אלגברה לינארית 2א

מאיה פרבר ברודסקי

סיכומי ההרצאות של פרופסור דוד גינסבורג, סמסטר א' תש"ף.

תוכן עניינים

חלק I
מבוא

1 חוגים

הגדרה 1.1 קבוצה לא ריקה R תקרא **חוג** אם ב־R מוגדרות שתי פעולות $+,\cdot$ כך שמתקיימים התנאים הבאים:

1. סגירות תחת חיבור: לכל $a,b\in R$, $a+b\in R$.
2. אסוציאטיביות ביחס לחיבור: לכל $a,b,c\in R$, $(a+b)+c=a+(b+c)$.
3. קומטטיביות ביחס לחיבור: לכל $a,b\in R$, $a+b=b+a$.
4. קיום איבר האפס: קיים $0\in R$ כך שלכל $a\in R$ מתקיים $0+a=a$.
5. קיום איבר נגדי: לכל $a\in R$ קיים $-a\in R$ כך ש־$a+(-a)=0$.
6. סגירות תחת כפל: לכל $a,b\in R$, $a\cdot b\in R$.
7. אסוציאטיביות ביחס לכפל: לכל $a,b,c\in R$, $(a\cdot b)\cdot c=a\cdot(b\cdot c)$.
8. חוק הפילוג: לכל $a,b,c\in R$, $a\cdot(b+c)=a\cdot b+a\cdot c$.

כלומר, חוג הינו שדה ללא תכונות הקומטטיביות ביחס לכפל, קיום איבר יחידה וקיום איבר הופכי.

הגדרה 1.2 אם קיים איבר $1\in R$ כך שלכל $a\in R$ מתקיים $1\cdot a=a\cdot 1=a$, נאמר ש־R הוא **חוג עם יחידה**.
אם לכל $a,b\in R$ מתקיים $a\cdot b=b\cdot a$, נאמר ש־R הוא **חוג חילופי**.

1.1 חוגי פולינומים

הגדרה 1.3 יהי $\mathbb{F}$ שדה. נגדיר את **חוג הפולינומים** במשתנה x מעל השדה $\mathbb{F}$ כאוסף כל הביטויים מהצורה $a_0+a_1x+\cdots+a_nx^n$, עבור $a_i\in\mathbb{F}$. נסמן חוג זה ב־$\mathbb{F}[x]$.

הגדרה 1.4 עבור זוג פולינומים $p(x)=a_0+a_1x+\cdots+a_nx^n, q(x)=b_0+b_1x+\cdots+b_mx^m$, נגדיר

1. $p(x),q(x)$ יקראו שווים ויסומנו $p(x)=q(x)$ אם לכל $i\geq 0$, $a_i=b_i$.
2. חיבור: בהנחה ש־$m\geq n$, $p(x)+q(x)=(a_0+b_0)+\ldots(a_n+b_n)x^n+b_{n+1}x^{n+1}+\cdots+b_mx^m$.
3. כפל: $p(x)\cdot q(x)=c_0+c_1x+\cdots+c_kx^k$, כאשר $c_t=\sum_{i=0}^{t}a_ib_{t-i}$.

טענה 1.5 בהינתן הגדרות אלו, $\mathbb{F}[x]$ הינו חוג חילופי עם יחידה.

הגדרה 1.6 אם $p(x)=a_0+a_1x+\cdots+a_nx^n$ פולינום ב־$\mathbb{F}[x]$ כך ש־$a_n\neq 0$, נאמר **שדרגת** $p(x)$ היא n, ונסמן זאת $\deg p(x)=n$. דרגת פולינום האפס אינה מוגדרת (ניתן גם להגדיר אותה להיות $-\infty$).

טענה 1.7 אם $f(x),g(x)$ שונים מאפס אז $\deg(f(x)+g(x))=\deg f(x)+\deg g(x)$.

הוכחה: נניח כי $f(x)=a_0+a_1x+\cdots+a_mx^m, g(x)=b_0+b_1x+\cdots+b_nx^n$ ו־$a_m,b_n\neq 0$.
לכן המקדם ה־$m+n$ של $f(x)\cdot g(x)$ הוא $\sum_{i=0}^{m+n}a_ib_{m+n-i}=a_mb_n\neq 0$, כלומר $\deg(f(x)\cdot g(x))\geq m+n$.
כדי להוכיח שוויון, צריך להוכיח שלכל $i>m+n$, המקדם ה־i של $f(x)\cdot g(x)$ הוא 0 כלומר $\sum_{j=0}^{i}a_jb_{i-j}=0$.
אכן, לכל j כיוון ש־$i=j+(i-j)>m+n$ מתקיים בהכרח $j>m$ או $i-j>n$, לכן $a_j=0$ או $b_{i-j}=0$, ובכל אופן $a_j\cdot b_{i-j}=0$. לכן גם הסכום הוא 0. ■

מסקנה 1.8 אם $f(x),g(x)\in\mathbb{F}[x]$ שונים מאפס אז $\deg f(x)\leq\deg(f(x)\cdot g(x))$.

מסקנה 1.9 ב־$\mathbb{F}[x]$ אין מחלקי אפס, כלומר אם $f(x)\cdot g(x)=0$ אז $f(x)=0$ או $g(x)=0$.

הגדרה 1.10 נאמר ש־$f(x)$ **מחלק את** $g(x)$ ונסמן $f(x)\mid g(x)$, אם קיים פולינום $p(x)$ כך ש־$g(x)=p(x)\cdot f(x)$.

הגדרה 1.11 בהינתן $f(x), g(x)$ שונים מאפס נגדיר את **המחלק המשותף הגדול ביותר** להיות הפולינום $p(x)$ המקיים $p(x) \mid f(x)$ וגם $p(x) \mid g(x)$, וכמו כן לכל $q(x)$ אם $q(x) \mid f(x)$ וגם $q(x) \mid g(x)$ אז $q(x) \mid p(x)$. נסמן זאת ב־$p(x) = (f(x), g(x))$, או gcd.

הערה 1.12 המחלק המשותף הגדול ביותר הוא יחיד עד כדי כפל בקבוע, ובדרך כלל ננרמל אותו לכדי פולינום מתוקן (פולינום בעל מקדם מוביל 1).

טענה 1.13 יהיו $f(x), g(x) \in \mathbb{F}[x]$ שונים מאפס. אזי קיימים $t(x), r(x) \in \mathbb{F}[x]$ כך ש־$f(x) = t(x) \cdot g(x) + r(x)$, כאשר $r(x) = 0$ או $\deg r(x) < \deg g(x)$.

הוכחה: אם $\deg f(x) < \deg g(x)$ אז נבחר $t(x) = 0$, $r(x) = f(x)$.

כעת, נניח $\deg f(x) \geq \deg g(x)$, ונסמן $f(x) = a_0 + \cdots + a_m x^m, g(x) = b_0 + \cdots + b_n x^n$ עבור $a_m, b_n \neq 0$, אז מההנחה $m \geq n$. נוכיח את הטענה באינדוקציה על m.

נגדיר $f_1(x) = f(x) - \frac{a_m}{b_n} x^{m-n} g(x)$, אז $\deg f_1(x) \leq m-1$. מהנחת האינדוקציה $f_1(x) = t_1(x) g(x) + r(x)$, כאשר $r(x) = 0$ או $\deg r(x) < \deg g(x)$. בסך הכל $f(x) - \frac{a_m}{b_n} x^{m-n} g(x) = t_1(x) g(x) + r(x)$, נעביר אגפים ונקבל

$$f(x) = t_1(x) g(x) + r(x) + \frac{a_m}{b_n} x^{m-n} g(x) = \left(t_1(x) + \frac{a_m}{b_n} x^{m-n}\right) g(x) + r(x)$$

כדרוש. ■

אלגוריתם 1.14 חילוק פולינומים ־ חילוק ארוך, מוצג בהוכחה.

אלגוריתם 1.15 מציאת gcd ־ **אלגוריתם אוקלידס**.

בהינתן $f(x), g(x) \in \mathbb{F}[x]$ נרצה לחשב את $(f(x), g(x))$. נניח ש־$\deg f(x) \geq \deg g(x)$, וממשפט החלוקה ניתן לרשום

$$f(x) = q_0(x) g(x) + r_1(x) \qquad (1)$$

כאשר $r_1(x) = 0$ או $\deg r_1(x) < \deg g(x)$. אם $r_1(x) = 0$, אז $g(x) \mid f(x)$ ומתקיים $g(x) = (f(x), g(x))$. אחרת, אם $\deg r_1(x) < \deg g(x)$, אז

$$\begin{aligned} g(x) &= q_1(x) r_1(x) + r_2(x) && (2) \\ r_1(x) &= q_2(x) r_2(x) + r_3(x) && (3) \\ &\vdots \\ r_{n-2}(x) &= q_{n-1}(x) r_{n-1}(x) + r_n(x) && (n) \\ r_{n-1}(x) &= q_n(x) r_n(x) && (n+1) \end{aligned}$$

דרגת השארית קטנה ממש בכל שלב, ולכן בהכרח השארית תתאפס לאחר מספר צעדים מסוים n. אז יתקיים $r_n = (f(x), g(x))$, כפי שנוכיח בטענה הבאה.

טענה 1.16 $r_n(x) = (f(x), g(x))$.

הוכחה: נוכיח באינדוקציה.

תחילה נוכיח $r_n(x) \mid g(x)$, $r_n(x) \mid f(x)$. ממשוואה $(n+1)$ נקבל ש־$r_n(x) \mid r_{n-1}(x)$. ממשוואה (n) נקבל ש־$r_n(x) \mid r_{n-2}(x)$. באינדוקציה נקבל מ־(2) ש־$r_n(x) \mid g(x)$, ומ־(1) ש־$r_n(x) \mid f(x)$.

כעת, יהי $p(x)$ פולינום כך ש־$p(x) \mid g(x), p(x) \mid f(x)$. מ־$(1)$ נקבל ש־$p(x) \mid r_1(x)$, ומ־(2) נקבל $p(x) \mid r_2(x)$, ובאינדוקציה נקבל $p(x) \mid r_n(x)$. ■

הגדרה 1.17 פולינום $f(x) \in \mathbb{F}[x]$ יקרא **אי־פריק** מעל $\mathbb{F}$ אם השוויון $f(x) = p(x) \cdot q(x)$ כאשר $p(x), q(x) \in \mathbb{F}[x]$, מתקיים רק אם $\deg p(x) = 0$ או $\deg q(x) = 0$.

הגדרה 1.18 בהינתן מספרים $a_1, \ldots, a_r \in \mathbb{Z}$ נסמן ב־$(a_1, \ldots, a_r)$ את המחלק המשותף הגדול ביותר שלהם.

פולינום $f(x) = a_0 + a_1 x + \cdots + a_n x^n$ שמקדמיו שלמים יקרא **פולינום פרימיטיבי** אם $(a_0, a_1, \ldots, a_n) = 1$.

משפט 1.19 הלמה של גאוס: אם $f(x) \in \mathbb{Z}[x]$ פולינום פרימיטיבי ניתן לפירוק מעל $\mathbb{Q}$ אז הוא ניתן לפירוק כמכפלה של שני פולינומים עם מקדמים שלמים.

משפט 1.20 קריטריון אייזנשטיין: יהי $f(x)=a_0+a_1x+\cdots+a_nx^n$ פולינום עם מקדמים שלמים. אם קיים מספר ראשוני p כך ש$p|a_0,p|a_1,\ldots,p|a_{n-1}$ אבל $p\nmid a_n$ וגם $p^2\nmid a_0$, אז $f(x)$ הוא אי פריק מעל $\mathbb{Q}$.

הערה 1.21 לא כל פולינום אי פריק מקיים את קריטריון אייזנשטיין.

הוכחה: יהי $d=(a_0,a_1,\ldots,a_n)$ ונרשום $f(x)=d\cdot q(x)$. כיוון ש$p\nmid a_n$ אז $p\nmid d$, לכן בלי הגבלת הכלליות ניתן להניח ש$f(x)$ פרימיטיבי. נניח בשלילה ש־$f(x)$ פריק מעל $\mathbb{Q}$, מהלמה של גאוס נקבל שניתן לרשום

$$f(x)=(b_0+b_1x+\cdots+b_rx^r)(c_0+c_1x+\cdots+c_sx^s)$$

כאשר $b_j,c_i\in\mathbb{Z}$ ו$r,s>0$. מתקיים $a_0=b_0c_0$, נתון כי $p\mid a_0$ אבל $p^2\nmid a_0$ לכן ניתן להניח בה"כ ש$p\mid b_0,p\nmid c_0$. p אינו מחלק את b_j לכל j, כי אם כן אז $p\mid a_n$ וזה לא יתכן. יהי b_k המקדם הראשון המקיים ש$p\nmid b_k$ ו$k\le r<n$, כלומר $p\mid b_j$ לכל $j<k$ אבל $p\nmid b_k$. מהגדרת כפל פולינומים $a_k=b_kc_0+b_{k-1}c_1+\cdots+b_0c_k$. מנתון $p\mid a_k$, ומבחירת b_k נובע $p\mid b_0,\ldots,b_{k-1}$. לכן $p\mid b_kc_0$ ולכן $p\mid c_0$ או $p\mid b_k$, סתירה. אז $f(x)$ אי פריק מעל $\mathbb{Q}$. ■

הגדרה 1.22 יהי $p(x)\in\mathbb{F}[x]$. נאמר ש$a\in\mathbb{F}$ הוא **שורש** של $p(x)$ אם $p(a)=0$.

טענה 1.23 יהיו $p(x)\in\mathbb{F}[x],b\in\mathbb{F}$ אז קיים $q(x)\in\mathbb{F}[x]$ כך ש$\deg q(x)=\deg p(x)-1$ עבורו מתקיים $p(x)=(x-b)q(x)+p(b)$.

הוכחה: ממשפט החלוקה קיים $q(x)\in\mathbb{F}[x]$ כך ש־$p(x)=(x-b)q(x)+r(x)$, כאשר $\deg r(x)<\deg(x-b)=1$ או $r(x)=0$. לכן בכל מקרה $r\in\mathbb{F}$ קבוע. נציב $x=b$, ונקבל $p(b)=(b-b)q(b)+r=r$. ■

מסקנה 1.24 יהי $p(x)\in\mathbb{F}[x]$, אז $a\in\mathbb{F}$ הוא שורש של $p(x)$ אם ורק אם $(x-a)\mid p(x)$.

הגדרה 1.25 יהי $p(x)\in\mathbb{F}[x]$ ו־$a\in\mathbb{F}$ שורש של $p(x)$. נאמר ש**הריבוי האלגברי** של a הוא m אם $(x-a)^m\mid p(x)$ אבל $(x-a)^{m+1}\nmid p(x)$.

טענה 1.26 יהי $p(x)\in\mathbb{F}[x]$. אם $\deg p(x)=n$ אז ל$p(x)$ יש לכל היותר n שורשים ב־$\mathbb{F}$ (כולל ריבוי).

הוכחה: נוכיח באינדוקציה על n. אם $n=1$, אז $p(x)=\alpha x+\beta$ ואז יש שורש אחד, או אין שורשים.
נניח $a\in\mathbb{F}$ הוא שורש עם ריבוי אלגברי m, אז $(x-a)^m\mid p(x)$ ולכן $m\le n$. כמו כן, $p(x)=(x-a)^mq(x)$, ו־$\deg q(x)=n-m$. כיוון ש$(x-a)^{m+1}\nmid p(x)$ אז $(x-a)\nmid q(x)$, לכן a אינו שורש של $q(x)$. אם $a\neq b\in\mathbb{F}$ הוא שורש של $p(x)$ אז $0=p(b)=(b-a)^mq(b)$. כיוון ש$\mathbb{F}$ שדה בהכרח $q(b)=0$, מהנחת האינדוקציה ל$q(x)$ יש לכל היותר $n-m$ שורשים, ולכן לp לכל היותר n שורשים בסך הכל. ■

הערה 1.27

1. לפולינומים מעל חוג שאינו שדה יתכנו יותר שורשים, לדוגמה אם $\mathbb{F}$ הוא חוג הקווטרניונים (quaternions) אז ל־x^2+1 יש לפחות שלושה שורשים שונים i,j,k.
2. המשפט היסודי של האלגברה: כל פולינום מעל $\mathbb{C}$ ניתן לרשום כמכפלה של גורמים לינאריים.
3. מסקנה מהמשפט היסודי של האלגברה: כל פולינום מעל $\mathbb{R}$ ניתן לרשום כמכפלה של גורמים לינאריים וגורמים ריבועיים.

חלק II
צורות קנוניות

2 לכסון

יהי V מרחב וקטורי ממימד סופי מעל שדה $\mathbb{F}$, $T\colon V \to V$ העתקה לינארית.

הגדרה 2.1 סקלר $\lambda \in \mathbb{F}$ יקרא **ערך עצמי** עבור T אם קיים וקטור $0 \neq v \in V$ כך ש-$T(v) = \lambda v$. הוקטור v יקרא **וקטור עצמי** השייך לערך העצמי λ.

טענה 2.2 λ הוא ערך עצמי של T אם ורק אם ההעתקה $\lambda I - T$ אינה הפיכה.

הוכחה: לכל $\lambda \in \mathbb{F}, 0 \neq v \in V$ מתקיים

$$T(v) = \lambda v \iff \lambda v - T(v) = 0 \iff (\lambda I - T)(v) = 0 \iff \ker(\lambda I - T) \neq \{0\}$$

כמו כן $\ker(\lambda I - T) \neq \{0\}$ אם ורק אם $\lambda I - T$ אינה חח"ע, וזה קורה אם ורק אם $\lambda I - T$ אינה הפיכה. ■

טענה 2.3 יהי $\lambda \in \mathbb{F}$ ערך עצמי של T, אז לכל $q(x) \in \mathbb{F}[x]$, $q(\lambda)$ הינו ערך עצמי של $q(T)$.

הוכחה: יהי $v \in V$ וקטור עצמי המתאים ל-$\lambda \in \mathbb{F}$. באינדוקציה

$$T^k v = T^{k-1} T v = \lambda T^{k-1} v \overset{\text{ind. hyp.}}{=} \lambda \lambda^{k-1} v = \lambda^k v$$

אם $q(x) = a_0 + a_1 x + \cdots + a_n x^n$ אז

$$\begin{aligned} q(T)v &= (a_0 I + a_1 T + \cdots + a_n T^n) v = a_0 v + a_1 T v + \cdots + a_n T^n v = a_0 v + a_1 \lambda v + \cdots + a_n \lambda^n v \\ &= (a_0 + a_1 \lambda + \cdots + a_n \lambda^n) v = p(\lambda) v \end{aligned}$$

■

הערה 2.4 אם v הוא וקטור עצמי של T עבור λ, אז v הוא וקטור עצמי של $q(T)$ עבור $q(\lambda)$.

טענה 2.5 תהי $T\colon V \to V$ ונניח $\dim V = n$, אז קיים $0 \neq q(x) \in \mathbb{F}[x]$ כך ש-$\deg q(x) \leq n^2$, עבורו $q(T) = 0$.

הוכחה: מימד מרחב העתקות הלינאריות מעל V הוא n^2, לכן $I, T, T^2, \ldots, T^{n^2}$ תלויים לינארית והתלות הלינארית ביניהם היא בדיוק פולינום ממעלה לכל היותר n^2 המאפס את T. ■

טענה 2.6 יהי $k \leq n^2$ המספר הקטן ביותר כך שקיים פולינום $q(x)$ שדרגתו k המקיים $q(T) = 0$. אם $p(x)$ הוא פולינום כך ש-$p(T) = 0$, אז $q(x) \mid p(x)$.

הוכחה: ממשפט החלוקה נקבל $p(x) = l(x) q(x) + r(x)$ עבור $r(x) = 0$ או $\deg r(x) < \deg q(x)$. **נציב** T ונקבל $0 = p(T) = l(T) q(T) + r(T) = l(T) \cdot 0 + r(T) = r(T)$. אם $\deg r(x) < \deg q(x)$ אז נקבל סתירה למינימליות של k, לכן $r(x) = 0$ כלומר $q(x) \mid p(x)$. ■

מסקנה 2.7 בין הפולינומים שדרגתם k המאפסים את T, $q(x)$ הוא יחיד עד כדי כפל בקבוע.

טענה 2.8 בהינתן $T\colon V \to V$ קיים מספר k וקיים פולינום $m(x)$ אחד ויחיד המקיים:

1. $m(T) = 0$
2. $m(x) = x^k + a_{k-1} x^{k-1} + \cdots + a_1 x + a_0$
3. אם $p(x) \in \mathbb{F}[x]$ מקיים $p(T) = 0$, אז $\deg p(x) \geq k$.

הגדרה 2.9 פולינום $m(x)$ זה ייקרא **הפולינום המינימלי** של T ויסומן ב-$m_T(x)$.

משפט 2.10 אם $\lambda \in \mathbb{F}$ הוא ערך עצמי של T אז λ הוא שורש של $m_T(x)$. בפרט, ל-T יש מספר סופי של ערכים עצמיים שונים. כיוון ש-$\deg m_T(x) \leq n^2$, אז ל-T יש לכל היותר n^2 ערכים עצמיים שונים.

הוכחה: נתון $m_T(T)=0$. אם $\lambda\in\mathbb{F}$ הינו ערך עצמי של T עם וקטור עצמי v, אז מטענה קודמת ניתן להסיק שv הוא וקטור עצמי של $m_T(T)$ עבור ערך עצמי $m_T(\lambda)$, כלומר $m_T(\lambda)v=m_T(T)v=0v=0$. כיוון ש$v\neq 0$ אז $m_T(\lambda)=0$. ■

טענה 2.11 יהיו $T,S\colon V\to V$ כך שS הפיכה. אז לT ולSTS^{-1} יש אותו פולינום מינימלי.

משפט 2.12 יהיו $\lambda_1,\dots,\lambda_k\in\mathbb{F}$ ערכים עצמיים שונים של T. אם $v_1,\dots,v_k$ הם הוקטורים העצמיים המתאימים לערכים עצמיים אלו בהתאמה, אז $v_1,\dots,v_k$ בלתי תלויים לינארית.

הוכחה: יהיו $a_1,\dots,a_k\in\mathbb{F}$ כך ש$a_1v_1+\dots+a_kv_k=0$.

נוכיח באינדוקציה על k. אם $k=1$, אז v_1 היא קבוצה בלתי תלויה. נניח שהוכחנו עבור $k-1$, ונוכיח עבור k.

נפעיל את T על התלות הלינארית ונקבל $0=T(a_1v_1+\dots+a_kv_k)=a_1\lambda_1v_1+\dots+a_k\lambda_kv_k$, וכיוון ש$k>1$ וכל הע"ע שונים אז בהכרח קיים i כך ש$\lambda_i\neq 0$. בה"כ על ידי ארגון מחדש ניתן להניח $\lambda_k\neq 0$.

נכפיל את התלות הלינארית בλ_k ונקבל $a_1\lambda_kv_k+\dots+a_k\lambda_kv_k=0$. נחסיר את תוצאה זו מהתוצאה הקודמת ונקבל $a_1(\lambda_1-\lambda_k)v_1+\dots+a_{k-1}(\lambda_{k-1}-\lambda_k)v_k=0$. לפי הנחת האינדוקציה $v_1,\dots,v_k$ בת"ל, כלומר לכל $1\leq i\leq k-1$ מתקיים $a_i(\lambda_i-\lambda_k)=0$. $\lambda_i\neq\lambda_k$, לכן בהכרח $a_i=0$. אז $a_1,\dots,a_{k-1}=0$, ומהתלות הלינארית נסיק $a_kv_k=0$ ולכן גם $a_k=0$. ■

מסקנה 2.13 אם $\dim V=n$ אז ל$T\colon V\to V$ יש לכל היותר n ערכים עצמיים שונים.

מסקנה 2.14 אם $\dim V=n$ ול$T\colon V\to V$ יש n ערכים עצמיים שונים אז לV יש בסיס המורכב מוקטורים עצמיים.

הגדרה 2.15 אם לV יש בסיס של וקטורים עצמיים של העתקה T, נאמר שT **ניתנת ללכסון**. כלומר, קיים בסיס e של וקטורים עצמיים שעבורו $[T]_e$ אלכסונית (ובפרט, הערכים על האלכסון הם הערכים העצמיים המתאימים).

טענה 2.16 עבור $T\colon V\to V$ העתקה לינארית וe בסיס של V, λ הוא ערך עצמי של T $\iff$ λ הוא ערך עצמי של $[T]_e$.

הוכחה: $[T]_e[v]_e=\lambda[v]_e \iff [Tv]_e=\lambda[v]_e \iff Tv=\lambda v$ ■

הגדרה 2.17 עבור מטריצה A, מתקיים $(\lambda I-A)v=0 \iff Av=\lambda v$. כלומר, λ הוא ערך עצמי ורק אם יש פתרון, כלומר אם ורק אם הוא שורש של הפולינום

$$\Delta_A(x)=\det(xI-A)$$

פולינום זה נקרא **הפולינום האופייני** של A.

אם T העתקה לינארית, נגדיר את הפולינום האופייני $\Delta_T(x)=\det(xI-[T]_e)$ כאשר e בסיס כלשהו.

טענה 2.18 $\deg\Delta_A(x)=n$. כמו כן, $\Delta_A(x)=x^n-\operatorname{tr}(A)x^{n-1}+\dots+ax+(-1)^n\det A$.

הוכחה: תרגיל. ■

הערה 2.19 אם נתונה מטריצה מסדר n שאיבריה הם פולינומים, אז ניתן לרשום אותה כפולינום שמקדמיו הם מטריצות מסדר n.

משפט 2.20 קיילי המילטון: כל מטריצה A היא שורש של הפולינום האופייני $\Delta_A(x)$.

הוכחה: על פי הגדרה, $\Delta_A(x)=\det(xI-A)=x^n+a_{n-1}x^{n-1}+\dots+a_0$.

נסמן $B(x)=\operatorname{adj}(xI-A)$, אז מהגדרת המטריצה הצמודה נקבל ש$B(x)$ הינה מטריצה שאיברה הם פולינומים שדרגתם הם לכל היותר $n-1$. נרשום את $B(x)$ כפולינום שמקדמיו הם מטריצות עם מקדמים ב$\mathbb{F}$, $B(x)=B_{n-1}x^{n-1}+\dots+B_1x+B_0$. מהמשפט $D\operatorname{adj}D=\det(D)I$ נקבל $(xI-A)B(x)=\det(xI-A)I$. לכן $(xI-A)\left(B_{n-1}x^{n-1}+\dots+B_1x+B_0\right)=\left(x^n+a_{n-1}x^{n-1}+\dots+a_0\right)I$. נפתח סוגריים ונשווה מקדמים

$$\begin{aligned} B_{n-1}&=I\\ B_{n-2}-AB_{n-1}&=a_{n-1}I\\ &\vdots\\ B_0-AB_1&=a_1I\\ -AB_0&=a_0I \end{aligned}$$

נכפיל משמאל ב$A^n, A^{n-1}, \dots, A, I$ ונקבל

$$\begin{aligned} A^n B_{n-1} &= A^n \\ A^{n-1}B_{n-2} - A^n B_{n-1} &= a_{n-1}A^{n-1} \\ &\vdots \\ AB_0 - A^2 B_1 &= a_1 A \\ -AB_0 &= a_0 I \end{aligned}$$

נחבר את כל המשוואות ונקבל

$$0 = A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I = \Delta_A(A)$$

■

מסקנה 2.21 $m_A(x) | \Delta_A(x)$, על פי הגדרת הפולינום המינימלי.

משפט 2.22 $\Delta_A(x) \,|\, (m_A(x))^n$, כאשר n הוא סדר המטריצה A.

הוכחה: נניח כי $m_A(x) = x^r + c_1 x^{r-1} + \cdots + c_r$. נגדיר

$$\begin{aligned} B_0 &= I \\ B_1 &= A + c_1 I \\ B_2 &= A^2 + c_1 A + c_2 I \\ &\vdots \\ B_{r-1} &= A^{r-1} + c_1 A^{r-2} + \cdots + c_{r-1} I \end{aligned}$$

מתקיים

$$\begin{aligned} B_0 &= I \\ B_1 - AB_0 &= c_1 I \\ B_2 - AB_1 &= c_2 I \\ &\vdots \\ B_{r-1} - AB_{r-2} &= c_{r-1} I \end{aligned}$$

וכמו כן

$$-AB_{r-1} = -\left(A^r + c_1 A^{r-1} + \cdots + c_{r-1}A\right) + c_r I - c_r I = -m_A(A)\, I + c_r I = c_r I$$

נגדיר $B(x) = x^{r-1}B_0 + x^{r-2}B_1 + \cdots + xB_{r-2} + B_{r-1}$, ונקבל

$$\begin{aligned} (xI - A)\, B(x) &= xB(x) - AB(x) \\ &= \left(x^r B_0 + x^{r-1}B_1 + \cdots + x^2 B_{r-2} + xB_{r-1}\right) - \left(x^{r-1}AB_0 + x^{r-2}AB_1 + \cdots + xAB_{r-2} + AB_{r-1}\right) \\ &= x^r B_0 + x^{r-1}(B_1 - AB_0) + \cdots + x(B_{r-1} - AB_{r-2}) - AB_{r-1} \\ &= x^r I + x^{r-1}c_1 I + \cdots + xc_{r-1}I + c_r I = \left(x^r + c_1 x^{r-1} + \cdots + c_{r-1}x + c_r\right) I = m_A(x)\, I \end{aligned}$$

לכן,

$$\Delta_A(x)\det(B(x)) = \det(xI - A)\det(B(x)) = \det((xI - A)\, B(x)) = \det(m_A(x)\, I) = (m_A(x))^n$$

כלומר $\Delta_A(x) \,|\, (m_A(x))^n$ כדרוש. ■

מסקנה 2.23 לΔ_A, m_A יש אותם גורמים אי פריקים.

הוכחה: יהי $f(x)$ פולינום אי פריק.
נניח ש$f(x) \mid \Delta_A(x)$, ממשפט קודם נסיק ש$f(x) \mid (m_A(x))^n$. כיוון ש$f(x)$ אי פריק, אז $f(x) \mid m_a(x)$.
נניח ש$f(x) \mid m_A(x)$, אז ממשפט קיילי המילטון $f(x) \mid \Delta_A(x)$. ■

הגדרה 2.24 מטריצה ריבועית מהצורה $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ עבור A, B ריבועיות נקראת **מטריצת בלוקים**.

טענה 2.25 אם $\begin{pmatrix} A_2 & \\ & B_2 \end{pmatrix}, \begin{pmatrix} A_1 & \\ & B_1 \end{pmatrix}$ הן שתי מטריצות בלוקים כך שA_1, A_2 מאותו הסדר וB_1, B_2 מאותו הסדר, אז $\begin{pmatrix} A_1 & \\ & B_1 \end{pmatrix} \begin{pmatrix} A_2 & \\ & B_2 \end{pmatrix} = \begin{pmatrix} A_1 A_2 & \\ & B_1 B_2 \end{pmatrix}$. לכן, אם $f(x)$ פולינום ו$M = \begin{pmatrix} A & \\ & B \end{pmatrix}$ אז $f(M) = \begin{pmatrix} f(A) & \\ & f(B) \end{pmatrix}$. כמו כן, $\det(M) = \det(A) \cdot \det(B)$. באינדוקציה ניתן להכליל למטריצות עם מספר סופי של בלוקים.

הגדרה 2.26 פולינום $p(x)$ יקרא **הגורם המשותף הקטן ביותר** (least common multiplier) של $f(x), g(x)$ אם $f(x) \mid p(x), g(x) \mid p(x)$ ואם גם $q(x)$ מקיים $f(x) \mid q(x), g(x) \mid q(x)$ אז $p(x) \mid q(x)$. נסמן זאת $p(x) = [f(x), g(x)]$.

דוגמה 2.27 אם $f(x), g(x)$ אי פריקים ושונים אז $[f(x), g(x)] = f(x) \cdot g(x)$ ו־$(f(x), g(x)) = 1$.

טענה 2.28 נניח $M = \begin{pmatrix} A & \\ & B \end{pmatrix}$ אז $m_M(x) = [m_A(x), m_B(x)]$ (באינדוקציה עבור מספר בלוקים גדול יותר).

הוכחה: מתקיים $\begin{pmatrix} m_M(A) & \\ & m_M(B) \end{pmatrix} = m_M \begin{pmatrix} A & \\ & B \end{pmatrix} = m_M(M) = 0$ לכן $m_M(A) = 0, m_M(B) = 0$. אזי, מהגדרת הפולינום המינימלי $m_A(x) | m_M(x), m_B(x) | m_M(x)$. כעת, יהי $f(x)$ פולינום כך ש$m_A(x) | f(x)$ וגם $m_B(x) | f(x)$, מנתון נובע $f(A) = 0, f(B) = 0$, לכן $f(M) = f\begin{pmatrix} A & \\ & B \end{pmatrix} = \begin{pmatrix} f(A) & \\ & f(B) \end{pmatrix} = 0$ ולכן מהגדרת הפולינום המינימלי $m_M(x) | f(x)$. ■

טענה 2.29 אם $M = \begin{pmatrix} A & \\ & B \end{pmatrix}$ אז $\Delta_M(x) = \Delta_A(x) \Delta_B(x)$.

הוכחה: תרגיל. ■

3 שילוש

3.1 מרחבי מנה

הגדרה 3.1 יהי V מרחב וקטורי ותהי $T\colon V\to V$ העתקה לינארית. תת מרחב $W\subseteq V$ יקרא **T־אינווריאנטי** אם $T\left(W\right)\subseteq W$.

דוגמה 3.2 אם λ הוא ערך עצמי של T עם וקטור עצמי v, אז $W=\mathrm{Sp}\left\{ v\right\}$ הוא T־אינווריאנטי.

הערה 3.3 אם קיים W_{1} שהוא T־אינווריאנטי ומתקיים $V=W\oplus W_{1}$ כאשר גם W הוא T־אינווריאנטי, אז קל למצוא בסיס לV שבו המטריצה המייצגת את T היא מטריצת בלוקים.

הגדרה 3.4 יהי V מ"ו מעל שדה $\mathbb{F}$. יהי $W\subseteq V$ תת מרחב. בהינתן $v\in V$, נסמן $v+W=\left\{ v+w\mid w\in W\right\}$. אוסף כל הקבוצות הנ"ל (לכל $v\in V$) יסומן V/W.

טענה 3.5

1. לכל $v\in V$, מתקיים $v\in v+W$.
2. לכל $v,w\in V$, מתקיים $u+W=v+W\iff u-v\in W$.
3. לכל $v,w\in V$, הקבוצות $u+W,v+W$ הן זרות או זהות.

הוכחה: ■

1. כיוון ש$v=v+0$ ו$0\in W$, מתקיים $v\in v+W$.
2. $\Leftarrow$ נניח $u+W=v+W$, אז לכל $w_{1}\in W$ קיים $w_{2}\in W$ כך ש$u+w_{1}=v+w_{2}$, אז $u-v=w_{2}-w_{1}\in W$.
$\Longrightarrow$ נניח $u-v=w\in W$, לכן לכל $w_{1}\in W$ נקבל $u+w_{1}=v+w+w_{1}\in V+W$ לכן $u+W\subseteq v+W$.
בכיוון השני, אם $u-v\in W$ אז $v-u\in W$ ונקבל באופן דומה $v+W\subseteq u+W$.
3. תרגיל.

הגדרה 3.6 יהי V מרחב וקטורי, W תת מרחב של V. לכל $u,v\in V$ ולכל $\alpha\in\mathbb{F}$ נגדיר

$$\left(u+W\right)+\left(v+W\right)=\left(u+v\right)+W$$
$$\alpha\left(u+W\right)=\alpha u+W$$

משפט 3.7 הקבוצה V/W עם הפעולות הנ"ל מגדירה מרחב וקטורי מעל השדה $\mathbb{F}$.

הוכחה: תרגיל. נציין שוקטור האפס בV/W הוא $0+W$. נסמן קוסט זה בW. ■

טענה 3.8 נניח ש$W\subseteq V$ תת מרחב, אז $\dim V=\dim W+\dim V/W$.

הוכחה: יהי $\left\{ w_{1},\ldots,w_{r}\right\}$ בסיס של W, ויהי $\left\{ v_{1}+W,\ldots,v_{s}+W\right\}$ בסיס של V/W. נוכיח $\left\{ w_{1},\ldots,w_{r},v_{1},\ldots,v_{s}\right\}$ בסיס של V.

B **פורשת**: יהי $u\in V$, אז $u+W=\alpha_{1}\left(v_{1}+W\right)+\cdots+\alpha_{s}\left(v_{s}+W\right)=\left(\alpha_{1}v_{1}+\cdots+\alpha_{s}v_{s}\right)+W$ ומטענה נובע $u-\left(\alpha_{1}v_{1}+\cdots+\alpha_{s}v_{s}\right)\in W$. לכן $u-\left(\alpha_{1}v_{1}+\cdots+\alpha_{s}v_{s}\right)=\beta_{1}w_{1}+\cdots+\beta_{r}w_{r}$, ונקבל $u=\alpha_{1}v_{1}+\cdots+\alpha_{s}v_{s}+\beta_{1}w_{1}+\cdots+\beta_{r}w_{r}$ כדרוש.

B **בת"ל**: נניח $\alpha_{1}v_{1}+\cdots+\alpha_{s}v_{s}+\beta_{1}w_{1}+\cdots+\beta_{r}w_{r}=0$, אז נעביר אגפים ונקבל $\alpha_{1}v_{1}+\cdots+\alpha_{s}v_{s}\in W$. אז מטענה נובע $\left(\alpha_{1}v_{1}+\cdots+\alpha_{s}v_{s}\right)+W=0+W$, ולכן $\alpha_{1}\left(v_{1}+W\right)+\cdots+\alpha_{s}\left(v_{s}+W\right)=W$. אגף שמאל הוא צירוף לינארי המתאפס של איברי הבסיס של V/W, ולכן $\alpha_{1}=\cdots=\alpha_{s}=0$. מכאן $\beta_{1}w_{1}+\cdots+\beta_{r}w_{r}=0$, ולכן $\beta_{1}=\cdots=\beta_{r}=0$. ■

מסקנה 3.9 אם $W,W_{1}\subseteq V$ הם תתי מרחב כך ש$V=W\oplus W_{1}$ אז $V/W\simeq W_{1}$.

טענה 3.10 תהי $T\colon V\to V$ ויהי $W\subseteq V$ תת מרחב T־אינווריאנטי.
עבור כל $u+W$, נגדיר $\overline{T}\left(u+W\right)=T\left(u\right)+W$. אז $\overline{T}\colon V/W\to V/W$ היא העתקה לינארית, ותיקרא **ההעתקה המושרית** על ידי T למרחב המנה.

הוכחה: נבדוק תחילה $\overline{T}$ מוגדרת היטב, כלומר אם $u_1+W=u_2+W$ אז $\overline{T}\left(u_1+W\right)=\overline{T}\left(u_2+W\right)$.
נניח $u_1+W=u_2+W$ אז מטענה נובע $u_1-u_2\in W$, ולכן כיוון שW הוא T־אינווריאנטי ניתן להסיק $T\left(u_1\right)-T\left(u_2\right)=T\left(u_1-u_2\right)\in W$, ולכן נובע $\overline{T}\left(u_1+W\right)=T\left(u_1\right)+W=T\left(u_2\right)+W=\overline{T}\left(u_2+W\right)$. נותר להוכיח לינאריות ־ תרגיל. ■

טענה 3.11 בתנאי הטענה הקודמת, אם $q\left(x\right)\in\mathbb{F}\left[x\right]$ כך ש$q\left(T\right)=0$ אז $q\left(\overline{T}\right)=0$.

הוכחה: נוכיח באינדוקציה עבור $q\left(x\right)=x^k$, והמקרה הכללי ינבע מלינאריות.
בסיס: עבור $k=1$, אכן מתקיים אם $T=0$ אז $\overline{T}=0$. עבור $k=2$, נוכיח תחילה כי $\overline{T^2}=\left(\overline{T}\right)^2$. יהי $v+W$ וקטור כלשהו בV/W, אז

$$\overline{T^2}\left(v+W\right)=T^2\left(v\right)+W=T\left(T\left(v\right)\right)+W=\overline{T}\left(T\left(v\right)+W\right)=\overline{T}\left(\overline{T}\left(v+W\right)\right)=\left(\overline{T}\right)^2\left(v+W\right)$$

מכאן, אם $T^2=0$ אז $\left(\overline{T}\right)^2=\overline{T^2}=0$ לכן אם T הוא שורש של $q\left(x\right)=x^2$ אז גם $\overline{T}$ הוא שורש של $q\left(x\right)=x^2$.
באינדוקציה מוכיחים עבור $q\left(x\right)=x^k$ (תרגיל). ■

מסקנה 3.12 $m_{\overline{T}}\left(x\right)|m_T\left(x\right)$.

הוכחה: ע"פ הגדרה $m_T\left(T\right)=0$ אז $m_T\left(\overline{T}\right)=0$, ולכן מהגדרת הפולינום המינימלי $m_{\overline{T}}\left(x\right)\mid m_T\left(x\right)$. ■

3.2 שילוש

הגדרה 3.13 העתקה $T\colon V\to V$ היא ***ניתנת לשילוש*** אם קיים בסיס e של V בו $\left[T\right]_e$ היא מטריצה משולשית.
כלומר, קיים בסיס $e=\left\{v_1,\ldots,v_n\right\}$ כך ש

$$\begin{aligned}
Tv_1&=\alpha_{11}v_1\\
Tv_2&=\alpha_{12}v_1+\alpha_{22}v_2\\
&\vdots\\
Tv_i&=\alpha_{1i}v_1+\alpha_{2i}v_2+\cdots+\alpha_{ii}v_i\\
&\vdots
\end{aligned}$$

משפט 3.14 תהי T העתקה לינארית על מרחב וקטורי V מעל שדה $\mathbb{F}$. אם כל הערכים העצמיים של T הם ב$\mathbb{F}$, אז T ניתנת לשילוש מעל $\mathbb{F}$.

הוכחה: באינדוקציה על $\dim V=n$. אם $\dim V=1$, אז הטענה ברורה.
נניח עבור $n-1$, ונוכיח עבור n. כיוון שכל הערכים העצמיים של T הם ב$\mathbb{F}$ אז לT יש וקטור עצמי v_1 עם ערך עצמי $\lambda_1\in\mathbb{F}$. נגדיר $W=\left\{\alpha v_1\mid\alpha\in\mathbb{F}\right\}$. $\dim W=1$ וגם W הוא T־אינווריאנטי, נבנה את מרחב המנה V/W ומטענה קודמת $\dim V/W=n-1$. ההעתקה $\overline{T}$ מוגדרת היטב, ומתקיים $m_{\overline{T}}\left(x\right)|m_T\left(x\right)$ ולכן כל הערכים העצמיים של $\overline{T}$ הם ב$\mathbb{F}$. לפי הנחת האינדוקציה ניתן לשלש את $\overline{T}$, כלומר קיים בסיס $\left\{v_2+W,v_3+W,\ldots,v_n+W\right\}$ כך שמתקיים

$$\begin{aligned}
\overline{T}\left(v_2+W\right)&=\alpha_{22}\left(v_2+W\right)\\
\overline{T}\left(v_3+W\right)&=\alpha_{23}\left(v_2+W\right)+\alpha_{33}\left(v_3+W\right)\\
&\vdots\\
\overline{T}\left(v_n+W\right)&=\alpha_{2n}\left(v_2+W\right)+\alpha_{3n}\left(v_3+W\right)+\cdots+\alpha_{nn}\left(v_n+W\right)
\end{aligned}$$

כפי שראינו, $e=\left\{v_1,\ldots,v_n\right\}$ הוא בסיס של V (בהוכחה לטענה של מימד מרחב המנה). נוכיח $\left[T\right]_e$ משולשית.
מתקיים $Tv_1=\alpha_{11}v_1$, מתקיים $Tv_2+W=T\left(v_2+W\right)=\alpha_{22}v_2+W$ ולכן מטענה $Tv_2-\alpha_{22}v_2\in W$ ולכן קיים $\alpha_{12}\in\mathbb{F}$ כך ש$Tv_2-\alpha_{22}v_2=\alpha_{12}v_1$ ולכן $Tv_2=\alpha_{12}v_1+\alpha_{22}v_2$. וכך הלאה באינדוקציה. ■

אלגוריתם 3.15 בפועל ניתן לחשב בצורה הבאה:
נבחר וקטור עצמי v_1 של T, ונבנה את $W=\mathrm{Sp}\left(v_1\right)$, נבנה את V/W ואת $\overline{T}$. נבחר וקטור עצמי v_2+W של $\overline{T}$ ונגדיר $W_1=\mathrm{Sp}\left(v_1,v_2\right)$. אז W_1 הוא T אינווריאנטי.
נבנה את V/W_1 ועבורו נחשב את $\overline{T}$, וכך נמשיך את התהליך.

דוגמה 3.16 נשלש את המטריצה $A=\begin{pmatrix}-1&3&0\\-2&2&1\\-4&2&3\end{pmatrix}$. נחשב ונקבל $\Delta_A(x)=(x-1)^2(x-2)$, אז הערכים העצמיים הם $1,2$. נפתור

$$Av=2v\iff(2I-A)v=0\iff\begin{pmatrix}1&-3&0\\2&0&-1\\4&-2&-1\end{pmatrix}v=0\iff v\in\mathrm{Sp}\begin{pmatrix}1\\1\\2\end{pmatrix}$$

אז נבחר לדוגמה $v_1=\begin{pmatrix}1\\1\\2\end{pmatrix}$, אז $W=\mathrm{Sp}\begin{pmatrix}1\\1\\2\end{pmatrix}$. נקבל $\overline{A}\colon {}^{V}\!/_{W}\to{}^{V}\!/_{W}$, זו העתקה דו־מימדית והערך העצמי שלה הוא 1. נפתור

$$\overline{A}\left(\begin{pmatrix}x\\y\\z\end{pmatrix}+W\right)=\begin{pmatrix}x\\y\\z\end{pmatrix}+W\iff A\begin{pmatrix}x\\y\\z\end{pmatrix}+W=\begin{pmatrix}x\\y\\z\end{pmatrix}+W\iff(I-A)\begin{pmatrix}x\\y\\z\end{pmatrix}\in W$$
$$\iff\begin{pmatrix}2&-3&0\\2&-1&-1\\4&-2&-2\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\alpha\begin{pmatrix}1\\1\\2\end{pmatrix}$$

אז נבחר לדוגמה $v_2=\begin{pmatrix}0\\1\\2\end{pmatrix}$, אז $W_1=\mathrm{Sp}\left\{\begin{pmatrix}1\\1\\2\end{pmatrix},\begin{pmatrix}0\\1\\2\end{pmatrix}\right\}$. נקבל $\overline{A}\colon {}^{V}\!/_{W_1}\to{}^{V}\!/_{W_1}$, זו העתקה חד־מימדית והערך העצמי שלה הוא 1. נפתור

$$\overline{A}\left(\begin{pmatrix}x\\y\\z\end{pmatrix}+W_1\right)=A\begin{pmatrix}x\\y\\z\end{pmatrix}+W_1\iff\begin{pmatrix}x\\y\\z\end{pmatrix}+W_1=\begin{pmatrix}x\\y\\z\end{pmatrix}+W_1\iff(I-A)\begin{pmatrix}x\\y\\z\end{pmatrix}\in W_1$$
$$\iff\begin{pmatrix}2&-3&0\\2&-1&-1\\4&-2&-2\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\alpha\begin{pmatrix}1\\1\\2\end{pmatrix}+\beta\begin{pmatrix}0\\1\\2\end{pmatrix}$$

אז נבחר לדוגמה $v_3=\begin{pmatrix}0\\0\\1\end{pmatrix}$. נסמן $e=\{v_1,v_2,v_3\}$, נקבל $Av_1=2v_1,Av_2=3v_1+v_2,Av_3=0v_1+v_2+v_3$ ולכן

$$[A]_e=\begin{pmatrix}2&3&0\\0&1&1\\0&0&1\end{pmatrix}$$

הערה 3.17 אם T ניתנת לשילוש מעל $\mathbb{F}$, אז במטריצה המשולשית $[T]_e$ איברי האלכסון הם הערכים העצמיים. כל אחד מופיע כריבוי האלגברי שלו ב־$\Delta_T(x)$.

הגדרה 3.18 תהי $T\colon V\to V$ העתקה לינארית. נסמן $V_\lambda=\{v\in V: Tv=\lambda v\}$. V_λ נקרא **המרחב העצמי** של λ. V_λ הוא תת מרחב של V, ו־$\dim V_\lambda$ נקרא **הריבוי הגיאומטרי** של λ.

הגדרה 3.19 תהי $T\colon V\to V$ העתקה לינארית. לכל ערך עצמי λ, **הריבוי האלגברי** של λ הוא הריבוי האלגברי של λ ב־$\Delta_T(x)$.

טענה 3.20 לכל ערך עצמי λ, הריבוי האלגברי גדול או שווה מהריבוי הגיאומטרי.

הוכחה: נסמן $r=\dim V_\lambda$. יהי $\{v_1,\dots,v_r\}$ בסיס של V_λ, ונשלים אותו בעזרת $v_{r+1},\dots,v_n$ לבסיס e של V. אז

$$[T]_e=\begin{pmatrix}\lambda&0&\cdots&0&*&*&*\\0&\lambda&\cdots&0&\vdots&\vdots&\vdots\\0&0&\ddots&\vdots&*&*&*\\\vdots&\vdots&\cdots&\lambda&*&*&*\\0&0&\cdots&0&*&*&*\\\vdots&\vdots&\cdots&\vdots&*&*&*\\0&0&\cdots&0&*&*&*\end{pmatrix}$$

לכן

$$\Delta_T\left(x\right)=\det\left(xI-\left[T\right]_e\right)=\det\begin{pmatrix}x-\lambda & 0 & \cdots & 0 & * & * & *\\ 0 & x-\lambda & \cdots & 0 & \vdots & \vdots & \vdots\\ 0 & 0 & \ddots & \vdots & * & * & *\\ \vdots & \vdots & \cdots & x-\lambda & * & * & *\\ 0 & 0 & \cdots & 0 & * & * & *\\ \vdots & \vdots & \cdots & \vdots & * & * & *\\ 0 & 0 & \cdots & 0 & * & * & *\end{pmatrix}=\left(x-\lambda\right)^r f_T\left(x\right)$$

וקיבלנו שהריבוי האלגברי של λ הוא לפחות r, מה שרצינו להוכיח. ■

4 פירוקי ז'ורדן

4.1 העתקות נילפוטנטיות

טענה 4.1 תהי $T\colon V\to V$, ונניח ש$V=V_1\oplus\cdots\oplus V_m$ כאשר כל V_i הוא T אינווריאנטי. אז קיים בסיס של V שבו

$$[T]=\begin{pmatrix}A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_m\end{pmatrix}$$

כאשר A_i היא מטריצה מסדר $\dim V_i$, והיא המטריצה המייצגת את צמצום T לV_i.

הוכחה: לכל $1\le i\le m$ נבחר בסיס לV_i, נחבר את כל הבסיסים ונקבל ש$[T]$ היא המטריצה המבוקשת. ∎

הגדרה 4.2 לכל $t\geq 1$ נגדיר M_t מסדר t באופן הבא:

$$M_1=\begin{pmatrix}0\end{pmatrix}$$
$$M_t=\begin{pmatrix}0 & 1 & \cdots & 0\\ 0 & 0 & \ddots & \vdots\\ \vdots & \ddots & \ddots & 1\\ 0 & \cdots & 0 & 0\end{pmatrix}$$

הגדרה 4.3 העתקה $T\colon V\to V$ תקרא **נילפוטנטית** אם $T^r=0$ עבור r טבעי. המספר k יקרא **אינדקס הנילפוטנטיות** של T אם $T^k=0$ אבל $T^{k-1}\neq 0$.

טענה 4.4 תהי $T\colon V\to V$ העתקה נילפוטנטית עם אינדקס נילפוטנטיות k, אז $m_T\left(x\right)=x^k$ ו־$\Delta_T\left(x\right)=x^{\dim V}$, בפרט כל הערכים העצמיים של T הם 0 ולכן ממשפט T ניתנת לשילוש מעל כל שדה.

משפט 4.5 תהי T העתקה נילפוטנטית עם אינדקס נילפוטנטיות n_1. אז קיים בסיס e של V כך ש

$$[T]_e=\begin{pmatrix}M_{n_1} & & & \\ & M_{n_2} & & \\ & & \ddots & \\ & & & M_{n_r}\end{pmatrix}$$

כאשר $n_1\geq n_2\geq\cdots\geq n_r$, $\sum_{i=1}^{r}n_i=\dim V$, ו$r=\dim V_0$.

הוכחה: כיוון ש$T^{n_1}=0$ אבל $T^{n_1-1}\neq 0$ קיים $v\in V$ כך ש$T^{n_1-1}v\neq 0$.
נוכיח שהקבוצה $\left\{v,Tv,\ldots,T^{n_1-1}v\right\}$ היא בלתי תלויה לינארית. נניח $\alpha_1 v+\alpha_2 Tv+\cdots+\alpha_{n_1}T^{n_1-1}v=0$ ונניח בשלילה שקיים i כך ש$\alpha_i\neq 0$, יהי s המינימלי כך ש$\alpha_s\neq 0$. אז נקבל

$$\left(\alpha_s I+\alpha_{s+1}T+\cdots+\alpha_{n_1}T^{n_1-s}\right)T^{s-1}v=0$$

מתרגיל בית ההעתקה $\alpha_s I+\alpha_{s+1}T+\cdots+\alpha_{n_1}T^{n_1-s}$ היא הפיכה. במקרה זה נקבל $T^{s-1}v=0$, בסתירה לכך ש$T^{n_1-1}v\neq 0$. $\left\{T^{n_1-1}v,\ldots,Tv,v\right\}$ היא בלתי תלויה, נשלים אותה לבסיס e של V. אז

$$T\left(T^{n_1-1}v\right)=0$$
$$T\left(T^{n_1-2}v\right)=T^{n_1-1}v$$
$$\vdots$$
$$Tv=Tv$$

ולכן

$$[T]_e = \begin{pmatrix} 0 & 1 & \cdots & 0 & * & \cdots & * \\ 0 & 0 & \ddots & \vdots & * & \cdots & * \\ 0 & 0 & \ddots & 1 & * & \cdots & * \\ 0 & 0 & \cdots & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & * & \cdots & * \end{pmatrix} = \begin{pmatrix} M_{n_1} & B \\ 0 & C \end{pmatrix}$$

כעת, נראה שקיים בסיס שבו $[T] = \begin{pmatrix} M_{n_1} & 0 \\ 0 & C \end{pmatrix}$. נניח שהראינו זאת, אז נסיים את הוכחת המשפט כי כיוון ש־T נילפוטנטית $[T]^{n_1} = 0$ ולכן $C^{n_1} = 0$ לכן C היא נילפוטנטית עם אינדקס נילפוטנטיות $n_2 \leq n_1$. באינדוקציה נובע שקיים בסיס e כך של־$[T]_e$ יש את הצורה המבוקשת. כמו כן למטריצה M_t יש ערך עצמי שהוא אפס ועבור מטריצה זו יש וקטור עצמי בלתי תלוי אחד ויחיד, ולכן $\dim V_0 = r$ (תרגיל).

נראה שקיים בסיס שבו $[T] = \begin{pmatrix} M_{n_1} & 0 \\ 0 & C \end{pmatrix}$. נסמן $p = n_1$. כדי להראות זאת, מספיק למצוא מטריצה הפיכה D כך ש־$D^{-1}\begin{pmatrix} M_p & B \\ 0 & C \end{pmatrix} D = \begin{pmatrix} M_p & 0 \\ 0 & C \end{pmatrix}$. נחפש D מהצורה $D = \begin{pmatrix} I_p & X \\ 0 & I_{n-p} \end{pmatrix}$, עבור $X \in Mat_{p\times(n-p)}(\mathbb{F})$. קל לראות $D^{-1} = \begin{pmatrix} I_p & -X \\ 0 & I_{n-p} \end{pmatrix}$. צריך להראות שלמשוואה

$$\begin{pmatrix} M_p & 0 \\ 0 & C \end{pmatrix} = \begin{pmatrix} I_p & X \\ 0 & I_{n-p} \end{pmatrix}\begin{pmatrix} M_p & B \\ 0 & C \end{pmatrix}\begin{pmatrix} I_p & -X \\ 0 & I_{n-p} \end{pmatrix}$$

יש פתרון בנעלם X. אם נבצע את הכפל באגף ימין נקבל

$$\begin{pmatrix} M_p & \\ & C \end{pmatrix} = \begin{pmatrix} M_p & -M_pX + B + XC \\ & C \end{pmatrix}$$

נראה שלמשוואה $-M_pX + B + XC$ יש פתרון.

נסמן $X = \begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix}$ ו $B = \begin{pmatrix} B_1 \\ \vdots \\ B_p \end{pmatrix}$ כאשר X_i הן שורות X ו־B_i הן שורות B. בסימונים אלו

$$-M_pX + B + XC = \begin{pmatrix} -X_2 + B_1 + X_1C \\ -X_3 + B_2 + X_2C \\ \vdots \\ -X_p + B_{p-1} + X_{p-1}C \\ B_p + X_pC \end{pmatrix}$$

נבחר $X_1 = 0, X_2 = B_1 + X_1C, \ldots, X_p = B_{p-1} + X_{p-1}C$. בבחירה זו נקבל

$$\begin{pmatrix} I_p & X \\ 0 & I_{n-p} \end{pmatrix}\begin{pmatrix} M_p & B \\ 0 & C \end{pmatrix}\begin{pmatrix} I_p & -X \\ 0 & I_{n-p} \end{pmatrix} = \begin{pmatrix} M_p & L' \\ & C \end{pmatrix}$$

כאשר L' היא $\begin{pmatrix} 0 \\ \vdots \\ 0 \\ L \end{pmatrix}$. מנתון $\begin{pmatrix} M_p & B \\ 0 & C \end{pmatrix}^p = 0$ ולכן $\begin{pmatrix} M_p & L' \\ & C \end{pmatrix}^p = 0$, מצד שני על ידי כפל מטריצות נקבל $\begin{pmatrix} M_p & L' \\ & C \end{pmatrix}^p = \begin{pmatrix} 0 & L'' \\ & 0 \end{pmatrix}$ כאשר ב־L'' השורה L היא השורה הראשונה. לכן בהכרח $L = 0$, אז גם $L' = 0$ וסיימנו.
■

הגדרה 4.6 המספרים $n_1, \ldots, n_r$ נקראים ה**אינווריאנטים** של T.

הגדרה 4.7 תהי T נילפוטנטית. תת מרחב $M \subseteq V$ T־אינווריאנטי ממימד n נקרא **ציקלי** ביחס ל־T אם קיים $z \in M$ כך ש־$z, Tz, \ldots, T^{m-1}z$ הוא בסיס ל־M. אם M הוא תת מרחב ציקלי, אז $T^m(M) = \{0\}$ וגם $T^{m-1}(M) \neq \{0\}$.

טענה 4.8 אם $\dim M=m$ ו־M תת מרחב ציקלי ביחס ל־T, אז עבור $k\leq m$ מתקיים $\dim\left(T^{k}\left(M\right)\right)=m-k$.

הוכחה: בסיס לתמונה של העתקה מוכל בתמונת איברי הבסיס. נבחר כבסיס של M את הקבוצה $z,Tz,\ldots,T^{m-1}z$, אז $T^{k}M$ נפרש על ידי $T^{k}z,T^{k+1}z,\ldots,T^{k+m-1}z$. כיוון ש־$T^{m}z=0$ אז $T^{k}z,T^{k+1}z,\ldots,T^{m-1}z$ הוא בסיס ל־$T^{k}M$. יש בבסיס זה $m-k$ וקטורים, ולכן $\dim\left(T^{k}\left(M\right)\right)=m-k$. ■

משפט 4.9 שתי העתקות (מטריצות) נילפוטנטיות הן דומות אמ"ם יש להן אותה קבוצת אינווריאנטים.

הוכחה: $\Longleftarrow$ נוכיח את יחידות הפירוק. ממשפט קודם ניתן לרשום $V=V_{1}\oplus\cdots\oplus V_{r}$ כאשר $\dim V_{i}=n_{i}$ ו־$n_{1}\geq n_{2}\geq\cdots\geq n_{r}$ כאשר כל אחד מה־V_{i} הוא ציקלי. אם גם $V=U_{1}\oplus\cdots\oplus U_{s}$ כאשר $\dim U_{i}=m_{i}$ ו־$m_{1}\geq m_{2}\geq\cdots\geq m_{s}$ כאשר U_{i} ציקליים אז נראה שבהכרח $r=s$ ו־$n_{i}=m_{i}$.

נניח בשלילה שלא, יהי i האינדקס הראשון כך ש־$m_{i}\neq n_{i}$ ונניח ש־$m_{i}<n_{i}$. נחשב את $\dim T^{m_{i}}V$ באופנים שונים. מצד אחד, מתקיים $T^{m_{i}}V=T^{m_{i}}V_{1}\oplus\cdots\oplus T^{m_{i}}V_{r}$. מהטענה הקודמת

$$\begin{aligned}\dim T^{m_{i}}V_{1}&=n_{1}-m_{i}\\ \dim T^{m_{i}}V_{2}&=n_{2}-m_{i}\\ &\vdots\\ \dim T^{m_{i}}V_{i}&=n_{i}-m_{i}\end{aligned}$$

מכאן נובע $\dim T^{m_{i}}V\geq\sum_{j=1}^{i}\dim T^{m_{i}}V_{j}=\sum_{j=1}^{i}\left(n_{j}-m_{i}\right)$.

מצד שני, הפירוק $V=U_{1}\oplus\cdots\oplus U_{s}$ ביחד עם העובדה $T^{m_{i}}U_{j}=0$ (כי $\dim T^{m_{i}}U_{j}=m_{j}-m_{i}\leq0$) לכל $j\geq i$ ייתן $T^{m_{i}}V=T^{m_{i}}U_{1}\oplus\ldots T^{m_{i}}U_{i-1}$. אז $\dim T^{m_{i}}V=\sum_{j=1}^{i-1}\dim T^{m_{i}}U_{j}=\sum_{j=1}^{i-1}\left(m_{j}-m_{i}\right)$. i הוא הראשון כך ש־$m_{i}\neq n_{i}$ לכן לכל $j<i$ מתקיים $m_{j}=n_{j}$, אז $\sum_{j=1}^{i-1}\left(m_{j}-m_{i}\right)=\sum_{j=1}^{i-1}\left(n_{j}-m_{i}\right)$.

אבל $\sum_{j=1}^{i-1}\left(n_{j}-m_{i}\right)<\sum_{j=1}^{i}\left(n_{j}-m_{i}\right)$, וזו סתירה כי שניהם שווים ל־$\dim T^{m_{i}}V$.

$\Longrightarrow$ נניח של־T,S יש אותה קבוצת אינווריאנטים, אז קיימים בסיסים $e=\left\{v_{1},\ldots,v_{n}\right\},f=\left\{w_{1},\ldots,w_{n}\right\}$ כך ש

$$\left[T\right]_{e}=\begin{pmatrix}M_{n_{1}}\\ & \ddots\\ & & M_{n_{r}}\end{pmatrix}=\left[S\right]_{f}$$

אז עבור מטריצת המעבר $\left[Id\right]_{f}^{e}$ יתקיים $\left[T\right]_{f}=\left[Id\right]_{f}^{e}\left[T\right]_{e}\left[Id\right]_{e}^{f}=\left[Id\right]_{f}^{e}\left[S\right]_{f}\left[Id\right]_{e}^{f}$, ולכן $\left[T\right]_{f},\left[S\right]_{f}$ דומות. ■

מסקנה 4.10 נסמן ב־$p\left(n\right)$ את מספר החלוקות של המספר n. אז לפי המשפט יש בדיוק $p\left(n\right)$ מטריצות נילפוטנטיות מסדר n עד כדי דמיון.

4.2 משפט הפירוק היסודי

טענה 4.11 תהי $T\colon V\to V$ לינארית ויהי $f\left(x\right)$ פולינום כלשהו, אז $\ker f\left(T\right)$ הוא תת מרחב T אינווריאנטי.

הוכחה: אם $v\in\ker f\left(T\right)$, אז נחשב את $f\left(T\right)Tv=Tf\left(T\right)v=0$ ולכן $Tv\in\ker f\left(T\right)$. ■

טענה 4.12 כאשר $f\left(x\right)=x-\lambda$, $\ker f\left(T\right)=\left\{v\in V\mid Tv=\lambda v\right\}$.

משפט 4.13 הפירוק היסודי: תהי $T\colon V\to V$ ונניח ש־$m_{T}\left(x\right)=q_{1}\left(x\right)^{l_{1}}q_{2}\left(x\right)^{l_{2}}\cdots q_{k}\left(x\right)^{l_{k}}$ כאשר $q_{i}\left(x\right)$ הם פולינומים אי פריקים. לכל $1\leq i\leq k$ נגדיר $V_{i}=\ker q_{i}\left(T\right)^{l_{i}}$. מהחישוב הנ"ל נקבל שכל V_{i} הוא תת מרחב T אינווריאנטי. נסמן $T_{i}\colon V_{i}\to V_{i}$ את צמצום T ל־V_{i}.

אזי לכל $1\leq i\leq k$, $V_{i}\neq\left\{0\right\}$ וגם $V=V_{1}\oplus\cdots\oplus V_{k}$. כמו כן, הפולינום המינימלי של T_{i} הוא $q_{i}\left(x\right)^{l_{i}}$.

טענה 4.14 תהי $T\colon V\to V$. נניח כי $f\left(x\right)=g\left(x\right)h\left(x\right)$ כאשר $f\left(T\right)=0$, וגם $\left(h,g\right)=1$. נסמן $U=\ker g\left(T\right),W=\ker h\left(T\right)$, אז $V=U\oplus W$.

הוכחה: נוכיח את הלמה. הוכחנו ש־U,W הם T אינווריאנטים, כיוון ש־$g\left(x\right),h\left(x\right)$ זרים אז קיימים פולינומים $r\left(x\right),s\left(x\right)$ כך ש־$r\left(x\right)g\left(x\right)+s\left(x\right)h\left(x\right)=1$.

נציב את T במשוואה זו ונקבל $r\left(T\right)g\left(T\right)+s\left(T\right)h\left(T\right)=I$, ולכן לכל $v\in V$ מתקיים $v=r\left(T\right)g\left(T\right)v+s\left(T\right)h\left(T\right)v$.

נשים לב ש$r(T)g(T)v\in\ker h(T), s(T)h(T)v\in\ker g(T)$, זאת כי

$$h(T)r(T)g(T)v=r(T)g(T)h(T)v=r(T)f(T)v=0$$
$$g(T)s(T)h(T)v=s(T)g(T)h(T)v=s(T)f(T)v=0$$

לכן $V=\ker g(T)+\ker h(T)=U+W$.

כעת נוכיח שהסכום הוא ישר, כלומר כל $v\in V$ ניתן להצגה יחידה כסכום $v=u+w$ כאשר $u\in U, w\in W$.

מהשוויון $v=u+w$ נקבל

$$r(T)g(T)v=r(T)g(T)u+r(T)g(T)w\overset{u\in\ker g(T)}{=}r(T)g(T)w$$

כמו כן

$$w=r(T)g(T)w+s(T)h(T)w\overset{w\in\ker h(T)}{=}r(T)g(T)w$$

ולכן $w=r(T)g(T)v$. לכן w נקבע באופן יחיד על ידי v, ובאותו האופן לגבי u. ■

טענה 4.15 בתנאי טענה 1 נניח ש$f(x)=m_T(x)$, נסמן בT_1 את צמצום T לU ובT_2 את צמצום T לW. נניח שהמקדם הגבוה של $q(x)$ ושל $h(x)$ הוא 1, אז $h(x)=m_{T_2}(x), g(x)=m_{T_1}(x)$. במילים אחרות, $m_T(x)=m_{T_1}(x)m_{T_2}(x)$.

הוכחה: כיוון ש$U=\ker g(T)$ אז $g(T_1)=0$. כמו כן $h(T_2)=0$. לכן $g(x)\mid m_{T_1}(x)$, $h(x)\mid m_{T_2}(x)$.

נזכיר כי אם $[T]_e=\begin{pmatrix}A&\\&B\end{pmatrix}$ תחת הפירוק $V=U\oplus W$, אז $m_T(x)=\operatorname{lcm}(m_{T_1}(x),m_{T_2}(x))$. כיוון ש$g(x),h(x)$ זרים אז $m_{T_1}(x),m_{T_2}(x)$ זרים, ולכן $m_T(x)=\operatorname{lcm}(m_{T_1}(x),m_{T_2}(x))=m_{T_1}(x)m_{T_2}(x)$. ■

דוגמה 4.16 $A=\begin{pmatrix}-1&3&0\\-2&2&1\\-4&2&3\end{pmatrix}$, אז $\Delta_A(x)=(x-1)^2(x-2)$. מתקיים $m_A(x)=(x-1)^m(x-2)$ עבור $m=1$ או $m=2$, כיוון ש$(A-I)(A-2I)\neq0$ אז בהכרח $m=2$, כלומר $m_A(x)=\Delta_A(x)$.

אז בתנאי משפט הפירוק היסודי $q_1(x)=x-1, l_1=2$ ו$q_2(x)=x-2, l_2=1$. נחשב

$$\ker q_1(x)^2=\ker(A-I)^2=\ker\begin{pmatrix}-2&-3&3\\-2&-3&2\\-4&-6&6\end{pmatrix}=\left\{\begin{pmatrix}-2&-3&3\\-2&-3&3\\-4&-6&6\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=0\right\}=\operatorname{Sp}\left\{\begin{pmatrix}3\\-2\\0\end{pmatrix},\begin{pmatrix}3\\0\\2\end{pmatrix}\right\}$$

$$\ker q_2(x)=\ker A-2I=\ker\begin{pmatrix}-3&3&0\\-2&0&1\\-4&2&1\end{pmatrix}=\left\{\begin{pmatrix}-3&3&0\\-2&0&1\\-4&2&1\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=0\right\}=\operatorname{Sp}\left\{\begin{pmatrix}1\\1\\2\end{pmatrix}\right\}$$

נסמן

$$v_1=\begin{pmatrix}1\\1\\2\end{pmatrix}, v_2=\begin{pmatrix}3\\-2\\0\end{pmatrix}, v_3=\begin{pmatrix}3\\0\\2\end{pmatrix}$$

אז

$$Av_1=\begin{pmatrix}2\\2\\4\end{pmatrix}=2\begin{pmatrix}1\\1\\2\end{pmatrix}$$
$$Av_2=\begin{pmatrix}-9\\-10\\-16\end{pmatrix}=5\begin{pmatrix}3\\-2\\0\end{pmatrix}+(-8)\begin{pmatrix}3\\0\\2\end{pmatrix}$$
$$Av_3=\begin{pmatrix}-3\\-4\\-6\end{pmatrix}=2\begin{pmatrix}3\\-2\\0\end{pmatrix}+(-3)\begin{pmatrix}3\\0\\2\end{pmatrix}$$

ולכן

$$[A]_e=\begin{pmatrix}2&&\\&5&2\\&-8&-3\end{pmatrix}$$

הוכחה: כעת נוכיח את משפט הפירוק היסודי, באינדוקציה על k.
עבור $k=1$ הוכחנו. נניח עבור $k-1$, נסמן $g(x)=q_1(x)^{l_1}, h(x)=q_2(x)^{l_2}\cdots q_k(x)^{l_k}$. מתקיים $(g,h)=1$ ואם $f(x)=g(x)h(x)$ אז $f(T)=m_T(T)=0$. אז מטענה 1 ניתן לרשום $V=V_1\oplus W_1$ כאשר $V_1=\ker q_1(T)^{l_1}, W_1=\ker h(T)$. לפי טענה 2 הפולינום המינימלי של T_1 (צמצום T לV_1) הינו $q_1(x)^{l_1}$ והפולינום המינימלי של T_2 (צמצום T לW_1) הוא $q_2(x)^{l_2}\cdots q_k(x)^{l_k}$, ואז לפי הנחת האינדוקציה ניתן לרשום $W_1=V_2\oplus\cdots\oplus V_k$ כאשר הפולינום המינימלי של צמצום T לV_i הוא $q_i(x)^{l_i}$ וסיימנו. ∎

משפט 4.17 העתקה $T\colon V\to V$ ניתנת ללכסון $\iff$ $m_T(x)$ הוא מכפלה של גורמים לינאריים שונים.

הוכחה: $\implies$ נניח ש$m_T(x)=(x-\lambda_1)\cdots(x-\lambda_r)$ כאשר $\lambda_i\neq\lambda_j$. אז לפי משפט הפירוק היסודי ניתן לרשום $V=V_1\oplus\cdots\oplus V_r$, כאשר V_i הוא $\ker(T-\lambda_i I)$. לכן אם $v\in V_i$ אז $(T-\lambda_i I)v=0$, ולכן $Tv=\lambda_i v$. לכן כל וקטור $0\neq v\in V_i$ הוא וקטור עצמי עבור הערך עצמי λ_i. מהשוויון $V=V_1\oplus\cdots\oplus V_r$ נובע שניתן למצוא בסיס לV המורכב מוקטורים עצמיים, ולפי בסיס זה T אלכסונית.
$\Longleftarrow$ נניח שT ניתנת ללכסון, ולכן לV יש בסיס של וקטורים עצמיים $v_1,\ldots,v_n$. יהי $\lambda_1,\ldots,\lambda_s$ הערכים העצמיים השונים של T. נגדיר $f(x)=(x-\lambda_1)\cdots(x-\lambda_s)$, אז

$$f(T)v_i=(T-\lambda_1 I)\cdots(T-\lambda_S I)v_i=\prod_{j\neq i}^{s}(T-\lambda_j I)(T-\lambda_i I)v_i=0$$

לכן $f(T)v=0$ לכל $v\in V$, לכן $m_T(x)\mid f(x)$ ובפרט $m_T(x)$ הוא מכפלה של גורמים לינאריים שונים. ∎

משפט 4.18 פירוק ז'ורדן: תהי $T\colon V\to V$ כך שכל הערכים העצמיים של T הם בשדה (זה תמיד קורה אם השדה הוא $\mathbb{C}$). נניח כי

$$\Delta_T(x)=(x-\lambda_1)^{n_1}\cdots(x-\lambda_r)^{n_r}$$
$$m_T(x)=(x-\lambda_1)^{m_1}\cdots(x-\lambda_r)^{m_r}$$

כאשר $\lambda_i\neq\lambda_j$. אז לV יש בסיס שבו T ניתנת לייצוג על ידי מטריצה J שאיברי האלכסון שלה הם מטריצות מהצורה

$$J_{i_j}=\begin{pmatrix}\lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1\\ & & & \lambda_i\end{pmatrix}$$

כאשר $1\leq i\leq r$ ו$1\leq j\leq\dim V_i$ מספר המטריצות המתאימות לערך עצמי λ_i.
לכל λ_i המטריצות J_{i_j} מקיימות את התנאים הבאים:

1. קיים לפחות J_{i_j} אחד מסדר m_i. כל שאר הJ_{i_j} הם מסדר שאינו עולה על m_i.
2. סכום כל הסדרים של J_{i_j} הוא n_i.
3. מספר וגודל כל הJ_{i_j} בכל סדר אפשרי נקבע באופן יחיד על ידי T.

הערה 4.19 אם סדר המטריצה J_{i_j} הוא k, אז $J_{i_j}=\lambda_i I_k+M_k$.

דוגמה 4.20

1. אם $\Delta_T(x)=(x-2)^4(x-3)^3$ ו$m_T(x)=(x-2)^2(x-3)^2$ אז $\dim V=7$. נסמן בA את מטריצת ז'ורדן המתאימה לT.
   על האלכסון יהיו 4 פעמים 2 ו3 פעמים 3. אז

$$A=\begin{pmatrix}2&1&&&&&\\0&2&&&&&\\&&3&1&&&\\&&0&3&&&\\&&&&3&&\\&&&&&2&1\\&&&&&0&2\end{pmatrix}\text{ or }A=\begin{pmatrix}2&1&&&&&\\0&2&&&&&\\&&3&1&&&\\&&0&3&&&\\&&&&3&&\\&&&&&2&\\&&&&&&2\end{pmatrix}$$

2. אם $\Delta_A(x) = (x-2)^3 (x-5)^2$, מהן צורות הז'ורדן האפשריות? סדר A הוא 5 ועל האלכסון יש 3 פעמים 2 ופעמיים 5.

$$m_A(x) = (x-2)^{m_1} (x-5)^{m_2}$$

כאשר $1 \leq m_i \leq 3, 1 \leq m_2 \leq 2$.

הוכחה: נוכיח את משפט פירוק ז'ורדן.
כיוון ש$m_T(x) = (x-\lambda_1)^{m_1} \cdots (x-\lambda_r)^{m_r}$, ממשפט הפירוק היסודי נקבל שניתן לרשום $V = V_1 \oplus \cdots \oplus V_r$ כאשר $V_i = \ker(T-\lambda_i I)^{m_i}$. כמו כן אם T_i הוא צמצום T לV_i, אז $m_{T_i}(x) = (x-\lambda_i)^{m_i}$. נגדיר $N_i = T_i - \lambda_i I$, ואז N_i נילפוטנטית על V_i עם אינדקס נילפוטנטיות m_i. אז ממשפט על העתקות נילפטונטיות קיים בסיס e_i כך ש

$$[T_i - \lambda_i I]_{e_i} = \begin{pmatrix} M_{m_{i_1}} & & & \\ & M_{m_{i_2}} & & \\ & & \ddots & \\ & & & M_{m_{i_l}} \end{pmatrix}$$

כאשר l הוא מימד מרחב העצמי של הערך העצמי אפס בהעתקה N_i, $m_{i_1} = m_1$ וגם $m_{i_1} \geq m_{i_2} \geq \cdots \geq m_{i_l}$, וכמו כן $n_i = m_{i_1} + m_{i_2} + \cdots + m_{i_l}$.
לכן

$$[T_i]_{e_i} = \lambda_i I + \begin{pmatrix} M_{m_{i_1}} & & & \\ & M_{m_{i_2}} & & \\ & & \ddots & \\ & & & M_{m_{i_l}} \end{pmatrix} = \begin{pmatrix} J_{i,1} & & & \\ & J_{i,2} & & \\ & & \ddots & \\ & & & J_{i,l} \end{pmatrix}$$

נובע גם $l = \dim V_i$. ■

5 צורה קנונית רציונלית

נתונה העתקה $T\colon V\to V$, ויהי $0\neq v\in V$. נבנה את קבוצת הוקטורים $v,Tv,T^2v,\ldots$, כיוון שהמימד של v הוא סופי קיים k מינימלי כך ש

$$T^kv=-a_{k-1}T^{k-1}v-\cdots-a_1Tv-a_0v$$

במקרה זה הקבוצה $\left\{v,Tv,\ldots,T^{k-1}v\right\}$ היא בלתי תלויה לינארית. נסמן $Z\left(v,T\right)=\operatorname{Sp}\left\{v,Tv,\ldots,T^{k-1}v\right\}$.

הגדרה 5.1 תת המרחב $Z\left(v,T\right)$ נקרא **תת מרחב** T **ציקלי הנפרש על ידי** v.

טענה 5.2 מתקיימות התכונות הבאות:

1. $Z\left(v,T\right)$ הינו T אינווריאנטי.
2. $\dim Z\left(v,T\right)=k$.
3. נגדיר $m_v\left(x\right)=x^k+a_{k-1}x^{k-1}+\cdots+a_1x+a_0$ אז $m_v\left(x\right)$ הוא הפולינום עם הדרגה הנמוכה ביותר המקיים $\left(m_v\left(T\right)\right)v=0$.

   אם נסמן בT_v את צמצום T ל$Z\left(v,T\right)$, אז $m_v\left(x\right)=m_{T_v}\left(x\right)$.
4. המטריצה המייצגת את T_v לפי הבסיס $e=\left\{v,Tv,\ldots,T^{k-1}v\right\}$ היא

$$\left[T_v\right]_e=\begin{pmatrix}0 & & & & -a_0\\ 1 & 0 & & & -a_1\\ & 1 & \ddots & & -a_2\\ & & \ddots & 0 & \vdots\\ & & & 1 & -a_{k-1}\end{pmatrix}$$

הגדרה 5.3 לכל פולינום מהצורה $f\left(x\right)=x^k+a_{k-1}x^{k-1}+\cdots+a_1x+a_0$ נתאים את המטריצה

$$C\left(f\left(x\right)\right)=\begin{pmatrix}0 & & & & -a_0\\ 1 & 0 & & & -a_1\\ & 1 & \ddots & & -a_2\\ & & \ddots & 0 & \vdots\\ & & & 1 & -a_{k-1}\end{pmatrix}$$

$C\left(f\left(x\right)\right)$ נקראת **המטריצה הנלווית** לפולינום $f\left(x\right)$.

משפט 5.4 תהי $T\colon V\to V$ עם פולינום מינימלי $m_T\left(x\right)=q\left(x\right)^n$ כאשר $q\left(x\right)$ אי פריק עם דרגה d. אז קיימים $v_1,\ldots,v_r\in V$ כך ש$V=Z\left(v_1,T\right)\oplus\cdots\oplus Z\left(v_r,T\right)$. אם T_i הוא צמצום T ל$Z\left(v_i,T\right)$, אז $m_{T_i}\left(x\right)=q\left(x\right)^{n_i}$ כאשר $n=n_1\geq n_2\geq\cdots\geq n_r$. יתר על כן, המספרים n_i תלויים בT בלבד ולא בבחירת v_i.

הערה 5.5 למעשה, כתוצאה ממשפט הפירוק היסודי ניתן להכליל את המשפט הנ"ל לכל העתקה.

מסקנה 5.6 קיים בסיס e של V כך ש

$$\left[T\right]_e=\begin{pmatrix}C\left(q\left(x\right)^{n_1}\right) & & & \\ & C\left(q\left(x\right)^{n_2}\right) & & \\ & & \ddots & \\ & & & C\left(q\left(x\right)^{n_r}\right)\end{pmatrix}$$

וכמו כן $\dim Z\left(v_i,T\right)=dn_i$, ולכן $\dim V=d\sum_{i=1}^{r}n_i$.

הוכחה: נוכיח את המשפט. יהי $v_1 \in V$ כך ש־$q(T)^{n_1-1} v_1 \neq 0$ ($n_1 = n$). נסמן

$$Z(v_1, T) = \mathrm{Sp}\left\{v_1, Tv_1, \ldots, T^{dn-1} v_1\right\}$$

אז $Z(v_1, T)$ הוא T־אינווריאנטי ונסמן ב־T_1 את צמצום T ל־$Z(v_1, T)$. אז $m_{T_1}(x) \mid m_T(x)$. כיוון ש־$q(x)$ אי פריק, נובע $m_{T_1}(x) = q(x)^r$, כאשר $r \le n_1$. כיוון ש־$q(T)^{n_1-1} v_1 \neq 0$ בהכרח $r = n_1$. לכן $m_{T_1}(x) = m_T(x)$, ולכן $\dim Z(v_1, T) = dn_1$.

נסמן $Z_1 = Z(v_1, T)$. נוכיח באינדוקציה על $\dim V$. נניח ש־$\dim V > 1$ ומתקיים $\dim {}^{V}/_{Z_1} < \dim V$. נסמן ב־$\overline{T}$ את ההעתקה המושרית של T על ${}^{V}/_{Z_1}$, כיוון ש־$m_{\overline{T}}(x) \mid m_T(x)$, וכיוון ש־$m_T(x) = q(x)^n$ ניתן להסיק $m_{\overline{T}}(x) = q(x)^{n_2}$ כאשר $n = n_1 \ge n_2$. מהנחת האינדוקציה נסיק

$${}^{V}/_{Z_1} = Z\left(v_2 + Z_1, \overline{T}\right) \oplus \cdots \oplus Z\left(v_r + Z_1, \overline{T}\right)$$

כאשר $n_2 \ge \cdots \ge n_r$ והפולינום המינימלי של $\overline{T_i}$ הינו $m_{\overline{T_i}}(x) = q(x)^{n_i}$. כמו כן, $\dim Z\left(v_i + Z_1, \overline{T}\right) = dn_i$.

כעת, נותר להוכיח שקיים $w_i \in v_i + Z_1$ כך ש־$m_{\overline{T_i}}(T)(w_i) = q(T)^{n_i}(w_i) = 0$ אבל $f(T) w_i \neq 0$ לכל פולינום שדרגתו נמוכה מדרגת $m_{\overline{T_i}}(x) = q(x)^{n_i}$ (כי במקרה זה קבוצת הוקטורים $\left\{w_i, Tw_i, \ldots, T^{dn_i-1} w_i\right\}$ תהיה בלתי תלויה ואז נגדיר $Z(w_i, T) = \mathrm{Sp}\left\{w_i, \ldots, T^{dn_i-1} w_i\right\}$).

יהי $u_i \in v_i + Z_1$. אז

$$q(T)^{n_i} u_i = q(T)^{n_i} v_i + Z_1 = \overline{q(T)^{n_i}}(v_i + Z_1) = q\left(\overline{T}\right)^{n_i}(v_i + Z_1) = 0 + Z_1$$

כאשר השוויון האחרון נובע מכך שהפולינום המינימלי של $\overline{T_i}$ הוא $q(x)^{n_i}$. בסך הכל נובע $q(T)^{n_i} u_i \in Z_1$. לכן,

$$q(T)^{n_i} u_i = \alpha_0 v_1 + \alpha_1 T v_1 + \cdots + \alpha_{dn_1-1} T^{dn_1-1} v_1$$

לכן קיים פולינום $g_i(x)$ כך ש־$q(T)^{n_i} u_i = g_i(T) v_1$. כעת, $q(x)^{n_1} = m_T(x)$ ולכן

$$0 = q(T)^{n_1} u_i = q(T)^{n_1-n_i} q(T)^{n_i} u_i = q(T)^{n_1-n_i} g_i(T) v_1$$

הפולינום המינימלי שמאפס את v_1 הוא $q(x)^{n_1}$ ולכן $q(x)^{n_1} \mid q(x)^{n_1-n_i} g_i(x)$. כלומר קיים $h_i(x)$ כך ש־$q(x)^{n_1} h_i(x) = q(x)^{n_1-n_i} g_i(x)$, ולכן $g_i(x) = q(x)^{n_i} h_i(x)$. נגדיר $w_i = u_i - h_i(T) v_1$.

כעת, נראה כי $q(T)^{n_i} w_i = 0$. אכן,

$$q(T)^{n_i} w_i = q(T)^{n_i} u_i - q(T)^{n_i} h_i(T) v_1 = q(T)^{n_i} u_i - g_i(T) v_1 = 0$$

ולסיום, נניח ש־$f(x)$ מקיים $f(T) w_i = 0$, ונראה ש־$f(x) \mid q(x)^{n_i}$. מ־$w = u_i - h_i(T) v_1$ נובע $u_i - w_i = h_i(T) v_1 \in Z_1$, ולכן $w_i \in u_i + Z_1 = v_i + Z_1$. $f(T) w_i = 0$ לכן

$$Z_1 = f(T) w_i + Z_1 = f\left(\overline{T}\right)(v_i + Z_1)$$

וכיוון שהפולינום המינימלי ב־$Z(v_i + Z_1, T)$ הוא $q(x)^{n_i}$ אז $f(x) \mid q(x)^{n_i}$ וסיימנו.

נותר להוכיח את יחידות הפירוק. נניח ש־$V = V_1 \oplus \cdots \oplus V_r$ כאשר V_i T־אינווריאנטים וקיים בסיס כך שהמטריצה המייצגת את T_i היא $C\left(q(x)^{n_i}\right)$, ו־$n = n_1 \ge \cdots \ge n_r$. אם גם $V = U_1 \oplus \cdots \oplus U_s$ הוא פירוק כנ"ל כאשר צמצום T ל־U_i מיוצג על ידי המטריצה $C\left(q(x)^{m_i}\right)$ כאשר $m = m_1 \ge \cdots \ge m_s$, צריך להוכיח ש־$r = s$ וגם $n_i = m_i$ לכל i.

נניח בשלילה שזה לא מתקיים, יהי k האינדקס הקטן ביותר כך ש־$n_k \neq m_k$ אבל $n_i = m_i$ לכל $i < k$. נניח כי $n_k > m_k$. מתקיים $q(T)^{m_k} U_j = 0$ לכל $j \ge k$, כי מההוכחה $q(x)^{m_k}$ הינו הפולינום המינימלי של צמצום T ל־U_k. לכן $q(T)^{m_k} V = q(T)^{m_k} U_1 \oplus \cdots \oplus q(T)^{m_k} U_{k-1}$. מתקיים $\dim q(T)^{m_k} = d(m_i - m_k)$ לכל $i \le k$ (תרגיל). אז $\dim\left(q(T)^{m_k} V\right) = d \sum_{j=1}^{k-1}(m_j - m_k)$.

מצד שני, $q(T)^{m_k} V \supseteq q(T)^{m_k} V_1 \oplus \cdots \oplus q(T)^{m_k} V_k$ ומהתרגיל $\dim q(T)^{m_k} V_i = d(n_i - m_k)$. זו סתירה, כי אז

$$\dim\left(q(T)^{m_k} V\right) \ge d \sum_{j=1}^{k}(n_j - m_k) = d \sum_{j=1}^{k-1}(m_j - m_k) + d(n_k - m_k) > d \sum_{j=1}^{k-1}(m_j - m_k) = \dim\left(q(T)^{m_k} V\right)$$

וסיימנו. ■

חלק III
מכפלה פנימית

6 הגדרות

בחלק זה נניח $\mathbb{F}=\mathbb{R}$ או $\mathbb{F}=\mathbb{C}$.

הגדרה 6.1 יהי V מרחב וקטורי מעל שדה $\mathbb{F}$. נאמר שV הוא **מרחב עם מכפלה פנימית** אם לכל $u,v\in V$ מוגדר סקלר $(u,v)\in\mathbb{F}$ כך שמתקיים:

1. $(u,v)=\overline{(v,u)}$
2. $(u,u)\geq 0$ **וגם** $(u,u)=0$ **אם ורק אם** $u=0$
3. $(\alpha u+\beta v,w)=\alpha\,(u,w)+\beta\,(v,w)$ **לכל** $u,v,w\in V$ **ו**$\alpha,\beta\in\mathbb{F}$.

דוגמה 6.2 אם $V=\mathbb{F}^n$, נגדיר עבור $u=(\alpha_1,\ldots,\alpha_n)\,,v=(\beta_1,\ldots,\beta_n)$ את המכפלה הפנימית הסטנדרטית (מכפלה סקלרית)

$$(u,v)=\sum_{i=1}^{n}\alpha_i\overline{\beta_i}$$

הערה 6.3 מתקיים

$$(u,\alpha v+\beta w)=\overline{(\alpha v+\beta w,u)}=\overline{\alpha\,(v,u)+\beta\,(w,u)}=\overline{\alpha}\overline{(v,u)}+\overline{\beta}\overline{(w,u)}=\overline{\alpha}\,(u,v)+\overline{\beta}\,(u,w)$$

טענה 6.4 יהי V מרחב וקטורי מעל $\mathbb{F}$, אז ניתן להגדיר על V מכפלה פנימית.

הוכחה: יהי $v_1,\ldots,v_n$ בסיס כלשהו של V. יהיו $u=\alpha_1v_1+\cdots+\alpha_nv_n$, $w=\beta_1v_1+\cdots+\beta_nv_n$ ונגדיר

$$(u,w)=\alpha_1\overline{\beta_1}+\cdots+\alpha_n\overline{\beta_n}$$

זו מכפלה פנימית על V. ■

הגדרה 6.5 יהי V מרחב וקטורי עם מכפלה פנימית, לכל $v\in V$ נגדיר את **הנורמה** של v

$$\|v\|=\sqrt{(v,v)}$$

הערה 6.6 מתקיים $\|v\|\geq 0$ **וגם** $\|v\|=0$ **אם ורק אם** $v=0$.

טענה 6.7 לכל $u\in V,\alpha\in\mathbb{F}$ מתקיים $\|\alpha u\|=|\alpha|\,\|u\|$.

הוכחה:

$$\|\alpha u\|^2=(\alpha u,\alpha u)=\alpha\overline{\alpha}\,(u,u)=|\alpha|^2\,\|u\|^2$$

■

משפט 6.8 אי שוויון קושי שוורץ: לכל $u,v\in V$ מתקיים $|(u,v)|\leq\|u\|\,\|v\|$.

הוכחה: אם $u=0$, אז $(u,v)=0$ לכל $v\in V$. כמו כן, $\|u\|=0$ ולכן אי השוויון מתקיים.
נניח $u\neq 0$, ונניח תחילה $(u,v)\in\mathbb{R}$. לכל $\lambda\in\mathbb{R}$ מתקיים

$$0\leq(\lambda u+v,\lambda u+v)=\lambda^2(u,u)+2\lambda(u,v)+(v,v)$$

ולכן הדיסקרימיננטה אי חיובית, כלומר $b^2-ac\leq 0$ ולכן $(u,v)^2\leq(u,u)(v,v)$, ולכן $|(u,v)|\leq\|u\|\|v\|$.
אם $(u,v)\notin\mathbb{R}$, $\alpha=(u,v)$, אז $\alpha\neq 0$ וגם $\left(\frac{1}{\alpha}u,v\right)=\frac{1}{\alpha}(u,v)=1$. הוכחנו את אי השוויון עבור ממשיים ובפרט עבור $\left(\frac{1}{\alpha}u,v\right)$, כלומר

$$1=\left|\left(\frac{1}{\alpha}u,v\right)\right|\leq\left\|\frac{1}{\alpha}u\right\|\|v\|=\left|\frac{1}{\alpha}\right|\|u\|\|v\|$$

ולכן $|(u,v)|=\alpha\leq\|u\|\|v\|$. ■

הגדרה 6.9 יהיו $u,v\in V$, נגדיר את **הזווית** θ ביניהם להיות הזווית המקיימת

$$\cos\theta=\frac{(u,v)}{\|u\|\|v\|}$$

זה מוגדר היטב מאי שוויון קושי שוורץ, שלפיו $\left|\frac{(u,v)}{\|u\|\|v\|}\right|\leq 1$.

7 וקטורים אורתוגונליים, היטלים ותהליך גראם־שמידט

הגדרה 7.1 $u,v\in V$ הם **מאונכים (אורתוגונליים)** אם $(u,v)=0$.

הגדרה 7.2 יהי $W\subseteq V$ תת מרחב, נגדיר את **המשלים האורתוגונלי של** W

$$W^{\perp}=\{u\in V\mid\forall w\in W.\,(u,w)=0\}$$

טענה 7.3 $W^{\perp}$ הינו תת מרחב של V.

הוכחה: אם $u,v\in W^{\perp}$, אז $(u,w)=(v,w)=0$ לכל $w\in W$. לכן, לכל $w\in W$

$$(\alpha u+\beta v,w)=\alpha(u,w)+\beta(v,w)=0$$

■

הגדרה 7.4 קבוצת וקטורים $\{v_1,\ldots,v_k\}$ נקראת **קבוצה אורתוגונלית** אם $(v_i,v_j)=0$ לכל $i\neq j$ וגם $(v_i,v_i)\neq 0$ לכל i.
קבוצה אורתוגונלית תקרא **קבוצה אורתונורמלית** אם בנוסף $(v_i,v_i)=1$ לכל i.

הערה 7.5 אם $\{v_1,\ldots,v_k\}$ אורתוגונלית, אז $\left\{\frac{1}{\|v_1\|}v_1,\ldots,\frac{1}{\|v_k\|}v_k\right\}$ היא אורתונורמלית ופורשת את אותו המרחב.

טענה 7.6 תהי $\{v_1,\ldots,v_k\}$ קבוצה אורתוגונלית, אז $\{v_1,\ldots,v_k\}$ היא בלתי תלויה.

הוכחה: נניח כי $\alpha_1v_1+\cdots+\alpha_kv_k=0$, אז לכל $1\leq i\leq k$ מתקיים

$$0=(0,v_i)=(\alpha_1v_1+\cdots+\alpha_kv_k,v_i)=\alpha_1(v_1,v_i)+\cdots+\alpha_k(v_k,v_i)=\alpha_i(v_i,v_i)$$

כאשר המעבר האחרון נובע מכך שהקבוצה אורתוגונלית, ולכן $(v_i,v_j)=0$ לכל $j\neq i$. כמו כן $(v_i,v_i)\neq 0$, ולכן בהכרח $\alpha_i=0$ וסיימנו. ■

מסקנה 7.7 בקבוצה אורתוגונלית יש לכל היותר $\dim V$ וקטורים.

משפט 7.8 תהליך גראם־שמידט: יהי V מרחב וקטורי עם מכפלה פנימית. אז ל־V יש בסיס אורתוגונלי (אורתונורמלי).

הוכחה: יהי $\{v_1,\dots,v_n\}$ בסיס כלשהו של V. נבנה מבסיס זה בסיס אורתוגונלי, שנסמנו ב$\{u_1,\dots,u_n\}$. נגדיר

$$\begin{aligned} u_1 &= v_1 \\ u_2 &= v_2 - \frac{(v_2,u_1)}{(u_1,u_1)}u_1 \\ u_{i+1} &= v_{i+1} - \frac{(v_{i+1},u_i)}{(u_i,u_i)}u_i - \dots - \frac{(v_{i+1},u_1)}{(u_1,u_1)}u_1 \\ &= v_{i+1} - \sum_{k=1}^{i} \frac{(v_{i+1},u_k)}{(u_k,u_k)}u_k \end{aligned}$$

כיוון שu_{i+1} הוא צירוף לינארי של $v_1,\dots,v_{i+1}$ אז $u_{i+1}\neq 0$, ולכן $(u_{i+1},u_{i+1})\neq 0$.
נניח שהוכחנו ש$\{u_1,\dots,u_j\}$ היא אורתוגונלית, ונראה ש$\{u_1,\dots,u_{j+1}\}$ אורתוגונלית. מהנחת האינדוקציה מספיק להוכיח ש$(u_{j+1},u_r)=0$ לכל $r\leq j$. נחשב:

$$\begin{aligned} (u_{j+1},u_r) &= \left(v_{j+1} - \sum_{k=1}^{j} \frac{(v_{j+1},u_k)}{(u_k,u_k)}u_k, u_r\right) = (v_{j+1},u_r) - \sum_{k=1}^{j} \frac{(v_{j+1},u_k)}{(u_k,u_k)}(u_k,u_r) \\ &= (v_{j+1},u_r) - \frac{(v_{j+1},u_r)}{(u_r,u_r)}(u_r,u_r) = 0 \end{aligned}$$

אז $\{u_1,\dots,u_n\}$ אורתוגונלית. לכן היא גם בלתי תלויה, וכמו כן יש בה n איברים, ולכן היא בסיס אורתוגונלי. ■

משפט 7.9 יהי $W\subseteq V$ תת מרחב, אז $V=W\oplus W^\perp$.

הוכחה: תחילה נראה שמתקיים $W\cap W^\perp=\{0\}$. יהי $w\in W\cap W^\perp$, אז $(w,w)=0$ ולכן $w=0$.
כעת נוכיח $V=W+W^\perp$. נניח ש$\dim W=r$, ויהי $\{w_1,\dots,w_r\}$ בסיס אורתונורמלי של W. יהי $v\in V$, נגדיר

$$v_0 = v-(v,w_1)w_1-\dots-(v,w_r)w_r = v-\sum_{i=1}^{r}(v,w_i)w_i$$

נוכיח $v_0\in W^\perp$. לכל $1\leq j\leq r$, נראה כי $(v_0,w_j)=0$. נחשב:

$$\begin{aligned} (v_0,w_j) &= \left(v-\sum_{i=1}^{r}(v,w_i)w_i, w_j\right) = (v,w_j)-\sum_{i=1}^{r}(v,w_i)(w_i,w_j) = (v,w_j)-(v,w_j)(w_j,w_j) \\ &= (v,w_j)-(v,w_j)\cdot 1 = 0 \end{aligned}$$

קיבלנו שכל $v\in V$ ניתן לפרק ל$v=v_0+\sum_{i=1}^{r}(v,w_i)w_i$, כאשר $\sum_{i=1}^{r}(v,w_i)w_i\in W, v_0\in W^\perp$ ולכן $V=W+W^\perp$.
■

הגדרה 7.10 יהי $W\subseteq V$ תת מרחב, ויהי $w_1,\dots,w_r$ בסיס אורתונורמלי של W. אז לכל $v\in V$, נגדיר את **ההיטל** של v על W

$$w=(v,w_1)w_1+\dots(v,w_r)w_r$$

כיוון שהסכום $V=W\oplus W^\perp$ הוא ישר, אז כל וקטור ניתן לרשום בצורה יחידה ולכן w הנ"ל אינו תלוי בבחירת הבסיס.

משפט 7.11 יהי $W\subseteq V$ תת מרחב, אז $\left(W^\perp\right)^\perp=W$.

הוכחה: על פי הגדרה $\left(W^\perp\right)^\perp=\{v\in V \mid \forall w'\in W^\perp.(v,w')=0\}$.
יהי $w\in W$, אז לכל $w'\in W^\perp$ מתקיים $(w,w')=0$. אז $W\subseteq\left(W^\perp\right)^\perp$.
מתקיים $V=W\oplus W^\perp$ וגם $V=W^\perp\oplus\left(W^\perp\right)^\perp$, ולכן $\dim W=\dim V-\dim W^\perp=\dim\left(W^\perp\right)^\perp$.
אז מהכלה ושוויון מימדים $W=\left(W^\perp\right)^\perp$. ■

הערה 7.12 זה לא בהכרח מתקיים עבור V שאינו נוצר סופית.

8 העתקות ולכסון אוניטרי

8.1 העתקות אוניטריות

עבור הטענות הבאות נניח כי V מרחב וקטורי מעל $\mathbb{C}$.

טענה 8.1 יהי V מרחב וקטורי מעל $\mathbb{C}$, תהי $T\colon V\to V$ ונניח כי $(Tv,v)=0$ לכל $v\in V$. אז $T=0$.

הוכחה: לכל $u,w\in V$ מתקיים

$$0=(T(u+w),u+w)=(Tu,u)+(Tu,w)+(Tw,u)+(Tw,w)=(Tu,w)+(Tw,u)$$

זהות זו נכונה לכל $w\in V$, ולכן נכונה גם אם נחליף את w ב-iw. אז

$$0=(Tu,iw)+(T(iw),u)=-i(Tu,w)+i(Tw,u)\implies-(Tu,w)+(Tw,u)=0$$

נחבר את המשוואות ונקבל $(Tw,u)=0$ לכל $u,w\in V$.

כעת, לכל $w\in V$ נבחר $u=Tw$ ונקבל $(Tw,Tw)=0$, ולכן מתכונות מכפלה פנימית $Tw=0$. אז $T=0$. ■

הערה 8.2 טענה זו אינה נכונה במרחב וקטורי מעל $\mathbb{R}$. לדוגמה, עבור $V=\mathbb{R}^2$ עם המכפלה הפנימית הסטנדרטית ו-$T\colon\mathbb{R}^2\to\mathbb{R}^2$ המוגדרת על ידי $T\begin{pmatrix}\alpha\\\beta\end{pmatrix}=\begin{pmatrix}-\beta\\\alpha\end{pmatrix}$, מתקיים לכל $\begin{pmatrix}\alpha\\\beta\end{pmatrix}\in\mathbb{R}^2$ ש-$\left(T\begin{pmatrix}\alpha\\\beta\end{pmatrix},\begin{pmatrix}\alpha\\\beta\end{pmatrix}\right)=\left(\begin{pmatrix}-\beta\\\alpha\end{pmatrix},\begin{pmatrix}\alpha\\\beta\end{pmatrix}\right)=0$ אבל $T\neq0$.

הערה 8.3 יהי $w\in V$, אם מתקיים $(u,w)=0$ לכל $u\in V$ אז $w=0$ (זה נכון גם במרחב וקטורי מעל $\mathbb{R}$).

הגדרה 8.4 יהי V מרחב וקטורי מעל $\mathbb{C}$. העתקה המקיימת $(Tu,Tv)=(u,v)$ לכל $u,v\in V$ נקראת **העתקה אוניטרית**.

העתקה אוניטרית שומרת על אורכים וזוויות, כלומר אם $v\in V$ ו-T אוניטרית אז $\|Tv\|=\|v\|$ וכנ"ל עבור זוויות.

טענה 8.5 T אוניטרית $\iff$ $\langle Tv,Tv\rangle=\langle v,v\rangle$ לכל $v\in V$.

הוכחה: $\Longleftarrow$ נובע ישירות מההגדרה.

$\Longrightarrow$ לכל $u,v\in V$ מתקיים

$$(Tu,Tu)+(Tu,Tv)+(Tv,Tu)+(Tv,Tv)=(T(u+v),T(u+v))=(u+v,u+v)=(u,u)+(u,v)+(v,u)+(v,v)$$

מתקיים $(Tu,Tu)=(u,u)$ וגם $(Tv,Tv)=(v,v)$, נצמצם ונקבל $(Tu,Tv)+(Tv,Tu)=(u,v)+(v,u)$.

נחליף את v ב-iv (השוויון מתקיים לכל v) ונקבל $-(Tu,Tv)+(Tv,Tu)=-(u,v)+(v,u)$, נחבר את המשוואות ונקבל $(Tv,Tu)=(v,u)$ ולכן T אוניטרית. ■

משפט 8.6 T אוניטרית $\iff$ T מעתיקה בסיס אורתונורמלי לבסיס אורתונורמלי. בפרט, אם T אוניטרית אז T הפיכה.

הוכחה: $\Longleftarrow$ יהי $\{v_1,\dots,v_n\}$ בסיס אורתונורמלי של V, כלומר $(v_i,v_j)=\begin{cases}0 & i\neq j\\1 & i=j\end{cases}$. נוכיח כי $\{Tv_1,\dots,Tv_n\}$ בסיס אורתונורמלי של V.

$\{Tv_1,\dots,Tv_n\}$ קבוצה אורתונורמלית כי מתכונות העתקה אוניטרית $(Tv_i,Tv_j)=(v_i,v_j)=\begin{cases}0 & i\neq j\\1 & i=j\end{cases}$, לכן ממשפט $\{Tv_1,\dots,Tv_n\}$ גם בלתי תלויה, יש בה n איברים ולכן היא בסיס אורתונורמלי.

$\Longrightarrow$ נניח כי $\{v_1,\dots,v_n\}$ ו-$\{Tv_1,\dots,Tv_n\}$ הם בסיסים אורתונורמליים, ונוכיח ש-T אוניטרית.

יהיו $u,w\in V$, נרשום $u=\alpha_1v_1+\dots+\alpha_nv_n,w=\beta_1v_1+\dots+\beta_nv_n$ ונקבל

$$(u,w)=(\alpha_1v_1+\dots+\alpha_nv_n,\beta_1v_1+\dots+\beta_nv_n)=\sum_{i,j}\alpha_i\overline{\beta_j}(v_i,v_j)=\alpha_1\overline{\beta_1}+\dots+\alpha_n\overline{\beta_n}$$

כעת, נחשב

$$(Tu,Tw)=(\alpha_1Tv_1+\dots+\alpha_nTv_n,\beta_1Tv_1+\dots+\beta_nTv_n)=\sum_{i,j}\alpha_i\overline{\beta_j}(Tv_i,Tv_j)=\alpha_1\overline{\beta_1}+\dots+\alpha_n\overline{\beta_n}$$

כאשר המעבר האחרון בשתי המשוואות נובע מכך ש-$v_1,\dots,v_n$ וגם $Tv_1,\dots,Tv_n$ בסיסים אורתונורמליים.

אז $(u,w)=(Tu,Tw)$ לכל $u,w\in V$, ולכן T אוניטרית. ■

8.2 הצמוד ההרמיטי

כעת נחזור לדון במרחב וקטורי V מעל $\mathbb{R}$ או $\mathbb{C}$.

משפט 8.7 תהי $T\colon V\to V$, אז לכל $v\in V$ **קיים** $w\in V$ **כך ש**$(Tu,v)=(u,w)$ **לכל** $u\in V$.
יתר על כן, הוקטור w נקבע באופן יחיד על ידי T,v.

הוכחה: תחילה נוכיח קיום. יהי $\{u_1,\ldots,u_n\}$ בסיס אורתונורמלי של V, ונגדיר

$$w=\overline{(Tu_1,v)}u_1+\cdots+\overline{(Tu_n,v)}u_n$$

כדי להוכיח $(Tu,v)=(u,w)$ לכל $u\in V$ מספיק להוכיח זאת עבור הבסיס $\{u_1,\ldots,u_n\}$. אכן, לכל $1\leq i\leq n$ מתקיים

$$(u_i,w)=\left(u_i,\overline{(Tu_1,v)}u_1+\cdots+\overline{(Tu_n,v)}u_n\right)=\sum_{j=1}^{n}\overline{\overline{(Tu_j,v)}}\,(u_i,u_j)=\sum_{j=1}^{n}(Tu_j,v)\,(u_i,u_j)=(Tu_i,v)$$

כעת נוכיח יחידות. נניח שw_1,w_2 מקיימים $(u,w_2)=(Tu,v)=(u,w_1)$ לכל $u\in V$. אז לכל $u\in V$ מתקיים $(u,w_1-w_2)=0$, ולכן מטענה שהוכחנו $w_1-w_2=0$ כלומר $w_1=w_2$. ■

הגדרה 8.8 בהינתן $T\colon V\to V$, הפונקציה השולחת כל $v\in V$ ל$w\in V$ המתאים מהמשפט הנ"ל נקראת **הצמוד ההרמיטי** של T, ומסומנת בT^* כלומר $T^*v=w$.
T,T^* מקיימות את הקשר $(Tu,v)=(u,T^*v)$ לכל $u,v\in V$.

טענה 8.9 ההעתקה $T^*\colon V\to V$ היא לינארית, ומתקיימות התכונות הבאות:

1. $(T^*)^*=T$.
2. $(S+T)^*=S^*+T^*$.
3. $(\alpha T)^*=\overline{\alpha}T^*$.
4. $(ST)^*=T^*S^*$.

הוכחה: נוכיח לינאריות. יהיו $u,v,w\in V$, אז

$$(u,T^*(v+w))=(Tu,v+w)=(Tu,v)+(Tu,w)=(u,T^*v)+(u,T^*w)=(u,T^*v+T^*w)$$

שוויון זה נכון לכל u, ולכן $T^*(v+w)=T^*v+T^*w$.
כפל בסקלר: תרגיל.
נוכיח את 1. לכל $u,v\in V$,

$$\left(u,(T^*)^*v\right)=(T^*u,v)=\overline{(v,T^*u)}=\overline{(Tv,u)}=(u,Tv)$$

שוויון זה נכון לכל u, ולכן $Tv=(T^*)^*v$ לכל $v\in V$ כלומר $T=(T^*)^*$. ■

דוגמה 8.10 נחשב את T^* עבור $T\colon\mathbb{C}^2\to\mathbb{C}^2$ המוגדרת על ידי $T(x,y)=(2x+3y,x-iy)$ והמכפלה הפנימית הסטנדרטית.
נבחר את הבסיס הסטנדרטי $e=\{e_1,e_2\}$, אז $Te_1=2e_1+e_2, Te_2=3e_1-ie_2$ ולכן $[T]_e=\begin{pmatrix}2&3\\1&-i\end{pmatrix}$. נניח

$$T^*e_1=\alpha_1e_1+\alpha_2e_2$$
$$T^*e_2=\beta_1e_1+\beta_2e_2$$

מהקשר $(Te_1,e_1)=(e_1,T^*e_1)$ נקבל $(2e_1+e_2,e_1)=(e_1,\alpha_1e_1+\alpha_2e_2)$ ולכן $\alpha_1=\overline{\alpha_1}=2$. מחישוב דומה ניתן להסיק $\alpha_2=3,\beta_1=1,\beta_2=i$. נקבל

$$T^*e_1=2e_1+3e_2$$
$$T^*e_2=e_1+ie_2$$

ולכן $[T^*]_e=\begin{pmatrix}2&1\\3&i\end{pmatrix}$, ומתקיים $[T^*]_e=\left(\overline{[T]_e}\right)^t$ (נוכיח בהמשך שתכונה זו נכונה לכל בסיס אורתונורמלי).
כמו כן, T^* היא

$$T^*(x,y)=xT^*e_1+yT^*e_2=2e_1+3e_2+e_1+ie_2=(2x+y,3x+iy)$$

טענה 8.11 T אוניטרית $\iff$ $T^*T=I$.

הוכחה: $\Leftarrow$ אם T אוניטרית, אז לכל $u,v\in V$ מתקיים $(u,T^*Tv)=(Tu,Tv)=(u,v)$ ולכן $T^*Tv=v$ לכל $v\in V$ כלומר $T^*T=I$.

$\Longrightarrow$ אם $T^*T=I$, אז לכל $u,v\in V$ מתקיים $(u,v)=(u,T^*Tv)=(Tu,Tv)$ ולכן T אוניטרית. ■

מסקנה 8.12 אם T אוניטרית אז T הפיכה, וגם $T^*=T^{-1}$.

משפט 8.13 יהי $e=\{v_1,\ldots,v_n\}$ בסיס אורתונורמלי, תהי T העתקה לינארית ונניח ש$[T]_e=A$, אז $[T^*]_e=\overline{A}^t$.

הוכחה: נניח $[T]_e=A=(\alpha_{ij})$ ו־$[T^*]_e=B=(\beta_{ij})$. אז לכל $1\le i\le n$ ולכל $1\le j\le n$

$$\begin{aligned}Tv_i&=\alpha_{1i}v_1+\cdots+\alpha_{ni}v_n\\ T^*v_j&=\beta_{1j}v_1+\cdots+\beta_{nj}v_n\end{aligned}$$

נחשב את (T^*v_j,v_i) בשני אופנים:

$$\begin{aligned}(T^*v_j,v_i)&=(\beta_{1j}v_1+\cdots+\beta_{nj}v_n,v_i)=\beta_{ij}\\ (T^*v_j,v_i)&=(v_j,Tv_i)=(v_j,\alpha_{1i}v_1+\cdots+\alpha_{ni}v_n)=\overline{\alpha_{ji}}\end{aligned}$$

ולכן $\beta_{ij}=\overline{\alpha_{ji}}$, כלומר $B=\overline{A}^t$. ■

מסקנה 8.14 נניח ש$[T]_e=A$ כאשר e בסיס אורתונורמלי. אז T אוניטרית $\iff$ שורות (עמודות) A מהוות קבוצה אורתונורמלית ב$\mathbb{F}^n$ ביחס למכפלה הפנימית הסטנדרטית.

הוכחה: נסמן $A=[T]_e=(\alpha_{ij})$ ואז מהמשפט הקודם $[T^*]_e=\overline{A}^t=(\overline{\alpha_{ji}})$. נניח כי $[TT^*]_e=(\beta_{ij})$, מתקיים $[TT^*]_e=[T]_e[T^*]_e$ ולכן

$$\beta_{ij}=\sum_{k=1}^{n}\alpha_{ik}\overline{\alpha_{jk}}=((\alpha_{i1},\ldots,\alpha_{in}),(\alpha_{j1},\ldots,\alpha_{jn}))=(R_i,R_j)$$

ממשפט קודם T אוניטרית $\iff TT^*=I$, כמו כן $TT^*=I$ $\iff$ $\beta_{ij}=\begin{cases}0 & i\neq j\\ 1 & i=j\end{cases}$ וכעת הוכחנו שזה קורה אם ורק אם $(R_i,R_j)=\begin{cases}0 & i\neq j\\ 1 & i=j\end{cases}$ וסיימנו.

אם נחשב באופן דומה את T^*T נקבל את הטענה בעבור עמודות במקום שורות. ■

הגדרה 8.15 מטריצה תקרא **אוניטרית** אם $A^*=A^{-1}$ כאשר $A^*=\overline{A}^t$.
מטריצה A היא אוניטרית אם ורק אם שורות (עמודות) A מהוות קבוצה אורתונורמלית ביחס למכפלה הפנימית הסטנדרטית.

8.3 העתקות אורתוגונליות

נעבור למקרה $F=\mathbb{R}$.

הגדרה 8.16 יהי V מרחב וקטורי עם מכפלה פנימית המוגדרת מעל $\mathbb{R}$. T נקראת **אורתוגונלית** אם $(Tu,Tv)=(u,v)$ לכל $u,v\in V$ (מקביל להעתקה אוניטרית, אך ב$\mathbb{R}$).

הערה 8.17 המקביל לטענה הראשונה על העתקות אוניטריות מעל $\mathbb{C}$, עבור העתקות מעל $\mathbb{R}$: תהי S כך ש$S=S^*$, אם $(Su,u)=0$ לכל $u\in V$ אז $S=0$. כלומר, נוסף התנאי $S=S^*$.

טענה 8.18 יהי V מרחב וקטורי מעל $\mathbb{R}$ או $\mathbb{C}$ עם מכפלה פנימית.
אזי התנאים הבאים שקולים:

1. $T^*=T^{-1}$, כלומר $TT^*=T^*T=I$.
2. לכל $u,v\in V$ מתקיים $(Tu,Tv)=(u,v)$, כלומר אם מעל $\mathbb{C}$ T אוניטרית ומעל $\mathbb{R}$ T אורתוגונלית.

3. לכל v $\|Tv\| = \|v\|$, כלומר T שומרת על אורכים.

הוכחה: מעל $\mathbb{F} = \mathbb{C}$ הוכחנו.
$2 \implies 1$: אם $T^*T = I$ אז $(Tu, Tv) = (u, T^*Tv) = (u, v)$.
$3 \implies 2$: ברור.
$1 \implies 3$: אם $(Tv, Tv) = (v, v)$ לכל v אז $(T^*Tv, v) = (Tv, Tv) = (v, v) = (Iv, v)$, נעביר אגפים ונקבל $((T^*T - I)\, v, v) = 0$. כיוון ש-$(T^*T - I)^* = T^*T - I$ אז מההערה (המקביל לטענה ב-$\mathbb{C}$) ניתן להסיק $T^*T - I = 0$ או $T^*T = I$. ■

המשפטים שהוכחנו על העתקות אוניטריות מתקיימים גם עבור העתקות אורתוגונליות.

משפט 8.19 T אורתוגונלית $\iff$ T מעתיקה בסיס אורתונורמלי לבסיס אורתונורמלי.

משפט 8.20 יהי e בסיס אורתונורמלי של V, אם $[T]_e = A$ אז $[T^*]_e = A^*$ כאשר $A^* = A^t$.

משפט 8.21 מטריצה היא אורתוגונלית (אם $A^* = A^{-1}$, או $A^t = A^{-1}$) אם ורק אם שורות (עמודות) A מהוות מערכת אורתונורמלית ביחס למכפלה הפנימית הסטנדרטית.

8.4 העתקות הרמיטיות (צמודות לעצמן)

נחזור למקרה $\mathbb{F} = \mathbb{C}$.

הגדרה 8.22 העתקה T תקרא **צמודה לעצמה (הרמיטית)** אם $T = T^*$. אם $T^* = -T$, אז T נקראת **אנטי הרמיטית**.

הערה 8.23 כל העתקה S ניתן לרשום בצורה הבאה: $S = \frac{S+S^*}{2} + i\left(\frac{S-S^*}{2i}\right)$. גם $\frac{S+S^*}{2}$ וגם $\frac{S-S^*}{2i}$ הן צמודות לעצמן, ולכן נסיק שכל העתקה S ניתן לרשום כ-$S = A + iB$ כאשר A, B צמודות לעצמן.

משפט 8.24 אם T צמודה לעצמה, אז כל הערכים העצמיים שלה הם ממשיים.

הוכחה: יהי λ ערך עצמי של T, ויהי v וקטור עצמי המתאים ל-λ. כלומר, $Tv = \lambda v$. אז

$$\lambda\,(v, v) = (\lambda v, v) = (Tv, v) = (v, T^*v) = (v, Tv) = (v, \lambda v) = \overline{\lambda}\,(v, v)$$

$v \neq 0$ לכן גם $(v, v) \neq 0$, אז $\lambda = \overline{\lambda}$ כלומר λ ממשי. ■

8.5 העתקות נורמליות ולכסון אוניטרי

נדון במרחב וקטורי מעל $\mathbb{C}$.

טענה 8.25 (1) תהי T לינארית, אם $T^*Tv = 0$ אז $Tv = 0$.

הוכחה: מתקיים

$$0 = (0, v) = (T^*Tv, v) = \left(Tv, (T^*)^*\, v\right) = (Tv, Tv) \implies Tv = 0$$

ולכן $Tv = 0$. ■

טענה 8.26 (2) אם T צמוד לעצמו $(T = T^*)$ ו-$T^k v = 0$ עבור $k \geq 1$ אז $Tv = 0$.

הוכחה: מספיק להוכיח שאם $T^{2^m} v = 0$ עבור m מסוים אז $Tv = 0$. נגדיר $S = T^{2^{m-1}}$, כיוון ש-$T^* = T$ אז $S^* = S$. כמו כן $S^*S = T^{2^{m-1}} \cdot T^{2^{m-1}} = T^{2^m}$ מנתון $T^{2^m} v = 0$ ולכן $S^*Sv = 0$, ומטענה (1) $Sv = 0$. קיבלנו $T^{2^{m-1}} v = Sv = 0$. באינדוקציה נקבל $Tv = 0$. ■

הגדרה 8.27 העתקה T נקראת **נורמלית** אם $TT^* = T^*T$.

הערה 8.28 כל ההעתקות האוניטריות / צמודות לעצמן הן נורמליות, אך לא כל ההעתקות הנורמליות הן אוניטריות / צמודות לעצמן.
לדוגמה, עבור $A = \begin{pmatrix} 1 & 1 \\ i & 3+2i \end{pmatrix}$ מתקיים $AA^* = A^*A$ אבל A אינה אוניטרית או צמודה לעצמה.

טענה 8.29 (3) אם N נורמלית ו$Nv=0$ עבור $v\in V$ אז $N^*v=0$.

הוכחה:

$$\left(N^*v,N^*v\right)=\left(NN^*v,v\right)=\left(N^*Nv,v\right)=\left(Nv,\left(N^*\right)^*v\right)=\left(Nv,Nv\right)=0$$

ולכן $N^*v=0$. ■

טענה 8.30 (4) אם λ הוא ערך עצמי של העתקה נורמלית N ואם $Nv=\lambda v$ אז $N^*v=\overline{\lambda}v$.

הוכחה: נחשב

$$\left(N-\lambda I\right)\left(N-\lambda I\right)^*=\left(N-\lambda I\right)\left(N^*-\overline{\lambda}I\right)=NN^*-\overline{\lambda}N-\lambda N^*+\lambda\overline{\lambda}I$$
$$\left(N-\lambda I\right)^*\left(N-\lambda I\right)=\left(N^*-\overline{\lambda}I\right)\left(N-\lambda I\right)=N^*N-\overline{\lambda}N-\lambda N^*+\lambda\overline{\lambda}I$$

N נורמלית לכן $NN^*=N^*N$, ולכן גם $N-\lambda I$ נורמלית.
לכן, אם $Nv=\lambda v$ אז $\left(N-\lambda I\right)v=0$ ואז מטענה (3) $\left(N^*-\overline{\lambda}I\right)v=\left(N-\lambda I\right)^*v=0$ ולכן $N^*v=\overline{\lambda}v$. ■

טענה 8.31 (5) אם T אוניטרית וλ ערך עצמי של T אז $|\lambda|=1$.

הוכחה: T אוניטרית ולכן נורמלית. λ ערך עצמי כלומר קיים וקטור $v\neq 0$ כך ש$Tv=\lambda v$, מטענה (4) מתקיים $T^*v=\overline{\lambda}v$ ולכן

$$v=Iv=T^*Tv=T^*\lambda v=\lambda T^*v=\lambda\overline{\lambda}v$$

ולכן $\lambda\overline{\lambda}=1$, כלומר $|\lambda|=1$. ■

טענה 8.32 (6) אם N נורמלית ו$N^kv=0$אז $Nv=0$.

הוכחה: נסמן $S=N^*N$, אז $S^*=S$, ומתקיים

$$S^kv=\left(N^*N\right)^kv\overset{(*)}{=}\left(N^*\right)^kN^kv=0$$

כאשר $(*)$ נובע מנורמליות. $S^kv=0$ וS צמודה לעצמה לכן לפי טענה (2) $Sv=0$. כלומר $N^*Nv=0$, ומטענה (1) נובע $Nv=0$. ■

טענה 8.33 (7) אם N נורמלית ואם $\left(N-\lambda I\right)^kv=0$ עבור $v\in V$, $\lambda\in\mathbb{C}$ אז $Nv=\lambda v$.

הוכחה: אם N נורמלית אז $N-\lambda I$ נורמלית, והטענה נובעת מטענה (6). ■

טענה 8.34 (8) תהי N נורמלית ויהיו λ,μ ערכים עצמיים שונים של N. אם $v,w\in V$ כך ש$Nv=\lambda v$ ו$Nw=\mu w$, אז $(v,w)=0$.

הוכחה: נחשב

$$\left(Nv,w\right)=\left(\lambda v,w\right)=\lambda\left(v,w\right)$$
$$\left(Nv,w\right)=\left(v,N^*w\right)\overset{(*)}{=}\left(v,\overline{\mu}w\right)=\mu\left(v,w\right)$$

כאשר $(*)$ נובע מטענה (4). לכן $\lambda\left(v,w\right)=\mu\left(v,w\right)$, וכיוון ש$\mu\neq\lambda$ בהכרח $(v,w)=0$. ■

משפט 8.35 ***כל העתקה נורמלית ניתנת ללכסון אוניטרי:*** תהי $N\colon V\to V$ העתקה נורמלית אז קיים בסיס אורתונורמלי המורכב מוקטורים עצמיים של N.
כלומר, אם N מטריצה נורמלית אז קיימת מטריצה אוניטרית U כך שUNU^{-1} היא אלכסונית.

הוכחה: תהי N נורמלית ויהיו $\lambda_1, \ldots, \lambda_k$ כל הערכים העצמיים השונים של N. ממשפט הפירוק היסודי $V = V_1 \oplus \cdots \oplus V_k$ כאשר $V_i = \ker (N - \lambda_i I)^{n_i}$.

אם $v \in V_i$ אז $(N - \lambda_i I)^{n_i} v = 0$, מטענה (6) נובע $(N - \lambda_i I) v = 0$ כלומר $Nv = \lambda_i v$. לכן כל וקטור $v \in V_i$ הוא וקטור עצמי עם ערך עצמי λ_i. מתהליך גראם־שמידט ל־V_i יש בסיס אורתונורמלי. מטענה (8) וקטורים השייכים ל־V_i ול־V_j כך ש־$i \neq j$ מאונכים זה לזה, ולכן מהפירוק $V = V_1 \oplus \cdots \oplus V_k$ ניתן למצוא ל־V בסיס אורתונורמלי המורכב כולו מוקטורים עצמיים של N. לכן, קיימת U אוניטרית כך ש־UNU^{-1} היא אלכסונית. ■

הערה 8.36 זו שקילות, כלומר אם A היא מטריצה שניתנת ללכסון אוניטרי (כלומר $A = UDU^*$ עבור U אוניטרית) אז A בהכרח נורמלית:

$$AA^* = UDU^* \left(UDU^*\right)^* = UDU^*UD^*U^* = UDD^*U^* = UD^*DU^* = UD^*U^*UDU^* = \left(UDU^*\right)^* UDU^* = A^*A$$

כעת נניח ש־$\mathbb{F} = \mathbb{R}$. נזכיר

1. T צמודה לעצמה: כלומר במטריצות $A^* = A$, בממשיים $A^t = A$ כלומר A סימטרית.
2. T אורתוגונלית: כלומר במטריצות $A^* = A^{-1}$, בממשיים $A^t = A^{-1}$.

T צמודה לעצמה (סימטרית)

טענה 8.37 תהי A מטריצה סימטרית, אז

1. $\Delta_A(x)$ מכפלה של גורמים לינאריים מעל $\mathbb{R}$.
2. ל־A יש וקטור עצמי מעל $\mathbb{R}$.
3. וקטורים עצמיים השייכים לערכים עצמיים שונים מאונכים זה לזה.

הוכחה: ■

1. נסתכל על A כעל מטריצה צמודה לעצמה מעל $\mathbb{C}$, הוכחנו בעבר שלמטריצה זו כל הערכים העצמיים הם ממשיים.
2. נובע מ־1.
3. מתקיים לפי טענה (8).

משפט 8.38 כל העתקה סימטרית (צמודה לעצמה מעל $\mathbb{R}$) ניתנת ללכסון מעל $\mathbb{R}$ על ידי העתקה אורתוגונלית.
כלומר, כל מטריצה סימטרית A ניתן ללכסן על ידי מטריצה אורתוגונלית P כלומר מתקיים $D = P^tAP = P^{-1}AP$ עבור D אלכסונית.

הוכחה: נוכיח באינדוקציה על $\dim V$. אם $\dim V = 1$, ברור.
נניח $\dim V = n > 1$. מהטענה הקודמת (חלק 2) ל־T ויש וקטור עצמי v_1 מעל $\mathbb{R}$. נגדיר $W = \mathrm{Sp}\left\{v_1\right\}$, ונסמן $u_1 = \frac{1}{\|v_1\|} v_1$. W הוא T אינווריאנטי וממשפט $V = W \oplus W^\perp$ אז T^* שומר על $W^\perp$ (תרגיל). אבל $T^* = T$ ולכן $W^\perp$ הוא T אינווריאנטי וממימדו $n-1$. מהנחת האינדוקציה נובע שקיים בסיס אורתונורמלי $\{u_2, \ldots, u_n\}$ של $W^\perp$ שבו צמצום של T ל־$W^\perp$ היא אלכסונית. אז $\{u_1, u_2, \ldots, u_n\}$ הוא הבסיס המבוקש של V. ■

T אורתוגונלית

טענה 8.39 תהי A מטריצה אורתוגונלית מסדר 2. אז $\det A = \pm 1$, אם $\det A = 1$ אז $A = A(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ (מטריצת סיבוב), אם $\det A = -1$ אז $A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.

הוכחה: כיוון ש־$A^t = A^{-1}$, $A^tA = I$ ולכן $(\det A)^2 = 1$ כלומר $\det A = \pm 1$.

אם $\det A = 1$: נסמן $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. A אורתוגונלית לכן שורותיה R_1, R_2 הן מערכת אורתונורמלית, ולכן

$$\begin{aligned} a^2 + b^2 &= 1 \\ c^2 + d^2 &= 1 \\ (R_1, R_2) = ac + bd &= 0 \\ \det A = ad - bc &= 1 \end{aligned}$$

אם $a = 0$, אז מהמשוואה הראשונה $b^2 = 1$ ולכן $b = \pm 1$ ומהמשוואה השלישית $d = 0$. אז מהמשוואה הרביעית $c = -b$. לכן $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ (כלומר $\theta = \frac{\pi}{2}$) או $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ (כלומר $\theta = -\frac{\pi}{2}$).

אם $a \neq 0$, אז מהמשוואה השלישית $c = -\frac{bd}{a}$, נציב במשוואה השנייה ונקבל $\frac{b^2d^2}{a^2} + d^2 = 1$ כלומר $\left(a^2 + b^2\right) d^2 = b^2d^2 + a^2d^2 = a^2$ ולכן מהמשוואה הראשונה $a^2 = d^2$ כלומר $a = \pm d$. אם $a = -d$ אז מהמשוואה השלישית נקבל $b = c$, מהמשוואה הרביעית נקבל $-a^2 - b^2 = 1$ וזו סתירה. לכן בהכרח $a = d$, ואז $b = -c$. בסך הכל $A = \begin{pmatrix} a & -c \\ c & a \end{pmatrix}$ וגם $a^2 + c^2 = 1$, תנאי זה מבטיח שקיים θ כך ש־$a = \cos\theta, c = \sin\theta$.

אם $\det A = -1$: נגדיר $A \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = B$ ואז $B^tB = A^t \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} A = A^tIA = A^tA = I$ כלומר B אורתוגונלית. כמו כן, $\det B = 1$. ■

משפט 8.40 תהי T העתקה אורתוגונלית מעל $\mathbb{R}$, אז קיים בסיס אורתונורמלי שבו T ניתנת לייצוג על ידי

$$\begin{pmatrix} I_p & & & & \\ & -I_k & & & \\ & & A(\theta_1) & & \\ & & & \ddots & \\ & & & & A(\theta_r) \end{pmatrix}$$

כאשר $A(\theta_1), \ldots, A(\theta_r)$ מטריצות סיבוב (מסדר 2).

הוכחה: נסמן $S = T + T^* = T + T^{-1}$ ואז $S^* = (T + T^*)^* = T^* + T = S$ אז S צמודה לעצמה, כלומר סימטרית. לכן ממשפט קיים ל־V בסיס אורתונורמלי המורכב מוקטורים עצמיים של S. יהיו $\lambda_1, \ldots, \lambda_m$ הערכים העצמיים השונים של S ואז $V = V_1 \oplus \cdots \oplus V_m$ כאשר V_i הוא המרחב העצמי של λ_i ביחס ל־S. אז V_i אינווריאנטי תחת T, כי אם $v \in V_i$ אז $Sv = \lambda_i v$ לכן $STv = \left(T + T^{-1}\right) Tv = T\left(T + T^{-1}\right) v = TSv = \lambda_i Tv$ כלומר גם Tv הוא וקטור עצמי של S עבור ערך עצמי λ_i ולכן $Tv \in V_i$. לכן מספיק להוכיח את המשפט לצמצום T ל־V_i.

על V_i מתקיים $\left(T + T^{-1}\right) v = Sv = \lambda_i v$, ולכן $\left(T^2 - \lambda_i T + I\right) v = 0$. נחלק למקרים:

אם $\lambda_i = \pm 2$, אז $(T \pm I)^2 v = 0$. T אורתוגונלית ולכן מעל $\mathbb{C}$ היא נורמלית, והוכחנו שאז $T \pm I$ גם היא נורמלית. הוכחנו שאם $T \pm I$ נורמלית אז $(T \pm I)^2 v = 0$ גורר $(T \pm I) v = 0$. לכן $(T \pm I) v = 0$, כלומר $Tv = \pm v$ ולכן במקרה זה צמצום T ל־V_i הוא I או $-I$.

אם $\lambda_i \neq \pm 2$, במקרה זה ל־T אין וקטורים עצמיים ב־V_i כי ממשפט להעתקות אוניטריות אם λ הוא ערך עצמי אז $|\lambda| = 1$, ואם λ ממשי אז $\lambda = \pm 1$. זה דורש $\lambda_i = \pm 2$ כפי שראינו, ולכן זה לא יכול להתקיים במקרה זה. לכן, אם $v \neq 0$ ב־V_i אז $\{v, Tv\}$ היא קבוצה בלתי תלויה. נסמן $W = \mathrm{Sp}\{v, Tv\}$, אז W הוא T אינווריאנטי כי $Tv \in W, T^2v = \lambda_i Tv - v \in W$. נרשום $V_i = W \oplus W^\perp$, אז $W^\perp$ גם הוא T אינווריאנטי (תרגיל). לכן $V_i = W_1 \oplus \cdots \oplus W_r$ כאשר $\dim W_j = 2$ לכל j וגם W_l, W_j מאונכים לכל $l \neq j$, וגם W_j הוא T אינווריאנטי לכל j. לכן מספיק להסתכל על צמצום T ל־W_j כאשר $1 \leq j \leq r$.

כיוון ש־$T^2 - \lambda_i T + I = 0$, אז אם נסמן ב־$\Delta(x)$ את הפולינום האופייני של צמצום T ל־W_j אז $\Delta(x) = x^2 - \lambda_i x + 1$ ולכן $\det T = 1$. לכן צמצום T ל־W_j היא העתקה אורתוגונלית כך ש־$\det T = 1$, לכן צמצום T ל־W_j ניתן לייצוג על ידי $A(\theta_j)$ מתאים. ■

מסקנה 8.41 סיכום התכונות של העתקות מעל $\mathbb{C}$ ומעל $\mathbb{R}$:

1. אם N נורמלית (מעל $\mathbb{C}$) או T סימטרית (מעל $\mathbb{R}$) אז הפולינום המינימלי של N (מעל $\mathbb{C}$) ושל T (מעל $\mathbb{R}$) הינו מכפלה של גורמים לינאריים שונים בריבוי אחד.
2. אם T אורתוגונלית (מעל $\mathbb{R}$) אז הפולינום המינימלי של T הוא מהצורה

$$m_T(x) = (x+1)^\alpha (x-1)^\beta \left[(x-a_1)^2 + b_1^2\right]^{c_1} \cdots \left[(x-a_r)^2 + b_r^2\right]^{c_r}$$

כאשר $\alpha, \beta, c_i = 0, 1$ ו־$b_i \neq 0$, וגם $a_i^2 + b_i^2 = 1$.

9 פירוק QR, בעיות קירוב ובעיית הריבועים הפחותים

פירוק QR

משפט 9.1 (פירוק QR) תהי B מטריצה הפיכה מסדר n המוגדרת מעל $\mathbb{C}$. אז קיימת מטריצה משולשית תחתונה M כך שאיברי האלכסון של M הם מספרים חיוביים, וכך שMB היא מטריצה אוניטרית. כמו כן, פירוק זה הוא יחיד.

הערה 9.2 אם M היא משולשית כנ"ל, אז גם M^{-1} היא משולשית כנ"ל. לכן, ניתן לרשום פירוק זה גם בצורה $B=M^{-1}U$ כאשר U אוניטרית.

הוכחה: יהיו $\beta_1,\ldots,\beta_n\in\mathbb{C}^n$ שורות המטריצה B. כיוון שB הפיכה, הקבוצה $\{\beta_1,\ldots,\beta_n\}$ מהווה בסיס. יהיו $\{\alpha_1,\ldots,\alpha_n\}$ הוקטורים המתקבלים מ$\{\beta_1,\ldots,\beta_n\}$ על ידי ביצוע תהליך גראם־שמידט. כלומר, לכל $1\le k\le n$ $\alpha_j=\beta_k-\sum_{j=1}^{k-1}\frac{(\beta_k,\alpha_j)}{(\alpha_j,\alpha_j)}\alpha_j$, וכמו כן לכל $1\le k\le n$ הקבוצה $\{\alpha_1,\ldots,\alpha_k\}$ היא בסיס אורתוגונלי לקבוצה $\mathrm{Sp}\{\beta_1,\ldots,\beta_k\}$. לכן, לכל $1\le k\le n$ קיימים סקלרים יחידים $c_{k_j}\in\mathbb{C}$ עבורם $\alpha_k=\beta_k\sum_{j=1}^{k-1}c_{k_j}\beta_j$.

נגדיר את המטריצה U להיות המטריצה האוניטרית ששורותיה הן $\frac{1}{\|\alpha_1\|}\alpha_1,\ldots,\frac{1}{\|\alpha_n\|}\alpha_n$. נגדיר את המטריצה M באופן הבא:

$$M_{kj}=\begin{cases}\frac{c_{k_j}}{\|a_k\|} & j\le k\\ \frac{1}{\|a_k\|} & j=k\\ 0 & j>k\end{cases}$$

אז M אכן מטריצה משולשית תחתונה עם איברי אלכסון ממשיים וחיוביים. כמו כן, אכן מתקיים

$$\frac{1}{\|\alpha_k\|}\alpha_k=\frac{1}{\|\alpha_k\|}\beta_k+\sum_{j=1}^{k-1}\frac{c_{k_j}}{\|\alpha_k\|}\beta_j=\sum_{j=1}^{n}M_{kj}B_j$$

כלומר שורות U שוות לשורות MB, ולכן $U=MB$.

נותר להראות שהפירוק הינו יחיד. כדי להוכיח זאת, נעיר תחילה שאוסף כל המטריצות המשולשיות התחתונות, ואוסף כל המטריצות האוניטריות הן חבורות - כלומר, סגורות תחת כפל והופכי. כעת, נניח שM_1,M_2 הן מטריצות משולשיות תחתונות כך שM_1B,M_2B הן מטריצות אוניטריות. מההערה הנ"ל נובע ש$(M_1B)\cdot(M_2B)^{-1}$ היא מטריצה אוניטרית. מתקיים $(M_1B)\cdot(M_2B)^{-1}=M_1BB^{-1}M_2^{-1}=M_1M_2^{-1}$, ומההערה הנ"ל נובע ש$M_1M_2^{-1}$ היא אוניטרית, וגם מטריצה משולשית תחתונה שאיברי האלכסון שלה חיובים. על כן, מאוניטריות מתקיים $\left(M_1M_2^{-1}\right)^{-1}=\left(M_1M_2^{-1}\right)^*=\left(\overline{M_1M_2^{-1}}\right)^t$. אגף שמאל היא מטריצה משולשית תחתונה, ואגף ימין היא מטריצה משולשית עליונה, ולכן $M_1M_2^{-1}$ היא מטריצה אלכסונית (וגם אוניטרית). לכן כל איברי האלכסון הם חיוביים, וכמו כן ערכים עצמיים של מטריצה אוניטרית. הוכחנו שכל הערכים העצמיים של מטריצה אוניטרית הם בעלי ערך מוחלט 1, ולכן בהכרח איברי האלכסון הם כולם 1. לכן $M_1M_2^{-1}=I$, כלומר $M_1=M_2$. ■

בעיות קירוב

יהי V מרחב וקטורי עם מכפלה פנימית מעל $\mathbb{F}=\mathbb{R},\mathbb{C}$. אם $v,w\in V$, אז $\|v-w\|$ הוא המרחק בין v לw.

הבעיה: יהי $W\subseteq V$ תת מרחב כלשהו, ויהי $v\in V$. אנו רוצים למצוא $w_0\in W$ שיהיה "הקרוב" ביותר לv, כלומר למצוא $w_0\in W$ כך ש$\|v-w_0\|\le\|v-w\|$ לכל $w\in W$.

השערה: הוכחנו ש$V=W\oplus W^{\perp}$. אם $v=w_0+w_0'$, אז אנו מצפים שw_0, ההיטל, יהיה הוקטור הקרוב ביותר.

משפט 9.3 יהי V מרחב וקטורי עם מכפלה פנימית, יהי $W\subseteq V$ תת מרחב ויהי $v\in V$. אז:

1. הוקטור $w_0\in W$ הוא "הקרוב" ביותר לv $\iff$ $v-w_0\in W^{\perp}$.
2. אם $w_0\in W$ כנ"ל קיים, אז w_0 יחיד.
3. אם $w_1,\ldots,w_k$ הינו בסיס אורתונורמלי של W, אז $w_0=(v,w_1)\,w_1+\cdots+(v,w_k)\,w_k$.

הוכחה: נוכיח את (1). $\Longrightarrow$ נניח $v-w_0\in W^{\perp}$, ונראה כי $\|v-w_0\|<\|v-w\|$ לכל $w_0\neq w\in W$. מתקיים

$$\begin{aligned}\|v-w\|^2&=(v-w,v-w)=(v-w_0+w_0-w,v-w_0+w_0-w)\\&=(v-w_0,v-w_0)+(v-w_0,w_0-w)+(w_0-w,v-w_0)+(w_0-w,w_0-w)\\&=(v-w_0,v-w_0)+(w_0-w,w_0-w)=\|v-w_0\|^2+\|w_0-w\|^2\overset{w\neq w_0}{>}\|v-w_0\|^2\end{aligned}$$

כאשר $(v-w_0,w_0-w)=(w_0-w,v-w_0)=0$ כי $v-w_0\in W^{\perp},w_0-w\in W$.

$\Longleftarrow$ נניח $\|v-w_0\|\geq\|v-w\|$ לכל $w\in W$, ונוכיח $v-w_0\in W^{\perp}$. כפי שחישבנו,

$$\|v-w\|^2=\|v-w_0\|^2+(v-w_0,w_0-w)+(w_0-w,v-w_0)+\|w_0-w\|^2$$
$$\|v-w\|^2-\|v-w_0\|^2=(v-w_0,w_0-w)+(w_0-w,v-w_0)+\|w_0-w\|^2$$

לפי הנתון אגף שמאל אינו שלילי, ולכן ניתן להסיק $(v-w_0,w_0-w)+(w_0-w,v-w_0)+\|w_0-w\|^2\geq 0$. נסמן אי שוויון זה ב$(*)$ץ

כדי להראות ש$v-w_0\in W^{\perp}$, מספיק להוכיח ש$(v-w_0,w_1)=0$ לכל $w_1\in W$ כך ש$\|w_1\|=1$. בהינתן $w_1\in W$ כך ש$\|w_1\|=1$ נבחר את w כך שיתקיים $w_0-w=\alpha w_1$, עבור $\alpha\in\mathbb{F}$ שיקבע בהמשך. לכן, כיוון שאי שוויון $(*)$ נכון לכל $w\in W$, נסיק

$$(v-w_0,\alpha w_1)+(\alpha w_1,v-w_0)+\|\alpha w_1\|^2\geq 0$$
$$\overline{\alpha}\,(v-w_0,w_1)+\alpha\overline{(v-w_0,w_1)}+|\alpha|^2\geq 0$$

נבחר $\alpha=-(v-w_0,w_1)$, ואז נקבל $-\alpha\overline{\alpha}-\alpha\overline{\alpha}+|\alpha|^2\geq 0$, כלומר $-|\alpha|^2\geq 0$. לכן, בהכרח $(v-w_0,w_1)=\alpha=0$ כדרוש.

נוכיח את (2). אם $v-w_0\in W^{\perp}$ וגם $v-w_0'\in W^{\perp}$, אז $(v-w_0',w)=(v-w_0,w)=0$ לכל $w\in W$, ולכן $(w_0-w_0',w)=0$ לכל $w\in W$. נבחר $w=w_0-w_0'$, ואז נובע $w_0-w_0'=0$ כלומר $w_0=w_0'$.

את (3) הוכחנו בעבר. ■

בעיית הריבועים הפחותים

הבעיה: נניח שנתונות n נקודות על המישור, מחפשים ישר שיהיה הקרוב ביותר לנקודות. נניח שהישר נתון על ידי $y=ax+b$, אם הנקודות הן $(x_1,y_1),\ldots,(x_n,y_n)$ המצב האופטימלי הוא

$$\begin{aligned}ax_1+b&=y_1\\ax_2+b&=y_2\\&\vdots\\ax_n+b&=y_n\end{aligned}$$

כלומר

$$\begin{pmatrix}x_1&1\\x_2&1\\\vdots&\vdots\\x_n&1\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix}=\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}$$

זוהי מערכת לינארית מהצורה $Ax=b$, למערכת זו בדרך כלל לא יהיה פתרון (הנקודות לא בהכרח נמצאות על אותו הישר). לכן, נחפש וקטור $\hat{x}$ כך ש$\|b-A\hat{x}\|$ יהיה מינימלי. כלומר, נחפש $\hat{x}\in\mathbb{R}^n$ כך ש$\|b-A\hat{x}\|\leq\|b-Ax\|$ לכל $x\in\mathbb{R}^n$.

פתרון: כל וקטור מהצורה Ax שייך ל$C(A)$, ממה שהוכחנו אם נסמן את הטלת b על $C(A)$ ב$\hat{b}$ אז למערכת $Ax=\hat{b}$ יש פתרון, ופתרון זה שנסמנו ב$\hat{x}$ הוא הקרוב ביותר.

במקרה זה, ניתן לרשום את הפתרון בדרך אחרת: נניח ש$A\hat{x}=\hat{b}$, אז $b-\hat{b}$ מאונך ל$C(A)$ כלומר אם נסמן את עמודות A בa_j אז מתקיים $\left(a_j,b-\hat{b}\right)=0$ לכל j, כאשר $(*,*)$ הינה המכפלה הפנימית הסטנדרטית ב$\mathbb{R}^n$. זה שקול לכפל המטריצות $a_j^t\left(b-\hat{b}\right)=0$, וזה שקול ל$a_j^t\,(b-A\hat{x})=0$ או $A^t\,(b-A\hat{x})=0$, כלומר $A^tA\hat{x}=A^tb$. לכן, כדי למצוא את $\hat{x}$ צריך לפתור את המערכת $A^tAx=A^tb$, למערכת זו יש פתרון.

דוגמה 9.4 נניח $b=\begin{pmatrix}2\\0\\11\end{pmatrix}$, $A=\begin{pmatrix}4&0\\0&2\\1&1\end{pmatrix}$, למערכת $\begin{pmatrix}4&0\\0&2\\1&1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}2\\0\\11\end{pmatrix}$ אין פתרון. נחשב:

$$A^tA=\begin{pmatrix}4&0&1\\0&2&1\end{pmatrix}\begin{pmatrix}4&0\\0&2\\1&1\end{pmatrix}=\begin{pmatrix}17&1\\1&5\end{pmatrix}$$

$$A^tb=\begin{pmatrix}4&0&1\\0&2&1\end{pmatrix}\begin{pmatrix}2\\0\\11\end{pmatrix}=\begin{pmatrix}19\\11\end{pmatrix}$$

למערכת $\begin{pmatrix}19\\11\end{pmatrix}=\begin{pmatrix}17&1\\1&5\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}$ יש פתרון, $\hat{x}=\begin{pmatrix}1\\2\end{pmatrix}$ הוקטור הקרוב ביותר.

10 היטלים

הגדרה 10.1 יהי V מרחב וקטורי מעל שדה $\mathbb{F}$, ונניח ש$V=W_1\oplus W_2$ כאשר W_1,W_2 הם שני תתי מרחב של V. נגדיר שתי העתקות $E_1\colon V\to V, E_2\colon V\to V$ באופן הבא: אם $v\in V$, נרשום אותו בצורה $v=w_1+w_2$ כאשר $w_1\in W_1, w_2\in W_2$ (כתיבה זו הינה יחידה) ואז נגדיר $E_1(v)=w_1, E_2(v)=w_2$.
ההעתקות E_i נקראות **ההטלות** של V על W_i.

טענה 10.2 תכונות של הטלות: E_i הן לינאריות, $E_i^2=E_i$, $E_1E_2=E_2E_1=0$, $I=E_1+E_2$, $\ker E_1=W_2$, $\ker E_2=W_1$, $\mathrm{Im}E_i=W_i$, וצמצום E_i לW_i היא העתקת הזהות של W_i. כמו כן, עבור בסיסים $\{w_1,\dots,w_r\}$ של W_1 ו$\{w_{r+1},\dots,w_n\}$ של W_2, נסמן $e=\{w_1,\dots,w_n\}$ ואז $[E_1]_e=\begin{pmatrix}I_r&0\\0&0\end{pmatrix}, [E_2]_e=\begin{pmatrix}0&0\\0&I_{n-r}\end{pmatrix}$.

טענה 10.3 יהי $T\colon V\to V$ העתקה לינארית כך ש$T^2=T$. אז $V=\ker T\oplus \mathrm{Im}T$, ו$T$ היא ההטלה על $\mathrm{Im}T$.

הוכחה: נוכיח תחילה ש$\ker T\cap \mathrm{Im}T=\{0\}$. יהי $v\in \ker T\cap \mathrm{Im}T$, כלומר $Tv=0$ וכמו כן קיים $u\in V$ כך ש$Tu=v$. אז $0=Tv=T^2u=Tu=v$. כעת, מתקיים $\ker T+\mathrm{Im}T\subseteq V$ וכמו כן $\dim(\ker T+\mathrm{Im}T)=\dim\ker T+\dim \mathrm{Im}T=\dim V$, ולכן $V=\ker T\oplus \mathrm{Im}T$.
כמו כן, אם $w\in \mathrm{Im}T$ אז $Tu=w$ ולכן $w=Tu=T^2u=Tw$, כלומר T פועלת כההעתקת הזהות על $\mathrm{Im}T$. ■

משפט 10.4 תכונות של הטלות כלליות: נניח ש$V=W_1\oplus\cdots\oplus W_k$, אז קיימות k העתקות לינאריות $E_1,\dots,E_k$ המוגדרות על V כך ש:

1. E_i הטלה, כלומר $E_i^2=E_i$.
2. $E_iE_j=0$ לכל $i\neq j$.
3. $I=E_1+\cdots+E_k$.
4. $\mathrm{Im}E_i=W_i$.

בכיוון ההפוך: אם $E_1,\dots,E_k$ הן k העתקות לינאריות על V המקיימות את (1), (2), (3) ונסמן $W_i=\mathrm{Im}E_i$ אז $V=W_1\oplus\cdots\oplus W_k$. **הוכחה:** נוכיח את הכיוון ההפוך (היתר נותר כתרגיל).
נראה כי $V=W_1\oplus\cdots\oplus W_k$.
תחילה נראה ש$V=W_1+\cdots+W_k$. יהי $v\in V$, אז לפי (3) מתקיים $v=Iv=E_1v+\cdots+E_kv\in W_1+\cdots+W_k$.
כעת נוכיח שהפירוק של כל $v\in V$ הינו יחיד. נרשום $v=v_1+\cdots+v_k$ כאשר $v_i\in W_i$, אז כיוון ש$v_i\in \mathrm{Im}E_i=W_i$ אז ניתן לרשום $v_i=E_iu_i$. אז לכל $1\le j\le k$

$$E_jv=E_j(v_1+\cdots+v_k)=E_j(E_1u_1+\cdots+E_ku_k)\overset{(2)}{=}E_j^2u_j\overset{(1)}{=}E_ju_j=v_j$$

אז הפירוק יחיד, ולכן הסכום ישר. ■

משפט 10.5 נניח ש$V=W_1\oplus\cdots\oplus W_k$, ותהי $T\colon V\to V$ העתקה לינארית.
יהיו E_i ההטלות על W_i, אז W_i הוא T אינווריאנטי $\iff TE_i=E_iT$.

הוכחה: $\Longrightarrow$ נניח $TE_i = E_iT$ ויהי $v \in W_i$, אז $v = E_iv$ ולכן $Tv = TE_iv = E_iTv$ כלומר $Tv \in \mathrm{Im}E_i = W_i$.
$\Longleftarrow$ נניח שW_i הוא T־אינווריאנטי ויהי $v \in V$, אז $v = E_1v + \cdots + E_kv$ ולכן $Tv = TE_1v + \cdots + TE_kv$.
כמו כן, כיוון ש$E_iv \in W_i$ ומנתון W_i הוא T אינווריאנטי, אז קיים $u_i \in W_i$ כך ש$TE_iv = E_iu_i$. לכן, $E_jTE_iv = E_jE_iu_i = \begin{cases} 0 & i \neq j \\ E_iu_i & i = j \end{cases}$. לכן

$$E_iTv = E_i\left(TE_1v + \cdots + TE_kv\right) = E_iTE_1v + \cdots + E_iTE_kv = E_iu_i = TE_iv$$

כלומר $E_iT = TE_i$. ■

משפט 10.6 תהי $T\colon V \to V$ העתקה לינארית. אם T ניתנת ללכסון ואם $c_1, \ldots, c_k$ הם הערכים העצמיים השונים של T, אז קיימות העתקות לינאריות $E_1, \ldots, E_k$ כך ש:

1. $T = c_1E_1 + \cdots + c_kE_k$.
2. $I = E_1 + \cdots + E_k$.
3. $E_iE_j = 0$ לכל $i \neq j$.
4. $E_i^2 = E_i$.
5. $\mathrm{Im}E_i$ הוא המרחב העצמי של הערך העצמי c_i.

בכיוון ההפוך: אם קיימים סקלרים $c_1, \ldots, c_k$ שונים והעתקות $E_1, \ldots, E_k$ המקיימות תכונות (1), (2), (3), אז T ניתנת ללכסון, $c_1, \ldots, c_k$ הם הערכים העצמיים השונים של T ותנאים (4), (5) מתקיימים.

הוכחה: הכיוון הראשון: נניח שT ניתנת ללכסון עם ערכים עצמיים שונים $c_1, \ldots, c_k$. יהי W_i המרחב העצמי המתאים לערך העצמי c_i, ואז $V = W_1 \oplus \cdots \oplus W_k$. יהיו $E_1, \ldots, E_k$ ההטלות על $W_1, \ldots, W_k$, ואז כפי שראינו (2), (3), (4), (5) מתקיימים. נוכיח ש(1) מתקיים. לכל $v \in V$,

$$v = E_1v + \cdots + E_kv \implies Tv = TE_1v + \cdots + TE_kv = c_1E_1v + \cdots + c_kE_kv$$

ולכן $T = c_1E_1 + \cdots + c_kE_k$.
הכיוון ההפוך: תהי $T\colon V \to V$ כך ש(1), (2), (3) מתקיימים. נכפיל את השוויון $I = E_1 + \cdots + E_k$ בE_i. כיוון ש$E_iE_j = 0$ לכל $i \neq j$, אז $E_i = E_i^2$ (כלומר תכונה (4) מתקיימת).
נכפיל את $T = c_1E_1 + \cdots + c_kE_k$ בE_i, אז נקבל $TE_i = c_iE_i^2 = c_iE_i$ כלומר $(T - c_iI)E_i = 0$. לכן אם $v \in \mathrm{Im}E_i$, אז $v \in \ker(T - c_iI)$ כי אם $v = E_iw$ אז $(T - c_iI)v = (T - c_iI)E_iw = 0$. אז הוכחנו $\ker(T - c_iI) \supseteq \mathrm{Im}E_i$.
כיוון ש$E_i \neq 0$, $\mathrm{Im}E_i \neq \{0\}$ ולכן $\ker(T - c_iI) \neq \{0\}$ כלומרc_i הוא ערך עצמי של T. כלומר, $c_1, \ldots, c_k$ הם ערכים עצמיים של T. נראה שלT אין ערכים עצמיים נוספים. יהי c ערך עצמי של T, אז קיים $v \neq 0$ כך ש

$$0 = (T - cI)v = Tv - cv \overset{(1),(2)}{=} (c_1 - c)E_1v + \cdots + (c_k - c)E_kv$$

לכל i, נפעיל את E_i על שוויון זה, ונקבל $0 = (c_i - c)E_iv$. בהינתן $0 \neq v \in V$, קיים E_i כך ש$E_iv \neq 0$ ולכן $c = c_i$ עבור i כלשהו. כלומר, $c_1, \ldots, c_k$ הם בדיוק כל הערכים העצמיים של T.
T ניתנת ללכסון, כי ראינו שכל וקטור ב$\mathrm{Im}E_i$ הוא וקטור עצמי של T ומהזהות $I = E_1 + \cdots + E_k$ נובע שכל $v \in V$ הוא צירוף לינארי של וקטורים עצמיים של T. לכן לV יש בסיס המורכב מוקטורים עצמיים של T.
נותר רק להוכיח את (5). נוכיח ש$\ker(T - c_iI) = \mathrm{Im}E_i$. הוכחנו קודם $\ker(T - c_iI) \supseteq \mathrm{Im}E_i$, ונותר להוכיח את ההכלה בכיוון השני. יהי $v \in \ker(T - c_iI)$, אז $Tv = c_iv$ וכפי שראינו $(c_1 - c_i)E_1v + \cdots + (c_k - c)E_kv = 0$. לכן, $(c_j - c_i)E_jv = 0$ לכל j. לכן $E_jv = 0$ לכל $i \neq j$. כיוון ש$v = E_1v + \cdots + E_kv$ וגם $E_jv = 0$ לכל $i \neq j$, אז $v = E_iv$ כלומר $v \in \mathrm{Im}E_i$. ■

טענה 10.7 בתנאי המשפט, $E_j = p_j(T)$ כאשר $p_j(x) = \prod_{i \neq j} \frac{x - c_i}{c_j - c_i}$ הינו **פולינום לגראנז'** המתאים לc_j.

הוכחה: נניח שתנאי המשפט מתקיימים, ונרשום $T = c_1E_1 + \cdots + c_kE_k$. יהי $g(x)$ פולינום, ואז

$$g(T) = g(c_1)E_1 + \cdots + g(c_k)E_k$$

מלינאריות, מספיק להוכיח זהות זו עבור $g(x) = x^r$ (תרגיל). מתקיים $p_j(c_i) = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$, לכן מהתרגיל מקבלים

$$p_j(T) = \delta_{1j}E_1 + \delta_{2j}E_2 + \cdots + \delta_{kj}E_k = E_j$$

■

משפט 10.8 המשפט הספקטרלי: תהי T העתקה נורמלית מעל מרחב וקטורי מרוכב, או העתקה סימטרית מעל מרחב וקטורי ממשי. יהיו c_j ($1 \leq j \leq k$) הערכים העצמיים השונים של T, יהיו W_j המרחבים העצמיים של c_j, ויהיו E_j ההטלות על W_j אז $V = W_1 \oplus \cdots \oplus W_k$, W_i מאונך ל-W_j אם $i \neq j$ ומתקיים $T = c_1E_1 + \cdots + c_kE_k$. פירוק זה נקרא **הפירוק הספקטרלי** של T, וקבוצת הערכים העצמיים $c_1, \ldots, c_k$ נקראת **הספקטרום** של T. כמו כן, $E_j = p_j(T)$ כאשר $p_j(x) = \prod_{i \neq j} \frac{x - c_i}{c_j - c_i}$ פולינום לגראנז' המתאים ל-c_j.

11 העתקות חיוביות

מוטיבציה להגדרה יהיו $x, y \in \mathbb{F}^n$ וקטורי עמודה, אז עבור המכפלה הסטנדרטית $(x,y) = y^*x = y^*Ix$. נרצה לשאול: עבור אילו מטריצות A הביטוי $(x,y) = y^*Ax$ מגדיר מכפלה פנימית?

במכפלה פנימית צריך להתקיים $(x,y) = \overline{(y,x)}$, ונשים לב $y^*A^*x = \overline{y}^t\overline{A}^t\overline{x^*}^t = \overline{x^*Ay}^t = \overline{x^*Ay} = \overline{(y,x)}$ לכן אנו מחפשים מטריצות כך ש-$y^*Ax = y^*A^*x$. התנאי $A^* = A$ יבטיח תנאי זה. כמו כן צריך להתקיים $0 \leq (x,x) = x^*Ax = x^*IAx = (Ax,x)_{ST}$ כאשר $(*,*)_{ST}$ היא המכפלה הפנימית הסטנדרטית.

הגדרה 11.1 תהי T העתקה צמודה לעצמה, כלומר $T^* = T$. מתרגיל נובע שהביטוי (Tv,v) הוא ממשי לכל $v \in V$.

נאמר ש-T **אי שלילית** אם $(Tv,v) \geq 0$ לכל $v \in V$, ונסמן זאת $T \geq 0$. נאמר ש-T **חיובית** אם $(Tv,v) > 0$ לכל $0 \neq v \in V$, ונסמן זאת $T > 0$.

טענה 11.2 העתקה צמודה לעצמה היא אי שלילית (חיובית) $\iff$ כל הערכים העצמיים שלה הם אי שליליים (חיוביים).

הוכחה: $\Longleftarrow$ נניח ש-$T \geq 0$ ונניח ש-$Tv = \lambda v$ עבור $v \neq 0$. אז

$$0 \leq (Tv,v) = \lambda (v,v)$$

$(v,v) \geq 0$, ולכן $\lambda \geq 0$.

$\Longrightarrow$ נניח ש-$T^* = T$ כך שכל הערכים העצמיים $\lambda_i \geq 0$. נראה ש-$T \geq 0$. ממשפט ל-T יש בסיס אורתונורמלי $v_1, \ldots, v_n$ המורכב מוקטורים עצמיים, כלומר $Tv_i = \lambda_i v_i$. יהי $v \in V$ ונרשום $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$, אז $Tv = \alpha_1\lambda_1 v_1 + \cdots + \alpha_n \lambda_n v_n$ ולכן

$$(Tv,v) = \left(\sum_{i=1}^{n} \alpha_i \lambda_i v_i, \sum_{i=1}^{n} \alpha_i v_i\right) = \lambda_1 \alpha_1 \overline{\alpha_1} + \cdots + \lambda_n \alpha_n \overline{\alpha_n} \geq 0$$

■

טענה 11.3 $T \geq 0 \iff T = A^*A$ עבור A העתקה לינארית.

הוכחה: $\Longrightarrow$ נניח ש-$T = A^*A$, אז $T^* = T$ וגם $(Tv,v) = (A^*Av,v) = (Av,Av) \geq 0$.

$\Longleftarrow$ נניח $T \geq 0$, וכאן נתייחס ל-T כמטריצה. אז קיימת מטריצה אוניטרית (אורתוגונלית) כך ש-$UTU^* = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ כאשר λ_i הם הערכים העצמיים של T, ולכן על פי הטענה הקודמת $\lambda_i \geq 0$. נגדיר $S = \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}$, אז $S^* = S$, ולכן U^*SU צמודה לעצמה ונגדיר $A = U^*SU$. ואז $A^* = A$ ולכן

$$A^*A = A^2 = U^*SUU^*SU = U^*S^2U = U^* \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U = T$$

■

מסקנה 11.4 אם $T \geq 0$, אז קיימת העתקה $N \geq 0$ יחידה כך ש-$N^2 = T$, ונסמן $N = \sqrt{T}$.

הוכחה: הוכחנו בטענה הקודמת שקיימת N כזו, כעת נותר להוכיח יחידות.

נפעיל את המשפט הספקטרלי על T. נניח ש-$T = c_1E_1 + \cdots + c_kE_k$, כאשר $I = E_1 + \cdots + E_k$ ו-$E_iE_j = 0$ לכל $i \neq j$ כאשר $c_1, \ldots, c_k$ הם הערכים העצמיים של T. תהי $P \geq 0$ כך ש-$P^2 = T$, ונסמן $P = d_1F_1 + \cdots + d_rF_r$

כאשר $I = F_1 + \cdots + F_r$ ו־$F_iF_j = 0$ לכל $i \neq j$ כאשר $d_1, \ldots, d_r$ הם הערכים העצמיים של P. מתקיים $T = P^2 = d_1^2F_1 + \cdots + d_r^2F_r$ ולכן מיחידות הפירוק הספקטרלי בהכרח d_i^2 הם הערכים העצמיים של T ו־F_i הם ההטלות על המרחבים העצמיים של T, ולכן $d_i^2 = c_i$ כלומר $P = \sqrt{c_1}E_1 + \cdots + \sqrt{c_k}E_k$ כאשר $I = E_1 + \cdots + E_k$, ו־$E_iE_j = 0$ לכל $i \neq j$. לכן P יחידה (נקבעת באופן יחיד על ידי T). ■

משפט 11.5 פירוק פולרי: יהי V מרחב וקטורי עם מכפלה פנימית, ותהי $T\colon V \to V$ העתקה לינארית. אז קיימות העתקה אוניטרית U והעתקה $N \geq 0$ כך ש־$T = UN$. כמו כן N נקבעת באופן יחיד, ואם T הפיכה אז גם U נקבעת באופן יחיד.

הוכחה: תחילה נראה ש־N נקבעת באופן יחיד. נניח ש־$T = UN$ כאשר U אוניטרית ו־$N \geq 0$. אז $T^* = (UN)^* = N^*U^* = NU^*$, ולכן $T^*T = NU^*UN = N^2$. כיוון ש־$T^*T \geq 0$ בהכרח $N = \sqrt{T^*T}$ כלומר נקבעת באופן יחיד.

כעת נוכיח שקיים פירוק. לכל $v \in V$ מתקיים $(Nv, Nv) = \left(v, N^2v\right) = (v, T^*Tv) = (Tv, Tv)$. לכן $\ker T = \ker N$, ובפרט N הפיכה $\iff$ T הפיכה.

אם T הפיכה, אז N הפיכה ולכן בהכרח $U = TN^{-1}$ (נקבעת באופן יחיד). כמו כן U זו אוניטרית שכן

$$U^* = \left(TN^{-1}\right)^* = \left(N^{-1}\right)^* T^* = (N^*)^{-1} T^* = N^{-1}T^*$$

ולכן $UU^* = TN^{-1}N^{-1}T^* = T\left(N^2\right)^{-1}T^* = T\left(T^*T\right)^{-1}T^* = TT^{-1}\left(T^*\right)^{-1}T^* = I$.

כעת נוכיח את המקרה הכללי. לצורך נוחות נניח ש־T, N הם מטריצות מסדר n. נזכיר כי $\ker T = \ker N$. כיוון ש־N צמודה לעצמה ל־$\mathbb{F}^n$ יש בסיס אורתונורמלי $v_1, \ldots, v_n$ המורכב מוקטורים עצמיים של N, נניח $Nv_i = \lambda_i v_i$. נניח ש־$l = \mathrm{rank}T = \mathrm{rank}N$ (הם שוויון משוויון ה־$\ker$) כאשר $l \leq n$, אז יש בדיוק l ערכים עצמיים שונים מ־0 וכיוון ש־$\lambda_i \geq 0$ לכל $1 \leq i \leq n$ על ידי סידור מתאים ניתן להניח ש־$\lambda_i > 0$ לכל $1 \leq i \leq l$.

הקבוצה $\left\{\frac{1}{\lambda_1}Tv_1, \ldots, \frac{1}{\lambda_l}Tv_l\right\}$ היא אורתונורמלית, שכן

$$\left(\frac{1}{\lambda_i}Tv_i, \frac{1}{\lambda_j}Tv_j\right) = \frac{1}{\lambda_i\lambda_j}\left(Tv_i, Tv_j\right) = \frac{1}{\lambda_i\lambda_j}\left(v_i, T^*Tv_j\right) = \frac{1}{\lambda_i\lambda_j}\left(v_i, N^2v_j\right) = \frac{\lambda_j^2}{\lambda_i\lambda_j}\left(v_i, v_j\right) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

ולכן ניתן להשלים אותה לבסיס אורתונורמלי של $\mathbb{F}^n$, נניח $B = \left\{\frac{1}{\lambda_1}Tv_1, \ldots, \frac{1}{\lambda_l}Tv_l, u_{l+1}, \ldots, u_n\right\}$. תהי U_1 המטריצה שעמודותיה הם הוקטורים בבסיס B, ותהי U_2 המטריצה ששורותיה הם הוקטורים $v_1^*, \ldots, v_n^*$. אז U_1, U_2 מטריצות אוניטריות, ולכן $U = U_1U_2$ גם כן אוניטרית.

נותר רק לבדוק $T = UN$, והיות ש־$v_1, \ldots, v_n$ בסיס מספיק להוכיח $Tv_i = UNv_i$ לכל $1 \leq i \leq n$:

$$UNv_i = \begin{cases} \lambda_i Uv_i & 1 \leq i \leq l \\ 0 & l < i \leq n \end{cases} = \begin{cases} \lambda_i U_1U_2v_i & 1 \leq i \leq l \\ 0 & l < i \leq n \end{cases} = \begin{cases} \lambda_i U_1e_i & 1 \leq i \leq l \\ 0 & l < i \leq n \end{cases} = \begin{cases} \lambda_i \frac{1}{\lambda_i}Tv_i & 1 \leq i \leq l \\ 0 & l < i \leq n \end{cases} = Tv_i$$

וסיימנו. ■

חלק IV
תבניות בילינאריות

12 הגדרות

הגדרה 12.1 יהי V מרחב וקטורי מעל שדה $\mathbb{F}$ (לא בהכרח $\mathbb{R}$ או $\mathbb{C}$). **תבנית בילינארית** הינה העתקה $f: V^2 \to \mathbb{F}$, המקיימת את התכונות הבאות:

1. לינאריות ברכיב הראשון: $f\left(c_1 v_1 + c_2 v_2, u\right) = c_1 f\left(v_1, u\right) + c_2 f\left(v_2, u\right)$
2. לינאריות ברכיב השני: $f\left(v, c_1 u_1 + c_2 u_2\right) = c_1 f\left(v, u_1\right) + c_2 f\left(v, u_2\right)$

נסמן את אוסף כל התבניות הבילינאריות ב$L\left(V, \mathbb{F}\right)$.

דוגמה 12.2 עבור $V = M_{m\times n}\left(\mathbb{F}\right)$, תהי $A \in M_m\left(\mathbb{F}\right)$ ונגדיר $f_A\left(X, Y\right) = \mathrm{tr}\left(X^t AY\right)$. זו תבנית בילינארית, כי

$$f_A\left(c_1 X + c_2 Z, Y\right) = \mathrm{tr}\left(\left(c_1 X + c_2 Z\right)^t AY\right) = \mathrm{tr}\left(\left(c_1 X^t\right) AY\right) + \mathrm{tr}\left(\left(c_2 Z^t\right) AY\right) = c_1 f_A\left(X, Y\right) + c_2 f_A\left(Z, Y\right)$$

ובאותו האופן ברכיב השני. אם $n = 1$, אז $V = \mathbb{F}^m$ ואז $f_A\left(x, y\right) = x^t Ay$.

הגדרה 12.3 אם $f, g \in L\left(V, \mathbb{F}\right)$ אז נגדיר $\left(f + g\right)\left(u, v\right) = f\left(u, v\right) + g\left(u, v\right)$. אם $f \in L\left(V, \mathbb{F}\right), c \in \mathbb{F}$ אז נגדיר $\left(cf\right)\left(u, v\right) = cf\left(u, v\right)$.

משפט 12.4 $L\left(V, \mathbb{F}\right)$ עם הפעולות הנ"ל מגדיר מרחב וקטורי מעל $\mathbb{F}$. כמו כן, $L\left(V, \mathbb{F}\right)$ איזומורפי ל$M_n\left(\mathbb{F}\right)$ ובפרט $\dim L\left(V, \mathbb{F}\right) = n^2$.

הוכחה: יהי $B = \left\{v_1, \ldots, v_n\right\}$ בסיס לV. בהינתן $u, v \in V$ נרשום $v = \sum_{i=1}^n \alpha_i v_i, u = \sum_{i=1}^n \beta_i v_i$ ואז

$$f\left(v, u\right) = f\left(\sum_{i=1}^n \alpha_i v_i, \sum_{i=1}^n \beta_i v_i\right) = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \beta_j f\left(v_i, v_j\right)$$

נגדיר $A = \left(a_{ij}\right) \in M_n\left(\mathbb{F}\right)$ כאשר $a_{ij} = f\left(v_i, v_j\right)$, ואז $f\left(v, u\right) = \left[v\right]_B^t A \left[u\right]_B$. לכל $f \in L\left(V, \mathbb{F}\right)$ התאמנו $A \in M_n\left(\mathbb{F}\right)$, הנקראת **המטריצה המייצגת את** f **לפי בסיס** B. ∎

הגדרה 12.5 A, B מטריצות מסדר n יקראו **מטריצות חופפות** אם קיימת P הפיכה כך ש$A = P^t BP$.

טענה 12.6 מטריצות A, B הן חופפות $\Longleftrightarrow$ הן מייצגות את אותה התבנית f.

הוכחה: $\Longleftarrow$ יהי C בסיס כלשהו ונניח כי A מייצגת את התבנית f לפי C. A, B חופפות אז קיימת מטריצה הפיכה P כך ש$B = P^t AP$, P הפיכה אז קיים בסיס C' כלשהו כך שP היא מטריצת המעבר מC' לC, כלומר $P\left[u\right]_{C'} = \left[u\right]_C$ לכל $u \in V$. אז מתקיים לכל u, v

$$f\left(u, v\right) = \left[u\right]_C^t A \left[v\right]_C = \left(P\left[u\right]_{C'}\right)^t A \left(P\left[v\right]_{C'}\right) = \left[u\right]_{C'}^t P^t AP \left[v\right]_{C'} = \left[u\right]_{C'}^t B \left[v\right]_{C'}$$

כלומר גם B מייצגת את f, כרצוי.

$\Longrightarrow$ נניח כי A, B מייצגות את אותה התבנית, נניח A מייצגת את f בבסיס C וB מייצגת את f בבסיס C'. אז מתקיים

$$f\left(u, v\right) = \left[u\right]_C^t A \left[v\right]_C$$
$$f\left(u, v\right) = \left[u\right]_{C'}^t B \left[v\right]_{C'}$$

נגדיר P להיות מטריצת המעבר מC' לC, כלומר $P\left[u\right]_{C'} = \left[u\right]_C$ לכל $u \in V$. אז מתקיים לכל u, v

$$\left[u\right]_{C'}^t B \left[v\right]_{C'} = f\left(u, v\right) = \left[u\right]_C^t A \left[v\right]_C = \left(P\left[u\right]_{C'}\right)^t A \left(P\left[v\right]_{C'}\right) = \left[u\right]_{C'}^t P^t AP \left[v\right]_{C'}$$

אז $P^t AP, B$ מייצגות את f לפי אותו הבסיס ולכן $B = P^t AP$ כלומר A, B חופפות. ∎

הגדרה 12.7 תבנית בילינארית f תקרא **לא מנוונת** אם $\mathrm{rank}A = n = \dim V$, כאשר A מטריצה כלשהי המייצגת את f לפי בסיס כלשהו. כיוון ש$\mathrm{rank}\left(P^t AP\right) = \mathrm{rank}A$ לכל P הפיכה, זו הגדרה טובה כלומר אינה תלויה בבחירת הבסיס. לכן, נסמן $\mathrm{rank}f = \mathrm{rank}A$ עבור A כלשהי המייצגת את f.

טענה 12.8 f לא מנוונת $\iff$ לכל $0 \neq u \in V$ קיים $v \in V$ כך ש$f\left(u, v\right) \neq 0$.

הוכחה: נניח ש$f\left(u, v\right) = \left[u\right]_B^t A \left[v\right]_B$ עבור בסיס B מסוים.

$\Leftarrow$ אם f אינה מנוונת, אז $\mathrm{rank}A = n$. כלומר, A היא מטריצה הפיכה ואם נתייחס לA כהעתקה $A\colon \mathbb{F}^n \to \mathbb{F}^n$ אז A העתקה חח"ע ועל. יהי $0 \neq u \in V$, ואז $\left[u\right]_B \neq 0$. יהי $w \in V$ כלשהו כך ש$\left[u\right]_B^t \left[w\right]_B \neq 0$. כיוון ש$A\colon \mathbb{F}^n \to \mathbb{F}^n$ היא על, קיים $0 \neq v \in V$ כך ש$A\left[v\right]_B = \left[w\right]_B$. ואז, $f\left(u, v\right) = \left[u\right]_B^t A \left[v\right]_B \neq 0$.

$\Longrightarrow$ נניח שלכל $0 \neq u \in V$ קיים $v \in V$ כך ש$f\left(u, v\right) \neq 0$, כלומר $\left[u\right]_B^t A \left[v\right]_B \neq 0$. נניח בשלילה שלמערכת $x^t A = 0$ יש פתרון לא טריויאלי, אז קיים $u \in V$ כך ש$\left[u\right]_B^T A = 0$. אז לכל $v \in V$ מתקיים $\left[u\right]_B^t A \left[v\right]_B = 0$, בסתירה לנתון. אז למערכת $x^t A = 0$ יש פתרון יחיד, כלומר A הפיכה ולכן f אינה מנוונת. ■

13 תבניות בילינאריות סימטריות ואנטי־סימטריות

תבניות בילינאריות סימטריות

הגדרה 13.1 תבנית בילינארית תקרא **סימטרית** אם לכל $u, v \in V$ מתקיים $f\left(u, v\right) = f\left(v, u\right)$.

טענה 13.2 f סימטרית $\iff$ המטריצה המייצגת אותה לפי בסיס כלשהו היא מטריצה סימטרית.

הוכחה: נניח ש$f\left(x, y\right) = x^t Ay$ סימטרית, כלומר $f\left(x, y\right) = f\left(y, x\right)$. לכן, $x^t Ay = y^t Ax$ לכל $x, y \in \mathbb{F}^n$.

כיוון ש$x^t Ay \in \mathbb{F}$, אז $y^t A^t x = \left(x^t Ay\right)^t = x^t Ay$ ולכן $x^t A^t y = x^t Ay$ לכל x, y, כלומר $A = A^t$. ■

טענה 13.3 נניח ש$\mathbb{F} = \mathbb{R}$. תהי f תבנית בילינארית סימטרית אז קיים בסיס של V שבו f מיוצגת על ידי מטריצה אלכסונית.

הוכחה: כל מטריצה סימטרית ניתנת ללכסון על ידי מטריצה אורתוגונלית, כלומר מתקיים $P^t AP = P^{-1}AP$. ■

הערה 13.4 נוכיח זאת בקרוב מעל כל שדה, על ידי אלגוריתם.

משפט 13.5 משפט ההתמדה של סילבסטר: נניח $\mathbb{F} = \mathbb{R}$ ותהי f תבנית סימטרית כך ש$\mathrm{rank}f = r \leq n = \dim V$. אז קיים בסיס $B = \left\{v_1, \ldots, v_n\right\}$ של V כך שהמטריצה המייצגת את f לפי B היא אלכסונית עם איברי אלכסון $0, \pm 1$. מספר ה1ים (ולכן מספר ה-1ים) אינו תלוי בבחירת הבסיס.

הוכחה: קיים בסיס $B' = \left\{u_1, \ldots, u_n\right\}$ כך שהמטריצה המייצגת את f לפיו אלכסונית, כלומר $f\left(u_i, u_j\right) = 0$ לכל $i \neq j$, $f\left(u_i, u_i\right) \neq 0$ כאשר $1 \leq i \leq r$ ו$f\left(u_i, u_i\right) = 0$ כאשר $i > r$. נגדיר בסיס $v_j = \begin{cases} \frac{1}{\sqrt{\left|f\left(u_j, u_j\right)\right|}} u_j & 1 \leq j \leq r \\ u_j & r < j \leq n \end{cases}$,
אז במטריצה המייצגת את f לפי בסיס זה יש רק $0, \pm 1$.

כעת נוכיח יחידות. יהי $\left\{u_1, \ldots, u_n\right\}$ בסיס שבו f מיוצגת על ידי מטריצה אלכסונית בעלת P 1ים וN -1ים, כאשר $P + N = r$. יהי $\left\{w_1, \ldots, w_n\right\}$ בסיס שבו יש P' 1ים וN' -1ים, כאשר $P' + N' = r$. נראה כי $P = P'$. נסמן $U = \mathrm{Sp}\left\{u_1, \ldots, u_p\right\}, W = \mathrm{Sp}\left\{w_{p'+1}, \ldots, w_n\right\}$. אם $0 \neq v \in U$, ונניח ש$v = \alpha_1 u_1 + \cdots + \alpha_p u_p$ אז $f\left(v, v\right) = \sum_{i=1}^{p} \alpha_i^2 f\left(u_i, u_i\right) > 0$. באותו האופן, אם $v \in W$ אז $f\left(v, v\right) \leq 0$. לכן $U \cap W = \left\{0\right\}$. אז

$$n = \dim V \geq \dim\left(U + W\right) = \dim U + \dim W - \dim\left(U \cap W\right) = p + \left(n - p'\right)$$

לכן $p \leq p'$, באותו האופן ניתן להוכיח גם $p \geq p'$ ובסך הכל נובע $p = p'$. ■

הגדרה 13.6 המספר $P - N$ נקרא ה**סיגנטורה** של f.

אלגוריתם 13.7 לכסון תבנית בילינארית סימטרית: (מעל כל שדה, מלבד שדה בעל מציין 2) תהי A מטריצה כלשהי ותהי E מטריצה אלמנטרית. אז EA היא המטריצה המתקבלת מA על ידי פעולת השורה האלמנטרית המתאימה למטריצה האלמנטרית. באופן דומה, AE^t היא המטריצה המתקבלת מA על ידי פעולת עמודה מתאימה. אם A סימטרית וE מטריצה אלמנטרית, אז EAE^t היא סימטרית.

נתונה מטריצה סימטרית $A = \left(a_{ij}\right)$.

1. אם $a_{11} \neq 0$, נחפוף ע״י מטריצה E מתאימה את A למטריצה $A' = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}$ ו־A_1 סימטרית.

מכאן, ההמשך באינדוקציה.

2. אם $a_{11} = 0$ ו־$a_{ii} \neq 0$ עבור $i > 1$, נבצע את הפעולות האלמנטריות $R_1 \leftrightarrow R_i, C_1 \leftrightarrow C_i$ ואז נקבל מטריצה שבה $a_{11} \neq 0$ ואנו חוזרים למקרה הראשון.

3. אם $a_{ii} = 0$ לכל i, בהנחה ש־$A \neq 0$ ו־$a_{ij} \neq 0$ עבור i, j כלשהם נבצע $R_i \to R_i + R_j, C \to C_i + C_j$ ואז נקבל $2a_{ij} \neq 0$ במקום ה־i, i במטריצה החדשה, ואנו חוזרים למקרה השני.

דוגמה 13.8 $A = \begin{pmatrix} 1 & -3 & 2 \\ -3 & 7 & -5 \\ 2 & -5 & 8 \end{pmatrix}$, נפעיל את האלגוריתם:

$$\left(\begin{array}{ccc|ccc} 1 & -3 & 2 & 1 & 0 & 0 \\ -3 & 7 & -5 & 0 & 1 & 0 \\ 2 & -5 & 8 & 0 & 0 & 1 \end{array}\right) \xrightarrow[R_3 \to -2R_1 + R_3]{R_2 \to 3R_1 + R_2} \left(\begin{array}{ccc|ccc} 1 & -3 & 2 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 1 & 4 & -2 & 0 & 1 \end{array}\right) \xrightarrow[C_3 \to -2C_1 + C_2]{C_2 \to 3C_2 + C_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 1 & 4 & -2 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{R_3 \to R_2 + 2R_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & 3 & 1 & 0 \\ 0 & 0 & 9 & -1 & 1 & 2 \end{array}\right) \xrightarrow{C_3 \to C_2 + 2C_3} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 0 & 3 & 1 & 0 \\ 0 & 0 & 18 & -1 & 1 & 2 \end{array}\right)$$

אם נגדיר $P = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -1 & 1 & 2 \end{pmatrix}$ אז מתקיים $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 18 \end{pmatrix} = \begin{pmatrix} 1 & 3 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & -3 & 2 \\ -3 & 7 & -5 \\ 2 & -5 & 8 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -1 & 1 & 2 \end{pmatrix} = PAP^t$.

משפט 13.9 תהי f תבנית סימטרית המוגדרת על V מרחב וקטורי מעל $\mathbb{C}$. אז קיים בסיס $B = \{v_1, \ldots, v_n\}$ כך שהמטריצה המייצגת את f היא אלכסונית, ומתקיים $f(v_j, v_j) = \begin{cases} 1 & 1 \leq j \leq r \\ 0 & j > r \end{cases}$ כאשר $r = \mathrm{rank} f$.

תבניות בילינאריות אנטי־סימטריות

הגדרה 13.10 תבנית $f\colon V \times V \to \mathbb{F}$ תקרא **אנטי־סימטרית** אם $f(u,v) = -f(v,u)$ לכל $u, v \in V$.

הערה 13.11 כל תבנית ניתן לרשום כסכום של תבנית סימטרית ותבנית אנטי־סימטרית, על ידי

$$g(u,v) = \frac{1}{2}\left(g(u,v) + g(v,u)\right) + \frac{1}{2}\left(g(u,v) - g(v,u)\right)$$

משפט 13.12 תהי f תבנית אנטי־סימטרית המוגדרת על V. נסמן $D = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, אז ל־$V$ יש בסיס שבו f מיוצגת על ידי המטריצה $\begin{pmatrix} D & & & \\ & D & & \\ & & \ddots & \\ & & & 0 \end{pmatrix}$, כאשר מספר ה־$D$ים הוא $\frac{1}{2}\mathrm{rank} f$. בפרט $\mathrm{rank} f$ הוא זוגי, ומספר האפסים הוא $n - \mathrm{rank} f$.

הוכחה: אם $f = 0$, אז הטענה ברורה. אם $\dim V = 1$, אז לכל $u \in V$, $f(c_1 u, c_2 u) = c_1 c_2 f(u,u) = 0$ כי מאנטי־סימטריות $f(u,u) = -f(u,u)$ ולכן $f(u,u) = 0$.

נניח $f \neq 0$ ו־$\dim V > 1$. כיוון ש־$f \neq 0$, קיימים $u_1, u_2 \in V$ כך ש־$f(u_1, u_2) \neq 0$. על ידי כפל מתאים, ניתן להניח $f(u_1, u_2) = 1$. נסמן $U = \mathrm{Sp}\{u_1, u_2\}$, אז $\dim U = 2$ כי אם $u_2 = cu_1$ אז $f(u_1, u_2) = f(u_1, cu_1) = cf(u_1, u_1) = 0$. מהקשר $f(u_1, u_2) = 1$, נובע ש־$f(u_2, u_1) = -1$. לכן, צמצום f ל־U מיוצג על ידי המטריצה $D = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ בבסיס $\{u_1, u_2\}$.

נגדיר $W = \{w \in V \mid \forall u \in U. f(w,u) = 0\}$, ונראה ש־$V = U \oplus W$.

תחילה נוכיח $U\cap W=\{0\}$. יהי $v\in U\cap W$, $v\in U$ אז ניתן לרשום $v=au_1+bu_2$, וגם $v\in W$ אז

$$0=f\left(v,u_1\right)=f\left(au_1+bu_2,u_1\right)=af\left(u_1,u_1\right)+bf\left(u_2,u_1\right)=-b$$
$$0=f\left(v,u_2\right)=f\left(au_1+bu_2,u_2\right)=af\left(u_1,u_2\right)+bf\left(u_2,u_2\right)=a$$

ולכן $a=b=0$ כלומר $v=0$.

כעת נוכיח $U+W=V$. לכל $v\in V$ נגדיר $u=f\left(v,u_2\right)u_1-f\left(v,u_1\right)u_2\in U$, $w=v-u$. נראה ש־$w\in W$, כלומר $f\left(w,u\right)=0$ לכל $u\in U$. מספיק להוכיח $f\left(w,u_1\right)=0, f\left(w,u_2\right)=0$.

$$\begin{aligned}
f\left(w,u_1\right)&=f\left(v-u,u_1\right)=f\left(v,u_1\right)-f\left(u,u_1\right)=f\left(v,u_1\right)-f\left(f\left(v,u_2\right)u_1-f\left(v,u_1\right)u_2,u_1\right)\\
&=f\left(v,u_1\right)-f\left(v,u_2\right)f\left(u_1,u_1\right)+f\left(v,u_1\right)f\left(u_2,u_1\right)=f\left(v,u_1\right)-0+f\left(v,u_1\right)\cdot(-1)=0\\
f\left(w,u_2\right)&=f\left(v-u,u_2\right)=f\left(v,u_2\right)-f\left(u,u_2\right)=f\left(v,u_2\right)-f\left(f\left(v,u_2\right)u_1-f\left(v,u_1\right)u_2,u_2\right)\\
&=f\left(v,u_2\right)-f\left(v,u_2\right)f\left(u_1,u_2\right)+f\left(v,u_1\right)f\left(u_2,u_2\right)=f\left(v,u_2\right)-f\left(v,u_2\right)\cdot 1+0=0
\end{aligned}$$

מכאן $V=U\oplus W$, ונפעיל אינדוקציה על צמצום f ל־W. צמצום f ל־W היא תבנית אנטי־סימטרית, ולכן קיים בסיס בו $f|_W$ מיוצגת על ידי $\begin{pmatrix}D&&&\\&D&&\\&&\ddots&\\&&&0\end{pmatrix}$. ■

חלק V
שימושים

14 מיון שניוניות

הגדרה 14.1 ביטוי מהצורה $a+2bx+2cy+dx^2+2exy+fy^2$ נקרא **שניונית** ב-$\mathbb{R}^2$. כאן $a,b,c,d,e,f\in\mathbb{R}$, ודורשים d,e,f לא כולם 0. ניתן לרשום שניונית בצורה

$$a+2\begin{pmatrix}x & y\end{pmatrix}\begin{pmatrix}b\\c\end{pmatrix}+\begin{pmatrix}x & y\end{pmatrix}\begin{pmatrix}d & e\\e & f\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=0$$

המטריצה $A_1=\begin{pmatrix}d & e\\e & f\end{pmatrix}$ נקראת **המטריצה המצומצמת** של השניונית. מהדרישה ש-d,e,f אינם כולם 0, נובע $A_1\neq 0$. ניתן לרשום גם בצורה

$$\begin{pmatrix}1 & x & y\end{pmatrix}\begin{pmatrix}a & b & c\\b & d & e\\c & e & f\end{pmatrix}\begin{pmatrix}1\\x\\y\end{pmatrix}=0$$

המטריצה $A=\begin{pmatrix}a & b & c\\b & d & e\\c & e & f\end{pmatrix}$ נקראת **המטריצה המורחבת** של השניונית.

העתקות שאינן משנות את השניונית הן סיבובים, שיקופים והזזות. נפעיל העתקות אורתוגונליות והזזות על השניוניות.

הגדרה 14.2 תהי T_0 העתקה אורתוגונלית, נגדיר העתקה $T\colon\mathbb{R}^2\to\mathbb{R}^2$ על ידי $Tv=T_0v+v_0$, כאשר $v,v_0\in\mathbb{R}^2$. T כזו נקראת **העתקה אורתוגונלית מוכללת**, ונדגיש שהעתקות אלו אינן בהכרח לינאריות.

טענה 14.3 אם T,S העתקות אורתוגונליות מוכללות, אז $S\circ T$ ו-T^{-1} הן העתקות אורתוגונליות מוכללות.

הוכחה: נניח ש-$Tv=T_0v+v_0, Su=S_0u+u_0$. נחשב:

$$(S\circ T)\,v=S\,(Tv)=S\,(T_0v+v_0)=S_0\,(T_0v+v_0)+u_0=S_0T_0v+(S_0v_0+u_0)$$

הרכבה של העתקות אורתוגונליות היא אורתוגונלית, ולכן $S\circ T$ העתקה אורתוגונלית מוכללת.
בהינתן T אורתוגונלית מוכללת, נגדיר $T'v=T_0^{-1}v-T_0^{-1}v_0$ ונקבל

$$(T\circ T')\,(v)=T\,(T'v)=T\left(T_0^{-1}v-T_0^{-1}v_0\right)=T_0\left(T_0^{-1}v-T_0^{-1}v_0\right)+v_0=v-v_0+v_0=v$$

ולכן $T\circ T'=I$, כלומר $T^{-1}=T'$ וכמו כן T' העתקה אורתוגונלית מוכללת כי T_0^{-1} אורתוגונלית. ■

הגדרה 14.4 שתי שניוניות X,Y הן **חופפות** אם קיימת העתקה אורתוגונלית מוכללת T כך ש-$TX=Y$.

טענה 14.5 תהי T העתקה אורתוגונלית מוכללת, ונניח $Y=TX$. אז המטריצה המורחבת של והמטריצה המצומצמת של Y חופפות לאלו של X. יתר על כן, המטריצות המצומצמות של X,Y חופפות על ידי מטריצות אורתוגונליות.

הוכחה: במטריצות, $\begin{pmatrix}x'\\y'\end{pmatrix}=T\begin{pmatrix}x\\y\end{pmatrix}=P_1\begin{pmatrix}x\\y\end{pmatrix}+\begin{pmatrix}x_0\\y_0\end{pmatrix}$ כאשר P_1 אורתוגונלית. מהנ"ל נקבל – $\begin{pmatrix}x\\y\end{pmatrix}=P_1^{-1}\begin{pmatrix}x'\\y'\end{pmatrix}$ $P_1^{-1}\begin{pmatrix}x_0\\y_0\end{pmatrix}=Q_1\begin{pmatrix}x'\\y'\end{pmatrix}+\begin{pmatrix}u\\v\end{pmatrix}$ כאשר $Q_1=P_1^{-1}$ הינה אורתוגונלית. אם נגדיר $Q=\begin{pmatrix}1 & 0 & 0\\u & & \\v & & Q_1\end{pmatrix}$ אז

$$\begin{pmatrix}1\\x\\y\end{pmatrix}=Q\begin{pmatrix}1\\x'\\y'\end{pmatrix}=\begin{pmatrix}1 & 0 & 0\\u & & \\v & & Q_1\end{pmatrix}\begin{pmatrix}1\\x'\\y'\end{pmatrix}.$$

נניח ש-A היא המטריצה המורחבת של X, אז

$$\begin{aligned}\begin{pmatrix}1 & x & y\end{pmatrix} A \begin{pmatrix}1\\x\\y\end{pmatrix} &= \left[\begin{pmatrix}1 & 0 \quad 0\\ u & \\ v & Q_1\end{pmatrix}\begin{pmatrix}1\\x'\\y'\end{pmatrix}\right]^t A \begin{pmatrix}1 & 0 \quad 0\\ u & \\ v & Q_1\end{pmatrix}\begin{pmatrix}1\\x'\\y'\end{pmatrix}\\ &= \begin{pmatrix}1 & x' & y'\end{pmatrix}\begin{pmatrix}1 & 0 \quad 0\\ u & \\ v & Q_1\end{pmatrix}^t A \begin{pmatrix}1 & 0 \quad 0\\ u & \\ v & Q_1\end{pmatrix}\begin{pmatrix}1\\x'\\y'\end{pmatrix}\end{aligned}$$

ולכן $\begin{pmatrix}1 & 0 \quad 0\\ u & \\ v & Q_1\end{pmatrix}^t A \begin{pmatrix}1 & 0 \quad 0\\ u & \\ v & Q_1\end{pmatrix}$ היא המטריצה המורחבת של Y. לכן המטריצות המורחבות חופפות.

כמו כן המטריצה המצומצמת של Y היא $Q_1^t A_1 Q_1$ ולכן היא חופפת ל-A_1 על ידי מטריצה אורתוגונלית כדרוש. ■

מסקנה 14.6 הגדלים הבאים נשמרים תחת העתקות אורתוגונליות מוכללות: הדרגה והסיגנטורה של המטריצה המורחבת והמצומצמת, והדטרמיננטה של המטריצה המצומצמת.

נחשב: תהי X שניונית נתונה, המטריצה המצומצמת A_1 היא סימטרית ולכן קיימת מטריצה אורתוגונלית Q כך ש-$Q^t A_1 Q = Q^{-1} AQ$ היא מטריצה אלכסונית. כיוון ש-$A_1 \neq 0$, ניתן להניח $A_1 = \begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix}$ כאשר לפחות אחד מבין λ_1, λ_2 אינו 0. בלי הגבלת הכלליות, ניתן להניח ש-$\lambda_1 > 0$. בסך הכל, ניתן להניח שהשניונית נתונה על ידי

$$a + 2bx + 2cy + \lambda_1 x^2 + \lambda_2 y^2 = 0$$

כאשר $\lambda_1 > 0$. כעת נבצע הזזה $(x', y') = (x, y) + (x_0, y_0)$ או $(x, y) = (x' - x_0, y' - y_0)$ כאשר x_0, y_0 יקבעו בהמשך. נקבל

$$a' + 2\left(b - \lambda_1 x_0\right) x' + 2\left(c - \lambda_2 y_0\right) y' + \lambda_1 x'^2 + \lambda_2 y'^2 = 0$$

מנתון $\lambda_1 \neq 0$ ולכן ניתן לאפס את המקדם של x' על ידי $x_0 = \frac{b}{\lambda_1}$. אם גם $\lambda_2 \neq 0$, אז ניתן לאפס את המקדם של y'. קיבלנו את האפשרויות הבאות:

1. אם $\lambda_1 > 0, \lambda_2 \neq 0$ אז מקבלים את הביטוי $\lambda_1 x^2 + \lambda_2 y^2 = a$.
2. אם $\lambda_1 > 0, \lambda_2 = 0, c \neq 0$ אז מקבלים את הביטוי $\lambda_1 x^2 + 2cy = a$. כאן על ידי הזזה נוספת ב-y ניתן לרשום $\lambda_1 x^2 + 2cy = 0$.
3. אם $\lambda_1 > 0, \lambda_2 = 0, c = 0$ אז מקבלים את הביטוי $\lambda_1 x^2 = a$.

על ידי סידור מתאים, ניתן לרשום

1. $\frac{x^2}{a^2} \pm \frac{y^2}{b^2} = m$ כאשר $m = 0, \pm 1$.
2. $\frac{x^2}{a^2} + 2cy = 0$.
3. $x^2 = a$.

בסך הכל, סוגי השניוניות הם

1. $y^2 = 4px$ או $x^2 = 4qy$ - פרבולות.
2. $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ - אליפסות.
3. $\frac{y^2}{a^2} - \frac{x^2}{b^2} = \pm 1$ - היפרבולות.
4. $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$ - ישרים מצטלבים.
5. $x^2 = a$ - ישר, ישרים מקבילים או קבוצה ריקה.

דוגמה 14.7 עבור השניונית $8+8x+16y-x^2-10xy-y^2$ מתקיים $a=8, b=4, c=8, e=-5, d=f=-1$ ולכן ניתן לרשום בצורה $\begin{pmatrix}1 & x & y\end{pmatrix}\begin{pmatrix}8 & 4 & 8\\ 4 & -1 & -5\\ 8 & -5 & -1\end{pmatrix}\begin{pmatrix}1\\ x\\ y\end{pmatrix}=0$, והמטריצה המצומצמת $A_1=\begin{pmatrix}-1 & -5\\ -5 & -1\end{pmatrix}$.

כעת נלכסן את A_1 אורתוגונלית. נחשב את הפולינום האופייני $\Delta_A(x)=(x-4)(x-6)$ ולכן הערכים העצמיים הם $\lambda_1=4, \lambda_2=-6$ והוקטורים העצמיים המתאימים הם $v_2=\begin{pmatrix}1\\ 1\end{pmatrix}, v_1=\begin{pmatrix}1\\ -1\end{pmatrix}$. אז $Q_1=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}$. אז

$$\frac{1}{\sqrt{2}}\begin{pmatrix}\sqrt{2} & & \\ & 1 & 1\\ & 1 & -1\end{pmatrix}\begin{pmatrix}8 & 4 & 8\\ 4 & -1 & -5\\ 8 & -5 & -1\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}\sqrt{2} & & \\ & 1 & 1\\ & 1 & -1\end{pmatrix}=\begin{pmatrix}8 & 6\sqrt{2} & -2\sqrt{2}\\ 6\sqrt{2} & -6 & 0\\ -2\sqrt{2} & 0 & 4\end{pmatrix}=M$$

נחפש $L=\begin{pmatrix}1 & & \\ u & 1 & \\ v & & 1\end{pmatrix}$ כך שL^tML אלכסונית (ניתן גם להשתמש באלגוריתם למציאת מטריצה חופפת). זה מתקיים עבור $L=\begin{pmatrix}1 & 0 & 0\\ \sqrt{2} & 1 & 0\\ \frac{\sqrt{2}}{2} & 0 & 1\end{pmatrix}$, ובמקרה זה $L^tML=\begin{pmatrix}18 & & \\ & -6 & \\ & & 4\end{pmatrix}$. לכן נקבל שהשניונית המקורית חופפת לשניונית $18-6x^2+4y^2=0$ או $3x^2-2y^2=9$, זו היפרבולה.

15 לכסון משותף

הגדרה 15.1 אם S הפיכה כך ש־$S^{-1}AS$ וגם $S^{-1}BS$ שתיהן אלכסוניות (משולשיות) נאמר של־A,B יש לכסון (שילוש) משותף. כאן A,B שתי מטריצות מאותו הסדר.

משפט 15.2 אם A,B מטריצות מסדר n מעל $\mathbb{C}$ כך ש־$AB=BA$, אז ל־A,B יש וקטור עצמי משותף.

הוכחה: יהי λ_1 ערך עצמי של A ונניח ש־$\{x_1,\ldots,x_k\}\subseteq\mathbb{C}^n$ הינה בסיס של וקטורים עצמיים עבור המרחב העצמי של λ_1.

אם עבור $1\leq i\leq k$ מסוים מתקיים $Bx_i=0$, אז x_i הוא וקטור עצמי עם ערך עצמי 0 עבור B והמשפט מוכח. לכן, ניתן להניח $Bx_i\neq 0$ לכל $1\leq i\leq k$. במקרה זה, $A\left(Bx_i\right)=B\left(Ax_i\right)=\lambda_1 Bx_i$ ולכן Bx_i הוא וקטור עצמי של A עם ערך עצמי λ_1. לכן $Bx_i\in\operatorname{Sp}\left\{x_1,\ldots,x_k\right\}$ לכל $1\leq i\leq k$.

נסמן ב־X את המטריצה מסדר $n\times k$ שעמודותיה הם הוקטורים $x_1,\ldots,x_k$. למערכת הלינארית $Xy=Bx_i$ יש תמיד פתרון במשתנה y, כי משתמשים במשפט שלמערכת לינארית $Ax=b$ יש פתרון אם ורק אם $b\in C\left(A\right)$. נסמן את הפתרון ב־$y_i\in\mathbb{C}^k$, כלומר לכל $1\leq i\leq k$ מתקיים $Xy_i=Bx_i$. נסמן ב־Y את המטריצה הריבועית מסדר k שעמודותיה הם הוקטורים $y_1,\ldots,y_k$. לכן אוסף השוויונים הנ״ל שקול לשוויון $XY=BX$.

יהי μ_1 ערך עצמי עבור המטריצה Y, ויהי $z=\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_k\end{pmatrix}$ וקטור עצמי מתאים. לכן, $BXz=XYz=X\left(\mu_1 z\right)=\mu_1 Xz$.

כדי להסיק ש־Xz הוא וקטור עצמי צריך להוכיח ש־$Xz\neq 0$. מבניית X נובע ש־$Xz=\alpha_1x_1+\cdots+\alpha_kx_k$. אם $Xz=0$, אז נקבל סתירה להנחה ש־$x_1,\ldots,x_k$ הוא בסיס למרחב העצמי. לכן, $Xz\neq 0$.

מהשוויון הנ״ל $B\left(Xz\right)=\mu Xz$ מקבלים ש־Xz הוא וקטור עצמי של B, מצד שני $Xz\in\operatorname{Sp}\left\{x_1,\ldots,x_k\right\}$ ולכן Xz וקטור עצמי של A, ולכן הוא וקטור עצמי משותף. ■

משפט 15.3 אם A,B שתי מטריצות מעל $\mathbb{C}$ כך ש־$AB=BA$, אז יש להן שילוש אוניטרי משותף.

הוכחה: נוכיח באינדוקציה על n, סדר המטריצות. מהמשפט הקודם קיים וקטור $x_1\in\mathbb{F}^n$ באורך 1 כך ש־$Ax_1=\lambda_1x_1, Bx_1=\mu_1x_1$. נשלים את x_1 לבסיס אורתונורמלי ונקבל מטריצה S אוניטרית המקיימת $S^*AS=\begin{pmatrix}\lambda_1 & z\\ 0 & A_1\end{pmatrix}, S^*BS=\begin{pmatrix}\mu_1 & w\\ 0 & B_1\end{pmatrix}$. מתקיים

$$\left(S^*AS\right)\left(S^*BS\right)=S^*ABS=S^*BAS=\left(S^*BS\right)\left(S^*AS\right)$$

אם נציב את המטריצות הנ״ל נקבל את הקשר $\begin{pmatrix}\lambda_1\mu_1 & *\\ 0 & B_1A_1\end{pmatrix}=\begin{pmatrix}\lambda_1\mu_1 & *\\ 0 & A_1B_1\end{pmatrix}$, ולכן בהכרח $A_1B_1=B_1A_1$ וניתן להמשיך באינדוקציה. ■

משפט 15.4 תהיינה A,B מטריצות נורמליות מסדר n. אז $AB=BA$ $\iff$ קיימת U אוניטרית כך ש־U^*AU,U^*BU הן אלכסוניות.

הוכחה: $\Leftarrow$ נניח ש־$AB=BA$. מהמשפט הקודם, קיימת U אוניטרית כך ש־U^*AU,U^*BU הן משולשיות. כמו כן U^*AU,U^*BU הן נורמליות (ניתן לבדוק ישירות). מכפל מטריצות מטריצה נורמלית משולשים חייבת להיות אלכסונית, ולכן U^*AU וגם U^*BU הן אלכסוניות ול־A,B יש לכסון משותף.

$\Longrightarrow$ נניח שקיימת U אוניטרית כך ש־$U^*AU=D_1, U^*BU=D_2$ אלכסוניות. אז

$$AB=\left(UD_1U^*\right)\left(UD_2U^*\right)=UD_1D_2U^*=UD_2D_1U^*=\left(UD_2U^*\right)\left(UD_1U^*\right)=BA$$

■

16 מטריצות סטוכסטיות

הגדרה 16.1 מטריצה ממשית P כך ש$0 \le p_{ij} \le 1$ וגם $\sum_{j=1}^{n} p_{ij} = 1$ לכל i נקראת **מטריצה סטוכסטית**.

דוגמה 16.2 נניח שבאוכלוסייה מסוימת מחלקים לאנשים גבוהים ונמוכים, נניח ש60% מהבנים של גברים נמוכים הם נמוכים ו20% מהבנים של גברים גבוהים הם נמוכים. ניתן לסדר את הנתונים במטריצה סטוכסטית $P = \begin{pmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{pmatrix}$, כאשר האיברים בשורה הראשונה מייצגים את ההסתברות שלגבר נמוך יהיה בן נמוך או בן גבוה בהתאמה, והאיברים בשורה השנייה מייצגים את ההסתברות שלגבר גבוה יהיה בן נמוך או בן גבוה בהתאמה.

הגדרה 16.3 מטריצה P תקרא **חיובית** אם $p_{ij} > 0$ לכל i, j.

משפט 16.4 פרון: לכל מטריצה חיובית P יש ערך עצמי דומיננטי שנסמנו ב$\lambda(P)$ המקיים את התכונות הבאות:

1. $\lambda(P)$ הינו מספר ממשי חיובי ויש לו וקטור עצמי h שהוא וקטור חיובי.
2. ל$\lambda(P)$ אין וקטורים עצמיים נוספים הבלתי תלויים בh.
3. כל ערך עצמי אחר μ של P מקיים $|\mu| < \lambda(P)$.
4. לP אין וקטורים עצמיים השונים מh שהם וקטורים אי שליליים.

סימון: אם $u, v \in \mathbb{R}^n$ נאמר ש$u \ge v$ אם לכל $1 \le i \le n$, $u_i \ge v_i$, ובאותו האופן עבור $u > v$.
נסמן ב$p(P)$ את קבוצת כל המספרים הלא שליליים λ שעבורם קיים וקטור $0 \ne x \ge 0$ כך ש$Px \ge \lambda x$.

טענה 16.5 תהי P מטריצה חיובית.

1. $p(P)$ אינה ריקה ומכילה מספר חיובי.
2. $p(P)$ קבוצה חסומה וסגורה.

הוכחה: יהי $x > 0$. כיוון שP חיובית אז $Px > 0$. אם ניקח λ מספיק קטן נקבל ש$Px \ge \lambda x$, ולכן טענה 1 מתקיימת.

חסומה: כעת נניח ש$Px \ge \lambda x$ עבור $0 \ne x \ge 0$. אז על ידי חלוקה ב$\sum x_i \ne 0$ ניתן להניח כי x מקיים $ex = \sum x_i = 1$ כאשר $e = (1, 1, \ldots, 1)$. לכן, אם $Px \ge \lambda x$ אז $ePx \ge \lambda ex = \lambda$. יהי b המספר הגדול ביותר בוקטור eP. אז $be \ge eP$, ולכן

$$b \overset{\sum x_i = 1}{=} bex \ge ePx \ge \lambda$$

ולכן $p(P)$ חסומה.

סגורה: אם $x_n \to x$ סדרה מתכנסת כך ש$x_n \ge 0$ וגם $\lambda_n \to \lambda$ ומתקיים $Px_n \ge \lambda_n x_n$ אז $Px \ge \lambda x$. ■