Algebra liniowa 2R

wersja beta

Tomasz Rzepecki

27 lutego 2022

Spis treści

Rozdział 1

Podstawowe pojęcia algebry liniowej

1.1 Ciała

Definicja 1.1. *Ciałem* nazywamy zbiór K wraz z działaniami $+, \cdot$ 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. dla każdego *niezerowego* elementu $x \in K$, istnieje element odwrotny $x^{-1} \in K$, tzn. taki że $x \cdot x^{-1} = 1$,

5. $0 \neq 1$.

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

Ćwiczenie 1.2. *Oznaczenia $-x$, x^{-1} są sensowne, tzn. jednoznacznie określają elementy K.*

Przykłady 1.3. (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 $\mathbf{F}_p = \{0, 1, 2, 3, \ldots, p-1\}$ z dodawaniem i mnożeniem modulo p jest ciałem[1],

(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.

Lemat 1.4. *Jeżeli K jest ciałem, to dla każdego $a \in K$ zachodzi* $0 \cdot a = 0$.

Dowód.

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

[1]jeszcze 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)

1.2 Przestrzenie i podprzestrzenie liniowe

Definicja 1.5. Ustalmy ciało K. *Przestrzenią liniową* (lub *wektorową*) nad K nazywamy zbiór V z wyróżnionym elementem $\vec{0} \in V$ (*wektorem zerowym*), oraz z działaniami $+_V : V \times V \to V$ oraz $\cdot_V : K \times V \to V$, spełniającymi warunki (*aksjomaty przestrzeni liniowej*):

1. $+_V$ jest przemienne, łączne, $\vec{0}$ jest jego elementem neutralnym, ma elementy przeciwne:
   - $v_1 +_V v_2 = v_2 +_V v_1$,
   - $v_1 +_V (v_2 +_V v_3) = (v_1 +_V v_2) +_V v_3$,
   - $v +_V \vec{0} = v$,
   - $v +_V (-v) = 0$
2. $\cdot$ 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. $\cdot$ jest rozdzielne względem $+$ (uwaga na indeksy przy plusach!):
   - $\alpha \cdot_V (v +_V w) = \alpha \cdot_V v +_V \alpha \cdot_V w$,
   - $(\alpha +_K \beta) \cdot_V v = \alpha \cdot_V v +_V \beta \cdot_V v$;
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 1.6. Podobnie jak dla ciał, elementy przeciwne są jedyne. Tak naprawdę ich istnienie, podobnie jak istnienie $\vec{0}$, jest konsekwencją pozostałych aksjomatów.

Przykłady 1.7.

(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 $A \to K$ 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.[2]

(e) Przestrzeń $C(\mathbf{R})$ wszystkich funkcji ciągłych $\mathbf{R} \to \mathbf{R}$ jest przestrzenią liniową nad $\mathbf{R}$.

Stwierdzenie 1.8.
- $0 \cdot v = \vec{0}$.
- *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,

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

Dowód jedyności jest podobny jak dowód jedyności elementu przeciwnego dla ciał. □

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 $+_V$).

---

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

Lemat 1.9. *Załóżmy 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). □

Definicja 1.10. Podzbiór $W \subseteq V$ spełniający warunki Lematu powyżej nazywamy *podprzestrzenią (liniową) V*, oznaczamy $W \leqslant V$ (lub czasami $W < V$).

Przykłady 1.11.
(a) proste przechodzące prez 0 w F^2 są podprzestrzeniami (rysunek: proste w $\mathbf{F}_2^2, \mathbf{F}_3^2$),

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

(c) zbiór funkcji różniczkowalnych jest podprzestrzenią w $C(\mathbf{R})$ (przestrzeni funkcji ciągłych $\mathbf{R} \to \mathbf{R}$),

(d) zbiór funkcji ciągłych jest podprzestrzenią $\mathbf{R}^{\mathbf{R}}$,

(e) zbiór funkcji takich że $\lim_{x \to x_0} f(x) = 0$ (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 1.12. *Jeżeli $W_1, W_2 \leqslant V$, to:*

- $W_1 \cap W_2 \leqslant V$,
- $W_1 + W_2 = \{ w_1 + w_2 \mid w_1 \in W_1, w_2 \in W_2 \} \leqslant V$ *(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 $a \in F, w_1 \in W_1, w_2 \in W_2$ 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. □

Stwierdzenie 1.13. *Jeżeli $W_1, W_1 \subseteq V$, to:*

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

Dowód. Ćwiczenie. □

1.3 Kombinacje liniowe i liniowa niezależność

Definicja 1.14. Niech V będzie przestrzenią liniową, a $A \subseteq V$ — dowolnym podzbiorem.

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

Stwierdzenie 1.15. $\mathrm{Lin}(A)$ *to dokładnie zbiór (wszystkich) wektorów postaci* $\sum_{k=1}^{n} \alpha_k v_k$*, gdzie* $\alpha_k \in F$*, 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 skończenie wielu).*

Dowód. Jest jasne, że $\mathrm{Lin}(A)$ 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 $\gamma_k = \begin{cases} \alpha_k & k \leqslant n \\ \beta_{k-n} & k > n \end{cases}$ i $u_k = \begin{cases} v_k & k \leqslant n \\ w_{k-n} & k > n \end{cases}$, a także $\alpha \cdot \sum_{k=1}^{n} \alpha_k v_k = \sum_{k=1}^{n} (\alpha\alpha_k) v_k$. □

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 $W = \mathrm{Lin}(A)$.

Jest jasne, że $A \subseteq W$, łatwo też zawuażyć, 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 $\gamma_k = \begin{cases} \alpha_k & k \leqslant n \\ \beta_{k-n} & k > n \end{cases}$ i $u_k = \begin{cases} v_k & k \leqslant n \\ w_{k-n} & k > n \end{cases}$, 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 W, 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 $W = \mathrm{Lin}(A)$. 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 $W = \mathrm{Lin}(A)$. □

Wniosek 1.16. $\mathrm{Lin}(A) = \bigcap\{\, W \leqslant V \mid A \subseteq W \,\}$

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

Definicja 1.17. *Kombinacja liniowa* wektorów $v_1, \ldots, v_n$ to element $\mathrm{Lin}(v_1, \ldots v_n)$ (czyli wektor postaci $\sum_{k=1}^{n}\alpha_k v_k$).

Przykłady 1.18. (a) $\mathrm{Lin}(\emptyset) = \{\vec{0}\}$.

(b) $\mathrm{Lin}(v) = \{\, \alpha v \mid \alpha \in F \,\}$.

(c) $\mathrm{Lin}\left(\left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbf{R}^2 \,\middle|\, xy = 1 \right\}\right) = \mathbf{R}^2$

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

Uwaga 1.19.
- $A \subseteq B$ implikuje $\mathrm{Lin}(A) \subseteq \mathrm{Lin}(B)$,
- $\mathrm{Lin}(\mathrm{Lin}(A)) = \mathrm{Lin}(A)$,
- $b \in \mathrm{Lin}(A)$ wtedy i tylko wtedy gdy $\mathrm{Lin}(A) = \mathrm{Lin}(A \cup \{b\})$.

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

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

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

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

Definicja 1.20. Układ wektorów $v_1, \ldots, v_n$ nazywamy *liniowo niezależnym* (w skrócie *lnz*), jeżeli $\sum_{k=1}^{n} \alpha_k v_k = 0$ zachodzi tylko gdy $\alpha_1 = \ldots = \alpha_n = 0$. 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 $\mathrm{Lin}(A) \neq \mathrm{Lin}(A \setminus \{a\})$ (żaden element A nie jest kombinacją liniową pozostałych). W przeciwnym wypadu mówimy że jest *liniowo zależny*

Rozdział 2

Bazy i wymiary

2.1 Baza przestrzeni liniowej

Twierdzenie 2.1. *Niech $B \subseteq V$. Następujące warunki są równoważne:*

1. *B jest lnz i* $\mathrm{Lin}(B) = V$,
2. *B jest maksymalnym zbiorem liniowo niezależnym,*
3. *każdy element V zapisuje się jednoznacznie jako kombinacja liniowa elementów B.*

Dowód. (1)⇒(2): wiemy że B jest liniowo niezależny. Z drugiej strony dla każdego $v \in V$ mamy $v \in \mathrm{Lin}(B)$, więc żaden zbiór $B \cup \{v\}$, $v \notin B$ nie jest liniowo niezależny, więc B jest maksymalny.

(2)⇒(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, o czym świadczy pewna kombinacja liniowa $\alpha \cdot v + \sum_{k=1}^{n} \alpha_k v_k = 0$. Gdyby $\alpha = 0$, to $\sum_{k=1}^{n} \alpha_k v_k = 0$, więc $\alpha_1 = \ldots = \alpha_k = 0$, co przeczy założeniu, że $\alpha \cdot v + \sum_{k=1}^{n} \alpha_k v_k$ świadczy o liniowej zależności. Wobec tego $\alpha \neq 0$, ale wtedy $v = \sum_{k=1}^{n} -\alpha^{-1} \alpha_k v_k$.

Aby pokazać jedyność, załóżmy 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)⇒(1): „zapisuje się” w oczywisty sposób implikuje $\mathrm{Lin}(B) = V$; 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. □

Definicja 2.2. *Baza* przestrzeni liniowej V to taki $B \subseteq V$ który spełnia warunki z powyższego twierdzenia.

Przykłady 2.3. (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 $\sum_a \alpha_a \delta_a = 0$ (czyli jest funkcją zerową), 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 $1 \notin \mathrm{Lin}\{\delta_a \mid a \in A\}$, bo 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 $f(k) = 0$, 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 2.4 (Lemat Kuratowskiego-Zorna)**.** *Jeżeli* $(P, \leqslant)$ *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* $\bar{l} \geqslant l$ *dla każdego* $l \in L$.

Wtedy P ma element maksymalny.

Dowód. Wstęp do matematyki. □

Twierdzenie 2.5 (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 $v_1, \ldots, v_n \in \bigcup L$, 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;
  - ustalmy dowolną taką n-kę $v_1, \ldots, v_n \in l^*$ (parami różnych elementów);
  - ponieważ $l^* = \bigcup L$, to dla każdego v_k możemy wybrać pewne $l_k \in L$ takie ze $v_k \in l_k$;
  - ponieważ L jest łańcuchem, pewien z l_k, powiedzmy l_m, jest największy z nich, czyli $l_m \supseteq l_1, \ldots, l_n$,
  - ponieważ $v_k \in l_k \subseteq l_m$, mamy $v_1, \ldots, v_n \in l_m$;
  - z tego i z liniowej niezależności l_m (a także z tego że $v_1, \ldots, v_n$ są parami różne) wynika, że $v_1, \ldots, v_n$ są liniowo niezależne.

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

[1] Blass 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.

Twierdzenie 2.6. *Niech V będzie przestrzenią liniową, a $N \subseteq G \subseteq V$ będą takie że:*

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

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 $\mathrm{Lin}(B) = V$.

Pokażemy najpierw, że $G \subseteq \mathrm{Lin}(B)$. Weźmy dowolny $v \in G$. Z z maksymalności B, $\mathrm{Lin}(B \cup \{v\})$ 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 $v \in \mathrm{Lin}(B)$.

Mamy zatem $G \subseteq \mathrm{Lin}(B)$. Kończymy stosując Uwagę 1.19: skoro $B \subseteq \mathrm{Lin}(B)$, to $\mathrm{Lin}\, G \subseteq \mathrm{Lin}\,\mathrm{Lin}\, B = \mathrm{Lin}\, B$, ale z założenia $\mathrm{Lin}\, G = V$. Z drugiej strony oczywiście $\mathrm{Lin}\, B \subseteq V$, co kończy dowód. □

Wniosek 2.7. *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 $G \subseteq V$ rozpina V (tzn. $V = \mathrm{Lin}(G)$), to G zawiera pewną bazę V. (Zastosuj poprzednie twierdzenie dla $N = \emptyset$.)*

2.2 Konstrukcja bazy

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

Stwierdzenie 2.8. *Załóżmy że $V = \mathrm{Lin}(v_1, \ldots, v_n)$. definiujmy rekurencyjnie ciąg podzbiorów V:*

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

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

Dowód. Najpierw przez łatwą indukcję pokazujemy że $v_k \in \mathrm{Lin}(B_k) \subseteq \mathrm{Lin}(B_n)$ (bo $B_k \subseteq B_n$.) Stąd $\mathrm{Lin}(B_n) \supseteq \mathrm{Lin}(v_1, \dots, v_n) = V$. 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 $v_{k+1} \notin \mathrm{Lin}(B_k)$, więc jeżeli $B_k = \{w_1, \dots, w_M\}$ i $\alpha \cdot v_{k+1} + \sum_{m=1}^{M} \alpha_m w_m = 0$, to musi być $\alpha = 0$, a z lnz B_k wynika wtedy, że pozostałe α_k też są zerowe, więc B_{k+1} jest lnz. □

Uwaga 2.9. Ze Stwierdzenia i jego dowodu wynika, że w sformułowaniu Stwierdzenia warunek „$v_{k+1} \notin \mathrm{Lin}(B_k)$” można zastąpić warunkiem „$\mathrm{Lin}(B_k \cup v_{k+1})$ jest lnz” lub warunkiem „$v_{k+1} \notin \mathrm{Lin}(v_1, \dots, v_k)$”.

Uwaga 2.10. Stwierdzenie powyżej prowadzi do bardziej konstruktywnego dowodu twierdzenia o istnieniu bazy dla przestrzeni rozpinanej przez skończony układ wektorów (bez wykorzystania lematu Kuratowskiego-Zorna).

Przykład 2.11. Wybieramy bazę w $V = \mathrm{Lin}(x^2, (1+x)^2, 2+4x, x^3, x^3+x) \leqslant \mathbf{R}[x]$

- $B_0 = \emptyset$,
- $B_1 = \{x^2\}$, 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 $1 \neq 2 \cdot 0$.

Uwaga 2.12. Jeżeli v jest niezerowym wektorem, to zbiór $\{v, v, v\} = \{v\}$ jest liniowo niezależny, ale ciąg v, v, v jest liniowo zależny!

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

2.3 Lemat Steinitza

Twierdzenie 2.13 (Lemat Steinitza o wymianie)**.** [2] *Załóżmy że B jest bazą V, a $a_1, \ldots, a_n$ jest liniowo niezależnym ciągiem wektorów. Wtedy:*

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

Wniosek 2.14 (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 $|B_1| \geqslant |B_2|$. Podobnie $|B_2| \geqslant |B_1|$, czyli $|B_1| = |B_2|$. □

Dowód. (przez indukcję pozaskończoną) Niech B, C będą bazami V. Ponumerujmy je jako $(b_\alpha)_{\alpha<\lambda}, (c_\alpha)_{\alpha<\varkappa}$. Możemy założyć bez zmniejszania ogólności, że $\lambda < \varkappa$. Jeżeli □

Definicja 2.15. *Wymiar* przestrzeni liniowej V nad K, oznaczany $\dim_K V$ (lub $\dim V$, kiedy K wynika z kontekstu) to moc jej bazy.

Przykłady 2.16. (a) W Przykładzie 2.11: pokazaliśmy, że $V = \mathrm{Lin}(x^2, (1+x)^2, 2+4x, x^3, x^3+x) \leqslant \mathbf{R}[x]$ ma bazę $B_5 = \{x^2, (1+x)^2, x^3, x^3+x\}$, więc $\dim V = 4$.

---

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

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

(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 nad $\mathbf{Q}$.

dowód Lematu Steinitza.

Claim. 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 $c \in \mathrm{Lin}((B \setminus \{c\}) \cup \{a\})$. To pokazuje, że $B \subseteq \mathrm{Lin}((B \setminus \{c\}) \cup \{a\})$, więc $\mathrm{Lin}((B \setminus \{c\}) \cup \{a\}) = V$.

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. □(claim)

Weźmy teraz dowolny liniowo niezależny układ $a_1, \ldots, a_n$. Skonstruujemy rekurencyjnie ciąg $c_1, \ldots, c_n$ elementów B, takich że $(B \setminus \{c_1, \ldots, c_k\}) \cup \{a_1, \ldots, a_k\}$ 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 $a_1, \ldots, a_n$ jest lnz), pewien α_c jest niezerowy, możemy więc wziać $c_1 = c$.

- 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 $c_2 = \alpha_{a_1} a_1 + \sum_{b \in B_1 \setminus \{c_1\}} \alpha_b b$. 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 $c_1, \ldots, c_k$. Oznaczmy $B'_k = B \setminus \{c_1, \ldots, c_k\}$, $B_k = B'_k \cup \{a_1, \ldots, a_k\}$. 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 $a_1, \ldots, a_{k+1}$ jest liniowo niezależny, więc dla pewnego $c_{k+1} \in B'_k$ mamy $\alpha_{c_{k+1}} \neq 0$. Z Claimu wynika, że to c_{k+1} działa.

To daje nam $c_1, \ldots, c_n$ 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 $\{c_1, \ldots, c_k\}$.

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

Uwaga 2.17. Lemat Steinitza jest prawdziwy również dla nieskończonych układów liniowo niezależnych (przy założeniu pewnika wyboru), co można udowodnić zastępując rekursję na końcu dowodu rekursją pozaskończoną[4].

2.4 Podstawowe własności wymiaru

Stwierdzenie 2.18. *(a) jeżeli $W \leqslant V$, to $\dim W \leqslant \dim V$,*

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

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

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 $c_1, \ldots, c_n \in C$, zbiór $C \setminus \{c_1, \ldots, c_n\} \cup B$ jest bazą V. Ale C ma dokładnie n elementów (bo $\dim V = n$), więc $C \setminus \{c_1, \ldots, c_n\} \cup B = B$, czyli $\operatorname{Lin} B = \operatorname{Lin} C = V$.

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

Definicja 2.19. Niech V, W będą przestrzeniami liniowymi. Funkcję $F : V \to W$ 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.

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

Stwierdzenie 2.20. *Relacja $\cong$ jest relacją równoważności.*

Dowód. Ćwiczenie. □

[4]Patrz np. wykład z wprowadzenia do teorii zbiorów

Twierdzenie 2.21. *Załóżmy że V, W to skończenie wymiarowe przestrzenie liniowe nad K, a n to liczba naturalna.*

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

Dowód. Załóżmy że $\dim V = n$. Niech $B = \{v_1, \ldots, v_n\}$ będzie bazą V. Chcemy zdefiniować $F\colon V \to K^n$ 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 $v_1, \ldots, v_n$ 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 $F(v) = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{pmatrix}^\top = F(w)$, to znaczy że $v = \sum_{k=1}^n \alpha_k v_k = w$.
- F jest „na”: dla dowolnego $y \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{pmatrix}^\top \in K^n$ mamy $y = F(v)$, gdzie $v = \sum_{k=1}^n \alpha_k v_k$.

To pokazuje pierwszą część.

Druga część: 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 $\Leftarrow$.

Załóżmy że $F\colon V \to W$ jest izomorfizmem i ustalmy bazę $v_1, \ldots, v_n$ przestrzeni V. Wtedy $F(v_1), \ldots, F(v_n)$ jest bazą W (ćwiczenie), co kończy dowód. $\square$

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

Przykład 2.22. 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 $V \to \mathbf{R}^3$, np. dla $2x^3 + 3x^2 = 2(x+1)^3 - 3(x+1)^2 + 1 \cdot 1$ mamy $[2x^3 + 3x^2]_B = \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix}$.

Uwaga 2.23.

- Pierwsza część Twierdzenia 2.21 literalnie nie jest prawdziwa dla przestrzeni nieskończenie wymiarowych — trzeba wpisać coś innego zamiast $K^{\dim V}$ — ta przestrzeń na ogół ma wymiar ostro większy niż $\dim V$.
- Zamiast pełnego produktu należy wziąć podzbiór złożony z elementów, które mają prawie wszystkie współrzędne zerowe.
- Po tej modyfikacji pierwsza część jest prawdziwa w dowolnym wymiarze. Druga część jest prawdziwa w dowolnym wymiarze (bez modyfikacji).
- Obydwie części w nieskończenie wymiarowym przypadku mają w zasadzie ten sam dowód, co powyżej.

Rozdział 3

Przekształcenia liniowe, ich jądra i obrazy

3.1 Przekształcenia liniowe

Definicja 3.1. Jeżeli V, W są przestrzeniami liniowymi nad ciałem K, to *przekształceniem liniowym* (lub *funkcją liniową*, nie mylić z funkcjami postaci $f(x) = ax + b$) z V w W nazywamy funkcję $F : V \to W$, która jest jednocześnie:

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

Ćwiczenie 3.2. Funkcja $F : V \to W$ jest liniowa wtedy i tylko wtedy gdy spełnia $F(\alpha v_1 + v_2) = \alpha F(v_1) + F(v_2)$.

Ćwiczenie 3.3. 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 3.4. 1. funkcje liniowe $F : \mathbf{R}^n \to \mathbf{R}^k$ to dokładnie funkcje postaci F_A, zadane wzorem:

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

gdzie

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

2. Analogicznie wyglądają funkcje liniowe $K^n \to K^k$ dla dowolnego ciała K.

3. $F\colon \mathbf{R}[x] \to \mathbf{R}[x]$, $F(P) = P'$, a także takie same funkcje $F\colon K[x] \to K[x]$ zadane analogicznym wzorem.

4. $F\colon \mathbf{R}_3[x] \to \mathbf{R}_3[x]$, $F(P) = 3P - X^2 \cdot P'$.

5. $F\colon \mathbf{R}_3[x] \to \mathbf{R}^2$, $F(P) = \begin{pmatrix} P(1) \\ P'(0) + \pi P(e) \end{pmatrix}$.

6. $F\colon C(\mathbf{R}) \to \mathbf{R}$, $F(f) = \int_0^1 f(t)\,\mathrm{d}t$.

7. $F\colon \mathbf{R}^{\mathbf{N}} \to \mathbf{R}^{\mathbf{N}}$, $F(a_0, a_1 \ldots,) = (a_1, a_2, \ldots)$ (lewy szift),

8. $F\colon \mathbf{R}^{\mathbf{N}} \to \mathbf{R}^{\mathbf{N}}$, $F(a_0, a_1 \ldots,) = (0, a_0, a_1, \ldots)$ (prawy szift),

9. $F\colon C([0, 2\pi]) \to \mathbf{C}^{\mathbf{Z}}$, $F(f)(k) = \int_0^{2\pi} f(x)e^{-ikx}\,\mathrm{d}x$.

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

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

   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

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

więc $F(\alpha f) = \alpha F(f)$.

Uwaga 3.5. Jeżeli V, W są dowolnymi przestrzeniami liniowymi, to zbiór $\mathrm{Hom}(V,W)$ wszystkich przekształceń liniowych $V \to W$ jest przestrzenią liniową. Jeżeli V, W są skończenie wymiarowe, to wymiar $\mathrm{Hom}(V,W)$ to iloczyn wymiarów V i W[1].

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

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

Definicja 3.6. Ustalmy liniowe $F\colon V \to W$.

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

(Rysunek.)

Fakt 3.7. *Jeżeli $F\colon V \to W$ jest przekształceniem liniowym, to:*

a) $\ker F \leqslant V$*,*

b) $\operatorname{im} F \leqslant W$.

Dowód.

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

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

Z liniowości jeżeli $F(v_1) = F(v_2) = 0$, 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 $F(v) = 0$, 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 ba 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. □

Uwaga 3.8. W każdej przestrzeni liniowej (i w każdym ciele) mamy:

$$v_1 - v_2 = 0 \iff v_1 = v_2$$

(to było na pierwszej liście zadań).

Z tego wynika „prawo skreśleń”:

$$v_1 + v_2 = v_1' + v_2 \iff v_1 = v_1',$$

Fakt 3.9. *Załóżmy że $F : V \to W$ jest liniowe. Wtedy:*

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

Dowód. Pierwsza część jest oczywista.

Załóżmy 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 $F(v_1) \neq \vec{0} + F(v_2) = F(v_2)$. □

Przykłady 3.10. a) Niech $A = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}$, $F_1 = F_A$.

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

$\ker F_A = \{\vec{0}\}$.

b) $F_2\colon \mathbf{R}_3[x] \to \mathbf{R}^2$, $F(P) = \begin{pmatrix}P'(2)\\P(-1)\end{pmatrix}$

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

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

c) $F_3\colon \mathbf{R}_{100}[x] \to \mathbf{R}$, $F_3(P) = \int_{-1}^{1} P(x)\,\mathrm{d}x$.

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

$\ker F_3 = \{\, P \in \mathbf{R}_{100}[x] \mid \int_{-1}^{1} P(x)\,\mathrm{d}x = 0 \,\}$.

d) $F_4\colon \mathbf{R}^{\mathbf{N}} \to \mathbf{R}^{\mathbf{N}}$, $F_4(a_0, \ldots, a_n, \ldots) = (a_1, \ldots, a_n, \ldots,)$.

$\operatorname{im} F_4 = \mathbf{R}^{\mathbf{N}}$, $\ker F_4 = \{\, (a, 0, 0, 0, \ldots) \mid a \in \mathbf{R} \,\}$

e) $F_5\colon \mathbf{R}^{\mathbf{N}} \to \mathbf{R}^{\mathbf{N}}$, $F_5(a_0, \ldots, a_n, \ldots) = (0, a_0, a_1, \ldots, a_n, \ldots,)$.

$\ker F_5 = \{\vec{0}\}$, $\operatorname{im} F_5 = \{\, (a_n)_{n\in\mathbf{N}} \in \mathbf{R}^{\mathbf{N}} \mid a_0 = 0 \,\}$

Stwierdzenie 3.11. *Załóżmy że V jest dowolną przestrzenią liniową, a $A, B \subseteq V$ są rozłączne i $A \cup B$ jest liniowo niezależny. Wtedy* $\operatorname{Lin}(A) \cap \operatorname{Lin}(B) = \{\vec{0}\}$

Dowód. Ćwiczenie. □

Definicja 3.12. *Rząd* przekształcenia liniowego F to $\operatorname{rk} F := \dim \operatorname{im} F$.

Twierdzenie 3.13 (twierdzenie o rzędzie)**.** *Jeżeli $F\colon V \to W$ jest liniowe, to*

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

Wniosek 3.14 (twierdzenie o indeksie)**.** *Jeżeli* $\dim V < \infty$*, to odejmując stronami dostajemy stąd:*

$$\dim \ker F = \dim V - \dim \operatorname{im} F$$
$$\dim \operatorname{im} F = \dim V - \dim \ker F$$

Przykład 3.15. Chcemy wyznaczyć wymiar $V = \{ P \in \mathbf{R}_{50}[x] \mid \int_{-1}^{1} e^{-x^2} P(x)\,\mathrm{d}x = 0 \}$

Weźmy $G\colon \mathbf{R}_{50}[x] \to \mathbf{R}$, zadane wzorem $G(P) = \int_{-1}^{1} e^{-x^2} P(x)\,\mathrm{d}x$. Wtedy $\ker G = V$.

$G(1) = \int_{-1}^{1} e^{-x^2}\,\mathrm{d}x > 0$, więc $\mathrm{im}(G) = \mathbf{R}$, zatem $\dim V = \dim\ker G = \dim R_{50}[x] - \dim\mathbf{R} = 51 - 1 = 50$.

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

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

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

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 $F(v) = \sum_{b\in B} \beta_b F(b) \in \mathrm{Lin}\, F[B]$.

Załóżmy następnie, że $\sum_{b\in B} \beta_b F(b) = \vec{0}$. Ale $\sum_{b\in B} \beta_b F(b) = F(\sum_{b\in B} \beta_b b)$, więc wtedy $v = \sum_{b\in B} \alpha_b b \in \ker F$. Ponieważ $v \in \mathrm{Lin}\, B$ i $v \in \ker F = \mathrm{Lin} A$, dostajemy stąd $v = \vec{0}$, czyli $\sum_{b\in B} \alpha_b b = \vec{0}$. Z liniowej niezależności B wnioskujemy, że wszystkie $\alpha_b = 0$, więc $F[B]$ faktycznie jest liniowo niezależny (i F jest 1-1 na B). $\square$

Wniosek 3.16. *Załóżmy że $F\colon V \to W$ jest przekształceniem liniowym i $\dim V = \dim W < \infty$. Wtedy następujące warunki są równoważne:*

- $\ker F = \{\vec{0}\}$,
- *F jest 1-1,*
- *F jest „na”,*
- *F jest izomorfizmem.*

Dowód. Równoważność pierwszych dwóch warunków już była.

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

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

Stąd F jest „na" wtedy i tylko wtedy gdy jest 1-1, więc wówczas jest też izomorfizmem. □

Wniosek 3.17. *Jeżeli V jest skończenie wymiarowa i $F\colon V \to V$ jest przekształceniem liniowym (takie przekształcenie nazywamy* endomorfizmem*), to następujące warunki są równoważne:*

- *F jest różnowartościowa,*
- *F jest „na",*
- *F jest izomorfizmem (endomorfizm który jest izomorfizmem nazywamy* automorfizmem*).*

Uwaga 3.18. Założenie skończonego wymiaru jest istotne: lewy i prawy szift na $\mathbf{R}^{\mathbf{N}}$ 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 3.19. *Jeżeli mamy „krótki ciąg dokładny (przestrzeni liniowych)"*

$$V_1 \to V_2 \to V_3,$$

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

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

Z twierdzenia o rzędzie dla F_2 dostajemy więc

$$\dim V_2 = \dim \ker F_2 + \dim \operatorname{im} F_2 = \dim V_1 + \dim V_3.$$ □

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 $V \oplus W$.

Przykłady 4.2.

1. $K \times K = K^2$,
2. $K \times (K \times K) = K \times K^2 = K^3$ (teoriomnogościowo są to trochę inne obiekty, ale w algebrze liniowej je utożsamiamy),
3. jeżeli $F \subseteq V \times W$ jest funkcją $V \to W$, to F jest liniowe wtedy i tylko wtedy gdy $F \leqslant V \times W$ (ćwiczenie).

Uwaga 4.3. Możemy rozważać też nieskończone produkty i sumy proste przestrzeni liniowej. Wtedy te dwa pojęcia się rozjeżdżają (o tym więcej może być na konwersatorium).

Fakt 4.4. $\dim(V \oplus W) = \dim V + \dim W$

Dowód. Mamy krótki ciąg dokładny:

$$V \to V \oplus W \to W,$$

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

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

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

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

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

Fakt 4.5 (/definicja)**.** *Jeżeli mamy liniowe $F_1 : V_1 \to W_1$, $F_2 : V_2 \to W_2$, to nam daje liniowe $F_1 \oplus F_2 : V_1 \oplus V_2 \to W_1 \oplus W_2$, zadane wzorem $(F_1 \oplus F_2)(v_1, v_2) = (F_1(v_1), F_2(v_2))$.*

Dowód. Ćwiczenie. □

Fakt 4.6. *Dla dowolnych $F_1 : V_1 \to W_1$, $F_2 : V_2 \to W_2$, $G_1 : W_1 \to U_1$, $G_2 : W_2 \to U_2$ (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 $(v_1, v_2) \in V_1 \oplus V_2$). □

Ćwiczenie 4.7. 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 4.8 (/definicja)**.** Jeżeli $W, U \leqslant V$ 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 $V = W \oplus U$).

4.2 Przestrzeń dualna

Definicja 4.9. Jeżeli V jest przestrzenią liniową nad K, to definiujemy *przestrzeń dualną do V* jako

$$V^* = V' = \{ f : V \to K \mid f \text{ jest liniowe} \}.$$

Elementy V^* nazywamy *funkcjonałami* na V.

Lemat 4.10. *V^* jest przestrzenią liniową, podprzestrzenią K^V (przestrzeni wszystkich funkcji $V \to K$).*

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 jednorodne: 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 4.11. *Jeżeli* $\dim V < \infty$*, to* $\dim V = \dim V^*$. [1]

Dowód. Niech $b_1, \ldots, b_n$ będzie bazą V. Dla $k = 1, \ldots, n$ 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 $b_1^*, \ldots, b_n^*$ nazywamy *bazą dualną* do $b_1, \ldots, b_n$.

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:

$$\sum_k \alpha_k b_k^*(b_j) = \alpha_j. \tag{*}$$

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

[1] Przy 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 $v^* = \sum_k \alpha_k b_k^*$ (kiedy $\dim V$ jest nieskończony, prawa strona zwykle nie ma sensu). Istotnie, dla dowolnego $v = \sum_j \beta_j b_j$ mamy

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

Uwaga 4.12. Jeżeli $\dim V > 1$ i $v \in V$, to nie istnieje „funkcjonał dualny” do v.

Dokładniej, jeżeli $B = b_1, \ldots, b_n$ i $C = c_1, c_2 = b_2, \ldots, c_n = b_n$ są bazami (różniącymi się tylko pierwszym elementem), to *nie* zachodzi $b_k^* = c_k^*$ (dla żadnego k, w tym $k > 1$).

Uwaga 4.13. Poza algebrą liniową V^* definiuje się czasami inaczej, na przykład jako przestrzeń liniową złożoną tylko z tych funkcjonałów, które są ciągłe (względem jakichś topologii na V i K). Wtedy V^* może mieć nawet mniejszy wymiar niż V, a nawet są nieskończenie wymiarowe V takie że $\dim V^* = 0$. (O tym możemy porozmawiać trochę na konwersatorium.)

Fakt 4.14 (/definicja). *Jeżeli $F: V \to W$ jest liniowe, to odwzorowanie $F^*: W^* \to V^*$ zadane wzorem $F^*(f)(v) = f(F(v))$ jest liniowe. Nazywamy je* przekształceniem dualnym do F

Dowód. Ćwiczenie. □

Ćwiczenie 4.15. Jeżeli $F_1: V_1 \to V_2$ i $F_2: V_2 \to V_3$ są liniowe (w skrócie: $V_1 \xrightarrow{F_1} V_2 \xrightarrow{F_2} V_3$), to $(F_2 \circ F_1)^* = F_1^* \circ F_2^*$.

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

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

Ćwiczenie 4.16. Jeżeli $V \cong W$, to $V^* \cong W^*$.

4.3 Przestrzeń ilorazowa

Definicja 4.17. Jeżeli $W \leqslant V$ są przestrzeniami liniowymi i $v \in V$, *warstwą* v względem W nazywamy zbiór:

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

Fakt 4.18. $v_1 + W = v_2 + W \iff v_1 - v_2 \in W$

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. □

Wniosek 4.19. *Zbiór warstw W w V to zbiór ilorazowy $V/\sim$, gdzie*

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

Definicja 4.20. Jeżeli $W \leqslant V$ są przestrzeniami liniowymi, to zbiór warstw $V/\sim$ oznaczamy V/W i nazywamy *przestrzenią ilorazową* (V nad W lub V przez W).

Zadajemy strukturę przestrzeni liniowej:

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

Fakt 4.21. *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:

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

Definicja 4.22. Odwzorowanie liniowe nazywamy *epimorfizmem* jeżeli jest „na”, a *monomorfizmem* jeżeli jest 1-1.

(W szczególności izomorfizm = epimorifzm i monomorfizm.)

Stwierdzenie 4.23. $\dim V = \dim V/W + \dim W$*, czyli* $\dim V/W = \dim V - \dim W$*, o ile* $\dim V < \infty$.

Dowód. Zdefiniujmy $F(v) = v+W$. Cel: F jest liniowe, $\ker V = W$, $\operatorname{im} F = V/W$.

- 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.

Stwierdzenie wynika z twierdzenia o rzędzie: $\dim V = \dim \operatorname{im} F + \dim \ker F = \dim V/W + \dim W$. □

Stwierdzenie 4.24 (Twierdzenie o izomorfizmie)**.** *Jeżeli $F\colon V \to W$ jest odwzorowaniem liniowym, to* $\operatorname{im} F \cong V / \ker F$.

$$\begin{tikzcd} V \arrow[r, "F"] \arrow[d] & \operatorname{im} F \\ V / \ker F \arrow[ur, "\bar{F}"] & \end{tikzcd}$$

Dowód. Zadajmy odwzorowanie $\bar{F}\colon V / \ker F \to \operatorname{im} F$ 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 $w = \bar{F}(v + \ker F)$. □

4.4 Przestrzeń bidualna

Stwierdzenie 4.25. *Dla dowolnej przestrzeni liniowej V mamy odwzorowanie liniowe $\Phi\colon V \to V^{**}$ (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 jednorodne.

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.

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

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

Stwierdzenie 4.26. *Jeżeli* $\dim V < \infty$, *to Φ zadaje izomorfizm $V \cong V^{**}$.*[2]

[2] Przy założeniu pewnika wyboru Φ zawsze jest różnowartościowe, ale w przypadku nieskończenie wymiarowym wymiar V^{**} jest dużo większy niż wymiar V.

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

Uwaga 4.27.

- Nawet jeżeli $\dim V$ jest nieskończony, to pierwsza część dowodu powyżej działa i pokazuje, że Φ jest monomorfizmem. (Na izomorfizm nie ma szans, bo zwykle $\dim V < \dim V^* < \dim V^{**}$.)

- 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 naturalny: dla dowolnego $F : V \to W$ mamy przemienny diagram:

$$\begin{tikzcd} V \arrow[r, "F"] \arrow[d, "\Phi_V"] & W \arrow[d, "\Phi_W"] \\ V^{**} \arrow[r, "F^{**}"] & W^{**}, \end{tikzcd}$$

  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 więcej na konwersatorium).

Rozdział 5

Macierze i układy równań

Macierze

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

Układy równań

Rozważamy układy równań

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

Definicja 5.2.
- a_{ij} — współczynniki,
- $y_1, \ldots, y_m$ — dane,
- $x_1, \ldots, x_n$ — niewiadome/szukane.

Ten sam układ równań inaczej (w postaci wektorowej):

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

$$x_1A_1 + x_2A_2 + \ldots + x_nA_n = Y \tag{2}$$

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

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

w skrócie

$$AX = Y \tag{3}$$

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,
- $Y \in \mathrm{Lin}(A_1, \ldots, A_n)$,
- $\mathrm{Lin}(A_1, \ldots, A_n) = \mathrm{Lin}(A_1, \ldots, A_n, Y)$,
- $\dim \mathrm{Lin}(A_1, \ldots, A_n) = \dim \mathrm{Lin}(A_1, \ldots, A_n, Y)$.

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

Z uwagi powyżej łatwo wynika następujące twierdzenie.

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

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.

Fakt 5.9. *Operacje kolumnowe K1-3 nie zmieniają obrazu* $F_A \colon K^n \to K^m$.

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. □

Fakt 5.10. *Operacje wierszowe W1-W3 nie zmieniają jądra* $F_A \colon K^n \to K^m$.

Dowód. W1: Niech $A \overset{W_1}{\rightsquigarrow} A'$. 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 $F_{A'}$).

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

Dla W2 i W3 rozumujemy analogicznie (raczej łatwiej). □

Wniosek 5.11. *Operacje wierszowe i kolumnowe nie zmieniają rzędu macierzy.*

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

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

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

$$\mathrm{rk} A = n - \dim \ker F_A = n - \dim \ker F_{A'} = \mathrm{rk} A'.$$ □

Przykład 5.12.

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

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

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

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

Czyli rząd = 2

Definicja 5.13. $\mathrm{rk}_W A$ = maksymalna liczba lnz wierszy $A = \mathrm{rk} A^\top$

Fakt 5.14. $\mathrm{rk}_W A$ *nie zmienia się przy W1-3 i K1-3.*

Lemat 5.15. *Każdą macierz można ciągiem operacji K1-3 i W1-3 sprowadzić do macierzy postaci*

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

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 $j = 2, 3, \ldots$) „czyścimy” pierwszą kolumnę.

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

7. To daje nam macierz następującej postaci

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

Powtarzamy procedurę dla macierzy B. □

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

Twierdzenie 5.16 (Eliminacja Gaussa)**.** *Każdą macierz można operacjami W1, W3 sprowadzić do macierzy w* postaci schodkowej*, to znaczy macierzy postaci:*

$$\begin{pmatrix} 0\cdots\cdots 0 & [\neq 0] & ? & & & & & \\ 0\cdots\cdots\cdots\cdots 0 & \cdots\cdots\cdots 0 & [\neq 0] & ? & & \text{\Large ?} & & \\ 0\cdots\cdots\cdots\cdots 0 & \cdots\cdots\cdots 0 & & 0\cdots\cdots\cdots 0 & [\neq 0] & ? & & \\ & & \text{\Large 0} & & & \ddots & & \end{pmatrix}$$

Wyrazy „$\neq 0$" (pierwsze niezerowe w wierszach) nazywamy wyrazami wiodącym.

Dowód. Przemy na wschód (tam musi być cywilizacja!).

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, czyścimy 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 5.17. *Korzystając z operacji W1, W2, W3 można sprowadzić każdą macierz do* zredukowanej postaci schodkowej*, tzn. postaci schodkowej z wyrazami wiodącymi 1.*

Dowód. Oczywiste — 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:

$$\begin{cases} x_{n_1} + b_{1(n_1+1)}x_{n_1+1} + \ldots = y_1 \\ x_{n_2} + b_{2(n_2+1)}x_{n_2+1} + \ldots = y_2 \\ \vdots \\ x_{n_k} + b_{m(n_m+1)}x_{n_k+1} + \ldots = y_k \\ 0 = y_{k+1} \\ \vdots \\ 0 = y_m \end{cases}, \tag{4}$$

gdzie $n_1 < n_2 < \ldots < n_k$

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 $y_{k+1} = y_{k+2} = \ldots = y_m = 0$.

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

$$\begin{cases} x_{n_1} = y_1 - b_{1(n_1+1)}x_{n_1+1} - \ldots \\ x_{n_2} = y_2 - b_{2(n_2+1)}x_{n_2+1} - \ldots \\ \vdots \\ x_{n_k} = y_k - b_{m(n_m+1)}x_{n_k+1} - \ldots \end{cases},$$

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 $x_{n_1}, \ldots, x_{n_k}$.

$$\begin{cases} x_{n_1} = y_1' + c_{1(n_1+1)}x_{n_1+1} + \ldots \\ x_{n_2} = y_2' + c_{2(n_2+1)}x_{n_2+1} + \ldots \\ \vdots \\ x_{n_k} = y_k' + c_{m(n_m+1)}x_{n_k+1} + \ldots \end{cases},$$

Przykład 5.18. Pewien układ równań zmiennych $x_1, x_2, \ldots, x_9$ sprowadziliśmy do zredukowanej postaci schodkowej, uzyskując układ:

$$\begin{cases} x_4 + 7x_5 - 3x_6 = 1 \\ x_6 + 2x_7 + 3x_8 = 2 \\ x_7 + 3x_8 + 4x_9 = 3 \\ \hfill 0 = 0 \\ \hfill \vdots \\ \hfill 0 = 0 \end{cases}, \tag{$*$}$$

Przenosząc zmienne związane na jedną stronę:

$$\begin{cases} x_4 = 1 - 7x_5 + 3x_6 \\ x_6 = 2 - 2x_7 - 3x_8 \\ x_7 = 3 - 3x_8 - 4x_9 \end{cases},$$

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

$$\begin{cases} x_7 = 3 - 3x_8 - 4x_9 \\ x_6 = 2 - 2(3 - 3x_8 - 4x_9) - 3x_8 \\ \quad = 2 - 6 + 6x_8 + 8x_9 - 3x_8 \\ \quad = -4 + 3x_8 + 8x_9 \\ x_4 = 1 - 7x_5 + 3(-4 + 3x_8 + 8x_9) \\ \quad = 1 - 7x_5 - 12 + 9x_8 + 24x_9 \\ \quad = -11 - 7x_5 + 9x_8 + 24x_9 \end{cases},$$

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:

$$X = x_1 \begin{pmatrix} 1\\0\\0\\0\\0\\0\\0\\0\\0 \end{pmatrix} + x_2 \begin{pmatrix} 0\\1\\0\\0\\0\\0\\0\\0\\0 \end{pmatrix} + x_3 \begin{pmatrix} 0\\0\\1\\0\\0\\0\\0\\0\\0 \end{pmatrix} + x_4 \begin{pmatrix} 0\\0\\0\\1\\0\\0\\0\\0\\0 \end{pmatrix} + x_5 \begin{pmatrix} 0\\0\\0\\0\\1\\0\\0\\0\\0 \end{pmatrix} + x_6 \begin{pmatrix} 0\\0\\0\\0\\0\\1\\0\\0\\0 \end{pmatrix} + x_7 \begin{pmatrix} 0\\0\\0\\0\\0\\0\\1\\0\\0 \end{pmatrix} + x_8 \begin{pmatrix} 0\\0\\0\\0\\0\\0\\0\\1\\0 \end{pmatrix} + x_9 \begin{pmatrix} 0\\0\\0\\0\\0\\0\\0\\0\\1 \end{pmatrix} =$$

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

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) stowarzyszonego z układem (∗), a wektor $X_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -11 \\ 0 \\ -4 \\ 3 \\ 0 \\ 0 \end{pmatrix}$ jest szczególnym rozwiązaniem (∗). 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 5.19. 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:

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

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

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

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:

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

Przykład 5.20. Mając dane liniowo niezależne wektory $X_1, X_2, \ldots, X_k \in K^n$, jak znaleźć układ równań, którego to jest fundamentalny układ rozwiązań?

Szukamy macierzy A takiej że $AX_1, AX_2, \ldots, AX_k = 0$. Jeżeli wiersze A to $A_1, A_2, \ldots, A_m$, to znaczy że $A_iX_j = 0$ dla $i = 1,\ldots,m$ i $j = 1,\ldots,k$. Transponując, szukamy rozwiązań jednorodnego układu równań

$$BY = 0,$$

gdzie $B = (X_1X_2X_3\ldots X_k)^\top$. Rząd B to k (bo $X_1,\ldots,X_k$ są liniowo niezależne). Macierz B ma szerokość n (więć 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 $Y_1,\ldots,Y_{n-k}$. Macierz $A = (Y_1\ldots Y_{n-k})^\top$ 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 $X_1,\ldots,X_k$ jest bazą $\ker F_A$, czyli fundamentalnym układem rozwiązań $AX = 0$.

(Konkretny przykład: zadanie z listy, rozwiązanie na kanale ćwiczeń.)

Uwaga 5.21. Metoda eliminacji Gaussa daje:

- paramateryzację zbioru rozwiązań postaci

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

  z minimialną liczbą parametrów,

- opis obrazu F_A przy użyciu minimalnej możliwej liczby równań liniowych

Dowód. Weźmy układ $AX = Y$, gdzie A jest wysokości m, szerokości n. Eliminacja Gaussa $(A|Y)$ prowadzi do $(S|Y)$, gdzie S jest macierzą schodkową o s schodkach. Wtedy $\operatorname{rk} A = \operatorname{rk} S = s$.

- Jest $n—s$ zmiennych wolnych, czyli $l = n—s$ parametrów w rozwiązaniu ogólnym. Ale $\dim\ker F_A = \dim n - \dim\operatorname{im} F_A = n - s = l$.

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

  $$F_B \colon K^m \to K^k \quad k = ?$$

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

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

Ale co gdy mamy dowolne (skończenie wymiarowe) V i W oraz odwzorowanie liniowe $F \colon V \to W$?

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

Przykład 5.22. $V = \{P \in \mathbf{R}_2[x] \mid P'(-1) = 0\}$, $W = \mathbf{R}_1[x]$, $F(P) = P' + P(1)x$.

Czym jest $F^{-1}\{x+1\}$?

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

$$\begin{tikzcd}
V \arrow[r, "F"] \arrow[d, "a+b(x^2+2x)\mapsto(a,b)"'] & W \arrow[d, "a+bx\mapsto(a,b)"] \\
\mathbf{R}^2 \arrow[r, "\begin{pmatrix} ? & ? \\ ? & ? \end{pmatrix}"'] & \mathbf{R}^2
\end{tikzcd}$$

Szukamy macierzy pasującej do diagramu.

$$F(a+b(x^2+2x)) = F(bx^2+2bx+a) = 2bx+2b+(b+2b+a)x = (a+5b)x+2b,$$

czyli mamy

$$\begin{tikzcd} V \arrow[r, "F"] \arrow[d, "a+b(x^2+2x)\mapsto(a,b)"'] & W \arrow[d, "a+bx\mapsto(a,b)"] \\ \mathbf{R}^2 \arrow[r, "\begin{pmatrix} 0 & 2 \\ 1 & 5 \end{pmatrix}"'] & \mathbf{R}^2 \end{tikzcd}$$

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

$$\begin{pmatrix} 0 & 2 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

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

$$\begin{tikzcd} v \arrow[r, equal] \arrow[d, mapsto] & \sum_i \beta_i b_i \\ {[v]_B} \arrow[r, equal] & \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \arrow[u, mapsto] \end{tikzcd} \qquad \begin{tikzcd} V \arrow[r, "F"] \arrow[d, leftrightarrow] & W \arrow[d, leftrightarrow] \\ K^n \arrow[r, dotted, "?"] & K^m \end{tikzcd} \qquad \begin{tikzcd} w \arrow[r, equal] \arrow[d, mapsto] & \sum_j \gamma_j c_j \\ {[w]_C} \arrow[r, equal] & \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix} \arrow[u, mapsto] \end{tikzcd}$$

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

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

(Na końcu zmieniona kolejność sumowania!)

Mamy więc

$$\begin{array}{ccccc} v & \longmapsto & F(v) & & \\ \updownarrow & & \updownarrow & & \\ \begin{pmatrix}\beta_1\\ \vdots \\ \beta_n\end{pmatrix} & \longmapsto & \begin{pmatrix}\sum_i \beta_i a_{1i}\\ \vdots \\ \sum_i \beta_i a_{mi}\end{pmatrix} & = & A\begin{pmatrix}\beta_1\\ \vdots \\ \beta_n\end{pmatrix}, \end{array}$$

gdzie $A = (a_{ji})$.

Definicja 5.23. 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)

$$\begin{tikzcd} V \arrow[r, "F"] \arrow[d, "{[\cdot]_B}"] & W \arrow[d, "{[\cdot]_C}"] \\ K^n \arrow[r, "m_C^B(F)"] & K^m, \end{tikzcd}$$

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

Uwaga 5.24. Jeżeli B, C są bazami V i W i A jest macierzą odpowiedniego rozmiaru, to istnieje (jedyne) takie $F: V \to W$, że $A = m_C^B(F)$.

Przykład 5.25. Jeszcze raz ten sam przykład.

$m_C^B(F) = \begin{pmatrix}0 & 2\\ 1 & 5\end{pmatrix}$. $F(b_1) = F(1) = 0 + 1\cdot x = x = 0\cdot 0 + 1\cdot x$, $F(b_2) = F(x^2+2x) = 2x + 2 + 3\cdot x = 5x + 2 = 2\cdot 1 + 5\cdot x$.

A co, gdy chcemy bazę $B' = (x^2 + 2x + 3, 1)$ (czyli chcemy policzyć $m_C^{B'}(F)$)?

Twierdzenie 5.26. *Załóżmy że mamy*

$$\begin{array}{ccccccc} Z & \xleftarrow[G]{} & W & \xleftarrow[F]{} & V & & \\ & & & & & & \\ D & & C & & B & & \textit{(bazy)} \end{array}$$

Wedty $m^B_D(G \circ F) = m^C_D(G) \cdot m^B_C(F)$. *(Mnemotechnika: C się skraca!)*

Dowód. i-ta kolumna $m^B_D(G \circ F)$ to

$$[(G \circ F)(b_i)]_D = [G(F(b_i))]_D = m^C_D(G)[F(b_i)]_C,$$

a $[F(b_i)]_C$ to i-ta kolumna $m^B_C(F)$. Ale to daje tezę (narysuj schemat mnożenia $m^C_D(G)$ i $m^B_C(F)$ żeby się przekonać!). □

Przykład 5.27. Ciąg dalszy.

$$\begin{array}{ccccc} W & \xleftarrow{F} & V & \xleftarrow{\mathrm{id}_V} & V \\ \\ C & & B & & B' \end{array}$$

Znamy $m^B_C(F) = \begin{pmatrix} 0 & 2 \\ 1 & 5 \end{pmatrix}$. Czym jest $m^{B'}_B(\mathrm{id}_V)$? 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 $m^{B'}_B(\mathrm{id}_V) = \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix}$, a więc

$$m^{B'}_C(F) = m^B_C(F) \cdot m^{B'}_B(\mathrm{id}_V) = \begin{pmatrix} 0 & 2 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 8 & 1 \end{pmatrix}.$$

Definicja 5.28. Jeżeli rozważamy $F \colon V \to V$, to oznaczamy $m_B(F) := m^B_B(F)$.

Przykład 5.29. Mamy $F \colon V \to V$ i bazy B, C przestrzeni V. Jak policzyć $m_C(F)$, znając $m_B(F)$?

$$m_C(F) = m^C_C(F) = m^B_C(\mathrm{id}_V) \cdot m^B_B(F) \cdot m^C_B(\mathrm{id}_V) = m^B_C(\mathrm{id}_V) \cdot m_B(F) \cdot m^C_B(\mathrm{id}_V)$$

Uwaga 5.30. Jeżeli B, C są bazami V i $v \in V$, to $[v]_C = m^B_C(\mathrm{id}_V)[v]_B$. ($B$ się skraca!)

Definicja 5.31. Mówimy że $A \in \mathrm{M}_{m\times n}(K)$ jest *odwracalna* jeżeli istnieje $B \in \mathrm{M}_{n\times m}(K)$ (*macierz odwrotna do* A) taka że $AB = \mathrm{I}_m$ i $BA = \mathrm{I}_n$.

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

(Dowód: jeżeli B, B' są takie, to $B' = B'\mathrm{I}_m = B'AB = \mathrm{I}_n B = B$.)

Fakt 5.32. $m^C_B(\mathrm{id})^{-1} = m^B_C(\mathrm{id})$.

Dowód. $m^C_B(\mathrm{id})\cdot m^B_C(\mathrm{id}) = m^B_B(\mathrm{id}\circ\mathrm{id}) = m^B_B(\mathrm{id}) = \mathrm{I}$ i podobnie odwrotnie. □

Jak liczyć A^{-1}?

Pierwszy sposób: 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

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

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 rząd A jest mniejszy niż liczba jej wierszy (lub kolumn), więc F_A nie jest „na” lub nie jest 1-1, a A^{-1} nie istnieje.

Przykład 5.33.

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

  nie ma macierzy odwrotnej

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

Przykład 5.34. $E = (e_1, e_2)$, $B = \left(\begin{pmatrix}1\\3\end{pmatrix}, \begin{pmatrix}2\\4\end{pmatrix}\right)$ — bazy $\mathbf{R}^2$. $F\colon \mathbf{R}^2 \to \mathbf{R}^2$, $m_E(F) = \begin{pmatrix}1 & -1\\1 & 0\end{pmatrix}$.

Wtedy

$$m_B(F) = m^E_B(\mathrm{id})m_E(F)m^B_E(\mathrm{id}).$$

Oczywiście $m^B_E(\mathrm{id}) = \begin{pmatrix}1 & 2\\3 & 4\end{pmatrix}$ i $m^E_B(\mathrm{id}) = m^B_E(\mathrm{id})^{-1} = \begin{pmatrix}-2 & 1\\ \frac{3}{2} & -\frac{1}{2}\end{pmatrix}$, czyli

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

Definicja 5.35. $F: V \to W$ nazywamy *odwracalnym* jeżeli istnieje takie liniowe $G: W \to V$, że $F \circ G = \mathrm{id}_W$ i $G \circ F = \mathrm{id}_V$.

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

Uwaga 5.36. F jest odwracalne $\iff$ F izomorfizmem. (To wiemy z ćwiczeń.)

Wniosek 5.37. *Jeżeli $F: V \to W$ jest odwracalne, to* $\dim V = \dim W$.

Fakt 5.38. *Niech $F: V \to W$ — 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 $m_C^B(F)$ (bo $\mathrm{I} = m_B^B(\mathrm{id}_V) = m_B^B(G \circ F) = m_B^C(G) m_C^B(F)$ i podobnie w drugą stronę).

Jeżeli $m_C^B(F)$ jest odwracalna, to weźmy $G: W \to V$ takie że $m_B^C(G) = m_C^B(F)^{-1}$. Dalej rachunek jak wyżej (tylko w odwrotnej kolejności). □

Wniosek 5.39. *Jeżeli macierz A jest odwracalna, to jest kwadratowa.*

Fakt 5.40.

1. $\mathrm{rk}(AB) \leqslant \mathrm{rk} A$*,*
2. $\mathrm{rk}(AB) \leqslant \mathrm{rk} B$*,*
3. *jeżeli A jest odwracalna, to* $\mathrm{rk}(AB) = \mathrm{rk} B$*,*
4. *jeżeli B jest odwracalna, to* $\mathrm{rk}(AB) = \mathrm{rk} A$*.*

Dowód. Rząd = wymiar obrazu.

1. $\mathrm{im}(F_A \circ F_B) \subseteq \mathrm{im}\, F_A$,
2. $\mathrm{rk}(AB) = \mathrm{rk}(AB)^\top = \mathrm{rk}(B^\top A^\top) \leqslant \mathrm{rk}\, B_T = \mathrm{rk}\, B$
3. wiemy że $\mathrm{rk}(AB) \leqslant \mathrm{rk}\, B$.

   $\mathrm{rk}\, B = \mathrm{rk} A^{-1}AB = \mathrm{rk}(A^{-1}(AB)) \leqslant \mathrm{rk}(AB)$.
4. $\mathrm{rk} A = \mathrm{rk}((AB)B^{-1}) \leqslant \mathrm{rk}(AB)$ □

Rozdział 6

Wyznacznik

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

- n-liniowe,
- alternujące (tzn. zerujące się, jeżeli jakiś wektor się powtarza).

Uwaga 6.1. Jeżeli $\operatorname{char} K \neq 2$, tzn. $1+1 \neq 0$ w K, to n-liniowe d jest alternujące $\iff$ jest antysymetryczne (zmienia znak, jeżeli zamienimy miejscami dwa argumenty).

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

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

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

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

Ta ostatnia wartość jest zerowa z alternowania.

W drugą stronę, jeżeli char $K \neq 2$, to jeżeli $d(X_1, X_1, X_3, \ldots) = -d(X_1, X_1, X_3, \ldots)$, więc $2d(X_1, X_1, X_3, \ldots) = 0$ i możemy podzielić stronami przez 2.

Dla $K = \mathbf{F}_2 = \{0,1\}$ i $n = 2$ funkcja $f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1y_1 + x_2y_2$ jest symetryczna (a więc antysymetryczna), ale nie alternująca. □

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

$$\begin{aligned}
d(A_1, A_2, \ldots) &= d\left(\sum_{i_1=1}^{n} a_{i_1 1} e_{i_1}, \sum_{i_2=1}^{n} a_{i_2 2} e_{i_2}, \ldots, \sum_{i_n=1}^{n} a_{i_n n} e_{i_n}\right) \\
&= \sum_{i_1=1}^{n} a_{i_1 i} d\left(e_{i_1}, \sum_{i_2=1}^{n} a_{i_2 2} e_{i_2}, \ldots, \sum_{i_n=1}^{n} a_{i_n n} e_{i_n}\right) \\
&= \sum_{i_1, i_2, \ldots, i_n = 1}^{n} a_{i_1 1} a_{i_2 2} \cdots a_{i_n n} d(e_1, e_2, \ldots, e_n) \\
&= \sum_{i_1, i_2, \ldots, i_n \text{ parami różne}} a_{i_1 1} a_{i_2 2} \cdots a_{i_n n} d(e_1, e_2, \ldots, e_n) \\
&= \sum_{\sigma \text{ permutacja} \circlearrowleft \{1, \ldots, n\}} a_{\sigma(1)1} \cdots a_{\sigma(n)n} d(e_{\sigma(1)}, \ldots, e_{\sigma(n)})
\end{aligned}$$

Ciąg $i_1, i_2, \ldots, i_n$ odpowiada permutacji $\sigma(j) = i_j$.

Zbiór permutacji (bijekcji) $\{1, \ldots, n\}$ oznaczamy S_n. $|S_n| = n!$.

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

Przykład 6.2.

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

Notacja: permutacje $\sigma \in S_n$ (czyli bijekcje $\{1, 2, \ldots, n\}\circlearrowleft$) reprezentuje się w następujący sposób:

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

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.

Składanie permutacji zapisujemy pisząc funkcje obok siebie.

Przykład 6.3.

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

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 6.4. *Każdą permutację można przedstawić w każdej z postaci:*

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

Definicja 6.5. 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 rysunek.)

2. parę (i,j) (lub $\{i,j\}$) taką jak po ostatniej równości powyżej *inwersją* permutacji σ

3. znak permutacji σ to $\operatorname{sgn}(\sigma) = (-1)^{\text{liczba inwersji } \sigma}$.

Wniosek 6.6. *$d(e_{\sigma(1)},\dots,e_{\sigma(n)}) = \operatorname{sgn}(\sigma) d(e_1,\dots,e_n)$, więc*

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

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

Załóżmy 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 podstępujemy 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. □

Definicja 6.7. *Wyznacznik* wektorów $A_1,\dots,A_n \in K^n$ to

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

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

Wniosek 6.8. *Każda alternująca, n-liniowa funkcja $d\colon K^n \to K$ jest postaci* $c\cdot\det$.

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 6.9. *Ustalmy permutację $\sigma = (\sigma(1), \ldots, \sigma(n))$ i $j < k \leqslant n$. Niech $\bar{\sigma}$ będzie permutacją powstałą z σ przez zamianę $\sigma(j)$ i $\sigma(k)$, tzn.*

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

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

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

- jeżeli $i_1 = i, i_2 = k$, to

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

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

- jeżeli $i_1, i_2 \neq j, k$, to

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

  (liczba inwersji jest ta sama)

- jeżeli $i_1 < j$, to

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

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

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

- jeżeli $j < i < k$, to

$$\begin{aligned}(j,i) \text{ jest inwersją } \bar{\sigma} &\iff \underset{\sigma(k)}{\underset{\|}{\bar{\sigma}(j)}} > \underset{\sigma(i)}{\underset{\|}{\bar{\sigma}(i)}} \\ &\iff (i,k) \textit{ nie} \text{ jest inwersją } \sigma,\end{aligned}$$

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

  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. □

Uwaga 6.10. Ogólnie jeżeli σ, τ są permutacjami, to $\operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\tau)$.

Wniosek 6.11. 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:

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

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

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

Alternowanie: załóżmy że $n > 1$ (przypadek $n = 0, 1$ jest trywialny). Chcemy pokazać, że

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

Istotnie, jeżeli $B = \sum_j b_j e_j$ i $A_i = \sum_j a_{ji} e_j$, to dla każdej $\sigma \in S_n$ (tzn. permutacji $\{1, 2, \ldots, n\}$) mamy:

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

Stąd suma

$$\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(j-1)(j-1)} b_{\sigma(j)} \cdots a_{\sigma(k-1)(k-1)} b_{\sigma(k)} \cdots a_{\sigma(n)n}$$

jest zerowa: wyrazy odpowiadające parze σ, $\bar{\sigma}$ się skracają (uwaga: to faktycznie są pary, co wynika z tego że $\bar{\bar{\sigma}} = \sigma$, $\bar{\sigma} \neq \sigma$ i dla $\sigma_1 \neq \sigma_2$ mamy $\bar{\sigma}_1 \neq \bar{\sigma}_2$!) $\square$

Definicja 6.12. Jeżeli $A \in \mathrm{M}_{n\times n}(K)$ o kolumnach $A_1, \ldots, A_n$, to

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

Lemat 6.13. *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 $i < j \leqslant n$. Jeżeli (i, j) jest inwersją σ, to $\sigma(i) > \sigma(j)$ i $(\sigma(j), \sigma(i))$ jest inwersją σ^{-1}: istotnie, $\sigma(\sigma^{-1}(j)) = j > i = \sigma^{-1}(\sigma(i))$. Podobnie jeżeli i, j nie jest inwersją, to $\sigma(i) < \sigma(j)$ i $(\sigma(i), \sigma(j))$ nie jest inwersją σ^{-1}. □

Twierdzenie 6.14. *Niech $A \in \mathrm{M}_{n\times n}(K)$,*

1. $\det(A_1, A_2, \ldots, A_n) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$ *(σ na innym indeksie!)*

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

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

4. *dla każdego $j \in \{1, \ldots, n\}$:*

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

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

(To A_{ij}, lub macierz którego jest wyznacznikiem, nazywamy minorem *macierzy A.)*

4'. *dla każdego $i \in \{1, \ldots, n\}$:*

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

(rozwinięcie Laplace'a względem i-tego wiersza)

5. • $\det A$ *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 (górno- lub dolno-)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.

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

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

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

   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 $d(\mathrm{I}) = \det(A\mathrm{I}) = \det(A) = \det(A) \cdot 1 = \det(A) \cdot \det(\mathrm{I})$, 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 $d(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} A_{ij}$. Chcemy pokazać, że $d = \det$. Pokażemy najpierw, że d jest n-liniowe i alternujące. n-liniowość jest łatwa (A_{ij} zależy liniowo od kolumn różnych od j-tej, ale $a_{ij} A_{ij}$ zależy liniowo od wszystkich kolumn).

   Alternowanie: załóżmy 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 żądaną równość.

   Stąd $d = c \cdot \det$, a łatwo sprawdzić, że $d(\mathrm{I}) = 1 = 1 \cdot \det(\mathrm{I})$, 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 trójkątną; 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.

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 zerowej kolumnie. Biorąc rozwinięcie Laplace'a A' względem zerowej kolumny wnioskujemy że $\det A = \det A' = 0$.

   Jeżeli kolumny A są liniowo niezależne, to $\mathrm{rk} A = n$ i stosując operacje W1, W3 (nie zmieniające wyznacznika ani rzędu) możemy sprowadzić A do macierzy diagonalnej A'. Ponieważ $\mathrm{rk} A' = n$, wyrazy na jej przekątnej są niezerowe, więc z 6. $\det A = \det A' \neq 0$. □

Przykład 6.15.

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

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

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

Przykład 6.16 (wyznacznik Vandermonde'a).

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

Wniosek 6.17. *Jeżeli* $\dim V = n < \infty$ *i* B *jest bazą* V*, to dla każdych* $v_1,\ldots,v_n \in V$ *zachodzi*

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

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

Przykład 6.18. Rozważmy wektory postaci $\begin{pmatrix} 1 \\ t \\ t^2 \end{pmatrix} \in \mathbf{R}^3$. Trzy takie wektory (dla t_1, t_2, t_3) są liniowo niezależne wtedy i tylko wtedy gdy

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

Definicja 6.19. Mówimy że macierz A jest *osobliwa*, jeżeli $\det A = 0$. W przeciwnym wypadku mówimy, że jest *nieosobliwa*.

Wniosek 6.20. *Macierz $A \in \mathrm{M}_{n\times n}(K)$ jest odwracalna $\iff$ A jest nieosobliwa.*

Dowód.

$$\begin{aligned} A \text{ jest odwracalna} &\iff F_A\colon K^n \to K^n \text{ jest izomorfizmem} \\ &\iff F_A \text{ jest „na”} \\ &\iff \operatorname{im} F_A = K^n \\ &\iff \dim \operatorname{im} F_A = n \\ &\iff \operatorname{rk} A = n \\ &\iff \text{kolumny } A \text{ są lnz} \\ &\iff \det A \neq 0 \end{aligned}$$

□

Definicja 6.21. *Minor* macierzy A (niekoniecznie kwadratowej) to macierz kwadratowa powstała z A przez wykreślenie pewnej liczby wierszy lub kolumn, lub wyznacznik takiej macierzy.

Przykład 6.22. (rysunek)

Stwierdzenie 6.23. *Niech A będzie dowolną macierzą (niekoniecznie kwadratową). Wtedy* $\operatorname{rk} A$ *= maksymalny rozmiar niezerowego minora.*

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

$\leqslant$: załóżmy że kolumny o indeksach $j_1, \ldots, j_k$ 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. □

Przykład 6.24. Chcemy wyznaczyć wymiar przestrzeni rozwiązań UJ

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

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

Stwierdzenie 6.25. *Każdy niezerowy minor można rozszerzyć do niezerowego minora maksymalnego wymiaru.*

Dowód. Weźmy niezerowy minor $k \times k$ macierzy A o $\operatorname{rk} A > k$. Występujące w nim kolumny są liniowo niezależnymi wektorami w obrazie F_A. Możemy do nich dobrać inne kolumny A, otrzymując bazę $\operatorname{im} F_A$. 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 □

Uwaga 6.26. Powyższe stwierdzenie można zinterpretować następująco: jeżeli $F : K^n \to K^m$ jest dowolnym odwzorowaniem liniowym, to dla pewnych $i_1 < i_2 < \ldots < i_k \leqslant n$ i $j_1 < j_2 < \ldots < j_k \leqslant m$, gdzie $k = \operatorname{rk} F$, odwzorowanie

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

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

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

Rozważamy równanie

$$AX = Y, \tag{$*$}$$

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

Twierdzenie 6.27 (Wzory Cramera)**.** *Dla A, Y jak powyżej, rozwiązaniem ($*$) jest $X = (x_1, \ldots, x_n)^\top$, gdzie*

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

Dowód. Sprawdzimy że $AX = Y$.

$$\begin{aligned}
(AX)_i &= \sum_{j=1}^n a_{ij} x_j \\
&= \sum_{j=1}^n a_{ij} \frac{1}{\det A} \cdot \sum_{l=1}^n (-1)^{l+j} y_l A_{l,j} \quad \text{(rozwinięcie Laplace'a w j-tej kolumnie)} \\
&= \sum_{l=1}^n \frac{y_l}{\det A} \underbrace{\sum_{j=1}^n (-1)^{l+j} a_{ij} A_{lj}}_{\begin{cases} 0 & l \neq i \\ \det A & l = i \end{cases}} \quad \text{(zmiana kolejności sumowania)} \\
&= \sum_{l=1}^n \begin{cases} \frac{y_l}{\det A} \cdot \det A & l = i \\ \frac{y_1}{\det A} \cdot 0 & l \neq i \end{cases} \\
&= y_i.
\end{aligned}$$

Wyjaśnienie „klamerki”:

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

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. □

Definicja 6.28. 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. $\mathrm{adj}(A) = (b_{ij})_{i,j}$, gdzie $b_{ij} = (-1)^{i+j}A_{ji}$.

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

(Uwaga na znaki i na transpozycję!)

Wniosek 6.29. *Niech* $A \in M_{n\times n}(K)$*. Wtedy:*

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

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

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

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

Przykład 6.30.

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

$$
\begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \frac{1}{\begin{vmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{vmatrix}} \cdot \operatorname{adj} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix} & -\begin{vmatrix} 0 & 1 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 0 & 2 \\ 0 & 0 \end{vmatrix} \\ -\begin{vmatrix} 2 & 3 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 0 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 0 & 0 \end{vmatrix} \\ \begin{vmatrix} 2 & 3 \\ 2 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 0 & 2 \end{vmatrix} \end{pmatrix}^{\top} =
$$

$$
= \frac{1}{2} \begin{pmatrix} 2 & 0 & 0 \\ -2 & 1 & 0 \\ -4 & -1 & 2 \end{pmatrix}^{\top} = \frac{1}{2} \begin{pmatrix} 2 & -2 & -4 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & -1 & -2 \\ 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix}
$$

Wyznacznik i wielomian charakterystyczny odwzorowania liniowego

Definicja 6.31. Niech V, W będą przestrzeniami liniowymi. Wtedy $\mathrm{Hom}(V, W)$ to zbiór odwzorowań liniowych $V \to W$.

Fakt 6.32. $\mathrm{Hom}(V, W)$ *jest przestrzenią liniową. (Podobnie jak* $\mathrm{Hom}(V, K) = V^*$*.)*

Fakt 6.33. *Zbiór* $\mathrm{M}_{k\times n}(K)$ *jest przestrzenią liniową (ze zwykłym dodawaniem macierzy i mnożeniem przez skalary), izomorficzną z* K^{kn}*.*

Fakt 6.34. *Jeżeli B jest bazą V, a C jest bazą W i V, W są skończenie wymiarowe, to odwzorowanie*

$$\mathrm{Hom}(V, W) \ni F \mapsto m^B_C(F) \in M_{\dim W \times \dim V}(K)$$

jest izomorfizmem.

Dowód. Bijekcja — było. Addytywność:

$$\underbrace{[(F+G)(b_i)]_C}_{\substack{\| \\ i\text{-ta kolumna } m^B_C(F+G)}} = [F(b_i) + G(b_i)]_C = \underbrace{[F(b_i)]_C}_{\substack{\| \\ i\text{-ta kolumna } B}} + \underbrace{[G(b_i)]_C}_{\substack{\| \\ i\text{-ta kolumna } C}}.$$

Jednorodność podobnie. □

Wniosek 6.35. *Jeżeli V, W są skończenie wymiarowe, to* $\dim \mathrm{Hom}(V, W) = \dim V \cdot \dim W$.

Uwaga 6.36. Ogólnie $\dim \mathrm{Hom}(V, W) = \dim W \cdot \dim V^*$ (to może być na konwersatorium, przy okazji iloczynów tensorowych).

Przykład 6.37. Rozważmy $F \colon V \to V$ (V: skończenie wymiarowa).

Wtedy

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

Dla dowolnej B — bazy V mamy:

$$m^B_B(F - \alpha\,\mathrm{id}_V) = m^B_B(F) - \alpha m^B_B(\mathrm{id}_V) = m^B_B(F) - \alpha \mathrm{I}.$$

Stwierdzenie 6.38. *Jeżeli* $F \in \mathrm{Hom}(V, V)$, *to* $\det m_B^B(F)$ *nie zależy od wyboru* B *— bazy* V.

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

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

□

Definicja 6.39. Jeżeli $F \colon V \to V$ i $\dim V < \infty$, definiujemy

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

gdzie B jest dowolną bazą V.

Definiujemy też *wielomian charakterystyczny* F jako

$$\chi_F(x) := \det(F - x \cdot \mathrm{id}_V) = \det(m_B^B(F) - x\mathrm{I})$$

Ślad F to suma wyrazów na przekątnej w $m_B(F)$, czyli

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

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

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

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

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

Rozdział 7

Sumy proste i liniowa niezależność podprzestrzeni

Definicja 7.1. Niech $V_1, V_2, \ldots, V_n \leqslant V$. Mówimy że V jest (wewnętrzną) *sumą prostą* $V_1, \ldots, V_n$, $V = V_1 \oplus V_2 \oplus \ldots \oplus V_n$, jeżeli każdy $v \in V$ przedstawia się *jednoznacznie* w postaci

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

gdzie $v_k \in V_k$.

Przykład 7.2. 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 (zewnętrzną!) sumą prostą/produktem dwóch kopii $\mathbf{R}$; zewnętrzna suma prosta to *operacja* na przestrzeniach liniowych, wewnętrzna to *relacja* pomiędzy przestrzenią a ciągiem jej podprzestrzeni!)

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

$$V = \mathrm{Lin}(B_1) \oplus \mathrm{Lin}(B_2) \oplus \mathrm{Lin}(B_3) \ldots \oplus \mathrm{Lin}(B_n).$$

Dowód dla $n = 2$ (ogólny przypadek jest analogiczny): niech $B_1 = \{b_1, \ldots, b_m\}$, $B_2 = \{b_{m+1}, b_{m+2}, \ldots, b_l\}$. Weźmy dowolne $v \in V$ i zapiszmy je w bazie $B = \{b_1, \dots, b_l\}$:

$$v = \sum_{k=1}^{l} \alpha_k b_k = \underbrace{\sum_{k=1}^{m} \alpha_k b_k}_{\substack{\| \\ v_1}} + \underbrace{\sum_{k=m+1}^{l} \alpha_k b_k}_{\substack{\| \\ v_2}},$$

co daje „zapisywanie się”. Jednoznaczność wynika z jednoznaczności przedstawienia v jako kombinacji liniowej elementów bazy B.

3.

$$\begin{aligned} \mathbf{R}_8[x] &= \{\, a + bx^2 + cx^4 + dx^6 + ex^8 \mid a, b, c, d, e \in \mathbf{R} \,\} \oplus \{\, ax + bx^3 + cx^5 + dx^7 \mid a, b, c, d \\ &= \{\, P \in \mathbf{R}_8[x] \mid P(-x) = P(x) \,\} \oplus \{\, P \in \mathbf{R}_8[x] \mid P(x) = -P(-x) \,\} \end{aligned}$$

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

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

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

$$v = \underset{\substack{\| \\ v_1}}{\begin{pmatrix} x \\ 0 \end{pmatrix}} + \underset{\substack{\| \\ v_2}}{\begin{pmatrix} 0 \\ y \end{pmatrix}} + \underset{\substack{\| \\ v_3}}{\begin{pmatrix} 0 \\ 0 \end{pmatrix}} = \underset{\substack{\| \\ v_1'}}{\begin{pmatrix} x \\ 0 \end{pmatrix}} + \underset{\substack{\| \\ v_2'}}{\begin{pmatrix} 0 \\ 0 \end{pmatrix}} + \underset{\substack{\| \\ v_3'}}{\begin{pmatrix} 0 \\ y \end{pmatrix}},$$

a więc niejednoznacznie.

Definicja 7.3. Niech $V_1, \dots, V_k \leqslant V$. Mówimy że $V_1, \dots, V_k$ są liniowo niezależne, jeżeli dla $v_i \in V_i$ mamy

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

Uwaga 7.4. 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 $\mathrm{Lin}(u_1), \mathrm{Lin}(u_2), \dots, \mathrm{Lin}(u_k)$ są lnz.

Dowód. Załóżmy że $u_1, \dots, u_k$ są lnz. Weźmy $v_i \in \mathrm{Lin}(u_i)$ 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 $u_i = 0$, czyli $v_i = 0$.

W drugą stronę, załóżmy że $\mathrm{Lin}(u_i)$ są lnz i weźmy dowolne $\alpha_1, \dots, \alpha_k$ takie że $\sum_i \alpha_i u_i = 0$. Wtedy $v_i = \alpha_i u_i \in \mathrm{Lin}(u_i)$, więc z założenia $v_i = 0$, więc (ponieważ $u_i \neq 0$), $\alpha_i = 0$, czyli u_i są lnz. □

Lemat 7.5. *$V_1, \dots, V_n$ 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 $v = -v_i = 0$.

Załóżmy 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 $0 \neq -v_i = \sum_{j \neq i} v_j \in V_i \cap \sum_{j \neq i} V_j$. □

Wniosek 7.6. *$V_1, V_2 \leqslant V$ są lnz $\iff$ $V_1 \cap V_2 = \{0\}$.*

Fakt 7.7. *Niech $V_1, \dots, V_n \leqslant V$. Wtedy NWSR:*

1. *$V = V_1 \oplus V_2 \dots \oplus V_n$.*
2. *$\Phi\colon V_1 \times V_2 \times \dots \times V_n \to V$ 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” $\iff V = \operatorname{im}\Phi = V_1 + \ldots + V_n$.

$2+3 \implies 4,5$ łatwe, bo $\dim(V_1 \times \ldots \times V_n = \dim V_1 + \ldots + \dim V_n$ i Φ zachowuje wymiar (jako izomorfizm).

$4 \implies 2$: skoro $V_1, \ldots, V_n$ 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 $\dim\operatorname{im}\Phi = \dim\operatorname{dom}\Phi - 0 = \dim V - 0 = \dim V$. Skoro $\dim V$ jest skończony, to $\operatorname{im}\Phi = V$.

$5 \implies 2$: skoro $V_1 + \ldots + V_n = V$, to Φ jest „na”, i podobnie z równości wymiarów i twierdzenia o indeksie wynika że $\dim\ker\Phi = 0$, czyli Φ jest 1-1. □

Wniosek 7.8. *($n = 2$, $\dim V < \infty$.)NWSR:*

1. $V = V_1 \oplus V_2$,
2. $V_1 \cap V_2 = \{0\}$ *i* $V_1 + V_2 = V$,
3. $V_1 \cap V_2$ *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 7.9. Jeżeli $W \leqslant V$, to przestrzeń $W' \leqslant V$ jest *dopełnicza* do W w V jeżeli $V = W \oplus W'$.[rysunek]

Lemat 7.10. *Jeżeli V jest przestrzenią liniową i $W \leqslant V$, to W jest dopełnialna w V, tzn. V ma podprzestrzeń W' dopełniczą do V.*

Dowód. Niech B_1 będzie bazą W. Wtedy możemy $B_1 \subseteq B$ dla pewnej bazy B przestrzeni V i $W' := \operatorname{Lin}(B \setminus B_1)$ działa (patrz Przykład 7.2.2 powyżej). □

Uwaga 7.11. Jeżeli rozważamy przestrzenie liniowo-topologiczne (TVS), tzn. przestrzenie liniowe z kompatybilną topologią, to powyższy lemat nie jest prawdziwy, tzn. jeżeli V jest (nieskończenie wymiarową) TVS i $W \leqslant V$ jest domkniętą podprzestrzenią, to zazwyczaj nie ma domkniętej $W' \leqslant V$ dopełniczej do W (intuicyjnie: przestrzenie dopełnicze są na ogół bardzo dziwne i trudno je zrozumieć). Wtedy pytanie o to, które podprzestrzenie są dopełnialne (mają domknięte dopełnienie) jest dużo ciekawsze.

Definicja 7.12. Niech $W \leqslant V$, 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 7.13. $W = V$ i $W = \{0\}$ zawsze są niezmiennicze.

Przykład 7.14. 0. Jeżeli $F(v) = \lambda v$, to $\mathrm{Lin}(v)$ jest F-niezmiennicza (bo dla $w \in \mathrm{Lin}(v)$ mamy $w = \alpha v$, więc $F(w) = \alpha\lambda v \in \mathrm{Lin}(v)$).

Uwaga 7.15. Ogólnie, jeżeli $F \in \mathrm{End}(V)$, to 1-wymiarowe przestrzenie F-niezmiennicze to dokładnie przestrzenie powyższej postaci (tzn. proste rozpinane przez wektory własne).

1. Rozważmy $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \in \mathrm{M}_{2\times 2}(\mathbf{R})$. Jakie są F_A-niezmiennicze podprzestrzenie $\mathbf{R}^2$?
   - $\{0\}$,
   - $\mathbf{R}^2$.
   - $\mathrm{Lin}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)$ (wektor własny).
   - Jeżeli $W = \mathrm{Lin}(v)$ dla $v = \begin{pmatrix} x \\ y \end{pmatrix}$ dla $y \neq 0$, 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 $A = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \in \mathrm{M}_{2\times 2}(\mathbf{R}), \mu \neq \lambda$, to przestrzenie F_A-niezmiennicze 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 $\mathrm{Lin}(v)$ nie jest F_A-niezmiennicza.

Fakt 7.16. *Jeżeli $W \leqslant V$ jest F-niezmiennicze, to mamy dobrze określone odwzorowanie liniowe $\bar{F}\colon V/W \to V/W$ 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. □

Lemat 7.17. *Niech F będzie endomorfizmem V,* $\dim(V) < \infty$.

1. *Jeżeli $V = V_1 \oplus V_2 \oplus \ldots \oplus V_n$, gdzie V_i są F-niezmiennicze, to*

$$\chi_F(x) = \chi_{F\restriction_{V_1}}(x) \cdot \chi_{F\restriction_{V_2}}(x) \cdots \chi_{F\restriction_{V_n}}(x)$$

2. *Niech $W \leqslant V$ będzie F-niezmiennicza. Określmy odwzorowanie $\bar{F}: V/W \to V/W$ wzorem $F(v+W) = v+W$. Wtedy $\bar{F}$ jest dobrze określone i $\chi_F(x) = \chi_{F\restriction_W}(x)\cdot\chi_{\bar{F}}(x)$.*

Dowód. Przypomnienie z ćwiczeń: 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

$$\left|\begin{array}{c|c|c|c} A_1 & * & * & * \\ \hline 0 & A_2 & * & * \\ \hline 0 & 0 & \ddots & \vdots \\ \hline 0 & 0 & \cdots & A_n \end{array}\right| = \prod_{k=1}^{n} \det(A_k)$$

1.: niech $B_1, B_2, \ldots, B_n$ będą bazami V i niech $B = B_1 B_2 \ldots B_n$ (w tej kolejności). Wtedy mamy

$$m_B(F) = \left(\begin{array}{c|c|c|c} m_{B_1}(F\restriction_{V_1}) & 0 & 0 & 0 \\ \hline 0 & m_{B_2}(F\restriction_{V_2}) & 0 & 0 \\ \hline 0 & 0 & \ddots & \vdots \\ \hline 0 & 0 & \cdots & m_{B_n}(F\restriction_{V_n}). \end{array}\right)$$

Istotnie, z założenia dla $b \in B_i$ mamy $F(b) \in V_i$, $F(b)$ wyraża się jako kombinacja liniowa wektorów z B_i, czyli $[F(b)]_B = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ [F(b)]_{B_i} \\ \vdots \\ 0 \end{pmatrix} (= [F\restriction_{V_i}(b)]_{B_i})$,

zatem kolumna odpowiadająca b jest taka jak podana, a teza wynika z obserwacji powyżej.

2.: niech $B_1 = (b_1^1, \ldots, b_n^1)$ będzie bazą W, a $B = (b_1^1, \ldots, b_n^1, b_1^2, \ldots, b_m^2)$ — rozszerzającą ją bazą V. Wtedy łatwo zauważyć, że $\bar{B}_2 = (b_1^2 + W, b_2^2 + W, \ldots, b_m^2 + W)$ są bazą V/W. Zachodzi wzór:

$$m_B(F) = \left(\begin{array}{c|c} m_{B_1}(F\restriction_W) & * \\ \hline 0 & m_{\bar{B}_2}(\bar{F}) \end{array}\right).$$

Dla kolumn pochodzących z B_1 wynika to z rozumowania jak powyżej. Dla kolumn pochodzących z $B_2 = B \setminus B_1$: jeżeli weźmiemy pewne b_j^2, to można zapisać $F(b_j^2) = (\sum_{i=1}^n \alpha_i b_i) + \sum_{i=1}^m \beta_i b_i^2$. Stąd

$$[\bar{F}(b_j^2 + W)]_{\bar{B}_2} = \Bigg[\underbrace{\left(\sum_{i=1}^{n} \alpha_i b_i^1 + W\right)}_{\substack{\|\\ W}} + \sum_{i=1}^{m} \beta_i b_i^2 + W\Bigg]_{\bar{B}_2} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix},$$

a więc

$$[F(b_j^2)]_B = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \\ \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} = \begin{pmatrix} * \\ [\bar{F}(b_j^2 + W)]_{\bar{B}_2} \end{pmatrix},$$

więc pozostałe kolumny też się zgadzają. Stąd wynika wzór z lematu. □

Przykład 7.18. Dla $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$, $W = OX$ mamy $m_{e_1}(F_A)\restriction_W = (2)$ i $m_{e_2+W}(\bar{F}) = (2)$, więc $\chi_A(x) = (2-x)\cdot(2-x)$.

Definicja 7.19. Niech F będzie endomorfizmem V, a α będzie skalarem. 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* (dla λ).

Zbiór $\sigma(F) = \sigma_p(F)\,\mathrm{Spec}(F) := \{\lambda \in K \mid V_\lambda \neq \{0\}\}$ nazywamy *spektrum [punktowym]* (lub *widmem*) F. (Dla nieskończenie wymiarowych przestrzeni do spektrum (nie punktowego) włącza się czasami też inne skalary.)

Przykład 7.20. 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.

Niech $V = \ell^2 = \{(a_n)_{n\in\mathbf{N}} \mid \sum_n |a_n|^2 < \infty\}$, a $F : \ell^2 \to \ell^2$ będzie prawym sziftem (dopisaniem zera na początku). Wtedy F nie ma wartości własnych, nawet zespolonych: $\sigma_p(F) = \emptyset$.

Uwaga 7.21.

- Każda V_λ jest niezmienniczą podprzestrzenią V.
- $\lambda \in \sigma(F) \iff \chi_F(\lambda) = 0$.

Dowód. Pierwszy punkt: $V_\lambda = \ker(F - \lambda \cdot \mathrm{id})$, więc jest podprzestrzenią. Niezmienniczość wynika łatwo z definicji.

Drugi punkt:

$$\begin{aligned} \lambda \in \sigma(F) &\iff \ker(F - \lambda \cdot \mathrm{id}) \neq \{0\} \\ &\iff F - \lambda \cdot \mathrm{id} \text{ nie jest odwracalne} \\ &\iff \underset{\chi_F(\lambda)}{\underset{\parallel}{\det(F - \lambda \cdot \mathrm{id})}} = 0 \end{aligned}$$ □

Wniosek 7.22. *$F : V \to V$ ma co najwyżej $\dim V$ wartości własnych.*

Dowód. $\chi_F(x)$ jest wielomianem stopnia $\dim V$, więc ma co najwyżej n pierwiastków (z twierdzenia Bezout). □

Fakt 7.23. *Załóżmy że $\lambda_1, \lambda_2, \ldots, \lambda_n \in K$ są parami różne. Wtedy przestrzenie $V_{\lambda_1}, \ldots, V_{\lambda_n}$ 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

$$\sum_{i=1}^n (\lambda_i - \lambda_1) v_i = \sum_{i=2}^n (\lambda_i - \lambda_1) v_i = 0 - 0 = 0.$$

Ponieważ $V_{\lambda_2}, \ldots V_{\lambda_n}$ 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 $0 = \sum_{i=1}^n v_i = v_1$. □

alternatywny dowód. Załóżmy że $0 = \sum_{i=1}^{n} v_i$, gdzie $v_i \in V_{\lambda_i}$. Wtedy

$$\begin{cases} 0 = v_1 + v_2 + \ldots + v_n \\ F(0) = 0 = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n \\ F^2(0) = 0 = \lambda_1^2 v_1 + \lambda_2^2 v_2 + \ldots + \lambda_n^2 v_n \\ \vdots \\ F^{n-1}(0) = 0 = \lambda_1^{n-1} v_1 + \lambda_2^{n-1} v_2 + \ldots + \lambda_n^{n-1} v_n \end{cases}$$

Sposób pierwszy: zauważyć że z powyższego wynika

$$\begin{pmatrix} \mathrm{I} & \mathrm{I} & \cdots & \mathrm{I} \\ \lambda_1 \cdot \mathrm{I} & \lambda_2 \cdot \mathrm{I} & \cdots & \lambda_n \cdot \mathrm{I} \\ \vdots & & \ddots & \vdots \\ \lambda_1^{n-1}\mathrm{I} & \lambda_2^{n-1}\mathrm{I} & \cdots & \lambda_n^{n-1}\mathrm{I} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = 0,$$

a wyznacznik dużej macierzy po prawej to wyznacznik Vandermonde'a podniesiony do potęgi $\dim V$, czyli $v_i = 0$.

Drugi sposób: kolumny $w_i = \begin{pmatrix} 1 \\ \lambda_i \\ \vdots \\ \lambda_i^{n-1} \end{pmatrix} \in K^n$ są liniowo niezależne (wyznacznik Vandermonde'a), więc są bazą K^n. Istnieje zatem baza dualna $\varphi_1, \ldots, \varphi_n$, 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

$$\vec{0} = \sum_{k=0}^{n-1} a_{ik} \underbrace{\sum_{j=1}^{n} \lambda_j^k v_j}_{\substack{\parallel \\ 0}} = \sum_{j=1}^{n} \Bigg(\underbrace{\sum_{k=0}^{n-1} (a_{ik} \lambda_j^k)}_{\substack{\parallel \\ \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}}} \Bigg) v_j = v_i$$

□

Definicja 7.24. Mówimy że $F\colon V \to V$ jest *diagonalizowalne* jeżeli $V = \bigoplus_{\lambda\in\sigma(F)} V_\lambda$.

(Wystarczy sprawdzić, że $V = \sum_{\lambda\in\sigma(F)} V_\lambda$, lub równoważnie, $\sum_{\lambda\in\sigma(\lambda)} \dim V_\lambda = \dim V$.)

Stwierdzenie 7.25. *Następujące warunki są równoważne dla $F\colon V \to V$:*

- *F jest diagonalizowalne,*
- *istnieje baza B przestrzeni V taka że $m_B(F)$ jest diagonalna (wtedy mówimy że 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 α_i $\iff$ b_i jest wektorem własnym.

Załóżmy że F jest diagonalizowalne. Niech $\lambda_1, \dots, \lambda_n$ będą (wszystkimi, różnymi) wartościami własnymi dla F i niech B_i będzie bazą V_{λ_i}. Wtedy z liniowej niezależności V_{λ_i} wynika, że $B_1 B_2 \dots B_n$ jest liniowo niezależny. Z drugiej strony łatwo zobaczyć, że rozpina V i składa się z wektorów własnych.

Załóżmy że V ma bazę B wektorów własnych dla F. Wtedy

$$\sum_{\lambda\in\sigma(F)} V_\lambda \supseteq \sum_{\lambda\in\sigma(F)} \operatorname{Lin}\underbrace{\{\, b\in B \mid \text{wartość własna } b \text{ to } \lambda \,\}}_{\substack{\| \\ B_\lambda}} = \operatorname{Lin}\Big(\bigcup_{\lambda\in\sigma(F)} B_\lambda\Big) = \operatorname{Lin}(B) = V,$$

więc F jest diagonalizowalna. □

Uwaga 7.26.
- Jeżeli $m_B(F)$ jest diagonalna, to dla dowolnej bazy C istnieje odwracalna macierz $P(= m_C^B(\mathrm{id}))$ taka że $m_C(F) = P \underbrace{m_B(F)}_{\substack{\| \\ D}} P^{-1}$.
- 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).

Przykład 7.27. $A = \begin{pmatrix} 2 & 1 & & \\ 0 & 2 & & \\ & & 3 & 0 \\ & & 0 & 3 \end{pmatrix}$ $\chi_A(x) = (2-x)^2(3-x)^2$, $\sigma(F_A) = \{2,3\}$,

$$V_2 = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in K^4 \;\middle|\; \begin{cases} 2x + y = 2x \\ 2y = 2y \\ 3z = 2z \\ 3t = 2t \end{cases} \right\} = \mathrm{Lin} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$V_2 = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in K^4 \;\middle|\; \begin{cases} 2x + y = 3x \\ 2y = 3y \\ 3z = 3z \\ 3t = 3t \end{cases} \right\} = \mathrm{Lin} \left(\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right)$$

$\dim V_2 + \dim V_3 = 1 + 2 < 4 = \dim V$, więc A nie jest diagonalizowalna.

Definicja 7.28. Mówimy że ciało K jest *algebraicznie domknięte* jeżeli każdy niestały wielomian o współczynnikach z K ma pierwiastek w K

Uwaga 7.29.

- Ciało $\mathbf{C}$ jest algebraicznie domknięte (to tzw. zasadnicze twierdzenie algebry[1], nietrywialne, ma wiele pięknych dowodów, można je poznać np. na wykładzie z funkcji analitycznych, topologii algebraicznej, lub topologii różniczkowej), ciało $\mathbf{R}$ nie jest, bo np. x^2+1 nie ma pierwiastka.

- Z twierdzenia Bézout wynika, że K jest algebraicznie domknięte $\iff$ każdy wielomian rozkłada się na iloczyn liniowych czynników (tzn. wielomianów postaci $x - \alpha$).

- Każde ciało można rozszerzyć do ciała algebraicznie domkniętego (łatwiejsze twierdzenie, ale nie z algebry liniowej).

Twierdzenie 7.30 (twierdzenie Jordana)**.** *Niech V będzie skończenie wymiarową przestrzenią liniową nad algebraicznie domkniętym ciałem K. Wtedy dla*

[1]wbrew nazwie to twierdzenie nie jest specjalnie zasadnicze i nie jest twierdzeniem algebry

każdego endomorfizmu $F\colon V \to V$ istnieje baza V, taka że

$$m_B(F) = \left(\begin{array}{c|c|c|c} J_1 & 0 & 0 & 0 \\ \hline 0 & J_2 & 0 & 0 \\ \hline 0 & 0 & \ddots & \vdots \\ \hline 0 & 0 & \cdots & J_n \end{array}\right)$$

*(*postać Jordana*).*

gdzie każde J_i jest macierzą postaci $J_i = J_{\lambda_i,d_i} := \underbrace{\begin{pmatrix} \lambda_i & 1 & & & \mathbf{0} \\ & \lambda_i & 1 & & \\ & & \lambda_i & 1 & \\ & & & \ddots & \ddots \\ & & & & 1 \\ \mathbf{0} & & & & \lambda_i \end{pmatrix}}_{d_i \text{ kolumn}}$

Dowód. Indukcja względem wymiaru V. Jeżeli $\dim V = 1$, to teza jest oczywista. Załóżmy ze $\dim V > 1$ i teza zachodzi dla przestrzeni niższych wymiarów.

Ponieważ K jest algebraicznie domknięte, $\chi_F(x)$ ma pierwiastek λ.

Możemy założyć bez zmniejszania ogólności, że $\lambda = 0$: $G := F - \lambda\,\mathrm{id}$ spełnia $\chi_G(0) = 0$ (wektor własny Fdla λ jest wektorem własnym G dla 0) i jeżeli

$$m_C(G\restriction_W) = \left(\begin{array}{c|c|c|c} J_{\mu_1,f_1} & 0 & 0 & 0 \\ \hline 0 & J_{\mu_2,f_2} & 0 & 0 \\ \hline 0 & 0 & \ddots & \vdots \\ \hline 0 & 0 & \cdots & J_{\mu_m,f_m} \end{array}\right)$$

to

$$m_C(F) = m_C(G + \lambda\,\mathrm{id}) = m_C(G) + \lambda\mathrm{I} = \left(\begin{array}{c|c|c|c} J_{\mu_1+\lambda,f_1} & 0 & 0 & 0 \\ \hline 0 & J_{\mu_2+\lambda,f_2} & 0 & 0 \\ \hline 0 & 0 & \ddots & \vdots \\ \hline 0 & 0 & \cdots & J_{\mu_m+\lambda,f_m} \end{array}\right)$$

Skoro $0 \in \sigma(F)$, to $W = \operatorname{im} F$ jest właściwą podprzestrzenią V, która jest też F-niezmiennicza, więc $F\restriction_W\colon W \to W$ spełnia założenie indukcyjne i w pewnej bazie C przestrzeni W jego macierz jest postaci

$$m_C(F\restriction_W) = \left(\begin{array}{c|c|c|c} J_1' & 0 & 0 & 0 \\ \hline 0 & J_2' & 0 & 0 \\ \hline 0 & 0 & \ddots & \vdots \\ \hline 0 & 0 & \cdots & J_m' \end{array}\right)$$

Zmieniając kolejność wektorów bazowych, możemy założyć że pierwsze k z klatek Jordana mają zera na przekątnej (być może $k = 0$, kiedy takich klatek nie ma). Zapisując $C = C_1C_2\ldots C_m$, gdzie $C_i = (c_{i1}, c_{i2}, \ldots, c_{if_i})$ dla $i = 1, \ldots, m$ to część bazy C odpowiadająca $J_i' = \underbrace{\begin{pmatrix} \mu_i & 1 & & & \mathbf{0} \\ & \mu_i & 1 & & \\ & & \mu_i & 1 & \\ & & & \ddots & \ddots & \\ & & & & \ddots & 1 \\ \mathbf{0} & & & & & \mu_i \end{pmatrix}}_{f_i \text{ kolumn}}$,

mamy wtedy $F(c_{i1}) = \mu_i c_{i1}$ oraz $F(c_{i(j+1)}) = \mu_i c_{i(j+1)} + c_{ij}$ dla $j < f_i$.

W szczególności dla $i \leqslant k$ mamy $\mu_i = k$, więc $c_{11}, \ldots, c_{1k} \in \ker F$. Są one liniowo niezależne, więc istnieje baza $c_{11}, \ldots, c_{k1} b_1, \ldots, b_l$ przestrzeni $\ker F$.

Ponieważ z założenia $C \subseteq W = \operatorname{im} F$, możemy znaleźć $c_{1(f_1+1)}, \ldots, c_{k(f_k+1)} \in V$ takie że $F(c_{i(f_i+1)}) = c_{if_i}$.

Pokażemy że układ $B = Cc_{1(f_1+1)}, \ldots, c_{k(f_k+1)} b_1, \ldots, b_l$ jest bazą V. Zauważmy że B ma $\dim V$ elementów: C ma $\dim W = \dim \operatorname{im} F$ elementów, $l = \dim \ker F - k$, więc łącznie mamy $|B| = |C| + k + \dim \ker F - k = \dim \operatorname{im} F + \dim \ker F = \dim V$.

Wystarczy więc pokazać, że B jest liniowo niezależny. Weźmy

$$\begin{aligned}
0 &= \sum_{b\in B}\alpha_b b\\
&= \sum_{i=1}^{m}\sum_{j=1}^{f_i}\alpha_{ij}c_{ij} + \sum_{i=1}^{k}\alpha_{i(f_i+1)}c_{i(f_i+1)} + \sum_{i=1}^{l}\beta_i b_i
\end{aligned}$$

$$0 = F(0) = \sum_{i=1}^{m}\sum_{j=1}^{f_i}\alpha_{ij}\underbrace{F(c_{ij})}_{\begin{cases}\mu_i c_{ij} + c_{i(j-1)} & j>1\\ \mu_i c_{ij} & j=1\end{cases}} + \sum_{i=1}^{k}\alpha_{if_i+1}\underbrace{F(c_{i(f_i+1)})}_{c_{if_i}} + \sum_{i=1}^{l}\beta_i\underbrace{F(b_i)}_{0}$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{f_i}\alpha_{i(j+1)}c_{ij} + \sum_{i=k+1}^{m}\sum_{j=1}^{f_i-1}(\alpha_{i(j+1)}+\mu_i\alpha_{ij})c_{ij} + \sum_{i=k+1}^{m}\mu_i\alpha_{if_i}c_{if_i}.$$

W ostatnim wierszu mamy już kombinację liniową elementów C, więc z jego liniowej niezależności wnioskujemy, że dla $i>k$ mamy $\mu_i\alpha_{if_i}=0$, więc (ponieważ wtedy $\mu_i\neq 0$) też $\alpha_{if_i}=0$. Podobnie też $0=(\alpha_{if_i}+\mu_i\alpha_{i(f_i-1)})$, podstawiając $\alpha_{if_i}=0$ otrzymujemy $\mu_i\alpha_{i(f_i-1)}=0$ itd., tak że $\alpha_{ij}=0$ dla $i>k$. Dla $i\leqslant k$ i $j>1$ dostajemy wprost $\alpha_{ij}=0$, pozostają więc jedynie α_{i1} dla $i\leqslant k$ oraz β_i. Stąd mamy

$$\begin{aligned}
0 &= \sum_{i=1}^{m}\sum_{j=1}^{f_i}\alpha_{ij}c_{ij} + \sum_{i=1}^{k}\alpha_{i(f_i+1)}c_{i(f_i+1)} + \sum_{i=1}^{l}\beta_i b_i\\
&= \sum_{i=1}^{k}\alpha_{i1}c_{i1} + \sum_{i=1}^{l}\beta_i b_i,
\end{aligned}$$

ale teraz mamy kombinację liniową elementów z bazy $\ker F$, więc liniowo niezależnych. Stąd pozostałe współczynniki również są zerowe, a więc B jest liniowo niezależny. Uporządkujmy B jako

$$B' = b_1 b_2 \dots b_l C_1 c_{1(f_1+1)} C_2 c_{2(f_2+1)} \dots C_k c_{k(f_k+1)} C_{k+1} C_{k+2} \dots C_m$$

Wtedy mamy

$$m_{B'}(F) = \left(\begin{array}{c|c|c|c|c}
0 & 0 & 0 & 0 & 0\\ \hline
0 & J_1'' & 0 & 0 & 0\\ \hline
0 & 0 & J_2'' & 0 & 0\\ \hline
0 & 0 & 0 & \ddots & \vdots\\ \hline
0 & 0 & 0 & \cdots & J_m''
\end{array}\right)$$

gdzie w lewym górnym rogu mamy macierz zerową rozmiaru $l \times l$, $J_i'' = J_i'$ (z $m_C(F\restriction_W)$) dla $i > k$, a dla $i \leqslant k$, J_i'' to J_i' powiększona o 1. Ta macierz jest w postaci Jordana: blok zer w lewym górnym rogu jest w postaci Jordana, składa się z klatek 1×1 z $\lambda = 0$. □

Przykład 7.31. 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 $(A-\mathrm{I})^2 \neq 0$, więc postać Jordana A to musi być $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ (w pozostałych przypadkach $(A-\mathrm{I})^2 = 0$).

Aby znaleźć bazę Jordana, musimy znaleźć wektor v taki że $(A-\mathrm{I})^2 v \neq 0$. Sprawdzimy, czy $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ działa (uwaga: $\ker(A-\mathrm{I})^2$ jest płaszczyzną, więc wystarczy że strzelimy wektor spoza tej płaszczyzny).

$$(A\text{–}\mathrm{I})^2 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 & -1 & 2 \\ 2 & -1 & 3 \\ -2 & 1 & -1 \end{pmatrix}^2 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 & -1 & 2 \\ 2 & -1 & 3 \\ -2 & 1 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \\ -2 \end{pmatrix} = \begin{pmatrix} -2 \\ -4 \\ 0 \end{pmatrix} \neq \vec{0}.$$

Udało się! Stąd baza Jordanizująca to np. $\begin{pmatrix} -2 \\ -4 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

Uwaga 7.32.

- Macierz w postaci Jordana jest macierzą górnotrójkątną, stąd wartości na przekątnej to dokładnie wartości własne (to wynika też z dowodu tw. Jordana) 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 nie zależne; patrz rysunek). Stąd $\dim\ker F$ = liczba klatek Jordana z 0 na przekątnej.

- Podobnie, dla dowolnego skalara λ zachodzi $\dim\ker(F-\lambda)$ = liczba klatek Jordana z λ na przekątnej.

- Jeżeli oznaczymy przez $j_k(\lambda)$, $j_{\geqslant k}(\lambda)$ liczbę klatek rozmiaru λ, lub rozmiaru co najmniej λ z λ na przekątnej, to zachodzi równość

$$\dim\ker(F-\lambda)^k = \sum_{i=1}^{k} j_{\geqslant i}(\lambda).$$

  Wynika to z postaci potęg klatki Jordana: jeżeli $J_{d,\mu}$ jest klatką Jordana, to $J_{d,\mu} - \lambda \mathrm{I}_d = J_{d,\mu-\lambda}$, czyli dla $\mu \neq \lambda$ macierz odwracalna (a więc wszystkie potęgi mają trywialne jądro), zaś gdy $\mu = \lambda$, to wymiar jądra $(J_{d,\mu} - \lambda\mathrm{I}_d)^k = J_{d,0}^k$ jest równy $\min(k,d)$. Z tego wynikają równości:

$$j_{\geqslant k}(\lambda) = \dim\ker(F-\lambda)^k - \dim\ker(F-\lambda)^{k-1}$$

$$j_k(\lambda) = j_{\geqslant k}(\lambda) - j_{\geqslant k+1}(\lambda) = 2\dim\ker(F-\lambda)^k - \dim\ker(F-\lambda)^{k-1} - \dim\ker(F-\lambda)^{k+1}.$$

  (Szczegóły: ćwiczenie.)

- Wektory odpowiadające klatkom z λ na przekątnej tworzą bazę przestrzeni pierwiastkowej $V^\lambda := \{ v \in V \mid (\exists k)(F-\lambda)^k(v) = 0 \}$. (Wynika to z podobnych obserwacji jak powyżej.)

- Ogólnie, $V^\lambda = \ker(F-\lambda)^k$, gdzie k to krotność pierwiastka λ w $\chi_F(x)$, a w szczególności $V^\lambda = \ker(F-\lambda)^{\dim V}$ (ćwiczenie).

- Przestrzenie pierwiastkowe są liniowo niezależne (z poprzeniej uwagi), ponadto $V = \bigoplus_\lambda V^\lambda$. Dokładniej, jeżeli oznaczymy przez B_λ fragment bazy Jordanowskiej odpowiadający klatkom z λ na przekątnej, mamy

$$\operatorname{Lin} V = \operatorname{Lin}(\bigcup_\lambda B_\lambda) = \bigoplus_\lambda \underbrace{\operatorname{Lin}(B_\lambda)}_{\substack{\parallel \\ V^\lambda}}$$

Uwaga 7.33. Twierdzenie Jordana można nieco wzmocnić: jeżeli $F \in \mathrm{End}(V)$ nad ciałem K jest takie że $\chi_F(x)$ rozkłada się na czynniki liniowe nad K (równoważnie, ma w K dokładnie $\deg \chi_F$ pierwiastków z krotnościami), to istnieje baza V w której F ma postać Jordana.

Dowód tego wariantu jest bardzo podobny. Dzięki założeniu $\chi_F(x)$ ma pierwiastek i wciąż możemy założyć, że tym pierwiastkiem jest 0. Wystarczy uzasadnić, że $F\restriction_{\operatorname{im} F}$ spełnia założenia, tzn. jego wielomian charakterystyczny rozkłada się na czynniki liniowe.

Ale ponieważ $\operatorname{im} F = W$ jest F-niezmiennicze, z lematu z poprzedniego wykładu wiemy, że $\chi_F(x) = \chi_{F\restriction_W}(x) \cdot \chi_{\bar{F}}(x)$, gdzie $\bar{F}$ to przekształcenie $V/W \to V/W$ indukowane z F (jak w lemacie). Co więcej, $\bar{F}$ jest przekształceniem zerowym (bo dla każdego v zachodzi $\bar{F}(v + \operatorname{im} F) = F(v) + \operatorname{im} F = \operatorname{im} F$), więc $\chi_{F\restriction_W}(x) = \frac{\chi_F(x)}{(-x)^{\dim \ker F}}$, skąd widać, że rozkłada się na czynniki liniowe.

Przykład 7.34. Przekształcenie $\mathbf{Q}^2 \to \mathbf{Q}^2$ 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 $\mathbf{R}^2 \to \mathbf{R}^2$ zadane tą samą macierzą się diagonalizuje:

$$\begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix} = \begin{pmatrix} \sqrt{2} \\ 2 \end{pmatrix} = \sqrt{2}\begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ -\sqrt{2} \end{pmatrix} = \begin{pmatrix} -\sqrt{2} \\ 2 \end{pmatrix} = -\sqrt{2}\begin{pmatrix} 1 \\ -\sqrt{2} \end{pmatrix},$$

Przykład 7.35. Rozważmy przekształcenie $F_A \colon \mathbf{R}^3 \to \mathbf{R}^3$ zadane macierzą $A = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 1 & 1 \\ -1 & -1 & 3 \end{pmatrix}$.

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 $V_2 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| \begin{cases} x = 2x \\ -x + y + z = 2y \\ -x - y + 3z = 2z \end{cases} \right\} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| z = x + y \right\}$, 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 $b_2 = (A - 2\mathrm{I})\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix}$, zaś $b_1 \in V_2$ musi być lnz z b_2, na przykład $b_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$.

Przykład 7.36. 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 $\geqslant 2$ i $1 = 8 - 7$ rozmiaru $\geqslant 3$ i $0 = 8 - 8$ klatek rozmiaru $\geqslant 4$, czyli są 3 klatki 1×1 i po jednej 2×2 i 3×3. Stąd odpowiedni fragment macierzy F w postaci Jordana ma postać

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

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 $b_5 \in \ker(F - \lambda)^2 \setminus \ker(F - \lambda)$, liniowo niezależny z b_6 i b_7. 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.

Ogólny algorytm wyznaczania postaci Jordana przekształcenia F:

1. znajdujemy pierwiastki wielomianu charakterystycznego F,

2. dla każdego pierwiastka λ wyznaczamy jądra $(F - \lambda)^k$ dla kolejnych k, aż dostaniemy dwa razy ten sam wynik,

3. badając wymiary jąder określamy rozmiary klatek,

4. najpierw wybieramy dowolny wektor z $\ker(F - \lambda)^k \setminus \ker(F - \lambda)^{k-1}$, gdzie k jest maksymalnym rozmiarem klatki dla λ, a jego obrazy przez $(F -$

λ) dadzą kolejne wektory z części bazy odpowiadającej największej klatce,

5. dla największej pozostałej klatki rozmiaru l (*może* być $l = k!$) wybieramy wektor z $\ker(F-\lambda)^l \setminus \ker(F-\lambda)^{l-1}$, dbając o to, żeby był liniowo niezależny z dotychczas wybranymi wektorami z $\ker(F-\lambda)^{l-1}$, a następnie dobieramy jego obrazy przez $(F-\lambda), \dots (F-\lambda)^{l-1}$, co daje nam część bazy odpowiadającą tej klatce,

6. powtarzamy poprzedni krok, aż obsłużymy wszystkie klatki.

Kompleksyfikacja i rzeczywiste tw. Jordana

Definicja 7.37. 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 zadanym wzorem

$$(a+bi)\cdot(v,w) = (a+bi)\cdot(v+iw) = (av-bw)+i(bv+aw) = (av-bw, bv+aw).$$

Intuicyjnie: kompleksyfikacja V to „rozszrzenie skalarów” V z $\mathbf{R}$ do $\mathbf{C}$.

Przykład 7.38.

- Kompleksyfikację $(\mathbf{R}^n)_{\mathbf{C}}$ utożsamiamy z $\mathbf{C}^n$.
- $(\mathbf{R}_n[x])_{\mathbf{C}}$ utożsamiamy z $\mathbf{C}_n[x]$.
- Dla $V = C(\mathbf{R},\mathbf{R})$, przestrzeni funkcji ciągłych $\mathbf{R} \to \mathbf{R}$, kompleksyfikacja to $(C(\mathbf{R},\mathbf{R}))_{\mathbf{C}} = C(\mathbf{R},\mathbf{C})$ (przestrzeń funkcji ciągłych $\mathbf{R} \to \mathbf{C}$).

Uwaga 7.39. Kompleksyfikacja jest przypadkiem szczególnym ogólnej konstrukcji: iloczynu tensorowego; $V_{\mathbf{C}}$ jest iloczynem tensorowym $\mathbf{C}$ i V nad $\mathbf{R}$, ozn. $\mathbf{C} \otimes_{\mathbf{R}} V$. O tym może być więcej na konwersatorium. W podobny sposób można rozszerzać skalary z $\mathbf{Q}$ do $\mathbf{R}$, z $\mathbf{F}_2$ do $\mathbf{F}_4$ itp.

Uwaga 7.40. 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 $riv = irv$ i nad $\mathbf{R}$ zachodzi istotnie $V_{\mathbf{C}} = V \oplus_{\mathbf{R}} iV$ (to wynika wprost z definicji).

Fakt 7.41. *Niech V będzie przestrzenią liniową nad* $\mathbf{R}$*. Wtedy* $\dim_{\mathbf{R}} V = \dim_{\mathbf{C}} V_{\mathbf{C}}$*; dokładniej, jeżeli B jest bazą V nad* $\mathbf{R}$*, to B jest też bazą $V_{\mathbf{C}}$ nad* $\mathbf{C}$ *(uwzględniając utożsamienie $V \subseteq V_{\mathbf{C}}$).*

Dowód. Dla uproszczenia załóżmy $\dim V < \infty$, $B = (v_1, v_2, \ldots, v_n)$.

Rozpinanie jest łatwe:

$$\begin{aligned}\mathrm{Lin}_{\mathbf{C}}(v_1, v_2, \ldots, v_n) = \mathrm{Lin}_{\mathbf{C}}(v_1, iv_1, v_2, iv_2, \ldots, v_n, iv_n) \supseteq \\ \supseteq \mathrm{Lin}_{\mathbf{R}}(v_1, \ldots, v_n) + \mathrm{Lin}_{\mathbf{R}}(iv_1, \ldots, iv_n) = V + iV = V_{\mathbf{C}}\end{aligned}$$

Załóżmy że dla pewnych $a_1 + ib_1, a_2 + ib_2, \ldots, a_n + ib_n$ (gdzie $a_k, b_k \in \mathbf{R}$) zachodzi $\sum_{k=1}^n (a_k + ib_k)v_k = 0$.

Możemy sumę przepisać jako $\underbrace{\left(\sum_{k=1}^n a_k v_k\right)}_{\in V} + i\underbrace{\left(\sum_{k=1}^n b_k v_k\right)}_{\in iV}$. Stąd $\sum_k a_k v = \sum_k b_k v = 0$, więc $a_k = b_k = 0$. □

Definicja 7.42. Dla przekształcenia $\mathbf{R}$-liniowego $F\colon V \to W$, jego kompleksyfikacja to $\mathbf{C}$-liniowe $F_{\mathbf{C}}\colon V_{\mathbf{C}} \to W_{\mathbf{C}}$ zadane oczywistym wzorem $F_{\mathbf{C}}(v + iv') = F(v) + iF(v')$.

Uwaga 7.43. Jeżeli $F\colon V \to W$ jest $\mathbf{R}$-liniowe i B, C to bazy V, W, to zachodzi $m_C^B(F) = m_C^B(F_{\mathbf{C}})$ (odczytane nad $\mathbf{C}$; to wynika wprost z definicji m_C^B).

W szczególności $F_{\mathbf{C}}$ ma w pewnej (ale być może nie w każdej) bazie $V_{\mathbf{C}}$ macierz o współczynnikach rzeczywistych.

Przykład 7.44.

- Jeżeli V jest przestrzenią liniową nad $\mathbf{R}$, to mamy $\mathbf{R}$-liniowe $\bar{\cdot}\colon V_{\mathbf{C}} \to V_{\mathbf{C}}$ zadane wzorem $\overline{v + iv'} = v - iv'$. Nie jest ono $\mathbf{C}$-liniowe, bo $\overline{iv} = -i\bar{v}$ (Nie jest więc kompleksyfikacją liniowego przekształcenia.)
- Odwzorowanie $F\colon V_{\mathbf{C}} \to V_{\mathbf{C}}$, $v \mapsto iv$ również nie jest kompleksyfikacją żadnego rzeczywistego przekształcenia liniowego, np. dlatego że jeżeli B jest (rzeczywistą) bazą V, to macierz $m_B(F) = i\mathrm{I}$ nie jest rzeczywistą macierzą.
- Odwzorowanie $F\colon \mathbf{R}^2 \to \mathbf{R}^2$ 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 7.45. *Jeżeli* $F \in \mathrm{End}(V)$*, gdzie* V *jest rzeczywista, to* $\chi_F(x) = \chi_{F_\mathbf{C}}(x)$*, a więc ten drugi jest wielomianem o współczynnikach rzeczywistych.*

Fakt 7.46. *Jeżeli* $F \in \mathrm{End}(V)$ *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. Weźmy dowolne $v \in V$ i $\lambda \in \mathbf{C}$. Wtedy

$$F_\mathbf{C}(\bar{v}) = F_\mathbf{C}(v_1 - iv_2) = F(v_1) - iF(v_2) = \overline{F(v_1) + iF(v_2)} = \overline{F_\mathbf{C}(v)}$$

Stąd mamy $F_\mathbf{C}(v) = \lambda v \iff F_\mathbf{C}(\bar{v}) = \overline{F_\mathbf{C}(v)} = \bar{\lambda}\bar{v}$, co daje $(V_\mathbf{C})_{\bar{\lambda}} = \overline{(V_\mathbf{C})_\lambda}$.
Podobnym rachunkiem $\overline{(F-\lambda)^k(v)} = (F-\bar{\lambda})^k(\bar{v})$, do daje $(V_\mathbf{C})^{\bar{\lambda}} = \overline{(V_\mathbf{C})^\lambda}$. □

Fakt 7.47. *Jeżeli* $F : V \to W$ *(gdzie* V, W *są rzeczywiste), to* $(\ker F)_\mathbf{C} = \ker(F_\mathbf{C})$.

Dowód. Załóżmy że $v = v_1 + iv_2$. Wtedy $F_\mathbf{C}(v) = F(v_1) + iF(v_2)$, przy czym $F(v_1) \in W$, $iF(v_2) \in iW$. Stąd

$$v \in \ker F_\mathbf{C} \iff 0 = F_\mathbf{C}(v_1 + iv_2) = F(v_1) + iF(v_2) \iff \begin{cases} F(v_1) = 0 \\ F(v_2) = 0 \end{cases}$$

$$\iff \begin{cases} v_1 \in \ker F \\ v_2 \in \ker F \end{cases} \iff v = v_1 + iv_2 \in (\ker F)_\mathbf{C}. \quad \square$$

Wniosek 7.48. *Załóżmy że* $r \in \sigma(F_\mathbf{C}) \cap \mathbf{R}$*. Wtedy* $(V_\mathbf{C})_r = (V_r)_\mathbf{C}$ *i* $(V_\mathbf{C})^r = (V^r)_\mathbf{C}$ *(w szczególności* $r \in \sigma(F)$*).*

Dowód. Z poprzedniego faktu dla $G = F - r$ dostajemy $(V_\mathbf{C})_r = (V_r)_\mathbf{C}$, a dla $G = (F-r)^{\dim V}$ $(V_\mathbf{C})^r = (V^r)_\mathbf{C}$. □

Stwierdzenie 7.49. *Jeżeli* $F \in \mathrm{End}(V)$ *jest taka, że* $F_\mathbf{C}$ *się diagonalizuje, to w pewnej (rzeczywistej) bazie* B_0 *przestrzeni* V *mamy*

$$m_B(F) = \begin{pmatrix} r_1 &&&&&& \\ & r_2 &&&&& \\ && \ddots &&&& \\ &&& r_M &&& \\ &&&& \begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix} && \\ &&&&& \ddots & \\ &&&&&& \begin{pmatrix} a_N & b_N \\ -b_N & a_N \end{pmatrix} \end{pmatrix}$$

gdzie r_j są rzeczywistymi wartościami własnymi F, a a_j i b_j są odpowiednio częściami rzeczywistymi i urojonymi nierzeczywistych pierwiastków $\chi_F(x)$, wszystkie występujące z takimi krotnościami, jak odpowiadające im pierwiastki $\chi_F(x)$.

Uwaga 7.50. Jeżeli rozważymy prostą zespoloną **C** jako rzeczywistą przestrzeń liniową, to przekształcenie „pomnóż przez $\lambda = a + bi$" jest oczywiście **R**- (a nawet **C**-)liniowe i ma w bazie e_1, ie_1 macierz

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

Wnioski pozostawiam czytelnikom/słuchaczom. [Koniec uwagi]

Przykład 7.51. Jeżeli $F\colon \mathbf{R}^3 \to \mathbf{R}^3$ 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. Zauważmy że dla dowolnej $\lambda = (\alpha + i\beta) \in \mathbf{C}$ i $v = (v_1 + iv_2) \in (V_{\mathbf{C}})_\lambda$ mamy

$$F_{\mathbf{C}}(v \pm \bar{v}) = \lambda v \pm \bar{\lambda}\bar{v} = \lambda v \pm \overline{\lambda v} = (av_1 - bv_2) + i(bv_1 + av_2) \pm (av_1 - bv_2) \mp i(bv_1 + av_2)$$

$$\begin{aligned} F(v_1) = F_{\mathbf{C}}(v_1) = F_{\mathbf{C}}\left(\frac{v + \bar{v}}{2}\right) = av_1 - bv_2 \\ F(v_2) = F_{\mathbf{C}}(v_2) = F_{\mathbf{C}}\left(\frac{v - \bar{v}}{2i}\right) = bv_1 + av_2 \end{aligned} \qquad (*)$$

czyli

$$[F(v_1)]_{v_1,v_2} = \begin{pmatrix} a \\ -b \end{pmatrix} \qquad\qquad [F(v_2)]_{v_1,v_2} = \begin{pmatrix} b \\ a \end{pmatrix}$$

Ponadto łatwo zauważyć, że $\mathrm{Lin}_{\mathbf{C}}(v, \bar{v}) = \mathrm{Lin}_{\mathbf{C}}\left(\frac{v+\bar{v}}{2}, \frac{v-i\bar{v}}{2}\right)$.

Stąd łatwo wywnioskować, że jeżeli $(b_1,\ldots,b_n)$ jest bazą $(V_{\mathbf{C}})_\lambda$, to $B_\lambda^+ = \left(\frac{b_1+\bar{b}_1}{2}, \frac{b_1-\bar{b}_1}{2i}, \ldots, \frac{b_n+\bar{b}_n}{2}, \frac{b_n-\bar{b}_n}{2i}\right)$ jest bazą $(V_{\mathbf{C}})_\lambda + (V_{\mathbf{C}})_{\bar{\lambda}}$ (składającą się z rzeczywistych wektorów!) i mamy

$$m_{B_\lambda^+}(F_{\mathbf{C}}\restriction_{V_\lambda+V_{\bar{\lambda}}}) = m_{B_\lambda^+}(F\restriction_{\mathrm{Lin}_{\mathbf{R}}(B_\lambda^+)}) = \begin{pmatrix} \begin{pmatrix} a & b \\ -b & a \end{pmatrix} & & & \\ & \begin{pmatrix} a & b \\ -b & a \end{pmatrix} & & \\ & & \ddots & \\ & & & \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \end{pmatrix}$$

Załóżmy teraz że $\sigma(F_{\mathbf{C}}) = \{r_1, r_2, \ldots, r_k, \lambda_1, \bar{\lambda}_1, \ldots, \lambda_l, \bar{\lambda}_l\}$.

Wybierzmy bazę V w następujący sposób: dla $j = 1,\ldots,k$ bierzemy dowolną bazę B_{r_j} przestrzeni V_{r_j}, a następnie dla $j = 1,2,\ldots,l$ bierzemy $B_{\lambda_i}^+$ postaci jak wyżej. Wtedy $B = B_{r_1}\ldots B_{r_k} B_{\lambda_1}^+ \ldots B_{\lambda_l}^+$ jest ($\mathbf{C}$-, więc i $\mathbf{R}$-)liniowo niezależny (bo przestrzenie własne są liniowo niezależne) i mamy

$$\begin{aligned} \mathrm{Lin}_{\mathbf{C}}(B) &= \mathrm{Lin}_{\mathbf{C}}(B_{r_1}) + \ldots + \mathrm{Lin}_{\mathbf{C}}(B_{r_k}) + \mathrm{Lin}_{\mathbf{C}}(B_{\lambda_1}^+) + \ldots + \mathrm{Lin}_{\mathbf{C}}(B_{\lambda_l}^+) \\ &= (V_{\mathbf{C}})_{r_1} + \ldots + (V_{\mathbf{C}})_{r_k} + (V_{\mathbf{C}})_{\lambda_1} + (V_{\mathbf{C}})_{\bar{\lambda}_1} + \ldots + (V_{\mathbf{C}})_{\lambda_l} + (V_{\mathbf{C}})_{\bar{\lambda}_l} \\ &= V_{\mathbf{C}}. \end{aligned}$$

Z drugiej strony, ponieważ $B \subseteq V$, mamy $\mathrm{Lin}_{\mathbf{C}}(B) = \mathrm{Lin}_{\mathbf{R}}(B) + i\,\mathrm{Lin}_{\mathbf{R}}(B)$, więc stąd $\mathrm{Lin}_{\mathbf{R}}(B) = V$, czyli B jest bazą V. Łatwo zauważyć, że $m_B(F)$ jest zadanej postaci. □

Co gdy $F_{\mathbf{C}}$ się *nie* diagonalizuje? Dla rzeczywistych $r \in \sigma(F_{\mathbf{C}})$ możemy wybrać bazy Jordana dla V^r. Dla nierzeczywistych λ postępujemy podobnie jak powyżej: wybieramy bazę B_λ Jordana dla $(V_{\mathbf{C}})^\lambda$, zauważamy że jej sprzężenie $\overline{B_\lambda}$ jest bazą Jordana dla $(V_{\mathbf{C}})^{\bar{\lambda}}$ i rozumując podobnie jak wyżej zauważamy, że dla B_λ^+ zdefiniowanego jak wyżej mamy

$$m_{B_\lambda^+}(F_{\mathbf{C}}\restriction_{V^\lambda+V^{\bar{\lambda}}}) = m_{B_\lambda^+}(F\restriction_{\mathrm{Lin}_{\mathbf{R}}(B_\lambda^+)}) = \begin{pmatrix} J_{\lambda,d_1}^+ & & & \\ & J_{\lambda,d_2}^+ & & \\ & & \ddots & \\ & & & J_{\lambda,d_{n_\lambda}}^+ \end{pmatrix}$$

$$J^+_{\lambda,d} = \left.\begin{pmatrix} \begin{pmatrix} a & b \\ -b & a \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & & & \\ & \begin{pmatrix} a & b \\ -b & a \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & & \\ & & \ddots & \ddots & \\ & & & \begin{pmatrix} a & b \\ -b & a \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ & & & & \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \end{pmatrix}\right\} 2d \text{ wierszy}$$

gdzie a, b to część rzeczywista i urojona λ, a d_j to rozmiary klatek Jordana odpowiadających λ.

W ten sposób otrzymujemy:

Twierdzenie 7.52 (rzeczywiste Jordana)**.** *Jeżeli $F \in \mathrm{End}(V)$, gdzie V jest rzeczywista, to istnieje baza B taka że*

$$m_B(F) = \begin{pmatrix} J_{r_1,d_1} & & & & & & & \\ & J_{r_2,d_2} & & & & & & \\ & & \ddots & & & & & \\ & & & J_{r_k,d_k} & & & & \\ & & & & J^+_{\lambda_1,e_1} & & & \\ & & & & & J^+_{\lambda_2,e_2} & & \\ & & & & & & \ddots & \\ & & & & & & & J^+_{\lambda_l,e_n} \end{pmatrix}$$

Uwaga 7.53. Można podobnymi metodami udowodnić wersję tw. Jordana dla dowolnych ciał, z tym że w ogólności rozmiary „podklatek” mogą być dowolnie duże, a rozmiary „podklatek” w postaci Jordana F zależą od stopni nierozkładalnych, na które rozkłada się $\chi_F(x)$.

Wszystkie wersje twierdzenia Jordana są szczególnym przypadkiem tzw. zasadniczego twierdzenia o skończenie generowanych modułach nad pierścieniami ideałów głównych. (O którym można usłyszeć np. na wykładzie z algebry przemiennej, na $\geqslant 3$ roku... Można też zapytać o nie np. mnie albo Gala, ale raczej po zrozumieniu co znaczą zaklęcia typu „pierścień ideałów głównych”, najlepiej po wykładzie z Algebry 1 R.)

Przykład 7.54. Chcemy rozwiązać równanie różniczkowe (liniowe, jednorodne) postaci

$$f^{(n)} = a_{n-1}f^{(n)} + \ldots + a_1 f' + a_0 f,$$

gdzie a_k są pewnymi rzeczywistymi (lub zespolonymi!) współczynnikami.

Zauważmy że jest ono równoważne równaniu

$$\begin{pmatrix} f' \\ \vdots \\ f^{(n)} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 1 \\ a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} \end{pmatrix}}_{A} \begin{pmatrix} f \\ \vdots \\ f^{(n-1)} \end{pmatrix}.$$

W skrócie $v'(t) = Av(t)$.

Szukamy rozwiązań postaci $v(t) = \begin{pmatrix} f(t) \\ \vdots \\ f^{(n-1)}(t) \end{pmatrix} = e^{tB}v$ dla pewnej macierzy B i wektora v_0 gdzie $e^{tB} = \sum_{k=0}^{\infty} \frac{(tA)^k}{k!} = \sum_{k=1}^{\infty} \frac{t^k}{k!}A^k$ (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 $v(t) = e^{tA}v_0$.

Jak wyznaczać efektywnie rozwiązania tego równania? Na przykład wyznaczamy postać Jordana $A = PJP^{-1}$. Wtedy $e^{tA} = Pe^{tJ}P^{-1}$.

Na przykład dla $f'' = -4f' + 4f$ mamy $A = \begin{pmatrix} 0 & 1 \\ -4 & 4 \end{pmatrix}$, $A = PJP^{-1}$, gdzie $P = \begin{pmatrix} -2 & 1 \\ -4 & 0 \end{pmatrix}$ i $J = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$. $\frac{t^k J^k}{k!} = \begin{pmatrix} \frac{(2t)^k}{k!} & \frac{k \cdot t^k 2^{k-1}}{k!} \\ 0 & \frac{(2t)^k}{k!} \end{pmatrix}$, $e^{tJ} = \begin{pmatrix} e^{2t} & te^{2t} \\ 0 & e^{2t} \end{pmatrix}$, co pozwala wyznaczyć jawny wzór na rozwiązanie.

(Szczegóły: wykład równania różniczkowe 1/1 R)

Rozdział 8

Formy dwuliniowe i kwadratowe

Przykład 8.1. Rozważmy iloczyn skalarny na $\mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$, zadany wzorem $\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \sum_{k=1}^n x_i y_i$.

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 $\varphi\colon V \times V \to K$ 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 $\varphi\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right) = \sum_{k=1}^n x_i y_i$
- det na K^2 jest formą dwuliniową

- jeżeli ustalimy wektory $c_1, c_2, \ldots, c_{n-2} \in K^n$, to $\varphi(v,w) = \det(c_1, c_2, \ldots, c_{n-2}, v, w)$ jest dwuliniowe
- na $C([0,1])$ (przestrzeń funkcji ciągłych na kwadracie jednostkowym) mamy symetryczną formę dwuliniową $\varphi(f,g) = \int_0^1 e^t f(t)g(t)\,\mathrm{d}t$.
- 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 $A = \mathrm{I}$ 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 $\mathrm{Hom}(V, V^*)$, $\varphi \mapsto (v \mapsto \varphi(v, -))$.

Uwaga 8.5. Jeżeli φ jest formą dwuliniową na V o bazie $B = (b_1, \ldots, b_n)$, to dla $v = \sum_k \alpha_k b_k$, $w = \sum_j \beta_j b_j$ mamy

$$\varphi(v,w) = \varphi\Big(\sum_k \alpha_k b_k, \sum_j \beta_j b_j\Big) = \sum_k \sum_j \alpha_k \varphi(b_k, b_j)\beta_j$$

Definicja 8.6. Niech $B = (b_1, \ldots, b_n)$ 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:

$$\begin{pmatrix}\alpha_1\\ \vdots \\ \alpha_n\end{pmatrix}^\top \begin{pmatrix} \varphi(b_1,b_1) & \varphi(b_1,b_2) & \cdots & \varphi(b_1,b_n) \\ \varphi(b_2,b_1) & \varphi(b_2,b_2) & & \vdots \\ \vdots & & \ddots & \vdots \\ \varphi(b_n,b_1) & \cdots & \cdots & \varphi(b_n,b_n)\end{pmatrix} \begin{pmatrix}\beta_1\\ \vdots \\ \beta_n\end{pmatrix} =$$

$$= \begin{pmatrix}\alpha_1 & \cdots & \alpha_n\end{pmatrix} \begin{pmatrix}\sum_j \varphi(b_1,b_j)\beta_j \\ \vdots \\ \sum_j \varphi(b_n,b_j)\beta_j\end{pmatrix} = \sum_k \sum_j \alpha_k \varphi(b_k,b_j)\beta_j \quad \square$$

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. $\square$

Załóżmy teraz, że mamy bazy B, C przestrzeni V i znamy $m^{BB}(\varphi)$. Jak wyznaczyć $m^{CC}(\varphi)$?

Wniosek 8.12. $m^{CC}(\varphi) = m_B^C(\mathrm{id})^\top \cdot m^{BB}(\varphi) \cdot m_B^C(\mathrm{id})$

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 $[v]_B = m_B^C(\mathrm{id}) \cdot [v]_C$. Stąd łatwo wynika:

$$\begin{aligned}\varphi(v,w) &= [v]_B^\top m^{BB}(\varphi)[w]_B \\ &= (m_B^C(\mathrm{id}) \cdot [v]_C)^\top m^{BB}(\varphi)(m_B^C(\mathrm{id}) \cdot [w]_C) \\ &= [v]_C^\top \cdot m_B^C(\mathrm{id})^\top m^{BB}(\varphi) m_B^C(\mathrm{id}) \cdot [w]_C\end{aligned}$$

Ponieważ v, w były dowolne, to kończy dowód. □

Uwaga 8.13. Z powyższego wzoru widać, że w ogólności nie działa wzór $m^{CC}(\varphi) = (m_B^C(\mathrm{id}))^{-1} \cdot m^{BB}(\varphi) \cdot m_B^C(\mathrm{id})$, chyba że $m_B^C(\mathrm{id})^{-1} = m_B^C(\mathrm{id})^\top$.

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 $B = PAP^\top$.*

Iloczyny skalarne

Od teraz do odwołania pracujemy nad $\mathbf{R}$ (chyba że podam inaczej).

Przykład 8.16. Macierz A reprezentuje iloczyn skalarny na $\mathbf{R}^n$ (względem pewnej bazy B $\mathbf{R}^n$) wtedy i tylko wtedy gdy spełnia $A = P^\top \mathrm{I} P = P^\top P$, gdzie P jest odwracalna. Co więcej, $P = m_E^B(\mathrm{id})$ jest macierzą której kolumny to wektory z bazy B.

- Jeżeli P jest *ortogonalna*, to znaczy $P^\top = P^{-1}$, to $A = \mathrm{I}$, 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 dwuliniowa, symetryczna i dodatnio określona (bo te własności nie zależą od wyboru bazy). Dwuliniowość nam nic nie da, ale symetryczność i dodatnia określoność już tak: wynika z nich, że jeżeli A reprezentuje iloczyn skalarny, to musi być macierzą symetryczną, a ponadto dla każdego niezerowego wektora v musi spełniać $v^T A v > 0$.

Definicja 8.17. Mówimy że macierz symetryczna $A \in \mathrm{M}_{n\times n}(\mathbf{R})$ jest *dodatnio określona* jeżeli dla każdego $v \in \mathbf{R}^n$, $v \neq 0$ mamy

$$v^T A v > 0.$$

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 $\varphi(f, g) = \int_0^1 f(t)g(t)\,\mathrm{d}t$ jest iloczynem skalarnym.

Pokażemy następujące twierdzenie.

Twierdzenie 8.20. *Następujące warunki są równoważne dla $A \in \mathrm{M}_{n\times n}(\mathbf{R})$.*

1. *A reprezentuje standardowy iloczyn skalarny na $\mathbf{R}^n$ (względem pewnej bazy),*

2. *istnieje macierz odwracalna P taka że $A = P^\top P$,*

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)**.** *Załóżmy że φ jest dodatnio określoną, symetryczną formą dwuliniową na skończenie 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 $b'_k \in \mathrm{Lin}(b_1, \dots, b_k)$ i $\mathrm{Lin}(b_1, \dots, b_k) = \mathrm{Lin}(b'_1, \dots, b'_k)$.

Będziemy oznaczać $\varphi(v, w)$ w skrócie przez $[v, w]$.

Niech $b'_1 = \frac{b_1}{\sqrt{[b_1, b_2]}}$. Wtedy

$$[b'_1, b'_1] = [\frac{b_1}{\sqrt{\varphi(b_1, b_1)}}, \frac{b_1}{\sqrt{\varphi(b_1, b_1)}}] = \frac{[b_1, b_1]}{\sqrt{[b_1, b_1]}^2} = 1.$$

Chcemy wyznaczyć $b'_2 \in \mathrm{Lin}(b_1, b_2) = \mathrm{Lin}(b'_1, b_2)$, ortogonalny do b'_1. Rozważmy wektor postaci $v = \alpha_1 b'_1 + b_2$. Wtedy

$$[b'_1, v] = [b'_1, \alpha_1 b'_1 + b_2] = [b'_1, \alpha_1 b'_1] + [b'_1, b_2] = \alpha_1 + [b'_1, b_2].$$

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'

Załóżmy że $k < n$ i mamy już ortonormalny układ $b_1', \dots, b_k'$ wektorów z $\mathrm{Lin}(b_1, \dots, b_k)$. Weźmy $b_{k+1}'' = b_{k+1} - \sum_{j=1}^{k}[b_{k+1}, b_j']b_j'$. Wtedy dla $l \leqslant k$ 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 niezależność 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. □

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 $m^{BB}(\varphi) = \mathrm{I}$.*

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 ij-ta współrzedna $m^{BB}(\varphi)$ to

$$\varphi(b_i, b_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases},$$

więc $m^{BB}(\varphi) = \mathrm{I}$.

$\Leftarrow$: ćwiczenie. □

To daje nam brakującą implikację z poprzedniego Twierdzenia.

Wniosek 8.26. *Jeżeli $A \in \mathrm{M}_{n\times n}(\mathbf{R})$ jest symetryczna i dodatnio określona, to forma $\varphi_A(v,w) = v^\top Aw$ jest równoważna ze standardowym iloczymem 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 i wierszy.

Twierdzenie 8.28 (Kryterium Sylvestera)**.** *Załóżmy że $A \in \mathrm{M}_{n\times n}(\mathbf{R})$. Niech A_k oznacza k-ty wiodący minor główny A, tj. $A_k = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$, gdzie $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$.*

Wtedy A jest dodatnio określona wtedy i tylko wtedy gdy $\det(A_k) > 0$ (dla wszystkich $k = 1, 2, \ldots, n$).

Twierdzenie to udowodnimy później.

Ortogonalność

Przykład 8.29. W szczególnej teorii względności ważną rolę odgrywa tzw. forma Lorentza:

$$\varphi((x,t),(y,s)) = x_1 y_1 + x_2 y_2 + x_3 y_3 - c^2 ts,$$

gdzie c to stała reprezentująca prędkość światła. Normalizując, możemy wybrać bazę B czasoprzestrzeni $\mathbf{R}^4 = \mathbf{R}^{3+1}$ taką że

$$m_B(\varphi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

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 żę 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 $\varphi(f,g) = \int_0^1 f(t)g(t)\,\mathrm{d}t$ składa się z funkcji spełniających $\int_0^1 f(t)\,\mathrm{d}t = 0$.
- 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 $A, B \subseteq V$:

- $A^{\perp} \leqslant V$,
- $A^{\perp} = (\mathrm{Lin} A)^{\perp}$,
- $(A \cup B)^{\perp} = A \perp \cap B^{\perp}$

Dowód. Ćwiczenie □

Fakt 8.36. *Niech v_0 będzie dowolnym nieizotropowym wektorem. Wtedy dla $W = \mathrm{Lin}(v)$ 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:

$$\varphi(v-w, v_0) = \varphi(v, v_0) - \varphi(\frac{\varphi(v, v_0)}{\varphi(v_0, v_0)} v_0, v_0) = \varphi(v, v_0) - \underbrace{\frac{\varphi(v, v_0)}{\varphi(v_0, v_0)} \varphi(v_0, v_0)}_{\varphi(v,v_0)} = 0,$$

więc

$$v = \underbrace{w}_{\in W} + \underbrace{v - w}_{\in W^\perp},$$

co daje tezę. $\square$

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 $B = (b_1, \ldots, b_n)$ 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 $\dim V = 1$ 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

$$\begin{aligned}
\varphi(b_1, b_1) &= \varphi\Big(\frac{v_0}{\sqrt{|\varphi(v_0, v_0)|}}, \frac{v_0}{\sqrt{|\varphi(v_0, v_0)|}}\Big) \\
&= \frac{\varphi(v_0, v_0)}{\sqrt{|\varphi(v_0, v_0)|}^2} \\
&= \frac{\varphi(v_0, v_0)}{|\varphi(v_0, v_0)|} \\
&= \pm 1
\end{aligned}$$

Rozważmy obcięcie φ do $W = b_1^\perp$.

Ponieważ $V = \mathrm{Lin}(b_1) \oplus b_1^\perp$, wiemy że $\dim W = \dim V - 1$. Z założenia indukcyjnego możemy więc znaleźć $b_2, b_3, \ldots, b_n$ — 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 $V = \mathrm{Lin}(b_1) \oplus W$, $B = (b_1, \ldots, b_n)$ 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 $\varphi(v) = 2$, więc bierzemy $v_1 = \frac{v}{\sqrt{2}}$.

Wśród wektorów v takich że $\varphi(v_1, v)$ 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 $v_2 = \frac{v_2}{\sqrt{2}}$.

Wśród v takich że $\varphi(v_1, v) = \varphi(v_2, v) = 0$ szukamy v takiego żę $\varphi(v,v) \neq 0$. 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

$$m^{BB}(\varphi) = m_E^B(\mathrm{id})^\top m^{EE}(\varphi) m_E^B(\mathrm{id})$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 2 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^\top \begin{pmatrix} 0 & 1 & 2 & 0 \\ 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 2 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

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:

$$m^{BB}(\varphi) = \begin{pmatrix} \mathrm{I}_p & & \\ & -\mathrm{I}_q & \\ & & 0_r \end{pmatrix}.$$

pokażemy następujące twierdzenie:

Twierdzenie 8.41 (prawo bezwładności Sylvestera)**.** *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 $B = (\frac{x}{\sqrt{2}}, \frac{-x+2}{\sqrt{2}}, x^2+2x-3, x^3-3x+2)$ to $m^{BB}(\varphi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. (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 $F \in \mathrm{End}(V)$ 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: załóżmy 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 $A[v]_B = 0$. □

Uwaga 8.46. Załóżmy że φ jest formą dwuliniową na przestrzeni V i $W \leqslant V$. Wtedy $\varphi\restriction_W$, 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 $W \leqslant V$ ($\dim V < \infty$). Załóżmy że obcięcie $\varphi\restriction_W$ jest niezdegenerowane (jako symetryczna forma dwuliniowa na W). Wtedy $V = W \oplus W^\perp$.*

Dowód. Niezdegenerowanie $\varphi\restriction_W$ łatwo implikuje $W^\perp \cap W = \{0\}$, wystarczy więc pokazać $W + W^\perp = V$.

Indukcja względem $\dim W$. Dla $\dim W = 1$ — było. Niech $w_1, \ldots, w_k$ będzie bazą W. Oznaczmy $W' = \mathrm{Lin}(w_1, \ldots, w_k)$. Z założenia indukcyjnego każdy wektor $v \in V$ przedstawia się w postaci $v = v_1 + v_2$, gdzie $v_1 \in W'$, $v_2 \in (W')^\perp$.

W szczególności $w_k = w_0 + w'_k$, gdzie $w_0 \in (W')^\perp, w'_k \in W'$.

Wtedy dla każdego v mamy $v_3 = v_2 - \varphi(v_2, w_0)w_0 \in (W')^\perp$ i podobnie jak w dowodzie Faktu 8.36 pokazujemy że $v_3 \in w_0^\perp$, więc $v_3 \in (W')^\perp \cap w_0^\perp = W^\perp$.

Stąd:

$$v = \underbrace{v_1 + \varphi(v_2, w_0)w_0}_{\in W' + \mathrm{Lin}(w_0) = W} + \underbrace{v_3}_{\in W^\perp}$$ □

dowód prawa bezwładności Sylvestera. Załóżmy że $B = (b_1, \ldots, b_n)$ oraz $B = (b'_1, \ldots, b'_n)$ 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 $C = b_1, b_2, \ldots, b_p, b'_{p'+1}, b'_{p'+2} \ldots, b'_n$ jest liniowo niezależny. To wystarczy, bo z tego będzie wynikało że $p \leqslant p'$, a więc (z symetrii) również $p' \leqslant p$, czyli $p = p'$

Załóżmy że $\alpha_1, \ldots, \alpha_p, \beta_{p'+1}, \ldots, \beta_n$ 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

$$\varphi(v, v) = \sum_{k=1}^{p} \alpha_k^2 \varphi(b_k, b_k) = \alpha_1^2 + \alpha_2^2 + \ldots + \alpha_p^2.$$

Z prawej zaś

$$\varphi(v,v)=\sum_{j=p'+1}^{n}\beta_j\underbrace{\varphi(b'_j,b'_j)}_{\begin{cases}-1 & j\leqslant p'+q'\\ 0 & j>p'+q'\end{cases}}=-\beta^2_{p'+1}-\beta^2_{p'+2}-\ldots-\beta^2_{p'+q'}.$$

Ponieważ α_k i β_j są liczbami rzeczywistymi, równość zachodzi tylko wtedy gdy wszystkie α_k i wszystkie β_j dla $j\leqslant p'+q'$ 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 Sylvestera. 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 $m^{BB}(\varphi)=\mathrm{I}$, więc istnieje odwracalna Q taka że $QAQ^\top=\mathrm{I}$, czyli $A=Q^{-1}Q^{-\top}$. Stąd $\det A=(\det Q)^{-2}$, więc $\det A>0$.

k-ty minor główny A to macierz obcięcia $\varphi\restriction_{V_k}$ dla $V_k=\mathrm{Lin}(e_1,e_2,\ldots,e_k)$. Ponieważ $\varphi\restriction_{V_k}$ jest dodatnio określona (jako obcięcie dodatnio określonej formy), odpowiedni wyznacznik jest dodatni jak wyżej.

Załóżmy 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.

Załóżmy że teza zachodzi dla macierzy mniejszych niż n. To implikuje że obcięcie $\varphi=\varphi_A$ do V_{n-1} jest dodatnio określone i istnieje odwracalna $Q'\in\mathrm{M}_{(n-1)\times(n-1)}(\mathbf{R})$ taka że $Q'A_{n-1}(Q')^\top=\mathrm{I}_{n-1}$. Biorąc macierz blokową $Q=\begin{pmatrix}Q'&0\\0&1\end{pmatrix}$ dostajemy

$$QAQ^\top=\begin{pmatrix}Q'&0\\0&1\end{pmatrix}\begin{pmatrix}A_{n-1}&*\\ *&*\end{pmatrix}\begin{pmatrix}(Q')^\top&0\\0&1\end{pmatrix}=\begin{pmatrix}\mathrm{I}_{n-1}&v\\v^\top&c\end{pmatrix}=A'$$

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 $\begin{pmatrix} \mathrm{I} & * \\ 0 & c' \end{pmatrix}$. Dokładniej, weźmy $E = \begin{pmatrix} \mathrm{I}_{n-1} & 0 \\ -v^\top & 1 \end{pmatrix}$.

Wtedy

$$EA' = \begin{pmatrix} \mathrm{I}_{n-1} & 0 \\ -v^\top & 1 \end{pmatrix} \begin{pmatrix} \mathrm{I}_{n-1} & v \\ v^\top & c \end{pmatrix} = \begin{pmatrix} \mathrm{I}_{n-1} & v \\ 0 & c' \end{pmatrix}:$$

dla wszystkich wierszy prócz ostatniego jest to zupełnie oczywiste. Wyliczmy i-tą współrzędną w ostatnim wierszu, gdzie $i < n$.

$$\begin{pmatrix} -v^\top & 1 \end{pmatrix} \underbrace{\begin{pmatrix} e_i \\ v_i \end{pmatrix}}_{e_i \in \mathbf{R}^{n-1}} = -v_i + v_i = 0.$$

Podobnie

$$EA'E^\top = \begin{pmatrix} \mathrm{I}_{n-1} & v \\ 0 & c' \end{pmatrix} \begin{pmatrix} \mathrm{I}_{n-1} & -v \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \mathrm{I}_{n-1} & 0 \\ 0 & c'' \end{pmatrix}$$

Łącznie dla $P = EQ$ mamy

$$PAP^\top = EQAQ^\top E^\top = \begin{pmatrix} \mathrm{I}_{n-1} & 0 \\ 0 & c'' \end{pmatrix},$$

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. $\square$

Formy kwadratowe

Teraz K jest dowolnym ciałem.

Definicja 8.48. Funkcję $Q\colon V \to K$ 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 $\operatorname{char} K \neq 2$ 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 $\operatorname{char} K \neq 2$, 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 $B = (b_1, \ldots, b_n)$ łatwo wyznaczyć wzór $\tilde{Q}$: jeżeli

$$Q(v) = \sum_{i \leqslant j} \alpha_{ij} v_i v_j,$$

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:

$$\alpha_{ij}(v_i+w_i)(v_j+w_j)-\alpha_{ij}v_iv_j-\alpha_{ij}w_iw_j = \alpha_{ij}(v_iv_j+v_iw_j+w_iv_j+w_iw_j-v_iv_j-w_iw_j) = \alpha_{ij}(v_iw_j+v_jw_i)$$

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* $\operatorname{char} K \neq 2$*, to* $Q \mapsto \tilde{Q}$*,* $\varphi \mapsto Q_\varphi$ *zadają wzajemnie odwrotne bijekcje miedzy 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. ($\operatorname{char} K \neq 2$)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 $v \neq 0$ (równoważnie, $m^{BB}(Q)$ jest dodatnio określona).

Uwaga 8.57. Jeżeli $Q \colon V \to K$ 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 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 z wartościami* $-1, 0, 1$ *na przekątnej. 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 Sylvestera.)

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 wcześniejszego 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

$$\begin{aligned} Q(v) &= 3(v_x^2 + \frac{2}{3}v_x v_y) + 4v_x v_z \\ &= 3\underbrace{(v_x + \frac{v_y}{3})}_{v_x'}{}^2 - \frac{1}{3}v_y^2 + 4 \underbrace{v_x}_{\substack{\| \\ v_x' - \frac{v_y}{3}}} v_z \\ &= 3(v_x')^2 - \frac{1}{3}v_y^2 + 4v_x' v_z - \frac{4}{3}v_y v_z \\ &= 3((v_x')^2 + \frac{4}{3}v_x' v_z) - \frac{1}{3}(v_y^2 + 4v_y v_z) \\ &= 3v_x'' - \frac{4}{3}v_z^2 - \frac{1}{3}(v_y')^2 + \frac{4}{3}v_z^2 = 3(v_x'')^2 - \frac{1}{3}(v_y')^2, \end{aligned}$$

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 $m^E_B(\mathrm{id}) = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ \frac{2}{3} & 2 & 1 \end{pmatrix}$ mamy $m^{BB}(Q) = \begin{pmatrix} 3 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & 0 \end{pmatrix}$, czyli

$$\begin{pmatrix} 3 & 1 & 2 \\ 1 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ \frac{2}{3} & 2 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ \frac{2}{3} & 2 & 1 \end{pmatrix}^{\top}$$

rówoważnie:

$$\begin{pmatrix} 3 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ \frac{2}{3} & 2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 3 & 1 & 2 \\ 1 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ \frac{2}{3} & 2 & 1 \end{pmatrix}^{-\top}$$

2. Co gdy nie ma kwadratów? Weźmy $Q(v) = 2v_xv_y + 4v_xv_z$. Podstawmy $v'_x = \frac{v_x+v_y}{2}$, $v'_y = \frac{v_x-v_y}{2}$. Wtedy $v_xv_y = (v'_x)^2 - (v'_y)^2$ i $v_x = v'_x + v'_y$, stąd

$$\begin{aligned} Q(v) &= 2v_xv_y + 4v_xv_z \\ &= 2(v'_x)^2 - 2(v'_y)^2 + 4v'_xv_z + 4v'_yv_z \\ &= 2((v'_x)^2 + 2v'_xv_z) - 2((v'_y)^2 - 2v'_yv_z) \\ &= 2(v''_x)^2 - v_z^2 - 2(v''_y)^2 + v_z^2 = 2(v''_x)^2 - 2(v''_y)^2, \end{aligned}$$

czyli

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

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.)

Rozdział 9

Przestrzenie euklidesowe i unitarne, twierdzenie spektralne i rozkład singularny

Przestrzenie euklidesowe

Uwaga 9.1. Forma dwuliniowa na **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 $\varphi(v,v) = [v]_B^\top [v]_B > 0$ dla $v \neq 0$. □

Definicja 9.2. *Przestrzeń euklidesowa* to skończenie wymiarowa przestrzeń liniowa V nad **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 każdych wektorów $v_1, v_2 \in V$ mamy

$$\langle v_1, v_2 \rangle = [v_1]^\top m^{BB}(\langle -,- \rangle)[v_2] = [v_1]_B^\top \mathrm{I}[v_2]_B = [v_1]_B^\top [v_2]_B,$$

czyli $[-]_B$ jest izomorfizmem $V \to \mathbf{R}^n$ 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 $W = \mathrm{Lin}(u_1, u_2)$ z $\mathbf{R}^2$ i tak zmierzyć kąt pomiędzy nimi.

Dokładniej, jeżeli w_1, w_2 jest bazą ortonormalną $\mathrm{Lin}(u_2, u_2)$, to proste $\mathrm{Lin}(w_1)$, $\mathrm{Lin}(w_2)$ 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 $\mathrm{Lin}(u_1, u_2)$)

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:* $\langle v, w \rangle \leqslant |v||w|$,
- *nierówność trójkąta:* $|v + w| \leqslant |v| + |w|$.

Zauważmy że jeżeli $W \leqslant V$ jest dowolną podprzestrzenią, to, ponieważ iloczyn skalarny jest niezdegenerowany (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 $W \leqslant V$ jest podprzestrzenią, to *rzut ortogonalny* na W to funkcja $\pi = \pi_W =: V \to W$, $\pi(v) = w$, (w zdefiniowane jak wyżej).

Stwierdzenie 9.11. *Załóżmy że $b_1, b_2, \ldots, b_k$ jest bazą ON $W \leqslant V$. 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 $b_1, b_2, \ldots, b_k$. 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

$$w - v \in b_1^{\perp} \cap b_2^{\perp} \cap \ldots \cap b_k^{\perp} = \mathrm{Lin}(b_1, \ldots, b_k)^{\perp} = W^{\perp}.$$

Z drugiej strony $\bar{w} \in W$, więc istotnie $\bar{w} = \pi_W(v)$. $\square$

Biorąc $W = V$ dostajemy:

Wniosek 9.13. *Jeżeli $b_1, \ldots, b_n$ 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 $v \in \mathbf{R}^3$ na płaszczyznę $W = \mathrm{Lin}\left(\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}1\\-1\\1\end{pmatrix}\right)$.

Najpierw wyznaczamy bazę ON W, np. przez proces Grama-Schmidta:
$\left|\begin{pmatrix}1\\2\\3\end{pmatrix}\right| = \sqrt{1+4+9} = \sqrt{14}$, więc $b_1 = \frac{1}{\sqrt{14}}\begin{pmatrix}1\\2\\3\end{pmatrix}$.

$$\begin{aligned}
b_2' &= \begin{pmatrix}1\\-1\\1\end{pmatrix} - \left\langle b_1, \begin{pmatrix}1\\-1\\1\end{pmatrix}\right\rangle b_1\\
&= \begin{pmatrix}1\\-1\\1\end{pmatrix} - \underbrace{\frac{1-2+3}{14}}_{\frac{1}{7}}\begin{pmatrix}1\\2\\3\end{pmatrix}\\
&= \frac{1}{7}\begin{pmatrix}6\\-9\\4\end{pmatrix},
\end{aligned}$$

więc $|b_2'| = \frac{\sqrt{36+81+16}}{7} = \frac{\sqrt{133}}{7}$, $b_2 = \frac{1}{\sqrt{133}}\begin{pmatrix}6\\-9\\4\end{pmatrix}$.

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 $b_1, \ldots, b_n$ jest bazą ortonormalną V, to dla każdego $v \in V$ zachodzi*

$$|v|^2 = \sum_{j=1}^{n} \langle v, b_j \rangle^2$$

Dowód.

$$|v|^2 = \langle v, v \rangle = \langle \sum_{j=1}^{n} \langle v, b_j \rangle b_j, \sum_{k=1}^{n} \langle v, b_k \rangle b_k \rangle = \sum_{j,k} \langle v, b_j \rangle \langle v, b_k \rangle \underbrace{\langle b_j, b_k \rangle}_{\begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases}} = \sum_{k} \langle v, b_j \rangle^2.$$

□

Wniosek 9.16 (nierówność Bessela)**.** *Jeżeli $b_1, b_2, \ldots, b_k$ jest układem ON (niekoniecznie bazą), to dla każdego $v \in V$ zachodzi nierówność:*

$$|v|^2 \geqslant \sum_{j=1}^{k} \langle v, b_j \rangle^2$$

Ćwiczenie 9.17. *Każdy układ ON w prestrzeni euklidesowej rozszerza się do bazy ON.*

Wniosek 9.18. *Dla każdego $W \leqslant V$ i $v \in V$ zachodzi $|\pi_W(v)| \leqslant |v|$ (równość zachodzi tylko gdy $v \in W$).*

Dowód. Ćwiczenie □

Wniosek 9.19. *Jeżeli $W \leqslant V$, i $v \in V$, to punkt $\pi_W(v)$ jest najbliższy v spośród punktów w W.*

Dowód. Ćwiczenie. □

Formy hermitowskie

Od teraz pracujemy nad $K = \mathbf{C}$.

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 $v = i \in \mathbf{C}^1$ 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):

$$\sqrt{a_1^2 + b_1^2 + a_2^2 \ldots + a_n^2 + b_n^2} = \sqrt{\bar{z}_1 z_1 + \bar{z}_2 z_2 + \ldots + \bar{z}_n z_n}.$$

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 nastepujące własności:

- liniowość w drugiej zmiennej:

$$\langle v, w_1 + w_2 \rangle = \langle v, w_1 \rangle + \langle v, w_2 \rangle \qquad \langle v, \alpha w \rangle = \alpha \langle v, w \rangle.$$

- antyliniowość w pierwszej zmiennej:

$$\langle v_1 + v_2, w\rangle = \langle v_1, w\rangle + \langle v_2, w\rangle \qquad \langle \alpha v, w\rangle = \bar{\alpha}\langle v, w\rangle.$$

- 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.

Definicja 9.20. *Forma hermitowska* na V nad $\mathbf{C}$ to funkcja $\varphi\colon V \times V \to \mathbf{C}$ która jest półtoraliniowa (liniowa w drugim, a antyliniowa w pierwszym argumencie) i hermitowsko symetryczna.

Uwaga 9.21. φ jest hermitowska wtedy i tylko wtedy gdy jest hermitowsko symetryczna i liniowa w drugim argumencie:

$$\varphi(\alpha v, w) = \overline{\varphi(w, \alpha v)} = \bar{\alpha}\overline{\varphi(w, v)} = \bar{\alpha}\varphi(v, w).$$

Formy hermitowskie nad $\mathbf{C}$ odpowiadają formom symetrycznym nad $\mathbf{R}$ (ale rozważa się też prawdziwe formy symetryczne na zespolonych przestrzeniach liniowych!; rozważa się też niekoniecznie hermitowskie formy półtoraliniowe, ale nie będziemy się tym zajmować).

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:

$$\begin{aligned}
&\varphi_{\mathbf{C}}(v_1 + iv_2, \underbrace{(a + bi)(w_1 + iw_2)}_{aw_1 - bw_2 + i(bw_1 + aw_2}) = \\
&= \varphi(v_1, aw_1 - bw_2) + i\varphi(v_1, bw_1 + aw_2) - i\varphi(v_2, aw_1 - bw_2) + \varphi(v_2, bw_1 + aw_2) = \\
&= \varphi(v_1, w_1)(a + ib) + \varphi(v_1, w_2)(\underbrace{-b + ai}_{i(a+bi)}) + \varphi(v_2, w_1)(\underbrace{-ia + b}_{-i(a+bi)}) + \varphi(v_2, w_2)(ib + a) = \\
&\qquad = (a + bi)\varphi_{\mathbf{C}}(v_1 + iv_2, w_1 + iw_2)
\end{aligned}$$

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
$$\varphi(iv,w) = i\varphi(v,w) \neq -i\varphi(v,w).$$
- $\varphi(z,w) = i\bar{z}w$ nie jest formą hermitowską na $\mathbf{C}$, bo $\varphi(1,1) = i \neq \overline{\varphi(1,1)}$.
- $\varphi(x_1 + ix_2, y_2 + iy_2) = x_1 y_1$ nie jest formą hermitowską na $\mathbf{C}$, bo nie jest półtoraliniowa: $\varphi(1,i) = 0 \neq i = i\varphi(1,1)$.
- $\varphi\left(\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}\right) = i\bar{z}_2 w_1 - i\bar{z}_1 w_2$ jest formą hermitowską na $\mathbf{C}^2$: półtoraliniowość jest łatwa, hermitowska symetryczność:
$$\overline{i\bar{z}_2 w_1 - i\bar{z}_1 w_2} = -iz_2\bar{w}_1 + iz_1\bar{w}_2 = i\bar{w}_2, z_1 - i\bar{w}_1 z_2.$$

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 $B = (b_1, \ldots, b_n)$ jest bazą V, to *macierz φ względem B* definiujemy analogicznie, jak dla form dwuliniowych: 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 $m^{BB}(\varphi) = m^{BB}(\varphi_{\mathbf{C}})$ (przypomnienie: B jest bazą $V_{\mathbf{C}}$!).

Przykład 9.27. Forma hermitowska $\varphi\left(\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}\right) = i\bar{z}_2 w_1 - i\bar{z}_1 w_2$ na $\mathbf{C}^2$ 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 dwuliniowych.)

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. □

Definicja 9.30. Jeżeli $A \in \mathrm{M}_{n\times n}(\mathbf{C})$, 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), ale tego raczej nie będziemy poruszać na tym wykładzie, przynajmniej na razie.

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:

$$\begin{pmatrix} r_1 & a_{12} & a_{13} \cdots\cdots & a_{1n} \\ \bar{a}_{12} & r_2 & a_{23} & \\ \bar{a}_{13} & \bar{a}_{23} & r_3 & \\ \vdots & & & \ddots \\ \bar{a}_{1n} & & & r_n \end{pmatrix}$$

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 samosprzężona wtedy i tylko wtedy gdy jest symetryczna.

Jeżeli B, C są różnymi bazami, to ponieważ mamy $[v]_C = m^B_C(\mathrm{id})[v]_B$, podobnie jak w przypadku form symetrycznych, dostajemy wzór na zmianę bazy.

Fakt 9.36. $m^{CC}(\varphi) = m^C_B(\mathrm{id})^* m^{BB}(\varphi) m^C_B(\mathrm{id})$.

Dowód.

$$\begin{aligned}
\varphi(v,w) &= [v]^*_B m^{BB}(\varphi)[w]_B \\
&= (m^C_B(\mathrm{id})[v]_C)^* m^{BB}(\varphi)(m^C_B(\mathrm{id})[w]_C) \\
&= [v]^*_C m^C_B(\mathrm{id})^* m^{BB}(\varphi) m^C_B(\mathrm{id})[w]_C. \qquad \square
\end{aligned}$$

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

$$m^{BB}(\varphi) = m^B_E(\mathrm{id})^* m^{EE}(\varphi) m^B_E(\mathrm{id}) = \begin{pmatrix} -i & 0 \\ 1 & -i \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} i & 1 \\ 0 & i \end{pmatrix} =$$
$$= \begin{pmatrix} -i & 0 \\ 1 & -i \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & i \end{pmatrix} = \begin{pmatrix} 0 & -i \\ i & 2 \end{pmatrix}.$$

Faktycznie, na przykład $\varphi(\begin{pmatrix} 1 \\ i \end{pmatrix}, \begin{pmatrix} 1 \\ i \end{pmatrix}) = i \cdot (-i) \cdot 1 - i \cdot 1 \cdot i = 1 + 1 = 2$.

Wniosek 9.38. *Hermitowskie* $A_1, A_2 \in \mathrm{M}_{n\times n}(\mathbf{C})$ *zadają tę samą formę hermitowską (może w różnych bazach) wtedy i tylko wtedy gdy istnieje odwracalna* $P \in \mathrm{M}_{n\times n}(\mathbf{C})$ *taka że*

$$A_1 = P^* A_2 P.$$

Definicja 9.39. Mówimy że $A \in M_{n\times n}(\mathbf{C})$ jest *unitarna* jeżeli $AA^* = \mathrm{I}$ (czyli A jest odwracalna i $A^{-1} = A^*$).

Jeżeli $A \in M_{n\times n}(\mathbf{R})$ jest unitarna (równoważnie: spełnia $AA^\top = \mathrm{I}$), to mówimy że jest *ortogonalna*.

Uwaga 9.40. A jest unitarna/ortogonalna wtedy i tylko wtedy gdy $A^*/A^\top$ 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 następujacy wniosek:

Wniosek 9.42. *Jeżeli B jest bazą $\mathbf{C}^n$, a $\langle -, - \rangle$ jest standardowym iloczynem skalarnym, to zachodzi równoważność:*

$$m^{BB}(\langle -, - \rangle) = \mathrm{I} \iff m_E^B(\mathrm{id}) \text{ jest unitarna}.$$

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). $\square$

Stwierdzenie 9.47. *Jeżeli φ jest formą hermitowską i jej obcięcie do $W \leqslant V$ jest niezdegenerowane, to $V = W \oplus W^{\perp}$.*

Dowód. Ćwiczenie (dowód jak dla form symetrycznych). $\square$

Stwierdzenie 9.48 (wzór polaryzacyjny dla form hermitowskich)**.** *Załóżmy 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

$$\begin{aligned}
\varphi(v+w, v+w) &= \varphi(v+w, v) + \varphi(v+w, w) \\
&= \varphi(v,v) + \varphi(w,w) + \varphi(v,w) + \underbrace{\varphi(w,v)}_{\overline{\varphi(v,w)}} \\
&= \varphi(v,v) + \varphi(w,w) + 2\operatorname{Re}\varphi(v,w)
\end{aligned}$$

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.

Rozdział 10

Przestrzenie unitarne i twierdzenie spektralne

Definicja 10.1. *Iloczyn skalarny* na zespolonej przestrzeni liniowej to dodatnio określona forma hermitowska.

Przestrzeń unitarna to skończenie wymiarowa przestrzeń liniowa nad $\mathbf{C}$ wyposażona w iloczyn skalarny.

Uwaga 10.2. Dla przestrzeni unitarnych zachodzą podobne fakty jak dla przestrzeni euklidesowych, z analogicznymi dowodami. Na przykład:

- rzut ortogonalny na $W = \mathrm{Lin}(\overbrace{b_1, \ldots, b_k}^{\text{układ ON}})$ 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 Parsevala

  $$\langle v, v\rangle = \sum_{j=1}^{n} |\langle b_j, v\rangle|^2$$

  (uwaga na moduły po prawej stronie!)

- zachodzi nierówność Bessela (też z modułami)

- zachodzą nierówność trójkąta i nierówność Cauchy'ego-Schwarza (też z modułem)

Uwaga 10.3. Kompleksyfikacja $V_{\mathbf{C}}$ przestrzeni euklidesowej V zawsze jest przestrzenią unitarną (ze skompleksyfikowanym iloczynem skalarnym) i bazy ON V są bazami ON $V_{\mathbf{C}}$.

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.)

Wszystkie fakty mają zazwyczaj zupełnie analogiczne dowody w przypadku rzeczywistym euklidesowym i zespolonym unitarnym, większość będzie miała dowody pisane pod ten drugi.

Co więcej, będziemy pracowali w zasadzie wyłącznie z bazami ortonormalnymi.

Ćwiczenie 10.4. *Każdy ON układ wektorów w V rozszerza się do bazy ON.*

Uwaga 10.5. Baza B jest ortonormalna wtedy i tylko wtedy gdy $m^{BB}(\langle -, - \rangle) = \mathrm{I}$.

Stąd (i wzoru na zmianę bazy) łatwo wynika, że jeżeli B jest bazą ortonormalną, to baza C jest ON $\iff$ macierz $m^B_C(\mathrm{id})$ jest unitarna/ortogonalna.

Z drugiej strony jeżeli U jest macierzą unitarną/ortogonalną, to jest odwracalna, więc istnieje taka baza C że $m^B_C(\mathrm{id}) = U$ i z powyższego wnioskujemy, że C musi być ortonormalna.

Rozważamy endomorfizm $F\colon V \to V$. Jeżeli B, C są bazami ortonormalnymi, to z powyższej uwagi dla $P = m^B_C(\mathrm{id})$ mamy $P^{-1} = P^*$, więc wzór na zmianę bazy przybiera następującą postać:

$$m_C(F) = P m_B(F) P^{-1} = P m_B(F) P^*.$$

(Czyli taką samą, jak wzór na zmianę bazy formy hermitowskiej/rzeczywistej symetrycznej!)

Stwierdzenie 10.6. *Jeżeli $F \in \mathrm{End}(V)$ jest dowolny, to dla $G \in \mathrm{End}(V)$ 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 $P = m_C^B(\mathrm{id})$, 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). □

Definicja 10.7. G jak powyżej mówimy że jest *sprzężony* do F i oznaczamy F^*.

Uwaga 10.8.
- to tylko pozornie kłóci z definicją odwzorowania dualnego $F^* \colon V^* \to V^*$, bo iloczyn skalarny pozwala utożsamić V i V^* (szczegóły: ćwiczenie na liście 13)
- ze stwierdzenia powyżej F^* istnieje i jest jedyny.

Stwierdzenie 10.9. *Niech* $F \in \mathrm{End}(V)$ *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 10.10. Mówimy że $F \in \mathrm{End}(V)$ 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 $A^*A = \mathrm{I}$. □

Przykład 10.11. Jeżeli P jest rzutem ortogonalnym na $\mathrm{Lin}(\overbrace{b_1, \ldots, b_k}^{\text{układ ON}})$, to jest samosprzężony (ale prawie nigdy nie jest ortogonalny, chyba że $P = \mathrm{id}$...): weźmy dowolne $v, w \in V$. Wtedy

$$\langle P(v), w\rangle = \langle \sum_{j=1}^{k} \langle b_j, v\rangle b_j, w\rangle = \sum_{j=1}^{k} \overline{\langle b_j, v\rangle} \langle b_j, w\rangle = \sum_{j=1}^{k} \langle v, b_j\rangle \langle b_j, w\rangle =$$
$$= \langle v, \sum_{j=1}^{k} \langle b_j, w\rangle b_j\rangle = \langle v, P(w)\rangle$$

Można powiedzieć więcej: jeżeli $P \in \mathrm{End}(V)$ jest rzutem (tzn. spełnia $P = P \circ P$), to P jest rzutem ortogonalnym (na swój obraz) wtedy i tylko wtedy gdy jest samosprzężony.

Stwierdzenie 10.12. *Jeżeli F jest nieujemnie określony na przestrzeni* unitarnej*, tzn. $\langle v, F(v)\rangle \geqslant 0$ 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:

$$0 \leqslant \langle \alpha v + w, F(\alpha v + w)\rangle = \bar{\alpha}\alpha \langle v, F(v)\rangle + \langle w, F(w)\rangle + \bar{\alpha}\langle v, F(w)\rangle + \alpha\langle w, F(v)\rangle$$
$$= |\alpha|^2 \langle v, F(w)\rangle + \langle w, F(w)\rangle + \overline{\alpha\langle F(w), v\rangle} + \alpha\langle w, F(v)\rangle.$$

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 $T = T^*$. □

Uwaga 10.13. Z dowodu widzimy, że wystarczy założyć, że $(\forall v)\langle v, T(v)\rangle \in$ **R**.

Przykłady 10.14.

- Rzeczywista wielokrotność przekształcenia samosprzężonego jest samosprzężona.
- Rzeczywista kombinacja liniowa przekształceń samosprzężonych jest samosprzężona.
- W szczególności rzeczywiste kombinacje liniowe rzutów ortogonalnych są samosprzężone

Zobaczymy wkrótce, że ostatni przykład powyżej opisuje *wszystkie* przekształcenia samosprzężone: są to dokładnie rzeczywiste kombinacje liniowe rzutów ortogonalnych!

Lemat 10.15. *Wartości własne zespolonych samosprzężonych endomorfizmów/macierzy są rzeczywiste.*

Dowód. Ćwiczenie. □

Wniosek 10.16. *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 10.17 (Twierdzenie spektralne)**.**

- *Załóżmy że V jest przestrzenią unitarną/euklidesową i $F \in \mathrm{End}(V)$ jest samosprzężony. Wtedy F diagonalizuje się w bazie ON.*
- *Załóżmy że $A \in \mathrm{M}_{n\times n}(\mathbf{C})/\,\mathrm{M}_{n\times n}(\mathbf{R})$ 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 10.18. Z twierdzenia spektralnego łatwo wynika, że samosprzężony $F \in \mathrm{End}(V)$ 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 10.19. Wariant twierdzenia spektralnego jest prawdziwy również dla nieskończenie wymiarowych przestrzeni (nad $\mathbf{R}$ lub $\mathbf{C}$) z iloczynem skalarnym.

Uwaga 10.20. Łatwo pokazać twierdzenie odwrotne do twierdzenia spektralnego: jeżeli F diagonalizuje się w bazie ON i ma rzeczywiste wartości własne, to jest samosprzężony (bo macierz w tej samej bazie jest rzeczywista i diagonalna, więc hermitowska).

Dowód. Druga część łatwo wynika z pierwszej.

Udowodnijmy pierwszą cześć 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 $B' = v_1, v_2', \ldots, v_n'$. 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 samosprzężony, $M = M^*$, więc $M = \begin{pmatrix} \lambda & 0 \\ 0 & N \end{pmatrix}$, gdzie $N = N^*$.

Rozważmy $W = \mathrm{Lin}(v_2', \ldots, v_n')$. Z postaci $m_B(F)$ widzimy, że W jest F-niezmiennicze, więc $F\restriction_W \in \mathrm{End}(W)$. Wtedy dla $B_0' = (v_2', \ldots, v_n')$ mamy $m_{B'}(F\restriction_W) = N = N^*$, więc $F\restriction_W$ jest samosprzężone. Stąd z założenia indukcyjnego istnieje baza ON $B_0 = v_2, v_3, \ldots, v_n$ przestrzeni W składająca się z wektorów własnych F. Stąd $B = (v_1, v_2, \ldots, v_n)$ jest bazą ON wektorów własnych F. $\square$

Przykład 10.21. 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 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 $m_E^B(\mathrm{id}) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$, więc $m_B^E(\mathrm{id}) = m_E^B(\mathrm{id})^{-1} = m_E^B(\mathrm{id})^* = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}$, 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 10.22. Jeżeli $M \in \mathrm{M}_{2\times 2}(\mathbf{C})$ jest hermitowska i $\chi_M(x)$ ma podwójny pierwiastek λ, to z twierdzenia spektralnego M się diagonalizuje, czyli $M = P(\lambda \mathrm{I})P^* = \lambda \mathrm{I}$ 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 10.23. Jeżeli $F \in \mathrm{End}(V)$ jest samosprzężony, 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 (jedyny!) taki samosprzężony $F = F_\varphi \in \mathrm{End}(V)$, że $\varphi(v,w) = \langle v, F(w)\rangle = \langle F(v), w\rangle$ (dowolny 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:

$$\begin{aligned} m_C(F) &= m_C^B(\mathrm{id})m_B(F_\varphi)m_B^C(\mathrm{id}) \\ &= m_B^C(\mathrm{id})^{-1}m^{BB}(\varphi)m_B^C(\mathrm{id}) \\ &= m_B^C(\mathrm{id})^{*}m^{BB}(\varphi)m_B^C(\mathrm{id}) \\ &= m^{CC}(\varphi) \end{aligned}$$

Stąd i z twierdzenia spektralnego wnioskujemy wzmocnienie twierdzenia Lagrange'a.

Wniosek 10.24.
- *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 $m^{BB}(\varphi_1) = \mathrm{I}$ 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 $\langle -,- \rangle = \varphi_1$.* □

Wniosek 10.25. *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 $C = c_1, \ldots, c_n$ 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 $B = (b_1, \ldots, b_n)$ działa: $\varphi(b_j, b_j) = \operatorname{sgn} \varphi(c_j, c_j)$. □

Fakt 10.26 (Prawo bezwładności Sylvestera 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 $\mathbf{R}$. □

Definicja 10.27. *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 10.28. *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 10.25 wynika teza. □

Przykład 10.29. 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 10.30. *Jeżeli F jest samosprzężone 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 10.31 (Rozkład spektralny)**.** *Jeżeli $F \in \mathrm{End}(V)$ jest samosprzężony, 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 $\sum_{\lambda \in \sigma(F)} P_\lambda = \mathrm{id}$ 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 10.32. 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

$$\begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} =$$
$$= \frac{3}{2} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} = 3 \underbrace{\begin{pmatrix} \frac{1}{2} & \frac{i}{2} \\ \frac{-i}{2} & \frac{1}{2} \end{pmatrix}}_{P_3} + \underbrace{\begin{pmatrix} \frac{1}{2} & \frac{-i}{2} \\ \frac{i}{2} & \frac{1}{2} \end{pmatrix}}_{P_1}$$

Uwaga 10.33. 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.

Rozkład singularny i rozkład biegunowy

Definicja 10.34. Dla $A \in \mathrm{M}_{n\times m}(\mathbf{C})$ (niekoniecznie kwadratowej) definiujemy A^* jako $\bar{A}^\top \in \mathrm{M}_{m\times n}(\mathbf{C})$ (tak samo jak dla macierzy kwadratowych).

(Własności * które mieliśmy dla macierzy kwadratowych wciąż działają — te, które mają sens; np. $(AB)^* = B^*A^*$, $(A+B)^* = A^* + B^*$.)

Ćwiczenie 10.35. *Załóżmy że $A \in \mathrm{M}_{n\times m}(\mathbf{C})$ jest dowolną macierzą. Wtedy dla dowolnego $v \in \mathbf{C}^n, w \in \mathbf{C}^m$ mamy*

$$\langle v, Aw\rangle = \langle A^*v, w\rangle.$$

(Gdzie po lewej i prawej stronie mamy standardowe iloczyny skalarne dla $\mathbf{C}^n$ i $\mathbf{C}^m$.)

Niech $A \in \mathrm{M}_{n\times m}(\mathbf{C})$ będzie dowolną macierzą. Rozważmy wartość $R(v) = \frac{|Av|}{|v|}$ dla $v \in \mathbf{C}^m \setminus \{0\}$. 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 $Ab_j = \lambda_j b_j$. Wtedy

$$R(v)^2 = \langle v, A^*Av\rangle = \sum_{j=1}^n \alpha_j \langle b_j, \lambda_j b_j\rangle = \sum_{j=1}^n \alpha_j \lambda_j.$$

Ponieważ ta liczba jest zawsze nieujemna, widzimy że $\lambda_j \geqslant 0$.

Wniosek 10.36. *Jeżeli $A \in \mathrm{M}_{n\times m}(\mathbf{C})$ jest dowolną macierzą, to A^*A jest nieujemnie określona.*

Możemy zmienić kolejność b_j tak że $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_m$.

Niech $k = \mathrm{rk} A^*A$. Wtedy $\lambda_1, \ldots \lambda_k > 0$ i $A^*Ab_j \neq 0$ dla $j \leqslant k$, a więc też $Ab_j \neq 0$. Ponadto wektory te są ortogonalne: jeżeli $j < l \leqslant k$, to

$$\langle Ab_j, Ab_l\rangle = \langle b_j, A^*Ab_l\rangle = \lambda_l \langle b_j, b_l\rangle = 0,$$

Połóżmy dla $j \leqslant k$:

$$c_j = \frac{Ab_j}{|Ab_j|} = \frac{Ab_j}{R(b_j)} = \frac{Ab_j}{\sqrt{\lambda}_j}.$$

Wtedy układ $c_1, \ldots, c_k \in \mathbf{C}^n$ jest ON, możemy więc rozszerzyć go do bazy ON $C = c_1, \ldots, c_n$ przestrzeni $\mathbf{C}^n$.

Twierdzenie 10.37 (rozkład singularny)**.** *Jeżeli $A \in \mathrm{M}_{n\times m}(\mathbf{C})$ jest dowolna, to istnieją macierze unitarne $U_1 \in \mathrm{M}_{n\times n}(\mathbf{C})$, $U_2 \in \mathrm{M}_{m\times m}(\mathbf{C})$ i diagonalna $D \in \mathrm{M}_{n\times m}(\mathbf{R}_{\geqslant 0})$ taka że*

$$A = U_1 D U_2^*,$$

przy czym

$$D = \left.\underbrace{\begin{pmatrix} \sqrt{\lambda_1} & & & & \\ & \sqrt{\lambda_2} & & \text{\Large 0} & \\ & & \ddots & & \\ & & & \sqrt{\lambda_k} & \\ & \text{\Large 0} & & & \text{\Large 0} \end{pmatrix}}_{m \text{ kolumn}}\right\} n \text{ wierszy},$$

gdzie $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_k$ to dodatnie (równoważnie, niezerowe) wartości własne A^*A.

Ponadto jeżeli A jest rzeczywista, to możemy wybrać U_1, U_2 rzeczywiste.

Dowód. Macierze $U_1 = (c_1, \ldots, c_n)$, $U_2 = (b_1, \ldots, b_m)$ działają. Istotnie, $U_2^* = U_2^{-1}$, więc dla każdego b_j mamy:

$$U_1 D U_2^* b_j = U_1 D e_j = \begin{cases} U_1 \sqrt{\lambda_j} e_j = \sqrt{\lambda_j} U_1 e_j = \sqrt{\lambda_j} c_j = A b_j & j \leqslant k \\ U_1 0 = 0 = A b_j & j > k. \end{cases}$$

Ponieważ $B = b_1, \ldots, b_m$ jest bazą $\mathbf{C}^m$, wynika z tego że $A = U_1 D U_2^*$. □

Uwaga 10.38. Jeżeli $A = U_1 D U_2^*$, to $A^* = U_2 D^\top U_1^*$.

Przykład 10.39. Weźmy macierz $A = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 0 & 1 \end{pmatrix}$. Szukamy jej rozkładu singularnego. Wyznaczmy 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

$c_1 = \frac{A^\top b_1}{|A^\top b_1|} = \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\0\\1\end{pmatrix}$, $c_2 = \frac{A^\top b_2}{|A^\top b_2|} = \frac{1}{\sqrt{6}}\begin{pmatrix}1\\2\\1\end{pmatrix}$. 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 10.40. *Jeżeli $F: V \to W$ 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 $U_2 = m_{B'}^{B}(\mathrm{id}_W), U_1 = m_{C'}^{C}(\mathrm{id}_V)$ i wtedy

$$m_C^B(F) = m_C^{C'}(\mathrm{id}) m_{C'}^{B'}(F) m_{B'}^{B}(\mathrm{id}) = U_1^* U_1 D U_2^* U_2 = D. \qquad \square$$

Definicja 10.41. Wartości $\sigma_j = \sqrt{\lambda_j}$ na przekątnej w D powyżej nazywamy *wartościami singularnymi F* (lub macierzy A).

Uwaga 10.42. Przypomnijmy sobie wzór dla v jednostkowych:

$$\frac{|Av|^2}{|v|^2} = R(v)^2 = \langle v, A^*Av\rangle = \sum_{j=1}^{n} \alpha_j \langle b_j, \lambda_j b_j\rangle = \sum_{j=1}^{n} \alpha_j \lambda_j = \sum_{j=1}^{k} \alpha_j \lambda_j = \sum_{j=1}^{k} \alpha_j \sigma_j^2.$$

Zauważmy że stąd:

- σ_1 to maksymalna wartość $R(v)$ dla $v \neq 0$,
- b_1 to wektor jednostkowy dla którego ta wartość jest osiągana, 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 osiągana, c_2 to wektor jednostkowy w kierunku Ab_2,
- σ_3 to maksymalna wartość $R(v)$ dla $v \perp b_1, b_2$...
- itd.

Ponadto dla $j \leqslant k$ zachodzą wzory:

$$Ab_j = \sigma_j c_j \qquad\qquad A^* c_j = \sigma_j b_j.$$

(Dla $j > k$ mamy $Ab_j = 0$ i $A^* c_j = 0$, o ile te napisy mają sens.)

Drugi wzór wynika z obserwacji, że

$$A^* c_j = A^* \sigma_j^{-1} Ab_j = \sigma_j^{-1} A^* Ab_j = \sigma_j^{-1} \sigma_j^2 b_j = \sigma_j b_j.$$

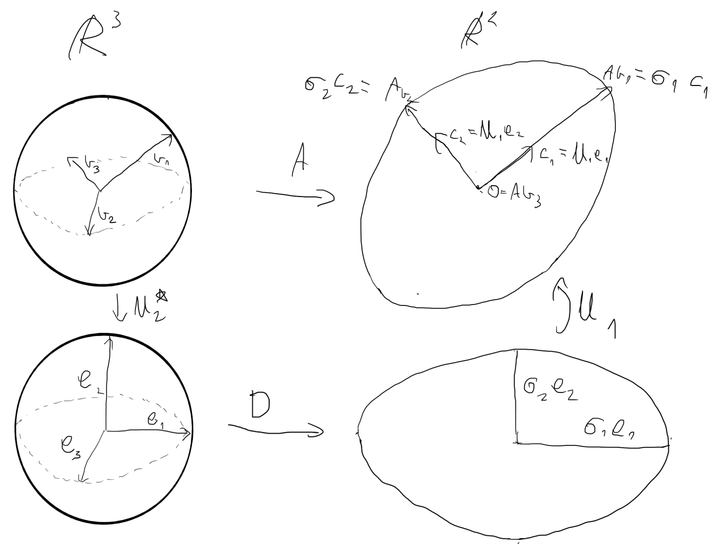
Stąd bierze się interpretacja geometryczna rozkładu singularnego (rysunek 1).

Definicja 10.43. Wektor b_j nazywamy *prawym wektorem singularnym* (odpowiadającym σ_j), a wektor c_j *lewym wektorem singularnym.* (Mnemotechnika: w $\sigma_j = \langle c_j, Ab_j \rangle$ 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 $\sigma_1 = \ldots = \sigma_n = 1$).

Przykład 10.44 (metoda najmniejszych kwadratów)**.** Załóżmy że mamy N punktów $P_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, P_2, \ldots, P_N$ na płaszczyźnie $V = \mathbf{R}^2$ (lub $\mathbf{C}^2$). Chcemy znaleźć prostą $W \leqslant V$ najlepiej przybliżającą te punkty w tym sensie, że minimalizuje sumę kwadratów odległości:

$$\sum_{j=1}^{N} d(P_j, W)^2$$



Rysunek 10.1: Interpretacja geometryczna rozkładu singularnego odwzorowania $\mathbf{R}^3 \to \mathbf{R}^2$.

Biorąc wektor jednostkowy v w kierunku W, mamy

$$
\begin{aligned}
d(P_j, W)^2 &= \langle P_j - \langle v, P_j\rangle v, P_j - \langle v, P_j\rangle v\rangle \\
&= \langle P_j, P_j - \langle v, P_j\rangle v_j\rangle - \langle\langle v, P_j\rangle v, P_j - \langle v, P_j\rangle v\rangle \\
&= |P_j|^2 - \langle P_j, v\rangle\langle v, P_j\rangle - \overline{\langle v, P_j\rangle}\langle v, P_j - \langle v, P_j\rangle v\rangle \\
&= |P_j|^2 - |\langle v, P_j\rangle|^2 - \langle P_j, v\rangle\langle v, P_j - \langle v, P_j\rangle v\rangle \\
&= |P_j|^2 - |\langle v, P_j\rangle|^2 - \langle P_j, v\rangle(\langle v, P_j\rangle - \langle v, P_j\rangle\langle v, v\rangle) \\
&= |P_j|^2 - |\langle v, P_j\rangle|^2.
\end{aligned}
$$

Stąd szukana prosta jest rozpinana przez jednostkowy v maksymalizujący $\sum_{j=1}^{N} |\langle v, P_j\rangle|^2$.

Rozważmy macierz $A = \overline{\begin{pmatrix} x_1 & y_1 \\ \vdots & \vdots \\ x_N & y_N \end{pmatrix}}$. 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 $A \in \mathrm{M}_{N\times n}(\mathbf{R})$, 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 10.45 (rozkład biegunowy).

- *Jeżeli $F \in \mathrm{End}(V)$ i B jest dowolną bazą ON, to istnieje taka baza ON C że $m_C^B(F)$ jest nieujemnie określona.*

- *Jeżeli $A \in \mathrm{M}_{n\times n}(\mathbf{C})$ 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 $A = U_1DU_2$, weźmy $P = U_2^*DU_2$, $U = U_1U_2$.

Pierwsza część wynika z drugiej: wychodząc z $A = m_B(F)$ musimy wybrać za C taką bazę V, że $m_C^B(\mathrm{id}) = U$. Taka baza zawsze istnieje, bo macierz U jest unitarna. □

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.*

Dowód. Niech v będzie wektorem własnym odpowiadającej wartości własnej5 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 $|\lambda| = 1$. □

Stwierdzenie 11.2. *(V nad dowolnym ciałem, być może nieskończenie wymiarowa.) Jeżeli $F \in \mathrm{End}(V)$ jest odwracalny i $W \leqslant V$ jest skończenie wymiarowa i F-niezmiennicza, to W jest F^{-1}-niezmiennicza.*

Dowód. Niech $W' = F[W] \leqslant W$. Wtedy F zadaje izomorfizm $W \to W'$, więc $\dim W = \dim W'$. Stąd $W = W'$, czyli każdy $w' \in W$ jest postaci $F(w)$ dla pewnego $w \in W$, zatem $F^{-1}(w') = w \in W$. □

Uwaga 11.3. Jeżeli $F \in \mathrm{End}(V)$ i $W \leqslant V$ 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). □

Stwierdzenie 11.5. *(V unitarna.) Jeżeli F jest unitarne, to diagonalizuje się w bazie ON.*

pierwzy dowód. Wynika z Twierdzenia A.6 (spektralnego dla normalnych endomorfizmów). □

drugi dowód. Indukcja względem $\dim V$. Gdy $\dim V = 1$, to teza jest oczywista. W przeciwnym wypadku weźmy jednostkowy wektor własny b_1. Wtedy $\mathrm{Lin}(b_1)$ jest F-niezmiennicze, więc $W = b_1^\perp$ też i $F\restriction_W$ jest oczywiście unitarne. Z założenia indukcyjnego W ma bazę B' ON wektorów własnych dla $F\restriction_W$. Wtedy $B = b_1B'$ działa. □

Twierdzenie 11.6 (twierdzenie spektralne dla przekształceń unitarnych)**.** *(V unitarna.) $F \in \mathrm{End}(V)$ 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:

$$m_B(F)m_B(F)^* = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}\begin{pmatrix} \bar\lambda_1 & & & \\ & \bar\lambda_2 & & \\ & & \ddots & \\ & & & \bar\lambda_n \end{pmatrix} = \begin{pmatrix} |\lambda_1| & & & \\ & |\lambda_2| & & \\ & & \ddots & \\ & & & |\lambda_n| \end{pmatrix} = \mathrm{I},$$

więc $m_B(F)$ jest unitarna, czyli F jest unitarne. □

Uwaga 11.7. Jeżeli F jest przekształceniem ortogonalnym przestrzeni euklidesowej V, to $F_\mathbb{C}$ jest przekształceniem unitarnym $V_\mathbb{C}$: istotnie, wystarczy wziąć bazę ON B dla V i wtedy B jest też bazą ON dla $V_\mathbb{C}$ i macierz $m_B(F) = m_B(F_\mathbb{C})$ jest ortogonalna i rzeczywista, a więc unitarna.

Ćwiczenie 11.8. *(V rzeczywista) Jeżeli $W \leqslant V_\mathbb{C}$, to $W \cap V \leqslant V$.*

Uwaga 11.9. Jeżeli $F \in \mathrm{End}(V)$ dla V — rzeczywistej przestrzeni liniowej i $W \leqslant V_\mathbb{C}$ jest $F_\mathbb{C}$-niezmiennicze, to $W \cap V \leqslant V$ jest F-niezmiennicze: istotnie, $F[W] \subseteq V$ i $F_\mathbb{C}[W] = F[W] \subseteq W$.

Ćwiczenie 11.10. *Jeżeli $F \in \mathrm{End}(V)$ dla rzeczywistej V i $v \in V_{\mathbf{C}}$ jest wektorem własnym dla $F_{\mathbf{C}}$ odpowiadającym nierzecywistej wartości własnej, to* $\mathrm{Lin}_{\mathbf{C}}(v, \bar{v}) \cap V = \mathrm{Lin}_{\mathbf{R}}(v + \bar{v}, i(v - \bar{v}))$ *oraz* $\dim_{\mathbf{R}}(\mathrm{Lin}_{\mathbf{C}}(v, \bar{v}) \cap V) = 2$.

Twierdzenie 11.11 (twierdzenie spektralne dla przekształceń ortogonalnych)**.** *(V euklidesowa.) $F \in \mathrm{End}(V)$ 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 11.12. 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.

Indukcja względem $\dim V$.

Gdy $\dim V = 1$, teza jest oczywista. Gdy $\dim V = 2$, to (utożsamiając V z $\mathbf{R}^2$) wiemy z poprzedniego semestru, że F jest symetrią lub obrotem. Jeżeli jest symetrią, to biorąc wektory b_1 równoległy i b_2 prostopadły do osi symetrii mamy $m_{b_1 b_2}(F) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Jeżeli jest obrotem o kąt θ, to w każdej bazie ON B macierz $m_B(F) = R_\theta$. Załóżmy więc że $\dim V > 2$.

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 $F\restriction_W$ ma żądaną postać i $B = b_1 B'$ działa.

Jeżeli F nie ma rzeczywistej wartości własnej, to weźmy wektor własny $v \in V_{\mathbf{C}}$ dla $F_{\mathbf{C}}$ odpowiadający nierzeczywistej wartości własnej λ. Wtedy $\mathrm{Lin}_{\mathbf{C}}(v, \bar{v})$ jest $F_{\mathbf{C}}$-niezmienniczy, więc $W_1 = \mathrm{Lin}_{\mathbf{C}}(v, \bar{v}) \cap V$ jest F-niezmienniczy. Z ćwiczenia wiemy że $\dim W_1 = 2$, więc z założenia indukcyjnego W_1 i $W_2 = W_0^\perp$ mają bazy ON B_1, B_2 takie że $m_{B_j}(F\restriction_{W_j})$ są jak w tezie (z tym że bez wartości ± 1, bo inaczej F miałoby wartość własną!). Stąd $B = B_1 B_2$ działa. □

11.2 Izometrie

Definicja 11.13. Mówimy że odwzorowanie $F\colon V \to V$ (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 11.14.

- translacje $T_{v_0}(v) = v + v_0$ są izometriami,
- obroty $\mathbf{R}^2$ wokół punktów, $\mathbf{R}^3$ wokół prostych są izometriami,
- sprzężenie w $\mathbf{C}^n$ jest izometrią (również sprzężenie tylko niektórych współrzędnych),
- przekształcenie $\mathbf{R}^n/\mathbf{C}^n$ zadane (w standardowej bazie) macierzą diagonalną postaci

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix} \begin{array}{l} \\ \\ \\ \leftarrow i\text{-te miejsce} \\ \\ \\ \\ \end{array}$$

  (symetria w hiperpłaszczyźnie $\mathrm{Lin}(e_1, \ldots, e_{i-1}, e_{i+1} \ldots, e_n)$) jest izometrią,

- symetrie w dowolnej hiperpłaszczyźnie $W \ni 0$: $S_W = 2\pi_W - \mathrm{id}$,
- przekształcenie $\mathbf{R}^n/\mathbf{C}^n$ zadane (w standardowej bazie) macierzą postaci
$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & R_\theta & & \\ & & & & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix} \leftarrow i\text{-te i } (i+1)\text{-sze miejsce}$$
(obrót wokół “hiperprostej” $\mathrm{Lin}(e_1, \ldots, e_{i-1}, e_{i+2}, \ldots, e_n)$) jest izometrią,
- podobnie symetrie w innych hiperpłaszczyznach/obroty wokół innych hiperprostych (niekoniecznie przechodzących przez 0).

Wniosek 11.15. *Odwzorowanie liniowe $F: V \to V$ jest izometrią wtedy i tylko wtedy gdy jest unitarne/ortogonalne.*

Dowód. Załóżmy że F jest unitarne/ortogonalne i weźmy dowolne $v_1, v_2 \in V$. 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 załóżmy, że F jest liniową izometrią.

Dowód w przypadku unitarnym: ćwiczenie. Dowód w przypadku euklidesowym.

Wtedy z wzoru polaryzacyjnego (i liniowości F) dla dowolnych v, w mamy
$$\begin{aligned} \langle v, w \rangle &= \frac{|v + w|^2 - |v|^2 - |w|^2}{2} \\ &= \frac{|F(v + w)| - |F(v)|^2 - |F(w)|^2}{2} \\ &= \frac{|F(v) + F(w)| - |F(v)|^2 - |F(w)|^2}{2} \\ &= \langle F(v), F(w) \rangle. \end{aligned}$$
$\square$

Do następnego stwierdzenia będzie potrzebna jeszcze inna postać wzoru polaryzacyjnego dla form symetrycznych.

Uwaga 11.16. 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 11.17. *(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 $B = (b_1, \ldots, b_n)$. Oznaczmy $b'_j := F(b_j)$. Wtedy $B' = b'_1, \ldots, b'_n$ 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. □

Uwaga 11.18. Ten sam dowód pokazuje, że przekształcenie przestrzeni unitarnej zachowujące iloczyn skalarny musi być liniowe.

Wniosek 11.19. *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 $F = F(0) + (F - F(0)) = T_{F(0)} \circ (F - F(0))$. □

Wniosek 11.20. *Każda izometria n-wymiarowej przestrzeni euklidesowej jest złożeniem $\leqslant 1$ translacji, k obrotów i m symetrii w hiperpłaszczyznach, przy czym $2k + m \leqslant n$.*

Dowód. Wynika natychmiast z poprzedniego wniosku z Twierdzenia 11.11 (spektralnego dla ortogonalnych endomorfizmów). □

Ćwiczenie 11.21. *Każda izometria n-wymiarowej przestrzeni euklidesowej jest złożeniem $\leqslant n+1$ symetrii w hiperpłaszczyznach (w tym najwyżej jednej nieliniowej/w hiperpłaszczyźnie nie zawierającej 0).*

Przykład 11.22. Mamy daną macierz izometrii

$$\begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{\sqrt{2}}{2} & -\frac{\sqrt{3}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ -\frac{\sqrt{3}}{2} & \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} & 0 \end{pmatrix}$$

Chcemy wyznaczyć jej postać kanoniczną. Najpierw szukamy wektorów własnych (koniecznie z wartościami własnymi 1 i -1).

Dla 1 mamy układ równań:

$$\begin{cases} \frac{\sqrt{3}}{2}x - \frac{\sqrt{2}}{2}y - \frac{\sqrt{3}}{2}z + \frac{\sqrt{2}}{2}t = x \\ \frac{\sqrt{2}}{2}x + \frac{\sqrt{3}}{2}y - \frac{\sqrt{2}}{2}z = y \\ -\frac{\sqrt{3}}{2}x + \frac{\sqrt{2}}{2}y + \frac{\sqrt{3}}{2}z + \frac{\sqrt{2}}{2}t = z \\ \frac{\sqrt{2}}{2}x + \frac{\sqrt{2}}{2}z = t \end{cases}$$

z czwartego równania $t = \frac{x+z}{\sqrt{2}}$. Z drugiego równania $(1 - \sqrt{3}/2)y = \frac{x-z}{\sqrt{2}}$. Wstawiając do pierwszego równania mamy

$$\begin{aligned} x &= \frac{\sqrt{3}}{2}x - \frac{y}{\sqrt{2}} - \frac{\sqrt{3}}{2}z + \frac{t}{\sqrt{2}} \\ &= \frac{\sqrt{3}}{2}x - (x-z)\underbrace{\frac{1}{2-\sqrt{3}}}_{\sqrt{3}+2} - \frac{\sqrt{3}}{2}z + \frac{x+z}{2} \\ &= x\left(\frac{\sqrt{3}}{2} - \sqrt{3} - 2 + \frac{1}{2}\right) + z\left(\sqrt{3} + 2 - \frac{\sqrt{3}-1}{2}\right) \\ &= -x\frac{\sqrt{3}+3}{2} + z\frac{\sqrt{3}+5}{2}, \end{aligned}$$

czyli $x = z$. Stąd $y = 0$, $t = \sqrt{2}x$. Faktycznie, wektor $\begin{pmatrix} 1 \\ 0 \\ 1 \\ \sqrt{2} \end{pmatrix}$ spełnia równanie, jest punktem stałym. Po unormowaniu dostajemy $b_1 = \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$.

Szukając wektorów własnych z wartością własną -1 możemy szukać tylko wśród wektorów ortogonalnych do b_1. To prowadzi do układu równań:

$$\begin{cases} \frac{\sqrt{3}}{2}x - \frac{\sqrt{2}}{2}y - \frac{\sqrt{3}}{2}z + \frac{\sqrt{2}}{2}t = -x \\ \frac{\sqrt{2}}{2}x + \frac{\sqrt{3}}{2}y - \frac{\sqrt{2}}{2}z = -y \\ -\frac{\sqrt{3}}{2}x + \frac{\sqrt{2}}{2}y + \frac{\sqrt{3}}{2}z + \frac{\sqrt{2}}{2}t = -z \\ \frac{\sqrt{2}}{2}x + \frac{\sqrt{2}}{2}z = -t \\ x + z + \sqrt{2}t = 0 \end{cases}$$

Znów z czwartego równania $t = -\frac{x+z}{\sqrt{2}}$. To jest równoważne z piątym równaniem, więc w tym przypadku ono nie pomaga. Znów z drugiego równania $(1 + \frac{\sqrt{3}}{2})y = \frac{z-x}{\sqrt{2}}$ i wstawiając do pierwszego równania mamy

$$\begin{aligned} -x &= \frac{\sqrt{3}}{2}x - \frac{y}{\sqrt{2}} - \frac{\sqrt{3}}{2}z + \frac{t}{\sqrt{2}} \\ &= \frac{\sqrt{3}}{2}x - (z-x)\underbrace{\frac{1}{2+\sqrt{3}}}_{2-\sqrt{3}} - \frac{\sqrt{3}}{2}z\frac{x+z}{2} \\ &= x(\frac{\sqrt{3}}{2} + \sqrt{3} + 2 - \frac{1}{2}) + z(-\sqrt{3} + 2 - \frac{\sqrt{3}+1}{2}) \\ &= x\frac{3\sqrt{3}+3}{2} - z\frac{3\sqrt{3}-3}{2}, \end{aligned}$$

co znów daje $x = z$. Stąd też $y = 0$ i $t = -\sqrt{2}x$ i faktycznie, wektor $\begin{pmatrix} 1 \\ 0 \\ 1 \\ -\sqrt{2} \end{pmatrix}$, po unormowaniu $b_2 = \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$ jest wektorem własnym z wartością własną -1.

Ponieważ nie ma więcej wektorów własnych, odwzorowanie na płaszczyźnie $b_1^\perp \cap b_2^\perp$ jest izometrią bez wektorów własnych, a więc obrotem. W tym przypadku za b_1, b_2 możemy wziąć dowolną bazę ON (macierz obrotu płaszczyzny w każdej jej bazie ON jest macierzą obrotu, zmienić się może najwyżej kierunek, w zależności od orientacji).

Stąd możemy dowolnie wybrać b_3, b_4 jako bazę ON $b_1^\perp \cap b_2^\perp$. Np. $b_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, $b_4 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \\ 0 \\ 0 \end{pmatrix}$. Wtedy $Ab_3 = \frac{\sqrt{3}}{2} b_3 - \frac{1}{2} b_4$ (i, choć tego nie musimy już sprawdzać, $Ab_4 = \frac{1}{2} b_3 + \frac{\sqrt{3}}{2} b_4$). Zatem postać kanoniczna A (jedna z możliwych) to

$$A = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & \frac{\sqrt{3}}{2} & \frac{1}{2} \\ & & -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \end{pmatrix}^\top$$

Przykład 11.23. Izometria $\mathbf{R}^4$ zadana macierzą

$$A = \begin{pmatrix} \frac{1}{6} & \frac{-2\sqrt{3}+1}{6} & \frac{-2\sqrt{3}-1}{6} & -\frac{1}{2} \\ \frac{2\sqrt{3}+1}{6} & \frac{1}{6} & \frac{2\sqrt{3}-1}{6} & -\frac{1}{2} \\ \frac{2\sqrt{3}-1}{6} & \frac{-2\sqrt{3}-1}{6} & \frac{1}{6} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

nie ma wektorów własnych z wartością własną 1 ani -1. Jest więc złożeniem dwóch obrotów wokół pewnych płaszczyzn. Jak je znaleźć?

- wyznaczyć zespoloną wartość własną λ dla A
- znaleźć wektor własny $v \in \mathbf{C}^4$ dla λ

- płaszczyzna w $\mathbf{R}^4$ rozpięta przez rzeczywistą i urojoną część v będzie płaszczyzną w której zachodzi obrót (o argument λ)

- płaszczyzna do niej ortogonalna będzie płaszczyzną w której zachodzi drugi obrót

Ogólnie rachunki są typowo dość skomplikowane (bo wyznaczanie wektorów własnych typowo jest rachunkowo skomplikowane).

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” $\mathbf{R}^n$, tzn. zbiór rozwiązań równania postaci

$$\sum_{i=1}^{n} a_{ii} v_i^2 + \sum_{i<j} 2a_{ij} v_i v_j = c,$$

ma w pewnym prostokątnym układzie współrzędnych postać kanoniczną

$$\sum_{i=1}^{n} b_{ii}(v_i')^2 = c,$$

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 $\mathbf{R}^3$.

Przykład 11.24. 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 izotropowego 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 11.25. *Macierz Grama* (lub *Gramian*) $G(v_1, \ldots, v_k)$ układu $v_1, v_2, \ldots, v_k$ w przestrzeni unitarnej/euklidesowej to macierz której ij-ty wyraz to $\langle v_i, v_j \rangle$.

Uwaga 11.26. W szczególności jeżeli $B = v_1, \ldots, v_k$ jest bazą, to jej macierz Grama to po prostu macierz standardowego iloczynu skalarnego względem tej bazy.

Uwaga 11.27. Jeżeli $v_1, \ldots, v_k \in \mathbf{R}^n/\mathbf{C}^n$, to $G(v_1, \ldots, v_k) = A^*A$, gdzie A to macierz o kolumnach v_j.

Ogólnie, dla dowolnej przestrzeni euklidesowej V i bazy ON B, $G(v_1, \ldots, v_k) = A^*A$, gdzie A to macierz o kolumnach $[v_j]_B$.

Sprawdziliśmy wcześniej, że macierze postaci A^*A są nieujemnie określone. Stąd:

Wniosek 11.28. *Macierz Grama jest hermitowska/symetryczna i nieujemnie określona.*

Stwierdzenie 11.29. *Jeżeli* $A \in \mathrm{M}_{n\times m}(\mathbf{C})$ *jest dowolną macierzą, to* $\mathrm{rk}A = \mathrm{rk}(A^*A) = \mathrm{rk}(AA^*)$.

Dowód. Pokażmy że $\mathrm{rk}A = \mathrm{rk}A^*A$. 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

$$\underline{Av = 0} \implies \underline{A^*Av = 0} \implies \langle v, A^*Av\rangle = 0 \iff \langle Av, Av\rangle = 0 \iff$$
$$\iff |Av| = 0 \iff \underline{Av = 0}.$$

czyli $Av = 0 \implies A^*Av = 0 \implies Av = 0$. □

Wniosek 11.30. *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* $\mathrm{Lin}(v_1, \dots, v_k)$.

Dowód. Wystarczy zastosować poprzednie stwierdzenie dla $A = ([v_1]_B, \dots, [v_k]_B)$, gdzie B jest bazą ON V. □

Chcemy za pomocą macierzy Grama zdefiniować objętość równoległościanu rozpiętego przez wektory $v_1, \dots, v_k$.

Stwierdzenie 11.31. *Weźmy dowolne wektory* liniowo niezależne $v_1, \dots, v_k$ *w przestrzeni euklidesowej* V *i bazę ON* B *dla* $\mathrm{Lin}(v_1, \dots, v_k)$*. 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. $V_k'(v_1, \dots, v_k) = V_{k-1}'(v_1, \dots, v_{k-1}) \cdot d(v_k, \mathrm{Lin}(v_1, \dots, v_{k-1}))$*, 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 $W \coloneqq \mathrm{Lin}(v_1, \dots, v_{k-1})$ (a b_k dopełnia B' do bazy ON $\mathrm{Lin}(v_1, \dots, v_n)$). Wtedy:

$$d(v_k, W) = |v - \pi_W(v)| = |\langle b_k, v_k\rangle b_k| = |\langle b_k, v_k\rangle|.$$

Oznaczmy $c \coloneqq \langle b_k, v_k\rangle$ i $A' \coloneqq ([v_1]_{B'}, \dots, [v_{k-1}]_{B'})$. 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. □

Uwaga 11.32. 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 11.33. 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 11.34. Chcemy policzyć odległość od płaszczyzny $\mathrm{Lin}\left(v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}\right)$ wektora $v_3 \begin{pmatrix} 1 \\ 3 \\ 0 \\ 1 \end{pmatrix}$. Policzmy iloczyny skalarne:

$$\left\langle \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} \right\rangle = 2 \quad \left\langle \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} \right\rangle = -1 \quad \left\langle \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 0 \\ 1 \end{pmatrix} \right\rangle = -2$$

$$\left\langle \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} \right\rangle = 7 \quad \left\langle \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 0 \\ 1 \end{pmatrix} \right\rangle = 8 \quad \left\langle \begin{pmatrix} 1 \\ 3 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 0 \\ 1 \end{pmatrix} \right\rangle = 11$$

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

$$d(v_3, \mathrm{Lin}(v_1, v_2)) = \frac{V_3(v_1, v_2, v_3)}{V_2(v_1, v_2)} = \sqrt{\frac{\begin{vmatrix} 2 & -1 & -2 \\ -1 & 7 & 8 \\ -2 & 8 & 11 \end{vmatrix}}{\begin{vmatrix} 2 & -1 \\ -1 & 7 \end{vmatrix}}} =$$

$$= \sqrt{\frac{14 \cdot 11 + 16 + 16 - 8 \cdot 16 - 11 - 2 \cdot 14}{13}} = \sqrt{\frac{9 \cdot 14 - 6 \cdot 16 - 11}{13}} =$$

$$= \sqrt{\frac{3 \cdot 14 - 12 - 11}{13}} = \sqrt{\frac{19}{13}}$$

Przykład 11.35. Chcemy policzyć wartość

$$\inf_{a,b \in \mathbf{R}} \int_0^1 |t^2 - at - b|^2 \, \mathrm{d}t.$$

Zauważmy że ta liczba to odległość x^2 od $\mathrm{Lin}(x, 1) \leqslant V = \mathbf{R}_2[x]$ z iloczynem skalarnym $\langle P, Q \rangle = \int_0^1 P(t)Q(t) \, \mathrm{d}t$.

Łatwo policzyć, że $\langle x^i, x^j \rangle = \int_0^1 t^{i+j} \, \mathrm{d}t = \frac{1}{i+j+1}$. 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

$$d(x^2, \mathrm{Lin}(1, x)) = \frac{V_3(1, x, x^2)}{V_2(1, x)} = \sqrt{\frac{12}{2160}} = \sqrt{\frac{1}{180}} = \frac{1}{6\sqrt{5}}.$$

W szczególności wynika stąd:

$$\inf_{a,b\in\mathbf{R}} \int_0^1 |t^2 - at - b|^2 \,\mathrm{d}t = \frac{1}{6\sqrt{5}}$$

Ćwiczenie 11.36. *Jeżeli macierz kwadratowa $A \in \mathrm{M}_{n\times n}(\mathbf{C})$ jest dodatnio półokreślona, to jest macierzą Grama.*

Wniosek 11.37. *Niech $A \in \mathrm{M}_{n\times n}(\mathbf{C})$ 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.

Załóżmy że A jest dodatnio określona. Wtedy z (1) mamy $A = G(v_1, \ldots, v_n)$. Wiemy, że odwracalność A jest równoważna z liniową niezależnością $v_1, \ldots, v_n$, a dla dodatnio półokreślonej macierzy jest równoważna z dodatnią określonością (z tym że 0 nie jest wartością własną). □

Wniosek 11.38. *Jeżeli $F \in \mathrm{End}(V)$, gdzie V jest euklidesowa lub unitarna i* $\dim V = n$*, to dla każdych wektorów $v_1, \ldots, v_n$ mamy*

$$V_n(F(v_1), \ldots, F(v_n)) = |\det F| V_n(v_1, \ldots, v_n).$$

Dowód. Jeżeli $v_1, \ldots, v_n$ są lz, to obydwie strony są zerowe i teza zachodzi. Załóżmy więc że są lnz.

Oznaczmy przez $A := ([v_1]_B, \ldots, [v_n]_B)$, $A' := ([F(v_1)]_B, \ldots, [F(v_n)]_B)$, $M := m_B(F)$. Wtedy $A' = MA$, skąd

$$\begin{aligned} V_n(F(v_1), \ldots, F(v_n)) &= |\det A'| \\ &= |\det(MA)| \\ &= |\det M \det A| \\ &= |\det M||\det A| \\ &= |\det F| V_n(v_1, \ldots, v_n). \end{aligned}$$

□

Przykład 11.39 (całkowanie przez podstawienie)**.** Mamy przekształcenie liniowe $F\colon \mathbf{R}^n \to \mathbf{R}^n$, funkcję ciągłą $f\colon \mathbf{R}^n \to \mathbf{R}$, a także podzbiór $V \subseteq \mathbf{R}^n$ który jest sumą równoległościanów.

Wtedy

$$
\begin{aligned}
\int_{F[V]} f(v)\,\mathrm{d}v &= \lim_{\mathscr{P} \text{ podział } F[V]} \sum_{P\in\mathscr{P}} V_n(P) f(v_P) \\
&= \lim_{\mathscr{P} \text{ podział } V} \sum_{P\in\mathscr{P}} V_n(F[P]) f(F(v_P)) \\
&= \lim_{\mathscr{P} \text{ podział } V} \sum_{P\in\mathscr{P}} |\det F| V_n(P) f(F(v_P)) \\
&= \int_V |\det F| f(F(v))\,\mathrm{d}v.
\end{aligned}
$$

gdzie $\lim_{\mathscr{P} \text{ podział } F[V]}$ przebiega skończone zbiory $\mathscr{P}$ n-wymiarowych prostopadłościanów o coraz mniejszych średnicach wypełniających $F[V]$, a v_P jest dowolnym wektorem z P.

Następnie możemy ten zbiór 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 funkcji liniowej i translacji (np. F może być klasy C^1) — wtedy $\det F$ należy zastąpić jakobianem F.

(Szczegóły: wykład analiza matematyczna/geometria różniczkowa.)

Dodatek A

Twierdzenie spektralne dla endomorfizmów normalnych

W tym dodatku V jest przestrzenią unitarną lub euklidesową.

Definicja A.1. Mówimy że macierz kwadratowa $A \in \mathrm{M}_{n\times n}(\mathbf{C})$ jest normalna, jeżeli $AA^* = A^*A$.

Mówimy że $F \in \mathrm{End}(V)$ jest normalny, jeżeli $F^* \circ F = F \circ F^*$.

Przykłady A.2.

- Endomorfizmy/operatory unitarne są normalne: $U^*U = \mathrm{I} = UU^*$.
- Endomorfizmy/macierze samosprzężone są normalne: $AA^* = AA = A^*A$.
- Zespolone wielokrotności normalnych endomorfizmów/macierzy są normalne: $(zA)^*(zA) = |z|^2A^*A = |z|^2AA^* = (zA)(zA)^*$.
- 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 A.3. Można pokazać, że macierz A jest normalna wtedy i tylko wtedy gdy istnieją samosprzężone $A_1 = A_1^*, A_2 = A_2^*$ takie że $A = A_1 + iA_2$ i $A_1A_2 = A_2A_1$. Podobnie dla endomorfizmów.

Uwaga A.4. $F \in \mathrm{End}(V)$ jest normalny wtedy i tylko wtedy gdy dla pewnej/każdej bazy ON B macierz $m_B(F)$ jest normalna:

$$FF^* = F^*F \iff m_B(FF^*) = m_B(F^*F) \iff$$
$$\iff m_B(F)m_B(F^*) = m_B(F^*)m_B(F) \iff m_B(F)m_B(F)^* = m_B(F)^*m_B(F).$$

Wniosek A.5. *Jeżeli U jest unitarna, to dla każdej macierzy A mamy równoważność: A jest normalna $\iff$ UAU^* jest normalna.*

Twierdzenie A.6 (Twierdzenie spektralne dla endomorfizmów normalnych)**.** *(wersja unitarna)*

- $F \in \mathrm{End}(V)$ *jest normalny* $\iff$ *diagonalizuje się w bazie ON.*
- $A \in \mathrm{M}_{n\times n}(\mathbf{C})$ *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 przemienne.

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 $N \in \mathrm{M}_{(n-1)\times(n-1)}(\mathbf{C})$ 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 $W = \operatorname{Lin}(b'_2, \dots, b'_n)$ jest F-niezmiennicze, czyli $F\restriction_W \in \operatorname{End}(W)$ jest normalny (bo N jest jego macierzą). Teza wynika wtedy wprost z założenia indukcyjnego. □

Dodatek B

Przestrzenie nieskończenie wymiarowe

B.1 Twierdzenie o wymiarze dla przestrzeni nieskończenie 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 B.1 (Halla o kojarzeniu małżeństw)**.** *Załóżmy 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 $|A_0| \leqslant \bigcup_{a \in A_0} B_a$.*

Wtedy istnieje funkcja różnowartościowa przekształcająca każdy element $a \in A$ na pewien element odpowiadającego mu B_a.[1]

Wniosek B.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 $|A| \leqslant |B|$, tzn. że istnieje funkcja różnowartościowa z $A \to B$: z symetrii wynikać będzie, że też $|B| \leqslant |A|$, a więc z twierdzenia Cantora-Bernsteina $|A| = |B|$.

[1] Twierdzenie 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 $A_0 \subseteq \mathrm{Lin}(B_0)$. Ponieważ A_0 i B_0 są liniowo niezależne, B_0 jest (skończoną) bazą $V = \mathrm{Lin}\, B_0$, a A_0 jego liniowo niezależnym podzbiorem, więc na mocy Twierdzenia 2.13, $|A_0| \leqslant |B_0|$.

Stąd na mocy twierdzenia Halla istnieje funkcja różnowartościowa $A \to \bigcup_{a \in A} B_a \subseteq B$, co należało dowieść. □

Wniosek B.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 $F: V_1 \to V_2$ 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. □

B.2 Lemat Steinitza o wymianie

Dowód twierdzenia o wymiarze (Wniosku B.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 B.4 (Lemat Steinitza o wymianie)**.** *Załóżmy 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 pozdbiorów $B_0 \subseteq B$ takich że $B_0 \cap A = \emptyset$ i $B_0 \cup A$ jest liniowo niezależny, uporządkowanych przez $\subseteq$.

Podobnie jak w dowodzie Twierdzenia 2.5, pokazujemy że porządek ten spełnia założenia lematu Kuratowskiego-Zorna, ma więc element maksymalny $B' \subseteq B$.

Z maksymalności B' wynika, że $B \subseteq \mathrm{Lin}(A \cup B')$ i oczywiście $A \cup B'$ jest liniowo niezależny, więc jest on bazą. Pozostaje pokazać, że $C = B \setminus B'$ jest równoliczny z A.

Rozważmy $W = \mathrm{Lin}(B'), W_1 = \mathrm{Lin}(A), W_2 = \mathrm{Lin}(C) \leqslant V$ i rozważmy odwzorowanie $F\colon V \to V/W$. 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 $F\restriction_{W_1}$ i $F\restriction_{W_2}$ 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. $\square$

Indeks

Indeks symboli i oznaczeń