Algebra
notatki do przedmiotu

Edycja 2019/20

Spis treści

Część I

Algebra Liniowa

Rozdział 1

Ciała, przestrzenie liniowe, liniowa niezależność, eliminacja Gaußa

1.1 Ciała

Przestrzenie liniowe to uogólnienie $\mathbb{R}^n$. W tym uogólnieniu najpierw chcemy uogólnić samo pojęcie liczb rzeczywistych $\mathbb{R}$, tak, aby obejmowało znane nam naturalne przykłady: $\mathbb{Q}$, $\mathbb{C}$, $\mathbb{Z}_p$ (dla pierwszego p). Takie wspólne uogólnienie to *ciało*, oznaczane ogólnie jako $\mathbb{F}$. Dokładne własności ciał omówimy w odpowiednim momencie, na razie pozostaniemy przy istotnych przykładach.

Przykład 1.1. Ciałami są: liczby rzeczywiste ($\mathbb{R}$), liczby wymierne ($\mathbb{Q}$), liczby zespolone ($\mathbb{C}$), reszty modulo p ($\mathbb{Z}_p$) dla p — liczby pierwszej.

Poza $\mathbb{Z}_p$ działania określamy w naturalny sposób. W $\mathbb{Z}_p$ działania $\cdot_p$ oraz $+_p$ określamy jako:

- $a +_p b = (a + b) \bmod p$
- $a \cdot_p b = (a \cdot b) \bmod p$

gdzie $a \bmod p$ oznacza resztę z dzielenia a przez p. (Dla przypomnienia, b jest resztą z dzielenia $a \in \mathbb{Z}$ przez p, jeśli $0 \leq a < p$ i istnieje liczba $c \in \mathbb{Z}$ taka że $bp + b = a$).

W ciele są dwie operacje: mnożenie "$\cdot$" i dodawanie "$+$", są one przemienne i zachowują się tak, jak intuicyjnie oczekujemy. Są też dwa wyróżnione elementy $0, 1$, które w naszych przykładach pokrywają się z tradycyjnie rozumianymi wyróżnionymi 0 i 1 i mają te same własności, tj. $1 \cdot \alpha = \alpha$ oraz $0 + \alpha = \alpha$.

W ciele przez $-\alpha$ rozumiemy element taki, że $\alpha + (-\alpha) = 0$ a przez α^{-1} dla $\alpha \neq 0$ (pisane też jako $\frac{1}{\alpha}$) taki, że $\alpha \cdot \alpha^{-1} = 1$. W ciałach $\mathbb{R}, \mathbb{Q}, \mathbb{C}$ oba te elementy wyglądają tak, jak się spodziewamy, w $\mathbb{Z}_p$ sytuacja jest trochę bardziej skomplikowana.

1.2 Przestrzenie liniowe

O przestrzeni liniowej chcemy myśleć, iż jest to uogólnienie $\mathbb{R}^n$. O jej elementach nazywamy *wektorami* i myślimy, że są to punkty w $\mathbb{R}^n$, ale traktowane jako wektory, tzn. możemy je dodawać i mnożyć przez elementy z $\mathbb{R}$, jest to mnożenie przez *skalary*.

Definicja 1.2. Zbiór $\mathbb{V}$ jest *przestrzenią liniową nad ciałem* $\mathbb{F}$, jeśli:

1. W $\mathbb{V}$ określone jest dodawanie
$$+ : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$$
2. Dodawanie w $\mathbb{V}$ jest przemienne, tj.:
$$\forall_{u,v \in \mathbb{V}} v + u = u + v$$
3. Dodawanie w $\mathbb{V}$ jest łączne:
$$\forall_{u,v,w \in \mathbb{V}} (u + v) + w = u + (v + w)$$
W związku z tym dodawanie w $\mathbb{V}$ zapisujemy bez nawiasów.

4. W $\mathbb{V}$ istnieje wyróżniony wektor $\vec{0}$:
$$\exists_{\vec{0}\in\mathbb{V}}\forall_{v\in\mathbb{V}}\vec{0}+v=v$$

5. Dla każdego elementu v istnieje element przeciwny $-v$:
$$\forall_{v\in\mathbb{V}}\exists_{-v\in\mathbb{V}}(-v)+v=\vec{0}$$

6. Zdefiniowane jest mnożenie (lewostronne) elementów $\mathbb{V}$ przez elementy z $\mathbb{F}$:
$$\cdot:\mathbb{F}\times\mathbb{V}\to\mathbb{V}$$

7. Zachodzi rozdzielność mnożenia względem dodawania (skalarów):
$$\forall_{\alpha,\beta\in\mathbb{F}}\forall_{v\in\mathbb{V}}(\alpha+\beta)\cdot v=\alpha v+\beta v$$

8. Zachodzi rozdzielność mnożenia względem dodawania (wektorów):
$$\forall_{\alpha\in\mathbb{F}}\forall_{v,u\in\mathbb{V}}\alpha\cdot(v+u)=\alpha v+\alpha u$$

9. Mnożenie jest łączne:
$$\forall_{\alpha,\beta\in\mathbb{F}}\forall_{v\in\mathbb{V}}\alpha\cdot(\beta\cdot v)=(\alpha\beta)\cdot v$$

10. Mnożenie przez „jedynkę” z ciała zachowuje wektor
$$\forall_{v\in\mathbb{V}}1\cdot v=v$$

Elementy $\mathbb{V}$ nazywamy *wektorami*, zaś elementy $\mathbb{F}$: *skalarami*.

Uwaga. Mnożymy tylko przez skalary, wektory możemy tylko dodawać.

Przykład 1.3. 1. $\mathbb{R}^n$, $\mathbb{C}^n$, $\{0\}$, $\mathbb{Q}^n$, $\mathbb{Z}_p^n$, każde nad odpowiednim ciałem: $\mathbb{R}$, $\mathbb{C}$, dowolnym, $\mathbb{Q}$, $\mathbb{Z}_p$.

2. Zbiory funkcji: $\mathbb{R}^\mathbb{R}$, $\mathbb{R}^\mathbb{Q}$, $\mathbb{Q}^\mathbb{R}$, $\mathbb{Z}_p^\mathbb{R}$, $\mathbb{R}^\mathbb{N}$. Zbiory funkcji o skończenie wielu (przeliczalnie wielu) wartościach niezerowych. Ale nie: zbiory funkcji o skończenie wielu wartościach równych 1.

3. $\mathbb{R}$, $\mathbb{C}$ nad $\mathbb{Q}$.

4. Zbiory ciągów o wartościach w $\mathbb{R}$, $\mathbb{Z}_p$, ... (czyli zbiory funkcji $\mathbb{R}^\mathbb{N}$, $\mathbb{Z}_p^\mathbb{N}$, ...)

5. Zbiory wielomianów o współczynnikach z $\mathbb{F}$ nad $\mathbb{F}$. Zbiory wielomianów określonego stopnia. Zbiory wielomianów zerujących się w jakichś punktach.

6. Punkty w $\mathbb{R}^2$ spełniające równanie $2x+y=0$. Punkty w $\mathbb{R}^3$ spełniające równanie $2x+y=0, x-y+3z=0$. Ale nie $2x+y=1, x-y+3z=0$.

Też mają dużo oczekiwanych własności.

Fakt 1.4. *1.* $\forall_{\vec{v}\in\mathbb{V}}0\cdot\vec{v}=\vec{0}$

2. $\forall_{\alpha\in\mathbb{F}}\alpha\cdot\vec{0}=\vec{0}$

3. $\forall_{\vec{v}\in\mathbb{V},\alpha\in\mathbb{F}}\alpha\cdot v=\vec{0}\iff v=\vec{0}\vee\alpha=0$

4. $\forall_{\vec{v}\in\mathbb{V}}(-1)v=-v$

5. wektor przeciwny jest dokładnie jeden

6. wektor zerowy jest dokładnie jeden

7. ...

1.3 Podprzestrzenie liniowe

Definicja 1.5 (Podprzestrzeń liniowa)**.** Dla przestrzeni liniowej $\mathbb{V}$ jej podzbiór $\mathbb{W} \subseteq \mathbb{V}$ jest *podprzestrzenią liniową*, gdy jest przestrzenią liniową nad tym samym ciałem i działania są określone tak, jak w $\mathbb{V}$. Zapisujemy to jako $\mathbb{W} \leq \mathbb{V}$.

Taki zbiór musi być niepusty (ale może zawierać tylko $\vec{0}$).

Przykład 1.6.

1. cała przestrzeń $\mathbb{V}$ jest swoją podprzestrzenią;
2. $\{\vec{0}\}$ jest podprzestrzenią;
3. w $\mathbb{R}^n$ zbiór wektorów mających 0 na ustalonych współrzędnych;
4. w $\mathbb{R}^n$ zbiór wektorów których o sumie współrzędnych równej 0;
5. dla zbioru wszystkich wielomianów o współczynnikach z $\mathbb{F}$, zbiór wielomianów o stopniu najwyżej k;
6. dla zbioru wszystkich wielomianów o współczynnikach z $\mathbb{F}$, zbiór wielomianów przyjmujących wartość 0 w ustalonym zbiorze punktów;
7. w $\mathbb{R}^n$ zbiór wektorów spełniających równania $x_1 + 2x_2 = 0$ i $x_3 - x_2 = 0$.

Lemat 1.7. *Niepusty podzbiór przestrzeni liniowej jest podprzestrzenią wtedy i tylko wtedy gdy jest zamknięty na dodawanie i mnożenie przez skalary.*

Dowód. Podprzestrzeń liniowa jest niepusta, zamknięta na dodawanie i mnożenie przez skalary.

Załóżmy, że $\emptyset \neq U \subseteq \mathbb{V}$ jest zamknięta na dodawanie i mnożenie przez skalary. Chcemy pokazać, że jest przestrzenią liniową; w oczywisty sposób zawiera się w $\mathbb{V}$.

Dodawanie i mnożenie w U określamy tak jak w $\mathbb{V}$. Ze względu na zamkniętość na dodawanie i mnożenie, jest to dobra definicja.

Dla każdego elementu istnieje przeciwny: wystarczy pomnożyć przez -1.

Wektor zerowy jest w U: otrzymujemy go jako sumę $v + (-v)$ (tu korzystamy z tego, że U jest niepusty); alternatywnie jako $0 \cdot v$ dla dowolnego v, ponownie korzystamy z niepustości.

Wszystkie pozostałe własności (łączność, przemienność) itp. są równościami pomiędzy pewnymi elementami U (to są elementy U, bo jest ono zamknięte na mnożenie i dodawanie). Ale te równości zachodzą w $\mathbb{V}$, a działania w U są takie same, jak w $\mathbb{V}$, czyli zachodzą też w U. □

Podprzestrzenie liniowe można generować używając pewnych standardowych operacji: przecięcia, sumy, iloczynu kartezjańskiego.

Definicja 1.8 (Suma, przecięcie, iloczyn kartezjański przestrzeni liniowych)**.** Niech $\mathbb{W}, \mathbb{W}' \leq \mathbb{V}$. Wtedy ich *suma* to

$$\mathbb{W} + \mathbb{W}' = \{w + w' \, : \, w \in \mathbb{W}, \, w' \in \mathbb{W}'\}.$$

Dla dowolnego zbioru podprzestrzeni liniowych $\{\mathbb{W}_i\}_{i \in I}$, gdzie $\mathbb{W}_i \leq \mathbb{V}$ dla każdego $i \in I$, przecięcie zdefiniowane jest naturalnie jako $\bigcap_{i \in I} \mathbb{W}_i$ (jako zbiór).

Dla dowolnego zbioru przestrzeni liniowych $\{\mathbb{V}_i\}_{i \in I}$ nad tym samym ciałem produkt kartezjański $\prod_{i \in I} \mathbb{V}_i$ zdefiniowany jest naturalnie. Działania zdefiniowane są po współrzędnych.

Lemat 1.9. *Suma, przecięcie oraz iloczyn kartezjański przestrzeni liniowych jest przestrzenią liniową.*

Suma przestrzeni liniowych $\mathbb{W} + \mathbb{W}'$ jest najmniejszą przestrzenią liniową zawierająca jednocześnie $\mathbb{W}$ i $\mathbb{W}'$.

Przekrój przestrzeni liniowych $\bigcap_i \mathbb{W}_i$ jest największą przestrzenią liniową zawartą jednocześnie we wszystkich podprzestrzeniach $\mathbb{W}_i$.

Dowód pozostawiony jest jako ćwiczenie.

1.4 Kombinacje liniowe wektorów

W przestrzeniach liniowych możemy w zwarty sposób reprezentować zbiory poprzez sumy.

Definicja 1.10 (Kombinacja liniowa). Dla wektorów $v_1, v_2, \ldots, v_k$ ich *kombinacja liniowa* to dowolny wektor postaci $\sum_{i=1}^{k} \alpha_i v_i$, gdzie $\alpha_1, \ldots, \alpha_k$ jest ciągiem skalarów z ciała $\mathbb{F}$.

Kombinacja liniowa jest *z definicji* skończona.

Przykład 1.11.
- W przestrzeni liniowej $\mathbb{R}^2$ prosta przechodząca przez $(0,0)$ i $(1,1)$ to kombinacja wektora $[1,1]$.
- Prosta przechodzące przez punkty $(1,1)$ i $(2,3)$ to kombinacja postaci $\alpha[1,1]+(1-\alpha)[2,3] = [1,1]+(1-\alpha)\cdot[1,2]$ dla $\alpha \in [0,1]$.
- Odcinek między $(1,1)$ a $(2,3)$ to ograniczona kombinacja postaci $\alpha[1,1]+(1-\alpha)[2,3]$ dla $\alpha \in [0,1]$.
- Równoległobok o wierzchołkach w punktach v_1, v_2, v_3, v_4, spełniających warunki $v_1 + v_4 = v_2 + v_3$ to zbiór punktów spełniających $v_1 + \alpha(v_2 - v_1) + \beta(v_3 - v_1)$ dla $\alpha, \beta \in [0,1]$. Obwód tego równoległościanu spełnia dodatkowo warunek, że przynajmniej jedna z liczb α, β należy do zbioru $\{0,1\}$.

Definicja 1.12. Niech $\mathbb{V}$ będzie przestrzenią liniową nad ciałem $\mathbb{F}$. Dla dowolnego zbioru wektorów (skończonego lub nie) $U \subseteq \mathbb{V}$ jego *otoczka liniowa*, oznaczana jako $\mathrm{LIN}(U)$, to zbiór kombinacji liniowych wektorów ze zbioru U:

$$\mathrm{LIN}(U) = \left\{ \sum_{i=1}^{k} \alpha_i \vec{v}_i \mid k \in \mathbb{N},\, \alpha_1, \ldots, \alpha_k \in \mathbb{F},\, \vec{v}_1, \ldots, \vec{v}_k \in \mathbb{V} \right\}. \tag{1.1}$$

$\mathrm{LIN}(U)$ nazywane jest też *podprzestrzenią rozpiętą* przez U lub *domknięciem liniowym* U.

Przykład 1.13. Dla zbioru wszystkich ciągów nieskończonych o wartościach z $\mathbb{R}$, niech e_i to ciąg mający na i-tym miejscu 1 i mający 0 na pozostałych pozycjach. Wtedy $\mathrm{LIN}(\{e_i\}_{i\in\mathbb{N}})$ to zbiór ciągów o skończenie wielu niezerowych współrzędnych.

Dla prostoty zapisu, nie zakładamy, że wektory $v_1, \ldots, v_k$ są różne, ale jeśli to wygodne, to bez zmniejszenia ogólności możemy to założyć. Dla układu wektorów $v_1, \ldots, v_k$ będziemy czasami pisać $\mathrm{LIN}(v_1, \ldots, v_k)$ na oznaczenie $\mathrm{LIN}(\{v_1, \ldots, v_k\})$.

Fakt 1.14. *Dla dowolnego zbioru wektorów $U \subseteq \mathbb{V}$ w przestrzeni liniowej $\mathbb{V}$ otoczka liniowa $\mathrm{LIN}(U)$ jest podprzestrzenią liniową $\mathbb{V}$. Jest to najmniejsza przestrzeń liniowa zawierającą U.*

Dowód. Skoro $U \subseteq \mathbb{V}$, to skoro $\mathbb{V}$ jest zamknięta na kombinacje liniowe, to również $\mathrm{LIN}(U) \subseteq \mathbb{V}$.

Sprawdźmy, że $\mathrm{LIN}(U)$ jest przestrzenią liniową: pokażemy, że jest zamknięta na dodawanie i mnożenie.

Jeśli $v, v' \in \mathrm{LIN}(U)$ to $v = \sum_{i=1}^{k} \alpha_i v_i$ i $v' = \sum_{i=k+1}^{\ell} \alpha_i v_i$ i tym samym $v + v' = \sum_{i=1}^{\ell} \alpha_i v_i$ oraz $\alpha v = \sum_{i=1}^{k} (\alpha\alpha_i) v_i$

Z drugiej strony, każda przestrzeń liniowa $\mathbb{W} \supseteq U$ jest zamknięta na dodawanie wektorów i mnożenie przez skalary, łatwo więc pokazać przez indukcję (po k w (1.1)), że musi zawierać też wszystkie elementy z $\mathrm{LIN}(U)$. □

Fakt 1.15. *Jeśli $U \subseteq U' \subseteq \mathbb{V}$, gdzie $\mathbb{V}$ jest przestrzenią liniową, to $\mathrm{LIN}(U) \leq \mathrm{LIN}(U')$.*

Dowód. Skoro $U \subseteq U'$ to każda kombinacja z U jest też kombinacją z U', czyli $\mathrm{LIN}(U) \subseteq \mathrm{LIN}(U')$, ale skoro obie są podprzestrzeniami liniowymi $\mathbb{V}$, to dostajemy tezę. □

Lemat 1.16. *Niech $\mathbb{V}$ będzie przestrzenią liniową, $U \subseteq \mathbb{V}$ układam wektorów. Wtedy:*

$$\mathrm{LIN}(U) = \mathrm{LIN}(\mathrm{LIN}(U)) \ .$$

Jeśli $U \subseteq U' \subseteq \mathrm{LIN}(U)$ to

$$\mathrm{LIN}(U') = \mathrm{LIN}(U) \ .$$

Dowód. Zauważmy, że z Faktu 1.14 wiemy, że $\mathrm{LIN}(\mathrm{LIN}(U))$ jest najmniejszą przestrzenią liniową zawierającą $\mathrm{LIN}(U)$. Ale $\mathrm{LIN}(U)$ jest przestrzenią liniową, czyli $\mathrm{LIN}(\mathrm{LIN}(U)) = \mathrm{LIN}(U)$.

Co do drugiego punktu, z Faktu 1.15 mamy:

$$\mathrm{LIN}(U) \leq \mathrm{LIN}(U') \leq \mathrm{LIN}(\mathrm{LIN}(U)) = \mathrm{LIN}(U) \ . \quad \square$$

Otoczka liniowa jest niezmiennicza na kombinacje liniowe.

Lemat 1.17. *Niech $\mathbb{V}$ będzie przestrzenią liniową nad ciałem $\mathbb{F}$, zaś $v_1, \ldots, v_k \in \mathbb{V}$ wektorami z tego ciała. Jeśli skalary $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ są niezerowe to*

$$\mathrm{LIN}(v_1, \ldots, v_k) = \mathrm{LIN}(\alpha_1 v_1, \ldots, \alpha_k v_k) \ .$$

Dla $i \neq j$ oraz skalara $\alpha \in \mathbb{F}$

$$\mathrm{LIN}(v_1, \ldots, v_k) = \mathrm{LIN}(v_1, \ldots, v_{i-1}, v_i + \alpha v_j, v_{i+1}, \ldots, v_k) \ .$$

Dowód. Dowód przy użyciu Lematu 1.16: niech $U_1 = (v_1, \ldots, v_k)$, $U_2 = (\alpha_1 v_1, \ldots, \alpha_k v_k)$, oraz $U_3 = U_1 \cup U_2$. Wtedy $U_1 \subseteq U_3 \subseteq \mathrm{LIN}(U_1)$, czyli $\mathrm{LIN}(U_1) = \mathrm{LIN}(U_3)$. Analogicznie $U_2 \subseteq U_3 \subseteq \mathrm{LIN}(U_2)$, co daje $\mathrm{LIN}(U_2) = \mathrm{LIN}(U_3)$.

Niech teraz $U_4 = (v_1, \ldots, v_{i-1}, v_i + \alpha v_j, v_{i+1}, \ldots, v_k)$ oraz $U_5 = U_1 \cup \{v_i + \alpha v_j\}$. Analogicznie, $U_1 \subseteq U_5 \subseteq \mathrm{LIN}(U_1)$ oraz $U_4 \subseteq U_5 \subseteq \mathrm{LIN}(U_4)$ co daje $\mathrm{LIN}(U_1) = \mathrm{LIN}(U_5) = \mathrm{LIN}(U_4)$. $\square$

Lemat 1.18. *Niech $\mathbb{V}$: przestrzeń liniowa nad ciałem $\mathbb{F}$, $\{v_1, v_2 \ldots, v_k\} \subseteq \mathbb{V}$: zbiór wektorów z $\mathbb{V}$, zaś $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$: ciąg skalarów, gdzie $\alpha_1 \neq 0$. Wtedy*

$$\mathrm{LIN}\left(\left\{\sum_{i=1}^{k} \alpha_i v_i, v_2 \ldots, v_k\right\}\right) = \mathrm{LIN}\left(\{v_1, v_2 \ldots, v_k\}\right). \tag{1.2}$$

Dowód pozostawiamy jako ćwiczenie.

1.5 Liniowa niezależność wektorów.

Definicja 1.19. Układ wektorów U jest *liniowo niezależny* gdy dla dowolnego $k \geq 0$, dowolnych różnych $v_1, \ldots, v_k \in U$ oraz ciągu współczynników $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$

$$\sum_{i=1}^{k} \alpha_i \cdot v_i = \vec{0}$$

implikuje

$$\alpha_1 = \alpha_2 = \cdots = \alpha_k = \vec{0} \ .$$

Uwaga. Uwaga, U traktujemy jako multizbiór: jeśli zawiera jakiś element m razy, to można go m razy użyć. W takim przypadku U jest liniowo zależny, bo $v + (-1) \cdot v = \vec{0}$.

Fakt 1.20. *Niech $\mathbb{V}$ będzie przestrzenią liniową. Układ wektorów $U \subseteq \mathbb{V}$ jest liniowo zależny wtedy i tylko wtedy jeden z nich można przedstawić jako liniową kombinację pozostałych.*

Równoważne sformułowanie: Układ wektorów $U \subseteq \mathbb{V}$ jest liniowo zależny wtedy i tylko wtedy, gdy istnieje $u \in U$ taki że $u \in \mathrm{LIN}(U \setminus \{u\})$.

Dowód. Jeśli układ jest liniowo zależny, to istnieje niezerowa kombinacja $\sum_i \alpha_i u_i = 0$. Bez zmniejszenia ogólności, niech $\alpha_1 \neq 0$. Wtedy $v_1 = \sum_{i>1} -\frac{\alpha_i}{\alpha_1} v_i$ i jest żądane przedstawienie.

Jeśli $u_1 = \sum_{i>1} \alpha_i u_i$ to $\sum_i \alpha_i u_i = \vec{0}$ dla $\alpha_1 = -1$. $\square$

Fakt 1.21. *Niech $\mathbb{V}$ będzie przestrzenią liniową. Układ wektorów $U \subseteq \mathbb{V}$ jest liniowo zależny wtedy i tylko wtedy, gdy istnieje $u \in U$ taki że*

$$\mathrm{LIN}(U) = \mathrm{LIN}(U \setminus \{u\}).$$

Jeśli U nie zawiera $\vec{0}$, to są przynajmniej dwa takie wektory.

(Uwaga: traktujemy U jako multizbiór, tzn. jeśli zawiera dwa razy ten sam wektor, to wyborem u mogą być dwie różne „kopie” tego samego wektora.)

Prosty dowód pozostawiamy jako ćwiczenie.

Uogólnijmy Lemat 1.18.

Lemat 1.22 (Porównaj Lemat 1.18). *Niech $U = (v_1, \ldots, v_k)$ będzie układem wektorów, rozpatrzmy układy*

$$\begin{aligned} U' &= (v_1, \ldots, v_{i-1}, \alpha v_i, v_{i+1}, \ldots, v_k) && \text{dla } \alpha \neq 0, 1 \leq i \leq k \\ U'' &= (v_1, \ldots, v_{i-1}, v_i + \alpha v_j, v_{i+1}, \ldots, v_k) && \text{dla } i \neq j. \end{aligned}$$

Wtedy U jest liniowo zależny wtedy i tylko wtedy gdy U' jest liniowo zależny, wtedy i tylko wtedy gdy U'' jest liniowo zależny.

Dowód. Skorzystamy z Faktu 1.20: załóżmy, że U jest liniowo zależne, niech $u_k \in \mathrm{LIN}(U \setminus \{u_k\})$. Jeśli $k \neq i$ to zauważmy, że $\mathrm{LIN}(U \setminus \{u_k\}) = \mathrm{LIN}(U' \setminus \{u_k\})$ (z Lematu 1.17) i w takim razie U' jest liniowo zależny. Jeśli $k = i$, to skoro $u_i \in \mathrm{LIN}(U \setminus \{u_i\}) = \mathrm{LIN}(U' \setminus \{\alpha u_i\})$ to oczywiście również $\alpha u_i \in \mathrm{LIN}(U' \setminus \{\alpha u_i\})$.

Dowód dla U'' jest analogiczny: jeśli $k \neq i$ to argument jest taki sam. Jeśli $k = i$ to korzystamy z tego, że $u_i \in \mathrm{LIN}(U'' \setminus \{u_i + \alpha u_k\})$ oraz $u_k \in \mathrm{LIN}(U'' \setminus \{u_i + \alpha u_k\})$ i w takim razie $u_i + \alpha u_k \in \mathrm{LIN}(U'' \setminus \{u_i + \alpha u_k\})$ i U'' jest liniowo zależny.

Implikacje w drugą stroną wynikają z symetrii. □

1.6 Metoda eliminacja Gaußa.

Chcemy mieć usystematyzowany sposób znajdowania dla (skończonego) zbioru wektorów $U \subseteq \mathbb{R}^n$ jego maksymalnego (względem zawierania) podzbioru niezależnego.

Chcemy uogólnić następujące obserwacje:

- jeśli każdy wektor ma współrzędną, na której tylko on jest niezerowy, to zbiór jest liniowo niezależny;
- układ wektorów zawierający $\vec{0}$ nie jest niezależny;
- używając Lematu 1.22 możemy „upraszczać wektory”.

Przykład 1.23.

$$\begin{bmatrix} 1 & 6 & 5 & 5 & 3 \\ 1 & 2 & 3 & 2 & 2 \\ 3 & 4 & 5 & 3 & 3 \\ 2 & 1 & 3 & 1 & 2 \end{bmatrix} \xrightarrow{(3)-(2)-(4)} \begin{bmatrix} 1 & 6 & 5 & 5 & 3 \\ 1 & 2 & 3 & 2 & 2 \\ 0 & 1 & -1 & 0 & -1 \\ 2 & 1 & 3 & 1 & 2 \end{bmatrix} \xrightarrow{(1)-(2),(4)-2\cdot(2)} \begin{bmatrix} 0 & 4 & 2 & 3 & 1 \\ 1 & 2 & 3 & 2 & 2 \\ 0 & 1 & -1 & 0 & -1 \\ 0 & -3 & -3 & -3 & -2 \end{bmatrix}$$

$$\xrightarrow{(1)-(3)+(4)} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 2 & 2 \\ 0 & 1 & -1 & 0 & -1 \\ 0 & -3 & -3 & -3 & -2 \end{bmatrix} \xrightarrow{(4)+3\cdot(3)} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 2 & 2 \\ 0 & 1 & -1 & 0 & -1 \\ 0 & 0 & -6 & -3 & -5 \end{bmatrix}$$

Czyli wejściowy układ wektorów był liniowo zależny. Jednocześnie układ wektorów bez pierwszego danego jest liniowo niezależny, co pokazujemy przy użyciu analogicznych rachunków.

Sformalizujmy postać wektorów, do której w ten sposób dojdziemy.

Definicja 1.24 (Postać schodkowa). Układ wektorów $v_1, \ldots, v_m \in \mathbb{F}^n$ jest w *postaci schodkowej*, jeśli istnieje ciąg pozycji $0 = i_0 < i_1 < i_1 < \cdots < i_m$ takich że dla każdego $j = 1, \ldots, m$:

- wektor v_j ma na pozycji i_j element niezerowy
- wektor v_j ma na pozycjach $< i_j$ same 0.

Lemat 1.25. *Jeśli układ wektorów w $\mathbb{F}^n$ jest w postaci schodkowej, to jest niezależny.*

Dowód. Niech te wektory to $v_1, \ldots, v_k$ a ich pozycje z definicji postaci schodkowej to $i_1, \ldots, i_m$. Rozważmy współczynniki $\alpha_1, \ldots, \alpha_k$ takie że $\sum_{j=1}^{k} \alpha_j v_j = \vec{0}$.

Niech α_j to najmniejszy niezerowy współczynnik w tej kombinacji liniowej. Wtedy liczba otrzymana na pozycji i_j jest niezerowa: wektory $v_1, \ldots, v_{j-1}$ są brane ze współczynnikami 0, wektory $v_{j+1}, \ldots, v_k$ mają na pozycji i_j same 0, czyli współczynnik na pozycji i_j w sumie $\sum_{\ell=1}^{k} \alpha_\ell v_\ell$ to α_j razy wartość w $(v_i)_j \neq 0$. Sprzeczność. □

W ogólności chcemy przekształcić dowolny układ wektorów używając operacji jak w Lemacie 1.22 do zbioru wektorów w postaci schodkowej i wektorów $\vec{0}$. Jeśli tych drugich nie ma, to wejściowy zbiór był niezależny, jeśli są, to był zależny.

Pokażemy teraz, że używając takich operacji zawsze można sprowadzić układ do postaci schodkowej.

W każdym kroku metody utrzymujemy dwa zbiory wektorów: U oraz U' oraz pozycję j. Początkowo U jest całym zbiorem wektorów, U' jest pusty, zaś $j = 0$. Jako niezmiennik utrzymujemy następujące własności:

- U' jest w postaci schodkowej oraz indeksy odpowiednich niezerowych pozycji są nie większe niż j
- wektory w U mają na pozycjach nie większych niż j same 0.

W każdym kroku wybieramy pozycję j' oraz wektor $v \in U$ takie że:

- $j' > j$ i j' jest najmniejsze, takie że któryś z wektorów z U ma niezerową współrzędną j'
- $v \in U$ oraz ma niezerową współrzędną j'

Dodajemy v do U', wybieramy j' jako nowe j.

Niech $(v)_{j'} = \alpha$. Dla każdego $v' \in U \setminus \{v\}$: Niech $(v')_{j'} = \alpha'$. Zastępujemy v' przez $v' - \frac{\alpha'}{\alpha} v$.

Łatwo pokazać, że po tym wyborze niezmienniki są zachowane.

Lemat 1.26. *Po zakończeniu otrzymujemy układ złożony z wektorów liniowo niezależnych oraz samych wektorów zerowych.*

Dowód. Skoro nie możemy kontynuować, to albo

- U jest pusty. Wtedy U' jest w postaci schodkowej, czyli z Lematu 1.25 jest liniowo niezależny, ma tyle samo wektorów, co zbiór wejściowy i z Lematu 1.22 wejściowy układ był liniowo niezależny.
- U jest niepusty, ale nie da się wybrać j'. Czyli wszystkie wektory w U mają zerowe współrzędne dla $j' > j$. Z założenia mają też zerowe współrzędne dla $j' \leq j$, czyli U zawiera same wektory $\vec{0}$. Czyli $U' \cup U$ jest liniowo zależny i z Lematu 1.22 również układ wejściowy jest liniowo zależny. □.

Fakt 1.27. *Jeśli w czasie eliminacji Gaußa używaliśmy do eliminowania jedynie wektorów $\vec{v}_1, \ldots, \vec{v}_n$, które na końcu są niezerowe, to odpowiadające im wektory początkowe tworzą bazę przestrzeni rozpiętej przez wszystkie wektory.*

W szczególności, oryginalny zbiór był niezależny wtedy i tylko wtedy gdy nie otrzymaliśmy żadnego wektora $\vec{0}$.

Dowód. Zauważmy, że są niezależne, bo gdy przeprowadzimy na nich eliminację Gaußa to uzyskamy te same wektory, co poprzednio, czyli niezerowe.

Jeśli na końcu nie ma wektora $\vec{0}$, to wszystkie początkowe wektory były niezależne. □

Rozdział 2

Baza przestrzeni liniowej, wymiar

2.1 Baza przestrzeni liniowej

Chcemy minimalny zbiór niezależny: bo po co więcej (i ma wiele innych, dobrych własności).

Definicja 2.1 (Baza)**.** B jest *bazą* przestrzeni liniowej $\mathbb{V}$ gdy $\mathrm{LIN}(B) = \mathbb{V}$ oraz B jest liniowo niezależny.

Alternatywnie, mówimy, że B jest *minimalnym zbiorem rozpinającym* $\mathbb{V}$.

Przykład 2.2.

- W przestrzeni $\mathbb{F}^n$ wektory (tzw. *baza standardowa*): $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0 \dots, 0)$, $\dots$, $e_{n-1} = (0, \dots, 0, 1, 0)$ $e_n = (0, \dots, 0, 1)$.
- W przestrzeni wielomianów stopnia $\leq n$: wielomiany $\{x^i\}_{i=0}^n$.
- W przestrzeni ciągów o wyrazach w $\mathbb{F}$, które mają skończenie wiele niezerowych wyrazów: $\{e_i\}$, gdzie e_i ma 1 na i-tej pozycji i 0 wszędzie indziej.

  Ta baza jest nieskończona.

2.1.1 Baza standardowa

Gdy pracujemy w $\mathbb{F}^n$ to jedna baza jest lepsza, niż inne: *baza standardowa*, składająca się wektorów $\vec{E_i} = (0, \dots, 0, \underbrace{1}_{i\text{-te miejsce}}, 0 \dots, 0)$.

Przykład 2.3. Rozważmy bazę $B = \{(1, 1, 1), (0, 1, 1), (0, 0, 1)\}$ przestrzeni $\mathbb{R}^3$; niech $\vec{E_1}, \vec{E_2}, \vec{E_3}$ będą wektorami bazy standardowej. Wtedy $(\vec{E_1})_B = (1, -1, 0)$, $(\vec{E_2})_B = (0, 1, -1)$ i $(\vec{E_3})_B = (0, 0, 1)$. Używając tej reprezentacji łatwo pokazać, np. że dla $v = (7, 4, 2)$ mamy $(v)_B = (7, -3, -2)$, bo

$$(v)_B = (7\vec{E_1} + 4\vec{E_2} + 2\vec{E_3})_B = 7(\vec{E_1})_B + 4(\vec{E_2})_B + 2(\vec{E_3})_B.$$

Bardziej interesują nas przestrzenie, które mają skończoną bazę. Prawie wszystko, co powiemy, jest też prawdą ogólnie, ale dowody są dużo bardziej techniczne.

Definicja 2.4 (Przestrzeń skończenie wymiarowa)**.** Przestrzeń jest *skończenie wymiarowa*, jeśli ma skończony zbiór rozpinający.

2.2 Wyrażanie wektora w bazie

Twierdzenie 2.5. *Każdy wektor ma jednoznaczne przedstawienie w bazie*

Dowód. Jeśli $\sum_{i=1}^k \alpha_i v_i$ oraz $\sum_{i=1}^k \beta_i v_i$ to dwa przedstawienia, to $\sum_{i=1}^k (\alpha_i - \beta_i) v_i = \vec{0}$ jest nietrywialną kombinacją dla wektora $\vec{0}$, co przeczy założeniu, że $\{v_1, v_2, \dots, v_k\}$ jest bazą. $\square$

Skoro każdy wektor można naturalnie wyrazić w bazie, to możemy uogólnić notację wektorową dla $\mathbb{F}^n$ na dowolne przestrzenie i bazy

Definicja 2.6 (Wyrażanie wektora w bazie)**.** Jeśli $B = \{v_1, v_2, \ldots, v_n\}$ jest bazą przestrzeni liniowej $\mathbb{V}$ oraz $v \in \mathbb{V}$ jest wektorem, to

$$(v)_B = (\alpha_1, \alpha_2, \ldots, \alpha_n)$$

gdzie $v = \sum_{i=1}^{n} \alpha_i v_i$. Liczby α_i to *współrzędne* wektora v w bazie B

Zauważmy, że po wyrażeniu wektorów $v_1, \ldots, v_n$ w ustalonej bazie B możemy traktować je podobnie jak wektory z $\mathbb{F}^n$. W pewnym sensie to jest „dokładne" odwzorowanie.

Definicja 2.7 (Izomorfizm przestrzeni liniowych)**.** Mówimy, że dwie przestrzenie $\mathbb{V}, \mathbb{W}$ nad ciałem $\mathbb{F}$ są *izomorficzne*, jeśli istnieją bijekcje $\varphi : \mathbb{V} \to \mathbb{W}$ oraz $\psi : \mathbb{W} \to \mathbb{V}$, takie że $\varphi(v +_{\mathbb{V}} v') = \varphi(v) +_{\mathbb{W}} \varphi(v')$ oraz $\varphi(\alpha \cdot_{\mathbb{V}} v) = \alpha \cdot_{\mathbb{W}} \varphi(v)$ i analogicznie dla ψ

Przykład 2.8.
- Przestrzeń wielomianów (o współczynnikach z $\mathbb{F}$) stopnia nie większego niż k oraz $\mathbb{F}^{k+1}$
- Przestrzeń wielomianów (o współczynnikach z $\mathbb{F}$) oraz przestrzeń $\{f \in \mathbb{F}^{\mathbb{N}} : |\{i \in \mathbb{N} : f(i) \neq 0\}|$ jest skończone$\}$ ciągów o wartościach w $\mathbb{F}$, takich że jedynie skończona liczba elementów ciągu jest niezerowa

Fakt 2.9. *Niech $\varphi : V \to W$ będzie izomorfizmem. Wtedy układ $\{v_1, \ldots, v_n\}$ jest liniowo niezależny wtedy i tylko wtedy, gdy układ $\{\varphi(v_1), \ldots, \varphi(v_n)\}$ jest liniowo niezależny.*

Twierdzenie 2.10. *Niech $\mathbb{V}$ nad $\mathbb{F}$ ma bazę n elementową. Wtedy $\mathbb{V}$ jest izomorficzna z $\mathbb{F}^n$.*

Dowolne dwie przestrzenie liniowe nad $\mathbb{F}$ mające bazy n elementowe są izomorficzne.

Dowód. Weźmy dowolną bazę $\mathbb{V}$. Wtedy wyrażenie $(v)_B$wektora v w bazie B jest takim izomorfizmem.

Co do drugiego punktu, to obie są izomorficzne z $\mathbb{F}^n$ i łatwo sprawdzić, że relacja bycia izomorficznymi przestrzeniami liniowymi jest relacją równoważności. □

Tak więc mając dowolny układ wektorów możemy wyrazić je w dowolnej bazie i zastosować na nich eliminację Gaußa.

Można w ten sposób udowodnić np. Twierdzenie 2.11: d-d na ćwiczeniach.

Naszym celem jest pokazanie, że rozmiar bazy nie zależy od wyboru bazy, lecz jest własnością przestrzeni liniowej.

Twierdzenie 2.11. *Każda przestrzeń (skończenie wymiarowa) $\mathbb{V}$ ma bazę.*

Każda baza przestrzeni (skończenie wymiarowej) $\mathbb{V}$ ma taką samą moc.

Dowód dla zainteresowanych, nie przedstawiany na wykładzie, nie wymagany. W skrócie polega on na rozważeniu dwóch baz różnej mocy i iteracyjnym przekształceniu jednej w drugą przy użyciu Lematu Steinitza.

Lemat 2.12 (Lemat Steinitza o wymianie)**.** *Niech $\mathbb{V}$ będzie przestrzenią liniową, $A \subsetneq \mathbb{V}$ liniowo niezależnym zbiorem wektorów, zaś B zbiorem rozpinającym $\mathbb{V}$. Wtedy albo A jest bazą, albo istnieje $v \in B$ taki że $A \cup \{v_i\}$ jest liniowo niezależny.*

Dowód. Rozważmy, czy dla każdego $v \in B$ mamy $v \in \mathrm{LIN}(A)$.

Tak Z Lematu 1.16 mamy

$$\mathrm{LIN}(A) = \mathrm{LIN}(B \cup A) \geq \mathrm{LIN}(B) = \mathbb{V}.$$

Czyli A jest bazą.

Nie Istnieje $v \in B$, taki że $\mathrm{LIN}(A \cup \{v\}) \neq \mathrm{LIN}(A)$. Załóżmy nie wprost, że $A \cup \{v\}$ jest liniowo zależny. Wtedy istnieje kombinacja liniowa

$$\sum_j \alpha_j u_j + \alpha v = 0$$

w której nie wszystkie współczynniki są zerowe, zaś $u_1, u_2 \ldots \in A$. Jeśli $\alpha \neq 0$ to to pokazuje, że $v \in \mathrm{LIN}(A)$, co nie jest prawdą. Jeśli $\alpha = 0$ to otrzymujemy, że A jest liniowo zależny, co z założenia nie jest prawdą, sprzeczność. □

dowód Twierdzenia 2.11. Punkt pierwszy wynika wprost z definicji przestrzeni skończenie wymiarowej i indukcji względem Lematu 2.12: rozpoczynamy ze zbiorem $B = \emptyset$ i dodajemy do niego kolejne wektory ze skończonego zbioru generującego $\mathbb{V}$, dbając, by był liniowo niezależny.

Niech $B_v = \{v_1, v_2, \dots, v_k\}$ oraz $B_u = \{u_1, u_2, \dots, u_\ell\}$ będą dwoma bazami, gdzie $\ell \leq k$. Pokażemy, że $k = l$. W tym celu będziemy zastępować kolejne elementy B_u przez $v_1, v_2, \dots$.

Dokładniej, pokażemy przez indukcję po $j = 0, \dots, \ell$, że istnieje podzbiór $\{v_{i_1}, \dots, v_{i_j}\} \subseteq B_v$ taki że $\{v_{i_1}, \dots, v_{i_j}\} \cup \{u_{j+1}, \dots, u_\ell\}$ jest bazą. Dla $j = \ell$ daje to tezę. Zauważmy, że dla $j = 0$ teza indukcyjna trywialnie zachodzi.

Pokażemy krok indukcyjny. Weźmy $B_j = \{u_{j+1}, \dots, u_\ell\} \cup \{v_{i_1}, \dots, v_{i_j}\}$ i usuńmy z niego u_{j+1}. Ten zbiór jest niezależny, nie jest bazą (bo wtedy B_j nie byłoby liniowe niezależne) $\{v_1, v_2, \dots, v_k\}$ jest bazą, z Lematu 2.12 istnieje $v_{i_{j+1}}$ taki że $B_{j+1} = \{u_{j+2}, \dots, u_\ell\} \cup \{v_{i_1}, \dots, v_{i_{j+1}}\}$ jest liniowo niezależny.

Przypuśćmy, że B_{j+1} nie jest bazą. Wtedy z Lematu 2.12 można go rozszerzyć o wektor z B_j do zbioru niezależnego. Jedynym takim możliwym wektorem jest u_{j+1} (bo pozostałe są już w B_{j+1}). Ale wtedy mamy, że $B_j \cup \{v_{j+1}\}$ jest niezależny, co nie jest możliwe, bo B_j było bazą. □

Wnioski z Lematu Steinitza:

Lemat 2.13. *Każdy zbiór niezależny skończenie wymiarowej przestrzeni liniowej $\mathbb{V}$ można rozszerzyć do bazy.*

Lemat 2.14. *Jeśli $\mathbb{V}$ jest przestrzenia skonczenie wyiarowa, to z każdego układu wektorów $A \subseteq \mathbb{V}$ można wybrać bazę przestrzeni* $\mathrm{LIN}(A)$.

2.3 Wymiar przestrzeni liniowej

Definicja 2.15 (Wymiar przestrzeni liniowej)**.** Dla przestrzeni skończenie wymiarowej $\mathbb{V}$ jej *wymiar* to moc jej bazy. Oznaczamy go jako $\dim(\mathbb{V})$.

Intuicja: to jest „n” w $\mathbb{R}^n$ (lub ogólnie n w $\mathbb{F}^n$).

Wniosek 2.16. Każde dwie przestrzenie liniowe n-wymiarowe nad $\mathbb{F}$ są izomorficzne i są izomorficzne z $\mathbb{F}^n$.

Lemat 2.17. *Jeśli $\mathbb{V}_1, \mathbb{V}_2 \leq \mathbb{V}$ są przestrzeniami skończenie wymiarowymi, to*

$$\dim(\mathbb{V}_1 + \mathbb{V}_2) = \dim(\mathbb{V}_1) + \dim(\mathbb{V}_2) - \dim(\mathbb{V}_1 \cap \mathbb{V}_2).$$

Dowód. Jeśli któraś z tych przestrzeni ma wymiar 0, to równość zachodzi w oczywisty sposób.

Niech B będzie bazą $\mathbb{V}_1 \cap \mathbb{V}_2$ lub puste, jeśli $\mathbb{V}_1 \cap \mathbb{V}_2 = \{\vec{0}\}$.

Rozszerzamy B do baz $\mathbb{V}_1, \mathbb{V}_2$, niech będą one $B \cup B_1$ oraz $B \cup B_2$.

Pokażemy, że $B \cup B_1 \cup B_2$ jest bazą $\mathbb{V}_1 + \mathbb{V}_2$. Zauważmy, że generują one $\mathbb{V}_1 + \mathbb{V}_2$: dla dowolnego $v \in \mathbb{V}_1 + \mathbb{V}_2$ mamy $v = v_1 + v_2$ dla pewnych $v_1 \in \mathbb{V}_1$ oraz $v_2 \in \mathbb{V}_2$. Wtedy $v_1 \in \mathrm{LIN}(B \cup B_1)$ oraz $v_2 \in \mathrm{LIN}(B \cup B_2)$, czyli $v_1, v_2 \in \mathrm{LIN}(B \cup B_1 \cup B_2)$ i w takim razie $v_1 + v_2 \in \mathrm{LIN}(B \cup B_1 \cup B_2)$, bo jest ona zamknięta na sumę wektorów (to jest przestrzeń liniowa).

Pozostało pokazać, że jest to zbiór liniowo niezależny. Rozpatrzmy dowolną kombinację liniową wektorów z $B \cup B_1 \cup B_2$, niech $B = \vec{b}_1, \dots, \vec{b}_n$, $B_1 = \vec{b}_{n+1}, \dots, \vec{b}_{n'}$, $B_2 = \vec{b}_{n'+1}, \dots, \vec{b}_{n''}$. Wtedy taka kombinacja jest postaci

$$\sum_{i=1}^{n''} \alpha_i b_i \quad .$$

Przenieśmy na drugą stronę wektory odpowiadające B_2:

$$\sum_{i=1}^{n'} \alpha_i b_i = \sum_{i=n'+1}^{n''} (-\alpha_i) b_i \quad .$$

Wektor po lewej stronie należy do $\mathbb{V}_1$, ten po prawej do $\mathbb{V}_2$, czyli należą do $\mathbb{V}_1 \cap \mathbb{V}_2$. W takim mają jednoznaczne przedstawienie w bazie B, ono jest takie samo w bazach $B \cup B_1$ oraz $B \cup B_2$, tj. takie przedstawienie w bazie $B \cup B_1$ używa tylko wektorów z B, analogicznie dla $B \cup B_2$. Jednocześnie, wektor po prawej stronie nie używa wektorów z B, czyli jest wektorem zerowym, czyli ma wszystkie współczynniki równe 0. W takim razie ten po lewej również jest $\vec{0}$ i w takim razie ma wszystkie współczynniki równe 0. □

Wzór ten służy głównie do liczenia wymiaru $\mathbb{V}_1 \cap \mathbb{V}_2$:

Fakt 2.18. *Jeśli B_1, B_2 są bazami dla $\mathbb{V}_1, \mathbb{V}_2 \leq \mathbb{V}$ to*

$$\mathbb{V}_1 + \mathbb{V}_2 = \mathrm{LIN}(B_1 \cup B_2)$$

W takim razie znamy $\dim(\mathbb{V}_1), \dim(\mathbb{V}_2)$ i umiemy policzyć moc bazy $\mathbb{V}_1 + \mathbb{V}_2$, czyli znamy wymiar $\mathbb{V}_1 + \mathbb{V}_2$. Czyli umiemy policzyć wymiar $\mathbb{V}_1 \cap \mathbb{V}_2$. (Przykład w kolejnym rozdziale.)

2.4 Zastosowanie eliminacji Gaussa do liczenia wymiaru

Gdy mamy dany zbiór A (skończony), to aby policzyć $\dim(\mathrm{LIN}(A))$ możemy zastosować eliminację Gaussa: wiemy, że po zakończeniu otrzymujemy zbiór wektorów liniowo niezależnych oraz wektory zerowe i generowana przestrzeń jest taka sama. Czyli otrzymany zbiór wektorów liniowo niezależnych to baza a jej liczność to liczba wymiarów przestrzeni.

Twierdzenie 2.19. *Eliminacja Gaussa zastosowana do układu wektorów U zwraca bazę $\mathrm{LIN}(U)$ (oraz wektory zerowe).*

Dowód. Z Lematu 1.26 po zakończeniu mamy zbiór wektorów niezależnych oraz zbiór wektorów zerowych. Czyli niezerowe wektory są bazą rozpinanej przez siebie przestrzeni.

Jednocześnie wiemy, że nie zmienia się rozpinana przez cały układ przestrzeń, czyli są bazą przestrzeni rozpinanej przez wejściowy układ wektorów. □

Fakt 2.20. *Jeśli po zakończeniu eliminacji Gaußa otrzymujemy zbiór złożony z k wektorów, to oryginalny zbiór zawierał dokładnie k wektorów niezależnych.*

Dowód. Komentarz: część z tych rzeczy już wiemy, ale można to prościej pokazać używając pojęcia wymiaru.

Wiemy już, że metoda eliminacji zachowuje przestrzeń rozpiętą przez przechowywany przez nią układ wektorów. W szczególności wymiar (=moc bazy tej przestrzeni) nie zmienia się. Na końcu jest to liczba niezerowych wektorów, na początku: moc maksymalnego (względem zawierania) zbioru wektorów liniowo niezależnych. Jeśli na końcu było jakieś $\vec{0}$ to początkowy zbiór miał mniejszy wymiar, niż liczba jego wektorów, czyli był liniowo zależny. □

Przykład 2.21. Rozważmy przestrzeni liniowe S, T, zadane jako $S = \mathrm{LIN}(\{(1, 6, 5, 5, 3), (1, 2, 3, 2, 2)\})$ oraz $T = \mathrm{LIN}(\{(3, 4, 5, 3, 3), (2, 1, 3, 1, 2)\})$. Obliczymy $\dim(S + T)$ oraz $\dim(S \cap T)$ i podamy bazę $S + T$.

Łatwo zauważyć, że podany zbiór generatorów S ma dwa wektory niezależne (są różne, a mają taką samą pierwszą współrzędną), podobnie T ma wymiar 2. Będziemy korzystać z zależności:

$$\dim(S + T) = \dim(S) + \dim(T) - \dim(S \cap T)$$

Czyli wystarczy, że policzymy wymiar $S + T$. Suma (mnogościowa) generatorów S oraz T generuje $S + T$, zastosujemy metodę eliminacji Gaussa w celu obliczenia wymiaru; odpowiednie rachunki zostały już przeprowadzone w Przykładzie 1.23.

$$\begin{matrix} 1 & 6 & 5 & 5 & 3 \\ 1 & 2 & 3 & 2 & 2 \\ 3 & 4 & 5 & 3 & 3 \\ 2 & 1 & 3 & 1 & 2 \end{matrix} \xrightarrow{(3)-(2)-(4)} \begin{matrix} 1 & 6 & 5 & 5 & 3 \\ 1 & 2 & 3 & 2 & 2 \\ 0 & 1 & -1 & 0 & -1 \\ 2 & 1 & 3 & 1 & 2 \end{matrix} \xrightarrow{(1)-(2),(4)-2\cdot(2)} \begin{matrix} 0 & 4 & 2 & 3 & 1 \\ 1 & 2 & 3 & 2 & 2 \\ 0 & 1 & -1 & 0 & -1 \\ 0 & -3 & -3 & -3 & -2 \end{matrix}$$

$$\xrightarrow{(1)-(3)+(4)} \begin{matrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 2 & 2 \\ 0 & 1 & -1 & 0 & -1 \\ 0 & -3 & -3 & -3 & -2 \end{matrix} \xrightarrow{(4)+3\cdot(3)} \begin{matrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 2 & 2 \\ 0 & 1 & -1 & 0 & -1 \\ 0 & 0 & -6 & -3 & -5 \end{matrix}$$

Wymiar $\mathrm{LIN}(S + T)$ wynosi więc 3. Tym samym wymiar $\mathrm{LIN}(S) \cap \mathrm{LIN}(T)$ wynosi 1.

Co do bazy $S + T$ zauważmy, że wektory uzyskane przez kombinacje liniowe generatorów $S + T$ (czyli naszych wektorów zapisanych w wierszach) dalej należą do $S + T$, tym samym trzy wektory

$$(1,2,3,2,2),(0,1,-1,0,-1),(0,0,-6,-3,-5)$$

są bazą tej przestrzeni.

W eliminacji używaliśmy jedynie wektorów $2,3,4$, tak więc odpowiednie wektory wejścia również są bazą, tj.:

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

są bazą $S+T$.

Przykład/Zastosowanie 2.22 (Rekurencje liniowe). Rekurencja na liczby Fibonacciego.

$$f_n = f_{n-1} + f_{n-2}, \quad f_1 = f_2 = 1.$$

Jak rozwiązać takie równanie (podać postać zwartą). To może za proste, bo wszyscy znają.

Albo na coś podobnego.

$$\begin{aligned} a_n &= a_{n-1} + 2a_{n-2} \qquad (2.1)\\ a_0 &= \alpha, a_1 = \beta \end{aligned}$$

Jeśli zapomnimy o warunkach początkowych, to zbiór ciągów o wartościach w $\mathbb{R}$ oraz spełniających równanie (2.1) tworzy przestrzeń liniową. Nasz ciąg to konkretny wektor w tej przestrzeni liniowej. Widać, że baza jest dwuelementowa (ciąg mający $a_0 = 1, a_1 = 0$ oraz drugi $a_0 = 0, a_1 = 1$). Czyli wystarczy przedstawić nasz ciąg jako kombinację wektorów z bazy.

Nic nie daje: jak wygląda baza?

Szukamy innej, bardziej nam przydatnej bazy. Najlepiej by było, gdyby składała się z ciągów, których elementy możemy jawnie zadać wzorem albo prosto policzyć.

Ciągi arytmetyczne? Nie działa.

Geometryczne? Działa!

$$a^n = a^{n-1} + 2a^{n-2}$$

Czyli szukamy rozwiązań równania (podzielenie przez a^{n-2} jest dopuszczalne, bo $a=0$ odpowiada trywialnemu przypadkowi wektora $\vec{0}$.)

$$x^2 - x - 2 = 0$$

Jeden to $x = 2$, drugi to $x = -1$. Czyli dwa ciągi stanowiące bazę to $(2^n)_{n\geq 1}$ oraz $((-1)^n)_{n\geq 1}$. Tylko trzeba dobrać współczynniki, tj. takie a, b, że

$$\begin{cases} a\cdot(-1)^0 + b\cdot 2^0 & = \alpha\\ a\cdot(-1)^1 + b\cdot 2^1 & = \beta \end{cases} .$$

2.5 Warstwy

Patrząc na $\mathbb{R}^2$ podprzestrzenie liniowe mają prostą i naturalną interpretację: są to dokładnie proste przechodzące przez 0. Niestety, żadna inna prosta nie jest podprzestrzenią liniową, choć ma podobne własności.

Takie proste odpowiadają intuicyjnie warstwom, które są zbiorami powstałymi przez „przesunięcie" podprzestrzeni liniowej o ustalony wektor.

Definicja 2.23 (Warstwa). Dla przestrzeni liniowej $\mathbb{V}$ i jej podprzestrzeni liniowej $\mathbb{W}$ zbiór U jest *warstwą* $\mathbb{W}$ w $\mathbb{V}$, jeśli jest postaci

$$U = u + \mathbb{W} = \{u + w \ : \ w \in \mathbb{W}\} .$$

Zauważ, że warstwy zwykle *nie są* przestrzeniami liniowymi.

Przykład 2.24.

1. Dla podprzestrzeni liniowej $\mathbb{R}^n$ takiej że trzecia współrzędna to 0, warstwami są zbiory wektorów o ustalonej trzeciej współrzędnej.

2. Dla zbioru wektorów spełniających równanie $2x_1 - x_3 = 0$ każda warstwa składa się z wektorów, dla których $2x_1 - x_3$ ma ustaloną wartość.

3. Dla przestrzeni liniowej wielomianów i podprzestrzeni składającej się z wielomianów zerujących się w 2 i 4, warstwy składają się z wektorów o ustalonej wartości w 2 i 4.

Lemat 2.25. *Niech $\mathbb{W} \leq \mathbb{V}$będą przestrzeniami liniowymi, zaś $U \subseteq \mathbb{V}$. Następujące warunki są równoważne:*

1. *istnieje wektor $\vec{u} \in \mathbb{V}$, taki że $U = \vec{u} + \mathbb{W}$*
2. *istnieje wektor $\vec{u} \in U$, taki że $U = \vec{u} + \mathbb{W}$*
3. *dla każdego wektora $\vec{u} \in U$ zachodzi $U = \vec{u} + \mathbb{W}$.*

Ponadto, następujące warunki są równoważne:

1. *istnieje wektor $\vec{u} \in \mathbb{V}$, taki że $U - \vec{u}$ jest przestrzenią liniową;*
2. *istnieje wektor $\vec{u} \in U$, taki że $U - \vec{u}$ jest przestrzenią liniową;*
3. *dla każdego wektora $\vec{u} \in U$ zbiór $U - \vec{u}$ jest przestrzenią liniową.*

Prosty dowód pozostawiamy jako ćwiczenie.

Lemat 2.26 (Wypukłość warstw)**.** *Załóżmy, że ciało $\mathbb{F}$ spełnia $1 + 1 \neq 0$.*
Niech $\mathbb{V}$ będzie przestrzenią liniową nad $\mathbb{F}$, zaś $U \subseteq \mathbb{V}$. Wtedy następujące warunki są równoważne

1. *U jest warstwą (odpowiedniej przestrzeni liniowej)*
2. $\forall_{\alpha \in \mathbb{F}, v, u \in U} \quad \alpha v + (1 - \alpha)u = u + \alpha(v - u) \in U$

Intuicja: na płaszczyźnie to są punkty na prostej wyznaczonej przez u, v.

Dowód. Jeśli U jest warstwą, to jest postaci $u + \mathbb{W}$, dla ustalonego u oraz pewnej przestrzeni liniowej $\mathbb{W}$, w szczególności, jej elementy są postaci $u + v$ dla $v \in \mathbb{W}$. Licząc $\alpha(u+v)+(1-\alpha)(u+v') = u+(\alpha v+(1-\alpha)v')$ i wtedy $\alpha v + (1 - \alpha)v' \in \mathbb{W}$.

W drugą stronę najlepiej przepisać $\alpha v + (1 - \alpha)u = u + \alpha(v - u)$ i tym samym zakładamy że

$$\forall_{u,v \in U, \alpha \in \mathbb{F}} u + \alpha(v - u) \in U. \tag{2.2}$$

Ustalmy wektor u, zdefiniujmy $\mathbb{W} = U - u$. Chcemy pokazać, że $\mathbb{W}$ jest podprzestrzenią liniową, czyli żę jest zamknięta na operacje.

mnożenie przez skalar jeśli $w \in \mathbb{W}$ to $w + u \in U$. Weźmy $\alpha \in \mathbb{F}$, chcemy pokazać, że $\alpha w \in \mathbb{W}$, czyli $u + \alpha w \in U$. Stosujemy (2.2) dla $u \leftarrow u$ oraz $v \leftarrow w + u$, oba wektory są w U, wtedy:

$$u + \alpha((w + u) - u) = u + \alpha w \in U.$$

Czyli $\alpha w \in \mathbb{W}$.

dodawanie wektorów Zauważmy najpierw, że jeśli $1 + 1 \neq 0$ to istnieje element odwrotny do $2 = 1 + 1$, oznaczmy go przez $\frac{1}{2}$. Wtedy $\frac{1}{2} + \frac{1}{2} = 1$:

$$\begin{aligned} \frac{1}{2} + \frac{1}{2} &= \frac{1}{2}(1 + 1) \\ &= \frac{1}{2} \cdot 2 \\ &= 1 \end{aligned}$$

Wracając do głównej części dowodu: jeśli $v, v' \in \mathbb{W}$ to $v + u, v' + u \in U$ i wtedy

$$\frac{1}{2}(v + u) + \frac{1}{2}(v' + u) = \frac{1}{2}(v + v') + u \in U \text{ i tym samym } \frac{1}{2}(v + v') \in \mathbb{W}.$$

Z punktu pierwszego mamy, że $v + v' \in \mathbb{W}$. □

Przykład/Zastosowanie 2.27 (Kontynuacja Przykładu 2.22). Chcemy zająć się ponownie rekurencjami, tym razem „prawie liniowymi”, np.

$$a_n = a_{n-1} + 2a_{n-2} - 1.$$

Łatwo sprawdzić, że zbiór rozwiązań *nie jest* przestrzenią liniową. Ale z Lematu 2.26 łatwo wynika, że jest on warstwą jakiejś przestrzeni liniowej. Z Lematu 2.25 różnica dwóch elementów z warstwy jest w odpowiadającej przestrzeni liniowej.

Tu są dwa możliwe podejścia.

- Szukamy dobrego wektora. Okazuje się, że wektor mający wszędzie tą samą wartość nadaje się; czyli szukamy α, takiego że

  $$\alpha = \alpha + 2\alpha - 1$$

  co daje $\alpha = \frac{1}{2}$. Wtedy $b_n = a_n - \frac{1}{2}$ spełnia

  $$\begin{aligned} a_n &= a_{n-1} + 2a_{n-2} - 1 & \iff \\ b_n + \frac{1}{2} &= b_{n-1} + \frac{1}{2} + 2b_{n-2} + 2 \cdot \frac{1}{2} - 1 & \iff \\ b_n &= b_{n-1} + 2b_{n-2}\ , \end{aligned}$$

  czyli uprościło się do równania liniowego, które rozwiązujemy używając poprzednich metod.

- Nie szukamy jednego wektora, lecz dla konkretnego ciągu dobieramy indywidualnie. Nasz wektor to oryginalny ciąg przesunięty (w indeksie) o jeden element, czyli $(a_{n-1})_{n\geq 0}$. Wtedy odjęcie daje

  $$a_{n+1} - a_n = a_n + 2a_{n-1} - a_{n-1} - 2a_{n-2}$$

  I to ponownie daje równanie liniowe, niestety wyższego (3.) stopnia. Wielomian dla niego jest dość skomplikowany, ale *wiemy*, że został on uzyskany jako

  $$x \cdot (x^2 - x - 2) - 1(x^2 - x - 2) = (x - 1)(x^2 - x - 2).$$

  To nam też mówi, dlaczego ciąg o wszystkich elementach takich samych zadziałał: bo w bazie ciągów nowej przestrzeni jest wektor odpowiadający ciągowi $(1^n)_{n\geq 0}$.

Rozdział 3

Przekształcenia liniowe

3.1 Przekształcenia liniowe

Definicja 3.1 (Przekształcenie liniowe)**.** Niech $\mathbb{V}, \mathbb{W}$ będą przestrzeniami liniowymi nad tym samym ciałem $\mathbb{F}$. Funkcja $F : \mathbb{V} \to \mathbb{W}$ jest *przekształceniem liniowym*, jeśli spełnia następujące warunki:

- $\forall_{v \in \mathbb{V}} \forall_{\alpha \in \mathbb{F}} F(\alpha v) = \alpha F(v)$
- $\forall_{v,w \in \mathbb{V}} F(v + w) = F(v) + F(w)$

Alternatywną nazwą dla „przekształcenie liniowe" jest *homomorfizm*, tj. mówimy, że F jest homomorfizmem między przestrzeniami liniowymi $\mathbb{V}, \mathbb{W}$ (nad tym samym ciałem) wtedy i tylko wtedy, gdy $F : \mathbb{V} \to \mathbb{W}$ jest przekształceniem liniowym. Nazwa ta jest podyktowana tym, że w ogólności „homomorfizm" oznacza przekształcenie między strukturami, które zachowujące działania. (w naszym przypadku: między przestrzeniami liniowymi, zachowują mnożenie przez skalar oraz sumę wektorów).

Przykład 3.2.
- $F : \mathbb{R}^n \to \mathbb{R}$: suma współrzędnych.
- $F : \mathbb{R}^n \to \mathbb{R}^n$: przemnożenie wszystkich współrzędnych przez stałą.
- $F : \mathbb{R}^n \to \mathbb{R}^{n-1}$ usunięcie i-tej współrzędnej.
- pochodna wielomianu (jako funkcja przestrzeni liniowej wszystkich wielomianów (o współczynnikach z $\mathbb{R}$) w nią samą)
- $F : \mathbb{Q}^3 \to \mathbb{Q}^2$, $F(x, y, z) = (2x + y, y - 3z)$
- całko (określona), tj. dla wielomianów ze współczynnikami z $\mathbb{R}$ przekształcenie $(F(f))(x) = \int\limits_0^x f(y)\mathrm{d}y$
- $F : \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = xy$ nie jest przekształceniem liniowym
- $F : \mathbb{R}^2 \to \mathbb{R}^2$, $F(x, y) = (y + 3, x - 2)$ nie jest przekształceniem liniowym

Na zbiorze przekształceń liniowych z $\mathbb{V}$ w W możemy w naturalny sposób zdefiniować dodawanie i mnożenie (przez skalar) „w punkcie":

$$(F + G)(v) = F(v) + G(v)$$
$$(\alpha F)(v) = \alpha F(v)$$

Lemat 3.3. *Zbiór przekształceń liniowych jest przestrzenią liniową.*

Dowód. Należy sprawdzić poprawność definicji, np. że gdy F jest liniowe to również αF jest liniowe:

$$(\alpha F)(v + u) = \alpha(F(v + u)) = \alpha(F(v) + F(u)) = \alpha F(v) + \alpha F(u) = (\alpha F)(v) + (\alpha F)(u)$$

Inne pokazujemy podobnie. □

Fakt 3.4. *Złożenie przekształceń liniowych jest przekształceniem liniowym.*

Lemat 3.5. *Każde przekształcenie liniowe jest jednoznacznie zadane poprzez swoje wartości na bazie. Każde takie określenie jest poprawne.*

Dowód. Niech F będzie zadane ma bazie $\vec{v}_1, \ldots, \vec{v}_n$ przestrzeni $\mathbb{V}$. Dla dowolnego v wiemy, że wyraża się ono w bazie, czyli jest postaci $v = \sum_i \alpha_i v_i$ dla pewnych skalarów $\alpha_1, \ldots, \alpha_n$. W takim razie wiemy, że wartość $F(v) = \sum_i \alpha_i F(v_i)$, co jest znane.

Poprawność określenia: trzeba sprawdzić, że jest to przekształcenie liniowe; to też wynika z jednoznaczności wyrażenia wektora w bazie. □

3.2 Jądro i obraz przekształcenia liniowego

Definicja 3.6 (Jądro i obraz przekształcenia liniowego)**.** Niech $\mathbb{V}, \mathbb{W}$ będą przestrzeniami linowymi, $F : \mathbb{V} \to \mathbb{W}$ przekształceniem liniowym.

Jądro przekształcenia to zbiór wektorów przekształcanych na $\vec{0}$:

$$\ker F = \{v \ : \ F(v) = \vec{0}\} \ .$$

Obraz przekształcenia to zbiór wektorów, które są wartościami F:

$$\mathrm{Im}(F) = \{u \ : \ \exists v F(v) = u\} \ .$$

Przykład 3.7.

- dla operacji różniczkowania i przestrzeni wielomianów stopnia nie większego niż 5, obrazem jest przestrzeń wielomianów stopnia niewiększego niż 4 a jądrem przestrzeń wielomianów stopnia nie większego niż 0.
- dla operacji całkowania przestrzeni wielomianów stopnia obrazem jest przestrzeń wielomianów stopnia różnego niż 0, a jądrem: wielomian zerowy.
- Dla przekształcenia $F : \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = x + y$ obrazem jest cała prosta $\mathbb{R}$ a jądrem prosta $x = -y$.

Lemat 3.8. *Jądro i obraz są przestrzeniami liniowymi.*

Dowód. Niech $F : \mathbb{V} \to \mathbb{W}$

Obraz jeśli $w, w' \in \mathrm{Im}\, F$ to istnieją $v, v' \in \mathbb{V}$ takie że $F(v) = w$ oraz $F(v') = w'$. Wtedy $F(v + v') = F(v) + F(v') = w + w'$ też jest w obrazie. Podobnie dla mnożenia przez skalar.

Jądro Jeśli $F(v) = \vec{0}$ to $F(\alpha v) = \alpha F(v) = \alpha\vec{0} = \vec{0}$.

Jeśli $F(v) = F(w) = \vec{0}$ to $F(v + w) = F(v) + F(w) = \vec{0} + \vec{0} = \vec{0}$. □

Fakt 3.9. *Jeśli* $F : \mathbb{V} \to \mathbb{W}$ *jest przekształceniem liniowym oraz* $\mathrm{LIN}(v_1, \ldots, v_k) = \mathbb{V}$ *to* $\mathrm{Im}(F) = \mathrm{LIN}(F(v_1), \ldots, F(v_k))$.

Dowód. Jeśli $w \in \mathrm{Im}\, F$ to $w = F(v)$ dla pewnego $v \in \mathbb{V} = \mathrm{LIN}(v_1, \ldots, v_k)$. Czyli $v = \sum_i \alpha_i v_i$ i tym samym $w = \sum_i \alpha_i F(v_i) \in \mathrm{LIN}(F(v_1), \ldots, F(v_k))$.

Jeśli $w \in \mathrm{LIN}(F(v_1), \ldots, F(v_k))$, to $w = \sum_i \alpha_i F(v_i) = F(\sum_i \alpha_i v_i) \in \mathrm{LIN}(v_1, \ldots, v_k)$. □

Twierdzenie 3.10. *Niech* $F : \mathbb{V} \to \mathbb{W}$ *będzie przekształceniem liniowym, gdzie* $\mathbb{V}, \mathbb{W}$*: skończenie wymiarowe przestrzenie liniowe. Wtedy*

$$\dim(\mathbb{V}) = \dim(\mathrm{Im}(F)) + \dim(\ker(F)).$$

Dowód. Niech $B = v_1, \ldots, v_n$ będzie bazą jądra. Zgodnie z Lematem 2.12 możemy rozszerzyć ją do bazy $\mathbb{V}$, niech te wektory to $u_1, \ldots, u_m$. Pokażemy, że $\{F(u_1), F(u_2), \ldots, F(u_m)\}$ jest bazą $\mathrm{Im}(F)$. Z Faktu 3.9 łatwo wynika, że generują obraz:

$$\begin{aligned}
\mathrm{Im}\, F &= \mathrm{LIN}(F(v_1), \ldots, F(v_n), F(u_1), \ldots, F(u_m)) \\
&= \mathrm{LIN}(\underbrace{\vec{0}, \ldots, \vec{0}}_{\text{nic nie wnoszą}}, F(u_1), \ldots, F(u_m)) \\
&= \mathrm{LIN}(F(u_1), \ldots, F(u_m)) \ .
\end{aligned}$$

Pozostaje sprawdzić, że są niezależne.

Niech $\sum_i \alpha_i F(u_i) = \vec{0}$. Wtedy $F(\sum_i \alpha_i u_i) = \vec{0}$ i tym samym $\sum_i \alpha_i u_i \in \ker F$. Ale to oznacza, że $\sum_i \alpha_i u_i \in \mathrm{LIN}(v_1, \ldots, v_n)$. Jeśli ten wektor jest niezerowy, to mamy dwa różne przedstawienia tego wektora w bazie: jedno przez wektory $v_1, \ldots, v_n$ a drugie $u_1, \ldots, u_m$, sprzeczność. Czyli $\sum_i \alpha_i u_i = \vec{0}$, co oznacza, że wszystkie współczynniki są równe 0. □

Uwaga. Dowód Twierdzenia 3.10 *nie zadziała*, jeśli weźmiemy na początku dowolną bazę $\mathbb{V}$, np. wszystkie wektory mogą przejść w to samo!

Definicja 3.11. *Rząd* przekształcenia liniowego F to $\mathrm{rk}(F) = \dim(\mathrm{Im}(F))$.

Fakt 3.12. *Jeśli* $F : \mathbb{V} \to \mathbb{W}$ *to* $\mathrm{rk}(F) \leq \min(\dim(\mathbb{V}), \dim(\mathbb{W}))$

Dowód. Ponieważ $\mathrm{Im}(F) \leq \mathbb{W}$ to $\mathrm{rk}(F) = \dim(\mathrm{Im}(F)) \leq \dim(\mathbb{W})$. Drugi punkt wynika z Twierdzenia 3.10. □

Fakt 3.13. *Jeśli* $F : \mathbb{V} \to \mathbb{V}'$ *oraz* $F' : \mathbb{V}' \to \mathbb{V}''$ *są przekształceniami liniowymi, to*

$$\mathrm{rk}(F'F) \leq \min(\mathrm{rk}(F), \mathrm{rk}(F')).$$

Dowód. W oczywisty sposób $\mathrm{Im}(F'F) \leq \mathrm{Im}(F')$, z czego mamy $\mathrm{rk}(F'F) \leq \mathrm{rk}(F')$.

Co do $\mathrm{rk}(F'F) \leq \mathrm{rk}(F)$, rozważmy przekształcenie F'' będące obcięciem F' do dziedziny będącej obrazem F, tj.

$$F'' = F' \restriction_{\mathrm{Im}(F)} \quad .$$

Wtedy $F''F$ jest dobrze określone i równe $F'F$, w szczególności $\mathrm{Im}(F''F) = \mathrm{Im}(F'F)$. Co więcej, $\mathrm{Im}(F''F) = \mathrm{Im}(F'')$, bo dziedzina F'' to dokładnie obraz F. Czyli $\mathrm{rk}(F'') = \mathrm{rk}(F''F) = \mathrm{rk}(F'F)$. Z Faktu 3.13 wiemy, że $\mathrm{rk}(F'')$ to najwyżej wymiar dziedziny, tj. $\dim(\mathrm{Im}(F'')) \leq \dim(\mathrm{Im}(F))$. □

Rozdział 4

Macierze

Chcemy operować na przekształceniach liniowych: składać je, dodawać, mnożyć itp. W tym celu potrzebujemy jakiegoś dobrego sposobu zapisu. Sposób ten jest formalizowany przy użyciu *macierzy.* Z technicznego punktu widzenia jest prościej najpierw zadać macierze a dopiero potem wyjaśnić, jak wiążą się z przekształceniami liniowymi.

Definicja 4.1. Macierzą M rozmiaru $m \times n$ nad ciałem $\mathbb{F}$ nazywamy funkcję $M : \{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\} \to \mathbb{F}$.

Zbiór wszystkich macierzy rozmiaru $m \times n$ nad ciałem $\mathbb{F}$ oznaczamy przez $M_{m,n}(\mathbb{F})$.

Zwykle macierz rozmiaru $m \times n$ oznaczamy jako tabelę:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} .$$

(Typem nawiasów za bardzo się nie przejmujemy). Zauważmy, że indeksy są zapisywane odwrotnie, niż w przypadku współrzędnych na płaszczyźnie.

Dla macierzy piszemy też $(A)_{ij}$ na oznaczenie a_{ij} i używamy podobnych konwencji. Gdy rozmiar macierzy nie jest jasny lub jest nieistotny, zapisujemy macierz jako (a_{ij})

Dla zwiększenia czytelności w zapisie macierzy używamy też przecinków między elementami a_{ij}, nawiasów okrągłych zamiast kwadratowych, przecinków między indeksami w $a_{i,j}$ itp.

4.1 Podstawowe operacje na macierzach

Definicja 4.2. Dodawanie macierzy określone jest po współrzędnych, tzn. dodawanie $A + B$ jest określone wtedy i tylko wtedy, gdy A, B są tego samego rozmiaru i wtedy

$$(A + B)_{ij} = (A)_{ij} + (B)_{ij} .$$

Mnożenie przez skalar również określone jest po współrzędnych, tzn. dla macierzy $A = (a_{ij})$ nad ciałem $\mathbb{F}$

$$(\alpha A)_{ij} = \alpha a_{ij} .$$

Tym samym macierze stanowią przestrzeń liniową (nad odpowiednim ciałem). Wektorem zerowym jest macierz złożona z samych zer.

4.1.1 Ważne i ciekawe macierze

Przykład 4.3. W poniższym przykładzie domyślnie zajmujemy się macierzami rozmiaru $m \times n$

1. *macierz zerowa* macierz składająca się z samych 0. Zwykle zapisujemy ją jako $\mathbf{0}$
2. macierz $\mathbf{1}_{ij}$: macierz, w której $a_{ij} = 1$ i wszystkie inne elementy są zerowe (Macierz ta zwana czasem *macierzą indykacyjną*, ale to nie jest dobra nazwa).

3. *macierz kwadratowa* Macierz rozmiaru $n \times n$

4. *macierz przekątniowa* macierz kwadratowa, która ma same zera poza przekątną $(a_{ii})_{i=1,\dots,n}$.

5. *macierz identycznościowa/jednostkowa* macierz przekątniowa, która ma jedynki na przekątnej $((a_{ii})_{i=1,\dots,n})$. Zapisywana jako Id_n.

$$\mathrm{Id}_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

6. *macierz górnotrójkątna* macierz kwadratowa, w której wszystkie elementy $(a_{ij})_{i>j}$ są zerowe

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

7. *macierz dolnotrójkątna* macierz kwadratowa, w której wszystkie elementy $(a_{ij})_{i<j}$ są zerowe

$$\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{nn} \end{bmatrix}$$

8. *macierz trójkątna* macierz dolnotrójkątna lub górnotrójkątna

4.1.2 Zestawianie macierzy

Mając dwie macierze M, M' rozmiaru $m \times n$ oraz $m \times n'$ (nad tym samym ciałem) będziemy pisać

$$\begin{bmatrix} M | M' \end{bmatrix}$$

na macierz rozmiaru $m \times (n+n')$ uzyskaną przez „zestawienie" macierzy M, M'. Rozszerzamy tę konwencję na wiele macierzy $M_1, M_2, \dots, M_k$ rozmiaru $m \times n_1$, $m \times n_2$, ..., $m \times n_k$ i piszemy $[M_1|M_2|\cdots|M_k]$. Jeśli macierze te są wymiaru $m \times 1$ to zwykle używamy liter $C_1, \dots, C_k$, jako że są to kolumny wynikowej macierzy.

Podobnie zestawiamy macierze w pionie: dla macierzy M, M' rozmiaru $m \times n$ i $m' \times n$ piszemy

$$\begin{bmatrix} M \\ \hline M' \end{bmatrix}$$

na „zestawienie" tych dwóch macierzy w pionie (w tym wypadku jest ono rozmiaru $(m+m') \times n$). Ponownie używamy tej notacji dla wielu macierzy $M_1, M_2, \dots, M_k$, jeśli macierz mają tylko jeden wiersz to zwykle oznaczamy je jako $R_1, R_2, \dots, R_m$ (bo są to wiersze).

4.1.3 Mnożenie macierzy

Mnożenie macierzy zdefiniujemy najpierw dla macierzy $1 \times n$ oraz $n \times 1$.

$$\begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \sum_{k=1}^{n} a_k b_k \ .$$

Wynik, w zależności od potrzeb, traktujemy jako liczbę (z ciała $\mathbb{F}$) lub jako macierz 1×1.

Mnożenie wektorów $m \times 1$ oraz $1 \times n$ definiujemy jako:

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \cdot \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} = \left[\begin{array}{c|c|c|c} b_1a_1 & b_1a_2 & \cdots & b_1a_n \\ b_2a_1 & b_2a_2 & \cdots & b_2a_n \\ \vdots & \vdots & \ddots & \vdots \\ b_ma_1 & b_ma_2 & \cdots & b_ma_n \end{array}\right] .$$

Następnie rozszerzamy mnożenie do macierzy rozmiaru $m \times k$ i $k \times n$ (wynikiem jest macierz rozmiaru $m \times n$). Mnożenie definiujemy tak, że dzielimy lewą macierz na wiersze a prawą na kolumny i mnożymy jak dwa wektory (odpowiednio: wierszy i kolumn), przy czym pojedyncze mnożenie wiersza i kolumny wykonujemy jak mnożenie wektorów.

$$\left[\begin{array}{c} R_1 \\ \hline R_2 \\ \hline \vdots \\ \hline R_m \end{array}\right] \cdot \left[\begin{array}{c|c|c|c} C_1 & C_2 & \cdots & C_n \end{array}\right] = \left[\begin{array}{c|c|c|c} R_1C_1 & R_1C_2 & \cdots & R_1C_n \\ \hline R_2C_1 & R_2C_2 & \cdots & R_2C_n \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline R_mC_1 & R_mC_2 & \cdots & R_mC_n \end{array}\right] .$$

Używając notacji z indeksami, jeśli $A = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,k}}$, $B = (b_{ij})_{\substack{i=1,\ldots,k \\ j=1,\ldots,n}}$, to $C = AB$ ma postać $(c_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$, gdzie

$$c_{ij} = \sum_{\ell=1}^{k} a_{i\ell} b_{\ell j} .$$

Uwaga. Zauważmy, że możliwy jest też odwrotny podział: lewa macierz jako wektor kolumn a prawa jako wektor wierszy. Wykonując bezpośrednie rachunki można łatwo sprawdzić, że wynik jest ten sam.

Fakt 4.4. *Mnożenie macierzy jest łączne.*

Dowód. Bo jest to funkcja, a składanie funkcji (w ogólności: relacji) jest łączne. □

Fakt 4.5. *Niech A, B, C będą macierzami nad tym samym ciałem $\mathbb{F}$, Id_n macierzą identycznościową $n \times n$, $\alpha \in \mathbb{F}$. Wtedy poniższe równości zachodzą, dla macierzy odpowiednich rozmiarów (tzn. takich, że odpowiednie mnożenie/dodawanie jest określone):*

1. $\mathrm{Id}_n A = A$, $B \,\mathrm{Id}_n = B$;
2. $A(B + C) = AB + BC$;
3. $(B + C)A = BA + CA$;
4. $\alpha(AB) = (\alpha A)B = A(\alpha B)$;
5. $A[B|C] = [AB|AC]$;
6. $\left[\begin{array}{c} B \\ \hline C \end{array}\right] A = \left[\begin{array}{c} BA \\ \hline CA \end{array}\right]$.

Dowód sprowadza się do prostych rachunków i zostanie pokazany na ćwiczeniach.

Przykład/Zastosowanie 4.6. Jak obliczać wyrazy ciągu Fibonacciego szybko?

Robienie tego przy użyciu wzorów z potęgami liczb niewymiernych nie jest praktyczne: powstają błędy zaokrągleń, mnożenie liczb rzeczywistych jest kosztowne. Choć wciąż działa to proporcjonalnie do $\log n$, a nie n, gdy chcemy policzyć n-ty wyraz.

Zapiszmy kolejne wartości jako wektory:

$$\begin{bmatrix} f_0 = 0 \\ f_1 = 1 \end{bmatrix}, \begin{bmatrix} f_1 = 1 \\ f_2 = 1 \end{bmatrix}, \begin{bmatrix} f_2 = 1 \\ f_3 = 2 \end{bmatrix}, \ldots, \begin{bmatrix} f_n \\ f_{n+1} \end{bmatrix} .$$

Zauważmy, że rekurencją możemy zapisać w postaci macierzy:

$$\begin{bmatrix} f_{n+1} \\ f_{n+2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} f_n \\ f_{n+1} \end{bmatrix} .$$

Wartości początkowe wpisujemy w wektor. Wtedy kolejne nałożenia to kolejne potęgi:

$$\begin{bmatrix} f_n \\ f_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n \cdot \begin{bmatrix} f_0 = 0 \\ f_1 = 1 \end{bmatrix} .$$

Zauważmy, że dzięki temu możemy policzyć f_n podnosząc naszą macierz do n-tej potęgi, co można wykonać w czasie proporcjonalnym do $\log n$.

4.1.4 Transpozycja

Definicja 4.7 (Transpozycja)**.** Dla macierzy $M = (m_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$ macierz M^T zdefiniowana jest jako

$$M^T = (m_{ji})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$$

to jest jako „obrót" wokół przekątnej.

Przykład 4.8.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

Lemat 4.9. *Dla macierzy M, N odpowiednich rozmiarów zachodzi*

$$\begin{aligned} (M+N)^T &= M^T + N^T \quad , \\ (MN)^T &= N^T M^T \quad , \\ (M^T)^T &= M \quad . \end{aligned}$$

Prosty dowód zostanie pokazany na ćwiczeniach.

4.2 Wartości na wektorach jednostkowych

Zdefiniujmy macierze rozmiaru $n \times 1$ (wektory) $\vec{E}_1, \dots, \vec{E}_n$, wektor $\vec{E}_i$ ma 1 na i-tej współrzędnej oraz 0 wszędzie poza tą pozycją (czyli inne spojrzenie na bazę standardową).

Lemat 4.10 (Bardzo ważny)**.**

$$M = \left[\ M\vec{E}_1 \mid M\vec{E}_2 \mid \cdots \mid M\vec{E}_n\ \right]$$

Dowód. Dowód można pokazać wprost z definicji, lub też zastosować trik:

$$\mathrm{Id}_n = \left[\ \vec{E}_1 \mid \vec{E}_2 \mid \cdots \mid \vec{E}_n\ \right]$$

i tym samym

$$M = M\,\mathrm{Id}_n = M \left[\ \vec{E}_1 \mid \vec{E}_2 \mid \cdots \mid \vec{E}_n\ \right] = \left[\ M\vec{E}_1 \mid M\vec{E}_2 \mid \cdots \mid M\vec{E}_n\ \right] \quad \square$$

Uwaga. To jest bardzo użyteczna własność: często zamiast pokazać równość macierzy czy też pomnożyć jakieś macierze będziemy liczyli wartości na wektorach $\vec{E}_i$.

Przykład/Zastosowanie 4.11 (Kontynuacja Zastosowania 4.6)*.* Teraz możemy też powiedzieć, skąd wzięliśmy macierz $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ służącą do liczenia wartości wyrazów ciągu Fibonacciego: jeśli weźmiemy iloczyn tej macierzy i wektorów $\vec{E}_1, \vec{E}_2$, a z drugiej wektory $\begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$ dla warunków początkowych $\begin{bmatrix} f_0 \\ f_1 \end{bmatrix} = \vec{E}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ oraz $\begin{bmatrix} f_0 \\ f_1 \end{bmatrix} = \vec{E}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

4.3 Operacje elementarne

Definicja 4.12 (Operacje elementarne.). Operacje elementarne (kolumnowe) to:

- zamiana kolumn;
- dodanie do jednej z kolumn wielokrotności innej;
- przemnożenie kolumny przez niezerowy skalar.

Analogicznie definiujemy operacje elementarne wierszowe.

Operacje elementarne można wyrazić jako macierze:

- macierz T_{ij} ma następujące wyrazy: na przekątnej 1, poza ii, jj, gdzie T_{ij} ma 0, oprócz przekątnej ma same 0, poza ij, ji, gdzie ma 1.

$$T_{3,6} = \begin{bmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & 0 & & & 1 & \\ & & & 1 & & & \\ & & & & 1 & & \\ & & 1 & & & 0 & \\ & & & & & & 1 \end{bmatrix}$$

- $\mathrm{Id}_n + 3 \cdot \alpha 1_{ij}$

$$\mathrm{Id}_6 + 3 \cdot 1_{3,6} = \begin{bmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & 1 & & & 3 & \\ & & & 1 & & & \\ & & & & 1 & & \\ & & & & & 1 & \\ & & & & & & 1 \end{bmatrix}$$

- $D_{i\alpha}$ to macierz przekątniowa, która na pozycji ii ma $\alpha \neq 0$ a pozostałe elementy na przekątnej to 1.

$$D_{3,2} = \begin{bmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & 2 & & & & \\ & & & 1 & & & \\ & & & & 1 & & \\ & & & & & 1 & \\ & & & & & & 1 \end{bmatrix}$$

Lemat 4.13 (Operacje elementarne jako macierze).

- *$M \cdot T_{ij}$ to macierz powstała przez zamianę i-tej oraz j-tej kolumny.*
- *$M \cdot (\mathrm{Id}_n + \alpha 1_{ij})$ to macierz powstała przez dodanie do j-tej kolumny α razy i-tej kolumny.*
- *$M \cdot (D_{i\alpha})$ to macierz powstała przez przemnożenie i-tej kolumny przez α .*

W szczególności:

- $T_{ij} \cdot T_{ij} = \mathrm{Id}$
- $(\mathrm{Id}_n + \alpha 1_{ij}) \cdot (\mathrm{Id}_n - \alpha 1_{ij}) = \mathrm{Id}_n$
- $D_{i\alpha} D_{i1/\alpha} = \mathrm{Id}_n$.

Dowód. Wszystkie fakty można pokazać przez bezpośrednie obliczenia, ale skorzystamy z Lematu 4.10 aby uzyskać ładną interpretację.

Rozważmy $M' = M \cdot T_{ij}$. Jej kolumny to wartości mnożenia kolejnych wektorów $\vec{E}_k$ przez M. Popatrzmy na $MT_{ij}\vec{E}_k$. Jeśli $k \notin \{i, j\}$ to $T_{ij}\vec{E}_k = \vec{E}_k$ i tym samym, $M'\vec{E}_k = M\vec{E}_k$, czyli M' oraz M mają te same kolumny $k \notin \{i, j\}$. Z drugiej strony $T_{ij}\vec{E}_i = \vec{E}_j$ i tym samym $M'\vec{E}_i = M\vec{E}_j$, czyli i-ta kolumna M' to j-ta kolumna M. Taka samo jest dla j. Czyli faktycznie M' powstaje przez zamianę i-tej oraz j-tej kolumny.

Rozumowanie dla $M \cdot (\mathrm{Id}_n + \alpha 1_{ij})$ jest analogiczne: rozważmy

$$M \cdot (\mathrm{Id}_n + \alpha 1_{ij})\vec{E}_k = M\vec{E}_k + M 1_{ij} \alpha \vec{E}_k \tag{4.1}$$

Zauważmy, że

$$1_{ij}\vec{E}_k = \begin{cases} \vec{0} & \text{dla } k \neq j \\ \vec{E}_i & \text{dla } k = j \end{cases} .$$

Wtedy (4.1) wynosi:

$$M\vec{E}_k + M 1_{ij} \alpha \vec{E}_k = \begin{cases} M\vec{E}_k & \text{dla} k \neq j \\ M\vec{E}_j + \alpha M \vec{E}_i & \text{dla } k = j \end{cases}$$

tj. w pierwszym przypadku jest to po prostu k-ta kolumna M, w drugim j-ta + α razy i-ta.

Dowód dla macierzy $D_{i\alpha}$ jest analogiczny, zauważmy, że możemy potraktować ją jako macierz $\mathrm{Id} + (\alpha - 1)1_{ii}$.

Podane równości łatwo udowodnić używając ich interpretacji: dla przykładu rozważmy $T_{ij}T_{ij}$ zinterpretowane jako $T_{ij}T_{ij}\,\mathrm{Id}$. Chcemy pokazać, że ich iloczyn wynosi Id. Wtedy $T_{ij}T_{ij}$ zamienia i-tą i j-tą kolumnę i potem znów zamienia te kolumny, czyli otrzymujemy macierz Id. Dowód dla pozostałych operacji jest podobny. □

Analogiczną interpretację można uzyskać też dla operacji wierszowych:

Lemat 4.14.
1. *Dla M odpowiedniego rozmiaru $T_{ij}M$ jest macierzą powstałą z M przez zamianę i-tego oraz j-tego wiersza.*
2. *Dla M odpowiedniego rozmiaru $(\mathrm{Id}_n + \alpha_{ij})M$ jest macierzą powstałą z M poprzez dodanie do i-tego wiersza α razy j-tego wiersza.*
3. *Dla M odpowiedniego rozmiaru $(D_{i\alpha}) \cdot M$ to macierz powstała przez przemnożenie i-tego wiersza M przez α .*

Zwróćmy uwagę, że dla macierzy $\mathrm{Id} + \alpha 1_{ij}$ zmienia się, który wiersz dodajemy do którego (w porównaniu z kolumnami). (Nie będzie to jednak istotne, zwykle używamy tylko, że taką operację da się wykonać macierzami tego typu.)

Dowód. Dowód można przeprowadzić wprost, przez bezpośrednie rachunki, ale prościej jest odwołać się do transpozycji, np.:

$$T_{ij}M = ((T_{ij}M)^T)^T = (M^T T_{ij}^T)^T = (M^T T_{ij})^T$$

przy czym $M^T T_{ij}$ jest macierzą M^T w której zamieniono i-tą oraz j-tą kolumnę, tak więc po transpozycji jest to M w której zamieniono i-ty i j-ty wiersz.

Podobnie dowodzimy pozostałych własności, warto przy tym zauważyć, że $1_{ij}^T = 1_{ji}$. □

Definicja 4.15 (Macierze elementarne)**.** Macierze odpowiadające operacjom elementarnym, tj. T_{ij} dla $i \neq j$, $(\mathrm{Id}_n + \alpha 1_{ij})$ dla $i \neq j$ oraz $D_{i\alpha}$ dla $\alpha \neq 0$ nazywamy *macierzami elementarnymi.*

Zauważmy, że tym samym możemy zinterpretować cały proces eliminacji Gaussa jako kolejne działania macierzy elementarnych.

Fakt 4.16. *Eliminację Gaußa można zinterpretować jako mnożenie macierzy powstałej przez zestawienie wektorów (w wierszach/kolumnach) z układu wejściowego przez macierze elementarne (odpowiednio z lewej lub prawej strony).*

4.4 Przekształcenie liniowe dla macierzy

Od teraz (w zasadzie do końca) wektory zapisujemy w pionie i identyfikujemy je z macierzami $n \times 1$.

Dla macierzy M rozmiaru $m \times n$ możemy zadać przekształcenie liniowe $F_M : \mathbb{F}^n \to \mathbb{F}^m$ przez

$$F_M(\vec{v}) = M\vec{v} \ .$$

Liniowość wynika z liniowości mnożenia macierzy.

Takie przekształcenie będziemy nazywać *przekształceniem indukowanym przez macierz* M.

Twierdzenie 4.17. *Przekształcenie* $M \mapsto F_M$ *jest izomorfizmem (przestrzeni liniowych) zbioru macierzy* $M_{m\times n}(\mathbb{F})$ *i zbioru przekształceń liniowych z* $\{F : F : \mathbb{F}^n \to \mathbb{F}^m, F \textit{ jest przekształceniem liniowym}\}$.

Pozostaje sprawdzić, jak wyraża się składanie tak zadanych przekształceń.

Twierdzenie 4.18. *Dla macierzy odpowiednich rzędów mamy*

$$F_{M'M} = F_{M'}F_M$$

tzn. przekształcenia zadane przez iloczyn macierzy $M'M$ *jest złożeniem przekształceń zadanych przez macierze* M' *i* M.

Dowód. Należy pokazać, że

$$F_{M'M}(v) = F_{M'}F_M(v)$$

Co jest oczywiste, bo obie strony to tylko inne nawiasowania mnożenia macierzy $M'Mv$. □

4.5 Rząd macierzy

Definicja 4.19 (Rząd macierzy)**.** Rząd macierzy to wymiar przestrzeni generowanej przez kolumny tej macierzy (traktowanych jako wektory w $\mathbb{F}^n$). Oznaczamy go przez $\mathrm{rk}(M)$. Tj. jeśli $M = [M_1|M_2|\cdots|M_n]$ to

$$\mathrm{rk}(M) = \dim \mathrm{LIN}(M_1, \dots, M_n)$$

Lemat 4.20. *Niech* M *będzie macierzą a* F_M *indukowanym przez nią przekształceniem liniowym. Wtedy*

$$\mathrm{rk}(M) = \mathrm{rk}(F_M)$$

Dowód. Rozważmy bazę standardową $\vec{E}_1, \dots, \vec{E}_n$. Wtedy wektory $F_M\vec{E}_1, \dots, F_M\vec{E}_n$ generują obraz $\mathrm{Im}\, F_M$. Jednocześnie są to kolumny macierzy M. □

Ta obserwacja pozwala przełożyć znane nam wyniki dotyczące rzędu przekształceń liniowych na macierze. Np.

Lemat 4.21. *Dla macierzy* M, N *odpowiednich rozmiarów zachodzi*

$$\mathrm{rk}(MN) \leq \min(\mathrm{rk}(M), \mathrm{rk}(N))$$

Dowód. Popatrzmy na przekształcenia $F_M, F_N, F_{MN} = F_M \circ F_N$. Odpowiednia nierówność zachodzi dla ich rzędów a zgodnie z Lematem 4.20 rzędy macierzy i ich indukowanych przekształceń są równe. □

Tak zdefiniowany rząd nazwiemy na potrzebę kolejnego dowodu *rzędem kolumnowym*, analogicznie można zdefiniować *rząd wierszowy*. Okazuje się, że są one równe.

Twierdzenie 4.22. *Rząd kolumnowy i wierszowy ustalonej macierzy* M *są sobie równe.*

W szczególności, $\mathrm{rk}(M) = \mathrm{rk}(M^T)$.

Pokażemy dowód tego faktu oparty na algorytmie eliminacji Gaussa.

Lemat 4.23. *Operacje elementarne kolumnowe (wierszowe) na macierzach nie zmieniają rzędu wierszowego i kolumnowego macierzy.*

Dowód. Pokażemy dowód w przypadku operacji kolumnowych, dla operacji wierszowych przebiega tak samo (lub możemy przejść przez transpozycję do przypadku operacji kolumnowych).

Z Lematu 1.18 operacje kolumnowe nie zmieniają otoczki liniowej i tym samym nie zmieniają rzędu kolumnowego.

Pozostaje nam pokazać, że operacje kolumnowe nie zmieniając rzędu wierszowego:

zamiana kolumn Zamiana kolumn to dla wektorów wierszowych zamiana kolejności współrzędnych. Nie wpływa ona na liniową niezależność układu wektorów: jeśli $\sum_i \alpha_i \vec{v}_i = \vec{0}$ to dla każdej współrzędnej j mamy $\sum_i \alpha_i (\vec{v}_i)_j = 0$ i odwrotnie. Tak więc zamiana kolejności kolumn nic nie zmienia.

dodanie wielokrotności innej kolumny Dodanie wielokrotności kolumny to dodanie wielokrotności którejś ze współrzędnych. Niech wektory przed tą operacją to $v_1, \ldots, v_n$ a po niej: $v'_1, \ldots, v'_n$. Twierdzimy, że

$$\sum_i \alpha_i v_i = \vec{0} \iff \sum_i \alpha_i v'_i = \vec{0} \ .$$

W szczególności, oryginalny układ jest liniowo zależny wtedy i tylko wtedy, gdy nowy jest liniowo zależny.

Wystarczy pokazać implikację w jedną stronę, bo w drugą da się uzyskać przez operację dodania kolumny.

Skoro $\sum_i \alpha_i v_i = \vec{0}$ to dla każdej współrzędnej j mamy, że

$$\sum_i \alpha_i (v_i)_j = 0 \ ,$$

czyli po dodanie wielokrotności innej współrzędnej to dodanie wielokrotności 0: powiedzmy, żę do j-tej kolumny dodajemy α razy j'-tą. Wtedy dla każdej innej współrzędnej $j'' \neq j$ mamy

$$\begin{aligned} \sum_i \alpha_i (v'_i)_{j''} &= \sum_i \alpha_i (v'_i)_j \\ &= 0 \end{aligned}$$

Zaś dla j-tej

$$\begin{aligned} \sum_i \alpha_i (v'_i)_j &= \sum_i \alpha_i ((v_i)_j + \alpha (v_i)_{j'}) \\ &= \sum_i \alpha_i (v_i)_j + \alpha \sum_i \alpha_i (v_i)_{j'} \\ &= 0 + 0 = 0 \end{aligned}$$

kombinacja dodanie wielokrotności tej współrzędnej też daje 0.

Ale to oznacza, że $v_1, \ldots, v_n$ są niezależne wtedy i tylko wtedy, kiedy niezależna są $v'_1, \ldots, v'_n$.

przemnożenie kolumny przez skalar Tak jak w przypadku powyżej, przemnożenie kolumny przez skalar to dla wektorów kolumnowych przemnożenie wszystkim wektorom jednej ze współrzędnych przez ten skalar. Tak jak powyżej, nie wpływa to na niezależność.

□

Lemat 4.24. *Dla macierzy w wierszowej postaci schodkowej rząd wierszowy i kolumnowy macierzy jest taki sam.*

Analogiczne stwierdzenie zachodzi dla macierzy w kolumnowej postaci schodkowej.

Dowód. Niech macierz $M = (m_{i,j})$ w wierszowej postaci schodkowej ma wiodące elementy na pozycjach $(1, j_1), (2, j_2), \ldots, (k, j_k)$, gdzie dla wiersza i oraz $j' < j_i$ mamy $m_{i,j'} = 0$ oraz k jest rzędem wierszowym.

$$\begin{bmatrix}
\underbrace{1}_{(1,j_1)=(1,1)} & & & & & & & & \\
0 & 0 & \underbrace{2}_{(2,j_2)=(2,3)} & & & & & & \\
0 & 0 & 0 & 0 & \underbrace{-1}_{(3,j_3)=(3,5)} & & & & \\
\vdots & \vdots & \vdots & \vdots & & \ddots & & & \\
0 & 0 & 0 & 0 & 0 & \ldots & 0 & \underbrace{5}_{(k,j_k)} & \\
0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0
\end{bmatrix}$$

Przeprowadzamy eliminację Gaußa na kolumnach: używając elementu $(1, j_1)$ usuwamy wszystkie niezerowe elementy w wierszu 1., potem elementu $(2, j_2)$ wszystkie w wierszu 2., itd. Po tej operacji w każdym wierszu i kolumnie mamy najwyżej jeden niezerowy element.

$$\begin{bmatrix}
\underbrace{1}_{(1,j_1)=(1,1)} & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 \\
0 & 0 & \underbrace{2}_{(2,j_2)=(2,3)} & 0 & 0 & \ldots & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \underbrace{-1}_{(3,j_3)=(3,5)} & \ldots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & \ldots & 0 & \underbrace{5}_{(k,j_k)} & 0 \\
0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0
\end{bmatrix}$$

Czyli jest k liniowo niezależnych kolumn, czyli rząd kolumnowy to k.

Dowód dla kolumnowej postaci schodkowej przeprowadzamy analogicznie (albo przechodzimy przez transpozycje i korzystamy z wierszowej postaci schodkowej). □

dowód Twierdzenia 4.22. Stosujemy eliminację Gaussa. Rząd obu się nie zmienia (Lemat 4.23). Dla macierzy w postaci schodkowej teza zachodzi z Lematu 4.24. □

Dlatego od tego momentu mówimy po prostu o rzędzie macierzy.

Uwaga. Przy liczeniu liniowej niezależności dla zbioru wektorów możemy wykonywać *zarówno* operacje wierszowe jak i kolumnowe. Proszę jednak pamiętać, że wykonywanie operacji kolumnowych nie zmienia przestrzeni rozpiętej przez kolumny, natomiast wykonanie operacji wierszowych może zmienić tę przestrzeń (i zwykle zmienia). Tym samym jeśli mieszamy te operacje, to nie umiemy powiedzieć, np. jaka jest baza przestrzeni rozpiętej przez układ wektorów.

Tym niemniej, jeśli stosujemy oba typy operacji, ale użyjemy tylko wierszy $\{i_1, \ldots, i_k\}$ do eliminacji (i jakichś kolumn) i na końcu odpowiadające wiersze są niezależne (a pozostałe zerami), to odpowiadające wiersze z wejścia są niezależne.

4.6 Obliczanie bazy jądra przekształcenia

Jako przykładowe zastosowaniem macierzy, pokażemy jak obliczyć bazę jądra przekształcenia indukowanego przez macierz M (zwanego dalej po prostu jądrem macierzy).

Napiszmy

$$M \operatorname{Id}_n = M$$

Wykonujemy teraz eliminację Gaussa (na kolumnach) tak długo, aż doprowadzimy M (po prawej stronie) do postaci schodkowej (kolumnowej). Zauważmy, że możemy myśleć o tych operacjach jak o macierzach, czyli odpowiadają one mnożeniu z prawej strony obu stron równości przez te same macierze, czyli wykonywaniu tych samych operacji kolumnowych na M oraz Id_n (czy też dokładniej macierzy, która tam jest).

Na końcu otrzymujemy zależność postaci:

$$MA = M'$$

gdzie M' jako pierwsze kolumny zawiera wektory niezależne, a potem same wektory zerowe. Ale to oznacza, że przy mnożeniu przez M odpowiednie wektory w macierzy A przechodzą na wektory 0. W czasie trwania procesu kolumny A pozostają niezależne (bo to jest eliminacja Gaussa), czyli odpowiednie kolumny stanowią bazę jądra.

Zauważmy, że M po lewej stronie potrzebne jest tylko do dowodu, w samym algorytmie możemy go nie używać. Innymi słowy, algorytm trzyma parę macierzy (początkowo: (M, Id_n)) i wykonujemy na obu z nich takie same operacje kolumnowe, tak by doprowadzić M do postaci schodkowej (kolumnowej). Wtedy kolumny w drugiej macierzy odpowiadające kolumnom $\vec{0}$ z pierwszej to baza jądra.

Przykład 4.25. Dla macierzy

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \xrightarrow{(2)+(3)-(1)} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Wykonując analogiczne operacje na ma cierzy Id_3:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{(2)+(3)-(1)} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} .$$

Łatwo sprawdzić, że faktycznie wektor $[-1, 1, 1]^T$ należy do jądra Z drugiej strony, wymiar jądra to $\dim(\mathbb{R}^3) - 2 = 1$, czyli faktycznie ten wektor stanowi bazę jądra.

4.7 Macierz odwrotna

Definicja 4.26. Macierz kwadratowa, która ma przekształcenie odwrotne, tj. istnieje macierz M' taka że

$$M \cdot M' = M' \cdot M = \mathrm{Id}_n \ ,$$

nazywamy macierzą *odwracalną* lub macierzą *nieosobliwą*. Macierz M' o właściwościach jak wyżej nazywamy macierzą odwrotną do M i oznaczamy przez M^{-1}.

Lemat 4.27. *Macierz M jest odwracalna $\iff$ przekształcenie F_M jest odwracalne. Co więcej, $F_{M^{-1}} = F_M^{-1}$.*

Dowód. Jeśli M^{-1} to macierz odwrotna do M, to $MM^{-1} = \mathrm{Id}$ i tym samym

$$F_M F_{M^{-1}} = F_{MM^{-1}} = F_{\mathrm{Id}} = \mathrm{Id}$$

analogicznie $M^{-1}M = \mathrm{Id}$ implikuje $F_{M^{-1}} F_M = \mathrm{Id}$.

W drugą stronę rachunki są analogiczne: jeśli $F_M F_{M'} = \mathrm{Id}$ to z jednej strony

$$F_M F_{M'} = F_{MM'}$$

zaś z drugiej

$$\mathrm{Id} = F_{\mathrm{Id}}$$

i tym samym $MM' = \mathrm{Id}$. Analogicznie $F_{M'} F_M = \mathrm{Id}$ implikuje, że $M'M = \mathrm{Id}$ i tym samym M' jest macierzą odwracalną. □

Twierdzenie 4.28. *Macierz A wymiaru $n \times n$ jest odwracalna $\iff$ $\mathrm{rk}(A) = n$.*

Dowód. $\Leftarrow$ Jeśli A jest odwracalna to jest w szczególności różnowartościowa, czyli $\dim \ker A = 0$ i tym samym

$$n = \mathrm{rk}(A) + \dim \ker A$$

implikuje, że $\mathrm{rk}(A) = n$.

$\Leftarrow$ Niech $A = [A_1|A_2|\cdots|A_n]$. Zgodnie z założeniem, wektory $A_1, \ldots, A_n$ są niezależne, czyli są bazą. Rozpatrzmy $F = F_A$. Zadajemy przekształcenie odwrotne F^{-1} na wektorach $(A_1, A_2, \ldots, A_n)$ jako $\vec{E}_1, \vec{E}_2, \ldots, \vec{E}_n$. Tak zadane F^{-1} istnieje, zgodnie z Lematem 3.5. Wtedy

$$FF^{-1}A_i = F\vec{E}_i = A_i$$

Czyli FF^{-1} jest identycznością na bazie, czyli jest identycznością. Analogicznie

$$F^{-1}F\vec{E}_i = F^{-1}A_i = \vec{E}_i$$

czyli $F^{-1}F$ też jest identycznością. W takim razie F jest odwracalne, czyli też A jest odwracalne, zgodnie z Lematem 4.27. □

Lemat 4.29. *Jeśli A jest macierzą kwadratową $n \times n$ to macierz kwadratowa B jest jej odwrotnością, jeśli $AB = \mathrm{Id}$ lub $BA = \mathrm{Id}$.*

Dowód. Niech $BA = \mathrm{Id}$, ustalmy $\vec{E}_1, \vec{E}_2, \ldots, \vec{E}_n$: baza standardowa $\mathbb{R}^n$. Niech $v_i = A\vec{E}_i$. Wtedy $Bv_i = \vec{E}_i$, z czego wnioskujemy, że $\vec{v}_1, \ldots, \vec{v}_n$ są bazą (ich obraz jest, a to ta sama przestrzeń).

Ale w takim razie F_{AB} jest identycznością na tej bazie, czyli jest identycznością.

Drugi przypadek pokazujemy analogicznie. □

Dowody poniższych prostych faktów pokażemy na ćwiczeniach.

Fakt 4.30. *Jeśli MN jest odwracalna a M, N są kwadratowe, to również M, N są odwracalne.*

Niech M, N będą odwracalne. Wtedy:

- $(M^T)^{-1} = (M^{-1})^T$
- $(M^{-1})^{-1} = M$
- $(MN)^{-1} = N^{-1}M^{-1}$

Proste dowody pozostawiamy jako ćwiczenia.

Fakt 4.31. *Jeśli A jest macierzą odwracalną a B, C są macierzami odpowiednich rozmiaróœ (tzn. takimi, że mnożenia AB oraz CA są określone) to*

$$\mathrm{rk}(AB) = \mathrm{rk}(B) \quad \textit{oraz} \quad \mathrm{rk}(CA) = \mathrm{rk}(C) \ .$$

4.7.1 Metoda algorytmiczna obliczania macierzy odwrotnej

Przedstawimy efektywny sposób obliczania macierzy odwrotnej.

Zapiszmy równanie:

$$A^{-1}A = \mathrm{Id}\,.$$

Dokonujemy diagonalizacji A używając metody eliminacji (dla kolumn). Wiemy, że każda operacja kolumnowa odpowiada przemnożeniu (z prawej strony) przez odpowiednią macierz elementarną. Tym samym w kroku pośrednim mamy równanie postaci

$$A^{-1}A' = B,$$

gdzie B jest macierzą uzyskaną przez zastosowanie tych samych operacji na Id, co na A.

Gdy A' jest macierzą diagonalną, to albo ma jakieś 0 (sprzeczność), albo nie i wtedy przekształcamy ją do macierzy Id mnożąc odpowiednio kolumny przez skalar. Te same operacje wykonujemy na macierzy B.

Na końcu uzyskujemy równanie

$$A^{-1}\,\mathrm{Id} = B$$

i tym samym mamy szukaną przez nas macierz A^{-1}.

Przykład 4.32. Obliczmy macierz odwrotną do

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

Sprowadzamy A do macierzy identycznościowej

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \xrightarrow{1/2((2)+(1)-(3))} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{(1)-(2)} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{(3)-(1)} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{(1)\leftrightarrow(2)} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Wykonajmy te same operacje na macierzy identycznościowej

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{1/2((2)+(1)-(3))} \begin{bmatrix} 1 & 0,5 & 0 \\ 0 & 0,5 & 0 \\ 0 & -0,5 & 1 \end{bmatrix} \xrightarrow{(1)-(2)} \begin{bmatrix} 0,5 & 0,5 & 0 \\ -0,5 & 0,5 & 0 \\ 0,5 & -0,5 & 1 \end{bmatrix} \xrightarrow{(3)-(1)}$$
$$\begin{bmatrix} 0,5 & 0,5 & -0,5 \\ -0,5 & 0,5 & 0,5 \\ 0,5 & -0,5 & 0,5 \end{bmatrix} \xrightarrow{(1)\leftrightarrow(2)} \begin{bmatrix} 0,5 & 0,5 & -0,5 \\ 0,5 & -0,5 & 0,5 \\ -0,5 & 0,5 & 0,5 \end{bmatrix}$$

Łatwo sprawdzić bezpośrednim rachunkiem, że faktycznie jest to macierz odwrotna do zadanej.

4.8 Jeszcze o eliminacji Gaußa

Lemat 4.33. *Jeśli macierz M jest odwracalna, to przy użyciu eliminacji Gaußa (na wierszach lub kolumnach) można doprowadzić ją do macierzy przekątniowej (bez zer na przekątnej).*

Używając eliminacji Gaußa zarówno na wierszach jak i na kolumnach można dowolną macierz kwadratową przekształcić do macierzy przekątniowej. Ponadto, można najpierw wykonać wszystkie operacje na wierszach a potem na kolumnach (lub odwrotnie).

Dowód. Postępujemy jak w Lemacie 4.24. Użyjmy eliminacji Gaußa na wierszach, dla operacji na kolumnach jest tak samo. Doprowadzamy macierz M do postaci schodkowej (wierszowej).

Jeśli jest odwracalna, to po przeprowadzeniu eliminacji Gaußa ma na przekątnej same niezerowe elementy. Idąc od dołu możemy kolejno eliminować niezerowe elementy poza przekątną dla kolumny nr $n, n-1, \ldots, 1$, analogicznie do wcześniejszego eliminowanie pod przekątną.

$$\begin{bmatrix} 1 & & & & & \\ 0 & 2 & & & & \\ 0 & 0 & -1 & & & \\ 0 & 0 & 0 & 3 & & \\ 0 & 0 & 0 & 0 & 2 & \\ 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$

Jeśli macierz nie była odwracalna, to używamy operacji na kolumnach: używając niezerowego elementu w pierwszym wierszu eliminujemy wszystkie niezerowe elementy na prawo, potem analogicznie dla kolejnych wierszy. Na koniec zamieniamy kolumny miejscami, żeby niezerowe elementy były na przekątnej.

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

□

Lemat 4.33 można zinterpretować jako mnożenie macierzy elementarnych.

Lemat 4.34. *Każdą macierz odwracalną A wymiaru $n \times n$ można przedstawić jako iloczyn (pewnej liczby) macierzy elementarnych Co więcej, macierze $D_{i\alpha}$ mogą być ostatnie lub pierwsze.*

Każdą macierz A wymiaru $n \times n$ można przedstawić jako iloczyn (pewnej liczby) macierzy elementarnych oraz macierzy postaci $D_{i,0}$ (lub jednej macierzy przekątniowej).

Dowód pozostawiamy jako ćwiczenie.

Rozdział 5

Przekształcenia liniowe i macierze

Wiemy, że każde przekształcenie liniowe z $\mathbb{F}^n$ w $\mathbb{F}^m$ można reprezentować w postaci macierzy rozmiaru $m \times n$ o współczynnikach z ciała $\mathbb{F}$. Z drugiej strony, dla dowolnych przestrzeni liniowych $\mathbb{V}, \mathbb{W}$ nad ciałem $\mathbb{F}$ o wymiarach n, m wiemy, że po wyborze ich baz $B_{\mathbb{V}}, B_W$ są one izomorficzne $\mathbb{F}^n$ i $\mathbb{F}^m$. Tym samym, mając dowolne przekształcenie $F : \mathbb{V} \to \mathbb{W}$ możemy reprezentować je jako macierz — ustalamy bazy $B_{\mathbb{V}}, B_W$, przekształcamy $\mathbb{V}, \mathbb{W}$ izomorficznie na $\mathbb{F}^n$ i $\mathbb{F}^m$ (przy użyciu reprezentacji w bazach $B_{\mathbb{V}}, B_W$) i potem wyrażamy przekształcenie F w tej reprezentacji.

$$\begin{tikzcd}
\mathbb{V} \arrow[r, "F"] \arrow[d, leftrightarrow, "(\cdot)_{B_{\mathbb{V}}}"] & W \arrow[d, leftrightarrow, "(\cdot)_{B_W}"] \\
\mathbb{F}^n \arrow[r, "M"] & \mathbb{F}^m
\end{tikzcd}$$

Okazuje się, że całość można zrobić dużo bardziej systematycznie.

5.1 Wprawka

Wiemy, że dla ustalonych $\mathbb{F}^n, \mathbb{F}^m$ zbiór macierzy M nad $\mathbb{F}$ rozmiaru $m \times n$ jest izomorficzny ze zbiorem wszystkich przekształceń liniowych z $\mathbb{F}^n$ w $\mathbb{F}^m$. W jedną stronę jest łątwo: mając M jego przekształcenie liniowe zadajemy jako

$$L_M(\vec{v}) = M\vec{v} \ .$$

A jak to zrobić w drugą stronę? Dla danego L chcemy skontruwać M takie że $M\vec{v} = L(v)$. Zauważmy, że i-ta kolumna M to $M\vec{E}_i$. Ale M ma spełniać $M\vec{v} = L(\vec{v})$ dla każdego $\vec{v}$, w szczególności dla $\vec{E}_i$. Czyli $L(\vec{E}_i) = M\vec{E}_i$. I tym samym

$$M = [L(\vec{E}_1)|L(\vec{E}_2)|\cdots|L(\vec{E}_n)] \ .$$

5.2 Wyrażanie przekształcenia liniowego w bazie

Definicja 5.1 (Macierz przekształcenia w bazie)**.** Dla przestrzeni liniowych $\mathbb{V}, \mathbb{W}$, przekształcenia liniowego $F : \mathbb{V} \to \mathbb{W}$ oraz $B_{\mathbb{V}}, B_W$: baz odpowiednio $\mathbb{V}$ oraz $\mathbb{W}$, gdzie $B_{\mathbb{V}} = \{v_1, \ldots, v_n\}$ oraz $B_W = \{w_1, \ldots, w_m\}$ macierz $M_{B_{\mathbb{V}} B_W}(F)$ (*macierz przekształcenia F wyrażona w bazach $B_{\mathbb{V}}$ i B_W*) to macierz zadana jako

$$[(F(v_1))_{B_W}|(F(v_2))_{B_W}|\cdots|(F(v_n))_{B_W}] \ .$$

Jest to macierz rozmiaru $m \times n$.

Uwaga. To formalizuje podejście z diagramu powyżej: startujemy z $\mathbb{F}^n$, bierzemy $\vec{E}_i$, przechodzimy do $\mathbb{V}$, czyli mamy v_i, nakładamy F, mamy $F(\vec{v}_i)$ i potem wracamy do $\mathbb{V}$ wyrażając $F(\vec{v}_i)$ w bazie $B_{\mathbb{W}}$.

Uwaga. Zwykle $W = \mathbb{V}$ oraz $B_{\mathbb{V}} = B_W$. Ponadto, dla $\mathbb{V} = \mathbb{F}^n$ i $W = \mathbb{F}^m$ bazami są zwykle bazy standardowe.

Przykład 5.2. Rozważmy przekształcenie $F : \mathbb{R}^3 \to \mathbb{R}^2$ określone jako: $F(x, y, z) = (x + y, y - z)$. Wtedy jego macierz w bazach standardowych dla $\mathbb{R}^3$ oraz $\mathbb{R}^2$ to

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$$

Rozważmy to samo przekształcenie w bazach $\{(1,1,1),(0,1,1),(1,1,0)\}$ oraz $\{(1,1),(1,-1)\}$. Wektory $\{(1,1,1),(0,1,1),(0,1,1)\}$ zostaną przekształcone na odpowiednio:

$$(2,0),(1,0),(2,1) \ ,$$

które wyrażają się w bazie $\{(1,1),(1,-1)\}$ jako

$$\begin{bmatrix} 1 & \frac{1}{2} & 1\frac{1}{2} \\ 1 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \ .$$

Lemat 5.3. *Niech $F : \mathbb{V} \to \mathbb{W}$: przekształcenie oraz $B_{\mathbb{V}}, B_W$ będą bazami odpowiednio $\mathbb{V}$ oraz W, gdzie $B_{\mathbb{V}} = \{v_1, \ldots, v_n\}$ oraz $B_W = \{w_1, \ldots, w_m\}$. Wtedy dla każdego wektora $v \in \mathbb{V}$:*

$$M_{B_{\mathbb{V}}B_W}(F)(v)_{B_{\mathbb{V}}} = (F\vec{v})_{B_W}$$

Dowód. Z definicji jest to prawda dla $v \in B_{\mathbb{V}}$: w takim przypadku $(v)_{B_{\mathbb{V}}}$ jest jednym z wektorów jednostkowych, powiedzmy $\vec{E}_i$, i tym samym $M_{B_{\mathbb{V}}B_W}(F)(v)_{B_{\mathbb{V}}}$ to odpowiednia kolumna $M(F)_{B_{\mathbb{V}}B_W}$, w naszym wypadku i-ta, która jest zadana jako $(F\vec{v})_{B_W}$.

W ogólności wynika to z liniowości: niech $v = \sum \alpha_i v_i$, wtedy

$$\begin{aligned}
M_{B_{\mathbb{V}}B_W}(F)(v)_{B_{\mathbb{V}}} &= M_{B_{\mathbb{V}}B_W}(F)\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{bmatrix} && \text{wyrażenie } v \text{ w bazie } B_{\mathbb{V}}\\
&= M_{B_{\mathbb{V}}B_W}(F)\left(\sum_i \alpha_i \vec{E}_i\right) && \text{Liniowość mnożenia macierzy}\\
&= \sum_i \alpha_i M_{B_{\mathbb{V}}B_W}(F)\vec{E}_i && \text{Liniowość mnożenia macierzy}\\
&= \sum_i \alpha_i (F\vec{v}_i)_{B_W} && i\text{-ta kolumna macierzy } M_{B_{\mathbb{V}}B_W}(F)\\
&= \left(\sum_i \alpha_i F\vec{v}_i\right)_{B_W} && \text{Liniowość wyrażania w bazie}\\
&= \left(F\left(\sum_i \alpha_i v_i\right)\right)_{B_W} && \text{Liniowość } F\\
&= (F\vec{v})_{B_W} && \text{Wyrażanie } v \ \square
\end{aligned}$$

Rozumowanie to przenosi się na macierze oraz na iloczyn macierzy, który odpowiada składaniu przekształceń liniowych.

Lemat 5.4. *Niech $\mathbb{V}, \mathbb{V}', \mathbb{V}''$ będą przestrzeniami liniowymi o bazach B, B', B'', zaś $F : \mathbb{V} \to \mathbb{V}'$, $F' : \mathbb{V}' \to \mathbb{V}''$ przekształceniami liniowymi. Wtedy*

$$M_{BB''}(F'F) = M_{B'B''}(F') \cdot M_{BB'}(F).$$

Dowód. Wystarczy sprawdzić, że na wektorach na wektorach $(v_1)_B, (v_2)_B, \ldots, (v_k)_B$ (czyli na wektorach $\vec{E}_1, \vec{E}_2, \ldots \vec{E}_k$ z bazy standardowej) zachodzi

$$M_{BB''}(F'F)(v_i)_B = M_{B'B''}(F')M_{BB'}(F)(v_i)_B$$

Policzmy prawą stronę:

$$\begin{aligned}
M_{B'B''}(F')M_{BB'}(F)(v_i)_B &= M_{B'B''}(F')(F\vec{v}_i)_{B'} && \text{z Lematu 5.3}\\
&= (F'F\vec{v}_i)_{B''} && \text{z Lematu 5.3}
\end{aligned}$$

Jednocześnie dla lewej strony:

$$M_{BB''}(F'F)(v_i)_B = (F'F\vec{v}_i)_{B''} \qquad \text{z Lematu 5.3}$$

Czyli obie strony są równe (i odpowiadają i-tej kolumnie macierzy $M_{BB''}(F'Fs)$). $\square$

Lemat 5.5. *Niech $F : \mathbb{V} \to \mathbb{W}$ będzie przekształceniem liniowym zaś $B_\mathbb{V}, B_W$ dowolnymi bazami $\mathbb{V}$ oraz W. Wtedy*

$$\mathrm{rk}(F) = \mathrm{rk}(M_{B_\mathbb{V} B_W}(F))$$

Dowód. Zauważmy, że układ wektorów $v_1, v_2, \ldots, v_k$ jest liniowo niezależny wtedy i tylko wtedy gdy dla (dowolnej) bazy B_W układ wektorów (zapisanych jako kolumny) $(v_1)_{B_W}, (v_2)_{B_W}, \ldots, (v_k)_{B_W}$ jest liniowo niezależny: $\sum_i \alpha_i v_i = \vec{0}$ wtedy i tylko wtedy, gdy $\sum_i \alpha_i (v_i)_{B_W} = \vec{0}$.

Pozostaje zaobserwować, że kolumny $M_{B_\mathbb{V} B_W}(F)$ generują obraz (wyrażony w bazie B_W). □

5.3 Macierz zmiany bazy

Jedną z rzeczy, którą możemy w ten sposób wyrazić, jest macierz zmiany bazy: chcemy mieć w miarę jednolity sposób na przejścia z macierzy w jednej bazie do macierzy w innej bazie.

Definicja 5.6 (Macierz zmiany bazy)**.** Dla baz B, B' przestrzeni wektorowej $\mathbb{V}$ *macierz zmiany bazy między B a B'* $M_{BB'}$ to macierz $M_{BB'}(\mathrm{Id})$.

Fakt 5.7. *Dla baz B' oraz $B = v_1, \ldots, v_n$ macierz zmiany bazy zadana jest jako*

$$M_{BB'} = [(v_1)_{B'}|(v_2)_{B'}|\cdots|(v_n)_{B'}] \ .$$

Dowód. Z definicji

$$M_{BB'}(\mathrm{Id}) = [(\mathrm{Id}\, v_1)_{B'}|(\mathrm{Id}\, v_2)_{B'}|\cdots|(\mathrm{Id}\, v_n)_{B'}] = [(v_1)_{B'}|(v_2)_{B'}|\cdots|(v_n)_{B'}] = M_{BB'} \ . \quad \square$$

Lemat 5.8. *Niech $B_\mathbb{V}, B'_\mathbb{V}$ będą bazami $\mathbb{V}$. Wtedy*

$$M_{BB'} M_{B'B} = \mathrm{Id},$$

tzn. są to macierze odwrotne.

Dowód. Policzmy

$$\begin{aligned} M_{BB'} M_{B'B} &= M_{BB'}(\mathrm{Id}) M_{B'B}(\mathrm{Id}) && \text{definicja} \\ &= M_{B'B'}(\mathrm{Id} \circ \mathrm{Id}) && \text{Lemat5.4} \\ &= M_{B'B'}(\mathrm{Id}) \end{aligned}$$

Teraz zgodnie z definicją i-tą kolumną $M_{B'B'}(\mathrm{Id})$ jest $(\mathrm{Id}\, v_i)_{B'} = (v_i)_{B'}$, gdzie v_i jest i-tym wektorem z bazy B'. Ale reprezentacja v_i w bazie B' to po prostu $\vec{E}_i$. Czyli

$$M_{B'B'} = [\vec{E}_1|\vec{E}_2|\cdots|\vec{E}_n] = \mathrm{Id} \quad \square$$

Lemat 5.9. *Niech $F : \mathbb{V} \to \mathbb{W}$ będzie przekształceniem liniowym, $B_\mathbb{V}, B'_\mathbb{V}$ bazami $\mathbb{V}$ zaś B_W, B'_W bazami W. Wtedy*

$$M_{B_\mathbb{V} B_W}(F) = M_{B'_W B_W} M_{B'_\mathbb{V} B'_W}(F) M_{B_\mathbb{V} B'_\mathbb{V}}.$$

Dowód. Korzystamy z Lematu 5.4 dla złożenia trzech przekształceń $\mathrm{Id} \circ F \circ \mathrm{Id}$ wyrażonych w odpowiednich bazach

$$\begin{aligned} M_{B'_W B_W} M_{B'_\mathbb{V} B'_W}(F) M_{B_\mathbb{V} B'_\mathbb{V}} &= M_{B'_W B_W}(\mathrm{Id}) M_{B'_\mathbb{V} B'_W}(F) M_{B_\mathbb{V} B'_\mathbb{V}}(\mathrm{Id}) \\ &= M_{B_\mathbb{V} B_W}(\mathrm{Id} \circ F \circ \mathrm{Id}) \\ &= M_{B_\mathbb{V} B_W}(F) \end{aligned}$$

□

Uwaga. Najczęściej będziemy zajmować się przypadkiem, gdy $\mathbb{W} = \mathbb{V}$ i $B_\mathbb{V} = B_W$ i $B'_\mathbb{V} = B'_W$.

Przykład 5.10. W $\mathbb{R}^3$ rozpatrzmy bazę standardową E oraz bazę B: $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$.

Wtedy

$$M_{BE} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, M_{EB} = \begin{bmatrix} 0,5 & 0,5 & -0,5 \\ 0,5 & -0,5 & 0,5 \\ -0,5 & 0,5 & 0,5 \end{bmatrix}$$

Można łatwo sprawdzić, że

$$M_{EB}M_{BE} = \mathrm{Id}$$

Rozpatrzmy przekształcenie F, (wyrażone w bazie standardowej) jako

$$M_{EE}(F) = \begin{bmatrix} 4 & 0 & 0 \\ -1 & 5 & 1 \\ -1 & 1 & 5 \end{bmatrix}$$

Wtedy

$$\begin{aligned} M_{BB}(F) &= M_{EB}M_{EE}(F)M_{BE} \\ &= \begin{bmatrix} 0,5 & 0,5 & -0,5 \\ 0,5 & -0,5 & 0,5 \\ -0,5 & 0,5 & 0,5 \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 \\ -1 & 5 & 1 \\ -1 & 1 & 5 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{bmatrix}. \end{aligned}$$

Oraz

$$\begin{aligned} M_{EE}(F) &= M_{BE}M_{BB}(F)M_{EB} \\ &= \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} 0,5 & 0,5 & -0,5 \\ 0,5 & -0,5 & 0,5 \\ -0,5 & 0,5 & 0,5 \end{bmatrix} \\ &= \begin{bmatrix} 4 & 0 & 0 \\ -1 & 5 & 1 \\ -1 & 1 & 5 \end{bmatrix}. \end{aligned}$$

Rozdział 6

Wyznacznik

6.1 Wyznacznik

Ważna funkcja na macierzach: *wyznacznik*. Uogólnienie objętości (ale ze znakiem).

Jakie własności powinna mieć objętość na zbiorze n wektorów $v_1, v_2, \ldots, v_n$ z $\mathbb{F}$ w $\mathbb{F}$? ($\det : (\mathbb{F}^n)^n \to \mathbb{F}$):

(W1) (liniowość) jest funkcją wielo-liniową, tj. liniową dla każdej kolumny:

$$\begin{aligned}
\det(v_1, v_2, \ldots, v_{i-1}, \alpha v_i, v_{i+1}, \ldots, v_n) &= \alpha \det(v_1, v_2, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n) \\
\det(v_1, v_2, \ldots, v_{i-1}, v_i + v_i', v_{i+1}, \ldots, v_n) &= \det(v_1, v_2, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n) \\
&\quad + \det(v_1, v_2, \ldots, v_{i-1}, v_i', v_{i+1}, \ldots, v_n)
\end{aligned}$$

W szczególności

$$\det(v_1, v_2, \ldots, v_{i-1}, \vec{0}, v_{i+1}, \ldots, v_n) = 0$$

(W2) zastąpienie v_i przez $v_i + \sum_{j \neq i} \alpha_j v_j$ nie powinno zmieniać wartości

$$\det(v_1, v_2, \ldots, v_{i-1}, v_i + \sum_{j \neq i} \alpha_j v_j, v_{i+1}, \ldots, v_n) = \det(v_1, v_2, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n)$$

(W3) zamiana kolejności dwóch wektorów zmienia znak (objętość ze znakiem)

$$\det(v_1, \ldots, v_n) = -\det(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1} \ldots, v_n)$$

(W4) na macierzy identycznościowej to jest 1

$$\det(\mathrm{Id}) = 1$$

Jest to tak zwana „aksjomatyczna definicja wyznacznika”.

Uargumentujemy, że taka funkcja jest najwyżej jedna oraz metodę jej liczenia. Formalnie, należałoby pokazać, że taka funkcja w ogóle istnieje. Jej definicja jest dość techniczna, zostanie przedstawiona później (ale już teraz poznamy wszystkie techniki, aby ją liczyć.)

Lemat 6.1. *Jest dokładnie jedna funkcja spełniająca warunki W1–W4.*

Dowód. Zauważmy, że warunki te oznaczają, że wartość macierzy zmienia się w prosty sposób przy stosowaniu operacji elementarnych. Czyli możemy stosować eliminację Gaussa (być może znak się zmienia przy zmianie kolejności). Jeśli układ wektorów jest zależny, to otrzymamy kolumnę zerową i tym samym wyznacznik to 0. Jeśli nie, to uzyskamy macierz górnotrójkątną bez 0 na przekątnej. Można ją przekształcić do macierzy przekątniowej przy użyciu operacji elementarnych. A dla niej to jest iloczyn wartości na przekątnej. □

Definicja 6.2 (Wyznacznik)**.** *Wyznacznik* macierzy kwadratowej $\det(A) = |A|$. To jedyna funkcja spełniająca warunki W1–W4. Oznaczamy go też przez przez $|A|$.

6.2 Własności i metody obliczania wyznacznika

Fakt 6.3. *Proste własności wyznacznika*

- *Jeśli* $v_i = v_j$ *to* $\det(v_1, v_2, \dots, v_n) = 0$.
- *Dla macierzy trójkątnej jest to iloczyn elementów na przekątnej.*
- $\det(A) \neq 0 \iff \mathrm{rk}(A) = n$

Definicja 6.4 (Minor macierzy). *Minorem macierzy M* nazywamy każdą macierz uzyskaną poprzez usunięcie z M pewnego zbioru wierszy i kolumn.

Zwyczajowo $A_{i,j}$ to macierz powstała z A poprzez usunięcie i-tego wiersza oraz j-tej kolumny.

Definicja 6.5 (Dopełnienie algebraiczne). Dopełnienie algebraiczne elementu $a_{i,j}$ to $(-1)^{i+j}\det(A_{i,j})$.

Fakt 6.6 (Rozwinięcie Laplace'a). *Dla macierzy kwadratowej $A = (a_{ij})_{i,j=1,\dots,n}$ mamy:*

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

Dowód. Ogólną wersję pozostawiamy jako ćwiczenie, tu pokażemy dowód dla $j = 1$. Z liniowości po pierwszej współrzędnej:

$$\begin{aligned}
\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}
&= \sum_{i=1}^{n} a_{i1} \begin{vmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{vmatrix} && \text{z liniowości} \\
&= \sum_{i=1}^{n} a_{i1} \begin{vmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{vmatrix} && \text{od kolumny } j > 1 \text{ odejmujemy } a_{ij} \text{ razy pierwszą} \\
&= \sum_{i=1}^{n} \begin{vmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{vmatrix} && \text{z liniowości}
\end{aligned}$$

Używając $i-1$ zamian możemy wprowadzić a_{i1} na przekątną.

$$\sum_{i=1}^{n} \begin{vmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \sum_{i=1}^{n} (-1)^{i+1} \begin{vmatrix} a_{12} & \cdots & 0 & \cdots & a_{1n} \\ a_{22} & \cdots & 0 & \cdots & a_{2n} \\ \vdots & \ddots & \vdots & \cdots & \vdots \\ 0 & \cdots & a_{i1} & \cdots & 0 \\ \vdots & \vdots & 0 & \ddots & \vdots \\ a_{n2} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} .$$

Zauważmy, że możemy na tej macierzy przeprowadzić eliminację Gaussa, ignorując i-ty wiersz i kolumnę. Dostajemy macierz górnotrójkątną, której wyznacznik to iloczyn elementów na przekątnej.

$$\sum_{i=1}^{n}(-1)^{i+1}\begin{vmatrix} a_{12} & \cdots & 0 & \cdots & a_{1n} \\ a_{22} & \cdots & 0 & \cdots & a_{2n} \\ \vdots & \ddots & \vdots & \cdots & \vdots \\ 0 & \cdots & a_{i1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n2} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} = \sum_{i=1}^{n}(-1)^{i+1}a_{i1}|A_{i1}| \ . \qquad \square$$

Rozwinięcie Laplace'a pozwala nam na podanie konkretnych wzorów na wyznacznik macierzy 2×2 oraz 3×3.

Przykład 6.7 (Obliczanie małych wyznaczników). Łatwo obliczyć, że wyznacznik macierzy 2×2, zadanej jako $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ to

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \ .$$

W przypadku macierzy 3×3 możemy zastosować metodę Sarrusa.

$$\begin{array}{|ccc|cc} + & + & + & & \\ a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \\ - & - & - & & \end{array}$$

Twierdzenie 6.8 (Cauchy).

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

Dowód. Teza łatwo zachodzi, jeśli $|A| = 0$ lub $|B| = 0$: odpowiada to sytuacji, w której $\text{rk}(A) < n$ lub $\text{rk}(B) < n$. A wtedy też $\text{rk}(AB) \leq \min(\text{rk}(A), \text{rk}(B)) < n$. Czyli $|AB| = 0$.

W dalszym dowodzie możemy zakładać, że macierze są rzędu n. W takim razie zgodnie z Lematem 4.34 B można przedstawić jako iloczyn macierzy elementarnych.

Pokażemy najpierw, że

$$|AB| = |A| \cdot |B| \ ,$$

gdzie B jest macierzą elementarną.

- Macierz T_{ij}: przez zamianę i-tej i j-tej kolumny dostajemy Id, czyli jej wyznacznik to -1. Jednocześnie przemnożenie A przez T_{ij} zamienia miejscami 2 kolumny, czyli zmienia znak wyznacznika na przeciwny.
- Macierz $\text{Id} + \alpha 1_{ij}$. Przemnożenie przez tą macierz dodaje wielokrotność kolumny do innej kolumny, czyli zgodnie z definicją nie zmienia wartości wyznacznika.

  Jednocześnie w $\text{Id} + \alpha 1_{ij}$ dodając do j-tej kolumny α razy i-tą usuwamy niezerowy element poza przekątną, otrzymując Id. Czyli $|\text{Id} + \alpha 1_{ij}| = 1$.
- Macierz $D_{i\alpha}$ przemnaża i-tą kolumnę α razy, jednocześnie $|D_{i\alpha}| = \alpha$.

Ale to jest proste, bo odpowiada to operacji elementarnej (kolumnowe) na macierzy A.

Wracając do dowodu. Przez indukcję łatwo stwierdzamy, że dla macierzy elementarnych $E_1, \ldots, E_m$

$$\left|\prod_{i=1}^{m} E_i\right| = \prod_{i=1}^{m} |E_i| \ . \qquad (6.1)$$

Zgodnie z Lematem 4.34 zarówno A jak i B są iloczynami macierzy elementarnych, niech

$$A = \prod_{i=1}^{m} E_i \quad B = \prod_{i=m+1}^{m+m'} E_i \ .$$

Wtedy

$$\begin{aligned}
|AB| &= \left| \left(\prod_{i=1}^{m} E_i\right) \cdot \left(\prod_{i=m+1}^{m+m'} E_i\right) \right| \\
&= \left| \prod_{i=1}^{m+m'} E_i \right| \\
&= \prod_{i=1}^{m+m'} |E_i| && \text{z (6.1)} \\
&= \left(\prod_{i=1}^{m} |E_i|\right) \cdot \left(\prod_{i=m+1}^{m+m'} |E_i|\right) \\
&= \left(\left|\prod_{i=1}^{m} E_i\right|\right) \cdot \left(\left|\prod_{i=m+1}^{m+m'} |E_i|\right|\right) && \text{z (6.1)} \\
&= |A| \cdot |B| \ .
\end{aligned}$$

□

Jest wiele innych dowodów, wszystkie wymagają trochę sprytu lub obserwacji.

Fakt 6.9. *Wyznacznik macierzy oraz macierzy transponowanej jest taki sam, tj.:*

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

Dowód. Jeśli $\det(A) = 0$ to $\mathrm{rk}(A) < n$ i wtedy $\mathrm{rk}(A^T) < n$ i $\det(A^T) = 0$. Czyli wystarczy rozważyć przypadek, gdy $\det(A) \neq 0$.

Dowód wynika z Tw. Cauchyego oraz Lematu 4.34: każdą nieosobliwą macierz A można przedstawić jako iloczyn macierzy elementarnych.

$$A = \prod_{i=1}^{k} M_i \ .$$

Wtedy $A^T = \prod_{i=1}^{k} M_{k-i+1}^T$. Łatwo sprawdzić, że macierzy elementarnej mamy

$$|M_i| = |M_i^T|.$$

Co daje

$$\begin{aligned}
|A^T| &= \left| \prod_{i=1}^{k} M_{k-i+1}^T \right| \\
&= \prod_{i=1}^{k} |M_{k-i+1}^T| \\
&= \prod_{i=1}^{k} |M_{k-i+1}| \\
&= \prod_{i=1}^{k} |M_i| \\
&= \left| \prod_{i=1}^{k} M_i \right| \\
&= |A| \ .
\end{aligned}$$

□

Zauważmy, że w konsekwencji operacje wierszowe zmieniają wartość wyznacznika tak samo, jak operacje kolumnowe: aby wykonać operację wierszową, możemy transponować macierz, wykonać odpowiadającą operację kolumnową i z powrotem transponować macierz; zaś transponowanie nie zmienia wyznacznika. W szczególności, w trakcie obliczania wyznacznika możemy używać jednych i drugich operacji.

Fakt 6.10.

- *Dodanie do wiersza macierzy wielokrotności innego wiersza nie zmienia wyznacznika.*
- *Wyznacznik macierzy z zerowym wierszem jest równy* 0.
- *Wyznacznik jest funkcją wieloliniową wierszy.*
- *Zamiana dwóch wierszy miejscami zmienia znak wyznacznika na przeciwny.*

Przykład 6.11 (Wyznacznik macierzy Vandermonde'a). Niech $q_1, q_2, \dots, q_n$ będą dowolnymi liczbami. Macierz $(n \times n)$ Vandermonde'a V_n ma wyrazy równe $v_{ij} = q_i^{j-1}$, tj.:

$$V_n = \begin{bmatrix} 1 & q_1 & q_1^2 & \dots & q_1^{n-1} \\ 1 & q_2 & q_2^2 & \dots & q_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & q_n & q_n^2 & \dots & q_n^{n-1} \end{bmatrix} .$$

Pokażemy, że

$$\det(V_n) = \prod_{1 \leq i < j \leq n} (q_j - q_i) .$$

W szczególności implikuje to, jeśli q_i są niezerowe i parami różne, to wyznacznik ten jest niezerowy.

Najpierw odejmujemy pierwszy rząd od każdego kolejnego, dostając

$$\begin{vmatrix} 1 & q_1 & q_1^2 & \dots & q_1^{n-1} \\ 0 & q_2 - q_1 & q_2^2 - q_1^2 & \dots & q_2^{n-1} - q_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & q_n - q_1 & q_n^2 - q_1^2 & \dots & q_n^{n-1} - q_1^{n-1} \end{vmatrix} .$$

Używamy rozwinięcia Laplace'a dla pierwszej kolumny: jedyny niezerowy wyraz w niej to $a_{11} = 1$, czyli

$$\begin{vmatrix} 1 & q_1 & q_1^2 & \dots & q_1^{n-1} \\ 0 & q_2 - q_1 & q_2^2 - q_1^2 & \dots & q_2^{n-1} - q_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & q_n - q_1 & q_n^2 - q_1^2 & \dots & q_n^{n-1} - q_1^{n-1} \end{vmatrix} = \begin{vmatrix} q_2 - q_1 & q_2^2 - q_1^2 & \dots & q_2^{n-1} - q_1^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ q_n - q_1 & q_n^2 - q_1^2 & \dots & q_n^{n-1} - q_1^{n-1} \end{vmatrix} .$$

Teraz od i kolumny odejmujemy q_1 razy $i - 1$-szą, zaczynajac od prawej strony (czyli eliminacja niezerowych elementów w górnym wierszu)

$$\begin{vmatrix} (q_2 - q_1) & (q_2^2 - q_1^2) - q_1(q_2 - q_1) & \dots & (q_2^{n-1} - q_1^{n-1}) - q_1(q_2^{n-2} - q_1^{n-2}) \\ \vdots & \vdots & \ddots & \vdots \\ (q_n - q_1) & (q_n^2 - q_1^2) - q_1(q_n - q_1) & \dots & (q_n^{n-1} - q_1^{n-1}) - q_1(q_n^{n-2} - q_1^{n-2}) \end{vmatrix} .$$

Po rozwinięciu odpowiednie wyrazy skracają się i dostajemy

$$\begin{vmatrix} q_2 - q_1 & q_2^2 - q_1 q_2 & \dots & q_2^{n-1} - q_1 q_2^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ q_n - q_1 & q_n^2 - q_1 q_n & \dots & q_n^{n-1} - q_1 q_n^{n-2} \end{vmatrix} = \begin{vmatrix} (q_2 - q_1) \cdot 1 & (q_2 - q_1) \cdot q_2 & \dots & (q_2 - q_1) \cdot q_2^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ (q_n - q_1) \cdot 1 & (q_n - q_1) \cdot q_n & \dots & (q_n - q_1) \cdot q_n^{n-2} \end{vmatrix} .$$

Teraz z liniowości wyjmujemy przed wyznacznik $(q_2 - q_1) \cdots (q_n - q_1)$ i dostajemy

$$\prod_{i=2}^{n} (q_i - q_1) \begin{vmatrix} 1 & q_2 & \dots & q_2^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & q_n & \dots & q_n^{n-2} \end{vmatrix}$$

i teraz przez indukcję.

6.3 Wyznacznik a macierz odwrotna

Fakt 6.12. *Jeśli M jest odwracalna, to*

$$\det(M^{-1}) = \frac{1}{\det(M)} \quad .$$

Lemat 6.13. *Macierz odwrotna do macierzy M jest równa*

$$\frac{1}{\det(A)} C^T, \text{ gdzie } c_{ij} = (-1)^{i+j}|A_{i,j}| \quad .$$

Dowód. Rozważmy element i, j w mnożeniu macierzy ze sformułowania lematu oraz macierzy A:

$$\frac{1}{|A|} \sum_{k=1}^{n} (-1)^{i+k}|A_{ki}|a_{kj} \quad .$$

Z rozwinięcia Laplace'a to jest wyznacznik macierzy A w której w i-tej kolumnie zastąpiliśmy C_i przez C_j (i przemnożyliśmy wszystko przez $\frac{1}{|A|}$). Dla $i = j$ to daje $\det(A)/\det(A)$, dla $i \neq j$ to daje 0. □

Przykład 6.14. Dzięki Lematowi 6.13 można np. łatwo policzyć macierz odwrotną do macierzy 2×2:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \quad .$$

6.4 Wyznacznik przekształcenia

Potrafimy zdefiniować wyznacznik dla macierzy, ale co z przekształceniem liniowym? Każde przekształcenie zadaje macierz, ale ta macierz zależy od bazy. Okazuje się, że wartość wyznacznika nie.

Lemat 6.15. *Niech $F : V \to V$ będzie przekształceniem liniowym, zaś M, M' będą macierzami dla tego przekształcenia wyrażonymi w rożnych bazach. Wtedy*

$$|M| = |M'|.$$

Dowód. Niech $M_{BB'}$ będzie macierzą przejścia z jednej bazy do drugiej, zaś $M_{B'B}$ z drugiej do pierwszej. Przypomnijmy, że $M_{BB'}M_{B'B} = \mathrm{Id}$. Wtedy

$$\begin{aligned}
\det(M') &= \det(M_{BB'} M M_{B'B}) \\
&= \det(M_{BB'}) \det(M) \det(M_{B'B}) \\
&= \det(M_{BB'}) \det(M_{B'B}) \det(M) \\
&= \frac{1}{\det(M_{BB'})} \det(M_{B'B}) \det(M) \\
&= \det(M) \quad .
\end{aligned}$$

□

Definicja 6.16. Dla przekształcenia liniowego $F : \mathbb{V} \to \mathbb{V}$ jego wyznacznik $\det(F)$ to $\det(M)$ gdzie M jest macierzą tego przekształcenia wyrażoną w dowolnej bazie $\mathbb{V}$.

Rozdział 7

Układy równań liniowych i ich rozwiązywanie

Rozważmy układ równań postaci

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

Będziemy zapisywać równania w postaci

$$A\vec{X} = \vec{B}\ , \tag{7.1}$$

gdzie

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \vec{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \vec{B} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

7.1 Bazowy przypadek: n zmiennych, n równań, macierz odwracalna

Intuicja: w najprostszym przypadku, gdy A jest macierzą kwadratową, możemy odwrócić A i nałożyć obustronnie na równanie, uzyskując

$$A^{-1}AX = \mathrm{Id}\, X = X = A^{-1}\vec{B}.$$

I tym samym mamy rozwiązanie. Można łatwo sprawdzić, że jest to jedyne rozwiązanie.

Pokażemy teraz, jak wygląda to rozwiązanie.

Twierdzenie 7.1 (Wzory Cramera)**.** *Jeśli w równaniu* (7.1) *macierz A jest kwadratowa i odwracalna, to jedyne rozwiązanie jest postaci $x_i = \frac{\det(A_{x_i})}{\det(A)}$, gdzie macierz A_{x_i} powstaje poprzez zastąpienie i-tej kolumny A przez $\vec{B}$.*

W szczególności, jeśli $\det(A) \neq 0$ *to równanie ma jedno rozwiązanie.*

Dowód. Chcemy policzyć

$$\det(A_{x_i}) = \det(A_1, \dots, \underbrace{B}_{i\text{-te miejsce}}, \dots, A_n)$$

Mamy

$$B = AX = A\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \sum_i x_i A\vec{E}_i = \sum_i x_i A_i$$

Czyli

$$\begin{aligned}
\det(A_{x_i}) &= \det(A_1, \dots, \underbrace{B}_{i\text{-te miejsce}}, \dots, A_n) \\
&= \det(A_1, \dots, \underbrace{\sum_j x_j A_j}_{i\text{-te miejsce}}, \dots, A_n) \\
&= \sum_j x_j \det(A_1, \dots, \underbrace{A_j}_{i\text{-te miejsce}}, \dots, A_n)
\end{aligned}$$

Zauważmy, że jeśli $i \neq j$ to w powyższej macierzy mamy dwie takie same kolumny,czyli wyznacznik się zeruje.

$$\begin{aligned}
\sum_j x_j \det(A_1, \dots, \underbrace{A_j}_{i\text{-te miejsce}}, \dots, A_n) &= x_i \det(A_1, \dots, \underbrace{A_i}_{i\text{-te miejsce}}, \dots, A_n) \\
&= x_i \det(A_i) \ .
\end{aligned}$$

□

Zauważmy, że wiemy już, że macierz odwrotną do A można wyrazić przez dopełnienia algebraiczne (Lemat 6.13). Nakładając ją na wektor B można otrzymać wzory Cramera bezpośrednio.

7.2 Ogólne układy równań liniowych

Chcemy jednak zająć się tym problemem w większej ogólności:

- co jeśli A nie jest odwracalna? Czy wtedy rozwiązań jest wiele, czy może 0?
- co jeśli A nie jest kwadratowa (w szczególności: nieodwracalna)? Czym różnią się przypadki:
  - jest więcej równań, niż zmiennych?
  - jest więcej zmiennych, niż równań?

7.2.1 Układy jednorodne

Zajmijmy się trochę mniej ogólnym problemem: co jeśli $B = \vec{0}$? Taki układ nazywamy *jednorodnym*. Jedno rozwiązanie na pewno jest.

Lemat 7.2 (Układ jednorodny). *Zbiór wszystkich rozwiązań równania*

$$AX = \vec{0}$$

jest przestrzenią liniową, jest to $\ker(A)$, *gdy A traktujemy jako przekształcenie liniowe z $\mathbb{F}^n$ w $\mathbb{F}^m$. Wymiar tej przestrzeni to $n - \mathrm{rk}(A)$.*

Dowód. Wystarczy potraktować A jako przekształcenie liniowe. Wtedy zbiór rozwiązań to dokładnie jądro tego przekształcenia. □

7.2.2 Układy niejednorodne

Fakt 7.3.

$$AX = B \text{ ma rozwiązanie} \iff B \in \mathrm{Im}(A) \ .$$

Jeśli równanie $AX = B$ ma rozwiązanie to zbiór wszystkich jego rozwiązań jest warstwą względem $\ker A$.

Uwaga. Jeśli ciało $\mathbb{F}$ jest nieskończone, to w tym przypadku jest nieskończenie wiele rozwiązań. W innym przypadku jest to $|\mathbb{F}|^k$, gdzie k jest wymiarem jądra.

Dowód. $\Rightarrow$ Jeśli X_0 jest rozwiązaniem, to $AX_0 = B$, czyli w szczególności $B \in \mathrm{Im}(A)$.

$\Leftarrow$ Jeśli $B \in \mathrm{Im}(A)$, to istnieje X_0, że $AX_0 = B$.

Ustalmy dowolne rozwiązanie $\vec{X_0}$. Wtedy dla dowolnego $\vec{X}$:

$$\begin{aligned}
AX = B \iff AX = AX_0 \iff A(X - X_0) = \vec{0} \iff X - X_0 \in \ker(A) \\
\iff X, X_0 \text{ są w tej samej warstwie } \ker(A) \ .
\end{aligned}$$

□

Fakt 7.4 (Tw. Kronecker-Capelli). *Układ*

$$AX = B$$

ma rozwiązanie $\iff \operatorname{rk}(A|B) = \operatorname{rk}(A)$.

Macierz $[A|B]$ nazywana jest czasem *macierzą rozszerzoną* układu $AX = B$.

Dowód. $\Leftarrow$ Jeśli $\operatorname{rk}(A|B) = \operatorname{rk}(A)$ to znaczy, że B jest kombinację kolumn z A, czyli jest w obrazie A.
$\Rightarrow$ Jeśli $\operatorname{rk}(A|B) > \operatorname{rk}(A)$ to B nie jest w obrazie A, czyli równanie nie ma rozwiązania. □

Uwaga. Liczenie *osobno* rzędów $A|B$ oraz A jest zwykle nadmiarowe: jeśli użyjemy eliminacji wierszowej, to automatycznie dostaniemy informację, jaki jest rząd A a jaki $A|B$. W eliminacji kolumnowej również jest to prawda, o ile nie użyjemy B do eliminowania innych kolumn.

Co więcej, jeśli zastosujemy eliminację Gaußa na wierszach lub kolumnach A (otrzymując A') to po zastosowaniu tych samych operacji na $A|B$ uzyskamy $A'|B'$ (i B' trzeba osobno policzyć).

W obu wypadkach nie ma potrzeby wykonywanie tych samych przekształceń wielokrotnie.

Przykład 7.5. Ile rozwiązań, w zależności od parametru λ, ma podany układ równań?

$$\begin{cases} 3x_1 & -x_2 & +4x_3 & = & 1 \\ 5x_1 & -2x_2 & +6x_3 & = & 1+\lambda \\ (6+\lambda^2)x_1 & -3x_2 & +(9-\lambda^2)x_3 & = & 3 \end{cases} .$$

Podany układ równań zapisany w postaci macierzowej wygląda następująco

$$\begin{bmatrix} 3 & -1 & 4 \\ 5 & -2 & 6 \\ (6+\lambda^2) & -3 & (9-\lambda^2) \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1+\lambda \\ 3 \end{bmatrix}$$

Jeśli wyznacznik macierzy głównej jest niezerowy, to ma on dokładnie jedno rozwiązanie. Policzmy więc wartość tego wyznacznika, użyjemy metody Sarrusa:

$$\left| \begin{matrix} 3 & -1 & 4 \\ 5 & -2 & 6 \\ (6+\lambda^2) & -3 & (9-\lambda^2) \end{matrix} \right| \begin{matrix} 3 & -1 \\ 5 & -2 \\ (6+\lambda^2) & -3 \end{matrix} =$$

$$3\cdot(-2)\cdot(9-\lambda^2) + (-1)\cdot 6\cdot(6+\lambda^2) + 4\cdot 5\cdot(-3) - 4\cdot(-2)\cdot(6+\lambda^2) - 3\cdot 6\cdot(-3) - (-1)\cdot 5\cdot(9-\lambda^2) =$$

$$-54 + 6\lambda^2 - 36 - 6\lambda^2 - 60 + 48 + 8\lambda^2 + 54 + 45 + 5\lambda^2 =$$

$$-13 + 13\lambda^2 = 13(\lambda^2 - 1)$$

Zauważmy, że wartość wyznacznika głównego tego układu równań jest niezerowa dla $\lambda \notin \{1, -1\}$, czyli dla takich wartości układ równań ma dokładnie jedno rozwiązanie.

Rozważmy więc pozostałe wartości. Zastosujemy w nich twierdzenia Kroneckera-Capellego: w tym celu musimy policzyć rząd macierzy głównej oraz rząd macierzy rozszerzonej tego układu. Rząd macierzy głównej jest taki sam dla $\lambda = 1$ oraz $\lambda = -1$:

$$\begin{bmatrix} 3 & -1 & 4 \\ 5 & -2 & 6 \\ 7 & -3 & 8 \end{bmatrix} \xrightarrow{(3)-(2),(2)-(1)} \begin{bmatrix} 3 & -1 & 4 \\ 2 & -1 & 2 \\ 2 & -1 & 2 \end{bmatrix}$$

Łatwo zauważyć, że ma ona rząd 2: wiersz drugi i trzeci są identyczne, zaś pierwszy i drugi różne (i mają tą samą drugą współrzędną).

Niech $\lambda = 1$, rozważamy macierz rozszerzoną, wykonujemy na niej takie same operacje, jak powyżej na macierzy głównej:

$$\begin{bmatrix} 3 & -1 & 4 & 1 \\ 5 & -2 & 6 & 2 \\ 7 & -3 & 8 & 3 \end{bmatrix} \xrightarrow{(3)-(2),(2)-(1)} \begin{bmatrix} 3 & -1 & 4 & 1 \\ 2 & -1 & 2 & 1 \\ 2 & -1 & 2 & 1 \end{bmatrix}$$

Rząd tej macierzy również wynosi 2, gdyż, jak powyżej, wiersz drugi i trzeci są identyczne, zaś pierwszy i drugi: różne i mają taką samą drugą współrzędną. Czyli rząd macierzy głównej i macierzy rozszerzonej jest taki sam i z tw. Kroneckera-Capellego ten układ równań ma nieskończenie wiele rozwiązań.

Dla $\lambda = -1$

$$\begin{bmatrix} 3 & -1 & 4 & 1 \\ 5 & -2 & 6 & 0 \\ 7 & -3 & 8 & 3 \end{bmatrix} \xrightarrow{(3)-(2),(2)-(1)} \begin{bmatrix} 3 & -1 & 4 & 1 \\ 2 & -1 & 2 & -1 \\ 2 & -1 & 2 & 1 \end{bmatrix} \xrightarrow{(3)-(2)} \begin{bmatrix} 3 & -1 & 4 & 1 \\ 2 & -1 & 2 & -1 \\ 0 & 0 & 0 & 2 \end{bmatrix} \xrightarrow{(2)-\frac{1}{2}(3),(1)-\frac{1}{2}(3)} \begin{bmatrix} 3 & -1 & 4 & 0 \\ 2 & -1 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$
$$\xrightarrow{(1)-(2)} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 2 & -1 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \xrightarrow{(2)-(1)} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

Łatwo zauważyć, że rząd wynosi 3. Czyli rząd macierzy głównej jest mniejszy niż rząd macierzy rozszerzonej i z tw. Kroneckera-Capellego ten układ równań nie ma rozwiązań.

7.3 Metoda eliminacji Gaussa

Definicja 7.6 (Układy równoważne)**.** Układy równań $AX = B$ oraz $A'X = B'$ są *równoważne*, jeśli mają ten sam zbiór rozwiązań.

Jak to policzyć wydajnie?

Lemat 7.7. *Rozważmy układ równań $AX = B$. Układ uzyskany przez następujące operacje przeprowadzone na macierzy rozszerzonej układu:*

- *zamianę i-tego oraz j-tego równania*
- *dodanie do j-tego równania wielokrotności i-tego*
- *przemnożenie i-tego równania przez stałą $\alpha \neq 0$*

dają układ równoważny wejściowemu.

Prosty dowód pozostawimy jako ćwiczenie.

Oznacza to, że możemy stosować metodę eliminacji (wierszowej) na równaniu. Na końcu dostajemy macierz w postaci schodkowej (wierszowo).

Wtedy

- Układ jest sprzeczny: Jeśli w wierszu z samymi współczynnikami zerowymi prawa strona jest niezerowa. Z tw. Kroneckera-Capelliego rząd (wierszowy) macierzy rozszerzonej jest większy, niż macierzy głównej.

  Intuicyjnie odpowiada to sytuacji, że mamy te same równania i różne wartości po prawej stronie. Nie ma nic więcej do zrobienia.

- Układ ma jedno rozwiązanie: Jeśli uzyskaliśmy macierz (równań) górnotrójkątną plus być może zerowe wiersze poniżej, ponadto na przekątnej nie ma zer oraz wartości odpowiadające wierszom zerowym to też zera, to jest dokładnie jedno rozwiązanie. Dowód wynika z tego, że możemy odrzucić zerowe równania (ukłąd pozostaje równoważny) i wtedy mamy macierz kwadratową i możemy nałożyć macierz odwrotną (alternatywnie: macierz jest odwracalna, czyli ma trywialne jądro, czyli warstwa ma jeden element). W tym przypadku łatwo podać rozwiązanie (wyliczamy kolejne wartości i wstawiamy do równań powyżej).

- W przeciwnym przypadku, już wcześniej powiedzieliśmy, ile tych rozwiązań jest (warstwa jądra). Umiemy policzyć to jądro, chcemy jeszcze jedno rozwiązanie szczególne. W tym celu możemy ustalić (dowolnie) wartość zmiennej, która nie odpowiada pierwszej wyróżnionej pozycji w wierszu. To przekształci macierz równania do postaci trójkątnej (pierwszy przypadek).

Przykład 7.8 (Kontynuacja Przykładu 7.5). Przypomnijmy, że chcemy sprawdzić, ile rozwiązań, w zależności od parametru λ, ma układ:

$$\begin{cases} 3x_1 & -x_2 & +4x_3 & = & 1 \\ 5x_1 & -2x_2 & +6x_3 & = & 1+\lambda \\ (6+\lambda^2)x_1 & -3x_2 & +(9-\lambda^2)x_3 & = & 3 \end{cases}.$$

Użyjemy tym razem eliminacji Gaußa: od trzeciego wiersza odejmujemy drugi, a od drugiego: pierwszy.

$$\begin{cases} 3x_1 & -x_2 & +4x_3 & = & 1 \\ 2x_1 & -x_2 & +2x_3 & = & \lambda \\ (1+\lambda^2)x_1 & -x_2 & +(3-\lambda^2)x_3 & = & 2-\lambda \end{cases}.$$

Teraz od trzeciego odejmujemy drugi:

$$\begin{cases} 3x_1 & -x_2 & +4x_3 & = & 1 \\ 2x_1 & -x_2 & +2x_3 & = & 1+\lambda \\ (\lambda^2-1)x_1 & & +(1-\lambda^2)x_3 & = & 2(1-\lambda) \end{cases}.$$

Łatwo zauważyć, że dla $\lambda = -1$ trzecie równanie jest sprzeczne ($0 = 4$), zaś dla $\lambda = 1$ jest puste ($0 = 0$).

Rozważmy dokładniej przypadek $\lambda = 1$.

$$\begin{cases} 3x_1 & -x_2 & +4x_3 & = & 1 \\ 2x_1 & -x_2 & +2x_3 & = & 2 \end{cases}.$$

Łatwo zauważyć, że rząd macierzy głównej wynosi 2, dlatego jądro ma wymiar 1. Czyli jest nieskończenie wiele rozwiązań.

Dla $\lambda \notin \{-1, 1\}$ dzielimy trzecie równanie przez $1 - \lambda$:

$$\begin{cases} 3x_1 & -x_2 & +4x_3 & = & 1 \\ 2x_1 & -x_2 & +2x_3 & = & 1+\lambda \\ (\lambda+1)x_1 & & +(1+\lambda)x_3 & = & 2 \end{cases}.$$

Tu już łatwo sprawdzić, że równanie ma dokładnie jedno rozwiązanie (np. licząc wyznacznik), ale można też dalej eliminacją Gaußa: od pierwszego równania odejmujemy drugie:

$$\begin{cases} x_1 & & +2x_3 & = & -\lambda \\ 2x_1 & -x_2 & +2x_3 & = & 1+\lambda \\ (\lambda+1)x_1 & & +(1+\lambda)x_3 & = & 2 \end{cases}.$$

Następnie od drugiego 2 razy pierwszy, od trzeciego $1 + \lambda$ razy pierwszy:

$$\begin{cases} x_1 & & +2x_3 & = & -\lambda \\ & -x_2 & -2x_3 & = & 1+3\lambda \\ & & -(1+\lambda)x_3 & = & 2(1+\lambda) \end{cases}.$$

Z czego wnioskujemy, że równanie ma dokładnie jedno rozwiązanie.

Rozdział 8

Wartości własne

8.1 Wartość własna, wektor własny

Definicja 8.1 (Wartość własna, wektor własny). λ jest *wartością własną* macierzy M (dla wektora $\vec{V} \neq 0$), gdy $M\vec{V} = \lambda\vec{V}$. $\vec{V}$ jest *wektorem własnym* tej macierzy.

λ jest *wartością własną* przekształcenia liniowego F, jeśli $F(v) = \lambda v$ dla pewnego $v \neq \vec{0}$. Taki wektor v jest *wektorem własnym* F.

Fakt 8.2. *Jeśli λ jest wartością własną przekształcenia F wtedy i tylko wtedy gdy jest wartością własną $M_{BB}(F)$, dla dowolnej bazy B.*

v jest wektorem własnym F dla wartości własnej λ wtedy i tylko wtedy, gdy $[v]_B$ jest wektorem macierzy $M_{BB}(F)$ dla wartości własnej λ.

Dowód. Zauważmy, że

$$[F(v)]_B = M_{BB}(F)[v]_B \ .$$

Jeśli v jest wektorem własnym F dla λ, to

$$[F(v)]_B = [\lambda v]_B = \lambda[v]_B$$

i tym samym

$$M_{BB}(F)[v]_B = \lambda[v]_B \ ,$$

czyli $[v]_B$ jest wektorem własnym dla wartości λ dla $M_{BB}(F)$.

Analogicznie, jeśli $[v]_B$ jest wektorem własnym dla wartości λ dla $M_{BB}(F)$ to

$$M_{BB}(F)[v]_B = \lambda[v]_B$$

czyli

$$[F(v)]_B = \lambda[v]_B,$$

tzn.

$$F(v) = \lambda v \ . \quad \square$$

Przykład 8.3. Przypomnijmy Przykład 5.10 i macierz

$$\begin{bmatrix} 4 & 0 & 0 \\ -1 & 5 & 1 \\ -1 & 1 & 5 \end{bmatrix} .$$

Łatwo sprawdzić, że

$$\begin{bmatrix} 4 & 0 & 0 \\ -1 & 5 & 1 \\ -1 & 1 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 0 & 0 \\ -1 & 5 & 1 \\ -1 & 1 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 0 & 0 \\ -1 & 5 & 1 \\ -1 & 1 & 5 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 6 \\ 6 \end{bmatrix}$$

Nie jest to zaskakujace: wiemy, że można przedstawić ją w postaci

$$\begin{bmatrix} 4 & 0 & 0 \\ -1 & 5 & 1 \\ -1 & 1 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} 0,5 & 0,5 & -0,5 \\ 0,5 & -0,5 & 0,5 \\ -0,5 & 0,5 & 0,5 \end{bmatrix} .$$

Co oznacza, że odpowiadające przekształcenie liniowe ma w bazie $\begin{bmatrix}1,1,0\end{bmatrix}^T; \begin{bmatrix}1,0,1\end{bmatrix}^T; \begin{bmatrix}0,1,1\end{bmatrix}^T$ macierz $\begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{bmatrix}$. No i wiemy już, że wartości własne przekształcenia nie zależą od wyboru bazy.

Wartości własne nie zawsze istnieją.

Przykład 8.4. Obrót $\mathbb{R}$ [2] o kąt 90^0 (w lewo). Jak wygląda macierz:

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} .$$

Geometrycznie „widać”, że przekształcenie to nie ma wektorów własnych, czyli nie ma też ich jego macierz.

Z drugiej strony, jeśli potraktujemy ją jako macierz nad $\mathbb{C}$, to wtedy

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} i \\ 1 \end{bmatrix} = i \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} -i \\ 1 \end{bmatrix} = -i \begin{bmatrix} 1 \\ i \end{bmatrix}$$

Czyli ma (zespolone) wartości własne i oraz $-i$.

8.2 Macierze podobne

Przedstawienie macierzy M w postaci $A^{-1}NA$ gdzie A to macierz zmiany bazy ma dla nas na razie sens tylko w przypadku przekształceń liniowych. Ale ta własność jest jakoś pomocna również bez rozważania konkretnych baz i zmian baz.

Definicja 8.5 (Macierze podobne)**.** Macierze kwadratowe A, B są *podobne*, jeśli istnieje macierz odwracalna C, taka że

$$A = C^{-1}BC \ .$$

Oznaczamy to jako $A \sim B$.

Lemat 8.6. *Rozpatrzmy macierz odwracalną $A = [A_1|A_2|\cdots|A_n]$. Jest to macierz zmiany bazy między bazą $B = \vec{A}_1, \ldots, \vec{A}_n$ oraz bazą standardową E:*

$$A = M_{BE} \ .$$

W szczególności, dla macierzy kwadratowej M oraz jej macierzy podobnej $M' = A^{-1}MA$ mamy

$$M' = M_{EB}MM_{BE} \ .$$

Oznacza to, że dla przekształcenia liniowego F_M indukowanego przez M macierz M' jest macierzą tego przekształcenia w bazie B.

$$M' = M_{EB}(M_{EE}(F))M_{BE} = M_{BB}(F_M) \ .$$

Dowód. Niech E: baza standardowa. Przypomnijmy, że dla bazy $B = v_1, \ldots, v_n$ oraz bazy B' mamy

$$M_{B,B'} = [(v_1)_{B'}|\cdots|(v_n)_{B'}]$$

W naszym przypadku

$$M_{BE} = [\vec{A}_1|\cdots|\vec{A}_n] \ .$$

.

Reszta to proste rachunki. □

Fakt 8.7. *Macierze podobne mają te same wartości własne.*

Dowód. Jeśli X jest wektorem własnym M dla wartości λ, to dla $M' = A^{-1}MA$ wektor $A^{-1}X$ jest wektorem własnym dla wartości λ. □

8.3 Wielomian charakterystyczny

Lemat 8.8. *λ jest wartością własną macierzy M* $\iff \det(M - \lambda \operatorname{Id}) = 0$

Dowód.

$$\lambda \text{ jest wartością własną } M \iff \exists v \neq \vec{0}\ Mv = \lambda v \iff \exists v \neq \vec{0} (M - \lambda \operatorname{Id})v = \vec{0} \iff \ker(M - \lambda \operatorname{Id}) \neq \{\vec{0}\} \iff \det(M - \lambda \operatorname{Id}) = 0 \ . \quad \square$$

Definicja 8.9 (Wielomian charakterystyczny)**.** *Wielomian charakterystyczny* macierzy kwadratowej to:

$$\varphi_M(x) = \det(A - x \operatorname{Id}) \ .$$

Wielomian charakterystyczny przekształcenia liniowego $F : V \to V$ to

$$\varphi_F(x) = \det(M_{BB}(F) - x \operatorname{Id}) \ ,$$

dla dowolnej bazy B przestrzeni $\mathbb{V}$.

Lemat 8.10. *Wielomian charakterystyczny dla macierzy $n \times n$ jest wielomianem stopnia n.*
λ jest wartością własną macierzy M wtedy i tylko wtedy gdy jest pierwiastkiem φ_M.

Dowód. Pokażemy przez indukcję trochę silniejszą tezę: dla macierzy, w które każdym wierszu i kolumnie najwyżej jeden element zależy liniowo od parametru x wyznacznik jest wielomianem stopnia najwyżej n. Jeśli w każdym wierszu i kolumnie jest taki wyraz, wielomian jest stopnia n.

Dowód to prosta indukcja względem rozwinięcia Laplace'a.

W drugiej części zauważmy, że $\varphi_M(\lambda) = \det(M - \lambda \operatorname{Id})$ jest dokładnie wartością z Lematu 8.8. $\square$

Lemat 8.11. *Wielomian charakterystyczny przekształcenia liniowego jest dobrze zdefiniowany.*

Dowód. Chcemy pokazać, że

$$\det(M_{BB}(F) - x \operatorname{Id}) = \det(M_{B'B'}(F) - x \operatorname{Id}) \ ,$$

dla dwóch dowolnych baz B, B'.

Policzmy

$$\begin{aligned}
\det(M_{BB}(F) - x \operatorname{Id}) &= \det(M_{BB'}(M_{BB}(F) - x \operatorname{Id})M_{B'B}) \\
&= \det(M_{BB'}(M_{BB}(F))M_{B'B} + M_{BB'}(-x \operatorname{Id})M_{B'B}) \\
&= \det(M_{B'B'}(F) - xM_{BB'} \operatorname{Id} M_{B'B}) \\
&= \det(M_{B'B'}(F) - x \operatorname{Id})
\end{aligned}$$
$\square$

Przykład 8.12 (Kontunuacja Przykładu 8.4)**.** Przypomnijmy, że obrót $\mathbb{R}$ [2] o kąt 90^0 (w lewo). Jak wygląda macierz:

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} .$$

Wielomian charakterystyczny tej macierzy to

$$\begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1.$$

Wielomian $\lambda^2 + 1$ nie ma pierwiastków rzeczywistych; ma pierwiastki zespolone: $i, -i$.

Rozwiązując układ równań (nad liczbami zespolonymi)

$$\begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} .$$

otrzymujemy wektor własny dla wartości własnej i. Analogicznie dla $-i$.

Przykład 8.13 (Kontynuacje Przykładu 8.3). Obliczmy wartości własne macierzy

$$\begin{bmatrix} 4 & 0 & 0 \\ -1 & 5 & 1 \\ -1 & 1 & 5 \end{bmatrix}$$

W tym celu obliczmy wyznacznik:

$$\begin{aligned} \begin{vmatrix} 4-\lambda & 0 & 0 \\ -1 & 5-\lambda & 1 \\ -1 & 1 & 5-\lambda \end{vmatrix} &= (4-\lambda)\begin{vmatrix} 5-\lambda & 1 \\ 1 & 5-\lambda \end{vmatrix} && \text{Rozwinięcie Laplace'a} \\ &= (4-\lambda)((5-\lambda)^2-1) \\ &= (4-\lambda)^2(6-\lambda) \end{aligned}$$

Czyli wartościami własnymi są 4, 6. Wektory własne obliczamy rozwiązując odpowiednie układy równań.

8.4 Krotności: algebraiczna i geometryczna.

Lemat 8.14. *Jeśli λ jest wartością własną dla M, to zbiór wektorów własnych dla M to* $\ker(M - \lambda \operatorname{Id})$*. W szczególności jest to przestrzeń liniowa.*

Uwaga. Formalnie wektor $\vec{0}$ nie jest wektorem własnym, ale wygodniej jest go zaliczyć tu do wektorów własnych, żeby wyszła podprzestrzeń.

Oznaczenie: dla ustalonej macierzy M oznaczamy

$$\mathbb{V}_\lambda = \{\vec{V} \ : \ M\vec{V} = \lambda\vec{V}\} \ .$$

Analogicznie dla przekształceń liniowych.

Tym samym, aby obliczyć wektory własne należy najpierw policzyć wielomian charakterystyczny, jego pierwiastki i dla ustalonego pierwiastka λ policzyć $\ker(M - \lambda \operatorname{Id})$. Można też oczywiście bezpośrednio próbować rozwiązać równanie

$$MX = \lambda X$$

w zmiennych $x_1, \ldots, x_n$.

Definicja 8.15 (Krotność algebraiczna, krotność geometryczna)**.** Dla wartości własnej λ *krotność geometryczna* to wymiar przestrzeni wektorów własnych dla λ, zaś *krotność algebraiczna* to krotność pierwiastka λ w wielomianie charakterystycznym.

Fakt 8.16. *Krotność geometryczna λ dla M to wymiar* $\ker(M - \lambda \operatorname{Id})$*.*

Lemat 8.17. *Krotność algebraiczna jest większa równa krotności geometrycznej.*

Dowód. Niech krotność geometryczna to k. Istnieje więc k niezależnych wektorów własnych $v_1, \ldots, v_k$ dla wartości własnej λ. Popatrzmy na przekształcenie liniowe F indukowane przez macierz M oraz na dowolną bazę B zawierającą $\vec{b}_1, \ldots, \vec{b}_k$. Wtedy $M_{BB}(F)$ jest podobna do M oraz jest postaci $\begin{bmatrix} D_\lambda & M'' \\ 0 & M' \end{bmatrix}$, gdzie D_λ jest macierzą diagonalną $k \times k$ której wszystkie elementy na przekątnej to λ. W szczególności wielomian charakterystyczny tej macierzy to

$$\begin{aligned} \left| \begin{bmatrix} D_\lambda & M'' \\ 0 & M' \end{bmatrix} - x \operatorname{Id} \right| &= \left| \begin{bmatrix} D_\lambda - x \operatorname{Id} & M'' \\ 0 & M' - x \operatorname{Id} \end{bmatrix} \right| \\ &= |D_\lambda - x \operatorname{Id}| \cdot |M' - x \operatorname{Id}| \\ &= (\lambda - x)^k \cdot |M' - x \operatorname{Id}|. \end{aligned}$$

Zawiera on $(\lambda - x)^k$, czyli λ jest k-krotnym pierwiastkiem, czyli krotność algebraiczna to przynajmniej k. $\square$

Uwaga. Jeśli krotność algebraiczna wynosi 1, to geometryczna też: nie może wynosić więcej, jednocześnie jeśli krotność algebraiczna λ wynosi 1 to $\det(M - \lambda \operatorname{Id}) = 0$, czyli λ jet wartością własną, czyli ma krotność geometryczną przynajmniej 1.

Przykład 8.18. $M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$ ma dwie wartości własne: 1 oraz 2. Krotność algebraiczna 2 to 2, ale geometryczna to 1: macierz $M - 2\,\mathrm{Id} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ ma rząd 2, więc wymiar jej jądra = wymiar przestrzeni wektorów własnych dla 2 to 1.

Przykład 8.19. Przypomnijmy Przykład 5.10 i macierz

$$\begin{bmatrix} 4 & 0 & 0 \\ -1 & 5 & 1 \\ -1 & 1 & 5 \end{bmatrix}$$

Ta macierz ma dwie wartości własne: 6, wymiar przestrzeni $\mathbb{V}_6$ to 1 (rozpięta przez wektor $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$); oraz 4, wymiar przestrzeni $\mathbb{V}_4$ to 2 (niezależne wektory własne to np. $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ i $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$).

W dalszej części będziemy pytać o to, ile wektorów własnych istnieje i czy może można z nich utworzyć bazę.

8.5 Przestrzenie niezmiennicze

Definicja 8.20 (Przestrzeń niezmiennicza)**.** Podprzestrzeń $\mathbb{V}' \leq \mathbb{V}$ przestrzeni liniowej $\mathbb{V}$ jest *przestrzenią niezmienniczą* dla $F : \mathbb{V} \to \mathbb{V}$, jeśli $F(\mathbb{V}') \subseteq \mathbb{V}'$.

Lemat 8.21. *Niech $\lambda_1, \ldots, \lambda_n$ będą różnymi wartościami własnymi macierzy M. Wtedy suma (mnogościowa) bez przestrzeni $\mathbb{V}_{\lambda_1}, \ldots, \mathbb{V}_{\lambda_k}$ jest zbiorem liniowo niezależnym.*

Dowód pozostawiamy jako ćwiczenie.

8.6 Macierze diagonalizowalne, przekształcenia diagonalne

Definicja 8.22 (Macierz diagonalizowalna, przekształcenie diagonalne)**.** Macierz M jest *diagonalizowalna* $\iff$ jest podobna do macierz przekątniowej.

Przekształcenie liniowe jest *diagonalne*, jeśli jego macierz (w jakiejś bazie) jest diagonalizowalna.

Lemat 8.23. *Następujące warunki są równoważne dla przekształcenia liniowego $F : \mathbb{V} \to \mathbb{V}$:*

1. *$M_{BB}(F)$ jest diagonalizowalna w pewnej bazie B;*
2. *$M_{BB}(F)$ jest diagonalizowalna w każdej bazie B;*
3. *$M_{BB}(F)$ jest diagonalna w pewnej bazie B.*

Dowód. Intuicja jest taka, że to są zmiany bazy.

$3 \Rightarrow 2$ Skoro $M_{BB}(F)$ jest diagonalna, to z definicji

$$M_{B'B'}(F) = M_{BB'} M_{BB}(F) M_{B'B}$$

i tym samym $M_{B'B'}(F)$ jest diagonalizowalna w każdej bazie B'.

$2 \Rightarrow 1$ Oczywiste.

$1 \Rightarrow 3$ Niech $M_{BB}(F)$ będzie diagonalizowalna, czyli $M_{BB}(F) = A^{-1}DA$. Zauważmy, że

$$D = AM_{BB}(F)A^{-1} \ .$$

Zdefiniujemy bazę B' tak, aby $M_{BB'} = A$. Wtedy

$$\begin{aligned} D &= AM_{BB}(F)A^{-1} \\ &= M_{BB'}M_{BB}(F)M_{B'B} \\ &= M_{B'B'}(F) \ . \end{aligned}$$

Taka baza istnieje: i-ta kolumna $A^{-1} = M_{B'B}$ to $(b'_i)_B$. □

Twierdzenie 8.24. *Następujące warunki są równoważne dla macierzy kwadratowej M rozmiaru $n \times n$:*

1. *M jest diagonalizowalna*
2. *M ma n niezależnych wektorów własnych*
3. *Suma wymiarów przestrzeni wartości własnych $\mathbb{V}_\lambda$ macierzy M wynosi n.*

Analogiczne twierdzenie zachodzi też dla przekształceń liniowych.

Dowód nieobowiązkowy, dla zainteresowanych.

Dowód. $1 \Rightarrow 2$ Skoro M jest diagonalizowalna, to istnieje A, A^{-1} oraz macierz przekątniowa D takie że

$$M = A^{-1}DA \ .$$

Oczywiście D ma n niezależnych wektorów własnych (konkretnie: $\vec{E}_1, \ldots, \vec{E}_n$) i w takim razie, analogicznie jak w Fakcie 8.7, wnioskujemy, że $A^{-1}\vec{E}_1, \ldots, A^{-1}\vec{E}_n$ są wektorami własnymi M dla odpowiadających wartości własnych. Łatwo sprawdzić, że są to po prostu kolumny A^{-1}.

$2 \Rightarrow 3$ Niech $\vec{V}_1, \ldots, \vec{V}_n$ to niezależne wektory własne M. Wystarczy zauważyć, że $\dim V_\lambda$ to liczba wektorów spośród $\vec{V}_1, \ldots, \vec{V}_n$, które odpowiadają wartości λ. Suma wymiarów $\mathbb{V}_\lambda$ po różnych λ wynosi przynajmniej n, jednocześnie z Lematu 8.21 nie może wynosić więcej niż n, bo suma baz przestrzeni $\mathbb{V}_\lambda$ jest zbiorem liniowo niezależnym, czyli ma wielkość najwyżej nn.

$3 \Rightarrow 1$ Rozważmy bazy poszczególnych przestrzeni $\mathbb{V}_\lambda$, niech w sumie dają one układ $A_1, \ldots, A_n$. Z Lematu 8.21 ten układ jest liniowo niezależny, czyli jest bazą. Zdefiniujmy $A^{-1} = [A_1|\cdots|A_n]$. Łatwo sprawdzić, jak w Fakcie 8.7, że

$$M = A^{-1}DA$$

gdzie D jest macierzą diagonalną mającą na pozycji ii wartość własną dla wektora A_i. □

8.7 Macierz Jordana

Zajmiemy się obecnie problemem, jak bardzo macierz może nie być diagonalizowalna. Zauważmy, że w przypadku liczb zespolonych każdy wielomian ma pierwiastek, w szczególności wielomian charakterystyczny każdej macierzy ma pierwiastek, czyli każda macierz zespolona ma wektor własny.

Definicja 8.25 (Klatka Jordana, macierz Jordana). *Klatką Jordana* nazywamy macierz postaci

$$\begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}$$

Macierzą Jordana nazywamy macierz postaci

$$\begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{bmatrix}$$

gdzie $J_1, J_2, \ldots, J_k$ są klatkami Jordana.

Ważne $\lambda \in \mathbb{C}$, tj. może być liczbą zespoloną.

Fakt 8.26. *Klatka Jordana J rozmiaru $k \times k$ ma jedną wartość własną: λ, o krotności algebraicznej k oraz geometrycznej* 1.

Dowód pozostawiamy jako ćwiczenie. Jest to w pewnym sensie najgorszy przypadek, jeśli chodzi o wartości własne.

Twierdzenie 8.27. *Każdą macierz M o wartościach w $\mathbb{C}$ można przedstawić w postaci*

$$M = A^{-1}JA$$

gdzie J jest macierzą Jordana a A jest macierzą odwracalną (o wartościach w $\mathbb{C}$).

Uwaga: różne klatki mogą być dla tej samej wartości λ.

Przykład/Zastosowanie 8.28. Przypomnijmy sobie macierz odpowiadającą rekurencji na liczny Fibonacciego.

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} .$$

Policzmy jej wielomian charakterystyczny:

$$\begin{vmatrix} -x & 1 \\ 1 & 1-x \end{vmatrix} = x^2 - x - 1 .$$

Pierwiastki to $\frac{\sqrt{5}+1}{2}$, $\frac{-\sqrt{5}+1}{2}$. Czyli macierz jest postaci

$$A^{-1}\begin{bmatrix} \frac{\sqrt{5}+1}{2} & 0 \\ 0 & \frac{-\sqrt{5}+1}{2} \end{bmatrix} A .$$

n-ta potęga tej macierzy to

$$A^{-1}\begin{bmatrix} \left(\frac{\sqrt{5}+1}{2}\right)^n & 0 \\ 0 & \left(\frac{-\sqrt{5}+1}{2}\right)^n \end{bmatrix} A .$$

To w szczególności mówi nam, jak wygląda wyraz ogólny: jest postaci $a\left(\frac{\sqrt{5}+1}{2}\right)^n + b\left(\frac{-\sqrt{5}+1}{2}\right)^n$, dla odpowiednich wartości a, b.

Zauważmy też, że A oraz A^{-1} można łatwo policzyć: kolumny A^{-1} to wektory własne: niech C_1 to wektor własny dla $\frac{\sqrt{5}+1}{2}$ zaś C_2 dla $\frac{-\sqrt{5}+1}{2}$. Zdefiniujmy $A^{-1} := [C_1|C_2]$. Aby pokazać, że jest to dobrze dobrane A i A^{-1}, wystarczy pokazać, że $A^{-1}\begin{bmatrix} \frac{\sqrt{5}+1}{2} & 0 \\ 0 & \frac{-\sqrt{5}+1}{2} \end{bmatrix} A$ oraz M są równe,

czyli wystarczy, że mają te same wartości na C_1, C_2:

$$\begin{aligned} A^{-1}\begin{bmatrix} \frac{\sqrt{5}+1}{2} & 0 \\ 0 & \frac{-\sqrt{5}+1}{2} \end{bmatrix} AC_1 &= A^{-1}\begin{bmatrix} \frac{\sqrt{5}+1}{2} & 0 \\ 0 & \frac{-\sqrt{5}+1}{2} \end{bmatrix} E_1 && \text{bo } A \text{ odwrotna do } A^{-1} \\ &= A^{-1}\begin{bmatrix} \frac{\sqrt{5}+1}{2} \\ 0 \end{bmatrix} \\ &= \frac{\sqrt{5}+1}{2} A^{-1}E_1 \\ &= \frac{\sqrt{5}+1}{2} C_1 && \text{bo } A^{-1} = [C_1|C_2] . \end{aligned}$$

Analogicznie liczymy dla C_2.

Używając macierzy Jordana możemy podać rozwiązanie ogólne dla każdej zależności tej postaci (tzn. rekurencji liniowej).

8.8 Macierze symetryczne

Twierdzenie 8.29. *Macierz symetryczna z* $M_{n\times n}(\mathbb{R})$ *ma* n *niezależnych wektorów własnych (nad* $\mathbb{R}$*).*

Rozdział 9

PageRank

Na podstawie pracy Kurt Bryan i Tanya Leise „The $25,000,000,000 eigenvector. The linear algebra behind Google.” *SIAM Review*, 48:3 (2006) 569–581.

9.1 Macierze sąsiedztwa, ranking

Modelujemy internet jako graf: zbiór wierzchołków to strony, (skierowane) krawędzie to linki miedzy nimi (krawędź z i do j oznacza, że jest link ze strony i do j). Naszym celem jest skonstruowanie rankingu, tj. przypisanie każdej stronie jej “ważności” w sieci. Chcemy to robić na podstawie linków, każdemu przypisujemy sumę głosów 1. Zakładamy, że graf nie ma „pętli”, tzn. krawędzi z i do i.

Definicja 9.1 (Znormalizowana macierz sąsiedztwa)**.** Dla grafu G o wierzchołkach $1, 2, \ldots, n$ niech $d_{i,j}$ oznacza liczbę krawędzi z j do i (może to być 0), zaś m_j liczbę krawędzi wychodzących z j ($= \sum_i d_{i,j}$).

Znormalizowana macierz sąsiedztwa $M(G)$ to macierz $(m_{i,j})_{i,j=1,\ldots,n}$, gdzie

$$m_{i,j} = \frac{d_{i,j}}{m_j} \quad .$$

Zauważmy, że liczby w kolumnie są nieujemne i jeśli istnieje choć jedna krawędź, to sumują się do 1. W dalszej części będziemy się zajmować grafami, które nie mają takich wierzchołków. Taką macierz nazywamy *macierzą stochastyczną*.

Definicja 9.2 (Macierz stochastyczna, wektor stochastyczny)**.** Wektor jest *stochastyczny*, jeśli jego współrzędne są nieujemne i sumują się do 1.

Macierz kwadratowa M jest *(kolumnowo) stochastyczna*, jeśli każda jej kolumna jest wektorem stochastycznym.

Fakt 9.3. *Iloczyn dwóch macierzy stochastycznych jest macierzą stochastyczną.*

Jeśli $M_1, \ldots, M_k$ są macierzami stochastycznymi oraz $\alpha_1, \ldots, \alpha_k$ są liczbami nieujemnymi, spełniającymi $\sum_i \alpha_i = 1$, to

$$\sum_{i=1}^{k} \alpha_i M_i$$

też jest macierzą stochastyczną.

Prosty dowód pozostawiamy na ćwiczenia.

Potęgi znormalizowanej macierzy sąsiedztwa mają naturalną interpretację: wyraz i, j macierzy M^k jest niezerowy wtedy i tylko wtedy, gdy istnieje scieżka długości k w grafie sąsiedztwa z j do i. To stwierdzenie ma dokładniejszą, ilościową wersję:

Lemat 9.4. *Niech M będzie znormalizowaną macierzą sąsiedztwa zaś $\vec{V}$ wektorem stochastycznym. Wtedy $M^k\vec{V}$ to rozkład prawdopodobieństwa procesu losowego:*

krok 0 *W kroku 0 losujemy wierzchołek początkowy wg. rozkładu wyznaczonego przez $\vec{V}$, tj. wierzchołek i jest wylosowany z prawdopodobieństwem v_i.*

krok k *W każdym kolejnym kroku, jeśli jesteśmy w wierzchołku v, wybieramy z takim samym prawdopodobieństwem jedną z krawędzi wychodzących z v.*

Definicja 9.5. *Ranking* dla macierzy stochastycznej M to wektor $\vec{R}$, taki, że $M\vec{R} = \vec{R}$.
Rankingiem wierzchołka grafu jest odpowiadająca współrzędna tego wektora.

Innymi słowy, jest to wektor własny dla wartości 1.

Uwaga. Ranking to „stabilny" rozkład prawdopodobieństwa, w tym sensie, że odpowiada prawdopodobieństwu znalezienia się w danym wierzchołku po dużej liczbie kroków (ta intuicja niestety jest zawodna z paru powodów).

Uwaga. Zauważmy, że zamiast $\sum_i r_i = 1$ moglibyśmy wziąć dowolną inną liczbę niż 1, ale dla 1 to daje ładną interpretację probabilistyczną.

Chcielibyśmy, żeby ranking istniał, był jedyny oraz był nieujemny.

Lemat 9.6 (Istnienie rankingu). *Macierz stochastyczna ma wartość własną* 1.

Dowód. Wiemy, że macierz M i M^T mają te same wartości własne. Popatrzmy więc na macierz M^T. Łatwo sprawdzić, że wektor $[1, 1, \ldots, 1]^T$ składający się z samych jedynej jest wektorem własnym dla wartości 1: i-ty element w $M^T[1, 1, \ldots, 1]^T = ([1, 1, \ldots, 1]M)^T$ to

$$\sum_j m_{j,i} \cdot 1 = \sum_j m_{j,i} = 1 \ . \qquad \square$$

Fakt 9.7. *Jeśli w grafie, który nie ma wierzchołków bez wychodzących krawędzi, istnieją dwa różne podzbiory wierzchołków, z których nie ma krawędzi wychodzących, to ranking nie jest jedyny.*

Uwaga. W praktyce, graf internetu nie był spójny (teraz być może już jest).
Poza tym wiszące wierzchołki są problemem.

Dowód. W języku wartości własnych: $\dim \mathbb{V}_1 > 1$.
Niech V_i będzie silnie spójną składową bez krawędzi wychodzących. Wtedy wektor mający 1 na współrzędnych z V_i oraz 0 gdzie indziej jest wektorem własnym dla wartości 1.
Graf spełniający warunek lematu ma przynajmniej dwie takie składowe. $\square$

9.2 Macierze dodatnie, PageRank

Aby zapewnić te warunki, zajmiemy się inną macierzą: dla znormalizowanej macierzy sąsiedztwa M rozmiaru $n \times n$ oraz liczby $0 < m < 1$ definiujemy

$$M' = (1-m)M + m \cdot \begin{bmatrix} \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \end{bmatrix}$$

Dla odpowiedniej wartości m ranking tej macierzy to PageRank.

Fakt 9.8. *Macierz M' jest macierzą stochastyczną.*

Uwaga. Macierz ta ma naturalną interpretację jako proces losowy: w każdym kroku z prawdopodobieństwem $1 - m$ losujemy krawędź wychodzącą, zaś z prawdopodobieństwem m losujemy jednorodnie jeden ze wszystkich wierzchołków.

Uwaga. Poniższe definicje oraz dowody dla macierzy dodatnich są prostszym wariantem ogólniejszego twierdzenia Frobeniusa-Perrona i jego dowodu.

Definicja 9.9. Mówimy, że macierz A jest *dodatnia*, co zapisujemy $A > 0$, jeśli wszystkie jej elementy są dodatnie.

Lemat 9.10. *Jeśli $A > 0$ i jest kolumnowo stochastyczna oraz $A\vec{V} = \vec{V}$ to $\vec{V} > 0$ lub $\vec{V} < 0$.*

Dowód. Załóżmy, że $\vec{V}$ ma współrzędne różnych znaków. Wtedy

$$\sum_i |v_i| > \left|\sum_i v_i\right| \quad . \tag{9.1}$$

Skoro $\vec{V} = A\vec{V}$ to

$$v_i = \sum_j a_{i,j} v_j \quad .$$

Zgodnie z wcześniejszą obserwacją (9.1) mamy

$$\begin{aligned} |v_i| &= \left|\sum_j a_{i,j} v_j\right| \\ &< \sum_j a_{i,j} |v_j| \end{aligned}$$

Sumując po i

$$\begin{aligned} \sum_i |v_i| &< \sum_i \sum_j a_{i,j} |v_j| \\ &= \sum_j |v_j| \underbrace{\sum_i a_{i,j}}_{\text{kolumna stochastyczna}} \\ &= \sum_j |v_j| \quad . \end{aligned}$$

Sprzeczność.

Pozostaje sprawdzić, że nie ma współrzędnej zerowej. Ale w sumie

$$v_i = \sum_j a_{i,j} v_j$$

wszystkie $a_{i,j}$ są dodatnie. Jeśli choć jeden v_j jest dodatni, to również v_i jest. A nie mogą być wszystkie zerowe (bo wtedy cały wektor $\vec{V}$ jest zerowy.) □

Lemat 9.11. *Dla dwóch niezależnych wektorów $\vec{S}, \vec{T} \in \mathbb{R}^n$ istnieje ich kombinacja liniowa, która ma pozycje różnych znaków.*

Dowód. Jeśli któryś z $\vec{S}, \vec{T}$ ma pozycje mieszanych znaków, to teza trywialnie zachodzi. W dalszej części zakładamy więc, że $\vec{S}, \vec{T} > 0$.

Nazwijmy składowe $\vec{S}$ i $\vec{T}$ przez $s_1, \ldots, s_n$ oraz $t_1, \ldots, t_n$. Jeśli dla którejś współrzędnej mamy $s_i = t_i = 0$, to usuwamy ją z obu wektorów; zauważmy, że musiały nam zostać przynajmniej dwie współrzędne. Oznaczmy $\alpha_i = \frac{s_i}{t_i}$, jeśli $t_i = 0$, to $\alpha_i = +\infty$. Weźmy teraz $\alpha \notin \{\alpha_1, \ldots, \alpha_n\}$, takie że istnieją α_i, α_j spełniające $\alpha_i < \alpha < \alpha_j$. Wtedy $\vec{S} - \alpha\vec{T}$ ma na współrzędnej i liczbę ujemną, a na j: dodatnią. □

Twierdzenie 9.12. *Dla stochastycznej macierzy dodatniej A mamy* $\dim \mathbb{V}_1 = 1$.

Dowód. Wiemy z Lematu 9.6, że $\dim \mathbb{V}_1 \geq 1$. Załóżmy więc, że wynosi przynajmniej 2. Wtedy istnieją $\vec{S}, \vec{T} \in \mathbb{V}_1 = 1$. Ale w takim razie z Lematu 9.11 istnieje $\vec{W} \in \mathbb{V}_1$, który ma zarówno dodatnie jak i ujemne współrzędne. Ale to jest sprzeczność z Lematu 9.10. □

9.3 Grafy silnie spójne

Jeśli dany na wejściu graf jest silnie spójny (czyli z każdego wierzchołka da się dojść do każdego innego), to można pokazać, że $\dim \mathbb{V}_1 = 1$ nawet dla znormalizowanej macierzy sąsiedztwa.

Definicja 9.13. Mówimy, że graf jest *silnie spójny*, jeśli dla każdej pary wierzchołków i, j istnieje ścieżka z i do j (oraz z j do i).

Choć wygląda niewinnie, w praktyce jest to bardzo silne założenie.

Lemat 9.14. *Dla znormalizowanej macierzy sąsiedztwa M grafu silnie spójnego o n wierzchołkach macierz $\frac{1}{n}\sum_{i=0}^{n-1} M^i$ jest dodatnią macierzą stochastyczną.*

Dowód. Z Faktu 9.3 mamy, że tak zdefiniowana macierz jest stochastyczna.

Przypomnijmy, że w M^k element ij jest niezerowy wtedy i tylko wtedy, gdy istnieje ścieżka z j do i długości dokładnie k. Skoro graf jest spójny, to między każdą parą wierzchołków j, i istnieje ścieżka długości najwyżej $n-1$. W takim razie dla pewnego $k \leq n-1$ mamy, że element ij macierzy M^k jest dodatni. (Dla $i = j$ korzystamy z tego, że $M^0 = \mathrm{Id}$) □

Lemat 9.15. *Jeśli $\vec{V}$ jest wektorem własnym znormalizowanej macierzy sąsiedztwa dla wartości 1, to jest nim też dla macierzy $\frac{1}{n}\sum_{i=0}^{n-1} M^i$.*

Dowód. Zauważmy najpierw, że $M^i\vec{V} = 1^i\vec{V} = \vec{V}$. Wtedy

$$\begin{aligned}
\left(\frac{1}{n}\sum_{i=0}^{n-1} M^i\right)\vec{V} &= \frac{1}{n}\sum_{i=0}^{n-1}\left(M^i\vec{V}\right)\\
&= \frac{1}{n}\sum_{i=0}^{n-1}\vec{V}\\
&= \frac{1}{n}\cdot n\cdot\vec{V}\\
&= \vec{V}
\end{aligned}$$

□

Twierdzenie 9.16. *Jeśli graf jest spójny, to jego znormalizowana macierz sąsiedztwa ma* $\dim \mathbb{V}_1 = 1$.

Dowód. Wiemy z Lematu 9.6, że $\dim \mathbb{V}_1 \geq 1$.

Rozpatrzmy macierz $\frac{1}{n}\sum_{i=0}^{n-1} M^i$. Oznaczmy przestrzeń jej wektorów własnych dla wartości własnej 1 przez $\mathbb{V}_1'$. Z Lematu 9.15 każdy wektor własny M dla wartości 1 jest też wektorem tej macierzy, czyli $1 \leq \dim \mathbb{V}_1 \leq \dim \mathbb{V}_1'$. Z Lematu 9.14 ta macierz jest stochastyczna dodatnia i z Twierdzenia 9.12 wymiar jej przestrzeni wektorów własnych dla wartości 1 to jeden, tj. $\dim \mathbb{V}_1' = 1$ i tym samym $1 = \dim \mathbb{V}_1' = \dim \mathbb{V}_1$. □

Ten wynik można wzmocnić, ale wymaga to głównie rozważań teorio-grafowych.

9.4 Obliczanie rankingu

Pozostaje powiedzieć, jak można policzyć ranking dla macierzy stochastycznej dodatniej.

Niech A będzie dodatnią macierzą kolumnowo stochastyczną.

9.4.1 Układ równań

Najprostsza obserwacja, to że skoro wymiar $\dim \mathbb{V}_1 = 1$,

Fakt 9.17. *Niech $A > 0$ będzie macierzą kolumnowo stochastyczną. Układ równań*

$$\left\{\begin{array}{lcl} (A - \mathrm{Id})\vec{X} & = & 0 \\ \sum_i x_i & = & 1 \end{array}\right. .$$

ma dokładnie jedno rozwiązanie.

Dowód. Zbiór rozwiązań równania $(A - \mathrm{Id})\vec{X} = 0$ ma wymiar 1, łatwo sprawdzić, że dokładnie jeden z tych wektorów spełnia dodatkowy warunek $\sum_i x_i = 1$: weźmy dowolne $\vec{V}$ spełniające pierwsze równanie, wszystkie inne są postaci $\alpha\vec{V}$ dla $\alpha \in \mathbb{R}$ i mają one wtedy sumę współrzędnych $\alpha\left(\sum_i v_i\right)$. Widać, że dla dokładnie jednego α $(= 1/\sum_i v_i)$ ta suma wynosi 1. □

Ten układ można więc rozwiązać problematyczny jednak jest jego rozmiar.

9.4.2 Metoda iteracyjna.

Alternatywnie, chcemy pokazać, że można to policzyć jako granicę $(M')^k\vec{V}$ (dla sensownie wybranego $\vec{V}$).

Weźmy dowolny wektor $\vec{V} > 0$ o sumie współrzędnych 1, niech $\vec{R}$ będzie rankingiem. Policzmy:

$$\begin{aligned} M^k\vec{V} &= M^k\vec{R} + M^k(\vec{V} - \vec{R}) \\ &= \vec{R} + M^k(\vec{V} - \vec{R}) \end{aligned}$$

Chcemy więc sprawdzić, jak się zachowuje $M^k(\vec{V} - \vec{R})$. W ogólności ciężko coś powiedzieć, ale zauważmy, że skoro suma współrzędnych $\vec{V}, \vec{R}$ to 1, to $\vec{V} - \vec{R}$ ma sumę współrzędnych równą 0; analogicznie, również suma współrzędnych $M^k\vec{V}$ oraz $M^k\vec{R} = \vec{R}$ jest równa 1, czyli $M^k(\vec{V} - \vec{R})$ ma sumę współrzędnych równą 0.

Zdefiniujmy $\mathbb{V}_{=0}$: przestrzeń liniową wektorów, których współrzędne sumują się do 0:

$$\mathbb{V}_{=0} = \{[v_1, \dots, v_n]^T \ : \ \sum_i v_i = 0\} \ .$$

Fakt 9.18. $M^k(\vec{V} - \vec{R}) \in \mathbb{V}_{=0}$.

Chcemy coś powiedzieć o „granicy” $M^k\vec{W}$ dla $\vec{W} \in \mathbb{V}_{=0}$. Skoro jest granica, to jest potrzebna jakaś odległość.

Definicja 9.19. Norma ℓ_1 $\|\cdot\|_1$ wektora $\vec{V} = [v_1, \dots, v_n]^T$ to

$$\|\vec{V}\|_1 = \sum_{i=1}^{n} |v_i| \ .$$

Lemat 9.20. *Niech A będzie dodatnia macierzą stochastyczną. Niech*

$$a = \max_{1 \leq j \leq\ n} (1 - 2 \min_{1 \leq i \leq\ n} a_{i,j}) \ .$$

Niech $\vec{0} \neq \vec{V} \in \mathbb{V}_{=0}$. Wtedy

$$\|AV\|_1 \leq a\|V\|_1 \ .$$

Uwaga. Niejawnie zakładamy, że $n > 1$, żeby a było dodatnie.

Dowód. Niech $\vec{V} = [v_1, \dots, v_n]^T$. Niech $\operatorname{sgn} x$ oznacza znak x, tj. $x \geq 0 \implies \operatorname{sgn}(x) = 1, x < 0 \implies \operatorname{sgn}(x) = -1$. Oznaczmy $\vec{W} = A\vec{V}$, niech $\vec{W} = [w_1, \dots, w_n]^T$. Jeśli $\vec{W} = \vec{0}$ to teza oczywiście zachodzi.

$$\begin{aligned} \|W\|_1 &= \sum_i |w_i| \\ &= \sum_i \operatorname{sgn}(w_i) w_i \\ &= \sum_i \operatorname{sgn}(w_i) \sum_j a_{i,j} v_j \\ &= \sum_j v_j \sum_i \operatorname{sgn}(w_i) a_{i,j} \\ &\leq \sum_j |v_j| \cdot \left| \sum_i \operatorname{sgn}(w_i) a_{i,j} \right| \end{aligned}$$

Zauważmy, że $\sum_i a_{i,j} = 1$ oraz że $w_1, \dots, w_n$ nie są wszystkie tego samego znaku, bo $\vec{0} \neq \vec{W} \in \mathbb{V}_{=0}$. Czyli $0 \leq |\sum_i \operatorname{sgn}(w_i) a_{i,j}| \leq 1 - 2\min_{1 \leq i \leq n} a_{i,j} \leq a$, bo od $\sum_i a_{i,j}$ odejmujemy przynajmniej dwa elementy.

$$\begin{aligned} \sum_j |v_j| \cdot \left| \sum_i \operatorname{sgn}(w_i) a_{i,j} \right| &\leq \sum_j |w_j| \cdot a \\ &= a\|\vec{V}\|_1 \end{aligned}$$

□

Niestety, wartość a może (w ogólności) być wielomianowo mała w stosunku do grafu (zwłaszcza przy wielokrotnych krawędziach). Sprawia to, że w ogólności trzeba policzyć wielomianowo wiele iteracji, by zmniejszyć dwukrotnie odległość od rozwiązania. W praktyce jednak nie jest to potrzebne. W PageRanku jest jednak stała, wystarczy więc policzyć logarytmicznie wiele iteracji.

Zauważmy też, że obliczanie $M'\vec{V}$ jest prostsze ze względu na strukturę M': nasza dodatnia macierz stochachastyczna jest w istocie macierzą

$$(1-m)M + m \begin{bmatrix} \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \end{bmatrix} .$$

Wtedy:

$$\begin{aligned} M'\vec{V} &= (1-m)M\vec{V} + m \begin{bmatrix} \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \end{bmatrix} \vec{V} \\ &= (1-m)M\vec{V} + \begin{bmatrix} \frac{m}{n} \\ \frac{m}{n} \\ \vdots \\ \frac{m}{n} \end{bmatrix} . \end{aligned}$$

Zauważmy, że ten iloczyn liczy się dużo prościej: macierz M jest dość rzadka. Co więcej, liczenie można zrównoleglić (każdy element $M\vec{V}$ może być liczony osobno).

Rozdział 10

Iloczyn skalarny

Chcemy uogólnić pojęcia odległości, prostopadłości, kąta na dowolną przestrzeń. W tym celu zajmiemy się *iloczynem skalarnym.*

10.1 Standardowy iloczyn skalarny

Przykład 10.1 (Standardowy iloczyn skalarny). Dla przestrzeni $\mathbb{R}^n$ oraz wielu $\mathbb{F}^n$ (ale nie $\mathbb{C}^n$) definiujemy iloczyn skalarny jako:

$$\langle (v_1, \dots, v_n), (u_1, \dots, u_n) \rangle = \sum_{i=1}^{n} v_i u_i$$

Dla $\mathbb{C}^n$ definiujemy zaś:

$$\langle (v_1, \dots, v_n), (u_1, \dots, u_n) \rangle = \sum_{i=1}^{n} v_i \overline{u_i}$$

Dla $\mathbb{R}^n$ czy $\mathbb{C}^n$ możemy go użyć do zdefiniowania (standardowej) długości, odległości oraz prostopadłości, kąta:

długość Długość (norma) wektora v to

$$\|v\| = \sqrt{\langle v, v \rangle}$$

odległość Odległość między wektorami u, v to

$$\|u - v\|$$

kąt Kąt między wektorami u, v to $\alpha \in [0, \pi]$ spełniające warunek

$$\cos \alpha = \frac{\langle v, u \rangle}{\|v\| \cdot \|u\|}$$

prostopadłość Dwa wektory u, v są prostopadłe, jeśli

$$\langle v, u \rangle = 0$$

Ta definicja prostopadłości okaże się przydatna nawet wtedy, kiedy kąt czy długość nie mają wiele sensu.

10.2 Ogólny iloczyn skalarny

Chcemy uogólnić to na ogólne przestrzenie. W zasadzie to rozważamy przestrzenie nad $\mathbb{R}$, informacyjnie nad $\mathbb{C}$.

Popatrzymy od innej strony: co musi spełniać funkcja dwóch zmiennych, by być iloczynem skalarnym.

Definicja 10.2 (Iloczyn skalarny). *Iloczyn skalarny* to funkcja $\langle \cdot, \cdot \rangle : \mathbb{V}^2 \mapsto \mathbb{F}$ (gdzie $\mathbb{V}$ jest przestrzenią liniową nad $\mathbb{F}$) spełniająca warunki:

(SK1) liniowa po pierwszej współrzędnej

(SK2) symetryczna, tj. $\langle u, v\rangle = \langle \vec{v}, \vec{u}\rangle$; (np. dla $\mathbb{F} = \mathbb{R}$) lub antysymetryczny $\langle u, v\rangle = \overline{\langle \vec{v}, \vec{u}\rangle}$ (np. dla $\mathbb{F} = \mathbb{C}$).

(SK3) $\langle v, v\rangle > 0$ dla $v \neq \vec{0}$.

Przestrzeń liniową, która ma tak określony iloczyn skalarny, nazywamy *przestrzenią Euklidesową* (jeśli $\mathbb{F} = \mathbb{R}$) lub *unitarną* (jeśli $\mathbb{F} = \mathbb{C}$).

Uwaga. Ostatni warunek ma sens dla $\mathbb{C}$, bo wartość jest samosprzężona. Dla innych ciał ostatni warunek może nie mieć sensu.

To pozwala na zdefiniowanie prostopadłości oraz długości.

Definicja 10.3 (Wektory prostopadłe)**.** Dwa wektory u, v są *prostopadłe*, gdy $\langle \vec{u}, \vec{v}\rangle = 0$. Zapisujemy to też jako $u \perp v$.

Definicja 10.4 (Długość i odległość)**.** W przestrzenie Euklidesowej (unitarnej):
Norma (*długość*) wektora u to $\|u\| = \sqrt{\langle u, u\rangle}$.
Odległość między u a v to norma z $(u - v)$.

Przykład 10.5.

- Tradycyjny iloczyn skalarny w $\mathbb{R}^n$, $\mathbb{C}^n$ spełnia te warunki.
- W przestrzeni wielomianów (nad $\mathbb{R}$) jako iloczyn skalarny można wziąć całkę (po odpowiednim zakresie):

$$\langle u, v\rangle = \int_I u(x)v(x)dx$$

Iloczyn skalarny ma wiele dobrych własności:

Lemat 10.6. *Jeśli $\mathbb{V}$ jest przestrzenia Euklidesową (unitarną), to:*

1. $\|tv\| = |t| \cdot \|v\|$
2. $|\langle \vec{u}, \vec{v}\rangle| \leq \|u\| \cdot \|v\|$ *(Nierówność Cauchy-Schwartz); równość* $\iff$ *są liniowo zależne*
3. $\|u + w\| \leq \|u\| + \|v\|$ *(Nierówność Minkowsky)*
4. $\|v\| - \|w\| \leq \|v - w\|$

Dowód. Ad 1: Oczywiste

Ad 2: Jak są liniowo zależne, to jasne. Rozważmy

$$f(t) = \|v - tw\|^2 > 0 \ .$$

Ma wartości ściśle dodatnie.

Po przekształceniu

$$f(t) = \|v\|^2 - 2t\langle v, w\rangle + t^2\|w\|^2 > 0 \ .$$

Patrzymy na

$$\Delta = 4\langle v, w\rangle^2 - 4\|w\|^2\|v\|^2 < 0 \ ,$$

co daje tezę.

Przy okazji: równość jest tylko wtedy, gdy są liniowo zależne.

Ad 3:

$$\begin{aligned}\|u + v\|^2 &= \langle u + v, u + v\rangle \\ &= \langle u, u\rangle + 2\langle \vec{u}, \vec{v}\rangle + \langle v, v\rangle \\ &\leq \|u\|^2 + 2\|u\| \cdot \|v\| + \|v\|^2 \\ &= (\|u\| + \|v\|)^2\end{aligned}$$

Ad 4: Wynika z punktu trzeciego. $\square$

Z nierówności Schwarza (dla liczb rzeczywistych) mamy

$$-1 \leq \frac{\langle \vec{u}, \vec{v} \rangle}{\|u\| \cdot \|v\|} \leq 1$$

I tym samym możemy zdefiniować *kąt* miedzy wektorami

Definicja 10.7. W przestrzeni Euklidesowej (unitarnej) dla wektorów u, v kąt między nimi to jedyne takie $\alpha \in [0, \pi]$, że

$$\cos \alpha = \frac{\langle \vec{u}, \vec{v} \rangle}{\|u\| \cdot \|v\|}.$$

10.3 Baza ortonormalna

Definicja 10.8 (Układ (baza) ortogonalny, układ (baza) ortonormalny)**.** Układ wektorów $\vec{v}_1, \ldots, \vec{v}_n$ jest *układem ortogonalnym*, jeśli dla $i \neq j$ mamy $\langle v_i, v_j \rangle = 0$. Jest układem *ortonormalnym*, jeśli dodatkowo $\langle v_i, v_i \rangle = 1$.

Analogicznie definiujemy bazę ortogonalną i ortonormalną.

To jest w pewnym sensie odpowiednik bazy standardowej w $\mathbb{R}^n$.

Twierdzenie 10.9. *Niech $\mathbb{V}$ będzie skończenie wymiarową przestrzenią Euklidesową (unitarną). Wtedy $\mathbb{V}$ ma bazę ortonormalną.*

Dowód wynika z bardziej technicznego lematu:

Lemat 10.10. *Niech $\mathbb{V}$ będziskończenie wymiarową przestrzenią Euklidesową (unitarną), niech B będzie niezależnym układem ortogonalnym. Wtedy $\mathrm{LIN}(B) = \mathbb{V}$ lub istnieje $\vec{b'} \in \mathbb{V} \setminus B$, taki że $B' = B \cup \{\vec{b'}\}$ jest ortogonalny i niezależny.*

Dowód. Zauważmy, że bez zmniejszenia ogólności, możemy założyć, że B jest układem ortonormalnym: wystarczy każde $\vec{b} \in B$ przemnożyć przez skalar $\|\vec{b}\|^{-1}$.

Jeśli $B = \emptyset$ to bierzemy dowolny niezerowy wektor z $\mathbb{V}$ (jeśli $\mathbb{V} = \{\vec{0}\}$ to teza zachodzi dla $B = \emptyset$).

Jeśli $\mathrm{LIN}(B) = \mathbb{V}$ to teza oczywiście zachodzi.

Niech więc $\mathrm{LIN}\, B \neq \mathbb{V}$. Niech $v \notin \mathbb{W}$. Rozpatrzmy wektor

$$\vec{b'} = \vec{v} - \sum_{\vec{b} \in B} \left\langle \vec{b}, \vec{v} \right\rangle \vec{b}$$

Nie należy on do $\mathbb{W}$ (bo jest sumą $\vec{v}$ oraz wektora z $\mathbb{W}$) i łatwo sprawdzić, że należy do $\mathbb{W}^\perp$: sprawdźmy, żę jest prostopadły dla każdego $\vec{b''} \in B$:

$$\begin{aligned}
\left\langle \vec{b''}, \vec{v} - \sum_{\vec{b} \in B} \left\langle \vec{b}, \vec{v} \right\rangle \vec{b} \right\rangle &= \left\langle \vec{b''}, \vec{v} \right\rangle - \left\langle \vec{b''}, \sum_{\vec{b} \in B} \left\langle \vec{b}, \vec{v} \right\rangle \vec{b} \right\rangle \\
&= \left\langle \vec{b''}, \vec{v} \right\rangle - \sum_{\vec{b} \in B} \left\langle \vec{b}, \vec{v} \right\rangle \underbrace{\left\langle \vec{b''}, \vec{b} \right\rangle}_{0 \text{ dla } \vec{b} \neq \vec{b''}} \\
&= \left\langle \vec{b''}, \vec{v} \right\rangle - \left\langle \vec{b''}, \vec{v} \right\rangle \left\langle \vec{b''}, \vec{b''} \right\rangle \\
&= \left\langle \vec{b'}, \vec{v} \right\rangle - \left\langle \vec{b'}, \vec{v} \right\rangle \\
&= 0
\end{aligned}$$

Wtedy $B \cup \{\vec{b'}\}$ jest niezależnym układem ortogonalnym. □

Dowód Twierdzenia 10.9 wynika z Lematu 10.10, przy czym na początku bierzemy pusty zbiór niezależny.

Lemat 10.11. *Niech $\mathbb{V}$ będzie przestrzenią Euklidesową (unitarną), $\vec{v}_1, \ldots, \vec{v}_n$ bazą ortonormalną a $\vec{v}$ wektorem wyrażanym w tej bazie jako*

$$\vec{v} = \sum_{i=1}^{n} \alpha_i \vec{v}_i \quad .$$

Wtedy

$$\alpha_i = \langle \vec{v}, \vec{v}_i \rangle \quad .$$

Dowód.

$$\begin{aligned}\langle \vec{v}, \vec{v}_i \rangle &= \left\langle \sum_{j=1}^{n} \alpha_j \vec{v}_j, \vec{v}_i \right\rangle \\ &= \sum_{j=1}^{n} \alpha_j \langle \vec{v}_j, \vec{v}_i \rangle \\ &= \alpha_i \|v_i\|^2 \\ &= \alpha_i \ .\end{aligned}$$

□

Lemat 10.12. *Niech $\mathbb{V}$ będzie przestrzenią Euklidesową (unitarną). Niech $F : \mathbb{V} \to \mathbb{V}$ będzie przekształceniem linowym, zaś $B = \vec{v}_1, \ldots, \vec{v}_n$ bazą ortonormalną. Wtedy*

$$M_{BB}(F) = (\langle F(\vec{v}_j), \vec{v}_i \rangle)_{i,j=1,\ldots,n} \ .$$

Dowód. Wiemy z definicji $M_{BB}(F)$, że

$$M_{BB}(F) = [(F\vec{v}_1)_B | (F\vec{v}_2)_B | \cdots | (F\vec{v}_n)_B]$$

Teraz pozostaje skorzystać z Lematu10.11:

$$(F\vec{v}_j)_B = [\langle F\vec{v}_j, \vec{v}_1 \rangle, \ldots, \langle F\vec{v}_j, \vec{v}_n \rangle]^T \ . \quad \square$$

10.4 Dopełnienie ortogonalne

Definicja 10.13 (Dopełnienie ortogonalne)**.** Niech $U \subseteq \mathbb{V}$ będzie podzbiorem przestrzeni Euklidesowej (lub unitarnej). Wtedy *dopełnienie ortogonalne* U to:

$$U^{\perp} = \{\vec{v} \in \mathbb{V} \ : \ \forall_{\vec{w} \in U} \ \vec{v} \perp \vec{w}\}$$

Fakt 10.14. *Jeśli B jest bazą $\mathbb{W}$ to $\vec{v} \in \mathbb{W}^{\perp}$ wtedy i tylko wtedy, gdy $\vec{v}$ jest prostopadły do każdego wektora z B.*

Dowód. ⇒ Jeśli $\vec{v} \in \mathbb{W}^{\perp}$ to w szczególności jest prostopadły do każdego wektora z $B \subseteq W$.

⇐ Załóżmy, że $\left\langle \vec{v}, \vec{b}_i \right\rangle$ dla każdego wektora $\vec{b}_i \in B$. Wtedy dowolne $\vec{w} \in W$ wyraża się jako $\sum_i \alpha_i \vec{b}_i$ i dlatego:

$$\begin{aligned}\langle \vec{v}, \vec{w} \rangle &= \left\langle \vec{v}, \sum_i \alpha_i \vec{b}_i \right\rangle \\ &= \sum_i \alpha_i \left\langle \vec{v}, \vec{b}_i \right\rangle \\ &= \sum_i \alpha_i \cdot 0 \\ &= 0 \ .\end{aligned}$$

□

Lemat 10.15. *Niech $U \subseteq \mathbb{V}$, gdzie $\mathbb{V}$ jest przestrzenią Euklidesową (lub unitarną). Wtedy*

- *$U^{\perp} \leq \mathbb{V}$ jest przestrzenią liniową;*
- *$U \cap (U^{\perp}) \subseteq \{\vec{0}\}$;*
- *$(U^{\perp})^{\perp} \supseteq U$;*

Dowód.

- Niech $v, v' \in U^{\perp}$, czyli

  $$\forall_{u \in U} \ \langle v, u \rangle = \langle v', u \rangle = 0 \ .$$

  Dodając te dwie równości uzyskujemy

  $$\forall_{u \in U} \ \langle v + v', u \rangle = 0 \ ,$$

  a mnożąc pierwszą przez α:

  $$\forall_{u \in U} \ \langle \alpha v, u \rangle = 0 \ .$$

  Czyli $U^{\perp}$ jest zamknięta na dodawanie i mnożenie przez skalar, oczywiście należy do niej wektor $\vec{0}$, czyli jest podprzestrzenią liniową.

- Jeśli $u \in U \cap U^\perp$ to $\langle u, u\rangle = 0$ i tym samym $u = \vec{0}$. Łatwo zauważyć, że $\vec{0} \in \mathbb{W} \cap \mathbb{W}^\perp$.
- Niech $u \in U$. Wtedy dla każdego $v \in U^\perp$ mamy $\langle u, v\rangle = 0$, czyli $u \in (U^\perp)^\perp$. □

Lemat 10.16. *Jeśli $\vec{b}_1, \dots, \vec{b}_n$ jest bazą ortonogonalną przestrzeni Euklidesowej lub unitarnej $\mathbb{V}$, to*

$$\mathrm{LIN}(\vec{b}_1, \dots, \vec{b}_k)^\perp = \mathrm{LIN}(\vec{b}_{k+1}, \dots, \vec{b}_n) \ .$$

W szczególności, jeśli $\mathbb{W} \leq \mathbb{V}$ to

$$\dim(\mathbb{W}^\perp) = \dim \mathbb{V} - \dim \mathbb{W} \ .$$

Dowód. Skoro każde z $\vec{b}_{k+1}, \dots, \vec{b}_n$ jest prostopadłe do wektorów bazy $\mathrm{LIN}(\vec{b}_1, \dots, \vec{b}_k)$, to są prostopadłe do całej rozpostartej przestrzeni $\mathrm{LIN}(\vec{b}_1, \dots, \vec{b}_k)$ i tym samym

$$\mathrm{LIN}(\vec{b}_1, \dots, \vec{b}_k)^\perp \geq \mathrm{LIN}(\vec{b}_{k+1}, \dots, \vec{b}_n) \ .$$

By sprawdzić równość, rozpatrzmy dowolny wektor $\vec{v}$ spoza $\mathrm{LIN}(\vec{b}_{k+1}, \dots, \vec{b}_n)$; ma on reprezentację w bazie B równą $\sum_{i=1}^n \alpha_i \vec{b}_i$ oraz $\alpha_j \neq 0$ dla pewnego $j \leq k$. Wtedy

$$\begin{aligned}
\left\langle \vec{b}_j, \sum_i \alpha_i \vec{b}_i \right\rangle &= \sum_i \alpha_i \underbrace{\left\langle \vec{b}_j, \vec{b}_i \right\rangle}_{0 \text{ dla } i \neq j} \\
&= \alpha_j \left\langle \vec{b}_j, \vec{b}_j \right\rangle \\
&\neq 0 \ ,
\end{aligned}$$

czyli $\vec{v} \notin \mathrm{LIN}(\vec{b}_1, \dots, \vec{b}_k)^\perp$. Co kończy dowód pierwszego punktu.

W drugim punkcie zauważmy, że dla $\mathbb{W}$ można wybrać bazę ortogonalną B (Twierdzenie 10.9), która można rozszerzyć do bazy ortogonalnej $\mathbb{V}$ (Lemat 10.10), z pierwszej części dostajemy wtedy tezę. □

Lemat 10.17. *Niech $\mathbb{V}$ będzie skończenie-wymiarową przestrzenia Euklidesową (lub unitarną) oraz niech $\mathbb{W} \leq \mathbb{V}$. Wtedy*

- $(\mathbb{W}^\perp)^\perp = \mathbb{W}$.
- $\mathbb{W} + \mathbb{W}^\perp = \mathbb{V}$
- *dla każdego wektora $v \in \mathbb{V}$ reprezentacja $\vec{v} = \vec{w} + \vec{w}^\perp$, gdzie $\vec{w} \in \mathbb{W}$ i $\vec{w}^\perp \in \mathbb{W}^\perp$ jest jedyna.*

Dowód.
- Wiemy już, że $\mathbb{W} \leq (\mathbb{W}^\perp)^\perp$. Reszta wynika z policzenia wymiaru:

$$\begin{aligned}
\dim(\mathbb{W}^\perp)^\perp &= \dim \mathbb{V} - \dim \mathbb{W}^\perp \\
&= \dim \mathbb{V} - (\dim \mathbb{V} - \dim \mathbb{W}) \\
&= \dim \mathbb{W}
\end{aligned}$$

- Ponownie, wynika to z rachunku wymiarów:

$$\begin{aligned}
\dim(\mathbb{W} + \mathbb{W}^\perp) &= \dim(\mathbb{W}) + \dim(\mathbb{W}^\perp) - \dim(\mathbb{W} \cap \mathbb{W}^\perp) \\
&= \dim(\mathbb{W}) + \dim(\mathbb{W}^\perp) \\
&= \dim(\mathbb{W}) + (\dim \mathbb{V} - \dim \mathbb{W}) \\
&= \dim \mathbb{V}
\end{aligned}$$

  Czyli

$$\mathbb{W} + \mathbb{W}^\perp = \mathbb{V} \ .$$

- Niech $B, B_\perp$ to bazy ortogonalne $\mathbb{W}, \mathbb{W}^\perp$, wtedy $B \cup B_\perp$ jest bazą $\mathbb{V}$.

  Załóżmy, że reprezentacja $\vec{v} = \vec{w} + \vec{w}_\perp$, gdzie $w \in \mathbb{W}$ i $w_\perp \in \mathbb{W}^\perp$ Nie jest jedyna. Wyrażając je jako kombinację wektorów z $B, B_\perp$ dostajemy dwie różne reprezentacje $\vec{v}$ w bazie $B \cup B_\perp$. □

10.5 Rzuty i rzuty prostopadłe.

Definicja 10.18 (Rzut, rzut prostopadły). *Rzutem* nazywamy przekształcenie liniowe $P : \mathbb{V} \to \mathbb{V}$ takie że $P^2 = P$. O rzucie P mówimy, że jest *rzutem na podprzestrzeń* $\operatorname{Im} P$.

Rzut jest rzutem *prostopadłym* jeśli dla każdego $\vec{v}$ mamy $P(\vec{v}) \perp (\vec{v} - P(\vec{v}))$.

Lemat 10.19. *Niech $\mathbb{V}$ będzie przestrzenią Euklidesową (unitarną) i $\mathbb{W} \leq \mathbb{V}$. Rzut prostopadły na $\mathbb{W}$ jest zdefiniowany jednoznacznie.*

Niech $P : \mathbb{V} \to \mathbb{V}$ będzie rzutem prostopadłym na $\mathbb{W}$. Jeśli $\vec{b}_1, \ldots, \vec{b}_k$ jest bazą ortogonalną $\mathbb{W}$ zaś $\vec{b}_1, \ldots, \vec{b}_n$: przestrzeni $\mathbb{V}$, to

$$P\left(\sum_{i=1}^{n} \alpha_i \vec{b}_i\right) = \sum_{i=1}^{k} \alpha_i \vec{b}_i \ .$$

Uwaga. Zauważmy, że skoro $\vec{b}_1, \ldots, \vec{b}_n$ jest bazą, to to definiuje P na całej przestrzeni $\mathbb{V}$.

Dowód. Pokarzemy najpierw, że rzut prostopadły jest zdefiniowany jednoznacznie (o ile istnieje). Niech P będzie rzutem prostopadłym na $\mathbb{W}$. Ponieważ $\operatorname{Im} P = \mathbb{W}$, to dla $\vec{v}_i$ dla $1 \leq i \leq k$ istnieje $\vec{w}_i$, takie że $P(\vec{w}_i) = \vec{v}_i$. Wynika z tego, że $P(\vec{v}_i) = \vec{v}_i$. Weźmy $\vec{v}_i$ dla $i > k$. Wtedy

$$\begin{aligned} 0 &= \langle P\vec{v}_i, \vec{v}_i - P\vec{v}_i \rangle \\ &= -\langle P\vec{v}_i, P\vec{v}_i \rangle + \langle P\vec{v}_i, \vec{v}_i \rangle \\ &= -\langle P\vec{v}_i, P\vec{v}_i \rangle \end{aligned}$$

Z tego wynika, że $P\vec{v}_i = \vec{0}$ dla $i > k$.

Niech P' będzie przekształceniem zdefiniowanym jak powyżej.

Ponieważ wyrażenie w bazie jest jednoznaczne, łatwo zobaczyć, że $(P')^2 = P'$, tj. P' jest rzutem: weźmy dowolne $\vec{v}$, wyraża się ono jednoznacznie jako suma $\sum_{i=1}^{n} \alpha_i \vec{b}_i$ dla pewnych $\alpha_1, \ldots, \alpha_n$. Wtedy

$$\begin{aligned} (P')^2(\vec{v}) &= (P')^2 \left(\sum_{i=1}^{n} \alpha_i \vec{b}_i\right) \\ &= P'\left(\sum_{i=1}^{k} \alpha_i \vec{b}_i\right) \\ &= \sum_{i=1}^{k} \alpha_i \vec{b}_i \\ &= P'\left(\sum_{i=1}^{n} \alpha_i \vec{b}_i\right) \\ &= P'(\vec{v}) \ . \end{aligned}$$

Ponieważ, $\vec{b}_1, \ldots, \vec{b}_k \in \mathbb{W}$, mamy, że $\operatorname{Im} P' \leq \mathbb{W}$ i jednocześnie $\operatorname{Im} P' \geq \mathbb{W}$, bo $P'(\vec{v}_i) = \vec{v}_i$ dla $1 \leq i \leq k$. Czyli $\operatorname{Im} P' = \mathbb{W}$.

Ponadto

$$\sum_{i=1}^{n} \alpha_i \vec{b}_i - P'\left(\sum_{i=1}^{n} \alpha_i \vec{b}_i\right) = \sum_{i=k+1}^{n} \alpha_i \vec{b}_i \in \mathbb{W}^{\perp} \ ,$$

czyli jest to rzut prostopadły. □

Wniosek 10.20. Dla wektora $\vec{v}$ oraz P — rzutu prostopadłego na $\mathbb{W}$ — para $P(\vec{v})$, $\vec{v} - P(\vec{v})$ jest rozkładem $\vec{v}$ na wektory z $\mathbb{W}$, $\mathbb{W}^{\perp}$.

Zauważmy, że używając bazy ortonormalnej można wyrazić (abstrakcyjny) iloczyn skalarny w analogiczny sposób jak iloczyn standardowy, trzeba tylko przejść przez reprezentację w odpowiedniej bazie:

Lemat 10.21. *Jeśli $\langle \cdot, \cdot \rangle$ jest iloczynem skalarnym na przestrzeni Euklidesowej (lub unitarnej) $\mathbb{V}$, $B = \vec{b}_1, \ldots, \vec{b}_n$ jest bazą ortonormalną, to*

$$\langle \vec{u}, \vec{v} \rangle = (\vec{u})_B^T (\vec{v})_B$$

tj. wartość iloczynu skalarnego $\langle \vec{u}, \vec{v} \rangle$ to standardowy iloczyn skalarny reprezentacji $\vec{u}$ oraz $\vec{v}$.

W szczególności

$$\|\vec{u}\| = \|(\vec{u})_B\|$$

przy czym długość po prawej to zwykła długość wektorów w $\mathbb{R}^n$ ($\mathbb{C}^n$).

Dowód. Obie strony są liniowe względem obu współrzędnych, dlatego wystarczy pokazać dla elementów z bazy, czyli $u, v \in B$. Wtedy $[b_i]_B = \vec{E}_i$ i mamy

$$\langle b_i, b_j \rangle = \begin{cases} 1 & \text{dla } i = j \\ 0 & \text{dla } i \neq j \end{cases} \quad \text{oraz} \quad \vec{E}_i^T \cdot \vec{E}_j = \begin{cases} 1 & \text{dla } i = j \\ 0 & \text{dla } i \neq j \end{cases} .$$

Dla długości zauważmy, że

$$\begin{aligned} \|\vec{u}\| &= \sqrt{\langle \vec{u}, \vec{u} \rangle} \\ &= \sqrt{(\vec{u})^T (\vec{u})_B} \\ &= \|(\vec{u})_B\| \ , \end{aligned}$$

co kończy dowód. □

10.6 Algorytm Grama-Schmidta ortonormalizacji bazy

Używając terminologii rzutów możemy podać algorytm konstrukcji bazy ortonormalnej (przez ortogonalizację istniejącej bazy).

Dla bazy $\vec{v}_1, \dots, \vec{v}_n$ przestrzeni $\mathbb{V}$ z iloczynem skalarnym $\langle \cdot, \cdot \rangle$ algorytm (Grama-Schmidta) ortonormalizacji bazy wygląda następująco:

Algorytm 1 Algorytm Gram-Schmidta ortonormalizacji bazy

Założenie: $\vec{v}_1, \dots, \vec{v}_n$ są niezależne

```
1: v⃗_1 ← v_1 / sqrt(⟨v_1, v_1⟩)                         ▷ Normowanie
2: for i ← 2..n do
3:     v⃗_i ← v⃗_i − Σ_{j=1}^{i−1} ⟨v⃗_i, v⃗_j⟩ v⃗_j        ▷ Odjęcie rzutu na przestrzeń rozpiętą przez v⃗_1, ..., v⃗_{i−1}
4:     if ⟨v⃗_i, v⃗_i⟩ = 0 then
5:         return Wektory są liniowo zależne.
6:     v⃗_i ← v⃗_i / sqrt(⟨v⃗_i, v⃗_i⟩)                     ▷ Normowanie
```

Uwaga. Ostatni krok, w którym ortonormalizujemy kolejne wektory, nie jest w zasadzie potrzebny (i możemy dostać bazę ortogonalną), jednak w takim przypadku musimy zmienić odpowiednio wyrażenie na rzut prostopadły.

Twierdzenie 10.22. *Jeśli układ na wejściu algorytmu Grama-Schmidta był niezależny, to uzyskane wektory są układem ortonormalnym.*

Jeśli układ $\vec{v}_1, \dots, \vec{v}_i$ był zależny i układ $\vec{v}_1, \dots, \vec{v}_{i-1}$ był niezależny, to w czasie algorytmu przekształcimy $\vec{v}_i$ na $\vec{0}$.

Dowód. Niech ${v'}_i$ oznacza wektor w czasie działania algorytmu, zaś v_i jego wartość na wejściu.

Pokażemy przez indukcję, że po i-tej iteracji pętli mamy

- po odrzuceniu wektorów zerowych, układ ${v'}_1, \dots, {v'}_i$ jest ortonormalny;
- dla każdego j mamy $\mathrm{LIN}(v_1, \dots, v_j) = \mathrm{LIN}({v'}_1, \dots, {v'}_j)$.

Z założenia indukcyjnego układ ${v'}_1, \dots, {v'}_{i-1}$ jest ortonormalny, i w takim razie algorytm wykonuje rzut v_i na przestrzeń $\mathrm{LIN}({v'}_1, \dots, {v'}_{i-1})$, również z założenia równą $\mathrm{LIN}(v_1, \dots, v_{i-1})$. Ta operacje jest poprawnie określona, bo ${v'}_1, \dots, {v'}_{i-1}$ to układ ortonarmalny. Jeśli $v_i \in \mathrm{LIN}(v_i, \dots, v_{i-1})$, to uzyskamy $v'_i = 0$. Jeśli nie, to uzyskamy wektor prostopadły do $\mathrm{LIN}({v'}_1, \dots, {v'}_{i-1})$ i następnie zmienimy jego długość na 1, czyli uzyskany wektor v'_i jest ortonormalny.

Co do drugiej części, to zauważmy, że w i-tej iteracji zamieniamy wektor v_i na kombinację liniową v_i (ze współczynnikime 1) oraz wektorów ${v'}_1, \ldots, {v'}_{i-1}$. Czyli nie zmieniamy przestrzeni rozpostartej przez dowolny podciąg wektorów ${v'}_1, \ldots, {v'}_j$. □

Przykład 10.23. Dla standardowego iloczynu skalarnego w $\mathbb{R}^4$ zortonormalizujemy układ wektorów

$$\{(4, 4, -2, 0); (1, 4, 1, 0); (5, -4, -7, 1)\}$$

i uzupełnimy go do bazy ortonormalnej.

Oznaczmy zadane wektory jako v_1, v_2, v_3. Dokonamy ortonormalizacji bazy metodą Grama-Schmidta; niech v'_1, v'_2, v'_3 to wektory po tym procesie.

Długość wektora v_1 to to $\sqrt{16 + 16 + 4} = 6$, czyli pierwszy wektor ortonormalny z bazy to

$$v'_1 = \frac{1}{6} \cdot v_1 = \left(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}, 0\right).$$

Liczymy iloczyn skalarny tego wektora (v'_1) i wektora drugiego (v_2):

$$\langle v'_1, v_2 \rangle = \left\langle \left(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}, 0\right); (1, 4, 1, 0) \right\rangle = \frac{2}{3} + \frac{8}{3} - \frac{1}{3} = \frac{9}{3} = 3 \ ,$$

i tym samym

$$v_2 - 3v'_1 = (1, 4, 1, 0) - (2, 2, -1, 0) = (-1, 2, 2, 0) \ .$$

Jego długość to $\sqrt{1 + 4 + 4} = 3$ i dlatego

$$v'_2 = \frac{1}{3}(-1, 2, 2, 0) = \left(-\frac{1}{3}, \frac{2}{3}, \frac{2}{3}, 0\right) \ .$$

Obliczamy teraz iloczyny skalarne $\langle v'_1, v_3 \rangle$ oraz $\langle v'_2, v_3 \rangle$:

$$\langle v'_1, v_3 \rangle = \left\langle \left(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}, 0\right); (5, -4, -7, 1) \right\rangle = \frac{1}{3}(10 - 8 + 7) = \frac{9}{3} = 3$$
$$\langle v'_2, v_3 \rangle = \left\langle \left(-\frac{1}{3}, \frac{2}{3}, \frac{2}{3}, 0\right); (5, -4, -7, 1) \right\rangle = \frac{1}{3}(-5 - 8 - 14) = -\frac{27}{3} = -9$$

Obliczamy $v_3 - 3v'_1 + 9v'_2$:

$$(5, -4, -7, 1) - 3 \cdot \left(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}, 0\right) + 9\left(-\frac{1}{3}, \frac{2}{3}, \frac{2}{3}, 0\right) = (5 - 2 - 3, -4 - 2 + 6, -7 + 1 + 6, 1) = (0, 0, 0, 1)$$

Wektor ten ma długość 1, czyli

$$v'_3 = (0, 0, 0, 1).$$

Aby rozszerzyć ten układ wektorów do bazy ortonormalnej, należy dodać do niej jeden wektor (niezależny) i następnie zortonormalizować cały układ. Weźmy wektor $v_4 = (1, 0, 0, 0)$: ma on niewiele współrzędnych i nie wygląda, żeby był liniowo zależny od pozostałych:

$$\langle v'_1, v_4 \rangle = \frac{2}{3}$$
$$\langle v'_2, v_4 \rangle = -\frac{1}{3}$$
$$\langle v'_3, v_4 \rangle = 0$$

Obliczamy $v_4 - \frac{2}{3}v'_1 + \frac{1}{3}v'_2$:

$$(1, 0, 0, 0) - \frac{2}{3} \cdot \left(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}, 0\right) + \frac{1}{3}\left(-\frac{1}{3}, \frac{2}{3}, \frac{2}{3}, 0\right) = \left(1 - \frac{4}{9} - \frac{1}{9}, 0 - \frac{4}{9} + \frac{2}{9}, 0 + \frac{2}{9} + \frac{2}{9}, 0\right) = \left(\frac{4}{9}, -\frac{2}{9}, \frac{4}{9}, 0\right)$$

Długość tego wektora to:

$$\sqrt{\frac{1}{81}(16 + 4 + 16)} = \frac{6}{9} = \frac{2}{3}$$

Po przemnożeniu dostajemy

$$v'_4 = \frac{3}{2} \cdot \left(\frac{4}{9}, -\frac{2}{9}, \frac{4}{9}, 0\right) = \left(\frac{2}{3}, -\frac{1}{3}, \frac{2}{3}, 0\right),$$

który to wektor jest dopełnieniem do bazy ortonormalnej.

10.7 Zastosowania: geometria

10.7.1 Reprezentacja przez dopełnienie ortogonalna

Dopełnienie ortogonalne jest dobrym sposobem reprezentacji płaszczyzn/prostych itp.: dla danej płaszczyzny $\mathbb{W}$ reprezentujemy ją jako bazę $\mathbb{W}^{\perp}$. Reprezentacja ta jest o tyle dobra, że można łatwo przecinać tak zadane przestrzenie: dla $\mathbb{W}_1, \mathbb{W}_2$ ich przecięcie to $(\mathbb{W}_1^{\perp} + \mathbb{W}_2^{\perp})$. Zwartą reprezentację otrzymujemy przez ortonormalizację sumy baz $\mathbb{W}_1^{\perp}$, $\mathbb{W}_2^{\perp}$.

10.7.2 Symetrie

Macierz symetrii względem prostej dość łatwo zadać używając rzutu: symetria względem $\mathbb{W}$ wyraża się jako $2P_{\mathbb{W}} - \mathrm{Id}$.

Rozdział 11

Izometrie, macierze ortogonalne

11.1 Izometrie

Definicja 11.1 (Izometria)**.** Przekształcenie liniowe $F : \mathbb{V} \to \mathbb{V}$ na przestrzeni liniowej $\mathbb{V}$ z iloczynem skalarnym $\langle \cdot, \cdot \rangle$, nazywamy *izometrią*, jeśli zachowuje iloczyn skalarny, tj. dla każdych dwóch wektorów $u, v \in \mathbb{V}$ zachodzi:

$$\langle F\vec{v}, F\vec{u} \rangle = \langle \vec{v}, \vec{u} \rangle \quad .$$

Przykład 11.2.

- obrót o kąt α (na płaszczyźnie)
- symetria względem prostej
- symetria względem płaszczyzny
- symetria względem punktu

Lemat 11.3. *Przekształcenie F jest izometrią wtedy i tylko wtedy gdy zachowuje długość, tj. dla każdego $v \in \mathbb{V}$ mamy $\|F(\vec{v})\| = \|\vec{v}\|$.*

Przekształcenie F jest izometrią wtedy i tylko wtedy gdy zachowuje iloczyn skalarny elementów z bazy.

Dowód. $\Rightarrow$ Jeśli F jest izometrią, to w szczególności $\langle v, v \rangle = \langle F(v), F(v) \rangle$, czyli $\|v\| = \|F(v)\|$.

$\Leftarrow$ Jeśli F zachowuje długość, to zachowuje iloczyn skalarny v z v, tj. dla każdego v mamy $\langle v, v \rangle = \langle F(v), F(v) \rangle$. Podstawiając za wektor $u + v$ dostajemy:

$$\langle u + v, u + v \rangle = \langle F(u + v), F(u + v) \rangle$$

Rozwijając obie strony z liniowości:

$$\|u\|^2 + \|v\|^2 + 2 \langle u, v \rangle = \|F(u)\|^2 + \|F(v)\|^2 + 2 \langle F(u), F(v) \rangle$$

Ponieważ $\|u\| = \|F(u)\|$ oraz $\|v\| = \|F(v)\|$ dostajemy

$$\langle u, v \rangle = \langle F(u), F(v) \rangle \, .$$

Druga część zostanie pokazana na ćwiczeniach. □

11.2 Macierze ortogonalne

Definicja 11.4. Macierz kwadratową nazywamy *ortogonalną*, jeśli jej kolumny są parami ortogonalne oraz są długości 1 (w standardowym iloczynie skalarnym).

Lemat 11.5. *M jest ortogonalna wtedy i tylko wtedy gdy $M^{-1} = M^T$.*

Dowód. Zauważmy, że wyraz ij iloczynu $M^T M$ to standardowy iloczyn skalarny i-tej oraz j-tej kolumny.

$\Rightarrow$ Jeśli M jest ortogonalna, to wyraz ij iloczynu MM^T wynosi 1 dla $i = j$ oraz 0 dla $i \neq j$. Czyli $MM^T = \mathrm{Id}$.

$\Leftarrow$ Jeśli $M^{-1} = M^T$ to $MM^T = \mathrm{Id}$ i tym samym iloczyn i-tej oraz j-tej kolumny M to 0 dla $i \neq j$ oraz 1 dla $i = j$. Czyli kolumny stanowią układ ortonormalny. □

Lemat 11.6. *Macierze ortogonalne są zamknięte na mnożenie, transponowanie i na branie macierzy odwrotnej.*

Dowód. Jeśli A, B są ortogonalne, to

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

i z tego wnioskujemy, że również AB jest ortogonalna.

Jeśli A jest ortogonalna to

$$(A^{-1})^{-1} = A = (A^T)^T = (A^{-1})^T$$

i tym samym A^{-1} też jest ortogonalna. □

Lemat 11.7. *Jeśli F jest izometrią a $B = \{\vec{b}_1, \ldots, \vec{b}_n\}$ bazą ortonormalną to $M_{BB}(F)$ jest macierzą ortogonalną. W szczególności, $M_{BB}(F)^{-1}$ to $M_{BB}(F)^T$.*

Dowód. Rozpatrzmy standardowy iloczyn skalarny i-tej oraz j-tej kolumny $M_{BB}(F)$. Zgodnie z definicją $M_{BB}(F)$ są to $[F(b_i)]_B$ oraz $[F(b_j)]_B$. Z Lematu 10.21 ich standardowy iloczyn skalarny $\langle [F(b_i)]_B, [F(b_j)]_B \rangle$ to $\langle F(b_i), F(b_j) \rangle$. Ponieważ F jest izometrią, to to jest równe $\langle b_i, b_j \rangle$. Czyli 0 dla $i \neq j$ oraz 1 dla $i = j$; czego należało dowieść. □

Lemat 11.8. *Jeśli M jest macierzą ortogonalną, to indukowane przez nią przekształcenie liniowe $L_M : \vec{V} \mapsto M\vec{V}$ jest izometrią.*

Dowód. Z Lematu 11.3 wystarczy pokazać, że L_M zachowuje iloczyn skalarny elementów z bazy standardowej. Dla $\vec{E}_i, \vec{E}_j$ wiemy, że $\left\langle \vec{E}_i, \vec{E}_j \right\rangle$ jest równy 1 dla $i = j$ oraz 0 dla $i \neq j$.

Niech $M = [M_1 | M_2 | \ldots | M_k]$. Wtedy $\left\langle M\vec{E}_i, M\vec{E}_j \right\rangle = \langle M_i, M_j \rangle$ i z tego, że M jest ortogonalna wnioskujemy, że ten iloczyn wynosi 1 dla $i = j$ oraz 0 dla $i \neq j$. □

Rozdział 12

Macierze dodatnio określone: zadawanie iloczynu skalarnego przez macierz

Jak zadawać iloczyn skalarny na przestrzeni? Dla zadanej bazy $B = \vec{e}_1, \dots, \vec{e}_n$ iloczyn skalarny jest jednoznacznie zadany przez macierz $M = (a_{ij})_{i,j=1,\dots,n}$, gdzie

$$a_{ij} = \langle \vec{e}_i, \vec{e}_j \rangle$$

Wtedy

$$\langle \vec{u}, \vec{v} \rangle = (\vec{u})_B^T M (\vec{v})_B$$

(Wystarczy sprawdzić z liniowości dla $u = \vec{e}_i$ oraz $v = \vec{e}_j$).

Definicja 12.1 (Macierz iloczynu skalarnego, macierz Grama)**.** Dla bazy $B = \vec{v}_1, \dots, \vec{v}_n$ oraz iloczynu skalarnego $\langle \cdot, \cdot \rangle$ określamy macierz tego iloczynu w bazie B jako

$$M^B = (\langle \vec{v}_i, \vec{v}_j \rangle)_{i,j=1,\dots,n}$$

Lemat 12.2. *Niech B: baza przestrzeni z iloczynem skalarnym $\langle \cdot, \cdot \rangle$. Wtedy*

$$\langle \vec{u}, \vec{v} \rangle = (\vec{u})_B^T M^B (\vec{v})_B$$

Dowód. Niech $B = \vec{b}_1, \dots, \vec{b}_n$. Z liniowości po obu argumentach wystarczy sprawdzić $u = \vec{b}_i$ oraz $v = \vec{b}_j$, co jest prostym rachunkiem. ☐

Lemat 12.3. *Niech A, B to dwie bazy przestrzeni z iloczynem skalarnym. Wtedy*

$$M^B = M_{BA}^T M^A M_{BA},$$

gdzie M_{BA} to macierz zmiany bazy.

Dowód. Niech $B = \vec{b}_1, \dots, \vec{b}_n$. Wystarczy pokazać, że dla dowolnych wektorów u, v mamy

$$\langle \vec{u}, \vec{v} \rangle = [u]_B^T M_{BA}^T M^A M_{BA} [v]_B \quad ,$$

bo ta własność zachodzi dla M^B.

Policzmy

$$\begin{aligned}
[u]_B^T (M_{BA}^T M^A M_{BA}) [v]_B &= (M_{BA}[u]_B)^T M^A (M_{BA}[v]_B) \\
&= ([u]_A)^T M^A ([v]_A) \\
&= \langle \vec{u}, \vec{v} \rangle \quad .
\end{aligned}$$

☐

Fakt 12.4. *Dla bazy ortnormalnej B dla iloczynu skalarnego $\langle \cdot, \cdot \rangle$ mamy $M^B = \mathrm{Id}$.*

Dla jakich macierzy to jest dobra definicja? Na pewno macierz musi być symetryczna. Tak zadana funkcja na pewno jest liniowa po obu argumentach. Tak w zasadzie to chodzi o to, żeby zachodziło

$$\langle \vec{v}, \vec{v} \rangle > 0.$$

Definicja 12.5 (Macierz dodatnio określona)**.** Macierz M wymiaru $n \times n$ jest *dodatnio określona*, jeśli funkcja $\langle \cdot, \cdot \rangle : (\mathbb{R}^n)^2 \to \mathbb{R}$ określona jako

$$(\vec{u}, \vec{v}) \mapsto \vec{u}^T M \vec{v}$$

jest iloczynem skalarnym na $\mathbb{R}^n$.

Uwaga. Jeśli to jest iloczyn skalarny, to dla bazy standardowej E dla $\mathbb{R}^n$ mamy $M^E = M$.

Fakt 12.6. *Macierz M jest dodatnio określona wtedy i tylko wtedy gdy:*

1. *jest symetryczna oraz*
2. *dla każdego wektora $\vec{v} \neq 0$ zachodzi*

$$\vec{v}^T M \vec{v} > 0 \ .$$

Lemat 12.7. *M jest dodatnio określona $\iff$ $M = A^T A$ dla pewnej odwracalnej macierzy A.*
Takie A można efektywnie uzyskać z M.

Dowód. $\Leftarrow$ Jeśli $M = A^T A$ to $M^T = (A^T A)^T = A^T A$, czyli jest symetryczna.
Zaś

$$v^T M v = v^T A^T A v = (Av)^T (Av) \ .$$

Ponieważ $v \neq \vec{0}$ oraz A jest odwracalna, to $Av \neq \vec{0}$ i dlatego ma choć jedną niezerową współrzędną i $(Av)^T(Av) > 0$ (bo zawiera kwadrat tej współrzędnej).

$\Rightarrow$ Skoro M jest dodatnio określona, to wyznacza iloczyn skalarny. Liczymy bazę ortonormalną A dla tego iloczynu. Wyrażamy w niej ten iloczyn, wtedy $M^A = \mathrm{Id}$. Wyrażamy $M = M^E$ przy pomocy M^A:

$$M = M^E = M_{EA}^T M^A M_{EA} \ . \quad \square$$

Ale jak to efektywnie sprawdzić? (W zasadzie to powyższy opis już nam to powiedział).

Dla macierzy M niech M_k oznacza macierz $k \times k$ która jest „w lewym górnym rogu” macierzy M.

Twierdzenie 12.8 (Kryterium Sylvestera)**.** *Symetryczna macierz M jest dodatnio określona $\iff$ dla każdego $k = 1, 2, \dots, n$ macierz M_k spełnia* $\det(M_k) > 0$.

Dowód dla zainteresowanych, nie został przedstawiony na wykładzie. $\Rightarrow$ Popatrzmy na macierz M_k oraz na przestrzeń $\mathbb{V}_k$ rozpiętą przez pierwsze k wektorów bazowych. Wtedy M_k to macierz iloczynu skalarnego dla przestrzeni $\mathbb{V}_k$. Czyli $M_k = A^T A$ i musi mieć dodatni wyznacznik.

$\Leftarrow$

Będziemy rozważać funkcjonał dwuliniowy zadany przez macierz M. Zauważmy, że dla funkcjonałów również zachodzi Lemat 12.3.

Pokazujemy przez indukcję. Dla $n = 1$ to jasne, no bo to jest iloczyn wektora v_1 samego ze sobą. Z założenia indukcyjnego dostajemy, że to jest iloczyn skalarny na przestrzeni $\mathbb{V}_{n-1}$ rozpiętej przez pierwsze $n - 1$ wektorów. Obliczamy bazę ortonormalną $B = \vec{b}_1, \dots, \vec{b}_{n-1}$ dla $\mathbb{V}_{n-1}$ i rozszerzamy ją o v_n: dowolny wektor spoza $\mathrm{LIN}(\vec{b}_1, \dots, \vec{b}_{n-1})$.

Nasz funkcjonał dwuliniowy (formalnie nie wiemy, że to iloczyn skalarny) wyrażony w tej bazie to

$$M' = M_{BE}^T M M_{BE}.$$

Zauważmy, że

$$\det M' = \det M_{BE}^T \det M \det M_{BE} = (\det M_{BE})^2 \det M > 0.$$

Ortonormalizujemy v_n do pozostałych wektorów, uzyskując bazę B'; możemy to zrobić, bo dla ortonormalizacji wystarczy, że $\vec{b}_1, \dots, \vec{b}_{n-1}$ są układem ortonormalnym. Ta operacja to kolejna zmiana bazy, dostajemy więc, że funkcjonał dwuliniowy wyrażony w bazie B' ma macierz

$$M'' = M_{B'B''}^T M' M_{B'B''}$$

i M'' jest macierzą przekątniową. Ponownie

$$\det M'' = \det M_{B'B''}^T \det M' \det M_{B'B''} = (M_{B'B''})^2 M' > 0 \ .$$

Warunek, że $\det M'' > 0$ mówi tyle, że iloczyn elementów na przekątnej jest dodatni. Ale wiemy, że iloczyn wszystkich poza ostatnim jest dodatni (z założenia indukcyjnego macierz M_{n-1} jest dodatnio określona). Czyli wszystkie elementy na przekątnej są dodatnie. W takim razie możemy wyrazić M'' jako iloczyn AA^T i jest ona dodatnio określona. W takim razie również macierz M jest dodatnio określona. $\square$

Część II

Algebra Abstrakcyjna

Rozdział 13

Grupy

13.1 Automorfizmy

Definicja 13.1 (Grupa przekształceń (automorfizmów) obiektu). Dla danego z obiektu kombinatorycznego S jego grupa przekształceń (symetrii, automorfizmów) $G = \mathrm{Aut}(S)$ powinna spełniać następująco warunki

- przekształcenie identycznościowe e jest w G
- jeśli $\varphi_1, \varphi_2 \in G$ to te przekształcenia można *złożyć* uzyskując $\varphi = \varphi_1 \circ \varphi_2 \in G$
- dla każdego $\varphi \in G$ istnieje φ^{-1} takie że $\varphi^{-1}\varphi = \varphi\varphi^{-1} = e$

Przykład 13.2.

1. kwadrat i jego obroty
2. kwadrat i jego symetrie
3. dwudziestościan foremny i jego obroty
4. macierz $n \times n$ i mnożenie przez macierze odwracalne
5. macierz $n \times n$ i mnożenie przez macierze odwracalne o wyznaczniku 1
6. macierz $n \times n$ i mnożenie przez macierze odwracalne o module wyznacznika równym 1
7. $\mathbb{Z}$ i dodawania elementów z $\mathbb{Z}$
8. $\mathbb{Z}_p$ i dodawanie elementów z $\mathbb{Z}_p$
9. $\mathbb{Z}_p \setminus \{0\}$ z mnożeniem przez niezerowe elementy w $\mathbb{Z}_p$
10. X i bijekcje z X w X
11. zbiór $\{1, 2, \dots, n\}$ i jego permutacje
12. $2n$-kąt foremny i jego symetrie
13. $2n$-kąt foremny i jego obroty

13.2 Grupa

Abstrahujemy od obiektu. Same przekształcenia.

Definicja 13.3 (Grupa). Zbiór $(G, \cdot)$, gdzie $\cdot : G \times G \to G$ jest działaniem dwuargumentowym jest *grupą*, gdy:

łączność działanie $\cdot$ jest łączne;

element neutralny istnieje element neutralny e taki że dla każdego $g \in G$ mamy $ge = eg = g$;

element odwrotny dla każdego $g \in G$ istnieje g^{-1} speniajacy $g^{-1}g = gg^{-1} = e$.

Jeśli $\cdot$ jest przemienne, to mówimy o grupie *przemiennej* (*abelowej*).

Uwaga. Alternatywnie możemy zdefiniować grupę tak, że ma ona dodatkowo jedną operację unarną: $^{-1}: G \to G$ (branie elementu odwrotnego) oraz jedną stałą: e. Te operacje mają spełniać warunki podane w Definicji 13.3.

Można pokazać, że:

Lemat 13.4.

- *Element odwrotny w grupie G jest jedyny.*
- *Element prawostronnie odwrotny jest też lewostronnie odwrotny.*
- *Identyczność jest jedyna.*
- *Równanie $ax = b$ oraz $xa = b$ mają dokładnie jedno rozwiązanie.*

Przykład 13.5.

- $\{1, 3, 5, 7\}$ z mnożeniem $\bmod 8$ [Grupa Kleina]
- obroty kwadratu
- symetrie kwadratu
- obroty dwudziestościanu foremnego
- odwracalne macierze $n \times n$ (z mnożeniem)
- macierze $n \times n$ o wyznaczniku 1 (z mnożeniem)
- macierze $n \times n$ o module wyznacznika równym 1 (z mnożeniem)
- ortogonalne macierze $n \times n$ (z mnożeniem)
- $\mathbb{Z}$ z dodawaniem
- $\mathbb{Z}_n$ z dodawaniem modulo n
- $\mathbb{Z}_p \setminus \{0\}$ z mnożeniem (p — liczba pierwsza)
- bijekcje z X w X
- permutacje zbioru $\{1, 2, \ldots, n\}$
- obroty i symetrie $2n$-kąta foremnego
- obroty $2n$-kąta foremnego

Uwaga 13.6. Teoria grup została rozwinięta przy okazji rozwiązywania równań stopnia ≥ 5. Ale prawdziwa eksplozja nastąpiła w czasie drugiej wojny światowej i związków z kryptografią. Do dziś stanowi podstawę przy projektowaniu i analizy sposobów szyfrowania oraz kryptoanalizy.

13.2.1 Półgrupy

W ogólności rozważa się też monoidy (półgrupy), w których nie zakładamy istnienia elementu odwrotnego (elementu odwrotnego ani identyczności).

13.3 Tabelka działań

Definicja 13.7 (Tabela działań). *Tabela działań* dla grupy G podaje wprost wszystkie możliwe $|G|^2$ wyników mnożenia.

Przykład 13.8 (Tabela działań dla grupy Klein'a).

$\cdot$	1	3	5	7
1	1	3	5	7
3	3	1	7	5
5	5	7	1	3
7	7	5	3	1

Fakt 13.9.
- *Każdy wiersz i każda kolumna w tabelce działań jest permutacją elementów z G.*
- *Dwa różne wiersze (dwie różne kolumny) są różne.*
- *Musi być dokładnie jeden wiersz (kolumna) w której permutacja jest identycznością.*

Definicja 13.10 (Iloczyn kartezjański grup; produkt prosty)**.** Dla grup G, H przez $G \times H$ oznaczamy grupę na zbiorze $G \times H$ i działaniu po współrzędnych

$$(g, h) \cdot (g', h') = (gg', hh') \ .$$

Definicję rozszerzamy naturalnie na iloczyn kartezjański dowolnej ilości grup.

13.4 Homomorfizm, Izomorfizm

Definicja 13.11 (Homomorfizm, izomorfizm grup)**.** Operację $\varphi : \mathbb{G} \mapsto \mathbb{H}$ nazywamy *homomorfizmem grup*, jeśli zachowuje działanie grupowe, tj. $\varphi(ab) = \varphi(a)\varphi(b)$.

φ jest *izomorfizmem*, jeśli istnieje φ^{-1} które jest przekształceniem odwrotnym i homomorfizmem (w szczególności: φ, φ^{-1} są bijekcjami).

Przykład 13.12.
- Izomorfizm: grupa Kleina oraz $\mathbb{Z}_2 \times \mathbb{Z}_2$.
- Homomorfizm: macierze odwracalne rozmiaru $n \times n$ nad $\mathbb{F}$ $M \mapsto \det M$
- Homomorfizm: $\mathbb{Z}_2 \times \mathbb{Z}_2$ na pierwszą współrzędną.
- Homomorfizm: Macierze odwracalne w macierze o wyznaczniku ± 1: $\varphi(M) = M/|\det(M)|$.
- Homomorfizm: Macierze o wyznaczniku o module 1 w macierze o wyznaczniku 1: $\varphi(M) = M/\det(M)$.
- Homomorfizm: Obroty i symetrie kwadratu w $\mathbb{Z}_2$: czy zmieniają orientację, czy nie (tzn. symetrie w -1, obroty w 1).

Lemat 13.13. *Homomorfizm przeprowadza element neutralny (odwrotny) w neutralny (odwrotny).*

Dowód. Wynika to z tego, że homomorfizm zachowuje równania: jeśli krotka $(a_1, \ldots, a_n)$ elementów z G spełnia jakieś równania, to $\varphi(a_1), \ldots, \varphi(a_n)$ też spełnia analogiczne równanie. Wtedy identyczność to jedyny element spełniający równanie $x^2 = x$. Natomiast para a, a^{-1} spełnia równanie $xy = e$ (i jeśli jakaś para jest spełnia to jest parą elementów do siebie odwrotnych). Zauważmy, że formalnie e w obu grupach to inny element; aby ominąć tę trudność możemy rozszerzyć naszą parę o element z i dopisać równanie $z^2 = z$ (czyli z jest identycznością w obu grupach) i początkowe równanie zastąpić przez $xy = z$.

Zwykle jednak postępujemy tak jakby branie elementu odwrotnego oraz element neutralny były dodatkowymi operacjami w grupie. □

13.5 Rząd elementu

Definicja 13.14 (Potęga, rząd)**.** Potęgą elementu a nazywamy dowolny element postaci a^n, gdzie $n \in \mathbb{Z}$. Dla $n = 0$ oznacza on e, dla $n > 1$: $a^n = \underbrace{a \cdot a \cdots a}_{n \text{ razy}}$, dla $n < 0$: $a^n = (a^{-1})^{-n}$.

Rząd elementu to najmniejsza dodatnia potęga n taka że $a^n = e$. Rząd elementu jest *nieskończony* (*nieokreślony*), jeśli nie ma takiego skończonego n.

Rząd grupy to ilość jej elementów (może, ale nie musi, być skończony).

Fakt 13.15. *W grupie skończonej każdy element ma rząd skończony.*

Lemat 13.16. *Jeśli $a \in G$ ma skończony rząd p, to $a^\ell = e \iff p|\ell$.*

Dowód. $\Leftarrow$ jest jasna.

$\Rightarrow$: założmy, że tak nie jest. Bez zmniejszenia ogólności możemy rozpatrzyć tylko dodatnie ℓ. Rozpatrzmy najmniejsze takie ℓ, że $a^\ell = e$ oraz $p \nmid \ell$. Wtedy $\ell > p$, bo inaczej p nie jest rzędem a. Ale wtedy $a^{\ell - p} = e$, sprzeczność z minimalnością ℓ. □

13.6 Podgrupy

Definicja 13.17 (Podgrupa). H jest podgrupą G, co zapisujemy jako $H \leq G$, gdy $H \subseteq G$ oraz jest grupą.

Uwaga. Nie wystarczy, że $H \subseteq G$ i że jest zamknięta na działanie: może nie zawierać elementu odwrotnego! (np. $\mathbb{N} \subseteq \mathbb{Z}$ jest zamknięte na działanie, ale nie ma elementów odwrotnych.)

Uwaga. Dla alternatywnej definicji (z dodatkową operacją branie elementu odwrotnego oraz elementem neutralnym) już wystarczy, bo wtedy to są formalnie działania.

Przykład 13.18.

- Grupa obrotów $2n$ kąta w grupie symetrii $2n$ kąta.
- Dodawanie liczb parzystych w $\mathbb{Z}$.
- $\mathbb{Z}_2 \times \{0\}$ w $\mathbb{Z}_2 \times \mathbb{Z}_2$.
- Macierze o wyznaczniku o module 1 w macierzach odwracalnych.
- Macierze o wyznaczniku 1 w macierzach odwracalnych.
- Macierze ortogonalne w macierzach.

Lemat 13.19. *W grupie skończonej G zbiór H jest podgrupą, gdy jest zamknięty na działanie.*

W grupie, w której rząd każdego elementu jest skończony, H jest podgrupą, gdy jest zamknięty na działanie.

Dowód. Zauważmy, że jeśli element a ma rząd k, to $a^{-1} = a^{k-1}$. W naszym przypadku oznacza to, że jeśli zbiór jest zamknięty na działanie, to jest też zamknięty na branie element odwrotnego i zawiera e. □

Definicja 13.20 (Generowanie). Dla grupy G oraz zbioru $A \subseteq G$ podgrupa generowana przez A, oznaczana jako $\langle A \rangle$, to najmniejsza podgrupa G zawierająca A. W takim wypadku mówimy, że A to zbiór *generatorów* tej podgrupy.

Przykład 13.21.

- $\mathbb{Z} = \langle 1 \rangle = \langle 3, 5 \rangle$
- $\mathbb{Z}_6 = \langle 1 \rangle = \langle 2, 3 \rangle$
- grupa obrotów kwadratu jest generowana przez obrót o 90^o
- grupa obrotów i symetrii kwadratu jest generowana przez obrót o 90^o i dowolną symetrię.

Fakt 13.22. $(x_1^{z_1} x_2^{z_2} \cdots x_k^{z_k})^{-1} = (x_k^{-1})^{z_k} (x_{k-1}^{-1})^{z_{k-1}} \cdots (x_1^{-1})^{z_1}$.

Prosty dowód pokazany zostanie na ćwiczeniach.

Definicja 13.23 (Postać zredukowana). O ciągu elementów $a_1^{\ell_1} a_2^{\ell_1} \cdots a_k^{\ell_k}$ mówimy, że są w *postaci zredukowanej*, jeśli $a_i \notin \{a_{i+1}^{-1}, a_{i+1}\}$ dla każdego możliwego i oraz $\ell_i \neq 0$ dla każdego i.

Lemat 13.24. *Dla każdego ciągu elementów $a_1^{\ell_1} a_2^{\ell_1} \cdots a_k^{\ell_k}$ istnieje $a_{i_1}^{\ell'_1} a_{i_2}^{\ell'_j} \cdots a_{i_j}^{\ell'_j}$ w postaci zredukowanej, taki że*

$$a_1^{\ell_1} a_2^{\ell_2} \cdots a_k^{\ell_k} = a_{i_1}^{\ell'_1} a_{i_2}^{\ell'_2} \cdots a_{i_j}^{\ell'_j}$$

oraz $a_{i_1}, a_{i_2}, \ldots, a_{i_j}$ jest podciągiem $a_1, a_2, \ldots, a_k$. Taka reprezentacja jest jedyna.

Dowód. Dowód jest intuicyjnie prosty: postać zredukowaną uzyskuje się przez kolejne wykreślanie aa^{-1}. Techniczne i ciut żmudne jest pokazanie, że postać zredukowana jest jedyna, tj. że wynik nie zależy od koleności wykonania skreśleń. (Dla tych, co znają pojecia: że ten system przepisywania termów jest silnie konfluentny). □

Lemat 13.25. *Dla zbioru generatorów X podgrupa $\langle X \rangle$ to dokładnie zbiór elementów postaci:*

$$\langle X \rangle = \{x_1^{z_1} x_2^{z_2} \cdots x_k^{z_k} \ : \ k \geq 0,\, x_1, \ldots, x_k \in X,\, z_1, \ldots, z_k \in \mathbb{Z}\} \ ,$$

bez zmniejszenia ogólności można dodatkowo założyć, że wszystkie elementy są w postaci zredukowanej.

Dowód. W oczywisty sposób $\langle X \rangle$ zawiera wszystkie elementy tej postaci.

W drugą stronę należy pokazać, że tak zadany zbiór jest grupą; co też jest proste, bo jest zamknięty na złączanie ciągów elementów oraz na branie elementów odwrotnych. □

13.7 Grupa cykliczna

Definicja 13.26 (Grupa cykliczna)**.** Grupa G jest *grupą cykliczną*, gdy $G = \langle\{a\}\rangle$ dla pewnego $a \in G$, tzn. jest generowana przez jeden element.

Uwaga. Grupa cykliczna nie musi być skończona: $\mathbb{Z} = \langle 1\rangle$. Dzieje się tak dlatego, że dla generatora dopuszczamy też ujemne potęgi.

Fakt 13.27. *Każda grupa cykliczna jest przemienna.*

Dowód. Z Lematu 13.25 wiemy, że każdy element w grupie cyklicznej jest postaci a^k lub $(a^{-1})^k$ dla pewnego k. A mnożenie takich elementów jest przemienne. □

Lemat 13.28. *Dla każdego $n < \infty$ wszystkie grupy cykliczne rzędu n są izomorficzne (z $(\mathbb{Z}_n, +)$). Wszystkie grupy cykliczne nieskończonego rzędu są izomorficzne (z $(\mathbb{Z}, +)$).*

Dowód. Dowód tego lematu polega głównie na zrozumieniu definicji oraz określeniu tego, co w zasadzie należy dowieść.

Niech $G = \langle g\rangle, H = \langle h\rangle$ będą grupami cyklicznymi tego samego rzędu p. Określamy $\varphi : G \to H$ jako $\varphi(g^k) = h^k$.

Po pierwsze, φ jest dobrze określone: jeśli $g^m = g^\ell$ to $p|(m - \ell)$ i w takim razie $h^m = h^\ell$.

Należy pokazać, że φ jest izomorfizmem. Jest to bijekcja i ma przekształcenie odwrotne (łatwo widać).

Pozostało pokazać, że φ jest homomorfizmem (bo przekształcenie odwrotne jest zdefiniowane analogicznie). $\varphi(g^k g^\ell) = h^{k+\ell}$ i jednocześnie $\varphi(g^k)\varphi(g^\ell) = h^k h^\ell = h^{k+\ell}$. □

13.8 Grupa wolna

Definicja 13.29 (Grupa wolna)**.** Niech Σ to zbiór różnych elementów (liter), $\Sigma^{-1} = \{a^{-1} : a \in \Sigma\}$ będzie rozłączny z Σ.

Grupa wolna o generatorach $\Sigma = \{a, b, \ldots, c\}$ to zbiór wszystkich słów

$$\{a_1^{k_1} a_2^{k_2} \cdot a_m^{k_m} \ : \ m \geq 0,\, a_1, \ldots, a_m \in \Sigma \cup \Sigma^{-1},\, a_{i+1} \neq a_i \neq a_{i+1}^{-1},\, k_1, \ldots, k_m > 0\} \ .$$

Mnożenie to konkatenacja po której następują wszystkie możliwe skracanie elementów $xx^{-1} \to \epsilon$ oraz $a^k a^\ell \to a^{k+\ell}$.

Można pokazać, że jest to dobrze zdefiniowane.

Rozdział 14

Grupy permutacji

Definicja 14.1. Grupa permutacji S_n to zbiór wszystkich bijekcji ze zbior $\{1, 2, \ldots, n\}$ w siebie; operacją jest składanie funkcji, tj.

$$(\sigma' \cdot \sigma)(i) = \sigma'(\sigma(i)).$$

Permutacje zapisujemy jako dwuwierszową tabelkę:

$$\begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \cdots & \sigma(n) \end{pmatrix} .$$

Przykład 14.2. Permutacje S_4:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 3 \end{pmatrix} .$$

Z takiej reprezentacji łatwo wyliczyć permutację odwrotną: wystarczy zamienić miejscami i przesortować kolumny.

Z dokładnością do izomorfizmu każda grupa jest grupą permutacji.

Twierdzenie 14.3 (Cayley). *Dla każdej grupy G (o n elementach) istnieje podgrupa S_n izomorficzna z G.*

Dowód. Niech $G = \{g_1, g_2, \ldots, g_n\}$. Po pierwsze, o grupie permutacji możemy myśleć, że operuje na elementach $\{g_1, g_2, \ldots, g_n\}$ a nie $\{1, 2, \ldots, n\}$. Zdefiniujmy grupę permutacji na G. Dla elementu g definiujemy permutację $\sigma_g(g_i) = gg_i$. Nasza podgrupa to $\{\sigma_g : g \in G\}$. A izomorfizm to $g \mapsto \sigma_g$.

na podgrupę definiujemy tak, że funkcja jest na (podgrupa to obraz tego przekształcenia)

różnowartościowość jeśli $g \neq g'$ to w szczególności: $\sigma_g(e) = g \neq g' = \sigma_{g'}(e)$, czyli $\sigma_g \neq \sigma_{g'}$.

homomorfizm weźmy g, g'. Wtedy $\sigma_g \circ \sigma'_g(h) = gg'h = \sigma_{gg'}(h)$.

grupa Jeśli $\sigma_g, \sigma_{g'}$ są w tej grupie, to $\sigma_g \sigma_{g'} = \sigma_{gg'}$ też tam jest, bo jest obrazem gg'. Analogicznie, jeśli σ_g jest w grupie, to jest też w niej $\sigma_{g^{-1}}$, które jest elementem odwrotnym do σ_g. □

14.1 Rozkład permutacji na cykle

Definicja 14.4 (cykl). *Cykl* σ to taka permutacja, że istnieją elementy $a_1, \ldots, a_n$, że $\sigma(a_i) = a_{i+1}$ (gdzie $\sigma(a_n) = a_1$) a na innych elementach jest identycznością. Cykl taki zapisujemy jako $(a_1, a_2, \ldots, a_n)$. Elementy $\{a_1, \ldots, a_n\}$ to *dziedzina cyklu* lub *nośnik cyklu*, mówimy też o cyklu *na elementach* $\{a_1, \ldots, a_n\}$.

Długość cyklu $(a_1, \ldots, a_n)$ to n.

Cykle są *rozłączne*, gdy ich nośniki nie mają takich wspólnego elementu.

Transpozycja to cykl dwuelementowy. *Transpozycja elementów sąsiednich* to transpozycja postaci $(i, i+1)$.

Lemat 14.5. *1. Rząd cyklu długości k to k.*

2. Dla cyklu $(a_1, \ldots, a_n)$ permutacja odwrotna to $(a_1, \ldots, a_n)^{-1} = (a_n, a_{n-1}, \ldots, a_2, a_1) = (a_1, \ldots, a_n)^{n-1}$.

3. *Jeśli* $\{c_i\}_{i=1}^k$ *są parami rozłącznymi cyklami, to* $c_{i_1}c_{i_2}\cdots c_{i_k}$ *jest tą samą permutacją, niezależnie od wyboru permutacji* $i_1,\ldots,i_k$ *liczb* $1,\ldots,k$.

4. *Jeśli* $\{c_i\}_{i=1}^k$ *są parami rozłącznymi cyklami, to rząd* $c_1\cdots c_k$ *to* nww *rzędów poszczególnych cykli* $c_1,\ldots,c_k$.

5. *Jeśli* $\sigma = c_1c_2\cdots c_k$, *gdzie* $c_1,\ldots,c_k$ *są parami rozłącznymi cyklami , to* $\sigma^{-1} = c_1^{-1}c_2^{-1}\cdots c_k^{-1}$.

Dowód. Ad 1) Oczywiste.

Ad 2) Oczywiste.

Ad 3) Jako że te cykle są parami rozłączne, to można zamienić każde możliwe dwa sąsiednie. Wystarczy ustawić na skrajnie lewym miejscu pierwszy, potem drugi itp.

Ad 4) Popatrzmy na

$$(c_{i_1}c_{i_2}\cdots c_{i_k})^\ell \ .$$

Jako że są to cykle rozłączne, to można zamieniać parami sąsiednie tak aby dostać grupowanie

$$(c_{i_1}c_{i_2}\cdots c_{i_k})^\ell = c_{i_1}^\ell c_{i_2}^\ell \cdots c_{i_k}^\ell \ .$$

Biorąc za ℓ najmniejszą wspólną wielokrotność uzyskujemy, że każde $c_i^\ell = e$. Jeśli ℓ nie jest wielokrotnością któregoś z rzędów, to któreś c_i^ℓ nie jest identycznością, z Lematu 13.16, Jako że inne cykle nie ruszają elementów z jego nośnika, ta permutacja nie jest wtedy identycznością.

Ad 5) Wystarczy zauważyć, że jeśli $c_1,\ldots,c_k$ są parami rozłącznymi cyklami, to również $c_1^{-1},\ldots,c_k^{-1}$ są parami rozłącznymi cyklami.

□

Twierdzenie 14.6. 1. *Każda permutacja* σ *jednoznacznie (z dokładnością do kolejności cykli) rozkłada się na rozłączne cykle.*

2. *Cykl długości* k *jest złożeniem* $k-1$ *transpozycji.*

3. *Każda transpozycja jest złożeniem nieparzystej liczby transpozycji elementów sąsiednich (niekoniecznie rozłącznych).*

4. *Każda permutacja da się przedstawić jako złożenie transpozycji (niekoniecznie rozłącznych).*

5. *Każda permutacja da się przedstawić jako złożenie transpozycji sąsiednich (niekoniecznie rozłącznych).*

6. *Grupa* S_n *jest generowana przez zbiór transpozycji (sąsiednich).*

Dowód. Ad 1) Bierzemy dowolny element i, obliczamy kolejno $\sigma^1(i), \sigma^2(i), \ldots$, aż coś się powtórzy. Musi to być i: w przeciwnym przypadku $\sigma(j) = \sigma(j')$, co nie jest możliwe. Czyli dostajemy cykl. Powtarzamy operację na kolejnych elementach spoza skonstruowanych już cykli. Nie możemy wejść do starego cyklu, bo każdy element w nim ma już ustalony przeciwobraz.

Ad 2) Pokażemy to dla cyklu $(1,2,\ldots,n)$. Załóżmy, że przedstawiliśmy już $(1,2,\ldots,n-1)$ jako złożenie transpozycji. Chcemy wydłużyć ten cykl: $n-1$ przesłać na n a n na 1. W tym celu trzeba nałożyć $(1,n)$ (z lewej):

$$(1,n)(1,2,\ldots,n-1) = (1,2,\ldots,n-1,n)$$

Czyli $(1,2,\ldots,n) = (1,n)(1,n-1)\cdots(1,2)$.

Ad 3) Popatrzmy na (i,j). Popatrzmy na złożenie $(j-1,j)(j-2,j-1)\cdots(i+1,i+2)(i,i+1)$. Łatwo zauważyć (albo pokazać przez indukcję po j), że i jest przekształcane kolejno na $i+1, i+2, \ldots, j$. Jednocześnie każda inna liczba ℓ z przedziału od i do j jest modyfikowana tylko raz, w transpozycji $(\ell-1,\ell)$, tym samym jest przekształcana na $\ell-1$.

Wracamy zamieniające wartość permutacji na j na kolejne elementy $j-1, j-2, \ldots$, przy okazji ustawiając właściwe wartości dla $j-1, j-2, \ldots$: nakładamy kolejno transpozycje $(j-1, j-2)$, $(j-2, j-3)$, ..., $(i+1, i)$, tj.

$$(i+1, i) \cdots (j-2, j-3)(j-1, j-2) \ .$$

Znowu łatwo zauważyć, lub pokazać indukcyjnie, że po nałożeniu $(\ell+1, \ell) \cdots (j-2, j-3)(j-1, j-2)$ j będzie przekształcane na ℓ, liczby z przedziału $[\ell+1, \ldots, j-1]$ same na siebie a pozostałe jak poprzednio. Czyli ostatecznie dostaniemy, że

$$(i, j) = (i+1, i) \cdots (j-2, j-3)(j-1, j-2) \cdot (j-1, j)(j-2, j-1) \cdots (i+1, i+2)(i, i+1) \ .$$

Pozostałe tezy wynikają bezpośrednio z tych pokazanych. □

Uwaga. Od teraz alternatywnym sposobem zapisu permutacji jest podanie jej jako iloczynu cykli rozłącznych. Np.:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 \\ 7 & 5 & 4 & 12 & 3 & 13 & 1 & 10 & 14 & 9 & 6 & 2 & 11 & 8 \end{pmatrix} = (1,7)(2,5,3,4,12)(6,13,11)(8,10,9,14) \ .$$

14.2 Permutacje parzyste i nieparzyste.

Ważna funkcja: parzystość permutacji (znak permutacji).

Definicja 14.7 (Inwersje, parzystość permutacji). Dla f będącej bijekcją z podzbioru liczb naturalnych w ten sam zbiór (czyli w szczególności permutacji) *inwersja* to para (i, j), taka że $i < j$ oraz $f(i) > f(j)$.

Parzystość permutacji to parzystość ilości jej inwersji.

Znak $\text{sgn}(\sigma)$ permutacji σ to $+1$, gdy σ jest parzysta i -1 gdy nieparzysta.

Uwaga. To jest własność w grupie permutacji, nie własność algebraiczna: grupy generowane przez cykl $(1, 2)$ oraz $(1, 2)(3, 4)$ są izomorficzne (z $\mathbb{Z}_2$), ale $(1, 2)$ jest nieparzysta, a $(1, 2)(3, 4)$ jest parzysta.

Lemat 14.8. *Niech $\sigma, \sigma' \in S_n$ będą permutacjami. Wtedy*

$$\text{sgn}(\sigma'\sigma) = \text{sgn}(\sigma') \, \text{sgn}(\sigma) \ .$$

Dowód. Pokażemy najpierw dla σ' będącego transpozycją sąsiednich elementów, tj. $\sigma' = (i, i+1)$. Popatrzmy na inwersje w

$$\sigma'\sigma \ .$$

Popatrzmy na czwórki $(k, \ell, \sigma(k), \sigma(\ell))$ i zastanówmy się, dla których z nich zmienia się, czy są inwersją, czy nie po nałożeniu σ'.

- jeśli $\sigma(k), \sigma(\ell) \notin \{i, i+1\}$ to dla pary k, ℓ nic się nie zmienia;
- jeśli $\{\sigma'(k), \sigma'(\ell)\} = \{i, i+1\}$ to dla pary k, ℓ zmienia się, czy jest inwersją (na przeciwny status)
- dla pozostałych par mamy, że jedno z $\sigma'(k)$, $\sigma'(\ell)$ jest w $\{i, i+1\}$ a jedno nie; niech to pierwsze to k a drugie ℓ. Ale wtedy $\sigma(\sigma'(k))$ zmienia się z i na $i+1$ (lub z $i+1$ na i) i tym samym dalej jest mniejsze/większe niż $\sigma(\sigma'(\ell)) = \sigma(\ell)$.

Dla dowolnego $\sigma'\sigma$: wyrażamy σ oraz σ' jako iloczyn transpozycji: $\sigma = \prod_{i=1}^{n} \sigma_i$, $\sigma' = \prod_{i=1}^{m} \sigma'_i$. Wtedy

$$\begin{aligned} \text{sgn}(\sigma) &= \prod_{i=1}^{n} \text{sgn}(\sigma_i) = (-1)^n \\ \text{sgn}(\sigma') &= \prod_{i=1}^{m} \text{sgn}(\sigma'_i) = (-1)^m \\ \text{sgn}(\sigma'\sigma) &= \prod_{i=1}^{n} \text{sgn}(\sigma_i) \prod_{i=1}^{m} \text{sgn}(\sigma'_i) = (-1)^{n+m} \end{aligned}$$

z czego widać, że $\text{sgn}(\sigma\sigma') = \text{sgn}(\sigma) \, \text{sgn}(\sigma')$. □

Wniosek 14.9. sgn jest homomorfizmem z S_n w $(\{-1, +1\}, \cdot)$.

Lemat 14.10.

- *Cykl parzysty jest permutacją nieparzystą.*
- *Cykl nieparzysty jest permutacją parzystą.*
- *Parzystość permutacji to parzystość ilości cykli parzystych w rozkładzie na cykle rozłączne.*
- *Permutacje parzyste stanowią podgrupę A_n, która ma $\frac{n!}{2}$ permutacji.*

Dowód. Dla pierwszych trzech punktów wystarczy skorzystać z charakteryzacji z Twierdzenia 14.6.

W ostatnim punkcie: Oczywiście jest to podgrupa; zauważmy, że przemnożenie przez (ustaloną) permutację nieparzystą to bijekcja między permutacjami parzystymi i nieparzystymi. Co daje, że $|A_n| = \frac{n!}{2}$. □

Przykład 14.11. Zwykle w celu policzenia parzystości danej explicite permutacji najłatwiej jest to zrobić licząc jej przedstawienie w postaci cykli rozłącznych. Przykładowo, dla rozważanej wcześniej permutacji

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 \\ 7 & 5 & 4 & 12 & 3 & 13 & 1 & 10 & 14 & 9 & 6 & 2 & 11 & 8 \end{pmatrix} = (1,7)(2,5,3,4,12)(6,13,11)(8,10,9,14)$$

łatwo sprawdzamy, że jest ona parzysta: ma dwa cykle długości parzystej.

14.3 Wyznacznik

Wartość wyznacznika macierzy $M = (a_{ij})_{i,j=1,\ldots,n}$ można zadać wprost jako:

$$|(a_{ij})_{i,j=1,\ldots,n}| = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i} \ .$$

Prosty dowód pozostawiamy jako ćwiczenie.

Rozdział 15

Działania grupy na zbiorze

15.1 Mnożenie podzbiorów grupy

W podzbiorach grupy G definiujemy działanie:

$$U \cdot W = \{uw \ : \ u \in U, w \in W\}.$$

Łatwo sprawdzić, że jest ono łączne, czyli

$$U \cdot (W \cdot V) = (U \cdot W) \cdot V = U \cdot W \cdot V.$$

z tak zdefiniowanym działaniem są monoidem (mają jedność, jest to $\{e\}$).

Dla zbiorów jednoelementowych zwykle będziemy opuszczać nawiasy oznaczające zbiór i pisać $a \cdot U$ lub po prostu aU, opuszczając znak działania grupowego.

Będziemy też korzystać z rozdzielności mnożenia względem sumy (mnogościowej), tzn.:

$$U(V \cup W) = UV \cup UW \quad \text{oraz} \quad (V \cup W)U = VU \cup WU \ .$$

Fakt 15.1. *Niech G grupa, $H \leq G$ to jej podgrupa. Wtedy:*

- $gG = Gg = G$;
- $gH = H \iff Hg = H \iff gH \subseteq H \iff Hg \supseteq H \iff gH \supseteq H \iff Hg \subseteq H \iff g \in H.$

Dowód. G jest zamknięta na mnożenie i dlatego

$$gG \subseteq G$$

W szczególności

$$g^{-1}G \subseteq G$$

Mnożąc obustronnie przez g dostajemy

$$G \subseteq gG$$

Dla mnożenia z prawej strony postępujemy analogicznie; co daje pierwszy punkt.

Jeśli $g \in H$ to z pierwszego punktu mamy $gH = Hg = H$. Jeśli $gH \subseteq H$ to w szczególności $ge = g \in H$, analogicznie postępujemy dla $Hg \subseteq H$. Jeśli $gH \supseteq H$ to mnożąc obustronnie przez g^{-1} otrzymujemy $g^{-1}H \subseteq H$, czyli $g^{-1} \in H$, czyli $g \in H$. □

15.2 Działanie grupy na zbiorze

Definicja 15.2 (Działania grupy na zbiorze)**.** Mamy zbiór obiektów kombinatorycznych $\mathcal{C}$ oraz grupę permutacji jego elementów $S(\mathcal{C})$, oznaczaną przez S.

Działanie grupy G na $\mathcal{C}$ to homomorfizm z G w S. Zwykle zapisujemy to działanie jako $g(c)$ lub nawet gc, pomijając homomorfizm.

Działanie będziemy też rozszerzali do podzbiorów G w naturalny sposób: jeśli G działa na $\mathcal{C}$ to dla $U \subseteq G$ definiujemy

$$Uc = \{gc \ : \ g \in U\}$$

Orbita elementu c: $Gc = \{g(c) \ : \ g \in G\}$

Stabilizator elementu c: $\{g \in G \ : \ g(c) = c\}$. Zauważmy, że G_c to największy zbiór taki że $G_c c = c$.

Przykład 15.3. Rozpatrzmy zbiór sześciennych kostek ze ścianami pomalowanymi na biało lub czarno. Działa na nim grupa obrotów (obrotów i odbić) sześcianu. Zauważmy, że grupa obrotów ma fizyczną interpretację, grupa obrotów i odbić już nie bardzo.

Rozpatrzmy zbiór możliwych kostek domina (pola od 0 do 6). Działa na niej grupa obrotów (obrotów i odbić) prostokąta. Podobnie, grupa obrotów ma sens fizyczny, obrotów i odbić mniej (chyba że są ze szkła).

Lemat 15.4. *Niech G działa na zbiorze $\mathcal{C}$, zaś $s \in \mathcal{C}$. Wtedy stabilizator G_s jest podgrupą G.*

Prosty dowód pozostawiamy jako ćwiczenie.

Lemat 15.5. *Niech G działa na zbiorze $\mathcal{C}$, zaś $c, c' \in \mathcal{C}$. Wtedy O_c, $O_{c'}$ są równe lub rozłączne.*

Dowód. Jeśli O_c, $O_{c'}$ są rozłączne to OK.

Jeśli O_c, $O_{c'}$ nie są rozłączne, to istnieje ich element wspólny, które jest postaci $gc = g'c'$ dla pewnych $g, g' \in G$. Wymnóżmy tę równość lewostronnie przez G:

$$Ggc = Gg'c' \ .$$

Jako że $Gg = G = Gg'$ oraz $Gc = O_c$ i $Gc' = O_{c'}$ dostajemy

$$O_c = O_{c'} \ . \quad \square$$

Lemat 15.6. *Niech G działa na zbiorze $\mathcal{C}$, zaś $s \in \mathcal{C}$. Wtedy $|O_s| \cdot |G_s| = |G|$.*

Dowód. Popatrzmy na gs dla różnych $g \in G$, takich elementów jest $|G|$, ale niektóre są takie same. W ogólności

$$\{gs \ : \ g \in G\} = O_s \ .$$

Pytanie, dla ilu $g \in G$ otrzymujemy ten sam element w O_s. Twierdzimy, że dla $|G_s|$, co da tezę.

Ustalmy $g_0 \in G$ i popatrzmy na zbiór $\{g \ : \ gs = g_0 s\}$. Twierdzimy, że jest on postaci $g_0 G_s$:

$\Leftarrow$

$$g_0 G_s s = g_0(G_s s) = g_0 s.$$

$\Rightarrow$ Jeśli $gs = g_0 s$ to

$$g_0^{-1} gs = s$$

i tym samym $g_0^{-1} g \in G_s$ i $g = g_0(g_0^{-1} g) \in g_0 G_s$. $\square$

Wniosek 15.7. Niech G działa na zbiorze $\mathcal{C}$, zaś $s \in \mathcal{C}$. Wtedy $|O_s|$ oraz $|G_s|$ dzielą $|G|$.

Przykład 15.8. Rozpatrzmy sześcian i grupy: obrotów oraz obrotów i symetrii, rozpatrujemy działanie na zbiorze ścian. Popatrzmy na ustaloną ścianę. Wielkość jej orbity to 6. Wielkość stabilizatora to 4 (dla obrotów) lub 8 (dla obrotów i odbić). Czyli rzędy tych grup to 24 lub 48.

Wyjdzie tyle samo, gdybyśmy rozpatrywali wierzchołki (orbita ma 8 elementów, stabilizator 3 lub 6).

15.3 Lemat Burnside'a

Zliczanie orbit działania grupy odpowiada zliczaniu „nierozróżnialnych” względem działania grupy obiektów, np. kostek nierozróżnialnych ze wzlęgu na obrót itp.

Przykład 15.9. Rozważmy planszę do gry w kółko i krzyżyk. Każde pole ma przypisany jeden z trzech symboli: kółko, krzyżyk lub nic. Naszą grupą będzie grupa symetrii kwadratu. Dwie plansze, które można na siebie przeprowadzić przy użyciu elementów tej grupy uważamy za „identyczne” (bo ten sam gracz wygrywa, taka sama jest strategia itp.). Takie „identyczne” plansze to dokładnie orbita planszy względem działania tej grupy. A czym są „różne” plansze? To dokładnie różne orbity. Pytając o różne plansze, pytamy o zbiór orbit.

Pokażemy, jak policzyć moc zbioru orbit.

Definicja 15.10 (Punkty stałe). Dla grupy G działającej na zbiorze $\mathcal{C}$ mówimy, że $c \in \mathcal{C}$ jest *punktem stałym* $g \in G$ jeśli $g(c) = c$. Zbiór punktów stałych g oznaczamy przez

$$\text{fix}(g) = \{c \in \mathcal{C} \ : \ g(c) = c\} \ .$$

Twierdzenie 15.11 (Lemat Burnside'a). *Niech G działa na zbiorze $\mathcal{C}$ a $\mathcal{O}$ będzie zbiorem orbit tego działania. Wtedy*

$$|\mathcal{O}| = \frac{1}{|G|} \sum_{g \in G} |\operatorname{fix}(g)| \quad .$$

Dowód. Popatrzmy na $|\{(g,x) \, : \, gx = x\}|$.

$$\begin{aligned}
|\{(g,x) \, : \, gx = x\}| &= \sum_x \sum_{g : gx = x} 1 \\
&= \sum_x |G_x| \\
&= \sum_x \frac{|G|}{|O_x|} \\
&= |G| \sum_x \frac{1}{|O_x|} \\
&= |G| \sum_{O \in \mathcal{O}} \sum_{x \in O} \frac{1}{|O|} \\
&= |G| \sum_{O \in \mathcal{O}} 1 \\
&= |G||\mathcal{O}| \\
|\{(g,x) \, : \, gx = x\}| &= \sum_g \sum_{x : gx = x} 1 \\
&= \sum_g |\operatorname{fix}(g)|
\end{aligned}$$

□

Przykład 15.12 (Kontynuacja Przykładu 15.9). Rozważmy planszę do gry w kółko i krzyżyk. Każde pole ma przypisany jeden z trzech symboli: kółko, krzyżyk lub nic. Naszą grupą będzie grupa symetrii kwadratu. Policzmy liczbę punktów stałych dla poszczególnych przekształceń:

e Każde z 3^9 ustawień jest punktem stałym.

o_{90^0} Cztery narożniki muszą być tego samego koloru (bo przechodzą cyklicznie na siebie), tak samo 4 pola zewnętrzne, możemy też dowolnie ustalić kolor pola środkowego. Czyli mamy 3^3 punktów stałych.

o_{270^0} Tak samo jak o_{90^0}.

o_{180^0} Przeciwległe narożniki są tego samego koloru, tak samo przeciwległe pola zewnętrzne. Czyli 3^5 punktów stałych.

symetria wzdłuż przekątnej (dwie takie symetrie)

pola na przekątnej przechodzą same na siebie, pozostałe 6 wymienia się parami. Czyli 3^6.

symetria przez bok Podobnie, jak wyżej: 3 pola przechodzą same na siebie, pozostałe grupują się parami.

Czyli w sumie

$$(3^9 + 2 \cdot 3^3 + 3^5 + 4 \cdot 3^6)/8$$

Warto sprawdzić, że naprawdę wyszła liczba całkowita...

Rozdział 16

Warstwy, Twierdzenie Lagrange'a

16.1 Warstwy

Najprostsze działanie grupy na sobie: przez mnożenie (z lewej strony). Co się dzieje z podzbiorami? A dokładniej: z podgrupą?

Definicja 16.1 (Warstwa)**.** Gdy $H \leq G$ to *warstwą lewostronną* H (w G) są zbiory postaci

$$aH = \{ah \, : \, h \in H\}$$

zaś prawostronnną

$$Ha = \{ha \, : \, h \in H\}$$

dla $a \in G$.

Zbiór warstw lewostronnych H w G oznaczamy przez G/H.

My będziemy myśleć głównie o warstwach lewostronnych.

Lemat 16.2.

- *Każde dwie warstwy są równoliczne.*
- *Każde dwie warstwy lewostronne (prawostronne) są rozłączne lub identyczne.*

Dowód. Działanie grupy G na zbiorze warstw lewostronnych przekształca dowolną warstwę w dowolną inną:

$$(g'g^{-1})gH = g'H$$

Czyli $|g'H| \leq |gH|$; z symetrii mamy równość (a

Lemat 16.3. *Niech $H \leq G$. Wtedy*

- $g_0H = g_1H \iff g_1^{-1}g_0 \in H \iff g_0^{-1}g_1 \in H$
- $Hg_0 = Hg_1 \iff g_1g_0^{-1} \in H \iff g_0g_1^{-1} \in H$

Dowód. Jeśli $g_0H = g_1H$ to mnożymy obie strony przez g_1^{-1}, otrzymując $g_1^{-1}g_0H = H$, co jest równoważne temu, że $g_1^{-1}g_0 \in H$. Resztę pokazuje się analogicznie. □

poza tym to to przekształcenie jest odwracalne).

Załóżmy, że $gH \cap g'H$ jest niepuste. W takim razie

$$gh = g'h'$$

dla jakichś $h, h' \in H$. Domnażając z prawej strony przez H dostajemy

$$ghH = gh'H$$

ale $h \in H$ oznacza, że $hH = H$, analogicznie $h'H = H$. Czyli

$$gH = g'H \ . \quad \square$$

Aby sprawdzić, czy dwa elementy sa w tej samej warstwie nie musimy liczyć ich warstw:

Wniosek 16.4. W grupie skończonej

- (Twierdzenie Legrange'a) Rząd podgrupy dzieli rząd grupy.
- Rząd elementu dzieli rząd grupy.
- Każda grupa o p-pierwszym elementach jest cykliczna i każdy jej element (poza e) jest generatorem.
- Dla każdego a zachodzi $a^{|G|} = e$.

Dowód. Niech $H \leq G$.

- Zbiór warstw względem H to partycja G, jednocześnie wszystkie są równoliczne i jedna z nich to H.
- Dla danego g rząd p to $|\langle p\rangle|$, korzystamy z poprzedniego punktu.
- Weźmy $g \neq e$; generuje podgrupę, jej rząd dzieli p i nie jest to 1, czyli to jest p. □

Wniosek 16.5 (Małe Twierdzenie Fermat'a). Jeśli $p \not| a$ to $a^{p-1} \bmod p = 1$.

Dowód. Wystarczy pokazać dla $a \in \{0, 1, \ldots, p-1\}$. Popatrzmy na $\mathbb{Z}_p \setminus 0$ z mnożeniem. To jest grupa, ma $p-1$ elementów. Czyli $a^{p-1} = e$ w $\mathbb{Z}_p \setminus 0$. Czyli to jest 1 modulo p. □

Definicja 16.6 (Indeks podgrupy)**.** Indeks podgrupy H względem grupy G to ilość warstw lewostronnych H w G, oznaczamy przez $G : H$.

Wartość jest taka sama, jeśli weźmiemy warstwy prawostronne. Zwykle zajmujemy się przypadkiem, kiedy indeks podgrupy jest skończony (a najczęśniej tym, że obie grupy są skończone).

Przykład 16.7. Naszą grupą będą obroty i odbicia kwadratu; niech wierzchołki kwadratu będą ponumerowane $1, 2, 3, 4$, w kolejności przeciwnej do ruchu wskazówek zegara, 1 w prawym dolnym rogu. Ta grupa ma 8 elementów (identyczność, obrót o 90^0, 180^0, 270^0, symetrie względem przekątnych, symetria pionowa i symetria pozioma) i możemy o niej myśleć jak o podgrupie S_4, czyli te elementy to $e; (1,2,3,4); (1,3)(2,4); (1,4,3,2); (1,3); (2,4$

Weźmy podgrupę obrotów, ma 4 elementy $e; (1,2,3,4); (1,3)(2,4); (1,4,3,2)$. Ma też dwie warstwy (warstwa lewostronna i prawostronna zgadzają się): sama ta grupa $\{e; (1,2,3,4); (1,3)(2,4); (1,4,3,2)\}$ oraz odbicia $\{(1,3); (2,4); (1,4)(2,3); (1,2)(3,4)\}$.

Weźmy grupę generowaną przez symetrię pionową, ta grupa ma dwa elementy (symetria pionowa $(1,4)(2,3)$ i identyczność e). Warstwy lewostronne:

- $e\{e, (1,4)(2,3)\} = \{e, (1,4)(2,3)\}$.
- $(1,2,3,4)\{e, (1,4)(2,3)\} = \{(1,2,3,4), (2,4)\}$.
- $(1,3)(2,4)\{e; (1,4)(2,3)\} = \{(1,3)(2,4); (1,2)(3,4)\}$.
- $(1,4,3,2)\{e, (1,4)(2,3)\} = \{(1,4,3,2); (1,3)\}$.

Warstwy prawostronne:

- $\{e, (1,4)(2,3)\}e = \{e, (1,4)(2,3)\}$.
- $\{e, (1,4)(2,3)\}(1,2,3,4) = \{(1,2,3,4); (1,3)\}$.
- $\{e; (1,4)(2,3)\}(1,3)(2,4) = \{(1,3)(2,4); (1,2)(3,4)\}$.
- $\{e, (1,4)(2,3)\}(1,4,3,2) = \{(1,4,3,2); (2,4)\}$.

Przykład 16.8. Grupa permutacji na 3 elementach (S_3). Podgrupa generowana przez cykl $(1,2,3)$ ma 3 elementy. Czyli ma dwie warstwy (ta podgrupa: permutacje parzyste i pozostałe elementy: permutacje nieparzyste).

Podgrupa generowana przez cykl $(1,2)$ (innymi słowy: wszystkie permutacje, które trzymają 3 w miejscu). Ma dwa elementy, czyli ma 3 warstwy lewostronne i 3 prawostronne.

Lewostronne

- $\{e, (1,2)\}$;

- $(1,3)\{e,(1,2)\} = \{(1,3);(1,2,3)\}$;
- $(2,3)\{e,(1,2)\} = \{(2,3);(1,3,2)\}$.

Opis: na co przechodzi 3; opis można wyprowadzić z Lematu 16.3 — zauważmy, że nasza podgrupa to zbiór elementów, które nie ruszają 3.

Prawostronne

- $\{e,(1,2)\}$;
- $\{e,(1,2)\}(1,3) = \{(1,3);(1,3,2)\}$;
- $\{e,(1,2)\}(2,3) = \{(2,3);(1,2,3)\}$.

Opis: co przechodzi na 3; jak wyżej — można go wyprowadzić z Lematu 16.3.

Rozdział 17

Homomorfizmy i grupy ilorazowe, podgrupy normalne.

17.1 Homomorfizmy

Definicja 17.1 (Jądro, obraz homomorfizmu)**.** Dla homomorfizmu $\varphi : G \to H$ jego *obraz* to $\operatorname{Im} \varphi = \{\varphi(g) : g \in G\} = \varphi(G)$ zaś *jądro* to $\ker \varphi = \{g : \varphi(g) = e\} = \varphi^{-1}(e)$.

Lemat 17.2. *Dla homomorfizmu $\varphi : G \to H$ jego jądro i obraz to podgrupy, odpowiednio G oraz H.*

Dowód. Jądro: jeśli $\varphi(a) = e$ to $\varphi(a^{-1}) = e^{-1} = e$. Ponadto, jeśli $\varphi(a) = \varphi(b) = e$ to $\varphi(ab) = e$.

Obraz. Jeśli $a, a' \in \operatorname{Im} \varphi$ to istnieją b, b' takie że $\varphi(b) = a, \varphi(b') = a'$ i wtedy $\varphi(bb') = aa'$. Ponadto $\varphi(b^{-1}) = a^{-1}$. □

Jaki jest związek między podgrupami a homomorfizmami? Miedzy podgrupami a jądrem jakiegoś homomorfizmu?

Definicja 17.3 (Podgrupa normalna)**.** H jest *podgrupą normalną* G, gdy $aH = Ha$ dla każdego elementu $a \in G$; zapisujemy to jako $H \trianglelefteq G$.

Przykład 17.4.

1. Trywialna podgrupa $\{e\}$ jest zawsze normalna.
2. Grupa alternująca A_n jest normalną podgrupą S_n.
3. Grupa obrotów kwadratu jest normalną podgrupą jego symetrii.
4. Wszystkie podgrupy grupy przemiennej są normalne.
5. Centrum każdej grupy jest podgrupą normalną.
6. Podgrupa grupy $S_4 : \{e, (1, 2)(3, 4), (1, 3)(2, 4), (1, 4)(2, 3)\}$ jest normalna.
7. Każda podgrupa indeksu 2 jest normalna.
8. Współrzędna w produkcie grup jest zawsze normalna.

Lemat 17.5. *Następujące warunki są równoważne dla podgrupy H*

1. *$aH = Ha$ dla każdego elementu a;*
2. *$aH \subseteq Ha$ dla każdego elementu a;*
3. *$aH \supseteq Ha$ dla każdego elementu a;*
4. *$aHa^{-1} = H$ dla każdego elementu a;*
5. *$aHa^{-1} \subseteq H$ dla każdego elementu a;*
6. *$aHa^{-1} \supseteq H$ dla każdego elementu a.*

Dowód. Pokażemy równoważność trzech pierwszych warunków a następnie równoważność warunku i oraz $i+3$.

$(1 \Rightarrow 2)$ Oczywiste.

$(2 \Rightarrow 3)$ Mnożąc $aH \subseteq Ha$ z lewej i prawej przez a^{-1} dostajemy $Ha^{-1} \subseteq a^{-1}H$.

$(1 \Rightarrow 2)$ Jak wyżej, mnożąc przez a^{-1} z lewej i prawej dostajemy 2, 3 i 2 to 1.

$(i \Leftrightarrow i+3)$ Należy pomnożyć z lewej przez a^{-1} lub z prawej przez a. □

Definicja 17.6 (Podgrupa sprzężona)**.** Dla $H \leq G$ podgrupa postaci gHg^{-1} to *podgrupa sprzężona* do H.

Fakt 17.7. *Podgrupy sprzężone są izomorficzne. W ogólności dla $g \in G$ przekształcenie $h \mapsto gxg^{-1}$ jest izomorfizmem grupy z samą sobą (może to być identyczność).*

Lemat 17.8. *Jeśli $\varphi : G \to H$ jest homomorfizmem, to $\ker\varphi$ jest podgrupą normalną.*

Dowód. Niech $N = \ker\varphi$. Wystarczy pokazać, że $gNg^{-1} \subseteq N$. W tym celu wystarczy pokazać, że $\varphi(gNg^{-1}) = e$, czyli że $\varphi(gng^{-1}) = e$ dla $n \in N$:

$$\begin{aligned}
\varphi(gng^{-1}) &= \varphi(g)\varphi(n)\varphi(g^{-1}) \\
&= \varphi(g)e\varphi(g^{-1}) \\
&= \varphi(g)\varphi(g^{-1}) \\
&= \varphi(gg^{-1}) \\
&= \varphi(e) \\
&= e
\end{aligned}$$

□

17.2 Działanie na warstwach

Popatrzmy na działanie mnożenia podzbiorów grupy w ograniczeniu do warstw (prawostronnych) $H \trianglelefteq G$. Wtedy

$$\begin{aligned}
(aH)(bH) &= (Ha)(bH) \\
&= (H(ab))H \\
&= ((ab)H)H \\
&= (ab)(HH) \\
&= (ab)H \quad .
\end{aligned} \tag{17.1}$$

I tym samym zbiór tych warstw jest zamknięty na tak zdefiniowane mnożenie.

Definicja 17.9 (Grupa ilorazowa)**.** Gdy H jest podgrupą normalną G, to zbiór warstw H w G, czyli G/H, ma strukturę grupy dla działania:

$$aH \cdot bH = abH$$

Grupę tę nazywamy *grupą ilorazową.*

Lemat 17.10. *„Grupa ilorazowa" jest grupą.*

Dowód. Zgodnie z (17.1) jest ona zamknięta na tak zdefiniowane mnożenie.

Łączność istnieje, bo jesteśmy w półgrupie podzbiorów G z mnożeniem.

Element neutralny to $eH = H$.

Element odwrotny łatwo podać: dla aH to $a^{-1}H$. □

17.3 Naturalny homomorfizm $G \mapsto G/H$.

Lemat 17.11. *Niech $H \trianglelefteq G$ będzie podgrupą normalną G. Wtedy naturalny rzut z G na warstwy G, tj. $\pi_H : G \mapsto G/H$, gdzie $\pi_H(a) = aH$, jest homomorfizmem; co więcej, $H = \ker \pi_H$.*

Dowód. Trzeba sprawdzić, że jest to homomorfizm: dla $g, g' \in G$:

$$\begin{aligned} \pi_H(gg') &= gg'H \\ &= gHg'H \\ &= \pi_H(g)\pi_H(g') \ . \end{aligned}$$

Analogicznie pokazujemy, że $\pi_H(g^{-1}) = \pi_H(g)^{-1}$.

Jądro to $\{g \ : \ gH = H\}$, czyli dokładnie H. □

Twierdzenie 17.12. *Niech $\varphi : G \to G'$ będzie homomorfizmem. Wtedy istnieje izomorfizm $\psi : G/\ker\varphi \to \operatorname{Im}\varphi$.*

Dowód. Oznaczmy $H = \ker\varphi$.

Izomorfizm definiujemy jako $\psi(aH) = \varphi(a)$.

Kwestia sprawdzenia definicji:

dobrze określone Jeśli $aH = bH$ to

$$\begin{aligned} \varphi(aH) &= \varphi(a)\varphi(H) \\ &= \varphi(a)e \\ &= \varphi(a) \ . \end{aligned}$$

W szczególności, wartość ψ nie zależy od wyboru reprezentanta warstwy.

na jasne, bo chcemy na $\operatorname{Im}\varphi$ i dla dowolnego $a \in G$ mamy $\varphi(aH) = \psi(a)$.

różnowartościowość Załóżmy, że $\psi(aH) = \psi(bH)$. Wtedy, jak dwa punkty temu: $\varphi(a) = \varphi(aH) = \varphi(bH) = \varphi(b)$ czyli $\varphi(a^{-1}b) = e$ i tym samym jest w jądrze. Czyli $a^{-1}b \in H$ i w takim razie $aH = bH$.

homomorfizm Weźmy $\psi(aH) = \varphi(a)$, $\psi(bH) = \varphi(b)$. Wtedy $\psi(aHbH) = \psi(abH) = \varphi(ab)$. □

17.4 Kongruencje, konstrukcja $\mathbb{Z}_n$

To pozwala na zdefiniowanie kongruencji dla podgrupy normalnej $H \trianglelefteq G$:

$$a \equiv_H b \leftrightarrow aH = bH \qquad \iff \qquad a \equiv_H b \leftrightarrow a^{-1}b \in H \qquad \iff \qquad a \equiv_H b \leftrightarrow ba^{-1} \in H$$

(Zauważmy też, że $aH = Ha$ oraz $bH = Hb$.)

To jest *kongruencja*:

Definicja 17.13 (Kongruencja w grupie)**.** Relacja $\equiv \subseteq G^2$ na grupie G jest *kongruencją*, jeśli:

relacja równoważności jest relacją równoważności *oraz*

zachowuje działania zachowuje działania, tzn. dla każdych $a, a', b, b' \in G$ zachodzi:

$$\begin{aligned} a \equiv b \wedge a' \equiv b' \to aa' \equiv bb' \\ a \equiv b \to a^{-1} \equiv b^{-1} \end{aligned}$$

Poprawność definicji kongruencji $\equiv_H$ można policzyć wprost, ale nie trzeba: wynika z tego, że przekształcenie $a \mapsto aH$ jest homomorfizmem.

17.4.1 Konstrukcja $\mathbb{Z}_m$

Ważny przykład: $\mathbb{Z}_n$: kongruencja na $\mathbb{Z}$ względem podgrupy „liczby podzielne przez n”, zwyczajowo określanej jako $n\mathbb{Z}$. Jako że $\mathbb{Z}$ jest przemienna, to ta podgrupa jest normalna. Czyli mamy podgrupę normalną, konstrukcję $\mathbb{Z}_n$ oraz kongruencję na $\mathbb{Z}$.

Rozdział 18

Pierścienie, ciała, arytmetyka modularna

18.1 Pierścienie

Definicja 18.1 (Pierścień)**.** Pierścień, oznaczany zwykle przez R, to zbiór z dwoma działaniami $+, \cdot$, spełniającymi warunki:

- $(R, \cdot)$ jest półgrupą (niekoniecznie przemienną)
- $(R, +)$ jest grupą przemienną

Ponadto zachodzi rozdzielność mnożenia względem dodawania

- $a(b + c) = ab + ac$, $(b + c)a = ba + ca$

Pierścień jest *z jednością*, jeśli ma element neutralny dla mnożenia. Pierścień jest *przemienny*, jeśli $ab = ba$ (czyli półgrupa ze względu na mnożenie jest półgrupą przemienną).

Dalej będziemy się zajmować w zasadzie tylko i wyłącznie pierścieniami przemiennymi z jednością.

Definicja 18.2. Ciało $\mathbb{F}$ to pierścień przemienny z jednością, w którym $(\mathbb{F}, \cdot)$ jest grupą, tzn. każdy element ma element odwrotny, oraz elementy neutralne dodawania i mnożenia są różne („$0 \neq 1$”).

Przykład 18.3.

- liczby całkowite $\mathbb{Z}$
- macierze o współczynnikach z dowolnego ciała (pierścień nieprzemienny!)
- $\mathbb{Z}_m$: liczby modulo m z dodawaniem i mnożeniem
- $R[x]$ wielomiany o współczynnikach z R
- $R[[x]]$ szeregi formalne o współczynnikach z R.

Twierdzenie 18.4. *$\mathbb{Z}_m$ jest ciałem $\iff$ m jest pierwsze.*

Dowód pokażemy w dalszej części rozdziału.

18.2 Arytmetyka modularna $\mathbb{Z}_m$

.

Definicja 18.5 (Liczenie modulo, $\mathbb{Z}_m$)**.** a przystaje do b modulo m gdy $m|(a - b)$. Oznaczenie:

$$a \equiv_m b \ .$$

Reszta z dzielenia przez m:

$$a \mod m = b \iff a \equiv_m b \wedge b \in \{0, 1, \dots, m - 1\} \ .$$

W zasadzie to liczymy tylko reszty z dzielenia itp. dla liczb dodatnich.

Lemat 18.6. *Dla dowolnego $m \in \mathbb{Z}_+$ relacja $\equiv_m$ jest* kongruencją ze względu na mnożenie i dodawanie*, tzn.:*

$$a \equiv_m b \wedge a' \equiv_m b' \Rightarrow aa' \equiv_m bb'$$
$$a \equiv_m b \wedge a' \equiv_m b' \Rightarrow a + a' \equiv_m b + b' \ .$$

Wniosek 18.7. Przekształcanie $n \mapsto n \bmod m$ jest homomorfizmem pierścieni $\mathbb{Z}$ i $\mathbb{Z}_m$.

To ważne o tyle, że wykonując działania $\bmod m$ możemy dowolnie przełączać się między $\mathbb{Z}$ i $\mathbb{Z}_m$.

W sumie to chcielibyśmy więcej: czy „prawa” przenoszą się między $\mathbb{Z}$ i $\mathbb{Z}_m$? Na pewno nie wszystkie: umiemy powiedzieć, że w $\mathbb{Z}$ są conajmniej 3 różne elementy, ale to nie jest prawda w $\mathbb{Z}_3$. Okazuje się, że prawa się przenoszą, jeśli nie używają negacji.

Definicja 18.8 (Formuła pozytywna)**.** Niech t_1, t_2 będą wyrażaniami zbudowanymi z nawiasów, zmiennych $x_1, x_2, \ldots, x_n$, elementów z A oraz działań $+, \cdot$. Wtedy formuła ψ składająca się spójników $\wedge, \vee$ oraz równości $t_1 = t_2$, gdzie t_1, t_2 są jak wyżej, nazywamy *formuła pozytywną.*

Lemat 18.9. *Niech ψ będzie formuła pozytywną, zaś $\varphi : A \mapsto B$ będzie homomorfizmem na pierścień B.*

Jeśli

$$Q_1 x_1 Q_2 x_2 \ldots Q_n x_n \psi(x_1, x_2, \ldots, x_n)$$

zachodzi w A, to w B zachodzi:

$$Q_1 x_1 Q_2 x_2 \ldots Q_n x_n \psi'(x_1, x_2, \ldots, x_n) \ ,$$

gdzie ψ' jest uzyskane z ψ przez zamianę stałych c w wyrażeniach przez $\varphi(c)$ zaś Q_i jest kwantyfikatorem (uniwersalnym lub egzystencjalnym).

Dowód to indukcja po strukturze. Podstawa indukcji wynika z tego, że to homomorfim i nie ma negacji.

Dowód nieobowiązkowy. Pokazujemy, że dla każdego wyrażenia t arytmetycznego, tj. zbudowanego ze zmiennych, stałych oraz operacji dodawania i mnożenia, zachodzi

$$t'(\varphi(a_1), \varphi(a_2), \ldots, \varphi(a_n)) = \varphi(t(a_1, a_2, \ldots, a_n)) \ .$$

Dowód przebiega przez standardową indukcję po strukturze t:

stała jeśli $t = c \in A$ to $t' = \varphi(c) \in B$ i jest OK.

$+$ jeśli $t = t_1 + t_2$ to $t' = t'_1 + t'_2$ i z założenia indukcyjnego $t'_i(\varphi(a_1), \varphi(a_2), \ldots, \varphi(a_n)) = \varphi(t_i(a_1, a_2, \ldots, a_n))$. Wtedy

$$\begin{aligned}\varphi(t(a_1, a_2, \ldots, a_n)) &= \varphi(t_1(a_1, a_2, \ldots, a_n)) + \varphi(t_2(a_1, a_2, \ldots, a_n)) \\ &= t'_1(\varphi(a_1), \varphi(a_2), \ldots, \varphi(a_n)) + t'_2(\varphi(a_1), \varphi(a_2), \ldots, \varphi(a_n)) \\ &= t'(\varphi(a_1), \varphi(a_2), \ldots, \varphi(a_n)) \ .\end{aligned}$$

$\cdot$ Analogicznie jak dodawanie.

Przechodząc dla dowodu dla formuł. Przejście przez spójniki jest podobne jak powyżej. Dla kwantyfikatorów używamy definicji spełnialności formuł z kwantyfikatorami. Jedyne, co istotne, to że jeśli równość $t_1(a_1, \ldots, a_n) = t_2(a_1, \ldots, a_n)$ zachodzi w A to w B zachodzi $t'_1(\varphi(a_1), \ldots, \varphi(a_n)) = t'_2(\varphi(a_1), \ldots, \varphi(a_n))$. Zauważmy, że przy kwantyfikatorze uniwersalnym używamy tego, że homomorfizm jest „na”.

To daje równość dla kwantyfikatorów (sprawdzamy semantykę kwantyfikatorów) □

Wniosek 18.10. W $\mathbb{Z}_m$ zachodzą wszystkie prawa, o których myślimy.

18.3 Algorytm Euklidesa

Wracamy do naszego ulubionego ciała: $\mathbb{Z}_p$. Kiedyś już powiedzieliśmy, że jest tam element odwrotny. A co w $\mathbb{Z}_m$? Jest? Nie ma? Dla którego jest, czy można efektywnie wyznaczyć?

Konstrukcyjna metoda używała będzie *algorytmu Euklidesa*. Opiera się on na obserwacji, że $\mathrm{nwd}(a, b) = \mathrm{nwd}(a - b, b)$ oraz $\mathrm{nwd}(0, b) = \mathrm{nwd}(b, 0) = b$. Można to przyspieszyć, poprzez $\mathrm{nwd}(a, b) = \mathrm{nwd}(a \bmod b, b)$.

Definicja 18.11. Liczba $0 \neq k \in \mathbb{N}$ jest największym wspólnym dzielnikiem $a, b \in \mathbb{Z}$, jeśli $k|a$, $k|b$ i dla każdego ℓ zachodzi $\ell|a, \ell|b \implies \ell|k$.

Oznaczenie: $\mathrm{nwd}(a, b)$.

Uwaga. nwd jest największy w sensie porządku częściowego zdefiniowanego przez podzielność.

Lemat 18.12. *1. Jeśli $k|a$ i $k|b$ to $k|(a + b)$ i $k|(a - b)$.*

2. Jeśli $k|a$ i $k|b$ to $k|(a \bmod b)$.

3. Jeśli $k|(a \bmod b)$ i $k|b$ to $k|a$.

Dowód. Pierwsze: trywialne, reszta to zastosowanie pierwszego. □

Algorytm 2 Algorytm Euklidesa

Założenie: a, b są nieujemne, choć jedna jest dodatnia

while $a > 0$ oraz $b > 0$ **do**
    if $a < b$ **then**
        zamień a, b
    $a \leftarrow a - b$ ▷ Może też być $a \bmod b$
if $a \geq b$ **then**
    return a
else
    return b

Wniosek 18.13. Algorytm Euklidesa zwraca największy wspólny dzielnik.

Lemat 18.14. *Algorytm Euklidesa (w wersji z modulo) działa w czasie wielomianowym (od długości zapisu liczb). To ograniczenie jest ścisłe.*

Dowód pozostawiamy jako zadanie.

Lemat 18.15. *W czasie algorytmu Euklidesa możemy przechowywane liczby reprezentować jako kombinacje liniowe a oraz b.*

Dowód. Przez indukcję. □

To pozwala na

Lemat 18.16. *Dla $a, b \in \mathbb{Z}_+$ istnieją $x, y \in \mathbb{Z}$ takie że*

$$\mathrm{nwd}(a, b) = xa + yb.$$

Dokładnie jedna z tych liczb jest dodatnia i jedna niedodatnia. Dodatkowo, liczby te można wybrać tak, że $|x| < b$, $|y| < a$. Jeśli $\mathrm{nwd}(a, b) = 1$ to są dokładnie dwa takie wyrażenia (w jednym x jest dodatnie a w drugim ujemne).

Proty dowód pozostawiamy jako ćwiczenie.

Lemat 18.17. *W pierścieniu $\mathbb{Z}_m$ element a ma element odwrotny $\iff$ $\mathrm{nwd}(a, m) = 1$.*

Dowód. Niech $m' = \mathrm{nwd}(a, m) > 1$, załóżmy, że a ma element odwrotny b. Wtedy $ab = km + 1$ dla pewnego $k \geq 0$. Ale $m'|a$, czyli też $m'|(km + 1)$, a jako że $m'|m$ dostajemy, że $m'|1$, sprzeczność.

Jeśli $\mathrm{nwd}(a, m) = 1$ to istnieją $x, y \in \mathbb{Z}$, takie że $ax + my = 1$. Elementem odwrotnym do a jest x: $ax = 1 - my$ i tym samym $ax \equiv_m 1$. □

Uwaga. Zauważmy, że Lemat 18.17 w szczególności daje dowód Twierdzenia 18.4.

18.4 Elementy odwracalne

Definicja 18.18 (elementy odwracalne). Element a pierścienia R nazywamy *odwracalnym*, jeśli istnieje $b \in R$ takie że $ab = 1$.

Zbiór elementów odwracalnych pierścienia R oznaczamy jako R^*.

Twierdzenie 18.19. *Dla dowolnego pierścienia R z jednością zbiór elementów odwracalnych R^* jest grupą na mnożenie.*

Dowód. Trzeba sprawdzić, że R^* jest zamknięte na branie elementu odwrotnego oraz na mnożenie.

1 jest odwracalne.

Jeśli a jest odwracalne to a^{-1} też.

Jeśli a, b są odwracalne, to elementem odwrotnym do ab jest $b^{-1}a^{-1}$. □

Uwaga. $\mathbb{Z}_m^*$ nie ma struktury pierścienia, w szczególności nie jest ciałem!

Twierdzenie 18.20. *Dla ciała skończonego $\mathbb{F}$ grupa $\mathbb{F}^*$ jest cykliczna.*

To twierdzenie jest dość trudne, Rozdział 21 zawiera dowód w przypadku $\mathbb{F} = \mathbb{Z}_p$.

Definicja 18.21 (Symbol Eulera). $\varphi(m)$ to liczba liczb względnie pierwszych z m mniejszych od m.

Wniosek 18.22 (Twierdzenie Eulera). Niech a, m są względnie pierwsze. Wtedy

$$a^{\varphi(m)} = 1 \mod m$$

Dowód. $\mathbb{Z}_p^*$ jest grupą o $\varphi(m)$ elementach. Rząd elementu dzieli rząd grupy $\varphi(m)$. □

18.5 Chińskie twierdzenie o resztach

Definicja 18.23 (Produkt pierścieni.). Produkt pierścieni definiujemy standardowo: dla pierścienie R, R' ich produkt $R \times R'$ ma jako zbiór iloczyn kartezjański zbiorów R, R' a działania są po współrzędnych.

Lemat 18.24. *Proste własności:*

- *$R \times R$ i $R' \times R$ są izomorficzne*
- *produkt kartezjański jest łączny (z dokładnością do izomorfizmu): $R_1 \times (R_2 \times R_3)$ i $(R_1 \times R_2) \times R_3$ są izomorficzne*
- *Jeśli R_1 jest izomorficzne z R_1' a R_2 z R_2', to $R_1 \times R_2$ jest izomorficzne z $R_1' \times R_2'$.*

Twierdzenie 18.25 (Chińskie Twierdzenie o resztach). *Jeśli $m_1, m_2, \ldots, m_k$ są parami względnie pierwsze, to naturalny homomorfizm z $\mathbb{Z}_{m_1 m_2 \cdots m_k}$ w $\prod_{i=1}^{k} \mathbb{Z}_{m_i}$, gdzie na i-tej współrzędej bierzemy modulo $\mathbb{Z}_{m_i}$, jest izomorfizmem.*

Dowód. Zauważmy, że oba ziory są skończone i mają tą samą liczność, tak więc wystarczy pokazać, że przekształcenie jest „na" i to już da też, że jest różnowartościowe.

Wystarczy pokazać, że dla $m = m_1 \cdot m_2$, dla m_1, m_2 jak w sformułowaniu twierdzenia, potrafimy wskazać liczby n_1, n_2, takie że ich rzuty na $\mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2}$ dają $(1, 0)$ oraz $(0, 1)$. Wtedy dowolny element $(\alpha, \beta) \in \mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2}$ otrzymujemy jako rzut $\alpha n_1 + \beta n_2$. Dowód dla dowolnego iloczynu $m_1 m_2 \cdots m_k$ wynika z prostej indukcji.

Ponieważ $\text{nwd}(m_1, m_2) = 1$ to z Algorytmu Euklidesa dostajemy liczby x, y, x', y' takie że $xm_1 + ym_2 = x'm_1 + y'm_2 = 1$ oraz $x, y' > 0 \geq x', y$. Nasze liczby to $n_1 = y'm_2$ oraz $n_2 = xm_1$. Wtedy $n_1 \bmod m_2 = 0$, $n_1 \bmod m_1 = (1 - x'm_1) \bmod m_1 = 1$; analogicznie dla n_2. □

Uwaga. Reprezentacja dużych liczb przy użyciu Chińskiego Twierdzenia o resztach jest jedną z najpraktyczniejszych.

18.6 Zastosowanie: Algorytm szyfrowania Rabina

Dane: dwie duże liczby pierwsze p, q znane właścicielowi. Publicznie znane jest jedynie: $n = pq$.

Traktujemy komunikat do zaszyfrowania jako element z $\mathbb{Z}_n$, oznaczamy go jako c (jeśli c jest większe niż n, to dzielimy je na kawałki). Nadawca wiadomości przesyła komunikat $c^2 \bmod n$.

Chcemy pokazać, że:

1. odbiorca umie odtworzyć c z c^2
2. jeśli ktoś umie odtworzyć c z c^2 to umie rozłożyć n na p, q, co uznajemy na trudny problem

Trzeba zrozumieć najpierw, jak wygląda mnożenie w $\mathbb{Z}_n$. Z chińskiego twierdzenia o resztach $\mathbb{Z}_n \simeq \mathbb{Z}_p \times \mathbb{Z}_q$. To najpierw w $\mathbb{Z}_p$ i $\mathbb{Z}_q$.

Skorzystamy z silnego twierdzenia, które udowodnimy potem:

Twierdzenie 18.26. *Grupa $\mathbb{Z}_p^*$ jest cykliczna.*

Ile jest liczb, które są kwadratami oraz jakiej są postaci?

Lemat 18.27. *Jeśli p jest liczbą pierwszą, to $a^2 \equiv_p b^2$ wtedy i tylko wtedy, gdy $a \equiv_p b$ lub $a \equiv_p -b$.*

Dowód.

$$\begin{aligned} a^2 \equiv_p b^2 &\iff a^2 - b^2 \equiv_p 0 \\ &\iff (a-b)(a+b) \equiv_p 0 \\ &\iff a - b \equiv_p 0 \text{ lub } a + b \equiv_p 0 \\ &\iff a \equiv_p b \text{ lub } a \equiv_p -b \end{aligned}$$ □

W szczególności, w $\mathbb{Z}_p$ dla zadanego c^2 mamy dwa możliwe odszyfrowania: c i $-c$. W $\mathbb{Z}_{pq}$ mamy 4.

Wniosek 18.28. W $\mathbb{Z}_p^*$ jest $(p-1)/2$ kwadratów i $(p-1)/2$ nie-kwadratów. Są to odpowiednio parzyste i nieparzyste potęgi generatora.

Dowód. Jest $p-1$ elementów, dwa przeciwne przechodzą przez kwadrat na to samo i żadne inne, czyli $(p-1)/2$ jest kwadratami, czyli $(p-1)/2$ nie jest. Potęgi parzyste oczywiście są kwadratami, są różne i jest ich $(p-1)/2$. Czyli nieparzyste to nie-kwadraty. □

Wniosek 18.29. Jeśli g jest generatorem w $\mathbb{Z}_p^*$ to $g^{(p-1)/2} = -1$.

Dowód. Wiemy, że $(g^{(p-1)/2})^2 = g^{p-1} = 1$. Z drugiej strony, są najwyżej dwie liczby x takie że $x^2 = 1$; łatwo sprawdzić, że są to -1 oraz 1. Jako że g jest generatorem, to nie może być, że $g^{(p-1)/2} = 1$ (bo wtedy g nie jest generatorem), czyli $g^{(p-1)/2} = -1$. □

Lemat 18.30. *Jeśli p jest pierwsza to w $\mathbb{Z}_p^*$ jeśli a jest kwadratem, to $a^{(p-1)/2} = 1$, w przeciwnym przypadku $a^{(p-1)/2} = -1$.*

Dowód. **kwadrat** Wtedy $a = g^{2k}$ i $a^{(p-1)/2} = g^{(p-1)k} = (g^{p-1})^k = 1^k = 1$.

nie-kwadrat Wtedy $a = g^\ell$ dla nieparzystego ℓ. Mamy $(g^\ell)^{(p-1)/2} = (g^{(p-1)/2})^\ell = (-1)^\ell = -1$. □

18.6.1 Odtwarzanie

Dla ułatwienia obliczeń, zakładamy, że $p = 3 \bmod 4$, czyli $p = 4k + 3$.

Lemat 18.31. *Dla jednej z liczb $c, -c$ w $\mathbb{Z}_p$ mamy $c^{(p-1)/2} = 1$ lub $(-c)^{(p-1)/2} = 1$*

Dowód. Niech g będzie generatorem. Wtedy $-1 = g^{(p-1)/2} = g^{2k+1}$, czyli jest nieparzystą potęgą generatora. Czyli dokładnie jedna z c, $-c$ jest nieparzystą potęgą generatora, a jedna parzystą. Dla tej parzystej zachodzi teza. □

Bez zmniejszenia ogólności w dalszej części zakładamy, że $c^{(p-1)/2} = 1$.

Z Lematu 18.31 mamy, że $c^{2k+2} = c^{(p-1)/2}c = c$. Czyli wystarczy podnieść komunikat c^2 do potęgi $k+1$ i dostajemy c. W drugim przypadku, gdy $c^{(p-1)/2} = -1$, otrzymujemy $-c$.

Robimy tak dla p, q i tym samym dla n.

18.6.2 Odtwarzanie implikuje rozkład liczby na czynniki

Zakładamy, że algorytm deszyfrujący jest deterministyczny, tj. dla zadanego m zwróci zawsze ten sam wynik (czyli komunikat c, taki że $c^2 = m$).

Algorytm 3 Algorytm faktoryzujący n używający dekodowania szyfru Rabina

Założenie: $n = pq$, n znane, p, q: liczby pierwsze

1: wylosuj (jednostajnie) $x \in \mathbb{Z}_n^*$
2: oblicz x^2
3: zdekoduj c z x^2 mod n
4: **if** $c = x$ lub $c = n - x$ **then**
5:     wróć do kroku 1
6: wyznacz $p = \mathrm{nwd}((c + x)/2, n)$
7: **return** $(p, n/p)$

Lemat 18.32. *Z prawdopodobieństwem 1/2 otrzymujemy dzielnik n*

Dowód. Zauważmy że x odpowiada parze (x_p, x_q) w $\mathbb{Z}_p \times \mathbb{Z}_q$. Możliwe zdekodowane wiadomości z (x_p^2, x_q^2) to (x_p, x_q), $(x_p, -x_q)$, $(-x_p, x_q)$ i $(-x_p, -x_q)$ i każda z nich jest równie prawdopodobna (bo dekoder jest deterministyczny a my losowaliśmy, czyli z równą szansą trafiliśmy na każdą z tych czwórek).

Jeśli dekoder zwróci (x_p, x_q) lub $(-x_p, -x_q)$, czyli x lub $-x$, to nic nie mamy. Ale jeśli jedną z pozostałych par, np. $(-x_p, x_q)$, to $((x_p, x_q) + (-x_p, x_q))/2 = (0, x_q)$, czyli liczbę podzielną przez p. Licząc nwd z $n = pq$ dostajemy p. □

Argument, gdy dekoder nie jest deterministyczny, lecz losowy, wygląda analogicznie.

Rozdział 19

Wielomiany

19.1 Pierścień wielomianów

Definicja 19.1 (Wielomian)**.** *Wielomian f to ciąg $(a_0, a_1, \ldots, a_n)$*, myślimy o nich jako o $\sum a_i x^i$. Zwykle zakładamy, że $a_n \neq 0$, w przeciwnym razie dla $n > 0$ utożsamiamy $a_0, \ldots, a_n$ z $a_0, \ldots, a_{n-1}$.

Mnożenie wielomianów definiujemy tak jak się spodziewamy, tzn. dla wielomianów $(a_0, \ldots, a_n)$ oraz $(b_0, \ldots, b_m)$ ich iloczyn to $(c_0, \ldots, c_{n+m})$, gdzie:

$$c_k = \sum_{i=0}^{k} a_i b_{k-i} \quad .$$

Zbiór wielomianów o współczynnikach z pierścienia R oraz naturalnym dodawaniem i mnożeniem (tj. po współrzędnych) to *pierścień wielomianów $R[x]$*. *Zerem* w tym pierścieniu jest wielomian (0).

Liczby $a_0, \ldots, a_n$ to *współczynniki wielomianu*, jeśli $a_n \neq 0$ to jest on *współczynnikiem wiodącym.*

Stopień wielomianu $\deg(a_0, \ldots, a_n)$ dla $a_n \neq 0$ to n. W przypadku wielomiany zerowego stopień to $-\infty$.

Zauważmy, że jest dobrze określone nawet dla pierścienia nieprzemiennego. Jeśli myślimy o wielomianach jak o ciągach, to tę operację nazywamy *splotem* dwóch ciągów (i często oznaczamy przez $*$).

Możemy też myśleć że wielomiany to ciągi nieskończone, które mają tylko skończenie wiele niezerowych wyrazów. Wynik mnożenia z dodanymi wiodącymi zerami jest taki sam.

Lemat 19.2 (Poprawność definicji)**.** *$R[x]$ z mnożeniem zdefiniowanym jako splot jest pierścieniem.*

Jeśli R jest pierścieniem przemiennym (z jednością), to $R[x]$ też jest pierścieniem przemiennym (z jednością).

Zwykle zajmujemy się wielomianami o współczynnikach z ciała.

Lemat 19.3. *Niech $f, g \in R[x]$. Wtedy*

$$\begin{aligned} \deg(f+g) &\leq \max(\deg(f), \deg(g)) \\ \deg(f \cdot g) &\leq \deg(f) + \deg(g) \quad . \end{aligned}$$

Jeśli R jest ciałem, to

$$\deg(f \cdot g) = \deg(f) + \deg(g) \quad .$$

Uwaga. W ostatnim punkcie założenie, że R jest ciałem jest istotne: np. w $\mathbb{Z}_6$ mamy $2 \cdot 3 = 0$ i iloczyn tych dwóch wielomianów stopnia 0 ma stopień $-\infty$.

Dowód. Niech $f = (f_0, \ldots, f_n)$, $g = (0_m, \ldots, 0_)$, gdzie $\deg(f) = n, \deg(g) = m$.

Wtedy $f+g$ ma same współczynniki 0 powyżej pozycji $\max(n, m)$, czyli $\deg(f+g) \leq \max(\deg(f), \deg(g)$.

W $f \cdot g$ zgodnie z definicją splotu dla $k > m + n$ w każdym iloczynie przynajmniej jeden współczynnik jest zerowy.

Jeśli R jest ciałem, to współczynnik przy x^{n+m} wynosi $f_n \cdot g_m$, przy czym $f_n \neq 0 \neq g_m$. Skoro R jest ciałem, to w takim razie $f_n \cdot g_m$ też jest niezerowe i tym samym $\deg(f \cdot g) = \deg(g) + \deg(f)$. $\square$

19.2 Ewaluacja (wartościowanie) wielomianów

Wielomian $f \in R[x]$ równy $(a_0, \ldots, a_n)$ możemy też potraktować jako funkcję z R w R, zdefiniowaną w naturalny sposób:

$$\overline{f}(p) = \sum_{k=0}^{n} a_k p^k$$

Uwaga 19.4. Różne wielomiany niekoniecznie definiują różne funkcje!

W skończonych ciałach to nie jest tak ogólnie możliwe; w nieskończonych tak jest.

Później omówimy to dokładniej.

Lemat 19.5. *Niech $f, g \in R[x]$ i $p \in R$. Wtedy*

$$\overline{f+g}(p) = \overline{f}(p) + \overline{g}(p)$$

Jeśli R jest przemienny, to dodatkowo

$$\overline{f \cdot g}(p) = \overline{f}(p) \cdot \overline{g}(p)$$

Dowód. Niech $f = (f_0, \ldots, f_m), g = (g_0, \ldots, g_m)$, jeżeli jeden ma mniejszą ilość współczynników, niż drugi, to uzupełniamy zerami. Wtedy dla sumy mamy

$$\overline{f+g}(p) = \sum_{i=0}^{n}(f_i + g_i)p^i = \sum_{i=0}^{n} f_i p^i + \sum_{i=0}^{n} g_i p^i = \overline{f}(p) + \overline{g}(p) \quad .$$

Dla iloczynu

$$\begin{aligned}
\overline{f \cdot g}(p) &= \sum_{i=0}^{2m} \sum_{k=0}^{i} f_k g_{i-k} p^k \\
&= \sum_{i=0}^{2m} \sum_{k=0}^{i} f_k p^k g_{k-i} p^{i-k} \\
&= \left(\sum_{k=0}^{m} f_k p^k\right)\left(\sum_{i=0}^{m} g_i p^i\right) \\
&= \overline{f}(p)\overline{g}(p)
\end{aligned}$$

□

19.3 Dzielenie, podzielność i największy wspólny dzielnik wielomianów

Patrzymy na $\mathbb{F}[x]$. Jest podzielność, podobnie jak dla liczb całkowitych.

Lemat 19.6 (Dzielenie wielomianów)**.** *Niech $\mathbb{F}$ będzie ciałem a $\mathbb{F}[x]$ pierścieniem wielomianów o współczynnikach z $\mathbb{F}$. Dla wielomianów f, g z tego pierścienia, o stopniach $m = \deg(f)$ oraz $n = \deg(g) \neq -\infty$ istnieje dokładnie jedna para wielomianów q, r, taka że $f = gq + r$, gdzie $\deg(r) < \deg(g)$.* [1] *Wielomiany te można efektywnie wyliczyć.*

Wielomiany q, r z Lematu 19.6 nazywamy *ilorazem* oraz *resztą z dzielenia* f przez g.

Dowód. Przez indukcję po stopniu f.

Jeśli $\deg(f) < \deg(g)$, to bierzemy $q = 0$ oraz $r = f$.

Jeśli $\deg(f) \geq \deg(g)$, to bierzemy odpowiednią potęgę: niech wiodący współczynnik g to g_m zaś wiodący współczynnik f to f_n. Wtedy $f - (f_n g_m^{-1})x^{n-m}g$ ma mniejszy stopień (tu korzystamy z tego, że współczynniki są z ciała i element g_m^{-1} istnieje) i z założenia indukcyjnego ma reprezentację

$$f - (f_n g_m^{-1})x^{n-m}g = qg + r \quad .$$

Wtedy

$$f = (q + (f_n g_m^{-1})x^{n-m})g + r \quad .$$

Łatwo sprawdzić, że $q + (f_n g_m^{-1})x^{n-m}$ spełnia warunki.

To jest de facto algorytm dzielenia.

Jedyność: jeśli są dwie reprezentacje, to je odejmujemy i dostajemy nietrywialną reprezentację wielomianu 0, sprzeczność. □

[1] Dla $\deg(g) = 0$ korzystamy z tego, że $\deg(0)$ to $-\infty$

Przykład 19.7. Podzielmy wielomiany $f = x^5 - 3x^4 - x^3 + 7x^2 - 4$ oraz $g = x^3 - 3x^2 + 2x$ z $\mathbb{R}[x]$ z resztą:

$$\begin{array}{r|l}
 & \phantom{x^5 - 3x^4 - {}} x^2 \qquad\qquad\quad - 3 \\ \hline
x^3 - 3x^2 + 2x) & x^5 - 3x^4 - x^3 + 7x^2 \qquad - 4 \\
 & -x^5 + 3x^4 - 2x^3 \\ \cline{2-2}
 & \qquad\quad -3x^3 + 7x^2 \\
 & \qquad\quad 3x^3 - 9x^2 + 6x \\ \cline{2-2}
 & \qquad\qquad\quad -2x^2 + 6x - 4
\end{array}$$

Czyli $x^5 - 3x^4 - x^3 + 7x^2 - 4 = (x^3 - 3x^2 + 2x)(x^2 - 3) + (-2x^2 + 6x - 4)$.

Definicja 19.8 (Podzielność wielomianów)**.** Wielomian f jest *podzielny przez wielomian* g, jeśli reszta dzielenia f przez g wynosi 0. Zapisujemy to jako $f|g$.

Fakt 19.9. $f|g \iff$ *istnieje wielomian* q *taki że* $g = fq$.

Lemat 19.10. *Każdy wielomian dzieli* 0.

Jeśli f *dzieli* $g \neq 0$, *to* $0 \leq \deg(f) \leq \deg(g)$.

Jeśli f *dzieli* g *i* g *ma stopień* 0, *to* f *też ma stopień* 0.

Jeśli f *dzieli* g *i* g *dzieli* f, *to* $\frac{f}{g}$ *jest stałą.*

Dowód. Oczywiście $f \cdot 0 = 0$.

Skoro $f|g$ to $g = fg'$ i $g', f \neq 0$ (bo $g \neq 0$). W takim razie $\deg(g) = \deg(f) + \deg(g') \geq \deg(f)$.

Skoro $f|g$ to $\deg(f) \leq \deg(g) = 0$. Przy czym $f = 0$ nie jest możliwe, bo wtedy $f|g$ implikuje $g = 0$, co nie jest prawdą (bo $\deg(g) = 0 \neq \deg(0)$.

Skoro $f|g$ i $g|f$ to $f = gf'$ oraz $g = fg'$. Czyli $f = f'g'f$. W takim razie $f'g' = 1$ i tym samym f', g' są stałymi. □

Definicja 19.11 (Wielomian nierozkładalny)**.** Wielomian $f \in R[x]$ jest *nierozkładalny* w $R[x]$, jeśli $\deg(f) > 0$ i nie istnieją wielomiany $g, h \in R[x]$ takie że $f = gh$ oraz $\deg(g), \deg(h) < \deg(f)$.

Wielomiany stopnia 1 są nierozkładalne. Ale mogą być też większego stopnia: np. wielomian $x^2 + 1$ w $\mathbb{R}[x]$

Definicja 19.12 (Największy wspólny dzielnik (nwd) wielomianów)**.** Największy wspólny dzielnik dwóch wielomianów f, g to taki wielomian h, że $h|f$, $h|g$ oraz jeśli $h'|f$, $h'|g$, to $h'|h$.

Zauważmy, że nwd wielomianów jest określone z dokładnością do stałej multiplikatywnej.

Liczymy to przy użyciu algorytmu Euklidesa. (Cały algorytm i dowód jego poprawności działa dokładnie tak jak w przypadku liczb całkowitych).

Lemat 19.13. *Każde dwa wielomiany* p, q *mają największy wspólny dzielnik. Jest on postaci* $ap + bq$ *dla pewnych wielomianów* a, b.

Dowód. Chcemy pokazać, że jeśli $f = qg + r$ to $\mathrm{nwd}(f, g) = \mathrm{nwd}(r, g)$ (i że oba istnieją). W tym celu pokazujemy, że

$$p|g \wedge p|f \iff p|g \wedge p|r$$

Obie implikacje to proste rachunki. Postępując w ten sposób dochodzimy do $\mathrm{nwd}(0, p)$, bo w każdym kroku suma stopni spada. który oczywiście istnieje i jest równy p.

Analogicznie jak dla liczb pokazujemy też, że para trzymanych wielomianów jest kombinacją wejściowych (współczynniki też są wielomianami). □

Przykład 19.14. Znajdźmy największy wspólny dzielnik wspomnianych już wielomianów $f = x^5 - 3x^4 - x^3 + 7x^2 - 4$ oraz $g = x^3 - 3x^2 + 2x$ z $\mathbb{R}[x]$ przy użyciu algorytmu Euklidesa. Pierwszy krok to jak poprzednio podzielenie tych wielomianów z resztą.

$$
\begin{array}{r l}
 & \overline{\quad\quad\quad\quad\quad\quad x^2 \quad\quad - 3} \\
x^3 - 3x^2 + 2x) & x^5 - 3x^4 - x^3 + 7x^2 \quad\quad - 4 \\
 & \underline{-x^5 + 3x^4 - 2x^3} \\
 & \quad\quad - 3x^3 + 7x^2 \\
 & \quad\quad \underline{3x^3 - 9x^2 + 6x} \\
 & \quad\quad\quad - 2x^2 + 6x - 4
\end{array}
$$

Czyli $x^5 - 3x^4 - x^3 + 7x^2 - 4 = (x^3 - 3x^2 + 2x)(x^2 - 3) + (-2x^2 + 6x - 4)$.

Dalej korzystamy z:

$$\mathrm{nwd}(af + b, f) = \mathrm{nwd}(b, f).$$

Tym samym pozostaje nam policzenie $\mathrm{nwd}(-2x^2 + 6x - 4, x^3 - 3x^2 + 2x)$.

$$
\begin{array}{r l}
 & \overline{\quad\quad\quad\quad -\tfrac{1}{2}x} \\
-2x^2 + 6x - 4) & x^3 - 3x^2 + 2x \\
 & \underline{-x^3 + 3x^2 - 2x} \\
 & \quad\quad\quad\quad 0
\end{array}
$$

Tj., $x^3-3x^2+2x = (-\frac{1}{2}x)(-2x^2+6x-4)$ i w takim razie $\mathrm{nwd}(-2x^2+6x-4, x^3-3x^2+2x)$ to $-2x^2+6x-4$.

Tym samym poszukiwany największy wspólny dzielnik f oraz g to

$$-2x^2 + 6x - 4 = 1 \cdot (x^5 - 3x^4 - x^3 + 7x^2 - 4) + (-x^2 + 3) \cdot (-2x^2 + 6x - 4).$$

Wyrażenie go przez f, g jest proste:

$$-2x^2 + 6x - 4 = x^5 - 3x^4 - x^3 + 7x^2 - 4 - (x^3 - 3x^2 + 2x)(x^2 - 3) \ .$$

Lemat 19.15. *Jeśli f jest nierozkładalny oraz $f|p_1p_2\dots p_k$ to $f|p_i$ dla pewnego i.*

Dowód. Dla dwóch, a potem przez indukcję.

$\mathrm{nwd}(f, p_2)|f$, czyli z dokładnością do przemnożenia przez stałą to jest f lub 1. Jeśli f to $f|p_2$ i ok, w przeciwnym razie

$$af + bp_2 = 1$$

Mnożymy przez p_1, dostajemy

$$afp_1 + bp_1p_2 = p_1 \ .$$

f dzieli lewą stronę, czyli też prawą. □

Lemat 19.16. *Jeśli f_i są nierozkładalne oraz $\mathrm{nwd}(f_i, f_j)$ jest stałą dla $i \neq j$ oraz $f_i|g$ to $f_1 \dots f_k|g$.*

Dowód. Przez indukcję.

Załóżmy, że $f_1 \cdots f_i|g$, czyli $g = f_1 \cdots f_i g'$. Czyli $f_{i+1}|f_1 \cdots f_i g'$. Czyli dzieli jeden z nich. Nie jest to żaden z f_j. Czyli g'. □

Twierdzenie 19.17 (Bézout)**.** *Jeśli $\mathbb{F}$ jest ciałem, $\mathbb{F}[x]$ pierścieniem wielomianów o współczynnikach z tego ciała zaś $f, (x - c) \in \mathbb{F}[x]$ wielomianami z tego pierścienia, to reszta z dzielenia f przez $(x - c)$ to $\overline{f}(c)$.*

W szczególności $(x - c)|f$ wtedy i tylko wtedy, gdy $\overline{f}(c) = 0$.

Dowód. Niech $f = q(x - c) + r$, gdzie $\deg(r) < \deg(x - c) = 1$, tj. r jest stałą. Obliczmy wartościowanie lewej i prawej strony w punkcie c:

$$
\begin{aligned}
\overline{f}(c) &= (\overline{q(x - c)} + r)(c) \\
&= \overline{q}(c)\overline{(x - c)}(c) + r \\
&= r \ ,
\end{aligned}
$$

co daje tezę. □

Definicja 19.18 (Pierwiastek, rozwiązanie wielomianu)**.** c nazywamy pierwiastkiem (rozwiązaniem) wielomianu f, gdy $(x - c)|f$; c jest pierwiastkiem k-krotnym, dla $k \geq 1$, gdy $(x - c)^k|f$.

Wniosek 19.19. c jest pierwiastkiem f wtedy i tylko wtedy gdy $\overline{f}(c) = 0$.

Twierdzenie 19.20. *Wielomian* $0 \neq f \in \mathbb{F}[x]$ *ma najwyżej* $\deg(f)$ *różnych pierwiastków.*

Dowód. Załóżmy, że ma $k > n$ różnych pierwiastków $p_1, \dots, p_k$. Wtedy jest podzielny przez każdy z wielomianów $(x - p_i)$. Ponieważ są to wielomiany nierozkładalne, to z Lematu19.16, f jest też podzielny przez $\prod_{i=1}^{k}(x - p_i)$. Stopień tego wielomianu jest większy niż stopień f, sprzeczność. □

Wniosek 19.21. Jeśli waluacje dwóch wielomianów stopnia conajwyżej n mają te same wartości w $n + 1$ punktach, to są równe.

W ciele nieskończonym dwa wielomiany mają skończoną liczbę wartości wspólnych.

Przykład/Zastosowanie 19.22 (Interpolacja wielomianu). Jeśli dla danego wielomianu $f \in \mathbb{F}[x]$ stopnia n mamy podane jego wartości $\overline{f}(p_i)$ dla różnych $p_0, \dots, p_n \in \mathbb{F}$, to jest on jednoznacznie wyznaczony.

Obliczenia wielomianu można dokonać przy użyciu macierzy Vandermonde'a: niech współczynniki wielomiany f to $f_0, \dots, f_n$. Wtedy

$$\begin{bmatrix} p_0 & p_0^1 & \cdots & p_0^n \\ p_1 & p_1^1 & \cdots & p_1^n \\ \vdots & \vdots & \ddots & \vdots \\ p_n & p_n^1 & \cdots & p_n^n \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_n \end{bmatrix} = \begin{bmatrix} f(p_0) \\ f(p_1) \\ \vdots \\ f(p_n) \end{bmatrix}$$

Macierz Vandermonde'a jest odwracalna, czyli układ ten można rozwiązać. Dla niektórych wyborów punktów można to zrobić szybko (tzw. szybka transformata Fouriera).

Wielomian ten można też podać bardziej „wprost”: powiedzmy, że podamy wielomiany $w_0, \dots, w_n$, takie że $\overline{w_i}(p_i) = 1$ oraz $w_i(p_j) = 0$ dla $j \neq i$ oraz $\deg(w_i) \leq n$. Wtedy $f = \sum_i \overline{f}(p_i) w_i$: zauważmy, że waluacja sumy po prawej stronie w każdym z punktów p_i to $\overline{f}(p_i)$ (bo $\overline{w_j}(p_i) = 1$ dla $i = j$ oraz 0 dla $i \neq j$). Zdefiniujmy dodatkowo wielomian

$$w = \prod_{j=0}^{n} (x - p_j) \quad .$$

Łatwo wtedy wyrazić w_i: niech $v_i = \frac{w_i}{x - p_i}$. Wtedy

$$w_i(x) = \frac{v_i}{\overline{v_i}(p_i)} \quad .$$

Wielomian ten zany jest jako *wielomian interpolacyjny Legrange'a.*

Przykład/Zastosowanie 19.23 (Dzielenie sekretu). Dla grupy n osób chcemy stworzyć protokół, który pozwala dowolnym $m + 1$ z nich poznać wiadomość, ale każdym m już nie.

Niech nasza wiadomość to c_0. Losujemy liczby $c_1, \dots, c_m$ i tworzymy wielomian $c = \sum_{i=0}^{m} c_i x^i$. Wyznaczamy teraz n różnych niezerowych punktów $p_1, \dots, p_n$, osoba i otrzymuje jako wiadomość wartość $c(p_i)$ oraz wartość punktu p_i.

Dzięki interpolacji $m + 1$ osób jest w stanie odtworzyć ten wielomian. Natomiast dla dowolnych m osób możemy dorzucić dowolną wartość w punkcie 0 (czyli dokładnie nasze c_0) i one wciąż są w stanie zinterpolować ten wielomian, do dowolnej wiadomości. Innymi słowy: dowodliwie nic nie wiedzą (każdy sekret jest możliwy i równie prawdopodobny).

Rozdział 20

Ciała skończone

Definicja 20.1 (Charakterystyka ciała; ciało proste)**.** Dla ciała $\mathbb{F}$ jego *charakterystyka* to rząd 1 w grupie addytywnej.

Ciało generowane przez 1 w ciele $\mathbb{F}$ to *ciało proste*.

Lemat 20.2. *Rząd ciała to albo $+\infty$ albo liczba pierwsza p. W pierwszym przypadku ciało proste to $\mathbb{Q}$, w drugim: $\mathbb{Z}_p$.*

Dowód. Dodajemy do siebie 1. Jeśli nigdy nie uzyskamy 0, to dostajemy kopię liczb naturalnych. W ciele istnieją elementy przeciwne, czyli mamy kopię liczb całkowitych. W ciele istnieją elementy odwrotne, czyli mamy liczby postaci $\{\frac{1}{n} : n \in \mathbb{Z} \setminus \{0\}\}$. Ciało jest zamknięte na mnożenie, czyli mamy wszystkie liczby postaci $\{\frac{p}{q} : p, q \in \mathbb{Z}, q \neq 0\} = \mathbb{Q}$. (Formalnie trzeba jeszcze pokazać, że operacje tam działają tak jak dla $\mathbb{Q}$, ale tak jest, bo one są wszystkie generowane przez 1.)

Jeśli po m dodaniach dostaliśmy 0, to m musi być pierwsze: w przeciwnym razie $m = m'm''$ i mamy równość $m'm'' = 0$ i żadne z nich nie jest 0.

Skoro dodane do siebie p razy 1 daje 0, to mamy $\mathbb{Z}_p$ (ponownie, powinniśmy pokazać, że operacje działają tak samo). □

Lemat 20.3. *Ciało jest przestrzenią liniową nad swoim ciałem prostym.*

Wniosek 20.4. Każde ciało skończone ma p^k elementów dla pewnych p—pierwsze i $k \in \mathbb{N} \setminus \{0\}$.

20.1 Konstrukcja ciał (skończonych)

Naszym celem obecnie jest konstrukcja ciała skończonego. Takie ciało uzyskamy przez wydzielenie pierścienia $\mathbb{F}[x]$ przez odpowiednią kongruencję. Jest to analogiczna konstrukcja do konstrukcji $\mathbb{Z}_p$ jako wydzielenia $\mathbb{Z}$ przez kongruencję podzielności przez liczbę pierwszą. Naszym ciałem zwykle jest ciało skończone (np. $\mathbb{Z}_p$), ale wszystko działa też dla ciał o charakterystyce $+\infty$.

Definicja 20.5 (Kongruencja w pierścieniu)**.** Relacja $\equiv \subseteq R^2$ jest *kongruencją* w pierścieniu R, jeśli

- jest relacją równoważności
- jest kongruencją w grupie $(R, +)$ oraz kongruencją w półgrupie $(R, \cdot)$.

Definicja 20.6 (Kongruencja modulo wielomian)**.** Dla ciała $\mathbb{F}$ oraz pierścienia wielomianów $\mathbb{F}[x]$ o współczynnikach z tego ciała oraz wielomianu $h \in \mathbb{F}[x]$ definiujemy kongruencję $\equiv_h$ na $\mathbb{F}[x]$:

$$f \equiv_h g \iff h|(f - g).$$

Lemat 20.7. *Dla ciała $\mathbb{F}$ oraz pierścienia wielomianów $\mathbb{F}[x]$ o współczynnikach z tego ciała oraz wielomianu $h \in \mathbb{F}[x]$ relacja $\equiv_h$ jest kongruencją na pierścieniu.*

Łatwo sprawdzić, że jest to relacja równoważności oraz że operacje dodawania oraz mnożenia są dobrze zdefiniowane (tj. nie zależą od wyboru reprezentanta). Ponadto uzyskany pierścień jest pierścieniem przemiennym z jednością.

Fakt 20.8. *Operacje* $+, \cdot$ *są dobrze zdefiniowane w* $\mathbb{F}[x]/\equiv_h$.
$\mathbb{F}[x]/\equiv_h$ *jest pierścieniem przemiennym z jednością.*

Lemat 20.9. *Jeśli wielomian* $h \in \mathbb{F}[x]$ *jest nierozkładalny, to w* $\mathbb{F}[x]/\equiv_h$ *istnieje element odwrotny dla* $f \not\equiv_h 0$.

Dowód. Weźmy $\mathrm{nwd}(f, h)$. Wtedy $af + bh = 1$. Wielomian a jest odwrotny do f w $\mathbb{F}[x]/\equiv_h$. □

Twierdzenie 20.10. *Jeśli wielomian* h *jest nierozkładalny, to ciało* $\mathbb{F}[x]/\equiv_h$ *(jako przestrzeń liniowa nad* $\mathbb{F}$*) ma wymiar* $\deg(h)$. *Jeśli* $\mathbb{F}$ *jest skończone, to takie rozszerzenie ma* $|\mathbb{F}|^{\deg h}$ *elementów.*

Dowód. Wielomiany $1, x, x^2, \ldots, x^{\deg(f)-1}$ są liniowo niezależne i są bazą tej przestrzeni. □

Twierdzenie 20.11 (bez dowodu). *Dwa ciała skończone o* p^k *elementach są izomorficzne.*

Przykład 20.12. Wielomian nierozkładalny $x^2 + 1 \in \mathbb{R}[x]$; wychodzi izomorficzne z $\mathbb{C}$.

Przykład 20.13. Zbudujmy ciało 4-elementowe. $4 = 2^2$, więc bierzemy $\mathbb{F} = \mathbb{Z}_2$ i potrzebujemy wielomianu nierozkładalnego stopnia 2. Jedynym takim wielomianem (w tym wypadku) jest $x^2 + x + 1$. Elementami ciała będą $0, 1, x, x + 1$ (albo ich klasy abstrakcji ze względu na $\equiv_{x^2+x+1}$). Działania są naturalne. Jedyne nietrywialne: mnożenie $x \cdot x$. Ale w tym wypadku mamy $x^2 \equiv x + 1$ (dokładniej, to $x^2 \equiv -(x + 1)$, ale $-(x + 1) = x + 1$ w $\mathbb{Z}_2[x]$).

Lemat 20.14. *W* $\mathbb{Z}_p[x]$ *jest wielomian nierozkładalny dowolnego stopnia większego niż* 0.

Dowód polega na podaniu konkretnego wielomianu lub na zliczaniu wielomianów rozkładalnych i nierozkładalnych. Szczegółów nie podamy.

Przykład/Zastosowanie 20.15 (Kody Reeda-Solomona). Ustalamy ciało $\mathbb{F}$, zwykle jest to cialo $\mathbb{F} = \mathbb{F}_{2^m}$. Kodujemy wiadomość $(a_0, a_1, \ldots, a_{k-1})$, gdzie $a_i \in \mathbb{F}$ jako wielomian

$$\sum_{i=0}^{k-1} a_i x_i \in \mathbb{F}[x]$$

Przekazujemy tę wiadomość jako wartości $\bar{f}$ w n różnych niezerowych punktach $p_0, p_1, \ldots, p_{n-1} \in \mathbb{F}$, gdzie $n \geq k$. Punkty mogą być wybrane dowolnie, ale zwykle ten wybór jest ustalony, bo dla pewnych wartości (pierwiastki z 1) łatwiej się liczy.

Zbiór możliwych wiadomości oznaczamy przez RS a jej elementy nazywamy *słowami kodowymi.*

Jeśli $n = k$ to nic nie zyskujemy. Jeśli więcej, to jest pewna nadmiarowość.

Kody liniowe

Kody Reeda Salomona są szczególnym przypadkiem *kodów liniowych*, w których kodowane słowo traktowane jest element $\mathbb{F}^k$ a kodowanie to mnożenie przez ustaloną macierz K rozmiaru $n \times k$ (w naszym przypadku: macierz a'la Vandermonde), gdzie $n \geq k$; nie jest to jedyne możliwe kodowanie. W szczególności obraz (tj. słowo kodowe) jest z przestrzeni $\mathbb{F}^n$.

Odległość

Odległością (Hamminga) jest dla nas ilość pozycji, na których różnią się dwa komunikaty. Oznaczenie: $d(c, c')$. To jest odległość.

Odległość kodu C to

$$d(C) = \min_{u,v \in C, u \neq v} d(u, v) \quad .$$

Lemat 20.16. *Odległość kodu Reeda-Solomona wynosi przynajmniej* $n - k + 1$.

Dowód. Dwa różne wielomiany stopnia $< k$ mają najwyżej $k - 1$ wartości wspólnych. Czyli dwa różne słowa kodowe mają nie więcej niż $k - 1$ wartości wspólnych, czyli przynajmniej $n - k + 1$ różnych. □

Optymalność odległości (ograniczenie Singletona)

Pokażemy teraz, że kody Reeda-Salomona są optymalne, tzn. jeśli kodujemy (dowolnym kodem) k-krotki elementów przy użyciu n-krotek, to któreś dwa mają odległość $\leq n-k+1$.

Twierdzenie 20.17 (Ograniczenie Singletona). *Jeśli w zbiorze $|\mathbb{F}|^n$ wybierzemy $|\mathbb{F}|^k$ wektorów, to któreś dwa mają odległość najwyżej $n-k+1$.*

Dowód. Podzielmy całe $\mathbb{F}^n$ na $|\mathbb{F}|^k$ „stożków”: jeden stożek to zbiór elementów o ustalonych k pierwszych współrzędnych. Jeśli któryś stożek zawiera dwa słowa kodowe, to różnią się one na najwyżej $n-k$ pozycjach (bo tyle ich w sumie jest). Czyli rozpatrujemy przypadek, że w każdym stożku jest dokładnie jedno słowo kodowe. Rozpatrzmy dwa stożki różniące się w k pierwszych pozycjach na jednym miejscu, mają one $k-1$ pozycji wspólnych, czyli najwyżej $n-k+1$ różnych. □

Poprawianie błędów

Naturalne poprawianie: poprawiamy otrzymane słowo do najbliższego słowa kodowego.

Twierdzenie 20.18. *Naturalne poprawianie poprawnie dekoduje słowo, jeśli ma ono mniej niż*

$$\frac{d(C)}{2}$$

błędów, czyli najwyżej

$$\left\lfloor \frac{d(C)-1}{2} \right\rfloor$$

błędów.

Dowód. Ponieważ słowa kodowe są odległe o najwyżej $d(C)$, to umiemy poprawić $\lfloor \frac{d(C)-1}{2} \rfloor$ błędów. □

Algorytm Berlekamp–Welch poprawiania błędów

Cel: dane $\vec{w} = [w_0, \ldots, w_{n-1}] \in \mathbb{F}^n$.
szukane: wielomian $f \in \mathbb{F}[x]$, t. że $\deg(f) < k$, $f(\alpha_i) \neq w_i$ dla najwyżej $e \leq \left\lfloor \frac{n-k}{2} \right\rfloor$ albo „?” jak nie ma takiego wielomianu.

Oznaczenie: $I = \{i : f(\alpha_i) \neq w_i\}$. Dobrze zdefiniowane, bo takie f jest jedyne. Jak nie ma wyniku, to trudno. Niech $e = |I|$.

Moglibyśmy po prostu wybrać te błędy, zinterpolować i rozwiązać...

Popatrzmy na wielomian

Definicja 20.19 (Error locator polynomial). Dla zbioru pozycji błędów I zdefiniujmy *error locator polynomial*:

$$E(x) = \prod_{i \in I}(x - \alpha_i) \ .$$

Dlaczego taki: głównie to wiedziano, że go należy użyć...

Idea chcemy:

$$Q = fE$$

Ten wielomian zeruje się tam, gdzie są błędy, i mówi coś o f tam, gdzie nie ma błędów.

A ściśle:

BW1 wielomian E, o wiodącym współczynniku 1, stopnia $e \leq \left\lfloor \frac{n-k}{2} \right\rfloor$

BW2 wielomian Q stopnia $\leq e + k - 1$

BW3 dla każdego i zachodzi $w_i E(\alpha_i) = Q(\alpha_i)$

Słowem kodowym ma być Q/E (jako wielomian).

Uwaga. Jeśli Q/E nie jest zdefiniowane, bo się dzieli z resztą, albo ma za duży stopień, to zwracamy błąd.

Lemat 20.20. *Jeśli dla danego $\vec{w}$ istnieje $\vec{w'} \in RS$ takie że $d(w, w') \leq e \leq \left\lfloor \frac{n-k}{2} \right\rfloor$ to istnieją Q, E spełniające BW.*

Dowód.

$$E(X) = \prod_{i \in I}(X - \alpha_i)$$
$$Q(X) = E(X)f(X)$$ □

Lemat 20.21. *Jeśli Q_1, E_1 oraz Q_2, E_2 spełniają BW, to $Q_1/E_1 = Q_2/E_2$.*

Uwaga. Zauważmy, że jest to równość ilorazów i reszt, tzn. może być, że oba dzielenia dają resztę. Ale jeśli jeden się dzieli bez reszty, to drugi też, tj. jeśli jest poprawny wynik algorytmu, to wszystkie zwracane dają to samo.

Dowód. Rozpatrzmy wielomian

$$Q_1E_2 - Q_2E_1$$

To jest wielomian stopnia $\leq 2e + k - 1 < n$. Rozpatrzmy jego wartość w α_i:

$$w_iQ_1(\alpha_i)E_2(\alpha_i) - w_iQ_2(\alpha_i)E_1(\alpha_i) = 0$$

Czyli wielomian $Q_1E_2 - Q_2E_1$ jest zerem. W takim razie

$$Q_1E_2 = Q_2E_1 \implies \frac{Q_1}{E_1} = \frac{Q_2}{E_2}$$ □

To jak to odtworzyć? Będziemy interpolować wielomiany. Zauważmy, że nie interesuje nas, czy ten układ jest jednoznacznie określony, może być nadokreślony lub niedookreślony — dowolne rozwiązanie jest OK i wiemy, że jakieś jest.

Czas działania: Trzeba rozwiązać układ równań: $\mathcal{O}(n^3)$ (da się może ciut szybciej w zależność od danych) oraz podzielić dwa wielomiany — $\mathcal{O}(n^3)$. Ponownie dla specyficznych wartości może być ciut lepiej.

20.2 Rozszerzenia ciał

Alternatywne podejście konstrukcji ciała skończonego (w pewnym sensie bardziej naturalne): dodawanie elementu do ciała. Najmniejsze ciało zawierające dany element: tak myślimy o $\mathbb{C}$: to jest najmniejsze ciało zawierające $\mathbb{R}$ oraz i.

Przykład 20.22. Liczby postaci $\{a + b\sqrt{3} : a, b \in \mathbb{Q}\}$ są ciałem. Jedyna nietrywialna operacja to odwrotność, ale łatwo sprawdzić, że $(a + b\sqrt{3})(a - b\sqrt{3}) = a^2 - 3b^2 \neq 0$ i tym samym łatwo podać element odwrotny do $a + b\sqrt{3}$.

Definicja 20.23 (Rozszerzenie ciała)**.** Dla ciała $\mathbb{F}$ przez $\mathbb{F}\langle S \rangle$ oznaczamy najmniejsze ciało zawierające $\mathbb{F}, S$.

Rozszerzenie $\mathbb{F}\langle a \rangle$ jest *przestępne*, jeśli a nie jest pierwiastkiem żadnego wielomianu z $\mathbb{F}[x]$ takie a również nazywamy przestępnym. Jest *algebraiczne*, jeśli a jest pierwiastkiem jakiegoś wielomianu z $\mathbb{F}[x]$.

Żeby ta definicja miała sens, to elementy S powinny być albo zupełnie „spoza” albo z jakiegoś ciała $\mathbb{F}' \supseteq \mathbb{F}$.

20.2.1 Rozszerzenie przestępne

Jak wygląda rozszerzenie przestępne? Możemy sobie wyobrazić, że dodajemy do $\mathbb{F}$ jakiś „świeży” element α. W nowym ciele muszą być też wszystkie wielomiany z $\mathbb{F}[\alpha]$ oraz ich odwrotności. Są więc też wszystkie ilorazy wielomian przez wielomian.

Definicja 20.24 (Ciało ułamków prostych)**.** Rozważmy ciało $\mathbb{F}$ oraz wielomiany nad nim $\mathbb{F}[x]$. Na zbiorze $\{\frac{f}{g} : f, g \in \mathbb{F}[x], g \neq 0\}$ wprowadzamy relację równoważności $\frac{f}{g} \sim \frac{f'}{g'} \iff fg' = f'g$. Tak określony zbiór jest ciałem z naturalnie zadanym dodawaniem oraz mnożeniem:

$$\frac{f}{g} + \frac{f'}{g'} = \frac{fg' + f'g}{gg'} \quad \frac{f}{g} \cdot \frac{f'}{g'} = \frac{ff'}{gg'}$$

Twierdzenie 20.25. *Ciało ułamków prostych dla $\mathbb{F}$ jest izomorficzne z $\mathbb{F}\langle a \rangle$ dla przestępnego α.*

20.2.2 Rozszerzenia algebraiczne

Rozważamy teraz przypadek $\mathbb{F}\langle a\rangle$ gdy a jest pierwiastkiem jakiegoś wielomianu w $\mathbb{F}[x]$. Chcielibyśmy powiedzieć, że w takim razie to rozszerzenie zawiera $a, a^2, \ldots, a^{k-1}$, gdzie wielomian nierozkładalny, którego a jest pierwiastkiem, ma stopień k. (Tak jak w konstrukcji ciał skończonych). Ale czy tak jest, w szczególności, czy takie wielomian istnieje?

Definicja 20.26. Dla ciała $\mathbb{F}$ oraz elementu a z jego rozszerzenia przez $I(a)$ (ideał a) oznaczamy

$$\{f \in \mathbb{F}[x] \,:\, \bar{f}(a) = 0\}$$

Lemat 20.27. *Dla ciała $\mathbb{F}$ oraz elementu a z jego rozszerzenia $I(a)$ jest zamknięty na dodawanie i mnożenie przez wielomiany z $\mathbb{F}[x]$.*

Lemat 20.28. *Dla ciała $\mathbb{F}$ oraz elementu a z jego rozszerzenia $I(a)$ jest postaci*

$$\{fg \,:\, g \in \mathbb{F}[x]\}$$

dla pewnego wielomianu nierozkładalnego $f \in \mathbb{F}[x]$. W szczególności $\bar{f}(a) = 0$.

Dowód. Skoro $I(a)$ jest zamknięty na dodawanie i mnożenie przez wielomiany z $\mathbb{F}[x]$, to jeśli $f, g \in I(a)$ to również $\mathrm{nwd}(f, g) \in I(a)$. Chcielibyśmy mieć „$\mathrm{nwd}(I(a))$”, ale czy to ma sens? Ma: możemy dodawać kolejne elementy i patrzeć, czy spada stopień. Może spaść skończoną ilość razy, czyli od pewnego momentu uzyskane nwd skończonego zbioru dzieli każdy wielomian w $I(a)$. □

Stopień rozszerzenia to stopień tego wielomianu.

Wniosek 20.29. Rozszerzenie algebraiczne $\mathbb{F}\langle a\rangle$ jest izomorficzne z $\mathbb{F}[x]/_{\sim_h}$, gdzie h generuje $I(a)$.

Dowód. W obu rozszerzeniach zachodzi $\bar{h}(a) = 0$; formalnie łatwo pokazać izomorfizm $a \mapsto x$. □

20.3 Ciała algebraicznie domknięte

Definicja 20.30 (Ciało algebraicznie domknięte)**.** Ciało $\mathbb{F}$ jest *algebraicznie domknięte*, jeśli każdy wielomian nierozkładalny jest stopnia 1.

Fakt 20.31. *Ciało $\mathbb{F}$ jest algebraicznie domknięte wtedy i tylko wtedy gdy każdy wielomian ma pierwiastek.*

Fakt 20.32. *Ciało algebraicznie domknięte jest nieskończone.*

Przykład 20.33. $\mathbb{C}$ jest ciałem algebraicznie domkniętym. Nie jest nim $\mathbb{R}$ ani żadne $\mathbb{Z}_p$.

Twierdzenie 20.34. *Dla ciała $\mathbb{F}$ istnieje $\mathbb{F}' \supseteq \mathbb{F}$, które jest algebraicznie domknięte oraz działania $\mathbb{F}'$ obcięte do $\mathbb{F}$ to działania $\mathbb{F}$.*

Rozdział 21

Skończone $\mathbb{F}^*$ jest cykliczne

Chcemy pokazać, że jeśli $\mathbb{F}$ jest skończone, to $\mathbb{F}^*$ jest cykliczna. Dowód opiera się na wykazaniu, że istnieje w niej element rzędu $n = |\mathbb{F}| - 1$, co daje, że jest on generatorem. Aby to pokazać, będziemy dla każdego $k \leq n$ zliczać w grupie cyklicznej n elementowej oraz w grupie $\mathbb{F}^*$ elementy, które są rzędu k. Zauważmy, że wystarczy pokazać, że w grupie $\mathbb{F}^*$ jest nie więcej, niż w C_n (grupa cykliczna o n elementach).

Lemat 21.1. *Niech $R(G, k)$ oznacza liczbę elementów rzędu k w grupie abelowej G. Jeśli dla grupy skończonej G o n elementach zachodzi dla każdego k*

$$R(G, k) \leq R(C_n, k)$$

to G jest izomorficzne z C_n.

Dowód. Zauważmy, że grupy te mają taką samą ilość elementów i każdy element ma dokładnie określony rząd. Czyli

$$\sum_k R(G, k) = \sum_k R(C_n, k)$$

W związku z tym wszystkie nierówności

$$R(G, k) \leq R(C_n, k)$$

są w istocie równościami, w szczególności G ma element rzędu n, czyli jest cykliczna. $\square$

Niestety, zliczanie elementów rzędu k jest dość kłopotliwe. Łatwiej jest zliczyć elementy, których rząd *dzieli* k.

21.1 Rzędy elementów w grupie cyklicznej

Lemat 21.2. *Niech g będzie generatorem grupy cyklicznej G o n elementach. Wtedy g^m jest jej generatorem* $\iff$ $\mathrm{nwd}(m, n) = 1$. *W szczególności G ma $\varphi(n)$ generatorów.*

Dowód. Z algorytmu Euklidesa $\mathrm{nwd}(m, n) = 1 = an + bm$, bez zmniejszania ogólności $b > 0$. Czyli

$$g^{bm} = g^{1-an} = gg^{-an} = g(g^n)^{-a} = ge^{-a} = g$$

Czyli podgrupa generowana przez g^m zawiera g, czyli zawiera też podgrupę generowaną przez g, czyli całą grupę.

Jeśli g^m jest generatorem, to w szczególności generuje g. Czyli $g^{am} = g^1$. Ale wiemy, że najmniejszą potęgą ℓ elementy g, taką że $g^\ell = e$ jest n. Czyli $g^{am} = g^1$ oznacza, że dla pewnego b

$$am = 1 + bn$$

Ale to daje, że $\mathrm{nwd}(n, m) = 1$.

Z definicji, ilość liczb względnie pierwszych z n mniejszych niż n to $\varphi(n)$. $\square$

Na ćwiczeniach pokazaliśmy lemat:

Lemat 21.3. *Jeśli G jest cykliczna, to każda jej podgrupa jest cykliczna.*

Lemat 21.4. *Niech G będzie grupą cykliczną rzędu n. W G istnieje element rzędu d wtedy i tylko wtedy, gdy $d|n$.*

Dla ustalonego rzędu d tych elementów jest $\varphi(d)$ i są one wszystkie elementami podgrupy rzędu d.

Dowód. $\Rightarrow$ Popatrzmy na podgrupę generowaną przez ten element. Rząd tej podgrupy to rząd tego elementu. Jednocześnie rząd podgrupy dzieli rząd grupy, czyli $d|n$.

$\Leftarrow$ Niech g będzie generatorem. Rozpatrzmy $g^{\frac{n}{d}}$ (ponieważ $d|n$, to $\frac{n}{d}$ jest liczbą naturalną). Wtedy $(g^{\frac{n}{d}})^d = g^n = e$ i rząd nie może być mniejszy, bo wtedy rząd g też byłby mniejszy.

Rozpatrzmy podgrupę generowaną przez wszystkie elementy rzędu d. Z Lematu 21.3 jest ona generowana przez jeden element: q. Wtedy $q^d = e$, bo grupa jest przemienna i rząd każdego z generatorów to d. Jednocześnie rząd nie może być mniejszy niż d, bo wtedy rząd każdego elementu w generowanej grupie też jest mniejszy niż d. Z Lematu 21.2 grupa ta ma $\varphi(d)$ generatorów. □

21.2 Rzędy elementów w $\mathbb{F}^*$

Pokazaliśmy wcześniej, że:

Lemat 21.5 (Przypomnienie)**.** *Równanie $x^k = 1$ ma w ciele skończonym $\mathbb{F}$ najwyżej k różnych pierwiastków.*

Lemat 21.6. *Niech G będzie grupą skończoną rzędu n. Jeśli dla dowolnego $k \in \mathbb{N}$ zbiór $\{g \in G : g^k = e\}$ ma najwyżej k elementów, to G jest cykliczna.*

Dowód. Chcemy użyć Lematu 21.1. Ustalmy rząd k. Rząd elementu dzieli rząd grupy, czyli $k|n$. Ile elementów rzędu k jest w G? Jeśli nie ma takiego elementu, to założenie Lematu 21.1 dla k zachodzi. Załóżmy więc, że jest taki element.

Rozpatrzmy grupę generowaną przez ten element, jest ona cykliczna i ma k elementów. Wszystkie elementy w tej podgrupie spełniają równanie

$$x^k = e$$

Z Lematu 21.4 w grupie cyklicznej C_n jest k elementów spełniających to równanie, czyli w G nie ma innych elementów spełniających to równanie poza generowanymi przez ustalony element.

Z Lematu 21.2 wiemy, że grupa cykliczna ma $\varphi(k)$ elementów rzędu k,

$$R(C_n, k) = \varphi(k)$$

To jest też prawda w podgrupie G generowanej przez ten ustalony element. Czyli w G jest też najwyżej $\varphi(k)$ elementów rzędu k w G. Czyli jest ich dokładnie $\varphi(k)$:

$$R(G, k) = \varphi(k) = R(C_n, k)$$

Czyli liczba elementów rzędu k w grupie cyklicznej oraz w G jest taka sama. Czyli założenie Lematu 21.1 jest też spełnione dla tego k. □

Twierdzenie 21.7. *Grupa $\mathbb{F}^*$ jest cykliczna.*

Dowód. Wiemy, że w $\mathbb{F}$ równanie $x^k = 1$ ma najwyżej k pierwiastków. Potraktujmy je jako równanie w $\mathbb{F}^*$. Z Lematu 21.6 otrzymujemy, że $\mathbb{F}^*$ jest cykliczna. □