

Algebra liniowa 2R

wersja beta

Tomasz Rzepecki

25 czerwca 2022

Spis treści

1	Podstawowe pojęcia algebry liniowej	5
1.1	Ciała	5
1.2	Przestrzenie i podprzestrzenie liniowe	7
1.3	Kombinacje liniowe i liniowa niezależność	11
2	Bazy i wymiary	14
2.1	Baza przestrzeni liniowej	14
2.2	Konstrukcja bazy	18
2.3	Lemat Steinitza	19
2.4	Wymiar a izomorfizmy liniowe	22
3	Przekształcenia liniowe	25
3.1	Przekształcenia liniowe	25
3.2	Jądro i obraz odwzorowania liniowego	27
3.3	Twierdzenia o rzędzie i o indeksie	30
4	Konstrukcje przestrzeni liniowych	33
4.1	Suma prosta	33
4.2	Przestrzeń dualna	34
4.3	Przestrzeń ilorazowa	37
4.4	Przestrzeń bidualna	40
5	Macierze i układy równań	42
5.1	Macierze układów równań	42
5.2	Operacje wierszowe i kolumnowe a rząd macierzy	44
5.3	Eliminacja Gaussa	48
5.4	Macierze przekształceń liniowych	56
5.5	Odwracalność macierzy	60

6	Wyznacznik	63
6.1	Permutacje i definicja wyznacznika	63
6.2	Podstawowe własności wyznacznika	72
6.3	Wzory Cramera	79
6.4	Wyznacznik i wielomian charakterystyczny odwzorowania liniowego	82
7	Rozkłady przestrzeni i ich endomorfizmów	85
7.1	Wewnętrzne sumy proste	85
7.2	Liniowa niezależność podprzestrzeni	86
7.3	Podprzestrzenie niezmiennicze	89
7.4	Przestrzenie własne	92
7.5	Diagonalizowalność	95
7.6	Twierdzenie Jordana	97
7.7	Kompleksyfikacja i rzeczywiste tw. Jordana	108
8	Formy dwuliniowe i kwadratowe	120
8.1	Formy dwuliniowe	120
8.2	Iloczynny skalarne	124
8.3	Ortogonalność	128
8.4	Formy kwadratowe	137
9	Przestrzenie z iloczynem skalarnym	142
9.1	Przestrzenie euklidesowe	142
9.2	Formy hermitowskie	146
9.3	Przestrzenie unitarne	154
10	Rozkład spektralny i singularny	156
10.1	Rozkład singularny i rozkład biegunowy	164
11	Izometrie i objętości	171
11.1	Reprezentacje przekształceń ortogonalnych i unitarnych . . .	171
11.2	Izometrie	174
11.3	Diagonalizacja form kwadratowych a hiperpowierzchnie kwadratowe	177
11.4	Macierz Grama i objętość	179

A Przestrzenie nieskończenie wymiarowe	185
A.1 Twierdzenie o wymiarze dla przestrzeni nieskończenie wymiarowych	185
A.2 Lemat Steinitza o wymianie	186
B Rzeczywiste twierdzenie Jordana	188
C Twierdzenie spektralne dla endomorfizmów normalnych	189
Indeks	192
Indeks symboli i oznaczeń	197

Rozdział 1

Podstawowe pojęcia algebry liniowej

1.1 Ciała

Definicja 1.1. *Ciałem* nazywamy zbiór K ¹ wraz z działaniami $+$, \cdot (formalnie: funkcjami $K \times K \rightarrow K$) oraz wyróżnionymi elementami 0 i 1 (formalnie: czwórkę $(K, +, \cdot, 0, 1)$), spełniającą warunki (*aksjomaty ciała*):

1. $+$, \cdot są dwuargumentowymi działaniami, które są łączne i przemienne, a \cdot jest rozdzielne względem $+$, tzn. dla każdych $x, y, z \in K$:
 - $x + y = y + x$,
 - $x \cdot y = y \cdot x$,
 - $x + (y + z) = (x + y) + z$,
 - $x \cdot (y \cdot z) = (x \cdot y) \cdot z$,
 - $x \cdot (y + z) = x \cdot y + x \cdot z$.
2. 0 jest elementem neutralnym dodawania, a 1 jest elementem neutralnym mnożenia, tzn. dla każdego $x \in K$ zachodzi $x + 0 = x \cdot 1 = x$.
3. dla każdego elementu $x \in K$, istnieje element przeciwny $-x$ względem $+$, tzn. taki że $x + (-x) = 0$,
- 4.

¹Ciała oznaczamy najczęściej literami K, k, F, L . Na tym wykładzie będziemy używać głównie litery K .

5. $0 \neq 1$.

Uwaga 1.2. $0, 1 \in K$ są po prostu wyróżnionymi elementami zbioru K , nie muszą być liczbami. Innymi słowy, dla K jako zbioru może być tak, że $K \cap \mathbf{C} = \emptyset$.

Ćwiczenie 1.3. Oznaczenia $-x$, x^{-1} są sensowne, tzn. jednoznacznie określają elementy K .

Przykłady 1.4. (a) \mathbf{Q} , \mathbf{R} , \mathbf{C} są ciałami,

(b) \mathbf{Z} nie jest ciałem, bo 2 nie ma elementu odwrotnego,

(c) każdy podzbiór \mathbf{C} (ogólniej: dowolnego ciała) zawierający 0 i 1, zamknięty na $+$, $-$, \cdot i $^{-1}$ (z wyjątkiem 0) jest ciałem,

(d) zbiór $\{0, 1, 2, 3, 4\}$ z dodawaniem i mnożeniem modulo 5 jest ciałem, gdzie $1^{-1} = 1$, $2^{-1} = 3$, $3^{-1} = 2$, $4^{-1} = 4$.

(e) ogólnie, jeżeli p jest liczbą pierwszą, to zbiór $\mathbf{F}_p = \{0, 1, 2, 3, \dots, p-1\}$ z dodawaniem i mnożeniem modulo p jest ciałem²,

(f) zbiór $\{0, 1, 2, 3, 4, 5\}$ z dodawaniem i mnożeniem modulo 6 nie jest ciałem, bo 2 nie ma elementu odwrotnego: gdyby k było odwrotne do 2, to by znaczyło że $k \cdot 2 \cdot 3 = 1 \cdot 3 = 3$, ale $k \cdot 2 \cdot 3 = 6k \equiv 0 \pmod{6}$.

(g) podobny argument pokazuje, że ciało nie może mieć *dzielników zera*, to znaczy takich elementów $a, b \neq 0$, że $ab = 0$.

(h) Dla $d \in \mathbf{Q}$ definiujemy $\mathbf{Q}[\sqrt{d}] = \{a + b\sqrt{d} \mid a, b \in \mathbf{Q}\}$. Dla każdego d ten zbiór jest ciałem.

(i) Jeżeli K jest ciałem, to zbiór $K[X]$ wielomianów zmiennej X , ze współczynnikami z K , nigdy nie jest ciałem (X nie ma elementu odwrotnego), ale jest nim *ciało funkcji wymiernych* $K(X)$, składające się z formalnych ilorazów $\frac{P(X)}{Q(X)}$, gdzie $P, Q \in K[X]$ i $Q \neq 0$.³

²jeszcze bardziej ogólnie, choć to trudniej zobaczyć: dla każdej liczby postaci $p^k > 1$, gdzie p jest pierwsza, istnieje (w zasadzie jedyne) ciało o p^k elementach (i nie ma innych skończonych ciał)

³Więcej o ciałach: trochę może na konwersatorium, ale zasadniczo to końcówka wykładu z Algebra 1 R i przede wszystkim — dla chętnych — Algebra 2 R.

Lemat 1.5. *Jeżeli K jest ciałem, to dla każdego $a \in K$ zachodzi $0 \cdot a = 0$.*

Dowód.

$$\begin{aligned} 0 \cdot a &= (0 + 0) \cdot a \\ 0 \cdot a &= 0 \cdot a + 0 \cdot a \\ 0 \cdot a + -(0 \cdot a) &= 0 \cdot a + 0 \cdot a + -(0 \cdot a) \\ 0 &= 0 \cdot a \end{aligned} \quad \square$$

1.2 Przestrzenie i podprzestrzenie liniowe

Definicja 1.6. Ustalmy ciało K . *Przestrzenią liniową (lub wektorową) nad K nazywamy zbiór V z wyróżnionym elementem $\vec{0} \in V$ (wektorem zerowym), oraz z działaniami $+_V : V \times V \rightarrow V$ oraz $\cdot_V : K \times V \rightarrow V$, spełniającymi warunki (aksjomaty przestrzeni liniowej):*

1. $+_V$ jest przemienne, łączne, $\vec{0}$ jest jego elementem neutralnym, ma elementy przeciwne:

$$\begin{aligned} &\bullet v_1 +_V v_2 = v_2 +_V v_1, & \bullet v +_V \vec{0} &= v, \\ &\bullet v_1 +_V (v_2 +_V v_3) = (v_1 +_V v_2) +_V v_3, & \bullet v +_V (-v) &= \vec{0} \end{aligned}$$

2. \cdot_V spełnia łączność mieszaną: dla $\alpha, \beta \in K$, $v \in V$ mamy $\alpha \cdot_V (\beta \cdot_V v) = (\alpha \cdot_K \beta) \cdot_V v$ (uwaga na indeksy przy kropkach!);

3. \cdot_V jest rozdzielne względem $+_V$ (uwaga na indeksy przy plusach!):

$$\begin{aligned} &\bullet \alpha \cdot_V (v +_V w) = \alpha \cdot_V v +_V \alpha \cdot_V w, \\ &\bullet (\alpha +_K \beta) \cdot_V v = \alpha \cdot_V v +_V \beta \cdot_V v; \end{aligned}$$

4. mnożenie przez jedynkę to funkcja identycznościowa: $1 \cdot_V v = v$. (Bez tego moglibyśmy zdefiniować \cdot_V stale równe $\vec{0}$, a tego nie chcemy.)

(Uwaga: zwykle zamiast $\vec{0}$, $+_V$, \cdot_V piszemy po prostu 0 , $+$ i \cdot , a \cdot często nie piszemy w ogóle.)

Dla danej przestrzeni liniowej V nad K , elementy V nazywamy *wektorami*, a elementy K nazywamy *skalarami*.

Uwaga 1.7. Podobnie jak dla ciał, elementy przeciwne są jedyne. Tak naprawdę ich istnienie, podobnie jak istnienie $\vec{0}$, jest konsekwencją pozostałych aksjomatów.

Przykłady 1.8.

- (a) $\mathbf{R}^2, \mathbf{R}^3, \mathbf{C}^2, \mathbf{C}^3$ są przestrzeniami liniowymi nad \mathbf{R} i \mathbf{C} odpowiednio.
- (b) Ogólnie, jeżeli K jest dowolnym ciałem, to K^n (z oczywistymi działaniami) jest przestrzenią liniową nad K .
- (c) Jeszcze bardziej ogólnie (?) jeżeli A jest dowolnym zbiorem, to zbiór K^A wszystkich funkcji $A \rightarrow K$ z działaniami zdefiniowanymi przez $(f_1 + f_2)(a) := f_1(a) + f_2(a)$, $(\alpha \cdot f)(a) := \alpha \cdot (f(a))$ jest przestrzenią liniową nad K .
- (d) Jeżeli K jest dowolnym ciałem, to zbiór $K[X]$ wszystkich wielomianów zmiennej X , o współczynnikach z K , jest przestrzenią liniową nad K , podobnie jak zbiory $K_n[X]$ wielomianów stopnia co najwyżej n .⁴
- (e) Przestrzeń $C(\mathbf{R})$ wszystkich funkcji ciągłych $\mathbf{R} \rightarrow \mathbf{R}$ jest przestrzenią liniową nad \mathbf{R} .

Stwierdzenie 1.9. • $0 \cdot v = \vec{0}$.

- Dla każdego wektora w i skalaru $a \neq 0$ istnieje jedyny wektor v taki że $av + w = 0$.

Dowód. Część pierwsza jak dla ciał: $0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v$, następnie odejmujemy stronami $0 \cdot v$.

Część druga: zauważmy że $(-a^{-1}) \cdot w$ działa: istotnie,

$$\begin{aligned} a \cdot ((-a^{-1}) \cdot w) + w &= (a \cdot (-a)^{-1})w + w \\ &= (-1) \cdot w + 1 \cdot w \\ &= (-1 + 1) \cdot w \\ &= 0 \cdot w \\ &= 0. \end{aligned}$$

⁴Nie mylić ze zbiorem wszystkich funkcji wielomianowych $K \rightarrow K$, np. są tylko cztery funkcje wielomianowe $\mathbf{F}_2 \rightarrow \mathbf{F}_2$ (wszystkie zadane wielomianami stopnia co najwyżej 1), ale $(\mathbf{F}_2)_n[X]$ ma 2^n elementów, zaś $\mathbf{F}_2[X]$ jest nieskończony.

Dowód jedyności jest podobny jak dowód jedyności elementu przeciwnego dla ciał. \square

Zauważmy że mamy w szczególności $(-1) \cdot w = -w$, czego przyjemną konsekwencją jest to, że $-(v + w) = -v + -w$ (ale to wynika też z przemienności $+_V$).

Lemat 1.10. *Założmy że V jest przestrzenią liniową nad K , a $W \subseteq V$ jest podzbiorem, takim że:*

1. $W \neq \emptyset$ (w praktyce sprawdzamy zwykle że $\vec{0} \in W$),
2. dla każdego $a \in K$ i $w \in W$ mamy $a \cdot_V w \in W$,
3. dla każdych $w_1, w_2 \in W$ mamy $w_1 +_V w_2 \in W$.

Wtedy W z tym samym $\vec{0}$ i z obcięciami $+_V, \cdot_V$ jest przestrzenią liniową.

Dowód. • Z założeń wynika, że obcięcia $+_V, \cdot_V$ faktycznie są takimi działaniami, jak w definicji przestrzeni liniowej (to znaczy: przyjmują wartości w W)

- Przemienność, łączność, łączność mieszana i rozdzielność są oczywiste.
- Pozostaje sprawdzić że $\vec{0} \in W$ i że W jest zamknięte na branie elementów przeciwnych.
- Skoro W jest niepuste, to ma pewien element $w \in W$. Ponieważ pokazaliśmy, że $0 \cdot w = \vec{0}$, wynika z tego (i z drugiego założenia), że $\vec{0} \in W$.
- Podobnie, jeżeli $w \in W$ jest dowolny, to $-w = (-1) \cdot w \in W$ (też z drugiego założenia). \square

Definicja 1.11. Podzbiór $W \subseteq V$ spełniający warunki Lematu powyżej nazywamy *podprzestrzenią (liniową) V* , oznaczamy $W \leq V$ (lub czasami $W < V$).

Przykłady 1.12. (a) proste przechodzące przez 0 w K^2 są podprzestrzeniami (rysunek: proste w $\mathbb{F}_2^2, \mathbb{F}_3^2$),

(b) naturalnie włożone K^n w K^m (przez dopisanie $(m - n)$ zer na końcu), $n < m$, jest podprzestrzenią

- (c) zbiór funkcji różniczkowalnych jest podprzestrzenią w $C(\mathbf{R})$ (prze-strzeni funkcji ciągłych $\mathbf{R} \rightarrow \mathbf{R}$),
- (d) zbiór funkcji ciągłych jest podprzestrzenią $\mathbf{R}^{\mathbf{R}}$,
- (e) zbiór funkcji takich że $\lim_{x \rightarrow x_0} f(x) = 0$ (gdzie x_0 jest ustalony) jest podprzestrzenią $\mathbf{R}^{\mathbf{R}}$,
- (f) przekrój powyższych jest podprzestrzenią $C(\mathbf{R})$ (złożoną z tych funkcji ciągłych, dla których $f(x_0) = 0$),
- (g) zbiór wszystkich ciągów o wyrazach rzeczywistych spełniających rekurencję $a_{n+2} = a_n + a_{n+1}$ jest podprzestrzenią $\mathbf{R}^{\mathbf{N}}$.

Lemat 1.13. *Jeżeli $W_1, W_2 \leq V$, to:*

- $W_1 \cap W_2 \leq V$,
- $W_1 + W_2 = \{w_1 + w_2 \mid w_1 \in W_1, w_2 \in W_2\} \leq V$ (suma kompleksowa)

Dowód. Pierwszy punkt — ćwiczenie.

Dowód drugiego punktu.

- Ponieważ W_1, W_2 są niepuste, to istnieje pewien $w_1 \in W_1$, $w_2 \in W_2$ i wtedy $w_1 + w_2 \in W_1 + W_2$, więc $W_1 + W_2 \neq \emptyset$.
- Dla dowolnych $a \in K, w_1 \in W_1, w_2 \in W_2$ mamy z założenia $aw_1 \in W_1, aw_2 \in W_2$, więc z definicji $W_1 + W_2 \ni (aw_1 + aw_2)$ i z rozdzielności $aw_1 + aw_2 = a(w_1 + w_2)$, więc drugi aksjomat jest spełniony.
- Dla dowolnych $w_1, w'_1 \in W_1, w_2, w'_2 \in W_2$, mamy z założenia $w_1 + w'_1 \in W_1, w_2 + w'_2 \in W_2$, więc z definicji $W_1 + W_2$ i z przemienności i łączności $+_V$ mamy $W_1 + W_2 \ni ((w_1 + w'_1) + (w_2 + w'_2)) = w_1 + w'_1 + w_2 + w'_2 = (w_1 + w_2) + (w'_1 + w'_2)$, co kończy dowód. \square

Stwierdzenie 1.14. *Jeżeli $W_1, W_2 \subseteq V$, to:*

- jeżeli $W_2 \leq V$ i $W_1 \leq W_2$, to $W_1 \leq V$, (uwaga: bez założenia $W_2 \leq V$ to drugie \leq nie ma sensu);
- jeżeli $W_1, W_2 \leq V$ i $W_1 \subseteq W_2$, to $W_1 \leq W_2$.

Dowód. Ćwiczenie. \square

1.3 Kombinacje liniowe i liniowa niezależność

Definicja 1.15. Niech V będzie przestrzenią liniową, a $A \subseteq V$ — dowolnym podzbiorem.

Otoczką liniową (lub liniowym domknięciem) A w V nazywamy najmniejszą podprzestrzeń V zawierającą A , i oznaczamy ją $\text{Lin}(A)$ (lub inaczej, np. $\text{Span}(A)$).

Stwierdzenie 1.16. $\text{Lin}(A)$ to dokładnie zbiór (wszystkich) wektorów postaci $\sum_{k=1}^n \alpha_k v_k$, gdzie $\alpha_k \in K$, a $v_k \in A$, lub równoważnie, wektorów postaci $\sum_{a \in A} \alpha_a a$, gdzie prawie wszystkie α_a są zerowe (to znaczy: wszystkie z wyjątkiem skończonego wielu; formalnie element który taką sumę opisuje to suma wszystkich wyrazów, przy których współczynniki są niezerowe).

Dowód. Jest jasne, że $\text{Lin}(A)$ zawiera wszystkie wektory podanej postaci.

Z drugiej strony $\sum_{k=1}^n \alpha_k v_k + \sum_{k=1}^m \beta_k w_k = \sum_{k=1}^{n+m} \gamma_k u_k$, gdzie $\gamma_k = \begin{cases} \alpha_k & k \leq n \\ \beta_{k-n} & k > n \end{cases}$ i $u_k = \begin{cases} v_k & k \leq n \\ w_{k-n} & k > n \end{cases}$, a także $\alpha \cdot \sum_{k=1}^n \alpha_k v_k = \sum_{k=1}^n (\alpha \alpha_k) v_k$. \square

dokładniejszy dowód. Oznaczmy przez W zbiór wszystkich sum postaci $\sum_{k=1}^n \alpha_k v_k$, gdzie $v_k \in A$. Chcemy pokazać, że $W = \text{Lin}(A)$.

Jest jasne, że $A \subseteq W$, łatwo też zauważyć, że W zawiera się w każdej podprzestrzeni V zawierającej A (bo każdy element jest sumą skalarnych wielokrotności elementów A). Wystarczy zatem pokazać, że W jest podprzestrzenią V .

Zauważmy że

$$\sum_{k=1}^n \alpha_k v_k + \sum_{k=1}^m \beta_k w_k = \sum_{k=1}^{n+m} \gamma_k u_k,$$

gdzie $\gamma_k = \begin{cases} \alpha_k & k \leq n \\ \beta_{k-n} & k > n \end{cases}$ i $u_k = \begin{cases} v_k & k \leq n \\ w_{k-n} & k > n \end{cases}$, a także

$$\alpha \cdot \sum_{k=1}^n \alpha_k v_k = \sum_{k=1}^n (\alpha \alpha_k) v_k.$$

Ponadto W jest niepusty (bo należy do niego $\vec{0}$, jako pusta suma). Stąd W jest podprzestrzenią V , co kończy dowód.

Inaczej: niech W będzie zbiorem wektorów postaci $\sum_{a \in A} \alpha_a a$ (gdzie tylko skończenie wiele $\alpha_a \neq 0$).

Chcemy pokazać, że $W = \text{Lin}(A)$. Istotnie, W oczywiście zawiera A , jest zamknięty na mnożenie przez skalary: $\beta \cdot \sum_{a \in A} \alpha_a a = \sum_{a \in A} (\beta \alpha_a) a$ wciąż ma prawie wszystkie współczynniki 0, podobnie jako $\sum_{a \in A} \alpha_a a + \sum_{a \in A} \beta_a a = \sum_{a \in A} (\alpha_a + \beta_a) a$. (Uwaga: ponieważ te sumy tak naprawdę są skończone — mają tylko skończenie wiele niezerowych wyrazów — możemy zmieniać kolejność sumowania i korzystać z praw rozdzielności tak jak to tutaj robimy!).

W jest więc podprzestrzenią V zawierającą A , a ponadto W zawiera się w każdej podprzestrzeni zawierającej A , bo każdy element jest sumą skalarnych wielokrotności elementów A , więc z definicji $W = \text{Lin}(A)$. \square

Wniosek 1.17. $\text{Lin}(A) = \bigcap \{ W \leq V \mid A \subseteq W \}$

Dowód. Z definicji $\text{Lin}(A) \subseteq W$ dla każdego $W \in \{ W \leq V \mid A \subseteq W \}$, więc $\text{Lin}(A) \subseteq \bigcap \{ W \leq V \mid A \subseteq W \}$. Z drugiej strony $\text{Lin}(A) \in \{ W \leq V \mid A \subseteq W \}$, więc $\text{Lin}(A) \supseteq \bigcap \{ W \leq V \mid A \subseteq W \}$. \square

Definicja 1.18. *Kombinacja liniowa* wektorów v_1, \dots, v_n to element $\text{Lin}(v_1, \dots, v_n)$ (czyli wektor postaci $\sum_{k=1}^n \alpha_k v_k$).

Przykłady 1.19. (a) $\text{Lin}(\emptyset) = \{\vec{0}\}$.

(b) $\text{Lin}(v) = \{ \alpha v \mid \alpha \in K \}$.

(c) $\text{Lin} \left(\left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbf{R}^2 \mid xy = 1 \right\} \right) = \mathbf{R}^2$

(d) $\text{Lin} \left(\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbf{R}^3 \mid x + y + z = 1 \right\} \right) = \mathbf{R}^3$

Uwaga 1.20. • $A \subseteq B$ implikuje $\text{Lin}(A) \subseteq \text{Lin}(B)$,

• $\text{Lin}(\text{Lin}(A)) = \text{Lin}(A)$,

• $b \in \text{Lin}(A)$ wtedy i tylko wtedy gdy $\text{Lin}(A) = \text{Lin}(A \cup \{b\})$.

Dowód. Pierwszy punkt: $\text{Lin}(B)$ jest podprzestrzenią V zawierającą B , a więc zawierającą A . Zawiera zatem z definicji najmniejszą podprzestrzeń V zawierającą A .

Drugi punkt: $\text{Lin}(A)$ jest podprzestrzenią V , więc najmniejsza podprzestrzeń zawierająca go to ona sama.

Trzeci punkt: jeżeli $b \in \text{Lin}(A)$, to $\text{Lin}(A \cup \{b\}) \subseteq \text{Lin}(\text{Lin}(A)) = \text{Lin}(A)$ (z pierwszego i drugiego punktu), i oczywiście $\text{Lin}(A) \subseteq \text{Lin}(A \cup \{b\})$, więc $\text{Lin}(A) = \text{Lin}(A \cup \{b\})$.

W drugą stronę, jeżeli $\text{Lin}(A) = \text{Lin}(A \cup \{b\})$, to (ponieważ z definicji $b \in \text{Lin}(A \cup \{b\})$), natychmiast $b \in \text{Lin}(A)$. \square

Definicja 1.21. Układ wektorów v_1, \dots, v_n nazywamy *liniowo niezależnym* (w skrócie *lnz*), jeżeli $\sum_{k=1}^n \alpha_k v_k = 0$ zachodzi tylko gdy $\alpha_1 = \dots = \alpha_n = 0$. W przeciwnym wypadku nazywamy go *liniowo zależnym* (w skrócie *lz*).

Mówimy że zbiór A wektorów jest *liniowo niezależny*, jeżeli każdy skończony układ różnych wektorów z A jest lnz, lub równoważnie, jeżeli dla każdego $a \in A$ zachodzi $\text{Lin}(A) \neq \text{Lin}(A \setminus \{a\})$ (żaden element A nie jest kombinacją liniową pozostałych). W przeciwnym wypadku mówimy że jest *liniowo zależny*.

Rozdział 2

Bazy i wymiary

2.1 Baza przestrzeni liniowej

Uwaga 2.1. Jeżeli v jest niezerowym wektorem, to zbiór $\{v, v, v\} = \{v\}$ jest liniowo niezależny, ale ciąg v, v, v jest liniowo zależny!

Ogólnie v_1, \dots, v_n jest liniowo niezależny jeżeli v_k są parami różne i zbiór $\{v_1, \dots, v_n\}$ jest liniowo niezależny.

Ćwiczenie 2.2. Jeżeli $v \in \text{Lin}(A) \setminus A$, to $A \cup \{v\}$ jest liniowo zależny.

Uwaga 2.3. Jeżeli zbiór A jest liniowo niezależny, to każdy jego podzbiór również liniowo niezależny. Podobnie dla układów wektorów.

Twierdzenie 2.4. Niech $B \subseteq V$. Następujące warunki są równoważne:

1. B jest lnz i $\text{Lin}(B) = V$,
2.
 - a. B jest lnz i dla każdego $v \in V \setminus B$ zbiór $B \cup \{v\}$ jest lz,
 - b. B jest maksymalnym zbiorem liniowo niezależnym,
3. każdy element V zapisuje się jednoznacznie jako kombinacja liniowa elementów B .

Dowód. (1) \Rightarrow (2a): wiemy że B jest liniowo niezależny. Weźmy dowolny $v \notin B$. Wtedy też (z założenia) $v \in \text{Lin}(B)$, więc z z Ćwiczenia 2.2 dostajemy tezę.

(2a) \Rightarrow (2b) wynika natychmiast z Uwagi 2.3.

(2b) \Rightarrow (3): ustalmy dowolny element $v \in V$.

Pokażemy najpierw, że v zapisuje się jako kombinacja liniowa elementów B . Istotnie, jeżeli $v \in B$, to jest to oczywiście prawda. Jeżeli $v \notin B$, to z założenia zbiór $B \cup \{v\}$ jest liniowo zależny, mamy więc kombinację liniową $\alpha \cdot v + \sum_{k=1}^n \alpha_k v_k = 0$, gdzie nie wszystkie z $\alpha, \alpha_1, \dots, \alpha_n$ są zerowe. Gdyby $\alpha = 0$, to $\sum_{k=1}^n \alpha_k v_k = 0$, co przeczyłoby liniowej niezależności B . Wobec tego $\alpha \neq 0$, ale wtedy $v = \sum_{k=1}^n -\alpha^{-1} \alpha_k v_k$.

Aby pokazać jedyność, założymy że $v = \sum_b \alpha_b b$ i $v = \sum_b \beta_b b$. Wtedy

$$v - v = 0 = \sum_b (\alpha_b - \beta_b) b,$$

więc (z liniowej niezależności) wszystkie $\alpha_b - \beta_b$ są zerowe, a więc $\alpha_b = \beta_b$.

(3) \Rightarrow (1): „zapisuje się” w oczywisty sposób implikuje $\text{Lin}(B) = V$; jeżeli $\sum_b \alpha_b b = 0$, to z jednoznaczności zapisu $\vec{0}$ jako kombinacji liniowej elementów B wynika, że $\alpha_b = 0$, czyli B jest lnz. \square

Definicja 2.5. Baza przestrzeni liniowej V to taki $B \subseteq V$ który spełnia warunki z powyższego twierdzenia.

Przykłady 2.6. (a) Bazą K^n jest $\{e_1, \dots, e_n\}$.

(b) Jeżeli A jest skończony, to bazą K^A jest zbiór funkcji postaci $\delta_a(x) = \begin{cases} 1 & x = a \\ 0 & x \neq a \end{cases}$.

Uzasadnienie:

- liniowa niezależność: $\sum_a \alpha_a \delta_a(b) = \alpha_b$, więc jeżeli $\sum_a \alpha_a \delta_a = \vec{0}$ (czyli jest funkcją stale równą 0), to wszystkie α_a są zerowe,
- dla dowolnej $f \in K^A$ mamy $f = \sum_{a \in A} f(a) \delta_a$, bo $\sum_a f(a) \delta_a(b) = f(b)$.

(c) Jeżeli A jest nieskończony, to taki zbiór nie jest bazą, bo np. funkcja stale równa 1 nie jest w jego liniowym domknięciu.

- Żeby to zobaczyć, zauważmy że dla dowolnego $f = \sum_{k=1}^n \alpha_k \delta_{a_k}$ możemy wziąć $b \neq a_1, \dots, a_k$ i wtedy wszystkie $\delta_{a_k}(b) = 0$, czyli $f(b) = 0$, a więc f nie jest stale równa 1.

- (d) W przestrzeni $K[X]$ (wielomianów zmiennej X o współczynnikach z K) jednomiany X^n tworzą bazę.

Przypomnienie z WDM:

Twierdzenie 2.7 (Lemat Kuratowskiego-Zorna). *Jeżeli (P, \leq) jest zbiorem częściowo uporządkowanym, który:*

- *jest niepusty,*
- *ma ograniczenie górne każdego łańcucha, tzn. dla każdego łańcucha $L \subseteq P$ istnieje $\bar{l} \in P$ takie że $\bar{l} \geq l$ dla każdego $l \in L$.*

Wtedy P ma element maksymalny.

Dowód. Wstęp do matematyki. □

Twierdzenie 2.8 (Twierdzenie o istnieniu bazy). *Każda przestrzeń liniowa ma bazę.¹*

Dowód. Ustalmy dowolną przestrzeń liniową V nad ciałem K . Rozważmy zbiór częściowo uporządkowany (P, \subseteq) , którego elementami są liniowo niezależne podzbiory V . Zauważmy że spełnia on założenia lematu Kuratowskiego-Zorna:

- jest niepusty, bo $\emptyset \in P$,
- każdy łańcuch w P ma ograniczenie górne: jeżeli $L \subseteq P$ jest łańcuchem, to $l^* = \bigcup L$ (zbiór takich $v \in V$, które należą do pewnego $l \in L$) jest jego ograniczeniem górnym; żeby sprawdzić liniową niezależność wystarczy zauważyć, że jeżeli $v_1, \dots, v_n \in \bigcup L$, to należą one też do pewnego $l \in L$.

Dokładniej:

- chcemy pokazać, że l^* jest lnz; w tym celu wystarczy pokazać, że każda n -ka różnych elementów l^* jest lnz;

¹Blass pokazał w 1984, że to twierdzenie (dla przestrzeni nieskończenie wymiarowych) jest równoważne (w ZF) pewnikowi wyboru. Niestety dowód wymaga znajomości algebry abstrakcyjnej wykraczającej za daleko poza zakres kursu z algebry liniowej, żeby umieścić go w tym skrypcie.

- ustalmy dowolną taką n -kę $v_1, \dots, v_n \in l^*$ (parami różnych elementów);
- ponieważ $l^* = \bigcup L$, to dla każdego v_k możemy wybrać pewne $l_k \in L$ takie że $v_k \in l_k$;
- ponieważ L jest łańcuchem, pewien z l_k , powiedzmy l_m , jest największy z nich, czyli $l_m \supseteq l_1, \dots, l_n$,
- ponieważ $v_k \in l_k \subseteq l_m$, mamy $v_1, \dots, v_n \in l_m$;
- z tego i z liniowej niezależności l_m (a także z tego że v_1, \dots, v_n są parami różne) wynika, że v_1, \dots, v_n są liniowo niezależne.

Zatem z LKZ wynika, że istnieje maksymalny zbiór liniowo niezależny B , czyli baza V . \square

Twierdzenie 2.9. *Niech V będzie przestrzenią liniową, a $N \subseteq G \subseteq V$ będą takie że:*

- N jest liniowo niezależny,
- $\text{Lin}(G) = V$.

Wtedy istnieje baza B taka że $N \subseteq B \subseteq G$.

Dowód. Podobny jak powyżej: rozważamy częściowy porządek (P, \subseteq) , gdzie $P = \{A \subseteq G \mid N \subseteq A, A \text{ jest liniowo niezależny}\}$.

P spełnia założenia Lematu Kuratowskiego-Zorna: niepustość wynika z tego, że $N \in P$.

Niech B będzie maksymalnym elementem P . Wtedy oczywiście $N \subseteq B \subseteq G$ i N jest liniowo niezależny. Zostaje pokazać, że $\text{Lin}(B) = V$.

Pokażemy najpierw, że $G \subseteq \text{Lin}(B)$. Weźmy dowolny $v \in G$. Jeżeli $v \notin B$, to z maksymalności B , $\text{Lin}(B \cup \{v\})$ jest liniowo zależny, a więc $\alpha v + \sum_{b \in B} \alpha_b b = 0$ dla pewnych α, α_b nie wszystkich równych 0. Argumentując jak wcześniej wnioskujemy, że $\alpha \neq 0$, czyli $\alpha v = -\sum_{b \in B} \alpha_b b$, czyli $v = \sum (-\alpha \alpha_b) b$, a więc $v \in \text{Lin}(B)$.

Mamy zatem $G \subseteq \text{Lin}(B)$. Kończymy stosując Uwagę 1.20: skoro $G \subseteq \text{Lin}(B)$, to $\text{Lin} G \subseteq \text{Lin Lin} B = \text{Lin} B$, ale z założenia $\text{Lin} G = V$. Z drugiej strony oczywiście $\text{Lin} B \subseteq V$, co kończy dowód. \square

Wniosek 2.10. *Niech V będzie przestrzenią liniową.*

- Jeżeli $N \subseteq V$ jest liniowo niezależny, to N można rozszerzyć do bazy V . (Zastosuj poprzednie twierdzenie dla $G = V$.)
- Jeżeli $G \subseteq V$ rozpiną V (tzn. $V = \text{Lin}(G)$), to G zawiera pewną bazę V . (Zastosuj poprzednie twierdzenie dla $N = \emptyset$.)

2.2 Konstrukcja bazy

Ćwiczenie 2.11. Jeżeli układ v_1, \dots, v_k są lnz i v_{k+1} nie jest ich kombinacją liniową, to układ v_1, \dots, v_k, v_{k+1} też jest lnz.

Poniższe stwierdzenie pozwala nam praktycznie uzyskać bazę ze skończonego zbioru rozpinającego przestrzeń liniową.

Stwierdzenie 2.12. Załóżmy że $V = \text{Lin}(v_1, \dots, v_n)$. definiujemy rekurencyjnie ciąg podzbiorów V :

- $B_0 = \emptyset$,
- $B_{k+1} = \begin{cases} B_k \cup \{v_{k+1}\} & v_{k+1} \notin \text{Lin}(B_k) \\ B_k & \text{w przeciwnym wypadku} \end{cases}$

Wtedy B_n jest bazą V . (Uwaga: $\text{Lin}(\emptyset) = \{0\}$.)

Dowód. Najpierw przez łatwą indukcję pokazujemy że $v_k \in \text{Lin}(B_k) \subseteq \text{Lin}(B_n)$ (bo $B_k \subseteq B_n$). Stąd $\text{Lin}(B_n) \supseteq \text{Lin}(v_1, \dots, v_n) = V$. Odwrotna inkluzja jest oczywista, bo $B_n \subseteq \{v_1, \dots, v_n\}$, pozostaje więc pokazać, że B_n jest liniowo niezależny.

Dowód tego też jest indukcyjny. B_0 oczywiście jest liniowo niezależny. Załóżmy teraz, że B_k jest liniowo niezależny. Jeżeli $B_{k+1} = B_k$, to oczywiście też jest liniowo niezależny. W przeciwnym wypadku $v_{k+1} \notin \text{Lin}(B_k)$, więc z ćwiczenia łatwo wynika, że B_{k+1} jest lnz. \square

Uwaga 2.13. Ze Stwierdzenia i jego dowodu wynika, że w sformułowaniu Stwierdzenia warunek „ $v_{k+1} \notin \text{Lin}(B_k)$ ” można zastąpić warunkiem „ $\text{Lin}(B_k \cup v_{k+1})$ jest lnz” lub warunkiem „ $v_{k+1} \notin \text{Lin}(v_1, \dots, v_k)$ ”.

Uwaga 2.14. Stwierdzenie powyżej prowadzi do bardziej konstruktywnego dowodu twierdzenia o istnieniu bazy dla przestrzeni rozpinanej przez skończony układ wektorów (bez wykorzystania lematu Kuratowskiego-Zorna).

Przykład 2.15. Wybieramy bazę w $V = \text{Lin}(X^2, (1+X)^2, 2+4X, X^3, X^3+X) \leq \mathbb{R}[X]$

- $B_0 = \emptyset$,
- $B_1 = \{X^2\}$, bo $X^2 \neq 0$,
- $B_2 = \{X^2, (1+X)^2\}$, bo $(1+X)^2$ nie jest wielokrotnością X^2 ,
- $B_3 = B_2$, bo $2+4X = 2(1+X)^2 - 2X^2$,
- $B_4 = \{X^2, (1+X)^2, X^3\}$, bo kombinacje liniowe X^2 i $(1+X)^2$ mają stopień co najwyżej 2,
- $B_5 = \{X^2, (1+X)^2, X^3, X^3+X\}$, bo współczynniki przy X i 1 w $aX^2 + b(1+X)^2 + cX^3$ są równe $2b$ i b (odpowiednio), a w X^3+X to 1 i 0, a $1 \neq 2 \cdot 0$.

2.3 Lemat Steinitza

Twierdzenie 2.16 (Lemat Steinitza o wymianie).² Załóżmy że B jest bazą V , a a_1, \dots, a_n jest liniowo niezależnym ciągiem wektorów. Wtedy:

- B ma co najmniej n elementów,
- możemy wybrać parami różne elementy $c_1, \dots, c_n \in B$ takie że $(B \setminus \{c_1, \dots, c_n\}) \cup \{a_1, \dots, a_n\}$ wciąż jest bazą V .

Wniosek 2.17 (twierdzenie o wymiarze). Każde dwie bazy ustalonej przestrzeni liniowej mają tyle samo elementów.

Dowód. Dowód dla przypadku gdy bazy są skończone.³

Jeżeli B_1, B_2 są bazami V , to B_2 jest liniowo niezależny, więc z lematu Steinitza zastosowanego dla $B = B_1$ mamy $|B_1| \geq |B_2|$. Podobnie $|B_2| \geq |B_1|$, czyli $|B_1| = |B_2|$. \square

²Prawdziwy jest również wariant lematu Steinitza dla nieskończonych zbiorów, patrz Twierdzenie A.4, ale dowód wymaga nieco silniejszych narzędzi

³Twierdzenie to ma też inny dowód, korzystający z tzw. twierdzenia Halla o kojarzeniu małżeństw. Ten drugi dowód działa również w przypadku nieskończonych baz; patrz Wniosek A.2.

Definicja 2.18. Wymiar przestrzeni liniowej V nad K , oznaczany $\dim_K V$ (lub $\dim V$, kiedy K wynika z kontekstu) to moc jej bazy.

Przykłady 2.19. (a) W Przykładzie 2.15: pokazaliśmy, że $V = \text{Lin}(X^2, (1+X)^2, 2+4X, X^3, X^3+X) \leq \mathbf{R}[X]$ ma bazę $B_5 = \{X^2, (1+X)^2, X^3, X^3+X\}$, więc $\dim V = 4$.

(b) K^n jest wymiaru n nad K .

(c) \mathbf{C}^n jest wymiaru $2n$ nad \mathbf{R} .

(d) $K[X]$ jest nieskończonego (ale przeliczalnego) wymiaru nad K .

(e) \mathbf{C} i \mathbf{R} są nieprzeliczalnego wymiaru (continuum) nad \mathbf{Q} .

Stwierdzenie 2.20. Jeżeli B jest bazą V , $a = \sum_{b \in B} \alpha_b b$ i $c \in B$ jest taki że α_c jest niezerowy, to $(B \setminus \{c\}) \cup \{a\}$ jest bazą V .

Dowód. Z założenia łatwo wynika, że $c = \alpha_c^{-1} a - \sum_{b \in B \setminus \{c\}} \alpha_c^{-1} \alpha_b b$, więc $c \in \text{Lin}((B \setminus \{c\}) \cup \{a\})$. To pokazuje, że $B \subseteq \text{Lin}((B \setminus \{c\}) \cup \{a\})$, więc $\text{Lin}((B \setminus \{c\}) \cup \{a\}) = V$.

Pozostaje pokazać, że $(B \setminus \{c\}) \cup \{a\}$ jest liniowo niezależny. Załóżmy więc, że $\beta_a a + \sum_{b \in B \setminus \{c\}} \beta_b b = 0$. Podstawiając $a = \sum_{b \in B} \alpha_b b$ i zmieniając kolejność sumowania dostajemy

$$\begin{aligned} 0 &= \beta_a a + \sum_{b \in B \setminus \{c\}} \beta_b b \\ &= \beta_a \sum_{b \in B} \alpha_b b + \sum_{b \in B \setminus \{c\}} \beta_b b \\ &= \beta_a \alpha_c c + \sum_{b \in B \setminus \{c\}} (\beta_a \alpha_b + \beta_b) b \end{aligned}$$

(Uwaga: możemy tu zmieniać kolejność sumowania, bo te sumy są tak naprawdę skończone, nawet jeżeli B nie jest!)

W tej ostatniej sumie mamy już kombinację liniową elementów B , więc współczynnik $\beta_a \alpha_c = 0$. Ponieważ założyliśmy, że $\alpha_c \neq 0$, to musi być $\beta_a =$

$0\alpha_c^{-1} = 0$, a stąd

$$\begin{aligned} 0 &= \beta_a \alpha_c c + \sum_{b \in B \setminus \{c\}} (\beta_a \alpha_b + \beta_b) b \\ &= 0\alpha_c c + \sum_{b \in B \setminus \{c\}} (0\alpha_b + \beta_b) b \\ &= 0c + \sum_{b \in B \setminus \{c\}} \beta_b b \end{aligned}$$

a to wprost z liniowej niezależności B implikuje, że wszystkie β_b są równe 0, czyli $(B \setminus \{c\}) \cup \{a\}$ jest liniowo niezależny, jest więc bazą B . \square

dowód Lematu Steinitza. Weźmy teraz dowolny liniowo niezależny układ a_1, \dots, a_n . Skonstruujemy rekurencyjnie ciąg c_1, \dots, c_n elementów B , takich że $(B \setminus \{c_1, \dots, c_k\}) \cup \{a_1, \dots, a_k\}$ jest bazą.

- Skoro B jest bazą V i $a_1 \in V$, to możemy zapisać $a_1 = \sum_{b \in B} \alpha_b b$. Ponieważ a_1 jest niezerowy (a jest, bo a_1, \dots, a_n jest lnz), pewien α_c jest niezerowy, możemy więc wziąć $c_1 = c$.
- Wtedy $B_1 = (B \setminus \{c_1\}) \cup \{a_1\}$ jest znowu bazą, moglibyśmy więc analogicznie wybrać $c_2 \in B_1$ taki że $(B_1 \setminus \{c_2\}) \cup \{a_2\}$ jest bazą, ale... to nie jest do końca dobrze: może się zdarzyć, że weźmiemy $c_2 = a_1$, a wtedy dostalibyśmy być może $c_2 \notin B$!
- Musimy wobec tego rozumować nieco sprytniej; zapisując $a_2 = \sum_{b \in B_1} \alpha_b b$ możemy wyłączyć wyraz z a_1 , otrzymując $a_2 = \alpha_{a_1} a_1 + \sum_{b \in B_1 \setminus \{c_1\}} \alpha_b b$. Wtedy któryś ze współczynników α_b musi być niezerowy (bo inaczej a_1, a_2 byłyby współliniowe!), możemy więc wybrać $c_2 \in B_1 \setminus \{c_1\} (= B \setminus \{b_1\})$.
- Powiedzmy, że mamy już c_1, \dots, c_k . Oznaczmy $B'_k = B \setminus \{c_1, \dots, c_k\}$, $B_k = B'_k \cup \{a_1, \dots, a_k\}$. Wiemy że B_k jest bazą, możemy więc zapisać

$$\begin{aligned} a_{k+1} &= \sum_{b \in B_k} \alpha_b b \\ &= \sum_{m=1}^k \alpha_{a_m} a_m + \sum_{b \in B'_k} \alpha_b b. \end{aligned}$$

Zauważmy że ta druga suma nie może być zerowa, ponieważ a_1, \dots, a_{k+1} jest liniowo niezależny, więc dla pewnego $c_{k+1} \in B'_k$ mamy $\alpha_{c_{k+1}} \neq 0$. Ze stwierdzenia wynika, że to c_{k+1} działa.

To daje nam c_1, \dots, c_n jak w drugiej części Twierdzenia — zauważmy, że są one faktycznie różne, bo c_{k+1} jest za każdym razem brany z B'_k , który jest rozłączny z $\{c_1, \dots, c_k\}$.

Pierwsza część wynika z drugiej, bo zbiór $\{c_1, \dots, c_n\} \subseteq B$ jest n -elementowy. \square

2.4 Wymiar a izomorfizmy liniowe

Stwierdzenie 2.21. (a) jeżeli $W \leq V$, to $\dim W \leq \dim V$,

(b) jeżeli $B \subseteq V$ jest n -elementowy i liniowo niezależny, a $\dim V = n < \infty$, to B jest bazą V ,

(c) jeżeli $W \leq V$ i $\dim W = \dim V < \infty$, to $W = V$.⁴

Dowód. Pierwsza część: każda baza W jest lnz w V , więc rozszerza się do bazy V .

Druga część: ustalmy bazę C przestrzeni V . Z lematu Steinitza wynika, że dla pewnych n różnych $c_1, \dots, c_n \in C$, zbiór $C \setminus \{c_1, \dots, c_n\} \cup B$ jest bazą V . Ale C ma dokładnie n elementów (bo $\dim V = n$), więc $(C \setminus \{c_1, \dots, c_n\}) \cup B = B$, czyli $\text{Lin } B = \text{Lin } C = V$.

Trzecia część: weź bazę B przestrzeni W i zastosuj drugą część. \square

Definicja 2.22. Niech V, W będą przestrzeniami liniowymi. Funkcję $F: V \rightarrow W$ nazywamy *izomorfizmem (liniowym)* jeżeli:

- F jest addytywne i jednorodne:

$$- F(v_1 + v_2) = F(v_1) + F(v_2),$$

$$- F(\alpha v) = \alpha F(v);$$

- F jest bijekcją.

⁴Ten punkt i poprzedni nie zachodzą przypadku nieskończenie wymiarowym. Np. $\dim K[X^2] = \dim K[X]$.

Mówimy że $V \cong W$ (słownie: V i W są *izomorficzne*) jeżeli istnieje izomorfizm liniowy $V \rightarrow W$.

Stwierdzenie 2.23. *Relacja \cong jest relacją równoważności.*

Dowód. Ćwiczenie. □

Twierdzenie 2.24. *Założmy że V, W to przestrzenie liniowe nad K , a n to liczba naturalna.*

- *Jeżeli $\dim V = n$, to $V \cong K^n$,*⁵
- *$V \cong W$ wtedy i tylko wtedy gdy $\dim V = \dim W$.*

Dowód. Założmy że $\dim V = n$. Niech $B = \{v_1, \dots, v_n\}$ będzie bazą V . Chcemy

zdefiniować $F: V \rightarrow K^n$ wzorem $F(v) = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \in K^n$, gdy $v = \sum_{k=1}^n \alpha_k v_k$. Co

trzeba sprawdzić:

- F jest dobrze określona: to wynika z tego, że każdy element V przedstawia się *jednoznacznie* w postaci $\sum_{k=1}^n \alpha_k v_k$ (czyli z tego, że v_1, \dots, v_n stanowią bazę).
- F jest addytywna: to wynika z tego, że

$$\sum_{k=1}^n \alpha_k v_k + \sum_{k=1}^n \beta_k v_k = \sum_{k=1}^n (\alpha_k + \beta_k) v_k.$$

(Formalnie: z przemienności $+$ i rozdzielności \cdot względem $+$.)

- F jest jednorodna: wynika z $\alpha \sum_{k=1}^n \alpha_k v_k = \sum_{k=1}^n (\alpha \alpha_k) v_k$ (formalnie: z rozdzielności — tej drugiej).
- F jest różnowartościowa: jeżeli $F(v) = (\alpha_1 \ \alpha_2 \ \dots \ \alpha_n)^\top = F(w)$, to znaczy że $v = \sum_{k=1}^n \alpha_k v_k = w$.

⁵To jest prawdą też gdy n jest potencjalnie nieskończoną liczbą kardynalną, jeżeli po prawej stronie napiszemy $K^{\oplus n}$. K^n dla nieskończonego n ma wymiar ostro większy.

- F jest „na”: dla dowolnego $y = (\alpha_1 \ \alpha_2 \ \cdots \ \alpha_n)^\top \in K^n$ mamy $y = F(v)$, gdzie $v = \sum_{k=1}^n \alpha_k v_k$.

To pokazuje pierwszą część.

Druga część (przypadek skończenie wymiarowy):⁶ jeżeli $\dim V = \dim W$, to z pierwszej części mamy $V \cong K^n$ i $W \cong K^n$. Ponieważ \cong jest relacją równoważności, to kończy dowód \Leftarrow .

Założmy że $F: V \rightarrow W$ jest izomorfizmem i ustalmy bazę v_1, \dots, v_n przestrzeni V . Wtedy $F(v_1), \dots, F(v_n)$ jest bazą W (ćwiczenie), co kończy dowód. \square

Izomorfizm $V \rightarrow K^n$ z poprzedniego zadania zapisujemy czasami w nawiasach kwadratowych $[v]_B$ (czytamy: *współrzędne (wektora) v w (bazie) B* (lub *względem bazy B*)).

Przykład 2.25. Przestrzeń $V = \mathbf{R}^2$ ma bazę $B = ((1, 1)^\top, (1, -1)^\top)$ i mamy np. $[(1, 0)^\top]_B = (\frac{1}{2}, \frac{1}{2})^\top$.

Przykład 2.26. Przestrzeń $V = \{P \in \mathbf{R}_3[X] \mid P'(-1) = 0\}$ ma bazę $1, (x+1)^2, (x+1)^3$. To nam zadaje izomorfizm $V \rightarrow \mathbf{R}^3$, np. dla

$$2x^3 + 3x^2 = 2(x+1)^3 - 3(x+1)^2 + 1 \cdot 1$$

$$\text{mamy } [2x^3 + 3x^2]_B = \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix}.$$

⁶Tak naprawdę to działa prawie tak samo w przypadku nieskończenie wymiarowym.

Rozdział 3

Przekształcenia liniowe

3.1 Przekształcenia liniowe

Definicja 3.1. Jeżeli V, W są przestrzeniami liniowymi nad ciałem K , to *przekształceniem liniowym* (lub *funkcją liniową*¹) z V w W nazywamy funkcję $F: V \rightarrow W$, która jest jednocześnie:

- addytywna: $F(v_1 + v_2) = F(v_1) + F(v_2)$, oraz
- jednorodna: $F(\alpha v) = \alpha F(v)$.

Ćwiczenie 3.2. Złożenie funkcji liniowych jest liniowe.

Ćwiczenie 3.3. Funkcja $F: V \rightarrow W$ jest liniowa wtedy i tylko wtedy gdy spełnia $F(\alpha v_1 + v_2) = \alpha F(v_1) + F(v_2)$.

Ćwiczenie 3.4. Jeżeli F jest liniowe, to $F(\vec{0}) = \vec{0}$ i $F(\sum_{k=1}^n \alpha_k v_k) = \sum_{k=1}^n \alpha_k F(v_k)$.

Przykłady 3.5. 1. funkcje liniowe $F: \mathbb{R}^n \rightarrow \mathbb{R}^k$ to dokładnie funkcje postaci F_A , zadane wzorem:

$$F_A(X) = AX = \left(\sum_{j=1}^n a_{ij} x_j \right)_{i=1}^k = \begin{pmatrix} \sum_{j=1}^n a_{1j} x_j \\ \vdots \\ \sum_{j=1}^n a_{kj} x_j \end{pmatrix},$$

¹nie mylić z funkcjami postaci $f(x) = ax + b$

gdzie

$$A = (a_{ij})_{\substack{i=1,\dots,k \\ j=1,\dots,n}} \in M_{k \times n}(\mathbf{R}), \quad x = (x_j)_{j=1}^n = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

2. Analogicznie wyglądają funkcje liniowe $K^n \rightarrow K^k$ dla dowolnego ciała K .
3. $F: \mathbf{R}[X] \rightarrow \mathbf{R}[X]$, $F(P) = P'$, a także takie same funkcje $F: K[X] \rightarrow K[X]$ zadane analogicznym wzorem.
4. $F: \mathbf{R}_3[X] \rightarrow \mathbf{R}_3[X]$, $F(P) = 3P - X^2 \cdot P'$.
5. $F: \mathbf{R}_3[X] \rightarrow \mathbf{R}^2$, $F(P) = \begin{pmatrix} P(1) \\ P'(0) + \pi P(e) \end{pmatrix}$.
6. $F: C(\mathbf{R}) \rightarrow \mathbf{R}$, $F(f) = \int_0^1 f(t) dt$.
7. $F: \mathbf{R}^{\mathbf{N}} \rightarrow \mathbf{R}^{\mathbf{N}}$, $F(a_0, a_1, \dots) = (a_1, a_2, \dots)$ (lewy szift),
8. $F: \mathbf{R}^{\mathbf{N}} \rightarrow \mathbf{R}^{\mathbf{N}}$, $F(a_0, a_1, \dots) = (0, a_0, a_1, \dots)$ (prawy szift),
9. $F: C([0, 2\pi]) \rightarrow \mathbf{C}^{\mathbf{Z}}$, $F(f)(k) = \int_0^{2\pi} f(x) e^{-ikx} dx$.

Sprawdzenie. Ustalmy dowolne dwie funkcje $f_1, f_2 \in C[0, 2\pi]$. Wtedy dla każdego k , z addytywności całki, mamy

$$\begin{aligned} F(f_1 + f_2)(k) &= \int_0^{2\pi} (f_1 + f_2)(x) e^{-ikx} dx \\ &= \int_0^{2\pi} f_1(x) e^{-ikx} dx + \int_0^{2\pi} f_2(x) e^{-ikx} dx \\ &= F(f_1)(k) + F(f_2)(k) \\ &= (F(f_1) + F(f_2))(k) \end{aligned}$$

Skoro jest tak dla każdego k , to znaczy że $F(f_1 + f_2) = F(f_1) + F(f_2)$.

Podobnie dla dowolnego $f \in C[0, 2\pi]$, $\alpha \in \mathbf{R}$ i $k \in \mathbf{Z}$ mamy

$$\begin{aligned} F(\alpha f)(k) &= \int_0^{2\pi} (\alpha f)(x) e^{-ikx} dx \\ &= \alpha \int_0^{2\pi} f(x) e^{-ikx} dx \\ &= \alpha F(f)(k), \end{aligned}$$

więc $F(\alpha f) = \alpha F(f)$.

Uwaga 3.6. Jeżeli V, W są dowolnymi przestrzeniami liniowymi (nad tym samym ciałem), to zbiór $\text{Hom}(V, W)$ wszystkich przekształceń liniowych $V \rightarrow W$ jest przestrzenią liniową. Jeżeli V, W są skończenie wymiarowe, to wymiar $\text{Hom}(V, W)$ to iloczyn wymiarów V i W ².

Jeżeli V jest przestrzenią liniową, to *endomorfizm* V to odwzorowanie liniowe $V \rightarrow V$.

Zbiór $\text{Hom}(V, V)$ endomorfizmów V oznaczamy $\text{End}(V)$.

3.2 Jądro i obraz odwzorowania liniowego

Definicja 3.7. Ustalmy liniowe $F: V \rightarrow W$.

- *Jądro* F to $\ker F := \{v \in V \mid F(v) = \vec{0}\} = F^{-1}[\{\vec{0}\}]$,
- *Obraz* F to $\text{im } F := \{F(v) \mid v \in V\} = F[V] = \{w \in W \mid \exists v \in V F(v) = w\}$ (czasami oznaczany też $\text{rng } F$)

(Rysunek.)

Ćwiczenie 3.8. Złożenie przekształceń liniowych jest liniowe, tzn. jeżeli $F: V_1 \rightarrow V_2$ i $G: V_2 \rightarrow V_3$ są liniowe, to $G \circ F: V_1 \rightarrow V_3$ jest liniowe.

Fakt 3.9. Jeżeli $F: V \rightarrow W$ jest przekształceniem liniowym, to:

- $\ker F \leq V$,
- $\text{im } F \leq W$.

²Ogólnie wymiar $\text{Hom}(V, W)$ to iloczyn wymiarów W i przestrzeni dualnej V^* , o której będzie mowa później

Dowód.

a) $0 \in \ker F$, więc $\ker F \neq \emptyset$.

Z liniowości jeżeli $F(v_1) = F(v_2) = \vec{0}$, to $F(v_1 + v_2) = F(v_1) + F(v_2) = \vec{0} + \vec{0} = \vec{0}$, więc $\ker F$ jest zamknięte na $+$.

Podobnie jeżeli $F(v) = \vec{0}$, to $F(\alpha v) = \alpha F(v) = \alpha \vec{0} = \vec{0}$, więc $\ker F$ jest zamknięte na mnożenie przez skalary.

b) $\vec{0} = F(\vec{0}) \in \operatorname{im} F$, więc $\operatorname{im} F \neq \emptyset$.

Jeżeli $w_1, w_2 \in \operatorname{im} F$, to dla pewnych v_1, v_2 mamy $F(v_k) = w_k$ i wtedy $F(v_1 + v_2) = F(v_1) + F(v_2) = w_1 + w_2$, więc $\operatorname{im} F$ jest zamknięty na sumy.

Jeżeli $w \in \operatorname{im} F$ i $F(v) = w$, to $F(\alpha v) = \alpha F(v) = \alpha w$, więc $\operatorname{im} F$ jest zamknięty na mnożenie przez skalary. \square

Uwaga 3.10. Podobnie, że jeżeli $V' \leq V$, $W' \leq W$ i $F: V \rightarrow W$ jest liniowe, to $F[V'] \leq W$ i $F^{-1}[W'] \leq V$.

Fakt 3.11. Załóżmy że $F: V \rightarrow W$ jest liniowe. Wtedy:

- F jest „na” $\iff \operatorname{im} F = W$,
- F jest 1-1 $\iff \ker F = 0$ ($:= \{\vec{0}\}$).

(Jeżeli $\ker F = 0$, to mówimy że jądro F jest trywialne lub zerowe.)

Dowód. Pierwsza część jest oczywista.

Założmy że F jest 1-1. Wtedy $F(\vec{0}) = \vec{0}$, więc $\vec{0} \in \ker F$ i z różnowartościowości, dla $v \neq \vec{0}$ mamy $F(v) \neq F(\vec{0}) = \vec{0}$, czyli $v \notin \ker F$. Stąd $\ker F = \{\vec{0}\}$.

Z drugiej strony, jeżeli $\ker F = \{\vec{0}\}$, to dla dowolnych $v_1 \neq v_2$ mamy

$$F(v_1) - F(v_2) = F(v_1 - v_2) \neq \vec{0}$$

(bo $v_1 - v_2 \neq \vec{0}$, więc $v_1 - v_2 \notin \ker F$).

Stąd $F(v_1) \neq \vec{0} + F(v_2) = F(v_2)$. \square

Przykłady 3.12. a) Niech $A = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}$, $F_1 = F_A \in \operatorname{End}(\mathbf{R}^3)$.

$$\operatorname{im} F_A = \operatorname{Lin} \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix} \right\} = \operatorname{Lin} \left(\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right),$$

$$\ker F_A \text{ składa się z } \begin{pmatrix} x \\ y \\ z \end{pmatrix} \text{ spełniających } \begin{cases} x + 4y + 7z = 0 \\ 2x + 5y + 8z = 0 \\ 3x + 6y + 9z = 0 \end{cases} . \text{ Trzecie}$$

równanie jest równoważne $x + 2y + 3z = 0$, czyli odejmując stronami

$$\text{równoważnie } \begin{cases} 2y + 4z = 0 \\ y + 2z = 0 \\ x + 2y + 3z = 0 \end{cases} . \text{ Pierwsze równanie wynika z dru-}$$

$$\text{giego, skąd widać że } \ker F_A = \operatorname{Lin} \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \times \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \operatorname{Lin} \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}.$$

$$\text{b) } F_2: \mathbf{R}_3[X] \rightarrow \mathbf{R}^2, F(P) = \begin{pmatrix} P'(2) \\ P(-1) \end{pmatrix}$$

$$\ker F_2 = \{ P \in \mathbf{R}_3 \mid P'(2) = P(-1) = 0 \}$$

$$\operatorname{im} F_2 = \mathbf{R}^2, \text{ bo } \begin{pmatrix} 1 \\ -1 \end{pmatrix} = F(X), \begin{pmatrix} 0 \\ 1 \end{pmatrix} = F(1) \in \operatorname{im} F_2 \leq \mathbf{R}^2.$$

$$\text{c) } F_3: \mathbf{R}_{100}[X] \rightarrow \mathbf{R}, F_3(P) = \int_{-1}^1 P(t) dt.$$

$$\operatorname{im} F_3 = \mathbf{R}, \text{ bo } 0 = \vec{0} \neq F_3(X^2) \in \operatorname{im} F_3 \leq \mathbf{R}$$

$$\ker F_3 = \{ P \in \mathbf{R}_{100}[X] \mid \int_{-1}^1 P(t) dt = 0 \}.$$

$$\text{d) Analogicznie jak wyżej możemy zdefiniować } F_4: C(\mathbf{R}) \rightarrow \mathbf{R}, F_4(f) = \int_{-1}^1 f(t) dt, \text{ obraz jest ten sam, a jądro ma podobny opis.}$$

$$\text{e) } F_5: \mathbf{R}^{\mathbf{N}} \rightarrow \mathbf{R}^{\mathbf{N}}, F_5(a_0, \dots, a_n, \dots) = (a_1, \dots, a_n, \dots).$$

$$\operatorname{im} F_5 = \mathbf{R}^{\mathbf{N}}, \ker F_5 = \{ (a, 0, 0, 0, \dots) \mid a \in \mathbf{R} \}$$

$$\text{f) } F_6: \mathbf{R}^{\mathbf{N}} \rightarrow \mathbf{R}^{\mathbf{N}}, F_6(a_0, \dots, a_n, \dots) = (0, a_0, a_1, \dots, a_n, \dots).$$

$$\ker F_6 = \{ \vec{0} \}, \operatorname{im} F_6 = \{ (a_n)_{n \in \mathbf{N}} \in \mathbf{R}^{\mathbf{N}} \mid a_0 = 0 \}$$

3.3 Twierdzenia o rzędzie i o indeksie

Stwierdzenie 3.13. Załóżmy że V jest dowolną przestrzenią liniową, a $A, B \subseteq V$ są rozłączne i $A \cup B$ jest liniowo niezależny. Wtedy $\text{Lin}(A) \cap \text{Lin}(B) = \{\vec{0}\}$

Dowód. Ćwiczenie. □

Definicja 3.14. Rząd przekształcenia liniowego F to $\text{rk } F := \dim \text{im } F$.

Twierdzenie 3.15 (twierdzenie o rzędzie). Jeżeli $F: V \rightarrow W$ jest liniowe, to

$$\dim V = \dim \ker F + \dim \text{im } F = \dim \ker F + \text{rk } F.$$

Wniosek 3.16 (twierdzenie o indeksie). Jeżeli $\dim V < \infty$, to odejmując stronami dostajemy stąd:

$$\dim \ker F = \dim V - \dim \text{im } F \quad \dim \text{im } F = \dim V - \dim \ker F.$$

Przykład 3.17. Chcemy wyznaczyć wymiar $V = \{P \in \mathbf{R}_{50}[X] \mid \int_{-1}^1 e^{-t^2} P(t) dt = 0\}$

Weźmy $G: \mathbf{R}_{50}[X] \rightarrow \mathbf{R}$, zadane wzorem $G(P) = \int_{-1}^1 e^{-t^2} P(t) dt$. Wtedy $\ker G = V$.

$G(1) = \int_{-1}^1 e^{-t^2} dt > 0$, więc $\text{im}(G) = \mathbf{R}$, zatem $\dim V = \dim \ker G = \dim \mathbf{R}_{50}[X] - \dim \mathbf{R} = 51 - 1 = 50$.

dowód twierdzenia o rzędzie. Wybierzmy dowolną bazę A przestrzeni $\ker F \leq V$. Wtedy A jest lnz w V , więc rozszerza się do pewnej bazy C przestrzeni V . Oznaczmy $B := C \setminus A$.

Pokażemy że F jest 1-1 na B i $F[B]$ jest bazą $\text{im } F$. To skończy dowód, ponieważ wtedy

$$\dim V = |C| = |A| + |B| = |A| + |F[B]| = \dim \ker F + \dim \text{im } F.$$

Istotnie, dla dowolnego $v \in V$ mamy (dla pewnych α_b)

$$v = \sum_{a \in A} \alpha_a a + \sum_{b \in B} \beta_b b,$$

więc

$$F(v) = \sum_{a \in A} \alpha_a F(a) + \sum_{b \in B} \beta_b F(b),$$

przy czym pierwsza z sum powyżej jest zerowa (bo dla $a \in A$ mamy $F(a) = \vec{0}$), więc $F(v) = \sum_{b \in B} \beta_b F(b)$, czyli $F[B]$ rozpina $\text{im } F$.

Zauważmy teraz, że ze stwierdzenia mamy $\text{Lin } B \cap \text{Lin } A = \text{Lin } B \cap \ker F = \{0\}$. W szczególności $\ker F|_{\text{Lin } B} = \{0\}$. Stąd F jest 1-1 na $\text{Lin } B$ (a więc i na B).

Pokażemy liniową niezależność $F[B]$. Niech $\sum_{b \in B} \beta_b F(b) = 0$. Wtedy $F(\sum_b \beta_b b) = 0$, czyli z różnowartościowości $\sum_b \beta_b b = 0$, a więc wszystkie $\beta_b = 0$. \square

Wniosek 3.18. Załóżmy że $F: V \rightarrow W$ jest przekształceniem liniowym i $\dim V = \dim W < \infty$. Wtedy następujące warunki są równoważne:

- $\ker F = 0$,
- F jest „na”,
- F jest 1-1,
- F jest izomorfizmem.

Dowód. $\ker F = 0 \iff F$ jest 1-1: było.

Jeżeli $\ker F = 0$, to $\dim \ker F = 0$, czyli $\dim \text{im } F = \dim V - 0 = \dim W$. Ale $\text{im } F \leq W$, więc wtedy $\text{im } F = W$.

Jeżeli F jest „na”, to $\text{im } F = W$, czyli $\dim \text{im } F = \dim W = \dim V$, więc $\dim \ker F = \dim V - \dim \text{im } F = \dim V - \dim V = 0$, więc $\ker F = \{\vec{0}\}$.

Stąd F jest „na” wtedy i tylko wtedy gdy jest 1-1, więc wówczas jest też izomorfizmem. \square

Wniosek 3.19. Jeżeli V jest skończenie wymiarowa³ i $F \in \text{End}(V)$ (tzn. jest endomorfizmem V , czyli liniowym odwzorowaniem $V \rightarrow V$) to następujące warunki są równoważne:

- F jest różnowartościowa,
- F jest „na”,
- F jest izomorfizmem.

Definicja 3.20. Takie odwzorowanie, tzn. izomorfizm, który jest jednocześnie endomorfizmem, nazywamy *automorfizmem*. Zbiór automorfizmów przestrzeni liniowej oznaczamy $\text{Aut}(V)$ lub (częściej) $\text{GL}(V)$.

³Założenie skończonego wymiaru jest istotne: lewy i prawy szift na $\mathbb{R}^{\mathbb{N}}$ pokazują, że przekształcenie liniowe przestrzeni nieskończenie wymiarowej w siebie samą może być „na”, nie będąc 1-1 (lewy szift), może też być 1-1, nie będąc „na” (prawy szift).

Wniosek 3.21. *Jeżeli mamy „krótki ciąg dokładny (przestrzeni liniowych)”*⁴

$$V_1 \rightarrow V_2 \rightarrow V_3,$$

to znaczy przekształcenia liniowe $F_1: V_1 \rightarrow V_2$ i $F_2: V_2 \rightarrow V_3$, takie że F_1 jest 1-1, F_2 jest „na” i $\text{im } F_1 = \ker F_2$, to $\dim V_2 = \dim V_1 + \dim V_3$.

Dowód. Skoro F_1 jest 1-1 i $\text{im } F_1 = \ker F_2$, to F_1 zadaje izomorfizm V_1 i $\text{im } F_1 = \ker F_2$, czyli $\dim V_1 = \dim \ker F_2$.

Z twierdzenia o rzędzie dla F_2 dostajemy więc

$$\dim V_2 = \dim \ker F_2 + \dim \text{im } F_2 = \dim V_1 + \dim V_3. \quad \square$$

⁴Ogólnie ciąg dokładny to taki ciąg przekształceń $\dots V_{-1} \rightarrow V_0 \rightarrow V_1 \rightarrow V_2 \rightarrow V_3 \rightarrow V_4 \rightarrow \dots$, że jądro każdej strzałki jest obrazem poprzedniej. Krótki ciąg dokładny $V_1 \rightarrow V_2 \rightarrow V_3$ można zapisać jako taki ciąg dokładny, że $V_k = 0$ dla $k < 1$ i dla $k > 3$, tzn. $\dots \rightarrow 0 \rightarrow 0 \dots 0 \rightarrow V_1 \rightarrow V_2 \rightarrow V_3 \rightarrow 0 \rightarrow 0 \dots \rightarrow 0 \rightarrow \dots$: w zapisie pomijamy zera.

Rozdział 4

Konstrukcje przestrzeni liniowych

4.1 Suma prosta

Definicja 4.1. Niech V, W będą przestrzeniami liniowymi. Na zbiorze $V \times W$ określamy działania:

- $(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2),$
- $\alpha(v, w) = (\alpha v, \alpha w).$

Powstałą przestrzeń liniową nazywamy *produktem* lub *sumą prostą* V i W , oznaczamy $V \times W$ lub $V \oplus W$.¹

Przykłady 4.2. 1. $K \times K = K^2,$

2. $K \times (K \times K) = K \times K^2 = K^3,$ ²

3. jeżeli $F \subseteq V \times W$ jest funkcją $V \rightarrow W$, to F jest liniowe wtedy i tylko wtedy gdy $F \leq V \times W$ (ćwiczenie).

Fakt 4.3. $\dim(V \oplus W) = \dim V + \dim W$

Dowód. Mamy krótki ciąg dokładny:

$$V \rightarrow V \oplus W \rightarrow W,$$

¹Możemy rozważać też nieskończone produkty i sumy proste przestrzeni liniowych. Wtedy te dwa pojęcia się rozjeżdżają (o tym więcej może być na konwersatorium).

²Teoriomnogościowo są to trochę inne obiekty, ale w algebrze liniowej zwykle je utożsamiamy.

więc $\dim V \oplus W = \dim V + \dim W$.

Nieco bardziej szczegółowo: mamy naturalne odwzorowanie $\pi_W: V \oplus W \rightarrow W$ (rzut), tzn. $\pi(v, w) = w$.

Łatwo sprawdzić, że π_W jest liniowe i $\text{im } \pi_W = W$, więc $\dim(V \oplus W) = \dim \ker \pi_W + \dim \text{im } \pi_W = \dim \ker \pi_W + \dim W$.

Z drugiej strony $\ker \pi_W$ jest izomorficzne z V przez odwzorowanie $V \rightarrow \ker \pi_W$, $v \mapsto (v, 0)$, więc $\dim \ker \pi_W = \dim V$, co daje tezę. \square

inny dowód. Niech v_1, \dots, v_n będzie bazą V , a w_1, \dots, w_m będzie bazą W . Wtedy $(v_1, 0), (v_2, 0), \dots, (v_n, 0), (0, w_1), \dots, (0, w_m)$ jest bazą $V \oplus W$ (ćwiczenie). \square

Fakt 4.4 (/definicja). *Jeżeli mamy liniowe $F_1: V_1 \rightarrow W_1$, $F_2: V_2 \rightarrow W_2$, to nam daje liniowe $F_1 \oplus F_2: V_1 \oplus V_2 \rightarrow W_1 \oplus W_2$, zadane wzorem $(F_1 \oplus F_2)(v_1, v_2) = (F_1(v_1), F_2(v_2))$.*

Dowód. Ćwiczenie. \square

Fakt 4.5. *Dla dowolnych $F_1: V_1 \rightarrow W_1$, $F_2: V_2 \rightarrow W_2$, $G_1: W_1 \rightarrow U_1$, $G_2: W_2 \rightarrow U_2$ (w skrócie: $V_1 \xrightarrow{F_1} W_1 \xrightarrow{G_1} U_1$, $V_2 \xrightarrow{F_2} W_2 \xrightarrow{G_2} U_2$) zachodzi*

$$(G_1 \oplus G_2) \circ (F_1 \oplus F_2) = ((G_1 \circ F_1) \oplus (G_2 \circ F_2)).$$

Dowód. Ćwiczenie (wystarczy rozpisać lewą i prawą stronę na dowolnym $(v_1, v_2) \in V_1 \oplus V_2$). \square

Ćwiczenie 4.6. *Jeżeli $V_1 \cong V_2$ i $W_1 \cong W_2$, to $V_1 \oplus W_1 \cong V_2 \oplus W_2$.*

Ćwiczenie 4.7 (/definicja). *Jeżeli $W, U \leq V$ są podprzestrzeniami takimi że $W \cap U = \{\vec{0}\}$, to $W \oplus U \cong W + U$. Jeżeli ponadto $V = W + U$, to mówimy że V jest sumą prostą W i U (i utożsamiamy $V = W \oplus U$).³*

4.2 Przestrzeń dualna

Definicja 4.8. *Jeżeli V jest przestrzenią liniową nad K , to definiujemy przestrzeń dualną do V jako*

$$V^* = V' = \text{Hom}(V, K) := \{f: V \rightarrow K \mid f \text{ jest liniowe}\}.$$

Elementy V^* nazywamy *funkcjonalami* na V .

³O tym będzie więcej w rozdziale 7.

Lemat 4.9. V^* jest przestrzenią liniową, podprzestrzenią K^V (przestrzeni wszystkich funkcji $V \rightarrow K$).

Dowód. Funkcja zerowa jest liniowa, więc V^* jest niepuste.

Niech $f_1, f_2 \in V^*$. Pokażemy że $f_1 + f_2 \in V^*$.

- $f_1 + f_2$ jest addytywne: jeżeli $v_1, v_2 \in V$, to:

$$\begin{aligned} (f_1 + f_2)(v_1 + v_2) &= f_1(v_1 + v_2) + f_2(v_1 + v_2) \\ &= f_1(v_1) + f_1(v_2) + f_2(v_1) + f_2(v_2) \\ &= f_1(v_1) + f_2(v_1) + f_1(v_2) + f_2(v_2) \\ &= (f_1 + f_2)(v_1) + (f_1 + f_2)(v_2). \end{aligned}$$

- $f_1 + f_2$ jest jednorodny: jeżeli $v \in V$ i $\alpha \in K$, to:

$$\begin{aligned} (f_1 + f_2)(\alpha v) &= f_1(\alpha v) + f_2(\alpha v) \\ &= \alpha f_1(v) + \alpha f_2(v) \\ &= \alpha(f_1(v) + f_2(v)) \\ &= \alpha((f_1 + f_2)(v)) \end{aligned}$$

Zamkniętość V^* na mnożenie przez skalary — ćwiczenie. □

Fakt 4.10. Jeżeli $\dim V < \infty$, to $\dim V = \dim V^*$.⁴

Dowód. Niech b_1, \dots, b_n będzie bazą V . Dla $k = 1, \dots, n$ definiujemy $b_k^*(v) := \alpha_k$, gdzie $v = \sum_{j=1}^n \alpha_j b_j$.

Ponieważ b_k tworzą bazę, łatwo sprawdzić że b_k^* są dobrze określone i liniowe. Układ b_1^*, \dots, b_n^* nazywamy *bazą dualną* do b_1, \dots, b_n .

Sprawdźmy, że faktycznie jest to baza.

Zauważmy że każde b_k^* spełnia $b_k^*(b_j) = \begin{cases} 1 & k = j \\ 0 & k \neq j \end{cases}$. Stąd dla każdego j

zachodzi:

$$\sum_k \alpha_k b_k^*(b_j) = \alpha_j. \quad (*)$$

Jeżeli $v^* = \sum_k \alpha_k b_k^* = 0$, to z (*) dostajemy dla każdego k , $\alpha_k = v^*(b_k) = 0$, czyli b_k^* są liniowo niezależne.

⁴Przy założeniu pewnika wyboru dla nieskończenie wymiarowych przestrzeni zachodzi $\dim V < \dim V^*$. Bez niego mogą istnieć nieskończenie wymiarowe V takie że $V^* = \{0\}$.

Z drugiej strony jeżeli $v^* \in V^*$, to dla $\alpha_k = v^*(b_k)$ mamy $v^* = \sum_k \alpha_k b_k^*$.⁵ Istotnie, dla dowolnego $v = \sum_j \beta_j b_j$ mamy

$$\begin{aligned}
 \left(\sum_k \alpha_k b_k^* \right) (v) &= \sum_k \alpha_k b_k^* \left(\sum_j \beta_j b_j \right) \\
 &= \sum_k \sum_j \alpha_k \beta_j b_k^*(b_j) \\
 &= \sum_k \alpha_k \beta_k \\
 &= \sum_k v^*(b_k) \beta_k \\
 &= \sum_k v^*(\beta_k b_k) \\
 &= v^* \left(\sum_k \beta_k b_k \right) \\
 &= v^*(v).
 \end{aligned}$$

□

Ćwiczenie 4.11. Jeżeli $\dim V > 1$ i $v \in V$, to nie istnieje „funkcjonał dualny” do v . (Vide zadanie 20 z listy 2.)

Uwaga 4.12. W analizie funkcjonalnej V^* definiuje się jako przestrzeń liniową złożoną tylko z tych funkcjonałów, które są ciągłe (względem ustalonych topologii na V i K). Wtedy V^* może mieć nawet mniejszy wymiar niż V , a nawet są nieskończone wymiarowe V takie że $\dim V^* = 0$. (O tym możemy porozmawiać trochę na konwersatorium.)

Fakt 4.13 (/definicja). Jeżeli $F: V \rightarrow W$ jest liniowe, to odwzorowanie $F^*: W^* \rightarrow V^*$ zadane wzorem $F^*(f)(v) = f(F(v))$ jest liniowe. Nazywamy je przekształceniem dualnym do F

Dowód. Ćwiczenie. □

Fakt 4.14. Jeżeli $F_1: V_1 \rightarrow V_2$ i $F_2: V_2 \rightarrow V_3$ są liniowe, to $(F_2 \circ F_1)^* = F_1^* \circ F_2^*$.

$$V_1 \xrightarrow{F_1} V_2 \xrightarrow{F_2} V_3$$

$$V_1^* \xleftarrow{F_1^*} V_2^* \xleftarrow{F_2^*} V_3^*$$

⁵Kiedy $\dim V$ jest nieskończony, prawa strona ostatniego równania zwykle nie ma sensu.

Dowód. $F_2 \circ F_1$ jest odwzorowaniem $V_1 \rightarrow V_3$, a więc $(F_2 \circ F_1)^*$ jest odwzorowaniem $V_3^* \rightarrow V_1^*$. Weźmy dowolny $f \in V_3^*$. Sprawdzimy że $(F_2 \circ F_1)^*(f) = F_1^* \circ F_2^*(f)$. Ustalmy dowolny $v \in V_1$. Wtedy:

$$\begin{aligned} (F_2 \circ F_1)^*(f)(v) &= f(F_2 \circ F_1(v)) \\ &= f(F_2(F_1(v))) \\ &= F_2^*(f)(F_1(v)) \\ &= F_1^*(F_2^*(f))(v) \\ &= F_1^* \circ F_2^*(f)(v). \end{aligned}$$

Ponieważ $v \in V_1$ był dowolny, istotnie zachodzi $(F_2 \circ F_1)^*(f) = F_1^* \circ F_2^*(f)$, co należało dowieść. \square

Ćwiczenie 4.15. Jeżeli $F: V \rightarrow W$ jest izomorfizmem, to F^* też nim jest. W szczególności jeżeli $V \cong W$, to $V^* \cong W^*$.

4.3 Przestrzeń ilorazowa

Definicja 4.16. Jeżeli $W \leq V$ są przestrzeniami liniowymi i $v \in V$, warstwą v względem W nazywamy zbiór:

$$v + W = \{v + w \mid w \in W\}.$$

Zbiór wszystkich warstw W oznaczamy przez V/W .

Fakt 4.17. $v_1 + W = v_2 + W \iff v_1 - v_2 \in W$

Dowód. Załóżmy że $v_1 + W = v_2 + W$. Wtedy z definicji $v_1 + 0 \in v_2 + W$, czyli dla pewnego $w \in W$ mamy $v_1 + 0 = v_2 + w$, czyli $v_1 - v_2 = w \in W$.

Z drugiej strony, jeżeli $v_1 - v_2 \in W$, to dla każdego w mamy

$$v_1 + w = (v_2 + (v_1 - v_2)) + w = v_2 + ((v_1 - v_2) + w) \in v_2 + W,$$

czyli $v_1 + W \subseteq v_2 + W$. Podobnie $v_2 + W \subseteq v_1 + W$, czyli te dwie warstwy są równe. \square

Wniosek 4.18. Zbiór V/W warstw W w V to zbiór ilorazowy V/\sim , gdzie

$$v_1 \sim v_2 \iff v_1 - v_2 \in W.$$

Przykłady 4.19. 1. Jeżeli $V = \mathbf{R}^2$ i $W = \text{Lin}(1, 1)^\top$, to elementy V/W to proste równoległe do W .

2. Podobnie jeżeli $V = \mathbf{R}^3$ i W jest prostą lub płaszczyzną przechodzącą przez 0, to elementy V/W to proste lub płaszczyzny (odpowiednio) równoległe do W .

Definicja 4.20. Jeżeli $W \leq V$ są przestrzeniami liniowymi, to zbiór warstw V/W nazywamy *przestrzenią ilorazową* (V nad W lub V przez W) z działaniami zadanymi następująco:

- $\vec{0}_{V/W} = \vec{0}_V + W$,
- $(v_1 + W) +_{V/W} (v_2 + W) = (v_1 +_V v_2) + W$,
- $\alpha \cdot_{V/W} (v + W) = (\alpha \cdot_V v) + W$.

Przykład 4.21. Jeżeli $V = \mathbf{R}^2$ i $W = \text{Lin}(1, 1)^\top$, zaś $\ell_1 \in V/W$ jest prostą zadaną równaniem $X = t(1, 1)^\top + (2, 3)^\top$ i $\ell_2 \in V/W$ jest zadaną równaniem $X = t(1, 1)^\top + (-3, 0)^\top$, to w V/W $\ell_1 + \ell_2$ jest prostą zadaną równaniem $X = t(1, 1)^\top + (-1, 3)^\top$.

Uwaga 4.22. Wyłączając mnożenie przez 0, definicje $+_{V/W}$ i $\cdot_{V/W}$ to dokładnie suma kompleksowa i mnożenie zbioru przez skalar, tzn. dla dowolnych warstw $\pi_1, \pi_2 \in V/W$ mamy

- $\pi_1 + \pi_2 = \{v_1 + v_2 \mid v_1 \in \pi_1, v_2 \in \pi_2\}$,
- $\alpha \pi = \{\alpha v \mid v \in \pi\}$.

Wyjątkowo $0 \cdot \pi = W \neq \{0\} = \{0 \cdot v' \mid v' \in \pi\}$ (chyba że $W = \{0\}$).

Fakt 4.23. Powyższe działania są dobrze określone i to jest struktura przestrzeni liniowej.

Dowód. $+_{V/W}$ jest dobrze określone: weźmy $v_1 + W = v'_1 + W$ i $v_2 + W = v'_2 + W$. Wtedy $v_1 - v'_1, v_2 - v'_2 \in W$, więc dodając stronami $(v_1 + v_2) - (v'_1 + v'_2) \in W + W = W$, czyli $(v_1 + v_2) + W = (v'_1 + v'_2) + W$, więc $+_{V/W}$ jest dobrze określone.

$\cdot_{V/W}$ dobrze określone — podobnie.

$0 + W + v + W = v + W, (v + W) + ((-v) + W) = 0 + W$ — oczywiste.

Przemienność, łączność i rozdzielność w V/W wynikają łatwo z tych samych własności w V . Na przykład:

$$\begin{aligned}
 \alpha(v_1 + W + v_2 + W) &= \alpha((v_1 + v_2) + W) \\
 &= (\alpha(v_1 + v_2)) + W \\
 &= (\alpha v_1 + \alpha v_2) + W \\
 &= (\alpha v_1) + W + (\alpha v_2) + W \\
 &= \alpha(v_1 + W) + \alpha(v_2 + W)
 \end{aligned}
 \quad \square$$

Definicja 4.24. Odwzorowanie liniowe nazywamy *epimorfizmem* jeżeli jest „na”, a *monomorfizmem* jeżeli jest 1-1.

(W szczególności izomorfizm = epimorfizm i monomorfizm.)

Stwierdzenie 4.25. $\dim V = \dim V/W + \dim W$, czyli $\dim V/W = \dim V - \dim W$, o ile $\dim V < \infty$.

Dowód. Zdefiniujmy $F(v) = v + W$. Cel: F jest liniowe, $\ker F = W$, $\operatorname{im} F = V/W$. Wtedy będziemy mieli krótki ciąg dokładny $W \rightarrow V \rightarrow V/W$, skąd wyniknie teza.

- F jest „na” V/W : oczywiste
- $\ker F = W$: jeżeli $v \in V$ jest takie że $v + W = 0 + W$, to z Faktu wcześniej wiemy, że $v - 0 = v \in W$.
- F jest addytywne: $F(v_1 + v_2) = (v_1 + v_2) + W$, co z definicji równa się $(v_1 + W) + (v_2 + W)$.
- Jednorodność — podobnie. \square

Stwierdzenie 4.26 (Twierdzenie o izomorfizmie). Jeżeli $F: V \rightarrow W$ jest

odwzorowaniem liniowym, to $\operatorname{im} F \cong V/\ker F$.

$$\begin{array}{ccc}
 V & \xrightarrow{F} & \operatorname{im} F \\
 \downarrow & \nearrow \bar{F} & \\
 V/\ker F & &
 \end{array}$$

Dowód. Zadajmy odwzorowanie $\bar{F}: V/\ker F \rightarrow \operatorname{im} F$ wzorem $\bar{F}(v + W) = F(v)$.

- \bar{F} jest dobrze określone: jeżeli $v_1 + \ker F = v_2 + \ker F$, to $v_1 - v_2 \in \ker F$, więc $F(v_1) = F(v_2) + F(v_1 - v_2) = F(v_2 + v_1 - v_2) = F(v_1)$.

- liniowość \bar{F} wynika łatwo z liniowości F .
- \bar{F} jest 1-1: jeżeli $v + \ker F \neq 0 + \ker F$, to $v \notin \ker F$, więc $\bar{F}(v + \ker F) = F(v) \neq 0$, czyli $\ker \bar{F}$ jest trywialne.
- \bar{F} jest „na”: jeżeli $w \in \operatorname{im} F$, to $w = F(v)$ i wtedy $w = \bar{F}(v + \ker F)$. \square

4.4 Przestrzeń bidualna

Stwierdzenie 4.27. Dla dowolnej przestrzeni liniowej V mamy odwzorowanie liniowe $\Phi: V \rightarrow V^{**}$ (przestrzeń dualna do V^* , przestrzeń bidualna), zadane wzorem $\Phi(v)(f) = f(v)$ (dla $f \in V^*$).

Dowód. Trzeba sprawdzić że $\Phi(v) \in V^{**}$, tzn. że jest addytywne i jedno-rodne.

Istotnie, jeżeli $\Phi(v)(f_1 + f_2) = (f_1 + f_2)(v)$, co z definicji dodawania w V^* równa się $f_1(v) + f_2(v) = \Phi(v)(f_1) + \Phi(v)(f_2)$.

Podobnie $\Phi(v)(\alpha f) = (\alpha f)(v) = \alpha(f(v)) = \alpha\Phi(v)(f)$.

Następnie musimy sprawdzić, że Φ jest liniowe.

$$\Phi(v_1 + v_2)(f) = f(v_1 + v_2) = f(v_1) + f(v_2) = \Phi(v_1)(f) + \Phi(v_2)(f)$$

$$\Phi(\alpha v)(f) = f(\alpha v) = \alpha(f(v)) = \alpha(\Phi(v)(f)).$$

\square

Stwierdzenie 4.28. Jeżeli $\dim V < \infty$, to Φ zadaje izomorfizm $V \cong V^{**}$.⁶

Dowód. Różnowartościowość: ustalmy niezerowy $v \in V$. Wtedy v jest liniowo niezależny, rozszerza się więc do bazy $b_1 = v, b_2, \dots, b_n$ przestrzeni V . Weźmy bazę dualną $b_1^*, \dots, b_n^* \in V^*$.

Wtedy $\Phi(v)(b_1^*) = \Phi(b_1)(b_1^*) = b_1^*(b_1) = 1$, więc $\Phi(b_1) \neq \vec{0}$. Stąd $\ker \Phi = \{0\}$, czyli Φ jest 1-1.

Skoro $\dim V < \infty$, to z wcześniejszego faktu wiemy że $\dim V = \dim V^*$, więc analogicznie $\dim V = \dim V^* = \dim V^{**}$.

Skoro $\Phi: V \rightarrow V^{**}$ jest liniowe i 1-1 i $\dim V = \dim V^{**}$, to z wniosku z twierdzenia o rzędzie wynika, że Φ jest izomorfizmem. \square

⁶Ten sam dowód pokazuje, że Φ zawsze jest różnowartościowe, ale w przypadku nieskończonego wymiaru wymiar V^{**} jest dużo większy niż wymiar V .⁷

⁷Zakładając pewnik wyboru: gdy V nie ma bazy, to Φ może nie być 1-1 i V^{**} może nie mieć tak dużego wymiaru.

Uwaga 4.29. Kiedy $\dim V < \infty$, to mamy $V \cong V^*$, ale ten izomorfizm nie jest „naturalny” — intuicyjnie: zależy od wyboru bazy, a zazwyczaj nie mamy naturalnego wyboru bazy. Φ natomiast jest naturalny⁸: dla dowolnego $F: V \rightarrow W$, pamiętając że F^{**} to odwzorowanie dualne do F^* , czyli zadane wzorem $F^{**}(v^{**})(f) = v^{**}(F^*(f))$, mamy przemienny diagram:

$$\begin{array}{ccc} V & \xrightarrow{F} & W \\ \downarrow \Phi_V & & \downarrow \Phi_W \\ V^{**} & \xrightarrow{F^{**}} & W^{**}, \end{array}$$

to znaczy dla każdego $v \in V$ mamy $\Phi_W(F(v)) = F^{**}(\Phi_V(v))$, dla dowolnego $f \in W^*$ mamy:

$$\begin{aligned} F^{**}(\Phi_V(v))(f) &= \Phi_V(v)(F^*(f)) \\ &= F^*(f)(v) \\ &= f(F(v)) \\ &= \Phi_W(F(v))(f) \end{aligned}$$

(o tym może być więcej na konwersatorium).

⁸Nawet gdy nie jest izomorfizmem!

Rozdział 5

Macierze i układy równań

5.1 Macierze układów równań

Konwencja:

- ij -ty współczynnik (współrzędna) macierzy to współczynnik w j -tej kolumnie, w i -tym wierszu,
- przez I oznaczamy macierz identycznościową (rozmiaru wynikającego z kontekstu),
- a przez E_{ij} oznaczamy macierz kwadratową, ij -ta współrzędna jest równa 1, a pozostałe współrzędne są zerowe,
- mnożenie macierzy $M \in M_{n \times m}(K), N \in M_{m \times p}(K)$ odbywa się zgodnie ze wzorem podanym w ubiegłym semestrze, tzn. MN jest macierzą $n \times p$, której ij -ty współczynnik to $\sum_{k=1}^m m_{ik}n_{kj}$.
- Transpozycja macierzy $A \in M_{m \times n}(K)$ o współczynnikach (a_{ij}) to macierz $A^T \in M_{n \times m}(K)$ o współczynnikach $b_{ij} = a_{ji}$. Innymi słowy transpozycja zamienia kolumny z wierszami.

Ćwiczenie 5.1. Jeżeli M jest dowolną macierzą, to $e_i^T M e_j$ to ij -ty wyraz M . W szczególności jeżeli macierze $M, N \in M_{n \times m}(K)$ spełniają $v^T M w = v^T N w$ dla każdych $v \in K^m, w \in K^n$, to $M = N$.

Rozważamy układy równań

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= y_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= y_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= y_m \end{cases} \quad (1)$$

Definicja 5.2. • a_{ij} — współczynniki,

- y_1, \dots, y_m — dane,
- x_1, \dots, x_n — niewiadome/szukane.

Ten sam układ równań inaczej (w postaci wektorowej):

$$x_1 \overbrace{\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}}^{A_1} + x_2 \overbrace{\begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}}^{A_2} + \dots + x_n \overbrace{\begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}}^{A_n} = \overbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}}^Y$$

$$x_1 A_1 + x_2 A_2 + \dots + x_n A_n = Y \quad (2)$$

Stąd widać, że rozwiązanie układu = przedstawienie Y jako kombinacji liniowej A_1, \dots, A_n . W postaci macierzowej:

$$\overbrace{(A_1 \ A_2 \ \dots \ A_n)}^A \overbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}^X = Y,$$

w skrócie

$$AX = Y \quad (3)$$

Definicja 5.3. • A — macierz główna układu (1)

- $(A|Y)$ — macierz rozszerzona układu (1)

Definicja 5.4. Mówimy że układ jest *jednorodny* jeżeli $Y = 0$.

Uwaga 5.5. Następujące warunki są równoważne:

- Układ (1) ma rozwiązanie,
- $Y \in \text{Lin}(A_1, \dots, A_n)$,
- $\text{Lin}(A_1, \dots, A_n) = \text{Lin}(A_1, \dots, A_n, Y)$,

- $\dim \text{Lin}(A_1, \dots, A_n) = \dim \text{Lin}(A_1, \dots, A_n, Y)$.

Definicja 5.6. *Rząd* A to $\text{rk} A = \dim \text{Lin}(A_1, \dots, A_n) =$ maksymalna liczba lnz układu kolumn A .
($\text{Lin}(A_1, \dots, A_n)$ to $\text{im } F_A$, więc $\text{rk} A = \text{rk } F_A$.)

Z uwagi powyżej łatwo wynika następujące twierdzenie.

Twierdzenie 5.7 (Kroneckera-Capelliego). *Układ (3) ma rozwiązanie $\iff \text{rk} A = \text{rk}(A|Y)$*

5.2 Operacje wierszowe i kolumnowe a rząd macierzy

Definicja 5.8. *Operacje kolumnowe na macierzach:*

(K1) Dodaj skalarną wielokrotność kolumny do *innej* kolumny.

(K2) Pomnóż kolumnę przez *niezerowy* skalar.

(K3) Zamień miejscami dwie kolumny.

Analogicznie definiujemy *operacje wierszowe* W1-3 na macierzach.

Przykład 5.9.

$$\begin{pmatrix} 3 & 5 & 1 \\ -2 & 3 & 3 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 5-2 \cdot 3 & 1 \\ -2 & 3-2 \cdot (-2) & 3 \\ 1 & 1-2 \cdot 1 & 0 \end{pmatrix}$$

Ćwiczenie 5.10. 1. Operacja (K1) odpowiada pomnożeniu macierzy z prawej przez macierz postaci $I + \alpha E_{ij}$ (gdzie i to numer kolumny, której wielokrotność dodajemy do j -tej kolumny),

2. Operacje K2-3 też odpowiadają pomnożeniu z prawej przez pewne macierze.

Operacje W1-3 odpowiadają mnożeniu przez te same macierze, tylko z lewej.

Fakt 5.11. *Operacje kolumnowe K1-3 nie zmieniają obrazu $F_A: K^n \rightarrow K^m$.*

Dowód. K1: Chcemy pokazać, że

$$\operatorname{im} F_A = V = \operatorname{Lin}(A_1, \dots, A_n) = \operatorname{Lin}(A_1, \dots, A_{i-1}, A_i + \alpha A_j, A_{i+1}, \dots, A_n) = V'.$$

Łatwo zauważyć, że $V \supseteq \operatorname{Lin}(A_i, A_j) = \operatorname{Lin}(A_i + \alpha A_j, A_j) \subseteq V'$.

Stąd $A_1, \dots, A_i, \dots, A_n \in V'$, więc $V \subseteq V'$ i podobnie $V' \subseteq V$.

Dla K2 — analogicznie (ale łatwiej), dla K3 — oczywiste. \square

Fakt 5.12. Operacje wierszowe W1-W3 nie zmieniają jądra $F_A: K^n \rightarrow K^m$.

Dowód. W1: Niech $A \xrightarrow{W_1} A'$. Pokażemy że $\ker F_A = \ker F_{A'}$.

Weźmy $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ (dziedzina F_A i $F_{A'}$).

$$\begin{array}{ccc} X \in \ker F_A & & X \in \ker F_{A'} \\ \Updownarrow & & \Updownarrow \\ AX = 0 & & A'X = 0 \\ \Updownarrow & & \Updownarrow \\ \left\{ \begin{array}{l} \vdots \\ \sum_{j=1}^n a_{ij} x_j = 0 \\ \vdots \\ \sum_{j=1}^n a_{lj} x_j = 0 \\ \vdots \end{array} \right. & \Longleftrightarrow & \left\{ \begin{array}{l} \vdots \\ \sum_{j=1}^n (a_{ij} + \alpha a_{lj}) x_j = 0 \\ \vdots \\ \sum_{j=1}^n a_{lj} x_j = 0 \\ \vdots \end{array} \right. \end{array}$$

Dla W2 i W3 rozumiemy analogicznie (raczej łatwiej). \square

Wniosek 5.13. Operacje wierszowe i kolumnowe nie zmieniają rzędu macierzy.

Dowód. Ustalmy macierz $A \in M_{m \times n}(K)$.

$\operatorname{rk} A = \operatorname{rk} F_A = \dim \operatorname{im} F_A$, więc K1-K3 oczywiście nie zmieniają $\operatorname{rk} A$, bo nie zmieniają $\operatorname{im} F_A$.

Z drugiej strony $\dim \operatorname{im} F_A = \dim K^n - \dim \ker F_A = n - \dim \ker F_A$, więc jeżeli $A \xrightarrow{W1-3} A'$, to $\ker F_A = \ker F_{A'}$, czyli

$$\operatorname{rk} A = n - \dim \ker F_A = n - \dim \ker F_{A'} = \operatorname{rk} A'. \quad \square$$

Przykład 5.14.

$$\begin{pmatrix} 1 & 1 & 2 & 1 \\ 2 & 0 & 2 & -2 \\ 3 & -1 & 2 & -5 \end{pmatrix} \begin{matrix} \\ -2[1] \\ -3[1] \end{matrix}$$

$$\begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & -2 & -2 & -4 \\ 0 & -4 & -4 & -8 \end{pmatrix} \begin{matrix} \\ \\ -2[2] \end{matrix}$$

$$\begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & -2 & -2 & -4 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -\frac{1}{2} \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{matrix} \\ -[1] \quad -2[1] \quad -[1] \\ \end{matrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{matrix} \\ -[2] \quad -2[2] \\ \end{matrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Czyli rząd = 2

Pokażemy teraz

Twierdzenie 5.15. *Jeżeli A jest dowolną macierzą, to $\text{rk}A$ jest równy maksymalnej liczbie liniowo niezależnych wierszy A , czyli $\text{rk}A^\top$.*

Oznaczmy roboczo:

Definicja 5.16. $\text{rk}_w A =$ maksymalna liczba lnz wierszy $A = \text{rk}A^\top$

Ponieważ transpozycja zamienia operacje W1-3 na operacje K1-3 i vice versa, natychmiast mamy

Wniosek 5.17. $\text{rk}_w A$ nie zmienia się przy W1-3 i K1-3.

Lemat 5.18. *Każdą macierz można ciągiem operacji K1-3 i W1-3 sprowadzić*

do macierzy postaci

$$\left(\begin{array}{ccc|c} 1 & & 0 & 0 \\ & \ddots & & \\ 0 & & 1 & \\ \hline & 0 & & 0 \end{array} \right)$$

Dowód. 1. Jeżeli macierz jest zerowa, to nie ma co robić.

2. W przeciwnym wypadku pewne $a_{i,j} \neq 0$.

3. Zamieniając i -ty wiersz z pierwszym i j -tą kolumnę z pierwszą, dostajemy nową macierz w której $a_{1,1} \neq 0$.

4. Wymnażając pierwszy wiersz przez $a_{1,1}^{-1}$ dostajemy nową macierz, w której $a_{1,1} = 1$.

5. Odejmując od j -tego wiersza $a_{1,j}$ -wielokrotność pierwszego (dla $j = 2, 3, \dots$) „czyścimy” pierwszą kolumnę.

6. Odejmując od i -tej kolumny $a_{i,1}$ -wielokrotność pierwszej (dla $i = 2, 3, \dots$) „czyścimy” pierwszy wiersz.

7. To daje nam macierz następującej postaci

$$\left(\begin{array}{c|cccccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & & \\ 0 & & & \end{array} \begin{array}{c} B \end{array} \right)$$

Powtarzamy procedurę dla macierzy B . □

dowód twierdzenia. Weźmy dowolną macierz A , niech A' będzie macierzą uzyskaną z A jak w lemacie. Wtedy oczywiście $\text{rk} A' = \text{rk}_W A'$ jest liczbą jedynek na przekątnej. Z drugiej strony z Faktów wiemy, że $\text{rk} A = \text{rk} A'$ i $\text{rk}_W A = \text{rk}_W A'$, więc to kończy dowód. □

5.3 Eliminacja Gaussa

Twierdzenie 5.19 (Eliminacja Gaussa). *Każdą macierz można operacjami $W1$, $W3$ sprowadzić do macierzy w postaci schodkowej, to znaczy macierzy postaci:*

$$\begin{pmatrix} 0 \dots 0 [\neq 0] ? & & & \\ 0 \dots 0 [\neq 0] ? & ? & & \\ 0 \dots 0 [\neq 0] ? & & ? & \\ & 0 & & \ddots \\ & & 0 & \end{pmatrix}$$

Wyrazy „ $\neq 0$ ” (pierwsze niezerowe w wierszach) nazywamy wyrazami wiodącymi.

Dowód. Przesuwamy się w prawo:

1. Zamieniając być może pierwszy wiersz z innym, możemy zagwarantować że najbardziej wysunięty na lewo niezerowy wyraz jest w pierwszym wierszu (ale jeszcze mogą być inne, równie daleko wysunięte).
2. Odejmując wielokrotność pierwszego wiersza od pozostałych, oczyścimy odpowiednią kolumnę, tak że w żadnym innym wierszu nie ma równie daleko na lewo wysuniętego niezerowego wyrazu.
3. Powtarzamy procedurę w dół i na prawo od wiodącego wyrazu pierwszego wiersza.

□

Wniosek 5.20. *Korzystając z operacji $W1$, $W2$, $W3$ można sprowadzić każdą macierz do zredukowanej postaci schodkowej, tzn. postaci schodkowej z wyrazami wiodącymi 1.*

Dowód. Oczywiście — wystarczy podzielić każdy niezerowy wiersz przez odwrotność wyrazu wiodącego. □

Jak to prowadzi do rozwiązania równania? Załóżmy że za pomocą eliminacji Gaussa sprowadziliśmy macierz rozszerzoną układu do zredukowanej postaci schodkowej. To daje macierz rozszerzoną układu równań, równoważnego wyjściowemu:

$$\left\{ \begin{array}{l} x_{n_1} + b_{1(n_1+1)}x_{n_1+1} + \dots = y_1 \\ x_{n_2} + b_{2(n_2+1)}x_{n_2+1} + \dots = y_2 \\ \vdots \\ x_{n_k} + b_{m(n_k+1)}x_{n_k+1} + \dots = y_k \\ 0 = y_{k+1} \\ \vdots \\ 0 = y_m \end{array} \right. , \quad (4)$$

gdzie $n_1 < n_2 < \dots < n_k$

Wtedy z twierdzenia Kroneckera-Capelliego (albo prosto z obserwacji powstałego układu równań) wynika, że ten układ równań (a więc i wyjściowy układ równań) ma rozwiązanie wtedy i tylko wtedy, gdy $y_{k+1} = y_{k+2} = \dots = y_m = 0$.

Zakładając że tak jest możemy uzyskać postać ogólną rozwiązań. Zmienne x_{n_i} traktujemy jako zmienne związane, a pozostałe jako zmienne wolne. Wartość zmiennych związanych wyliczamy przenosząc zmienne wolne na drugą stronę:

$$\left\{ \begin{array}{l} x_{n_1} = y_1 - b_{1(n_1+1)}x_{n_1+1} - \dots \\ x_{n_2} = y_2 - b_{2(n_2+1)}x_{n_2+1} - \dots \\ \vdots \\ x_{n_k} = y_k - b_{m(n_k+1)}x_{n_k+1} - \dots \end{array} \right. ,$$

a następnie podstawiając za x_{n_k} w pierwszych $k-1$ równaniach prawą stronę ostatniego równania, za $x_{n_{k-1}}$ w pierwszych $k-2$ równaniach prawą stronę przedostatniego równania itd., co ostatecznie prowadzi do układu postaci jak powyżej, z tym że po prawej stronie nie występują już zmienne związane x_{n_1}, \dots, x_{n_k} .

$$\left\{ \begin{array}{l} x_{n_1} = y'_1 + c_{1(n_1+1)}x_{n_1+1} + \dots \\ x_{n_2} = y'_2 + c_{2(n_2+1)}x_{n_2+1} + \dots \\ \vdots \\ x_{n_k} = y'_k + c_{m(n_k+1)}x_{n_k+1} + \dots \end{array} \right. ,$$

Przykład 5.21. Pewien układ równań zmiennych x_1, x_2, \dots, x_9 sprowadziliśmy do zredukowanej postaci schodkowej, uzyskując układ:

$$\left\{ \begin{array}{l} x_4 + 7x_5 - 3x_6 = 1 \\ x_6 + 2x_7 + 3x_8 = 2 \\ x_7 + 3x_8 + 4x_9 = 3 \\ \qquad \qquad \qquad 0 = 0 \\ \qquad \qquad \qquad \vdots \\ \qquad \qquad \qquad 0 = 0 \end{array} \right. \quad (*)$$

Przenosząc zmienne związane na jedną stronę:

$$\left\{ \begin{array}{l} x_4 = 1 - 7x_5 + 3x_6 \\ x_6 = 2 - 2x_7 - 3x_8, \\ x_7 = 3 - 3x_8 - 4x_9 \end{array} \right.$$

Podstawiając za x_7 , a następnie x_6 i x_4 uzyskujemy równoważny układ

$$\left\{ \begin{array}{l} x_7 = 3 - 3x_8 - 4x_9 \\ x_6 = 2 - 2(3 - 3x_8 - 4x_9) - 3x_8 \\ \qquad = 2 - 6 + 6x_8 + 8x_9 - 3x_8 \\ \qquad = -4 + 3x_8 + 8x_9 \\ x_4 = 1 - 7x_5 + 3(-4 + 3x_8 + 8x_9) \\ \qquad = 1 - 7x_5 - 12 + 9x_8 + 24x_9 \\ \qquad = -11 - 7x_5 + 9x_8 + 24x_9 \end{array} \right. ,$$

i to jest postać ogólna rozwiązania układu (zmienne wolne $x_1, x_2, x_3, x_5, x_8, x_9$ mogą przyjmować dowolne wartości). W postaci wektorowej rozwiązanie

jest więc postaci:

$$\begin{aligned}
 X = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_6 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_7 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_8 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_9 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \\
 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -11 \\ 0 \\ -4 \\ 3 \\ 0 \\ 0 \end{pmatrix} + x_1 e_1 + x_2 e_2 + x_3 e_3 + x_5 \begin{pmatrix} 0 \\ 0 \\ 0 \\ -7 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_8 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 9 \\ 0 \\ 3 \\ -3 \\ 1 \\ 0 \end{pmatrix} + x_9 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 24 \\ 0 \\ 8 \\ -4 \\ 0 \\ 1 \end{pmatrix}
 \end{aligned}$$

$$\text{Wektory } X_1 = e_1, X_2 = e_2, X_3 = e_3, X_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -7 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, X_5 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 9 \\ 0 \\ 3 \\ -3 \\ 1 \\ 0 \end{pmatrix}, X_6 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 24 \\ 0 \\ 8 \\ -4 \\ 0 \\ 1 \end{pmatrix}$$

stanowią *fundamentalny układ rozwiązań* układu jednorodnego (UJ) stowa-

rzyszonego z układem (*), a wektor $X_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -11 \\ 0 \\ -4 \\ 3 \\ 0 \\ 0 \end{pmatrix}$ jest szczególnym rozwią-

zaniem (*). Rozwiązania (*) to wszystkie sumy rozwiązania szczególnego i dowolnej kombinacji liniowej *fundamentalnego układu rozwiązań*:

$$X = X_0 + aX_1 + bX_2 + cX_3 + dX_4 + eX_5 + fX_6,$$

gdzie a, b, c, d, e, f to dowolne skalary.

Przykład 5.22. Jak wyznaczyć obraz odwzorowania F_A zadanego macierzą A , lub innymi słowy wyznaczyć wektory $Y = (y_1, \dots, y_m)^\top$, dla których równanie

$$AX = Y$$

ma rozwiązanie? Patrzymy na macierz rozszerzoną $(A|Y)$ układu powyżej z parametrami y_1, \dots, y_m . Eliminacja Gaussa prowadzi nas do macierzy w postaci schodkowej:

$$\left(\begin{array}{c|c} 0 \dots \dots [\neq 0] \dots \dots \dots & \text{(pewna kombinacja liniowa } y_1, \dots, y_m) \\ 0 \dots \dots \dots [\neq 0] \dots \dots \dots & \text{(pewna kombinacja liniowa } y_1, \dots, y_m) \\ \vdots & \\ 0 \dots \dots \dots [\neq 0] \dots \dots \dots & \text{(pewna kombinacja liniowa } y_1, \dots, y_m) \\ 0 \dots \dots \dots & \text{(pewna kombinacja liniowa } y_1, \dots, y_m) \\ \vdots & \\ 0 \dots \dots \dots & \text{(pewna kombinacja liniowa } y_1, \dots, y_m) \end{array} \right)$$

W świetle twierdzenia Kroneckera-Capelliego, dla danych y_1, \dots, y_m ten układ ma rozwiązanie dokładnie wtedy gdy ich kombinacje w wierszach, które mają zerowe współczynniki w części głównej, są zerowe. To daje nam układ równań opisujący obraz F_A .

Na przykład jeżeli dostalibyśmy macierz postaci

$$\left(\begin{array}{cccccccc} 0 & \dots & \pi & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \sqrt{3} - \sqrt{\pi} & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 8 & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{array} \middle| \begin{array}{l} \text{(pewna kombinacja liniowa } y_1, \dots, y_m) \\ \text{(pewna kombinacja liniowa } y_1, \dots, y_m) \\ \text{(pewna kombinacja liniowa } y_1, \dots, y_m) \\ y_4 + \pi y_6 - \sqrt{7} y_2 3 \\ y_9 \end{array} \right)$$

to układ ma rozwiązanie dokładnie wtedy gdy $y_4 + \pi y_6 - \sqrt{7} y_2 3 = 0 = y_9$ (i te równania opisują obraz F_A).

Słowniczek:

$$\begin{array}{llll} X \text{ jest rozwiązaniem UJ } AX = 0 & \iff & X \in \ker F_A \\ AX = Y \text{ ma rozwiązanie} & \iff & Y \in \operatorname{im} F_A \\ \text{Fundamentalny układ rozwiązań UJ } AX = 0 & = & \text{baza } \ker F_A \\ \text{rozwiązanie szczególne } X_0 \text{ równania } AX = Y & = & X_0 \in F_A^{-1}\{Y\} \end{array}$$

Przykład 5.23. Mając dane liniowo niezależne wektory $X_1, X_2, \dots, X_k \in K^n$, jak znaleźć układ równań, którego to jest fundamentalny układ rozwiązań?

Szukamy macierzy A takiej że $AX_1, AX_2, \dots, AX_k = 0$. Jeżeli wiersze A to A_1, A_2, \dots, A_m , to znaczy że $A_i X_j = 0$ dla $i = 1, \dots, m$ i $j = 1, \dots, k$. Transponując, dostajemy równania $X_j^\top A_i^\top$, a więc A_i^\top mają być rozwiązaniami jednorodnego układu równań

$$BY = 0,$$

gdzie $B = (X_1 X_2 \dots X_k)^\top$. Rząd B to k (bo X_1, \dots, X_k są liniowo niezależne). Macierz B ma szerokość n (więc to jest wymiar dziedziny F_B), więc z twierdzenia o rzędzie $\dim \ker F_B = n - k$.

Fundamentalny układ rozwiązań $BY = 0$ składa się z liniowo niezależnych wektorów Y_1, \dots, Y_{n-k} . Macierz $A = (Y_1 \dots Y_{n-k})^\top$ spełnia $AX_j = 0$ dla każdego j . Ponadto jej rząd to $n - k$ (bo wiersze są liniowo niezależne), a szerokość to n (bo szerokość = długość Y_i = szerokość B), więc $\dim \ker F_A = n - (n - k) = k$, więc X_1, \dots, X_k jest bazą $\ker F_A$, czyli fundamentalnym układem rozwiązań $AX = 0$.

Przykład 5.24. Weźmy liniowo niezależne wektory $(1, 1, 0, 0)^\top, (3, 1, -1, 1)^\top \in \mathbb{R}^4$. Szukamy układu równań, którego to jest fundamentalny układ rozwiązań. Rozważamy więc równanie jednorodne

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 3 & 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = 0.$$

Eliminacja Gaussa macierzy powyższego układu daje macierz

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

Stąd równoważnie mamy $y_2 = \frac{y_4 - y_3}{2}$, $y_1 = -y_2 = \frac{y_3 - y_4}{2}$, więc rozwiązania są zadane równaniem $y_3(\frac{1}{2}, -\frac{1}{2}, 1, 0)^\top + y_4(-\frac{1}{2}, \frac{1}{2}, 0, 1)^\top$ i wektory $(\frac{1}{2}, -\frac{1}{2}, 1, 0)^\top, (-\frac{1}{2}, \frac{1}{2}, 0, 1)^\top$ stanowią fundamentalny układ rozwiązań.

Stąd szukany układ jednorodny ma macierz główną $\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 & 1 \end{pmatrix}$, czyli jest układem

$$\begin{cases} \frac{1}{2}x_1 - \frac{1}{2}x_2 + x_3 = 0 \\ -\frac{1}{2}x_2 + \frac{1}{2}x_2 + x_4 = 0. \end{cases}$$

Podsumowanie

Uwaga 5.25. Metoda eliminacji Gaussa daje, dla danego układu równań liniowych postaci $AX = Y$:

- parametryzację zbioru rozwiązań postaci

$$X = X_0 + t_1X_1 + \dots + t_lX_l$$

z minimalną liczbą parametrów,

- opis obrazu F_A przy użyciu minimalnej możliwej liczby równań liniowych.

Dowód. Załóżmy że A jest wysokości m , szerokości n . Eliminacja Gaussa $(A|Y)$ prowadzi do $(S|Y')$, gdzie S jest macierzą schodkową o s schodkach. Wtedy $\text{rk} A = \text{rk} S = s$.

- Jest n zmiennych, z których s zmiennych jest związanych.
- Jest $n - s$ zmiennych wolnych, czyli $l = n - s$ parametrów w rozwiązaniu ogólnym. Ale (z twierdzenia o indeksie) $\dim \ker F_A = \dim n - \dim \text{im} F_A = n - s = l$; stąd więc zbiór rozwiązań jest warstwą l -wymiarowej podprzestrzeni K^m , skąd wynika pierwszy punkt.

- $\text{im } F_A = \{Y \in K^m \mid BY = 0\}$. Ile wierszy może mieć B ?

$$F_B: K^m \rightarrow K^k \quad k = ?$$

$s = \dim \text{im } F_A = \dim \ker F_B = m - \dim \text{im } F_B \geq m - k$, czyli $k \geq m - s$. Ale eliminacja Gaussa dostarcza dokładnie $m - s$ równań opisujących $\text{im } F_A$. \square

Uwaga 5.26. Jeżeli zobaczymy co eliminacja Gaussa robi z układem równań liniowych (zadany daną macierzą), to łatwo zobaczyć, że jest to w istocie (tylko nieco usystematyzowana) zupełnie szkolna metoda — wyliczamy kolejne zmienne z pozostałych, korzystając z kolejnych równań.

Uwaga 5.27. W szczególności jeżeli X_1, X_2, \dots, X_s są dowolnymi liniowo niezależnymi wektorami, to biorąc macierz $C = (X_1, X_2, \dots, X_s)$ otrzymujemy opis $\text{Lin}(X_1, X_2, \dots, X_s) = \text{im } F_C$ za pomocą (minimalnego) układu równań (równoważnie, jako zbioru rozwiązań równania postaci $AX = 0$).

Uwaga 5.28. Mamy też konkurencyjny sposób znalezienia tego układu równań: dla danego układu $X_1, X_2, \dots, X_s \in K^n$ lnz wektorów, możemy rozważyć macierz $B = (X_1, X_2, \dots, X_s)^\top = C^\top$. Ma ona wysokość i rząd s (bo jej wiersze są lnz) i szerokość n .

Postępując jak powyżej, znajdujemy liniowo niezależne wektory Y_1, Y_2, \dots, Y_{n-s} rozpinające $\ker F_B$ (bo $\dim \ker F_B = n - \text{rk } B = n - s$). Za-uważmy że wektory Y_j spełniają $X_i^\top Y_j = 0$ (dla każdych i, j).

Łatwo sprawdzić, że $A = (Y_1, Y_2, \dots, Y_{n-s})^\top$ jest macierzą szerokości n , wysokości i rzędu $n - s$ i spełnia $AX_i = 0$ dla każdego i (ponieważ $X_i^\top Y_j = 0$), czyli $X_i \in \ker F_A$.

Ponieważ $\text{rk } A = n - s$, $\dim \ker F_A = n - s$, czyli wektory X_1, \dots, X_s są fundamentalnym układem rozwiązań $AX = 0$, tj. bazą $\ker F_A$.

Gdyby A' była macierzą $n \times m$, gdzie $m < n - s$, to $\dim \text{im } A' \leq m < n - s$, a więc $\dim \ker A' \geq n - m > n - (n - s) = s$, a więc $\text{Lin}(X_1, \dots, X_n)$ nie mógłby być zbiorem rozwiązań $A'X = 0$. Stąd A faktycznie daje minimalny układ równań o zadanej własności.

Ćwiczenie 5.29. Używając uwag powyżej, dla dowolnego wektora X_0 i lnz układu $X_1, X_2, \dots, X_s \in K^n$, możemy wyprodukować podobnie optymalny opis warstwy $X_0 + \text{Lin}(X_1, X_2, \dots, X_n)$ jako zbioru rozwiązań równania $AX = Y$ /niejednorodnego układu równań.

Uwaga 5.30. Powyższe konstrukcje działają też wtedy, gdy wektory X_1, \dots, X_s nie są liniowo niezależne — jednak wtedy macierz A ma wysokość $n - \dim \text{Lin}(X_1, \dots, X_s) > n - s$. Wynika to np. z tego, że macierz B będzie miała odpowiednio mniejszy rząd.

5.4 Macierze przekształceń liniowych

Metody które mamy pozwalają nam wyznaczyć obraz i jądro odwzorowania $K^n \rightarrow K^m$ zadanego macierzą A , jak również przeciwobraz wektora $Y \in K^m$ przez takie odwzorowanie.

Ale co gdy mamy dowolne (skończone wymiarowe) V i W oraz odwzorowanie liniowe $F: V \rightarrow W$?

Możemy wybrać bazy V i W i utożsamić V i W z $K^{\dim V}$ i $K^{\dim W}$.

Przykład 5.31. $V = \{P \in \mathbf{R}_2[X] \mid P'(-1) = 0\}$, $W = \mathbf{R}_1[X]$, $F(P) = P' + P(1)X$.

Czym jest $F^{-1}\{X+1\}$ (czyli zbiór rozwiązań równania niejednorodnego $F(P) = X+1$)?

$\dim V = 2 = \dim W$, przykładowa baza V : $1, X^2 + 2X$. Przykładowa baza W : $1, X$.

$$\begin{array}{ccc} V & \xrightarrow{F} & W \\ \downarrow a+b(X^2+2X) \mapsto (a,b) & & \downarrow c+dX \mapsto (c,d) \\ \mathbf{R}^2 & \xrightarrow{\begin{pmatrix} ? & ? \\ ? & ? \end{pmatrix}} & \mathbf{R}^2 \end{array}$$

Szukamy macierzy pasującej do diagramu.

$$F(a+b(X^2+2X)) = F(bX^2+2bX+a) = 2bX+2b+(b+2b+a)X = 2b+(a+5b)X,$$

czyli mamy

$$\begin{array}{ccc} V & \xrightarrow{F} & W \\ \downarrow a+b(X^2+2X) \mapsto (a,b) & & \downarrow c+dX \mapsto (c,d) \\ \mathbf{R}^2 & \xrightarrow{\begin{pmatrix} 0 & 2 \\ 1 & 5 \end{pmatrix}} & \mathbf{R}^2 \end{array}$$

Szukamy rozwiązania $F(P) = X + 1$, czyli w naszym tłumaczeniu

$$\begin{pmatrix} 0 & 2 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = [X + 1]_{(1, X)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

to znaczy:

$$\begin{cases} 2b = 1 \\ a + 5b = 1 \end{cases}$$

czyli $b = \frac{1}{2}$, $a = 1 - 5b = -\frac{3}{2}$, a więc jedyne rozwiązanie to $P = a + b(X^2 + 2X) = \frac{1}{2}X^2 + X - \frac{3}{2}$.

Ogólnie: jeżeli $B = (b_1, \dots, b_n)$ to baza V , a $C = (c_1, \dots, c_m)$ to baza W , a $F: V \rightarrow W$ jest odwzorowaniem liniowym, to mamy

$$\begin{array}{ccccc} v & \xlongequal{\quad} & \sum_i \beta_i b_i & & V \xrightarrow{F} W \\ \downarrow & & \uparrow & & \downarrow \\ [v]_B & \xlongequal{\quad} & \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} & & K^n \xrightarrow{\quad ? \quad} K^m \\ & & & & \uparrow \\ w & \xlongequal{\quad} & \sum_j \gamma_j c_j & & W \\ \downarrow & & \uparrow & & \downarrow \\ [w]_C & \xlongequal{\quad} & \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix} & & K^m \end{array}$$

Dla $v = \sum_i \beta_i b_i$ mamy $F(v) = \sum_i \beta_i F(b_i)$. Zapisujemy $F(b_i) = \sum_j a_{ji} c_j$. Wtedy

$$F(v) = \sum_i \beta_i F(b_i) = \sum_i \beta_i \sum_j a_{ji} c_j = \sum_j \left(\sum_i a_{ji} \beta_i \right) c_j$$

(Na końcu zmieniona kolejność sumowania!)

Mamy więc

$$\begin{array}{ccc} v & \xrightarrow{\quad} & F(v) \\ \updownarrow & & \updownarrow \\ \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} & \xrightarrow{\quad} & \begin{pmatrix} \sum_i a_{1i} \beta_i \\ \vdots \\ \sum_i a_{mi} \beta_i \end{pmatrix} \xlongequal{\quad} A \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}, \end{array}$$

gdzie $A = (a_{ji})$. Zauważmy że i -ta kolumna macierzy A to dokładnie $[F(b_i)]_C$.

Definicja 5.32. Macierz A jak powyżej nazywamy *macierzą przekształcenia F względem baz B i C* , oznaczamy $m_C^B(F)$ (ale czasami też niestety $m_{BC}(F)$, $m_{CB}(F)$,...?).

(i-ta kolumna $m_C^B(F)$ to $[F(b_i)]_C$.)

Innymi słowy, $m_C^B(F)$ to taka macierz, że poniższy diagram jest przemienny (komutuje)

$$\begin{array}{ccc} V & \xrightarrow{F} & W \\ \downarrow [\cdot]_B & & \downarrow [\cdot]_C \\ K^n & \xrightarrow{m_C^B(F)} & K^m, \end{array}$$

to znaczy $[F(v)]_C = m_C^B(F)[v]_B$ (mnemotechnika: B się “skraca”).

Jeszcze inaczej, $m_C^B(F)$ oznacza takie przekształcenie $K^n \rightarrow K^m$: weź współrzędne $[v]_B$ dla pewnego v , policz $F(v)$, a następnie odczytaj jego współrzędne $[F(v)]_C$.

Uwaga 5.33. Jeżeli B, C są bazami V i W i A jest macierzą odpowiedniego rozmiaru, to istnieje (jedyne) takie $F: V \rightarrow W$, że $A = m_C^B(F)$: z diagramu powyżej widzimy, że musi być $F = [\cdot]_C^{-1} \circ F_A \circ [\cdot]_B$.

Innymi słowy, dana macierz zadaje jedyne przekształcenie $V \rightarrow W$ (względem ustalonych baz V, W), czyli dla każdego wyboru baz B, C , funkcja $F \mapsto m_C^B(F)$ jest bijekcją:

$$m_C^B: \text{Hom}(V, W) \rightarrow M_{\dim W \times \dim V}(K).$$

(A nawet izomorfizmem liniowym!)

Przykład 5.34. Jeszcze raz ten sam przykład.

$m_C^B(F) = \begin{pmatrix} 0 & 2 \\ 1 & 5 \end{pmatrix}$. $F(b_1) = F(1) = 0 + 1 \cdot X = X = 0 \cdot 0 + 1 \cdot X$, $F(b_2) = F(X^2 + 2X) = 2X + 2 + 3 \cdot X = 5X + 2 = 2 \cdot 1 + 5 \cdot X$: więc faktycznie, kolumny $m_C^B(F)$ to współrzędne w bazie C obrazów wektorów z bazy B .

A co, gdy chcemy zamiast B wziąć inną bazę, np.: $B' = (X^2 + 2X + 3, 1)$ (czyli chcemy policzyć $m_C^{B'}(F)$)?

Twierdzenie 5.35. Załóżmy że mamy

$$Z \xleftarrow{G} W \xleftarrow{F} V$$

$$D \quad C \quad B \quad (\text{bazy})$$

Wtedy $m_D^B(G \circ F) = m_D^C(G) \cdot m_C^B(F)$. (Mnemotechnika: C się skraca!)

Dowód. i -ta kolumna $m_D^B(G \circ F)$ to

$$[(G \circ F)(b_i)]_D = [G(F(b_i))]_D = m_D^C(G)[F(b_i)]_C,$$

a $[F(b_i)]_C$ to i -ta kolumna $m_C^B(F)$. Ale to daje tezę (narysuj schemat mnożenia $m_D^C(G)$ i $m_C^B(F)$ żeby się przekonać!). \square

Przykład 5.36. Ciąg dalszy.

$$W \xleftarrow{F} V \xleftarrow{\text{id}_V} V$$

$$C \quad B \quad B'$$

Znamy $m_C^B(F) = \begin{pmatrix} 0 & 2 \\ 1 & 5 \end{pmatrix}$. Czym jest $m_B^{B'}(\text{id}_V)$? Kolumny = wektory bazy B' w bazie B , czyli $X^2 + 2X + 3 = 3 \cdot 1 + 1 \cdot (X^2 + 2X)$, $1 = 1 \cdot 1 + 0 \cdot (X^2 + 2X)$, więc $m_B^{B'}(\text{id}_V) = \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix}$, czyli

$$m_C^{B'}(F) = m_C^B(F) \cdot m_B^{B'}(\text{id}_V) = \begin{pmatrix} 0 & 2 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 8 & 1 \end{pmatrix}.$$

Definicja 5.37. Dla $F: V \rightarrow V$ (czyli endomorfizmu V) oznaczamy $m_B(F) := m_B^B(F)$.

Przykład 5.38. Mamy $F: V \rightarrow V$ i bazy B, C przestrzeni V . Jak policzyć $m_C(F)$, znając $m_B(F)$? Ponieważ $F = \text{id}_V \circ F \circ \text{id}_V$, mamy:

$$m_C(F) = m_C^C(F) = m_C^B(\text{id}_V) \cdot m_B^B(F) \cdot m_B^C(\text{id}_V) = m_C^B(\text{id}_V) \cdot m_B(F) \cdot m_B^C(\text{id}_V).$$

Uwaga 5.39. Jeżeli B, C są bazami V i $v \in V$, to $[v]_C = m_C^B(\text{id}_V)[v]_B$. (B się skraca!)

5.5 Odwracalność macierzy

Definicja 5.40. Mówimy że $A \in M_{m \times n}(K)$ jest *odwracalna* jeżeli istnieje $B \in M_{n \times m}(K)$ (macierz odwrotna do A) taka że $AB = I_m$ i $BA = I_n$.

Taka macierz, o ile istnieje, jest jedyna, i oznaczamy ją przez A^{-1} .

(Dowód: jeżeli B, B' są takie, to $B' = B'I_m = B'AB = I_n B = B$.)

Fakt 5.41. $m_B^C(\text{id})^{-1} = m_C^B(\text{id})$.

Dowód. $m_B^C(\text{id}) \cdot m_C^B(\text{id}) = m_B^B(\text{id} \circ \text{id}) = m_B^B(\text{id}) = I$ i podobnie odwrotnie. \square

Wniosek 5.42. Jeżeli $F \in \text{End}(V)$ i B, C są bazami V , to zachodzi wzór:

$$\begin{aligned} m_C(F) &= m_C^B(\text{id}) m_B(F) m_C^B(\text{id})^{-1} \\ &= P m_B(F) P^{-1} \\ &= m_B^C(\text{id})^{-1} m_B(F) m_B^C(\text{id}), \end{aligned}$$

gdzie $P = m_C^B(\text{id})$ (macierz przejścia z bazy B do bazy C , macierz wektorów “starej” bazy B w “nowej” bazie C).

Jak liczyć A^{-1} ?

Pierwszy sposób¹: eliminacja Gaussa. Zauważmy że $AX = Y \iff X = A^{-1}Y$.

Wychodząc od $(A|I)$ przez operacje W1–W3 uzyskujemy $(I|A^{-1})$, co w języku układów równań oznacza przejście między równoważnymi układami równań:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = y_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = y_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = y_n \end{cases} \rightsquigarrow \begin{cases} x_1 = a_{11}^{-1}y_1 + \dots \\ x_2 = -a_{11}^{-1}a_{21}y_1 + \dots \\ \vdots \\ x_n = \dots \end{cases}$$

Jeżeli to się nie uda, to dojdziemy do postaci schodkowej o liczbie schodków po lewej stronie mniejszej niż liczba wierszy (lub kolumn), to znaczy że $\text{rk} A$ jest mniejszy niż liczba jej wierszy (lub kolumn). Przekonamy się wkrótce, że wtedy A^{-1} nie istnieje.

¹Drugi sposób będzie trochę później.

Przykład 5.43. • $\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 2 & 4 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 0 & -2 & 1 \end{array}\right)$: nie ma macierzy odwrotnej

$$\bullet \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & -2 & -3 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array}\right)$$

Przykład 5.44. $E = (e_1, e_2)$, $B = \left(\begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \end{pmatrix}\right)$ — bazy \mathbf{R}^2 . $F: \mathbf{R}^2 \rightarrow \mathbf{R}^2$,

$$m_E(F) = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}.$$

Wtedy

$$m_B(F) = m_B^E(\text{id}) m_E(F) m_E^B(\text{id}).$$

Oczywiście $m_E^B(\text{id}) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ i $m_B^E(\text{id}) = m_E^B(\text{id})^{-1} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$, czyli

$$m_B(F) = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \dots$$

Definicja 5.45. $F: V \rightarrow W$ nazywamy *odwracalnym* jeżeli istnieje takie liniowe $G: W \rightarrow V$, że $F \circ G = \text{id}_W$ i $G \circ F = \text{id}_V$.

Takie G (gdy istnieje) jest jedyne i oznaczamy je F^{-1} (dowód jak dla macierzy).

Ćwiczenie 5.46. F jest odwracalne $\iff F$ izomorfizmem. (To wiemy z ćwiczeń.)

Wniosek 5.47. Jeżeli $F: V \rightarrow W$ jest odwracalne, to $\dim V = \dim W$.

Fakt 5.48. Niech $F: V \rightarrow W$ — liniowe, a B, C to bazy V i W .

Wtedy F jest odwracalne $\iff m_C^B(F)$ jest odwracalna.

Dowód. Jeżeli F jest odwracalne i $G = F^{-1}$, to $m_B^C(G)$ jest odwrotna do $m_C^B(F)$:

$$I = m_B^B(\text{id}_V) = m_B^B(G \circ F) = m_B^C(G) m_C^B(F)$$

i podobnie w odwrotnej kolejności.

Jeżeli $m_C^B(F)$ jest odwracalna, to weźmy $G: W \rightarrow V$ takie że $m_B^C(G) = m_C^B(F)^{-1}$. Dalej rachunek jak wyżej (tylko czytany od końca). \square

Z tego faktu i z twierdzenia o indeksie otrzymujemy:

Wniosek 5.49. *Macierz A jest odwracalna $\iff A$ jest kwadratowa i $\text{rk} A$ równa się liczbie jej kolumn/wierszy.*

Fakt 5.50. 1. $\text{rk}(AB) \leq \text{rk} A$,

2. $\text{rk}(AB) \leq \text{rk} B$,

3. jeżeli A jest odwracalna, to $\text{rk}(AB) = \text{rk} B$,

4. jeżeli B jest odwracalna, to $\text{rk}(AB) = \text{rk} A$.

Dowód. Rząd = wymiar obrazu.

1. $\text{im}(F_A \circ F_B) \subseteq \text{im} F_A$,

2. $\text{rk}(AB) = \text{rk}(AB)^\top = \text{rk}(B^\top A^\top) \leq \text{rk} B^\top = \text{rk} B$

3. wiemy że $\text{rk}(AB) \leq \text{rk} B$.

$$\text{rk} B = \text{rk} A^{-1}AB = \text{rk}(A^{-1}(AB)) \leq \text{rk}(AB).$$

4. $\text{rk} A = \text{rk}((AB)B^{-1}) \leq \text{rk}(AB)$

□

Rozdział 6

Wyznacznik

6.1 Permutacje i definicja wyznacznika

Przykład 6.1. Wyznacznik $\det: \mathbf{R}^2 \times \mathbf{R}^2 \rightarrow \mathbf{R}$ jest funkcją

- dwuliniową, tzn. przy ustalonym X_2 , funkcja $\det(X_1, X_2)$ jest liniowa w X_1 i podobnie przy ustalonym X_1 ;
- antysymetryczną, tzn. $\det(X_1, X_2) = -\det(X_2, X_1)$

Przykład 6.2. Podobnie wyznacznik $\det: (\mathbf{R}^3)^3 \rightarrow \mathbf{R}$ jest:

- trójliniowy (gdy ustalimy dwa z X_1, X_2, X_3 , to $\det(X_1, X_2, X_3)$ jest liniowy jako funkcja trzeciego wektora),
- antysymetryczny: jeżeli zamienimy ze sobą miejscami dwa argumenty, to znak $\det(X_1, X_2, X_3)$ się zmieni.

Uwaga 6.3. Obydwa wyznaczniki są *alternujące*, tzn. zerują się, jeżeli któreś dwa argumenty są równe (a nawet wtedy, kiedy argumenty są tylko liniowo zależne).

Chcemy uogólnić pojęcie wyznacznika na dowolne $(K^n)^n$.

Problem: wyznaczyć wszystkie funkcje $d: (K^n)^n \rightarrow K$, które są:

- n -liniowe: przy jeżeli ustalimy dowolne $n - 1$ argumentów funkcji $d(X_1, X_2, \dots, X_n)$, to uzyskana funkcja $K^n \rightarrow K$ jest liniowa.
- alternujące: jeżeli $X_i = X_j$ dla pewnych $i \neq j$, to $d(X_1, X_2, \dots, X_n) = 0$.

Uwaga 6.4. Jeżeli $\text{char } K \neq 2$ (tzn. $1 + 1 \neq 0$ w K), to n -liniowe d jest alternujące \iff jest antysymetryczne (zmienia znak, jeżeli zamienimy miejscami dwa argumenty).¹

Dowód. Załóżmy że d jest n -liniowa i alternująca. Pokażemy że $d(X_1, X_2, X_3, \dots) = -d(X_2, X_1, X_3, \dots)$.

To jest równoważne $d(X_1, X_2, X_3, \dots) + d(X_2, X_1, X_3, \dots) = 0$. Z alternowania wiemy, że $d(X_1, X_1, X_3, \dots) = d(X_2, X_2, X_3, \dots) = 0$, możemy więc dodać to do lewej strony, co daje równoważną nierówność

$$\begin{aligned} 0 &= d(X_1, X_2, X_3, \dots) + \\ &\quad + d(X_2, X_1, X_3, \dots) + \\ &\quad + d(X_1, X_1, X_3, \dots) + \\ &\quad + d(X_2, X_2, X_3, \dots) \\ &= d(X_1 + X_2, X_2, X_3, \dots) + \\ &\quad + d(X_1 + X_2, X_1, X_3, \dots) \\ &= d(X_1 + X_2, X_1 + X_2, X_3, \dots). \end{aligned}$$

Ta ostatnia wartość jest zerowa z alternowania.

W drugą stronę, równość $d(X_1, X_1, X_3, \dots) = -d(X_1, X_1, X_3, \dots)$ (wynikająca z antysymetryczności) implikuje $2d(X_1, X_1, X_3, \dots) = 0$ i (gdy $\text{char } K \neq 2$) możemy podzielić stronami przez 2.² \square

¹Jeżeli $\text{char } K = 2$, to alternowanie implikuje antysymetryczność (równoważnie: symetryczność, bo jeżeli $1 + 1 = 0$, to $-1 = 1$, czyli $-\alpha = (-1)\alpha = 1\alpha = \alpha$), ale nie odwrotnie.

²Dla $K = \mathbf{F}_2 = \{0, 1\}$ i $n = 2$ funkcja $f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1 y_1 + x_2 y_2$ jest antysymetryczna, ale nie alternująca.

Wracamy do problemu. Weźmy $A_1, A_2, \dots, A_n \in K^n$. Wtedy

$$\begin{aligned}
 d(A_1, A_2, \dots) &= d\left(\sum_{i_1=1}^n a_{i_1 1} e_{i_1}, \sum_{i_2=1}^n a_{i_2 2} e_{i_2}, \dots, \sum_{i_n=1}^n a_{i_n n} e_{i_n}\right) \\
 &= \sum_{i_1=1}^n a_{i_1 1} d\left(e_{i_1}, \sum_{i_2=1}^n a_{i_2 2} e_{i_2}, \dots, \sum_{i_n=1}^n a_{i_n n} e_{i_n}\right) \\
 &= \sum_{i_1, i_2, \dots, i_n=1}^n a_{i_1 1} a_{i_2 2} \cdots a_{i_n n} d(e_{i_1}, e_{i_2}, \dots, e_{i_n}) \\
 &= \sum_{i_1, i_2, \dots, i_n \text{ parami różne}} a_{i_1 1} a_{i_2 2} \cdots a_{i_n n} d(e_{i_1}, e_{i_2}, \dots, e_{i_n}) \\
 &= \sum_{\sigma \text{ permutacja } \mathbb{C}\{1, \dots, n\}} a_{\sigma(1)1} \cdots a_{\sigma(n)n} d(e_{\sigma(1)}, \dots, e_{\sigma(n)})
 \end{aligned}$$

Ciąg i_1, i_2, \dots, i_n odpowiada permutacji $\{1, 2, \dots, n\}$ zadanej przez $\sigma(j) = i_j$.

Zbiór permutacji (bijekcji w siebie) zbioru $\{1, \dots, n\}$ oznaczamy S_n , ma on oczywiście $n!$ elementów.

Z antysymetryczności wynika że $d(e_{\sigma(1)}, \dots, e_{\sigma(n)}) = \pm d(e_1, \dots, e_n)$. Od czego zależy znak?

Przykład 6.5.

$$\begin{aligned}
 d(e_3, e_4, e_2, e_1) &= -d(e_3, e_4, e_1, e_2) \\
 &= d(e_3, e_1, e_4, e_2) \\
 &= -d(e_1, e_3, e_4, e_2) \\
 &= d(e_1, e_3, e_2, e_4) \\
 &= -d(e_1, e_2, e_3, e_4)
 \end{aligned}$$

Notacja: permutacje $\sigma \in S_n$ (=bijekcje $\{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$) można reprezentować w postaci tabelki, tj. następujący sposób:

$$\begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$$

Ponadto dla $i \neq j \in \{1, \dots, n\}$ oznaczamy przez (i, j) transpozycję i i j , tzn. permutację zamieniającą i i j miejscami (i nie ruszającą innych punktów),

tzn. następującą permutację:

$$(i, j)(k) = \begin{cases} j & k = i \\ i & k = j \\ k & k \neq i, j \end{cases}$$

Podobnie dla parami różnych i_1, i_2, \dots, i_k przez $(i_1, i_2, i_3, \dots, i_k)$ oznaczamy cykl, tj. permutację przekształcającą i_1 na i_2 , i_2 na i_3 itd., aż do i_k , które jest przekształcane na i_1 (pozostałe liczby nie są ruszane).

Składanie permutacji zapisujemy pisząc funkcje obok siebie (nie piszemy \circ).

Przykład 6.6.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 1 & 3 & 4 \end{pmatrix} = (1, 2, 5, 4, 3) = (1, 3)(1, 4)(1, 5)(1, 2)$$

Istotnie,

$$(1, 3)(1, 4)(1, 5)(1, 2)(1) = (1, 3)(1, 4)(1, 5)(2) = 2$$

$$(1, 3)(1, 4)(1, 5)(1, 2)(2) = (1, 3)(1, 4)(1, 5)(1) = (1, 3)(1, 4)(5) = 5$$

$$(1, 3)(1, 4)(1, 5)(1, 2)(3) = (1, 3)(3) = 1$$

$$(1, 3)(1, 4)(1, 5)(1, 2)(4) = (1, 3)(1, 4)(4) = (1, 3)(1) = 3$$

$$(1, 3)(1, 4)(1, 5)(1, 2)(5) = (1, 3)(1, 4)(1, 5)(5) = (1, 3)(1, 4)(1) = (1, 3)(4) = 4$$

Ćwiczenie 6.7. Każdą permutację można przedstawić w każdej z postaci:

1. iloczynu transpozycji postaci $(i, i + 1)$,
2. iloczynu rozłącznych cykli.

Definicja 6.8. 1. *Liczba inwersji (nieporządków)* permutacji $\sigma \in S_n$ to:

$$\begin{aligned} & \text{liczba liczb } > 1 \text{ występujących przed } 1 + \\ & + \text{liczba liczb } > 2 \text{ występujących przed } 2 + \\ & \dots \\ & + \text{liczba liczb } > n \text{ występujących przed } n \\ & = \text{liczba par } (i, j) \text{ takich że } i < j \text{ ale } \sigma(i) > \sigma(j). \end{aligned}$$

(Równoważnie: „liczba skrzyżowań” — patrz przykład poniżej.)

2. parę (i, j) (lub $\{i, j\}$) taką jak po ostatniej równości powyżej nazywamy *nieporządkiem* lub *inwersją* permutacji σ .
3. znak permutacji σ to $\text{sgn}(\sigma) = (-1)^{\text{liczba inwersji } \sigma}$.

Przykład 6.9. Chcemy policzyć znak i liczbę inwersji permutacji $\sigma = (1, 2, 3)(2, 3)(6, 1)(7, 8, 9) \in S_9$.

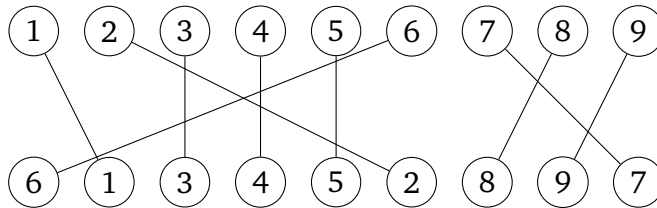
Chcemy przedstawić najpierw σ w postaci tabelkowej:

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

Wyliczając wartości na kolejnych liczbach łatwo dochodzimy do tego, że

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

Następnie łączymy ze sobą liniami wystąpienia każdej liczby na dole i na górze:



Każda inwersja jest reprezentowana przez skrzyżowanie linii (uwaga: skrzyżowania na rysunkach mogą się pokrywać, trzeba je liczyć dla każdej pary linii osobno).

Widać, że mamy 10 inwersji i $\text{sgn}(\sigma) = (-1)^{10} = 1$.

Wniosek 6.10. $d(e_{\sigma(1)}, \dots, e_{\sigma(n)}) = \text{sgn}(\sigma)d(e_1, \dots, e_n)$, więc

$$d(A_1, A_2, \dots) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} d(e_1, \dots, e_n).$$

Dowód. Z rachunku wykonanego wcześniej wynika, że druga część łatwo wynika z pierwszej.

Założmy że $\sigma(i) = 1$. Wtedy $i - 1$ to liczba liczb > 1 występujących przed 1.

$$\begin{aligned} d(e_{\sigma(1)}, \dots, e_{\sigma(n)}) &= d(e_{\sigma(1)}, \dots, e_{\sigma(i-1)}, e_1, e_{\sigma(i+1)}, \dots, e_{\sigma(n)}) \\ &= -d(e_{\sigma(1)}, \dots, e_{\sigma(i-2)}, e_1, e_{\sigma(i-1)}, e_{\sigma(i+1)}, \dots, e_{\sigma(n)}) \\ &\vdots \\ &= (-1)^{i-1} d(e_1, e_{\sigma(1)}, \dots, e_{\sigma(i-1)}, e_{\sigma(i+1)}, \dots, e_{\sigma(n)}) \end{aligned}$$

Następnie postępujemy indukcyjnie: przesuwamy 2 w lewo, następnie 3 itd. Łączna liczba kroków (zmian znaku) jest równa liczbie inwersji w σ , co daje wzór, który chcieliśmy uzyskać. \square

Definicja 6.11. Wyznacznik wektorów $A_1, \dots, A_n \in K^n$ to

$$\det(A_1, A_2, \dots, A_n) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n},$$

gdzie $A_i = \sum_{j=1}^n a_{ji} e_j$ (Tzn. jak wyżej, dla $\det(e_1, \dots, e_n) = 1$.)

Wniosek 6.12. Każda alternująca, n -liniowa funkcja $d: K^n \rightarrow K$ jest postaci $c \cdot \det$ (gdzie $c = d(e_1, e_2, \dots, e_n)$).

Pozostaje sprawdzić, że \det zadany powyższym wzorem jest n -liniowy i alternujący. (Gdyby nie był, to by znaczyło, że jedyna n -liniowa funkcja alternująca na $(K^n)^n$ to funkcja zerowa). Wpierw udowodnimy lemat.

Lemat 6.13. Ustalmy permutację $\sigma = (\sigma(1), \dots, \sigma(n))$ i $j < k \leq n$. Niech $\bar{\sigma}$ będzie permutacją powstałą z σ przez zamianę $\sigma(j)$ i $\sigma(k)$, tzn.

$$\bar{\sigma}(i) = \begin{cases} \sigma(i) & i \neq j, k \\ \sigma(j) & i = k \\ \sigma(k) & i = j. \end{cases}$$

(Innymi słowy, $\bar{\sigma} = \sigma(j, k)$.)

Wtedy $\text{sgn}(\sigma) = -\text{sgn}(\bar{\sigma})$.

Dowód. Weźmy dowolną parę indeksów $i_1 < i_2 \leq n$. Badamy, czy jest ona inwersją σ i $\bar{\sigma}$.

- jeżeli $i_1 = j, i_2 = k$, to

$$(j, k) = (i_1, i_2) \text{ jest inwersją } \bar{\sigma} \iff \bar{\sigma}(j) > \bar{\sigma}(k) \\ \qquad \qquad \qquad \parallel \qquad \parallel \\ \qquad \qquad \qquad \sigma(k) \qquad \sigma(j) \\ \iff (j, k) \text{ nie jest inwersją } \sigma$$

(ubywa lub przybywa 1 inwersja, parzystość liczby inwersji się *zmienia*),

- jeżeli $i_1, i_2 \neq j, k$, to

$$(i_1, i_2) \text{ jest inwersją } \bar{\sigma} \iff \bar{\sigma}(i_1) > \bar{\sigma}(i_2) \\ \qquad \qquad \qquad \parallel \qquad \parallel \\ \qquad \qquad \qquad \sigma(i_1) \qquad \sigma(i_2) \\ \iff (i_1, i_2) \text{ jest inwersją } \sigma$$

(liczba inwersji jest ta sama)

- jeżeli $i_1 < j$, to

$$(i_1, j) \text{ jest inwersją } \bar{\sigma} \iff \bar{\sigma}(i_1) > \bar{\sigma}(j) \\ \qquad \qquad \qquad \parallel \qquad \parallel \\ \qquad \qquad \qquad \sigma(i_1) \qquad \sigma(k) \\ \iff (i_1, k) \text{ jest inwersją } \sigma,$$

i podobnie $(i_1, k) \text{ jest inwersją } \bar{\sigma} \iff (i_1, j) \text{ jest inwersją } \sigma$ (liczba inwersji jest ta sama)

- podobnie jeżeli $i_2 > k$, to $(j, i_2) \text{ jest inwersją } \bar{\sigma} \iff (k, i_2) \text{ jest inwersją } \sigma$ (liczba się nie zmienia),

- jeżeli $j < i < k$, to

$$(j, i) \text{ jest inwersją } \bar{\sigma} \iff \bar{\sigma}(j) > \bar{\sigma}(i) \\ \qquad \qquad \qquad \parallel \qquad \parallel \\ \qquad \qquad \qquad \sigma(k) \qquad \sigma(i) \\ \iff (i, k) \text{ nie jest inwersją } \sigma,$$

i analogicznie $(i, k) \text{ jest inwersją } \bar{\sigma} \iff (j, i) \text{ nie jest inwersją } \sigma$.

Parzystość liczby inwersji zmienia się dwa razy, a więc pozostaje ta sama.

Łącznie zatem parzystość liczby inwersji się zmienia, a więc znak zmienia się na przeciwny. \square

Ćwiczenie 6.14. Ogólnie jeżeli σ, τ są permutacjami, to $\text{sgn}(\sigma \circ \tau) = \text{sgn}(\sigma) \cdot \text{sgn}(\tau)$.

Ćwiczenie 6.15. Jeżeli $\sigma = (a_1, a_2, a_3, \dots, a_n)$ jest cyklem, to $\text{sgn}(\sigma) = (-1)^{n-1}$.

Przykład 6.16. Znak permutacji $(1, 2, 3)(2, 3)(6, 1)(7, 8, 9)$ to $(-1)^2 \cdot (-1)^1 \cdot (-1)^1 \cdot (-1)^2 = 1$ (tak jak wyliczyliśmy wcześniej).

Wniosek 6.17. \det jest n -liniowy i alternujący.

Dowód. n -liniowość:

- zauważmy że jeżeli $A_i = \sum_{j=1}^n a_{ij}e_j$, to $\alpha A_i = \sum_{j=1}^n (\alpha a_{ij})e_j$ (współrzędne αA_i to współrzędne A_i pomnożone przez α). Stąd:

$$\begin{aligned} \det(A_1, A_2, \dots, \alpha A_i, \dots) &= \sum_{\sigma \in S_n} \text{sgn}(\sigma) a_{\sigma(1)1} \cdots (\alpha a_{\sigma(i)i}) \cdots a_{\sigma(n)n} \\ &= \alpha \sum_{\sigma \in S_n} \text{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(i)i} \cdots a_{\sigma(n)n} \\ &= \alpha \det(A_1, \dots, A_i, \dots, A_n) \end{aligned}$$

- jeżeli $A_i = \sum_{j=1}^n a_{ji}e_j$, a $B_i = \sum_{j=1}^n b_{ji}e_j$, to $A_i + B_i = \sum_{j=1}^n (a_{ji} + b_{ji})e_j$, stąd:

$$\begin{aligned} \det(A_1, A_2, \dots, A_i + B_i, \dots) &= \sum_{\sigma \in S_n} \text{sgn}(\sigma) a_{1\sigma(1)} \cdots (a_{\sigma(i)i} + b_{\sigma(i)i}) \cdots a_{\sigma(n)n} \\ &= \sum_{\sigma \in S_n} \text{sgn}(\sigma) \begin{pmatrix} a_{\sigma(1)1} \cdots a_{\sigma(i)i} \cdots a_{\sigma(n)n} + \\ + a_{\sigma(1)1} \cdots b_{\sigma(i)i} \cdots a_{\sigma(n)n} \end{pmatrix} \\ &= \left(\sum_{\sigma \in S_n} \text{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(i)i} \cdots a_{\sigma(n)n} \right) + \\ &\quad + \left(\sum_{\sigma \in S_n} \text{sgn}(\sigma) a_{\sigma(1)1} \cdots b_{\sigma(i)i} \cdots a_{\sigma(n)n} \right) \\ &= \det(A_1, \dots, A_i, \dots, A_n) + \\ &\quad + \det(A_1, \dots, B_i, \dots, A_n) \end{aligned}$$

Alternowanie: załóżmy że $n > 1$ (przypadek $n = 0, 1$ jest trywialny). Chcemy pokazać, że dla $j < k$ mamy:

$$\det(A_1, \dots, A_{j-1}, B, A_{j+1}, \dots, A_{k-1}, B, A_{k+1}, \dots, A_n) = 0.$$

Istotnie, powiedzmy że $B = \sum_j b_j e_j$ i $A_i = \sum_j a_{ji} e_j$.

Oznaczmy $d_\sigma := \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(j-1)(j-1)} b_{\sigma(j)} \cdots a_{\sigma(k-1)(k-1)} b_{\sigma(k)} \cdots a_{n\sigma(n)}$,
tak że

$$\det(A_1, \dots, A_{j-1}, B, A_{j+1}, \dots, A_{k-1}, B, A_{k+1}, \dots, A_n) = \sum_{\sigma \in S_n} d_\sigma.$$

to dla każdej $\sigma \in S_n$ mamy (biorąc $\bar{\sigma} = \sigma(j, k)$ jak w lemacie):

$$\begin{aligned} d_\sigma &= \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(j-1)(j-1)} b_{\sigma(j)} \cdots a_{\sigma(k-1)(k-1)} b_{\sigma(k)} \cdots a_{n\sigma(n)} = \\ &= -\operatorname{sgn}(\bar{\sigma}) a_{\sigma(1)1} \cdots a_{\sigma(j-1)(j-1)} b_{\sigma(j)} \cdots a_{\sigma(k-1)(k-1)} b_{\sigma(k)} \cdots a_{\sigma(n)n} \\ &= -\operatorname{sgn}(\bar{\sigma}) a_{\bar{\sigma}(1)1} \cdots a_{\bar{\sigma}(j-1)(j-1)} b_{\bar{\sigma}(k)} \cdots a_{\bar{\sigma}(k-1)(k-1)} b_{\bar{\sigma}(j)} \cdots a_{\bar{\sigma}(n)n} \\ &= -\operatorname{sgn}(\bar{\sigma}) a_{\bar{\sigma}(1)1} \cdots a_{\bar{\sigma}(j-1)(j-1)} b_{\bar{\sigma}(j)} \cdots a_{\bar{\sigma}(k-1)(k-1)} b_{\bar{\sigma}(k)} \cdots a_{\bar{\sigma}(n)n} \\ &= -d_{\bar{\sigma}}. \end{aligned}$$

W szczególności $d_\sigma + d_{\bar{\sigma}} = 0$.

Stąd nietrudno wywnioskować, że

$$\det(A_1, \dots, A_{j-1}, B, A_{j+1}, \dots, A_{k-1}, B, A_{k+1}, \dots, A_n) = \sum_{\sigma \in S_n} d_\sigma = 0.$$

Dokładniej, jeżeli przez \sim oznaczymy relację na S_n zadaną przez $\sigma \sim \sigma'$ gdy $\sigma' = \sigma$ lub $\sigma' = \bar{\sigma}$, to \sim jest relacją równoważności na S_n o klasach $[\sigma] = \{\sigma, \bar{\sigma}\}$ (przy czym $\sigma \neq \bar{\sigma}$) i mamy:

$$\begin{aligned} \sum_{\sigma \in S_n} d_\sigma &= \sum_{[\sigma] \in S_n / \sim} \left(\sum_{\tau \in [\sigma]} d_\tau \right) \\ &= \sum_{[\sigma] \in S_n / \sim} (d_\sigma + d_{\bar{\sigma}}) \\ &= \sum_{[\sigma] \in S_n / \sim} 0 \\ &= 0. \end{aligned}$$

□

Definicja 6.18. Jeżeli $A \in M_{n \times n}(K)$ ma kolumny A_1, \dots, A_n , to

$$\det A := \det(A_1, \dots, A_n)$$

6.2 Podstawowe własności wyznacznika

Lemat 6.19. *Jeżeli σ^{-1} jest permutacją odwrotną do $\sigma \in S_n$ (tzn. jest funkcją odwrotną, czyli $\sigma^{-1}(i) = j$ wtedy i tylko wtedy gdy $\sigma(j) = i$), to $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)$, a nawet mają one tyle samo inwersji.*

Dowód. Rysunkowo oczywiste — σ^{-1} jest lustrzanym odbiciem σ , więc ma tyle samo skrzyżowań.

Ustalmy $i < j \leq n$. Jeżeli (i, j) jest inwersją σ , to $\sigma(i) > \sigma(j)$ i $(\sigma(j), \sigma(i))$ jest inwersją σ^{-1} : istotnie, $\sigma^{-1}(\sigma(j)) = j > i = \sigma^{-1}(\sigma(i))$. Podobnie jeżeli i, j nie jest inwersją, to $\sigma(i) < \sigma(j)$ i $(\sigma(i), \sigma(j))$ nie jest inwersją σ^{-1} . \square

Twierdzenie 6.20. *Niech $A \in M_{n \times n}(K)$,*

$$1) \det(A_1, A_2, \dots, A_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} \quad (\sigma \text{ na innym indeksie!})$$

$$2) \det(A) = \det(A^T).$$

$$3) \det(AB) = \det A \cdot \det B.$$

$$4) \text{ dla każdego } j \in \{1, \dots, n\}:$$

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} A_{ij},$$

gdzie A_{ij} jest wyznacznikiem macierzy powstałej z A przez wykreślenie kolumny i i wiersza w którym jest a_{ij} (rozwiniecie Laplace'a względem j -tej kolumny).

(Mówimy że A_{ij} , lub macierz którego jest wyznacznikiem, jest minorem macierzy A . A ma też inne minory, o tym później.)

$$4') \text{ dla każdego } i \in \{1, \dots, n\}:$$

$$\det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} A_{ij}.$$

(rozwiniecie Laplace'a względem i -tego wiersza)

$$5) \quad \bullet \det A \text{ nie zmienia się, gdy do dowolnego wiersza [kolumny] dodamy skalarną wielokrotność dowolnego innego wiersza [kolumny],}$$

- jeżeli pomnożymy dowolny wiersz [kolumnę] A przez skalar, to $\det A$ pomnoży się przez ten sam skalar,
 - zamiana wierszy [kolumn] A miejscami zmienia znak $\det A$.
- 6) Jeżeli A jest macierzą (górną- lub dolną-)trójkątną (np. diagonalną), to $\det A$ to iloczyn wyrazów na przekątnej.
- 7) $\det A = 0$ wtedy i tylko wtedy gdy kolumny [wiersze] A są liniowo zależne.

Dowód. 1)

$$\begin{aligned}
 \det A &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} && \text{(definicja det)} \\
 &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma^{-1}) a_{\sigma(1)1} \cdots a_{\sigma(n)n} && \text{(lemat)} \\
 &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma^{-1}) a_{\sigma(1)\sigma^{-1}(\sigma(1))} \cdots a_{\sigma(n)\sigma^{-1}(\sigma(n))} && \text{(definicja } \sigma^{-1}) \\
 &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma^{-1}) a_{1\sigma^{-1}(1)} \cdots a_{n\sigma^{-1}(n)} && \text{(zmiana kolejności czynników)} \\
 &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} && \text{(przeindeksowanie)}
 \end{aligned}$$

- 2) Niech $A^T = (b_{ij})$. Wtedy $b_{ij} = a_{ji}$.

$$\det A^T = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \underset{a_{1\sigma(1)}}{\underset{\parallel}{b_{\sigma(1)1}}} \cdots \underset{a_{n\sigma(n)}}{\underset{\parallel}{b_{\sigma(n)n}}}$$

więc teza wynika z 1.

- 3) Ustalmy dowolne A . Funkcja $d(B_1, \dots, B_n) = \det(AB) = \det(AB_1, \dots, AB_n)$ jest n -liniowa i alternująca, co łatwo sprawdzić, np.:

$$\begin{aligned}
 d(B_1 + B'_1, B_2, \dots, B_n) &= \det(A(B_1 + B'_1), B_2, \dots, B_n) \\
 &= \det(AB_1 + AB'_1, AB_2, \dots, AB_n) \\
 &= \det(AB_1, AB_2, \dots, AB_n) + \det(AB'_1, AB_2, \dots, AB_n).
 \end{aligned}$$

Stąd $d = c \cdot \det$ dla pewnej stałej c . Ale $d(I) = \det(AI) = \det(A) = \det(A) \cdot 1 = \det(A) \cdot \det(I)$, więc $c = \det A$, czyli dla każdej macierzy B zachodzi

$$\det(AB) = d(B) = c \det(B) = \det(A) \det(B).$$

Ponieważ A było dowolne, to kończy dowód 3.

- 4) Z 2. wnioskujemy że 4 i 4' są równoważne (transpozycja zamienia wiersze na kolumny i vice versa).

Pokażemy 4'. Ustalmy i . Niech $d(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} A_{ij}$. Chcemy pokazać, że $d = \det$. Pokażemy najpierw, że d jest n -liniowe i alternujące.

n -liniowość: A_{ij} zależy liniowo od kolumn różnych od j -tej, ale zupełnie nie zależy od j -tej kolumny, zaś a_{ij} zależy liniowo *tylko* od j -tej kolumny (od pozostałych nie zależy). Stąd każde $a_{ij} A_{ij}$ zależy liniowo od wszystkich kolumn.

Alternowanie: założmy że kolumny A : j_1 -ta i j_2 -ta (gdzie $j_1 < j_2$), są równe. Wtedy $A_{ij} = 0$ dla $j \neq j_1, j_2$, bo w odpowiednim minorze są dwie takie same kolumny. Stąd $d(A) = (-1)^{i+j_1} a_{ij_1} A_{ij_1} + (-1)^{i+j_2} a_{ij_2} A_{ij_2} = 0$. Chcemy pokazać, że $d(A) = 0$. Ponieważ z założenia $a_{ij_1} = a_{ij_2}$, wystarczy pokazać że:

$$A_{ij_1} = (-1)^{j_2-j_1-1} A_{ij_2}.$$

Tak jest faktycznie, co wynika z antysymetryczności \det : macierze których wyznacznikami są A_{ij_1}, A_{ij_2} różnią się tylko pozycją, na której znajduje się kolumna pochodząca od j_1 -szej = j_2 -giej kolumny A . Można jedną uzyskać z drugiej przesuwając tę kolumnę $j_2 - j_1 - 1$ razy w prawo, a więc tyleż razy zmieniając znak wyznacznika, co daje żadaną równość.

Stąd $d = c \cdot \det$, a łatwo sprawdzić, że $d(I) = 1 = 1 \cdot \det(I)$, więc $d = 1 \cdot \det = \det$.

- 5) dla kolumn wynika wprost z n -liniowości i antysymetryczności \det , dla wierszy zaś z tego oraz z 2.
- 6) Indukcja względem n : niech A będzie macierzą górnotrójkątną; z rozwinięcia Laplace'a względem pierwszej kolumny wynika, że $\det A$ jest iloczynem wyrazu w lewym górnym rogu oraz minora A_{11} . Ten ostatni z założenia indukcyjnego jest iloczynem wyrazów na przekątnej, co łatwo daje tezę. Przypadek dolnotrójkątny jest analogiczny (albo wynika z 2.).
- 7) Jeżeli kolumny A są liniowo zależne, to pewna kolumna jest kombinacją liniową innych kolumn. To oznacza, że odejmując od tej kolumny

wielokrotności pozostałych kolumn możemy dostać macierz A' o kolumnie złożonej z samych zer. Biorąc rozwinięcie Laplace'a A' względem tej kolumny wnioskujemy że $\det A = \det A' = 0$.

Jeżeli kolumny A są liniowo niezależne, to $\text{rk} A = n$ i stosując operacje W1, W3 (a więc nie zmieniając rzędu i zmieniając co najwyżej znak wyznacznika) możemy sprowadzić A do macierzy diagonalnej A' . Ponieważ $\text{rk} A' = n$, wyrazy na jej przekątnej są niezerowe, więc z 6. $\det A = \pm \det A' \neq 0$. \square

Przykład 6.21.

$$\begin{vmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 1 \\ 2 & 1 & 2 & 0 \\ 3 & -1 & 0 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 2 & 1 & 4 & 0 \\ 3 & -1 & 3 & 3 \end{vmatrix} = (-1)^{1+1} \cdot 1 \cdot \begin{vmatrix} 1 & 1 & 1 \\ 1 & 4 & 0 \\ -1 & 3 & 3 \end{vmatrix} =$$

$$= \begin{vmatrix} 1 & 1 & 1 \\ 0 & 3 & -1 \\ 0 & 4 & 4 \end{vmatrix} = 1 \cdot \begin{vmatrix} 3 & -1 \\ 4 & 4 \end{vmatrix} = 12 + 4 = 16$$

$$\begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 & 1 \\ 1 & 1 & 3 & 1 & 1 \\ 1 & 1 & 1 & 4 & 1 \\ 1 & 1 & 1 & 1 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{vmatrix} = 1 \cdot 2 \cdot 3 \cdot 4 = 24.$$

Przykład 6.22 (wyznacznik Vandermonde'a).

$$\begin{aligned}
 V(x_1, x_2, \dots, x_n) &= \begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \cdots & x_3^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{vmatrix} \\
 &= \begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 0 & x_2 - x_1 & x_2^2 - x_1^2 & \cdots & x_2^{n-1} - x_1^{n-1} \\ 0 & x_3 - x_1 & x_3^2 - x_1^2 & \cdots & x_3^{n-1} - x_1^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & x_n - x_1 & x_n^2 - x_1^2 & \cdots & x_n^{n-1} - x_1^{n-1} \end{vmatrix} \\
 &= 1 \cdot \begin{vmatrix} x_2 - x_1 & x_2^2 - x_1^2 & \cdots & x_2^{n-1} - x_1^{n-1} \\ x_3 - x_1 & x_3^2 - x_1^2 & \cdots & x_3^{n-1} - x_1^{n-1} \\ \vdots & \vdots & & \vdots \\ x_n - x_1 & x_n^2 - x_1^2 & \cdots & x_n^{n-1} - x_1^{n-1} \end{vmatrix} \\
 &= (x_2 - x_1)(x_3 - x_2) \cdots (x_n - x_1) \begin{vmatrix} 1 & x_2 + x_1 & x_2^2 + x_1x_2 + x_1^2 & \cdots & x_2^{n-1} - x_1^{n-1} \\ 1 & x_3 + x_1 & x_3^2 + x_1x_3 + x_1^2 & \cdots & x_3^{n-1} - x_1^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n + x_1 & x_n^2 + x_1x_n + x_1^2 & \cdots & x_n^{n-1} - x_1^{n-1} \end{vmatrix} \\
 &= (x_2 - x_1)(x_3 - x_2) \cdots (x_n - x_1) \begin{vmatrix} 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \cdots & x_3^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{vmatrix} \\
 &= (x_2 - x_1)(x_3 - x_2) \cdots (x_n - x_1)(x_3 - x_2)(x_4 - x_2) \cdots (x_n - x_{n-1}) \\
 &= \prod_{i < j} (x_j - x_i)
 \end{aligned}$$

Przykład 6.23. Rozważmy wektory postaci $\begin{pmatrix} 1 \\ t \\ t^2 \end{pmatrix} \in \mathbb{R}^3$. Trzy takie wektory

(dla t_1, t_2, t_3) są liniowo niezależne wtedy i tylko wtedy gdy

$$\det \left(\begin{pmatrix} 1 \\ t_1 \\ t_1^2 \end{pmatrix} + \begin{pmatrix} 1 \\ t_2 \\ t_2^2 \end{pmatrix} + \begin{pmatrix} 1 \\ t_3 \\ t_3^2 \end{pmatrix} \right) = (t_2 - t_1)(t_3 - t_1)(t_3 - t_2) \neq 0$$

Wniosek 6.24. Jeżeli $\dim V = n < \infty$ i B jest bazą V , to dla każdych $v_1, \dots, v_n \in V$ zachodzi

$$v_1, v_2, \dots, v_n \text{ są lnz} \iff \det([v_1]_B, \dots, [v_n]_B) \neq 0.$$

Dowód. $[\cdot]_B$ jest izomorfizmem między V i K^n , więc v_1, \dots, v_n są lnz wtedy i tylko wtedy gdy $[v_1]_B, \dots, [v_n]_B$ są lnz, a te są lnz wtedy i tylko wtedy gdy ich wyznacznik jest niezerowy (z twierdzenia). \square

Przykład 6.25. Chcemy sprawdzić, że funkcje $f(x) = \sin(x) + \cos(x)$, $g(x) = 6\sin(x) - 7\cos(x)$ są lnz w $V = \text{Lin}(\sin, \cos) \leq \mathbb{R}^{\mathbb{R}}$ (wiedząc, że \sin i \cos są liniowo niezależne).

Biorąc bazę $B = (\sin, \cos)$ widzimy, że $[f]_B = (1, 1)^T$, $[g]_B = (6, -7)^T$ i $\det \begin{pmatrix} 1 & 6 \\ 1 & -7 \end{pmatrix} = -7 - 6 = -13 \neq 0$, więc f i g faktycznie są liniowo niezależne.

Definicja 6.26. Mówimy że macierz A jest *osobliwa*, jeżeli $\det A = 0$. W przeciwnym wypadku mówimy, że jest *nieosobliwa*.

Wniosek 6.27. Macierz $A \in M_{n \times n}(K)$ jest odwracalna $\iff A$ jest nieosobliwa.

Dowód.

A jest odwracalna $\iff F_A: K^n \rightarrow K^n$ jest izomorfizmem

$$\iff F_A \text{ jest „na”}$$

$$\iff \text{im } F_A = K^n$$

$$\iff \dim \text{im } F_A = n$$

$$\iff \text{rk } A = n$$

$$\iff \text{kolumny } A \text{ są lnz}$$

$$\iff \det A \neq 0$$

\square

Definicja 6.28. *Minor* macierzy A (niekoniecznie kwadratowej) to macierz kwadratowa powstała z A przez wykreślenie pewnej liczby wierszy lub kolumn, lub wyznacznik takiej macierzy. *Rozmiar* minora to liczba pozostałych wierszy/kolumn.

Przykład 6.29. Przykładowe minory macierzy $A = \begin{pmatrix} 1 & 5 & 1 & -2 \\ 3 & 3 & 3 & 1 \\ 2 & 1 & 2 & -5 \end{pmatrix}$:

$$\begin{aligned}
& \bullet \begin{pmatrix} 1 & 5 & -2 \\ 3 & 3 & 1 \\ 2 & 1 & -5 \end{pmatrix} = \begin{pmatrix} 1 & 5 & 1 & -2 \\ 3 & 3 & 3 & 1 \\ 2 & 1 & 2 & -5 \end{pmatrix}, \\
& \bullet \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 5 & 1 & -2 \\ 3 & 3 & 3 & 1 \\ 2 & 1 & 2 & -5 \end{pmatrix}, \\
& \bullet (-5) = \begin{pmatrix} 1 & 5 & 1 & -2 \\ 3 & 3 & 3 & 1 \\ 2 & 1 & 2 & -5 \end{pmatrix},
\end{aligned}$$

Uwaga 6.30. Jeżeli macierz A jest minorem macierzy B , która jest minorem macierzy C , to A jest również minorem C .

Stwierdzenie 6.31. Niech A będzie dowolną macierzą (niekoniecznie kwadratową). Wtedy $\text{rk} A =$ maksymalny rozmiar niezerowego minora (wyznacznika).

Dowód. \geq : załóżmy że pewien minor $k \times k$ jest niezerowy. Niech j_1, \dots, j_k będą indeksami kolumn, które w nim występują (nie są wykreślone). Wtedy te same kolumny są liniowo niezależne (nawet pomijając część współrzędnych), więc $\text{rk} A \geq k$.

\leq : załóżmy że kolumny o indeksach j_1, \dots, j_k są liniowo niezależne. Skreślamy pozostałe kolumny, a następnie stosujemy kolumnową eliminację Gaussa (operacje kolumnowe $K1, K3$) do powstałej macierzy A' .

To daje nam macierz schodkową szerokości k , o k wyrazach wiodących. Skreślając wszystkie wiersze prócz tych, w których są wyrazy wiodące, otrzymujemy macierz trójkątną $k \times k$ o niezerowych wyrazach na przekątnej, a więc niezerowym wyznaczniku. Skreślając odpowiednie wiersze z A' otrzymujemy minor $k \times k$ (który jest też minorem A !) o tym samym, niezerowym wyznaczniku. \square

Uwaga 6.32. Stwierdzenie 6.31 można zinterpretować następująco: jeżeli $F: K^n \rightarrow K^m$ jest dowolnym odwzorowaniem liniowym, to dla pewnych $i_1 < i_2 < \dots < i_k \leq n$ i $j_1 < j_2 < \dots < j_k \leq m$, gdzie $k = \text{rk} F$, odwzorowanie

$$\pi_{\vec{j}} \circ F|_{\vec{i}}$$

jest izomorfizmem, gdzie $F|_{\vec{i}}$ to obcięcie F do podprzestrzeni K^n rozpiętej przez wektory e_{i_1}, \dots, e_{i_k} , a $\pi_{\vec{j}}$ to rzut „prostokątny” na współrzędne j_1, j_2, \dots, j_k .

W szczególności jeżeli $V \leq K^n$ jest k -wymiarową podprzestrzenią, to rzut V na pewne k współrzędnych jest izomorfizmem.

x

Przykład 6.33. Chcemy wyznaczyć wymiar przestrzeni rozwiązań UJ

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 - x_2 + x_3 - x_4 = 0 \\ x_2 - x_3 + x_4 = 0 \end{cases}$$

Szukamy wymiaru jądra odwzorowania opisanego przez macierz $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 0 & 1 & -1 & 1 \end{pmatrix}$, czyli $4 - \text{rk} A$. Mamy

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 1 & -1 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ 0 & -2 & 0 \\ 0 & 0 & -1 \end{vmatrix} = 2 \neq 0,$$

więc $\text{rk} A = 3$, a $\dim \ker F_A = 4 - 3 = 1$.

Stwierdzenie 6.34. *Każdy niezerowy minor można rozszerzyć do niezerowego minora maksymalnego wymiaru.*

Dowód. Weźmy niezerowy minor $k \times k$ macierzy A o $\text{rk} A > k$. Występujące w nim kolumny są liniowo niezależnymi wektorami w obrazie F_A . Możemy do nich dobrać inne kolumny A , otrzymując bazę $\text{im } F_A$. Wykreślamy pozostałe kolumny, a następnie transponujemy A , otrzymując macierz B której wiersze są liniowo niezależne.

Kolumny B odpowiadające wierszom wyjściowego minora są liniowo niezależne, więc można do nich dobrać inne kolumny B , otrzymując \square

6.3 Wzory Cramera

Rozważamy równanie

$$AX = Y, \quad (*)$$

gdzie $A = (A_1, \dots, A_n) \in M_{n \times n}(K)$ jest nieosobliwa, tzn. $\det A \neq 0$. Wtedy oczywiście $(*)$ ma jedyne rozwiązanie $X = A^{-1}Y$.

Twierdzenie 6.35 (Wzory Cramera). Dla A, Y jak powyżej, rozwiązaniem $(*)$ jest $X = (x_1, \dots, x_n)^\top$, gdzie

$$x_k = \frac{\det(A_1, \dots, A_{k-1}, Y, A_{k+1}, \dots, A_n)}{\det A}.$$

Dowód. Sprawdzimy że $AX = Y$.

$$\begin{aligned} (AX)_i &= \sum_{j=1}^n a_{ij} x_j \\ &= \sum_{j=1}^n a_{ij} \frac{1}{\det A} \cdot \sum_{l=1}^n (-1)^{l+j} y_l A_{l,j} \quad (\text{rozwiniecie Laplace'a w } j\text{-tej kolumnie}) \\ &= \sum_{l=1}^n \frac{y_l}{\det A} \underbrace{\sum_{j=1}^n (-1)^{l+j} a_{ij} A_{l,j}}_{\parallel} \quad (\text{zmiana kolejności sumowania}) \\ &\quad \begin{cases} 0 & l \neq i \\ \det A & l = i \end{cases} \\ &= \sum_{l=1}^n \begin{cases} \frac{y_l}{\det A} \cdot \det A & l = i \\ \frac{y_l}{\det A} \cdot 0 & l \neq i \end{cases} \\ &= y_i. \end{aligned}$$

Wyjaśnienie „klamerki”:

$$\sum_{j=1}^n (-1)^{l+j} a_{ij} A_{l,j} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{(l-1)1} & a_{(l-1)2} & \cdots & a_{(l-1)n} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{(l+1)1} & a_{(l+1)2} & \cdots & a_{(l+1)n} \\ \vdots & & & \end{vmatrix} = \begin{vmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(l-1)} \\ A^{(i)} \\ A^{(l+1)} \\ \vdots \end{vmatrix}$$

to rozwinięcie Laplace'a względem l -tego wiersza wyznacznika macierzy powstałej z A przez zastąpienie l -tego wiersza przez i -ty wiersz A (pozostałe wiersze, w tym i -ty, pozostają bez zmian).

- Jeżeli $i = l$, to to wciąż A .

- Jeżeli $i \neq l$, to jest to macierz o dwóch takich samych wierszach, więc jej wyznacznik jest zerowy. \square

Definicja 6.36. Dla macierzy $A = (a_{ij})$, iloczyn $(-1)^{i+j}A_{ij}$ (gdzie A_{ij} to ij -ty minor A) nazywamy *dopełnieniem algebraicznym* wyrazu a_{ij} macierzy A .

Macierz dołączona do A to macierz transponowana do macierzy dopełnień algebraicznych A , tzn. $\text{adj}(A) = A^\vee = (b_{ij})_{i,j}$, gdzie $b_{ij} = (-1)^{i+j}A_{ji}$.

$$\begin{aligned} A^\vee &= \text{adj}(A) = ((-1)^{i+j}A_{ij})_{i,j}^\top = ((-1)^{i+j}A_{ji})_{i,j} \\ &= \begin{pmatrix} +A_{11} & -A_{21} & +A_{31} & \cdots & (-1)^{n+1}A_{n1} \\ -A_{12} & +A_{22} & -A_{32} & \cdots & \\ \vdots & & & & \\ (-1)^{n+1}A_{1n} & (-1)^{n+2}A_{2n} & (-1)^{n+3}A_{3n} & \cdots & +A_{nn} \end{pmatrix} \end{aligned}$$

(Uwaga na znaki i na transpozycję!)

Wniosek 6.37. Niech $A \in M_{n \times n}(K)$. Wtedy:

- $A \cdot \text{adj}(A) = \text{adj}(A) \cdot A = \det A \cdot I$.
- Jeżeli $\det A \neq 0$, to $A^{-1} = \frac{1}{\det A} \text{adj} A$.

Dowód. Druga część wynika natychmiast z pierwszej. Pierwsza część: oznaczmy $C = (c_{ij}) = A \cdot \text{adj}(A)$, $\text{adj}(A) = B = (b_{ij})$. Wtedy

$$\begin{aligned} c_{ij} &= \sum_{l=1}^n a_{il} b_{lj} \\ &= \sum_{l=1}^n a_{il} (-1)^{j+l} A_{jl} && \text{(definicja } B = \text{adj}(A)) \\ &= \begin{vmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(j-1)} \\ A^{(i)} \\ A^{(j+1)} \\ \vdots \end{vmatrix} && \text{(rozwinięcie Laplace'a względem } j\text{-tego wiersza)} \\ &= \begin{cases} \det A & j = i \\ 0 & j \neq i, \end{cases} \end{aligned}$$

stąd $C = \det A \cdot I$. Podobnie dla $\text{adj}(A) \cdot A$ (z tym że tam pojawia się rozwinięcie kolumnowe zamiast wierszowego). \square

Wniosek 6.38. *Odwracanie macierzy wyraża się wielomianem. W szczególności np. gdy $K = \mathbf{R}$ lub $K = \mathbf{C}$, to funkcja $A \mapsto A^{-1}$ jest funkcją ciągłą (a nawet C^∞) na zbiorze macierzy odwracalnych (jako otwartym podziorze K^{n^2}).*

Przykład 6.39.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \text{adj} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 0 \cdot I$$

$$\begin{aligned} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}^{-1} &= \frac{1}{\begin{vmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{vmatrix}} \cdot \text{adj} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix} & -\begin{vmatrix} 0 & 1 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 0 & 2 \\ 0 & 0 \end{vmatrix} \\ -\begin{vmatrix} 2 & 3 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 0 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 0 & 0 \end{vmatrix} \\ \begin{vmatrix} 2 & 3 \\ 2 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 0 & 2 \end{vmatrix} \end{pmatrix}^T = \\ &= \frac{1}{2} \begin{pmatrix} 2 & 0 & 0 \\ -2 & 1 & 0 \\ -4 & -1 & 2 \end{pmatrix}^T = \frac{1}{2} \begin{pmatrix} 2 & -2 & -4 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & -1 & -2 \\ 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

6.4 Wyznacznik i wielomian charakterystyczny odwzorowania liniowego

Definicja 6.40. Niech V, W będą przestrzeniami liniowymi. Wtedy $\text{Hom}(V, W)$ to zbiór odwzorowań liniowych $V \rightarrow W$.

Fakt 6.41. $\text{Hom}(V, W)$ jest przestrzenią liniową. (Podobnie jak $\text{Hom}(V, K) = V^*$.)

Fakt 6.42. Zbiór $M_{k \times n}(K)$ jest przestrzenią liniową (ze zwykłym dodawaniem macierzy i mnożeniem przez skalary), izomorficzną z K^{kn} .

6.4. WYZNACZNIK I WIELOMIAN CHARAKTERYSTYCZNY ODWZOROWANIA LINIOWEGO 83

Fakt 6.43. Jeżeli B jest bazą V , a C jest bazą W i V, W są skończenie wymiarowe, to odwzorowanie

$$\text{Hom}(V, W) \ni F \mapsto m_C^B(F) \in M_{\dim W \times \dim V}(K)$$

jest izomorfizmem.

Dowód. Bijekcja — było. Addytywność:

$$\underbrace{[(F+G)(b_i)]_C}_{\parallel \text{ } i\text{-ta kolumna } m_C^B(F+G)} = [F(b_i) + G(b_i)]_C = \underbrace{[F(b_i)]_C}_{\parallel \text{ } i\text{-ta kolumna } m_C^B(F)} + \underbrace{[G(b_i)]_C}_{\parallel \text{ } i\text{-ta kolumna } m_C^B(G)}.$$

Jednorodność podobnie. □

Wniosek 6.44. Jeżeli V, W są skończenie wymiarowe, to $\dim \text{Hom}(V, W) = \dim V \cdot \dim W$.³

Przykład 6.45. Rozważmy $F: V \rightarrow V$ (V : skończenie wymiarowa).

Wtedy

$$(F - \alpha \text{id}_V)(v) = F(v) - \alpha \text{id}_V(v) = F(v) - \alpha v.$$

Dla dowolnej B — bazy V mamy:

$$m_B^B(F - \alpha \text{id}_V) = m_B^B(F) - \alpha m_B^B(\text{id}_V) = m_B^B(F) - \alpha I.$$

Stwierdzenie 6.46. Jeżeli $F \in \text{Hom}(V, V)$, to $\det m_B^B(F)$ nie zależy od wyboru B — bazy V .

Dowód. Weźmy dowolne bazy B, C przestrzeni V . Wtedy:

$$\begin{aligned} \det m_C^C(F) &= \det(m_C^B(\text{id}) \cdot m_B^B(F) \cdot m_B^C(\text{id})) \\ &= \det m_C^B(\text{id}) \cdot \det m_B^B(F) \cdot \det m_B^C(\text{id}) \\ &= \det m_B^B(F) \cdot (\det m_C^B(\text{id}) \cdot \det(m_B^C(\text{id}))) \\ &= \det m_B^B(F) \cdot \det(m_C^B(\text{id}) \cdot m_B^C(\text{id})) \\ &= \det m_B^B(F) \cdot \det m_C^C(\text{id}) \\ &= \det m_B^B(F) \cdot \det I \\ &= \det m_B^B(F) \cdot 1 \\ &= \det m_B^B(F) \end{aligned}$$

□

³Ogólnie $\dim \text{Hom}(V, W) = \dim W \cdot \dim V^*$. Można to wywnioskować z izomorfizmu $\text{Hom}(V, W) \cong V^* \otimes W$ — patrz konwersatorium.

Definicja 6.47. Jeżeli $F: V \rightarrow V$ i $\dim V < \infty$, definiujemy

$$\det F := \det m_B^B(F),$$

gdzie B jest dowolną bazą V .

Definiujemy też *wielomian charakterystyczny* F jako

$$\chi_F(x) := \det(F - x \cdot \text{id}_V) = \det(m_B^B(F) - xI).^4$$

Ślad F to suma wyrazów na przekątnej w $m_B^B(F)$, a równoważnie

$$\text{tr}(F) = (-1)^{n-1} \cdot (\text{współczynnik przy } x \text{ w } \chi_F(x)).$$

(Z tego wynika, że ślad nie zależy od wyboru bazy B !)

Dla macierzy $A = (a_{ij}) \in M_{n \times n}(K)$ podobnie definiujemy:

$$\begin{aligned} \chi_A(x) &:= \chi_{F_A}(x) = \det(A - x \cdot I) = \begin{vmatrix} a_{11} - x & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} - x & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} - x & \cdots & a_{3n} \\ \vdots & & & & \\ a_{nn} & a_{n2} & a_{n3} & \cdots & a_{nn} - x \end{vmatrix} = \\ &= \det A + \dots + (-1)^{n-1} (a_{11} + a_{22} + \dots + a_{nn}) x^{n-1} + (-1)^n x^n, \end{aligned}$$

oraz $\text{tr}(A)$ = suma wyrazów na przekątnej w A .

Ćwiczenie 6.48. Dla dowolnych macierzy $A \in M_{n \times m}(K)$, $B \in M_{m \times n}(K)$ zachodzi $\text{tr}(AB) = \text{tr}(BA)$.

⁴Tu jest małe oszustwo: tak naprawdę nie pokazaliśmy do końca, że taka definicja nie zależy od B , bo x jest zmienną, a nie skalar.

Rozdział 7

Rozkłady przestrzeni i ich endomorfizmów

7.1 Wewnętrzne sumy proste

Fakt 7.1. *Jeżeli $A, B \subseteq V$ są liniowo niezależne, to*

$$\text{Lin}(A) \cap \text{Lin}(B) = \{0\} \iff AB \text{ jest liniowo niezależny.}$$

Dowód. \Leftarrow było na ćwiczeniach. \Rightarrow : weźmy dowolne skalary α_a, β_b takie że $0 = \sum_{a \in A} \alpha_a a + \sum_{b \in B} \beta_b b$. Wtedy $v := \sum_{a \in A} \alpha_a a$ równa się $\sum_{b \in B} (-\beta_b) b$, więc z założenia $v = 0$, a więc z liniowej niezależności A i B wnioskujemy, że wszystkie $\alpha_a = \beta_b = 0$, czyli AB jest lnz. \square

Definicja 7.2. Niech $V_1, V_2, \dots, V_n \leq V$. Mówimy że V jest (wewnętrzną) sumą prostą V_1, \dots, V_n , $V = V_1 \oplus V_2 \oplus \dots \oplus V_n$, jeżeli każdy $v \in V$ przedstawia się jednoznacznie w postaci

$$v = v_1 + v_2 + \dots + v_n,$$

gdzie $v_k \in V_k$.

Przykład 7.3. 1. $\mathbf{R}^2 = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \middle| x \in \mathbf{R} \right\} \oplus \left\{ \begin{pmatrix} 0 \\ y \end{pmatrix} \middle| y \in \mathbf{R} \right\}$ (Porównaj to ze stwierdzeniem, że $\mathbf{R}^2 \cong \mathbf{R} \times \mathbf{R} = \mathbf{R} \oplus \mathbf{R}$, tzn. \mathbf{R}^2 jest izomorficzne z (zewnątrzną!) sumą prostą/produktem dwóch kopii \mathbf{R} ; zewnętrzną sumą prostą to operacja na przestrzeniach liniowych, wewnętrzną to relacja pomiędzy przestrzenią a ciągiem jej podprzestrzeni!)

2. Jeżeli V jest dowolną przestrzenią liniową o bazie B i $B = B_1 \sqcup B_2 \sqcup \dots \sqcup B_n$ (tzn. B_k są parami rozłączne i sumują się do B), to

$$V = \text{Lin}(B_1) \oplus \text{Lin}(B_2) \oplus \text{Lin}(B_3) \dots \oplus \text{Lin}(B_n).$$

Dla $n = 2$: ponieważ $V = \text{Lin}(B_1 \cup B_2)$, mamy $V = \text{Lin}(B_1) + \text{Lin}(B_2)$, co daje „zapisywanie się”. Jednoznaczność: założmy że $v = v_1 + v_2 = v'_1 + v'_2$. Wtedy $v_1 - v'_1 = v_2 - v'_2 \in \text{Lin}(B_1) \cap \text{Lin}(B_2)$, więc z faktu $v_1 - v'_1 = v_2 - v'_2 = 0$. Dla $n > 2$ dowodzimy indukcyjnie (lub analogicznie).

3.

$$\begin{aligned} \mathbf{R}_8[X] &= \{a + bX^2 + cX^4 + dX^6 + ex^8 \mid a, b, c, d, e \in \mathbf{R}\} \oplus \{aX + bX^3 + cX^5 + dX^7 \mid a, b, c, d \in \mathbf{R}\} \\ &= \{P \in \mathbf{R}_8[X] \mid P(-X) = P(X)\} \oplus \{P \in \mathbf{R}_8[X] \mid P(X) = -P(-X)\} \end{aligned}$$

4. $V = \mathbf{R}^2$, $V_1 = OX$, $V_2 = OY$,

$$\mathbf{R}^2 \neq V_1 \oplus V_2 \oplus V_2 :$$

istotnie, dowolny $v = \begin{pmatrix} x \\ y \end{pmatrix}$ zapisuje się w postaci

$$v = \underbrace{\begin{pmatrix} x \\ 0 \end{pmatrix}}_{v_1} + \underbrace{\begin{pmatrix} 0 \\ y \end{pmatrix}}_{v_2} + \underbrace{\begin{pmatrix} 0 \\ 0 \end{pmatrix}}_{v_3} = \underbrace{\begin{pmatrix} x \\ 0 \end{pmatrix}}_{v'_1} + \underbrace{\begin{pmatrix} 0 \\ 0 \end{pmatrix}}_{v'_2} + \underbrace{\begin{pmatrix} 0 \\ y \end{pmatrix}}_{v'_3},$$

a więc niejednoznacznie.

7.2 Liniowa niezależność podprzestrzeni

Definicja 7.4. Niech $V_1, \dots, V_k \leq V$. Mówimy że V_1, \dots, V_k są liniowo niezależne, jeżeli dla $v_i \in V_i$ mamy

$$\sum_{i=1}^k v_i = 0 \iff v_1 = v_2 = \dots = v_k = 0.$$

Uwaga 7.5. Jeżeli $u_1, \dots, u_k \in V$ są niezerowe, to układ u_1, u_2, \dots, u_k jest lnz wtedy i tylko wtedy gdy $\text{Lin}(u_1), \text{Lin}(u_2), \dots, \text{Lin}(u_k)$ są lnz.

Dowód. Załóżmy że u_1, \dots, u_k są lnz. Weźmy $v_i \in \text{Lin}(u_i)$ takie że $\sum_{i=1}^k v_i = 0$. Wtedy $v_i = \alpha_i u_i$, czyli $\sum_i \alpha_i u_i = 0$, więc z lnz wektorów u_i wnioskujemy $\alpha_i = 0$, czyli $v_i = 0$.

W drugą stronę, załóżmy że $\text{Lin}(u_i)$ są lnz i weźmy dowolne $\alpha_1, \dots, \alpha_k$ takie że $\sum_i \alpha_i u_i = 0$. Wtedy $v_i := \alpha_i u_i \in \text{Lin}(u_i)$, więc z założenia $v_i = 0$, więc (ponieważ $u_i \neq 0$), $\alpha_i = 0$, czyli u_i są lnz. \square

Lemat 7.6. V_1, \dots, V_n są lnz wtedy i tylko wtedy gdy dla każdego i zachodzi $V_i \cap \sum_{j \neq i} V_j = \{0\}$.

Dowód. Załóżmy że V_1, \dots, V_n są lnz. Niech $v \in V_i \cap \sum_{j \neq i} V_j$ (dla pewnego i). Wtedy z definicji $\sum_j V_j$ mamy $v = \sum_{j \neq i} v_j$ dla pewnych $v_j \in V_j$ i dla $v_i = -v$ mamy $0 = \sum_i v_i$, czyli (z lnz) $v_i = 0$, więc $v = -v_i = 0$.

Założmy że V_1, \dots, V_n są lz, czyli $0 = \sum_j v_j$ dla pewnych $v_j \in V_j$, takich że pewien $v_i \neq 0$. Wtedy $0 \neq -v_i = \sum_{j \neq i} v_j \in V_i \cap \sum_{j \neq i} V_j$. \square

Wniosek 7.7. $V_1, V_2 \leq V$ są lnz $\iff V_1 \cap V_2 = \{0\}$.

Fakt 7.8. Niech $V_1, \dots, V_n \leq V$. Wtedy NWSR:

1. $V = V_1 \oplus V_2 \dots \oplus V_n$.
2. $\Phi: V_1 \times V_2 \times \dots \times V_n \rightarrow V$ zadana wzorem $\Phi(v_1, \dots, v_n) = v_1 + \dots + v_n$ jest izomorfizmem.
3. V_1, \dots, V_n są liniowo niezależne i $V_1 + V_2 + \dots + V_n = V$.
4. (Dla $\dim V < \infty$.) V_1, \dots, V_n są lnz i $\dim V = \dim V_1 + \dots + \dim V_n$.
5. (Dla $\dim V < \infty$.) $V_1 + \dots + V_n = V$ i $\dim V = \dim V_1 + \dots + \dim V_n$.

Dowód. $1 \iff 2$: łatwo sprawdzić że Φ jest zawsze liniowe i

- Φ jest „na” \iff każdy $v \in V$ przedstawia się jako suma wektorów z V_i ,
- Φ jest „1-1” \iff każdy $v \in V$ przedstawia się na co najwyżej jeden sposób jako suma wektorów z V_i .

$2 \iff 3$:

- Φ jest 1-1 $\iff \ker \Phi = \{0\} \iff V_1, \dots, V_n$ są lnz

• Φ jest „na” $\iff V = \text{im } \Phi = V_1 + \dots + V_n$.

$2 + 3 \implies 4, 5$ łatwe, bo $\dim(V_1 \times \dots \times V_n) = \dim V_1 + \dots + \dim V_n$ i Φ zachowuje wymiar (jako izomorfizm).

$4 \implies 2$: skoro V_1, \dots, V_n są lnz, to Φ jest 1-1, a z równości wymiarów wynika, że wymiar dziedziny i przeciwdziedziny Φ są równe, więc z twierdzenia o rzędzie $\dim \text{im } \Phi = \dim \text{dom } \Phi - 0 = \dim V - 0 = \dim V$. Skoro $\dim V$ jest skończony, to $\text{im } \Phi = V$.

$5 \implies 2$: skoro $V_1 + \dots + V_n = V$, to Φ jest „na”, i podobnie z równości wymiarów i twierdzenia o indeksie wynika że $\dim \ker \Phi = 0$, czyli Φ jest 1-1. \square

Wniosek 7.9. ($n = 2$, $\dim V < \infty$.) NWSR:

1. $V = V_1 \oplus V_2$,
2. $V_1 \cap V_2 = \{0\}$ i $V_1 + V_2 = V$,
3. $V_1 \cap V_2 = \{0\}$ i $\dim V = \dim V_1 + \dim V_2$,
4. $V_1 + V_2 = V$ i $\dim V = \dim V_1 + \dim V_2$.

Definicja 7.10. Jeżeli $W \leq V$, to mówimy że przestrzeń $W' \leq V$ jest dopełnicza do W w V jeżeli $V = W \oplus W'$. [rysunek]

Lemat 7.11. Jeżeli $W_1, W_2 \leq V$ są liniowo niezależne i $W_2'' \supseteq W_2$ jest taka że $W_1 + W_2'' = V$, to istnieje W_2' dopełnicza do W_1 , taka że $W_2 \subseteq W_2' \subseteq W_2''$

Dowód. Niech B_1 będzie bazą W_1 , a B_2 będzie bazą W_2 . Z Faktu 7.1 wiemy, że $B_1 \cup B_2$ jest lnz i oczywiście $B_1 \cup B_2 \cup W_2'' = B_1 \cup W_2''$ rozpina V , można więc znaleźć bazę C taką że $B_1 \cup B_2 \subseteq C \subseteq B_1 \cup W_2''$. Wtedy $B_2' = C \setminus B_1 \subseteq W_2''$ i $B_2' \supseteq B_2$, więc

$$\text{Lin}(B_2) = W_2 \leq W_2' \leq W_2'',$$

$W_1 + W_2' = \text{Lin}(C) = V$, oraz W_1, W_2' są liniowo niezależne (bo suma ich baz jest liniowo niezależna), więc $V = W_1 \oplus W_2'$. \square

Biorąc $W_2 = \{0\}$ i $W_2'' = V$ dostajemy wniosek.

Wniosek 7.12. Jeżeli V jest przestrzenią liniową i $W \leq V$, to W jest dopełnialna w V , tzn. V ma podprzestrzeń W' dopełniczą do W .¹

Ćwiczenie 7.13. Jeżeli $V = V_1 \oplus V_2$, to mamy odwzorowania liniowe $\pi_1: V \rightarrow V_1$, $\pi_2: V \rightarrow V_2$ takie że każdy $v \in V$ zapisuje się jako $v = \pi_1(v) + \pi_2(v)$. Podobnie dla większej liczby składników prostych. W szczególności z poprzedniego wniosku wynika, że jeżeli $V' \leq V$ jest dowolna, to istnieje rzut na V' , tzn. liniowe $\pi: V \rightarrow V'$ takie że $\pi|_{V'} = \text{id}_{V'}$.

7.3 Podprzestrzenie niezmiennicze

Definicja 7.14. Niech $W \leq V$, a F będzie endomorfizmem V . Mówimy że W jest $[F]$ -niezmiennicza, jeżeli $F[W] \subseteq W$ (innymi słowy, $(\forall w \in W) F(w) \in W$).

Uwaga 7.15. $W = V$ i $W = \{0\}$ zawsze są niezmiennicze.

Przykład 7.16. 0. Jeżeli $F(v) = \lambda v$, to $\text{Lin}(v)$ jest F -niezmiennicza (bo dla $w \in \text{Lin}(v)$ mamy $w = \alpha v$, więc $F(w) = \alpha \lambda v \in \text{Lin}(v)$).

Uwaga 7.17. Ogólnie, jeżeli $F \in \text{End}(V)$, to 1-wymiarowe przestrzenie F -niezmiennicze to dokładnie przestrzenie powyższej postaci (tzn. proste rozpinane przez wektory własne).

1. Rozważmy $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \in M_{2 \times 2}(\mathbf{R})$. Jakie są F_A -niezmiennicze podprzestrzenie \mathbf{R}^2 ?

- $\{0\}$,
- \mathbf{R}^2 .
- $\text{Lin}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)$ (wektor własny).

¹Jeżeli rozważamy przestrzenie liniowo-topologiczne (TVS), tzn. przestrzenie liniowe z kompatybilną topologią, to powyższy lemat nie jest prawdziwy, tzn. jeżeli V jest (nie-skończenie wymiarową) TVS i $W \leq V$ jest domkniętą podprzestrzenią, to zazwyczaj nie ma domkniętej $W' \leq V$ dopełniczej do W (intuicyjnie: przestrzenie dopełnicze są na ogół bardzo dziwne i trudno je zrozumieć). Wtedy pytanie o to, które domknięte podprzestrzenie są dopełnialne (mają domknięte dopełnienie) jest dużo ciekawsze.

- Jeżeli $W = \text{Lin}(v)$ dla $v = \begin{pmatrix} x \\ y \end{pmatrix}$ przy $y \neq 0$, to $F_A(v) = \begin{pmatrix} 2x+y \\ 2y \end{pmatrix}$ nie jest równoległy do v , więc W nie jest niezmiennicza, więc powyższe 3 to wszystkie.
- 2. Jeżeli $A = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \in M_{2 \times 2}(\mathbf{R})$, $\mu \neq \lambda$, to przestrzeń F_A -niezmiennicza to $\{0\}, \mathbf{R}^2$ oraz osie: jeżeli $v = \begin{pmatrix} x \\ y \end{pmatrix}$, gdzie $x \neq 0 \neq y$, to v nie jest współliniowy z $F(v) = \begin{pmatrix} \lambda x \\ \mu y \end{pmatrix}$, więc $\text{Lin}(v)$ nie jest F_A -niezmiennicza.
- 3. Jeżeli F, k są dowolne, to $\text{im } F^k$ jest zawsze F -niezmiennicza, bo $F(F^k(v)) = F^k(F(v))$.
- 4. Podobnie $\ker F^k$ jest zawsze F -niezmiennicza: jeżeli $v \in \ker F^k$, to $F^k(F(v)) = F(F^k(v)) = F(0) = 0$.

Fakt 7.18. Jeżeli $W \leq V$ jest F -niezmiennicza, to mamy dobrze określone odwzorowanie liniowe $\bar{F}: V/W \rightarrow V/W$ zadane wzorem $\bar{F}(v+W) = F(v)+W$.

Dowód. Jeżeli $v_1 + W = v_2 + W$, to $v_2 = v_1 + w$ dla pewnego $w \in W$, czyli $F(v_2) + W = F(v_1) + F(w) + W = F(v_1) + W$. Stąd \bar{F} jest dobrze określone. Liniowość jest oczywista. \square

Lemat 7.19. Niech $F \in \text{End}(V)$.

1. Jeżeli $V = V_1 \oplus V_2 \oplus \dots \oplus V_n$, gdzie V_i są F -niezmiennicze i B_1, \dots, B_n są bazami odpowiednich V_i , to $B = B_1 B_2 \dots B_n$ jest bazą V i

$$m_B(F) = \left(\begin{array}{c|c|c|c} m_{B_1}(F|_{V_1}) & 0 & 0 & 0 \\ \hline 0 & m_{B_2}(F|_{V_2}) & 0 & 0 \\ \hline 0 & 0 & \ddots & \vdots \\ \hline 0 & 0 & \cdots & m_{B_n}(F|_{V_n}) \end{array} \right)$$

2. Niech $W \leq V$ będzie F -niezmiennicza, niech $B_1 = (b_1, \dots, b_k)$ będzie bazą W , a $B = (b_1, b_2, \dots, b_n)$ dowolnym rozszerzeniem B_1 do bazy V .

Położmy $B_2 = (b_{k+1} + W, \dots, b_n + W)$.

Wtedy B_2 jest bazą V/W oraz

$$m_B(F) = \left(\begin{array}{c|c} m_{B_1}(F|_W) & * \\ \hline 0 & m_{B_2}(\bar{F}) \end{array} \right).$$

Dowód. 1. B jest bazą, dla $n = 2$: jeśli mamy $0 = \sum_{k=1}^{m_1} \beta_{k,1} b_{k,1} + \sum_{k=1}^{m_2} \beta_{k,2} b_{k,2}$, to z liniowej niezależności $\sum_{k=1}^{m_1} \beta_{k,1} b_{k,1} = 0$, więc wszystkie $\beta_{k,1}$ są zerowe (z $\ln z B_1$) i podobnie dla drugiej sumy. Dla $n > 2$ dowód jest analogiczny.

Ustalmy i . Z niezmienniczości V_i , dla dowolnego $b_0 \in B_i$ mamy $F(b_0) \in V_i$, czyli

$$F(b_0) = \sum_{b \in B_i} \beta_b b = \sum_{b \in B} \gamma_b b,$$

gdzie $\gamma_b = 0$ dla $b \notin B_i$ i $\gamma_b = \beta_b$ dla $b \in B_i$. Wobec tego kolumna odpowiadająca b_0 po lewej i prawej stronie jest taka sama. Z dowolności b_0, i wynika więc teza.

2. B_2 jest bazą: $\dim V/W = n - k = |B_2|$, wystarczy więc pokazać, że B_2 rozpiną V/W . Istotnie, jeżeli $v = \sum_{j=1}^n \beta_j b_j$ jest dowolny, to

$$v + W = \left(\sum_{j=1}^n \beta_j b_j \right) + W = \sum_{j=1}^n (\beta_j b_j + W),$$

i wyrazy z $j \leq k$ po prawej stronie są równe W , co daje tezę.

Rozumowanie jak wyżej pokazuje, że kolumny odpowiadające wektorom z B_1 po lewej i prawej stronie się zgadzają. Dla kolumn odpowiadających wektorom $B \setminus B_1$: jeżeli weźmiemy pewne b_j , $j > k$, to $F(b_j) = \sum_{i=1}^n \beta_i b_i$. Stąd

$$[\bar{F}(b_j + W)]_{B_2} = \left[\underbrace{\left(\sum_{i=1}^k \beta_i b_i + W \right)}_{\parallel W} + \sum_{i=k+1}^m \beta_i b_i + W \right]_{B_2} = \begin{pmatrix} \beta_{k+1} \\ \beta_{k+2} \\ \vdots \\ \beta_n \end{pmatrix},$$

a więc

$$[F(b_j)]_B = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \\ \beta_{k+1} \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} * \\ [\bar{F}(b_j + W)]_{B_2} \end{pmatrix},$$

więc pozostałe kolumny też się zgadzają. □

Ćwiczenie 7.20. Jeżeli A, B są macierzami kwadratowymi, to

$$\left| \begin{array}{c|c} A & * \\ \hline 0 & B \end{array} \right| = \det A \cdot \det B,$$

skąd łatwo indukcyjnie wynika ogólny wzór

$$\left| \begin{array}{c|c|c|c} A_1 & * & * & * \\ \hline 0 & A_2 & * & * \\ \hline 0 & 0 & \ddots & \vdots \\ \hline 0 & 0 & \cdots & A_n \end{array} \right| = \prod_{k=1}^n \det(A_k)$$

Wniosek 7.21. W sytuacji jak w lemacie mamy:

1. $\chi_F(x) = \chi_{F|_{V_1}}(x) \cdot \chi_{F|_{V_2}}(x) \cdots \chi_{F|_{V_n}}(x)$
2. $\chi_F(x) = \chi_{F|_W}(x) \cdot \chi_{\bar{F}}(x).$

Przykład 7.22. Dla $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$, $W = OX$ mamy $m_{e_1}(F_A|_W) = (2)$ i $m_{e_2+W}(\bar{F}) = (2)$, więc $\chi_A(x) = (2-x) \cdot (2-x).$

7.4 Przestrzenie własne

Definicja 7.23. Niech F będzie endomorfizmem V , a λ będzie skalar. Wtedy *przestrzeń własna F dla λ* to

$$V_\lambda = V_\lambda(F) = \{v \in V \mid F(v) = \lambda v\}.$$

Jeżeli $V_\lambda \neq \{0\}$, to λ nazywamy *wartością własną F* , a elementy V_λ nazywamy *wektorami własnymi F* (odpowiadającymi lub dla λ).

Zbiór $\sigma(F) = \sigma_p(F) = \text{Spec}(F) := \{\lambda \in K \mid V_\lambda \neq \{0\}\}$ nazywamy *spektrum [punktowym] (lub widmem) F* .²

Przykład 7.24. Niech $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$. Wtedy $\sigma(F_A) = \{1, 2, 3\}$ i V_1, V_2, V_3 to osie układu współrzędnych.

²Dla nieskończenie wymiarowych przestrzeni do spektrum (nie punktowego) włącza się czasami też inne skalary.

Niech $V = \ell^2 = \{(a_n)_{n \in \mathbb{N}} \mid \sum_n |a_n|^2 < \infty\}$, a $F: \ell^2 \rightarrow \ell^2$ będzie prawym sziftem (dopisaniem zera na początku). Wtedy F nie ma wartości własnych, nawet zespolonych: $\sigma_p(F) = \emptyset$.³

Definicja 7.25. Oznaczenie: dla dowolnego $F \in \text{End}(V)$ i $\lambda \in K$ będziemy oznaczać $F - \lambda := F - \lambda \text{id}_V$. Podobnie definiujemy $F + \lambda$ oraz $A \pm \lambda$ dla macierzy kwadratowej A .

Uwaga 7.26. Dla F, λ jak wyżej, jeżeli $\dim V < \infty$ i B jest dowolną bazą V , to mamy $m_B(F \pm \lambda) = m_B(F) \pm \lambda$.

Uwaga 7.27. • Każda V_λ jest niezmienniczą podprzestrzenią V .

- $\lambda \in \sigma(F) \iff \chi_F(\lambda) = 0$ (wartości własne F to dokładnie pierwiastki wielomianu charakterystycznego).

Dowód. Pierwszy punkt: $V_\lambda = \ker(F - \lambda)$, więc jest podprzestrzenią. Niezmienniczość wynika łatwo z definicji.

Drugi punkt:

$$\begin{aligned} \lambda \in \sigma(F) &\iff \ker(F - \lambda) \neq \{0\} \\ &\iff F - \lambda \text{ nie jest odwracalne} \\ &\iff \det(F - \lambda) = 0 \\ &\quad \parallel \\ &\quad \chi_F(\lambda) \end{aligned} \quad \square$$

Wniosek 7.28. $F: V \rightarrow V$ ma co najwyżej $\dim V$ wartości własnych.

Dowód. $\chi_F(x)$ jest wielomianem stopnia $\dim V$, więc ma co najwyżej n pierwiastków (z twierdzenia Bézout). \square

Fakt 7.29. Załóżmy że $\lambda_1, \lambda_2, \dots, \lambda_n \in K$ są parami różne. Wtedy przestrzenie $V_{\lambda_1}, \dots, V_{\lambda_n}$ są liniowo niezależne.

Dowód. Indukcja względem n . Dla $n = 0, 1$ teza jest oczywista. Załóżmy że $n > 1$ i teza zachodzi dla mniejszej liczby λ_i .

Weźmy $v_i \in V_{\lambda_i}$ takie że $\sum_{i=1}^n v_i = 0$. Wymnażając przez λ_1 dostajemy $\sum_{i=1}^n \lambda_1 v_i = 0$. Nakładając F dostajemy $\sum_{i=1}^n \lambda_i v_i = 0$. Po odjęciu stronami otrzymujemy

$$\sum_{i=1}^n (\lambda_i - \lambda_1) v_i = \sum_{i=2}^n (\lambda_i - \lambda_1) v_i = 0 - 0 = 0.$$

³Ale na analizie funkcjonalnej można się dowiedzieć, że $\sigma(F) = [-1, 1]$.

Ponieważ $V_{\lambda_2}, \dots, V_{\lambda_n}$ są lnz z założenia indukcyjnego, wnioskujemy stąd że wszystkie $(\lambda_i - \lambda_1)v_i = 0$. Ale ponieważ $\lambda_i \neq \lambda_1$ dla $i > 1$, wynika stąd że $v_i = 0$ dla $i > 1$. Stąd $0 = \sum_{i=1}^n v_i = v_1$. \square

alternatywny dowód. Załóżmy że $0 = \sum_{i=1}^n v_i$, gdzie $v_i \in V_{\lambda_i}$. Wtedy

$$\begin{cases} 0 = v_1 + v_2 + \dots + v_n \\ F(0) = 0 = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n \\ F^2(0) = 0 = \lambda_1^2 v_1 + \lambda_2^2 v_2 + \dots + \lambda_n^2 v_n \\ \vdots \\ F^{n-1}(0) = 0 = \lambda_1^{n-1} v_1 + \lambda_2^{n-1} v_2 + \dots + \lambda_n^{n-1} v_n \end{cases}$$

Z powyższego wynika

$$\begin{pmatrix} I & I & \dots & I \\ \lambda_1 \cdot I & \lambda_2 \cdot I & \dots & \lambda_n \cdot I \\ \vdots & & \ddots & \vdots \\ \lambda_1^{n-1} I & \lambda_2^{n-1} I & \dots & \lambda_n^{n-1} I \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = 0,$$

a wyznacznik dużej macierzy po lewej to wyznacznik Vandermonde'a podniesiony do potęgi $\dim V$, czyli $v_i = 0$. \square

alternatywny alternatywny dowód. Zaczynając jak w poprzednim sposobie,

kolumny $w_i = \begin{pmatrix} 1 \\ \lambda_i \\ \vdots \\ \lambda_i^{n-1} \end{pmatrix} \in K^n$ są liniowo niezależne (wyznacznik Vander-

monde'a), więc są bazą K^n . Istnieje zatem baza dualna $\varphi_1, \dots, \varphi_n$, czyli

$\varphi_i(w_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$. Niech $\varphi_i \begin{pmatrix} x_0 \\ \vdots \\ x_{n-1} \end{pmatrix} = \sum_{k=0}^{n-1} a_{ik} x_k$. Wtedy dla każdego i

mamy

$$\vec{0} = \sum_{k=0}^{n-1} a_{ik} \underbrace{\sum_{j=1}^n \lambda_j^k v_j}_{\parallel 0} = \sum_{j=1}^n \underbrace{\left(\sum_{k=0}^{n-1} (a_{ik} \lambda_j^k) \right)}_{\parallel \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}} v_j = v_i. \quad \square$$

7.5 Diagonalizowalność

Definicja 7.30. Mówimy że $F: V \rightarrow V$ jest *diagonalizowalne* jeżeli $V = \bigoplus_{\lambda \in \sigma(F)} V_\lambda$.

(Wystarczy sprawdzić, że $V = \sum_{\lambda \in \sigma(F)} V_\lambda$, lub równoważnie, $\sum_{\lambda \in \sigma(F)} \dim V_\lambda = \dim V$.)

Stwierdzenie 7.31. Następujące warunki są równoważne dla $F: V \rightarrow V$:

- F jest diagonalizowalne,
- istnieje baza B przestrzeni V taka że $m_B(F)$ jest diagonalna (wtedy mówimy że F diagonalizuje się w bazie B lub baza B diagonalizuje F),
- istnieje baza B przestrzeni V złożona z wektorów własnych dla F .

Dowód. Dwa ostatnie punkty są równoważne z definicji $m_B(F)$: $m_B(F)$ jest diagonalna \iff (dla każdego i) w i -tej kolumnie tylko i -ty wyraz może być niezerowy $\iff F(b_i) = \alpha_i b_i$ dla pewnego $\alpha_i \iff b_i$ jest wektorem własnym.

Założmy że F jest diagonalizowalne. Zauważmy że $F|_{V_\lambda} = \lambda \cdot \text{id}_{V_\lambda}$ ma w każdej bazie V_λ macierz $\lambda \cdot \text{I}_{\dim V_\lambda}$. Jeżeli wartości własne to $\lambda_1, \dots, \lambda_k$, to biorąc (dowolne) bazy B_1, \dots, B_k dla $V_{\lambda_1}, \dots, V_{\lambda_k}$ widzimy z Lematu 7.19, że $B = B_1 B_2 \dots B_k$ jest bazą V i $m_B(F)$ spełnia

$$m_B(F) = \left(\begin{array}{c|c|c|c} m_{B_1}(F|_{V_{\lambda_1}}) & 0 & 0 & 0 \\ \hline 0 & m_{B_2}(F|_{V_{\lambda_2}}) & 0 & 0 \\ \hline 0 & 0 & \ddots & \vdots \\ \hline 0 & 0 & \cdots & m_{B_k}(F|_{V_{\lambda_k}}) \end{array} \right)$$

$$= \left(\begin{array}{c|c|c|c} \lambda_1 \text{I}_{\dim V_{\lambda_1}} & 0 & 0 & 0 \\ \hline 0 & \lambda_2 \text{I}_{\dim V_{\lambda_2}} & 0 & 0 \\ \hline 0 & 0 & \ddots & \vdots \\ \hline 0 & 0 & \cdots & \lambda_k \text{I}_{\dim V_{\lambda_k}} \end{array} \right)$$

W drugą stronę, założmy że V ma bazę B wektorów własnych dla F . Wtedy

$$\sum_{\lambda \in \sigma(F)} V_\lambda \supseteq \underbrace{\sum_{\lambda \in \sigma(F)} \text{Lin} \{ b \in B \mid \text{wartość własna } b \text{ to } \lambda \}}_{B_\lambda} = \text{Lin} \left(\bigcup_{\lambda \in \sigma(F)} B_\lambda \right) = \text{Lin}(B) = V,$$

więc F jest diagonalizowalna. \square

Uwaga 7.32. • Jeżeli $m_B(F)$ jest diagonalna, to dla dowolnej bazy C istnieje odwracalna macierz $P(=m_C^B(\text{id}))$ taka że $m_C(F) = P m_B(F) P^{-1}$.
 \parallel
 D

- Mówimy że macierz $A \in M_{n \times n}(K)$ jest *diagonalizowalna*, jeżeli F_A jest diagonalizowalne, lub równoważnie, jeżeli istnieje taka odwracalna P że PAP^{-1} jest diagonalna (wtedy mówimy że P *diagonalizuje* A).
- Podobnie mówimy że baza B *diagonalizuje* przekształcenie F jeżeli $m_B(F)$ jest diagonalna.

Przykład 7.33. $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ $\chi_A(x) = (2-x)^2(3-x)^2$, $\sigma(F_A) = \{2, 3\}$,

$$V_3 = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in K^4 \mid \begin{cases} 2x + y = 2x \\ 2y = 2y \\ 3z = 2z \\ 3t = 2t \end{cases} \right\} = \text{Lin} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$V_2 = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in K^4 \mid \begin{cases} 2x + y = 3x \\ 2y = 3y \\ 3z = 3z \\ 3t = 3t \end{cases} \right\} = \text{Lin} \left(\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right)$$

$\dim V_2 + \dim V_3 = 1 + 2 < 4 = \dim V$, więc A nie jest diagonalizowalna.

7.6 Twierdzenie Jordana

Definicja 7.34. *Klatkę Jordana dla λ rozmiaru m nazywamy następującą macierz kwadratową:*

$$J_{\lambda,m} := \underbrace{\begin{pmatrix} \lambda & 1 & & & 0 \\ & \lambda & 1 & & \\ & & \lambda & 1 & \\ & & & \ddots & \ddots \\ 0 & & & & \lambda \end{pmatrix}}_{m \text{ kolumn}}$$

Uwaga 7.35. Jeżeli macierz F w bazie $B = (b_1, b_2, \dots, b_n)$ jest równa $J_{0,n}$, to znaczy że $F(b_1) = 0$ i $F(b_j) = b_{j-1}$ dla $j > 1$. Podobnie, jeżeli $m_B(F) = J_{\lambda,n}$, to znaczy że $F(b_1) = \lambda b_1$ i $F(b_j) = b_{j-1} + \lambda b_j$ dla $j > 1$.

Uwaga 7.36. Łatwo indukcyjnie udowodnić, że dla $\lambda = 0$ mamy $J_{0,m}^m = 0$ i $J_{0,m}^{m-1} \neq 0$, a dla $\lambda \neq 0$ macierz $J_{\lambda,m}$ jest odwracalna (o wyznaczniku λ^m).

Ćwiczenie 7.37. Wyznaczyć jawne wzory na $J_{\lambda,m}^k$ (wskazówka: można to zrobić najpierw dla $\lambda = 0$, a następnie skorzystać z wzoru dwumianowego na $J_{\lambda,m}^k = (J_{0,m} + \lambda)^k$).

Cel: udowodnić następujące twierdzenie Jordana.

Twierdzenie 7.38 (Rozkład Jordana). *Jeżeli V jest skończenie wymiarową przestrzenią liniową nad \mathbb{C} i $F \in \text{End}(V)$, to istnieje baza B dla V taka że*

$$m_B(F) = \begin{pmatrix} J_{\lambda_1, m_1} & & & 0 \\ & J_{\lambda_2, m_2} & & \\ & & J_{\lambda_3, m_3} & \\ & & & \ddots \\ 0 & & & & J_{\lambda_k, m_k} \end{pmatrix}$$

gdzie $\lambda_1, \dots, \lambda_k$ są wartościami własnymi F (być może z powtórzeniami).

Definicja 7.39. Bazę taką jak w Twierdzeniu (również nad ciałami innymi niż \mathbb{C}) nazywamy *bazą jordanowską* dla F , a postać macierzy, którą otrzymujemy (tzn. blokową, z klatkami Jordana na przekątnej i zerami poza przekątną), nazywamy *postacią Jordana*.

Analogicznie definiujemy postać Jordana macierzy kwadratowej.

Przykład 7.40. Szukamy postaci Jordana $A = \begin{pmatrix} 3 & -1 & 2 \\ 2 & 0 & 3 \\ -2 & 1 & 0 \end{pmatrix}$. Łatwo policzyć, że $\chi_A(x) = (1-x)^3$. Z drugiej strony $(A-I)^2 \neq 0$, więc postać Jordana A to musi być $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ (w pozostałych przypadkach $(A-I)^2 = 0$).

Aby znaleźć bazę Jordana, musimy znaleźć wektor v taki że $(A-I)^2 v \neq 0$. Sprawdźmy, czy $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ działa (uwaga: $\ker(A-I)^2$ jest płaszczyzną, więc wystarczy że strzelimy wektor spoza tej płaszczyzny, więc jest spora szansa, że losowo wybrany wektor zadziała).

$$(A-I)^2 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 & -1 & 2 \\ 2 & -1 & 3 \\ -2 & 1 & -1 \end{pmatrix}^2 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 & -1 & 2 \\ 2 & -1 & 3 \\ -2 & 1 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \\ -2 \end{pmatrix} = \begin{pmatrix} -2 \\ -4 \\ 0 \end{pmatrix} \neq \vec{0}.$$

Udało się! Stąd baza jordanowska to np. $\begin{pmatrix} -2 \\ -4 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

Uwaga 7.41. Ogólnie jeżeli wiemy, że macierz ma tylko jedną klatkę w postaci Jordana, to łatwo zgadnąć jej bazę jordanowską postępując jak wyżej.

Uwaga 7.42. • Macierz w postaci Jordana jest macierzą górnątrójkątną, stąd wartości na przekątnej to dokładnie wartości własne i liczba wystąpień to dokładnie krotność pierwiastka wielomianu charakterystycznego.

- Innymi słowy, krotność pierwiastka wielomianu charakterystycznego = suma rozmiarów klatek Jordana odpowiadających danej wartości własnej.

- Patrząc się na postać Jordana macierzy przekształcenia F widzimy, że jądro F jest rozpinane przez wektory bazowe odpowiadające „początkom” klatek Jordana odpowiadającym wartości własnej 0 (bo obrazy pozostałych wektorów z bazy Jordana są liniowo niezależne — to wynika z odwracalności klatek Jordana). Stąd $\dim \ker F =$ liczba klatek Jordana z 0 na przekątnej.
- Podobnie, dla dowolnego skalaru λ zachodzi $\dim \ker(F - \lambda) =$ liczba klatek Jordana z λ na przekątnej.
- Oznaczając przez $j_{\geq k}(\lambda)$ liczbę klatek rozmiaru co najmniej k z λ na przekątnej, to $j_{\geq k}(\lambda) = \dim \ker(F - \lambda)^k - \dim \ker(F - \lambda)^{k-1}$. Stąd łatwo wyliczyć liczbę klatek rozmiaru k . (Szczegóły na ćwiczeniach.)

Przykład 7.43. Załóżmy że F ma wartość własną λ i $\dim \ker(F - \lambda)^k = 5, 7, 8, 8$ dla $k = 1, 2, 3, 4$. Stąd wnioskujemy że w części postaci Jordana odpowiadającej λ jest 5 klatek, z czego $2 = 7 - 5$ rozmiaru ≥ 2 i $1 = 8 - 7$ rozmiaru ≥ 3 i $0 = 8 - 8$ klatek rozmiaru ≥ 4 , czyli są 3 klatki 1×1 i po jednej 2×2 i 3×3 . Stąd odpowiedni fragment macierzy F w postaci Jordana ma postać

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

Od teraz do końca dowodu twierdzenia Jordana V jest przestrzenią liniową skończonego wymiaru nad K , a $F \in \text{End}(V)$.

Definicja 7.44. Dla $\lambda \in K$ przestrzeń pierwiastkową dla λ definiujemy jako

$$V^\lambda := \{v \in V \mid \exists k(F - \lambda)^k(v) = 0\}.$$

Mówimy że elementy V^λ są uogólnionymi wektorami własnymi dla λ .

Uwaga 7.45. Nietrudno zobaczyć, że w bazie jordanowskiej wektory odpowiadające klatkom z λ na przekątnej stanowią bazę przestrzeni pierwiastkowej V^λ .

Uwaga 7.46. Zauważmy że V^λ dla $F = V^0$ dla $G = F - \lambda$.

Ćwiczenie 7.47. Jeżeli $\lambda, \mu \in K$, to potęgi $(F - \lambda)$ i $(F - \mu)$ są przemienne, tzn. dla każdych n, m zachodzi $(F - \lambda)^n (F - \mu)^m = (F - \mu)^m (F - \lambda)^n$. (Wskazówka: dla $m = n = 1$ teza jest łatwa do sprawdzenia, dla większych m, n możemy rozumować indukcyjnie.)

Ćwiczenie 7.48. Jeżeli F, G są przemienne, to zachowują wzajemnie swoje przestrzenie własne i pierwiastkowe.

Wniosek 7.49. Przestrzenie V^λ są F -niezmiennicze.

Dowód. Ponieważ F jest przemienne z F , wynika to wprost z ćwiczenia. \square

Ćwiczenie 7.50. Jeżeli $F^k(v) \neq F^{k+1}(v) = 0$, to układ $v, F(v), \dots, F^k(v)$ jest liniowo niezależny.

Wniosek 7.51. Niech $v \in V$. Wtedy $V^\lambda = \ker(F - \lambda)^{\dim V} = \ker(F - \lambda)^{\dim V^\lambda}$.

Dowód. Obydwie \supseteq są jasne. Wystarczy pokazać, że $V^\lambda \subseteq \ker(F - \lambda)^{\dim V^\lambda}$.

Weźmy dowolny $v \in V^\lambda$. Wtedy dla pewnego d zachodzi $(F - \lambda)^d(v) = 0$. Niech d będzie najmniejsze możliwe. Wtedy z ćwiczenia układ $v, (F - \lambda)(v), \dots, (F - \lambda)^{d-1}(v)$ jest liniowo niezależny, zawiera się w V^λ i składa się z d wektorów, więc $d \leq \dim V^\lambda$. \square

Uwaga 7.52. Z twierdzenia Jordana można wywnioskować, że V^λ jest równe $\ker(F - \lambda)^k$, gdzie k to maksymalny rozmiar klatki Jordana odpowiadającej λ .

Fakt 7.53 (Rozkład Fittinga). $V = V^0 \oplus V'$, gdzie $V' = \operatorname{im} F^{\dim V^0}$; ponadto $F|_{V'}$ jest odwracalne.

Dowód. Oznaczmy $k = \dim V^0$. Z poprzedniego wniosku wiemy, że $V^0 = \ker F^k$. Weźmy dowolny $v \in \ker F^k \cap \operatorname{im} F^k$ oraz $w \in V$ taki że $F^k(w) = v$. Wtedy z poprzedniego wniosku $F^{2k}(w) = F^k(v) = 0$, więc $w \in V^0$, a więc (znów z poprzedniego wniosku) $v = F^k(w) = 0$. Stąd $V^0 = \ker F^k$ i $\operatorname{im} F^k$ są liniowo niezależne. Wobec tego tym bardziej $\ker F \cap \operatorname{im} F^k = \{0\}$, a więc $F|_{\operatorname{im} F^k}$ jest 1-1.

Ponieważ $\operatorname{im} F^k$ oczywiście jest F -niezmiennicze (i skończenie wymiarowe), wynika stąd że $F|_{\operatorname{im} F^k}$ jest też „na” $\operatorname{im} F^k$, jest więc odwracalne (jako element $\operatorname{End}(\operatorname{im} F^k)$).

Z drugiej strony z twierdzenia o rzędzie $\dim \ker F^k + \dim \operatorname{im} F^k = \dim V$, więc stąd $V' = \ker F^k \oplus \operatorname{im} F^k$. \square

Uwaga 7.54. Nietrudno zobaczyć, że w poprzednim fakcie mamy $V' = \text{im } F^l$ dla każdego $l > k$, czyli $V' = \bigcap_j \text{im } F^j$.

Wniosek 7.55. *Jeżeli $\lambda \neq 0$, to $V^\lambda \subseteq \text{im } F^k$ dla każdego k (a więc też $V^\lambda \subseteq V'$, gdzie V' jest jak w rozkładzie Fittinga).*

Dowód. Ponieważ V^λ jest skończenie wymiarowa i F -niezmiennicza, wystarczy pokazać, że $F|_{V^\lambda}$ jest 1-1, czyli $V^\lambda \cap \ker F = \{0\}$.

Weźmy dowolne $v \in \ker F \cap V^\lambda$ i najmniejsze d takie że $(F - \lambda)^d(v) = 0$. Jeżeli $d = 0$, to $v = 0$ i teza jest spełniona. W przeciwnym razie mamy

$$\begin{aligned} 0 &= (F - \lambda)^d(v) = (F - \lambda)^{d-1}(F - \lambda)(v) \\ &= (F - \lambda)^{d-1}(F(v) - \lambda v) = (F - \lambda)^{d-1}(0 - \lambda v) = -\lambda(F - \lambda)^{d-1}(v), \end{aligned}$$

czyli (ponieważ $\lambda \neq 0$) $(F - \lambda)^{d-1}(v) = 0$, przecząc minimalności d . \square

Wniosek 7.56. *Przestrzenie pierwiastkowe dla różnych wartości własnych są liniowo niezależne. Dokładniej, jeżeli $\lambda_1, \dots, \lambda_k$ są parami różne, to $V^{\lambda_1}, \dots, V^{\lambda_k}$ są lnz.*

Dowód. Z Lematu 7.6 wystarczy pokazać, że dla każdego j_0 zachodzi $V^{\lambda_{j_0}} \cap \sum_{j \neq j_0} V^{\lambda_j} = \{0\}$. Bez zmniejszania ogólności możemy założyć, że $j_0 = 1$. Ponadto zastępując F przez $G = F - \lambda_1$ możemy założyć, że $\lambda_1 = 0$.

Wówczas chcemy pokazać, że $V^0 \cap \sum_{j>1} V^{\lambda_j} = \{0\}$ (przy czym $\lambda_j \neq 0$ dla $j > 1$). Ale biorąc rozkład Fittinga $V = V^0 \oplus V'$ wiemy, że każda $V^{\lambda_j} \subseteq V'$, a więc też $\sum_{j>1} V^{\lambda_j} \subseteq V'$, co daje tezę. \square

Stwierdzenie 7.57. *Ustalmy $F \in \text{End}(V)$, $\dim V < \infty$. Wtedy $\dim V^\lambda =$ krotkość λ jako pierwiastka $\chi_F(x)$ (czyli największa potęga $(x - \lambda)$ dzieląca $\chi_F(x)$).*

Dowód. Rozważmy przypadek $\lambda = 0$. Weźmy rozkład Fittinga $V = V^0 \oplus V'$ dla F . Ponieważ $F|_{V'}$ jest odwracalne, 0 nie jest jego wartością własną, a więc nie jest też pierwiastkiem jego wielomianu charakterystycznego.

Z drugiej strony $F|_{V^0}$ nie ma niezerowych wartości własnych: gdyby v było wektorem własnym dla niezerowej λ , to $F^k(v) = \lambda^k v \neq 0$. Stąd jego wielomian charakterystyczny to $(-x)^{\dim V^0}$.

Ponieważ $V = V^0 \oplus V'$ i obydwie przestrzenie są niezmiennicze, z Lematu 7.21 wiemy, że $\chi_F = \chi_{F|_{V^0}} \cdot \chi_{F|_{V'}}$, przy czym $\chi_{F|_{V'}}(0) \neq 0$, skąd łatwo wynika teza.

Dla $\lambda \neq 0$ wystarczy rozważyć $G = F - \lambda$. Wtedy z rozumowania powyżej widzimy, że krotność 0 jako pierwiastka $\chi_G(x) = \chi_F(x + \lambda)$ to $\dim V^\lambda$, tzn. $\chi_F(x + \lambda) = (-x)^{\dim V^\lambda} P(x)$, gdzie $P(0) \neq 0$. Podstawiając $y = x + \lambda$ mamy $\chi_F(y) = (\lambda - y)^{\dim V^\lambda} P(y - \lambda)$, skąd wynika teza. \square

Definicja 7.58. Mówimy że niestały wielomian $P \in K[x]$ rozszczepia się nad K , jeżeli jest iloczynem jednomianów z $K[x]$, tzn. istnieją $a, x_1, \dots, x_n \in K$ takie że

$$P(x) = a(x - x_1)(x - x_2)(x - x_3) \cdots (x - x_n).$$

Wniosek 7.59. Jeżeli wielomian charakterystyczny $F \in \text{End}(V)$ rozszczepia się nad K , to

$$V = \bigoplus V^\lambda,$$

gdzie λ przebiega różne pierwiastki $\chi_F(x)$.

Dowód. Niech $\lambda_1, \dots, \lambda_k$ będą pierwiastkami $\chi_F(x)$. Z Wniosku 7.56 wiemy, że $V^{\lambda_1}, \dots, V^{\lambda_k}$ są liniowo niezależne. Z drugiej strony z założenia i ze stwierdzenia mamy:

$$\dim V = \deg \chi_F(x) = \deg \left(\prod_{j=1}^k (x - \lambda_j)^{\dim V^{\lambda_j}} \right) = \sum_{j=1}^k \dim V^{\lambda_j}$$

Stąd wynika teza. \square

Definicja 7.60. Endomorfizm $F \in \text{End}(V)$ nazywamy *nilpotentnym* (stopnia d) jeżeli $F^d = 0$ (przy $\dim V < \infty$, F jest nilpotentny gdy $V = V^0$).⁵

Definicja 7.61. Mówimy że podprzestrzeń $W \leq V$ jest *cykliczna* gdy jest postaci $K[F] \cdot v := \text{Lin}(v, F(v), F^2(v), \dots)$. (Taka podprzestrzeń zawsze jest F -niezmiennicza.)

Uwaga 7.62. Weźmy nilpotentny F , dowolne $v \in V$ i najmniejsze d takie że $F^d(v) = 0$.

Wtedy $B = (F^{d-1}(v), \dots, F(v), v)$ jest bazą $W = K[F] \cdot v$ i $m_B(F|_W) = J_{0,d}$. W szczególności istnieje baza, w której macierz $F|_W$ jest tej postaci.

Lemat 7.63. Załóżmy że F jest nilpotentny, $d \in \mathbb{N}_+$ jest minimalne takie że $F^d = 0$, zaś $v_0 \in V \setminus \ker F^{d-1}$. Wtedy istnieje niezmiennicza podprzestrzeń $V' \leq V$ dopełnicza do $K[F] \cdot v_0$.

Dowód. Indukcja ze względu na d . Jeżeli $d = 1$, to teza jest oczywista: $F = 0$, więc każda podprzestrzeń jest niezmiennicza, w szczególności dowolna V' dopełnicza do $K[F] \cdot v = \text{Lin}(v_0)$ spełnia założenia.

Założmy że $d > 1$ i weźmy takie v_0 . Oznaczmy $W = \text{im } F$, $G = F|_W$. W jest F -niezmiennicza, więc $G \in \text{End}(W)$. Ponadto $G^{d-1} = 0$ i $w_0 = F(v) \in W$ spełnia $G^{d-2}(w_0) \neq 0$. Stąd z założenia indukcyjnego $W = K[G] \cdot w_0 \oplus W'$ dla pewnej G -niezmienniczej $W' \leq W$.

Weźmy $V'' = F^{-1}[W']$. Wtedy $V'' \supseteq W'$: biorąc dowolne $w \in W'$ mamy (z G -niezmienniczości W') $F(w) = G(w) \in W'$, więc $w \in F^{-1}[W']$.

Zauważmy że $K[G] \cdot w_0 = F[K[F] \cdot v_0] = (K[F] \cdot v_0) \cap \text{im } F$: pierwsza równość jest oczywista, w drugiej \subseteq jest oczywiste (z niezmienniczości $K[F] \cdot v_0$), \supseteq zaś wynika z obserwacji, że wymiar prawej strony jest ostro mniejszy niż $d = \dim K[F] \cdot v_0$ (bo v_0 do niej nie należy), a wymiar lewej strony to $d - 1$ (bo $w_0, F(w_0), \dots, F^{d-2}(w_0)$ jest bazą).

Stąd $(K[F] \cdot v_0) \cap W' = \{0\}$: jeżeli v jest wektorem w przekroju, to w szczególności $v \in \text{im } F$, czyli $v \in K[G] \cdot w_0 \cap W' = \{0\}$. Stąd $K[F] \cdot v_0, W'$ są liniowo niezależne.

Ponadto ponieważ $\text{im } F = K[G] \cdot w_0 \oplus W'$, to dla każdego $v \in V$ mamy $F(v) = w_1 + w_2$, gdzie $w_1 \in K[G] \cdot w_0$ i $w_2 \in W'$. Biorąc $v_1 \in K[F] \cdot v_0$ taki że $F(v_1) = w_1$ widzimy, że $F(v - v_1) = w_2 \in W'$, więc $v - v_1 \in V''$. Stąd $v \in K[F] \cdot v_0 + V''$, więc $V = K[F] \cdot v_0 + V''$.

W związku z tym założenia Lematu 7.11 są spełnione i możemy rozszerzyć W' do $V' \subseteq V''$ dopełniczej do $K[F] \cdot v_0$. Ponadto V' jest F -niezmiennicza, bo $F[V'] \subseteq F[V''] \subseteq W' \subseteq V'$. \square

Wniosek 7.64. *Założmy że F jest nilpotentny. Wtedy V jest sumą prostą cyklicznych podprzestrzeni.*

Dowód. Indukcja ze względu na $\dim V$. Jeżeli $\dim V = 0$, to teza jest oczywista. W przeciwnym razie z poprzedniego lematu mamy $V = V_1 \oplus V_2$, gdzie $V_1 \neq \{0\}$ jest cykliczna, a V_2 jest F -niezmiennicza. Wtedy $\dim V_2 < \dim V$, więc z założenia indukcyjnego V_2 jest sumą prostą cyklicznych podprzestrzeni. Biorąc te przestrzenie wraz z V_1 otrzymujemy rozkład V na cykliczne podprzestrzenie. \square

Wniosek 7.65. *F jest nilpotentny wtedy i tylko wtedy gdy istnieje baza B taka*

że dla pewnych m_j zachodzi:

$$m_B(F) = \begin{pmatrix} J_{0,m_1} & & & 0 \\ & J_{0,m_2} & & \\ & & J_{0,m_3} & \\ & 0 & & \ddots & \\ & & & & J_{0,m_k} \end{pmatrix}$$

Ogólnie, $F - \lambda$ jest nilpotentny (tzn. $V = V^\lambda$) wtedy i tylko wtedy gdy istnieje taka baza B , że dla pewnych m_j zachodzi:

$$m_B(F) = \begin{pmatrix} J_{\lambda,m_1} & & & 0 \\ & J_{\lambda,m_2} & & \\ & & J_{\lambda,m_3} & \\ & 0 & & \ddots & \\ & & & & J_{\lambda,m_k} \end{pmatrix}$$

Dowód. Niech $V = \bigoplus_{j=1}^k V_j$, gdzie V_j są cykliczne. Wtedy z Uwagi dla każdego j mamy możemy wybrać bazę B_j dla V_j taką że $m_{B_j}(F|_{V_j}) = J_{0,\dim V_j}$. Biorąc $B = B_1 B_2 \dots B_k$ dostajemy:

$$m_B(F) = \begin{pmatrix} m_{B_1}(F|_{V_1}) & & & 0 \\ & m_{B_2}(F|_{V_2}) & & \\ & & m_{B_3}(F|_{V_3}) & \\ & 0 & & \ddots & \\ & & & & m_{B_k}(F|_{V_k}) \end{pmatrix}$$

Aby wywnioskować „ogólnie” wystarczy rozważyć $G = F - \lambda$ i zauważyć, że

$$m_B(F) = m_B(G + \lambda) = m_B(G) + m_B(\lambda \text{id}_V) = m_B(G) + \lambda I. \quad \square$$

Twierdzenie 7.66. Niech V będzie przestrzenią liniową skończonego wymiaru nad K , a $F \in \text{End}(V)$. Wtedy następujące warunki są równoważne:

- $\chi_F(x)$ rozszczepia się nad K ,

- istnieje baza jordanowska dla F ,

tzn. istnieje baza B dla V taka że (dla pewnych λ_j, m_j)

$$m_B(F) = \begin{pmatrix} J_{\lambda_1, m_1} & & & 0 \\ & J_{\lambda_2, m_2} & & \\ & & J_{\lambda_3, m_3} & \\ & & & \ddots \\ 0 & & & & J_{\lambda_k, m_k} \end{pmatrix}.$$

Dowód. Macierz jordanowska jest trójkątna, co łatwo implikuje, że wielomian charakterystyczny się rozszczepia (wartości na przekątnej są pierwiastkami wielomianu charakterystycznego). Załóżmy teraz, że $\chi_F(x)$ się rozszczepia.

Niech $\lambda_1, \dots, \lambda_k$ będą różnymi pierwiastkami $\chi_F(x)$. Wtedy $V = \bigoplus_j V^{\lambda_j}$ i dla każdego j , $F|_{V^{\lambda_j}} - \lambda_j$ jest nilpotentny, więc z poprzedniego Wniosku otrzymujemy bazę B_j przestrzeni V^{λ_j} , taką że

$$m_{B_j}(F|_{V^{\lambda_j}}) = \begin{pmatrix} J_{\lambda_j, m_1} & & & 0 \\ & J_{\lambda_j, m_2} & & \\ & & J_{\lambda_j, m_3} & \\ & & & \ddots \\ 0 & & & & J_{\lambda_j, m_{k_j}} \end{pmatrix}$$

Ponieważ $V = \bigoplus_j V^{\lambda_j}$, wnioskujemy stąd, że $B = B_1 B_2 \dots B_k$ jest bazą V i wtedy:

$$m_B(F) = \begin{pmatrix} m_{B_1}(F|_{V^{\lambda_1}}) & & & 0 \\ & m_{B_2}(F|_{V^{\lambda_2}}) & & \\ & & m_{B_3}(F|_{V^{\lambda_3}}) & \\ & & & \ddots \\ 0 & & & & m_{B_k}(F|_{V^{\lambda_k}}) \end{pmatrix},$$

skąd łatwo wynika teza. □

Definicja 7.67. Mówimy że ciało K jest *algebraicznie domknięte* jeżeli każdy niestały wielomian o współczynnikach z K ma pierwiastek w K .

Przykład 7.68. Ciało liczb rzeczywistych (ani żadne podciało, jak np. ciało liczb wymiernych) nie jest algebraicznie domknięte, bo wielomian $x^2 + 1$ nie ma pierwiastków.

Przykład 7.69. Ciało \mathbf{F}_p nie jest algebraicznie domknięte, bo wielomian $x^p - x + 1$ nie ma pierwiastków.⁴

Uwaga 7.70. • Ciało \mathbf{C} jest algebraicznie domknięte (to tzw. zasadnicze twierdzenie algebry⁵, nietrywialne, ma wiele pięknych dowodów, można je poznać np. na wykładzie z funkcji analitycznych, topologii algebraicznej, lub topologii różniczkowej).

- Z twierdzenia Bézout wynika, że K jest algebraicznie domknięte \iff każdy wielomian w $K[x]$ się rozszczepia.
- Każde ciało można rozszerzyć do ciała algebraicznie domkniętego (dowód można zobaczyć np. na wykładzie Algebra 2R).

Wniosek 7.71 (Twierdzenie Jordana). *Jeżeli $K = \mathbf{C}$ (lub ogólniej, jeżeli K jest ciałem algebraicznie domkniętym) i $F \in \text{End}(V)$, to istnieje baza jordanowska dla F .*

Przykład 7.72. Przekształcenie $\mathbf{Q}^2 \rightarrow \mathbf{Q}^2$ zadane macierzą $\begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}$ ma wielomian charakterystyczny $x^2 - 2$, który nie ma pierwiastków w \mathbf{Q} , więc nie ma wymiernych wartości własnych, a więc i wymiernej postaci Jordana. Natomiast przekształcenie $\mathbf{R}^2 \rightarrow \mathbf{R}^2$ zadane tą samą macierzą się diagonalizuje:

$$\begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix} = \begin{pmatrix} \sqrt{2} \\ 2 \end{pmatrix} = \sqrt{2} \begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -\sqrt{2} \end{pmatrix} = \begin{pmatrix} -\sqrt{2} \\ 2 \end{pmatrix} = -\sqrt{2} \begin{pmatrix} 1 \\ -\sqrt{2} \end{pmatrix},$$

Przykład 7.73. Rozważmy przekształcenie $F_A: \mathbf{R}^3 \rightarrow \mathbf{R}^3$ zadane macierzą

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

⁴Ogólnie żadne ciało skończone nie jest algebraicznie domknięte: z zasady szufladkowej można wywnioskować, że w ciele o q elementach wielomian $x^q - x + 1$ nie ma pierwiastków.

⁵wbrew nazwie to twierdzenie nie jest specjalnie zasadnicze, a na pewno nie dla algebry (ani w ogóle nie jest twierdzeniem algebry)

Mamy $\chi_A(x) = (2-x) \cdot \begin{vmatrix} 1-x & 1 \\ -1 & 3-x \end{vmatrix} = (2-x) \cdot (4-4x+x^2) = (2-x)^3$. Mamy $V_2 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \left| \begin{cases} x = 2x \\ -x + y + z = 2y \\ -x - y + 3z = 2z \end{cases} \right. \right\} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \left| z = x + y \right. \right\}$, więc $\dim V_2 = 3 - 1 = 2$. Szukamy bazy jordanowskiej b_1, b_2, b_3 . Dwa z tych wektorów muszą być w $\ker(F-2)$, a trzeci jest przeprowadzany przez $(F-2)$ na jeden z nich. Z tego wynika że $(F-2)^2 = 0$ (bo zeruje każdy wektor z bazy). Szukamy wektora b_3 takiego że $(F-2)b_3 \neq 0$, czyli $b_3 \notin V_2$; wystarczy rozwiązanie $z \neq x + y$, np. $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. Wtedy musimy wziąć $b_2 = (A - 2I) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix}$, zaś $b_1 \in V_2$ musi być lnz z b_2 , na przykład $b_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$.

Przykład 7.74. Załóżmy że mamy macierz 8×8 , o której wiemy, że ma następującą postać Jordana:

$$\begin{pmatrix} \lambda & & & & & & & \\ & \lambda & & & & & & \\ & & \lambda & & & & & \\ & & & \lambda & 1 & & & \\ & & & & \lambda & & & \\ & & & & & \lambda & 1 & \\ & & & & & & \lambda & 1 \\ & & & & & & & \lambda \end{pmatrix}.$$

Szukanie bazy jordanowskiej zaczynamy od ostatniego wektora. Jest to dowolny $b_8 \in \ker(F-\lambda)^3 \setminus \ker(F-\lambda)^2$.

Znalazłszy b_8 łatwo wyznaczamy $b_7 = (F-\lambda)b_8$, $b_6 = (F-\lambda)b_7$. Następnie wyznaczamy $b_5 \in \ker(F-\lambda)^2 \setminus \ker(F-\lambda)$, tak żeby b_5 i $(F-\lambda)b_5$ tworzyły układ lnz z b_6 i b_7 (o b_8 nie musimy się martwić — dlaczego?). Wtedy $b_4 = (F-\lambda)b_5$ jest liniowo niezależny z b_5, b_6, b_7, b_8 i $b_4 \in \ker(F-\lambda)$. Następnie wybieramy $b_1, b_2, b_3 \in \ker(F-\lambda)$, liniowo niezależne z b_4, b_6 .

7.7 Kompleksyfikacja i rzeczywiste tw. Jordana

Od teraz do końca rozdziału 7. V jest skończone wymiarową przestrzenią liniową nad \mathbf{R} , a $F \in \text{End}(V)$.

Definicja 7.75. Niech V będzie rzeczywistą przestrzenią liniową. *Kompleksyfikacją* V nazywamy zespoloną przestrzeń liniową:

$$V_{\mathbf{C}} := V \oplus V = „V \oplus iV”,$$

ze zwykłym dodawaniem oraz z mnożeniem przez skalary zadany wzorem

$$(a+bi) \cdot (v, w) = (a+bi) \cdot (v+iw) = (av-bw) + i(bv+aw) = (av-bw, bv+aw).$$

Intuicyjnie: kompleksyfikacja V to „rozszerzenie skalarów” V z \mathbf{R} do \mathbf{C} .⁶

Przykład 7.76. • Kompleksyfikację $(\mathbf{R}^n)_{\mathbf{C}} = \mathbf{C}^n$.

- $(\mathbf{R}_n[X])_{\mathbf{C}} = \mathbf{C}_n[X]$.
- Dla $V = C(\mathbf{R})_{\mathbf{C}} = C(\mathbf{R}, \mathbf{C})$ (przestrzeń funkcji ciągłych $\mathbf{R} \rightarrow \mathbf{C}$).

Uwaga 7.77. Zgodnie z notacją powyżej, utożsamiamy V z rzeczywistą podprzestrzenią $V_{\mathbf{C}}$ (V nie jest zespoloną podprzestrzenią, bo nie jest zespoloną przestrzenią liniową!).

Mnożenie przez rzeczywiste skalary w $V_{\mathbf{C}}$ działa tak jak powinno (czyli tak jak w $V \oplus V$). W szczególności dla dowolnego $v \in V$ (a nawet $v \in V_{\mathbf{C}}$) i $r \in \mathbf{R}$ mamy $rv = irv$ i nad \mathbf{R} zachodzi istotnie $V_{\mathbf{C}} = V \oplus_{\mathbf{R}} iV$ (to wynika wprost z definicji).

Definicja 7.78. Dla $v = v_1 + iv_2 \in V_{\mathbf{C}}$ ($v_1, v_2 \in V$) definiujemy *sprzężenie* $\bar{v} := v_1 - iv_2$.

(Uwaga: jest to \mathbf{R} -liniowe odwzorowanie.)

Fakt 7.79. Jeżeli $A \subseteq V$, to

$$\text{Lin}_{\mathbf{R}}(A) = V \cap \text{Lin}_{\mathbf{C}}(A).$$

⁶ Za pomocą tzw. iloczynu tensorowego, kompleksyfikację można zdefiniować też jako $V_{\mathbf{C}} := V \otimes_{\mathbf{R}} \mathbf{C}$. W podobny sposób można rozszerzać skalary z \mathbf{Q} do \mathbf{R} , z \mathbf{F}_2 do \mathbf{F}_4 itp. O tym może być więcej na konwersatorium.

Dowód. Inkluzja \subseteq jest oczywista. Pokażemy \supseteq .

Niech $v_j \in A$, $\alpha_j \in \mathbb{C}$ będą takie że $v = \sum_{j=1}^k \alpha_j v_j \in V$. Wtedy $v = \bar{v}$ i $\bar{v}_j = v_j$, więc:

$$v = \frac{1}{2}v + \frac{1}{2}\bar{v} = \frac{1}{2} \sum_{j=1}^k \alpha_j + \frac{1}{2} \sum_{j=1}^k \bar{\alpha}_j v_j = \sum_{j=1}^k \frac{\alpha_j + \bar{\alpha}_j}{2} v_j = \sum_{j=1}^k \operatorname{Re}(\alpha_j) v_j \in \operatorname{Lin}_{\mathbb{R}}(A). \square$$

Fakt 7.80. Niech V będzie przestrzenią liniową nad \mathbb{R} . Wtedy $\dim_{\mathbb{R}} V = \dim_{\mathbb{C}} V_{\mathbb{C}}$; dokładniej, dla $B \subseteq V$ mamy:

- B rozpina $V \iff B$ rozpina $V_{\mathbb{C}}$,
- B jest lnz w $V \iff B$ jest lnz w $V_{\mathbb{C}}$,
- B jest bazą $V \iff B$ jest bazą $V_{\mathbb{C}}$.

Dowód. Jeżeli B jest lnz nad \mathbb{C} , to oczywiście B jest też lnz nad \mathbb{R} . W drugą stronę, jeżeli B jest lnz nad \mathbb{C} , to znaczy, że dla pewnego $b \in B$ mamy $b \in \operatorname{Lin}_{\mathbb{C}}(B \setminus \{b\})$. Ale $b \in V$, więc z poprzedniego faktu wnioskujemy, że $b \in \operatorname{Lin}_{\mathbb{R}}(B \setminus \{b\})$, czyli B jest lnz nad \mathbb{R} .

Jeżeli B rozpina V nad \mathbb{R} , to nad \mathbb{C} rozpina $V + iV = \mathbb{C}V$. Z drugiej strony jeżeli $\operatorname{Lin}_{\mathbb{C}}(B) = V_{\mathbb{C}}$, to z poprzedniego faktu $\operatorname{Lin}_{\mathbb{R}}(B) = V_{\mathbb{C}} \cap V = V$.

Trzeci punkt wynika z pierwszych dwóch. \square

Definicja 7.81. Dla przekształcenia \mathbb{R} -liniowego $F: V \rightarrow W$, jego kompleksyfikacją to \mathbb{C} -liniowe $F_{\mathbb{C}}: V_{\mathbb{C}} \rightarrow W_{\mathbb{C}}$ zadane oczywistym wzorem $F_{\mathbb{C}}(v + iv') = F(v) + iF(v')$.

Uwaga 7.82. Jeżeli $F: V \rightarrow W$ jest \mathbb{R} -liniowe i B, C to bazy V, W , to zachodzi $m_C^B(F) = m_C^B(F_{\mathbb{C}})$ (odczytane nad \mathbb{C} ; to wynika wprost z definicji m_C^B).

W szczególności $F_{\mathbb{C}}$ ma w pewnej (ale być może nie w każdej) bazie macierz o współczynnikach rzeczywistych.

Przykład 7.83. • $v \mapsto \bar{v}$ jest \mathbb{R} -liniowe $V_{\mathbb{C}} \rightarrow V_{\mathbb{C}}$, nie jest jednak \mathbb{C} -liniowe (bo $i\bar{v} = -i\bar{v} \neq i\bar{v}$ gdy $v \neq 0$), a więc nie jest też kompleksyfikacją żadnego endomorfizmu V .

- Odwzorowanie $F: V_{\mathbb{C}} \rightarrow V_{\mathbb{C}}$, $v \mapsto iv$ również nie jest kompleksyfikacją żadnego rzeczywistego przekształcenia liniowego, np. dlatego że jeżeli B jest (rzeczywistą) bazą V , to macierz $m_B(F) = iI$ nie jest rzeczywistą macierzą.

- Mamy \mathbf{R} -liniowe odwzorowania $\operatorname{Re}, \operatorname{Im}: V_{\mathbf{C}} \rightarrow V$, $\operatorname{Re}(v + iw) = v$, $\operatorname{Im}(v + iw) = w$. Nie są one oczywiście \mathbf{C} -liniowe (V nie jest przestrzenią liniową nad \mathbf{C}).
- Odwzorowanie $F: \mathbf{R}^2 \rightarrow \mathbf{R}^2$ zadane macierzą $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ w standardowej bazie nie diagonalizuje się, ale w bazie $B = (e_1 + ie_2, e_1 - ie_2)$ mamy $m_B(F_{\mathbf{C}}) = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$.

Wniosek 7.84. Jeżeli $F \in \operatorname{End}(V)$, gdzie V jest rzeczywista, to $\chi_F(x) = \chi_{F_{\mathbf{C}}}(x)$, a więc ten drugi jest wielomianem o współczynnikach rzeczywistych i $F, F_{\mathbf{C}}$ mają te same rzeczywiste wartości własne.

Fakt 7.85. Jeżeli $F: V \rightarrow W$ (gdzie V, W są rzeczywiste), to $(\ker F)_{\mathbf{C}} = \ker(F_{\mathbf{C}})$.

Dowód. Weźmy $v = v_1 + iv_2 \in V_{\mathbf{C}}$. Wtedy z definicji $F_{\mathbf{C}}$:

$$\begin{aligned} v \in \ker F_{\mathbf{C}} &\iff 0 = F_{\mathbf{C}}(v_1 + iv_2) = F(v_1) + iF(v_2) \iff \begin{cases} F(v_1) = 0 \\ F(v_2) = 0 \end{cases} \\ &\iff \begin{cases} v_1 \in \ker F \\ v_2 \in \ker F \end{cases} \iff v = v_1 + iv_2 \in (\ker F)_{\mathbf{C}}. \quad \square \end{aligned}$$

Wniosek 7.86. Załóżmy że $r \in \mathbf{R}$. Wtedy $(V_{\mathbf{C}})_r = (V_r)_{\mathbf{C}}$ i $(V_{\mathbf{C}})^r = (V^r)_{\mathbf{C}}$.

Dowód. Z poprzedniego faktu dla $G = F - r$ dostajemy $(V_{\mathbf{C}})_r = (V_r)_{\mathbf{C}}$, a dla $G = (F - r)^{\dim V}$ $(V_{\mathbf{C}})^r = (V^r)_{\mathbf{C}}$. \square

Definicja 7.87. Przez $V_{\mathbf{C}, \lambda}$ i $V_{\mathbf{C}}^{\lambda}$ będziemy oznaczać przestrzenie własne i pierwiastkowe dla $F_{\mathbf{C}}$ (z poprzedniego wniosku widzimy, że nie ma tu dwuznaczności gdy $\lambda \in \mathbf{R}$).

Uwaga 7.88. • $F_{\mathbf{C}}(\bar{v}) = F_{\mathbf{C}}(v_1 - iv_2) = F(v_1) - iF(v_2) = \overline{F(v_1) + iF(v_2)} = \overline{F_{\mathbf{C}}(v)}$,

$$\bullet F(\operatorname{Re}(v)) = F_{\mathbf{C}}\left(\frac{v+\bar{v}}{2}\right) = \frac{F_{\mathbf{C}}(v) + F_{\mathbf{C}}(\bar{v})}{2} = \operatorname{Re}(F_{\mathbf{C}}(v)),$$

$$\bullet F(\operatorname{Im}(v)) = F_{\mathbf{C}}\left(\frac{v-\bar{v}}{2i}\right) = \operatorname{Im}(F_{\mathbf{C}}(v)).$$

Fakt 7.89. Jeżeli $F \in \operatorname{End}(V)$ i λ jest wartością własną $F_{\mathbf{C}}$, to $\bar{\lambda}$ też nią jest, a ponadto $V_{\mathbf{C}, \bar{\lambda}} = \overline{V_{\mathbf{C}, \lambda}}$ i $V_{\mathbf{C}}^{\bar{\lambda}} = \overline{V_{\mathbf{C}}^{\lambda}}$.

Dowód. Z poprzedniej uwagi dla dowolnych v, λ mamy $F_C(v) = \lambda v \iff F_C(\bar{v}) = \overline{F_C(v)} = \bar{\lambda} \bar{v}$, co daje $V_{C, \bar{\lambda}} = \overline{V_{C, \lambda}}$ (w szczególności zerowość tych dwóch przestrzeni jest równoważna).

Podobnym rachunkiem $\overline{(F_C - \lambda)^k(v)} = (F - \bar{\lambda})^k(\bar{v})$, co daje $V_C^{\bar{\lambda}} = \overline{V_C^{\lambda}}$. \square

Lemat 7.90. *Jeżeli $W \leq V_C$ (zespółona podprzestrzeń!) i $B = (b_1, \dots, b_n)$ jest bazą W , to, oznaczając $W_R = (W + \bar{W}) \cap V$ mamy:*

- $W_R = \operatorname{Re}[W] + i \operatorname{Im}[W]$,
- $(W_R)_C = W + \bar{W}$,
- jeżeli $W \cap \bar{W} = \{0\}$, to $B_R = (\operatorname{Re}(b_1), \operatorname{Im}(b_1), \operatorname{Re}(b_2), \dots, \operatorname{Im}(b_n))$ jest bazą W_R (nad \mathbf{R}) i $W + \bar{W}$ (nad \mathbf{C}).

Dowód. Dla każdego $w \in W$ mamy $\operatorname{Re}(w) = \frac{w + \bar{w}}{2} \in W_R$, co pokazuje $W_R \supseteq \operatorname{Re}[W]$. Z drugiej strony jeżeli $w_1 \in W, w_2 \in \bar{W}$ spełniają $w = w_1 + w_2 \in V$, to $w = \operatorname{Re}(w_1 + w_2) = \operatorname{Re}(w_1 + \bar{w}_2) \in \operatorname{Re}[W]$. Ponadto dla każdego $w \in W$ mamy $\operatorname{Im}(w) = \operatorname{Re}(iw)$, więc $\operatorname{Re}[W] = \operatorname{Im}[W]$.

Drugi punkt łatwo wynika z pierwszego: z jego pomocą łatwo zobaczyć, że obydwie strony są równe $\operatorname{Re}[W] + i \operatorname{Im}[W]$.

Trzeci punkt: jeżeli $W \cap \bar{W} = \{0\}$, to $\dim(W + \bar{W}) = 2n$. Ponieważ B_R ma $2n$ elementów i oczywiście rozpinają (nad \mathbf{C}) $\operatorname{Re}[W] + i \operatorname{Im}[W] = W + \bar{W}$, to B_R jest bazą $W + \bar{W}$, a więc (z drugiego punktu) jest też bazą W_R . \square

Definicja 7.91. Dla $\lambda \in \mathbf{C}$, oznaczmy przez R_λ rzeczywistą macierz 2×2 postaci

$$\begin{pmatrix} \operatorname{Re}(\lambda) & \operatorname{Im}(\lambda) \\ -\operatorname{Im}(\lambda) & \operatorname{Re}(\lambda) \end{pmatrix}.$$

(W szczególności jeżeli $\lambda = e^{i\theta}$, to R_λ jest macierzą zegarowego obrotu o θ względem standardowej bazy \mathbf{R}^2 .)

Stwierdzenie 7.92. *Jeżeli $F \in \operatorname{End}(V)$ jest taka, że F_C się diagonalizuje, to w pewnej (rzeczywistej) bazie B_0 przestrzeni V mamy*

$$m_B(F) = \begin{pmatrix} r_1 & & & & \\ & r_2 & & & \\ & & \ddots & & \\ & & & r_M & \\ & & & & R_{\lambda_1} & & \\ & & & & & \ddots & \\ & & & & & & R_{\lambda_N} \end{pmatrix}$$

gdzie r_j są rzeczywistymi, a λ_j nierzeczywistymi wartościami własnymi F (z dokładnością do sprzężenia; można przyjąć dla ustalenia uwagi, że $\text{Im}(\lambda_j) > 0$). Wszystkie wartości występują z takimi krotnościami, jak ich krotności w $\chi_F(x)$.

Uwaga 7.93. Nietrudno pokazać, że (zmieniając bazę przez zamianę miejscami dwóch wektorów bazowych) można każde R_{λ_j} zamienić na $R_{\bar{\lambda}_j}$.

Przykład 7.94. Jeżeli $F: \mathbf{R}^3 \rightarrow \mathbf{R}^3$ jest nietrywialnym obrotem wokół prostej. Wtedy 1 jest wartością własną i $\det F = 1$, więc $\chi_F(x)$ ma 3 pierwiastki: λ i $\bar{\lambda}$, gdzie $|\lambda| = 1$. Stąd $\lambda = \cos \theta + i \sin \theta$ dla pewnego $\theta \in [0, 2\pi)$. Powyższe stwierdzenie mówi, że istnieje baza B \mathbf{R}^3 taka że

$$m_B(F) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{pmatrix}$$

Dowód. Ustalmy nierzeczywistą λ .

Niech $B_\lambda = (b_1, \dots, b_n)$ będzie bazą $V_{C,\lambda}$. Wtedy

$$B_{\lambda,R} = (\text{Re}(b_1), \text{Im}(b_1), \text{Re}(b_2), \dots, \text{Im}(b_n))$$

(jak w lemacie) jest bazą $V_{\lambda,R} := \text{Re}[V_{C,\lambda}] \oplus V_{C,\bar{\lambda}}$ (bo $V_{C,\lambda}$ i $V_{C,\bar{\lambda}}$ są lnz!) i korzystając z uwagi widzimy, że

$$F(\text{Re}(b_j)) = \text{Re}(F_C(b_j)) = \text{Re}(\lambda b_j) = \text{Re}(\lambda) \text{Re}(b_j) - \text{Im}(\lambda) \text{Im}(b_j),$$

i podobnie $F(\text{Im}(b_j)) = \text{Im}(\lambda) \text{Re}(b_j) + \text{Re}(\lambda) \text{Im}(b_j)$. Stąd

$$m_{B_{\lambda,R}}(F_C|_{V_{C,\lambda}+V_{C,\bar{\lambda}}}) = m_{B_{\lambda,R}}(F|_{V_{\lambda,R}}) = \begin{pmatrix} R_\lambda & & \\ & R_\lambda & \\ & & \ddots \\ & & & R_\lambda \end{pmatrix}$$

Dla rzeczywistej wartości własnej r wiemy że $V_{C,r} = (V_r)_C$, więc możemy znaleźć rzeczywistą bazę $B_r \subseteq V_r$ dla $V_{C,r}$ i w niej macierz $m_{B_r}(F_C|_{V_{C,r}})$ jest diagonalna (z r na przekątnej).

Powtarzamy tę procedurę dla wszystkich zespolonych pierwiastków $\chi_F(x)$ (czyli wartości własnych F_C), dbając przy tym, żeby wziąć dokładnie jedną z każdej pary $\lambda, \bar{\lambda}$.

Sklejając uzyskane bazy do $B = B_{r_1} B_{r_2} \dots B_{r_M} B_{\lambda_1} \dots B_{\lambda_N}$ zauważamy, że z liniowej niezależności $V_{C,\lambda}$ wynika, że B jest (rzeczywistą!) bazą V i V_C i wtedy $m_B(F) = m_B(F_C)$ równa się

$$\begin{pmatrix} m_{B_{r_1}}(F_C \upharpoonright_{V_{C,r_1}}) & & & & \\ & \ddots & & & \\ & & m_{B_{r_l}}(F_C \upharpoonright_{V_{C,r_l}}) & & \\ & & & m_{B_{\lambda_1,R}}(F_C \upharpoonright_{V_{C,\lambda_1} + V_{C,\bar{\lambda}_1}}) & \\ & & & & \ddots \\ & & & & & m_{B_{\lambda_k,R}}(F_C \upharpoonright_{V_{C,\lambda_k} + V_{C,\bar{\lambda}_k}}) \end{pmatrix}$$

i macierze w blokach na przekątnej są zadanej postaci, co daje tezę. \square

Co gdy F_C się *nie* diagonalizuje? Dla rzeczywistych $r \in \sigma(F_C)$ możemy wybrać bazy Jordana dla V^r . Dla nierzeczywistych λ potrzebujemy pomocniczej definicji.

Definicja 7.95. *Rzeczywista klatka Jordana* rozmiaru $2m$ odpowiadająca $\lambda \in \mathbb{C} \setminus \mathbb{R}$ (nierzeczywistej!) to macierz blokowa postaci

$$J_{\lambda,m} := \underbrace{\begin{pmatrix} R_\lambda & I & & & 0 \\ & R_\lambda & I & & \\ & & R_\lambda & I & \\ & & & \ddots & I \\ 0 & & & & R_\lambda \end{pmatrix}}_{2m \text{ kolumn}}$$

gdzie $R_\lambda = \begin{pmatrix} \operatorname{Re}(\lambda) & \operatorname{Im}(\lambda) \\ -\operatorname{Im}(\lambda) & \operatorname{Re}(\lambda) \end{pmatrix}$ a I nad przekątną to macierze identycznościowe 2×2 .

Następnie postępujemy podobnie jak powyżej: dla rzeczywistych wartości własnych r postępujemy jak w zwykłym dowodzie twierdzenia Jordana znajdując bazy V^r , a dla nierzeczywistej λ wybieramy bazę B_λ Jordana dla V_C^λ , tworzymy z niej (rzeczywistą!) bazę $B_{\lambda,R}$ dla $V_C^\lambda + V_C^{\bar{\lambda}}$. Wtedy $m_{B_{\lambda,R}}(F_C \upharpoonright_{V_C^\lambda + V_C^{\bar{\lambda}}})$ ma postać Jordana. Powtarzamy tę procedurę dla wszystkich wartości własnych (z dokładnością do sprzężenia) i sklejamy powstałe bazy.

Ten tok rozumowania prowadzi nas do

Twierdzenie 7.96 (Rzeczywiste twierdzenie Jordana). *Jeżeli $F \in \text{End}(V)$ i V jest skończone wymiarową przestrzenią liniową nad \mathbf{R} , to V ma bazę B taką że*

$$m_B(F) = \begin{pmatrix} J_{\lambda_1, m_1} & & & \\ & J_{\lambda_2, m_2} & & \\ & & \ddots & \\ & & & J_{\lambda_k, m_k} \end{pmatrix},$$

gdzie λ_j są (wszystkimi, z dokładnością do sprzężenia) zespolonymi pierwiastkami $\chi_F(x)$.

Ponadto rozmiary klatek odpowiadają dokładnie rozmiarom klatek w rozkładzie Jordana dla $F_{\mathbf{C}}$ (tzn. są takie same w przypadku rzeczywistych λ , podwojone w przypadku nierzeczywistych λ).

Szczegóły: Dodatek B do skryptu (kiedyś).

Definicja 7.97. Powyższą postać macierzy nazywamy *rzeczywistą postacią Jordana*.

Przykład 7.98. Chcemy znaleźć zespoloną i rzeczywistą postać Jordana macierzy

$$A = \begin{pmatrix} 3 & 1 & 1 & -4 & 0 & 0 \\ 0 & 6 & 0 & -10 & 0 & 0 \\ -1 & 3 & 3 & -10 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & -1 \\ 0 & 0 & 0 & 0 & 10 & -1 \end{pmatrix}$$

Wielomian charakterystyczny to iloczyn wyznaczników

$$\begin{vmatrix} 3-x & 1 & 1 & -4 \\ 0 & 6-x & 0 & -10 \\ -1 & 3 & 3-x & -10 \\ 0 & 1 & 0 & -x \end{vmatrix} \begin{vmatrix} 5-x & -1 \\ 10 & -1-x \end{vmatrix},$$

Drugi wyznacznik to $x^2 - 4x + 5$, jego pierwiastki to $\frac{4 \pm \sqrt{16-20}}{2} = \frac{4 \pm 2i}{2} = 2 \pm i$. Biorąc rozwinięcie Laplace'a względem ostatniego wiersza, pierwszy z

wyznaczników to

$$\begin{aligned}
 \begin{vmatrix} 3-x & 1 & -4 \\ 0 & 0 & -10 \\ -1 & 3-x & -10 \end{vmatrix} - x \begin{vmatrix} 3-x & 1 & 1 \\ 0 & 6-x & 0 \\ -1 & 3 & 3-x \end{vmatrix} &= \\
 = 10 \begin{vmatrix} 3-x & 1 \\ -1 & 3-x \end{vmatrix} - x(6-x) \begin{vmatrix} 3-x & 1 \\ -1 & 3-x \end{vmatrix} &= \\
 = (10-x(6-x))((3-x)^2+1) &= \\
 = (x^2-6x+10)(x^2-6x+10), &
 \end{aligned}$$

jego pierwiastki (obydwa podwójne) to $\frac{6 \pm \sqrt{36-40}}{2} = \frac{6 \pm 2i}{2} = 3 \pm i$.

Ze struktury blokowej wynika, że wektor własny dla $2+i$ będzie miał niezerowe tylko dwie ostatnie współrzędne. Musi spełniać $5x_5 - x_6 = (2+i)x_5$,

czyli $x_6 = (3-i)x_5$, czyli przykładowo $c_5 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 3-i \end{pmatrix}$ działa, a ponieważ

macierz jest rzeczywista, $c_6 = \bar{c}_5 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 3+i \end{pmatrix}$ jest wektorem własnym dla $2-i$.

Szukamy teraz uogólnionych wektorów własnych dla $3+i$. Podobnie, wystarczy rozważać pierwsze 4 współrzędne szukamy więc elementów

$$\ker \begin{pmatrix} -i & 1 & 1 & -4 \\ 0 & 3-i & 0 & -10 \\ -1 & 3 & -i & -10 \\ 0 & 1 & 0 & -3-i \end{pmatrix}^2 = \ker \begin{pmatrix} -2 & 2-2i & -2i & -8+8i \\ 0 & -2-6i & 0 & 20i \\ 2i & -2-6i & -2 & 4+20i \\ 0 & -2i & 0 & -2+6i \end{pmatrix}$$

Stosując metodę Gaussa, odejmujemy drugi wiersz od trzeciego, a następnie czwarty od pierwszego i trzy razy czwarty od drugiego, co daje nam

$$\begin{pmatrix} -2 & 2 & -2i & -6+2i \\ 0 & -2 & 0 & 6+2i \\ 2i & 0 & -2 & 4 \\ 0 & -2i & 0 & -2+6i \end{pmatrix}$$

teraz dodajemy drugi wiersz do pierwszego i odejmujemy i razy drugi od czwartego, co daje

$$\begin{pmatrix} -2 & 0 & -2i & 4i \\ 0 & -2 & 0 & 6+2i \\ 2i & 0 & -2 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Łatwo zobaczyć, że pierwszy i trzeci wiersz są współliniowe. Szukane wektory to są więc rozwiązania $x_2 = (3+i)x_4$, $x_1 = -ix_3 + 2ix_4$. Przykładowo

$$\begin{pmatrix} 2i \\ 3+i \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

Nakładając $(A-3-i)$ otrzymujemy wektor $\begin{pmatrix} 1+i \\ 0 \\ -1+i \\ 0 \\ 0 \\ 0 \end{pmatrix}$. Znów, zważywszy że

A jest rzeczywistą macierzą, wektory sprzężone są uogólnionymi wektorami własnymi dla sprzężonych wartości własnych, tak więc

$$A = P \begin{pmatrix} 3+i & 1 & & & & \\ & 3+i & & & & \\ & & 3-i & 1 & & \\ & & & 3-i & & \\ & & & & 2+i & \\ & & & & & 2-i \end{pmatrix} P^{-1}$$

dla

$$P = \begin{pmatrix} 1+i & 2i & 1-i & -2i & 0 & 0 \\ 0 & 3+i & 0 & 3-i & 0 & 0 \\ -1+i & 0 & -1-i & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 3-i & 3+i \end{pmatrix}$$

oraz

$$A = Q \begin{pmatrix} 3 & -1 & 1 & 0 & & \\ 1 & 3 & 0 & 1 & & \\ & & 3 & -1 & & \\ & & 1 & 3 & & \\ & & & & 2 & 1 \\ & & & & -1 & 2 \end{pmatrix} Q^{-1}$$

dla

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

Uwaga 7.99. Można sformułować też wersję tw. Jordana dla dowolnych ciał, z tym że w ogólności rozmiary „podklatek” mogą być dowolnie duże — zależą od stopni nierozkładalnych wielomianów, na które rozkłada się $\chi_F(x)$.

Wszystkie wersje twierdzenia Jordana uogólniają się na tzw. zasadnicze twierdzenie o skończeniu generowanych modułów nad pierścieniami ideałów głównych. (O którym można usłyszeć np. na wykładzie z algebry przemiennej, na raczej ≥ 3 roku...⁷)

Przykład 7.100. Chcemy rozwiązać równanie różniczkowe (liniowe, jednorodne) postaci

$$f^{(n)} = a_{n-1}f^{(n-1)} + \dots + a_1f' + a_0f,$$

gdzie a_k są pewnymi rzeczywistymi (lub zespolonymi!) współczynnikami.

⁷W wersji dla odważnych, można już teraz zapytać o to np. Gala.

Zauważmy że jest ono równoważne równaniu

$$\begin{pmatrix} f' \\ \vdots \\ f^{(n)} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & 1 \\ a_0 & a_1 & \dots & a_{n-2} & a_{n-1} \end{pmatrix}}_A \begin{pmatrix} f \\ \vdots \\ f^{(n-1)} \end{pmatrix}.$$

W skrócie $v'(t) = Av(t)$, gdzie $v(t) = (f(t), f'(t), \dots, f^{(n-1)}(t))^T$.

Szukamy rozwiązań postaci $v(t) = \begin{pmatrix} f(t) \\ \vdots \\ f^{(n-1)}(t) \end{pmatrix} = e^{tB} v_0$ dla pewnej macierzy B i wektora (warunku początkowego) v_0 gdzie $e^{tB} = \sum_{k=0}^{\infty} \frac{(tA)^k}{k!} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k$ (ten szereg zawsze jest bezwzględnie zbieżny), czyli rozwiązań równania

$$(e^{tB} v_0)' = Ae^{tB} v_0.$$

Różniczkując $e^{tB} v_0$ względem t dostajemy

$$\frac{d}{dt}(e^{tB} v_0) = Be^{tB} v_0,$$

chcemy zatem mieć

$$Ae^{tB} v_0 = Be^{tB} v_0,$$

skąd widać, że rozwiązaniami są wszystkie funkcje postaci $v(t) = e^{tA} v_0$, czyli $f(t)$ jest pierwszą współrzędną tegoż.

Jak wyznaczać efektywnie rozwiązania tego równania? Na przykład wyznaczamy postać Jordana $A = PJP^{-1}$. Wtedy $e^{tA} = Pe^{tJ}P^{-1}$. Korzystając z jawnych wzorów na potęgi klatek Jordana nietrudno wyznaczyć jawne wzory na exponensy klatek Jordana (także pomnożone przez t).

Na przykład dla $f'' = 4f' - 4f$ mamy $A = \begin{pmatrix} 0 & 1 \\ -4 & 4 \end{pmatrix}$, $A = PJP^{-1}$, gdzie

$$P = \begin{pmatrix} -2 & 1 \\ -4 & 0 \end{pmatrix} \text{ i } J = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}. \quad \frac{t^k J^k}{k!} = \begin{pmatrix} \frac{(2t)^k}{k!} & \frac{k \cdot t^k 2^{k-1}}{k!} \\ 0 & \frac{(2t)^k}{k!} \end{pmatrix}.$$

Wtedy

$$e^{tJ} = \sum_{k=0}^{\infty} \begin{pmatrix} \frac{(2t)^k}{k!} & k t \frac{(2t)^{k-1}}{k!} \\ 0 & \frac{(2t)^k}{k!} \end{pmatrix} = \begin{pmatrix} e^{2t} & t e^{2t} \\ 0 & e^{2t} \end{pmatrix},$$

skąd $e^{tA} = Pe^{tJ}P^{-1} = \begin{pmatrix} -2te^{2t} + e^{2t} & te^{2t} \\ -4te^{2t} & 2te^{2t} + e^{2t} \end{pmatrix}$, więc rozwiązanie ogólne jest postaci $f(t) = a(1-2t)e^{2t} + bte^{2t}$ gdzie $a = f(0)$, $b = f'(0)$ są dowolnymi stałymi; grupując wyrazy otrzymujemy wzór ogólny $f(t) = ce^{2t} + dte^{2t}$. Bezpośrednim rachunkiem można sprawdzić, że to faktycznie są rozwiązania równania $f'' = 4f' - 4f$.

(Szczegóły: wykład równania różniczkowe 1/1 R)

Rozdział 8

Formy dwuliniowe i kwadratowe

Od teraz przez jakiś czas znów wracamy do dowolnego ciała K .

8.1 Formy dwuliniowe

Przykład 8.1. Rozważmy iloczyn skalarny na $\mathbf{R}^n \times \mathbf{R}^n \rightarrow \mathbf{R}$, zadany wzorem

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \sum_{k=1}^n x_k y_k.$$

Ma on następujące własności:

- jest dwuliniowy, tzn. liniowy w obydwu argumentach,
- jest symetryczny, tzn. spełnia $\langle v, w \rangle = \langle w, v \rangle$,
- jest dodatnio określony, tzn. spełnia $\langle v, v \rangle > 0$, gdy $v \neq \vec{0}$.

Będziemy badać inne podobne funkcje na przestrzeniach liniowych nad różnymi ciałami.

Definicja 8.2. Forma dwuliniowa na V (p. lin. nad K) to odwzorowanie $\varphi: V \times V \rightarrow K$ liniowe w każdym argumencie.

Mówimy że φ jest symetryczna, jeżeli spełnia $\varphi(v, w) = \varphi(w, v)$, a antysymetryczna (lub skośnie symetryczna), jeżeli $\varphi(v, w) = -\varphi(w, v)$.

Przykłady 8.3. • na K^n mamy symetryczną formę dwuliniową $\varphi \left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right) =$

$$\sum_{i=1}^n x_i y_i$$

- \det na K^2 jest formą dwuliniową
- jeżeli ustalimy wektory $c_1, c_2, \dots, c_{n-2} \in K^n$, to $\varphi(v, w) = \det(c_1, c_2, \dots, c_{n-2}, v, w)$ jest dwuliniowe
- na $C([0, 1])$ (przestrzeń funkcji ciągłych na kwadracie jednostkowym) mamy symetryczną formę dwuliniową $\varphi(f, g) = \int_0^1 e^t f(t) g(t) dt$.
- dla każdej $A = (a_{kj})_{kj} \in M_{n \times n}(K)$ mamy formę dwuliniową na K^n zadaną wzorem

$$\varphi_A(v, w) = v^\top \cdot A \cdot w = \sum_{k=1}^n \sum_{j=1}^n \underbrace{v_k}_{k\text{-ta współrzędna } v} a_{kj} \underbrace{w_j}_{j\text{-ta współrzędna } w}$$

Np. dla $A = I$ i $K = \mathbf{R}$ dostajemy zwykły iloczyn skalarny.

Uwaga 8.4. Zbiór wszystkich form dwuliniowych jest zamknięty na dodawanie i mnożenie przez skalary, stanowi więc przestrzeń liniową. Podobnie zbiory form symetrycznych i antysymetrycznych.

Można to sprawdzić bezpośrednim rachunkiem, bądź zauważając, że zbiór form dwuliniowych na V ma naturalną bijekcję z $\text{Hom}(V, V^*)$, $\varphi \mapsto (v \mapsto \varphi(v, -))$.

Uwaga 8.5. Jeżeli φ jest formą dwuliniową na V o bazie $B = (b_1, \dots, b_n)$, to dla $v = \sum_k \alpha_k b_k$, $w = \sum_j \beta_j b_j$ mamy

$$\varphi(v, w) = \varphi\left(\sum_k \alpha_k b_k, \sum_j \beta_j b_j\right) = \sum_k \sum_j \alpha_k \varphi(b_k, b_j) \beta_j$$

Definicja 8.6. Niech $B = (b_1, \dots, b_n)$ będzie bazą V , a φ — formą dwuliniową na V .

Macierz V względem bazy B to

$$m^B(\varphi) = m^{BB}(\varphi) = \begin{pmatrix} \varphi(b_1, b_1) & \varphi(b_1, b_2) & \cdots & \varphi(b_1, b_n) \\ \varphi(b_2, b_1) & \varphi(b_2, b_2) & & \vdots \\ \vdots & & \ddots & \vdots \\ \varphi(b_n, b_1) & \cdots & \cdots & \varphi(b_n, b_n) \end{pmatrix}$$

Wniosek 8.7.

$$\varphi(v, w) = [v]_B^\top m^{BB}(\varphi) [w]_B$$

(Jeżeli B jest standardową bazą K^n , to $\varphi = \varphi_{m^{BB}(\varphi)}$.)

Dowód. Jeżeli $v = \sum_k \alpha_k b_k$, $w = \sum_j \beta_j b_j$, to $[v]_B = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$ $[w]_B = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$ i

mamy:

$$\begin{aligned} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}^\top \begin{pmatrix} \varphi(b_1, b_1) & \varphi(b_1, b_2) & \cdots & \varphi(b_1, b_n) \\ \varphi(b_2, b_1) & \varphi(b_2, b_2) & & \vdots \\ \vdots & & \ddots & \vdots \\ \varphi(b_n, b_1) & \cdots & \cdots & \varphi(b_n, b_n) \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \\ = (\alpha_1 \quad \cdots \quad \alpha_n) \begin{pmatrix} \sum_j \varphi(b_1, b_j) \beta_j \\ \vdots \\ \sum_j \varphi(b_n, b_j) \beta_j \end{pmatrix} = \sum_k \sum_j \alpha_k \varphi(b_k, b_j) \beta_j \quad \square \end{aligned}$$

Przykład 8.8. Rozważmy formę dwuliniową $\varphi(P, Q) = P(1)Q(2)$ na $\mathbf{R}_2[X]$. W

bazie $B = (X(X-1), (X-1)(X-2), X(X-2))$ mamy $m^{BB}(\varphi) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -2 & 0 & 0 \end{pmatrix}$.

Z drugiej strony w bazie $C = (1, X, X^2)$ mamy $m^{CC}(\varphi) = \begin{pmatrix} 1 & 2 & 4 \\ 1 & 2 & 4 \\ 1 & 2 & 4 \end{pmatrix}$.

Uwaga 8.9. Jeżeli B jest bazą przestrzeni n -wymiarowej V , to $\varphi \mapsto m^{BB}(\varphi)$ zadaje bijekcję między zbiorem wszystkich form dwuliniowych na V , a $M_{n \times n}(K)$.

Definicja 8.10. Mówimy że macierz A jest symetryczna jeżeli $A^\top = A$.

Fakt 8.11. Następujące warunki są równoważne:

- φ jest symetryczna,
- macierz φ w (każdej \iff pewnej) bazie jest symetryczna.

Dowód. Załóżmy że $m^{BB}(\varphi)$ jest symetryczna. Wtedy

$$\varphi(v, w) = \varphi(v, w)^\top = ([v]_B^\top m^{BB}(\varphi) [w]_B)^\top = [w]_B^\top m^{BB}(\varphi)^\top [v]_B^{\top\top} = [w]_B^\top m^{BB}(\varphi) [v]_B = \varphi(w, v),$$

więc φ jest symetryczna.

Z drugiej strony wprost z definicji wynika, że jeżeli φ jest symetryczna, to jej macierze są symetryczne. \square

Założmy teraz, że mamy bazy B, C przestrzeni V i znamy $m^{BB}(\varphi)$. Jak wyznaczyć $m^{CC}(\varphi)$?

Wniosek 8.12. $m^{CC}(\varphi) = m_B^C(\text{id})^\top \cdot m^{BB}(\varphi) \cdot m_B^C(\text{id})$

Dowód. Z uwagi wynika, że wystarczy sprawdzić, że $m^{CC}(\varphi)$ zadaje φ (zgodnie ze wzorem z Wniosku 8.7). Weźmy dowolne $v, w \in V$. Przypomnijmy sobie, że $[v]_B = m_B^C(\text{id}) \cdot [v]_C$. Stąd łatwo wynika:

$$\begin{aligned} \varphi(v, w) &= [v]_B^\top m^{BB}(\varphi) [w]_B \\ &= (m_B^C(\text{id}) \cdot [v]_C)^\top m^{BB}(\varphi) (m_B^C(\text{id}) \cdot [w]_C) \\ &= [v]_C^\top \cdot m_B^C(\text{id})^\top m^{BB}(\varphi) m_B^C(\text{id}) \cdot [w]_C \end{aligned}$$

Ponieważ v, w były dowolne, to kończy dowód. \square

Uwaga 8.13. Z powyższego wzoru widać, że w ogólności nie działa wzór $m^{CC}(\varphi) = (m_B^C(\text{id}))^{-1} \cdot m^{BB}(\varphi) \cdot m_B^C(\text{id})$, chyba że $m_B^C(\text{id})^{-1} = m_B^C(\text{id})^\top$.

Definicja 8.14.

równoważność form dwuliniowych Mówimy że dwie formy dwuliniowe na V są równoważne jeżeli mają te same macierze (niekoniecznie względem tych samych baz).

Wniosek 8.15. Macierze $A, B \in M_{n \times n}(K)$ zadają równoważne formy dwuliniowe wtedy i tylko wtedy gdy istnieje taka odwracalna P , że $B = PAP^\top$.

8.2 Iloczyny skalarne

Od teraz do odwołania pracujemy nad \mathbf{R} (chyba że podano inaczej).

Przykład 8.16. $A = m^{BB}(\langle -, - \rangle)$ (tu $\langle -, - \rangle$ jest standardowym iloczynem skalarnym na \mathbf{R}^n) dla pewnej bazy B wtedy i tylko wtedy gdy spełnia $A = P^\top IP = P^\top P$, gdzie P jest odwracalna. Co więcej, $P = m_E^B(\text{id})$ jest macierzą której kolumny to wektory z bazy B .

- Jeżeli P jest ortogonalna, to znaczy $P^\top = P^{-1}$, to $A = I$, czyli w nowych współrzędnych zachodzi

$$\langle x, y \rangle = \sum_{j=1}^n x'_j y'_j$$

(tu x'_j, y'_j to współrzędne x, y względem nowej bazy).

- w ogólności tak nie jest; np. dla nowej bazy $B = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ mamy $P = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, więc $A = P^\top P = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, czyli we współrzędnych pochodzących od B mamy $\langle x, y \rangle = 2x'_1 y'_1 + x'_1 y'_2 + x'_2 y'_1 + x'_2 y'_2$.

To daje nam teoretyczny opis wszystkich możliwych form dwuliniowych równoważnych ze standardowym iloczynem skalarnym: są to wszystkie formy, których macierze są postaci $A = P^\top P$ dla pewnej odwracalnej P . Niestety nie wiemy jeszcze, jak stwierdzić, czy dana A jest tej postaci, albo jak wyznaczyć P , znając A ...

Zauważmy że każda forma dwuliniowa równoważna z iloczynem skalarnym musi być symetryczna i dodatnio określona (bo te własności nie zależą od wyboru bazy). Z tego wynika, że jeżeli A reprezentuje iloczyn skalarny, to musi być macierzą symetryczną, a ponadto dla każdego niezerowego wektora v musi spełniać $v^\top A v > 0$.

Definicja 8.17. Mówimy że macierz symetryczna $A \in M_{n \times n}(\mathbf{R})$ jest *dodatnio określona* jeżeli dla każdego $v \in \mathbf{R}^n$, $v \neq 0$ mamy

$$v^\top A v > 0.$$

Ogólniej:

Definicja 8.18. Mówimy że forma symetryczna φ na V (nad \mathbf{R}) jest *dodatnio określona*, jeżeli $\varphi(v, v) > 0$ dla $v \neq 0$ (gdy $\dim(V) < \infty$, równoważnie: jej macierz względem pewnej/każdej bazy jest dodatnio określona).

Mówimy wtedy że φ jest *iloczynem skalarnym*.

Przykład 8.19. Na przestrzeni $C([0, 1])$ forma $\varphi(f, g) = \int_0^1 f(t)g(t) dt$ jest iloczynem skalarnym.

Pokażemy następujące twierdzenie.

Twierdzenie 8.20. *Następujące warunki są równoważne dla $A \in M_{n \times n}(\mathbf{R})$.*

1. *A reprezentuje standardowy iloczyn skalarny na \mathbf{R}^n (względem pewnej bazy),*
2. *istnieje macierz odwracalna P taka że $A = P^T P$,*
3. *A jest symetryczna i dodatnio określona.*

Równoważność pierwszych dwóch warunków została już pokazana, podobnie jak fakt, że implikują trzeci. Brakującą implikację będzie wynikała z bardziej ogólnego twierdzenia.

Definicja 8.21. Mówimy że para wektorów $v, w \in V$ jest *ortogonalna* względem formy symetrycznej dwuliniowej φ jeżeli $\varphi(v, w) = 0$. Piszemy $v \perp w$.

Mówimy że wektor v jest *unormowany* (lub *jednostkowy*) względem φ jeżeli $\varphi(v, v) = 1$.

Mówimy że układ (v_1, \dots, v_n) wektorów w przestrzeni V jest *ortogonalny* (względem φ) jeżeli jego elementy są parami ortogonalne, a *ortonormalny*, jeżeli są ponadto unormowane, tzn. dla dowolnych i, j mamy:

$$\varphi(v_i, v_j) = \begin{cases} 0 & i \neq j \\ 1 & i = j. \end{cases}$$

Uwaga 8.22. Każdy ortonormalny układ wektorów jest liniowo niezależny. Ogólnie, każdy ortogonalny układ wektorów nieizotropowych (takich że $\varphi(v, v) \neq 0$) jest liniowo niezależny.

Dowód. Ćwiczenie

□

Twierdzenie 8.23 (Ortogonalizacja Grama-Schmidta). *Założmy że φ jest dodatnio określoną, symetryczną formą dwuliniową na skończonej wymiarowej przestrzeni liniowej V nad \mathbf{R} . Wtedy istnieje baza ortonormalna względem φ .*

Dowód. Niech $B = (b_1, b_2, \dots, b_n)$ będzie dowolną bazą V . Zdefiniujemy rekurencyjnie bazę $B' = (b'_1, \dots, b'_n)$ ortonormalną względem φ , tak że $b'_k \in \text{Lin}(b_1, \dots, b_k)$ i $\text{Lin}(b_1, \dots, b_k) = \text{Lin}(b'_1, \dots, b'_k)$.

Będziemy oznaczać $\varphi(v, w)$ w skrócie przez $[v, w]$.

Niech $b'_1 = \frac{b_1}{\sqrt{[b_1, b_1]}}$. Wtedy

$$[b'_1, b'_1] = \left[\frac{b_1}{\sqrt{\varphi(b_1, b_1)}}, \frac{b_1}{\sqrt{\varphi(b_1, b_1)}} \right] = \frac{[b_1, b_1]}{\sqrt{[b_1, b_1]}^2} = 1.$$

Chcemy wyznaczyć $b'_2 \in \text{Lin}(b_1, b_2) = \text{Lin}(b'_1, b_2)$, ortogonalny do b'_1 . Rozważmy wektor postaci $v = \alpha_1 b'_1 + b_2$. Wtedy

$$[b'_1, v] = [b'_1, \alpha_1 b'_1 + b_2] = [\alpha_1 b'_1, b'_1] + [b'_1, b_2] = \alpha_1 + [b'_1, b_2].$$

Stąd wektor $b''_2 = b_2 - [b'_1, b_2]b'_1$ jest ortogonalny do b'_1 i jest niezerowy (bo b'_1, b_2 są liniowo niezależne), więc z dodatniej określoności $[b''_2, b''_2] > 0$ i wektor $b'_2 = \frac{b''_2}{\sqrt{[b''_2, b''_2]}}$ jest unormowany, ortogonalny do b'_1 .

Założmy że $k < n$ i mamy już ortonormalny układ b'_1, \dots, b'_k wektorów z $\text{Lin}(b_1, \dots, b_k)$. Weźmy $b''_{k+1} = b_{k+1} - \sum_{j=1}^k [b_{k+1}, b'_j]b'_j$. Wtedy dla $l \leq k$ mamy:

$$\begin{aligned} [b''_{k+1}, b'_l] &= [b_{k+1} - \sum_{j=1}^k [b_{k+1}, b'_j]b'_j, b'_l] \\ &= [b_{k+1}, b'_l] - \sum_{j=1}^k [b_{k+1}, b'_j] \underbrace{[b'_j, b'_l]}_{\begin{cases} 1 & j=l \\ 0 & j \neq l \end{cases}} \\ &= [b_{k+1}, b'_l] - [b_{k+1}, b'_l] \\ &= 0. \end{aligned}$$

Z liniowej niezależności wyjściowej bazy widzimy że b''_{k+1} jest niezerowy, możemy więc go unormować, uzyskując $b'_{k+1} = \frac{b''_{k+1}}{\sqrt{[b''_{k+1}, b''_{k+1}]}}$.

Uzyskany układ b'_1, \dots, b'_n jest ortonormalny, więc jest liniowo niezależny, jest więc bazą V . \square

Definicja 8.24. Mówimy że baza B diagonalizuje formę dwuliniową φ gdy $m^{BB}(\varphi)$ jest diagonalna.

Wniosek 8.25. Jeżeli φ jest iloczynem skalarnym (tzn. jest dwuliniowa, dodatnio określona i symetryczna) \iff istnieje baza B spełniająca $m^{BB}(\varphi) = I$.
W szczególności każdy iloczyn skalarny diagonalizuje się w pewnej bazie.

Dowód. \Rightarrow : dowolna baza ortonormalna działa: istotnie, jeżeli b_1, \dots, b_n jest bazą ortonormalną, to z definicji i -ta współrzędna $m^{BB}(\varphi)$ to

$$\varphi(b_i, b_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases},$$

więc $m^{BB}(\varphi) = I$.

\Leftarrow : ćwiczenie. □

To daje nam brakującą implikację z poprzedniego Twierdzenia.

Wniosek 8.26. Jeżeli $A \in M_{n \times n}(\mathbf{R})$ jest symetryczna i dodatnio określona, to forma $\varphi_A(v, w) = v^T A w$ jest równoważna ze standardowym iloczynem skalarnym na \mathbf{R}^n .

Dowód. Z symetryczności i dodatniej określoności wynika łatwo, że φ_A jest iloczynem skalarnym, więc teza wynika natychmiast z poprzedniego wniosku. □

Wiemy zatem, że macierze iloczynu skalarnego to dokładnie dodatnio określone, symetryczne macierze. Symetryczność łatwo sprawdzić, ale co z dodatnią określonością?

Można to sprawdzić „brutalnie”: spróbować zastosować proces Grama-Schmidta. Jeżeli się uda, to znaczy że forma (a więc macierz) jest iloczynem skalarnym, a więc w szczególności jest dodatnio określona. Jeżeli nie, to znaczy że nie jest.

Kryterium nie jest takie proste.

Definicja 8.27. Dla macierzy kwadratowej A *minor główny* to minor w którym występują wiersze o tych samych indeksach, co występujące w nim kolumny („symetryczny”).

k -ty wiodący *minor główny* to minor złożony z pierwszych k kolumn pierwszych k wierszy.

Twierdzenie 8.28 (Kryterium Sylwestera). *Założmy że $A \in M_{n \times n}(\mathbf{R})$. Niech A_k oznacza k -ty wiodący minor główny A , tj. $A_k = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$, gdzie*

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

Wtedy A jest dodatnio określona wtedy i tylko wtedy gdy $\det(A_k) > 0$ (dla wszystkich $k = 1, 2, \dots, n$).

Twierdzenie to udowodnimy później.

8.3 Ortogonalność

Przykład 8.29. W szczególnej teorii względności ważną rolę odgrywa tzw. forma Lorentza (czasami definiowana z przeciwnymi znakami):

$$\varphi((x, t), (y, s)) = x_1 y_1 + x_2 y_2 + x_3 y_3 - c^2 ts,$$

gdzie c to stała reprezentująca prędkość światła. Normalizując, możemy wybrać bazę B czasoprzestrzeni $\mathbf{R}^4 = \mathbf{R}^{3+1}$ taką że

$$m^{BB}(\varphi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

ma ona więc bardzo prosty opis (względem odpowiednio dobranej bazy).

Cel: znaleźć podobnie prosty opis *każdej* symetrycznej formy dwuliniowej (np. w postaci macierzy diagonalnej). W szczególności chcemy znaleźć ortogonalną bazę dla dowolnej takiej formy.

W tej części wykładu φ jest ustaloną (ale dowolną) symetryczną formą dwuliniową na rzeczywistej przestrzeni V .

Zauważmy że jeżeli forma φ nie jest dodatnio określona, to mogą istnieć wektory v ortogonalne do siebie, tzn. spełniające $\varphi(v, v) = 0$, np. wektor $((c, 0, 0), 1)$ jest taki względem formy Lorentza.

Definicja 8.30. Niezerowy wektor ortogonalny do siebie nazywamy *izotropowym*.

To wydaje się istotną przeszkodą, gdybyśmy chcieli zastosować argument podobny do procesu Grama-Schmidta.

Na szczęście nie jest tak całkiem źle.

Stwierdzenie 8.31. Załóżmy że φ jest niezerową formą symetryczną. Wtedy istnieje niezerowy i nieizotropowy wektor.

Dowód. Ćwiczenie. □

Uwaga 8.32. Dla form niesymetrycznych to nie musi być prawdą, co pokazuje np. \det na \mathbf{R}^2 .

Definicja 8.33. Jeżeli $A \subseteq V$ (wyposażonej w ustaloną formę symetryczną φ), to jego *dopełnienie ortogonalne* A^\perp to zbiór wektorów prostopadłych do wszystkich elementów A :

$$A^\perp = \{v \in V \mid (\forall a \in A) \varphi(a, v) = 0\}.$$

Przykłady 8.34. • dopełnienie ortogonalne prostej w \mathbf{R}^3 (względem standardowego iloczynu skalarnego) to płaszczyzna ortogonalna do niej (i vice versa);

- ogólnie dopełnienie ortogonalne k -wymiarowej podprzestrzeni \mathbf{R}^n względem standardowego iloczynu skalarnego to $n - k$ -wymiarowa podprzestrzeń dopełnicza;
- dopełnienie ortogonalne $1^\perp \subseteq C([0, 1])$ względem $\varphi(f, g) = \int_0^1 f(t)g(t) dt$ składa się z funkcji spełniających $\int_0^1 f(t) dt = 0$.
- jeżeli V jest dowolną przestrzenią liniową i φ jest formą zerową, to dla każdego $A \subseteq V$ mamy $A^\perp = V$.

Uwaga 8.35. Dla dowolnych $A, B \subseteq V$:

- $A^\perp \leq V$,
- $A^\perp = (\text{Lin} A)^\perp$,
- $(A \cup B)^\perp = A^\perp \cap B^\perp$

Dowód. Ćwiczenie □

Fakt 8.36. Niech v_0 będzie dowolnym nieizotropowym wektorem. Wtedy dla $W = \text{Lin}(v_0)$ zachodzi $V = W \oplus W^\perp$.

Dowód. Z uwagi wynika że $W^\perp = \{v_0\}^\perp$.

Z tego że v_0 nie jest izotropowy łatwo wynika, że $W \cap W^\perp = \{0\}$, więc W, W^\perp są liniowo niezależne. Wystarczy więc pokazać, że $W + W^\perp = V$.

Weźmy dowolny wektor $v \in V$. Niech $w = \frac{\varphi(v, v_0)}{\varphi(v_0, v_0)} v_0$. Wtedy mamy:

$$\varphi(v - w, v_0) = \varphi(v, v_0) - \varphi\left(\frac{\varphi(v, v_0)}{\varphi(v_0, v_0)} v_0, v_0\right) = \varphi(v, v_0) - \underbrace{\frac{\varphi(v, v_0)}{\varphi(v_0, v_0)} \varphi(v_0, v_0)}_{\varphi(v, v_0)} = 0,$$

więc

$$v = \underbrace{w}_{\in W} + \underbrace{v - w}_{\in W^\perp},$$

co daje tezę. □

Twierdzenie 8.37 (twierdzenie Lagrange’a). • Niech φ będzie formą symetryczną na rzeczywistej przestrzeni liniowej V . Wtedy istnieje baza ortogonalna dla φ (a więc diagonalizująca ją); dokładniej, istnieje taka baza $B = (b_1, \dots, b_n)$ przestrzeni V , że dla $i \neq j$ zachodzi $\varphi(b_i, b_j) = 0$, a przy tym $\varphi(b_i, b_i) \in \{-1, 0, 1\}$; w szczególności w tej bazie

$$\varphi(v, w) = \sum_{i=1}^n d_i v_i w_i,$$

gdzie d_i to wartości na przekątnej, a v_i, w_i to współrzędne v, w w bazie B .

- Niech A będzie rzeczywistą macierzą symetryczną; wtedy istnieje taka macierz odwracalna Q że QAQ^\top jest macierzą diagonalną

Dowód. Część druga łatwo wynika z pierwszej.

Dowód części pierwszej: indukcja względem $\dim V$. Jeżeli $\dim V = 0$ lub jeżeli $\varphi = 0$, to teza jest oczywiście prawdziwa.

W przeciwnym wypadku istnieje nieizotropowy $v_0 \in V$ różny od 0. Wtedy wektor $b_1 = \frac{v_0}{\sqrt{|\varphi(v_0, v_0)|}}$ spełnia

$$\begin{aligned}\varphi(b_1, b_1) &= \varphi\left(\frac{v_0}{\sqrt{|\varphi(v_0, v_0)|}}, \frac{v_0}{\sqrt{|\varphi(v_0, v_0)|}}\right) \\ &= \frac{\varphi(v_0, v_0)}{\sqrt{|\varphi(v_0, v_0)|}^2} \\ &= \frac{\varphi(v_0, v_0)}{|\varphi(v_0, v_0)|} \\ &= \pm 1\end{aligned}$$

Rozważmy obcięcie φ do $W = b_1^\perp$.

Ponieważ $V = \text{Lin}(b_1) \oplus b_1^\perp$, wiemy że $\dim W = \dim V - 1$. Z założenia indukcyjnego możemy więc znaleźć b_2, b_3, \dots, b_n — bazę W taką że $\varphi(b_i, b_j) = 0$ dla $i \neq j$ i $\varphi(b_i, b_i) \in \{-1, 0, 1\}$.

Z drugiej strony oczywiście $\varphi(b_1, b_i) = 0$ dla $i \neq 1$, a ponieważ $V = \text{Lin}(b_1) \oplus W$, $B = (b_1, \dots, b_n)$ jest bazą. \square

Przykład 8.38. Chcemy zdiagonalizować formę symetryczną $\varphi(v, w) = v_x w_y + v_y w_x + 2v_x w_z + 2v_z w_x$ na \mathbf{R}^4 . Jej macierz (w standardowej bazie) to

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

Wybieramy dowolny $v \in \mathbf{R}^4$ taki że $\varphi(v, v) \neq 0$, np. $v = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$. Mamy

$\varphi(v, v) = 2$, więc bierzemy $v_1 = \frac{v}{\sqrt{2}}$.

Wśród wektorów v takich że $\varphi(v_1, v) = 0$ szukamy v takiego że $\varphi(v, v) \neq 0$. $\varphi(v_1, v) = 0$ oznacza $v_x + v_y + 2v_z = 0$. Na przykład $v = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$ działa:

$\varphi(v, v) = -2$, stąd $v_2 = \frac{v}{\sqrt{2}}$.

Wśród v takich że $\varphi(v_1, v) = \varphi(v_2, v) = 0$ szukamy v takiego że $\varphi(v, v) \neq 0$. Mamy więc układ równań:

$$\begin{cases} v_x + v_y + 2v_z = 0 \\ -v_x + v_y + 2v_z = 0 \end{cases}$$

stąd $v_x = 0$ (przez odjęcie stronami). Dla każdego takiego wektora zachodzi $\varphi(v, v) = 0$, więc nie znajdziemy rozwiązania: forma obcięta do tej podprzestrzeni jest zerowa.

Stąd możemy za v_3, v_4 wziąć dowolne liniowo niezależne wektory spełniające ten układ równań, np. $v_3 = \begin{pmatrix} 0 \\ 2 \\ -1 \\ 0 \end{pmatrix}, v_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$.

Stąd macierz φ w bazie $B = (v_1, v_2, v_3, v_4)$ to

$$m^{BB}(\varphi) = m_E^B(\text{id})^\top m^{EE}(\varphi) m_E^B(\text{id})$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 2 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^\top \begin{pmatrix} 0 & 1 & 2 & 0 \\ 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 2 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

i we współrzędnych $[v]_B = \begin{pmatrix} v'_x \\ v'_y \\ v'_z \\ v'_t \end{pmatrix}$ mamy

$$\varphi(v, w) = v'_x w'_x - v'_y w'_y.$$

Uwaga 8.39. Nad dowolnym ciałem takim że $1 + 1 \neq 0$ można w ten sam sposób pokazać, że dla symetrycznej φ zawsze istnieje baza B taka że $m^{BB}(\varphi)$ jest diagonalna, więc w pewnej bazie

$$\varphi(v, w) = \sum_{i=1}^n d_i v_i w_i,$$

(z tym że d_i mogą być różne od $-1, 0, 1$).

Uwaga 8.40. Z twierdzenia powyżej wynika łatwo (przez permutację bazy), że dla każdej symetrycznej φ (nad \mathbf{R}) mamy bazę B w której φ ma strukturę blokową:

$$m^{BB}(\varphi) = \begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0_r \end{pmatrix}.$$

pokażemy następujące twierdzenie:

Twierdzenie 8.41 (prawo bezwładności Sylwestera). *Liczby p, q, r występujące powyżej nie zależą od wyboru bazy B (względem której φ ma macierz tej postaci).*

Definicja 8.42. Trójkę (p, q, r) (lub czasami parę (p, q)) nazywamy *sygnaturą* formy symetrycznej φ .

Przykłady 8.43. • Standardowy iloczyn skalarny na \mathbf{R}^n ma sygnaturę $(n, 0, 0)$; podobnie, każdy iloczyn skalarny na n -wymiarowej przestrzeni ma sygnaturę $(n, 0, 0)$ (to wynika np. z istnienia bazy ortonormalnej).

- Forma z przykładu powyżej ma sygnaturę $(1, 1, 2)$.
- Forma Lorentza $x_1y_1 + x_2y_2 + x_3y_3 - c(ts)$ ma sygnaturę $(3, -1, 0)$.
- Forma na $\mathbf{R}_3[X]$ zadana wzorem $\varphi(P, Q) = P(1)Q'(-1) + P'(-1)Q(1)$ ma sygnaturę $(1, 1, 2)$: jej macierz w bazie $B = (\frac{X}{\sqrt{2}}, \frac{-X+2}{\sqrt{2}}, X^2 + 2X -$

$$3, X^3 - 3X + 2) \text{ to } m^{BB}(\varphi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \text{ (Wystarczy sprawdzić ze}$$

wartości tych wielomianów w 1 to odpowiednio $\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0$, a wartości ich pochodnych w -1 to $\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0, 0$.)

Definicja 8.44. *Podprzestrzeń zerowa* dla formy symetrycznej φ to zbiór wektorów zerujących φ , tzn.:

$$N = N_\varphi = \{v \in V \mid v^\perp = V\} = \{v \in V \mid (\forall w \in V) \varphi(v, w) = 0\}$$

Mówimy że φ jest *niezdegenerowana* jeżeli jej przestrzeń zerowa jest trywialna, tzn. $N_\varphi = \{0\}$.

Stwierdzenie 8.45. Niech $A = m^{BB}(\varphi)$ dla pewnej symetrycznej formy φ i bazy B . Wtedy:

1. N_φ składa się z rozwiązań równania jednorodnego $A[v]_B = 0$ (tzn. jest jądrem $F \in \text{End}(V)$ takiego że $m_B(F) = A$),
2. φ jest niezdegenerowana wtedy i tylko wtedy gdy A jest odwracalna.

Dowód. Druga część natychmiast wynika z pierwszej. Pierwsza część: założmy że $A[v]_B = 0$. Wtedy dla każdego w mamy

$$\varphi(v, w) = \varphi(w, v) = [w]_B^\top A[v]_B = [w]_B^\top 0 = 0.$$

Z drugiej strony jeżeli $v \in N_\varphi$, to dla wektora bazowego $b_i \in B$ mamy:

$$\varphi(b_i, v) = [b_i]_B^\top A[v]_B = e_i^\top A[v]_B,$$

ale ta ostatnia liczba to i -ta współrzędna $A[v]_B$, więc wtedy $A[v]_B = 0$. \square

Uwaga 8.46. Załóżmy że φ jest formą dwuliniową na przestrzeni V i $W \leq V$. Wtedy $\varphi|_W$, obcięcie φ do $W \times W$ jest formą dwuliniową na W . Ponadto symetryczność i dodatnia określoność (ale już nie niezdegenerowanie) φ pociąga to samo dla obcięcia.

Fakt 8.36 ma następujące uogólnienie.

Stwierdzenie 8.47. Niech $W \leq V$ ($\dim V < \infty$), a φ będzie formą symetryczną. Jeżeli $\varphi|_W$ jest niezdegenerowana, to $V = W \oplus W^\perp$.

Założmy że obcięcie $\varphi|_W$ jest niezdegenerowane (jako symetryczna forma dwuliniowa na W). Wtedy $V = W \oplus W^\perp$.

Dowód. Rozważmy odwzorowanie $\alpha: W \rightarrow W^*$ zadane wzorem $\alpha(w_1)(w) = \varphi(w_1, w)$. α jest oczywiście liniowe i $\ker \alpha = W \cap W^\perp = \{0\}$, więc α jest 1-1, a więc jest izomorfizmem (bo $\dim W < \infty$).

W szczególności dla każdego $v \in V$ istnieje taki $w_1 \in W$ że funkcjonały $\varphi(v, -)$ i $\varphi(w_1, -)$ na W są równe, czyli $\varphi(v, w) = \varphi(w_1, w)$, a więc $\varphi(v - w_1, w) = 0$ dla $w \in W$, co znaczy że $w_2 = v - w_1 \in W^\perp$.

Przedstawiliśmy tym samym v jako sumę $w_1 + w_2$, pokazując tym samym że $V = W + W^\perp$. Ponieważ oczywiście $W \cap W^\perp = \{0\}$, to pokazuje tezę. \square

dowód prawa bezwładności Sylwestera. Załóżmy że $B = (b_1, \dots, b_n)$ oraz $B' = (b'_1, \dots, b'_n)$ są bazami jak w Uwadze 8.40, dającymi liczby p, q, r, p', q', r' .

Zauważmy że r jest wymiarem jądra odwzorowania zadanego przez $m^{BB}(\varphi)$ w bazie B , czyli przestrzeni zerowej N_φ . Ta zaś nie zależy od wyboru bazy, więc $r = r'$

Ponieważ $p + q + r = n$, to wynika z tego natychmiast, że $p + q = p' + q'$. Pokażemy że układ $C = b_1, b_2, \dots, b_p, b'_{p'+1}, b'_{p'+2}, \dots, b'_n$ jest liniowo niezależny. To wystarczy, bo z tego będzie wynikało że $p \leq p'$, a więc (z symetrii) również $p' \leq p$, czyli $p = p'$

Założmy że $\alpha_1, \dots, \alpha_p, \beta_{p'+1}, \dots, \beta_n$ są takie że

$$v = \sum_{k=1}^p \alpha_k b_k = \sum_{j=p'+1}^n \beta_j b'_j.$$

Policzymy $\varphi(v, v)$ na dwa sposoby. Z lewej strony wynika że

$$\varphi(v, v) = \sum_{k=1}^p \alpha_k^2 \varphi(b_k, b_k) = \alpha_1^2 + \alpha_2^2 + \dots + \alpha_p^2.$$

Z prawej zaś

$$\varphi(v, v) = \sum_{j=p'+1}^n \beta_j \underbrace{\varphi(b'_j, b'_j)}_{\begin{cases} -1 & j \leq p' + q' \\ 0 & j > p' + q' \end{cases}} = -\beta_{p'+1}^2 - \beta_{p'+2}^2 - \dots - \beta_{p'+q'}^2.$$

Ponieważ α_k i β_j są liczbami rzeczywistymi, równość zachodzi tylko wtedy gdy wszystkie α_k i wszystkie β_j dla $j \leq p' + q'$ są zerowe. Podstawiając dostajemy:

$$v = 0 = \sum_{j=p'+q'+1}^n \beta_j b'_j,$$

więc z liniowej niezależności pozostałe β_j też są zerowe, co kończy dowód. \square

dowód kryterium Sylwestera. Załóżmy że macierz symetryczna A jest dodatnio określona. Chcemy pokazać że wiodące minory główne A są dodatnie. Oznaczmy przez $\varphi = \varphi_A$ stowarzyszoną z nią formę symetryczną.

Skoro A jest dodatnio określona, to istnieje baza B taka że $m^{BB}(\varphi) = I$, więc istnieje odwracalna Q taka że $QAQ^\top = I$, czyli $A = Q^{-1}Q^{-\top}$. Stąd $\det A = (\det Q)^{-2}$, więc $\det A > 0$.

k -ty minor główny A to macierz obcięcia $\varphi|_{V_k}$ dla $V_k = \text{Lin}(e_1, e_2, \dots, e_k)$. Ponieważ $\varphi|_{V_k}$ jest dodatnio określona (jako obcięcie dodatnio określonej formy), odpowiedni wyznacznik jest dodatni jak wyżej.

Założmy że macierz A jest taka że wiodące minory główne są dodatnie. Chcemy pokazać, że A jest dodatnio określona. Dowodzimy przez indukcję. Gdy A jest macierzą 1×1 , teza jest oczywista.

Założmy że teza zachodzi dla macierzy rozmiarów mniejszych niż n . To implikuje że obcięcie $\varphi = \varphi_A$ do V_{n-1} jest dodatnio określone i istnieje odwracalna $Q' \in M_{(n-1) \times (n-1)}(\mathbf{R})$ taka że $Q'A_{n-1}(Q')^\top = I_{n-1}$. Biorąc macierz blokową $Q = \begin{pmatrix} Q' & 0 \\ 0 & 1 \end{pmatrix}$ dostajemy

$$QAQ^\top = \begin{pmatrix} Q' & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A_{n-1} & * \\ * & * \end{pmatrix} \begin{pmatrix} (Q')^\top & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} I_{n-1} & v \\ v^\top & c \end{pmatrix} = A'$$

dla pewnej $c \in \mathbf{R}$ i $v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_{n-1} \end{pmatrix} \in \mathbf{R}^{n-1}$ (bo macierz QAQ^\top jest symetryczna!).

Przez elementarne operacje wierszowe możemy macierz A' sprowadzić do macierzy postaci $\begin{pmatrix} I & * \\ 0 & c' \end{pmatrix}$. Dokładniej, weźmy $E = \begin{pmatrix} I_{n-1} & 0 \\ -v^\top & 1 \end{pmatrix}$.

Wtedy

$$EA' = \begin{pmatrix} I_{n-1} & 0 \\ -v^\top & 1 \end{pmatrix} \begin{pmatrix} I_{n-1} & v \\ v^\top & c \end{pmatrix} = \begin{pmatrix} I_{n-1} & v \\ 0 & c' \end{pmatrix}:$$

dla wszystkich wierszy prócz ostatniego jest to zupełnie oczywiste. Wyliczmy i -tą współrzędną w ostatnim wierszu, gdzie $i < n$.

$$(-v^\top \quad 1) \underbrace{\begin{pmatrix} e_i \\ v_i \end{pmatrix}}_{e_i \in \mathbf{R}^{n-1}} = -v_i + v_i = 0.$$

Podobnie

$$EA'E^\top = \begin{pmatrix} I_{n-1} & v \\ 0 & c' \end{pmatrix} \begin{pmatrix} I_{n-1} & -v \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} I_{n-1} & 0 \\ 0 & c'' \end{pmatrix}$$

Łącznie dla $P = EQ$ mamy

$$PAP^\top = EQAQ^\top E^\top = \begin{pmatrix} I_{n-1} & 0 \\ 0 & c'' \end{pmatrix},$$

przy czym $c'' = \det(PAP^\top) = \det A \det P^2 > 0$ (bo P jest odwracalna). Stąd łatwo wynika że PAP^\top jest dodatnio określona, a więc A też. \square

8.4 Formy kwadratowe

Teraz K jest dowolnym ciałem.

Definicja 8.48. Funkcję $Q: V \rightarrow K$ nazywamy *formą kwadratową* jeżeli:

- jest jednorodna stopnia 2, to znaczy spełnia $Q(\alpha v) = \alpha^2 Q(v)$,
- funkcja $Q(v + w) - Q(v) - Q(w)$ jest formą dwuliniową na V

Przykłady 8.49. • $Q(x) = x^2$ jest formą kwadratową na $V = K$: jednorodność jest łatwa, funkcja $\varphi(x, y) = (x + y)^2 - x^2 - y^2 = xy$ jest dwuliniowa.

- $Q\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 x_2$ jest formą kwadratową na $V = K^2$: łatwo sprawdzić, że jest jednorodna, a funkcja $\varphi(x, y) = (x_1 + y_1)(x_2 + y_2) - x_1 x_2 - y_1 y_2 = x_1 y_2 + y_1 x_2$ jest dwuliniowa.
- Ogólnie: wielomian kwadratowy n zmiennych, bez wyrazów liniowych i wolnych, zadaje formę kwadratową na K^n .

Uwaga 8.50. Załóżmy że $\varphi(v, w)$ jest formą dwuliniową. Wtedy łatwo sprawdzić, że $Q_\varphi(v) = \varphi(v, v)$ jest formą kwadratową.

Mówimy że $\text{char } K \neq 2$ jeżeli w K jest prawdziwe $1 + 1 \neq 0$, tzn. 2^{-1} jest dobrze określone. Od teraz zakładamy, że tak jest.

Definicja 8.51. Jeżeli $\text{char } K \neq 2$, to dla formy kwadratowej Q definiujemy $\tilde{Q}(v, w) = \frac{Q(v+w) - Q(v) - Q(w)}{2}$.

Uwaga 8.52. \tilde{Q} zawsze jest symetryczną formą dwuliniową.

Uwaga 8.53. Mając dany wzór na Q w bazie $B = (b_1, \dots, b_n)$ łatwo wyznaczyć wzór \tilde{Q} : jeżeli

$$Q(v) = \sum_{i \leq j} \alpha_{ij} v_i v_j,$$

gdzie v_i to współrzędne v w bazie B , to współczynnik $\tilde{Q}(v, w)$ przy $v_i w_j$ (dla $i \neq j$) to $\frac{\alpha_{ij}}{2}$, a współczynnik przy $v_i w_i$ to α_{ii} . Istotnie:

$$\begin{aligned} \alpha_{ij}(v_i + w_i)(v_j + w_j) - \alpha_{ij}v_i v_j - \alpha_{ij}w_i w_j &= \\ &= \alpha_{ij}(v_i v_j + v_i w_j + w_i v_j + w_i w_j - v_i v_j - w_i w_j) = \alpha_{ij}(v_i w_j + v_j w_i); \end{aligned}$$

jeżeli $i = j$, to dzieląc przez 2 otrzymujemy $\alpha_{ii}v_i w_i$, a jeżeli $i \neq j$ — $\frac{\alpha_{ij}}{2}v_i w_j + \frac{\alpha_{ij}}{2}v_j w_i$.

Wniosek 8.54 (wzór polaryzacyjny). *Jeżeli $\text{char } K \neq 2$, to $Q \mapsto \tilde{Q}$, $\varphi \mapsto Q_\varphi$ zadają wzajemnie odwrotne bijekcje między formami kwadratowymi, a symetrycznymi formami dwuliniowymi.*

W szczególności jeżeli φ jest symetryczną formą dwuliniową, to mamy wzór:

$$\varphi(v, w) = \frac{1}{2}(\varphi(v + w, v + w) - \varphi(v, v) - \varphi(w, w)).$$

Dowód. Ćwiczenie (było). □

Uwaga 8.55. Czasami utożsamiamy formę kwadratową z \tilde{Q} i piszemy po prostu $Q(v, w)$

Przez utożsamienie form kwadratowych i symetrycznych dostajemy łatwo następujące fakty i definicje.

Definicja 8.56. ($\text{char } K \neq 2$) Macierz formy kwadratowej Q względem bazy B to $m^{BB}(Q) = m^{BB}(\tilde{Q})$.

($K = \mathbf{R}$) Q jest dodatnio określona gdy $Q(v) > 0$ dla $v \neq 0$ ($\iff m^{BB}(Q)$ jest dodatnio określona $\iff \tilde{Q}$ jest dodatnio określona).

Uwaga 8.57. Jeżeli $Q: V \rightarrow K$ jest formą kwadratową, to jej macierz względem każdej bazy jest symetryczna (bo \tilde{Q} jest, a macierze symetrycznych form dwuliniowych są symetryczne).

Wniosek 8.58. Jeżeli Q jest formą kwadratową nad K (charakterystyki $\neq 2$), to istnieje baza B taka że $m^{BB}(Q)$ jest macierzą diagonalną, tzn. w pewnej bazie Q przedstawia się jako

$$Q(v) = \sum_{i=1}^n \alpha_i v_i^2,$$

gdzie $[v]_B = \begin{pmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{pmatrix}$.

Jeżeli $K = \mathbf{R}$, to można ponadto założyć, że wartości $m^{BB}(\varphi)$ na przekątnej (czyli α_i) to $-1, 0, 1$. Liczba $-1, 0$ i 1 na przekątnej nie zależy od B .

Definicja 8.59. Sygnatura formy kwadratowej nad \mathbf{R} to trójka liczb jedynek, minus jedynek i zer na przekątnej w postaci z poprzedniego wniosku (jak dla form symetrycznych). (Jest dobrze określona z prawa Sylwestera.)

Przykład 8.60. Forma kwadratowa $Q(v) = 2v_x v_y + 4v_x v_z$ na \mathbf{R}^4 jest stowarzyszona z formą dwuliniową $\tilde{Q}(v, w) = v_x w_y + v_y w_x + 2v_x w_z + 2v_z w_x$ z wcze-

śniejszego przykładu. W bazie B o macierzy przejścia $\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 2 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

jej $m^{BB}(Q)$ jest diagonalna z 1, -1 i dwoma zerami na przekątnej, więc w tej bazie ma wzór

$$Q(v) = (v'_1)^2 - (v'_2)^2,$$

a jej sygnatura to 1, 1, 2.

Przykład 8.61. (Algorytm Lagrange'a.) Formy kwadratowe można diagonalizować tak jak w dowodzie twierdzenia Lagrange'a. Można też postępować w następujący sposób.

1. Weźmy $Q(v) = 3v_x^2 + 2v_x v_y + 4v_x v_z$ na \mathbf{R}^3 . Chcemy ją zdiagonalizować.

Korzystając z dodatniego współczynnika przy v_x^2 , piszemy

$$\begin{aligned}
 Q(v) &= 3(v_x^2 + \frac{2}{3}v_x v_y) + 4v_x v_z \\
 &= 3 \underbrace{(v_x + \frac{v_y}{3})^2}_{v_x'} - \frac{1}{3}v_y^2 + 4 \underbrace{v_x v_z}_{v_x' - \frac{v_y}{3}} \\
 &= 3(v_x')^2 - \frac{1}{3}v_y^2 + 4v_x' v_z - \frac{4}{3}v_y v_z \\
 &= 3((v_x')^2 + \frac{4}{3}v_x' v_z) - \frac{1}{3}(v_y^2 + 4v_y v_z) \\
 &= 3v_x'' - \frac{4}{3}v_z^2 - \frac{1}{3}(v_y')^2 + \frac{4}{3}v_z^2 = 3(v_x'')^2 - \frac{1}{3}(v_y')^2,
 \end{aligned}$$

gdzie $v_x'' = v_x' + \frac{2}{3}v_z = v_x + \frac{1}{3}v_y + \frac{2}{3}v_z$ i $v_y' = v_y + 2v_z$.

Stąd dla bazy B takiej że $m_B^E(\text{id}) = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ \frac{2}{3} & 2 & 1 \end{pmatrix}$ mamy $m^{BB}(Q) = \begin{pmatrix} 3 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & 0 \end{pmatrix}$, czyli

$$\begin{pmatrix} 3 & 1 & 2 \\ 1 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ \frac{2}{3} & 2 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ \frac{2}{3} & 2 & 1 \end{pmatrix}^T$$

równoważnie:

$$\begin{pmatrix} 3 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ \frac{2}{3} & 2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 3 & 1 & 2 \\ 1 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ \frac{2}{3} & 2 & 1 \end{pmatrix}^{-T}$$

2. Co gdy nie ma kwadratów? Weźmy $Q(v) = 2v_x v_y + 4v_x v_z$. Podstawmy

$v'_x = \frac{v_x + v_y}{2}, v'_y = \frac{v_x - v_y}{2}$. Wtedy $v_x v_y = (v'_x)^2 - (v'_y)^2$ i $v_x = v'_x + v'_y$, stąd

$$\begin{aligned} Q(v) &= 2v_x v_y + 4v_x v_z \\ &= 2(v'_x)^2 - 2(v'_y)^2 + 4v'_x v_z + 4v'_y v_z \\ &= 2((v'_x)^2 + 2v'_x v_z) - 2((v'_y)^2 - 2v'_y v_z) \\ &= 2(v''_x)^2 - v_z^2 - 2(v''_y)^2 + v_z^2 = 2(v''_x)^2 - 2(v''_y)^2, \end{aligned}$$

czyli

$$\begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ 1 & -1 & 1 \end{pmatrix}^T$$

3. ogólny algorytm:

- (a) tak długo jak jest jakiś kwadrat, któremu towarzyszy czynnik mieszany, zwijamy je razem do kwadratu;
- (b) jeżeli nie ma, a są jakieś wyrazy mieszane typu xy (bez x^2 ani y^2), to podstawiamy $x' = \frac{x+y}{2}, y' = \frac{x-y}{2}$, po czym wracamy do poprzedniego punktu;
- (c) jeżeli nie ma wyrazów mieszanych, to kończymy algorytm.

(To działa dla dowolnego ciała charakterystyki różnej od 2.¹⁾)

¹W charakterystyce 2 np. $Q(x, y) = xy$ się nie diagonalizuje.

Rozdział 9

Przestrzenie z iloczynem skalarnym

9.1 Przestrzenie euklidesowe

Uwaga 9.1. Forma dwuliniowa na \mathbf{R} jest dodatnio określona wtedy i tylko wtedy, gdy dopuszcza bazę ON.

Dowód. W jedną stronę: proces Grama-Schmidta. W drugą stronę wniosek jest oczywisty: dla bazy ON B mamy $\varphi(v, v) = [v]_B^T [v]_B > 0$ dla $v \neq 0$. \square

Definicja 9.2. *Przestrzeń euklidesowa* to skończenie wymiarowa przestrzeń liniowa V nad \mathbf{R} wraz z iloczynem skalarnym (tzn. dodatnio określoną, symetryczną formą dwuliniową).

W tej części wykładu (do odwołania) V jest przestrzenią euklidesową. (Ale większość faktów stosunkowo łatwo uogólnia się na rzeczywiste przestrzenie nieskończenie wymiarowe z iloczynem skalarnym.)

Iloczyn skalarny w przestrzeni euklidesowej oznaczamy $\langle -, - \rangle$ (tak samo jak standardowy iloczyn skalarny).

Definicja 9.3. *Długość* wektora v w przestrzeni euklidesowej to

$$|v| = \sqrt{\langle v, v \rangle}.$$

Uwaga 9.4. Z dodatniej określoności wynika, że $|v| = 0 \iff v = 0$.

Uwaga 9.5. $|-|$ spełnia $|\alpha v| = |\alpha| |v|$.

Uwaga 9.6. Z istnienia bazy ortonormalnej wynika, że n -wymiarowa przestrzeń euklidesowa zachowuje się pod każdym względem jak \mathbf{R}^n ze standardowym iloczynem skalarnym.

Dokładniej, jeżeli B jest bazą ortonormalną przestrzeni euklidesowej V , to dla każdego wektora $v_1, v_2 \in V$ mamy

$$\langle v_1, v_2 \rangle = [v_1]_B^\top m^{BB}(\langle -, - \rangle)[v_2]_B = [v_1]_B^\top I[v_2]_B = [v_1]_B^\top [v_2]_B,$$

czyli $[-]_B$ jest izomorfizmem $V \rightarrow \mathbf{R}^n$ zachowującym iloczyn skalarny.

Uwaga 9.7. Obcięcie iloczynu skalarnego do podprzestrzeni jest iloczynem skalarnym, więc liniowa podprzestrzeń przestrzeni euklidesowej jest przestrzenią euklidesową.

Za pomocą tej obserwacji możemy zdefiniować *nieskierowany* kąt między wektorami w dowolnej przestrzeni euklidesowej (a nawet nieskończenie wymiarowej przestrzeni z iloczynem skalarnym!): mając dane wektory u_1, u_2 , mówimy że kąt między nimi jest równy 0 jeżeli są liniowo zależne, a w przeciwnym wypadku (za pomocą obcięcia iloczynu skalarnego) możemy utożsamić $W = \text{Lin}(u_1, u_2)$ z \mathbf{R}^2 i tak zmierzyć kąt pomiędzy nimi.

Dokładniej, jeżeli w_1, w_2 jest bazą ortonormalną $\text{Lin}(u_1, u_2)$, to proste $\text{Lin}(w_1), \text{Lin}(w_2)$ utożsamiamy z osiami OX, OY , a kąt θ między wektorami u_1, u_2 można wyliczyć korzystając ze wzoru

$$\langle u_1, u_2 \rangle = |u_1||u_2| \cos \theta.$$

(uwaga: kąt skierowany ogólnie nie ma sensu, nawet w \mathbf{R}^3 — wymagałby on wybrania orientacji $\text{Lin}(u_1, u_2)$)

Definicja 9.8. Kąt między wektorami v, w w przestrzeni euklidesowej to

$$\angle(v, w) = \arccos \frac{\langle v, w \rangle}{|v||w|}.$$

Ćwiczenie 9.9. Przez ograniczenie się do dwuwymiarowych podprzestrzeni sprawdzamy, że dla dowolnych wektorów v, w zachodzą:

- nierówność Cauchy'ego-Schwarza: $\langle v, w \rangle \leq |v||w|$,
- nierówność trójkąta: $|v + w| \leq |v| + |w|$.

Zauważmy że jeżeli $W \leq V$ jest dowolną podprzestrzenią, to, ponieważ iloczyn skalarny jest nie zdegenerowany (bo jest dodatnio określony!), mamy $V = W \oplus W^\perp$. Stąd każdy $v \in V$ zapisuje się jednoznacznie w postaci $w + w'$, gdzie $w \in W$, $w' \in W^\perp$.

Definicja 9.10. Jeżeli $W \leq V$ jest podprzestrzenią, to rzut ortogonalny na W to funkcja $\pi = \pi_W =: V \rightarrow W$, $\pi(v) = w$, (w zdefiniowane jak wyżej).

Stwierdzenie 9.11. Załóżmy że b_1, b_2, \dots, b_k jest bazą ON $W \leq V$. Wtedy π_W jest zadane wzorem

$$\pi_W(v) = \sum_{j=1}^k \langle b_j, v \rangle b_j.$$

Uwaga 9.12. W szczególności to znaczy że $v - \pi_W(v)$ (zadane prawą stroną równania powyżej) jest ortogonalne do b_1, b_2, \dots, b_k . To tłumaczy sens geometryczny procesu Grama-Schmidta: konstruując bazę ON, odejmujemy od kolejnego wektora rzut na podprzestrzeń rozpiętą przez poprzednie.

Dowód. Ustalmy v , oznaczmy prawą stronę przez \bar{w} . Wtedy dla każdego l mamy

$$\langle b_l, \bar{w} \rangle = \sum_{j=1}^n \langle b_l, \langle b_j, v \rangle b_j \rangle = \langle b_l, v \rangle.$$

Odejmując stronami dostajemy $\langle b_l, \bar{w} - v \rangle = 0$. Stąd

$$w - v \in b_1^\perp \cap b_2^\perp \cap \dots \cap b_k^\perp = \text{Lin}(b_1, \dots, b_k)^\perp = W^\perp.$$

Z drugiej strony $\bar{w} \in W$, więc istotnie $\bar{w} = \pi_W(v)$. □

Biorąc $W = V$ dostajemy:

Wniosek 9.13. Jeżeli b_1, \dots, b_n jest bazą ortonormalną V , to dla każdego $v \in V$ zachodzi

$$v = \sum_{j=1}^n \langle b_j, v \rangle b_j.$$

Przykład 9.14. Chcemy wyznaczyć wzór na rzut wektora $v \in \mathbb{R}^3$ na płaszczyznę $W = \text{Lin} \left(\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right)$.

Najpierw wyznaczamy bazę ON W , np. przez proces Grama-Schmidta:

$$\left| \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right| = \sqrt{1+4+9} = \sqrt{14}, \text{ więc } b_1 = \frac{1}{\sqrt{14}} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

$$\begin{aligned} b'_2 &= \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} - \left\langle b_1, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\rangle b_1 \\ &= \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} - \underbrace{\frac{1-2+3}{14}}_{\frac{1}{7}} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \\ &= \frac{1}{7} \begin{pmatrix} 6 \\ -9 \\ 4 \end{pmatrix}, \end{aligned}$$

$$\text{więc } |b'_2| = \frac{\sqrt{36+81+16}}{7} = \frac{\sqrt{133}}{7}, \quad b_2 = \frac{1}{\sqrt{133}} \begin{pmatrix} 6 \\ -9 \\ 4 \end{pmatrix}.$$

Stąd dla $v = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ mamy

$$\begin{aligned} \pi_W(v) &= \langle b_1, v \rangle b_1 + \langle b_2, v \rangle b_2 \\ &= \frac{x+2y+3z}{14} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \frac{6x-9y+4z}{133} \begin{pmatrix} 6 \\ -9 \\ 4 \end{pmatrix} \\ &= \frac{1}{266} \begin{pmatrix} 19(x+2y+3z) + 12(6x-9y+4z) \\ 38(x+2y+3z) - 18(6x-9y+4z) \\ 57(x+2y+3z) + 8(6x-9y+4z) \end{pmatrix} \\ &= \frac{1}{266} \begin{pmatrix} 91x - 70y + 105z \\ -70x + 238y + 42z \\ 105x + 42y + 203z \end{pmatrix} \end{aligned}$$

Mamy też następujący wzór (uogólniający twierdzenie Pitagorasa).

Wniosek 9.15 (tożsamość Parsevala). *Jeżeli b_1, \dots, b_n jest bazą ortonormalną V , to dla każdego $v \in V$ zachodzi*

$$|v|^2 = \sum_{j=1}^n \langle b_j, v \rangle^2$$

Dowód.

$$|v|^2 = \langle v, v \rangle = \left\langle \sum_{j=1}^n \langle b_j, v \rangle b_j, \sum_{k=1}^n \langle b_k, v \rangle b_k \right\rangle = \sum_{j,k} \langle b_j, v \rangle \langle b_k, v \rangle \underbrace{\langle b_j, b_k \rangle}_{\begin{cases} 1 & j=k \\ 0 & j \neq k \end{cases}} = \sum_k \langle b_k, v \rangle^2.$$

□

Wniosek 9.16 (nierówność Bessela). *Jeżeli b_1, b_2, \dots, b_k jest układem ON (niekoniecznie bazą), to dla każdego $v \in V$ zachodzi nierówność:*

$$|v|^2 \geq \sum_{j=1}^k \langle b_j, v \rangle^2$$

Ćwiczenie 9.17. Każdy układ ON w przestrzeni euklidesowej rozszerza się do bazy ON.

Wniosek 9.18. *Dla każdego $W \leq V$ i $v \in V$ zachodzi $|\pi_W(v)| \leq |v|$ (równość zachodzi tylko gdy $v \in W$).*

Dowód. Ćwiczenie

□

Wniosek 9.19. *Jeżeli $W \leq V$, i $v \in V$, to punkt $\pi_W(v)$ jest najbliższy v spośród punktów w W .*

Dowód. Ćwiczenie.

□

9.2 Formy hermitowskie

Od teraz pracujemy nad $K = \mathbb{C}$.

Chcielibyśmy w tej sytuacji również mieć pojęcie długości wektora. Zwykły wzór na iloczyn skalarny niestety nie jest dobry, bo nie jest dodatnio określony, np. dla $v = i \in \mathbb{C}^1$ mamy $v^\top v = ii = -1$.

Przestrzeń \mathbf{C}^n możemy oczywiście utożsamić z \mathbf{R}^{2n} i wtedy długość zespolonego wektora $\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$, gdzie $z_j = a_j + ib_j$ to (z twierdzenia Pitagorasa albo tożsamości Parsevala):

$$\sqrt{a_1^2 + b_1^2 + a_2^2 + b_2^2 + \dots + a_n^2 + b_n^2} = \sqrt{\bar{z}_1 z_1 + \bar{z}_2 z_2 + \dots + \bar{z}_n z_n}.$$

To sugeruje, że zamiast $v^\top w$, dla wektorów w \mathbf{C}^n lepszy byłby wzór:

$$\langle v, w \rangle = \bar{v}^\top w.$$

Funkcja określona tym wzorem jest dodatnio określona:

$$\langle v, v \rangle > 0, \text{ gdy } v \neq 0.$$

Ponadto wzór ten zgadza się ze zwykłym iloczynem skalarnym dla rzeczywistych wektorów.

$\langle -, - \rangle$ zadany powyższym wzorem nazywamy *standardowym (hermitowskim) iloczynem skalarnym* na \mathbf{C}^n . Oprócz dodatniej określoności ma on następujące własności:

- liniowość w drugiej zmiennej:

$$\langle v, w_1 + w_2 \rangle = \langle v, w_1 \rangle + \langle v, w_2 \rangle \quad \langle v, \alpha w \rangle = \alpha \langle v, w \rangle.$$

- antyliniowość w pierwszej zmiennej:¹

$$\langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle \quad \langle \alpha v, w \rangle = \bar{\alpha} \langle v, w \rangle.$$

- hermitowska symetryczność:

$$\langle v, w \rangle = \overline{\langle w, v \rangle}.$$

Możemy więc odzyskać dodatnią określoność kosztem lekkiego osłabienia liniowości i symetryczności; hermitowski iloczyn skalarny jest właściwy, jeżeli chcemy rozważać długości wektorów nad \mathbf{C} .

¹Czasami przyjmuje się odwrotnie — liniowo w pierwszej, antyliniowo w drugiej.

Definicja 9.20. Forma hermitowska na V nad \mathbf{C} to funkcja $\varphi: V \times V \rightarrow \mathbf{C}$ która jest półtoraliniowa (liniowa w drugim, a antyliniowa w pierwszym argumencie) i hermitowsko symetryczna.

Uwaga 9.21. $\varphi: V \times V \rightarrow \mathbf{C}$ jest formą hermitowską wtedy i tylko wtedy gdy jest hermitowsko symetryczna i liniowa w drugim argumencie:

$$\varphi(\alpha v, w) = \overline{\varphi(w, \alpha v)} = \overline{\bar{\alpha} \varphi(w, v)} = \bar{\alpha} \varphi(v, w).$$

Formy hermitowskie nad \mathbf{C} odpowiadają formom symetrycznym nad \mathbf{R} .²

Przykład 9.22. Jeżeli φ jest formą symetryczną na rzeczywistej przestrzeni liniowej V , to mamy formę hermitowską $\varphi_{\mathbf{C}}$ na $V_{\mathbf{C}}$ (kompleksyfikację φ) zadaną wzorem (dla rzeczywistych $v_1, v_2, w_1, w_2 \in V$)

$$\varphi_{\mathbf{C}}(v_1 + iv_2, w_1 + iw_2) = \varphi(v_1, w_1) + i\varphi(v_1, w_2) - i\varphi(v_2, w_1) + \varphi(v_2, w_2) :$$

hermitowska symetryczność jest oczywista; liniowość w drugim argumencie:

$$\begin{aligned} \varphi_{\mathbf{C}}(v_1 + iv_2, \underbrace{(a + bi)(w_1 + iw_2)}_{aw_1 - bw_2 + i(bw_1 + aw_2)}) &= \\ &= \varphi(v_1, aw_1 - bw_2) + i\varphi(v_1, bw_1 + aw_2) - i\varphi(v_2, aw_1 - bw_2) + \varphi(v_2, bw_1 + aw_2) = \\ &= \varphi(v_1, w_1)(a + ib) + \varphi(v_1, w_2)\underbrace{(-b + ai)}_{i(a + bi)} + \varphi(v_2, w_1)\underbrace{(-ia + b)}_{-i(a + bi)} + \varphi(v_2, w_2)(ib + a) = \\ &= (a + bi)\varphi_{\mathbf{C}}(v_1 + iv_2, w_1 + iw_2) \end{aligned}$$

Jeżeli $\varphi = \langle -, - \rangle$ jest dodatnio określona, to łatwo sprawdzić, że $\langle -, - \rangle_{\mathbf{C}}$ też:

$$\langle v_1 + iv_2, v_1 + iv_2 \rangle_{\mathbf{C}} = \langle v_1, v_1 \rangle + i\langle v_1, v_2 \rangle - i\langle v_2, v_1 \rangle + \langle v_2, v_2 \rangle = \langle v_1, v_1 \rangle + \langle v_2, v_2 \rangle.$$

Przykłady 9.23. • $\varphi(z, w) = zw$ nie jest formą hermitowską na \mathbf{C} , bo nie jest półtoraliniowa; ogólnie żadna niezerowa forma dwuliniowa nie jest hermitowska, bo nie jest półtoraliniowa: jeżeli $\varphi(v, w) \neq 0$, to

$$\varphi(iv, w) = i\varphi(v, w) \neq -i\varphi(v, w).$$

²Rozważa się też prawdziwe formy symetryczne na zespolonych przestrzeniach liniowych, jak również niekoniecznie hermitowskie formy półtoraliniowe, ale nie będziemy się tym zajmować na tym wykładzie.

- $\varphi(z, w) = i\bar{z}w$ nie jest formą hermitowską na \mathbb{C} , bo $\varphi(1, 1) = i \neq \overline{\varphi(1, 1)}$.
- $\varphi(x_1 + ix_2, y_2 + iy_2) = x_1y_1$ nie jest formą hermitowską na \mathbb{C} , bo nie jest półtoraliniowa: $\varphi(1, i) = 0 \neq i = i\varphi(1, 1)$.
- $\varphi\left(\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}\right) = i\bar{z}_2w_1 - i\bar{z}_1w_2$ jest formą hermitowską na \mathbb{C}^2 : półtoraliniowość jest łatwa, hermitowska symetryczność:

$$\overline{i\bar{z}_2w_1 - i\bar{z}_1w_2} = -i\bar{z}_2\bar{w}_1 + i\bar{z}_1\bar{w}_2 = i\bar{w}_2z_1 - i\bar{w}_1z_2.$$

Uwaga 9.24. Ogólnie, jeżeli φ jest formą hermitowską, to dla każdego v liczba $\varphi(v, v)$ jest rzeczywista, bo z hermitowskiej symetrii $\varphi(v, v) = \overline{\varphi(v, v)}$.

Definicja 9.25. Jeżeli φ jest formą hermitowską, a $B = (b_1, \dots, b_n)$ jest bazą V , to macierz φ względem B definiujemy analogicznie, jak dla form dwulinio-
wych: ij -ty współczynnik $m^{BB}(\varphi)$ to $\varphi(b_i, b_j)$.

Przykład 9.26. Jeżeli φ jest formą symetryczną na rzeczywistej V i B jest bazą V , to $m^{BB}(\varphi) = m^{BB}(\varphi_{\mathbb{C}})$ (przypomnienie: B jest bazą $V_{\mathbb{C}}$!).

Przykład 9.27. Forma hermitowska $\varphi\left(\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}\right) = i\bar{z}_2w_1 - i\bar{z}_1w_2$ na \mathbb{C}^2 ma w standardowej bazie macierz $m^{EE}(\varphi) = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, więc nie jest kompleksyfikacją żadnej rzeczywistej formy symetrycznej.

Fakt 9.28. Jeżeli φ jest formą hermitowską na V o bazie B , to zachodzi wzór:

$$\varphi(v, w) = \overline{[v]_B}^{\top} m^{BB}(\varphi) [w]_B.$$

(Dowód analogiczny jak dla form dwulinio-
wych.)

Fakt 9.29. Dla dowolnej φ i B macierz $A = m^{BB}(\varphi)$ spełnia wzór $A = \bar{A}^{\top}$.

Dowód. ij -ta współrzędna \bar{A}^{\top} to $\overline{\varphi(b_j, b_i)}$, a ij -ta współrzędna A to $\varphi(b_i, b_j)$.
Teza wynika z hermitowskiej symetrii φ . \square

Definicja 9.30. Jeżeli $A \in M_{n \times n}(\mathbb{C})$, to macierz sprzężona do A to $A^* := \bar{A}^{\top}$.

Mówimy że A jest samosprzężona lub hermitowska jeżeli $A = A^*$.

Uwaga 9.31. Tu jest pozorny konflikt oznaczeń między sprzężeniem a odwzorowaniem dualnym. Okazuje się jednak, że w pewnym sensie macierz A^* to macierz F_A^* (odwzorowania dualnego do F_A).

Powyższy wzór przyjmuje zatem postać

$$\varphi(v, w) = [v]_B^* m^{BB}(\varphi) [w]_B.$$

Hermitowskość macierzy implikuje, że współczynniki na przekątnej są rzeczywiste! Macierz hermitowska ma zatem postać:

$$\begin{pmatrix} r_1 & a_{12} & a_{13} & \cdots & a_{1n} \\ \bar{a}_{12} & r_2 & a_{23} & & \\ \bar{a}_{13} & \bar{a}_{23} & r_3 & & \\ \vdots & & & \ddots & \\ \bar{a}_{1n} & & & & r_n \end{pmatrix}$$

gdzie r_j są rzeczywiste, a a_{kj} są zespolone.

Przykład 9.32. Hermitowskie macierze diagonalne to dokładnie te, które mają rzeczywiste wyrazy na przekątnej. Ogólnie: zespolona macierz symetryczna jest hermitowska wtedy i tylko wtedy gdy jest rzeczywista.

Uwaga 9.33. Łatwo sprawdzić tożsamości:

- $(A + B)^* = A^* + B^*$,
- $(AB)^* = B^*A^*$,
- $(A^*)^{-1} = (A^{-1})^*$ (w szczególności sprzężenie macierzy odwracalnej jest odwracalne),
- $A^{**} = A$.

Uwaga 9.34. Jeżeli $B = (b_1, \dots, b_n)$ jest bazą V , to $\varphi \mapsto m^{BB}(\varphi)$ zadaje bijekcję między formami hermitowskimi na V a macierzami hermitowskimi $n \times n$.

Uwaga 9.35. Rzeczywista macierz jest samosprężona wtedy i tylko wtedy gdy jest symetryczna.

Jeżeli B, C są różnymi bazami, to ponieważ mamy $[v]_C = m_C^B(\text{id})[v]_B$, podobnie jak w przypadku form symetrycznych, dostajemy wzór na zmianę bazy.

Fakt 9.36. $m^{CC}(\varphi) = m_B^C(\text{id})^* m^{BB}(\varphi) m_B^C(\text{id})$.

Dowód.

$$\begin{aligned}\varphi(v, w) &= [v]_B^* m^{BB}(\varphi) [w]_B \\ &= (m_B^C(\text{id})[v]_C)^* m^{BB}(\varphi) (m_B^C(\text{id})[w]_C) \\ &= [v]_C^* m_B^C(\text{id})^* m^{BB}(\varphi) m_B^C(\text{id}) [w]_C. \quad \square\end{aligned}$$

Przykład 9.37. Forma $\varphi(v, w) = i\bar{v}_2 w_1 - i\bar{v}_1 w_2$ w bazie $B = \begin{pmatrix} i \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ i \end{pmatrix}$ ma macierz

$$\begin{aligned}m^{BB}(\varphi) &= m_E^B(\text{id})^* m^{EE}(\varphi) m_E^B(\text{id}) = \begin{pmatrix} -i & 0 \\ 1 & -i \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} i & 1 \\ 0 & i \end{pmatrix} = \\ &= \begin{pmatrix} -i & 0 \\ 1 & -i \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & i \end{pmatrix} = \begin{pmatrix} 0 & -i \\ i & 2 \end{pmatrix}.\end{aligned}$$

Faktycznie, na przykład $\varphi\left(\begin{pmatrix} 1 \\ i \end{pmatrix}, \begin{pmatrix} 1 \\ i \end{pmatrix}\right) = i \cdot (-i) \cdot 1 - i \cdot 1 \cdot i = 1 + 1 = 2$.

Wniosek 9.38. Hermitowskie $A_1, A_2 \in M_{n \times n}(\mathbf{C})$ zadają tę samą formę hermitowską (może w różnych bazach) wtedy i tylko wtedy gdy istnieje odwracalna $P \in M_{n \times n}(\mathbf{C})$ taka że

$$A_1 = P^* A_2 P.$$

Definicja 9.39. Mówimy że $A \in M_{n \times n}(\mathbf{C})$ jest *unitarna* jeżeli $AA^* = I$ (czyli A jest odwracalna i $A^{-1} = A^*$).

Jeżeli $A \in M_{n \times n}(\mathbf{R})$ jest unitarna (równoważnie: spełnia $AA^T = I$), to mówimy że jest *ortogonalna*.

Uwaga 9.40. A jest unitarna/ortogonalna wtedy i tylko wtedy gdy A^*/A^T jest unitarna/ortogonalna. (Bo $A^{**} = A$.)

Ćwiczenie 9.41. Dla dowolnej zespolonej/rzeczywistej macierzy kwadratowej A następujące warunki są równoważne:

- A jest unitarna/ortogonalna,
- wiersze A tworzą układ ON (względem standardowego iloczynu skalarnego),

- kolumny A tworzą układ ON.

Ze wzoru na zmianę bazy dostajemy też następujący wniosek:

Wniosek 9.42. *Jeżeli B jest bazą \mathbf{C}^n , a $\langle -, - \rangle$ jest standardowym iloczynem skalarnym, to zachodzi równoważność:*

$$m^{BB}(\langle -, - \rangle) = I \iff m_E^B(\text{id}) \text{ jest unitarna.}$$

Uwaga 9.43. Za jakiś czas sprawdzimy, że macierze unitarne/ortogonalne to dokładnie macierze liniowych izometrii. W szczególności moduły ich wartości własnych są zawsze równe 1.

Ortogonalność, znormalizowanie i ortonormalność względem form hermitowskich definiujemy tak samo jak dla form symetrycznych.

Definicja 9.44. Niech φ będzie formą hermitowską.

- v, w są ortogonalne, $v \perp w$ (względem φ) gdy $\varphi(v, w) = 0$ (to jest symetryczna relacja, z hermitowskiej symetrii),
- v jest znormalizowany/jednostkowy (względem φ) gdy $\varphi(v, v) = 1$,
- układ wektorów jest ortogonalny gdy jego wektory są parami ortogonalne, a ortonormalny gdy są też znormalizowane,
- φ jest dodatnio określona gdy $\varphi(v, v) > 0$ dla $v \neq 0$,
- φ jest niezdegenerowana, gdy nie istnieje $v \neq 0$ taki że dla każdego w mamy $\varphi(v, w) = 0$,
- $W^\perp = \{v \in V \mid (\forall w \in W) v \perp w\}$.
- itd.

Ćwiczenie 9.45. Ortogonalny układ wektorów jest liniowo niezależny.

Twierdzenie 9.46. V ma bazę ON względem φ wtedy i tylko wtedy gdy φ jest dodatnio określona.

Dowód. Ćwiczenie (proces Grama-Schmidta działa). □

Stwierdzenie 9.47. *Jeżeli φ jest formą hermitowską i jej obcięcie do $W \leq V$ jest niezdegenerowane, to $V = W \oplus W^\perp$.*

Dowód. Ćwiczenie (dowód jak dla form symetrycznych). \square

Stwierdzenie 9.48 (wzór polaryzacyjny dla form hermitowskich). *Założmy że φ jest formą hermitowską. Oznaczmy $Q(v) = Q_\varphi(v) := \varphi(v, v)$. Wtedy zachodzi wzór:*

$$\varphi(v, w) = \frac{Q(v+w) - iQ(v+iw) - (1-i)(Q(v) + Q(w))}{2}.$$

Dowód. Zauważmy że

$$\begin{aligned}\varphi(v+w, v+w) &= \varphi(v+w, v) + \varphi(v+w, w) \\ &= \varphi(v, v) + \varphi(w, w) + \varphi(v, w) + \underbrace{\varphi(w, v)}_{\overline{\varphi(v, w)}} \\ &= \varphi(v, v) + \varphi(w, w) + 2\operatorname{Re} \varphi(v, w)\end{aligned}$$

$$\text{Stąd } \operatorname{Re}(\varphi(v, w)) = \frac{Q(v+w) - Q(v) - Q(w)}{2}.$$

Z drugiej strony, podstawiając $w = iw$ otrzymujemy:

$$\operatorname{Im}(\varphi(v, w)) = -\operatorname{Re}(i\varphi(v, w)) = -\operatorname{Re}(\varphi(v, iw)) = \frac{Q(v+iw) - Q(v) - \overbrace{Q(iw)}^{Q(w)}}{2}.$$

Stąd

$$\varphi(v, w) = \operatorname{Re}(\varphi(v, w)) + i\operatorname{Im}(\varphi(v, w)) = \frac{Q(v+w) - iQ(v+iw) - (1-i)(Q(v) + Q(w))}{2}.$$

\square

Ćwiczenie 9.49 (inna postać wzoru polaryzacyjnego). Dla form symetrycznych zachodzi następujący wzór (gdy $1+1 \neq 0$):

$$\varphi(v, w) = \frac{Q(v+w) - Q(v-w)}{4}.$$

Dla form hermitowskich zachodzi zaś wzór:

$$\varphi(v, w) = \frac{Q(v+w) + iQ(v+iw) - Q(v-w) - iQ(v-iw)}{4} = \frac{\sum_{k=0}^3 i^k Q(v+i^k w)}{4}$$

Wskazówka: zauważ że $\varphi(v, w) = \frac{\varphi(v, w) - \varphi(v, -w)}{2}$ i $Q(w) = Q(-w)$, skorzystaj z uzyskanych wcześniej wzorów.

9.3 Przestrzenie unitarne

Definicja 9.50. *Iloczyn skalarny* na zespolonej przestrzeni liniowej to dodatkowo określona forma hermitowska.

Przestrzeń unitarna to skończenie wymiarowa przestrzeń liniowa nad \mathbb{C} wyposażona w iloczyn skalarny.

Uwaga 9.51. Dla przestrzeni unitarnych zachodzą podobne fakty jak dla przestrzeni euklidesowych, z analogicznymi dowodami. Na przykład:

- rzut ortogonalny na $W = \text{Lin}(\overbrace{b_1, \dots, b_k}^{\text{układ ON}})$ jest zadany wzorem

$$\pi_W(v) = \sum_{j=1}^k \langle b_j, v \rangle b_j,$$

(Uwaga: tu kolejność argumentów w $\langle b_j, v \rangle$ jest ważna, aby to było liniowe względem v !)

- zachodzi tożsamość Parsewala

$$\langle v, v \rangle = \sum_{j=1}^n |\langle b_j, v \rangle|^2$$

(uwaga na moduły po prawej stronie!)

- zachodzi nierówność Bessela (też z modułami)
- zachodzą nierówność trójkąta i nierówność Cauchy'ego-Schwarza (też z modułem)

Uwaga 9.52. Kompleksyfikacja $V_{\mathbb{C}}$ przestrzeni euklidesowej V zawsze jest przestrzenią unitarną (ze skompleksyfikowanym iloczynem skalarnym) i bazy ON V są bazami ON $V_{\mathbb{C}}$.

Wszystkie fakty mają zazwyczaj zupełnie analogiczne dowody w przypadku rzeczywistym euklidesowym i zespolonym unitarnym, większość będzie miała dowody pisane pod ten drugi.

Ćwiczenie 9.53. Każdy ON układ wektorów w V rozszerza się do bazy ON.

Uwaga 9.54. Jeżeli V jest przestrzenią euklidesową, to mamy naturalny izomorfizm $V \cong V^*$ zadany przez $\varphi: v \mapsto (w \mapsto \langle v, w \rangle)$. Podobnie każda niezdegenerowana forma dwuliniowa zadaje izomorfizm $V \rightarrow V^*$.

Podobnie, jeżeli V jest przestrzenią unitarną, to mamy naturalny *anty-izomorfizm* $V \cong V^*$ zadany tym samym wzorem: odwzorowanie to nie jest już liniowe: spełnia za to warunek $\varphi(\alpha v) = \bar{\alpha} \varphi(v)$.

Rozdział 10

Rozkład spektralny i singularny

Od teraz (do końca semestru) V to przestrzeń unitarna/euklidesowa, a jej iloczyn skalarny oznaczamy przez $\langle -, - \rangle$. (Można myśleć, że $V = \mathbf{C}^n / \mathbf{R}^n$ ze standardowym zespolonym/rzeczywistym iloczynem skalarnym.)

Co więcej, będziemy pracowali w zasadzie wyłącznie z bazami ortonormalnymi.

Uwaga 10.1. Baza B jest ortonormalna wtedy i tylko wtedy gdy $m^{BB}(\langle -, - \rangle) = I$.

Stąd (i wzoru na zmianę bazy) łatwo wynika, że jeżeli B jest bazą ortonormalną, to baza C jest ON \iff macierz $m_C^B(\text{id})$ jest unitarna/ortogonalna.

Z drugiej strony jeżeli U jest macierzą unitarną/ortogonalną, to jest odwracalna, więc istnieje taka baza C że $m_C^B(\text{id}) = U$ i z powyższego wnioskuje się, że C musi być ortonormalna.

Rozważamy endomorfizm $F: V \rightarrow V$. Jeżeli B, C są bazami ortonormalnymi, to z powyższej uwagi dla $P = m_B^C(\text{id})$ mamy $P^{-1} = P^*$, więc wzór na zmianę bazy przybiera następującą postać:

$$m_C(F) = P^{-1} m_B(F) P = P^* m_B(F) P.$$

(Czyli taką samą, jak wzór na zmianę bazy formy hermitowskiej/rzeczywistej symetrycznej!)

Stwierdzenie 10.2. Jeżeli $F \in \text{End}(V)$ jest dowolny, to dla $G \in \text{End}(V)$ następujące warunki są równoważne:

1. $(\forall v, w) \langle v, F(w) \rangle = \langle G(v), w \rangle$

$$2. (\forall v, w) \langle F(v), w \rangle = \langle v, G(w) \rangle$$

3. dla pewnej (równoważnie: każdej) bazy ON B zachodzi $m_B(G) = m_B(F)^*$.

Dowód. Jeżeli B, C są różnymi bazami ON i $m_B(G) = m_B(F)^*$, to $m_C(G) = Pm_B(G)P^* = Pm_B(F)^*P^* = (Pm_B(F)P^*)^*$ dla $P = m_C^B(\text{id})$, stąd warunek (3) nie zależy od wyboru bazy.

Ustalmy dowolną bazę ON B . Wtedy dla każdych v i w mamy $\langle v, w \rangle = [v]_B^*[w]_B$, więc

$$\langle v, F(w) \rangle = [v]_B^*(m_B(F)[w]_B) = (m_B(F)^*[v]_B)^*[w]_B$$

i

$$\langle F(v), w \rangle = (m_B(F)[v]_B)^*[w]_B = [v]_B^*(m_B(F)^*[w]_B).$$

Stąd łatwo wynika implikacja (3) do (1,2). Z drugiej strony jeżeli $m_B(G) \neq m_B(F)^*$, to podobny rachunek pokazuje, że G nie może spełniać żadnego z warunków (1) i (2). \square

Definicja 10.3. G jak powyżej mówimy że jest *sprzężony* do F i oznaczamy F^* .

Uwaga 10.4. • to tylko pozornie kłóci z definicją odwzorowania dualnego $F^*: V^* \rightarrow V^*$, bo iloczyn skalarny pozwala utożsamić V i V^* (szczegóły: ćwiczenie na liście 13)

- ze stwierdzenia powyżej F^* istnieje i jest jedyny (wystarczy wybrać bazę ON B i wziąć przekształcenie o macierzy $m_B(F)^*$).

Stwierdzenie 10.5. Niech $F \in \text{End}(V)$ i niech $A = m_B(F)$ dla pewnej ON bazy B .

1. A jest hermitowska/symetryczna wtedy i tylko wtedy gdy zachodzi wzór:

$$\langle F(v), w \rangle = \langle v, F(w) \rangle$$

2. A jest unitarna/ortogonalna wtedy i tylko wtedy gdy zachodzi wzór:

$$\langle v, w \rangle = \langle F(v), F(w) \rangle.$$

Definicja 10.6. Mówimy że $F \in \text{End}(V)$ jest *samosprzężony* (czasami: *hermitowski* lub *symetryczny*) gdy spełnia $F = F^*$, lub równoważnie $\langle Fv, w \rangle = \langle v, Fw \rangle$.

Mówimy że F jest *unitarny* (w przypadku zespolonym) lub *ortogonalny* (w przypadku rzeczywistym), gdy spełnia $F^* = F^{-1}$, lub równoważnie $\langle v, w \rangle = \langle Fv, Fw \rangle$.

Dowód. Pierwsza część wynika łatwo z poprzedniego stwierdzenia: podany warunek oznacza, że $F = F^*$, czyli $m_B(F) = m_B(F^*) = m_B(F)^*$.

Drugi punkt dowodzimy podobnie: równość $\langle v, w \rangle = \langle F(v), F(w) \rangle$ jest równoważna tożsamości $[v]_B^* A^* A [w]_B = [v]_B^* [w]_B$, co łatwo daje $A^* A = I$. \square

Przykład 10.7. Jeżeli P jest rzutem ortogonalnym na $\text{Lin}(\overbrace{b_1, \dots, b_k}^{\text{układ ON}})$, to jest samosprzężony (ale prawie nigdy nie jest ortogonalny, chyba że $P = \text{id} \dots$): weźmy dowolne $v, w \in V$. Wtedy

$$\begin{aligned} \langle P(v), w \rangle &= \left\langle \sum_{j=1}^k \langle b_j, v \rangle b_j, w \right\rangle = \sum_{j=1}^k \overline{\langle b_j, v \rangle} \langle b_j, w \rangle = \sum_{j=1}^k \langle v, b_j \rangle \langle b_j, w \rangle = \\ &= \langle v, \sum_{j=1}^k \langle b_j, w \rangle b_j \rangle = \langle v, P(w) \rangle \end{aligned}$$

Można powiedzieć więcej: jeżeli $P \in \text{End}(V)$ jest rzutem (tzn. spełnia $P = P \circ P$), to P jest rzutem ortogonalnym (na swój obraz) wtedy i tylko wtedy gdy jest samosprzężony.

Stwierdzenie 10.8. Jeżeli F jest nieujemnie określony na przestrzeni unitarnej, tzn. $\langle v, F(v) \rangle \geq 0$ dla każdego v (w szczególności to jest rzeczywiste!), to jest samosprzężony. (W przypadku rzeczywistym to nie jest prawda, co pokazuje np. obrót o kąt nie większy niż $\pi/2$.)

Dowód. Załóżmy że F jest nieujemnie określony, weźmy dowolne v, w . Rozważmy $\alpha v + w$. Wtedy:

$$\begin{aligned} 0 \leq \langle \alpha v + w, F(\alpha v + w) \rangle &= \bar{\alpha} \alpha \langle v, F(v) \rangle + \langle w, F(w) \rangle + \bar{\alpha} \langle v, F(w) \rangle + \alpha \langle w, F(v) \rangle \\ &= |\alpha|^2 \langle v, F(v) \rangle + \langle w, F(w) \rangle + \overline{\alpha \langle F(w), v \rangle} + \alpha \langle w, F(v) \rangle. \end{aligned}$$

Stąd $\overline{\alpha \langle F(w), v \rangle} + \alpha \langle w, F(v) \rangle$ jest liczbą rzeczywistą. Dla $\alpha = 1$ wnioskujemy że części urojone $\langle F(w), v \rangle$ i $\langle w, F(v) \rangle$ są równe, a dla $\alpha = i$, że części rzeczywiste są równe. Stąd $\langle F(w), v \rangle = \langle w, F(v) \rangle$, a z dowolności v, w mamy $T = T^*$. \square

Uwaga 10.9. Z dowodu widzimy, że wystarczy założyć, że $(\forall v)\langle v, T(v) \rangle \in \mathbf{R}$.

Przykłady 10.10. • Rzeczywista wielokrotność przekształcenia samosprzężonego jest samosprzężona.

- Rzeczywista kombinacja liniowa przekształceń samosprzężonych jest samosprzężona.
- W szczególności rzeczywiste kombinacje liniowe rzutów ortogonalnych są samosprzężone

Zobaczmy wkrótce, że ostatni przykład powyżej opisuje *wszystkie* przekształcenia samosprzężone: są to dokładnie rzeczywiste kombinacje liniowe rzutów ortogonalnych!

Lemat 10.11. *Wartości własne zespolonych samosprzężonych endomorfizmów/macierzy są rzeczywiste.*

Dowód. Ćwiczenie. □

Wniosek 10.12. *Jeżeli F jest rzeczywistym samosprzężonym endomorfizmem, jego wielomian charakterystyczny rozkłada się na czynniki liniowe nad \mathbf{R} .*

Dowód. $F_{\mathbf{C}}$ też jest samosprzężony (np. dlatego że ma tę samą macierz). Stąd wszystkie zespolone pierwiastki $\chi_F(x) = \chi_{F_{\mathbf{C}}}(x)$ są rzeczywiste, a stąd łatwo wynika teza. □

Twierdzenie 10.13 (Twierdzenie spektralne). • *Założmy że V jest przestrzenią unitarną/euklidesową i $F \in \text{End}(V)$ jest samosprzężony. Wtedy F diagonalizuje się w bazie ON.*

- *Założmy że $A \in M_{n \times n}(\mathbf{C})/M_{n \times n}(\mathbf{R})$ jest macierzą hermitowską/symetryczną. Wtedy istnieje macierz unitarna/ortogonalna P taka że $A = PDP^*/PDP^{\top}$, gdzie D jest macierzą diagonalną (rzeczywistą, z poprzedniego wniosku).*

Uwaga 10.14. Z twierdzenia spektralnego łatwo wynika, że samosprzężony $F \in \text{End}(V)$ jest nieujemnie określony $\iff \sigma(F) \subseteq [0, +\infty)$ (ćwiczenie), a dodatnio określony $\iff \sigma(F) \subseteq (0, +\infty)$.

W szczególności nieujemnie określony F jest dodatnio określony \iff jest odwracalny.

Uwaga 10.15. Wariant twierdzenia spektralnego jest prawdziwy również dla nieskończone wymiarowych przestrzeni (nad \mathbf{R} lub \mathbf{C}) z iloczynem skalarnym.

Uwaga 10.16. Łatwo pokazać twierdzenie odwrotne do twierdzenia spektralnego: jeżeli F diagonalizuje się w bazie ON i ma rzeczywiste wartości własne, to jest samosprężony (bo macierz w tej samej bazie jest rzeczywista i diagonalna, więc hermitowska).

Dowód. Druga część łatwo wynika z pierwszej.

Udowodnijmy pierwszą część przez indukcję ze względu na $\dim V$.

Jeżeli $\dim V = 1$, to teza jest oczywista. Załóżmy że $\dim V = n > 1$. Weźmy dowolne $\lambda \in \sigma(F)$ (w przypadku euklidesowym λ istnieje z poprzedniego wniosku) i wektor własny $v_1 \in V_\lambda$ dla F . Możemy założyć bez zmniejszania ogólności, że $|v_1| = 1$. Wtedy v_1 rozszerza się do bazy ON $B' = v_1, v'_2, \dots, v'_n$. Ponieważ v_1 jest wektorem własnym dla λ , mamy

$$m_B(F) = M = \begin{pmatrix} \lambda & * \\ 0 & N \end{pmatrix}$$

dla pewnej macierzy $N \in M_{(n-1) \times (n-1)}(\mathbf{C})$. Ponieważ F jest samosprężony, $M = M^*$, więc $M = \begin{pmatrix} \lambda & 0 \\ 0 & N \end{pmatrix}$, gdzie $N = N^*$.

Rozważmy $W = \text{Lin}(v'_2, \dots, v'_n)$. Z postaci $m_B(F)$ widzimy, że W jest F -niezmiennicze, więc $F|_W \in \text{End}(W)$. Wtedy dla $B'_0 = (v'_2, \dots, v'_n)$ mamy $m_{B'_0}(F|_W) = N = N^*$, więc $F|_W$ jest samosprężone. Stąd z założenia indukcyjnego istnieje baza ON $B_0 = v_2, v_3, \dots, v_n$ przestrzeni W składająca się z wektorów własnych F . Stąd $B = (v_1, v_2, \dots, v_n)$ jest bazą ON wektorów własnych F . \square

Przykład 10.17. Weźmy macierz hermitowską $M = \begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix}$. Jej wielomian charakterystyczny do $x^2 - 4x + 3 = (x - 3)(x - 1)$, wartości własne to 3 i 1, a odpowiadające im wektory własne to np. $\begin{pmatrix} 1 \\ -i \end{pmatrix}$ i $\begin{pmatrix} 1 \\ i \end{pmatrix}$.

Obydwa są długości $\sqrt{2}$, więc to daje nam bazę ON $B = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}$, czyli $m_E^B(\text{id}) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$, więc $m_B^E(\text{id}) = m_E^B(\text{id})^{-1} = m_E^B(\text{id})^* = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}$,

czyli

$$m_B(F_M) = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$$

Przykład 10.18. Jeżeli $M \in M_{2 \times 2}(\mathbf{C})$ jest hermitowska i $\chi_M(x)$ ma podwójny pierwiastek λ , to z twierdzenia spektralnego M się diagonalizuje, czyli $M = P(\lambda I)P^* = \lambda I$ i $\lambda \in \mathbf{R}$. Faktycznie, jeżeli M jest hermitowska, to jest postaci

$$\begin{pmatrix} a & b \\ \bar{b} & d \end{pmatrix},$$

gdzie $a, b \in \mathbf{R}$, a jej wielomian charakterystyczny to $x^2 - (a+d)x + ad - |b|^2$ i ma podwójny pierwiastek dokładnie gdy

$$0 = \Delta = (a+d)^2 - 4(ad - |b|^2) = (a-d)^2 + 4|b|^2,$$

co (dla $a, d \in \mathbf{R}$) zachodzi tylko gdy $a = d$ i $b = 0$.

Uwaga 10.19. Jeżeli $F \in \text{End}(V)$ jest samosprężony, to $\varphi(v, w) = \langle v, F(w) \rangle = \langle F(v), w \rangle$ jest formą hermitowską/symetryczną i dla dowolnej bazy ON B mamy $m_B(\varphi) = m^{BB}(\varphi)$.

Z drugiej strony jeżeli B jest bazą ON i φ jest formą hermitowską/symetryczną, to istnieje (jedyne!) taki samosprężony $F = F_\varphi \in \text{End}(V)$, że $\varphi(v, w) = \langle v, F(w) \rangle = \langle F(v), w \rangle$ (taki że $m_B(F) = m^{BB}(\varphi)$ dla pewnej bazy ON B).

Wtedy $m_C(F_\varphi) = m^{CC}(\varphi)$ dla każdej bazy ON C :

$$\begin{aligned} m_C(F) &= m_C^B(\text{id}) m_B(F_\varphi) m_B^C(\text{id}) \\ &= m_B^C(\text{id})^{-1} m^{BB}(\varphi) m_B^C(\text{id}) \\ &= m_B^C(\text{id})^* m^{BB}(\varphi) m_B^C(\text{id}) \\ &= m^{CC}(\varphi) \end{aligned}$$

Stąd i z twierdzenia spektralnego wnioskujemy wzmocnienie twierdzenia Lagrange'a.

Wniosek 10.20. • Każda forma hermitowska/rzeczywista symetryczna diagonalizuje się w bazie ON (tzn. istnieje taka baza ON B że $m^{BB}(\varphi)$ jest diagonalna).

- Jeżeli V jest przestrzenią liniową nad \mathbf{C}/\mathbf{R} i φ_1, φ_2 są hermitowskie/symetryczne, a ponadto φ_1 jest dodatnio określona, to istnieje taka baza B że $m^{BB}(\varphi_1) = I$ i $m^{BB}(\varphi_2)$ jest diagonalna. W szczególności φ_1, φ_2 diagonalizują się we wspólnej bazie.

Dowód. Ćwiczenie. Wskazówki:

- Pierwsza część: skorzystaj z poprzedniej uwagi i zastosuj twierdzenie spektralne do F_φ .
- Druga część: φ_1 jest dodatnio określona, rozważ V jako przestrzeń unitarną/euklidesową z iloczynem skalarnym $\langle -, - \rangle = \varphi_1$. \square

Wniosek 10.21. Jeżeli φ jest formą hermitowską, to istnieje taka baza B (niekoniecznie ON), że $m^{BB}(\varphi)$ jest diagonalna z 1, -1 i 0 na przekątnej.

Jeżeli A jest macierzą hermitowską, to istnieje taka macierz odwracalna P że PAP^* jest diagonalna z 1, -1 i 0 na przekątnej.

Dowód. Dowód dla pierwszej części (druga część jest łatwo równoważna).

Weźmy bazę ON $C = c_1, \dots, c_n$ taką że $m^{CC}(\varphi)$ jest diagonalna. Weźmy

$$b_j = \begin{cases} \frac{c_j}{\sqrt{\varphi(c_j, c_j)}} & \varphi(c_j, c_j) \neq 0 \\ c_j & \varphi(c_j, c_j) = 0. \end{cases}$$

Wtedy $B = (b_1, \dots, b_n)$ działa: z dwuliniowości $\varphi(b_j, b_j) = \frac{1}{\sqrt{\varphi(c_j, c_j)}} \cdot \frac{1}{\sqrt{\varphi(c_j, c_j)}} \varphi(c_j, c_j) = \frac{1}{|\varphi(c_j, c_j)|^2} \varphi(c_j, c_j) = \text{sgn}(\varphi(c_j, c_j))$. \square

Fakt 10.22 (Prawo bezwładności Sylwestera dla form hermitowskich). Liczba 1, -1 i 0 na przekątnej $m^{BB}(\varphi)$ z poprzedniego wniosku nie zależy od wyboru B .

Dowód. Analogiczny jak dla form symetrycznych nad \mathbf{R} . \square

Definicja 10.23. Sygnatura formy hermitowskiej zdefiniowana jest tak samo jak dla form symetrycznych rzeczywistych: to trójka (p, q, r) liczb 1, -1 i zer na przekątnej $m^{BB}(\varphi)$ z wniosku powyżej.

Wniosek 10.24. Jeżeli φ jest formą hermitowską/rzeczywistą symetryczną sygnatury (p, q, r) , a A jest macierzą φ względem dowolnej bazy, to p = liczba

dodatnich pierwiastków $\chi_A(x)$ (z krotnościami), q = liczba ujemnych pierwiastków $\chi_A(x)$ (z krotnościami), r = liczba zerowych pierwiastków $\chi_A(x)$.

W szczególności φ jest dodatnio określona wtedy i tylko wtedy gdy $\chi_A(x)$ ma tylko dodatnie pierwiastki.

Dowód. Z twierdzenia spektralnego $A = PDP^* = PDP^{-1}$ dla pewnej unitarnej P i wartości na przekątnej D to pierwiastki $\chi_A(x)$. Stąd i z dowodu poprzedniego Wniosku 10.21 wynika teza. \square

Przykład 10.25. Dla formy hermitowskiej $\varphi(v, w) = 2v_1w_1 + iv_1w_2 - iv_2w_1 + 2v_2w_2$ (o macierzy w standardowej bazie $\begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix}$) znaleźliśmy bazę ON w której macierz to $\begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$. Wszystkie znaki na przekątnej są dodatnie, więc φ ma sygnaturę $(2, 0, 0)$ i jest dodatnio określona.

Ćwiczenie 10.26. Jeżeli F jest samosprężone i v, w są wektorami własnymi odpowiadającymi różnym wartościom własnym, to $v \perp w$. (To oczywiście nie jest prawda, gdy v, w odpowiadają tej samej wartości własnej!)

Wniosek 10.27 (Rozkład spektralny). Jeżeli $F \in \text{End}(V)$ jest samosprężony, to zachodzi wzór:

$$F = \sum_{\lambda \in \sigma(F)} \lambda P_\lambda,$$

gdzie P_λ jest rzutem ortogonalnym na przestrzeń własną V_λ dla F , ponadto $\sum_{\lambda \in \sigma(F)} P_\lambda = \text{id}$ i $P_\lambda P_\mu = 0$ dla $\lambda \neq \mu$.

Dowód. Weźmy dowolny $v \in V$. Ponieważ F się diagonalizuje, $V = \bigoplus_\lambda V_\lambda$, $v = \sum_{\lambda \in \sigma(F)} v_\lambda$, gdzie $F(v_\lambda) = \lambda v_\lambda$. Z poprzedniego wniosku widzimy, że $P_\lambda(v_\mu) = \begin{cases} v_\mu & \mu = \lambda \\ 0 & \mu \neq \lambda \end{cases}$, więc $v_\lambda = P_\lambda(v)$. Stąd łatwo wynika teza. \square

Przykład 10.28. Mamy

$$\begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix},$$

czyli

$$\begin{aligned} \begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} = \\ &= \frac{3}{2} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} = 3 \underbrace{\begin{pmatrix} \frac{1}{2} & \frac{i}{2} \\ \frac{-i}{2} & \frac{1}{2} \end{pmatrix}}_{P_3} + \underbrace{\begin{pmatrix} \frac{1}{2} & \frac{-i}{2} \\ \frac{i}{2} & \frac{1}{2} \end{pmatrix}}_{P_1} \end{aligned}$$

Uwaga 10.29. Twierdzenie spektralne i rozkład spektralny mają też wariant dla tzw. normalnych endomorfizmów i macierzy, tzn. takich, które spełniają $AA^* = A^*A$. Konkluzja (w przypadku zespolonym) jest taka sama, poza tym że wartości własne nie muszą być rzeczywiste. W przypadku rzeczywistym wymaga modyfikacji podobnej jak w tw. Jordana.

10.1 Rozkład singularny i rozkład biegunowy

Ćwiczenie 10.30. Załóżmy że $A \in M_{n \times m}(\mathbb{C})$ jest dowolną macierzą. Wtedy dla dowolnego $v \in \mathbb{C}^n, w \in \mathbb{C}^m$ mamy

$$\langle v, Aw \rangle = \langle A^*v, w \rangle.$$

(Gdzie po lewej i prawej stronie mamy standardowe iloczyny skalarne dla \mathbb{C}^n i \mathbb{C}^m .)

Niech $A \in M_{n \times m}(\mathbb{C})$ będzie dowolną macierzą. Rozważmy wartość $R(v) = \frac{|Av|}{|v|}$ dla $v \in \mathbb{C}^m \setminus \{0\}$. Zauważmy że zastąpienie v przez $\frac{v}{|v|}$ nie zmienia wartości $R(v)$ i wtedy

$$R(v) = |Av| = \sqrt{\langle Av, Av \rangle} = \sqrt{\langle v, A^*Av \rangle}.$$

Macierz A^*A jest oczywiście hermitowska, diagonalizuje się więc w bazie ON $B = b_1, \dots, b_m$ i możemy rozpisać $v = \sum_{j=1}^m \alpha_j b_j$, gdzie $A^*Ab_j = \lambda_j b_j$. Wtedy

$$R(v)^2 = \langle v, A^*Av \rangle = \sum_{j=1}^m \alpha_j^2 \langle b_j, \lambda_j b_j \rangle = \sum_{j=1}^m \alpha_j^2 \lambda_j.$$

Ponieważ ta liczba jest zawsze nieujemna, widzimy że $\lambda_j \geq 0$.

Wniosek 10.31. Jeżeli $A \in M_{n \times m}(\mathbb{C})$ jest dowolną macierzą, to A^*A jest nieujemnie określona.

Możemy zmienić kolejność b_j tak że $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$.

Niech $k = \text{rk} A^*A$. Wtedy $\lambda_1, \dots, \lambda_k > 0$ i $A^*Ab_j \neq 0$ dla $j \leq k$, a więc też $Ab_j \neq 0$. Ponadto wektory te są ortogonalne: jeżeli $j < l \leq k$, to

$$\langle Ab_j, Ab_l \rangle = \langle b_j, A^*Ab_l \rangle = \lambda_l \langle b_j, b_l \rangle = 0,$$

Położmy dla $j \leq k$:

$$c_j = \frac{Ab_j}{|Ab_j|} = \frac{Ab_j}{R(b_j)} = \frac{Ab_j}{\sqrt{\lambda_j}}.$$

Wtedy układ $c_1, \dots, c_k \in \mathbb{C}^n$ jest ON, możemy więc rozszerzyć go do bazy ON $C = c_1, \dots, c_n$ przestrzeni \mathbb{C}^n .

Twierdzenie 10.32 (rozkład singularny). *Jeżeli $A \in M_{n \times m}(\mathbb{C})$ jest dowolna, to istnieją macierze unitarne $U_1 \in M_{n \times n}(\mathbb{C})$, $U_2 \in M_{m \times m}(\mathbb{C})$ i diagonalna $D \in M_{n \times m}(\mathbb{R}_{\geq 0})$ taka że*

$$A = U_1 D U_2^*,$$

przy czym

$$D = \left(\begin{array}{cccc} \sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_k} \\ & 0 & & & 0 \end{array} \right) \left. \vphantom{\begin{array}{c} \sqrt{\lambda_1} \\ \sqrt{\lambda_2} \\ \ddots \\ \sqrt{\lambda_k} \\ 0 \end{array}} \right\} \begin{array}{l} n \text{ wierszy,} \\ m \text{ kolumn} \end{array}$$

gdzie $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k$ to dodatnie (równoważnie, niezerowe) wartości własne A^*A .

Ponadto jeżeli A jest rzeczywista, to możemy wybrać U_1, U_2 rzeczywiste.

Dowód. Macierze $U_1 = (c_1, \dots, c_n)$, $U_2 = (b_1, \dots, b_m)$ działają. Istotnie, $U_2^* = U_2^{-1}$, więc dla każdego b_j mamy:

$$U_1 D U_2^* b_j = U_1 D e_j = \begin{cases} U_1 \sqrt{\lambda_j} e_j = \sqrt{\lambda_j} U_1 e_j = \sqrt{\lambda_j} c_j = Ab_j & j \leq k \\ U_1 0 = 0 = Ab_j & j > k. \end{cases}$$

Ponieważ $B = b_1, \dots, b_m$ jest bazą \mathbb{C}^m , wynika z tego że $A = U_1 D U_2^*$. □

Uwaga 10.33. Jeżeli $A = U_1 D U_2^*$, to $A^* = U_2 D^\top U_1^*$.

Przykład 10.34. Weźmy macierz $A = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 0 & 1 \end{pmatrix}$. Szukamy jej rozkładu singularnego. Wyznamy najpierw rozkład singularny $A^* = A^\top$: to będzie łatwiejsze, bo $A^{**}A^* = AA^*$ jest macierzą 2x2, a A^*A macierzą 3x3.

$$AA^* = AA^\top = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 2 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & 2 \end{pmatrix}.$$

Widzimy że AA^* jest już w postaci diagonalnej i wartości na przekątnej są uporządkowane malejąco, więc możemy wziąć $b_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $b_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Wtedy

$$c_1 = \frac{A^\top b_1}{|A^\top b_1|} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad c_2 = \frac{A^\top b_2}{|A^\top b_2|} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}.$$

Dopełniamy c_1, c_2 do bazy ON, na przykład przez $c_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$. Stąd mamy

$$A^\top = \begin{pmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} \sqrt{6} & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^\top,$$

więc szukany rozkład singularny to:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{6} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix}^\top$$

Wniosek 10.35. Jeżeli $F: V \rightarrow W$ jest odwzorowaniem między przestrzeniami unitarnymi/euklidesowymi, to istnieją takie bazy ON B dla V i C dla W , że $m_C^B(F)$ jest macierzą diagonalną z nieujemnymi rzeczywistymi wyrazami na przekątnej. Możemy ponadto założyć, że są one uporządkowane nierosnąco.

Dowód. Weźmy dowolne bazy ON B', C' dla V, W odpowiednio. Z poprzedniego twierdzenia możemy wybrać macierze unitarne U_1, U_2 oraz diagonalną D taką że

$$m_{C'}^{B'}(F) = U_1 D U_2^*.$$

Wtedy możemy wziąć bazy ON B, C dla V, W takie że $U_2 = m_{B'}^B(\text{id}_W), U_1 = m_{C'}^C(\text{id}_V)$ i wtedy

$$m_C^B(F) = m_C^{C'}(\text{id})m_{C'}^{B'}(F)m_{B'}^B(\text{id}) = U_1^*U_1DU_2^*U_2 = D. \quad \square$$

Definicja 10.36. Wartości $\sigma_j = \sqrt{\lambda_j}$ na przekątnej w D powyżej nazywamy *wartościami singularnymi* F (lub macierzy A).

Uwaga 10.37. Przypomnijmy sobie wzór dla v jednostkowych:

$$\frac{|Av|^2}{|v|^2} = R(v)^2 = \langle v, A^*Av \rangle = \sum_{j=1}^n \alpha_j^2 \langle b_j, \lambda_j b_j \rangle = \sum_{j=1}^n \alpha_j^2 \lambda_j = \sum_{j=1}^k \alpha_j^2 \lambda_j = \sum_{j=1}^k \alpha_j^2 \sigma_j^2.$$

Zauważmy że stąd:

- σ_1 to maksymalna wartość $|R(v)|$ dla $v \neq 0$ (np. z nierówności między średnimi),
- b_1 to wektor jednostkowy dla którego ta wartość jest osiągnięta, a c_1 to wektor jednostkowy w kierunku Ab_1 .
- σ_2 to maksymalna wartość $R(v)$ dla niezerowych v ortogonalnych do b_1 ,
- b_2 to wektor jednostkowy na którym jest ona osiągnięta, c_2 to wektor jednostkowy w kierunku Ab_2 ,
- σ_3 to maksymalna wartość $R(v)$ dla $v \perp b_1, b_2, \dots$
- itd.

Ponadto dla $j \leq k$ zachodzą wzory:

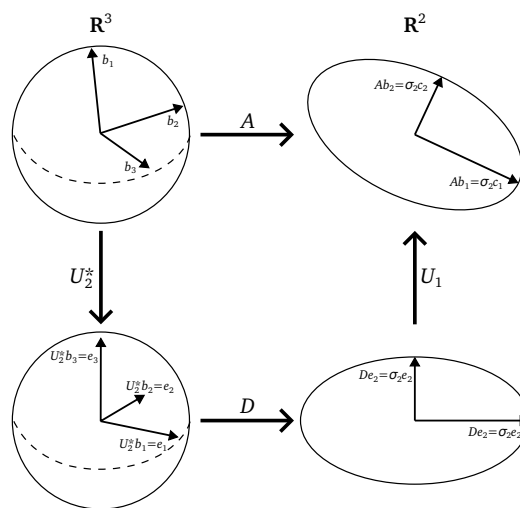
$$Ab_j = \sigma_j c_j \quad A^*c_j = \sigma_j b_j.$$

(Dla $j > k$ mamy $Ab_j = 0$ i $A^*c_j = 0$, o ile te napisy mają sens.)

Drugi wzór wynika z obserwacji, że

$$A^*c_j = A^*\sigma_j^{-1}Ab_j = \sigma_j^{-1}A^*Ab_j = \sigma_j^{-1}\sigma_j^2b_j = \sigma_j b_j.$$

Stąd bierze się interpretacja geometryczna rozkładu singularnego (rysunek 1).



Rysunek 10.1: Interpretacja geometryczna rozkładu singularnego odwzorowania $\mathbf{R}^3 \rightarrow \mathbf{R}^2$.

Definicja 10.38. Wektor b_j nazywamy *prawym wektorem singularnym* (odpowiadającym σ_j), a wektor c_j *lewym wektorem singularnym*. (Mnemotechnika: w $\sigma_j = \langle c_j, Ab_j \rangle = c_j^* Ab_j$ wektor b_j jest po prawej, c_j po lewej od A .)

Uwaga: wektory singularne *nie* są jedyne, nawet z dokładnością do znaku (np. dla id każdy wektor jest lewy i prawy singularny dla $\sigma_1 = \dots = \sigma_n = 1$).

Przykład 10.39 (metoda najmniejszych kwadratów). Załóżmy że mamy N punktów $P_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, P_2, \dots, P_N$ na płaszczyźnie $V = \mathbf{R}^2$ (lub \mathbf{C}^2). Chcemy znaleźć prostą $W \leq V$ najlepiej przybliżającą te punkty w tym sensie, że minimalizuje sumę kwadratów odległości:

$$\sum_{j=1}^N d(P_j, W)^2$$

Biorąc wektor jednostkowy v w kierunku W , mamy

$$\begin{aligned}
 d(P_j, W)^2 &= \langle P_j - \langle v, P_j \rangle v, P_j - \langle v, P_j \rangle v \rangle \\
 &= \langle P_j, P_j - \langle v, P_j \rangle v \rangle - \langle \langle v, P_j \rangle v, P_j - \langle v, P_j \rangle v \rangle \\
 &= |P_j|^2 - \langle P_j, v \rangle \langle v, P_j \rangle - \overline{\langle v, P_j \rangle} \langle v, P_j - \langle v, P_j \rangle v \rangle \\
 &= |P_j|^2 - |\langle v, P_j \rangle|^2 - \langle P_j, v \rangle \langle v, P_j - \langle v, P_j \rangle v \rangle \\
 &= |P_j|^2 - |\langle v, P_j \rangle|^2 - \langle P_j, v \rangle (\langle v, P_j \rangle - \langle v, P_j \rangle \langle v, v \rangle) \\
 &= |P_j|^2 - |\langle v, P_j \rangle|^2.
 \end{aligned}$$

Stąd szukana prosta jest rozpinana przez jednostkowy v maksymalizujący $\sum_{j=1}^N |\langle v, P_j \rangle|^2$.

Rozważmy macierz $A = \begin{pmatrix} x_1 & y_1 \\ \vdots & \vdots \\ x_N & y_N \end{pmatrix}$. Zauważmy że:

$$Av = \begin{pmatrix} \langle P_1, v \rangle \\ \vdots \\ \langle P_N, v \rangle \end{pmatrix},$$

czyli $|Av| = \sqrt{\sum_{j=1}^n |\langle P_j, v \rangle|^2} = \sqrt{\sum_{j=1}^n |\langle v, P_j \rangle|^2}$. Szukamy więc wektora jednostkowego maksymalizującego $|Av|$ — ale to jest z definicji pierwszy wektor singularny A , tzn. pierwsza kolumna V dla rozkładu singularnego $A = UDV^*$.

Podobnie dla ciągu N punktów w \mathbf{R}^n również możemy wyznaczyć prostą (lub ogólniej k -wymiarową podprzestrzeń dla pewnego $k < n$) najlepiej je przybliżającą (minimalizującą sumę kwadratów odległości) — będzie ona rozpięta przez pierwszy wektor singularny/pierwsze k wektorów singularnych macierzy $A \in M_{N \times n}(\mathbf{R})$, której wiersze to współrzędne danych punktów (w przypadku zespolonym: ich sprzężenia; można też wziąć lewe wektory singularne macierzy której kolumny to współrzędne danych punktów — wtedy nie trzeba sprzęgać).

Wniosek 10.40 (rozkład biegunowy). • Jeżeli $F \in \text{End}(V)$ i B jest dowolną bazą ON , to istnieje taka baza ON C że $m_C^B(F)$ jest nieujemnie określona.

- Jeżeli $A \in M_{n \times n}(\mathbf{C})$ jest dowolną macierzą, to istnieje macierz unitarna U i dodatnio określona P taka że $A = UP$. Jeżeli A jest rzeczywista, to U, P też można wybrać takie.

Dowód. Druga część: wychodząc z rozkładu singularnego $A = U_1 D U_2$, weźmy $P = U_2^* D U_2$, $U = U_1 U_2$.

Pierwsza część wynika z drugiej: wychodząc z $A = m_B(F)$ musimy wybrać za C taką bazę V , że $m_C^B(\text{id}) = U$. Taka baza zawsze istnieje i jest ON, bo macierz U jest unitarna. \square

Rozdział 11

Izometrie i objętości

11.1 Reprezentacje przekształceń ortogonalnych i unitarnych

Stwierdzenie 11.1. *Jeżeli F jest unitarne/ortogonalne, to wartości własne F (również nierzeczywiste) są co do modułu równe 1.*

Ponadto jeżeli v, w są wektorami własnymi odpowiadającymi wartościom własnym $\lambda \neq \mu$, to $v_1 \perp v_2$.

Dowód. Niech v będzie wektorem własnym odpowiadającej wartości własnej λ . Wtedy:

$$|v|^2 = \langle v, v \rangle = \langle F(v), F(v) \rangle = \langle \lambda v, \lambda v \rangle = \bar{\lambda} \lambda \langle v, v \rangle = |\lambda|^2 |v|^2.$$

Stąd $|\lambda| = 1$.

Teraz niech w będzie jak w stwierdzeniu. Wtedy podobnie:

$$\langle v, w \rangle = \langle \lambda v, \mu w \rangle = \bar{\lambda} \mu \langle v, w \rangle.$$

Ponieważ $|\mu| = 1$, $\mu^{-1} = \bar{\mu} \neq \bar{\lambda}$ (bo z założenia $\mu \neq \lambda$), więc $\bar{\lambda} \mu \neq 1$, czyli $\langle v, w \rangle = 0$. \square

Stwierdzenie 11.2. *(V nad dowolnym ciałem, być może nieskończenie wymiarowa.) Jeżeli $F \in \text{End}(V)$ jest odwracalny i $W \leq V$ jest skończenie wymiarowa i F -niezmiennicza, to W jest F^{-1} -niezmiennicza.*

Dowód. Niech $W' = F[W] \leq W$. Wtedy F zadaje izomorfizm $W \rightarrow W'$, więc $\dim W = \dim W'$. Stąd $W = W'$, czyli każdy $w' \in W$ jest postaci $F(w)$ dla pewnego $w \in W$, zatem $F^{-1}(w') = w \in W$. \square

Uwaga 11.3. Jeżeli $F \in \text{End}(V)$ i $W \leq V$ jest F -niezmiennicza, to W^\perp jest F^* -niezmiennicza: dla każdego $v \in W^\perp$ i $w \in W$ mamy:

$$\langle F^*(v), w \rangle = \langle v, F(w) \rangle = 0.$$

Wniosek 11.4. Jeżeli F jest unitarne i W jest F -niezmiennicza, to W^\perp też jest F -niezmiennicza.

Dowód. Z poprzedniego stwierdzenia W^\perp jest $F^* = F^{-1}$ -niezmiennicza, a więc też F -niezmiennicza (bo jest skończenie wymiarowa). \square

Stwierdzenie 11.5. (V unitarna.) Jeżeli F jest unitarne, to diagonalizuje się w bazie ON.

pierwszy dowód. Wynika z Twierdzenia C.6 (spektralnego dla normalnych endomorfizmów). \square

drugi dowód. Indukcja względem $\dim V$. Gdy $\dim V = 1$, to teza jest oczywista. W przeciwnym wypadku weźmy jednostkowy wektor własny b_1 . Wtedy $\text{Lin}(b_1)$ jest F -niezmiennicze, więc $W = b_1^\perp$ też i $F|_W$ jest oczywiście unitarne. Z założenia indukcyjnego W ma bazę B' ON wektorów własnych dla $F|_W$. Wtedy $B = b_1 B'$ działa. \square

Twierdzenie 11.6 (twierdzenie spektralne dla przekształceń unitarnych). (V unitarna.) $F \in \text{End}(V)$ jest unitarny wtedy i tylko wtedy gdy istnieje baza ON B taka że $m_B(F)$ jest diagonalna i ma na przekątnej wartości o module 1.

Dowód. Implikacja z lewej do prawej to poprzednie stwierdzenie. Implikacja w prawej do lewej:

$$m_B(F)m_B(F)^* = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_n \end{pmatrix} \begin{pmatrix} \bar{\lambda}_1 & & \\ & \bar{\lambda}_2 & \\ & & \ddots \\ & & & \bar{\lambda}_n \end{pmatrix} = \begin{pmatrix} |\lambda_1| & & \\ & |\lambda_2| & \\ & & \ddots \\ & & & |\lambda_n| \end{pmatrix} = I,$$

więc $m_B(F)$ jest unitarna, czyli F jest unitarne. \square

Uwaga 11.7. Jeżeli F jest przekształceniem ortogonalnym przestrzeni euklidesowej V , to F_C jest przekształceniem unitarnym V_C : istotnie, wystarczy wziąć bazę ON B dla V i wtedy B jest też bazą ON dla V_C i macierz $m_B(F) = m_B(F_C)$ jest ortogonalna i rzeczywista, a więc unitarna.

11.1. REPREZENTACJE PRZEKSZTAŁCEŃ ORTOGONALNYCH I UNITARNYCH 173

Uwaga 11.8 (przypomnienie). Jeżeli $F \in \text{End}(V)$ dla rzeczywistej przestrzeni liniowej V i λ jest nierzeczywistą wartością własną $F_{\mathbb{C}}$ i $F_{\mathbb{C}}(v) = \lambda v$ dla pewnego niezerowego $v \in V_{\mathbb{C}}$, to:

- $\text{Lin}_{\mathbb{R}}(\text{Re}(v), \text{Im}(v)) = V \cap \text{Lin}_{\mathbb{C}}(v, \bar{v})$ jest dwuwymiarową F -niezmienniczą podprzestrzenią V ,
- $F(\text{Re}(v)) = F_{\mathbb{C}}((v + \bar{v})/2) = \frac{1}{2}(\lambda v + \bar{\lambda} \bar{v}) = \text{Re}(\lambda v) = \text{Re}(\lambda) \text{Re}(v) - \text{Im}(\lambda) \text{Im}(v)$ i podobnie $F(\text{Im}(v)) = \text{Im}(\lambda v) = \text{Im}(\lambda) \text{Re}(v) + \text{Re}(\lambda) \text{Im}(v)$,
- w szczególności macierz F obciętego do tej podprzestrzeni w bazie $\text{Re}(v), \text{Im}(v)$ to $\begin{pmatrix} \text{Re}(\lambda) & \text{Im}(\lambda) \\ -\text{Im}(\lambda) & \text{Re}(\lambda) \end{pmatrix}$,
- w szczególności gdy $|\lambda| = 1$, to $\lambda = \cos \theta + i \sin \theta$, więc macierz to $\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$.

Twierdzenie 11.9 (twierdzenie spektralne dla przekształceń ortogonalnych). (V euklidesowa.) $F \in \text{End}(V)$ jest ortogonalne \iff istnieje baza ON B taka że

$$m_B(F) = \begin{pmatrix} \pm 1 & & & & & \\ & \pm 1 & & & & \\ & & \ddots & & & \\ & & & \pm 1 & & \\ & & & & R_{\theta_1} & \\ & & & & & R_{\theta_2} \ddots \\ & & & & & & R_{\theta_k} \end{pmatrix}$$

$$= \begin{pmatrix} \pm 1 & & & & & \\ & \pm 1 & & & & \\ & & \ddots & & & \\ & & & \pm 1 & & \\ & & & & \cos \theta_1 & -\sin \theta_1 \\ & & & & \sin \theta_1 & \cos \theta_1 \\ & & & & & \ddots \\ & & & & & & \cos \theta_k & -\sin \theta_k \\ & & & & & & \sin \theta_k & \cos \theta_k \end{pmatrix}$$

Definicja 11.10. Wyrażenie z konkluzji twierdzenia spektralnego powyżej nazywamy *postacią kanoniczną* F .

Dowód. Implikacja z prawej do lewej jest oczywista (macierz zadanej postaci jest ortogonalna). Udowodnimy implikację z lewej do prawej.

Dowodzimy przez indukcję względem $\dim V$. Gdy $\dim V = 1$, teza jest oczywista, a gdy $\dim V = 2$, to jest znana z ubiegłego semestru (bo wtedy V jest płaszczyzną euklidesową).

Jeżeli F ma (rzeczywistą) wartość własną, to jest ona równa ± 1 (bo jej moduł to 1) i jak w przypadku unitarnym, wybieramy jednostkowy wektor własny b_1 . Wtedy $W = b_1^\perp$ jest F -niezmiennicza i z założenia indukcyjnego mamy bazę ON B' dla W w której $F|_W$ ma żadaną postać i $B = b_1 B'$ działa.

Jeżeli F nie ma rzeczywistej wartości własnej, to weźmy wektor własny $v \in V_{\mathbb{C}}$ dla $F_{\mathbb{C}}$ odpowiadający nierzeczywistej wartości własnej λ . Wtedy $\text{Lin}_{\mathbb{C}}(v, \bar{v})$ jest $F_{\mathbb{C}}$ -niezmienniczny, więc $W_1 = \text{Lin}_{\mathbb{C}}(v, \bar{v}) \cap V$ jest F -niezmienniczny. Z ćwiczenia wiemy że $\dim W_1 = 2$, więc z założenia indukcyjnego W_1 i $W_2 = W_1^\perp$ mają bazy ON B_1, B_2 takie że $m_{B_j}(F|_{W_j})$ są jak w tezie (z tym że bez wartości ± 1 , bo inaczej F miałoby wartość własną!). Stąd $B = B_1 B_2$ działa. \square

11.2 Izometrie

Definicja 11.11. Mówimy że odwzorowanie $F: V \rightarrow V$ (niekoniecznie liniowe) jest *izometrią* jeżeli zachowuje odległości, tzn. dla każdych $v_1, v_2 \in V$ zachodzi $|v_1 - v_2| = |F(v_1) - F(v_2)|$.

Przykłady 11.12. • translacje $T_{v_0}(v) = v + v_0$ są izometriami,

- obroty \mathbb{R}^2 wokół punktów, \mathbb{R}^3 wokół prostych są izometriami,
- sprzężenie w \mathbb{C}^n jest izometrią (również sprzężenie tylko niektórych współrzędnych),
- przekształcenie $\mathbb{R}^n / \mathbb{C}^n$ zadane (w standardowej bazie) macierzą diago-

nalną postaci

$$\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & -1 & \leftarrow i\text{-te miejsce} \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix}$$

(symetria w hiperpłaszczyźnie $\text{Lin}(e_1, \dots, e_{i-1}, e_{i+1}, \dots, e_n)$) jest izometrią,

- symetrie w dowolnej hiperpłaszczyźnie $W \ni 0$: $S_W = 2\pi_W - \text{id}$,
- przekształcenie $\mathbf{R}^n/\mathbf{C}^n$ zadane (w standardowej bazie) macierzą postaci

$$\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & R_\theta & \leftarrow i\text{-te i } (i+1)\text{-sze miejsce} \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix}$$

(obróć wokół “hiperprostej” $\text{Lin}(e_1, \dots, e_{i-1}, e_{i+2}, \dots, e_n)$) jest izometrią,¹

- podobnie symetrie w innych hiperpłaszczyznach/obroty wokół innych hiperprostych (niekoniecznie przechodzących przez 0).

Wniosek 11.13. *Odwzorowanie liniowe $F: V \rightarrow V$ jest izometrią wtedy i tylko wtedy gdy jest unitarne/ortogonalne.*

Dowód. Załóżmy że F jest unitarne/ortogonalne i weźmy dowolne $v_1, v_2 \in V$.

¹Podobnie jak w wymiarze 3, ustalenie hiperprostej nie wystarczy do wyznaczenia obrotu — aby odróżnić obrót o kąt θ od obrotu o kąt $-\theta$, trzeba ustalić jeszcze jej orientację (w przypadku $n = 3$ orientacja odpowiadała zwrotowi wektora kierunkowego).

Wtedy mamy

$$\begin{aligned}
 |v_1 - v_2|^2 &= \langle v_1 - v_2, v_1 - v_2 \rangle \\
 &= \langle F(v_1 - v_2), F(v_1 - v_2) \rangle \\
 &= \langle F(v_1) - F(v_2), F(v_1) - F(v_2) \rangle \\
 &= |F(v_1) - F(v_2)|^2.
 \end{aligned}$$

Z drugiej strony założmy, że F jest liniową izometrią.

Dowód w przypadku euklidesowym.

Wtedy z wzoru polaryzacyjnego (i liniowości F) dla dowolnych v, w mamy

$$\begin{aligned}
 \langle v, w \rangle &= \frac{|v + w|^2 - |v|^2 - |w|^2}{2} \\
 &= \frac{|F(v + w)|^2 - |F(v)|^2 - |F(w)|^2}{2} \\
 &= \frac{|F(v) + F(w)|^2 - |F(v)|^2 - |F(w)|^2}{2} \\
 &= \langle F(v), F(w) \rangle. \quad \square
 \end{aligned}$$

Dowód w przypadku unitarnym: ćwiczenie (przebiega analogicznie, tylko trzeba użyć wzoru polaryzacyjnego dla form hermitowskich).

Do następnego stwierdzenia będzie potrzebna jeszcze inna postać wzoru polaryzacyjnego dla form symetrycznych.

Uwaga 11.14. Jeżeli $\varphi(v, w)$ jest formą symetryczną i $Q(v) = \varphi(v, v)$, to mamy

$$\varphi(v, w) = -\varphi(v, -w) = -\frac{Q(v + (-w)) - Q(v) - Q(w)}{2} = \frac{Q(v) + Q(w) - Q(v - w)}{2}.$$

Stwierdzenie 11.15. (V euklidesowa.) Jeżeli F jest izometrią V i $F(0) = 0$, to F jest liniowe.

Dowód. Z założenia łatwo wynika, że dla każdego $v \in V$ mamy $|F(v)| = |v|$. Stąd i z poprzedniej uwagi wnioskujemy, że F zachowuje iloczyn skalarny:

$$\langle F(v), F(w) \rangle = \frac{1}{2}(|F(v)|^2 + |F(w)|^2 - |F(v) - F(w)|^2) = \frac{1}{2}(|v|^2 + |w|^2 - |v - w|^2) = \langle v, w \rangle.$$

Weźmy bazę ON $B = (b_1, \dots, b_n)$. Oznaczmy $b'_j := F(b_j)$. Wtedy $B' = b'_1, \dots, b'_n$ też jest ON i dla każdego $v = \sum_j \alpha_j b_j \in V$ mamy

$$\langle b'_j, F(v) \rangle = \langle b_j, v \rangle = \alpha_j,$$

a więc (ponieważ B' jest bazą ON) $F(v) = \sum_{j=1}^n \alpha_j b'_j$, więc F jest liniowe. \square

Uwaga 11.16. Ten sam dowód pokazuje, że przekształcenie przestrzeni unitarnej zachowujące iloczyn skalarny musi być liniowe.

Wniosek 11.17. Każda izometria przestrzeni euklidesowej jest złożeniem translacji i izometrii liniowej.

Dowód. Jeżeli F jest izometrią, to $F - F(0)$ też jest izometrią i zachowuje 0, więc jest liniową izometrią. Stąd $F = F(0) + (F - F(0)) = T_{F(0)} \circ (F - F(0))$. \square

Wniosek 11.18. Każda izometria n -wymiarowej przestrzeni euklidesowej jest złożeniem ≤ 1 translacji, k obrotów i m symetrii w hiperpłaszczyznach, przy czym $2k + m \leq n$.

Dowód. Wynika natychmiast z poprzedniego wniosku z Twierdzenia 11.9 (spektralnego dla ortogonalnych endomorfizmów). \square

Ćwiczenie 11.19. Każda izometria n -wymiarowej przestrzeni euklidesowej jest złożeniem $\leq n+1$ symetrii w hiperpłaszczyznach (w tym najwyżej jednej nieliniowej/w hiperpłaszczyźnie nie zawierającej 0).

11.3 Diagonalizacja form kwadratowych a hiperpowierzchnie kwadratowe

Widzieliśmy już (z twierdzenia spektralnego), że symetryczne formy dwuliniowe nad \mathbf{R} diagonalizują się w bazie ON. To pociąga za sobą oczywiście analogiczne stwierdzenie o formach kwadratowych.

Oznacza to że jeżeli Q jest formą kwadratową na \mathbf{R}^n , tzn. wyrażeniem postaci

$$Q(v) = \sum_{i=1}^n a_{ii} v_i^2 + \sum_{i < j} 2a_{ij} v_i v_j,$$

to istnieje taka baza ON B , że

$$Q(v) = \sum_{i=1}^n b_{ii}(v'_i)^2,$$

gdzie $\begin{pmatrix} v'_1 \\ v'_2 \\ \vdots \\ v'_n \end{pmatrix} = [v]_B$. Ponadto wartości b_{ii} to dokładnie wartości własne macierzy (a_{ij}) .

W szczególności oznacza to że każda „kwadryka” lub „hiperpowierzchnia kwadratowa” w \mathbf{R}^n , tzn. zbiór rozwiązań równania postaci

$$\sum_{i=1}^n a_{ii}v_i^2 + \sum_{i < j} 2a_{ij}v_iv_j = c,$$

ma w pewnym prostokątnym układzie współrzędnych postać kanoniczną

$$\sum_{i=1}^n b_{ii}(v'_i)^2 = c,$$

przy czym można założyć, że $c \in \{0, 1\}$.

Ponieważ ortogonalna zmiana współrzędnych odpowiada nałożeniu macierzy ortogonalnej, a to są dokładnie macierze izometrii, wnioskujemy stąd, że każda hiperpowierzchnia kwadratowa jest izometryczna z hiperpowierzchnią w postaci kanonicznej. Dzięki temu nietrudno zbudować pełną klasyfikację kwadryk (z dokładnością do izometrii), analogiczną do klasyfikacji krzywych stożkowych na płaszczyźnie i powierzchni kwadratowych w \mathbf{R}^3 .

Jeżeli ponadto zaniedbamy skalowanie osi, to „kształt” kwadryki jest całkowicie zdeterminowany przez sygnaturę odpowiadającą jej formy kwadratowej oraz znak drugiej strony równania. W szczególności z tego punktu widzenia każda kwadryka „wygląda jak” kwadryka w przestrzeni $n = (p+q+r)$ -wymiarowej zadana równaniem postaci

$$\sum_{i=1}^p x_i^2 - \sum_{i=p+1}^{p+q} x_i^2 = 1$$

lub takim samym, ale z prawą stroną równą 0.

Przykład 11.20. Niepusta kwadryka zadana równaniem $Q(v) = 1$ jest ograniczona wtedy i tylko wtedy gdy Q jest dodatnio określona — przykład widzieliśmy na ćwiczeniach. (W przeciwnym razie zawiera prostą w kierunku wektora z przestrzeni zerowej lub hiperbolę zawartą w płaszczyźnie rozpiętej przez wektory v_1, v_2 takie że $Q(v_1) > 0 > Q(v_2)$.)

W tym wypadku z postaci kanonicznej widzimy, że mamy do czynienia z (wielowymiarową) elipsoidą, liniowym obrazem sfery.

11.4 Macierz Grama i objętość

Definicja 11.21. *Macierz Grama* (lub *Gramian*) $G(v_1, \dots, v_k)$ układu v_1, v_2, \dots, v_k w przestrzeni unitarnej/euklidesowej to macierz której ij -ty wyraz to $\langle v_i, v_j \rangle$.

Uwaga 11.22. W szczególności jeżeli $B = v_1, \dots, v_k$ jest bazą, to jej macierz Grama to po prostu macierz standardowego iloczynu skalarnego względem tej bazy.

Uwaga 11.23. Jeżeli $v_1, \dots, v_k \in \mathbf{R}^n / \mathbf{C}^n$, to $G(v_1, \dots, v_k) = A^*A$, gdzie A to macierz o kolumnach v_j .

Ogólnie, dla dowolnej przestrzeni unitarnej/euklidesowej V i bazy ON B , $G(v_1, \dots, v_k) = A^*A$, gdzie A to macierz o kolumnach $[v_j]_B$.

Sprawdziliśmy wcześniej, że macierze postaci A^*A są nieujemnie określone. Stąd:

Wniosek 11.24. *Macierz Grama jest hermitowska/symetryczna i nieujemnie określona.*

Stwierdzenie 11.25. *Jeżeli $A \in M_{n \times m}(\mathbf{C})$ jest dowolną macierzą, to $\text{rk} A = \text{rk}(A^*A) = \text{rk}(AA^*)$.*

Dowód. Pokażmy że $\text{rk} A = \text{rk} A^*A$. Wystarczy pokazać, że $\ker F_A = \ker F_{A^*A}$ (bo F_A i F_{A^*A} mają tę samą dziedzinę). Weźmy dowolny $v \in \mathbf{C}^m$. Wtedy

$$\begin{aligned} \underline{Av = 0} &\implies \underline{A^*Av = 0} \implies \langle v, A^*Av \rangle = 0 \iff \langle Av, Av \rangle = 0 \iff \\ &\iff |Av| = 0 \iff \underline{Av = 0}. \end{aligned}$$

czyli $Av = 0 \implies A^*Av = 0 \implies Av = 0$. □

Wniosek 11.26. Układ v_1, \dots, v_k jest liniowo niezależny wtedy i tylko wtedy gdy $\det G(v_1, \dots, v_k) \neq 0$. Dokładniej, rząd macierzy Grama jest równy wymiarowi $\text{Lin}(v_1, \dots, v_k)$.

Dowód. Wystarczy zastosować poprzednie stwierdzenie dla $A = ([v_1]_B, \dots, [v_k]_B)$, gdzie B jest bazą ON V . \square

Chcemy za pomocą macierzy Grama zdefiniować objętość równoległościanu rozpiętego przez wektory v_1, \dots, v_k .

Stwierdzenie 11.27. Weźmy dowolne wektory liniowo niezależne v_1, \dots, v_k w przestrzeni euklidesowej V i bazę ON B dla $\text{Lin}(v_1, \dots, v_k)$. Wtedy następujące liczby są równe:

1. $V_k(v_1, \dots, v_k) = |\det([v_1]_B, [v_2]_B, \dots, [v_k]_B)|$,
2. $V'_k(v_1, \dots, v_k) = V'_{k-1}(v_1, \dots, v_{k-1}) \cdot d(v_k, \text{Lin}(v_1, \dots, v_{k-1}))$, gdzie $V'_1(v_1) = |v_1|$,
3. $\sqrt{\det(G(v_1, \dots, v_k))}$.

Dowód. Oznaczmy przez A macierz z (1).

(1) i (3): Rozumując jak powyżej widzimy, że $G(v_1, \dots, v_k) = A^*A$, skąd:

$$\det G(v_1, \dots, v_k) = \det(A^*A) = \det(A^*) \det(A) = \overline{\det A} \det A = |\det A|^2.$$

(1) i (2): Dowodzimy przez indukcję względem k . Dla $k = 1$ teza jest oczywista. Załóżmy że teza zachodzi dla $k - 1$, tzn. $V'_{k-1}(v_1, \dots, v_{k-1}) = V_{k-1}(v_1, \dots, v_{k-1})$.

Z równości (1) i (3) wynika, że wartość w (1) nie zależy od wyboru bazy ON. Możemy założyć więc, że $B' = (b_1, \dots, b_{k-1})$ jest bazą ON $W := \text{Lin}(v_1, \dots, v_{k-1})$ (a b_k dopełnia B' do bazy ON $\text{Lin}(v_1, \dots, v_k)$). Wtedy:

$$d(v_k, W) = |v - \pi_W(v)| = |\langle b_k, v_k \rangle b_k| = |\langle b_k, v_k \rangle|.$$

Oznaczmy $c := \langle b_k, v_k \rangle$ i $A' := ([v_1]_{B'}, \dots, [v_{k-1}]_{B'})$. Wtedy:

$$\det A = \det([v_1]_B, \dots, [v_{k-1}]_B, [v_k]_B) = \det([v_1]_{B'}, \dots, [v_{k-1}]_{B'}, [v_k - \pi_W(v_k)]_B) = \begin{vmatrix} A' & 0 \\ * & c \end{vmatrix}.$$

Z eliminacji Gaussa łatwo wnioskujemy, że ten ostatni wyznacznik to $c \cdot \det A'$. Ponieważ $|\det A'|$ to z definicji $V_{k-1}(v_1, \dots, v_{k-1})$ i $|c| = d(v_k, W)$, stąd łatwo wynika teza. \square

Uwaga 11.28. Liczby w (2) i (3) są równe nawet wtedy, gdy v_1, \dots, v_k są lz (wtedy obydwie są równe 0). Liczba (1) nie ma wtedy sensu, bo macierz po prawej stronie nie jest kwadratowa.

Definicja 11.29. Liczbę $V_k(v_1, \dots, v_k) = \sqrt{\det G(v_1, \dots, v_k)}$ nazywamy (k -wymiarową) *objętością* równoległościanu rozpiętego przez v_1, \dots, v_k (nawet wtedy, gdy są one lz).

(Wzór (2) pokazuje, że ta definicja faktycznie ma intuicyjny sens geometryczny.)

Przykład 11.30. Chcemy policzyć odległość od płaszczyzny $\text{Lin} \left(v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} \right)$

wektora $v_3 = \begin{pmatrix} 1 \\ 3 \\ 0 \\ 1 \end{pmatrix}$. Policzmy iloczyny skalarne:

$$\begin{aligned} \left\langle \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} \right\rangle &= 2 & \left\langle \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} \right\rangle &= -1 & \left\langle \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 0 \\ 1 \end{pmatrix} \right\rangle &= -2 \\ \left\langle \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} \right\rangle &= 7 & \left\langle \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 0 \\ 1 \end{pmatrix} \right\rangle &= 8 & \left\langle \begin{pmatrix} 1 \\ 3 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 0 \\ 1 \end{pmatrix} \right\rangle &= 11 \end{aligned}$$

Stąd mamy macierz Grama:

$$G(v_1, v_2, v_3) = \begin{pmatrix} 2 & -1 & -2 \\ -1 & 7 & 8 \\ -2 & 8 & 11 \end{pmatrix}$$

i możemy wyliczyć odległość

$$\begin{aligned}
 d(v_3, \text{Lin}(v_1, v_2)) &= \frac{V_3(v_1, v_2, v_3)}{V_2(v_1, v_2)} = \frac{\begin{vmatrix} 2 & -1 & -2 \\ -1 & 7 & 8 \\ -2 & 8 & 11 \end{vmatrix}}{\begin{vmatrix} 2 & -1 \\ -1 & 7 \end{vmatrix}} = \\
 &= \sqrt{\frac{14 \cdot 11 + 16 + 16 - 8 \cdot 16 - 11 - 2 \cdot 14}{13}} = \sqrt{\frac{9 \cdot 14 - 6 \cdot 16 - 11}{13}} = \\
 &= \sqrt{\frac{3 \cdot 14 - 12 - 11}{13}} = \sqrt{\frac{19}{13}}
 \end{aligned}$$

Przykład 11.31. Chcemy policzyć wartość

$$\inf_{a, b \in \mathbf{R}} \int_0^1 |t^2 - at - b|^2 dt.$$

Zauważmy że ta liczba to kwadrat odległości x^2 od $\text{Lin}(x, 1) \leq V = \mathbf{R}_2[x]$ z iloczynem skalarnym $\langle P, Q \rangle = \int_0^1 P(t)Q(t) dt$.

Łatwo policzyć, że $\langle x^i, x^j \rangle = \int_0^1 t^{i+j} dt = \frac{1}{i+j+1}$. Stąd

$$V_2(1, x)^2 = \begin{vmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{vmatrix} = \frac{1}{3} - \frac{1}{4} = \frac{1}{12},$$

oraz

$$V_3(1, x, x^2)^2 = \begin{vmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{vmatrix} = \frac{1}{2160},$$

czyli odległość to

$$d(x^2, \text{Lin}(1, x)) = \frac{V_3(1, x, x^2)}{V_2(1, x)} = \sqrt{\frac{12}{2160}} = \sqrt{\frac{1}{180}} = \frac{1}{6\sqrt{5}}.$$

W szczególności wynika stąd:

$$\inf_{a, b \in \mathbf{R}} \int_0^1 |t^2 - at - b|^2 dt = \frac{1}{180}$$

Ćwiczenie 11.32. Jeżeli macierz kwadratowa $A \in M_{n \times n}(\mathbb{C})$ jest dodatnio półokreślona, to jest macierzą Grama.

Wniosek 11.33. Niech $A \in M_{n \times n}(\mathbb{C})$ będzie hermitowska/rzeczywista symetryczna. Wtedy:

1. A jest macierzą Grama (pewnego układu n wektorów) \iff jest dodatnio półokreślona,
2. A jest macierzą Grama pewnego liniowo niezależnego układu wektorów \iff jest dodatnio określona.

Dowód. Wiemy, że jeżeli A jest macierzą Grama, to jest dodatnio półokreślona, więc (1) wynika z ćwiczenia.

Założmy że A jest dodatnio określona. Wtedy z (1) mamy $A = G(v_1, \dots, v_n)$. Wiemy, że odwracalność A jest równoważna z liniową niezależnością v_1, \dots, v_n , a dla dodatnio półokreślonej macierzy jest równoważna z dodatnią określonością (z tym że 0 nie jest wartością własną). \square

Wniosek 11.34. Jeżeli $F \in \text{End}(V)$, gdzie V jest euklidesowa lub unitarna i $\dim V = n$, to dla każdego wektorów v_1, \dots, v_n mamy

$$V_n(F(v_1), \dots, F(v_n)) = |\det F| V_n(v_1, \dots, v_n).$$

Dowód. Jeżeli v_1, \dots, v_n są lz, to obydwie strony są zerowe i teza zachodzi. Założmy więc że są lnz.

Oznaczmy przez $A := ([v_1]_B, \dots, [v_n]_B)$, $A' := ([F(v_1)]_B, \dots, [F(v_n)]_B)$, $M := m_B(F)$. Wtedy $A' = MA$, skąd

$$\begin{aligned} V_n(F(v_1), \dots, F(v_n)) &= |\det A'| \\ &= |\det(MA)| \\ &= |\det M \det A| \\ &= |\det M| |\det A| \\ &= |\det F| V_n(v_1, \dots, v_n). \end{aligned} \quad \square$$

Przykład 11.35 (całkowanie przez podstawienie). Mamy odwracalne przekształcenie liniowe $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$, funkcję ciągłą $f: \mathbb{R}^n \rightarrow \mathbb{R}$, a także podzbiór $V \subseteq \mathbb{R}^n$ który jest sumą równoległościów.

Wtedy

$$\begin{aligned}
 \int_{F[V]} f(v) dv &= \lim_{\mathcal{P} \text{ podział } F[V]} \sum_{P \in \mathcal{P}} V_n(P) f(v_P) \\
 &= \lim_{\mathcal{P} \text{ podział } V} \sum_{P \in \mathcal{P}} V_n(F[P]) f(F(v_P)) \\
 &= \lim_{\mathcal{P} \text{ podział } V} \sum_{P \in \mathcal{P}} |\det F| V_n(P) f(F(v_P)) \\
 &= \int_V |\det F| f(F(v)) dv.
 \end{aligned}$$

gdzie $\lim_{\mathcal{P} \text{ podział } F[V]}$ przebiega skończone zbiory \mathcal{P} n -wymiarowych prostopadłościanów o coraz mniejszych średnicach wypełniających $F[V]$, a v_P jest dowolnym wektorem z P .

Następnie możemy ten wzór poprawić:

- V nie musi być zbiorem, który nie jest sumą prostopadłościanów, ale tylko dobrze się takimi sumami przybliża (np. jest jordanowsko mierzalny),
- f nie musi być ciągła, a tylko (na przykład) riemannowsko całkowalna,
- F nie musi być liniowa, a tylko dobrze przybliżać się przez złożenie odwracalnej funkcji liniowej i translacji (np. F może być klasy C^1 z odwracalną pochodną) — wtedy $\det F$ należy zastąpić jakobianem F . V należy ponadto wybrać tak, by $F|_V$ była 1-1.

(Szczegóły: wykład analiza matematyczna/geometria różniczkowa.)

Dodatek A

Przestrzenie nieskończone wymiarowe

A.1 Twierdzenie o wymiarze dla przestrzeni nieskończone wymiarowych

Do udowodnienia w ogólności twierdzenia o wymiarze potrzebujemy następującego twierdzenia (pozostawiamy je bez dowodu; wynika on stosunkowo łatwo z twierdzenia Tichonowa z topologii oraz ze skończonego twierdzenia Halla o kojarzeniu małżeństw — o tym samym sformułowaniu, z tym że z założeniem, że zbiór A jest skończony).

Twierdzenie A.1 (Halla o kojarzeniu małżeństw). *Założmy że A jest zbiorem i dla każdego $a \in A$ mamy dany skończony zbiór B_a , taki że dla każdego skończonego $A_0 \subseteq A$ zachodzi $|A_0| \leq \bigcup_{a \in A_0} B_a$.*

*Wtedy istnieje funkcja różnowartościowa przekształcająca każdy element $a \in A$ na pewien element odpowiadającego mu B_a .*¹

Wniosek A.2 (Twierdzenie Löwiga o wymiarze). *Jeżeli V jest dowolną przestrzenią liniową i A, B są bazami V , to A i B są równoliczne.*

Dowód. Wystarczy pokazać, że $|A| \leq |B|$, tzn. że istnieje funkcja różnowartościowa z $A \rightarrow B$: z symetrii wynikać będzie, że też $|B| \leq |A|$, a więc z twierdzenia Cantora-Bernsteina $|A| = |B|$.

¹Twierdzenie Halla w tej formie nie jest twierdzeniem ZF, ale jest istotnie słabsze od pewnika wyboru.

Ponieważ B jest bazą, każdy element $a \in A$ przedstawia się jednoznacznie jako kombinacja liniowa (skończenie wielu) elementów B . Niech $B_a \subseteq B$ składa się dokładnie z tych elementów, które występują w niej z niezerowymi współczynnikami.

Pokażemy że spełniają one założenia twierdzenia Halla. Istotnie, każdy B_a jest skończony, i jeżeli $A_0 \subseteq A$ jest skończony, to $B_0 = \bigcup_{a \in A_0} B_a$ spełnia $A_0 \subseteq \text{Lin}(B_0)$. Ponieważ A_0 i B_0 są liniowo niezależne, B_0 jest (skończoną) bazą $V = \text{Lin}(B_0)$, a A_0 jego liniowo niezależnym podzbiorem, więc na mocy Twierdzenia 2.16, $|A_0| \leq |B_0|$.

Stąd na mocy twierdzenia Halla istnieje funkcja różnowartościowa $A \rightarrow \bigcup_{a \in A} B_a \subseteq B$, co należało dowieść. \square

Wniosek A.3. *Izomorfizm zachowuje wymiar przestrzeni liniowej. Jeżeli $V_1 \cong V_2$, i B_1, B_2 są bazami B_1 i B_2 odpowiednio, to mają tyle samo elementów.*

Dowód. Izomorfizm przekształca bazy na bazy (bo zachowuje kombinacje liniowe i liniową niezależność), więc jeżeli $F : V_1 \rightarrow V_2$ jest izomorfizmem, to $F[B_1]$ ma tyle samo elementów co B_1 (bo F jest 1-1) i jest bazą. Stąd łatwo wynika teza. \square

A.2 Lemat Steinitza o wymianie

Dowód twierdzenia o wymiarze (Wniosku A.2) korzystał z lematu Steinitza dla skończonych baz. Aby uzyskać lemat Steinitza dla nieskończonych baz, postąpimy odwrotnie: wykorzystamy twierdzenie o wymiarze (oraz lemat Kuratowskiego-Zorna).

Twierdzenie A.4 (Lemat Steinitza o wymianie). *Założmy że V jest przestrzenią liniową o bazie B , zaś $A \subseteq V$ jest liniowo niezależny.*

Wtedy istnieje $C \subseteq B$ równoliczny z A taki że $A \cup (B \setminus C)$ jest bazą V .

Dowód. Rozważmy porządek częściowy złożony z podzbiorów $B_0 \subseteq B$ takich że $B_0 \cap A = \emptyset$ i $B_0 \cup A$ jest liniowo niezależny, uporządkowanych przez \subseteq .

Podobnie jak w dowodzie Twierdzenia 2.8, pokazujemy że porządek ten spełnia założenia lematu Kuratowskiego-Zorna, ma więc element maksymalny $B' \subseteq B$.

Z maksymalności B' wynika, że $B \subseteq \text{Lin}(A \cup B')$ i oczywiście $A \cup B'$ jest liniowo niezależny, więc jest on bazą. Pozostaje pokazać, że $C = B \setminus B'$ jest równoliczny z A .

Rozważmy $W = \text{Lin}(B')$, $W_1 = \text{Lin}(A)$, $W_2 = \text{Lin}(C) \leq V$ i rozważmy odwzorowanie $F: V \rightarrow V/W$. Z założeń łatwo wynika że $W \cap W_1 = W \cap W_2 = \{0\}$, a ponadto $F[W_1] = F[W_1 + W] = F[V] = F[W_2 + W] = F[W_2] = V/W$, więc $F|_{W_1}$ i $F|_{W_2}$ są liniowymi bijekcjami, a więc są izomorfizmami.

Stąd W_1 i W_2 są izomorficzne, więc ich bazy mają tyle samo elementów. Ale A jest bazą W_1 i C jest bazą W_2 , czyli A i C są równoliczne, co należało dowieść. \square

Dodatek B

Rzeczywiste twierdzenie Jordana

Dodatek C

Twierdzenie spektralne dla endomorfizmów normalnych

W tym dodatku V jest przestrzenią unitarną lub euklidesową.

Definicja C.1. Mówimy że macierz kwadratowa $A \in M_{n \times n}(\mathbb{C})$ jest normalna, jeżeli $AA^* = A^*A$.

Mówimy że $F \in \text{End}(V)$ jest normalny, jeżeli $F^* \circ F = F \circ F^*$.

Przykłady C.2. • Endomorfizmy/operators unitarne są normalne:
 $U^*U = I = UU^*$.

- Endomorfizmy/macierze samosprężone są normalne: $AA^* = AA = A^*A$.
- Zespolone wielokrotności normalnych endomorfizmów/macierzy są normalne: $(zA)^*(zA) = |z|^2 A^*A = |z|^2 AA^* = (zA)(zA)^*$.
- Suma dwóch normalnych endomorfizmów/macierzy ogólnie *nie* jest normalna: np. $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ i $\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$ są normalne, ale

$$\begin{pmatrix} 0 & i \\ i & 1 \end{pmatrix}^* \begin{pmatrix} 0 & i \\ i & 1 \end{pmatrix} = \begin{pmatrix} 1 & -i \\ i & 2 \end{pmatrix}$$

podczas gdy

$$\begin{pmatrix} 0 & i \\ i & 1 \end{pmatrix} \begin{pmatrix} 0 & i \\ i & 1 \end{pmatrix}^* = \begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix}$$

(ale nietrudno pokazać, że $A + B$ jest normalna, gdy A, B są normalne i $AB^* = A^*B$).

Uwaga C.3. Można pokazać, że macierz A jest normalna wtedy i tylko wtedy gdy istnieją samosprężone $A_1 = A_1^*, A_2 = A_2^*$ takie że $A = A_1 + iA_2$ i $A_1A_2 = A_2A_1$. Podobnie dla endomorfizmów.

Uwaga C.4. $F \in \text{End}(V)$ jest normalny wtedy i tylko wtedy gdy dla pewnej/każdej bazy ON B macierz $m_B(F)$ jest normalna:

$$\begin{aligned} FF^* = F^*F &\iff m_B(FF^*) = m_B(F^*F) \iff \\ &\iff m_B(F)m_B(F^*) = m_B(F^*)m_B(F) \iff m_B(F)m_B(F)^* = m_B(F)^*m_B(F). \end{aligned}$$

Wniosek C.5. Jeżeli U jest unitarna, to dla każdej macierzy A mamy równoważność: A jest normalna $\iff UAU^*$ jest normalna.

(To łatwo też sprawdzić bezpośrednim rachunkiem.)

Twierdzenie C.6 (Twierdzenie spektralne dla endomorfizmów normalnych).
(wersja unitarna)

- $F \in \text{End}(V)$ jest normalny \iff diagonalizuje się w bazie ON.
- $A \in M_{n \times n}(\mathbb{C})$ jest normalna \iff istnieje unitarna macierz U i diagonalna macierz D taka że $A = UDU^*$.

Dowód. Druga część łatwo wynika z pierwszej.

Pierwsza część: implikacja z prawej do lewej jest łatwa: jeżeli F diagonalizuje się w bazie ON, tj. $m_B(F) = D$ dla pewnej bazy ON B i macierzy diagonalnej D , to $m_B(F^*) = D^* = \bar{D}$, więc F^* diagonalizuje się w tej samej bazie, a macierze diagonalne są przemiennie.

Implikacja w przeciwną stronę: indukcja względem $\dim V$. Przypadek $\dim V = 1$ jest oczywisty. Załóżmy więc że $\dim V > 1$.

Niech $b_1 \in V$ będzie jednostkowym wektorem własnym dla F odpowiadający wartości własnej λ_1 . Uzupełnijmy go dowolnie do bazy ON $B' = (b_1, b'_2, \dots, b'_n)$.

Wtedy mamy

$$A := m_{B'}(F) = \begin{pmatrix} \lambda_1 & v^\top \\ 0 & N \end{pmatrix}$$

dla pewnego wektora $v \in \mathbf{C}^{n-1}$ i macierzy $N \in M_{(n-1) \times (n-1)}(\mathbf{C})$ i macierz A jest normalna. Zbadajmy wyraz w lewym górnym rogu macierzy:

$$A^*A = \begin{pmatrix} \bar{\lambda}_1 & 0 \\ \bar{v} & N^* \end{pmatrix} \begin{pmatrix} \lambda_1 & v^\top \\ 0 & N \end{pmatrix} = \begin{pmatrix} \lambda_1 & v^\top \\ 0 & N \end{pmatrix} \begin{pmatrix} \bar{\lambda}_1 & 0 \\ \bar{v} & N^* \end{pmatrix} = AA^*.$$

Z wyrażenia po lewej jest on równy $\bar{\lambda}_1 \lambda_1 = |\lambda_1|^2$, zaś z wyrażenia po prawej $-\bar{\lambda}_1 \lambda_1 + v^\top \bar{v} = |\lambda_1|^2 + |\bar{v}|^2$. Stąd $|\bar{v}| = |v| = 0$, czyli $v = 0$.

Z powyższego łatwo wynika że $A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & N \end{pmatrix}$, więc też

$$A^*A = \begin{pmatrix} |\lambda_1|^2 & 0 \\ 0 & N^*N \end{pmatrix} = \begin{pmatrix} |\lambda_1|^2 & 0 \\ 0 & NN^* \end{pmatrix} = AA^*,$$

więc macierz N jest normalna, a ponadto $W = \text{Lin}(b'_2, \dots, b'_n)$ jest F -niezmiennicze, czyli $F|_W \in \text{End}(W)$ jest normalny (bo N jest jego macierzą). Teza wynika wtedy wprost z założenia indukcyjnego. \square

Indeks

- algorytm
 - Gram-Schmidta, *zob.*
 - ortogonalizacja Gram-Schmidta
 - Lagrange'a, 139
- alternowanie, 63
- antyliniowość, 147
- antysymetryczność
 - formy dwuliniowej, *zob.* forma antysymetryczna
 - funkcji, 64
- automorfizm, 31
- baza
 - przestrzeni liniowej, 15
- całkowanie przez podstawienie, 183
- ciało, 5
 - algebraicznie domknięte, 106
 - skończone, 6
- ciąg dokładny, 32
- cykl, 66
- diagonalizacja
 - endomorfizmu, 95
 - formy dwuliniowej, 127
 - macierzy, 96
- diagonalizowalność
 - endomorfizmu, 95
 - macierzy, 96
- dodatnia określoność
 - formy hermitowskiej, 152
 - formy kwadratowej, 138
 - macierzy rzeczywistej, 124
 - rzeczywistej formy symetrycznej, 125
- dopełnialność, 89
- dopełniczość, *zob.* podprzestrzeń dopełnicza
- dopełnienie
 - algebraiczne, 81
 - ortogonalne, 129, 152
- długość wektora, 142
- eliminacja Gaussa, 48, 54
- endomorfizm, 27
 - nieujemnie określony, 158
 - normalny, 189
 - ortogonalny, 158
 - samosprężony, 158
 - sprężony, 157
 - unitarny, 158
- epimorfizm, 39
- forma
 - antysymetryczna, 120

- dwuliniowa, 120
- hermitowska, 148
- kwadratowa, 137
- Lorentza, 128
- skośnie symetryczna, *zob.*
 - forma antysymetryczna
- symetryczna, 120
- zadana przez macierz, 121
- funkcja liniowa, *zob.*
 - przekształcenie liniowe
- Gramian, *zob.* macierz Grama
- hermitowska symetryczność, 147
- homomorfizm (przestrzeni liniowych), *zob.*
 - przekształcenie liniowe
- iloczyn
 - skalary, 125, 154
 - standardowy, 120, 147
 - tensorowy, 108
- inwersja, 67
- izometria, 174
- izomorfizm, 22
- Jordana
 - baza, 97
 - klatka, 97
 - postać, 97
 - rzeczywista, 114
 - twierdzenie, 97
- jordana
 - klatka
 - rzeczywista, 113
- jądro, 27
- kombinacja liniowa, 12
- kompleksyfikacja
 - przekształcenia liniowego, 109
 - przestrzeni liniowej, 108
 - rzeczywistej formy
 - symetrycznej, 148
- kryterium Sylvestera, 128
- krótki ciąg dokładny, 32
- kąt nieskierowany między wektorami, 143
- lemat
 - Kuratowskiego-Zorna, 16
 - Steinitza o wymianie, 19
- liczba inwersji, 66
- liniowa
 - niezależność
 - podprzestrzeni, 86
 - wektorów, 13
 - zależność, 13
- liniowe domknięcie, *zob.* otoczka liniowa
- macierz
 - dołączona, 81
 - formy dwuliniowej, 121
 - formy hermitowskiej, 149
 - formy kwadratowej, 138
 - Gram, 179
 - hermitowska, 149
 - nieosobliwa, 77
 - normalna, 189
 - odwrotna, 60
 - ortogonalna, 151
 - osobliwa, 77
 - przekształcenia liniowego, 58
 - samosprzężona, *zob.* macierz hermitowska
 - schodkowa, *zob.* postać schodkowa macierzy

- sprężona, 149
- unitarna, 151
- metoda najmniejszych kwadratów, 168
- minor, 77
 - główny, 127
 - wiodący, 127
- monomorfizm, 39
- morfizm (przestrzeni liniowych),
 - zob. przekształcenie liniowe
- n -liniowość, 63
- nieporządek, zob. inwersja
- nierówność
 - Bessela, 146, 154
 - Cauchy'ego-Schwarza, 143, 154
 - trójkąta, 143, 154
- niezdegenerowana
 - forma hermitowska, 152
 - forma symetryczna, 133
- objętość, 181
- obraz, 27
- obróć (wokół hiperprostej), 175
- odbicie (w hiperpłaszczyźnie), 174
- odwracalność
 - macierzy, 60
 - przekształcenia liniowego, 61
- operacje
 - kolumnowe, 44
 - wierszowe, 44
- ortogonalizacja
 - Gram-Schmidta, 126
 - Lagrange'a, 130
- ortogonalność
 - macierzy, 124
 - pary wektorów, 125, 152
 - układu wektorów, 125, 152
- ortonormalność
 - układu wektorów, 125, 152
- otoczka liniowa, 11
- permutacja, 65
- podprzestrzeń
 - dopełnicza, 88
 - liniowa, 9
 - niezmiennicza, 89
 - zerowa, 133
- postać
 - kanoniczna
 - endomorfizmu
 - ortogonalnego, 174
 - formy kwadratowej, 178
 - schodkowa macierzy, 48
 - schodkowa zredukowana macierzy, 48
- prawo bezwładności Sylwestera, 133, 162
- proces Gram-Schmidta, zob. ortogonalizacja Gram-Schmidta
- produkt, 33
- przekształcenie
 - dualne, 36
- przestrzeń
 - bidualna, 40
 - dualna, 34
 - euklidesowa, 142
 - funkcji, 8
 - funkcji ciągłych, 8
 - ilorazowa, 38
 - liniowa, 7
 - macierzy, 82

- pierwiastkowa, 99
- unitarna, 154
- wektorowa, 7
- wielomianów, 8
- własna, 92
- przeszktałenie
 - liniowe, 25
- półtoraliniowość, 148
- rozkład
 - biegunowy, 169
 - singularny, 165
 - spektralny, 163
- rzut ortogonalny, 144, 154
- rząd
 - macierzy, 44
 - przekształcenia liniowego, 30
- równanie różniczkowe, 117
- skalar, 7
- spektrum, 92
- sprzężenie, 108
- suma
 - kompleksowa, 10
 - prosta
 - wewnętrzna, 85
 - zewnętrzna, 33
- sygnatura
 - formy hermitowskiej, 162
 - rzeczywistej formy
 - kwadratowej, 139
 - rzeczywistej formy symetrycznej, 133
- symetryczność
 - formy dwuliniowej, *zob.* forma symetryczna
 - macierzy, 123
- szift, 26
- tożsamość Parsevala, 146, 154
- translacja, 174
- transpozycja, 65
- twierdzenie
 - Jordana, 104, 106
 - rzeczywiste, 114
 - Kroneckera-Capelliego, 44
 - Lagrange'a, 130
 - o indeksie, 30
 - o istnieniu bazy, 16
 - o izomorfizmie, 39
 - o rzędzie, 30
 - spektralne, 159
 - wariant normalny, 190
 - wariant ortogonalny, 173
 - wariant unitarny, 172
 - zasadnicze algebry, 106
- warstwa, 37
- wartość singularna, 167
- wartość własna, 92
- wektor, 7
 - izotropowy, 129
 - jednostkowy, 125
 - singularny, 168
 - unormowany, 125
 - własny, 92
 - znormalizowany, 152
- widmo, *zob.* spektrum
- wieloliniowość, *zob.* *n*-liniowość
- wielomian charakterystyczny, 84
- współrzędne wektora w bazie, 24
- wymiar (przestrzeni liniowej), 20
- wyznacznik
 - endomorfizmu, 84
 - macierzy, 71
 - Vandermonde'a, 75
 - wektorów, 68

wzór

Cramera, 80

na odwrócenie macierzy, 81

na rzut ortogonalny, 144

na wyliczenie wartości formy
w ustalonej bazie, 122

na zmianę bazy

dla formy dwuliniowej, 123

dla endomorfizmu, 59

dla formy hermitowskiej,

151

na zmianę bazy dla

endomorfizmu

dla baz ON, 156

polaryzacyjny, 138, 153, 176

znak permutacji, 67

śląd, 84

Indeks symboli i oznaczeń

- $<$, zob. podprzestrzeń liniowa
 \leq , zob. podprzestrzeń liniowa
 \oplus , 33
 \perp , 125, 152
 \times , 33
 $\langle -, - \rangle$, 120
 \cong , zob. izomorfizm
 $[v]_B$, zob. współrzędne wektora w bazie
 A^* , 149
 A^{-1} , 60
 $\text{adj}(A)$, 81
 A^V , 81
 $C(\mathbf{R})$, 8
 $\chi_F(x)$, zob. wielomian charakterystyczny
 δ_a , 15
 \det , 68, 71, zob. wyznacznik endomorfizmu
 \dim , zob. wymiar
 $\text{End}(V)$, 27
 F^* , zob. przesłtalenie dualne, 157
 $F_{\mathbf{C}}$, zob. kompleksyfikacja przesłtalenia liniowego
 φ_A , 121
 $\varphi_{\mathbf{C}}$, 148
 \mathbf{F}_p , 6
 $\text{Hom}(V, W)$, 27
 $\text{Hom}(V, W)$, 82
 im , 27
 K_{1-3} , 44
 $K[X]$, 8
 K^A , 8
 \ker , 27
 $K_n[X]$, 8
 Lin , zob. otoczka liniowa
 $M_{k \times n}(K)$, 82
 $m_B(F)$, 59
 $m^{BB}(\varphi)$, zob. macierz formy dwuliniowej
 $m_C^B(F)$, 58
 perp , zob. dopełnienie ortogonalne
 $\mathbf{Q}[\sqrt{d}]$, 6
 rk , zob. rząd przesłtalenia liniowego, zob. rząd macierzy
 rng , 27

sgn , zob. znak permutacji

$\sigma(F)$, 92

S_n , 65

Span , zob. otoczka liniowa

tr , 84

V^* , zob. przestrzeń dualna

V^{**} , zob. przestrzeń bidualna

$v + W$, 37

V/W , zob. przestrzeń ilorazowa

$V_{\mathbb{C}}$, zob. kompleksyfikacja

V^{λ} , 99

V_{λ} , 92

$W_1 + W_2$, zob. suma kompleksowa

$W1-3$, 44