכן יחידים.

בכיוון השני, נניח שלכל $v \in V$ קיימים סקלרים יחידים עבורם הצירוף הלינארי המתאים שווה ל- v. אז ברור מהנתון שהקבוצה $v_1, ..., v_n$ פורשת את V. כדי להראות שהיא גם בת"ל, ניקח $v = 0_V$. אז לפי הנתון קיימים סקלרים יחידים $\alpha_1, ..., \alpha_n \in \mathbb{F}$ כך ש-

$$\alpha_1 v_1 + \alpha_n v_n = 0_V.$$

מהיחידות נובע שסקלרים אלה הם בהכרח אפסים, כלומר $\alpha_j = 0$ לכל $1 \le j \le n$ כנדרש.
□

8.5.1 אפיון של בסיס לפי מטריצה

כעת נתמקד ב- $V = \mathbb{F}^n$ כדי להבין את הקשר בין המושג של בסיס לבין החומר של הפרקים הקודמים (הפיכות של מטריצה). נניח כי $v_1, ..., v_k$ וקטורי עמודה ב- $\mathbb{F}^n$. ראינו בטענה 8.32 שלכל $w \in \mathbb{F}^n$ מתקיים $w \in \mathrm{Span}(v_1, ..., v_k)$ אם ורק אם קיים פתרון לממ"ל $Ax = w$, כאשר

$$A = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix}.$$

ליתר דיוק, ראינו כי לכל $\alpha_1, ..., \alpha_k \in \mathbb{F}$ מתקיים

$$Ax = w \iff \alpha_1 v_1 + ... + \alpha_k v_k = w$$

כאשר

$$.x = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix}$$

עבור $w = 0$ ניתן לראות את הקשר לתלות לינארית. מכאן הטענה הבאה:

טענה 8.56. *יהיו $v_1, ..., v_k \in \mathbb{F}^n$. אז הוקטורים הנ״ל הם בת״ל אם ורק אם קיים פתרון יחיד לממ״ל $Ax = 0$ כאשר*

$$, , A = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_k \\ | & | & & | \end{pmatrix}$$

כלומר $.\mathrm{N}(A) = \{ 0 \}$

הוכחה. לכל $\alpha_1, ..., \alpha_k \in \mathbb{F}$ מתקיים

$$.\alpha_1 v_1 + ... + \alpha_k v_k = 0 \iff Ax = 0$$

אם הפתרון לממ״ל יחיד, אז בהכרח $\alpha_1 = ... = \alpha_k = 0$ ולכן $v_1, ..., v_k$ בת״ל. לעומת זאת, אם הפתרון לממ״ל אינו יחיד זה בדיוק אומר שלא כל הסקלרים הם 0, ולכן הוקטורים ת״ל. □

מסקנה 8.57. *יהיו* $.v_1, ..., v_k \in \mathbb{F}^n$

א. *אם* $\{ v_1, ..., v_k \}$ *בת״ל, אז* $.k \leq n$

ב. *אם* $\{ v_1, ..., v_k \}$ *פורשת את* $\mathbb{F}^n$*, אז* $.k \geq n$

ג. *אם* $\{ v_1, ..., v_k \}$ *היא בסיס ל-* $\mathbb{F}^n$*, אז* $.k = n$

הוכחה. נסתכל על

$$.A = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_k \\ | & | & & | \end{pmatrix} \in \mathbb{M}_{n \times k}(\mathbb{F})$$

א. לממ״ל $Ax = 0$ קיים פתרון יחיד אם ורק אם $\mathrm{rank}(A) = k$ (כי אחרת יש משתנה חופשי). באופן כללי, הדרגה לא עולה על מספר השורות ולכן $.k \leq n$

ב. לממ"ל $Ax = b$ קיים פתרון לכל $b \in \mathbb{F}^n$ אם ורק אם $\operatorname{rank}(A) = n$ (כי אחרת תתקבל שורת סתירה עבור b מסוים). לכן, במקרה של קבוצה פורשת נובע כי $k \geq n$ כי הדרגה לא עולה על מספר העמודות.

ג. זהו שילוב של שני המקרים הקודמים, ולכן $k = n$.

$\square$

הערה 8.58. אז הוכחנו כי כל בסיס ל- $\mathbb{F}^n$ הוא קבוצה בגודל n. נראה בהמשך שגם במרחבים וקטוריים נוספים גודל הבסיס קבוע. גודל זה נקרא מימד והוא אחד מהמאפיינים של מרחב וקטורי.

המקרה $k = n$ הוא המקרה של מטריצה ריבועית, ומכאן הקשר למטריצה הפיכה. נרחיב את משפט **6.75** ע"י הוספת תנאים שקולים:

משפט 8.59. *לכל $A \in \mathbb{M}_{n\times n}(\mathbb{F})$ התנאים הבאים שקולים:*

(א) A *שקולת שורה ל-* I_n.

(ב) $\operatorname{rank}(A) = n$

(ג) *לממ"ל* $Ax = 0$ *יש פתרון יחיד, כלומר* $\mathrm{N}(A) = \{ 0 \}$.

(ד) *לכל* $b \in \mathbb{F}^n$ *יש פתרון יחיד לממ"ל* $Ax = b$.

(ה) *לכל* $b \in \mathbb{F}^n$ *יש פתרון לממ"ל* $Ax = b$.

(ו) A *הפיכה משמאל.*

(ז) A *הפיכה מימין.*

(ח) A *הפיכה.*

(ט) $\det(A) = 0$

(י) *קבוצת העמודות של* A *היא בת"ל.*

(יא) *קבוצת העמודות של* A *פורשת את* $\mathbb{F}^n$*, כלומר* $\operatorname{Col}(A) = \mathbb{F}^n$.

(יב) *קבוצת העמודות של* A *היא בסיס ל-* $\mathbb{F}^n$.

(יג) *קבוצת השורות של* A *היא בת"ל.*

(יד) *קבוצת השורות של* A *פורשת את* $\mathbb{F}^n$*, כלומר* $\operatorname{Row}(A) = \mathbb{F}^n$.

(טו) קבוצת השורות של A היא בסיס ל- $\mathbb{F}^n$.

הוכחה. תשעת הסעיפים הראשונים שקולים לפי 6.75 ו- 7.15. לפי 8.32 ו- 8.56 מתקיים

$$(\text{יא}) \iff (\text{ה}) \iff (\text{ג}) \iff (\text{י})$$

מכאן שגם (יב) שקול לכל התנאים הקודמים, כי הוא שילוב של (י) ו- (יא).

בשביל שלושת הסעיפים האחרונים נסתכל על A^t. (יג) – (טו) שקולים להפיכות של A^t, ולפי 6.81, A^t הפיכה אם ורק אם A הפיכה. לכן, (יג) – (טו) שקולים לכל שאר הסעיפים. $\square$

המשפט לעיל מראה את הקשר בין כל המושגים העיקריים שלמדנו: קבוצת הפתרונות של ממ"ל, הפיכות מטריצה, דטרמיננטה ותכונות של קבוצת וקטורים במרחב וקטורי. נרצה להתמקד רק בחלק האחרון, ולכן ננסח משפט שניסוחו לא מתייחס למטריצה:

משפט 8.60. *תהי $B \subseteq \mathbb{F}^n$. אז התנאים הבאים שקולים:*

א. B בת"ל בגודל n.

ב. B פורשת את $\mathbb{F}^n$ וגודלה n.

ג. B היא בסיס ל- $\mathbb{F}^n$.

הערה 8.61. מקובל לקרוא למשפט זה "שלישי חינם" כי למעשה יש פה שלושה תת-תנאים: קבוצה בת"ל, קבוצה פורשת, וגודל n. לפי משפט זה, כל שניים מתת-התנאים אלה גוררים את תת-התנאי השלישי, והמשמעות של כל השלושה היא בסיס ל- $\mathbb{F}^n$.

הוכחה. כל אחד מתנאי המשפט קובע שהגודל של B הוא n (זהו הגודל של כל בסיס ל- $\mathbb{F}^n$ לפי מסקנה 8.57). אז נסמן $B = \{ v_1, ..., v_n \}$ ונסתכל על המטריצה

$$A = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix}.$$

שלושת התנאים הם בדיוק סעיפים (יא) – (יג) במשפט 8.59. אז הם שקולים.

$\square$

8.5.2 חילוץ קבוצה בת"ל והשלמה לבסיס

בהינתן $\{ v_1, ..., v_k \} \subseteq \mathbb{F}^n$ נרצה לפתח מתכון לחילוץ תת-קבוצה בת"ל והשלמתה לבסיס. ראינו דרך אחת לחלץ קבוצה בת"ל, וזאת לפי המטריצה שעמודותיה הן הוקטורים האלה.

דוגמה 8.62. עבור

$$\{ v_1, ..., v_5 \} = \{ (1,1,0,0), (0,1,1,0), (0,0,1,1), (1,3,2,0), (3,3,-1,-1) \}$$

נרצה למצוא בסיס ל- $W = \mathrm{Span}(v_1, ..., v_5)$. נכתוב את הוקטורים כעמודות מטריצה (לפי הסדר) ונדרג אותה:

$$A = \begin{pmatrix} 1 & 0 & 0 & 1 & 3 \\ 1 & 1 & 0 & 3 & 3 \\ 0 & 1 & 1 & 2 & -1 \\ 0 & 0 & 1 & 0 & -1 \end{pmatrix} \xrightarrow{\mathbb{R}_2 - R_1 \to R_2} \begin{pmatrix} 1 & 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & 2 & 0 \\ 0 & 1 & 1 & 2 & -1 \\ 0 & 0 & 1 & 0 & -1 \end{pmatrix}$$
$$\xrightarrow{\mathbb{R}_3 - R_2 \to R_3} \begin{pmatrix} 1 & 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 \end{pmatrix} \xrightarrow{R_4 - R_3 \to R_4} \begin{pmatrix} 1 & 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

קיבלנו שלושה איברים מובילים רק בשלוש העמודות הראשונות, ולכן הדרגה היא 3. זה מראה כי מלכתחילה $v_4, v_5 \in \mathrm{Span}(v_1, v_2, v_3)$, ולמעשה $v_4 = v_1 + 2v_2, v_5 = 3v_1 - v_3$.

לכן $\{ v_1, v_2, v_3 \}$ היא בסיס ל- W.

החסרון של שיטה זו היא שאם נרצה להשלים את $\{ v_1, v_2, v_3 \}$ לבסיס ל- $\mathbb{R}^4$, הדירוג לא מראה איך לעשות זאת. למעשה, פעולות הדירוג משנות את מרחב העמודות $\mathrm{Col}(A)$. אפשר "לנחש" וקטור נוסף, כמו $(0,0,0,1)$, ולבדוק שאכן מתקבל בסיס. זה פתרון לגיטימי כל עוד הבדיקה נכונה, אבל נרצה למצוא שיטה חלופית שלא דורשת שום ניחוש. למזלנו, ניתן גם לדרג את המטריצה A^t ששורותיה הן הוקטורים המקוריים.

$$\begin{pmatrix}1&1&0&0\\0&1&1&0\\0&0&1&1\\1&3&2&0\\3&3&-1&-1\end{pmatrix}\xrightarrow{R_4-R_1\to R_4}\begin{pmatrix}1&1&0&0\\0&1&1&0\\0&0&1&1\\0&2&2&0\\3&3&-1&-1\end{pmatrix}\xrightarrow{R_5-3R_1\to R_5}\begin{pmatrix}1&1&0&0\\0&1&1&0\\0&0&1&1\\0&2&2&0\\0&0&-1&-1\end{pmatrix}$$

$$\begin{pmatrix}1&1&0&0\\0&1&1&0\\0&0&1&1\\0&2&2&0\\0&0&-1&-1\end{pmatrix}\xrightarrow{R_1-R_2\to R_1}\begin{pmatrix}1&0&-1&0\\0&1&1&0\\0&0&1&1\\0&2&2&0\\0&0&-1&-1\end{pmatrix}\xrightarrow{R_4-2R_2\to R_4}\begin{pmatrix}1&0&-1&0\\0&1&1&0\\0&0&1&1\\0&0&0&0\\0&0&-1&-1\end{pmatrix}$$

$$\begin{pmatrix}1&0&-1&0\\0&1&1&0\\0&0&1&1\\0&0&0&0\\0&0&-1&-1\end{pmatrix}\xrightarrow{R_1+R_3\to R_1}\begin{pmatrix}1&0&0&1\\0&1&1&0\\0&0&1&1\\0&0&0&0\\0&0&-1&-1\end{pmatrix}\xrightarrow{R_2-R_3\to R_2}\begin{pmatrix}1&0&0&1\\0&1&0&-1\\0&0&1&1\\0&0&0&0\\0&0&-1&-1\end{pmatrix}$$

$$\begin{pmatrix}1&0&0&1\\0&1&0&-1\\0&0&1&1\\0&0&0&0\\0&0&-1&-1\end{pmatrix}\xrightarrow{R_5+R_3\to R_5}\begin{pmatrix}1&0&0&1\\0&1&0&-1\\0&0&1&1\\0&0&0&0\\0&0&0&0\end{pmatrix}$$

ההתאפסות של שתי השורות האחרונות מראה כי הן נפרשות ע"י שלוש השורות הראשונות, שמתאימות לאיברים מובילים בסוף הדירוג. לכן, שוב אנחנו רואים כי $\{\, v_1, v_2, v_3 \,\}$ היא בסיס ל-W. כדי להשלים קבוצה זו לבסיס לכל $\mathbb{R}^4$ ניתן (וזה נוח מאוד) להוסיף את וקטור היחידה e_4, כי אין איבר מוביל בעמודה הרביעית של המטריצה המדורגת קנונית. נדגיש שאין פה ניחוש: אם ניפטר משורות האפסים ונוסיף את e_4, נקבל את המטריצה

$$\begin{pmatrix}1&0&0&1\\0&1&0&-1\\0&0&1&1\\0&0&0&1\end{pmatrix}.$$

זו מטריצה הפיכה (הדרגה מלאה), ולפי משפט 8.59 נובע כי קבוצת שורותיה היא בסיס ל-$\mathbb{R}^4$. באותה מידה, זה גם מוצדק להחליף את שתי השורות הראשונות בוקטורים v_1, v_2 המקוריים ולקבל בסיס $\{\, v_1, v_2, v_3, e_4 \,\}$ ל-$\mathbb{R}^4$.

אז למה המסקנה מוצדקת לפי הדירוג החדש (עם שורות) ולא הקודם (עם עמודות)? כי דירוג לא משפיע על מרחב השורות (הגדרה 8.30), בניגוד למרחב העמודות.

טענה 8.63. *לכל* $A \in \mathbb{M}_{m\times n}(\mathbb{F})$ *ולכל פש"א* S*, מתקיים* $\mathrm{Row}(S(A)) = \mathrm{Row}(A)$.

הוכחה. תהי $A \in \mathbb{M}_{m\times n}(\mathbb{F})$ עם שורות $R_1, ..., R_m \in \mathbb{F}^n$, ותהי S פש"א. אם מדובר בהחלפה בין שורות, ברור כי $\mathrm{Row}(S(A)) = \mathrm{Row}(A)$ שהרי אין חשיבות לסדר הוקטורים הפורשים (לפי חוק החילוף). בכל מקרה אחר, S משפיעה רק על שורה אחת R_i.

נוכיח כי $\mathrm{Row}(S(A)) \subseteq \mathrm{Row}(A)$: לפי טענה 8.26, מספיק לבדוק שהשורה החדשה R_i' במקום R_i מקיימת $R_i' \in \mathrm{Span}(R_1, ..., R_m)$. נשארו שני מקרים:

- $S = (\alpha R_i \to R_i)$: במקרה זה מתקיים

$$R_i' = \alpha R_i \in \mathrm{Span}(R_i) \subseteq (R_1, ..., R_m).$$

- $S = (R_i + \alpha R_j \to R_i)$: במקרה זה מתקיים

$$R_i' = R_i + \alpha R_j \in \mathrm{Span}(R_i, R_j) \subseteq (R_1, ..., R_m).$$

נוכיח באופן דומה כי $\mathrm{Row}(S(A)) \supseteq \mathrm{Row}(A)$: לפי טענה 8.26, מספיק לבדוק שהשורה המקורית R_i מקיימת $R_i \in \mathrm{Span}(R_1, ..., R_i', ..., R_m)$. נבדוק את שני המקרים:

- $S = (\alpha R_i \to R_i)$: במקרה זה מתקיים

$$R_i = \frac{1}{\alpha} R_i' \in \mathrm{Span}(R_i') \subseteq (R_1, ..., R_i', ..., R_m).$$

- $S = (R_i + \alpha R_j \to R_i)$: במקרה זה מתקיים

$$R_i = R_i' - \alpha R_j \in \mathrm{Span}(R_i', R_j) \subseteq (R_1, ..., R_i', ..., R_m).$$

□

מסקנה 8.64. *יהיו* $A, B \in \mathbb{M}_{m\times n}(\mathbb{F})$ *שקולות שורה. אז* $\mathrm{Row}(A) = \mathrm{Row}(B)$.

הוכחה. לפי הנתון קיימות פש"אות $S_1, ..., S_k$ כך ש- $B = S_k(...(S_1(A))$, או באופן שקול קיימות מטריצות אלמנטריות $E_1, ..., E_k \in \mathbb{M}_{m\times m}$ כך ש- $B = E_k \cdots E_1 A$. אז לפי טענה 8.63 נובע כי

$$\mathrm{Row}(A) = \mathrm{Row}(E_1 A) = ... = \mathrm{Row}(E_k \cdots E_1 A) = \mathrm{Row}(B).$$

□

תרגיל 8.65. יהיו $A, B \in \mathbb{M}_{n\times n}(\mathbb{F})$ מטריצות ריבועיות. הוכיחו כי:

א. $\text{Row}(BA) \subseteq \text{Row}(A)$.

ב. אם B הפיכה, אז מתקיים $\text{Row}(BA) = \text{Row}(A)$.

פתרון.

א. נראה כי כל שורה של BA היא צירוף לינארי של שורות A, ולכן ינבע מטענה 8.26 כי $\text{Row}(BA) \subseteq \text{Row}(A)$. נסמן את שורות A ב- $R_1, ..., R_n$, ובנוסף נסמן

$$B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \vdots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix}.$$

אז לפי הגדרת הכפל, לכל $1 \leq i \leq n$ השורה ה- i של BA היא בדיוק הוקטור שנתון ע"י הצירוף הלינארי הבא:

$$\sum_{j=1}^{n} b_{ij} R_j \in \text{Span}(R_1, ..., R_n) = \text{Row}(A)$$

ב. כאשר B הפיכה, B היא מכפלת מטריצות אלמנטריות ובפרט BA שקולת שורה ל- A. אז לפי מסקנה 8.64 נובע כי $\text{Row}(BA) = \text{Row}(A)$.

8.6 קוארדינטות

יש לנו כבר הבנה טובה לגבי פרישה, תלות לינארית ובסיסים ב- $\mathbb{F}^n$, אבל מה לגבי מרחב וקטורי כללי? נרצה להחיל את הטענות והמשפטים שהוכחנו לגבי $\mathbb{F}^n$ על מרחבים וקטוריים נוספים. לשם כך, נגדיר כלי שיאפשר לנו לייצג וקטורים במ"ו V כוקטורים ב- $\mathbb{F}^n$ עבור $n \in \mathbb{N}$ מתאים (שתלוי ב- V).

טענה 8.66. *בסיס סדור* $B = \{ v_1, ..., v_n \}$ *למ"ו* V *הוא בסיס יחד עם סדר על איבריו (כלומר יש חשיבות לסדר כתיבתם).*

הערה 8.67. נזכיר שבקבוצה רגילה אין חשיבות לסדר האיברים. אז מקודם שכדיברנו על בסיס אמנם השתמשנו באינדקסים, אבל לא הייתה חשיבות לסדר. עכשיו זה ישתנה, ומכאן הצורך בהגדרה החדשה.

הסדר של הוקטורים בבסיס סדור מזכיר את סדר הקוארדינטות של וקטור ב- $\mathbb{F}^n$. בפרט, לכל שני בסיסים $B_1 = \{ v_1, ..., v_n \}, C = \{ w_1, ..., w_n \}$ מתקיים

$$B = C \iff \forall 1 \leq i \leq n \quad v_i = w_i.$$

דוגמה 8.68. נגדיר $B_1 = \{ (1,0), (0,1) \}$ וגם $B_2 = \{ (0,1), (1,0) \}$. אין הבדל ביניהם כבסיסים, אבל $B_1 \neq B_2$ כבסיסים סדורים. אז יש חשיבות להקשר.

ניזכר בטענה **8.54**. לאור טענה זו, ההגדרה הבאה מתבקשת:

הגדרה 8.69. יהי V מ"ו מעל $\mathbb{F}$, ויהי $B = \{ v_1, ..., v_n \}$ בסיס סדור ל- V. לכל $v \in V$ קיימים $\alpha_1, ..., \alpha_n \in \mathbb{F}$ יחידים כך ש- $v = \alpha_1 v_1 + ... + \alpha_n v_n$. אז עבור v נגדיר את וקטור הקואורדינטות ביחס ל- B ע"י

$$[v]_B = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{F}^n.$$

דוגמה 8.70.

א. ב- $\mathbb{R}^3$ נסתכל על הבסיסים הסדורים

$$E = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}, B = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

E הוא הבסיס הסטנדרטי. הקואורדינטות ביחס אליו הן הקואורדינטות הרגילות: לכל $(a,b,c) \in \mathbb{R}^3$ מתקיים

$$\left[\begin{pmatrix} a \\ b \\ c \end{pmatrix} \right]_E = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

לעומת זאת, הסדר ב- B שונה. הפעם מתקיים

$$\left[\begin{pmatrix} a \\ b \\ c \end{pmatrix} \right]_B = \begin{pmatrix} c \\ a \\ b \end{pmatrix}$$

כי

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = c \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + a \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

ב. ב- $\mathbb{M}_{2\times 2}(\mathbb{R})$ נסתכל על הבסיס הסטנדרטי

$$E = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

ועל הבסיס הסדור

$$.B = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

אז לכל $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{M}_{2\times 2}(\mathbb{R})$ **מתקיים**

$$\left[\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right]_E = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$$

וגם

$$.\left[\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right]_B = \begin{pmatrix} a \\ c \\ b \\ d \end{pmatrix}$$

השוויון השני נובע מהשוויון הבא:

$$.\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + c\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + d\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

ג. ב- $\mathbb{F}^n$ **הבסיס הסטנדרטי הוא**

$$.E = \{ e_1, ..., e_n \}$$

לכל

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{F}^n$$

מתקיים

$$.\left[\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}\right]_E = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

תרגיל 8.71. ב- $\mathbb{R}^2$ **נסתכל על הבסיס הסדור**

$$B = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \end{pmatrix} \right\}$$

חשבו את $.\left[\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right]_B$

פתרון. נמצא $\alpha_1, \alpha_2 \in \mathbb{R}$ כך ש-

$$\begin{pmatrix}1\\1\end{pmatrix} = \alpha_1 \begin{pmatrix}1\\2\end{pmatrix} + \alpha_2 \begin{pmatrix}2\\-1\end{pmatrix} \iff \begin{pmatrix}1 & 2\\2 & -1\end{pmatrix}\begin{pmatrix}\alpha_1\\\alpha_2\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix}.$$

ניתן לפתור את הממ"ל ע"י דירוג, אבל נוח להשתמש בנוסחה מטענה 7.10 לחישוב מטריצה הופכית. כך נקבל

$$\begin{pmatrix}1 & 2\\2 & -1\end{pmatrix}^{-1} = -\frac{1}{5}\begin{pmatrix}-1 & -2\\-2 & 1\end{pmatrix} \implies \begin{pmatrix}\alpha_1\\\alpha_2\end{pmatrix} = -\frac{1}{5}\begin{pmatrix}-1 & -2\\-2 & 1\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}\frac{3}{5}\\\frac{1}{5}\end{pmatrix}$$

אז המסקנה מהחישוב היא שמתקיים

$$\left[\begin{pmatrix}1\\1\end{pmatrix}\right]_B = \begin{pmatrix}\frac{3}{5}\\\frac{1}{5}\end{pmatrix}.$$

לוקטור קוארדינטות יש תכונות בסיסיות אך חשובות:

טענה 8.72. *יהיו V מ"ו מעל $\mathbb{F}$ ו- $B = \{ v_1, ..., v_n \}$ בסיס סדור ל- V. אז לכל $v, w \in V$ ו- $\alpha \in F$. מתקיים:*

א.

$$[v+w]_B = [v]_B + [w]_B$$

ב.

$$[\alpha v]_B = \alpha[v]_B$$

ג.

$$[v]_B = [w]_B \iff v = w$$

ד.

$$[0_V]_B = \begin{pmatrix}0\\\vdots\\0\end{pmatrix}$$

הוכחה. נסמן

$$[v]_B = \begin{pmatrix}a_1\\\vdots\\a_n\end{pmatrix}, [w]_B = \begin{pmatrix}b_1\\\vdots\\b_n\end{pmatrix}.$$

א. לפי הסימונים מתקיים

$$v = a_1v_1 + ... + a_nv_n, w = b_1v_1 + ... + b_nv_n,$$

ולכן

$$v + w = (a_1 + b_1)v_1 + ... + (a_n + b_n)v_n.$$

אז נובע כי

$$[v + w]_B = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = [v]_B + [w]_B.$$

ב. מתקיים

$$\alpha v = \alpha(a_1v_1 + ... + a_2v_2) = \alpha a_1v_1 + ... + \alpha a_2v_2,$$

ולכן

$$[\alpha v]_B = \begin{pmatrix} \alpha a_1 \\ \vdots \\ \alpha a_n \end{pmatrix} = \alpha \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \alpha[v]_B.$$

ג. אם $v = w$, אז מתקיים $[v]_B = [w]_B$ לפי ההגדרה של וקטור קואורדינטות (הסקלרים נקבעים ביחידות). להיפך, אם $[v]_B = [w]_B$ נובע כי

$$v = a_1v_1 + ... + a_nv_n = b_1v_1 + ... + b_nv_n = w.$$

ד. מתקיים

$$0_V = 0 \cdot v_1 + ... + 0 \cdot v_n \implies [0_V]_B = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

$\square$

מסקנה 8.73. *יהיו V מ"ו מעל $\mathbb{F}$, $B = \{ v_1, ..., v_n \}$ בסיס סדור ל- V, ו- $u, w_1, ..., w_k \in V$. אז מתקיים*

$$u \in \mathrm{Span}(w_1, ..., w_k) \iff [u]_B \in \mathrm{Span}([w_1]_B, ..., [w_k]_B).$$

הוכחה. לפי טענה 8.72, כל $\alpha_1, ..., \alpha_k \in \mathbb{F}$ מתקיים:

$$\begin{aligned} u = \alpha_1 w_1 + ... + \alpha_k w_k &\iff [u]_B = [\alpha_1 w_1 + ... + \alpha_k w_k]_B \\ &\iff [u]_B = \alpha_1 [w_1]_B + ... + \alpha_k [w_k]_B \end{aligned}$$

אז מתקיים

$$u \in \mathrm{Span}(w_1, ..., w_k) \iff [u]_B \in \mathrm{Span}([w_1]_B, ..., [w_k]_B).$$

□

משפט 8.74. *יהיו V מ"ו מעל $\mathbb{F}$, $B = \{ b_1, ..., b_n \}$ בסיס סדור ל- V, ו- $v_1, ..., v_k \in V$.*

א. *$\{ v_1, ..., v_k \}$ בת"ל ב- V אם ורק אם $\{ [v_1]_B, ..., [v_k]_B \}$ בת"ל ב- $\mathbb{F}^n$.*

ב. *$\{ v_1, ..., v_k \}$ פורשת את V אם ורק אם $\{ [v_1]_B, ..., [v_k]_B \}$ פורשת את $\mathbb{F}^n$.*

ג. *$\{ v_1, ..., v_k \}$ בסיס ל- V אם ורק אם $\{ [v_1]_B, ..., [v_k]_B \}$ בסיס ל- $\mathbb{F}^n$.*

הוכחה. **א.** לפי טענה 8.72, לכל $\alpha_1, ..., \alpha_k \in \mathbb{F}$ מתקיים:

$$\begin{aligned} \alpha_1 v_1 + ... + \alpha_k v_k = 0_V &\iff [\alpha_1 v_1 + ... + \alpha_k v_k]_B = [0_V]_B \\ &\iff \alpha_1 [v_1]_B + ... + \alpha_k [v_k]_B = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \end{aligned}$$

לכן, $\{ v_1, ..., v_k \}$ בת"ל ב- V אם ורק אם $\{ [v_1]_B, ..., [v_k]_B \}$ בת"ל ב- $\mathbb{F}^n$.

ב. לפי 8.73, לכל $v \in V$ מתקיים

$$v \in \mathrm{Span}(v_1, ..., v_k) \iff [v]_B \in \mathrm{Span}([v_1]_B, ..., [v_k]_B).$$

כל וקטור ב- $\mathbb{F}^n$ הוא מהצורה $[v]_B$, ולכן נובע כי

$$V = \mathrm{Span}(v_1, ..., v_k) \iff \mathbb{F}^n = \mathrm{Span}([v_1]_B, ..., [v_k]_B).$$

ג. סעיף זה נובע ישירות משילוב שני הסעיפים הקודמים.

□

דוגמה 8.75. נסתכל על הקבוצה הבאה ב- $\mathbb{M}_{2\times 2}(\mathbb{R})$:

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \right\}$$

נבדוק אם היא בסיס ל- $\mathbb{M}_{2\times 2}(\mathbb{R})$ לפי משפט **8.74**. לשם כך נצטרך להשתמש בבסיס ידוע, הלא הוא הבסיס הסטנדרטי

$$E = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

נחשב את וקטורי הקוארדינטות של הקבוצה הנתונה לפי הבסיס E:

$$\left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right]_E = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \left[\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\right]_E = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right]_E = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad \left[\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}\right]_E = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

כדי לקבוע אם וקטורי הקוארדינטות מהווים בסיס ל- $\mathbb{R}^4$, נסתכל על המטריצה שאלה עמודותיה (אפשר גם שורותיה) ונחשב דטרמיננטה:

$$\det \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} \overset{\substack{R_1+R_3\to R_1 \\ R_2+R_4\to R_2}}{=\!=} \det \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} \overset{R_1-R_2\to R_1}{=\!=} \det \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$
$$= 0$$

לפי משפט **8.59**, וקטורי הקוארדינטות (עמודות המטריצה) לא מהווים בסיס ל- $\mathbb{R}^4$. לפי משפט **8.74**, נובע כי הקבוצה המקורית של המטריצות לא מהווה בסיס ל- $\mathbb{M}_{2\times 2}(\mathbb{R})$. יותר מכך, קבוצה זו אינה בת"ל וגם אינה פורשת.

8.7 מימדים

המסקנה הבאה היא הכללה של מסקנה 8.57 למרחבים וקטוריים כלליים:

מסקנה 8.76. *יהיו V מ"ו מעל $\mathbb{F}$, $B = \{ b_1, ..., b_n \}$ בסיס ל- V ו- $v_1, ..., v_k \in V$.*

א. *אם $\{ v_1, ..., v_k \}$ בת"ל, אז $k \leq n$.*

ב. *אם $\{ v_1, ..., v_k \}$ פורשת את V, אז $k \geq n$.*

ג. *אם $\{ v_1, ..., v_k \}$ היא בסיס ל- V, אז $k = n$.*

הוכחה. נסתכל על וקטורי הקואורדינטות ביחס ל- B ונשתמש במשפט 8.74 ומסקנה 8.57.

א. נניח כי $\{ v_1, ..., v_k \}$ בת"ל. אז $\{ [v_1]_B, ..., [v_k]_B \}$ בת"ל ב- $\mathbb{F}^n$, ולכן $k \leq n$.

ב. נניח $\{ v_1, ..., v_k \}$ פורשת את V. אז $\{ [v_1]_B, ..., [v_k]_B \}$ פורשת את $\mathbb{F}^n$, ולכן $k \geq n$.

ג. נניח כי $\{ v_1, ..., v_k \}$ בסיס ל- V. אז משני הסעיפים הקודמים נובע כי $k \leq n$ וגם $k \geq n$, ולכן $k = n$.

□

מסקנה 8.77. *יהי V מ"ו נוצר סופית, כלומר קיים לו בסיס. אז כל שני בסיסים ל- V הם שווי גודל.*

לאור המסקנה, נוכל להגדיר מימד באופן כללי:

הגדרה 8.78. גודל הבסיס של מ"ו V נקרא המימד של V. מסמנים $\dim(V)$.

דוגמה 8.79. לפי 8.28, לכל $m, n \in \mathbb{N}$ מתקיים $\dim(\mathbb{F}^n) = n$ וגם $\dim(\mathbb{M}_{m\times n}(\mathbb{F})) = mn$.

תרגיל 8.80.

א. האם קיימת $\{ A_1, A_2, A_3, A_3, A_4, A_5 \} \subseteq \mathbb{M}_{2\times 2}(\mathbb{F})$ שהיא בת"ל?

ב. האם קיימת $\{ B_1, B_2, B_3 \} \subseteq \mathbb{M}_{2\times 2}(\mathbb{R})$ שפורשת את $\mathbb{M}_{2\times 2}(\mathbb{R})$?

פתרון. נשים לב שמתקיים $\dim(\mathbb{M}_{2\times 2}(\mathbb{R})) = 4$.

א. התשובה היא לא. לפי מסקנה 8.76, הגודל המקסימלי של קבוצה בת"ל ב- $\mathbb{M}_{2\times 2}(\mathbb{R})$ הוא 4.

ב. התשובה היא לא. לפי מסקנה 8.76, הגודל המינימלי של קבוצה שפורשת את $\mathbb{M}_{2\times 2}(\mathbb{R})$ הוא 4.

תרגיל 8.81. נגדיר תת-מרחב ב- $\mathbb{M}_{2\times 2}(\mathbb{C})$ ע"י

$$W = \operatorname{Span}\left(\begin{pmatrix}1 & i\\ i & 1\end{pmatrix}, \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \begin{pmatrix}2 & 1+i\\ 1+i & 2\end{pmatrix}\right).$$

מהו $\dim(W)$? מצאו בסיס ל- W והשלימו אותו לבסיס ל- $\mathbb{M}_{2\times 2}(\mathbb{C})$.

פתרון. נשתמש בוקטורי קוארדינטות לפי הבסיס הסטנדרטי E. נכתוב אותם כשורות מטריצה (כי המטרה היא להשלים את הבסיס), נדרג אותה ונקבל

$$\begin{pmatrix}1 & i & i & 1\\ 0 & 1 & 1 & 0\\ 2 & 1+i & 1+i & 2\end{pmatrix} \xrightarrow{R_1-iR_2\to R_3} \begin{pmatrix}1 & 0 & 0 & 1\\ 0 & 1 & 1 & 0\\ 2 & 1+i & 1+i & 2\end{pmatrix}$$
$$\xrightarrow{R_3-2R_1-(1+i)R_2\to R_3} \begin{pmatrix}1 & 0 & 0 & 1\\ 0 & 1 & 1 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}$$

זה מראה כי שתי השורות הראשונות פורשות את השורה השלישית, ושתיהן מהוות בסיס למרחב השורות. בנוסף, ניתן להוסיף את וקטורי היחידה e_3, e_4 כדי להשלים את בסיס זה לבסיס ל- $\mathbb{C}^4$. נחזור חזרה למרחב הוקטורי $\mathbb{M}_{2\times 2}$ ונקבל ש- $\left\{ \begin{pmatrix}1 & i\\ i & 1\end{pmatrix}, \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} \right\}$ היא בסיס ל- W, ולכן $\dim(W) = 2$. לאחר השלמה נקבל את הבסיס הבא ל- $\mathbb{M}_{2\times 2}(\mathbb{C})$:

$$\left\{ \begin{pmatrix}1 & i\\ i & 1\end{pmatrix}, \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix} \right\}$$

כעת נוכיח הכללה של משפט 8.60 ("שלישי חינם") למרחבים וקטוריים כלליים:

משפט 8.82. *יהי V מ"ו מעל $\mathbb{F}$ כך ש- $\dim(V) = n$, ותהי $S \subseteq V$. אז התנאים הבאים שקולים:*

א. S בת"ל בגודל n.

ב. S פורשת את V וגודלה n.

ג. S היא בסיס ל- V.

הוכחה. נשים לב שלפי כל אחד מהתנאים הגודל של S הוא בהכרח n. אז מלכתחילה ניתן לכתוב $S = \{ v_1, ..., v_n \}$. בנוסף, מהנתון $\dim(V) = n$ נובע שקיים בסיס B בגודל n ל- V. נסתכל על הקבוצה $S_B = \{ [v_1]_B, ..., [v_n]_B \} \subseteq \mathbb{F}^n$. אז לפי משפט 8.60, התנאים הבאים שקולים:

א. S_B בת"ל בגודל n.

ב. S_B פורשת את $\mathbb{F}^n$ וגודלה n.

ג. S_B היא בסיס ל- $\mathbb{F}^n$.

לבסוף, לפי משפט 8.74, כל תנאי לעיל שקול לתנאי המתאים לגבי S. לכן, כל התנאים שקולים. $\square$

מסקנה 8.83. *יהי V מ"ו נוצר סופית מעל $\mathbb{F}$, ויהי W תת-מרחב. אז מתקיים:*

א. W נוצר סופית ומתקיים $\dim(W) \leq \dim(V)$.

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

הוכחה. נסמן $\dim(V) = n$.

א. אם $W = \{ 0_V \}$, אז הוא נוצר סופית ומתקיים $\dim(W) = 0 \leq \dim(V)$. אחרת, קיים $0_V \neq w_1 \in W$ כך ש- $\operatorname{Span}(w_1) \subseteq W$. אם מתקיים שוויון, אז $\dim(W) = 1$ וסיימנו. לחילופין, נוסיף $w_2 \in W$ כך ש- $w_2 \notin \operatorname{Span}(W_1)$, ונקבל $\{ w_1, w_2 \}$ בת"ל כך ש- $\operatorname{Span}(w_1, w_2) \subseteq W$. אם עדיין אין שוויון, נמשיך באופן דומה ע"י הוספת וקטור בכל שלב עד שנקבל $B = \{ w_1, ..., w_k \}$ בת"ל כך ש- $\operatorname{Span}(w_1, ..., w_k) = W$. התהליך חייב להסתיים עם שוויון עבור $k \leq n$ וקטורים, כי לא קיימת קבוצה בת"ל ב- V בגודל $n+1$ לפי מסקנה 8.76. אז W נוצר סופית כי B היא בסיס שלו, ומתקיים

$$\dim(W) = k \leq n = \dim(V)\,.$$

ב. נניח כי $k = n$. אז B זו קבוצה בת"ל ב- V בגודל n, ולכן לפי משפט 8.82 היא גם בסיס ל- V. אז בפרט, מתקיים $W = \operatorname{Span}(w_1, ..., w_n) = V$.

$\square$

דוגמה 8.84. יהי V מ"ו מעל $\mathbb{F}$. נניח כי $B = \{ v_1, v_2, v_3, v_4 \}$ הוא בסיס ל- V. נגדיר את תת-המרחבים הבאים:

$$W_1 = \operatorname{Span}(v_1 + v_2 + v_3 + v_4, v_1 - v_2 - v_3 + v_4, v_1 + v_4)$$
$$W_2 = \operatorname{Span}(v_1 + v_4, v_2 + v_3)$$

נראה כי $W_2 = W_1$. בשביל הכיוון $W_2 \supseteq W_1$ נראה כי

$$v_2 + v_3 \in \operatorname{Span}(v_1 + v_2 + v_3 + v_4, v_1 - v_2 - v_3 + v_4, v_1 + v_4)\,,$$

וזה יספיק שהרי v_1+v_4 מופיע בשתי הקבוצות. במקום "לנחש" את הסקלרים שיראו זאת (לפעמים זה אפשרי עם בדיקה), הרעיון הוא להשתמש בוקטורי קוארדינטות לפי B ומסקנה 8.73. מתקיים

$$\begin{aligned}[v_1+v_2+v_3+v_4]_B &= (1,1,1,1)\\ [v_1-v_2-v_3+v_4]_B &= (1,-1,-1,1)\\ [v_1+v_4]_B &= (1,0,0,1)\\ [v_2+v_3]_B &= (0,1,1,0)\end{aligned}$$

נכתוב את הוקטורים כעמודות מטריצה ונדרג אותה:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0\\ 1 & -1 & 0 & 1\\ 1 & -1 & 0 & 1\\ 1 & 1 & 1 & 0\end{array}\right) \xrightarrow[R_3-R_2\to R_3]{R_4-R_1\to R_4} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0\\ 1 & -1 & 0 & 1\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{array}\right) \xrightarrow{R_2-R_1\to R_2} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0\\ 0 & -2 & -1 & 1\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{array}\right)$$

אין צורך בדירוג קנוני כי כבר ניתן לקבוע שקיים פתרון (לא יחיד), ולכן נובע כי $W_1 \supseteq W_2$. בשביל הכיוון $W_2 \subseteq W_1$ לכאורה צריך לעשות דירוג נוסף, כאשר העמודות שמתאימות ל- v_1+v_4, v_2+v_3 יהיו בצד שמאל. אבל יש קיצור דרך: נראה כי $\dim(W_1)=\dim(W_2)$ ואז ינבע כי $W_1=W_2$ לפי מסקנה 8.83. ובכן, הדירוג הקודם כבר מראה זאת כי בלי העמודה הימנית, קיבלנו דירוג עבור הוקטורים בהגדרת W_1. העמודה השלישית נפרשת ע"י שתי העמודות הראשונות, שמהוות קבוצה בת"ל. לכן, הקבוצה הבאה היא בסיס ל- W_1:

$$\{\, v_1+v_2+v_3+v_4, v_1-v_2-v_3+v_4 \,\}$$

אז $\dim(W_1)=2$. באופן דומה, אם נתמקד רק בשתי העמודות הימניות, הדירוג מראה כי $\{\, v_1+v_4, v_2+v_3 \,\}$ היא בסיס ל- W_2. לכן מתקיים $\dim(W_2)=2=\dim(W_1)$, ומכאן $W_2=W_1$.

8.7.1 משפט הדרגה

הדוגמה הבאה היא הכנה לקראת משפט שיתאר את הקשר בין המימדים של תת-מרחבים הקשורים למטריצה: מרחב השורות, מרחב העמודות ומרחב הפתרונות.

דוגמה 8.85. תהי

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

נחשב את המימדים של $\mathrm{N}(A),\mathrm{Col}(A),\mathrm{Row}(A)$ ע״י מציאת בסיסים. נדרג אותה ונקבל

$$\begin{pmatrix}1&2&0&3\\2&1&-1&0\\5&4&-2&3\\0&-3&-1&-6\end{pmatrix}\xrightarrow[R_3-5R_1\to R_3]{R_2-2R_1\to R_2}\begin{pmatrix}1&2&0&3\\0&-3&-1&-6\\0&-6&-2&-12\\0&-3&-1&-6\end{pmatrix}$$

$$\xrightarrow[R_4-R_2\to R_4]{R_3-2R_2\to R_3}\begin{pmatrix}1&2&0&3\\0&-3&-1&-6\\0&0&0&0\\0&0&0&0\end{pmatrix}\xrightarrow{-\frac{1}{3}R_2\to R_2}\begin{pmatrix}1&2&0&3\\0&1&\frac{1}{3}&2\\0&0&0&0\\0&0&0&0\end{pmatrix}$$

$$\xrightarrow{-\frac{1}{3}R_2\to R_2}\begin{pmatrix}1&0&-\frac{2}{3}&-1\\0&1&\frac{1}{3}&2\\0&0&0&0\\0&0&0&0\end{pmatrix}$$

יש שני איברים מובילים בשתי העמודות הראשונות, ולכן במטריצה המקורית שתי העמודות האחרונות נפרשות ע״י שתי העמודות הראשונות. אז הקבוצה

$$\left\{\begin{pmatrix}1\\2\\5\\0\end{pmatrix},\begin{pmatrix}2\\1\\4\\-3\end{pmatrix}\right\}$$

היא בסיס ל- $\mathrm{Col}(A)$, ולכן $\dim(\mathrm{Col}(A))=2$.

לגבי מרחב השורות: הדירוג לא משנה אותו, ולכן ניתן להשתמש בשורות המטריצה המדורגת קנונית (בלי שורות האפסים). אז $\left\{(1,0,-\frac{2}{3},-1),(0,1,\frac{1}{3},2)\right\}$ היא בסיס ל- $\mathrm{Row}(A)$, ולכן $\dim(\mathrm{Row}(A))=2$. נשים לב כי עד עכשיו שני המימדים שווים לדרגת המטריצה, וזה נכון באופן כללי לפי משפט שנוכיח בקרוב.

במקרה זה מתקיים גם $\dim(\mathrm{N}(A))=2$, אבל הסיבה שונה. באופן אינטואיטיבי, המימד של מרחב הפתרונות אמור להיות מספר המשתנים החופשיים, שהוא ההפרש בין מספר כל המשתנים והדרגה (מספר המשתנים התלויים). עוד לא הוכחנו את זה, אז נמצא בסיס ל- $\mathrm{N}(A)$. נפתור את הממ״ל השקולה:

$$\begin{cases} x_1 - \frac{2}{3}x_3 - x_4 = 0 \\ x_2 + \frac{1}{3}x_3 + 2x_4 = 0 \end{cases}$$

נציב $x_3 = s, x_4 = t$ ונקבל

$$\begin{aligned} \mathrm{N}(A) &= \left\{ \left(\frac{2}{3}s + t, -\frac{1}{3}s - 2t, s, t\right) \;\middle|\; s, t \in \mathbb{R} \right\} \\ &= \left\{ s\left(\frac{2}{3}, -\frac{1}{3}, 1, 0\right) + t(1, -2, 0, 1) \;\middle|\; s, t \in \mathbb{R} \right\} \\ &= \mathrm{Span}\left(\left(\frac{2}{3}, -\frac{1}{3}, 1, 0\right), (1, -2, 0, 1) \right) \end{aligned}$$

אז $\left\{ (\frac{2}{3}, -\frac{1}{3}, 1, 0), (1, -2, 0, 1) \right\}$ פורשת את $\mathrm{N}(A)$. קל לבדוק שהיא בת"ל כי שני הוקטורים אינם קו-לינאריים, אבל יש דרך אחרת לראות זאת. אם נניח שקיימים $\alpha_1, \alpha_2 \in \mathbb{R}$ כך ש-

$$\alpha_1\left(\frac{2}{3}, -\frac{1}{3}, 1, 0\right) + \alpha_2(1, -2, 0, 1) = (0, 0, 0, 0),$$

אז לפי שתי הקוארדינטות האחרונות נקבל $\alpha_1 = \alpha_2 = 0$. לכן, קבוצה זו היא בסיס ומתקיים $\dim(\mathrm{N}(A)) = 2$.

תרגיל 8.86. יהיו $\{ v_1, ..., v_k \} \subseteq \mathbb{F}^n$ ו- $A \in \mathbb{M}_{n \times n}(\mathbb{F})$ הפיכה. הוכיחו כי $\{ Av_1, ..., Av_k \}$ בת"ל אם ורק אם $\{ v_1, ..., v_k \}$ בת"ל.

פתרון. בכיוון הראשון, נניח כי $\{ v_1, ..., v_k \}$ בת"ל וקיימים $\alpha_1, .., \alpha_k \in \mathbb{F}$ כך ש-

$$\alpha_1 Av_1 + ... + \alpha_k Av_k = 0 \iff A(\alpha_1 v_1 + ... + \alpha_k v_k) = 0.$$

מהפיכות A נובע כי $\alpha_1 v_1 + ... + \alpha_k v_k = 0$, ומהנתון על $\{ v_1, ..., v_k \}$ נובע כי בהכרח $\alpha_1 = ... = \alpha_k = 0$. אז $\{ Av_1, ..., Av_k \}$ בת"ל.

בכיוון השני (שלא דורש הפיכות), נניח כי $\{ Av_1, ..., Av_k \}$ בת"ל וקיימים $\alpha_1, .., \alpha_k \in \mathbb{F}$ כך ש-

$$\alpha_A v_1 + ... + \alpha_k v_k = 0.$$

נכפיל את שני האגפים ב- A ונקבל

$$A(\alpha_1 v_1 + ... + \alpha_k v_k) = A0 \iff \alpha_1 Av_1 + ... + \alpha_k Av_k = 0.$$

לכן, בהכרח מתקיים $\alpha_1 = ... = \alpha_k = 0$ לפי ההנחה. אז $\{ v_1, ..., v_k \}$ בת"ל.

משפט 8.87. *תהי $A \in \mathbb{M}_{m \times n}(\mathbb{F})$. אז מתקיים:*

א.

$$\dim(\mathrm{Row}(A)) = \dim(\mathrm{Col}(A)) = \mathrm{rank}(A)$$

ב.

$$\dim(\mathrm{N}(A)) + \mathrm{rank}(A) = n$$

הערה 8.88. הסעיף השני של המשפט נקרא *"משפט הדרגה"*. שימו לב ש- $\dim(\mathrm{N}(A))$ קטן ככל ש- $\mathrm{rank}(A)$ גדלה. כבר ראינו שמתקיים $\mathrm{rank}(A) = n$ אם ורק אם $\mathrm{N}(A) = \{ 0 \}$, וזה שקול לתנאי $\dim(\mathrm{N}(A)) = 0$. במקרה זה הסכום הוא אכן n, והמשפט מכליל זאת למקרים אחרים. רעיון ההוכחה הוא פשוט: הדרגה שווה למספר המשתנים התלויים, ומימד מרחב הפתרונות שווה למספר המשתנים החופשיים (והסכום הכולל הוא n). צריך רק להצדיק את הקביעות האלו.

הוכחה. A שקולת שורה למטריצה מדורגת קנונית A'. מספיק להוכיח את המשפט עבור A', כי לפי הגדרת הדרגה $\mathrm{rank}(A) = \mathrm{rank}(A')$ ולפי טענה 8.63 מתקיים:

$$\begin{aligned} \mathrm{Row}(A) = \mathrm{Row}(A') &\implies \dim(\mathrm{Row}(A)) = \dim(\mathrm{Row}(A')) \\ \mathrm{N}(A) = \mathrm{N}(A') &\implies \dim(\mathrm{N}(A)) = \dim(\mathrm{N}(A')) \end{aligned}$$

בנוסף, נראה כי $\dim(\mathrm{Col}(A)) = \dim(\mathrm{Col}(A'))$. אכן, אם $v_1, ..., v_n \in \mathbb{F}^m$ הן עמודות A, אז קיימת תת-קבוצה שהיא בסיס ל- $\mathrm{Col}(A)$. ניתן להניח (בלי הגבלת הכלליות, כי הסדר לא משפיע על המימד) שזו תת-קבוצה מהצורה $\{ v_1, ..., v_k \}$ כאשר $k \leq n$. אז לפי מסקנה 6.68, קיימת מטריצה הפיכה $B \in \mathbb{M}_{n\times n}(\mathbb{F})$ כך ש- $A' = BA$. לפי תרגיל 8.86, נובע כי $\{ Bv_1, ..., Bv_k \}$ היא בת"ל. היא גם פורשת את $\mathrm{Col}(A') = \mathrm{Span}(Bv_1, ..., Bv_n)$, כי אחרת קיים $j > k$ כך ש- $\{ Bv_1, ..., Bv_k, Bv_j \}$ בת"ל ואז נובע לפי תרגיל 8.86 כי $\{ v_1, ..., v_k, v_j \}$ בת"ל וזו סתירה להנחה ש- $\{ v_1, ..., v_k \}$ היא בסיס ל- $\mathrm{Col}(A)$. לכן, $\{ Bv_1, ..., Bv_k \}$ היא בסיס ל- $\mathrm{Col}(A')$ ומתקיים $\dim(\mathrm{Col}(A)) = k = \dim(\mathrm{Col}(A'))$.

אז אנחנו מוכנים להתמקד בהוכחת המשפט עבור A', שהיא מהצורה

$$A' = \begin{pmatrix} 1 & 0 & \cdots & 0 & b_{1,r+1} & \cdots & b_{1,n} \\ 0 & 1 & \cdots & 0 & b_{2,r+1} & \cdots & b_{2,n} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & b_{r,r+1} & \cdots & b_{r,n} \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

כאשר $r = \mathrm{rank}(A)$. למעשה, נראה כי $k = r$.

א. נסמן ב- $w_1, ...w_r$ את השורות של A' שאינן שורות אפסים. ברור ש- $\{ w_1, ..., w_r \}$ פורשת את $\mathrm{Row}(A)$. נראה כי זו גם קבוצה בת"ל. כבר הנחנו בלי הגבלת הכלליות שהאיברים המובילים מופיעים ב- r העמודות הראשונות (אם יש קפיצות, הן משפיעות רק על האינדקסים ולא על המימדים). כעת נניח שקיימים $\alpha_1, ..., \alpha_r \in \mathbb{F}$ כך ש-

$$\begin{aligned}\alpha_1 w_1 + ... + \alpha_r w_r = 0 &\implies (\alpha_1, ..., \alpha_n, *, ..., *) = (0, ..., 0, 0 ..., 0)\\ &\implies \alpha_1 = ... = \alpha_n = 0\end{aligned}$$

אז $\{ w_1, ..., w_r \}$ היא בסיס ל- $\mathrm{Row}(A)$, ולכן $\dim(\mathrm{Row}(A)) = r = \mathrm{rank}(A)$.

לגבי העמודות, ברור כי $v_i = e_i$ לכל $1 \leq i \leq r$. ברור גם ש- $\{ e_1, ..., e_r \}$ פורשת את שאר העמודות ולכן את $\mathrm{Col}(A')$. ידוע שזו קבוצה בת"ל, ולכן מדובר בבסיס. אז קיבלנו $\dim(\mathrm{Col}(A')) = r$, ולכן $\dim(\mathrm{Col}(A)) = \mathrm{rank}(A)$.

ב. הממ"ל המתאימה ל- A' (אחרי העברת אגפים ומחיקת משוואות מהצורה $0 = 0$) היא

$$\begin{cases} x_1 &= & -b_{1,r+1}x_{r+1} - \cdots - b_{1,n}x_n \\ x_2 &= & -b_{2,r+1}x_{r+1} - \cdots - b_{2,n}x_n \\ \vdots & \vdots & \vdots \\ x_r &= & -b_{r,r+1}x_{r+1} - \cdots - b_{r,n}x_n \end{cases}$$

המשתנים החופשיים הם $x_{r+1}, ..., x_n$. נציב $x_i = t_i$ לכל $r + 1 \leq i \leq n$ ונקבל

$$\mathrm{N}(A)=\left\{\begin{pmatrix}-b_{1,r+1}t_{r+1}-\cdots-b_{1,n}t_n\\ \vdots\\ -b_{r,r+1}t_{r+1}-\cdots-b_{r,n}t_n\\ t_{r+1}\\ \vdots\\ t_n\end{pmatrix}\middle|\ t_{r+1},...,t_n\in\mathbb{F}\right\}$$

$$=\left\{t_{r+1}\begin{pmatrix}-b_{1,r+1}\\ \vdots\\ -b_{r,r+1}\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix}+...+t_n\begin{pmatrix}-b_{1,n}\\ \vdots\\ -b_{r,n}\\ 0\\ \vdots\\ 0\\ 1\end{pmatrix}\middle|\ t_{r+1},...,t_n\in\mathbb{F}\right\}$$

$$=\mathrm{Span}\left(\begin{pmatrix}-b_{1,r+1}\\ \vdots\\ -b_{r,r+1}\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix},...,\begin{pmatrix}-b_{1,n}\\ \vdots\\ -b_{r,n}\\ 0\\ \vdots\\ 0\\ 1\end{pmatrix}\right)$$

$\left\{\begin{pmatrix}-b_{1,r+1}\\ \vdots\\ -b_{r,r+1}\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix},...,\begin{pmatrix}-b_{1,n}\\ \vdots\\ -b_{r,n}\\ 0\\ \vdots\\ 0\\ 1\end{pmatrix}\right\}$ היא גם בת"ל ולכן היא בסיס ל- $\mathrm{N}(A)$. אכן, אם

נניח שקיימים $\alpha_{r+1},...,\alpha_n \in \mathbb{F}$ **כך ש-**

$$\alpha_1 \begin{pmatrix} -b_{1,r+1} \\ \vdots \\ -b_{r,r+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + ... + \alpha_n \begin{pmatrix} -b_{1,n} \\ \vdots \\ -b_{r,n} \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

נסתכל רק על $n-r$ **הקוארדינטות האחרונות ונקבל** $\alpha_{r+1} = ... = \alpha_n = 0$**. אז קיבלנו בסיס עם** $n-r$ **וקטורים (כמספר המשתנים החופשיים), ולכן** $\dim(\mathrm{N}(A')) = n-r$.
באופן שקול, נזכיר כי $\mathrm{N}(A) = \mathrm{N}(A')$ **ולכן מתקיים**

$$\dim(\mathrm{N}(A)) + \mathrm{rank}(A) = n \,.$$

□

תרגיל 8.89. **תהי** $A \in \mathbb{M}_{m\times n}(\mathbb{F})$**. הוכיחו כי:**

א. $\mathrm{rank}(A^t) = \mathrm{rank}(A)$

ב. $\dim(\mathrm{N}(A)) \geq n-m$

פתרון.

א. **לפי סעיף א' של משפט 8.87, מתקיים**

$$\mathrm{rank}(A^t) = \dim(\mathrm{Col}(A^t)) = \dim(\mathrm{Row}(A)) = \mathrm{rank}(A) \,.$$

ב. **לפי משפט הדרגה ואי-השוויון** $\mathrm{rank}(A) \leq m$**, מתקיים**

$$\dim(\mathrm{N}(A)) = n - \mathrm{rank}(A) \geq n-m \,.$$

תרגיל 8.90. **נסתכל על תת-המרחב** $W \subseteq \mathbb{R}^4$ **שמוגדר ע"י**

$$W = \mathrm{Span}\left(\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}\right).$$

א. נניח שקיימת מטריצה A כך ש- $\mathrm{N}(A)=W$. מהו מספר העמודות שלה? כמה שורות יש לה לכל הפחות? כמה שורות יש לה לכל היותר?

ב. מצאו מטריצה A כנ"ל.

פתרון. א. מספר השורות של A הוא בהכרח 4 בהתאם ל- $\mathbb{R}^4$. הקבוצה $\left\{ \begin{pmatrix}1\\2\\3\\4\end{pmatrix}, \begin{pmatrix}1\\0\\-1\\0\end{pmatrix} \right\}$ היא בת"ל כי שני הוקטורים אינם קו-לינאריים. לכן, קבוצה זו היא בסיס ל- W ומתקיים $\dim(\mathrm{N}(A))=\dim(W)=2$. לפי משפט הדרגה נובע כי

$$\dim(\mathrm{N}(A))+\mathrm{rank}(A)=4 \implies \mathrm{rank}(A)=2 .$$

אז מספר השורות המינימלי הוא 2, אבל אין שום חסם עליון כי אפשר לחזור על אחת משתי השורות (או שתיהן) מספר שרירותי של פעמים. אפשר גם לקחת צירופים לינאריים שונים של שתי השורות ולקבל מספר שרירותי של שורות שונות, ועדיין הדרגה תישאר 2.

ב. יש שתי דרכי פתרון.

<u>דרך א'</u>: נשים לב כי

$$\begin{pmatrix}x\\y\\z\\w\end{pmatrix}\in W \iff \mathrm{N}\left(\begin{array}{cc|c}1&1&x\\2&0&y\\3&-1&z\\4&0&w\end{array}\right)\neq\emptyset$$

או במילים: הוקטור שייך ל- W אם ורק אם קיים פתרון לממ"ל שמתארת את תנאי הפרישה. נדרג את מטריצת המקדמים המורחבת ונמצא תנאים על x,y,z,w לכך שלא תופיע אף שורת סתירה בסוף הדירוג. תנאים אלה יהוו ממ"ל חדשה, שתוביל למטריצה A.

$$\left(\begin{array}{cc|c}1&1&x\\2&0&y\\3&-1&z\\4&0&w\end{array}\right)\xrightarrow[R_4-4R_1\to R_4]{\substack{R_2-2R_1\to R_2\\R_3-3R_1\to R_3}}\left(\begin{array}{cc|c}1&1&x\\0&-2&y-2x\\0&-4&z-3x\\0&-4&w-4x\end{array}\right)$$

$$\xrightarrow{-\frac{1}{2}R_2\to R_2}\left(\begin{array}{cc|c}1&1&x\\0&1&\frac{2x-y}{2}\\0&-4&z-3x\\0&-4&w-4x\end{array}\right)\xrightarrow[R_4+4R_2\to R_4]{R_3+4R_2\to R_3}\left(\begin{array}{cc|c}1&1&x\\0&1&\frac{2x-y}{2}\\0&0&x-2y+z\\0&0&w-2y\end{array}\right)$$

לכן:

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} \in W \iff \begin{cases} x-2y+z=0 \\ \quad -2y+w=0 \end{cases}$$

אין תשובה יחידה ל- A, כי ניתן לבצע פש"אות על הממ"ל. אבל בלי להסתבך עם כל האפשרויות, ניקח את הממ"ל שקיבלנו ונכתוב את מטריצת המקדמים המצומצמת:

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

לפי הדירוג הקודם, אכן מתקיים $\mathrm{N}(A) = W$.

<u>דרך ב'</u>: נסמן

$$B = \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 3 & -1 \\ 4 & 0 \end{pmatrix}$$

ונשים לב כי $\dim(W) = \mathrm{rank}(B) = 2$. לפי סעיף א', נמצא A מסדר 2×4 (אפשר להוסיף עוד שורות, אך אין סיבה לעשות זאת). התנאי $AB = \mathbf{0}_{2\times 2}$ פירושו $W \subseteq \mathrm{N}(A)$. אז אם נמצא $A \in \mathbb{M}_{2\times 4}(\mathbb{R})$ כך ש- $\mathrm{rank}(A) = 2$ ו- $AB = \mathbf{0}_{2\times 2}$, ינבע כי $\dim(\mathrm{N}(A)) = 2$ לפי משפט הדרגה ומכאן $\mathrm{N}(A) = W$ לפי מסקנה **8.83**.
נשים לב כי

$$AB = \mathbf{0}_{2\times 2} \iff (AB)^t = \mathbf{0}_{2\times 2} \iff B^tA^t = \mathbf{0}_{2\times 2}.$$

זה אומר שכדי למצוא את A, צריך למצוא בסיס ל- $\mathrm{N}(B^t)$ ולכתוב את איבריו כשורות של A, כי הן העמודות של A^t. כמובן, יש אינסוף בסיסים אפשריים וכל אחד מהם ייתן תשובה אחרת. אז נדרג את B^t:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 0 & -1 & 0 \end{pmatrix} \xrightarrow{R_2 - R_1 \to R_1} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -2 & -4 & -4 \end{pmatrix}$$
$$\xrightarrow{-\frac{1}{2}R_2 \to R_2} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 2 \end{pmatrix} \xrightarrow{-\frac{1}{2}R_2 \to R_2} \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 2 \end{pmatrix}$$

לכן, מתקיים

$$\mathrm{N}(B^t) = \{\, (s, -2s-2t, s, t) \mid s, t \in \mathbb{R} \,\} = \mathrm{Span}\left((1,-2,1,0),(0,-2,0,1)\right).$$

אז שוב קיבלנו

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

פרק 9

תת-מרחבים

9.1 פעולות החיתוך והסכום

בהינתן מ"ו וקטורי V, ראינו את ההגדרה של תת-מרחב $W \subseteq V$ וגם את הקשר בין המימדים במסקנה 8.83. בפרק זה נרצה להבין מהן הפעולות שאפשר לבצע על שני תת-מרחבים (או יותר), כאשר המטרה היא לקבל תת-מרחבים נוספים.

קודם כל ניזכר בהגדרה של פעולת החיתוך מפרק 2, אבל הפעם נתאים אותה לתת-מרחבים במקום קבוצות כלליות.

הגדרה 9.1. יהי V מ"ו מעל $\mathbb{F}$, ויהיו $W_1, W_2 \subseteq V$ תת-מרחבים. אז החיתוך $W_1 \cap W_2$ הוא קבוצת כל הוקטורים השייכים גם ל- W_1 וגם ל- W_2, כלומר

$$W_1 \cap W_2 = \left\{ v \in V \mid v \in W_1, v \in W_2 \right\}.$$

דוגמה 9.2. למעשה, כבר דיברנו על חיתוך של תת-מרחבים בהקשר של פרק 4 ופרק 5.

א. נסתכל על ישרים ב- $\mathbb{R}^2$ שעוברים דרך הראשית ונתונים ע"י

$$L_1 = \left\{ (x, y) \in \mathbb{R}^2 \;\middle|\; a_1 x + b_1 y = 0 \right\}, L_2 = \left\{ (x, y) \in \mathbb{R}^2 \;\middle|\; a_2 x + b_2 y = 0 \right\}.$$

אם הם אינם מתלכדים, מתקיים $L_1 \cap L_2 = \left\{ (0,0) \right\}$. אם הם מתלכדים (כלומר וקטורי הנורמל הם קו-לינאריים), אז $L_1 \cap L_2 = L_1 = L_2$.

ב. נסתכל על מישורים ב- $\mathbb{R}^3$ שעוברים דרך הראשית ונתונים ע"י:

$$H_1 = \left\{ (x, y, z) \in \mathbb{R}^3 \;\middle|\; a_1 x + b_1 y + c_1 z = 0 \right\}$$
$$H_2 = \left\{ (x, y, z) \in \mathbb{R}^3 \;\middle|\; a_2 x + b_2 y + c_2 z = 0 \right\}$$

מתקיים

$$.H_1 \cap H_2 = \left\{ (x,y) \in \mathbb{R}^2 \;\middle|\; a_1x + b_1y + c_1z = 0, a_2x + b_2y + c_2z = 0 \right\}$$

אם הישרים אינם מתלכדים, אז הנורמלים $(a_1,b_1,c_1),(a_2,b_2,c_2)$ מהווים קבוצה בת"ל והדרגה של מטריצת המקדמים של הממ"ל היא 2. לפי משפט הדרגה, המימד של קבוצת הפתרונות הוא $3-2=1$. לכן, החיתוך של שני מישורים כנ"ל הוא אכן ישר. אבל אם המישורים מתלכדים, החיתוך ביניהם הוא המישור $H_1 = H_2$.

ג. נכליל את הדוגמאות הקודמות: עבור $A \in \mathbb{M}_{m\times n}(\mathbb{F})$, נסתכל על $\mathrm{N}(A)$ - מרחב הפתרונות של הממ"ל $Ax = 0$. נסתכל גם על $\mathrm{N}(B)$ עבור $B \in \mathbb{M}_{k\times n}(\mathbb{F})$. אז $\mathrm{N}(A)\cap\mathrm{N}(B) = \mathrm{N}\begin{pmatrix} A \\ B \end{pmatrix}$, כאשר $\begin{pmatrix} A \\ B \end{pmatrix} \in \mathbb{M}_{(m+k)\times n}(\mathbb{F})$ היא המטריצה שמתקבלת ע"י כתיבת שורות A למעלה ושורות B למטה. הסיבה לשוויון היא החישוב הבא:

$$v \in \mathrm{N}\begin{pmatrix} A \\ B \end{pmatrix} \iff \begin{pmatrix} A \\ B \end{pmatrix} v = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \iff \begin{pmatrix} Av \\ Bv \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$
$$\iff v \in \mathrm{N}(A) \cap \mathrm{N}(B)$$

ד. ב- $\mathbb{M}_{n\times n}(\mathbb{F})$ נסתכל על תת-המרחב W_1 של כל המטריצות המשולשיות עליונות, ותת-המרחב W_2 של כל המטריצות המשולשיות תחתונות. אז $W_1 \cap W_2$ הוא תת-המרחב של כל המטריצות האלכסוניות.

עבור $n=2$ ניתן למצוא את הבסיסים הבאים:

$B_1 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ היא בסיס ל- W_1, ולכן $\dim(W_1) = 3$.

$B_2 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ היא בסיס ל- W_2, ולכן $\dim(W_2) = 3$.

ו- $B = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ היא בסיס ל- $W_1 \cap W_2$, ולכן $\dim(W_1 \cap W_2) = 2$.

עבור $n \in \mathbb{N}$ כללי, במטריצה מסדר $n \times n$ יש n איברים על האלכסון הראשי. לכן, מתקיים $\dim(W_1 \cap W_2) = n$ כמספר המטריצות בבסיס הבא:

$$\left\{ \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} \right\}$$

תרגיל 9.3. בסעיף ד' של הדוגמה האחרונה, מהם $\dim(W_1), \dim(W_2)$ עבור $n \in \mathbb{N}$ כללי?

פתרון. ראינו שמימד החיתוך הוא n, כמספר המשבצות על האלכסון הראשי. בשביל $\dim(W_1)$ צריך להוסיף את מספר המשבצות מעל האלכסון הראשי. דרך אחת לחשב אותו הוא ע"י הנוסחה לסכום סדרה חשבונית, אבל נשתמש בדרך אחרת. משיקולי סימטריה, מספר זה שווה למספר המשבצות מתחת לאלכסון הראשי. אם נסמן מספר זה ב- k, הרי שניתן לספור את כלל המשבצות בשתי דרכים שונות ולקבל

$$2k + n = n^2 \implies k = \frac{n^2 - n}{2}.$$

מכאן נובע כי

$$\dim(W_1) = \dim(W_2) = k + n = \frac{n^2 - n}{2} + n = \frac{n(n+1)}{2}.$$

בכל הדוגמאות לעיל, חיתוך של תת-מרחבים הוא תת-מרחב. נוכיח זאת באופן כללי.

טענה 9.4. *יהי V מ"ו מעל $\mathbb{F}$, ויהיו $W_1, W_2 \subseteq V$ תת-מרחבים. אז $W_1 \cap W_2$ הוא תת-מרחב של V.*

הוכחה. לפי טענה 8.13, צריך להראות כי $0_V \in W_1 \cap W_2$ ובנוסף שיש סגירות לחיבור וכפל בסקלר.

- מתקיים $0_V \in W_1$ וגם $0_V \in W_2$. לכן, לפי הגדרת החיתוך נובע כי $0_V \in W_1 \cap W_2$.
- סגירות לחיבור: יהיו $v_1, v_2 \in W_1 \cap W_2$. בפרט, $v_1, v_2 \in W_1$ ולכן $v_1 + v_2 \in W_1$ כי W_1 סגור לחיבור. באופן דומה, מתקיים $v_1 + v_2 \in W_2$ ולכן $v_1 + v_2 \in W_1 \cap W_2$.
- סגירות לכפל סקלר: יהיו $v \in W_1 \cap W_2$ ו- $\alpha \in \mathbb{F}$. בפרט, $v \in W_1$ ולכן $\alpha v \in W_1$ מסגירות W_1 לכפל בסקלר. באופן דומה, גם מתקיים $\alpha v \in W_2$, ולכן $\alpha v \in W_1 \cap W_2$.

□

תרגיל 9.5. נסמן את תת-המרחבים הבאים ב- $\mathbb{R}^4$:

$$W_1 = \operatorname{Span}\left(\begin{pmatrix}1\\1\\0\\1\end{pmatrix}, \begin{pmatrix}2\\1\\1\\0\end{pmatrix}\right), \quad W_2 = \operatorname{Span}\left(\begin{pmatrix}0\\1\\-1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\\1\end{pmatrix}\right).$$

מצאו את $W_1 \cap W_2$.

פתרון. דרך פתרון אחת היא למצוא מטריצות $A, B \in \mathbb{M}_{2\times 4}(\mathbb{R})$ כך ש-

$$\mathrm{N}(A) = W_1, \mathrm{N}(B) = W_2,$$

ואז לחשב את

$$W_1 \cap W_2 = \mathrm{N}\begin{pmatrix} A \\ B \end{pmatrix}.$$

דרך זו ארוכה כי מדובר בפתרון של שלוש ממ"ליות. בדרך הבאה נצטרך לפתור ממ"ל אחת בלבד: מאחר שהמטרה היא למצוא איבר כללי בחיתוך, נרצה לדעת אילו וקטורים הם גם ב- W_1 וגם ב- W_2. כלומר, אילו צירופים לינאריים של $\begin{pmatrix}2\\1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\\1\end{pmatrix}$ הם גם צירופים לינאריים של $\begin{pmatrix}1\\0\\1\\1\end{pmatrix}, \begin{pmatrix}0\\1\\-1\\0\end{pmatrix}$? לשם כך, נחפש $a, b, c, d \in \mathbb{R}$ כך ש-

$$a\begin{pmatrix}1\\1\\0\\1\end{pmatrix} + b\begin{pmatrix}2\\1\\1\\0\end{pmatrix} = c\begin{pmatrix}0\\1\\-1\\0\end{pmatrix} + d\begin{pmatrix}1\\0\\1\\1\end{pmatrix}. \tag{9.1}$$

העברת אגפים מובילה לממ"ל הומוגנית המיוצגת ע"י המטריצה הבאה:

$$\begin{pmatrix} 1 & 2 & 0 & -1 \\ 1 & 1 & -1 & 0 \\ 0 & 1 & 1 & -1 \\ 1 & 0 & 0 & -1 \end{pmatrix}$$

לאחר דירוג מקבלים את הצורה המדורגת קנונית הבאה:

$$\begin{pmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

ולכן, הקשר בין הסקלרים הוא בהכרח:

$$a = 2d, b = 0, c = d.$$

נציב זאת באגף שמאל של משוואה 9.1. (אפשר גם באגף ימין) ונקבל

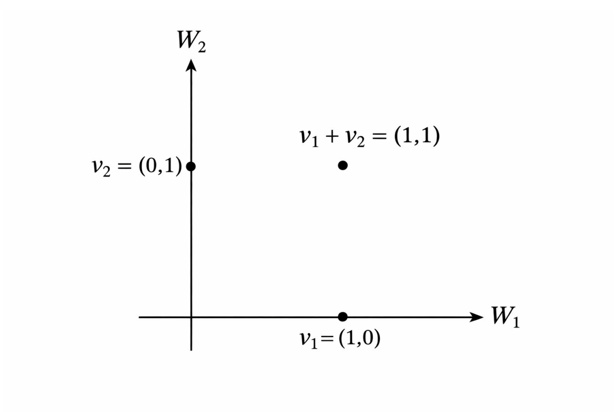
$$.W_1 \cap W_2 = \left\{ 2d \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \;\middle|\; d \in \mathbb{R} \right\} = \operatorname{Span}\left(\begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right)$$

ראינו בפרק 2 שבתורת הקבוצות הפעולה הטבעית הנלווית לחיתוך היא איחוד. עבור שני תת-מרחבים $W_1, W_2 \subseteq V$ האיחוד $W_1 \cup W_2$ הוא קבוצת כל הוקטורים השייכים לפחות לאחד מ- W_1 ו- W_2. נדגיש שפעולה זו אינה שימושית עבור תת-מרחבים, כי איחוד של שני תת-מרחבים אינו בהכרח תת-מרחב. אמנם אין בעיה עם וקטור האפס וגם לא עם סגירות לכפל בסקלר, אבל לא בהכרח מתקיימת סגירות לחיבור. אינטואיטיבית, האיחוד "לא יודע" על פעולת החיבור.

דוגמה 9.6. ב- $\mathbb{R}^2$ נסתכל על $W_1 = \operatorname{Span}((1,0)), W_2 = \operatorname{Span}((0,1))$. אז מתקיים

$$,W_1 \cup W_2 = \{\, (t,0) \mid t \in \mathbb{R} \,\} \cup \{\, (0,s) \mid s \in \mathbb{R} \,\}$$

שזה איחוד של שני ישרים (ציר x וציר y). מתקיים $(1,0),(0,1) \in W_1 \cup W_2$, אך עבור החיבור נקבל $(1,0)+(0,1) = (1,1) \notin W_1 \cup W_2$ כפי שרואים באיור למטה. לכן, אין סגירות לחיבור ו- $W_1 \cup W_2$ אינו תת-מרחב.



איור 9.1: שני הוקטורים שייכים לאיחוד של תת-המרחבים (שני הצירים), אך סכומם נמצא מחוץ לאיחוד

תרגיל 9.7. יהי V מ"ו מעל $\mathbb{F}$ ויהיו $W_1, W_2 \subseteq V$ תת-מרחבים. אז $W_1 \cup W_2$ הוא תת-מרחב אם ורק אם $W_1 \subseteq W_2$ או $W_2 \subseteq W_1$.

רמז: בשביל הכיוון הקשה יותר כדאי להניח בשלילה שקיימים $w_1 \in W_1, w_2 \in W_2$ כך כ- $.w_1 \notin W_2, w_2 \notin W_1$

פתרון. בכיוון אחד, נניח כי $W_1 \subseteq W_2$ (המקרה של $W_2 \subseteq W_1$ דומה משיקולי סימטריה). אז מתקיים $W_1 \cup W_2 = W_2$, ובפרט זהו תת-מרחב.

בכיוון השני, נניח כי $W_1 \cup W_2$ הוא תת-מרחב. נניח בשלילה שקיימים w_1, w_2 כמו ברמז. אז לא ייתכן כי $w_1 + w_2 \in W_1$, כי אחרת $w_2 = (w_1 + w_2) + (-w_1) \in W_1$ מסגירות לחיבור של W_1. זו סתירה.

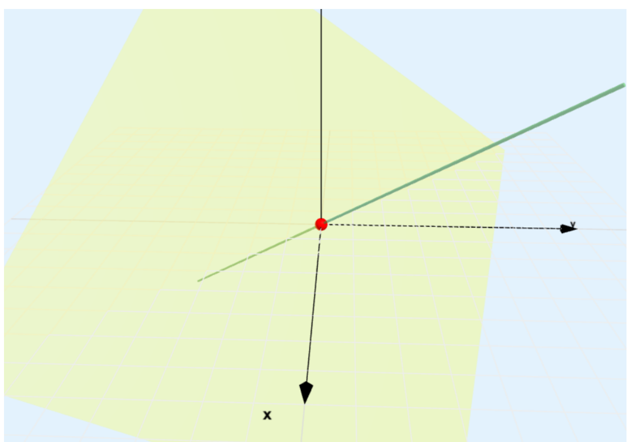
אז $w_1 + w_2 \notin W_1$, ובאופן דומה $w_1 + w_2 \notin W_2$. לכן, מתקיים $w_1 + w_2 \notin W_1 \cup W_2$ וזו סתירה לסגירות לחיבור של $W_1 \cup W_2$.

איזו פעולה על תת-מרחבים W_1, W_2 אפשר להגדיר כך שהתוצאה תהיה תת-המרחב הקטן ביותר שמכיל את שניהם? הבנו שיש צורך לקחת בחשבון את פעולת החיבור. למעשה, לכל $w_1 \in W_1$ ולכל $w_2 \in W_2$ צריך לדרוש שהחיבור $w_1 + w_2$ יהיה בתת-המרחב החדש. זה מוביל להגדרה הבאה:

הגדרה 9.8. יהי V מ"ו מעל $\mathbb{F}$, ויהיו $W_1, W_2 \subseteq V$ תת-מרחבים. אז הסכום $W_1 + W_2$ מוגדר ע"י

$$W_1 + W_2 = \{ w_1 + w_2 \mid w_1 \in W_1, w_2 \in W_2 \}.$$

במילים אחרות, בקבוצה החדשה $W_1 + W_2$ נמצאים כל הוקטורים שהם סכום של וקטור מ-W_1 עם וקטור מ-W_2. הרעיון דומה לפרישה ויש קשר, אך כאן הפעולה יותר כללית.



איור 9.2: תת-מרחב אחד הוא ישר (בירוק) ותת-המרחב השני הוא מישור (בצהוב), כאשר החיתוך (נקודה) מופיע באדום והסכום (כל המרחב) מופיע בכחול

טענה 9.9. *יהי V מ"ו מעל $\mathbb{F}$, ויהיו $W_1, W_2 \subseteq V$ תת-מרחבים. אז $W_1 + W_2$ הוא תת-מרחב של V, המקיים $W_1 \subseteq W_1 + W_2$ וגם $W_2 \subseteq W_1 + W_2$.*

הוכחה. לפי טענה 8.13, כדי להוכיח כי $W_1 + W_2$ הוא תת-מרחב, יש להראות כי $0_V \in W_1 + W_2$ ובנוסף שיש סגירות לחיבור וכפל בסקלר.

- מתקיים $0_V \in W_1$ וגם $0_V \in W_2$. לכן, נובע כי

$$.0_V = 0_V + 0_V \in W_1 + W_2$$

- סגירות לחיבור: יהיו $u_1, u_2 \in W_1 + W_2$. לכן, קיימים $v_1, v_2 \in W_1, w_1, w_2 \in W_2$ עבורם

$$.u_1 = v_1 + w_1, u_2 = v_2 + w_2$$

W_1 ו- W_2 סגורים לחיבור, ולכן מתקיים $v_1 + v_2 \in W_1, w_1 + w_2 \in W_2$. מכאן נובע כי

$$.u_1 + u_2 = (v_1 + w_1) + (v_2 + w_2) = (v_1 + v_2) + (w_1 + w_2) \in W_1 + W_2$$

- סגירות לכפל בסקלר: יהי $u \in W_1 + W_2$ ויהי $\alpha \in \mathbb{F}$. קיימים $v \in W_1, w \in W_2$ עבורם $u = v + w$. מסגירות לכפל בסקלר של W_1 ו- W_2 מתקיים $\alpha u_1 \in W_1$ וגם $\alpha u_2 \in W_2$. לכן, נובע כי

$$.\alpha(u_1 + u_2) = \alpha u_1 + \alpha u_2 \in W_1 + W_2$$

הוכחנו ש- $W_1 + W_2$ הוא אכן תת-מרחב. לבסוף, לכל $v \in W_1, w \in W_2$ מתקיים:

$$v = v + 0_V \in W_1 + W_2$$
$$w = 0_V + w \in W_1 + W_2$$

כלומר, מתקיים $W_1 \subseteq W_1 + W_2$ וגם $W_2 \subseteq W_1 + W_2$, כנדרש. $\square$

דוגמה 9.10.

א. נסמן ב- $\mathbb{R}^2$ את $W_1 = \mathrm{Span}((1,2))$ ו- $W_2 = \mathrm{Span}((2,1))$. אז מתקיים

$$.W_1 + W_2 = \{\, s(1,2) + t(2,1) \mid s,t \in \mathbb{R} \,\} = \mathrm{Span}((1,2),(2,1)) = \mathbb{R}^2$$

השוויון האחרון נובע מכך שהוקטורים הם בת"ל (לא קו-לינאריים). השוויון לפניו נכון מהגדרת הפרישה.

ב. נסמן ב- $\mathbb{R}^3$ את $W_1 = \mathrm{Span}((1,0,0),(1,1,1))$ ו- $W_2 = \mathrm{Span}((0,1,0),(1,1,1)))$. לא קשה לבדוק כי $W_1 \neq W_2$, ולכן באופן אינטואיטיבי $W_1 + W_2$ אמור להיות ממימד גבוה יותר מ- W_1 וגם מ- W_2 (במקרה זה הם שווי-מימד). אז אמור להתקיים $\dim(W_1 + W_2) = 3$, כי אין אפשרות למימד גבוה מזה ב- $\mathbb{R}^3$. בהמשך נראה טיעון פורמלי מסוג זה, המסתמך על שיקולי מימד, אבל בינתיים נראה כי $W_1 + W_2 = \mathbb{R}^3$ לפי

הגדרת הסכום. נשים לב כי $(1,1,1) \in W_1 \cap W_2$ ונשתמש בעובדה זו כדי להציג את הסכום בצורה פשוטה יותר, במקום לחזור על $(1,1,1)$.

$$\begin{aligned} W_1 + W_2 &= \{\, q(1,0,0) + r(1,1,1) + s(0,1,0) + t(1,1,1) \mid q,r,s,t \in \mathbb{R} \,\} \\ &= \{\, q(1,0,0) + s(0,1,0) + (r+t)(1,1,1) \mid q,r,s,t \in \mathbb{R} \,\} \\ &= \{\, q(1,0,0) + s(0,1,0) + p(1,1,1) \mid p,q,s \in \mathbb{R} \,\} \\ &= \operatorname{Span}((1,0,0),(0,1,0),(1,1,1)) = \mathbb{R}^3 \end{aligned}$$

השתמשנו בהצבה $p = r + t$. נלמד בהמשך דרך כללית יותר למצוא בסיסים לחיתוך וסכום של תת-מרחבים. השוויון האחרון נכון לפי משפט 8.59, למשל כי

$$.\det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} = 1^3 = 1 \neq 0$$

ג. נסמן ב- $\mathbb{M}_{2\times 2}(\mathbb{R})$ את

$$.W_1 = \operatorname{Span}(I_2), W_2 = \left\{ \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix} \middle| a, b \in \mathbb{R} \right\}$$

ניקח את הסכום ונקבל:

$$\begin{aligned} W_1 + W_2 &= \left\{ \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix} + \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix} \middle| t, a, b \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} t & a \\ b & t \end{pmatrix} \middle| t, a, b \in \mathbb{R} \right\} \\ &= \operatorname{Span}\left(I_2, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right) \end{aligned}$$

התרגיל הבא חשוב כי הוא מראה איך למצוא סכום של שני תת-מרחבים שנתונים ע"י Span, כמו ברוב המקרים שנעסוק בהם.

תרגיל 9.11. יהי V מ"ו מעל $\mathbb{F}$, ויהיו $w_1, ..., w_k, v_1, ..., v_m \in V$. נסמן

$$.W_1 = \operatorname{Span}(w_1, ..., w_k), \quad W_2 = \operatorname{Span}(v_1, ..., v_m)$$

הראו כי

$$.W_1 + W_2 = \operatorname{Span}(w_1, ..., w_k, v_1, ..., v_m)$$

פתרון. מתקיים

$$\begin{aligned} w\in W_1+W_2 &\iff \exists w_1\in W_1, w_2\in W_2 \quad w=w_1+w_2 \\ &\iff \exists \alpha_1,...,\alpha_k,\beta_1,...,\beta_m\in\mathbb{F} \quad w=\alpha_1w_1+...\alpha_kw_k+\beta_1v_1+...+\beta_mv_m \\ &\iff w\in \mathrm{Span}(w_1,...,w_k,v_1,...,v_m) \end{aligned}$$

לכן יש שוויון בין הקבוצות.

הערה 9.12. סכום בין תת-מרחבים לא מקיים את כל התכונות הרגילות של חיבור וקטורי (בין וקטורים בודדים). אמנם מתקיים חוק החילוף:

$$W_1+W_2=W_2+W_1$$

ואפילו יש איבר ניטרלי $\{\, 0_V \,\}$:

$$W+\{\, 0_V \,\}=W$$

אבל לתת-מרחב $W\neq\{\, 0_V \,\}$ אין איבר נגדי. נראה זאת: נניח בשלילה שקיים איבר נגדי, שהוא תת-מרחב $U\subseteq V$. אבל מתקיים $W\subseteq W+U$, ובפרט $W+U\neq\{\, 0_V \,\}$ וזו סתירה.

תרגיל 9.13. הוכיחו או הפריכו את הטענה הבאה: יהיו V מ"ו מעל $\mathbb{F}$ ויהיו $W_1,W_2,W_3\subseteq V$ תת-מרחבים. אם $W_1+W_3=W_2+W_3$, אז $W_1=W_2$.

פתרון. הטענה אינה נכונה. נפריך אותה ע"י הדוגמה הנגדית הבאה:
נבחר $V=\mathbb{R}^2$ ונסמן

$$W_1=\mathrm{Span}(e_1),\quad W_2=\mathrm{Span}(e_1+e_2),\quad W_3=\mathrm{Span}(e_2).$$

אז מתקיים $W_1\neq W_2$, כי e_1,e_1+e_2 אינם קו-לינאריים. ובכל זאת:

$$W_1+W_3=\mathrm{Span}(e_1,e_2)=\mathbb{R}^2=\mathrm{Span}(e_1+e_2,e_2)=W_2+W_3.$$

9.2 נוסחת המימדים

נרצה להבין את הקשר בין המימדים של שני תת-מרחבים, חיתוכם וסכומם. קשר אחד כבר קל להסיק:

מסקנה 9.14. *יהיו V מ"ו מעל $\mathbb{F}$ ו- $W_1,W_2\subseteq V$. אז מתקיים:*

$$\dim(W_1\cap W_2)\leq\dim(W_1)\leq\dim(W_1+W_2)\leq\dim(V)$$

ובאופן דומה:

$$\dim(W_1\cap W_2)\leq\dim(W_2)\leq\dim(W_1+W_2)\leq\dim(V)$$

הוכחה. זו מסקנה ישירה משילוב של מסקנה 8.83 וההכלות הבאות:

$$W_1 \cap W_2 \subseteq W_1 \subseteq W_1 + W_2 \subseteq V$$
$$W_1 \cap W_2 \subseteq W_2 \subseteq W_1 + W_2 \subseteq V$$

□

תרגיל 9.15. כהמשך למסקנה לעיל, הוכיחו שאם מתקיים $\dim(W_1 + W_2) = \dim(W_2)$, אז $W_1 \subseteq W_2$.

פתרון. נניח כי $\dim(W_1 + W_2) = \dim(W_2)$. מאחר שמתקיים $W_2 \subseteq W_1 + W_2$, שוויון מימדים גורר $W_2 = W_1 + W_2$ לפי מסקנה 8.83. אבל גם מתקיים $W_1 \subseteq W_1 + W_2$, ולכן $W_1 \subseteq W_2$.

הנוסחה במשפט הבא נקראת נוסחת המימדים, והיא מבטאת את קשר יותר מדויק בין המימדים של שני תת-מרחבים, חיתוכם וסכומם.

משפט 9.16. *יהי V מ"ו נוצר סופית מעל $\mathbb{F}$, ויהיו $W_1, W_2 \subseteq V$ תת-מרחבים. אז מתקיים*

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

הערה 9.17. זו הנוסחה האנלוגית לנוסחה שמבטאת את הקשר בין הגדלים של שתי קבוצות סופיות A, B, חיתוכן ואיחודן. אם $|X|$ מייצג את הגודל של קבוצה X, אז מתקיים

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

במשפט מדובר על מימד במקום גודל, ועל סכום במקום איחוד. אבל נבין דרך ההוכחה את האנלוגיה.

הוכחה. נסמן $k = \dim(W_1), l = \dim(W_2), m = \dim(W_1 \cap W_2)$.

קיים בסיס $B_0 = \{\, w_1, ..., w_m \,\}$ ל- $W_1 \cap W_2$. זהו תת-מרחב לא רק של V, אלא גם של W_1. לכן, ניתן להשלים את B_0 לבסיס של W_1 ע"י הוספת $k - m$ וקטורים. כך נקבל את $B_1 = \{\, w_1, ..., w_m, w_{m+1}, ..., w_k \,\}$. באופן דומה, $W_1 \cap W_2$ הוא גם תת-מרחב של W_2, ולכן ניתן להשלים את B_0 לבסיס של W_2 ע"י הוספת $l - m$ וקטורים. כך נקבל את הבסיס $B_2 = \{\, w_1, ..., w_m, u_1, ..., u_{l-m} \,\}$.

נסמן $B = B_1 \cup B_2 = \{\, w_1, ..., w_k, u_1, ..., u_{l-m} \,\}$. זו קבוצה הפורשת את $W_1 + W_2$ כי מתקיים

$$W_1 = \mathrm{Span}(B_1), W_2 = \mathrm{Span}(B_2) \implies W_1 + W_2 = \mathrm{Span}(B_1 \cup B_2) = \mathrm{Span}(B).$$

יותר מכך, נראה כי B היא בת"ל ולכן היא בסיס ל- $W_1 + W_2$.

נניח שקיימים $\alpha_1, ..., \alpha_k, \beta_1, ..., \beta_{l-m} \in \mathbb{F}$ כך שמתקיים

$$\alpha_1 w_1 + ... + \alpha_k w_k + \beta_1 u_1 + ... + \beta_{l-m} u_{l-m} = 0. \tag{9.2}$$

נחסיר משני האגפים את $\beta_1 u_1 + ... + \beta_{l-m} u_{l-m}$, ונקבל

$$\alpha_1 w_1 + ... + \alpha_k w_k = -\beta_1 u_1 + ... - \beta_{l-m} u_{l-m}.$$

נשים לב כי אגף שמאל שייך ל- W_1, ואילו אגף ימין שייך ל- W_2. לכן, שניהם שייכים ל- $W_1 \cap W_2$. בפרט, שניהם שווים לצירוף לינארי של איברי הבסיס B_0. אז קיימים $\gamma_1, ..., \gamma_m \in \mathbb{F}$ כך שמתקיים

$$\alpha_1 w_1 + ... + \alpha_m w_m + \alpha_{m+1} w_{m+1} + ... + \alpha_k w_k = \gamma_1 w_1 + ... + \gamma w_m$$

הקבוצה $B_1 = \{ w_1, ..., w_m, w_{k+1}, ..., w_m \}$ היא בסיס, ולכן הסקלרים נקבעים ביחידות לפי טענה **8.54**. מכאן

$$\alpha_i = \begin{cases} \gamma_i, & 1 \leq i \leq m \\ 0, & m+1 \leq i \leq k \end{cases}.$$

לכל $m+1 \leq i \leq k$ נציב $\alpha_i = 0$ במשוואה **9.2**, ונקבל

$$\alpha_1 w_1 + ... + \alpha_m w_m + \beta_1 u_1 + ... + \beta_{l-m} u_{l-m} = 0.$$

אגף שמאל הוא צירוף לינארי של איברי הבסיס B_2 (נפטרנו מהוקטורים שלא מופיעים בו), ולכן בהכרח כל הסקלרים מתאפסים. אז B בת"ל ולכן מדובר בבסיס ל- $W_1 + W_2$. לבסוף, נשים לב כי

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

$\square$

דוגמה 9.18.

א. ב- $\mathbb{R}^2$, לכל שני ישרים מהצורה $L_1 = \operatorname{Span}(v_1), L_2 = \operatorname{Span}(v_2)$ כך ש- $L_1 \neq L_2$ מתקיים:

$$\dim(L_1) = \dim(L_2) = 1$$
$$\dim(L_1 \cap L_2) = \dim(\{ 0 \}) = 0$$

נציב את המספרים בנוסחת המימדים ונקבל

$$\dim(L_1 + L_2) = 2 = 1 + 1 - 0 = 2,$$

ולכן $L_1 + L_2 = \mathbb{R}^2$ כי זהו תת-המרחב היחיד של $\mathbb{R}^2$ ממימד 2.

לעומת זאת, אם $L_1 = L_2$ נקבל:

$$\dim(L_1 + L_2) = \dim(L_1) = 1$$
$$\dim(L_1 \cap L_2) = \dim(L_1) = 1$$

ועדיין הנוסחה מתקיימת:

$$\dim(L_1 + L_2) = 1 = 1 + 1 - 1 = \dim(L_1) + \dim(L_2) - \dim(L_1 \cap L_2)$$

ב. ב- $\mathbb{R}^3$, כל שני מישורים מהצורה $H_1 = \mathrm{Span}(v_1, v_2), H_2 = \mathrm{Span}(v_3, v_4)$ כך ש- $H_1 \neq H_2$ נחתכים לאורך ישר L כלשהו. אז מתקיים:

$$\dim(H_1) = \dim(H_2) = 2$$
$$\dim(H_1 \cap H_2) = \dim(L) = 1$$

נציב את המספרים בנוסחת המימדים ונקבל

$$\dim(H_1 + H_2) = \dim(H_1) + \dim(H_2) - \dim(H_1 \cap H_2) = 2 + 2 - 1 = 3,$$

ולכן $H_1 + H_2 = \mathbb{R}^3$.

אם $H_1 = H_2$ נקבל

$$\dim(H_1 + H_2) = \dim(H_1) = 2$$
$$\dim(H_1 \cap H_2) = \dim(H_1) = 2$$

שוב הנוסחה מתקיימת:

$$\dim(H_1 + H_2) = 2 = 2 + 2 - 2 = \dim(H_1) + \dim(H_2) - \dim(H_1 \cap H_2)$$

ג. ב- $\mathbb{F}^n$, נסתכל על שני על-מישורים שונים H_1, H_2. הכוונה היא לתת-מרחבים ממימד $n - 1$. מהו מימד הסכום? מימד החיתוך? יותר קל לענות על השאלה הראשונה, כי מתקיים

$$H_1 \subseteq H_1 + H_2 \subseteq \mathbb{F}^n \implies n - 1 = \dim(H_1) \leq \dim(H_1 + H_2) \leq n.$$

לא ייתכן כי $\dim(H_1) = \dim(H_1 + H_2)$, שהרי זה גורר $H_1 = H_1 + H_2$ ומכאן $H_2 \subseteq H_1$. זו סתירה כי $\dim(H_2) = \dim(H_1)$, אך $H_2 \neq H_1$. לכן, מתקיים $\dim(H_1 + H_2) \neq n - 1$, ואז בהכרח נובע כי

$$\dim(H_1 + H_2) = n = \dim(\mathbb{F}^n) \implies H_1 + H_2 = \mathbb{F}^n.$$

נציב את המימדים הידועים בנוסחת המימדים ונקבל

$$n = n - 1 + n - 1 - \dim(H_1 \cap H_2) \implies \dim(H_1 \cap H_2) = n - 2.$$

ד. ב- $\mathbb{M}_{n\times n}(\mathbb{F})$ נסמן את תת-המרחבים של מטריצות סימטריות ואנטי-סימטריות:

$$W_1 = \left\{ A \in \mathbb{M}_{n\times n}(\mathbb{F}) \;\middle|\; A^t = A \right\}$$
$$W_2 = \left\{ A \in \mathbb{M}_{n\times n}(\mathbb{F}) \;\middle|\; A^t = -A \right\}$$

לפי טענה 6.53 מתקיים:

$$W_1 \cap W_2 = \{ \mathbf{0}_{n\times n} \} \implies \dim(W_1 \cap W_2) = 0$$
$$W_1 + W_2 = \mathbb{M}_{n\times n}(\mathbb{F}) \implies \dim(W_1 + W_2) = n^2$$

כיצד זה מתיישב עם נוסחת המימדים? נציין מראש שכדאי לנסות לחשוב לבד על המקרה $n = 3$ כדי להבין יותר טוב את הפרטים הטכניים של המקרה הכללי, שדורש סימונים חדשים.

נמצא בסיס ל- W_1:

$$B_1 = \{ D(i) \mid 1 \le i \le n \} \cup \{ S(i,j) \mid 1 \le i < j \le n \}$$

כאשר $D(i)$ היא המטריצה שכל איבריה 0 חוץ מ- 1 במשבצת ה- (i,i), ו- $S(i,j)$ היא המטריצה שכל איבריה 0 חוץ מ- 1 במשבצות ה- (i,j) וה- (j,i). קל לראות ש- B_1 היא קבוצה בת"ל, כי אין חיתוך בין המשבצות בהן המטריצות לא מתאפסות (לכל מטריצה יש את המשבצות שלה, או רק אחת). גם קל לראות שהיא פורשת את W_1: לכל $A \in W_1$ מתקיים

$$A = (A)_{11}D(1) + \ldots + (A)_{nn}D(n) + (A)_{12}S(1,2) + \ldots + (A)_{n-1,n}S(n-1,n)$$

אז מהו $\dim(W_1)$? יש n מטריצות מהצורה $D(i)$. כדי למנות את המטריצות מהצורה $S(i,j)$, נשים לב שלכל $2 \le j \le n$ יש $j-1$ מספרים טבעיים הקטנים ממנו (עבור $j = 1$ אין מספר טבעי כזה). נשתמש בנוסחה לסכום סדרה חשבונית ונקבל

$$\sum_{j=2}^{n}(j-1) = 1 + \ldots + (n-1) = \frac{n(n-1)}{2},$$

ובסך הכל

$$\dim(W_1) = n + \frac{n(n-1)}{2} = \frac{n(n+1)}{2}.$$

נמצא בסיס ל- W_2:

$$B_2 = \{\, T(i,j) \mid 1 \le i < j \le n \,\}$$

כאשר $T(i,j)$ היא המטריצה שכל איבריה 0 חוץ מ- 1 במשבצת ה- (i,j), ו- (-1) במשבצת ה- (j,i). בדומה לבדיקה עבור B_1, ניתן לבדוק כי B_2 היא אכן בת"ל ופורשת את W_2. מתקיים

$$\dim(W_2) = 1 + ... + (n-1) = \frac{n(n-1)}{2},$$

וכצפוי (לפי נוסחת המימדים)

$$n^2 = \frac{n(n+1)}{2} + \frac{n(n-1)}{2} - 0.$$

תרגיל 9.19. יהיו W_1, W_2 תת-מרחבים של $\mathbb{C}^4$.

א. נניח כי $\dim(W_1) = \dim(W_2) = 2$. הוכיחו כי $W_1 + W_2 = \mathbb{C}^4$ אם ורק אם $W_1 \cap W_2 = \{\, 0 \,\}$.

ב. נניח כי $\dim(W_1) = 3, \dim(W_2) = 2$. מהן האפשרויות ל- $\dim(W_1 \cap W_2)$? מצאו דוגמה ל- W_1, W_2 עבור כל אפשרות.

פתרון.

א. לפי מסקנה 8.83, מתקיים

$$W_1 + W_2 = \mathbb{C}^4 \iff \dim(W_1 + W_2) = 4,$$

ולפי נוסחת המימדים נובע כי

$$\dim(W_1+W_2) = 4 \iff 2+2-\dim(W_1 \cap W_2) = 4 \iff \dim(W_1 \cap W_2) = 0.$$

לכן, מתקיים

$$W_1 + W_2 = \mathbb{C}^4 \iff W_1 \cap W_2 = \{\, 0 \,\}.$$

ב. מתקיים

$$3 = \dim(W_1) \le \dim(W_1 + W_2) \le 4,$$

ולפי נוסחת המימדים (לאחר העברת אגפים) נובע כי

$$1 = 5 - 4 \le \dim(W_1 \cap W_2) = 3 + 2 - \dim(W_1 + W_2) \le 5 - 3 = 2.$$

האפשרות ל- $\dim(W_1 \cap W_2) = 1$ מתקבלת (למשל) עבור

$$W_1 = \mathrm{Span}(e_1, e_2, e_3), W_2 = (e_3, e_4) \implies W_1 \cap W_2 = \mathrm{Span}(e_3).$$

באופן דומה, האפשרות ל- $\dim(W_1 \cap W_2) = 2$ מתקבלת (למשל) עבור

$$W_1 = \mathrm{Span}(e_1, e_2, e_3), W_2 = \mathrm{Span}(e_2, e_3) \implies W_1 \cap W_2 = W_2.$$

כדוגמה לאפשרות האחרונה, ניתן לבחור כל W_1 ממימד 3 וכל $W_2 \subseteq W_1$ ממימד 2.

9.3 מציאת בסיסים

יהיו W_1, W_2 תת-מרחבים של מ"ו V מעל $\mathbb{F}$. הדבר החשוב ביותר שיש לדעת על תת-מרחב הוא בסיס. נניח שמצאנו בסיס $B_1 = \{ w_1, ..., w_m \}$ ל- W_1 ובסיס $B_2 = \{ u_1, ..., u_k \}$ ל- W_2. אז איך נמצא בסיסים ל- $W_1 \cap W_2, W_1 + W_2$? מתקיים

$$W_1 + W_2 = \mathrm{Span}(w_1, ..., w_m, u_1, ..., u_k),$$

ולכן כבר יש לנו קבוצה פורשת - אך היא לא בהכרח בת"ל. למעשה, היא בת"ל אם ורק אם $W_1 \cap W_2 = \{ 0_V \}$. אכן, לפי משפט 8.82 ("שלישי חינם") קבוצה פורשת היא בת"ל אם ורק אם גודלה שווה למימד תת-המרחב, ולפי נוסחת המימדים מתקיים:

$$\begin{aligned} \dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) &\iff \dim(W_1 \cap W_2) = 0 \\ &\iff W_1 \cap W_2 = \{ 0_V \} \end{aligned}$$

אם זה לא המקרה, צריך לחלץ בסיס מתוך $\{ w_1, ..., w_m, u_1, ..., u_k \}$. נניח כי $V = \mathbb{F}^n$ (אחרת נשתמש בוקטורי קואורדינטות) ונסתכל על המטריצה הבאה:

$$\begin{pmatrix} | & & | & | & & | \\ w_1 & \cdots & w_m & u_1 & \cdots & u_k \\ | & & | & | & & | \end{pmatrix}$$

בתת-פרק 8.5.2 ראינו שניתן למצוא בסיס למרחב העמודות ע"י דירוג קנוני, לפיו העמודות בהן יש איברים מובילים מתאימות לעמודות המקוריות שמהוות יחד בסיס למרחב העמודות (אין לקחת את העמודות החדשות, כי הדירוג משנה את מרחב העמודות). במקרה זה, מרחב העמודות הוא $W_1 + W_2$.

מה לגבי בסיס ל- $W_1 \cap W_2$? נשאלת השאלה: אילו צירופים לינאריים של $w_1, ..., w_m$ הם גם צירופים לינאריים של $u_1, ..., u_k$? באופן שקול, צריך למצוא $\alpha_1, ..., \alpha_m, \gamma_1, ..., \gamma_k \in \mathbb{F}$ כך ש-

$$\alpha_1 w_1 + ... + \alpha_m w_m = \gamma_1 u_1 + ... + \gamma_k u_k. \tag{9.3}$$

לאחר העברת אגפים מקבלים ממ"ל הומוגנית שמיוצגת ע"י המטריצה הבאה:

$$\begin{pmatrix} | & & | & | & & | \\ w_1 & \cdots & w_m & -u_1 & \cdots & -u_k \\ | & & | & | & & | \end{pmatrix}$$

זו לא בדיוק המטריצה הקודמת, ולכאורה דרוש דירוג קינוני חדש. אך יש דרך קיצור, כי במשוואה **9.3** ניתן להציב $\gamma_1 = -\beta_1, ..., \gamma_k = -\beta_k$, ואז המטריצה המתאימה לממ"ל החדשה בנעלמים $\alpha_1, ..., \alpha_m, \beta_1, ..., \beta_k$ היא המטריצה שכבר דירגנו:

$$\begin{pmatrix} | & & | & | & & | \\ w_1 & \cdots & w_m & u_1 & \cdots & u_k \\ | & & | & | & & | \end{pmatrix}$$

כך נמצא את הפתרון הכללי עבור $\alpha_1, ..., \alpha_m$ ונציב אותו בביטוי $\alpha_1 w_1 + ... + \alpha_m w_m$ כדי לקבל את הצורה הכללית של איבר ב- $W_1 \cap W_2$. זה יוביל לקבוצה פורשת, שהיא למעשה בסיס.

דוגמה 9.20. נסתכל על תת-המרחבים הבאים של $\mathbb{R}^4$:

$$W_1 = \text{Span}\left(\begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\0\\2\\0\end{pmatrix}\right), W_2 = \text{Span}\left(\begin{pmatrix}2\\1\\3\\1\end{pmatrix}, \begin{pmatrix}1\\-1\\1\\-1\end{pmatrix}\right)$$

מתקיים

$$W_1 + W_2 = \text{Span}\left(\begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\0\\2\\0\end{pmatrix}, \begin{pmatrix}2\\1\\3\\1\end{pmatrix}, \begin{pmatrix}1\\-1\\1\\-1\end{pmatrix}\right),$$

אך צריך לבדוק אם הקבוצה בת"ל. נכתוב את ארבעת הוקטורים כעמודות מטריצה, ולאחר דירוג קנוני נקבל:

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

אין איבר מוביל בעמודה השלישית, ולכן הוקטור $\begin{pmatrix}2\\1\\3\\1\end{pmatrix}$ מיותר. אז הקבוצה הבאה היא בסיס ל- W_1+W_2:

$$\left\{\begin{pmatrix}1\\1\\1\\1\end{pmatrix},\begin{pmatrix}1\\0\\2\\0\end{pmatrix},\begin{pmatrix}1\\-1\\1\\-1\end{pmatrix}\right\}$$

הדירוג גם מראה שאם מתקיים

$$\alpha_1\begin{pmatrix}1\\1\\1\\1\end{pmatrix}+\alpha_2\begin{pmatrix}1\\0\\2\\0\end{pmatrix}=-\beta_1\begin{pmatrix}2\\1\\3\\1\end{pmatrix}-\beta_2\begin{pmatrix}1\\-1\\1\\-1\end{pmatrix},$$

אז בהכרח קיים $t\in\mathbb{R}$ כך ש-

$$\alpha_1=\alpha_2=-t,\beta_1=t,\beta_2=0.$$

לכן:

$$W_1\cap W_2=\left\{-t\begin{pmatrix}1\\1\\1\\1\end{pmatrix}-t\begin{pmatrix}1\\0\\2\\0\end{pmatrix}\;\middle|\;t\in\mathbb{R}\right\}=\operatorname{Span}\left(\begin{pmatrix}2\\1\\3\\1\end{pmatrix}\right)$$

לכן, הקבוצה הבאה היא בסיס ל- $W_1\cap W_2$:

$$\left\{\begin{pmatrix}2\\1\\3\\1\end{pmatrix}\right\}$$

נשים לב כי נוסחת המימדים אכן מתקיימת:

$$3=\dim(W_1+W_2)=\dim(W_1)+\dim(W_2)-\dim(W_1\cap W_2)=2+2-1$$

תרגיל 9.21. נסתכל על תת-המרחבים הבאים של $\mathbb{R}^4$:

$$W_1=\mathrm{N}(A),W_2=\operatorname{Span}\left(\begin{pmatrix}1\\0\\0\\-1\end{pmatrix},\begin{pmatrix}1\\-1\\1\\-1\end{pmatrix}\right),$$

כאשר

$$.A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \end{pmatrix}$$

מצאו בסיסים ל- $W_1 \cap W_2, W_1 + W_2$.

פתרון. דרך אחת לפתור את התרגיל היא למצוא בסיס ל- $\mathrm{N}(A)$. בבסיס זה יש שני וקטורים, ולכן יחד עם הוקטורים בקבוצה הפורשת את W_2, נקבל מטריצה של 4 עמודות. אין שום בעיה עם דרך זו, אך היא דורשת שני דירוגים וכדאי להכיר גישה קצת שונה.

אמנם המרחב הוא $\mathbb{R}^4$ ולא $\mathbb{R}^3$, אך יש דמיון לחיתוך בין מישור שנתון ע״י ממ״ל לישר שנתון בהצגה פרמטרית. כאן W_1 נתון ע״י ממ״ל ו- W_2 נתון בהצגה פרמטרית (אבל שניהם ממימד 2). גם כאן ניתן למצוא את החיתוך ע״י הצבת ההצגה הפרמטרית של W_2 בממ״ל המתאימה ל- A. נשים לב כי

$$.W_2 = \left\{ s \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix} + t \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} \;\middle|\; s, t \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} s+t \\ -t \\ t \\ -s-t \end{pmatrix} \;\middle|\; s, t \in \mathbb{R} \right\}$$

הממ״ל המתאימה ל- A היא:

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 - x_2 + x_3 - x_4 = 0 \end{cases}$$

נציב $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} s+t \\ -t \\ t \\ -s-t \end{pmatrix}$ ונקבל:

$$\begin{cases} s + t - t + t - s - t = 0 \\ s + t + t + t + s + t = 0 \end{cases} \implies s = -2t$$

נחזור להצגה של W_2 ונציב $s = -2t$:

$$W_2 \left\{ \begin{pmatrix} -t \\ -t \\ t \\ t \end{pmatrix} \;\middle|\; s, t \in \mathbb{R} \right\} = \mathrm{Span}\left(\begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix} \right)$$

קיבלנו שהקבוצה $\left\{\begin{pmatrix}-1\\-1\\1\\1\end{pmatrix}\right\}$ היא בסיס ל- $W_1 \cap W_2$, ולכן $\dim(W_1 \cap W_2) = 1$.

בנוסף, $\dim(W_2) = 2$ כי הקבוצה הפורשת אותו היא בת"ל (הוקטורים אינם קו-לינאריים) ולכן בסיס. באופן דומה, $\dim(W_1) = 4 - \operatorname{rank}(A) = 4 - 2 = 2$ לפי משפט 8.87 (משפט הדרגה). לכן, לפי נוסחת המימדים נובע כי

$$\dim(W_1 + W_2) = 2 + 2 - 1 = 3\ .$$

כדי למצוא בסיס ל- $W_1 + W_2$ מספיק להוסיף לקבוצה הפורשת את W_2 כל וקטור ב- W_1 שאינו בחיתוך, כי זה בדיוק אומר שהוא לא נפרש ע"י וקטורי הבסיס של W_2. כך נקבל קבוצה בת"ל ב- $W_1 + W_2$ שגודלה שווה למימדו, ולכן היא בהכרח בסיס.

מספיק למצוא וקטור אחד ב- $\mathrm{N}(A)$ שאינו כפל בסקלר של $\begin{pmatrix}-1\\-1\\1\\1\end{pmatrix}$. דוגמה לוקטור כזה היא $\begin{pmatrix}1\\0\\-1\\0\end{pmatrix}$, שאפשר למצוא ע"י הצבת $x_2 = x_4 = 0$ במשוואות או ע"י דירוג A. למען הסר ספק, צורתה המדורגת קנונית של A היא

$$\begin{pmatrix}1 & 0 & 1 & 0\\0 & 1 & 0 & 1\end{pmatrix}\ .$$

בכל אופן, נוסיף את הוקטור שמצאנו לבסיס הנתון ל- W_2 ונקבל את הבסיס הבא ל- $W_1 + W_2$:

$$\left\{\begin{pmatrix}1\\0\\0\\-1\end{pmatrix}, \begin{pmatrix}1\\-1\\1\\-1\end{pmatrix}, \begin{pmatrix}1\\0\\-1\\0\end{pmatrix}\right\}$$