

Notatki z Wprowadzenia do matematyki III

Mateusz Łelyk, Bartosz Wcisło

9 stycznia 2016

1 Wstęp

Niniejszy skrypt zawiera notatki z konwersatorium prowadzonego w Instytucie Filozofii w roku akademickim 2015/2016. Części odpowiadające poszczególnym zajęciom zostały oddzielone seriami zadań. Są one dwójakiego typu: te, które warto rozwiązać, żeby upewnić się, że zrozumiało się treść wykładu oraz pomocne przy przygotowaniu do kolokwium nie zostały w żaden sposób wyróżnione, a dodatkowe, do których dobrze przejść po zrobieniu tych pierwszych, zostały oznaczone *.

2 Przestrzenie i odwzorowania liniowe

Zajęcia 5. X

Rozpocniemy od zarysu algebry liniowej. Algebra liniowa to dziedzina dotycząca przestrzeni liniowych i ich odwzorowań. Struktury te pozwalają zdefiniować czym są wektory w bardzo daleko posuniętej ogólności. W naszym wypadku dokonamy połowicznego uogólnienia, gdyż będziemy rozważali jedynie rzeczywiste przestrzenie liniowe.

Definicja 1. Przez **przestrzeń liniową** nad $\mathbb{R}$ rozumiemy strukturę $(V, +, \cdot)$, gdzie V jest niepustym zbiorem, zaś $+: V^2 \rightarrow V$ oraz $\cdot: \mathbb{R} \times V \rightarrow V$ operacjami spełniającymi następujące warunki:

1. $\forall u, v, w \in V \quad u + (v + w) = (u + v) + w$. (Łączność +)
2. $\exists v \in V \forall w \in V \quad v + w = w + v = w$. (Istnienie elementu neutralnego)

Można udowodnić, że istnieje dokładnie jeden element neutralny spełniający powyższy warunek. Będziemy go oznaczać przez $\vec{0}$, a z czasem po prostu przez 0.

3. $\forall v \in V \exists w \in V \ v + w = w + v = \vec{0}$. (Istnienie elementu odwrotnego)

Element w wprowadzony przez powyższą definicję będziemy oznaczać przez $-v$. Będziemy stosować standardową konwencję notacyjną i pisać $u - w$ zamiast $u + (-w)$.

4. $\forall v, w \in V \ v + w = w + v$. (Przemienność)
 5. $\forall x, y \in \mathbb{R} \forall v \in V \ (x + y)v = xv + yv$. (Rozdzielność działań)
 6. $\forall x \in \mathbb{R} \forall v, w \in V \ x(w + v) = xw + xv$. (Rozdzielność działań)
 7. $\forall x, y \in \mathbb{R} \forall v \in V \ x(yv) = (xy)v$. (Łączność $\cdot$)
 8. $\forall v \in V \ 1 \cdot v = v$.

Elementy przestrzeni V nazywamy **wektorami**. Przestrzenie liniowe nazywamy czasem **przestrzeniami wektorowymi**.

Wprowadzone aksjomaty pozwalają udowodnić wszystkie standardowe własności, których oczekiwaliśmy po wektorach w przestrzeni euklidesowej. Dla przykładu wymienimy dwie proste równości.

Fakt 2. *Ustalmy dowolną przestrzeń liniową V i dowolny wektor $v \in V$. Wówczas:*

1. $0 \cdot v = \vec{0}$.
2. $(-1)v = -v$.

Dowód. Udowodnimy punkt pierwszy, drugi pozostawiając jako ćwiczenie. Weźmy więc dowolną przestrzeń V oraz $v \in V$. Na mocy łączności mnożenia wektora przez liczby mamy:

$$2 \cdot (0 \cdot v) = (2 \cdot 0)v = 0 \cdot v.$$

Na mocy przemienności dodawania liczb z mnożeniem liczb przez wektory mamy:

$$2 \cdot (0 \cdot v) = (1 + 1)(0 \cdot v) = 0 \cdot v + 0 \cdot v.$$

Zatem uzyskujemy równość:

$$0 \cdot v = 0 \cdot v + 0 \cdot v,$$

z której po dodaniu stronami wektora $-0 \cdot v$ otrzymujemy:

$$\vec{0} = 0 \cdot v.$$

□

Od tej pory umówmy się oznaczać wektor zerowy przez 0. Wprowadźmy kolejne podstawowe pojęcie budowanej teorii.

Definicja 3. Podprzestrzenią przestrzeni V nazwiemy dowolny niepusty podzbiór $W \subseteq V$ taki, że dla dowolnych $w_1, w_2 \in W$ oraz $x \in \mathbb{R}$ wektory $w_1 + w_2$ oraz xw_1 również są elementami W .

Zauważmy, że z dowolności elementu x w powyższej definicji wynika, że $0 \in W$ dla dowolnej podprzestrzeni W dowolnej przestrzeni liniowej.

Podamy teraz garść przykładów przestrzeni liniowych i ich podprzestrzeni.

Przykład 4. 1. Liczby rzeczywiste $\mathbb{R}$ ze zwykłym dodawaniem i mnożeniem stanowią przestrzeń liniową nad $\mathbb{R}$. Geometrycznie można myśleć o tej przestrzeni liniowej jako o prostej.

2. Płaszczyzna $\mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}$ z dodawaniem: $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ oraz $a(x, y) = (ax, ay)$ jest przestrzenią liniową. Jest to kanoniczny przykład.

3. Analogicznie możemy zdefiniować przestrzenie euklidesowe wyższych wymiarów $\mathbb{R}^3, \mathbb{R}^4, \dots$

4. Prosta $L = \{(x, y) \mid y = 3x\}$ jest podprzestrzenią płaszczyzny $\mathbb{R}^2$. Istotnie: jeśli $y_1 = 3x_1$ oraz $y_2 = 3x_2$, to $(y_1 + y_2) = 3(x_1 + x_2)$ oraz $ay_1 = 3(ax_1)$, co dowodzi, że L jest domknięta na dodawanie wektorów i mnożenie wektora przez liczbę.

5. Prosta $L' = \{(x, y) \mid y = 3x + 5\}$ nie jest podprzestrzenią liniową $\mathbb{R}^2$, dlatego że nie jest domknięta ani na dodawanie ani na mnożenie wektorów przez liczby różne od ± 1 .

6. Parabola $P = \{(x, y) \mid y = x^2\}$ nie jest podprzestrzenią liniową $\mathbb{R}^2$.

7. Wielomiany jednej zmiennej z działaniami zwykłego dodawania wielomianów i mnożenia wielomianów przez liczbę są przestrzenią liniową.

8. Wielomiany jednej zmiennej stopnia ≤ 7 są podprzestrzenią przestrzeni wszystkich wielomianów. Wielomiany stopnia ≥ 7 nie są podprzestrzenią, bo np. $(x^8 + 1) - (x^8)$ nie jest wielomianem stopnia ≥ 8 , mimo że stanowi różnicę dwóch takich wielomianów.

9. Funkcje ciągłe $[0, 1] \rightarrow \mathbb{R}$ są przestrzenią liniową ze zwykłymi operacjami dodawania funkcji i mnożenia funkcji przez liczbę. (W ogólności, dla $A \subseteq \mathbb{R}$, przestrzeń funkcji ciągłych $A \rightarrow \mathbb{R}$ oznaczamy przez $C(A)$.)

Warto poczynić jedną uwagę, która może zapobiec nieporozumieniom: w wypadku wektorów w szkole średniej (przynajmniej tych na fizyce) dodawanie $v + w$ było dobrze określone tylko w wypadku, gdy punkt zaczepienia w był punktem końcowym v . W szczególności $w + v$ niekoniecznie było dobrze określonym działaniem. Bliższy naszej aksjomatyce byłby następujący obrazek: wszystkie wektory zaczepione są w początku układu współrzędnych, a dodawanie polega na odczepianiu jednego z wektorów, przyczepianiu go na końcu drugiego i przeprowadzaniu sumowania z fizyki licealnej. Zapobiegłszy ewentualnemu zamętowi, możemy kontynuować rozbudowę słownika.

Definicja 5. Powiemy, że v jest **kombinacją liniową** wektorów $v_1, \dots, v_n$, jeśli istnieją takie $x_1, \dots, x_n \in \mathbb{R}$, że

$$v = x_1 v_1 + x_2 v_2 + \dots + x_n v_n.$$

Jeśli wszystkie współczynniki x_i są równe 0, to kombinację nazywamy **trywialną**. Dla dowolnej przestrzeni V i dowolnych $v_1, \dots, v_n \in V$ przez $\text{lin}(v_1, \dots, v_n)$ oznaczamy podprzestrzeń kombinacji liniowych tych wektorów. Powiemy, że wektory $v_1, \dots, v_n$ są **liniowo zależne**, gdy 0 jest ich nietrywialną kombinacją liniową. W przeciwnym wypadku nazwiemy je **liniowo niezależnymi**. Jeśli $V = \text{lin}(v_1, \dots, v_n)$, to mówimy, że $v_1, \dots, v_n$ **rozpinają** przestrzeń V .

Zilustrujmy te pojęcia kilkoma przykładami:

- Przykład 6.**
1. Wektory $(1, 0), (0, 1), (1, 1) \in \mathbb{R}^2$ są liniowo zależne, bo $(1, 0) + (0, 1) - (1, 1) = 0$.
 2. Wektory $(1, 0, 0)$ oraz $(0, 0, 1) \in \mathbb{R}^3$ są liniowo niezależne, bo jeśli $a_1(1, 0, 0) + a_2(0, 0, 1) = (a_1, 0, a_2) = (0, 0, 0)$, to $a_1 = 0$ oraz $a_2 = 0$. Wkrótce poznamy bardziej systematyczne metody badania liniowej zależności układu wektorów.
 3. To raczej uwaga, która łączy dwa wprowadzone pojęcia niż przykład: otóż gdy $v_1, \dots, v_n$ stanowią układ liniowo niezależny, to $v \in \text{lin}(v_1, \dots, v_n)$ wtw układ $v, v_1, \dots, v_n$ jest liniowo zależny.

Jesteśmy już gotowi, by udowodnić główne twierdzenie pierwszych zajęć, które pozwoli nam zrozumieć strukturę przestrzeni liniowych. W szczególności zdefiniujemy pojęcie wymiaru i przekonamy się, że wszystkie przestrzenie liniowe skończonego wymiaru d są izomorficzne z przestrzenią $\mathbb{R}^d$.

Twierdzenie 7. Jeśli $v_1, \dots, v_k \in \text{lin}(w_1, \dots, w_n)$, to $k \leq n$.

Dowód. Dowód przeprowadzimy przez indukcję po n . Dla $n = 1$ twierdzenie jest oczywiste. Przypuśćmy więc, że zostało już udowodnione dla n i rozważmy

przypadek $n+1$. W takim wypadku bez utraty ogólności możemy zakładać, że $k > 1$, a więc zmieniawszy numerację rozważać wektory $v_0, v_1, \dots, v_k$ w przestrzeni $\text{lin}(w_0, w_1, \dots, w_n)$. Z definicji kombinacji liniowych istnieją takie liczby a_{ij} , że

$$\begin{aligned} v_0 &= a_{0,0}w_0 + a_{0,1}w_1 + \dots + a_{0,n}w_n \\ v_1 &= a_{1,0}w_0 + a_{1,1}w_1 + \dots + a_{1,n}w_n \\ &\vdots \\ v_k &= a_{k,0}w_0 + a_{k,1}w_1 + \dots + a_{k,n}w_n. \end{aligned}$$

Skoro wektory v_i są liniowo niezależne, to w szczególności żaden nie jest wektorem zerowym (w przeciwnym wypadku biorąc dowolny niezerowy współczynnik przy tym wektorze i 0 przy pozostałych otrzymalibyśmy 0 jako nietrywialną kombinację liniową). W szczególności któryś ze współczynników $a_{0,j}$ jest różny od 0. Ewentualnie zmieniając oznaczenia możemy przyjąć, że jest to współczynnik $a_{0,0}$. Niech teraz:

$$u_i = v_i - \frac{a_{i0}}{a_{00}}v_0.$$

Wówczas dla $i > 0$ mamy:

$$u_i = \left(a_{i0} - \frac{a_{i0}}{a_{00}}a_{00}\right)w_0 + \left(a_{i1} - \frac{a_{i0}}{a_{00}}a_{01}\right)w_1 + \dots + \left(a_{i,n} - \frac{a_{i0}}{a_{00}}a_{0,n}\right)w_n.$$

Współczynnik przy w_0 zeruje się, a zatem dla $i > 0$ mamy $u_i \in \text{lin}(w_1, \dots, w_n)$. Aby skorzystać z założenia indukcyjnego dla k i n wystarczy pokazać, że wektory $w_1, \dots, w_k$ są liniowo niezależne. Jednak wniosek ten łatwo uzyskać. Gdyby

$$b_1u_1 + \dots + b_ku_k = 0,$$

to mielibyśmy również z definicji w_i :

$$b_1v_1 + \dots + b_kv_k - \left(b_1\frac{a_{10}}{a_{00}} + \dots + b_k\frac{a_{k0}}{a_{00}}\right)v_0 = 0,$$

co przeczy liniowej niezależności $v_0, v_1, \dots, v_n$. Widzimy więc k wektorów liniowo niezależnych w przestrzeni rozpiętej przez n wektorów, a zatem $k < n$, stąd też i $k+1 < n+1$, co należało udowodnić.

□

Zadania

Zadanie 1. Udowodnij, że dla dowolnej przestrzeni liniowej V i dowolnego $v \in V$ zachodzi równość $(-1)v = -v$.

Zadanie 2. Pokaż, że dla dowolnej przestrzeni liniowej V , wektora $v \in V \setminus \{0\}$ i $c \in \mathbb{R}$

$$\text{jeśli } cv = 0, \text{ to } c = 0$$

Zadanie 3. Udowodnij, że jeśli W jest podprzestrzenią $\mathbb{R}^2$, to albo $W = \{(0, 0)\}$, albo $W = \mathbb{R}^2$, albo $W = \{(x, y) \mid y = ax\}$ dla pewnego $a \in \mathbb{R}$.

Zadanie 4. Niech V będzie przestrzenią liniową i $v_0, \dots, v_n \in V$. Niech $u_0, \dots, u_k \in \text{lin}(v_0, \dots, v_n)$. Udowodnij, że $\text{lin}(u_0, \dots, u_k) \subseteq \text{lin}(v_0, \dots, v_n)$. Kiedy zachodzi równość?

Zadanie 5. Wskaż zbiór przeliczalnej mocy rozpinający przestrzeń wielomianów o współczynnikach rzeczywistych. Pokaż, że żaden skończony zbiór wektorów nie rozpinają tej przestrzeni.

***Zadanie 6.** Podaj nieprzeliczalny zbiór liniowo niezależnych wektorów w $C([0, 1])$. Wskaż jakikolwiek zbiór rozpinający $C([0, 1])$.

Zadanie 7. Sprawdź, czy poniższe układy wektorów są liniowo niezależne:

1. $(1, 1), (0, 1)$.
2. $(1, 2, 3), (0, 1, 0), (0, 0, 1)$.
3. $(1, 2, 3), (2, 3, 4), (3, 4, 5), (4, 5, 6)$.
4. $(0, 0, 0, 1), (1, 0, 0, 0)$.
5. $(1, 5, 2, 4), (0, 9, 4, 2), (4, 7, 2, 1), (0, 8, 5, 4), (9, 3, 7, 4)$.

Zadanie 8. 1. Czy wektory $(1, 2), (0, 1)$ rozpinają $\mathbb{R}^2$?

2. Czy wektory $(1, 1, 0), (1, 0, 1), (0, 1, 1)$ rozpinają $\mathbb{R}^3$?

3. Czy wektory $(1, 4, 5, 0), (9, 2, 7, 1), (8, 4, 5, 2)$ rozpinają $\mathbb{R}^4$?

Zadanie 9. Załóżmy, że $v_1, \dots, v_n$ jest układem liniowo niezależnym. Wówczas $v \in \text{lin}(v_1, \dots, v_n)$ wtw $v_1, \dots, v_n, v$ jest liniowo zależny.

Zadanie 10. Pokaż, że układ wektorów $v_1, \dots, v_n$ jest liniowo zależny wtw dla pewnego i mamy $v_i \in \text{lin}(v_1, \dots, v_n)$.

Zajęcia 12. X

Zajmiemy się teraz następującym problemem: czy każda przestrzeń liniowa V jest postaci $\text{lin}(A)$ dla pewnego zbioru *liniowo niezależnych* wektorów z V . Jak się okaże, odpowiedź na to pytanie jest twierdząca. W tym momencie już powinno dla Czytelnika być jasne, że jeśli tak jest, to wybór A na pewno nie jest jednoznaczny. Wyprowadźmy ważny wniosek z Twierdzenia 7.

Wniosek 8. *Przypuśćmy, że $v_0, \dots, v_n \in V, w_0, \dots, w_m \in V$ są dwoma liniowo niezależnymi układami wektorów w V . Jeśli $\text{lin}(v_0, \dots, v_n) = \text{lin}(w_0, \dots, w_m)$, to $n = m$.*

Wprowadzimy teraz dwie kluczowe dla tej części wykładu definicje

Definicja 9. Powiemy, że przestrzeń liniowa V ma **wymiar** n (ozn. $\dim V = n$) wtedy i tylko wtedy, gdy w V istnieje n wektorów liniowo niezależnych, które rozpinają V . Powiemy, że przestrzeń V jest skończenie wymiarowa jeśli dla pewnego n $\dim V = n$.

Podczas zajęć skupimy się na przestrzeniach skończenie wymiarowych - jeśli nie stwierdzamy inaczej, to "przestrzeń liniowa" znaczy "skończenie wymiarowa przestrzeń liniowa". Następująca uwaga wynika natychmiast z definicji wymiaru.

Uwaga 10. Niech V będzie przestrzenią liniową i $v_1, \dots, v_n \in V$. Jeśli $v_1, \dots, v_n$ są liniowo niezależne, to $\dim V \geq n$.

Następująca uwaga wynika natomiast z Twierdzenia 7:

Uwaga 11. Niech V będzie przestrzenią liniową i $v_1, \dots, v_n \in V$. Jeśli $v_1, \dots, v_n$ rozpinają V , to $\dim V \leq n$.

Z powyższych dwóch uwag wypływa ważny:

Wniosek 12. *Niech V będzie przestrzenią liniową. Jeśli istnieją $v_1, \dots, v_n$ takie, że $V = \text{lin}(v_1, \dots, v_n)$, oraz $v_1, \dots, v_n$ są liniowo niezależne, to $\dim V = n$.*

Przykład 13. $\dim \mathbb{R}^n = n$. Faktycznie, łatwo się przekonać, że wektory e_i dla $i \leq n$, gdzie

$$e_i = \begin{matrix} (0, \dots, 1, \dots, 0) \\ \text{na } i\text{-tym miejscu 1, na pozostałych 0} \end{matrix}$$

są liniowo niezależne i rozpinają $\mathbb{R}^n$.

Przykład 14. Niech V będzie przestrzenią liniową wielomianów stopnia co najwyżej n . Wtedy $\dim V = n + 1$. Niech teraz W będzie przestrzenią liniową

wszystkich wielomianów o współczynnikach rzeczywistych zmiennej x . Wtedy W nie ma skończonego wymiaru, bo dla dowolnego k wektory

$$x^0, x, \dots, x^k$$

są liniowo niezależne.

Definicja 15. Bazą przestrzeni liniowej V nazwiemy dowolny układ wektorów $v_1, \dots, v_n$, który jest liniowo niezależny i rozpiną V .

Przykład 16. W $\mathbb{R}^n$ wektory e_i (wprowadzone w Przykładzie 13) stanowią bazę. Jest to tzw. **baza standardowa**. Niech $v = (1, \dots, 1) \in \mathbb{R}^n$. Inną bazą w $\mathbb{R}^n$ jest np. $v - e_0, v - e_1, \dots, v - e_n$.

Przykład 17. W przestrzeni wielomianów stopnia co najwyżej n wektory

$$x^0, x, x^2, \dots, x^n$$

stanowią bazę.

Uwaga 18. W przestrzeni n -wymiarowej dowolne dwie bazy są równoliczne i mają moc n .

Zachęcamy do próby udowodnienia poniższego stwierdzenia samodzielnie i przejrzenia poniższego dowodu tylko dla pewności, że zrobiło się to poprawnie.

Stwierdzenie 19. Niech V będzie przestrzenią liniową i $v_1, \dots, v_n \in V$. Następujące warunki są równoważne:

1. $v_1, \dots, v_n$ jest maksymalnym układem liniowo niezależnym w V , tzn. dla dowolnego $v \in V \setminus \{v_1, \dots, v_n\}$, $v_1, \dots, v_n, v$ jest liniowo zależny.
2. $v_1, \dots, v_n$ jest minimalnym układem rozpinającym V , tzn. dla dowolnego $i \leq n$ $\{v_1, \dots, v_n\} \setminus \{v_i\}$ nie rozpiną V .
3. $\dim V \leq n$ i $v_1, \dots, v_n$ są liniowo niezależne.
4. $\dim V \geq n$ i $v_1, \dots, v_n$ rozpinają V .
5. $v_1, \dots, v_n$ są bazą V .

Dowód. (1) $\Rightarrow$ (2) Przypuśćmy, że $v_1, \dots, v_n$ jest maksymalnym układem liniowo niezależnym w V . Pokażemy, że $v_1, \dots, v_n$ rozpiną V . Ustalmy dowolne $v \in V$. Z maksymalności układu $v_1, \dots, v_n$ układ $v_1, \dots, v_n, v$ jest liniowo zależny, a zatem, na mocy Zadania 9 $v \in \text{lin}(v_1, \dots, v_n)$. Wykażemy, że $v_1, \dots, v_n$ jest minimalny. Gdyby bowiem, dla pewnego $i \leq n$, $\{v_1, \dots, v_n\} \setminus \{v_i\}$ rozpinął V , to mielibyśmy $v_i \in \text{lin}(\{v_1, \dots, v_n\} \setminus \{v_i\})$ i $v_1, \dots, v_n$ byłby liniowo zależny,

na mocy Zadania 9.

(2) $\Rightarrow$ (3) Skoro $v_1, \dots, v_n$ rozpina V , to $\dim V \leq n$, na mocy Uwagi 11. Musimy pokazać, że $v_1, \dots, v_n$ jest liniowo niezależny. Gdyby był jednak liniowo zależny, to na mocy Zadania 10 mielibyśmy $v_i \in \text{lin}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n)$ i $(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n)$ rozpinałby V wbrew minimalności $v_1, \dots, v_n$.

(3) $\Rightarrow$ (4) Jeśli $v_1, \dots, v_n$ są liniowo niezależne, to $\dim V \geq n$ na mocy Uwagi 10. Gdyby istniał $v \in V \setminus \text{lin}(v_1, \dots, v_n)$ to układ $v_1, \dots, v_n, v$ byłby liniowo niezależny na mocy Zadania 9. Wtedy jednak musilibyśmy mieć

$$\dim V \geq n + 1 > n,$$

co stoi w sprzeczności z tym, że na mocy (3) $\dim V \leq n$.

(4) $\Rightarrow$ (5) Przypuśćmy, że $\dim V \geq n$ i $v_1, \dots, v_n$ rozpinają V . Gdyby były liniowo zależne, to rozumując podobnie jak w dowodzie (2) $\Rightarrow$ (3) dostalibyśmy, że (b.o.o.) $v_2, \dots, v_n$ rozpinają V . Wtedy jednak na mocy Uwagi 11 mielibyśmy

$$\dim V \leq n - 1 < n,$$

co przeczy założeniu, że $\dim V \geq n$.

(5) $\Rightarrow$ (1) Niech $v_1, \dots, v_n$ będzie liniowo niezależny i rozpina V . Musimy pokazać, że $v_1, \dots, v_n$ jest maksymalny, tzn. dla dowolnego v

$$v_1, \dots, v_n, v$$

jest liniowo zależny. Tak jednak jest, bo z definicji $v \in \text{lin}(v_1, \dots, v_n)$, a zatem układ $v_1, \dots, v_n, v$ jest liniowo zależny na mocy Zadania 9. $\square$

Zauważmy, że na mocy definicji wymiaru przestrzeni liniowej każda (skończenie wymiarowa) przestrzeń liniowa ma bazę. Bazę danej przestrzeni liniowej możemy traktować jako tymczasową reprezentację tej przestrzeni. Poniższe twierdzenie dobrze obrazuje to stwierdzenie: mając wybraną jakąkolwiek bazę przestrzeni V każdemu wektorowi możemy w jednoznaczny sposób przyporządkować krotkę liczb rzeczywistych "reprezentującą" ten wektor w tej bazie. Napisaliśmy *tymczasową* dlatego, że dana przestrzeń w żaden kanoniczny sposób nie wyróżnia żadnej ze swoich baz, dlatego też można je swobodnie zmieniać w zależności od tego, z którą wygodniej jest w danej sytuacji pracować. Druga część twierdzenia, zwana też *Twierdzeniem Steinitza o wymianie*, pokazuje (m.in.), że każdy wektor jest elementem jakiejś bazy. Sens tej nazwy objaśnimy na przykładzie po przeprowadzeniu dowodu.

Twierdzenie 20. Niech V będzie przestrzenią liniową z bazą $v_1, \dots, v_n$.

1. każdy wektor $v \in V$ można przedstawić jako kombinację liniową $v_1, \dots, v_n$ w dokładnie jeden sposób, tzn. dla dowolnego $v \in V$ istnieją $a_1, \dots, a_n$ takie, że

$$v = \sum_{i=1}^n a_i v_i$$

i dla dowolnych $b_1, \dots, b_n$ jeśli

$$v = \sum_{i=1}^n b_i v_i,$$

to dla dowolnego $i \leq n$, $a_i = b_i$.

2. każdy liniowo niezależny układ wektorów $w_1, \dots, w_k$ dla $k \leq n$ można uzupełnić do bazy wektorami z bazy $v_1, \dots, v_n$.

Dowód. (1) Ustalmy wektor $v \in V$. Istnienie liczb a_i o których mówi ten punkt wynika z tego, że $V = \text{lin}(v_1, \dots, v_n)$. Ustalmy jakiegokolwiek takie liczby a_i . Przypuśćmy, że dla pewnych $b_i \in \mathbb{R}$

$$v = \sum_{i=1}^n b_i v_i,$$

Wtedy jednak

$$v = \sum_{i=1}^n b_i v_i = \sum_{i=1}^n a_i v_i$$

czyli

$$\sum_{i=1}^n (b_i - a_i) v_i = 0$$

Z liniowej niezależności $v_1, \dots, v_n$ wynika, że dla dowolnego i , $a_i - b_i = 0$, czyli dla dowolnego i , $a_i = b_i$.

- (2) Ustalmy liniowo niezależny układ $w_1, \dots, w_k$ dla $k \leq n$. Rozpatrzmy następujący układ wektorów rozpinający V

$$w_1, \dots, w_k, v_1, \dots, v_n \tag{1}$$

Usuńmy z tego układu wszystkie wektory, które są kombinacjami liniowymi poprzednich wektorów w ciągu. Z założenia o liniowej niezależności $w_1, \dots, w_k$ żaden wektor z tego układu nie został usunięty. Dostajemy zatem układ wektorów rozpinający V :

$$w_1, \dots, w_k, v_{i_1}, \dots, v_{i_m} \tag{2}$$

Wystarczy pokazać, że układ ten jest liniowo niezależny. Ustalmy $a_1, \dots, a_k$, $b_{i_1}, \dots, b_{i_m}$ takie, że

$$a_1 w_1 + \dots + a_k w_k + b_{i_1} v_{i_1} + \dots + b_{i_m} v_{i_m} = 0$$

Z liniowej niezależności $w_1, \dots, w_n$ wynika, że jeśli wszystkie b_{i_j} są równe zero, to wszystkie powyższe współczynniki są równe zero i kombinacja jest trywialna. Jeśli zaś któreś z b_{i_j} jest niezerowe, to istnieje też takie o największym współczynniku j . Wtedy jednak v_{i_j} wyraża się jako kombinacja liniowa poprzednich wektorów, co przeczy konstrukcji układu (2). Stąd musimy mieć

$$a_1 = \dots = a_k = b_{i_1} = \dots = b_{i_m} = 0.$$

□

Przykład 21. Ustalmy przestrzeń liniową V z bazą $v_1, \dots, v_n$. Ustalmy $k < n$ i popatrzmy na

$$V_1 = \text{lin}(v_1, \dots, v_k)$$

V_1 jest podprzestrzenią V (Definicja 3) rozpiętą przez pierwsze k wektorów wybranej bazy. Ustalmy inną bazę $w_1, \dots, w_k$ przestrzeni V_1 . Wtedy na mocy punktu 2 powyższego twierdzenia możemy uzupełnić $w_1, \dots, w_k$ do bazy przestrzeni V elementami bazy $v_1, \dots, v_n$. To uzupełnienie *musi* składać się z $n - k$ wektorów $v_{k+1}, \dots, v_n$, bowiem $v_1, \dots, v_k$ da się uzyskać jako kombinacje liniowe z układu $w_1, \dots, w_k$. Wymieniliśmy zatem bazę w podprzestrzeni V_1 , nie ruszając pozostałych wektorów wybranej w pierw bazą V .

Sensowność poniższej definicji uzasadniliśmy we właśnie przeprowadzonym dowodzie:

Definicja 22. Niech V będzie przestrzenią liniową z bazą $v_1, \dots, v_n$ i $v \in V$. **Współrzędnymi** wektora v w bazie $v_1, \dots, v_n$ nazywamy jedyne takie liczby $a_1, \dots, a_n \in \mathbb{R}$, że

$$v = \sum_{i=1}^n a_i v_i$$

Przykład 23. Rozpatrzmy $V = \mathbb{R}^3$ i $v = (1, 2, 3)$. Wektor ten ma w bazie standardowej przestrzeni $\mathbb{R}^3$ (tzn. złożonej z wektorów $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$) współrzędne

$$1, 2, 3$$

Jeśli jednak zmienimy kolejność wektorów bazy standardowej rozważając np. bazę $(0, 1, 0)$, $(1, 0, 0)$, $(0, 0, 1)$, to *ten sam* wektor w tej bazie będzie miał współrzędne

$$2, 1, 3$$

Jeśli rozważymy bazę tej samej przestrzeni złożoną z wektorów $(1, 1, 0)$, $(1, 0, 1)$, $(0, 1, 1)$, to v będzie miał w niej współrzędne

$$0, 1, 2$$

Uwaga 24. Z Twierdzenia 20 wynika, że dla dowolnego wektora $v \in V \setminus \{0\}$ istnieje taka baza V , że v ma w tej bazie współrzędne

$$1, 0, \dots, 0$$

$n-1$ zer

gdzie $n = \dim V$. Istotnie, jeśli $v \neq 0$, to jednoelementowy układ

$$v$$

jest liniowo niezależny, zatem można uzupełnić go do bazy V . W tej bazie v ma żądane współrzędne.

Uwaga 25. Niech $v_1, \dots, v_n$ będzie bazą przestrzeni V . Jeśli v ma w tej bazie współrzędne

$$a_1, \dots, a_n$$

a w ma w niej współrzędne

$$b_1, \dots, b_n$$

to $v + w$ ma w niej współrzędne $a_1 + b_1, \dots, a_n + b_n$ i dla dowolnego $c \in \mathbb{R}$, cv ma w niej współrzędne $ca_1, \dots, ca_n$.

Zadania

Zadanie 11. Znaleźć dwie różne, nowe (tzn. różne od podanych w powyższym wykładzie) bazy przestrzeni $\mathbb{R}^3$. Znaleźć współrzędne wektora $(1, 1, 1)$ w tych bazach.

Zadanie 12. Znaleźć bazy podanych podprzestrzeni $V_1 \subseteq V$ dla

1. $V = \mathbb{R}^3, V_1 = \{(x_1, x_2, x_3) \mid x_1 + 2x_2 - x_3 = 0\}$

2. $V = \mathbb{R}^3, V_1 = \text{lin}((1, 2, 1), (0, 1, 1), (1, 3, 2))$.

Zadanie 13. Niech $V = \mathbb{R}^4$. Znaleźć współrzędne wektorów $(1, 0, 8, 9), (1, 7, 1, 7)$, w bazie $(1, 1, 1, 0), (1, 1, 0, 1), (1, 0, 1, 1), (0, 1, 1, 1)$.

***Zadanie 14.** Niech $v \in \mathbb{R}^3 \setminus \{0\}$ i niech $a, b, c \in \mathbb{R}$ nie wszystkie będą zerowe. Udowodnij, że istnieje baza $\mathbb{R}^3$, w której v ma współrzędne a, b, c .

Zadanie 15. Udowodnij, że $v_1, \dots, v_n$ jest bazą przestrzeni V wtedy i tylko wtedy, gdy każdy wektor da się przedstawić jako kombinację liniową $v_1, \dots, v_n$ w dokładnie jeden sposób. (Implikację $(\Rightarrow)$ zrobiliśmy w Twierdzeniu 20. Pozostaje implikacja przeciwna.)

Zajęcia 19. X

Wprowadzimy kolejne kluczowe pojęcie rozważanej teorii – przekształcenia liniowego.

Definicja 26. Niech V, W będą dwiema przestrzeniami liniowymi. Odwzorowanie $\phi : V \rightarrow W$ nazwiemy **liniowym**, gdy spełnione są następujące dwa warunki:

1. $\phi(v + v') = \phi(v) + \phi(v')$ dla dowolnych $v, v' \in V$.
2. $\phi(cv) = c\phi(v)$ dla dowolnego $v \in V$ oraz $c \in \mathbb{R}$.

Uwaga 27. Z drugiego warunku od razu wynika, że jeśli $\phi : V \rightarrow W$ jest liniowe, to $\phi(0) = 0$.

Zilustrujmy to pojęcie szeregiem przykładów i kontrprzykładów. Dla uproszczenia zapisu będziemy funkcję o argumentach z przestrzeni $\mathbb{R}^n$ zapisywać jako funkcję n zmiennych rzeczywistych. Tak więc $\phi(x, y, z)$ jest skrótowym zapisem $\phi((x, y, z))$.

1. Dla dowolnej przestrzeni liniowej V odwzorowanie $\text{id} : V \rightarrow V$, takie że $\phi(v) = v$ dla dowolnego $v \in V$ jest liniowe.
2. Odwzorowanie $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}$ dane wzorem $\phi(x, y, z) = z$ jest liniowe.
3. Odwzorowanie $\phi : \mathbb{R} \rightarrow \mathbb{R}^2$ dane wzorem $\phi(t) = (3t, -\frac{1}{2}t)$ jest liniowe.
4. Odwzorowanie $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dane wzorem $\phi(x, y) = (\frac{x+y}{\sqrt{2}}, \frac{-x+y}{\sqrt{2}})$ jest liniowe.
5. Różniczkowanie jako odwzorowanie przestrzeni $\mathbb{R}[X]$ wielomianów jednej zmiennej o współczynnikach rzeczywistych w siebie jest liniowe.

We wszystkich powyższych przypadkach liniowość sprawdzamy bezpośrednio. Przerachujmy drugi z nich:

$$\begin{aligned}\phi((x, y, z) + (x', y', z')) &= \phi(x + x', y + y', z + z') \\ &= z + z' \\ &= \phi(x, y, z) + \phi(x', y', z').\end{aligned}$$

Drugi warunek sprawdzamy bardzo podobnie:

$$\begin{aligned}\phi(c(x, y, z)) &= \phi(cx, cy, cz) \\ &= cz \\ &= c\phi(x, y, z).\end{aligned}$$

Być może lepsze intuicje na temat odwzorowań liniowych pozwolą wyrobić poniższe kontrprzykłady:

6. Odwzorowanie $\phi : \mathbb{R} \rightarrow \mathbb{R}$ dane wzorem $\phi(x) = x^2$ nie jest liniowe. Istotnie: $\phi(x + x') = (x + x')^2 = x^2 + (x')^2 + 2xx' = \phi(x) + \phi(x') + 2xx'$, co na ogół różni się od $\phi(x) + \phi(x')$, o ile tylko x, x' oraz 2 są różne od zera.

7. Odwzorowanie $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ dane wzorem $\phi(x, y) = xy$ nie jest liniowe. Faktycznie: $\phi((x, y) + (x', y')) = \phi(x + x', y + y') = (x + x')(y + y') = xy + x'y' + xy' + x'y = \phi(x, y) + \phi(x', y') + xy' + x'y$, zaś ostatni wyraz jest równy $\phi(x, y) + \phi(x', y')$ tylko o ile w obu parach $(x, y), (x', y')$ przynajmniej jedna ze współrzędnych jest równa 0.
8. Odwzorowanie $\phi : \mathbb{R} \rightarrow \mathbb{R}$ zadane $\phi(x) = x + 1$ nie jest liniowe, gdyż np. $\phi(0) = 1$.

Okazuje się, że odwzorowania liniowe dopuszczają bardzo prosty opis. Poniższe stwierdzenie stanowi uzasadnienie tego faktu. Przy okazji sformułujemy kluczową własność baz, która mogłaby w zasadzie posłużyć jako ich równoważna definicja. Dla niewielkiego uproszczenia zapisu ograniczymy się do przypadku skończonego wymiarowego.

Stwierdzenie 28. Niech V, W będą skończone wymiarowymi przestrzeniami liniowymi. Niech $(v_1, \dots, v_n)$ będzie dowolną bazą V . Wówczas dla dowolnych wektorów $w_1, \dots, w_n$ istnieje dokładnie jedno $\phi : V \rightarrow W$ liniowe, takie że dla dowolnego i :

$$\phi(v_i) = w_i.$$

Dowód. Pokażemy najpierw jedyność odwzorowania ϕ , a następnie zastanowimy się nad jego ewentualnym istnieniem. Niech v będzie dowolnym wektorem z przestrzeni V . Na mocy Twierdzenia 20, punkt 1. wektor v ma w bazie $(v_1, \dots, v_n)$ jednoznacznie określone współrzędne, czyli istnieje dokładnie jeden ciąg liczb $(a_1, \dots, a_n)$, taki że

$$v = a_1 v_1 + a_2 v_2 + \dots + a_n v_n.$$

Na mocy liniowości odwzorowania ϕ zachodzą następujące równości:

$$\begin{aligned} \phi(v) &= \phi\left(\sum_{j=1}^n a_j v_j\right) \\ &= \sum_{j=1}^n \phi(a_j v_j) \\ &= \sum_{j=1}^n a_j \phi(v_j) \\ &= \sum_{j=1}^n a_j w_j. \end{aligned}$$

To pokazuje, że ustalenie wartości odwzorowania ϕ na wektorach bazowych wyznacza wartości odwzorowania ϕ także i na wszystkich innych wektorach V .

Wystarczy więc pokazać, że odwzorowanie liniowe ϕ o zadanych wartościach istnieje. Zdefiniujmy więc ϕ w następujący sposób. Niech v będzie dowolnym wektorem z przestrzeni V . Niech $a_1, \dots, a_n$ będą takimi liczbami, że

$$v = a_1 v_1 + a_2 v_2 + \dots + a_n v_n.$$

Definiujemy:

$$\phi(v) = \sum_{j=1}^n a_j w_j.$$

Odwzorowanie ϕ jest określone na całym V , bo $v_1, \dots, v_n$ rozpinają V . Powyższy wzór rzeczywiście wyznacza funkcję (tzn. dla dowolnego v tylko jeden wektor $w \in W$ spełnia definicję $\phi(v)$), bo ciąg liczb $(a_1, \dots, a_n)$ jest wyznaczony jednoznacznie. Wystarczy więc sprawdzić liniowość. Weźmy dowolne dwa wektory v, v' i niech w rozważanej bazie mają współrzędne odpowiednio $(a_1, \dots, a_n)$ oraz $(b_1, \dots, b_n)$. Wówczas:

$$\phi(v + v') = \sum_{j=1}^n (a_j + b_j) w_j = \sum_{j=1}^n a_j w_j + \sum_{j=1}^n b_j w_j = \phi(v) + \phi(v').$$

Podobnie widzimy, że:

$$\phi(cv) = \sum_{j=1}^n c a_j w_j = c \sum_{j=1}^n a_j w_j = c \phi(v).$$

Co dowodzi, że tak zdefiniowane ϕ jest liniowe. □

Ostatnie twierdzenie pozwoli nam na wprowadzenie użytecznego sposobu zapisu odwzorowań liniowych, który umożliwi nam systematyczne znajdowanie współrzędnych (w dowolnej bazie) wektora stanowiącego wartość odwzorowania liniowego w ustalonym punkcie.

Definicja 29. Przez **macierz** (o wymiarach) $n \times m$ o wyrazach rzeczywistych rozumiemy tablicę

$$\begin{pmatrix} a_{11} & \dots & a_{1m} \\ a_{21} & \dots & a_{2m} \\ \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix},$$

gdzie dla dowolnych indeksów i, j wartości a_{ij} są rzeczywiste. Macierz powyższej postaci oznaczamy też $(a_{i,j})_{i \leq n, j \leq m}$, a czasem niezbyt precyzyjnie po prostu $(a_{i,j})$.

Zauważmy, że powyższa definicja używa niezdefiniowanego pojęcia tablicy. W istocie moglibyśmy napisać, że macierz to funkcja ze zbioru $\{1, 2, \dots, n\} \times \{1, 2, \dots, m\}$ w $\mathbb{R}$, która parze (i, j) przypisuje liczbę $a_{i,j}$. Będziemy jednak konsekwentnie mówić o macierzach jako o tablicach.

Kolejna definicja wiąże z każdym odwzorowaniem liniowym (po ustaleniu bazy) pewną macierz, co jak się niedługo przekonamy stanowi punkt wyjścia do ważnych technik rachunkowego rozwiązywania problemów geometrycznych.

Definicja 30. Niech $\phi : V \rightarrow W$ będzie przekształceniem liniowym. Niech $\mathcal{A} = (v_1, \dots, v_n)$, $\mathcal{B} = (u_1, \dots, u_k)$ będą dowolnymi bazami tych przestrzeni. **Macierzą przekształcenia** ϕ w bazach $\mathcal{A}, \mathcal{B}$ nazwiemy taką macierz $(a_{i,j})_{i \leq k, j \leq n}$, że dla dowolnych indeksów i, j wartość $a_{i,j}$ to i -ta współrzędna wektora $\phi(v_j)$ w bazie $\mathcal{B}$.

Macierz tę oznaczamy $M(\phi)_{\mathcal{A}}^{\mathcal{B}}$.

Powyższa definicja może wywoływać u czytelnika wrażenie, że faktyczne znalezienie macierzy danego przekształcenia w danej bazie jest sprawą trudną. Jak się jednak przekonamy, jest to dość nietrafne poczucie.

Przykład 31. Niech st oznacza bazę standardową w $\mathbb{R}^2$. Niech $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ oznacza odbicie symetryczne względem osi OX . Znajdziemy macierz ϕ w bazach st, st . Niech a_{ij} oznacza wyraz tej macierzy o indeksach i, j . Wektor e_1 przechodzi przy odwzorowaniu ϕ na tenże wektor. Czyli współrzędne $\phi(e_1)$ w bazie st to po prostu $(1, 0)$. Zatem $a_{1,1} = 1$, zaś $a_{2,1} = 0$. Dalej zauważmy, że $\phi(e_2) = -e_2$. Czyli współrzędne $\phi(e_2)$ w bazie st to $(0, -1)$. Stąd $a_{2,1} = 0$, natomiast $a_{2,2} = -1$. Wnioskujemy więc, że szukana macierz to:

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

Niech $\phi : V \rightarrow V$. Od tej pory przez **macierz przekształcenia ϕ w bazie $\mathcal{A}$** będziemy rozumieli macierz tego przekształcenia w bazach $\mathcal{A}, \mathcal{A}$.

Przykład 32. Macierz obrotu o 30° w $\mathbb{R}^2$ zapisuje się jako:

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

Do tej pory nie widać jeszcze, jak korzystamy z faktu, że macierz zdefiniowaliśmy jako tablicę. Zdefiniujemy teraz operację na macierzach, której opanowanie znacznie ułatwia graficzny sposób zapisu macierzy.

Definicja 33. Niech $A = (a_{i,j}), B = (b_{i,j})$ będą macierzami o wymiarach $n \times m$ oraz $m \times l$. Wówczas przez AB rozumiemy macierz $n \times l$ o wyrazach $(c_{i,j})$, gdzie dla dowolnych indeksów i, j zachodzi równość:

$$c_{i,j} = \sum_{k=1}^m a_{i,k} b_{k,j}.$$

Innymi słowy: macierz AB to macierz, w której wyraz na miejscu (i, j) wyznaczamy w następujący sposób: bierzemy i -ty wiersz macierzy A oraz j -tą kolumnę macierzy B , przykładamy jedno do drugiego tak, żeby lewy kraniec wiersza znajdował się na górze kolumny, mnożymy po współrzędnych i dodajemy uzyskane iloczyny. W razie problemów ze zrozumieniem, na czym polega ta operacja, warto być może spróbować obejrzeć animację w sieci.

Przykład 34.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 1 \cdot 5 + 2 \cdot 7 & 1 \cdot 6 + 2 \cdot 8 \\ 3 \cdot 5 + 4 \cdot 7 & 3 \cdot 6 + 4 \cdot 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}$$

Na koniec podamy jedno stwierdzenie, które wyjaśnia związek między macierzą przekształcenia liniowego a samą macierzą.

Definicja 35. Niech $v \in V$ będzie dowolnym wektorem, zaś $\mathcal{A} = (v_1, \dots, v_n)$ – skończoną bazą przestrzeni V . **macierzą współrzędnych** wektora v w bazie $\mathcal{A}$ nazwiemy macierz o wymiarach $n \times 1$, w której na miejscu $(i, 1)$ stoi i -ta współrzędna wektora v w tej bazie.

Przykładowo, macierz współrzędnych wektora e_2 w bazie standardowej w $\mathbb{R}^3$, to

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

Stwierdzenie 36. Niech $\phi : V \rightarrow W$. Niech $\mathcal{A}$ będzie dowolną bazą w V , natomiast $\mathcal{B}$ – dowolną bazą w W . Jeśli A jest macierzą ϕ w bazach $\mathcal{A}, \mathcal{B}$, zaś $\mathbf{v}$ – macierzą współrzędnych v w bazie $\mathcal{A}$, to $A\mathbf{v}$ jest macierzą współrzędnych wektora $\phi(v)$ w bazie $\mathcal{B}$.

Powyższe stwierdzenie łatwo zastosować w konkretnej sytuacji.

Przykład 37. Znajdziemy obraz wektora $(15, 17)$ po obrocie o 30° . Sprowadza się to do ustalenia, jakie są jego współrzędne w bazie standardowej. Wektor ten ma w bazie standardowej współrzędne 15, 17. Wiemy też, jak wygląda macierz obrotu o 30° w bazie standardowej. Wystarczy więc wykonać mnożenie:

$$\begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \cdot \begin{pmatrix} 15 \\ 17 \end{pmatrix} = \begin{pmatrix} \frac{15\sqrt{3}-17}{2} \\ \frac{15+17\sqrt{3}}{2} \end{pmatrix}$$

Zadania

Zadanie 16. Niech $\phi : V \rightarrow W$ będzie odwzorowaniem liniowym oraz na W . Niech $\psi : W \rightarrow U$ będzie odwzorowaniem liniowym. Udowodnij, że $\psi \circ \phi$ też jest liniowe.

Zadanie 17. Znajdź macierze następujących przekształceń w bazie standardowej:

1. Identytyność w $\mathbb{R}^k$.
2. Obrót o 45° wokół osi OX w $\mathbb{R}^3$.
3. Przekształcenie dane wzorem $\phi(x, y, z) = z$.
4. Przekształcenie dane wzorem $\phi(x, y, z) = (z, y, x)$.
5. Obrót o dowolny kąt α w $\mathbb{R}^2$.
6. Symetria względem płaszczyzny $\text{lin}(e_2, e_3)$ w $\mathbb{R}^3$.

***Zadanie 18** (Dość ważny przykład). Znajdź macierz obrotu o kąt θ wokół prostej $\text{lin}(v)$ w bazie (v, v', v'') przestrzeni $\mathbb{R}^3$, gdzie v, v', v'' są wzajemnie prostopadłe i tej samej długości.

Zadanie 19. Pomnóż macierze:

1. $\begin{pmatrix} 10 & 5 \\ 12 & -7 \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 \\ 3 & 8 \end{pmatrix}$.
2. $\begin{pmatrix} 0 & 7 & 11 \\ 1 & 23 & -2 \end{pmatrix} \cdot \begin{pmatrix} 5 & -e \\ 2 & -2 \\ 1 & 3 \end{pmatrix}$.
3. $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.
4. $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.
5. $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.
6. $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} -1 & 2 \\ 3 & 8 \end{pmatrix}$.

Zadanie 20. Pokaż, że dowolne przekształcenie liniowe $\phi : \mathbb{R} \rightarrow \mathbb{R}$ jest zadane wzorem

$$\phi(x) = cx$$

dla pewnego $c \in \mathbb{R}$.

Zadanie 21. Znajdź obraz następujących wektorów w $\mathbb{R}^3$ po obrocie o kąt $\frac{\pi}{3}$ wokół osi OZ .

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

***Zadanie 22.** Niech V będzie przestrzenią liniową. Pokaż, że jeśli układ wektorów $v_1, \dots, v_n \in V$ ma następującą własność:

dla dowolnej przestrzeni liniowej W i dowolnej funkcji $\phi : \{v_1, \dots, v_n\} \rightarrow W$ istnieje *dokładnie jedno* przekształcenie liniowe $\phi' : V \rightarrow W$ takie, że

$$\phi' \upharpoonright_{\{v_1, \dots, v_n\}} = \phi$$

to $v_1, \dots, v_n$ jest bazą V .

2.1 Zajęcia 26.X

Konwencja 1. Od tej pory pisząc $\phi : V \rightarrow W$ będziemy mieli zawsze na myśli, że ϕ jest przekształceniem liniowym przestrzeni liniowych V i W . Wszelkie odejścia od tej umowy będą sygnalizowane.

Wprowadźmy jedno z ważniejszych pojęć, służących do badania przekształceń liniowych.

Definicja 38. Niech $\phi : V \rightarrow W$.

1. **Jądrem** ϕ nazywamy przeciwobraz wektora zerowego przestrzeni W przy przekształceniu ϕ , tj.

$$\phi^{-1}(0) = \{v \in V \mid \phi(v) = 0\}$$

Jądro ϕ oznaczamy przez $\ker \phi$.

2. **Obrazem** ϕ nazywamy obraz przestrzeni V przy przekształceniu ϕ , tj.

$$\phi[V] = \{w \in W \mid \exists v \in V \ \phi(v) = w\}$$

Obraz ϕ oznaczamy przez $\text{im} \phi$.

Przykład 39. Przypomnijmy, że przez e_i oznaczamy wektory bazy standardowej $\mathbb{R}^n$ (zob. Przykład 13)

- Jeśli $\pi : \mathbb{R}^4 \rightarrow \mathbb{R}^3$,

$$\pi(x, y, z, v) = (x, y, z),$$

to $\ker \pi = \text{lin}\{e_4\}$, a $\text{im} \pi = \mathbb{R}^3$.

- Jeśli $\pi : \mathbb{R}^3 \rightarrow \mathbb{R}^4$,

$$\pi(x, y, z) = (x, y, z, 0),$$

to $\ker \pi = \{0\}$ i $\text{im} \pi = \text{lin}\{e_1, e_2, e_3\}$.

- Jeśli ρ jest obrotem $\mathbb{R}^3$ wokół osi $\text{lin}\{(1, 0, 0)\}$ o kąt θ , to $\ker \phi = \{0\}$ i $\text{im} \phi = \mathbb{R}^3$.

Uwaga 40. Niech $\phi : V \rightarrow W$. $\ker \phi$ i $\text{im} \phi$ to podprzestrzenie liniowe V i W odpowiednio. Faktycznie, sprawdźmy, że $\text{im} \phi$ jest domknięte na dodawanie: niech w_1, w_2 będą wektorami z $\text{im} \phi$. Wtedy z definicji istnieją $v_1, v_2 \in V$ takie, że

$$\phi(v_1) = w_1, \phi(v_2) = w_2$$

Wtedy na mocy liniowości $\phi(v_1 + v_2) = \phi(v_1) + \phi(v_2) = w_1 + w_2$, zatem $v_1 + v_2$ świadczy o tym, że $w_1 + w_2 \in \text{im} \phi$. Sprawdzenie pozostałych warunków pozostawimy jako ćwiczenie.

Uwaga 41. Niech $\phi : V \rightarrow W$. Jeśli $v_1, \dots, v_n$ rozpina V , to $\phi(v_1), \dots, \phi(v_n)$ rozpina $\text{im} \phi$. Faktycznie, niech $w \in \text{im} \phi$. Wtedy istnieje $v \in V$ takie, że $\phi(v) = w$. Z założenia, istnieją $a_1, \dots, a_n$ takie, że

$$v = a_1 v_1 + \dots + a_n v_n$$

Stąd, z liniowości ϕ , mamy

$$\begin{aligned} w = \phi(v) &= \phi(a_1 v_1 + \dots + a_n v_n) = \\ &= a_1 \phi(v_1) + \dots + a_n \phi(v_n) \end{aligned}$$

liniowość ϕ

Zatem $w \in \text{lin}\{\phi(v_1), \dots, \phi(v_n)\}$.

Poniższe stwierdzenie jest jednym z ważniejszych w tym rozdziale:

Stwierdzenie 42. Niech $\phi : V \rightarrow W$ i niech $v_1, \dots, v_n, v_{n+1}, \dots, v_{n+k}$ będzie bazą V taką, że $v_1, \dots, v_n$ jest bazą V . Wtedy $\phi(v_{n+1}), \dots, \phi(v_{n+k})$ jest bazą dla $\text{im} \phi$

Dowód. Ustalmy wszystko jak powyżej. Musimy pokazać, że $\phi(v_{n+1}), \dots, \phi(v_{n+k})$ rozpina $\text{im} \phi$ i jest liniowo niezależny. Dla dowodu tego pierwszego ustalmy dowolny $w \in \text{im} \phi$. Wtedy z definicji istnieje $v \in V$ taki, że

$$\phi(v) = w$$

Skoro $v_1, \dots, v_{n+k}$ jest bazą V , to istnieją $a_1, \dots, a_{n+k}$ takie, że

$$v = a_1 v_1 + \dots + a_{n+k} v_{n+k}$$

Na mocy liniowości ϕ i tego, że $v_1, \dots, v_n$ są z jądra ϕ dostajemy:

$$\begin{aligned} w = \phi(v) &= \phi(a_1 v_1 + \dots + a_{n+k} v_{n+k}) = \\ &= \underbrace{a_1 \phi(v_1) + \dots + a_n \phi(v_n)}_{=0} + a_{n+1} \phi(v_{n+1}) + \dots + a_{n+k} \phi(v_{n+k}) = \\ &= a_{n+1} \phi(v_{n+1}) + \dots + a_{n+k} \phi(v_{n+k}) \end{aligned}$$

Zatem $w \in \text{lin}\{\phi(v_{n+1}), \dots, \phi(v_{n+k})\}$.

To teraz liniowa niezależność: przypuśćmy, że

$$b_1 \phi(v_{n+1}) + \dots + b_k \phi(v_{n+k}) = 0$$

Wtedy z liniowości ϕ , $\phi(b_1 v_{n+1} + \dots + b_k v_{n+k}) = 0$, stąd

$$b_1 v_{n+1} + \dots + b_k v_{n+k} \in \ker \phi$$

Skoro $v_1, \dots, v_n$ rozpinają $\ker \phi$ to istnieją $a_1, \dots, a_n$ takie, że

$$b_1 v_{n+1} + \dots + b_k v_{n+k} = a_1 v_1 + \dots + a_n v_n$$

Ponieważ jednak $v_1, \dots, v_{n+k}$ jest bazą, to

$$a_1 = \dots = a_n = b_1 = \dots = b_k = 0,$$

co było do wykazania. □

Wyciągnijmy bardzo ważny

Wniosek 43. Niech $\phi : V \rightarrow W$. Wtedy

$$\dim V = \dim \text{im} \phi + \dim \ker \phi$$

Dowód. Faktycznie, zauważmy, że na mocy powyższego stwierdzenia, mamy, że baza $\text{im} \phi$ łączy $\dim V - \dim \ker \phi$ wektorów. To wystarczy. □

Przykład 44. Niech $\sigma : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ będzie zadane

$$\sigma(u, v) = (u + v, 2u + 2v)$$

Łatwo sprawdzić, że $\ker \phi = \{(u, v) \in \mathbb{R}^2 \mid u = v\} = \text{lin}\{(1, 1)\}$. Łatwo sprawdzić, że wektor $(1, 0)$ dopełnia tę bazę jądra do bazy $\mathbb{R}^2$. Zatem $\sigma(1, 0) = (1, 2)$ jest bazą $\text{im} \phi$.

Definicja 45. Niech $\phi : V \rightarrow W$

1. ϕ jest **monomorfizmem** jeśli ϕ jest różnowartościowe.
2. ϕ jest **epimorfizmem** jeśli $\text{im}\phi = W$.
3. ϕ jest **izomorfizmem** jeśli jest monomorfizmem i epimorfizmem zarazem.

Przykład 46. • $\pi : \mathbb{R}^3 \rightarrow \mathbb{R}$ zdefiniowane $\pi(x, y, z) = z$ jest epimorfizmem.

- Niech $R[X]$ będzie przestrzenią wielomianów. Przekształcenie $\sigma : \mathbb{R}^2 \rightarrow R[X]$ zdefiniowane

$$\sigma(a, b) = aX + b$$

jest monomorfizmem.

- Obrót $\mathbb{R}^2$ o kąt θ jest izomorfizmem.

Sformułujemy teraz banalną uwagę, która, w połączeniu z dotychczasową naszą wiedzą, ma bardzo niebanalne konsekwencje.

Uwaga 47. Jeśli $\phi : V \rightarrow W$ i $\ker \phi = \{0\}$, to ϕ jest monomorfizmem. Istotnie, przypuśćmy, że

$$\phi(v) = \phi(u)$$

dla pewnych $u, v \in V$. Wtedy z liniowości

$$\phi(u - v) = 0$$

Stąd $u - v \in \ker \phi$. Na mocy założenia, że $\ker \phi = \{0\}$ dostajemy, że $u - v = 0$, czyli $u = v$.

Wniosek 48. Jeśli $\phi : V \rightarrow W$ jest epimorfizmem i $\dim V = \dim W$, to ϕ jest izomorfizmem.

Dowód. Faktycznie, wystarczy sprawdzić, że ϕ jest monomorfizmem, a zatem, że $\ker \phi = \{0\}$. Ale jeśli ϕ jest epimorfizmem, to $\text{im}\phi = W$, stąd na mocy Wniosku 43 i założenia

$$\dim V = \dim \ker \phi + \dim \text{im}\phi = \dim \ker \phi + \dim W = \dim \ker \phi + \dim V$$

Stąd $\dim \ker \phi = 0$, a zatem $\ker \phi = \{0\}$. □

Wniosek 49. Niech V będzie przestrzenią liniową i $\dim V = n$. Wtedy V jest izomorficzne z $\mathbb{R}^n$.

Dowód. Oczywiście $\dim \mathbb{R}^n = n$. Na mocy poprzedniej uwagi wystarczy znaleźć epimorfizm $V \rightarrow \mathbb{R}^n$. Ustalmy bazę $v_1, \dots, v_n$ przestrzeni V i zdefiniujmy funkcję

$$\phi'(v_i) = e_i, 1 \leq i \leq n$$

gdzie $e_1, \dots, e_n$ jest bazą standardową $\mathbb{R}^n$. Na mocy Stwierdzenia 28 ϕ rozszerza się do przekształcenia liniowego $\phi : V \rightarrow W$. Ponieważ dla dowolnego i , $e_i \in \text{im} \phi$, więc

$$\text{im} \phi = \mathbb{R}^n,$$

stąd ϕ jest epimorfizmem, a nasz dowód jest zakończony. $\square$

Wniosek 50. *Dowolne dwie przestrzenie liniowe tego samego wymiaru są izomorficzne.*

Wniosek 51. *Niech $\dim V = n$ i $\dim W = k$. Przestrzeń przekształceń liniowych z V do W jest izomorficzna z przestrzenią macierzy o wymiarach $k \times n$.*

Dowód. Na mocy powyższego wniosku wystarczy sprawdzić, że obydwie przestrzenie mają ten sam wymiar. Zauważmy, że rodzina macierzy

$$e_{k,l} = (a_{i,j})_{i \leq k, j \leq n},$$

gdzie

$$a_{i,j} = \begin{cases} 1 & \text{jeśli } i = k, j = l \\ 0 & \text{w p.p.} \end{cases}$$

jest bazą przestrzeni macierzy (pozostawimy to jako łatwe ćwiczenie). Ustalmy bazy $v_1, \dots, v_n$ przestrzeni V i $w_1, \dots, w_k$ przestrzeni W . Zdefiniujmy przekształcenie $\phi_{i,j}$ zadane na bazie

$$\begin{aligned} \phi_{i,j}(v_i) &= w_j \\ \phi_{i,j}(v_l) &= 0 \text{ dla } l \neq i \end{aligned}$$

Sprawdzenie, że tak zdefiniowane $\phi_{i,j}$ stanowią bazę przestrzeni przekształceń $V \rightarrow W$ pozostawimy jako ćwiczenie. Stąd obydwie przestrzenie mają wymiar równy $n \cdot k$. Zatem są izomorficzne. $\square$

Konwencja 2. Od tej pory i na wiek wieków, przestrzeń przekształceń liniowych z V do W będziemy oznaczać przez $L(V, W)$, a przestrzeń macierzy $M_{n \times k}$.

Uwaga 52. Ustalmy V i W wymiaru n i k odpowiednio i bazami $\mathcal{V}$ i $\mathcal{W}$. Niech $\phi_{i,j}, e_{i,j}$ będą bazami przestrzeni $L(V, W)$ i $M_{k \times n}$ jak w powyższym twierdzeniu. Oczywiście istnieje wiele (ile dokładnie?) izomorfizmów $L(V, W)$ na $M_{k \times n}$, które przeprowadzają bazę złożoną z $\phi_{i,j}$ na bazę złożoną z $e_{i,j}$. Wśród nich jeden jednak się narzuca, mianowicie ten zadany przez

$$\phi_{i,j} \mapsto e_{i,j}$$

Zauważmy, że przy tym izomorfizmie przekształcenie $\phi \in L(V, W)$ przechodzi na macierz $M(\phi)_{\mathcal{V}}^{\mathcal{W}}$. Dowód tej obserwacji pozostawimy jako ćwiczenie.

Definicja 53. Niech $\phi : V \rightarrow W$ i $\psi : W \rightarrow Y$. **Złożenie przekształceń** ϕ i ψ , to przekształcenie

$$\begin{aligned}\psi \circ \phi : V &\rightarrow Y \\ \psi \circ \phi(v) &= \psi(\phi(v))\end{aligned}$$

We Wniosku 51 widzieliśmy, że przestrzeń macierzy jest izomorficzna z przestrzenią liniową przekształceń liniowych. Pozostaje pytanie: jakie działanie na macierzach odpowiada składaniu przekształceń. Poprzedzimy to jednak jedną uwagą dotyczącą algebry macierzy

Uwaga 54. Wypisując odpowiednio długie napisy można przekonać się, że mnożenie macierzy jest łączne, tj. dla dowolnych macierzy A, B, C rozmiaru $n \times k, k \times l, l \times m$ odpowiednio, mamy

$$A \cdot (B \cdot C) = (A \cdot B) \cdot C$$

Ponadto mnożenie macierzy jest rozdzielne względem ich dodawania: dla dowolnych A rozmiaru $n \times k$ i B, C rozmiaru $k \times l$ mamy

$$A \cdot (B + C) = A \cdot B + A \cdot C \quad (*)$$

Podobnie, jeśli $c \in \mathbb{R}$, to mamy

$$A \cdot (c \cdot B) = (c \cdot A) \cdot B \quad (**)$$

Ustalmy teraz przestrzenie liniowe V i W z bazami $\mathcal{V}, \mathcal{W}$ i wymiaru odpowiednio k i n . W szczególności, jeśli $l = 1$, to B i C reprezentują pewne wektory przestrzeni V , a $A \cdot B$ i $A \cdot C$ pewne wektory przestrzeni W . Możemy zatem myśleć o A jako o funkcji

$$\begin{aligned}V &\rightarrow W \\ v &\mapsto A \cdot v\end{aligned}$$

(*) i (**) stwierdzają wtedy, że takie przekształcenie jest liniowe. Łatwo przekonać się, że jest to dokładnie przekształcenie $\phi : V \rightarrow W$, takie, że $A = M(\phi)_{\mathcal{V}}^{\mathcal{W}}$.

Stwierdzenie 55. Niech $\phi : V \rightarrow W, \psi : W \rightarrow Y$ i $\mathcal{V}, \mathcal{W}, \mathcal{Y}$ będą bazami V, W i Y odpowiednio. Wtedy

$$M(\psi \circ \phi)_{\mathcal{V}}^{\mathcal{Y}} = M(\psi)_{\mathcal{W}}^{\mathcal{Y}} \cdot M(\phi)_{\mathcal{V}}^{\mathcal{W}}$$

Dowód. Ustalmy wszystko jak powyżej. Z definicji złożenia dla dowolnego $v \in V$ mamy

$$\psi \circ \phi(v) = \psi(\phi(v))$$

W bazie $\mathcal{Y}$, na mocy Stwierdzenia 36 współrzędne wektora v w bazie $\mathcal{Y}$, to

$$M(\psi)_{\mathcal{W}}^{\mathcal{Y}} \cdot (M(\phi)_{\mathcal{V}}^{\mathcal{W}} \cdot \mathbf{v})$$

Na mocy łączności (Uwaga 54) powyższy wektor jest równy

$$(M(\psi)_{\mathcal{W}}^{\mathcal{Y}} \cdot M(\phi)_{\mathcal{V}}^{\mathcal{W}}) \cdot \mathbf{v}$$

Stąd, na mocy jednoznaczności odwzorowania $\phi \mapsto M(\phi)_{\mathcal{V}}^{\mathcal{Y}}$ musi być

$$M(\psi)_{\mathcal{W}}^{\mathcal{Y}} \cdot M(\phi)_{\mathcal{V}}^{\mathcal{W}} = M(\psi \circ \phi)_{\mathcal{V}}^{\mathcal{Y}}$$

□

2.1.1 Zadania

Zadanie 23. Dokończ dowód Uwagi 40.

Zadanie 24. Znaleźć macierze przekształceń z Przykładu 39

1. w bazach standardowych.
2. w bazie standardowej i bazie $v_i = (1, \dots, 1) - e_i, i \leq n$.

Zadanie 25. Znaleźć macierz przekształcenia σ z punktu drugiego Przykładu 46 w bazach $\{e_1, e_2\} \subseteq \mathbb{R}^2$ i $\{1, X\} \subseteq R[X]$.

Zadanie 26. Udowodnij Wniosek 50.

Zadanie 27. Uzupełnij dowód Wniosku 51.

Zadanie 28. Policz izomorfizmy z Uwagi 52.

Zadanie 29. Łatwo przekonać się o tym, o czym łatwo przekonać się w Uwadze 54.

Zadanie 30. Wskaż przekształcenie liniowe $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, dla którego $\text{im } \phi \subseteq \ker \phi$.

Zadanie 31. Niech ϕ^n oznacza

$$\underbrace{\phi \circ \dots \circ \phi}_n$$

Udowodnij, że $\text{im } \phi \subseteq \ker \phi$ wtedy i tylko wtedy, gdy $\phi^2 = 0$. Znajdź macierz $A \in M_{2 \times 2}$ taką, że $A^2 = 0$.

Zajęcia 09.XI

Zanim przejdziemy do dalszego rozwijania teorii przestrzeni liniowych i ich przekształceń, czeka nas krótkie interludium poświęcone operacjom, które można wykonywać na samych macierzach. Rozwinięcie teorii tych operacji pozwoli nam potem dokonywać użytecznych rachunków służących badaniu pytań geometrycznych. Rozpocznijmy od zdefiniowania pewnego szczególnego kształtu macierzy:

Definicja 56. Powiemy, że macierz A jest w **postaci schodkowej**, jeśli każdy zerowy wiersz tej macierzy znajduje się poniżej wszystkich wierszy niezerowych oraz dla dowolnego i , jeśli j jest najmniejszym wskaźnikiem, dla którego $a_{i,j} \neq 0$, zaś j' jest najmniejszym wskaźnikiem, dla którego $a_{i+1,j'} \neq 0$, to $j' > j$. Innymi słowy, pierwszy niezerowy wyraz w danym wierszu jest położony na lewo od pierwszego niezerowego wyrazu w kolejnym wierszu.

Powiemy, że macierz A jest w **zredukowanej postaci schodkowej**, jeśli jest w postaci schodkowej, a ponadto w każdym wierszu pierwszym niezerowym wyrazem jest 1 i jest to jedyny niezerowy wyraz w danej kolumnie.

Bardzo łatwo zrozumieć, czym są macierze w postaci schodkowej na przykładzie:

Przykład 57. 1. Macierz $\begin{pmatrix} 10 & 5 & 4 & 1 \\ 0 & 0 & 12 & -7 \\ 0 & 0 & 0 & -17 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ jest w postaci schodkowej, ale nie zredukowanej.

2. Macierz $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ jest w postaci schodkowej zredukowanej.

3. Macierz $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ nie jest w postaci schodkowej.

Definicja 58. Przez **operacje elementarne** na wierszach macierzy rozumiemy następujące działania:

1. Dodanie do siebie dwóch *różnych* wierszy.
2. Pomnożenie wiersza przez liczbę rzeczywistą różną od zera.
3. Zamiana dwóch wierszy miejscami.

Za pomocą operacji elementarnych można każdą macierz sprowadzić do postaci schodkowej zredukowanej. Rozważenie kilku przykładów całkowicie wyjaśnia dlaczego tak jest. Formalnie fakt ten można uzasadnić przez indukcję po liczbie wierszy.

Fakt 59. Każdą macierz można sprowadzić do postaci schodkowej zredukowanej za pomocą operacji elementarnych.

Przykład 60. Sprowadzimy do postaci schodkowej zredukowanej macierz:

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

Odejmijmy od drugiego wiersza pierwszy wiersz pomnożony przez 4, a od trzeciego wiersza pierwszy wiersz pomnożony przez 7. Otrzymujemy:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{pmatrix}.$$

Odejmijmy od trzeciego wiersza drugi wiersz pomnożony przez 2. Mamy:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{pmatrix}.$$

Pomnóżmy drugi wiersz przez $(-\frac{1}{3})$. Otrzymujemy:

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

Odejmijmy wreszcie od pierwszego wiersza drugi wiersz pomnożony przez 2. Dostajemy:

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

Ostatnia macierz jest już w postaci schodkowej zredukowanej.

Nietrudno zauważyć, że niezerowe wiersze macierzy w postaci schodkowej zredukowanej są liniowo niezależne: dowolna liniowa zależność między nimi pozwoliłaby określić liniową zależność między tymi fragmentami wierszy, które mają wyraz 1 na dokładnie jednym miejscu i 0 na wszystkich innych. W związku z tym schodkowanie macierzy pozwala szybko ustalać, ile jest wektorów liniowo niezależnych w danych zbiorze skończonym.

Jest też jasne, że jeśli pomyślimy o wierszach macierzy jako o wektorach pewnej przestrzeni, to operacje elementarne nie zmieniają wymiaru podprzestrzeni rozpiętej przez te wektory.

Możemy podsumować te uwagi poniższym stwierdzeniem:

Stwierdzenie 61. Niech $v_1, \dots, v_n$ będą dowolnymi wektorami w pewnej przestrzeni skończonej wymiarowej V . Niech A będzie macierzą, której wierszami są współrzędne wektorów $v_1, \dots, v_n$ w pewnej ustalonej bazie $\mathcal{A}$. Wówczas $\dim \text{lin}(v_1, \dots, v_n)$ to dokładnie liczba niezerowych wierszy macierzy A' powstałej przez zeszkokowanie macierzy A .

Schodkowanie macierzy daje też efektywną metodę rozwiązywania układów równań liniowych.

Przykład 62. Powiedzmy, że chcemy znaleźć wszystkie takie (x, y, z) , które stanowią rozwiązanie układu równań liniowych:

$$\begin{cases} x + 2y + 3z = 0 \\ 4x + 5y + 6z = 0. \end{cases}$$

Układy równań takie jak powyżej, w których wszystkie wyrazy wolne wynoszą 0 nazywamy **jednorodnie**. Opiszemy najpierw na przykładzie, jak rozwiązuje się właśnie takie układy równań liniowych.

Zauważmy, że jeśli (x, y, z) spełnia powyższy układ równań, to spełnia również każde z równań, jak również każde z równań powstałych przez przemnożenie powyższych równań przez liczbę rzeczywistą.

Co więcej, możemy zauważyć, że dowolne (x, y, z) , które będzie spełniało oba powyższe równania będzie też spełniać ich sumę. Stwierdzamy więc na końcu, że jeśli rozważymy macierz współczynników:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix},$$

to operacje elementarne na wierszach nie zmieniają zbioru rozwiązań powyższego układu równań.

Powyższa macierz po sprowadzeniu do postaci schodkowej przybiera następującą postać:

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

A to pozwala nam na prosty opis zbioru rozwiązań. Mianowicie otrzymujemy układ równań:

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

Czyli $x = z$, a $y = -2z$. Wynika stąd, że rozwiązania powyższego układu równań tworzą jednowymiarową przestrzeń liniową, w której skład wchodzi wszystkie wektory postaci $(t, -2t, t)$, dla $t \in \mathbb{R}$, innymi słowy $\text{lin}((1, -2, 1))$.

Rozwiązania układów równań tworzą przestrzeń liniową tylko dla układów jednorodnych, ale na szczęście rozwiązywanie układów równań niejednorodnych jest równie proste. Powiedzmy, że chcemy znaleźć wszystkie takie (x, y, z) , dla których:

$$\begin{cases} x + 2y + 3z = 7 \\ 4x + 5y + 6z = 8. \end{cases}$$

Ponownie zapisujemy to równanie jako macierz:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 7 \\ 4 & 5 & 6 & 8 \end{array} \right).$$

Ponownie możemy ją sprowadzić do postaci schodkowej zredukowanej:

$$\left(\begin{array}{ccc|c} 1 & 0 & -1 & \frac{-19}{3} \\ 0 & 1 & 2 & \frac{20}{3} \end{array} \right).$$

Znów otrzymujemy zgrabny opis rozwiązań powyższego układu równań, mianowicie $x = z + \frac{-19}{3}$, zaś $y = -2z + \frac{20}{3}$. Czyli rozwiązania to dokładnie wektory postaci $(t + \frac{-19}{3}, -2t + \frac{20}{3}, t)$ dla dowolnych $t \in \mathbb{R}$. Innymi słowy jest to zbiór wektorów $v + (\frac{-19}{3}, \frac{20}{3}, 0)$, gdzie $v \in \text{lin}((1, -2, 1))$. Zbiór ten będziemy też oznaczać przez $\text{lin}((1, -2, 1)) + (\frac{-19}{3}, \frac{20}{3}, 0)$.

Odnotujmy na końcu, że rozwiązania jednorodnych układów równań liniowych stanowią zawsze przestrzeń liniową. Podobnie rozwiązania dowolnych układów równań są zawsze postaci $V + w$, gdzie V jest przestrzenią liniową, a w jest pewnym wektorem.

Kontynuujemy badanie algebraicznej struktury macierzy. Wiemy już, że po ustaleniu bazy przekształcenia w danej przestrzeni skończonej wymiarowej V odpowiadają jednoznacznie macierzom kwadratowym o wymiarze $\dim V \times \dim V$. (Wniosek 51) Fakt ten pozwala nam łatwo uzyskać przydatne informacje o operacjach na macierzach.

Stwierdzenie 63. Dla dowolnego n istnieje taka macierz I , że dla dowolnej macierzy A rozmiaru $n \times n$ zachodzi równość: $AI = IA = A$.

Dowód. Ustalmy n i rozważmy dowolną przestrzeń n -wymiarową V . Niech id oznacza przekształcenie tożsamościowe tej przestrzeni (czyli takie przekształcenie, że $\text{id}(v) = v$ dla dowolnego $v \in V$). Dla dowolnego przekształcenia $\phi : V \rightarrow V$ zachodzi wówczas równość:

$$\phi \circ \text{id} = \text{id} \circ \phi = \phi,$$

co wynika w oczywisty sposób z definicji przekształcenia tożsamościowego. Ustalmy dowolną bazę $v_1, \dots, v_n$ przestrzeni V . Niech I będzie macierzą id w tej bazie. Z odpowiedniości między mnożeniem macierzy w ustalonej bazie, a składaniem przekształceń (Stwierdzenie 55) wynika, że macierz I spełnia tożsamość w tezie stwierdzenia. $\square$

Chcemy jeszcze sprawdzić, czy macierz I jest określona jednoznacznie. *Prima facie* do jej ustalenia konieczny był uprzedni wybór bazy. Przyjmijmy jednak, że wybraliśmy bazę $(v_1, \dots, v_n)$. Wówczas dla dowolnego j mamy $\text{id}(v_j) = v_j$, czyli $\text{id}(v_j)$ ma w wybranej przez nas bazie współrzędne $(0, \dots, 0, 1, 0, \dots, 0)$, gdzie jedynka znajduje się na j -tym miejscu. Na mocy definicji macierzy przekształcenia znaczy to, że macierz przekształcenia id w tej bazie miała postać:

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

niezależnie od wyboru bazy.

Skoro wiemy już, że w zbiorze macierzy kwadratowych rozmiaru $n \times n$ istnieje jedynka względem mnożenia macierzy, możemy spytać, czy operacja mnożenia macierzy jest odwracalna.

Definicja 64. Niech A będzie dowolną macierzą rzeczywistą wymiaru $n \times n$. Powiemy, że macierz B jest **odwrotna** do A , jeśli spełnione są równości:

$$AB = BA = I.$$

Macierz odwrotną do A oznaczamy A^{-1} .

Operację odwracania można też zdefiniować dla przekształceń liniowych w sposób znany z kursów logiki.

Definicja 65. Niech $\phi : V \rightarrow V$ będzie dowolnym przekształceniem liniowym. Powiemy, że ψ jest **odwrotne** do ϕ , jeśli

$$\phi \circ \psi = \psi \circ \phi = \text{id}.$$

Łatwo można pokazać, że dla każdego izomorfizmu istnieje przekształcenie odwrotne.

Stwierdzenie 66. Niech $\phi : V \rightarrow V$ będzie dowolnym izomorfizmem liniowym. Wówczas istnieje przekształcenie liniowe $\psi : V \rightarrow V$ odwrotne do ϕ .

Zanim przejdziemy do dowodu poczyńmy dwie uwagi: po pierwsze, skoro ϕ jest bijekcją, to istnienie pewnej funkcji ψ , która spełnia tezę stwierdzenia wynika z czysto teoriomnogościowych rozważań. Nieoczywista jest tylko liniowość. Po drugie, zastrzeżmy, że niekiedy teza powyższego twierdzenia służy za definicję izomorfizmu liniowego.

Dowód. Niech V będzie przestrzenią liniową, a $v_1, \dots, v_n$ pewną jej bazą. Wiemy, że izomorfizm przekształca dowolną bazę na bazę, a zatem $\phi(v_1), \dots, \phi(v_n)$ też jest bazą.

Wiemy, że przekształcenie liniowe można jednoznacznie określić zadając jego wartości na dowolnej bazie. Niech więc ψ będzie jedynym takim przekształceniem liniowym, że dla dowolnego j :

$$\psi(\phi(v_j)) = v_j.$$

Twierdzimy, że $\phi \circ \psi = \text{id}$. Istotnie, widzimy, że dla dowolnego j

$$\text{id}(v_j) = v_j,$$

a określenie wartości na bazie definiuje przekształcenie liniowe *jednoznacznie*. Czyli

$$\psi \circ \phi = \text{id}.$$

Podobnie pokazujemy, że $\phi \circ \psi = \text{id}$. □

Ponownie możemy skorzystać z odpowiedniości między mnożeniem macierzy a składaniem przekształceń w ustalonej bazie, żeby wysunąć wniosek na temat algebraicznych własności macierzy.

Wniosek 67. Niech A będzie macierzą izomorfizmu $\phi : V \rightarrow V$ w bazie $\mathcal{A}$ (tzn. niech A będzie macierzą ϕ w bazie $\mathcal{A}$, a ϕ niech będzie izomorfizmem.) Wówczas istnieje macierz odwrotna do A .

Dowód. Wystarczy zauważyć, że macierz przekształcenia odwrotnego do ϕ w bazie $\mathcal{A}$ spełnia definicję A^{-1} . □

Wiemy już, że w wielu przypadkach można odwracać macierze. Nie jest wciąż jasne, jak to robić. Aby dokonywać rachunków na macierzach, będziemy więc musieli ustalić dwie kwestie:

1. Jak sprawdzić, czy macierz A jest macierzą izomorfizmu w pewnej bazie?
2. Jeśli jest, to jak znaleźć macierz odwrotną?

Odpowiedź na oba pytania okazuje się prosta. W tym celu będziemy musieli powrócić do schodkowania macierzy. Zaczniemy od prostej obserwacji, że operacje elementarne same odpowiadają mnożeniu przez pewną macierz.

Stwierdzenie 68. *Dla każdej z trzech operacji elementarnych i dowolnego n istnieją takie macierze Q , że dla dowolnego k i dowolnej macierzy A rozmiaru $n \times k$ macierz QA jest równa macierzy A'*

Dowód. Pokażemy, że istnieje macierz E_{kl} , taka, że dla dowolnej macierzy A rozmiaru $n \times n$ macierz $E_{kl}A$ powstaje z macierzy A przez dodanie k -tego wiersza do l -tego wiersza.

Ustalmy dowolną bazę $v_1, \dots, v_n$ w przestrzeni $\mathbb{R}^n$. Niech $v_k^* : \mathbb{R}^n \rightarrow \mathbb{R}$ będzie zdefiniowane w następujący sposób: $v_k^*(v)$ to k -ta współrzędna wektora v we wspomnianej bazie. Niech $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ będzie odwzorowaniem określonym w następujący sposób:

$$\phi(v) = v_k^*(v)v_l.$$

Niech ψ będzie dowolnym innym odwzorowaniem $\mathbb{R}^n \rightarrow \mathbb{R}^n$. Zauważmy, że $\phi \circ \psi(v_j)$ to wektor $v_k^*(\psi(v_j))v_l$, czyli wektor, którego l -ta współrzędna to k -ta współrzędna wektora $\psi(v_j)$, a wszystkie pozostałe współrzędne są zerowe. Na mocy definicji macierzy przekształcenia znaczy to, że macierz $\phi \circ \psi$ w bazie $v_1, \dots, v_n$ to macierz, która jedyne niezerowe wyrazy ma w l -tym wierszu, który stanowił k -ty wiersz macierzy przekształcenia ψ .

Stwierdzamy więc, że dla dowolnego przekształcenia ψ macierz przekształcenia

$$(\text{id} + \phi) \circ \psi = \psi + \phi \circ \psi$$

to macierz przekształcenia ψ z k -tym wierszem dodanym do l -tego wiersza. Jeśli więc zdefiniujemy E_{kl} jako macierz odwzorowania $(\text{id} + \phi)$ w bazie $v_1, \dots, v_n$, to stwierdzamy, że E_{kl} ma żądaną własność.

Dowód istnienia macierzy pozostałych operacji elementarnych jest podobny i pozostawiamy go jako ćwiczenie. $\square$

Szukając metody odwracania macierzy skorzystamy z następującego stwierdzenia.

Stwierdzenie 69. *Dla dowolnej macierzy A liczba liniowo niezależnych kolumn tej macierzy jest równa liczbie liniowo niezależnych wierszy tej macierzy.*

Z powyższego stwierdzenia możemy wywnioskować, że istnieje prosty test, który pozwala ustalić, czy dana macierz jest macierzą izomorfizmu.

Stwierdzenie 70. *Niech A będzie dowolną macierzą $n \times n$. Wówczas A jest macierzą izomorfizmu dokładnie wtedy, gdy jej postać schodkowa zredukowana to macierz identyczności I .*

Dowód. Każdy izomorfizm przeprowadza bazę na bazę, a zatem wartości wektorów dowolnej bazy przy izomorfizmie stanowią układ liniowo niezależny. Skoro A jest macierzą pewnego izomorfizmu ϕ w bazie $v_1, \dots, v_n$, to kolumny macierzy A muszą być liniowo niezależne, gdyż są to po prostu współrzędne wektorów $\phi(v_j)$.

Skoro liczba liniowo niezależnych wierszy jest taka, jak liczba niezależnych kolumn, to macierz A ma n liniowo niezależnych kolumn, a zatem w postaci schodkowej zredukowanej ma postać I . $\square$

Łatwo pokazać, że macierze, które mają odwrotność, to dokładnie macierze izomorfizmów.

Definicja 71. Macierz A rozmiaru $n \times n$ nazwiemy **odwracalną**, gdy istnieje macierz A^{-1} .

Zauważmy, że bez wprowadzania stwierdzenia 70 moglibyśmy podać równie proste kryterium ustalania, czy dana macierz jest odwracalna: wystarczyłoby ustalić, czy można daną macierz doprowadzić do postaci I operacjami elementarnymi na kolumnach. Podaliśmy to stwierdzenie, gdyż w praktyce operacje elementarne przeprowadzamy raczej na wierszach (choć bez szczególnie głębokich powodów.)

Jesteśmy gotowi podać metodę odwracania macierzy: założmy, że A jest dowolną macierzą $n \times n$. Doprowadzamy ją operacjami elementarnymi do postaci I . Niech $Q_1, \dots, Q_N$ oznaczają macierze kolejnych operacji elementarnych. Na mocy związku między mnożeniem przez macierze operacji elementarnych a stosowaniem tych operacji stwierdzamy, że

$$Q_N Q_{N-1} \cdots Q_1 A = I.$$

Ale znaczy to dokładnie, że $Q_N Q_{N-1} \cdots Q_1$ jest szukaną macierzą A^{-1} . Zauważmy ponadto, że

$$Q_N Q_{N-1} \cdots Q_1 = Q_N Q_{N-1} \cdots Q_1 I,$$

ale prawa strona powyższej równości to zarazem efekt zastosowania kolejno operacji elementarnych odpowiadających macierzom $Q_1, Q_2, \dots, Q_N$ do macierzy I . Czyli wystarczy sprowadzić macierz A do postaci schodkowej zredukowanej, a równocześnie stosować te same operacje do macierzy I . Gdy A zostanie sprowadzone do postaci I , macierz I zostanie sprowadzona do postaci A^{-1} .

Przykład 72. Odwróćmy macierz

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

Wypiszmy powyższą macierz oraz macierz jednostkową I . W celu łatwiejszego śledzenia wykonywanych operacji elementarnych wypiszemy je obok siebie:

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

Odejmujemy trzykrotność pierwszego wiersza od drugiego uzyskując:

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

Dzielimy drugi wiersz przez (-2) . Otrzymujemy:

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

Odejmujemy dwukrotność drugiego wiersza od pierwszego:

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

Sprowadziliśmy naszą macierz do macierzy jednostkowej I . Stwierdzamy zatem, że efekt zastosowania tych samych operacji elementarnych do macierzy I , czyli szukana przez nas macierz odwrotna, to:

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

Zadania

Zadanie 32. Opisz wszystkie rozwiązania układu równań:

$$\begin{cases} x_1 + 17x_2 - 4x_3 - x_4 = 1 \\ 4x_1 - 6x_3 = 6 \\ 2x_2 + 8x_4 = 4 \end{cases}$$

Zadanie 33. Znajdź macierze odwrotne do następujących macierzy:

$$1. \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$$

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

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

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

$$5. \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

$$6. \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & 1 \end{pmatrix}$$

$$7. \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$8. \begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Zadanie 34. Udowodnij, że dla dowolnych macierzy A, B rozmiaru $n \times n$ zachodzi wzór:

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Zadanie 35. Udowodnij, że macierz A jest odwracalna wtedy i tylko wtedy, gdy jest macierzą izomorfizmu w pewnej bazie.

Zadanie 36. Udowodnij, że dla dowolnej macierzy odwracalnej A i dowolnej macierzy C istnieje taka macierz B , że

$$AB = C.$$

Zadanie 37. Udowodnij, że istnieją macierze wszystkich operacji elementarnych. Znajdź jawny opis macierzy operacji elementarnych.

Zajęcia 16. XI

Wiemy już, jak ustaliwszy bazę przypisać odwzorowaniu liniowemu macierz. Często (na ogół) bardziej niż macierz interesuje nas samo odwzorowanie, chcielibyśmy więc zrozumieć zależność między macierzami przekształceń a rozmaitymi bazami. Zauważmy na początek, że gdy znamy współrzędne wektora w jednej bazie, to znalezienie jego współrzędnych w innej bazie sprowadza się do przemnożenia kolumny współrzędnych tego wektora przez pewną zdefiniowaną już przez nas macierz.

Definicja 73. Macierzą zmiany bazy (zamiany współrzędnych) od bazy $\mathcal{A}$ do $\mathcal{B}$ nazywamy macierz $M(\text{id})_{\mathcal{A}}^{\mathcal{B}}$.

Zanim uzasadnimy, czemu warto nazywać wprowadzoną powyżej macierz macierzą zmiany bazy przypomnijmy stwierdzenie 36:

Stwierdzenie 74. Niech $\phi : V \rightarrow W$. Niech $\mathcal{A}$ będzie dowolną bazą w V , natomiast $\mathcal{B}$ – dowolną bazą w W . Jeśli A jest macierzą ϕ w bazach $\mathcal{A}, \mathcal{B}$, zaś $\mathbf{v}$ – macierzą współrzędnych v w bazie $\mathcal{A}$, to $A\mathbf{v}$ jest macierzą współrzędnych wektora $\phi(v)$ w bazie $\mathcal{B}$.

Dowód. Niech $\mathcal{A} = (v_1, \dots, v_n), \mathcal{B} = (w_1, \dots, w_n)$ będą dwiema bazami w przestrzeni V . Niech $A = (a_{ij})$ będzie macierzą ϕ w bazach $\mathcal{A}, \mathcal{B}$. Niech v będzie kolumną współrzędnych i -tego wektora bazy $\mathcal{A}$ w tejże bazie, czyli kolumną z jedyną na i -tym miejscu i zerem w pozostałych miejscach.

Na mocy definicji a_{ij} to j -ta współrzędna wektora $\phi(v_i)$ w bazie $\mathcal{B}$. Pomnożenie opisanej wyżej kolumny współrzędnych przez A daje w wyniku macierz składającą się z jednej kolumny, która na j -tym miejscu ma liczbę a_{ij} . Jest to dokładnie kolumna współrzędnych wektora $\phi(v_i)$ w bazie $\mathcal{B}$. Udowodniliśmy więc stwierdzenie dla wektorów bazy $\mathcal{A}$.

Wystarczy zauważyć, że jeśli v jest dowolnym wektorem, to istnieją takie współczynniki a_k , że

$$v = \sum_{k=1}^n a_k v_k.$$

A wówczas

$$Av = \sum_{k=1}^n a_k Av_k,$$

z czego już łatwo wynika nasze stwierdzenie. $\square$

Biorąc w szczególności $\phi = \text{id}$ możemy łatwo przekonać się, jaki jest związek macierzy zmiany bazy ze zmianą bazy.

Wniosek 75. Niech A będzie macierzą zmiany bazy od $\mathcal{A} = (v_1, \dots, v_n)$ do $\mathcal{B} = (w_1, \dots, w_n)$. Wówczas dla dowolnego wektora v jeśli w jest kolumną współrzędnych wektora v w bazie $\mathcal{A}$, to Aw jest kolumną współrzędnych wektora v w bazie $\mathcal{B}$.

Przykład 76. Niech $\mathcal{A} = ((1, 0), (5, 7))$. Niech st oznacza bazę standardową w $\mathbb{R}^2$. Wówczas macierz przejścia z $\mathcal{A}$ do st to

$$\begin{pmatrix} 1 & 5 \\ 0 & 7 \end{pmatrix}.$$

Z drugiej strony $(1, 0) = 1(1, 0) + 0(5, 7)$, a $(0, 1) = -\frac{5}{7}(1, 0) + \frac{1}{7}(5, 7)$, czyli macierz przejścia z st do $\mathcal{A}$ to

$$\begin{pmatrix} 1 & -\frac{5}{7} \\ 0 & \frac{1}{7} \end{pmatrix}.$$

Jeśli znamy macierz przejścia z bazy $\mathcal{A}$ do bazy $\mathcal{B}$ to łatwo też znaleźć macierz przejścia z bazy $\mathcal{B}$ do bazy $\mathcal{A}$. Zanim pokażemy jak to zrobić sformułujemy nieco ogólniejszy fakt.

Stwierdzenie 77. Dla dowolnych $\phi : U \rightarrow V, \psi : V \rightarrow W$ oraz baz $\mathcal{A}, \mathcal{B}, \mathcal{C}$ w przestrzeniach odpowiednio U, V, W mamy:

$$M_{\mathcal{B}}^{\mathcal{C}}(\psi)M_{\mathcal{A}}^{\mathcal{B}}(\phi) = M_{\mathcal{A}}^{\mathcal{C}}(\psi \circ \phi).$$

Dowód. Niech $u \in U$ będzie dowolnym wektorem, a c kolumną jego współrzędnych w bazie $\mathcal{A}$. Wiemy już na mocy stwierdzenia 36, że

$$M_{\mathcal{A}}^{\mathcal{B}}(\phi)c$$

to kolumna współrzędnych wektora $\phi(u)$ w bazie $\mathcal{B}$. Stosując raz jeszcze to samo stwierdzenie widzimy, że

$$M_{\mathcal{B}}^{\mathcal{C}}(\psi)\left(M_{\mathcal{A}}^{\mathcal{B}}(\phi)c\right)$$

jest kolumną współrzędnych wektora $\psi(\phi(u))$ w bazie $\mathcal{C}$. Ale skoro dla dowolnego u oraz kolumny c jego współrzędnych w bazie $\mathcal{A}$ w wyniku mnożenia:

$$M_{\mathcal{B}}^{\mathcal{C}}(\psi)M_{\mathcal{A}}^{\mathcal{B}}(\phi)$$

otrzymujemy kolumnę współrzędnych $\psi \circ \phi(u)$ w bazie $\mathcal{C}$, to stwierdzamy biorąc za u wektory bazowe bazy $\mathcal{A}$, że powyższy iloczyn macierzy jest macierzą $M_{\mathcal{A}}^{\mathcal{C}}(\psi \circ \phi)$. $\square$

Wniosek 78. Dla dowolnych baz $\mathcal{A}, \mathcal{B}$ tej samej przestrzeni mamy:

$$M_{\mathcal{A}}^{\mathcal{B}}(\text{id}) = M_{\mathcal{B}}^{\mathcal{A}}(\text{id})^{-1}.$$

Dowód.

$$M_{\mathcal{A}}^{\mathcal{B}}(\text{id})M_{\mathcal{B}}^{\mathcal{A}}(\text{id}) = M_{\mathcal{B}}^{\mathcal{B}}(\text{id}),$$

przekonaliśmy się już, że dla dowolnej bazy $\mathcal{B}$ macierz identyczności $M_{\mathcal{B}}^{\mathcal{B}}(\text{id})$ to macierz jednostkowa I . Skoro $M_{\mathcal{B}}^{\mathcal{A}}(\text{id})$ pomnożona przez $M_{\mathcal{A}}^{\mathcal{B}}(\text{id})$ daje w wyniku macierz jednostkową I , to z definicji jest równa macierzy $M_{\mathcal{A}}^{\mathcal{B}}(\text{id})^{-1}$. $\square$

Przykład 79. Niech $\mathcal{A} = ((1, 2, 3), (4, 5, 6), (0, 1, 1))$. Znajdziemy macierz przejścia z bazy st do bazy $\mathcal{A}$, czyli znajdziemy macierz, która w kolumnach ma współrzędne wektorów bazowych st w bazie $\mathcal{A}$. Aby to zrobić wypisujemy najpierw łatwo macierz $M_{\mathcal{A}}^{\text{st}}(\text{id})$:

$$\begin{pmatrix} 1 & 4 & 0 \\ 2 & 5 & 1 \\ 3 & 6 & 1 \end{pmatrix}.$$

Szukana macierz $M_{\text{st}}^{\mathcal{A}}(\text{id})$ to odwrotność powyższej macierzy, czyli

$$\begin{pmatrix} -3 & 4 & -\frac{4}{3} \\ 1 & -1 & \frac{1}{3} \\ 1 & 2 & 1 \end{pmatrix}.$$

Przypuśćmy, że znamy macierz odwzorowania ϕ w pewnej bazie i chcemy poznać jego macierz w pewnej innej bazie. Możemy to łatwo zrobić opierając się na dotychczasowych rozważaniach.

Stwierdzenie 80. Dla dowolnego $\phi : V \rightarrow V$ oraz dowolnych baz $\mathcal{A}, \mathcal{B}$ przestrzeni V zachodzi równość:

$$\begin{aligned} M_{\mathcal{B}}^{\mathcal{B}}(\phi) &= M_{\mathcal{A}}^{\mathcal{B}}(\text{id})M_{\mathcal{A}}^{\mathcal{A}}(\phi)M_{\mathcal{B}}^{\mathcal{A}}(\text{id}) \\ &= M_{\mathcal{A}}^{\mathcal{B}}(\text{id})M_{\mathcal{A}}^{\mathcal{A}}(\phi)M_{\mathcal{A}}^{\mathcal{B}}(\text{id})^{-1}. \end{aligned}$$

Dowód. Jest to bezpośredni wniosek ze stwierdzenia 77. $\square$

Przykład 81. Niech ϕ oznacza obrót o $\frac{\pi}{6}$ wokół prostej $\text{lin}((2, 1, 0))$ w $\mathbb{R}^3$. Znajdziemy macierz ϕ w bazie st czyli $M_{\text{st}}^{\text{st}}(\phi)$.

Rozszerzmy $((2, 1, 0))$ do bazy, w której wszystkie wektory są prostopadłe, na przykład do $\mathcal{A} = ((2, 1, 0), (-1, 2, 0), (0, 0, 1))$. Powtarzając geometryczne rozumowanie z przykładu 32 stwierdzamy, że macierz ϕ w bazie $\mathcal{A}$ to po prostu:

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

Wiemy też, że

$$M_{\mathcal{A}}^{\text{st}}(\text{id}) = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Odwracając wypisaną wyżej macierz wnioskujemy, że:

$$M_{\text{st}}^{\mathcal{A}}(\text{id}) = \begin{pmatrix} \frac{2}{5} & \frac{1}{5} & 0 \\ -\frac{1}{5} & \frac{2}{5} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

A zatem wnioskujemy ostatecznie, że:

$$M_{\text{st}}^{\text{st}}(\phi) = \begin{pmatrix} \frac{2}{5} & \frac{1}{5} & 0 \\ -\frac{1}{5} & \frac{2}{5} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} 2 & -1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Zadania

Zadanie 38. Czy dla dowolnych $\phi, \psi : V \rightarrow V$ i dowolnych baz $\mathcal{A}, \mathcal{B}$ przestrzeni V zachodzi wzór: $M_{\mathcal{A}}^{\mathcal{B}}(\phi) = M_{\mathcal{B}}^{\mathcal{A}}(\phi)^{-1}$?

Zadanie 39. Znajdź współrzędne wektora $(1, 0, 0)$ w bazach:

1. $(1, 2, 3), (4, 5, 8), (1, 3, 1)$.
2. $(0, 1, 1), (1, 0, 1), (1, 1, 0)$.

Zadanie 40. Znajdź macierz przejścia z bazy $\mathcal{A}$ do bazy $\mathcal{B}$, gdzie:

1. $\mathcal{A} = ((1, 1, 0), (1, 0, -1), (0, -1, 1)), \mathcal{B} = ((2, 3, 4), (0, 5, 6), (1, 2, 3))$.
2. $\mathcal{A} = ((1, 0, 0), (0, 1, 0), (-2, 0, 1)), \mathcal{B} = ((1, 1, 1), (1, 0, 1), (1, 2, 2))$.

Zadanie 41. Znajdź macierz obrotu o $\frac{\pi}{3}$ w $\mathbb{R}^2$ w bazie $((1, 0), (0, 3))$ a następnie w bazie $((1, 1), (3, -3))$.

Zadanie 42. Wiedząc, że wektory $(1, 1, 0, 0), (1, -1, 0, 0), (0, 0, 1, 1), (0, 0, 1, -1)$ są wzajemnie prostopadłe znajdź macierz obrotu o $\frac{\pi}{4}$ w przestrzeni $\mathbb{R}^4$ wokół prostej $\text{lin}((0, 0, 1, 1))$ bazy standardowej.

***Zadanie 43.** Załóżmy, że $\phi : V \rightarrow V$ jest przekształceniem, które ma taką samą macierz w każdej bazie (tzn. $M_{\mathcal{A}}^{\mathcal{A}}(\phi)$ nie zależy od wyboru bazy $\mathcal{A}$). Udowodnij, że wówczas $\phi = a \cdot \text{id}$ dla pewnego $a \in \mathbb{R}$.

3 Wyznaczniki

3.1 Zajęcia 23. XI

Zacniemy tym razem od krótkiej gawędy, która (mam nadzieję) pomoże lepiej zrozumieć geometrię stojącą za wprowadzanymi w tym rozdziale pojęciami. Zachęcamy do sporządzania rysunków w trakcie lektury.

Uwaga 82. Sposób podejścia do problemu, który tu prezentujemy, a który jest kluczowy dla zrozumienia metod algebry liniowej, jest w pewnym sensie przeciwny do "szkolnego" sposobu patrzenia. "W szkole" najpierw rysujemy sobie układ współrzędnych, a potem badamy figury, za pomocą współrzędnych, które zadaje nasz układ. W podejściu, które tu prezentujemy wychodzimy od abstrakcyjnych pojęć "wektora", "równoległoboku" itp., *których definicja jest niezależna od tego jak dobierzemy nasze współrzędne*. Dobierając na różne sposoby bazę naszej przestrzeni wybieramy "sposób patrzenia" na te obiekty.

Równoległościanem w $\mathbb{R}^n$ nazywamy wielościan, który ma n par równoległych ścian. (W $\mathbb{R}^2$ to po prostu równoległobok, a co to dokładniej znaczy w $\mathbb{R}^n$, że dwie płaszczyzny $n - 1$ wymiarowe są równoległe pozostawimy do zastanowienia.) Formalnie: jeśli mamy zadanych n liniowo niezależnych wektorów $v_1, \dots, v_n$ w przestrzeni $\mathbb{R}^n$, to równoległościanem nazywamy zbiór wektorów zadanych przez sumy postaci

$$\sum_{i=1}^n t_i v_i$$

dla wszystkich t_i takich, że $0 \leq t_i \leq 1$. Naszym celem jest znalezienie wzoru na liczenie objętości równoległościanów o dobrych własnościach algebraicznych. Zaczniemy od przypadku przestrzeni jednowymiarowej, czyli prostej $\mathbb{R}$ (równoległościany w $\mathbb{R}$ to po prostu odcinki). Każdy punkt $w \in \mathbb{R}$ wyznacza odcinek pomiędzy punktem 0 a w . Chcielibyśmy żeby długością takiego odcinka była "odległość" pomiędzy 0 a w , w szczególności powinna to być jakaś liczba (biorąc po prostu w dostaniemy wektor). Możemy zatem ustalić jakąś bazę $\mathbb{R}$, tzn. niezerowy wektor, i popatrzeć na współrzędną w w tej bazie. Niektóre wybory bazy przyporządkują jednak w ujemną współrzędną (np., jeśli jako bazę weźmiemy $-w$, to w w tej bazie będzie miał współrzędną -1). Moglibyśmy zastosować rozwiązanie brutalne i po prostu wziąć moduł (wartość bezwzględna) współrzędnej wektora w w danej bazie. Zamiast tego, zastanówmy się jaką informację niesie ze sobą ujemny znak długości takiego odcinka. Zauważmy, że to czy w dostanie ujemną, czy dodatnią współrzędną zależy od tego, czy wybierzemy wektor o tym samym znaku, co w , czy przeciwnym. Przyjmijmy ponadto na potrzeby tego przykładu następującą definicję: wektory w, v na prostej wskazują w tę samą stronę jeśli mają ten sam znak (jako liczby rzeczywiste).

Wychodzi na to, że na prostej w wskazuje w tę samą stronę, co v wtedy i tylko wtedy, gdy odcinek wyznaczany przez w ma w bazie v dodatnią długość. Dokładnie po to będzie nam dopuszczenie możliwości ujemnej objętości: poza informacją ile miejsca zajmuje dany obiekt w przestrzeni będziemy chcieli wiedzieć *jak na niego patrzymy*. Na prostej odpowiada to wskazaniu, w którą stronę jest "prawo" (wybraniu zwrotu na osi współrzędnych). Ponadto wzięcie po prostu współrzędnej wektora w w danej bazie, a nie modułu tej współrzędnej, ma tę zaletę, że przyporządkowanie wektorowi jego objętości jest *przekształceniem liniowym*.

Zanim przejdziemy do sytuacji z $\mathbb{R}^2$ zastanówmy się jak można prostym wzorem opisać pole równoległoboku wyznaczanego przez liniowo niezależne wektory v_1, v_2 . Nietrudno się przekonać, że jeśli, w bazie standardowej (e_1, e_2) , v_1 leży "na prawo" od v_2 (wyobrażamy sobie, że stoimy w początku układu współrzędnych i obydwa wektory są przed nami) i wektory te mają współrzędne (x_1, y_1) i (x_2, y_2) odpowiednio, to pole naszego równoległoboku jest równe

$$x_1 y_2 - x_2 y_1 \quad (3)$$

Oznaczmy powyższą liczbę przez v . Jeśli wykonamy tę samą operację na tych samych wektorach, ale zapisanych w bazie (e_2, e_1) , to otrzymamy

$$x_2 y_1 - x_1 y_2$$

a zatem $-v$. Oczywiście jedna z tych liczb musi być ujemna. Gdybyśmy jednak umówili się, że pierwszy wektor w naszej bazie wskazuje "gdzie jest prawo" (a zatem i gdzie jest lewo, bo to w przeciwnym kierunku), a drugi, "gdzie jest góra" (a zatem i gdzie jest dół), to zauważymy, że jako prawy wektor musielibyśmy wziąć v_2 , którego współrzędne w tej bazie wynoszą (y_2, x_2) , a zatem obliczając pole dostalibyśmy znów liczbę v . Podobnie jak w wypadku $\mathbb{R}$ zauważyliśmy, że to czy dana para wektorów wyznacza tak samo kierunki, jak druga ma związek z tym, czy pola równoległoboków wychodzą dodatnie, czy ujemne. To, co zmieniliśmy zamieniając wektory miejscami, nazwiemy *orientacją* naszej przestrzeni liniowej. W dalszej części powiemy, że dwa układy wektorów $v_1, \dots, v_n, w_1, \dots, w_n$ wyznaczają tę samą orientację przestrzeni, jeśli pole równoległościanu wyznaczanego przez $v_1, \dots, v_n$ w bazie $w_1, \dots, w_n$ jest dodatnie. Zanim to zrobimy, musimy jednak zobaczyć jak wzór (3) uogólnia się na wyższe wymiary. Do tego potrzebujemy jeszcze drobnego przygotowania.

3.1.1 Permutacje

Definicja 83. Ustalmy zbiór $[n] := \{1, \dots, n\}$. **Permutacją** $[n]$ będziemy nazywać dowolną bijekcję tego zbioru. Zbiór permutacji $[n]$ oznaczamy przez S_n .

Jeśli σ jest permutacją $[n]$, taką, że dla pewnych $k < i < n$

$$\begin{aligned}\sigma(k) &= i \\ \sigma(i) &= k\end{aligned}$$

i dla pozostałych $l \in [n]$ mamy $\sigma(l) = l$, to σ nazywamy **transpozycją**. Transpozycję elementów k, i zapisujemy w skrócie (k, i) . Przez id oznaczamy permutację identycznościową.

Przykład 84. Rozpatrzmy permutacje zbioru $[3]$. (Oczywiście jest ich 6). Używając notacji z powyższej definicji łatwo wskazać cztery z nich:

1. id
2. $\sigma_2 = (1, 2)$ - permutacja zamieniająca 1 i 2, a trzymająca 3 w miejscu.
3. $\sigma_3 = (2, 3)$ - permutacja zamieniająca 2 i 3, a trzymająca 1 w miejscu.
4. $\sigma_4 = (1, 3)$ - permutacja zamieniająca 1 i 3, a trzymająca 2 w miejscu.

Pozostałe permutacje, to *cykle trójelementowe*

$$\begin{aligned}\sigma_5(1) &= 2 \\ \sigma_5(2) &= 3 \\ \sigma_5(3) &= 1\end{aligned}$$

oraz

$$\begin{aligned}\sigma_6(1) &= 3 \\ \sigma_6(2) &= 1 \\ \sigma_6(3) &= 2\end{aligned}$$

σ_5 i σ_6 wygodnie jest zapisać w postaci ciągów, odpowiednio $(1, 2, 3)$ oraz $(1, 3, 2)$ (pierwszy element przechodzi na drugi, drugi, na trzeci, a trzeci na pierwszy, oczywiście można zapisać tak dowolnie długie cykle). Zauważmy, że σ_5 jest złożeniem dwóch transpozycji, mianowicie

$$(1, 2, 3) = (3, 2) \circ (1, 3)$$

Podobnie σ_6

$$(1, 3, 2) = (2, 3) \circ (1, 2)$$

Sprawdzenie tego pozostawimy jako łatwe ćwiczenie na oswojenie się z notacją.

W powyższym przykładzie widzieliśmy, że każdy element możemy zapisać jako złożenie pewnej liczby transpozycji. Da się z niego wydobyć zresztą

ogólną metodę jak znajdować taki rozkład: weźmy cykl, $(1, 2, 3)$, dla którego takiego rozkładu nie widać od razu. Cykl ten przesuwając 3 na 1. Składając go z transpozycją $(3, 1)$ z prawej strony widzimy, że

$$(1, 2, 3) \circ (3, 1) = (2, 3)$$

Poniższe stwierdzenie pokazuje, że ta sytuacja uogólnia się na dowolne permutacje. Na jego dowód składa się powyższy pomysł + indukcja.

Stwierdzenie 85. *Dowolną permutację σ zbioru $[n]$ można przedstawić jako złożenie pewnej liczby transpozycji.*

Dowód. Dowodzimy przez indukcję ze względu na $[n]$. Dla $n = 1$ (podobnie jak dla $n = 2, = 3$) zarówno istnienie takiego przedstawienia, jak i jednoznaczność liczby czynników w rozkładzie są oczywiste. Przypuśćmy, że $n > 1$ i twierdzenie prawdziwe jest dla $k < n$. Ustalmy dowolną permutację σ zbioru $[n]$. Jeśli σ jest identycznością, to jest złożeniem transpozycji. Jeśli nie jest identycznością, to istnieją $m, l \in [n]$ takie, że

$$\sigma(m) = l \neq m$$

Wtedy permutacja $\sigma \circ (m, l)$ spełnia

$$\sigma \circ (m, l)(l) = l$$

zatem możemy ją traktować jako permutację zbioru $n - 1$ elementów. Z założenia indukcyjnego $\sigma \circ (m, l)$ możemy przedstawić jako złożenie transpozycji. Zatem σ też da się przedstawić jako złożenie transpozycji. $\square$

Oczywiście rozkład permutacji na transpozycje nie jest jednoznaczny, nawet z dokładnością do liczby czynników. W dowolnym miejscu takiego rozkładu możemy zawsze wstawić dwa razy tę samą transpozycję i otrzymać zapis tej samej permutacji z większą liczbą transpozycji.

Przykład 86. $(1, 2, 3) = (1, 3)(2, 3) = (1, 3)(2, 3)(1, 2)(1, 2) = (1, 3)(1, 2)(1, 2)(2, 3)$

W powyższym przykładzie liczba czynników w różnych rozkładach będzie zawsze parzysta. Jak się okazuje parzystość liczby transpozycji w rozkładzie jest stała dla danej permutacji. Dowód tego stwierdzenia pominiemy (dociekliwi mogą go znaleźć [na tej stronie](#) w wykładzie "Permutacje i podziały" pod postacią Obserwacji 6.7.).

Stwierdzenie 87. *Jeśli $\tau_0 \circ \dots \circ \tau_n$ i $\tau'_0 \circ \dots \circ \tau'_k$ są dwoma rozkładami tej samej permutacji na transpozycje, to $n - k$ jest parzysta.*

Czyli parzystość liczby czynników w rozkładach danej permutacji na transpozycje jest stała. Powyższe stwierdzenie było nam potrzebne żeby uzasadnić poprawność poniższej definicji

Definicja 88. Permutacja σ jest **parzysta** jeśli da się przedstawić jako złożenie parzystej liczby transpozycji. W przeciwnym przypadku permutacja σ jest **nieparzysta**.

Uwaga 89. Niech $\sigma \in S_n$ i niech σ^{-1} oznacza permutację odwrotną do σ . Zauważmy, że σ jest parzysta wtedy i tylko wtedy, gdy σ^{-1} jest parzysta.

3.2 Permutacyjna definicja wyznacznika

Zauważmy, że zamiast mówić, o przypisywaniu pola jakiemuś równoległoscianowi, możemy mówić o przypisywaniu liczby pewnej *macierzy*. Faktycznie, widzieliśmy, że w $\mathbb{R}^n$, n liniowo niezależnych wektorów wyznacza równoległoscian i dowolny równoległoscian o niezerowym polu jest wyznaczany przez n liniowo niezależnych wektorów. Możemy zatem myśleć o n -wymiarowym równoległoscianie jako o macierzy $n \times n$, której kolumny to wektory wyznaczające nasz równoległoscian. To podejście będzie miało dodatkowo jeszcze tę zaletę, że macierze mogą reprezentować inne obiekty (np., co już dobrze wiemy, przekształcenia liniowe).

Przykład 90. Równoległoscian w $\mathbb{R}^3$ wyznaczany przez wektory $(1, 2, 3)$, $(0, 2, 0)$, $(3, 2, 0)$ reprezentujemy w macierzy jako

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

Pamiętajmy, że kolejność w jakiej podajemy wektory ma znaczenie (co więcej: ma mieć).

Zauważmy, że w $\mathbb{R}^2$ zoorientowana objętość równoległoboku zadanego macierzą

$$\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$$

wynosiła $a_{1,1}a_{2,2} + (-1)a_{2,1}a_{1,2}$. W obydwu składnikach tej sumy bierzemy po dokładnie jednym wyrazie z każdego wiersza. Ponadto, jeśli popatrzymy na pary współczynników, w różnych składnikach to wyznaczają one permutacje zbioru dwuelementowego: w pierwszym składniku widzimy identyczność (pary $(1, 1)$, $(2, 2)$), w drugim transpozycję (pary $(1, 2)$, $(2, 1)$). Ponadto, każda permutacja zbioru dwuelementowego jest reprezentowana, a (-1) dostawiliśmy

przy składniku, który reprezentuje permutację nieparzystą. Poniższa definicja uogólnia zatem wyprowadzony w przypadku $\mathbb{R}^2$ wzór na wyższe wymiary. Przypomnijmy, że S_n to zbiór wszystkich permutacji zbioru n -elementowego.

Definicja 91. Wyznacznikiem macierzy $A = (a_{i,j})_{i \leq n, j \leq n}$ nazywamy liczbę

$$\sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)} \quad (\det)$$

gdzie

$$\varepsilon(\sigma) = \begin{cases} 1 & \text{jeśli } \sigma \text{ jest parzysta} \\ -1 & \text{w.p.p.} \end{cases}$$

Wyznacznik A oznaczamy przez $\det A$.

Przykład 92. Zastosujemy wzór $(\det)$ do obliczenia wyznacznika macierzy

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}$$

W prawej kolumnie podpisujemy od jakiej permutacji pochodzi dany składnik

$$\begin{array}{ll} \det A = a_{1,1}a_{2,2}a_{3,3} & \text{id} \\ - a_{1,1}a_{2,3}a_{3,2} & (2, 3) \\ - a_{1,3}a_{2,2}a_{3,1} & (1, 3) \\ - a_{1,2}a_{2,1}a_{3,3} & (1, 2) \\ + a_{1,2}a_{2,3}a_{3,1} & (1, 2, 3) \\ + a_{1,3}a_{2,1}a_{3,2} & (1, 3, 2) \end{array}$$

Przykład 93. Niech id oznacza macierz identyczności $\mathbb{R}^n$ w bazach standardowych (tzn. macierz, która na przekątnej ma 1, a na pozostałych miejscach 0.) Wtedy $\det \text{id} = 1$. Faktycznie, we wzorze $(\det)$ jedynym niezerowym składnikiem jest ten pochodzący od permutacji identycznościowej i oczywiście $\varepsilon(\text{id}) = 1$, a także $a_{1,1} = \dots = a_{n,n} = 1$. Jak myślimy o wyznaczniku jako o zoorientowanej objętości równoległościanu, to udowodniliśmy właśnie, że objętość n -wymiarowej kostki o krawędzi 1, zoorientowanej zgodnie z bazą standardową, wynosi 1.

Wprowadźmy jedno oznaczenie: niech $v^1, \dots, v^n$ będą n -tkami liczb rzeczywistych (myślimy o nich jako o współrzędnych wektorów z $\mathbb{R}^n$ w pewnej bazie). k -tą współrzędną wektora v^i oznaczamy przez v_k^i . Macierz o wierszach $v^1, \dots, v^n$ będziemy oznaczać przez $M(v^1, \dots, v^n)$. Zauważmy, że wyraz w i -tym wierszu i j -tej kolumnie takiej macierzy jest równy dokładnie v_j^i . Możemy się teraz zapytać czy nasza konwencja żeby wektory ustawiać w wierszach

macierzy, a nie w kolumnach jest w jakimkolwiek sensie istotna dla wartości wyznacznika. Okazuje się, że tak nie jest.

Definicja 94 (Macierz transponowana). Niech $A = (a_{i,j})$ i $B = (b_{i,j})$ będą macierzami $n \times n$. B nazywamy macierzą transponowaną do A jeśli dla dowolnych $i \leq n, j \leq n$

$$b_{i,j} = a_{j,i}$$

Macierz transponowaną do A oznaczamy przez A^T .

Przykład 95. Niech

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

Wtedy

$$A^T = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}$$

Stwierdzenie 96. Dla dowolnej macierzy A rozmiaru $n \times n$, $\det A = \det A^T$.

Dowód. Niech $A = (a_{i,j})$, $A^T = (b_{i,j})$. Pokażemy, że każdy składnik sumy opisującej $\det A$ jest składnikiem sumy opisującej $\det A^T$ i na odwrót. Przypuśćmy, że

$$\varepsilon(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$$

jest w rozwinięciu $\det A$. Niech σ^{-1} będzie permutacją odwrotną do σ . Zauważmy, że, przestawiając ewentualnie czynniki, powyższy iloczyn możemy zapisać też

$$\varepsilon(\sigma) a_{\sigma^{-1}(1),1} \cdots a_{\sigma^{-1}(n),n}$$

Ponadto dla dowolnego i , $a_{\sigma^{-1}(i),i} = b_{i,\sigma^{-1}(i)}$ a także $\varepsilon(\sigma) = \varepsilon(\sigma^{-1})$ (na mocy Uwagi 89), stąd powyższy składnik jest w rozwinięciu $\det A^T$. W drugą stronę rozumiemy analogicznie. $\square$

Mówiąc nie do końca precyzyjnie, powyższe stwierdzenie gwarantuje nam, że jeśli wyznacznik ma jakąś własność ze względu na wiersze macierzy A , to ma ją też ze względu na kolumny A .

Stwierdzenie 97. Wyznacznik jest funkcją wierszy (kolumn) macierzy, liniową ze względu na każdy argument.

Dowód. Musimy pokazać, że dla dowolnych $\alpha, \beta \in \mathbb{R}$ oraz $v^1, \dots, v^i, w^i, v^{i+1}, v^n$ mamy

$$\begin{aligned} \det M(v^1, \dots, \alpha v^i + \beta w^i, v^{i+1}, v^n) &= \\ &= \alpha \det M(v^1, \dots, v^i, v^{i+1}, v^n) + \beta \det M(v^1, \dots, w^i, v^{i+1}, v^n) \quad (*) \end{aligned}$$

Jak spojrzymy na wzór (det), zobaczymy, że żeby udowodnić powyższą równość trzeba po prostu rozbić każdy ze składników dużej sumy na sumę w której jednym składniku występuje αv^i i, a w drugim βw^i . Następnie pogrupować składniki, w których występuje α i wyciągnąć α przed sumę. Zobaczymy przykład dla $n = 3, i = 2$: każdy składnik rozwinięcia wyznacznika po lewej stronie (*) jest postaci

$$\varepsilon(\sigma)v_{\sigma(1)}^1(\alpha v_{\sigma(2)}^2 + \beta w_{\sigma(2)}^2)v_{\sigma(3)}^3$$

Powyższe wyrażenie jest zaś równe

$$\alpha\varepsilon(\sigma)v_{\sigma(1)}^1v_{\sigma(2)}^2v_{\sigma(3)}^3 + \beta\varepsilon(\sigma)v_{\sigma(1)}^1w_{\sigma(2)}^2v_{\sigma(3)}^3$$

Sumując po wszystkich permutacjach i grupując wyrazy z α w jedną sumę, a wyrazy z β w drugą dochodzimy do wniosku, że powyższe jest równe:

$$\alpha \sum_{\sigma \in S_3} \varepsilon(\sigma)v_{\sigma(1)}^1v_{\sigma(2)}^2v_{\sigma(3)}^3 + \beta \sum_{\sigma \in S_3} \varepsilon(\sigma)v_{\sigma(1)}^1w_{\sigma(2)}^2v_{\sigma(3)}^3$$

Powyższe to zaś na mocy definicji

$$\alpha \det M(v^1, v^2, v^3) + \beta \det M(v^1, w^2, v^3)$$

□

Wypowiadając poniższe twierdzenie po ludzku: jeśli w macierzy A zamienimy dwa wiersze miejscami, to wyznacznik zmieni znak. Taką własność *zorientowanej objętości* zaobserwowaliśmy już na przykładzie równoległoboku w $\mathbb{R}^2$.

Stwierdzenie 98. Wyznacznik jest funkcją antysymetryczną wierszy (kolumn) macierzy, tzn. jeśli τ jest transpozycją $[n]$, to

$$\det M(v^1, \dots, v^n) = -\det M(v^{\tau(1)}, \dots, v^{\tau(n)})$$

Dowód. Dowód jest analogiczny do dowodu powyższych stwierdzeń i korzysta z faktu, że jeśli τ jest transpozycją, to dla dowolnej permutacji σ

$$\varepsilon(\sigma \circ \tau) = \varepsilon(\tau \circ \sigma) = -\varepsilon(\sigma)$$

□

Wniosek 99. Jeśli $v_1, \dots, v_n$ jest liniowo zależnym układem wektorów, to

$$\det M(v_1, \dots, v_n) = 0.$$

Dowód. Zauważmy, że z antysymetrii wyznacznika, o której mówiło Stwierdzenie 98, wynika, że dla dowolnych $v_1, \dots, v_{n-1}$ mamy

$$\det M(v_1, \dots, v_{n-1}, v_1) = 0$$

Faktycznie, zamieniając ze sobą pierwszy i ostatni wiersz $M(v_1, \dots, v_{n-1}, v_1)$ dostajemy tę samą macierz, stąd musi być

$$\det M(v_1, \dots, v_{n-1}, v_1) = -\det M(v_1, \dots, v_{n-1}, v_1)$$

a to jest prawdą tylko jeśli $\det M(v_1, \dots, v_{n-1}, v_1) = 0$. Resztę szczegółów pozostawiamy jako ćwiczenie (należy skorzystać z liniowości wyznacznika). $\square$

Poniższe stwierdzenie jest jedną z najważniejszych własności wyznacznika. Pozostawimy je bez dowodu.

Twierdzenie 100 (Cauchy). *Niech A, B będą dwoma macierzami $n \times n$. Wtedy*

$$\det(AB) = \det(A) \det(B)$$

Poniższy wniosek będzie bardzo ważny w kontekście definicji, tego, co to znaczy, że dwie bazy danej przestrzeni n -wymiarowej określają tę samą orientację tej przestrzeni. Na początku tego rozdziału stwierdziliśmy, że w $\mathbb{R}^2$ możemy powiązać to ze znakiem objętości równoległoboku wyznaczanego przez jeden układ względem bazy wyznaczonej przez drugi układ. Zauważmy, że w opisanej sytuacji macierz takiego równoległościanu to po prostu macierz zmiany bazy. Czy może się zdarzyć, że wyznacznik macierzy zmiany bazy będzie równy 0? Poniższy wniosek z Twierdzenia Cauchy'ego wyklucza taką sytuację.

Wniosek 101. *Niech A będzie macierzą $n \times n$. Wtedy A jest macierzą izomorfizmu wtedy i tylko wtedy, gdy $\det A \neq 0$.*

Dowód. Zauważmy, że jeśli A nie jest macierzą izomorfizmu, to A ma liniowo zależne kolumny (pozostawimy to jako ćwiczenie). Zatem $\det A = 0$. W drugą stronę: przypuśćmy, że A jest macierzą izomorfizmu. Wtedy istnieje macierz A^{-1} taka, że

$$AA^{-1} = \text{id}$$

Na mocy Twierdzenia Cauchy'ego mamy

$$\det A \det A^{-1} = \det AA^{-1} = \det \text{id} = 1$$

Stąd $\det A \neq 0$. $\square$

Wprowadźmy teraz definicję, wokół której snuliśmy rozważania na początku tego rozdziału:

Definicja 102 (Orientacja). Niech $\mathcal{V}$ i $\mathcal{W}$ będą dwiema bazami $\mathbb{R}^n$. Powiemy, że $\mathcal{V}$ i $\mathcal{W}$ wyznaczają tę samą orientację jeśli $\det M(\text{id})_{\mathcal{W}}^{\mathcal{V}} > 0$.

Na koniec wprowadźmy jeszcze jedną metodę liczenia wyznacznika, która często przydaje się w praktyce. Niech dla macierzy A , $A_{i,j}$ oznacza macierz powstającą przez usunięcie i -tego wiersza i j -tej kolumny.

Przykład 103. Jeśli

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & \pi & e & \sqrt{2} \\ 17 & 13 & \pi^e & e^\pi \end{pmatrix}$$

to

$$A_{3,2} = \begin{pmatrix} 1 & 3 & 4 \\ 5 & 7 & 8 \\ 17 & \pi^e & e^\pi \end{pmatrix}$$

Stwierdzenie 104. Niech A będzie macierzą $n \times n$. Wtedy dla dowolnego $i \leq n$

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

Wzór podany w powyższym stwierdzeniu nosi nazwę *rekurencyjnej definicji wyznacznika*. Uzależniamy w nim bowiem wyznacznik macierzy $n \times n$ od wyznaczników pewnych mniejszych macierzy.

Przykład 105. Zastosujemy powyższy wzór do policzenia wyznacznika macierzy

$$A = \begin{pmatrix} 1 & 3 & 4 \\ 5 & 7 & 8 \\ 17 & \pi^e & e^\pi \end{pmatrix}$$

Policzymy to przy tym na dwa sposoby: dla 1-ego wiersza (w powyższym stwierdzeniu $i = 1$ i "lewa" suma) i 2-giej kolumny (w powyższym stwierdzeniu $i = 2$ i "prawa" suma).

1.

$$\begin{aligned} \det A &= (-1)^{1+1} a_{1,1} \det A_{1,1} + (-1)^{1+2} a_{1,2} \det A_{1,2} + (-1)^{1+3} a_{1,3} \det A_{1,3} = \\ &= 1 \det A_{1,1} - 3 \det A_{1,2} + 4 \det A_{1,3} = \\ &= (7e^\pi - 8\pi^e) - 3(5e^\pi - 136) + 4(5\pi^e - 119) = \\ &= -8e^\pi + 12\pi^e - 68 \end{aligned} \tag{4}$$

2.

$$\begin{aligned}
 \det A &= (-1)^{1+2} a_{1,2} \det A_{1,2} + (-1)^{2+2} a_{2,2} A_{2,2} + (-1)^{2+3} A_{2,3} = \\
 &= -3 \det A_{1,2} + 7 A_{2,2} - \pi^e A_{2,3} \\
 &= -3(5e^\pi - 136) + 7(e^\pi - 68) - \pi^e(-12) \\
 &= -e^\pi + 12\pi^e - 68
 \end{aligned}$$

3.3 Zadania

Zadanie 44. Oblicz wyznaczniki następujących macierzy:

1.

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

2.

$$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 5 & 1 & 0 & 0 \\ 4 & 3 & 6 & 0 \\ 7 & 9 & 2 & 8 \end{pmatrix}$$

3.

$$\begin{pmatrix} 2 & 0 & 0 & 1 \\ 0 & 3 & 1 & 0 \\ 2 & 1 & 1 & 2 \\ 3 & 4 & 5 & 4 \end{pmatrix}$$

4.

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

5.

$$\begin{pmatrix} 1 & 2 & 0 & 0 \\ 3 & 4 & 0 & 0 \\ 0 & 0 & 5 & 6 \\ 0 & 0 & 7 & 8 \end{pmatrix}$$

6.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 8 & 9 & 10 & 11 & 12 & 13 & 14 \\ 15 & 16 & 17 & 18 & 19 & 20 & 21 \\ 22 & 23 & 24 & 25 & 26 & 27 & 27 \\ 16 & 18 & 20 & 22 & 24 & 26 & 28 \\ 29 & 30 & 31 & 32 & 33 & 34 & 35 \\ 1 & 5 & 7 & 11 & 13 & 17 & 19 \end{pmatrix}$$

Zadanie 45. Niech $v_1, \dots, v_i, v_{i+1}, v_n$ będą n -tkami liczb rzeczywistych. Pokaż, że $M(v_1, \dots, v_n) = M(v_1, \dots, v_i + v_{i+1}, v_{i+1}, \dots, v_n)$. (dodanie do siebie dwóch wierszy macierzy nie zmienia wyznacznika).

Zadanie 46. To zadanie otwarte: opisz jak zmienia się wyznacznik macierzy przy stosowaniu operacji elementarnych.

Zadanie 47. Macierzą górnątrójkątną (dolnątrójkątną) nazywamy macierz $(a_{i,j})$ taką, że jeśli $i > j$ (odpowiednio $i < j$), to $a_{i,j} = 0$. Pokaż, że jeśli $A = (a_{i,j})$ jest macierzą górnątrójkątną, to $\det A = a_{1,1}a_{2,2} \dots a_{n,n}$.

Zadanie 48. Niech $A = (a_{i,j})$, będzie macierzą taką, że

$$a_{i,j} = \begin{cases} k, & \text{jeśli } i = j = k \\ 1 & \text{w.p.p} \end{cases}$$

Udowodnij, że $\det A = n!$.

4 Struktura euklidesowa w $\mathbb{R}^n$

4.1 Zajęcia 30. XI

W tej części zdefiniujemy pojęcia odległości i kąta, a więc temat naszych zainteresowań będzie miał coraz wyraźniejszy związek z geometrią znaną ze szkoły średniej. Zanim to zrobimy, wprowadzimy jeszcze jedno pojęcie czysto algebraiczne, które ułatwi nam omawianie nowo wprowadzanych struktur.

Definicja 106. **Forma dwuliniowa** o wartościach w $\mathbb{R}$ to dowolna funkcja $g : V \times V \rightarrow \mathbb{R}$, dla której zachodzą następujące własności:

1. $g(v + v', w) = g(v, w) + g(v', w)$.
2. $g(av, w) = ag(v, w)$ dla dowolnego $a \in \mathbb{R}$.
3. $g(v, w + w') = g(v, w) + g(v, w')$.
4. $g(v, aw) = ag(v, w)$ dla dowolnego $a \in \mathbb{R}$.

Innymi słowy, forma dwuliniowa to funkcja dwóch zmiennych o wartościach w $\mathbb{R}$ liniowa ze względu na każdą zmienną.

Przykład 107. Mnożenie dwóch liczb rzeczywistych $\times : \mathbb{R}^2 \rightarrow \mathbb{R}$ jest formą dwuliniową.

Przykład 108. Wyznacznik macierzy 2×2 jest formą dwuliniową (zarówno jako funkcja kolumn, jak i wierszy).

Definicja 109. Powiemy, że forma dwuliniowa g jest **symetryczna** jeśli dla dowolnych v, w

$$g(v, w) = g(w, v).$$

Przykładem formy dwuliniowej symetrycznej jest ponownie mnożenie liczb rzeczywistych. Wyznacznik macierzy 2×2 nie jest formą symetryczną. Formy dwuliniowe, podobnie jak odwzorowania liniowe, są wyznaczone jednoznacznie przez swoje wartości na wektorach dowolnej bazy.

Stwierdzenie 110. Niech $v_1, \dots, v_n$ będzie dowolną bazą w V . Wówczas dla dowolnych liczb $a_{i,j} \in \mathbb{R}$ istnieje dokładnie jedna forma dwuliniowa $g : V^2 \rightarrow \mathbb{R}$, taka że dla dowolnych $i, j \leq n$:

$$g(v_i, v_j) = a_{i,j}.$$

Dowód powyższego stwierdzenia jest analogiczny do dowodu stwierdzenia 28 i czytelnik nie powinien mieć już problemu z jego samodzielnym przeprowadzeniem.

Stwierdzenie 111. *Standardowym iloczynem skalarnym na $\mathbb{R}^n$ nazywamy formę dwuliniową symetryczną $\langle v, w \rangle$ taką, że*

$$\langle e_i, e_j \rangle = \begin{cases} 1, & \text{jeśli } i = j, \\ 0 & \text{w przeciwnym wypadku,} \end{cases}$$

gdzie $e_1, \dots, e_n$ stanowią bazę standardową w $\mathbb{R}^n$.

Uwaga 112. Zauważmy, że dla dowolnych v, w o współrzędnych w bazie standardowej odpowiednio $(a_1, \dots, a_n), (b_1, \dots, b_n)$ zachodzi wzór

$$\langle v, w \rangle = a_1 b_1 + a_2 b_2 + \dots + a_n b_n.$$

Przykład 113. $\langle (1, 2, 3), (4, 5, 6) \rangle = 1 \cdot 4 + 2 \cdot 5 + 3 \cdot 6 = 32$.

Postaramy się teraz przedstawić geometryczną interpretację iloczynu skalarnego. Niestety, aby to uczynić posłużymy się pojęciami kąta między prostymi oraz długości wektorów. Tych zaś pojęć nigdzie przedtem nie zdefiniowaliśmy, co byłoby konieczne, gdybyśmy chcieli zadbać o ścisłość wywodu. Problematyczne pojęcia zostaną wprowadzone jeszcze w tej części notatek, ale trafność ich definicji będzie opierać się właśnie na geometrycznej interpretacji iloczynu skalarnego.

Rozważmy wektory $v \in \mathbb{R}^n$ długości 1. Zdefiniujemy następującą funkcję $f : \mathbb{R}^n \rightarrow \mathbb{R}$ określoną na takich wektorach: $f(v, w)$ to długość rzutu prostopadłego wektora w na prostą rozpiętą przez wektor v . Liniowość funkcji f po ustaleniu v jest w miarę jasna: ewentualnie wybierając inną bazę $\mathcal{A}$, której wektory są wzajemnie prostopadłe, możemy przyjąć, że v jest jednym z wektorów tej bazy, a wówczas dla dowolnego w' , $f(v, w')$ to po prostu współrzędna wektora w' odpowiadająca wektorowi bazowemu v , skąd wynika już liniowość bo przy dodawaniu wektorów oczywiście dodają się ich współrzędne w dowolnej bazie.

Zauważmy dalej, że długość rzutu prostopadłego wektora w na prostą rozpiętą przez wektor v jest taka sama niezależnie od tego, czy rozpatrujemy ją jako funkcję określoną na $\mathbb{R}^n$, czy na $\text{lin}(v, w, \cdot)$. Będziemy ją więc rozpatrywać na płaszczyźnie $\text{lin}(v, w)$. Bez utraty ogólności v jest pierwszym wektorem bazy prostopadłej tej płaszczyzny. Naszą funkcję f można więc obliczyć w następujący sposób: wektorowi w o długości 1 przypisujemy jego pierwszą współrzędną. Jest jasne, że wynosi ona dokładnie $\cos \alpha$, gdzie α jest kątem między $\text{lin}(w)$, a $\text{lin}(v)$.

Podobnie, jeśli rozważymy dowolne trzy wektory v, w, w' , aby obliczyć funkcję f wystarczy ją rozważać na $\text{lin}(v, w, w')$, przy czym można założyć, że wybraliśmy pewną bazę tej przestrzeni składającą się z wektorów wzajemnie prostopadłych, z których v jest pierwszym. Jest jasne, że pierwsza współrzędna sumy dwóch wektorów, to suma ich pierwszych współrzędnych. Zdefiniowaliśmy funkcję dwuargumentową f określoną na wektorach długości 1, dla której zachodzą własności:

1. $f(v + v', w) = f(v, w) + f(v', w)$.
2. $f(v, w + w') = f(v, w) + f(v, w')$.

Jest jasne, że dla każdego wektora $v \in \mathbb{R}^n$ istnieje taki v' , że v' jest długości 1 oraz $v = av'$ dla pewnego $a \in \mathbb{R}$. Zatem możemy zdefiniować formę dwulinioową g wzorem:

$$g(v, w) = af(v', w'),$$

gdzie v', w' są długości 1, zaś $v = av', w = bw'$. Zauważmy też, że dla dowolnych wektorów v, w wartość $g(v, w)$ to wartość $f(v, w)$ razy długość v razy długość w .

Podsumowując dotychczasowy wywód: zdefiniowaliśmy formę dwulinioową $g : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ następującym wzorem:

$$\begin{aligned} g(v, w) &= \text{długość rzutu prostopadłego wektora } v \text{ na prostą} \\ &\quad \text{rozpiętą przez wektor } w \text{ razy długość wektora } w \\ &= \text{długość } v \text{ razy długość } w \text{ razy kosinus kąta między nimi.} \end{aligned}$$

Niektórzy czytelnicy rozpoznają tu definicję iloczynu skalarnego w stylu fizyki licealnej. Równoważnie można powiedzieć, że

$$\frac{\langle v, w \rangle}{\|v\| \|w\|}$$

(gdzie $\|v\|$ to długość wektora v) to kosinus kąta między wektorami v a w .

Zauważmy na koniec, że zdefiniowana przez nas forma f przyjmuje na wektorach bazy standardowej dokładnie te same wartości co zdefiniowany przez nas iloczyn skalarny, zatem na mocy stwierdzenia 110, powinna to być ta sama forma, o ile umielibyśmy zdefiniować formę f bez odwoływania się do pojęć, których znaczenia nie objaśniliśmy.

Wykorzystamy teraz powyższe rozważania, żeby zdefiniować pojęcie długości wektora oraz miary kąta w $\mathbb{R}^n$.

Definicja 114. Niech $v \in \mathbb{R}^n$. Przez **normę** wektora v mamy na myśli wielkość $\sqrt{\langle v, v \rangle}$.

Aby upewnić się o formalnej poprawności definicji kąta będziemy potrzebowali następującego ważnego faktu z analizy.

Twierdzenie 115 (Nierówność Cauchy'ego). *Dla dowolnych $v, w \in \mathbb{R}^n$ zachodzi nierówność:*

$$|\langle v, w \rangle| \leq \|v\| \|w\|.$$

W szczególności powyższa nierówność gwarantuje, że

$$\left| \frac{\langle v, w \rangle}{\|v\| \|w\|} \right| \leq 1.$$

Przypomnijemy, że intuicyjnie o powyższej wielkości możemy myśleć jako o kosinusie kąta między v a w . To pozwala nam wprowadzić następującą definicję:

Definicja 116. Przez **kąt** między wektorami v a w w $\mathbb{R}^n$ mamy na myśli

$$\arccos \left(\frac{\langle v, w \rangle}{\|v\| \|w\|} \right).$$

Zgodnie z powyższą definicją kąt między dowolnymi dwoma wektorami wynosi $\pm 90^\circ$ wtedy i tylko wtedy, gdy $\langle v, w \rangle = 0$. Pary wektorów spełniających tę równość nazywamy **prostopadłymi** lub **ortogonalnymi**. Powiemy, że układ wektorów jest **ortogonalny**, jeśli wszystkie pary elementów tego układu są ortogonalne. Powiemy, że układ wektorów jest **ortonormalny**, jeśli jest ortogonalny i wszystkie jego elementy mają długość 1. Zakończymy tę część notatek twierdzeniem o dopełnianiu dowolnych układów wektorów ortonormalnych do baz ortonormalnych, którego dowód daje prosty algorytm znajdowania brakujących wektorów.

Twierdzenie 117. Niech $v_1, \dots, v_k \in \mathbb{R}^n$ będzie dowolnym układem wektorów ortonormalnych. Wówczas istnieją takie $w_{k+1}, \dots, w_n$, że układ $v_1, \dots, v_k, w_{k+1}, \dots, w_n$ jest ortonormalny i stanowi bazę.

Dowód. Pokażemy, że jesteśmy w stanie rozszerzyć $v_1, \dots, v_k$ o jeden wektor w , tak żeby otrzymany układ był ortonormalny. Sprawdzenie, że każdy układ ortogonalny jest liniowo niezależny pozostawiamy jako ćwiczenie. Otrzymamy zatem $k + 1$ niezależnych liniowych wektorów, które stanowią układ ortonormalny.

Przypuśćmy więc, że $k < n$ oraz $v_1, \dots, v_k$ stanowią układ liniowo niezależny w $\mathbb{R}^n$ oraz ortonormalny. Niech v będzie dowolnym takim wektorem, że układ $(v_1, \dots, v_k, v)$ jest liniowo niezależny. Taki wektor istnieje, bo wymiar $\mathbb{R}^n$ jest większy niż k .

Niech

$$v' = v - \langle v_1, v \rangle v_1 - \langle v_2, v \rangle v_2 - \dots - \langle v_k, v \rangle v_k.$$

Zauważmy, że $v' \neq 0$, bo w przeciwnym wypadku przedstawilibyśmy v jako kombinację liniową wektorów $v_1, \dots, v_k$. Ponadto v' jest prostopadły do wektorów $v_1, \dots, v_k$, bo dla dowolnego j mamy:

$$\begin{aligned} \langle v', v_j \rangle &= \langle v, v_j \rangle - \langle v_1, v \rangle \langle v_1, v_j \rangle - \langle v_2, v \rangle \langle v_2, v_j \rangle - \dots - \langle v_k, v \rangle \langle v_k, v_j \rangle \\ &= \langle v, v_j \rangle - 0 - 0 - \dots - 0 - \langle v_j, v \rangle \langle v_j, v_j \rangle \\ &= \langle v, v_j \rangle - \langle v_j, v \rangle \\ &= 0. \end{aligned}$$

Druga równość zachodzi, ponieważ $\langle v_i, v_j \rangle = 0$ dla $i \neq j$. Trzecia równość zachodzi, bo z założenia $\langle v_j, v_j \rangle = \|v_j\|^2 = 1$. Pokazaliśmy zatem, że v' jest ortogonalny do każdego wektora v_j .

Niech

$$w = \frac{v'}{\|v'\|}.$$

Łatwo się przekonać, że w jest prostopadły do wektorów $v_1, \dots, v_k$, oraz że jego długość wynosi 1. Rozszerzyliśmy więc dowolny ortonormalny układ wektorów, które nie rozpinają całego $\mathbb{R}^n$ do ortonormalnego układu wektorów mającego o jeden element więcej. Przez prostą indukcję możemy więc uzyskać tezę twierdzenia.

□

Korzystając z powyższego twierdzenia łatwo się przekonać, że również każdy układ wektorów ortogonalnych można rozszerzyć do bazy ortogonalnej. Algorytm rozszerzania opisany w dowodzie powyższego twierdzenia nosi miano **ortogonalizacji Grama-Schmidta**.

Zadania

Zadanie 49. Czy kąt między wektorami $(1, 2, 3)$ a $(4, 5, 6)$ jest większy, czy mniejszy niż $\frac{\pi}{2}$? A między $(1, 2, 3, 4)$ a $(4, 3, 2, 1)$?

Zadanie 50. Rozszerz do bazy ortonormalnej następujące układy wektorów:

1. $\frac{1}{\sqrt{3}}(1, 1, 1)$ w $\mathbb{R}^3$.
2. $\frac{1}{\sqrt{2}}(1, 1, 0, 0), \frac{1}{\sqrt{2}}(0, 0, 1, 1)$ w $\mathbb{R}^4$.

Zadanie 51. Pokaż, że dowolny układ $v_1, \dots, v_k$ wektorów ortogonalnych jest liniowo niezależny.

Zadanie 52. Niech ϕ będzie obrotem o $\frac{\pi}{3}$ wokół $\text{lin}((1, -1, 1))$ w $\mathbb{R}^3$. Znajdź macierz ϕ w bazie standardowej.

Zadanie 53. Niech ϕ będzie obrotem w $\mathbb{R}^3$, czyli przekształceniem, które w pewnej bazie ortonormalnej ma macierz

$$\begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Pokaż, że ϕ zachowuje iloczyn skalarny, to znaczy dla dowolnych v, w zachodzi równość $\langle v, w \rangle = \langle \phi(v), \phi(w) \rangle$. W szczególności znaczy to, że ϕ zachowuje kąty między wektorami i ich normy.

Zadanie 54. Niech $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ będzie przekształceniem liniowym. Udowodnij, że istnieje $M \in \mathbb{R}, M > 0$ takie, że dla dowolnego $v \in \mathbb{R}^n$

$$\|\phi(v)\| \leq M\|v\|$$

5 Podstawowe pojęcia analizy w $\mathbb{R}^n$

Konwencja 3. Jeśli $x \in \mathbb{R}^n$, to przez x_i oznaczamy i -tą współrzędną x .

5.1 Zajęcia 07.XII

W poprzednim wykładzie, dla dowolnego $n \in \mathbb{N}$ zdefiniowaliśmy funkcję $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$, którą interpretujemy jako "długość wektora w bazie standardowej". Przypomnijmy, że dla $v = (v_1, \dots, v_n) \in \mathbb{R}^n$:

$$\|v\| = \sqrt{\langle v, v \rangle} = \sqrt{\sum_{i=1}^n v_i^2} \quad (\text{Norma})$$

Norma wektora w naturalny sposób zadaje w przestrzeni $\mathbb{R}^n$ funkcję mierzącą odległość pomiędzy punktami x, y jako "długość wektora o początku w x i końcu w y ." Formalnie określimy funkcję $d(x, y)$ wzorem

$$d(x, y) = \|x - y\| \quad (\text{Metryka})$$

Łatwo sprawdzić, że powyższa funkcja ma poniższe własności dla dowolnych $x, y, z \in \mathbb{R}^n$

1. $d(x, y) = d(y, x)$
2. $d(x, y) \geq 0$ oraz $d(x, y) = 0$ wtw. $x = y$
3. $d(x, y) \leq d(x, z) + d(z, y)$

Faktycznie, sprawdzimy 2: ze wzoru (Norma) jest oczywiste, że dla dowolnych x, y , $d(x, y) \geq 0$ i ponieważ $\|0\| = 0$ to także $d(x, x) = 0$. Przypuśćmy zatem, że $x \neq y$. Wtedy istnieje i takie, że $x_i \neq y_i$. Wtedy $|x_i - y_i| > 0$, stąd

$$\|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \geq |x_i - y_i| > 0$$

co kończy dowód. Dowód pozostałych punktów pozostawimy jako ćwiczenie.

Uwaga 118. W przestrzeni liniowej dowolną funkcję spełniającą 1-3 nazywa się *metryką*. W $\mathbb{R}^n$ istnieje wiele różnych metryk - przykład innej metryki niż zdefiniowana powyżej podaje Zadanie 57

Poniższa definicja wprowadza odpowiednik "odcinka otwartego" na prostej i "koła bez brzegu" na płaszczyźnie w dowolnej przestrzeni $\mathbb{R}^n$

Definicja 119. Niech $x \in \mathbb{R}^n$ i $r > 0$. **Kulą otwartą** o środku w x i promieniu r nazywamy zbiór

$$B(x, r) = \{y \in \mathbb{R}^n \mid d(x, y) < r\}$$

Dysponując pojęciem "odległości punktów" w $\mathbb{R}^n$ możemy w naturalny sposób uogólnić definicję granicy ciągu z przypadku jednowymiarowego na wielowymiarowy.

Definicja 120. Niech $(x^k)_{k \in \mathbb{N}}$ będzie ciągiem punktów z $\mathbb{R}^n$ i $x \in \mathbb{R}^n$. Powiemy, że x **jest granicą** x^k (co będziemy zapisywać $\lim_k x^k = x$, lub $x^k \rightarrow x$) jeśli

$$\forall \epsilon > 0 \exists j \forall k > j \quad d(x^k, x) < \epsilon$$

Stwierdzenie 121 (Przeformułowania definicji granicy). Niech $(x^k), x \in \mathbb{R}^n$. Następujące warunki są równoważne

1. x jest granicą (x_k)
2. $\lim_k d(x^k, x) = 0$
3. $\forall \epsilon > 0 \exists j \forall k > j \ x^k \in B(x, \epsilon)$

Przykład 122. Niech $x^k = (\frac{1}{n} \cos(\frac{6\pi}{\sqrt{n}}), \frac{1}{n} \sin(\frac{6\pi}{\sqrt{n}}), \frac{1}{n})$. Wtedy $\lim x^k = (0, 0, 0)$.

W poniższym stwierdzeniu wykazujemy, że zbieżność w $\mathbb{R}^n$ to zbieżność "po współrzędnych": żeby zbadać do czego zbiega dany ciąg (x^k) wystarczy sprawdzić do jakich liczb rzeczywistych zbiegają poszczególne ciągi współrzędnych (x_k) .

Stwierdzenie 123. Niech $x^k = (x_1^k, \dots, x_n^k) \in \mathbb{R}^n$ i $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. Wtedy

$$x^k \rightarrow x$$

wtedy i tylko wtedy, gdy dla dowolnego $i \leq n$,

$$x_i^k \rightarrow x_i.$$

Dowód. Przypomnijmy, że $d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$.

($\Leftarrow$) Jeśli dla dowolnego i $x_i^k \rightarrow x_i$, to także dla dowolnego i ("granica iloczynu to iloczyn granic")

$$(x_i^k - x_i)^2 \rightarrow 0$$

Stąd ("granica sumy to suma granic")

$$\sum_{i=1}^n (x_i^k - x_i)^2 \rightarrow 0$$

Zatem także (dokładając pierwiastek) $d(x^k, x) \rightarrow 0$ i x jest granicą x^k .

($\Rightarrow$) Zauważmy, że dla dowolnych $w, v \in \mathbb{R}^n$ oraz $\epsilon > 0$, jeśli $|w_i - v_i| > \epsilon$ dla pewnego $j \leq n$, to $d(w, v) > \sqrt{\epsilon}$. Faktycznie

$$d(w, v) = \sqrt{\sum_{i=1}^n (w_i - v_i)^2} \geq \sqrt{(w_j - v_j)^2} = |w_j - v_j| > \epsilon$$

(Zachęcamy do narysowania sobie poniższej sytuacji w przypadku $n = 2$). Przypuśćmy zatem, że dla pewnego $i \leq n$, $x_i^k \not\rightarrow x_i$. Wtedy istnieje $\epsilon > 0$ takie, że dla dowolnego j istnieje $k > j$ takie, że $|x_i^k - x_i| > \epsilon$. Na mocy powyższej obserwacji mamy także, że dla dowolnego j istnieje $k > j$ takie, że

$$d(x^k, x) > \epsilon$$

Stąd $x^k \not\rightarrow x$. □

Przejdziemy teraz do definicji granicy funkcji w punkcie. Wprowadźmy najpierw jedną pożyteczną definicję:

Definicja 124. Niech $A \subseteq \mathbb{R}^n$.

1. Zbiór A jest **domknięty** jeśli dla dowolnego ciągu $(x^k) \in A$, $\lim x^k \in A$.
2. Zbiór B jest **domknięciem** A (ozn. $\bar{A}$) jeśli B jest najmniejszym zbiorem domkniętym zawierającym A .

Przykład 125. 1. $[0, 1] \subseteq \mathbb{R}$ jest domknięty. Faktycznie, mamy bowiem, że jeśli dla dowolnego k $0 \leq x^k \leq 1$ to także $0 \leq \lim x^k \leq 1$. $(0, 1)$ nie jest domknięty, gdyż $\lim \frac{1}{n} = 0 \ni (0, 1)$. Mamy za to $\overline{(0, 1)} = [0, 1]$.

2. Przez $\mathbb{Q}^n$ oznaczamy podzbiór $\mathbb{R}^n$ złożony z punktów o wszystkich współrzędnych wymiernych. Mamy $\overline{\mathbb{Q}^n} = \mathbb{R}^n$. Łatwo wynika to z faktu, że $\overline{\mathbb{Q}} = \mathbb{R}$ i Stwierdzenia 123.

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

Definicja 126. Niech $A \subseteq \mathbb{R}^n$, $x \in \bar{A}$, $y \in \mathbb{R}^m$ i $f : A \rightarrow \mathbb{R}^m$. Powiemy, że y jest **granica** f w x wtedy i tylko wtedy, gdy

$$\forall \epsilon \exists \delta \forall z \quad (0 < d(x, z) < \delta \rightarrow d(f(z), y) < \epsilon)$$

Jeśli y jest granicą f w punkcie x , to piszemy $\lim_{z \rightarrow x} f(z) = y$.

Przypomnijmy, że jeśli $f : B \rightarrow C$ i $A \subseteq B$, to przez $f[A]$ oznaczamy obraz zbioru A przy funkcji f .

Stwierdzenie 127 (Równoważne definicje granicy). *Przy oznaczeniach jak w powyższej definicji, następujące warunki są równoważne*

1. $\lim_{z \rightarrow x} f(z) = y$
2. dla dowolnego ciągu $(z^k) \in \mathbb{R}^n$ takiego, że $\lim z^k = x$ mamy $\lim f(z^k) = y$.
3. dla dowolnego $\epsilon > 0$, istnieje $\delta > 0$ taka, że $f[B(x, \delta)] \subseteq B(y, \epsilon)$.

Przykład 128. 1. Niech $f : \mathbb{R} \rightarrow \mathbb{R}^3$ będzie zadana

$$f(x) = (|x| \cos(\pi|x|), |x| \sin(\pi|x|), x)$$

Wtedy $\lim_{x \rightarrow 0} f(x) = 0$. Faktycznie, mamy bowiem

$$d(f(x), 0) = \|f(x)\| = \sqrt{x^2 \cos^2(\pi|x|) + x^2 \sin^2(\pi|x|) + x^2} = |x|\sqrt{2}$$

Stąd jeśli $x^k \rightarrow 0$, to $d(f(x^k), 0) = |x^k|\sqrt{2} \rightarrow 0$.

2. Niech $f : \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}$ będzie zadana

$$f(x, y) = \frac{xy}{x^2 + y^2}$$

Wtedy $(0, 0) \in \overline{\mathbb{R}^2 \setminus \{(0, 0)\}}$, ale f nie ma granicy w $(0, 0)$. Faktycznie zauważmy, że dla ciągów

(a) $x^k = (\frac{1}{k}, \frac{1}{k})$

(b) $y^k = (\frac{1}{k}, 0)$

mamy $\lim x^k = \lim y^k = 0$ lecz dla dowolnego k

$$f(x^k) = f\left(\left(\frac{1}{k}, \frac{1}{k}\right)\right) = \frac{\frac{1}{k} \cdot \frac{1}{k}}{\frac{1}{k^2} + \frac{1}{k^2}} = \frac{1}{2}$$

$$f(y^k) = f\left(\frac{1}{k}, 0\right) = \frac{\frac{1}{k} \cdot 0}{\frac{1}{k^2} + 0} = 0$$

Stąd $\lim f(x^k) = \frac{1}{2} \neq 0 = \lim f(y^k)$.

3. Funkcja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ zdefiniowana

$$f((x, y)) = x \cdot y$$

ma granicę w dowolnym punkcie $\mathbb{R}^2$. Wynika to z tego, że zbieżność w $\mathbb{R}^n$ to zbieżność po współrzędnych, a także granica iloczynu to iloczyn granic.

4. Niech $\mathbb{R}_+$ oznacza zbiór dodatnich liczb rzeczywistych i $A = \mathbb{R}_+ \times \mathbb{R}$. Funkcja $f : A \rightarrow \mathbb{R}$ zdefiniowana

$$f((x, y)) = x^y (:= e^{y \ln(x)})$$

nie ma granicy w punkcie $(0, 0)$. Faktycznie, pokażemy, że dla dowolnego $\epsilon > 0$ funkcja f przyjmuje na zbiorze

$$B((0, 0), \epsilon) \cap A$$

zarówno wartości dowolnie bliskie 0 jak i dowolnie duże, co będzie przeczyło warunkowi 3 ze Stwierdzenia 127. Ustalmy $\epsilon > 0$. Wtedy dla dowolnego $n \in \mathbb{N}$ takiego, że $\frac{1}{n} < \frac{\epsilon}{2}$ mamy

$$\left(\frac{1}{n}, \frac{\epsilon}{2}\right) \in B((0, 0), \epsilon) \cap A$$

oraz

$$\left(\frac{1}{n}, -\frac{\epsilon}{2}\right) \in B((0, 0), \epsilon) \cap A$$

Dla dowolnego $n \in \mathbb{N}$ mamy

$$f\left(\frac{1}{n}, \frac{\epsilon}{2}\right) = \frac{1}{n^{\frac{\epsilon}{2}}}$$

oraz

$$f\left(\frac{1}{n}, -\frac{\epsilon}{2}\right) = \frac{1}{n^{-\frac{\epsilon}{2}}} = n^{\frac{\epsilon}{2}}$$

Stąd

$$\lim_{n \rightarrow \infty} f\left(\frac{1}{n}, \frac{\epsilon}{2}\right) = 0$$

oraz

$$\lim_{n \rightarrow \infty} f\left(\frac{1}{n}, -\frac{\epsilon}{2}\right) = +\infty$$

Ustalając k takie, że dla dowolnego $n > k$ zarówno $(\frac{1}{n}, \frac{\epsilon}{2}) \in B((0, 0), \epsilon) \cap A$ jak i $(\frac{1}{n}, -\frac{\epsilon}{2}) \in B((0, 0), \epsilon) \cap A$ wnioskujemy, że na dowolnie małych kulach wokół $(0, 0)$ f zwraca zarówno wartości dowolnie bliskie 0, jak i dowolnie duże.

Uwaga 129. Analogicznie, co w przypadku funkcji jednej zmiennej, możemy zdefiniować co to znaczy, że dla $A \subseteq \mathbb{R}^n$ $x \in \bar{A}$, $f : A \rightarrow \mathbb{R}$ mamy

$$\lim_{z \rightarrow x} f(z) = \infty$$

Nasz dowód z przypadku 4 pokazuje jednak, że funkcja x^y w punkcie $(0, 0)$ nie ma nawet granicy równej ∞ . Wartości dla symbolu 0^0 nie można zatem zdefiniować nawet uwagłędniając "symbole nieoznaczone". (Przynajmniej jeśli chcemy żeby był spełniony naturalny warunek, żeby x^y była ciągła).

Definicja 130. Niech $A \subseteq \mathbb{R}^n$. Funkcja $f : A \rightarrow \mathbb{R}^m$ jest **ciągła** jeśli dla dowolnego $x \in A$

$$\lim_{z \rightarrow x} f(z) = f(x)$$

Zauważmy, że jeśli $A \subseteq \mathbb{R}^n$ i $f : A \rightarrow \mathbb{R}^m$, to istnieją funkcje $f_i : A \rightarrow \mathbb{R}$, $i \leq m$ takie, że dla dowolnego $x \in A$

$$f(x) = (f_1(x), \dots, f_m(x))$$

Faktycznie, funkcję f_i można zdefiniować np. poprzez złożenie f z funkcją rzutowania $\mathbb{R}^m$ na i -tą oś.

Przykład 131. Niech $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ będzie zadane

$$f((x, y)) = (|xy| \cos(\pi|x|), |xy| \sin(\pi|x|), xy)$$

Wtedy $f_1((x, y)) = |xy| \cos(\pi|x|)$, $f_2((x, y)) = |xy| \sin(\pi|x|)$, $f_3((x, y)) = xy$.

Następujące stwierdzenie jest oczywistą konsekwencją Stwierdzenia 123.

Stwierdzenie 132. Funkcja $f : \mathbb{R}^n \supseteq A \rightarrow \mathbb{R}^m$ jest ciągła wtedy i tylko wtedy, gdy dla dowolnego $i \leq m$, $f_i : \mathbb{R}^n \supseteq A \rightarrow \mathbb{R}$ jest ciągła.

Na koniec tego wykładu wprowadźmy jeszcze jedną ważną definicję:

Definicja 133. Niech $A \subseteq \mathbb{R}^n$.

1. Zbiór A jest **otwarty** jeśli dla dowolnego $x \in A$ istnieje takie $\epsilon > 0$, że

$$B(x, \epsilon) \subseteq A$$

2. Wnętrzem zbioru A nazywamy maksymalny zbiór otwarty zawarty w A . Wnętrze zbioru A oznaczamy $\text{Int}(A)$.

Uwaga 134. Łatwo sprawdzić, że wnętrze zbioru A jest równe

$$\bigcup \{B(x, r) \mid B(x, r) \subseteq A\}$$

Przykład 135. 1. Dla dowolnego $x \in \mathbb{R}^n$ i $r > 0$ kula $B(x, r)$ jest otwarta. Faktycznie, niech $y \in B(x, r)$, wtedy $d(x, y) < r$, a zatem dla $\epsilon = r - d(x, y)$

$$B(y, \epsilon) \subseteq B(x, r)$$

2. $(0, 1] \subseteq \mathbb{R}$ nie jest otwarty, gdyż dla dowolnego $r > 0$, $B(1, r)$ zawiera liczby > 1 .
3. $(0, 1)$ jest otwarty jako podzbiór $\mathbb{R}$, ale zbiór $(0, 1) \times \{0\} \subseteq \mathbb{R}^2$ nie jest otwarty, bo dowolna kula o środku w $(0, 1) \times \{0\}$ zawiera punkty o niezerowej drugiej współrzędnej.
4. $\{(x, y) \in \mathbb{R}^2 \mid xy < 2\}$ jest otwarty.
5. $\mathbb{R}^n, \emptyset$ są otwarte.
6. $\text{Int}(\mathbb{R}^n) = \mathbb{R}^n$, ogólniej, jeśli A jest otwarty, to $\text{Int}(A) = A$.
7. Niech $B = \{(x, y) \in \mathbb{R}^2 \mid y = x^2\}$. Wtedy $\text{Int}(B) = \emptyset$ (zbiór B nie zawiera żadnej kuli otwartej.)

Uwaga 136. $A \subseteq \mathbb{R}^n$ jest otwarty wtedy i tylko wtedy, gdy $\mathbb{R}^n \setminus A$ jest domknięty.

5.2 Zadania

Zadanie 55. Pokaż, że funkcja $f : \mathbb{R} \setminus \{0\} \times \mathbb{R} \rightarrow \mathbb{R}$ zadana wzorem

$$f((x, y)) = \frac{y}{x}$$

nie ma granicy w punkcie $(0, 0)$.

Zadanie 56. Pokaż, że funkcja $f: \mathbb{R}^2 \setminus \{(0,0)\} \rightarrow \mathbb{R}$ zadana wzorem

$$f(x, y) = \frac{x|y|}{\sqrt{x^2 + y^2}}$$

ma granicę w $(0,0)$.

Zadanie 57. Dla dowolnych $x, y \in \mathbb{R}^n$ określmy

$$d_M(x, y) = \max\{|x_i - y_i| \mid i \leq n\}$$

Pokaż, że d_M jest metryką (zob. Uwaga 118). Przypomnijmy, że kulą o środku x i promieniu $r > 0$ w metryce d nazywamy zbiór

$$B(x, r) = \{y \in \mathbb{R}^n \mid d(x, y) < r\}$$

Jak wyglądają kule w metryce d_M ?

Zadanie 58. Jeśli w definicji 120 zastąpimy d przez d_M , to otrzymamy definicję granicy ciągu w metryce d_M . Niech (x^k) będzie ciągiem punktów z $\mathbb{R}^n$ i $x \in \mathbb{R}^n$. Pokaż, że x jest granicą (x_k) w metryce d wtedy i tylko wtedy, gdy jest granicą (x_k) w metryce d_M .

***Zadanie 59.** Pokazać, że $B = \bar{A}$ wtedy i tylko wtedy, gdy B jest największym zbiorem takim, że dla dowolnego $x \in B$ istnieje $(x^k) \in A$ taki, że $x^k \rightarrow x$.

Zajęcia 14. XII

W niniejszej części zaczniemy definiować podstawowe pojęcia rachunku różniczkowego wielu zmiennych. Rzopoczniemy od pojęcia pochodnej cząstkowej funkcji rzeczywistej.

Definicja 137. Niech $f: \mathbb{R}^d \rightarrow \mathbb{R}$. **Pochodną cząstkową** funkcji f w punkcie $a = (a_1, \dots, a_d) \in \mathbb{R}^d$ względem zmiennej x_j nazywamy granicę:

$$\lim_{h \rightarrow 0} \frac{f(a_1, \dots, a_{j-1}, a_j + h, a_{j+1}, \dots, a_d) - f(a_1, \dots, a_d)}{h}$$

i oznaczamy przez $\frac{\partial f}{\partial x_j}(a)$ lub $D_j f(a)$.

Podobnie jak na kursie Wprowadzenia do matematyki II, jeśli dla dowolnego $a \in \mathbb{R}^d$ istnieje $\frac{\partial f}{\partial x_i}(a)$ to przez $\frac{\partial f}{\partial x_i}$ oznaczamy funkcję

$$a \mapsto \frac{\partial f}{\partial x_i}(a)$$

W praktyce obliczanie pochodnych cząstkowych jest bardzo proste: w ramach kursu z Wprowadzenia do matematyki II obliczaliśmy nieraz pochodne funkcji

jednej zmiennej, w których występował parametr rzeczywisty, np. dla $f(x) = e^{\alpha x}$ zachodził wzór:

$$f'(x) = \alpha e^{\alpha x}.$$

Pochodna cząstkowa względem zmiennej x_j to zwykła pochodna funkcji jednej zmiennej x_j , w której zmienne x_i dla $i \neq j$ występują jako parametry. Podajmy kilka przykładów, które pozwolą się oswoić z tym pojęciem.

Przykład 138. 1. Niech $f(x, y) = xy$. Wówczas $\frac{\partial f}{\partial x} = y$, a $\frac{\partial f}{\partial y} = x$.

2. Niech $f(x, y) = e^{xy}$. Wówczas $\frac{\partial f}{\partial x} = ye^{xy}$, a $\frac{\partial f}{\partial y} = xe^{xy}$.

3. Niech $f(x, y) = x^2y$. Wówczas $\frac{\partial f}{\partial x} = 2xy$.

Pochodna cząstkowa spełnia szereg własności pochodnej funkcji jednej zmiennej:

Stwierdzenie 139. Dla dowolnych funkcji $f, g : \mathbb{R}^d \rightarrow \mathbb{R}$, $a \in \mathbb{R}$, które mają pochodne cząstkowe względem x_j w punkcie p zachodzą następujące własności:

1. $\frac{\partial af}{\partial x_j}(p) = a \frac{\partial f}{\partial x_j}(p)$.
2. $\frac{\partial f+g}{\partial x_j}(p) = \frac{\partial f}{\partial x_j}(p) + \frac{\partial g}{\partial x_j}(p)$.
3. $\frac{\partial fg}{\partial x_j}(p) = \frac{\partial f}{\partial x_j}(p)g(p) + f(p) \frac{\partial g}{\partial x_j}(p)$.
4. $\frac{\partial \frac{f}{g}}{\partial x_j}(p) = \frac{\frac{\partial f}{\partial x_j}(p)g(p) - f(p) \frac{\partial g}{\partial x_j}(p)}{g^2(p)}$, o ile $g(p) \neq 0$.

Pochodna złożenia też zachowuje się zgodnie z naszymi oczekiwaniami.

Stwierdzenie 140. Niech $f : \mathbb{R}^d \rightarrow \mathbb{R}$, $g : \mathbb{R} \rightarrow \mathbb{R}$ mają pochodne cząstkowe w punktach odpowiednio p i $f(p)$. Wówczas:

$$\frac{\partial g \circ f}{\partial x_j}(p) = g'(f(p)) \frac{\partial f}{\partial x_j}(p).$$

Obliczanie pochodnych cząstkowych nie stanowi więc istotnie nowego zagadnienia w stosunku do materiału z poprzednich lat. Dla czytelnika nowy może być natomiast wyraźnie geometryczny punkt widzenia, którego przyjęcia wymagają zagadnienia analizy wielu zmiennych rzeczywistych.

Geometrycznie o pochodnej cząstkowej można myśleć tak: funkcja f przypisuje jakąś wielkość punktom przestrzeni $\mathbb{R}^d$. O funkcji $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ możemy na przykład myśleć jako o mapie z zaznaczoną wysokością różnych punktów terenu, a o funkcji $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ jako o przypisaniu każdemu punktowi przestrzeni

trójwymiarowej temperatury w tym punkcie. Pochodna cząstkowa w kierunku x wskazuje wówczas tempo zmiany wielkości f , gdy poruszamy się w kierunku osi OX . Gdy jesteśmy na przykład na północnym zboczu grani ciągnącej się wzdłuż osi wschód–zachód, to idąc wzdłuż tej osi zrazu ani nie nabieramy ani nie tracimy wysokości. Jeśli zatem wyobrazimy sobie, że osi tej odpowiada oś OX , to pochodna cząstkowa funkcji wysokości względem współrzędnej x będzie wynosiła 0. Z drugiej strony jeśli poruszmy się choć odrobinę na północ tracimy wysokość, a jeśli na południe – to ją zyskujemy, zatem jeśli osi północ–południe odpowiada oś OY , to pochodna cząstkowa funkcji wysokości względem zmiennej y jest ujemna. Jest jasne, że kierunki wyznaczone przez osie układu współrzędnych nie mają wyróżnionego statusu z czysto geometrycznego punktu widzenia, warto więc zdefiniować pojęcia tempa zmiany funkcji w dowolnym kierunku.

Definicja 141. Niech $v \in \mathbb{R}^d \setminus \{0\}$. Niech $f : \mathbb{R}^d \rightarrow \mathbb{R}$. **Pochodną kierunkową** funkcji f w kierunku v w punkcie p nazywamy pochodną $g'(0)$, gdzie funkcję $g : \mathbb{R} \rightarrow \mathbb{R}$ definiujemy wzorem:

$$g(t) = f(p + tv).$$

Wielkość tę oznaczamy przez $\frac{\partial f}{\partial v}(p)$ lub $D_v f(p)$.

Zauważmy, że w myśl powyższej definicji pochodna cząstkowa względem zmiennej x_j to to samo co pochodna kierunkowa względem wektora bazy standardowej e_j . Okazuje się, że w wypadku funkcji, z którymi będziemy mieli na ogół do czynienia, pochodną kierunkową można łatwo wyrazić w terminach pochodnych cząstkowych:

Stwierdzenie 142. Niech $v = \sum_{j=1}^d a_j e_j$ będzie dowolnym wektorem w $\mathbb{R}^d$. Niech $p \in \mathbb{R}^d$. Przypuśćmy, że istnieje takie ϵ , że wszystkie funkcje $\frac{\partial f}{\partial x_j}$ są ciągłe na kuli o środku w p i promieniu ϵ . Wtedy

$$\frac{\partial f}{\partial v} = a_1 \frac{\partial f}{\partial x_1} + a_2 \frac{\partial f}{\partial x_2} + \dots + a_d \frac{\partial f}{\partial x_d}.$$

Założenia powyższego twierdzenia można istotnie osłabić, ale nam to osłabienie nie będzie potrzebne. Jego ogólniejszą postać poznamy na kolejnych zajęciach, gdy wprowadzimy pojęcie różniczki. Z powyższego twierdzenia wynika, że ilekroć rozważamy wystarczająco porządne funkcje, to pochodna kierunkowa w kierunku, który jest kombinacją liniową pewnych wektorów ze współczynnikami $a_1, \dots, a_n$ to kombinacja pochodnych w kierunku tychże wektorów z takimi samymi współczynnikami.

Przykład 143. Niech $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ będzie dana wzorem $f(x, y) = xy$. Niech $v = (1, 1)$. Wówczas:

$$\frac{\partial f}{\partial v} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} = y + x.$$

Wprowadźmy kolejny istotny obiekt geometryczny.

Definicja 144. Niech $f : \mathbb{R}^d \rightarrow \mathbb{R}$. **Gradientem** funkcji f w punkcie p nazywamy wektor

$$\left(\frac{\partial f}{\partial x_1}(p), \frac{\partial f}{\partial x_2}(p), \dots, \frac{\partial f}{\partial x_d}(p) \right).$$

Oznaczamy go przez $\text{grad} f(p)$ lub $\nabla f(p)$.

Zauważmy, że gdy funkcja f ma ciągłe pochodne cząstowe w małej kuli o środku w punkcie p , to wprost ze stwierdzenia 142 wynika, że dla dowolnego wektora v zachodzi wzór:

$$\frac{\partial f}{\partial v}(p) = \langle v, \nabla f(p) \rangle.$$

Stąd możemy przez bezpośrednie zastosowanie nierówności Cauchy'ego (twierdzenie 115) wyciągnąć bardzo istotny wniosek:

Stwierdzenie 145. Przy oznaczeniach i założeniach jak wyżej, dla dowolnego wektora v o długości 1 zachodzi nierówność:

$$\left| \frac{\partial f}{\partial v}(p) \right| \leq \left| \frac{\partial f}{\partial \nabla f(p)}(p) \right|.$$

Dowód. Mamy bowiem na mocy nierówności 115:

$$\begin{aligned} \left| \frac{\partial f}{\partial v}(p) \right| &= |\langle v, \nabla f(p) \rangle| \\ &\leq \|v\| \|\nabla f(p)\| \\ &= \|v\| |\langle \nabla f(p), \nabla f(p) \rangle| \\ &= \left| \frac{\partial f}{\partial \nabla f(p)}(p) \right|. \end{aligned}$$

□

Przypomnijmy, że pochodna kierunkowa funkcji w punkcie p kierunku wektora v odpowiada tempu zmiany wartości funkcji, gdy z punktu p będziemy się przesuwać w kierunku wyznaczonym przez wektor v . Powyższe stwierdzenie mówi zatem, że gradient funkcji wyznacza kierunek jej najszybszego wzrostu.

Przykład 146. Niech $f(x, y) = xy$, $p = (0, 1)$. Wówczas:

$$\nabla f = (y, x), \nabla f(p) = (1, 0).$$

Innymi słowy, jeśli chcemy jak najbardziej zwiększyć iloczyn xy , gdzie $x = 0$, $y = 1$ zmieniając oba współczynniki w sumie o ϵ , powinniśmy zwiększać x .

Zadania

Zadanie 60. Znajdź pochodne cząstkowe i gradient następujących funkcji:

1. $f(x, y) = x + y$.
2. $f(x, y) = x^y$, gdzie $x, y > 0$.
3. $f(x, y) = x \cos y$.
4. $f(x, y, z) = e^z(x \cos y + x \sin y)$.

Zadanie 61. Znajdź pochodne kierunkowe funkcji f w punkcie $(0, 0)$ w kierunku v :

1. $f(x, y) = e^{x+y}$, $v = (1, 1)$.
2. $f(x, y) = \sin(-x + 3y)$, $v = (-1, 3)$.

Zajęcia 21. XII

W niniejszej części uogólnimy pojęcie pochodnej na funkcje $f : \mathbb{R}^d \rightarrow \mathbb{R}^k$. Na początek warto zastanowić się, *jakiego typu obiektem* powinna być różniczka odwzorowania $\mathbb{R}^d \rightarrow \mathbb{R}^k$. W wypadku funkcji jednej zmiennej rzeczywistej, pochodna funkcji w punkcie była po prostu pewną liczbą. Jest to miarę jasne, że ta sytuacja nie uogólnia się na większą liczbę wymiarów. Intuicyjnie chcielibyśmy, żeby pochodna zawierała informacje, jak zmieniają się wartości danej funkcji w małym otoczeniu danego punktu. Już w wypadku funkcji $f : \mathbb{R}^n \rightarrow \mathbb{R}$ brak naturalnego kandydata na *jedną* liczbę, która miałaby w sensowny sposób mierzyć takie zmiany. Okazuje się, że naturalnie jest zdefiniować różniczkę odwzorowania w punkcie jako pewne przekształcenie liniowe.¹ Fakt ten, być może dość zaskakujący, będziemy się starali bliżej objaśnić w dalszych częściach notatek.

Definicja 147. Niech $f : \mathbb{R}^d \rightarrow \mathbb{R}^m$ i niech $p \in \mathbb{R}^d$ będzie dowolnym punktem. **Różniczką** odwzorowania f w punkcie p nazywamy takie przekształcenie liniowe A , że

$$\lim_{\|v\| \rightarrow 0} \frac{\|f(p+v) - f(p) - A(v)\|}{\|v\|} = 0.$$

Wprowadzone wyżej odwzorowanie A oznaczamy $Df(p)$

¹Zauważmy, że odwzorowania liniowe $\mathbb{R} \rightarrow \mathbb{R}$ to "to samo", co macierze o wymiarach 1×1 , a te ostatnio to "to samo", co liczby rzeczywiste. Przez "to samo" mamy oczywiście na myśli, że odpowiednie przestrzenie są izomorficzne.

Postaramy się objaśnić powyższą definicję. Zaczniemy od uwagi notacyjnej. $Df(p)$ jest oznaczeniem pewnego odwzorowania liniowego $\mathbb{R}^d \rightarrow \mathbb{R}^m$. W szczególności znaczy to, że efekt zastosowania odwzorowania $Df(p)$ do wektora $v \in \mathbb{R}^d$ możemy zapisać $Df(p)(v)$.

Dla uproszczenia dalszych rozważań przyjmijmy, że rozpatrujemy punkt $p = 0 \in \mathbb{R}^d$ oraz odwzorowanie f takie, że $f(p) = 0$. Wówczas powyższa definicja głosi, że

$$\frac{\|f(v) - Df(0)(v)\|}{\|v\|} \rightarrow 0.$$

Innymi słowy, dla wektorów v o bardzo małej normie odwzorowanie f zachowuje się bardzo podobnie do odwzorowania liniowego $Df(0)$.

Definicja 148. Niech $f : \mathbb{R}^d \rightarrow \mathbb{R}^m$. Niech $p \in \mathbb{R}^d$. Powiemy, że f jest **różniczkowalna** w p dokładnie wtedy, gdy istnieje różniczka odwzorowania f w punkcie p .

Czytelnikowi może się wydawać, że różniczkowalność funkcji wielu zmiennych to bardzo rzadko spełniony warunek. Zauważmy, że różniczkowalność funkcji $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ w punkcie 0 oznacza między innymi, że dla odpowiednio małych liczb dodatnich ϵ_1 oraz ϵ_2 wartość funkcji $f(\epsilon_1, \epsilon_2)$ to mniej więcej $f(\epsilon_1, 0) + f(0, \epsilon_2)$. Okazuje się jednak, że funkcje, które, mówiąc nieścisłe, można opisać "jednym wzorem" na ogół czynią zadość temu wymaganiu. Bardzo ważny warunek dostateczny różniczkowalności można ująć w następującym twierdzeniu:

Twierdzenie 149. Niech $f : \mathbb{R}^d \rightarrow \mathbb{R}^m$, $p \in \mathbb{R}^d$ i niech dla dowolnego $x \in \mathbb{R}^d$ $f(x) = (f_1(x), \dots, f_m(x))$. Załóżmy, że istnieje takie $\epsilon > 0$, że dla dowolnego $x \in B(p, \epsilon)$ wszystkie pochodne cząstkowe

$$\frac{\partial f_i}{\partial x_j}(x)$$

istnieją. Załóżmy ponadto, że dla dowolnego i, j funkcje

$$\frac{\partial f_i}{\partial x_j}$$

są ciągłe w kuli $B(p, \epsilon)$. Wówczas funkcja f jest różniczkowalna w p . Co więcej, macierz różniczki $Df(p)$ w bazie standardowej ma w i -tym wierszu i j -tej kolumnie wyraz

$$\frac{\partial f_i}{\partial x_j}.$$

Macierz różniczki w bazie standardowej nazywamy **macierzą Jacobiego**. Zauważmy, że powyższe twierdzenie podaje w wielu typowych przypadkach efektywny sposób znajdowania różniczki funkcji: mianowicie wiemy, że do opisu przekształcenia liniowego wystarczy podanie jego macierzy w dowolnej bazie. Zobaczmy, jak to działa na konkretnym przykładzie.

Przykład 150. Niech $f : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}^2$ będzie zdefiniowana wzorem

$$f(r, \phi) = (r \cos \phi, r \sin \phi).$$

Innymi słowy, funkcja f liczbie dodatniej r oraz dowolnej liczbie ϕ przyporządkowuje punkt $f(r, \phi)$, który jest odległy od początku układu współrzędnych o r , a kąt między osią OX a odcinkiem łączącym początek układu współrzędnych z punktem $f(r, \phi)$ wynosi ϕ .

Pochodne cząstkowe wynoszą odpowiednio:

$$\begin{aligned} \frac{\partial r \cos \phi}{\partial r} &= \cos \phi & \frac{\partial r \cos \phi}{\partial \phi} &= -r \sin \phi \\ \frac{\partial r \sin \phi}{\partial r} &= \sin \phi & \frac{\partial r \sin \phi}{\partial \phi} &= r \cos \phi. \end{aligned}$$

Wszystkie pochodne cząstkowe są ciągłe w całej dziedzinie. Zatem funkcja f jest różniczkowalna w każdym punkcie dziedziny, a macierz jej różniczki w bazie standardowej to:

$$\begin{pmatrix} \cos \phi & -r \sin \phi \\ \sin \phi & r \cos \phi \end{pmatrix}.$$

Aby przekonać się, jaka jest różniczka funkcji f w danym punkcie należy podstawić odpowiednie wartości w miejsce r oraz ϕ . Na przykład $Df((1, \pi))$ ma w bazie standardowej macierz:

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

Na koniec tej części zastanówmy się jeszcze, jak można myśleć o różniczkach funkcji $f : \mathbb{R}^d \rightarrow \mathbb{R}$. Ustalmy dowolną taką funkcję f . Przypuśćmy, że jej pochodne cząstkowe są ciągłe w całej dziedzinie. W takim wypadku macierz jej różniczki w bazie standardowej jest w każdym punkcie taka sama i wynosi:

$$\left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_d} \right).$$

Zauważmy, że pochodna kierunkowa funkcji f w punkcie p i w kierunku dowolnego wektora v to dokładnie:

$$Df(p)(v).$$

Zatem, mówiąc nieściśle, różniczka koduje informacje dotyczące pochodnych kierunkowych we wszystkich możliwych kierunkach. Dzięki liniowości różniczki, ta informacja dopuszcza w miarę prosty opis w postaci macierzy. Jeśli współrzędne wektora v w bazie standardowej to $(v_1, \dots, v_d)$, to pochodna kierunkowa wyraża się wzorem:

$$\frac{\partial f}{\partial v}(p) = \left(\frac{\partial f}{\partial x_1}(p) \quad \dots \quad \frac{\partial f}{\partial x_d}(p) \right) \begin{pmatrix} v_1 \\ \vdots \\ v_d \end{pmatrix} = \sum_{j=1}^d v_j \frac{\partial f}{\partial x_j}(p).$$

Czytelnik może się czuć trochę zdezorientowany rozróżnieniem między różniczką funkcji o wartościach w $\mathbb{R}$ a jej gradientem. Są to istotnie dwa różne obiekty. Różniczka to *odwzorowanie liniowe* $A : \mathbb{R}^d \rightarrow \mathbb{R}$, natomiast gradient, to jedyny taki wektor $v \in \mathbb{R}^d$, że dla dowolnego $w \in \mathbb{R}^d$:

$$\langle v, w \rangle = A(w).$$

Powyższe równanie można równoważnie zapisać w postaci:

$$\langle \nabla f(p), w \rangle = Df(p)(w).$$

Innymi słowy, odwzorowanie liniowe A można opisać jako branie iloczynu skalarnego danego wektora z gradientem. Obiekty te warto rozróżniać, choć niestety trudno podać przekonujące argumenty za sensownością tego rozróżnienia w tym momencie kursu.

Zadania

Zadanie 62. Znajdź macierze Jacobiego następujących funkcji:

1. $f(x, y) = xy$.
2. $f(x, y) = (e^x + e^{-y}, e^y + e^{-x})$.
3. $f(t) = (\cos t, \sin t, t)$.
4. $f(x, y) = (\sin(x \cos y), \sin(y \cos x))$.
5. $f(x, y, z) = (e^z(x + y), e^z(x - y), z)$.

Zadanie 63. Podaj przykład funkcji $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, która w punkcie 0 ma obie pochodne cząstkowe, ale nie jest różniczkowalna.

Zadanie 64. Podaj przykład funkcji $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, która w punkcie 0 ma wszystkie pochodne kierunkowe, ale nie jest różniczkowalna.

6 Zastosowania pochodnych funkcji wielu zmiennych

6.1 Zajęcia 11.I

Przypomnijmy sobie podstawowe zastosowanie rachunku różniczkowego, które poznaliśmy na zajęciach WdM II: pochodne służyły nam do badania przebiegu zmienności funkcji, tj. znajdowania przedziałów, na których funkcja jest rosnąca, malejąca, gdzie osiąga ekstrema, gdzie jest wklęsła, a gdzie wypukła. Teraz pokażemy, że podobnie jak w jednowymiarowym przypadku, pochodne funkcji wielu zmiennych dostarczają niektórych informacji o zachowaniu funkcji. Warto tutaj od razu zauważyć, że nie ma co liczyć na takie same informacje jak w przypadku jednowymiarowym: niektóre pojęcia nie mają bowiem sensu dla funkcji wielu zmiennych. Jednym z nich jest *monotoniczność* funkcji określonej na podzbiorze $\mathbb{R}^n$, dla $n > 1$. Na $\mathbb{R}^n$ nie ma bowiem naturalnego liniowego porządku, takiego, jaki jest na $\mathbb{R}$. Możemy pytać o monotoniczność funkcji, gdy sprowadzimy nasz przypadek do przypadku jednowymiarowego, np. wyróżniając w $\mathbb{R}^n$ jakiś kierunek (tzn. pytając jak funkcja zachowuje się na podprzestrzeni $\mathbb{R}^n$ rozpinanej przez jeden konkretny wektor). Robiliśmy tak np. gdy uzasadnialiśmy, że gradient wskazuje kierunek najszybszego wzrostu funkcji. Pytania, które przenoszą się niemalże bez zmian na przypadek wielowymiarowy dotyczą wartości ekstremalnych funkcji. Zauważmy jednak, że znów, żeby takie pytanie miało sens, musimy umieć porównywać wartości funkcji o którą pytamy. Dlatego będziemy zajmować się o funkcjami o wartościach w $\mathbb{R}$. Definicja minimum (maksimum) funkcji o wartościach w $\mathbb{R}$ jest w pełni analogiczna do przypadku funkcji jednej zmiennej, o ile pamiętamy, że rolę przedziałów otwartych w przestrzeniach wielowymiarowych pełnią kule.

Definicja 151. Niech $A \subseteq \mathbb{R}^n$ i $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Powiemy, że f ma w punkcie $a \in A$ minimum lokalne (odp. maksimum lokalne) jeśli istnieje $\epsilon > 0$, taki, że dla dowolnego $b \in B(a, \epsilon)$

$$f(b) \geq f(a)$$

(odpowiednio w przypadku maksimum $f(b) \leq f(a)$). Powiemy, że minimum (odp. maksimum) jest **właściwe** jeśli odpowiednia nierówność jest ostra poza przypadkiem $b = a$. Powiemy, że f ma w punkcie a **ekstremum** jeśli ma w tym punkcie minimum lub maksimum. Punkt a stanowi maksimum (minimum) **globalne**, jeśli $f(a) > f(b)$ ($f(a) \leq f(b)$) dla dowolnego b należącego do dziedziny funkcji.

Dobrze jest prześledzić poniższe przykłady szkicując wykresy odpowiednich funkcji.

Przykład 152. Funkcja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ określona

$$f(x, y) = x^2 + y^2$$

ma *minimum lokalne* (a także globalne) w punkcie $(0, 0)$. Funkcja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ zadana

$$f(x, y) = \sin(xy)$$

osiąga w punkcie $(\frac{\pi}{2}, 1)$ maksimum lokalne. Nie jest to maksimum lokalne właściwe, gdyż na całej hiperboli określonej wzorem

$$xy = \frac{\pi}{2}$$

funkcja f przyjmuje wartość 1. Funkcja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ zadana

$$f(x, y) = xy$$

nie ma ekstremum w punkcie $(0, 0)$. Faktycznie, na krzywej

$$y = x$$

mamy $f(x, y) = x^2$ i dla punktów z tej krzywej punkt $(0, 0)$ (należący do niej) jest minimum lokalnym. Jednakże na krzywej

$$y = -x$$

(do której też należy punkt $(0, 0)$.) mamy $f(x, y) = -x^2$, a zatem dla punktów z tej krzywej, f przybiera maksimum lokalne w punkcie $(0, 0)$. Po naszkicowaniu wykresu tej funkcji w $\mathbb{R}^3$ dość łatwo przekonać się dlaczego punkt $(0, 0)$ na wykresie bywa nazywany "siodłem".

Znajdowanie ekstremów funkcji wielu zmiennych na danym zbiorze jest zatem zadaniem subtelniejszym niż w wypadku funkcji jednej zmiennej. Pokażemy teraz jakim rozumowanie może w niektórych przypadkach pomóc w znalezieniu ekstremów funkcji. Zaczniemy od definicji:

Definicja 153. Powiemy, że $A \subseteq \mathbb{R}^n$ jest *ograniczony* jeśli dla pewnego $r \in \mathbb{R}$, $A \subseteq B(0, r)$. Powiemy, że zbiór $A \subseteq \mathbb{R}^n$ jest *zwarty* jeśli jest domknięty (zob. Definicja 124) i ograniczony. *Brzegiem* zbioru zwartego A nazywamy zbiór

$$A \setminus \text{Int}(A).$$

Przykład 154. Oczywiście każda kula $B(x, r) \subseteq \mathbb{R}^n$ jest zbiorem ograniczonym. Dla dowolnego $x \in \mathbb{R}^n$ domknięcie kuli $B(x, r)$ (tzw. *kula domknięta*), tzn. zbiór

$$\overline{B(x, r)} = \{y \in \mathbb{R}^n \mid d(x, y) \leq r\}$$

jest zbiorem zwartym. Jego brzegiem jest sfera

$$\{y \in \mathbb{R}^n \mid d(x, y) = 1\}$$

Wykres paraboli w $\mathbb{R}^2$, tj. zbiór

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

nie jest zbiorem ograniczonym.

Znamy wiele przykładów funkcji, które na swojej dziedzinie nie osiągają ani globalnego maksimum, ani minimum. Takie przykłady można podać już w wypadku funkcji jednej zmiennej (np. tg). Taka sytuacja nie może jednak zdarzyć się w wypadku funkcji określonych na zbiorach zwartych, o czym mówi poniższe (bardzo ogólne) twierdzenie:

Twierdzenie 155 (Weierstrassa o przyjmowaniu kresów). *Niech $A \subseteq \mathbb{R}^n$ będzie zwarty, a $f : A \rightarrow \mathbb{R}$ będzie ciągła. Wtedy istnieją $a, b \in A$ takie, że f przyjmuje w a minimum globalne, a w b maksimum globalne.*

Nasza strategia znajdowania ekstremów funkcji $f : \mathbb{R}^n \rightarrow \mathbb{R}$ będzie następująca: przypuśćmy, że szukamy maksimum funkcji f . Nasze zadanie sprowadzamy do badania funkcji na pewnym zbiorze zwartym $K \subseteq \mathbb{R}^n$. Na tym zbiorze wiemy, że funkcja przybiera swoje kresy. Jeśli poza tym zbiorem funkcja przyjmuje wyłącznie wartości *mniejsze* od pewnej wartości, którą przyjmuje na zbiorze K , to znaczy, że maksimum, które, z tw. Weierstrassa, osiąga na zbiorze K to jest jej maksimum globalne. Jak znaleźć ekstrema funkcji f na zbiorze K ? Do tego przydaje się "wielowymiarowy" lemat Fermata:

Stwierdzenie 156 (lemat Fermata). *Niech $A \subseteq \mathbb{R}^n$ będzie otwarty. Jeśli funkcja $f : A \rightarrow \mathbb{R}$ ma ekstremum lokalne w punkcie $a \in A$ i jest w tym punkcie różniczkowalna, to*

$$Df(a) = 0$$

tzn. dla dowolnego $i \leq n$, $\frac{\partial f}{\partial x_i}(a) = 0$.

Dowód. Ustalmy A , a i f jak wyżej. Przypuśćmy, że f osiąga ekstremum lokalne w punkcie a . Niech ϵ będzie taki, że

$$B(a, \epsilon) \subseteq A$$

a jest ekstremum f na zbiorze A . Wtedy dla dowolnego $i \leq n$ funkcja

$$\begin{aligned} g_i : \mathbb{R} \supseteq (-\epsilon, \epsilon) &\rightarrow \mathbb{R} \\ g_i(t) &= f(a + te_i) \end{aligned}$$

osiąga ekstremum w punkcie $t = 0$. g_i jest funkcją zmiennej rzeczywistej, zatem na mocy "jednowymiarowego" lematu Fermata $g'_i(0) = 0$. Dowód kończy obserwacja, że $g'_i(0) = \frac{\partial f}{\partial x_i}(a)$. $\square$

Założenie otwartości zbioru A było istotne: funkcja

$$f(x, y) = x^2 + y^2$$

osiąga na zbiorze

$$\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$$

maksimum w punktach okręgu jednostkowego

$$\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$$

ale pochodne cząstkowe w tych punktach nie zerują się łącznie. Faktycznie

$$\begin{aligned}\frac{\partial f}{\partial x}(x_1, y_1) &= 2x_1 \\ \frac{\partial f}{\partial y}(x_1, y_1) &= 2y_1\end{aligned}$$

Zatem jedynym punktem (x_1, y_1) takim, że $\frac{\partial f}{\partial x}(x_1, y_1) = \frac{\partial f}{\partial y}(x_1, y_1) = 0$ jest punkt $x_1 = y_1 = 0$.

Warto też pamiętać, że implikacja w lemacie Fermata zachodzi **tylko w jedną stronę**. Nie jest prawdą, że jeśli pochodne cząstkowe funkcji zerują się w danym punkcie, to funkcja w tym punkcie osiąga ekstremum. Rozpatrzmy jeszcze jeden przykład:

Przykład 157. Niech $g(x, y) = x^3 - 3y^2x$. Wtedy

$$\begin{aligned}\frac{\partial g}{\partial x}(x_1, y_1) &= 3x_1^2 - 3y_1^2 = 3(x_1^2 - y_1^2) \\ \frac{\partial g}{\partial y}(x_1, y_1) &= -6y_1x_1\end{aligned}$$

Znów jedynym punktem (x_1, y_1) , dla którego $\frac{\partial g}{\partial x}(x_1, y_1) = \frac{\partial g}{\partial y}(x_1, y_1) = 0$ jest $(0, 0)$, dla którego mamy

$$g(0, 0) = 0.$$

Dowolnie blisko tego punktu funkcja g osiąga jednak zarówno wartości ujemne jak i dodatnie (np. patrząc na wartości tej funkcji na prostej $y = 0$.)

Podsumujmy metodę znajdowania ekstremów funkcji: jeśli umiemy pokazać, że (np.) maksimum f musi leżeć we *wnętrzu* pewnego zbioru zwartego K , to wystarczy znaleźć miejsca, w których na K znika różniczka f (w pozostałych punktach K f nie może mieć ekstremum lokalnego, zatem tym bardziej globalnego) i wybrać największą z tych wartości (taka istnieje na mocy tw. Weierstrassa). Zobaczmy to na przykładzie:

Przykład 158. Niech $f(x, y) = xe^{-(x^2+y^2-2)}$. Zauważmy najpierw, że

$$\lim_{\|x\| \rightarrow \infty} f(x) = 0$$

tzn. w miarę oddalania się od początku układu współrzędnych wartości f stają się coraz bliższe zeru. Najłatwiej to zobaczyć badając zachowanie f na okręgach wokół punktu $(0, 0)$ o coraz większych promieniach, tzn. na zbiorach

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

dla coraz większych R . Łatwo przekonać się, że największą wartością jaką f przyjmuje na $S(R)$ jest

$$\max_R = \sqrt{R}e^{-R-2}$$

a najmniejszą

$$\min_R = -\sqrt{R}e^{-R-2}$$

Oczywiście zarówno $\max_R$, jak $\min_R$ w granicy dają 0, więc dostaliśmy to, co chcieliśmy. Znajdźmy kandydatów na ekstrema funkcji f , tzn. miejsca, w których zerują się pochodne:

$$\begin{aligned}\frac{\partial f}{\partial x}(x, y) &= e^{-(x^2+y^2-2)} + x(-2x)e^{-(x^2+y^2-2)} = e^{-(x^2+y^2-2)}(1 - 2x^2) \\ \frac{\partial f}{\partial y}(x, y) &= -2xye^{-(x^2+y^2-2)}\end{aligned}$$

Zatem jedynymi punktami, w którym obie pochodne się zerują są $a_1 = (\frac{1}{\sqrt{2}}, 0)$ i $a_2 = (-\frac{1}{\sqrt{2}}, 0)$. W a_1 funkcja przyjmuje wartość

$$f(a_1) = \frac{1}{\sqrt{2}}e^{(2-\frac{1}{2})},$$

w a_2 –

$$f(a_2) = -\frac{1}{\sqrt{2}}e^{(2+\frac{1}{2})}.$$

$f(a_1)$ jest zatem kandydatem na maksimum funkcji, $f(a_2)$ – na minimum. Wystarczy pokazać, że poza pewnym zbiorem zwartym, zawierającym punkty a_1 i a_2 , wartości f zawierają się w przedziale $(f(a_2), f(a_1))$. Jako taki zbiór wystarczy wziąć domknięcie kuli $B(0, r)$ dla dostatecznie dużego r .

Przykład 159. Teraz udowodnimy jedną z podstawowych analitycznych nierówności, tzw. *nierówność między średnimi*. Tytułem wprowadzenia: *średnią arytmetyczną* n -liczb $x_1, \dots, x_n$ nazywamy liczbę

$$\frac{x_1 + \dots + x_n}{n}.$$

Średnią geometryczną n -liczb dodatnich $x_1, \dots, x_n$ nazywamy liczbę

$$(x_1 \cdot \dots \cdot x_n)^{\frac{1}{n}}.$$

A oto stwierdzenie, którego będziemy dowodzić stosując poznane metody:

Stwierdzenie 160. Dla dowolnych $x_1 \geq 0, \dots, x_n \geq 0$

$$\frac{x_1 + \dots + x_n}{n} \geq (x_1 \cdot \dots \cdot x_n)^{\frac{1}{n}}$$

Dowód. Zauważmy, że bez ograniczenia ogólności możemy założyć, że $x_1 + \dots + x_n = n$. Tak jest, bo jeśli każdą z liczb x_i przemnożymy przez tę samą liczbę $k > 0$, to zarówno prawa, jak i lewa strona zostaną pomnożone przez k . Wystarczy zatem pokazać, że jeśli $x_1 + \dots + x_n = n$, to

$$x_1 \cdot \dots \cdot x_n \leq 1.$$

Niech teraz

$$\begin{aligned} f : \mathbb{R}^{n-1} &\rightarrow \mathbb{R} \\ f(x_1, \dots, x_{n-1}) &= x_1 \cdot \dots \cdot x_{n-1} (n - x_1 - \dots - x_{n-1}) \end{aligned}$$

A także

$$K = \{(x_1, \dots, x_{n-1}) \mid x_i \geq 0 \text{ oraz } x_1 + \dots + x_{n-1} \leq n\}$$

Łatwo pokazać, że zbiór K jest zwarty.² Zatem f osiąga na K minimum i maksimum. Jesteśmy zainteresowani pokazaniem, że maksimum f na K jest mniejsze niż 1. Ponieważ na brzegu K , tzn. zbiorze

$$\begin{aligned} \partial K &:= \{(x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1} \mid x_i \geq 0 \text{ oraz } x_1 + \dots + x_{n-1} = n\} \\ &\cup \{(x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1} \mid \text{dla pewnego } i, x_i = 0\} \end{aligned}$$

funkcja f jest stale równa 0, a w jego wnętrzu, np. punkcie $(1, \dots, 1)$ przyjmuje wartość dodatnią 1, to musi osiągać maksimum w $\text{Int}(K)$, w punkcie, w którym zerują się pochodne cząstkowe. Obliczmy je (niech $a = (a_1, \dots, a_{n-1})$):

$$\begin{aligned} \frac{\partial f}{\partial x_i}(a) &= \frac{\partial x_1 \cdot \dots \cdot x_{n-1}}{\partial x_i}(a) (n - a_1 - \dots - a_{n-1}) + a_1 \cdot \dots \cdot a_{n-1} \cdot \frac{\partial (n - (x_1 + \dots + x_{n-1}))}{\partial x_i}(a) \\ &= a_1 \cdot \dots \cdot a_{i-1} \cdot a_{i+1} \cdot \dots \cdot a_{n-1} (n - (a_1 + \dots + a_{n-1})) - a_1 \cdot \dots \cdot a_{n-1} \end{aligned} \quad (5)$$

Zatem przyrównując każdą z pochodnych cząstkowych do zera dostajemy układ $n - 1$ równań. Wyciągając w i -tym równaniu $a_1 \cdot \dots \cdot a_{i-1} \cdot a_{i+1} \cdot \dots \cdot a_{n-1}$ przed nawias i skracając dostajemy układ $n - 1$ równań liniowych

$$n - \sum_{j=1}^{n-1} a_j - a_i = 0, \quad i \leq n - 1 \quad (6)$$

²Dobrze jest narysować sobie ten zbiór w przypadku $n = 3$.

Dodając te równania do siebie i przerzucając wyrazy z a_i na prawą stronę dostajemy

$$n(n-1) = (n-1) \sum_{j=1}^{n-1} a_j + a_1 + \dots + a_n$$

a zatem

$$n(n-1) = n \sum_{j=1}^{n-1} a_j$$

Stąd $n-1 = \sum_{j=1}^{n-1} a_j$. Podstawiając tę tożsamość do równania 6 stwierdzamy, że zachodzi ona wtedy i tylko wtedy, gdy $a_1 = \dots = a_{n-1} = 1$. Mamy zatem następującą sytuację: na zbiorze K funkcja f musi przyjmować maksimum. Nie może przyjmować go na brzegu K , bo tam jest stale równa 0, a we wnętrzu K jest dodatnia. Zatem musi przyjmować je we wnętrzu K , w punkcie, w którym zerują się pochodne. Stąd w $(1, \dots, 1)$ f osiąga maksimum na K , a zatem na tym zbiorze wartości funkcji są ≤ 1 , tak jak chcieliśmy. $\square$

Nie zawsze jest tak, że jeśli funkcja na swojej dziedzinie ma dokładnie jedno ekstremum, to jest to jej maksimum lub minimum (tutaj sytuacja jest inna niż w przypadku jednowymiarowym, gdzie powyższa implikacja jest prawdziwa). Ciekawego kontrprzykładu dostarcza poniższa funkcja:

Przykład 161. Niech $f(x, y) = x^2(1+y)^3 + y^2$. Wtedy f ma dokładnie jedno ekstremum lokalne, które jest jej minimum lokalnym, ale nie jest ograniczona ani z góry, ani z dołu. Istotnie, f nie jest ograniczona z góry, ani z dołu, bo na zbiorze

$$x = y$$

jest równa wielomianowi 5-tego stopnia

$$x^2(1+x)^3 + x^2,$$

który (jak to wielomian nieparzystego stopnia) nie jest ograniczony z góry ani z dołu. Zbadajmy zachowanie pochodnych cząstkowych:

$$\begin{aligned} \frac{\partial f}{\partial x}(x_1, y_1) &= 2x_1(1+y_1)^3 \\ \frac{\partial f}{\partial y}(x_1, y_1) &= 3x_1^2(1+y_1)^2 + 2y_1 \end{aligned}$$

Stąd $\text{grad} f(x_1, y_1) = 0$ wtedy i tylko wtedy, gdy $x_1 = y_1 = 0$. Łatwo sprawdzić, że na $B(0, \frac{1}{2})$, f przyjmuje tylko dodatnie wartości. Stąd $(0, 0)$ jest minimum lokalnym funkcji f i innego nie ma, bo to jedyne miejsce, w którym zeruje się gradient f .